

Characterization of the Secondary Metabolome of a Lichenizing Fungus

By

ROBERT L. BERTRAND

A Thesis submitted to the Faculty of Graduate Studies
of The University of Manitoba in partial fulfillment
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Department of Chemistry
Faculty of Science
University of Manitoba
Winnipeg, Manitoba, Canada

May 2019

© Copyright 2019, Robert Bertrand

Front Matter

Abstract

Lichens are traditionally described as symbionts of fungi and algae and are renowned for their diverse secondary metabolites. How lichens produce these natural products, such as their biosynthetic pathways and the genes that are required to produce these molecules, remains unknown. The genome of the fungal partner of the lichen *Cladonia uncialis* was *de novo* sequenced and its genetic secondary metabolome was annotated. This work revealed 48 secondary metabolite biosynthetic gene clusters, providing a first glimpse into the genetic programming of lichen polyketides, terpenes, and non-ribosomal peptides. A ‘deductive approach’ employing retro-biosynthetic reasoning was applied to find the genes responsible for the biosynthesis of usnic acid, the most extensively studied lichen secondary metabolite. A ‘homology mapping’ approach, combining phylogenetics and rational pathway deductions, provided putative assignments of biosynthetic function to nine gene clusters in *C. uncialis*, including what appears to be gene clusters encoding patulin, the betaenones, 6-hydroxymellein, and grayanic acid. The development of a reliable heterologous expression protocol is a prerequisite to an advanced understanding of lichen secondary metabolite biosynthesis. The heterologous expression of seven *C. uncialis* polyketide synthases was investigated using the NSAR1 *A. oryzae* platform. These experiments revealed that *A. oryzae* can transcribe lichen genes and remove introns to produce translation-coherent mRNA. Though *de novo* biosynthesis of lichen metabolites was not observed in *A. oryzae*, this work propels lichen natural product studies into the genomic era and provide a foundation upon which future studies may strive to uncover the rich biosynthetic potential of these enigmatic organisms.

Acknowledgements

There are many people to whom I owe the success of this research program. Many of the most formative experiences of my doctoral program arose through the efforts of my supervisor, John Sorensen, to afford me with the opportunities that have allowed me to develop myself to my potential. His enthusiasm and thorough familiarity with natural products chemistry are what had inspired me to pursue this field of research. I thank him immensely for his mentorship and patience throughout these years. I thank my committee members, Sean McKenna, Silvia Cardona, Joe O'Neil, and the late Phil Hultin, for their encouragement and guidance, and for helping me pass through several critical junctures of my doctoral program. I also extend my thanks to Ikuro Abe for his invitation and hospitality at the University of Tokyo, and for providing our research group with the fungal expression platform that we needed to conduct our heterologous expression trials. The training that I received from Yudai Matsuda, Takayoshi Awakawa, and Zhi Yang Quan was immeasurably helpful in conducting this research. I thank them for providing this training and for their patience in answering my many questions. I thank Evan Booy, of the Department of Chemistry, who is thoroughly knowledgeable in biochemistry and molecular biology techniques. His advice allowed me to overcome several frustrating obstacles during my research program. His service to the chemistry department by operating the 'BioBar' has also allowed me to conduct my research at a much faster pace. Nidhi Shah and Manu Singh, of the Department of Microbiology, are well-versed in bioinformatics and genome assembly, and I thank them for the advice they provided to me during this phase of my project. I also wish to thank Teresa de Kievit, also from the Department of Microbiology, for kindly providing me with access to her powerful computer – an instrument that was necessary for completing the genome assembly. Scott Legare, of the Sorensen group, is an aspiring organic chemist who kindly prepared for me chemical standards of

several of the molecules that we expected to observe within our heterologous expression trials. It would have been difficult to conduct this expression work without these chemical standards, and I thank him for this contribution. I have been blessed to have worked with several wonderful undergraduate project students and visiting researchers. I thank Nichdali Nava (Monterrey Institute of Technology and Higher Education, Mexico), Snimar Bali (University of Manitoba), Vienna Peters (UofM), and Sabina Ozog (UofM), for going above and beyond the call of duty. Their contributions towards preparing and processing experiments have allowed our group to conduct our heterologous expression trials at a pace that would have otherwise been impossible. I have been lucky to be able to lean upon Keith Travis, the academic programs administrator in the Department of Chemistry, for helping me navigate through the many administrative concerns that inevitably arise within this line of work. I have had a good experience as a teaching assistant in the Department of Chemistry. I thank Horace Luong, senior instructor for the organic chemistry teaching laboratories, for his inexhaustible patience and his sincere desire to develop his teaching assistants. I have become a better instructor because of him. I am fortunate to have been a member of the Canadian Forces. I am grateful for the training I had received, the many friends that I had made along the way, and for the charity that I had received that allowed me to balance my military and academic duties throughout this great journey. Edwin Othen, of the Sorensen lab, is a great friend to have, and he was there to help me traverse many of the most calamitous moments of my doctoral program. I thank Mona Abdel-Hameed, also of the Sorensen lab, for welcoming me as a member of her family, and for the many wonderful times we have shared together. I owe my marriage to Mona, as she is the one who introduced me to my wife. Lucile, I love you. You are my support, my guidance, my angel, and my best friend. Your love has transformed me into a better person. I am also grateful for the love and support that I have received from my family –

Joe, Guy, Annette, Gary, Lynn, Trevor, Brett, Pierrick, Marie-Odile, Lucas, and Charlène. Don't worry Zack and Lilly, I haven't forgotten you! Mom and Dad, though you are no longer with me here on this Earth, I know that you love me, and I hope I have made both of you proud. If there is Providence in my life – and I believe there is a case to be argued – then I thank God the Father and His Son for providing me with the gift of education and for guiding my life towards the pursuit of honourable works. Lastly, if there is a case to be made for a being in this universe who has more control over one's research career than God, then I must finally extend my gratitude to the Natural Sciences and Engineering Research Council (NSERC) for providing me with funding I needed to conduct this doctoral research, and for their recent announcement of their commitment to my research in the form of a post-doctoral fellowship.

Dedication

To the father who looked upon his toddler, and said:

“This kid’s gonna stay in school ‘til he rots!”

Reproduction of Texts

Figures, tables, and passages within this thesis have been adapted from works that I have authored and published in academic journals. To ensure faithful and legal reproduction of works, I have requested and received permissions from the relevant publishers to reproduce these works (in parts or in whole) within this thesis. The permission documents are found in the appendix, and pertain to the following articles:

1. Abdel-Hameed M, Bertrand RL, Piercey-Normore M, Sorensen JL. 2016. Putative identification of the usnic acid biosynthetic gene cluster by *de novo* whole-genome sequencing of a lichen-forming fungus. *Fungal Biology* 120: 306-16. <http://dx.doi.org/10.1016/j.funbio.2015.10.009>
2. Abdel-Hameed M, Bertrand RL, Piercey-Normore M, Sorensen JL. 2016. The identification of 6-hydroxymellein synthase and accessory genes in the lichen *Cladonia uncialis*. *Journal of Natural Products* 79: 1645-1650. <http://dx.doi.org/10.1021/acs.jnatprod.6b00257>
3. Bertrand RL, Abdel-Hameed M, Sorensen JL. 2018. Lichen biosynthetic gene clusters part I: Genome sequencing reveals a rich biosynthetic potential. *Journal of Natural Products* 81: 723-731. <http://dx.doi.org/10.1021/acs.jnatprod.7b00769>
4. Bertrand RL, Abdel-Hameed M, Sorensen JL. 2018. Lichen biosynthetic gene clusters part II: Homology mapping suggests a functional diversity. *Journal of Natural Products* 81: 732-748. <http://dx.doi.org/10.1021/acs.jnatprod.7b00770>
5. Abdel-Hameed M, Bertrand RL, Donald LJ, Sorensen JL. 2018. Lichen ketosynthase domains are not responsible for inoperative polyketide synthases in *Ascomycota* hosts. *Biochemical and Biophysical Research Communications* 503: 1228-1234. <http://dx.doi.org/10.1016/j.bbrc.2018.07.029>
6. Bertrand RL, Sorensen JL. 2018. A comprehensive catalogue of polyketide synthase gene clusters in lichen fungi. *Journal of Industrial Microbiology and Biotechnology* 45: 1067-1081. <http://dx.doi.org/10.1007/s10295-018-2080-y>

For the purpose of writing the introductory material (see Chapter 1), several figures have also been adapted from published works of which I am not an author. Permissions from the relevant publishers have also been acquired for the purpose of reproducing these figures within this thesis. These permission documents are also found in the Appendix.

Contribution of Authors

Chapter 1: *Non-applicable*

Chapter 2: *Non-applicable*

Chapter 3: Mona-Abdel-Hameed is the primary contributor of the genetic taxonomic identification of the lichen, the sub-culturing of the fungal partner, and the isolation of DNA from the sub-cultured partner for purpose of *de novo* whole-genome sequencing. I am the primary contributor of other all presented work within this chapter.

Chapter 4: Mona Abdel-Hameed performed the degenerate primer amplification and sequencing of the transcriptionally active PKS. I am the primary contributor of all other presented work within this chapter.

Chapter 5: I am the primary contributor of the presented work within this chapter.

Chapter 6: I am the primary contributor of the presented work within this chapter.

Chapter 7: I am the primary contributor of the presented work within this chapter.

Chapter 8: I am the primary contributor of the presented work within this chapter.

Chapter 9: *Non-applicable*

Contents

Front Matter	I
Abstract	I
Acknowledgements	II
Dedication	V
Reproduction of Texts	VI
Contribution of Authors	VII
Contents	VIII
List of Tables	XII
List of Figures	XIII
List of Supporting Information Items	XVII
Non-IUPAC Abbreviations	XXV
Chapter 1: An Introduction to Lichens and their Secondary Metabolites.....	1
1.1. Introduction – Lichens and their natural products	1
1.2. Biosynthesis of secondary metabolites	7
1.2.1. Polyketides	8
1.2.2. Non-ribosomal peptides	16
1.2.3. Hybrid PKS-NRPS	19
1.2.4. Terpenes	20
1.3. Biosynthetic pathways, gene clusters, and biotechnological applications	21
1.4. What did we know about lichen secondary metabolite biosynthesis in 2012?	26
1.5. Summary of research objectives	29
Chapter 2: Materials and Methods	31
2.1. Materials and methods related to Chapter 1	31
2.2. Materials and methods related to Chapter 2	31
2.3. Materials and methods related to Chapter 3	31
2.3.1. Collection and taxonomic identification of <i>C. uncialis</i>	31
2.3.2. Sub-culturing of <i>C. uncialis</i> from the algal partner	31
2.3.3. Genome sequencing and assembly	32
2.3.4. Annotation of secondary metabolite gene clusters	33

2.3.5. Phylogenetics	33
2.4. Materials and methods related to Chapter 4	33
2.4.1. Detection and sequencing of a transcriptionally active PKS	33
2.4.2. Annotation of gene cluster	33
2.4.3. Phylogenetics	34
2.5. Materials and methods related to Chapter 5	34
2.5.1. Annotation of lichen gene clusters	34
2.5.2. Proofreading of annotations	34
2.6. Materials and methods related to Chapter 6	35
2.6.1. Identification of genetically-similar gene clusters	35
2.6.2. Phylogenetics	35
2.7. Materials and methods related to Chapter 7	36
2.7.1. <i>In silico</i> experiments	36
2.7.2. Polymerase chain reaction	36
2.7.3. Preparation of plasmids	37
2.7.4. Transformation of NSAR1 <i>A. oryzae</i>	38
2.7.5. Incubation conditions and media extractions	39
2.7.6. Reverse transcriptase-polymerase chain reaction (RT-PCR)	40
2.7.7. Phylogenetics	40
2.7.8. HPLC	41
2.7.9. SDSPAGE	41
2.7.10. Western Blot	42
2.8. Materials and methods related to Chapter 8	43
2.9. Materials and methods related to Chapter 9	43
Chapter 3: Identification of Usnic Acid Gene Cluster by Deductive Method	44
3.1. Selecting a suitable lichen and secondary metabolite to study	44
3.2. <i>De novo</i> whole-genome sequencing of an usnic acid-producing species of lichen	52
3.3. Identifying the usnic acid biosynthetic gene cluster in <i>C. uncialis</i>	55
3.4. Summary	58
Chapter 4: 6-Hydroxymellein and the ‘Homology Mapping’ Method	60
4.1. Introduction - Searching for the usnic acid gene cluster, <i>circa</i> 2008-2012.	60
4.2. What does this PKS do?	61

4.3. Summary	65
Chapter 5: Gene Annotation of a Lichen Secondary Metabolome	67
5.1. Introduction - Annotating the lichen secondary metabolome	67
5.2. Type III PKS gene clusters.....	68
5.3. Multi-module NRPS gene clusters	69
5.4. Single-module NRPS gene clusters.....	70
5.5. Hybrid PKS-NRPS gene clusters	71
5.6. Terpene synthase gene clusters	72
5.7. Type I PKS gene clusters	74
5.8. Interpretation of findings and importance of contributions	78
5.9. Summary	79
Chapter 6: Assignment of Biosynthetic Pathways to Lichen Gene Clusters	81
6.1. Introduction – Applying a ‘homology mapping’ approach to the secondary metabolome	81
6.2. A proposed complete grayanic acid pathway.....	83
6.3. A proposed complete betaenone pathway	85
6.4. A proposed complete patulin pathway	88
6.5. A proposed azaphilone-based pathway	93
6.6. A proposed cytochalasin-based pathway	96
6.7. A proposed fusarubin-based pathway	100
6.8. A proposed pestheic acid-based pathway.....	105
6.9. A proposed mycophenolic acid-based pathway	108
6.10. Summary	111
Chapter 7: Preliminary Experiments Towards Heterologous Expression Trials.....	112
7.1. Introduction – Choosing a heterologous host.....	112
7.2. ‘Dry’ control experiments	115
7.2.1. Introduction	115
7.2.2. Phylogenetic analysis of KS domains	115
7.2.3. Multiple sequence alignment of KS domains.....	117
7.2.4. Protein structure modelling of KS domains	119
7.3. ‘Wet’ control experiments.....	123
7.3.1. Introduction	123
7.3.2. Testing the selection markers of NSAR1 <i>A. oryzae</i>	123

7.3.3. Testing toxicity of methylphloroacetophenone and usnic acid on <i>A. oryzae</i>	124
7.3.4. Testing heterologous transcription of mpas and mpao in <i>A. oryzae</i>	124
7.3.5. Choosing an incubation end-point	128
7.3.6. Mock extractions	128
7.3.7. Functional expression of a characterized non-lichen PKS in <i>A. oryzae</i>	129
7.3.8. Functional expression of an uncharacterized non-lichen PKS in <i>A. oryzae</i>	131
7.3.9. Testing heterologous translation of mpas in <i>A. oryzae</i>	133
7.4. Summary	135
Chapter 8: Heterologous Expression Trials	136
8.1. Introduction – heterologous expression trials	136
8.2. Heterologous expression trials of the usnic acid gene cluster (MPAS and MPAO).....	136
8.2.1. Trials with native mpas and mpao	136
8.2.2. Trials with codon-optimized mpas and mpao.....	137
8.2.3. Trials using phosphopantetheinyl transferase (PPTase).....	138
8.2.4. Trials using multiple accessory genes	140
8.3. Heterologous expression trials of 6-hydroxymellein gene cluster (6HMS and KRDH)..	142
8.4. Heterologous expression trials of two orsellinic acid synthases (OAS and GAS).....	144
8.5. Heterologous expression trials of 6-methylsalicylic acid synthase (6MSAS)	146
8.6. Heterologous expression trials of two type III PKS (T3PKS1 and T3PKS2).....	147
8.7. Comparison of codon adaptability index (CAI) scores	149
8.8. Summary	151
Chapter 9: Conclusions and Future Prospects.....	153
9.1. Current knowledge on the genetic programming of lichen secondary metabolite biosynthesis	153
9.2. Proposed approaches to functionally expressing lichen genes in heterologous hosts	155
9.3. Proposed approaches to discovering cryptic secondary metabolites in lichens	158
9.4. Summary	161
Appendix.....	163
References	163
Supporting information items.....	194
Permissions Documents	304

List of Tables

Table 1: All literature reports of sequenced biosynthetic genes from lichenizing fungi prior to September 2012.

Table 2: BLAST similarity statistics between genes within *C. uncialis* and *A. terreus*.

Table 3: BLAST statistics and proposed functions of the *C. uncialis* genes that are genetically similar to genes within the grayanic acid cluster of *C. grayi*.

Table 4: BLAST statistics and proposed functions of the *C. uncialis* genes that are genetically similar to genes of the betaenone pathway of *P. betae*.

Table 5: BLAST statistics and proposed functions of the *C. uncialis* genes that are genetically similar to genes of the patulin cluster of *A. clavatus*.

Table 6: BLAST statistics and proposed functions of the *C. uncialis* genes that are genetically similar to genes of the azaphilone cluster of *M. pilosus*.

Table 7: BLAST statistics and proposed functions of the *C. uncialis* genes that are genetically similar to genes of the cytochalasin cluster of *A. clavatus*.

Table 8: BLAST statistics and proposed functions of the *C. uncialis* genes that are genetically similar to genes of the fusarubin pathway of *F. fujikuroi*.

Table 9: BLAST statistics and proposed functions of the *C. uncialis* genes that are genetically similar to genes of the pestheic acid cluster of *P. fici*.

Table 10: BLAST statistics and proposed functions of the *C. uncialis* genes that are genetically similar to genes of the mycophenolic acid cluster of *P. brevicompactum*.

Table 11: Codon adaptation index (CAI) scores of lichen versus non-lichen PKS genes in *A. oryzae*. Adapted from Bertrand et al. (2019), manuscript in preparation.

Table 12: Complete PKS genes from lichen fungi reported in the literature (updated to 2018).

List of Figures

Figure 1: Cross-section of a lichen thallus.

Figure 2: Fluorescence hybridization of *Basidiomycetes* yeasts within the cortex of lichens.

Figure 3: Selected lichen secondary metabolites and their bioactive properties.

Figure 4: General scheme of polyketide elongation by polyketide synthases (PKS).

Figure 5: Termination release mechanisms by PKS.

Figure 6: ‘Non-reducing’, ‘partially reducing’, and ‘highly reducing’ PKS.

Figure 7: 6-deoxyerythronolide synthase (DEBS), a model modular PKS.

Figure 8: 6-methylsalicylic acid synthase (MSAS), a model iterative PKS.

Figure 9: Summary of polyketide assembly by type I, II, and III PKS.

Figure 10: Schematic of peptide elongation by non-ribosomal peptide synthetases (NRPS).

Figure 11: Biosynthesis of surfactin by surfactin synthetase, a model NRPS.

Figure 12: Dihydromaltophilin, an example of a metabolite produced by a hybrid PKS-NRPS.

Figure 13: Carbocation-mediated polymerization of DMAPP and IPP by terpene synthases.

Figure 14: Three examples of terpen(oids).

Figure 15: The patulin gene cluster and biosynthetic pathway.

Figure 16: Summary of isotope feeding experiments performed by Taguchi et al. (1969) that led to the elucidation of the usnic acid biosynthetic pathway in lichens.

Figure 17: Radical coupling mechanism of methylphloracetophenone.

Figure 18: Aldol-style versus Claisen-style folding of C₈-polyketides by CLC and TE domains.

Figure 19: The hypothesized usnic acid biosynthetic pathway.

Figure 20: General experimental approach for the heterologous expression of the usnic acid gene cluster.

Figure 21: Assembly programs used in this work

Figure 22: The five non-reducing PKS that possessed C-methyltransferase (cMeT) domains in *C. uncialis*.

Figure 23: The proposed usnic acid gene cluster and biosynthetic pathway.

Figure 24: Truncated phylogenetic trees of (A) the KS domain of *mpas*, and (B) the CLC domain of *mpas*.

Figure 25: The gene cluster of unknown function, containing a transcriptionally active PKS.

Figure 26: Terrein gene cluster and biosynthetic pathway.

Figure 27: Truncated phylogenetic trees of monooxygenase, KR-DH-like gene, and PKS.

Figure 28: Proposed biosynthetic pathway in *C. uncialis*.

Figure 29: Type III polyketide synthase (PKS) gene clusters in the *C. uncialis* mycobiont genome.

Figure 30: A multi-module non-ribosomal peptide synthetase (NRPS) gene cluster in the *C. uncialis* mycobiont genome.

Figure 31: Single-module non-ribosomal peptide synthetase (NRPS) gene clusters in the *C. uncialis* mycobiont genome.

Figure 32: Non-ribosomal peptide synthetase – polyketide synthase (PKS-NRPS) gene clusters identified in the *C. uncialis* mycobiont genome.

Figure 33: Terpene synthase gene clusters in the *C. uncialis* mycobiont genome.

Figure 34: Type I non-reducing polyketide synthase (PKS) gene clusters in the *C. uncialis* mycobiont genome.

Figure 35: Type I reducing polyketide synthase (PKS) gene clusters in the *C. uncialis* mycobiont genome.

Figure 36: The *C. grayi* grayanic acid gene cluster and *C. uncialis* uncharacterized gene cluster.

Figure 37: Truncated phylogenetic trees illustrating degree of relationship between genes of *C. uncialis* and *C. grayi* pertinent to grayanic acid biosynthesis.

Figure 38: Experimentally supported pathway for grayanic acid biosynthesis in *C. grayi*, and juxtaposed, the biosynthetic pathway that is proposed to be encoded in *C. uncialis*.

Figure 39: The *P. betae* betaenone gene cluster and *C. uncialis* uncharacterized gene cluster.

Figure 40: Truncated phylogenetic trees illustrating degree of relationship between genes of *C. uncialis* and *P. betae* pertinent to betaenone biosynthesis.

Figure 41: The experimentally supported pathway for betaenone A-C biosynthesis in *P. betae*, and juxtaposed, the biosynthetic pathway that is proposed to be encoded in *C. uncialis*.

Figure 42: The *A. clavatus* patulin gene cluster and *C. uncialis* uncharacterized gene cluster.

Figure 43: Truncated phylogenetic trees illustrating degree of relationship between genes of *C. uncialis* and *A. clavatus* pertinent to patulin biosynthesis.

Figure 44: The experimentally supported pathway for patulin biosynthesis in *A. clavatus*, and juxtaposed, the biosynthetic pathway that is proposed to be encoded in *C. uncialis*.

Figure 45: The *M. pilosus* azaphilone gene cluster and *C. uncialis* uncharacterized gene cluster.

Figure 46: Truncated phylogenetic trees illustrating degree of relationship between genes of *C. uncialis* and *M. pilosus* pertinent to azaphilone biosynthesis.

Figure 47: The experimentally supported pathway for azaphilone biosynthesis in *M. pilosus*, and juxtaposed, the biosynthetic pathway that is proposed to be encoded in *C. uncialis*.

Figure 48: The *A. clavatus* cytochalasin gene cluster and *C. uncialis* uncharacterized gene cluster.

Figure 49: Truncated phylogenetic trees illustrating degree of relationship between genes of *C. uncialis* and *A. clavatus* pertinent to cytochalasin biosynthesis.

Figure 50: The experimentally supported pathway for cytochalasin biosynthesis in *A. clavatus*, and juxtaposed, the biosynthetic pathway that is proposed to be encoded in *C. uncialis*.

Figure 51: Examples of secondary metabolites possessing the cyclized phenylalanine moiety from *Talaromyces sp.*, *Penicillium sp.*, and *Aspergillus sp.*

Figure 52: The *F. fujikuroi* fusarubin gene cluster and *C. uncialis* uncharacterized gene cluster.

Figure 53: The experimentally supported pathway for fusarubin biosynthesis in *F. fujikuroi*, and juxtaposed, the biosynthetic pathway that is proposed to be encoded in *C. uncialis*.

Figure 54: A plausible biosynthetic pathway leading from the fusarubin intermediate 8-*O*-methylanthydrofusarubinlactol to the lichen metabolites haemoventosin and coronataquinone.

Figure 55: The *P. fici* pestheic acid gene cluster and *C. uncialis* uncharacterized gene cluster.

Figure 56: The experimentally confirmed pathway for pestheic acid biosynthesis in *P. fici*, and juxtaposed, the biosynthetic pathway that is proposed to be encoded in *C. uncialis*.

Figure 57: Proposed biosynthesis of anthraquinone lichen natural products from emodin.

Figure 58: The *P. brevicompactum* mycophenolic acid gene cluster and *C. uncialis* unassigned gene cluster.

Figure 59: The experimentally confirmed pathway for mycophenolic acid biosynthesis in *P. brevicompactum*, and juxtaposed, the proposed biosynthetic pathway encoded in *C. uncialis*.

Figure 60: Sketch of experimental scheme for heterologous expression trials.

Figure 61: Expression plasmids pUSA (methionine-based selection), pTAex3 (arginine-based selection), and pAdeA (adenine-based selection).

Figure 62: Phylogenetic tree containing KS of methylphloroacetophenone synthase, KS of 6-hydroxymellein synthase, and three representative KS domains from each of the 14 groups of the KS3 family of ketosynthases.

Figure 63: Multiple sequence alignments of ketosynthase domains.

Figure 64: Protein modelling of KS domains of MPAS and 6HMS.

Figure 65: Control experiment confirming triple auxotrophy of NSAR1 *A. oryzae*.

Figure 66: Control experiment testing toxicity of methylphloroacetophenone and usnic acid on *A. oryzae*.

Figure 67: Outline of procedure for testing heterologous transcription of lichen genes in *A. oryzae*, and interpretation of results by agarose electrophoresis.

Figure 68: Agarose gels demonstrating transcription of *mpas* and *mpao* in *A. oryzae* and in *C. uncialis*.

Figure 69: Schematic indicating location of introns in *mpas* and *mpao*.

Figure 70: The incubation procedure that was adopted for heterologous expression trials.

Figure 71: Mock extraction trials of methylphloroacetophenone and usnic acid in media with and without NSAR1 *A. oryzae*.

Figure 72: HPLC trace of *A. oryzae* transformed with *fu-oas* demonstrating *de novo* orsellinic acid biosynthesis in this fungus.

Figure 73: HPLC trace of *A. oryzae* transformed with *pe-6msas* demonstrating *de novo* 6-methylsalicylic acid synthase in this fungus.

Figure 74: Western blots of protein extracts of *A. oryzae* using a polyhistidine antibody.

Figure 75: An excerpt of a two-gene nBLAST alignment between the native *mpas* (query sequence) and the codon-optimized *mpas* (subject sequence).

Figure 76: Functional heterologous expression trials of the putative usnic acid-producing genes methylphloroacetophenone synthase (*mpas*) and methylphloroacetophenone synthase (*mpao*).

Figure 77: Outline of method of creating ‘cassette’ containing a total of five lichen genes, for heterologous expression trials in *A. oryzae*.

Figure 78: Mock extraction trial of 6-hydroxymellein in media with and without *A. oryzae*.

Figure 79: Functional heterologous expression trials of 6-hydroxymellein.

Figure 80: Functional heterologous expression trials of orsellinic acid.

Figure 81: Functional heterologous expression trials of 6-methylsalicylic acid.

Figure 82: Heterologous expression trials of two type III PKS from *C. uncialis*.

List of Supporting Information Items

(In order of appearance, beginning on page 194)

Table S1: Primers used throughout this work.

Table S2: Plasmids used throughout this work

Table S3: The 20 highest BLASTn similarity results for the internal transcribed spacer (ITS) sequence of native lichen specimen used in the study

Table S4: The 20 highest BLASTn similarity results for the internal transcribed spacer (ITS) sequence of axenic sub-cultured fungal partner of the lichen specimen used in the study

Figure S1: Phylogenetic tree constructed from translated amino acid sequences of the KS domain of the five non-reducing PKS genes possessing C-methyltransferase domains.

Figure S2: Phylogenetic tree constructed from translated amino acid sequences of the CLC domain of the five non-reducing PKS genes possessing C-methyltransferase domains.

Figure S3: Phylogenetic tree of the monooxygenase (6-hydroxymellein oxidase) from *Cladonia uncialis*.

Figure S4: Phylogenetic tree of the ketoreductase-dehydratase-like (KR-DH) peptide (2,3-dehydro-6-hydroxymellein reductase) from *Cladonia uncialis*.

Figure S5: Phylogenetic tree of the polyketide synthase from *Cladonia uncialis*.

Table S5: BLAST statistics for Cluster 1.

Table S6: BLAST statistics for Cluster 2.

Table S7: BLAST statistics for Cluster 3.

Table S8: BLAST statistics for Cluster 4.

Table S9: BLAST statistics for Cluster 5.

Table S10: BLAST statistics for Cluster 6.

Table S11: BLAST statistics for Cluster 7.

Table S12: BLAST statistics for Cluster 8.

Table S13: BLAST statistics for Cluster 9.

Table S14: BLAST statistics for Cluster 10.

Table S15: BLAST statistics for Cluster 11.

Table S16: BLAST statistics for Cluster 12.

Table S17: BLAST statistics for Cluster 13.

Table S18: BLAST statistics for Cluster 14.

Table S19: BLAST statistics for Cluster 15.

Table S20: BLAST statistics for Cluster 16.

Table S21: BLAST statistics for Cluster 17.

Table S22: BLAST statistics for Cluster 18.

Table S23: BLAST statistics for Cluster 19.

Table S24: BLAST statistics for Cluster 20.

Table S25: BLAST statistics for Cluster 21.

Table S26: BLAST statistics for Cluster 22.

Table S27: BLAST statistics for Cluster 23.

Table S28: BLAST statistics for Cluster 24.

Table S29: BLAST statistics for Cluster 25.

Table S30: BLAST statistics for Cluster 26.

Table S31: BLAST statistics for Cluster 27.

Table S32: BLAST statistics for Cluster 28.

Table S33: BLAST statistics for Cluster 29.

Table S34: BLAST statistics for Cluster 30.

Table S35: BLAST statistics for Cluster 31.

Table S36: BLAST statistics for Cluster 32.

Table S37: BLAST statistics for Cluster 33.

Table S38: BLAST statistics for Cluster 34.

Table S39: BLAST statistics for Cluster 35.

Table S40: BLAST statistics for Cluster 36.

Table S41: BLAST statistics for Cluster 37.

Table S42: BLAST statistics for Cluster 38.

Table S43: BLAST statistics for Cluster 39.

Table S44: BLAST statistics for Cluster 40.

Table S45: BLAST statistics for Cluster 41.

Table S46: BLAST statistics for Cluster 42.

Table S47: BLAST statistics for Cluster 43.

Table S48: BLAST statistics for Cluster 44.

Table S49: BLAST statistics for Cluster 45.

Table S50: BLAST statistics for Cluster 46.

Table S51: BLAST statistics for Cluster 47.

Table S52: BLAST statistics for Cluster 48.

Table S53: Summary of contigs containing *Cladonia uncialis* biosynthetic genes.

Figure S6: Phylogenetic relationship between the putative 4-*O*-demethyl sphaerophorin synthase of *Cladonia uncialis* and a genetically similar gene encoding a putative 4-*O*-demethyl sphaerophorin synthase that is proposed to be a part of the grayanic acid biosynthetic gene cluster of *Cladonia grayi*.

Figure S7: Phylogenetic relationship between the putative 4-*O*-demethyl sphaerophorin oxidase of *Cladonia uncialis* and a genetically similar gene encoding a putative 4-*O*-demethyl sphaerophorin oxidase that is proposed to be a part of the grayanic acid biosynthetic gene cluster of *Cladonia grayi*.

Figure S8: Phylogenetic relationship between the putative 4-*O*-demethyl-grayanic acid *O*-methyltransferase of *Cladonia uncialis* and a genetically similar gene encoding a putative 4-*O*-demethyl-grayanic acid *O*-methyltransferase that is proposed to be a part of the grayanic acid biosynthetic gene cluster of *Cladonia grayi*.

Figure S9: Phylogenetic relationship between a putative dehydroprobetaenone reductase of *Cladonia uncialis* and a genetically similar gene encoding a dehydroprobetaenone reductase that is part of the betaenone biosynthetic gene cluster of *Phoma betae*.

Figure S10: Phylogenetic relationship between a putative betaenone C oxidase of *Cladonia uncialis* and a genetically similar gene that is proposed to encode a betaenone C oxidase that participates in betaenone biosynthesis in *Phoma betae*.

Figure S11: Phylogenetic relationship between a putative dehydroprobetaenone oxidase of *Cladonia uncialis* and a genetically similar gene encoding a dehydroprobetaenone oxidase that is part of the betaenone biosynthetic gene cluster of *Phoma betae*.

Figure S12: Phylogenetic relationship between a putative *trans*-enoylreductase of *Cladonia uncialis* and a genetically similar gene encoding a *trans*-enoylreductase that is part of the betaenone biosynthetic gene cluster of *Phoma betae*.

Figure S13: Phylogenetic relationship between a putative dehydroprobetaenone synthase of *Cladonia uncialis* and a genetically similar gene encoding a dehydroprobetaenone synthase that is part of the betaenone biosynthetic gene cluster of *Phoma betae*.

Figure S14: Phylogenetic relationship between a putative ABC transporter of *Cladonia uncialis* and a genetically similar gene encoding an ABC transporter found with the patulin biosynthetic gene cluster of *Aspergillus clavatus*.

Figure S15: Phylogenetic relationship between a putative isoeopoxydon dehydrogenase of *Cladonia uncialis* and a genetically similar gene encoding an isoeopoxydon dehydrogenase that is part of the patulin biosynthetic gene cluster of *Aspergillus clavatus*.

Figure S16: Phylogenetic relationship between a putative *m*-cresol hydroxylase of *Cladonia uncialis* and a genetically similar gene encoding *m*-cresol hydroxylase that is part of the patulin biosynthetic gene cluster of *Aspergillus clavatus*.

Figure S17: Phylogenetic relationship between a putative isoamyl alcohol oxidase of *Cladonia uncialis* and a genetically similar gene encoding an isoamyl alcohol oxidase found with the patulin biosynthetic gene cluster of *Aspergillus clavatus*.

Figure S18: Phylogenetic relationship between a putative 6-methylsalicylic acid decarboxylase of *Cladonia uncialis* and a genetically similar gene encoding 6-methylsalicylic acid decarboxylase that is part of the patulin biosynthetic gene cluster of *Aspergillus clavatus*.

Figure S19: Phylogenetic relationship between a putative C6 transcription factor of *Cladonia uncialis* and a genetically similar gene encoding a C6 transcription factor found with the patulin biosynthetic gene cluster of *Aspergillus clavatus*.

Figure S20: Phylogenetic relationship between a putative 6-methylsalicylic acid synthase of *Cladonia uncialis* and a genetically similar gene encoding 6-methylsalicylic acid synthase that is part of the patulin biosynthetic gene cluster of *Aspergillus clavatus*.

Figure S21: Phylogenetic relationship between a gene of unknown function (no. 1) found in *Cladonia uncialis* and a genetically similar gene of unknown function found with the patulin biosynthetic gene cluster of *Aspergillus clavatus*.

Figure S22: Phylogenetic relationship between a putative *m*-hydroxybenzyl alcohol hydroxylase of *Cladonia uncialis* and a genetically similar gene encoding *m*-hydroxybenzyl alcohol hydroxylase that is part of the patulin biosynthetic gene cluster of *Aspergillus clavatus*.

Figure S23: Phylogenetic relationship between a putative carboxylesterase of *Cladonia uncialis* and a genetically similar gene encoding a carboxylesterase found with the patulin biosynthetic gene cluster of *Aspergillus clavatus*.

Figure S24: Phylogenetic relationship between a putative MFS transporter of *Cladonia uncialis* and a genetically similar gene encoding a MFS transporter found with the patulin biosynthetic gene cluster of *Aspergillus clavatus*.

Figure S25: Phylogenetic relationship between a gene of unknown function (no. 2) found in *Cladonia uncialis* and a genetically similar gene of unknown function found with the patulin biosynthetic gene cluster of *Aspergillus clavatus*.

Figure S26: Phylogenetic relationship between a putative GMC oxidoreductase of *Cladonia uncialis* and a genetically similar gene encoding a GMC oxidoreductase found with the patulin biosynthetic gene cluster of *Aspergillus clavatus*.

Figure S27: Phylogenetic relationship between a putative polyketide synthase of *Cladonia uncialis* and a genetically similar gene encoding a polyketide synthase that is part of the azaphilone biosynthetic gene cluster of *Monascus pilosus*.

Figure S28: Phylogenetic relationship between a putative oxidoreductase (no. 1) of *Cladonia uncialis* and a genetically similar gene that is proposed to encode an oxidoreductase that participates in azaphilone biosynthesis in *Monascus pilosus*.

Figure S29: Phylogenetic relationship between a gene of unknown function (no. 1) of *Cladonia uncialis* and a genetically similar gene of unknown function of which its product is proposed to participate in azaphilone biosynthesis in *Monascus pilosus*.

Figure S30: Phylogenetic relationship between a gene of unknown function (no. 2) of *Cladonia uncialis* and a genetically similar gene of unknown function of which its product is proposed to participate in azaphilone biosynthesis in *Monascus pilosus*.

Figure S31: Phylogenetic relationship between a putative efflux transporter of *Cladonia uncialis* and a genetically similar gene that is proposed to be an efflux transporter associated with the azaphilone biosynthetic gene cluster of *Monascus pilosus*.

Figure S32: Phylogenetic relationship between a putative transcriptional regulator of *Cladonia uncialis* and a genetically similar gene that is proposed to be a transcriptional regulator that is associated with the azaphilone biosynthetic gene cluster of *Monascus pilosus*.

Figure S33: Phylogenetic relationship between a putative hydroxylase of *Cladonia uncialis* and a genetically similar gene encoding a hydroxylase that is part of the azaphilone biosynthetic gene cluster of *Monascus pilosus*.

Figure S34: Phylogenetic relationship between a putative acyltransferase of *Cladonia uncialis* and a genetically similar gene encoding an acyltransferase that is part of the azaphilone biosynthetic gene cluster of *Monascus pilosus*.

Figure S35: Phylogenetic relationship between a putative amino-oxidase of *Cladonia uncialis* and a genetically similar gene encoding an amino-oxidase that is part of the azaphilone biosynthetic gene cluster of *Monascus pilosus*.

Figure S36: Phylogenetic relationship between a putative oxidoreductase (no. 2) of *Cladonia uncialis* and a genetically similar gene that is proposed to encode an oxidoreductase that participates in azaphilone biosynthesis in *Monascus pilosus*.

Figure S37: Phylogenetic relationship between a putative *trans*-enoylreductase of *Cladonia uncialis* and a genetically similar gene encoding a *trans*-enoylreductase that is part of the cytochalasin biosynthetic gene cluster of *Aspergillus clavatus*.

Figure S38: Phylogenetic relationship between a putative polyketide synthase-non-ribosomal peptide synthetase (NRPS-PKS) of *Cladonia uncialis* and a genetically similar gene encoding a NRPS-PKS that is part of the cytochalasin biosynthetic gene cluster of *Aspergillus clavatus*.

Figure S39: Phylogenetic relationship between a putative 6-*O*-demethylfusarubinaldehyde synthase of *Cladonia uncialis* and a genetically similar gene encoding a 6-*O*-demethylfusarubinaldehyde synthase that is part of the fusarubin biosynthetic gene cluster of *Fusarium fujikuroi*.

Figure S40: Phylogenetic relationship between a putative 6-*O*-demethylfusarubinaldehyde-*O*-methyltransferase of *Cladonia uncialis* and a genetically similar gene encoding a 6-*O*-demethylfusarubinaldehyde-*O*-methyltransferase that is part of the fusarubin biosynthetic gene cluster of *Fusarium fujikuroi*.

Figure S41: Phylogenetic relationship between a putative multi-functional monooxygenase of *Cladonia uncialis* and a genetically similar gene that is proposed to encode a multi-functional monooxygenase that participates in fusarubin biosynthesis in *Fusarium fujikuroi*.

Figure S42: Phylogenetic relationship between a putative short-chain dehydrogenase / reductase of *Cladonia uncialis* and a genetically similar gene that is proposed to encode a short-chain dehydrogenase/reductase that participates in fusarubin biosynthesis in *Fusarium fujikuroi*.

Figure S43: Phylogenetic relationship between a putative endocrocin anthrone monooxygenase of *Cladonia uncialis* and a genetically similar gene encoding an endocrocin anthrone monooxygenase that is part of the pestheic acid biosynthetic gene cluster of *Pestalotiopsis fici*.

Figure S44: Phylogenetic relationship between a putative atrochrysone carboxylic acid thioesterase of *Cladonia uncialis* and a genetically similar gene encoding atrochrysone carboxylic acid thioesterase that is part of the pestheic acid biosynthetic gene cluster of *Pestalotiopsis fici*.

Figure S45: Phylogenetic relationship between a putative atrochrysone carboxylic acid synthase of *Cladonia uncialis* and a genetically similar gene encoding atrochrysone carboxylic acid synthase that is part of the pestheic acid biosynthetic gene cluster of *Pestalotiopsis fici*.

Figure S46: Phylogenetic relationship between a putative 5-methylorsellinic acid synthase of *Cladonia uncialis* and a genetically similar gene encoding a 5-methylorsellinic acid synthase that is part of the mycophenolic acid biosynthetic gene cluster of *Penicillium brevicompactum*.

Figure S47: Phylogenetic relationship between a putative 5-methylorsellinic acid oxidase of *Cladonia uncialis* and a genetically similar gene encoding a 5-methylorsellinic acid oxidase that is part of the mycophenolic acid biosynthetic gene cluster of *Penicillium brevicompactum*.

Figure S48: Phylogenetic relationship between a putative 4,6-dihydroxy-2-hydroxymethyl-3-methylbenzoic acid lactamase of *Cladonia uncialis* and a genetically similar gene encoding a 4,6-dihydroxy-2-hydroxymethyl-3-methylbenzoic acid lactamase that is part of the mycophenolic acid biosynthetic gene cluster of *Penicillium brevicompactum*.

Table S54: nBLAST alignment of ITS sequence of unknown species of non-lichenizing fungus identified in this study with entries deposited in GenBank.

Table S55: nBLAST alignment of DNA sequence of *6msas* gene identified in *Penicillium sp.* with entries in GenBank.

Figure S49: Control experiment verifying presence of *N*-terminal His(6)-tagged PE-6MSAS.

Figure S50: Coomassie stain of Western blot membrane, performed after blot imaging, to confirm effective transfer of high molecular weight proteins.

Figure S51: Coomassie stain of SDSPAGE gel, performed after transfer to Western blot membrane, to verify that a suitable concentration of protein was loaded into the gel.

Sequence S1: DNA sequence of *mpas* (5' – 3').

Sequence S2: DNA sequence of *mpao* (5' – 3').

Sequence S3: Results of amplification and sequencing of ITS region of fungus that we have identified as member of the *Penicillium* genus (5' – 3').

Figure S52: Uncropped agarose gels, pertaining to Figure 68 (see main text).

Figure S53: HPLC traces of a total of nine replicates of organic extracts of NSAR1 *A. oryzae* cultures (not transformed with plasmids), following incubation in Czapek-Dox (with added starch) for five days.

Figure S54: HPLC traces of organic extracts of 11 *A. oryzae* colonies that were transformed with plasmid vectors, following incubation in Czapek-Dox (with added starch) for five days.

Figure S55: HPLC traces of organic extracts of ten *A. oryzae* colonies that were transformed with codon-optimized *mpas*, codon-optimized *mpao*, and *pptase* (from *C. uncialis*), following incubation in starch-based Czapek-Dox media for five days.

Figure S56: HPLC traces of organic extracts of ten *A. oryzae* colonies that were transformed with codon-optimized *mpas*, codon-optimized *mpao*, and ‘cassette’ (cassette comprises *pptase*, *cpr1*, *cpr2*, *b5r1*, and *b5r2*, all originating from *C. uncialis*) following incubation in Czapek-Dox (with added starch) for five days.

Figure S57: HPLC traces of organic extracts of eight *A. oryzae* colonies that were transformed with *6hms*, *krdh*, and *pptase* (all genes originating from *C. uncialis*), following incubation in Czapek-Dox (with added starch) for five days.

Figure S58: HPLC traces of organic extracts of eight *A. oryzae* colonies that were transformed with *oas* and *pptase* (both genes originating from *C. uncialis*), following incubation in Czapek-Dox (with added starch) for five days.

Figure S59: HPLC traces of organic extracts of eight *A. oryzae* colonies that were transformed with *gas* and *pptase* (both genes originating from *C. uncialis*), following incubation in Czapek-Dox (with added starch) for five days.

Figure S60: HPLC traces of organic extracts of eight *A. oryzae* colonies that were transformed with *6msas* and *pptase* (both genes originating from *C. uncialis*), following incubation in Czapek-Dox (with added starch) for five days.

Figure S61: HPLC traces of organic extracts of ten *A. oryzae* colonies that were transformed with *t3pks1* from *C. uncialis*, following incubation in Czapek-Dox (with added starch) for five days.

Figure S62: HPLC traces of organic extracts of ten *A. oryzae* colonies that were transformed with *t3pks3* from *C. uncialis*, following incubation in Czapek-Dox (with added starch) for five days.

Sequence S4: The codon-optimized variant of *mpas* that was commissioned for this work (5' – 3').

Sequence S5: The codon-optimized variant of *mpao* that was commissioned for this work (5' – 3').

Sequence S6: DNA sequence of *pptase* from *Cladonia uncialis* (5' – 3').

Sequence S7: DNA sequence of *cpr1* from *Cladonia uncialis* (5' – 3').

Sequence S8: DNA sequence of *cpr2* from *Cladonia uncialis* (5' – 3').

Sequence S9: DNA sequence of *b5r1* from *Cladonia uncialis* (5' – 3').

Sequence S10: DNA sequence of *b5r2* from *Cladonia uncialis* (5' – 3').

Sequence S11: DNA sequence of *6hms* from *Cladonia uncialis* (5' – 3').

Sequence S12: DNA sequence of *oas* from *Cladonia uncialis* (5' – 3').

Sequence S13: DNA sequence of *gas* from *Cladonia uncialis* (5' – 3').

Sequence S14: DNA sequence of *6msas* from *Cladonia uncialis* (5' – 3').

Sequence S15: DNA sequence of *t3pks1* from *Cladonia uncialis* (5' – 3').

Sequence S16: DNA sequence of *t3pks2* from *Cladonia uncialis* (5' – 3').

Non-IUPAC Abbreviations

6HMS:	6-hydroxymellein synthase
6MSAS:	6-methylsalicylic acid synthase
A:	Adenylation
ACP:	Acyl carrier protein
AT:	Acyltransferase
B5R:	Cytochrome B5 reductase
C:	Condensation
CLC:	Claisen cyclase
CPR:	Cytochrome p450 reductase
GAS:	Grayanic acid synthase
ITS:	Internal transcribed spacer (May also be referred to in the literature as “nuclear ribosomal internal transcribed spacer” (NRITS))
KRDH:	Ketoreductase-dehydratase-like protein
KS:	Ketosynthase
MPAO:	Methylphloroacetophenone oxidase
MPAS:	Methylphloroacetophenone synthase
MT:	C-methyltransferase (An alternative abbreviation is “cMeT”)
NRPS:	Non-ribosomal peptide synthetase
NSAR1:	Refers to <i>A. oryzae</i> strain: Δ <i>niaD</i> , Δ <i>sC</i> , Δ <i>adeA</i> , Δ <i>argB</i>
OAS:	Orsellinic acid synthase
PKS:	Polyketide synthase
PPTase:	Phosphopantetheinyl transferase
PT:	Product template
R:	Reductase
SAT:	Starter acyltransferase (May also be referred to in the literature as “starter unit-acyl carrier protein transacylase” (SAT))
T:	Thiolation (May also be referred to in the literature as “Peptidyl carrier protein” (PCP))
TE:	Thioesterase

Chapter 1

An introduction to lichens and their secondary metabolites

1.1. Introduction – Lichens and their natural products

Lichens are a subset of *Fungi* that are traditionally described as symbiotic partners of algae and cyanobacteria (Nash III, 2008). Lichens are also known to be ecological hubs of dozens of species of bacteria and yeast that live on (or even within!) the lichen thallus (Bates et al. 2011; Grube et al. 2009; Spribille et al. 2016; Cernava et al. 2016). Lichenization first occurred around 600 million years ago (Yuan et al. 2005). About 20 percent of the five million species of fungi are lichens (Blackwell, 2011; Lutzoni & Miadlikowska, 2009). The fungal partner, known as the ‘mycobiont’, provides protection and an anchor to the ground from which the lichen may draw water and nutrients. The algal or cyanobacterial partner, known as the ‘photobiont’, generates food by photosynthesis (**Figure 1**). By convention, the taxonomic name of the fungal partner is also the name of the lichen partnership. For example, the name “*Cladonia uncialis*” may refer to the fungus as it exists within in its lichenized partnership in natural settings, or it may refer to this same fungus if it was isolated from its algal partner within the laboratory and no longer in a lichenized state. There are about one hundred species of photobionts that are known to associate with the more than ten thousand identified lichen-forming fungal species (Lutzoni & Miadlikowska, 2009).

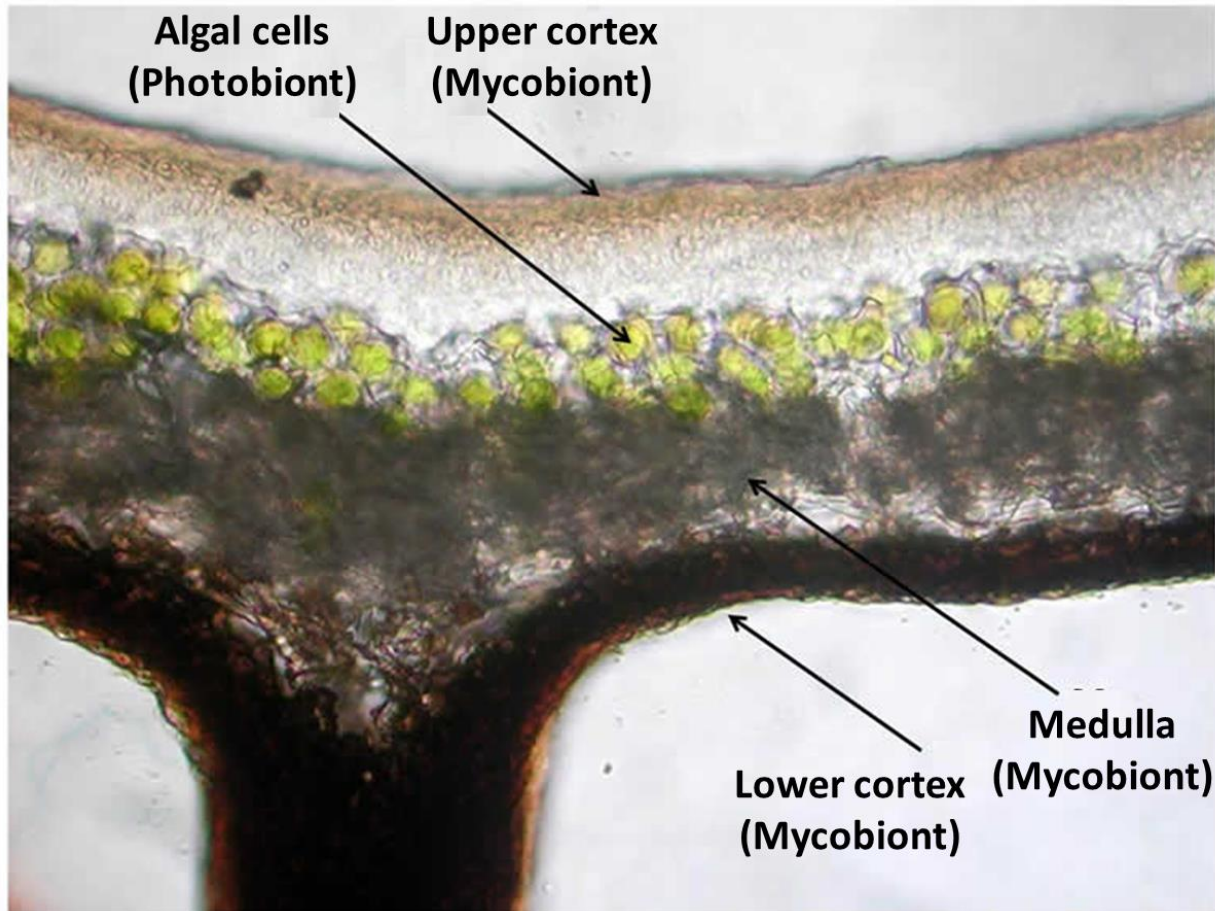


Figure 1: Cross-section of the lichen thallus of *Parmelia saxatilis*. Adapted with minor modification from the Royal Botanical Garden of Edinburgh, available at: <https://rbg-web2.rbge.org.uk/lichen/> (Accessed 12 May 19).

Lichens can be found occupying every terrestrial habitat on Earth including Antarctica (Øvstedal & Lewis-Smith, 2001). This fact is owed to their ability to colonize virtually any surface, to tolerate extreme environments, and to reproduce by both sexual and asexual means. For example, lichens can be found growing on rocks, wood, soil, mosses, leaves, concrete, glass, metals, and plastics (Lutzoni & Miadlikowska, 2009). Experiments on the International Space Station have demonstrated that lichens can withstand an extended exposure (18 months!) to the vacuum and radiation of space (Brandt et al. 2014). Lichens can also survive simulated Martian conditions, leading astrobiologists to speculate that lichens could be used as a photosynthetic primer for the terraforming of Mars (de Vera et al. 2010). Lichens may reproduce sexually and asexually. Sexual reproduction requires the release of spores that must germinate in the presence

of a compatible photobiont, whereas asexual reproduction requires the dispersal of fragments of lichenized tissue (Lutzoni & Miadlikowska, 2009).

Recreating symbioses *in vitro* is a pre-condition to an advanced understanding of how lichenization occurs. A longstanding problem in lichenology is the inability of scientists to recreate the symbioses that are found in nature by combining mycobionts and photobionts together within the laboratory. This problem suggested the existence of some factor that was absent from these experiments that would presumably be required for symbiosis to occur (Spribille, 2018). An advancement in the understanding of this problem was recently provided with the discovery that many lichens are symbiotic relationships of *three* species. This third species, a *Basidiomycetes* yeast, was found to be living within the thallus of common lichens (Spribille et al. 2016). This discovery was provided using fluorescence hybridization of *Basidiomycetes* rRNA (**Figure 2**). This experiment was necessary because microscopy and culturing attempts failed to detect these yeasts (Spribille et al. 2016). This may explain how a third species eluded detection by lichenologists even though the lichen symbiosis was discovered more than 140 years ago. An interesting observation is the correlation between the density of the yeast within the lichenized tissue and the presence or absence of vulpinic acid, a naturally-occurring molecule that is responsible for a yellowing of the lichen tissue (**Figure 2**). Though more experiments are clearly required, it is possible that a role of these yeasts is to modulate the production of naturally-occurring lichen molecules such as vulpinic acid (Spribille et al. 2016).

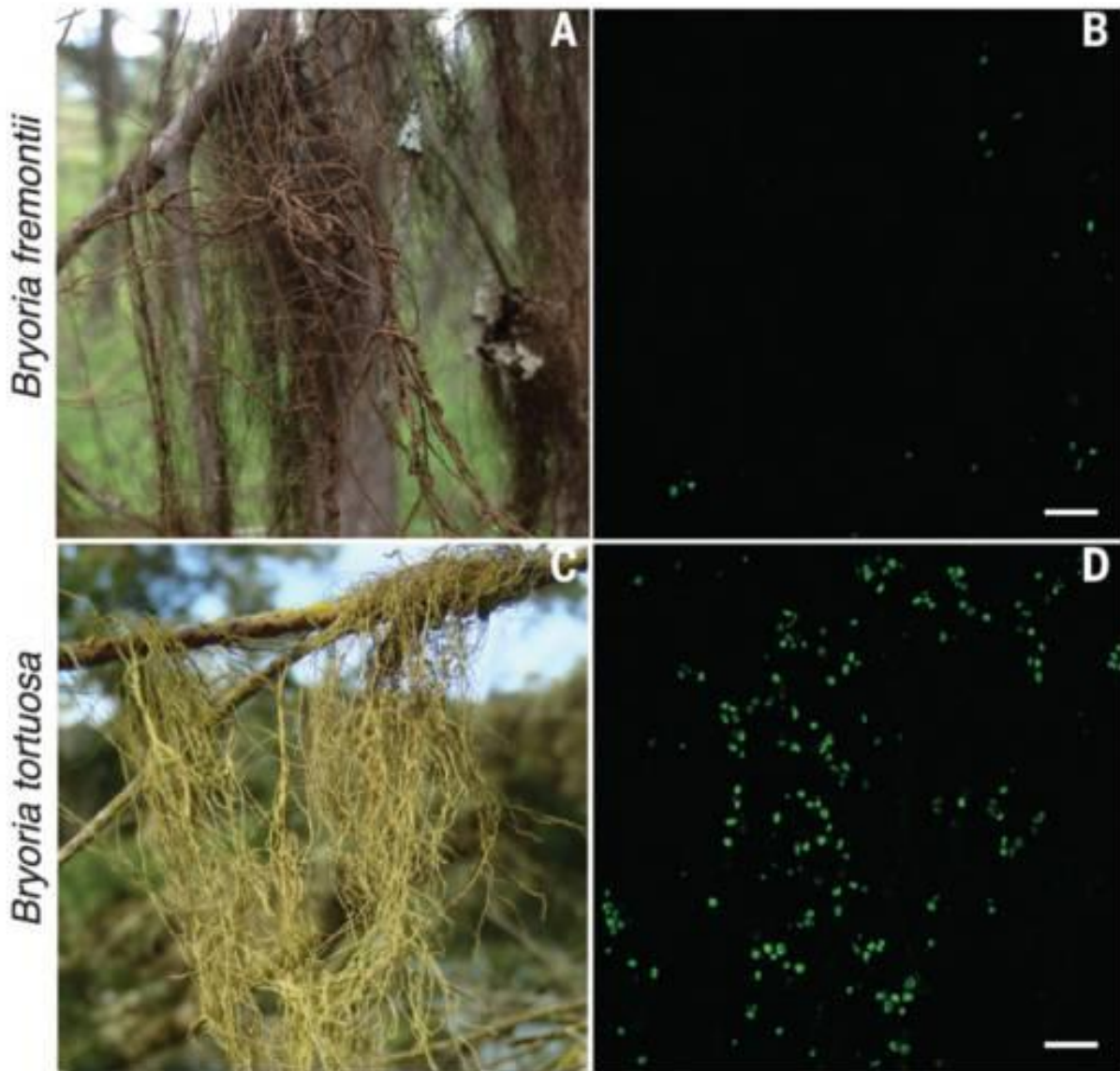


Figure 2: Fluorescence hybridization of *Cyphobasidiales* yeasts in *B. fremontii* and *B. tortuosa*. (A) Vulpinic acid-poor *B. fremontii* with (B) low-density yeast; (C) Vulpinic acid-rich *B. tortuosa* with (D) high-density yeast. Reproduced from Spribille et al. 2016, with permission from the *American Association for the Advancement of Science* (AAAS).

Primary metabolites are molecules that originate from biogenic sources and are required by organisms to live, grow, reproduce, and otherwise carry-out the customary functions of life. Examples include proteinogenic amino acids, carbohydrates, lipids, and nucleotides. In contrast, secondary metabolites are molecules that are produced by organisms that provide a competitive advantage with respect to surviving and proliferating within an ecosystem. The toxic secretions

and aposematic pigments produced by some species of frogs are examples of secondary metabolites because these molecules are not required to carry-out obligate biochemical functions but instead promote survival, in this case, by discouraging predation.

In lichens, these secondary metabolites are usually small phenolic or aliphatic molecules and can comprise between 0.1 to 30 percent of the dry weight of lichens (Molnár & Farkas, 2010). The brilliant colours that distinguish lichens from the surrounding foliage arise from the accumulation of secondary metabolites. More than one thousand secondary metabolites have been identified from lichens (Calcott et al. 2018). Selected examples of lichen secondary metabolites are shown (**Figure 3**). Some of these secondary metabolites possess intriguing biological activities and are being investigated as drug candidates or commercial products (Nguyen et al. 2013; Shrestha & St. Clair, 2013; Zambare & Christopher, 2012; Shukla et al. 2010). For example, five lichen compounds - physodic acid, norlobaric acid, salazinic acid, parellic acid, and virensic acid – inhibit HIV-1 integrase with an IC_{50} of 100 μ M or less (Neamati et al. 1997). Three lichen compounds - pannarin, chloropannarin, and sphaerophorin – are also known for their cytotoxic effects against rat lymphocytes (Correch   et al. 2002). The antiproliferative activity of three lichen secondary metabolites - gyrophoric acid, usnic acid, and diffractaic acid - have been reported to have an IC_{50} of 2.6 μ M or less on human keratinocyte cells (Kumar & M  ller, 1999b). Administration of 100 mg/kg of diffractaic acid to rats possessing indomethacin-induced gastric lesions reduced the size of these lesions by 96.7 percent (Bayir et al. 2006). Leukotriene B₄ is an inflammatory mediator that is known to promote insulin resistance in obese mice (Li et al. 2015). The biosynthesis of Leukotriene B₄ appears to be inhibited by a non-redox mechanism by atranorin, diffractaic acid, and protolichesterinic acid (Kumar & M  ller, 1999a). Derivatives of orsellinic acid isolated from Oakmoss lichen possessed nematocidal activity with evident lethality

occurring at 13 μM (Ahad et al. 1991). Barbatic acid is known to inactivate photosystem II in plants, suggesting that this natural product could be developed into a pesticide (Takahagi et al. 2006). These are only a few of the many examples of the biological effects of lichen molecules.

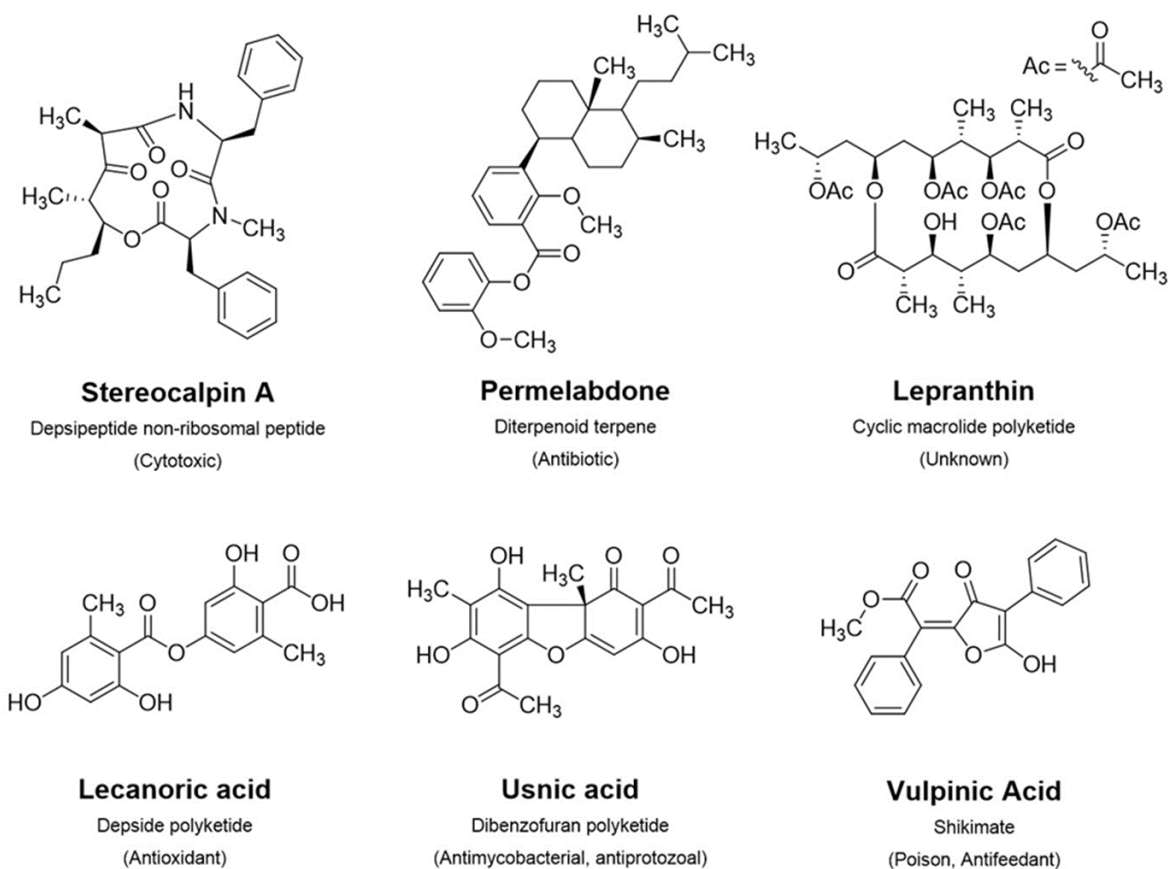


Figure 3: Selected lichen secondary metabolites and their bioactive properties, if known: Stereocalpin A, from *Stereocaulon alpinum* (Seo et al. 2008); Permellabdone, from *Parmelia perlata* (Abdullah et al. 2007); Lepranthin, from *Arthonia impolita* (Polborn et al. 2014); Lecanoric acid, from *Umbilicaria antarctica* (Luo et al. 2009); Usnic acid, from *Usnea longissima* (Araújo et al. 2015); Vulpinic acid, from *Letharia vulpina* (Stephenson & Rundel, 1979).

Humans for millennia have been well-aware of the biological properties of Nature's secondary metabolome. One of the earliest examples of a proto-pharmaceutical is the use of salicylate-rich plants by ancient Egyptians and Greeks to alleviate pain and inflammation (Crawford, 2015). Historians would be hard-pressed to identify any ancient civilization that did not discover the inebriating and analgesic effects of alcohol. Lichen materials have been used as

remedies to ailments (Crawford, 2015). *Usnea longissima*, of which its eponymous antibiotic usnic acid is its major chemical substituent, was commonly used as an antiseptic. Among the First Nations of Canada it was common practice to wrap moistened *U. longissima* around a wound to prevent an infection from setting in. A similar practice was observed in India and in the Himalayas to prevent infection arising from open bone fractures (Crawford, 2015). Sometimes the applications of lichens diverged among cultures: *Cladonia rangiferina* was used by the Cree (Quebec region) to treat inflammation. In Finland, the broth of boiled *C. rangiferina* was consumed to treat coughs and tuberculosis. The ancient Chinese drank or applied a powdered form of this lichen to affected areas to treat a host of conditions (Crawford, 2015). *Letharia vulpine*, known as the ‘wolf lichen’, was used to poison wolves in Scandinavia, prepare poison arrows by the Achomawi nation (California region), and to treat festering wounds in ancient China (Crawford, 2015)¹.

1.2. Biosynthesis of secondary metabolites

To comprehensively review how Nature produces its secondary metabolome would be a Herculean task requiring the filling of several volumes. Instead, I will provide a brief but focused overview of the chemistry that is pertinent to navigating through the later chapters of this thesis.

¹Finding the secondary metabolites that produce biological effects reported by traditional healers is a field of study known as ‘ethnopharmacology’. Bioactive secondary metabolites have been discovered through collaborations between traditional healers and chemists. For example, extracts from the leaves of a tree (*Azadirachta indica*) have been used as a traditional remedy to alleviate gastrointestinal complications associated with malaria. Gedunin has been identified as the secondary metabolite that produces the biological effects attributed to these extracts (Mackinnon et al. 1997; Omar et al. 2003). Extracts from this tree also appear to possess activity against HIV, with the effect of increasing the prevalence of CD4⁺ immune cells during an HIV infection (Mbah et al. 2007; Udeinya et al. 2004). Fractionated extracts of these tree leaves are now packaged as 250 mg capsules and are sold in Nigeria as a crude drug under the brand name of IRACARP (Anyahie, 2009). I am well-acquainted with IRACARP: I was given this drug to alleviate a bout of malaria that I contracted during a field study in Nigeria in 2010, to enormous relieving effect. It was this experience that inspired me to explore natural products chemistry.

1.2.1. Polyketides

Polyketides are polymers of acyl units that are assembled to form phenolic or aliphatic structures (Weissman, 2009). It was the study of polyketide biosynthesis that allowed natural products chemistry to emerge as a mature scientific discipline (Birch & Donovan, 1953; Birch et al. 1955). Polyketides are the most extensively studied class of secondary metabolite (Staunton & Weissman, 2001). Polyketide biosynthetic pathways are shown in chemistry classrooms as examples of “Nature at work” (Puel et al. 2010; Roze et al. 2013; Campbell & Vederas, 2010). Polyketides comprise numerous pharmaceuticals, for example, the anti-hypercholesterolemic lovastatin (Campbell & Vederas, 2010), the antibacterial erythromycin (Zhang et al. 2010), the antimycobacterial rifamycin (Qi et al. 2018).

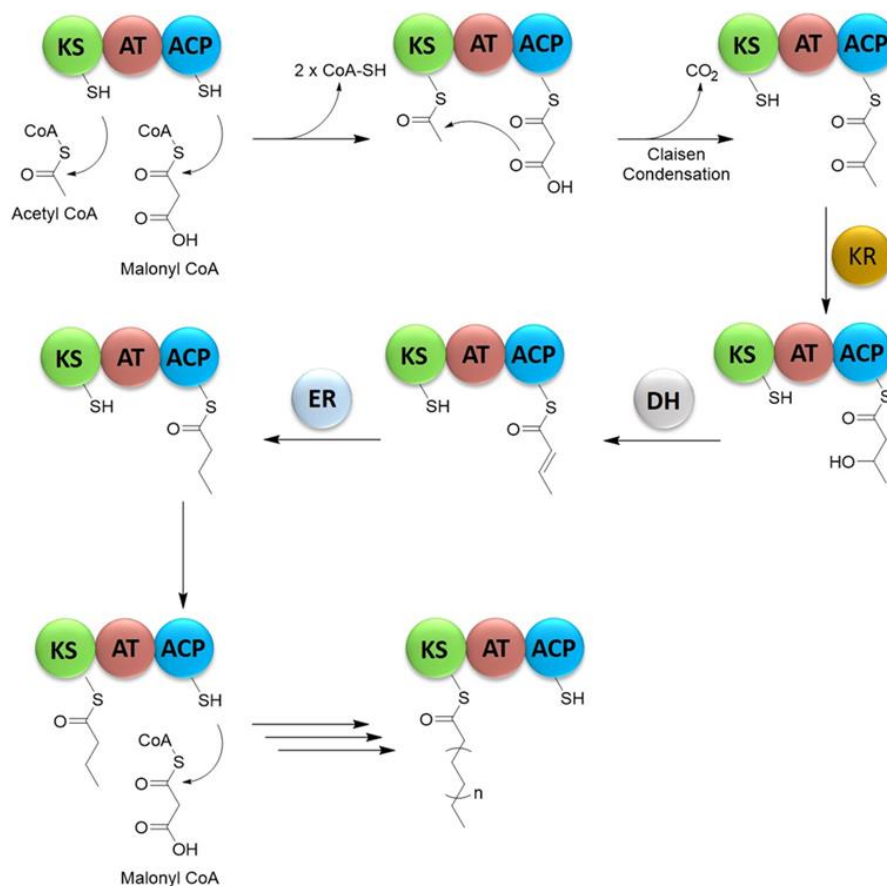


Figure 4: General scheme of polyketide elongation by polyketide synthases (PKS). *Abbreviations:* Ketosynthase (KS), Acyltransferase (AT), Acyl carrier protein (ACP), Ketoreductase (KR), Dehydratase (DH), Enoylreductase (ER).

Polyketides are produced by polyketide synthases (PKS), multi-catalytic domain enzymes that facilitate carbon-carbon bond formation through decarboxylative Claisen condensation reactions (Hertweck, 2009). The basic principle of polyketide assembly by PKS is illustrated (**Figure 4**). Three catalytic domains are required for polyketide elongation to occur. The ketosynthase (KS) domain facilitates the carbon-carbon bond-forming step essential to polyketide chemistry. The acyl carrier protein (ACP) domain carries polyketide intermediates to and from other catalytic domains. The ACP domain has a wide berth of movement within the PKS superstructure and possesses a phosphopantetheine ‘arm’ that carries intermediates to the other catalytic domains. The acyltransferase (AT) domain serves as a ‘gatekeeper’ that recruits the appropriate priming or extension unit to the KS and ACP domains (Park et al. 2014).

To initiate polyketide biosynthesis, the PKS must first be ‘primed’. The AT domain recruits an initial acyl unit (typically acetyl CoA) to the KS domain. An extension unit (typically malonyl CoA) is then recruited by the AT domain to the ACP domain. Numerous exceptions to the use of acetyl CoA and malonyl CoA as priming and extension units have been observed (Chan et al. 2009; Zhang & Tang, 2007). The KS domain performs a decarboxylative Claisen condensation to extend the polyketide chain by two carbon atoms. Transferring this intermediate to the KS domain allows the AT domain to recruit another malonyl CoA to the ACP. This process can be repeated more than once to extend the backbone chain length by two carbon atoms per iteration (**Figure 4**).

Other catalytic domains may also be present within a PKS. These domains are not strictly required for chain elongation to occur but provide chemical modifications. Three reducing domains - ketoreductase (KR), dehydratase (DH), and enoylreductase (ER) – are available to progressively reduce a carbonyl group from a ketone residue down to a fully reduced carbon

(Keatinge-Clay & Stroud, 2006; Akey et al. 2010; Kwan & Leadlay, 2010). A C-methyltransferase (MT) domain (not shown in **Figure 5**) is another ‘optional’ domain and is responsible for adding methyl groups to the α -carbons of the polyketide backbone (Skiba et al. 2016). This methyl group comes from S-adenosyl methionine (SAM) (Skiba et al. 2016). A PKS that produces highly reduced polyketides could be conceived as similar in function to fatty acid synthases, the group of enzymes responsible for producing fatty acids (Smith & Tsai, 2007).

Once the polyketide has been fully extended, the reaction must be terminated. Three possible terminal domains are available to terminate the reaction (**Figure 5**). The most common is the thioesterase (TE), which removes the polyketide from the ACP domain by thioester hydrolysis, leaving a terminal carboxylic acid on the polyketide (Du & Lou, 2010). The second domain, appearing less commonly than TE domains, is a reductase (R) domain. This operates similarly to the TE domain but uses a reducing equivalent of NADPH to produce an aldehyde (Bailey et al. 2007). The third domain, also appearing less commonly than TE domains, is a Claisen cyclase (CLC; may also be abbreviated in the literature as CYC). This domain simultaneously cleaves and cyclizes polyketides by Claisen condensation (Korman et al. 2010). An interesting implication of these different termination methods is that by folding polyketide intermediates in different ways it is possible to produce different molecules from the same substrate (**Figure 5**). This is an important feature of polyketide biosynthesis that we will revisit in Chapter 3. It is also possible that the PKS does not possess a terminal domain. In these cases, a reductive enzyme independent from the PKS is responsible for cleaving the polyketide from the ACP domain (Xu et al. 2014). This is a common setup among PKS that perform reductive functions found in *Fungi*. The PKS that have reducing domains are colloquially referred to as

‘reducing PKS’ – these can be sub-divided into ‘highly reducing’ and ‘partially reducing’ PKS. Those that do not perform reductive functions are called ‘non-reducing PKS’ (**Figure 6**).

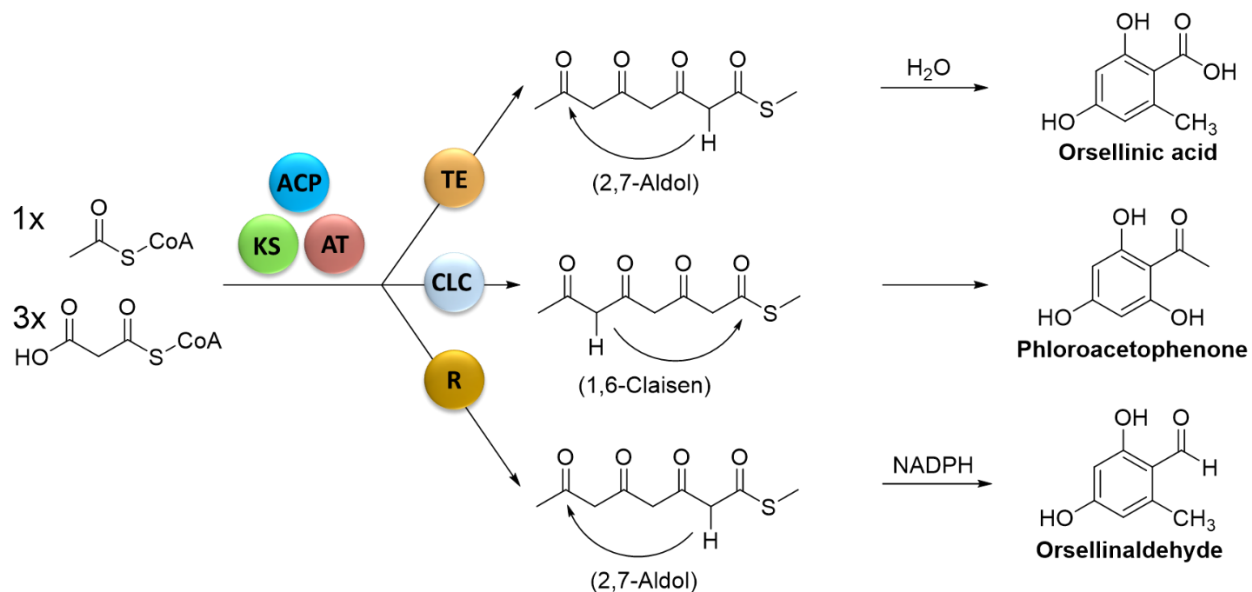


Figure 5: Termination release mechanisms by PKS. *Abbreviations:* Ketosynthase (KS), acyltransferase (AT), Acyl carrier protein (ACP), Thioesterase (TE), Claisen cyclase (CLC), Reductase (R).

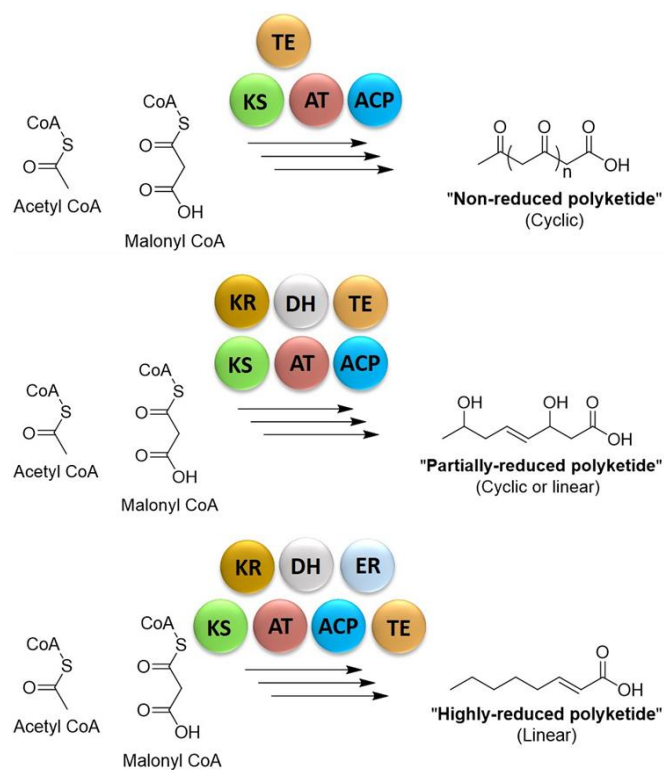


Figure 6: ‘Non-reducing’, ‘partially reducing’, and ‘highly reducing’ PKS. *Abbreviations:* Ketosynthase (KS), Acyltransferase (AT), Acyl carrier protein (ACP), Ketoreductase (KR), Dehydratase (DH), Enoylreductase (ER), Thioesterase (TE)

The deoxyerythronolide PKS, known as ‘DEBS’, provides an example of how these various components come together to produce a molecule (**Figure 7**). The domains of this PKS are arranged in a series of modules. Each module elongates the polyketide backbone by two carbon atoms in a manner analogous to a factory assembly line. In this example, methylmalonyl-CoA is used instead of malonyl-CoA for chain elongation. To prime polyketide biosynthesis, an acetyl-CoA is recruited to the first KS domain by a dedicated ACP-AT didomain located at the beginning of the assembly line. At some stages of the process, reducing domains are present to perform reductions at the carbonyl group. This is an example of a ‘modular PKS’ (Dutta et al. 2014; Robbins et al. 2016). This style of assembly is commonly observed in bacteria.

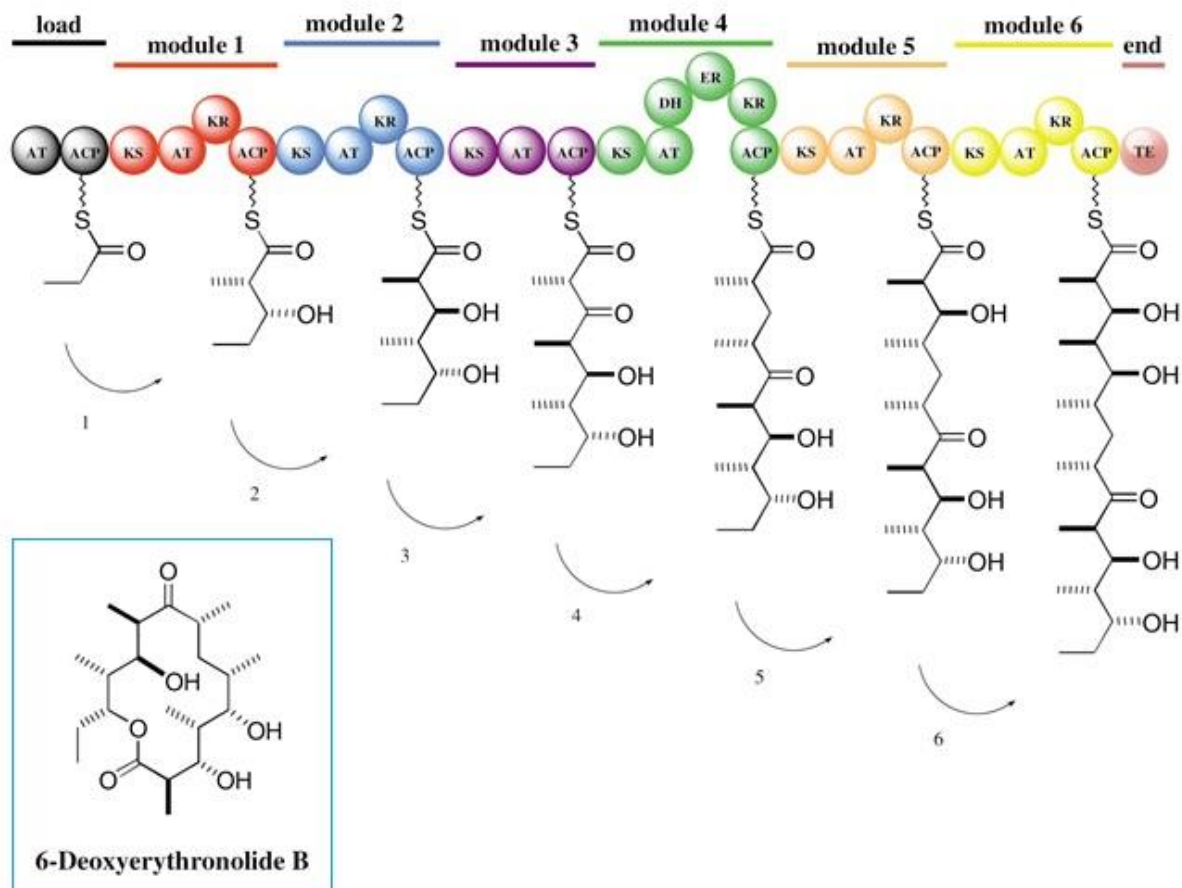


Figure 7: 6-deoxyerythronolide synthase (DEBS), a model modular PKS. Reproduced with minor modifications from Staunton & Weissman, 2001, with permission from the Royal Society of Chemistry. *Abbreviations:* Ketosynthase (KS), Acyltransferase (AT), Acyl carrier protein (ACP), Ketoreductase (KR), Dehydratase (DH), Enoylreductase (ER), Thioesterase (TE)

The 6-methylsalicylic acid synthase provides a second example of a PKS in action (**Figure 8**). No modules are present. Instead, each domain is used multiple times to elongate a polyketide through iterative condensation cycles. A starter unit acyltransferase (SAT) domain (not shown in **Figure 8**) is responsible for initiating polyketide assembly by recruiting an acetyl-CoA or other priming unit to the KS domain (Crawford et al. 2006). An additional domain known as the product template (PT) domain (not shown in **Figure 8**) directs the number of elongation steps and the folding/cyclization of phenolic polyketides (Crawford et al. 2009). A remarkable feature of these PKS is the ability to reduce carbonyl groups at the correct location during polyketide chain

elongation. This is an example of an ‘iterative PKS’ (Tsai, 2018; Herbst et al. 2018). This style of polyketide assembly is commonly observed in *Fungi*.

Modular and iterative PKS comprise a class of PKS known as ‘Type I PKS’. The second class, known as ‘Type II PKS’, could be broadly conceived as functioning in a similar manner to type I PKS. A key difference is that domains of type II PKS are encoded as separate proteins and must associate non-covalently to perform polyketide assembly. As type II PKS are not relevant to this thesis, the reader is instead directed to this informative review for more information (Hertweck et al. 2007). Unlike type I and type II PKS, the third class of PKS, ‘Type III PKS’ do not have distinct catalytic domains. A single active site facilitates the priming, extension, and cyclization. The length of the polyketide chain appears to be controlled by the size of this catalytic pocket (Yu et al. 2012; Shimizu et al. 2017). Although type III PKS were initially thought to occur only in plants, studies within the past two decades have revealed that type III PKS also exist in bacteria and fungi (Funa et al. 1999; Funa et al. 2007) (**Figure 9**). Perhaps the most common subtype of type III PKS is the chalcone synthase (Austin & Noel, 2003). This type III PKS uses 4-coumaroyl CoA as an assembly primer in lieu of acetyl CoA, followed by several extensions with malonyl-CoA (Austin & Noel, 2003).

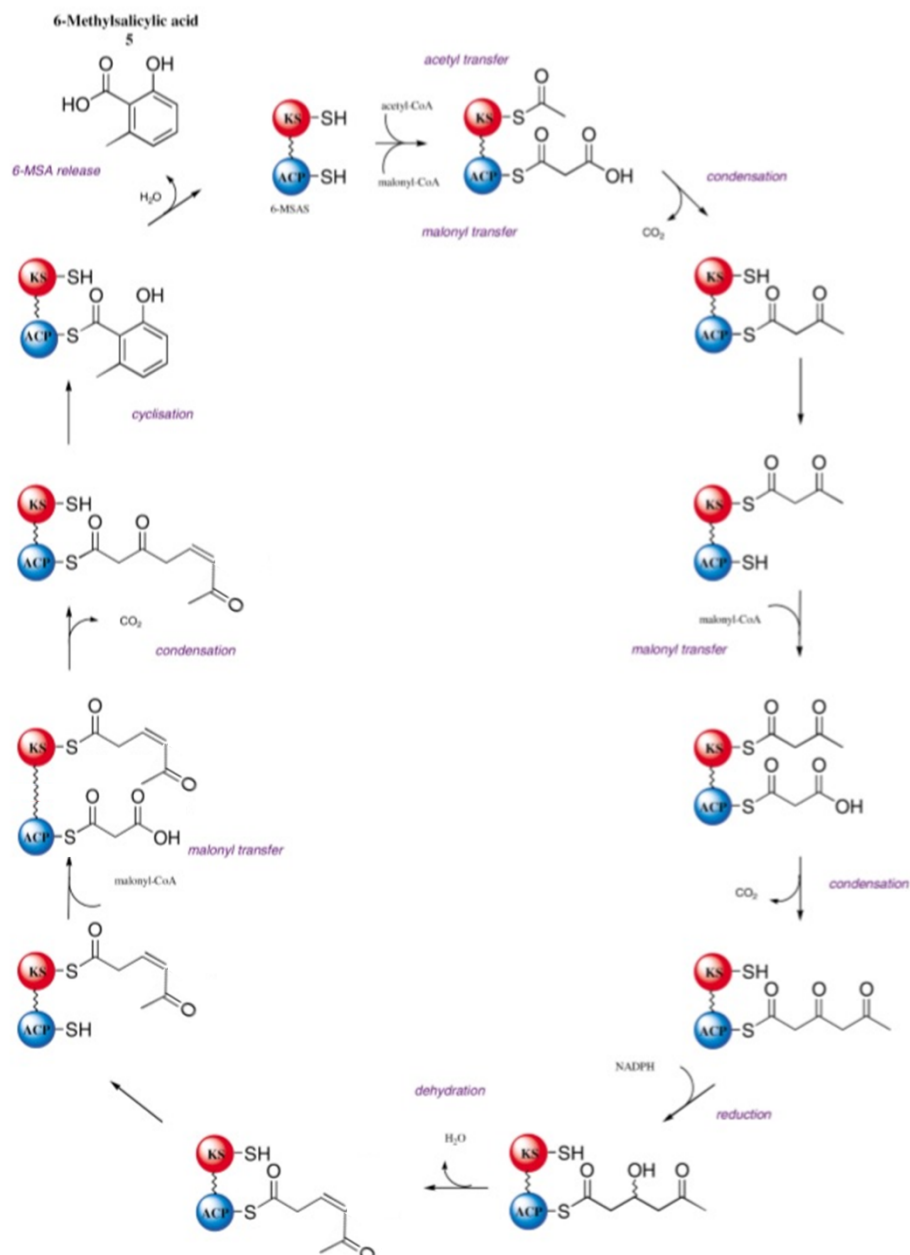


Figure 8: 6-methylsalicylic acid synthase (MSAS), a model iterative PKS. Reproduced with minor modifications from Staunton & Weissman (2001), with permission from the Royal Society of Chemistry. *Abbreviations:* Ketosynthase (KS), Acyl carrier protein (ACP).

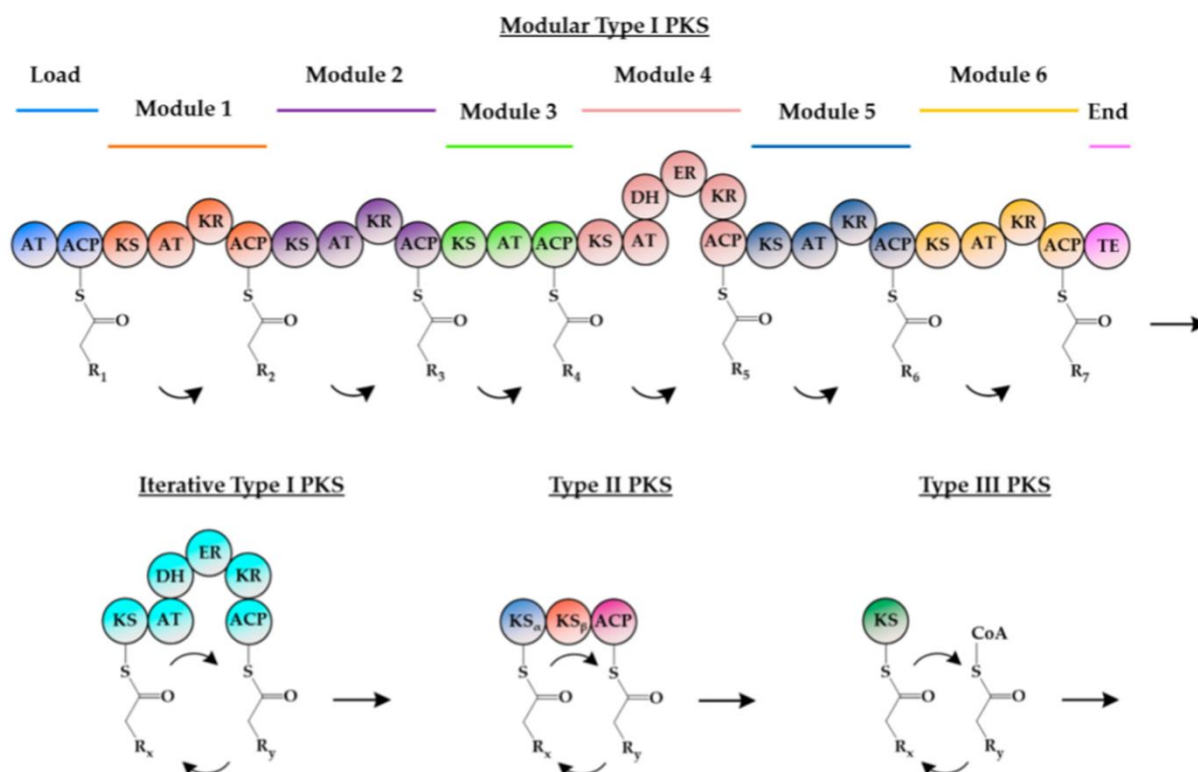


Figure 9: Summary of polyketide assembly by type I, II, and III PKS. Reproduced from Lim et al. (2016), in compliance with the CC-BY Creative Commons Attribution License. *Abbreviations:* Ketosynthase (KS), Acyltransferase (AT), Acyl carrier protein (ACP), Ketoreductase (KR), Dehydratase (DH), Enoylreductase (ER), Thioesterase (TE).

1.2.2. Non-ribosomal peptides

Non-ribosomal peptides are short polymers of amino acids that are produced independently of ribosomes (Strieker et al. 2010). The enzymes that produce these secondary metabolites, the non-ribosomal peptide synthetases (NRPS), may use more than one thousand proteinogenic and non-proteinogenic amino acids to build the peptide (Caboche et al. 2008; Caboche et al. 2009). Such monomer variability confers an enormous degree of potential structural diversity. Although ribosome-independent peptides were discovered relatively recently in the history of the natural products sciences, extensive biochemical and mechanistic analyses have since provided an in-depth understanding of the biosynthesis of this class of secondary metabolite (Kristjan-Bloudoff & Schmeing, 2017; Miller & Gulick, 2016; Tarry et al. 2017; Weissman, 2015). Several dozen non-ribosomal peptides have been marketed as pharmaceuticals. Notable examples include the

antibiotics penicillin, gramicidin, bacitracin and vancomycin (Süssmuth & Mainz, 2017). Other non-ribosomal peptides have found non-medical commercial applications, for example, surfactants for soil restoration (Martinez-Núñez & López y López, 2016).

Three catalytic domains are minimally required (**Figure 10**). The peptidyl carrier protein (PCP; alternatively referred to as a thiolation (T) domain), carries the growing peptide chain. The T domain functions similarly to the ACP domains of PKS such that the T domains are also charged with a phosphopantetheine ‘arm’ and possess a wide range of movement to deliver nascent peptides to other catalytic domains (Kittilä & Cryle, 2016). The adenylation (A) domain is a ‘gatekeeping’ domain that recruits an amino acid to the PCP domain and activates the amino acid by adenylation using one equivalent of ATP (Lee et al. 2010). The condensation domain (C) binds two amino acids that have been recruited at PCP domains (Kittilä & Cryle, 2016). Many optional catalytic domains are available to perform chemical modifications to the growing polypeptide chain (Walsh et al. 2001). One example of a commonly occurring domain is the epimerase (E) domain, which is responsible for inverting the stereochemistry of the incorporated amino acid. Though PKS and NRPS catalyze distinct chemical reactions, the C, A, and T domains of NRPS could be broadly conceived as parallels of the KS, AT, and ACP domains of PKS. A second parallel is that the domains of NRPS are arranged in a series of modules akin to ‘assembly line-style’ PKS (**Figure 10**).

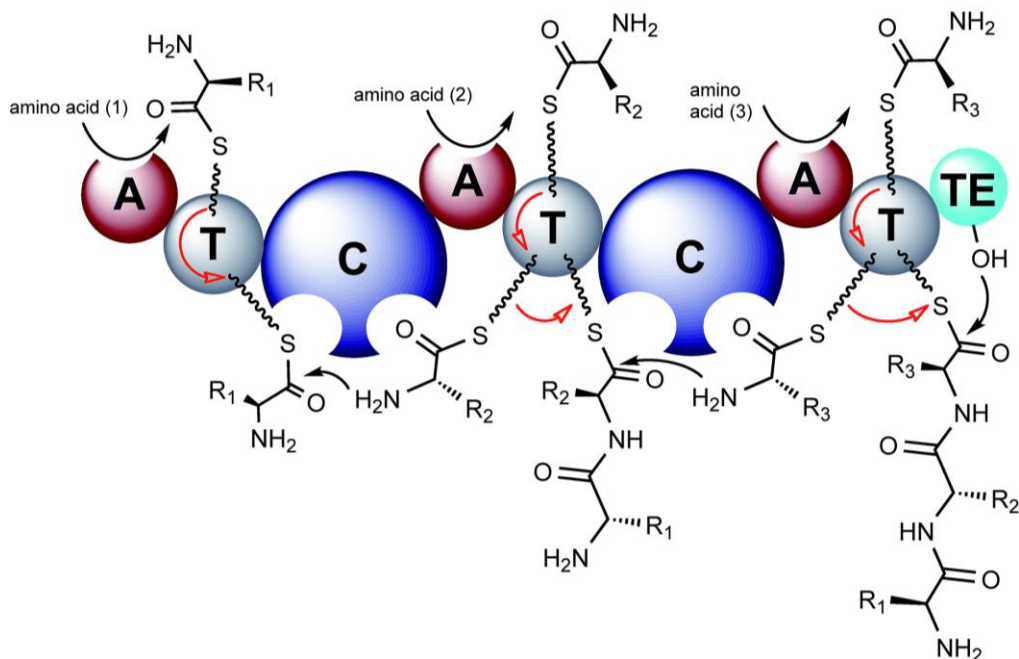


Figure 10: Schematic of peptide elongation by non-ribosomal peptide synthetases (NRPS). Reproduced with minor modifications from Winn et al. (2016), in compliance with the CC-BY Creative Commons Attribution License. *Abbreviations:* Condensation (C), Adenylation (A), Thiolation (T), Thioesterase (TE).

An example of an NRPS is surfactin synthetase (Wu et al. 2017) (**Figure 11**). Seven amino acids are incorporated into the peptide. One interesting feature of surfactin biosynthesis is the addition of a fatty acid chain early during the assembly (indicated by ‘FA’ in **Figure 11**). A catalytic domain located at the terminal end of the assembly line (not illustrated in **Figure 11**) appears to be responsible for adding this fatty chain (Steller et al. 2004). The NRPS may terminate chain extension by thioester hydrolysis of a thioester bond by a thioesterase (TE) domain. Termination may also be achieved by nucleophilic attack and macrocyclization of the peptide. This alternative means of terminating peptide assembly is used by surfactin synthetase (Kopp & Marahiel, 2007).

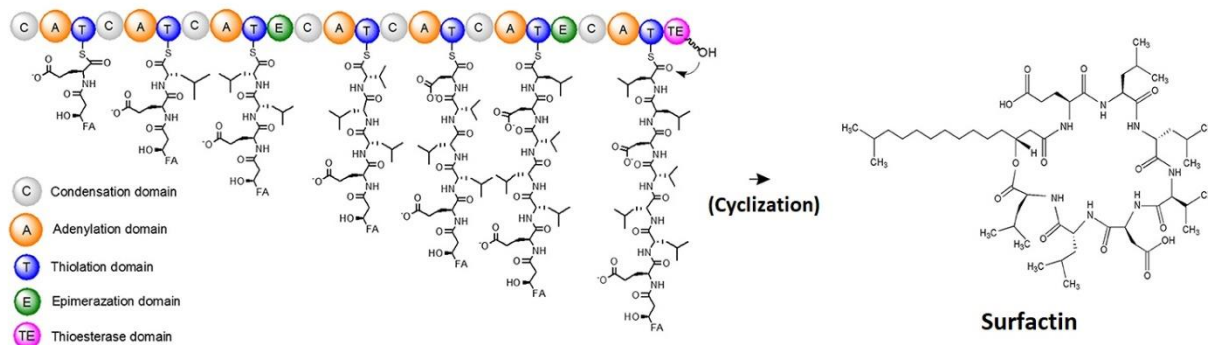


Figure 11: Biosynthesis of surfactin by surfactin synthetase, a model NRPS. Reproduced with minor modifications from Wu et al. (2017), in compliance with the CC-BY Creative Commons Attribution Licence. *Abbreviations:* Condensation (C), Adenylation (A), Thiolation (T), Epimerase (E), Thioesterase (TE), Fatty acid (FA).

1.2.3. Hybrid PKS-NRPS

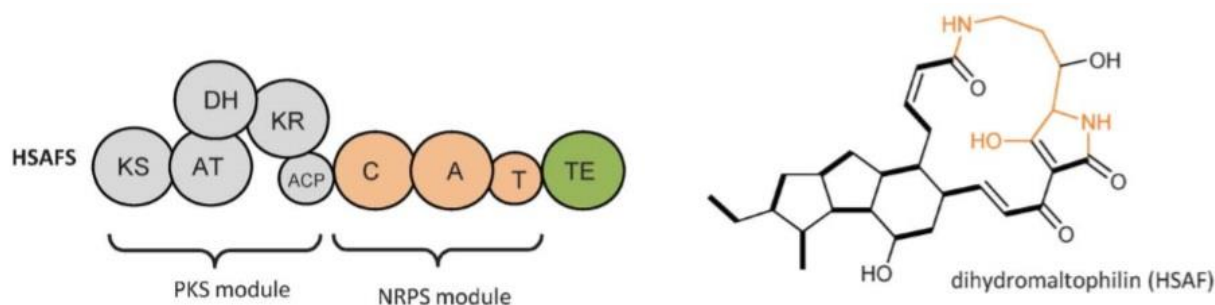


Figure 12: Dihydromaltophilin, an example of a metabolite produced by a hybrid PKS-NRPS. Portion of the molecule highlighted in orange is the amino acid incorporated by the NPRS module. Reproduced from Fisch (2013), with permission from the Royal Society of Chemistry. *Abbreviations:* Ketoynthase (KS), Acyltransferase (AT), Dehydratase (DH), Ketoreductase (KR), Acyl carrier protein (ACP), Condensation (C), Adenylation (A), Thiolation (T), Thioesterase (TE).

Given the similarities between PKS and NRPS, it is perhaps unsurprising that Nature has found a way to combine the biosynthetic machinery of both enzymes to form what are referred to as PKS-NRPS (Fisch, 2013; Boettger & Hertweck, 2013). In these hybrid enzymes, the PKS-like portion builds a polyketide intermediate that is then passed to the NRPS-like portion for peptide addition. Triple chimeras of polyketide/non-ribosomal peptide/fatty acid secondary metabolites have also been observed (Masschelien et al. 2013). Dihydromaltophilin is an example of a metabolite produced by a hybrid PKS-NRPS (**Figure 12**).

1.2.4. Terpenes

Terpenes are hydrocarbon polymers of isopentenyl pyrophosphate (IPP) and dimethylallyl pyrophosphate (DMAPP) (**Figure 13**). The enzyme isopentenyl pyrophosphate isomerase catalyses the reversible isomerization between these two molecules to ensure a steady-state balance of both monomers (Berthelot et al. 2012). Two pathways, the mevalonic acid pathway, and the mevalonic acid-independent pathway, are responsible for providing these monomers. Although only one pathway is present in most organisms, some species of bacteria and plants have both. Each organism including humans produces dozens of terpenes (Pichersky & Raguso, 2018; Gershenzon & Dudareva, 2007; Cho et al. 2017; Vattekkatte et al. 2018; Yamada et al. 2015). Examples of terpenes include the anticancer drug taxol (Croteau et al. 2006), the anti-malarial drug artemisinin (Xie et al. 2016), and the flavouring agent limonene (Jongedijk et al. 2016).

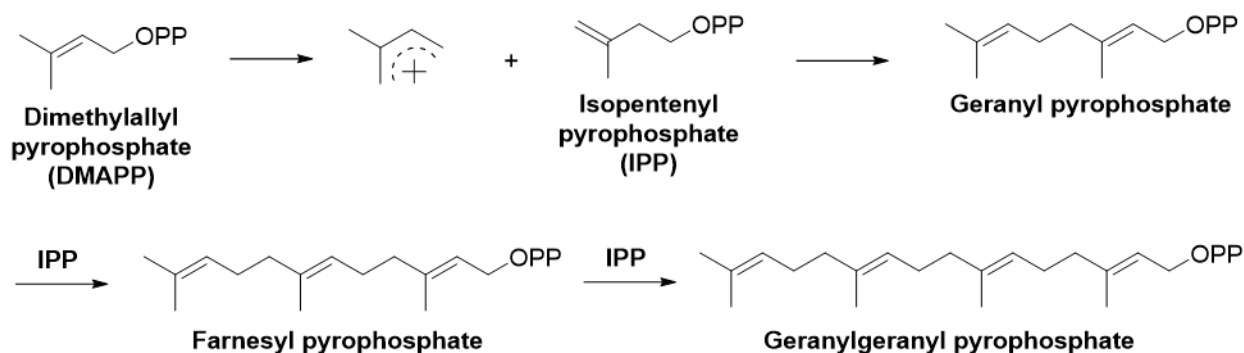


Figure 13: Carbocation-mediated polymerization of DMAPP and IPP by terpene synthases.

Terpenes are produced by terpene synthases. Terpene polymerization by terpene synthases is shown (**Figure 13**). Pyrophosphate acts as a leaving group, turning the dimethylallyl moiety into an electrophile for nucleophilic attack by the isopentenyl pyrophosphate. Each repetition of this reaction extends the chain by five carbon atoms (Oldfield & Lin, 2012). The discovery that terpenes could be formed through this process of polymerization was foundational to the natural products sciences (Ružička, 1953). A subset of terpenes are the terpenoids – terpenes that possess

at least one oxygen residue. Elegant folding of the linear chain creates polycyclic molecules. Three examples of terpenes or terpenoids are shown (**Figure 14**).

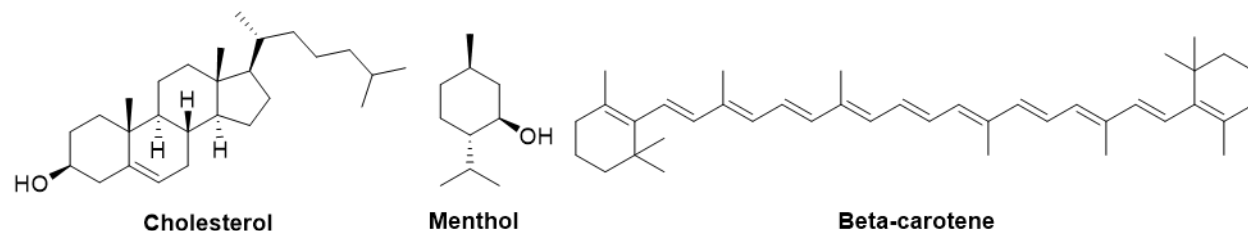


Figure 14: Three examples of terpen(oids).

1.3. Biosynthetic pathways, gene clusters, and biotechnological applications

The structural diversity of Nature's secondary metabolites does not end with the formation of a polyketide, non-ribosomal peptide, or terpene. Structural elements may be added or modified by other enzymes. Examples of modifications include the addition of hydroxyl or epoxy groups, oxidations and reductions of bonds, halogenations, *O*-methylations, carbon-carbon coupling or decoupling, et cetera. In some cases, the final product is so thoroughly transformed as to be unrecognizable from the starting molecule. An example of this is the fungal toxin patulin (Puel et al. 2010). The polyketide is produced in ten chemical steps beginning with 6-methylsalicylic acid (**Figure 15**). These 'tailoring' enzymes, together with the PKS, NRPS, or TS, are encoded as genes within the genome of organisms. With exceptions to some forms of marine bacteria, enzymes that contribute to the formation of a specific natural product are usually encoded together within a common region of DNA. This group of genes is known as a 'biosynthetic gene cluster'. A complete gene cluster therefore encodes the biosynthetic machinery necessary for the biosynthesis of a single secondary metabolite. The patulin gene cluster and biosynthetic pathway is provided as an example (**Figure 15**).

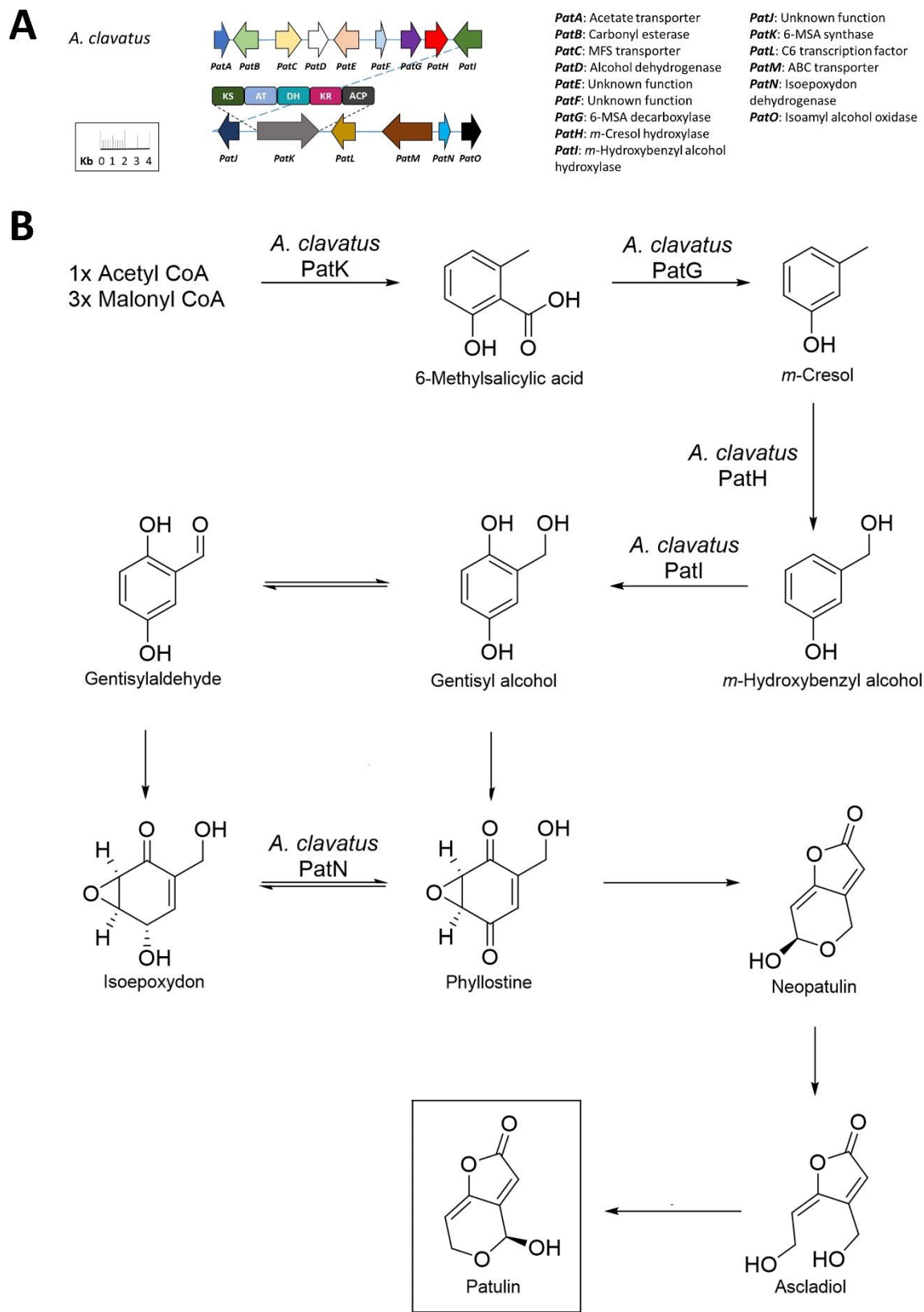


Figure 15: The patulin gene cluster and biosynthetic pathway – an example of a complex metabolic pathway.

The amount of genetic information that is available on biosynthetic gene clusters is growing at an astonishing rate. This has been made possible by advancements in genome sequencing technology that allow researchers to sequence larger genomes faster and at lower costs (Goodwin et al. 2016). This progress is also aided by the development and improvement of programs designed to detect genes associated with secondary metabolism, for example, AntiSMASH (Blin et al. 2017), SMURF (Khaldi et al. 2010), and PRISM (Skinnider et al. 2017). The ‘assembly line’ arrangement of modular PKS and NRPS, combined with the structural elements that allow adenylation and acyltransferase domains to selectively recruit monomers, also enables programs to generate predictions on the structural features of the metabolites that the encoded PKS or NRPS produces (Starcevic et al. 2008).

Genome sequencing projects in bacteria, plants, and fungi have revealed that the number of biosynthetic gene clusters encoded within organisms vastly exceeds the number of secondary metabolites known to be produced by these organisms (Schorn et al. 2016; Huang et al. 2017; Nielsen et al. 2017). Most gene clusters cannot be associated with any known metabolite. In these cases, these gene clusters are ‘cryptic’ and are speculated to encode undiscovered secondary metabolites. Investigating these ‘cryptic’ clusters using biotechnological and bioinformatic tools is a process known as ‘genome mining’. This process has led to the identification of new natural products (Cacho et al. 2015; Ziemert et al. 2016).

There are several means of investigating the biosynthetic function of gene clusters. In the case of ‘cryptic’ gene clusters, one method is to activate their expression and observe for any metabolic changes. If silence is due to a regulatory ‘switch’ that is only ‘turned on’ given a specific stimulus (e.g., UV radiation, co-culturing with other organisms, et cetera), applying that stimulus could result in gene expression and metabolite biosynthesis (Pettit, 2009; Bode et al. 2002; Tanaka

et al. 2010). Alternatively, the regulatory ‘switches’ can be engineered within the organism to either induce or increase the rate of gene expression (Aigle & Corre, 2012; Suroto et al. 2017; Saha et al. 2017). This strategy is becoming increasingly widespread due to the emergence of CRISPR Cas9 gene-editing technology (Zhang et al. 2017). In cases where the metabolite is known but the associated gene cluster is speculative, a second method of investigation is to knock-out candidate genes and observe whether the studied metabolite is no longer being produced. This is a particularly useful technique for dissecting the function of encoded enzymes and how they individually contribute towards the formation of a complete biosynthetic pathway (Alhawatema et al. 2017; Fuller et al. 2015). A third method of investigation is to insert biosynthetic genes into a host and observe for *de novo* metabolite biosynthesis within the host, a process known as heterologous expression (Zhang et al. 2016; Alberti et al. 2017). A variety of expression hosts have been developed specifically for the functional heterologous expression of biosynthetic gene clusters from prokaryotes and eukaryotes alike (Anyagou & Mortensen, 2015; Yaegashi et al. 2014; Billingsley et al. 2016; Li & Neubauer, 2014; Loeschcke & Thies, 2015; Gomez-Escribano & Bibb, 2014; Baltz, 2010; Ikram et al. 2015).

The motivation to find and functionally characterize biosynthetic gene clusters rests in the biotechnological applications arising from its cloning and heterologous expression. One common barrier to the commercialization of natural products concerns their mass production. The most obvious way to produce a biogenic molecule is to cultivate the producing organism at commercial scale, then extract the molecule from it. However, this approach is often untenable because cultivation is difficult, expensive, or because the native organism produces the molecule in miniscule quantities. Although one may initially suggest that this problem could be overcome through chemical synthesis, such a task is greatly complicated by the numerous stereocenters and

functional groups that are often present on complex secondary metabolites. Case in point, the major barrier to the commercialization of taxol, an FDA-approved anticancer drug produced by Pacific yew trees, was that it required cutting down three trees to produce enough drug to treat one patient. This was an ecologically devastating process that was met with adverse public opinion. The development of a protocol for the complete chemical synthesis of taxol, justifiably heralded as a breakthrough, nonetheless required 17 (!) chemical steps (Nicolaou et al. 1994).

These problems could be overcome by harnessing the molecular machinery responsible for producing these complex molecules. The enzymes that are responsible for producing secondary metabolites have been ‘programmed’ by evolution to perform precise chemical modifications. If the genes encoding these biosynthetic enzymes were expressed in a heterologous host, and the enzymes purified, a chemist could use these enzymes to resolve difficult steps via a semi-synthetic approach. Alternatively, if the complete biosynthetic gene cluster were to be expressed in a fast-growing host such as yeast or bacteria, commercially-relevant amounts of the molecule could be produced in large bioreactors. For example, the mass production of the potent anti-malarial drug artemisinin by heterologous expression of its biosynthetic gene cluster in yeast is arguably one of the greatest success stories within the field of biotechnology (Paddon et al. 2013).

The ‘assembly line’ nature of modular PKS and NRPS also presents the exciting possibility that components of these molecular factories could be swapped to produce ‘unnatural’ derivatives of natural products. This emerging sub-field of natural products chemistry, known as ‘combinatorial biosynthesis’, aims to develop enzymes that can produce libraries of derivatives of naturally-occurring molecules with improved drug-like properties (Sun et al, 2015; Bayly & Yadav, 2017; Winn et al. 2016). Although pharmaceutical companies have in recent decades

shied-away from seeking naturally-sourced drug candidates, biotechnological developments such as these have reinvigorated interest in the prospecting of Nature's molecules (Shen, 2015).

1.4. What did we know about lichen secondary metabolite biosynthesis in 2012?

Not much. When I began my research in September 2012, the isolation and structural elucidation of hundreds of secondary metabolites from lichens had been reported. However, this body of work was not complimented by an in-depth understanding of the genetically-encoded biosynthetic machinery that is responsible for producing these molecules. Degenerate primer 'fishing' experiments combined with phylogenetic analyses revealed that genes encoding type III PKS and 6-methylsalicylic acid-producing PKS are widely distributed throughout lichens (Muggia & Grube, 2010; Schmitt et al. 2008). However, concrete information, such as complete DNA sequences of biosynthetic genes or their associated clustered genes, remained conspicuously absent from the literature. At the outset of my research, only 13 biosynthetic genes were known and completely sequenced. Of these 13 examples, information on the accessory tailoring genes was available for only four of these PKS (**Table 1**)².

Table 1: All literature reports of sequenced biosynthetic genes from lichenizing fungi prior to September 2012. Adapted from Bertrand & Sorensen (2018), in accordance with authors' retained rights.

Entry	Name	Accession	Species	Architecture	Clustered genes	Detection Method	Predicted product	Prediction Method	Ref(s)
1	<i>Cgrpks1</i>	HQ823618	<i>Cladonia grayi</i>	SAT-KS-AT-PT-ACP-ACP-MT-CYC	?	cDNA - genome assembly	?	?	Armaleo et al. 2011
2	<i>Cgrpks2</i>	GU930714	<i>Cladonia grayi</i>	SAT-KS-AT-PT-ACP-MT-CYC	?	cDNA - genome assembly	?	?	Armaleo et al. 2011
3	<i>Cgrpks13</i>	HQ823619	<i>Cladonia grayi</i>	SAT-KS-AT-PT-ACP-ACP-TE	?	cDNA - genome assembly	?	?	Armaleo et al. 2011
4	<i>Cgrpks14</i>	HQ823620	<i>Cladonia grayi</i>	SAT-KS-AT-PT-ACP-ACP-TE	?	cDNA - genome assembly	?	?	Armaleo et al. 2011
5	<i>Cgrpks15</i>	HQ823621	<i>Cladonia grayi</i>	SAT-KS-AT-PT-ACP-TE	?	cDNA - genome assembly	?	?	Armaleo et al. 2011
6	<i>Cgrpks16</i>	GU930713	<i>Cladonia grayi</i>	SAT-KS-AT-PT-ACP-ACP-TE	p450, MT	cDNA - genome assembly	Grayanic acid	Transcription profiling	Armaleo et al. 2011
7	<i>Dnpks1</i>	EU872212	<i>Dirinaria applanata</i>	SAT-KS-AT-PT-ACP-ACP-TE	?	gDNA - library screening	Melanin/anthraquinone-related	Phylogenetics	Valarmathi et al. 2009
8	<i>Ulpks1</i>	JN408682	<i>Usnea longissima</i>	SAT-KS-AT-PT-ACP-ACP-TE	Red, Unk, Unk	gDNA - library screening	Melanin/orsellinic acid-related	Phylogenetics	Wang et al. 2012
9	<i>Cmpks1</i>	HQ413098	<i>Cladonia metacoraillifera</i>	KS-AT-DH-KR-ACP	LAAO, Mono, Unk	gDNA - library screening	Methylsalicylic acid-related	Phylogenetics	Kim et al. 2012
10	<i>Ulpks3</i>	HQ824546	<i>Usnea longissima</i>	KS-AT-DH-ER-KR-ACP	Hyd, Unk	gDNA - library screening	Resorcylic acid lactone-related	Phylogenetics	Wang et al. 2011
11	<i>Scpks1</i>	EF554834	<i>Solorina crocea</i>	KS-AT-DH-MT-ER-KR-ACP	?	gDNA - library screening	?	?	Gagunashvili et al. 2009
12	<i>Xsepks1</i>	AB558604	<i>Xanthoparmelia semiviridis</i>	SAT-KS-AT-PT-ACP-MT	?	gDNA - library screening	Methylorsellinic acid-related	Phylogenetics	Chooi et al. 2008
13	<i>Xepks1</i>	DQ660910	<i>Xanthoria elegans</i>	SAT-KS-AT-PT-ACP-ACP-TE	?	cDNA - RACE	Anthraquinone-related	Phylogenetics	Brunauer et al. 2009

Abbreviations: Starter unit acyl carrier protein transacylase (SAT), Ketosynthase (KS), Acyltransferase (AT), Product template domain (PT), Acyl carrier protein (ACP), C-methyltransferase (MT), Thioesterase (TE), Claisen cyclase (CYC), No information available (?), Cytochrome p450 (p450), Reductase (Red), Unknown function (Unk), L-amino acid oxidase (LAAO), Monooxygenase (Mono), Hydratase (Hyd).

² For information updated to 2018, see Bertrand & Sorensen (2018). Information on sequenced NPRS and terpene synthases from lichens can also be found in these articles (Bertrand et al. 2018a; Wang et al. 2014).

This absence of data on lichen biosynthetic genes was in part due to the technological limitations at the time. For example, genome sequencing by next-generation sequencing platforms had yet to have become both widespread and affordable to small research groups such as ours. Consequently, most research groups must have instead relied upon low-throughput techniques to find new genes such as degenerate primer ‘fishing’ combined with conventional Sanger-style DNA sequencing. Experiments to discover new natural products in lichens by stimulating dormant metabolic pathways (e.g., UV radiation, carbon source, etc) have by-and-large been conducted ‘blind’ such that no information was available on the ‘regulatory’ switches that were available to be turned ‘on’ or ‘off’ by these stimulants (Brunauer et al. 2007; Deduke et al. 2012; Elshobary et al. 2016; Stocker-Wörgötter et al. 2004; Timsina et al. 2013; Culberson & Armaleo, 1992). The declining cost and expanding availability of ‘next-generation’ platforms have since created an impressive volume of new entries: Between 2012 and 2018, the number of lichen biosynthetic genes reported in the literature has expanded by ten-fold (Bertrand & Sorensen, 2018).

There is a second explanation for the relative lack of data on lichen biosynthetic genes: Lichens grow extremely slowly. A growth rate of 1 cm/year of diameter expansion is considered typical (Grube et al. 2014). Lichenologists have even taken advantage of this unusual quality of lichens by developing a biometric dating method known as ‘lichenometry’ that can be used to estimate the age of ancient ruins or monitor climate changes, with a time-scale on the order of several thousand years (Benedict, 2009; Armstrong, 2004). Lichens appear to be able to grow virtually anywhere: Readers who have hiked in the Canadian Shield will have noticed lichens growing on exposed bedrock, and lichens are among a select few forms of flora that are observable within the Arctic and Antarctic. Such desolation has an advantage: Living where nothing else wants to live practically eliminates nutrient competition. Perhaps a slow growth rate is the means

by which lichens can survive in such harsh environments. Their slow growth rate persists even if cultivated in ideal conditions in laboratory, suggesting that slow growth is an evolutionary-engrained phenotype and is not merely a response to the current milieu. This feature of lichens greatly complicates laboratory tasks that would otherwise be considered mundane for other forms of life, for example, DNA isolation or time-course experiments. Elucidation of the function of gene clusters via gene knockout is also considered impractical because of their extended lifespans. Heterologous expression is therefore considered a more practical option, and *Ascomycota* hosts such as *Aspergillus oryzae* and *Aspergillus nidulans* are available to conduct such expression studies (Anyagou & Mortensen, 2015). However, experiments to functionally express lichen PKS in hosts have only been met with frustration. At the outset of my doctoral program, three unsuccessful attempts had been reported in the literature. These experiments include the insertion of a PKS gene (*xsepks1*) from *Xanthoparmelia semiviridis* into *Aspergillus nidulans* (Chooi et al. 2008), the insertion of a PKS gene (*scpk1*) from *Solorina crocea* into *Aspergillus oryzae* (Gagunashvili et al. 2009), and the insertion of a PKS gene from *Cladonia grayi* (*cgrpk2*) into *A. nidulans* (Armaleo et al. 2011). Until functional heterologous expression of lichen gene clusters becomes possible, our understanding of secondary metabolite biosynthesis in lichens will lag far behind comparable studies among plants, bacteria, and non-lichenizing fungi. For this reason, I considered the testing of heterologous expression protocols a priority of my doctoral program.

In lieu of heterologous expression and gene knockout, lichenologists have reverted to non-definitive approaches to link genes to metabolites. One such approach example is phylogenetics – the study of the evolutionary relationships between organisms by comparison of DNA sequences. A core premise of phylogenetics is that genes with common ancestry that evolved within a recent timeframe are more likely to encode proteins with similar functions as compared to related genes

within a distant timeframe. In secondary metabolite biosynthetic studies, observation of common and recent ancestry between a subject and reference gene is used as supporting evidence that the encoded proteins possess similar roles in metabolite biosynthesis (Ziemert & Jensen, 2012). This process provided general structural predictions about the metabolites that are produced by six of the encoded PKS (**Table 1**). The most significant advancement in lichen gene-metabolite linkages *circa* 2012 was the identification of the grayanic acid gene cluster in *C. grayi* via a transcription profiling approach (Armaleo et al. 2011). Incubation conditions were first established to reproducibly induce grayanic acid production in *C. grayi* (Culberson & Armaleo, 1992). After several candidate PKS genes were identified, quantitative PCR was used to reveal a correlation between increased transcription of one of these candidate PKS genes (*Cgrpks16*) and grayanic acid bioaccumulation over a period of 36 days. A phylogenetic analysis demonstrated genetic similarity between *Cgrpks16* and a PKS from a non-lichenizing fungus known to produce orsellinic acid-based metabolites. As grayanic acid biosynthesis requires orsellinic acid as the first intermediate, this phylogenetic analysis supports the assignment of CgrPKS16 as the PKS responsible for grayanic acid biosynthesis. Genes encoding a cytochrome p450 and an *O*-methyltransferase were found to be located near *Cgrpks16*, enabling the authors to propose a complete grayanic acid biosynthetic pathway (Armaleo et al. 2011) (**Table 1**).

1.5. Summary of research objectives

To explore the lichen secondary metabolome, we need a way to functionally express lichen biosynthetic genes within a heterologous host. Due to the limited genetic information available *circa* 2012, we decided that genome sequencing and annotation of lichen biosynthetic gene clusters was an appropriate first step. If heterologous expression could be achieved, we could then use the developed method to functionally characterize our annotated lichen genes. It would also be useful

to generate putative assignments of function of the annotated gene clusters to assist us in these characterization experiments. This would in turn provide the methodological groundwork for the mass production and commercialization of interesting secondary metabolites from lichens. This can be summarized as three objectives:

Objective 1: Sequence the genome of a lichenizing fungus, annotate its genetic secondary metabolome, and provide putative assignments of biosynthetic function.

Objective 2: Develop a reliable protocol for the heterologous expression of lichen biosynthetic genes and/or gene clusters within a host.

Objective 3: Use this protocol to characterize lichen biosynthetic gene clusters.

To pursue these three objectives, I investigated the biosynthesis of usnic acid in *Cladonia uncialis*. In Chapter 3, I will discuss why we chose to study this particular secondary metabolite and this particular species of lichen.

Chapter 2

Materials and methods

2.1. Materials and methods related to Chapter 1

Not applicable.

2.2. Materials and methods related to Chapter 2

For a list of primers used in all chapters, see **Table S1** in Appendix. For a list of plasmids constructed in all chapters, see **Table S2** in Appendix.

2.3. Materials and methods related to Chapter 3

*2.3.1. Collection and taxonomic identification of *C. uncialis**

As described by Abdel-Hameed (2015). The internal transcribed spacer (ITS) sequence that was acquired by Abdel-Hameed (2015) matched entries for *C. uncialis* with >95 % sequence identity (**Table S3** in Appendix). Within **Table S3** and others throughout this thesis, BLAST alignment statistics will be reported in terms of “Identity % / Coverage %”. The “identity” is the percentage of nucleotides or amino acids within a queried sequence that are identical to that of the reference gene. The “coverage” is the percentage of overlap between the analyzed portions of the queried gene and the reference gene. These statistics are reported as XX / YY throughout this thesis.

*2.3.2. Sub-culturing of *C. uncialis* from the algal partner*

As described by Abdel-Hameed (2015). The ITS sequence that was acquired by Abdel-Hameed (2015) matched entries for *C. uncialis* with >95 % sequence identity (**Table S4** in Appendix).

2.3.3. Genome sequencing and assembly

The genome of the fungal partner was *de novo* sequenced by MICB DNA sequencing services (Manitoba Institute of Cell Biology, CancerCare Manitoba, University of Manitoba), using an Illumina MiSeq sequencer with a MiSeq Micro V2 sequencing kit. The genomic DNA sample from the fungal culture was used to generate a paired-end DNA library for 150 bp paired-end sequencing reads. The genome was estimated to be 30 MB in length, an estimate based on studies of genome length of *Cladonia grayi* (Armaleo & May, 2009). The average length of the raw reads was 147 nucleotides with a standard deviation of 16 nucleotides. A total of 33664396 raw reads were generated, to provide an estimated genomic coverage of 165x. A Phred quality score of 30 or greater was achieved for at least 85 % of the nucleotides generated, signifying that these nucleotides were correctly identified as A, T, G, or C with confidence of 99.9 % or greater (Ewing et al. 1998; Ewing & Green, 1998). Raw sequence reads were deposited in GenBank and are available under accession number SRR4418292. The raw data obtained from Illumina MiSeq sequencing were assembled into contigs using four DNA assembly programs: DNASTAR, Geneious, SPAdes, and Velvet (Kearse et al. 2012; Bankevich et al. 2012; Zerbino & Birney, 2008). The results generated by SPAdes was chosen for subsequent analysis. This assembly generated a total of 2109 contigs (≥ 1 KB) with a combined total of 32.9 million nucleotides. This result was consistent with our prior estimation of 30 million nucleotides based on the estimated length of the *C. grayi* genome (Armaleo & May, 2009). The GC/AT ratio of the assembled DNA was 46.38 %. The contig N50 (defined as the value at which half of the genomic DNA is contained within contigs of X length or greater) was 34.7 KB. The longest contig was 143.1 KB. The 2109 contigs were deposited in GenBank under accession number NAPT000000000. 33664396

2.3.4. Annotation of secondary metabolite gene clusters

The rapid *in silico* identification of putative polyketide synthase (PKS) genes was performed using Antibiotics and Secondary Metabolites Analysis Shell (AntiSMASH Version 2.0) (Blin et al. 2013), freely available for academic use at <http://www.secondarymetabolites.org/>

2.3.5. Phylogenetics

A total of 100 amino acid sequences of KS and CLC domains that were similar to the those of the usnic acid-associated PKS were compiled from GenBank. Amino acid sequences were aligned using Clustalx V. 2.0 (Larkin et al. 1985). Maximum likelihood trees were produced using the option tree bisection and reconnection branch swapping. Heuristic searches were conducted using 1000 random addition replicates with a limit of 10 trees per search and bootstrap searches of 1000 re-samplings (Felsenstein 1985). Bootstrap tests with a score of 60 % or greater are reported above each node. The sequence of the *Trichoderma atroviride* (EHK44445) PKS gene was selected as an out-group for the KS domain tree while *Trichoderma virens* (EHK19347) hypothetical protein was chosen for the CLC domain tree.

2.4. Materials and methods related to Chapter 4

2.4.1. Detection and sequencing of a transcriptionally active PKS

As described by Abdel-Hameed (2015).

2.4.2. Annotation of gene cluster

Annotation of gene clusters was performed using the Antibiotics and Secondary Metabolites Analysis Shell (AntiSMASH version 3.0) (Weber et al. 2015), freely available for academic use at <http://www.secondarymetabolites.org/>

2.4.3. Phylogenetics

Genes of interest were analyzed by BLAST (Altschul et al. 1990). The 35 highest similar amino acid sequences found in GenBank were used to construct a phylogenetic tree using the neighbor-joining method (Saitou & Nei, 1987) with MEGA6 software (Tamura et al. 2013). Confidence probability was estimated using 1000 replicates of the interior branch test (Rzhetsky & Nei, 1992). Interior branch tests with a score of 60 % or greater are reported above each node. Evolutionary distances were computed using the Poisson correction method (Zuckerkandl & Pauling, 1965) and are reported in units of amino acid substitutions per site.

2.5. Materials and methods related to Chapter 5

2.5.1. Annotation of lichen gene clusters

Annotation of the gene clusters was performed using Antibiotics and Secondary Metabolites Analysis Shell (AntiSMASH version 3.0) (Weber et al. 2015), freely available for academic use at <http://www.secondarymetabolites.org/>.

2.5.2. Proofreading of annotations

The domain architectures of PKS and NRPS identified by AntiSMASH (V. 4.0) were verified with BLAST (Altschul et al. 1990). The broad functional roles of all other genes (e.g., *O*-methyltransferase) were predicted with BLAST and are based on consensus assignment of genetically similar sequences deposited in GenBank. In some cases, it was necessary to suggest more than one putative function due to lack of consensus. BLAST statistics including the most closely similar gene for each annotated gene is provided (**Tables S5-S52** in Appendix). Any gene with coverage values below 60 % and/or identity values below 25 % was listed as ‘no significant similarity’. The open reading frame of each gene (including introns) was determined via AUGUSTUS (Stanke et al. 2004; Stanke & Morgenstern, 2005), freely available for academic use

at <http://augustus.gobics.de/>. These genes are deposited in GenBank and the accession numbers are provided (**Tables S5-S52** in Appendix). Contigs containing putative biosynthetic gene clusters were also deposited in GenBank and the accession numbers are provided (**Table S53** in Appendix)

2.6. Materials and methods related to Chapter 6

2.6.1. Identification of genetically-similar gene clusters

Cladonia uncialis gene clusters were uploaded to AntiSMASH (v. 4.0) (Blin et al. 2017), freely available for academic use at <http://www.secondarymetabolites.org/>. The KnownClusterBLAST module embedded within AntiSMASH (V. 4.0) provided cluster similarity statistics to characterized gene clusters. A pBLAST analysis was performed on each genetically similar gene to determine percent similarity scores (Altschul et al. 1990). Lichen genes not identified as genetically similar to genes of the reference cluster were manually analyzed using comparative pBLAST to confirm absence of similarity. The reporting of ‘no significant similarity found’ was interpreted as a negative result.

2.6.2. Phylogenetics

Genes were analyzed via pBLAST (Altschul et al. 1990) and the 38 most genetically similar entries in GenBank were compiled for phylogenetic analysis. Out-group genes were selected from organisms known to be more distantly related than organisms of the in-group genes that also share the general function of in-group genes (e.g., *O*-methyltransferase). The hypothesized fungal homologue found via the KnownClusterBLAST module of AntiSMASH (Blin et al. 2017), and one out-group gene, was included for a total of 40 entries per phylogenetic tree. Multiple sequence alignments were performed with MEGA (v. 7.0) (Kumar et al. 2016), freely available for download at <http://www.megasoftware.net/>. Phylogenetic trees were constructed in MEGA (v. 7.0) (Kumar et al. 2016) using the neighbour-joining method (Saitou &

Nei, 1987). Branch confidence was estimated using 1000 replicates of the interior branch test (Dopazo, 1994; Rzhetsky & Nei, 1992). Nodes with confidence of 70 percent or greater are shown on each tree. The Poisson correction method was used to estimate evolutionary distance (Zuckerkindl & Pauling, 1965), and is reported on each tree in units of amino acid substitutions per site.

2.7. Materials and methods related to Chapter 7

2.7.1. In silico experiments

Amino acid sequences of three representative KS domains from each of the 14 sub-families of the KS3 family were selected at random for phylogenetic analysis (Chen et al. 2011). The KS of *Arabidopsis thaliana* (NP_179113), a member of the KS2 family, was chosen as an outgroup. Multiple sequence alignment was performed using Clustal W (Thompson et al. 1994). Phylogenetic analysis was performed using MEGA7 (Kumar et al. 2016). The neighbour-joining method was used to infer evolutionary history (Saitou & Nei, 1987). The resultant tree was evaluated using 1000 replicates of the bootstrap test (Felsenstein, 1985). The estimated evolutionary distance was determined using the Poisson correction method (Zuckerkindl & Paulin, 1965). Protein modelling was performed with Swiss Model (Waterhouse et al. 2018)), freely available at <https://swissmodel.expasy.org/>. Models were processed with DeepView/Swiss-PdbViewer (V.4.1) (Guex & Peitsch, 1997), freely available at <https://spdbv.vital-it.ch/>.

2.7.2. Polymerase chain reaction

All PCR reactions were performed using Phusion polymerase and reagents (ThermoScientific). Per 20 µL: 4 µL Phusion HF Buffer (containing 7.5 mM MgCl₂), 0.8 µL of 10 mM dNTP mix, 0.6 µL of 50 mM EDTA, 0.8 µL of 10 µM forward primer, 0.8 µL of 10 µM reverse primer, template DNA (variable concentration), 0.2 µL of 2 U/µL Phusion polymerase,

11.8 μ L of water. Primers are provided in **Table S1** in Appendix. PCR was performed with a SimpliAmp thermocycler (*Applied Biosystems*): Initial denaturation was at 94°C for 60 seconds, 25 rounds of denaturation at 94°C for 10 seconds, 25 rounds of annealing at 3-5°C below primer T_m for 15 seconds, 25 rounds of extension at 72°C for 30 seconds per KB, and final extension at 72°C for 30 seconds per KB. Amplicons were electrophoresed on a 1 % agarose gel containing 1.25 μ L of CYBR Safe DNA Gel Stain (*Invitrogen*), and electrophoresed for 45-60 minutes using a gel box (*BioRad*) and Powerpack (VWR) programmed at 230 mA and 200 V. Amplicons were visualized using blue light, excised from the gel, and the DNA purified using the GeneJet Gel Purification Kit (*ThermoScientific*) in accordance with the manufacturer's protocols. The amplified DNA was prepared for plasmid construction by concentrating the template using a SpeedVac. DNA concentration and quality were evaluated using a model 2000c NanoDrop (*ThermoScientific*).

2.7.3. Preparation of plasmids

Transformation plasmids (pTAex3, pUSA, and pAdeA) were linearized by restriction digestion with FastDigest SmaI (*ThermoScientific*) in accordance with the manufacturer's instructions. Linearized plasmids were electrophoresed on agarose gel, purified, concentrated, and evaluated for concentration and quality (see 'PCR', above). Genes amplified from *C. uncialis* were bound to linearized plasmid using InFusion (*Takara*) in accordance with the manufacturer's instructions. Plasmid constructs were screened by transformation into Stellar *E. coli* (*Takara*) using the manufacturer's recommended heat-shock method, then plated on LB media (Per 1 L of water: 10 g tryptone, 5 g yeast extract, 5 g NaCl, 0.1 g ampicillin; pH unadjusted). Transformants were inoculated by sterile toothpick into liquid LB media containing ampicillin and grown overnight. Plasmids were extracted from *E. coli* suspension using the GenJet Plasmid Miniprep

Kit (*ThermoScientific*), in accordance with the manufacturer's instructions. The extracted plasmids were evaluated by DNA sequencing using a commercial service provider (Eurofins MWG Operon). A list of all plasmids prepared for this work is provided in **Table S2** in Appendix.

2.7.4. Transformation of NSARI *A. oryzae*

Spores of *A. oryzae* were inoculated into 20 mL of DPY media (Per 1 L of water: 20 g dextrose, 10 g peptone, 5 g yeast extract, 0.5 g $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 5 g KH_2PO_4 ; pH adjusted to 5.5) and incubated in a sterile Erlenmeyer flask for 3-4 days on a shaking incubator at 160 RPM and 30°C. The *A. oryzae* tissue was pressed through a sterile syringe with cotton, and the pellet placed into 10 mL of 'TF1' solution (Per 10 mL of water: 0.058 g maleic acid, 0.79 g $(\text{NH}_4)_2\text{SO}_4$, 0.1 g yatalase (*Clontech*); pH adjusted to 5.5). This was incubated for two hours at 30°C and 160 RPM, and the tissue was pressed within a sterile syringe with cotton to collect the protoplasts within the supernatant. The protoplast solution was centrifuged at 1500 RPM for 10 minutes to discard the supernatant. The protoplasts were washed by resuspension in 10 mL of 'TF2' solution pre-warmed at 30°C (Per 1L of water: 218.5 g sorbitol, 10.95 g $\text{CaCl}_2 \cdot 6\text{H}_2\text{O}$, 2.05 g NaCl, 1.21 g Tris, pH adjusted to 7.5), then centrifuged at 1500 RPM for 10 minutes to discard the supernatant. The protoplasts were resuspended in 1-2 mL of TF2 solution, and 200 μL of protoplast solution was placed into 15 mL volume centrifuge tubes. A total of 10 μL of 1000 ng/ μL of each transformation plasmid was added to the 200 μL protoplast solution and then left to incubate at room temperature for 30 minutes. Sequentially and slowly, 250 μL , 250 μL , and 850 μL aliquots of 'TF3' solution (Per 1 L of water: 600 g Polyethylene glycol (4000), 10.95 g $\text{CaCl}_2 \cdot 6\text{H}_2\text{O}$, 1.21 g Tris, pH adjusted to 7.5) were added to the protoplast solution, then left to incubate at room temperature for 30 minutes. A total of 5 mL of TF2 solution was added to each protoplast solution then centrifuged for 10 minutes at 1500 RPM to discard the supernatant. The protoplast pellet was resuspended in

500 μ L of TF2. To create a two-layer agar plate, a ‘bottom layer’ was first prepared and left to solidify (Per 1 L of water: 2 g NH_4Cl , 1 g $(\text{NH}_4)_2\text{SO}_4$, 0.5 g KCl , 0.5 g NaCl , 1 g KH_2PO_4 , 0.5 g $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 0.02 g FeSO_4 , 218.6 g sorbitol, 15 g agar [Depending on selection markers used, add 1 g arginine, 1.5 g methionine and/or 0.1 g adenine], pH adjusted to 5.5). The 500 μ L of protoplast suspension was then mixed with 5 mL of a liquid ‘top layer’ pre-warmed to 50°C (Identical in composition to ‘bottom layer’, but use 8 g/L of agar), then poured onto the ‘bottom layer’ and left to solidify. These plates were incubated upside-down in a stationary incubator at 30°C for 4-7 days until *A. oryzae* colonies appeared on plates. Spores from transformed *A. oryzae* were inoculated into plates lacking sorbitol (Identical in composition to ‘bottom layer’ but omit sorbitol) using a sterile toothpick. These plates were incubated upside-down at 30°C for 4-7 days until *A. oryzae* colonies appeared on plates. These *A. oryzae* colonies, each colony considered a replicate in experiments described in text, are considered ready for metabolite screening trials.

2.7.5. Incubation conditions and media extractions

Colonies of transformed *A. oryzae* was inoculated by sterile toothpick into Czapek-Dox media with added starch and adenine (Per 1L of water: 3 g NaNO_3 , 2 g KCl , 10 g peptone, 0.5 g $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 1 g KH_2PO_4 , 0.02 g FeSO_4 , 20 g starch, 0.1 g adenine, pH adjusted to 5.5). To dissolve starch, bring about 800 mL of water to a boil, then slowly add a starch-water slurry to the boiling water while stirring. Once the solution becomes clear, add all other ingredients, then bring final volume of water to 1 L. Do not attempt to dissolve the starch by adding starch powder to a suspension followed by autoclaving – the starch will form a gelatinous ‘puck’ at the bottom of the flask. A total of 100 mL of such Czapek-Dox media was placed in 500 mL baffled Erlenmeyer flasks, inoculated with *A. oryzae* spores, and grown on a shaking incubator for five days at 30°C and 160 RPM. To extract metabolites, the *A. oryzae* tissue was first filtered from the broth using

cheesecloth. The broth was acidified with two drops of concentrated HCl, twice extracted with 75 mL of ethyl acetate, dried with sodium sulfate, evaporated on rotary evaporator, and the metabolites resuspended in methanol. The weight of the dry extracts varied from 5 mg to 50 mg.

2.7.6. Reverse transcriptase-polymerase chain reaction (RT-PCR)

Transformed colonies of *A. oryzae*, or *C. uncialis*, were incubated in Czapek-Dox media (See ‘*Incubation Conditions and Metabolite Extractions*’, above). Approximately 0.05 g of fungal tissue was frozen with liquid nitrogen and ground with mortar and pestle. Total mRNA was extracted using the E.Z.N.A. Fungal RNA Mini Kit (*Omega*), in accordance with the manufacturer’s instructions. An aliquot of mRNA extract was set aside for analysis, and the remainder was treated with DNase (*ThermoScientific*). An aliquot of DNase-treated mRNA was set aside for analysis, and reverse transcription was applied to the remainder using the RevertAid H Minus First Strand cDNA Synthesis Kit (*ThermoScientific*) in accordance with the manufacturer’s instructions. Aliquots of the DNA-contaminated mRNA and DNase-treated mRNA were cleaned using a PCR cleanup kit (*ThermoScientific*). The efficacy of DNase was evaluated by using the cleaned aliquots as templates for PCR reactions (See ‘*PCR*’ and ‘*Plasmid Construction*’, above, and Chapters 7 and 8, in main text).

2.7.7. Phylogenetics

Amino acid sequences of *C. uncialis* PKS bearing domain architecture of SAT-KS-AT-PT-ACP-TE were selected for analysis (Accession no. AUW31040, AUW31256, AUW31152, AUW31210, AUW31200, AUW31177, AUW31139, AUW31121). A reducing PKS from *C. uncialis* was selected as an out-group (Accession no. AUW31355). Five non-lichen PKS that are experimentally characterized as orsellinic acid-producing PKS were included (Accession no. CBF73505, XP_001835415, CEF78872, AST08390, BAZ95871) (Yu et al. 2017; Ishiuchi et al.

2012; Okada et al. 2017; Jørgensen et al. 2014; Sanchez et al. 2010). Multiple amino acid sequence alignment was performed using Clustal W (Thompson et al. 1994). Phylogenetic tree was constructed using MEGA7 (Kumar et al. 2016). Evolutionary history was inferred using the neighbour-joining method (Saitou & Nei, 1987). The assembled tree was evaluated with 1000 replicates of the bootstrap test (Felsenstein, 1985). Bootstrap tests with a score of 60 % or greater are reported above each node. The evolutionary distance was estimated using the Poisson correction method (Zuckerkandl & Pauling, 1965).

2.7.8. HPLC

Analysis by HPLC on extracts of *A. oryzae* media was performed using a Waters HPLC Separations Module 2695 combined with a PDA Detector Model 2996. The column was a μ Bondapak® Waters C₁₈ (3.9 X 300 mm) column, with a particle diameter of 15-20 μ m with 125 Å pores. The flow rate was 1 mL/min and the eluent was monitored continuously at 210-600 nm. The elution was done at 20 % methanol and 80 % water containing 0.075 % aqueous trifluoroacetic acid for 10 minutes, following which a linear gradient was applied to 80 % methanol and held at that composition for 20 minutes, followed by application of a linear gradient back to 20 % methanol for 10 minutes and held there for 10 minutes. The total run time was 60 minutes.

2.7.9. SDSPAGE

A stacking gel was prepared within Mini-Protean System Glass Plates (*BioRad*) to a final acrylamide concentration of 7.5 % or 10.0 % (Resolving layer, per 20 mL: 4.68 mL or 6.24 mL of 40% 29/1 acrylamide/bisacrylamide, 7.50 mL of 1 M Tris (pH 8.8), 0.20 mL of 10 % ammonium persulfate, 0.02 mL of TEMED, 7.86 mL or 9.42 mL of water; Stacking layer, per 10 mL: 0.93 mL or 1.24 mL of 40% 29/1 acrylamide/bisacrylamide, 3.76 mL of 1M Tris (pH 6.8), 0.10 mL of 10 % ammonium persulfate, 0.01 mL of TEMED, 5.20 mL or 5.51 mL of water). The

concentration of protein samples was measured using the Bradford technique (Bradford, 1976). Protein samples were mixed in a 3:1 ratio with 4x loading buffer (2 mL of 1 M Tris (pH 6.8), 0.8 g of sodium dodecyl sulfate, 4 mL of glycerol, 0.4 mL of 14.7 M β -mercaptoethanol, 1 mL of 0.5 M EDTA, 0.008 g of bromophenol blue, water to 10 mL). Proteins were denatured by incubation at 95°C for 5 minutes, and 15 mg of total protein extract was loaded into each well. Molecular weight standards used were either 5 μ L of PageRuler Plus Pre-stained Protein Ladder (*ThermoScientific*) or 10 μ L of HiMark Pre-stained Protein Standard (*Invitrogen*). Electrophoresis was performed using a Mini-Protean Tetra System (*BioRad*) and a Powerpack (VWR) set to 230 mA and 200 V. Samples were electrophoresed under constant current for 1 to 3 hours.

2.7.10. Western Blot

Two Trans-Blot Turbo Mini-size Transfer Stacks (*BioRad*) and the SDSPAGE gel were washed/wetted with Trans-Blot Turbo Transfer Buffer (*BioRad*) and assembled into a tower along with a methanol-activated Immun-Blot PVDF membrane (*BioRad*), in accordance with the manufacturer's instructions. Blotting was performed using a Trans-Blot Turbo system (*Biorad*) set to 1.3 A, 25 V, and 7 min. The tower was disassembled, and the membrane was washed for five minutes on an end-to-end rotator in 10 mL of TBS solution (Per 1 L of water: 3 g Tris, 0.2 g KCl, 8 g NaCl, pH adjusted to 7.4). Blocking was performed by incubating the membrane in a 5 % milk powder-TBS solution for one hour on rotator. Primary antibody application was provided by applying for one hour on rotator 10 ml of 5 % milk powder-TBS solution containing 5 μ L of Novex 6x-His Epitope Tag Antibody (*Invitrogen*). Membrane was washed three times with TBS solution in ten-minute applications. Secondary antibody application was provided by applying for one hour on rotator 10 ml of 5 % milk powder-TBS solution containing 1 μ L of Goat Anti-Mouse IgG (*EMD Millipore*). Membrane was washed three times with TBS solution in ten-minute

applications. Chemiluminescence was performed by applying 1.5 mL of Immobilon Forte Western HRP Substrate (*EMD Millipore*) for five minutes on rotator. Membranes were imaged using a FluorChem Q system (*Cell Biosciences*) in accordance with manufacturer's instructions.

2.8. Materials and methods related to Chapter 8

Refer to 'Materials and methods related to Chapter 7' and subsections therein.

2.9. Materials and methods related to Chapter 9

Not applicable

Chapter 3

Identification of Usnic Acid Gene Cluster by Deductive Method

3.1. Introduction - Selecting a suitable lichen and secondary metabolite to study

To explore secondary metabolite biosynthesis in lichens, it would be necessary to identify the enzyme-encoding genes that are responsible for the biosynthesis of lichen secondary metabolites. The precise biosynthetic function of genes can be elucidated by inserting genes into a host with the expectation of observing *de novo* biosynthesis of a molecule within the host. These experiments would allow us to conclusively determine what the encoded enzymes do, and how the secondary metabolite is produced via a step-by-step series of biosynthetic steps. This is a process known as ‘functional heterologous expression’ and has been widely used to explore secondary metabolite biosynthesis in bacteria, plants, and fungi (Zhang et al. 2016; Alberti et al. 2017; Anyaogu & Mortensen, 2015; Yaegashi et al. 2014; Billingsley et al. 2016; Li & Neubauer, 2014; Loeschcke & Thies, 2015; Gomez-Escribano & Bibb, 2014; Baltz, 2010; Ikram et al. 2015). As functional heterologous expression has not been successfully reported in lichens, we would first need to develop a procedure for the heterologous expression of lichen genes. The development of this procedure would be informed by methods well-established for expressing non-lichen fungal genes in heterologous hosts. However, to test this procedure, we would need to identify one or more biosynthetic genes from a species of lichen that can serve as “test subjects”. If the lichen metabolite was observed within the heterologous host following the application of this procedure, this would confirm that the procedure is working. This procedure could then be used to elucidate the biosynthetic pathways of other lichen metabolites and their associated genes within the lichen secondary metabolome.

We therefore first sought to identify one or more candidate biosynthetic genes to serve as “test subjects” for developing this procedure. The following points were taken into consideration:

- (1) These genes should encode a known and well-characterized metabolite, as it would be foolhardy to attempt to use genes that have no association with a known secondary metabolite.
- (2) The associated metabolite is ideally produced by a species of lichen that is easily accessible and identifiable.
- (3) This metabolite would also ideally have a commercial supplier so that a chemical reference standard would be available for all heterologous expression trials.
- (4) The metabolite should have a deciphered biosynthetic pathway containing a small number of biochemical steps.
- (5) This deciphered pathway should allow us to easily identify its gene cluster.

Meeting the two latter considerations would greatly assist with the identification of the genes responsible for its biosynthesis and would also limit potential complications owing to host compatibility issues.

Usnic acid, a polyketide first isolated from lichens in 1844, was earmarked as an obvious candidate for these experiments (**Figure 16**). Usnic acid is the most widely studied lichen secondary metabolite. Firstly, this molecule possesses a variety of pharmacological properties including anti-inflammatory, antioxidant, antimicrobial, antiprotozoal, antiviral, and larvicidal effects (Araújo et al. 2015). This provides us with a rich literature from which to draw upon for consultation and an imperative for its scientific value to lichenology. Secondly, usnic acid is widely distributed among common species of lichens. In some cases, such as the eponymous *Usnea* genus of lichens, usnic acid is the major chemical substituent (Liao et al. 2010). Acquiring species of lichen that produce usnic acid is therefore a trivial matter. Thirdly, chemical standards of usnic acid are available from commercial supplies such as *Sigma-Aldrich*.

Summary of isotope-feeding experiments

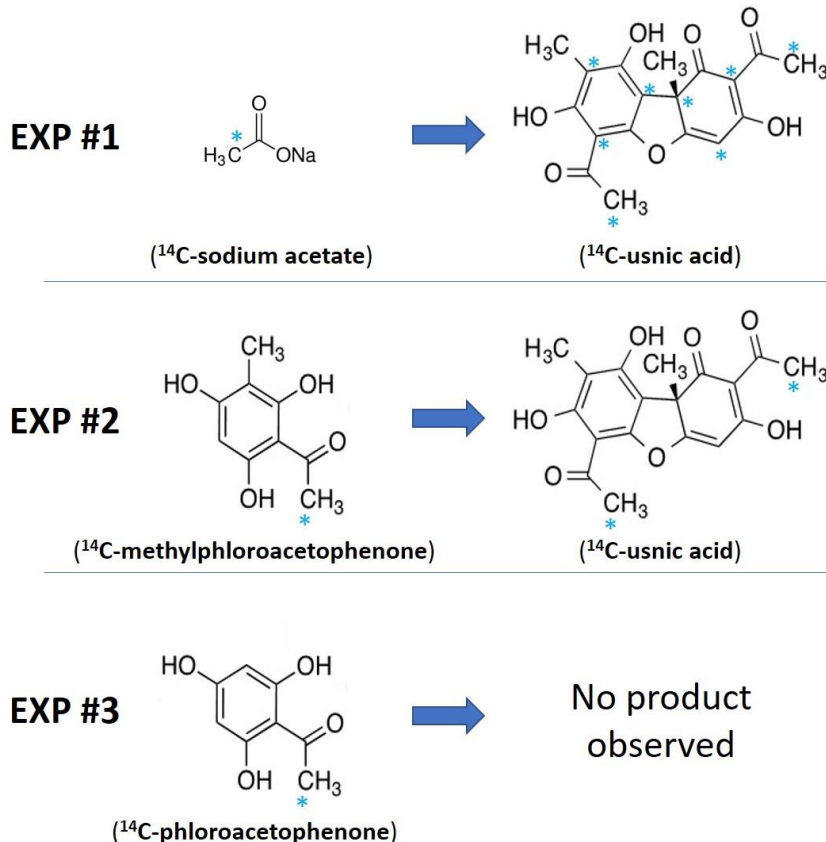


Figure 16: Summary of the isotope-feeding performed by Taguchi et al. (1969) that led to the elucidation of the usnic acid biosynthetic pathway in lichens.

Fourthly, thanks to elegant isotope-labelling by Taguchi et al. (1969), the biosynthetic pathway of usnic acid is known (**Figure 16**). Feeding lichens with isotope-labelled acetate results in an enrichment pattern consistent with polyketide biosynthesis ('EXP #1' of **Figure 16**). If isotope-labelled methylphloroacetophenone is fed to lichens, enriched usnic acid can be re-isolated, demonstrating that usnic acid is a dimer of two molecules of methylphloroacetophenone ('EXP #2' of **Figure 16**). If isotope-labelled phloroacetophenone is fed to lichens, enriched usnic acid cannot be re-isolated. This demonstrates that the methyl group is added during the biosynthesis of methylphloroacetophenone by the PKS ('EXP #3' of **Figure 16**). As we now know

today, such additions are performed by PKS using a C-methyltransferase (MT) domain using S-adenosyl methionine (SAM) as substrate.

Fifthly, oxidation experiments suggest that usnic acid is most likely formed through the oxidative coupling of two units of methylphloroacetophenone (**Figure 17A**). Single-electron oxidants such as horseradish peroxidase (in the presence of peroxide) or ferricyanide may dimerize methylphloroacetophenone into usnic acid *in vitro* (Barton et al. 1956; Penttilä & Fales, 1966; Hawranik et al. 2009). These experiments lend support toward a radical theory of dimerization (**Figure 17A**). In nature, this dimerization could be plausibly performed by a single oxidative enzyme. A cytochrome p450 oxidase would be an obvious candidate for such a reaction as this class of enzyme is known to perform carbon-carbon oxidative coupling reactions (Bernhardt 2006; Guengerich & Munro, 2013; Isin and Guengerich 2007).

Two enantiomers of usnic acid are observable in nature and have been observed to have enantiospecific pharmacological effects (Galanty et al. 2019) (**Figure 17B**). Some genera of lichens exclusively produce (+) usnic acid (e.g., *Usnea* sp.) whereas the (-) enantiomer is found in other genera (e.g., *Alectoria* sp.). A mixture of both enantiomers can be found in some lichens (e.g., *Vulpicidia*). In other genera, such as *Cladonia* sp., no specific pattern is observed between members of this genus (Galanty et al. 2019). It is possible that the hypothetical oxidative enzyme that is responsible for dimerizing methylphloroacetophenone into usnic may display variable enantioselectivity across lichen species. For example, the enzyme in *Usnea* sp. would enantioselectively produce (+) usnic acid, whereas this same enzyme in *Alectoria* sp. would instead enantioselectively produce (-) usnic acid, and in other genera such as in *Vulpicidia* sp., no or limited stereoselectivity would be observed (**Figure 17B**). Should it be true that only a single oxidative enzyme is required, it would be interesting (from the standpoint of structural

enzymology) how such an enzyme can radically oxidize two molecules then dimerize them together in a stereoselective manner.

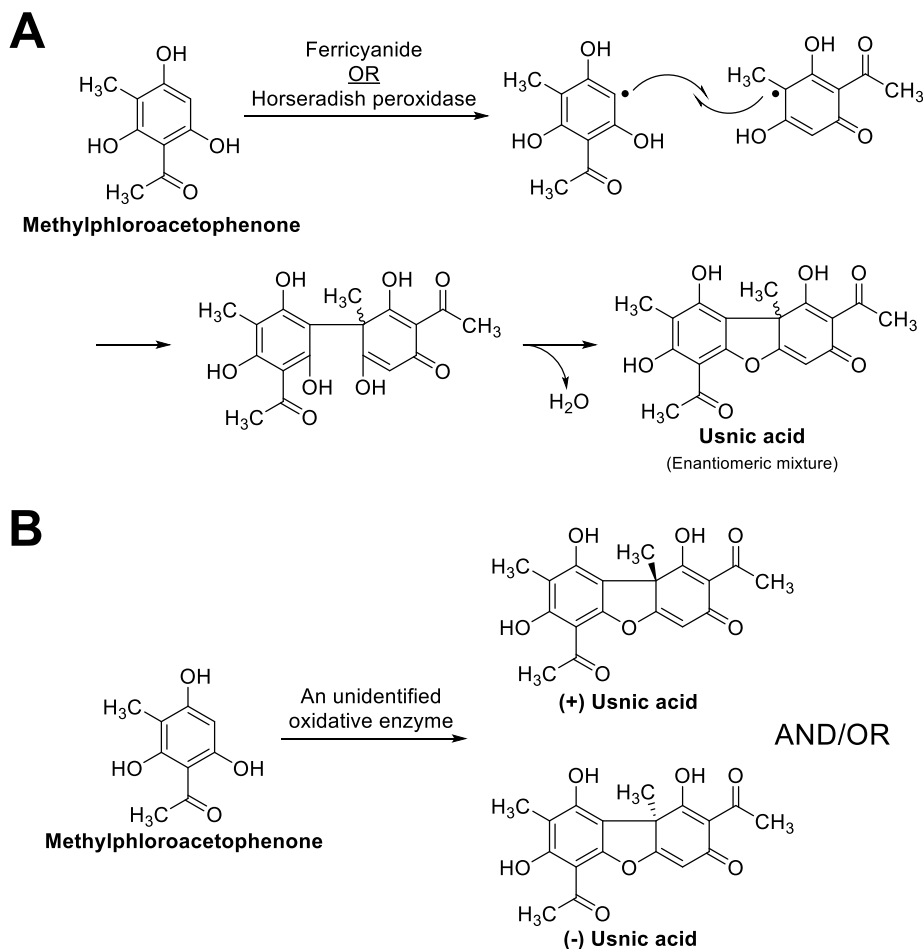


Figure 17: Radical coupling mechanism of methylphloroacetophenone. (A) Single-electron oxidants such as horseradish peroxidase (with H_2O_2) and ferricyanide are observed to dimerize methylphloroacetophenone into usnic acid *in vitro*, lending support to a radical mechanism of usnic acid formation *in vivo* (Barton et al. 1956; Penttilä & Fales, 1966; Hawranik et al. 2009). (B) Biosynthesis of usnic acid in nature is proposed to be facilitated by a single oxidative enzyme. Such radical coupling may variably proceed using either an enantiospecific or non-specific mechanism, in accordance with the distribution of (+) and (-) enantiomers of usnic acid found throughout lichen taxa (Gelantý et al. 2019).

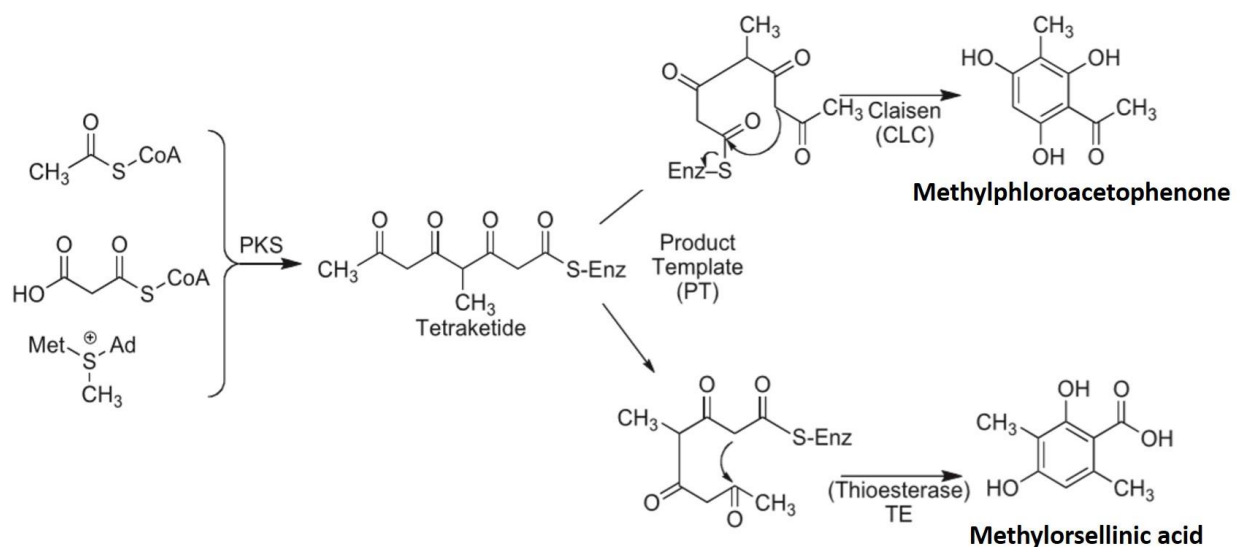


Figure 18: Aldol-style versus Claisen-style folding of C_8 -polyketides by CLC and TE domains. Reproduced from Abdel-Hameed et al. (2016a), in accordance with authors' retainment of privileges.

Lastly, the biosynthetic pathway elucidated by Taguchi et al. (1969) allows us to predict the architecture of the gene cluster that would be responsible for producing usnic acid. The placement and oxidation state of oxygen atoms present on methylphloroacetophenone suggests that the PKS must be a non-reducing PKS. The isotope-labelling experiments by Taguchi et al. (1969) also demonstrate that the PKS must also have a *C*-methyltransferase domain. To produce methylphloroacetophenone, a relatively rare terminal domain known as a Claisen cyclase (CLC) is required. Whereas C_8 -polyketides are typically folded using an Aldol reaction via a thioesterase (TE) domain, the production of methylphloroacetophenone requires Claisen condensation (**Figure 18**). Hence, this PKS must have a terminal CLC domain. In consideration of the oxidative coupling experiments described previously (Barton et al. 1956; Penttilä & Fales, 1966; Hawranik et al. 2009), we would also expect an oxidative enzyme to be encoded near the PKS. As will be shown later in this chapter, a compilation of candidate PKS genes can be evaluated based on these expected characteristics, with the expectation of finding at least one gene cluster that meets the

expected requirements of usnic acid biosynthesis. We assigned names to the hypothetical PKS and p450 as methylphloroacetophenone synthase (MPAS) and methylphloroacetophenone oxidase (MPAO), respectively (**Figure 19**).

Figure 19: The hypothesized usnic acid biosynthetic pathway. One unit of acetyl CoA, three units of malonyl CoA, and one unit of *S*-adenosyl methionine are assembled by a polyketide synthase named methylphloroacetophenone synthase (MPAS). Two units of the resultant methylphloroacetophenone are then oxidatively dimerized into usnic acid by a hypothetical oxidative enzyme which we have named methylphloroacetophenone oxidase (MPAO). Adapted from Abdel-Hameed et al. (2016a), in accordance with authors' retainment of privileges.

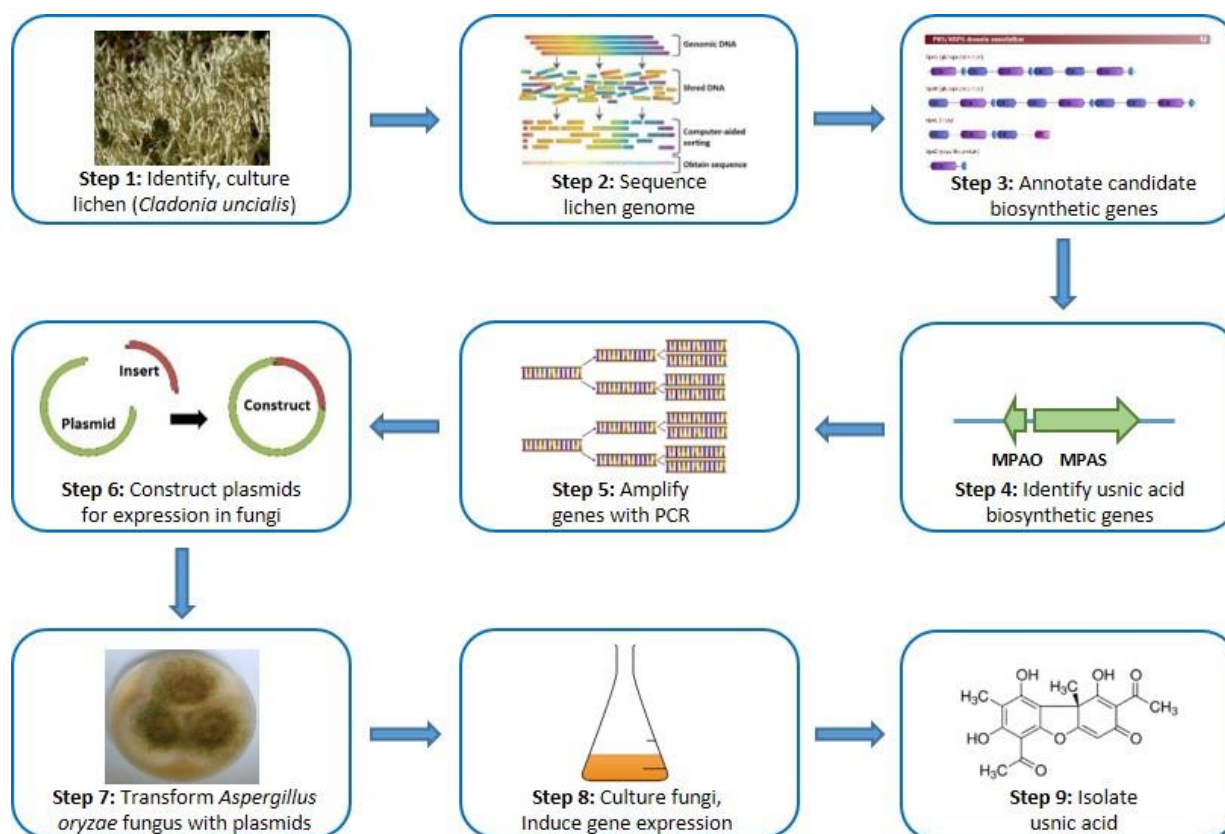


Figure 20: General experimental approach for the heterologous expression of the usnic acid gene cluster. For steps 1-4, read onwards in this chapter. For steps 5-9, see Chapters 7 and 8.

A general experimental scheme began to emerge (**Figure 20**). We would first need to sub-culture the *C. uncialis* fungus from its photosynthetic partner ('STEP 1' of **Figure 20**). We would then need to sequence its genome, assemble the raw reads, and annotate its biosynthetic gene clusters ('STEP 2' and 'STEP 3' of **Figure 20**). Using the architectural clues outlined above, we would then need to identify the usnic acid gene cluster ('STEP 4' of **Figure 20**). We would then need to prepare these genes for heterologous expression ('STEP 5' to 'STEP 9' of **Figure 20**). These latter steps will be addressed in Chapters 7 and 8. For now, we will focus on the first four steps of this experimental outline (**Figure 20**).

3.2. *De novo* whole-genome sequencing of an usnic acid-producing species of lichen

Dr. Piercey-Normore morphologically identified the lichen that she provided to us as *Cladonia uncialis*. To provide a genetic identification of species, we amplified by PCR the internal transcriber spacer (ITS) region and sequenced the amplicon using a conventional Sanger platform (see Chapter 2). The ITS region is used as a ‘DNA barcode’ to identify species within kingdom *Fungi* (Schoch et al. 2012; Xu et al. 2016). The reason for this convention among mycologists is because the ITS region is easy to amplify by PCR and because it has a high degree of variation between species, thereby making species identification a simple process (Schoch et al. 2012; Xu et al. 2016). By comparing our sequencing results to ITS sequences available in GenBank, we can conclusively determine whether the specimen provided to us was indeed *C. uncialis*. The resultant ITS sequence (GenBank accession no. KY80361) was revealed to be 95 percent similar to GenBank-deposited ITS sequences assigned to *C. uncialis* (**Table S3** in Appendix). This was interpreted as evidence that the lichen was indeed *C. uncialis*. We now turned our attention towards sequencing its genome.

Lichens are a partnership between a fungus and an algae or cyanobacterium. Although the photosynthetic partner is known to produce a few of the secondary metabolites arising from the lichen partnership (e.g., see Kassalainen et al. (2012), Kampa et al. (2013)), the fungal partner is responsible for the overwhelming majority of the lichen secondary metabolome. It will therefore be necessary to sequence the genome of the fungal partner to find the usnic acid gene cluster. To avoid downstream complications arising from contaminating species, we needed to separate the fungal partner from the algal partner to create an isolated specimen of *C. uncialis* (an “axenic culture”). Spores from *C. uncialis* were collected onto a water-agar media, then transferred to nutrient-rich agar (see Chapter 2). To generate enough biomass for *de novo* whole-genome

sequencing, this fungus was grown on agar for a period of one year. After one year, the ITS region was again sequenced using an identical procedure to provide a redundant confirmation of species identity. The resultant ITS sequence (Accession no. KY80362) confirmed that the species of the isolated fungus was *C. uncialis* (Table S4 in Appendix). We interpreted this result as evidence that we successfully isolated *C. uncialis* fungus from its photosynthetic partner.

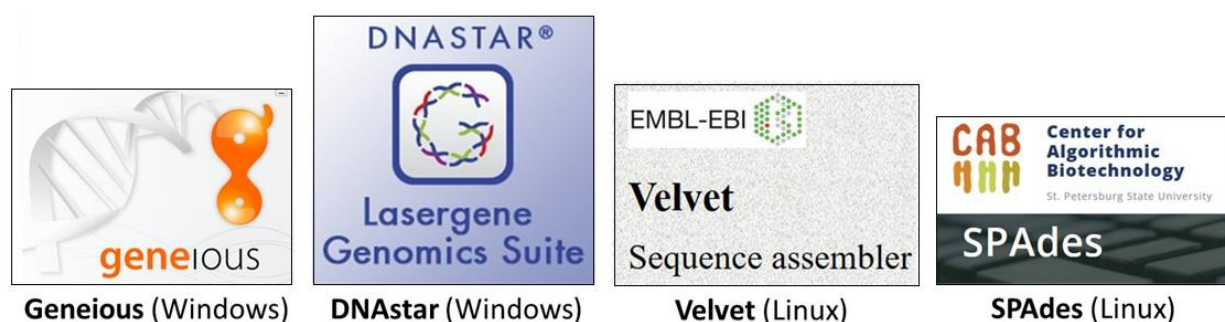


Figure 21: Assembly programs used in this work (Kearse et al. 2012; Bankevich et al. 2012; Zerbino & Birney, 2008). The assembly provided by SPAdes was used for the annotation of secondary metabolite gene clusters.

The genomic DNA of *C. uncialis* was isolated and sequenced using an Illumina platform, then assembled using SPAdes (see Chapter 2) (Kearse et al. 2012). Contigs generated from SPAdes were submitted to the Antibiotic and Secondary Metabolites Analysis Shell (AntiSMASH V. 2.0) (Blin et al. 2013). This program identifies genes associated with secondary metabolism. The creators of AntiSMASH compiled a library of protein motifs that are known to be associated with secondary metabolism, such as the protein motifs of PKS and NRPS (Medema et al. 2011). This library also includes motifs from genes associated with primary metabolism (e.g. fatty acid synthases) to mitigate false-positives generated by AntiSMASH. AntiSMASH predicts genes within queried DNA sequences, then searches the amino acid translations of those genes for protein motifs that are consistent with the library. When motifs indicative of a biosynthetic gene (e.g., a polyketide synthase) are discovered, the program then expands its analysis of the nucleotides flanking the biosynthetic gene to search for accessory genes (e.g., P450, *O*-methyltransferase,

halogenase). When an accessory gene is identified, AntiSMASH repeats the process by searching the nucleotides adjacent to the last known accessory gene to attempt to find more accessory genes. This ‘greedy methodology’ ensures that a complete biosynthetic gene cluster is likely captured by the AntiSMASH analysis process (Medema et al. 2011). Subsequent versions of AntiSMASH have since been developed (Blin et al. 2013; Weber et al. 2015; Blin et al. 2017). These later versions can process more information, detect a greater range of genes, and have built-in additional analysis tools that assists the user in generating predictions of biosynthetic function. This latter feature of AntiSMASH will be used in Chapters 4 and 6. Although the original analysis of the *C. uncialis* secondary metabolome was reported using AntiSMASH V. 2.0 (Blin et al. 2013), the *C. uncialis* contig assemblies were re-analyzed using the latest version of AntiSMASH (V. 4.0) (Blin et al. 2017) to ensure that gene clusters reported within this thesis are comprehensive and accurate.

A total of 32 type I PKS genes were identified within the assembled contigs. To avoid redundancy within this thesis, these 32 PKS gene clusters are displayed and discussed in detail in Chapter 5. To provide as thorough an annotation as possible, the contigs containing these 32 type I PKS were manually reviewed. This manual review was performed by reviewing the open reading frames (“ORFs”) of genes identified by AntiSMASH. As programs may on occasion miss an ORF or merge two ORFs together, a separate program (AUGUSTUS) was used to review the ORFs generated by AntiSMASH and provide corrections to the annotation wherever necessary (Stanke et al. 2004; Stanke & Morgenstern, 2005). Each gene identified was also submitted to the Basic Local Alignment Search Tool (BLAST) (Altschul et al. 1990) and the probable function of each gene was noted based on consensus similarity to characterized genes deposited in GenBank. If an anomalous result was observed, e.g. with respect to estimated gene length, the putative ORF was re-evaluated. This iterative process was applied to the entire contig containing a putative

biosynthetic gene cluster. This process resulted in the production of contigs that were completely annotated of all putative genes. This process allowed us to confidently apply our deductive approach on the PKS-containing contigs.

3.3. Identifying the usnic acid biosynthetic gene cluster in *C. uncialis*

We know from isotope-labelling experiments (Taguchi et al. 1969) that usnic acid must be produced by a non-reducing PKS possessing methyltransferase (cMeT) and Claisen cyclase (CLC) domains. A single oxidative enzyme, most probably a cytochrome p450, should be found near the PKS. We therefore evaluated the 32 PKS genes to determine if any of these PKS met these criteria. A total of 14 non-reducing and 18 reducing PKS genes were observed (see Chapter 5). As the usnic acid-associated PKS must be a non-reducing PKS, we focused our attention on the 14 candidate non-reducing PKS. Of these 14 remaining genes, only five had the requisite MT domain (**Figure 22**). Of these five remaining PKS, genes #2 to #5 possessed terminal thioesterase (TE) domains, an architectural feature that was inconsistent with usnic acid biosynthesis. In addition, these four genes possessed post-synthetic tailoring genes encoding enzymes that performed extraneous functions, for example, nitrogen-adding enzymes and halogenases (**Figure 22**). Genes #2 to #5 were therefore ruled out as plausible usnic acid-associated genes. Only one PKS, gene #1, was consistent with the biosynthetic requirements of usnic acid (**Figure 22**). This gene cluster was therefore identified as the hypothesized usnic acid gene cluster (**Figure 23**). The cluster-bearing contig (MG777490), *mpas* (AUW31052) and *mpao* (AUW31051) have been deposited in GenBank and are accessible through their respective accession numbers.

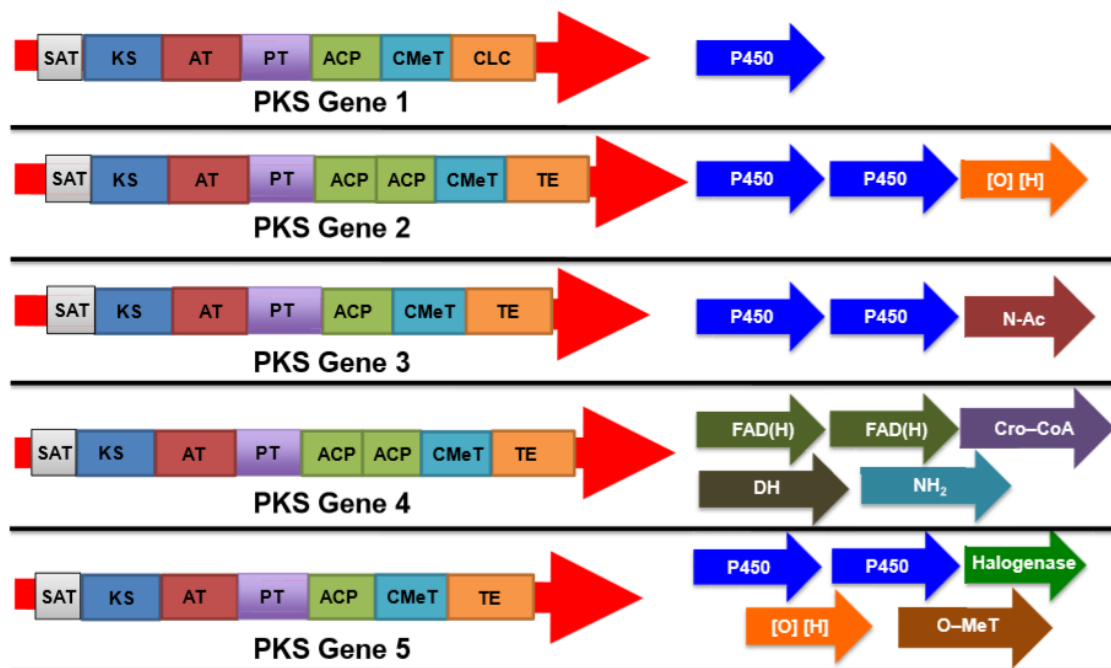


Figure 22: Of the 32 type I PKS genes identified in the assembled contigs of *C. uncialis*, only five were non-reducing PKS that possessed C-methyltransferase (cMeT) domains. From these five, ‘PKS Gene 1’ was deduced to be the usnic acid gene cluster based on the presence of a terminal Claisen cyclase (CLC) domain and a single cytochrome p450 found near this PKS. Figure reproduced from Abdel-Hameed et al. (2016a), in accordance with authors’ retainment of privileges.

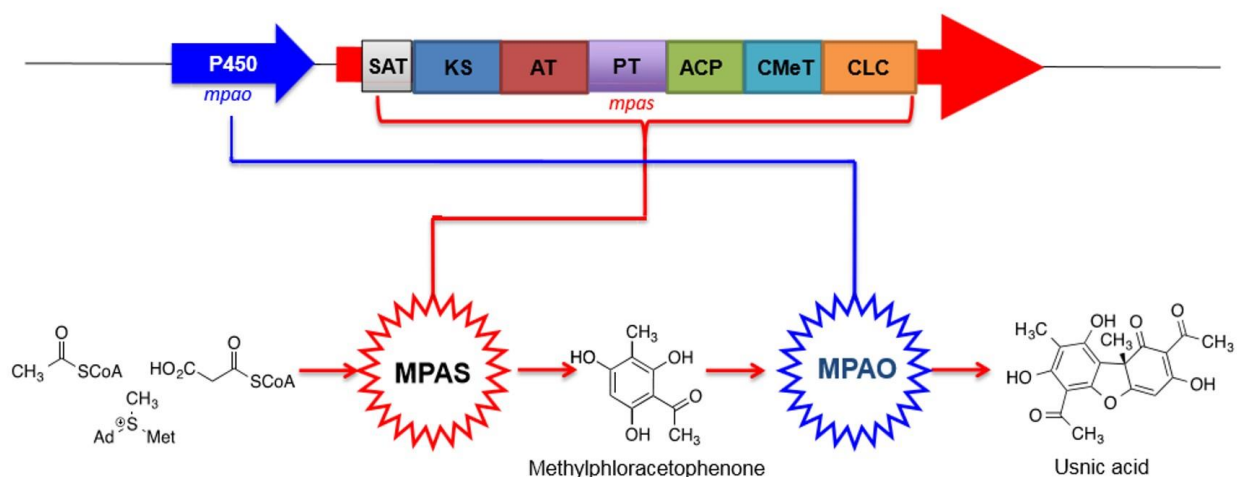


Figure 23: The proposed usnic acid gene cluster and biosynthetic pathway. Figure reproduced from Abdel-Hameed et al. (2016a), in accordance with authors’ retainment of privileges.

We were interested in determining whether *mpas* could be used to identify genetically homologous genes. This analysis could reveal whether usnic acid-related PKS have already been

sequenced in other species of lichen but have yet to be functionally assigned as the usnic acid-related PKS. This could be done using a phylogenetic analysis. Phylogenetics is the study of the evolutionary relationship between organisms. This is done by comparing nucleotide or amino acid sequences between compilations of genes or genomes.³ A core premise of phylogenetics (and of evolutionary biology) is that genes with a more recent and common evolutionary origin are more likely to retain similar biosynthetic functions than genes that are distantly related. In the case of secondary metabolite biosynthesis studies, identifying closely related genes could be used as supporting evidence that the two genes share a common biosynthetic function. Phylogenetics is therefore commonly used to predict the metabolites that are produced by experimentally uncharacterized lichen PKS (Schmitt et al. 2008; Timsina et al. 2012; Timsina et al. 2014).

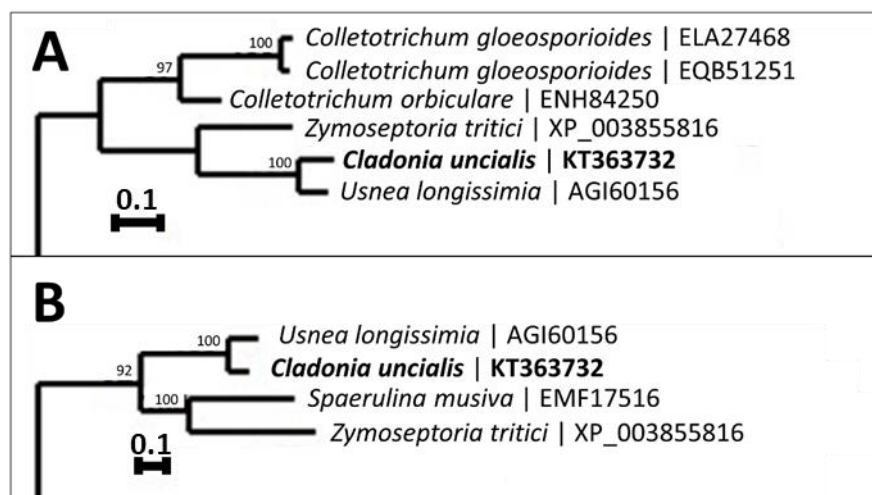


Figure 24: Truncated phylogenetic trees of (A) the KS domain of *mpas*, and (B) the CLC domain of *mpas*. Scale bar indicates amino acid substitutions per site. For the complete phylogenetic trees, see **Figure S1** and **Figure S2** in the Appendix. Adapted from Abdel-Hameed et al. (2016a), in accordance with authors' retainment of privileges.

Two sequence regions of *mpas* were used for the phylogenetic analysis. The KS domain is the most conserved of PKS domains and is often used in phylogenetics to probe the functions and evolutionary relationships of PKS genes (Timsina et al. 2014). We therefore used the KS

³ For an excellent beginner's review on performing and interpreting phylogenetic trees, please see Baldauf (2003).

domain to construct our first phylogenetic tree. The CLC domain of PKS is required to produce methylphloroacetophenone. We therefore used the CLC domain to construct our second phylogenetic tree because this domain was likely to identify other methylphloroacetophenone-producing PKS. In separate analyses, a total of 100 entries of amino acid sequences that were similar to the KS and CYC domains of *mpas* were compiled from GenBank and placed into a phylogenetic tree (see Chapter 2). The expected outcome was that *mpas* would aggregate within the phylogenetic tree with one or more of these GenBank entries. This result would be interpreted as supporting evidence of a homologue of *mpas* within another organism. In the resultant KS tree, we observed that *mpas* clustered with a single PKS gene (**Figure 24**). This gene is from the lichen species *Usnea longissima*, a known producer of usnic acid (Wang et al. 2014). In the resultant CLC tree, we observed that *mpas* also clustered with this same PKS gene from *U. longissima* (**Figure 24**). The identification of the same PKS gene from *U. longissima* using two phylogenetic analyses establishes a concordance of outcomes. We conclude that this gene from *U. longissima* likely encodes a PKS that produces the same molecule as that of *mpas* from *C. uncialis*, namely, methylphloroacetophenone⁴.

3.4. Summary

To study secondary metabolite biosynthesis in lichens we will need to functionally express lichen biosynthetic gene clusters within a heterologous host. As functional heterologous expression of lichen biosynthetic genes has yet to be successfully demonstrated by lichenologists,

⁴In 2018, a transcriptional profiling experiment on the lichen *Nephromopsis pallescens* was reported (Wang et al. 2018). A comparison of the transcription rates of seven candidate PKS genes identified one of these genes to be highly transcriptionally active under conditions wherein usnic acid was being produced. The approach that was reported is similar to that of Armaleo et al. (2011) in pursuit of the grayanic acid gene cluster (see Chapter 1). A phylogenetic analysis of the PKS from *N. pallescens* revealed this PKS to form a sub-clade with *mpas* from *C. uncialis* as well as the same PKS from *U. longissima* that we previously identified to be a homologue to *mpas* (this chapter). The fact that two independent approaches conducted by independent research groups have identified the same kin genes as the usnic acid-producing PKS provides compelling evidence that we have correctly identified the usnic acid gene cluster.

we first needed to choose a ‘test’ molecule (and its associated gene cluster) to use in developing the heterologous expression protocol. We chose usnic acid, a lichen polyketide with a biosynthetic pathway that is well-informed by isotope-labelling and oxidation experiments. To find its associated gene cluster, we sequenced the genome of *C. uncialis* and annotated its PKS genes. A total of 32 PKS gene clusters were identified. Using a deductive approach, one of these 32 PKS gene clusters was identified as a probable usnic acid gene cluster. Phylogenetic analysis suggests that a functional homologue of the PKS gene exists in *U. longissima*.

Chapter 4

6-Hydroxymellein and the ‘Homology Mapping’ Method

4.1. Introduction - Searching for the usnic acid gene cluster, *circa* 2008-2012.

Genome sequencing is expensive. Or, at least, it *was* expensive. One of the most dramatic developments in natural products research is the ability to sequence small genomes on a budget. This was achieved by the development of ‘next-generation’ sequencing platforms combined with the maturation of this technology within the marketplace (Goodwin et al. 2016). A consequence of this transformation is the emergence of genomics as a common element within the natural product sciences (Cacho et al. 2015).⁵ An example of the utility of genomics in natural products chemistry is provided in Chapter 3. However, before this marketplace transformation occurred, alternative approaches to finding genes were used. One alternative approach was ‘primer fishing’: By attempting PCR reactions using degenerate primers on cDNA generated from the total mRNA of an organism, it is possible to amplify transcriptionally active genes. The nucleotide sequence of those genes can then be determined using conventional Sanger platforms. To attempt to find the usnic acid-associated PKS, the previous Ph.D. candidate of our group, (now Dr.) Mona Abdel-Hameed, used primer fishing to produce a whole DNA sequence of a transcriptionally active PKS from *C. uncialis* (Abdel-Hameed, 2015). However, it became apparent that this was not the usnic acid-associated PKS. This was made evident by the absence of an MT domain and the presence of a terminal TE domain, two architectural characteristics inconsistent with usnic acid biosynthesis

⁵ To illustrate the importance of this commercial development, consider the following anecdote: We requested a quote from a service provider to sequence the *C. uncialis* genome. In 2008, this service was quoted for 100,000 (CAD). In 2013, this service was quoted for 1,000 (CAD).

(see Chapter 3). Our efforts to find the usnic acid gene cluster therefore stalled, at least until 2013, when the declining cost of ‘next-generation’ platforms permitted us to use this technology.

4.2. What does this PKS do?

We were intrigued – if not usnic acid, what molecule does the encoded PKS produce? Our genome assembly provided the first clues. This assembly revealed that this PKS was associated with four accessory genes to comprise a biosynthetic gene cluster (**Figure 25**). In addition to an *O*-methyltransferase and a monooxygenase, this cluster was notable for also encoding a non-heme iron-dependent halogenase: Halogenated secondary metabolites are rare and often associated with potent biological effects (Wagner et al. 2009). At the time of writing, only 13 halogenated natural products have been isolated from lichens, and this remains the only report of a halogenase gene from a lichen. In addition, a PKS-like gene that appeared to encode *only* a ketoreductase and dehydratase domain was also observed. Type I PKS employing domains that are encoded as separate enzymes are known as *trans*-PKS (Helfrich & Piel, 2016). This is a relatively rare configuration among type I PKS, and in our case, this was the first time anyone observed a *trans*-PKS within a lichenizing fungus. This gene cluster is indeed intriguing!

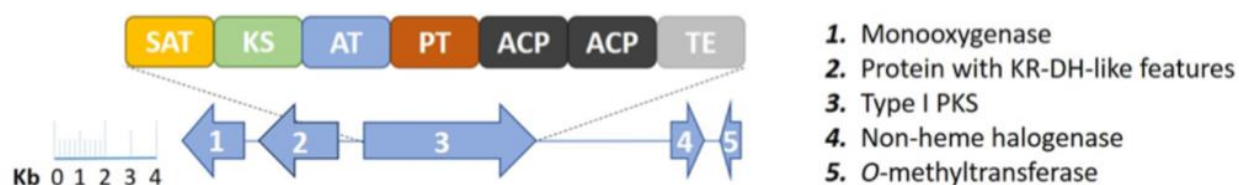


Figure 25: The gene cluster of unknown function, containing a transcriptionally active PKS, discovered by a combination of ‘primer fishing’ and *de novo* whole-genome sequencing. Reproduced from Abdel-Hameed et al. 2016b, in accordance with authors’ retention of privileges.

Though we now know that this gene cluster is associated with a halogenated product, we were unable to propose a specific molecule. We made further inroads to solving this mystery when the third version of AntiSMASH (V. 3.0) was publicly released (Weber et al. 2015). The previous

version, AntiSMASH V. 2.0 (Blin et al. 2013), was used to identify the PKS genes within our draft assembly of the *C. uncialis* genome (see Chapter 3). An added feature of AntiSMASH V. 3.0 was the integration of a ‘KnownClusterBLAST’ algorithm that compares the gene clusters identified within an imputed sequence to a compilation of well-characterized gene clusters (Medema et al. 2015). By submitting the lichen gene cluster to AntiSMASH (V. 3.0), this algorithm may find genetically similar gene clusters that have been experimentally associated with secondary metabolites. If a ‘hit’ was found, we could therefore use this information to propose a molecule that is associated with this lichen gene cluster. It was also possible that we could propose a biosynthetic pathway based on the information found.

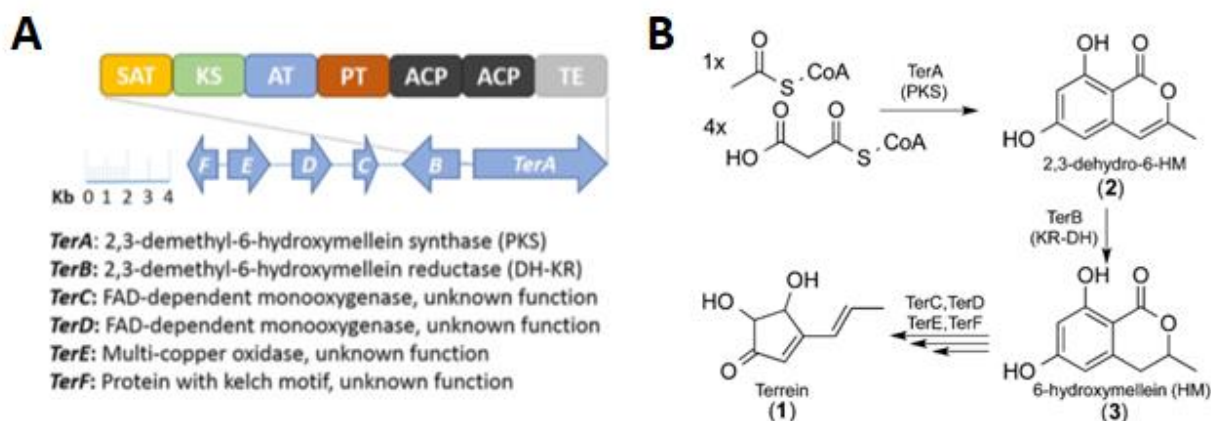


Figure 26: (A) Gene cluster associated with terrein biosynthesis in *A. terreus*. (B) Biosynthetic pathway of terrein in *A. terreus*. Reproduced from Abdel-Hameed et al. (2016b), in accordance with authors’ retainment of privileges.

AntiSMASH (V. 3.0) suggested a homologous relationship between our gene cluster in *C. uncialis* and a gene cluster in *A. terreus* known to produce terrein (Zaehle et al. 2014). Terrein biosynthesis requires six enzymes, encoded by genes *terA* to *terF* (**Figure 26**). The *A. terreus* pathway is notable for also possessing a *trans*-KR-DH-like protein (encoded as *terB*). The role of this PKS in *A. terreus* was deduced by Zaehle et al. (2014) to produce 2,3-dehydro-6-hydroxymellein. The deduced function of the *trans*-KR-DH-like protein was to perform a two-

electron reduction, yielding 6-hydroxymellein (**Figure 26**). The remaining enzymes (TerC to TerF) are monooxygenases or oxidases responsible for converting 6-hydroxymellein to terrein. Due to the instability of these latter intermediates, Zaehle et al. (2014) were unable to elucidate a complete biosynthetic pathway. What is known about terrein biosynthesis and its associated gene cluster is shown in **Figure 26**.

Table 2: BLAST similarity statistics between genes within *C. uncialis* and *A. terreus*. Reproduced from Abdel-Hameed et al. (2016b), in accordance with authors' retainment of privileges.

gene in cluster	accession no.	putative function	species highest homology (accession no.)	identity/query coverage	terrein pathway homologue (accession no.)	identity/query coverage
(1)	KU740324	monooxygenase	<i>C. metacorrallifera</i> (ADR00967)	79/99	<i>TerD</i> (XP_001210228)	68/99
(2)	KU740325	KR-DH peptide	<i>P. scopiformis</i> (KUJ09199)	67/99	<i>TerB</i> (XP_001210230)	46/79
(3)	KU740326	type I PKS	<i>C. grayi</i> (ADX36086)	73/99	<i>TerA</i> (XP_001210231)	59/99
(4)	KU740327	halogenase	<i>S. borealis</i> (ESZ96411)	56/98	N/A	N/A
(5)	KU740328	O-methyltransferase	<i>P. scopiformis</i> (KUJ06291)	77/99	N/A	N/A

To determine the degree of genetic similarity between the two gene clusters, we performed a comparative pBLAST analysis (Altschul et al. 1990). Three genetically similar pairs of genes were observed: (1) *C. uncialis* monooxygenase / *terD*; (2) *C. uncialis* KR-DH-like protein / *terB*; and (3) the *C. uncialis* PKS / *terA* (**Table 2**). No counterparts to the *C. uncialis* O-methyltransferase nor halogenase were observed within the *A. terreus* gene cluster.

We decided to conduct phylogenetic analyses to determine whether these three pairs of genes were genetically homologous. This would also lend support to the proposition that the products of these three pairs of genes share similar biosynthetic functions. A total of 30 genes that were genetically similar to the queried lichen gene were compiled from GenBank. These genes, as well as the *C. uncialis* / *A. terreus* gene-pairs, were then assembled into a phylogenetic tree (see Chapter 2). The expected outcome was that the pair of genes would cluster together within a common phylogenetic clade, lending supporting evidence to a conclusion of homology. This was indeed the observed result (**Figure 27**).

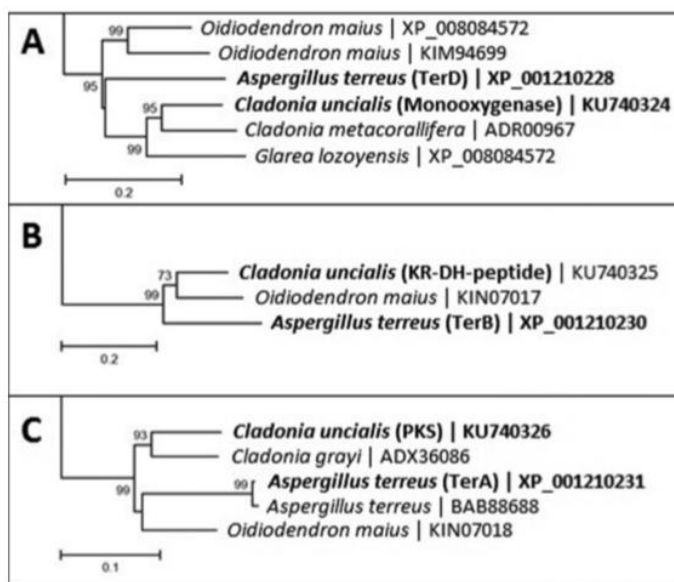


Figure 27: Truncated phylogenetic trees of (A) the monooxygenase, (B) the KR-DH-like gene, and (C) the PKS, from *C. uncialis* and *A. terreus*. Scale bars denote average number of substitutions per site. The complete phylogenetic trees are available as **Figure S3-S5** in the Appendix. Reproduced from Abdel-Hameed et al. (2016b), in accordance with authors' retainment of privileges.

This information allowed us to propose a biosynthetic pathway associated with the lichen gene cluster. We propose that the PKS and KR-DH-like protein in *C. uncialis* are responsible for producing 2,3-dehydro-6-hydroxymellein and 6-hydroxymellein, respectively (**Figure 28**). The 6-hydroxymellein would then be hydroxylated by the homologous monooxygenase. Based on speculative chemistry provided by Zaehle et al. (2014), this hydroxylation would most likely occur at either C-7 or C-5 (**Figure 28**). This hydroxylated derivative of 6-hydroxymellein would be the last common intermediate between the two pathways. In *A. terreus*, this last common intermediate would be derivatized by additional oxidation events to produce terrein. In *C. uncialis*, we propose that the product is subsequently *O*-methylated and halogenated (**Figure 28**).

To determine whether lichen metabolites similar to the hypothesized end-product have been reported in the literature, we searched the *Dictionary of Natural Products* – an online global compendium of naturally-occurring molecules. Three similar molecules were observed (**Figure 28**, inset). Compounds A and B were isolated from *Lachnum papyraceum* (Stadler et al. 1995)

and compound C was isolated from *Graphis sp.* (Takenaka et al. 2011). These compounds have not been observed in *C. uncialis*. Although only two of the three total modifications are seen in each of these isolated metabolites (the three modifications being *O*-methylation, hydroxylation, and halogenation), the fact that molecules similar to our hypothesized end-product exist lends credence to our claim that our proposed end-product is a plausible molecule.

We refer to this approach as ‘homology mapping’. In Chapter 6, we will revisit this ‘homology mapping’ approach to propose functions for other gene clusters within *C. uncialis*.

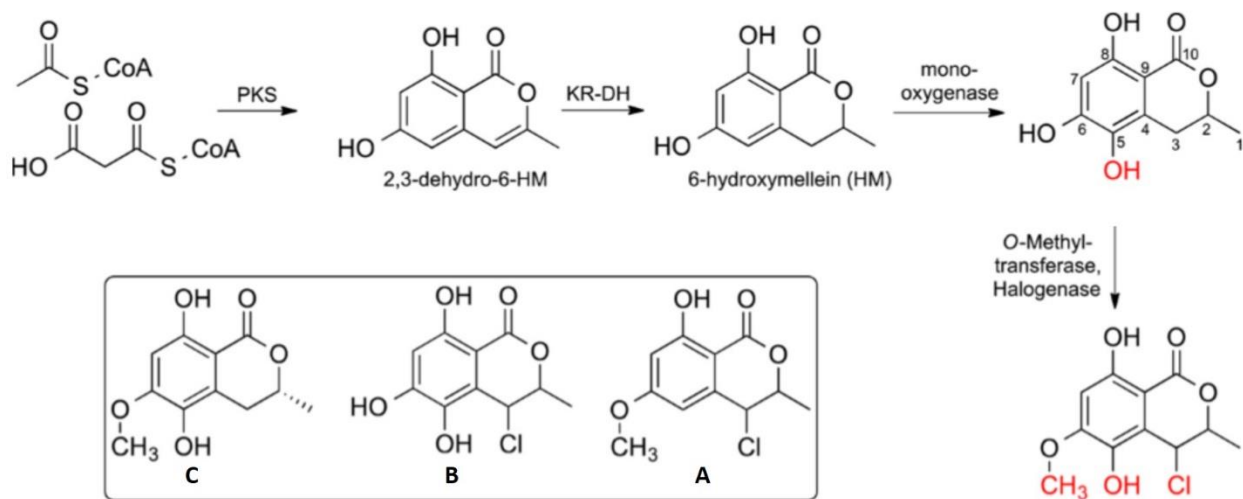


Figure 28: Proposed biosynthetic pathway in *C. uncialis*. The position of the moieties highlighted in red are variable. (Inset) Three molecules isolated from lichens that are similar in structure to the hypothesized end-product of the proposed biosynthetic pathway: (A) 4-chloro-6-methoxymellein (B) 4-chloro-5,6-dihydroxymellein (C) 5-hydroxy-6-methoxymellein. Reproduced from Abdel-Hameed et al. (2016b), in accordance with authors’ retainment of privileges.

4.3. Summary

A gene cluster encoding *trans*-PKS-like domains and a halogenase was observed within the *C. uncialis* genome. Genes within this cluster were observed to be genetically similar to genes within the terrein cluster of *A. terreus*. Phylogenetic analyses provide evidence of homology between three pairs of genes between *C. uncialis* and *A. terreus*. Based on the known biosynthetic functions of the cluster in *A. terreus*, we propose that the function of the three genes in *C. uncialis*

is to produce a hydroxylated derivative of 6-hydroxymellein. The remaining genes in the *C. uncialis* cluster suggests that this hypothesized intermediate is derivatized by *O*-methylation and halogenation. This proposition is supported by molecules isolated from lichens that possess structures similar to the hypothesized end-product.

Chapter 5

Gene Annotation of a Lichen Secondary Metabolome

5.1. Introduction - Annotating the lichen secondary metabolome

At the outset of my doctoral program, little information was available on the genetic programming of lichen secondary metabolite biosynthesis. A parallel objective of my project was to therefore expand the amount of genetic information available on the lichen secondary metabolome. As we needed to sequence the genome of *C. uncialis* to find the usnic acid gene cluster, this latter objective was congruent with the former. The assembled contigs were submitted to AntiSMASH (3.0) (Weber et al. 2015) and a draft compendium of biosynthetic gene clusters was generated (see Chapter 3). A total of 48 secondary metabolite biosynthetic gene clusters were identified. In this chapter, we will discuss these clusters. Within each figure of this chapter, a broad functional role of each gene is displayed. These proposals are based upon consensus similarity with genes deposited in GenBank (Altschul et al. 1990). All genes discussed in this chapter have been deposited in GenBank. The GenBank accession numbers and the similarity scores for each *C. uncialis* gene are provided in **Tables S5 to S52** in the Appendix. The accession numbers of the cluster-bearing contigs are provided in **Table S53** in the Appendix. For general information on polyketide synthases (PKS), non-ribosomal peptide synthetases (NRPS), hybrid PKS-NRPS, and terpene synthases, please refer to Chapter 1.

5.2. Type III PKS gene clusters

Type III PKS are distinct from type I and II PKS because they lack distinct catalytic domains. A single active site instead performs priming, extension, and cyclization reactions through iterative condensation cycles (Yu et al. 2012). As Type III PKS have only recently been discovered in *Fungi* (Seshime et al. 2005), relatively little is known about their function in *Fungi* as compared to type I PKS (Hashimoto et al. 2014). Experiments using degenerate primers have demonstrated that type III PKS genes are common among lichens (Muggia & Grube, 2010). Prior to our genome sequencing of *C. uncialis*, there were no reports of complete type III PKS gene sequences within the literature.

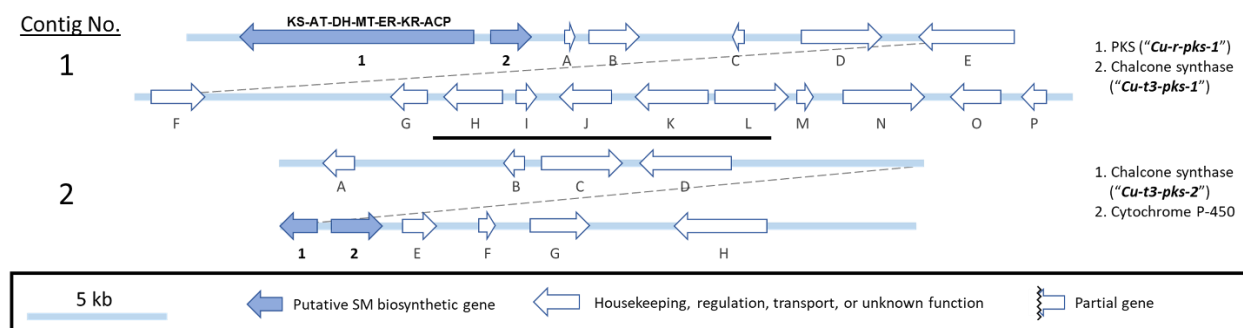


Figure 29: Type III polyketide synthase (PKS) gene clusters in the *C. uncialis* mycobiont genome. Abbreviations: Ketosynthase (KS), acyltransferase (AT), Dehydratase (DH), C-methyltransferase (MT), enoylreductase (ER), ketoreductase (KR), acetyl carrier protein (ACP). Reproduced from Bertrand et al. (2018a), in accordance with authors' retention of privileges.

Two complete type III PKS genes from *C. uncialis* were found, named *Cu-t3pks-1* and *cu-t3pks-2* (**Figure 29**). BLAST alignment of both type III PKS revealed consensus similarity to chalcone synthase genes deposited in GenBank. Chalcone synthases are a sub-group of type III PKS that use *p*-coumaroyl-CoA or cinnamoyl-CoA as starter units and typically carry out three condensations with malonyl-CoA (Austin & Noel, 2003). If functional, these PKS likely produce small aromatic polyketides. The *Cu-t3pks-1* gene is conspicuous for its proximity to a type I reducing PKS (named *Cu-r-pks-1*), perhaps suggesting CU-T3PKS-1 and CU-R-PKS-1 cooperate

in a single biosynthetic pathway. A cytochrome p450 is observed next to *Cu-t3-pks-2* suggesting an oxidative tailoring function (**Figure 29**).

5.3. Multi-module NRPS gene clusters

Non-ribosomal peptide synthetases (NRPS) operate independently of mRNA but assemble polypeptides through an assembly line of catalytic domains using proteinogenic and non-proteinogenic amino acids (Süssmuth & Mainz, 2017). These enzymes minimally require adenylation (A), condensation (C), and peptidyl carrier protein (PCP) domains, though numerous other functional domains are known. Two multi-modular NRPS have been previously reported in lichens (Wang et al. 2014).

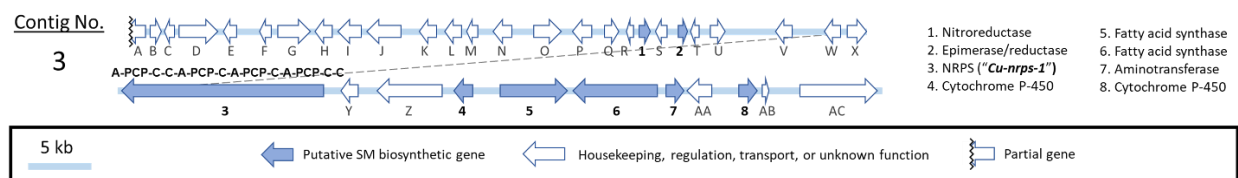


Figure 30: A multi-module non-ribosomal peptide synthetase (NRPS) gene cluster in the *C. uncialis* mycobiont genome. Abbreviations: Adenylation (A), peptidyl carrier protein (PCP), condensation (C). Reproduced from Bertrand et al. (2018a), in accordance with authors' retainment of privileges.

One multi-module NRPS from *C. uncialis* was found, named *Cu-nrps-1* (**Figure 30**). This gene contains four [A-PCP-C] modules and a terminal condensation (C) domain. No terminal thioesterase was observed. The absence of a TE domain is an alternative architectural style that is common among fungal NRPS when macrocyclization is used to release the peptide chain (Gao et al. 2012. Note that some TE-possessing NRPS may also release cyclized products (Trauger et al. 2000; Kohli et al. 2001). The presence of four modules and a terminal condensation domain therefore suggests that this NRPS may produce a cyclized tetrapeptide. AntiSMASH contains an algorithm for predicting the amino acids that are incorporated into a peptide by NRPS. This algorithm predicted that the peptide produced by CU-NRPS-1 is composed of leucine, isovaline,

homoalanine, and glutamine. These predictions await experimental confirmation. This NRPS was located with two cytochrome p450 oxidases, an aminotransferase, and two fatty acid synthases (FAS). The presence of two FAS is notable because FAS can participate in secondary metabolite biosynthesis in cooperation with PKS and NRPS by producing highly-reduced acyl products (Masschelein et al. 2013). A dedicated FAS producing hexanoic acid for aflatoxin biosynthesis is a well-known example in *Fungi* (Hitchman et al. 2001). The architecture of the NRPS and the types of accessory tailoring genes lead us to hypothesize that the cluster is responsible for producing a macrocyclized tetrapeptide attached to a fatty acid.

5.4. Single-module NRPS gene clusters

Eight single-module NRPS genes were identified, named *Cu-nrps-2* to *Cu-nrps-9* (**Figure 31**). These NRPS are conspicuous because they lack condensation ('C') domains. Single module NRPS lacking condensation domains are known to appear in fungi but are poorly understood (Pel et al. 2007). Two examples in bacteria include IndC, responsible for the conversion of glutamine to the blue pigment indigoidine in *Photorhabdus luminescens* (Brachmann et al. 2012), and a ϵ -poly-L-lysine synthetase, responsible for the iterative polymerization of L-lysine in *Streptomyces albulus* (Yamanaka et al. 2008). Six of the eight single-module NRPS genes found in *C. uncialis* appear to be associated with post-synthetic tailoring genes. We speculate that the encoded NRPS produce small peptides that are then altered by one or more distinct chemical steps by the accessory gene products. AntiSMASH V. 4.0 (Blin et al. 2017) was unable to predict the amino acids that are incorporated.

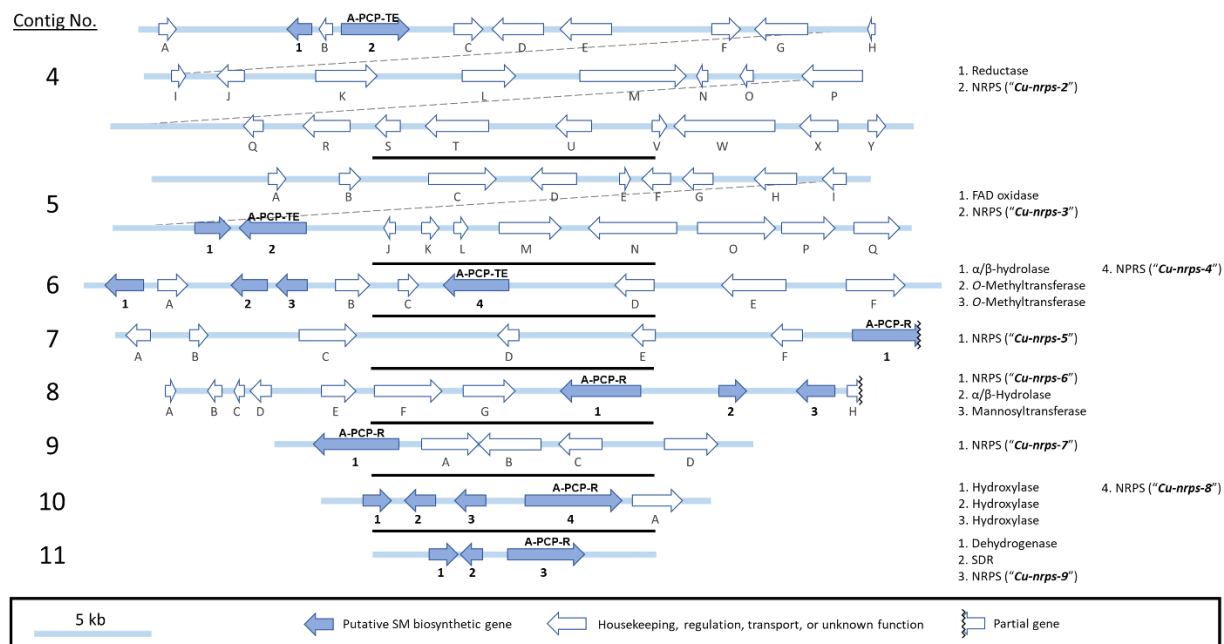


Figure 31: Single-module non-ribosomal peptide synthetase (NRPS) gene clusters in the *C. uncialis* mycobiont genome. Abbreviations: Adenylation (A), peptidyl carrier protein (PCP), thioesterase (TE), reductase (R). Reproduced from Bertrand et al. (2018a), in accordance with authors' retainment of privileges.

5.5. Hybrid PKS-NRPS gene clusters

Hybrid PKS-NRPS combine the biosynthetic capabilities of PKS and NRPS to produce structurally complex natural products (Miyanaga et al. 2018; Fisch, 2013). Two PKS-NRPS genes have been previously reported in lichenizing fungi (Wang et al. 2014). Three PKS-NRPS genes were observed in *C. uncialis*, named *Cu-pks-nrps-1* to *Cu-pks-nrps-3*, complete with information on the 'tailoring' genes associated with these PKS-NRPS genes (**Figure 32**). The PKS component of all hybrid PKS-NRPS contain KR and DH reducing domains, suggesting that the polyketide is reduced prior to peptidylation. The presence of *cis*-ER with *Cu-pks-nrps-2* and *trans*-ER with *Cu-pks-nrps-3* also suggest that the polyacyl moieties are highly reduced. A *trans*-TE was observed near *Cu-pks-nrps-2* whereas terminal reductase (R) domains are present in the other two genes. Regrettably, AntiSMASH V. 4.0 (Blin et al. 2017) was unable to provide confident predictions of

amino acid incorporations among these PKS-NRPS. Two PKS-NRPS genes, *Cu-pks-nrps-1* and *Cu-pks-nrps-2*, are noteworthy for their proximity to three terpene synthase genes, herein named *Cu-terp-1* to *Cu-terp-3*. It is possible that these terpene synthases cooperate with the PKS-NRPS to produce terpene-polyketide-peptide triple chimeric natural products (**Figure 32**).

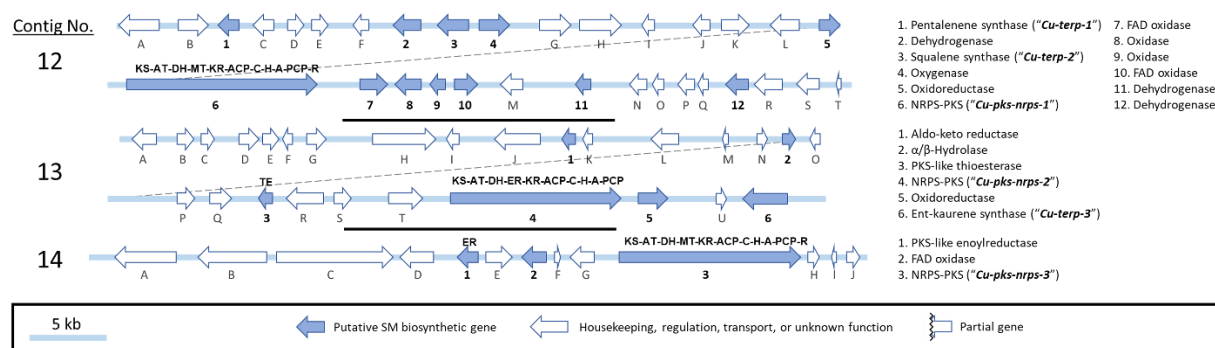


Figure 32: Non-ribosomal peptide synthetase – polyketide synthase (PKS-NRPS) gene clusters identified in the *C. uncialis* mycobiont genome. Abbreviations: Ketosynthase (KS), acyltransferase (AT), C-methyltransferase (MT), ketoreductase (KR), dehydratase (DH), enoylreductase (ER), acetyl carrier protein (ACP), thioesterase (TE), reductase (R), condensation (C), HxxPF domain (H), adenylation (A), peptidyl carrier protein (PCP). Reproduced from Bertrand et al. (2018a), in accordance with authors' retainment of privileges.

5.6. Terpene synthase gene clusters

Terpenes are chains of activated $(C_5H_8)_n$ isoprene units and are a diverse class of natural product with both structural and functional roles in nature (Oldfield & Lin, 2012; Gonzalez-Burgos & Gomez-Serranillos, 2012). Nine terpene synthase genes were observed in *C. uncialis*, named *Cu-terp-1* to *Cu-terp-9*. Three of these genes were found with PKS-NRPS genes and are displayed in **Figure 32**. The remaining six genes are shown in **Figure 33**. One of these terpene synthases, *Cu-terp-9*, was found next to a reducing type I PKS named *Cu-r-pks-2* (**Figure 33**).

Putative functions of terpene synthase gene products can be proposed based on consensus similarity to characterized genes deposited in GenBank. The gene product of *Cu-terp-1* is proposed to be pentalenene synthase, a sesquiterpene synthase that cyclizes farnesyl diphosphate into pentalenene (Seemann et al. 2002). The gene product of *Cu-terp-2* was identified as a phytoene

synthase. Phytoene is a 40-carbon tetraterpene formed from two molecules of geranylgeranyl pyrophosphate (Linnemannstons et al. 2002). Phytoene is the intermediate to the carotenoid class of natural products, which are highly conjugated terpenoids typically acting as pigments and photo-protectants. The identification of what appears to be a phytoene dehydrogenase proximal to phytoene synthase suggests that these two genes comprise a gene cluster responsible for the biosynthesis of carotenoids or carotenoid precursors. The gene product of *Cu-terp-3* is proposed to be an ent-kaurene synthase, an enzyme that converts copalyl diphosphate to the tetracyclic diterpene *ent*-kaurene, an intermediate in the biosynthesis of gibberellins (Kawaide et al. 1997; Hedden et al. 2002). Two genes, *Cu-terp-4* and *Cu-terp-8*, are suggested to be squalene-hopene cyclases. This enzyme converts squalene into the pentacyclic triterpenoid hopene (Hoshino & Sato, 2002; Wendt et al. 1997). Consensus similarity identified *Cu-terp-5* as a possible dimethylallyl tryptophan synthase. This enzyme binds dimethylallyl pyrophosphate and L-tryptophan to form dimethylallyltryptophan, an intermediate in ergot biosynthesis (Metzger et al. 2009). *Cu-terp-6* is proposed to be a farnesyl-diphosphate farnesyl transferase (squalene synthase). This enzyme dimerizes two units of farnesylpyrophosphate into squalene, a 30-carbon triterpene that is an intermediate in sterol biosynthesis (Tansey & Shechter, 2000). Two aristolochene synthases, *Cu-terp-7* and *Cu-terp-9*, were also suggested to be encoded in the *C. uncialis* genome. Aristolochene is a bicyclic sesquiterpene that is derivatized to form a range of mycotoxins, for example, PR toxin from the cheese mold *Penicillium roqueforti* (Proctor & Hohn, 1993; Chen et al. 1982).

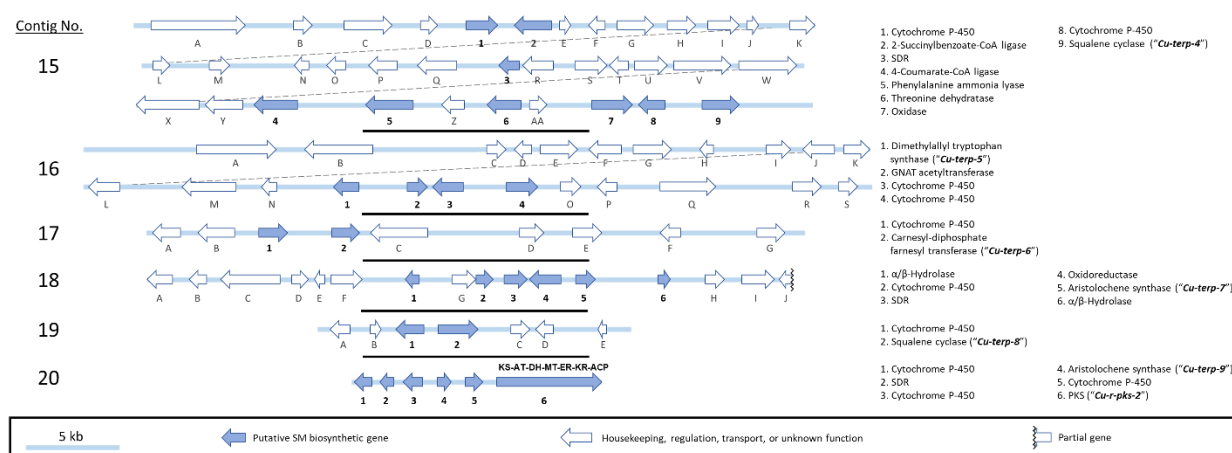


Figure 33: Terpene synthase gene clusters in the *C. uncialis* mycobiont genome. Abbreviations: Ketosynthase (KS), acyltransferase (AT), dehydratase (DH), C-methyltransferase (MT), enoylreductase (ER), ketoreductase (KR), acetyl carrier protein (ACP), Gcn5-related n-acetyltransferase (GNAT). Reproduced from Bertrand et al. (2018a), in accordance with authors' retention of privileges.

5.7. Type I PKS gene clusters

Type I PKS are multi-catalytic domain proteins that assemble polyketides through decarboxylative Claisen condensations of acyl units (Staunton & Weissman, 2001). Degenerate primer and genome sequencing studies have previously shown that PKS genes are ubiquitous among lichenizing fungi and that dozens of PKS genes could be present in a single lichen genome (Wang et al. 2014; Schmitt et al. 2008; Junttila & Rudd, 2012; Kampa et al. 2013). However, information on complete PKS genes as well as their accompanying 'accessory' genes is lacking. A total of 14 non-reducing and 18 reducing PKS were observed in *C. uncialis*. These are encoded as *Cu-nr-pks-1* to *Cu-nr-pks-14* (**Figure 34**) and *Cu-r-pks-1* to *Cu-r-pks-18* (**Figure 35**), respectively. A few of these PKS genes are displayed in alternative figures within this chapter because they were located near other classes of biosynthetic genes. The cluster containing *Cu-nr-pks-2* is the usnic acid gene cluster, and the cluster containing *Cu-nr-pks-8* is the 6-hydroxymellein gene cluster (**Figure 34**).

Some PKS use domains that are encoded in separate genes. These are known as *trans*-PKS (Helfrich & Piel, 2016). Two *trans*-PKS were observed: An ACP-R-like gene near *Cu-nr-pks-10* and the KR-DH-like gene near *Cu-nr-pks-8* (**Figure 34**). In Chapter 4, we proposed that the function of the KR-DH-like protein was to provide a two-electron reduction of 2,3-dehydro-6-hydroxymellein into 6-hydroxymellein. The function of the ACP-R gene is unclear. It is possible that this protein is responsible for cleaving a product from the ACP domain of its target PKS to produce an aldehyde-containing polyketide. However, the associated PKS already possesses a terminal TE domain, necessitating that the TE domain be inactive for this hypothesis to be viable.

Gcn5-related *N*-acetyltransferases (GNAT) use acetyl-CoA to acetylate substrates including secondary metabolites. This enzyme may also prime ACP domains with acetyl-CoA in lieu of a SAT domain or loading module (Vetting et al. 2005; Gu et al. 2007). Consensus similarity with genes deposited in GenBank suggests that a GNAT is encoded near *Cu-nr-pks-12* (**Figure 34**). A second putative GNAT was encoded next to *Cu-terp-5* (**Figure 33**). Experimental demonstration of GNAT activity would confirm a role for the GNAT superfamily in lichen natural product biosynthesis.

Half of the 14 non-reducing PKS genes were observed to possess two ACP domains (**Figure 34**). This is not an unusual architectural feature of PKS, as double ACP domains have been observed in several well-studied PKS from *Fungi* (Fujii et al. 2000; Yu & Leonard, 1995; Huang et al. 2001). Inactivation studies have demonstrated that double ACP domains operate in parallel to increase the catalytic turnover of the PKS (Gu et al. 2011; Jiang et al. 2008; Rahman et al. 2005). How it is possible for two or more ACP domains to increase the catalytic turnover of a PKS or FAS remains unclear. Nonetheless, we speculate that the role of double ACP domains in these lichen PKS is to increase the catalytic turnover of the encoded PKS.

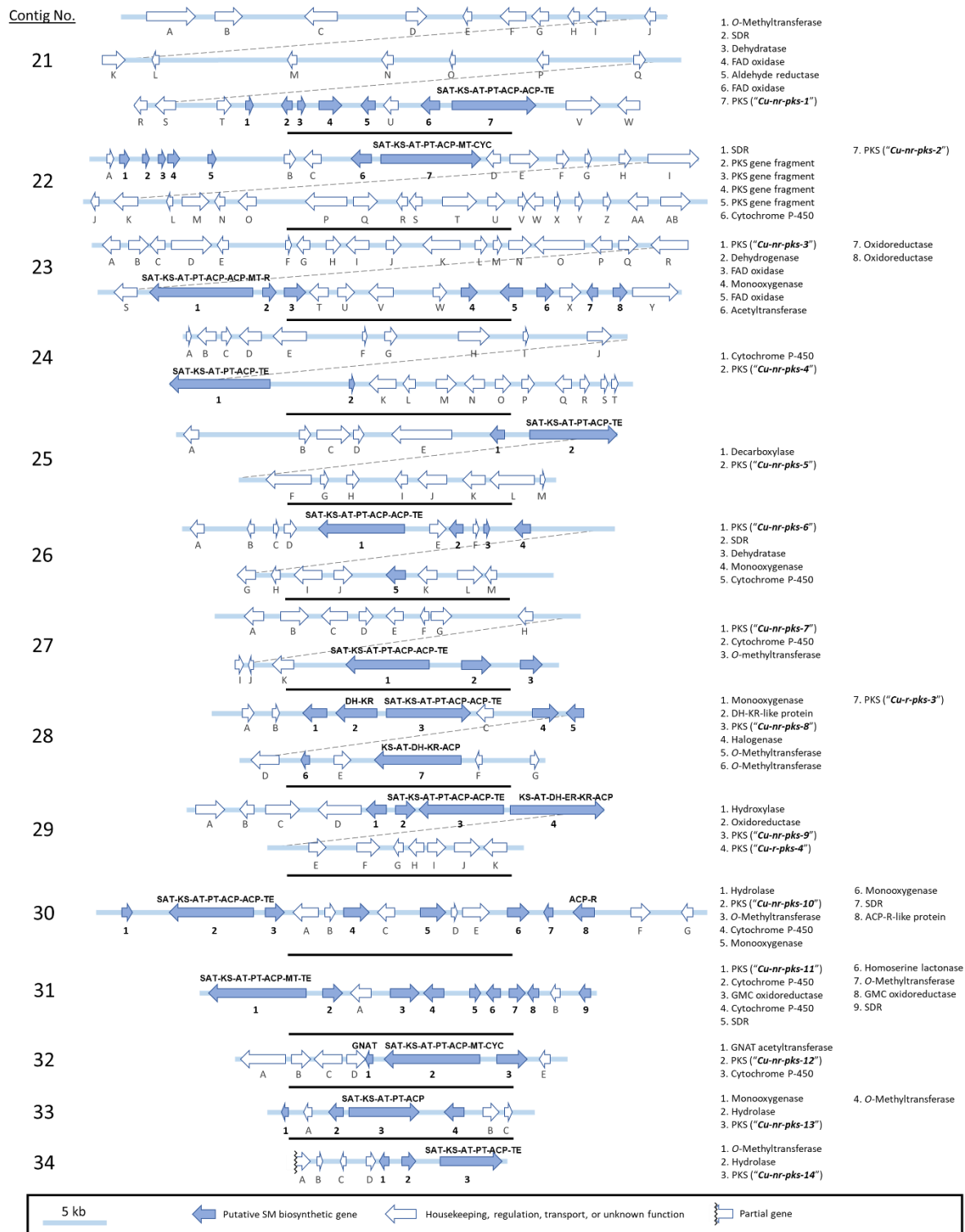


Figure 34: Type I non-reducing polyketide synthase (PKS) gene clusters in the *C. uncialis* mycobiont genome. Abbreviations: Starter acyltransferase (SAT), ketosynthase (KS), acyltransferase (AT), product template domain (PT), C-methyltransferase (MT), acetyl carrier protein (ACP), ketoreductase (KR), dehydratase (DH), enoylreductase (ER), Claisen cyclase (CYC), thioesterase (TE), reductase (R), Gcn5-related N-acetyltransferase (GNAT). Reproduced from Bertrand et al. (2018a), in accordance with authors' retainment of privileges.

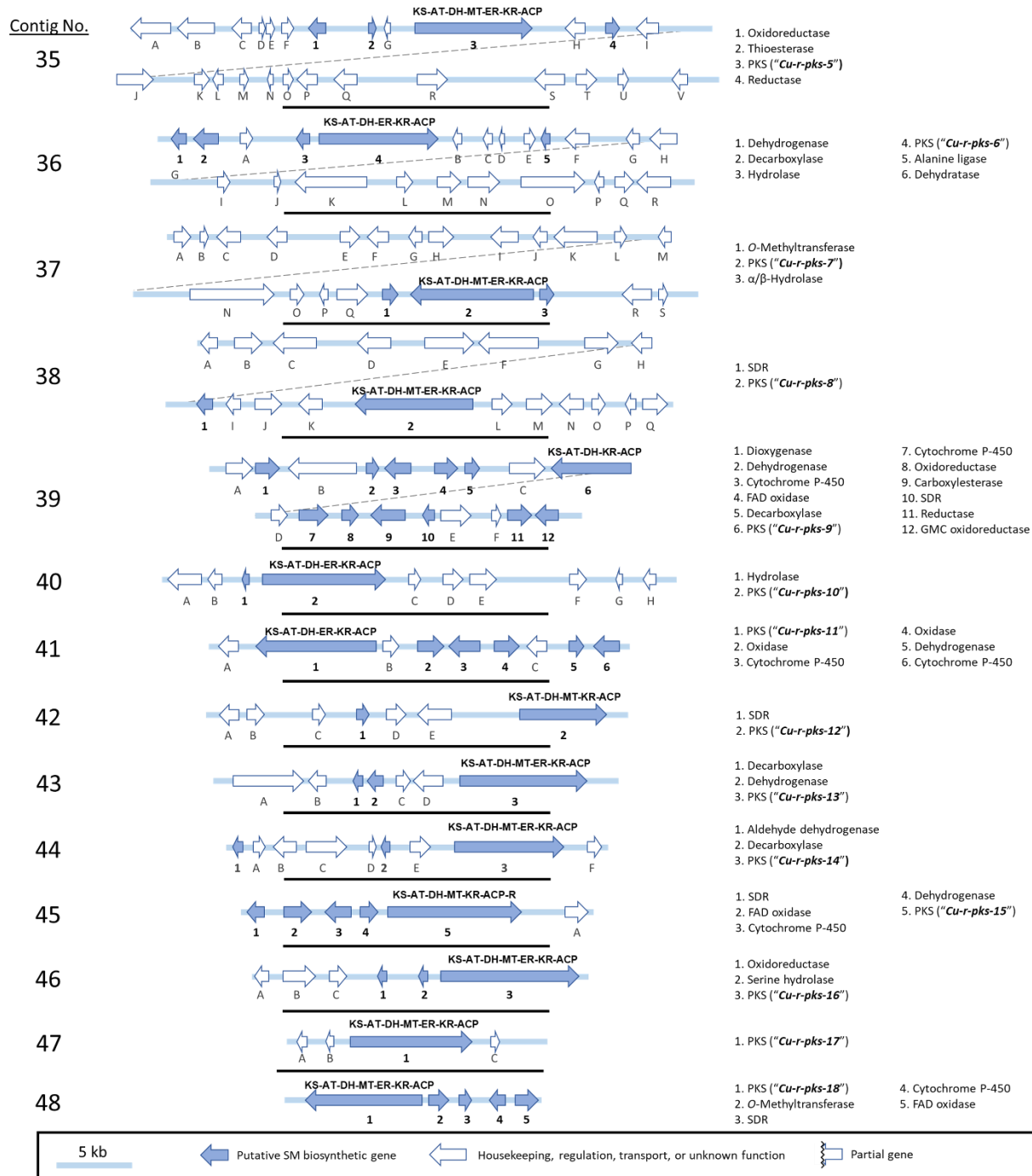


Figure 35: Type I reducing polyketide synthase (PKS) gene clusters in the *C. uncialis* mycobiont genome. Ketosynthase (KS), acyltransferase (AT), dehydratase (DH), C-methyltransferase (MT), enoylreductase (ER), ketoreductase (KR), acetyl carrier protein (ACP), reductase (R). Reproduced from Bertrand et al. (2018a), in accordance with authors' retention of privileges.

Fungal reducing PKS lack SAT, PT, and terminal domains (Simpson & Cox, 2012; Chiang et al. 2010). With the exception of *Cu-r-pks-15* (**Figure 35**), this architecture is consistent with reducing PKS genes observed in this study and other documented examples of reducing PKS genes from lichens (Gagunashvili et al. 2009; Kim et al. 2012; Wang et al. 2011; Wang et al. 2014). A terminal reductase domain found in *Cu-r-pks-15* instead suggests that the encoded PKS requires reductive termination. The domain architecture of *Cu-nr-pks-13* is notable for lacking a terminal domain (**Figure 34**). In this case, it is possible that a reducing enzyme encoded near this PKS is responsible for terminating polyketide assembly.

5.8. Interpretation of findings and importance of contributions

It may have become apparent to the reader that some of the gene clusters reported here appear incomplete. In some cases, such as *Cu-t3-pks-2* (**Figure 29**), the gene cluster is flanked by long regions of DNA that do not contain genes related to secondary metabolism. This suggests that a complete gene cluster was captured by the DNA assembly. In other cases, such as *Cu-r-pks-18* (**Figure 35**), secondary metabolite genes appear at the edges of the contig. It is therefore possible that other biosynthetic genes associated with this cluster are present but were not captured within the contig. For this reason, I illustrated the entire cluster-bearing contig so that the reader may discern the relative completeness of each of the proposed gene clusters. It is also possible that some of these shorter contigs are fragments of a much larger gene cluster. If this is the case, then the true number of biosynthetic gene clusters in *C. uncialis* is less than 48. To determine which genes within a speculative biosynthetic cluster are associated with a known metabolite, one commonly used technique is to induce the production of the metabolite, then observe whether there is an increase in the transcription of the studied genes (E.g., see Zaehle et al. 2014). This technique can also be used to rule-out the involvement of ‘housekeeping’ genes that are located near or

sandwiched between secondary metabolism genes. To apply this technique to the 48 gene clusters of *C. uncialis* is well-beyond the intended scope of my doctoral program.

An important discovery borne by this work is the incredible biosynthetic potential that is evidently hidden away within the lichen genome. Whereas the cluster of *Cu-nr-pks-2* is speculated to produce usnic acid (Chapter 3) and the cluster of *Cu-nr-pks-8* is associated with 6-hydroxymellein biosynthesis (Chapter 4), this work nonetheless leaves 46 gene clusters without a description of their potential biosynthetic functions. In this regard, this work parallels genome sequencing studies in bacteria and fungi such that these latter works have similarly revealed numerous and unaccountable biosynthetic gene clusters hidden away within an organism's DNA (Nielsen et al. 2017; Schorn et al. 2016; Guo & Wang, 2014; Aigle et al. 2014). In the case of *C. uncialis*, what do these other 46 gene clusters do? In the upcoming chapter, we will use the 'homology mapping' approach (Chapter 4) to attempt to answer this question for some of the gene clusters presented. Functional heterologous expression of these cryptic gene clusters will ultimately be required to gain a definitive understanding of what molecules the encoded enzymes produce. This development of a reliable method to express lichen genes in heterologous hosts will be required to bridge this gap between gene surveying and metabolite discovery. Our efforts to establish a protocol will be detailed in Chapters 7 and 8.

5.9. Summary

The *C. uncialis* was sub-cultured from its photosynthetic partner and its DNA was isolated (Chapter 3). The genomic DNA was *de novo* sequenced and assembled into contigs. Annotation of the secondary metabolite genes revealed 48 biosynthetic gene clusters – the first such instance of attempting to capture a complete genetic secondary metabolome within a species of lichenizing fungi. This analysis provided us with hitherto unparalleled insight to the secondary metabolism

of lichens. This is the first report of gene clusters within a species of lichen encoding type III PKS, NRPS, hybrid PKS-NRPS, and terpene synthases. The number of gene clusters uncovered by this work suggests a vast biosynthetic potential. The development of a reliable method to express lichen biosynthetic genes within heterologous hosts will be required to explore this vast biosynthetic potential.

Chapter 6

Assignment of Biosynthetic Pathways to Lichen Gene Clusters

6.1. Introduction – Applying a ‘homology mapping’ approach to the secondary metabolome

In Chapter 3, a deductive approach was used to identify *Cu-nr-pks-2* as the PKS gene associated with usnic acid biosynthesis in *C. uncialis*. In Chapter 4, I described a ‘homology mapping’ approach that was used to associate a hypothesized halogenated derivative of 6-hydroxymellein with the gene cluster of *Cu-nr-pks-8*. In Chapter 5, I detailed a total of 48 secondary metabolite gene clusters that were discovered as a result of our gene surveying in *C. uncialis*. The *Cu-nr-pks-2* and *Cu-nr-pks-8* clusters aside, this leaves 46 clusters without predictions of biosynthetic functions. We therefore decided to apply our ‘homology mapping’ approach to determine probable functions for these remaining 46 clusters. The result of this work, described in this chapter, is the assignment of eight additional biosynthetic pathways.

The procedure for identifying a plausible biosynthetic function via the ‘homology mapping’ approach is summarized by addressing the following three questions:

1. Are the genes similar to those of a known biosynthetic gene cluster?
2. Do the identified genes form a coherent biosynthetic pathway?
3. Is genetic homology supported by phylogenetic analysis?

Predictions that contain many paired genes with a high degree of evolutionary homology and form a coherent series of chemical steps comprise a more useful prediction of function as

compared to the opposite. Therefore, the value of each assessment can be weighted by these two additional questions:

4. For items 1-3, above, for how many genes is this the case?
5. Are there metabolites reported in the literature that are similar to the predicted product?

The eight pathways that we have proposed below are described below, in order of highest degree of confidence to the most speculative. Note that it is not obligatory for all genes in a pathway to be assigned with a specific biosynthetic function. As shown in the example of terrein (Chapter 4), we assigned specific functions to only three of the five genes of the *Cu-nr-pks-8* gene cluster. In this case, we propose that those three encoded enzymes produce a hydroxylated derivative of 6-hydroxymellein. Whereas in *A. terreus* this is further derivatized into terrein, in *C. uncialis* we proposed that the end-product is halogenated and *O*-methylated at speculative positions on the molecule. The strength of this prediction rests in the genetic similarity of these three pairs of genes, their genetic homology as support by phylogenetics, and their ability to form a coherent chemical pathway. From our work on 6-hydroxymellin, and from some of the examples shown below, it is evident that evolution has shuffled components of ancestral gene clusters to create new biosynthetic pathways and new molecules.

Throughout this work, phylogenetic analyses will be conducted to ascertain the degree of evolutionary homology between the *C. uncialis* and identified reference genes (see Chapter 2). Complete phylogenetic trees, including trees that do not support a homologous relationship, are shown as **Figures S6 to S48** in the Appendix.

6.2. A proposed complete grayanic acid pathway

Grayanic acid is a polyketide metabolite produced exclusively by lichenizing fungi. Reproducible induction of grayanic acid biosynthesis combined with transcriptional profiling of candidate polyketide synthase genes led to the putative identification of the grayanic acid gene cluster in *C. grayi* (Armaleo et al. 2011). This PKS gene was found next to genes encoding a cytochrome p450 oxidase and an *O*-methyltransferase, two genes that are consistent with the expected biosynthetic pathway of grayanic acid (Armaleo et al. 2011). The grayanic acid gene cluster and proposed biosynthetic pathway are shown in **Figure 36** and **Figure 38**.

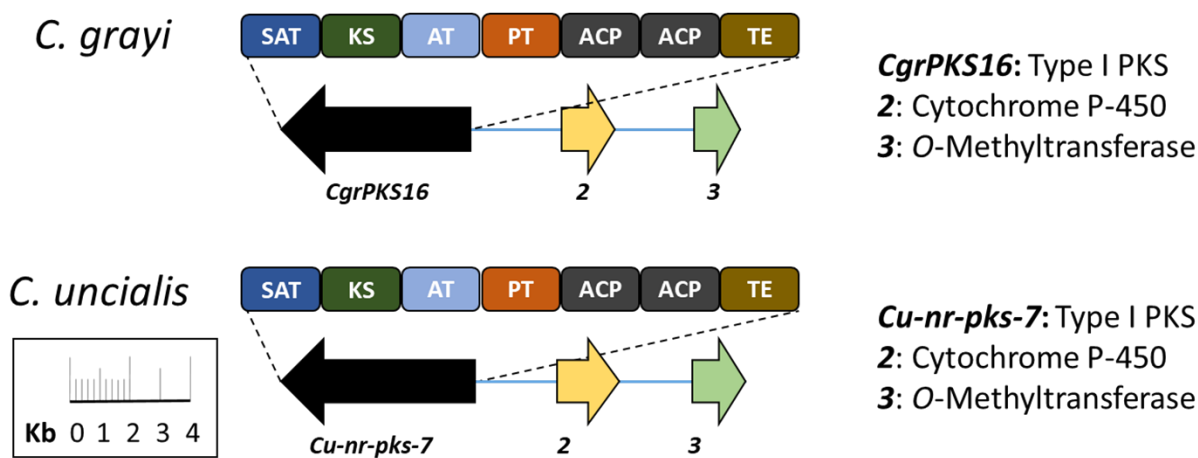


Figure 36: The *C. grayi* grayanic acid gene cluster (Accession no. GU930713) and *C. uncialis* uncharacterized gene cluster (Accession no. MG777495). Colour coding indicates homology pairing. Domain abbreviations: Starter acyltransferase (SAT), ketosynthase (KS), acyltransferase (AT), product template domain (PT), acetyl carrier protein (ACP), thioesterase (TE). Reproduced from Bertrand et al. (2018b), in accordance with authors' retainment of privileges.

Table 3: BLAST statistics and proposed functions of *C. uncialis* genes that are genetically similar to genes within the grayanic acid cluster of *C. grayi*. Reproduced from Bertrand et al. (2018b), in accordance with authors' retainment of privileges.

<i>C. uncialis</i> gene	Accession no.	<i>C. grayi</i> gene	Accession no.	Proposed function of <i>C. uncialis</i> gene	Similarity (%) / Coverage (%)
<i>Nr-pks-7</i>	AUW31177	<i>CgrPKS16</i>	ADM79459	4- <i>O</i> -Demethyl-sphaerophorin synthase	91/97
2	AUW31178	2	ADM79460	4- <i>O</i> -Demethyl-sphaerophorin oxidase	86/88
3	AUW31179	3	ADM79461	4- <i>O</i> -Demethylgrayanic acid <i>O</i> -methyltransferase	74/99

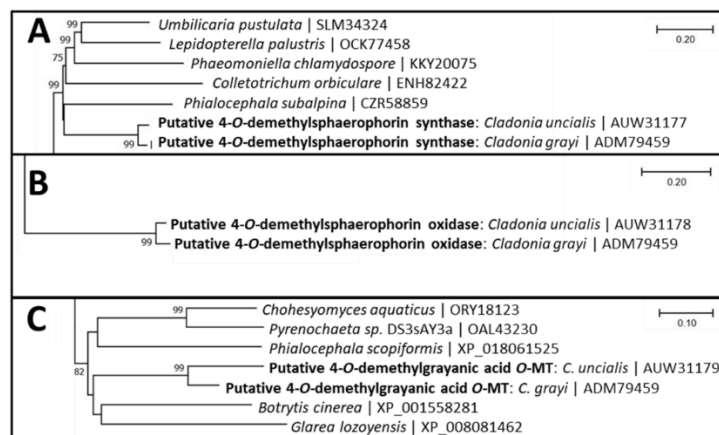


Figure 37: Truncated phylogenetic trees illustrating degree of relationship between genes of *C. uncialis* and *C. grayi* pertinent to grayanic acid biosynthesis: (A) 4-*O*-demethylsphaerophorin synthase; (B) 4-*O*-demethylsphaerophorin oxidase; (C) 4-*O*-demethylgrayanic acid *O*-methyltransferase. Bar signifies number of amino acid substitutions per site. Complete phylogenetic trees are provided as **Figures S6-S8** in the Appendix. Reproduced from Bertrand et al. (2018b), in accordance with authors' retainment of privileges.

One gene cluster in *C. uncialis* was identified as a possible functional homologue of the *C. grayi* grayanic acid cluster (**Figure 36**). A pBLAST search on each of the three genes within the *C. uncialis* cluster revealed robust similarity scores for each of the three genes in the grayanic acid gene cluster of *C. grayi* (**Table 3**). These gene-pairs are illustrated by colour-coding (**Figure 36**). The domain architecture of the two PKS genes are identical (**Figure 36**). Clustering of gene-pairs in the resultant phylogenetic trees infers a recent and common evolutionary homology (**Figure 37**). This is perhaps unsurprising because *C. uncialis* and *C. grayi* are members of the *Cladonia* genus. These findings are interpreted as evidence that a grayanic acid gene cluster is encoded in *C. uncialis*. A biosynthetic pathway identical to that of *C. grayi* was proposed (**Figure 38**). The *C. uncialis* PKS named CU-NR-PKS-7 is proposed to synthesize 4-*O*-demethylsphaerophorin. The *C. uncialis* cytochrome P-450 and *O*-methyltransferase are then proposed to oxidize and methylate this intermediate, producing 4-*O*-demethylgrayanic acid and grayanic acid, respectively. Grayanic acid is not known to be produced by *C. uncialis*.

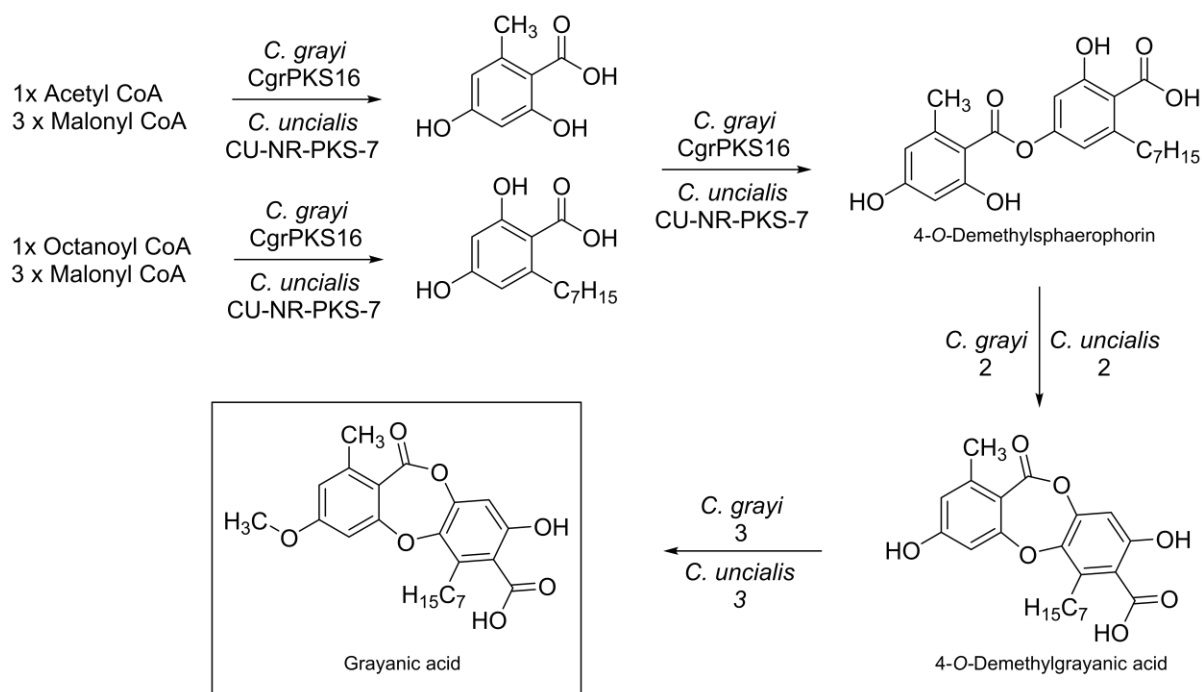


Figure 38: Experimentally supported pathway for grayanic acid biosynthesis in *C. grayi*, and juxtaposed, the biosynthetic pathway that is proposed to be encoded in *C. uncialis*. Reproduced from Bertrand et al. (2018b), in accordance with authors' retainment of privileges.

6.3. A proposed complete betaenone pathway

Betaenones A-C are phytotoxic polyketides produced by the fungus *Phoma betae*. Heterologous expression experiments demonstrated that the PKS (*Bet1*) and a *trans*-enoylreductase (*Bet3*) produce dehydroprobetaenone I in *P. betae* (Ugai et al. 2015). The gene product of *Bet4* reduces dehydroprobetaenone I to probetaenone I. The gene product of *Bet2* is a biofunctional oxidase that may convert probetaenone I to betaenone B, and dehydroprobetaenone I to betaenone C. The function of *Bet5* could not be conclusively determined by Ugai et al. (2015) but is hypothesized to convert betaenone C to betaenone A through an aldol reaction. The gene cluster and biosynthetic pathway of betaenone A-C are shown in **Figure 39** and **Figure 41**.

The KnownClusterBLAST module embedded within AntiSMASH (V. 4.0) (Blin et al. 2017) identified a *C. uncialis* cluster with genetic similarities to the betaenone cluster of *P. betae*.

A pBLAST analysis demonstrated that each of the five genes in *P. betae* cluster are genetically similar to five genes in *C. uncialis*. Similarity scores and gene-pairings are shown in **Table 4** and **Figure 39**. The domain architecture of the lichen PKS is identical to *Bet1*, including the presence of a *trans*-enoylreductase (**Figure 39**). Phylogenetic analyses of each of these gene pairs support close genetic homology (**Figure 40**). These data suggest that the *C. uncialis* gene cluster is a complete betaenone gene cluster. Functional roles of the each of the five genes can be proposed from the homology pairings. In this case, an identical biosynthetic pathway is proposed (**Figure 41**). Lichens are not known to produce betaenones A-C.

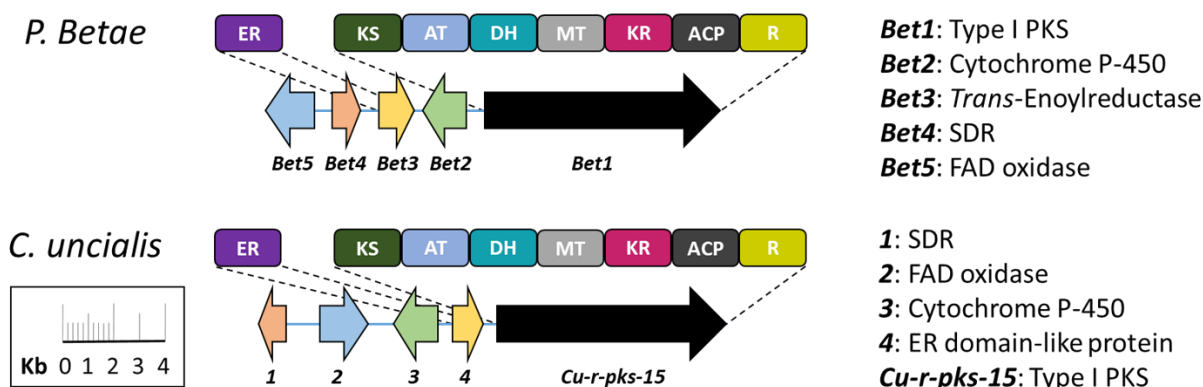


Figure 39: The *P. betae* betaenone gene cluster (Accession no. LC011911) and *C. uncialis* uncharacterized gene cluster (Accession no. MG777513). Genes with matching colours are genetically similar. *Domain abbreviations:* Enoylreductase (ER), ketosynthase (KS), acyltransferase (AT), dehydratase (DH), C-methyltransferase (MT), ketoreductase (KR), acetyl carrier protein (ACP), reductase (R). *Gene abbreviations:* Polyketide synthase (PKS), short-chain dehydrogenase/reductase (SDR), flavin adenine dinucleotide-dependent oxidase (FAD oxidase). Reproduced from Bertrand et al. (2018b), in accordance with authors' retainment of privileges.

Table 4: BLAST statistics and proposed functions of the *C. uncialis* genes that are genetically similar to genes of the betaenone pathway of *P. betae*. Reproduced from Bertrand et al. (2018b), in accordance with authors' retainment of privileges.

<i>C. uncialis</i> gene	Accession no.	<i>P. betae</i> gene	Accession no.	Proposed function of <i>C. uncialis</i> gene	Similarity (%) / Coverage (%)
1	AUW31407	<i>Bet4</i>	BAQ25463	Dehydroprobetaenone reductase	58/98
2	AUW31408	<i>Bet5</i>	BAQ25462	Putative betaenone C oxidase	56/98
3	AUW31409	<i>Bet2</i>	BAQ25465	Dehydroprobetaenone oxidase	60/96
4	AUW31410	<i>Bet3</i>	BAQ25464	<i>trans</i> -Enoylreductase	64/98
<i>R-pks-15</i>	AUW31411	<i>Bet1</i>	BAQ25466	Dehydroprobetaenone synthase	61/99

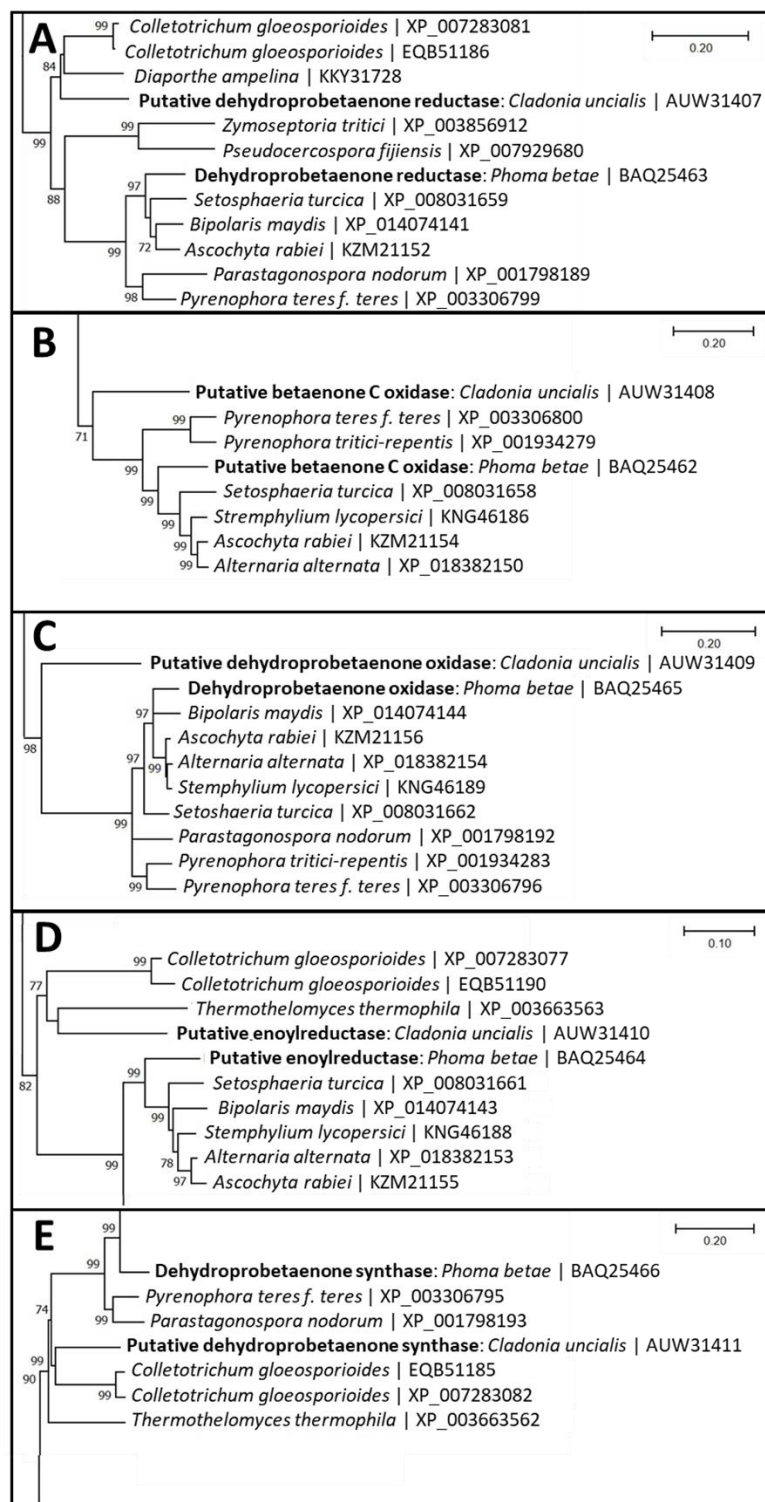


Figure 40: Truncated phylogenetic trees illustrating degree of relationship between genes of *C. uncialis* and *P. betae* pertinent to betaenone biosynthesis: (A) Dehydroprobetaenone reductase; (B) Betaenone C oxidase; (C) Dehydroprobetaenone oxidase; (D) Enoylreductase; (E) Dehydroprobetaenone synthase. Bar signifies number of amino acid substitutions per site. Complete phylogenetic trees are provided as **Figures S9-S13** in the Appendix Reproduced from Bertrand et al. (2018b), in accordance with authors' retainment of privileges.

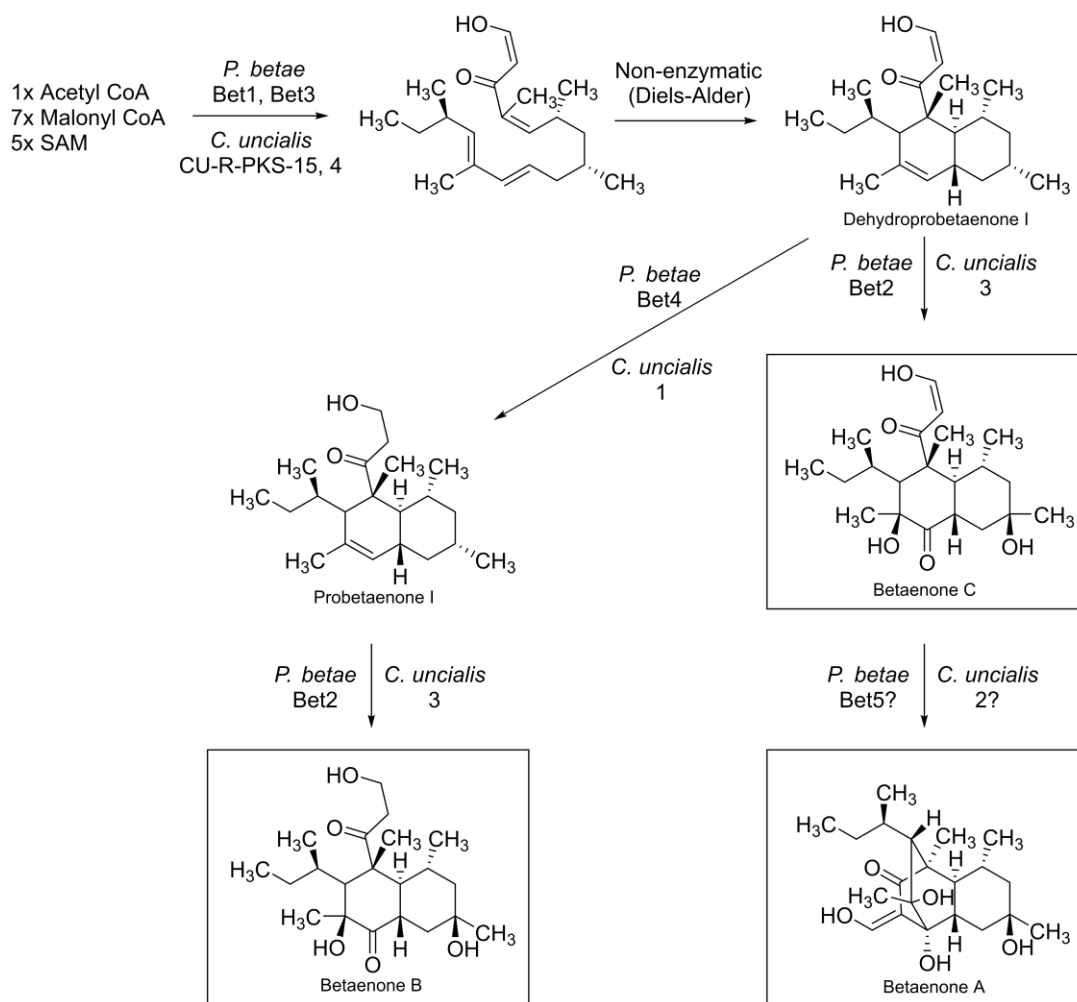


Figure 41: The experimentally supported pathway for betaenone A-C biosynthesis in *P. betae*, and juxtaposed, the biosynthetic pathway that is proposed to be encoded in *C. uncialis*. Reproduced from Bertrand et al. (2018b), in accordance with authors' retention of privileges.

6.4. A proposed complete patulin pathway

Patulin is a mycotoxin subject to strict regulation in the food industry because of its acute toxicity. The study of patulin biosynthesis was formative in the Birch acetate hypothesis and our understanding of polyketide biosynthesis in *Fungi* (Birch et al. 1955). Patulin biosynthesis requires 10 chemical steps facilitated by 15 gene products (Puel et al. 2010; Tannous et al. 2014). The function of five of these genes have been experimentally characterized. These include *PatK*, encoding 6-methylsalicylic acid synthase, *PatG*, encoding 6-methylsalicylic acid decarboxylase, *PatH* encoding *m*-cresol hydroxylase, *PatI*, encoding *m*-hydroxybenzyl alcohol hydroxylase, and

PatN, encoding isoeopoxydon dehydrogenase (Puel et al. 2007; Snini et al. 2014; Artigot et al. 2009; Dombrink-Kurtzman & Engberg, 2006). The patulin gene cluster and biosynthetic pathway are displayed in **Figure 42** and **Figure 44**.

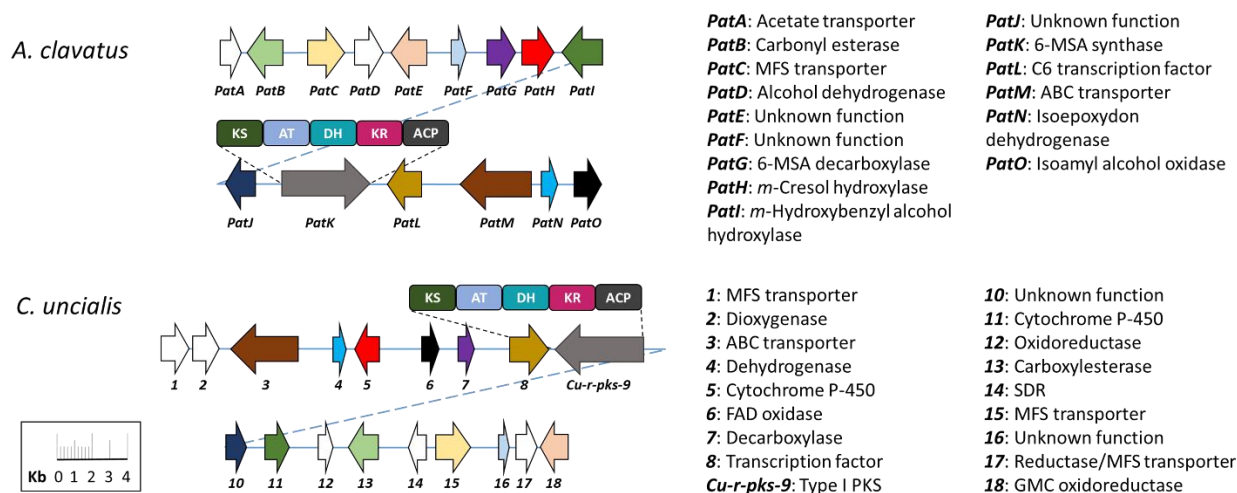


Figure 42: The *A. clavatus* patulin gene cluster (Accession no. A1CFL8) and *C. uncialis* uncharacterized gene cluster (Accession no. MG777507). Genes with matching colours are genetically similar. *Domain abbreviations:* Ketosynthase (KS), acyltransferase (AT), dehydratase (DH), ketoreductase (KR), acetyl carrier protein (ACP). *Gene abbreviations:* Polyketide synthase (PKS), major facilitator superfamily transporter (MFS transporter), 6-methylsalicylic acid (6-MSA), ATP-binding cassette transporter (ABC transporter), flavin adenine dinucleotide-dependent oxidase (FAD oxidase), short-chain dehydrogenase/reductase (SDR), glucose-methanol-choline oxidoreductase (GMC oxidoreductase). Reproduced from Bertrand et al. (2018b), in accordance with authors' retainment of privileges.

Table 5: BLAST statistics and proposed functions of the *C. uncialis* genes that are genetically similar to genes of the patulin cluster of *A. clavatus*. Reproduced from Bertrand et al. (2018b), in accordance with authors' retainment of privileges.

<i>C. uncialis</i> gene	<i>C. uncialis</i> Accession no.	<i>A. clavatus</i> gene	<i>A. clavatus</i> Accession no.	Proposed function of <i>C. uncialis</i> gene	Similarity (%) / Coverage (%)
3	AUW31349	<i>PatM</i>	XM_001273094	Putative ABC transporter	66/94
4	AUW31350	<i>PatN</i>	XM_001273095	Isoeopoxydon dehydrogenase	68/99
5	AUW31351	<i>PatH</i>	XM_001273089	<i>m</i> -Cresol hydroxylase	66/94
6	AUW31352	<i>PatO</i>	XM_001273096	Isoamyl alcohol oxidase	65/93
7	AUW31353	<i>PatG</i>	XM_001273088	6MSA decarboxylase	63/100
8	AUW31354	<i>PatL</i>	XM_001273093	Putative C6 transcription factor	46/97
<i>R-pks-9</i>	AUW31355	<i>PatK</i>	XM_001273092	6-Methylsalicylic acid synthase	60/99
10	AUW31356	<i>PatJ</i>	XM_001273091	Unknown function	77/87
11	AUW31357	<i>PatI</i>	XM_001273090	<i>m</i> -Hydroxybenzyl alcohol hydroxylase	65/89
13	AUW31359	<i>PatB</i>	XM_001273083	Putative carboxylesterase	49/91
15	AUW31361	<i>PatC</i>	XM_001273084	Putative MFS transporter	52/92
16	AUW31362	<i>PatF</i>	XM_001273087	Unknown function	59/74
18	AUW31364	<i>PatE</i>	XM_001273086	Putative GMC oxidoreductase	29/97

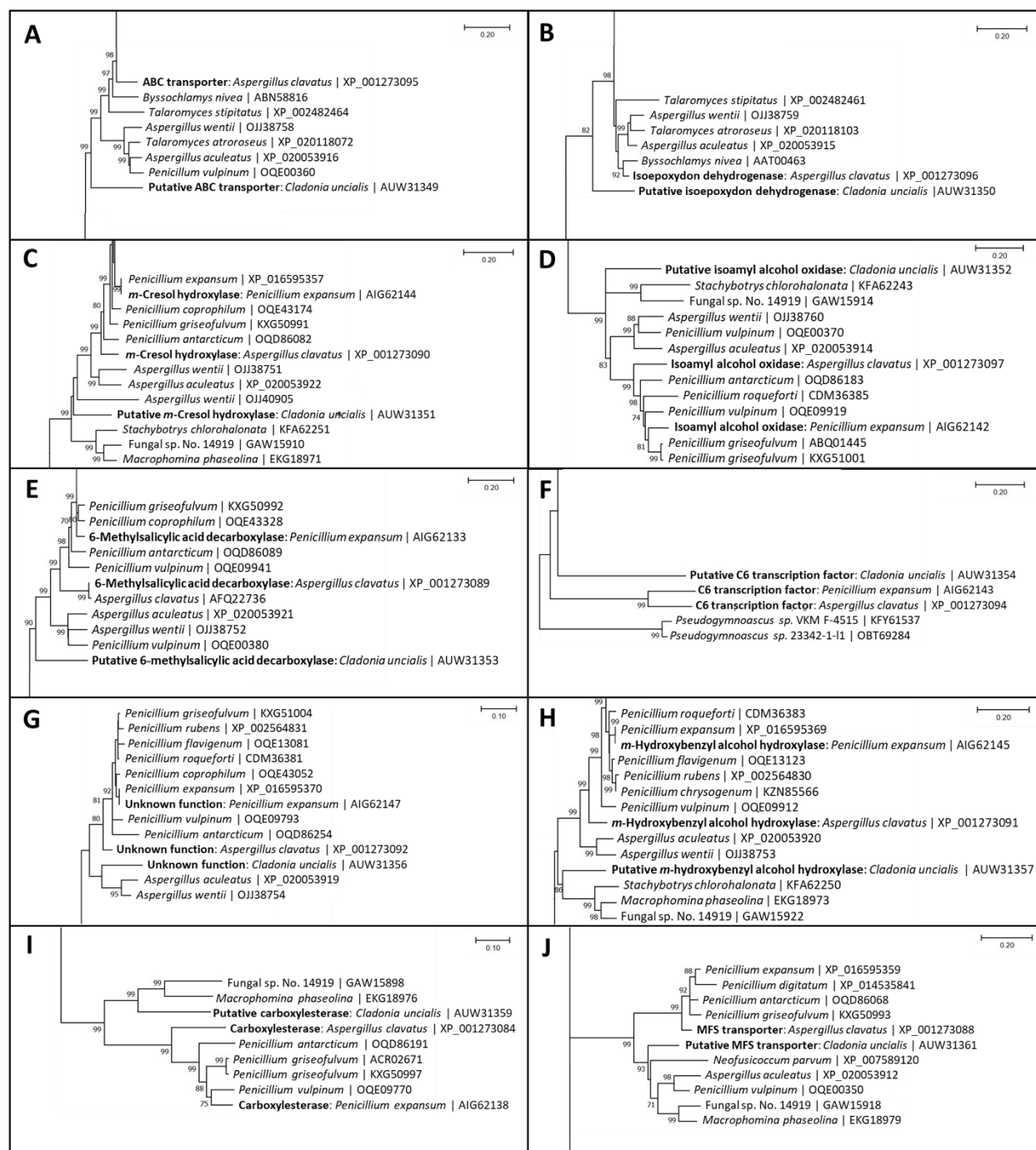


Figure 43: Truncated phylogenetic trees illustrating degree of relationship between genes of *C. uncialis* and *A. clavatus* pertinent to patulin biosynthesis. Functional homologues of *A. clavatus* genes in *P. expansum* are included: (A) ABC transporter; (B) Isoepoxydon dehydrogenase; (C) *m*-Cresol hydroxylase; (D) Isoamyl alcohol oxidase; (E) 6-Methylsalicylic acid decarboxylase; (F) C6 transcription factor; (G) Unknown function; (H) *m*-Hydroxybenzyl alcohol hydroxylase; (I) Carboxylesterase; (J) MFS transporter. Bar signifies number of amino acid substitutions per site. Complete phylogenetic trees are provided as **Figures S14-S26** in the Appendix. Reproduced from Bertrand et al. (2018b), in accordance with authors' retainment of privileges.

The KnownClusterBLAST module of AntiSMASH (V. 4.0) (Blin et al. 2017) indicated that a *C. uncialis* gene cluster was genetically similar to the patulin cluster of *Aspergillus clavatus*. Of the 18 genes present in the lichen cluster, pBLAST analysis revealed 13 gene-pairs (**Figure 42; Table 5**). The domain architectures of the *C. uncialis* and *A. clavatus* PKS are identical (**Figure 42**). Similarity pairings were not established for *PatA*, encoding a putative acetate transporter, and *PatD*, encoding a putative alcohol dehydrogenase. The role of *PatA* and *PatD* in patulin biosynthesis are unknown. Five genes within the lichen cluster were not similar to any gene member of the patulin cluster. These include a putative MFS transporter (gene *1*), a dioxygenase (gene *2*), an oxidoreductase (gene *12*), a short-chain dehydrogenase/reductase (gene *14*), and a gene ambiguously assigned as a reductase or an MFS transporter (gene *17*). Phylogenetic trees of the 13 gene pairings were generated to assess genetic homology. Functional homologues of the patulin genes of *Penicillium expansum* are also included in the phylogenetic trees. Of the 13 genetically similar gene pairs, 10 are present in a common clade. In most cases, the corresponding *P. expansum* homologue was also present (**Figure 43**). A close phylogenetic relationship could not be established for the lichen PKS (*Cu-r-pks-9*), a gene of unknown function (gene *16*), and a GMC oxidoreductase (gene *18*). It is possible that close ancestry could not be established for the PKS because its ancestry is distinct from other patulin-related genes, because it is not a 6-methylsalicylic acid synthase, or because the phylogenetic analysis was anomalous. Truncated versions of the supporting phylogenetic trees are displayed in **Figure 43** and all phylogenetic trees are provided in the Appendix. Though the function of every gene product in the patulin biosynthetic pathway in *A. clavatus* has yet to be elucidated, the high number of homology pairings lends support to the conclusion that *C. uncialis* likely has the requisite biosynthetic programming to produce patulin or a close derivative of patulin. We have therefore proposed a scheme by which

the gene products of *C. uncialis* can plausibly produce patulin in a manner similar to *A. clavatus* and *P. expansum* (**Figure 44**). Patulin has not been isolated from any species of lichen.

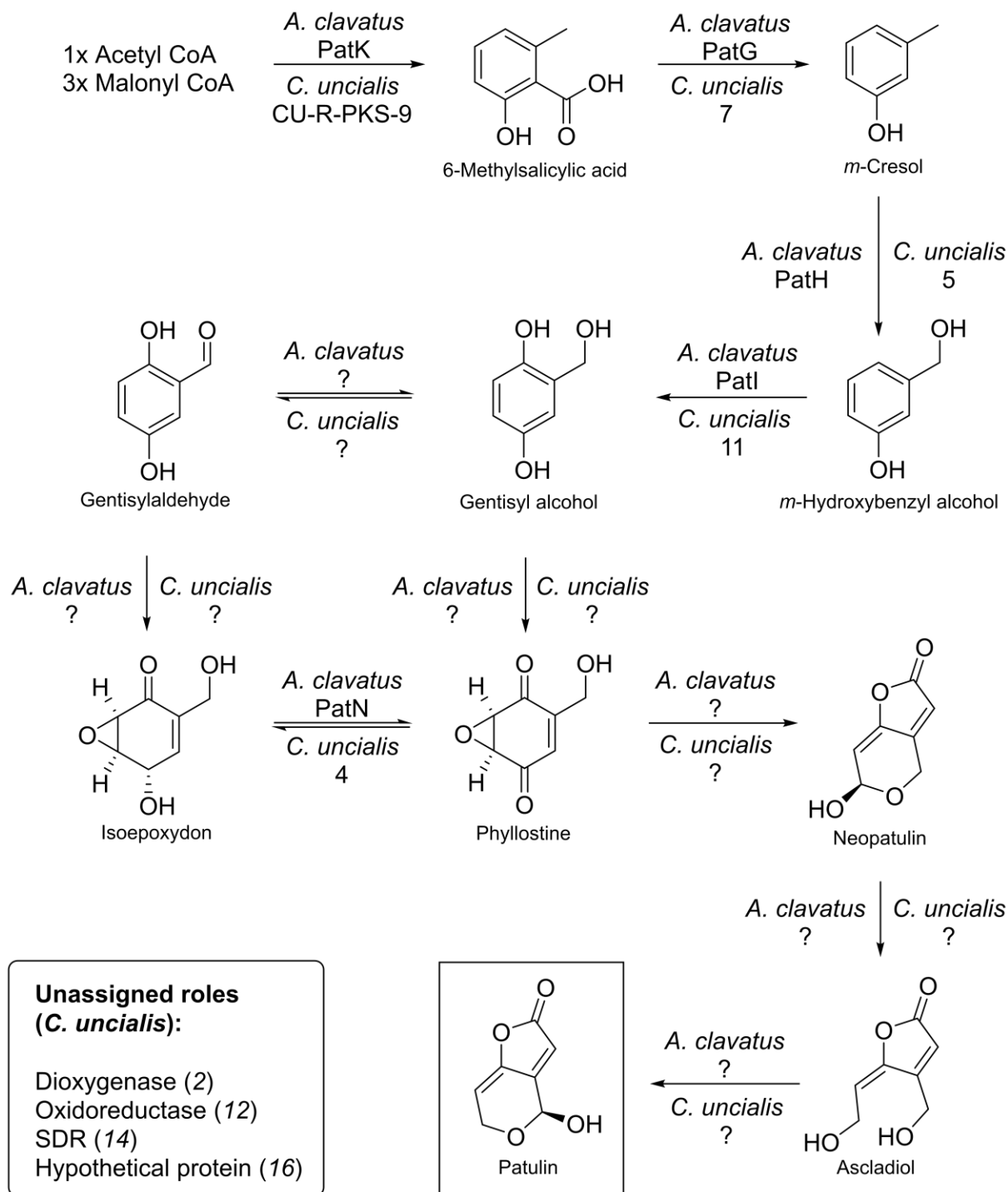


Figure 44: The experimentally supported pathway for patulin biosynthesis in *A. clavatus*, and juxtaposed, the biosynthetic pathway that is proposed to be encoded in *C. uncialis*. Reproduced from Bertrand et al. (2018b), in accordance with authors' retainment of privileges.

6.5. A proposed azaphilone-based pathway

Azaphilones are a class of fungal polyketide pigments with pyrone-quinone structures. Azaphilones display antimicrobial, antifungal, antiviral, cytotoxic, nematocidal, and anti-inflammatory properties (Osmanova et al. 2010). In *Monascus pilosus*, the gene cluster of an azaphilone pigment was identified through T-DNA random mutagenesis and selection of pigment-defective strains (Balakrishnan et al. 2013). Inactivation experiments demonstrated that the gene product of *MpPKS5* is responsible for the biosynthesis of a substituted 3-methylorcinaldehyde. The authors hypothesize that the SAT domain of MpPKS5 selects crotonyl-CoA in lieu of acetyl-CoA (Balakrishnan et al. 2013). Tailoring enzymes MppB, MppD, and MppF are then proposed to facilitate hydroxylation, esterification, and acylation, yielding rubropunctatin (Balakrishnan et al. 2013). The function of the remaining genes in the *M. pilosus* cluster are unknown. The azaphilone cluster and biosynthetic pathway are shown in **Figure 45** and **Figure 47**.

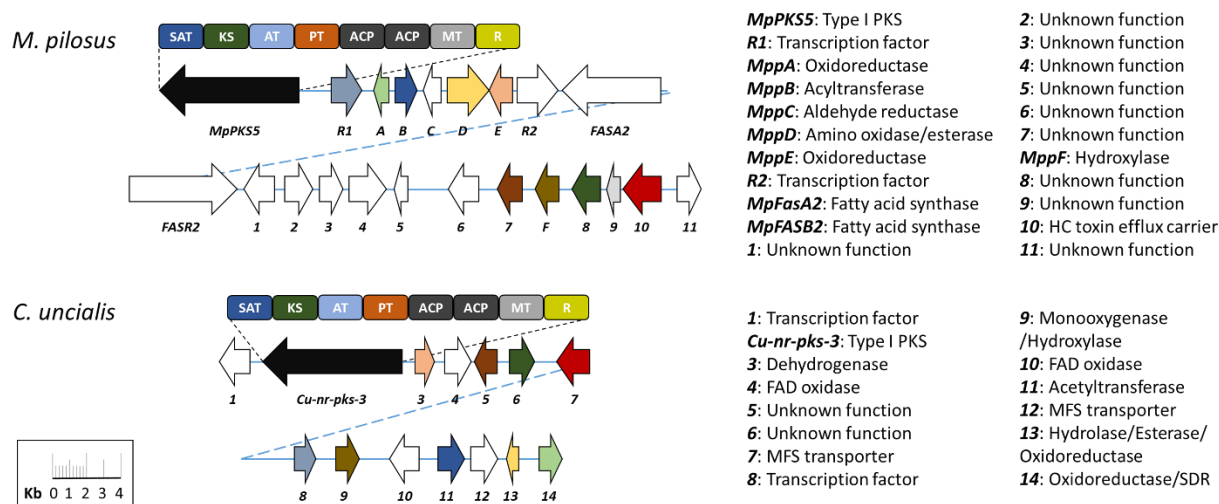


Figure 45: The *M. pilosus* azaphilone gene cluster (KC148521) and *C. uncialis* uncharacterized gene cluster (MG777491). Genes with matching colours are genetically similar. **Domain abbreviations:** Starter acyltransferase (SAT), ketosynthase (KS), acyltransferase (AT), product template domain (PT), acetyl carrier protein (ACP), C-methyltransferase (MT), reductase (R). **Gene abbreviations:** Polyketide synthase (PKS), flavin adenine dinucleotide-dependent oxidase (FAD oxidase), major facilitator superfamily transporter (MFS transporter), short-chain dehydrogenase/reductase (SDR). Reproduced from Bertrand et al. (2018b), in accordance with authors' retainment of privileges.

Table 6: BLAST statistics and proposed functions of the *C. uncialis* genes that are genetically similar to genes of the azaphilone cluster of *M. pilosus*. Reproduced from Bertrand et al. (2018b), in accordance with authors' retainment of privileges.

<i>C. uncialis</i> gene	Accession no.	<i>M. pilosus</i> gene	Accession no.	Proposed function of <i>C. uncialis</i> gene	Similarity (%) / Coverage (%)
<i>Nr-pks-3</i>	AUW31097	<i>MpPKS5</i>	AGN71604	Polyketide synthase	65/100
3	AUW31098	<i>E</i>	AGN71610	Oxidoreductase	53/100
5	AUW31100	7	AGN71622	Unknown	54/96
6	AUW31101	8	AGN71624	Unknown	59/99
7	AUW31102	10	AGN71625	Efflux transporter	54/94
8	AUW31103	<i>R1</i>	AGN71605	Transcriptional regulator	50/90
9	AUW31104	<i>F</i>	AGN71623	Hydroxylase	64/98
11	AUW31106	<i>B</i>	AGN71607	Acyltransferase	52/100
13	AUW31108	<i>D</i>	AGN71609	Amino oxidase/esterase	62/97
14	AUW31109	<i>A</i>	AGN71606	Oxidoreductase	64/96

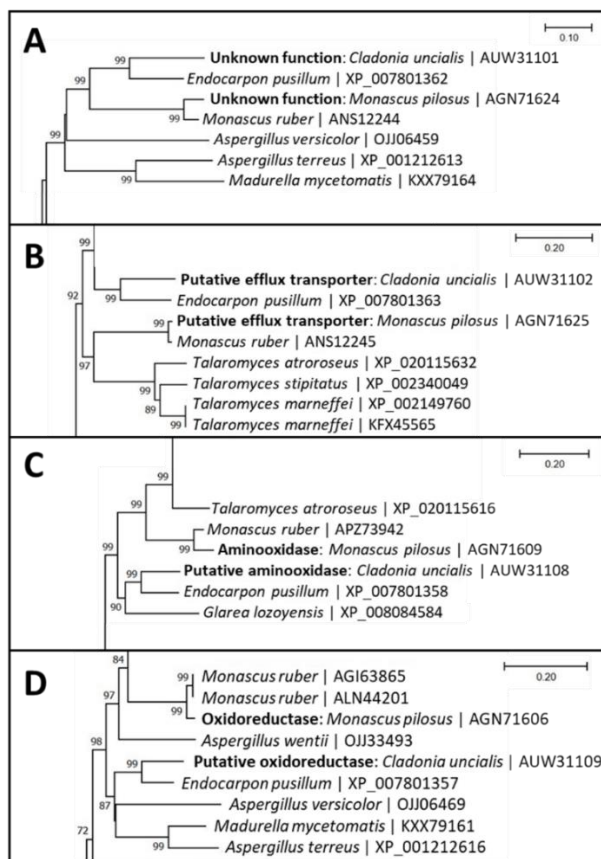


Figure 46: Truncated phylogenetic trees illustrating degree of relationship between genes of *C. uncialis* and *M. pilosus* pertinent to azaphilone biosynthesis: (A) Unknown function; (B) Efflux transporter; (C) Amino-oxidase; (D) Oxidoreductase. Bar signifies number of amino acid substitutions per site. Complete phylogenetic trees are provided as **Figures S27-S36** in the Appendix. Reproduced from Bertrand et al. (2018b), in accordance with authors' retainment of privileges.

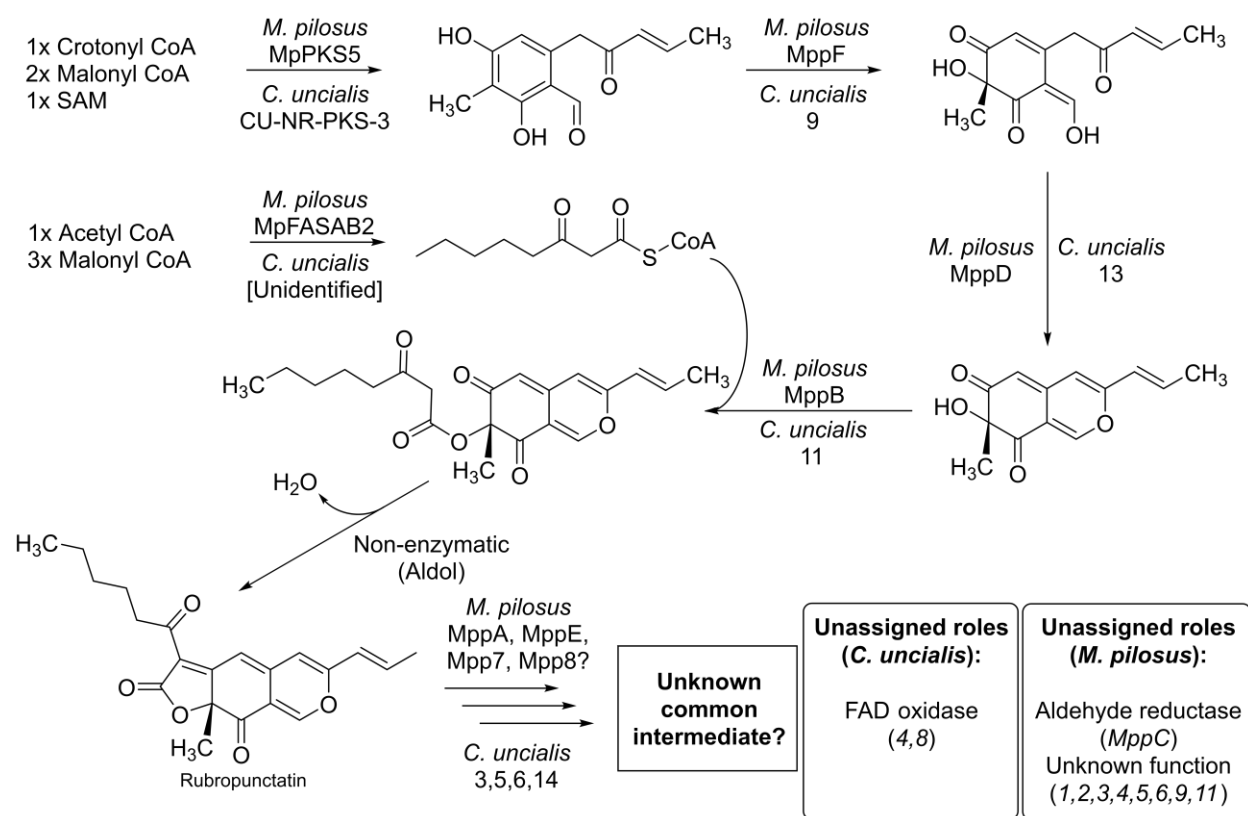


Figure 47: The experimentally supported pathway for azaphilone biosynthesis in *M. pilosus*, and juxtaposed, the biosynthetic pathway that is proposed to be encoded in *C. uncialis*. Reproduced from Bertrand et al. (2018b), in accordance with authors' retainment of privileges.

The KnownClusterBLAST module of AntiSMASH (V. 4.0) (Blin et al. 2017) suggested that a *C. uncialis* gene cluster was genetically similar to the azaphilone pathway of *M. pilosus*. A total of 10 similarities were observed (Table 6; Figure 45). Phylogenetic support of genetic homology was only limited because four of 10 gene-pairs were found to cluster within a common clade (Figure 46). The four genetically homologous genes include an amino-oxidase, an oxidoreductase, an efflux transporter, and a gene of unknown function. Although the lack of phylogenetic support does not automatically invalidate predictions of function, the viability of this prediction is weakened by inference that these genes are evolutionarily distant and therefore are more likely to have divergent functions. Nonetheless, a biosynthetic pathway will be proposed.

With the exception of a fatty acid synthase not found in the lichen cluster, all requisite genes for rubropunctatin biosynthesis are present. A plausible biosynthetic pathway can be proposed leading to rubropunctatin (**Figure 47**). The fatty acid could be provided by a non-dedicated fatty acid synthase. Four genes in *C. uncialis* were found to be genetically similar to genes of *M. pilosus* though their product roles in *M. pilosus* have not been determined. It is possible that the products of these gene pairings function similarly to produce a derivative of rubropunctatin that is common to both *M. pilosus* and *C. uncialis* (**Figure 47**).

6.6. A proposed cytochalasin-based pathway

The cytochalasins are a group of fungal polyketide-amino acid chimeric natural products produced by a hybrid PKS-NRPS. The macrocyclic ring and isoindolone moieties characteristic of this group of natural products are derivatized from a highly-reduced polyketide backbone that is fused to a phenylalanine (Scherlach et al. 2010). The gene cluster responsible for the biosynthesis of cytochalasins E and K have been identified in *Aspergillus clavatus* through genome surveying and gene inactivation (Qiao et al. 2011). Biosynthesis is proposed to proceed through seven rounds of Claisen condensation between an acetyl-CoA primer and seven units of malonyl-CoA to form a C₁₆-polyketide backbone that is methylated three times. The polyketide chain is fused to a phenylalanine by the NRPS module and released by an NADPH-dependent terminal reductase (R) domain. The NRPS-PKS (*CcsA*) lacks a functional ER domain required for complete reduction of β -keto sites. A *trans*-enoylreductase (*CcsC*) encoded downstream of *CcsA* is proposed to fulfill this role. The released aminoaldehyde intermediate is readily cyclized to form a pentaene intermediate. This intermediate then undergoes a series of modifications performed by five tailoring enzymes to produce cytochalasins E and K (Qiao et al. 2011). The cytochalasin gene

cluster and an abbreviated cytochalasin pathway illustrating the role of CcsA and CcsC are displayed in **Figure 48** and **Figure 50**.

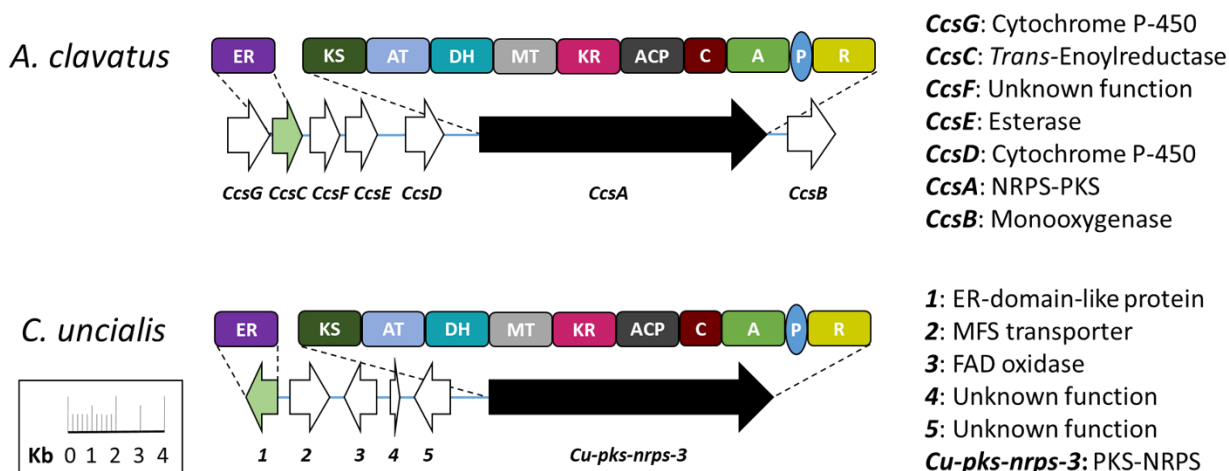


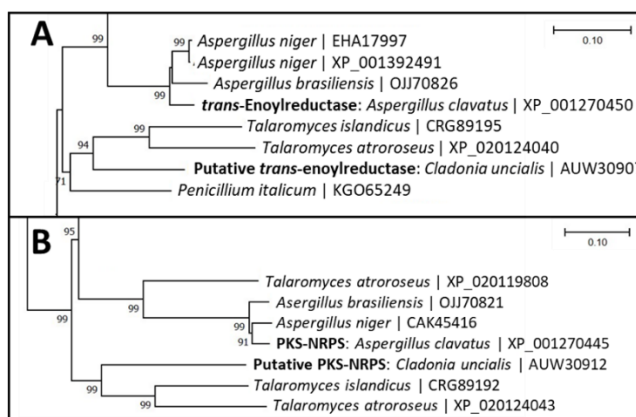
Figure 48: The *A. clavatus* cytochalasin gene cluster (Accession no. DS027057) and *C. uncialis* uncharacterized gene cluster (Accession no. MG777482). Genes with matching colours are genetically similar. *Domain abbreviations:* Enoylreductase (ER), ketosynthase (KS), acyltransferase (AT), dehydratase (DH), C-methyltransferase (MT), ketoreductase (KR), acetyl carrier protein (ACP), condensation (C) adenylation (A), peptidyl carrier protein (PCP), reductase (R). *Gene abbreviations:* Hybrid polyketide synthase-non-ribosomal peptide synthetase (PKS-NRPS), major facilitator superfamily transporter (MFS transporter), flavin adenine dinucleotide-dependent oxidase (FAD oxidase). Reproduced from Bertrand et al. (2018b), in accordance with authors' retainment of privileges.

Table 7: BLAST statistics and proposed functions of the *C. uncialis* genes that are genetically similar to genes of the cytochalasin cluster of *A. clavatus*. Reproduced from Bertrand et al. (2018b), in accordance with authors' retainment of privileges.

<i>C. uncialis</i> gene	Accession no.	<i>A. clavatus</i> gene	Accession no.	Proposed function of <i>C. uncialis</i> gene	Similarity (%) / Coverage (%)
1	AUW30907	CcsC	XP_001270547	<i>Trans</i> -enoyl reductase	42/92
Pks-nrps-3	AUW30912	CcsA	XP_001270543	Cytochalasin synthase	52/95

A pBLAST analysis suggests that two genes in *C. uncialis* are similar to cluster members of the *A. clavatus* cytochalasin pathway. These include a *C. uncialis* NRPS-PKS (*Cu-pks-nrps-3*) and a *trans*-enoylreductase. BLAST statistics and similarity pairings are shown in **Table 7** and **Figure 48**. The domain architecture of *Cu-pks-nrps-3* is identical to *CcsA* including the requisite 'R' terminal domain and the presence of a *trans*-enoylreductase encoded near the NRPS-PKS (**Figure 48**). Phylogenetic analyses support close genetic homology because both gene pairs were

found to cluster within common clades (**Figure 49**). Four genes in *C. uncialis* did not have genetic similarity. These include an MFS transporter (2), a FAD oxidase (3), and two genes of unknown function (4,5). It is proposed that the first two chemical steps in cytochalasin biosynthesis are encoded in *C. uncialis*. Derivatization of this common intermediate may then occur to produce an unknown final product (**Figure 50**).



It is interesting that the phylogenetic analysis of both cytochalasin-associated genes of *C. uncialis* also clustered with genes originating from *Talaromyces sp.*, *Penicillium sp.*, and *Aspergillus sp.* (**Figure 49**). The cyclized phenylalanine moiety of cytochalasin is also present in metabolites isolated from these species. Talaroconvolutin A, citridone C, carbonarin C, and preaspyridone A are four examples (Suzuki et al. 2000; Fukuda et al. 2005; Alfatafta et al. 1998; Xu et al. 2010). Though no prediction could be offered of the final metabolite arising from the lichen cluster, the existence of these fungal metabolites lends *a posteriori* support for the proposition that the encoded NRPS-PKS in *C. uncialis* produces a metabolite possessing the core phenylalanine moiety (**Figure 51**).

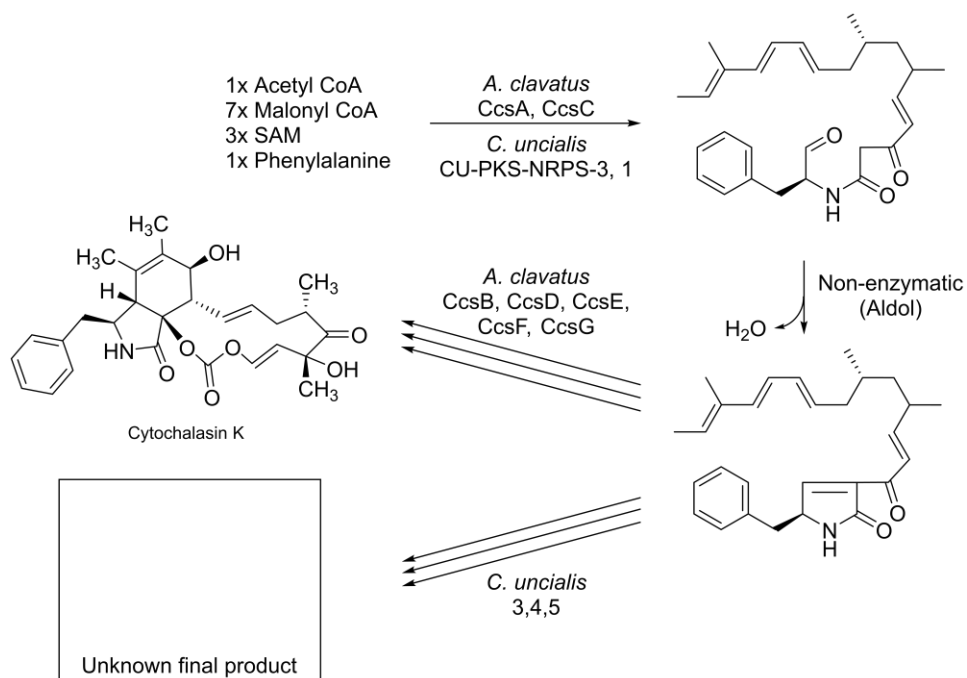


Figure 50: The experimentally supported pathway for cytochalasin biosynthesis in *A. clavatus*, and juxtaposed, the biosynthetic pathway that is proposed to be encoded in *C. uncialis*. Reproduced from Bertrand et al. (2018b), in accordance with authors' retainment of privileges.

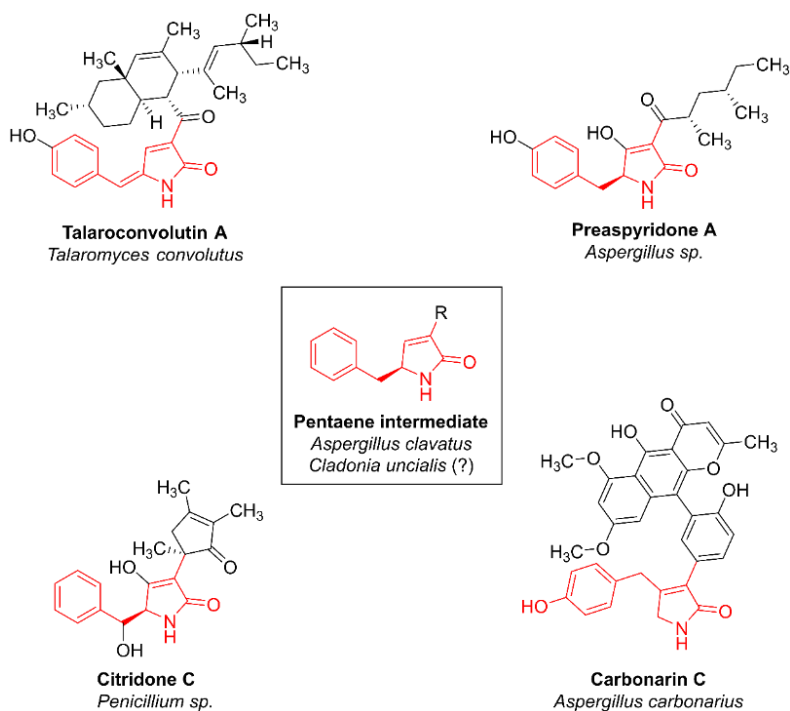


Figure 51: Examples of secondary metabolites possessing the cyclized phenylalanine moiety from *Talaromyces sp.*, *Penicillium sp.*, and *Aspergillus sp.* Reproduced from Bertrand et al. (2018b), in accordance with authors' retainment of privileges.

6.7. A proposed fusarubin-based pathway

Fusarubin is a naphthoquinone pigment produced by the fungus *Fusarium fujikuroi*. Deletion of *fsr1* abated fusarubin production confirming Fsr1 as the non-reducing PKS devoted to fusarubin biosynthesis (Studt et al. 2012). Northern blot analyses demonstrated co-regulation of *Fsr1* with *Fsr2* to *Fsr6* suggesting that these six genes comprise the fusarubin gene cluster. Deletion experiments on each of the tailoring genes and chemical identification of the intermediate compounds elucidated the biosynthetic pathway. Two compounds similar in structure to 8-*O*-methylfusarubinaldehyde were identified in culture extracts, namely, 8-*O*-methylanhydrofusarubinlactol and 8-*O*-methyl-13-hydroxynorjavanicin (**Figure 53**). Although it remains unclear how these structures are formed, 8-*O*-methylfusarubinic acid is a hypothesized intermediate that could lead to both compounds (Studt et al. 2012). It remains unclear what gene products are responsible for facilitating these latter reactions. Deletion of *fsr4* (putative short-chain dehydrogenase/reductase) and *fsr5* (putative oxidoreductase) did not disrupt fusarubin biosynthesis. The authors hypothesize that these gene products have regulatory roles or are responsible for the derivatization of fusarubin an unidentified products (Studt et al. 2012). The fusarubin gene cluster and biosynthetic pathway are shown in **Figure 52** and **Figure 53**.

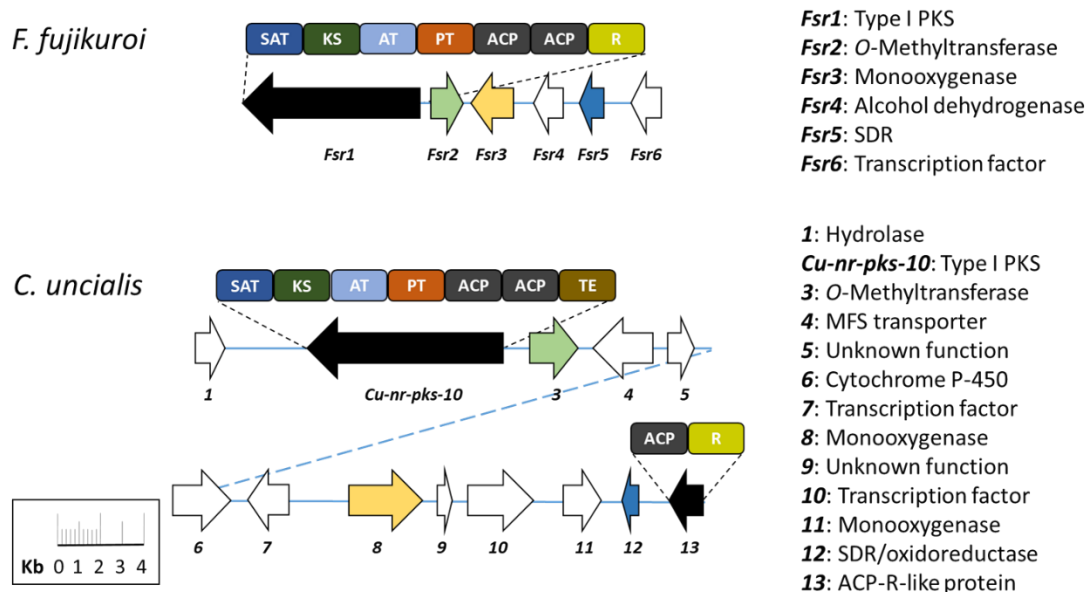


Figure 52: The *F. fujikuroi* fusarubin gene cluster (Accession no. HE613440) and *C. uncialis* uncharacterized gene cluster (Accession no. MG777498). Genes with matching colours are genetically similar. *Domain abbreviations:* Starter acyltransferase (SAT), ketosynthase (KS), acyltransferase (AT), product template domain (PT), acetyl carrier protein (ACP), reductase (R), thioesterase (TE). *Gene abbreviations:* Polyketide synthase (PKS), short-chain dehydrogenase/reductase (SDR), major facilitator superfamily transporter (MFS transporter). Reproduced from Bertrand et al. (2018b), in accordance with authors' retainment of privileges.

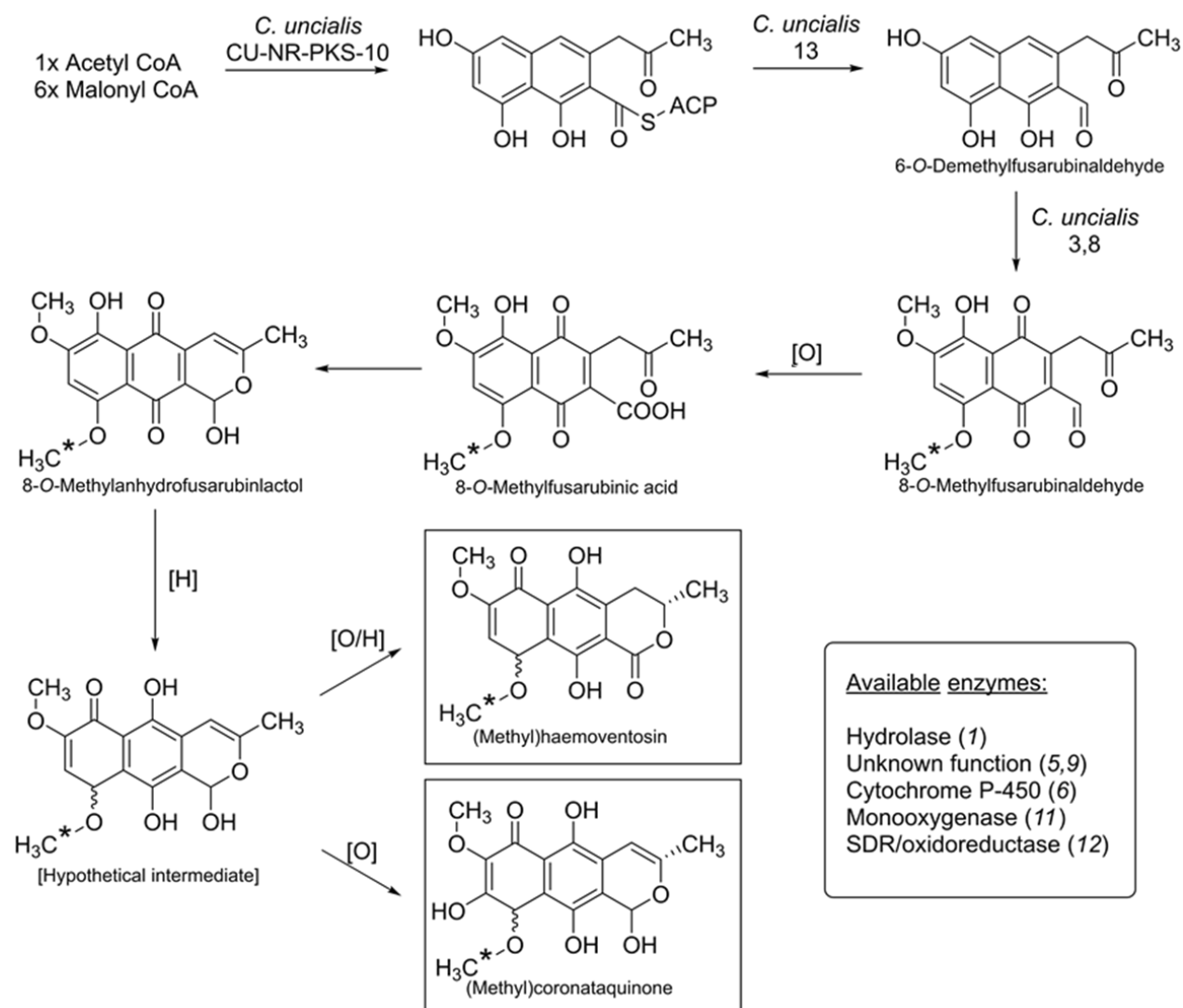
Table 8: BLAST statistics and proposed functions of the *C. uncialis* genes that are genetically similar to genes of the fusarubin pathway of *F. fujikuroi*. Reproduced from Bertrand et al. (2018b), in accordance with authors' retainment of privileges.

<i>C. uncialis</i> gene	Accession no.	<i>F. fujikuroi</i> gene	Accession no.	Proposed function of <i>C. uncialis</i> gene	Similarity (%) / Coverage (%)
<i>Nr-pks-10</i>	AUW31210	<i>Fsr1</i>	CCE67070	6-O-demethylfusarubinaldehyde synthase	40/80
3	AUW31211	<i>Fsr2</i>	CCE67071	6-O-demethylfusarubinaldehyde O-methyltransferase	43/85
8	AUW31216	<i>Fsr3</i>	CCE67072	Multi-functional monooxygenase	52/93
12	AUW31220	<i>Fsr5</i>	CCE67074	Putative SDR	36/97
13	AUW31221	<i>Fsr1</i>	CCE67070	Putative <i>trans</i> -reductase domain	22/97

Four genes in *C. uncialis* presented similarity pairings to the *F. fujikuroi* fusarubin pathway suggesting that a possible function of the lichen cluster involves biosynthesis of a naphthoquinone. BLAST similarity scores and pairings are shown (**Figure 52; Table 8**). The identification of the lichen PKS, *Cu-nr-pks-10*, as a possible genetic homologue of the *F. fujikuroi* PKS *fsr1* is supported by a pBLAST similarity score of 40 percent (**Table 8**). However, the suggestion that *Cu-nr-pks-10* is a plausible functional homologue of *fsr1* is not supported by the domain

architecture of the lichen PKS. Whereas the lichen PKS possesses a terminal thioesterase (TE), the *F. fujikuroi* PKS possesses a terminal reductase (R) domain (**Figure 52**). One gene in the lichen cluster is unusual because it appears to have acyl carrier protein (ACP) and terminal reductase (R) domain-like features (**Figure 52**). It is possible that this gene product serves as a *trans*-reductase and that the terminal thioesterase of CU-NR-PKS-10 is inoperative. Inactive domains are commonly found in PKS and can be identified by determining the integrity of conserved active-site motifs. Three examples include an inactive KR in the stigmatellin PKS (Gaitatzis et al. 2002), an inactive ER in the lovastatin PKS (Kennedy et al. 1999), and two inactive DH in the tacrolimus PKS (Motamed & Shafiee, 1998). Should the terminal thioesterase of CU-NR-PKS-10 be inactive, then a biosynthetic route to the first intermediate in fusarubin biosynthesis would be feasible. Genetic similarities extend to an *O*-methyltransferase (3 | *fsr2*), a monooxygenase (8 | *fsr3*), and a short-chain dehydrogenase/reductase (12 | *fsr5*), extending possible chemical transformations in *C. uncialis* to those known to occur in *F. fujikuroi* (**Figure 53**). Three lichen genes, a hydrolase (1), a cytochrome P-450 oxidase (6), and a monooxygenase (11), remain unassigned. Phylogenetic analysis suggests that these gene clusters have distant relationships and therefore close genetic homology cannot be inferred (see **Figures S39-S42** in Appendix).

11), an SDR/oxidoreductase (gene 12) and two genes of unknown function (genes 5 and 9). These gene products could plausibly perform the transformations leading to both metabolites. The presence or absence of the *O*-methyl residue (indicated by CH₃*) on *O*-methylhaemoventosin and *O*-methylcoronataquinone are variable (**Figure 54**). We propose that the end-product of the *C. uncialis* gene cluster is derived from fusarubin pathway intermediates and may also be structurally similar to coronataquinone or haemoventosin.



Scheme 52: A plausible biosynthetic pathway leading from the fusarubin intermediate 8-*O*-methylanhydrofusarubinlactol to the lichen metabolites haemoventosin and coronataquinone. The presence or absence of a methyl group (indicated as CH₃*) is variable. Reproduced from Bertrand et al. (2018b), in accordance with authors' retainment of privileges.

6.8. A proposed pestheic acid-based pathway

Pestalotiopsis fici is an endophytic fungus that produces the plant growth regulator pestheic acid. The polyketide gene cluster for pestheic acid was recently identified in *P. fici* via gene deletion (Xu et al. 20130. The gene cluster consists of a PKS (*PtaA*), 11 tailoring genes (*PtaB* to *PtaM*), and three regulatory genes (*R1* to *R3*). The domain architecture of *PtaA* is notable for lacking a terminal domain. The β -lactamase (*PtaB*) is proposed to replace the function of the absent terminal domain of *PtaA* by facilitating hydrolysis and Claisen cyclization of the polyketide, yielding atrochrysone carboxylic acid. The spontaneous dehydration of the unstable intermediate produces endocrocin anthrone. An oxygenase (*PtaC*) is then proposed to oxidize endocrocin anthrone, yielding endocrocin. Spontaneous decarboxylation forms emodin. A series of additional modifications by the remaining tailoring enzymes on endocrocin produces pestheic acid (Xu et al. 2013). The pestheic acid gene cluster and an abridged pathway illustrating the roles of PtaA to PtaC in pestheic acid biosynthesis is shown in **Figure 55** and **Figure 56**.

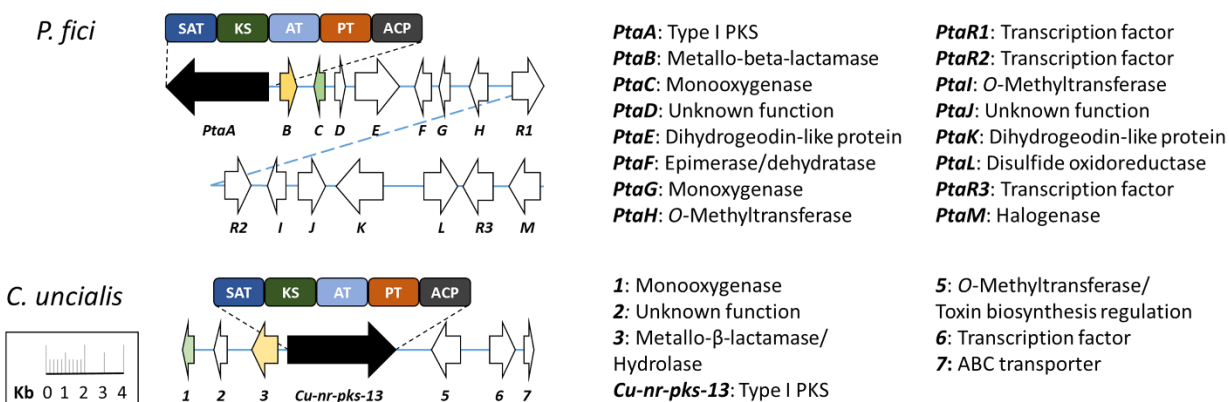


Figure 55: The *P. fici* pestheic acid gene cluster (Accession no. KC145148) and *C. uncialis* uncharacterized gene cluster (Accession no. MG777501). Genes with matching colours are genetically similar. **Domain abbreviations:** Starter acyltransferase (SAT), ketosynthase (KS), acyltransferase (AT), product template domain (PT), acetyl carrier protein (ACP). **Gene abbreviations:** Polyketide synthase (PKS), ATP-binding cassette transporters (ABC transporter). Reproduced from Bertrand et al. (2018b), in accordance with authors' retainment of privileges.

Table 9: BLAST statistics and proposed functions of the *C. uncialis* genes that are genetically similar to genes of the pestheic acid cluster of *P. fici*. Reproduced from Bertrand et al. (2018b), in accordance with authors' retainment of privileges.

<i>C. uncialis</i> gene	Accession no.	<i>P. fici</i> gene	Accession no.	Proposed function of <i>C. uncialis</i> gene	Similarity (%) / Coverage (%)
1	AUW31243	<i>PtaC</i>	AGO59043	Endocrocin anthrone monooxygenase	47/99
3	AUW31245	<i>PtaB</i>	AGO59041	Atrochrysone carboxylic acid thioesterase	55/96
<i>Nr-pks-13</i>	AUW31246	<i>PtaA</i>	AGO59040	Atrochrysone carboxylic acid synthase	71/88

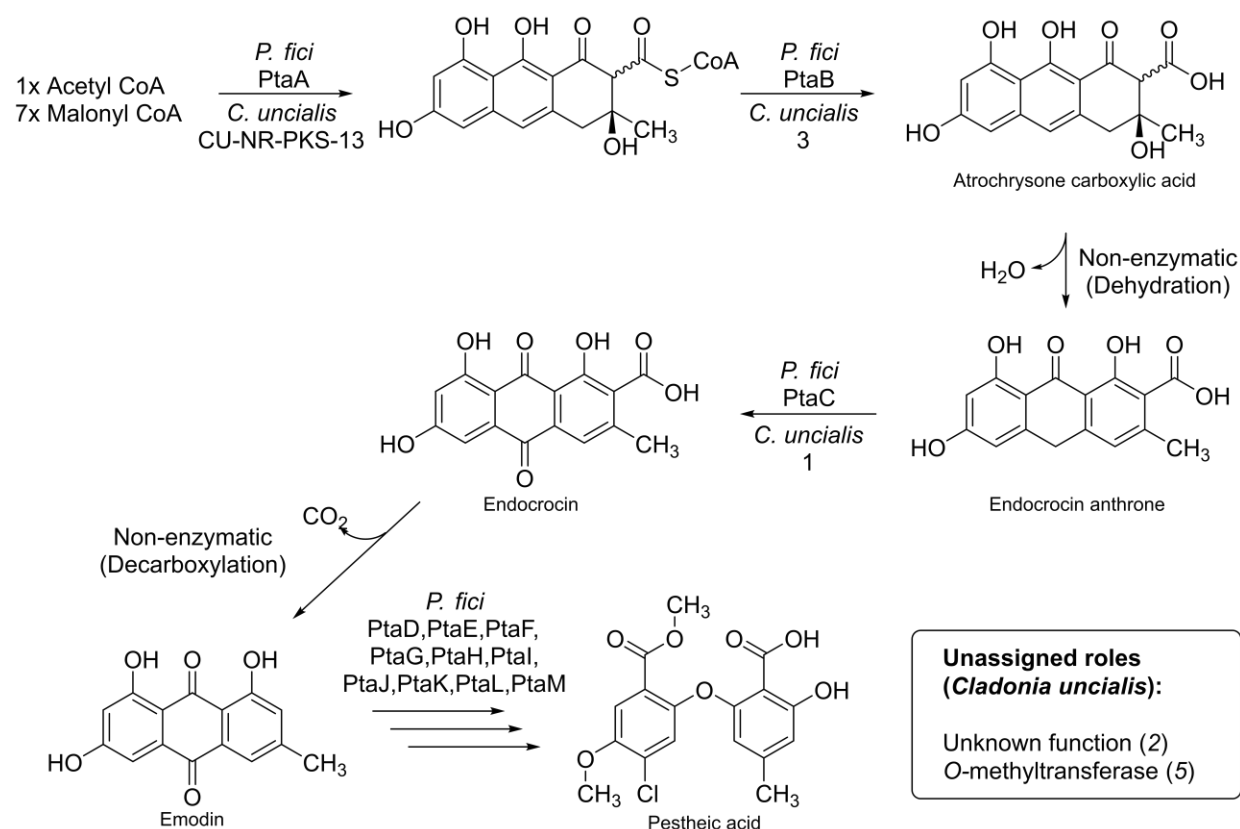


Figure 56: The experimentally confirmed pathway for pestheic acid biosynthesis in *P. fici*, and juxtaposed, the biosynthetic pathway that is proposed to be encoded in *C. uncialis*. Reproduced from Bertrand et al. (2018b), in accordance with authors' retainment of privileges.

A pBLAST analysis suggested that the first three genes involved in the *P. fici* pestheic acid pathway (*PtaA* to *PtaC*) are genetically similar to three genes in *C. uncialis*. The putative assignment of the *C. uncialis* PKS (*Cu-nr-pks-13*) as atrochrysone carboxylic acid synthase is

strengthened by a robust similarity score of 71 percent (**Table 9**). This putative assignment is also supported by similar PKS domain architecture (including the absent terminal domain) and the proximity of a metallo- β -lactamase/hydrolase similar to that of *P. fici* (**Figure 55**). The putative assignment is however weakened by the absence of convincing phylogenetic support from any of the three gene pairings (see **Figures S43-S45** in Appendix). A metallo- β -lactamase/hydrolase, homologous to *PtaB*, is proposed to facilitate chain cleavage from CU-NR-PKS-13 yielding atrochrysone carboxylic acid (**Figure 56**). The gene product of *l*, similar to *PtaC*, is proposed to oxidize endocrocin to emodin. Two lichen genes are not genetically similar to *P. fici* genes. These include a gene of unknown function (gene 2) and a gene ambiguously identified as an *O*-methyltransferase or a toxin regulator (gene 5). A biosynthetic pathway is proposed (**Figure 56**).

Rather than derivatizing emodin to some final product, this gene cluster could be devoted to producing emodin itself. Emodin is an intermediate in the biosynthesis of many lichen polyketides, and it is possible that the purpose of this gene cluster is to supply emodin as starting substrate in other pathways. Several lichen metabolites possessing an emodin scaffold could be produced with a minimal number of tailoring steps from emodin. Examples include 5-chloroemodin, isolated from the lichen *Nephroma laevigatum*, 7-chloro-4-hydroxyemodin, isolated from *Lasallia papulosa*, 7-chloro-1-*O*-methyl-*w*-hydroxyemodin, isolated from the lichen *Nephroma laevigatum*, and skyrin, isolated from the lichen *Rinodina peloleuca* (Cohen et al. 1996; Matzer et al. 1998; Bohman, 1969) (**Figure 57**).

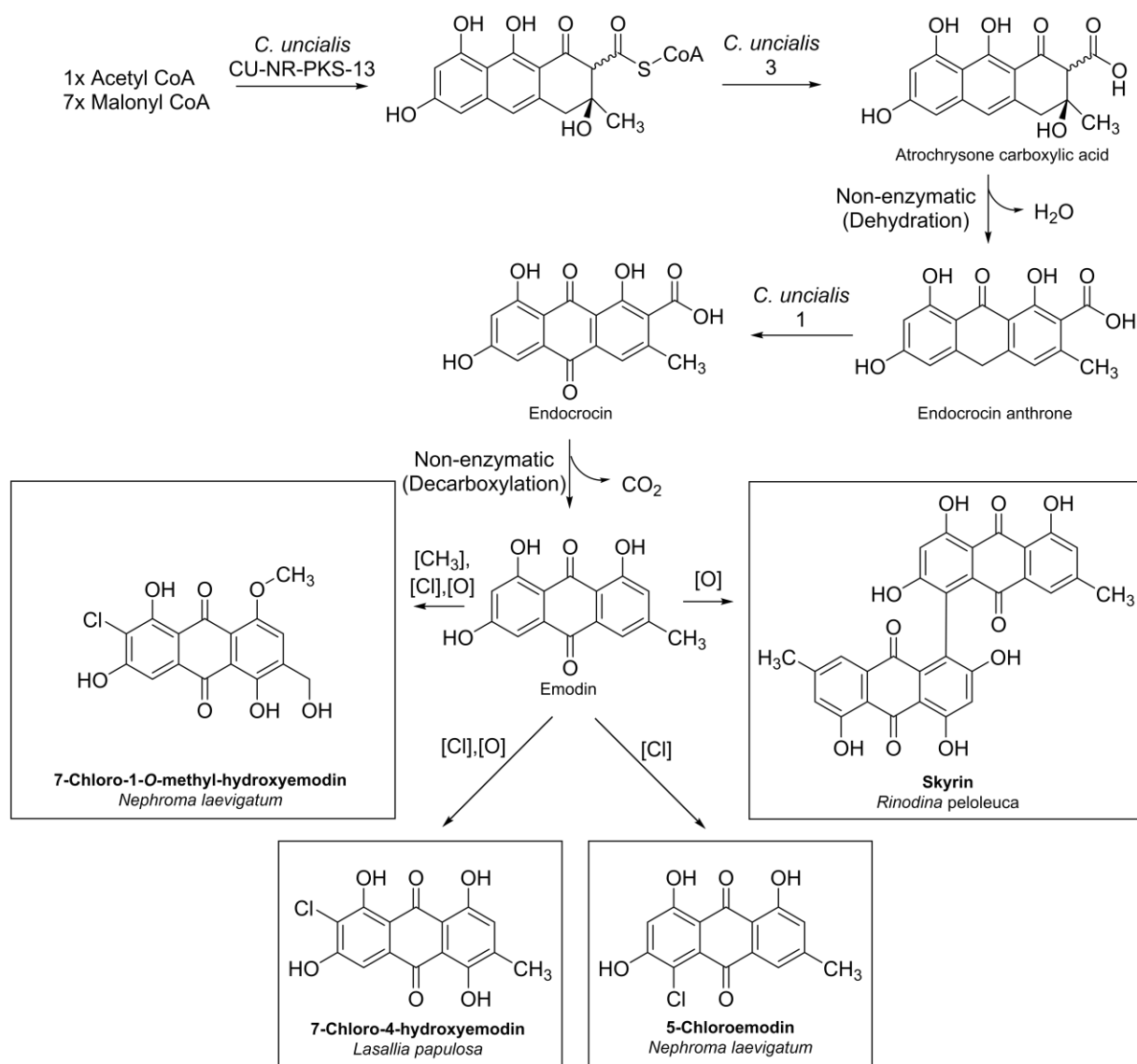


Figure 57: Proposed biosynthesis of anthraquinone lichen natural products from emodin. Reproduced from Bertrand et al. (2018b), in accordance with authors' retainment of privileges.

6.9. A proposed mycophenolic acid-based pathway

Mycophenolic acid is a fungal metabolite consisting of a phthalide derivatized from 5-methylorsellinic acid and a sesquiterpene. It was first discovered for its effects against *Bacillus anthracis* (Bentley, 2000). The gene cluster devoted to mycophenolic acid biosynthesis was discovered in *Penicillium brevicompactum* (Regueira et al. 2011). Deletion of the PKS (*MpaC*) conclusively demonstrated that biosynthesis of mycophenolic acid begins with 5-methylorsellinic

acid as the first intermediate (Regueira et al. 2011). *MpaD* is hypothesized to hydroxylate 5-methylorsellinic acid because *MpaD* is the only candidate gene with similarity to cytochrome P-450 monooxygenases (Regueira et al. 2011). Formation of the lactone moiety is likely performed by the gene product of *MpaE* because *MpaD* and *MpaE* form a heterodimer that facilitates lactonization *in vivo* (Hansen et al. 2012). The remaining tailoring enzymes are proposed to fixate the farnesyl pyrophosphate to the phthalide intermediate, methylate a hydroxy group, and perform an oxidative cleavage, yielding mycophenolic acid. As no terpene synthase is located in the mycophenolic acid gene cluster of *P. brevicompactum*, the farnesyl pyrophosphate is hypothesized to be produced by a non-dedicated terpene synthase. The mycophenolic acid gene cluster and an abridged biosynthetic pathway are shown in **Figure 58** and **Figure 59**.

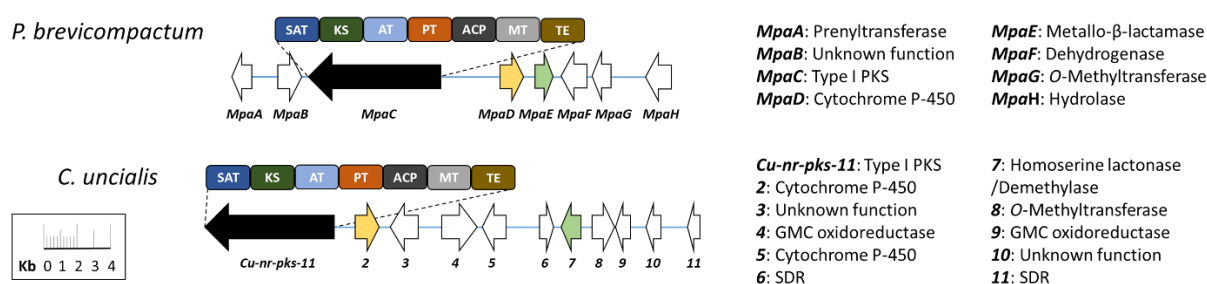


Figure 58: The *P. brevicompactum* mycophenolic acid gene cluster (Accession no. HQ731031) and *C. uncialis* unassigned gene cluster (Accession no. MG777499). Colour coding indicates similarity pairings. **Domain abbreviations:** Starter acyltransferase (SAT), ketosynthase (KS), acyltransferase (AT), product template domain (PT), acetyl carrier protein (ACP), C-methyltransferase (MT), thioesterase (TE). **Gene abbreviations:** Polyketide synthase (PKS), glucose-methanol-choline oxidoreductase (GMC oxidoreductase), short-chain dehydrogenase/reductase (SDR). Reproduced from Bertrand et al. (2018b), in accordance with authors' retainment of privileges.

Table 10: BLAST statistics and proposed functions of the *C. uncialis* genes that are genetically similar to genes of the mycophenolic acid cluster of *P. brevicompactum*. Reproduced from Bertrand et al. (2018b), in accordance with authors' retainment of privileges.

<i>C. uncialis</i> gene	Accession no.	<i>P. brevicompactum</i> gene	Accession no.	Proposed function of <i>C. uncialis</i> gene	Similarity (%) / Coverage (%)
<i>Nr-pks-11</i>	AUW31224	<i>MpaC</i>	ADY00130	5-methylorsellinic acid synthase	41/98
2	AUW31225	<i>MpaD</i>	ADY00131	5-methylorsellinic acid oxidase	60/67
7	AUW31230	<i>MpaE</i>	ADY00132	4,6-dihydroxy-2-hydroxymethyl-3-methylbenzoic acid lactamase	50/97

Three similarity pairings were observed in *C. uncialis* suggesting that functional homologues of the *P. brevicompactum* mycophenolic acid pathway could be encoded in the lichen. BLAST similarity scores and pairings are shown in **Table 10** and **Figure 58**. The lichen PKS possesses identical domain architecture to *MpaC* (**Figure 58**). This PKS is proposed to produce 5-methylorsellinic acid (**Figure 59**). *C. uncialis* genes, 2 and 7, similar to *MpaD* and *MpaE*, respectively, are proposed to facilitate lactonization, yielding the last common intermediate, 5,7-dihydroxy-4-methylphthalide (**Figure 59**). The lichen cluster also contains two GMC oxidoreductases (4,9), a cytochrome p450 (5), two short-chain dehydrogenases/reductases (6,11), an *O*-methyltransferase (8), and two genes of unknown function (3,10). A role of these gene products may include derivatization of 5,7-dihydroxy-4-methylphthalide to form a distinct final product. Phylogenetic analyses of the three gene pairings did not support genetically homologous relationships, weakening the predictive utility of the proposal (see **Figure S46-S48** in Appendix).

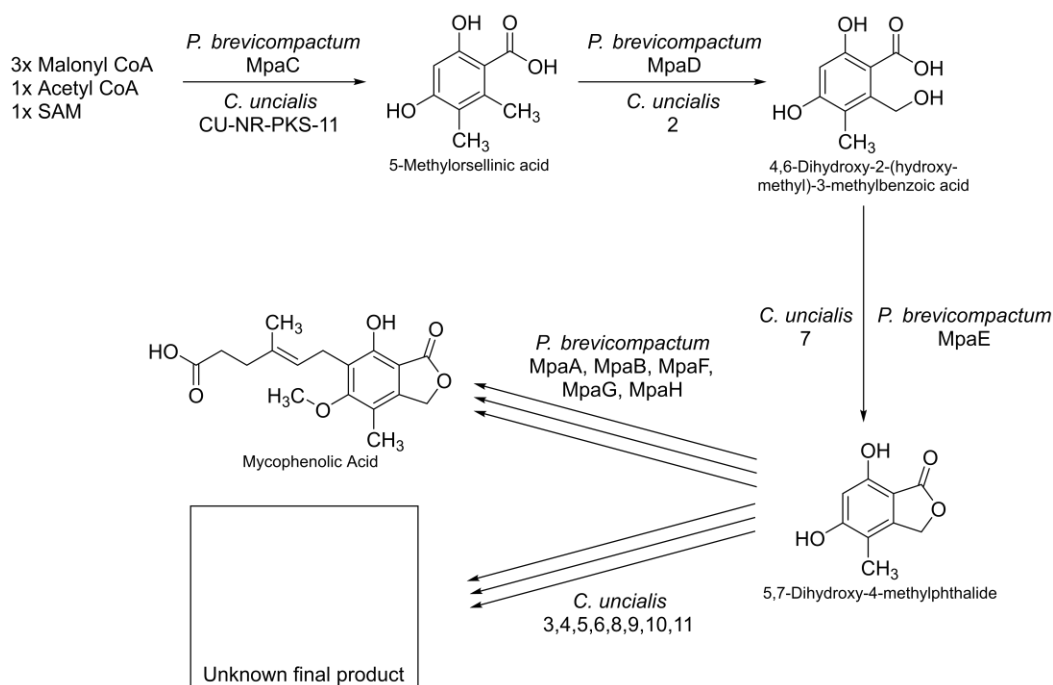


Figure 59: The experimentally confirmed pathway for mycophenolic acid biosynthesis in *P. brevicompactum*, and juxtaposed, the proposed biosynthetic pathway encoded in *C. uncialis*. Reproduced from Bertrand et al. (2018b), in accordance with authors' retainment of privileges.

6.10. Summary

In Chapter 5, I described the annotation of 48 secondary metabolite gene clusters in *C. uncialis*. By applying the ‘homology mapping’ approach described in Chapter 4, proposals of function for eight additional gene clusters were generated. These proposals include what appear to be complete biosynthetic pathways for grayanic acid, patulin, and betaenones A-C. Some gene clusters appear only partially related to a reference cluster. In these cases, we have proposed that *C. uncialis* and the reference fungi share two or more common biochemical steps leading to a last common intermediate. Whereas these reference organisms derivatize the last common intermediate to produce azaphilone, cytochalasin, fusarubin, pestheic acid, and mycophenolic acid, *C. uncialis* is hypothesized to produce different and unknown final products. To confirm these predictions, it will be necessary to express these genes and search for *de novo* production of a metabolite in a heterologous host. In the upcoming chapters, I will describe our efforts to explore the lichen secondary metabolome by expressing lichen genes within a heterologous host.

Chapter 7

Preliminary experiments towards heterologous expression trials

7.1. Introduction – Choosing a heterologous host

In Chapter 3, I described how we identified a putative usnic acid gene cluster in *C. uncialis* using a deductive approach. We now set ourselves to the task of providing a definitive assignment of function. The simplest approach in most circumstances would be to knock-out the suspected genes and observe for the cessation of usnic acid biosynthesis in *C. uncialis*. However, as lichens have lifespans often in excess of one thousand years, lichens grow too slowly for gene knockout protocols to be efficiently applied. As the metabolic profile of lichens are also highly sensitive to their environmental milieus (Grube et al. 2014), we were concerned that a gene-knockout approach may produce false-positives. In addition, loss-of-function techniques such as gene knockout would not be useful for exploring the biosynthetic functions of gene clusters that do not appear to be producing any detectable metabolite within *C. uncialis*. We therefore elected to use heterologous expression - a gain-of-function technique that is commonly applied to characterize the function of biosynthetic genes in *Fungi* (Alberti et al. 2016; Luo et al. 2015; Luo et al. 2017). Fungal platforms such as *Aspergillus oryzae* and *Aspergillus nidulans* are available to conduct heterologous expression studies and are routinely used to characterize the biosynthetic functions of PKS and other enzymes (Anyagwu & Mortensen, 2015). In this case, inserting the proposed usnic acid-related genes into a suitable host and observing *de novo* usnic acid biosynthesis within the host would provide definitive proof that these genes are involved in usnic acid biosynthesis in *C. uncialis*. An outline of this experimental scheme is provided in **Figure 60**.

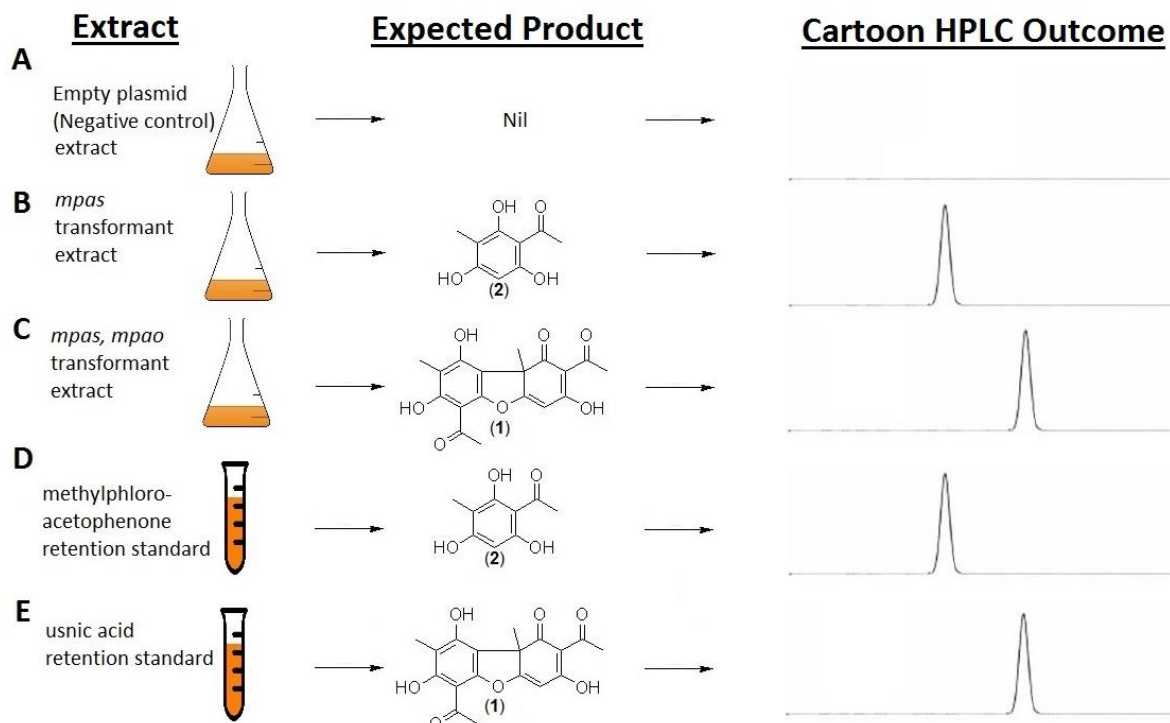


Figure 60: Sketch of experimental scheme. By transforming a host with *mpas* and *mpao*, it is possible to ascertain the biosynthetic functions of MPAS and MPAO by observing the metabolite profile of the transformed organism.

We selected the NSAR1 *Aspergillus oryzae* as the platform for heterologous expression, a choice that is based on abundant examples of its successful use in the functional elucidation of fungal biosynthetic genes including PKS (For recent examples, see: Nofiani et al. 2018; Matsuda et al. 2018; Schor et al. 2018). This strain of *A. oryzae* is a triple auxotroph - It cannot produce methionine, arginine, and adenine. Three plasmids, each encoding a gene complementing one of the three auxotrophic mutations, provides a selection marker for the simultaneous transformation of up to three plasmids into *A. oryzae* (Gomi et al. 1987; Iimura et al. 1987; Jin et al. 2004) (**Figure 61**). These plasmids are chromosomally integrated into *A. oryzae* protoplasts via a polyethylene glycol-based approach (see Chapter 2). An amylase promoter and terminator region flank the site where biosynthetic genes are to be inserted, allowing the induction of expression of the inserted

gene (**Figure 61**). Induction of expression is achieved by incubating transformed *A. oryzae* in the presence of maltose or starch (Tsuchiya et al. 1992).⁶

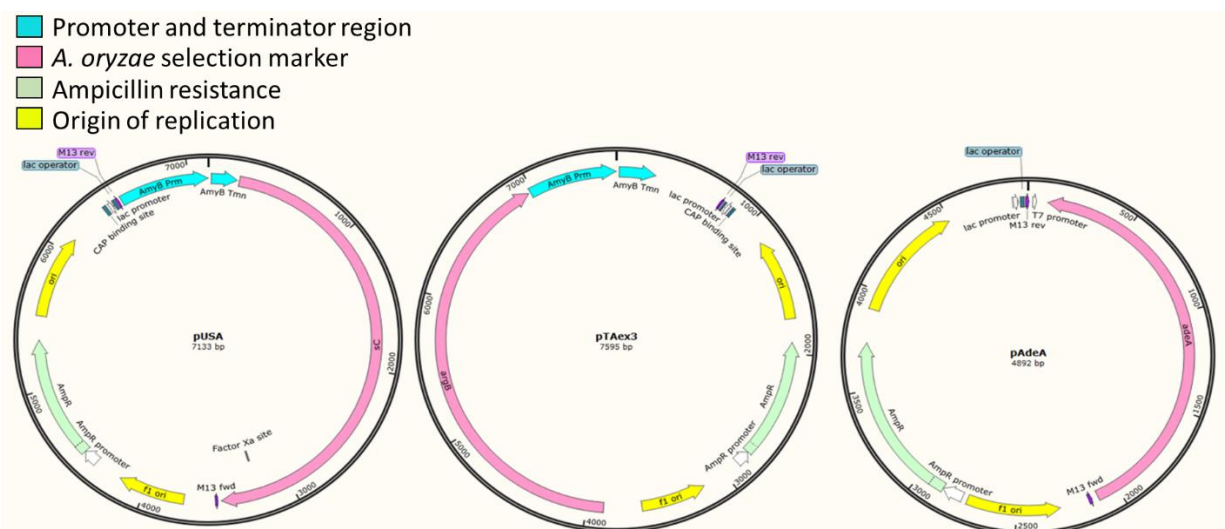


Figure 61: Expression plasmids pUSA (methionine-based selection), pTAex3 (arginine-based selection), and pAdeA (adenine-based selection). Genes providing the *A. oryzae* selection markers are indicated in pink. The amylase promoter and terminator regions, indicated by blue regions, allow the induction of an inserted gene by incubating transformed *A. oryzae* in starch or maltose. An ampicillin resistance gene, indicated in green, allows shuttling of initial plasmid constructs through *E. coli*. The attuned reader may notice that pAdeA lacks a promoter/terminator region: This is pAdeA in its original state as provided to our group by Ikuro Abe and coworkers. A preliminary task was to therefore insert a promoter/terminator region. This was easily accomplished using routine biochemistry.

To confirm that *A. oryzae* was a suitable host for functionally expressing MPAS and MPAO, numerous control experiments were performed. As will be shown later in Chapter 8, our best effort to functionally express MPAS and MPAO in *A. oryzae* was ultimately met with failure. Some of these control experiments were performed before we attempted heterologous expression, whereas others were performed concurrently with the work described in Chapter 8 in an ongoing effort to troubleshoot the problem. These control experiments are described below in order to provide the reader with a detailed understanding of the nature of the problem. These control experiments are not provided in chronological order.

⁶ To familiarize myself with the NSAR1 *Aspergillus oryzae* expression system, I spent two months in the laboratory of Ikuro Abe (Graduate School of Pharmaceutical Sciences, University of Tokyo) wherein I was trained in the use of this expression system. This is the research group that created the NSAR1 system.

7.2. ‘Dry’ control experiments

7.2.1. Introduction

The KS domain catalyses the formation of carbon-carbon bonds and is therefore central to all polyketide chemistry (Keatinge-Clay, 2012). If the KS domains of lichen PKS are unusual (in some respect) as compared to KS domains from *Ascomycota* hosts such as *A. oryzae*, this could explain why difficulties are encountered when attempting to functionally express lichen PKS such as MPAS in *A. oryzae*. For that reason, we decided to test three hypotheses related to evolution, structure, and mechanisms of the KS domains within MPAS (Chapter 3) and 6HMS (Chapter 4).

7.2.2. Phylogenetic analysis of KS domains

Lichens are not merely symbionts of fungi and algae but micro-communities encompassing dozens of cohabitating organisms (Grube et al. 2009; Bates et al. 2011; Spribille et al. 2016). Lichenizing fungi have also displayed a propensity during their evolutionary history for switching algal partners (Piercey-Normore & Depriest, 2001). Although vertical gene transfers predominate the evolutionary history of lichens, horizontal gene transfers have also played a role in shaping the secondary metabolome (Schmitt & Lumbsch, 2009; Tunjić & Korać, 2013; Beck et al. 2015). Although *Cladonia uncialis* is taxonomically classified under phylum *Ascomycota*, it is possible that *mpas* and *6hms* arose through horizontal gene transfer. It is therefore possible that *mpas* and *6hms* do not originate from an *Ascomycota* lineage even though *C. uncialis* is an *Ascomycota* fungus. This distinction, if true, could explain why we have encountered difficulty in expressing MPAS in *A. oryzae*, an *Ascomycota* fungus.

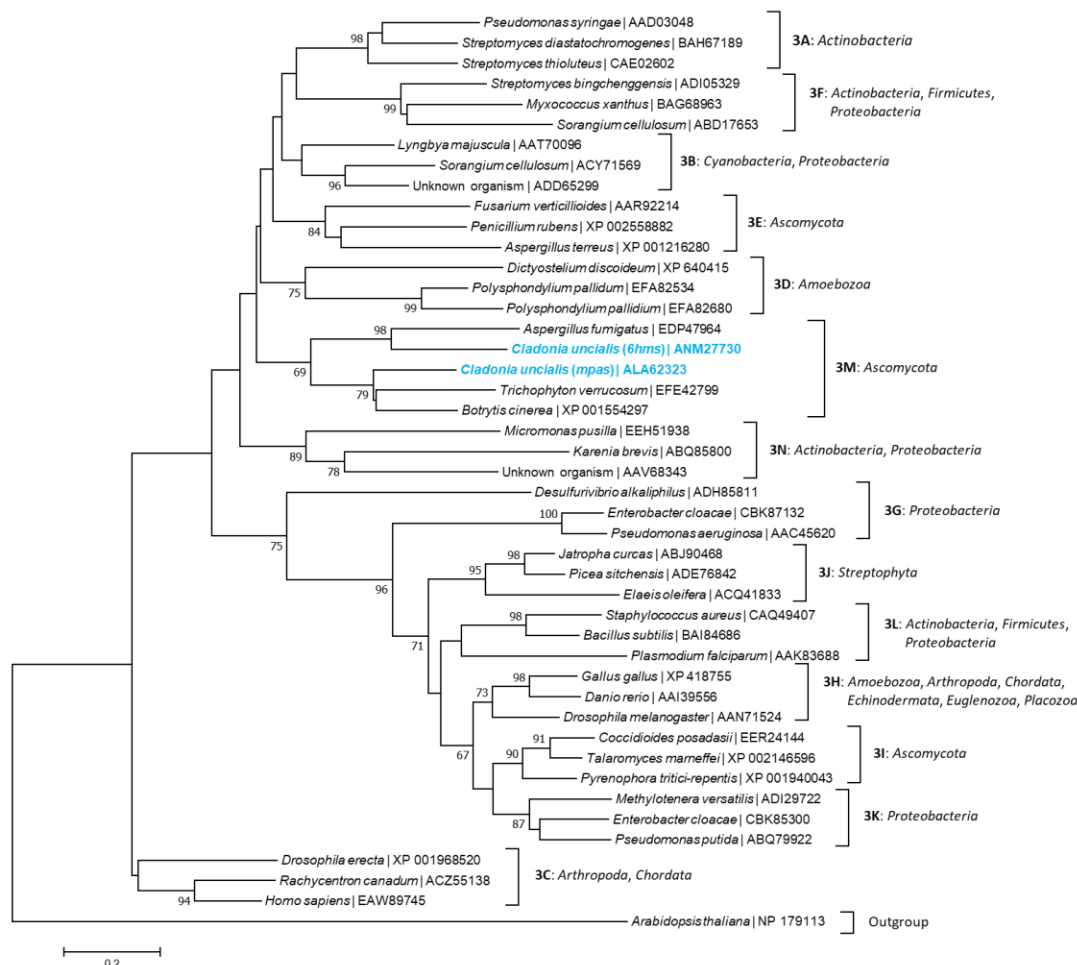


Figure 62: Phylogenetic tree containing KS of methylphloroacetophenone synthase (MPAS), KS of 6-hydroxymellein synthase (6HMS), and three representative KS domains from each of the 14 groups ('3A' to '3N') of the KS3 family of ketosynthases (Chen et al. 2011). A ketosynthase of *Arabidopsis thaliana*, a member of the KS2 family of ketosynthases (Chen et al. 2011), was chosen as an outgroup. Percent bootstrap replications at or above 60% of 1000 replicates are shown next to the branches. Bar denotes number of amino acid substitutions per site. Reproduced from Abdel-Hameed et al. 2018, in accordance with authors' retention of privileges.

The KS is the most conserved domain in PKS and is therefore widely used to study the evolution of PKS in lichenizing fungi (Timsina et al. 2014). We therefore used the KS of *mpas* and *6hms* to test the hypothesis that either PKS has an evolutionary origin beyond *Ascomycota*. The ketosynthase, a broad term for enzymes catalyzing condensations of acyl-coenzyme A and related substrates, have been classified into five families (Chen et al. 2011). The third family, KS3, is composed of modular and iterative PKS as well as fatty acid synthases. The KS3 family is

sub-divided into 14 groups (Named ‘3A’ to ‘3N’) comprising both prokaryotes and eukaryotes (Chen et al. 2011).

To test the hypothesis that these genes have a non-*Ascomycota* origin, three representative KS domains of each of these 14 groups were selected randomly and a phylogenetic tree was constructed using these genes as well as the KS sequences of *mpas* and *6hms* (See Chapter 2). The expected outcome was that the KS of *mpas* and *6hms* would cluster within one of the 14 groups indicating a phylogenetic relationship. The resultant phylogenetic tree revealed that *mpas* and *6hms* both clustered within group ‘3M’, comprising *Ascomycota* (**Figure 62**). These results suggest that *mpas* and *6hms* are indeed *Ascomycota* PKS genes. The hypothesis that these PKS were not *Ascomycota*-inherited genes was therefore invalidated.

7.2.3. Multiple sequence alignment of KS domains

Decarboxylative Claisen condensation by KS is performed by a Cys-His-His catalytic triad. These residues are found within conserved TACSSS, EHGTGT, and KSNIGHH motifs (Xu et al. 2013). Mutational and biochemical analyses indicate that the purpose of histidine (EHGTGT) is to facilitate the transfer of the ACP-polyacyl intermediate to cysteine by acidifying the ACP-thioester intermediate (Robbins et al. 2016). The other histidine (KSNIGHH) supports this transfer by serving as a general base to increase the nucleophilicity of the cysteine thiol (Robbins et al. 2016). These residues also support decarboxylation of methylmalonyl-ACP and C-C bond formation by transition state stabilization and stereoselective orientation of the substrates. The purpose of cysteine (TACSSS) is to provide a nucleophilic anchor for polyketide intermediates (Robbins et al. 2016).



Figure 63: Multiple sequence alignments of ketosynthase domains. Asterisks denote TACSSS, EAHGTG, and KSNIGH conserved motifs. Abbreviations and accession numbers: mPKS | Type I modular PKS 6-deoxyerythronolide B synthase, 2QO3 (Tang et al. 2007); FAS | Iterative fatty acid synthase, 2VZ8 (Maier et al. 2008); T2-PKS | Type II PKS daunorubicin synthase, 5TT4 (Jackson et al. 2018) transPKS | *trans*-AT PKS bacillaene synthase, 4NA1 (Gay et al. 2014). Iterative PKS from selected *Ascomycota* species: AY495646 | HM486910 | KX519720 | JX125042. MPAS | Type I iterative PKS methylphloroacetophenone synthase, ALA62323 (Abdel-Hameed et al. 2016a); 6HMS | Type I iterative PKS 6-hydroxymellein synthase, ANM27730 (Abdel-Hameed et al. 2016b). Reproduced from Abdel-Hameed et al. (2018), in accordance with authors' retainment of privileges.

Though this Cys-His-His triad is highly conserved, it is not absolutely required for polyketide assembly to occur nor are ketosynthases restricted to C-C bond formation as a biological role (Kwon et al. 2002). What residues are present may therefore have broad functional implications. For example, chalcone synthases (Type III PKS) and some fatty acid synthases employ Cys-His-Asn triads (Scarsdale et al. 2001; Morita et al. 2011). Whereas type I PKS recruit acetyl-CoA primer via AT-ACP loading module (modular PKS) or via starter unit-ACP transacylase (iterative PKS) (Wang et al. 2015; Crawford et al. 2006), type II PKS employ a ketosynthase-like protein known as chain length factor (CLF). In CLF, glutamine replaces cysteine and facilitates priming by decarboxylating malonyl-ACP to acetyl-ACP (Bisang et al. 1999). Serine- and tyrosine-containing triads have also been observed in KS-like loading modules servicing modular type I PKS (Aparicio et al. 1999; Brautaset et al. 2000; Tang et al. 2000).

We hypothesized that the KS of MPAS and 6HMS possess atypical catalytic triads. If true, perhaps *Ascomycota* platforms (utilizing Cys-His-His triads) are ill-suited to expressing these

proteins functionally. To test this hypothesis, a multiple sequence alignment was performed by aligning the two lichen KS with four representative examples from PKS of non-lichen *Ascomycota* fungi. Four non-*Ascomycota* KS with resolved crystal structures were also included (Tang et al. 2007; Maier et al. 2008; Jackson et al. 2018; Gay et al. 2014). The expected outcome was that atypical catalytic domains of lichen KS would be clearly reflected through the absence of TACSSSS, EHGTGT, and KSNIGHH motifs within their aligned amino acid sequences.

The produced alignment revealed that both lichen KS domains possess canonical Cys-His-His catalytic triads (**Figure 63**). It is interesting to note that in both lichen KS the conserved serine (S) in the KSNIGH motif is replaced with glycine (G) in MPAS and alanine (A) in 6HMS. This suggests that the replacement of serine with small non-polar residues could be necessary for proper polyketide folding within the catalytic pocket. The observation that both lichen KS possess typical Cys-His-His triads suggests that these lichen KS are likely to function in a similar manner to those of fungi and bacteria. The hypothesis that an atypical active site explains the inability of *Ascomycota* to functionally express lichen PKS was therefore invalidated.

7.2.4. Protein structure modelling of KS domains

Current knowledge on KS structure and function is provided by crystal structures of KS acquired from diverse organisms (Xu et al. 2013). The KS is a homodimer, and this homodimerism maintains the structural coherency of the PKS (Witkowski et al. 2004). Each monomer possesses a thiolase fold, comprised of alternating layers of α -helixes and β -sheets with $\alpha/\beta/\alpha/\beta/\alpha$ -layered architecture (Xu et al. 2013). The dimerization is maintained by hydrogen bonding between two anti-parallel strands at the monomer interface. The KS catalytic triad is located within a deep pocket that must be accessed by the phosphopantetheine arm attached to the ACP (Xu et al. 2013). The ketosynthase influences polyketide chain length and steric constraints provided by the size

and shape of the KS active site pocket partially explain how polyketide chain length is controlled by iterative PKS (Liu et al. 2014; Murphy et al. 2016).

We hypothesized that lichen KS domains possess a tertiary protein structure or catalytic site architecture that is atypical of KS domains found elsewhere in nature. If true, these proteins may prove difficult to translate and fold correctly in *Ascomycota* hosts. To evaluate this hypothesis, Swiss Model (Waterhouse et al. 2018) was used to construct protein models of both lichen KS proteins. Swiss Model is a modelling facility that integrates expert modelling knowledge into an automated process (Waterhouse et al. 2018). The user first inputs the amino acid sequence of the protein he or she wants to model. The program then conducts a template search based on the input data provided by the user. Results are arranged in order of amino acid similarity to the input sequence. The user may then select a template or allow Swiss Model to select a template automatically. Swiss Model then extracts structural information from the template structure in 3D space. Insertions and deletions are resolved by searching a structural database for candidate structures. Model building proceeds using conserved atom coordinates first, followed by resolution of motifs with high flexibility (e.g., loops). Model building concludes with energy minimization (Waterhouse et al. 2018).

Crystal structures of KS were used as a template to model the structure of KS domains from *C. uncialis*. A validation experiment was first conducted to determine whether the modelling program can reproduce a known KS structure. The well-studied KS from 6-deoxyerythronolide B synthase (DEBS) (Tang et al. 2007) was chosen for this validation experiment. The most important factor in the accuracy of homology-based modelling is the use of a reference structure with high amino acid sequence similarity (Fiser et al. 2010). This is because protein structures are more evolutionarily conserved than amino acid sequences. Selecting reference structures with high

amino acid sequence similarity is therefore likely to produce models that are highly accurate. In cases where sequence similarity is exceptional (>50%), only minor errors in side chain packing and rotameric state of individual residues are expected, whereas in cases where similarity is low (<30%), important errors in structure prediction may arise (Fiser et al. 2010). For this reason, the crystal structure of KS from bacillaene synthase (BACS) (Gay et al. 2016) was selected as a reference structure for Swiss Model because KS of BACS possessed reasonable amino acid similarities to the KS of DEBS (39%), MPAS (37%) and 6HMS (34%). The expected outcome was that modelled KS of DEBS would be superimposable with the known crystal structure of DEBS. This was indeed the case (**Figure 64A**). The root-mean-square deviation (RMSD), the average distance between the atoms of the superimposed proteins, was found to be 1.18 Å. When the amino acid similarity is between 30 to 40%, such as in this control experiment (39%), a RMSD error of 1.50 Å or less would be expected (Fiser et al. 2010). An observed RMSD of 1.18 Å between the modeled and known structure of KS of DEBS would therefore suggest that Swiss Model accurately reproduced this KS structure.

We then modelled the KS of MPAS and 6HMS, again using the KS of BACS (Gay et al. 2016) as the reference structure. The modelled lichen KS were then superimposed with the crystal structure of KS of DEBS (Tang et al. 2007). The expected outcome was that if either lichen KS possessed unusual structural features this would be reflected in the superimposed structures. These overlays revealed that modelled lichen KS possess tertiary architecture consistent with KS domains characterized in other organisms (**Figures 62B, 62C**). Consistencies include the presence of thiolase-like $\alpha/\beta/\alpha/\beta/\alpha$ folds and two antiparallel β -sheets within the monomer interface. Whereas the tertiary protein structure of KS of MPAS aligned remarkably well with that of DEBS (RMSD of 1.29 Å) (**Figure 64B**), the KS of 6HMS was imperfectly superimposable (RMSD of

1.93 Å) (**Figure 64C**). However, all expected tertiary structural elements were present in both lichen KS. The imperfect alignment of the latter KS did not result in imperfect alignment of the active site residues of the three proteins. As shown in **Figure 64D**, an overlay of three KS revealed superb alignment of the Cys-His-His catalytic triad. This alignment suggests that the position of active site residues have been optimized by evolution and that lichen KS use a polyketide elongation mechanism identical to other KS found in nature (Keatinge-Clay, 2012). We therefore concluded that the structure of lichen KS are similar to those of non-lichen KS. The hypothesis that lichen KS possess unusual architectural features contributing to lack of functional activity in *Ascomycota* was therefore invalidated.

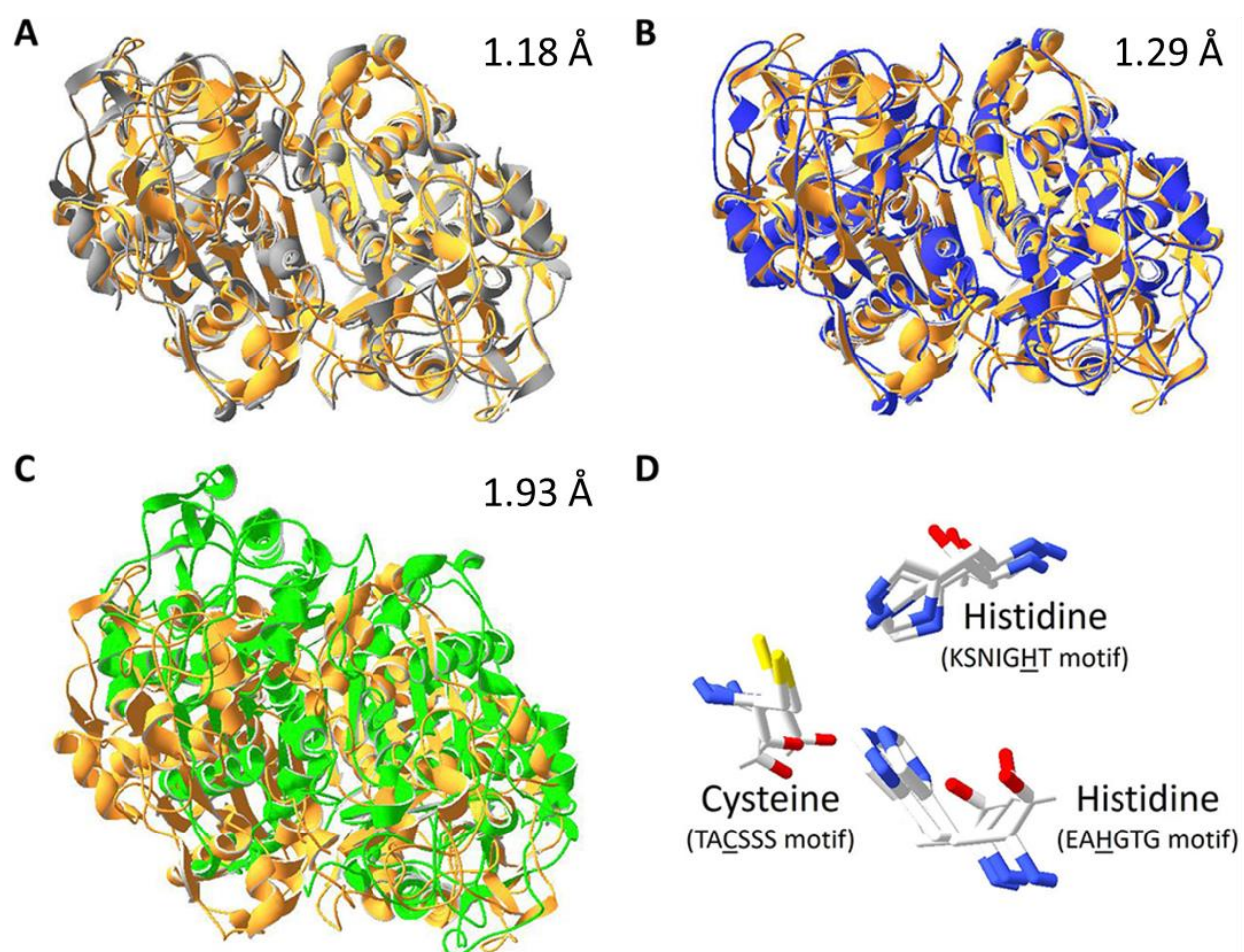


Figure 64: **A** Superimposed crystal structure of DEBS KS (Tang et al. 2007) (gold) and modelled DEBS KS (gray); **B** Superimposed crystal structure of KS of DEBS (Tang et al. 2007) (gold) and modelled KS of MPAS (Abdel-Hameed

et al. 2016a) (blue); **C** Superimposed crystal structure of KS of DEBS (Tang et al. 2007) (gold) and modelled KS of 6HMS (Abdel-Hameed et al. 2016b) (green); **D** Superimposed Cys-His-His active sites of crystal structure of KS of DEBS (Tang et al. 2007) with modelled KS of MPAS (Abdel-Hameed et al. 2016a) and modelled KS of 6HMS (Abdel-Hameed et al. 2016b). The root-mean-square deviation (RMSD) of all superimposed atoms are indicated in the top right of each panel and are provided in units of Å. Adapted from Abdel-Hameed et al. (2018), in accordance with authors' retainment of privileges.

In conclusion, these experiments were unable to rule-out any evolutionary, structural, or mechanistic abnormalities in the KS domains of MPAS and 6HMS that would be indicative of a cause for concern regarding the potential for the functional heterologous expression of these enzymes within *Ascomycota* hosts such as *A. oryzae*.

7.3. 'Wet' control experiments

7.3.1. Introduction

To rule out possible issues related to host compatibility, several hypotheses were examined related to metabolite toxicity, transcription, and optimization of the experimental time-course.

7.3.2. Testing the selection markers of NSAR1 *A. oryzae*

To confirm that the fungus provided to us was indeed a triple auxotroph of arginine, adenine, and methionine, we prepared minimal media plates containing these selection markers, and inoculated *A. oryzae* onto each plate. As shown in **Figure 65**, the fungus may only grow when all three selection markers are provided.

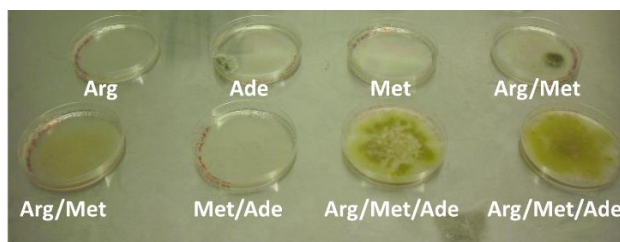


Figure 65: Control experiment confirming triple auxotrophy of NSAR1 *A. oryzae*. In this experiment, two contaminating colonies of fungi, not *A. oryzae*, were also observed on the 'Ade' and 'Arg/Met' plates.

7.3.3. Testing toxicity of methylphloroacetophenone and usnic acid on *A. oryzae*

We evaluated the toxicity of methylphloroacetophenone and usnic acid on *A. oryzae* to verify that these molecules do not adversely affect the fungus. A chemical standard of usnic acid was acquired from a commercial source (Abcam Biochemicals), and methylphloroacetophenone was prepared by a co-worker (see Acknowledgements). DPY agar (see Chapter 2) was prepared containing usnic acid or methylphloroacetophenone to a final concentration of 1 g/L or 10 g/L. These concentrations are well-above the expected outcome of the heterologous expression of a PKS in *A. oryzae* without tooling the titre using metabolic engineering. Spores of *A. oryzae* were inoculated onto the plate via sterile toothpick and allowed to grow for several days. As compared to an untreated control, no difference in the morphology, colour, germination time, or growth rate was observed in response to usnic acid (**Figure 66**). The germination of *A. oryzae* exposed to 1 g/L of methylphloroacetophenone was delayed by approximately one day. No other adverse effects were observed in response to methylphloroacetophenone as *A. oryzae* otherwise appeared to grow normally. We therefore proceeded to insert *mpas* and *mpao* into the transformation plasmids and chromosomally integrate these genes into *A. oryzae* (see Chapter 2).

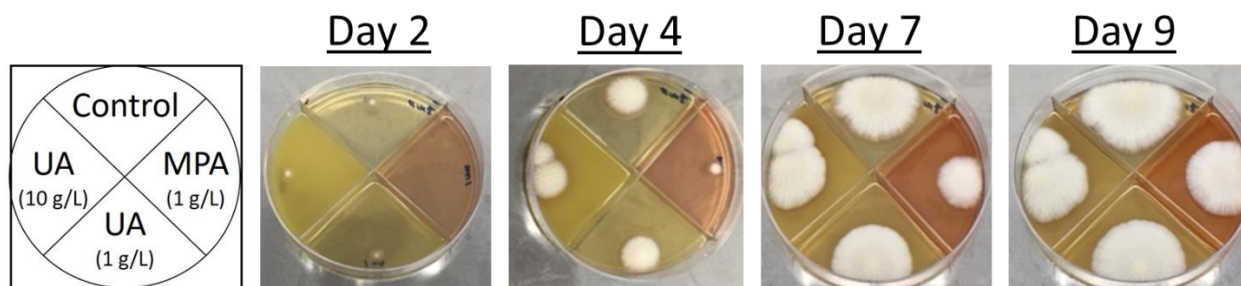


Figure 66: Control experiment testing toxicity of methylphloroacetophenone (1 g/L) and usnic acid (1 g/L and 10 g/L) on *A. oryzae* growing on DPY agar. Adapted from Bertrand et al. (2019), manuscript in preparation.

7.3.4. Testing heterologous transcription of *mpas* and *mpao* in *A. oryzae*

The *mpas* and *mpao* genes were loaded into transformation plasmids, and NSAR1 *A. oryzae* was then transformed with these genes (see Chapter 2). Transformants of *A. oryzae* were

then incubated in Czapek-Dox media containing starch (see Chapter 2). To determine whether transcription of *mpas* and *mpao* was occurring within *A. oryzae*, we isolated total mRNA and evaluated transcription by RT-PCR using primers targeting *mpao* and *mpas* (see Chapter 2). As the amplification of long genes is challenging, we amplified *mpas* cDNA in four portions (“cDNA”). Amplifications were also performed on the mRNA extract before and after DNase application (“Before DNase” and “After DNase”), the transformation plasmid (“+ PCR ctrl”), and the PCR reagents (“- PCR ctrl”). The absence of an amplification product arising from template that was treated with DNase signifies that all contaminating genomic DNA was denatured preceding the reverse transcription of mRNA into cDNA. The absence of an amplification product in the negative control verifies the absence of cross-contaminating genomic DNA within the PCR reagents. A scheme of this procedure and its possible outcomes are provided in **Figure 67**.

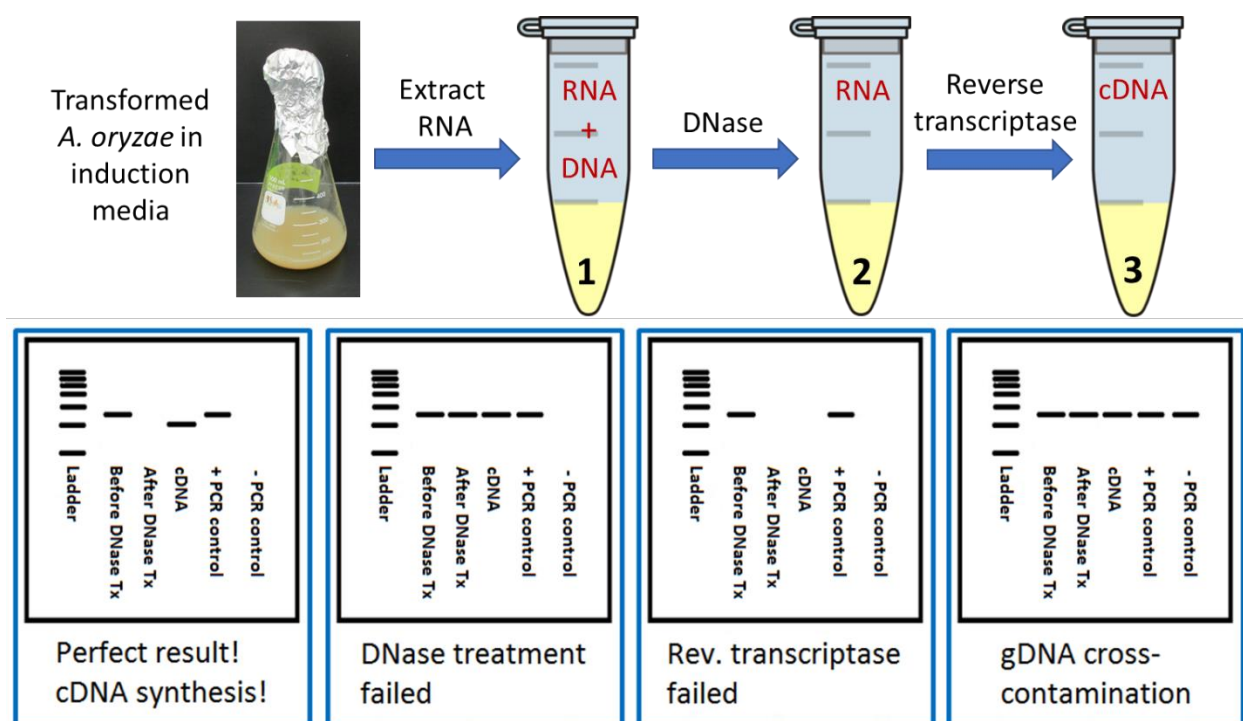


Figure 67: Outline of procedure for testing heterologous transcription of lichen genes in *A. oryzae*, and interpretation of results by agarose electrophoresis. If numerous introns were removed to produce the mature mRNA, the cDNA amplicon will be a different size and may be observed to migrate slightly farther down the agarose gel.

As shown in **Figure 68**, cDNA of *mpas* and *mpao* was successfully generated from mRNA isolated from *A. oryzae*. Sequencing these amplicons and comparing the DNA sequences to the *C. uncialis* genome revealed that four introns were removed from *mpas* and that two introns were removed from *mpao* (**Figure 68**). When the translation of the exonic DNA was simulated by an online program (ExPASy translate tool), a coherent reading frame without in-frame stop codons was produced (not shown). This suggests that *A. oryzae* can transcribe lichen genes and remove introns to produce translation-coherent mature mRNA.

To verify that the mature mRNA of these two lichen genes in *A. oryzae* was prepared in a manner that is identical to the native organism, we repeated this experiment by isolating the mRNA from *C. uncialis* and amplifying *mpas* and *mpao* using the identical RT-PCR procedure (see Chapter 2). As shown in **Figure 68**, cDNA of *mpas* and *mpao* was successfully created from mRNA isolated from *C. uncialis*. The number and positions of the introns were observed to be identical between *C. uncialis* and *A. oryzae* (**Figure 69**). The gDNA sequence of *mpas* and *mpao* (including exon/intron boundaries) is provided as **Sequence S1** and **S2** in the Appendix. In summary, *A. oryzae* can transcribe lichen genes, remove introns to produce translationally-coherent mature mRNA, and do so in an identical manner to *C. uncialis*.

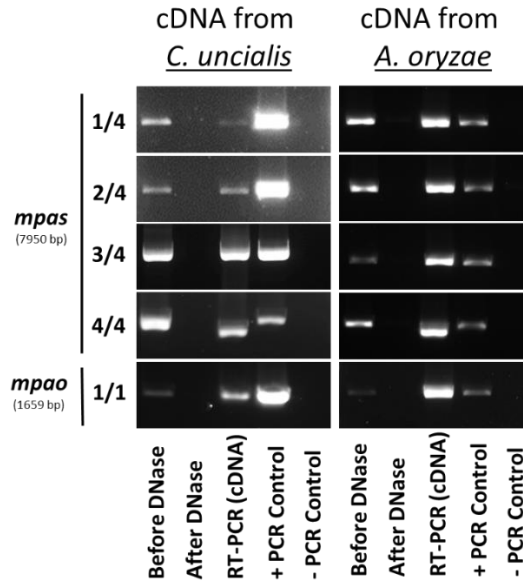


Figure 68: Agarose gels demonstrating (A) transcription of *mpas* and *mpao* in *A. oryzae* (B) transcription of *mpas* and *mpao* in *C. uncialis*. Adapted from Bertrand et al. (2019), manuscript under preparation. Uncropped agarose gels of this figure have been made available as **Figure S52** in Appendix.

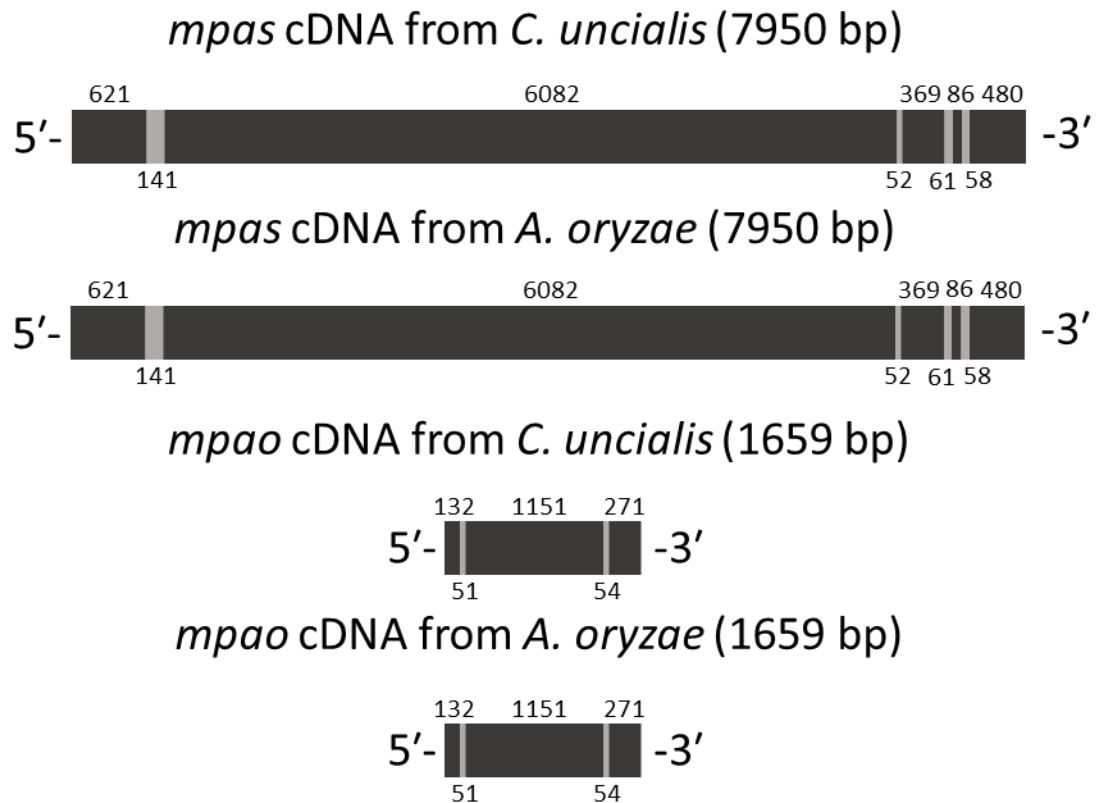


Figure 69: Schematic indicating location of exons (black) and introns (grey) in the genomic DNA of *mpas* and *mpao*, as revealed by the transcription of these genes in *A. oryzae* and *C. uncialis*. Adapted from Bertrand et al. (2019), manuscript under preparation.

7.3.5. Choosing an incubation end-point

To screen *A. oryzae* transformed with *mpas* and *mpao*, it will be necessary to grow the transformed *A. oryzae* for several days within induction media. The metabolites within the media will then need to be extracted to determine whether MPAS and MPAO are actively producing methylphloroacetophenone and/or usnic acid within *A. oryzae*. Incubation periods ranging between 3 to 10 days, and temperatures ranging between 28°C to 30°C, have been reported in the literature (He & Cox, 2016; Fujii et al. 2011; Bailey et al. 2016; Matsuda et al. 2016; Matsuda et al. 2015; Matsuda et al. 2013). We elected to use a five-day procedure at 30°C to remain consistent with the literature (**Figure 70**). As will be shown later in this chapter, this procedure was proven to be remarkably effective for producing orsellinic acid and 6-methylsaliylic acid from two non-lichen PKS that were transformed into *A. oryzae*.

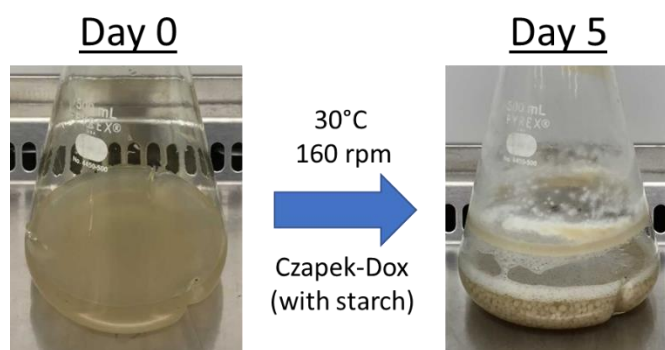


Figure 70: The incubation procedure that was adopted for heterologous expression trials. A total of 100 mL of starch-based Czapek-Dox media was placed into a 500 mL Baffled Erlenmeyer flask. Spores of transformed *A. oryzae* were then inoculated into culture using a sterile toothpick. The culture was incubated for five days on a shaking incubator at 30°C and 160 rpm.

7.3.6. Mock extractions

To confirm that the proposed extraction procedure was appropriate for the isolation of methylphloroacetophenone and usnic acid from media, we conducted a mock extraction trial by spiking both metabolites into Czapek-dox media. This media was then incubated with no other additives, or with *A. oryzae*, for a period of five days, then extracted via an ethyl acetate-based

approach (see Chapter 2). Usnic acid was observed in HPLC traces in extracts of both experiments (**Figure 71**). Our laboratory has long-observed that methylphloroacetophenone is an oxidatively unstable molecule. In this mock trial, we were only able to recover this molecule in extracts of the fungal-free media (**Figure 71**). For this reason, we decided to always co-transform *mpas* with *mpao* so that usnic acid could be immediately produced *in situ*. Nonetheless, the instability of methylphloroacetophenone lends concern to the possibility that this metabolite could be degrading *in situ* before it is dimerized by MPAO into usnic acid.

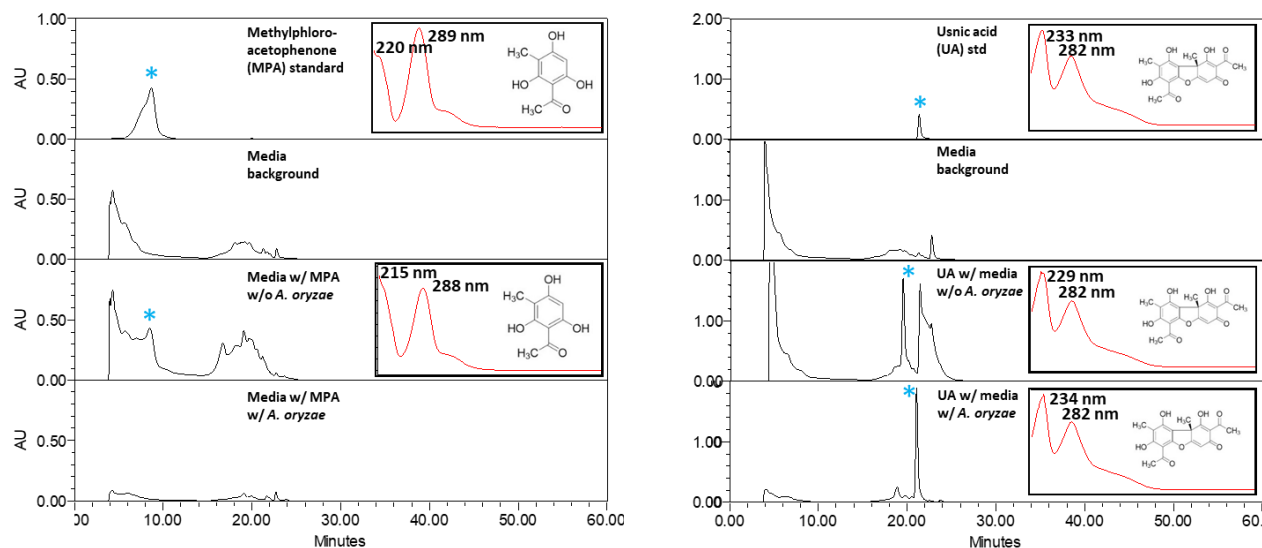


Figure 71: Mock extraction trials of methylphloroacetophenone and usnic acid in media with and without NSAR1 *A. oryzae*. Traces are displayed at 280 nm.

7.3.7. Functional expression of a characterized non-lichen PKS in *A. oryzae*

To verify that we can correctly use the NSAR1 expression system and produce *de novo* metabolite from an exogenous gene, we tested the system using a known PKS gene. Our collaborators (Ikuro Abe *et al.*, University of Tokyo) recently used *A. oryzae* to characterize an orsellinic acid synthase (*fu-oas*) from *Fusarium sp.*, a non-lichen fungus (Okada et al. 2017). Our collaborators kindly provided us with a transformation plasmid containing *fu-oas*. If we were to transform *A. oryzae* with this plasmid containing *fu-oas* and demonstrate *de novo* orsellinic acid

production in *A. oryzae*, this would demonstrate that we can correctly use the *A. oryzae* platform. We elected to grow transformants of *fu-oas* for five days in Czapek-Dox media at 30°C. The media was then extracted using ethyl acetate. These extracts were then run on HPLC (see Chapter 2). The expected outcome of this experiment was that orsellinic acid would be observable within the HPLC traces.

As shown in the resultant HPLC trace, transformation of *fu-oas* in *A. oryzae* immediately resulted in copious production of orsellinic acid (**Figure 72**). The relevant HPLC signal, indicated by an arrow, possesses a retention time and absorbance spectrum that is identical to the chemical standard of orsellinic acid. To exemplify the efficacy of the NSAR1 system, to produce the results shown in **Figure 72** it was first necessary to dilute this extract approximately 10-fold as the undiluted extract was saturating the HPLC detector. This experiment demonstrates that we can correctly use the NSAR1 system to produce *de novo* metabolite from a transformed PKS.

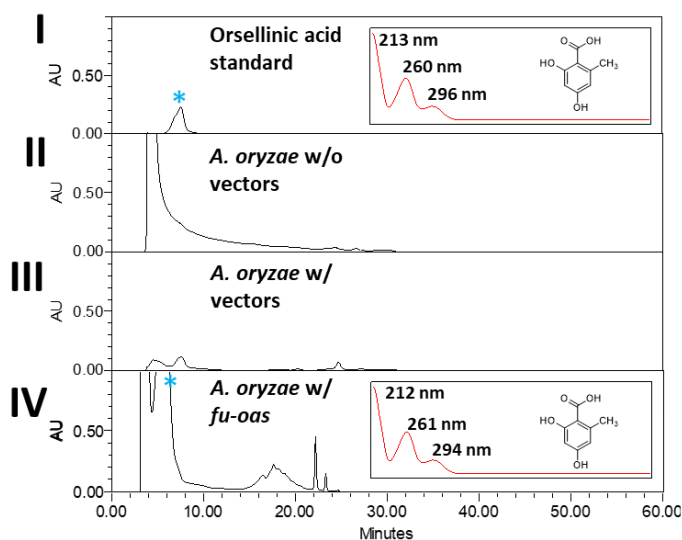


Figure 72: HPLC trace of *A. oryzae* transformed with *fu-oas* demonstrating *de novo* orsellinic acid biosynthesis in this fungus. (I) Chemical standard of orsellinic acid (II) Native NSAR1 grown in media (III) NSAR1 transformed with plasmids that do not contain genes (IV) NSAR1 transformed with *fu-oas*-containing plasmid. Traces shown are at 260 nm. Adapted from Bertrand et al. (2019), manuscript in preparation.

7.3.8. Functional expression of an uncharacterized non-lichen PKS in *A. oryzae*

We were also interested in functionally expressing an experimentally uncharacterized PKS from a non-lichenizing species of fungi. The intent of this experiment was to test my ability to correctly follow-through with the complete transformation protocol from start to finish. This would begin with the amplification of a PKS gene from the genome of a fungus. The gene would then need to be correctly inserted into an expression plasmid. The *A. oryzae* would then be transformed, and the gene induced under appropriate conditions to functionally express the PKS. A successful conclusion of this experiment would conclusively demonstrate our expertise with this expression system.

We selected 6-methylsalicylic acid for this experiment. 6-methylsalicylic acid is a polyketide produced by 6-methylsalicylic acid synthase (Bedford et al. 1995). This polyketide is the first intermediate in the biosynthesis of patulin, a mycotoxin of paramount importance to the history and development of the natural products sciences (Birch & Donovan, 1953; Birch et al. 1955). We were kindly provided with a patulin-producing fungus which was identified by our collaborators as *Penicillium griseofulvum* (George Hausner, Department of Microbiology, University of Manitoba) (Engelberg et al. 2018). To confirm the species of this fungus, we isolated its gDNA and amplified its internal transcribed spacer (ITS), a region of DNA encoding nuclear ribosomal RNA that is widely used as a ‘DNA barcode’ for fungi (Schoch et al. 2012). We sequenced the ITS amplicon and compared it to ITS sequences deposited in GenBank (see **Sequence S3** in Appendix). We identified this fungus as a member of the *Penicillium* genus, a result was consistent with our collaborators’ prior observation (**Table S54** in Appendix). Using known 6MSAS genes such as *6msas* from *P. patulum*, we designed primers to target 6MSAS genes in our *Penicillium* specimen. A single amplification product was produced. DNA sequencing and

BLAST analysis (Altschul et al. 1990) revealed this gene to be similar to two 6MSAS genes deposited in GenBank originating from *P. patulum* or *P. griseofulvum* (Beck et al. 1990; Dombink-Kurtzman, 2009) (**Table S55** in Appendix). We named this gene *pe-6msas* and prepared it for transformation into *A. oryzae*.

The *pe-6msas* was transformed into *A. oryzae*, and the transformed fungus was grown in media for five days at 30°C on a shaking incubator. The media was then extracted and run on HPLC. A commercial standard of 6-methylsalicylic acid (*Accela*) was acquired to assist us in interpreting the results. The resultant HPLC trace demonstrates copious 6-methylsalicylic acid production as a result of this experiment (**Figure 73**). As was the case of *fu-oas* in *A. oryzae*, to produce the trace shown in **Figure 73** it was first necessary to dilute the extract about 10-fold as the undiluted extract was saturating the HPLC detector. This experiment demonstrates that we can correctly carry-out the transformation procedure from start to finish to functionally express an exogenous PKS within a heterologous host.

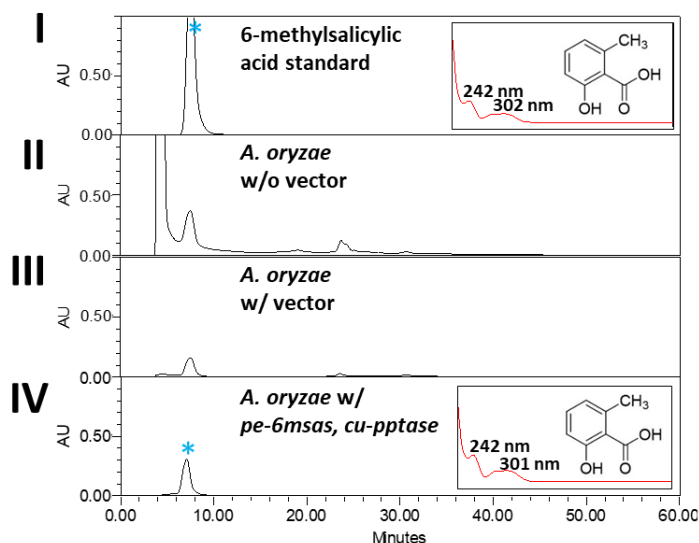


Figure 73: HPLC trace of *A. oryzae* transformed with *pe-6msas* demonstrating *de novo* 6-methylsalicylic acid synthase in this fungus (I) Chemical standard of 6-methylsalicylic acid (II) Native NSAR1 grown in media (III) NSAR1 transformed with plasmids that do not contain genes (IV) NSAR1 transformed with *pe-6msas*-containing plasmid. Traces are shown at 300 nm. Adapted from Bertrand et al. (2019), manuscript in preparation.

7.3.9. Testing heterologous translation of *mpas* in *A. oryzae*

The final preliminary experiment was to test whether heterologous translation of *mpas* into MPAS was occurring in *A. oryzae*. We explored this problem by attaching *N*-terminal His(6)-tags to genes by PCR using sense primers encoding a His(6)-tag (see Chapter 2). DNA sequencing confirmed the amplification of the His(6)-tagged variants of genes constructed (data not shown). The intent of this experiment was to screen for heterologous protein production in extracts of transformed *A. oryzae* using a His(6)-antibody and Western Blot. To provide a positive control of subsequent analyses, His(6)-tags were also attached to *fu-oas* from *Fusarium sp.* (*fu-oas*) and *pe-6msas* from *Penicillium sp.* The His(6)-tagged versions of *fu-oas* and *pe-6msas* were then re-transformed into *A. oryzae*. To confirm that the tagged PKS were functional, we screened transformants for orsellinic acid and 6-methylsalicylic acid biosynthesis using a screening process that was identical to the process described in previous sections. 6-methylsalicylic acid production was observed from His(6)-tagged 6MSAS, demonstrating functional heterologous expression of the His(6)-tagged PKS (data not shown). However, orsellinic acid production was not observed from His(6)-tagged OAS (data not shown). We speculate that the His(6)-tag interfered with the function of FU-OAS. The His(6)-tagged version of *mpas* from *C. uncialis* was also prepared, then transformed into *A. oryzae*. Total protein extract of the 6-methylsalicylic acid-producing transformant of *A. oryzae*, and total protein extracts of five transformants of *A. oryzae* containing His(6)-tagged *mpas*, were acquired. The concentration of protein extracts was quantified using the Bradford assay (Bradford, 1976) then electrophoresed and blotted (see Chapter 2). A Coomassie stain, applied to the membrane after blot imaging, confirmed that high molecular weight proteins were effectively transferred onto the membrane and that a suitable concentration of total protein extract was loaded into the SDSPAGE gel (**Figures S50 and S51** in Appendix).

The detection of the positive antibody control within the molecular weight ladder confirmed that the blotting procedure was effective (**Figure 74A**). No His(6)-tagged MPAS was detectable from any of the five transformants of *A. oryzae* containing this gene. This would suggest that the lichen protein was not translated in *A. oryzae*. However, the His(6)-6MSAS that was constructed as an internal control for this experiment also remained undetected by Western Blot (**Figure 74A**). This was a surprising result because we had observed copious 6-methylsalicylic acid production in extracts of *A. oryzae*. Although the quantity of PE-6SMAS in *A. oryzae* is unknown, we know that the enzyme must be present because we can observe its metabolic product. The His(6)-6MSAS was similarly undetectable in Western blots of the insoluble lipid fractions that were processed during the protein isolation (not shown).

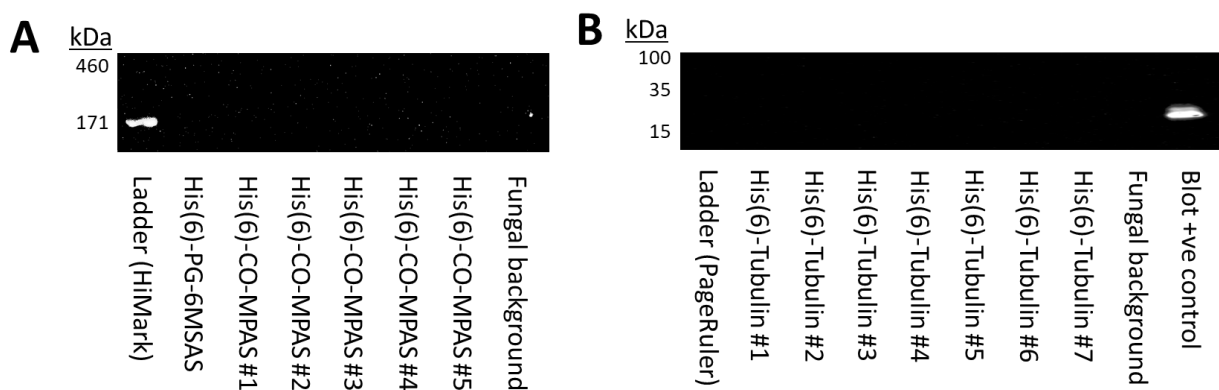


Figure 74: Western blots of protein extracts of *A. oryzae* using a polyhistidine antibody. (A) Western blot of protein extracts of *A. oryzae* transformed with plasmids encoding His(6)-tagged variants of *6msas* from *Penicillium sp.* and of *mpas* from *C. uncialis*. The fungal culture from which the proteins were extracted was confirmed by HPLC to produce copious 6-methylsalicylic acid (data not shown). ‘Fungal background’ is *A. oryzae* transformed with empty vectors. The 171 kDa protein in the HiMark ladder is responsive to His(6) antibody and is used as a blotting control. (B) Western blot of protein extracts of *A. oryzae* transformed with plasmids encoding His(6)-tagged β -tubulin from *A. oryzae*. ‘Fungal background’ is *A. oryzae* transformed with empty vectors. ‘Blot +ve control’ is a crude protein extract from *E. coli* containing a His(6)-tagged protein with a molecular weight of approximately 25 kDa.

To attempt to generate an alternative positive control, we amplified the β -tubulin gene from *A. oryzae* and attached an *N*-terminal His(6) tag using the appropriate sense primer (see Chapter 2). DNA sequencing confirmed the amplification of the His(6)-tagged variant of the β -tubulin

gene (data not shown). This His(6)-tagged tubulin gene from *A. oryzae* was then transformed into *A. oryzae*, with the expectation of observing (in this case, homologous) expression of this protein. A total of seven transformants of *A. oryzae* containing this gene were grown in Czapek-Dox media containing starch, and whole-protein was extracted from these samples. Though the blot control indicated successful blotting, we could not detect His(6)-tagged β -tubulin in *A. oryzae* (**Figure 74B**). In summary, we were unable to use His(6)-tagged variants of PE-6MSAS nor β -tubulin as internal controls. Although it is possible that lichen proteins are not being translated in *A. oryzae*, the failure to establish a reliable internal control renders any conclusion on this matter untenable.

7.4. Summary

The NSAR1 *Aspergillus oryzae* expression system was chosen for heterologous expression experiments on the usnic acid gene cluster. Control experiments indicate that this species of *Ascomycota* appears to be a suitable host for the expression of *Ascomycota* PKS such as *mpas* and *6hms*. Control experiments also indicate that neither methylphloroacetophenone nor usnic acid is toxic to *A. oryzae*, that the extraction protocol (see Chapter 2) is suitable for extracting these metabolites from media, that *A. oryzae* can transcribe these genes and correctly remove introns, and that we are competent with the full range of procedures required of the expression system. Experiments attempting to demonstrate translation of lichen proteins within *A. oryzae* remain inconclusive.

Chapter 8

Heterologous expression trials

8.1. Introduction – heterologous expression trials

In this chapter, I will describe our heterologous expression trials on the usnic acid gene cluster. Thereafter, I will also describe heterologous expression trials on other genes from *C. uncialis*. Regrettably, these experiments are marred by frustrations. In Chapter 9, I will offer some commentary on possible ways to overcome this conundrum and advance our exploration of the lichen secondary metabolome.

8.2. Heterologous expression trials of the usnic acid gene cluster (MPAS and MPao)

8.2.1. Trials with native mpas and mpao

Aspergillus oryzae transformed with *mpas* and *mpao* was grown in starch-based Czapek-Dox media and incubated on a shaking incubator at 30°C for five days. The media was extracted with ethyl acetate and the extracts were analyzed on HPLC (see Chapter 2). Neither methylphloroacetophenone nor usnic acid were observed in *A. oryzae* (Data not shown).⁷ Aside from starch, research groups using *A. oryzae* have also reported the successful expression of fungal genes using maltose-based media or DPY media (He & Cox, 2016; Bailey et al. 2016; Matsuda et al. 2016; Okada et al. 2017). Using DPY media or Czapek-Dox with maltose did not work (data not shown). We therefore began to explore other hypotheses to explain why we were unable to detect lichen secondary metabolites in *A. oryzae*.

⁷ Due to a computer malfunction with a laboratory desktop computer, these first HPLC traces were irrecoverably lost. As I am claiming that functional heterologous expression was *not* observed in these experiments, I hope that the absence of data will not endanger the reader's acceptance of this claim.

8.2.2. Trials with codon-optimized *mpas* and *mpao*

With the exception of methionine and tryptophan, the 20 canonical amino acids are encoded by at least two synonymous codons. Enumeration of the codon composition of protein-encoding genes has revealed that organisms possess biases with respect to which codon within a synonymous set of codons is preferentially used to encode an amino acid. These biases reflect the bioavailability of complementary tRNA (Quax et al. 2015; Hanson & Collier, 2018). Codon biases are present in PKS genes of *C. uncialis* (Bertrand et al. 2015). As different organisms may have different codon biases, problems arise during heterologous expression trials when a gene from one organism has codon biases that are incongruous with the tRNA pools of the host organism. This problem can be overcome by chemically synthesizing a gene to contain synonymous codons that are optimized for translation within the host. Codon optimization is widely used to enhance the expression of heterologous genes (Elena et al. 2014; Webster et al. 2016; Tian et al. 2017).

Score	Expect	Identities	Gaps	Strand
5505 bits(6104)	0.0	5796/7625(76%)	0/7625(0%)	Plus/Plus
Query 1	ATGGCTCTGCCCTCTCTGATCGCTTTTCGGCGCTCTCGCTCCCTGGCCGCTTCTGATCGC	60		
Sbjct 1	ATGGCATTACCTTCGCTGATCGCATTTCGGCGCTTTAGCAGCTTGGCCAGCATCAGATCGC	60		
Query 61	CTCGATCAGCTCCGCAATGCCCCCAGCATCACAACTCCCTGAAGCCCAACCAAGGCC	120		
Sbjct 61	CTGGATCAGCTTCGAAACGCTCTGAGCAGCCCAACTCCTTGAACCAATCACAAAGGCA	120		
Query 121	ATCCAGGAGCTGCCTCTCTCTGTGGAAGGCCCTCTCCAGCAGGACCACTCTCTGCACCT	180		
Sbjct 121	ATTGAGGAGCTTCGGTTGCTCTGGAAGGCTCTTCCAGCAGGATCAGAGTTTACATAGC	180		
Query 181	ATCGCCGCGAGGCTCCGCTGATCAGCTGGCTCAGTGGATCTCTGGCGCTGGTACCGCC	240		
Sbjct 181	ATCGCCGCGAGGCGGCTGAGACCACTTAGCGCAATGGATAGCGGAGCTGGCAGCGCT	240		
Query 241	CAGCTCGTCGATGATAAGGGCAATGTGACCGCATGCCTCTCACCACCATCGCCAGATC	300		
Sbjct 241	CAGCTTGTGGACGACAGGGTATGTGCAAGAAATGCCCTTAACAACCATGCGCAGATT	300		
Query 301	GCCAGTATGTCTTATCTCTGCCAGTACGAGGAGCCCTCCGCCAGAGTCCATCATC	360		
Sbjct 301	GCGCAATATGTTAGCTATCTGTGTAATGAGGAGCCCTTCGTACGAAATCAATCATA	360		
Query 361	AAGTCTGCTGCCATCGGCGGCGGTATCCAGGGCTTCTGTATCGGCTCTCTCCGCTCTG	420		
Sbjct 361	AAAAGCGCTGCGATCGGCGGAGGTATCCAAGGGTTTGTATTGGCTCTTAAGCGCTCTT	420		
Query 421	GCCGTGCTTCTGGCAAGACCGAGGATGACGTGGGCAATTCGCCGCCAATGTCGGTCCGC	480		
Sbjct 421	GCASTGGCAGCGGAAAGACGGAAGACGATGTTGCAATTCGCACTATGTCCGTGCGA	480		
Query 481	CTCGCTTCTGTGTGGGCGCTTACGTGACCTCGATCGCCACGCAACGGCGGTGATTCC	540		
Sbjct 481	TTAGCCTTTTCGCTCGGGGCTACGTTGATCTTGATCGACATCGCAACGGTGGCGACTCC	540		

Figure 75: An excerpt of a two-gene nBLAST alignment between the native *mpas* (query sequence) and the codon-optimized *mpas* (subject sequence). Only a portion of the alignment is shown. This process altered 24 % of the nucleotides in *mpas* and 23 % of the nucleotides in *mpao*.

We therefore hypothesized that the codon usage biases between *C. uncialis* and *A. oryzae* were too divergent to effectively translate MPAS and MPAO. We therefore commissioned codon-

optimized variants of *mpas* and *mpao* for expression in *A. oryzae*. A portion of the alignment between the native and optimized versions of *mpas* is provided to illustrate the extent of the nucleotide changes provided by this process (**Figure 75**). These codon-optimized sequences are available as **Sequences S4** and **S5** in the Appendix. These optimized genes were inserted into *A. oryzae* with the expectation of *de novo* metabolite biosynthesis. No new metabolites were observed in *A. oryzae* (data not shown). We therefore sought other explanations.

8.2.3. Trials using phosphopantetheinyl transferase (PPTase)

Phosphopantetheine, a molecule derived from Coenzyme A, is a prosthetic group that is required for the operation of acyl carrier protein (ACP) domains of polyketide synthases (PKS), fatty acid synthases (FAS), and non-ribosomal peptide synthetases (NRPS). This post-translational modification is facilitated by a phosphopantetheinyl transferase (PPTase) (Beld et al. 2013). Although PPTase are often promiscuous, some PPTase display specificity with respect to the activation of ACP domains, thereby complicating the functional heterologous expression of PKS, FAS, and NRPS. Co-expressing the native PPTase (or the PPTase of a closely related organism) within the heterologous host has been observed to activate or increase metabolite turnover, particularly in cases where the heterologous host is a yeast or bacterium (Ku et al. 1997; Kealey et al. 1998; Wattanachaisaereekul et al. 2007; Gidijala et al. 2009; Rugbjerg et al. 2013).

We were therefore interested in testing whether co-expression of the *C. uncialis* PPTase with MPAS and MPAO would activate metabolite production in *A. oryzae*. We used gene sequences of experimentally-characterized PPTase deposited in GenBank to perform a BLAST survey (Altschul et al. 1990) on our draft genome of *C. uncialis*. A single gene with consensus similarity to PPTase was identified by this approach. This gene was amplified by PCR and placed into a transformation plasmid. The codon-optimized versions of *mpas* and *mpao*, as well as the

native *pptase* from *C. uncialis*, were chromosomally integrated into *A. oryzae*. Isolation of mRNA of transformed *A. oryzae* grown on starch media and RT-PCR confirmed transcriptional activity of these three genes (**Figure 76; Sequence S6**). A total of 10 triple-transformants of *A. oryzae* were incubated in starch media for five days, then screened for metabolite production. Neither methylphloroacetophenone nor usnic acid were observable in these extracts (**Figure 76**). We therefore sought other solutions to this problem.

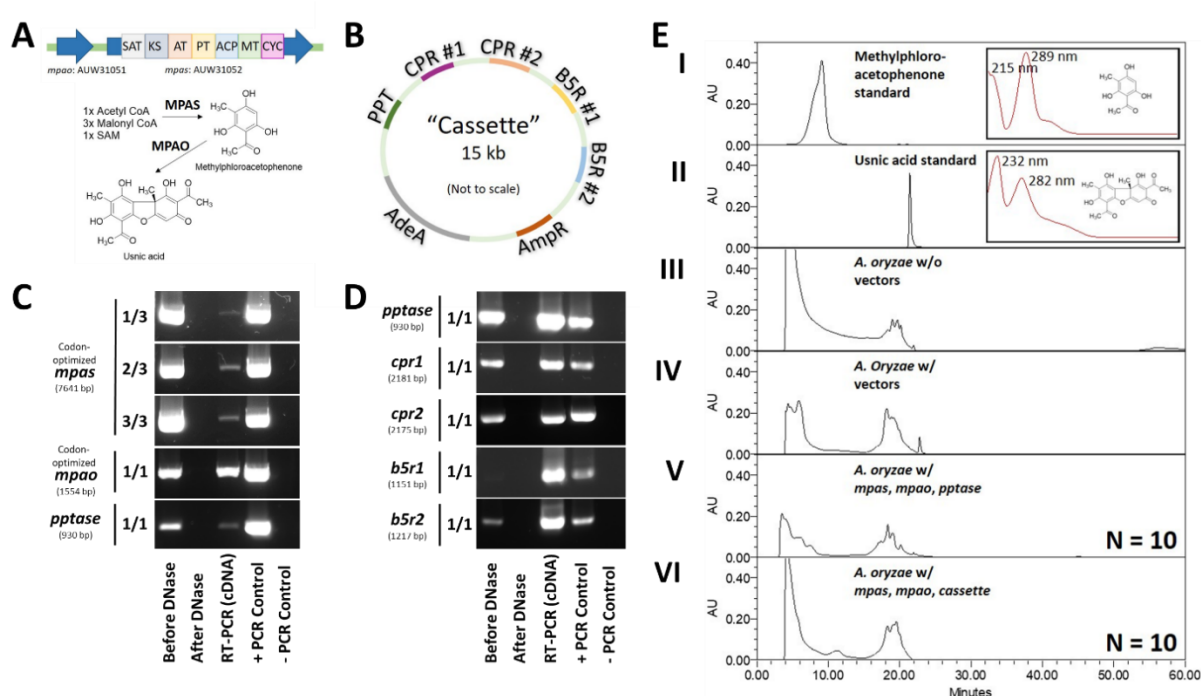


Figure 76: Functional heterologous expression trials of the putative usnic acid-producing genes methylphloroacetophenone synthase (*mpas*) and methylphloroacetophenone oxidase (*mpao*). **(A)** The usnic acid gene cluster and biosynthetic pathway in *C. uncialis* **(B)** Schematic of 'cassette' constructed in pAdeA containing five genes: Phosphopantetheinyl transferase (*pptase*) two cytochrome p450 reductases (*cpr1* and *cpr2*), and two cytochrome b₅ reductases (*b5r1* and *b5r2*). **(C)** RT-PCR of codon-optimized *mpas*, codon-optimized *mpao*, and *pptase* from *C. uncialis*, from mRNA of *A. oryzae* after transformation with these genes. To simplify the amplification of long templates, the cDNA of *mpas* was amplified in three portions. **(D)** RT-PCR of the cassette from mRNA of *A. oryzae* after transformation with the cassette. **(E)** HPLC traces of whole-metabolite extracts of *A. oryzae* after incubation in starch-imbued Czapek-Dox for five days: **(I)** Methylphloroacetophenone (MPA) standard **(II)** Usnic acid (UA) standard **(III)** Extract of *A. oryzae* without plasmid transformation. All such traces produced throughout this work are provided in **Figure S53** in Appendix **(IV)** Extract of *A. oryzae* when transformed with plasmids not containing lichen genes. All such traces produced throughout this work are provided in **Figure S54** in Appendix **(V)** Representative extract of *A. oryzae* when transformed with codon-optimized *mpas*, codon-optimized *mpao*, and *pptase*. The remaining nine traces are shown in **Figure S55** in Appendix **(VI)** Representative extract of *A. oryzae* when transformed with codon-optimized *mpas*, codon-optimized *mpao*, and cassette. Remaining nine traces are shown in **Figure S56** in Appendix. All traces displayed at 280 nm. Adapted from Bertrand et al. (2019), manuscript under preparation.

8.2.4. Trials using multiple accessory genes

Cytochrome p450 reductase (CPR) facilitates electron transfer from NADPH to cytochrome p450 and is therefore required as part of the p450 redox cycle. A second reductase, cytochrome B5 reductase (B5R), is also known to support the redox cycle of p450 (Waskell & Kim, 2015). Co-expression of CPR and/or B5R with cytochrome p450 in heterologous hosts can activate or increase the catalytic turnover of the cytochrome p450 (Hatakeyama et al. 2016; Brown et al. 2015; Kim et al. 2005; Jennewein et al. 2005; Li et al. 2016). We speculated that either *C. uncialis* CPR and/or B5R are required to complete the redox cycle of MPAO and that these functions were not being fulfilled in *A. oryzae*. Mock extractions of methylphloroacetophenone and usnic acid have shown that usnic acid is a stable molecule whereas methylphloroacetophenone appears to decompose over time (see Chapter 7). It is therefore possible that MPAS is producing trace amounts of methylphloroacetophenone but that it is decomposing in the absence of oxidative coupling to usnic acid. We were therefore interested in determining whether co-expression of *C. uncialis* CPR and B5R in *A. oryzae* would activate MPAO, with the expected outcome of accumulation of usnic acid.

Using characterized CPR and B5R entries available on GenBank, we searched the *C. uncialis* draft genome for genetically similar sequences. We identified two candidate CPR genes, named *cpr1* and *cpr2*, and two candidate B5R genes, named *b5r1* and *b5r2*. Including *pptase*, we used routine molecular biology techniques to construct a transformation plasmid that contains all five genes. This construct is referred to as the ‘cassette’ (**Figure 77**). Each of the five *C. uncialis* genes have a dedicated amylose promoter and terminator region. Isolation of mRNA from *A. oryzae* transformed with this cassette and RT-PCR demonstrated that these five genes were transcriptionally active including intron removal (**Figure 76**; see **Sequences S6-S10** in Appendix).

This cassette was co-transformed into *A. oryzae* with codon-optimized *mpas* and *mpao*. A total of 10 transformants were screened for metabolite production following five days of shaking incubation in starch-based media (see Chapter 2). No new natural products were observed in these trials (**Figure 76**). At this time, we re-directed our attention away from usnic acid biosynthesis and towards attempting the heterologous expression of other gene clusters from *C. uncialis*.

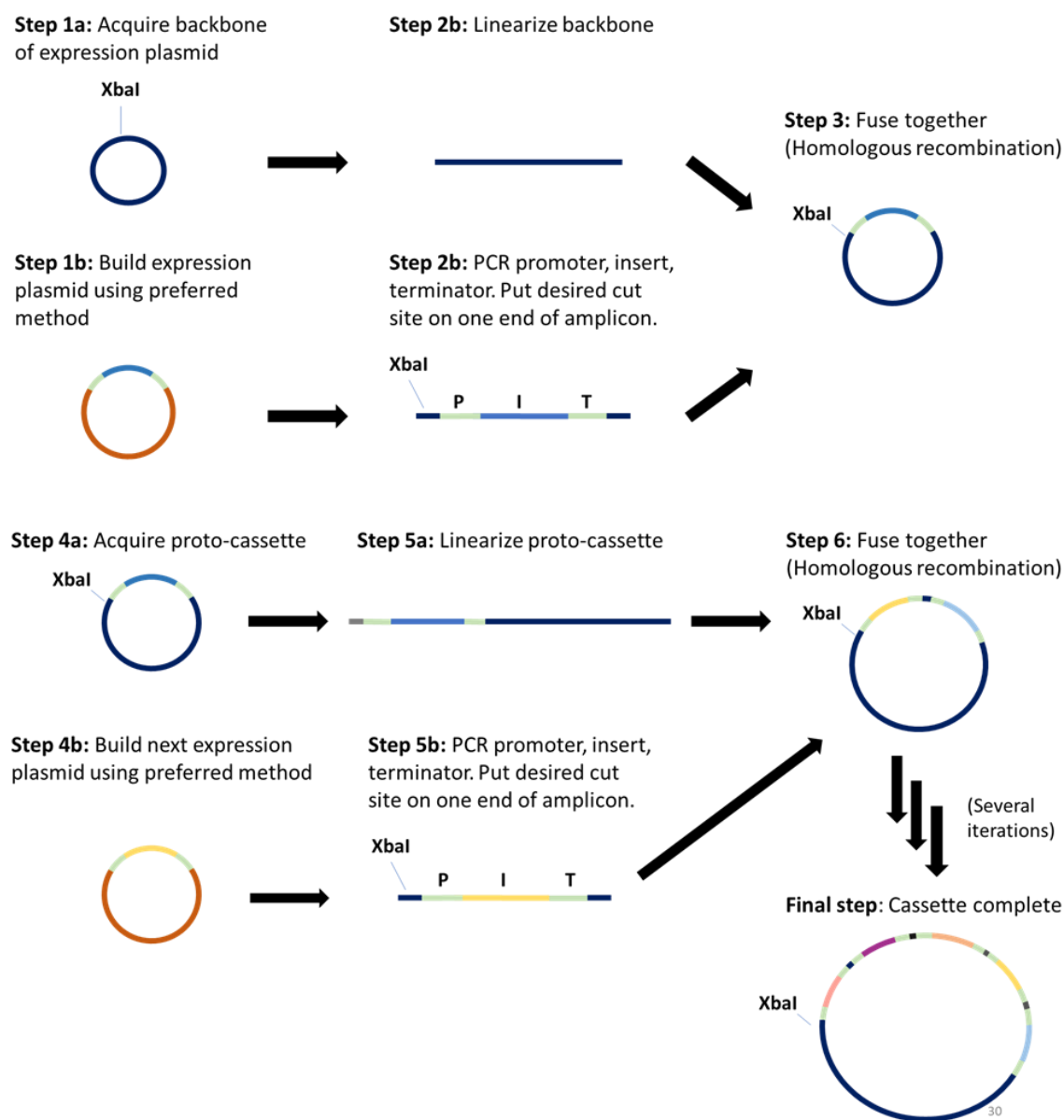


Figure 77: Outline of method of creating 'cassette' containing a total of five lichen genes, for heterologous expression trials in *A. oryzae*. Each of the five genes were demonstrated to be transcriptionally active (including intron removal) when transformed in *A. oryzae* (see **Figure 76**).

8.3. Heterologous expression trials of 6-hydroxymellein gene cluster (6HMS and KRDH)

The aromatic polyketide 6-hydroxymellein is the primary intermediate in the biosynthesis of terrein, a plant anti-germinant produced by *Aspergillus terreus* (Zaehle et al. 2014). As shown in Chapter 4, a non-reducing PKS produces 2,3-dehydro-6-hydroxymellein. A reductase, appearing to have evolved from a *trans*-ketoreductase/dehydratase (KRDH) didomain, performs a reduction, yielding 6-hydroxymellein. Several additional chemical steps transform 6-hydroxymellein into terrein (Zaehle et al. 2014). We have previously identified in *C. uncialis* what appears to be a homolog of this PKS and didomain, named *6hms* and *krdh* (see Chapter 4). Although the final product of this lichen gene cluster is unknown, several genes including a halogenase were observed near *6hms* and *krdh*, leading us to speculate that the final product is a halogenated isocoumarin (see Chapter 4). The fact that 6-hydroxymellein is an intermediate that appears to be common to terrein and our hypothesized isocoumarin pathway is an example of how evolution derivatizes ancestral biosynthetic pathways to produce new natural products.

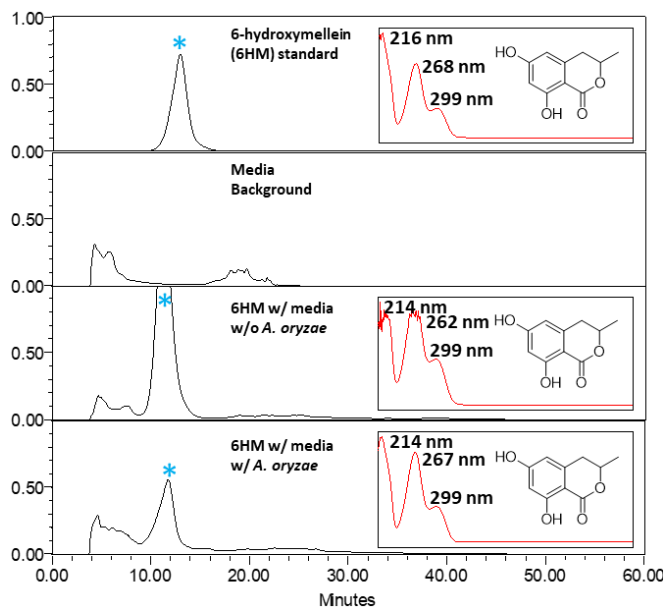


Figure 78: Mock extraction trial of 6-hydroxymellein in media with and without *A. oryzae*. Adapted from Bertrand et al. (2019), Supporting Information file, manuscript in preparation.

We were interested in testing these putative assignments of function by expressing the 6HMS and KRDH from *C. uncialis* in *A. oryzae*. The expected outcome was *de novo* biosynthesis of one or both of 2,3-dehydro-6-hydroxymellein and 6-hydroxymellein in *A. oryzae*. A chemical standard of 6-hydroxymellein was prepared by a co-worker to assist me in interpreting the results (see Acknowledgements). To avoid potential problems with *A. oryzae* PPTase, we decided to co-transform the *pptase* from *C. uncialis* with *6hms* and *krdh*. A mock extraction trial involving media spiked with 6-hydroxymellein demonstrated that 6-hydroxymellein is effectively recovered from media incubated alone or inoculated with *A. oryzae* (**Figure 78**). A total of eight transformants of *A. oryzae* containing these three lichen genes were incubated in starch media for five days on a rotating incubator at 30°C. Isolation of mRNA and RT-PCR demonstrated that transcription of *6hms* was occurring in *A. oryzae* (**Figure 79**; see **Sequence S11** in Appendix). However, we were unable to demonstrate heterologous transcription of *krdh* in *A. oryzae* despite three attempts to do so (data not shown). Nonetheless heterologous expression of both genes was attempted. No new metabolites were observed (**Figure 79**). We therefore tried a third gene cluster from *C. uncialis*.

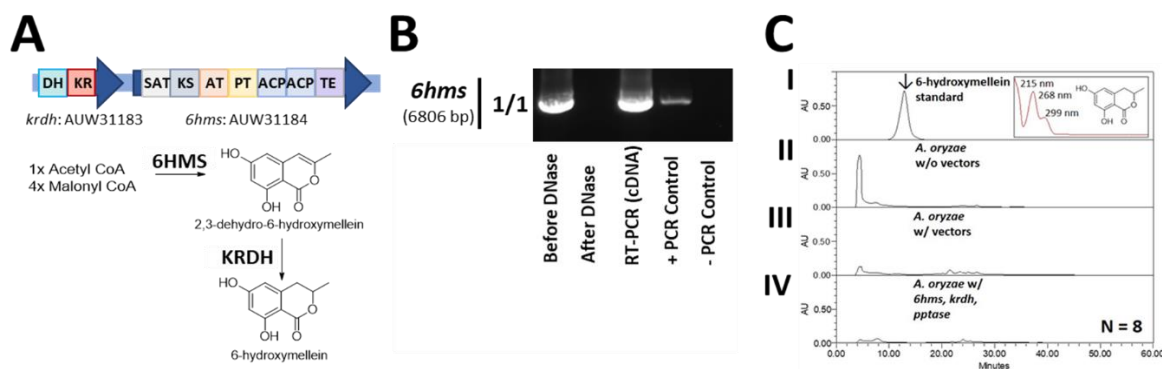


Figure 79: Functional heterologous expression trials of 6-hydroxymellein. (A) Biosynthetic pathway and gene cluster of 6-hydroxymellein in *C. uncialis* (B) RT-PCR of mRNA of *6hms* produced by *A. oryzae* (C) HPLC traces of whole-metabolite extracts of *A. oryzae* after incubation in Czapek-Dox with starch for five days: (I) 6-hydroxymellein standard; (II) Extract of *A. oryzae* not transformed with plasmids. All such traces produced throughout this work are provided in **Figure S53** in Appendix (III) Extract of *A. oryzae* transformed with plasmids that do not contain lichen genes. All such traces produced throughout this work are provided in **Figure S54** in Appendix (IV) Representative extract of *A. oryzae* transformed with *6hms*, *krdh*, and *pptase*. The remaining seven traces are shown in **Figure S57** in Appendix. Traces shown at 300 nm. Reproduced from Bertrand et al. (2019), manuscript in preparation.

8.4. Heterologous expression trials of two orsellinic acid synthases

In our preliminary experiments (Chapter 7), we demonstrated functional heterologous expression of an orsellinic acid synthase (OAS) from *Fusarium sp.* We decided to use this successful example to explore the heterologous expression of OAS genes from *C. uncialis*. This experiment would allow us to directly compare the ability of *A. oryzae* to express a lichen PKS versus a non-lichen PKS of identical biosynthetic function.

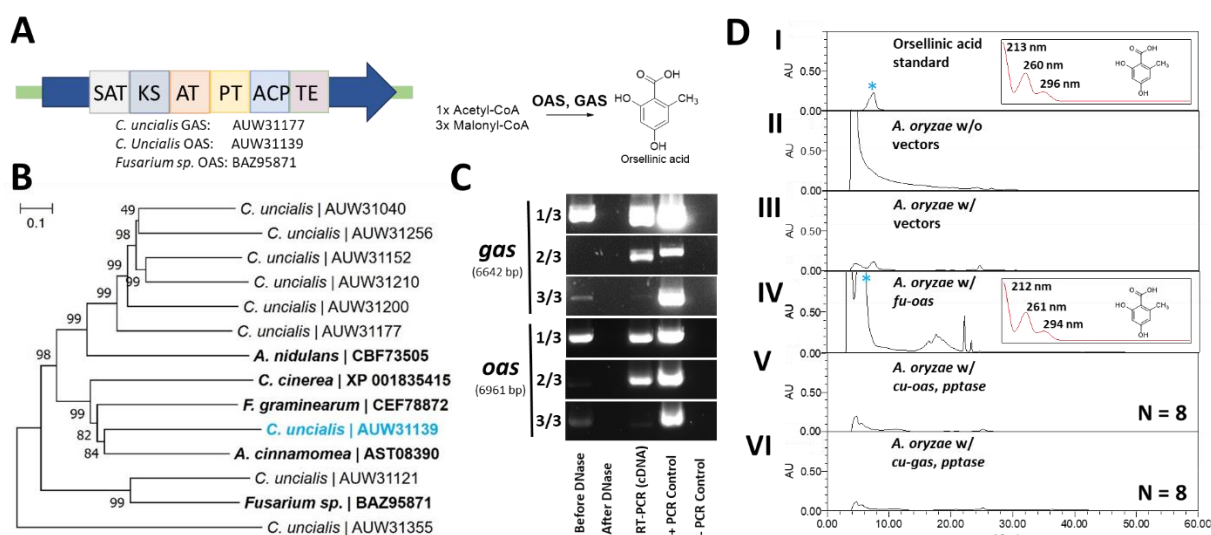


Figure 80: Functional heterologous expression trials of orsellinic acid. (A) Proposed orsellinic acid-producing PKS in *C. uncialis*, and known orsellinic acid-producing PKS in *Fusarium sp.* (B) Phylogenetic analysis of *C. uncialis* PKS with SAT-KS-AT-PT-ACP-TE domain architecture with known orsellinic acid-producing PKS from non-lichenizing fungi (highlighted in bold). Lichen PKS highlighted in blue was selected for heterologous expression studies. Estimated evolutionary distance is drawn to scale and unit bar represents units of amino acid substitutions per site. (C) RT-PCR of *gas* and *oas* of mRNA from *A. oryzae* transformed with these genes. To simplify the amplification of long templates, this cDNA was amplified in three portions. (D) HPLC traces of whole-metabolite extracts of *A. oryzae* after incubation in Czapek-Dox with starch for five days: (I) Orsellinic acid standard. (II) Extract of *A. oryzae* not transformed with plasmids. All such traces produced throughout this work are provided in **Figure S53** in Appendix (III) Extract of *A. oryzae* transformed with plasmids that do not contain genes from *C. uncialis* or *Fusarium sp.* All such traces produced throughout this work are provided in **Figure S54** in Appendix (IV) Extract of *A. oryzae* transformed with *oas* from *Fusarium sp.* and *pptase* from *C. uncialis*. Trace peak with retention time and absorption spectrum identical to standard orsellinic acid is shown. (V) Representative extract of *A. oryzae* transformed with *oas* and *pptase* from *C. uncialis*. The remaining seven traces are displayed as **Figure S58** in Appendix (VI) Representative extract of *A. oryzae* transformed with *gas* and *pptase* from *C. uncialis*. The remaining seven traces are displayed as **Figure S59** in Appendix. Traces shown at 260 nm. Reproduced from Bertrand et al. (2019), manuscript in preparation.

We were already aware of one candidate OAS from our previous work. Grayanic acid is a polyketide that is produced by the lichen *Cladonia grayi* (Armaleo et al. 2011). Orsellinic acid is required as an intermediate in the biosynthesis of grayanic acid. Transcriptional profiling experiments in *C. grayi* identified the PKS gene that is responsible for grayanic acid biosynthesis (Armaleo et al. 2011). In our work to annotate the genetic secondary metabolome of *C. uncialis*, we identified a gene cluster that appears to be functionally homologous to the grayanic acid gene cluster of *C. grayi*, an association that is supported by phylogenetic evidence (see Chapter 5 and Chapter 6). This grayanic acid synthase from *C. uncialis* (GAS, encoded as *cu-gas*) was therefore prepared for transformation into *A. oryzae*.

We sought a second OAS from *C. uncialis* for this experiment. These PKS have domain architecture of SAT-KS-AT-PT-ACP-TE. We compiled all PKS genes in *C. uncialis* that encoded this architecture (*6hms* was omitted from this analysis) and constructed a phylogenetic tree (see Chapter 2). Five PKS genes from non-lichen fungi, all of which have been experimentally characterized as orsellinic acid-producing PKS, were included in the phylogenetic tree (Yu et al. 2017; Ishiuchi et al. 2012; Okada et al. 2017; Jørgensen et al. 2014; Sanchez et al. 2010). The expected outcome was that the non-lichen PKS would aggregate with one or more lichen PKS within the phylogenetic tree, lending support to a putative assignment of function. This phylogenetic tree revealed that three of the five characterized OAS aggregated with one of our lichen PKS genes (**Figure 80**). This PKS, named CU-OAS (encoded as *cu-oas*) was selected as a second candidate OAS.

The *cu-oas* and *cu-gas* were transformed into *A. oryzae* in parallel experiments. To maintain consistency between heterologous expression trials, the *C. uncialis pptase* was co-transformed with each gene. These transformants were grown for five days in starch media.

Extraction of mRNA and RT-PCR confirmed that *cu-oas* and *cu-gas* are both transcriptionally active in *A. oryzae* (**Figure 80**; see **Sequences S12** and **S13** in Appendix). A standard of orsellinic acid was chemically prepared by a co-worker to assist us in interpreting the results (see Acknowledgements). As shown in the accompanying HPLC trace, neither orsellinic acid nor other metabolite were observed in a total of eight transformants of *cu-oas* and *cu-gas*. The process that was used to previously screen *fu-oas* for function activity in *A. oryzae* was identical to the process used to screen *cu-oas* and *cu-gas* (see Chapter 7). These experiments suggest that *A. oryzae* can functionally express OAS from non-lichenizing fungi such as *Fusarium sp.*, but cannot functionally express OAS from lichens such as *C. uncialis*.

8.5. Heterologous expression trials of 6-methylsalicylic acid synthase (6MSAS)

To further explore this problem, we elected to produce a parallel of the preceding experiment by attempting to functionally express a *C. uncialis* 6MSAS in *A. oryzae*. As we have previously demonstrated the functional heterologous expression of *pe-6msas* from *Penicillium sp.* (see Chapter 7), this provided us with another opportunity to compare the efficacy of lichen versus non-lichen PKS of identical function.

The *cu-6msas* was transformed into *A. oryzae*. To maintain consistency between trials, the *C. uncialis* *pptase* was co-transformed with *cu-6msas*. Isolation of mRNA from a transformant of *cu-6msas* confirmed that this gene was transcriptionally active in *A. oryzae* (**Figure 81**; see **Sequence S14** in Appendix). A total of eight transformants of *cu-6msas* were screened. Neither 6-methylsalicylic acid nor other metabolite were observed (**Figure 81**). These experiments suggest that *A. oryzae* can functionally express 6MSAS from non-lichenizing fungi such as *Penicillium sp.* but cannot functionally express 6MSAS from lichens such as *C. uncialis*.

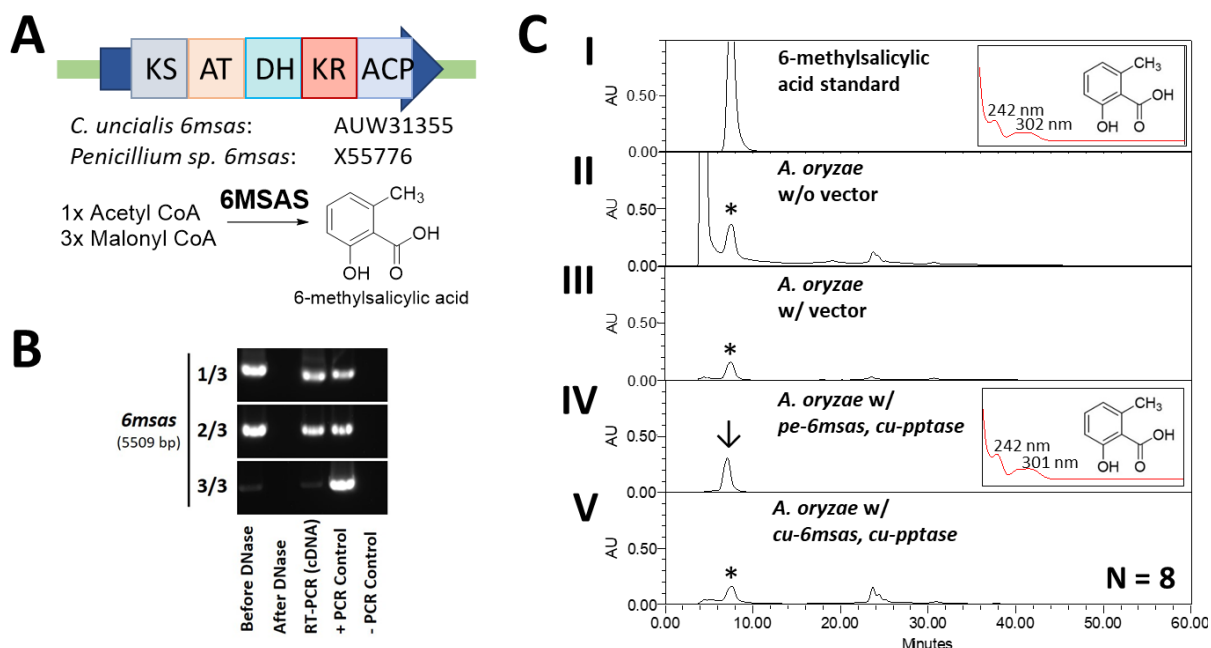


Figure 81: Functional heterologous expression trials of 6-methylsalicylic acid. (A) Proposed 6-methylsalicylic acid-producing PKS in *C. uncialis* and *Penicillium sp.* (B) RT-PCR of 6msas from *C. uncialis* of mRNA from *A. oryzae* transformed with this gene. To simplify the amplification of long templates, this cDNA was amplified in three portions. (C) HPLC traces of whole-metabolite extracts of *A. oryzae* after incubation in Czapek-Dox with starch for five days: (I) 6-methylsalicylic acid standard. (II) Extract of *A. oryzae* not transformed with plasmids. All such traces produced throughout this work are provided in **Figure S53** in Appendix (III) Extract of *A. oryzae* transformed with plasmids that do not contain genes from *C. uncialis* or *Penicillium sp.* All such traces produced throughout this work are provided in **Figure S54** in Appendix (IV) Extract of *A. oryzae* transformed with 6msas from *Penicillium sp.* and pptase from *C. uncialis*. Trace peak with retention time and absorption spectrum identical to standard 6-methylsalicylic acid is shown. (V) Representative extract of *A. oryzae* transformed with 6msas and pptase from *C. uncialis*. The remaining seven traces are shown in **Figure S60** in Appendix. Traces displayed at 300 nm. Asterisks in HPLC traces identify a peak produced by an endogenous metabolite with a retention time that is identical to 6-methylsalicylic acid but does not possess an identical absorption spectrum. Adapted from Bertrand et al. (2019), manuscript in preparation.

8.6. Heterologous expression trials of two type III PKS

Having now witnessed what appears to be a systemic inability of *A. oryzae* to functionally express type I PKS from lichens, we were interested in examining whether *A. oryzae* can process other types of biosynthetic enzymes from *C. uncialis*. In Chapter 5, I reported the annotation of two type III PKS genes in *C. uncialis*. In consideration of our findings, we decided to pivot away from type I PKS and instead attempt functional heterologous expression of these two type III PKS in *A. oryzae*. As we do not know the product of these type III PKS, no HPLC retention standard could be provided. However, based on our finding that transforming *fu-oas* from *Fusarium sp.*

and *pe-6msas* from *Penicillium* sp. in *A. oryzae* results in sufficient product to saturate the HPLC detector, we concluded that any metabolites produced from these type III PKS ought to be immediately observable through this approach. If successful, we could then begin work on elucidating the structure of the resultant molecule.

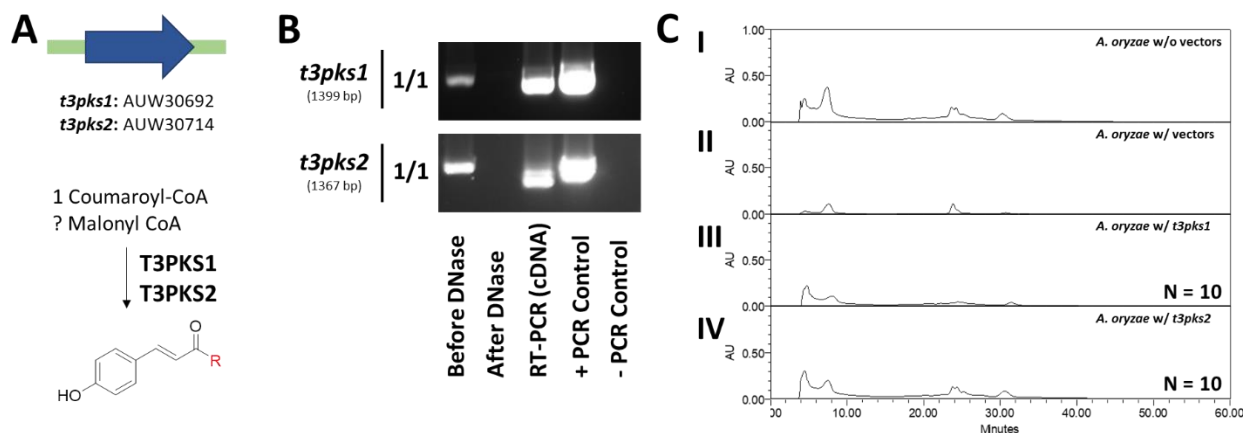


Figure 82: Heterologous expression trials of two type III PKS from *C. uncialis*. (A) The two type III PKS genes from *C. uncialis* (B) RT-PCR of *t3pks1* and *t3pks2* of mRNA from *A. oryzae* transformed with these genes. (C) HPLC traces of whole-metabolite extracts of *A. oryzae* after incubation in Czapek-Dox with starch for five days. (I) Extract of *A. oryzae* not transformed with plasmids. All such traces produced throughout this work are provided in **Figure S53** in Appendix (II) Extract of *A. oryzae* transformed with plasmids that do not contain genes from *C. uncialis*. All such traces produced throughout this work are provided in **Figure S54** in Appendix. (III) Representative extract of *A. oryzae* transformed with *t3pks1* from *C. uncialis*. The remaining nine traces are shown in **Figure S61** in Appendix. (IV) Representative extract of *A. oryzae* transformed with *t3pks2* from *C. uncialis*. The remaining nine traces are shown in **Figure S62** in Appendix. Traces displayed at 300 nm. Adapted from Bertrand et al. (2019), manuscript in preparation.

Both type III PKS were amplified by PCR from the genomic DNA of *C. uncialis*, placed into expression plasmids, and transformed into *A. oryzae* in parallel experiments (see Chapter 2). Isolation of mRNA from *A. oryzae* transformed either with *cu-t3pks-1* or *cu-t3pks-2* and RT-PCR demonstrated transcription of these genes including intron removal in *A. oryzae* (**Figure 82**; see **Sequence S15** and **S16** in Appendix). A total of ten transformants of *cu-t3pks-1* and ten transformants of *cu-t3pks-2* were prepared. These transformants were incubated in starch-based Czapek-Dox media for five days on a shaking incubator at 30°C. the media was subsequently extracted, and extractions analyzed by HPLC. These analyses did not reveal *de novo* production

of any new metabolite indicative of successful function expression of either type III PKS in *A. oryzae* (**Figure 82**).

8.7. Comparison of codon adaptability index (CAI) scores

In this Chapter, I have provided evidence of what appears to be a systemic problem affecting the functional expression of lichen PKS in *Aspergillus* hosts. This assessment is based on seven unsuccessful heterologous expression experiments of PKS from *C. uncialis* into *A. oryzae*. This assessment is also based on four reports provided by independent research groups detailing non-compliance of PKS genes (from species of lichen other than *C. uncialis*) when using either *A. oryzae* or *A. nidulans* as heterologous hosts (Chooi et al. 2008; Gagunashvili et al. 2009; Armaleo et al. 2011; Wang et al. 2016). There are no examples of the successful functional heterologous expression of lichen proteins involved in secondary metabolism.

One interesting outcome of this work is that we were able to readily functionally express two PKS from non-lichenizing fungi whereas we could not do so for PKS of identical biosynthetic function from *C. uncialis*. We speculated that differences in codon usage biases could explain this divergent outcome. As previously explained in the case of MPAS, organisms have biases with respect to which synonymous codons are preferentially used to encode amino acids, and that trouble may arise when attempting heterologous expression in organisms that have incongruent tRNA pools. We therefore hypothesized that the PKS of *Penicillium sp.* and *Fusarium sp.* have codon usage biases that are well-suited to translation in *A. oryzae* whereas the PKS of *C. uncialis* do not. This hypothesis may also explain the non-compliance of the other lichen PKS tested.

To evaluate this hypothesis, we calculated the codon adaptation index (CAI) scores of each PKS. The CAI is a measure of the similarity of the codon usage bias between genes and reference organisms (Sharp & Li, 1987). The CAI scores range from 0.0 to 1.0. Genes with scores of 0.8

or above are typically considered to have congruent codon biases with respect to the heterologous host. The CAI is therefore used as a proxy value to evaluate the likelihood of a successful heterologous expression experiment.

Table 11: Codon adaptation index (CAI) scores of lichen versus non-lichen PKS genes in *A. oryzae*. Adapted from Bertrand et al. (2019), manuscript in preparation.

Biosynthetic function of PKS	PKS Gene Species of Origin	CAI (OPTIMIZER)	CAI (E-CAI)	CAI (Average)	% CAI Difference
Methylphloroaceto- phenone synthase (MPAS)	<i>C. uncialis</i>	0.837	0.818	0.823	14.6
	[Codon-optimized]	0.964	0.963	0.964	N/A
6-methylsalicylic acid synthase (6MSAS)	<i>C. uncialis</i>	0.837	0.826	0.832	4.6
	<i>Penicillium sp.</i>	0.876	0.868	0.872	N/A
Orsellinic acid synthase (OAS)	<i>C. uncialis (gas)</i>	0.841	0.825	0.833	4.7
	<i>C. uncialis (oas)</i>	0.855	0.841	0.848	3.0
	<i>Fusarium sp.</i>	0.881	0.866	0.874	N/A
6-hydroxymellein synthase (6HMS)	<i>C. uncialis</i>	0.814	0.832	0.823	N/A
Chalcone synthase (T3PKS)	<i>C. uncialis (t3pks1)</i>	0.812	0.828	0.820	N/A
Chalcone synthase (T3PKS)	<i>C. uncialis (t3pks2)</i>	0.805	0.822	0.814	N/A

To calculate CAI, the codon usage table for the *A. oryzae* genome was acquired from the Codon Usage Database (Available at: <https://www.kazusa.or.jp/codon/>). This table was then fed into two calculators, OPTIMIZER (Puigbò et al. 2007) and E-CAI (Puigbò et al. 2008). The CAI of PKS genes was evaluated by submitting the exonic portions of PKS genes into these two calculators, and the resultant CAI scores were averaged. The CAI of the native versus codon-optimized variant of *mpas* was also calculated. The expected outcome was that lichen PKS genes would have CAI scores lower than 0.8, which would indicate unsuitability for heterologous

expression in *A. oryzae*. A second expected outcome was that a considerable difference would be observed between the CAI scores of lichen versus non-lichen PKS of identical function.

Contrary to the expected outcome, each *C. uncialis* PKS gene (*mpas*, *6msas*, *gas*, *oas*, *6hms*, *t3pks1*, and *t3pks2*) all possessed CAI scores greater than 0.8, indicating a high likelihood of heterologous expression in *A. oryzae* (**Table 11**). Moreover, although non-lichen PKS possessed greater CAI scores than lichen PKS, the difference was small: In no case was the percent difference between lichen versus non-lichen PKS greater than 5.0% (**Table 11**). Should a proposition be forwarded that this small (>5.0%) difference in CAI scores is nonetheless paramount to functional expression, the reader ought to consider that the CAI of codon-optimized *mpas* was improved by nearly 15% as compared to its non-optimized variant, yet still did not result in functional expression of this enzyme (**Table 11**). I must therefore conclude that differences in codon usage biases cannot adequately explain the divergent outcomes seen in our heterologous expression trials.

8.8. Summary

A total of five type I PKS genes (*mpas*, *6hms*, *oas*, *gas*, *6hms*) and two type III PKS genes (*t3pks1*, *t3pks2*) from *C. uncialis* were transformed into *A. oryzae*. With exception of *krdh*, extraction of mRNA from transformed *A. oryzae* and RT-PCR confirmed that lichen genes were transcriptionally active in *A. oryzae* and that introns were removed to produce mature mRNA. Production of the expected metabolite was not observed in any instance. With respect to *mpas* and *mpao*, this problem was not resolved by codon-optimization of these genes nor by co-expression of accessory genes such as *pptase*, *cpr1*, *cpr2*, *b5r1*, and *b5r2*. Two genes from non-lichenizing fungi, *fu-oas* from *Fusarium sp.* and *pe-6msas* from *Penicillium sp.* were transformed into *A. oryzae*, and the result of these experiments was the copious production of orsellinic acid and 6-

methylsalicylic acid, respectively. We conclude that *A. oryzae* may functionally express non-lichen PKS, but for reasons presently unknown to us, cannot do so with lichen PKS of equivalent functions.

Chapter 9

Conclusions and Future Prospects

9.1. Current knowledge on the genetic programming of lichen secondary metabolite biosynthesis

Although lichens are known to produce many secondary metabolites, few of these molecules have been linked to their originating gene clusters. This is in large part because of the challenges of studying lichens, such as their slow rates of growth (see Chapter 1). An objective of my doctoral program was to therefore expand the amount of genetic information available on lichen secondary metabolite gene clusters. When I began this research in September 2012, only 13 complete PKS genes from lichens were reported (see **Table 1** in Chapter 1). Since that time, information on lichen biosynthetic genes has expanded tremendously. Consider, for example, that by 2018 the number of PKS genes reported in the literature increased to 81 (**Table 12**). This total does not include gene annotation projects published after late 2018 (Dal Grande et al. 2018; Wang et al. 2018; Calchera et al. 2019), nor does it include reports of draft genomes of lichen fungi wherein dozens of PKS genes have been identified in each organism (Park et al. 2013a; 2013b; 2014a; 2014b; 2014c; Wang et al. 2018).

Table 12: Complete PKS genes from lichen fungi reported in the literature (compiled 2018). Adapted from Bertrand & Sorensen (2018), in accordance with authors' retainment of privileges.

Entry	Name	Accession	Species	Architecture	Detection Method	Putative Function	Prediction Method
1	<i>Cgrpks15</i>	HQ823621	<i>Cladonia grayi</i>	SAT-KS-AT-PT-ACP-TE	cDNA - genome assembly	-	-
2	<i>Cmapks1</i>	JQ340775	<i>Cladonia macilenta</i>	SAT-KS-AT-PT-ACP-TE	gDNA - library screening	Melanin/orsellinic acid-related	Phylogenetics
3	<i>Cu-nr-pks-4</i>	MG777492	<i>Cladonia uncialis</i>	SAT-KS-AT-PT-ACP-TE	Genome sequencing	-	-
4	<i>Cu-nr-pks-5</i>	MG777493	<i>Cladonia uncialis</i>	SAT-KS-AT-PT-ACP-TE	Genome sequencing	-	-
5	<i>Cu-nr-pks-14</i>	MG777502	<i>Cladonia uncialis</i>	SAT-KS-AT-PT-ACP-TE	Genome sequencing	-	-
6	<i>Ulpks6</i>	JX232188	<i>Usnea longissima</i>	SAT-KS-AT-PT-ACP-TE	gDNA - library screening	Orsellinic acid-related	Phylogenetics
7	<i>Cgrpks13</i>	HQ823619	<i>Cladonia grayi</i>	SAT-KS-AT-PT-ACP-ACP-TE	cDNA - genome assembly	-	-
8	<i>Cgrpks14</i>	HQ823620	<i>Cladonia grayi</i>	SAT-KS-AT-PT-ACP-ACP-TE	cDNA - genome assembly	-	-
9	<i>Cgrpks16</i>	GU930713	<i>Cladonia grayi</i>	SAT-KS-AT-PT-ACP-ACP-TE	cDNA - genome assembly	Grayanic acid	Transcription profiling
10	<i>Cu-nr-pks-1</i>	MG777489	<i>Cladonia uncialis</i>	SAT-KS-AT-PT-ACP-ACP-TE	Genome sequencing	-	-
11	<i>Cu-nr-pks-6</i>	MG777494	<i>Cladonia uncialis</i>	SAT-KS-AT-PT-ACP-ACP-TE	Genome sequencing	-	-
12	<i>Cu-nr-pks-7</i>	MG777495	<i>Cladonia uncialis</i>	SAT-KS-AT-PT-ACP-ACP-TE	Genome sequencing	Grayanic acid	Homology mapping
13	<i>Cu-nr-pks-8</i>	MG777496	<i>Cladonia uncialis</i>	SAT-KS-AT-PT-ACP-ACP-TE	cDNA - RACE	Terrein-related	Homology mapping
14	<i>Cu-nr-pks-9</i>	MG777497	<i>Cladonia uncialis</i>	SAT-KS-AT-PT-ACP-ACP-TE	Genome sequencing	-	-
15	<i>Cu-nr-pks-10</i>	MG777498	<i>Cladonia uncialis</i>	SAT-KS-AT-PT-ACP-ACP-TE	Genome sequencing	Fusarubin-related	Homology mapping
16	<i>Dnpks1</i>	EU872212	<i>Dirinaria applanata</i>	SAT-KS-AT-PT-ACP-ACP-TE	gDNA - library screening	Melanin/anthraquinone-related	Phylogenetics
17	<i>Eppks3</i>	ERF69996	<i>Endocarpon pusillum</i>	SAT-KS-AT-PT-ACP-ACP-TE	Genome sequencing	-	-
18	<i>Eppks4</i>	ERF77221	<i>Endocarpon pusillum</i>	SAT-KS-AT-PT-ACP-ACP-TE	Genome sequencing	-	-
19	<i>Pmpks4</i>	HM180407	<i>Peltigera membranacea</i>	SAT-KS-AT-PT-ACP-ACP-TE	Genome sequencing	-	-
20	<i>Pmpks6</i>	HM180409	<i>Peltigera membranacea</i>	SAT-KS-AT-PT-ACP-ACP-TE	Genome sequencing	-	-
21	<i>Ulpks1</i>	JN408682	<i>Usnea longissima</i>	SAT-KS-AT-PT-ACP-ACP-TE	gDNA - library screening	Melanin/orsellinic acid-related	Phylogenetics
22	<i>Ulpks2</i>	JX232185	<i>Usnea longissima</i>	SAT-KS-AT-PT-ACP-ACP-TE	gDNA - library screening	Orsellinic acid-related	Phylogenetics
23	<i>Xsmpks1</i>	KJ501919	<i>Xanthoparmelia substrigosa</i>	SAT-KS-AT-PT-ACP-ACP-TE	cDNA - RACE	-	-
24	<i>Xepks1</i>	DQ660910	<i>Xanthoria elegans</i>	SAT-KS-AT-PT-ACP-ACP-TE	cDNA - RACE	Anthraquinone-related	Phylogenetics
25	<i>Cu-nr-pks-11</i>	MG777499	<i>Cladonia uncialis</i>	SAT-KS-AT-PT-ACP-MT-TE	Genome sequencing	Mycophenolic acid-related	Homology mapping
26	<i>Ulpks4</i>	JX232186	<i>Usnea longissima</i>	SAT-KS-AT-PT-ACP-MT-TE	gDNA - library screening	Methylorsellinic acid-related	Phylogenetics
27	<i>Cu-nr-pks-2</i>	MG777490	<i>Cladonia uncialis</i>	SAT-KS-AT-PT-ACP-MT-CYC	Genome sequencing	Usnic acid	Gene surveying
28	<i>Cu-nr-pks-12</i>	MG777500	<i>Cladonia uncialis</i>	SAT-KS-AT-PT-ACP-MT-CYC	Genome sequencing	-	-
29	<i>Cgrpks2</i>	GU930714	<i>Cladonia grayi</i>	SAT-KS-AT-PT-ACP-MT-CYC	cDNA - genome assembly	-	-
30	<i>Cgrpks1</i>	HQ823618	<i>Cladonia grayi</i>	SAT-KS-AT-PT-ACP-ACP-MT-CYC	cDNA - genome assembly	-	-
31	<i>Eppks1</i>	ERF73753	<i>Endocarpon pusillum</i>	SAT-KS-AT-PT-ACP-R	Genome sequencing	-	-
32	<i>Hypopks1</i>	XK067626	<i>Hypogymnia physodes</i>	SAT-KS-AT-PT-ACP-MT-R	gDNA - library screening	Methylorcinolaldehyde-related	Phylogenetics
33	<i>Hypopks2</i>	XK067627	<i>Hypogymnia physodes</i>	SAT-KS-AT-PT-ACP-MT-R	gDNA - library screening	Methylorcinolaldehyde-related	Phylogenetics
34	<i>Cu-nr-pks-3</i>	MG777491	<i>Cladonia uncialis</i>	SAT-KS-AT-PT-ACP-ACP-MT-R	Genome sequencing	Azaphilone-related	Homology mapping
35	<i>Eppks2</i>	ERF73358	<i>Endocarpon pusillum</i>	SAT-KS-AT-PT-ACP-ACP-MT-R	Genome sequencing	-	-
36	<i>Cu-nr-pks-13</i>	MG777501	<i>Cladonia uncialis</i>	SAT-KS-AT-PT-ACP	Genome sequencing	Pestheic acid-related	Homology mapping
37	<i>Ulpks5</i>	JX232187	<i>Usnea longissima</i>	SAT-KS-AT-PT-ACP	gDNA - library screening	Anthraquinone-related	Phylogenetics
38	<i>Xseps1</i>	AB558604	<i>Xanthoparmelia semiviridis</i>	SAT-KS-AT-PT-ACP-MT	gDNA - library screening	Methylorsellinic acid-related	Phylogenetics
39	<i>Cmpks1</i>	HQ413098	<i>Cladonia metacorrallifera</i>	KS-AT-DH-KR-ACP	gDNA - library screening	Methylsalicylic acid-related	Phylogenetics
40	<i>Cu-r-pks-3</i>	MG777496	<i>Cladonia uncialis</i>	KS-AT-DH-KR-ACP	Genome sequencing	-	-
41	<i>Cu-r-pks-9</i>	MG777507	<i>Cladonia uncialis</i>	KS-AT-DH-KR-ACP	Genome sequencing	Patulin	Homology mapping
42	<i>Eppks5</i>	ERF77015	<i>Endocarpon pusillum</i>	KS-AT-DH-KR-ACP	Genome sequencing	-	-
43	<i>Eppks14</i>	ERF69189	<i>Endocarpon pusillum</i>	KS-AT-DH-KR-ACP	Genome sequencing	-	-
44	<i>Cu-r-pks-12</i>	MG777510	<i>Cladonia uncialis</i>	KS-AT-DH-MT-KR-ACP	Genome sequencing	-	-
45	<i>Cu-r-pks-4</i>	MG777497	<i>Cladonia uncialis</i>	KS-AT-DH-ER-KR-ACP	Genome sequencing	-	-
46	<i>Cu-r-pks-6</i>	MG777504	<i>Cladonia uncialis</i>	KS-AT-DH-ER-KR-ACP	Genome sequencing	-	-
47	<i>Cu-r-pks-10</i>	MG777508	<i>Cladonia uncialis</i>	KS-AT-DH-ER-KR-ACP	Genome sequencing	-	-
48	<i>Cu-r-pks-11</i>	MG777509	<i>Cladonia uncialis</i>	KS-AT-DH-ER-KR-ACP	Genome sequencing	-	-
49	<i>Eppks11</i>	ERF75982	<i>Endocarpon pusillum</i>	KS-AT-DH-ER-KR-ACP	Genome sequencing	-	-
50	<i>Eppks12</i>	ERF73411	<i>Endocarpon pusillum</i>	KS-AT-DH-ER-KR-ACP	Genome sequencing	-	-
51	<i>Eppks13</i>	ERF68064	<i>Endocarpon pusillum</i>	KS-AT-DH-ER-KR-ACP	Genome sequencing	-	-
52	<i>Eppks15</i>	ERF76568	<i>Endocarpon pusillum</i>	KS-AT-DH-ER-KR-ACP	Genome sequencing	-	-
53	<i>Ulpks3</i>	HQ824546	<i>Usnea longissima</i>	KS-AT-DH-ER-KR-ACP	gDNA - library screening	Resorcylic acid lactone-related	Phylogenetics
54	<i>Cu-r-pks-1</i>	MG777470	<i>Cladonia uncialis</i>	KS-AT-DH-MT-ER-KR-ACP	Genome sequencing	-	-
55	<i>Cu-r-pks-2</i>	MG777488	<i>Cladonia uncialis</i>	KS-AT-DH-MT-ER-KR-ACP	Genome sequencing	-	-
56	<i>Cu-r-pks-5</i>	MG777503	<i>Cladonia uncialis</i>	KS-AT-DH-MT-ER-KR-ACP	Genome sequencing	-	-
57	<i>Cu-r-pks-7</i>	MG777505	<i>Cladonia uncialis</i>	KS-AT-DH-MT-ER-KR-ACP	Genome sequencing	-	-
58	<i>Cu-r-pks-8</i>	MG777506	<i>Cladonia uncialis</i>	KS-AT-DH-MT-ER-KR-ACP	Genome sequencing	-	-
59	<i>Cu-r-pks-13</i>	MG777511	<i>Cladonia uncialis</i>	KS-AT-DH-MT-ER-KR-ACP	Genome sequencing	-	-
60	<i>Cu-r-pks-14</i>	MG777512	<i>Cladonia uncialis</i>	KS-AT-DH-MT-ER-KR-ACP	Genome sequencing	-	-
61	<i>Cu-r-pks-16</i>	MG777514	<i>Cladonia uncialis</i>	KS-AT-DH-MT-ER-KR-ACP	Genome sequencing	-	-
62	<i>Cu-r-pks-17</i>	MG777515	<i>Cladonia uncialis</i>	KS-AT-DH-MT-ER-KR-ACP	Genome sequencing	-	-
63	<i>Cu-r-pks-18</i>	MG777516	<i>Cladonia uncialis</i>	KS-AT-DH-MT-ER-KR-ACP	Genome sequencing	-	-
64	<i>Eppks6</i>	ERF74612	<i>Endocarpon pusillum</i>	KS-AT-DH-MT-ER-KR-ACP	Genome sequencing	-	-
65	<i>Eppks7</i>	ERF70234	<i>Endocarpon pusillum</i>	KS-AT-DH-MT-ER-KR-ACP	Genome sequencing	-	-
66	<i>Eppks8</i>	ERF71956	<i>Endocarpon pusillum</i>	KS-AT-DH-MT-ER-KR-ACP	Genome sequencing	-	-
67	<i>Eppks9</i>	ERF77204	<i>Endocarpon pusillum</i>	KS-AT-DH-MT-ER-KR-ACP	Genome sequencing	-	-
68	<i>Pmpks8</i>	HM180411	<i>Peltigera membranacea</i>	KS-AT-DH-MT-ER-KR-ACP	Genome sequencing	-	-
69	<i>Pmpks9</i>	HM180412	<i>Peltigera membranacea</i>	KS-AT-DH-MT-ER-KR-ACP	Genome sequencing	-	-
70	<i>Scpks1</i>	EF554834	<i>Solorina crocea</i>	KS-AT-DH-MT-ER-KR-ACP	gDNA - library screening	-	-
71	<i>Cu-r-pks-15</i>	MG777513	<i>Cladonia uncialis</i>	KS-AT-DH-MT-KR-ACP-R	Genome sequencing	Betaenone A-C	Homology mapping
72	<i>Pmpks1</i>	GU441232	<i>Peltigera membranacea</i>	KS-AT-DH-ER-KR	Genome sequencing	-	-
73	<i>Pmpks2</i>	GU460164	<i>Peltigera membranacea</i>	KS-AT-DH-ER-KR	Genome sequencing	-	-
74	<i>Pmpks3</i>	GU477350	<i>Peltigera membranacea</i>	KS-AT-DH-ER-KR	Genome sequencing	-	-
75	<i>Pmpks10</i>	HM189674	<i>Peltigera membranacea</i>	KS-AT-DH-ER-KR	Genome sequencing	-	-
76	<i>Eppks10</i>	ERF73749	<i>Endocarpon pusillum</i>	KS-AT-DH-MT-ER-KR	Genome sequencing	-	-
77	<i>Pmpks5</i>	HM180408	<i>Peltigera membranacea</i>	KS-AT-DH-MT-ER-KR	Genome sequencing	-	-
78	<i>Pmpks7</i>	HM180410	<i>Peltigera membranacea</i>	KS-AT-DH-MT-ER-KR	Genome sequencing	-	-
79	<i>Pmpks11</i>	HM189675	<i>Peltigera membranacea</i>	KS-AT-DH-MT-KR-ACP-C	Genome sequencing	-	-
80	<i>Cu-t3-pks-1</i>	MG777469	<i>Cladonia uncialis</i>	N/A	Genome sequencing	Chalcone synthase	BLAST consensus
81	<i>Cu-t3-pks-2</i>	MG777470	<i>Cladonia uncialis</i>	N/A	Genome sequencing	Chalcone synthase	BLAST consensus

Where is this new information coming from? The use of degenerate primers to amplify PKS genes from either genomic DNA or cDNA templates has historically been the most commonly applied method to discover new PKS genes in lichen fungi (Armaleo et al. 2011; Bingle et al. 1999; Brunauer et al. 2009, Miao et al. 2001; Nicholson et al. 2001). However, this process is slow, low throughput, and as I have explained in the introductory paragraph of Chapter 4, can lead to mis-targeting of sought genes. The maturation of ‘next-generation’ sequencing technologies has since enabled the rapid detection of large numbers of biosynthetic genes within lichen genomes (Cacho et al. 2015; Doroghazi et al. 2014; Goodwin et al. 2016; Ziemert et al. 2016). In addition to our own reports using this technology (see Chapter 3 and 5), whole-genome sequencing platforms have been used by other research groups to annotate lichen gene clusters (Kampa et al. 2013; Wang et al. 2014; Dal Grande et al. 2018; Wang et al. 2018; Calchera et al. 2019). Studies such as these are responsible for generating most information presently available on the genetic programming of lichen secondary metabolite biosynthesis. Next-generation platforms will undoubtedly continue to serve a role in the study of lichen secondary metabolites in the foreseeable future.

9.2. Proposed approaches to functionally expressing lichen genes in heterologous hosts

In Chapters 7 and 8, I have presented our ultimately unsuccessful efforts to functionally express lichen PKS genes in *A. oryzae*. Functional expression was attempted for: (1) *mpas* – methylphloacetophenone; (2) *6hms* – 6-hydroxymellein; (3) *oas* – orsellinic acid; (4) *gas* – orsellinic acid; (5) *6msas* – 6-methylsalicylic acid; (6) *t3pks1* – unknown product; and (7) *t3pks2* – unknown product. In contrast to these unsuccessful attempts, we were able to immediately observe functional expression of two non-lichen fungal PKS (*fu-oas* from *Fusarium sp.* and *pe-6msas* from *Penicillium sp.*). Various research groups attempting to express PKS from other species

of lichen in *A. oryzae* and *A. nidulans* have reported similarly negative outcomes with PKS genes from other species of lichens (Chooi et al. 2008; Gagunashvili et al. 2009; Armaleo et al. 2011; Wang et al. 2016). We conclude, for reasons unknown to us, that these expression systems can readily functionally express PKS from non-lichenizing fungi but cannot do so for lichen PKS of identical biosynthetic functions.

Although the cause of this apparently systemic failure is unknown, there are several factors that can be ruled-out as potential causes. We know from our *in silico* work on *mpas* and *6hms* that these PKS genes evolved by vertical gene transfer from ancestral *Ascomycota* fungi (Chapter 7). The catalytic triad and tertiary architecture of the KS domains of *mpas* and *6hms*, the key carbon-carbon bond-forming domain universal to PKS, appear to be similar in structure and function as compared to KS from other *Ascomycota* fungi. (Chapter 7). Choosing *A. oryzae* (an *Ascomycota* fungus) as a heterologous host for these PKS would be a logical choice based on these findings. With the exception of a single gene from *C. uncialis* (*krdh*), all genes that we have inserted into *A. oryzae* were observed to be transcriptionally active (Chapters 7 and 8). Introns are removed from the pre-mRNA to produce translation-coherent mRNA. From our detailed study of *mpas* transcription in *C. uncialis* and *A. oryzae*, we also know that the host and native organisms are removing introns at identical nucleotide positions (Chapter 7). We can therefore rule-out aberrant transcription as an explanation of findings for all but one of the PKS tested. Codon biases do not appear to be a factor because codon-optimized versions of *mpas* and *mpao* also failed to produce detectable metabolite in host (Chapter 7). Due to the cost of optimizing genes, we only explored the ‘codon bias hypothesis’ with respect to the usnic acid gene cluster. It is therefore possible that other PKS were not functionally expressed in *A. oryzae* because of mismatched codons. The problem does not appear to be caused by a failure to post-translationally modify PKS or

cytochrome p450s: Co-expression of *C. uncialis* phosphopantetheinyl transferases (PPTase), cytochrome p450 reductases (CPR1 and CPR2), and B5 reductases (B5R1 and B5R2) with MPAS and MPAO did not result in *de novo* metabolite biosynthesis. Neither did co-expression of PPTase with other PKS result in new metabolites arising in *A. oryzae* (Chapter 8). Whether translation of exogenous genes into protein was occurring in *A. oryzae* remains an open question. This is because we were unable to create a reliable internal control despite two attempts using either an endogenous or an exogenous protein (Chapter 7). It is therefore possible that protein translation of lichen genes is not occurring, though we have no evidence for or against this proposition. Lastly, errors by me (e.g., incorrect use of the NSAR1 *A. oryzae* expression system) is also ruled-out due to the successful functional heterologous expression of two PKS from *Fusarium sp.* and *Penicillium sp.* (Chapter 7).

What else can be done to functionally express lichen PKS? One option is to try *Aspergillus nidulans* as a host. As seemingly illogical as it sounds to propose another member of the *Aspergillus* genus after our failures with *A. oryzae*, this proposal is supported by one very crucial observation: This host was used to functionally express a lichen enzyme. PyrG is a decarboxylase found in the lichen *Solorina crocea* and is responsible for the decarboxylation of orotidine 5'-monophosphate to the primary metabolite uridine monophosphate. Functional heterologous expression of PyrG including product turnover was reported using *Aspergillus nidulans* (Sinnemann et al. 2000). This singular example of functional heterologous expression, in this case an enzyme involved in primary metabolism rather than secondary metabolism, provides proof-of-concept evidence that functional expression of enzymes involved in secondary metabolism ought to be possible. Other expression platforms could also be used: Numerous examples of successful heterologous expression of fungal PKS have been reported using yeast (Tsunematsu et al. 2013;

Billingsley et al. 2016). Prokaryotic hosts such as *E. coli* may also be used to functionally express fungal PKS although codon optimization and removing intronic regions from protein-encoding sequences are usually required as preliminary steps (Li & Neubauer 2014). An alternative to these intact cellular approaches is cell-free systems. In its crudest form, cell-free synthesis is an *in vitro* method that combines the lysate of an organism with RNA template, amino acids, and an energy supply to produce exogenous proteins (Lee & Kim, 2018). A clear advantage of this approach over *in vivo* methods (whole cells) is the ability to fine-tune environmental conditions. This could be useful for troubleshooting protein translation and elucidating the biosynthetic function of lichen enzymes. However, a broad goal of this research is to not only elucidate the function of these enzymes but to create a platform for the scalable production of useful biomolecules from lichens. Using whole cells provides that scalability. The best example is the reconstruction of the complete biosynthetic pathway of artemisinin in *S. cerevisiae* (Paddon et al. 2013). This yeast platform is now grown in industrial bioreactors and has thus far produced an estimated 100 million doses of this potent antimalarial drug. Though cell-free systems could be used as a ‘bridging’ experiment that connects phylogenetic predictions to experimental deduction of biosynthetic function, heterologous expression will inevitably be required to commercially exploit the lichen secondary metabolome. For this reason, I advocate that research remains focused on developing reliable methodologies for the heterologous production of lichen secondary metabolites.

9.3. Proposed approaches to discovering cryptic secondary metabolites in lichens

We were unable to demonstrate functional heterologous expression of lichen biosynthetic genes in a heterologous host. What else could be done to explore its secondary metabolome? One option is to continue applying the indirect methods that I have described throughout this thesis. For example, phylogenetics is useful for predicting the biosynthetic function of encoded enzymes

(Ziemert & Jensen, 2012). Numerous examples exist of the applications of phylogenetics towards the prediction of function of lichen PKS genes (Schmitt et al. 2008; Timsina et al. 2014; Timsina et al. 2012; Wang et al. 2014). The work of Armaleo et al. (2011), and more recently, the work of Wang et al. (2018), demonstrated the utility of transcriptional profiling experiments in the identification of gene clusters associated with lichen secondary metabolites. In Chapter 3, I described a gene surveying approach that identified the usnic acid gene cluster by applying retro-biosynthetic deductions on a list of compiled PKS clusters. The ‘homology mapping’ approach that I described in Chapter 4, and expanded upon in Chapter 6, associated a total of nine partial or complete biosynthetic pathways to gene clusters in *C. uncialis*. Given the rapid expansion of genetic information available through next-generation sequencing platforms, and continued successes in pathway deduction in non-lichen species, applying these indirect methods may allow us to expand what we know about lichen secondary metabolite biosynthesis. Of course, definitive assignment of function by functional heterologous expression would be the preferable course of action, but with that option presently unavailable, these indirect methods may provide some recourse to advancing our understanding of secondary metabolite biosynthesis in lichens.

Activating the expression of cryptic gene clusters is a possibility that may allow us to identify new natural products and link those products to their genetic origin. In model organisms, metabolic engineering techniques such as ribosome engineering or overproducing malonyl-CoA have been used to discover new natural products (Hosaka et al. 2009; Ye et al. 2017; Zabala et al. 2013). Manipulating transcriptional regulators to activate cryptic biosynthetic genes is another example (Aigle & Corre, 2012; Suroto et al. 2017). Recently, a proof-of-concept experiment demonstrating the use of CRISPR-Cas9 to activate a silent biosynthetic gene cluster demonstrates the possibility of using this technology towards natural products discovery (Zhang et al. 2017).

Alas, to apply any of these techniques will require precision manipulation of the lichen genome – an accomplishment that has yet to be reported in the literature. We are still a long way away from developing a genetic toolbox for lichens that is comparable to the vast repertoire of techniques and procedures currently available for model bacteria and fungi. However, ‘low-tech’ tools are also available. The sensitivity of an organism’s chemical profile to a new stimulus can be exploited to activate otherwise silent genes (Bode et al. 2002). Successful examples of metabolite discovery in non-lichen fungi and bacteria have been reported by culturing multiple organisms together or by culturing organisms in the presence of rare-earth elements (Onaka et al. 2011; Tanaka et al. 2010). Examples of stimulus-based natural products discovery among lichens include changes to carbon source or applying stressors such as desiccation or UV radiation (Brunauer et al. 2007; Deduke et al. 2012; Elshobary et al. 2016; Stocker- Wörgötter et al. 2004; Stocker-Wörgötter & Hager, 2008; Timsina et al. 2013). Finding a means of reproducibly inducing grayanic acid biosynthesis (a technique that was later exploited by Armaleo et al. (2011) to find the grayanic acid gene cluster) is an example of the utility of this work (Culberson & Armaleo, 1992).

Although examples of precisely manipulating the lichen genome have yet to have been reported in the literature, there is a singular example of a technique employing random mutagenesis. A relatively fast-growing species of lichen named *Umbilicaria muehlenbergii* was modified by random insertional mutagenesis using a tumour-inducing plasmid produced by the plant pathogen *Agrobacterium tumefaciens* (Park et al. 2013c). In this study, a total of 918 transformants were produced. Phenotype changes in this library were immediately observable with respect to colony colour, size, and morphology. I hypothesize that such a mutational technique could also be used to discover new natural products. If genes remain silent because of a repressor, the disruption of that repressor would cause the expression of that gene to activate.

One consequence of random mutagenesis could therefore be the activation of otherwise silent biosynthetic gene clusters and the emergence of new natural products within the lichen's chemical portfolio. For example, if a similar library of transformants were to be generated for *C. uncialis*, it would then be possible to screen those transformants for new secondary metabolites and elucidate their molecular structures. Equipped with our annotated biosynthetic gene clusters, we could then use a deductive approach similar to our approach to usnic acid to find its originating genes. As a proof of concept, it would be a useful exercise to screen the *U. muehlenbergii* library generated by Park et al. (2013) for evidence of natural products that were biosynthesized as a consequence of random mutagenesis.

9.4. Summary

The Manitoba lichen *Cladonia uncialis* and the secondary metabolite usnic acid was chosen as the ideal species and metabolite to explore heterologous expression protocols (Chapter 1). Thanks to the maturation of next-generation sequencing platforms, we were able to sequence the whole-genome of *C. uncialis* and find the usnic acid gene cluster (Chapter 3). Using a 'homology mapping' approach we were able to propose biosynthetic pathways associated with 9 of the 46 gene clusters identified within *C. uncialis* (Chapters 4-6). This work provided us with a rich understanding of the genetic secondary metabolome of *C. uncialis* and a foundation upon which we were able to explore its metabolome through functional heterologous expression. The NSAR1 *A. oryzae* expression system is well-regarded among natural products chemists for its use in the characterization of secondary metabolite gene clusters in *Fungi* and was therefore chosen to explore the secondary metabolome of *C. uncialis*. However, we were unable to meet the latter objectives outlined in Chapter 1 as we were unable to observe *de novo* metabolite production in *A. oryzae* despite our efforts to screen a total of seven PKS genes from *C. uncialis* (Chapters 7-8).

These failures, combined with our successes using PKS originating from non-lichen fungi, led us to conclude that some presently unidentifiable quality among lichen PKS precludes their functional expression in *A. oryzae*. Numerous control experiments were therefore conducted in our attempt to troubleshoot the problem and rule out its potential causes (Chapter 7-8). Although this research has greatly expanded the amount of genetic information available on secondary metabolite biosynthesis in lichens, we were unable to achieve our core objective of exploring its secondary metabolome through functional heterologous expression. I must conclude, for reasons presently inexplicable to me, that NSAR1 *A. oryzae* is an ill-suited host to express lichen biosynthetic genes. Alternative hosts such as *A. nidulans*, *S. cerevisiae* or *E. coli* may produce positive outcomes.

Appendix

References

1. Abdel-Hamed M. 2015. *Polyketide biosynthesis in lichen fungi Cladonia uncialis*. Ph.D. thesis, University of Manitoba.
2. Abdel-Hameed M, Bertrand RL, Donald LJ, Sorensen JL. 2018. Lichen ketosynthase domains are not responsible for inoperative polyketide synthases in *Ascomycota* hosts. *Biochem. Biophys. Res. Commun.* 503: 1228-1234
3. Abdel-Hameed M, Bertrand RL, Piercey-Normore M, Sorensen JL. 2016b. The identification of 6-hydroxymellein synthase and accessory genes in the lichen *Cladonia uncialis*. *J. Nat. Prod.* 79: 1645-1650.
4. Abdel-Hameed M, Bertrand RL, Piercey-Normore MD, Sorensen JL. 2016a. Putative identification of the usnic acid biosynthetic gene cluster by *de novo* whole-genome sequencing of a lichen-forming fungus. *Fungal Biol.* 120: 306-316.
5. Abdullah ST, Hamid H, Ali M, Ansari SH, Alam MS. 2007. The new terpenes from the lichen *Parmelia perlata*. *Ind. J. Chem.* 46: 173-176.
6. Ahad AM, Goto Y, Kiuchi F, Tsuda Y, Kondo K, Sato T. 1991. Nematocidal principles in “oakmoss absolute” and nematocidal activity of 2, 4-dihydroxybenzoates. *Chem. Pharm. Bull.* 39: 1043–1046.
7. Aigle B, Corre C. 2012. Waking up *Streptomyces* secondary metabolism by constitutive expression of activators or genetic disruption of repressors. *Methods Enzymol.* 517: 343-366.
8. Aigle B, Lautru S, Spiteller D, Dickschat JS, Challis GL, Leblond P, Pernodet JL. 2014. Genome mining of *Streptomyces ambofaciens*. *J. Ind. Microbiol. Biotechnol.* 41: 251-263.
9. Akey DL, Razelun JR, Tehranisa J, Sherman DH, Gerwick WH, Smith JL. 2010. Crystal structure of dehydratase domains from the curacin polyketide biosynthetic pathway. *Structure* 18: 94-105.
10. Alberti F, Foster GD, Bailey AM. 2017. Natural products from filamentous fungi and production by heterologous expression. *Appl. Microbiol. Biotechnol.* 101: 493-500.
11. Alfatafta AA, Dowd PF, Gloer JB, Wicklow DT. 1998. *Carbonarin Antiinsectan Metabolites*. US Patent 5519052.

12. Alhawatemala MS, Gebril S, Cook D, Creamer R. 2017. RNAi-mediated down-regulation of a melanin polyketide synthase (pks1) gene in the fungus *Slafractonia leguminicola*. *World J. Microbiol. Biotechnol.* 33: 179.
13. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J. Mol. Biol.* 215: 403-410.
14. Anyaehie UB. 2009. Medicinal properties of fractionated acetone/water neem (*Azadirachta indica*) leaf extract from Nigeria: A review. *Nigerian J. Physiol. Sci.* 24: 157-159.
15. Anyaogu DC, Mortensen UH. 2015. Heterologous production of fungal secondary metabolites in *Aspergilli*. *Front. Microbiol.* 6: 77.
16. Aparicio JF, A.J. Colina, E. Ceballos, Martin JF. 1999. The biosynthetic gene cluster for the 26-membered ring polyene macrolide pimaricin: A new polyketide synthase organization encoded by two subclusters separated by functionalization genes. *J. Biol. Chem.* 274: 10133-10139.
17. Araújo AA, de Melo MG, Rabelo TK, Nunes PS, Santos SL, Serafini MR, Santos MR, Quintans-Júnior LJ, Gelain DP. 2015. Review of the biological properties and toxicity of usnic acid. *Nat. Prod. Res.* 29: 2167-2180.
18. Armaleo D, May S. 2009. Sizing the fungal and algal genomes of the lichen *Cladonia grayi* through quantitative PCR. *Symbiosis* 49: 43-51.
19. Armaleo D, Sun X, Culberson C. 2011. Insights from the first putative biosynthetic gene cluster for a lichen depside and depsidone. *Mycologia* 103: 741-754.
20. Armstrong R. 2004. Lichens, lichenometry and global warming. *Microbiologist* 2004: 32-35.
21. Artigot MP, Loiseau N, Laffitte J, Mas-Requieq L, Tadriss S, Oswald IP, Puel O. 2009. Molecular cloning and functional characterization of two CYP619 cytochrome P450s involved in biosynthesis of patulin in *Aspergillus clavatus*. *Microbiology* 155: 1738-1747.
22. Austin MB, Noel JP. 2003. The chalcone synthase superfamily of type III polyketide synthases. *Nat. Prod. Rep.* 20: 79-110.
23. Bailey AM, Alberti F, Kilaru S, Collins CM, de Mattos-Shipley K, Hartley AJ, Hayes P, Griffin A, Lazarus CM, Cox RJ, Willis CL, O'Dwyer K, Spence DW, Foster GD. 2016. Identification and manipulation of the pleuromutilin gene cluster from *Clitopilus passeckerianus* for increased rapid antibiotic production. *Sci. Rep.* 6: 25202

24. Bailey AM, Cox RJ, Harley K, Lazarus CM, Simpson TJ, Skellam E. 2007. Characterization of 3-methylorcinolaldehyde synthase (MOS) in *Acremonium strictum*: First observation of a reductive release mechanism during polyketide biosynthesis. *Chem. Commun.* 39: 4053-4055.
25. Balakrishnan B, Karki S, Chiu SH, Kim HJ, Suh JW, Nam B, Yoon YM, Chen CC, Kwon HJ. 2013. Genetic localization and *in vivo* characterization of a *Monascus* azaphilone pigment biosynthetic gene cluster. *Appl. Microbiol. Biotechnol.* 97: 6337-6345.
26. Baldauf SL. 2003. Phylogeny for the faint of heart: A tutorial. *Trends Genet.* 19: 345-351.
27. Baltz RH. 2010. *Streptomyces* and *Saccharopolyspora* hosts for heterologous expression of secondary metabolite gene clusters. *J. Ind. Microbiol. Biotechnol.* 37: 759-772.
28. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* 19: 455-477.
29. Barton DHR, Deflorin AM, Edwards OE. 1956. The synthesis of usnic acid. *J. Chem. Soc.* 1956: 530-534.
30. Bates ST, Cropsey GWG, Caporaso JG, Knight R, Fierer N. 2011. Bacterial communities associated with the lichen symbiosis. *Appl. Environ. Microbiol.* 77: 1309-1314.
31. Bayir Y, Odabasoglu F, Cakir A, Aslan A, Suleyman H, Halici M, Kazaz C. 2006. The inhibition of gastric mucosal lesion, oxidative stress and neutrophil- infiltration in rats by the lichen constituent diffractaic acid. *Phytomedicine* 13:584–590.
32. Bayly CL, Yadav VG. 2017. Towards precision engineering of canonical polyketide synthase domains: Recent advances and future prospects. *Molecules* 22: e235
33. Beck A, P.K. Divakar, N. Zhang, Molina MC, Struwe K. 2015. Evidence of ancient horizontal gene transfer between fungi and the terrestrial alga *Trebouxia*. *Org. Divers. Evol.* 15: 235-248.
34. Beck J, Ripka S, Siegner A, Schiltz E, Schweizer E. 1990. The multifunctional 6-methylsalicylic acid synthase gene of *Penicillium patulum*. Its gene structure relative to that of other polyketide synthases. *Eur. J. Biochem.* 192: 487-498.
35. Bedford DJ, Schweizer E, Hopwood DA, Khosla C. 1995. Expression of a functional fungal polyketide synthase in the bacterium *Streptomyces coelicolor* A3(2). *J. Bacteriol.* 177: 4544-4548.

36. Beiggi S, Piercey-Normore MD, 2007. Evolution of ITS ribosomal RNA secondary structures in fungal and algal symbionts of selected species of *Cladonia* sect. *Cladonia* (Cladoniaceae, Ascomycotina). *J. Mol. Evol.* 64: 528–542.
37. Beld J, Sonnenschein EC, Vickery CR, Noel JP, Burkart MD. 2013. The phosphopantetheinyl transferases: Catalysis of a post-translational modification crucial for life. *Nat. Prod. Rep.* 31: 61-108.
38. Benedict JB. 2009. A review of lichenometric dating and its applications to archaeology. *Am. Antiquity* 74: 143-172.
39. Bentley R. 2000. Mycophenolic acid: A one hundred year odyssey from antibiotic to immunosuppressant. *Chem. Rev.* 100: 3801-3826.
40. Bernhardt R. 2006. Cytochrome p450 as versatile biocatalysts. *J. Biotechnol.* 124:128-145.
41. Berthelot K, Estevez Y, Deffieux A, Peruch F. 2012. Isopentenyl diphosphate isomerase: A checkpoint to isoprenoid biosynthesis. *Biochimie* 94: 1621-1634.
42. Bertrand RL, Abdel-Hameed M, Sorensen JL. 2015. Limitations of the ‘ambush hypothesis’ at the single-gene scale: What codon biases are to blame? *Mol. Genet. Genomics* 290: 493-504.
43. Bertrand RL, Abdel-Hameed M, Sorensen JL. 2018a. Lichen biosynthetic gene clusters part I: Genome sequencing reveals a rich biosynthetic potential. *J. Nat. Prod.* 81: 723-731
44. Bertrand RL, Abdel-Hameed M, Sorensen JL. 2018b. Lichen biosynthetic gene clusters part II: Homology mapping suggests a functional diversity. *J. Nat. Prod.* 81: 732-748
45. Bertrand RL, Sorensen JL. 2018. A comprehensive catalogue of polyketide synthase gene clusters in lichenizing fungi. *J. Ind. Microbiol. Biotechnol.* 45: 1067-1081.
46. Billingsley JM, DeNichola AB, Tang Y. 2016. Technology development for natural product biosynthesis in *Saccharomyces cerevisiae*. *Curr. Opin. Biotechnol.* 42: 74-83.
47. Bingle LE, Simpson TJ, Lazarus CM. 1999. Ketosynthase domain probes identify two subclasses of fungal polyketide synthase genes. *Fungal Genet. Biol.* 26: 209-223.
48. Birch AJ, Donovan FW. 1953. Studies in relation to biosynthesis. I. Some possible routes to derivatives of orcinol and phloroglucinol. *Aus. J. Chem.* 6: 360-368.
49. Birch AJ, Massy-Westropp RA, Moye CJ. 1955. Studies in relation to biosynthesis. VII. 2-hydroxy-6-methylbenzoic acid in *Penicillium griseofulvum* Dierckx. *Aus. J. Chem.* 8: 539-544.

50. Bisang C, Long PF, Cortés J, Westcott J, Crosby J, Matharu AL, Cox RJ, Simpson TJ, Staunton J, Leadlay PF. 1999. A chain initiation factor common to both modular and aromatic polyketide synthases. *Nature* 401: 502-505.
51. Blackwell M. 2011. The fungi: 1,2,3,...5.1 million species? *Am. J. Bot.* 98: 426-438.
52. Blin K, Medema MH, Kazempour D, Fischbach MA, Breitling R, Takano E, Weber T. 2013. AntiSMASH 2.0 – a versatile platform for genome mining of secondary metabolite producers. *Nucleic Acids Res.* 41: 204-212.
53. Blin K, Wolf T, Chevrette MG, Lu X, Schwalen CJ, Kautsar SA, Suarez-Duran HG, de los Santos EL, Kim HU, Nave M, Dickschat JS, Mitchell DA, Shelest E, Breitling R, Takano E, Lee SY, Weber T, Medema MH. 2017. AntiSMASH 4.0 – Improvements in chemistry prediction and gene cluster boundary identification. *Nucleic Acids Res.* 45: W36-W41.
54. Bode HB, Bethe B, Höfs R, Zeeck A. 2002. Big effects from small changes: possible ways to explore nature's chemical diversity. *ChemBioChem* 3: 619–627.
55. Boettger D, Hertweck C. 2013. Molecular diversity sculpted by fungal PKS-NRPS hybrids. *ChemBioChem* 14: 28-42.
56. Bohman G. 1969. Chemical Studies on Lichens. 22. Anthraquinones from the lichen *Lasallia papulosa* var. *rubiginosa* and the fungus *Valsaria rubricosa*. *Acta Chem. Scand.* 23: 2241-2244.
57. Brachmann AO, Kirchner F, Kegler C, Kinski SC, Schmitt I, Bode HB. 2012. Triggering the production of the cryptic blue pigment indigoidine from *Photorhabdus luminescens*. *J. Biotechnol.* 157: 96-99.
58. Bradford MM. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* 72: 248-254.
59. Brandt A, de Vera JP, Onofri S, Ott S. 2014. Viability of the lichen *Xanthoria elegans* and its symbionts after 18 months of space exposure and simulated Mars conditions on the ISS. *Int. J. Astrobiol.* 14: 411-425.
60. Brautaset T, Sekurova ON, Sletta H, Ellingsen TE, Strøm AR, Valla S, Zotchev SB. 2000. Biosynthesis of the polyene antifungal antibiotic nystatin in *Streptomyces noursei* ATCC 11455: Analysis of the gene cluster and deduction of the biosynthetic pathway. *Chem. Biol.* 7: 395-403.
61. Brown S, Clastre M, Courdavault V, O'Connor SE. 2015. *De novo* production of the plant-derived alkaloid strictosidine in yeast. *Proc. Natl. Acad. Sci. U.S.A.* 112: 3205-3210.

62. Brunauer G, Hager A, Grube M, Türk R, Stocker-Wörgötter E. 2007. Alterations in secondary metabolism of aposymbiotically grown mycobionts of *Xanthoria elegans* and cultured resynthesis stages. *Plant Physiol. Biochem.* 45: 146-151.
63. Brunauer G, Muggia L, Stocker-Wörgötter E, Grube M. 2009. A transcribed polyketide synthase gene from *Xanthoria elegans*. *Mycol. Res.* 113: 82–92.
64. Bruun T, Lamvik A. 1971. Haemoventosin. *Acta Chem. Scand.* 25: 483-486.
65. Caboche S, Pupin M, Leclère V, Jacques P, Kuchero G. 2009. Structural pattern matching of nonribosomal peptides. *BMC Struct. Biol.* 9: 15.
66. Caboche S, Pupin M, Leclère, Fontaine A, Jacques P, Kuchero G. 2008. NORINE: A database of nonribosomal peptides. *Nucleic Acids Res.* 36: D326-D331.
67. Cacho RA, Tang Y, Chooi YH. 2015. Next-generation sequencing approach for connecting secondary metabolites to biosynthetic gene clusters in fungi. *Front. Microbiol.* 5: 774.
68. Calchera A, Dal Grande F, Bode HB, Schmitt I. 2019. Biosynthetic gene content of the ‘perfume lichens’ *Evernia prunastri* and *Pseudevernia furfuracea*. *Molecules* 24: 203
69. Calcott MJ, Ackerley DF, Knight A, Keyzers RA, Owen JG. 2018. Secondary metabolism in the lichen symbiosis. *Chem. Soc. Rev.* 47: 1730-1760.
70. Campbell CD, Vederas JC. 2010. Biosynthesis of lovastatin and related metabolites formed by fungal iterative PKS enzymes. *Biopolymers.* 93: 755-763.
71. Cernava T, Berg G, Grube M. 2016. High life expectancy of bacteria on lichens. *Microb. Ecol.* 72: 510-513.
72. Chan YA, Podevels AM, Kevany BM, Thomas MG. 2009. Biosynthesis of polyketide synthase extender units. *Nat. Prod. Rep.* 26: 90-114.
73. Chen FC, Chen CF, Wei RD. 1982. Acute toxicity of PR toxin, a mycotoxin from *Penicillium roqueforti*. *Toxicon* 20: 433-441.
74. Chen Y, Kelly EE, Masluk RP, Nelson CL, Cantu DC, Reilly PJ. 2011. Structural classification and properties of ketoacyl synthases. *Protein Sci.* 20: 1659-1667.
75. Chiang YM, Oakley BR, Keller NP, Wang CC. 2010. Unraveling polyketide synthesis in members of the genus *Aspergillus*. *Appl. Microbiol. Biotechnol.* 86: 1719-1736.
76. Cho KS, Lim Y, Lee K, Lee J, Lee JH, Lee IS. 2017. Terpenes from forests and human health. *Toxicol. Res.* 33: 97-106.

77. Chooi YH, Stalker DM, Davis MA, Fujii I, Elix JA, Louwhoff SH, Lawrie AC. 2008. Cloning and sequence characterization of a non-reducing polyketide synthase gene from the lichen *Xanthoparmelia semiviridis*. *Mycol. Res.* 112: 147-161.
78. Cohen PA, Towers GHN. 1996. Biosynthetic studies on chlorinated anthraquinones in the lichen *Nephroma laevigatum*. *Phytochemistry* 42: 1325-1329.
79. Correché ER, Carrasco M, Giannini F, Piovano M, Garbarino J, Enriz D. 2002. Cytotoxic screening activity of secondary lichen metabolites. *Act. Farm. Bon.* 21: 273-278.
80. Crawford JM, Dancy BCR, Hill EA, Udvary DW, Townsend CA. 2006. Identification of a starter unit acyl-carrier protein transacylase domain in an iterative type I polyketide synthase. *Proc. Natl. Acad. Sci. U.S.A.* 103: 16728-17633.
81. Crawford JM, Korman TP, Labonte JW, Vagstad AL, Hill EA, Kamari-Bidkorpheh O, Tsai SC, Townsend CA. 2009. Structural basis for biosynthetic programming of fungal aromatic polyketide cyclization. *Nature* 461: 1139-1143.
82. Crawford SD. 2015. Lichens used in traditional medicine. In: *Lichen secondary metabolites: Bioactive properties and pharmaceutical potential*. Branislav Ranković (ed). Springer, New York.
83. Croteau R, Ketchum REB, Long RM, Kaspera R, Wildung MR. 2006. Taxol biosynthesis and molecular genetics. *Phytochem. Rev.* 5: 75-97.
84. Culberson CF, Armaleo D. 1992. Induction of a complete secondary-product pathway in a cultured lichen fungus. *Exp. Mycol.* 16: 52-63.
85. Dal Grande F, Meiser A, Greshake Tzovaras B, Otte J, Ebersberger I, Schmitt I. 2018. The draft genome sequence of the lichen-forming fungus *Lasallia hispanica* (Frey) Sancho & A. Crespo. *Lichenologist* 50: 329–340.
86. De Vera JP, Möhlmann D, Butina F, Lorek A, Wernecke R, Ott S. 2010. Survival potential and photosynthetic activity of lichens under Mars-like conditions: A laboratory study. *Astrobiology* 10: 215-227.
87. Deduke C, Timsina B, Piercey-Normore MD. 2012. Effect of environmental change on secondary metabolite production in lichen-forming fungi. In: Young SS, Silvern SE (ed) *International perspectives on global environmental change*. InTech, Croatia, pp 197-230.
88. Dombrink-Kurtzman MA, Engberg AE. 2006. *Byssoschlamys nivea* with patulin-producing capability has an isoepoxydon dehydrogenase gene (*idh*) with sequence homology to *Penicillium expansum* and *P. griseofulvum*. *Mycol. Res.* 110: 1111-1118.

89. Dombrink-Kurtzman, M. 2009. Molecular Characterization of *Penicillium Griseofulvum* Genes Involved in Biosynthesis of the Mycotoxin Patulin [abstract]. *American Chemical Society National Meeting*. Ref. No. P-023. p. 445.
90. Dopazo J. 1994. Estimating errors and confidence intervals for branch lengths in phylogenetic trees by a bootstrap approach. *J. Mol. Evol.* 38: 300-304.
91. Doroghazi JR, Albright JC, Goering AW, Ju KS, Haines RR, Tchalukov KA, Labeda DP, Kelleher NL, Metcalf WW. 2014. A roadmap for natural product discovery based on large-scale genomics and metabolomics. *Nat. Chem. Biol.* 10: 963-968.
92. Du L, Lou L. 2010. PKS and NRPS release mechanisms. *Roy. Soc. Chem.* 27: 255-278.
93. Dutta S, Whicher JR, Hansen DA, Hale WA, Chemler JA, Congdon GR, Narayan ARH, Håkansson K, Sherman DH, Smith JL, Skinotis G. 2014. Structure of a modular polyketide synthase. *Nature* 510: 512-517.
94. Elena C, Ravasi P, Castelli ME, Peirú S, Menzella HG. 2014. Expression of codon optimized genes in microbial systems: Current industrial applications and perspectives. *Front. Microbiol.* 5: 21.
95. Elshobary ME, Osman ME, Abo-Shady AM, Komatsu E, Perreault H, Sorensen J, Piercey-Normore MD. 2016. Algal carbohydrates affect polyketide synthesis of the lichen-forming fungus *Cladonia rangiferina*. *Mycologia* 108: 646-656.
96. Engelberg R, Danielson A, Wang S, Singh M, Wai A, Sorensen J, Duan K, Hausner G, Kumar A. Creation of a drug-sensitive reporter strain of *Pseudomonas aeruginosa* as a tool for the rapid screening of antimicrobial products. *J. Microbiol. Methods* 152: 1-6.
97. Ernst-Russell MA, Elix JA, Chai CLL, Rive MJ, Wardlaw JH. 2000. The structure and stereochemistry of coronatoquinone, a new pyranonaphthazarin from the lichen *Pseudocyphellaria coronata*. *Aust. J. Chem.* 53: 303-306.
98. Ewing B, Green P. 1998. Base-calling of automated sequencer tracers using phred. II. Error probabilities. *Genome Res.* 8: 186-194.
99. Ewing B, Hillier L, Wendl MC, Green P. 1998. Base-calling of automated sequencing traces using phred. I. Accuracy assessment. *Genome Res.* 8: 175-185.
100. Felsenstein J. 1985a. Phylogenies and the comparative method. *Am. Nat.* 125: 1-15.
101. Felsenstein J. 1985b. Confidence limits on phylogenies: An approach using the bootstrap, *Evolution* 39: 783-791.
102. Fisch KM. 2013. Biosynthesis of natural products by microbial iterative hybrid PKS-NRPS. *RSC Adv.* 3: 18228-18247.

103. Fiser A. 2010. Template-based protein structure modeling. *Methods Mol. Biol.* 673: 73-94.
104. Fujii I, Mori Y, Watanabe A, Kubo Y, Tsuji G, Ebizuka Y. 2000. Enzymatic synthesis of 1,3,6,8-tetrahydroxynaphthalene solely from malonyl coenzyme A by a fungal iterative type I polyketide synthase PKS1. *Biochemistry* 39: 8853-8858.
105. Fujii R, Minami A, Tsukagoshi T, Sato N, Sahara T, Ohgiya S, Gomi K, Oikawa H. 2011. Total biosynthesis of diterpene aphidicolin, a specific inhibitor of DNA polymerase α : Heterologous expression of four biosynthetic genes in *Aspergillus oryzae*. *Biosci. Biotechnol. Biochem.* 75: 1813-1817.
106. Fukuda T, Tomoda H, Omura S. 2005. Citridones, new potentiators of antifungal miconazole activity, produced by *Penicillium* sp. FKI-1938. II. Structure elucidation. *J. Antibiot.* 58: 315-321.
107. Fuller KK, Chen S, Loros JJ, Dunlap JC. 2015. Development of the CRISPR/Cas9 system for targeted gene disruption in *Aspergillus fumigatus*. *Eukaryot. Cell* 14: 1073-1080.
108. Funa N, Awakawa T, Horinouchi S. 2007. Pentaketide resorcylic acid synthesis by type III polyketide synthase from *Neurospora crassa*. *J. Biol. Chem.* 282: 14476-14461.
109. Funa N, Ohnishi Y, Fujii I, Shibuya M, Ebuzuka Y, Horinouchi S. 1999. A new pathway for polyketide synthesis in microorganisms. *Nature* 400: 897-899.
110. Gagunashvili AN, Davidsson SP, Jónsson ZO, Andrésson OS. 2009. Cloning and heterologous transcription of a polyketide synthase gene from the lichen *Solorina crocea*. *Mycol. Res.* 113: 354-363.
111. Gaitatzis N, Silakowski B, Kunze B, Nordsiek G, Blöcker H, Höfle G, Müller R. 2002. The biosynthesis of the aromatic myxobacterial electron transport inhibitor stigmatellin is directed by a novel type of modular polyketide synthase. *J. Biol. Chem.* 277: 13082-13090.
112. Galanty A, Paško P, Podolak I. 2019. Enantioselective activity of usnic acid: A comprehensive review and future perspectives. *Phytochem. Rev.* [Epub ahead of print]
113. Gao X, Haynes SW, Ames BD, Wang P, Vien LP, Walsh CT, Tang Y. 2012. Cyclization of fungal nonribosomal peptides by a terminal condensation-like domain. *Nat. Chem. Biol.* 8: 823-830.
114. Gay DC, Gay G, Axelrod AJ, Jenner M, Kohlhaas C, Kampa A, Oldham NJ, Piel J, Keatinge-Clay AT. 2014. A close look at a ketosynthase from a *trans*-acyltransferase modular polyketide synthase. *Structure* 22: 444-451.

115. Gay DC, Wagner DT, Meinke JL, Zogzas CE, Gay GR, Keatinge-Clay AT. 2016. The LINKS motif zippers *trans*-acyltransferase polyketide synthase assembly lines into a biosynthetic megacomplex. *J. Struct. Biol.* 193: 196-205.
116. Gershenzon J, Dudareva N. 2007. The function of terpene natural products in the natural world. *Nat. Chem. Biol.* 3: 408-414.
117. Gidijala L, Kiel JA, Douma RD, Seifar RM, van Gulik WM, Bovenberg RA, Veenhuis M, van der Klei IJ. 2009. An engineered yeast efficiently secreting penicillin. *PLoS ONE* 4:e8317.
118. Gomez-Escribano JP, Bibb MJ. 2014. Heterologous expression of natural product biosynthetic gene clusters in *Streptomyces coelicolor*: From genome mining to manipulation of biosynthetic pathways. *J. Ind. Microbiol. Biotechnol.* 41: 425-431.
119. Gomi K, Iimura Y, Hara S. 1987. Integrative transformation of *Aspergillus oryzae* with a plasmid containing the *Aspergillus nidulans argB* gene. *Agric. Biol. Chem.* 51: 2549-2555.
120. González-Burgos E, Gómez-Serranillos MP. 2012. Terpene compounds in nature: A review of the potential antioxidant activity. *Curr. Med. Chem.* 19: 5319-5341.
121. Goodwin S, McPherson JD, McCombie R. 2016. Coming of age: Ten years of next-generation sequencing technologies. *Nat. Rev. Genet.* 17: 333-351.
122. Grube M, Cardinale M, de Castro JV Jr, Müller H, Berg G. 2009. Species-specific structural and functional diversity of bacterial communities in lichen symbioses. *ISME J.* 3: 1105-1115.
123. Grube M, DePriest PT, Gargas A, Hafellner J. 1995. DNA isolation from lichen ascomata. *Mycol. Res.* 99: 1321-1324.
124. Grube M, G. Berg, ÓS Andrésson, Dyer PS, Miao VPW, Vilhelmsson O. 2014. Lichen genomics: Prospects and progress. In: F. Martin (Ed.), *The Ecological Genomics of Fungi*. John Wiley and Sons, New York, pp 191-212.
125. Gu L, Eisman EB, Dutta S, Franzmann TM, Walter S, Gerwick WH, Skiniotis G, Sherman DH. 2011. Tandem acyl carrier proteins in the curacin biosynthesis pathway promotes consecutive multienzyme reactions with a synergistic effect. *Angew. Chem. Int. Ed.* 50: 2795-2798.
126. Gu L, Geders TW, Wang B, Gerwick WH, Håkansson K, Smith JL, Sherman DH. 2007. GNAT-like strategy for polyketide chain initiation. *Science* 318: 970-974.
127. Guengerich FP, Munro AW. 2013. Unusual cytochrome p450 enzymes and reactions. *J. Biol. Chem.* 288:17065-17073.

128. Guex N, Peitsch MC. 1997. SWISS MODEL and Swiss-PdbViewer: An environment for comparative protein modeling. *Electrophoresis* 18: 2714-2723.
129. Guo CJ, Wang CC. 2014. Recent advances in genome mining of secondary metabolites in *Aspergillus terreus*. *Front. Microbiol.* 5: 717.
130. Hansen BG, Mnich E, Nielsen KF, Nielsen JB, Nielsen MT, Mortensen UH, Larsen TO, Patil KR. 2012. Involvement of a natural fusion of a cytochrome P450 and a hydrolase in mycophenolic acid biosynthesis. *Appl. Environ. Microbiol.* 78: 4908-4913.
131. Hanson G, Collier J. 2018. Codon optimality, bias and usage in translation and mRNA decay. *Nat. Rev. Mol. Cell Biol.* 19: 20-30.
132. Hashimoto M, Nonaka T, Fujii I. 2014. Fungal type III polyketide synthases. *Nat. Prod. Rep.* 31: 1306-1317.
133. Hatakeyama M, Kitaoka T, Ichinose H. 2016. Heterologous expression of fungal cytochromes P450 (CYP5136A1 and CYP5136A3) from the white-rot basidiomycete *Phanerochaete chrysosporium*: Functionalization with cytochrome b5 in *Escherichia coli*. *Enzyme Microb. Technol.* 89: 7-14.
134. Hawranik DJ, Anderson KS, Simmonds R, Sorensen JL. 2009. The chemoenzymatic synthesis of usnic acid. *Bioorg. Med. Chem. Lett.* 19: 2383-2385.
135. He Y, Cox RJ. 2016. The molecular steps in citrinin biosynthesis in fungi. *Chem. Sci.* 7: 2119-2127.
136. Hedden P, Phillips AL, Rojas MC, Carrera E, Tudzynski B. 2001. Gibberellin biosynthesis in plants and fungi: A case of convergent evolution? *J. Plant Growth Regul.* 20: 319-331.
137. Helfrich EJ, Piel J. 2016. Biosynthesis of polyketides by *trans*-AT polyketide synthases. *Nat. Prod. Rep.* 33: 231-316.
138. Herbst DA, Townsend CA, Maier T. 2018. The architectures of iterative type I PKS and FAS. *Nat. Prod. Rep.* 35: 1046-1069.
139. Hertweck C, Luzhetskyy A, Rebets Y, Bechthold A. 2007. Type II polyketide synthases: Gaining a deeper insight into enzymatic teamwork. *Nat. Prod. Rep.* 24: 162-190.
140. Hertweck C. 2009. The biosynthetic logic of polyketide diversity. *Angew. Chem. Int. Ed.* 48: 4688-4716.
141. Hitchman TS, Schmidt EW, Trail F, Rarick MD, Linz JE, Townsend CA. 2001. Hexanoate synthase, a specialized type I fatty acid synthase in aflatoxin B1 biosynthesis. *Bioorg. Chem.* 29: 293-307.

142. Hosaka T, Ohnishi-Kaneyama M, Muramatsu H, Murakami K, Tsurumi Y, Kodani S, Yoshida M, Fujie A, Ochi K. 2009. Antibacterial discovery in actinomycetes strains with mutations in RNA polymerase or ribosomal protein S12. *Nat. Biotechnol.* 27: 462-464.
143. Hoshino T, Sato T. 2002. Squalene-hopene cyclase: Catalytic mechanism and substrate recognition. *Chem. Commun.* 2002: 291-301.
144. Huang AC, Kautsar SA, Hong YJ, Medema MH, Bond AD, Tantillo DJ, Osbourn A. 2017. Unearthing a sesterterpene biosynthetic repertoire in the *Brassicaceae* through genome mining reveals convergent evolution. *Proc. Natl. Acad. Sci. U.S.A.* 114: E6005-E6014.
145. Huang G, Zhang L, Birch RG. 2001. A multifunctional polyketide-peptide synthetase essential for albicidin biosynthesis in *Xanthomonas albilineans*. *Microbiology* 147: 631-642.
146. Huneck S, Yoshimura Y. 1996. *Identification of lichen substances*. Springer, New York
147. Iimura Y, Gomi K, Uzu H. 1987. Transformation of *Aspergillus oryzae* through plasmid-mediated complementation of the methionine-auxotrophic mutation. *Agric. Biol. Chem.* 51: 323-328.
148. Ikram NK, Zhan X, Pan XW, King BC, Simonsen HT. 2015. Stable heterologous expression of biologically active terpenoids in green plant cells. *Front. Plant Sci.* 6: 129.
149. Ishiuchi K, Nakazawa T, Ookuma T, Sugimoto S, Sato M, Tsunematsu Y, Ishikawa N, Noguchi H, Hotta K, Moriya H, Watanabe K. 2012. Establishing a new methodology for genome mining and biosynthesis of polyketides and peptides through yeast molecular genetics. *ChemBioChem* 13: 846-854.
150. Isin EM, Guengerich FP. 2007. Complex reactions catalyzed by cytochrome p450 enzymes. *Biochim. Biophys. Acta* 1770: 314-329.
151. Jackson DR, Shakya G, Patel AB, Mohammed LY, Vasilakis K, Wattana-Amorn P, Valentic TR, Milligan JC, Crump MP, Crosby J, Tsai SC. 2018. Structural and functional studies of the daunorubicin priming ketosynthase DpsC. *ACS Chem. Biol.* 13: 141-151.
152. Jennewein S, Park H, DeJong JM, Long RM, Bollon AP, Croteau RB. 2005. Coexpression in yeast of taxus cytochrome p450 reductase with cytochrome p450 oxygenases involved in taxol biosynthesis. *Biotechnol. Bioeng.* 89: 588-598.
153. Jiang H, Zirkle R, Metz JG, Braun L, Richter L, van Lanen SG, Shen B. 2008. The role of tandem-acyl carrier protein domains in polyunsaturated fatty acid biosynthesis. *J. Am. Chem. Soc.* 130: 6336-6337.

154. Jin FJ, Maruyama J, Juvvadi PR, Arioka M, Kitamoto K. 2004. Adenine auxotrophic mutants of *Aspergillus oryzae*: Development of a novel transformation system with triple auxotrophic hosts. *Biosci. Biotech. Biochem.* 68: 656-662.
155. Jongedijk E, Cankar K, Buchhaupt M, Schrader J, Bouwmeester H, Beekwilder J. 2016. Biotechnological production of limonene in microorganisms. *Appl. Microbiol. Biotechnol.* 100: 2927-2938.
156. Jørgensen SH, Frandsen RJ, Nielsen KF, Lysøe E, Sondergaard TE, Wimmer R, Giese H, Sørensen JL. 2014. *Fusarium graminearum* PKS14 is involved in orsellinic acid and orcinol biosynthesis. *Fungal Genet. Biol.* 70: 24-31.
157. Junttila S, Rudd S. 2012. Characterization of a transcriptome from a non-model organism, *Cladonia rangiferina*, the grey reindeer lichen, using high-throughput next generation sequencing and EST sequence data. *BMC Genomics* 13: 575.
158. Kaasalainen U, Fewer DP, Jokela J, Wahlsten M, Sivonen K, Rikkinen J. 2012. Cyanobacteria produce a high variety of hepatotoxic peptides in lichen symbiosis. *Proc. Natl. Acad. Sci. U.S.A.* 109: 5886-5891.
159. Kampa A, Gagunashvili AN, Gulder TA, Morinaka BI, Daolio C, Godejohann M, Miao VP, Piel J, Andrésson Ó. 2013. Metagenomic natural product discovery in lichen provides evidence for a family of biosynthetic pathways in diverse symbioses. *Proc. Natl. Acad. Sci. U.S.A.* 110: E3129-E3137.
160. Kawaide H, Imai R, Sassa T, Kamiya Y. 1997. Ent-kaurene synthase from the fungus *Phaeosphaeria* sp. L487. cDNA isolation, characterization, and bacterial expression of a bifunctional diterpene cyclase in fungal gibberellin biosynthesis. *J. Biol. Chem.* 272: 21706-21712.
161. Kealey JT, Liu L, Santi DV, Betlach MC, Barr PJ. 1998. Production of a polyketide natural product in nonpolyketide-producing prokaryotic and eukaryotic hosts. *Proc. Natl. Acad. Sci. U.S.A.* 95: 505-509.
162. Kears M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Mentjies P, Drummond A. 2012. Geneious basic: An integrated and extendable desktop software program for the organization and analysis of sequence data. *Bioinformatics* 28: 1647-1649.
163. Keatinge-Clay AT, Stroud RM. 2006. The structure of a ketoreductase determines the organization of the β -carbon processing enzymes of modular polyketide synthases. *Structure* 14: 737-748.
164. Keatinge-Clay AT. 2012. The structures of type I polyketide synthases. *Nat. Prod. Rep.* 29: 1050-1073.

165. Kennedy J, Auclair K, Kendrew SG, Park C, Vederas JC, Hutchinson R. 1999. Modulation of polyketide synthase activity by accessory proteins during lovastatin biosynthesis. *Science* 284: 1368-1372.
166. Khaldi N, Seifuddin FT, Turner G, Haft D, Nierman WC, Wolfe KH, Fedorova ND. 2010. SMURF: Genomic mapping of fungal secondary metabolite clusters. *Fungal Genet. Biol.* 47: 736-741.
167. Kim DH, Kim BG, Lee HJ, Lim Y, Hur HG, Ahn JH. 2005. Enhancement of isoflavone synthase activity by co-expression of p450 reductase from rice. *Biotechnol. Lett.* 27: 1291-1294.
168. Kim JA, Hong SG, Cheong YH, Koh YJ, Hur JS. 2012. A new reducing polyketide synthase gene from the lichen-forming fungus *Cladonia metacorallifera*. *Mycologia* 104: 362-370.
169. Kittilä T, Cryle MJ. 2016. Capturing the structure of the substrate bound condensation domain. *Cell Chem. Biol.* 23: 315-316.
170. Kittilä T, Mollo A, Charkoudian LK, Cryle MJ. 2016. New structural data reveal the motion of carrier proteins in nonribosomal peptide synthesis. *Angew. Chem. Int. Ed.* 55: 9834-9840.
171. Kohli RM, Trauger JW, Schwarzer D, Marahiel MA, Walsh CT. 2001. Generality of peptide cyclization catalyzed by isolated thioesterase domains of nonribosomal peptide synthetases. *Biochemistry* 40: 7099-7108.
172. Kopp F, Marahiel MA. 2007. Macrocyclization strategies in polyketide and nonribosomal peptide biosynthesis. *Nat. Prod. Rep.* 24: 735-749.
173. Korman TP, Crawford JM, Labonte JW, Newman AG, Wong J, Townsend CA, Tsai SC. 2010. Structure and function of an iterative polyketide synthase thioesterase domain catalyzing Claisen cyclization in aflatoxin biosynthesis. *Proc. Natl. Acad. Sci. U.S.A.* 107: 6246-6251.
174. Kristjan-Bloudoff T, Schmeing M. 2017. Structural and functional aspects of the nonribosomal peptide synthetase condensation domain superfamily: Discovery, dissection and diversity. *Biochim. Biophys. Acta* 1865: 1587-1604.
175. Ku J, Mirmira RG, Liu L, Santi DV. 1997. Expression of a functional non-ribosomal peptide synthetase module in *Escherichia coli* by coexpression with a phosphopantetheinyl transferase. *Chem. Biol.* 4: 203-207.
176. Kumar KCS, Müller K. 1999a. Lichen metabolites. 1. Inhibitory action against leukotriene B4 biosynthesis by a nonredox mechanism. *J Nat. Prod.* 62: 817-820.

177. Kumar KCS, Müller K. 1999b. Lichen metabolites. 2. Antiproliferative and cytotoxic activity of gyrophoric, usnic, and diffractaic acid on human keratinocyte growth. *J. Nat. Prod.* 62: 821–823.
178. Kumar S, Stecher G, Tamura K. 2016. MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33: 1870-1874.
179. Kwan DH, Leadlay PF. 2010. Mutagenesis of a modular polyketide synthase enoylreductase domain reveals insights into catalysis and stereospecificity. *ACS Chem. Biol.* 5: 829-838.
180. Kwon HJ, Smith WC, Scharon AJ, Hwang SH, Kurth MJ, Shen B. 2002. C-O bond formation by polyketide synthases. *Science* 297: 1327-1330.
181. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG. 2007. Clustal W and Clustal X version 2.0. *Bioinformatics* 23: 2947-2948.
182. Lee KH, Kim DM. 2018. Recent advances in development of cell-free protein synthesis systems for fast and efficient production of recombinant proteins. *FEMS Microbiol. Lett.* 365: fny174.
183. Lee TV, Johnson LJ, Johnson RD, Koulman A, Lane GA, Lott JS, Arcus VL. 2010. Structure of a eukaryotic non-ribosomal peptide synthetase adenylation domain that activates a large hydroxamate amino acid in siderophore biosynthesis. *J. Biol. Chem.* 285: 2415-2427.
184. Li D, Zhang Q, Zhou Z, Zhao F, Lu W. 2016. Heterologous biosynthesis of triterpenoid dammarenediol-II in engineered *Escherichia coli*. *Biotechnol. Lett.* 38: 603-609.
185. Li J, Neubauer P. 2014. *Escherichia coli* as a cell factory for heterologous production of nonribosomal peptides and polyketides. *New Biotechnol.* 31: 579-585.
186. Li P, Oh DY, Bandyopadhyay G, Lagakos WS, Talukdar S, Osborn O, Johnson A, Chung H, Maris M, Ofrecio JM, Taguchi S, Lu M, Olefsky JM. 2015. LTB4 promotes insulin resistance in obese mice by acting on macrophages, hepatocytes and myocytes. *Nat. Med.* 21: 239-247.
187. Liao C, Piercey-Normore MD, Sorensen JL, Gough K. 2010. *In situ* imaging of usnic acid in selected *Cladonia* spp. by vibrational spectroscopy. *Analyst* 135: 3242-3248.
188. Lim YP, Go MK, Yew WS. 2016. Exploiting the biosynthetic potential of type III polyketide synthases. *Molecules* 21: 806.

189. Linnemannstöns P, Prado MM, Fernández-Martin R, Tudzynski B, Avalos J. 2002. A carotenoid biosynthesis gene cluster in *Fusarium fujikuroi*: The genes *carB* and *carRA*. *Mol. Genet. Genomics* 267: 593-602.
190. Liu T, Sanchez JF, Chiang YM, Oakley BR, Wang CC. 2014. Rational domain swaps reveal insights about chain length control by ketosynthase domains in fungal nonreducing polyketide synthases. *Org. Lett.* 16: 1676-1679.
191. Loeschcke A, Thies S. 2015. *Pseudomonas putida*: A versatile host for the production of natural products. *Appl. Microbiol. Biotechnol.* 99: 6197-6214.
192. Luo H, Yamamoto Y, Kim JA, Jung JS, Koh YJ, Hur JS. 2009. Lecanoric acid, a secondary lichen substance with antioxidant properties from *Umbilicaria antarctica* in maritime Antarctica (King George Island). *Polar Biol.* 32: 1033-1040.
193. Luo Y, Enghiad B, Zhao H. 2017. New tools for reconstruction and heterologous expression of natural product biosynthetic gene clusters. *Nat. Prod. Rep.* 33: 174-182.
194. Luo Y, Li BZ, Liu D, Zhang L, Chen Y, Jia B, Zeng BX, Zhao H, Yuan YJ. 2015. Engineered biosynthesis of natural products in heterologous hosts. *Chem. Soc. Rev.* 44: 5265-5290.
195. Lutzoni F, Miadlikowska J. 2009. Lichens. *Curr. Biol.* 19: R502-R503.
196. Mackinnon S, Durst T, Arnason JT, Angerhofer C, Pezzuto J, Sanchez-Winday PE, Poveda LJ, Gbeassor M. 1997. Antimalarial activity of tropical meliaceae extracts and gedunin derivatives. *J. Nat. Prod.* 60: 336-341.
197. Maier T, Leibundgut M, Ban N. 2008. The crystal structure of a mammalian fatty acid synthase. *Science* 321: 1315-1322.
198. Martinez-Núñez MA, López y López VE. 2016. Nonribosomal peptides synthetases and their applications in industry. *Sust. Chem. Proc.* 4: 13.
199. Masschelien J, Mattheus W, Gao LJ, Moons P, van Houdt R, Uytterhoeven B, Lamberigts C, Lescrinier E, Rozenski J, Herdewijn P, Aertsen A, Michiels C, Lavigne R. 2013. A PKS/NPRS/FAS hybrid gene cluster from *Serratia plymuthica* RVH1 encoding the biosynthesis of three broad spectrum, zeamine-related antibiotics. *PLoS ONE* 8: e54143.
200. Matsuda Y, Awakawa T, Abe I. 2013. Reconstituted biosynthesis of fungal meroterpenoid andrastin A. *Tetrahedron* 69: 8199-9204.
201. Matsuda Y, Bai T, Phippen CBW, Nødvig CS, Kjaerbølling, Vesth TC, Andersen MR, Mortensen UH, Gotfredsen CH, Abe I, Larsen TO. 2018. Novofumigatonin biosynthesis

- involves a non-heme iron-dependent endoperoxide isomerase for orthoester formation. *Nat. Commun.* 9: 2587.
202. Matsuda Y, Iwabuchi T, Fujimoto T, Awakawa T, Nakashima Y, Mori T, Zhang H, Hayashi F, Abe I. 2016. Discovery of key dioxygenases that diverged the paraherquonin and acetoxylhydroaustin pathways in *Penicillium brasilianum*. *J. Am. Chem. Soc.* 138: 12671-12677.
 203. Matsuda Y, Iwabuchi T, Wakimoto T, Awakawa T, Abe I. 2015. Uncovering the unusual D-ring construction in terretinin biosynthesis by collaboration of a multifunctional cytochrome p450 and a unique isomerase. *J. Am. Chem. Soc.* 137: 3393-3401.
 204. Matzer M, Mayrhofer H, Elix JA. 1998. *Rinodina peloleuca* (Physciaceae), a maritime lichen with a distinctive austral distribution. *N.Z. J. Bot.* 36: 175-188.
 205. Mbah AU, Udeinya II, Shu EN, Chijioke CP, Nubila T, Udeinya F, Muobuikwe A, Mmuobieri A, Obioma MS. 2007. Fractionated neem leaf extract is safe and increased CD4⁺ cell levels in HIV/AIDS patients. *Am. J. Ther.* 14: 369-374.
 206. Medema MH, Blin K, Cimermancic P, de Jager V, Zakrzewski P, Fischbach MA, Weber T, Takano E, Breitling R. 2011. AntiSMASH: Rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. *Nucleic Acids Res.* 39: W339-W346.
 207. Medema MH, Kottmann R, Yilmaz P, Cummings M, Biggins JB, Blin K, de Bruijn I, Chooi YH, Claesen J, Coates RC, Cruz-Morales P, Duddela S, Dusterhus S, Edwards DJ, Fewer DP, Garg N, Geiger C, Gomez-Escribano JP, Greule A, Hadjithomas M, Haines AS, Helfrich EJ, Hillwig ML, Ishida K, Jones AC, Jones CS, Jungmann K, Kegler C, Kim HU, Kötter P, Krug D, Masschelein J, Melnik AV, Mantovani SM, Monroe EA, Moore M, Moss N, Nützmänn HW, Pan G, Pati A, Petras D, Reen FJ, Rosconi F, Rui Z, Tian Z, Tobias NJ, Tsunematsu Y, Wiemann P, Wyckoff E, Yan X, Yim G, Yu F, Xie Y, Aigle B, Apel AK, Balibar CJ, Balskus EP, Barona-Gómez F, Bechthold A, Bode HB, Borriss R, Brady SF, Brakhage AA, Caffrey P, Cheng YQ, Clardy J, Cox RJ, De Mot R, Donadio S, Donia MS, van der Donk WA, Dorrestein PC, Doyle S, Driessen AJ, Ehling-Schulz M, Entian KD, Fischbach MA, Gerwick L, Gerwick WH, Gross H, Gust B, Hertweck C, Höfte M, Jensen SE, Ju J, Katz L, Kaysser L, Klassen JL, Keller NP, Kormanec J, Kuipers OP, Kuzuyama T, Kyrpides NC, Kwon HJ, Lautru S, Lavigne R, Lee CY, Lincuan B, Liu X, Liu W, Luzhetskyy A, Mahmud T, Mast Y, Méndez C, Metsä-Ketelä M, Micklefield J, Mitchell DA, Moore BS, Moreira LM, Müller R, Neilan BA, Nett M, Nielsen J, O'Gara F, Oikawa H, Osbourn A, Osburne MS, Ostash B, Payne SM, Pernodet JL, Petricek M, Piel J, Ploux O, Raaijmakers JM, Salas JA, Schmitt EK, Scott B, Seipke RF, Shen B, Sherman DH, Sivonen K, Smanski MJ, Sosio M, Stegmann E, Süßmuth RD, Tahlan K, Thomas CM, Tang Y, Truman AW, Viaud M, Walton JD, Walsh CT, Weber T, van Wezel GP, Wilkinson B, Willey JM, Wohlleben W, Wright GD, Ziemert N,

- Zhang C, Zotchev SB, Breitling R, Takano E, Glöckner FO. 2015. Minimum information about a biosynthetic gene cluster. *Nat. Chem. Biol.* 11:625-631.
208. Metzger U, Schall C, Zocher G, Unsöld I, Stec E, Li SM, Heide L, Stehle T. 2009. The structure of dimethylallyl tryptophan synthase reveals a common architecture of aromatic prenyltransferases in fungi and bacteria. *Proc. Natl. Acad. Sci. U.S.A.* 106: 14309-14314.
 209. Miao V, Coëffet-LeGal M-F, Brown D, Sinnemann S, Donaldson G, Davies J. 2001. Genetic approaches to harvesting lichen products. *Trends Biotechnol.* 19: 349-355.
 210. Miller BR, Gulick AM. 2016. Structural biology of non-ribosomal peptide synthetases. *Methods Mol. Biol.* 1401: 3-29.
 211. Miyanaga A, Kudo F, Eguchi T. 2018. Protein-protein interactions in polyketide synthase-nonribosomal peptide synthetase hybrid assembly lines. *Nat. Prod. Rep.* 35: 1185-1209.
 212. Molnár K, Farkas E. 2010. Current results on biological activities of lichen secondary metabolites. A review. *Z. Naturforsch C.* 65: 157-173.
 213. Morita H, Shimokawa Y, Tanio M, Kato R, Noguchi H, Sugio S, Kohno T, Abe I. 2010. A structure-based mechanism for benzalacetone synthase from *Rheum palmatum*. *Proc. Natl. Acad. Sci. U.S.A.* 107: 669-673.
 214. Motamedi H, Shafiee A. 1998. The biosynthetic gene cluster for the macrolactone ring of the immunosuppressant FK506. *Eur. J. Biochem.* 256: 528-534.
 215. Muggia L, Grube M. 2010. Type III polyketide synthases in lichen mycobionts. *Fungal Biol.* 114: 379-385.
 216. Murphy AC, Hong H, Vance S, Broadhurst RW, Leadlay PF. 2016. Broadening substrate specificity of a chain-extending ketosynthase through a single active-site mutation. *Chem. Commun.* 52: 8373-8376.
 217. Nash III TH 2008. *Lichen Biology*, 2nd Ed Cambridge University Press, Cambridge.
 218. Neamati N, Hong H, Mazumder A, Wang S, Sunder S, Nicklaus MC, Milne GWA, Proksa B, Pommier Y. 1997. Depsides and depsidones as inhibitors of HIV-1 integrase: discovery of novel inhibitors through 3D database searching. *J. Med. Chem.* 40: 942-951.
 219. Nguyen KH, Chollet-Krugler M, Gouault N, Tomasi S. 2013. UV-protectant metabolites from lichens and their symbiotic partners. *Nat. Prod. Rep.* 30: 1490-1508.

220. Nicholson TP, Rudd BAM, Dawson M, Lazarus CM, Simpson TJ, Cox RJ. 2001. Design and utility of oligonucleotide gene probes for fungal polyketide synthases. *Chem. Biol.* 8: 157-178.
221. Nicolaou KC, Yang Z, Liu JJ, Ueno H, Nantermet PG, Guy RK, Claiborne CF, Renaud J, Couladouros EA, Paulvannan K, Sorensen EJ. 1994. Total synthesis of taxol. *Nature* 367: 630-634.
222. Nielsen JC, Grijseels S, Prigent S, Ji B, Dainat J, Nielsen KF, Frisvad JC, Workman M, Nielsen J. 2017. Global analysis of biosynthetic gene clusters reveals vast potential of secondary metabolite production in *Penicillium species*. *Nat. Microbiol.* 2: 17044.
223. Nofiani R, de Mattos-Shipley K, Lebe KE, Han LC, Iqbal Z, Bailey AM, Willis CL, Simpson TJ, Cox RJ. 2018. Strobilurin biosynthesis in Basidiomycete fungi. *Nat. Commun.* 9: 3940.
224. Okada M, Saito K, Wong CP, Li C, Wang D, Iijima M, Taura F, Kurosaki F, Awakawa T, Abe I. 2017. Combinatorial biosynthesis of (+)-daurichromenic acid and its halogenated analogue. *Org. Lett.* 19: 3183-3186.
225. Oldfield E, Lin FY. 2012. Terpene biosynthesis: Modularity rules. *Angew. Chem. Int. Ed.* 51: 1124-1137.
226. Omar S, Godard K, Ingham A, Hussain H, Wongpanich V, Pezzuto J, Durst T, Eklun C, Gbeassor M, Sanchez-Vindas P, Poveda L, Philogene BJR, Arnason JT. 2003. Antimalarial activities of gedunin and 7-methoxygedunin and synergistic activities with dillaiol. *Ann. Appl. Biol.* 143: 135-141.
227. Onaka H, Mori Y, Igarashi Y, Furumai T. 2011. Mycolic acid-containing bacteria induce natural-product biosynthesis in *Streptomyces* species. *Appl. Environ. Microbiol.* 77: 400-406.
228. Osmanova N, Schultze W, Ayoub N. 2010. Azaphilones: A class of fungal metabolites with diverse biological activities. *Phytochem. Rev.* 9: 315-342.
229. Øvstedal DO, Lewis-Smith RI. 2001. *Lichens of Antarctica and South Georgia: A Guide to their Identification and Ecology*. Cambridge: Cambridge University Press.
230. Paddon CJ, Westfall PJ, Pitera DJ, Benjamin K, Fisher K, McPhee D, Leavell MD, Tai A, Main A, Eng D, Polichuk DR, Teoh KH, Reed DW, Treynor T, Lenihan J, Fleck M, Bajad S, Dang G, Dengrove D, Diola D, Dorin G, Ellens KW, Fickes S, Galazzo J, Gaucher SP, Geistlinger T, Henry R, Hepp M, Horning T, Iqbal T, Jiang H, Kizer L, Lieu B, Melis D, Moss N, Regentin R, Secrest S, Tsuruta H, Vazquez R, Westblade LF, Xu L, Yu M, Zhang Y, Zhao L, Lievens J, Covello PS, Keasling JD, Reiling KK, Renninger NS, Newman JD. 2013. High-level semi-synthetic production of the potent antimalarial artemisinin. *Nature* 496: 528-532.

231. Park H, Kevany BM, Dyer DH, Thomas MG, Forest KT. 2014. A polyketide synthase acyltransferase domain structure suggests a recognition mechanism for its hydroxymalonlyl-acyl carrier protein substrate. *PLoS ONE* 9: e110965.
232. Park SY, Choi J, Kim JA, Jeong MH, Kim S, Lee YH, Hur JS. 2013a. Draft genome sequence of *Cladonia macilenta* KoLRI003786, a lichen-forming fungus producing biruloquinone. *Genome Announc.* 1: e00695-13.
233. Park SY, Choi J, Kim JA, Yu NH, Kim S, Kondratyuk SY, Lee YH, Hur JS. 2013b. Draft genome sequence of lichen-forming fungus *Caloplaca flavorubescens* strain KoLRI002931. *Genome Announc* 1: e00678-13.
234. Park SY, Choi J, Lee GW, Jeong MH, Kim JA, Oh SO, Lee YH, Hur JS. 2014a. Draft genome sequence of *Umbilicaria muehlenbergii* KoLRILF000956, a lichen-forming fungus amenable to genetic manipulation. *Genome Announc.* 2: e00357-14.
235. Park SY, Choi J, Lee GW, Kim JA, Oh SO, Jeong MH, Yu NH, Kim S, Lee YH, Hur JS. 2014b. Draft genome sequence of lichen-forming fungus *Cladonia metacorallifera* strain KoLRI002260. *Genome Announc.* 2: e01065-13.
236. Park SY, Choi J, Lee GW, Park CH, Kim JA, Oh SO, Lee YH, Hur JS. 2014c. Draft genome sequence of *Endocarpon pusillum* strain KoLRILF000583. *Genome Announc.* 2: e00452-14
237. Park SY, Jeong MH, Wang HY, Kim JA, Yu NH, Kim S, Cheong YH, Kang S, Lee YH, Hur JS. 2013c. *Agrobacterium tumefaciens*-mediated transformation of the lichen fungus, *Umbilicaria muehlenbergii*. *PLoS ONE* 8: e83896.
238. Pel HJ, de Winde JH, Archer DB, Dyer PS, Hofmann G, Schaap PJ, Turner G, de Vries RP, Albang R, Albermann K, Andersen MR, Bendtsen JD, Benen JA, van den Berg M, Breestraat S, Caddick MX, Contreras R, Cornell M, Coutinho PM, Danchin EG, Debets AJ, Dekker P, van Dijck PW, van Dijk A, Dijkhuizen L, Driessen AJ, d'Enfert C, Geysens S, Goosen C, Groot GS, de Groot PW, Guillemette T, Henrissat B, Herweijer M, van den Hombergh JP, van den Hondel CA, van der Heijden RT, van der Kaaij RM, Klis FM, Kools HJ, Kubicek CP, van Kuyk PA, Lauber J, Lu X, van der Maarel MJ, Meulenberg R, Menke H, Mortimer MA, Nielsen J, Oliver SG, Olsthoorn M, Pal K, van Peij NN, Ram AF, Rinas U, Roubos JA, Sagt CM, Schmoll M, Sun J, Ussery D, Varga J, Vervecken W, van de Vondervoort PJ, Wedler H, Wösten HA, Zeng AP, van Ooyen AJ, Visser J, Stam H. 2007. Genome sequencing and analysis of the versatile cell factory *Aspergillus niger* CBS 513.88. *Nat. Biotechnol.* 25: 221-231.
239. Penttilä A, Fales HM. 1966. On the biosynthesis in vitro of usnic acid. *Chem. Commun.* 1966: 656-657.
240. Pettit RK. 2009. Mixed fermentation for natural product drug discovery. *Appl. Microbiol. Biotechnol.* 83: 19–25.

241. Pichersky E, Raguso RA. 2018. Why do plants produce so many terpenoid compounds? *New Phytologist* 220: 692-702.
242. Piercey-Normore MD, Depriest PT. 2001. Algal switching among lichen symbioses. *Am. J. Bot.* 88: 1490-1498.
243. Polborn K, Steglich W, Connolly JD, Huneck S. 2014. Structure of the Macrocyclic Bis-lactone Lepranthin from the Lichen *Arthonia impolita*: an X-Ray Analysis. *Z. Naturforschung B.* 50: 1111-1114.
244. Proctor RH, Hohn TM. 1993. Aristolochene synthase. Isolation, characterization, and bacterial expression of a sesquiterpenoid biosynthetic gene (*AriI*) from *Penicillium roqueforti*. *J. Biol. Chem.* 268: 4543-4548.
245. Puel O, Galtier P, Oswald IP. 2010. Biosynthesis and toxicological effects of patulin. *Toxins* 2: 613-631.
246. Puel O, Tadrist S, Delaforge M, Oswald IP, Lebrihi A. 2007. The inability of *Byssoschlamys fulva* to produce patulin is related to absence of 6-methylsalicylic acid synthase and isoeopoxygen dehydrogenase genes. *Int. J. Food Microbiol.* 115: 131-139.
247. Puigbò P, Bravo IG, Garcia-Vallvé S. 2008. E-CAI: A novel server to estimate an expected value of Codon Adaptation Index (eCAI). *BMC Bioinformatics* 9: 65.
248. Puigbò P, Guzman E, Romeu A, Garcia-Vallvé S. 2007. OPTIMIZER: A web server for optimizing the codon usage of DNA sequences. *Nucleic Acids Res* 35: W126-W131.
249. Qi F, Lei C, Li F, Zhang X, Wang J, Zhang W, Fan Z, Li W, Tang G, Xiao Y, Zhao G, Li S. 2018. Deciphering the late steps of rifamycin biosynthesis. *Nat. Commun.* 9: 2342.
250. Qiao K, Chooi YH, Tang Y. 2011. Identification and engineering of the cytochalasin gene cluster from *Aspergillus clavatus* NRRL 1. *Metab. Eng.* 13: 723-732.
251. Quax TEF, Claassens NJ, Söll D, van der Oost J. 2015. Codon bias as a means to fine-tune gene expression. *Mol. Cell* 59: 149-161.
252. Rahman AS, Hothersall J, Crosby J, Simpson TJ, Thomas CM. 2005. Tandemly duplicated acyl carrier proteins, which increase polyketide antibiotic production, can apparently function either in parallel or in series. *J. Biol. Chem.* 280: 6399-6408.
253. Rausch C, Weber T, Kohlbacher O, Wohllensben W, Huson DH. 2005. Specificity prediction of adenylation domains in nonribosomal peptide synthetases (NRPS) using transductive support vector machines (TSVMs). *Nucleic Acids Res.* 33: 5799-5808.

254. Regueira TB, Kildegaard KR, Hansen BG, Mortensen UH, Hertweck C, Nielsen J. 2011. Molecular basis for mycophenolic acid biosynthesis in *Penicillium brevicompactum*. *Appl. Environ. Microbiol.* 77: 3035-3043.
255. Robbins T, Kapilivsky J, Cane DE, Khosla C. 2016. Roles of conserved active site residues in the ketosynthase domain of an assembly line polyketide synthase. *Biochemistry* 55: 4476-4484.
256. Robbins T, Liu YC, Cane DE, Khosla C. 2016. Structure and mechanism of assembly line polyketide synthases. *Curr. Opin. Struct. Biol.* 41: 10-18.
257. Roze LV, Hong SY, Linz JE. 2013. Aflatoxin biosynthesis: Current frontiers. *Annu. Rev. Food Sci. Technol.* 4: 293-311.
258. Rugbjerg P, Naesby M, Mortensen UH, Frandsen RJ. 2013. Reconstruction of the biosynthetic pathway for the core fungal polyketide scaffold rubrofusarin in *Saccharomyces cerevisiae*. *Microb. Cell Fact.* 12: 31.
259. Ružička L. 1953. The isoprene rule and the biogenesis of terpenic compounds. *Cell Mol. Life Sci.* 9: 357-367.
260. Rycroft DS, Connolly JD, Huneck S, Himmelreich UZ. 1995. Revised structure of haemoventosin. *Z. Naturforsch B.* 50: 1557-1563.
261. Rzhetsky A, Nei M. 1992. A simple method for estimating and testing minimum-evolution trees. *Mol. Biol. Evol.* 9: 945-967.
262. Saha S, Zhang W, Zhang G, Zhu Y, Chen Y, Liu W, Yuan C, Zhang Q, Zhang H, Zhang L, Zhang W, Zhang C. 2017. Activation and characterization of a cryptic gene cluster reveals a cyclization cascade for polycyclic tetramate macrolactams. *Chem. Sci.* 8: 1607-1612.
263. Saitou N, Nei M. 1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4: 406-425.
264. Sanchez JF, Chiang YM, Szewczyk E, Davidson AD, Ahuja M, Oakley CE, Woo Bok J, Keller N, Oakley BR, Wang CCC. 2010. Molecular genetic analysis of the orsellinic acid/F9775 gene cluster of *Aspergillus nidulans*. *Mol. Biosyst.* 6: 587-593.
265. Scarsdale JN, Kazanina G, He X, Reynolds KA, Wright HT. 2001. Crystal structure of the *Mycobacterium tuberculosis* β -ketoacyl-acyl carrier protein synthase III. *J. Biol. Chem.* 276: 20516-20522.
266. Scherlach K, Boettger D, Remme N, Hertweck C. 2010. The chemistry and biology of cytochalasans. *Nat. Prod. Rep.* 27: 869-886.

267. Schmitt I, Kautz S, Lumbsch HT. 2008. 6-MSAS-like polyketide synthase genes occur in lichenized ascomycetes. *Mycol. Res.* 112: 289-296.
268. Schmitt I, Lumbsch HT. 2009. Ancient horizontal gene transfer from bacteria enhances biosynthetic capabilities of fungi. *PLoS ONE* 4: e4437.
269. Schoch CL, Seifert KA, Huhndorf S, Robert V, Spouge JL, Levesque CA, Chen W, Fungal Barcoding Consortium. 2012. Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for *Fungi*. *Proc. Natl. Acad. Sci. U.S.A.* 109: 6241-6246.
270. Schor R, Schotte C, Wibberg D, Kalinowski J, Cox RJ. 2018. Three previously unrecognized classes of biosynthetic enzymes revealed during the production of xenovulene A. *Nat. Commun.* 9: 1963.
271. Schorn MA, Alanjary MM, Aquinaldo K, Korobeynikov A, Podell S, Patin N, Linecum T, Jensen PR, Ziemert N, Moore BS. 2016. Sequencing rare marine actinomycete genomes reveals high density of unique natural product biosynthetic gene clusters. *Microbiology* 162: 2075-2086.
272. Seemann M, Zhai G, de Kraker JW, Paschall CM, Christianson DW, Cane DE. 2002. Pentalenene synthase. Analysis of active site residues by site-directed mutagenesis. *J. Am. Chem. Soc.* 124: 7681-7689.
273. Seo C, Yim JH, Lee HK, Park SM, Sohn JH, Oh H. 2008. Stereocalpin A, a bioactive cyclic depsipeptide from the Antarctic lichen *Stereocaulon alpinum*. *Tetrahedron Lett.* 49: 29–31.
274. Seshime Y, Juvvadi PR, Fujii I, Kitamoto K. 2005. Discovery of a novel superfamily of type III polyketide synthases in *Aspergillus oryzae*. *Biochem. Biophys. Res. Commun.* 331: 253-260.
275. Sharp PM, Li WH. 1987. The codon adaptation index – A measure of directional synonymous codon usage bias, and its potential applications. *Nucleic Acids Res.* 15: 1281-1295.
276. Shen B. 2015. A new golden age of natural products drug discovery. *Cell* 163: 1297-1300.
277. Shimizu Y, Ogata H, Goto S. 2017. Type III polyketide synthases: Functional classification and phylogenomics. *ChemBioChem* 18: 50-65.
278. Shrestha G, St. Clair LL. 2013. Lichens: a promising source of antibiotic and anticancer drugs. *Phytochem. Rev.* 12: 229-224.
279. Shukla V, Joshi GP, Rawat MSM. 2010. Lichens as a potential natural source of bioactive compounds: A review. *Phytochem. Rev.* 9: 303-314.

280. Simpson TJ, Cox RJ. 2012. Polyketides in fungi. In: *Natural Products in Chemical Biology*, Civjan N (ed.) John Wiley & Sons: New Jersey, p. 141-161
281. Sinnemann SJ, Andrésson OS, Brown DW, Miao VP. 2000. Cloning and heterologous expression of *Solorina crocea* *pyrG*. *Curr. Genet.* 37: 333-338.
282. Skiba MA, Sikkema AP, Fiers WD, Gerwick WH, Sherman DH, Aldrich CC, Smith JL. 2016. Domain organization and active site architecture of a polyketide synthase C-methyltransferase. *ACS Chem. Biol.* 11: 3319-3327.
283. Skinnider MA, Merwin NJ, Johnston CW, Magarvey NA. 2017. PRISM 3: Expanded prediction of natural product chemical structures from microbial genomes. *Nucleic Acids Res.* 45: W49-W54.
284. Smith S, Tsai SC. 2007. The type I fatty acid and polyketide synthases. A tale of two megasynthases. *Nat. Prod. Rep.* 24: 1041-1072.
285. Snini SP, Tadriss S, Laffitte J, Jamin EL, Oswald IP, Puel O. 2014. The gene *patG* involved in the biosynthesis pathway of patulin, a food-borne mycotoxin, encodes a 6-methylsalicylic acid decarboxylase. *Int. J. Food Microbiol.* 171: 77-83.
286. Spribille T, Tuovinen V, Resl P, Vanderpool D, Wolinski H, Aime MC, Schneider K, Stabentheiner E, Toome-Heller M, Thor G, Mayrhofer H, Johannesson H, McCutcheon JP. 2016. Basidiomycete yeasts in the cortex of ascomycete macrolichens. *Science* 353: 488-492.
287. Spribille T. 2018. Relative symbiont input and the lichen symbiotic outcome. *Curr. Opin. Plant Biol.* 44: 57-63.
288. Stadler M, Anke H, Sterner OJ. 1995. New metabolites with nematocidal and antimicrobial activities from the ascomycete *Lachnum papyraceum* (Karst.) Karst. IV. Structural elucidation of novel isocoumarin derivatives. *J. Antibiot.* 48: 261– 266.
289. Stanke M, Morgenstern B. 2005. AUGUSTUS: A web server for gene prediction in eukaryotes that allows user-defined constraints. *Nucleic Acids Res.* 33: W465-W467.
290. Stanke M, Steinkamp R, Waack S, Morgenstern B. 2004. AUGUSTUS: A web server for gene finding in eukaryotes. *Nucleic Acids Res.* 32: W309-W312.
291. Starcevic A, Zucko J, Simunkovic J, Long PF, Cullum J, Hranueli D. 2008. ClustScan: An integrated program package for the semi-automated annotation of modular biosynthetic gene clusters and *in silico* prediction of novel chemical structures. *Nucleic Acids Res.* 36: 6882-6892.
292. Staunton J, Weissman KJ. 2001. Polyketide biosynthesis: A millennium review. *Nat. Prod. Rep.* 18: 380-416.

293. Steller S, Sokoll A, Bernhard F, Franke P, Vater J. 2004. Initiation of surfactin biosynthesis and the role of the SrfD-thioesterase protein. *Biochemistry* 43: 11331-11343.
294. Stephenson NL, Rundel PW. 1979. Quantitative variation and the ecological role of vulpinic acid and atranorin in thallus of *Letharia vulpina*. *Biochem. Syst. Ecol.* 7: 263-267.
295. Stocker-Wörgötter AE, Elix JA, Grube M. 2004. Secondary chemistry of lichen-forming fungi: Chemosyndromic variation and DNA-analyses of cultures and chemotypes in the *Ramalina farinacea* complex. *Bryologist* 107: 152-162.
296. Stocker-Wörgötter E, Hager A. 2008. Appendix: Culture methods for lichens and lichen symbionts. In: TH Nash III (ed) *Lichen Biology*, 2nd Ed. Cambridge University Press, Cambridge, p. 353-363.
297. Strieker M, Tanović A, Marahiel MA. 2010. Nonribosomal peptide synthetases: Structures and dynamics. *Curr. Opin. Struct. Biol.* 20: 234-240.
298. Studt L, Wiemann P, Kleigrew K, Humpf HU, Tudzynski B. 2012. Biosynthesis of fusarubins accounts for pigmentation of *Fusarium fujikuroi* perithecia. *Appl. Environ. Microbiol.* 78: 4468-4480.
299. Sun H, Liu Z, Zhao H, Ang EL. 2015. Recent advances in combinatorial biosynthesis for drug discovery. *Drug. Des. Devel. Ther.* 9: 823-833.
300. Suroto DA, Kitani S, Miyamoto KT, Sakihama Y, Arai M, Ikeda H, Nihira T. 2017. Activation of cryptic phthoxazolin A production in *Streptomyces avermitilis* by the disruption of autoregulator-receptor homologue AvaR3. *J. Biosci. Bioeng.* 124: 611-617.
301. Süssmuth RD, Mainz A. 2017. Nonribosomal peptide synthesis – Principles and prospects. *Angew. Chem. Int. Ed.* 56: 3770-3821.
302. Suzuki S, Hosoe T, Nozawa K, Kawai K, Yaguchi T, Udagawa S. 2000. Antifungal substances against pathogenic fungi, talaroconvolutins, from *Talaromyces convolutus*. *J. Nat. Prod.* 63: 768-772.
303. Taguchi H, Sankawa U, Shibata S. 1969. Biosynthesis of natural products. VI. Biosynthesis of usnic acid in lichens. A general scheme of biosynthesis of usnic acid. *Chem. Pharm. Bull.* 17: 2054-2060.
304. Takahagi T, Ikezawa N, Endo T, Ifuku K, Yamamoto Y, Kinoshita Y, Takeshita S, Sato F. 2006. Inhibition of PSII in atrazine-tolerant tobacco cells by barbatic acid, a lichen-derived depside. *Biosci. Biotechn. Biochem.* 70: 266-268.

305. Takenaka Y, Hamada N, Tanahashi T. 2011. Aromatic compounds from cultured lichen mycobionts of three *Graphis* species. *Heterocycles* 83: 2157–2164.
306. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: Molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* 30: 2725–2729.
307. Tanaka Y, Hosaka T, Ochi K. 2010. Rare earth elements activate the secondary metabolite-biosynthetic gene clusters in *Streptomyces coelicolor* A3(2). *J. Antibiot.* 63:477-481.
308. Tang L, Shah S, Chung L, Carney J, Katz L, Khosla C, Julien B. 2000. Cloning and heterologous expression of the epothilone gene cluster. *Science* 287: 640-642.
309. Tang Y, Chen AY, Kim CY, Cane DE, Khosla C. 2007. Structural and mechanistic analysis of protein interactions in module 3 of the 6-deoxyerythronolide B synthase. *Chem. Biol.* 14: 931-943.
310. Tannous J, El Khoury R, Snini SP, Lippi Y, El Khoury A, Atoui A, Lteif R, Oswald IP, Puel O. 2014. Sequencing, physical organization and kinetic expression of the patulin biosynthetic gene cluster from *Penicillium expansum*. *Int. J. Food Microbiol.* 189: 51-60.
311. Tansey TR, Shechter I. 2000. Structure and regulation of mammalian squalene synthase. *Biochim. Biophys. Acta* 1529: 49-62.
312. Tarry MJ, Hague AS, Bui KH, Schmeing TM. 2017. X-ray crystallography and electron microscopy of cross- and multi-module nonribosomal peptide synthetase proteins reveal a flexible architecture. *Structure* 25: 783-793.
313. Thompson JD, Higgins DG, Gibson TJ. 1994. Clustal W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22: 4673-4680.
314. Tian J, Yan Y, Yue Q, Liu X, Chu X, Wu N, Fan Y. 2017. Predicting synonymous codon usage and optimizing the heterologous gene for expression in *E. coli*. *Sci. Rep.* 7: 9926.
315. Timsina BA, Hausner G, Piercey-Normore MD. 2014. Evolution of ketosynthase domains of polyketide synthase genes in the *Cladonia chlorophaea* species complex (*Cladoniaceae*). *Fungal Biol.* 118: 896-909.
316. Timsina BA, Sorensen JL, Weihrauch D, Piercey-Normore MD. 2013. Effect of aposymbiotic conditions on colony growth and secondary metabolite production in the lichen-forming fungus. *Ramalina dilacerata*. *Fungal Biol.* 117: 731-743.

317. Timsina BA, Stocker-Wörgötter E, Piercey-Normore MD. 2012. Monophyly of some North American species of *Ramalina* and inferred polyketide synthase gene function. *Botany* 90: 1295-1307.
318. Trauger JW, Kohli RM, Mootz HD, Marahiel MA, Walsh CT. 2000. Peptide cyclization catalyzed by the thioesterase domain of tyrocidine synthetase. *Nature* 407: 215-218.
319. Tsai SC. 2018. The structural enzymology of iterative aromatic polyketide synthases: A critical comparison with fatty acid synthases. *Ann. Rev. Biochem.* 87: 503-531.
320. Tsuchiya K, Tada S, Gomi K, Kitamoto K, Kumagai C, Tamura G. 1992. Deletion analysis of the Taka-amylase A gene promoter using a homologous transformation system in *Aspergillus oryzae*. *Biosci. Biotechnol. Biochem.* 56: 1849-1853.
321. Tsunematsu Y, Ishiuchi K, Hotta K, Watanabe K. 2013. Yeast-based genome mining, production and mechanistic studies of the biosynthesis of fungal polyketide and peptide natural products. *Nat. Prod. Rep.* 30: 1139-1149.
322. Tunjić M, Korać P. 2013. Vertical and horizontal gene transfer in lichens. *Period Biol.* 115: 321-329.
323. Udeinya IJ, Mbah AU, Chijioke CP, Shu EN. 2004. An anti-malarial extract from neem leaves is antiretroviral. *Trans. Royal Soc. Trop. Med. Hyg.* 98: 453-457.
324. Ugai T, Minami A, Fujii R, Tanaka M, Oguri H, Gomi K, Oikawa H. 2015. Heterologous expression of highly reducing polyketide synthase involve in betaenone biosynthesis. *Chem. Commun.* 51: 1878-1881.
325. Valarmathi R, Hariharan GN, Venkataraman G, Parida A. 2009. Characterization of a non-reducing polyketide synthase gene from lichen *Dirinaria applanata*. *Phytochemistry* 70: 721–729.
326. Vattekkatte A, Garms S, Brandt W, Boland W. 2018. Enhanced structural diversity in terpenoid biosynthesis: Enzymes, substrates and cofactors. *Org. Biomol. Chem.* 16: 348-362.
327. Vetting MW, S de Carvalho LP, Yu M, Hegde SS, Magnet S, Roderick SL, Blanchard JS. 2005. Structure and functions of the GNAT superfamily of acetyltransferases. *Arch. Biochem. Biophys.* 433: 212-226.
328. Wagner C, El Omari M, König GM. 2009. Biohalogenation: Nature's way to synthesize halogenated metabolites. *J. Nat. Prod.* 72: 540-553.

329. Walsh CT, Chen H, Keating TA, Hubbard BK, Losey HC, Luo L, Marshall CG, Miller DA, Patel HM. 2001. Tailoring enzymes that modify nonribosomal peptides during and after chain elongation on NRPS assembly lines. *Curr. Opin. Chem. Biol.* 5: 525–534.
330. Wang F, Wang Y, Ji J, Zhou Z, Yu J, Zhu H, Su Z, Zhang L, Zheng J. 2015. Structural and functional analysis of the loading acyltransferase from avermectin modular polyketide synthase. *ACS Chem. Biol.* 10: 1017-1025.
331. Wang Y, Geng C, Yuan X, Hua M, Tian F, Li C. 2018. Identification of a putative polyketide synthase gene involved in usnic acid biosynthesis in the lichen *Nephromopsis pallescens*. *PLoS ONE* 13: e0199110.
332. Wang Y, Kim JA, Cheong YH, Joshi Y, Koh YJ, Hur JS. 2011. Isolation and characterization of a reducing polyketide synthase gene from the lichen-forming fungus *Usnea longissima*. *J. Microbiol.* 49: 473–480.
333. Wang Y, Kim JA, Cheong YH, Koh YJ, Hur JS. 2012. Isolation and characterization of a non-reducing polyketide synthase gene from the lichen-forming fungus *Usnea longissima*. *Mycol. Prog.* 11:75–83.
334. Wang Y, Liu B, Zhang X, Zhou Q, Zhang T, Li H, Yu Y, Zhang X, Hao X, Wang M, Wang L, Wei J. 2014. Genome characteristics reveal the impact of lichenization on lichen-forming fungus *Endocarpon pusillum* Hedwig (Verrucariales, Ascomycota). *BMC Genomics* 15:34.
335. Wang Y, Neng Z, Xiaolong Y, Mei H, Hur JS, Yuming Y, Juan W. 2016. Heterologous transcription of a polyketide synthase gene from the lichen forming fungus *Usnea longissima*. *Res. J. Biotechnol.* 11: 16-21.
336. Wang Y, Wang J, Cheong YH, Hur JS. 2014. Three new non-reducing polyketide synthase genes from the lichen-forming fungus *Usnea longissima*. *Mycobiology* 42: 34-40.
337. Wang Y, Yuan X, Chen L, Wang X, Li C. 2018. Draft genome sequence of the lichen-forming fungus *Ramalina intermedia* strain YAF0013. *Genome Announc.* 6 :e00478-18.
338. Waskell L. Kim JJP. 2015. Electron transfer partners of cytochrome p450. In: Paul R, Montellano OD (Eds) *Cytochrome p450: Structure, mechanism and biochemistry*, 4th ed. Springer International Publishing, p. 33-68.
339. Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, Heer FT, de Beer TAP, Rempfer C, Bordoli L, Lepore R, Schwede T. 2018. SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Res.* 46: W296-W303.

340. Wattanachaisaereekul S, Lantz AE, Nielsen ML, Andrésson OS, Nielsen J. 2007. Optimization of heterologous production of the polyketide 6-MSA in *Saccharomyces cerevisiae*. *Biotechnol. Bioeng.* 97: 893-900.
341. Weber T, Blin K, Duddela S, Krug D, Kim HU, Bruccoleri R, Lee SY, Fischbach MA, Müller R, Wohlleben W, Breitling R, Takano E, Medema MH. 2015. AntiSMASH 3.0 – a comprehensive resource for the genome mining of biosynthetic gene clusters. *Nucleic Acids Res.* 43: W237-W243.
342. Webster GR, The AYH, Ma JKC. 2016. Synthetic gene design – The rationale for codon optimization and implications for molecular pharming in plants. *Biotechnol. Bioeng.* 114: 492-502.
343. Weissman KJ. 2009. Introduction to polyketide biosynthesis. *Methods Enzymol.* 459: 3-16.
344. Weissman KJ. 2015. The structural biology of biosynthetic megaenzymes. *Nat. Chem. Biol.* 11: 660–670.
345. Wendt KU, Poralla K, Schulz GE. 1997. Structure and function of a squalene cyclase. *Science* 277: 1811-1815.
346. Winn M, Fyans JK, Zhuo Y, Micklefield J. 2016. Recent advances in engineering nonribosomal peptide assembly lines. *Nat. Prod. Rep.* 33: 317-347.
347. Witkowski A, Ghosal A, Joshi AK, Witkowska HE, Asturias FJ, Smith S. 2004. Head-to-head coiled arrangement of the subunits of the animal fatty acid synthase. *Chem. Biol.* 11: 1667-1676.
348. Wu YS, Ngai SC, Goh BH, Chan KG, Lee LH, Chuah LH. 2017. Anticancer activities of surfactin and potential application of nanotechnology assisted surfactin delivery. *Front. Pharmacol.* 8: 761.
349. Xie DY, Ma DM, Judd R, Jones AL. 2016. Artemisinin biosynthesis in *Artemisia annua* and metabolic engineering: Questions, challenges, and perspectives. *Phytochem. Rev.* 15: 1093-1114.
350. Xu J. 2016. Fungal DNA barcoding. *Genome* 59: 913-932.
351. Xu W, Cai X, Jung ME, Tang Y. 2010. Analysis of intact and dissected fungal polyketide synthase-nonribosomal peptide synthetase *in vitro* and in *Saccharomyces cerevisiae*. *J. Am. Chem. Soc.* 132: 13604-13607.
352. Xu W, Qiao K, Yang Y. 2013. Structural analysis of protein-protein interactions in type I polyketide synthases. *Crit. Rev. Biochem. Mol. Biol.* 48: 98-122.

353. Xu X, Liu L, Zhang F, Wang W, Li J, Guo L, Che Y, Liu G. 2014. Identification of the first diphenyl ether gene cluster for pestheic acid biosynthesis in plant endophyte *Pestalotiopsis fici*. *ChemBioChem* 15: 284-292.
354. Yaegashi J, Oakley BR, Wang CCC. 2014. Recent advances in genome mining of secondary metabolite biosynthetic gene clusters and the development of heterologous expression systems in *Aspergillus nidulans*. *J. Ind. Microbiol. Biotechnol.* 41: 433-442.
355. Yamada Y, Kuzuyama T, Komatsu M, Shin-ya K, Omura S, Cane DE, Ikeda H. 2015. Terpene synthases are widely distributed in bacteria. *Proc. Natl. Acad. Sci. U.S.A.* 112: 857-862.
356. Yamanaka K, Maruyama C, Takagi H, Hamano Y. 2008. Epsilon-poly-L-lysine dispersity is controlled by a highly unusual nonribosomal peptide synthetase. *Nat. Chem. Biol.* 4: 766-772.
357. Ye S, Molloy B, Braña AF, Zabala D, Olano C, Cortés J, Moris F, Salas JA, Méndez C. 2017. Identification by genome mining of a type I polyketide gene cluster from *Streptomyces argillaceus* involved in the biosynthesis of pyridine and piperidine alkaloids argimycins P. *Front. Microbiol.* 8: 194.
358. Yoshimura I, Yamamoto Y, Nakano T, Finnie J. 2002. Isolation and culture of lichen photobionts and mycobionts. In: I Kranner, Beckett R, Varma A, (Eds) *Protocols in Lichenology: Culturing, biochemistry, ecophysiology and use in biomonitoring*. Springer, New York, NY, p. 3-33.
359. Yu D, Xu F, Zeng J, Zhan J. 2012. Type III polyketide synthases in natural product biosynthesis. *IUBMB Life* 64: 285-295.
360. Yu JH, Leonard TJ. 1995. Sterigmatocystin biosynthesis in *Aspergillus nidulans* requires a novel type I polyketide synthase. *J. Bacteriol.* 177: 4792-4800.
361. Yu PW, Cho TY, Liou RF, Tzean SS, Lee TH. 2017. Identification of the orsellinic acid synthase PKS63787 for the biosynthesis of antroquinonols in *Antrodia cinnamomea*. *Appl. Microbiol. Biotechnol.* 101: 4701-4711.
362. Yuan X, Xiao S, Taylor TN. 2005. Lichen-like symbiosis 600 million years ago. *Science* 308: 1017-1020.
363. Zabala D, Braña AF, Flórez AB, Salas JA, and Méndez C. 2013. Engineering precursor metabolite pools for increasing production of antitumor mithramycins in *Streptomyces argillaceus*. *Metab. Eng.* 20: 187-197.
364. Zaehle C, Gressler M, Shelest E, Geib E, Hertweck C, Brock M. 2014. Terrein biosynthesis in *Aspergillus terreus* and its impact on phytotoxicity. *Chem. Biol.* 21: 719-731.

365. Zambare VP, Christopher LP. 2012. Biopharmaceutical potential of lichens. *Pharm. Biol.* 50: 778-798.
366. Zerbino DR, Birney E. 2008. Velvet: Algorithms for *de novo* short read assembly using de Bruijn graphs. *Genome Res.* 18: 821-829.
367. Zhang H, Wang Y, Wu J, Skalina K, Pfeifer BA. 2010. Complete biosynthesis of erythromycin A and designed analogs using *E. coli* as a heterologous host. *Chem. Biol.* 17: 1232-1240.
368. Zhang MM, Wang Y, Ang EL, Zhao H. 2015. Engineering microbial hosts for production of bacterial natural products. *Nat. Prod. Rep.* 33: 963-987.
369. Zhang MM, Wong FT, Wang Y, Luo S, Lim YH, Heng E, Yeo WL, Cobb RE, Enghiad B, Ang EL, Zhao H. 2017. CRISPR-Cas9 strategy for activation of silent *Streptomyces* biosynthetic gene clusters. *Nat. Chem. Biol.* 13: 607-607.
370. Zhang W, Tang Y. 2007. Engineering starter units in aromatic polyketides. *ACS Symposium Series* 955: 231-245.
371. Ziemert N, Alanjary M, Weber T. 2016. The evolution of genome mining in microbes – a review. *Nat. Prod. Rep.* 33: 988-1005.
372. Ziemert N, Jensen PR. 2012. Phylogenetic approaches to natural product structure prediction. *Methods Enzymol.* 517: 161-182.
373. Zuckerkandl E, Pauling L. 1965. Evolutionary divergence and convergence in proteins. In: Bryson V, Vogel HJ (Eds.) *Evolving Genes and Proteins*, Academic Press, New York, p. 97-166.

Supporting information items

Chapter 1

Not applicable

Chapter 2

Table S1: Primers used throughout this work.

Name	Sequence (5' to 3')	Purpose(s)
CU-MPAO-F	TCGAGCTCGGTACCCGGGATGA-TTTCACCAAGTGTCTTCTATTCTAG	Amplifies <i>mpao</i> (<i>C. uncialis</i>); contains InFusion element for plasmid construction
CU-MPAO-R	CTACTACAGATCCCCGGGTCAA-GCCTTTTCATCCGGAAC	Amplifies <i>mpao</i> (<i>C. uncialis</i>); contains 15nt to insert gene into plasmid
CU-MPAS-R1	GATGCTCTCCATCTCGATG	Amplifies the first of four portions of <i>mpas</i> cDNA. CU-MPAS-F is sense primer.
CU-MPAS-F2	CATCGAGATGGAGAGCATC	Amplifies the second of four portions of <i>mpas</i> cDNA. CU-MPAS-R2 is anti-sense primer.
CU-MPAS-R2	GAGCGATTCACTTCGAGG	Amplifies the second of four portions of <i>mpas</i> cDNA. CU-MPAS-F2 is sense primer.
CU-MPAS-F3	CCTCGAAGTGAATCGCTC	Amplifies the third of four portions of <i>mpas</i> cDNA. CU-MPAS-R3 is anti-sense primer.
CU-MPAS-R3	GCAGTCAAATGTCATGAATCT	Amplifies the third of four portions of <i>mpas</i> cDNA. CU-MPAS-F3 is sense primer.
CU-MPAS-F4	AGATTCATGACATTTGACTGC	Amplifies the fourth of four portions of <i>mpas</i> cDNA. CU-MPAS-R is anti-sense primer.
AO-Tubulin-qPCR-R	AACGACATCCACAACAGAATC	qPCR of β -tubulin (<i>A. oryzae</i>)
CO-MPAS-F	TCGAGCTCGGTACCCGGGAT-GGCTCTGCCCTCTCTGAT	Amplifies <i>mpas</i> (codon optimized for translation in <i>A. oryzae</i>); contains InFusion element for plasmid construction
His6-CO-MPAS-F	TCGAGCTCGGTACCCATGC-ACCATCACCATCACCATGCT-CTGCCCTCTCTGATCGCTTC	Amplifies <i>mpas</i> (codon-optimized for translation in <i>A. oryzae</i>); attaches N-terminal His(6)-tag; InFusion element for plasmid construction

CO-MPAS-R	CTACTACAGATCCCCGGGT- CAGGACATGCCCAGCTG	Amplifies <i>mpas</i> (codon-optimized for translation in <i>A. oryzae</i>); contains InFusion element for plasmid construction
CO-MPAS-R1	GCGCTGGCAATACTGGTC	Amplifies the first of three portions of codon-optimized <i>mpas</i> cDNA. CO-MPAS-F is sense primer.
CO-MPAS-F2	CAGTATTGCCAGCGCCTGCG- CTTCACCCAGGATA	Amplifies the second of three portions of codon-optimized <i>mpas</i> cDNA. CO-MPAS-R2 is anti-sense primer.
CO-MPAS-R2	GATGACCTCGATCATCATCAGG- GAATCGACGCC	Amplifies the second of three portions of codon-optimized <i>mpas</i> cDNA. CO-MPAS-F2 is sense primer.
CO-MPAS-F3	ATGATCGAGGTCATCTCCGAG	Amplifies the third of three portions of codon-optimized <i>mpas</i> cDNA. CO-MPAS-R is anti-sense primer.
PPTase-R	CTACTACAGATCCCCGGG- TCAAGGATAGACATCCTT- TTCAACATCAAGCTC	Amplifies <i>pptase</i> (<i>C. uncialis</i>); contains InFusion element for plasmid construction
CPR1-F	TCGAGCTCGGTACCCGGG- ATGGATCGCTCGAAGAATGC	Amplifies <i>cpr1</i> (<i>C. uncialis</i>); contains InFusion element for plasmid construction
CPR1-R	CTACTACAGATCCCCGGGC- TAGCTCGACCATACATCCTCTG	Amplifies <i>cpr1</i> (<i>C. uncialis</i>); contains InFusion element for plasmid construction
CPR2-F	TCGAGCTCGGTACCCGGGA- TGGCGCAACTGGACAC	Amplifies <i>cpr2</i> (<i>C. uncialis</i>); contains InFusion element for plasmid construction
CPR2-R	CTACTACAGATCCCCGGGCT- AAGACCAACATCTTCTTGGTACT	Amplifies <i>cpr2</i> (<i>C. uncialis</i>); contains InFusion element for plasmid construction
B5R1-F	TCGAGCTCGGTACCCGGGAT- GTCTTCTCCTCACCAAGTCCG	Amplifies <i>b5r1</i> (<i>C. uncialis</i>); contains InFusion element for plasmid construction
B5R1-R	CTACTACAGATCCCCGGGTTA- GAAGCAGAAGACTTGGTCGACAA	Amplifies <i>b5r1</i> (<i>C. uncialis</i>); contains InFusion element for plasmid construction
B5R2-F	TCGAGCTCGGTACCCGGGATGT- TTGCACGACAGTTGTT	Amplifies <i>b5r2</i> (<i>C. uncialis</i>); contains InFusion element for plasmid construction
B5R2-R	CTACTACAGATCCCCGGGCTAA- AACTTATAAACCTGCTCCTTCTC	Amplifies <i>b5r2</i> (<i>C. uncialis</i>); contains InFusion element for plasmid construction
P-Gene-T- (All)-R	GCAGGTCGACTCTAGATGGTG- TTTTGATCATTTTAAATT	Amplifies promoter-gene-terminator to construct 'cassette'. This primer is used to insert all five genes.
P-Gene-T- (1 st)-F	TAGTAGATCCTCTAGCTTTCCTA- TAATAGACTAGCGTGC	Amplifies promoter-gene-terminator to construct 'cassette'. This primer is used to insert the first gene.

P-Gene-T-(Rem)-F	AAAACACCATTCTAGCTTTCCTA-TAATAGACTAGCGTGC	Amplifies promoter-gene-terminator to construct 'cassette'. This primer is used to insert the 2 nd -5 th genes.
6HMS-F	TCGAGCTCGGTACCCGGGATGG-CTGAGGCATTTCAAGT	Amplifies <i>6hms</i> (<i>C. uncialis</i>); contains InFusion element for plasmid construction
6HM-R	CTACTACAGATCCCCGGGTAA-GTAATGGCGATAGATACTTACATCG	Amplifies <i>6hms</i> (<i>C. uncialis</i>); contains InFusion element for plasmid construction
KRDH-F	TCGAGCTCGGTACCCGGGATG-ATTAGAACAGCTGTTGCACAAG	Amplifies <i>krdh</i> (<i>C. uncialis</i>); contains InFusion element for plasmid construction
KRDH-R	CTACTACAGATCCCCGGGTCA-GGTTGAGTTCTTCCCAGTC	Amplifies <i>krdh</i> (<i>C. uncialis</i>); contains InFusion element for plasmid construction
CU-GAS-R	CTACTACAGATCCCCCTACTT-CTCCGGAAGGCCC	Amplifies <i>gas</i> (<i>C. uncialis</i>); contains InFusion element for plasmid construction
CU-GAS-F	TCGAGCTCGGTACCCATGACC-CTACCAAACAATGTTGTC	Amplifies <i>gas</i> (<i>C. uncialis</i>); contains InFusion element for plasmid construction
CU-GAS-R1	CTTGTCGCATGAGATTCTG	Amplifies the first of three portions of <i>cu-gas</i> cDNA. CU-GAS-F is sense primer.
CU-GAS-F2	ATCTCATGCGACAAGTTCTAC	Amplifies the second of three portions of <i>cu-gas</i> cDNA. CU-GAS-R2 is the anti-sense primer.
CU-GAS-R2	GAGTTGTTTCACCTTTTCGA	Amplifies the second of three portions of <i>cu-gas</i> cDNA. CU-GAS-F2 is the sense primer.
CU-GAS-F3	AAGGTGAAACAACCTCGCAC	Amplifies the third of three portions of <i>cu-gas</i> cDNA. CU-GAS-R is the anti-sense primer.
His6-FU-OAS- PKS-F	TCGAGCTCGGTACCCATGCAC-CATCACCATCACCATGGTGCC-ACAACGAGGCCGAG	Amplifies <i>oas</i> (<i>Fusarium sp.</i>); attaches N-terminal His(6)-tag; contains InFusion element for plasmid construction
FU-OAS- PKS-R	CTACTACAGATCCCCTTATTC-CCTCGGTGAGCATTCTCAATCTG	Amplifies <i>oas</i> (<i>Fusarium sp.</i>); contains InFusion element for plasmid construction
CU-OAS-R1	CACAACCGGAGGCAGCTTC	Amplifies the first of three portions of <i>cu-oas</i> cDNA. CU-OAS-F is sense primer.
CU-OAS-F2	GAAGCTGCCTCCGGTTGTG	Amplifies the second of three portions of <i>cu-oas</i> cDNA. CU-OAS-R2 is the anti-sense primer.
CU-OAS-R2	TTCTTGCGCAGAGATTGGC	Amplifies the second of three portions of <i>cu-oas</i> cDNA. CU-OAS-F2 is the sense primer.
CU-OAS-F3	GCCAATCTCTGCGCCAAGAA	Amplifies the third of three portions of <i>cu-oas</i> cDNA. CU-OAS-R is the anti-sense primer.

CU-6MSAS-F	TCGAGCTCGGTACCCGGGATG- GCTGGAATGGTAGATGACAAG	Amplifies <i>6msas</i> (<i>C. uncialis</i>); contains InFusion element for plasmid construction
CU-6MSAS-R	CTACTACAGATCCCCGGGTTAT- TGTTTCATCTGCTCCGTGAAC	Amplifies <i>6msas</i> (<i>C. uncialis</i>); contains InFusion element for plasmid construction
CU-6MSAS-R1	AGTAGCCAGCGTTGCTGCGATTG	Amplifies the first of three portions of <i>cu-6msas</i> cDNA. CU-6MSAS-F is sense primer.
CU-6MSAS-F2	CAATCGCAGCAACGCTGGCTACT	Amplifies the second of three portions of <i>cu-6msas</i> cDNA. CU-6MSAS-R2 is anti-sense primer.
CU-6MSAS-R2	CGAGCCAATCGAGGTCGCTGC	Amplifies the second of three portions of <i>cu-6msas</i> cDNA. CU-6MSAS-F2 is sense primer.
CU-6MSAS-F3	GCAGCGACCTCGATTGGCTCG	Amplifies the third of three portions of <i>cu-6msas</i> cDNA. CU-6MSAS-R is ant-sense primer.
T3PKS1-F	TCGAGCTCGGTACCCGGG- ATGGCAGCAGTCGATATGTC	Amplifies <i>t3pks1</i> (<i>C. uncialis</i>); contains InFusion element for plasmid construction
T3PKS1-R	CTACTACAGATCCCCGGG- TCACTCTTCTCTTTTCTCACG	Amplifies <i>t3pks1</i> (<i>C. uncialis</i>); contains InFusion element for plasmid construction
T3PKS2-F	TCGAGCTCGGTACCCGGG- ATGTCGCCATCTGCTCTCAACG	Amplifies <i>t3pks2</i> (<i>C. uncialis</i>); contains InFusion element for plasmid construction
T3PKS2-R	CTACTACAGATCCCCGGGTTAAT- CCACGTCTTCAGTGGGTAATCC	Amplifies <i>t3pks2</i> (<i>C. uncialis</i>); contains InFusion element for plasmid construction
His6-AO-Tubulin-F	TCGAGCTCGGTACCCATGCAC- CATCACCATCACCATCGTGAGA- TCGTGCACCTC	Amplifies β -tubulin gene from <i>A. oryzae</i> ; attaches N-terminal His(6)-tag; contains InFusion element for plasmid construction
AO-Tubulin-R	CTACTACAGATCCCCCTACTGC- TCCTCGATCTCCTCCTC	Amplifies β -tubulin gene from <i>A. oryzae</i> ; contains InFusion element for plasmid construction

Table S2: Plasmids used throughout this work

Plasmid name	Marker	Description
MPAS-pUSA	Methionine	Contains <i>mpas</i> from <i>C. uncialis</i>
MPAO-pTAex3	Arginine	Contains <i>mpao</i> from <i>C. uncialis</i>
CO-MPAS-pUSA	Methionine	Contains <i>mpas</i> codon-optimized for <i>A. oryzae</i>
CO(His6)-MPAS-pUSA	Methionine	Contains <i>mpas</i> codon-optimized for <i>A. oryzae</i> ; N-terminal His(6)-tagged variant of entry above
CO-MPAO-pTAex3	Arginine	Contains <i>mpao</i> codon-optimized for <i>A. oryzae</i>
PPTase-pUSA	Methionine	Contains <i>pptase</i> from <i>C. uncialis</i>
PPTase-pTAex3	Arginine	Contains <i>pptase</i> from <i>C. uncialis</i>
PPTase-pAdeA	Adenine	Contains <i>pptase</i> from <i>C. uncialis</i>
CPR1-pUSA	Methionine	Prepared as a PCR template for 'cassette' construction. Contains <i>cpr1</i> from <i>C. uncialis</i>
CPR2-pUSA	Methionine	Prepared as a PCR template for 'cassette' construction. Contains <i>cpr2</i> from <i>C. uncialis</i>
B5R1-pUSA	Methionine	Prepared as a PCR template for 'cassette' construction. Contains <i>b5r1</i> from <i>C. uncialis</i>
B5R2-pUSA	Methionine	Prepared as a PCR template for 'cassette' construction. Contains <i>b5r2</i> from <i>C. uncialis</i>
Cassette	Adenine	pAdeA backbone. Contains <i>pptase</i> , <i>cpr1</i> , <i>cpr2</i> , <i>b5r1</i> , <i>b5r2</i> . Each gene possesses a promoter and terminator.
6HMS-pUSA	Methionine	Contains <i>6hms</i> from <i>C. uncialis</i>
KRDH-pTAex3	Arginine	Contains <i>krdh</i> from <i>C. uncialis</i>
OAS(CU)-pAdeA	Adenine	Contains <i>oas</i> from <i>C. uncialis</i>
GAS(CU)-pAdeA	Adenine	Contains <i>gas</i> from <i>C. uncialis</i>
OAS(FU)-pAdeA	Adenine	Contains <i>oas</i> from <i>Fusarium sp.</i>
OAS(FU)(His6)-pAdeA	Adenine	Contains <i>oas</i> from <i>Fusarium sp.</i> ; N-terminal His(6)-tagged variant of entry above.
6MSAS(CU)-pUSA	Methionine	Contains <i>6msas</i> from <i>C. uncialis</i>
6MSAS(PE)-pUSA	Methionine	Contains <i>6msas</i> from <i>Penicillium sp.</i>
6MSAS(PE)(His6)-pUSA	Methionine	Contains <i>6msas</i> from <i>Penicillium sp.</i> ; N-terminal His(6)-tagged variant of entry above.
T3PKS1-pTAex3	Arginine	Contains <i>t3pks1</i> from <i>C. uncialis</i>
T3PKS2-pTAex3	Arginine	Contains <i>t3pks2</i> from <i>C. uncialis</i>
Tub(AO)(His6)-pAdeA	Adenine	Contains <i>tubulin</i> from <i>A. oryzae</i> ; with N-terminal His(6)-tag

Chapter 3

Table S3: The 20 highest BLASTn similarity results for the internal transcribed spacer (ITS) sequence of native lichen specimen used in the study (Accession no. KY80361). Reproduced from Bertrand et al. (2018a), Supporting Information file, in accordance with authors' retainment of privileges.

Accession no.	Summary	Identity/ Coverage
KR019404	Cladonia uncialis subsp. uncialis isolate CL204 internal transcribed spacer	95/96
AF455247	Cladonia uncialis subsp. uncialis internal transcribed spacer	95/96
KR019420	Cladonia uncialis subsp. uncialis isolate FH324 internal transcribed spacer	95/96
KR019415	Cladonia uncialis subsp. uncialis isolate DW7 internal transcribed spacer	95/96
KR019412	Cladonia uncialis subsp. uncialis isolate DW3 internal transcribed spacer	95/96
KR019409	Cladonia uncialis subsp. uncialis isolate DW26 internal transcribed spacer	95/96
KR019407	Cladonia uncialis subsp. uncialis isolate DW2 internal transcribed spacer	95/96
KR019387	Cladonia uncialis subsp. uncialis isolate FH232 internal transcribed spacer	95/96
AF455249	Cladonia uncialis subsp. uncialis internal transcribed spacer	95/96
KR019400	Cladonia uncialis subsp. uncialis isolate FH328 internal transcribed spacer	94/98
AF455248	Cladonia uncialis subsp. uncialis internal transcribed spacer	94/96
KR019417	Cladonia uncialis subsp. uncialis isolate FH234 internal transcribed spacer	94/96
KR019405	Cladonia uncialis subsp. biuncialis isolate DW12 internal transcribed spacer	94/96
KR019374	Cladonia uncialis subsp. biuncialis isolate CL210 internal transcribed spacer	94/96
AF456396	Cladonia uncialis subsp. biuncialis internal transcribed spacer	94/96
AF456395	Cladonia uncialis subsp. biuncialis internal transcribed spacer	94/96
KR019419	Cladonia uncialis subsp. uncialis isolate FH237 internal transcribed spacer	93/96
KR019418	Cladonia uncialis subsp. uncialis isolate FH235 internal transcribed spacer	93/96
KR019401	Cladonia uncialis subsp. uncialis isolate FH332 internal transcribed spacer	93/96
KR019423	Cladonia uncialis subsp. uncialis isolate FH350 internal transcribed spacer	93/98

Table S4: The 20 highest BLASTn similarity results for the internal transcribed spacer (ITS) sequence of axenic sub-cultured fungal partner of the lichen specimen used in the study (Accession no. KY80362). Reproduced from Bertrand et al. (2018a), Supporting Information file, in accordance with authors' retainment of privileges.

Accession no.	Summary	Identity/ Coverage
KR019419	Cladonia uncialis subsp. uncialis isolate FH237 internal transcribed spacer	99/99
KR019418	Cladonia uncialis subsp. uncialis isolate FH235 internal transcribed spacer	99/99
KR019401	Cladonia uncialis subsp. uncialis isolate FH332 internal transcribed spacer	99/99
AF456392	Cladonia uncialis subsp. uncialis internal transcribed spacer	99/100
KR019421	Cladonia uncialis subsp. uncialis isolate FH326 internal transcribed spacer	99/99
KR019413	Cladonia uncialis subsp. uncialis isolate DW4 internal transcribed spacer	99/99
KR019408	Cladonia uncialis subsp. uncialis isolate DW25 internal transcribed spacer	98/99
KR019422	Cladonia uncialis subsp. uncialis isolate FH349 internal transcribed spacer	98/99
KR019414	Cladonia uncialis subsp. uncialis isolate DW5 internal transcribed spacer	98/99
KR019411	Cladonia uncialis subsp. uncialis isolate DW28 internal transcribed spacer	98/99
AF456395	Cladonia uncialis subsp. biuncialis internal transcribed spacer	98/100
KR019417	Cladonia uncialis subsp. uncialis isolate FH234 internal transcribed spacer	98/99
AF456396	Cladonia uncialis subsp. biuncialis internal transcribed spacer	98/100
KR019405	Cladonia uncialis subsp. biuncialis isolate DW12 internal transcribed spacer	98/99
KR019374	Cladonia uncialis subsp. biuncialis isolate CL210 internal transcribed spacer	98/99
AF455249	Cladonia uncialis subsp. uncialis internal transcribed spacer	98/100
KR019415	Cladonia uncialis subsp. uncialis isolate DW7 internal transcribed spacer	98/99
KR019409	Cladonia uncialis subsp. uncialis isolate DW26 internal transcribed spacer	98/99
KR019407	Cladonia uncialis subsp. uncialis isolate DW2 internal transcribed spacer	98/99
KR019400	Cladonia uncialis subsp. uncialis isolate FH328 internal transcribed spacer	98/99

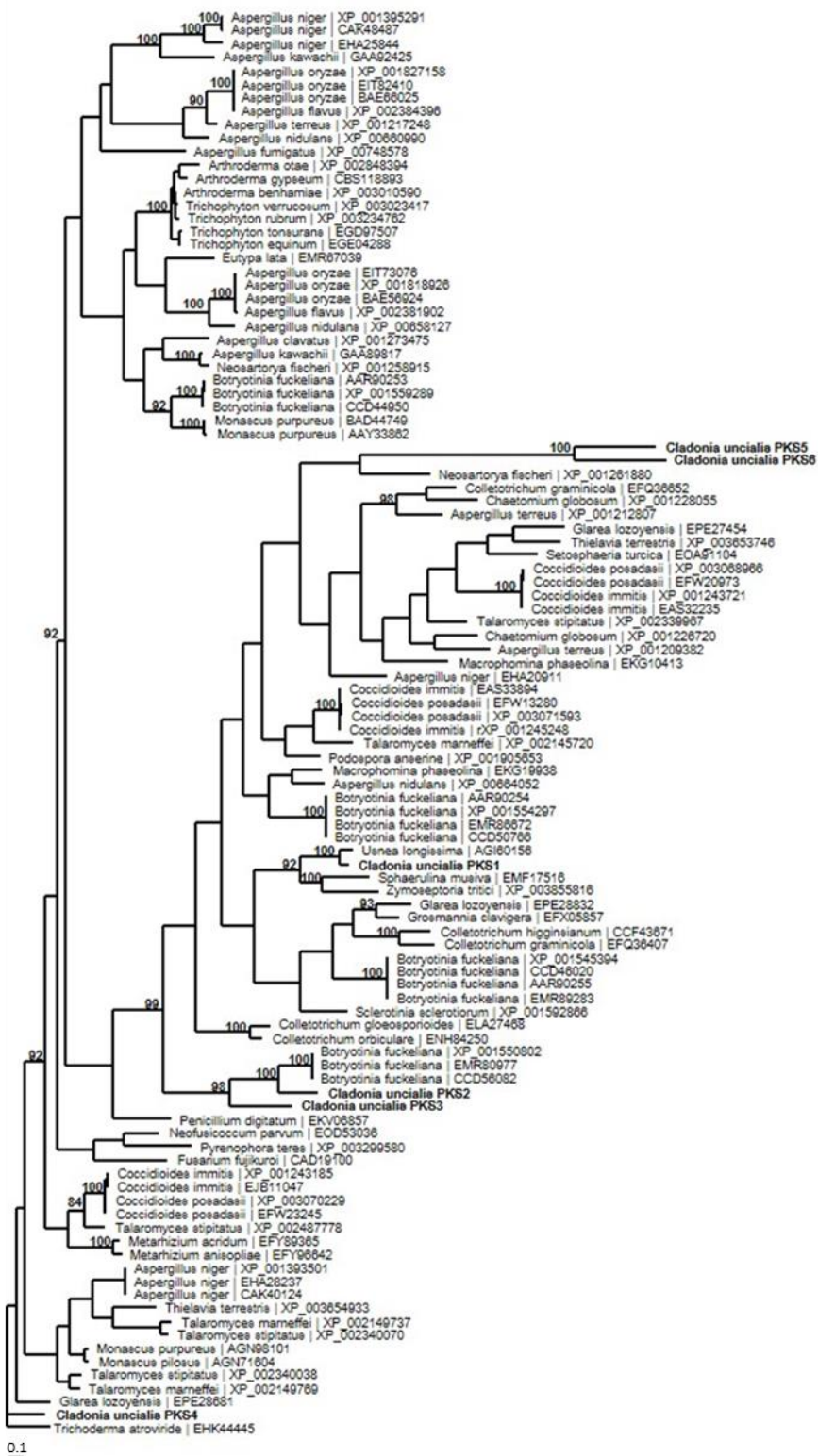


Figure S1: Phylogenetic tree constructed from translated amino acid sequences of the KS domain of the five non-reducing PKS genes possessing C-methyltransferase domains. The “*C. uncialis* PKS1” gene encodes MPAS. Unit bar denotes average number of amino acid substitutions per site. Reproduced from Abdel-Hameed et al. (2016a), Supporting Information file, in accordance with authors’ retainment of privileges.

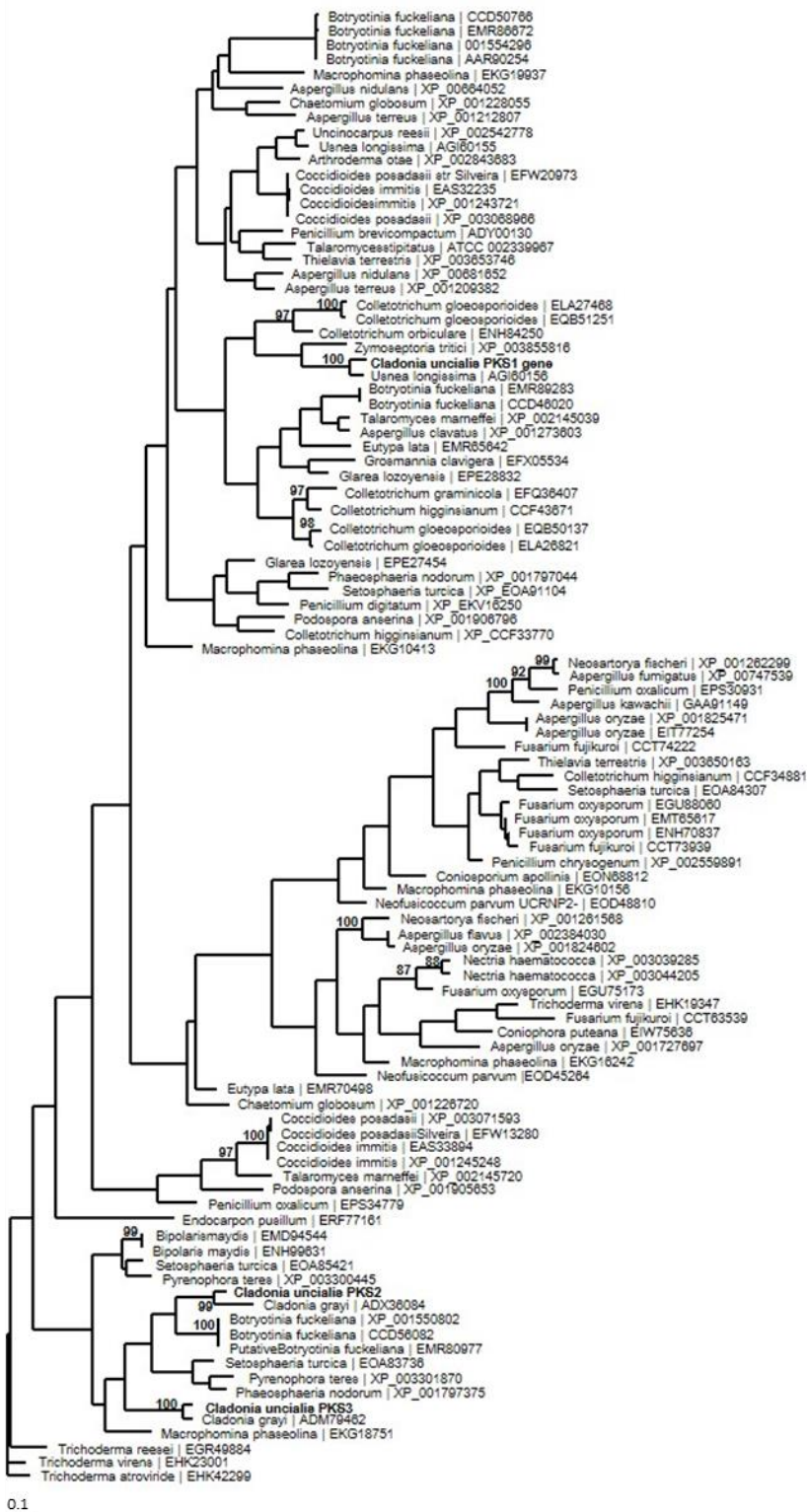


Figure S2: Phylogenetic tree constructed from translated amino acid sequences of the CLC domain of the five non-reducing PKS genes possessing *C*-methyltransferase domains. The “*C. uncialis* PKS1” gene encodes MPAS. Unit bar denotes average number of amino acid substitutions per site. Reproduced from Abdel-Hameed et al. (2016a), Supporting Information file, in accordance with authors’ retainment of privileges.

Chapter 4

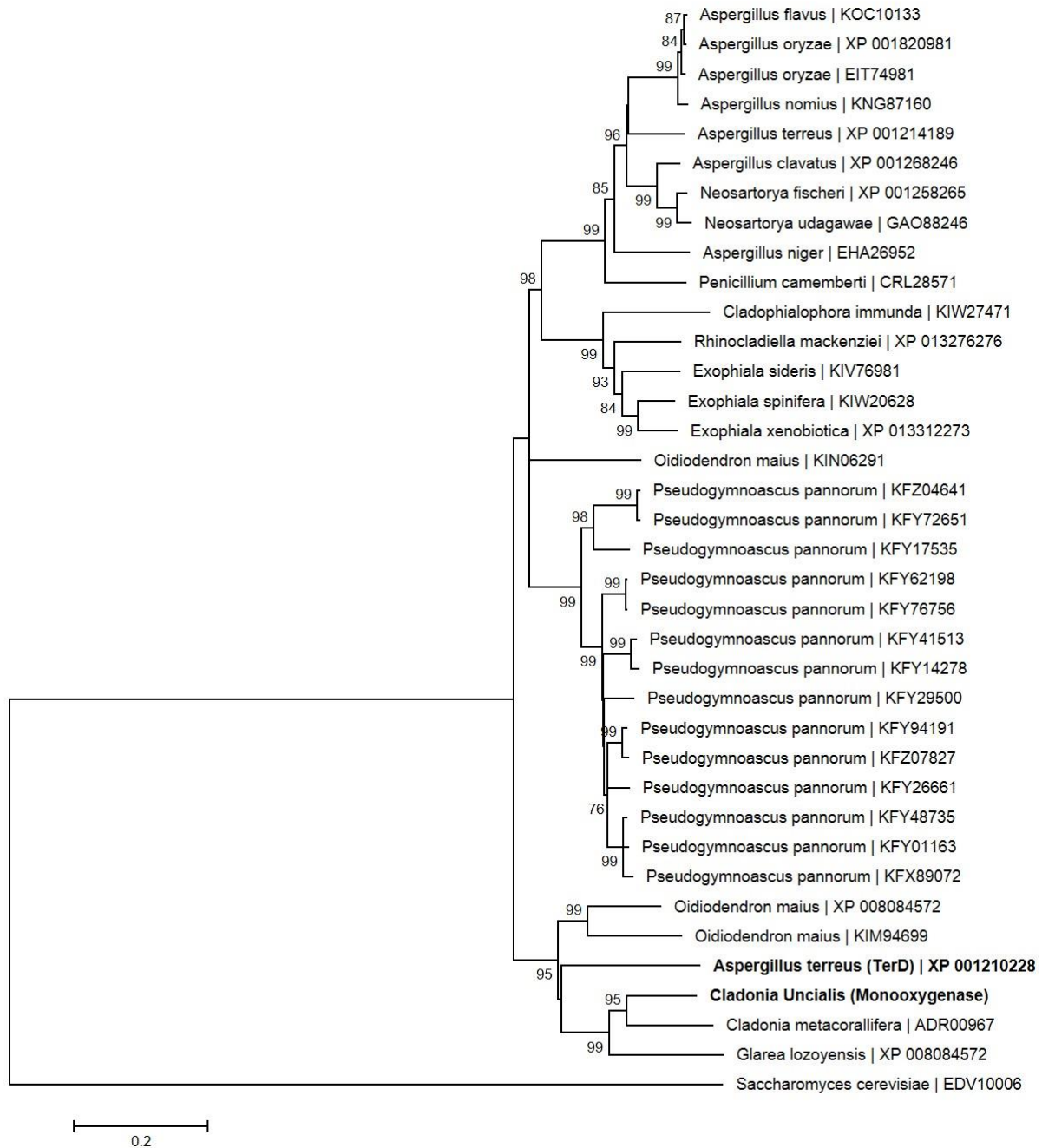


Figure S3: Phylogenetic tree of the monooxygenase (6-hydroxymellein oxidase) from *Cladonia uncialis*. Evolutionary distances are drawn to scale and are in units of number of amino acid substitutions per site. Nodes with confidence probabilities of 70 percent or greater are displayed. This tree was constructed with 37 amino acid sequences and were chosen based on greatest homology to the *C. uncialis* monooxygenase as determined by BLAST analysis. The species and accession numbers are displayed. The *C. uncialis* monooxygenase and the identified *A. terreus* homolog (TerD) are highlighted in bold. A monooxygenase from *Saccharomyces cerevisiae* (Accession no. EDV10006) was chosen as an out-group. Reproduced from Abdel-Hameed et al. (2016b), Supporting Information file, in accordance with authors' retainment of privileges.

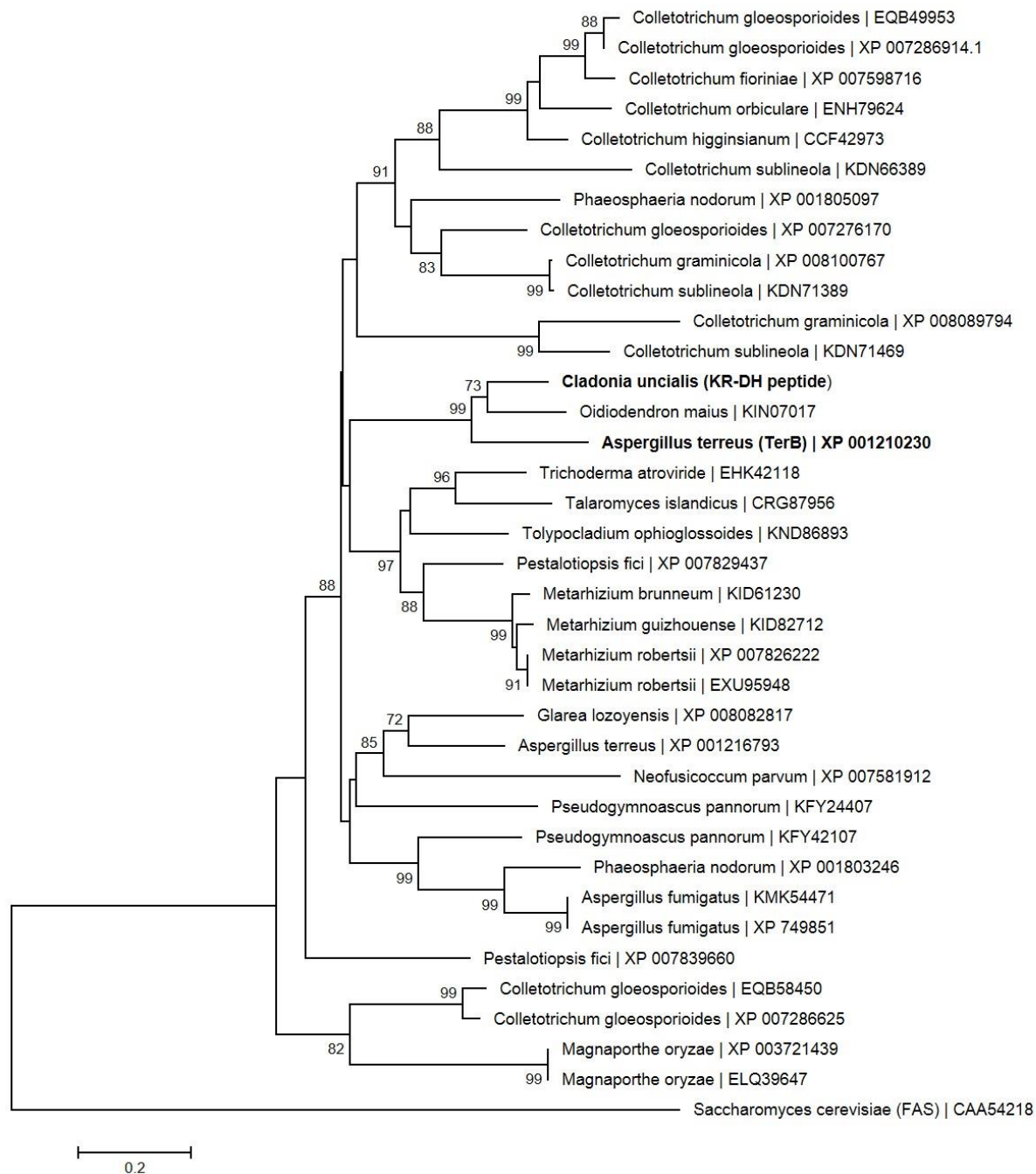


Figure S4: Phylogenetic tree of the ketoreductase-dehydratase-like (KR-DH) peptide (2,3-dehydro-6-hydroxymellein reductase) from *Cladonia uncialis*. Evolutionary distances are drawn to scale and are in units of number of amino acid substitutions per site. Nodes with confidence probabilities of 70 percent or greater are displayed. This tree was constructed with 37 amino acid sequences and were chosen based on greatest homology to the *C. uncialis* KR-DH pepetide as determined by BLAST analysis. The species and accession numbers are displayed. The *C. uncialis* KR-DH pepetide and the identified *A. terreus* homolog (TerB) are highlighted in bold. A fatty acid synthase from *Saccharomyces cerevisiae* (Accession no. CAA54218) was chosen as an out-group. Reproduced from Abdel-Hameed et al. (2016b), Supporting Information file, in accordance with authors' retainment of privileges.

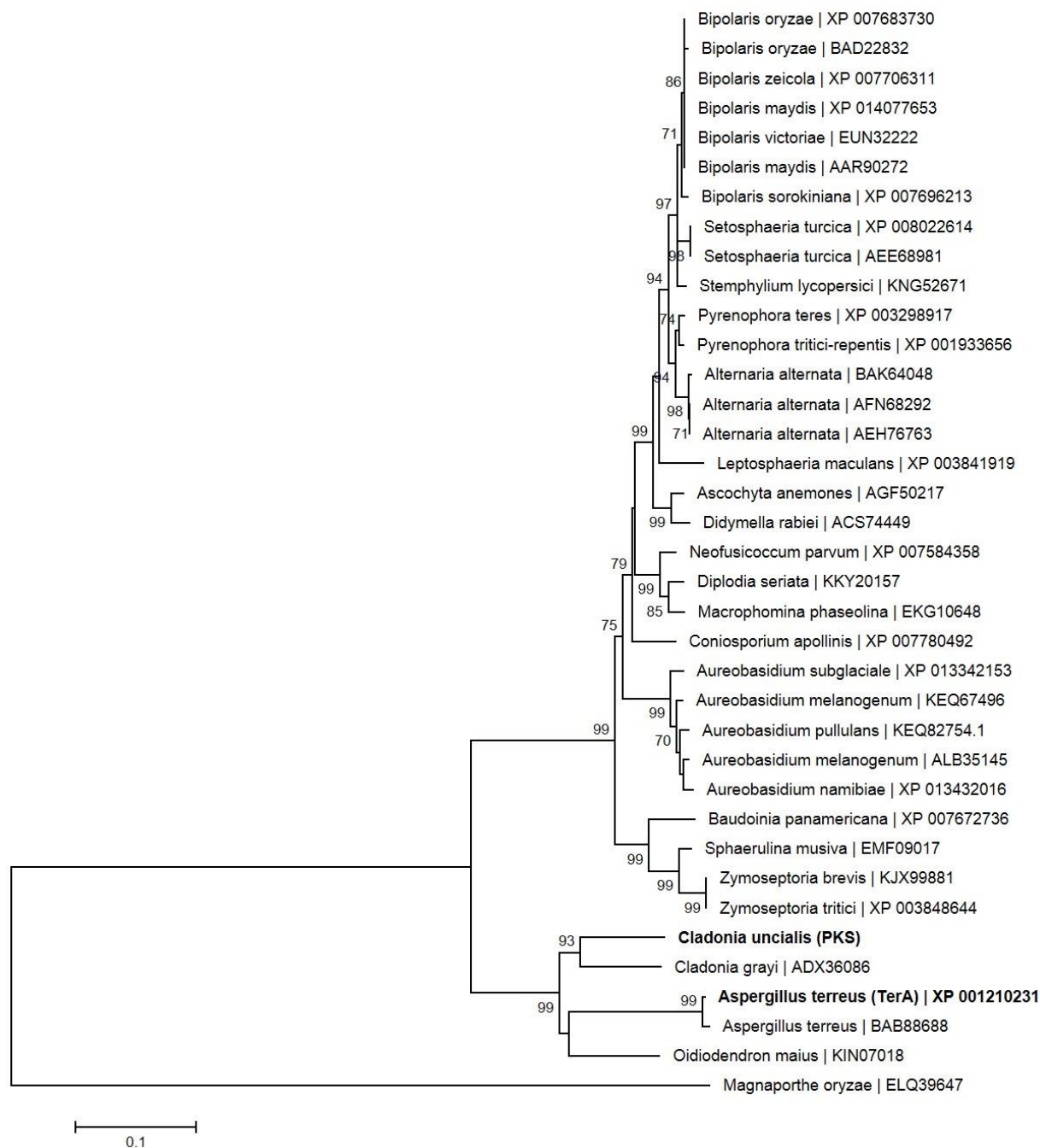


Figure S5: Phylogenetic tree of the polyketide synthase from *Cladonia uncialis*. Evolutionary distances are drawn to scale and are in units of number of amino acid substitutions per site. Nodes with confidence probabilities of 70 percent or greater are displayed. This tree was constructed with 37 amino acid sequences and were chosen based on greatest homology to the *C. uncialis* PKS as determined by BLAST analysis. The species and accession numbers are displayed. The *C. uncialis* PKS and the identified *A. terreus* homolog (TerA) are highlighted in bold. A PKS from *Magnaporthe oryzae* (Accession no. ELQ39647) was chosen as an out-group. Reproduced from Abdel-Hameed et al. (2016b), Supporting Information file, in accordance with authors' retainment of privileges.

Chapter 5

Table S5: BLAST statistics for Cluster 1. Reproduced from Bertrand et al. (2018a), Supporting Information file, in accordance with authors' retainment of privileges.

Gene (No.)	Protein Accession No.	Protein Putative function(s)	Species Closest homolog	Homolog Accession No.	Identity (%) / Coverage (%)
1	AUW30692	Type I PKS (Cu-r-pks-1)	<i>F. gaditjirri</i>	ALQ32871	54/99
2	AUW30693	Chalcone synthase (Cu-t3-pks-1)	<i>B. bassiana</i>	XP_008597692	53/84
A	AUW30694	Hypothetical protein	<i>G. lozoyensis</i>	XP_008086870	53/83
B	AUW30695	No significant similarity	-	-	-
C	AUW30696	No significant similarity	-	-	-
D	AUW30697	No significant similarity	-	-	-
E	AUW30698	Serine/threonine kinase	<i>C. yegresii</i>	XP_007757863	30/97
F	AUW30699	Hypothetical protein	<i>E. pusillum</i>	XP_007804569	41/98
G	AUW30700	Glutamate decarboxylase	<i>S. borealis</i>	ESZ90106	65/98
H	AUW30701	ARM repeat-containing protein	<i>L. palustris</i>	OCK84907	52/91
I	AUW30702	TrkA-N-domain dehydrogenase	<i>G. stellatum</i>	OCL11131	43/99
J	AUW30703	No significant similarity	-	-	-
K	AUW30704	RINT-1 family protein	<i>U. pustulata</i>	SLM36544	57/98
L	AUW30705	DNA Ligase 1	<i>U. pustulata</i>	SLM36543	72/98
M	AUW30706	Hypothetical protein	<i>P. chlamydospore</i>	KKY15877	51/98
N	AUW30707	Betaine lipid synthase	<i>E. parva</i>	PGG98678	53/90
O	AUW30708	Hypothetical protein	<i>P. zonata</i>	XP_022582027	54/71
P	AUW30709	Ras GTPase	<i>M. phaseolina</i>	EKG10611	73/50

Table S6: BLAST statistics for Cluster 2. Reproduced from Bertrand et al. (2018a), Supporting Information file, in accordance with authors' retainment of privileges.

Gene (No.)	Protein Accession No.	Protein Putative function(s)	Species Closest homolog	Homolog Accession No.	Identity (%) / Coverage (%)
A	AUW30710	90S preribosome ssu processome component krr1	<i>U. pustulata</i>	SLM33697	88/75
B	AUW30711	20S proteasome alpha subunit	<i>B. cinerea</i>	XP_001547757	82/97
C	AUW30712	Spindle pole body component	<i>U. pustulata</i>	SLM33712	79/81
D	AUW30713	C6 finger domain protein	<i>B. spectabilis</i>	GAD93882	46/70
1	AUW30714	Chalcone synthase (Cu-t3-pks-2)	<i>E. dermatitidis</i>	XP_009159865	57/91
2	AUW30715	Cytochrome P-450	<i>P. scopiformis</i>	KUJ12289	57/94
E	AUW30716	Peptidase	<i>P. omphalodes</i>	CCX16186	51/89
F	AUW30717	No significant similarity	-	-	-
G	AUW30718	Metallo-dependent phosphatase	<i>M. phaseolina</i>	EKG18441	63/93
H	AUW30719	RING finger domain	<i>L. palustris</i>	OCK7893	65/96

Table S7: BLAST statistics for Cluster 3. Reproduced from Bertrand et al. (2018a), Supporting Information file, in accordance with authors' retention of privileges.

Gene (No.)	Protein Accession No.	Protein Putative function(s)	Species Closest homolog	Homolog Accession No.	Identity (%) / Coverage (%)
A	AUW30720	Thiolase	<i>C. geophilum</i>	OCK92611	63/99
B	AUW30721	Pyridoxine	<i>A. Clavatus</i>	XP_001267783	62/96
C	AUW30722	Nuclear cap binding protein subunit 2	<i>B. gilchristii</i>	OAT12789	95/60
D	AUW30723	Nuclear pore complex protein An-Nup82	<i>B. spectabilis</i>	GAD93109	49/82
E	AUW30724	Hypothetical protein	<i>H. minnesotensis</i>	KJZ6354	26/81
F	AUW30725	Hypothetical protein	<i>C. apollinis</i>	XP_007782562	61/54
G	AUW30726	ABC multidrug transporter	<i>B. spectabilis</i>	GAD95620	75/80
H	AUW30727	Nucleoside transporter	<i>L. palustris</i>	OCK82911	65/95
I	AUW30728	No significant similarity	-	-	-
J	AUW30729	Hypothetical protein	<i>X. heveae</i>	XP_018187302	33/78
K	AUW30730	DUF1682-domain protein	<i>L. palustris</i>	OCK85594	61/85
L	AUW30731	Hypothetical protein	<i>P. decumbens</i>	OQD75607	31/97
M	AUW30732	Gluconate 5-dehydrogenase	<i>E. dermatitidis</i>	XP_009161506	66/85
N	AUW30733	Sorbitol dehydrogenase	<i>S. lycopersici</i>	KNG50061	59/79
O	AUW30734	Amino acid transporter	<i>R. emersonii</i>	XP_01333153	61/91
P	AUW30735	Hypothetical protein	<i>N. parvum</i>	XP_007589192	61/83
Q	AUW30736	Inositol-pentakisphosphate 2-kinase	<i>U. pustulata</i>	SLM41096	42/91
R	AUW30737	Hypothetical protein	<i>Umbilicaria</i>	SLM41103	64/99
1	AUW30738	Nitroreductase	<i>D. robiginosus</i>	WP_050180489	39/79
S	AUW30739	Histone acetyltransferase	<i>U. pustulata</i>	SLM41101	60/86
2	AUW30740	Epimerase/dehydratase	<i>N. parvum</i>	XP_007583116	66/100
T	AUW30741	Hypothetical protein	<i>E. crescens</i>	KKZ60000	72/100
U	AUW30742	No significant similarity	-	-	-
V	AUW30743	Hypothetical protein	<i>A. fumigatus</i>	EDP49293	27/86
W	AUW30744	Hypothetical protein	<i>E. aquamarina</i>	XP_013260036	28/78
X	AUW30745	Hypothetical protein	<i>F. multimorphosa</i>	KIY00977	41/87
3	AUW30746	NRPS (<i>Cu-nrps-1</i>)	<i>M. robertsii</i>	XP_007817828	45/99
Y	AUW30747	Hypothetical protein	<i>P. nodorum</i>	XP_001799385	52/89
Z	AUW30748	ABC transporter	<i>M. majus</i>	XP_014574609	49/97
4	AUW30749	Cytochrome P-450	<i>A. kawachii</i>	GAA91881	49/98
5	AUW30750	Fatty acid synthase	<i>B. graminis</i>	EPQ63903	50/99
6	AUW30751	Fatty acid synthase	<i>B. cinerea</i>	CCD55624	52/96
7	AUW30752	Aminotransferase	<i>A. oryzae</i>	XP_001825739	60/90
AA	AUW30753	Hypothetical protein	<i>S. lycopersici</i>	KNG47193	43/99
8	AUW30754	Cytochrome P-450	<i>P. scopiformis</i>	KUJ18102	41/98
AB	AUW30755	Hypothetical protein	<i>P. italicum</i>	KGO72703	55/95
AC	AUW30756	No significant similarity	-	-	-

Table S8: BLAST statistics for Cluster 4. Reproduced from Bertrand et al. (2018a), Supporting Information file, in accordance with authors' retention of privileges.

Gene (No.)	Protein Accession No.	Protein Putative function(s)	Species Closest homolog	Homolog Accession No.	Identity (%) / Coverage (%)
A	AUW30757	No significant similarity	-	-	-
1	AUW30758	2-dehydropantoate 2-reductase	<i>L. maculans</i>	XP_003833957	31/93
B	AUW30759	No significant similarity	-	-	-
2	AUW30760	NRPS (<i>Cu-nrps-2</i>)	<i>L. maculans</i>	XP_003833959	43/95
C	AUW30761	Capsule polysaccharide biosynthesis protein	<i>S. borealis</i>	ESZ97657	63/98
D	AUW30762	Serine/threonine-protein phosphatase	<i>C. platani</i>	KKF92403	27/98
E	AUW30763	No significant similarity	-	-	-
F	AUW30764	No significant similarity	-	-	-
G	AUW30765	No significant similarity	-	-	-
H	AUW30766	No significant similarity	-	-	-
I	AUW30767	F-box domain	<i>U. pustulata</i>	SLM39731	95/84
J	AUW30768	Hypothetical protein	<i>C. apollinis</i>	XP_007779599	45/95
K	AUW30769	Heme-dependent catalase	<i>A. melanogenum</i>	KEQ63647	81/74
L	AUW30770	No significant similarity	-	-	-
M	AUW30771	No significant similarity	-	-	-
N	AUW30772	Ribonuclease	<i>A. arachidicola</i>	PIG89255	60/96
O	AUW30773	Hypothetical protein	<i>O. maius</i>	KIN08130	48/90
P	AUW30774	C6 finger domain protein	<i>U. pustulata</i>	SLM40476	49/65
Q	AUW30775	No significant similarity	-	-	-
R	AUW30776	2-methylcitrate dehydratase	<i>H. capsulatum</i>	XP_001539627	85/74
S	AUW30777	Hypothetical protein	<i>U. pustulata</i>	SLM40455	76/75
T	AUW30778	Glucosyltransferase	<i>E. weberi</i>	KOS21086	77/81
U	AUW30779	Glucan 1,3-beta-glucosidase	<i>D. rosae</i>	PBP24414	68/89
V	AUW30780	Hypothetical protein	<i>L. palustris</i>	OCK75302	46/98
W	AUW30781	No significant similarity	-	-	-
X	AUW30782	Glycoside hydrolase	<i>B. cinerea</i>	CCD44408	61/92
Y	AUW30783	No significant similarity	-	-	-

Table S9: BLAST statistics for Cluster 5. Reproduced from Bertrand et al. (2018a), Supporting Information file, in accordance with authors' retention of privileges.

Gene (No.)	Protein Accession No.	Protein Putative function(s)	Species Closest homolog	Homolog Accession No.	Identity (%) / Coverage (%)
A	AUW30784	No significant similarity	-	-	-
B	AUW30785	No significant similarity	-	-	-
C	AUW30786	No significant similarity	-	-	-
D	AUW30787	No significant similarity	-	-	-
E	AUW30788	Hypothetical protein	<i>C. geophilum</i>	OCK95795	65/85
F	AUW30789	Glycosyltransferase	<i>D. corticola</i>	XP_020124919	40/87
G	AUW30790	Hypothetical protein	<i>Fungal sp. 14919</i>	GAW17984	39/89
H	AUW30791	Hypothetical protein	<i>P. chlamydospora</i>	KKY14618	57/96
I	AUW30792	Hypothetical protein	<i>F. verticillioides</i>	EWG54928	36/98
1	AUW30793	FAD oxidoreductase	<i>P. tritici-repentis</i>	XP_001930724	58/97
2	AUW30794	NRPS (<i>Cu-nrps-3</i>)	<i>B. cinerea</i>	CCD46068	54/99
J	AUW30795	6-phosphogluconate dehydrogenase	<i>B. cinerea</i>	CCD46066	66/57
K	AUW30796	Isy-1-like splicing factor	<i>P. scopiformis</i>	KUJ09333	84/99
L	AUW30797	Hypothetical protein	<i>O. maius</i>	KIM93913	30/85
M	AUW30798	Hypothetical protein	<i>C. gigas</i>	XP_011414905	35/92
N	AUW30799	Rho guanyl nucleotide exchange factor	<i>R. emersonii</i>	XP_013328390	66/96
O	AUW30802	No significant similarity	-	-	-
P	AUW30800	SAP domain-containing protein	<i>U. pustulata</i>	SLM34974	43/84
Q	AUW30801	Organic solute transporter	<i>U. pustulata</i>	SLM34975	45/80

Table S10: BLAST statistics for Cluster 6. Reproduced from Bertrand et al. (2018a), Supporting Information file, in accordance with authors' retention of privileges.

Gene (No.)	Protein Accession No.	Protein Putative function(s)	Species Closest homolog	Homolog Accession No.	Identity (%) / Coverage (%)
1	AUW30803	α/β -hydrolase	<i>D. seriata</i>	KKY19860	69/98
A	AUW30804	Hypothetical protein	<i>B. spectabilis</i>	GAD92735	57/96
2	AUW30805	O-methyltransferase	<i>G. lozoyensis</i>	XP_008085565	52/97
3	AUW30806	O-methyltransferase	<i>P. scopiformis</i>	KUJ18479	53/99
B	AUW30807	Hypothetical protein	<i>E. pusillum</i>	XP_007803213	38/80
C	AUW30808	Hypothetical protein	<i>E. pusillum</i>	XP_007801142	65/74
4	AUW30809	NRPS (<i>Cu-nrps-4</i>)	<i>B. cinerea</i>	CCD46068	52/99
D	AUW30810	Galactose oxidase	<i>M. phaseolina</i>	EKG13329	34/93
E	AUW30811	No significant similarity	-	-	-
F	AUW30812	Glycoside hydrolase	<i>X. heveae</i>	XP_018189030	71/70

Table S11: BLAST statistics for Cluster 7. Reproduced from Bertrand et al. (2018a), Supporting Information file, in accordance with authors' retention of privileges.

Gene (No.)	Protein Accession No.	Protein Putative function(s)	Species Closest homolog	Homolog Accession No.	Identity (%) / Coverage (%)
A	AUW30813	No significant similarity	-	-	-
B	AUW30814	No significant similarity	-	-	-
C	AUW30815	Transcription factor	<i>L. maculans</i>	XP_003844206	33/83
D	AUW30816	Hypothetical protein	<i>V. mali</i>	KUI67861	43/71
E	AUW30817	Hypothetical protein	<i>T. virens</i>	XP_013956227	27/95
F	AUW30818	Hypothetical protein	<i>T. ophioglossoides</i>	KND87668	42/69
1	AUW30819	NRPS (<i>Cu-nrps-5</i>)	<i>P. scopiformis</i>	KUJ07624	38/99

Table S12: BLAST statistics for Cluster 8. Reproduced from Bertrand et al. (2018a), Supporting Information file, in accordance with authors' retention of privileges.

Gene (No.)	Protein Accession No.	Protein Putative function(s)	Species Closest homolog	Homolog Accession No.	Identity (%) / Coverage (%)
A	AUW30820	Sucrase/ferredoxin-like protein	<i>A. aleyrodis</i>	KZZ91358	69/99
B	AUW30821	No significant similarity	-	-	-
C	AUW30822	Hypothetical protein	<i>P. tritici-repentis</i>	XP_001933264	30/61
D	AUW30823	No significant similarity	-	-	-
E	AUW30824	No significant similarity	-	-	-
F	AUW30825	V-type proton ATPase subunit	<i>E. crescens</i>	KKZ66119	70/99
G	AUW30826	Glycerol-3-phosphate O-acetyltransferase	<i>U. pustulata</i>	SLM35940	69/94
1	AUW30827	NRPS (<i>Cu-nrps-6</i>)	<i>D. ampelina</i>	KKY35466	45/97
2	AUW30828	α/β -hydrolase	<i>G. lozoyensis</i>	XP_008078773	40/97
3	AUW30829	Mannosyltransferase	<i>N. parvum</i>	XP_007587195	54/83
H	AUW30830	MFS transporter	<i>P. scopiformis</i>	KUJ21827	54/94

Table S13: BLAST statistics for Cluster 9. Reproduced from Bertrand et al. (2018a), Supporting Information file, in accordance with authors' retention of privileges.

Gene (No.)	Protein Accession No.	Protein Putative function(s)	Species Closest homolog	Homolog Accession No.	Identity (%) / Coverage (%)
1	AUW30831	NRPS (<i>Cu-nrps-7</i>)	<i>A. pullulans</i>	KEQ85122	48/96
A	AUW30832	ABC transporter	<i>D. seriata</i>	KKY21038	64/97
B	AUW30833	Integrase – catalytic core Gag-Pol polyprotein	<i>P. italicum</i>	KGO69328	46/60
			<i>M. quizhouense</i>	KID81852	42/60
C	AUW30834	Integrase – catalytic core Gag-Pol polyprotein	<i>P. italicum</i>	KGO69328	41/81
			<i>L. niger</i>	KMQ86003	26/72
D	AUW30835	ABC transporter	<i>E. dermatitidis</i>	XP_009154738	64/98

Table S14: BLAST statistics for Cluster 10. Reproduced from Bertrand et al. (2018a), Supporting Information file, in accordance with authors' retainment of privileges.

Gene (No.)	Protein Accession No.	Protein Putative function(s)	Species Closest homolog	Homolog Accession No.	Identity (%) / Coverage (%)
1	AUW30836	Hydroxylase	<i>B. cinerea</i>	CCD46850	56/89
2	AUW30837	Hydroxylase	<i>T. hemipterigena</i>	CEJ87713	47/97
3	AUW30838	Hydroxylase	<i>C. coronate</i>	XP_007726192	36/96
4	AUW30839	NRPS (<i>Cu-nrps-8</i>)	<i>B. cinerea</i>	CCD46857	43/93
A	AUW30840	Choline transport protein	<i>F. oxysporum</i>	ENH75846	54/87

Table S15: BLAST statistics for Cluster 11. Reproduced from Bertrand et al. (2018a), Supporting Information file, in accordance with authors' retainment of privileges.

Gene (No.)	Protein Accession No.	Protein Putative function(s)	Species Closest homolog	Homolog Accession No.	Identity (%) / Coverage (%)
1	AUW30841	Alcohol dehydrogenase	<i>A. parasiticus</i>	KJK63939	47/99
2	AUW30842	Short-chain dehydrogenase / reductase	<i>M. brunnea</i>	XP_007291571	37/95
3	AUW30843	NRPS (<i>Cu-nrps-9</i>)	<i>T. marneffeii</i>	XP_002143254	43/96

Table S16: BLAST statistics for Cluster 12. Reproduced from Bertrand et al. (2018a), Supporting Information file, in accordance with authors' retainment of privileges.

Gene (No.)	Protein Accession No.	Protein Putative function(s)	Species Closest homolog	Homolog Accession No.	Identity (%) / Coverage (%)
A	AUW30844	DNA helicase	<i>R. emersonii</i>	XP_013325547	49/90
B	AUW30845	DNA helicase	<i>R. emersonii</i>	XP_013325547	49/90
1	AUW30846	Pentalenene synthase (Cu-terp-1)	<i>Termitomyces sp.</i>	KNZ73784	38/61
C	AUW30847	No significant similarity	-	-	-
D	AUW30848	No significant similarity	-	-	-
E	AUW30849	No significant similarity	-	-	-
F	AUW30850	G-protein-coupled receptor	<i>G. lozoyensis</i>	XP_008088280	69/88
2	AUW30851	Phytoene dehydrogenase	<i>E. pusillum</i>	XP_007800418	73/90
3	AUW30852	Squalene/phytoene synthase (Cu-terp-2)	<i>M. phaseolina</i>	EKG11672	61/99
4	AUW30853	Retinal pigment epithelial membrane protein	<i>M. brunnea</i>	XP_007288547	66/97
		Carotenoid oxygenase	<i>P. scopiformis</i>	KUJ14379	65/96
G	AUW30854	DEAD/DEAH box RNA helicase	<i>H. capsulatum</i>	EER45752	60/98
H	AUW30855	No significant similarity	-	-	-
I	AUW30856	No significant similarity	-	-	-
J	AUW30857	No significant similarity	-	-	-
K	AUW30858	No significant similarity	-	-	-
L	AUW30859	HC-toxin efflux carrier	<i>M. mycetomatis</i>	KOP45986	70/98
5	AUW30860	D-xylose 1-dehydrogenase	<i>E. weberi</i>	KOS18113	74/99
		Oxidoreductase	<i>T. stipitatus</i>	XP_002479371	66/99
6	AUW30861	PKS-NRPS (Cu-pks-nrps-1)	<i>E. weberi</i>	KOS18114	64/100
7	AUW30862	FAD oxidase	<i>F. avenaceum</i>	KIL88412	65/98
8	AUW30863	6-hydroxy-D-nicotine oxidase	<i>M. quizhouense</i>	KID84875	71/99
9	AUW30864	Fe(II)-oxygenase	<i>F. avenaceum</i>	KIL88413	81/98
10	AUW30865	FAD oxidase	<i>F. avenaceum</i>	KIL88414	69/99
		Hydroxylase	<i>M. mycetomatis</i>	KOP45991	58/87
M	AUW30866	Zinc C6 fungal DNA regulatory protein	<i>M. quizhouense</i>	KID84879	38/66
11	AUW30867	Dehydrogenase	<i>A. flavus</i>	KJJ37178	73/99
N	AUW30868	Regulator-vacuolar morphogenesis	<i>C. coronate</i>	XP_007727281	58/100
O	AUW30869	Microtubule-associated protein	<i>P. minimum</i>	XP_007917371	71/100
P	AUW30870	F-box and WD40 domain- containing protein	<i>A. calidoustus</i>	CEL07183	65/99
Q	AUW30871	No significant similarity	-	-	-
12	AUW30872	Dehydrogenase	<i>T. cellulolyticus</i>	GAM39238	78/100
		NAD-binding oxidoreductase	<i>B. spectabilis</i>	GAD96687	77/100
R	AUW30873	Hypothetical protein	<i>B. spectabilis</i>	GAD92385	43/97
S	AUW30874	No significant similarity	-	-	-
T	AUW30875	Hypothetical protein	<i>P. musae</i>	KXT00830	45/62

Table S17: BLAST statistics for Cluster 13. Reproduced from Bertrand et al. (2018a), Supporting Information file, in accordance with authors' retainment of privileges.

Gene (No.)	Protein Accession No.	Protein Putative function(s)	Species Closest homolog	Homolog Accession No.	Identity (%) / Coverage (%)
A	AUW30876	Hypothetical protein	<i>R. mackenziei</i>	XP_013273914	33/91
B	AUW30877	Hypothetical protein	<i>G. stellatum</i>	OCL09011	28/89
C	AUW30878	Hypothetical protein	<i>O. maius</i>	KIM96162	52/90
D	AUW30879	No significant similarity	-	-	-
E	AUW30880	Transcription factor	<i>B. cinerea</i>	CCD50292	39/89
F	AUW30881	Hypothetical protein	<i>S. borealis</i>	ESZ92534	39/95
G	AUW30882	No significant similarity	-	-	-
H	AUW30883	Hypothetical protein	<i>P. brasiliense</i>	CEJ62700	49/67
I	AUW30884	Hypothetical protein	<i>Z. tritici</i>	SMR44684	29/94
J	AUW30885	No significant similarity	-	-	-
1	AUW30886	Aldo-keto reductase	<i>A. otae</i>	XP_002842630	28/99
K	AUW30887	No significant similarity	-	-	-
L	AUW30888	Amino acid transporter	<i>C. geophilum</i>	OCL00641	62/89
M	AUW30889	No significant similarity	-	-	-
N	AUW30890	Hypothetical protein	<i>A. namibiae</i>	XP_013431221	41/72
2	AUW30891	Alpha/beta-hydrolase	<i>G. lozoyensis</i>	XP_008076678	56/88
		SAM-methyltransferase	<i>Pyrenochaeta sp.</i>	OAL55343	49/76
O	AUW30892	No significant similarity	-	-	-
P	AUW30893	B-(1-6) glucan synthase	<i>I. fumosorosea</i>	OAR00096	61/79
Q	AUW30894	No significant similarity	-	-	-
3	AUW30895	PKS-like thioesterase	<i>A. niger</i>	OWW37439	42/98
R	AUW30896	MFS transporter	<i>G. lozoyensis</i>	XP_008085945	51/70
S	AUW30897	RTA-like protein	<i>A. oryzae</i>	OOO12024	43/75
T	AUW30898	Hypothetical protein	<i>E. pusillum</i>	XP_007802714	39/75
4	AUW30899	PKS-NRPS (<i>Cu-pks-nrps-2</i>)	<i>G. lozoyensis</i>	XP_008085943	56/98
5	AUW30900	Sulfite oxidase	<i>A. namibiae</i>	XP_013432137	73/96
		Oxidoreductase	<i>A. pullulans</i>	KEQ99319	74/96
U	AUW30901	Zinc C6 transcription factor	<i>O. piceae</i>	EPE07108	40/76
6	AUW30902	Ent-kaurene synthase	<i>M. oryzae</i>	ELQ58520	31/95
		(<i>Cu-terp-3</i>)			

Table S18: BLAST statistics for Cluster 14. Reproduced from Bertrand et al. (2018a), Supporting Information file, in accordance with authors' retainment of privileges.

Gene (No.)	Protein Accession No.	Protein Putative function(s)	Species Closest homolog	Homolog Accession No.	Identity (%) / Coverage (%)
A	AUW30903	Arginine biosynthesis protein	<i>N. gypsea</i>	XP_003174163	78/99
B	AUW30904	No significant similarity	-	-	-
C	AUW30905	1-phosphatidylinositol-3- phosphate 5-kinase	<i>U. pustulata</i>	SLM37344	67/81
D	AUW30906	No significant similarity	-	-	-
1	AUW30907	PKS Enoylreductase	<i>P. italicum</i>	KGO65249	68/86
		Alcohol dehydrogenase	<i>A. niger</i>	EHA17997	66/86
E	AUW30908	MFS transporter	<i>D. seriata</i>	KKY16916	55/97
2	AUW30909	FAD oxidase	<i>P. scopiformis</i>	KUJ08736	47/98
F	AUW30910	Hypothetical protein	<i>T. islandicus</i>	CRG89197	51/84
G	AUW30911	No significant similarity	-	-	-
3	AUW30912	PKS-NRPS (<i>Cu-pks-nrps-3</i>)	<i>A. clavatus</i>	XP_001270445	47/98
H	AUW30913	Glycoside hydrolase	<i>O. maius</i>	KIM99643	56/99
I	AUW30914	Hypothetical protein	<i>C. carrionii</i>	XP_008722044	36/63
J	AUW30915	Hypothetical protein	<i>E. pusillum</i>	XP_007800149	57/99

Table S19: BLAST statistics for Cluster 15. Reproduced from Bertrand et al. (2018a), Supporting Information file, in accordance with authors' retainment of privileges.

Gene (No.)	Protein Accession No.	Protein Putative function(s)	Species Closest homolog	Homolog Accession No.	Identity (%) / Coverage (%)
A	AUW30916	Hypothetical protein	<i>Pseudogymnoascus</i>	KFY65896	52/90
B	AUW30917	Thiosulfate sulfurtransferase	<i>U. pustulata</i>	SLM34737	71/91
C	AUW30918	No significant similarity	-	-	-
D	AUW30919	PLC-like phosphodiesterase	<i>P. scopiformis</i>	XP_018066270	66/92
1	AUW30920	Cytochrome P-450	<i>U. pustulata</i>	SLM34165	59/98
2	AUW30921	2-succinylbenzoate-CoA ligase	<i>S. lycopersici</i>	KNG47290	67/96
E	AUW30922	BolA-like protein	<i>U. pustulata</i>	SLM34753	80/52
F	AUW30923	GTP-binding protein Ypt5	<i>U. pustulata</i>	SLM38500	90/72
G	AUW30924	Phosphoinositide phosphatase	<i>U. pustulata</i>	SLM38501	75/97
H	AUW30925	Hypothetical protein	<i>A. oryzae</i>	XP_001821260	32/69
I	AUW30926	Nicotinate phosphoribosyltransferase	<i>U. pustulata</i>	SLM34755	61/94
J	AUW30927	No significant similarity	-	-	-
K	AUW30928	Hypothetical protein	<i>Pseudogymnoascus</i>	KFX93959	63/61
L	AUW30929	Hypothetical protein	<i>Fungal sp. 11243</i>	GAM88438	57/74
M	AUW30930	Translation initiation factor	<i>X. heveae</i>	XP_018189790	60/78
N	AUW30931	No significant similarity	-	-	-
O	AUW30932	Hypothetical protein	<i>M. phaseolina</i>	EKG13067	30/76
P	AUW30933	Hypothetical protein	<i>C. gloeosporioides</i>	EQB59441	34/77
Q	AUW30934	Hypothetical protein	<i>P. chrysogenum</i>	KZN85741	37/81
3	AUW30935	Short-chain dehydrogenase / reductase	<i>U. pustulata</i>	SLM34748	73/98
R	AUW30936	No significant similarity	-	-	-
S	AUW30937	No significant similarity	-	-	-
T	AUW30938	40S ribosomal protein	<i>P. brasilianum</i>	CEJ55024	91/83
U	AUW30939	UPF0183 domain protein	<i>P. brasilianum</i>	CEJ55020	49/98
V	AUW30940	Isoleucyl-tRNA synthetase	<i>T. stipitatus</i>	XP_002479535	51/98
W	AUW30941	No significant similarity	-	-	-
X	AUW30942	No significant similarity	-	-	-
Y	AUW30943	Hypothetical protein	<i>B. cinerea</i>	CCD51123	58/99
4	AUW30944	4-coumarate-CoA ligase	<i>E. dermatitidis</i>	XP_009156000	47/79
		AMP-dependent synthetase	<i>P. camemberti</i>	CRL23284	47/80
5	AUW30945	Phenylalanine ammonia lyase	<i>L. vulpine</i>	BAN29055	64/94
Z	AUW30946	Opsin-1	<i>P. attae</i>	KPI45543	70/98
6	AUW30947	Threonine dehydratase	<i>B. spectabilis</i>	GAD96042	74/97
AA	AUW30948	Hypothetical protein	<i>M. brunnea</i>	XP_007288818	47/67
7	AUW30949	Multicopper oxidase	<i>P. chlamydospora</i>	KKY23553	53/88
		Laccase-2	<i>Fusarium fujikuroi</i>	KLO79025	51/90
8	AUW30950	Cytochrome P-450	<i>E. lata</i>	XP_007795124	38/98
9	AUW30951	Squalene cyclase(Cu-terp-4)	<i>A. oryzae</i>	EIT72300	53/98

Table S20: BLAST statistics for Cluster 16. Reproduced from Bertrand et al. (2018a), Supporting Information file, in accordance with authors' retainment of privileges.

Gene (No.)	Protein Accession No.	Protein Putative function(s)	Species Closest homolog	Homolog Accession No.	Identity (%) / Coverage (%)
A	AUW30952	Hypothetical protein	<i>C. immunda</i>	XP_016245517	43/92
B	AUW30953	STU1-mitotic spindle protein	<i>D. seriata</i>	OMP82116	46/96
C	AUW30954	HET-domain-containing protein	<i>C. geophilum</i>	OCL00662	55/78
D	AUW30955	Plant basic secretory protein	<i>D. corticola</i>	XP_020125420	59/73
E	AUW30956	Phosphatases II	<i>P. scopiformis</i>	XP_018071140	57/78
F	AUW30957	Adenylyl cyclase-like protein	<i>C. geophilum</i>	OCK99921	57/95
G	AUW30958	Ribokinase-like protein	<i>U. pustulata</i>	SLM33835	40/68
H	AUW30959	No significant similarity	-	-	-
I	AUW30960	Mitochondrial 54S ribosomal protein	<i>A. udagawae</i>	GAO88988	53/99
J	AUW30961	No significant similarity	-	-	-
K	AUW30962	Ribosomal protein	<i>U. pustulata</i>	SLM33838	52/89
L	AUW30963	Protein-tyrosine phosphatase receptor/non-receptor type	<i>M. phaseolina</i>	EKG14049	45/92
M	AUW30964	No significant similarity	-	-	-
N	AUW30965	Hypothetical protein	<i>N. parvum</i>	XP_007587955	47/95
1	AUW30966	Dimethylallyl tryptophan synthase (Cu-terp-5)	<i>B. cinerea</i>	CCD48756	44/93
2	AUW30967	GNAT acetyltransferase	<i>A. flavus</i>	KOC10883	35/77
3	AUW30968	Cytochrome P-450	<i>B. cinerea</i>	EMR90165	57/94
4	AUW30969	Cytochrome P-450	<i>G. lozoyensis</i>	XP_008079960	45/97
O	AUW30970	Integral membrane protein	<i>B. spectabilis</i>	GAD95310	31/91
P	AUW30971	No significant similarity	-	-	-
Q	AUW30972	Minichromosome maintenance protein	<i>C. apollinis</i>	XP_007784541	76/99
		DNA-replication licencing factor	<i>N. parvum</i>	XP_007588344	75/100
R	AUW30973	Hypothetical protein	<i>U. pustulata</i>	SLM39090	64/68
S	AUW30974	Hypothetical protein	<i>A. rabiei</i>	KZM18494	35/63

Table S21: BLAST statistics for Cluster 17. Reproduced from Bertrand et al. (2018a), Supporting Information file, in accordance with authors' retainment of privileges.

Gene (No.)	Protein Accession No.	Protein Putative function(s)	Species Closest homolog	Homolog Accession No.	Identity (%) / Coverage (%)
A	AUW30975	Molecular chaperone, heat shock protein Hsp40, DnaJ	<i>P. camemberti</i>	CRL27427	74/94
B	AUW30976	Hypothetical protein	<i>T. marneffeii</i>	XP_002143165	33/81
1	AUW30977	Cytochrome P-450	<i>A. flavus</i>	XP_002378813	40/92
2	AUW30978	Farnesyl transferase (<i>Cu-terp-6</i>)	<i>S. borealis</i>	ESZ91924	70/100
C	AUW30979	Small nucleolar ribonucleo-protein complex subunit	<i>A. fumigatus</i>	KMK57755	45/99
D	AUW30980	G1/S-specific cyclin Cln1	<i>A. oryzae</i>	XP_001818433	62/89
E	AUW30981	Cysteine protease	<i>C. higginsianum</i>	CCF41055	27/82
F	AUW30982	GPR/FUN34 family protein	<i>T. marneffeii</i>	XP_002150578	72/95
G	AUW30983	Png1p protein	<i>U. reesii</i>	XP_002542827	67/92

Table S22: BLAST statistics for Cluster 18. Reproduced from Bertrand et al. (2018a), Supporting Information file, in accordance with authors' retainment of privileges.

Gene (No.)	Protein Accession No.	Protein Putative function(s)	Species Closest homolog	Homolog Accession No.	Identity (%) / Coverage (%)
A	AUW30984	Hypothetical protein	<i>P. chlamydospora</i>	KKY29237	44/99
B	AUW30985	Rheb small monomeric GTPase RhbA	<i>P. tritici-repentis</i>	XP_001931327	78/100
C	AUW30986	Wings apart-like protein	<i>M. phaseolina</i>	EKG15885	33/67
D	AUW30987	Uracil phosphoribosyltransferase	<i>P. fici</i>	XP_007838875	84/100
E	AUW30988	Transcription elongation factor	<i>A. flavus</i>	XP_002382925	55/100
F	AUW30989	GTP cyclohydrolase II RibA	<i>M. phaseolina</i>	EKG15882	75/98
1	AUW30990	α/β -hydrolase	<i>M. phaseolina</i>	EKG15887	63/98
G	AUW30991	No significant similarity	-	-	-
2	AUW30992	Cytochrome P-450	<i>A. alternata</i>	OWY47057	53/94
3	AUW30993	Short-chain dehydrogenase/reductase	<i>P. rogueforti</i>	CDM31317	52/93
4	AUW30994	Oxidoreductase	<i>A. nidulans</i>	CBF83051	40/99
5	AUW30995	Aristolochene synthase (<i>Cu-terp-7</i>)	<i>A. terreus</i>	Q9UR08	55/98
6	AUW30996	Alpha/beta-hydrolase	<i>D. misasensis</i>	WP_084571630	34/87
H	AUW30997	Hypothetical protein	<i>U. pustulata</i>	SLM38976	51/94
I	AUW30998	Transport protein Sec61 alpha subunit	<i>A. clavatus</i>	XP_001274237	72/84
J	AUW30998	Arginine biosynthesis protein	<i>N. gypsea</i>	XP_003174163	78/99

Table S23: BLAST statistics for Cluster 19. Reproduced from Bertrand et al. (2018a), Supporting Information file, in accordance with authors' retainment of privileges.

Gene (No.)	Protein Accession No.	Protein Putative function(s)	Species Closest homolog	Homolog Accession No.	Identity (%) / Coverage (%)
A	AUW31000	Glycerol dehydrogenase	<i>A. clavatus</i>	XP_001269267	82/100
B	AUW31001	No significant similarity	-	-	-
1	AUW31002	Cytochrome P-450	<i>P. tritici-repentis</i>	XP_001942168	40/92
2	AUW31003	Squalene cyclase (Cu-terp-8)	<i>F. hepatica</i>	KIY45656	60/97
C	AUW31004	Hypothetical protein	<i>N. ditissima</i>	KPM42694	44/96
D	AUW31005	Hypothetical protein	<i>P. brasillianum</i>	CEJ57697	47/65
E	AUW31006	No significant similarity	-	-	-

Table S24: BLAST statistics for Cluster 20. Reproduced from Bertrand et al. (2018a), Supporting Information file, in accordance with authors' retainment of privileges.

Gene (No.)	Protein Accession No.	Protein Putative function(s)	Species Closest homolog	Homolog Accession No.	Identity (%) / Coverage (%)
1	AUW31007	Cytochrome P-450	<i>P. rogueforti</i>	CDM31319	44/96
2	AUW31008	Short-chain dehydrogenase / reductase	<i>C. gloeosporioides</i>	XP_007286745	64/100
3	AUW31009	Cytochrome P-450	<i>M. phaseolina</i>	EKG12700	62/93
4	AUW31010	Aristolochene synthase (Cu-terp-9)	<i>A. ustus</i>	KIA75982	59/99
5	AUW31011	Cytochrome P-450	<i>C. fioriniae</i>	XP_007596781	35/98
		Transferase	<i>P. expansum</i>	KGO70752	41/98
6	AUW31012	Type I PKS	<i>P. membranacea</i>	AEE65373	53/99

Table S25: BLAST statistics for Cluster 21. Reproduced from Bertrand et al. (2018a), Supporting Information file, in accordance with authors' retainment of privileges.

Gene (No.)	Protein Accession No.	Protein Putative function(s)	Species Closest homolog	Homolog Accession No.	Identity (%) / Coverage (%)
A	AUW31013	No significant similarity	-	-	-
B	AUW31014	No significant similarity	-	-	-
C	AUW31015	No significant similarity	-	-	-
D	AUW31016	No significant similarity	-	-	-
E	AUW31017	No significant similarity	-	-	-
F	AUW31018	No significant similarity	-	-	-
G	AUW31019	No significant similarity	-	-	-
H	AUW31020	No significant similarity	-	-	-
I	AUW31021	No significant similarity	-	-	-
J	AUW31022	No significant similarity	-	-	-
K	AUW31023	Kinase	<i>U. pustulata</i>	SLM39639	25/86
L	AUW31024	No significant similarity	-	-	-
M	AUW31025	No significant similarity	-	-	-
N	AUW31026	Retrotransposable element	<i>U. pustulata</i>	SLM36746	42/83
O	AUW31027	Hypothetical protein	<i>Pseudogymnoascus</i>	KFZ14235	37/78
P	AUW31028	No significant similarity	-	-	-
Q	AUW31029	3-ketoacyl-acyl carrier protein reductase	<i>P. subalpine</i>	CZR56198	68/88
R	AUW31030	Hypothetical protein	<i>C. coronate</i>	XP_007719734	74/100
S	AUW31031	Hypothetical protein	<i>E. xenobiotica</i>	XP_013310093	71/000
T	AUW31032	Transcriptional regulator	<i>P. camemberti</i>	CRL18704	27/87
1	AUW31033	O-methyltransferase	<i>N. parvum</i>	XP_007587219	41/99
2	AUW31034	Short-chain dehydrogenase / reductase	<i>M. phaseolina</i>	EKG20376	55/89
3	AUW31035	Sclyatone dehydratase	<i>C. apollinis</i>	XP_007776594	54/94
4	AUW31036	FAD oxidase	<i>O. sinensis</i>	EQL03992	40/96
5	AUW31037	Aldehyde reductase	<i>T. harzianum</i>	KKO98764	50/85
U	AUW31038	Hypothetical protein	<i>S. commune</i>	XP_003035435	65/99
6	AUW31039	FAD oxidase	<i>N. parvum</i>	XP_007587222	57/95
7	AUW31040	PKS (<i>Cu-nr-pks-1</i>)	<i>C. grayi</i>	ADX36085	89/100
V	AUW31041	Hypothetical protein	<i>C. bantiana</i>	KIW95639	30/95
W	AUW31042	Na ⁺ /H ⁺ antiporter	<i>B. spectabilis</i>	GAD97611	43/99

Table S26: BLAST statistics for Cluster 22. Reproduced from Bertrand et al. (2018a), Supporting Information file, in accordance with authors' retainment of privileges.

Gene (No.)	Protein Accession No.	Protein Putative function(s)	Species Closest homolog	Homolog Accession No.	Identity (%) / Coverage (%)
A	AUW31043	Hypothetical protein	<i>C. epimyces</i>	XP_007738320	45/90
1	AUW31044	Short-chain dehydrogenase / reductase	<i>A. Clavatus</i>	XP_001276386	64/98
2	AUW31045	PKS-like gene fragment	<i>C. Posadasii</i>	XP_003071594	39/95
3	AUW31046	PKS-like gene fragment	<i>P. omphalodes</i>	CCX04911	50/93
4	AUW31047	PKS-like gene fragment	<i>C. posadasii</i>	XP_003071594	41/96
5	AUW31048	PKS-like gene fragment	<i>M. phaseolina</i>	EKG19938	66/100
B	AUW31049	Hypothetical protein	<i>P. nodorum</i>	XP_001798604	62/97
C	AUW31050	C6 zinc transcription factor	<i>T. marneffeii</i>	XP_002149521	24/94
6	AUW31051	Cytochrome P-450 ("MPAO")	<i>C. uncialis</i>	ALA62324	100/100
7	AUW31052	PKS ("MPAS") (<i>Cu-nr-pks-2</i>)	<i>C. uncialis</i>		100/100
D	AUW31053	Hypothetical protein	<i>O. maius</i>	KIM99259	62/90
E	AUW31054	Hypothetical protein	<i>O. maius</i>	KIM99260	61/99
F	AUW31055	Hypothetical protein	<i>B. oryzae</i>	XP_007692284	58/84
G	AUW31077	No significant similarity	-	-	-
H	AUW31056	No significant similarity	-	-	-
I	AUW31057	Ankyrin repeat protein	<i>A. niger</i>	GAQ41134	37/66
J	AUW31058	RNA-binding domain-containing protein	<i>G. stellatum</i>	OCL02225	58/82
K	AUW31059	RING-6	<i>A. otae</i>	XP_002845934	91/66
L	AUW31060	Leucine-rich repeat protein	<i>M. roreri</i>	XP_007843355	32/73
M	AUW31061	Ubiquitin C-terminal hydrolase	<i>C. immitis</i>	XP_001242887	58/99
N	AUW31062	Malic enzyme	<i>U. pustulata</i>	SLM39588	63/93
O	AUW31063	Metallo-depedenent phosphatase	<i>U. pustulata</i>	SLM39586	59/91
P	AUW31064	Mannosyl-oligosaccharide-alpha-mannosidase	<i>U. pustulata</i>	SLM39597	64/94
Q	AUW31065	Phenylalanyl-tRNA synthetase	<i>U. pustulata</i>	SLM39598	87/88
R	AUW31066	No significant similarity	-	-	-
S	AUW31067	Lysophospholipase II	<i>E. dermatitidis</i>	XP_009152259	43/77
T	AUW31068	Glycoside hydrolase	<i>Hypoxylon sp.</i>	OTA66092	66/93
U	AUW31069	No similarity found			
V	AUW31070	IgE-binding domain	<i>C. higginsianum</i>	XP_018163109	82/99
W	AUW31071	Hypothetical protein	<i>B. percursus</i>	OJD26432	38/60
X	AUW31072	No significant similarity	-	-	-
Y	AUW31073	No significant similarity	-	-	-
Z	AUW31074	No significant similarity	-	-	-
AA	AUW31075	Telomere capping protein 1	<i>M. brunnea</i>	XP_007291825	69/96
AB	AUW31076	Hypothetical protein	<i>U. pustulata</i>	SLM39581	50/82

Table S27: BLAST statistics for Cluster 23. Reproduced from Bertrand et al. (2018a), Supporting Information file, in accordance with authors' retainment of privileges.

Gene (No.)	Protein Accession No.	Protein Putative function(s)	Species Closest homolog	Homolog Accession No.	Identity (%) / Coverage (%)
A	AUW31078	DnaJ domain	<i>U. pustulata</i>	SLM33480	56/95
B	AUW31079	No significant similarity	-	-	-
C	AUW31080	Tyrosine/serine phosphatase-like protein	<i>G. stellatum</i>	OCL06553	51/86
D	AUW31081	No significant similarity	-	-	-
E	AUW31082	Fraq1/DRAM/Sfk1	<i>U. pustulata</i>	SLM33472	71/79
F	AUW31083	No significant similarity	-	-	-
G	AUW31084	Integral membrane protein	<i>T. cellulolyticus</i>	GAM43264	34/93
H	AUW31085	Homoisocitrate dehydrogenase	<i>C. platani</i>	KKF94604	80/99
I	AUW31086	Pyruvate kinase	<i>L. maculans</i>	XP_003837979	81/100
J	AUW31087	Hypothetical protein	<i>M. brunnea</i>	XP_007296613	57/91
K	AUW31088	ABC transporter	<i>A. nomius</i>	KNG87138	76/100
L	AUW31089	Aspartate aminotransferase	<i>P. brasiliense</i>	CEJ55728	66/97
M	AUW31090	ER-derived vesicles protein	<i>A. melanogenum</i>	KEQ63097	96/100
N	AUW31091	mRNA splicing factor	<i>H. capsulatum</i>	EGC42727	64/99
O	AUW31092	No significant similarity	-	-	-
P	AUW31093	Mitochondria protoheme IX farnesyltransferase	<i>X. heveae</i>	XP_018191753	73/70
Q	AUW31094	No significant similarity	-	-	-
R	AUW31095	No significant similarity	-	-	-
S	AUW31096	Zinc transcription factor	<i>G. lozoyensis</i>	XP_008084588	49/98
1	AUW31097	PKS (<i>Cu-nr-pks-3</i>)	<i>M. ruber</i>	ALN44200	64/99
2	AUW31098	Alcohol dehydrogenase	<i>R. emersonii</i>	XP_013328612	65/100
3	AUW31099	Monooxygenase	<i>G. lozoyensis</i>	XP_008084594	64/99
T	AUW31100	Calcium channel	<i>M. mycetomatis</i>	KOP42940	51/98
U	AUW31101	Hypothetical protein	<i>E. pusillum</i>	XP_007801362	73/100
V	AUW31102	MFS transporter	<i>P. italicum</i>	KIM97024	66/99
W	AUW31103	Transcription factor	<i>M. purpureus</i>	AGN98100	53/97
4	AUW31104	Monooxygenase	<i>T. stipitatus</i>	XP_002340035	62/99
		Hydroxylase	<i>T. marneffeii</i>	XP_002149772	62/97
5	AUW31105	FAD oxidase	<i>G. lozoyensis</i>	XP_008084587	53/99
6	AUW31106	Trichothecene 3-O- acetyltransferase	<i>M. mycetomatis</i>	KOP48766	58/99
X	AUW31107	MFS transporter	<i>G. lozoyensis</i>	XP_008084593	63/99
7	AUW31108	Hydrolase	<i>G. lozoyensis</i>	XP_008084584	69/96
		Esterase	<i>M. mycetomatis</i>	KOP48768	62/98
		Oxidoreductase	<i>A. oryzae</i>	XP_001823616	60/97
8	AUW31109	Oxidoreductase	<i>M. ruber</i>	ALN44201	71/93
		Short-chain dehydrogenase / reductase	<i>T. cellulolyticus</i>	GAM33974	61/95
Y	AUW31110	26S proteasome regulatory subunit	<i>C. immitis</i>	XP_001241676	66/99

Table S28: BLAST statistics for Cluster 24. Reproduced from Bertrand et al. (2018a), Supporting Information file, in accordance with authors' retainment of privileges.

Gene (No.)	Protein Accession No.	Protein Putative function(s)	Species Closest homolog	Homolog Accession No.	Identity (%) / Coverage (%)
A	AUW31111	No significant similarity	-	-	-
B	AUW31112	Serine hydroxymethyltransferase	<i>U. pustulata</i>	SLM37171	86/93
C	AUW31113	Hypothetical protein	<i>A. candida</i>	CCI43743	29/73
D	AUW31114	Hypothetical protein	<i>B. spectabilis</i>	GAD95252	31/62
E	AUW31115	Hypothetical protein	<i>U. pustulata</i>	SLM33372	41/65
F	AUW31116	Hypothetical protein	<i>G. lozoyensis</i>	XP_008078674	46/70
G	AUW31117	No significant similarity	-	-	-
H	AUW31118	Six-hairpin glycosidase	<i>G. lozoyensis</i>	XP_008083366	51/95
I	AUW31119	No significant similarity	-	-	-
J	AUW31120	No significant similarity	-	-	-
1	AUW31121	PKS (<i>Cu-nr-pks-4</i>)	<i>S. borealis</i>	ESZ91464	58/99
2	AUW31122	Cytochrome P-450	<i>V. dahliae</i>	XP_009654521	38/97
K	AUW31123	Sugar transporter	<i>T. reesei</i>	ETR99312	63/91
L	AUW31124	No significant similarity	-	-	-
M	AUW31125	Ribonuclease	<i>N. parvum</i>	XP_007580708	62/83
N	AUW31126	Caspase-like protein	<i>G. lozoyensis</i>	XP_008086178	84/71
O	AUW31127	Isopenicillin N synthetase	<i>B. spectabilis</i>	GAD93339	61/91
		Clayaminate synthase-like protein	<i>G. lozoyensis</i>	XP_008077427	61/88
		1-aminocyclopropane-1- carboxylate oxidase	<i>P. attae</i>	XP_018002059	63/88
		Oxoglutarate/iron- dependent dioxygenase	<i>P. accitanis</i>	PCH03280	60/87
		Gibberellin 20-oxidase	<i>R. emersonii</i>	XP_013327414	58/90
		Thymine dioxygenase	<i>D. corticola</i>	XP_020130585	58/88
P	AUW31128	Glycoside hydrolase	<i>L. palustris</i>	OCK80932	58/92
Q	AUW31129	Hypothetical protein	<i>E. spinifera</i>	XP_016237774	41/66
R	AUW31130	No significant similarity	-	-	-
S	AUW31131	Hypothetical protein	<i>P. sporulosa</i>	XP_018031724	61/60
T	AUW31132	Hypothetical protein	<i>P. arizonense</i>	XP_022483471	40/82

Table S29: BLAST statistics for Cluster 25. Reproduced from Bertrand et al. (2018a), Supporting Information file, in accordance with authors' retainment of privileges.

Gene (No.)	Protein Accession No.	Protein Putative function(s)	Species Closest homolog	Homolog Accession No.	Identity (%) / Coverage (%)
A	AUW31133	C2h2 conidiation transcription factor	<i>U. pustulata</i>	SLM34383	78/90
B	AUW31134	Hypothetical protein	<i>P. subalpine</i>	CZR67757	39/87
C	AUW31135	No significant similarity	-	-	-
D	AUW31136	No significant similarity	-	-	-
E	AUW31137	No significant similarity	-	-	-
1	AUW31138	2-amino-3-carboxymuconate -6-semialdehyde decarboxylase Amidohydrolase	<i>B. bassiana</i>	XP_008601571	52/86
2	AUW31139	PKS (<i>Cu-nr-pks-5</i>)	<i>A. parasiticus</i>	KJK68130	53/85
F	AUW31140	Ankyrin repeat protein	<i>A. oryzae</i>	EIT78482	39/98
G	AUW31141	No significant similarity	<i>N. fischeri</i>	XP_001258750	49/81
H	AUW31142	Hypothetical protein	-	-	-
I	AUW31143	Transcription elongation s-ii	<i>P. fici</i>	XP_007839463	39/99
J	AUW31144	mRNA export factor mex67	<i>U. pustulata</i>	SLM34221	58/92
K	AUW31145	No significant similarity	<i>H. capsulatum</i>	EEH11278	47/95
L	AUW31146	Hypothetical protein	-	-	-
M	AUW31147	No significant similarity	<i>U. pustulata</i>	SLM34369	64/93
			-	-	-

Table S30: BLAST statistics for Cluster 26. Reproduced from Bertrand et al. (2018a), Supporting Information file, in accordance with authors' retainment of privileges.

Gene (No.)	Protein Accession No.	Protein Putative function(s)	Species Closest homolog	Homolog Accession No.	Identity (%) / Coverage (%)
A	AUW31148	Ribosome biogenesis protein	<i>E. pusillum</i>	XP_007802122	87/100
B	AUW31149	No significant similarity	-	-	-
C	AUW31150	Hypothetical protein	<i>C. europaea</i>	XP_008718010	50/68
D	AUW31151	Hypothetical protein	<i>G. lozoyensis</i>	XP_008078151	38/95
1	AUW31152	PKS (<i>Cu-nr-pks-6</i>)	<i>C. macilenta</i>	AFB81352	94/99
E	AUW31153	Transcription factor	<i>C. lagenaria</i>	BAE98094	32/97
2	AUW31154	Short-chain dehydrogenase / reductase	<i>M. phaseolina</i>	EKG20376	68/80
F	AUW31155	No significant similarity	-	-	-
3	AUW31156	Scytalone dehydratase	<i>E. aquamarina</i>	XP_013257419	70/92
4	AUW31157	Monooxygenase	<i>R. secalis</i>	CZT49571	31/87
G	AUW31158	Chromate transport protein	<i>P. secalis</i>	CZT47223	55/98
H	AUW31159	No significant similarity	-	-	-
I	AUW31160	Hypothetical protein	<i>G. stellatum</i>	OCL07820	37/69
J	AUW31161	MFS transporter	<i>P. rogueforti</i>	CDM36445	52/73
5	AUW31162	Cytochrome P-450	<i>T. stipitatus</i>	XP_002341512	53/95
K	AUW31163	Hypothetical protein	<i>A. fumigatus</i>	KMK54740	51/66
L	AUW31164	GABA permease	<i>A. fumigatus</i>	EDP50382	65/91
M	AUW31165	No significant similarity	-	-	-

Table S31: BLAST statistics for Cluster 27. Reproduced from Bertrand et al. (2018a), Supporting Information file, in accordance with authors' retainment of privileges.

Gene (No.)	Protein Accession No.	Protein Putative function(s)	Species Closest homolog	Homolog Accession No.	Identity (%) / Coverage (%)
A	AUW31166	No significant similarity	-	-	-
B	AUW31167	MFS transporter	<i>B. spectabilis</i>	GAD95989	57/77
C	AUW31168	Hypothetical protein	<i>U. pustulata</i>	SLM36183	50/72
D	AUW31169	No significant similarity	-	-	-
E	AUW31170	L-galactose dehydrogenase	<i>C. apollinis</i>	XP_007780585	63/86
F	AUW31171	No significant similarity	-	-	-
G	AUW31172	Pex24p-domain-containing protein	<i>C. geophilum</i>	OCK92340	56/99
H	AUW31173	Hypothetical protein	<i>C. europaea</i>	XP_008720218	58/93
I	AUW31174	Hypothetical protein	<i>Pseudogymnoascus sp.</i>	KFY25369	29/90
J	AUW31175	No significant similarity	-	-	-
K	AUW31176	Glucan endo-1,3-alpha-glucosidase agn1	<i>F. oxysporum</i>	ENH75409	46/94
1	AUW31177	PKS (<i>Cu-nr-pks-7</i>)	<i>C. grayi</i>	ADM79459	92/99
2	AUW31178	Cytochrome P-450	<i>C. grayi</i>	ADM79460	83/100
3	AUW31179	O-methyltransferase	<i>C. grayi</i>	ADM79461	80/100

Table S32: BLAST statistics for Cluster 28. Reproduced from Bertrand et al. (2018a), Supporting Information file, in accordance with authors' retainment of privileges.

Gene (No.)	Protein Accession No.	Protein Putative function(s)	Species Closest homolog	Homolog Accession No.	Identity (%) / Coverage (%)
A	AUW31180	No significant similarity	-	-	-
B	AUW31181	MFS transporter	<i>G. lozoyensis</i>	XP_008084574	62/98
1	AUW31182	Phenol-2-monooxygenase	<i>C. metacoralifera</i>	ADR00967	79/99
2	AUW31183	KR-DH domain-like	<i>C. gloeosporioides</i>	XP_007286914	34/98
3	AUW31184	PKS (<i>Cu-nr-pks-8</i>)	<i>C. grayi</i>	ADX36086	73/99
C	AUW31185	C6 transcription factor	<i>C. orbiculare</i>	ENH79621	29/78
4	AUW31186	Tryptophan-2-halogenase	<i>T. ophioglossoides</i>	KND87338	59/93
		Non-heme iron-dependent halogenase	<i>S. borealis</i>	ESZ96411	56/98
5	AUW31187	O-methyltransferase	<i>P. scopiformis</i>	KUJ06291	77/99
D	AUW31188	Hypothetical protein	<i>C. metacoralifera</i>	ADR00966	86/90
6	AUW31189	O-methyltransferase	<i>P. sporulosa</i>	XP_018038216	29/79
E	AUW31190	Integral membrane protein	<i>M. anisopliae</i>	KFG84816	42/83
7	AUW31191	PKS (<i>Cu-r-pks-3</i>)	<i>C. metacoralifera</i>	ADQ27444	85/100
F	AUW31192	2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase	<i>S. tuberosum</i>	XP_006343929	25/68
G	AUW31193	No significant similarity	-	-	-

Table S33: BLAST statistics for Cluster 29. Reproduced from Bertrand et al. (2018a), Supporting Information file, in accordance with authors' retainment of privileges.

Gene (No.)	Protein Accession No.	Protein Putative function(s)	Species Closest homolog	Homolog Accession No.	Identity (%) / Coverage (%)
A	AUW31194	Laccase-2	<i>F. oxysporum</i>	ENH72863	62/80
B	AUW31195	Hypothetical protein	<i>U. pustulata</i>	SLM38157	63/68
C	AUW31196	Hypothetical protein	<i>B. panamericana</i>	XP_007680224	45/81
D	AUW31197	Hypothetical protein	<i>X. heveae</i>	XP_018188437	40/68
1	AUW31198	Salicylate hydroxylase	<i>T. stipitatus</i>	XP_002483592	45/78
		Oxidoreductase	<i>P. subalpine</i>	CZR60184	43/84
		Monooxygenase	<i>M. majus</i>	XP_014576521	42/77
2	AUW31199	Oxidoreductase	<i>T. aphioglossoides</i>	KND93558	53/84
3	AUW31200	PKS (<i>Cu-nr-pks-9</i>)	<i>P. scopiformis</i>	KUJ11557	52/99
4	AUW31201	PKS (<i>Cu-r-pks-4</i>)	<i>P. camemberti</i>	CRL21974	45/99
E	AUW31202	C6 transcription factor	<i>T. cellulolyticus</i>	GAM39985	30/79
F	AUW31203	Glucose dehydrogenase	<i>B. cinerea</i>	EMR80543	52/96
G	AUW31204	No significant similarity	-	-	-
H	AUW31205	DUF636 domain protein	<i>R. emersonii</i>	XP_013327853	46/95
I	AUW31206	Protein with kelch motif	<i>M. brunnea</i>	XP_007287917	49/97
J	AUW31207	No significant similarity	-	-	-
K	AUW31208	MFS multidrug transporter	<i>A. ruber</i>	EYE92134	54/94

Table S34: BLAST statistics for Cluster 30. Reproduced from Bertrand et al. (2018a), Supporting Information file, in accordance with authors' retainment of privileges.

Gene (No.)	Protein Accession No.	Protein Putative function(s)	Species Closest homolog	Homolog Accession No.	Identity (%) / Coverage (%)
1	AUW31209	Hydrolase	<i>M. mycetomatis</i>	KOP42977	33/86
2	AUW31210	PKS (<i>Cu-nr-pks-10</i>)	<i>R. emersonii</i>	XP_013329845	55/99
3	AUW31211	Sterigmatocystin 8-O-methyltransferase	<i>N. udagawae</i>	GAO83070	53/89
A	AUW31212	MFS transporter	<i>P. solitum</i>	KJJ28016	65/93
B	AUW31213	Hypothetical protein	<i>B. oryzae</i>	XP_007689788	41/92
4	AUW31214	Cytochrome P-450	<i>P. solitum</i>	KJJ28025	66/96
C	AUW31215	Transcription activator	<i>A. parasiticus</i>	AAS66018	25/83
		AFLR	<i>A. parasiticus</i>	AAW32179	25/83
5	AUW31216	6-hydroxynicotinate 3-monooxygenase	<i>T. islandicus</i>	CRG89874	54/98
D	AUW31217	Dimeric alpha-beta barrel	<i>M. robertsii</i>	XP_007824558	55/92
E	AUW31218	Fungal transcription factor	<i>A. otae</i>	XP_002849816	36/96
6	AUW31219	Monooxygenase	<i>P. tritici-repentis</i>	XP_001933064	51/94
7	AUW31220	Short-chain dehydrogenase / reductase	<i>A. melanogenum</i>	KEQ57944	54/97
		Oxidoreductase	<i>T. ophioglossoides</i>	KND87425	53/98
8	AUW31221	NRPS-like enzyme (ACP-R domain-like protein)	<i>A. flavus</i>	KOC11045	33/99
F	AUW31222	No significant similarity	-	-	-
G	AUW31223	HET domain-bearing protein	<i>M. gypseum</i>	XP_003169660	25/71

Table S35: BLAST statistics for Cluster 31. Reproduced from Bertrand et al. (2018a), Supporting Information file, in accordance with authors' retainment of privileges.

Gene (No.)	Protein Accession No.	Protein Putative function(s)	Species Closest homolog	Homolog Accession No.	Identity (%) / Coverage (%)
1	AUW31224	PKS (<i>Cu-nr-pks-11</i>)	<i>G. lozoyensis</i>	XP_008084813	57/99
2	AUW31225	Cytochrome P-450	<i>P. camemberti</i>	CRL19965	62/99
A	AUW31226	Hypothetical protein	<i>A. rambellii</i>	KKK20309	36/97
3	AUW31227	GMC oxidoreductase	<i>P. camemberti</i>	CRL19964	56/97
4	AUW31228	Cytochrome P-450	<i>V. mali</i>	KUI73797	56/94
5	AUW31229	Short-chain dehydrogenase / reductase	<i>P. camemberti</i>	CRL26106	62/100
6	AUW31230	Homoserine lactamase	<i>V. mali</i>	KUI58326	57/95
		Pistatin demethylase	<i>V. mali</i>	KUI69904	58/95
7	AUW31231	O-methyltransferase	<i>A. parasiticus</i>	KJK66687	38/97
8	AUW31232	GMC oxidoreductase	<i>P. camemberti</i>	CRL19964	57/95
B	AUW31233	No significant similarity	-	-	-
9	AUW31234	Short-chain dehydrogenase / reductase	<i>P. minimum</i>	XP_007919583	58/100

Table S36: BLAST statistics for Cluster 32. Reproduced from Bertrand et al. (2018a), Supporting Information file, in accordance with authors' retainment of privileges.

Gene (No.)	Protein Accession No.	Protein Putative function(s)	Species Closest homolog	Homolog Accession No.	Identity (%) / Coverage (%)
A	AUW31235	Kinesin	<i>S. borealis</i>	ESZ93150	65/97
B	AUW31236	Tetrahydrofolylpolyglutamate synthase	<i>B. spectabilis</i>	GAD96053	62/98
		Tetrahydrofolate synthase	<i>R. emersonii</i>	XP_013329221	63/97
C	AUW31237	Ribosome biogenesis protein	<i>C. posadasii</i>	EFW17846	62/98
D	AUW31238	GPI transamidase component	<i>A. otae</i>	XP_002845857	64/100
1	AUW31239	GNAT acetyltransferase	<i>B. spectabilis</i>	GAD94037	49/88
2	AUW31240	PKS (<i>Cu-nr-pks-12</i>)	<i>C. grayi</i>	ADM79462	82/99
3	AUW31241	Cytochrome P-450	<i>T. islandicus</i>	CRG89931	45/68
E	AUW31242	No significant similarity	-	-	-

Table S37: BLAST statistics for Cluster 33. Reproduced from Bertrand et al. (2018a), Supporting Information file, in accordance with authors' retainment of privileges.

Gene (No.)	Protein Accession No.	Protein Putative function(s)	Species Closest homolog	Homolog Accession No.	Identity (%) / Coverage (%)
1	AUW31243	Noranthrone monooxygenase	<i>V. mali</i>	KUI68499	53/88
A	AUW31244	Hypothetical protein	<i>Aspergillus sp.</i>	ADM34150	70/94
2	AUW31245	Hydroxyacylglutathione hydroxylase	<i>M. gypseum</i>	XP_00316994	70/95
		Metallo-beta-lactamase	<i>T. marneffeii</i>	XP_002144864	68/95
3	AUW31246	PKS (<i>Cu-nr-pks-13</i>)	<i>P. scopiformis</i>	KUJ13605	75/99
4	AUW31247	O-methyltransferase	<i>T. cellulolyticus</i>	GAM33825	51/97
		Toxin biosynthesis regulatory protein	<i>A. nidulans</i>	CBF90109	52/97
B	AUW31248	C6 transcription factor	<i>R. emersonii</i>	XP_013324718	59/98
C	AUW31249	ABC transporter	<i>C. graminicola</i>	XP_008089355	65/95

Table S38: BLAST statistics for Cluster 34. Reproduced from Bertrand et al. (2018a), Supporting Information file, in accordance with authors' retainment of privileges.

Gene (No.)	Protein Accession No.	Protein Putative function(s)	Species Closest homolog	Homolog Accession No.	Identity (%) / Coverage (%)
A	AUW31250	No significant similarity	-	-	-
B	AUW31251	Hypothetical protein	<i>U. pustulata</i>	SLM38200	57/90
C	AUW31252	Hypothetical protein	<i>F. avenaceum</i>	KIL83654	41/83
D	AUW31253	Hypothetical protein	<i>O. maius</i>	KIM97379	42/100
1	AUW31254	O-methyltransferase	<i>P. lutea</i>	WP_037018229	51/88
2	AUW31255	Metallo-beta-lactamase	<i>T. stipitatus</i>	XP_002482927	56/97
		Hydrolase	<i>M. gypseum</i>	XP_003169334	52/100
3	AUW31256	PKS (<i>Cu-nr-pks-14</i>)	<i>P. scopiformis</i>	KUJ13605	56/99

Table S39: BLAST statistics for Cluster 35. Reproduced from Bertrand et al. (2018a), Supporting Information file, in accordance with authors' retainment of privileges.

Gene (No.)	Protein Accession No.	Protein Putative function(s)	Species Closest homolog	Homolog Accession No.	Identity (%) / Coverage (%)
A	AUW31257	Hypothetical protein	<i>B. gilchristii</i>	XP_002623822	36/95
B	AUW31258	Hypothetical protein	<i>C. apollinis</i>	XP_007782785	53/98
C	AUW31259	Phosphoglycerate kinase	<i>P. chlamydospore</i>	KKY19995	84/98
D	AUW31260	Hypothetical protein	<i>O. maius</i>	KIM96182	59/99
E	AUW31261	Hypothetical protein	<i>C. aquaticus</i>	ORX95829	39/64
F	AUW31262	SPRY domain-containing protein	<i>M. brunnea</i>	XP_007296958	30/77
1	AUW31263	Oxidoreductase	<i>A. niger</i>	GAQ44047	48/98
2	AUW31264	Thioesterase	<i>A. clavatus</i>	XP_001273696	36/97
G	AUW31265	DUF341 domain protein	<i>M. acridum</i>	XP_007807556	53/97
3	AUW31266	PKS (<i>Cu-r-pks-5</i>)	<i>A. clavatus</i>	XP_001268490	42/98
H	AUW31267	Hypothetical protein	<i>A. nomius</i>	KNG82959	48/96
4	AUW31268	Tetrahydroxynaphthalene reductase	<i>P. subalpine</i>	CZR65788	63/86
		Trihydroxynaphthalene reductase	<i>F. pedrosoi</i>	XP_013281129	55/86
I	AUW31269	No significant similarity	-	-	-
J	AUW31270	Hypothetical protein	<i>F. poae</i>	OBS29650	30/92
K	AUW31271	Cysteine proteinase	<i>C. ligniaria</i>	OIW25441	47/92
L	AUW31272	No significant similarity	-	-	-
M	AUW31273	Ankryin repeat-containing protein	<i>O. borbonicus</i>	KRT84714	38/95
N	AUW31274	Hypothetical protein	<i>P. croceum</i>	KIM77068	35/62
O	AUW31275	Cupin	<i>S. iranensis</i>	WP_078956578	79/81
P	AUW31276	Membrane transporter	<i>D. corticola</i>	XP_020134205	43/88
Q	AUW31277	No significant similarity	-	-	-
R	AUW31278	Hypothetical protein	<i>C. apollinis</i>	XP_007785410	31/89
S	AUW31279	No significant similarity	-	-	-
T	AUW31280	Mitochondrial carrier	<i>G. lozoyensis</i>	XP_008075965	56/83
U	AUW31281	No significant similarity	-	-	-
V	AUW31282	Heme oxygenase-like protein	<i>P. strigosozonata</i>	XP_007389506	34/72

Table S40: BLAST statistics for Cluster 36. Reproduced from Bertrand et al. (2018a), Supporting Information file, in accordance with authors' retainment of privileges.

Gene (No.)	Protein Accession No.	Protein Putative function(s)	Species Closest homolog	Homolog Accession No.	Identity (%) / Coverage (%)
1	AUW31283	2-hydroxyacid dehydrogenase	<i>Bradyrhizobium sp.</i>	WP_024513000	70/97
2	AUW31284	Thiamine pyrophosphate Pyruvate decarboxylase	<i>C. immunda</i> <i>E. dermatitidis</i>	OQV10385 XP_009159300	72/98 70/98
A	AUW31285	Hypothetical protein	<i>A. richmondensis</i>	KXL45205	37/96
3	AUW31286	Serine hydrolase	<i>V. mali</i>	KUI67061	31/80
4	AUW31287	PKS (Cu-r-pks-6)	<i>P. verrucosus</i>	XP_018129701	36/98
B	AUW31288	D-alanine -D-alanine ligase	<i>A. niger</i>	GAQ42953	40/96
C	AUW31289	No significant similarity	-	-	-
D	AUW31290	Hypothetical protein	<i>U. pustulata</i>	SLM35821	64/84
E	AUW31291	Phosphatidylethanolamine- binding protein	<i>U. pustulata</i>	SLM35322	34/75
5	AUW31292	Scytalone dehydratase	<i>M. brunnea</i>	XP_007296032	73/63
F	AUW31293	Hypothetical protein	<i>P. zonata</i>	XP_022580981	35/83
G	AUW31294	No significant similarity	-	-	-
H	AUW31295	High affinity glucose transporter RGT2	<i>U. pustulata</i>	SLM41166	87/87
I	AUW31296	No significant similarity	-	-	-
J	AUW31297	RNase adaptor RapZ	<i>C. luteum</i>	WP_092107835	31/69
K	AUW31298	No significant similarity	-	-	-
L	AUW31299	No significant similarity	-	-	-
M	AUW31300	Hypothetical protein	<i>E. aquamarina</i>	XP_013257405	35/68
N	AUW31301	Lipase	<i>D. seriata</i>	OMP84603	51/81
O	AUW31302	No significant similarity	-	-	-
P	AUW31303	60S ribosomal protein	<i>C. apollinis</i>	XP_007777610	84/58
Q	AUW31304	rRNA methyltransferase	<i>E. necator</i>	KHJ30510	73/63
R	AUW31305	No significant similarity	-	-	-

Table S41: BLAST statistics for Cluster 37. Reproduced from Bertrand et al. (2018a), Supporting Information file, in accordance with authors' retainment of privileges.

Gene (No.)	Protein Accession No.	Protein Putative function(s)	Species Closest homolog	Homolog Accession No.	Identity (%) / Coverage (%)
A	AUW31306	RAS small monomeric GTPase	<i>T. marneffeii</i>	ACH42500	57/63
B	AUW31307	Cytochrome P-450	<i>U. pustulata</i>	SLM26344	42/85
C	AUW31308	Amidase signature enzyme	<i>P. scopiformis</i>	XP_018074719	73/95
D	AUW31309	Hypothetical protein	<i>C. geophilum</i>	OCK97656	56/91
E	AUW31310	Hypothetical protein	<i>F. poae</i>	OBS27948	33/86
F	AUW31311	Alpha/beta-hydrolase	<i>A. alternata</i>	XP_018385998	58/72
G	AUW31312	YjeF N-terminal domain-like protein	<i>P. scopiformis</i>	XP_018064223	43/80
H	AUW31313	C6 transcription factor	<i>U. pustulata</i>	SLM34042	27/75
I	AUW31314	Amino acid transporter	<i>U. pustulata</i>	SLM34781	68/94
J	AUW31315	Hypothetical protein	<i>X. heveae</i>	XP_018190186	36/58
K	AUW31316	No significant similarity	-	-	-
L	AUW31317	No significant similarity	-	-	-
M	AUW31318	MFS general substrate transporter	<i>M. brunneum</i>	XP_014548836	32/64
N	AUW31319	DNA helicase, C-terminal	<i>P. solitum</i>	KJJ24275	34/87
O	AUW31320	No significant similarity	-	-	-
P	AUW31321	No significant similarity	-	-	-
Q	AUW31322	Hypothetical protein	<i>O. maius</i>	KIM97070	35/79
1	AUW31323	O-methyltransferase	<i>M. brunnea</i>	XP_007289784	57/95
2	AUW31324	PKS (<i>Cu-r-pks-7</i>)	<i>C. gloeosporioides</i>	XP_007278751	58/99
3	AUW31325	α/β -hydrolase	<i>G. lozoyensis</i>	XP_008078704	33/97
R	AUW31326	Hypothetical protein	<i>E. pusillum</i>	XP_007805358	35/96
S	AUW31327	Hypothetical protein	<i>C. geophilum</i>	OCK92546	73/91

Table S42: BLAST statistics for Cluster 38. Reproduced from Bertrand et al. (2018a), Supporting Information file, in accordance with authors' retainment of privileges.

Gene (No.)	Protein Accession No.	Protein Putative function(s)	Species Closest homolog	Homolog Accession No.	Identity (%) / Coverage (%)
A	AUW31328	Hypothetical protein	<i>F. monophora</i>	XP_022509238	32/68
B	AUW31329	Choline transport protein	<i>A. lentulus</i>	GAQ11825	40/79
C	AUW31330	No significant similarity	-	-	-
D	AUW31331	No significant similarity	-	-	-
E	AUW31332	Hypothetical protein	<i>X. heveae</i>	XP_018187872	55/90
F	AUW31333	VPS9	<i>G. lozoyensis</i>	XP_008087854	56/98
G	AUW31334	Hypothetical protein	<i>F. pseudograminearum</i>	XP_009259243	30/97
H	AUW31335	Zinc finger transcriptional regulator	<i>T. virens</i>	ABV48713	25/75
1	AUW31336	Short-chain dehydrogenase / reductase	<i>N. udagawae</i>	GAO90497	50/95
I	AUW31337	Hypothetical protein	<i>G. lozoyensis</i>	XP_008077270	63/99
J	AUW31338	MFS transporter	<i>T. marneffeii</i>	KFX45277	53/98
K	AUW31339	Pyridoxal phosphate- dependent transferase	<i>M. quizhouense</i>	KID84201	58/96
2	AUW31340	PKS (<i>Cu-r-pks-8</i>)	<i>F. bulbicola</i>	ALQ32802	54/96
L	AUW31341	Ribosomal protein	<i>A. niger</i>	GAQ37462	33/79
M	AUW31342	MFS transporter	<i>E. lata</i>	XP_007791731	59/90
N	AUW31343	GTP-binding protein	<i>C. apollinis</i>	XP_007776753	88/100
O	AUW31344	Hag1 (peptidase)	<i>M. acridum</i>	XP_007814652	56/88
P	AUW31345	Mitochondrial ribosomal subunit S18	<i>L. maculans</i>	XP_003836080	46/94
Q	AUW31346	Mitochondrial inner membrane translocase	<i>A. nidulans</i>	CBF87801	73/84

Table S43: BLAST statistics for Cluster 39. Reproduced from Bertrand et al. (2018a), Supporting Information file, in accordance with authors' retainment of privileges.

Gene (No.)	Protein Accession No.	Protein Putative function(s)	Species Closest homolog	Homolog Accession No.	Identity (%) / Coverage (%)
A	AUW31347	MFS transporter	<i>A. nomius</i>	KNG81528	71/98
1	AUW31348	Dioxygenase	<i>S. borealis</i>	ESZ99522	79/93
B	AUW31349	ABC transporter	<i>A. clavatus</i>	XP_001273095	59/99
2	AUW31350	Isoepoxydon dehydrogenase	<i>B. nivea</i>	AAT00463	76/100
3	AUW31351	M-cresol methyl hydroxylase (P-450)	<i>A. clavatus</i>	XP_001273090	70/99
4	AUW31352	Isoamyl alcohol oxidase	<i>P. expansum</i>	AIG62142	64/98
5	AUW31353	6-methylsalicylic acid decarboxylase	<i>P. expansum</i>	KGO43532	64/97
C	AUW31354	Transcription factor	<i>A. otae</i>	XP_002842634	51/96
6	AUW31355	PKS (Cu-r-pks-9)	<i>B. nivea</i>	AAK48943	60/97
D	AUW31356	Cupin RmIC-type protein	<i>M. phaseolina</i>	EKG18984	69/95
7	AUW31357	m-Hydroxybenzyl alcohol hydroxylase (P-450)	<i>A. clavatus</i>	XP_001273091	67/98
8	AUW31358	Oxidoreductase	<i>B. cinerea</i>	CCD45693	56/99
9	AUW31359	Carboxylesterase	<i>A. clavatus</i>	XP_001273084	53/86
10	AUW31360	Short-chain dehydrogenase / reductase	<i>M. phaseolina</i>	EKG18977	71/100
E	AUW31361	MFS transporter	<i>M. phaseolina</i>	EKG18978	68/88
F	AUW31362	Hypothetical protein	<i>M. phaseolina</i>	EKG18979	71/68
11	AUW31363	Ketoreductase	<i>C. immitis</i>	XP_001248190	67/91
12	AUW31364	GMC oxidoreductase	<i>B. cinerea</i>	CCD44943	69/99

Table S44: BLAST statistics for Cluster 40. Reproduced from Bertrand et al. (2018a), Supporting Information file, in accordance with authors' retainment of privileges.

Gene (No.)	Protein Accession No.	Protein Putative function(s)	Species Closest homolog	Homolog Accession No.	Identity (%) / Coverage (%)
A	AUW31365	Exonuclease	<i>H. capsulatum</i>	EGC40676	49/99
B	AUW31366	Hypothetical protein	<i>E. pusillum</i>	XP_007786021	39/83
1	AUW31367	Serine hydrolase	<i>N. udagawae</i>	GAO87671	30/100
2	AUW31368	PKS (Cu-r-pks-10)	<i>F. euwallaceae</i>	ALQ32844	34/99
C	AUW31369	Glycoside hydrolase	<i>D. ampelina</i>	KKY37913	48/99
D	AUW31370	No significant similarity	-	-	-
E	AUW31371	Ankyrin	<i>P. scopiformis</i>	KUJ21848	53/89
F	AUW31372	Isocitrate dehydrogenase	<i>A. otae</i>	XP_002848642	51/88
G	AUW31373	Chloramphenicol acetyltransferase-like domain protein	<i>A. aleurodis</i>	KZZ97445	35/90
H	AUW31374	No significant similarity	-	-	-

Table S45: BLAST statistics for Cluster 41. Reproduced from Bertrand et al. (2018a), Supporting Information file, in accordance with authors' retainment of privileges.

Gene (No.)	Protein Accession No.	Protein Putative function(s)	Species Closest homolog	Homolog Accession No.	Identity (%) / Coverage (%)
A	AUW31375	No significant similarity	-	-	-
1	AUW31376	PKS (<i>Cu-r-pks-11</i>)	<i>V. mali</i>	KUI64629	59/98
B	AUW31377	Hypothetical protein	<i>V. mali</i>	KUI64496	71/96
2	AUW31378	6-hydroxy-D-nicotine oxidase	<i>V. mali</i>	KUI65223	54/93
3	AUW31379	Trichodiene oxygenase (P-450)	<i>V. mali</i>	KUI64499	66/93
4	AUW31380	FAD oxidase	<i>P. minimum</i>	XP_007916293	57/98
C	AUW31381	Integral membrane protein	<i>A. flavus</i>	KOC16459	28/89
5	AUW31382	Alcohol dehydrogenase	<i>P. subalpine</i>	CZR68690	58/99
6	AUW31383	Trichodiene oxygenase (P-450)	<i>V. mali</i>	KUI64500	48/95

Table S46: BLAST statistics for Cluster 42. Reproduced from Bertrand et al. (2018a), Supporting Information file, in accordance with authors' retainment of privileges.

Gene (No.)	Protein Accession No.	Protein Putative function(s)	Species Closest homolog	Homolog Accession No.	Identity (%) / Coverage (%)
A	AUW31384	Tetracycline resistance protein	<i>P. minimum</i>	XP_007913391	47/100
B	AUW31385	Glycosidase	<i>P. scopiformis</i>	KUJ20814	54/94
C	AUW31386	Integral membrane protein	<i>R. necatrix</i>	GAP89890	42/83
1	AUW31387	Short-chain dehydrogenase / reductase	<i>M. phaseolina</i>	EKG11161	49/97
D	AUW31388	Phosphoribosylamine-glycine ligase	<i>C. coronate</i>	XP_007723050	54/99
E	AUW31389	Hypothetical protein	<i>V. longisporum</i>	CRK12658	37/82
2	AUW31390	PKS (<i>Cu-r-pks-12</i>)	<i>R. necatrix</i>	GAP92219	39/98

Table S47: BLAST statistics for Cluster 43. Reproduced from Bertrand et al. (2018a), Supporting Information file, in accordance with authors' retainment of privileges.

Gene (No.)	Protein Accession No.	Protein Putative function(s)	Species Closest homolog	Homolog Accession No.	Identity (%) / Coverage (%)
A	AUW31391	No significant similarity	-	-	-
B	AUW31392	Integral membrane protein	<i>A. fumigatus</i>	XP_00754114	29/92
1	AUW31393	Carboxymuconolactone decarboxylase	<i>S. xylophagus</i>	WP_043662492	33/95
2	AUW31394	Dehydrogenase	<i>A. namibiae</i>	XP_013429252	56/96
C	AUW31395	Hypothetical protein	<i>Pseudogymnoascus sp.</i>	KFY24242	36/97
D	AUW31396	MFS transporter	<i>C. graminicola</i>	XP_008099014	49/88
3	AUW31397	PKS (<i>Cu-r-pks-13</i>)	<i>G. lozoyensis</i>	XP_008082817	50/97

Table S48: BLAST statistics for Cluster 44. Reproduced from Bertrand et al. (2018a), Supporting Information file, in accordance with authors' retainment of privileges.

Gene (No.)	Protein Accession No.	Protein Putative function(s)	Species Closest homolog	Homolog Accession No.	Identity (%) / Coverage (%)
1	AUW31398	Aldehyde dehydrogenase	<i>M. notabilis</i>	XP_010112250	35/72
A	AUW31399	No significant similarity	-	-	-
B	AUW31400	Ribosome biogenesis protein	<i>A. niger</i>	GAQ33854	76/99
C	AUW31401	Hypothetical protein	<i>C. apollinis</i>	XP_007780362	65/92
D	AUW31402	Hypothetical protein	<i>G. trabeum</i>	XP_007865805	48/88
2	AUW31403	Carboxymuconolactone decarboxylase	<i>A. rifamycini</i>	WP_051301640	33/85
E	AUW31404	Integral membrane protein	<i>R. necatrix</i>	GAP88685	39/81
3	AUW31405	PKS (<i>Cu-r-pks-14</i>)	<i>C. gloeosporioides</i>	XP_007278670	46/95
F	AUW31406	Hypothetical protein	<i>B. panamericana</i>	XP_007681552	28/69

Table S49: BLAST statistics for Cluster 45. Reproduced from Bertrand et al. (2018a), Supporting Information file, in accordance with authors' retainment of privileges.

Gene (No.)	Protein Accession No.	Protein Putative function(s)	Species Closest homolog	Homolog Accession No.	Identity (%) / Coverage (%)
1	AUW31407	Short-chain dehydrogenase / reductase	<i>C. gloeosporioides</i>	EQB51186	73/99
2	AUW31408	FAD oxidase	<i>S. lycopersici</i>	KNG46186	55/99
3	AUW31409	Cytochrome P-450	<i>C. gloeosporioides</i>	EQB51189	61/99
4	AUW31410	Zinc-binding dehydrogenase	<i>C. gloeosporioides</i>	XP_007283077	73/100
5	AUW31411	PKS (<i>Cu-r-pks-15</i>)	<i>C. gloeosporioides</i>	XP_007283082	67/99
A	AUW31412	Hypothetical protein	<i>G. stellatum</i>	OCL13418	34/90

Table S50: BLAST statistics for Cluster 46. Reproduced from Bertrand et al. (2018a), Supporting Information file, in accordance with authors' retainment of privileges.

Gene (No.)	Protein Accession No.	Protein Putative function(s)	Species Closest homolog	Homolog Accession No.	Identity (%) / Coverage (%)
A	AUW31413	No significant similarity	-	-	-
B	AUW31414	Cation channel	<i>B. cinerea</i>	EMR82051	57/92
C	AUW31415	Hypothetical protein	<i>F. graminearum</i>	XP_011319371	24/75
1	AUW31416	Oxidoreductase	<i>A. nomius</i>	KNG88689	59/97
2	AUW31417	Serine hydrolase	<i>C. aquaticus</i>	ORY17863	46/91
3	AUW31418	PKS (<i>Cu-r-pks-16</i>)	<i>A. nomius</i>	KNG88668	45/94

Table S51: BLAST statistics for Cluster 47. Reproduced from Bertrand et al. (2018a), Supporting Information file, in accordance with authors' retainment of privileges.

Gene (No.)	Protein Accession No.	Protein Putative function(s)	Species Closest homolog	Homolog Accession No.	Identity (%) / Coverage (%)
A	AUW31419	MFS transporter	<i>S. borealis</i>	ESZ91288	45/89
B	AUW31420	No significant similarity	-	-	-
1	AUW31421	PKS (<i>Cu-r-pks-17</i>)	<i>P. membranacea</i>	AEE65375	41/99
C	AUW31422	No significant similarity	-	-	-

Table S52: BLAST statistics for Cluster 48. Reproduced from Bertrand et al. (2018a), Supporting Information file, in accordance with authors' retainment of privileges.

Gene (No.)	Protein Accession No.	Protein Putative function(s)	Species Closest homolog	Homolog Accession No.	Identity (%) / Coverage (%)
1	AUW31423	PKS (<i>Cu-r-pks-18</i>)	<i>T. islandicus</i>	CRG92723	75/100
2	AUW31424	O-methyltransferase	<i>M. mycetomatis</i>	KOP50440	85/99
3	AUW31425	Short-chain dehydrogenase / reductase	<i>G. clavigera</i>	XP_014172074	68/99
4	AUW31426	Cytochrome P-450	<i>T. islandicus</i>	CRG92717	79/97
5	AUW31427	FAD oxidase	<i>M. mycetomatis</i>	KOP50437	71/98

Table S53: Summary of contigs containing *Cladonia uncialis* biosynthetic genes. Reproduced from Bertrand et al. (2018a), Supporting Information file, in accordance with authors' retainment of privileges.

Cluster No.	Cluster Accession No.	Putative encoded biosynthetic enzymes
1	MG777469	CU-T3-PKS-1 CU-R-PKS-1
2	MG777470	CU-T3-PKS-2
3	MG777471	CU-NRPS-1
4	MG777472	CU-NRPS-2
5	MG777473	CU-NRPS-3
6	MG777474	CU-NRPS-4
7	MG777475	CU-NRPS-5
8	MG777476	CU-NRPS-6
9	MG777477	CU-NRPS-7
10	MG777478	CU-NRPS-8
11	MG777479	CU-NRPS-9
12	MG777480	CU-TERP-1 CU-TERP-2 CU-PKS-NRPS-1
13	MG777481	CU-PKS-NRPS-2 CU-TERP-3
14	MG777482	CU-PKS-NRPS-3
15	MG777483	CU-TERP-4
16	MG777484	CU-TERP-5
17	MG777485	CU-TERP-6
18	MG777486	CU-TERP-7
19	MG777487	CU-TERP-8
20	MG777488	CU-TERP-9 CU-R-PKS-2
21	MG777489	CU-NR-PKS-1
22	MG777490	CU-NR-PKS-2
23	MG777491	CU-NR-PKS-3
24	MG777492	CU-NR-PKS-4
25	MG777493	CU-NR-PKS-5
26	MG777494	CU-NR-PKS-6
27	MG777495	CU-NR-PKS-7
28	MG777496	CU-NR-PKS-8 CU-R-PKS-3
29	MG777497	CU-NR-PKS-9 CU-R-PKS-4
30	MG777498	CU-NR-PKS-10
31	MG777499	CU-NR-PKS-11
32	MG777500	CU-NR-PKS-12
33	MG777501	CU-NR-PKS-13
34	MG777502	CU-NR-PKS-14
35	MG777503	CU-R-PKS-5
36	MG777504	CU-R-PKS-6
37	MG777505	CU-R-PKS-7
38	MG777506	CU-R-PKS-8
39	MG777507	CU-R-PKS-9
40	MG777508	CU-R-PKS-10
41	MG777509	CU-R-PKS-11
42	MG777510	CU-R-PKS-12
43	MG777511	CU-R-PKS-13
44	MG777512	CU-R-PKS-14
45	MG777513	CU-R-PKS-15
46	MG777514	CU-R-PKS-16
47	MG777515	CU-R-PKS-17
48	MG777516	CU-R-PKS-18

Chapter 6

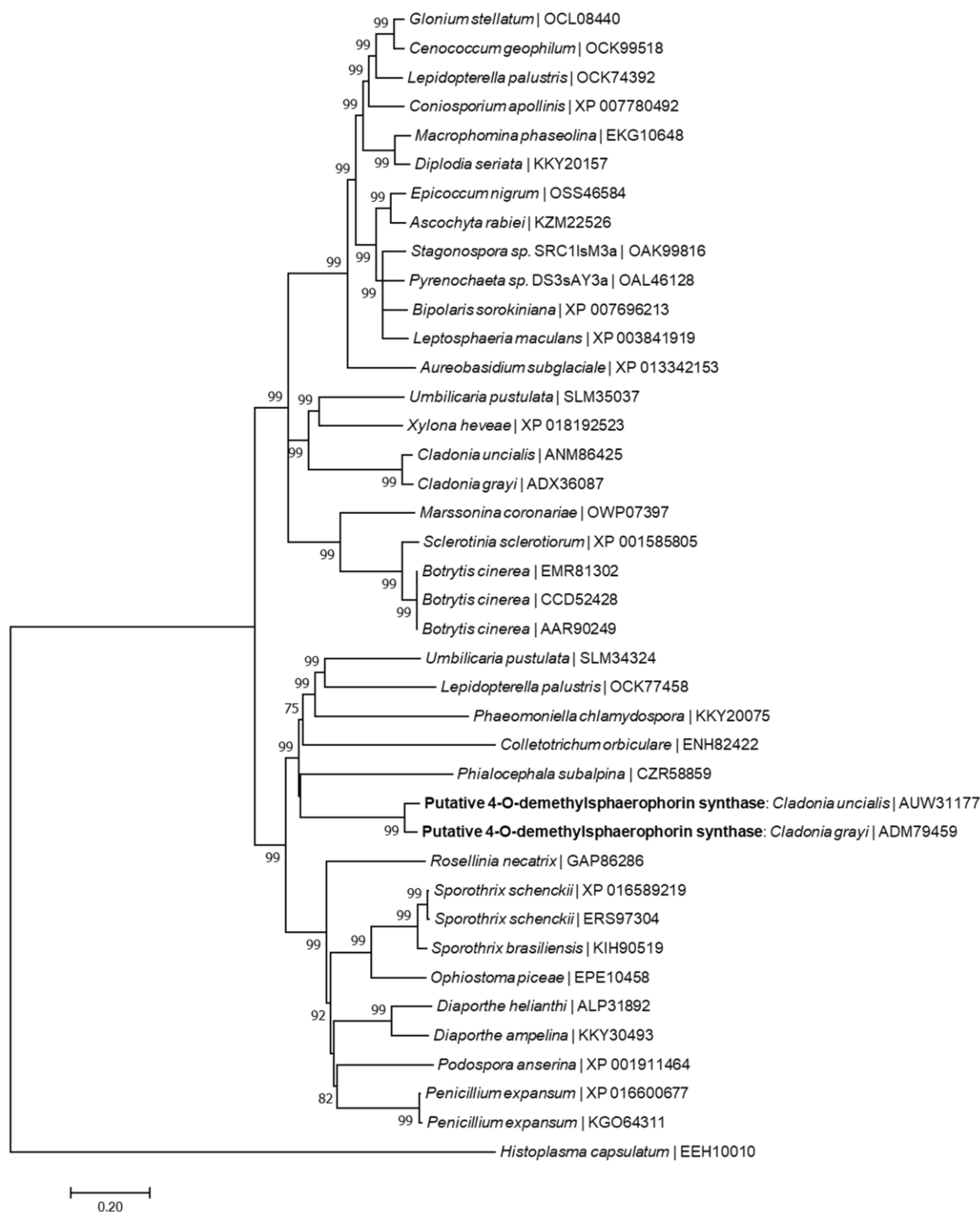


Figure S6: Phylogenetic relationship between the putative **4-O-demethyl sphaerophorin synthase** of *Cladonia uncialis* and a genetically similar gene encoding a putative 4-O-demethyl sphaerophorin synthase that is proposed to be a part of the **grayanic acid** biosynthetic gene cluster of *Cladonia grayi*. The PKS of *Histoplasma capsulatum* (EEH10010) was chosen as an out-group. Reproduced from Bertrand et al. (2018b), Supporting Information file, in accordance with authors' retainment of privileges.

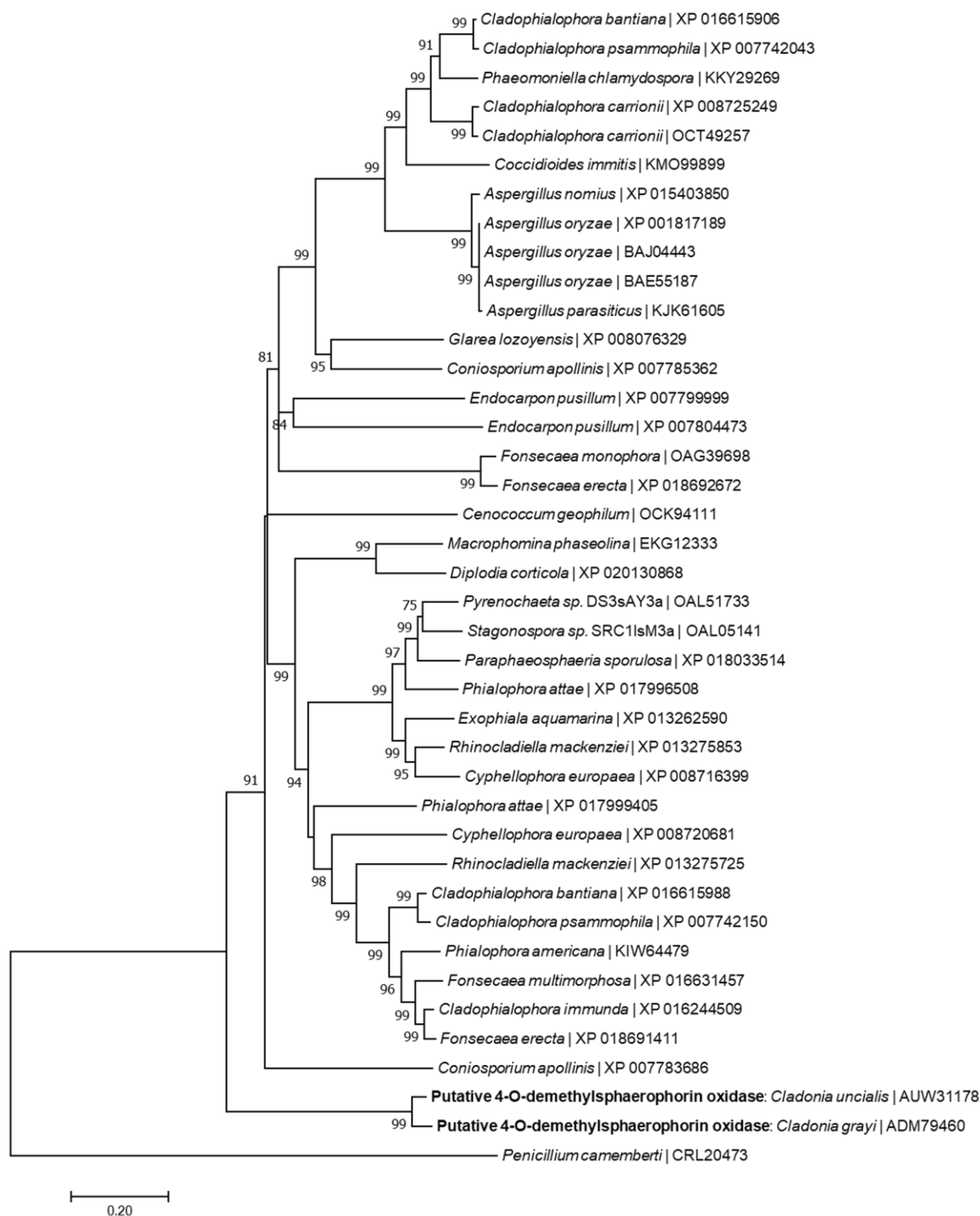


Figure S7: Phylogenetic relationship between the putative 4-O-demethyl sphaerophorin oxidase of *Cladonia uncialis* and a genetically similar gene encoding a putative 4-O-demethyl sphaerophorin oxidase that is proposed to be a part of the **grayanic acid** biosynthetic gene cluster of *Cladonia grayi*. The cytochrome p450 of *Penicillium camemberti* (CRL20473) was chosen as an out-group. Reproduced from Bertrand et al. (2018b), Supporting Information file, in accordance with authors' retainment of privileges.

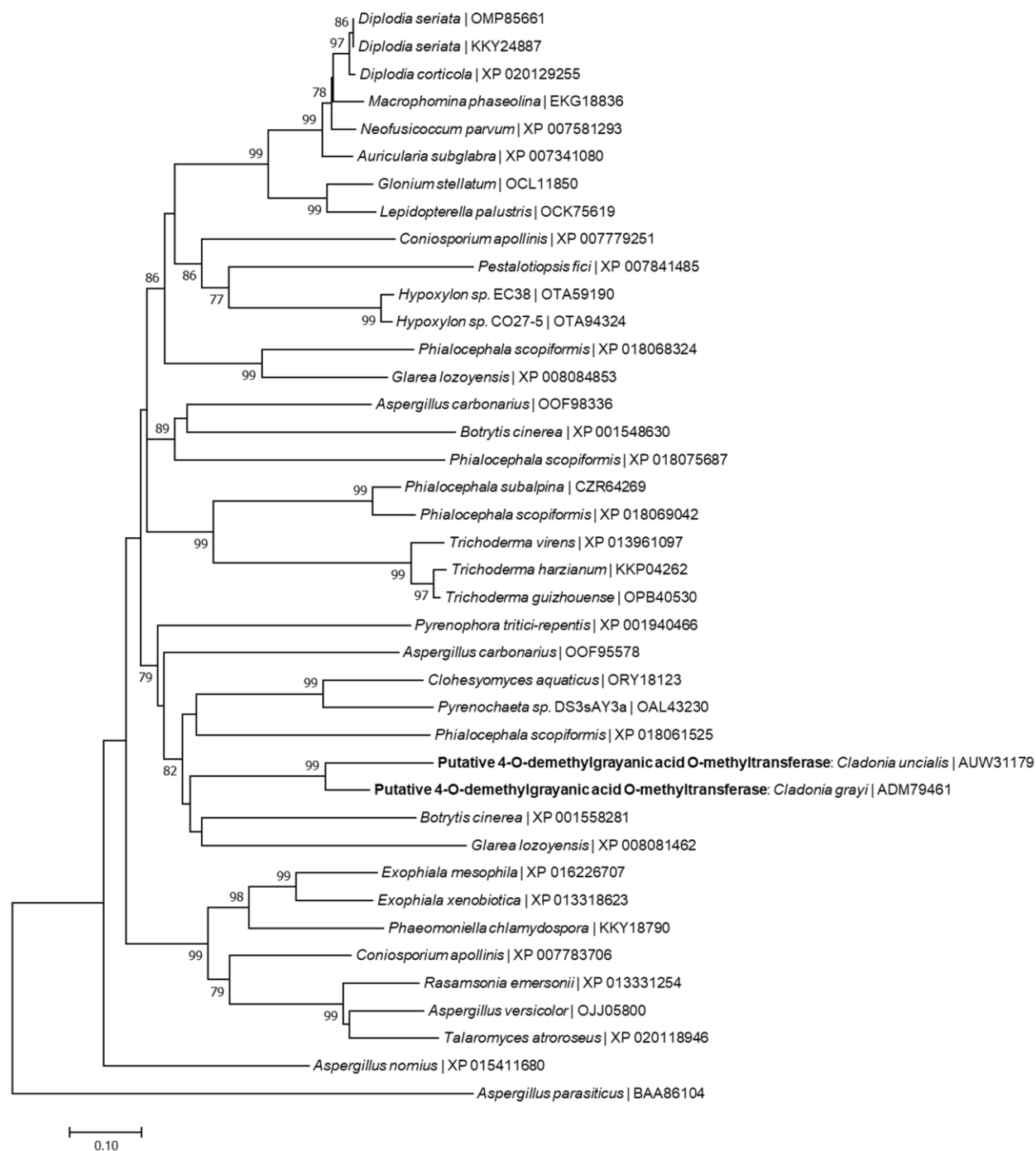


Figure S8: Phylogenetic relationship between the putative 4-*O*-demethyl-grayanic acid *O*-methyltransferase of *Cladonia uncialis* and a genetically similar gene encoding a putative 4-*O*-demethyl-grayanic acid *O*-methyltransferase that is proposed to be a part of the **grayanic acid** biosynthetic gene cluster of *Cladonia grayi*. The *O*-methyltransferase of *Aspergillus parasiticus* (BAA86104) was chosen as an out-group. Reproduced from Bertrand et al. (2018b), Supporting Information file, in accordance with authors' retainment of privileges.

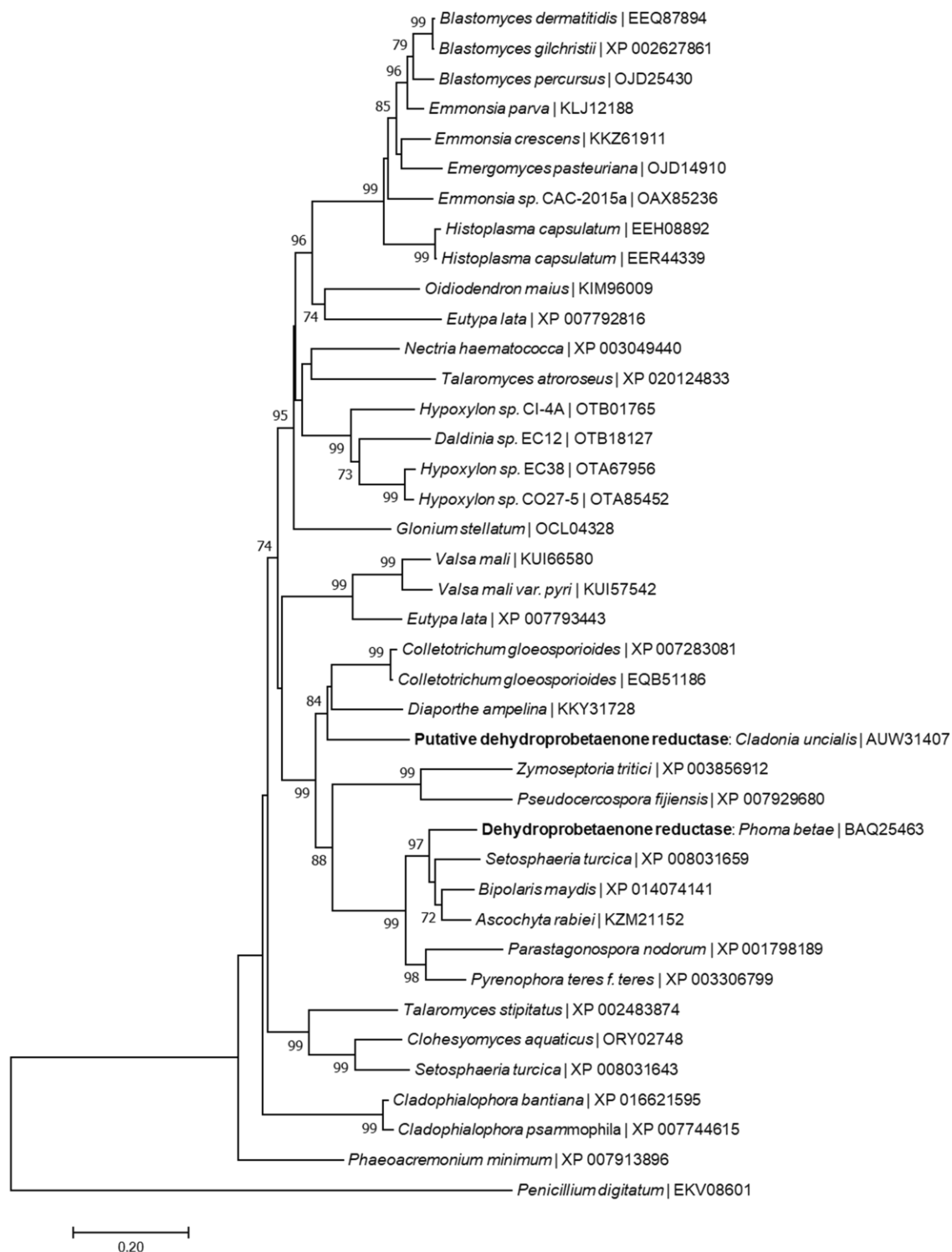


Figure S9: Phylogenetic relationship between a putative **dehydroprobetaenone reductase** of *Cladonia uncialis* and a genetically similar gene encoding a dehydroprobetaenone reductase that is part of the **betaenone** biosynthetic gene cluster of *Phoma betae*. The short-chain dehydrogenase/reductase of *Penicillium digitatum* (EKV08601) was chosen as an out-group. Reproduced from Bertrand et al. (2018b), Supporting Information file, in accordance with authors' retainment of privileges.

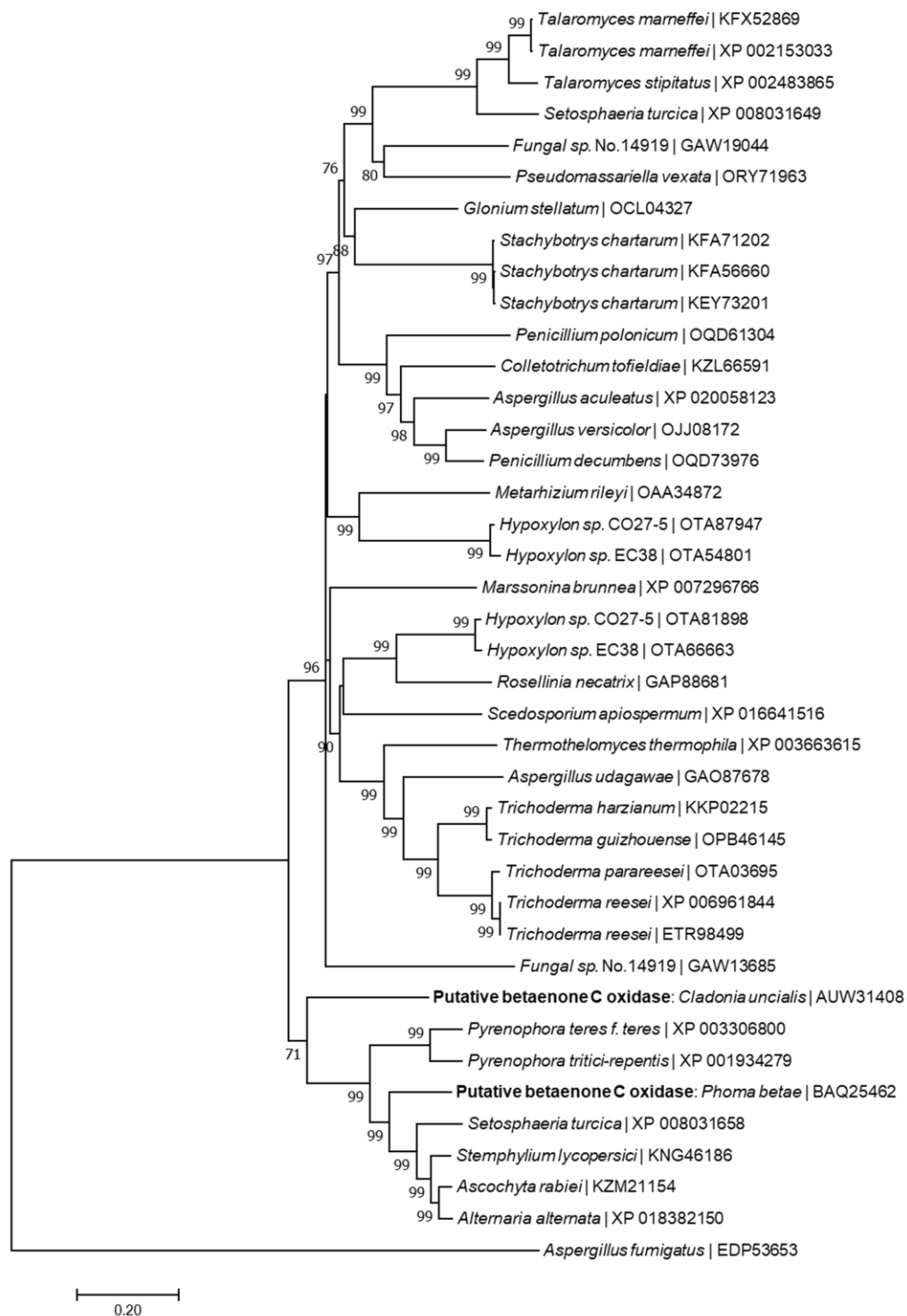


Figure S10: Phylogenetic relationship between a putative **betaenone C oxidase** of *Cladonia uncialis* and a genetically similar gene that is proposed to encode a betaenone C oxidase that participates in **betaenone** biosynthesis in *Phoma betae*. The FAD oxidase of *Aspergillus fumigatus* (EDP53653) was chosen as an out-group. Reproduced from Bertrand et al. (2018b), Supporting Information file, in accordance with authors' retainment of privileges.

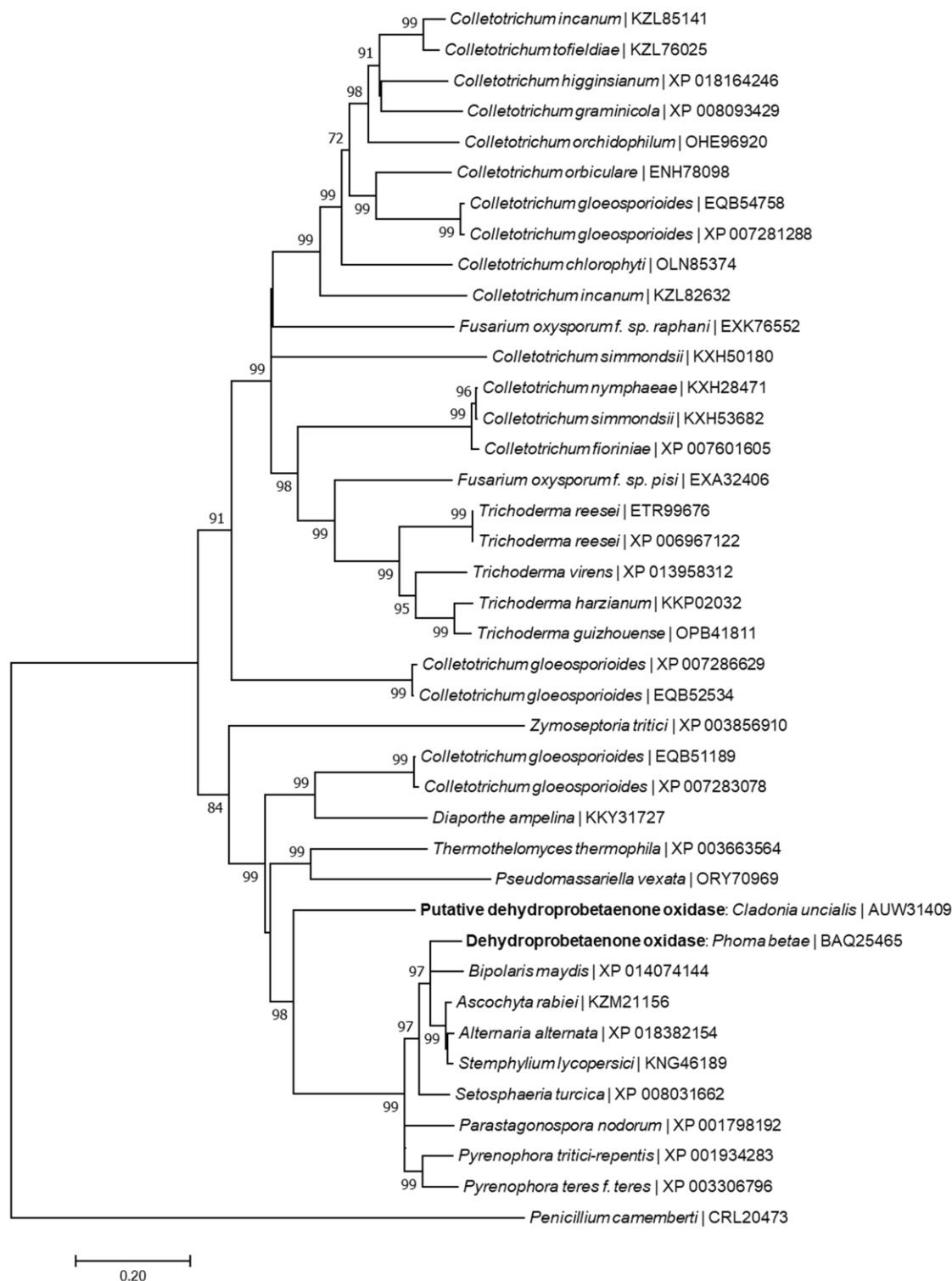


Figure S11: Phylogenetic relationship between a putative **dehydroprobetaenone oxidase** of *Cladonia uncialis* and a genetically similar gene encoding a dehydroprobetaenone oxidase that is part of the **betaenone** biosynthetic gene cluster of *Phoma betae*. The cytochrome p450 of *Penicillium camemberti* (CRL20473) was chosen as an out-group. Reproduced from Bertrand et al. (2018b), Supporting Information file, in accordance with authors' retainment of privileges.

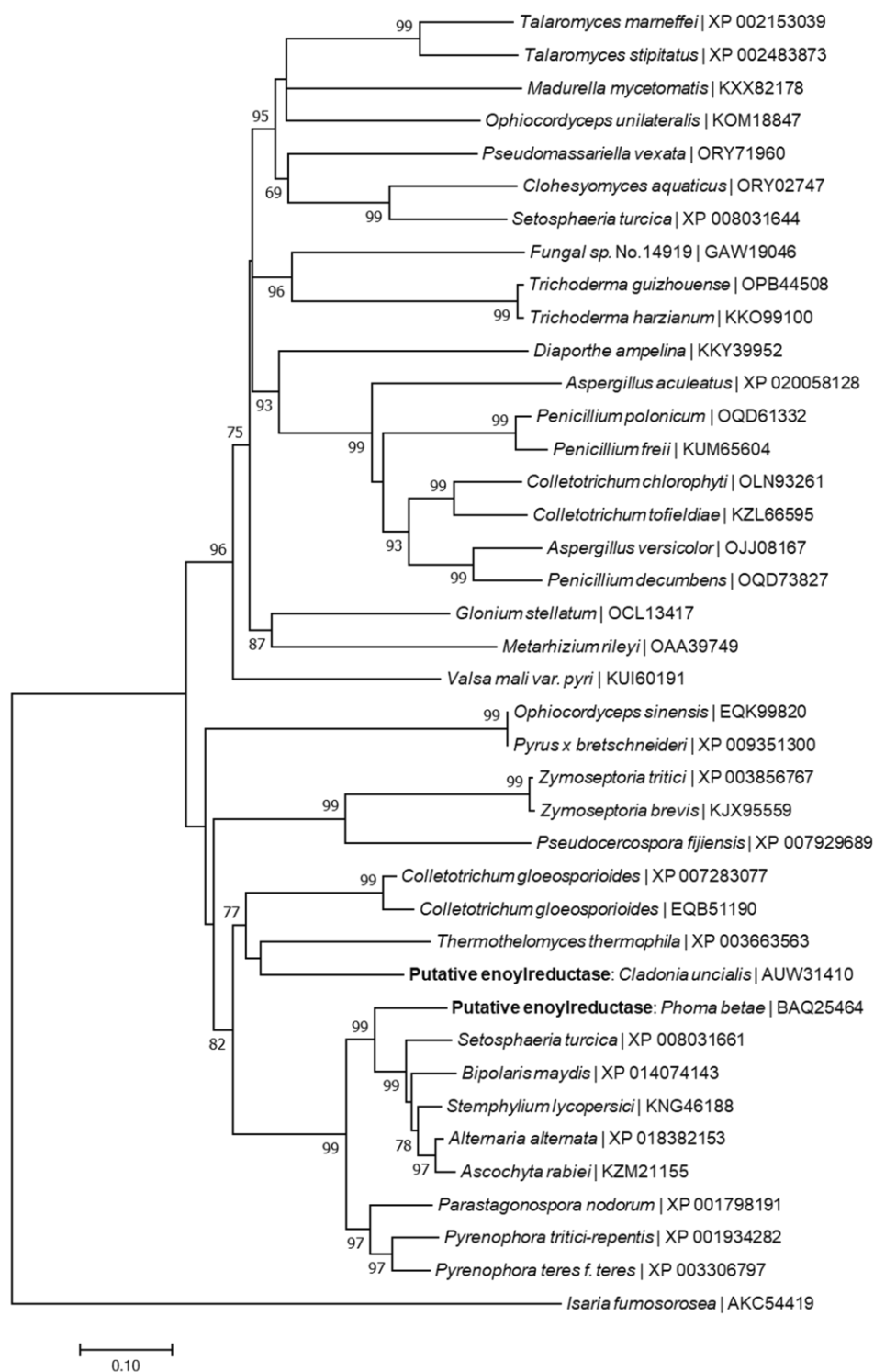


Figure S12: Phylogenetic relationship between a putative *trans*-enoylreductase of *Cladonia uncialis* and a genetically similar gene encoding a *trans*-enoylreductase that is part of the **betaenone** biosynthetic gene cluster of *Phoma betae*. The *trans*-enoylreductase of *Isaria fumosorosea* (AKC54419) was chosen as an out-group. Reproduced from Bertrand et al. (2018b), Supporting Information file, in accordance with authors' retainment of privileges.

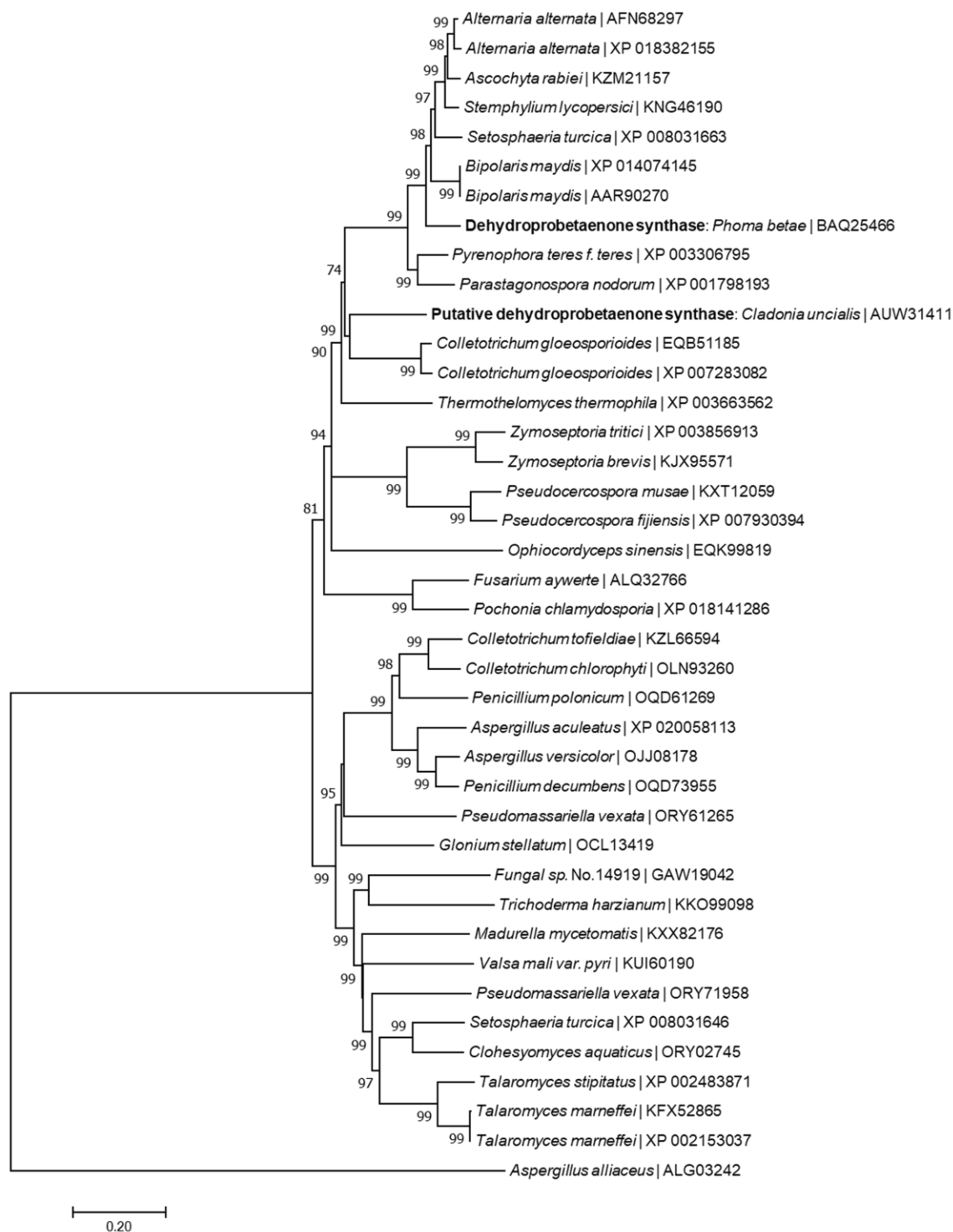


Figure S13: Phylogenetic relationship between a putative **dehydroprobetaenone synthase** of *Cladonia uncialis* and a genetically similar gene encoding a dehydroprobetaenone synthase that is part of the **betaenone** biosynthetic gene cluster of *Phoma betae*. The PKS of *Aspergillus alliaceus* (ALG03242) was chosen as an out-group. Reproduced from Bertrand et al. (2018b), Supporting Information file, in accordance with authors' retainment of privileges.

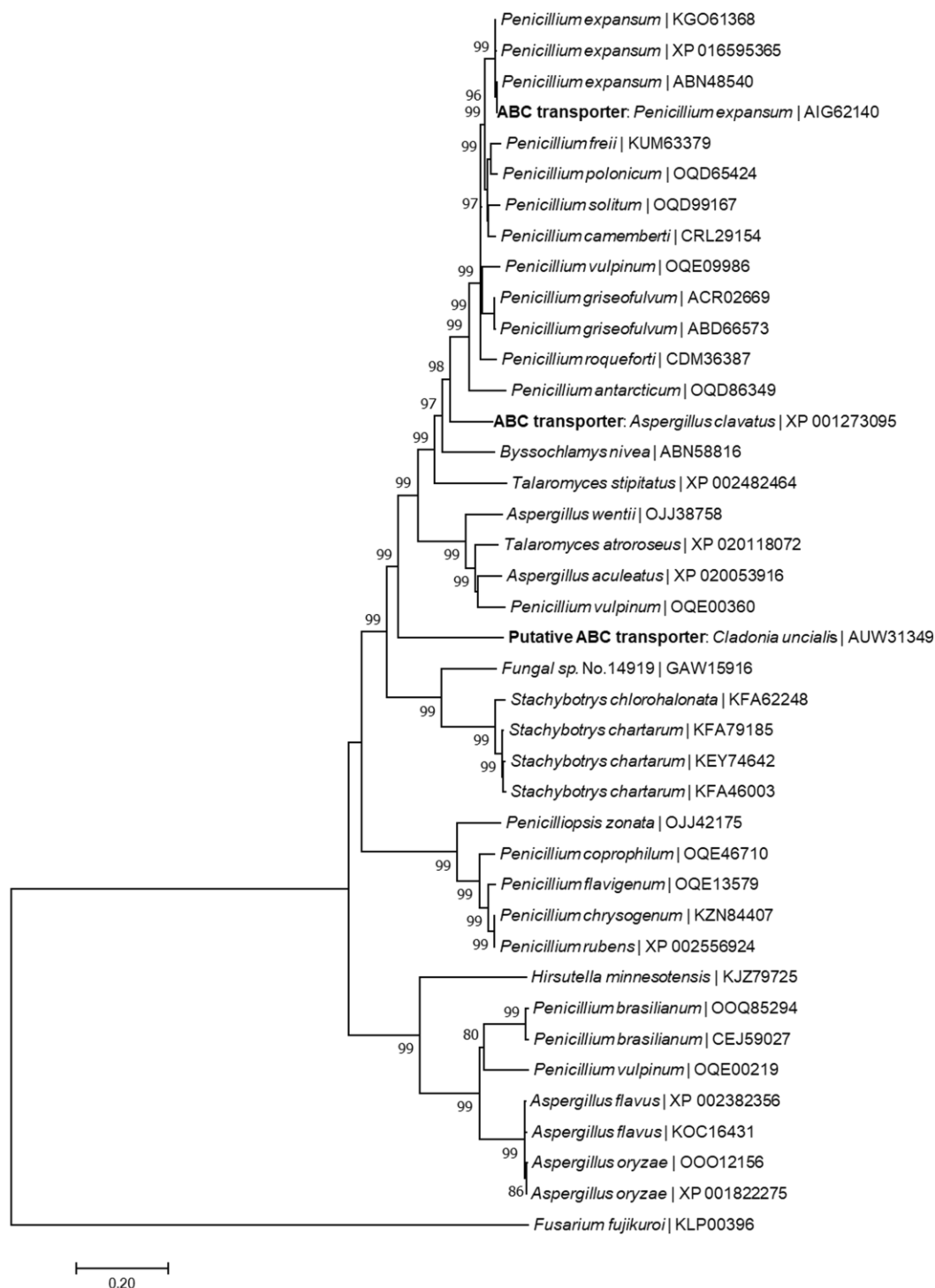


Figure S14: Phylogenetic relationship between a putative **ABC transporter** of *Cladonia uncialis* and a genetically similar gene encoding an ABC transporter found with the **patulin** biosynthetic gene cluster of *Aspergillus clavatus*. Highlighted in blue is a second ABC transporter found with the patulin gene cluster of *Penicillium expansum*. The ABC transporter of *Fusarium fujikuroi* was chosen as an out-group. Reproduced from Bertrand et al. (2018b), Supporting Information file, in accordance with authors' retainment of privileges.

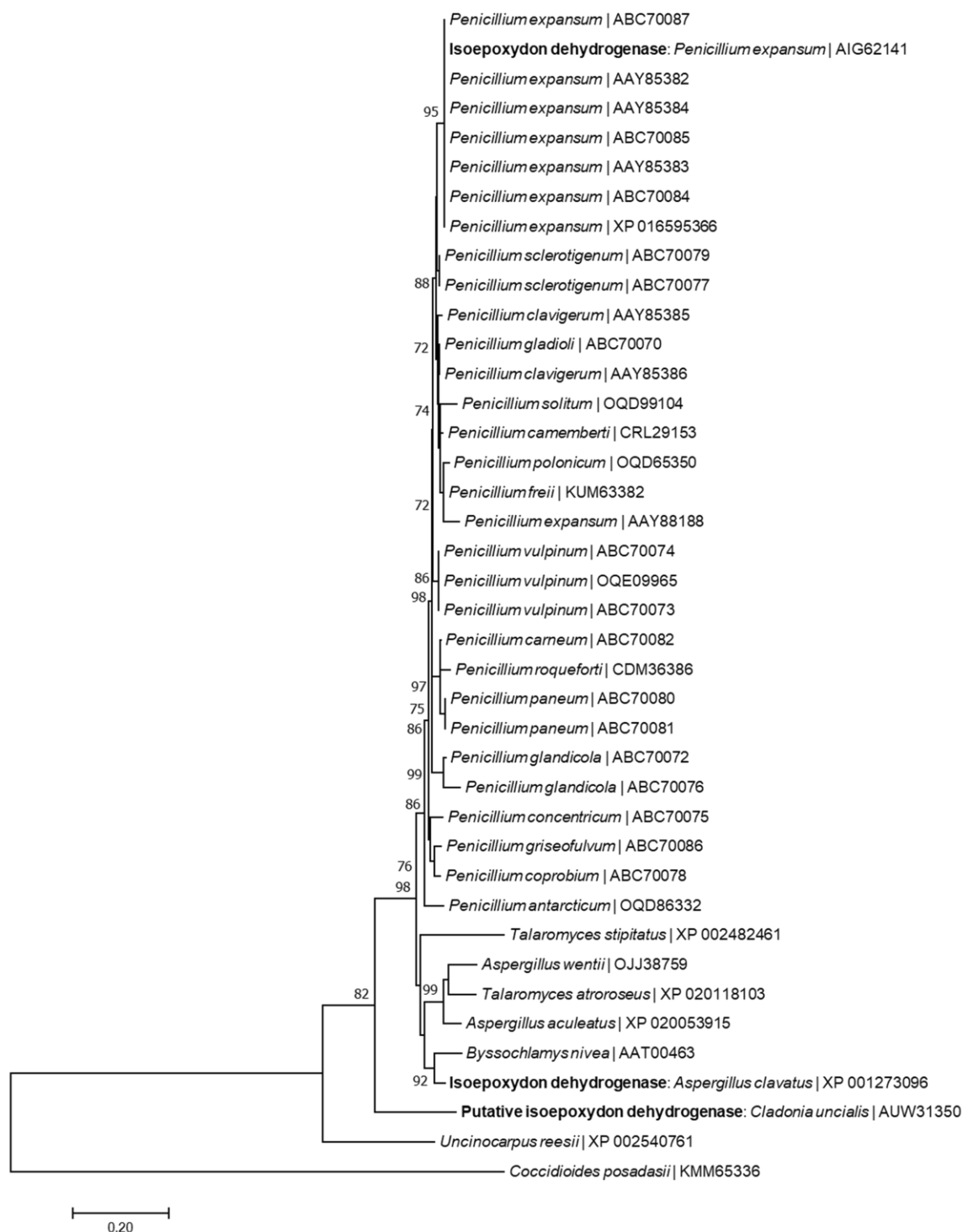


Figure S15: Phylogenetic relationship between a putative **isoepoxydon dehydrogenase** of *Cladonia uncialis* and a genetically similar gene encoding an isoepoxydon dehydrogenase that is part of the **patulin** biosynthetic gene cluster of *Aspergillus clavatus*. Highlighted in blue is a second patulin-related isoepoxydon dehydrogenase encoded in *Penicillium expansum*. The dehydrogenase of *Coccidioides posadasii* (KMM65336) was chosen as an out-group. Reproduced from Bertrand et al. (2018b), Supporting Information file, in accordance with authors' retainment of privileges.

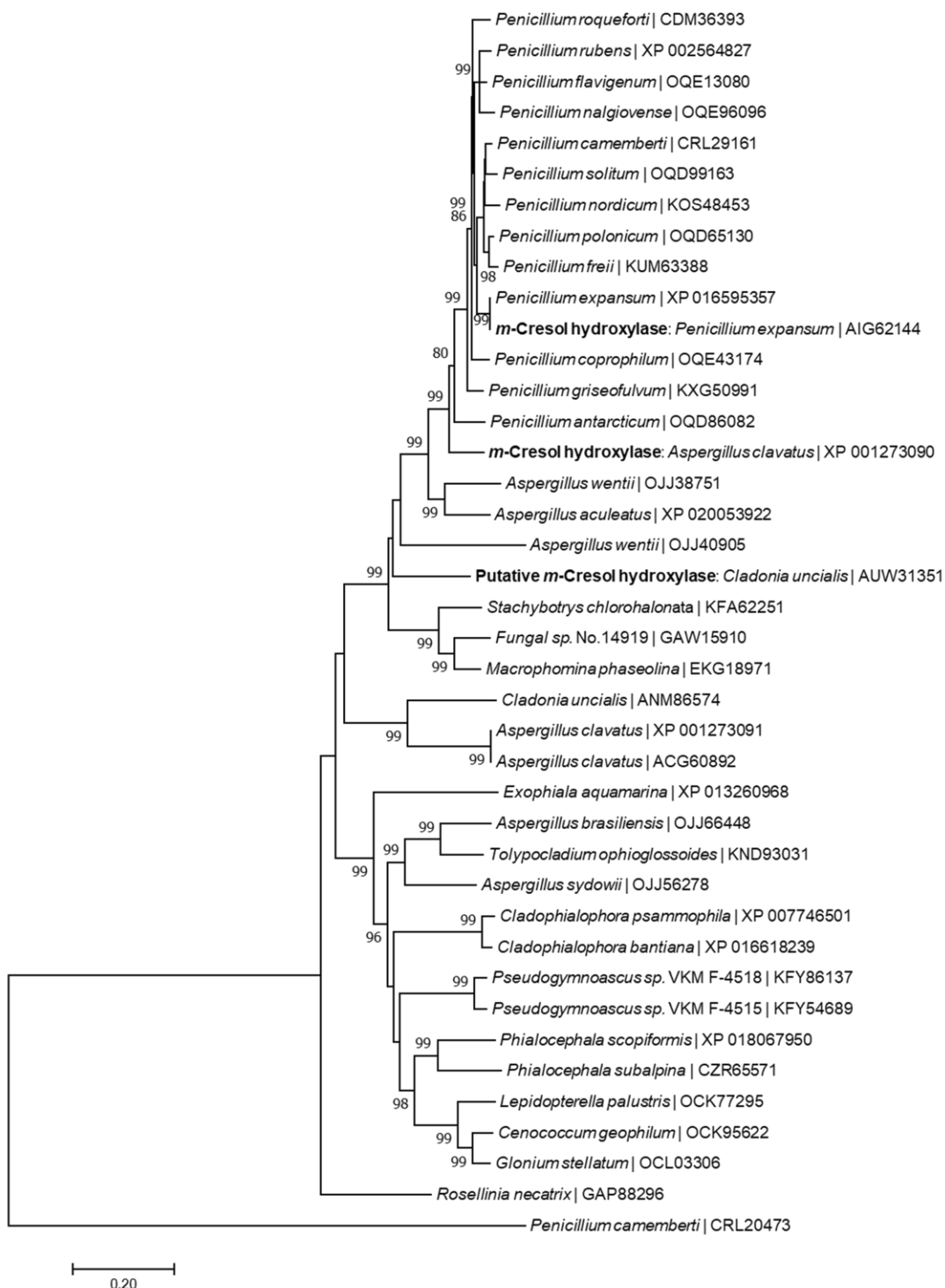


Figure S16: Phylogenetic relationship between a putative *m-cresol hydroxylase* of *Cladonia uncialis* and a genetically similar gene encoding *m-cresol hydroxylase* that is part of the **patulin** biosynthetic gene cluster of *Aspergillus clavatus*. Highlighted in blue is a second patulin-related *m-cresol hydroxylase* encoded in *Penicillium expansum*. The cytochrome p450 of *Penicillium camemberti* (CRL20473) was chosen as an out-group. Reproduced from Bertrand et al. (2018b), Supporting Information file, in accordance with authors' retainment of privileges.

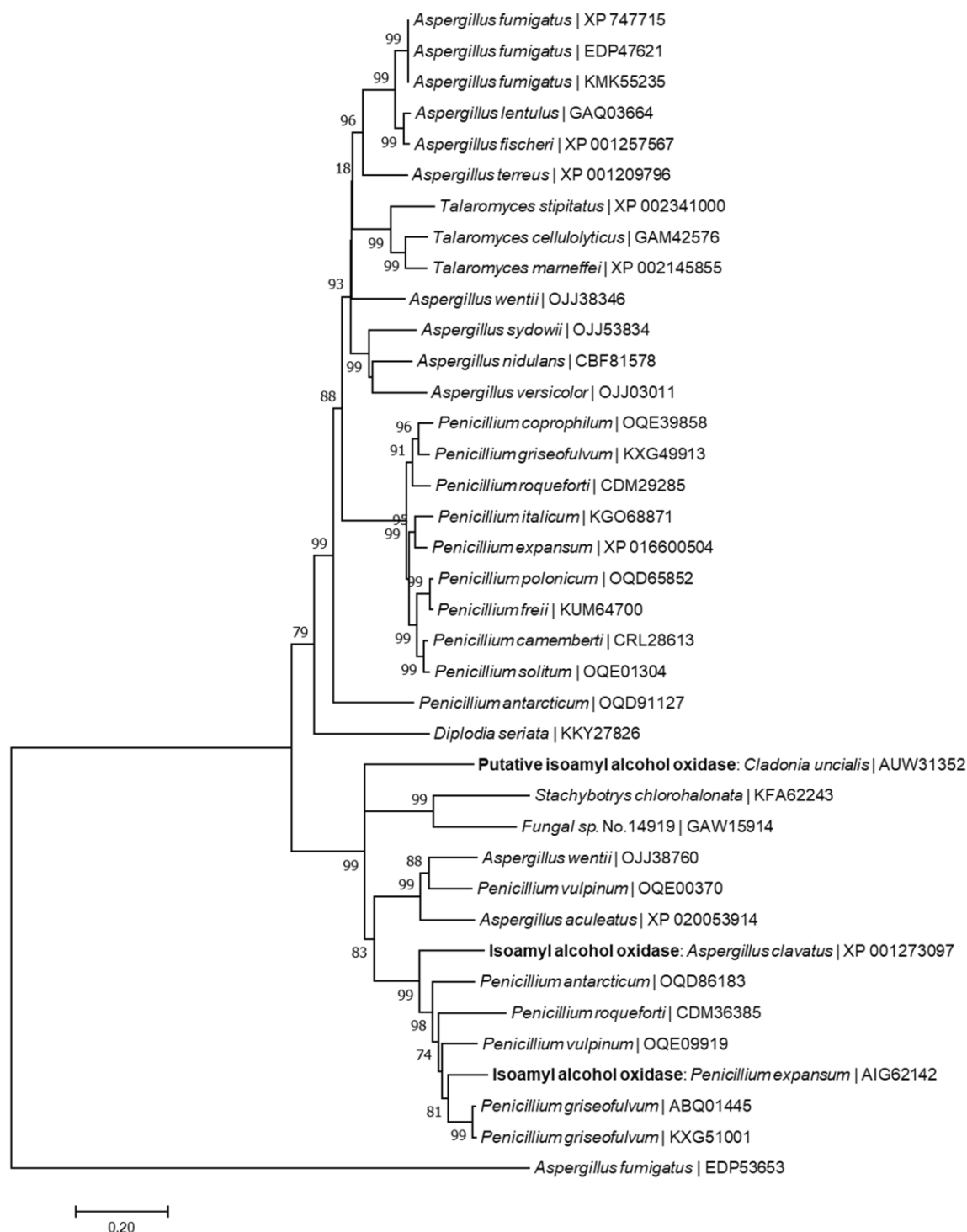


Figure S17: Phylogenetic relationship between a putative **isoamyl alcohol oxidase** of *Cladonia uncialis* and a genetically similar gene encoding an isoamyl alcohol oxidase found with the **patulin** biosynthetic gene cluster of *Aspergillus clavatus*. Highlighted in blue a second patulin-associated isoamyl alcohol oxidase found in *Penicillium expansum*. The FAD oxidase of *Aspergillus fumigatus* (EDP53653) was chosen as an out-group. Reproduced from Bertrand et al. (2018b), Supporting Information file, in accordance with authors' retainment of privileges.

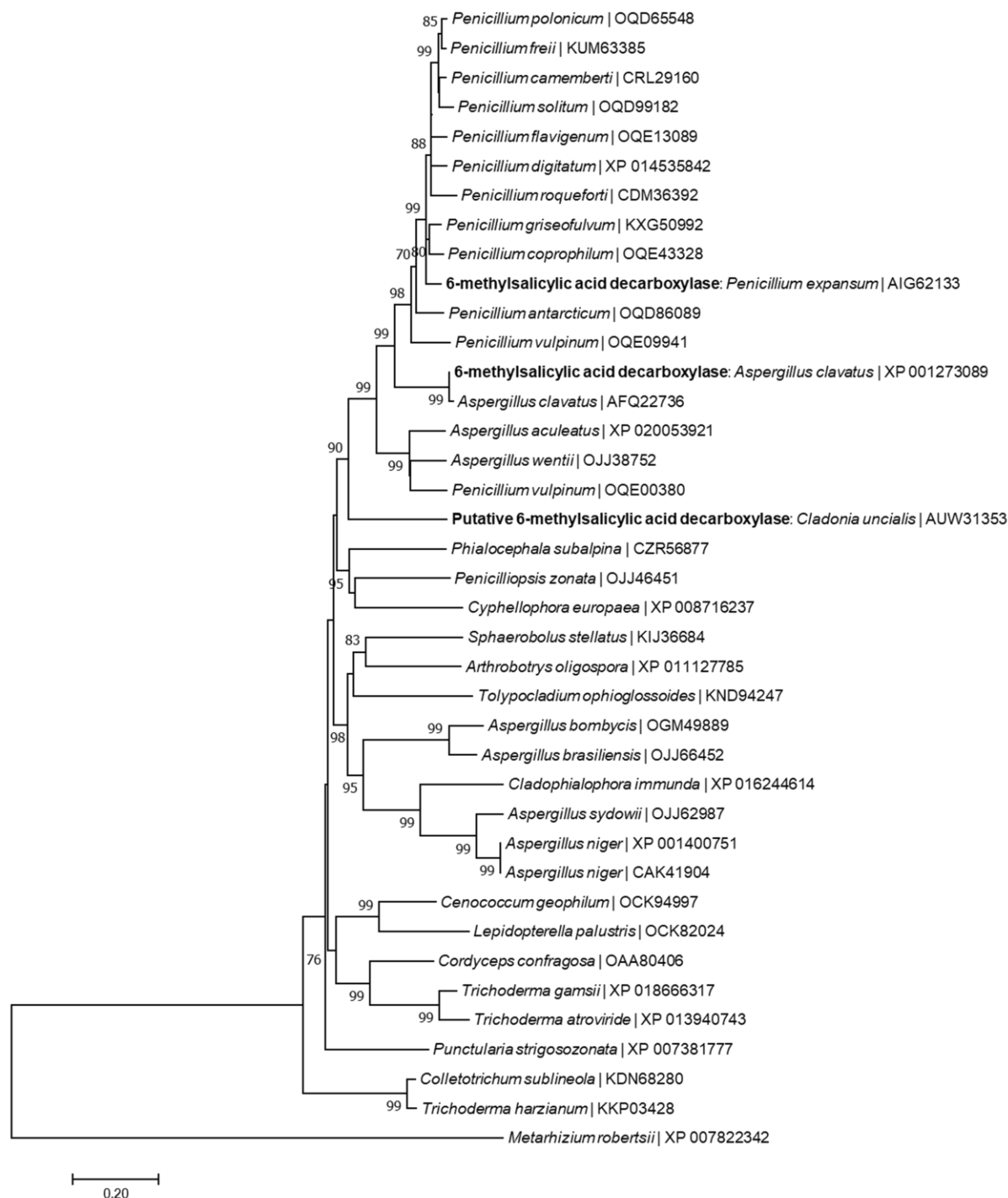


Figure S18: Phylogenetic relationship between a putative **6-methylsalicylic acid decarboxylase** of *Cladonia uncialis* and a genetically similar gene encoding 6-methylsalicylic acid decarboxylase that is part of the **patulin** biosynthetic gene cluster of *Aspergillus clavatus*. Highlighted in blue is a second patulin-related 6-methylsalicylic acid decarboxylase encoded in *Penicillium expansum*. The decarboxylase of *Metarhizium robertsii* (XP_007822342) was chosen as an out-group. Reproduced from Bertrand et al. (2018b), Supporting Information file, in accordance with authors' retainment of privileges.

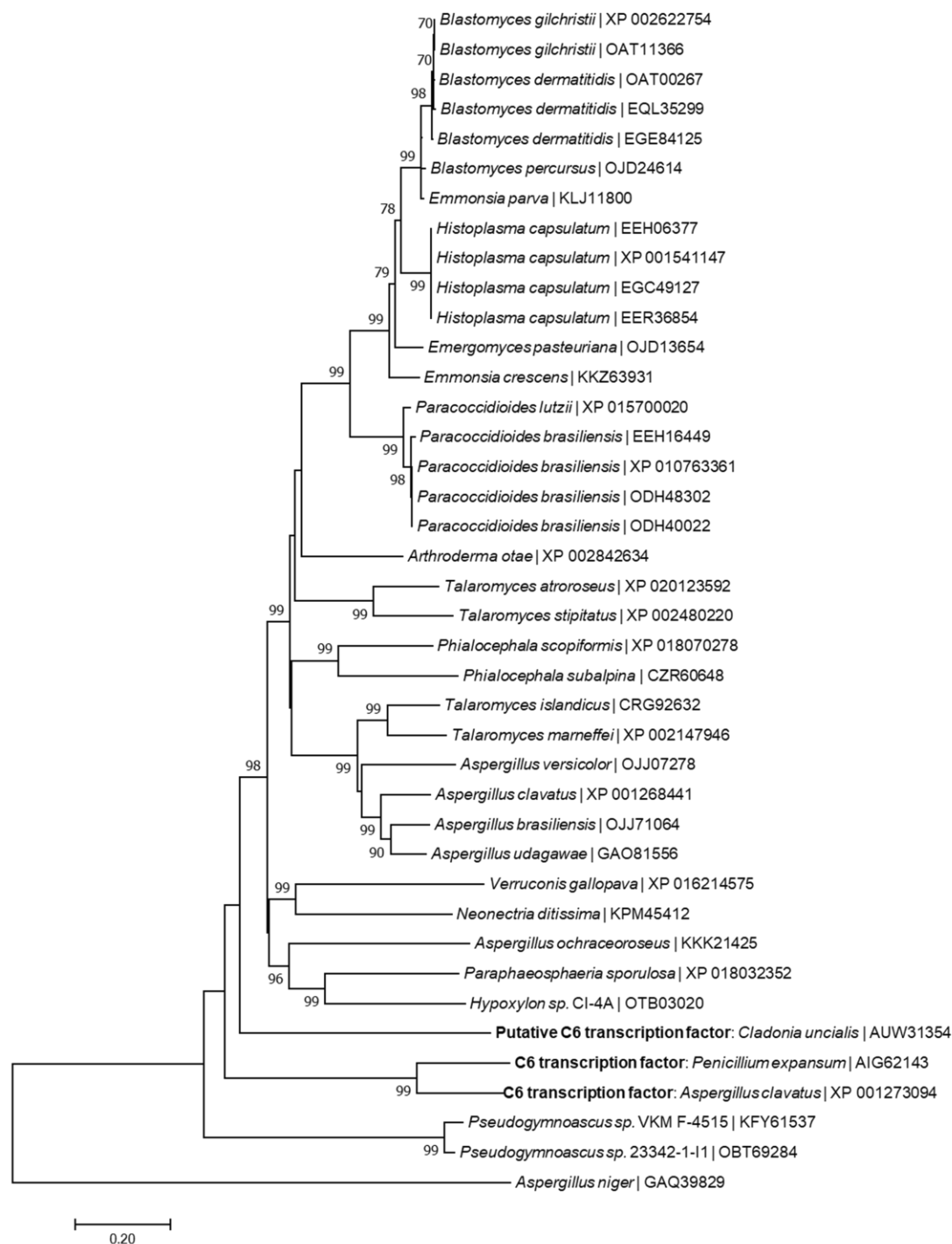


Figure S19: Phylogenetic relationship between a putative **C6 transcription factor** of *Cladonia uncialis* and a genetically similar gene encoding a C6 transcription factor found with the **patulin** biosynthetic gene cluster of *Aspergillus clavatus*. Highlighted in blue is a second C6 transcription factor found with the patulin gene cluster of *Penicillium expansum*. The C6 transcription factor of *Aspergillus niger* (GAQ39829) was chosen as an out-group. Reproduced from Bertrand et al. (2018b), Supporting Information file, in accordance with authors' retainment of privileges.

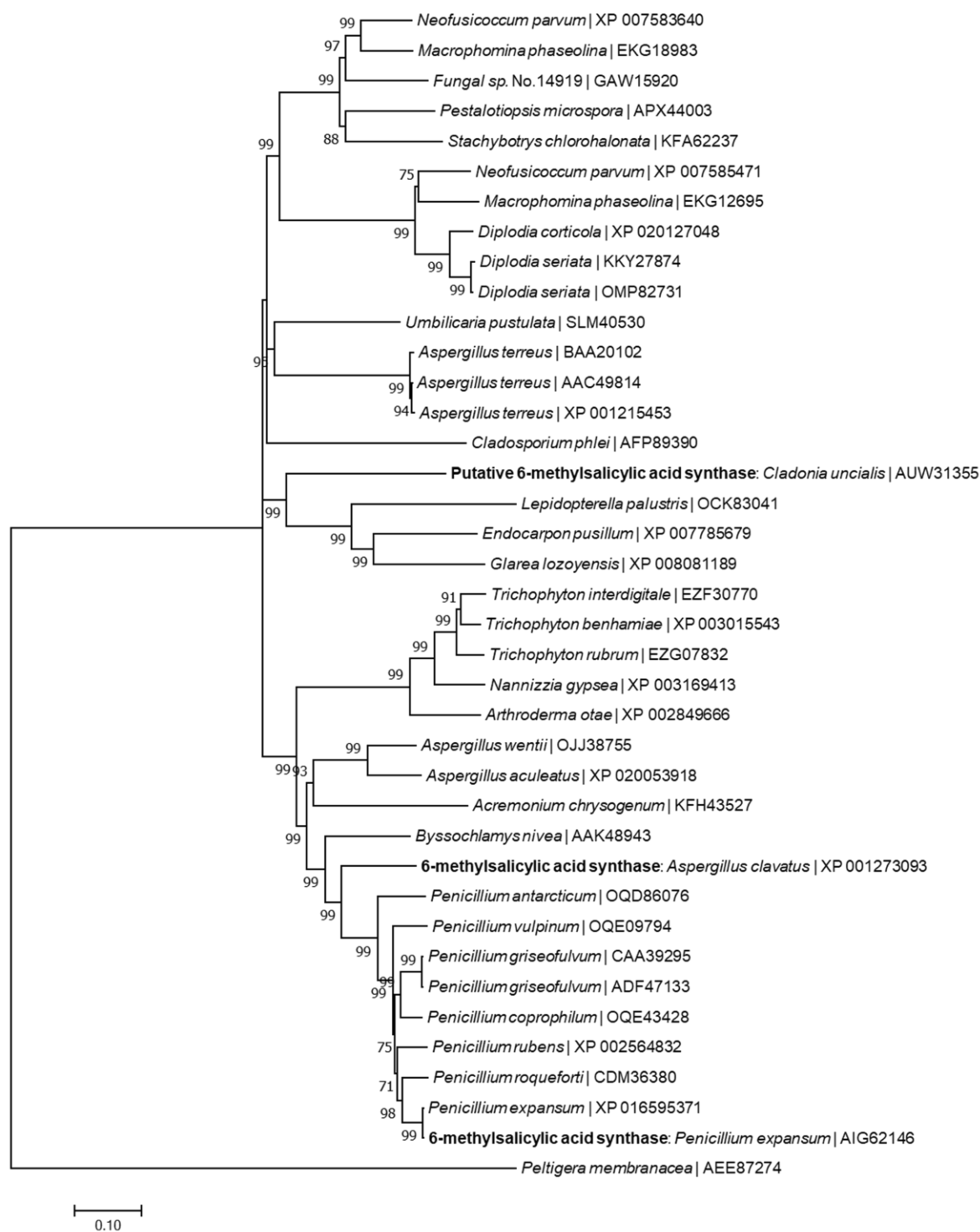


Figure S20: Phylogenetic relationship between a putative **6-methylsalicylic acid synthase** of *Cladonia uncialis* and a genetically similar gene encoding 6-methylsalicylic acid synthase that is part of the **patulin** biosynthetic gene cluster of *Aspergillus clavatus*. Highlighted in blue is a second patulin-related 6-methylsalicylic acid synthase encoded in *Penicillium expansum*. The reducing PKS of *Peltigera membranacea* (AEE87274) was chosen as an out-group. Reproduced from Bertrand et al. (2018b), Supporting Information file, in accordance with authors' retainment of privileges.

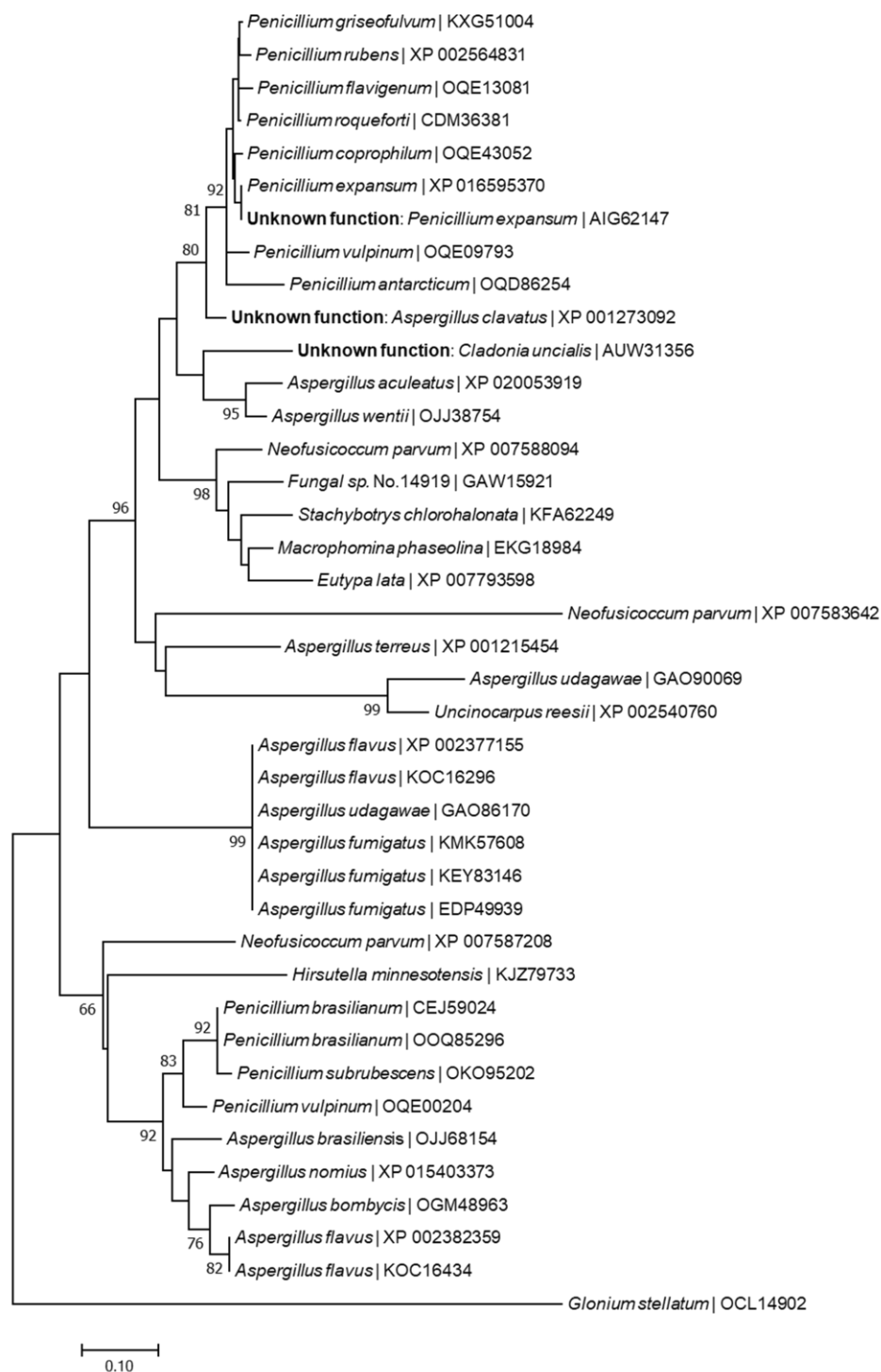


Figure S21: Phylogenetic relationship between a gene of **unknown function (no. 1)** found in *Cladonia uncialis* and a genetically similar gene of unknown function found with the **patulin** biosynthetic gene cluster of *Aspergillus clavatus*. Highlighted in blue is a second gene of unknown function found with the patulin gene cluster of *Penicillium expansum*. The gene encoded in *Glonium stellatum* (OCL14902) was chosen as an out-group. Reproduced from Bertrand et al. (2018b), Supporting Information file, in accordance with authors' retainment of privileges.

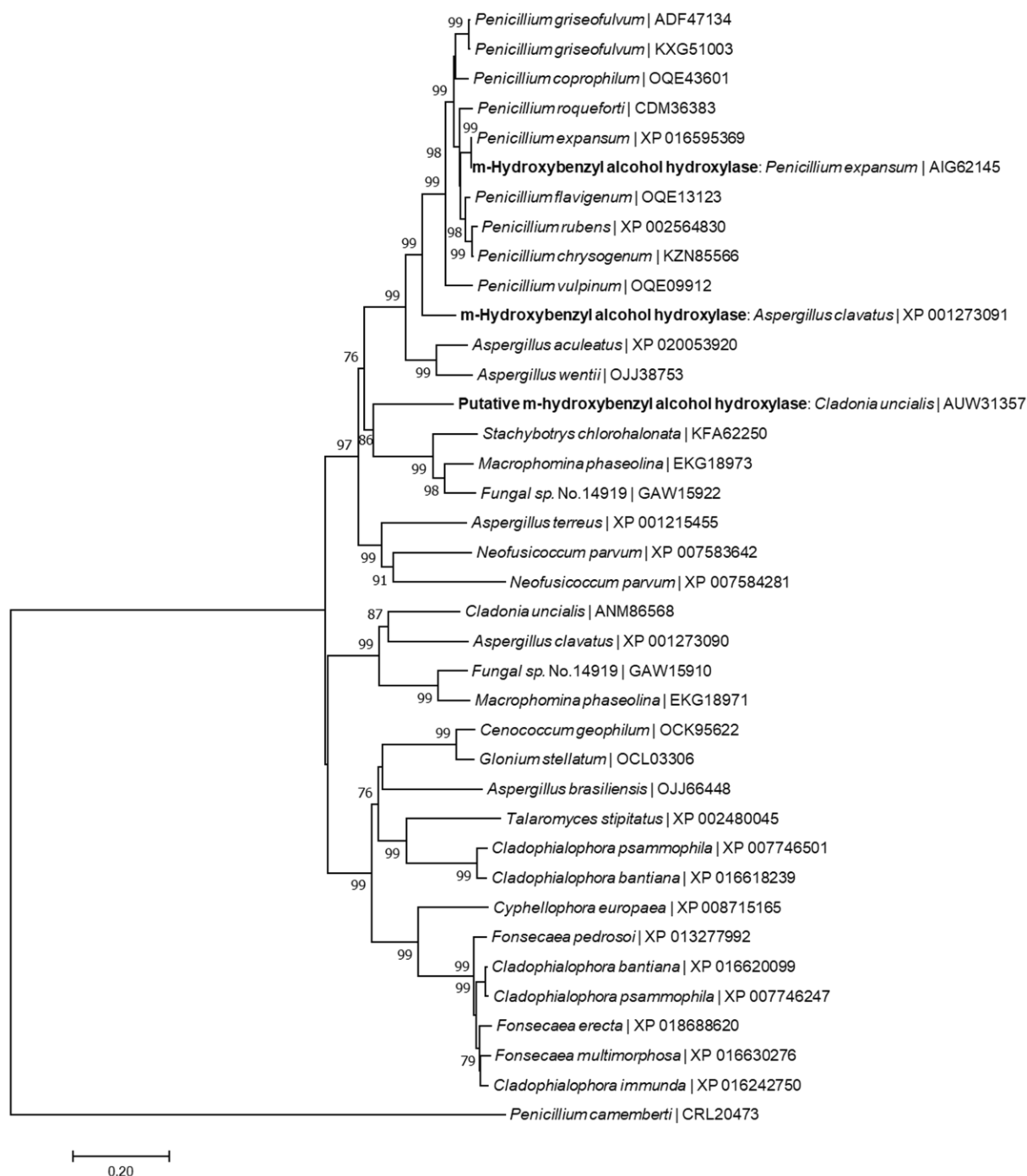


Figure S22: Phylogenetic relationship between a putative *m*-hydroxybenzyl alcohol hydroxylase of *Cladonia uncialis* and a genetically similar gene encoding *m*-hydroxybenzyl alcohol hydroxylase that is part of the **patulin** biosynthetic gene cluster of *Aspergillus clavatus*. Highlighted in blue is a second patulin-related *m*-hydroxybenzyl alcohol hydroxylase encoded in *Penicillium expansum*. The cytochrome p450 of *Penicillium camemberti* (CRL20473) was chosen as an out-group. Reproduced from Bertrand et al. (2018b), Supporting Information file, in accordance with authors' retainment of privileges.

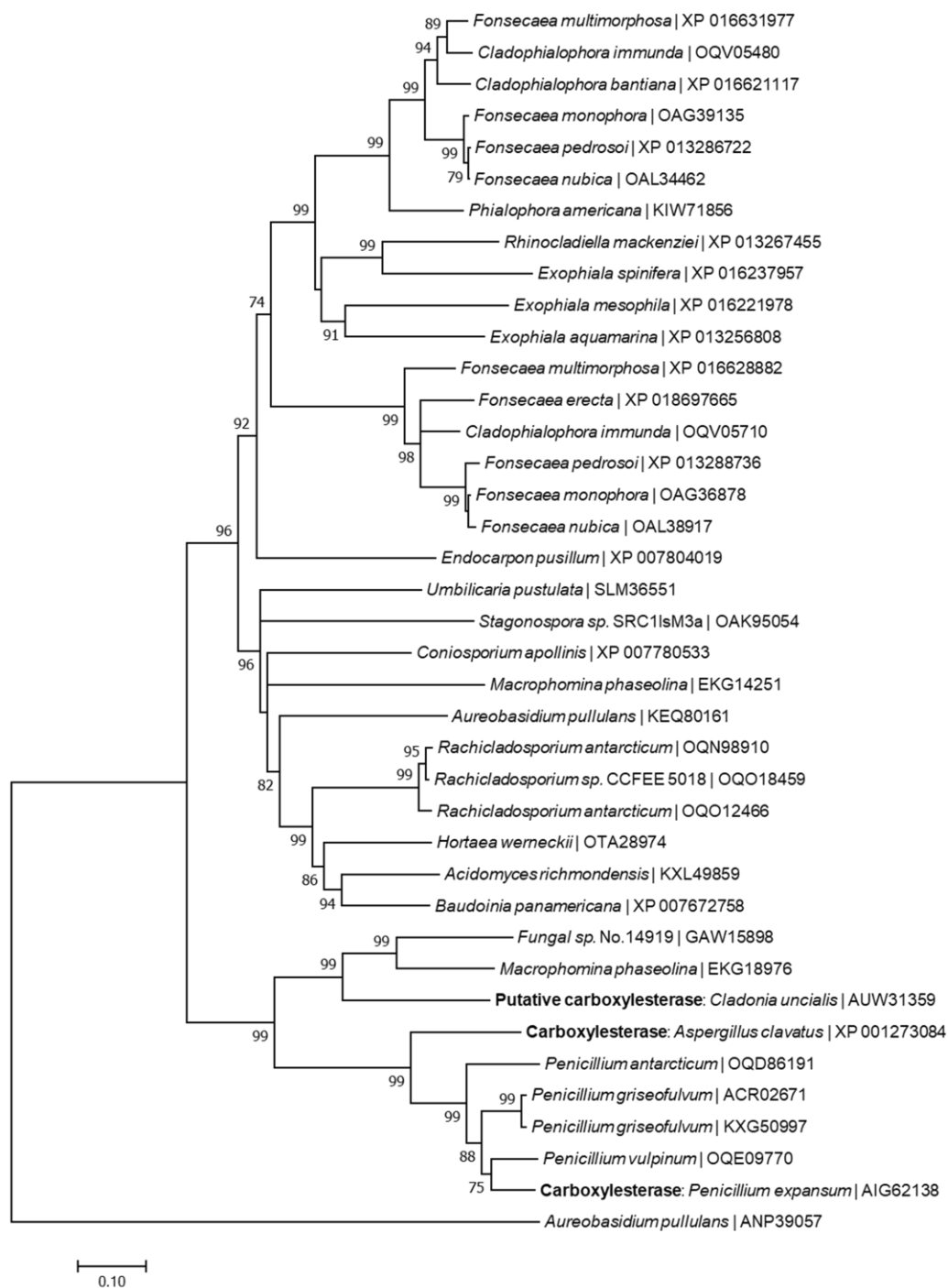


Figure S23: Phylogenetic relationship between a putative **carboxylesterase** of *Cladonia uncialis* and a genetically similar gene encoding a carboxylesterase found with the **patulin** biosynthetic gene cluster of *Aspergillus clavatus*. Highlighted in blue is a second carboxylesterase found with the patulin gene cluster of *Penicillium expansum*. The carboxylesterase of *Aureobasidium pullulans* was selected as an out-group. Reproduced from Bertrand et al. (2018b), Supporting Information file, in accordance with authors' retainment of privileges.

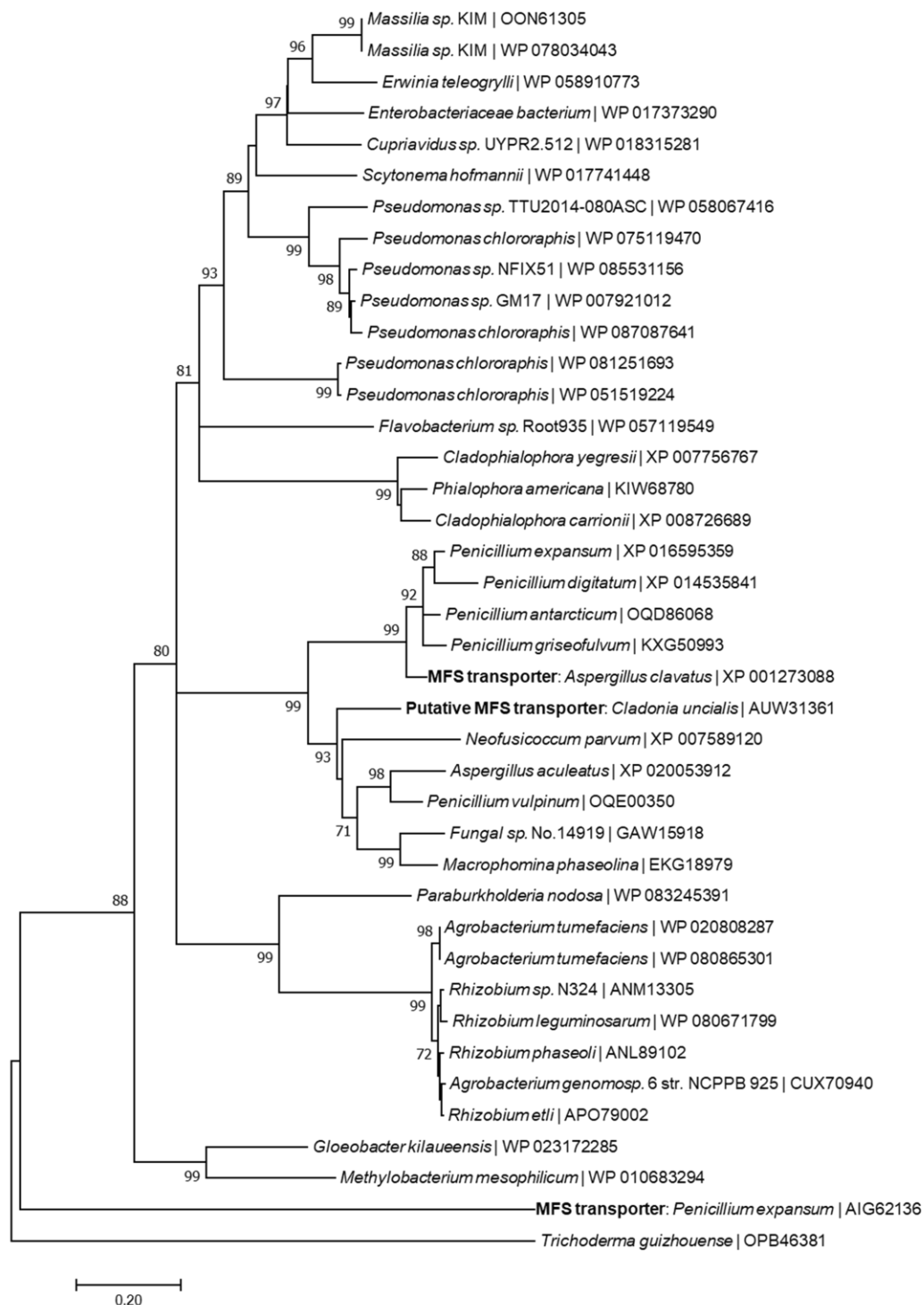


Figure S24: Phylogenetic relationship between a putative **MFS transporter** of *Cladonia uncialis* and a genetically similar gene encoding a MFS transporter found with the **patulin** biosynthetic gene cluster of *Aspergillus clavatus*. Highlighted in blue is a second MFS transporter found with the patulin gene cluster of *Penicillium expansum*. The MFS transporter of *Trichoderma guizhouense* (OPB46381) was chosen as an out-group. Reproduced from Bertrand et al. (2018b), Supporting Information file, in accordance with authors' retainment of privileges.

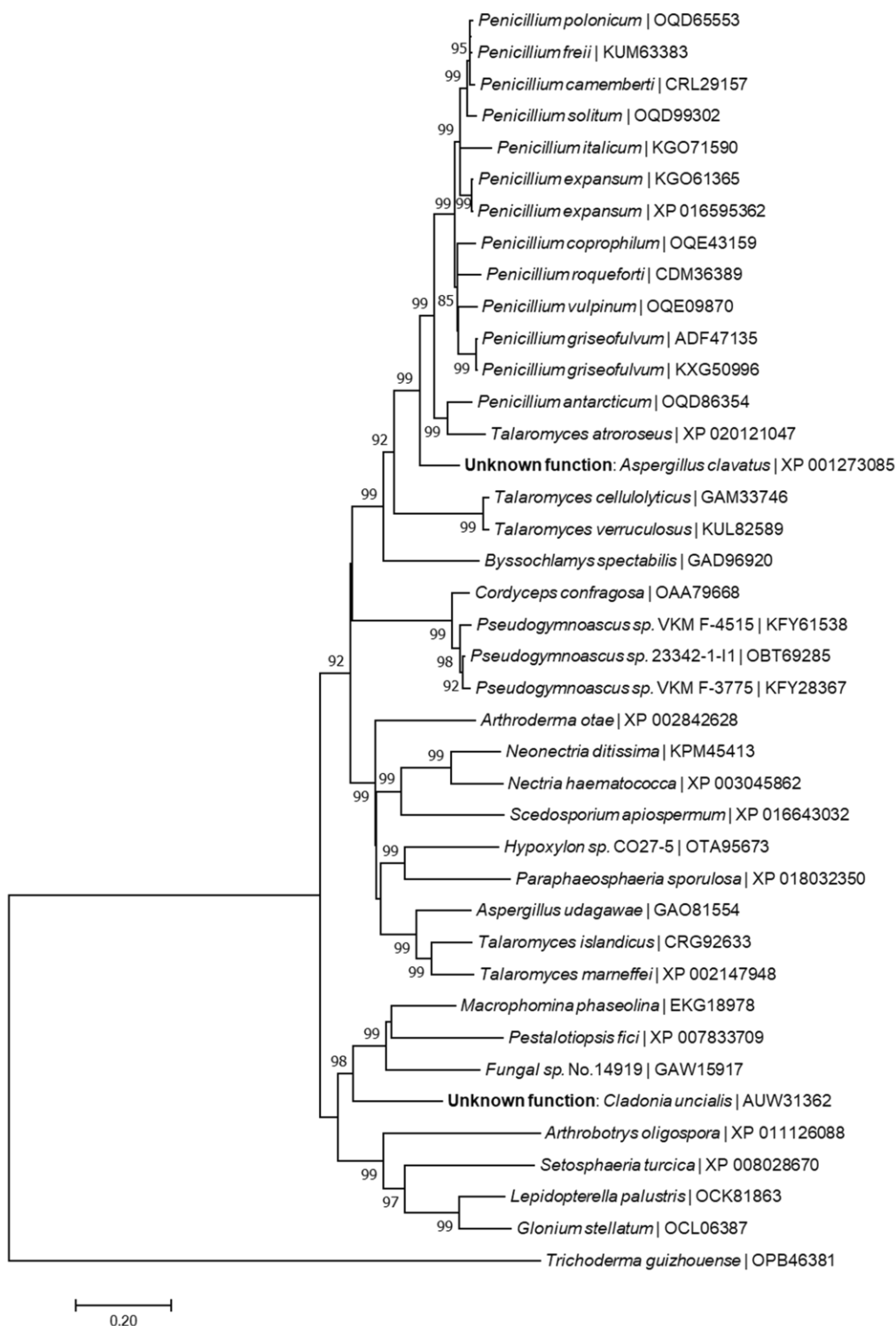


Figure S25: Phylogenetic relationship between a gene of **unknown function (no. 2)** found in *Cladonia uncialis* and a genetically similar gene of unknown function found with the **patulin** biosynthetic gene cluster of *Aspergillus clavatus*. The gene encoded in *Trichoderma guizhouense* (OPB46381) was chosen as an out-group. Reproduced from Bertrand et al. (2018b), Supporting Information file, in accordance with authors' retainment of privileges.

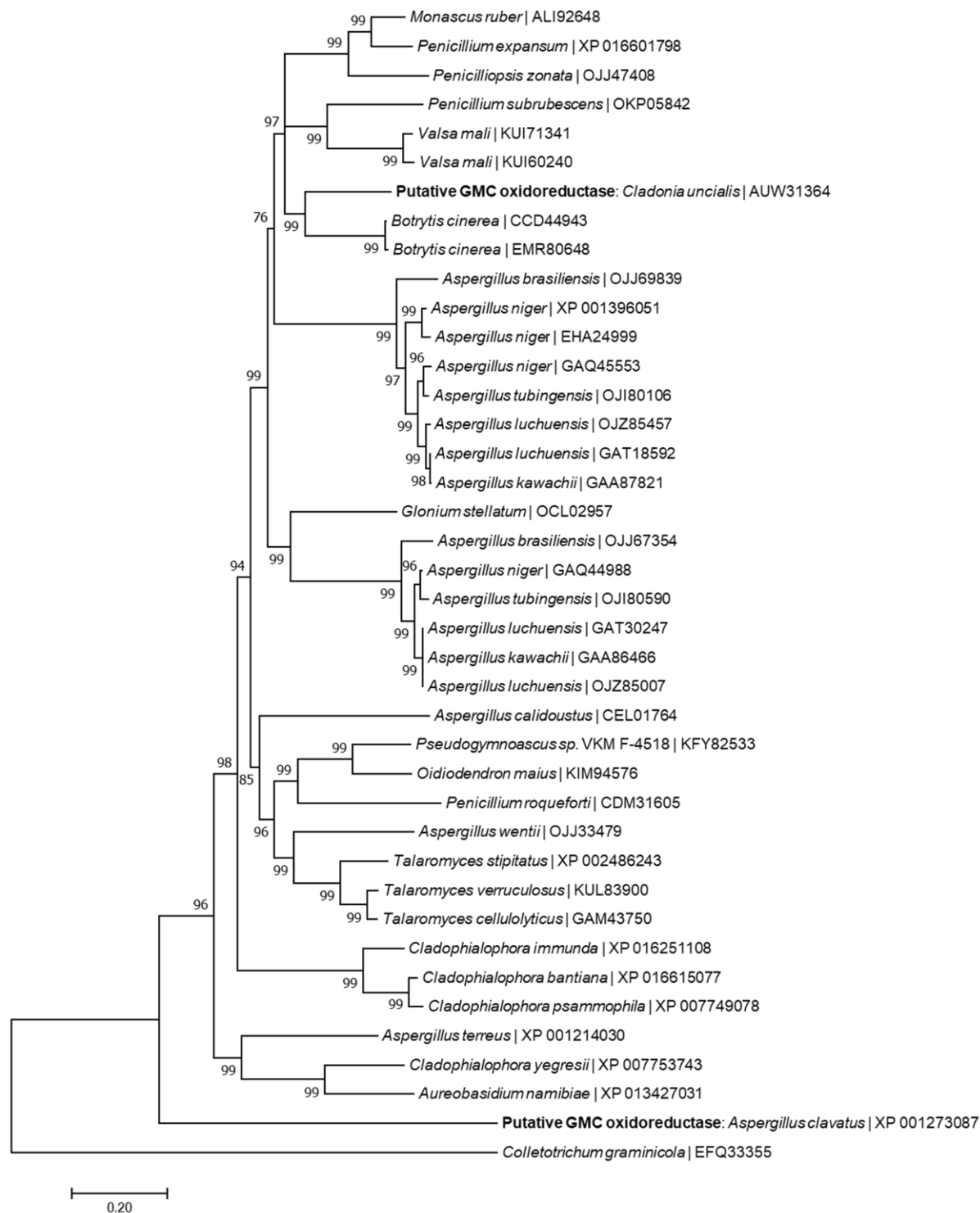


Figure S26: Phylogenetic relationship between a putative **GMC oxidoreductase** of *Cladonia uncialis* and a genetically similar gene encoding a GMC oxidoreductase found with the **patulin** biosynthetic gene cluster of *Aspergillus clavatus*. The GMC oxidoreductase of *Colletotrichum graminicola* (EFQ33355) was chosen as an out-group. Reproduced from Bertrand et al. (2018b), Supporting Information file, in accordance with authors' retainment of privileges.

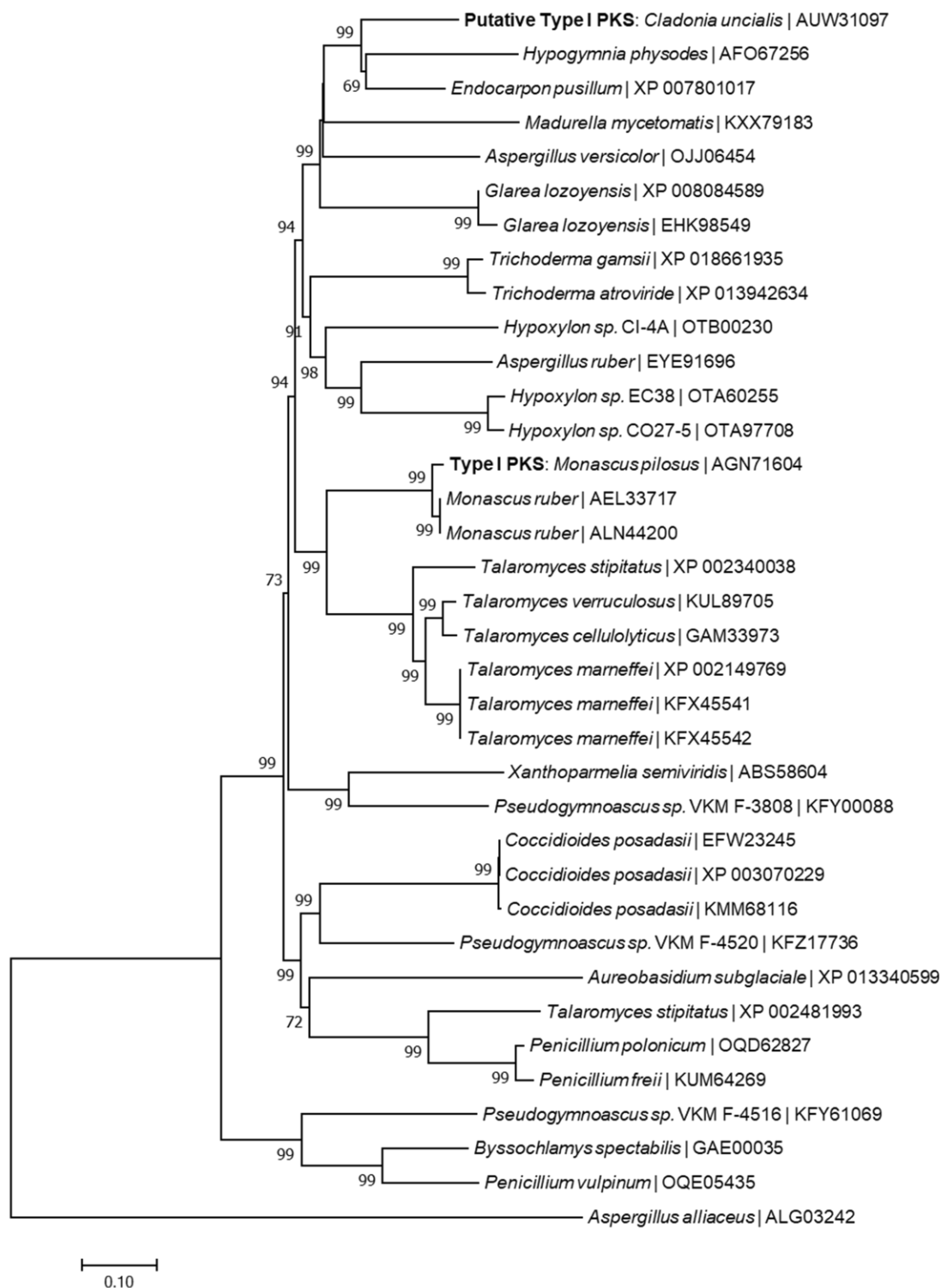


Figure S27: Phylogenetic relationship between a putative **polyketide synthase** of *Cladonia uncialis* and a genetically similar gene encoding a polyketide synthase that is part of the **azaphilone** biosynthetic gene cluster of *Monascus pilosus*. The PKS of *Aspergillus alliaceus* (ALG03242) was chosen as an out-group. Reproduced from Bertrand et al. (2018b), Supporting Information file, in accordance with authors' retainment of privileges.

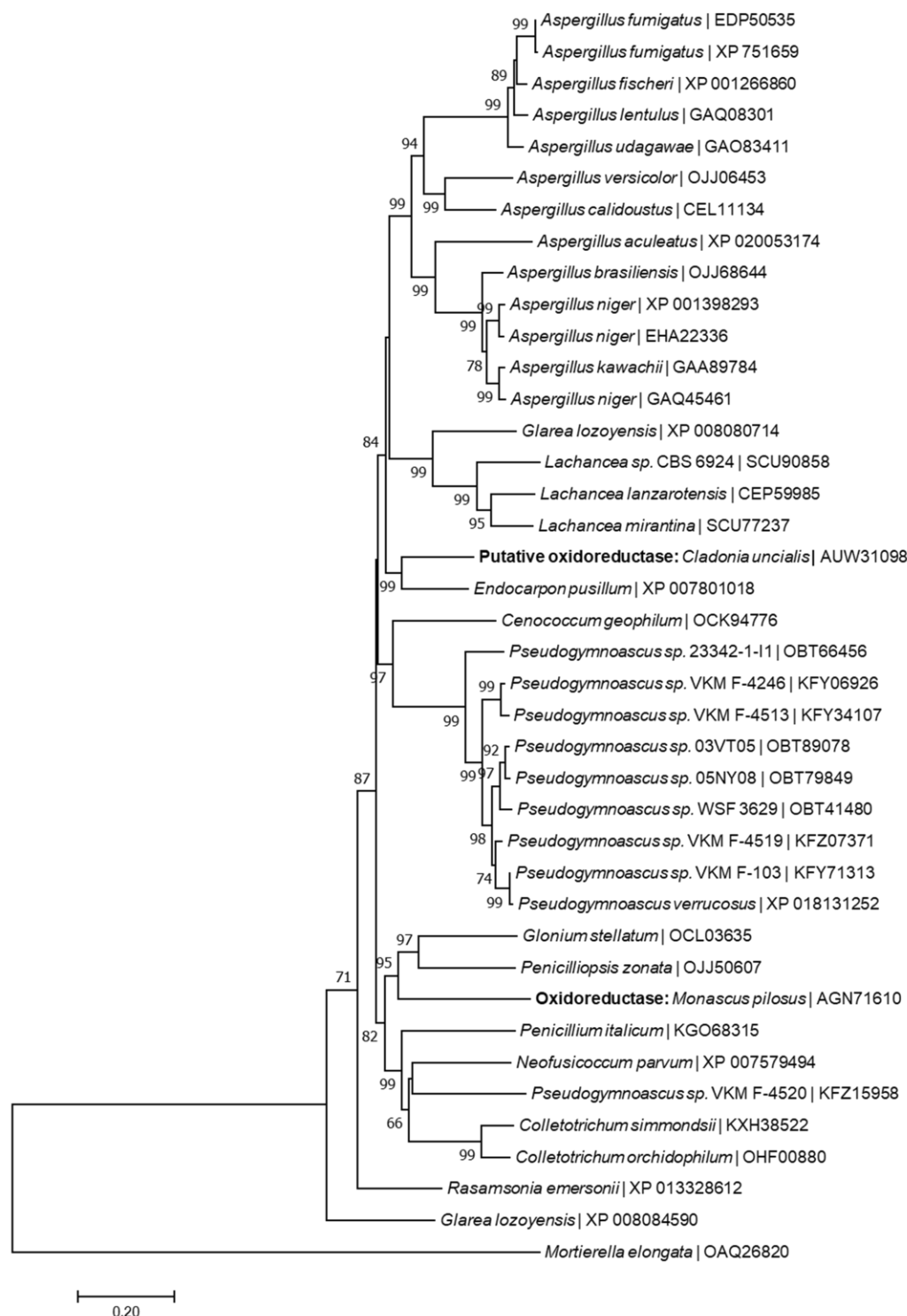


Figure S28: Phylogenetic relationship between a putative **oxidoreductase (no. 1)** of *Cladonia uncialis* and a genetically similar gene that is proposed to encode an oxidoreductase that participates in **azaphilone** biosynthesis in *Monascus pilosus*. The oxidoreductase of *Mortierella elongata* (OAQ26820) was chosen as an out-group. Reproduced from Bertrand et al. (2018b), Supporting Information file, in accordance with authors' retainment of privileges.

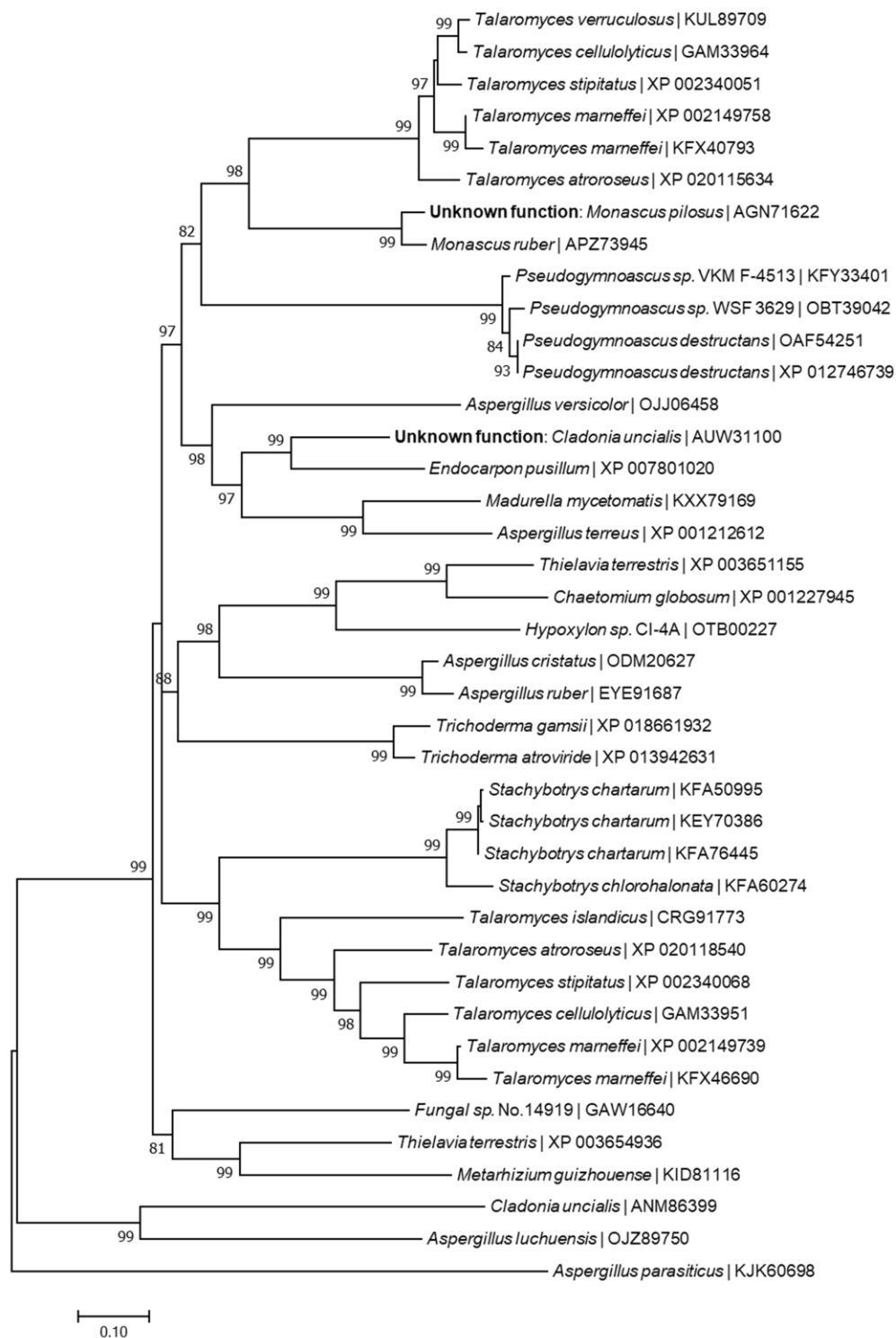


Figure S29: Phylogenetic relationship between a gene of **unknown function (no. 1)** of *Cladonia uncialis* and a genetically similar gene of unknown function of which its product is proposed to participate in **azaphilone** biosynthesis in *Monascus pilosus*. The gene encoded in *Aspergillus parasiticus* (KJK60698) was chosen as an out-group. Reproduced from Bertrand et al. (2018b), Supporting Information file, in accordance with authors' retainment of privileges.

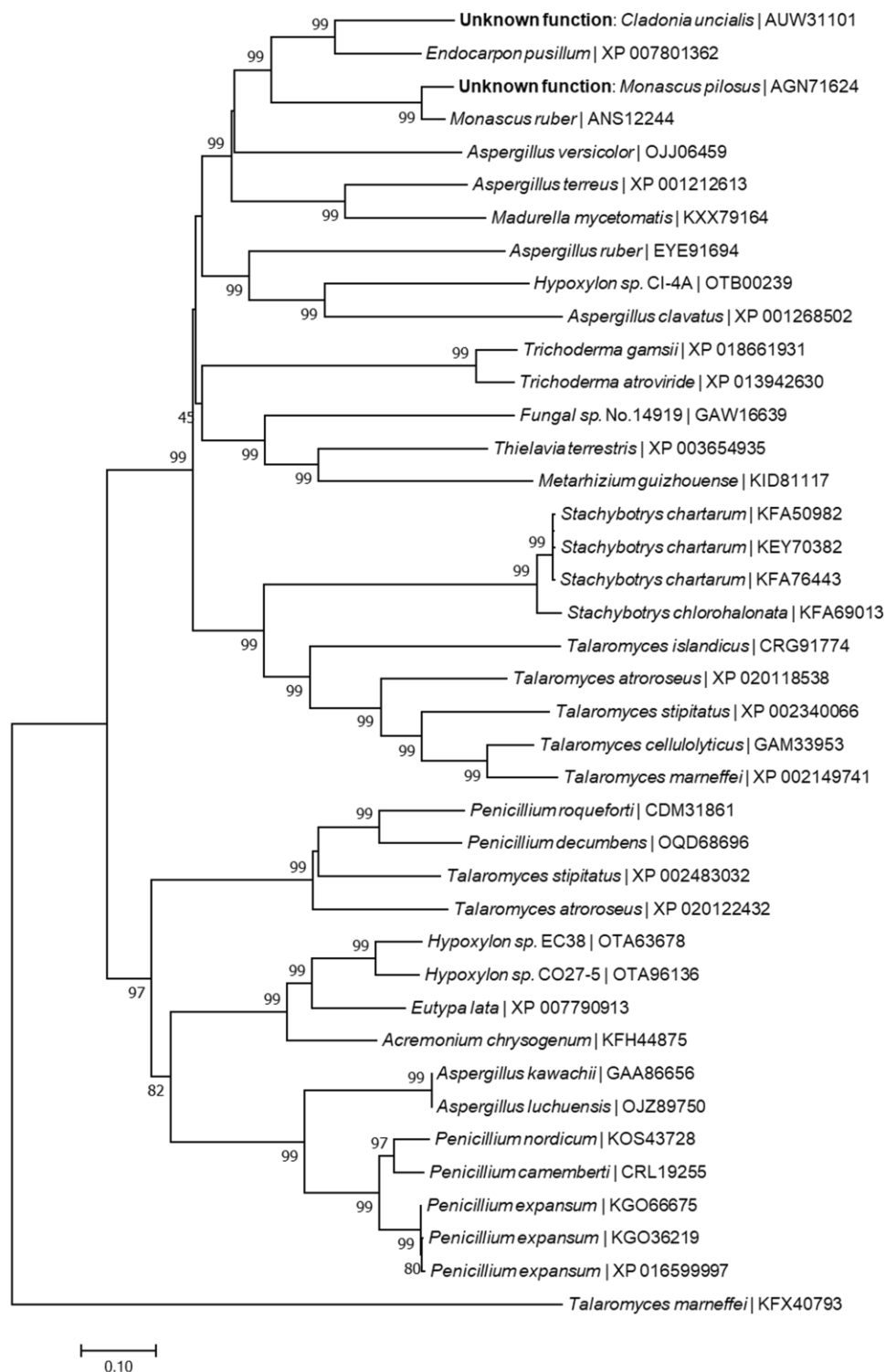


Figure S30: Phylogenetic relationship between a gene of **unknown function (no. 2)** of *Cladonia uncialis* and a genetically similar gene of unknown function of which its product is proposed to participate in **azaphilone** biosynthesis in *Monascus pilosus*. The gene encoded in *Talaromyces marneffei* (KFX40793) was chosen as an out-group. Reproduced from Bertrand et al. (2018b), Supporting Information file, in accordance with authors' retainment of privileges.

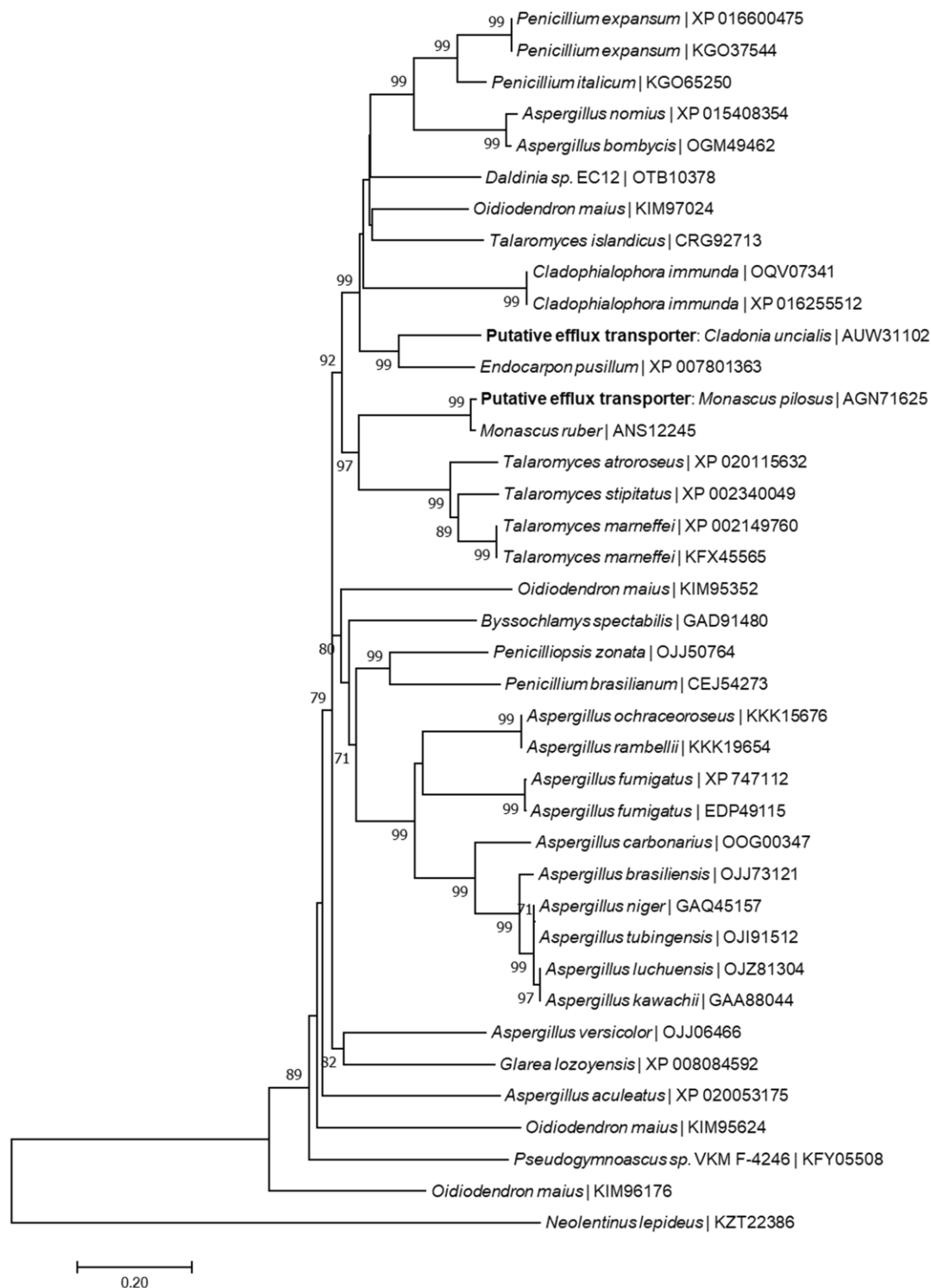


Figure S31: Phylogenetic relationship between a putative **efflux transporter** of *Cladonia uncialis* and a genetically similar gene that is proposed to be an efflux transporter associated with the **azaphilone** biosynthetic gene cluster of *Monascus pilosus*. The efflux transporter of *Neolentinus lepideus* (KZT22386) was chosen as an out-group. Reproduced from Bertrand et al. (2018b), Supporting Information file, in accordance with authors' retainment of privileges.

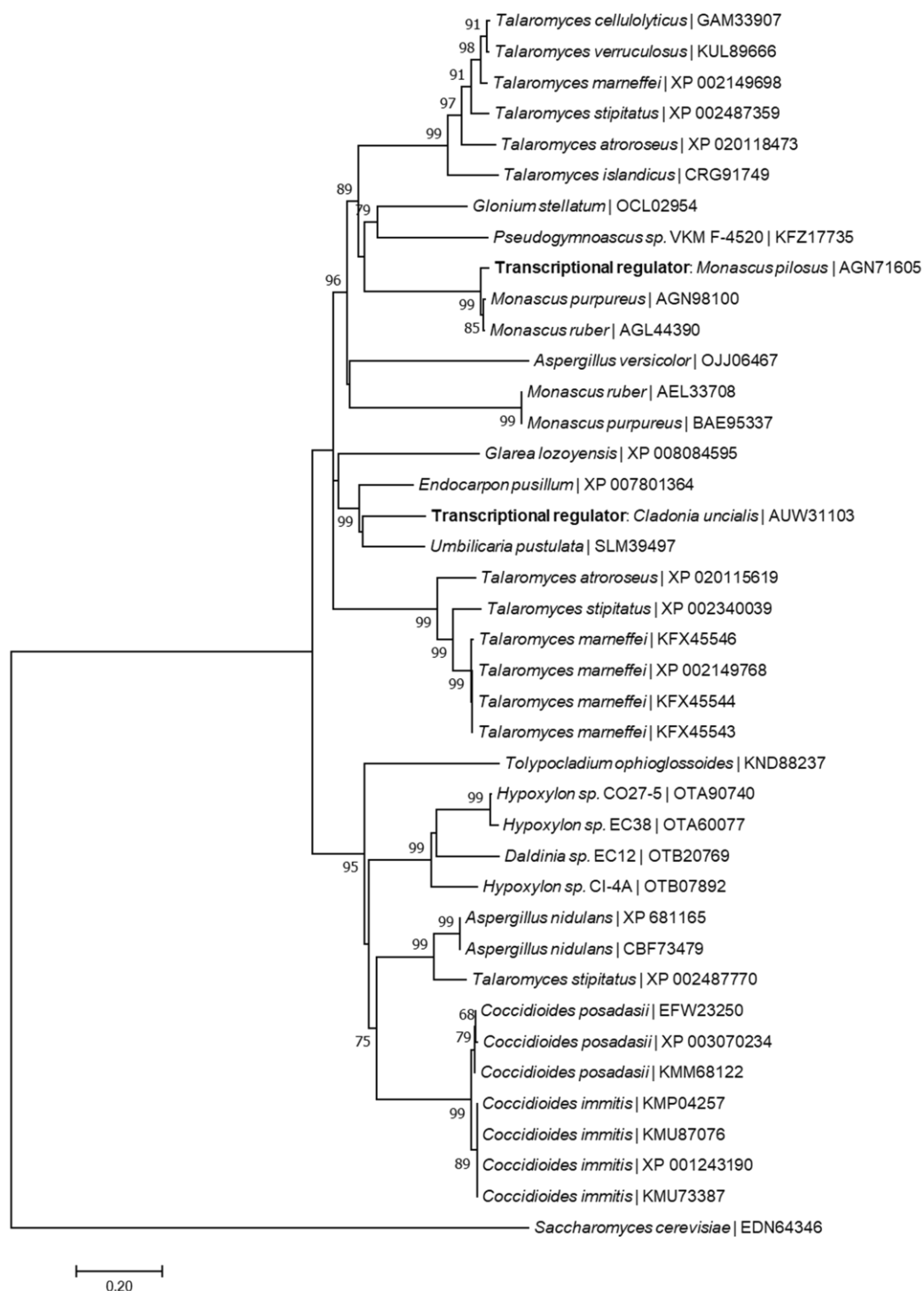


Figure S32: Phylogenetic relationship between a putative **transcriptional regulator** of *Cladonia uncialis* and a genetically similar gene that is proposed to be a transcriptional regulator that is associated with the **azaphilone** biosynthetic gene cluster of *Monascus pilosus*. The transcriptional regulator of *Saccharomyces cerevisiae* was chosen as an out-group. Reproduced from Bertrand et al. (2018b), Supporting Information file, in accordance with authors' retainment of privileges.

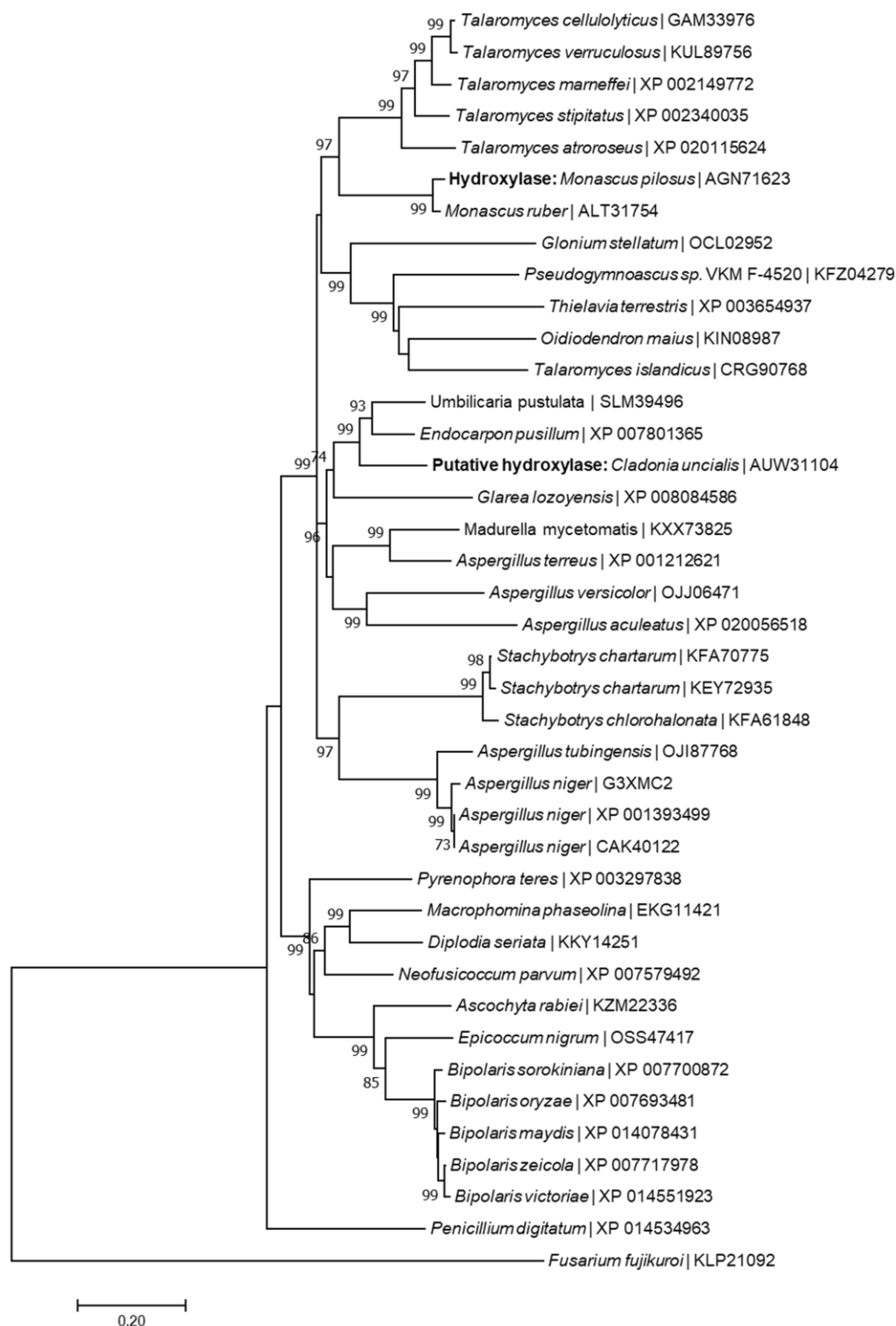


Figure S33: Phylogenetic relationship between a putative **hydroxylase** of *Cladonia uncialis* and a genetically similar gene encoding a hydroxylase that is part of the **azaphilone** biosynthetic gene cluster of *Monascus pilosus*. The hydroxylase of *Fusarium fujikuroi* (KLP21092) was chosen as an out-group. Reproduced from Bertrand et al. (2018b), Supporting Information file, in accordance with authors' retainment of privileges.

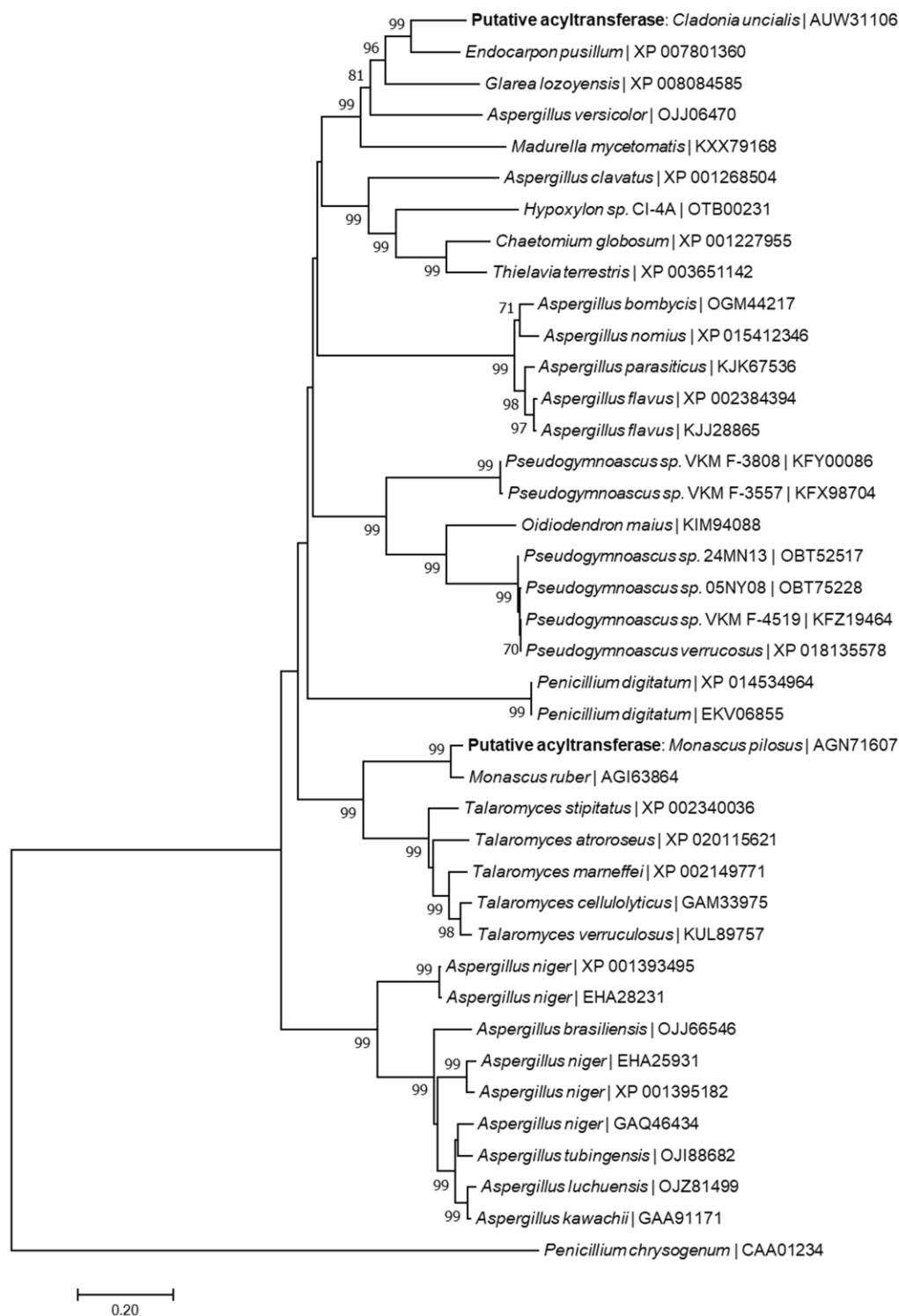


Figure S34: Phylogenetic relationship between a putative **acyltransferase** of *Cladonia uncialis* and a genetically similar gene encoding an acyltransferase that is part of the **azaphilone** biosynthetic gene cluster of *Monascus pilosus*. The acyltransferase of *Penicillium chrysogenum* (CAA01234) was chosen as an out-group. Reproduced from Bertrand et al. (2018b), Supporting Information file, in accordance with authors' retainment of privileges.

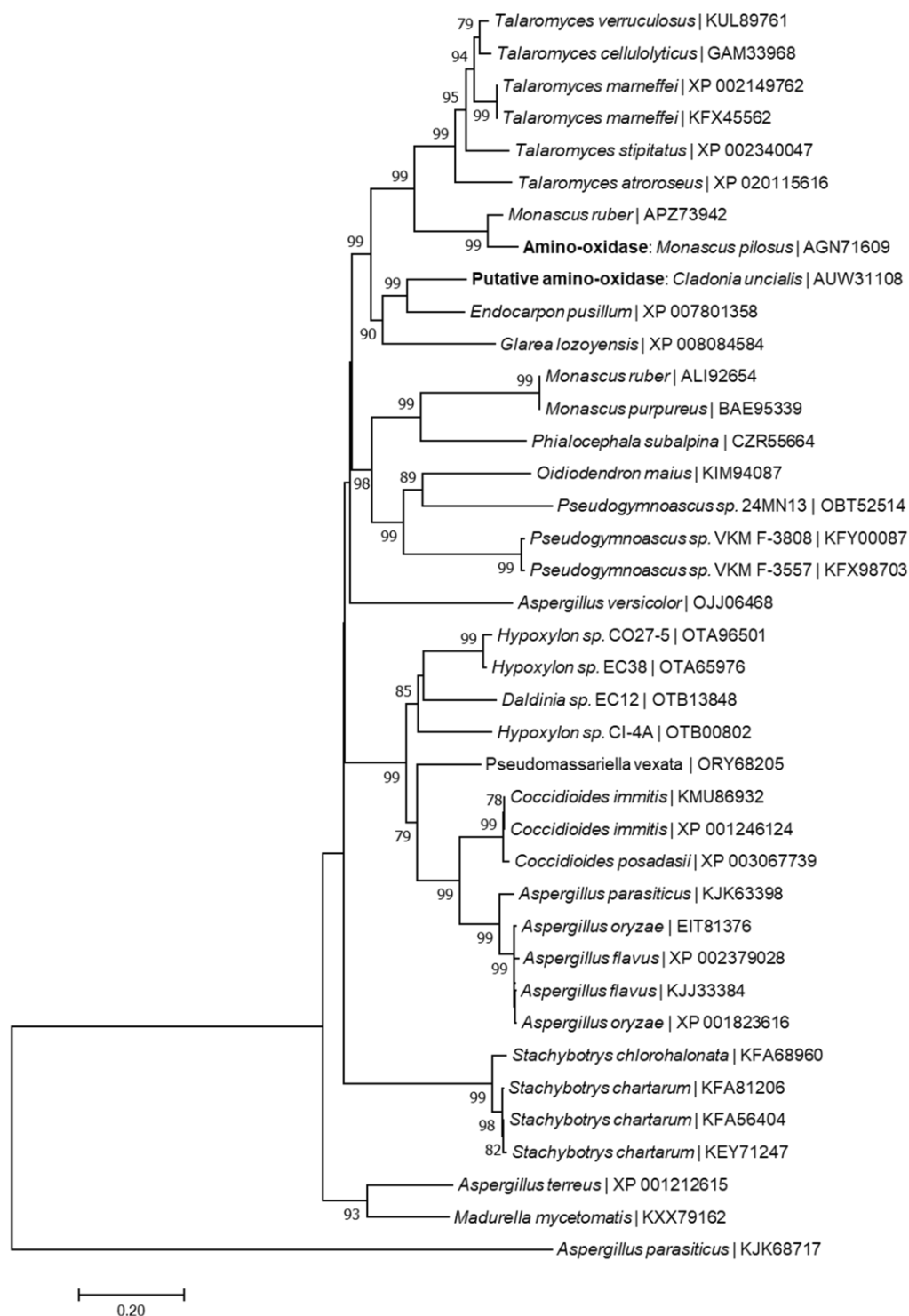


Figure S35: Phylogenetic relationship between a putative **amino-oxidase** of *Cladonia uncialis* and a genetically similar gene encoding an amino-oxidase that is part of the **azaphilone** biosynthetic gene cluster of *Monascus pilosus*. The amino-oxidase of *Aspergillus parasiticus* (KJK68717) was chosen as an out-group. Reproduced from Bertrand et al. (2018b), Supporting Information file, in accordance with authors' retainment of privileges.

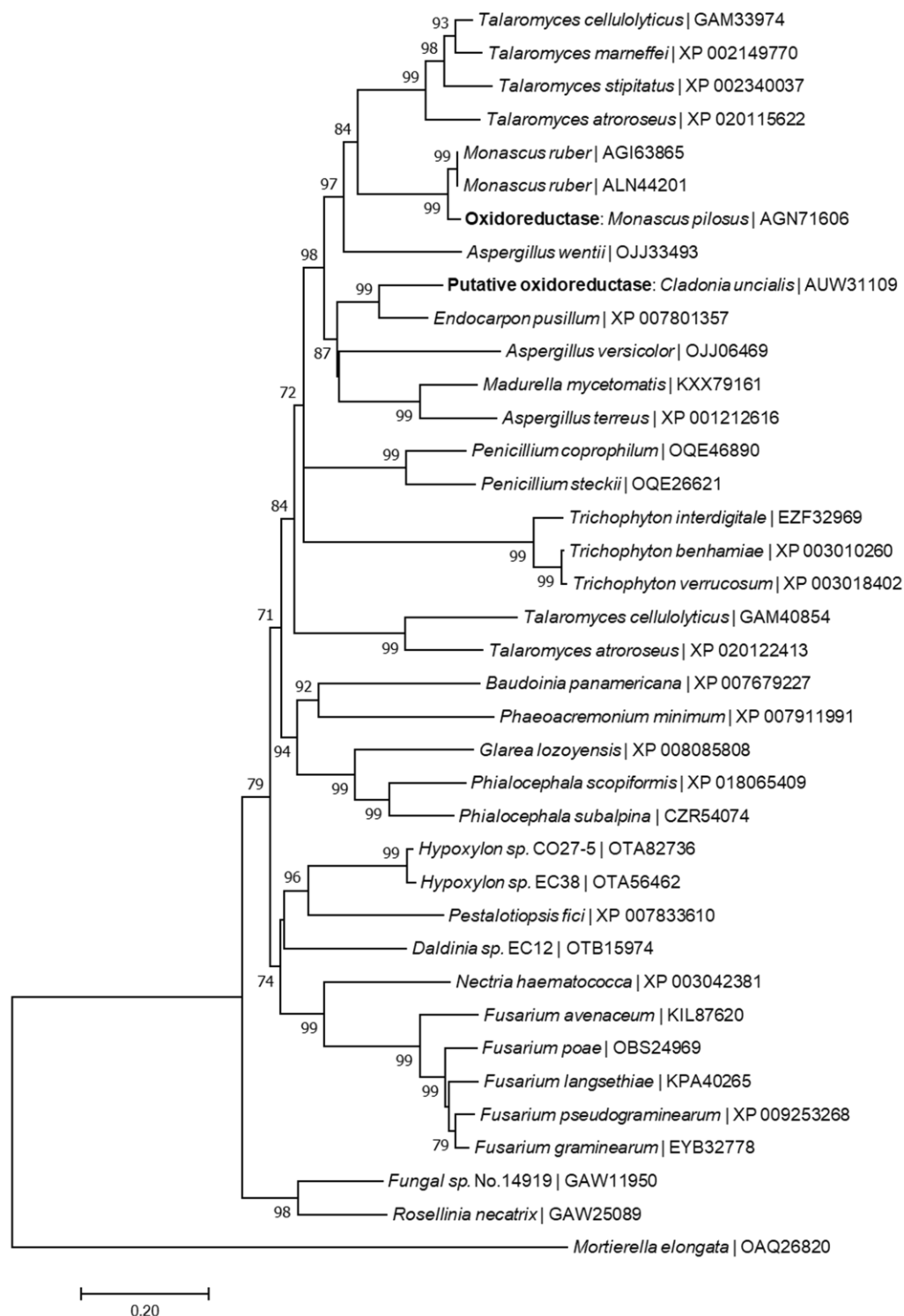


Figure S36: Phylogenetic relationship between a putative **oxidoreductase (no. 2)** of *Cladonia uncialis* and a genetically similar gene that is proposed to encode an oxidoreductase that participates in **azaphilone** biosynthesis in *Monascus pilosus*. The oxidoreductase of *Mortierella elongata* (OAQ26820) was chosen as an out-group. Reproduced from Bertrand et al. (2018b), Supporting Information file, in accordance with authors' retainment of privileges.

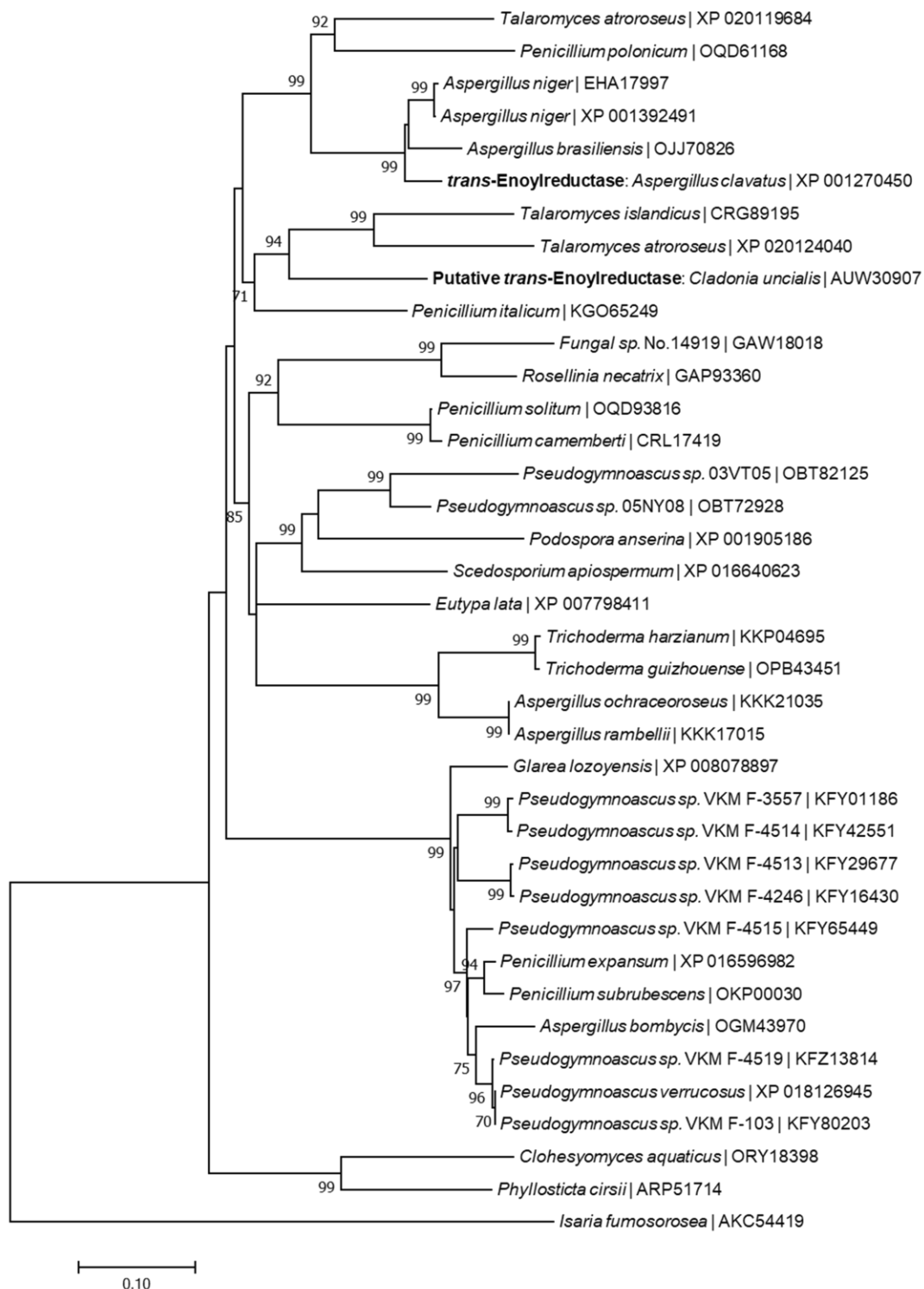


Figure S37: Phylogenetic relationship between a putative *trans*-enoylreductase of *Cladonia uncialis* and a genetically similar gene encoding a *trans*-enoylreductase that is part of the **cytochalasin** biosynthetic gene cluster of *Aspergillus clavatus*. The *trans*-enoylreductase of *Isaria fumosorosea* (AKC54419) was chosen as an out-group. Reproduced from Bertrand et al. (2018b), Supporting Information file, in accordance with authors' retainment of privileges.

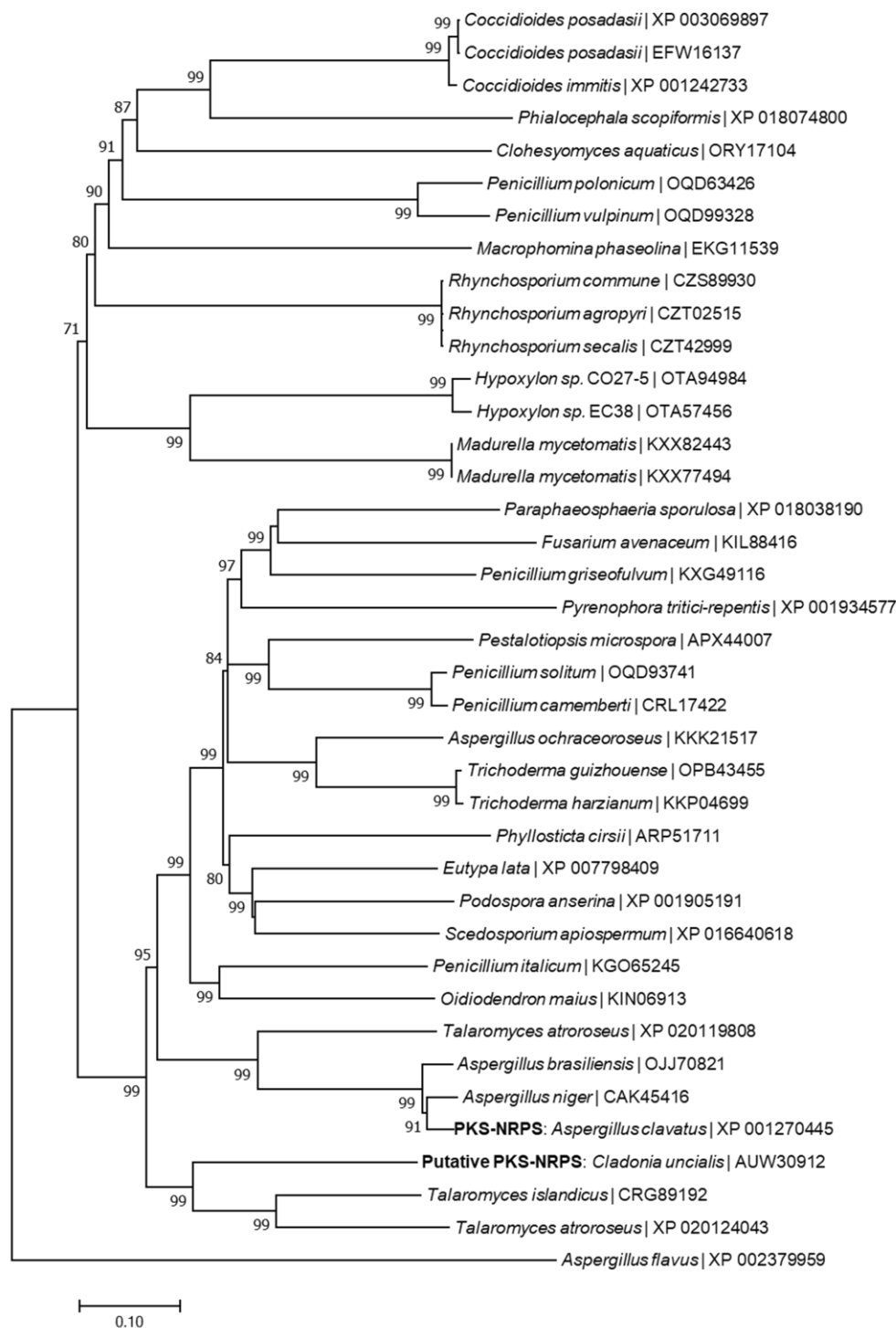


Figure S38: Phylogenetic relationship between a putative **polyketide synthase-non-ribosomal peptide synthetase (NRPS-PKS)** of *Cladonia uncialis* and a genetically similar gene encoding a NRPS-PKS that is part of the **cytochalasin** biosynthetic gene cluster of *Aspergillus clavatus*. The NRPS-PKS of *Aspergillus flavus* (XP_002379959) was chosen as an out-group. Reproduced from Bertrand et al. (2018b), Supporting Information file, in accordance with authors' retainment of privileges.

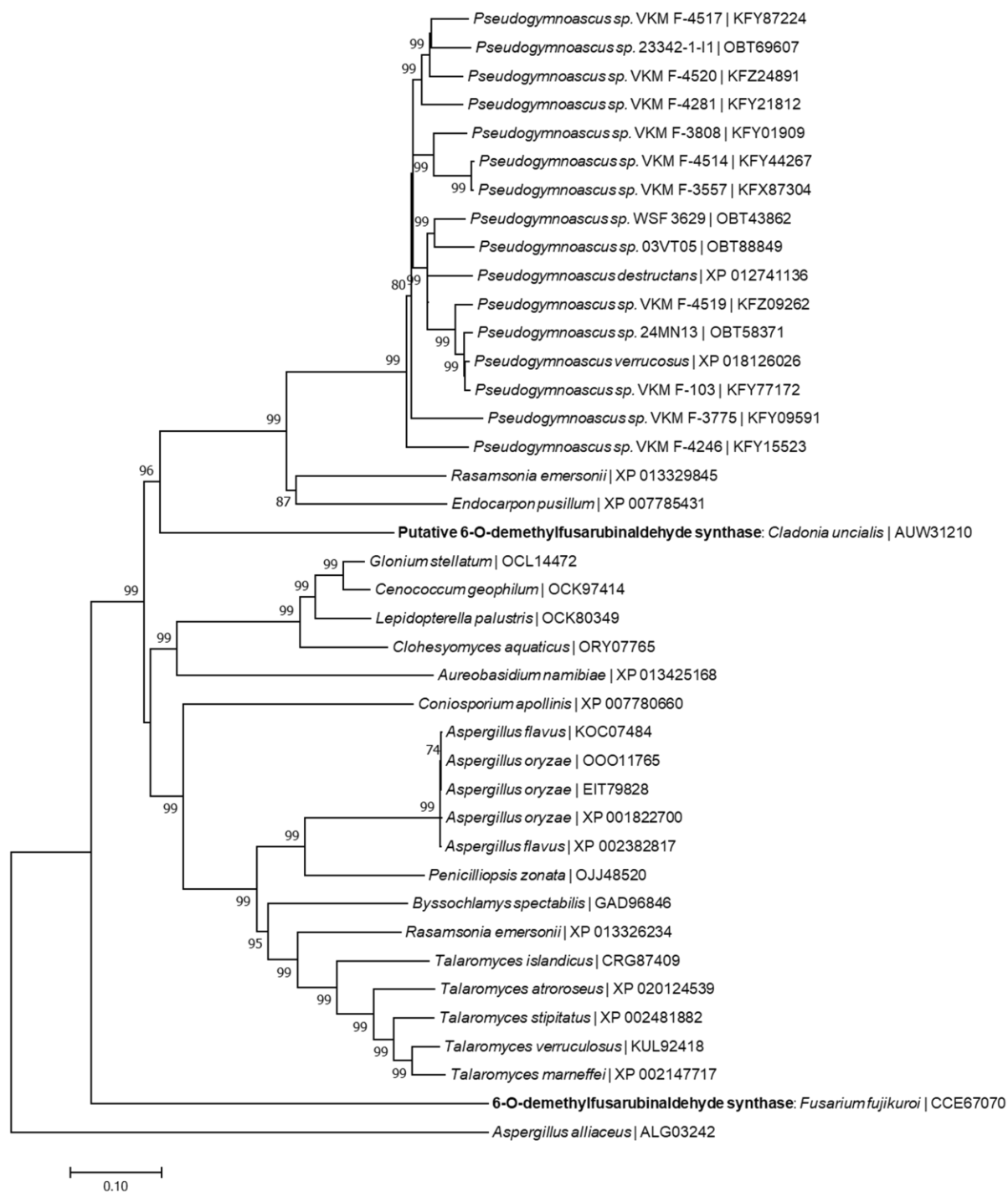


Figure S39: Phylogenetic relationship between a putative **6-O-demethylfusarubinaldehyde synthase** of *Cladonia uncialis* and a genetically similar gene encoding a 6-O-demethylfusarubin-aldehyde synthase that is part of the **fusarubin** biosynthetic gene cluster of *Fusarium fujikuroi*. The PKS of *Aspergillus alliaceus* (ALG03242) was chosen as an out-group. Reproduced from Bertrand et al. (2018b), Supporting Information file, in accordance with authors' retainment of privileges.

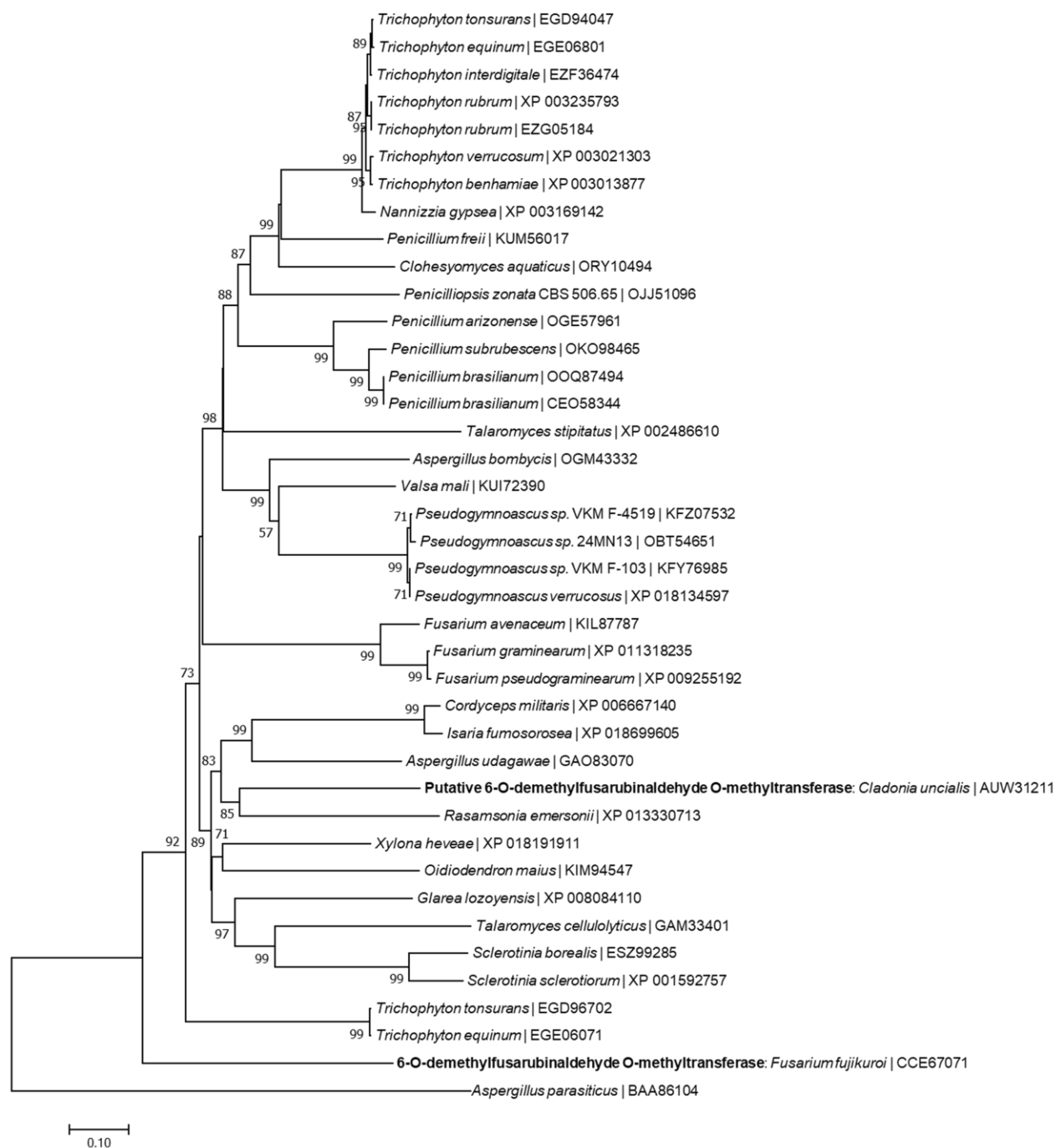


Figure S40: Phylogenetic relationship between a putative **6-O-demethylfusarubinaldehyde-O-methyltransferase** of *Cladonia uncialis* and a genetically similar gene encoding a 6-O-demethylfusarubinaldehyde-O-methyltransferase that is part of the **fusarubin** biosynthetic gene cluster of *Fusarium fujikuroi*. The O-methyltransferase of *Aspergillus parasiticus* (BAA86104) was chosen as an out-group. Reproduced from Bertrand et al. (2018b), Supporting Information file, in accordance with authors' retainment of privileges.

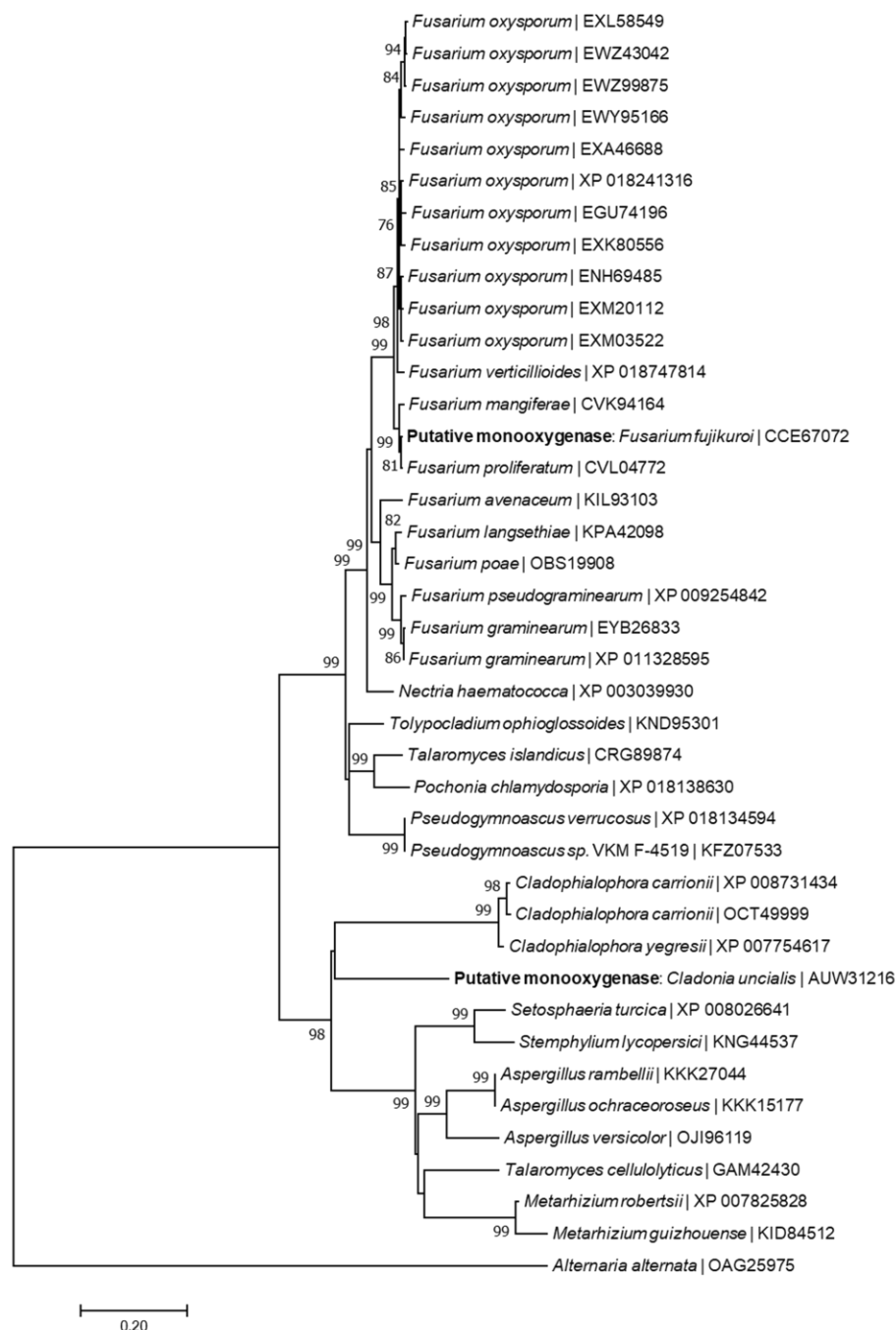


Figure S41: Phylogenetic relationship between a putative **multi-functional monooxygenase** of *Cladonia uncialis* and a genetically similar gene that is proposed to encode a multi-functional monooxygenase that participates in **fusarubin** biosynthesis in *Fusarium fujikuroi*. The monooxygenase of *Alternaria alternata* (OAG25975) was chosen as an out-group. Reproduced from Bertrand et al. (2018b), Supporting Information file, in accordance with authors' retainment of privileges.

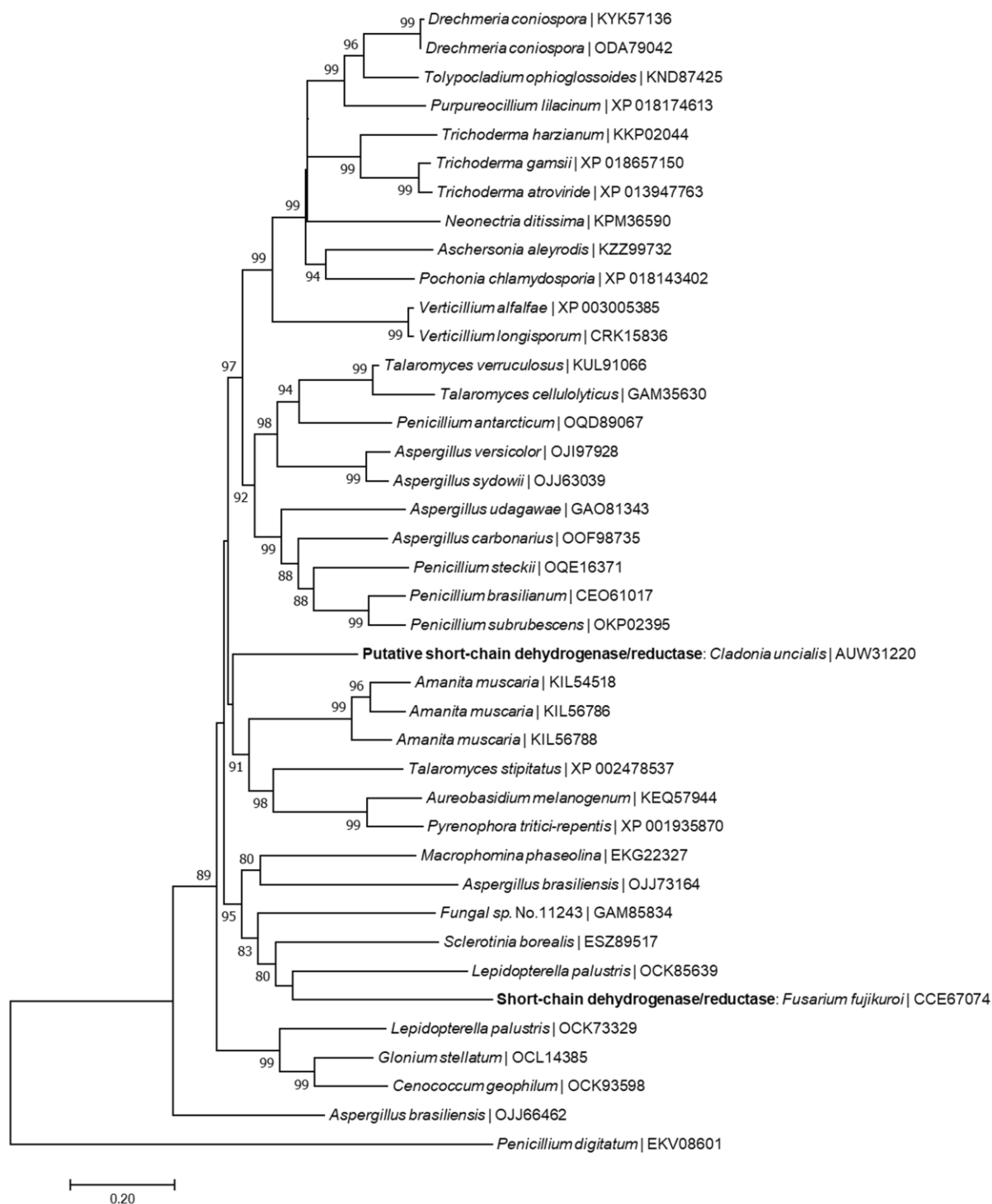


Figure S42: Phylogenetic relationship between a putative **short-chain dehydrogenase / reductase** of *Cladonia uncialis* and a genetically similar gene that is proposed to encode a short-chain dehydrogenase/reductase that participates in **fusarubin** biosynthesis in *Fusarium fujikuroi*. The SDR of *Penicillium digitatum* (EKV08601) was chosen as an out-group. Reproduced from Bertrand et al. (2018b), Supporting Information file, in accordance with authors' retainment of privileges.

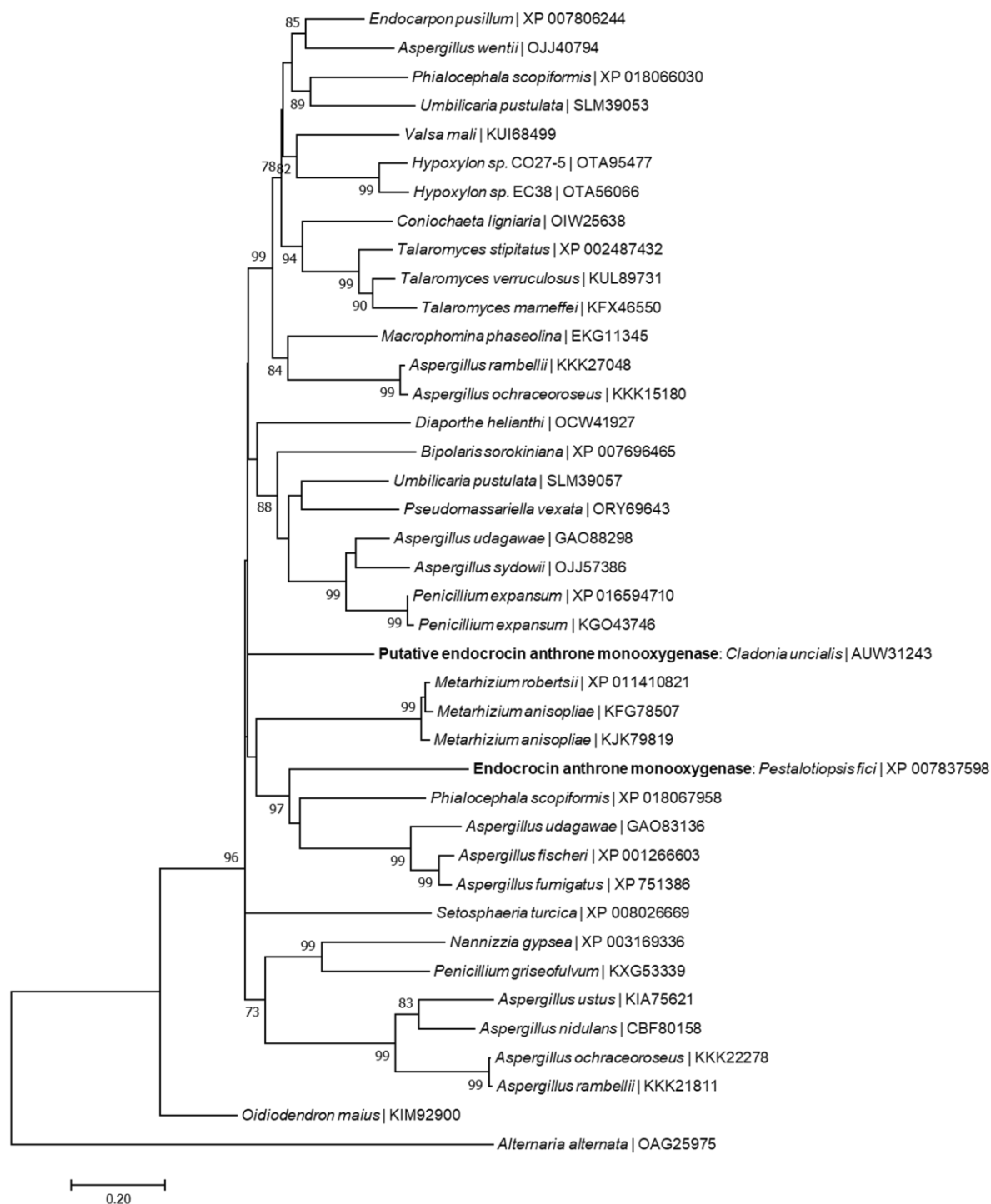


Figure S43: Phylogenetic relationship between a putative **endocrocin anthrone monooxygenase** of *Cladonia uncialis* and a genetically similar gene encoding an endocrocin anthrone monooxygenase that is part of the **pestheic acid** biosynthetic gene cluster of *Pestalotiopsis fici*. The monooxygenase of *Alternaria alternate* (OAG25975) was chosen as an out-group. Reproduced from Bertrand et al. (2018b), Supporting Information file, in accordance with authors' retainment of privileges.

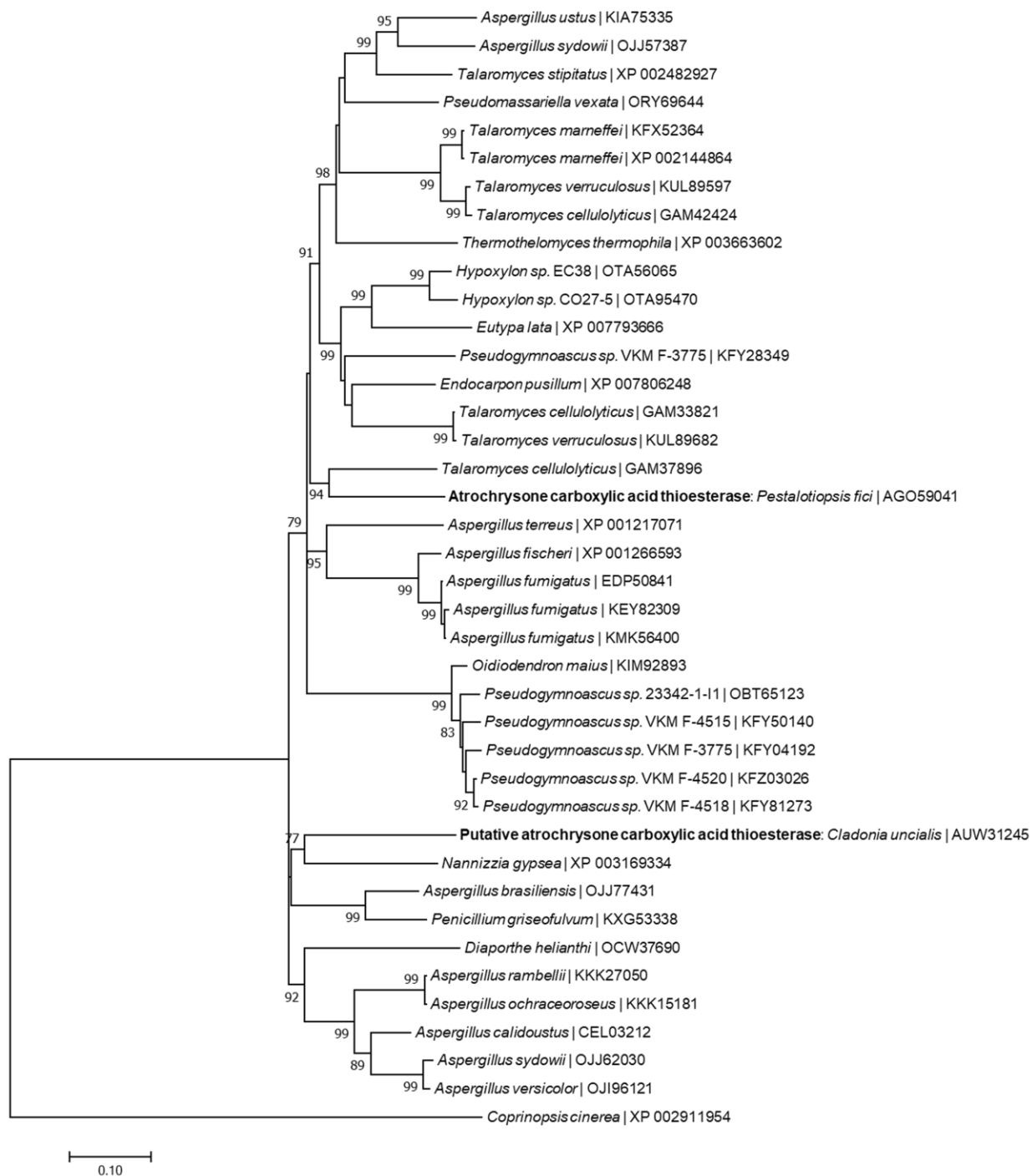


Figure S44: Phylogenetic relationship between a putative **atrochrysone carboxylic acid thioesterase** of *Cladonia uncialis* and a genetically similar gene encoding atrochrysone carboxylic acid thioesterase that is part of the **pestheic acid** biosynthetic gene cluster of *Pestalotiopsis fici*. The metallo- β -lactamase of *Coprinopsis cinerea* (XP_002911954) was chosen as an out-group. Reproduced from Bertrand et al. (2018b), Supporting Information file, in accordance with authors' retainment of privileges.

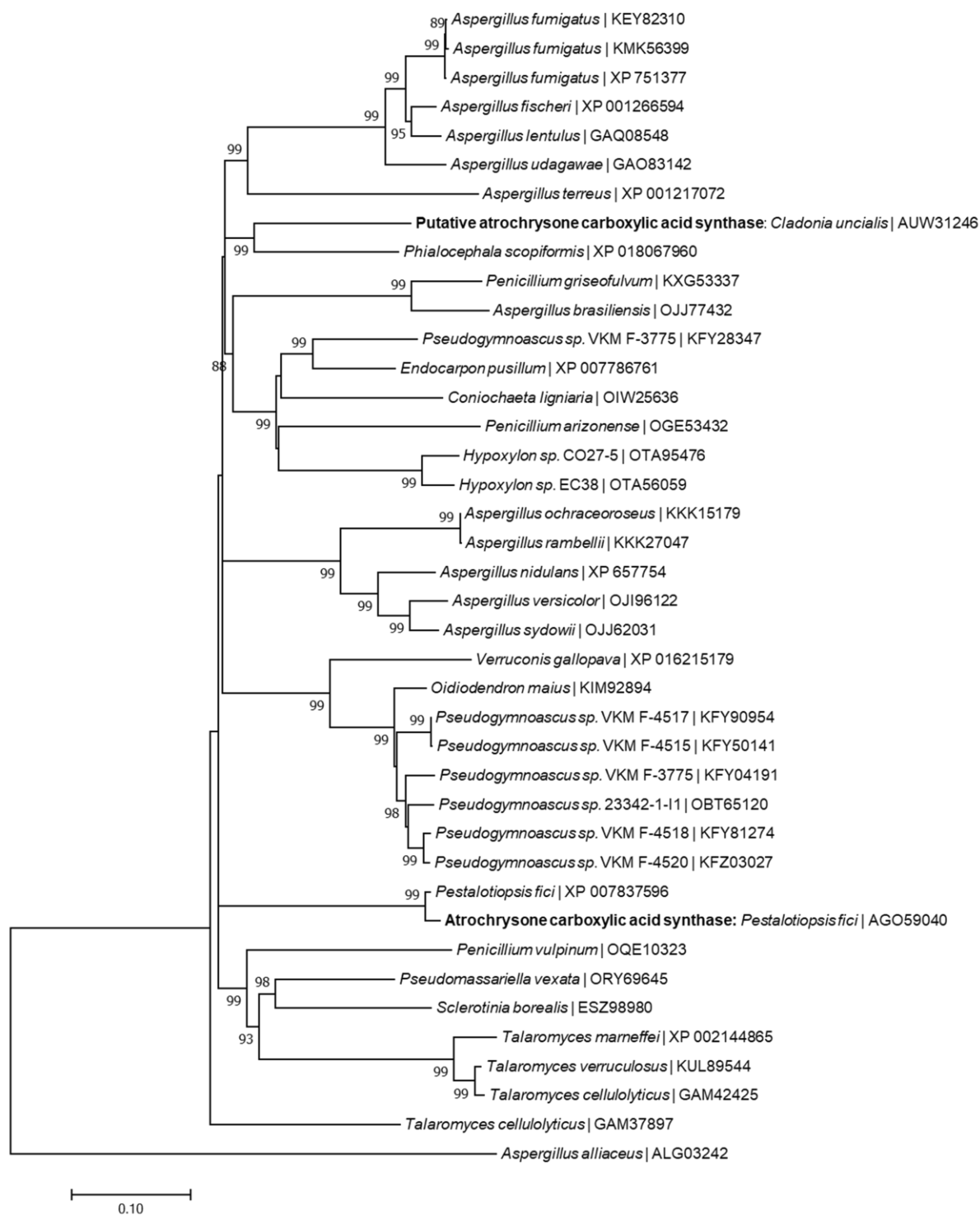


Figure S45: Phylogenetic relationship between a putative **atrochrysone carboxylic acid synthase** of *Cladonia uncialis* and a genetically similar gene encoding atrochrysone carboxylic acid synthase that is part of the **pestheic acid** biosynthetic gene cluster of *Pestalotiopsis fici*. The PKS of *Aspergillus alliaceus* (ALG03242) was chosen as an out-group. Reproduced from Bertrand et al. (2018b), Supporting Information file, in accordance with authors' retainment of privileges.

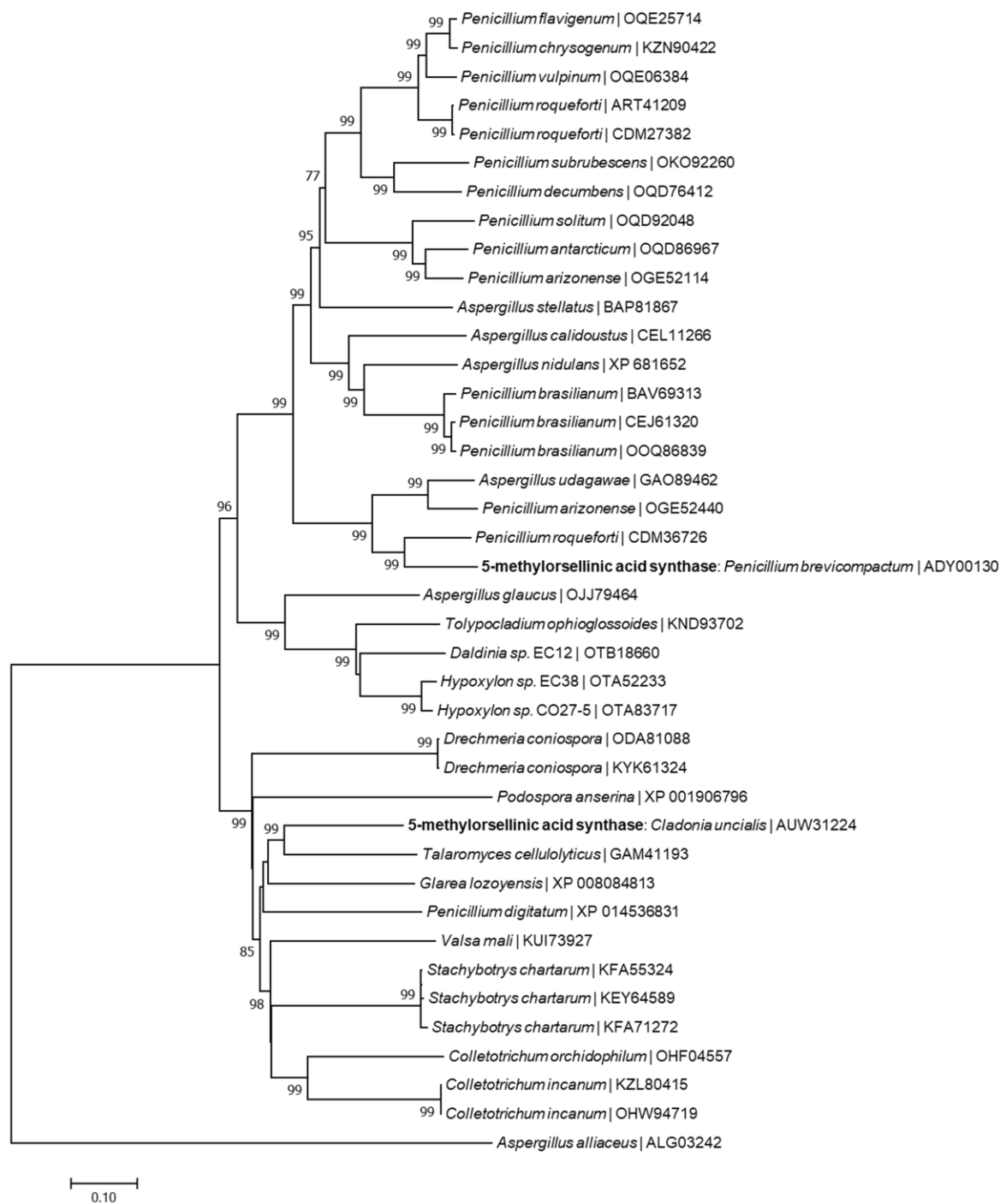


Figure S46: Phylogenetic relationship between a putative **5-methylorsellinic acid synthase** of *Cladonia uncialis* and a genetically similar gene encoding a 5-methylorsellinic acid synthase that is part of the **mycophenolic acid** biosynthetic gene cluster of *Penicillium brevicompactum*. The PKS of *Aspergillus alliaceus* (ALG03242) was chosen as an out-group. Reproduced from Bertrand et al. (2018b), Supporting Information file, in accordance with authors' retainment of privileges.

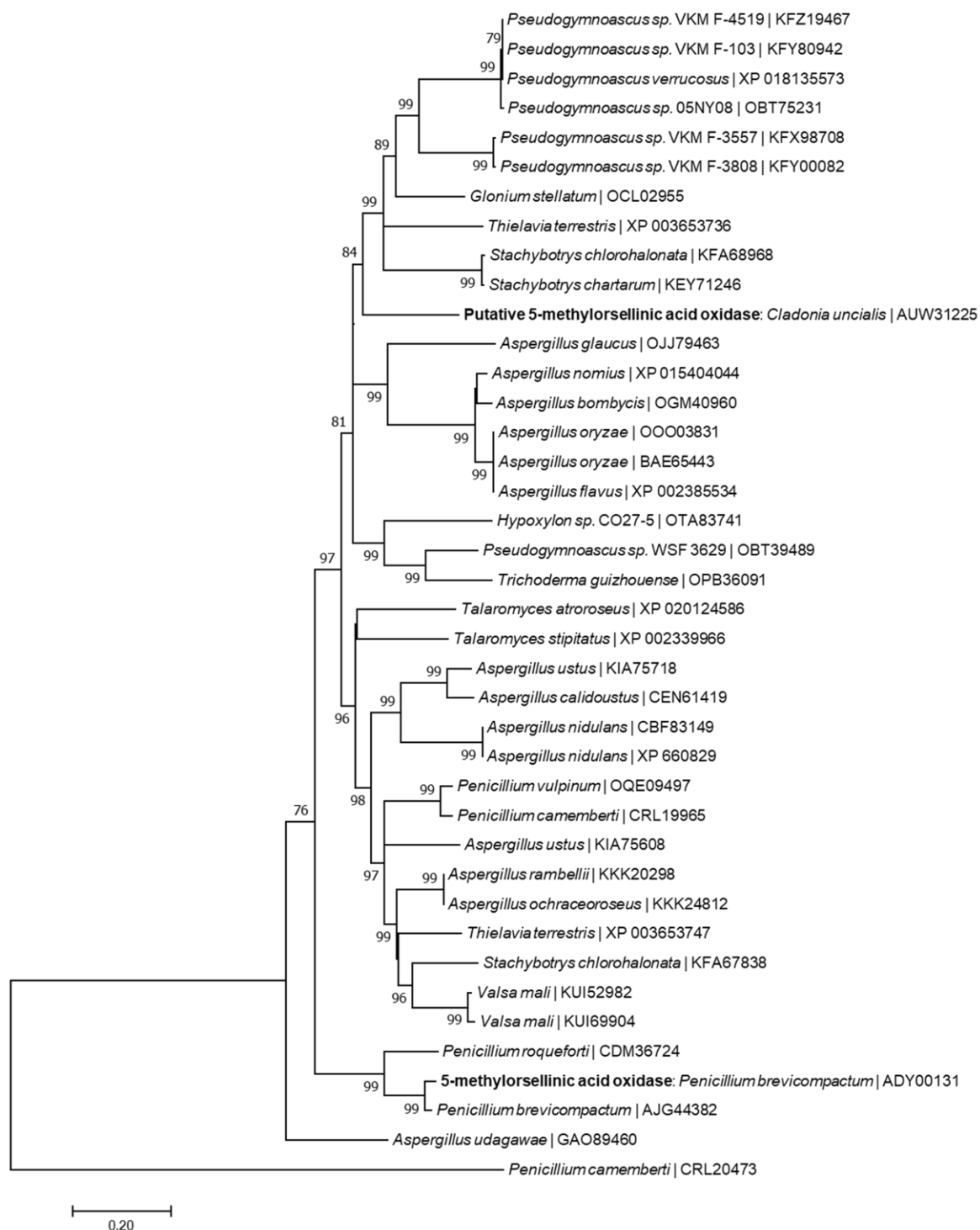


Figure S47: Phylogenetic relationship between a putative **5-methylorsellinic acid oxidase** of *Cladonia uncialis* and a genetically similar gene encoding a 5-methylorsellinic acid oxidase that is part of the **mycophenolic acid** biosynthetic gene cluster of *Penicillium brevicompactum*. The cytochrome p450 of *Penicillium camemberti* (CRL20473) was chosen as an out-group. Reproduced from Bertrand et al. (2018b), Supporting Information file, in accordance with authors' retainment of privileges.

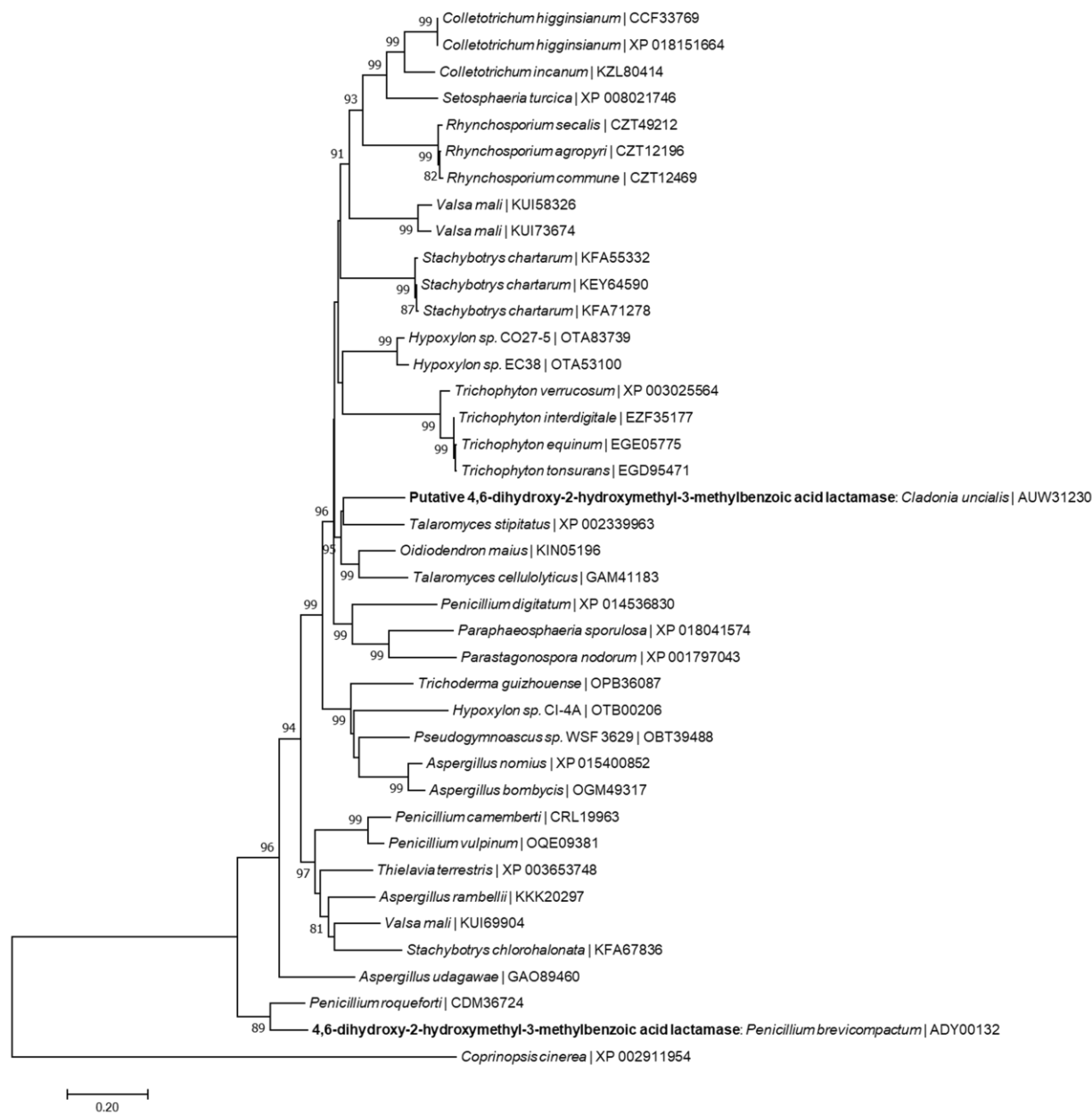


Figure S48: Phylogenetic relationship between a putative **4,6-dihydroxy-2-hydroxymethyl-3-methylbenzoic acid lactamase** of *Cladonia uncialis* and a genetically similar gene encoding a 4,6-dihydroxy-2-hydroxymethyl-3-methylbenzoic acid lactamase that is part of the **mycophenolic acid** biosynthetic gene cluster of *Penicillium brevicompactum*. The lactamase of *Coprinopsis cinerea* (XP_002911954) was chosen as an out-group. Reproduced from Bertrand et al. (2018b), Supporting Information file, in accordance with authors' retainment of privileges.

Chapter 7

Table S54: nBLAST alignment of ITS sequence of unknown species of non-lichenizing fungus identified in this study with entries deposited in GenBank. Reproduced from Bertrand et al. (2019), Supporting Information file, manuscript in preparation.

GenBank Entry	Query coverage / Identity score	Accession number
<i>Penicillium dipodomyicola</i> NRRL 35583 18S ribosomal RNA	94/99	DQ339570
<i>Penicillium griseofulvum</i> NRRL 5256 18S ribosomal RNA	94/99	DQ339557
<i>Penicillium griseofulvum</i> NRRL 35258 18S ribosomal RNA	94/99	DQ339553
<i>Penicillium griseofulvum</i> NRRL 2159A 18S ribosomal RNA	94/99	DQ339551
<i>Penicillium dipodomyicola</i> NRRL 35582 18S ribosomal RNA	94/99	DQ339550
<i>Penicillium griseofulvum</i> NRRL 3523 18S ribosomal RNA	94/99	DQ339549
<i>Penicillium concentricum</i> CBS 285.36 ribosomal RNA	94/99	MH855801
<i>Penicillium concentricum</i> NRRL 2034 18S ribosomal RNA	94/99	DQ339561
<i>Penicillium glandicola</i> NRRL 2036 18S ribosomal RNA	94/99	DQ339565
<i>Penicillium gladioli</i> CBS 278.47 ribosomal RNA	94/99	MH856256

Table S55: nBLAST alignment of DNA sequence of *6msas* gene identified in *Penicillium* sp. with entries in GenBank. Reproduced from Bertrand et al. (2019), Supporting Information file, manuscript in preparation.

GenBank Entry	Query coverage / Identity score	Accession number
<i>Penicillium patulum</i> 6-methylsalicylic acid synthase gene	100/99	X55776
<i>Penicillium griseofulvum</i> NRRL 2159A 6-MSAS gene	99/99	GQ258970
Cloning vector pTac-Pg6MSAS-sfp	98/99	MH488924
Cloning vector pTac-Pg6MSAS	98/99	MH488922
<i>Penicillium expansum</i> NRRL 35695 patulin biosynthetic...	100/86	kF899892
<i>Penicillium expansum</i> acyl transferase/acyl hydrolase...	97/87	XP_016745558
<i>Penicillium chrysogenum</i> Wisconsin 54-1255	100/85	AM920437
<i>Penicillium chrysogenum</i> Wisconsin 54-1255	97/85	XM_002564786
<i>P. freii</i> gene for polyketide synthase	29/85	X95884
<i>Aspergillus aculeatinus</i> CBS 121060 6-MSAS	23/74	XM_025653259

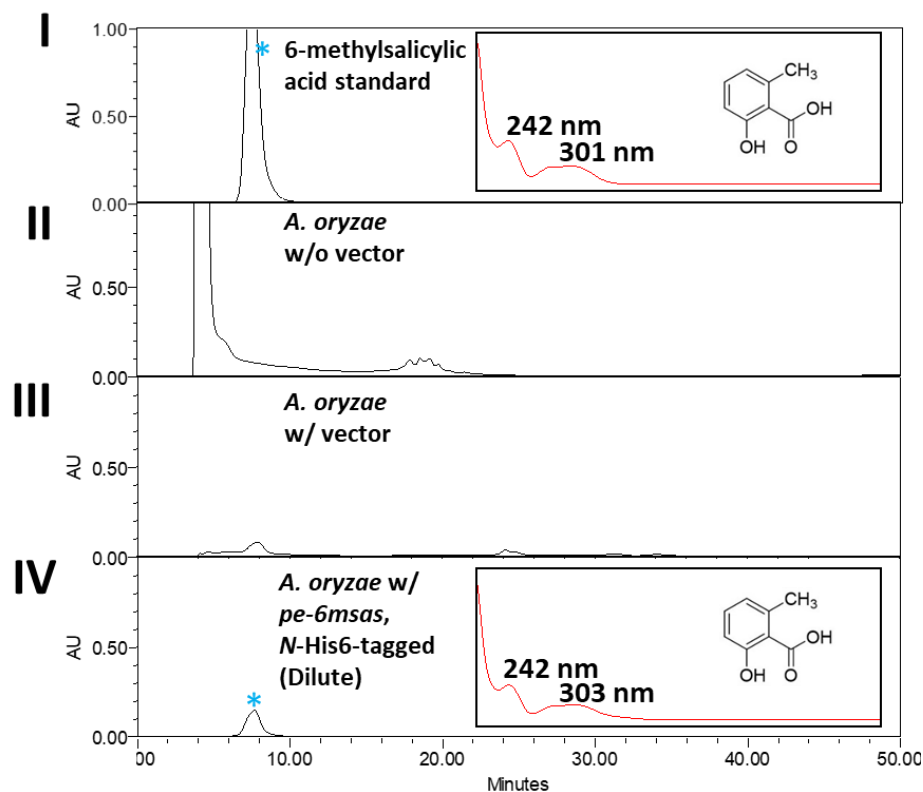


Figure S49: HPLC analyses of control experiment to verify activity of His(6)-tagged PE-6MSAS in *A. oryzae*. I) 6-methylsalicylic acid standard. II) Extract of untransformed *A. oryzae*. All such traces generated in this work are shown in **Figure S53**, below. III) Extract of *A. oryzae* transformed with plasmids that do not contain exogenous genes. All such traces generated in this work are shown in **Figure S54**, below. IV) Diluted extract of *A. oryzae* transformed with *N*-terminal His(6)-tagged *pe-6msas* in *A. oryzae*. Traces are shown at 300 nm.

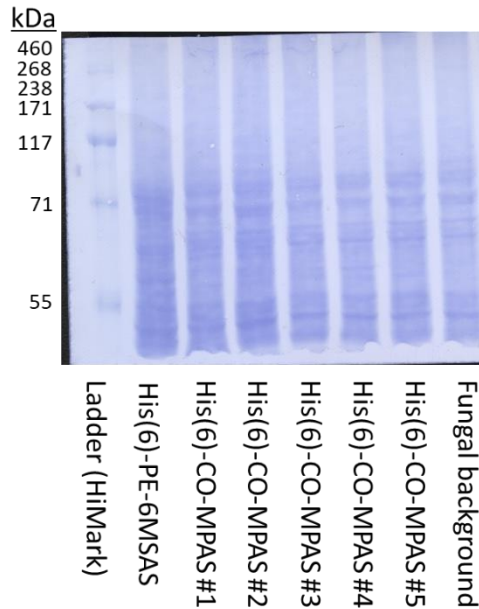


Figure S50: Coomassie stain of Western blot membrane, performed after blot imaging, to confirm effective transfer of high molecular weight proteins. Protein samples are acquired from *A. oryzae* transformed with plasmids encoding His(6)-tagged variants of *pe-6msas* and codon-optimized *mpas*. Fungal background is *A. oryzae* transformed with plasmid vectors (not containing a gene). The band labeled 171 (kDa) in the HiMark ladder is responsive to His(6) antibody and was used as a blotting positive control. Reproduced from Bertrand et al. (2019), Supporting Information file, manuscript in preparation.

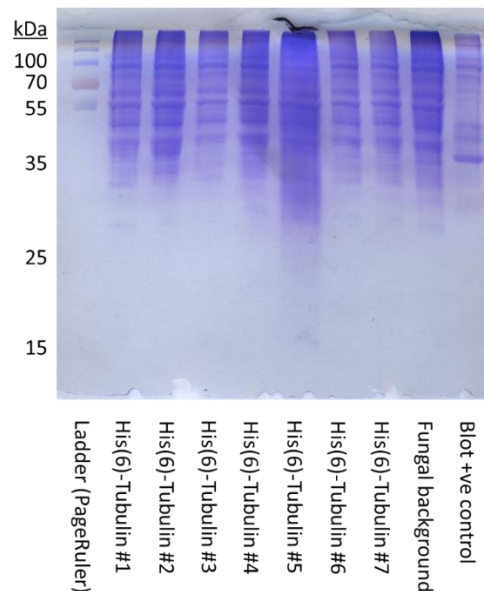


Figure S51: Coomassie stain of SDSPAGE gel, performed after transfer to Western blot membrane, to verify that a suitable concentration of protein was loaded into the gel. Proteins acquired from protein extracts of *A. oryzae* transformed with plasmids encoding His(6)-tagged β -tubulin from *A. oryzae*. Fungal background is *A. oryzae* transformed with plasmid vectors (not containing a gene). Blot +ve control is a crude protein extract of *E. coli* expressing a His(6)-tagged protein with an approximate molecular weight of 25 kDa. Low molecular weight proteins are not visible on this gel because these proteins were completely transferred to the blotting membrane. Reproduced from Bertrand et al. (2019), Supporting Information file, manuscript in preparation.

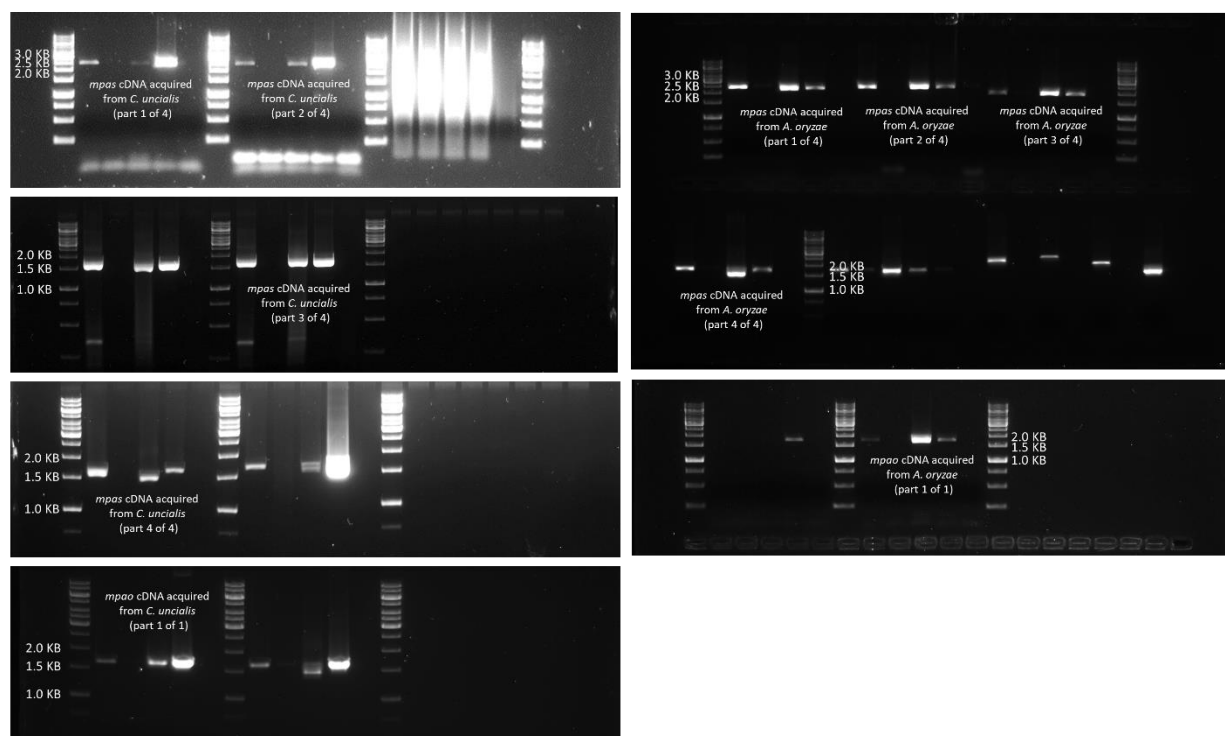


Figure S52: Uncropped version of **Figure 68** (see main text). Agarose gels of cDNA of *mpas* and *mpao* extracted from (L) *C. uncialis* and (R) *mpas*- and *mpao*-transformed *A. oryzae*. Unlabelled portions of agarose gels are experiments unrelated to **Figure 68**.

Sequence S1: DNA sequence of *mpas* (5' – 3'). DNA highlighted in blue are the exonic portions. Exonic regions are observable by performing RT-PCR on mRNA isolated from *C. uncialis*. A parallel experiment performed on mRNA from *A. oryzae* that was transformed with *mpas* revealed identical exonic regions. Accession number: AUW31052. Reproduced from Bertrand et al. (2019), Supporting Information file, manuscript in preparation.

ATGGCATTACCCTCGCTGATCGCATTGGCGCTTTAGCACCTTGCCAG
CATCAGATCGCCTGGATCAGCTTCGAAACGCTCTGCAGCACCACAACCTCCTTGAAACCAATCACAAAGGCAATTCAGGA
GCTTCCGTTGCTCTGGAAGGCTCTTTCCAACCAGGATCAGAGTTTACATAGCATCGCCGGCGAAGCGGCTGCAGACCAG
TTAGCGCAATGGATAAGCGGAGCTGGTACCGCTCAGCTTGTGGACGACAAGGTAATGTCAACAAGATGCCCTTAACAA
CCATTGCCAGATTGCGCAATATGTAGCTATCTGTGTCATATGAGGAGCCCCCTCGTCACGAATCAATCATAAAAAG
CGCTGCGATCGGCGGAGGTATCCAAGGGTTTTGTATTGGCCTCCTAAGCGCTCTTCAGTGGCAAGCGGAAAGACGGAA
GACGATGTTGGCAATTTGCGAGCTATGTCCGTGCGATTAGCCTTTTGGCTCGGGGCTACGTTGATCTTGATCGACATC
GCAACGGTGGCGACTCCAAAGCATCGACTATCGCCGTGAGGTGGAAGACGCCAACGACATTGGAGGATATTCAGCGCTT
GCTTTCCAGACATCCGGATGTGAGTCTCATTATAATATTTTCATATCTCGATACTTCATGCCAGCATCATCACCTGTTGC
TGGAAGTCGGATCTCTAGGCATACGGTGCCGTTCCGTCCTTGAGTTGTGGTAAGAGGCTAGCTACTAATACTACGACAGC
AGACATATATCGCTGTGGTCCGAGATATCAGGGATGTGACAATAACGGTGCCAGCTTCTGTCATGGAGCATTGTGATTGA
GGATTTGTACAGATAGGCGCGTCTTGCCTGACACTGGTGTCTCTGGACGGTACCACGTCGCCATCCATGAAGGAATT
CCGCAAAAAATCTTGGAGACGTGCCAAGCACAATTCAGTCCGACCATCAATGGGCAACCCCTCGTGAGGTCCAACACTG
ATGCACATCTGTTCTCTGGTGAGGACACCGCTCTGTTGGCTCTGGAGTGCATATTGGGAGAACGTGCAGACTGGTACTC
GACCATATCTACAGCCGCTTCTGCATTGAACCAATCAGCGCCAATCCTTTTCATCTCTCCATAGGTACCGATCCAGTG
CCGCAATCAGTCGCAAGGAGCTTTCCAGTGGTGAAAGCGACAATGATTGCCGACCGAGTCAACGGTATCGTAGAACCGG
AGATACCTGCCCTTGACACGACCCCTTTGGATCGGTTTCTCAAGGATATCCAAGGATGCCATCGCGATCATTGGTAT
GGGTTGTAGATTCCCGGGTGTGATTCCATTGACGAGTATTGGAAATTTGCTCACTGAGGGAAGTCCATGCTATCAGAA
ATTCGGAAGCGAGGTTTGGACGAGGTGCGCCTGCCAGATCCAACAGCAGCCTACGATTCTGGGGCAACTTTCTGCGGG
ACATCGAAGCTTTTGACCATGGATTTTCAAAAAGTCAACACGGGAAGCCGTTTCTATGGATCCTCAGCAGCGTGTGCT
TCTCCAGGTCGCTTATGAAGCTTTGGAGTCGTCTGGGTATTTTGTGATTTCATCGAGGCCTGAGGACGTTGGCTGCTAT
ATTGGAGCTTGTGCAACAGACTACGATTTCAATGTGGCATCCACCCCTCCATCGGCATACTCAGCAATTGGCACCCCTAC
GATCTTTCTAAGTGGCAAGCTGTGCGATTACTTTGGCTGGTCTGGTCCCTCTCTTGTCTGGATACCGCCTGCTCTTC
GTCGGCGGTGCGCATCCATACAGCATGTACTGCTTTGAGGACTGGTCAATGTTCTCAGGCCTTAGCGGGCGGGATCACC
TTGATGACAAGCCCTATCTCTACGAAAACCTCGCCGCTGCTCACTTCTTGAGTCCAACCTGGAAGCTCAAAGCCGTTCA

GCGCAGATGCGGATGGGTATTGTGCGAGGAGAAGGTGGTGGACTCGTTGTTTTGAAGCGACTCTCAGACGCTCTGAGGGA
TAATGACCATATTTCTCGGTGTCATCGCCGGCTCGCGGGTCAATCAGAATGACAAC TGCGTGCCAA TCACCGTTCCTCAT
ACCTCGTCTCAGGGAAATCTGTATGAACGAGTTACTGAACAGGCAGGGGTGAGACCTAGCGAGGTCACCTTTGTAGAAG
CCCACGGTACTGGCACACCAGTTGGCGACCCCATCGAGATGGAGAGCATCCGTCGTGTTTTTGGTGGACTACATCGAGT
TGCTCCATTGATTGTATCGTCAGCCAAAGGAAATATTGGACATCTTGAAGGTGCCCTCTGGGGTGGCAGCATTAATAAAA
GCGCTCTTGAGATGGAACATCATCTTGCGCCCCGCCAAGCGTCATTCAAGACCTTAAATCCAAAAATACCGGCACTTG
AGCCAGATAATCTTTGCATACCGACGTCAAATCTTGCAATTGTGCGGAGAACGATTAGCCGCTTGCAATCAATAACTACGG
TGCCGCTGGCAGTAATGCCGCTATGATTGTACTTGAACCTCCTCGAAAGAGTGTACCTACCACGACAAATCCAAGATG
AGCATCTCTTCCCAGCTAAGATACATCCAATTAGCTGGCGGCCGCGTCAC TAGGGGGTCTTCTTGCACTGTGTAG
CCCTTGACCAAGTATTGCCAGCGATTGCGGTTTACTCAAGACACATCCGAACAGCCGCAAGTATTATCAGACTTGGCCTA
TAGTCTGTCAACAAGGCTTAACCAGGAGTTCGGTTCACGTTGACAATGACTGTACAGATCTGGATCAGCTACAAGCT
CAGCTTCGTCAACAGACTGTACAGGTAACAATATCAAGAACGCTTCAAAAGCACCCCTGTAGTACTGTGTTTTGGGG
GCCAAGTATCCGACAGAGTTGCACTTGACAAGTGTGTTGGCAAGAATCAACGTTGCTGCGTTCTACTTGACATCTG
CGACAACACCCTACGCGTATTGGGATATCCCGGGCTCTATCCTAGCATTTTTTCAAATGAGGCCGTGACTGATGTTGTC
CTATTACACTCCATGATCTTCGCTTGCAATATTCTGTGCGCAGGCATGGTTAGAATCTGGCCTGAAGGTGGATGCTC
TCGTTGGGCACAGCTTTGGTCAGCTTACTGCGCTGTGTGTCTCCGGGATATTGTCCCTGCGGGACGGTCTGAGACTGGT
GGCTGGAAGAGCATCTTTGATGCAAAAGCATTGGGGCCCCGAGTCAGGTAATGATTGCTATCGAGACAGACCAGCAG
ACCCTAGAAGAGCTTCAAAAGGTTATCTGTGAATCCAACGCCAGCTACAACCTTTGAGATCGCCTGTTTTAACGGGCCAA
CATCCCACGTTGTGGTGAGCGACAGATGTTCTGTAGCGAGCTCGAAACGAAGTTGATGGAGCGGGCAATAAGACACAG
GAGCCTGGACGTTCCCTACGGCTTCCACTCTCGGTTACCGAACCCCTTGTTACCACATCTCGAGGATTGGCATCATCT
CTGACCTTCCATGAGCCTAAGATCCCCTTAGAAACCTGTACAGACATGGGCACTTGACAGAGCCCACTTCAAGCTCA
TTGCTGCTCATACAGGGAGCCAGTTTTCTTTGGTAAGGCGATTGAGCGTTTACAGGCTAGATTAGGTCCATGTACTTG
GCTCGAAGCCGGTTGACACTCGAGTATAGTCAATATGGTGCGACGGGCTCTAGGGCAAGCCTCCGCTACGGCCAATAAC
TTTGCTCTCTTGCAACTAAACAAGCCTAACTCTAGTAAGCTTGTTGTCGATGCAACAGTCGCTCTTTGGGACGCCAGCC
ATCGGGCACAGTTTTTGAATTTTCATAGACTGCAAAGCGCCCAATATGATCATCTACGCTTGCCCTCTTATGCGTGGA
GAAGTCTAAGCACTGGCTGGAATTGGACATGTCTGCAGCGCTGAATTCCGACAAAACAAATACGCCTCCCCAACAAAT
ACAGCAGCGCAAGTGGAGCTACCCGCTGTGCTCATACGTCGAAATCTTTTGACTCGCAGGGACATCATTTTCGTATCA
ATCCGAGTAGCGAGGAATATCAGACTATCGTCAAAGATCTTGAGTCGTTGGGTAGCGCTGTGTGCCCTCTACATTATA
TG TAGAGCTGGCATCCAGAGCCGT CAGGGTGGCAGAGGAGGACAAAGGAAACGGTTTGCTCTCAATAAAAGATCTCAGA
GTGCACTCTCTTCTCGGAGTAAACGTACATCAGACCATCAGCCCTTGATCTCCAACGGCTTGCTCAAAGCTGGAGATTTA
GAATCAACATGCGGACGGGTCAATTAGTGGTTCAAACCCCGAGAATCGTTCTGTATGCTGAGGGTACAGTCAATTT
GAAAGTAGCGGACGATAGTCTGGAGGAAGAGTTCTGTGATATGAGAGACTAACAGGTCAACAACAGATCATCTCTATC
GCCGATGATCCTCGAAGTGAATCGCTCCGAGGCAACGTGCTATATAATATGCTTGCCGAGTTGTCAATTATCCGGACT
GGTACCGAGGAGTTAAGAGTGTGGCTGCACCTTGACCTGCGGGTTGTGCCAAGGTTACATGCCAGTGGGCATCCCGGA
GATTTGTTTCAAAGAAAGCACGACTCAACTACCCATCCTGGAGAGCTTCATCCAATCGCTAGCTTACATGCAAATGT
TTGCAAGAGTGTGCGGGCGGCGAAGTGTTCAGTTCACTAGAGCAGACCATATGCAGTGGGCACCGGGCTTTGACTTGC
ATGGCTATGGAGACTCTGCAGAAGCATCATGGGACGTGTTAGCTTACAATTCTACAAATGCCGAAAACGTCGTGATCGA
CATATTTGTCCACGATGCTGAGTCACGGGTAGACTGCTAGCTTACTGGGAGCCAACTTTACAGATATCTTCGCCGACCA
GTCCCCATCTCGGCTGGCTTGAATACTAGTTTGGCGTCTGAGAAGGACATACCTATGCTCAAGAACGCCAATGCAGAGA
GAGCCGAAATTAGCCTAAACAGCCAATTACCAGCAGAATCACATTACAGGCCAACCTGACCAGGCCAGGTAAGGATGC
GAAAACGAGCATTTACGAAGACATCTGTGGCCTTCTGGAGAAGCTGGCAGACATACCTGGAGACCAAGTATCCGGTGAA
GCTACTTTTCGACAATCTTGGGGTGGATTGCTTGATGATGATTGAAGTTATTAGCGAGCTTTGACTCTGTTCCGGGTG
ATCTACCAATCCACGAACTGGAGGAGCTCACCAGACATTAACCTCGCTCGTCGATTATCTGCATGAAAAAGGCTGCGTAGG
TAGCTTATATGTAGAGGACAGTGCAATGCATCATCCCTAAGCTCCAGCCACGCCATATCCACTGGGGCAAGCTCTCCG
CCGACTCATCGGGGGCGAGTGCAATGACAACGCCACCTGAGACGCTATCGCTTGTTGGACTATCCTGTTCACTACCA
CCAACAAGAATCACGCGCAGCGCTGCGATATCAAATGGAACAGGCAGGCAACCGTTGGATATGGGCCATATGGCAT
TCAACAGGTTTTTACACGCTGCGGTTGATTTTCAAAAAGTATGCTGAACAGACAGGGGCTAAAGGATTCTGGACAAAC
GTTTTACCCCGAGCAGGCAGACCTTGTTGTCGCGTACGTCGTGGACGCTTATCGCAAGCTAGGCTGCGACCTGGCAACTC
TTGCCCGCGGTGAGCAACTACCATCGATGAACACATTACCCAGGCACAAGCACCTTGTTGCTCAGCTGCGCAATATCCT
GGTGGAAAGCGGGCTCCTGGAGCTCCGAGGCAATCAAGTACATGTTAGAACTGCCAAGACGGTTGACTCGACACCGAGC
GCCATTGATACGAGCAAAATGTTGACGCGACACCCCTTTGGAGCTTCCGAGACCAAGCTTTTGAACGTGACCGCCCCC
GTCTTGCTGACTGCTTAACCGGACAAAAGGAGCCTTTGTCAATTGCTTTTTCGGCGACAAGCACAAACCGCGATCTCTAGC
CGACTTCTACGCCAACTCGCAGATGTTAAAGGCAGCCACGCGCTCCTTGCAAGATTCGTGAGTTCTACATTCTCGGCC
GCGCAATCTGGAGATACTCTCTGCATCCTAGAGTGGGTGCTGGCACGGGAGGCACCACTAGGTATCTCGTGGACGTTT
TGAACCGTTGCGGTATACCTATGAGTATACGTTACGGATATTTACAGTCACCTGTTACTCAAGCCAAGAGGAACTT
CGCAAGTCTCCACAGATGAGATTGATGACATTTGACTGCGACCGGCCGCGCGCAAGAGCTCCTAGGCAAAATCCAT
ATTGTTATCTCGACCAATTGTATTCATGCTACAAGCAACATCACAACATCTACCACAAACATTTCTGCCACACTTCGTG
ACGACGGCGTCTCTGCTTACCGAGAACCTTTACTGTTGATCTAGTCTTTGGTCTTCTGGAGGGTTG
GTGGCTCTTACGCGATGGTGTGTCAGCACGCCCTGGCAAAATGAATGGTTTTGGGATCGCAGCCTCCGCGCGGCTGGGTT
AAGCATGTATCTTGGACAGATGGCAACACTGAGGAGGCAAGACTCTGCGGTTGATCTGCGCTTTAGGGGCGAGGCCA
AAGAAGATCGTAATCTGGCCGACCAAAATGGAGCTATTACGAAGAGACCGCGGTACCGATGGAGGAAGTAGTGTGGA
ACGTGTGCGGTACGCTTGACCTCTCCGCTGATATATACTTCCGGAAGACACCTGATCCGCCAGGAAAGAAAAGACCCATA
GGTACGATCAACTCTCAACTGTTCTGACTCGCCTGCGGCTAACATGACACAGCTCTCCTCATACATGGCGGAGGGCAT

TTCTTGTGGACGAAAAGATGTGCCTATGAAGCACATTCGAACCTCTGATTGAGCGTGGCTTTCTCCAATAAGCACTG
ATTATCGACTTTGTCCAGAAACGAATCTTTTTGAGGGCCCAATGACGGATTGTTGCGACGCTCTGAAATGGGCGACAGA
GACGTTGCCAACTCTGCCGCTGTCGGGACCTACCGTTAGGCCCGATCCAACCAAGTGGTGTGCGTTGGTTGGTCTAGC
GGTGGACAGCTGGCCATGAGCTTAGGATACACGGCTCCAGTCAAAGGGATCAAAGCACCTGACGCCATATTCGCACTGT
ATCCGCTAGTGATATGGAATCTAACCGTACGTGTTGGCCTGAATCGACTCCATACGCGAGGTGGAGCCTGCTGACATT
TGAAACCAGACTGGCATCAACCGTGCTATCCTCTAGCTGCTGAAGAAGAGCCAACAGAGATCCTGGACATTCTGGCAGG
CGTTCCGAGAAAGTCCGGTAAGTATCTTCTGCTCGGCAGCACCCAGCGTGACAGAGGTTGACAATCCATTTTAGATTGT
GGAATATGCTCCTGTGAGTGAAGAGAACCATGGCTCTTTTCGCTGACTTTGAAAGACGACCGCGCCAGTATCATTTTG
CACATGTGCTGGAATACAAACAGTGCCCATACTTGTCCATGGTCTACCATACAGAAGAACCTTCCAGACACAGACA
AGACCGATTGGAATATAGGCCACCTGCGTCCGCTGAGCAGGTCCAAGCCATCAGCCCGCTGTGGCAGATCAGGCAGGG
TAATTACAAGACGCTACCTTTATAGTCCACGGCAACGGTGACGATTGGCTTCTCTTTTCGATGAGTGAGCGGACAGTT
GAAGAGTTGAAACGCTCGCGGCATACCCGCTAGTCTAGCTGTTCCAGAACAGTGTGGGCATGCCTTCGACCTATTTCCCG
TAGGGGATCCGCTAGGTGTTGGTTGGACATCACTTGAGCAGGGTTACGATTTTCATCTGCCGTCAATTAGGAATGAGTTA
G

Sequence S2: DNA sequence of *mpao* (5' – 3'). DNA highlighted in blue are the exonic portions. Exonic regions are observable by performing RT-PCR on mRNA isolated from *C. uncialis*. A parallel experiment performed on mRNA from *A. oryzae* that was transformed with *mpao* revealed identical exonic regions. Accession number: AUW31051. Reproduced from Bertrand et al. (2019), Supporting Information file, manuscript in preparation.

ATGATTTACCAGTGTCTTCTATTCTAGCATCCATTTGGGATAACAGCAAATTGCTTTTGGACCACAGTCGGTTTAAAGCATTGC
CTTGATCGGTGTTGCATGCGCGATATCTATCCGATCGATATTATATGTAAGGCTTGCCGTGCGCCAATTACTCAATACTACTGCTGA
CGCATGCCAGAGACTCCGGCTCGCATACTCAACGCCGCTCCGCCAGTGCCAGGACCATGGTATGCCAAATTTACCGCCCTGGGC
CTAAGGGCGAATGACGTTGCTGGCAACCGATGGTACTATGTCCAGGGATTGCATAAGAAGTATGGCTCAATCGTTTCGCATCGCGCC
GGAGGAAGTGGCCATCAGCGATCCCAAGGTAGTCAGCAAAGTCCACGCGCTCGGCACCGAATTCAGAAAACGCCAACAGCCCGGCA
CACCATTCAACATTTTCAGTATTAGTGATCCTAAGGCACACCGGACTCGGCAGCGATTCTACGCAAAAGCCTTCTCCGACGAGACT
TTGAAGGCGAGCACTGAGCCAGCAGTTCGACAACCTATTAACACAGCTGTGGCAAGCATCAAACGAGATGCGGCGCTAAGAAAAGA
CCACACAGCCGATGTGTACAAGTGGTGCATGCTTTTCGGGAGCGATGTGCGCTTCCAGGTGATCTACGGCAACAGCAACACAGAGG
GGCTGATGGCAACCCAGAAGACGACTGATGAGGTGATCATGGGTGCGTATCTGCAGAGAATGAACGCGTGGGCGCAATTCTGCTTC
CCCGTGTTCCTGCTCGGTGCGTGGCTTTACCTCTTTCCCAACTCTGCACAACATTTTTCGCGTCGAGGAGAAGTATGGGGACTT
TTGGCAGGAGGGTCAGCGACAACGCGAGATTGCAGCCAGGACTGTCTTCGTCCAGAATACCAAGTACTCGAAAAATGATGGTGTCT
TTAGCGTAAGTGACGAAGTTGAAGTTGAGTGACGTGGACATTGCTCATGATATCACCACATTCCCTTGGTGGGGTGGGGAGCCCTGA
GGTGGCTCTCTGGTCTTTTGTATCTGGCAGGTCTTCGCGATGCCAGATCTGCAGAGGGAGCTTGAAGCGGAGGTGAGGCGCTCAC
CGAGCCAATCACGGATGCCACAACCGCGCAGCTACCTATCCTGAACGGGGTTATATACGAGACTTTACGTCTTTACGGTGGAGGCG
TCACGCAAATGCCACGATACGCTCCTATTGCTACCGAGCTGGGCGCTACGTCAATCCACCAGGCACGGCAGTGACGACGCATACC
GGCGCTTTCGATCGCAATCCGGCAGCGTGGGATGACCCAGAGAAATGAGTCTGTTTCCAGGCCTCAGCATTCATTTCTATTGACGT
ATATTGGGGCAGTTTGTATCATACAGTTGGCTGTGCGCGTCGCGAGCAAACTTTGGTCCAGGATAAGTTCCAGCCGTTTTCCT
CTGGTGGCGGGGCTTGCATCGCAATCCACCTTGTATGATTGAATTGCGTGTCTTCGTTTCTGTCTTCTTTAGAGAATGTGCCGGC
GTCAAGTTGGCACCCTCGACTACAGATGCGAGCATGAGGTCCTTGATCGCTTCCATATCAGCCCGAACGCTAAGCGTTGCGAGGT
CATAGTTCCGGATGAAAAGGCTTGA

Sequence S3: Results of amplification and sequencing of ITS region of fungus that we have identified as member of the *Penicillium* genus (5' – 3'). This sequence is intentionally unhighlighted.

NNNNNNNNCGGGNGCGGNNNNCCCCATCCCNACGGGATTCTCACCTCTATGACGTCCCGTTCCAGGGCACTTAGAT
GGGGACCGCTCCCGAAGCATCCTCTACAAATTACAATGCGGACCCCGAAGGAGCCAGCTTTCAAATTTGAGCTCTTG
CCGCTTCACTCGCCGTTACTGGGGCAATCCCTGTTGGTTTTCTTTTCTCCGCTTATTGATATGCTTAAGTTTCAGCGG
GTATCCCTACCTGATCCGAGGTCAACCTGGATAAAAAATTTGGGTTGATCGGCAAGCGCCGGCCGGGCTACAGAGCG
GGTGACAAAGCCCCATACGCTCGAGGACCGGACGCGGTGCCGCCGCTGCCTTTTCGGGCCCGTCCCCCGGAATCGGAG
GACGGGGCCCCAACACACAAGCCGTGCTTGAGGGCAGCAATGACGCTCGGACAGGCATGCCCCCGGAATACCAGGGG
GCGCAATGTGCGTTCAAAGACTCGATGATTCACTGAATTTGCAATTCACATTACGTATCGCATTTTCGCTGCGTTCTT
CATCGATGCCGGAACCAAGAGATCCGTTGTTGAAAGTTTTAAATAATTTATATTTTCACTCAGACTACAATCTTCAG
ACAGAGTTCGAGGGTGTCTTCGGCGGGCGCGGGCCCGGGGGCGTAAGCCCCCGGCGGCCAGTTAAGGCGGGCCCGC
CGAAGCAACAAGGTAAATAAACACGGGTGGGAGGTTGGACCCAGAGGGCCCTCACTCAGTAATGATCCTTCCGCAGG
TTCACCTACGAAACCTTGTACGACTTTTACTTCTCTAAATGACCGAGTTTGACCAAGTTTTCGGCTCTGGGGGG
TCGTTGCCAACCTCCTAAGCCAATCCCAAGGCCTCACTGAGCCATTCAATCGGTAGTAGCGACGGGCGGNGTGTAC
AAAGGGCAGGGACGTAATCGGCACGAGCTGATGACTCGTGCCTACTAGGCATTCTCGTTGAAGAGCAATAATTGCA
ATGCTCNNTCCCCNNCCCGANNGGGGTAAACAAAAAANNNNNNNNNNNNNNNNNNNNN

Chapter 8

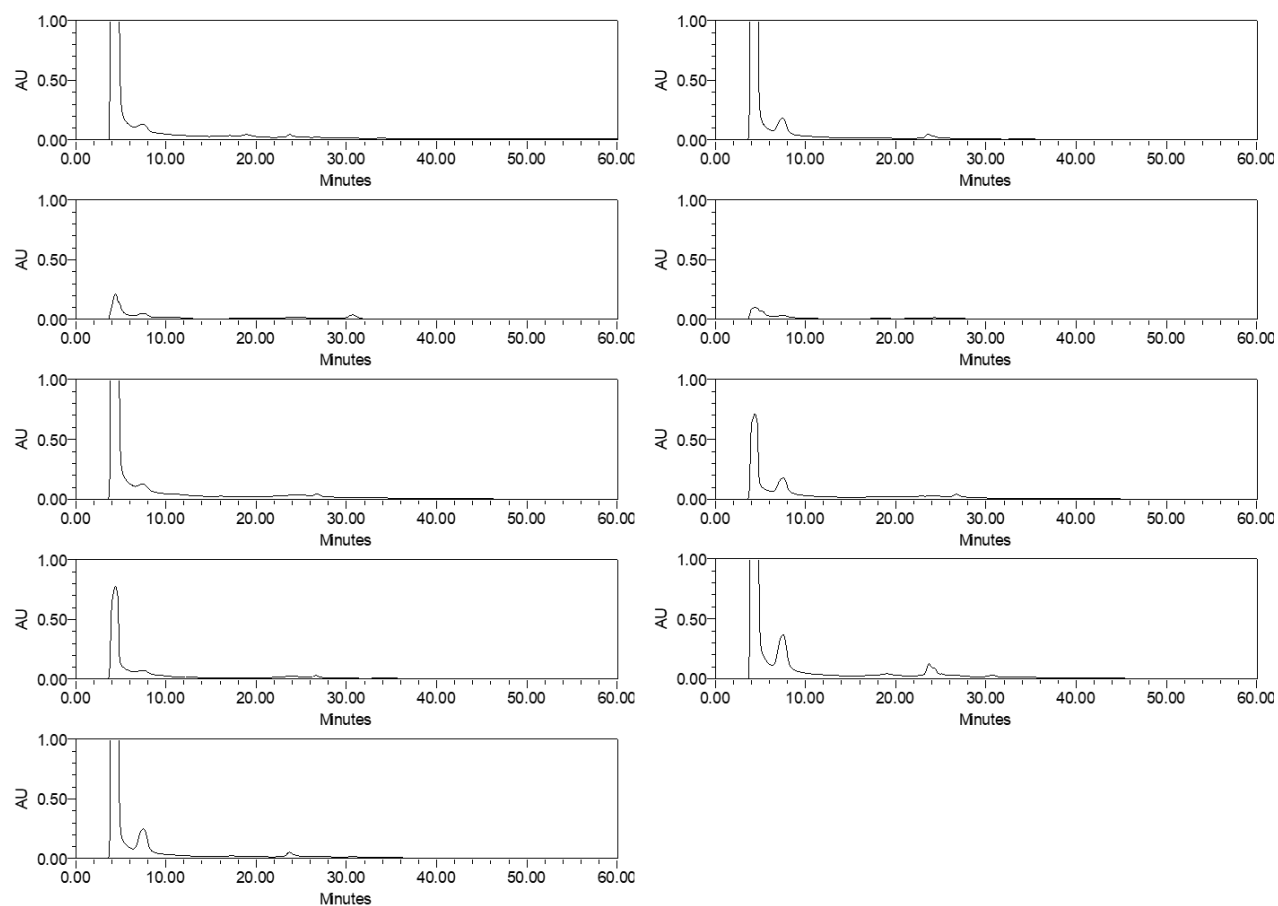


Figure S53: HPLC traces of a total of nine replicates of organic extracts of NSAR1 *A. oryzae* cultures (not transformed with plasmids), following incubation in Czapek-Dox (with added starch) for five days. Traces are displayed at 300 nm. Reproduced from Bertrand et al. (2019), Supporting Information file, manuscript in preparation.

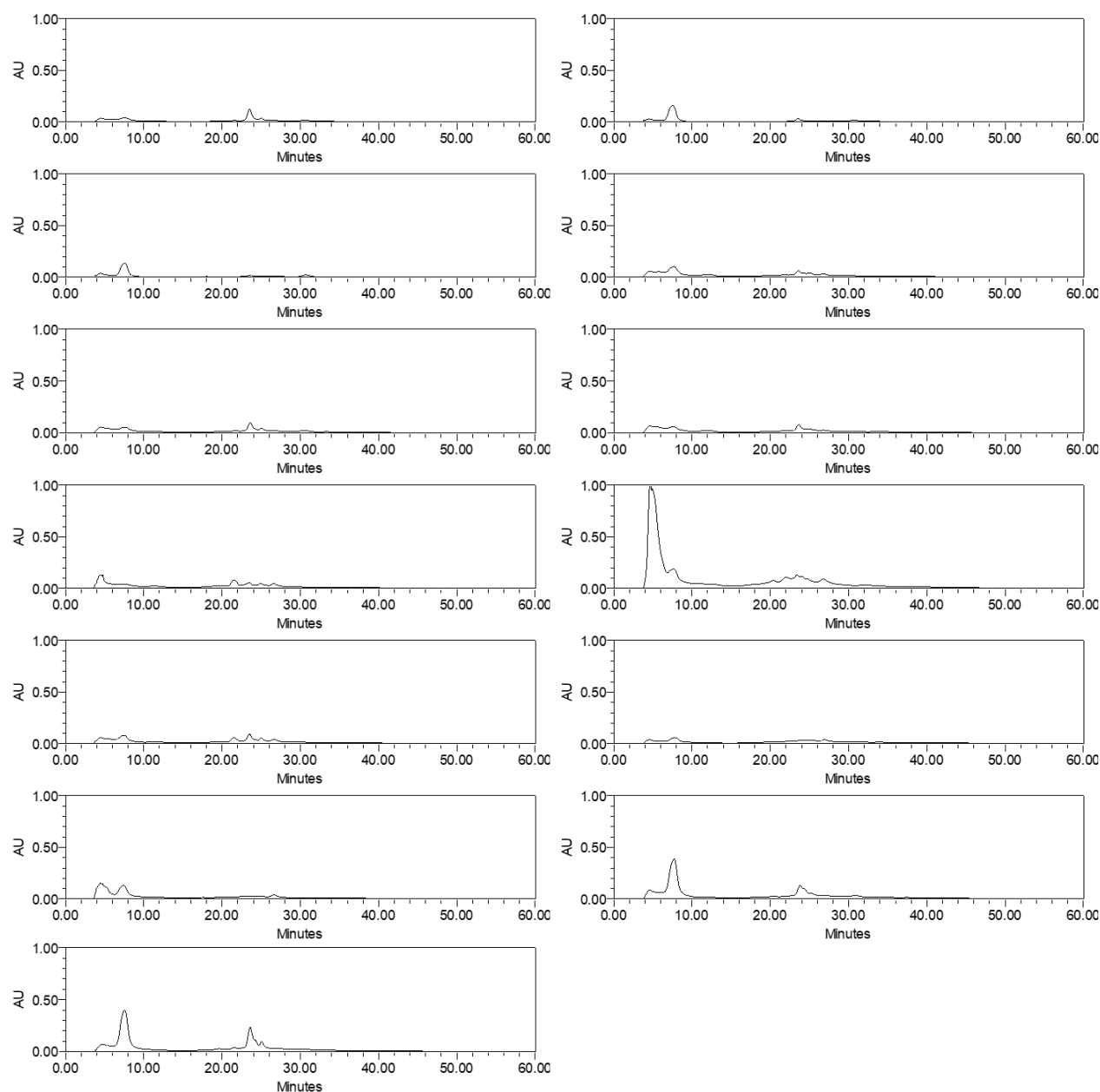


Figure S54: HPLC traces of organic extracts of 11 *A. oryzae* colonies that were transformed with plasmid vectors, following incubation in Czapek-Dox (with added starch) for five days. Traces are displayed at 300 nm. Reproduced from Bertrand et al. (2019), Supporting Information file, manuscript in preparation.

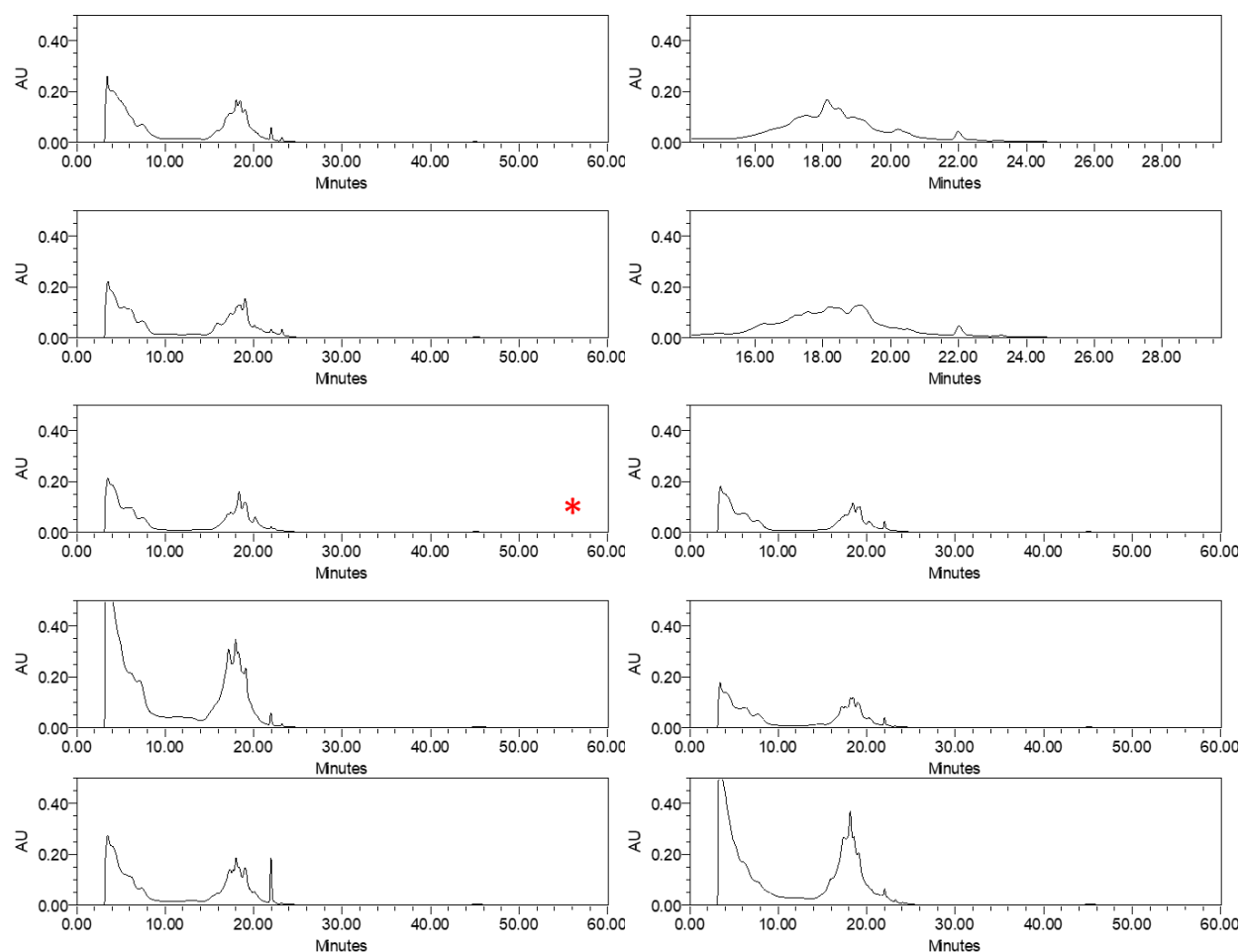


Figure S55: HPLC traces of organic extracts of ten *A. oryzae* colonies that were transformed with codon-optimized *mpas*, codon-optimized *mpao*, and *pptase* (from *C. uncialis*), following incubation in starch-based Czapek-Dox media for five days. Red asterisk indicates the trace that was selected for the main text. Traces are displayed at 280 nm. Reproduced from Bertrand et al. (2019), Supporting Information file, manuscript in preparation.

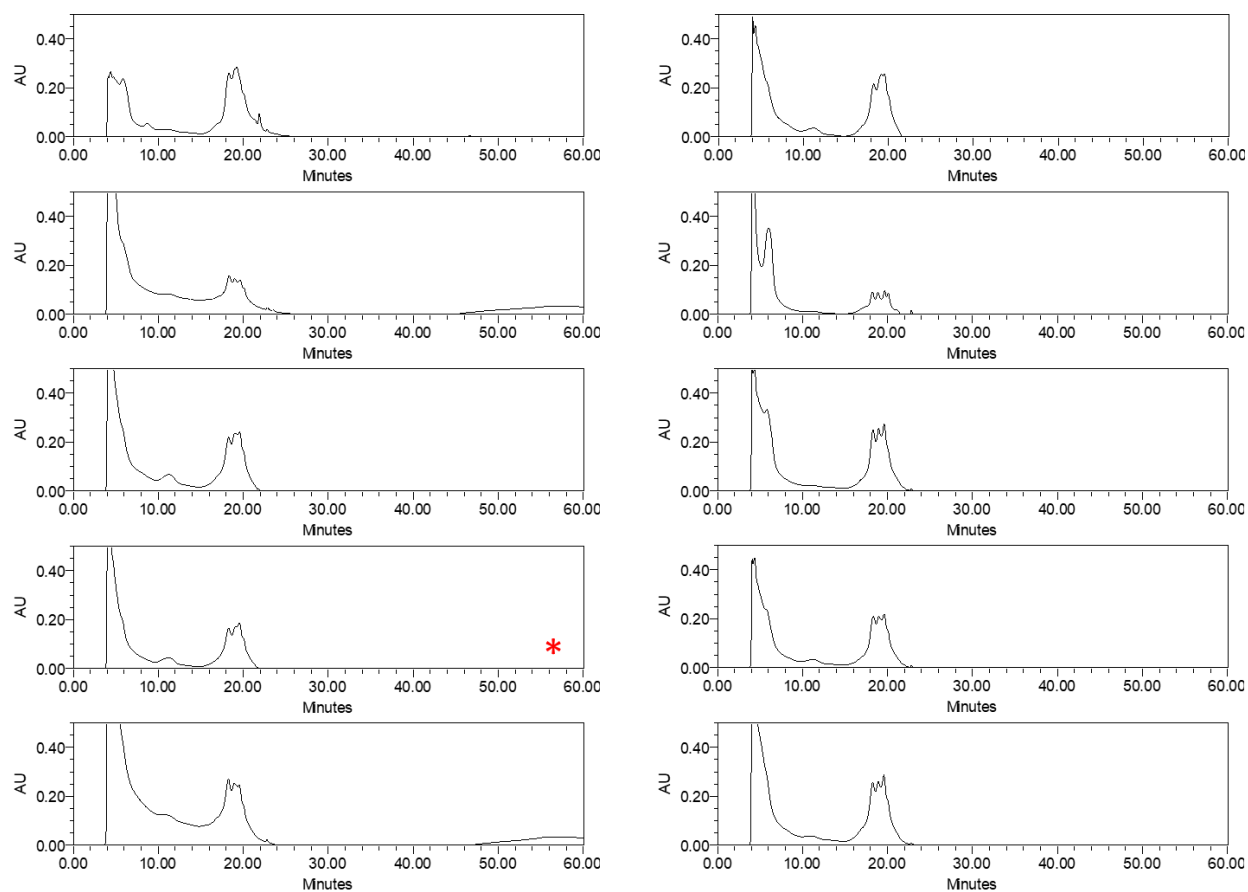


Figure S56: HPLC traces of organic extracts of ten *A. oryzae* colonies that were transformed with codon-optimized *mpas*, codon-optimized *mpao*, and ‘cassette’ (cassette comprises *pptase*, *cpr1*, *cpr2*, *b5r1*, and *b5r2*, all originating from *C. uncialis*) following incubation in Czapek-Dox (with added starch) for five days. Red asterisk indicates the trace that selected for the main text. Traces are displayed at 280 nm. Reproduced from Bertrand et al. (2019), Supporting Information file, manuscript in preparation.

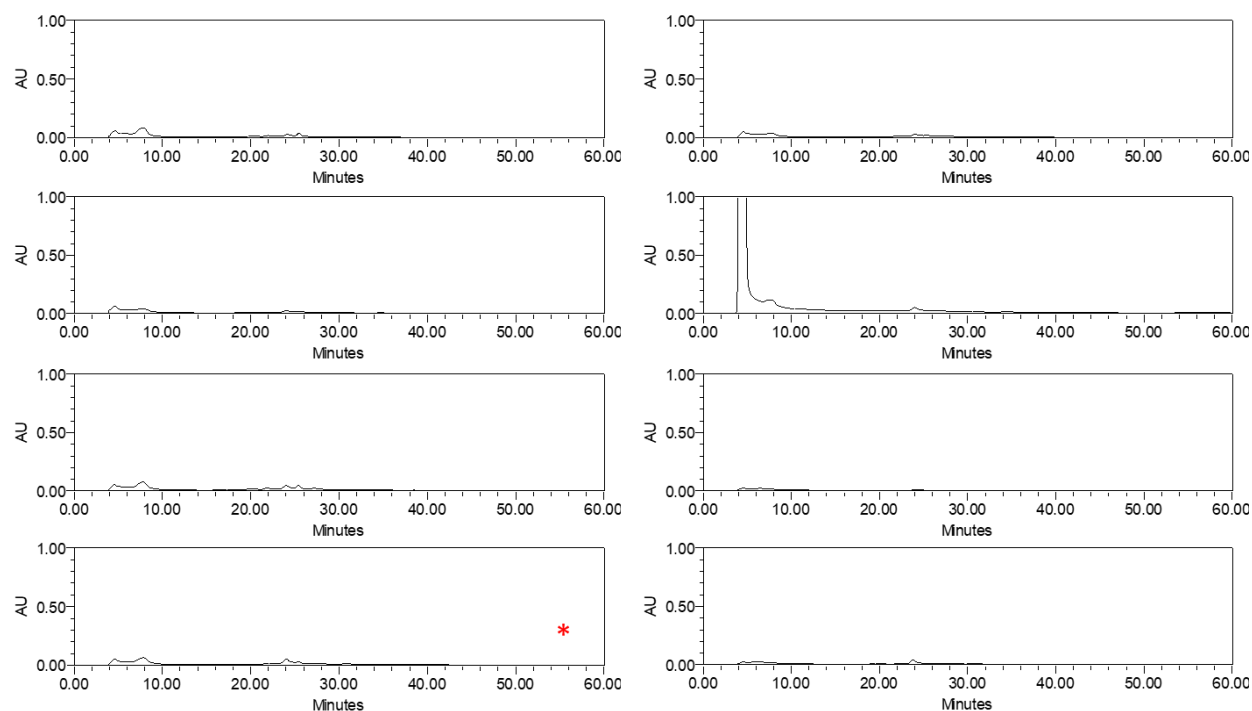


Figure S57: HPLC traces of organic extracts of eight *A. oryzae* colonies that were transformed with *6hms*, *krdh*, and *pptase* (all genes originating from *C. uncialis*), following incubation in Czapek-Dox (with added starch) for five days. Red asterisk indicates the trace that selected for the main text. Traces are displayed at 300 nm. Reproduced from Bertrand et al. (2019), Supporting Information file, manuscript in preparation.

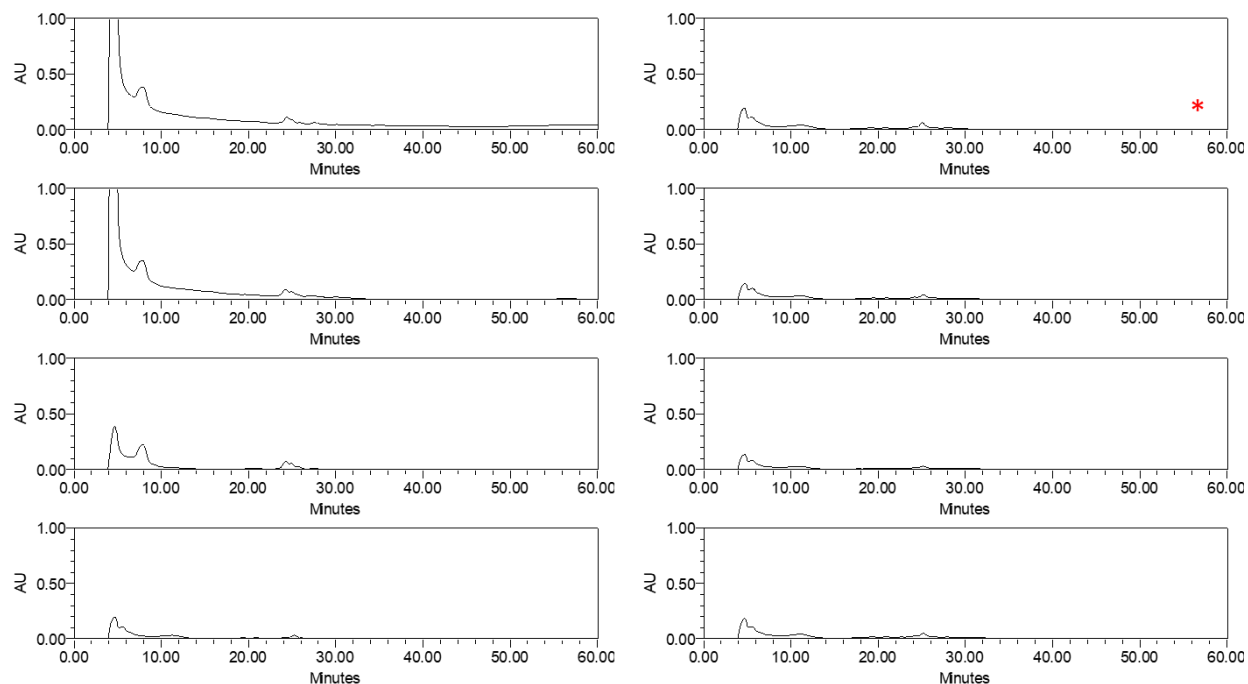


Figure S58: HPLC traces of organic extracts of eight *A. oryzae* colonies that were transformed with *oas* and *pptase* (both genes originating from *C. uncialis*), following incubation in Czapek-Dox (with added starch) for five days. Red asterisk indicates the trace that selected for the main text. Traces are displayed at 260 nm. Reproduced from Bertrand et al. (2019), Supporting Information file, manuscript in preparation.

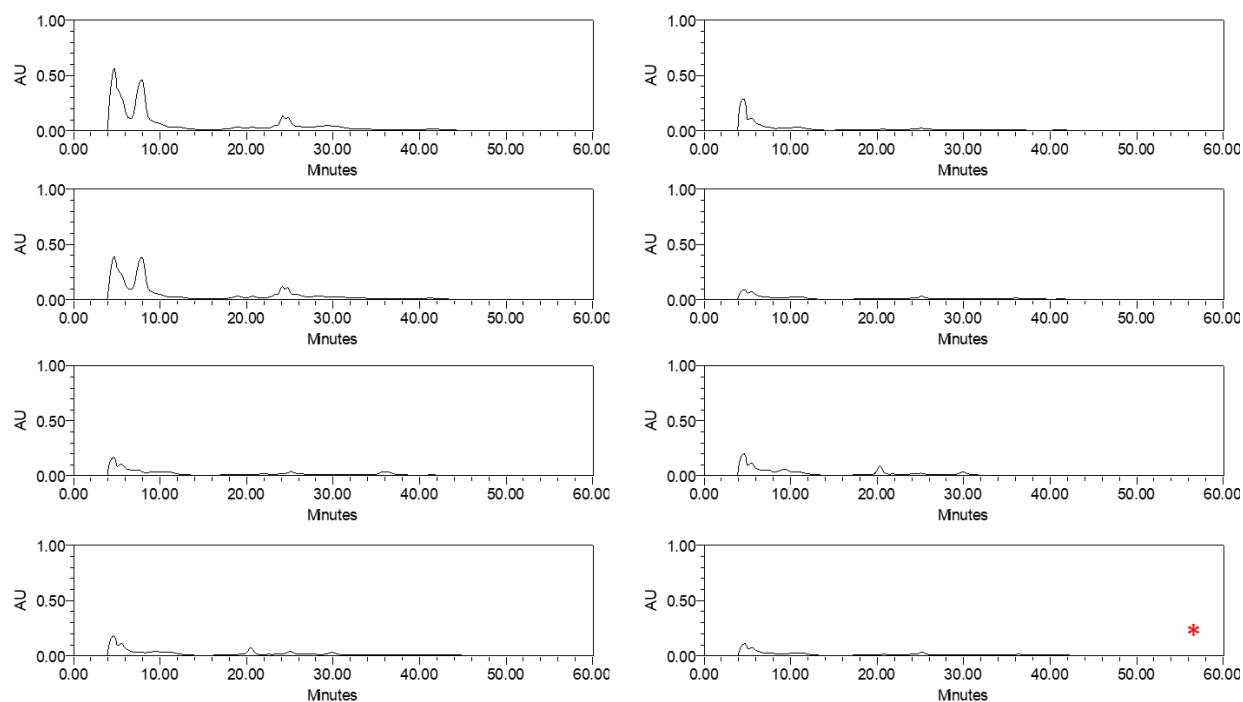


Figure S59: HPLC traces of organic extracts of eight *A. oryzae* colonies that were transformed with *gas* and *pptase* (both genes originating from *C. uncialis*), following incubation in Czapek-Dox (with added starch) for five days. Red asterisk indicates the trace that selected for the main text. Traces are displayed at 260 nm. Reproduced from Bertrand et al. (2019), Supporting Information file, manuscript in preparation.

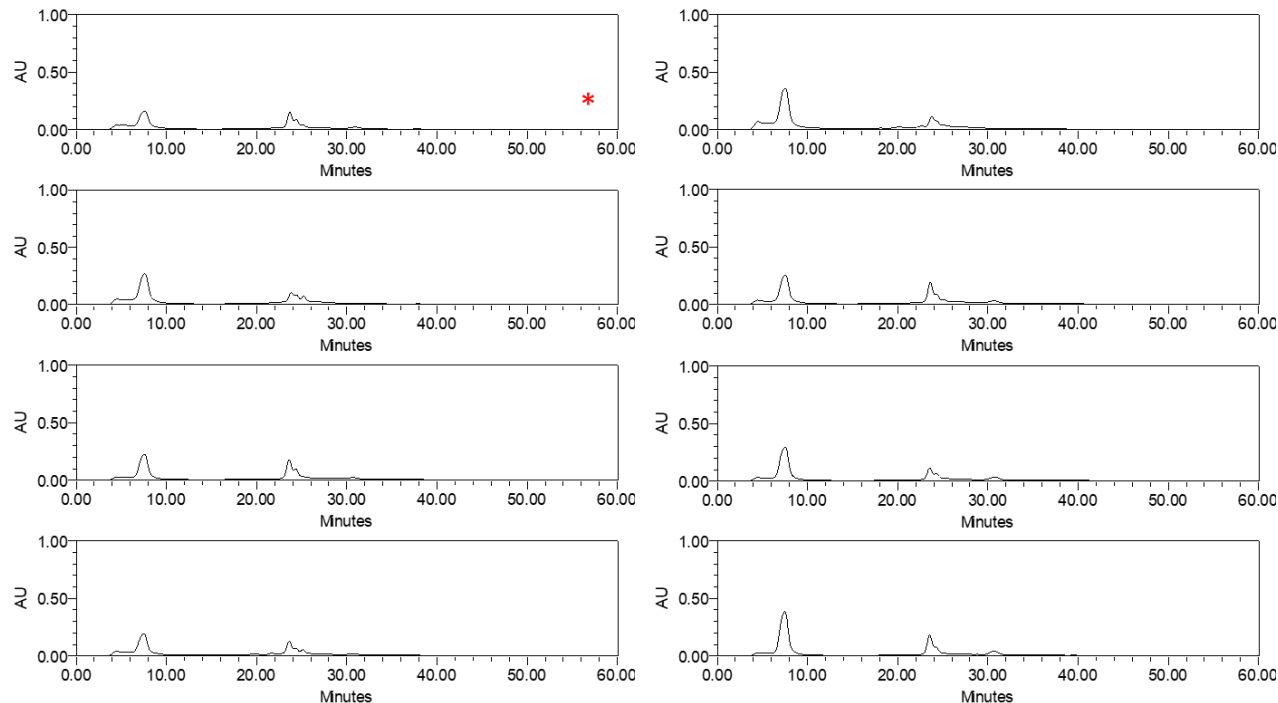


Figure S60: HPLC traces of organic extracts of eight *A. oryzae* colonies that were transformed with *6msas* and *pptase* (both genes originating from *C. uncialis*), following incubation in Czapek-Dox (with added starch) for five days. Red asterisk indicates the trace that selected for the main text. Traces are displayed at 300 nm. Reproduced from Bertrand et al. (2019), Supporting Information file, manuscript in preparation.

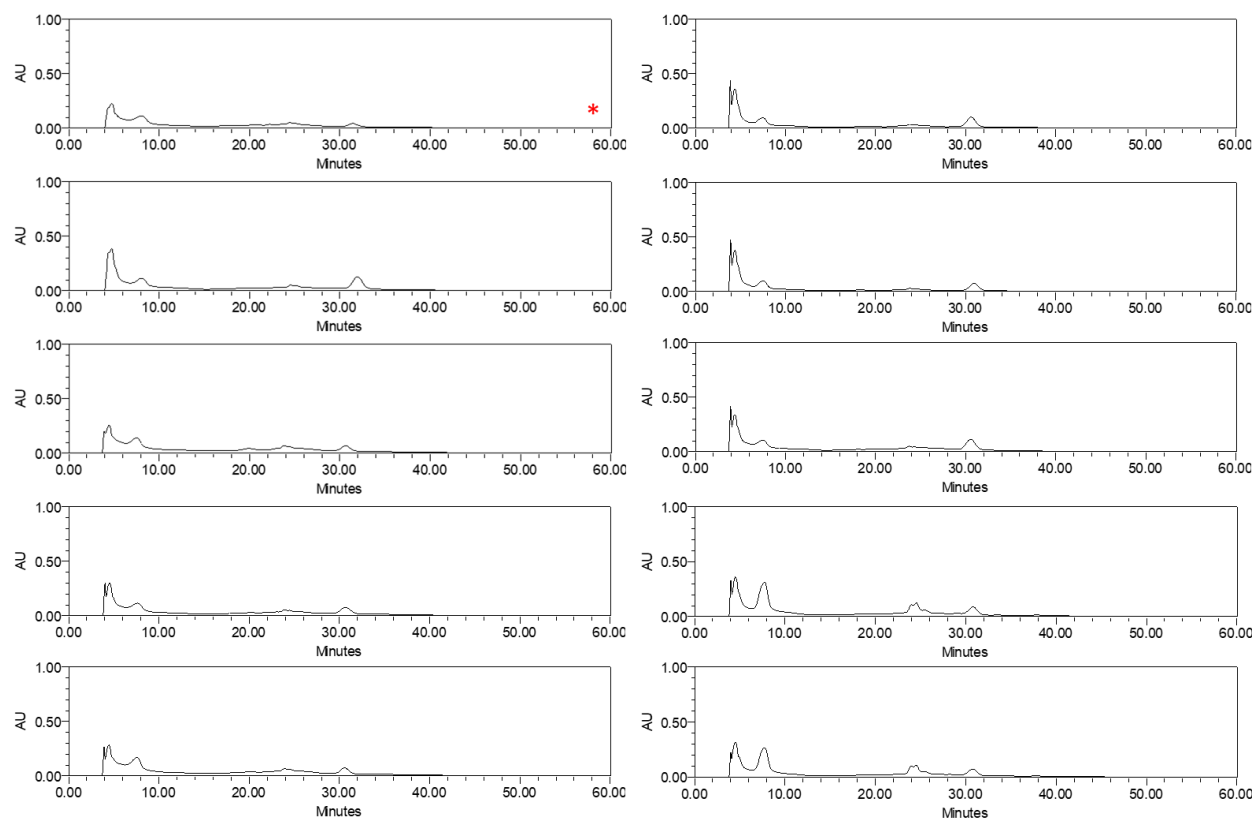


Figure S61: HPLC traces of organic extracts of ten *A. oryzae* colonies that were transformed with *t3pks1* from *C. uncialis*, following incubation in Czapek-Dox (with added starch) for five days. Red asterisk indicates the trace that selected for the main text. Traces are displayed at 300 nm. Reproduced from Bertrand et al. (2019), Supporting Information file, manuscript in preparation.

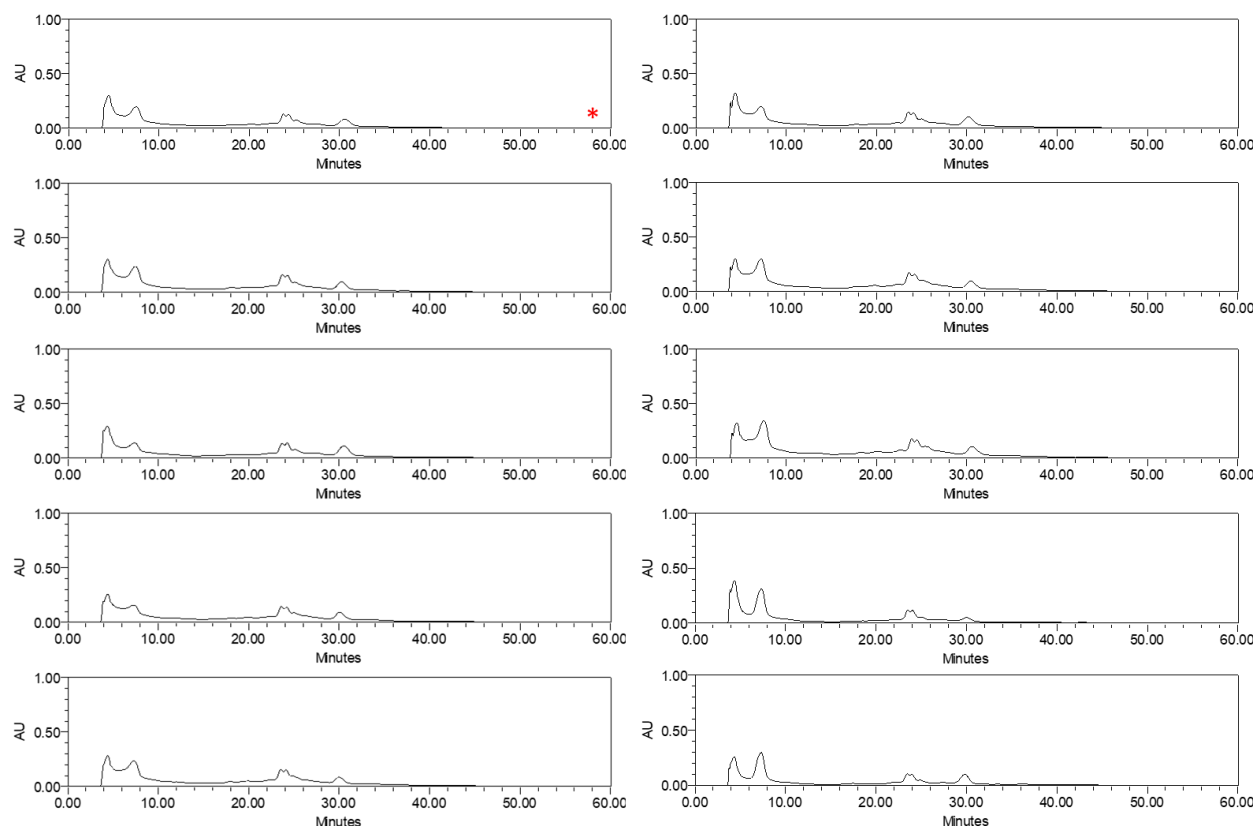


Figure S62: HPLC traces of organic extracts of ten *A. oryzae* colonies that were transformed with *t3pks2* from *C. uncialis*, following incubation in Czapek-Dox (with added starch) for five days. Red asterisk indicates the trace that selected for the main text. Traces are displayed at 300 nm. Reproduced from Bertrand et al. (2019), Supporting Information file, manuscript in preparation.

Sequence S4: The codon-optimized variant of *mpas* that was commissioned for this work (5' – 3'). This sequence only contains exonic DNA. RT-PCR of mRNA isolated from *A. oryzae* that was transformed with codon-optimized *mpas* demonstrated transcriptional activity of the complete gene. Reproduced from Bertrand et al. (2019), Supporting Information file, manuscript in preparation. Accession number: MK481045

ATGGCTCTGCCCTCTCTGATCGCTTTCGGCGCTCTCGCTCCCTGGCCCGCTTCTGATCGCCTCGATCAGCTCCGCAATGCCCTCCA
GCATCACAACCTCCCTGAAGCCCATCACCAAGGCCATCCAGGAGCTGCCTCTCCTGTGGAAGGCCCTCTCCAACCAGGACCAGTCTC
TGCACTCTATCGCCGCGGAGGCTGCCGCTGATCAGCTGGCTCAGTGGATCTCTGGCGCTGGTACCGCCAGCTCGTCGATGATAAG
GGCAATGTGACCCGATGCCTCTCACCACCATCGCCAGATCGCCAGTATGTCTCTTATCTCTGCCAGTACGAGGAGCCCTCCG
CCACGAGTCCATCATCAAGTCTGCTGCCATCGCGCGCGGTATCCAGGGCTTCTGTATCGGCCCTCTCTCCGCTCTGGCCGTCGCTT
CTGGCAAGACCGAGGATGACGTGGGCAATTCGCCGCCATGTCCGTCCGCTCGCTTCTGTGTGGGCGCTTACGTCGACCTCGAT
CGCCACCGCAACGGCGGTGATTCCAAGGCTTCTACCATCGCTGTGCGCTGGAAGACCCCTACCACCTCGAGGATATCCAGCGCCT
GCTCTCCCGCCATCCCGATCAGACCTACATCGCTGTGGTCCGCGACATCCGCGATGTCACCATCACCCTGCCGCTCCGTCATGG
AGCACCTCATCGAGGATCTGTCTCAGATCGGCGCCTCTCTCCGCGACACCGGTGTGTCTGGTTCGCTACCAGTCGCCATCCACGAG
GGTATCCCCAGAAGATCCTGGAGACCTGCCAGGCCAGTTCTCTCTACCATCAATGGCCAGCCCTCGTCCGCTCCAACACCGA
TGCCACCTGTTCTCGGCGAGGATACCGCCCTGCTCGCCCTTGAGTATCTTGGGCGAGCGCGCCGACTGGTACTCCACTATCT
CTACCGCCGCTTCTGCTCTCAACCAGATCTCTGCCAACCCCTTTCATCCTGTCTATCGGTACCGATCCCGTCCCCAGTCCGTCGCT
CGTTCCTTCCCTGTGTCGAAGGCCACCATGATCGCCGACCGCGTCAACGGTATCGTCGAGCCCGAGATCCCTGCTCTGGCTCCTGA
CCCCCTCGGTTCCGTGTCAGGGCTACCCTAAGGACGCTATCGTATCATCGGCATGGGTTGCCGCTTCCCTGGCGCCGATTCTA
TCGACGAGTATTGGAATCTCCTGACCGAGGGCAAGTCCATGCTGTCCGAGATCCCCGAGGCTCGCTTCGGCCGTGGTTCCTGCT
CGTTCCAACTCTCCCTCCGCTTCTGGGGTAATTCCTGCGCGACATCGAGGCTTCGACCATGGTTTCTTCAAGAAGTCCCCCG
CGAGGCCGTCTCCATGACCCCTCAACAGCGCGTGTGCTCAGTGGCTATGAGGCTCTGGAGTCTCTGTTACTTTCGCTGATT
CTTCTCGCCCCGAGGATGTCGGCTGTATATCGGTGCTTGCGCCACCGATTATGATTTCAACGTCCGCTCTCATCCCCCTCTGCC
TATTTGCTATCGGCACCCTGCGCTCTTCTCTCCGGCAAGCTGTCTCATTACTTCGGTTGGTCTGGCCCTTCCCTGGTCTGGA
TACCGCTGTTCTTCTCTGCCGTGCCATCCACACCGCTTGACCGCTCTCCGACCGGTGAGTCTCCAGGCTCTCGCTGGCG
GTATACCCCTCATGACCTCTCCTTACCTGTACGAGAATTCGCTGCCGCCACTTCTCTCCCCACCGGTTCTTCTAAGCCTTTC

TCCGCCGACGCCGATGGTTATTGCCGCGGCGAGGGCGGTGGTCTGGTGGTCTCAAGCGCCTCTCTGACGCCCTCCGCGATAATGA
TCACATCCTCGGCGTGATCGCCGGTCTGCTGTGAATCAGAACGATAACTGCGTGCCTATCACCGTCCCCACACCTCCTCCCAGG
GCAACCTCTACGAGCGCGTCACCGAGCAGGCTGGTGTCCGTCCTTCTGAGGTCACCTTCGTGGAGGCCCATGGTACCGGCACCCCT
GTCGGTGACCCATATCGAGATGGAGTCCATCCGCCGCGTCTTCGGTGGCCTGCATCGCGTTGCCCTCTGATCGTCTCCTCCGCTAA
GGCAATATCGGTCACTCGAGGGTGGTTCGCCGCGTGGTCTCATCAAGGCCCTGCTCCAGATGGAGACACCATCTGGCTCCCC
GCCAGGCTTCTTTCAAGACCCTCAACCCTAAGATCCCTGCCCTGGAGCCCGATAATCTGTGTATCCCTACCTCCAATCTCGCTCTC
TCCGCGAGCGCCTGGCCGCTTGCATTAACAACATATGGCGCTGCTGGTTCCAACGCTGCCATGATCGTGTCTCGAGCCTCCTCGCAA
GTCCGTGACCTATCATGATAAGTCTAAGATGTCCATCTCTTCTCGCCCAAAGATCCATCCTATCCAGCTGGCCGCTGCCTCCCTGG
GTGGTTTGTGTGGCTTACTGTGTGGCCCTGGACCAGTATTGCCAGCGCCTGCGCTTCACCCAGGATACCTCCGAGCAGCCTCAGGTG
CTGTCTGATCTCGCCTACTCCCTGTCTACCCGCCCTCAACCAGGAGCTGCCATTACCCCTGACCATGACCGTACCCGATCTGGACCA
GCTGGAGGCTCAGCTGCGCCGAGACCGCTCACCTCCAACAATATCAAGCAGCGCTCCAAGGCTCCTCTGTCTGTCTGTGTTTTCG
GTGGCAAGGTCTCCGATCGCGTGGCCTTGGATAAGTCTGTGGCAGGAGTCTACCCCTGCTGCGCTCCTATCTGTGATATCTGCGAC
AACACCCTGCGCGTCTCGGCTATCCTGGTCTCTATCCCTCCATCTTCCAGAATGAGGCTGTACCCGATGTCTGCTCCTGCACTC
CATGATCTTCTGCTCTGAGTATTCTTGGCTCAGGCCTGGCTCGAGTCCGGCCTTAAGGTGATGCCCTGGTGGGTCAATTCTTTTCG
GCCAGCTCACCGCTCTCTGTGTGTCCGGCATCCTGTCTCTGCGCGACGGTCTGCGCCTCGTTCGCTGGTAGAGCTTCTCTGATGCAG
AAGCACTGGGGCCCTGAGTCCGGCAAGATGATCGCCATCGAGACCGATCAGCAGACCCTCGAGGAGCTGCAGAAGGTGATCTGCGA
GTCTAACGCTTCTTACAACCTCGAGATCGCTTGGCTTCAATGGCCCCACCTCTCATGTGGTCTGTCTGTATCGCTGTCTCGCCCTCTG
AGCTGGAGACCAAGCTCATGGAGCGGCCATCCGCCACCGCTCTCTCGATGTGCCCTACGGTTTCCATTCCCGCTTACCCGAGCCC
CTGCTGCGCTCATCTGGAGGATCTCGCTTCCCTCCCTCACCTTCCACGAGCCCAAGATCCCTCTCGAGACCTGTACCGACATGGGCAC
CTGGACCGAGCCTACCTCCAAGCTGATCGCCGCTCATACCCGCGAGCCTGTGTCTCTCGGTAAGGCCATCCAACGCTCCAGGCTC
GCCTCGGTCCCTGTACTTGGCTCGAGGCCGCTTCCGATTCTCCATCGTGAACATGGTGCGCCGCGCTCTGGGTGAGCCCTCTGCT
ACTGCCAACAACCTTCTGTGCTCCCTGAGCTCAACAAGCCCAACTCTTCCAAGCTGGTGGTGGATGCTACCGTCTGCTCTCTGGGACGC
CTCCACCGTGTCTCAGTTCTGGAATTTCCATCGCCTCCAGCGCGCCAGTATGATCACCTCCGCCCTCCCTCCCTATGCCTGGGAGA
AGTCTAAGCACTGGCTCGAGCTGGATATGTCCGCCGCTCTGAATTTCTGACAAGACCAATACCCCTCCTCCACCAACACCGCCGCT
CAGGTCGAGCTGCCGCTGTTCTGATCCGCTCAAGTCTTTTCGATTTCTCAGGGTCATCATTTTCGTATCAACCCCTCCTCCGAGGA
GTATCAGACCATCGTCAAGGATCTGGAGTCTCTCGGCTCTGCTGTGTGCCCTCCACCCTGTATGTGGAGCTCGCCTCTCGCGCTG
TCCGCGTTGCTGAGGAGGACAAGGGCAACGGCCTGCTGTCCATCAAGGATCTCCGCGTGCATTCTCTCTGGGCGTGAACGTGCAT
CAGACCATCTCCCTCGATCTGCAGCGCCTGGCTCAGTCTGGCGCTTCCGTATCACCACGCTGATGGCTCTATCTCTGGTTCCAA
CCCCGGCGAGTCTTCTGTCTGTCATGCTGAGGGTACCGTCAATCTCAAGGTGCTGATGATTCTCTCGAGGAGGAGTTCTGCCGCTATG
AGCGCCTGACCGGTCTATAACAAGATCATCTCTATCGCCGACGACCCCCGCTCCGAGTCCCTTAGAGGCAATGTCTCTATAATATG
CTGGGCGCGCTCTCAACTATCTGATTGGTACCGCGCGTGAAGTCCGTTGGCCGCTCTGGATCTCCGCGTTGTGGCTAAGGTAC
CTGCCCGTGGGTATCCCTGAGATCGTCTCCAAGGAGTCCACCACCGAGTCCCTATCCTGAGTCTTTTCATCCAGATCGCTTCCC
TGCATGCTAATTGCCGTGCATGAGTGCCGCGGTGGCGAGGTCTTCCAGTTACCCGCGCCGACCATATGCAGTGGGCTCCTGGTTTC
GACCTGCATGGTTACGGCGATTCCGCCGAGGCCCTCCTGGGATGTGCTGGCTTACAACCTCTACCAATGCTGAGAAATGTGGTCTACGA
CATCTTCGTCCACGATGCTGTACCGGTGCGCTCGCTCCTGCTGCTCGGTGCTAACTTCACCGATATCCGCCGCCCCGTGCCTA
TCTCTGCCGGTCTTAATACCTCTCTCGCTTCCGAGAAGGACATCCCTATGCTCAAGAACGCCAATGCCGAGCGCGCTGAGATCTCC
CTCAACTCTCAGCTGCCCGCGAGTCCCACTCTCAGGCTAACCCTGACCCGCCCGGCAAGGACGCTAAGACCTCTATCTACGAGGA
CATCTCGCGCCTCTTCGAGAAGCTCGCCGACATCCCGGTGACCGGTGCTGTGGTGAGGCCACCTTCCAGACCTTGGGCGTTCGATT
CCCTGATGATGATCGAGGTATCTCCGAGCTCTCCACCCTCTTCCGCGTTCGATCTGCCCATCCATGAGCTCGAGGAGCTACCCGAC
ATCAACTCTCTGGTCGACTACCTGCATGGCAAGGGTGTGTGGGCTCTCTGTACGTCGAGGATTTCTGGTAATGCTTCTTCCCTCTC
CTCCTCCCATGCTATCTCCACCGGTGCTCTTCCCCCCCCGACTCTTCTGGTGCTTCCGCCATGACCACCCCTCCTGAGACCCCTCT
CTCTGGTTCGATTACCTGGCTCCCTCACCACCAAGCAGGAGTCCCGCGCTGCCCTGCTATCTCCAACGGTACCGGCCGCCAGCCC
CTGGATATGGGTCCATACGGTATCCAGCAGGTGTTACCCGCCCTCCGCTTCGATTTCGAGAAGTATGCCGAGCAGACCGGCCGCAA
GGGCTTCTGGACCAATGTGTACCCTCAGCAGGCTGACCTGGGTGTGTCCTACGTGGTTCGACGCTTACCGCAAGCTCGGTTCGCATC
TCGCCACCTTCGCTGCCGTCAGCAGCTTCCATCCATGAACACCTTCCGCCCATAGACACCTGGTGCCCGAGCTCCGCAACATC
CTCGTCGAGTCCGGCTGCTCGAGCTGCGTGGTAATCAGGTCCACGTCCGCACCGCCAAGACCGTTCGACTCCACCCTACCGCCAT
CCGCTATGAGCAGATGCTCCAGCGCCACCCCTTCCGTGCTTCCGAGACCAAGCTGCTCAACGTACCCGGCCCTCGCCTGGCTGATT
GTCTACCCGCCAGAGGAGCCCTCTCTCTGCTCTTCGGTGACAAGCATAATCGCGATCTGCTGGCTGACTTCTACGCCAACTCC
CAGATGCTGAAGGTGCCACCCGCTGCTGGCCGAATTCGTGTCTTCCACCTTCTCTGCCGCCAGTCCGGTGATACCCCTGTGTAT
CCTGGAGGTTCGGTGCCGGCACCGGCGGTACCACCCGTTATCTCTGTCGATGTCTGAATCGCTGTGGCATCCCTACGAGTACACCT
TCACCGACATCTCCAGTCCCTCGTCACCCAGGCTAAGCGCAACTTCGCTCCTTCCAGATGCGCTTACATGACCTTCGACTGC
GACCGCCCCGCCCCACAAGAGCTTCTCGGTAAGTTCCACATCGTCATCTCTACCAACTGCATCCATGCTACCTCTAACATACCAC
CTCCACCACCAACATCCTGCCACCCTGCGCGACGATGGCGTGTGTGCTCGTGGAGTTACCCGCAATCTGTATTGGTTTCGACC
TGGTGTTCGGTCTCCTGGAGGGCTGGTGGCTGTTCTCTGATGGCCGCCAGCATGCTCTGGCCAATGAGTGGTTCTGGGACCGCTCT
CTCCGCGCTGCCGTTTCAAGCACGTGTCTGGACCGACGGCAACACCGAGGAGCTAAGACCTCCGCTGATCTGCGCCTTCCG
CGGTGAGGCTAAGGAGGACCGCAACCTCGCCGCTCCCAACGGTGTATACCAAGCGCGCTGGTGTCCCTATGGAGGAGGTCTGTG
GGAAGCGCTCGGCACCCCTGGATCTCTCCGCCGATATCTACTTCCCTAAGACCCCCGATCTCTCTGGTAAGAAGCGCCCTATCGCC
CTGCTGATCCATGGCGCGGCCACTTCCGTGTTCCGCTCGCAAGGAGTCCCTATGAGCATATCCGACACCTCATCGAGCGCGCTT
CCTGCTATCTCTACCGATTATCGCCTCTGCCCGAGACCAATCTGTTCGAGGGTCCCATGACCGATTGCTGCGATGCCCTGAAGT
GGGCTACCGAGACCCCTGCCACCCTTCTCTCTCCGGTCCACCGTCCGCCCTGATCCAACCAAGGTGGTGTCTCTCGGCTGGTCC
TCCGTTGGCCAGCTGGCTATGTCTCTCGGTTACACCGCCCCCGTGAAGGGTATCAAGGCCCGGACGCCATCTTCCGCCCTCTATCC
TCCCTCCGATATGGAGTCTAACCATTGGCATCAGCCCTGTTACCTCTCGCCGCTGAGGAGGAGCTACCGAGATCCTGGATATCC
TGGCTGGTGTCCGCGAGTCTCCCATCGTGGAGTATGCCCTGTCTCTGAGAAGCGCACCATGGCTCTGTCCCTGACCCTCAAGGAT

GATCGCGCTTCCATCATCCTCCATATGTGCTGGAAGTCCCAGACCGTCCCCATCCTGGTCCACGGTCTCCCTTATAAGAAGAACCT
 CCCCATACCGACAAGACCGACTGGAAGTACCGCCCCCTGCTTCTGCCGAGCAGGTCCAAGCTATCTCCCTCTCTGGCAGATCC
 GCCAGGGCAACTACAAGACCCCTACTTTTCATCGTGCATGGTAATGGCGACGACTGGCTGCCTCTCTCCATGTCTGAGCGCACCGTC
 GAGGAGCTGAAGCGCCGTGGTATCCCCGCTTCTCTCGCCGTGCCGAGCAGTGTGGCCATGCTTTGACCTCTTCCCTGTGGCGA
 CCCTCTCGGCGTCGGTTGGACTTCTCTCGAGCAGGGTTACGATTTTCATCTGCCGCCAGCTGGGCATGTCTGA

Sequence S5: The codon-optimized variant of *mpao* that was commissioned for this work (5' – 3'). This sequence only contains exonic DNA. RT-PCR of mRNA isolated from *A. oryzae* that was transformed with codon-optimized *mpao* demonstrated transcriptional activity of the complete gene. Reproduced from Bertrand et al. (2019), Supporting Information file, manuscript in preparation. Accession number: MK481046

ATGATCTCTCTGTCTCCTCCATCCTCGCCTCTATCTGGGACAATTCCAAGCTGCTCCTCGATCATACCTCTGTCTCTCCATCGC
 CCTCATCGGCGTGGCTGTGCCATCTCTATCCGCTCCATCCTGTATCGCCTGCGCCTGCGCTACTCCACCCCTTTGCGCCATGTCC
 CCGGTCCCTGGTATGCCAAGTTCACCGCCCTCGGTCTGCGCGCTAACGACGTGGCTGGCAACCGCTGGTACTATGTCCAGGGCCTG
 CATAAGAAGTATGGCTCCATCGTGCATCGCCCCGTGAGGAGGTGGCCATCTCCGACCCAAAGGTCGTCTCCAAGTCCACGCTCT
 GGGCACCGAGTTCCGCAAGCGCCAGCAGCCAGGCACCCATTCAACATCTTCTCCATCTCTGATCCCAAGGCCCATCGCACCCGCC
 AGCGTTTCTATGCCAAGGCCCTTCTCTGATGAGACCTGAAGGCCCTCACCGAGCCTGCGGTGCGTCAGTTGATCAAGACCGCCGTG
 GCCTCCATCAAGCGCGACGCTGCTCTCCGCAAGGATCATACCGCCGATGTCTACAAGTGGTGTATGCTCTTCGGCTCTGATGTGGC
 CTTCCAGGTGATCTACGGCAATTCCAACACCGAGGGTCTCATGGCCACCCAGAACACCGATGAGGTATCATGGGCGCCTACC
 TGCAGCGCATGAATGCCTGGGCCAGTTCTGCTTCCCCGTGTTCTCTCGGTGCGTGGCTGTCCCTCTCTCTCCTACCCTGCAT
 AACATCTTCCGCGTCGAGGAGAAGTACGGCGATTCTTGGCAGGAGGGCCAGCGCCAGCGGAAATCGCTGCTCGTACCGTCTTCGT
 GCAGAATACCAAGTATTCCAAGAATGACGGTGTCTTCTCTGTCTCCGACGAGGTGAAGCTCTCCGATGTGACATCGCTCACGACA
 TCACCACCTTCTGGGCGCTGGCGGCGAGCCTGTTGGTGTCTCTCGTCTTCTGATCTGGCAGGTGCTGCGCATGCCTGATCTC
 CAGCGCGAGCTGGAGGCCGAGGTGCTGTTTACCGAGCCTATACCGACGCCACCACCGCCAGCTGCCAATCCTCAACGGCGT
 CATCTATGAGACCTCCGCTGTATGGTGGTGGTGTACCCAGATGCCCCGCTATGCCCCCATCGCCACCGAGTTGGGCGGTTATG
 TCATCCCTCCCGGCACCGCTGTGACCACCCACACTGGTGTCTGCACCGCAACCCCGCTGCCTGGGATGATCCCGAGAAGTTCGAC
 CACACCGCTGGCTCTCTCCCTCTCAGTCTAAGACCCTCGTCCAGGACAAGTTCAGCCTTTCTCTCCGGTGTCTGCGCCTGCAT
 CGCTATCCACCTGGTATGATCGAGCTGCGCGTGTTCGTCTCCGTCTTCTTCCGCGAGTGTGCCGGTGTCAAGCTGGCCCCCTTCTA
 CCACCGATGCTTCTATGCGCGTCTCGATCGCTTCCATATCTCCCCAATGCTAAGCGCTGTGAGGTGATCGTGCCCGACGAGAAG
 GCTTGA

Sequence S6: DNA sequence of *pptase* from *Cladonia uncialis* (5' – 3'). DNA highlighted in blue are the exonic regions. Exonic regions were observed by performing RT-PCR on mRNA from *A. oryzae* that was transformed with *pptase*. In this case, no introns were observed. Reproduced from Bertrand et al. (2019), Supporting Information file, manuscript in preparation. Accession number: MK481047

ATGGCCATGGATGGGCCCCAAGTCTTTCGTTGGGTTCTTGATGTTCAATCACTCTGGCCCACTCCCCAGACGGAACAGCGGTCT
 TCAGCCTTCAGCGCGTGAAGCTACAGCTCGATGGGCAAGTGGTAAAGAAGCCCAACATGCTTTAAGTCTTCTTACCCCGGAAGAAC
 AGGCCAAGGTCTCCGCTTCTACCGACCAAGTATGCAAAGCTTTCTCTCGCCTCCTGTCTTCTCAAACATCGTGCCATAGCAACC
 ACGTGTGAAATTCCCTGGTCTGAAGCCACCATTGGCGAGGATAGCAATCGAAAACCTGTTACAAGCCTTCCACTTCAGAAGGCCAA
 GGCCCTGGAATTCAACGTCTCGCATCATGGTACCCTCGTTGCGCTGGTGGGCTGCCCGGAAAAGCTGTGAGACTTGGGGTTGATA
 TTGTGAAGATGAACTGGGACAAGGACTATGCTACAGTAATGAAAGAGGGCTTTTCACTCTTGGGCTAGGACATACAGTCCGTGTTT
 TCGGATCGGGAAGTGCAGGACATTGCGCATTATGAGGCACCAAGCACGACGACGTACAAGACACAATCAGAGCTAAACTTCGCCA
 CTTCTATGCCCACTGGTGCCTCAAGGAAGCGTATGTTAAGATGACTGGGGAAGCTTTGTTGGCGCCTTGGCTGAAGGATGTGGAGT
 TTCGGAATGTGCAAGTGCCTCTCCCTGGTAGTCTCGCGGTGGATGGAGTTCCAGAGGTCAACCTGTGGGGCCAGAGTGTACAGAT
 GTTGAGATCTGGGCTCATGGAATCGGGTGACGGATGTGCAGCTGGAGATACAAGCCTTTCTGTGATGACTATATGATTGCAACGGC
 ATCTTCTCATATTGGCGCTAAGTTCTCCGCATTCAAGGAGCTTGATGTTGAAAAGGATGTCTATCCTTGA

Sequence S7: DNA sequence of *cpr1* from *Cladonia uncialis* (5' – 3'). DNA highlighted in blue are the exonic regions. Exonic regions were observed by performing RT-PCR on mRNA from *A. oryzae* that was transformed with *cpr1*. In this case, no introns were observed. Reproduced from Bertrand et al. (2019), Supporting Information file, manuscript in preparation. Accession number: MK481048

ATGGATCGCTCGAAGAATGCATGGCCGCTGTACCTCACGACACATTCTACCTATCGCCCAAGGTGTTACAACACCTCGACCTGGA
 TGACATTGTTGCGCTGTTATTCTAGCCTTCCTCTCTTTAGCGTACCTTTTCTATGGCTATCTTTACGCAAAGCCGGAGCCCAACA
 ACCATCTTTTCTATGATATCCCTCAGGATACGTACGAGACTTGTGAGGAAGTCACAAGATCGTCGCAATATCGTGGAGAAGATGAGA
 GAGTTGGACAAGGACGTCGCCATCTTCTTCGGTTTCGAGAGTGGGAGAGCACACGGATTTGCAAGCAGCCTGGCCCCGGAATGCCG

TGCACGGTTTGGCATAAAGGCTCTTGTGGCTGATCTTGACGACTACGACCATAGACATCTGAGCACCTTCCCAGCGAACAAAATCA
 TGATCCTCATTCTGGCCACATACGGCGAAGGAGGTCCCTCAGATAATGCTAATACATTCCACGAATACCTTTTCGGTTTACACCAT
 GAAAAGACTTCAAAGCTCGCTAGCCTTCGATACTTTGTCATATGGACTTGGCGATAGTAATTATCGTCTTTACAACCGATTTGTGGA
 TGTGGTTGATGAAAGGATGCAGGCTGCTGGTGCCTTTGCGTTTAGGCTACATAGGGAAAGGTGACGCAGCACTAGGCAATACGGCTA
 CAGAAGACTCCTTCGGCAAATGGAAGTCGACATGCTCAAAGCGTTGGGTGAAATTATGGGCAAAGAAGCAAGACCAATGACATAC
 GAGCCAAGCTTGAAGTATTGCAACAGTGCCTGCGAAGCTCGGGACATATATTTGGGCGAGCCGAGCCCGCAGCATCTCACTGG
 TTCGCCCCAAGTCGAATATCAACCTGCAGAACCCCTACGCTGCTCCCGTAGCACTAAGCAAGGAACCTTTTACAGAGGGGAGATCGCA
 ATTGTGTCCATATGGAATTGACATTTCCAAAGTACCCCTCGCTGAAATACGTCTCTGGCGACCACCTGGCAGTGTGGCCTATCAAC
 CCTGAAGGCGAAGTCGCACGCCTCATTGAAGTACTCGGTTGGGACGAGACAATGCGCAATGCCTGTATCGACATAGTCCCCAAAGA
 GAACGGAACGCCAGTGCTTCTGCCAACCCCGACCACCCGTGAGACTGTTCTGAGATACTACCTAGAGATTTGTGGTCCCGTATCTT
 ATGACTTGGTGGGCTGCATGAAGGAACTCGCGCCAGTGAGAATGGCAAAAACATGCTGGAAGCATTGCATCGTACTGGACCAGA
 ATCGGTGAAGAGCTCTCCACAAAGTACCTCAACGTTGCCAAAGTGATGCAGCTTGGCGAACCCAGCCAACCCCTGGACCAACGTTCC
 GCTCGCCCTGCTGATCGAGTCGCTTCCCAAGCTCCGGCCGCGCTACTACTCCATCGCCTCCTCGCCAGTGTTGCTGCCCGCAAAC
 CCGCTATCACGGCCGTGGTCAACGCCATGCCCTCAACCCCAACCACTCGCAGCCCGACCGCCGAGCGCTTCCACGGTGTAGCCACC
 AACTACCTCCTCGCCCTCAAACGTCACCTTGACGCGGCAACTCGACTCCCTCACGGCGGACCACCCCAACTACGCCCTTTTCGGGTCC
 GCGTAACAAACTCGAGGGACCCAAAGTCTCTGTACATCCGTCTTCGACCTTCAAGCTGCCTACCAACCCAGCCCGTCCCATCA
 TCATGGTTCGCGCCGGAACAGGCATCGCGCCCTTCCGCGGCTTTATCCAGGAACGTGGCAAGCTCGCCTCCATGGGCAAGGATGTC
 GGGCGCGCTCTGCTCTTCTTCGGCTGCAGACATCCGGACGAGGATTTCTTATACC GCGACGAGTGGGAGGCATTCAGCAACAGCA
 GCCTGGTATGCTGGACCTGGTGAAGTGCATTCTCACGGCTCGACTCAGCGGAAGGTAAAGTGTATGTGCAGCAGCGGGTACGTGAGA
 GAGGGCAAGAAGTGGCAGGTTGCTTCTCGAGGAGCAGGCTTACTTTTACATCTGTGGGAGCGTGCACATGGCTCGCGGATGTGCGT
 GTGGCGTTGTTGGAGATGTTGGCGCAGTGGGGGAGCAAGAGCGGAGAGGAAGCGGACAGGTATCTGAAAATGCTGAAGAAGATTAA
 GCGATATCAGGAGGATGTATGGTCGAGCTAG

Sequence S8: DNA sequence of *cpr2* from *Cladonia uncialis* (5' – 3'). DNA highlighted in blue are the exonic regions. Exonic regions were observed by performing RT-PCR on mRNA from *A. oryzae* that was transformed with *cpr2*. Reproduced from Bertrand et al. (2019), Supporting Information file, manuscript in preparation. Accession number: MK481049

ATGGCGCAACTGGACACCCTCGACCTCATCGTCCTTGCCATTCTCCTGGCGGGGACAGCGGCATACTTCACCAAAGGAACCTTATTG
 GGGTATACCGAAAGACCCCTTATGCAGGCGCTTACTCCTTAGCCAAACGGGGCGAAAGCGGGCAAACAAGGAATATATTGGAGAAGA
 TGGACGAATCGGAGAAGAACTGTGTGGTGTTTTACGGTTTCGCAACGGGCACAGCTGAAGACTACGCATCACGACTAGCCAAGGAA
 GGCTCACAGAGATTTGGGTTGAAGACCATGGTGGCCGATTTGGAGGACTATGATTATGACAACCTGGATACGTTTCCCGAAGACAA
 GCTAGCAATGTTCTGCTTGGCACTTATGGTGAAGCGCAACCGACGACAACGCTGTGGACTTCTATGAGTTTGTGACTGGCCAGG
 ATATTCTTTTACGAGTGGGGCCACTGCGGATGAGAAGCCTCTCGAGAACTGAAATTCATAGCATTCCGGTCTTGGCAACAATACA
 TACGAGCACTATAATTCTATGGTCCGCCGCTAAATGCCGCCCTAACTAAGCTTGGGGCCACGAGGATTGGGAAAGCTGGAGAGGG
 TGACGACGGCGCGGGAACATATGGAGGAAGACTTCTTAGCTTGGAAAGAACCCATGTGGGCCGAACCTGACTGAAAAGATGGGTCTTG
 AGGAACGAGAAGCCGTCTATGAGCCTGTCTTAAGTATTACAGAGCGAGAAGACTTGGATCCAGAGGCTGATGAGGTTTATCTTGGT
 GAACCGAAGAAAAATCACCTTGAAGGTAGCCAAAAGGGACCCATAATTCTACAATCCGTACATCGCGCCAATCGTGGAAATCGAG
 AGAAAATTTTACCCTCAAGGACAGGAATTGTCTGCATATCCGATGCTGATGCTCAAGGGCTCCAATTTAACAATACCACTAGGTTGACC
 ACATCGCCGTATGGCCCAAAATGCAGGGAAGAACTGCATGCTGCTCAATATTTCTCGGTCTAGTAGAGAAAACGCCACAAGGTC
 ATTAGCGTGAAGGGCTGGATCCTACTGCCAAAGTGCCCTTTTCCACGCCCACAACCTACGATGCAGCGATTCCGGTATCATATGGA
 AATATGTGGCGCTGTCTCCCGTCAATTCATCTCATCTTTAGCACCCTTGCACCTGACGATAAGTCGAAAGCAGAGTTGACAAGAC
 TTGGCAATGATAAAGACTATTTCTCGGAAAAGGTCTCAAGCCAATATCTGAACATCGCACAAGTACTGGAAAGGATACACAGCTCG
 ACTCCATGGACAAGCATCCCGTTCTCCATCATGATTGAAGGCATACACAAGATCCAACCACGTTATTATTCATATCATCTCTCTC
 GCTGGTACAAAAGGACAAGATCTCGATTACAGCCGTCGTGGAGGAAATTCGTATTCCGGGCAGTGAGAACATCGTGAAGGGAGTCA
 CGACTAACTATCTTTTAGCCTTGAAACAGAAACAACATGGTGACCCAGATCCAGATCCCCATGGCTTAACGTACGCCATTACAGGG
 CCGAGGAACAAGTACGATGGCATTACGTCCTGTACATGTTCTGCTCACTCAAACCTTAAACTACCATCGGACCCATCGAAGCCCAT
 CATAATGGTGGGTCAGGTTTCAGGCGTTGCACCTTTCAGAGGATTATCCAAGAGCGAGCTGCTCAATCTCAAGCTGGCAAGAACG
 TAGGACAGACAATACTCTTCTACGGCTGCAGGAGAGCTTCAGAGACTACTTATATAAGGAAGAGTGGGAAGTAAGCATACATTTT
 CAATCTTTCTCGATCCATGATGGACGCGCTTGCAGTAAGGGATAAAGACTGACAGAGCCTTTTACAGCACTACAAGCAGATCTCT
 TGGTGACAAGTTTTACTCTCATCACCGCCTTCTCTCGCGAGGGGCCAAGAAGGTCTATGTCCAGTATAGACTTCAAGAGAATGCTG
 CCCTTGTCAATGAGCTGCTCGAGAAGGAGAAGGCAAACTTCTACGTTTTCGGGGACGCGGCAAAATATGGCTCGCGCCGTACGCGAC
 ACGCTAGGGAAAATCATAAGTGAGCACCAGGCGTGAGTGCAGAGAAAGCAGAAGAAATCGTGAAGGCGATGAGGTGAGTAATCA
 GTACCAAGAAGATGTTTGGTCTTAG

Sequence S9: DNA sequence of *b5r1* from *Cladonia uncialis* (5' – 3'). DNA highlighted in blue are the exonic regions. Exonic regions were observed by performing RT-PCR on mRNA from *A. oryzae* that was transformed with *b5r1*. Reproduced from Bertrand et al. (2019), Supporting Information file, manuscript in preparation. Accession number: MK481050

ATGTCCTTCTCTCACCAAGTCCGGCCTACTTCTCTCTATCCCCGAATTCATTACCGGAGTTTATATTCCCTCGGCCTTGTTGCT
AGCTGGAGCGTGCTTTATTAATATCAACTTGCTTCCATATGCAGTAGCAATCTCCTTGGTCCTAGGGACGTGGAAGATCTACACCA
ATGGTAATTGAAGCTTACGAGAAATTTTCCAAAGCTCTTGCTAACCTTCGGTATTACAGCACCCGATATACCAGGAGTCGCCCCTA
CGGCTGTGATACCGGCTAAAGCACTCAATCCCAATGAATTCAGGAATTTCCACTCAAAGAAAAGACCGAAATTTCTCACAATGTA
GCAATGTCAGTCAACAATCGAATATTCAAGACGAACAACAGCTCACTCTCAGCAGTTACCGATTTAGTCTACAAAGTCCAACACAG
ATATTGGGTCTCCCCATTGGCCAGCATATGTCTTTGGCAGCCACACTAGAGGTCAAAGGTGAATCGAAGGAAATTTGTTTCGATCATA
TACCCCAATTTTCATCCGACGAGAACCCCGGTACTTCGACCTCCTCATCAAGTCTACCCCGAGGGAACATATCAAGGCACATGG
CAACGCTGAATATTGGAGACGTAATGAAAGTCAAAGGACCTAAAGGTGCTATGGTCTATACACCGGTCTGGTCAAGCGCTTCGGT
ATGATCGCTGGTGGGACTGGCATAACGCCAATGCTGCAGATTATTCGGGCAATCATCAGGGGTAGACCTAGAAATGGAGGCAATGA
TACAACGGAAGTGGATCTGATATTTGCTAATGTCAATTACGAGGATATTCTACTGAAAGAAGATTGGATAGCCTAACAAAAGAGG
ATCCGGGTTTTAGGTGTACTATGTTCTCAACAATCCTCCGGAAGAATGGGATGGTGGTGTGGGATTCTGCACGGCTGATATGATG
AAGGTAGGTTGTATCTTTGGAGTCTCCATGTCCATGGCGTGCTCAGCAGTATAGCGCAAATTACCTCCTCCAGGATCAGACATGA
AACTGCTCATGTGCGGTCCGCTCCCATGATAAGTGCCATGAAGAAGGCAGCTGATTCTCTGGGCTATGAGAAGCCAAGACCTGTC
AGCAAACCTTGTCGACCAAGTCTTCTGCTTCTAA

Sequence S10: DNA sequence of *b5r2* from *Cladonia uncialis* (5' – 3'). DNA highlighted in blue are the exonic regions. Exonic regions were observed by performing RT-PCR on mRNA from *A. oryzae* that was transformed with *b5r2*. Reproduced from Bertrand et al. (2019), Supporting Information file, manuscript in preparation. Accession number: MK481051

ATGTTTGCACGACAGTTGTTTAGATCAGCGCAGCCGCTGAAGCAAATAAGTGTATGGATGTTGCCAGTTTCAATCGCCTTCAAAA
GGACCTCTCCGCCGCCATCGAACCTCTATATGTCGGTCAACGTGCTGACGTTTCTACTTCCTAGAGTGTCCGCCGCTATGCTTC
TCCCCCGGGCCAGCATCAGGTGGCTCCAACACTGCCATCTATGCTGTCTGGGTGCGGCCGCTTAGGAGGAGGTGGATACTACT
ACTACCAGCGCACTGGCATCGGCAAGCCAGAAGCCTACGCACCTCCCTCAAAGACGAACGAAAGACCGTCCCCTCGAGTCCCGCA
CCCAGCGCACCAGCTCCCCAGTCTGCTGCCTTCTCAGGTGGCGACCAAGGCTTCATAGACCTAAAGCTTGAATCCGTCGAAAACAT
CAACCACAACACGAAGAAGTTTCGCTTTGCTCTTCCAAACAGAGATGATGTATCCGGACTGCAAAATGCCTCGGCGCTCTTGACGA
AGTACAAGGGTCCGGAGATGGAGAAGCCGGTTATAAGGCCATACACGCCCCGTTAGCGACGAGGACCAGCAAGGTTACATTGATTTG
GCTGTGAAGCGGTATCCCAATGGGCAATGAGCGACATCTGCACAATATGAATCCAGGGCAGCGGTTGGAATTCAGGGACCAAT
ACCTAAATATCCATGGGAGCCGAACAAGCAGACCATATTGCATTGATTGCTGGTGGTACCGGAATCACTCCGTACGTTCCCATCC
TCCTACAACATTTTCCATACACAAACCCCAATAAATGGGACTGATTGGATGACAGTATGTACCAACTCGCGCGGCCATCTTCAAA
AACCCCGCCGACAAAACAAAAGTCAACCCTTGCTTTCGGCAACATCACCAGAGCAGATCCTCCTCAAGCGCGAATTCGAAGACCT
CGAGAACACCTACCCGCAACGCTTCCGCGCCTTCTACCTCCTCGACAAGCCTCCGGAGGGCTGGACCCAGGGGAAGGGTTTTATCA
CGAAAGATTTATTGAAGACGGTGTGCCGGAGGCGAAAGAGGAGAATGTAAAGATCTTTGTGTGCGGACCGCCGGGATTGTATAAG
GCTGTGAGTGGGGCGAAGAAGAGCCGAGTGATCAAGGGGAGTTGATGGGGTATTTGAAGGAGTTGGGGTATGAGAAGGAGCAGGT
TTATAAGTTTTAG

Sequence S11: DNA sequence of *6hms* from *Cladonia uncialis* (5' – 3'). DNA highlighted in blue are the exonic regions. Exonic regions were observed by performing RT-PCR on mRNA from *A. oryzae* that was transformed with *6hm*. Reproduced from Bertrand et al. (2019), Supporting Information file, manuscript in preparation. Accession number: AUW31184.

ATGGCTGA
GGCATTTCAGTTCTTCTATTTGGCGATGAGACAGGAGATTTCCAAGACCCGCTGCGAAAGCTTTGCGAGCGCCAAAAG
GGAGTTCTATTCTTACATTTTATTGATAGACTAAATCAGGTTTTGCACGATGAATTGCGCCGACACCCCTCGCCACGCGA
AGAAACAAATTCACCGTTTACAGACGTTTTAGATCTCGTGAAACAGTACCAGGACTCGGCTTCCCGAAATCAAATTTCT
TGAGACTACCCTCGCTTGTATATGTCAGTTGGGCAGCATTATTAGGTGTGCATCCTGTATTCCGGCGTTCTTAGTGAG
CATACAAAGTAATACGTGGCATAGCTTTTACGATGATCATCCGGCCCAACATGTGCTGCCCTCAAATACCGTACTTGTG
GGCTTCTGCACCGGCTTGCTAGCCGCGGTAGCTGTATCAGCTAGCCAAAGCACTCTTGACTTGGTAGATAATGCGCTGA
ACATTGTCCGAGTTGCGTTTAGGATTGGTGTCAAGGTAAACGATGCTGCACAGCGCTTATCCACAGAAGTCAACCAGAG
ATGGTCACGACTTGTGTTGGGTGCACAGAAGGAGGCTCGATTGCGGAAGTTAGGCATTTTAAACGAAAAGAAGGTGAGA
ATAGCTTCCGCAAGTTTTTCCCTGTTTTACCTGTTCAGAGACACTCTCGGCTCGAAAAGACAGCCAGGAAGGTCCCC
CACAAAGGGCGGGCGCTAGGCAGAGGCTGTTGTCCGAATGGAGAGAAGTAGCACCATTGATCGTCATCCTCAGCATAGA
GCTAAAAATTATCTTAGGCTTTACGCGGGCTAGCCATGCCTATGTAAGTCCAGCAGTAGAGAAGTCGTCACCATTAG
TGGACCCGCGGCAACCCCTCAAAGCTCTATTCTCCGAATCAGAATACTTTCGCGAACGCAAGACACGACCTCTTGATATC
TTTGGTCCGTTCCATGCGTCACATTTGTACAGCGATGTTGGTATTGAAGAAGTTCTCCAGCCGATGACAGAGACGGACG
TCAACCTATCTTCTCCACCAAGTCGTGTCATCTCGGGTACCACGGACACGTCTCTTCCAGATGTGGGCTGGATCAACT
GTTAAGACAAGTTGTGGGCGATATCTTGGTTCGACCACTCTCATTTGACAATGTACTCACTGCGGTGCTCTCGGAGGCC
AGAGCTTCGGGCAGAAAAAATGCAGAATCTTCTATAGGGCCTTCGAATGCCGTGTCCAGCTTAGCATCGATCTTAA
GAGCCGAAACAGAGCTTGAAGTGAGTGTGGGGATCCGTTTGGTCCGACGAAACTGGTAAAAGACGCATCAGGTACCTC
CGGAAGATCGCCGTTGTGGGAATGCTGGACGCTTCCCAATGCCGACAATCTCGAGAGTCTCTGGAGCTTGTAGAA

CAGGGATTGGATGTACATCGAAAAGTTCTCCGGATCGATTTGACGTTGACGCACATTACGATCCGACTGGGAAAAGGA
AAAATACTAGTCACACTCCTTACGGTTGCTTTGTAAAGAGCCAGGCCTGTTTCGACCCCGTTTCTTCAACATGTCACC
TCGCGAAGCCTACCAAACCGATCCTATGGGCAGACTGGCCTTGGTGACAGCCTATGAGGCCCTTGAGATGTCAGGTTTC
ATGCCAATAGGACACCGTCGTCAATGTTAGACCGAGTTGGCACCTTCTATGGCCAGACGAGTGATGATTGGCGACAGG
TCAATGCAGCACAAAACATTCGACACATATTACATTTCCCGGACATACGTGCTTTTGGCTTCAGGACGTATCAATTATTA
TTTCAAGTTCAAAGGACCGAGCTACAACGTCGATACGGCTTGTTCATCGAGCTTTTCCGCGATTCAATTAGCCTGCACA
TCTCTCCAGGCGAAAGAATGCGATACAGTAGTTGCAGGTGGCCTTAATATTATGACAACTCCAGATATCTTTTCCGGAC
TCAGCCGTGCACAGTTCTTATCTAAGACTGGCTCCTGCAGGACATTTGACGATACTGCGGACGGTTTCTGTGCGAGGTGA
TGGAGTTGCTACGGTCATTTTAAAGAGGCTGGAGGATGCGGAGGCTGATAACGATCCAATCCTAGGCATTATATTAGGG
ACAGCAACAAACCATTTCTTCCGAAGCGGTGTCAATTACTCAACCCCATGCTCCAACCTCAAGAGCTCCTCTATAGGAAGA
TATTGAGAAACACTGGTGTGCGACGCTCGTGACATTAGCTACGTGGAAATGCACGGGACAGGAACGCAAGCTGCGGACGG
AGCTGAAATGGAATCTATTTCCAACGTCTTTGTCTCCACGATCCAAAGGACGTTTACCAGGCCAAACCGTTGCTGTGCGGA
GCTCTCAAAGCAAATATTGGCCATGGGGAAGCCTCTGCAGGGTGCATCTCTAATCAAAGTCTTGCTTATGCTGCAAC
GAAACGCCATCCCTCCACATATCGGTATCAAAGGCATTATGAATAAAACCTTCCCCACCGACCTCGAAGAGAGGGGAAT
CCGTATCGCGCTCCAGAAACCTCGTGGCCTCATCCGACTGGTGAAAGAGGCGAGCATACTTGAACAACCTTTAGTGCA
GCAGGTGGAACACTGGCCTTCTACTCGAAGACGCTCTGAAGCCCGCACGAATGCGCGAAAATGACCCTAGAACGACTT
TCGTTGTTTCGGTAACGTCAAAGTCGGCCTGGTCTCTGGAGATGAACATCCAGAACCTTATCTCTTACCTGGAGATGAA
GCCGACACATCGTTACCAAGTCTTCTTATACCTACTACTACCCGAAGGACGACGAGCCTCTCCGTGTCTCCTTTGCT
GTATCAAGTATTTCCGAGGCTACCGAAGCCTTGAGATCAGTCCAAGCTGAGATCATCAAACCTGTTTCTATAAAACCTC
CCAAGGTTGCATTTGTATTACTTGCCAAAGGCGCACATTATCCATCACTAGGGAAACAACCTTTTGAACCTCAAGACA
GTTTAGATCTGATATTCTAGACTTCGATAGAATTGGTCGCAGCGAGGGCTTCCCATCTTTCTTACCTCTTCTAGATGGC
ACTATTACAAATGTACATCGTTTTCGCCCTTCGTGCTACAGCTTGGACAAACCTGCATTAGATGGCGTTGGCGCGGT
TATGGATTTTCATGGGGCGTCTTACCAGTGCTGTATCGGCCACAGCCTGGGCGAGTATGCTGCGTTGAATGTGGCTGG
TGTTCTATCCGTTAGCGATACCATATTTCTCGTCGGGCAAGAGCAGAACCTTCTCGAGAAGCATTGTACAGCCGGAACC
CATTCATTATTGGCGCCGCTGCATCTGTGGCTTCCACAGCCAAGTCAATAGGCGGCGACAAACCAACATTGCTTGTA
TTAACGGACCTCGAGAAACTGTGATCAGTGGTCCAACAGAACAGCTAATGTCTTACTCAAAGACTCTCAAGGCAGCCGG
CATCAAATCTGTTCTGTTGCCCTGCGCTTACGCTTTTCATTTCGGCCCAAGTGACACCGATTCTGGAATCGTTTAAAGGAG
TCAGCCAGCTCTGTTTCTTTTCGCAAATCCAGCAATACCCGTCATCTCCCCTCTGCTTCAAGAGGTAGTGACTTGTGGAG
ATGTTTTTGGTCCAGAATACTTGGCTCGCCACGCCCAAGAAACGGTCAGTTTTCTAGGCGGTATTACGGCTGCGAAAAT
GGAAGAGCTAGTGGACCAGAACATGATCTGGGTTGAGATTGGTCCAACACCCGCTGCTCTGCTCTTATCAAATCGATT
CTTGGAATAGAGACATTAGTTGTGCCCTCACTCCACAAAAGGAAGACCTTGGCGGACTACGTCAAGTGGACTCAGTC
TCCACACCGCAAGGTTTACATATTGACTGGGATGAAGTCCATAACGAGTATGAATCCTCCCATGTGTTGCTTGGCT
ACCAAGTTATACTTTTGAAGACAAAACCTACTGGCTTGAGTATACCAACAATTGGTGTCTAACTAAAGGGGAGGTCAATC
GAGGCACCTACAGAGAAACGACAGAGCCGTTGTTTATCTACTTCTCTGTCCAGAACTAGTCAAAGAAGATTATGGAG
GCACGATCAGACTGGTAGCTGAATCAGATCTTTCCGACCCCGATCTTAACGACGCTATCTCTGGGCACTTGGTCAATGG
CTCAGCCCTCTGCCCATCAGTACGTACATATAGCCAACCTTCTGACTAGAGCTAACAAAACCTTTCAGGGTGTATTTCGCCGA
TATGGCACTGACAAATAGCAGGCCATCTGCATAGAAAGGTGAACAAGAGCGTAGATTATGTTCCCATGAACGTGCGCCAA
CTAGAAATTGAAAAGCCTTTGATCGCAAAGGTCGAGATGAGAAAGAAACAAGTGCTTCAATGACTGCCACTGCTG
AGCGACCTCTCAGGCTTGTCAAAATTGGTTACAAGAGCATATCTCGGACGATAGTGTTGAGGAGTTGCATGCCTTGTG
CGAGGTTGAATACGGGATGCAAAGTCTTGGCTTTTCAGACTGGTCGCGATTGATTACTTGTTCGGTACAGGATCGAT
GTCTTGTCTCAAGGAGCTGATGTTGGCAAATACAAAAGGTCAACCGTGAGAAAGCGTATGAAGCCTTCTCATCTTTGG
TCCAGTATGACAAGAGATACCACGGAATGAAAGAAGTTATTCTCGACAGCAAGAAATTCGAGGCAAGCTCAGTCATCGA
GTTTCGGGCTACTGAGCGCGACGGCGATTTCGAGACGAGTCCCTACTGGATTGACAATATTGCTCATCTTTCTGGATT
GTCTCAATGGCTCCGACGCGCTCGACTCAAAGAAACAAGTCTACATTTCTCATGGCTGGGAGTCTTTGCAAATGTTGC
GTCCCTGTGACGACGACCTCTTACCACAATCATGTCAAGATGTTTCAAGGACCAAGGAGAAACAATTTGTTGGCAGCT
ACACGTATTTGACGGAGATGACATGGTTGCTCTGGTTGGGGGTGTAAAGTTCCAAGCCATTCCACGATCTCTTCTCAAT
AGGCTTCTACCTCCCACCCAGGGAGCCGCTGTTTCATCCGAGGCGGCTGCTCTCTCCCTCGTGAATGGCATCTCTGGAG
ATGCAGAGGTAAGGCAATCTAAAGTTGAAAAAAGAAATTGAATTTCAGCTGTGAACGATACGGCGCATCAACTCAAGAC
ATCAATATCTCATCTCAAAGGGAAATGGTGATAGACCTGTTCTTACTAAATTTATGAGTATCATTTGTGAAGAGCTA
GGGGTGGAACGTTTCGAACCTCTCAGACGAAGCCGCTTTTGTGATATCGGACTCGACTCACTTATGTCGCTTACAATTA
CAGGGCGAATGCGCGAAGTGCTCGAAATGGATGTCCCTACATCTTTTATCTCGGACAACACAACCGTTGGTGGCGGAA
ATCTGCAATCCTTGCTATTGACCGTGATGATTCGAATGAATGTCTCTCAATTGAAAGTGTTGCATGGTAGTAGCCTA
ACTGCAGACGCAATGACCGGCGCTGACACTACGGAGAATACCTCAATTGAGAGTGGTCCAGCGAAACACTCCACCGAGG
ACACTACAGAGAAGCTTCTAAGTATAGTCTCTGAAGAGTTGGGCATCGAGCAATCTGAACCTTCTTGAGAGGGGAGTTT
TGCTGATATGGGTGTCGATTCACTCATGTCACTTACTATTACGGGGAGAATTTCGTGAGGAAATGAATTAGACGTTCCC
ACGTCAATTTTTCGCCGACCATCTGATATCAATAAGGTGAAGGTGGCGATCTCGGCTCTGATTGGCCCGAATTCAGCG
GAAGTGCTACGCCATCTAGCTACACCGGAATACCAGTCCAGGTGAGACTGACACGCCAGCAAGCATGCCGAATTCGGA
CGATGAGATCTTGACGGGTGAGATTGTGGAATTACGAGCTCAAGCAAGGAGAACTTCCCCGGCTACATCGATATTGTTG
TCAGGCGACCCCAAAGCCTGCTCGAAGACTTTGTTCTCTTCCCTGACGGCTCCGGTTCAGCAACTTCGTACGCATTGC
TTCCAGCATTTTCCCGACGTTTGTGTTTACGGTCTTAATTGTCTTTCATGAAGATTCCCGCCGACTATACGAACGG
CATCGATGGTGTCTCGGCACAATATCTTGCTGAAATCAAGAAACGACAGCCCCAGGGTCCATATTATCTTGATGGATGG
TCCGCCGGCGGCGTGATTGCATATCAAGTCGCGTACAAGTTGTTGCAAACGGGTGAAAAGACTGAGCGTCTCTTCTTGA
TCGATTTCGCCCTGTCTATTGATCTGGAGCCACTACCATCCTCGTTGCTTCAATTCTTCAATTCCGCTGGATTACTTGG

CACTGAAAGCACACCTCCTGACTGGTTGATTCCACATTTTCGAGGCCTCAATTGCCAACCTTGCGGCCTACACAACCTGT
GCTATGGGTCCGTCGGAAGCACCTAAAACCTTGATCATTTGGGCCCGCGATGGACTGTGTAAAGACCTTAAAGATCATC
AGTACGTAAGATCCACAGCGGAGCCCAAGAGCGTGAAGTTCTTGCTTGATGATAGACGTAACCTTCGGATGTTATGGATG
GGAAACGCTCTTAGGAGAGGAAAACATCATCACAGCATTTGTTTCAGGGCAATCATTTCACAGTGATAAGGGAACCCGAT
GTAAGTATCTATCGCCATTACTTAAAAGAATTCATATACTGA

Sequence S12: DNA sequence of *oas* from *Cladonia uncialis* (5' – 3'). DNA highlighted in blue are the exonic regions. Exonic regions were observed by performing RT-PCR on mRNA from *A. oryzae* that was transformed with *oas*. Reproduced from Bertrand et al. (2019), Supporting Information file, manuscript in preparation. Accession number: AUW31139.

ATGTCTCGTCTCCAGTCTTTGCCGGTCTTGCGTCA
GAGACTGTATTCTCGAGCTCTGTTGTCGATCAGGCTGTAAAAGACGCCGACAACCTTGAGGCCCAATTGTTGATCAAGG
CTTGCTTCCAGATATTTGTGGAGGAAGCTCGATTCTGCAAGTGGCAGATCTTGCGGAGTTGATGAAAAAGATTTCCCTAC
CGCTGAAAGCCTTCTCCGACCATCGACATATTATCACCACAATGCCGCCATCCAACATGCCTTCATTTGCCTGATACAA
CTCCTCAGATATTTAAGTGTGAGGGTTCGATCAGCTCCAAGCTTCTTGGGGCCTCTGGCTTCTGCGCGGGTCTTTTGC
CTGCGGTGGCCGTTGCGACTTCAAAGTCATCTCTCGAGTATCTCTTAGGGCACAAGATTGCTTCCGCGTTGCATTCAA
GATCGGTATTGCAAGCGAGCAATCCCGAGGGCCACATGCCAATGCGGATGGGGAATTGCCGTGGAGCTTGATCGTCGAC
AAACTGAGCGAGGAAGAGGCCAAAAGGATTATCTCTGAGGTCTGTACCTCATATTCACCTACCAAAGTATTCTCCTAATC
GATCATTTTCAGGCTTCTCCAGCTCTACGTCACTGCAGCAACACCGAGAAGTGTGTGAGTTTGGCGGCGAGGGTGC
TACCTTGGCGCAGTTTGTGACACAAAAGCTGCCCAACGTTGTGCAACAAGGTCGACCAACATTTTCACACTTTACCAT
AATAGCCATGCCTTGAGTGAAATCTCGGCACGCGTCTTTCAAGAGATCAAAGAGGGCGGTTTCTCATTTCCAGACGCGT
CGGAGCTGGTGTGCCAATCTTCCCACTCATGATGGGTACAGGTACACAGTGTGAAAGCTCCTCTTCCGACCAGCT
GCTGAAGCGAGTCTTGAAATGATTTTCTCGAATCAGTGGACTGGGTGGCCGTACAAGACGAGATACTCAAATCCGGC
AGGGAAGCACTTGCGGCAGGAGTTCCATCTGTTTCGAGTCCAAAATTTTCGGCCCCGGATACGGCGCCTTGGTGTATCGCA
AAGATCTGCCTAAAGGCATGCAAGTAGTCGATGTCGCGGTCTCTAAAACCTCTACTCAAAGTGGGCCAAGGACGATAT
CGCGATTGTAGGCGCAGGACTCGATGTTCCGGGGCGCAATGGACTACAACATCCTGTGGGAGAACATTATGAACGGG
ATCAATACCTGCTCAAAGGTATGTCTTACTTGGCGCCCTTGACAGAAACCTTAAGATCATTGAGCTAATTTTCGACAACTT
AGATACCCGCTGACCGATTCCGCATCGAAGATTACCACCAAGACGACAAGTCTAAGTTGGGTAACAAAAATCGATCCAT
GGATACACAATACGGTAGCTTTTTGAAAACCCGTGGCTTTTCGACAACACTTTCTTTGGAATCTCGCCTAGAGAAGCC
CGCAGCATGGATCCACAGCAGCGCATCTTGTTACAAACGGCTCAACGGGCGCTCGAAGATGCAGGCTACTCTGGGATT
CTACGCCATCCAATCAGATCGCTTCTTTCCGGTGTACATTGGCGCTGCGACAGGAGACTACGACAGAACTTGAAAGA
AGAGATGGATGTGTACTGGGTTCCTGGTGTGCTACGTTCTGTTGAGTGGACGGATATCCTATGCCTATGGCTGGGGC
GGTCCATCCCTAACGGTCGACACCGCTTGCTCTTCTTATTGTTGCCATTTGGCAAGCTGCCAGAGCGATTGCAACCG
GTGATTGCAAGCGTGCATTGCAGGTGGTGTGAACATTGTACAAAGTCCAGACATGTATATCGGACTCGATCGAGCACA
TTTCTTAAGTCCCACGGGGCCTTGCCATTCTTTTCGACAGCCCATCTGATGGTTATTGCCGGGCAGATGGATCTGCGCTA
GTTGTGATGAAGAGGATTGAAGATGCACTGGCCGAGAATGATCGTATTCTTGGTGTGATCAAAGGAGTAGAGTGAACC
AGAGTGGTACTGCTCACTCGATTACTCATCTCACACTCTACTCAAGAGAGGCTTTTCAAGAAAGTCTGGCAAAGAA
CAACATCAACCCGCACCACGTTTCCGTAGTGGATGCCACGGAACCGGAACCGGGGCTGGAGATCCAAGTGAGCTTCA
GGTATCACAACCTCTTTGCGTTAATGGTCGGGATGCATCAAATCCGCTCTTTATTGGCTCAATCAAAGCCAAATATTG
GCCACAGCGAAGTGCCTCCGTTGTGCATCATTGATCAAAATTTCTCATGATGATGAAGAACAAGACTTTGCCCTCCACA
GCCGCGATTTAATAAGGAAAATCACAATGTTGCAGACCTCGCGCCGAGGGGCGTCGTCGTTAACCGAGAATCCAAACCT
TGGATTCCGAAGGAAGGCATTCCCTCGTATGGCTATGCTCAACAACCTTCGGTGGAGCTGGCTCTAACGGCGCTCTTATTG
TGCAAGAATACGAAACATCATCGCACCAGCTGGAAAGCAACTCACTACTGTTTGGAAATCTCTACAAAGAGTACAGA
GACTCTTGCGACAATGCGGGACGAGTTCGTGACCCTTCTTGACGCTCAACAGGACATGCTCAACGTCCGCGATGTTTGC
TACTCTAGCACAGCTCGAAAGATTCTGTACCCCTATCGCTTGTCCACTACTGCAAGCTCCGTACAAGAGCTTATACAAA
ACTTTGCAAAATGTTAATCATCGATGCAATTAAGCCCCAAAAGGTTGTGATGAGAAACCAAGGCTGCTTCTCTGTTCTC
TGGACAAGGAAGTCAGTACCTGGGAATGGGCAAAGAGCTCATGACTATGTTCCCTCAATTCGCGGAAACTGTCAACAAA
TGCCACCTGCTGTTGAGCAAATGGGGCATGCCAGCTGCCTTGAGGTCATCAATGCCCCAGAAGCTGAGGCGAACCCAG
AAGATACCACCAACTCCAAGCATTCCAGACTGGCGTGTTCGTTTTGGAGGTTGCACTCGCCCGGCTGGTGTATGCTTT
TGGCGTACTCCAAAAGCGGTAGCTGGCCATAGCCTTGGAGAATATGCGGCTTTAGTCATCGCTGGCGTTATGGACCTT
ACAACAGGATTGAAATTTGGTGGCACAGAGGGCGGCATTGATTGTCATCGTTGTGGACTTCGTGAGACTTGATGTTGG
CGGTCAGCAGCCGACCAAGAAAATCCAGGAAGTGTGAAAGAAAGCCGATTTTAATTCTCTCAGCATCTCTTGCGA
CAATAGCCCCGATGATTGTGCTGTTGGTGGACCGACGCACTTCTGCAAAAGTTTGAAGAGTACGTCTCAGAGTCCATG
AAGGTCAAGTCAAAGCTGTTGGATGTTCCGGTTCGCTACACACCGACGCTCTTGCCCCAGTGACGGAGATTTTGACCG
AGCTTGCCCCAAGAGGTGAAGCTTTCCGGCACCATCTATTCCAGTTGCATCAAATCTCCTTGGTCTGTGCTCCCTGCCGG
TGAAGACGCTTTTCAAACCCGACTACTTTGCTCGTCATTGCGTGCACATGGTGCAGTTTCGATCAATCCATGGAATCTCTG
ATGGCCATGGACAGCACCGCCGCTGACGGCTGCTGGATCGAAATTTGGACCTCACCAATCGTAACGCCAATGCTTAAC
CGAGGCACAAGGACAAGCAAGCCCAGCGACTCTTCAGCTTGAAGAAGAAGAGTCATCCATCGGCCCTCGATGTCTCAGCT
TCTGAGCGGCTATACAGACCACTTCGAATGTCAACTTTCGAAAAGCTTTTCGAAGGCCTCTCGCAAAAACCTTCTCTG
GTGGCTGTGCTGTTGTTACCTTTCGCTCGACAAAACCTTCTTCGTCGAGTACCCACGCGATTATGAATCAGTGGTGAGC
AAATTGCTTTGAGTCAGCATGGTAGCCATGATGCAGGCTACGTCGGGACACCGTATCAATTCTCTACGAAGGTTGTTCA
GGCACCTGCTGCTAACGAGTCTGGTATCGCTGTTTTCGATACTCCAATCCAGACTCTTTACGAGTATATTACGGGTAC

ACTGTATGCGGCTACGCTTTGTGCCCCGCCTCGGTGTACTGCGAACTTGCTCTCGCCGCTGTACACGCGCTGGAGCAGC
CAGGATGGCCCTCAGCTTGTGCGCAATTGACCATTCCCAAGTCGTTGGTCTACAAAGAGGATGCCATCTCTCGAGAAGA
CCTCAGCATCGTTGTACGGGTCACTGTCAAGCCCTCTGATCAAGTTCATGGTCTACGCGAAGTCGAGTGTCTTCGTAC
GATCGACTCAGGGTACAGAGAGAAGTCAGGTGCATGCCCCGCGCCTCGTCAAGTCACGTTTCAGTGGCTGAGGTGGCCG
AGAAATACGCCAATCTCGGCATTTATTGGATCAGCAGAAGTGGAACTTCCTGAACCTCGGCAGATTCAGACATGATTGA
GAAATTCTCCTCGTCCGTCTTCTATGACAAAGTCTTCACAAGGACCGTCAACTACGCGGACATGTACCGCTCCATCGAG
TCTATCAAGATCAATACACCTGCCAAGGAGGCCCTTGTCATCTTGCAAGCTCCCCCTTGACTCGGACTCCCAACAAATTTG
TCGCCGACCCCATCATGGTGGACGCCCTCTTCCAGCTTGGTGGTTTCATGGCCAATCTCTGCGCCAAGAATGATGAGAT
TTGCATCTGCAACTTTGTCAAGTCAGCAACCGTTTTTGCGTGACATCGCCAGCACAGACCGTAAATTCAGATCCACGCC
AGCAACTATGAGATCGCAGGTGGGAAGACGATCTTGTCTGATGCGCAAGCGATCGACTCTCACGGCGTGTTCGTGTCA
TCAAAGCTATGAAGTTTAAGCGCGTGAAAGTATCTGATCGACTTGATGCCTACGCTTTGCTGTCAAAAGATCCAAAGT
TGCGGAAGTTTCTCTCAAGATATGAAATGAACCTTTGCAAGCGAGCTCCCAAGCAGGAAGAGACTCAGCTCTCTTG
GAAGGGGCACCACAGTGAAGCAAGCTTCATCGACTCCTTCGACCGCACAGATCTCTCCACCTTCCTCGCCGGATCTCT
CCTCAGCCAGCAGCCAGTCTATCTCCATCTGAGCAAGAGAATGAGGTACCCACGCCATAAATGGTGGGTTTTTCCACTGA
GGCCGTCATAGCTGAACTTGTGGCACCAGTGAAGTGAAGTAAACCAACACAGTGCTAGAGTCTCTGGGCATGAT
TCTCTCATGGTGCATGAGTTGAGGGACAAGCTCTCCACGCCTTCCAAAACGGACCTTTTCAGTCTCCGAACTGACCAATT
GCACGATCGTAGAAGACATCGAACGGCTGATCAATGAGCAAGGTGCGCTCCGCCCACTGCTTTTCCCATTCATCTATC
CGTCCCATCTCAGATTGCTTGAGCCAACACGATGATTTGCTAATCTGTAGATTATATGATCAGGCTCTTCCCTGAAAC
CACATCAACCACCATAGCCGAGTTGAAAGCTTATCATCGATTACAGAACTCCCGCGTCGCTCAAAAGCGACAGATTT
GATATCGACTCACCAACTATGTACCCGGTATGTACCCATAATGACGGCTGAAGAAGGCGCATTGTGAGAAACATCGA
CCCTCGCCCCGACAGTCCATTAGACATGCCGTCGACTCCCTCGACGTGTGCCCTGATTCTCCCCAAAATTTACCGTC
CATGCTTAAAGAAACCATGGACCCCTTTCACCTTACCACCTCCACTGATCGATGCCGAAGTTCGACAAACTTCAATTAAA
GCATCTAATTTTCAATTTAGCCCAACGCCATCACAGCCAGCTCCGATTGAACCGCCACTGCCAGATCCTTATGATCCAA
CTGGATTTCGACGAGGCTACACTCTGCAATTAAGGAAGAAAAGACACTAGCCTCACCGCCTCTACCGCCAAGATCCGA
ACCTGAACAGAAACCATCGGTGCCCTCCCGTTCCCCCGTAGCTGAAGAGCCAAAGGAGACATCATCCGATCCTCCAAAA
GTACAAAAGGCCAGGATCCTCAAAGCCCTTGAAGGCTAAGACTGTCGAGTGCACCTCAAGCTATCAAAGACTCAGTGCCTC
AAATCTTACGACTGGATAACAAATCCAATGTATCCATTGCCCCAAATGCGAGTCCCTAGTCGAAACAGCGGAGCCATT
ACCTCTTTTCTCGTCCATCCTGGTAGTGGTCTCATTTGTTGAGTATCACAAATGAATGTTTCCAGCGACCCATTGG
ACAATTCACAACCCTGAGTTTCATGAATCCAGGCGCTGGGATGGTATCCCATCCATGGCTCGAGCTTACGCCTCTCTTG
TGTCCCAACAAATCAACAGGCCGTGCATCTTAGGAGGGTGGTCTTTGGAGGAGTTGTGGCTTTCGAAGCAGCCAGGAT
GTTAATTAATAATGGGACTCGAGTTGCTGGCGTTACCTTGATTGACTCACCACCCTGTGGACACCCTTGACATTTCA
CTCCCAACAGAAACCATGAGGCTGCAATCACACCATCCGGGCCCTCTAGAGACCCCAACGGCGTCCGAAAAGCCG
TACGAGACTTTGTGAGAGAGAACTACGTCGCTAGCTCTGGCATGTTAGGGGACTTTACCATCAAGAGCGAAGGGAAGAT
GCCTCCTCGCGTCGCTCTTTCGCGCAAAAGACAATTTTGAAGTGTCCCGGGCGGAGCTTCCATACCACCTTTCTCTC
GATCATCAAGGAGACCCGAAACTAGTCGTAAAGGATGGGAAGCCATTGTCCGTGACCGGGTCACAATGCACGAAATTC
CAGGAAATCATTTTCGATCTTTTCGACGAGAGCAATATTGCTTCCACTTCGGAACAGCTGTCTGATGCTTGGCGAGAGAT
CCAGAAAGACTATACCAGGCGCGGTGAGCTTCTCGGGAGGCCAGGGTCTGA

Sequence S13: DNA sequence of *gas* from *Cladonia uncialis* (5' – 3'). DNA highlighted in blue are the exonic regions. Exonic regions were observed by performing RT-PCR on mRNA from *A. oryzae* that was transformed with *gas*. Note the presence of an unusual GC/AG intron boundary in the third intron. Reproduced from Bertrand et al. (2019), Supporting Information file, manuscript in preparation. Accession number: AUW31177.

ATGACCTTACCAACAATGTTGTCTGTTTGGGGACCAGACTGTTGACCCCTGCCCTATCATCAAACAGCTGTACCGCCAGTCAAG
AGATTCTCTGACACTTCAGGCACCTTCCGACAGAGCTATGACGCGTTCGAAGAGAGATTGCAACCTCGGAATATTCTGATCGCA
CACTGTTCCCTTCCCTTCGATTTCGATTCAAGGCTTAGCTGAGAAACAACTGAAAGACATAACGAAGCTGTATCGACAGTCTTACAT
TGTATAGCGCAATTGGGTTTATTATTGATGTAATTATTACCACATGGCGTTTAAACGGAAGCTGATCTTTATACGCTCCTTTAGT
CATGCGGATCAGGATGACTTTCGGCTTGACGCCAGGCCATCGCGGACATACTTGGTGGGGTTGTGTAAGTGGTATGCTGCCGGCAGC
CGCTTTGGCAGCATCCAGTTCGGCTAGTCAACTGCTCAGACTAGCACCAGAAATGTTCTGTTGGCCTTCGCTTTGGGTCTAGAAG
CAAACCGTTCGTCGACAGATCGAGGCATCGACTGAAAGCTGGCGCTCTGCTGCGGGCATGCGTCCCAAGAACAGCAAGAA
GCGCTCGCTCAATTTAATGATGAATTTGATGATCCAGATATGTACGATTTTCGCGTAACTGACAGAAACCCCCATAGATGATACC
CACCAGCAAAACAGCGTACATCAGTGCAGAATCAGACTCAACCGCGACCCCTTAGTGGACACCTTCGACGTTGGTCTCTCTTTTCT
CATTTGTCAGATTCTTTTCAGGAAAGCTCGCAGAATAAAGCTCCCGATCACTGCCGCGTTCATGCCCTCATTTACGGCTTCCCAAT
GTGGAGAAGATCATTGGTTCCCTGTACACAGTGTAGTATCCCTTAGGAATGATGTAGTCATCATATCGACGAGGTCTGGAAA
GCCATAACAGCCCAAAGTCTTGGCGACGCATTGCAGCATATCATCTTGACATCTACGAGAGCCGATACGCTGGTGCAGAGTTG
TTGAAGAAATGATCAACAATTTTGAAGACCAGGGCGCAAAATTTGACAAGCGTGGGCCCTGTACGTGCAGCCGATAGTCTTCGGCAA
CGAATGGCCACTGCGGGTATCGAGATTTTAAAGAGCACTGAATTGCAGCCACAGCAGGAGCCGCTACCAAACTCGCTCGAATGA
CATTTGCGATTATCGGGTACGCGGCTCGATTACCAGAAAGCGAAACCCCTTGAAGAGGCGTGGAAAACTCCTCGAAGACGGTAGAGACG
TTCATAAGAAGGCAAGTTCTGAATTAACCGGGGGGAAGTTGTGATCTAATTTTATAGATCCCCAGTGATCGTTTTTGACGTGGACA
CTCATTGTGATCCGTGGGTAATAAAGAAATACGACTTACACTCCATACGGATGCTTCTTGACCGGCCCGGCTTTTTTGACGCG
AGGCTTTTTAATATGAGCCCGAGGAGGCATCGCAGACCGATCCGGCCAGCGATTGCTACTATTAACACCTACGAAGCCCTTGA

GATGGCAGGCTACACTCCTGATGGAACCTCCATCCACCGCCGGCGATAGAATTGGCACTTTTTTTGGGCAGACACTCGACGACTACC
GCGAAGCGAATGCATCACAAAACATCGAGATGTACTATGTCTCTGGCGGGATCCGAGCATTGGCGCTGGCCGGCTTAACTACCAC
TTCAAATGGGAAGGTCCATCTTATGTGTAGATGCCGCCTGCTCTTCCAGTACACTGTCTATCCAAATGGCAATGTCATCGCTTCG
GGCGCATGAATGCGATACCGCAGTCGCTGGTGGTACAAATGTGCTTACTGGCGTGGACATGTTTTCGGGCCTGTCGCGCGGAAGTT
TCCTATCGCCACAGGGTCATGCAAAACCTTTGACAATGATGCAGATGGATATTGCCGAGGCGACGGGGTGGGTAGCGTGATCCTG
AAGCGTTTGGATGATGCAGTTGCCGATGGAGACAACATCCAAGCGGTTCATCAAGTCTGCCGCCACTAATCATTCGGCACATGCAGT
ATCGATCACGCATCCACATGCTGGAGCGCAGCAGAATCTCATGCGACAAGTTCTACGGGAAGCAGATGTTGAACCGTCTGAAATTG
ACTACGTGGAATGCATGGGACAGGTACCCAGGCAGGAGACGCCACAGAATTCACCTTCGGTCACAAATGTTATTTTCAGGTCGGACG
AGGGACAACCTTTGTACGTGGGTGCAGTTAAGGCAAACTTTGGGCACGCTGAGGCGGTAAGGCTCACCATCCAATTCTCACTGAA
ATCTATGCGCTAAACGTCAAAATCCAGGCTGCCGGCACCAATTCTTTGGTAAAAGTATTGATGATGATGCGAAAAACGCCATACC
ACCACACATTGGTATCAAAGGGCGTATCAATGAAAAGTTCCCACCATTAGACAAAAATCAATGTTTCGTATTAACCGTACGATGACCC
CATTTGTCGCTCGGGCGGGAGGTGATGGAAAAAGGCGCGTGCTACTTAATAATTTCAACGCCACAATAAGTTATGCCTCGGAAACA
GAAGAGAAAATCCAAGGCTGAGTACTGCTACAGGGCGGCAACACAAGTCTTCTGCTAGAAAGATGCTCCAAAAACGACATTTCGAGG
TCACGATCCACGTAGCGCCCATGTGATCGCCATCTCAGCCAAGACTCCCTACTCTTTCAGGCAAAAATACGCAACGCCTTTTAGAGT
ACCTCCAGCAAAACCCCGATACTCAGCTACAGAACTTATCTTACACTACTACTGCGCGCCGAATGCATCATGCCATACGCAAAAGCA
TACGCAGTCCAAAGCATAGAGGAGCTTGTTTCAGTCAATGAAGAAGGACGTTTCAAACCTCATCAGAGCTAGGTGCCACCACCTGAGCA
CTCAACCGCAATCTTCTGTTTACCAGGCCAAGGCTCACAATATCTTGGTATGGGACGTCAACTGTTCCAGACTAACACATCATTTT
GAAAAAGCATCTCAGACTCTGACAACTCTGTATTAGACAAGGTCTGCCATCGTTTGAATGGATAGTGAGTGCCGAACCTTCGGAG
GAGCGTGTCCCTACACCGAGTGAGTCACAGTTAGCTTTGGTTCGAATTTGCTCTCGCGCTTGCTTCCCTCTGGCAGTCGTGGGGCAT
TACGCCCCAAGGCGGTTCATTGGCCACAGCCTGGGCGAATATGCTGCTCTATGTGTTGCTGGAGTACTTTCCATTAGCGACACACTTT
ACCTTGTGTTGTTAAACGGGCGGAAATGATGGAGAAGAAGTGTATCGGAATACTCACTCCATGCTAGCGGTTCAATCCGCTTCGGAC
TCCATCCAACAGATCATTTCCGGCGGACAAATGCCATCCTGCGAGATTGCATGTCTCAATGGCCCACTAACACTGTGTAAGCGG
CTCGTCAAGGATATTCAATTCGTTGAAAGAAAAGCTGGACACCATGGGAACGAAGACGACTTTGCTTAAACTTCCCTTCGCATTTT
ATTCACTGTCAGATGGATCCAATCTTAGAGGATATCCGCGCCCTCGCGCAAAATGTACAGTTCGGTAAACCCATTGTTCCAATAGCT
TCGACACTTTTAGGCACGTTGGTCAAAGACCATGGTATCATCACCGCCGACTACCTGACTCGACAAGCGAGGACAGGCTGTGAGATT
TCAAGGAGCTTTGCAAGCTTGCAAGGCTGAGAGCATTGCCGGTGACGACACTTTATGGATCGAACTTGGTCCACATCCACTGTGTC
ACGGAATGGTGCGGTCAACACTAGGTGTAAGCCAGCAAAGGCCCTGCCATCACTCAAGCGGGACGAGGACTGTTGGTCAACATTA
TCCAGGAGTATTGCAATGCCTATAACAGTGGTGTAAAAATGAGCTGGATCGACTATCACAGGGATTTTCAGGCGCTCTCAAGCT
CCTGGAGCTACCTTCATATGCATTTGATCTCAAGAACTACTGGATTCAACATGAAGGAGACTGGTCTCTTCGAAAAGGTGAAACAA
CTCGCACTACTGCACCACCTCCACAAGCGAGCTTCTCCACCACATGTCTCCAAGTCGTGCAAAACGAAACTTTCAACCAGGATAGT
GCCCTCGGTGACGTTTTTCGTACAACTATCGGAGCCGAAACTGAACGCTGCCATCCGCGGGCACCTCGTTAGCGGCATCGGCCTATG
CCCAAGCTCTGTCTACGCTGATGTGGCTTTTACAGCAGCTTGGTATATTGCGTCTCACATGACACCTTCTGACCCTGTACCAGCAA
TGGATTTATCGACCATGGAAGTGTTCAGACCCCTCATTGTGGACTCAAACGAAACTCCACAGCTTCTTAAAGTATCTGCATCCAAG
AATTTCAACGAGCAAGTGGTGAATATCAAGATCAGCTCTCGGGATGACAAAGGACGCCAGGAACATGCTCATTGTACTGTGATGTA
TGGCGATGGTCATCAGTGGATAGATGAGTGGCAAAGGAATGCCTACCTTTTCGAATCGAGGATTGCCAAGCTTACACAGCCCAGCT
CACCTGGTATCCACAGGATGCTGAAAGAGATGATCTACAAGCAATTCCAGACTGTTGTGACATACAGTCGGGAATACCACAACATT
GATGAGATCTTCATGGACTGCGACTTGAATGAAACAGCAGCAAACATCAAACCTCAATCCATGGCCGGAAATGGCGAATTCATATA
CAGCCCGTATTGGATTGACACAATAGCACATTTAGCTGGTTTCTATCCTTAATGCTAACGTGAAGACACCTGCAGACACAGTCTTCA
TATCCCACGGCTGGCAGAGCTTCCGCATCGCAGCCCCACTCTCTGCCGAGAAGAAATACCGCGTTACGTTGTGATGCAACCGTCC
AGTGCCCGTGGTGTATGCTGGGATGTGTACCTCTTTGATGGCGACCAGATTGTCGTGGTCTGTAAAGGGTCAAGTTCAGCA
AATGAAGCGCACAACTCCAAAGCCTGCTCGGTGTAAGTCCAGCTGCTACGCCAATGTCAAAGCCTATTACCGCCAAATCGACTA
GACCGCACCTGTGGCCGTGAGGAAGGTTGTAGTCACGCAAAGCCCCGGTGTGGTTTCTCGAAAGTCTTGACACAATTGCCTCT
GAAGTTGGTGTGGATGCCAGCGAGCTCAGCGATGATGTAAAGATCAGTGACATCGGTGTAGATTCCCTGCTCAGCATCTCTATCTT
GGGTAGACTGCGGCGGAGACTGTTTGGACTTGTCAAGTTTCATTATTTCATTGAACACCCACTATAGCGGAGCTTCGAGCCTTCT
TCCTTGACAAGATGGTTCGTGCCTCAAGCTACTGTCAATGACGATGATAGCGATGATTTCGTCCGAAGATGGGGGTCCAGGGTTTTCA
AGGTCTCAAAGCAACTCGACCATCTCAACCCAGAAAGAGCCAGACGTCGTTAGTATCCCTTATGTCCATCATCGCACGTGAAGTGGG
CGTTGAAGAGTCAGAAATTCAGCTCTCCACTCCATTCGCCGAAATTTGGCGTAGATTCCCTCTTGACCATCTCCATCTTGGATGCTT
TCAAAACAGAGATAGGAATGAACCTCTCCGCGAATTTCTTCCAGACCATCCCACTGTGCGAGACGTCCAGAAAGCTCTTGGCACC
GCATCAACACCCCAAGGCCACTCGATCTGCCCCCTCACCGGCTTGAACAAAATTCGAAACCATTAAGCCAGAACCTTAGAGCCAA
ATCCGTTCTTCTCCAAAGGCGCCCGAAAAAGGCAAGCAGCATTAATTCCTTCTCCCTGATGGCGCTGGCTCCCTATTCTCTCTACA
TCAGCCTGCCCTCTCTCCCTTCCGGCCTCCCCGTTTACGGTCTCGATTCTCCCTTCCACCACAACCCCTTCCGAGTATACCATCTCC
TTCGCGCGCGTCGCAACCATCTATATCGCCGCCATCCGCGCCATACAGCCCAAGGCCCTACATGCTCGGCGGCTGGTCCCTCGG
CGGCATCCACGCCATACGAAACAGCCCGCCAGCTCATCGAGCAGGGCGAAACTATCTCCAACCTAATCATGATTGACAGCCCTGCC
CTGGCACTCTGCCACCTTTACCAGCCCTACTCTCTCACTCCTGGAGAAAGCTGGCATCTTTGACGGCCTGTCAACATCTGGCGCA
CCTATACCCGAACGCACCCGCTCCATTTCTTAGGCTGTGTGAGAGCGCTGGAGAATTATACCGTTACGCCCTTACCACCAGGTAA
GAGCCCCGGAAGGTCACCGTCATCTGGGCACAAGAAGGCGTGTGGAAGGAAGGAAGAGCAAGGAAAGGAGTATATGGCTGCTA
CGTCGTCTGGTGTCTGAATAAAGATATGGATAAGGCGAAGGAGTGGTTGACGGGGAAGCGGACGAGCTTCGACCCGAGTGGATGG
GATAAATTGACGGGACGGATGTGCATTGTGATGTTGTTGGTGGTAATCACTTTTCTATTATGTTTCCGCTAAGGTATGCTGGAG

ATCCACATTTTCCCTCTCTTCATCGATCGACAATGATACTAACGCTTATAACCTCCAGATTGCAGCCGTTGCTAAAGCTGTGGCTA
CGGGCCTTCCGGAGAAGTAG

Sequence S14: DNA sequence of *6msas* from *Cladonia uncialis* (5' – 3'). DNA highlighted in blue are the exonic regions. Exonic regions were observed by performing RT-PCR on mRNA from *A. oryzae* that was transformed with *6msas*. Reproduced from Bertrand et al. (2019), Supporting Information file, manuscript in preparation. Accession number: AUW31355.

ATGGCTGGAATGGTAGATGACAAGGTCATCTTACCTCGGCCTAAAACGAAATCGCATGGCCCAGCTGCACCGAGCCAACTGGCAAC
TCCTCCATTGACGGCCAGTGCCACACCAATCGATAGTCCGACGGGAACTGCATCCCCGGGTACCACAATGTTGAAATAAGTAGTG
CGCTCTTCTATCCACTTCTCATATGGTCTGACAGCGAACAAGACTGCTCCCTCGGACATTGCTGTTGTGGGCATGGCGTGTGAGT
GGCTGGCGGGAATGACACGCCGGAGAAGCTGTGGGAGTTTATTATGAACCAAAAGGATGGTTCAGCGCAAAATCGACCCCTTTGAGGT
GGGAGCCATATCGCCGAAGAGACCCAAGGAACGCCAAGGAACCTCGACAAGACAACCTTCACGTGGATATTTTGTAGAAAATCTGGAG
CATTTTCGACAGTAGCTTCTTCGGGATATCTCCGAAAAGAGGCCGAGATGATGGATCCTCAACAGAGGTTGTCCCTGGAAGTTTGTCTG
GGAGGCACTCGAAAGCGCTGGTATTCCCCACAAAGCCTAGCAGGCTCGAATACGGCAGTTTACATGGGAGTCAACTCAGACGATT
ATGGCAAGCTGTTACTTGAAGATCTCCCCGGTGTGCAAGCATGGATGGGTATAGGCACGGCATAATTGTGGGATACCTAATCGAATC
TCTTATCACTTAGACCTTATGGGCCCTAGTACAGCTGTTGATGCTGCGTGTGCTTCATCGCTGGTTGCAATTCATCATGGACGCCA
AAGCCTTCTCGCGGGCGAAACAGATCTCGCCATTGTTGGAGGTGTGAATGCATTATATGGTCCCGGATTGACAAAAGTACTCGACC
AGGCTGGTGCAGTGTCCCCAGAAGGCTTGTGCCGGTCTTTTGACGATGAGTCGAAAGGATACGGTCGAGGTGAGGGTGCGGGCATT
GTCATTCTGAGGAGGATGGATGACGCCCTGGACAACGGAGACAACATTCTTGCTGTCTGAAGGGCAGTGCCGTGGGGCAAGATGG
GCACACAAATGGCATCATGGCTCCCAACGGAAAAGCTCAAGAGCTCGTGGCAAAGAATGCTCTCAACACAGCAGGGATCGACCCAA
GTACTGTTGGCTATGTGCAAGCGCATGCCACATCCACTCCTGTGCGTGACCCGTGTGAGGTAAACGCAGTTTCAAATGTCTATGGT
AAAGGGAGGAGTGGCCACGAGCCTTGTACATAGGCTCTATCAAGCCGAACATTGGCCATCTCGAAGCTGGAGCCGGTGTGATGGG
CTTTATCAAGACTGTAATGGCCGTTAACAAAGGAATCATGCCGCCCAAGCAAACTCTCAATAAACTGAATTCAGGGTCAATTGGG
ACGAGGCCGGTGTGCAAGTCGTGCGTGAACCCACACGCTGGAAATCTCCCGAAGCACGCCGTGAGCAGGGATCTGCAGTTACGGA
TATGGCGGGTCTGTTTCGCACGCTGTCATTGAAGAATTCGCTGCCGAACGTGAGGAGTTTCCCGCTCTTGCGCGTATGGATGGCGA
GACTAGTCAGGGACCAAGTATCTTGGTCATCTCAGCACCGCAACAAAAGCGGCTTGCTTCTCAAGCGAAAGCATAACCAATCGTGGG
TCCTTGGGGACGGTAACTTTACAGCCTTTCTTCAATCGCAGCAACGCTGCTACTCGCCGAGGTATCATGATTATCGAGCGCGC
TTTGTGGTTAATTTCTCATGAGGATGCAGCTGATACTCTCAAAGCCTTTTTCAGAAAGTGATAGTGGTGAATGGAGCGTTTCTAGCAA
AGTGTCTGAGAGTGGCGATAGCAGAAAATCTGTCTGGGTATTCTCTGGCCATGGAGCTCAATGGAAGAGATGGGCCAAGAGATGT
TGCAGAACACGGTATTCCGTGAGGCAATAGCTCCGTGGATCCAGTTGTGGTTGACGAAGCAGGCTTTTTCAGCCTTGAAGCTTTG
AAGGTGCGCGATTTTGGGGCTTCGGACGAGGTTCAAATCTCACTTTCGTGATGCAGATTGGGTGGCCGCTGTGCTGCAATCCAA
AGGCGCTGCCCCCTACGCGGTTCATTGGCCATTCACTAGGCGAAATTCAGCTTCTGTGGTTGCTGGAGCTTTGACAGCGAAGGAGG
GAGCCATTATCGTCTGTAGACGCGCGGTCTTATATCGCAAGGTCAATGGGTCTCGGTTTCGATGGTCTGGTTAATGTTCCATTTTCG
GAGATGCAGCAAGAACTCGGTGATCGGCAGGATATTGCTGCAGCCATCGACTCATCCCCATCGTCTTGTGTGGTCTCAGGCGCCGC
CAATGTTGTGCGCGATTTTGCAGAGGTCTGGAAGACCGTGGCATCAGGGCAATTCAAGTGAAGACAGACGTTGGTTTTTCACAGTC
CACTTCTAAACAGCCTTGTGGAGCCACTTTCCAAGGCTCTTGCGCAGTCTCTGAAGCCGAGATCTCCCTCTATCAAACCTTTACTCA
ACGTCGCTCGCAGACCCAGTGGCCAGGATCCCCGCGATATCCGACTAGGACAAACACATGGTCAAGCCTGTTCTACTAACCCCG
AGCAATACAAGCTGCGGCAGAGATGGCTTCCGGTCTTCTTGGAGCCTTTTCGCACCCCATCGTCTCCCAATCAATCGAAGAAA
CGATAGTACAAACAGAGTTGAAGATTATGCAGTCACTCCCAACAATGACTCGAGGAAGTCTGCTGAAAAGAGCATTTCTTACAGC
CTTGCTCAACTCCATTGCAAAGGCGTTCCAATCTCTTGAAGAGCCAAATGCCTCGCGACTGGGCGCCAGGGGTACCCACGATGAG
TTGGAATCATCGACGTTACGTGAAACAGATAGAACTGGACCCATGAACTCAGCGGAAGTACATGACAGCGAGAAGCACACTCTGC
TGGGTGACGCTATTCCAATATTTGGGGAATAATGACGGTGTATACGACAAAGCTCGATCAAAACACCAACCCCTTCTGGAAGT
CACCCGCTGCATGGCACAGAAATCGTCCCTGCTGCAGTGTGATCAATACGTTTACCATGGTACCGGAAAGCGCAACTTGCACGA
TATCGTACTCCATGTCCCTGTTGCAATAAGTGCCCCACGAAGCATTCAGGTCTCCATCCACCAAGCACAGGTGAAGTTGACATCTC
GGCTAATCAAAGCTGAAGAACAATCCTCCGACGAAGAGTCAATGGGTCACTCATACAACCTAGCCGTTTCGCTGATCAAGACGATTCTG
CTGACTGATGGCGTGAAGAGCCAGTTGACATTCAGCCCTCAAGTCTGAGATCGGTACCCAATTGAAAGACACGTTTTCTATCGA
TTATCTCGATGGTGTGGAGTTTTCAGCCATGGGCTTCCCATGGAGTGTGACAGAGCACTTTGGAAACGAAAAGCAATGATAGCCA
AAGTGAACGTGGCTCCAGACGTTGCTGTAGGTTCTGCTCTGCTTGGCATGCTCAATCGTGGGCCCAATATTTCGACGCGAGCACC
TCGATTGGCTCGACGTTATTCTTCAATGAGCCTAGACTGCGCATGCCAGCTCAGATAGGTAGGGTGACCGTTCGATGTCTAATGCTAC
ACCACCTCATATTGGATATATTCTATGTCCGAGAGGCTACCCAGACTGAGTTAGCAGTGCACCTTGATGTTCTCAGCGAGAGTGGCC
AACTATTGGCAAGTATGGAAGACATGAGATTGCGGGAATGGAAGTACGATAGGCGCTAGTGGGAGCGTTGAGAGTCTGGTGCAC
CAAATTTCTTGGACGCCCATGCCTTTGGCTGAACAGCCATTGGCGCTTGGCCACGTGATGCTCATCTCAGAGGACAAGTGGGCAGT
GGAGAGATGCAAGAAGGCCTTGAAGGATCAGGCTACCTCCGTCGTATCCTATCCTTCTATAAATGACCTTCGTCGAGATAACGTGC
AGGCAACCGAGGAGGGCAAAACGGCTCTTCTAATATATATCCCAACCATGTGGAATATTTTGTAGGACGTACCAGCGCGGCTAGG
AAATCTGTACGGAGCTTGTGACAGGTCAAATACGCTGCGGAACAATCCAGCGCATCAAAGCATTGCCATCACTAGCAGCGT
CTCGAAAGGTCAATCTCCAACAGCTCTCGCACAGCGCCCTTACATGGCCTCGGACGAATCATTCGATCGGAGCAATCAGCGACCT
GGGGTGCCTTATAGATAATGAAGATGATGAATGGACGTTCCAGTCCAAGCGATCAAGTACGCCCAAGGCGCAGATGTAATCAGG
CTTGATGATGGCGTAGCAAGAACAGCAGCTTTCGACGGCTTCCAGTGTAGAAATTCATCCGGCCGGCCATCAAGCACACGTTCT

CCCACGTCCCGAGGGAACGTATCTGATCACCGGCGGGCTCGGTGCTCTTGGTCTAGAGGTTGCCAACTTCCTTGTGGAAAATGGCG
CAAGGAGACTCGTGCTGGTAGCACGCCGACATTTACCGCCACGCACGAATGGAAGACGGCCACGGGTCCAGTCGAGAACGATAT
AAAGCGATTGAGCGCTTGGAGGGTCTTGGCGCCACAGTGCATGTTCTTTTCGATCGACCTTACTGCCGAAGGATCGAAGCAAAAGCT
TGAAGCCTCTCTCAATGCTTGTGCTTCTCCATGCTAGGAGTTGTCCACGCGGCAGGCGTGCTTGAAGACCAGCTTATCTGA
ACTCCACAGCCGAATCATTGATCGAGTCTCGTCCCAAGATCGCAGGAAGTTTGGCCTTACACCAGCTCTTTCTGTGGGACT
CTCGACTTCTTTATTCTCTTCTCTTCTGTTGGCCAGCTCTTTGGCTTCCCCGGGCAAGGCTCCTATGCCAGTGGAAACGCCTTTCT
CGATACTCTTGGCAAGCCACGGCGTGAGCAAGGGGACAATTGTGTGGCGGTGCAGTGGACGAGCTGGCGTGGAATGGGTATGGCAG
CTAGCACCGACTTCAATCAAGTACAGAGCTAGAGTCAAAGGGGATCACTGACGTTACGCGTGATGAGGCTTTCCGCGCTGGACACAC
CTGGCCAAATATGATCTTGACCATGGTGTGCTATTGCGCAGTCTGCTCTGGACGAGGATGAGCCCTACCAGCCCCAATCCTAGC
TGATATCGCGGTTCTGAAGCGTCTGCTGGTGGTGTGCTGGTGACGGCGAGACCAAGCAAGGGCAGGGATCTGCCAGCGAGG
TCCCTAAGTCAGGTCCCGAGCTTGATGCATATGTCGATACGGCCATTGCGACAAGCGTTGCTGAAACGCTTCAGATGCCGTCTGCG
GACGAGTGGATCCTCAGGCAGCGTTGTCGGACCTGGGCATGGACTCTGTGATGACTGTTGCGCTGAGACGCCAGCTGCAAAAGGC
CCTTGGCGTCAAGTACCGCCGACACTGACGTGGTCTCATCCAGCGCATCGCATCTACAGAAATGGTTACGGAGCAGATGAAAC
AATAA

Sequence S15: DNA sequence of *t3pks1* from *Cladonia uncialis* (5' – 3'). DNA highlighted in blue are the exonic regions. Exonic regions were observed by performing RT-PCR on mRNA from *A. oryzae* that was transformed with *t3pks1*. Reproduced from Bertrand et al. (2019), Supporting Information file, manuscript in preparation. Accession number: AUW30693.

ATGGCAGCAGTCGATATGTCAGCCCGTCACACCAACATGGTATCCCAGC
CCGACTTCGAAGCTCAGCCACGTACGAAGTACTGCCCTGATGGGACCAAGCATTGCGACGAAAAAGCTTTTCGCGGA
GAGTGATTAGATATTGGGTCCAGTAGATCATCTAGCACGCTTTCAGTTCCAAGGCTGTACCTGGGCTCTACATC
ACTGGCCTAGGATCACAGTATCCGCCATTCTCTTTACGCAGATAGACTCGCAGGTTTCATTAGCAAATGGTACGACC
TTGAGACGCCAGGGTAAGTGAACAATAACCAATACTGGGTTCCGTAATGAGACTGATAATGAACCCAGTATAAAGAAAC
TCTTAACATATCAACCAAAGTACAACCATTCAGACACGTTCAACTATCAAACATCCTGATGATCCCTTCTGGAACCGCTC
CAACGCACCATCGATCTCGGAATTGGACGTGTTCTTTTCGCAAAGCCGGTGTGACCTGACAGTTCAAGCCTGCAAGAAG
GCGCTGCGTGATGCAAATGTCGGCCCTGAATCAATACCCACATTGTAGCCGTGACATGCACAAACGCTGGCAATCCTG
GGTTTGACCTTTTGGTATCCAGAACTCAAACCAAGCCGAAACCGACAGGACACTGCTCCATGGCGTGGGTGTGTC
TGGTGGCCTCTCCGTATGCGCGCTGCAGCCGCCATGGCGCAATCAGCATCCATGCGTGGTTCGCGCAGCCCGTGTGCTC
GTTTTTGCATGTGAAATTTGCAGCATCAACGTTTCGTTGCGACCTCGAAGAAGTAGTCGAACATCCCGAAGATACTAGAA
TATCTCCGGTACTGTTCTCTGACGGCGCTGCAGCTTTCGTTCTCTGCAACGAATTAGCTTCTGAAGCCGAAGACAGGC
CGTTTACTCCCTCAATAATTGGGAAACCGCAACTATGCCCGAGACATCCAAAGAAGTTGAATTTATGACTGACCTCTA
GGTTTCCGCGCTACATTGACAAAAGAAGTGCCCTAACCTAGCTGTCAAGGCCATTGGCCCCATGTTTGAGAAGTTGCGCC
ATTACCTTCCCGCCGACAGCCAGCTCGCCGCGAGCGGTGACTCCCGGAAGCTAAGGAATTGATTGGGCTTTACATCC
AGGTGGCATTGCAATCCTAAAAGGTGTTCAAGATAAATTGCAACTGGAGAAAGAACAGTTACGTGCATCGTATAAGGTG
TATGAAGGCCATGGAATTCATCAAGTCCCACCGTGCTAGTAGTGTTAGATCAGCTGAGGAAGATGGGGGAAGGGAGGG
AGAATGTGATGGCATGTAGTTTTGGACCGGCATGACGATCGAAATGGCAATGTTGAAGCGATGCGTGAGAAAAGAGGA
AGAGTGA

Sequence S16: DNA sequence of *t3pks2* from *Cladonia uncialis* (5' – 3'). DNA highlighted in blue are the exonic regions. Exonic regions were observed by performing RT-PCR on mRNA from *A. oryzae* that was transformed with *t3pks2*. Reproduced from Bertrand et al. (2019), SI file, manuscript in preparation. Accession number: AUW30714.

ATGTCGCCATCTGCTCTCAACGGTAGCACATCACAGCATCGTTTTTGAACAACCTCAACCTTTCTATCGTGGGGTTGGGGACCGA
ATACCCCCCGTTTCAATTGGGACCAGAGGCCCTCGAGACATTGGCCCAGAGATTCTATCCACCGTCTACAGCGTGAGGCCTTCCCG
TGCGACCTCACTTTCTTATCGTCGTCCTCATCCACTCTTCATTGCATTCTTATCGCACATATGCGCTGAAACATCAGCATCTAGG
CTCTCTAAAGTCCCTCCATCAATCGCTTTACCGGTATCGATACTCGTGCTGCAATTGGCACTGTGACCATCCTCTCGTTAACCA
TCCCACGCCCCAGTATCGCTGAATTGAATCAATGTTTTTCGCCAAGAAGGCGTCCGACTTTCCATCAGTGCATGTAAGAAAGCAA
TTGCAGAATGGGGAGGAAGTGTAGATGAAATCACACATGTTGTTTCAACCACTTGTACGAACCTAGCTAACCCTGGCTTCGATCAT
TATGTCACGAAAGGACTGGGTCTGCGGAGCGGAGTTGAGAAAATACTACTGCATGGAATTGGATGCTCAGGGGGACTAGCTGCGAT
AAGAGCGGCTTCGAACCTTGCACTCGGGTCATCGTTTCGGAAGAAGCCTGCAAGAATTCTTGACTGGCATGTGAAATTTCCAGTC
TGCTTGTACGATCGGAACCTTGATTGATCGACAAAGAGCAGAAGACACGCATTGGCGTTTGCCTTTTACGACTGCGCGTCTGCA
GCTGTGCTGAGCAATGGCATGGGGGAATATGAGGTTTACTCTCCAAAATATGAGATACTGGGCTGGCGGCATGAGATATTAGAAGA
CTCTGATCAAGACCTTGGCTTTGATGTTGATCCATTTCGGTATGCCTTAGCGAATGAGAATATTGGATGAATGGGACATGGTTGACA
CCTAACAGGCTGGAAAGTCGTTCTTACGCCTCGAGTGCCATCAATGGCCTCCGCTGCGGTCTCGCCAGCCTTTTACAGGACCTGATCA
AGTCCCTTCCGGAACCTCACCAAGATGGTAGGCTCCCTACCGCGGCCGACTTTGATTGGGCTTTACATCCAGGCGGATCAACGATT
ATTACTGGGCTGGAACAGGCAATGAACCTGACACAAGATCATTGCGAGCAAGCTACGAGATCTATGTCAACTATGGGAACAGCTC
GAGTGTCTACCATAATGTCTGTGATGGACAAACTACCGGATATGAGCGAGGGTAGAGAGTATGTGGTCTGCGCTTGGGCCCTG
GAATATCACTCGAAATGATGATACTCAGAAGGCCAAGGACCATAGTAGATGGATTACCCACTGAAGACGTGGATTAA

Permissions Documents



RightsLink®

Home

Account
Info

Help



Science
AAAS

Title: Basidiomycete yeasts in the
cortex of ascomycete
macrolichens
Author: Toby Spribille, Veera
Tuovinen, Philipp Resl, Dan
Vanderpool, Heimo Wolinski, M.
Catherine Aime, Kevin
Schneider, Edith
Stabentheiner, Merje Toome-
Heller, Göran Thor, Helmut
Mayrhofer, Hanna
Johannesson, John P.
McCutcheon

Publication: Science

Publisher: The American Association for
the Advancement of Science

Date: Jul 29, 2016

Copyright © 2016, Copyright © 2016, American
Association for the Advancement of Science

Logged in as:
Robert Bertrand

LOGOUT

Order Completed

Thank you for your order.

This Agreement between Mr. Robert Bertrand ("You") and The American Association for the Advancement of Science ("The American Association for the Advancement of Science") consists of your license details and the terms and conditions provided by The American Association for the Advancement of Science and Copyright Clearance Center.

Your confirmation email will contain your order number for future reference.

Pertains to: Spribille T, Tuovinen V, Resl P, Vanderpool D, Wolinski H, Aime MC, Schneider K, Stabentheiner E, Toome-Heller M, Thor G, Mayrhofer H, Johannesson H, McCutcheon JP. 2016. Basidiomycete yeasts in the cortex of ascomycete macrolichens. *Science* 353: 488-492. <http://dx.doi.org/10.1126/science.aaf8287>



Title: Putative identification of the usnic acid biosynthetic gene cluster by de novo whole-genome sequencing of a lichen-forming fungus

Author: Mona Abdel-Hameed, Robert L. Bertrand, Michele D. Piercey-Normore, John L. Sorensen

Publication: Fungal Biology

Publisher: Elsevier

Date: March 2016

Copyright © 2015 The British Mycological Society.
Published by Elsevier Ltd. All rights reserved.

Logged in as:
Robert Bertrand

LOGOUT

Please note that, as the author of this Elsevier article, you retain the right to include it in a thesis or dissertation, provided it is not published commercially. Permission is not required, but please ensure that you reference the journal as the original source. For more information on this and on your other retained rights, please visit: <https://www.elsevier.com/about/our-business/policies/copyright#Author-rights>

BACK

CLOSE WINDOW

Copyright © 2019 Copyright Clearance Center, Inc. All Rights Reserved. [Privacy statement](#). [Terms and Conditions](#).

Comments? We would like to hear from you. E-mail us at customercare@copyright.com

Pertains to: Abdel-Hameed M, Bertrand RL, Piercey-Normore M, Sorensen JL. 2016. Putative identification of the usnic acid biosynthetic gene cluster by *de novo* whole-genome sequencing of a lichen-forming fungus. *Fungal Biol.* 120:306-16. <http://dx.doi.org/10.1016/j.funbio.2015.10.009>

Regarding Incident 2501878 Permissions request

support@services.acs.org

Thu 1/17/2019 6:18 PM

To: Robert Bertrand <Robert.Bertrand@umanitoba.ca>;

Dear Dr. Bertrand:

Thank you for contacting ACS Publications Support.

Your permission request is granted and there is no fee for this reuse. In your planned reuse, you must cite the ACS as the source, add this direct link (<https://pubs.acs.org/doi/abs/10.1021%2Facs.jnatprod.6b00257>) and include a notice to readers that further permissions related to the material excerpted should be directed to the ACS.

I hope this information helped. Please let me know if I can be of further assistance.

Sincerely,

Kryxie J. Ramirez
ACS Customer Services & Information
<https://help.acs.org/>

Dear Robert.Bertrand@umanitoba.ca,

Incident Information:

Incident #: 2501878
Date Created: 2019-01-18T07:30:55
Priority: 3
Customer: Robert.Bertrand@umanitoba.ca
Title: Permissions request
Description: Good evening,

I am an author of the following publication in Journal of Natural Products:

M. Abdel-Hameed, R.L. Bertrand, M. Piercey-Normore, J.L. Sorensen (2016) The identification of 6-hydroxymellein synthase and accessory genes in the lichen *Cladonia uncialis*. *Journal of Natural Products* 79: 1645-1650.

I am requesting permission to use this whole article (or parts thereof) in my doctoral thesis (Department of Chemistry, University of Manitoba). The article is marked 'Editor's choice', and for this reason I believe that the Permissions portal has re-directed me to contact you at this e-mail address.

Robert L. Bertrand
Ph.D. Candidate
Department of Chemistry
University of Manitoba

{CMI: MCID611492}

Pertains to: Abdel-Hameed M, Bertrand RL, Piercey-Normore M, Sorensen JL. 2016. The identification of 6-hydroxymellein synthase and accessory genes in the lichen *Cladonia uncialis*. *J. Nat. Prod.* 79: 1645-1650. <http://dx.doi.org/10.1021/acs.jnatprod.6b00257>



RightsLink®

Home

Account Info

Help



ACS Publications
Most Trusted. Most Cited. Most Read.

Title: Lichen Biosynthetic Gene Clusters. Part I. Genome Sequencing Reveals a Rich Biosynthetic Potential
Author: Robert L. Bertrand, Mona Abdel-Hameed, John L. Sorensen
Publication: Journal of Natural Products
Publisher: American Chemical Society
Date: Apr 1, 2018
Copyright © 2018, American Chemical Society

Logged in as:

Robert Bertrand

LOGOUT

PERMISSION/LICENSE IS GRANTED FOR YOUR ORDER AT NO CHARGE

This type of permission/license, instead of the standard Terms & Conditions, is sent to you because no fee is being charged for your order. Please note the following:

- Permission is granted for your request in both print and electronic formats, and translations.
- If figures and/or tables were requested, they may be adapted or used in part.
- Please print this page for your records and send a copy of it to your publisher/graduate school.
- Appropriate credit for the requested material should be given as follows: "Reprinted (adapted) with permission from (COMPLETE REFERENCE CITATION). Copyright (YEAR) American Chemical Society." Insert appropriate information in place of the capitalized words.
- One-time permission is granted only for the use specified in your request. No additional uses are granted (such as derivative works or other editions). For any other uses, please submit a new request.

BACK

CLOSE WINDOW

Copyright © 2019 Copyright Clearance Center, Inc. All Rights Reserved. [Privacy statement](#). [Terms and Conditions](#).
Comments? We would like to hear from you. E-mail us at customercare@copyright.com

Pertains to: Bertrand RL, Abdel-Hameed M, Sorensen JL. 2018. Lichen biosynthetic gene clusters part I: Genome sequencing reveals a rich biosynthetic potential. *J. Nat. Prod.* 81: 723-731. <http://dx.doi.org/10.1021/acs.jnatprod.7b00769>



Title:

Lichen Biosynthetic Gene
Clusters Part II: Homology
Mapping Suggests a Functional
Diversity

Author:

Robert L. Bertrand, Mona Abdel-
Hameed, John L. Sorensen

Publication:

Journal of Natural Products

Publisher:

American Chemical Society

Date:

Apr 1, 2018

Copyright © 2018, American Chemical Society

Logged in as:

Robert Bertrand

LOGOUT

PERMISSION/LICENSE IS GRANTED FOR YOUR ORDER AT NO CHARGE

This type of permission/license, instead of the standard Terms & Conditions, is sent to you because no fee is being charged for your order. Please note the following:

- Permission is granted for your request in both print and electronic formats, and translations.
- If figures and/or tables were requested, they may be adapted or used in part.
- Please print this page for your records and send a copy of it to your publisher/graduate school.
- Appropriate credit for the requested material should be given as follows: "Reprinted (adapted) with permission from (COMPLETE REFERENCE CITATION). Copyright (YEAR) American Chemical Society." Insert appropriate information in place of the capitalized words.
- One-time permission is granted only for the use specified in your request. No additional uses are granted (such as derivative works or other editions). For any other uses, please submit a new request.

BACK

CLOSE WINDOW

Copyright © 2019 Copyright Clearance Center, Inc. All Rights Reserved. [Privacy statement](#). [Terms and Conditions](#).

Comments? We would like to hear from you. E-mail us at customercare@copyright.com

Pertains to: Bertrand RL, Abdel-Hameed M, Sorensen JL. 2018. Lichen biosynthetic gene clusters part II: Homology mapping suggests a functional diversity. *J. Nat. Prod.* 81: 732-748. <http://dx.doi.org/10.1021/acs.jnatprod.7b00770>



Title: Lichen ketosynthase domains are not responsible for inoperative polyketide synthases in Ascomycota hosts

Author: Mona E. Abdel-Hameed, Robert L. Bertrand, Lynda J. Donald, John L. Sorensen

Publication: Biochemical and Biophysical Research Communications

Publisher: Elsevier

Date: 10 September 2018

© 2018 Published by Elsevier Inc.

Logged in as:
Robert Bertrand

LOGOUT

Please note that, as the author of this Elsevier article, you retain the right to include it in a thesis or dissertation, provided it is not published commercially. Permission is not required, but please ensure that you reference the journal as the original source. For more information on this and on your other retained rights, please visit: <https://www.elsevier.com/about/our-business/policies/copyright#Author-rights>

BACK

CLOSE WINDOW

Copyright © 2019 Copyright Clearance Center, Inc. All Rights Reserved. [Privacy statement](#). [Terms and Conditions](#).

Comments? We would like to hear from you. E-mail us at customercare@copyright.com

Pertains to: Abdel-Hameed M, Bertrand RL, Donald LJ, Sorensen JL. 2018. Lichen ketosynthase domains are not responsible for inoperative polyketide synthases in *Ascomycota* hosts. *Biochem. Biophys. Res. Commun.* 503: 1228-1234. <http://dx.doi.org/10.1016/j.bbrc.2018.07.029>

Thank you for your order! A confirmation for your order will be sent to your account email address. If you have questions about your order, you can call us 24 hrs/day, M-F at +1.855.239.3415 Toll Free, or write to us at info@copyright.com. This is not an invoice.

Confirmation Number: 11783500
Order Date: 01/17/2019

If you paid by credit card, your order will be finalized and your card will be charged within 24 hours. If you choose to be invoiced, you can change or cancel your order until the invoice is generated.

Payment Information

Robert Bertrand
robert.bertrand@umanitoba.ca
+1 (204) 804-5569
Payment Method: n/a

Order Details

Journal of industrial microbiology & biotechnology

Order detail ID: 71770655
Order License Id: 4511570434653

ISSN: 13675435
Publication Type: Journal

Volume:

Issue:

Start page:

Publisher: SPRINGER-VERLAG

Author/Editor: Society for Industrial Microbiology (U.S.) ; SOCIETY FOR INDUSTRIAL MICROBIOLOGY

Permission Status:  **Granted**

Permission type: Republish or display content

Type of use: Thesis/Dissertation

[View details](#)

Note: This item will be invoiced or charged separately through CCC's **RightsLink** service. [More info](#)

\$ 0.00

Total order items: 1

This is not an invoice.

Order Total: 0.00 USD

Pertains to: Bertrand RL, Sorensen JL. 2018. A comprehensive catalogue of polyketide synthase gene clusters in lichen fungi. *J. Ind. Microbiol. Biotechnol.* 45: 1067-1081. <http://dx.doi.org/10.1007/s10295-018-2080-y>

Thank you for your order! A confirmation for your order will be sent to your account email address. If you have questions about your order, you can call us 24 hrs/day, M-F at +1.855.239.3415 Toll Free, or write to us at info@copyright.com. This is not an invoice.

Confirmation Number: 11783490
Order Date: 01/17/2019

If you paid by credit card, your order will be finalized and your card will be charged within 24 hours. If you choose to be invoiced, you can change or cancel your order until the invoice is generated.


Payment Information

Robert Bertrand
robert.bertrand@umanitoba.ca
+1 (204) 804-5569
Payment Method: n/a

Order Details

RSC advances

Order detail ID: 71770630
Order License Id: 4511551224431
ISSN: 2046-2069
Publication Type: e-Journal
Volume:
Issue:
Start page:
Publisher: RSC Publishing

Permission Status:  **Granted**
Permission type: Republish or display content
Type of use: Thesis/Dissertation
[View details](#)

Note: This item will be invoiced or charged separately through CCC's **RightsLink** service. [More info](#)

\$ 0.00

Total order items: 1

This is not an invoice.

Order Total: 0.00 USD

Pertains to: Fisch KM. 2013. Biosynthesis of natural products by microbial iterative hybrid PKS-NRPS. *RSC Adv.* 3: 18228-18247. <http://dx.doi.org/10.1039/C3RA42661K>

Thank you for your order! A confirmation for your order will be sent to your account email address. If you have questions about your order, you can call us 24 hrs/day, M-F at +1.855.239.3415 Toll Free, or write to us at info@copyright.com. This is not an invoice.

Confirmation Number: 11783442
Order Date: 01/17/2019

If you paid by credit card, your order will be finalized and your card will be charged within 24 hours. If you choose to be invoiced, you can change or cancel your order until the invoice is generated.

Payment Information

Robert Bertrand
robert.bertrand@umanitoba.ca
+1 (204) 804-5569
Payment Method: n/a

Order Details

Natural product reports

Order detail ID: 71770525
Order License Id: 4511530756309

ISSN: 1460-4752

Publication Type: e-Journal


Volume:

Issue:

Start page:

Publisher: Royal Society of Chemistry

Author/Editor: Royal Society of Chemistry (Great Britain)

Permission Status:  **Granted**

Permission type: Republish or display content

Type of use: Thesis/Dissertation

[View details](#)

Note: This item will be invoiced or charged separately through CCC's **RightsLink** service. [More info](#)

\$ 0.00

Total order items: 1

This is not an invoice.

Order Total: 0.00 USD

Pertains to: Staunton J, Weissman KJ. 2001. Polyketide biosynthesis: A millennium review. *Nat. Prod. Rep.* 18: 380-416. <http://dx.doi.org/10.1039/A909079G>