SHY FUNGI | NEW DRUGS

Functional elucidation
of unassigned biosynthetic pathways
for novel natural products

Olga V. Mosunova

05.06.2024

Propositions

01. The chemical space of non-reduced polyketides in the Lecanoromycetes is nearly exhausted.

(this thesis)

02. Phylogenetic dereplication and comparative genomics of nrPKS pathways is suitable for discovering new chemicals through the identification of uncharacterized pathways.

(this thesis)

03. Combinatorial biocatalysis relying on catalogued functions of tailoring genes from different biosynthetic pathways offers a more promising method for producing new drug leads compared to combinatorial chemistry.

04. Many biosynthetic pathways encoded in the genomes of lichens are evolutionary junk.

05. Art is essential to understand and appreciate science.

06. Restricted use of genetically modified fungi for food manufacturing negatively impacts the environment.

07. The good concept of open access publishing has transformed into a channel of public tax funds to private business.

Propositions belonging to the thesis, entitled:

SHY FUNGI | NEW DRUGS:
Functional elucidation of unassigned biosynthetic pathways
for novel natural products

Olga V. Mosunova
Wageningen, 05.06.2024
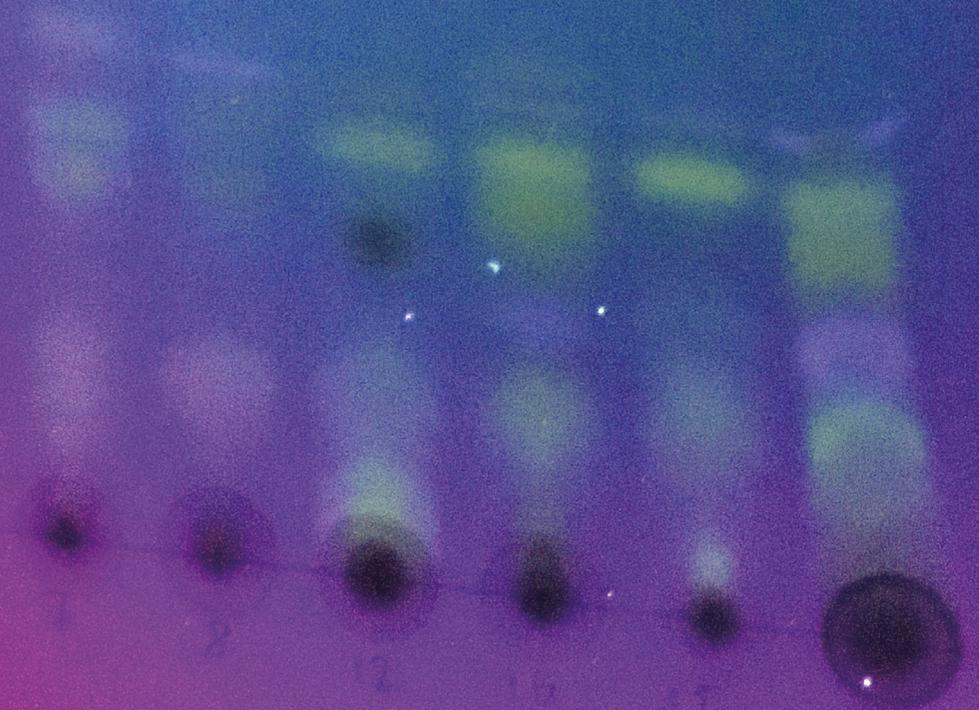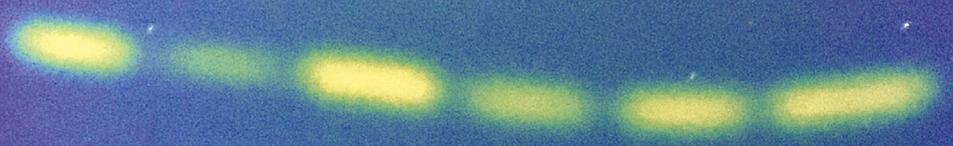
SHY FUNGI | NEW DRUGS:

Functional elucidation
of unassigned biosynthetic pathways
for novel natural products

Olga V. Mosunova

SHY FUNGI | NEW DRUGS:

Functional elucidation
of unassigned biosynthetic pathways
for novel natural products

Olga V. Mosunova

Thesis committee

Promotor

Prof. Dr P. (Pedro) Crous
Special professor Evolutionary Phytoparhology
Wageningen University & Research
Westerdijk Fungal Biodiversity institute

Co-promotor

Dr J. (Jérôme) Collemare
Westerdijk Fungal Biodiversity institute

Other members

Prof. Dr M (Maria) Suarez Diez, Wageningen University and Research
Prof. Dr R. (Russel) J. Cox, Leibniz Universität Hannover, Germany
Dr K. (Kristina) Haslinger, University of Groningen
Dr R. (Robin) A. Ohm, Utrecht University

Olga V. Mosunova

SHY FUNGI | NEW DRUGS:

functional elucidation of unassigned biosynthetic pathways for novel natural products

1. General introduction and scope of the thesis

Humanity and limits to growth

Around 5,000 years ago, the global human population is estimated to have reached 12 million people [01]. By the year 1, this number had increased to approximately 190 million people worldwide, marking a gradual growth. However, in the last 400 years, we have witnessed an unprecedented growth of the human population, from 600 million in the year 1600 to 8 billion in 2022 [02].

The projected future growth of the population presents new challenges for agriculture, medicine, and food manufacturing. The interconnected global society faces threats such as the spread of diseases, exemplified by recent global pandemics. The emergence of multi-resistant bacteria, as highlighted in the World Health Organization's warning list, poses a significant concern. If no changes are made by 2050, antimicrobial resistance (AMR) is estimated to claim 10 million lives annually, surpassing the toll of SARS-CoV-2 by three times (WHO, https://www.who.int/data/stories/the-true-death-toll-of-covid-19-estimating-global-excess-mortality). About 1.2 million humans died from infections caused by multi-resistant microbes in 2019 [03].

Over the past three decades, we have encountered a "discovery void" in the realm of antibiotics, with no structurally new antibiotics introduced to the medical market. This void is partly attributed to research strategies that were relying on reusing the same panel of source strains, and combinatorial chemistry [04]. Exploration of the chemical space of Streptomyces bacteria often yielded the same molecules repeatedly. The pharmaceutical industry invested in combinatorial chemistry to generate large numbers of compounds by combining scaffolds with functional groups that react with each other. The limitations of combinatorial chemistry became evident, and it provided unacceptably low number of molecules that could have been developed

into marketed drugs [05]. These strategies failed to deliver structurally new chemical moieties, lacking reliance on evolution. Evolution sets biology aside from exact sciences like mathematics and physics. Under evolutionary pressure plants, bacteria and fungi produce of bioactive molecules, which requires trade-off of building blocks with primary metabolism, and can arguably be deemed energetically costly [06,07]. Often their molecular structure violates the "Lipinski rule of five" used in pharmacology for selecting potential drug leads [04], indicating that naturally occurring bioactive molecules could be a valuable source for the development of clinically successful compounds.

Among other actions, finding new antimicrobials that structurally differ from the existing ones provide a better chance to get effective medical treatment. Plants and microbes, such as bacteria and fungi are potential source of bioactive molecules, or natural products (NPs). Around 40% of the developed therapeutics drugs approved by the US Food and Drug Administration (FDA) are NP-derived or NP-inspired chemicals. Recent genomic studies have revealed that fungal genomes encode for a spectrum of pathways that offer them evolutionary benefit, which can be exploited as a source of biologically active molecules [08].

NPs refer to compounds that are not directly required for growth and reproduction called secondary metabolites (SMs). Biosynthetic pathways for SMs in fungi consist of genes that are often physically colocalized in the genome, and co-regulated [09]. These clusters of genes are called biosynthetic gene clusters (BGCs). Typically, a BGC contains one or more core genes which assemble the backbone of the molecule, and one or several tailoring genes that modify the molecule scaffold further. Core genes are classified based on the precursors they recognize and incorporate into the backbone: polyketide synthases (PKSs) assemble polyketides from acetyl- and malonyl-CoA, non-

ribosomal peptide synthases (NRPSs) recognize proteinogenic and non-proteinogenic amino acids and condensate them into non-ribosomal peptides. Terpene cyclases (TCs) recognize molecules assembled by condensation of isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP) from acetyl-CoA, and assemble diverse terpenes [10]. Dimethylallyltryptophan synthases (DMATs) assemble indole alcaloids from tryptophan, tryptamine and indole-3-glycerol phosphate 4-dimethylallyltryptophan (4-DMAT).

Genomes of Lecanoromycetes as a source of potentially novel bioactive molecules

Lichens are symbiotic organisms that typically consist of a photobiont (algae or cyanobacteria), and a mycobiont (fungus) mostly belonging to Lecanoromycetes. Lichens can be found in diverse and extreme environments across the globe. Lichens possess unique capability for adaptation, which has led to significant physiological differences between the fungal and algal partners, and their free-living counterparts [11]. The spectrum of molecules involved in complex interactions within a lichen and its environment is remarkable. Lichens compounds have been a subject of extensive research and are a focal point in traditional Chinese medicine, dyeing and fragrance industry [12, 13]. Lichen compounds are produced by mycobionts, and most of those chemicals structurally relate to polyketides [14, 15]. Lichen depsides and depsidones such as atranorin, usnic acid and evernic acid exhibit antimicrobial [16], antiviral [17], and anti-inflammatory activities [18]. Usnic acid and its derivatives are used as a photoprotective agents [19], and it also has shown potential as an antibiotic and antitumor agent [20]. The lichens *Pseudevernia furfuracea* and *Evernia prunastri* are extensively used in perfume

industry [21].

Despite that more than 1000 compounds have been reported for lichens [22], the full compound-producing potential of Lecanoromycetes remains largely unexplored. At the time this PhD work was performed, only about 28 species of Lecanoromycetes had been sequenced (now 98 genomes can be found accessible on NCBI), but the available data on their genomes revealed that they possess high amount of pathways encoded per genome, compared to the rest of Ascomycota [23]. Genome of *Usnea florida* contains striking amount of 120 pathways [23]. This high abundance of pathways which encode PKSs in genomes of lichen forming fungi (LFF) does not directly reflect the structural diversity of the polyketides (PKs) produced by LFF: most of the produced PKs bear aromatic ring and are assembled by non-reducing PKSs (nrPKSs). The genomes of lichens number more PKSs than other types or core genes [24], which is consistent with metabolomic studies [15]. Fewer molecules of other classes have been identified for Lecanoromycetes, although their genomes indicate that more reduced polyketides, non-ribosomal peptides, terpenes and alkaloids can be theoretically produced by these fungi [24]. Such discrepancy suggests that there are many unexplored pathways in lichen genomes that can potentially produce previously unreported molecules.

The abundance of pathways encoded in the genomes of LFF requires strategies for their prioritization for functional studies. The first step required for prioritization of pathways is linking already known compounds to their respective pathways. As the majority of extrolites of LFF structurally relate to polyketides [14], most existing strategies were employed to delimit pathways with PKSs as core genes. Phylogenetic analyses of fungal PKSs is a useful approach for exploring the chemical diversity of fungal polyketides, as enzymes in the same monophyletic clade often exhibit similar enzymatic activities,

leading to the production of identical polyketide backbones. The phylogenetic position of PKSs in regards to already characterized PKSs together with comparative genomics, metabolomics and gene expression studies provide useful information for linking gene clusters to compounds [23]. Expression analysis of nrPKSs concurrent with production of usnic acid, and features of molecular structure of usnic acid allowed identifying a single candidate nrPKS in the genome of *C. uncialis* that is consistent with usnic acid structure [25]. Correlation of gene expression to grayanic acid production revealed putative grayanic acid BGC in *Cladonia grayi* with a core nrPKS *CgrPKS16* [26]. Comparative genomics studies demonstrated that the genomes of *E. prunastri, P. furfuracea* [14] and *C. uncialis* [27] also encode for grayanic acid BGC, and the candidate usnic acid BGC was identified in genomes of *E. prunastri, P. furfuracea* [14]. A BGC for the depside atranorin was identified by profiling the genomes of *Stereocaulon alpinum* and *Cladonia* strains that produce atranorin, for a common BGC, which was functionally characterized using heterologous expression [28]. Genetic studies indicate that genomes of LFF may encode for polyketides that have not been reported for a given species yet, and allow proposing their putative structures. Genome of *C. uncialis* encodes for BGC with a putative 6-hydoxymellein synthase *cu-pks-4*, which was linked to the respective backbone by homology search, revealing *C. uncialis* BGC similarity to the terrein BGC of *Aspergillus terreus* [29]. Based on the comparison of genetic context around *cu-pks-4* and *terA*, the authors proposed a structure to the hypothetical oxidized, methylated, and halogenated derivative of 6-hydroxymellein that may be produced by the identified BGC of *C. uncialis*.

Lichen mycobionts can be grown in culture, although extremely slow growth of fungal partners in the lab hampers functional studies. Spectrum of metabolites that are produced in culture often differ from those of the original lichen, and biosynthetic pathways of the cultured fungi remain inactive, or silent. It is challenging to identify and simulate specific environmental stimuli to trigger expression of silent pathways, and no universal stimulant has been found [30]. Together, these factors make functional studies of BGCs from lichens difficult. No successful genetic manipulation or a pathway functional study performed directly in a lichen mycobionts has been reported to the date. Conventional tools like promoter swap, gene deletion or disruption are therefore not developed for LFF. It must be noted that a proof of principle protocols for *Agrobacterium*-mediated transformation has been developed for *Umbilicaria muehlenbergii* [31] and *C. macilenta* [32]. Heterologous expression of biosynthetic pathways in a fungal host of choice is a reliable strategy that enables linking genes to respective compounds, allowing to circumvent constraints of the native organism. *Saccharomyces cerevisiae* [33, 34], and *Ascochyta rabiei* [28] hosts have been employed for elucidation of lecanoric acid and atranorin pathways from LFF. *Aspergillus oryzae* is frequently used as a heterologous host for elucidation of pathways from fungi [35,36]. However, attempts for heterologous expression of putative 6-methylsalicylic acid and orsellinic acid PKSs from *C. uncialis* in *A. oryzae* only achieved heterologous transcription, and no compounds has been found to be produced [37]. The same host did produce orsellinic acid and 6-MSAS when orthologs of the initially selected lichen genes from *Penicillium* sp. and *Fusarium* sp. were expressed [37]. Therefore, expression of orthologs of genes from mycobiont BGCs could provide access to the potential of lichens to produce compounds, even if expression of BGCs directly from lichens is not achievable for a chosen heterologous host.

Project aims and thesis outline

This thesis narrates to the chemical space of lichenizing fungi, exploring their genomes and compounds that could be produced by the encoded biosynthetic pathways. We recognize lichens as potential source of bioactive molecules, including structurally unique and/or potentially new antibiotics. We explore nrPKSs encoded in lichen genomes, dereplicate pathways and characterize selected BGCs using heterologous expression in *A. oryzae*.

In **Chapter 2** we provide current understanding of secondary metabolism in fungi, summarize existing bioinformatic tools for detection/prediction of biosynthetic gene clusters (BGCs), provide an overview of engineering strategies used to activate silent pathways, heterologously express selected pathways, and discuss synthetic biology future perspectives in relation to biosynthetic pathways in fungi.

In **Chapter 3** we employ a powerful strategy of evolution-guided pathway prioritization to the genomes of eight lichenizing fungi, which provided with a set of two phylogenetic clades containing nrPKS. These nrPKSs are phylogenetically unrelated to earlier identified clades, or groups. We identify a gene cluster family (GCF) that consists of set of genes conserved across Ascomycota pathways that contain orthologs of the identified nrPKSs from Lecanoromycetes. We characterize the orthologous nrPKS of *Aspergillus parvulus* by heterologous expression in *A. oryzae* NSAR1, which together with metabolic profiling of *A. parvulus* enabled us to assign the GCF to biosynthesis of naphthalenones in fungi. We further predict pathway boundaries as defined by gene expression in conducive compared to non-conducive conditions. A

putative pathway to naphthalenones is proposed, and involvement of tailoring genes in modifications of the chemical backbone is discussed.

In **Chapter 4** we further elucidated the naphthalenone pathway of *A. parvulus* using heterologous expression. We expressed combinations of tailoring genes that correspond to steps of the proposed pathway and yields asparvenone and parvulenone derivatives. Structures of selected resulting molecules are characterized using 2D NMR. We have identified previously unreported derivatives, and characterized stereoselective properties of FAD-binging oxidoreductase Apr2. Proposed chemical steps of naphthalenone pathway are discussed in the light of the obtained data.

In **Chapter 5** we examined the genomes of nine LFF and identified one phylogenetically distinct clade consisting of three nrPKSs. We used comparative genomics to identify tailoring genes that are conserved within the GCF. Characterization of homologous pathway of *Aspergillus melleus* yields orsellinic acid and orcinol, leading to the assignment of the corresponding GCF to the biosynthesis of orsellinates. We characterize individual steps of the pathway by co-expressing combinations of tailoring genes, and identify pathway intermediate with antibiotic and demelanizing activity. We discuss the putative mode of action and future perspectives.

Chapter 6 considers most important findings of the present thesis and suggests further development.

References

01. **Adam D. (2022).** World population hits eight billion - here's how researchers predict it will grow. *Nature*.

https://doi.org/10.1038/d41586-022-03720-6

02. **Taagepera, R., & Nemčok, M. (2023).** World population growth over millennia: Ancient and present phases with a temporary halt in-between. *The Anthropocene Review*, 0(0).

https://doi.org/10.1177/20530196231172423

03. **Antimicrobial Resistance Collaborators (2022).** Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *Lancet (London, England)*, 399(10325), 629–655.

https://doi.org/10.1016/S0140-6736(21)02724-0

04. **Silver L. L. (2011).** Challenges of antibacterial discovery. *Clinical microbiology reviews*, 24(1), 71–109.

https://doi.org/10.1128/CMR.00030-10

05. **Kennedy J. P., Williams L., Bridges T.M., Daniels R.N., Weaver, D. Lindsley C.W. (2008).** Application of combinatorial chemistry science on modern drug discovery. *Journal of Combinatorial Chemistry,* 10 (3), 345-354.

https://doi.org/10.1021/cc700187t

06. **Gershenzon, J. (1994).** Metabolic costs of terpenoid accumulation in higher plants. *J Chem Ecol* 20, 1281–1328

https://doi.org/10.1007/BF02059810

07. **Shwab, E. K., & Keller, N. P. (2008).** Regulation of secondary metabolite production in filamentous ascomycetes. *Mycological research,* 112(Pt 2), 225–230.

https://doi.org/10.1016/j.mycres.2007.08.021

08. **Robey, M. T., Caesar, L. K., Drott, M. T., Keller, N. P., & Kelleher, N. L. (2021).** An interpreted atlas of biosynthetic gene clusters from 1,000 fungal genomes. *Proceedings of the National Academy of Sciences of the United States of America*, 118(19), e2020230118.

https://doi.org/10.1073/pnas.2020230118

09. **Keller N.P. (2019).** Fungal secondary metabolism: regulation, function and drug discovery. *Nat Rev Microbiol*. 17(3):167-180.

https://doi.org/10.1038/s41579-018-0121-1

10. **González-Hernández, R.A., Valdez-Cruz, N.A., Macías-Rubalcava, M.L. et al. (2023).** Overview of fungal terpene synthases and their regulation. *World J Microbiol Biotechnol* 39, 194

https://doi.org/10.1007/s11274-023-03635-y

11. **Kono, M., Kon, Y., Ohmura, Y. et al. (2020).** In vitro resynthesis of lichenization reveals the genetic background of symbiosis-specific fungal-algal interaction in *Usnea hakonensis*. *BMC Genomics* 21, 671.

https://doi.org/10.1186/s12864-020-07086-9

12. **Joulain D., Tabacchi R. (2008).** Lichen extracts as raw materials in perfumery. Part 1: Oakmoss. *Flavour Fragr. J. 2009, 24, 49–61*.

https://doi.org/10.1002/ffj.1916

13. **Purvis, O.W., & Pawlik-Skowrońska, B. (2008).** Chapter 12 Lichens and metals. *British Mycological Society Symposia Series,Academic Press,27,2008,175-200*.

https://doi.org/10.1016/S0275-0287(08)80054-9

14. **Calchera, A.; Dal Grande, F.; Bode, H.B.; Schmitt, I. (2019).** Biosynthetic Gene Content of the 'Perfume Lichens' *Evernia prunastri* and *Pseudevernia furfuracea*. *Molecules* 2019, 24, 203.

https://doi.org/10.3390/molecules24010203

15. **Ren, M., Jiang, S., Wang, Y., Pan, X., Pan, F., & Wei, X. (2023).** Discovery and excavation of lichen bioactive natural products. *Frontiers in microbiology,* 14, 1177123.

https://doi.org/10.3389/fmicb.2023.1177123

16. **Harikrishnan, A., Veena, V., Lakshmi, B., Shanmugavalli, R., Theres, S., Prashantha, C. N., Shah, T., Oshin, K., Togam, R., & Nandi, S. (2021).** Atranorin, an antimicrobial metabolite from lichen *Parmotrema rampoddense* exhibited *in vitro* anti-breast cancer activity through interaction with Akt activity. *Journal of biomolecular structure & dynamics*, 39(4), 1248–1258.

https://doi.org/10.1080/07391102.2020.1734482

17 . **Oh, E., Wang, W., Park, KH. et al. (2022).** (+)-Usnic acid and its salts, inhibitors of SARS-CoV-2, identified by using in silico methods and in vitro assay. *Sci Rep* 12, 13118

https://doi.org/10.1038/s41598-022-17506-3

18 . **Lee, S., Suh, Y. J., Yang, S., Hong, D. G., Ishigami, A., Kim, H., Hur, J. S., Chang, S. C., & Lee, J. (2021).** Neuroprotective and anti-inflammatory effects of evernic acid in an MPTP-induced Parkinson's disease model. *International journal of molecular sciences*, 22(4), 2098.

https://doi.org/10.3390/ijms22042098

19 . **Galanty, A., Popiół, J., Paczkowska-Walendowska, M., Studzińska-Sroka, E., Paśko, P., Cielecka-Piontek, J., Pękala, E., & Podolak, I. (2021).** (+)-Usnic Acid as a Promising Candidate for a Safe and Stable Topical Photoprotective Agent. *Molecules* (*Basel, Switzerland*), 26(17), 5224.

https://doi.org/10.3390/molecules26175224

20 . **Araújo, A. A., de Melo, M. G., Rabelo, T. K., Nunes, P. S., Santos, S. L., Serafini, M. R., Santos, M.R., Quintans-Júnior, L. J., & Gelain, D. P. (2015).** Review of the biological properties and toxicity of usnic acid. *Natural product research*, 29(23), 2167–2180.

https://doi.org/10.1080/14786419.2015.1007455

21 . **Lutzoni, F., & Miadlikowska, J. (2009).** Lichens. *Current biology* : CB, 19(13), R502–R503.

https://doi.org/10.1016/j.cub.2009.04.034

22 . **Shrestha G., St. Clair L. L. (2013).** Lichens: a promising source of antibiotic and anticancer drugs. *Phytochem. Rev.* 12, 229–244.

https://doi.org/10.1007/s11101-013-9283-7

23 . **Gill, H., Sorensen, J.L., Collemare, J. (2023).** Lichen Fungal Secondary Metabolites: Progress in the Genomic Era Toward Ecological Roles in the Interaction. In: Scott, B., Mesarich, C. (eds) Plant Relationships. The Mycota, vol 5. Springer, Cham.

https://doi.org/10.1007/978-3-031-16503-0_7

24 . **Mosunova O., Navarro-Muñoz J.C., Collemare J. (2021).** The biosynthesis of fungal secondary metabolites: from fundamentals to biotechnological applications, p 458–476. In Zaragoza Ó, Casadevall A (ed), Encyclopedia of mycology. Elsevier, Oxford, United Kingdom.

https://doi.org/10.1016/ B978-0-12-809633-8.21072-8

25 . **Abdel-Hameed, M., Bertrand, R. L., Piercey-Normore, M. D., & Sorensen, J. L. (2016).** Putative identification of the usnic acid biosynthetic gene cluster by de novo whole-genome sequencing of a lichen-forming fungus. *Fungal biology,* 120(3), 306–316.

https://doi.org/10.1016/j.funbio.2015.10.009

26 . **Armaleo D, Sun X., Culberson C. (2011).** Insights from the first putative biosynthetic gene cluster for a lichen depside and depsidone. *Mycologia, 103:4, 741-754.*

https://doi.org/10.3852/10-335

27 . **Bertrand R.L., Abdel-Hameed M., and Sorensen J.L. (2018).** *Journal of Natural Products* 81 (4), 723-731.

https://doi.org/10.1021/acs.jnatprod.7b00769

28 . **Kim, W., Liu, R., Woo, S., Kang, K. B., Park, H., Yu, Y. H., Ha, H. H., Oh, S. Y., Yang, J. H., Kim, H., Yun, S. H., & Hur, J. S. (2021).** Linking a Gene Cluster to Atranorin, a Major Cortical Substance of Lichens, through Genetic Dereplication and Heterologous Expression. *mBio*, 12(3), e0111121.

https://doi.org/10.1128/mBio.01111-21

29 . **Abdel-Hameed, M., Bertrand, R. L., Piercey-Normore, M. D., & Sorensen, J. L. (2016).** Identification of 6-hydroxymellein synthase and accessory genes in the lichen *Cladonia uncialis*. *Journal of Natural Products* 2016 79 (6), 1645-1650.

https://doi.org/10.1021/acs.jnatprod.6b00257

30 . **Covington, B. C., Xu, F., & Seyedsayamdost, M. R. (2021).** A Natural Product Chemist's Guide to Unlocking Silent Biosynthetic Gene Clusters. *Annual review of biochemistry,* 90, 763–788.

https://doi.org/10.1146/annurev-biochem-081420-102432

31 . **Park, S. Y., Jeong, M. H., Wang, H. Y., Kim, J. A., Yu, N. H., Kim, S., Cheong, Y. H., Kang, S., Lee, Y. H., & Hur, J. S. (2013).** *Agrobacterium tumefaciens*-mediated transformation of the lichen fungus, *Umbilicaria muehlenbergii*. *PloS one, 8*(12), e83896.

https://doi.org/10.1371/journal.pone.0083896

32 . **Liu, R., Kim, W., Paguirigan, J. A., Jeong, M. H., & Hur, J. S. (2021).** Establishment of *Agrobacterium tumefaciens*-Mediated Transformation of *Cladonia macilenta*, a Model

Lichen-Forming Fungus. Journal of fungi (*Basel, Switzerland*), 7(4), 252.

https://doi.org/10.3390/jof7040252

33. **Harvey, C. J. B., Tang, M., Schlecht, U., Horecka, J., Fischer, C. R., Lin, H. C., Li, J., Naughton, B., Cherry, J., Miranda, M., Li, Y. F., Chu, A. M., Hennessy, J. R., Vandova, G. A., Inglis, D., Aiyar, R. S., Steinmetz, L. M., Davis, R. W., Medema, M. H., Sattely, E., Khosla, C., St Onge, R.P., Tang, Y., Hillenmeyer, M. E. (2018).** HEx: A heterologous expression platform for the discovery of fungal natural products. *Science advances*, 4(4), eaar5459.

https://doi.org/10.1126/sciadv.aar5459

34. **Kealey, J. T., Craig, J. P., & Barr, P. J. (2021).** Identification of a lichen depside polyketide synthase gene by heterologous expression in *Saccharomyces cerevisiae*. *Metabolic engineering communications*, 13, e00172.

https://doi.org/10.1016/j.mec.2021.e00172

35. **de Mattos-Shipley, K. M. J., Lazarus, C. M., & Williams, K. (2022). I**nvestigating Fungal Biosynthetic Pathways Using Heterologous Gene Expression: *Aspergillus oryzae* as a Heterologous Host. *Methods in molecular biology (Clifton, N.J.), 2489*, 23–39.

https://doi.org/10.1007/978-1-0716-2273-5_2

36. **Yang, H., Song, C., Liu, C., & Wang, P. (2024).** Synthetic Biology Tools for Engineering *Aspergillus oryzae*. *Journal of fungi* (*Basel, Switzerland*), 10(1), 34.

https://doi.org/10.3390/jof10010034

37. **Bertrand RL, Sorensen JL. (2019).** Lost in translation: challenges with heterologous expression of lichen polyketide synthases. *ChemistrySelect 4*: 6473–6483.

https://doi.org/10.1002/slct.201901762

2. The biosynthesis of fungal secondary metabolites: from fundamentals to biotechnological applications

Olga V. Mosunova[01]

Jorge C Navarro-Muñoz[01]

Jérôme Collemare[01]

[01] Westerdijk Fungal Biodiversity Institute, Utrecht, The Netherlands

Introduction

Secondary metabolites (SMs) are natural products synthesized mainly by bacteria, fungi and plants. They are molecules of low molecular weight with diverse chemical structures and biological activities. The name secondary metabolite originates from the initial observation that their production is not necessary for the growth and reproduction of organisms, in contrast to primary metabolites which include lipids, amino acids, carbohydrates and nucleic acids. However, SMs are far from being secondary and the term "specialized metabolites" is emerging to describe them. It is now accepted that SMs play key roles in the survival of the organisms that produce them because SMs determine interactions within their environment. Nowadays, SM production is a major research field for organic chemists, molecular biologists and bioinformaticians alike. In this review, we will focus on SMs produced by fungi to introduce the fundamental concepts of their biosynthesis, present how the genomics era has impacted the study of fungal SMs, and report the biotechnological tools that have been developed to engineer SM biosynthetic pathways.

Fungi Produce Bioactive Secondary Metabolites That Impact Human Societies

Fungal Secondary Metabolites in Human Daily Life

Fungal SMs have impacted human societies both positively and negatively since ancient times. The episode known as Saint Anthony's fire in 1039 was an outbreak of ergotism, a disease caused by the consumption of rye that

we know now was infected by the fungus *Claviceps purpurea* (van Dongen and de Groot, 1995). This disease is actually caused by SMs called ergot alkaloids that are secreted by the fungus. Many other so-called mycotoxins have been reported and their presence in food and feed is strictly regulated and monitored. The most frequently encountered mycotoxins worldwide are produced by *Fusarium* species on cereals and include deoxynivalenol (DON), fumonisin and zearalenone (Biomin mycotoxin survey see "Relevant Websites Section"). Aflatoxins produced by *Aspergillus* species are also an important threat, which are famous to have caused in England the death of 100,000 turkeys that consumed contaminated peanut meal (Richard, 2008).
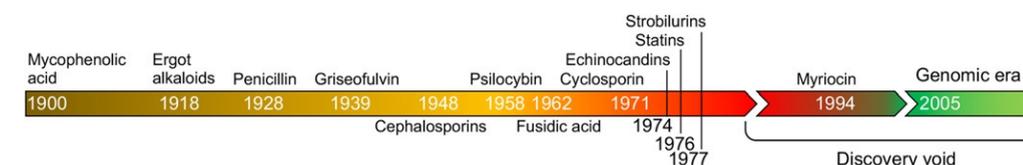
Fig. 01.

Discovery of fungal secondary metabolites with commercial application. Since the end of the 1980s, we have entered a so-called discovery void due to companies stopping the prospection of natural products and investing in chemical synthesis. Since the release of the first genomes of filamentous fungi at the beginning of the 21st century, the genomic era has revived interest in fungal secondary metabolites.

Although some fungal SMs are mycotoxins, many others exhibit beneficial activities to humans. Hundreds of fungal species, mostly basidiomycetous edible mushroom-forming species, have been used in traditional Chinese medicine for thousands of years. Their beneficial properties are linked to the production of mixtures of SMs with diverse biological activities, for example against cancer, diabetes and cardiovascular

diseases (Hyde et al., 2019). In the Western world, fungal SMs have revolutionized modern medicine more than once **(Fig. 01)**. The most famous example is the discovery of the first large-spectrum antibiotic penicillin produced by *Penicillium rubens* (formerly *P. chrysogenum*; (Houbraken et al., 2011)). Penicillin belongs to a family of compounds known as beta-lactams, which still accounts for the largest antibiotic market size. The discovery of cyclosporin A produced by *Tolypocladium inflatum* brought another revolution in the second half of the 20th century: this immunosuppressive SM allowed successful organ transplantation (Colombo and Ammirati, 2011). The statin family of compounds, used as cholesterol-lowering agents, was also discovered in the 1970s and several compounds of fungal origin, like monacolin K, actually represent the largest market of drugs, reaching an annual turn-over circa US$50 billion (Hyde et al., 2019). Fungal SMs find many other applications in diverse fields, including agriculture (e.g., strobilurin fungicides are all derived from SMs produced by the basidiomycetous mushroom *Strobilurus tenacellus* (Bartlett et al., 2002)) and industry (e.g., *Monascus purpureus* is exploited for the production of red rice powder in Asia because it produces red pigments like rubropunctamine (Mukherjee and Singh, 2011)).

Since the 1990s, the development of natural products into marketed active molecules has entered a so-called discovery void **(Fig. 01)** because the same molecules of natural origin were rediscovered, which turned companies to redirect their efforts towards chemical synthesis of analogs of existing compounds (Li and Vederas, 2009). Despite having provided key compounds to human societies, the fungal kingdom had been neglected in the search for natural products. But the genomic era promises to bring fungi in the spotlight again, especially to find solutions to the emergence of multi-resistant bacterial and fungal pathogens (Harvey et al., 2015). Academic laboratories and small and medium enterprises have initiated new exploitation of the fungal kingdom

for bioactive compounds, and these initiatives represent the beginning of a long-term research effort.

Fungal Secondary Metabolites in Fungal Daily Life

While fungal SMs exhibit biological activities that are useful or detrimental to humans, these activities are most often side effects and do not reflect the biological function for the organisms that produce them. It is now well accepted that SMs have important functions in the interactions between fungi and their environments, especially by contributing to the colonization of specific ecological niches. Fungal pigments like dihydroxynaphthalene (DHN) melanin and cladofulvin provide protection against environmental stresses such as UV light, desiccation and extreme temperatures (Langfelder et al., 2003; Dadachova et al., 2007; Griffiths et al., 2018). Fungal SMs are particularly important for pathogens to establish disease. DHN melanin protects human pathogens from the immune system (Chai et al., 2010; Thywißen et al., 2011), while this SM is required for certain plant pathogens to mechanically penetrate plant tissues (Chumley and Valent, 1990). Toxic SMs are particularly well known to contribute to plant infection. Host-specific toxins (e.g., HC-toxin, AK-toxin) produced by *Cochliobolus* and *Alternaria* species determine their host range by targeting plant cultivars that carry a sensitivity gene (Stergiopoulos et al.,2013). Non-specific toxins are more widely distributed and also contribute to the virulence of many fungi. One of the most studied class of toxins are perylenequinones, photoactive toxins that make *Cercospora nicotianae* pathogenic on tobacco (Choquer et al., 2005; Ebert et al., 2019). More recently, it is becoming clear that SMs produced by plant pathogens play more diverse and subtle roles in virulence which are not
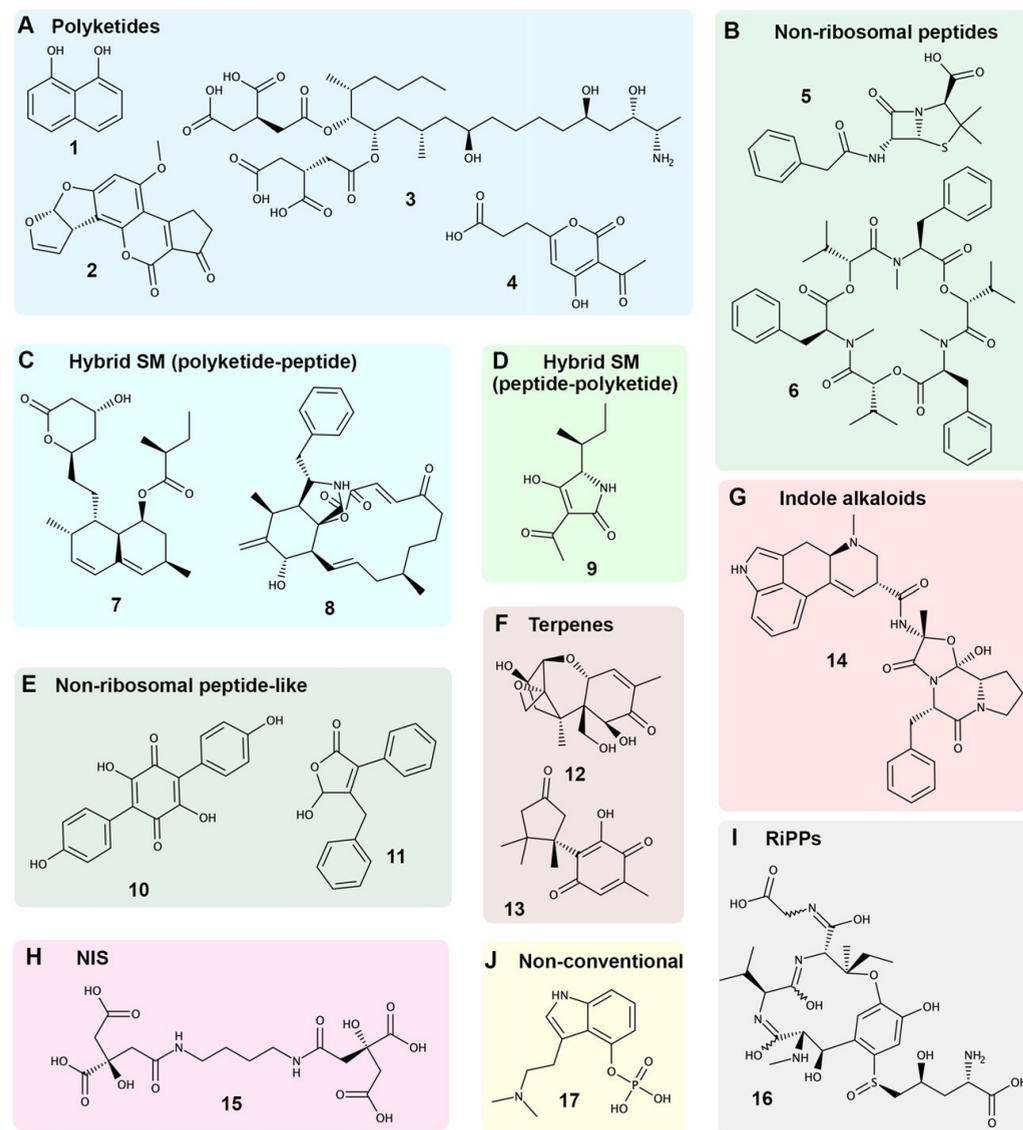
Fig. 02.

Diversity of chemical structures observed for fungal secondary metabolites (SMs). Examples of fungal secondary metabolites. (A) polyketides (1) 1,8-dihydroxynaphthalene; (2) aflatoxin B1; (3) fumonisin B; (4) cyperone B1, (B) non-ribosomal peptides (5) penicillin G; (6) beauvericin, (C) hybrid polyketide-peptides (7) monacolin K; (8) cytochalasin A, (D) hybrid peptide-polyketides (9) tenuazonic acid, (E) non-ribosomal peptide-like (10) atromentin; (11) microperfuranone, (F) terpenes (12) deoxynivalenol; (13) lagopodin B, (G) indole alkaloids (14) ergotamine, (H) NRPS-independent siderophore (NIS) (15) rhizoferrin, (I) Ribosomally synthesized and post-translationally modified peptides (RiPPs) (16) ustiloxin, and (J) unconventional compounds (17) psilocybin.

straightforward to study (Collemare et al., 2019). Symbiotic fungi also appear to rely on the production of SMs to establish a mutualistic interaction with their hosts. The ectomycorrhizal fungus *Laccaria bicolor* secretes terpenes that were shown to be required for symbiosis establishment (Ditengou et al., 2015), while the symbiont *Epichloë festucae* produces indole alkaloid SMs that protect its ryegrass host from herbivorous animals (Schardl et al., 2013). Such a defensive biological function can be extended to antimicrobial compounds produced by fungi when they compete with other microorganisms living in the same ecological niche (O'Brien and Wright, 2011). This function was recently experimentally proven with bikaverin and beauvericin produced by *Fusarium fujikuroi* against the bacterium *Ralstonia solanacearum* (Spraker et al., 2018) and with lagopodin B produced by *Coprinopsis cinerea* in response to the presence of bacteria (Stöckli et al., 2019).

The Fundamentals of Secondary Metabolite Production in Fungi

A Diversity of Biosynthetic Enzymes

Fungal SMs are classified into a few major chemical classes depending on the precursors (either acyl-coAs, amino acids, or prenyl diphosphates) and enzymes used for their biosynthesis (Keller, 2019). Polyketides, nonribosomal peptides, terpenes and indole alkaloids are synthesized by core enzymes named polyketides synthases (PKSs), non-ribosomal peptide synthetases (NRPSs), terpene cyclases (TCs) and dimethylallyltryptophan synthases (DMATSs), respectively **(Fig. 02 and Fig. 03(A))**. These core enzymes are characterized by highly conserved domains that exhibit diverse catalytic activities. TCs and DMATSs are enzymes which are composed of one or two

A Core enzymes

Type I non-reducing PKSs

Type I reducing PKSs

Type III PKSs

NRPSs

Hybrid NRPS-PKSs

DMATSs

NRPS-like enzymes

Hybrid PKS-NRPSs

TCs

NIS synthetases

RiPPs

B Biosynthetic gene clusters

non-biosynthetic functions

unknown or not involved

decoy

transcription factor

transporter

biosynthetic functions

precursors

core enzyme

raw backbone

tailoring enzymes

Final compound(s)

Fig. 03.

Secondary metabolite (SM) biosynthetic pathways in fungal genomes. (A) Core genes and enzymes involved in the biosynthesis of the SM chemical backbone. The module organization of the core enzymes is shown with the arrangement of conserved domains. SAT: Starter Acylcarrier protein Transacylase; KS: Keto-Synthase: AT: Acyl-Transferase; PT: Product Template; ACP: Acyl-Carrier Protein; TE: ThioEsterase; CLC: Claisen Cyclase; MT: Methyl Transferase; R: Reductase; DH: DeHydratase; ER; Enoyl Reductase; KR: Keto Reductase; A: Adenylation; T: Thiolation; C: Condensation; E: Epimerization; TC: Terpene Cyclase; SQCY: Squalene Cyclase; PPt: PolyPrenyl transferase; ABBA-PT: Aromatic Prenyl Transferase; SP: Signal Peptide. Two architectures are found for iterative type I non reducing polyketide synthases (PKSs), but both are characterized by the presence of the SAT and PT domains. The brackets indicate that the ACP domain is duplicated in certain enzymes. Type I reducing PKSs are characterized by the presence of at least one of the DH, ER and KR domain. The brackets indicate the domains that are not found in partially reducing PKSs. The most common structure of NRPS-like enzymes is represented, but diverse domains, including ferric reductase, polynucleotidyl transferase ribonuclease H and LPS-induced tumor necrosis alpha factor, linked to a single A domain are also found. Terpene Cyclases (TCs) are classified into different groups according to the conserved domain they harbor. Bifunctional TCs are characterized by the presence of a PPt domain. NRPS-independent siderophore (NIS) synthetases contain a single IucA_IucC domain. Ribosomally synthesized and post-translationally modified peptides (RiPPs) are encoded by genes. In many cases, the precursor peptide contains a SP followed by a repetitive sequence, which consists of the core peptide embedded between N- and C-terminal sequences and KR recognition motif for the kexin Golgi protease. (B) The biosynthetic gene cluster organization. Reproduced from Bushley, K.E., Turgeon, B.G., 2010. Phylogenomics reveals subfamilies of fungal nonribosomal peptide synthetases and their evolutionary relationships. BMC Evolutionary Biology 10 (1), 26.
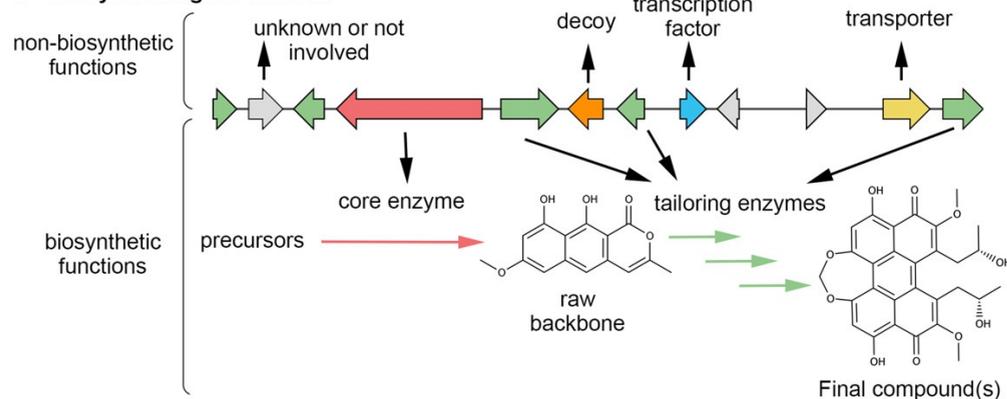
conserved domains (Schmidt-Dannert, 2015). In contrast, PKSs and NRPSs are multidomain mega-enzymes with complex catalytic activities **(Fig. 03(A))**.

In fungi, two types of PKSs are found: iterative type I PKSs and type III PKSs. Type I PKSs are mega-enzymes with several domains that are used iteratively to elongate the polyketide backbone mostly from acetyland malonyl-CoA (Cox, 2007; Herbst et al., 2018). Type I iterative PKSs are further classified into three groups depending on the presence of specific conserved domains **(Fig. 03(A))**. Non-reducing PKSs (nrPKSs) produce aromatic compounds while reducing (rPKS) and partially reducing (prPKS) enzymes produce reduced polyketides due to the presence of one or several reducing conserved domains (dehydratase (DH), keto-reductase (KR) and enoyl-reductase (ER) domains) (Cox, 2007; Herbst et al., 2018). The release of the polyketide chain from the PKS can be performed through different mechanisms depending on the presence of a final thiolesterase (TE) or reductase (R) domain (Cox, 2007). Certain PKSs do not carry such a releasing domain and have co-evolved with other enzymes, including beta-lactamases and hydrolases, to release the polyketide chain (Cox, 2007; Griffiths et al., 2016). While progress has been made towards understanding the programming of iterative type I PKSs, predicting the compound they will produce in terms of acyl-CoA specificity, chain length, cyclisation and reduction steps, remains challenging.

Type III PKSs are small enzymes consisting of a single ketosynthase (KS) domain **(Fig. 03(A))**. They catalyze the iterative condensation of a starter fatty acyl-coA and several extender units, mostly malonyl-CoA (Shimizu et al., 2017). Type III PKSs show considerable flexibility, accepting a wide range of starter unit, from short to long linear or cyclic acyl-coA, resulting in the production of a variety of compounds (Shimizu et al., 2017; Kaneko et al., 2019). Compared to type I PKSs, fungal type III PKSs have been neglected enzymes and more research effort is needed to understand how they function.

NRPSs are large multi-modular enzymes, each module usually consisting of three conserved domains **(Fig. 03(A))**: adenylation (A; recruits an amino acid), condensation (C; catalyses the peptide bond) and thiolation (T; carries the peptide chain) domains (Singh et al., 2017). Eventually, amino acid epimerization and N-methylation domains can be found within modules Singh et al., 2017). Often, the NRPS organization ends with a single C domain that is responsible for the release of the peptide chain through cyclisation (Zhang et al., 2016). In many cases, NRPS modules are sequentially used, so that the number and order of the modules determine the length and nature of the peptide. However, some fungal NRPSs are iterative and contain modules that reuse the A domain of another module to incorporate a new amino acid (Schwecke et al., 2006; Yu et al., 2017). Similarly to fungal PKSs, it is possible to classify NRPSs into large phylogenetic groups (Bushley and Turgeon, 2010), but predicting which amino acids are incorporated remains difficult. A recent study revealed that the activity of NRPS modules also depends on the presence of other surrounding modules (Degen et al., 2019), making these predictions even more uncertain.

Hybrid enzymes between PKSs and NRPSs are also commonly found in fungi, either as PKS-NRPS or NRPS-PKS hybrid enzymes **(Fig. 03(A))**. The former consists of an rPKS fused to a single NRPS module and a final releasing domain, leading to the production of a polyketide which contains a single amino acid (Fisch, 2013). These enzymes are notably known for the production of cytochalasins, which are inhibitors of actin polymerization (Skellam, 2017). Statins are produced by hybrid PKS-NRPSs that are truncated and have lost the A and T domains, consistent with the absence of an amino acid in the statin chemical structure (Boettger et al., 2012). Hybrid NRPS-PKSs are less widespread and thus less studied. The only two characterized fungal NRPS-PKSs either carry a single NRPS module upstream of a KS domain or an A-T

module upstream of a prPKS, and they are respectively responsible for the production of tenuazonic acid in the rice pathogen *Pyricularia oryzae* (Yun et al., 2015), and of swainsonine in the insect pathogen *Metarhizium robertsii* (Cook et al., 2017).

Several other classes have been less studied in fungi **(Fig. 02 and Fig. 03(A))**. NRPS-like enzymes consist of a single A-T module linked to various extra domains, including thioester reductase, ADH short chain dehydrogenase and ferric reductase domains (Bushley and Turgeon, 2010). Only a few compounds, such as atromentin and microperfuranone, have been identified as the product of NRPS-like enzymes (Schneider et al., 2008; Yeh et al., 2012). While most fungi produce siderophores, ironbinding compounds, through an NRPS pathway, it was reported that *Rhizopus delemar* produces the siderophore rhizoferrin using an NRPSindependent siderophore (NIS) synthetase that is characterized by an lucA/lucC conserved domain (Carroll et al., 2017). Such synthetases have been characterized in bacteria (Carroll and Moore, 2018), but genome analyzes suggest that NISs could actually be produced by many fungi.

Certain SMs are not synthesized by conserved core enzymes and are therefore difficult to identify in fungal genomes. Ribosomally synthesized and post-translationally modified peptides (RiPPs) represent a minor class of SMs in fungi (Vogt and Künzler, 2019). In contrast to the previous pathways, RiPPs are directly encoded by genes in the form of a precursor peptide (Yang and van der Donk, 2013). Remarkably, the precursors of the ustiloxin, phomopsin, asperipin-2a and epichloëcyclin RiPPs consist of a signal peptide followed by repeats of the core peptide in between KR signal sequences that are recognized by the kexin Golgi protease for cleavage **(Fig. 03(A))** (Vogt and Künzler, 2019). Finally, SMs like kojic acid and psilocybin are also produced by unconventional biosynthetic pathways that do not include any core enzyme, but only enzymes for the modification of the respective precursors (Terabayashi et al., 2010; Fricke et al., 2017).

The Gene Cluster Organization

The core enzymes described above are involved in the production of the first stable intermediate. However, biosynthetic pathways are usually much more complex and can involve many additional modifications of the produced backbone. For example, the production of aflatoxin involves at least 15 tailoring steps after release of the norsorolinic acid precursor from the PKS (Bhatnagar et al., 2003). These modifications are catalyzed by socalled tailoring enzymes with very diverse catalytic activities such as methylation, monooxygenation, decarboxylation, etc. (Walsh and Tang, 2017). The number and type of modifications steps after the release of the initial chemical backbone from the core enzymes expand the diversity of biosynthetic pathways and chemical structures **(Fig. 02)**.

In fungi, the genes that encode core and tailoring enzymes in a given pathway often co-localize in the genome and are co-regulated, criteria that define a gene cluster organization **(Fig. 03(B))** (Keller and Hohn, 1997). Additionally, these biosynthetic gene clusters (BGCs) can also comprise genes that encode MFS or ABC transporters for SM export or self-protection (Gardiner et al., 2005; Wiemann et al., 2009; Brown et al., 2015; Dolan et al., 2015), transcription factors for BGC regulation (Lyu et al., 2019) and protein decoys for self-protection (Bushley et al., 2013; Yeh et al., 2016). Certain BGCS contain more than one core gene, leading to the production of hybrid molecules. For example, macrolide lactones are produced by a conserved pair of nrPKS and rPKS (Kim et al., 2005), and meroterpenoids are the products of pathways involving PKSs and TCs (Itoh et al., 2010). The diversity of core genes, tailoring genes and their combinations in BGCs makes it possible to produce a large diversity of bioactive compounds from a very limited set of precursors issued from primary metabolism.

How BGCs are formed in fungal genomes remains a matter of debate. A BGC organization has been suggested to facilitate the coregulation of genes and interactions between enzymes from a given biosynthetic pathway (Hurst et al., 2004; Santoni et al., 2013; Rokas et al., 2018). Another hypothesis suggests that the partial loss of genes from a given pathway would result in the production of toxic intermediates (McGary et al., 2013). Finally, this organization could be the result of accumulated horizontal gene transfers (HGTs) (Rokas et al., 2018). None of these hypotheses has been experimentally tested and thus requires further investigation.

Tight Regulation of Secondary Metabolite Production

Given that SMs often exhibit biological activities that could be detrimental to the producing organisms and that their production costs energy, SMs should be produced only when required. BGCs are tightly regulated to be expressed under very specific conditions. The first level of regulation involves pathway-specific transcription factors, which are encoded by genes within the BGC they regulate (Lyu et al., 2019) **(Fig. 04)**. The best characterized example of such a local transcription factor is AflR that regulates the production of aflatoxins (Woloshuk et al., 1994). AflR binds to the promoter of the other genes in the BGC and activate their transcription (Fernandes et al., 1998; Ehrlich et al., 1999).

The second level of regulation involves global regulators that control the expression of a number of BGCs in response to developmental and environmental signals **(Fig. 04)**. CreA, AreA and PacC are transcription factors that control BGC expression in response to the carbon source, nitrogen source, and pH, respectively (Lyu et al., 2019). The Velvet and LaeA protein
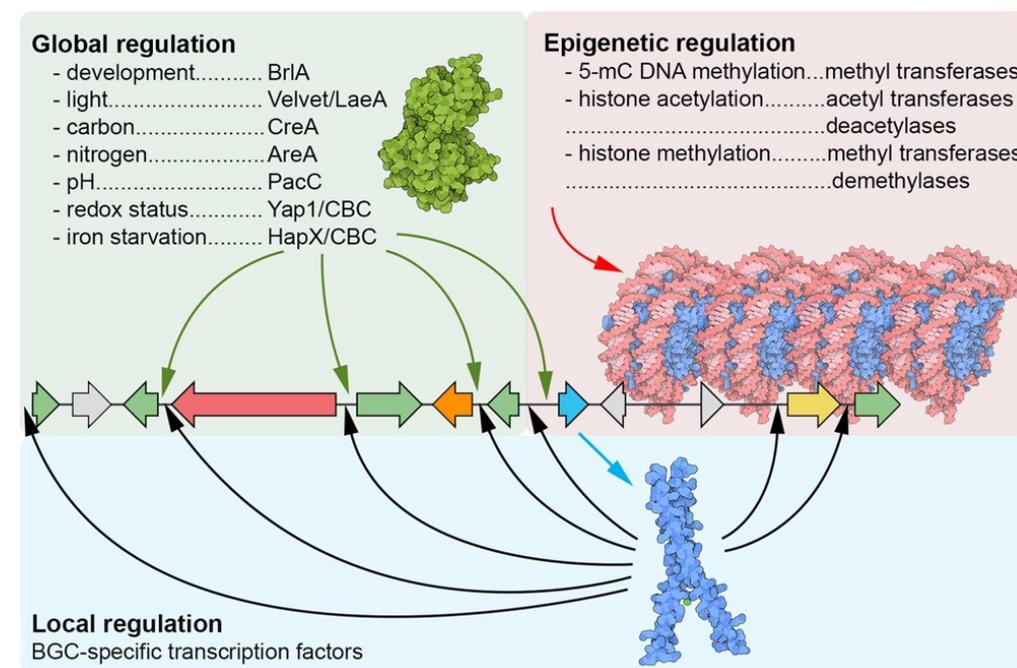
Fig. 04.

Tight regulation of secondary metabolite production in fungi. The expression of biosynthetic gene clusters (BGCs) is controlled by several interconnected levels of regulation. BGCs can contain a gene that encode transcription factors that regulate the expression of the other genes in the BGC by binding to their promoters. Global regulators can also directly regulate the expression of fungal BGCs in response to environmental and developmental signals. Expression of fungal BGCs is also regulated through the modification of chromatin, the complex of DNA and histone proteins. Both DNA methylation and post-translational modifications of histones are epigenetic modifications that affect SM production in fungi. The protein structures are representations obtained with the Illustrate program. Reproduced from Goodsell, D.S., Autin, L., Olson, A.J., 2019. Illustrate: Software for biomolecular illustration. Structure 27 (11), 1716–1720. (Cell Press). doi : 10.1016/J.STR.2019.08.011.

complex regulates BGC expression in response to light (Bayram et al., 2008; Fischer, 2008), while Hap transcription factors control the response to redox status and iron starvation (Reverberi et al., 2012; Wiemann and Keller, 2014; Hortschansky et al., 2017). SM production is also interconnected with fungal development, often linked to conidiation and formation of survival structures like sclerotia (Calvo and Cary, 2015). This coordination involves specific signaling pathways, which are also activated in response to environmental signals (Lind et al., 2018).

The third level of global regulation corresponds to epigenetic modifications **(Fig. 04)**. The conformation of chromatin, the complex of DNA and histone proteins, determine the accessibility of DNA to regulatory proteins (Armeev et al., 2019). Changes between euchromatin (open state) and heterochromatin (closed state) are especially determined by posttranslational modifications of histones, mainly acetylation and methylation (Turner, 2007). As a result, BGCs located in euchromatin can be activated while BGCs located in heterochromatin remain silent (Strauss and Reyes-Dominguez, 2011). This research field has seen much progress in the last decade and a histone code for active and silent BGCs could be determined. BGC expression is mainly linked to acetylation of lysine residues in H3 and H4 histone proteins (Nützmann et al., 2011; Gacek and Strauss, 2012; Soukup et al., 2012; Nützmann et al., 2013). In contrast, trimethylation of lysine residues in H3 histones are a hallmark of BGC repression (Reyes-Dominguez et al., 2010; Gacek and Strauss, 2012; Connolly et al., 2013; Studt et al., 2016). Several histone acetyltransferases and deacetylases, as well as histone methyltransferases and demethylases, have been characterized in fungi and were shown to impact the expression of several BGCs in a given fungus (Collemare and Seidl, 2019). However, BGC regulation by these enzymes is complex because altering the acetylation or methylation status of histones usually results in both activating and repressing

different BGCs (Lee et al., 2009; Connolly et al., 2013; Rösler et al., 2016). Although less studied because of its low abundance in fungi, DNA methylation was also found to modify SM production in several fungi, including *Aspergillus* and *Fusarium* species (Williams et al., 2008; Fisch et al., 2009; Yang et al., 2016). More detailed information about each kind of regulation can be found on recent extensive reviews dedicated to this topic (Collemare and Seidl, 2019; Pfannenstiel and Keller, 2019).

Genomics of Fungal Secondary Metabolism

The study of fungal SMs had been focused on their diverse biological activities and functions until the advent of genomes. Since then, the genomics era has provided new research questions, especially regarding the origin and evolution of BGCs, and new tools to investigate fungal SMs. The number and diversity of available fungal genomes is increasing every day thanks to the development of Next Generation Sequencing (NGS) technologies and the consequential drop in sequencing cost. The public repositories NCBI and Joint Genome Institute MycoCosm portal (Grigoriev et al., 2014) currently host around 2000 and 1400 fungal genomes, respectively. This wealth of genomic information allows for the exploitation of a treasure trove in terms of biochemical reactions and biological activities.

Abundance of Biosynthetic Pathways in Fungal Genomes

The regain of interest in fungal SMs is certainly due to the availability of fungal genomes. For a given fungus, only a few SMs can be detected under standard laboratory conditions. However, genome analyzes have revealed that fungal

species possess a much higher SM production capacity with many BGCs in their genomes remaining silent under laboratory conditions. These silent BGCs encode so-called cryptic biosynthetic pathways. For example, the genome of the well-studied model fungus *Aspergillus nidulans* contains 58 predicted BGCs, but we know the corresponding SM for only 26 of them (Romsdahl and Wang, 2019). The explanation for this discrepancy is the tight regulation of SM production under specific environmental conditions that are difficult to identify and reproduce in the laboratory.

Predicted numbers of core genes highlighted the very high SM production capacity of Ascomycota **(Fig. 05)**. Apart from Saccharomycotina and Taphrinomycotina, the average number of core genes in Ascomycota genomes varies between 12 and 68 **(Fig. 05)**. Basidiomycota fungi, which includes edible mushrooms and wood-decaying species, exhibit a more moderate production capacity with a predominance of NRPS-like and TC genes. With the exception of the Neocallimastigomycetes that are especially rich in NRPS genes, the other fungal lineages show a reduced SM production capacity **(Fig. 05)**.

Even though more than 15,000 fungal SMs are reported in the Natural Product Atlas (van Santen et al., 2019) (an open database of SM compounds curated by the scientific community), characterization of the BGCs behind their synthesis is still scarce. The largest collection of characterized BGCs is the Minimum Information about a Biosynthetic Gene cluster (MIBiG) (Medema et al., 2015; Kautsar et al., 2019), which currently contains around 270 fungal BGCs. This limited number of characterized BGCs and the observation that most of fungal BGCs predicted in fungal genomes are silent, highlight the task that is ahead. Especially, the correct identification and prioritization of silent BGCs using bioinformatic approaches are crucial before embarking on lengthy characterization efforts.
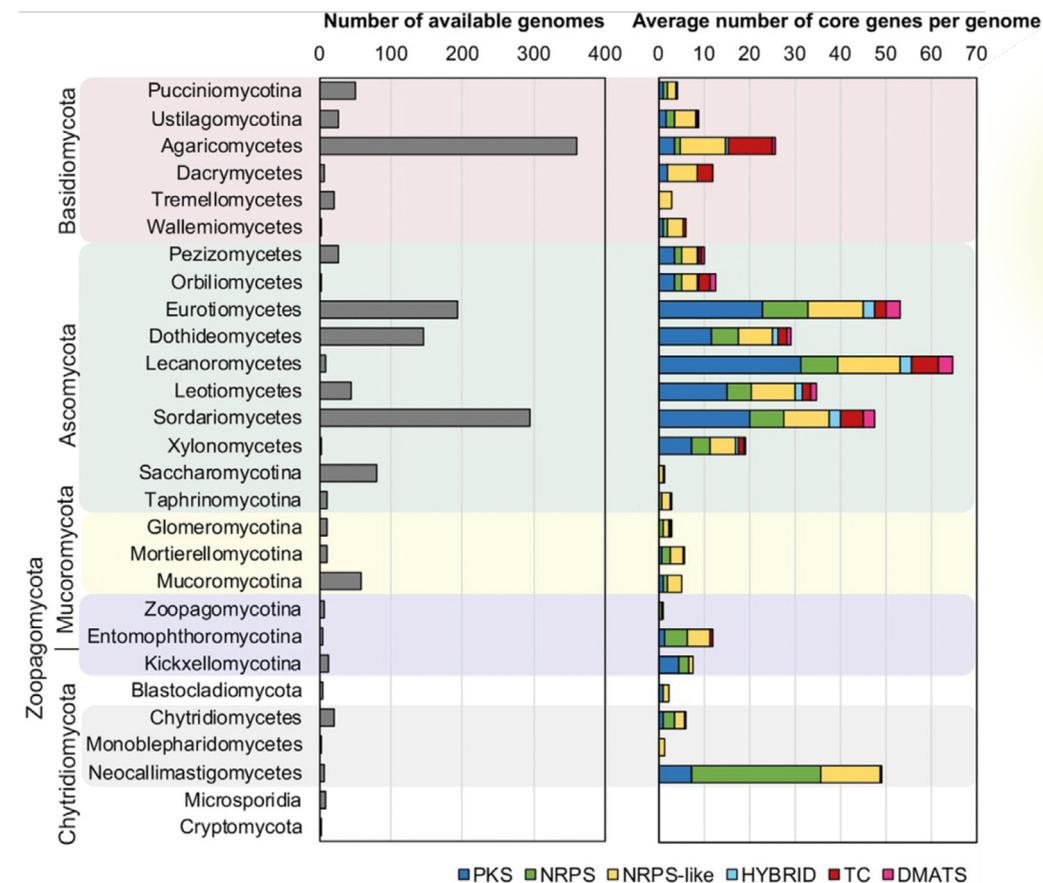
Fig. 05.

Abundance of biosynthetic pathways in fungal genomes. Number of fungal genomes and predicted core genes in these genomes were retrieved from the Joint Genome Institute MycoCosm repository (November 7, 2019). Reproduced from Grigoriev, I.V., et al., 2014. MycoCosm portal: Gearing up for 1000 fungal genomes. Nucleic Acids Research 42 (D1), D699–D704. doi:10.1093/nar/gkt1183.

In Silico Prediction of Biosynthetic Gene Clusters

In silico prediction of BGCs usually starts with the detection of the core genes, which requires good quality genome assemblies and, more importantly,

Table 01.

Software dedicated to the prediction and analysis of biosynthetic gene clusters (BGCs) in fungi

Name	Main Function	Latest publication	Availability	Remarks
SMURF	BGC prediction	(Khaldi et al., 2010)	Webserver: http://smurf.jcvi.org	Motif-dependent prediction. Flexible BGC borders. Not updated anymore
fungiSMASH	BGC prediction	(Blin et al., 2019)	Webserver: https://fungismash.secondarymetabolites.org. Open source: https://docs.antismash.secondarymetabolites.org/install	Motif-dependent prediction. Fixed BGC borders. Active development
CASSIS	Fungal BGC border detection	(Wolf et al., 2016)	Webserver: https://sbi.hki-jena.de/cassis/ Open source: https://sbi.hki-jena.de/smips/Download.php fungiSMASH module	Predict co-regulation by searching shared motifs in gene promoters Incorporated in fungiSMASH 4 (but currently not available in fungiSMASH webserver version)
MIDDAS-M	BGC prediction	(Umemura et al., 2013)	Not available	Motif-independent prediction using gene co-expression
MIPS-CG	BGC prediction	(Takeda et al., 2014)	Not available	Motif-independent prediction using nonsyntenic blocks
FunGeneClusterS	BGC prediction	(Vesth et al., 2016)	Webserver and open source: https://fungiminions.shinyapps.io/FunGeneClusterS/	BGC identification from gene coexpression
NaPDos	Phylogenetic analysis	(Ziemert et al., 2012)	Webserver: http://napdos.ucsd.edu	Phylogeny of key biosynthetic domains
MultiGeneBlast	Comparative genomics	(Medema et al., 2013)	Open source: http://multigeneblast.sourceforge.net	Used as a base for the KnownClusterBlast module in fungiSMASH
BiG-SCAPE	Comparative genomics	(Navarro-Muñoz et al., 2019)	Open source: https://git.wur.nl/medemagroup/BiG-SCAPE	Creates sequence similarity networks and defines BGC families
none	Comparative genomics	(Theobald et al., 2018)	Open source: https://github.com/RoerdamAndersenLab/gene_cluster_networks_and_genetic_dereplication	Creates sequence similarity networks and defines BGC families

accurate gene prediction. While this can be accomplished with simple BLAST searches using reference data, several *in silico* BGC mining tools have been created to facilitate this process (Ziemert et al., 2016), with most of them falling into two categories: motif-dependent and motif-independent detection. It is noteworthy to mention that many tools have been superseded or are not updated anymore, and in some cases, they only exist as a web tool with no open source code **(Table 01)**.

Motif-dependent detection

This approach is based on the conserved sequences (domains) that characterize the SM core enzymes of the most commonly found biosynthetic classes. Protein conserved domains can be encoded as probabilistic mathematical models (hmm: hidden Markov model) that are built from multiple sequence alignments. The frequencies of every amino acid at each position in the alignment are recorded into a so-called profile and then used by software like HMMER (hmmer.org) to search and score similar sequences (Eddy, 1998). Nowadays, two tools that make use of hmm profiles (from the Pfam (El-Gebali et al., 2019), TIGRFAM (Haft et al., 2001) or custom-made databases) are predominantly employed to predict BGCs in fungal genomes with high confidence (Medema and Fischbach, 2015).

SMURF was the first tool specifically designed to detect fungal BGCs (Khaldi et al., 2010). This tool uses domain detection not only to find NRPS, PKS, PKS/NRPS hybrid, NRPS-like, PKS-like and DMATS core enzymes in predicted proteomes, but also to find tailoring enzymes. These proteins are also commonly characterized by conserved domains and SMURF relies on a dataset of conserved domains found in 22 BGCs from *Aspergillus fumigatus*.

Once a core gene is found, 20 genes upstream and downstream are scanned for the presence of tailoring SM domains and are either tagged as "SM domain positive" or "SM domain negative". The algorithm keeps scanning genes until (a) the 20-gene window is over; (b) a maximum intergenic distance to the next gene is found; or (c) a maximum number of "SM domain negative" genes are found. While this strategy is flexible to account for the different BGC configurations, its parameters are based on a very limited dataset *A. fumigatus* BGCs. Although SMURF is available online, it is not actively developed anymore.

Just after the release of SMURF, the first version of antiSMASH was published (Medema et al., 2011), which follows the same principle of detecting core genes using profile hmms. antiSMASH was originally developed for mining bacterial BGCs, but it is nevertheless capable of finding all the usual fungal BGC types due to the fact that core enzymes are conserved in both kingdoms. Then, a region defined by a simple extension of the up- and downstream sequences from the core gene(s) position(s) is reported as containing a BGC. Currently, version 5 of antiSMASH brought a complete revision of the code base, a new conceptual organization for the locus with biosynthetic genes (core, neighborhood, candidate gene cluster and region), and detection of fungal RiPP BGCs that resemble the ustiloxin BGC (Blin et al., 2019). antiSMASH is actively maintained and can be used through the online web version, installed locally or used through Docker. It is worth noting that the fungal version of antiSMASH, named fungiSMASH, does not accept unannotated genomes, meaning that accurate gene prediction in the genome of interest has to be performed beforehand.

While detecting core genes is relatively straightforward, finding borders of a given BGC is the major challenge in BGC prediction. Both SMURF and antiSMASH show low accuracy in predicting BGC borders. The key difference between fungiSMASH and the regular antiSMASH was the

incorporation of the CASSIS algorithm to predict BGC borders (Wolf et al., 2016; Blin et al., 2017). In a given predicted BGC region, CASSIS reports genes that share a motif in their promoter region as this motif would indicate their potential co-regulation. A set of 38 characterized BGCs from *Aspergillus* spp. was used to determine CASSIS parameters.

Both SMURF and antiSMASH suffer from several limitations because they can only detect the major known core enzymes and they cannot report complete BGCs when they are dispersed at more than a single locus (e.g., cephalosporin (Singh et al., 2019), melanin (Langfelder et al., 2003), dothistromin (Schwelm and Bradshaw, 2010)) (Medema and Fischbach, 2015). In addition, the prediction of BGC borders is in most cases inaccurate and manual curation of the results is highly recommended, including validation of functional annotations. Many tools to predict substrate specificities and produced compounds are available either as standalone or integrated into antiSMASH. However, most of them are based on bacterial data and are most often inaccurate when applied to fungal sequences. Functional characterization of many more fungal BGCs is needed in order to develop accurate prediction tools.

Motif-independent detection

Considering the limitations of motif-dependent detection tools, other strategies have been employed to comprehensively identify BGCs in fungal genomes, including non-classical pathways (Umemura et al., 2015). A first strategy is based on the observation that genes from the same biosynthetic pathway are co-regulated, even when they do not contain a classical conserved core gene. Such a strategy was developed in the MIDDAS-M algorithm which

averages gene expression over a genewindow and normalizes expression ratios to yield significantly larger values for BGCs (Umemura et al., 2013). This algorithm applied to several fungal species not only accurately reported characterized BGCs, but also identified a new BGC for the RiPP ustiloxin B in *Aspergillus flavus*, which was functionally validated through targeted gene deletion (Umemura et al., 2013). Limitations of this approach are its dependency to experimental data and its ability to detect only the few BGCs that are expressed under laboratory conditions.

Comparative genomic analyzes revealed that fungal BGCs tend to be located in non-syntenic blocks (NSBs) (Machida et al., 2005; Amselem et al., 2011; Collemare et al., 2014; Lind et al., 2017). This observation led to the development of MIPS-CG, which attempted to detect novel BGC candidates by searching for conserved clustered genes between genomes in NSBs (Takeda et al., 2014). This approach is independent of any prior knowledge on fungal BGCs, allowing the successful prediction of nonclassical biosynthetic pathways as exemplified with the detection of the kojic acid BGC in *Aspergillus oryzae* (Takeda et al., 2014).

Both MIDDAS-M and MIPS-CG tools are actually not available anymore, but the logic behind these tools can be easily implemented in tailor-made methods to study genomes of interest. The combination of motif-dependent and motif-independent strategies was also tested in *A. nidulans* by linking SMURF predictions to microarray expression data from diverse conditions (Andersen et al., 2013). In this study, the expression data was used to calculate a Cluster Score as a measure for gene coexpression, which accurately predicted that two distant BGCs are responsible for the production of the hybrid peptide-terpene compound nidulanin A. FunGeneClusterS, a graphical online tool with an improved algorithm, was released in a follow up work (Vesth et al., 2016).

Strategies for Prioritizing the Study of Biosynthetic Gene Clusters

Considering the large amount of predicted BGCs in fungal genomes, prioritization of BGCs for functional characterization is needed. Several strategies can be employed to select candidate BGCs that are likely to produce novel compounds or variations of known compounds (Bertrand and Sorensen, 2018; Kjærbølling et al., 2019).

Phylogeny-informed prediction

General phylogenetic studies of core enzymes have been used to elucidate their evolutionary histories (nrPKS (Koczyk et al., 2015)) and to define broad evolutionary clades that relate to enzymatic activities, often by focusing on a single domain (sesterterpenes (Narita et al., 2017), nrPKS (Liu et al., 2015; Throckmorton et al., 2015), NRPS (Bushley and Turgeon, 2010)). Phylogenetic analyzes of core enzymes indicated that closely related core enzymes show functional homology as exemplified with nrPKSs involved in the production of emodin-like anthraquinones (Collemare et al., 2014; Griffiths et al., 2016) or siderophore synthetases (Bushley and Turgeon, 2010). This principle can be used to link BGCs with known metabolites (e.g., (Gibson et al., 2014)) and to directly target novel BGCs whose core enzymes are divergent from characterized ones (Harvey et al., 2018). NaPDoS is a bioinformatic tool that was developed to perform functional predictions based on the phylogeny of key domains from characterized PKS (KS domain) and NRPS (C domain) enzymes (Ziemert et al., 2012). While this phylogenetic approach is very useful, predictions of the final chemical structures is not possible because tailoring enzymes modify the backbone produced by the core enzymes. It is therefore crucial to employ also strategies that consider whole BGCs.

Strategies for Comparative genomics and gene cluster families

Comparative genomics is based on pairwise comparisons between a query and a set of BGCs ("one-to-many"). One of the classical tools for pairwise BGC comparison is MultiGeneBlast (Medema et al., 2013), which has been incorporated into antiSMASH in two analysis modules. A first module (ClusterBlast) compares each query BGC against predicted BGCs from antiSMASH DB (Blin et al., 2018) (currently holding bacterial BGCs only), while the other module (KnownClusterBlast) compares the query against the set of characterized clusters from MIBiG (Medema et al., 2015; Kautsar et al., 2019). These tools remain of limited interest for fungal BGCs until a significant higher number of BGCs will become functionally characterized.

A more integrative approach makes use of sequence similarity networks (SSN) to provide a "many-to-many" overview, automatically defining gene cluster families (GCFs). GCFs corresponds to groups of BGCs that share a common origin and have diverged, resulting in the production of diverse, yet related, SMs. For example, BiG-SCAPE processes antiSMASH results, annotating protein domains in all regions containing a predicted BGC to define a distance matrix based on three components (domain content similarity, synteny and sequence identity) (Navarro-Muñoz et al., 2019). Clustering into GCFs is performed using an affinity propagation algorithm on the resulting distance matrix. BiG-SCAPE accuracy was validated by comparing the GCF calling to metabolomics data for a bacterial case. In principle, BiG-SCAPE could also be used with fungal BGCs.

A similar approach was employed using BGC prediction with a modified version of SMURF in 32 *Aspergillus* spp. (Theobald et al., 2018). The developed method aggregated percentages of identity of best bidirectional hits for each

protein in the predicted BGCs, resulting in similarity scores that are translated into a BGC network. A random-walk algorithm is then applied to the network to define GCFs. In combination with expert knowledge of the malformin chemical structure, this method allowed to link malformin to a GCF in *Aspergillus* species, of which one BGC in *Aspergillus brasiliensis* was experimentally confirmed.

Function-guided prioritization

A different approach for BGC prioritization is not based on sequence comparisons but on researcher's interest in specific enzymatic functions. For example, new chlorinated polyketides from *Bipolaris sorokiniana* were discovered by selecting predicted BGCs which contain genes encoding halogenases (Han et al., 2019).

The same strategy was successfully applied to search for BGCs that would produce SMs with specific biological activities. Several characterized BGCs contain a gene that encodes the enzymatic target of the SM, acting as a decoy and thus providing resistance to the toxic compound (Almabruk et al., 2018). For example, the cyclosporine A BGC contains a gene encoding cyclophilin, the target of cyclosporine (Yang et al., 2018). Such a strategy successfully identified the BGCs involved in the production of mycophenolic acid (MPA) in *Penicillium brevicompactum* (Regueira et al., 2011) and aspterric acid in *Aspergillus terreus* (Yan et al., 2018). The MPA BGC was identified in a cosmid library searching for a meroterpenoid BGC that would contain a copy of IMP dehydrogenase, the known target of MPA in B and T lymphocytes (Regueira et al., 2011). The aspterric acid BGC was identified in the search for herbicides that target the enzyme dihydroxyacid dehydratase (DHAD) involved

in the plantessential branched chain amino acid pathway. Mining of fungal genomes for BGCs that contain a DHAD copy retrieved the aspterric acid BGC, which was functionally validated (Yan et al., 2018).

Resistance gene-directed discovery has been implemented in ARTS (Alanjary et al., 2017), a web tool for discovery of bacterial BGCs with different resistance mechanisms. Very recently, a similar pipeline called FRIGG has been designed for fungi (Kjærbølling et al., 2019). This tool combines precomputed homolog protein families from different genomes with predicted BGCs. In this case, BGCs that contain genes homologous to genes present outside of the cluster may indicate they encode the target of the corresponding SM. Experimental validation of the FRIGG pipeline is still awaiting.

Linking Compounds to Biosynthetic Genes

Screenings for biological activities have resulted in the identification of many compounds from diverse fungi. In most cases, the BGCs behind these compounds is not known. When the genome of a given fungus becomes available, it is possible to implement the strategies mentioned above to select candidate BGCs to study. For this purpose, expert knowledge of the compound chemical structure is necessary to target specific biosynthetic classes using a "retro-biosynthetic" approach. The chemical structure will provide hints about the major chemical classes (polyketide, peptide, etc.) and the needed tailoring reactions (methylation, halogenation, dimerization, etc.). For example, the BGC involved in the production of aspirochlorine was identified thanks to the presence of an halogenase gene that correlate with the chlorinated residue found in this SMs (Chankhamjon et al., 2014). Such a method can also apply to nonclassical biosynthetic pathways. The genomes of psilocybin-

producers *Psilocybe cubensis* and *Psilocybe cyanescens* were interrogated for the presence of clustered genes that would encode a methyltransferase, a hydroxylase, and a kinase, which are expected for psilocybin biosynthesis. This search retrieved a conserved locus between both species harboring all the expected genes and the predicted BGC was subsequently proven to be responsible for psilocybin production (Fricke et al., 2017).

If the target compound is a chemical analog of a SM with a characterized BGC, it may be inferred that both compounds share the same biosynthetic origin. In this case, the characterized BGC may be used to guide a comparative genomics analysis. This approach have been applied in the *Aspergillus* genus and candidate BGCs could be identified for the production of novo-funigatonin, ent-cycloechinulin and epiaszonalenin A and C in *A. fumigatus* and ochrindol in *Aspergillus steynii* (Kjærbølling et al., 2018).

Genomics of Fungal Secondary Metabolism

Bioinformatics has revealed a large number of predicted BGCs in fungal genomes and has made it possible to explore effectively the diversity of naturally encoded pathways *in silico*. At the same time, only a very small fraction of these predicted BGCs are associated with characterized SMs and most pathways remain cryptic because they are not expressed under standard laboratory conditions. In the last 15 years, much effort has focused on methods to induce the expression of these silent BGCs and get access to the wealth of SMs they could produce. Several successful strategies have been developed, from the modification of culture conditions to genetic modifications, taking advantage of the accumulated knowledge about BGC regulation.

Modification of Culture Conditions

The tight regulation of fungal BGCs make them produce SMs under very specific conditions during their lifecycle. The laboratory conditions are usually adjusted for optimal growth and thus lack the specific signals for BGC activation. Considering that environmental signals are known to affect SM production (Takahashi et al., 2013), one approach is to grow the fungus of interest under a variety of conditions (light, temperature, oxygenation, etc.) and substrates (different carbon, nitrogen sources, etc.), so that one of the conditions actually contains a signal required for the activation of a BGC. This approach is known as OSMAC (One Strain, Many Compounds) and has been successfully applied to many fungi. For example, cultivating *A. nidulans* under more than 40 different conditions resulted in the discovery of the compounds aspoquinolone (Scherlach and Hertweck, 2006) and aspernidine A and B (Scherlach et al., 2010). Since the initial OSMAC study that discovered 15 SMs produced by *Aspergillus ochraceus* (Bode et al., 2002), this approach has resulted in the discovery of many SMs as recently reviewed in (Romano et al., 2018). Despite its success, the OSMAC approach can be quite tedious because there is no common signal that could always be used to activate the production of several SMs in many fungi. The multiplication of the conditions to test makes this approach difficult to employ on a large number of fungal species at once.

Nowadays, OSMAC approaches also include the use of chemicals to test their effect on SM production. In particular, chromatin remodeling agents are commonly employed because they provide a more consistent and broader effect on BGC activation due to their epigenetic regulation. Growing fungi in the presence of histone deacetylase inhibitors like suberoylanilide hydroxamic acid (SAHA) and of DNA methyltransferase inhibitors like 5-azacytidine resulted in successful activation of SM production in many studies (Henrikson et al., 2009; Wang et al., 2010; Vervoort et al., 2011; Zutz et al., 2013).

Co-Cultivation of Microorganisms

In contrast to competition-deprived laboratory conditions, fungi reside in natural habitats that are normally occupied by many other organisms, including bacterial and fungal competitors. The production of SMs is important to interact or fight with these other organisms. Co-cultivation of organisms (for example, bacteria with fungi or fungi with fungi) is used to mimic competition in natural habitats and induce SM biosynthesis. For example, co-culture of two extremophile *Penicillium* species from the Berkeley Pit Lake yielded eight new macrolide lactones (Stierle et al., 2017). Similarly, co-culture of *A. nidulans* with the soil-dwelling actinomycete *Streptomyces rapamycinicus* demonstrated elevated synthesis of several polyketides, such as orsellinic acid, its derivative lecanoric acid, and two cathepsin K inhibitors, F-9775A and F-9775B (Schroeckh et al., 2009). Co-cultivation of the human pathogen *A. fumigatus* with soil dwelling actinobacteria also induced the synthesis of the meroterpenoid fumicycline (König et al., 2013). Physical contact with living bacteria seems necessary to trigger SM production because in both cases induction was not accomplished by supplementing fungal axenic culture with bacterial media, sterilized bacteria cultivation supernatant or co-cultivation with physical separation of organisms with dialysis tube.

The induction of SM production in co-cultures appears very specific because a single species out of 58 different *Streptomyces* bacteria induced SM production in *A. nidulans* (Schroeckh et al., 2009). Such a specificity makes the outcome of the co-culture approach difficult to predict. The

underlying molecular mechanism(s) has not been deeply investigated yet, but the co-culture of *A. nidulans* with *S. rapamycinicus* was showed to trigger modifications of fungal histone acetylation (Nützmann et al., 2011). Higher acetylation at the locus of the transcription factor gene basR induced its expression, which in turn activated the expression of BGCs (Fischer et al., 2018). Recently, it was shown that the bacterium *Pseudomonas piscium* suppressed SM production in *Fusarium graminearum* by secreting the histone acetyltransferase inhibitor phenazine-1-carboxamide (Chen et al., 2018). Overall, it appears that the specific effect of co-cultures relies on the active secretion of SMs that trigger a general response in the fungus, either activation or suppression of SM production.

Genetic Modifications

Pathway-specific activation

Biosynthetic pathways can be regulated by BGC-specific transcription factors that are encoded within the corresponding BGC [(Fig. 04)]. A strategy to activate a silent BGC which contains a transcription factor gene is thus to overexpress this regulatory gene. Such an overexpression can be achieved by introducing an extra copy of the transcription factor gene that is under the control of a constitutive or inducible promoter at another locus in the genome. This strategy successfully activated a silent PKS-NRPS BGC in *A. nidulans*, leading to the discovery of aspyridone (Bergmann et al., 2007). Instead of adding an extra copy of the transcription factor gene, an alternative strategy is to replace the native promoter of the regulatory gene. For example, replacement by the inducible alcA promoter of the promoter of the

transcription factor gene found in a silent gene cluster with two PKS genes in *A. nidulans*, successfully induced the targeted BGC and yielded the new polyketide asperfuranone (Chiang et al., 2009). However, in some cases, over-expression of the transcription factor gene does not result in BGC activation as shown for the fellutamide B BGC in *A. nidulans* (Yeh et al., 2016). Activation of this silent BGC and discovery of fellutamide B was accomplished by serial promoter swap for each gene of this BGC.

Genetic manipulation of the global regulation

The study of BGC global regulation, involving both global regulators and epigenetic modifications [(Fig. 04)], through targeted gene deletion and over-expression resulted in the activation or repression of many BGCs at once. Thus, these functional studies were further developed as a strategy to activate silent BGCs. One of the most studied global regulators is the complex formed by Velvet proteins and LaeA. In *A. nidulans*, deletion of *LaeA* impairs the production of penicillin, sterigmatocystin and lovastatin, while over-expression increases penicillin and lovastatin production (Bok and Keller, 2004). Similarly, manipulation of other Velvet complex components leads to diverse effects on different BGCs. For example, deletion of FfVel1 in *F. fujikuroi* resulted in the stimulation and downregulation of bikaverin and fumonisin, respectively, while it did not influence fusarin C production (Wiemann et al., 2010). Application of LaeA over-expression in the non-model fungus *Aspergillus fumisynnematus* successfully activated the production of a new compound in this fungus, which was identified as the known SM cyclopiazonic acid (Hong et al.,2015).

Preferred targets for genetic modifications are genes involved

in post-translational modifications of histones. Acetylation of histone proteins is commonly associated with active gene transcription, while their methylation is commonly associated with repression of gene expression (Strauss and Reyes-Dominguez, 2011). Consistently, the deletion of histone deacetylases (HDACs) and histone methyltransferases activates silent BGCs (Shwab et al., 2007; Bok et al., 2009). Similarly, over-expression of histone acetyltransferases activates SM production as shown in *A. nidulans* (Soukup et al., 2012). In contrast, deletion of histone acetyltransferases and histone demethylases mostly results in reducing SM production (Gacek-Matthews et al., 2015; Rösler et al., 2016). Because the deletion of HDACs results in hyperacetylated chromatin and thus active gene transcription (Eberharter, and Becker, 2002), these enzymes have been the preferred target to activate silent BGCs and SM production in a number of fungi (Lee et al., 2009; Studt et al., 2013; Maeda et al., 2017; Pidroni et al., 2018). A detailed overview of fungal chromatin-modifying enzymes that can be used as targets for deletion or overexpression has recently been published (Pfannenstiel and Keller, 2019).

Despite the success of this strategy to induce the expression of silent BGCs, their effect is more complex and cannot be predicted. For example, while deletion of the Gcn5 histone acetyltransferase in *F. fujikuroi* decreased the expression of most BGCs, the production of bikaverin was unexpectedly increased (Rösler et al., 2016). In all examples mentioned above, both activation and repression of BGCs were actually found. In other fungi, manipulation of these global regulators only alters the production of already active BGCs and do not activate the production of any new SM (Griffiths et al., 2015).

Heterologous Expression of Biosynthetic Gene Clusters

A major limitation to the genetic modifications mentioned above is that they can be efficiently implemented only in tractable fungal species. Thus, their use remains limited to mostly model fungi and does not allow exploiting the full SM production potential of fungi. An alternative that has been increasingly employed is to express BGCs in heterologous hosts (Skellam, 2019).

The cloning of large PKS and NRPS genes, as well as complete BGCs, is a challenge that was overcome, thanks to cloning technologies like Gateway (Reece-Hoyes and Walhout, 2018) and Golden Gate (Engler and Marillonnet, 2014), as well as the use of transformation-associated recombination (TAR), a method that relies on homologous recombination to assemble several overlapping DNA fragments in *Saccharomyces cerevisiae* (Kouprina and Larionov, 2016) and in *Escherichia coli* (Jacobus and Gross, 2015). Several cloning systems have been developed for heterologous expression of multiple genes in a fungal host, giving the possibility to express complete BGCs (Ishiuchi et al., 2012; Pahirulzaman et al., 2012; Nielsen et al., 2013; Unkles et al., 2014; Bok et al., 2015; Gressler et al., 2015; Geib and Brock, 2017). Both yeast and filamentous fungi have been used as suitable heterologous hosts. *S. cerevisiae* benefits from decades of biotechnological development for protein and metabolite production. Engineering *S. cerevisiae* for successful SM production is required, especially for the production of polyketides and non-ribosomal peptides, which core enzymes need to be activated by a 4-phosphopantetheinyl transferase (Ishiuchi et al., 2012; Billingsley et al., 2016; Harvey et al., 2018). In addition, incorrect splicing of foreign genes in yeast is commonly encountered, meaning that intron-less genes should be amplified for successful expression (Billingsley et al., 2016; Harvey et al., 2018). Many different species of filamentous fungi have been used for

heterologous expression, mostly from the *Aspergillus* genus (*A. nidulans*, *A. niger*, *A. oryzae*) (Pahirulzaman et al., 2012; Chiang et al., 2013; Richter et al., 2014). *Penicillium* and *Fusarium* species have also been reported as suitable host for heterologous expression of fungal BGCs (Kindinger et al., 2019; Nielsen et al., 2019). So far, the most used host has been *A. oryzae* as it benefits from a long biotechnological development for fermentation processes, which makes this fungus suitable for safe production of enzymes and metabolites. Currently, the *A. oryzae* NSAR1 strain is becoming the main host for the expression of fungal BGCs, because its four auxotrophic markers make it suitable to transform with several plasmids harboring different selection markers and safe to use as it cannot survive if unintentionally released to the environment (Jin et al., 2004; Skellam, 2019). Intron splicing in filamentous fungal hosts is less of an issue than in yeast, but the foreign genes are not always accurately processed as exemplified by the attempts to express the ACE1 gene from *P. oryzae* in *A. oryzae* (Song et al., 2015).

Heterologous expression has mainly been applied to single targeted BGCs. A scaled-up heterologous expression system was designed, making use of fungal artificial chromosomes and metabolomic scoring (FAC-MS) (Clevenger et al., 2017). A library of 156 FACs from *Aspergillus wentii*, *Aspergillus aculeatus* and *A. nidulans* was transformed into *A. nidulans*. Analysis of SMs produced by the transformants yielded 15 new SMs and their corresponding BGCs. Although successful, this method requires conserved signaling pathways and thus can only be applied in closely related fungal species.

Nowadays, we are entering a new era in which DNA synthesis is allowing large-scale targeted heterologous expression of silent BGCs. In a recent study, 41 fungal BGCs selected to cover a wide phylogenetic diversity were synthesized and expressed in yeast (Harvey et al., 2018). The production of a SM was detected for 22 of these BGCs, of which 10 were new compounds

(Harvey et al., 2018). Incorrect gene structure prediction was suggested to be a major reason for failed SM production in this study. Indeed, manual inspection of intron prediction in the terpene cyclase TC5 identified a wrongly predicted intron and the corrected gene structure expressed in yeast successfully yielded new sesquiterpenoids (Harvey et al., 2018). It is likely that DNA synthesis combined with heterologous expression will become the strategy of choice thanks to the development of new technologies that will lower its cost. The generation of high-quality genomic data and accurate gene prediction will be crucial for the success of large-scale heterologous expression.

Engineering Biosynthetic Pathways: Chimeric Enzymes, Pathway Refactoring, Combinatorial Biosynthesis, Artificial Pathways

The development of heterologous expression methods has been instrumental in elucidating biosynthetic pathways and activating silent BGCs. However, these methods are also opening new biotechnological opportunities towards the production of non-natural SMs.

Studies to understand the biosynthetic mechanisms of core enzymes, especially PKSs and NRPSs, made use of chimeric enzymes that combined conserved domains from different enzymes. For example, a chimeric PKS made of the SAT-KS-AT-PT domains from *Colletotrichum lagenarium* Pks1 and of the ACP-CLC domains from *A. nidulans* wA yielded the production of a novel hexaketide different from the precursor produced by Pks1 and wA (Watanabe and Ebizuka, 2002). Rational domain swapping between core enzymes often result in the production of new SMs which provide information about how these enzymes function, especially about chain length control in PKSs (Fisch et al., 2011; Liu et al., 2014).

Nearly a decade ago, the concept of plug-and-play synthetic biology was suggested as a method to engineer or optimize biosynthetic pathways (Medema et al., 2011). In this concept, all core and tailoring genes are considered as pieces that can be rearranged and optimized in order to increase the production of a given SM or produce new SMs. Such a strategy was recently applied to the austinoid pathway. Austinol is a meroterpenoid compound with insecticide activity produced by *A. nidulans* under favorable fermentation conditions (Mattern et al., 2017). *Aspergillus calidoustus* produces another derivative with higher insecticide activity, calidodehydroaustin, but only under conditions that are not favorable for industrial fermentation (Mattern et al., 2017; Valiante et al., 2017). The comparison of the austinoid BGCs in *A. nidulans*, *A. calidoustus* and *Penicillium brasilianum* identified four candidate genes for the conversion of austinol into calidodehydroaustin (Mattern et al., 2017). Expression of the candidate *A. calidoustus* genes in *A. nidulans* under control of an inducible promoter resulted in successfully rewiring the austinol biosynthetic pathway in *A. nidulans* to produce calidodehydroaustin (Mattern et al., 2017).

The plug-and-play concept allows envisaging combinatorial to produce new compounds. This possibility was exemplified with the study of BGCs involved in the production of macrolide lactones and which encode both an nrPKS and an rPKS. Different combinations of nrPKS and rPKS pairs from related BGCs in different fungal species resulted in the biosynthesis of diverse SMs, including unnatural lactone variants with altered biological activity (Xu et al., 2014).

Theoretically, it is conceivable to create completely artificial pathways using genes from unrelated BGCs. While such an example with fungal sequences has not yet been published, an artificial biosynthetic pathway to produce carminic acid, a red food colorant of insect origin, in *A. nidulans* was recently reported. In this case, a type III PKS gene of plant origin was expressed in *A. nidulans* together with cyclase and aromatase genes of bacterial origin, resulting in the production of the expected carminic acid backbone (Frandsen et al., 2018). This backbone was modified by endogenous monooxygenases, yielding the needed next precursor. The final step of carminic biosynthesis was completed with the expression of a C-glucosyltransferase from the insect *Dactylopius coccus*.

Fungal Secondary Metabolite Production: Future Perspectives

Our fundamental knowledge of fungal SM production has impressively improved in the first two decades of the 21st century, notably thanks to the development of genomics and biotechnological tools. Despite this progress, the study and exploitation of fungal SMs remains behind compared to bacterial SMs as clearly shown by the restricted number of characterized fungal BGCs. Further development of biotechnological tools to fully exploit the SM production capacity of fungi is still hampered from lack of fundamental knowledge.

Active development of bioinformatics tools dedicated to fungal genomes is needed to address several key challenges such as accurate BGC definition, integration of expression data, detection of BGCs split over different loci and evolution of biosynthetic pathways in GCFs. With the increasing availability of both fungal genomes and characterized data, including transcriptomics and metabolomics data, there is now a timely opportunity to develop new fungal-specific approaches.

Although a better understanding of BGC regulation could provide new tools to activate silent BGCs in the future, biotechnological strategies

relying on heterologous expression appear to be the most promising for the efficient production of new fungal SMs. Such strategies will allow obtaining bioactive compounds in industrial-scale quantities, developing semi-synthetic approaches to produce chemical structures that are otherwise difficult to obtain with synthetic chemistry (Asai et al., 2015; Alberti et al., 2017), and building artificial BGCs. These developments require the functional characterization of many more fungal core and tailoring enzymes, especially regarding substrate specificities and interactions between tailoring enzymes. The combination of functional characterizations, bioinformatics and biotechnological tools is promising to lead to the efficient and less labor-intensive production of diverse bioactive SMs of fungal origin.

Acknowledgments

References

* **Alanjary, M., et al., (2017).** The antibiotic resistant target seeker (ARTS), an exploration engine for antibiotic cluster prioritization and novel drug target discovery. Nucleic Acids Research 45 (W1), W42–W48. doi : 10.1093/nar/gkx360.

* **Alberti, F., et al., (2017).** Heterologous expression reveals the biosynthesis of the antibiotic pleuromutilin and generates bioactive semi-synthetic derivatives. Nature Communications 8 (1), 1831. doi : 10.1038/s41467-017-01659-1.

* **Almabruk, K.H., Dinh, L.K., Philmus, B., (2018).** Self-resistance of natural product producers: Past, present, and future focusing on self-resistant protein variants. ACS Chemical Biology 13 (6), 1426–1437. doi : 10.1021/acschembio.8b00173.

* **Amselem, J., et al., (2011).** Genomic analysis of the necrotrophic fungal pathogens *Sclerotinia sclerotiorum* and *Botrytis cinerea*. PLoS Genetics 7 (8), e1002230. doi : 10.1371/journal.pgen.1002230.

* **Andersen, M.R., et al., (2013).** Accurate prediction of secondary metabolite gene clusters in filamentous fungi. Proceedings of the National Academy of Sciences of the United States of America 110 (1), doi : 10.1073/pnas.1205532110.

* **Armeev, G.A., et al., (2019).** Linking chromatin composition and structural dynamics at the nucleosome level. Current Opinion in Structural Biology 56, 46–55. doi : 10.1016/J. SBI.2018.11.006.

* **Asai, T., et al., (2015).** Use of a biosynthetic intermediate to explore the chemical diversity of pseudo-natural fungal polyketides. Nature Chemistry 7 (9), 737–743. doi : 10.1038/ nchem.2308.

* **Bartlett, D.W., et al., (2002).** The strobilurin fungicides. Pest Management Science 58 (7), 649–662. doi : 10.1002/ps.520.

* **Bayram, O., et al., (2008).** VelB/VeA/LaeA complex coordinates light signal with fungal development and secondary metabolism. Science 320 (5882), 1504–1506. doi : 10.1126/ science.1155888.

* **Bergmann, S., et al., (2007).** Genomics-driven discovery of PKS-NRPS hybrid metabolites from *Aspergillus nidulans*. Nature Chemical Biology 3 (4), 213–217. doi : 10.1038/nchembio869.

* **Bertrand, R.L., Sorensen, J.L., (2018).** A comprehensive catalogue of polyketide synthase gene clusters in lichenizing fungi. Journal of Industrial Microbiology & Biotechnology 45 (12), 1067–1081. doi : 10.1007/s10295-018-2080-y.

* **Bhatnagar, D., Ehrlich, K.C., Cleveland, T.E., (2003).** Molecular genetic analysis and regulation of aflatoxin biosynthesis. Applied Microbiology and Biotechnology 61 (2), 83–93. doi : 10.1007/s00253-002-1199-x.

* **Billingsley, J.M., DeNicola, A.B., Tang, Y., (2016).** Technology development for natural product biosynthesis in *Saccharomyces cerevisiae*. Current Opinion in Biotechnology. 74–83. doi : 10.1016/j.co pbio.2016.02.033.

* **Blin, K., et al., (2017).** AntiSMASH 4.0 – Improvements in chemistry prediction and gene cluster boundary identification. Nucleic Acids Research 45 (W1), doi : 10.1093/nar/gkx319.

* **Blin, K., et al., (2018).** The antiSMASH database version 2: A comprehensive resource on secondary metabolite biosynthetic gene clusters. Nucleic Acids Research 47, 625–630. doi : 10.1093/nar/gky1060.

* **Blin, K., et al., (2019).** AntiSMASH 5.0: Updates to the secondary metabolite genome mining pipeline. Nucleic Acids Research. 1–7. doi : 10.1093/nar/gkz310.

* **Bode, H.B., et al., (2002).** Big effects from small changes: Possible ways to explore nature's chemical diversity. ChemBioChem 3 (7), 619. doi : 10.1002/1439-7633(20020703)3:7o619::AID-CBIC61943.0.CO;2-9.

* **Boettger, D., et al., (2012).** Evolutionary imprint of catalytic domains in fungal PKSNRPS hybrids. ChemBioChem 13 (16), 2363–2373. doi : 10.1002/cbic.201200449.

* **Bok, J.W., et al., (2009).** Chromatin-level regulation of biosynthetic gene clusters. Nature Chemical Biology 5 (7), 462–464. doi : 10.1038/nchembio.177.

* **Bok, J.W., et al., (2015).** Fungal artificial chromosomes for mining of the fungal secondary metabolome. BMC Genomics 16 (1), 343. doi : 10.1186/s12864-015-1561-x.

Bok, J.W., Keller, N.P., (2004). LaeA, a regulator of secondary metabolism in *Aspergillus* spp. Eukaryotic cell 3 (2), 527–535. doi : 10.1128/ec.3.2.527-535.2004.

Brown, D.W., et al., (2015). Identification of a 12-gene fusaric acid biosynthetic gene cluster in *Fusarium* species through comparative and functional genomics. Molecular Plant-Microbe Interactions 28 (3), 319–332. doi : 10.1094/MPMI-09-14-0264-R.

Bushley, K.E., et al., (2013). The genome of *Tolypocladium inflatum*: Evolution, organization, and expression of the cyclosporin biosynthetic gene cluster. PLoS Genetics 9 (6), e1003496. doi : 10.1371/journal.pgen.1003496.

Bushley, K.E., Turgeon, B.G., (2010). Phylogenomics reveals subfamilies of fungal nonribosomal peptide synthetases and their evolutionary relationships. BMC Evolutionary Biology 10 (1), 26. doi : 10.1186/1471-2148-10-26.

Calvo, A.M., Cary, J.W., (2015). Association of fungal secondary metabolism and sclerotial biology. Frontiers in Microbiology 6, 62. doi : 10.3389/fmicb.2015.00062.

Carroll, C.S., et al., (2017). The rhizoferrin biosynthetic gene in the fungal pathogen *Rhizopus delemar* is a novel member of the NIS gene family. The International Journal of Biochemistry & Cell Biology 89, 136–146. doi : 10.1016/J.BIOCEL.2017.06.005.

Carroll, C.S., Moore, M.M., (2018). Ironing out siderophore biosynthesis: A review of non-ribosomal peptide synthetase (NRPS)-independent siderophore synthetases. Critical Reviews in Biochemistry and Molecular Biology 53 (4), 356–381. doi : 10.1080/10409238.2018.1476449.

Chai, L.Y.A., et al., (2010). *Aspergillus fumigatus* conidial melanin modulates host cytokine response. Immunobiology 215 (11), 915–920. doi : 10.1016/J.IMBIO.2009.10.002.

Chankhamjon, P., et al., (2014). Biosynthesis of the halogenated mycotoxin Aspirochlorine in Koji mold involves a cryptic amino acid conversion. Angewandte Chemie International Edition 53 (49), 13409–13413. doi : 10.1002/anie.201407624.

Chen, Y., et al., (2018). Wheat microbiome bacteria can reduce virulence of a plant pathogenic fungus by altering histone acetylation. Nature Communications 9 (1), 3429. doi : 10.1038/s41467-018-05683-7.

Chiang, Y.-M., et al., (2009). A gene cluster containing two fungal polyketide synthases encodes the biosynthetic pathway for a polyketide, Asperfuranone, in *Aspergillus nidulans*. Journal of the American Chemical Society 131 (8), 2965–2970. doi : 10.1021/ja8088185.

Chiang, Y.-M., et al., (2013). An efficient system for heterologous expression of secondary metabolite genes in *Aspergillus nidulans*. Journal of the American Chemical Society 135 (20), 7720–7731. doi : 10.1021/ja401945a.

Choquer, M., et al., (2005). The CTB1 gene encoding a fungal polyketide synthase is required for cercosporin biosynthesis and fungal virulence of *Cercospora nicotianae*. Molecular Plant-Microbe Interactions 18 (5), 468–476. doi : 10.1094/MPMI-18-0468.

Chumley, F.G., Valent, B., (1990). Genetic analysis of melanin-deficient, nonpathogenic mutants of *Magnaporthe grisea*. Molecular Plant-Microbe Interactions 3 (3), 135. doi : 10.1094/MPMI-3-135.

Clevenger, K.D., et al., (2017). A scalable platform to identify fungal secondary metabolites and their gene clusters. Nature Chemical Biology 13 (8), 895–901. doi : 10.1038/nchembio.2408.

Collemare, J., et al., (2014). Secondary metabolism and biotrophic lifestyle in the tomato pathogen *Cladosporium fulvum*. PLoS One 9 (1), doi : 10.1371/journal.pone.0085877.

Collemare, J., Seidl, M.F., (2019). Chromatin-dependent regulation of secondary metabolite biosynthesis in fungi: Is the picture complete? FEMS Microbiology Reviews. doi : 10.1093/femsre/fuz018.

Collemare, J., O'Connell, R., Lebrun, M., (2019). Nonproteinaceous effectors: The terra incognita of plant – Fungal interactions. New Phytologist 223 (2), 590–596. doi : 10.1111/nph.15785.

Colombo, D., Ammirati, E., (2011). Cyclosporine in transplantation – A history of converging timelines. Journal of Biological Regulators and Homeostatic Agents 25 (4), 493–504.

Connolly, L.R., Smith, K.M., Freitag, M., et al., (2013). The *Fusarium graminearum* histone H3 K27 methyltransferase KMT6 regulates development and expression of secondary metabolite gene clusters. PLoS Genetics 9 (10), e1003916. doi : 10.1371/journal.pgen.1003916.

* **Cook, D., et al., (2017).** Swainsonine biosynthesis genes in diverse symbiotic and pathogenic fungi. G3 7 (6), 1791–1797. doi : 10.1534/g3.117.041384.

* **Cox, R.J., (2007).** Polyketides, proteins and genes in fungi: Programmed nanomachines begin to reveal their secrets. Organic & Biomolecular Chemistry 5 (13), 2010. doi : 10.1039/b704420h.

* **Dadachova, E., et al., (2007).** The radioprotective properties of fungal melanin are a function of its chemical composition, stable radical presence and spatial arrangement. Pigment Cell & Melanoma Research 21 (2), 192–199. doi : 10.1111/j.1755-148X.2007.00430.x.

* **Degen, A., et al., (2019).** Context-dependent activity of A domains in the tyrocidine synthetase. Scientific Reports 9 (1), 5119. doi : 10.1038/s41598-019-41492-8.

* **Ditengou, F.A., et al., (2015).** Volatile signalling by sesquiterpenes from ectomycorrhizal fungi reprogrammes root architecture. Nature Communications 6 (1), 6279. doi : 10.1038/ncomms7279.

* **Dolan, S.K., et al., (2015).** Resistance is not futile: Gliotoxin biosynthesis, functionality and utility. Trends in Microbiology 23 (7), 419–428. doi : 10.1016/J.TIM.2015.02.005.

* **Eberharter, A., Becker, P.B., (2002).** Histone acetylation: A switch between repressive and permissive chromatin: Second in review series on chromatin dynamics. EMBO Reports 3 (3), 224–229. doi : 10.1093/embo-reports/kvf053.

* **Ebert, M.K., et al., (2019).** Gene cluster conservation identifies melanin and perylenequinone biosynthesis pathways in multiple plant pathogenic fungi. Environmental Microbiology 21 (3), 913–927. doi : 10.1111/1462-2920.14475.

* **Eddy, S.R., (1998).** Profile hidden Markov models. Bioinformatics 14 (9), 755–763. doi : 10.1093/bioinformatics/14.9.755.

* **Ehrlich, K.C., Montalbano, B.G., Cary, J.W., (1999).** Binding of the C6-zinc cluster protein, AFLR, to the promoters of aflatoxin pathway biosynthesis genes in *Aspergillus parasiticus*. Gene 230 (2), 249–257. doi : 10.1016/s0378-1119(99)00075-x.

* **El-Gebali, S., et al., (2019).** The Pfam protein families database in 2019. Nucleic Acids Research 47 (D1), D427–D432. doi : 10.1093/nar/gky995.

* **Engler, C., Marillonnet, S., (2014).** Golden Gate Cloning. Totowa, NJ: Humana Press, pp. 119–131. 10.1007/978-1-62703-764-8_9.

* **Fernandes, M., Keller, N.P., Adams, T.H., (1998).** Sequence-specific binding by *Aspergillus nidulans* AflR, a C6 zinc cluster protein regulating mycotoxin biosynthesis. Molecular Microbiology 28 (6), 1355–1365. doi : 10.1046/j.1365-2958.1998.00907.x.

* **Fisch, K.M., et al., (2009).** Chemical induction of silent biosynthetic pathway transcription in *Aspergillus niger*. Journal of Industrial Microbiology & Biotechnology 36 (9), 1199–1213. doi : 10.1007/s10295-009-0601-4.

* **Fisch, K.M., et al., (2011).** Rational domain swaps decipher programming in fungal highly reducing polyketide synthases and resurrect an extinct metabolite. Journal of the American Chemical Society 133 (41), 16635–16641. doi : 10.1021/ja206914q.

* **Fisch, K.M., (2013).** Biosynthesis of natural products by microbial iterative hybrid PKS–NRPS. RSC Advances 3 (40), 18228. doi : 10.1039/c3ra42661k.

* **Fischer, J., et al., (2018).** Chromatin mapping identifies BasR, a key regulator of bacteria-triggered production of fungal secondary metabolites. eLife 7. doi : 10.7554/eLife.40969.

* **Fischer, R., (2008).** Sex and poison in the dark. Science 320 (5882), doi : 10.1126/science.1160123.

* **Frandsen, R.J.N., et al., (2018).** Heterologous production of the widely used natural food colorant carminic acid in *Aspergillus nidulans*. Scientific Reports 8 (1), 12853. doi : 10.1038/s41598-018-30816-9.

* **Fricke, J., Blei, F., Hoffmeister, D., (2017).** Enzymatic synthesis of Psilocybin. Angewandte Chemie International Edition 56 (40), 12352–12355. doi : 10.1002/anie.201705489.

* **Gacek, A., Strauss, J., (2012).** The chromatin code of fungal secondary metabolite gene clusters. Applied Microbiology and Biotechnology 95 (6), 1389–1404. doi : 10.1007/s00253-012-4208-8.

* **Gacek-Matthews, A., et al., (2015).** KdmA, a histone H3 demethylase with bipartite function, differentially regulates primary and secondary metabolism in *Aspergillus nidulans*.

Molecular Microbiology 96 (4), 839–860. doi : 10.1111/mmi.12977.

Gardiner, D.M., Jarvis, R.S., Howlett, B.J., (2005). The ABC transporter gene in the sirodesmin biosynthetic gene cluster of *Leptosphaeria maculans* is not essential for sirodesmin production but facilitates self-protection. Fungal Genetics and Biology 42 (3), 257–263. doi : 10.1016/j.fgb.2004.12.001.

Geib, E., Brock, M., (2017). ATNT: An enhanced system for expression of polycistronic secondary metabolite gene clusters in *Aspergillus niger*. Fungal Biology and Biotechnology 4 (1), 13. doi : 10.1186/s40694-017-0042-1.

Gibson, D.M., et al., (2014). Discovering the secondary metabolite potential encoded within entomopathogenic fungi. Natural Product Reports 31 (10), 1287–1305. doi : 10.1039/C4NP00054D.

Gressler, M., et al., (2015). A new high-performance heterologous fungal expression system based on regulatory elements from the *Aspergillus terreus* terrein gene cluster. Frontiers in Microbiology 6, 184. doi : 10.3389/fmicb.2015.00184.

Griffiths, S., et al., (2015). Regulation of secondary metabolite production in the fungal tomato pathogen *Cladosporium fulvum*. Fungal Genetics and Biology 84, 52–61. doi : 10.1016/j.fgb.2015.09.009.

Griffiths, S., et al., (2016). Elucidation of cladofulvin biosynthesis reveals a cytochrome P450 monooxygenase required for anthraquinone dimerization. Proceedings of the National Academy of Sciences of the United States of America 113 (25), 6851–6856. doi : 10.1073/pnas.1603528113.

Griffiths, S., et al., (2018). Down-regulation of cladofulvin biosynthesis is required for biotrophic growth of *Cladosporium fulvum* on tomato. Molecular Plant Pathology 19 (2), 369–380. doi : 10.1111/mpp.12527.

Grigoriev, I.V., et al., (2014). MycoCosm portal: Gearing up for 1000 fungal genomes. Nucleic Acids Research 42 (D1), D699–D704. doi : 10.1093/nar/gkt1183.

Haft, D.H., et al., (2001). TIGRFAMs: A protein family resource for the functional identification of proteins. Nucleic Acids Research 29 (1), 41–43. doi : 10.1093/nar/29.1.41.

Han, J., et al., (2019). Genome- and MS-based mining of antibacterial chlorinated chromones and xanthones from the phytopathogenic fungus *Bipolaris sorokiniana* strain 11134. Applied Microbiology and Biotechnology 103 (13), 5167–5181. doi : 10.1007/s00253-019-09821-z.

Harvey, A.L., Edrada-Ebel, R., Quinn, R.J., (2015). The re-emergence of natural products for drug discovery in the genomics era. Nature Reviews Drug Discovery 14 (2), 111–129. doi : 10.1038/nrd4510.

Harvey, C.J.B., et al., (2018). HEx: A heterologous expression platform for the discovery of fungal natural products. Science Advances 4 (4), eaar5459. doi : 10.1126/sciadv.aar5459.

Henrikson, J.C., et al., (2009). A chemical epigenetics approach for engineering the in situ biosynthesis of a cryptic natural product from *Aspergillus niger*. Organic and Biomolecular Chemistry 7 (3), 435–438. doi : 10.1039/b819208a.

Herbst, D.A., Townsend, C.A., Maier, T., (2018). The architectures of iterative type I PKS and FAS. Natural Product Reports 35 (10), 1046–1069. doi : 10.1039/C8NP00039E.

Hong, E.J., et al., (2015). Overexpression of the *laeA* gene leads to increased production of cyclopiazonic acid in *Aspergillus fumisynnematus*. Fungal Biology 119 (11), 973–983. doi : 10.1016/J.FUNBIO.2015.06.006.

Hortschansky, P., et al., (2017). The CCAAT-binding complex (CBC) in *Aspergillus* species. Biochimica et Biophysica Acta (BBA) – Gene Regulatory Mechanisms 1860 (5), 560–570. doi : 10.1016/J.BBAGRM.2016.11.008.

Houbraken, J., Frisvad, J.C., Samson, R.A., (2011). Fleming's penicillin producing strain is not *Penicillium chrysogenum* but *P. rubens*. IMA Fungus 2 (1), 87–95. doi : 10.5598/imafungus.2011.02.01.12.

Hurst, L.D., Pál, C., Lercher, M.J., (2004). The evolutionary dynamics of eukaryotic gene order. Nature Reviews Genetics 5 (4), 299–310. doi : 10.1038/nrg1319.

Hyde, K.D., et al., (2019). The amazing potential of fungi: 50 ways we can exploit fungi industrially. Fungal Diversity 97 (1), 1–136. doi : 10.1007/s13225-019-00430-9.

Ishiuchi, K., et al., (2012). Establishing a new methodology for genome mining and biosynthesis of polyketides and peptides through yeast molecular genetics. ChemBioChem 13 (6), 846–854.

doi : 10.1002/cbic.201100798.

Itoh, T., et al., (2010). Reconstitution of a fungal meroterpenoid biosynthesis reveals the involvement of a novel family of terpene cyclases. Nature Chemistry 2 (10), 858–864. doi : 10.1038/nchem.764.

Jacobus, A.P., Gross, J., (2015). Optimal cloning of PCR fragments by homologous recombination in *Escherichia coli*. PLoS One 10 (3), e0119221. doi : 10.1371/journal.pone.0119221.

Jin, F.J., et al., (2004). Development of a novel quadruple auxotrophic host transformation system by *argB* gene disruption using *adeA* gene and exploiting adenine auxotrophy in *Aspergillus oryzae*. FEMS Microbiology Letters 239 (1), 79–85. doi : 10.1016/j.femsle.2004.08.025.

Kaneko, A., et al., (2019). Post-genomic approach based discovery of alkylresorcinols from a cricket-associated fungus, *Penicillium soppi*. Organic & Biomolecular Chemistry 17 (21), 5239–5243. doi : 10.1039/C9OB00807A.

Kautsar, S.A., et al., (2019). MIBiG 2.0: A repository for biosynthetic gene clusters of known function. Nucleic Acids Research. doi : 10.1093/nar/gkz882.

Keller, N.P., (2019). Fungal secondary metabolism: Regulation, function and drug discovery. Nature Reviews Microbiology 17 (3), 167–180. doi : 10.1038/s41579-018-0121-1.

Keller, N.P., Hohn, T.M., (1997). Metabolic pathway gene clusters in filamentous fungi. Fungal Genetics and Biology 21 (1), 17–29. doi : 10.1006/FGBI.1997.0970.

Khaldi, N., et al., (2010). SMURF: Genomic mapping of fungal secondary metabolite clusters. Fungal Genetics and Biology 47 (9), 736–741. doi : 10.1016/J.FGB.2010.06.003.

Kim, Y.-T., et al., (2005). Two different polyketide synthase genes are required for synthesis of zearalenone in *Gibberella zeae*. Molecular Microbiology 58 (4), 1102–1113. doi : 10.1111/j.1365-2958.2005.04884.x.

Kindinger, F., et al., (2019). Genomic locus of a *Penicillium crustosum* pigment as an integration site for secondary metabolite gene expression. ACS Chemical Biology 14 (6), 1227–1234. doi : 10.1021/acschembio.9b00164.

Kjærbølling, I., et al., (2018). Linking secondary metabolites to gene clusters through genome sequencing of six diverse *Aspergillus* species. Proceedings of the National Academy of Sciences of the United States of America 115 (4), E753–E761. doi : 10.1073/pnas.1715954115.

Kjærbølling, I., et al., (2019). Strategies to establish the link between biosynthetic gene clusters and secondary metabolites. Fungal Genetics and Biology 130, 107–121. doi : 10.1016/J.FGB.2019.06.001.

Kjærbølling, I., Vesth, T., Andersen, M.R., (2019). Resistance gene-directed genome mining of 50 *Aspergillus* species. mSystems 4 (4), doi : 10.1128/MSYSTEMS.00085-19.

Koczyk, G., Dawidziuk, A., Popiel, D., (2015). The distant siblings – A phylogenomic roadmap illuminates the origins of extant diversity in fungal aromatic polyketide biosynthesis. Genome Biology and Evolution 7 (11), 3132–3154. doi : 10.1093/gbe/evv204.

König, C.C., et al., (2013). Bacterium induces cryptic meroterpenoid pathway in the pathogenic fungus *Aspergillus fumigatus*. ChemBioChem 14 (8), 938–942. doi : 10.1002/cbic.201300070.

Kouprina, N., Larionov, V., (2016). Transformation-associated recombination (TAR) cloning for genomics studies and synthetic biology. Chromosoma 125 (4), 621–632. doi : 10.1007/s00412-016-0588-3.

Langfelder, K., et al., (2003). Biosynthesis of fungal melanins and their importance for human pathogenic fungi. Fungal Genetics and Biology 38 (2), 143–158. doi : 10.1016/S1087-1845(02)00526-1.

Lee, I., et al., (2009). HdaA, a class 2 histone deacetylase of *Aspergillus fumigatus*, affects germination and secondary metabolite production. Fungal Genetics and Biology 46 (10), 782–790. doi : 10.1016/J.FGB.2009.06.007.

Li, J.W.-H., Vederas, J.C., (2009). Drug discovery and natural products: End of an era or an endless frontier? Science 325 (5937), 161–165. doi : 10.1126/science.1168243.

Lind, A.L., et al., (2017). Drivers of genetic diversity in secondary metabolic gene clusters within a fungal species. PLoS Biology 15 (11), 1–26. doi : 10.1371/journal. pbio.2003583.

Lind, A.L., et al., (2018). An LaeA- and BrlA-dependent cellular network governs

tissuespecific secondary metabolism in the Human pathogen *Aspergillus fumigatus*. mSphere 3 (2), doi : 10.1128/mSphere.00050-18.

- **Liu, L., et al., (2015).** Bioinformatical analysis of the sequences, structures and functions of fungal polyketide synthase product template domains. Scientific Reports 5 (1), 10463. doi : 10.1038/srep10463.

- **Liu, T., et al., (2014).** Rational domain swaps reveal insights about chain length control by ketosynthase domains in fungal nonreducing polyketide synthases. Organic Letters 16 (6), 1676–1679. doi : 10.1021/ol5003384.

- **Lyu, H.-N., et al., (2019).** Harnessing diverse transcriptional regulators for natural product discovery in fungi. Natural Product Reports. doi : 10.1039/C8NP00027A.

- **Machida, M., et al., (2005).** Genome sequencing and analysis of *Aspergillus oryzae*. Nature 438 (7071), 1157–1161. doi : 10.1038/nature04300.

- **Maeda, K., et al., (2017).** Increased metabolite production by deletion of an HDA1-type histone deacetylase in the phytopathogenic fungi, *Magnaporthe oryzae* (*Pyricularia oryzae*) and *Fusarium asiaticum*. Letters in Applied Microbiology 65 (5), 446–452. doi : 10.1111/lam.12797.

- **Mattern, D.J., et al., (2017).** Rewiring of the Austinoid biosynthetic pathway in filamentous fungi. ACS Chemical Biology 12 (12), 2927–2933. doi : 10.1021/acschembio.7b00814.

- **McGary, K.L., Slot, J.C., Rokas, A., (2013).** Physical linkage of metabolic genes in fungi is an adaptation against the accumulation of toxic intermediate compounds. Proceedings of the National Academy of Sciences of the United States of America 110 (28), 11481–11486. doi : 10.1073/pnas.1304461110.

- **Medema, M.H., et al., (2015).** Minimum Information about a Biosynthetic Gene cluster. Nature Chemical Biology 11 (9), doi : 10.1038/nchembio.1890.

- **Medema, M.H., Blin, K., et al., (2011).** antiSMASH: Rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. Nucleic Acids Research 39 (suppl_2), W339–W346. doi : 10.1093/nar/gkr466.

- **Medema, M.H., Breitling, R., et al., (2011).** Exploiting plug-and-play synthetic biology

for drug discovery and production in microorganisms. Nature Reviews Microbiology 9 (2), 131–137. doi : 10.1038/nrmicro2478.

- **Medema, M.H., Fischbach, M.A., (2015).** Computational approaches to natural product discovery. Nature Chemical Biology 11 (9), 639–648. doi : 10.1038/nchembio.1884.

- **Medema, M.H., Takano, E., Breitling, R., (2013).** Detecting sequence homology at the gene cluster Level with multigeneblast. Molecular Biology and Evolution 30 (5), 1218–1223. doi : 10.1093/molbev/mst025.

- **Mukherjee, G., Singh, S.K., (2011).** Purification and characterization of a new red pigment from *Monascus purpureus* in submerged fermentation. Process Biochemistry 46 (1), 188–192. doi : 10.1016/J.PROCBIO.2010.08.006.

- **Narita, K., et al., (2017).** Focused genome mining of structurally related sesterterpenes: Enzymatic formation of enantiomeric and diastereomeric products. Organic Letters 19 (24), 6696–6699. doi : 10.1021/acs.orglett.7b03418.

- **Navarro-Muñoz, J.C., et al., (2019).** A computational framework to explore large-scale biosynthetic diversity. Nature Chemical Biology 16 (1), 60–68. doi : 10.1038/s41589-019- 0400-9.

- **Nielsen, M.R., et al., (2019).** Heterologous expression of intact biosynthetic gene clusters in *Fusarium graminearum*. Fungal Genetics and Biology 132, 103248. doi : 10.1016/J.FGB.2019.103248.

- **Nielsen, M.T., et al., (2013).** Heterologous reconstitution of the intact Geodin gene cluster in *Aspergillus nidulans* through a simple and versatile PCR based approach. PLoS One 8 (8), e72871 doi : 10.1371/journal.pone.0072871.

- **Nützmann, H.-W., et al., (2011).** Bacteria-induced natural product formation in the fungus *Aspergillus nidulans* requires Saga/Ada-mediated histone acetylation. Proceedings of the National Academy of Sciences of the United States of America 108 (34), 14282–14287. doi : 10.1073/pnas.1103523108.

- **Nützmann, H.-W., et al., (2013).** Distinct amino acids of histone H3 control secondary metabolism in *Aspergillus nidulans*. Applied and Environmental Microbiology 79 (19), 6102–

6109. doi : 10.1128/AEM.01578-13.

* **O'Brien, J., Wright, G.D., (2011).** An ecological perspective of microbial secondary metabolism. Current Opinion in Biotechnology 22 (4), 552–558. doi : 10.1016/J.CO PBIO.2011.03.010.

* **Pahirulzaman, K.A., Williams, K., Lazarus, C.M., (2012).** A toolkit for heterologous expression of metabolic pathways in *Aspergillus oryzae*. Methods in Enzymology. 241–260. doi : 10.1016/B978-0-12-404634-4.00012-7.

* **Pfannenstiel, B.T., Keller, N.P., (2019).** On top of biosynthetic gene clusters: How epigenetic machinery influences secondary metabolism in fungi. Biotechnology Advances 37 (6), 107345. doi : 10.1016/j.biotechadv.2019.02.001.

* **Pidroni, A., et al., (2018).** A class 1 histone deacetylase as major regulator of secondary metabolite production in *Aspergillus nidulans*. Frontiers in Microbiology 9, 2212. doi : 10.3389/fmicb.2018.02212.

* **Reece-Hoyes, J.S., Walhout, A.J.M., (2018).** Gateway Recombinational Cloning. Cold Spring Harbor protocols. Cold Spring Harbor Laboratory Press. pdb.top094912. doi : 10.1101/pdb.top094912.

* **Regueira, T.B., et al., (2011).** Molecular basis for mycophenolic acid biosynthesis in *Penicillium brevicompactum*. Applied and Environmental Microbiology 77 (9), 3035–3043. doi : 10.1128/AEM.03015-10.

* **Reverberi, M., et al., (2012).** Aoyap1 regulates OTA synthesis by controlling cell redox balance in *Aspergillus ochraceus*. Applied Microbiology and Biotechnology 95 (5), 1293–1304. doi : 10.1007/s00253-012-3985-4.

* **Reyes-Dominguez, Y., et al., (2010).** Heterochromatic marks are associated with the repression of secondary metabolism clusters in *Aspergillus nidulans*. Molecular Microbiology 76 (6), 1376–1386. doi : 10.1111/j.1365-2958.2010.07051.x.

* **Richard, J.L., (2008).** Discovery of aflatoxins and significant historical features. Toxin Reviews 27 (3–4), 171–201. doi : 10.1080/15569540802462040.

* **Richter, L., et al., (2014).** Engineering of *Aspergillus niger* for the production of secondary metabolites. Fungal Biology and Biotechnology 1 (1), 4. doi : 10.1186/s40694-014-0004-9.

* **Rokas, A., Wisecaver, J.H., Lind, A.L., (2018).** The birth, evolution and death of metabolic gene clusters in fungi. Nature Reviews Microbiology 16 (12), 731–744. doi : 10.1038/s41579-018-0075-3.

* **Romano, S., et al., (2018).** Extending the "One strain many compounds" (OSMAC) principle to marine microorganisms. Marine Drugs 16 (7), 244. doi : 10.3390/md16070244.

* **Romsdahl, J., Wang, C.C.C., (2019).** Recent advances in the genome mining of: *Aspergillus* secondary metabolites (covering 2012-2018). MedChemComm. 840–866. doi : 10.1039/c9md00054b.

* **Rösler, S.M., et al., (2016).** The SAGA complex in the rice pathogen *Fusarium fujikuroi*: Structure and functional characterization. Molecular Microbiology 102 (6), 951–974. doi : 10.1111/mmi.13528.

* **Santoni, D., Castiglione, F., Paci, P., (2013).** Identifying correlations between chromosomal proximity of genes and distance of their products in protein-protein interaction networks of yeast. PLoS One 8 (3), e57707. doi : 10.1371/journal.pone.0057707.

* **Schardl, C.L., et al., (2013).** The Epichloae: Alkaloid diversity and roles in symbiosis with grasses. Current Opinion in Plant Biology 16 (4), 480–488. doi : 10.1016/J. PBI.2013.06.012.

* **Scherlach, K., et al., (2010).** Aspernidine A and B, prenylated isoindolinone alkaloids from the model fungus *Aspergillus nidulans*. The Journal of Antibiotics 63 (7), 375–377. doi : 10.1038/ja.2010.46.

* **Scherlach, K., Hertweck, C., (2006).** Discovery of aspoquinolones A–D, prenylated quinoline-2-one alkaloids from *Aspergillus nidulans*, motivated by genome mining. Organic and Biomolecular Chemistry 4 (18), 3517–3520. doi : 10.1039/B607011F.

* **Schmidt-Dannert, C., (2015).** Biosynthesis of terpenoid natural products in fungi. Advances in Biochemical Engineering/Biotechnology 148, 19–61. doi : 10.1007/10_2014_283.

* **Schneider, P., Bouhired, S., Hoffmeister, D., (2008).** Characterization of the atromentin biosynthesis genes and enzymes in the homobasidiomycete *Tapinella panuoides*. Fungal Genetics and Biology 45 (11), 1487–1496. doi : 10.1016/J.FGB.2008.08.009.

* **Schroeckh, V., et al., (2009).** Intimate bacterial-fungal interaction triggers biosynthesis of archetypal polyketides in *Aspergillus nidulans*. Proceedings of the National Academy of Sciences of the United States of America 106 (34), 14558–14563. doi : 10.1073/pnas.0901870106.

* **Schwecke, T., et al., (2006).** Nonribosomal peptide synthesis in *Schizosaccharomyces pombe* and the architectures of ferrichrome-type siderophore synthetases in fungi. ChemBioChem 7 (4), 612–622. doi : 10.1002/cbic.200500301.

* **Schwelm, A., Bradshaw, R.E., (2010).** Genetics of dothistromin biosynthesis of *Dothistroma septosporum*: An update. Toxins 2 (11), 2680–2698. doi : 10.3390/toxins2112680.

* **Shimizu, Y., Ogata, H., Goto, S., (2017).** Type III polyketide synthases: Functional classification and phylogenomics. ChemBioChem 18 (1), 50–65. doi : 10.1002/cbic.201600522.

* **Shwab, E.K., et al., (2007).** Histone deacetylase activity regulates chemical diversity in *Aspergillus*. Eukaryotic Cell 6 (9), 1656–1664. doi : 10.1128/EC.00186-07.

* **Singh, K., et al., (2019).** Genetics and molecular biology of genes encoding Cephalosporin biosynthesis in microbes. New and Future Developments in Microbial Biotechnology and Bioengineering. 25–34. doi : 10.1016/B978-0-444-63503-7.00002-4.

* **Singh, M., Chaudhary, S., Sareen, D., (2017).** Non-ribosomal peptide synthetases: Identifying the cryptic gene clusters and decoding the natural product. Journal of Biosciences 42 (1), 175–187. doi : 10.1007/s12038-017-9663-z.

* **Skellam, E., (2017).** The biosynthesis of cytochalasans. Natural Product Reports 34 (11), 1252–1263. doi : 10.1039/C7NP00036G.

* **Skellam, E., (2019).** Strategies for engineering natural product biosynthesis in fungi. Trends in Biotechnology 37 (4), 416–427. doi : 10.1016/j.tibtech.2018.09.003.

* **Song, Z., et al., (2015).** Heterologous expression of the avirulence gene ACE1 from the fungal rice pathogen *Magnaporthe oryzae*. Chemical Science 6 (8), 4837–4845. doi : 10.1039/c4sc03707c.

* **Soukup, A.A., et al., (2012).** Overexpression of the *Aspergillus nidulans* histone 4 acetyltransferase EsaA increases activation of secondary metabolite production. Molecular Microbiology 86 (2), 314–330. doi : 10.1111/j.1365-2958.2012.08195.x.

* **Spraker, J.E., et al., (2018).** Conserved responses in a war of small molecules between a plant-pathogenic bacterium and fungi. mBio 9 (3), doi : 10.1128/mBio.00820-18.

* **Stergiopoulos, I., et al., (2013).** Phytotoxic secondary metabolites and peptides produced by plant pathogenic Dothideomycete fungi. FEMS Microbiology Reviews 37 (1), 67–93. doi : 10.1111/j.1574-6976.2012.00349.x.

* **Stierle, A.A., et al., (2017).** The Berkeleylactones, antibiotic macrolides from fungal coculture. Journal of Natural Products 80 (4), 1150–1160. doi : 10.1021/acs.jnatprod.7b00133.

* **Stöckli, M., et al., (2019).** Bacteria-induced production of the antibacterial sesquiterpene lagopodin B in *Coprinopsis cinerea*. Molecular Microbiology 112 (2), 605–619. doi : 10.1111/mmi.14277.

* **Strauss, J., Reyes-Dominguez, Y., (2011).** Regulation of secondary metabolism by chromatin structure and epigenetic codes. Fungal Genetics and Biology 48 (1), 62–69. doi : 10.1016/j.fgb.2010.07.009.

* **Studt, L., et al., (2013).** Two histone deacetylases, FfHda1 and FfHda2, are important for *Fusarium fujikuroi* secondary metabolism and virulence. Applied and Environmental Microbiology 79 (24), 7719–7734. doi : 10.1128/AEM.01557-13.

* **Studt, L., et al., (2016).** Knock-down of the methyltransferase Kmt6 relieves H3K27me3 and results in induction of cryptic and otherwise silent secondary metabolite gene clusters in *Fusarium fujikuroi*. Environmental Microbiology 18 (11), 4037–4054. doi : 10.1111/1462-2920.13427.

* **Takahashi, J.A., et al., (2013).** Classical and epigenetic approaches to metabolite diversification in filamentous fungi. Phytochemistry Reviews 12 (4), 773–789. doi : 10.1007/s11101-013-9305-5.

* **Takeda, I., et al., (2014).** Motif-independent prediction of a secondary metabolism gene cluster using comparative genomics: Application to sequenced genomes of *Aspergillus* and ten other filamentous fungal species. DNA Research 21 (4), 447–457. doi : 10.1093/dnares/dsu010.

2

* **Terabayashi, Y., et al., (2010).** Identification and characterization of genes responsible for biosynthesis of kojic acid, an industrially important compound from *Aspergillus oryzae*. Fungal Genetics and Biology 47 (12), 953–961. doi : 10.1016/J.FGB.2010.08.014.

* **Theobald, S., et al., (2018).** Uncovering secondary metabolite evolution and biosynthesis using gene cluster networks and genetic dereplication. Scientific Reports 8 (1), 17957. doi : 10.1038/s41598-018-36561-3.

* **Throckmorton, K., Wiemann, P., Keller, N., (2015).** Evolution of chemical diversity in a group of non-reduced polyketide gene clusters: Using phylogenetics to inform the search for novel fungal natural products. Toxins 7 (9), 3572–3607. doi : 10.3390/toxins7093572.

* **Thywißen, A., et al., (2011).** Conidial Dihydroxynaphthalene Melanin of the human pathogenic fungus *Aspergillus fumigatus* interferes with the host endocytosis pathway. Frontiers in Microbiology 2, 96. doi : 10.3389/fmicb.2011.00096.

* **Turner, B.M., (2007).** Defining an epigenetic code. Nature Cell Biology 9 (1), 2–6. doi : 10.1038/ncb0107-2.

* **Umemura, M., et al., (2013).** MIDDAS-M: Motif-Independent de novo detection of secondary metabolite gene clusters through the integration of genome sequencing and transcriptome data. PLoS One 8 (12), e84028 doi : 10.1371/journal.pone.0084028.

* **Umemura, M., Koike, H., Machida, M., (2015).** Motif-independent de novo detection of secondary metabolite gene clusters – Toward identification from filamentous fungi. Frontiers in Microbiology 6, 371. doi : 10.3389/fmicb.2015.00371.

* **Unkles, S.E., et al., (2014).** Synthetic biology tools for bioprospecting of natural products in eukaryotes. Chemistry & Biology 21 (4), 502–508. doi : 10.1016/j.chembiol.2014.02.010.

* **Valiante, V., et al., (2017).** Discovery of an extended Austinoid biosynthetic pathway in *Aspergillus calidoustus*. ACS Chemical Biology 12 (5), 1227–1234. doi : 10.1021/acschembio.7b00003.

* **van Dongen, P.W., de Groot, A.N., (1995).** History of ergot alkaloids from ergotism to ergometrine. European Journal of Obstetrics & Gynecology and Reproductive Biology 60 (2), 109–116. doi : 10.1016/0028-2243(95)02104-Z.

* **van Santen, J.A., et al., (2019).** The natural products atlas: An open access knowledge base for microbial natural products discovery. ACS Central Science 5 (11), 1824–1833. doi : 10.1021/acscentsci.9b00806.

* **Vervoort, H.C., Draskovic, M., Crews, P., (2011).** Histone deacetylase inhibitors as a tool to up-regulate new fungal biosynthetic products : Isolation of EGM-556, a cyclodepsipeptide, from *Microascus* sp. Organic Letters 13 (3), 410–413. doi : 10.1021/ol1027199.

* **Vesth, T.C., Brandl, J., Andersen, M.R., (2016).** FunGeneClusterS: Predicting fungal gene clusters from genome and transcriptome data. Synthetic and Systems Biotechnology 1 (2), 122–129. doi : 10.1016/j.synbio.2016.01.002.

* **Vogt, E., Künzler, M., (2019).** Discovery of novel fungal RiPP biosynthetic pathways and their application for the development of peptide therapeutics. Applied Microbiology and Biotechnology 103 (14), 5567–5581. doi : 10.1007/s00253-019-09893-x.

* **Walsh, C., Tang, Y., (2017).** Natural product biosynthesis : Chemical logic and enzymatic machinery. Available at: https://pubs.rsc.org/en/content/ebook/978-1-78801-076-4

* **Wang, X., et al., (2010).** Chemical epigenetics alters the secondary metabolite composition of guttate excreted by an Atlantic-forest-soil-derived *Penicillium citreonigrum*. Journal of Natural Products 73 (5), 942–948. doi : 10.1021/np100142h.

* **Watanabe, A., Ebizuka, Y., (2002).** A novel hexaketide naphthalene synthesized by a chimeric polyketide synthase composed of fungal pentaketide and heptaketide synthases. Tetrahedron Letters 43 (5), 843–846. doi : 10.1016/S0040-4039(01)02251-1.

* **Wiemann, P., et al., (2009).** Biosynthesis of the red pigment bikaverin in *Fusarium fujikuroi* : Genes, their function and regulation. Molecular Microbiology 72 (4), 931–946. doi : 10.1111/j.1365-2958.2009.06695.x.

* **Wiemann, P., et al., (2010).** FfVel1 and FfLae1, components of a velvet-like complex in *Fusarium fujikuroi*, affect differentiation, secondary metabolism and virulence. Molecular Microbiology 77 (4), no-no. doi : 10.1111/j.1365-2958.2010.07263.x.

* **Wiemann, P., Keller, N.P., (2014).** Strategies for mining fungal natural products. Journal of Industrial Microbiology & Biotechnology 41 (2), 301–313. doi : 10.1007/s10295-013-1366-3.

* **Williams, R.B., et al., (2008).** Epigenetic remodeling of the fungal secondary metabolome. Organic & Biomolecular Chemistry 6 (11), 1895. doi : 10.1039/b804701d.

* **Wolf, T., et al., (2016).** CASSIS and SMIPS: Promoter-based prediction of secondary metabolite gene clusters in eukaryotic genomes. Bioinformatics 32 (8), 1138–1143. doi : 10.1093/bioinformatics/btv713.

* **Woloshuk, C.P., et al., (1994).** Molecular characterization of aflR, a regulatory locus for aflatoxin biosynthesis. Applied and Environmental Microbiology 60 (7), 2408–2414. Available at: http://www.ncbi.nlm.nih.gov/pubmed/8074521.

* **Xu, Y., et al., (2014).** Diversity-oriented combinatorial biosynthesis of benzenediol lactone scaffolds by subunit shuffling of fungal polyketide synthases. Proceedings of the National Academy of Sciences of the United States of America 111 (34), 12354–12359. doi : 10.1073/pnas.1406999111.

* **Yan, Y., et al., (2018).** Resistance-gene-directed discovery of a natural-product herbicide with a new mode of action. Nature 559 (7714), 415–418. doi : 10.1038/s41586-018-0319-4.

* **Yang, K., et al., (2016).** The DmtA methyltransferase contributes to *Aspergillus flavus* conidiation, sclerotial production, aflatoxin biosynthesis and virulence. Scientific Reports 6 (1), 23259. doi : 10.1038/srep23259.

* **Yang, X., et al., (2018).** Cyclosporine biosynthesis in *Tolypocladium inflatum* benefits fungal adaptation to the environment. mBio 9 (5), doi : 10.1128/mBio.01211-18.

* **Yang, X., van der Donk, W.A., (2013).** Ribosomally synthesized and post-translationally modified peptide natural products: New insights into the role of leader and core peptides during biosynthesis. Chemistry – A European Journal 19 (24), 7662–7677. doi : 10.1002/chem.201300401.

* **Yeh, H.-H., et al., (2012).** Molecular genetic analysis reveals that a nonribosomal peptide synthetase-like (NRPS-like) gene in *Aspergillus nidulans* is responsible for microperfuranone biosynthesis. Applied Microbiology and Biotechnology 96 (3), 739–748. doi : 10.1007/s00253-012-4098-9.

* **Yeh, H.-H., et al., (2016).** Resistance gene-guided genome mining: Serial promoter

exchanges in *Aspergillus nidulans* reveal the biosynthetic pathway for Fellutamide B, a proteasome inhibitor. ACS Chemical Biology 11 (8), 2275–2284. doi : 10.1021/acschembio.6b00213.

* **Yu, D., et al., (2017).** Decoding and reprogramming fungal iterative nonribosomal peptide synthetases. Nature Communications 8 (1), 15349. doi : 10.1038/ncomms15349.

* **Yun, C.-S., Motoyama, T., Osada, H., (2015).** Biosynthesis of the mycotoxin tenuazonic acid by a fungal NRPS–PKS hybrid enzyme. Nature Communications 6 (1), 1–9. doi : 10.1038/ncomms9758.

* **Zhang, J., et al., (2016).** Structural basis of nonribosomal peptide macrocyclization in fungi. Nature Chemical Biology 12 (12), 1001–1003. doi : 10.1038/nchembio.2202.

* **Ziemert, N., et al., (2012).** The natural product domain seeker NaPDoS: A phylogeny based bioinformatic tool to classify secondary metabolite gene diversity. PLoS One 7 (3), e34064. doi : 10.1371/journal.pone.0034064.

* **Ziemert, N., Alanjary, M., Weber, T., (2016).** The evolution of genome mining in microbes – A review. Natural Product Reports 33 (8), 988–1005. doi : 10.1039/c6np00025h.

* **Zutz, C., et al., (2013).** Small chemical chromatin effectors alter secondary metabolite production in *Aspergillus clavatus*. Toxins 5 (10), 1723–1741. doi : 10.3390/toxins5101723.

2

Relevant Websites

- **antiSMASH fungal version.**

  https://fungismash.secondarymetabolites.org

- **BiG-SCAPE.**

  https://git.wur.nl/medema-group/BiG-SCAPE

- **Conserved Domain Search Service (CD Search).**

  https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi

- **Fungal Gene Clustering.**

  https://fungiminions.shinyapps.io/FunGeneClusterS/

- **HMMER.**

  http://hmmer.org

- **InterPro**

  EMBL-EBI.

  http://www.ebi.ac.uk/interpro/

- **MIBiG: Minimum Information about a Biosynthetic Gene cluster.**

  https://mibig.secondarymetabolites.org/

- **MultiGeneBlast: Combined Blast search for multigene modules.**

  http://multigeneblast.sourceforge.net/

- **Mycotoxin Survey.**

  https://www.biomin.net/solutions/mycotoxin-survey/

- **NaPDoS.**

  http://napdos.ucsd.edu

- **Natural Products Atlas.**

  https://www.npatlas.org

- **Pfam**.

  https://pfam.xfam.org/

- **RoerdamAndersenLab.**

  https://github.com/RoerdamAndersenLab/gene_cluster_networks_and_genetic_dereplication

- **Secondary Metabolite Unique Regions Finder (SMURF).**

  https://smurf.jcvi.org/index.php

- **The Secondary Metabolite Bioinformatics Portal.**

  https://secondarymetabolites.org/

- **TIGRFAMS | J. Craig Venter Institute.**

  https://www.jcvi.org/tigrfams

3. Evolution-informed discovery of the naphthalenone biosynthetic pathway in fungi

Olga V. Mosunova[01]

Jorge C. Navarro-Muñoz[01]

Diksha Haksar[01]

Jacq van Neer[01]

Jelmer Hoeksma[02,03]

Jeroen den Hertog[03,04]

Jérôme Collemare[01]

[01] Westerdijk Fungal Biodiversity Institute, Utrecht, The Netherlands

[02] Hubrecht Institute-KNAW, Utrecht, The Netherlands

[03] University Medical Center Utrecht, Utrecht, The Netherlands

[04] Institute Biology Leiden, Leiden University, Leiden, The Netherlands

Abstract

Fungi produce a wide diversity of secondary metabolites with interesting biological activities for the health, industrial, and agricultural sectors. While fungal genomes have revealed an unexpectedly high number of biosynthetic pathways that far exceeds the number of known molecules, accessing and characterizing this hidden diversity remain highly challenging. Here, we applied a combined phylogenetic dereplication and comparative genomics strategy to explore eight lichenizing fungi. The determination of the evolutionary relationships of aromatic polyketide pathways resulted in the identification of an uncharacterized biosynthetic pathway that is conserved in distant fungal lineages. The heterologous expression of the homologue from *Aspergillus parvulus* linked this pathway to naphthalenone compounds, which were detected in cultures when the pathway was expressed. Our unbiased and rational strategy generated evolutionary knowledge that ultimately linked biosynthetic genes to naphthalenone polyketides. Applied to many more genomes, this approach can unlock the full exploitation of the fungal kingdom for molecule discovery.

Importance

Fungi have provided us with life-changing small bioactive molecules, with the best-known examples being the first broad-spectrum antibiotic penicillin, immunosuppressive cyclosporine, and cholesterol-lowering statins. Since the 1980s, exploration of chemical diversity in nature has been highly reduced. However, the genomic era has revealed that fungal genomes are concealing an unexpected and largely unexplored chemical diversity. So far, fungal genomes have been exploited to predict the production potential of bioactive compounds or to find genes that control the production of known molecules of interest. But accessing and characterizing the full fungal chemical diversity require rational and, thus, efficient strategies. Our approach is to first determine the evolutionary relationships of fungal biosynthetic pathways in order to identify those that are already characterized and those that show a different evolutionary origin. This knowledge allows prioritizing the choice of the pathway to functionally characterize in a second stage using synthetic biology tools like heterologous expression. A particular strength of this strategy is that it is always successful: it generates knowledge about the evolution of bioactive molecule biosynthesis in fungi, it either yields novel molecules or links the studied pathway to already known molecules, and it reveals the chemical diversity within a given pathway, all at once. The strategy is very powerful to avoid studying the same pathway again and can be used with any fungal genome. Functional characterization using heterologous expression is particularly suitable for fungi that are difficult to grow or not genetically tractable. Thanks to the decreasing cost of gene synthesis, ultimately, only the genome sequence is needed to identify novel pathways and characterize the molecules that they produce. Such an evolution-informed strategy allows the efficient exploitation of the chemical diversity hidden in fungal genomes and is very promising for molecule discovery.

Keywords

biosynthetic gene clusters, comparative genomics, phylogeny, heterologous expression, *Aspergillus parvulus*, *Aspergillus oryzae*, lichen, polyketide, nonreducing polyketide synthase, acetyl tetrahydroxynaphthalene

Fungi are an illustrious source of small bioactive compounds, called natural products or secondary metabolites (SMs), which are not strictly required for growth and reproduction. Instead, they serve as chemical mediators of interactions with the physical environment and with other organisms. At present, fungal SMs are experiencing a renewed interest in drug discovery after the high-throughput screening of libraries of synthetic compounds showed its limitation in yielding new bioactive molecules [01]. Fungi have provided us with key antimicrobial compounds, like the first broad-spectrum antibiotic, penicillin, and the first antifungal compound of the caspofungin class, demonstrating their life-changing potential [02, 03]. The genomic era has revealed that the fungal kingdom has been underexploited because fungal genomes encode an outstanding number of biosynthetic pathways that is far higher than the number of known fungal molecules [04].

The ever-increasing number of fungal genomes provides an exciting opportunity to identify novel bioactive molecules, but at the same time, such a vast amount of data represents a significant challenge to successful exploitation. So far, analyses of fungal genomes have mostly been restricted to searching the genes involved in the production of already known molecules [05-08] or to surveying the global biosynthetic potential of a given fungus or fungal lineage [09-11]. While providing interesting frameworks for functional analyses, such approaches do not allow the rational exploitation of fungal genomes for molecule discovery.

Phylogenetic studies have revealed that closely related SM biosynthetic enzymes tend to share similar catalytic activities and produce similar chemical backbones [12, 13]. Prioritizing functional studies using so-called phylogenetic dereplication to discover novel backbones have been rudimentarily employed only. Using this approach, Harvey and coworkers selected 41 biosynthetic pathways from diverse fungal species for functional studies, leading to the

detection of 22 compounds, including a few novel ones [14]. This previous report showed the potential of this approach to prioritize functional analyses for molecule discovery, but it did not make full use of the evolutionary information underlying phylogenetic relationships, and it did not make use of comparative genomics. Comparing evolutionarily related biosynthetic pathways is necessary to resolve complex metabolic patterns in distinct fungi. Thus, evolutioninformed analysis of fungal biosynthetic pathways is still an unexplored ground for molecule discovery and the full exploitation of fungal genomes [15].

Fungal SMs are highly structurally diverse and exhibit various biological activities, yet SM biosynthesis relies on a few classes of core enzymes, including polyketide synthases (PKSs), nonribosomal peptide synthetases (NRPSs), and terpene cyclases (TCs) [16]. In addition to the core enzyme, the biosynthesis of a given SM typically involves other socalled tailoring enzymes, which are encoded by genes that are usually organized into biosynthetic gene clusters (BGCs), meaning that they colocalize in the genome and are coregulated [16]. Within the fungal kingdom, the Ascomycota show the greatest potential, with 12 to 68 BGCs on average per fungal genome [17]. Within the Ascomycota, the Lecanoromycetes class is particularly interesting because it comprises species with the highest number of PKSs per genome. Most known polyketides from the Lecanoromycetes are structurally related and comprise an orsellinic acid-like backbone, as in depsides (atranorin, lecanoric acid, and sekikaic acid), depsidones (lobaric acid), and depsones (picrolichenic acid) [18]. Other known Lecanoromycetes polyketides exhibit chemical structures like anthra- and naphthoquinones (parietin, rhodocladonic acid, and cristazarin), dibenzofurans (usnic acid and pannaric acid), or chromones (lepraric acid) [19]. Despite many known compounds, to date, only atranorin and lecanoric acid have been functionally linked to their respective BGCs [20, 21]. A few compounds have been putatively assigned to BGCs based

on genomic and transcriptomic information [06, 21-23], but they remain to be functionally validated, and thus, the vast majority of lichen compounds remain unassigned to BGCs. In addition, genetic and chemoinformatic analyses of BGCs in Ascomycota fungi suggest that the Lecanoromycetes comprise a reservoir of fairly dissimilar biosynthetic pathways compared to other classes of the Ascomycota, and therefore, their genetic and chemical potential has remained far from characterized [09].

In this study, we present how an evolution-informed strategy that combines phylogenetic dereplication and comparative genomics can be used to prioritize BGC functional characterization, link BGCs to molecules, and, ultimately, fully exploit fungal chemical diversity. The phylogenetic dereplication of nonreducing polyketide synthases (nrPKSs) encoded in eight Lecanoromycetes genomes revealed a novel biosynthetic pathway in *Lobaria pulmonaria* and *Umbilicaria pustulata* lichen mycobionts. A comparative genomics approach identified a homologous pathway in *Aspergillus parvulus*, and its nrPKS was functionally characterized using heterologous expression in *Aspergillus oryzae*. Thanks to the determination of the evolutionary relationships between polyketide BGCs, we were able to predict chemical diversity in distinct fungal lineages, and we suggest that *L. pulmonaria* and *U. pustulata* may produce SMs unreported for these species, which might play a role during their interaction with their respective photobionts.

Results

Mining and phylogenetic dereplication of biosynthetic pathways encoded in Lecanoromycetes genomes identify an uncharacterized aromatic polyketide pathway.

Although Lecanoromycetes fungi have been extensively screened for bioactive molecules, most of these compounds belong to very few polyketide chemical classes, and the few published genome analyses indicated an unexpectedly high potential to produce diverse polyketides [9]. While this potential has been explored to link BGCs to known molecules, it has not been used to identify novel biosynthetic pathways. For this purpose, a maximum likelihood phylogenetic tree of 79 predicted nrPKSs retrieved from eight Lecanoromycetes genomes (*Cladonia grayi*, *Xanthoria parietina*, *Usnea florida*, *Lobaria pulmonaria*, *Acarospora strigata*, *Dibaeis baeomyces*, *Graphis scripta*, and *Umbilicaria pustulata*) was built together with 87 characterized nrPKSs from the Minimum Information about a Biosynthetic Gene Cluster (MIBiG) database [24] and from the literature [Fig. 01; see also Data Sets S1 to S3 in the supplemental material].

The tree is divided into 13 strongly supported clades, 8 of which correspond to previously defined nrPKS groups based on phylogeny and cyclization patterns [12, 21, 25]. However, our analyses indicate that groups I, II, and IV should each be split into two subgroups that are consistent with the precursors released from the nrPKSs [Fig. 01]. All nrPKS groups (IIa, IIb, III, IVa, IVb, and V) that produce polyketides with two or more aromatic rings share an origin, while other nrPKS groups that produce polyketides with a single aromatic ring seem to form three different clades (groups Ia and Ib; group VIII; and groups VI, VII, and IX). The most common Lecanoromycetes compounds, depsides, depsidones, depsones, and dibenzofurans, are thus likely produced by such single-ring-polyketideproducing nrPKSs. Consistently, the nrPKS involved in the production of atranorin belongs to group IX (21). Anthraquinones and xanthones like parietin and lichexanthone are likely produced by nrPKSs from group V because this clade comprises characterized nrPKSs involved in the biosynthesis of emodin-derived anthraquinones [Fig. 01].

Fig. 01.

Phylogenetic dereplication of nonreducing polyketide synthases (nrPKSs) from eight
Lecanoromycetes. A maximum likelihood phylogenetic tree was built with reference
characterized nrPKSs from the MIBiG database **(24)** and from the literature **(see Data Set S4 in
the supplemental material)**. Ultrafast bootstrap values of .95 are shown. The tree is midpoint
rooted. The nine phylogenetic groups as reported in previous analyses are indicated, with
backbone molecules produced by the characterized nrPKSs shown next to each clade. Two
new phylogenetic groups, X and XI, are highlighted in red. DMOA, 3,5-demethylorsellinic acid;
MOA, methylorsellinic acid; OA, orsellinic acid; T4HN, 1,3,6,8-tetrahydroxynaphthalene; AT4HN,
2-acetyl-1,3,6,8-tetrahydroxynaphthalene; YWA1, 2,5,6,8-tetrahydroxy-2-methyl-2,3-dihydro-
4H-naphthol[2,3-b]pyran-4-one.

In addition to these nine previously reported groups, two additional phylogenetic clades are strongly supported and are referred to as new phylogenetic groups X and XI [Fig. 01]. Group X appears basal to groups Ia and Ib and thus likely comprises enzymes that produce polyketides with a single aromatic ring. The nrPKS involved in the production of the depside lecanoric acid in *Pseudevernia furfuracea* belongs to this clade [20], confirming that nrPKSs in this group produce orsellinic acid derivatives. In contrast, group XI does not comprise any characterized nrPKS. The basal position to group V suggests that the polyketides released by nrPKSs from group XI contain several aromatic rings. This group contains only four nrPKSs from *L. pulmonaria*, *C. grayi*, and *U. pustulata*, but manual curation of gene models revealed that Clagr3_1822 is actually a pseudogene because it contains a disruptive mutation [Data Set S04]. A single nrPKS from *D. baeomyces* is not related to any group and forms an outgroup to groups II to V and XI. Based on this phylogenetic dereplication, the vast majority of nrPKSs in the Lecanoromycetes fall into groups for which the chemical backbone can be predicted, which will be useful to assign known molecules to BGCs. We then embarked on characterizing the new group XI of fungal nrPKSs as it may be involved in the production of new polyketides in lichenizing and other fungi.

Group XI nrPKSs belong to a novel conserved biosynthetic gene cluster.

Because the phylogenetic dereplication was performed with functionally characterized nrPKSs only, it was not known whether group XI nrPKSs are restricted to the Lecanoromycetes or are present in other distant fungal species. To answer this question, we sought close homologues of the four
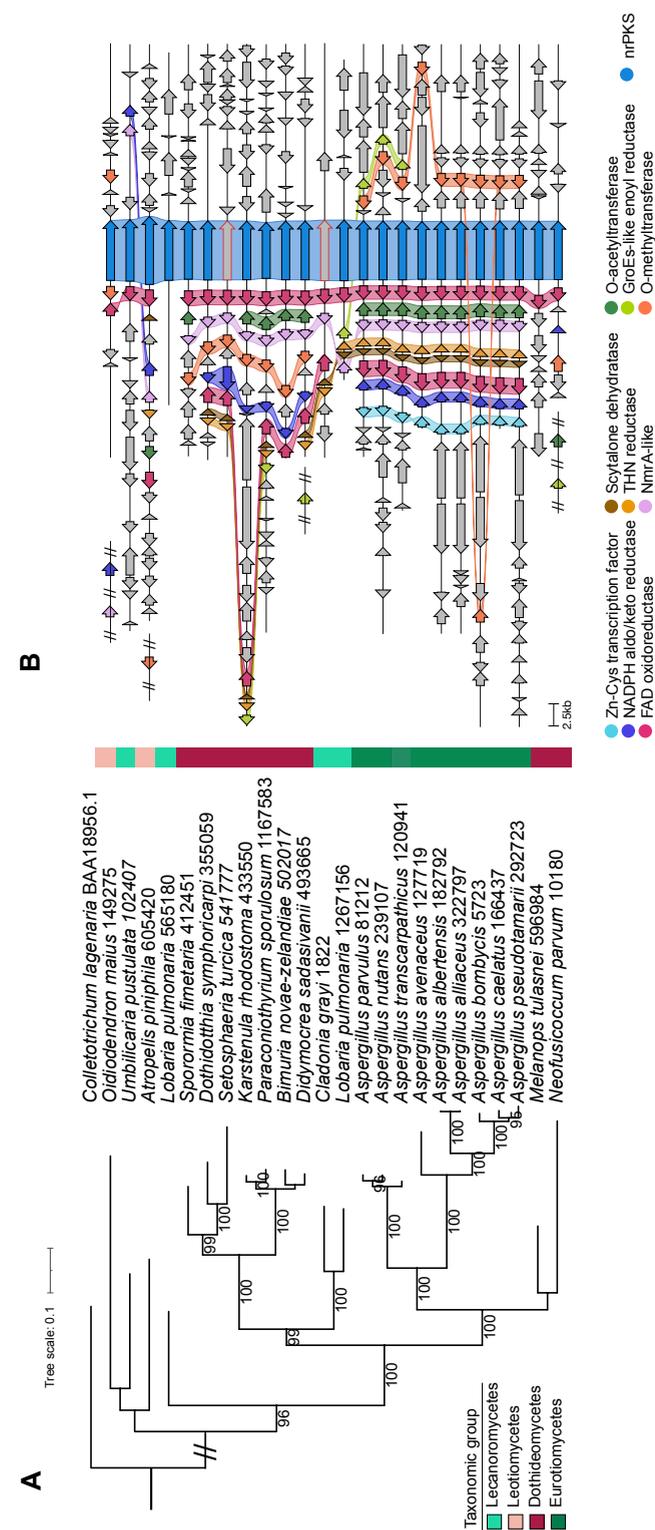
Fig. 02.

A novel biosynthetic gene cluster (BGC) from *Lobaria pulmonaria* and *Umbilicaria pustulata* is conserved in distant fungal species. (A) A maximum likelihood phylogenetic tree was built with homologues of group XI nonreducing polyketide synthases (nrPKSs), and the melanin nrPKS from *Colletotrichum lagenarium* was used as an outgroup. Ultrafast bootstrap values of .95 are shown. (B) Loci of nrPKS genes as predicted by fungiSMASH [42] were compared using Clinker [51]. Colored arrows and connections show conserved genes that are predicted to be part of the BGC. Orange nrPKS arrows correspond to genes with disruptive mutations (see Table S1 in the supplemental material). Homologous genes that were found at another genomic location are depicted with double slashes. THN, tetrahydroxynaphthalene.

group XI nrPKSs in available Ascomycota predicted proteomes. In total, we identified 20 other closely related nrPKSs, which expanded group XI to distant fungal lineages, including Leotiomycetes, Dothideomycetes, and Eurotiomycetes (Fig. 02A and Data Sets S01 to S03). The phylogeny of these homologues indicates that Lobpul1_1267156, Lobpul1_565180, and Umbpus1_102407 are actually paralogues (Fig. 02A). The Lobpul1_1267156 clade is further divided into two well-supported branches, both of which comprise Dothideomycetes sequences, also suggesting paralogy or horizontal transfer from Eurotiomycetes to Dothideomycetes (Fig. 02A).

We then compared the genomic loci of all Lobpul1_1267156 homologues in order to predict the borders of a putative conserved BGC. Genes that encode putative tailoring enzymes were identified, based on their functional conserved domains, at all loci but one (Fig. 02B and Table S01). Although the Lobpul1_565180 paralogue is predicted to be functional, this locus does not comprise any putative tailoring gene. A flavin adenine dinucleotide (FAD)-binding oxidoreductase tailoring gene is located upstream of the nrPKS gene in all species, and both genes appear to share a bidirectional promoter, except in *Oidiodendron maius*, in which an O-methyltransferase gene is inserted in between (Fig. 02B). A set of six other tailoring genes (encoding an O-acyltransferase, an O-methyltransferase, a tetrahydroxynaphthalene [T4HN] reductase, a scytalone dehydratase, an aldo-keto reductase, and a second FAD-binding oxidoreductase) and one putative regulatory gene encoding an NmrA-like protein are conserved in most of the fungal species and form a predicted BGC (Fig. 02B and Table S01). In addition, a gene encoding a GroES-like alcohol dehydrogenase enzyme is present at the locus in eight distant species, and a close homologue was found at a different locus in *Didymocrea sadasivanii* and *Neofusicoccum parvum* (Fig. 02B and Table S01). Similarly, close homologues of the aldo-keto reductase and NmrA-like protein-encoding

genes were found at another locus in *O. maius*. Close homologues of the O-methyltransferase and O-acyltransferase genes are also found at a different locus in *Atropellis piniphila* and *N. parvum*, respectively. In *Aspergillus* species exclusively, a gene encoding a putative transcription factor is also found at the locus. Disruptive mutations were detected in the nrPKS gene not only in *C. grayi* but also in *Setosphaeria turcica*, suggesting a nonfunctional pathway in both species (Fig. 02B, Table S01, and Data Set S04). Similarly, disruptive mutations were found in a few tailoring genes, including the T4HN reductase and scytalone dehydratase genes in *L. pulmonaria*, and the insertion of a long sequence in an intron of the keto reductase gene in *S. turcica* likely makes it nonfunctional (Table S01 and Data Set S04). Although the phylogeny suggests a complex evolutionary history with several paralogues, the comparative genomics analysis indicates that nrPKS group XI belongs to a conserved BGC with little diversification between fungal species.

Ancestral duplication of T4HN reductase and scytalone dehydratase genes.

Two tailoring genes from the predicted group XI BGC encode a T4HN reductase and a scytalone dehydratase. These two enzymes are well characterized and act together to dehydroxylate intermediates in the conserved dihydroxynaphthalene (DHN) melanin [26] and anthraquinone cladofulvin [27] biosynthetic pathways. Especially, both genes in group XI appear to share a bidirectional promoter (Fig. 02B), a gene organization also found in the *Aspergillus fumigatus* DHN (ARP1 and ARP2) [28, 29] and *Cladosporium fulvum* cladofulvin [27] BGCs, suggesting a common origin. A previous report indicated that *ARP1* and *ARP2* are distant paralogues of *claB*

A

B
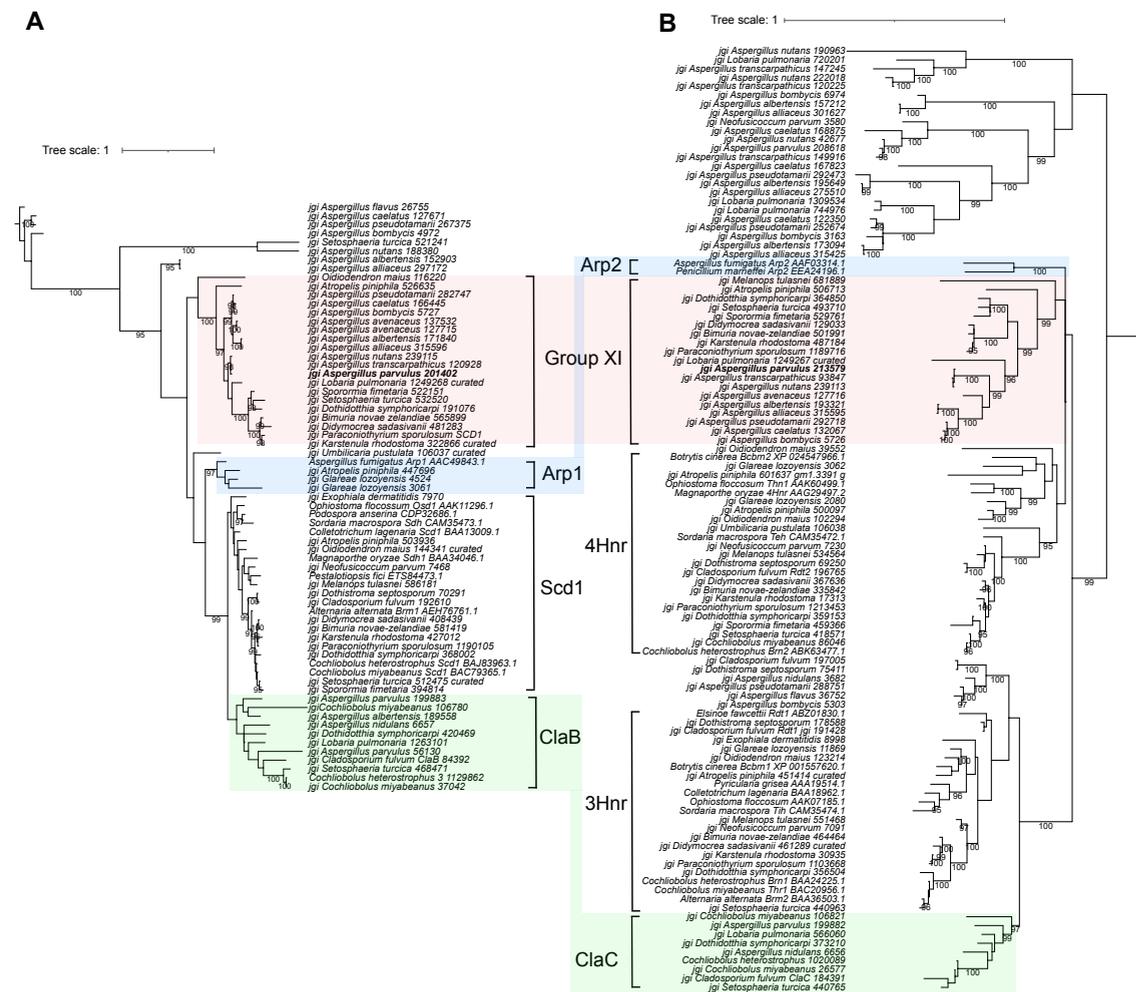
Fig. 03.

Coevolution between hydroxynaphthalene reductases (Hnr) and scytalone dehydratases. Maximum likelihood phylogenetic trees of homologues of scytalone dehydratases (A) and tetra- and tri-Hnr proteins (B) are shown [52]. Paralogues that share a bidirectional promoter in the group XI, DHN melanin, and cladofulvin biosynthetic gene clusters (BGCs) are highlighted. Ultrafast bootstrap values of >95 are shown.
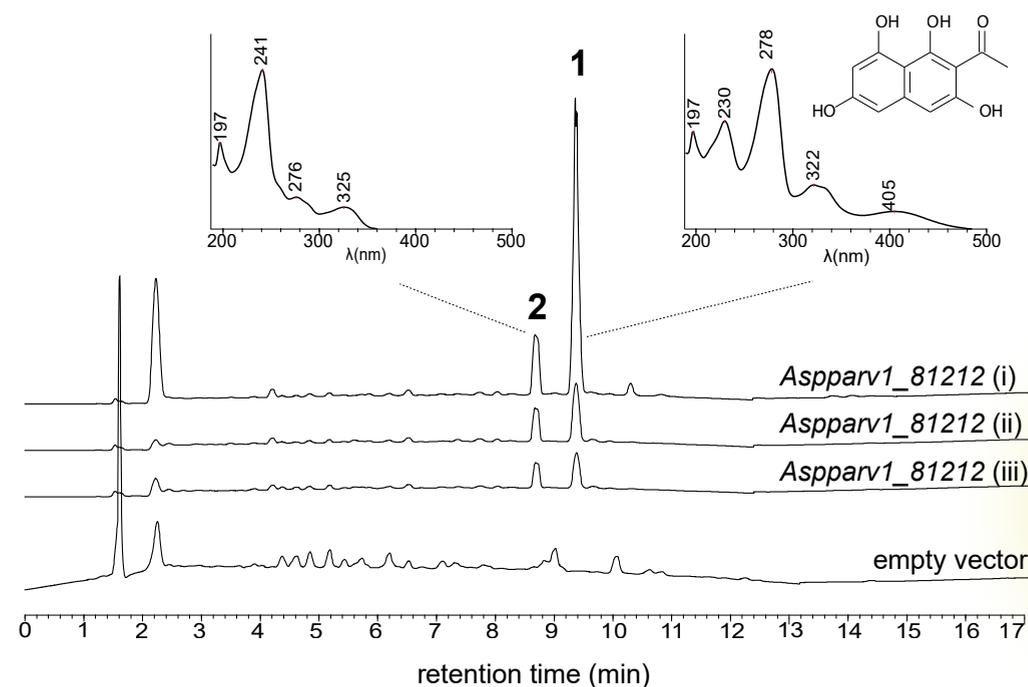
Fig. 04.

Heterologous expression of the group XI nonreducing polyketide synthase from *Aspergillus parvulus*. Organic extracts of 4-day-old *Aspergillus oryzae* NSAR1 transformants carrying an empty vector or expressing *Aspparv1_81212* (three independent transformants) were analyzed using UV-HPLC. Major product **1** was identified as acetyl-tetrahydroxynaphthalene (AT4HN). Minor product **2** was identified as 6,8-dihydroxy-3-methylisocoumarin.

and *claC* that have been recruited in different biosynthetic pathways [30]. Phylogenetic analyses of these enzymes, including the group XI sequences, are consistent with these previous findings [Fig. 03 and Data Sets S01 to S03] and reveal that group XI T4HN reductase and scytalone dehydratase are distinct paralogues with a similar evolutionary history. While *claB* and *claC* are most closely related to *SCD1* and *3HNR*, respectively, from the DHN melanin pathway found in Dothideomycetes and Sordariomycetes, the group XI gene pair could originate from an ancestral duplication of the *ARP1-ARP2* gene pair [Fig. 03].

Aspparv1_81212 is a hexaketide synthase that produces AT4HN.

Although many biosynthetic pathways from other fungi have been successfully characterized in *Aspergillus oryzae* [31, 32], no PKS from the Lecanoromycetes has so far been successfully expressed in this heterologous host for unknown reasons [33]. Because the Lobpul1_1267156 BGC is well conserved [Fig. 02], we chose to functionally characterize in *A. oryzae* the homologous nrPKS from *A. parvulus*, *Aspparv1_81212*, as it belongs to a complete predicted BGC [Fig. 02B]. *A. oryzae* transformants expressing *Aspparv1_81212* produced two novel compounds compared to transformants carrying an empty vector [Fig. 04 and Fig. S01]. Product **1** (retention time [RT] = 9.35 min; maximum UV absorption [UV max] = 197, 230, 278, 322, and 405 nm; m/z = 233 [M −H]−) exhibits a yellow colour and has an exact mass of 235.0577 [M+H]+ as determined by high-resolution mass spectrometry (HRMS) [Fig. S02]. Product **1** was identified as 2-acetyl-1,3,6,8-tetrahydroxynaphthalene (AT4HN) [Fig. 04] using nuclear magnetic resonance (NMR) [Data Set S05] and by comparing mass and UV spectra to previously published data [34]. Product **2** (RT = 8.65 min; UV max = 197, 241, 276, and 325 nm; m/z = 191 [M−H]−) was identified as the pyrone 6,8-dihydroxy-3-methylisocoumarin based on NMR analyses [Data Set S05] and published data [35]. Such pyrones are known shunt metabolites from nrPKSs when the final thioesterase (TE) domain is inactivated [34]. However, such shunt pyrones harbour the same length as that of the polyketide released by the fully functional nrPKS [34]. While product **1** is a hexaketide released through TE catalyzed Claisen cyclization, product **2** is a pentaketide that seems to be released after the incorporation of four malonyl-CoAs through spontaneous O-C cyclization [34]. Analysis of the transcripts in *A. oryzae* transformants showed that a single mature nrPKS is expressed [Fig. S01],

Table 01.

Functions of proteins encoded in the predicted group XI biosynthetic gene cluster and flanking genes in *Aspergillus parvulus*

Protein ID	Gene name	Protein function	Pfam domain, E value
171943		Protein kinase	PF00069.27, 4.4e−49 PF07714.19, 2.3e−23
171942		Unknown	No hit
171941		Mitochondrial carrier protein	PF00153.29, 2.9e−07
201406		Cupin superfamily protein	PF06172.13, 2.0e−51
201405	APR9	Zn/Cys transcription factor	PF00172, 7.37e−09
171936	APR8	Aldo-keto reductase	PF00248, 6.40e−144
201403	APR7	FAD-binding oxidoreductase	PF01565, 7.49e−19 PF08031, 2.42e−11
201402	APR5	Scytalone dehydratase	PF02982, 6.74e−57
213579	APR6	T4HN reductase	PF13561, 1.26e−91
193045		NmrA-like protein	PF05368, 4.70e−91
171934	APR3	O-Acyltransferase	PF13813, 6.04e−20
81217	APR2	FAD-binding oxidoreductase	PF01565, 4.34e−29 PF08031, 4.12e−06
81212	APR1	Nonreducing polyketide synthase	SAT (PF16073), 1.04e−57 KS (PF00109), 1.07e−82 KS_C (PF02801), 8.67e−41 AcT (PF00698), 3.56e−38 PT (TIGR04532), 8.65e−13 PP-b (PF00550), 21.89e−07 PP-b (PF00550), 2.06e−10 TE (PF00975), 3.61e−23
171930		Thioredoxin	PF00085.22, 3.8e219
193042	APR4	S-Adenosylmethioninedependent O-methyltransferase	PF00891, 3.87e220
81192		GroES-like alcohol dehydrogenase	PF08240, 4.76e−07 PF00107, 6.14e−06
171927		WD domain-containing protein	PF11816.10, 6.4e−59 PF00400.3, 1.0e−05
81165		PP loop family protein	PF01171.22, 1.6e−48
81156		RhoGEF protein	PF00621.22, 4.1e−25 PF12015.10, 1.7e−11

meaning that this nrPKS can release two polyketides of different lengths. The three transformants yielded AT4HN as the major product (83, 56, and 350 mg/L) compared to the pyrone (7, 13, and 219 mg/mL). The yield difference between the transformants most likely reflects differences in gene expression due to the integration of the plasmid at different genomic loci. These results demonstrate that the *Aspparv1_81212* homologue encodes a PKS that releases the hexaketide AT4HN as the first stable intermediate. Thus, group XI PKSs produce the same backbone as those of certain PKSs in group II **(Fig. 01)**.

The group XI biosynthetic gene cluster is linked to the production of 6-O-methylasparvenone and ethylparvulenone in *Aspergillus parvulus*.

A. parvulus is known to produce several polyketides that could structurally derive from compound **1**, namely, the naphthalenone asparvenone compound **3**, parvulenone compound **5**, and methylated or ethylated derivatives **4**, **6**, and **7** [36] **(Fig. 05A)**. Analysis of organic extracts from *A. parvulus* grown under conditions conducive to naphthalenone production [37] detected the presence of compound **4** (RT = 15.25 min; m/z = 237 [M +H]+; UV max = 219 and 288 nm) and compound **7** (RT = 12.6 min; m/z = 249 [M−H]−; UV max = 216, 262, and 308 nm) **(Fig. 05B)**. Product **4** was confirmed to be 6-O-methylasparvenone by NMR **(Data Set S05)**.

We then assessed if the predicted *Aspparv1_81212* BGC is expressed when compounds **4** and **7** are produced. With the exception of the predicted and not conserved thioreductase *Aspparv1_171930* gene, all genes predicted to be part of the BGC were found to be significantly expressed compared to the housekeeping gene *H2B* **(Fig. 05C)**. The *H2B* transcript exhibit similar
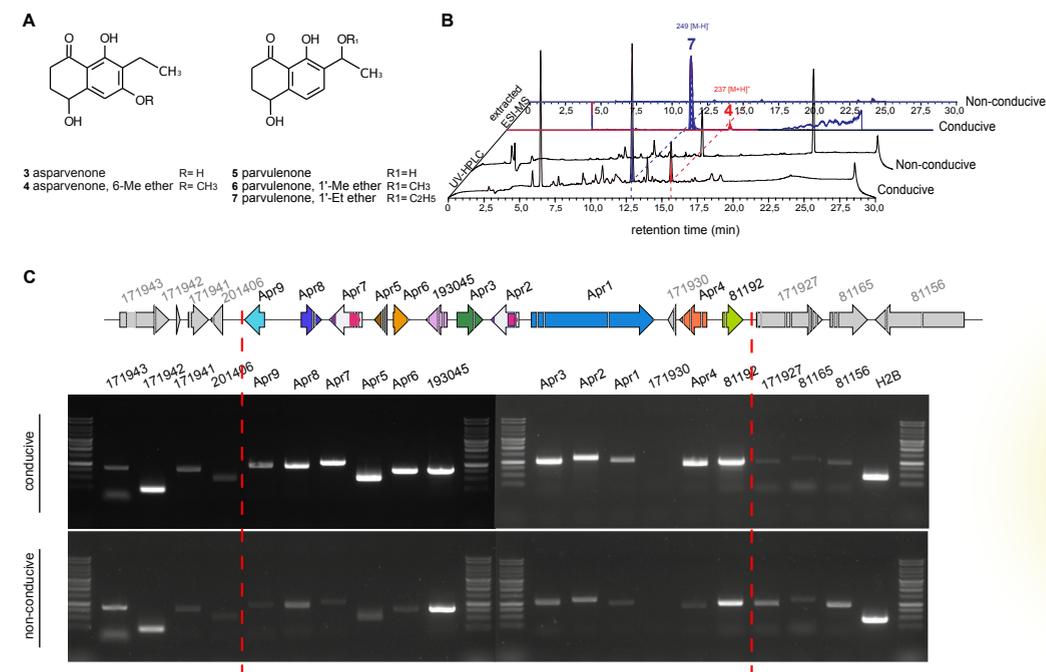
Fig. 05.
Correlation between the expression of the group XI biosynthetic gene cluster and the production of naphthalenone compounds in *Aspergillus parvulus*. (A) Chemical structures of naphthalenone compounds reported for *A. parvulus*. (B) UV-HPLC traces of organic extracts obtained from a 6-day-old *A. parvulus* culture in malt extract (conducive) or Czapek-Dox (nonconducive) liquid medium. Analysis of UV spectra and extraction of mass ions (electrospray ionization-mass spectrometry [ESI-MS]) of naphthalenone compounds shown in panel A identified products 4 and 7, provided as representative of results from three replicates. (C) Expression of predicted genes at the *Aspparv1_81212* locus in *A. parvulus* grown under conducive and nonconducive conditions as determined by RT-PCR. The *H2B* housekeeping gene was used as an expression control.

signal intensities in both cultures. In comparison to *H2B* signals, a condition non-conducive to naphthalenone production shows the very limited expression of most genes apart from the genes encoding an NmrA-like protein (*Aspparv1_193045*) and a GroES-like alcohol dehydrogenase (*Aspparv1_81192*),

which exhibit similar or slightly lower expression levels. These results make it uncertain whether these two genes belong to the BGC or not. The flanking genes *Aspparv1_171941* and *Aspparv1_201406* seem to be expressed at slightly higher levels under conducive conditions, but they encode a mitochondrial carrier protein and a cupin protein, respectively [Table 01], which are unlikely to be involved in naphthalenone biosynthesis. The other flanking and housekeeping genes do not show coregulation [Fig. 05C]. These results show that most of the genes from the predicted BGC in *A. parvulus* are coregulated and that their expression correlates with the production of naphthalenones. Thus, the predicted BGC comprises nine genes likely involved in the production of asparvenone and derivatives, and the corresponding genes were named *APR1* to *APR9* [Table 1].

Discussion

Evolutionary relationships of biosynthetic gene clusters inform fungal chemical diversity.

The phylogenetic dereplication of nrPKSs encoded in the genomes of eight lichenizing fungi has revealed two new phylogenetic groups, of which group XI was not functionally characterized. Comparative genomics allowed the prediction of a BGC that was found to be expressed in *A. parvulus* when naphthalenone compounds were produced. These findings demonstrate that the combination of phylogenetic dereplication and comparative genomics is a powerful strategy that not only provides new insights into the evolution of fungal BGCs but also allows relevant prioritization of functional characterization without a *priori* knowledge about chemical structures.

Although chemical backbone **1** produced by Apr1 is already known [38], this result was unexpected because group XI is a sister clade of the anthraquinone-producing group V nrPKSs and is not related to group IIa, which comprises nrPKSs that are known to produce compound **1** [Fig. 01]. Similar to orsellinic acid that is produced by nrPKSs from distantly related clades [Fig. 01], compound **1** appears to be a common chemical backbone produced by enzymes that have diverged long ago.

Compound **1** is a precursor of DHN melanin in certain species like *Exophiala dermatitidis*, in which it is further converted to T4HN via the removal of the acetyl group by the polyketide-shortening enzyme WdYg1p (38). A similar reaction is catalyzed by the homologue Ayg1 to convert YWA1 into T4HN in *A. fumigatus* [39]. No homologue of WdYg1p could be identified in the *A. parvulus* genome, indicating that compound **1** could not be converted to T4HN in this fungus. Similarly, several homologues of T4HN reductase and

scytalone dehydratase were found encoded in the *A. parvulus* genome, but none of them correspond to paralogues involved in the DHN melanin pathway **(Fig. 03)**. Similarly, other *Aspergillus* species included in this study lack these paralogues **(Fig. 03)**, suggesting that the DHN melanin pathway has been lost in all of them. In contrast, Dothideomycetes and Leotiomycetes species carry both DHN melanin and naphthalenone BGCs, and a few appear to also contain a BGC related to the cladofulvin one. Noteworthy, in *O. maius*, the closest homologues of both the hydroxynaphthalene reductase and scytalone dehydratase genes form an outgroup to the group XI clades **(Fig. 03; see also Table S1 in the supplemental material)** and are located at another locus next to genes encoding an nrPKS (*Oidma1_51005*) and a cytochrome P450 monooxygenase (*Oidma1_157855*). Consistent with the phylogenetic position, these paralogues correspond to a different biosynthetic pathway in this species. Altogether, the phylogenetic analyses and genomic organizations suggest that the group XI BGC has an ancestral origin, possibly with the recruitment of paralogues from the DHN or cladofulvin pathways after gene duplication and of other tailoring genes that have resulted in the BGC for naphthalenone production in fungi.

Gene content diversification of the group XI biosynthetic gene cluster correlates with the chemical diversity of naphthalenone compounds.

Our results suggest that product 1 is the initial chemical backbone to produce products **4** and **7**, which is consistent with the previous proposition that a hexaketide precursor is the starting molecule of product **4** [36]. We also found that the pentaketide pyrone product **2** is released at the same time
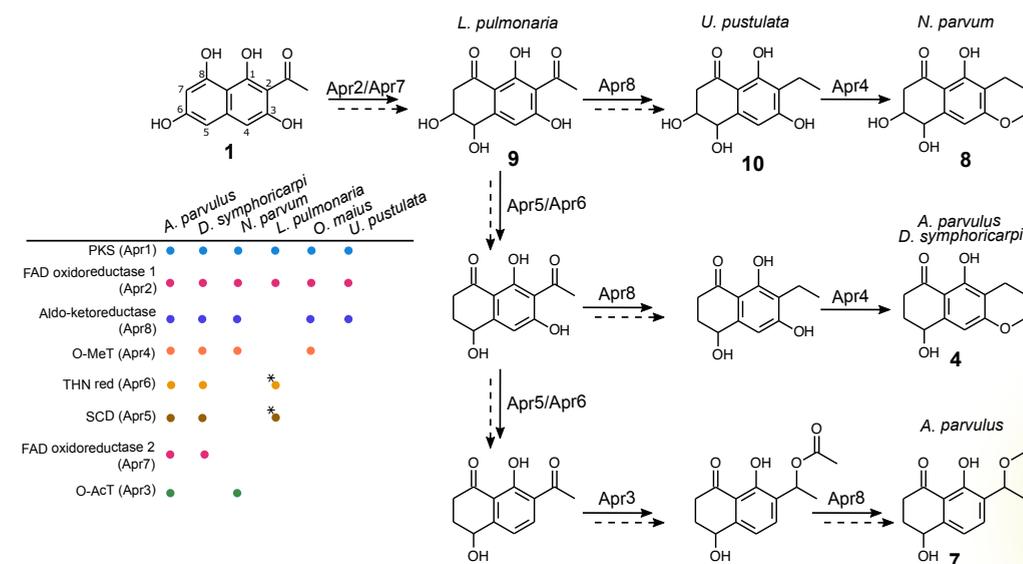
Fig. 06.

Prediction of enzymatic requirements for the production of diverse naphthalenones in fungi. *Aspergillus parvulus* produces compounds **4** and **7**, while *Neofusicoccum parvum* produces compound **8**, consistent with the gene component of the group XI biosynthetic gene cluster (BGC). Based on the BGC content, Apr2 and Apr7 are likely responsible for the hydroxylation at C-5. Apr5 and Apr6 are expected to jointly remove the hydroxyl groups at C-3 and C-6. Apr8 is predicted to reduce and Apr3 is predicted to acetylate the keto group, respectively. The combination of Apr3 and Apr8 is expected to be responsible for yielding compound **7**. Apr4 is predicted to be responsible for the methylation of the oxygen at C-3. Dashed arrows indicate that several steps are most likely required between hypothetical intermediates, either spontaneously or through the action of enzymes that remain to be identified. The predicted BGC in *Lobaria pulmonaria* suggests that this species produces the hypothetical compound **9**, while *Umbilicaria pustulata* and *Oidiodendron maius* are predicted to produce derivatives of the hypothetical compound 10. Genes with identified disruptive mutations are labeled with a star. O-MeT, O-methyltransferase; red, reductase; OAcT, O-acetyltransferase; SCD, scytalone dehydratase.

as product **1**. Traces of product **2** were detected in organic extracts of *A. parvulus* **(Fig. S03)**, which indicates that the release of product **2** is likely not a misfunction of the nrPKS in *A. oryzae*. However, the release of product **2** in

A. parvulus needs to be confirmed to validate this hypothesis. In addition, the predicted catalytic activities of tailoring enzymes encoded at the *APR1* locus are consistent with necessary modifications to convert product **1** into product **4** or **7**. DHN biosynthesis in fungi involves the removal of two hydroxyl groups through the sequential action of a hydroxynaphthalene reductase (4Hnr/Arp2 and 3Hnr) and a scytalone dehydratase (Scd1/Arp1) [26]. Similarly, ClaB and ClaC in *C. fulvum* are paralogues of 3Hnr and Scd1, which catalyze the removal of a hydroxyl group from emodin to yield chrysophanol hydroquinone in the cladofulvin pathway [27]. In *A. parvulus*, the removal of the C-3 and C-6 hydroxyl groups, as found in products **4** and **7** compared to product **1**, likely involves Apr5 and Apr6, the paralogues of Arp1 and Arp2, respectively [Fig. 03 and 06]. The presence of the predicted Omethyltransferase (Apr4) and O-acyltransferase (Apr3) encoded in the BGC is consistent with the methylation and ethylation observed in products **4** and **7**, respectively [Fig. 06]. A particular feature of product **4** is the reduction of the ketone, which could involve Apr8, a predicted aldo-keto reductase [Fig. 06], and possibly another enzyme to remove the resulting hydroxyl group. In a similar reaction to produce product **4**, the acetyl group added by Apr3 is likely reduced by Apr8, followed by the removal of the hydroxyl group to yield the ethyl group. The oxidoreductases Apr2 and Apr7 could be involved in the complete reduction of the aromatic ring and/or the selective oxidation of C-5 [Fig. 06]. Because the BGC is fully conserved in all *Aspergillus* species in which an APR1 orthologue was found as well as in *Atropellis piniphila*, *Sporormia fimetaria*, *Dothidotthia symphoricarpi*, *Karstenula rhodostoma*, *Paraconiothyrium sporulosum*, *Bimuria novae-zelandiae*, and *D. sadasivanii* [Fig. 02B and Table S01], we expect all these species to produce naphthalenone molecules related to products **4** and **7**. The absence of *APR3* in *D. symphoricarpi* indicates that derivatives of product **4** may be produced only [Fig. 02B and Table S01].

The fungus *N. parvum* is known to produce diverse naphthalenones, including botryosphaerone D compound **8** [40], which differs from compound **4** only by the presence of the C-6 hydroxyl group [Fig. 06]. The group XI BGC in *N. parvum* comprises five genes only [Fig. 02], including *APR8* and *APR4*, which encode the predicted aldo-keto reductase and O-methyltransferase, consistent with the chemical structure of compound **8**. The presence of the O-acyltransferase Apr3, although encoded at a different locus in the genome [Fig. 02B], suggests that *N. parvum* could also produce a molecule related to compound **7**. Orthologues of *APR5* and *APR6* could not be identified in the genome of *N. parvum*, which likely explains why both the C-3 and C-8 hydroxyl groups remain in compound **8** [Fig. 03 and Table S01]. Because the *N. parvum* genome also lacks the oxidoreductase gene *APR7* [Table S01], this suggests that both Apr2 and Apr7 may be functionally redundant. Further functional validation of these tailoring enzymes is being performed to address such biosynthetic hypotheses.

The BGC composition in *L. pulmonaria* and *Melanops tulasnei* is limited to *APR1* and *APR2* only, as in *L. pulmonaria*, both the *APR5* and *APR6* genes contain disruptive mutations that make the proteins likely not functional [Table S01 and Data Sets S01 to S03]. Thus, these two species could produce compounds related to compound **9** [Fig. 06]. In addition to *APR1* and *APR2*, *U. pustulata* and *O. maius* also comprise the aldo-keto reductase *APR8* gene [Fig. 02B and Table S01], suggesting that they might produce compounds related to compound **10** [Fig. 06]. The presence of the O-methyltransferase *APR4* gene in *O. maius* suggests that a methylated derivative of compound **10** is produced [Fig. 06]. Assigning functions of tailoring genes to biosynthetic steps of the naphthalenone pathway in *A. parvulus* will allow the reconstruction of pathways encoded in the genomes of other fungi, including the lichen mycobionts, and validate these hypotheses.

Conclusions.

Combining phylogenetic dereplication and comparative genomics is a powerful strategy to prioritize the functional characterization of BGCs without any a priori knowledge other than evolutionary novelty. While most fungal genomes are used to either list their production potential or find BGCs for known molecules of interest, our approach generates knowledge on the evolution of fungal BGCs, potentially finds novel molecules, and otherwise links genes to already known molecules, as found here. This strategy is very promising to unlock the full rational exploitation of fungal genomes for BGC characterization and molecule discovery.

Materials and methods

- **Fungal genomes and sequences.** Genome assemblies and gene predictions of *Cladonia grayi* Cgr/ DA2myc/ss v2.0, *Xanthoria parietina* 46-1-SA22 v1.1, *Usnea florida* ATCC 18376 v1.0, Lobaria pulmonaria Scotland reference genome v1.0, *Acarospora strigata* CBS 132363, *Dibaeis baeomyces*, *Graphis scripta* CBS 132367, and *Umbilicaria pustulata* were retrieved from the Joint Genome Institute (JGI) MycoCosm repository [41]. BGCs were predicted using fungiSMASH 4 with default parameters [42]. Predicted nrPKSs were selected based on the presence of the signature SAT starter unit:ACP transacylase (PF16073) and PT product template (TIGR04532) conserved domains. Gene models for homologues of *APR1*, *APR5*, *APR6*, and *APR8* were curated manually **(see Data Set S4 in the supplemental material)**. Characterized nrPKSs were retrieved from the Minimum Information about a Biosynthetic Gene Cluster (MIBiG) database [24] and from the literature **(Data Sets S01 to S03)**. Close homologues of *A. parvulus* Apr1 were retrieved from the MycoCosm repository using BLASTp. Another BLASTp search with each *A. parvulus* protein predicted in the pathway was performed on filtered proteins of each species containing the group XI BGC, allowing the identification of genes located at a different locus **(Table S01)**.

- **Phylogenetic trees.** Protein alignments were performed using Clustal Omega v1.2.4 [43], with the KS keto-synthase domain PF00195.19 as a guide (parameters –hmm –in) in the case of nrPKS sequences. Poorly aligned regions were removed using trimaL 1.4.rev15 (build 2013-12-17; parameter -automated1) [44]. Maximum likelihood trees were built with IQ-TREE v1.6.11-he860b03_0 bioconda [45] with model finder [30] and ultrafast bootstrapping as well as an approximate Bayes test [46] and a Shimodaira-Hasegawa approximate likelihood-ratio test [47] (parameters –bb 1000 –nt AUTO -mset LG -alrt 1000 -abayes -m MFP). The resulting trees were visualized using iTOL [48]. All curated alignments and phylogenetic tree files are provided in Data Sets S1 to S3.

- **Fungal strains and growth conditions.** *A. parvulus* CBS 136.61 from the CBS collection (Westerdijk Fungal Biodiversity Institute, The Netherlands) was grown for 6 days in malt

extract broth (MB) (filtered malt extract at 400 mL/L [pH 7.0]) liquid medium at 25°C under constant agitation at 200 rpm or on malt extract agar (MEA) (50 g/L [pH 5.4]; Oxoid) plates. For total RNA isolation and secondary-metabolite extraction, *A. parvulus* was grown for 6 days in 50 mL Difco Czapek-Dox (CZD) broth (BD, Franklin Lakes, NJ) or MB (pH 3.5) liquid medium at 25°C under constant agitation at 200 rpm. *A. oryzae* NSAR1 was grown on MEA plates for 5 days at 30°C. *A. oryzae* transformants were grown on selective CZD medium without arginine (35 g/L Difco CZD broth [BD, Franklin Lakes, NJ], 1 g/L ammonium sulfate [Sigma-Aldrich, St. Louis, MO], 0.5 g/L adenine [Sigma-Aldrich, St. Louis, MO], 1.5 g/L methionine [Sigma-Aldrich, St. Louis, MO]). For the induction of the PamyB promoter and polyketide production [49], transformants were grown at 30°C for 5 days in Yeast Malt Agar (YMA) (3 g/L Difco yeast extract, 3 g/L Difco malt extract, 5 g/L Difco Bacto peptone, 10 g/L glucose [Merck, Kenilworth, NJ]) liquid or agar (Ferwo 700 agar) medium. The mycelium from transformants grown in liquid YMA cultures was used to extract total RNA.

Nucleic acid extraction and RT-PCR. The mycelium of *A. parvulus* or *A. oryzae* from liquid cultures was filtered through a paper filter, frozen in liquid nitrogen, and ground using a mortar and pestle. Genomic DNA was isolated using the DNeasy plant minikit (Qiagen, Hilden, Germany) according to the manufacturer's recommendations. For total RNA extraction, 100 mg of the ground mycelium was mixed with 1 mL of Invitrogen TRIzol reagent (Thermo Fisher Scientific, Waltham, MA) in a 1.5-mL microcentrifuge tube and incubated for 5 min at 25°C. The resulting lysate was mixed with 0.2 mL chloroform, gently mixed by hand, and incubated for 5 min. Samples were centrifuged at 12,000 *g* for 15 min at room temperature. The aqueous phase was transferred into a new microcentrifuge tube, mixed with 0.5 volumes of 100% ethanol, and loaded into a column from the NucleoSpin RNA extraction kit (Macherey-Nagel, Allentown, PA). Downstream steps were performed according to the manufacturer's protocol. Five hundred nanograms of total RNA was used to synthesize cDNA using oligo(dT) primers and GoScript reverse transcription (RT) mix (Promega, Madison, WI) according to the manufacturer's protocol. PCR was performed for each *A. parvulus* gene at the APR1 locus and the housekeeping control gene *H2B* [Table S02] using GoTaq DNA polymerase

(Promega, Madison, WI). To confirm *APR1* expression in *A. oryzae* transformants, primers specific for the *A. parvulus* APR1 and *A. oryzae H2B* genes [Table S02] were used with GoTaq DNA polymerase (Promega, Madison, WI).

Gene amplification and plasmid digestion. The five exons of *APR1* (JGI protein identifier 81212) were amplified from genomic DNA of *A. parvulus* using primers that harbor 15-bp sequences homologous to the previous and next exons [Table S02], with the exception of the forward and reverse primers used to amplify the first and last exons, respectively, which harbor 30-bp sequences homologous to the pEYA2 plasmid [49] [Table S02]. The last exon was amplified to include a 100-bp downstream terminator sequence. All PCR fragments were amplified using Phusion high-fidelity DNA polymerase (Thermo Fisher Scientific, Waltham, MA) according to the manufacturer's protocol. One microgram of the pEYA2 plasmid was digested overnight with 10 U NotI at 37°C (Promega, Madison, WI). Fragments of the expected size and the linearized plasmid were purified from a 0.8% agarose gel or directly from the PCR mix using a Geneclean II kit (MP Biomedicals, Santa Ana, CA).

Transformation-associated recombination in Saccharomyces cerevisiae. A strain of *Saccharomyces cerevisiae* BMA 64 with a ura32 auxotrophic marker was used for transformationassociated recombination according to a protocol adapted from the one described previously [50]. *S. cerevisiae* was grown overnight at 30°C in 3 mL yeast extract-peptone-dextrose (YPD) medium (20 g/L D-glucose [Sigma-Aldrich, St. Louis, MO], 20 g/L Difco peptone, 10 g/L Difco yeast extract). Two milliliters containing 108 cells was transferred into 50 mL YPD medium and incubated at 30°C under agitation at 200 rpm for about 5 h until reaching an optical density at 600 nm (OD600) of 1 to 1.5. Yeast cells were centrifuged for 5 min at 2,500 rpm at 4°C. The cells were resuspended in 20 mL of a filter-sterilized lithium acetate (LiAc)-dithiothreitol (DTT) solution (100 mM LiAc, 10 mM DTT, 0.6 M sorbitol, 10 mM Tris-HCl [pH 7.5]) and incubated at room temperature for 30 min under agitation at 100 rpm. The cells were centrifuged for 5 min at 2,500 rpm at 4°C, and the supernatant was discarded. Cells were washed in 7 mL of ice-cold 1 M sorbitol and centrifuged at 2,500 rpm three times. Finally, the cells were resuspended in 400 mL

of ice-cold 1 M sorbitol. Eight microliters of the cell suspension was mixed with 6 mL of the NotIlinearized pEYA2 plasmid and 2 mL of each APR1 exon fragment in a prechilled electroporation cuvette. After a 5-min incubation on ice, cells were electroporated (1,500 V and 200 Ω) and mixed immediately with 1 mL of ice-cold 1 M sorbitol. Cells were incubated for 1.5 h at 30°C without agitation and then centrifuged, and 200 mL was plated onto synthetic dropout agar medium (SDM) (20 g/L agar, 20 g/L D-glucose [Sigma-Aldrich, St. Louis, MO], 1.92 g/L yeast dropout supplements without uracil [Sigma-Aldrich, St. Louis, MO], 6.7 g/L yeast nitrogen base without amino acids [Sigma-Aldrich, St. Louis, MO]). Plates were incubated for 3 to 7 days at 30°C. Yeast transformants were transferred to a new selective plate and grown overnight. Single colonies were transferred into a microcentrifuge tube in 30 mL of 25 mM NaOH and boiled for 10 min at 100°C. Next, 1 mL was used for PCR screening with GoTaq DNA polymerase (Promega, Madison, WI) and primers APR1_F and APR1_R [Table S02]. Positive transformants were grown overnight in liquid SDM to isolate the pEYA2::APR1 plasmid using the Zymoprep yeast plasmid miniprep kit (Zymo Research, Irvine, CA), and the obtained plasmid was subsequently introduced into electrocompetent *Escherichia coli* DH5α cells (Thermo Fisher Scientific, Waltham, MA) using an electroporation method according to the manufacturer's protocol. PCR screening was performed by transferring individual colonies into the PCR mixture with GoTaq DNA polymerase (Promega, Madison, WI). The pEYA2:: *APR1* plasmid was isolated from confirmed positive clones using the Zyppy plasmid miniprep kit (Zymo Research, Irvine, CA), and the plasmid was validated by sequencing (Macrogen, Seoul, South Korea).

Construction of the expression vector. Seventy nanograms of the pEYA2::*APR1* entry vector and 100 ng of the pTYGSarg destination vector [49] were mixed with 1 mL of the Gateway LR Clonase II enzyme (Thermo Fisher Scientific, Waltham, MA) in a 5 mL final volume, and the reaction mixture was incubated at 25°C for 2 h. The total reaction mixture was introduced into chemically competent *E. coli* DH5α cells (Thermo Fisher Scientific, Waltham, MA) using a heat shock protocol. The pTYGSarg::*APR1* expression vector was isolated from positive colonies using the Zyppy plasmid miniprep kit (Zymo Research, Irvine, CA).

Transformation *A. oryzae* NSAR1. Spores from *A. oryzae* NSAR1 were harvested from MEA plates in 5 mL of sterile water, and 1 mL of this spore suspension was inoculated into 50 mL of MB liquid medium and grown overnight at 28°C with shaking at 200 rpm. Germinating spores were collected by centrifugation at room temperature for 10 min at 3,500 rpm and resuspended in 25 mL of 0.8 M NaCl. After centrifugation for 10 min at 3,500 rpm at room temperature, germinated spores were resuspended in 10 mL of a freshly made filter-sterilized protoplasting solution (200 mg Trichoderma lysing enzyme [Thermo Fisher Scientific, Waltham, MA] and 50 mg Driselase [Thermo Fisher Scientific, Waltham, MA] in 0.8 M NaCl) and incubated at 30°C for 2 to 2.5 h with shaking at 100 rpm. Protoplasts were filtered through sterile Miracloth and then centrifuged for 5 min at 3,000 rpm at 4°C. Protoplasts were resuspended in 200 mL of solution 1 (0.8 M NaCl, 10 mM CaCl2, and 50 mM Tris-HCl [pH 7.5]) and aliquoted to 100 mL in 2 mL microcentrifuge tubes. Ten micrograms of the pTYGSarg::APR1 expression plasmid or the empty vector pTYGSarg was added to protoplasts, and the mixture was incubated on ice for 2 min. One milliliter of solution 2 (60% [wt/vol] polyethylene glycol 3350 [PEG 3350], 0.8 M NaCl, 10 mM CaCl2, and 50 mM Tris-HCl [pH 7.5]) was added, and the tubes were gently inverted before incubation at room temperature for 20 min. Protoplasts were then mixed with 25 mL of cooled selective CZD top 1.5% agar without arginine supplemented with 1 M sorbitol and immediately plated onto selective CZD bottom 0.8% agar without arginine supplemented with 1 M sorbitol. Transformation plates were incubated at 30°C for 3 to 10 days.

Secondary-metabolite extraction and HPLC–MS analyses. Secondary metabolites from 6-day-old *A. parvulus* or 5-day-old *A. oryzae* transformant liquid culture filtrates were isolated with a 1:1 volume of ethyl acetate (VWR Chemicals, Radnor, PA). After shaking on an orbital shaker for at least 1 h, the organic phase was transferred to a 50 mL tube and evaporated under nitrogen flow. The resulting solid was dissolved in acetonitrile. Organic extracts were analyzed with a Shimadzu LC-2030 3D-Prominence-i PDA system coupled to a Shimadzu LCMS-2020 mass spectrometer and equipped with a Shimadzu Shim-pack GIST C18-HP reversed-phase column (3 mm, 4.6 by 100 mm). The following method was used:

a linear gradient of buffer B (0 to 95%) for 20 min, 5 min of 95% buffer B, and then 100% buffer A for 5 min. Water with 0.1% trifluoroacetic acid (TFA) for high-performance liquid chromatography (HPLC) or 0.05% formic acid for mass spectrometry (MS)-coupled analyses was used as buffer A, and acetonitrile (LCMS grade) with 0.1% TFA for HPLC or 0.05% formic acid for MS-coupled analyses was used as buffer B. The flow rate was 1 mL/min or 0.5 mL/min for HPLC or MS-coupled analyses, respectively. The equipment was controlled and results were analyzed using Shimadzu LabSolutions LCMS software.

- **Compound purification.** Crude organic extracts were fractionated using a Shimadzu preparative HPLC system consisting of a CBM-20A controller, an LC-20AP pump, an SPD-20A detector, and an FRC-10A fraction collector, equipped with a C18 reversed-phase Reprosil column (10 mm, 120 Å, 250 by 22 mm). The system was controlled with Shimadzu LabSolutions software. A 12.5-mL flow was used with a linear gradient of buffer B (0 to 95%), 5 min of 95% buffer B, and then 100% buffer A for 5 min.

- **HRMS and NMR.** HRMS was performed using an LCT instrument (Micromass Ltd., Manchester, UK). Calibration was done with sodium formate, and measurements were acquired for samples mixed with sodium formate. ^{1}H (600 MHz) and ^{13}C (151 MHz) NMR analyses were performed for samples reconstituted in dimethyl sulfoxide (DMSO) on a Bruker 600 spectrometer and analyzed using MNOVA software. Chemical shifts for protons are reported in parts per million downfield from tetramethylsilane and are referenced to residual protium in the solvent (^{1}H NMR, DMSO-d6 at 2.50 ppm). Chemical shifts for carbons are reported in parts per million downfield from tetramethylsilane and are referenced to the carbon resonances of the residual solvent peak (^{13}C NMR, DMSO-d6 at 39.52 6 0.06 ppm). NMR data are represented as follows: chemical shift, multiplicity (s, singlet; bs, broad singlet; d, doublet; dd, doublet of doublet; t, triplet; q, quartet; ddd, doublet of doublet of doublets; dtd, doublet of triplet of doublets; m, multiplet), coupling constants (hertz), and integration.

Acknowledgements

References

01. **Atanasov AG, Zotchev SB, Dirsch VM (2021)**. International Natural Product Sciences Taskforce, Supuran CT. Natural products in drug discovery: advances and opportunities. Nat Rev Drug Discov 20:200–216. https://doi .org/10.1038/s41573-020-00114-z.

02. **Quinn R. (2013).** Rethinking antibiotic research and development: World War II and the penicillin collaborative. Am J Public Health 103:426–434. https://doi.org/10.2105/AJPH.2012.300693.

03. **Schueffler A, Anke T. (2014).** Fungal natural products in research and development. Nat Prod Rep 31:1425–1448. https://doi.org/10.1039/c4np00060a.

04. **Scherlach K, Hertweck C. (2021).** Mining and unearthing hidden biosynthetic potential. Nat Commun 12:3864. https://doi.org/10.1038/s41467 -021-24133-5.

05. **Singh G, Armaleo D, Dal Grande F, Schmitt I. (2021).** Depside and depsidone synthesis in lichenized fungi comes into focus through a genome-wide comparison of the olivetoric acid and physodic acid chemotypes of *Pseudevernia furfuracea*. Biomolecules 11:1445. https://doi.org/10.3390/ biom11101445.

06. **Armaleo D, Sun X, Culberson C. (2011).** Insights from the first putative bio-synthetic gene cluster for a lichen depside and depsidone. Mycologia 103:741–754. https://doi.org/10.3852/10-335.

07. **Kakule TB, Sardar D, Lin Z, Schmidt EW. (2013).** Two related pyrrolidine-dione synthetase loci in *Fusarium heterosporum* ATCC 74349 produce divergent metabolites. ACS Chem Biol 8:1549–1557. https://doi.org/10 .1021/cb400159f.

08. **Chen L, Yue Q, Zhang X, Xiang M, Wang C, Li S, Che Y, Ortiz-López FJ, Bills GF, Liu X, An Z. (2013).** Genomics-driven discovery of the pneumocandin biosynthetic gene cluster in the fungus *Glarea lozoyensis*. BMC Genomics 14:339. https://doi.org/10.1186/1471-2164-14-339.

09. **Robey MT, Caesar LK, Drott MT, Keller NP, Kelleher NL. (2021).** An interpreted atlas of biosynthetic gene clusters from 1,000 fungal genomes. Proc Natl Acad Sci U S A 118:e2020230118. https://doi.org/10.1073/pnas .2020230118.

10. **Tang S, Zhang W, Li Z, Li H, Geng C, Huang X, Lu X. (2020).** Discovery and characterization of a PKS-NRPS hybrid in *Aspergillus terreus* by genome mining. J Nat Prod 83:473–480. https://doi.org/10.1021/acs.jnatprod .9b01140.

11. **Kjærbølling I, Vesth TC, Frisvad JC, Nybo JL, Theobald S, Kuo A, Bowyer P, Matsuda Y, Mondo S, Lyhne EK, Kogle ME, Clum A, Lipzen A, Salamov A, Ngan CY, Daum C, Chiniquy J, Barry K, LaButti K, Haridas S, Simmons BA, Magnuson JK, Mortensen UH, Larsen TO, Grigoriev IV, Baker SE, Andersen MR. (2018).** Linking secondary metabolites to gene clusters through genome sequencing of six diverse *Aspergillus* species. Proc Natl Acad Sci U S A 115:E753–E761. https:// doi.org/10.1073/pnas.1715954115.

12. **Throckmorton K, Wiemann P, Keller NP. (2015).** Evolution of chemical diversity in a group of non-reduced polyketide gene clusters: using phylogenetics to inform the search for novel fungal natural products. Toxins (Basel) 7:3572–3607. https://doi.org/10.3390/toxins7093572.

13. **Theobald S, Vesth TC, Andersen MR. (2019).** Genus level analysis of PKS-NRPS and NRPS-PKS hybrids reveals their origin in Aspergilli. BMC Genomics 20:847. https://doi.org/10.1186/s12864-019-6114-2.

14. **Harvey CJB, Tang M, Schlecht U, Horecka J, Fischer CR, Lin H-C, Li J, Naughton B, Cherry J, Miranda M, Li YF, Chu AM, Hennessy JR, Vandova GA, Inglis D, Aiyar RS, Steinmetz LM, Davis RW, Medema MH, Sattely E, Khosla C, Onge RPS, Tang Y, Hillenmeyer ME. (2018).** HEx: a heterologous expression platform for the discovery of fungal natural products. Sci Adv 4:eaar5459. https://doi.org/10.1126/sciadv.aar5459.

15. **Adamek M, Alanjary M, Ziemert N. (2019).** Applied evolution: phylogeny-based approaches in natural products research. Nat Prod Rep 36: 1295–1312. https://doi.org/10.1039/c9np00027e.

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

17. **Mosunova O, Navarro-Muñoz JC, Collemare J. (2021).** The biosynthesis of fungal secondary metabolites: from fundamentals to biotechnological applications, p 458–476. In

Zaragoza Ó, Casadevall A (ed), Encyclopedia of mycology. Elsevier, Oxford, United Kingdom. https://doi.org/10.1016/ B978-0-12-809633-8.21072-8.

18. **Rankovic B, Kosanic M. (2015).** Lichens as a potential source of bioactive secondary metabolites, p1–26.InRankovicB(ed), Lichensecondary metabolites: bioactive properties and pharmaceutical potential. Springer International Publishing, Cham, Switzerland. https://doi.org/10.1007/978 -3-319-13374-4_1.

19. **Culberson CF. (1970).** Supplement to "Chemical and botanical guide to lichen products." Bryologist 73:177–377. https://doi.org/10.2307/3241261.

20. **Kealey JT, Craig JP, Barr PJ. (2021).** Identification of a lichen depside polyketide synthase gene by heterologous expression in *Saccharomyces cerevisiae*. Metab Eng Commun 13:e00172. https://doi.org/10.1016/j.mec.2021 .e00172.

21. **Kim W, Liu R, Woo S, Kang KB, Park H, Yu YH, Ha H-H, Oh S-Y, Yang JH, Kim H, Yun S-H, Hur JS, Turgeon BG. (2021).** Linking a gene cluster to atranorin, a major cortical substance of lichens, through genetic dereplication and heterologous expression. mBio 12:e01111-21. https://doi.org/10.1128/ mBio.01111-21.

22. **Jeong M-H, Park C-H, Kim JA, Choi ED, Kim S, Hur J-S, Park S-Y. (2021).** Production and activity of cristazarin in the lichen-forming fungus *Cladonia metacorallifera*. J Fungi (Basel) 7:601. https://doi.org/10.3390/jof7080601.

23. **Bertrand RL, Sorensen JL. (2018).** A comprehensive catalogue of polyketide synthase gene clusters in lichenizing fungi. J Ind Microbiol Biotechnol 45: 1067–1081. https://doi.org/10.1007/s10295-018-2080-y.

24. **Kautsar SA, Blin K, Shaw S, Navarro-Muñoz JC, Terlouw BR, van der Hooft JJJ, van Santen JA, Tracanna V, Suarez Duran HG, Pascal Andreu V, Selem-Mojica N, Alanjary M, Robinson SL, Lund G, Epstein SC, Sisto AC, Charkoudian LK, Collemare J, Linington RG, Weber T, Medema MH. (2020).** MIBiG 2.0: a repository for biosynthetic gene clusters of known function. Nucleic Acids Res 48:D454–D458. https://doi.org/10.1093/nar/gkz882.

25. **Liu L, Zhang Z, Shao C-L, Wang J-L, Bai H, Wang C-Y. (2015).** Bioinformatical analysis of the sequences, structures and functions of fungal polyketide synthase product template domains. Sci Rep 5:10463. https://doi.org/10 .1038/srep10463.

26. **Schumacher J. (2016).** DHN melanin biosynthesis in the plant pathogenic fungus *Botrytis cinerea* is based on two developmentally regulated key enzyme (PKS)-encoding genes. Mol Microbiol 99:729–748. https://doi .org/10.1111/mmi.13262.

27. **Griffiths S, Mesarich CH, Saccomanno B, Vaisberg A, De Wit PJGM, Cox R, Collemare J. (2016).** Elucidation of cladofulvin biosynthesis reveals a cytochrome P450 monooxygenase required for anthraquinone dimerization. Proc Natl Acad Sci U S A 113:6851–6856. https://doi.org/10.1073/pnas .1603528113.

28. **Jackson JC, Higgins LA, Lin X. (2009).** Conidiation color mutants of *Aspergillus fumigatus* are highly pathogenic to the heterologous insect host *Galleria mellonella*. PLoS One 4:e4224. https://doi.org/10.1371/journal.pone .0004224.

29. **Tsai H-F, Washburn RG, Chang YC, Kwon-Chung KJ. (1997).** *Aspergillus fumigatus* arp1 modulates conidial pigmentation and complement deposition. Mol Microbiol 26:175–183. https://doi.org/10.1046/j.1365-2958.1997 .5681921.x.

30. **Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS. (2017).** ModelFinder: fast model selection for accurate phylogenetic estimates. Nat Methods 14:587–589. https:// doi.org/10.1038/nmeth.4285.

31. **Skellam E. (2019).** Strategies for engineering natural product biosynthesis in fungi. Trends Biotechnol 37:416–427. https://doi.org/10.1016/j.tibtech .2019.03.014.

32. **Meng X, Fang Y, Ding M, Zhang Y, Jia K, Li Z, Collemare J, Liu W. (2022).** Developing fungal heterologous expression platforms to explore and improve the production of natural products from fungal biodiversity. Biotechnol Adv 54: 107866. https://doi.org/10.1016/ j.biotechadv.2021.107866.

33. **Bertrand RL, Sorensen JL. (2019).** Lost in translation: challenges with heterologous expression of lichen polyketide synthases. ChemistrySelect 4: 6473–6483. https://doi.org/ 10.1002/slct.201901762.

34. **Vagstad AL, Hill EA, Labonte JW, Townsend CA. (2012).** Characterization of a fungal thioesterase having Claisen cyclase and deacetylase activities in melanin biosynthesis. Chem

Biol 19:1525–1534. https://doi.org/10.1016/j .chembiol.2012.10.002.

35. **Feng L-X, Zhang B-Y, Zhu H-J, Pan L, Cao F. (2020).** Bioactive metabolites from
Talaromyces purpureogenus, an endophytic fungus from *Panax notoginseng*. Chem Nat
Compd 56:974–976. https://doi.org/10.1007/s10600 -020-03206-9.

36. **Simpson TJ, Stenzel DJ. (1981).** 13C and 2H n.m.r. studies on the biosynthesis of
Omethylasparvenone, a hexaketide metabolite of *Aspergillus parvulus*. J Chem Soc Chem
Commun 5:239–240.

37. **Bartman CD, Campbell IM. (1979).** Naphthalenone production in *Aspergillus parvulus*. Can J
Microbiol 25:130–137. https://doi.org/10.1139/m79-021.

38. **Wheeler MH, Abramczyk D, Puckhaber LS, Naruse M, Ebizuka Y, Fujii I, Szaniszlo PJ.
(2008).** New biosynthetic step in the melanin pathway of *Wangiella* (*Exophiala*) *dermatitidis*:
evidence for 2-acetyl-1,3,6,8-tetrahydroxynaphthalene as a novel precursor. Eukaryot Cell
7:1699–1711. https://doi.org/10.1128/EC.00179-08.

39. **Tsai H-F, Wheeler MH, Chang YC, Kwon-Chung KJ. (1999).** A developmentally regulated
gene cluster involved in conidial pigment biosynthesis in *Aspergillus fumigatus*. J Bacteriol
181:6469–6477. https://doi.org/10.1128/ JB.181.20.6469-6477.1999.

40. **Burruano S, Giambra S, Mondello V, Dellagreca M, Basso S, Tuzi A, Andolfi A. (2016).**
Naphthalenone polyketides produced by *Neofusicoccum parvum*, a fungus associated with
grapevine Botryosphaeria dieback. Phytopathol Mediterr 55:197–206.

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

42. **Blin K, Wolf T, Chevrette MG, Lu X, Schwalen CJ, Kautsar SA, Suarez Duran HG, de Los
Santos ELC, 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.
https://doi.org/10.1093/ nar/gkx319.

43. **Sievers F, Higgins DG. (2018).** Clustal Omega for making accurate alignments of many
protein sequences. Protein Sci 27:135–145. https://doi .org/10.1002/pro.3290.

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

45. **Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. (2015).** IQ-TREE: a fast and effective
stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol 32:268–274.
https://doi.org/10.1093/molbev/ msu300.

46. **Anisimova M, Gil M, Dufayard J-F, Dessimoz C, Gascuel O. (2011).** Survey of branch
support methods demonstrates accuracy, power, and robustness of fast likelihood-based
approximation schemes. Syst Biol 60:685–699. https://doi.org/10.1093/sysbio/syr041.

47. **Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. (2010).** New
algorithms and methods to estimate maximum-likelihood phylogenies: assessing the
performance of PhyML 3.0. Syst Biol 59: 307–321. https://doi.org/10.1093/sysbio/syq010.

48. **Letunic I, Bork P. (2021).** Interactive Tree of Life (iTOL) v5: an online tool for phylogenetic
treedisplay and annotation. Nucleic Acids Res 49: W293–W296.
https://doi.org/10.1093/nar/gkab301.

49. **Lazarus CM, Williams K, Bailey AM. (2014).** Reconstructing fungal natural product
biosynthetic pathways. Nat Prod Rep 10:1339–1347. https://doi .org/10.1039/c4np00084f.

50. **Kombrink A. (2012).** Heterologous production of fungal effectors in *Pichia pastoris*.
Methods Mol Biol 835:209–217. https://doi.org/10.1007/978-1 -61779-501-5_13.

51. **Gilchrist CLM, Chooi Y-H. (2021).** Clinker & clustermap.js: automatic generation of gene
cluster comparison figures. Bioinformatics 37:2473–2475.
https://doi.org/10.1093/bioinformatics/btab007.

52. **Griffiths SA, Cox RJ, Overdijk EJR, Mesarich CH, Saccomanno B, Lazarus CM, de Wit
PJGM, Collemare J. (2018).** Assignment of a dubious gene cluster to melanin biosynthesis in
the tomato fungal pathogen *Cladosporium fulvum*. PLoS One 13:e0209600.
https://doi.org/10.1371/journal.pone.0209600.

3

Supplementary materials

- **Supplementary Data Set S01.** Protein sequences of nonreducing polyketide synthases, hydroxynaphthalene reductases, and scytalone dehydratases.

- **Supplementary Data Set S02.** Protein alignments of nonreducing polyketide synthases, hydroxynaphthalene reductases, and scytalone dehydratases.

- **Supplementary Data Set S03.** Phylogenetic trees of nonreducing polyketide synthases, hydroxynaphthalene reductases, and scytalone dehydratases.

- **Supplementary Data Set S04.** Curated gene models.

Supplementary Table S01.
BLASTp search for proteins from the Aspergillus parvulus group XI biosynthetic pathway.

- BlastP search was performed on filtered models of fungal species harbouring the group XI non-reducing polyketide synthases at the Joint Genome Institute MycoCosm repository

- Queries were protein sequences from A. parvulus group XI predicted pathway except for the predicted transcription factor

- For each hit, the protein id, score, e-value, identity and coverage are provided.

- Hits in italics are distant homologues that are predicted to belong to a different pathway (phylogenetic clade as in Figure 03 is indicated in between bracket for Apr5 and Apr6 homologues).

- Close homologues found at a different locus are indicated in bold. Incorrect protein models due to disruptive mutations are indicated with a star.

- Cells highlighted in yellow indicate the absence of a close homologue

Find on P. 130 – 137

Supplementary Table S01.

BLASTp search for proteins from *Aspergillus parvulus* group XI biosynthetic pathway.

A. parvulus protein id	Predicted function	Protein name	Oidiodendron maius Zn	Atropellis piniphila CBS 197.64
81212	non-reducing polyketide synthase	Apr1	149275|5017|0.0|49.7|90.7	60542|0|5212|0.0|53.1|89.2
81217	oxidoreductase	Apr2	135662|100|13.01e-110|51.7|88.3	60542|1|120|13.38e-133|51.7|94
171936	aldo-keto reductase	Apr8	115438|675|1.80e-074|48.9|88.5	352615|82|13.35e-099|56.9|82.3
193045	NmrA-like protein		182281|94|65.35e-116|56.7|101.6	60490|3|910|4.43e-111|55.4|99.4
193042	O-methyltransferase	Apr4	151812|83|19.46e-088|42.8|84.8	90173|83|19.58e-088|42.8|70.7
213579	hydroxynaphthalene reductase	Apr6	39552|669|14.11e-079|52.3|93.7 (unique) 102294|656|1.74e-44|50|94.4 (4Hnr) 123214|1595|3.16e-64|46.2|86.1 (3Hnr)	50671|3180|518.45e-098|60.5|88.9 50009|7|646|1.96e-42|49.6|93.2 (4Hnr) 45141|4|6311.36e-64|48.8|73.7 (3Hnr) 60163|7|6241|1.55e-50|48.6|92.1 (4Hnr)
201402	scytalone dehydratase	Apr5	116220|446|1e-048|50.3|96.2 (unique) 144341|405|4.08e-43|43.7|44.8 (Scd1)	52663|6221|4.2e-068|66.7|101.8 50393|6|395|9.72e-42|40.7|79 (Scd1) 44769|6|366|9.01e-38|51.2 (Arp1)
201403	oxidoreductase	Apr7	74151|399|1.96e-025|40.4|43.3	65015|8111|84|4.98e-120|46.2|89.9
171934	O-acetyltransferase	Apr3	61211|4.31e-037|34.7|40.3	35271|0|78|12.30e-078|43.6|73.3
81192	GroES-like alcohol dehydrogenase		17034|156|913.79e-051|39.1|82.4	52789|5|42|13.42e-025|34.1|77.5

Supplementary Table S01.

BLASTp search for proteins from *Aspergillus parvulus* group XI biosynthetic pathway **(continued)**.

A. parvulus protein id	Predicted function	Protein name	Umbilicaria pustulata	Sporormia fimetaria	Setosphaeria turcica NY001
81212	non-reducing polyketide synthase	Apr1	10240|7|5356|0.0|52.5|98	41245|16|39|0.0|56.6|93.6	*54177|7|5809|0.0|54.7|93.3
81217	oxidoreductase	Apr2	10240|81|134411.23e-143|56|93.2	51113|41|14221|3.03e-163|61.9|88	44062|6|139|6.31e-154|58.2|96.7
171936	aldo-keto reductase	Apr8	10240|41|916|1.79e-107|56.9|92.2	41242|6|100|111.34e-129|60.1|96.4	*49371|8|99|6|3.01e-123|60.5|91.8
193045	NmrA-like protein		10249|5|978|1.4e-120|57|97.8	41235|1|100|311.51e-129|61.11|100	54177|5|89|811.46e-109|58.6|101.4
193042	O-methyltransferase	Apr4	10210|5|518|15.45e-054|33.9|56.3 (unique)	49049|41|399|12.64e-180|62|96.9	49372|1|13|10|6.96e-158|6.0.7|97.5
213579	hydroxynaphthalene reductase	Apr6	10603|8|69714.07e-045|50.8|96 (4Hnr)	52976|5|192|2.94e-118|66.7|99.6 45936|6|716|6.30e-81|49.6|104.5 (4Hnr)	49370|1|903|5.96e-117|67.4|99.6 41857|1|73|719.78e-84|50.1|104.5 (4Hnr) 44076|5|1622|5.08e-68|45.8|98.1 (ClaC) 44096|3|598|3.11e-60|47.2|93.6 (3Hnr)
201402	scytalone dehydratase	Apr5	10603|7|406|6.08e-39|53.5|56.8 (unique)	52215|1|66|1.70e-084|72.7|91.2 39481|4|389|4.27e-41|41.2|95.5 (Scd1)	53252|0|644|2.12e-081|69.5|95.4 46847|1|3|344|1.61e-33|42.1|89.5 (ClaB) 51247|5|287|4.16e-27|43.6|90.2 (Scd1) 52124|1|2201|7.54e-14|31.7|79 (outgroup)
201403	oxidoreductase	Apr7	10240|81|33|18.58e-026|31.6|47.3	41241|3|130|711.13e-160|53.4|83.9	49371|5|11|235|18.96e-151|49.1|91.9
171934	O-acetyltransferase	Apr3	no hit	49049|6|110|911.68e-111|51.4|82	53436|1|333|11.5e-029|38.3|40.9
81192	GroES-like alcohol dehydrogenase		10764|1|3971.2e-027|38.5|67.8	41241|4|1109|3.31e-144|62.7|99.7	41992|7|32|6|7.3e-026|40.9|54.4

3

Supplementary Table S01.

BLASTp search for proteins from the *Aspergillus parvulus* group XI biosynthetic pathway **(continued)**.

A. parvulus protein id	Predicted function	Protein name	Dothidotthia symphoricarpi	Bimuria novae-zelandiae CBS 107.79	Didymocrea sadasivanii CBS 438.65
81212	non-reducing polyketide synthase	Apr1	355059|6148|0.0|53.7|99.5	502017|5989|0.0|E000|55.99|92.35	493665|6005|0.0|E000|56.77|93.23
81217	oxidoreductase	Apr2	335398|1373|1.66e-156|62.3|86.7	528606|1151|4.88E-117|53.6|91.28	416795|1173|4.12E-120|54.31|91.28
171936	aldo-keto reductase	Apr8	364847|981|3.29e-121|60.9|93	551685|964|2.16E-124|57.94|97.27	416784|995|1.05E-128|59.81|97.27
193045	NmrA-like protein		383339|1050|6.23e-136|64.6|99.7	502006|985|2.31E-121|61.29|98.41	353526|1010|7.79E-125|62.58|98.41
193042	O-methyltransferase	Apr4	311485|1443|0.0|63.8|96.2	456270|1382|7.94E-178|62.35|96.24	481287|1377|3.41E-177|61.37|96.24
213579	hydroxynaphthalene reductase	Apr6	364850|926|3.85e-120|68.6|99.6 359153|728|1.6|e-82|51.1|1103 (4Hnr) 356504|1595|7.68e-60|46.7|95.5 (3Hnr) 373210|1590|1.15e-63|44.6|99.2 (ClaC)	501991|784|1.88E-100|62.20|93.18 335842|751|1.39e-85|52.21|103 (4Hnr) 464464|600|1.94e-60|48|92.5 (3Hnr)	*129033|834|2.03E-107|64.17|92.70 367636|747|4.24e-85|52.21|103 (4Hnr) 461289|563|2.22e-55|47.21|89.1 (3Hnr)
201402	scytalone dehydratase	Apr5	191076|653|1.17e-082|69.5|92.8 368002|401|1.08e-42|43.5|87.5 (Scdl) 420469|349|3.15e-31|42.6|94.6 (ClaB)	565899|650|3.66E-082|72.22|91.53 581419|372|1.20e-38|41.1|92.4 (Scdl)	481283|564|2.36E-070|73.38|78.53 408439|412|3.52e-44|42.4|96.7 (Scdl)
201403	oxidoreductase	Apr7	284731|1319|2.9e-162|50.1|90.8	578616|1173|1.15e-132|46.56|95.35	444021|1211|2.77E-133|48.25|85.55
171934	O-acetyltransferase	Apr3	365365|3299|1.94e-033|36.6|42.3	456274|994|6.86E-122|51.59|87.63	416793|1050|5.50E-125|50.40|83.04
81192	GroES-like alcohol dehydrogenase		435522|3533|8.71e-027|42|52.9	551683|1063|3.83E-132|64.10|89.66	353683|1059|1.18E-131|64.10|89.66

Supplementary Table S01.

BLASTp search for proteins from the *Aspergillus parvulus* group XI biosynthetic pathway **(continued)**.

A. parvulus protein id	Predicted function	Protein name	Paraconiothyrium sporulosum AP3s5-JAC2a	Karstenula rhodostoma CBS 690.94	Lobaria pulmonaria Scotland reference
81212	non-reducing polyketide synthase	Apr1	116758|3157|8310.00|E000|56.19|93.11	433550|5990|0.00|E000|55.60|92.95	126715|6162|03|0.00E000|55.79|97.32
81217	oxidoreductase	Apr2	115307|911|73|6.45E-115|54.81|88.51	433548|1235|14.32E-122|54.04|94.89	108924|3113|44|1.74E-152|57.34|90.59
171936	aldo-keto reductase	Apr8	126245|1194|3|1.64E-121|56.17|98.18	433541|1607|12.16E-075|62.37|83.98	12468|1549|2.78E-051|45.16|77.02
193045	NmrA-like protein		116758|1199|014.78E-122|61.41|85.21	363269|1993|1.89E-122|61.41|98.73	108925|8|1141|1.15E-142|68.91|95.41
193042	O-methyltransferase	Apr4	115307|0|1313|15.07E-174|60.45|93.41	585851|322|2.90E-175|60.71|93.41	118199|1916|1.03E-103|43.70|92.40
213579	hydroxynaphthalene reductase	Apr6	118971|6|782|3.53E-100|62.60|96.09 121345|317|3916.05e-84|51.18|1103 (4Hnr) 103686|1591|3.2e-59|44.2|89.6 (3Hnr)	487184|828|1.59E-106|64.17|95.49 17313|736|1.58e-83|52.21|103 (4Hnr) 309351|608|1.61e-61|46.7|95.9 (3Hnr)	*124926|7|855|9.6|6E-105|68.62|99.58 566060|6|1111.81e-66|43.7|1101.5 (ClaC) 744976|429|1.13e-40|45.11|68.4 (outgroup) 130953|4142|415.56e-40|45.11|67.4 (outgroup)
201402	scytalone dehydratase	Apr5	Not predicted, but present 119010|5142|6|14.90E-046|42.94|96.72 (Scdl)	328866|1597|17.65E-075|73.47|100.00 427012|1411|15.61e-44|41.8|96.7 (Scdl)	*124926|1162|11.79E-073|80.71|95.24 126310|1141|213.92e-44|44.41|101.8 (ClaB)
201403	oxidoreductase	Apr7	117944|3511|931|2.22E-135|50.47|80.68	416712|1187|11.46E-134|49.77|82.18	107898|0|965|1.64E-104|40.43|87.24
171934	O-acetyltransferase	Apr3	187614|1062|1.31E-126|51.20|82.06	433547|1060|12.54E-126|51.20|82.06	107812|4|382|1.38E-034|39.47|46.80
81192	GroES-like alcohol dehydrogenase		118971|5|1046|8.04E-130|62.82|89.66	433532|1066|11.51E-132|64.74|89.66	106044|71|1058|1.18E-121|70.92|85.98

3

Supplementary Table S01.

BLASTp search for proteins from the *Aspergillus parvulus* group XI biosynthetic pathway **(continued)**.

A. parvulus protein id	Predicted function	Protein name	Cladonia grayi Cgr/DA2myc/ss	Neofusicoccum parvum UCRNP2	Melanops tulasnei CBS 116805
81212	non-reducing polyketide synthase	Apr1	*1822|6085|0.00E000|57.9|92.4	10180|5088|0.00E000|63.47|85.45	5969846|918|0.00E000|63.78|95.89
81217	oxidoreductase	Apr2	1821|1371|1.20E-151|58.0|85.3	10179|1517|0.00E000|64.37|91.77	5969861|603|0.00E000|66.82|92.83
171936	aldo-keto reductase	Apr8	4185|1498|3.44E-046|41.1|70.9	10177|673|1.29E-084|68.23|81.01	83000|516|1.21E-048|40.07|17.17
193045	NmrA-like protein		18201|1194|1.24E-155|70.6|100.0	23671|6611|.40E-072|44.14|96.03	146951|356|2.83E-023|43.95|46.59
193042	O-methyltransferase	Apr4	2762|1731|2.95E-066|44.7|59.7	10176|1249|1.71E-154|64.90|87.15	3351441|9021|.25E-106|41.41|73.20
213579	hydroxynaphthalene reductase	Apr6	1817|992|3.04e-129|72.3|99.6 / 454|726|3.11e-82|53.81|98.5 / 10393|7251|6.38e-87|52.8|84.4 / 4947|6611|3.98e-78|49.6|96	7230|741|2.22E-084|51.43|104.48 (4Hnr) / 7091|597|1.05e-64|47.1|89.3 (3Hnr) / 3580|467|1.78e-38|47.2|83.1 (outgroup)	68889|6201|.55e-72|49.2|94.7 (outgroup) / 5345|6471|73412.79E-068|50.53|106.34 (4Hnr) / 55146|603|1.8e-65|47|91.9 (3Hnr)
201402	scytalone dehydratase	Apr5	1818|711|1.19e-90|74.(1|00 / 1423|436|1.79e-47|43.71|70.7 / 4943|4191|1.33e-40|47.8|61.1 / 10394|4185|1.19e-45|52.9|36.1 / 4941|3351|1.16e-33|50.8|54.9	7468|3806|.71E-040|43.79|189.47 (Scd1)	58618|1139|211.73E-041|42.14|78.71 (Scd1)
201403	oxidoreductase	Apr7	1819|1467|10.00E000|53.1|95.7	7049|3616|.43E-025|40.70|34.61	54048|5133|18.99E-026|44.53|29.46
171934	O-acetyltransferase	Apr3	8152|4181|6.75E-046|40.0|58.2	10178|5591|8.60E-066|60.24|88.30	64940|6|2451|.63E-017|42.72|112.86
81192	GroES-like alcohol dehydrogenase		4902|420|1.23E-029|36.7|65.6	1816|1179|17.75E-154|65.60|98.28	65940|4|32912|.96E-031|34.74|62.46

Supplementary Table S01.

BLASTp search for proteins from the *Aspergillus parvulus* group XI biosynthetic pathway **(continued)**.

A. parvulus protein id	Predicted function	Protein name	Aspergillus nutans CBS 121.56	Aspergillus transcarpathicus CBS 423.68	Aspergillus avenaceus IBT 18842
81212	non-reducing polyketide synthase	Apr1	23910|710409|0.00E000|95.1|99.1	12094|1103351|0.00E000|93.99|99.11	12771|91676|10.00E000|62.32|96.01
81217	oxidoreductase	Apr2	2391|0912226|0.00E000|93.04|95.04	12093|712238|0.00E000|93.04|95.04	17000|311513|0.00E000|62.06|92.87
171936	aldo-keto reductase	Apr8	64632|114601|0.00E000|87.16|100.00	12091|9143610|.00E000|85.32|100.00	17000|01|0631|2.35E-137|61.16|100.00
193045	NmrA-like protein		2391|111428|0.00E000|87.03|100.00	14714|11503|0.00E000|91.46|100.00	54909|1223|16.53E-159|72.78|100.00
193042	O-methyltransferase	Apr4	23910|3121071|0.00E000|98.53|96.45	12094|62|105|0.00E000|98.53|96.45	15048|71|612|0.00E000|72.41|96.67
213579	hydroxynaphthalene reductase	Apr6	2391|131325|3.32E-175|99.62| / 426|77|4391|.23e-34|43.4|83.5 (outgroup) 100.00	93847|113103|.6OE-173|98.49|99.25 / 147245|4341|4.07E-038|40.4|91.0 (outgroup)	17000|21|02415|.69E-133|75.09|100.00 / 15484|41396|2.88E-036|45.1|69.7 (outgroup)
201402	scytalone dehydratase	Apr5	23911|5|887|6.29E-115|97.62|100.00 / 18838|01|204|1.02e-111|34.7|71.7 (outgroup)	12092|81899|11.28E-116|98.81|100.00	13753|21741|4.15E-094|81.82|100.00 / 14664|1213|6.30E-013|30.2|81.8 (outgroup)
201403	oxidoreductase	Apr7	23917|126610|0.00E000|93.94|100.00	9387|1256|0.00E000|91.15|100.76	13753|1627|0.00E000|59.24|92.74
171934	O-acetyltransferase	Apr3	1890|46|215|0.00E000|94.20|96.64	12093|5|2124|0.00E000|93.04|96.64	13753|51026|2.98E-116|47.64|89.16
81192	GroES-like alcohol dehydrogenase		64540|1599|0.00E000|91.01|100.00	12095|11542|0.00E000|87.01|100.00	16404|3|396|3.10E-035|35.1|167.0

Supplementary Table S01.

BLASTp search for proteins from the *Aspergillus parvulus* group XI biosynthetic pathway **(continued)**.

A. parvulus protein id	Predicted function	Protein name	Aspergillus alliaceus CBS 536.65	Aspergillus albertensis	Aspergillus bombycis NRRL 26010
81212	non-reducing polyketide synthase	Apr1	322797I66381I0.00E000I95.71I61.38	182792I66351I0.00E000I61.66I95.01	57231I67081I0.00E000I61.30I97.13
81217	oxidoreductase	Apr2	303631I1528I0.00E000I91.70I65.16	1933.231I535I0.00E000I65.38I91.70	57241I15420.00E000I64.16I91.31
171936	aldo-keto reductase	Apr8	322802I1068I4.55E-133I99.39I61.73	171838I106618.52E-133I61.73I99.39	3831I10341I5.14E-134I58.72I95.61
193045	NmrA-like protein		315594I122714.04E-160I100.00I71.84	1933.221I1232I8.33E-161I72.47I100.00	57251I2131I3.50E-158I71.52I100.00
193042	O-methyltransferase	Apr4	315589I15071I0.00E000I96.67I67.00	153707I1504I0.00E000I67.00I96.67	57211I1505I0.00E000I67.24I96.67
213579	hydroxynaphthalene reductase	Apr6	315595I953I794E-124I98.47I72.09 / 315425I4244.97E-040I40.11I78.7 (outgroup)	1933.21I952I1.09E-123I71.71I98.47 / 173094I42011.67E-039I42.11I75.3 (outgroup)	572619836.00E-128I75.00I97.71 / 5303I6321.05E-051I46.91I40.5 (unique) / 316314701.84E-046I49.21I72.5 (outgroup)
201402	scytalone dehydratase	Apr5	315596I726I1.08E-092I99.39I78.66 / 297172I3621.257E-037I50.41I73.7 (outgroup)	171840I7261I.08E-092I78.66I99.39 / 152903I431I9.32E-047I51.41I88.9 (outgroup) / 189558I42012.99E-045I45.11I94.7 (ClaB)	5727I17351I6.51E-094I80.61I100.00 / 49721I26318.60E-024I31.51I96.1 (outgroup)
201403	oxidoreductase	Apr7	315597I16653I0.00E000I92.55I61.57	1933.201I66610.00E000I60.71I94.61	5728I16653I0.00E000I61.20I92.94
171934	O-acetyltransferase	Apr3	315593I1005I5.46E-114I87.06I47.61	182790I1007I2.86E-114I47.86I87.06	manually predicted
81192	GroES-like alcohol dehydrogenase		171698I4241I.56E-021I36.01I78.4	171698I4241I.56E-021I36.01I78.4	9505I72811.83E-085I50.51I85.76

Supplementary Table S01.

BLASTp search for proteins from the *Aspergillus parvulus* group XI biosynthetic pathway **(continued)**.

A. parvulus protein id	Predicted function	Protein name	Aspergillus pseudotamarii CBS 117625	Aspergillus caelatus CBS 763.97
81212	non-reducing polyketide synthase	Apr1	2927236I7780I0.00E000I60.75I99.06	16437I167110.00E000I95.02I61.92
81217	oxidoreductase	Apr2	271113I152110.00E000I64.06I90.87	16441I15220.00E000I91.31I63.27
171936	aldo-keto reductase	Apr8	119583I1011I1.54E-131I58.41I100.00	160085I1061I4.53E-132I100.31I59.63
193045	NmrA-like protein		282749I121817.32E-159I71.84I100.00	132066I1223I1.57E-159I86.81I72.15
193042	O-methyltransferase	Apr4	271130I148110.00E000I66.50I96.67	160076I1492I0.00E000I96.67I67.00
213579	hydroxynaphthalene reductase	Apr6	292718I982I8.35E-128I74.13I98.85 / 28875I6221I9.16E-051I46.21I00.0 (unique) / 252674I446I5.28E-043I45.31I73.4 (outgroup)	132067I98414.57E-128I97.71I75.00 / 122350I4481I2.92E-043I47.01I71.4 (outgroup)
201402	scytalone dehydratase	Apr5	282747I72816.10E-093I80.00I100.00 / 267375I231I1.94E-019I34.81I86.3 (outgroup)	164451726I1.19E-092I80.01I100.00 / 127671I2341I7.87E-020I35.61I86.3 (outgroup)
201403	oxidoreductase	Apr7	282746I1662I0.00E000I0.84I92.74	22425I1674110.00E000I92.74I61.85
171934	O-acetyltransferase	Apr3	282750I1007I1.79E-112I47.86I92.11	160080I107419.70E-113I92.11I48.10
81192	GroES-like alcohol dehydrogenase		268556I696I4.34E-081I48.81I85.01	169909I6881I5.54E-080I85.01I49.15
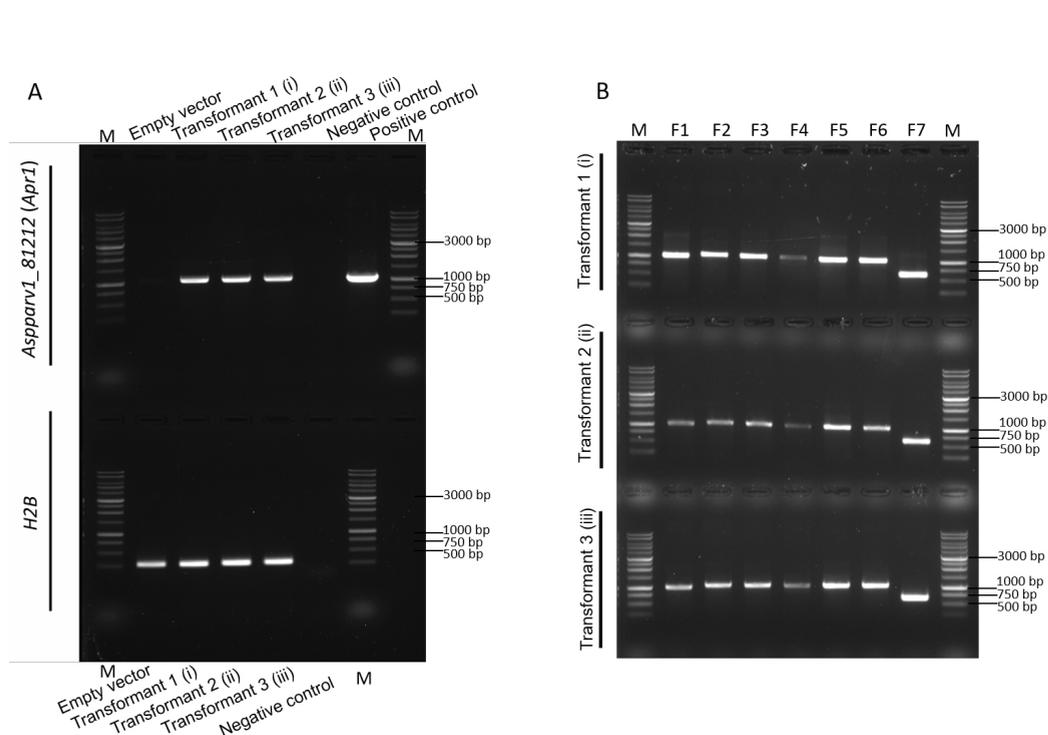
Supplementary Figure S01.

Expression of Aspparv1_81212 in *Aspergillus oryzae* transformants.

(A) Expression of *Aspparv1_81212* and H2B genes in *Aspergillus oryzae* transformants. M indicates
the BenchTop 1-kb ladder Promega), the negative control contains water, and the positive control
for *Aspparv1_81212* contains plasmid pTYGSarg::Aspparv1_81212.

(B) Analysis of *Aspparv1_81212* expression by *A. oryzae* transformants. F1 to F7 designate fragments
of the *Aspparv1_81212* gene; M designates the BenchTop 1-kb ladder (Promega). All fragments have
the expected sizes (F1, 1,007 bp; F2, 1,059 bp; F3, 1,044 bp; F4, 1,030 bp; F5, 1,035 bp; F6, 1,029
bp; F7, 628 bp).

Supplementary Figure S02.

High Resolution Mass Spectrometry (HRMS) data for metabolite 1 produced by *Aspergillus oryzae*
NSAR1 transformants expressing *Aspparv1_81212*. Compound with molecular formula $C_{12}H_{11}O_5$ and
mass 235.0577 (highlighted in green) is consistent with acetyl tetrahydroxynaphthalene.

Supplementary Data Set S05.

¹H and ¹³C NMR spectra of products

Product 01: ATHN

¹H NMR (600 MHz, DMSO)

¹³C NMR (151 MHz, DMSO)

¹H NMR (600 MHz, DMSO) δ 11.24 (s, O*H* x2), 10.10 (s, O*H* x2), 6.27 (d, *J* = 4.0 Hz, 2H), 6.13 (dd, *J* = 4.2, 2.1 Hz, 1H), 2.68 (s, *J* = 3.9 Hz, 3H, COC*H*$_3$).

¹³C NMR (151 MHz, DMSO) δ 203.85 (*C*O-CH$_3$), 167.43 (*C*-OH), 161.68 (*C*-OH), 159.44 (C-OH), 156.67 (*C*-OH), 141.76, 105.45, 103.73, 100.44, 99.50, 99.17, 32.51 (CO*C*H$_3$).

Supplementary Data Set S05.

¹H and ¹³C NMR spectra of products

Product 02: 6,8-dihydroxy-3-methylisocoumarin
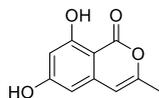
¹H NMR (600 MHz, DMSO)
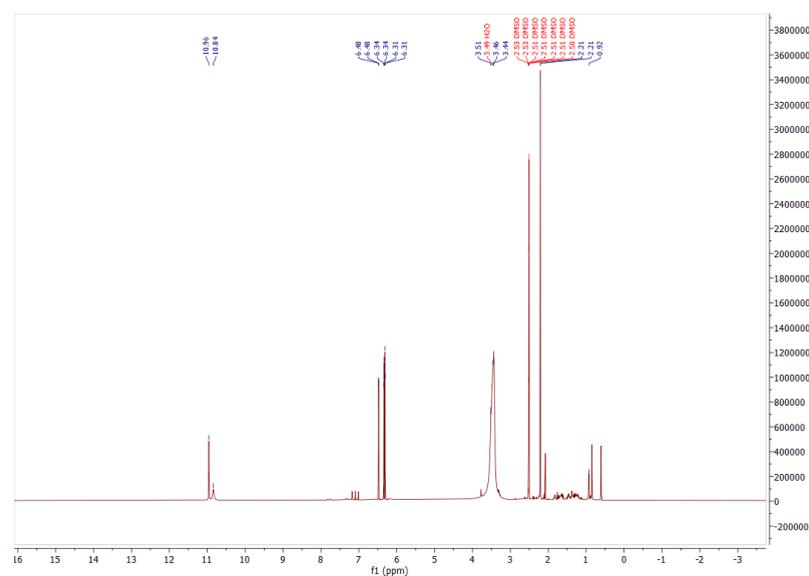
¹H NMR (600 MHz, DMSO) δ 10.96 (s, 1H, O*H*), 10.84 (s, 1H, O*H*), 6.48 (d, J = 1.2 Hz, 1H), 6.34 (d, J = 2.2 Hz, 1H), 6.31 (d, J = 2.2 Hz, 1H), 2.21 (d, J = 1.0 Hz, 3H, CH₃).

¹³C NMR (151 MHz, DMSO)
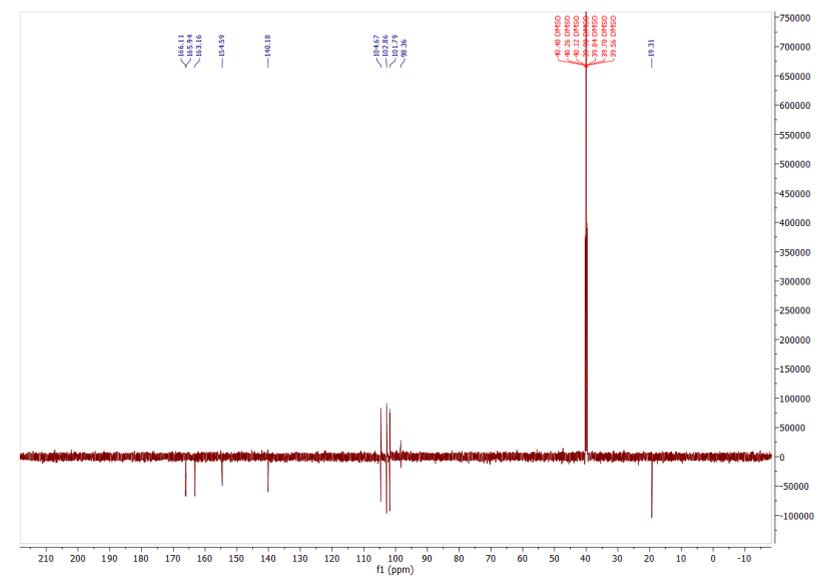
¹³C NMR (151 MHz, DMSO) δ 166.11 (*C*O, carbonyl), 165.94 (*C*-OH, aromatic), 163.16 (*C*-OH, aromatic), 154.59 (*C*-CH₃), 140.18, 104.67, 102.86, 101.79(CH, aromatic), 98.36 (CH, aromatic), 19.31 (CH₃).

HSQC (600 MHz, DMSO)
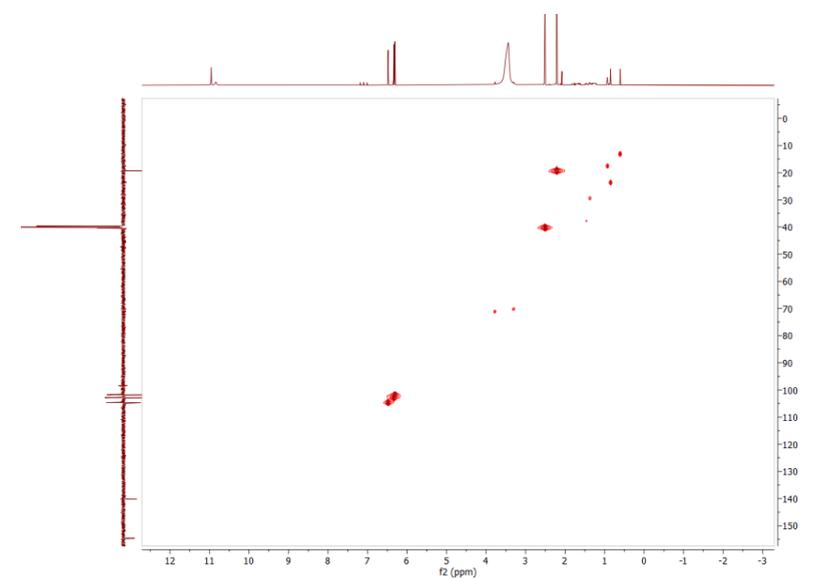
Supplementary Data Set S05.

¹H and ¹³C NMR spectra of products

Product 04: 6-O-methylasparvenone

¹H NMR (600 MHz, DMSO)
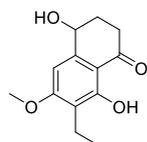
¹³C NMR (151 MHz, DMSO)
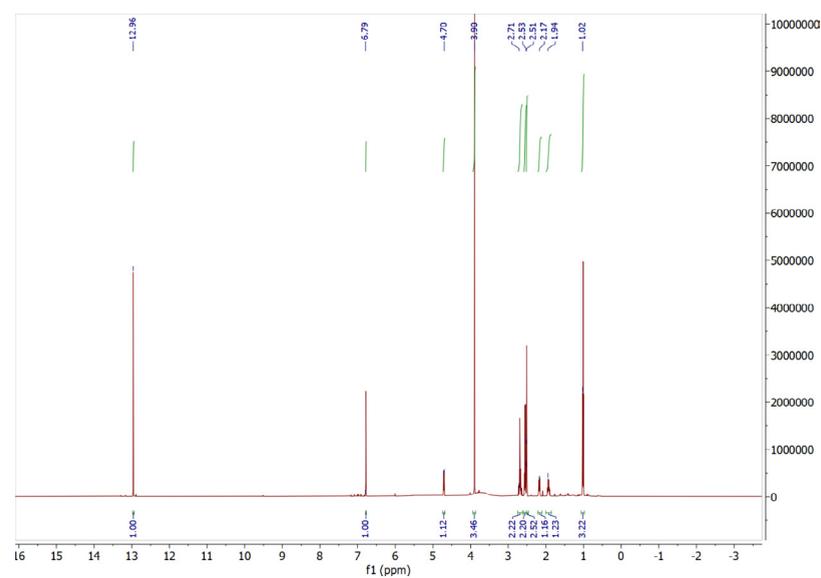
¹H NMR (600 MHz, DMSO) δ 12.96 (s, 1H), 6.78 (d, J = 0.9 Hz, 1H), 4.71 (ddd, J = 9.3, 4.1, 1.0 Hz, 1H), 3.90 (s, 3H, OCH₃), 2.76 – 2.62 (m, 2H), 2.55 (t, J = 7.4 Hz, 2H), 2.17 (dtd, J = 12.6, 5.2, 4.1 Hz, 1H), 1.93 (dtd, J = 12.6, 9.6, 5.4 Hz, 1H), 1.01 (t, J = 7.4 Hz, 3H, CH₂-CH₃).
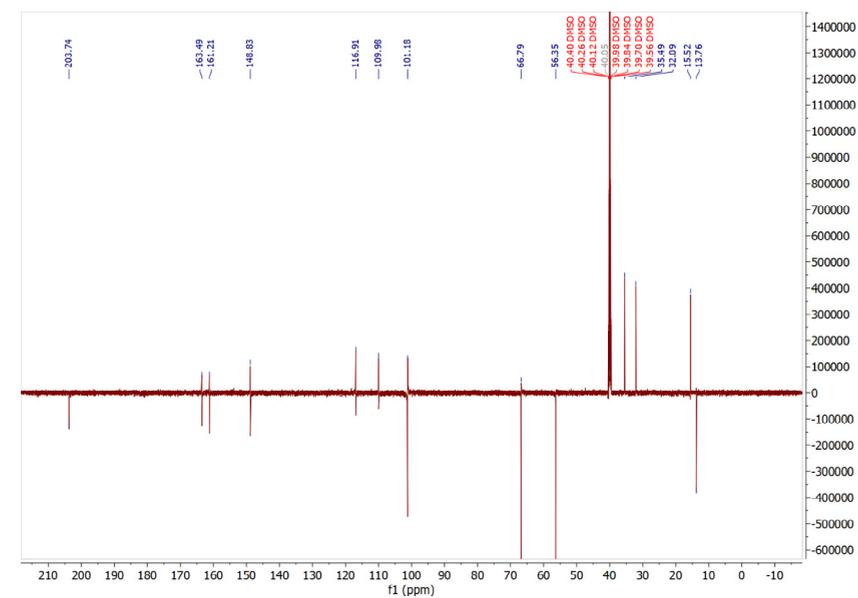
¹³C NMR (151 MHz, DMSO) δ 203.74 (*C*O, carbonyl), 163.49 (*C*-OCH₃), 161.21 (*C*OH, aromatic), 148.83, 116.91 (*C*H-CH₂-CH₃)., 109.98, 101.18 (*C*H, aromatic), 66.79 (C-OH, cyclohexane), 56.35 (OCH₃), 35.49 (*C*H₂, cycohexane), 32.09 (*C*H₂, cycohexane), 15.52 (*C*H₂-CH₃)., 13.76 (CH₂-*C*H₃).

Supplementary Table S02.

Oligonucleotides used in this study.

- Flanks homologous to pEYA2 entry vector are highlighted in red.
- Primers 1-10 were used to produce DNA fragments for transformation-associated recombination in yeast.
- Primers 11-52 were used for gene expression evaluation.
- All primers were used with an annealing temperature of 60°C.

Number	Name	Fragment size (bp)	Sequence (5' to 3')
01	PKS_exon 1_F	326	TAATGCCAACTTTGTACAAAAAAGCAGGCTATGACTGTAT TAATCTTTCCAGATC
02	PKS_exon 1_R		TGGCTCTCAATTAACTGGATAAACGTAGCCAGTTGATACA CACAATG
03	PKS_exon 2_F	319	TGGCTACGTTTATCCAGTTAATTGAGAGCCACCCAGATCG ATTCC
04	PKS_exon 2_R		TGATTTGGGAATATCCAGTTCGGATTGGATGCTGTCTAAG
05	PKS_exon 3_F	3444	ATCCAATCCGAACTGGATATTCCCAAATCAGGCAAAGTAT AC
06	PKS_exon 3_R		TATCGGCGTAGAGGGAAGAAGGACAAAGCCCAATGTCG TTGACC
07	PKS_exon 4_F	2406	TTGGGCTTTGTCCTTCTTCCCTCTACGCCGATATGGCCAT GAC
08	PKS_exon 4_R		AGGCGGCCATTTCGTGTAACTTTTCCCCATCCTTCATCATG GTG
09	PKS_exon 5_F	157	TGAAGGATGGGGAAAAGTTACACGAAATGGCCGCCTATT TG
10	PKS_native_ terminator_R		TAATGCCAACTTTGTACAAGAAAGCTGGGTAGTGAATTCT AACAATCTAGTTAAC
11	171943_F	gDNA 853 cDNA 853	CTCCTACCACCCCGAAAACC
12	171943_R		AAATCTCGATGAGCCAGGCC
13	171942_F	gDNA 209 cDNA 209	TCTCGACAACTGAACTCCCG
14	171942_R		CCAAGCTATCCTGTCCCAGTG
15	171941_F	gDNA 771 cDNA 771	CTCCCCAACTCTCCCTGAGA
16	171941_R		AACTGCCAAGGAATACCGCA
17	201406_F	gDNA 550 cDNA 474	GAACCCGCACCAATCCAAAC
18	201406_R		TTCAGCCCGCATGAAGTCAT

Number	Name	Fragment size (bp)	Sequence (5' to 3')
19	APR9_F	gDNA 929 cDNA 929	CCGTTTTAGGATCGTCGCCA
20	APR9_R		GGCACTGACCTTGGCTAGTT
21	APR8_F	gDNA 1023 cDNA 902	TCCACGCATCGTCTTTGGAA
22	APR8_R		ACCCGTCATCTAGTGGACCA
23	APR7_F	gDNA 1179 cDNA 1053	GGAGCGCGAGTAGTCTTCTC
24	APR7_R		AGTCGGCCTGGTTCTCATTG
25	APR5_F	gDNA 645 cDNA 477	ACCTCAGAAGCCCACCTTTG
26	APR5_R		CCCGCACTTCCCAAAAAGAC
27	APR6_F	gDNA 751 cDNA 751	CACACCTATCCGTACAGCCC
28	APR6_R		AAGCCATCCAGCATCCTCAC
29	193045_F	gDNA 906 cDNA 773	TGACCAAGCAGACCATCGTC
30	193045_R		AGCCATGATGCCCTTGAACA
31	APR3_F	gDNA 1076 cDNA 1011	CATACGCTGTCTCCGCGAT
32	APR3_R		CGTCCAGATCAACCACACCA
33	APR2_F	gDNA 1224 cDNA 1169	CGCGACTCTCAAGGATTCCA
34	APR2_R		CCCGAGAGAATATCCGTGCC
35	APR1_F	gDNA 1029 cDNA 1029	GCCGAGGTTTGAAGAGGACA
36	APR1_R		AGCAAATGCACGAACACCAC

Supplementary Table S02.

Oligonucleotides used in this study **(continued)**.

Number	Name	Fragment size (bp)	Sequence (5' to 3')
37	171930_F	gDNA 418 cDNA 310	ATGGTACTAAAAAATATTACCGA
38	171930_R		GGTTCTGTTCAATAGCTGC
39	APR4_F	gDNA 1001 cDNA 859	ACCAACAACAGCTCCCACAA
40	ARR4_R		ATGGGCTGGACGTTGTTCAT
41	81192_F	gDNA 1020 cDNA 970	CAGGCTGCGTGGATTAAGGA
42	81192_R		GATCAAGCGCATGTTGGACC
43	171927_F	gDNA 1082 cDNA 953	GGTCGCAGCCGGTGATTATA
44	171927_R		CTGGCGATGTCAAAGACCCT
45	81165_F	gDNA 1130 cDNA 1130	TTGCTCGCATCCCTGAATGT
46	81165_R		AGTTTGCCCTGGAGCTTCTC
47	81156_F	gDNA 994 cDNA 926	CGTTTCGGCTTCAGTTTGCA
48	81156_R		CCGATAAGGTCCAGCTCGTC
49	A.parvulus H2B_F	gDNA 529 cDNA 423	ATGCCTCCCAAAGCCGCTGAG
50	A.parvulus H2B_R		CTATTTGGCAGAGGAGGAGTAC
51	A.oryzae H2B_F	gDNA 532 cDNA 381	GCTGCTGCCTCTGGTGAC
52	A.oryzae H2B_R		GTGCCTTCCGACACAGCATGC
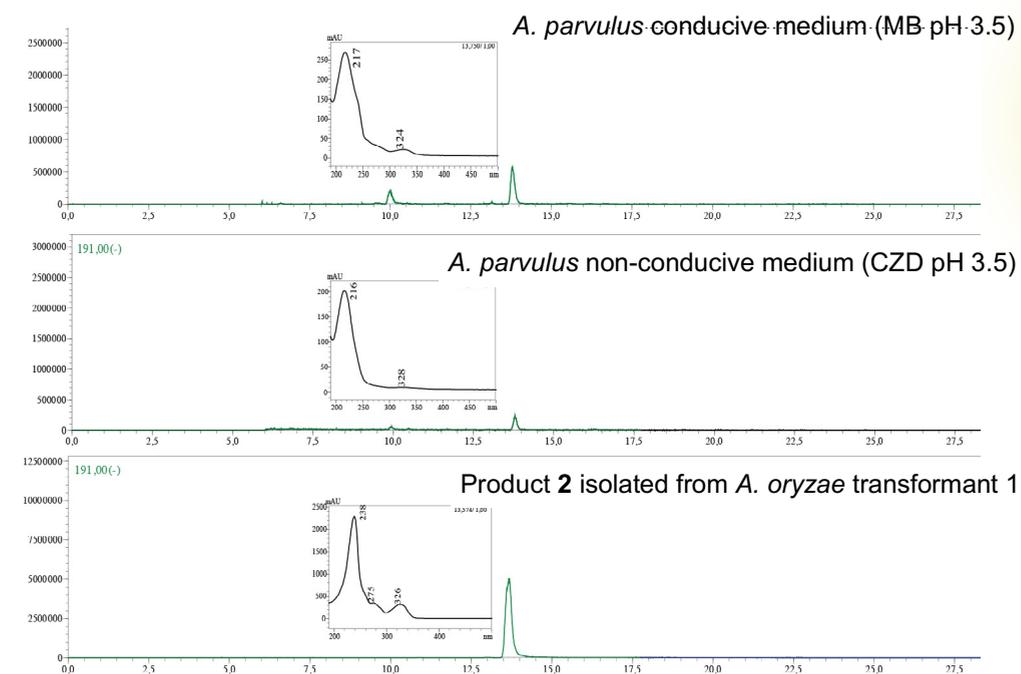
Supplementary Figure S03.

Search for product 2 in organic extracts from *Aspergillus parvulus*. Extracted mass of product 2 (191 in negative mode) in extracts from conducive and non-conducive conditions, with UV spectra. Purified product 2 produced by *Aspergillus oryzae* transformant is showed as reference.

4. Functional elucidation of the naphthalenone biosynthetic pathway of *Aspergillus parvulus*

Olga V. Mosunova[01]

Janieke Klusener[01]

Jérôme Collemare[01]

[01] Westerdijk Fungal Biodiversity Institute, Utrecht, The Netherlands

Abstract

Fungi are able to produce a diverse array of polyketides, including naphthalenones. In this chapter we continue the investigation of the naphthalenone pathway identified in *Aspergillus parvulus*. We functionally verify proposed steps of the pathway by heterolgous expression in Aspergillus oryzae. Our findings confirm previously suggested function of a FAD-oxidoreductase Apr2 and unravel stereoselectivity of this enzyme. Additionally, we identify a novel previously undocumented naphthalenone derivative, and discuss possible mechanisms leading to its occurrence. Based on the obtained data on chemical intermediates obtained by coexpression of several combination of tailoring genes and elucidation of chemical structures of respective chemicals, we propose a more comprehensive biosynthetic pathway for synthesizing asparvenone, parvulenone, botryosphaerone, and their derivatives.

Introduction

Naphthalenones are a group of polyketides that derive their chemical structure from naphthalene, an organic molecule comprised of two fused benzene rings, and a ketone functional group. Naphthalenones are known to be produced by plants [01, 02] and fungi [03], they often have biological properties that can be of interest for pharmaceutical (4-hydroxyvermelone, [04]), agricultural chemistry [05] or other fields of biotechnology [03]. Approximately 159 naphthalenones have been identified and reported in the literature for fungi [03], and in a recent study Mosunova with colleagues [06] have identified a biosynthetic gene cluster (BGC) in *Aspergillus parvulus* which was linked to the production of naphthalenone compounds.

Expression of genes of the identified BGC was correlated with the presence of naphthalenones 6-methylasparvenone and 1- ethylparvulenone in a liquid culture of *A. parvulus*. The core non-reducing polyketide synthase (nrPKS) gene *APR1* was heterologously expressed in *Aspergillus oryzae* and the transformants produced acetyltetrahydroxynaphthalene (AT4HN), a polyketide that is known to be related to biosynthesis of dihydroxynaphthalene (DHN) melanin in some fungi [07]. Thus, AT4HN serves as a backbone for naphthalenones produced by *A. parvulus*. Predictions were made about the *APR1* biosynthetic pathway and putative tailoring genes, but detailed elucidation of this naphthalenone pathway was not conducted.

Mosunova with coworkers [06] demonstrated that the putative naphthalenone BGC in *A. parvulus* is comprised of 8 tailoring genes, of which two (*APR5* and *APR6*) are homologs of genes that are present in pathways for DHN melanin and anthraquinones, namely scytalone reductase and T4HN reductase genes [08, 09]. It was previously reported that scytalone reductase and tetrahydroxynaphthalene reductase catalyze the removal of two or one hydroxyl groups from T4HN or emodin hydroquinone, respectively [10]. It is sensible to suggest that Apr6 and Apr5 could catalyze the conversion of AT4HN to acetylscytalone and acetylvermelone, making these reactions early steps in the naphthalenone pathway in *A. parvulus* [Fig. 01]. However, to the date there is no direct evidence that THNr and SCD1 enzymes or their homologs could act on AT4HN moiety directly.

Orthologs of *APR1* of *A. parvulus* were identified in a number of ascomycetous fungi [06]. This indicates that these fungi possess the capability to produce compounds belonging to the naphthalenones class, and suggests the existence of a Gene Cluster Family (GCF) for naphthalenones. Among these fungi *Neofusicoccum parvum* is reported to be producing naphthalenones known as botryosphaerones A-D [11]. The structure of botryosphaerones

4

bears similarity to that of 6- methylasparvenone **(Fig. 01)**. Specifically, the structure of botryosphaerones suggests the preservation of meta-hydroxyl groups of the AT4HN backbone, consistent with the absence of Apr5 and Apr6 orthologues in *N. parvum* **(06)**. Both botryosphaerones and asparvenone-parvulenone-like compounds undergo introduction of *para*-hydroxygroup to the A ring, which at that point becomes non-aromatic. Among the common enzymes between *A. parvulus* and *N. parvum* pathways, the most likely candidate responsible for this structural modification is the FAD-oxidoreductase APR2 and its homolog in *N. parvum*. Notably, the pathway of *A. parvulus* contains a second FAD-binding oxidoreductase **(06)**, and it remains unclear whether these two enzymes are redundant or fulfil distinct roles within the pathway. Given that the entire set of tailoring genes of *N. parvum* pathway has homologs in *A. parvulus* pathway, it is possible to reconstruct the
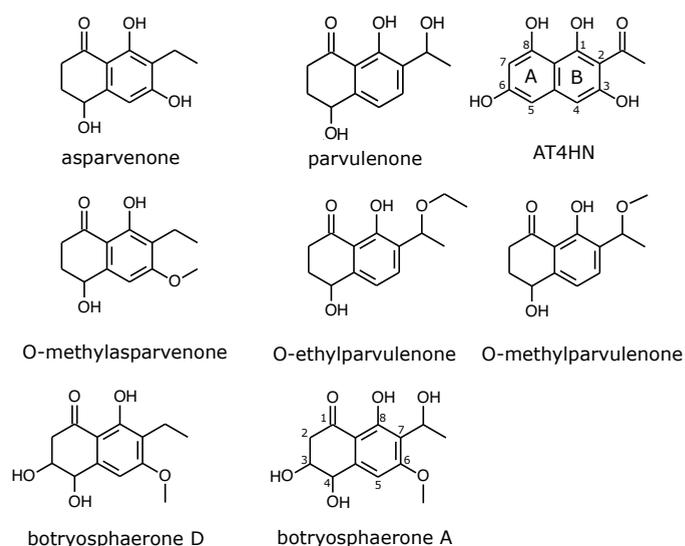
Figure 01.
Structures of naphthalenones produced by *Aspergillus parvulus* and *Neofusicoccum parvum*.

botryosphaerones pathway by expressing certain set of genes from the 6-methylasparvenone pathway in *A. oryzae*.

The O-acetyltransferase Apr3 and O-methyltransferase Apr4 are likely to target oxygen atoms of a hydroxyl group **(06)**. Structure of 6-methylasparvenone suggests that Apr4 acts on metahydroxyl of a right B ring and introduces methyl group in that position **(Fig.01)**. It is not clear from an enzymatic standpoint how the keto group of AT4HN is derivatized as in 6-methylasparvenone and 1-ethylparvulenone structures, but it seems that action of aldo-ketoreductase Apr8 and Apr3 is required to obtain the 1-ethylparvulenone structure.

The objective of this study was to functionally characterize specific steps in the naphthalenone pathway of *A. parvulus*. Different gene combinations were expressed heterologously in *A. oryzae* NSAR1, followed by verification of their expression and analysis of the resulting molecules using LC-MS and NMR techniques. This approach provided valuable insights into the pathway. Importantly, this methodology offers the potential to reconstruct pathways of other fungi in a heterologous host like *A. oryzae* using their available orthologs. This strategy can be used not only for elucidation of a single pathway, but for entire families of pathways, or Gene Cluster Families (GCFs).

Results

Expression of the entire set of genes of *A.parvulus* pathway yields 6-methylasparvenone-like molecule

In order to achieve expression of the entire set of genes of *A. parvulus*

naphthalenone pathway, we generated an *A. oryzae* NSAR1 strain expressing *APR1* and producing AT4HN, which we downstream refer to as a parental strain. The parental strain was generated using integration of the pTYGSmet plasmid carrying *APR1* into an undefined genomic locus of *A. oryzae*. The parental strain expressing *APR1* when cultivated with starch accumulates compound **1** identified as AT4HN **(Fig. 02A)** based on previously obtained data **(06) (Supplementary Figure 01)**.

Next, we performed integration of FAD-binding oxidoreductase *APR2*, O-acyltransferase *APR3*, O-methyltransferase *APR4* and GroES-like alcohol dehydrogenase *81192* cloned into pTYGSarg plasmid **(Supplementary Figure 02)**, and scytalone dehydratase *APR5*, T4HN reductase *APR6*, FAD-binding oxidoreductase *APR7*, aldo-keto reductase *APR8* cloned into pTYGSade backbone, into undefined loci in *A. oryzae* NSAR1 host strain and parental strain genomes. GroES-like alcohol dehydrogenase *81192* was not found to be co-regulated with the rest of the gene in the cluster, but it is conserved in pathways in the GCF **(06)**, and was therefore included in a panel of genes for expression of the entire pathway. Obtained fungal colonies were screened for the production of compounds different from AT4HN backbone, and four putative transformants were identified. Three of the four transformants (i, iii, and iv) had similar chemical profiles, producing compound **3** (RT = 5.95 min; UV max = 219 and 287 nm; major detected ionization peaks are 235 and 253 [TIC+] and 227, 251, 296 and 365 [TIC−]; exhibits a pale-yellow colour) and compound **4** (RT = 7.73 min; UV max = 196, 222, 286 nm; major detected ionization peaks are 235, 253 and 298 [TIC+]). Additionally, transformants (i) and (iii) produced seemingly lower amounts of compound **5** (RT = 3.23 min; UV max = 215 and 270 nm; no ionization observed when using ESIMS). Transformant (ii) produced seemingly more of the compound **5** and was also the only one producing compounds **1** and **2**.

A
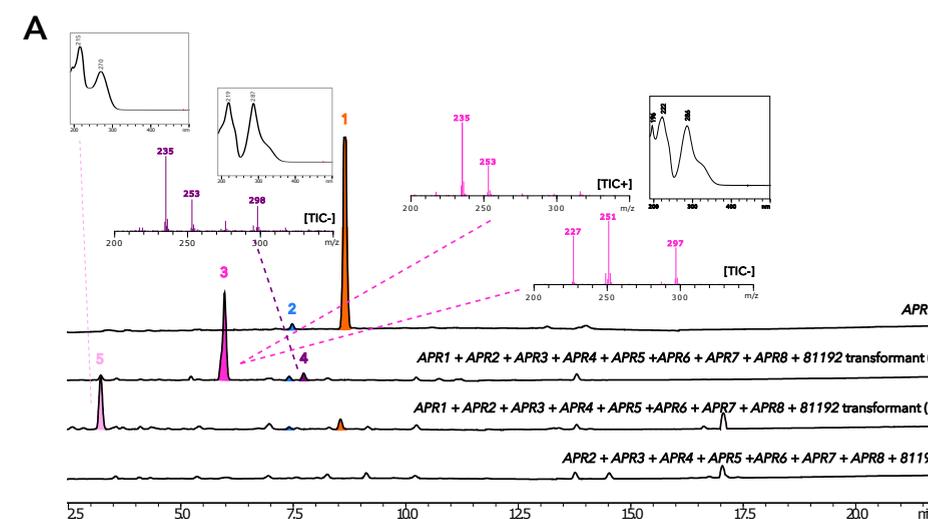
B

Transformant id/ Gene expression	i	ii	iii	iv
Apr1	Expressed	Expressed	Expressed	Expressed
Apr2	Expressed	Not expressed	Expressed	Expressed
Apr3	Expressed	Expressed	Expressed	Expressed
Apr4	Expressed	Expressed	Expressed	Expressed
Apr5	Expressed	Expressed	Expressed	Expressed
Apr6	Expressed	Expressed	Expressed	Expressed
Apr7	Not expressed	Not expressed	Not expressed	Not expressed
Apr8	Expressed	Not expressed	Expressed	Expressed
81192	Expressed	Expressed	Not expressed	Expressed
Compounds produced	3, 4, 5 (trace)	1, 2, 5	3, 4, 5 (trace)	3, 4

Figure 02.
HPLC and LC/MS profiles of extracts from cultural liquid of *A. oryzae* expressing entire naphthalenone pathway. UV profiles for individual peaks are shown. B - summary of the gene expression of the four transformants of *A. oryzae* carrying the full naphthalenone BGC.

4

From the gene expression profile [Fig. 02B] we deducted that transformants (i) and (iv) share the same expression profile, namely *APR1, APR2, APR3, APR4, APR5, APR6, APR8, 81192,* and both are producing compounds **3** and **4**. Transformant (iii) did not express *APR7,* and *81192,* but did express *APR1, APR2, APR3, APR4, APR5, APR6* and *APR8*. Transformant (ii), which produced **1**, **2**, and **5**, it seemed to not express *APR2* and *APR8*. Supposedly the proteins that resulted in the production of **3** and **4** could be attributed to either Apr2 or Apr8.

We attempted the chemical structure elucidation of compound **3** produced by transformant (i). The NMR data for compound **3** is represented on Fig. 3. Based on the NMR data, the identified compound is a chimera of 6-methylasparvenone and ethylparvulenone, therefore it was coined 1-ethyl,6-O-methylasparvulenone with a calculated mass of 280.320 [Fig. 03]. The compounds expected to be produced by expressing *APR1, APR2, APR3, APR4, APR5, APR6, APR7, APR8, 81192* in *A. oryzae* were asparvenone, parvulenone and/or their derivatives. We did not identify asparvenone or parvulenone in the extrolites of *A. oryzae*, and compound **3** is a new derivative showing structural features of both asparvenone and parvulenone. For the compound **3** we observe the addition of a hydroxyl group on C4 (presumably by Apr2), and the hydroxyl group on C9 became OCH_3 (predicted to be performed by Apr4), as expected. We previously hypothesized that 81192 performs removal of the hydroxygroup on C11, but in fact is seems like it does not bear this function, as both transformants i and iii produce identical compounds while *81192* is not expressed in transformant iii. The structure of the compound **3** confirms that the studies pathway was previously correctly assigned to the production of naphthalenones [06].

A

B

δ_H 600 MHz DMSO-d_6	δ_C 151 MHz DMSO-d_6
6.89 – 6.61 (m, 1H)	203.59
5.17 (q, J = 6.7 Hz, 1H)	163.31
5.01 (dp, J = 13.5, 6.1 Hz, 2H)	160.91
4.69 (dt, J = 11.7, 5.8 Hz, 1H)	149.64
3.95 – 3.66 (m, 3H)	118.06
2.74 – 2.58 (m, 2H)	109.57
2.22 – 2.10 (m, 1H)	101.14
1.98 – 1.85 (m, 1H)	66.40
1.40 (dd, J = 6.8, 3.3 Hz, 6H)	60.42
	56.30 – 55.21
	35.13
	31.79 – 31.26
	22.23
	19.76

Figure 03.

A - Putative structure of molecule 3 based on NMR signatures.

B - NMR peaks.

APR8 is unlikely to act on AT4HN backbone directly

Earlier, it was suggested that Apr8 plays a role in altering the acetyl group
of the AT4HN backbone [06]. Expression of the *APR8* aldo/ketoreductase
with *APR1* by a single identified transformant did not result in production
of new molecules [Supplementary Fig. 03, Table S01]. We did however observe a
slight elevation of the peak located right to **1**, but did not
pursue with further investigation, because this compound can be found in
lower amounts in chromatograms of extract from *A. oryzae* expressing
APR1 alone. It is plausible that Apr8 does not recognize AT4HN directly,
but instead acts on already modified molecule. This hypothesis is
supported by the fact that when *APR1, APR2, APR3, APR4, APR8, 81192* are
co-expressed in *A. oryzae,* HPLC chromatogram peaks that are observed
for *Apr1+APR8* expression are not present, suggesting they were fully
converted to unique peaks **20** and **21** [Supplementary Fig. 04]. Careful investigation
of these molecules and more transformants presenting same
profile of metabolites would provide information on possible activity of the Apr8.

Expression of *APR1, APR5, APR6* yields a range of ATHN-related molecules

Scytalone dehydratase and T4HN reductase (Apr5 and Apr6 homologs,
respectively) are involved in the first steps of the DHN melanin pathway,
namely they convert T4HN into T3HN by removing the *m*-hydroxy group. We
hypothesized that Apr5 and Apr6 perform a similar modification and act directly
on the AT4HN backbone. We integrated *APR5* and *APR6* in pTYGSade backbone
into a random locus in *A.oryzae* genome, and screened chemical profiles of the

obtained colonies. One transformant carried out expression of *APR1* and *APR5,
APR6,* and we make speculations based on the information that we obtained
from it. Compared to transformant expressing *APR1* only, this transformant
produces 6 novel compounds, some of which UV spectrum indicates a
relation to AT4HN [Fig. 04]. Compounds **6** (RT = 3.07 min; ionised at 316
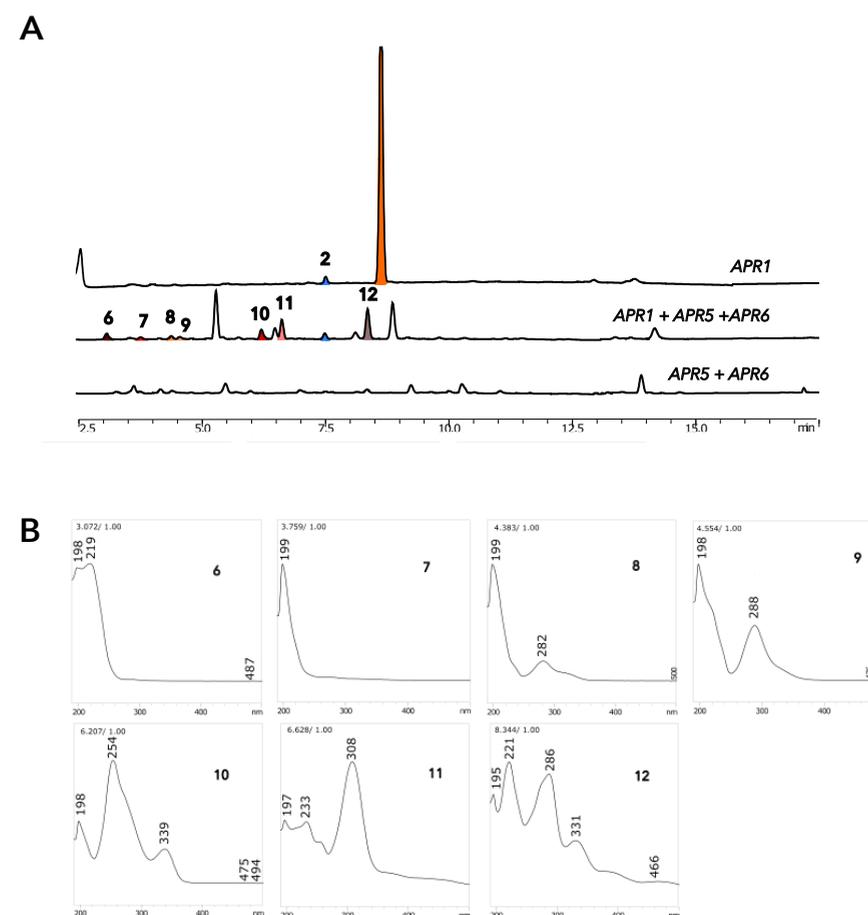
Figure 04.

A – HPLC chromatogram of extracts from *A. oryzae* NSAR1 expressing *APR1, APR5* and *APR6.*
B – UV profiles of identified peaks.

[TIC+]) and **7** (RT = 3.75 min; major detected ionization peaks are 227, 268 and 300 [TIC+]) only have a single UV absorption peak at about 200 nm.

Compounds **8** (RT = 4.38 min; ionised at 302 [TIC+]) and **9** (RT = 4.55 min; major detected ionization peaks are 252 and 357 [TIC+]) have an additional peak at 280 nm, generally following the UV profile of **3**, where there is a distinct gap between the two peaks. Compound **10** (RT = 6.20 min; ionised at 253 [TIC+] and strongly at 251 [TIC−]) has three peaks where the maximum peak is at 254 nm. Compound **11** (RT = 6.62 min; ionised at 248 and 311 [TIC+] and 246 [TIC−]) has one major peak at 308 nm and two smaller ones at 197 and 233 nm. The UV profile of **12** (RT = 8.34 min; major detected ionization peaks are 248 [TIC+] and 246 [TIC−]) has four sharper peaks (195, 221, 286, 331 nm) and a small peak at 466 nm.

Because expression of *APR1* together with *APR5* and *APR6* produced a range of molecules that were absent for the control strain that expressed *APR5* and *APR6* only, we suggest that Apr5 and Apr6 are able to act on AT4HN backbone directly. This evidence is supported by the fact that peak corresponding to **1** is almost entirely converted to other products, while **2** remains at a comparable level. Additionally, we speculate that other molecules produced by *A. oryzae* are likely not related to the conversions of AT4HN. In a course of the current study we did not obtain conclusive data to be able to resolve structures of compounds **6–12**.

Expression of *APR1, APR2, APR5, APR6* produced *trans-* and *cis*-(3R, 4R)-7-acetyl-3,4,6,8-tetrahydroxy-1,2,3,4-tetrahydronaphthalen-1-one

Previously we suggested that oxidoreductase Apr2 is involved in the
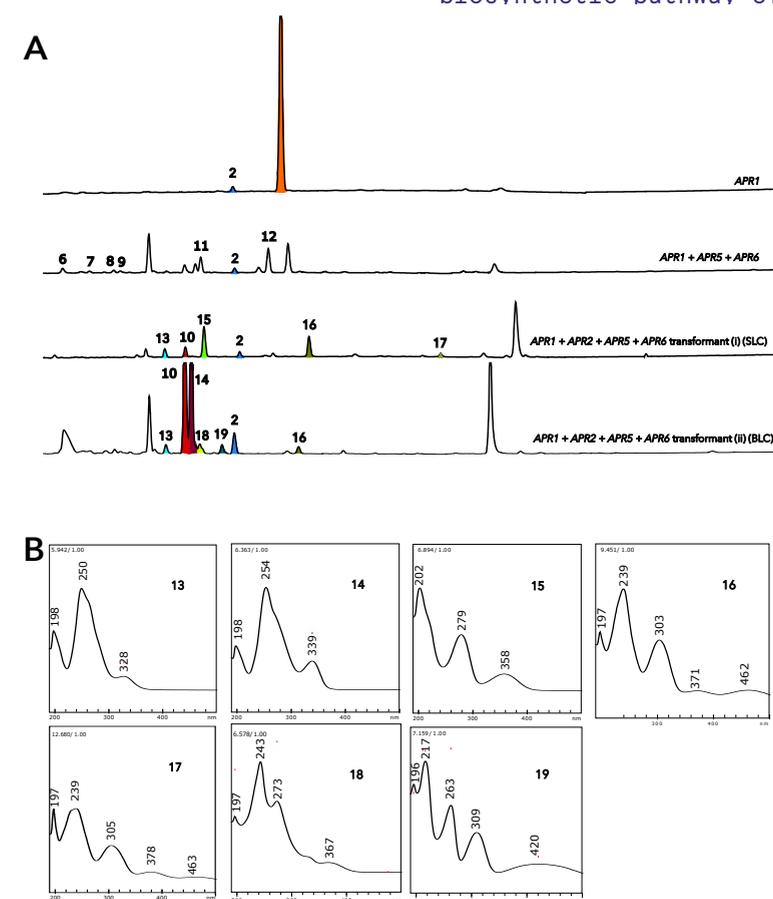
Gene expression	Transformant id		
	i	ii	iii
Apr1	Expressed	Expressed	Expressed
Apr2	Expressed	Expressed	Expressed
Apr5	Expressed	Expressed	Expressed
Apr6	Not expressed	Expressed	Not expressed
Compounds produced	2, 10, 13, 14, 15, 16, 17	2, 10, 13, 14, 15, 16, 18, 19	2, 10, 13, 14, 15, 16, 17

Figure 05.

A – HPLC profiles of transformants expressing *APR1, APR2, APR5,* and *APR1, APR2, APR5, APR6.*
SCL-small liquid culture, BLC- big liquid culture.

B – UV profiles of identified peaks.

Table S02.

NMR data for compound **10**

(3R,4R)-7-acetyl-3,4,6,8-tetrahydroxy-1,2,3,4-tetrahydronaphthalen-1-one.

δ_H 600 MHz DMSO-d$_6$	δ_C 151 MHz DMSO-d$_6$
6.64 (d, J = 1.2 Hz, 1H)	203.58
4.72 (dd, J = 2.8, 1.2 Hz, 1H)	202.16
4.18 (dt, J = 4.6, 2.9 Hz, 1H)	167.73
2.99 (dd, J = 17.4, 3.0 Hz, 1H)	165.41
2.69 (dd, J = 17.4, 4.6 Hz, 1H)	153.45
2.63 (s, 3H)	109.81
	108.21
	107.21
	69.46
	69.14
	43.90
	32.78

Table S02.

NMR data for compound **14**

(3R,4S)-7-acetyl-3,4,6,8-tetrahydroxy-1,2,3,4-tetrahydronaphthalen-1-one.

δ_H 600 MHz DMSO-d$_6$	δ_C 151 MHz DMSO-d$_6$
6.63 (s, 1H)	167.32
4.42 – 4.38 (m, 1H)	110.49
3.89 (ddd, J = 8.3, 7.1, 4.1 Hz, 1H)	108.40
2.93 – 2.87 (m, 1H)	107.87
2.69 – 2.63 (m, 1H)	71.39
2.61 (s, 3H)	69.22
	43.16
	32.77

incorporation of the hydroxygroup in para position of the A ring [06]. The transformant expressing *APR1*, *APR5* and *APR6* was used as a host strain for expressing *APR2* using the same methodology as above. We identified two transformants expressing *APR1*, *APR2*, *APR5*, and one transformant expressing all four genes [Fig. 05]. In comparison to the strain expressing *APR1*, *APR5*, *APR6*, peaks **1**, **6**, **7**, **8**, **9**, **12** were not present, and 5 new peaks were identified. We speculate that because compound **1** was not present in the HPLC trace, it was likely derivatized into the new peaks [Fig. 05A]. Compound **10** seemed to be present in both *APR1*, *APR5*, *APR6* strain and three newly obtained strains expressing *APR1*, *APR5*, *APR6* and *APR2*. New compounds were registered: **13** (RT = 5.64 min; UV max = 198, 250 and 328 nm; ionised at 221 [TIC−]), **14** (RT = 6.36 min; UV max = 198, 254 and 339 nm; ionised at 253 [TIC+] and 251 [TIC−]), **15** (RT = 6.69 min; UV max = 202, 279 and 358 nm; ionised at 237 [TIC−]), **16** (RT = 9.45 min; UV max = 197, 239, 303, 371 and 462 nm; ionised at 247 and 464 [TIC−]), and **17** (RT = 12.68 min; UV max = 197, 239, 305, 378 and 463 nm; ionised at 493 [TIC−]). Compounds **10** and **14** are identical in UV spectrum and mass but not in retention time, suggesting their high structural similarity.

Transformant (ii) expressing *APR1*, *APR2*, *APR5*, *APR6* was selected for further examination. We found that a volume of cultivation flask impacts HPLC profile of this transformant. We observed that in liquid culture with greater volume (BLC: 300 mL medium, 1.5 L Erlenmeyer flask) peaks **15** and **17** were not present anymore, and following new peaks were formed [Fig. 06B]: **18** (RT = 6.57 min; UV max = 197, 243, 273 and 367 nm; no ionization), and **19** (RT = 7.15 min; UV max = 196, 217, 263, 309 and 420 nm; ionised at 249 [TIC−]). The UV absorption peak pattern of **15**, **18**, and **19** are similar to that of AT4HN. UV profiles of **16** and **17** look similar and may be more related to **2**, due to the UV maxima following the 197, 239 nm pattern.

According to the UV-HPLC profile, compounds **10** and **14** seem to be

produced in equal amounts. When isolated, both **10** and **14** presented a red-brown colour. NMR data for compound **10** and **14** produced by transformant (ii) expressing *APR1, APR2, APR5, APR6*, presented in Tables 1 and 2, respectively. Both compounds showed similarities with the backbone of AT4HN [(Fig. 06)]. 13C-NMR revealed that a hydroxyl group was introduced in the *para*-position of the A ring. The orientation of this group is what distinguishes compound **10** from **14**, where for compound the C4- hydroxygroup 10 is in the *trans* position, while for **14** it is in the *cis* formation. Diastereoisomers **10** and **14** were identified as trans- and cis- (3R,4R(S))-7-acetyl-3,4,6,8-tetrahydroxy-1,2,3,4-tetrahydronaphthalen-1- one, respectively.

As **10** can also be found in the *APR1, APR5, APR6* expressing strain in trace amounts, which does not carry *APR2*, we speculate that the addition of the hydroxyl group occurred spontaneously. Apr6 is predicted to reduce ATHN into acetylscytalone, and Apr5 is thought to remove mhydroxygroup, both reactions are analogous to those of the melanin pathway [(12)]. For **14**, the addition of the hydroxyl group at C4 is likely to be enzymatically done by Apr2 after modification done by Apr6, but before Apr5, as *m*-hydroxygroup remains in the structure of **14**. The acetyl group of the backbone remained unmodified,

which indicates that the enzymes Apr2 and Apr6 can act on AT4HN directly. The fact that the A rings of **10** and **14** are saturated, should have been sufficient for Apr5 to perform the dehydration [(13)]. This, however, did not happen for compounds **10** and **14**, but might have occurred and resulted in compounds **13**, **18** and **19**.
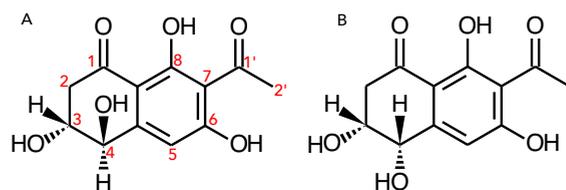
Figure 06.
Molecular structures of compound **10** (A) and **14** (B), with the corresponding positions of the carbons and hydrogens based on the NMR data. The orientation of the hydroxyl group at C4 results in the two different compounds.

Figure 07.

A - proposed pathway to asparvenone, parvulenone and their derivatives.

B - hypothetical pathway to botryosphaerones. C - hypothetical pathway to cristazarin (adapted from [15]). Molecules identified in the current study highlighted in green, reactions that are inferred form the melanin pathway highlighted in orange.

Discussion

Proposed pathway to naphthalenones

The assembly process of asparvulenone initiates with the production of AT4HN catalyzed by Apr1 [06]. Subsequently, a hydroxyl group is incorporated into the AT4HN structure, either spontaneously or facilitated by an enzyme. Spontaneous hydroxylation may occur when AT4HN hydroquinone hypothetically undergoes spontaneous oxidation, forming compound **10** [Fig. 08A]. This is supported by the fact that compound **10** is observed in extrolites of strains that do not express oxidase Apr2 that is responsible for enzymatic hydroxygroup incorporation. Asparvulenone pathway encodes for homologs of THNr and SCD1, enzymes that are found in melanin pathway and cladofulvin pathway. In melanin pathway, T4HN is converted into scytalone by THNr, and then to T3HN by SCD1 [12]. During this conversion, hydroxygroup of the A ring is removed. Structures of **14** suggest that Apr6 may have been able to act on AT4HN backbone directly, what was not reported for THNr before. We speculate that Apr6 and Apr5 are able to convert AT4HN to ADHN likewise THNr and SCD1 convert T4HN to DHN [8].

Secondly, AT4HN could be converted to **14** by Apr2 and Apr6 [Fig. 08A] similarly to the mode of reduction of hydroquinone and lawsone described by Saha with colleagues [14], where hydroquinone can tautomerize to 1,4 – diketo compound, which in turn can be reduced to cisketodiol by T4HNR. This reaction mechanism translates well to the possible asparvulenone pathway, where Apr2 first introduces hydroxygroup to C4 of the AT4HN, and resulting compound is serving as a substrtate for the Apr6 [Fig. 07A]. The opposite order of enzymatic actions is also possible, when Apr6 first performs reduction of A ring of AT4HN, and Apr2 introduces the hydroxy group to C4. The possibility of either option should be carefully investigated.
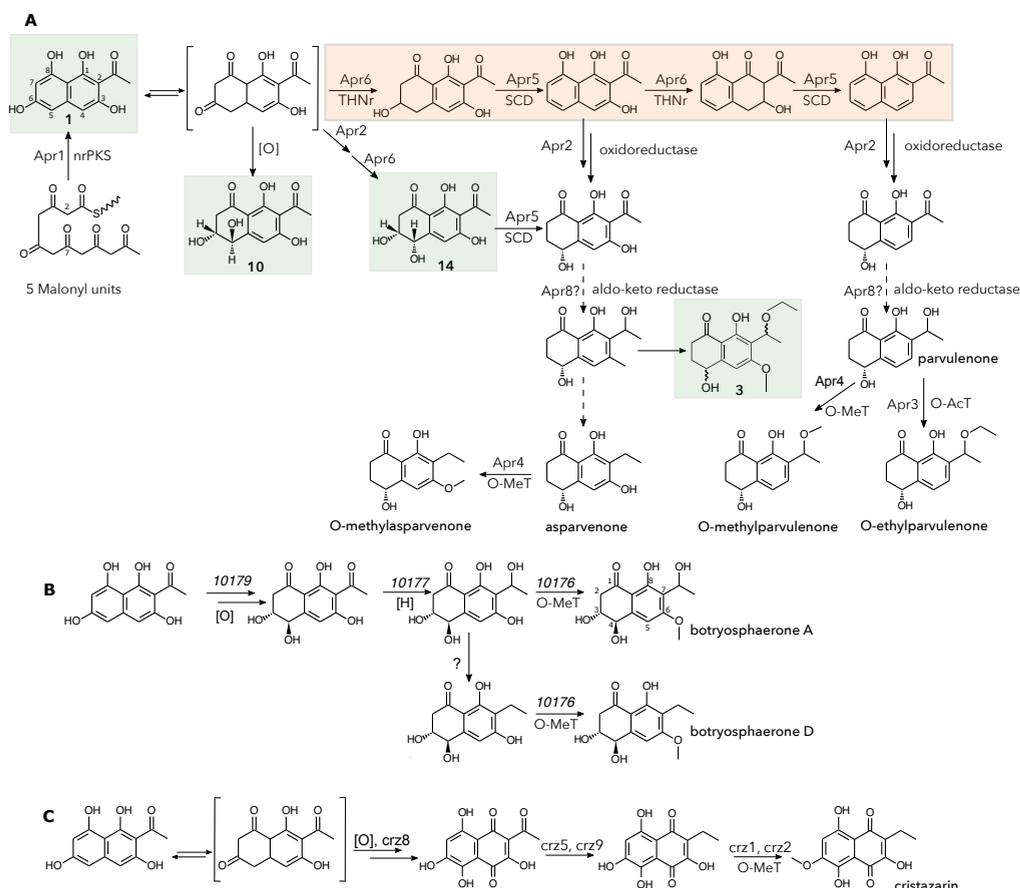
Molecule formed after saturation of A ring and hydroxylation could hypothetically serve as a substrate for Apr5, since its structure is not conflicting with required elements for the catalytic cycle of the SCD1 (13). SCD1 requires ketogroup in order to perform dehydroxylation [13], and we speculate that its homolog Apr5 can hypothetically act on both **10**, **14** or acetylscytalone. It is plausible that acetylscytalone could potentially be recognized by Apr2, eventually resulting into compound **14**, or modified by Apr5 into AT3HN and further, resembling conversion steps of DHN melanin pathway [Fig. 07A]. Structural elucidation of compounds **6**, **7**, **8**, **9**, **11**, **12**, was not performed here, but is valuable for establishing primary conversion steps of AT4HN in the asparvulenone pathway.

Previously we have identified a core nrPKS gene of a putative pathway for botryosphaerones [06], naphthalenone compounds that are produced by *N. parvum* [11], and discussed similarities to the asparvulenone pathway. We initially hypothesized that introduction of the *p*-hydroxygroup is likely to occur due to the spontaneous oxidation, or with help of Apr2 homolog 10179. Orthologs of Apr5 and Apr6 are absent in the genome of *N. parvum* [06], which suggests that Apr2 homolog 10179 does not require orthologs Apr6 and Apr5 for introduction of the hydroxygroup into the *p*-position. Structure of botryosphaerones features dearomatized A ring, suggesting that activity of Apr2 homolog 10179 alone could be sufficient to introduce -OH group and perform ring saturation simultaneously. Interestingly, orientation of the hydroxy group of compound **10** is the same as for botryosphaerones A and D [11], whereas compound **14** demonstrates thesame orientation of the hydroxy group as (+)-O-methylasparvenone [16]. The orientation of the -OH group is therefore the structural branchpoint to asparvulenone and botryopshaerones. To summarize, we hypothesize that stereochemistry of the hydroxygroup introduction may depend on action of two enzymes, Apr6 and Apr2, and

presence of both may be required to obtain **14**.

Based on asparvulenone assembly steps elucidated in the present study, we thereby suggest a pathway to botryosphaerones A and D [Fig. 07B]. Most probable first intermediate of botryosphaerones pathway after AT4HN is compound **10**. Next, a homolog of Apr8, aldo-ketoreductase 10177 is likely reducing ketogtoup to hydroxygroup, and with the help Omethyltransferase 10176 botryosphaerone A is obtained. Alternatively, if 10177 is capable of a complete removal of a ketogroup, Omethyltransferase 10176 is acting on a resulting molecule yielding botryosphaerone D.

Naphthalenones have been isolated from a number of fungal species, including lichens. Recent study discussed a relation between the asparvulenone pathway and the cristazarin pathway of *Cladonia metacorallifera* [15], which was not functionally verified yet. Cristazarin is structurally similar to asparvenone and exhibits antibacterial, antitumor, and anticancer activities, and is produced by *C. metacorallifera* mycobiont after 3 weeks of incubation using Lilly and Barnett!s (LB) medium supplemented with fructose as a carbon source, at 15°C and fluorescent light (6500 k, 18 wattages). Cristazarin belongs to naphthaquinones as its structure bears two carbonyl groups and has both of its rings aromatic, while asparvulenone is a naphtalenone with one saturated ring A. Paguirigan with colleagues [15] identified the nrPKS crz7, which is a paralog of *APR1*, and suggested plausible pathway to cristazarin based on domain architecture of coregulated tailoring genes around crz7 and pathway similarity to the asparvulenone pathway [Fig. 07 panels A and C]. The authors suggested that crz7 produces AT4HN backbone, that is later modified directly by oxidase crz8, and made parallel with reaction that is catalyzed by Apr2/Apr7 in asparvulenone pathway [15]. Although assembly of asparvulenone and cristazarin is similar, reciprocal BLAST analysis of crz genes versus *APR* genes performed by Paguirigan with colleagues [15], revealed that none of the *APR* tailoring genes are best

4

hits for crz genes, and the opposite is likewise. Oxidase Apr2 is a FAD-binding oxidorecutase (PF01565), and crz8 is a FAD-dependent oxidoreductase with PF01494 domain. Based on the data we obtained on the activity of Apr2, we could hypothesize that oxidases Apr2 or crz8 perform introduction of the *para*-hydroxygroup into the backbone of AT4HN utilizing distinct enzymatic mechanisms. It is of great interest to functionally verify activity of crz8 and compare its action to Apr2.

Expression of the entire pathway of *A. parvulus* in *A. oryzae* produces a new derivative

Expression of the entire set of genes of *A. parvulus* naphthalenone pathway was expected to result in molecules and their derivatives that are reported for the *A. parvulus* strain, namely asparvenone, parvulenone, and their derivatives 6-methylasparvenone and 1-ethylparvulenone. Surprisingly, simple expression of all genes of the cluster does not produce two distinct compounds as in *A. parvulus*, but a derivative with structural features characteristic for both asparvenone and parvulenone. This discrepancy suggests that in the natural setup of the parental strain, there might be a specific timing for gene expression enabling the production of these distinct molecules. Mosunova with colleagues [06] utilized end-point PCR at a single time point to assess gene expression, which lacked substantial evidence to confirm differential expression. Expression of the incomplete set of genes does not address why individual compounds that are normally produced by *A. parvulus* were not identified. It must be noted that since not all compounds were purified and structurally characterized, it is not possible to draw strong conclusions and outrule the absence of 6-methylasparvenone and 1-ethylparvulenone in the extrolites of transformants.

Genes of this pathway were expressed in *Aspergillus oryzae* under the control of strong constitutive promoters: *PgpdA*, *Peno*, and *Padh* [17, 18, 19]. The purpose of this strong expression was to ensure an adequate level of protein for the conversion of one substrate to another in the host cell. However, this high expression strength may not resemble the expression strength ratios that may be seen for native promoters. This imbalance could steer the pathway towards the complete conversion of intermediates into a new derivative end product and hinder the accumulation of intermediary products.

4

Materials and methods

- **Strains and culture conditions.** *A. parvulus* CBS 136.61 and *A. oryzae* NSAR1 strains were routinely cultured on malt extract agar plates (50 g/L, pH 5.4; Oxoid). *A. oryzae* NSAR1 transformants were sustained on selective Difco Czapek-Dox agar supplemented with specific nutrients based on the genetic construct: arginine (1 g/ L), adenine hemisulfate salt (0.5 g/L), methionine (1.5 g/L), and ammonium sulfate (1 g/L).

- **Nucleic acid extraction and RT-PCR.** Fungal mycelium from liquid cultures was filtered from cultural liquid, washed with nuclease-free water, frozen in liquid nitrogen, and ground using a mortar and pestle. Genomic DNA was isolated using the DNeasy plant minikit (Qiagen, Hilden, Germany) in line with the manufacturer's protocol. To extract total RNA, 100 mg of finely ground mycelium was combined with 1 mL of Invitrogen TRIzol reagent (Thermo Fisher Scientific, Waltham, MA) in a 1.5-mL microcentrifuge tube and incubated at 25°C for 5 minutes. Following this, the lysate was mixed with 0.2 mL of chloroform, gently hand-mixed, and incubated for an additional 5 minutes. After centrifugation at 12,000 g for 15 minutes at room temperature, the resulting aqueous phase was carefully transferred to a new microcentrifuge tube and mixed with 0.5 volumes of 100% ethanol. This mixture was then loaded into a column from the NucleoSpin RNA extraction kit (Macherey- Nagel, Allentown, PA), and subsequent steps were carried out according to the manufacturer's instructions. Finally, 500 nanograms of total RNA was utilized to synthesize cDNA using oligo(dT) primers and GoScript reverse transcription (RT) mix (Promega, Madison, WI) following the manufacturer's protocol.

- **Gene amplification and plasmid digestion.** Genes *Apr2-Apr8* and 81192 were amplified from cDNA of *A.parvulus* grown on liquid malt broth medium (filtered malt extract at 400 mL/L [pH 7.0]) at 25°C with constant agitation at 200 rpm. PCR amplification was as performed using Phusion® high-fidelity polymerase (Thermo Fisher Scientific, Waltham, MA) in accordance with manufacturer's protocol,extention time and primer sequences are specified in Supplementary Table S2.

- **Gene expression verification.** Gene expression of *Apr2-Apr8* and in *A. oryzae* transformants

was assessed using gene-specific primers for *A. parvulus* genes and *A. oryzae H2B* gene as control [Table S02]. Primers were used with GoTaq DNA polymerase (Promega, Madison, WI) in accordance with manufacturer's protocol, and DNA isolated from transformants was used as template DNA.

- **Transformation-associated recombination in *Saccharomyces cerevisiae*.** TAR in *S. cerevisiae* was used to assemble *APR2*, *APR3*, *APR4* into pTYGSade destination vector, and *APR5*, *APR6*, *APR7* into pTYGSade destination vector. A modified protocol from [20] was used. In brief, *S. cerevisiae* BMA 64 with ura3- auxotrophic marker was grown in 5 mL of YPD (Yeast extract - Difco 212759 at 8 g/L, Bacto peptone - Difco 211677 at 16 g/L, D(+)-Glucose - Merck 1.08337 at 16 g/L) overnight at 30°C 200 rpm. Then, two milliliters, containing approximately $1*10^8$ cells, were transferred to a 50 mL YPD solution in a 250 mL Erlenmeyer flask and incubated at 30°C with agitation at 200 rpm for about 5 hours. The yeast biomass was pelleted for 5 minutes at 2000 rpm at room temperature, washed with 10 mL of sterile water, pelleted once more, and reconstituted in 300 μL of sterile water. From this, 50 microliters of the resulting cell suspension were combined with 250 μL of 100mM DTT (dithiothreitol) and incubated for 10 minutes at room temperature. The yeasts were pelleted for 15 seconds at 12,000 g, and the supernatant was discarded. The pellet was reconstituted with 500 μL of PLTE solution (comprising 800 μL of 50% PEG 4000, 0.1 mL of 1 M LiAc, 20 μL of 50 mM EDTA, 10 μL of 1M TrisHCL at pH 7.5, and 70 μL of H_2O), 8-10 μL of DNA (plasmid and gene fragments in equal proportions), and 50 μL of recently boiled salmon sperm DNA (2 mg/mL). This mixture was incubated for 1 hour at 30°C without shaking. The samples were exposed to heat for 15 minutes at 45°C, spun for 15 seconds at 12,000 g, and then resuspended in 1 mL of YPD. After a 30-minute incubation at 30°C without shaking, the cells were pelleted for 15 seconds at 12,000 g, reconstituted in 200 μL of sterile water, and plated onto synthetic dropout media SDM (comprising 20 g/L agar, 20 g/ L D-glucose - Sigma-Aldrich, St. Louis, MO, 1.92 g/L yeast dropout supplements without uracil - Sigma-Aldrich, St. Louis, MO, and 6.7 g/L yeast nitrogen base without amino acids - Sigma-Aldrich, St. Louis, MO).

- **Seamless ligation cloning Extract (SLiCE) cloning in *E.coli* JM109.** SLiCE was used to

assemble *APR8* and *81192* into pEYA2 entry vector. SLiCE was performed in accordance with published protocol [21]. Briefly, cell extract of *E. coli* JM109 was prepared. First, 5 mL of *E. coli* JM109 was precultured at 37°C, 200 rpm overnight. Then, 1 mL of the overnight culture was transferred to 50 mL LB media in a 250 mL Erlenmeyer flask and grown at 37°C, 200 rpm until the OD600 reached approximately 2.0-3.0, taking around 6 hours. Following this, the culture was centrifuged at 5,000 *g* for 10 minutes at 4°C. The resulting pellet was washed with 50 ml of ice-cold sterile MiliQ water, then centrifuged again at 5,000 *g* for 10 minutes at 4°C. After discarding the supernatant, the pellet was resuspended in 1.2 ml of diluted and buffered Cell Lytic B cell lysis reagent and incubated for 10 minutes at room temperature. The mixture was centrifuged at 20,000 *g* for 2 minutes at 4°C. From this point, the downstream steps were carried out on ice. The supernatant was transferred to a 1.5 ml tube and mixed in a 1:1 ratio with 87% glycerol. Then, aliquots were prepared in 1.5 ml tubes. These tubes were snap-frozen in liquid nitrogen and stored at -80°C.

In a clean Eppendorf tube, 1 μL of SLICE 10x buffer (500 mM Tris-HCl pH 7.5, 100 mM MgCl$_2$, 10 mM ATP, 10 mM dithiothreitol (DTT)), was mixed with 1 μL of *E. coli* cell extract and DNA comprising 10-100 ng of linear vector (CIAP-treated) and 20-200 ng of fragments, and adjusted to reach a final volume of 10 μL with nuclease free water. The mixture was incubated at 37°C for 30 minutes and transformed in heat shock competent *E. coli* DH5α (Invitrogen) according to manufacturer's protocol.

Gateway cloning. Gateway® cloning (Thermo Fisher Scientific, Waltham, MA) was used to obtain expression vectors. Seventy nanograms of entry vector pEYA2_APR1, pEYA2_APR8 or pEYA2_81192 were mixed with 100 ng of destination vector *pTYGSmet*, *pTYGAade_Apr2_Apr3_Apr4* or pEYA2_*Apr5_Apr6_Apr7*, respectively. Vector mixture was combined with 1 μL of the Gateway LR Clonase II enzyme in 10 μL final volume Reaction mixture was incubate at 25oC overnight, followed by the inactivation procedure as per the manufacturer!s protocol, and the entire reaction was transformed into heat-shock competent *E. coli* DH5α.

Transformation of *A. oryzae* NSAR1. The transformation of *A. oryzae* NSAR1 was carried out following a previously documented procedure [06]. In this process, approximately 10 μg of vector was employed to transform around 1*10^7 protoplasts per mL. Each transformation reaction was distributed across four square plates (Greiner) and then incubated at 28°C for a period of 5-8 days until the transformants germinated.

Secondary metabolite extraction and HPLC-MS analyses. The collected cultural liquid was introduced into a flask separator, and an equivalent volume of ethyl acetate was introduced. The flask underwent repeated shaking and was subsequently placed on a stand to facilitate the separation of solvents for at least 1 hour. The resulting upper organic phase was then transferred to an evaporation flask. Ethyl acetate was evaporated using BUCHI Rotavapor R-100 connected to the BUCHI Heating Bath B-100 and VWR RC-10 Digital Chiller. The resulting pellet was reconstituted in 1-4 mL of acetonitrile, divided into aliquots in Eppendorf tubes, and stored at -80°C.

The resultant solid was dissolved in acetonitrile. Organic extracts underwent analysis using a Shimadzu LC-2030 3D-Prominence-i PDA system connected to a Shimadzu LCMS-2020 mass spectrometer, which was equipped with a Shimadzu Shim-pack GIST C18-HP reversed-phase column (3 mm, 4.6 by 100 mm). The employed method included a linear gradient of buffer B, ranging from 0 to 95% over 20 minutes, followed by 5 minutes at 95% buffer B, and concluding with 5 minutes at 100% buffer A.

Buffer A was composed of water with 0.1% trifluoroacetic acid (TFA) for high performance liquid chromatography (HPLC) or 0.05% formic acid for mass spectrometry (MS)-coupled analyses. Buffer B consisted of acetonitrile (LCMS grade) with 0.1% TFA for HPLC or 0.05% formic acid for MS-coupled analyses. The flow rate maintained was 1 mL/min for HPLC or 0.5 mL/min for MS-coupled analyses. Control and analysis of results were executed using the Shimadzu LabSolutions LCMS software, overseeing the equipment's operation.

Compound purification. Crude organic extracts underwent fractionation employing a Shimadzu preparative HPLC system. The setup included a CBM-20A controller, an LC-20AP pump, an SPD-20A detector, and an FRC-10A fraction collector, all integrated with a C18 reversed-phase Reprosil column (10 mm, 120 Å, 250 by 22 mm). The operation of the

4

system was managed through Shimadzu LabSolutions software.

A 12.5-milliliter flow rate was utilized, featuring a linear gradient progression of buffer B from 0 to 95%. This was followed by 5 minutes at 95% buffer B and then a subsequent 5-minute phase with 100% buffer A.

- **NMR.** Samples reconstituted in dimethyl sulfoxide (dDMSO) were subjected to ^{1}H NMR analysis at 600 MHz and ^{13}C NMR analysis at 151 MHz using a Bruker 600 spectrometer. The obtained data were analyzed utilizing MNOVA software.

For proton NMR, chemical shifts are reported in parts per million (ppm) downfield from tetramethylsilane and referenced to the residual protium in the solvent (^{1}H NMR, DMSO-d6 at 2.50 ppm). Meanwhile, carbon NMR chemical shifts are reported in parts per million downfield from tetramethylsilane and referenced to the carbon resonances of the residual solvent peak (^{13}C NMR, DMSO-d6 at 39.52 ± 0.06 ppm).

Acknowledgements

References

01. **Chen, G., Pi, X. M., & Yu, C. Y. (2015).** A new naphthalenone isolated from the green walnut husks of *Juglans mandshurica* Maxim. *Natural product research, 29(2),* 174–179. https://doi.org/10.1080/14786419.2014.971789

02. **Selvi S., Polat R., Çakilcioğlu U., Celep F., Dirmenci T., Ertuğ Z.F. (2022).** An ethnobotanical review on medicinal plants of the Lamiaceae family in Turkey. *Turkish Journal of Botany*: Vol. 46: No. 4, Article 1. https://doi.org/10.55730/1300-008X.2712

03. **Ibrahim, S. R. M., Fadil, S. A., Fadil, H. A., Eshmawi, B. A., Mohamed, S. G. A., & Mohamed, G. A. (2022).** Fungal Naphthalenones; Promising Metabolites for Drug Discovery: Structures, Biosynthesis, Sources, and Pharmacological Potential. Toxins, 14(2), 154. https://doi.org/10.3390/toxins14020154

04. **Nong, X. H., Zheng, Z. H., Zhang, X. Y., Lu, X. H., & Qi, S. H. (2013).** Polyketides from a marine-derived fungus *Xylariaceae* sp. *Marine drugs*, 11(5), 1718–1727. https://doi.org/10.3390/md11051718

05. **Xu, D., Xue, M., Shen, Z., Jia, X., Hou, X., Lai, D., & Zhou, L. (2021).** Phytotoxic Secondary Metabolites from Fungi. *Toxins, 13(4),* 261. https://doi.org/10.3390/toxins13040261

06. **Mosunova, O. V., Navarro-Muñoz, J. C., Haksar, D., van Neer, J., Hoeksma, J., den Hertog, J., & Collemare, J. (2022).** Evolution-Informed Discovery of the Naphthalenone Biosynthetic Pathway in Fungi. *mBio, 13(3), e0022322.* https://doi.org/10.1128/mbio.00223-22

07 . **Wheeler M., Abramczyk D, Puckhaber L, Naruse M, Ebizuka Y, Fujii I, Szaniszlo P. (2008).** New Biosynthetic Step in the Melanin Pathway of *Wangiella (Exophiala) dermatitidis*: Evidence for 2-Acetyl-1,3,6,8-Tetrahydroxynaphthalene as a Novel Precursor. *Eukaryotic Cell 7: 1699 - 1711.* https://doi.org/10.1128/ec.00179-08

08 . **Schumacher J. (2016).** DHN melanin biosynthesis in the plant pathogenic fungus *Botrytis cinerea* is based on two developmentally regulated key enzyme (PKS)-encoding genes. *Molecular microbiology, 99(4), 729–748.* https://doi.org/10.1111/mmi.13262

09 . **Andersson, A., Jordan, D., Schneider, G., & Lindqvist, Y. (1996).** Crystal structure of the ternary complex of 1,3,8-trihydroxynaphthalene reductase from *Magnaporthe grisea* with NADPH and an active-site inhibitor. *Structure (London, England : 1993), 4(10), 1161–1170.* https://doi.org/10.1016/s0969-2126(96)00124-4

10 . **Griffiths, S., Mesarich, C. H., Saccomanno, B., Vaisberg, A., De Wit, P. J., Cox, R., & Collemare, J. (2016).** Elucidation of cladofulvin biosynthesis reveals a cytochrome P450 monooxygenase required for anthraquinone dimerization. *Proceedings of the National Academy of Sciences of the United States of America, 113(25), 6851–6856.* https://doi.org/10.1073/pnas.1603528113

11 . **Burruano, S., Giambra, S., Mondello, V., Dellagreca, M., Basso, S., Tuzi, A., & Andolfi, A. (2016).** Naphthalenone polyketides produced by *Neofusicoccum parvum*, a fungus associated with grapevine *Botryosphaeria* dieback. *Phytopathologia Mediterranea, 55(2), 197–206.* http://www.jstor.org/stable/44809327

12 . **Gao, J., Wenderoth, M., Doppler, M., Schuhmacher, R., Marko, D., & Fischer, R. (2022).** Fungal Melanin Biosynthesis Pathway as Source for Fungal Toxins. *mBio, 13(3), e0021922.* https://doi.org/10.1128/mbio.00219-22

13 . **Zheng, Y. J., & Bruice, T. C. (1998).** Role of a critical water in scytalone dehydratasecatalyzed reaction. *Proceedings of the National Academy of Sciences of the United States of America, 95(8), 4158–4163.* https://doi.org/10.1073/pnas.95.8.4158

14 . **Saha, N., Müller, M., & Husain, S. M. (2019).** Asymmetric Synthesis of Natural *cis*-Dihydroarenediols Using Tetrahydroxynaphthalene Reductase and Its Biosynthetic Implications. *Organic letters, 21(7), 2204–2208.* https://doi.org/10.1021/acs.orglett.9b00500

15 . **Paguirigan, J. A. G., Kim, J. A., Hur, J. S., & Kim, W. (2023).** Identification of a biosynthetic gene cluster for a red pigment cristazarin produced by a lichen-forming fungus *Cladonia metacorallifera. PloS one, 18(6), e0287559.* https://doi.org/10.1371/journal.pone.0287559

16 . **Lafleur-Lambert, R., & Boukouvalas, J. (2016).** Asymmetric total synthesis of (+)-Omethylasparvenone, a rare nitrogen-free serotonin 2C receptor antagonist. *Organic & biomolecular chemistry, 14(37), 8758–8763.* https://doi.org/10.1039/c6ob01678b

17 . **Lu, Y., Zheng, X., Wang, Y., Zhang, L., Wang, L., Lei, Y., Zhang, T., Zheng, P., & Sun, J. (2022).** Evaluation of *Aspergillus niger* Six Constitutive Strong Promoters by Fluorescent-Auxotrophic Selection Coupled with Flow Cytometry: A Case for Citric Acid Production. *Journal of fungi (Basel, Switzerland), 8(6), 568.* https://doi.org/10.3390/jof8060568

18 . **Li, J., Wang, J., Wang, S. et al. (2012).** Achieving efficient protein expression in *Trichoderma reesei* by using strong constitutive promoters. *Microb Cell Fact 11, 84.* https://doi.org/10.1186/1475-2859-11-84

19 . **Fleißner, A., Dersch, P. (2010).** Expression and export: recombinant protein production systems for *Aspergillus . Appl Microbiol Biotechnol 87, 1255–1270.* https://doi.org/10.1007/s00253-010-2672-6

20 . **Tripp, J. D., Lilley, J. L., Wood, W. N., & Lewis, L. K. (2013).** Enhancement of plasmid DNA transformation efficiencies in early stationary-phase yeast cell cultures. *Yeast (Chichester, England), 30(5), 191–200.* https://doi.org/10.1002/yea.2951

21 . **Zhang, Y., Werling, U., & Edelmann, W. (2014).** Seamless Ligation Cloning Extract (SLiCE) cloning method. *Methods in molecular biology (Clifton, N.J.), 1116, 235–244.* https://doi.org/10.1007/978-1-62703-764-8_16

4

Supplementary materials

A

B

A

B

C

4
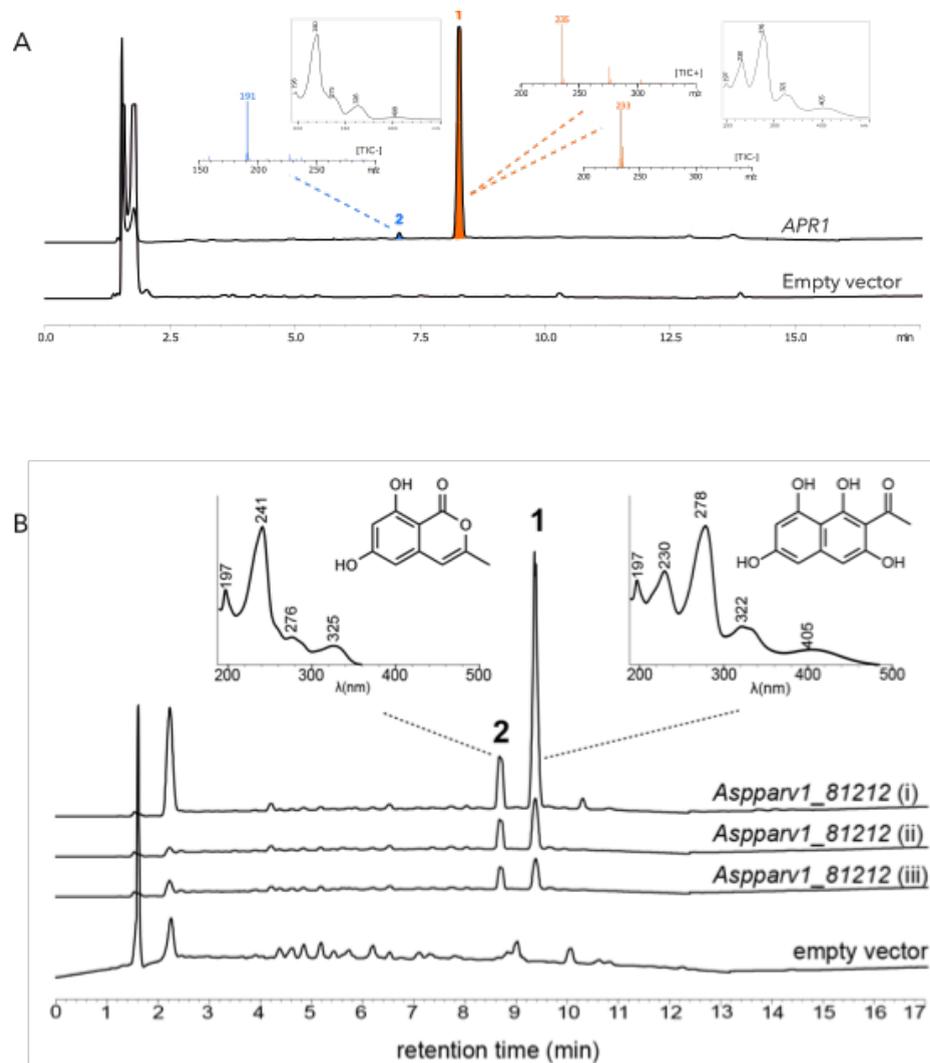
Supplementary Figure 01.

A – HPLC trace of an extract from 5-day old *A. oryzae* NSAR1 transformant expressing APR1
(pTYGSmet backbone).

B – Graph adapted from (6) depicting extracts of *A. oryzae* NSAR1 expressing APR1 (pTYGSarg
backbone).
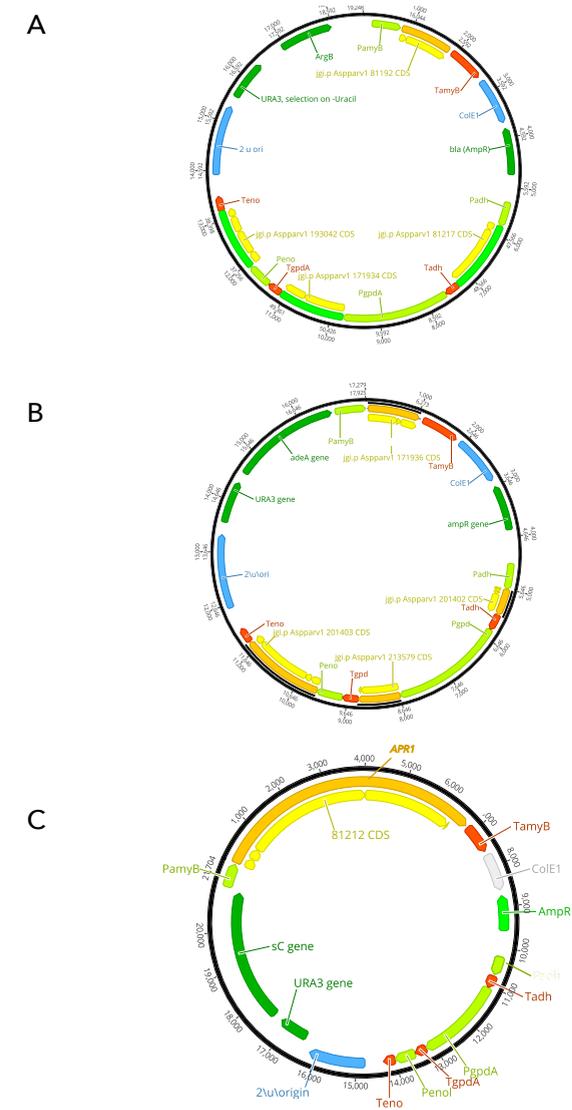
Supplementary Figure 02.

A – Plasmid map for *pTYGSarg* with *APR2, APR3, APR4* and *81192*.

B – Plasmid map for *pTYGSade* with *APR5, APR6, APR7, APR8*.

C – Plasmid map for *pTYGSsC* with *APR1*.
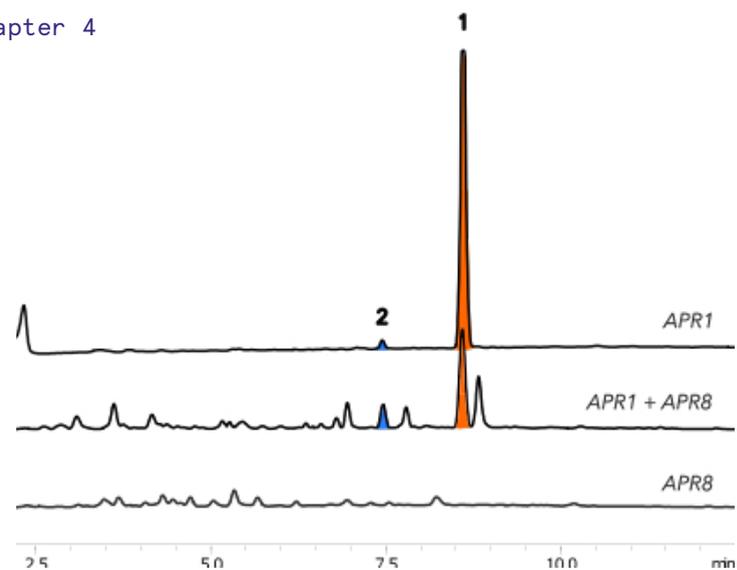
Supplementary Figure 03.

HPLC chromatogram of extracts from *A. oryzae* NSAR1 expressing *Apr1* and *Apr8*.
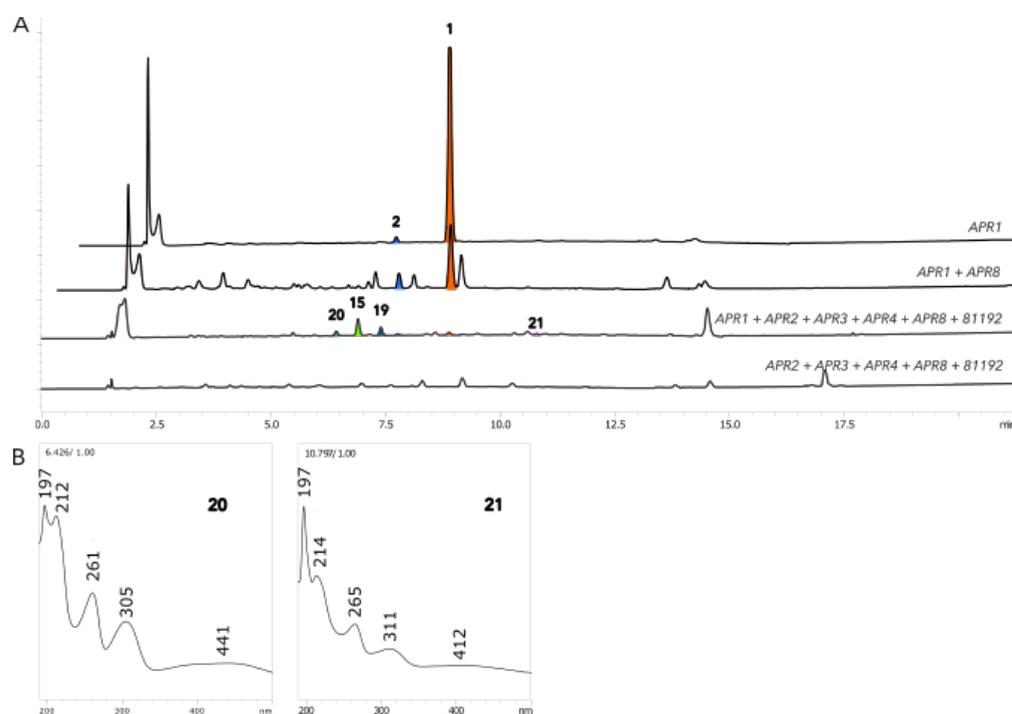
Supplementary Figure 04.

A – HPLC of extracts of *APR1, APR2, APR3, APR4, APR8, 81192* transformants. This gene combination represents pathway of *N. parvum* to botryosphaerones. Compared to APR1 expressed alone, two new peaks were identified, **20** and **21**.

B – UV spectra of both compounds.

Product 3
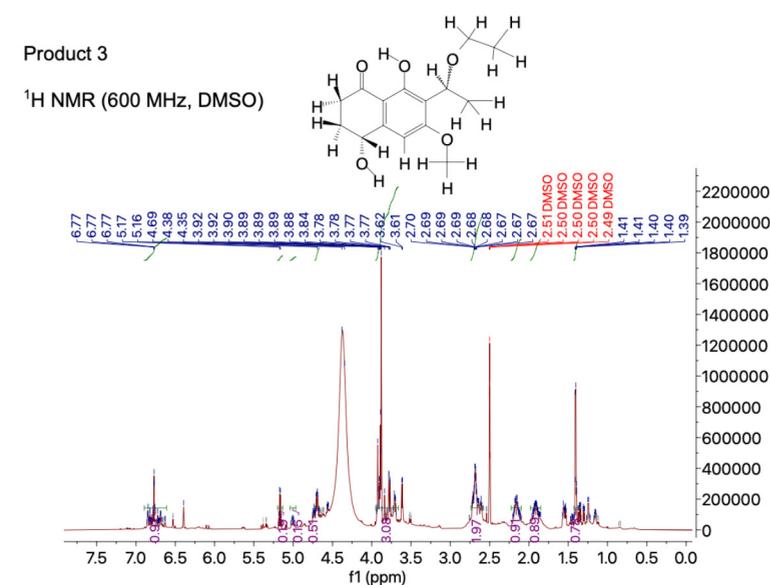
¹H NMR (600 MHz, DMSO)

¹H NMR (600 MHz, DMSO) δ 6.89 – 6.61 (m, 1H), 5.17 (q, *J* = 6.7 Hz, 1H), 5.01 (dp, *J* = 13.5, 6.1 Hz, 1H), 4.69 (dt, *J* = 11.7, 5.8 Hz, 1H), 3.95 – 3.66 (m, 3H), 2.74 – 2.58 (m, 2H), 2.22 – 2.10 (m, 1H), 1.98 – 1.85 (m, 1H), 1.40 (dd, *J* = 6.8, 3.3 Hz, 1H).
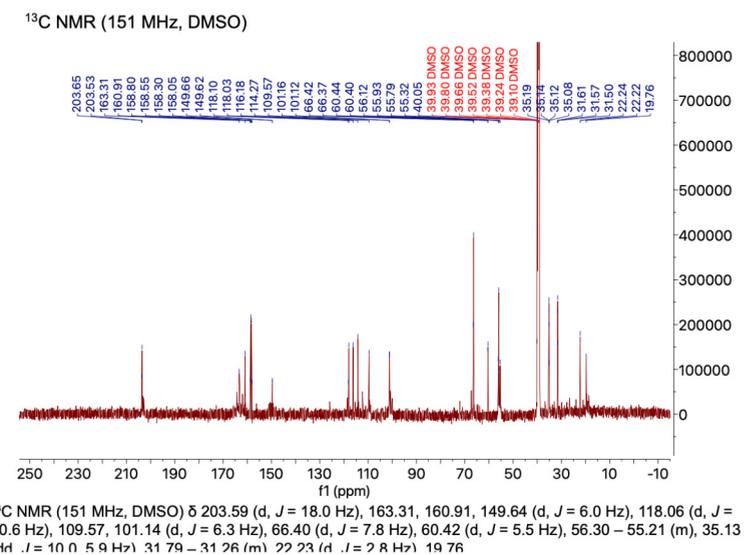
¹³C NMR (151 MHz, DMSO)

¹³C NMR (151 MHz, DMSO) δ 203.59 (d, *J* = 18.0 Hz), 163.31, 160.91, 149.64 (d, *J* = 6.0 Hz), 118.06 (d, *J* = 10.6 Hz), 109.57, 101.14 (d, *J* = 6.3 Hz), 66.40 (d, *J* = 7.8 Hz), 60.42 (d, *J* = 5.5 Hz), 56.30 – 55.21 (m), 35.13 (dd, *J* = 10.0, 5.9 Hz), 31.79 – 31.26 (m), 22.23 (d, *J* = 2.8 Hz), 19.76.

Supplementary Data Set S01.

¹H and ¹³C NMR spectra of products **3**, **10**, **14**.

Product 14: (3R,4S)-7-acetyl-3,4,6,8-tetrahydroxy-1,2,3,4-tetrahydronaphthalen-1-one

¹H NMR (600 MHz, DMSO)

¹H NMR (600 MHz, DMSO) δ 6.63 (s, 1H), 4.42 – 4.38 (m, 1H), 3.89 (ddd, *J* = 8.3, 7.1, 4.1 Hz, 1H), 2.93 – 2.87 (m, 1H), 2.69 – 2.63 (m, 1H), 2.61 (s, 3H).

¹³C NMR (151 MHz, DMSO)

¹³C NMR (151 MHz, DMSO) δ 167.32, 110.49, 108.40, 107.87, 71.39, 69.22, 43.16, 32.77.

Product 10: (3R,4R)-7-acetyl-3,4,6,8-tetrahydroxy-1,2,3,4-tetrahydronaphthalen-1-one

¹H NMR (600 MHz, DMSO)

intramolecular
hydrogen bond

¹H NMR (600 MHz, DMSO) δ 6.64 (d, *J* = 1.2 Hz, 1H), 4.72 (dd, *J* = 2.8, 1.2 Hz, 1H), 4.18 (dt, *J* = 4.6, 2.9 Hz, 1H), 2.99 (dd, *J* = 17.4, 3.0 Hz, 1H), 2.69 (dd, *J* = 17.4, 4.6 Hz, 1H), 2.63 (s, 3H).

¹³C NMR (151 MHz, DMSO)

¹³C NMR (151 MHz, DMSO) δ 203.58, 202.16, 167.73, 165.41, 153.45, 109.81, 108.21, 107.21, 69.46, 69.14, 43.90, 32.78.

Supplementary Data Set S01.

¹H and ¹³C NMR spectra of products **3**, **10**, **14 (continued)**.

Supplementary Table S01.

Genetic background of transformants and compounds identified in their cultural liquid.-1

Genotype	Genotype (Based on Expression)	Compounds identified
APR1	APR1	1, 2
	APR1+APR2+APR3+APR4+APR5+APR6+APR8	3, 4, 5
	APR1+ APR3+APR4+ APR5+APR6	1, 2, 5
APR1+APR2+APR3+APR4+APR5+APR6+APR8	APR1+APR2+APR3+APR4+APR5+APR6+APR8	3, 4, 5
	APR1+APR2+APR3+APR4+APR5+APR6+APR8	3, 4
	APR1+APR5+APR6	2, 6, 7, 8, 9, 10, 11, 12
	APR1+APR2+ APR5	2, 10, 13, 14, 15, 16, 17
APR1+APR2+APR5+ APR6	APR1+APR2+ APR5+APR6	2, 10, 13, 14, 15, 16
	APR1+APR2+APR5+ APR6	2, 10, 13, 14, 16, 18, 19
	APR1+APR2+ APR5	2, 10, 13, 14, 15, 16, 17
	APR1+APR8	1, 2
APR1+APB8+81192	APR1+APR8+81192	1, 2
	APR1+APR8+81192	1, 2
	APR1+APR8+81192	1, 2
	APR1+APR2+APR3 +APR8+81192	1, 2, 15, 19, 20, 21
APR1+APR2+APR3+APR4+APR8+81192	APR1+APR2+APR3+APR4+APR8+81192	1, 2, 15, 19, 20, 21
	APR1+APR2+APR3+APR4+APR8+81192	1, 2, 15, 19, 20, 21
	APR1+APR2 +81192	2, 13, 15, 16
APR1+APR2+APR5+APR6+APR8+81192	APR1+APR2 +APR8+81192	2, 15, 16, 19, 20, 21
	APR1+APR2+ APR5 +81192	2, 13, 14, 15, 16, 22

Supplementary Table S02.

Primers used in this study.

(Primers used for screening of gene expression – see Chapter 3.)

Primers used for amplification of genes from cDNA for downstream cloning

Number	Name	Sequence (5' to 3')
01	Apr2 F	AGATCCCAAAGTCAAATGATGCCTGTAAGGGATGACCATTCCCACCGATCCGTCCG
02	Apr2 R	TGTCGAAAGATCCACTAGAGTAAATCTGGGCTACCGGTGCAAAGCATACGATTTGTC
03	Apr3 F	TGAGCAGACATCACCGTAAGAGGTATATGGATGGCAATCCCCCTGTCTGAAGTC
04	Apr3 R	ACACCTACAAACACACACATATATACATAATCACACGGAATAATTCTGTGTAACGTAGCG
05	Apr4 F	CGACTGACCAATTCCGCAGCTCGTCAAAGGATGGCTGCCTTGACTGATCTTGCGGCTAT
06	Apr4 R	ATCCATATACTGTCAGTTTCCTCAACCCTCTTAGATATCCCAAACTACTTCAATAATCChnf
07	Apr5 F	TCTTTCAACACAAGATCCCAAAGTCAAAGGATGGCATCACCTCAGAAGCCCACCTTTGAA
08	Apr5 R	TTCATTCTATGCGTTATGAACATGTTCCCTGGCTACTCTTCATCATCAAATTTGCCCGCA
09	Apr6 F	CAGCTACCCCGCTTGAGCAGACATCACCGGATGTCCACACCTATCCGTACAGCCCGTCTC
10	Apr6 R	ACGACAATGTCCATATCATCAATCATGACCGGTTATGCAGACGCTCCACCAGAGATG
11	Apr7 F	CGACTGACCAATTCCGCAGCTCGTCAAAGGATGTTGTTCATTTCTCTGACCACGGCTCTT
12	Apr7 R	TTGGCTGGTAGACGTCATATAATCATACGGTTAAGTTGCAGGGTTGATTGAGAATGGGTT
13	Apr8 F	CCAACTTTGTACAAAAAAGCAGGCTCCGCATGTCTCCTCCAGAGCATCCACGCATCGTC
14	Apr8 R	AAGCTGGGTCGGCGCGCCTGTTTAAACTGCTCAGAGATGGTACGGATCAGCATCGTCTTT
15	81192 F	GCCAACTTTGTACAAAAAAGCAGGCTCCGCATGACCAACCAGGCTGCGTGGATTAAGGAG
16	81192 R	AAGCTGGGTCGGCGCGCCTGTTTAAACTGCCTATTCAACCCATGGGTTCAGGACTAGCTT

4

5. Evolution-informed discovery of a novel antibiotic compound from *Aspergillus melleus*

Olga V. Mosunova[01]

Ella Schunselaar[01]

Elske J. Dwars[01]

Jorge C. Navarro-Muñoz[01]

Jelmer Hoeksma[02,03]

Ronnie J.M. Lubbers[02]

Jeroen den Hertog[03,04]

Jérôme Collemare[01]

[01] Westerdijk Fungal Biodiversity Institute, Utrecht, The Netherlands

[02] Hubrecht Institute-KNAW, Utrecht, The Netherlands

[03] University Medical Center Utrecht, Utrecht, The Netherlands

[04] Institute Biology Leiden, Leiden University, Leiden, The Netherlands

Abstract

In this chapter we continued exploring the chemical space of non-reduced polyketides in a different subset of genomes of Lecanoromycetes. We utilized phylogenetic dereplication in order to identify standalone nrPKSs and identified three phylogenetically distant nrPKSs. Following this, we conducted heterologous expression experiments on a pathway from *Aspergillus melleus* containing a closely related nrPKS to those identified in lichen-forming fungi. Our investigations linked the nrPKS AmoA to the synthesis of orcinol and orsellinic acid. Through heterologous expression of the entire pathway, we obtained previously unreported compound exhibiting intriguing antibiotic activity agains range of gram-negative bacteria and demelanizing properties when assayed using zebrafish embryos. Finally, based on obtained findings we proposed a potential biosynthetic pathway.

Introduction

Lichenizing fungi are known to produce diverse compounds with orsellinic acid (OA) as a core structure. The presence of specific orsellinates, along with other secondary metabolites, is often used for chemotaxonomic purposes [01]. OA itself is a precursor of other industrially valuable monoaromatic orsellinates such as orcinol, resorcinol and 3,5-dimethylorcinol, which exhibits rose petal smell. Compounds like lecanoric acid, usnic acid, evernic acid, atranorin can be found in cortex/medulla of lichen body, and belong to the polyaromatic orsellinates. They encompass depsides and depsidones, and typically consist of two or more monoaromatic units linked by an ester group (depsides) or ester and ether bonds (depsidones). Depsides are typically formed through the esterification of orsellinic acid with another aromatic compound, often a phenolic acid (e.g. lecanoric acid) or another orsellinic acid (e.g. atranorin) molecule [02]. Depsides are commonly found to be produced by Lecanoromycetes, although they are not restricted to lichenizing fungi and are also found to be produced by other fungal species [03]. Depsides are precursors for the biosynthesis of depsidones through oxidative cyclization [04, 05, 06]. Depsidones typically feature two aromatic rings joined together by ester and ether linkages, exemplified by compounds like physodic acid, norstictic acid and evernic acid [05]. Orsellinates derived from lichens often demonstrate antimicrobial properties, and their structural diversity contributes to variations in biological activity. Dibenzofuran usnic acid exhibits antimicrobial, antiprotozoal and antiviral properties, among others [07]. Depsidone alpha-alectoronic acid demonstrates activity against B16 mouse melanoma cells [08]. Depsidone lobaric acid is active against methicillin-resistant clinical isolates of *Staphylococcus aureus* (MIC90=64 μg/mL) [09]. Depside lecanoric acid arrests cell cycle in colon cancer cells [10] and demonstrates antifungal activity [11].

Despite potential interest for usage of OA derivatives as medicinal compounds, limited growth of lichenizing fungi in nature and in culture hampers investigation of orsellinates and their biosynthetic patwhays using methods involving gene manipulation. Several biosynthetic pathways to orsellinates have been functionally characterized in other fungal species. Notable examples include lecanoric acid pathway that was expressed in *Saccharomyces cerevisiae* [12], as well as characterization of OA pathway of *Aspergillus nidulans* [13] and *Aspergillus niger* [14], PKS7 in *Claviceps purpurea* [15], PKS14 of *Fusarium graminearum* [16].

Isolation of metabolites from lichens allows identifying major compounds deposited in cortex or medulla of a lichen body, but extraction

5

and identification of possible intermediates that lead to the production of final products is not facilitated by direct metabolite extraction. With the rising availability of fungal genomes, including those of lichen mycobionts [17, 18, 19, 20, 21, 22, 23, 24, 25], analyses of BGCs with core polyketide synthase genes from Lecanoromycetes have been conducted [26, 27, 28].

Several studies perform functional characterization of unassigned pathways from Lecanoromycete strains. Kim with colleagues [29] aim at identification of the atranorin biosynthetic pathway. They functionally characterized the atranorin pathway encoded in *Stereocaulon alpinum*, with its PKS falling into Group IX. They performed a phylogenetic analysis of nrPKSs derived from 30 genomes of lichen forming fungi (LFF) with respect to characterized nrPKSs, and reported 9 phylogenetic clades, or groups. At the same time, Kealey with colleagues [30] reports functional assignment of *Pseudevernia furfuracea* PKS PFUR17_02294 to lecanoric acid, and the nrPKS protein phylogenetically belongs to Group X as we later found [31]. Last, Mosunova with colleagues [31] identified another standing out phylogenetic clade, group XI using an evolution-guided approach. Newly identified group XI comprised three PKSs derived from genomes of LFF, and their ortholog from *Aspergillus parvulus* was functionally characterized by heterologous expression in *Aspergillus oryzae*. Lichen PKSs of group XI were then assigned to the production of naphthalenones.

In this study we explore the diversity of nrPKSs in 11 lichen genomes, identify new standalone phylogenetic clade and assign it to the new phylogenetic Group XII. We identify Gene Cluster Family (GCF) consisting of BGCs conserved in fungal species across Ascomycota, and heterologously express a pathway from *Aspergillus melleus* that contains nrPKS paralogue to lichen nrPKSs of Group XII. We characterize selected OA derivatives with antibiotic and demelanizing activity.
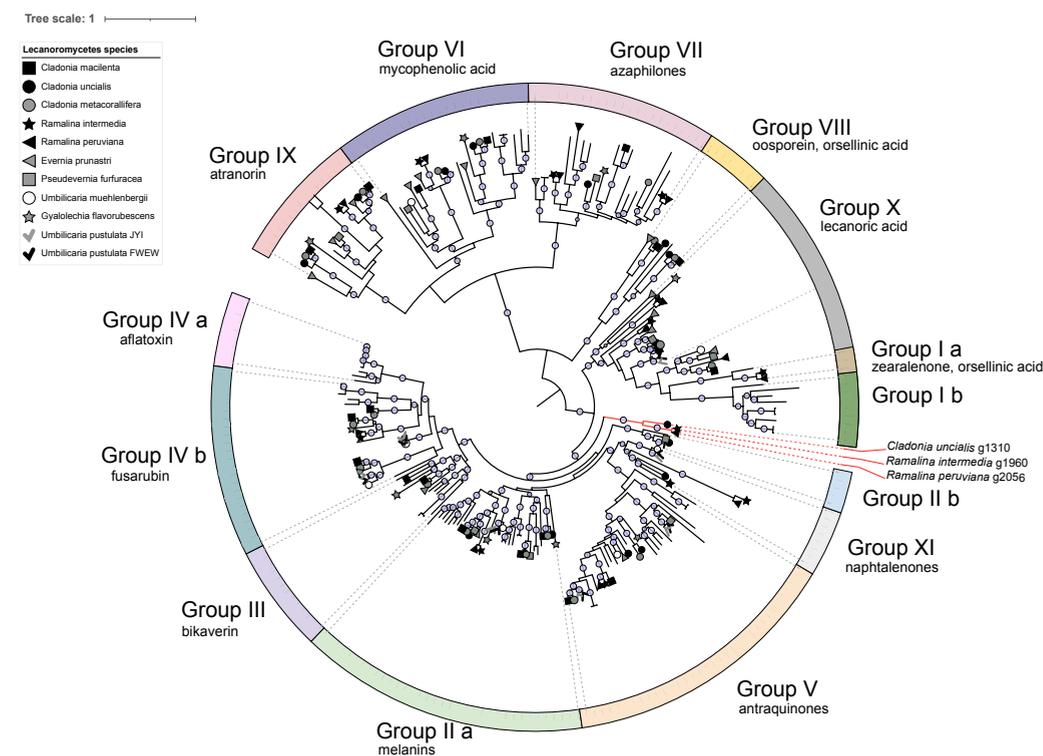
Figure 01.

Phylogenetic dereplication of nrPKSs from 9 genomes of LFF. Maximum likelihood phylogenetic tree was constructed with curated full protein sequences of nrPKS from LFF and 87 reference protein sequences from MiBIG database [32], and from publicly available sources (Supplementary Data Set 01). The tree is midpoint rooted, and bootstrap values calculated with Ultrafast bootstrap exceeding 95 are shown.

Results

Genome mining in lichens reveals unexplored GCF

In the present study, full protein sequences of nrPKSs from 9 LFF genomes along with 87 reference nrPKSs from the MiBIG database [32] were used to evaluate their evolutionary relationships. The maximum likelihood tree

obtained in this study agrees well with previous research [31]: all of the
11 reported groups are strongly supported (Figure 01). Group X does not
appear monophyletic and overlaps with Group I. Group I, however, is split
into two distinct clades, labeled as a and b. While Group X appears narrow,
the nrPKSs outside of Groups I and X exhibit notable differences, suggesting
they likely belong to other phylogenetic groups. However, due to limitations in
tree resolution, clear assignment is challenging. Three nrPKSs from genomes
of Lecanoromycetes form a new, strongly supported standalone clade,
designated as Group XII. Group XII is positioned basally to nrPKSs from Groups
II, III, IV, V, and XI, which typically assemble polyketide backbones featuring
multiple aromatic rings and hydroxymellein-like structures (Fig. 01, highlighted
in red). This positioning makes it hard to predict the possible backbone
produced by *Cladonia uncialis g1310*, *Ramalina intermedia* g1960 or *Ramalina
peruviana* g2056, or to suggest their involvement to any particular pathway
conserved for these species.

Comparative genomics reveals conserved GCF with a PKS from group XII

Previously described phylogenetic analysis resolves evolutionary relationships
within restricted set of proteins, while orthologs of these proteins may be
conserved in other fungal lineages. We have identified 23 close homologs of
group XII proteins in available Ascomycota proteomes. We then sought for
gene conservation in putative BGCs containing these orthologues. The resulting
Gene Cluster Family (GCF) consists of two main branches (Fig. 02). Phylogeny
of nrPKSs suggests that the tree mostly contains orthologues of PKSs from LFF,
or their close homologues in the lower part of the tree. Position of sequences
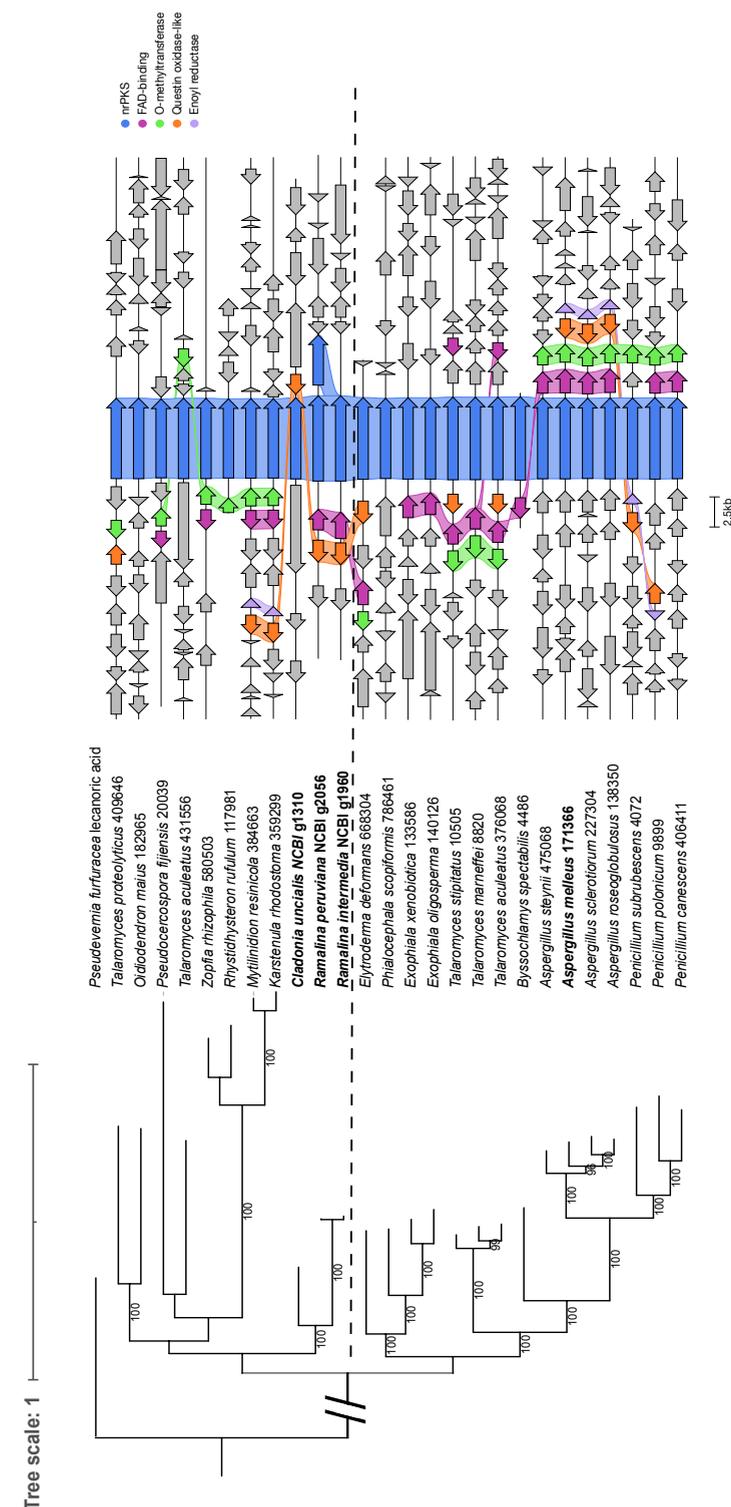
Figure 02.
Comparative analysis of genomic loci around nrPKS-orthologues of *C. uciclis, R. peruviana* and *R. intermedia*. Maximum likelihood phylogenetic tree was constructed with curated full protein sequences of identified homologous nrPKS available to the date of conducting this study. Lecanoric acid PKS is chosen as an outgroup. Ultrafast bootstrap values >95 are shown.

of *Talaromyces proteolyticus*, *Oidiodendron maius*, *Pseudocercospora fijiensis*, *Talaromyces aculeatus* and two *Exophiala xenobiotica* and *Exophiala oligosperma* sequences suggest either paralogy or horizontal gene transfer **(Fig. 02A)**. Each of the GCF branches contains specific but partially overlapping subset of tailoring genes **(Fig. 02B)**. Based on analysis of conserved domains in encoded tailoring genes, clusters from upper branch can be characterized by presence of FAD-binding protein and questin oxidase-like protein, while lower branch clusters contain two additional genes: enoyl reductase and methyl transferase.

Because paralogues of the lower branch contain the full set of conserved genes, we have chosen to characterize a pathway from *Aspergillus melleus* CBS112787, which consists of nrPKS, FAD-binding protein, methyl transferase, questin oxidase, and enoyl reductase. Absence of the enoyl reductase in pathways of the upper branch will likely yield different molecules, although the backbone product of the PKS is expected to be the same. Genes of *A. melleus* pathway were designated as follows: nrPKS *AmoA* (jgi.p_Aspneoa1_171366), FAD-binding domain containing protein *AmoB* (jgi.p_Aspneoa1_140189), O-methyltransferase *AmoC* (jgi.p_Aspneoa1_ 1140190), questin oxidase-like *AmoD* (jgi.p_Aspneoa1_128829), enoyl reductase *AmoE* (jgi.p_Aspneoa1_171370).

Heterologous expression of the AmoA nrPKS of *A. melleus* pathway yields tetraketides

Expression of nrPKS *AmoA* alone resulted in the production of two molecules by *Aspergillus oryzae* NSAR1, coined **1** and **2** **(Figure 03)**. Compound **1** has a retention time (RT) = 9.59 min, UV max of 199 and 275 nm (Supplementary **Fig. 01)**, and does not ionize when using ESI MS. Although the exact mass
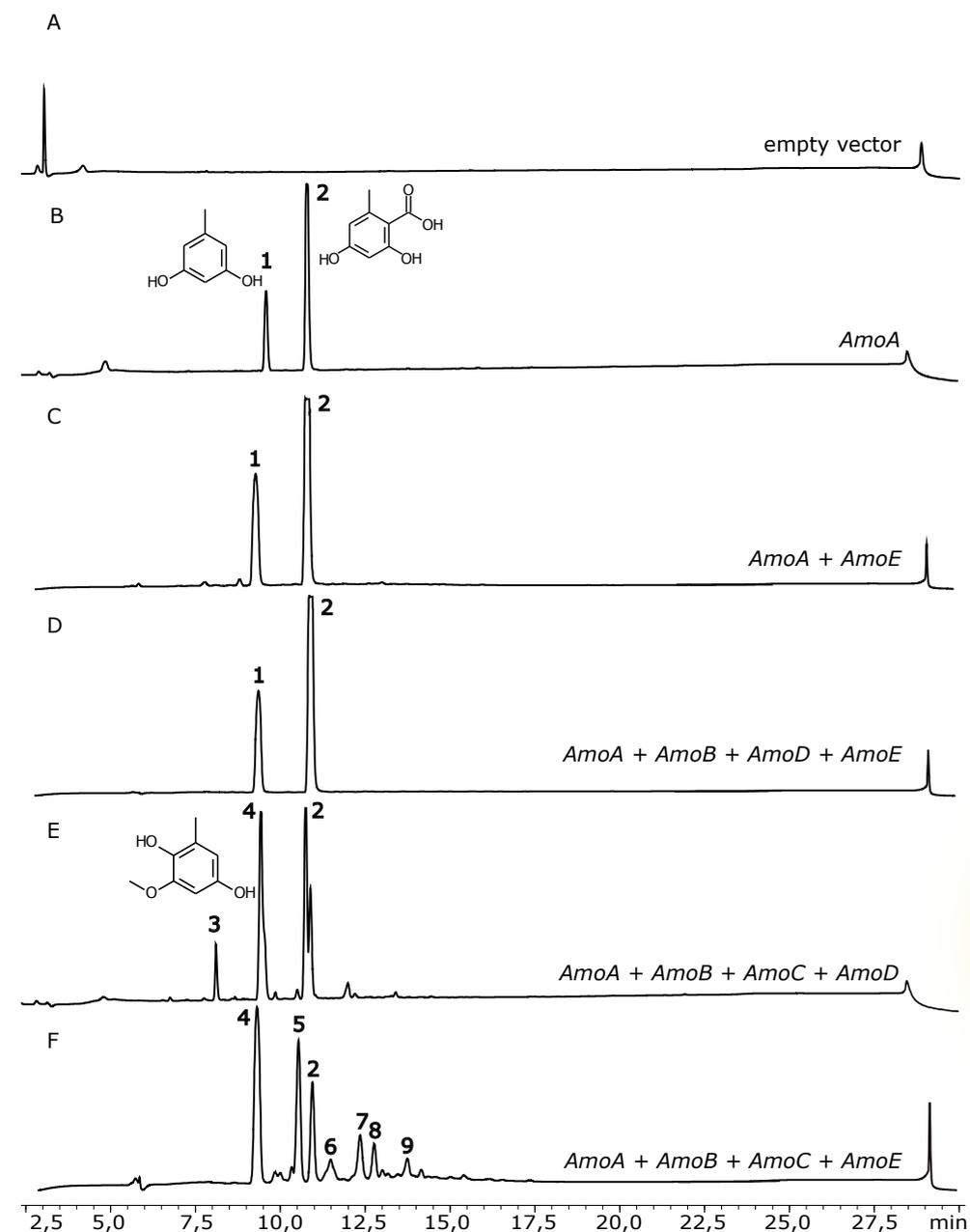
Figure 03.

HPLC chromatogram of extracts from *A. oryzae* NSAR1 expressing combinations of genes of the studied pathway.

of **1** has not been determined by high-resolution mass spectrometry either (HRMS), its UV and NMR data corresponds to published data for orcinol **(Supplementary Table S01, S2)**, **(34)**. Compound **2** (RT = 10.8 min, UV max of 213, 260 and 297 nm, *m/z* = 167 [M−H]−) was identified as orsellinic acid, with its UV, m/z and NMR signatures consistent with available data (Musharraf et al., 2015) **(Supplementary Table S03, S04)**. Overexpression of *PKS14* of *Fusarium graminearum* that was preformed by S. Jørgensen with colleagues has resulted in the production of orcinol and orsellinic acid **(33)**, however, in their experimental setup a lot more of orcinol was obtained compared to OA. It is not entirely clear whether endogenous enzymes of *A. oryzae* are performing decarboxylation of OA to orcinol, or it is an intrinsic activity of the expressed nrPKS AmoA. Given functional characterization of *AmoA*, we link nrPKS of the studied GCF to the production of orsellinates. Production of OA is commonly reported for the phylogenetic groups Ia, VIII and IX, and derivatives of OA (3-MOA, 5-MOA and DMOA) by groups VI, VII **(31)**.

Expression of combinations of pathway genes results in seven more molecules

Co-expression of nrPKS *AmoA* with enoyl reductase *AmoE,* and nrPKS *AmoA* with FAD-binding domain containing *AmoB,* questin oxidase-like *AmoD,* enoyl reductase *AmoE* **(Supplementary Table 05)** results in compounds **1** and **2** only **(Figure 03, panels C and D)**, indicating that these tailoring genes might not be able to act on the orcinol and OA polyketide backbones directly, but do so after modifications done by *AmoC.*

Co-expression of *AmoA* with *AmoB*, O-methyltransferase *AmoC* and *AmoD* provided two new compounds, coined **3** and **4** **(Figure 03E)**. Since

compound **1** is not observed in this sample, it is likely that **3** and **4** derive from **1**, while orsellinic acid **2** remains unmodified by the expressed enzymes. Compound **4** demonstrated absence of ionization and the same RT (9.59 min) as **1**, but UV max of 202 and 287 nm, and was elucidated using NMR as 1,4-dihydroxy-2-methoxy-6-methylbenzene **(Figure 04, Data Set S02)**. The NMR data for **4** **(Supplementary Data Set S02)** shows two meta-coupled aromatic protons, a methyl-group, a methoxy-group, two non-carbon bound protons (likely to be OH-groups), and 4 quaternary bound carbons.

Compound **3** (RT = 8.16 min; UVmax = 201, 274 nm; ionized at 293 [M-H]− and 295 [M+H]+) has the same UV spectrum as **1** but different RT. The exact mass of compound **3** is 294 as determined by HRMS **(Supplementary Figure 02)**. The elemental composition report (ECR) suggests the chemical formula $C_{14}H_{14}O_7$, indicative of a molecule with two benzene rings. The exact structure of **3** was not elucidated. The NMR data there are two benzene rings of which one corresponds to compound **4**, the modified orcinol.

Transformant expressing *AmoA (*nrPKS), *AmoB* (Fad-binding oxidoreductase), *AmoC* (methyl transferase), *AmoE* (enoyl reductase)
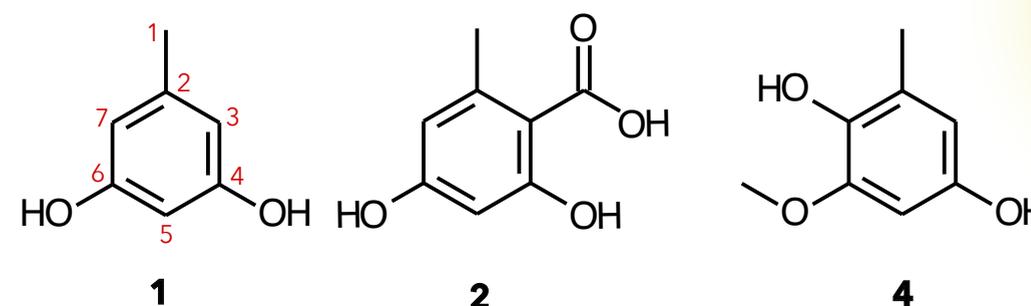
Figure 04.
Compounds identified in this study. **1** - orcinol, **2** - orsellinic acid, **4** - 1,4- dihydroxy-2-methoxy-6-methylbenzene.

produced **2** and **4**, and unique molecules **5, 6, 7, 8, 9** **(Figure 03F)**. **2** remains

unmodified by tailoring genes, which is consistent with data obtained for

AmoA, AmoB, AmoC, AmoD gene combination described above **(Figure 03E)**.

Molecule **4** originates from orcinol that undergoes O-methylation by *AmoC*,

and hydroxylation to C7 by *AmoB*. Compound **3** was not found in this sample,

indicating that **3** is a product of co-expression with *AmoD*. Compound **5**

bears unique UV spectrum (Uvmax = 193, 266, 373) and ionizes in both modes

(TIC+: 194, TIC-: 275), likely being a mixture of two compounds based on the

ionization pattern. Compound **6** bears UV profile similar, but not identical, to **3**

(UVmax = 204, 265) and ionizes in negative mode (TIC-: 319). Compounds **7** and

8 have the same UV profile with minor differences (UVmax=204, 269, 357 for

7 and UVmax = 217, 266 for **8**). Both **7** and **8** ionize the same way ([M+H]+=319,

[M-H]—=317, *m/z* = 318), taken together indicating that these molecules have

minimal structural differences. Compound **9** (UVmax = 205, 273, 350) ionized

in both modes, and based on ionization patterns, may actually be a mixture

of two molecules (TIC+: 304, 425, TIC- : 423), *m/z* of one of them being 424

(Supplementary Figure 01).

1,4-dihydroxy-2-methoxy-6-methylbenzene is responsible for antimicrobial activity against range of gram-negative bacteria, and inhibits melanization of zebrafish embryos

Disc diffusion assay performed with crude extracts of *A. oryzae* NSAR1

transformants demonstrated that extract from culture releasing compounds

assembled by *AmoA+AmoB+AmoC+AmoD* inhibited growth of Gram-negative

and Gram-positive bacteria, namely *Escherichia coli* DH5α and *Bacillus subtilis*,
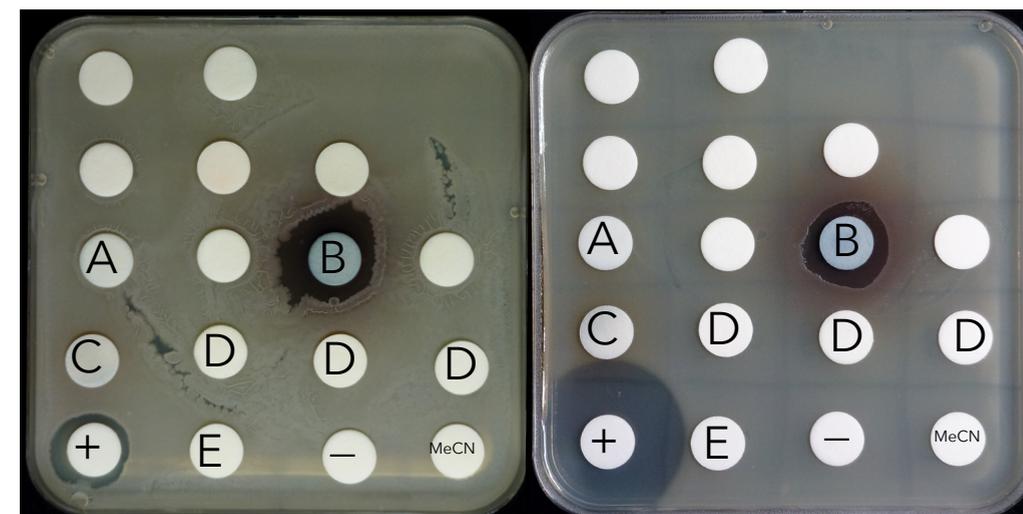
Figure 05.

Disc diffusion assay of crude extracts of *A. oryzae* NSAR1 assayed against *E. coli* DH5α cells on the
left plate and versus *B. subtilis* on the right plate. Crude extracts correspond to gene combinations
expressed by *A. oryzae* are as follows: A - *AmoA+ AmoE*; B - *AmoA + AmoB + AmoC + AmoD*; C - *AmoA*;
D - *AmoE*; E - *AmoB+AmoC+AmoD*; '+' corresponds to 50 ug/mL ampicillin; '—' corresponds to an
extract from cultivation media with no fungus inoculated; 'MeCN' corresponds to acetonitrile control.

respectively **(Figure 05)**. Interestingly, crude extracts of strains expressing

AmoA alone, or *AmoA+AmoE* did not demonstrate antibiotic activity. None of

the tested extracts demonstrated antifungal activity when assayed against

Candida albicans CBS562 **(Supplementary Figure 03)** and *Penicillium rubens*

(Supplementary Figure 04).

The bioactive crude extract was subjected to preparative HPLC,

and individual compounds were tested for biological activity. It was found

that **4** corresponds to the antibiotic activity. Bioactive fraction containing **4**

was tested on a panel of clinically relevant strains of bacteria **(Figure 06A)**.

Compound **4** completely inhibited growth of *Stenotrophomonas maltophilia*,

Staphylococcus epidermidis, both *Staphylococcus aureus* strains MSSA476 and

A

B

| Control | Compound 4 1:8000 |
| Compound 3 1:500 | Compound 4 1:4000 |

Figure 06.
A – Growth profile of bacteria with compound **4** overnight, compound **4** is tested in duplicate (F1 and F2), 2% DMSO without compound serves as a negative control. B – zebrafish (*D. rerio*) embryos 48 hpf exposed to compounds **3** and **4**, from top to bottom: control (solvent), **3** 1:500 dilution, **4** 1:8000 dilution, **4** 1:4000 dilution. Representatives from 10 fertilized embryos are shown.

USA300, *A. calcoaceticus*. It delayed growth of *Enterococcus faecium* 15A623, *E. faecium* 16D030, *Acinetobacter baumanii*, *E. coli*, and did not affect growth of *Klebsiella pneumoniae, Pseudomonas aeruginosa, Enterobacter cloacae*.

We additionally assessed toxicity of **3** and **4** on a zebrafish (*Danio rerio*) embryo model **(Figure 06B)**. Embryos were exposed to compounds **3** or **4** from 8 hours post fertilization (hpf) to 48 hpf. Both compounds lead to demelanized phenotype of the fish embryo, alhough for compound **3** the effect is not visible in higher dilutions **(Supplementary Figure 05)**. Compound **4** leads to pigmentation defect in both dilutions and to protrusion at the tail at 1:4000 dilution. Additionally, all embryos exposed to compounds **3** and **4** have enlarged cardial sac (round formation at the ventral side of the embryo).
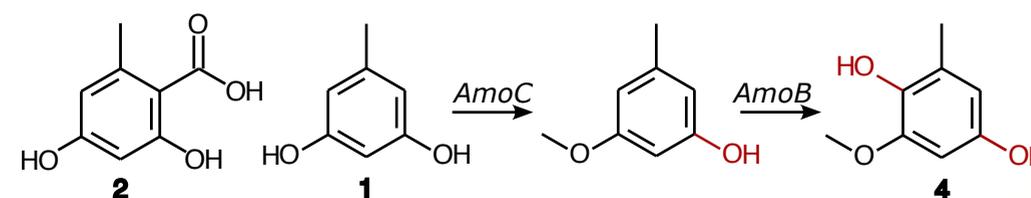
Figure 07.
Proposed biosynthetic pathway. Positions in molecules that could be possible targets for *AmoE* are highlighted in red.

Discussion

Proposed pathway from orcinol to 1,4-dihydroxy-2-methoxy-6-methylbenzene

Expression of *AmoA* alone results in production of individual molecules **1** and **2**, with no detected dimers **(Figure 07)**. Orcinol **1** has previously been reported

to be found when orsellinic acid is produced in the study of Jørgensen with colleagues [33]. PKS14 of *Fusarium graminearum* was shown to be expressed during plant infection, and promoter swap to a constitutive *PgpdA* achieved production of orsellinic acid and orcinol [33]. It was hypothesized, however, that decarboxylase FGSG_03965 that is co-regulated with PKS14 is likely to catalyze the decarboxylation of orsellinic acid to orcinol. Orsellinic acid BGC of *A. nidulans* contains PKSs AN7907.4 and AN7909.4, decarboxylase AN7911.4 and tyrosinase AN7912.4, it was demonstrated that sole AN7907.4 activity is sufficient to generate OA, although no monoaromatic orcinol was detected [35]. It was also hypothesized that orcinol is produced by decarboxylation of OA in LFF *Umbilicaria papulosa* and *Gliocladium roseum* [36]. When *G. roseum* was incubated with OA supplemented in the medium, orcinol was found to be secreted [36]. Moreover, it was demonstrated that during incubation of medium containing OA over period of 5 days at 27°C, 20% of OA was spontaneously converted to orcinol [36].

Our findings on expression of *AmoA* alone suggest spontaneous decarboxylation of orsellinic acid. We cannot rule out possible activity of an endogenous decarboxylase during heterologous expression in *A. oryzae*. Feeding experiments of OA to *A. oryzae* cultures would be helpful to provide more insight on involvement of the endogenous decarboxylase, and consequently the mechanism of assembly of OA and orcinol by AmoA. Alternatively, bifunctional TE domain of the AmoA may explain simultaneous production of orsellinic acid and orcinol by the nrPKS.

Based on structures of molecules identified in this study, mainly orcinol becomes a substrate for tailoring genes of the pathway, which yields **3**, **4**, although we cannot state that **5**, **6**, **7**, **8** and **9** derive from orcinol exclusively. Due to the fact that gene combinations *AmoA+AmoE* and *AmoA+AmoB+AmoD+AmoE* result in no modifications to either **1** or **2**, these

results suggest that backbone modifications start after a methyl group is attached to oxygen bound to C6 of **1** by O-methyltransferase *AmoC*. Next, FAD-binding domain containing protein *AmoB* is able to introduce hydroxyl into C7 position of O-methylorcinol, resulting in **4** (Fig. 07).

It is not clear which molecules are modified by the questin oxidase AmoD. Previous studies for enzymes annotated as questin oxidase characterized them to be involved in ring-cleaving of anthraquinone-like chemical moieties [37]. Although structure of **3** was not elucidated completely, it indicates possible dimerization of two molecules, one of them being **4**. We speculate that AmoD may be involved in formation of **3**, but due to the absence of conclusive data on structure of **3** no further speculations can be suggested.

Co-expression of AmoA, AmoB, AmoC with enoyl reductase *AmoE* is resulting in production of **5**, **6**, **7**, **8** and **9**. Interestingly, *m/z* value of compounds **7** and **8** (Supplementary Fig. 01) corresponds to the mass of the depside lecanoric acid, and their UV profile is close to the UV profile of lecanoric acid standard [30]. Because no structural elucidation was performed for **7** and **8**, no speculations are possible. Further structural studies on molecules **3**, **5**, **6**, **7**, **8** and **9** are important for establishing roles of individual enzymes in the pathway.

1,4-dihydroxy-2-methoxy-6-methylbenzene is a putative tyrosinase inhibitor with antimicrobial activity

Lichen orsellinates belonging to the group of depsides and depsidones demonstrate a spectrum of biological activities, from photoprotective and antioxidant to antimicrobial, antiviral and anticancerous [04]. Monoaromatic

compounds from lichens also possess various biological activities, including cytotoxic [38], antiprotozoal [39], antifungal [39] and antibacterial [40]. A number of such compounds demonstrated antibiotic activity against antibiotic-resistant, pathogenic bacteria *E. faecium*, *S. aureus* and *A. baumannii* (41). Methyl orsellinate and orcinol are reported to have antibiotic activity against *S. aureus* and *E. coli* [42] and methicillin resistant *S. aureus* and *Enterococcus faecalis*, respectively [43]. Interestingly, orcinol was not found to be active against *A. baumannii*. We speculate that structural rearrangements of **4** may be responsible for making this molecule more potent towards *A. baumannii*.

It has been demonstrated that orcinol decreased intracellular tyrosinase activity, melanin content, and the expression of melanogenesis-related genes, including tyrosinase, in B16F10 murine melanoma cells [44]. Tyrosinase inhibiting activity was also reported for a fungal metabolite kojic acid [45] and hydroquinone [46]. Orcinol cytotoxicity on B16F10 cells was not pronounced up to 1.0 mM, when it decreased melanin content and tyrosinase activity [44], making this compound promising for skin lightening formulations. The mechanism of demelanizing activity of orcinol is mediated by downregulation of expression of MITF, a transcription factor that regulates melanogenesis, mediated by activation of the extracellular signal-regulated kinase (ERK) pathway in cells.

We did not evaluate mechanisms of demelanization of *D. rerio* embryos mediated by **4** in this study. Because structure of **4** is derived from orcinol, we speculate that similar mode of action can be underlying demelanizing activity of **4**. Detailed evaluation of potency for demelanizing activity and cytotoxicity of **4** can divert its potential application to comsetics and dermal care for skin pigmentation disorders, or as anticancerous agent for treating melanoma.

Materials and methods

- **Fungal genomes and gene curation.** Genome assemblies of Lecanoromycetes were retrieved from NCBI database, namely: *Cladonia macilenta* KoLRI003786 (GCA_000444155.1), *Cladonia metacorallifera* KoLRI002260 (GCA_000482085.2), *Cladonia uncialis* strain Normore 8774 (GCA_002927785.1), *Evernia prunastri* strain FR SP7-11 epruFC11 (GCA_003184365.1), *Gyalolechia flavorubescens* KoLRI002931 (GCA_000442125.1), *Pseudevernia furfuracea* strain AKPM 0122M pfurFC1 (GCA_003184345.1), *Ramalina intermedia* strain YAF0013 (GCA_003073195.1), *Ramalina peruviana* strain YAF0012 (GCA_001956345.1), *Umbillicaria muehlenbergii* strain KoLRI No. LF000956 (GCA_000611775.1), *Umbillicaria (Lasallia) pustulata* (GCA_900169345.1), *Umbilicaria pustulata* isolate Sardinia_28052013 (GCA_000938525.1). Gene prediction for these assemblies was made using Augustus [47] with default parameters, gff3 output file type and *Aspergillus fumigatus* as a training set. BGCs were predicted using antiSMASH 4 (https://pubmed.ncbi.nlm.nih.gov/28460038/) with following parameters: antismash *.fasta -- gff3 *.gff3 --minimal --verbose --taxon fungi. Predicted nrPKSs with conserved domains signature containing starter SAT (PF16073) and protein template PT (TIGR04532) were subtracted from the dataset and used for further analysis. nrPKS genes were manually curated (Supplementary Data Set 01).

- **Alignment and phylogenetic analysis.** KS-domain (PF00195.19) guided protein alignment was done using Clustal Omega v1.2.4 [48], poorly aligned regions were trimmed using TrimAl 1.4.rev15 [49]. Maximum-likelihood tree was constructed using iq-tree (https://academic.oup.com/mbe/article/32/1/268/2925592) with following parameters: -bb 1000 -nt AUTO -mset LG -alrt 1000 -abayes -m MFP. Ultrafast bootstrapping, approximate Bayes test [50] and a Shimodaira-Hasegawa approximate likelihood-ratio test [51] were employed for evaluation of tree topology. Resulting trees were visualised using iTOL [52]. Alignment and tree files are provided in Supplementary Data Set 1. Gene models of *C. uncialis g1310*, *R. intermedia g1960* or *R. peruviana g2056* were curated to contain correct exon-intron boundaries and are likely to be functional.

Strains and culture conditions. *Aspergillus melleus* CBS112787 and *A. oryzae* NSAR1 were routinely maintained on malt extract agar plates (50 g/L [pH 5.4]; Oxoid). *A. oryzae* NSAR1 transformants were maintained on selective Difco Czapek-Dox agar supplied with nutrients in accordance with genetic construct: arginine (1 g/L), adenine hemisulfate salt (0.5 g/L), methionine (1.5 g/L) and ammonium sulphate (1 g/L). For the induction of genes under *PamyB* control, transformants were grown at 30°C for 5 days in liquid yeast malt extract (YM) medium (3 g/L Difco yeast extract, 3 g/L Difco malt extract, 5 g/L Difco Bacto peptone, 10 g/L glucose [Merck, Kenilworth, NJ]), or solid (YMA) with agar (Ferwo 700 agar).

Gene fragments amplification. nrPKS *AmoA* of *A. melleus* CBS112787 (JGI accession: jgi.p_Aspneoa1_17136, 6,270 bp) was constructed synthetically (Twist Bioscience, CA, USA) without introns and provided in two parts. Larger 5′ end (4,604 bp) was pre-cloned into Gateway® compatible pTWIST vector, and remaining 1,666 bp were supplied as overlapping fragments. Complete gene was assembled into pTWIST vector using SLiCE [53].

Tailoring genes *AmoB* (jgi.p_Aspneoa1_140189, 1,473 bp), *AmoC* (jgi.p_Aspneoa1_1140190, 1,335 bp), *AmoD* (jgi.p_Aspneoa1_128829, 1,482 bp) were synthesised without introns into pTWIST vectors (Twist Bioscience, CA, USA), which served as PCR template for genes amplification. ER gene (jgi.p_Aspneoa1_171370, 855 bp) contains no introns and therefore was amplified from gDNA of *A. melleus*.

Genes were amplified using polymerase chain reactions (PCR) using PhusionTM High-Fidelity DNA polymerase (Thermofisher Scientific, Wilmington DE), according to the manufacturer's protocol and with annealing of the primers at 60°C for 1 minute (Supplementary Table S06). The PCR products from the amplified genes were purified using either the Geneclean® II Kit (MPbio, Solon, OH, USA) for those intended for yeast transformation, or the NucleoSpin® Gel and PCR Clean-up Kit (Macherey-Nagel) for those designated for SLiCE or BP Gateway®-cloning (Invitrogen), following the respective manufacturer's protocols.

Vector construction. Genes were cloned into vectors were cloned into plasmids using either transformation-associated recombination (TAR) in *S. cerevisiae*, or Seamless Ligation

Cloning Extract (SLiCE). Prior to that, one microgram of the pEYA2 entry vector was digested overnight with 10 U NotI at 37°C, or pTYGS destination vector was digested overnight with 1U SgsI and inactivated for 20 minutes at 65 °C (Thermo Fisher Scientific, Waltham, MA). Fragments of the expected size and the linearized plasmid were purified from a 0.8% agarose gel or directly from the PCR or digestion mix using a Geneclean II kit (MP Biomedicals, Santa Ana, CA). For TAR, *S. cerevisiae* BMA 64 with a ura3 auxotrophic marker was used. The protocol was adapted from [54] with following modifications. Five mL of yeast culture was grown at 30°C in YPD (Yeast extract (Difco 212759) - 8 g/L, Bacto peptone (Difco 211677) - 16 g/L, D(+)-Glucose (Merck 1.08337) - 16 g/L) overnight. Two mL containing approximately $1*10^8$ cells were transferred to 50 mL YPD in 250 mL Erlenmeyer flask and incubated at 30°C and agitation 200 rpm for about 5 hours. Yeast biomass was pelleted for 5 min at 2,000 rpm at room temperature, washed with 10 mL sterile water, pelleted again and reconstituted in 300 µL sterile water. Fifty microliters of resulting cell suspension were combined with 250 µL of 100 mM DTT (dithiothreitol) and incubated for 10 minutes at room temperature. Yeasts were pelleted for 15 sec at 12,000 g and the supernatant was discarded. Pellet was reconstituted with 500 µL PLTE solution (for 1 mL: 800 µL 50% PEG 4000, 0.1 mL 1 M LiAc, 20 µL 50 mM EDTA, 10 µL 1M TrisHCL pH 7.5 , 70 µL H_2O), 8-10 µL of DNA (plasmid and gene fragments in equal proportions) and 50 µL of recently boiled salmon sperm DNA (2 mg/mL) and incubated for 1 hour at 30°C without shaking. Samples were exposed to heat for 15 min at 45°C, spun for 15 sec at 12,000 g and resuspended in 1 mL of YPD. After 30 min incubation at 30°C without shaking, cells were pelleted for 15 sec at 12,000 g, reconstituted in 200 µL of sterile water and plated to synthetic dropout media SDM (20 g/L agar, 20 g/L D-glucose (Sigma-Aldrich, St. Louis, MO), 1.92 g/L yeast dropout supplements without uracil (Sigma-Aldrich, St. Louis, MO), 6.7 g/L yeast nitrogen base without amino acids (Sigma-Aldrich, St. Louis, MO)).

SLiCE was performed in accordance with published protocol [53]. Initially, *E. coli* JM109 cell extract was prepared by overnight preculturing at 37°C, 200 rpm with subsequent transfer of 1 mL to 50 mL LB media for further growth until reaching an OD600 of approximately 2.0-3.0, taking approximately 6 hours. The cell culture was then centrifuged at 5,000 g for 10

5

minutes at 4°C, and the resulting pellet was washed and resuspended in diluted and buffered Cell Lytic B cell lysis reagent. After centrifugation at 20,000 g for 2 minutes at 4°C, further steps were conducted on ice. The supernatant was mixed with glycerol, aliquoted, snap-frozen in liquid nitrogen, and stored at -80°C. For SLiCE reaction, 1 μL of SLiCE 10x buffer was mixed with 1 μL of *E. coli* cell extract and DNA containing 10-100 ng of linear vector (CIAP-treated) and 20-200 ng of DNA fragments. The mixture was adjusted to a final volume of 10 μL with nuclease-free water, incubated at 37°C for 30 minutes, and then transformed into heat-shock competent *E. coli* DH5α (Invitrogen) according to the manufacturer's protocol.

SLiCE was used to clone *AmoE* gene (ER, jgi.p_Aspneoa1_171370, 855 bp) into *pEYA2* entry vector, which was propagated in *E.coli* DH5α. Homologous recombination in yeast was used for genes *AmoB*, *AmoC*, and *AmoD* to create destination vector *pTYGSarg::AmoB::AmoC::AmoD*. One Shot ccdB Survival 2T1 *E. coli* cells were used to propagate *pTYGS* vector, unless indicated otherwise. Expression vectors were constructed using LR Gateway® cloning (Invitrogen), using *pEYA2::AmoA* or *pEYA2::AmoE* as entry vectors and *pTYGSarg::AmoB::AmoC::AmoD* or *pTYGSade* as destination vectors, respectively.

Construction of expression vectors. Expression vectors were obtained via Gateway® cloning (Thermo Fisher Scientific, Waltham, MA). Seventy nanograms of the *pEYA2::AmoA* entry vector and 100 ng of the *pTYGSarg::AmoB::AmoC::AmoD* destination vector were mixed with 1 μL of the Gateway LR Clonase II enzyme in 10 μL final volume, and the reaction mixture was incubated at 25°C overnight. After inactivation according to manufacturer's protocol, entire reaction was transformed into heat-shock competent *E.coli* DH5α.

Transformation of *A.oryzae* NSAR1. *A.oryzae* NSAR1 was transformed according to previously described procedure [31]. 10 μg of vector was used to transform approximately $1*10^7 * mL^{-1}$ protoplasts. Each transformation reaction was spread to four square plates (Greiner) and incubated at 28°C for 5-8 days until germination of transformants.

Nucleic acid isolation, RT-PCR. RNA was isolated using Trizol (Thermo Fisher Scientific, Waltham, MA) following previously described procedure [31], and cleanup was performed according to NucleoSpin RNA extraction kit (Macherey- Nagel, Allentown, PA). Five hundred

nanograms of total RNA underwent cDNA synthesis with oligo(dT) primers and GoScript reverse transcription (RT) mix (Promega, Madison, WI) following the manufacturer's instructions. To validate gene expression in *A. oryzae* transformants, primers specific for the *A. melleus* genes and *A. oryzae* H2B gene **(Supplementary Table S06)** were used with GoTaq DNA polymerase (Promega, Madison, WI) according to the manufacturer's protocol.

Extraction of metabolites and HPLC-MS analysis. Liquid culture after 5 days of cultivation (300 mL YM medium supplemented with starch in 1L Erlenmeyer flask, 28°C, 200 rpm) were separated to supernatant and mycelium by filtration. Supernatant was mixed with ethyl acetate in 1:1 proportion. To facilitate the migration of compounds, the medium was acidified by the addition of 0.1 mL of 37% HCl per 100 mL. After at least 1 hour of extraction with ethyl acetate and shaking, samples were centrifuged at 6,000 rpm for 10 minutes to separate the organic and aqueous phases. The organic phase was carefully transferred to a sterile 50 mL tube. Following the removal of the organic phase, the extraction process was repeated twice, first with ethyl acetate and then with 2-butanone. The organic phase from each extraction round was evaporated under nitrogen flow. The dried compounds were pooled together and reconstituted in 500 μL of acetonitrile. These crude extracts were then stored at -80°C until further use.

Compounds purification. Preparative HPLC (prepHPLC) was conducted utilizing a Shimadzu CBM-20A controller, a Shimadzu LC-20AP pump, and a Shimadzu FRC-10A fraction collector, all equipped with a C18 reversed-phase Reprosil column (10 μm, 120 Å, 250 × 22 mm). Buffer A consisted of MQ with 0.1% TFA, while buffer B comprised HPLC-grade acetonitrile with 0.1% TFA. Organic extracts from fungal cultures were subjected to prepHPLC following this program: 5% buffer B for 5 min, a linear gradient from 0% to 95% B over 25 min, 95% B for 5 min, and finally 100% buffer A for 5 min, wuth a flow rate of 12.5 mL/min. Analysis of results was performed using LabSolutionsTM software (Shimadzu Corporation).

Structural elucidation.

HRMS. High-resolution mass spectrometry (HRMS) was conducted using an LCT instrument (Micromass Ltd, Manchester, UK) to ascertain the precise mass of the extracted

5

compounds. Calibration was achieved using sodium formate (NaFor), and samples were measured in conjunction with NaFor.

NMR. ^{1}H NMR (600 MHz) and ^{13}C NMR spectroscopy (151 MHz) were conducted on a Bruker 600 spectrometer. Proton chemical shifts are reported in parts per million relative to tetramethylsilane and referenced to residual protium in the solvent (^{1}H NMR: DMSO-d6 at 2.50 ppm). Carbon chemical shifts are reported in parts per million relative to tetramethylsilane and referenced to the carbon resonances of the residual solvent peak (^{13}C NMR: DMSO-d6 at 39.52±0.06 ppm). NMR spectra were analyzed using Mnova software 14.2.3 (MestreLab).

Bioassays: bacteria, zebrafish embryos. Bacterial strains of *Enterococcus faecium* 16D030 (vancomycin resistant), *Enterococcus faecium* 15A623 (vancomycin succeptible), *Enterobacter cloacae*, *Escherichia coli*, *Pseudomonas aeruginosa*, *Klebsiella pneumoniae*, *Acinetobacter baumanii*, *Acinetobacter calcoaceticus*, *Staphylococcus aureus* MSSA476 (methicillin resistant), *Staphylococcus aureus* USA300 (methicillin succeptible), *Staphylococcus epidermis*, *Stenotrophomonas maltophilia* were used. *E. coli, Staphylococcus sp., Pseudomonas sp.* and *Klebsiella sp.* were grown in MHB (Oxoid, CM0405B), the other strains were grown in TSB (Oxoid, BO0351R).

50 µL of fraction containing compound **4** was mixed with 150 µL DMSO, and 10 µL of this mixture was loaded to the well of 96-well plate (Corning) containing 90 µL of media and 100 µL of culture grown till OD 0.4-0.6 and diluted 100 times. Growth of bacteria was performed in a Multiskan FC plate reader (ThermoFisher) and monitored in a course of 16h using OD600 as a readout. Plane media with DMSO was used as control.

Disc diffusion assay. To evaluate the antibacterial activity of the crude organic extracts, disc diffusion assays were conducted. Overnight cultures (500 µL) of *E. coli* DH5α or *B. subtilis* were spread into Luria Agar (LA) plates (per 1L: 10g tryptone (Oxoid LP0042), 5 g yeast extract (Difco 212750), 5 g sodium chloride (NaCl) (Baker 0278), 0.02 g thymine (Sigma T-0376), 15 g agar bacteriological (Oxoid L11)), and air-dried. Sterile Whatman discs soaked with 75 uL of extract were applied to the surface of the plate. Cultures were placed to

37°C incubator overnight. For anti-fungal activity assay, 500 µL of *P. ruben*s spores in sterile H_2O were plated onto MEA plates (Malt extract agar, Oxoid cm59), or 550 µL of *C. albicans* CBS562 cells were plated onto Sabouraud dextrose agar (Difco 210950). Following this, the plates were air-dried, and crude extract disks were placed onto the surface of agar. *P. rubens* plates were incubated at room temperature for 3 days, while *C. albicans* plates were incubated at 35°C for 1 day.

5

Acknowledgements

References

01. **Pol, C.S., Savale, S.A., Khare, R., Verma, N., Behera, B.C. (2017).** Antioxidative, Cardioprotective, and Anticancer Potential of Two Lichenized Fungi, *Everniastrum cirrhatum* and *Parmotrema reticulatum,* from Western Ghats of India. J*ournal of Herbs, Spices & Medicinal Plants, 23:2, 142-156.* https://doi.org.10.1080/10496475.2017.1280578

02. **Singh, G., Armaleo, D., Dal Grande, F., & Schmitt, I. (2021).** Depside and Depsidone Synthesis in Lichenized Fungi Comes into Focus through a Genome-Wide Comparison of the Olivetoric Acid and Physodic Acid Chemotypes of *Pseudevernia furfuracea*. *Biomolecules, 11(10), 1445.* https://doi.org/10.3390/biom11101445

03. **Norouzi, H., Sohrabi, M., Yousefi, M., & Boustie, J. (2023).** Tridepsides as potential bioactives: a review on their chemistry and the global distribution of their lichenic and non-lichenic natural sources. *Frontiers in fungal biology, 4, 1088966.* https://doi.org/10.3389/ffunb.2023.1088966

04. **Ureña-Vacas, I., González-Burgos, E., Divakar, P. K., & Gómez-Serranillos, M. P. (2022).** Lichen Depsidones with Biological Interest. *Planta medica, 88(11), 855–880.* https://doi.org/10.1055/a-1482-6381

05. **Yang, J., Zhou, Z., Chen, Y., Song, Y., & Ju, J. (2023).** Characterization of the depsidone gene cluster reveals etherification, decarboxylation and multiple halogenations as tailoring steps in depsidone assembly. *Acta pharmaceutica Sinica. B, 13(9), 3919–3929.* https://doi.org/10.1016/j.apsb.2023.05.036

06. **Zhao, X., Chen, Y., Long, T., Liu, Z., Zhang, Q., Zhang, H., Yan, Y., Zhang, C., & Zhu, Y.** (2023). Genome Mining and Biosynthetic Reconstitution of Fungal Depsidone Mollicellins Reveal a Dual Functional Cytochrome P450 for Ether Formation. *Journal of natural products, 86(8), 2046–2053.* https://doi.org/10.1021/acs.jnatprod.3c00609

07. **Araújo, A. A., de Melo, M. G., Rabelo, T. K., Nunes, P. S., Santos, S. L., Serafini, M. R., Santos, M. R., Quintans-Júnior, L. J., & Gelain, D. P. (2015).** Review of the biological properties and toxicity of usnic acid. *Natural product research, 29(23), 2167–2180.* https://doi.org/10.1080/14786419.2015.1007455

08. **Millot, M., Tomasi, S., Articus, K., Rouaud, I., Bernard, A., & Boustie, J. (2007).** Metabolites from the Lichen *Ochrolechia parella* growing under two different heliotropic conditions. J*ournal of natural products, 70(2), 316–318.* https://doi.org/10.1021/np060561p

09. **Bellio, P., Segatore, B., Mancini, A., Di Pietro, L., Bottoni, C., Sabatini, A., Brisdelli, F., Piovano, M., Nicoletti, M., Amicosante, G., Perilli, M., & Celenza, G. (2015).** Interaction between lichen secondary metabolites and antibiotics against clinical isolates methicillin-resistant *Staphylococcus aureus* strains. *Phytomedicine : international journal of phytotherapy and phytopharmacology, 22(2), 223–230.* https://doi.org/10.1016/j.phymed.2014.12.005

10. **Roser, L. A., Erkoc, P., Ingelfinger, R., Henke, M., Ulshöfer, T., Schneider, A. K., Laux, V., Geisslinger, G., Schmitt, I., Fürst, R., & Schiffmann, S. (2022).** Lecanoric acid mediates anti-proliferative effects by an M phase arrest in colon cancer cells. Biomedicine & pharmacotherapy = Biomedecine & pharmacotherapie, 148, 112734. https://doi.org/10.1016/j.biopha.2022.112734

11. **Rajendran K, Ponmurugan P, Gnanamangai BM, Karuppiah P, Shaik MR, Khan M, Khan M, Shaik B. (2023).** Bioefficacy of Lecanoric Acid Produced by *Parmotrema austrosinense* (Zahlbr.) Hale against Tea Fungal Pathogens. *Horticulturae 9(6):705.* https://doi.org/10.3390/horticulturae9060705

12. **Kealey, J. T., Craig, J. P., & Barr, P. J. (2021).** Identification of a lichen depside polyketide synthase gene by heterologous expression in *Saccharomyces cerevisiae*. *Metabolic engineering communications, 13, e00172.* https://doi.org/10.1016/j.mec.2021.e00172

5

13. **Sanchez, J. F., Chiang, Y. M., Szewczyk, E., Davidson, A. D., Ahuja, M., Elizabeth Oakley, C., Woo Bok, J., Keller, N., Oakley, B. R., & Wang, C. C. (2010).** Molecular genetic analysis of the orsellinic acid/F9775 gene cluster of *Aspergillus nidulans. Molecular bioSystems, 6(3),* 587–593. https://doi.org/10.1039/b904541d

14. **Gressler, M., Hortschansky, P., Geib, E., & Brock, M. (2015).** A new high-performance heterologous fungal expression system based on regulatory elements from the *Aspergillus terreus* terrein gene cluster. *Frontiers in microbiology, 6, 184.* https://doi.org/10.3389/fmicb.2015.00184

15. **Lünne, F., Niehaus, E. M., Lipinski, S., Kunigkeit, J., Kalinina, S. A., & Humpf, H. U. (2020).** Identification of the polyketide synthase PKS7 responsible for the production of lecanoric acid and ethyl lecanorate in *Claviceps purpurea. Fungal genetics and biology : FG & B, 145,* 103481. https://doi.org/10.1016/j.fgb.2020.103481

16. **Jørgensen, S. H., Frandsen, R. J., Nielsen, K. F., Lysøe, E., Sondergaard, T. E., Wimmer, R., Giese, H., & Sørensen, J. L. (2014).** *Fusarium graminearum* PKS14 is involved in orsellinic acid and orcinol synthesis. *Fungal genetics and biology : FG & B, 70, 24–31.* https://doi.org/10.1016/j.fgb.2014.06.008

17. **Armaleo, D., Müller, O., Lutzoni, F. et al. (2019).** The lichen symbiosis re-viewed through the genomes of *Cladonia grayi* and its algal partner *Asterochloris glomerata. BMC Genomics 20, 605.* https://doi.org/10.1186/s12864-019-5629-x

18. **Allen, J. L., McKenzie, S. K., Sleith, R. S., & Alter, S. E. (2018).** First genome-wide analysis of the endangered, endemic lichen *Cetradonia linearis* reveals isolation by distance and strong population structure. *American journal of botany, 105(9), 1556–1567.* https://doi.org/10.1002/ajb2.1150

19. **Allen, J. L., Jones, S. J. M., & McMullin, R. T. (2021).** Draft Genome Sequence of the Lichenized Fungus *Bacidia gigantensis. Microbiology resource announcements, 10(44),* e0068621. https://doi.org/10.1128/MRA.00686-21

20. **Armstrong, E. E., Prost, S., Ertz, D., Westberg, M., Frisch, A., & Bendiksby, M. (2018).** Draft Genome Sequence and Annotation of the Lichen-Forming Fungus *Arthonia radiata. Genome announcements, 6(14), e00281-18.* https://doi.org/10.1128/genomeA.00281-18

21. **Duong, T.A., Aylward, J., Ametrano, C.G. et al. (2021).** IMA Genome - F15. Draft genome assembly of *Fusarium pilosicola, Meredithiella fracta, Niebla homalea, Pyrenophora teres* hybrid WAC10721, and *Teratosphaeria viscida. IMA Fungus 12, 30.* https://doi.org/10.1186/s43008-021-00077-9

22. **Greshake Tzovaras, B., Segers, F. H. I. D., Bicker, A., Dal Grande, F., Otte, J., Anvar, S. Y., Hankeln, T., Schmitt, I., & Ebersberger, I. (2020).** What Is in *Umbilicaria pustulata*? A Metagenomic Approach to Reconstruct the Holo-Genome of a Lichen. *Genome biology and evolution, 12(4), 309–324.* https://doi.org/10.1093/gbe/evaa049

23. **Al Grande F., Meiser A., Greshake Tzovaras B., Otte J., Ebersberger I., Schmitt I. (2018).** The draft genome of the lichen-forming fungus *Lasallia hispanica* (Frey) Sancho & A. Crespo. The Lichenologist. 2018;50(3):329-340. https://doi.org/10.1017/S002428291800021X

24. **Mead, O. L., & Gueidan, C. (2020).** Complete Genome Sequence of an Australian Strain of the Lichen-Forming Fungus *Endocarpon pusillum* (Hedwig). *Microbiology resource announcements, 9(50), e01079-20.* https://doi.org/10.1128/MRA.01079-20

25. **Wang, Y., Yuan, X., Chen, L., Wang, X., & Li, C. (2018).** Draft Genome Sequence of the Lichen-Forming Fungus *Ramalina intermedia* Strain YAF0013. *Genome announcements, 6(23), e00478-18.* https://doi.org/10.1128/genomeA.00478-18

26. **Calchera, A., Dal Grande, F., Bode, H. B., & Schmitt, I. (2019).** Biosynthetic Gene Content of the 'Perfume Lichens' *Evernia prunastri* and *Pseudevernia furfuracea. Molecules (Basel, Switzerland), 24(1), 203.* https://doi.org/10.3390/molecules24010203

27. **Yuan, X., Li, Y., Luo, T., Bi, W., Yu, J., & Wang, Y. (2023).** Genomic Analysis of the *Xanthoria elegans* and Polyketide Synthase Gene Mining Based on the Whole Genome. *Mycobiology, 51(1), 36–48.* https://doi.org/10.1080/12298093.2023.2175428

28. **Pizarro, D., Divakar, P. K., Grewe, F., Crespo, A., Dal Grande, F., & Lumbsch, H. T. (2020).** Genome-Wide Analysis of Biosynthetic Gene Cluster Reveals Correlated Gene Loss

5

with Absence of Usnic Acid in Lichen-Forming Fungi. *Genome biology and evolution, 12(10),* 1858–1868. https://doi.org/10.1093/gbe/evaa189

29. **Kim, W., Liu, R., Woo, S., Kang, K. B., Park, H., Yu, Y. H., Ha, H. H., Oh, S. Y., Yang, J. H., Kim, H., Yun, S. H., & Hur, J. S. (2021).** Linking a Gene Cluster to Atranorin, a Major Cortical Substance of Lichens, through Genetic Dereplication and Heterologous Expression. *mBio, 12(3),* e0111121. https://doi.org/10.1128/mBio.01111-21

30. **Kealey, J. T., Craig, J. P., & Barr, P. J. (2021).** Identification of a lichen depside polyketide synthase gene by heterologous expression in *Saccharomyces cerevisiae*. *Metabolic engineering communications, 13,* e00172. https://doi.org/10.1016/j.mec.2021.e00172

31. **Mosunova, O. V., Navarro-Muñoz, J. C., Haksar, D., van Neer, J., Hoeksma, J., den Hertog, J., & Collemare, J. (2022).** Evolution-Informed Discovery of the Naphthalenone Biosynthetic Pathway in Fungi. *mBio, 13(3),* e0022322. https://doi.org/10.1128/mbio.00223-22

32. **Kautsar, S. A., Blin, K., Shaw, S., Navarro-Muñoz, J. C., Terlouw, B. R., van der Hooft, J. J. J., van Santen, J. A., Tracanna, V., Suarez Duran, H. G., Pascal Andreu, V., Selem-Mojica, N., Alanjary, M., Robinson, S. L., Lund, G., Epstein, S. C., Sisto, A. C., Charkoudian, L. K., Collemare, J., Linington, R. G., Weber, T., Medema, M. H. (2020).** MIBiG 2.0: a repository for biosynthetic gene clusters of known function. *Nucleic acids research, 48(D1),* D454–D458. https://doi.org/10.1093/nar/gkz882

33. **Jørgensen, S. H., Frandsen, R. J., Nielsen, K. F., Lysøe, E., Sondergaard, T. E., Wimmer, R., Giese, H., & Sørensen, J. L. (2014).** *Fusarium graminearum* PKS14 is involved in orsellinic acid and orcinol synthesis. *Fungal genetics and biology : FG & B, 70,* 24–31. https://doi.org/10.1016/j.fgb.2014.06.008

34. **Yang, L., Zhuang, Q., Wu, M., Long, H., Lin, C., Lin, M., & Ke, F. (2021).** Electrochemical-induced hydroxylation of aryl halides in the presence of Et 3 N in water. *Organic & Biomolecular Chemistry, 19(29),* 6417–6421.

35. **Sanchez, J. F., Chiang, Y. M., Szewczyk, E., Davidson, A. D., Ahuja, M., Elizabeth Oakley, C., Woo Bok, J., Keller, N., Oakley, B. R., & Wang, C. C. (2010).** Molecular genetic analysis

of the orsellinic acid/F9775 gene cluster of *Aspergillus nidulans*. *Molecular bioSystems, 6(3),* 587–593. https://doi.org/10.1039/b904541d

36. **Packter, N. M., & Steward, M. W. (1967).** Studies on the biosynthesis of phenols in fungi. Biosynthesis of 3,4-dimethoxy-6-methyltoluquinol and gliorosein in *Gliocladium roseum* I.M.I. 93 065. *The Biochemical journal, 102(1),* 122–132. https://doi.org/10.1042/bj1020122

37. **Throckmorton, K., Lim, F. Y., Kontoyiannis, D. P., Zheng, W., & Keller, N. P. (2016).** Redundant synthesis of a conidial polyketide by two distinct secondary metabolite clusters in *Aspergillus fumigatus*. *Environmental microbiology, 18(1),* 246–259. https://doi.org/10.1111/1462-2920.13007

38. **Gomes, A. T., Honda, N. K., Roese, F. M., Muzzi, R. M., & Sauer, L. (2006).** Cytotoxic activity of orsellinates. *Zeitschrift fur Naturforschung. C, Journal of biosciences, 61(9-10),* 653–657. https://doi.org/10.1515/znc-2006-9-1007

39. **Schmeda-Hirschmann, G., Tapia, A., Lima, B., Pertino, M., Sortino, M., Zacchino, S., Arias, A. R., & Feresin, G. E. (2008).** A new antifungal and antiprotozoal depside from the Andean lichen *Protousnea poeppigii*. *Phytotherapy research : PTR, 22(3),* 349–355. https://doi.org/10.1002/ptr.2321

40. **Ingólfsdóttir, K., Bloomfield, S. F., & Hylands, P. J. (1985).** In vitro evaluation of the antimicrobial activity of lichen metabolites as potential preservatives. *Antimicrobial agents and chemotherapy, 28(2),* 289–292. https://doi.org/10.1128/AAC.28.2.289

41. **Do, T. H., Duong, T. H., Nguyen, H. T., Nguyen, T. H., Sichaem, J., Nguyen, C. H., Nguyen, H. H., & Long, N. P. (2022).** Biological Activities of Lichen-Derived Monoaromatic Compounds. *Molecules (Basel, Switzerland), 27(9),* 2871. https://doi.org/10.3390/molecules27092871

42. **Gomez, A.T., Smania Junior, A., Seidel, S., Albino Smania, E.F., Hinda E.K., Roese F.M., Muzzi R.M. (2003).** Antibacterial activity of orsellinates. Brazilian Journal of Microbiology, 34:194-196. https://doi.org/10.1590/S1517-83822003000300002

43. **Zorrilla, J. G., D'Addabbo, T., Roscetto, E., Varriale, C., Catania, M. R., Zonno, M. C., Altomare, C., Surico, G., Nimis, P. L., & Evidente, A. (2022).** Antibiotic and

5

Nematocidal Metabolites from Two Lichen Species Collected on the Island of Lampedusa
(Sicily). *International journal of molecular sciences, 23(15), 8471.*
https://doi.org/10.3390/ijms23158471

44 . **Yu C-L, Wu H, Chen Y-P, Chen F, Wang G-H. (2023).** Orcinol Inhibits Melanogenesis in
B16F10 Cells via the Upregulation of the MAPK/ERK Signaling Pathway. *Natural Product
Communications. 2023;18(3).* https://doi.org/10.1177/1934578X231156704

45 . **Lajis, A. F., Hamid, M., & Ariff, A. B. (2012).** Depigmenting effect of Kojic acid esters in
hyperpigmented B16F1 melanoma cells. *Journal of biomedicine & biotechnology, 2012,*
952452. https://doi.org/10.1155/2012/952452

46 . **Smith, C. J., O'Hare, K. B., & Allen, J. C. (1988).** Selective cytotoxicity of hydroquinone
for melanocyte-derived cells is mediated by tyrosinase activity but independent of melanin
content. *Pigment cell research, 1(6), 386–389.*
https://doi.org/10.1111/j.1600-0749.1988.tb00140.x

47 . **Stanke, M., & Morgenstern, B. (2005).** AUGUSTUS: a web server for gene prediction in
eukaryotes that allows user-defined constraints. *Nucleic acids research, 33(Web Server
issue), W465–W467.* https://doi.org/10.1093/nar/gki458

48 . **Sievers, F., & Higgins, D. G. (2018).** Clustal Omega for making accurate alignments of many
protein sequences. *Protein science : a publication of the Protein Society, 27(1), 135–145.*
https://doi.org/10.1002/pro.3290

49 . **Capella-Gutiérrez, S., Silla-Martínez, J. M., & Gabaldón, T. (2009).** trimAl: a tool for
automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics (Oxford,
England), 25(15), 1972–1973.* https://doi.org/10.1093/bioinformatics/btp348

50 . **Anisimova, M., Gil, M., Dufayard, J. F., Dessimoz, C., & Gascuel, O. (2011).** Survey of
branch support methods demonstrates accuracy, power, and robustness of fast likelihood-
based approximation schemes. *Systematic biology, 60(5), 685–699.*
https://doi.org/10.1093/sysbio/syr041

51 . **Guindon, S., Dufayard, J. F., Lefort, V., Anisimova, M., Hordijk, W., & Gascuel, O. (2010).**
New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the
performance of PhyML 3.0. *Systematic biology, 59(3), 307–321.*
https://doi.org/10.1093/sysbio/syq010

52 . **Letunic, I., & Bork, P. (2021).** Interactive Tree Of Life (iTOL) v5: an online tool for
phylogenetic tree display and annotation. *Nucleic acids research, 49(W1), W293–W296.*
https://doi.org/10.1093/nar/gkab301

53 . **Zhang, Y., Werling, U., & Edelmann, W. (2014).** Seamless Ligation Cloning Extract (SLiCE)
cloning method. *Methods in molecular biology (Clifton, N.J.), 1116, 235–244.*
https://doi.org/10.1007/978-1-62703-764-8_16

54 . **Tripp, J. D., Lilley, J. L., Wood, W. N., & Lewis, L. K. (2013).** Enhancement of plasmid DNA
transformation efficiencies in early stationary-phase yeast cell cultures. Yeast (Chichester,
England), 30(5), 191–200. https://doi.org/10.1002/yea.2951

55 . **Musharraf, S.G., Kanwal, N., Thadhani, V.M., & Choudhary, M I. (2015).** Rapid
identification of lichen compounds based on the structure–fragmentation relationship using
ESI-MS/MS analysis. Analytical Methods, 7(15), 6066-6076. DOI : 10.1039/C5AY01091H

5

Supplementary materials

Supplementary Table S01.

^{1}H NMR data for compound **1**, orcinol.

Position	δ_H 600 MHz DMSO-d$_6$	Literature data[a] δ_h 400 MHz DMSO-d$_6$
A	2.10, s, 3H	2.10, s, 3H
B	5.99, t (2.2 Hz), 1H	5.97, t, 1H
C	6.01, d (2.1 Hz), 1H	6.05, d, 1H
—OH	9.01, s, 2H	9.03, s, 2H

[a](34)

^{1}H NMR (600 MHz, DMSO) δ 2.10 (s, 3H), 5.99 (t, *J* = 2.2 Hz, 1H), 6.01 (d, *J* = 2.1 Hz, 1H), 9.01 (s, 1H).

Supplementary Table S02.

^{13}C NMR data for compound **1**, orcinol.

Position	δ_c 150 MHz DMSO-d$_6$	Literature data[a] δ_c 100 MHz DMSO-d$_6$
1	21.22	21.22
2	99.70	99.75
3-4	107.03	107.09
5	139.11	139.19
6-7	158.21	158.22

[a](34)

^{13}C NMR (151 MHz, DMSO) δ 21.22, 99.70, 107.03, 139.11, 158.21.

Supplementary Data Set 01. Curated protein sequences of reference and newly identified nrPKSs from genomes of LFF, alignment and tree files. Available on request.

Supplementary Table S03.

^{1}H NMR data for compound **2**, orsellinic acid.

Position	δ_H 600 MHz DMSO-d$_6$	Literature data[b] 300 MHz DMSO
A	2.39, s, 3H	2.39, s, 3H
B	6.12, d (2.4 Hz), 1H	6.10, d, 1H
C	6.18, dd (2.5 Hz), 1H	6.16, d, 1H

^{1}H NMR (600 MHz, DMSO-d$_6$) δ 2.39 (s, 3H), 6.12 (d, *J* = 2.4 Hz,, 1H), 6.18 (d, *J* = 2.5, 0.8 Hz, 1H), 10.05 (s, 1H), 12.00 (s, 1H), 13.33 (s, 1H)

[b] Musharraf *et al.*, 2015 (55)

Supplementary Table S04.

^{13}C NMR data for compound **2**, orsellinic acid.

Position	δ_c 151 MHz DMSO-d$_6$	Literature data[b] 75 MHz DMSO
1	23.37	23.4
2	100.39	100.3
3	104.71	104.9
4	110.90	110.7
5	142.80	142.6
6	161.88	161.5
7	164.35	164.2
8	173.17	172.9

^{13}C NMR (151 MHz, DMSO-d$_6$, δ 23.37, 100.39, 104.71, 110.90, 142.80, 161.88, 164.35, 173.17

[b] Musharraf *et al.*, 2015 (55)

5

Supplementary Table S05.

Genetic background of transformants and compounds identified in their cultural liquid.

ID	Genotype	Genotype (Based on Expression)	Compounds identified
B	AmoA	AmoA	1, 2
C	AmoA, AmoE	AmoA, AmoE	1, 2
D	AmoA, AmoB, AmoC, AmoD, AmoE	AmoA, AmoB, ,AmoD, AmoE	1, 2
E	AmoA, AmoB, AmoC, AmoD	AmoA, AmoB, AmoC, AmoD	3, 4, 2
F	AmoA, AmoB, AmoC, AmoD, AmoE	AmoA, AmoB, AmoC, ,AmoE	2, 4, 5, 6, 7, 8, 9

Supplementary Table S06.

Primers used in this study.

Primers used for amplification of genes from Twist plasmids for cloning

Number	Name	Sequence
01	AmoB-F	TCTTTCAACACAAGATCCCAAAGTCAAAGGATGCCCCTCTCTATCGCGATCGTCGGC
02	AmoB-R	CATTCTATGCGTTATGAACATGTTCCCTGGTCACGCCGTAGTCTCCTGTCCATAAGCCG
03	AmoC-F	AACAGCTACCCCGCTTGAGCAGACATCACCGGATGTCGCTCGAGTCTCTAGCATCCACCATC
04	AmoC-R	ACGACAATGTCCATATCATCAATCATGACCGGTCACTGAAATTGGAACTCTAACAACGACTG
05	AmoD-F	TCGACTGACCAATTCCGCAGCTCGTCAAAGGATGGCGCTCGAATTAGCCTCCACCTCTGAA
06	AmoD-R	AGGTTGGCTGGTAGACGTCATATAATCATACGGCTACAACTTATTCACCAACCCCCCAGTTCGC
07	AmoE-F	ACTTTGTACAAAAAAGCAGGCTCCGCGGCCACCTCCTTCCAACACCGATCCAGAATTG
08	AmoE-R	TGGGTCGGCGCGCCTGTTTAAACTGCGGCCCTACAGCTTCATGCCCGCATTCGTCTCCGC

Primers used for screening of gene expression

Number	Name	Sequence
09	E-AmoA-F	CGCACTCGATGAAAAGCTCG
10	E-AmoA-R	TCTGCCTTCGGAATCAACCC
11	E-AmoB-F	AGCTCGATGCGGAAGAG
12	E-AmoB-R	CCACCCTGACGCTAGACATC
13	E-AmoC-F	CCGAAGACATCCTCTACCGC
14	E-AmoC-R	GCCGCCAACATGAAGTGATC
15	E-AmoD-F	CCACCACATCTTCACCACGA
16	E-AmoD-R	GAGGGAGGGTTCAGCTAGAG
17	E-AmoE-F	CTTCCAACACCGATCCAGAATTG
18	E-AmoE-R	CTTTGTGCAGCTCAGAGGTT
19	*A. oryzae* H2B F	GCTGCTGCCTCTGGTGAC
20	*A. oryzae* H2B R	GTGCCTTCCGACACAGCATGC
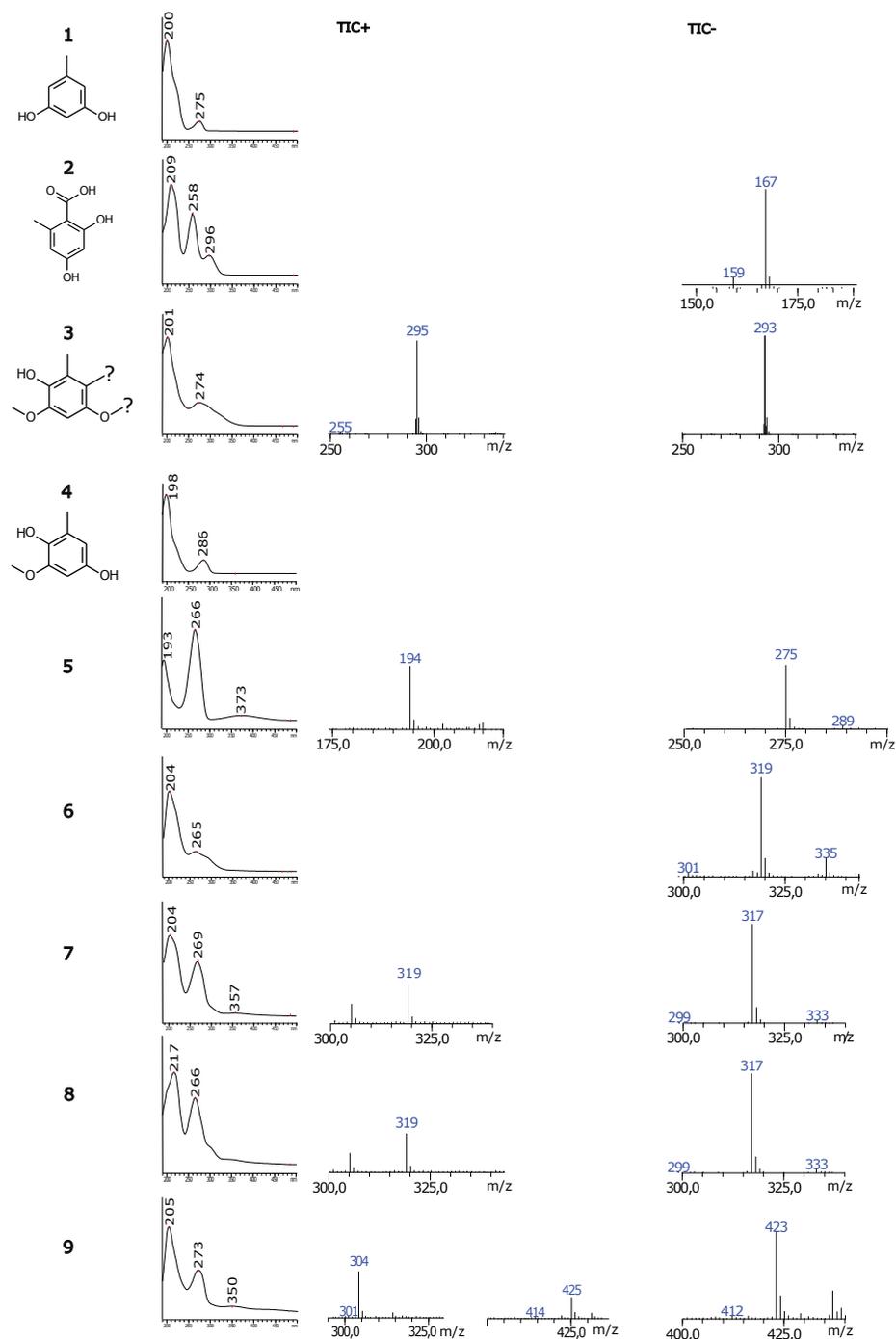
Supplementary Figure 01.

Structures, UV maxima and MS data for compounds identified in this study.

A

TOF MS ES+
3.18e4
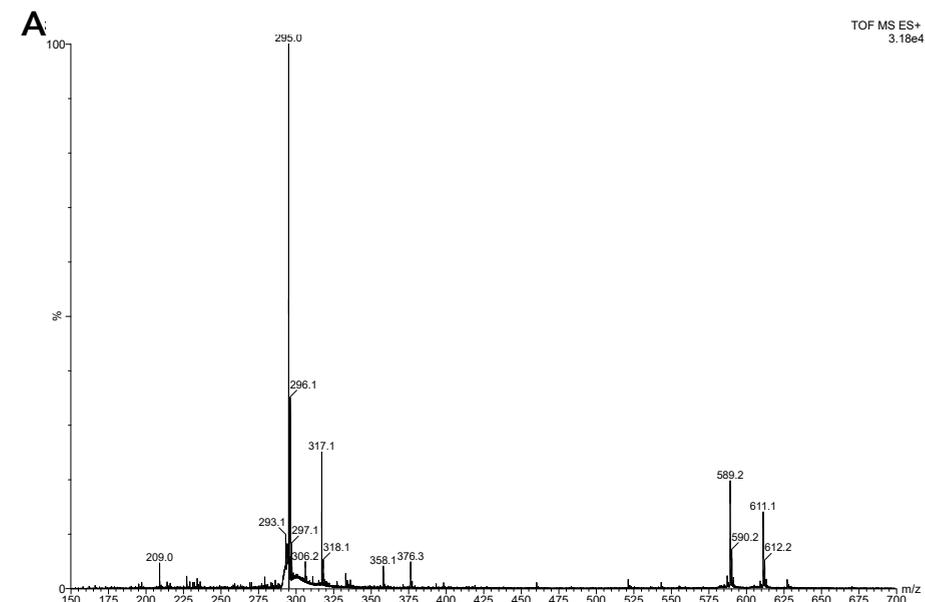
B

Single Mass Analysis
Tolerance = 10.0 mDa   /   DBE: min = -0.5, max = 50.0
Isotope cluster parameters: Separation = 1.0   Abundance = 1.0%

Monoisotopic Mass, Even Electron Ions
870 formula(e) evaluated with 31 results within limits (up to 50 closest results for each mass)

Minimum:				-0.5		
Maximum:		10.0	5.0	50.0		
Mass	Calc. Mass	mDa	PPM	DBE	Score	Formula
317.0621	317.0570	5.1	16.0	4.5	30	C5 H10 N8 O7 Na
	317.0610	1.1	3.4	8.5	19	C10 H10 N6 O5 Na
	317.0637	-1.6	-5.1	7.5	15	C14 H14 O7 Na
	317.0709	-8.8	-27.9	3.5	25	C8 H14 N4 O8 Na
	317.0597	2.4	7.6	3.5	22	C9 H14 N2 O9 Na
	317.0651	-3.0	-9.3	12.5	12	C15 H10 N4 O3 Na
	317.0538	8.3	26.1	12.5	10	C16 H10 N2 O4 Na
	317.0696	-7.5	-23.6	9.5	28	C5 H6 N14 O2 Na
	317.0584	3.7	11.8	9.5	26	C6 H6 N12 O3 Na
	317.0691	-7.0	-22.0	16.5	2	C20 H10 N2 O Na
	317.0552	6.9	21.9	17.5	8	C17 H6 N6 Na
	317.0578	4.3	13.4	16.5	1	C21 H10 O2 Na
	317.0624	-0.3	-0.9	13.5	17	C11 H6 N10 O Na
	317.0661	-4.0	-12.7	10.5	9	C16 H13 O7
	317.0562	5.9	18.5	15.5	3	C18 H9 N2 O4
	317.0675	-5.4	-16.9	15.5	7	C17 H9 N4 O3
	317.0720	-9.9	-31.2	1.5	23	C9 H17 O12
	317.0693	-7.2	-22.8	2.5	31	C5 H13 N6 O10
	317.0707	-8.6	-27.0	7.5	27	C6 H9 N10 O6
	317.0594	2.7	8.5	7.5	24	C7 H9 N8 O7
	317.0648	-2.7	-8.5	16.5	13	C13 H5 N10 O
	317.0535	8.6	27.0	16.5	11	C14 H5 N8 O2
	317.0603	1.8	5.8	19.5	4	C23 H9 O2
	317.0715	-9.4	-29.6	19.5	5	C22 H9 N2 O
	317.0608	1.3	4.2	12.5	20	C8 H5 N12 O3
	317.0634	-1.3	-4.2	11.5	16	C12 H9 N6 O5
	317.0720	-9.9	-31.2	12.5	21	C7 H5 N14 O2
	317.0522	9.9	31.2	11.5	14	C13 H9 N4 O6
	317.0576	4.5	14.3	20.5	6	C19 H5 N6
	317.0621	0.0	0.0	6.5	18	C11 H13 N2 O9

Supplementary Figure 02.

A: HRMS of compound **3**. Observed masses: 295 [M+H], 317 [M+Na], 589 [2xM+H], 611 [2xM+Na].

B: Elemental composition report.
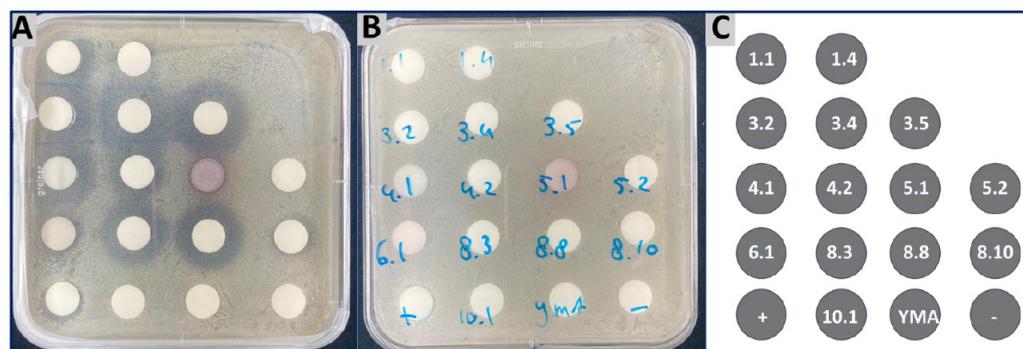
Supplementary Figure 03.

Disc diffusion assay of crude extracts from *A. oryzae* transformants against *Candida albicans*, incubation for 1 day at 35°C. Disk diffusion assays against *C. albicans* were performed using extracts on disks derived from A) the initial metabolite extraction and B) the subsequent metabolite extraction. C) An overview of the extracted compounds on disks is provided. 5.1 correspond to *AmoA+AmoB+AmoC+AmoD* co-expression, 4.1 and 4.2: *AmoA+AmoE*, 6.1: *AmoA*, 8.3, 8.8 nd 8.8 to *AmoE* expressed alone, 10.1 to *AmoB+AmoC+AmoD*; + denotes the positive control (amphotericin B), YMA (extract from uninoculated YMA medium), and - indicates the negative control (acetonitrile).
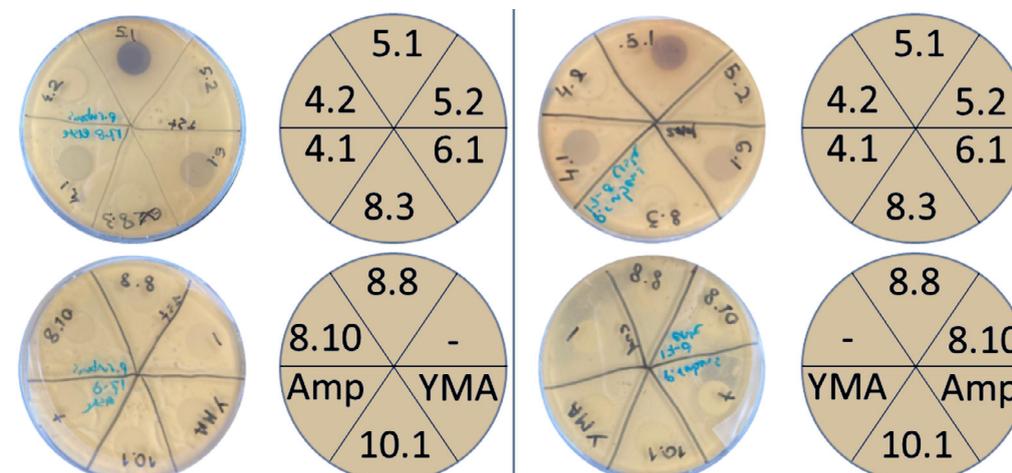
Supplementary Figure 04.

Biological assays of crude extracts from *A. oryzae* transformants against *Penicillium rubens* for 2 days at room temperature. Disk diffusion assays were performed against *P. rubens* using extracts from A) the first metabolite extraction and B) the second metabolite extraction. Extracts were applied onto disks, layout of disc distribution is represented in blue. 5.1 correspond to *AmoA+AmoB+AmoC+AmoD* co-expression, 4.1 and 4.2: *AmoA+AmoE*, 6.1: *AmoA*, 8.3, 8.8 nd 8.8 to *AmoE* expressed alone, 10.1 to *AmoB+AmoC+AmoD*. The controls: Amp (50 mg/mL ampicillin in LB), FZ (fungizone, positive control), YMA (extract from uninoculated YMA medium), and - (acetonitrile).

5
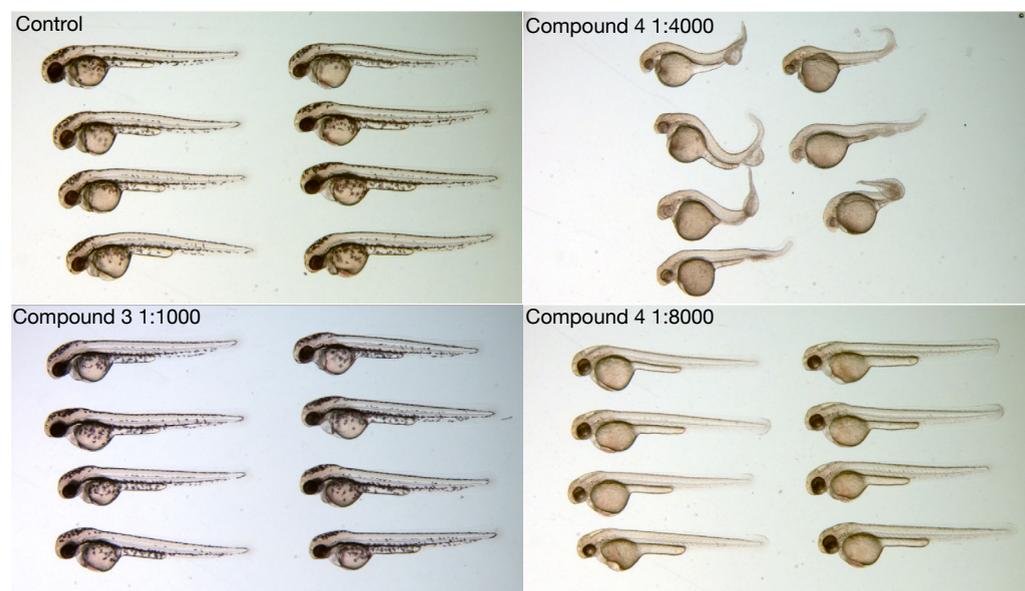
Supplementary Figure 05.

Zebrafish (*D. rerio*) embryos 48 hpf exposed to dilutions of compounds **3** and **4**, or control (solvent). Compound **3** does not demonstrate demelanizing activity at 1:1000 dilution (left). Compound **4** does not demonstrate protrusion of the tail at 1:8000 dilution, but still exhibits demelanizing activity (right).

Product 1: orcinol

¹H NMR (600 MHz, DMSO)

¹H NMR (600 MHz, DMSO) δ 9.00 (s, *OH* x2), 6.01 (d, *J* = 2.1 Hz, 2H), 5.99 (t, *J* = 2.2 Hz, 1H), 2.10 (s,

¹³C NMR (151 MHz, DMSO)

¹³C NMR (151 MHz, DMSO) δ 158.21, 139.11, 107.03, 99.70, 21.22.
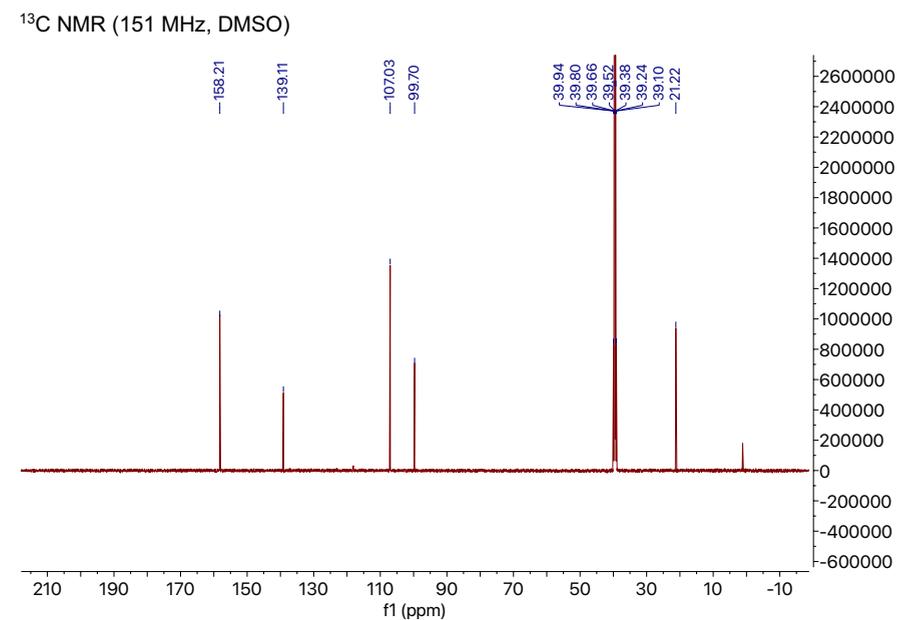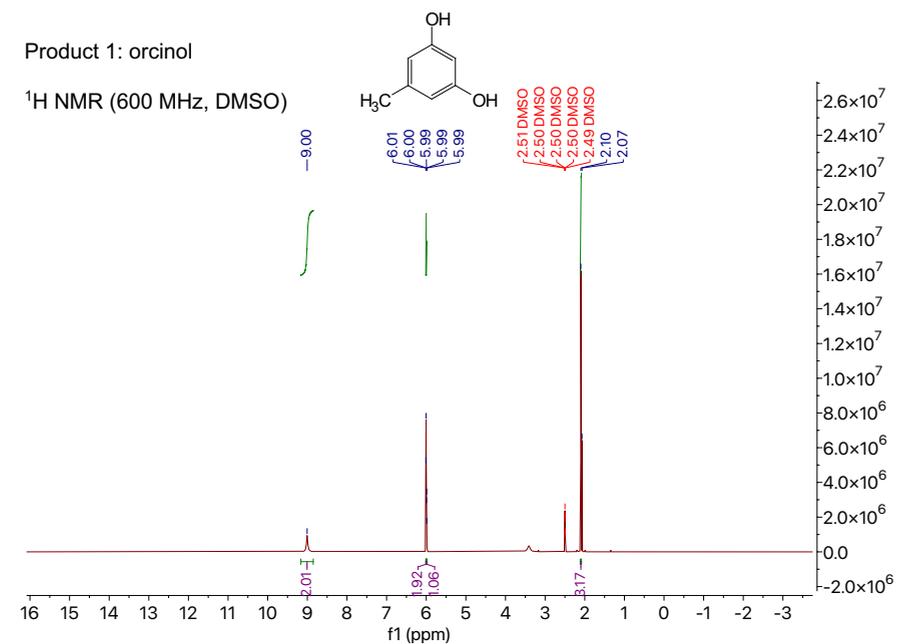
Supplementary Data Set S02.

NMR data for compounds **1**, **2**, **4**.

Product 2: orsellinic acid

^{1}H NMR (600 MHz, DMSO)
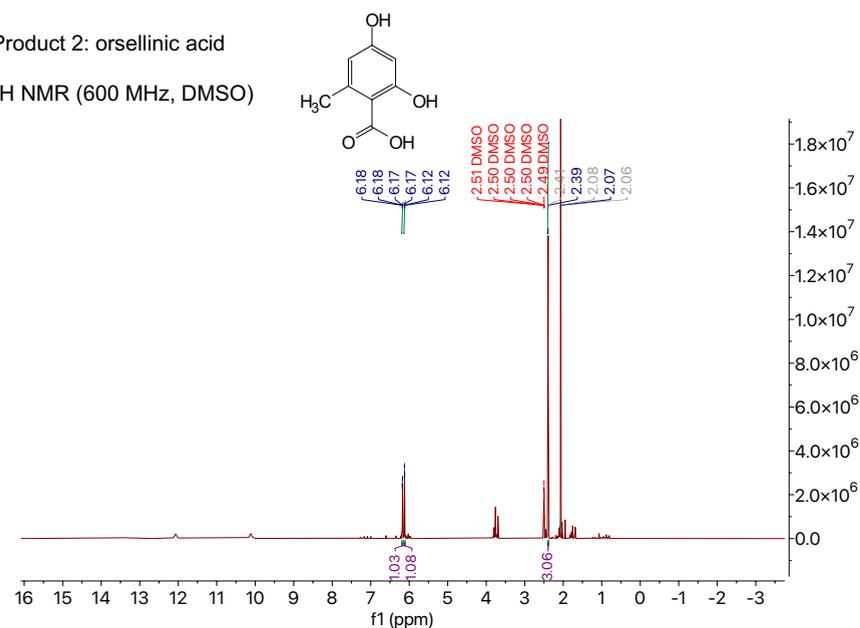
^{1}H NMR (600 MHz, DMSO) δ 6.18 (dd, *J* = 2.4, 0.9 Hz, 1H), 6.12 (d, *J* = 2.4 Hz, 1H), 2.39 (s, 3H).

^{13}C NMR (151 MHz, DMSO)
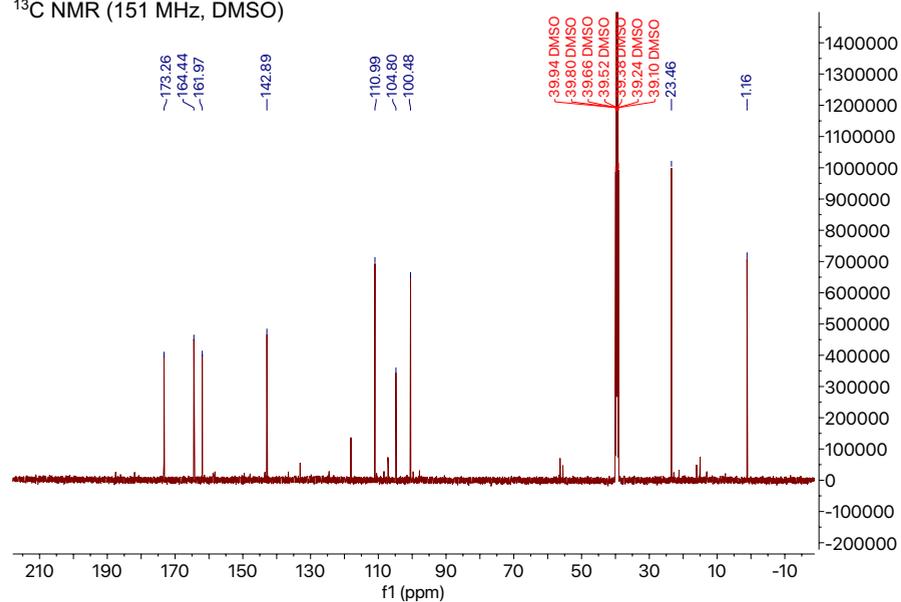
^{13}C NMR (151 MHz, DMSO) δ 173.26, 164.44, 161.97, 142.89, 110.99, 104.80, 100.48, 23.46.

Product 4: 1,4-dihydroxy-2-methoxy-6-methylbenzene

^{1}H NMR (600 MHz, DMSO)
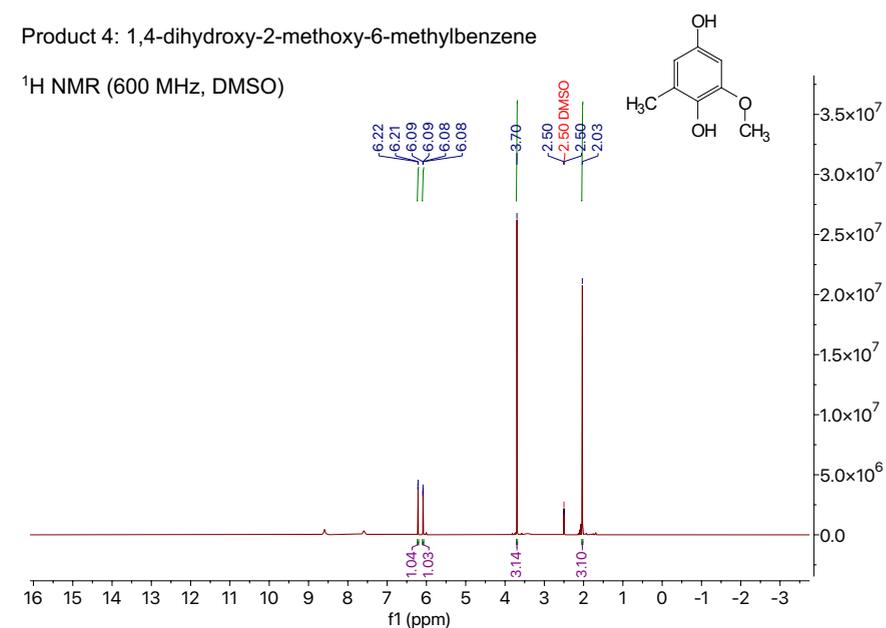
^{1}H NMR (600 MHz, DMSO) δ 6.21 (d, *J* = 2.7 Hz, 1H), 6.09 (dd, *J* = 2.8, 0.8 Hz, 1H), 3.70 (s, 3H), 2.03 (s, 3H).

^{13}C NMR (151 MHz, DMSO)
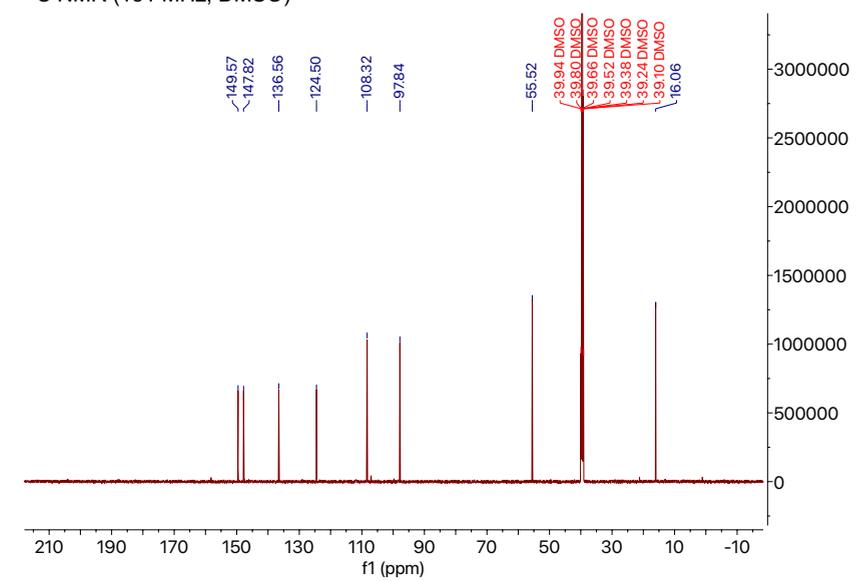
^{13}C NMR (151 MHz, DMSO) δ 149.57, 147.82, 136.56, 124.50, 108.32, 97.84, 55.52, 16.06.

Supplementary Data Set S02.

NMR data for compounds **1**, **2**, **4 (continued)**.
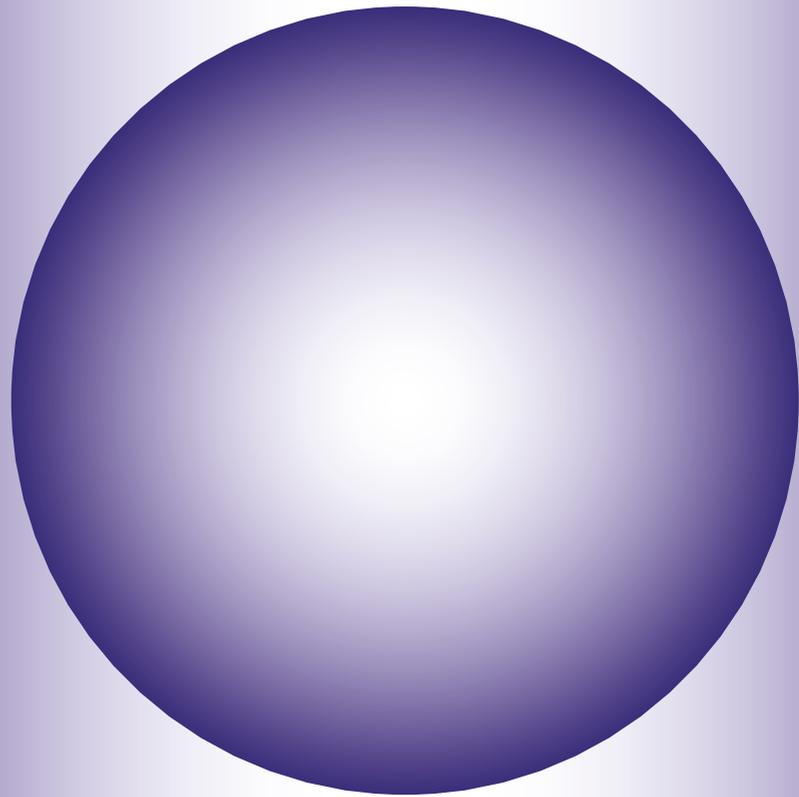
6. General discussion and summary

The present thesis aimed at the discovery of novel chemistries derived from genomes of lichenizing fungi. Our findings highlight the vast biosynthetic potential of lichen-forming fungi and underscore the importance of genomics in unraveling the diversity and functionality of their secondary metabolites **(Chapters 2-5)**. Through phylogenetic analyses and comparative genomics, we revealed evolutionary relationships and established functional annotations for two of these gene clusters **(Chapters 3, 5)**. Our suggestions on possible roles of studied enzymes from biosynthetic pathways would hopefully contribute to a comprehensive database of reactions and associated proteins from these fungal biosynthetic pathways. This catalogue could be instrumental in biochemical engineering and the rational design of valuable molecules.

Fungi as a source of fine chemicals

Fungi exhibit a remarkable chemical diversity, and their secondary metabolites often possess unique and complex structures that are challenging to synthesize using conventional chemical methods [01]. Biological synthesis of a complete or even partial path of chemical rearrangements necessary to obtain a target molecule offers sustainability benefit over traditional, large-scale chemical processes. The microbial production of drugs such as some antibiotics, statins, insulin, and enzymes has proven to be economically favourable [02]. From the 1940's of past century until now large scale production of penicillin is achieved by fermentation of strains of *Penicillium sp.*, which is still advantageous compared to the chemical synthesis [03]. An increasing number of fungal genomes are being sequenced annually, including the efforts within the frame of the 1000 fungal genomes project (https://1000.fungalgenomes.org/). Genomic studies on biosynthetic pathways of fungi [05, 06, 07] revealed

that fungi tend to harbour more biosynthetic pathways than the reported molecules for any given species. This trend extends to entire fungal taxa, revealing an unexplored and untapped biosynthetic potential. **Chapter 2** illustrates that this potential particularly applies to Lecanoromycetes. This fungal class not only possess highest amount of BGCs predicted per genome [8], but also analyses of the fungal chemical space show that ions registered by LCMS/MS in lichen extracts [04] further confirm that the molecules produced by lichens differ from those of other fungal groups **(Fig. 01)**.
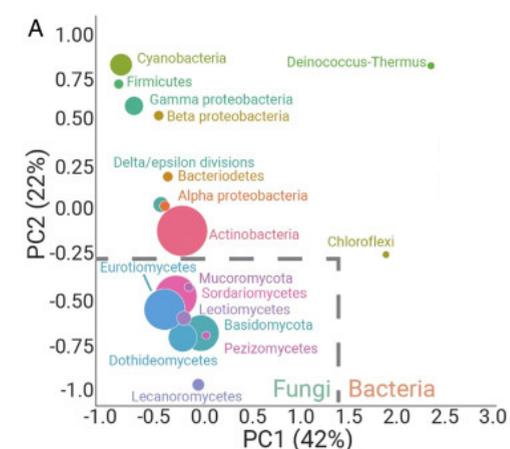
Figure 01.
(adapted from [04]). 24,595 known compounds of fungal and bacterial origin organized in PCA, coloured regions sized in proportion to the number of compounds. Natural product scaffolds are distinct for bacteria and fungi, but also they are distinct between fungal taxonomic groups.

Evolution-guided prioritization of pathways for functional studies

Given the abundance of predicted BGCs in fungal genomes, including

genomes of Lecanoromycetes, there is a necessity for strategic prioritization in selecting BGCs for functional characterization. Phylogenetic studies are often used to link a pathway to a compound in LFF [09], as limited growth of mycobionts in culture hampers genetic manipulations, which in turn prevents using conventional methods like gene deletion or overexpression for gene-to-function assignment. In this thesis, we utilized phylogeny-informed predictions to deduce the function of a non-reducing polyketide synthase (nrPKS) for dereplication of nrPKS pathways. The selected pathway was considered part of a broader set of related pathways, a gene cluster family (GCF). Our strategy extended beyond predicting what a pathway could assemble based solely on phylogeny, as it also involved functional characterization of pathway genes using heterologous expression.

We used *A. oryzae* NSAR1 as a heterolgous host. This strain has a proven track record of successful cases of usage as a heterologous host for functional characterization of SM pathways in fungi [10, 11, 12]. Additionally, a convenient toolkit for heterolgous expression has been developed for this species [13, 14]. Furthermore, *A. oryzae* NSAR1 has a silent metabolite background, which significantly simplifies identification of a newly produced compound and downstream purification. Heterologous expression as a methodology to investigate lichen pathways has proven difficult [15], and no successful expression of a pathway from LFF was achieved using *A. oryzae* as a host. To avoid possible challenges with heterolgous expression in *A. oryzae* NSAR1, we sought for orthologs of the identified lichen nrPKS within Ascomycota, retrieved entire BGCs that contain those orthologs, and chose a pathway that originates from a non-lichen genomes for heterologous expression. This resulted in the discovery of phylogenetically distinct nrPKSs that were successfully characterized and assigned to the naphthalenone pathway **(Chapters 3 and 4)** and orsellinates pathway **(Chapter 5)**. It provided

insight into the encoded potential of lichenizing fungi, enabling the putative assignment of several nrPKSs belonging to the same gene cluster family (GCF) to their respective backbones.

Evolutionary perspective on SM pathways allows dereplicating pathways efficiently, and also provides perspective on evolutionary relationship between pathways. As described in **Chapter 3**, we found that LFF nrPKSs of Group XI are positioned as a sister clade to the group V that is producing antraquinones. Group XI nrPKSS are distant to group IIa that are known to assemble ATHN and are involved in the production of melanin **(Chapter 3, Fig. 01)**. Relation to melanin pathway is also evident from the fact that Group XI BGCs, including asparvulenone pathway contain paralogs of SCD1 and THNr tailoring genes that also occur in DHN melanin pathway [16] and cladofulvin pathway [17]. In **Chapter 4**, we further speculated about possible relationship between the group XI asparvulenone pathway and the botryosphaerones pathway identified in *Neofusicoccum parvum* [18, 19]. Additionally, Kim with co-authors [20] discussed similarities between the asparvulenone pathway and the naphthoquinone cristazarin pathway of *Cladonia metacorallifera*. nrPKS of cristazarin pathway *crz7* is distinct but close to group IIa nrPKSs, unlike *APR1*. The genetic context around *crz7* is similar to *APR1* and includes an NMR-like transcriptional regulator *crz3*, O-methyltransferases *crz1* and *crz2*, an enoyl reductase *crz5*, an oxidase *crz8* and a short-chain dehydrogenase *crz9* (https://doi.org/10.1371/journal.pone.0287559). However, BLAST analysis demonstrated that none of the *crz* genes emerged as the top hits for the *APR* genes within the naphthalenone BGC, and vice versa [20]. We cautiously propose that similarity of gene context around *APR1* and *crz7* could be a result of a convergent evolution, or an ancestral duplication followed by diversification and recruitment of similar kinds of tailoring genes.

We speculate that nrPKSs of melanin, asparvulenone and cladofulvin

6

pathways are the outcomes of ancestral duplication which was followed by pathway specialization. However, the manner of recruitment of tailoring genes in different pathways remains unclear. We suggest that *APR5* and *APR6* genes are likely the result of ancestral duplications of *ARP1–ARP2* gene pair of the *A. fumigatus* DHN pathway **(Chapter 3, Fig. 03)**.

While our goal was to discover novel chemical backbones through phylogeny-informed dereplication we did find already reported backbones AT4HN **(Chapter 3)**, orcinol and orsellinic acid **(Chapter 5)**. This suggests that the chemical space of non-reduced polyketides in LFF may have been exhausted. However, our investigation of the *A. melleus* pathway in **Chapter 5** yielded a novel molecule with antimicrobial properties. Consequently, we propose a shift in focus towards exploring the functions of tailoring genes encoded within polyketide pathways. Libraries of tailoring genes acting on diverse backbones produced by fungi could be used as an alternative to combinatorial chemistry, a combinatorial biocatalysis. Apart from delivering new potentially bioactive molecules, this approach will generate knowledge on functions of diverse groups of tailoring genes, their substrate specificity, as structure-to-activity data for modified chemical scaffolds. This knowledge can be used to engineer synthetic pathways to valuable molecules in commercially established strains for microbial production of drugs.

Two products released by the same nrPKS

In **Chapter 3**, the pathway identified in *A. parvulus* genome was assigned to the production of naphthalenones asparvenone and parvulenone. nrPKS Apr1 produced the hexaketide AT4HN and a pentaketide 6,8-dihydroxy-3-methylisocoumarin simultaneously when expressed in *A. oryzae,* and we identified AT4HN to be a major product. We speculate that the deacetylating activity of the TE domain of Apr1 coupled with spontaneous C–O cyclization is a possible mechanism underlying the formation of 6,8-dihydroxy-3-methylisocoumarin. Chain length variation occurs for fungal nrPKSs. During heterologous expression of *TerA* of *Aspergillus terreus* in *Aspergillus niger* FGSC A1144, the nrPKS under control of P*amyB* promoter simultaneously produced three polyketides with varying chain length: orsellinic acid, 6,7-dihydroxymellein, and 4-hydroxy-6-methylpyranone [21]. A study by Watanabe and Ebizuka [22] demonstrated that Pks1 of *Colletotrichum lagenarium* is producing the tetraketide orsellinic acid, two pentaketides α-acetylorsellinic acid and THN, and a hexaketide ATHN when heterologously expressed in *Aspergillus oryzae*. It was hypothesized that during the condensation reaction, TE domain is performing interception of the polyketide chain from the acyl carrier protein (ACP) domain, resulting in polyketides with varying chain length [22]. Domain swapping studies on the Pks1 gene of *C. lagenarium* have demonstrated that Pks1 is a hexaketide ATHN synthase with a bifunctional TE domain. Deacetylation coupled with cyclization by this domain releases a pentaketide THN, but not a pyrone. At the same time, samples with truncated Pks1 enzyme with no TE domain did accumulate a pyrone exhibiting C–O cyclization [23]. Pyrones have been reported to be produced in smaller proportions to the main backbone by PKSs with inactivated thioesterase (TE) domains, however their polyketide chain length matches that of the main backbone [24], unlike the chain length of 6,8-dihydroxy-3-methylisocoumarin in regards to AT4HN.

Alternatively, it can be speculated that the nrPKS Apr1 may have produced 6,8-dihydroxy-3-methylisocoumarin due to the expression strength that is disproportional to the available pool of malonyl-CoA. This disproportionality resulted in the incorrect incorporation of acetyl-CoA and the

6

production of the pyrone. Increased acetyl-CoA precursor supply in *S. cerevisiae* resulted in a 60% increase in the production of 6-MSA [25]. To explore this hypothesis, Apr1 could be expressed under a promoter with reduced strength compared to PamyB, or intracellular malonyl-CoA pool availability could be optimized.

In **Chapter 5**, we consider the potential decarboxylating function of the TE domain in the *A. melleus* nrPKS AmoA. AmoA releases orcinol and orsellinic acid polyketides, which was earlier reported for nrPKS of *F. graminearum*, PKS14 [26]. However, it is hypothesized that a decarboxylase encoded in the *F. graminearum* pathway is required for decarboxylation of the orsellinic acid into orcinol. In case of *AmoA* expression, possible activity of an endogenous decarboxylase in *A. oryzae* during heterologous expression cannot be ruled out. At the same time, it was reported that orsellinic acid can be converted into orcinol spontaneously, or by supplementing cultures of *Gliocladium roseum* with orsellinic acid [27]. As with nrPKS Apr1, a possible dual function of the TE domain may explain simultaneous production of two backbones by a single nrPKS.

Heterologous expression as an efficient strategy that gives access to the encoded potential of fungi

Power and limitations of the approach.

Identifying the desired bioactive compounds not only can pose challenges, but obtaining an ample amount of biological material for the isolation and characterization of bioactive natural products may present additional difficulties [28]. Producers of compounds can lose their producing potential over time [29], or their biosynthetic machinery may not support production of high amounts of the product of interest. Additionally, not all fungi can be successfully cultivated and manipulated in laboratory conditions. Heterologous expression allows accessing to pathways from culturally non-amenable strains of fungi, and pathway refactoring enables high expression of genes in a validated heterologous host. A spectrum of hosts have been engineered for heterologous expression, such as *Saccharomyces cerevisiae* [30], *Pichia pastoris* [31], *Aspergillus niger* [32, 33], *Aspergillus oryzae* [34, 35] including *A. oryzae* NSAR1 [36], *Aspergillus nidulans* [37]. Moreover, engineering of a completely novel pathway from genetic parts of organisms from different kingdoms is a viable approach to generate molecules of interest [38]. Catalogues of tailoring genes activities could be beneficial when designing biosynthetic pathways for molecules that, for example, require to be derivatised in a non-natural way.

6

Figure 02.
Purple coloured molecule produced during heterologous expression of *A. melleus* pathway in *A. oryzae*.

Polyketides and terpenes require acetyl-CoA and malonyl-CoA as their building blocks. The cellular pool of malonyl-CoA could be rate-limiting for production of the metabolites [39]. It is plausible that under circumstances such as unnaturally high expression, PKSs may derail the assembly process by incorporating one acetyl-CoA instead of one malonyl-CoA, leading to the generation of pyrones upon completion of cyclization. When dealing with expression of NRPSs, availability of non-proteinogenic amino acids could impact efficiency of production of a target molecule.

Cloning of selected pathways into vectors may appear seemingly simple and straightforward, but could easily become complicated and delay functional studies. This challenge could be addressed by reducing the cost of gene synthesis, making it more financially feasible. Ideally, vendors could supply constructs assembled into vectors that are ready for cloning and then for transformation in reasonable time.

Artifacts during heterologous expression.

In **Chapter 4**, when we achieved heterologous expression of the entire gene set from this pathway in *A. oryzae* NSAR1, we discovered a novel derivative, asparvulenone, among other metabolites that were not further characterized. This could potentially be an artifact resulting from the disproportionally strong expression of the genes due to the refactoring of the pathway for expression in a heterologous host. While a comprehensive investigation of all metabolites produced by *A. parvulus* in a time course was not conducted, we cannot rule out the possibility that a the identified intermediate molecule, 6-methyl-1-ethylasparvulenone, may be produced in smaller quantities by *A. parvulus*. Notably, this compound has not been reported in databases like Natural Products Atlas [40], the Dictionary of Natural Products [41], and other databases that were sought. Likewise, observed production of orcinol from orsellinic acid in

Chapter 5 could be attributable to the activity of an endogenous decarboxylase of *A. oryzae* NSAR1.

Molecules produced through pathway dissection, such as by employing heterologous expression, may not necessarily represent the intended products of a given pathway. However, these molecules can still hold value as they might exhibit valuable biological activity, or pigmentation. During purification of molecules generated during heterologous expression of the *A. melleus* pathway in **Chapter 5**, we have identified a fraction containing a bright purple metabolite (Fig. 02). Unfortunately, despite bright coloration when dissolved in a solvent, the quantity of the actual molecule was minute and obtained NMR signal was not robust enough to propose any structural features for this pigmented molecule.

Would BGC mining provide us new chemistries?

Studies on SMs of fungi often report new molecules and characterize their biological activities against a few common targets. Although these studies undoubtedly have a scientific value, it is often not possible to get physical access to or reuse those molecules for assays against uncommon targets elsewhere, or physically incorporate into high-throughput screening programs. This represents a significant missed opportunity. This valuable data can be curated and compiled into a database, which can then be leveraged by artificial intelligence (AI) algorithms to analyze correlations between bioactivity and chemical structure. By identifying molecules registered as natural products and previously tested for activity, AI could suggest potential bioactive compounds for further investigation. Liu with co-authors trained a neural network on ~7500 molecules that demonstrated growth inhibiting activity towards *Acinetobacter baumannii*, a pathogenic bacteria that often presents with multi-resistance

6

[42]. This has led to the discovery of abaucin, a molecule with narrow-spectrum activity against *A. baumannii,* which structure was proposed by the neural network. Deep learning approach utilised by Wong with co-authors [43] screened curated database of 39312 molecules with reported antibiotic activity and cell toxicity, and predicted antibiotic activity and human cell toxicity for 12076365 molecules. After 283 of those molecules were physically tested, a new structural class of antibiotics was proposed based on common structural features of assayed molecules [43]. Current developments and challenges in AI-assisted natural product drug discovery are reviewed in [44].

Although we used evolution-guided approach to dereplicate biosynthetic pathways, we seem to discover pathways leading to already reported backbone. This suggests that the chemical diversity within non-reduced polyketides sourced from the genomes of lichenizing fungi is nearly exhausted. However, even known backbones can yield new molecules with potential antibiotic properties by altering their structural arrangement. Finding and utilizing a structurally less novel molecule may delay the onset of antibiotic resistance if the new derivative affects different, or even new, cellular target.

The information on pathways, activity and chemical structures of fungal natural products could be used as a training dataset for AI applications in order to identify and develop potential NP-derived bioactive structures. It could be advantageous to integrate proposals for new antibiotic structures with pathways for their production by biocatalysis. This requires a comprehensive catalog of tailoring genes with defined functions, thereby directing attention from exploring natural product backbones and mature molecules towards the chemical space of derivatives enabled by tailoring reactions. This approach essentially represents AI-assisted combinatorial chemistry *in vivo,* which has potential to deliver novel bioactive molecules sustainably, and is leveraging input curated by evolutionary pressures.

References

01. **Fazili, M. A., Bashir, I., Ahmad, M., Yaqoob, U., & Geelani, S. N. (2022).** In vitro strategies for the enhancement of secondary metabolite production in plants: a review. Bulletin of the National Research Centre, 46(1), 35. https://doi.org/10.1186/s42269-022-00717-z

02. **Gober, C. M., & Joullié, M. M. (2016).** Joining Forces: Fermentation and Organic Synthesis for the Production of Complex Heterocycles. *The Journal of organic chemistry*, *81*(21), 10136–10144. https://doi.org/10.1021/acs.joc.6b01308

03. **Zhgun AA. (2023).** Industrial Production of Antibiotics in Fungi: Current State, Deciphering the Molecular Basis of Classical Strain Improvement and Increasing the Production of High-Yielding Strains by the Addition of Low-Molecular Weight Inducers. Fermentation 9(12):1027. https://doi.org/10.3390/fermentation9121027

04. **Robey, M. T., Caesar, L. K., Drott, M. T., Keller, N. P., & Kelleher, N. L. (2021).** An interpreted atlas of biosynthetic gene clusters from 1,000 fungal genomes. *Proceedings of the National Academy of Sciences of the United States of America, 118(19), e2020230118.* https://doi.org/10.1073/pnas.2020230118

05. **Lacovelli, R., He, T., Allen, J.L. et al. (2024).** Genome sequencing and molecular networking analysis of the wild fungus *Anthostomella pinea* reveal its ability to produce a diverse range of secondary metabolites. Fungal Biol Biotechnol 11, 1 (2024). https://doi.org/10.1186/s40694-023-00170-1

06. **Tsunematsu Y. (2021).** Genomics-directed activation of cryptic natural product pathways deciphers codes for biosynthesis and molecular function. *Journal of natural medicines, 75(2), 261–274.* https://doi.org/10.1007/s11418-020-01466-x

07. **Skellam E. (2019).** Strategies for Engineering Natural Product Biosynthesis in Fungi. *Trends in biotechnology, 37(4), 416–427.* https://doi.org/10.1016/j.tibtech.2018.09.003

08. **Mosunova O., Navarro-Muñoz J.C., Collemare J. (2021).** The biosynthesis of fungal secondary metabolites: from fundamentals to biotechnological applications, p 458–476. *In Zaragoza Ó, Casadevall A (ed), Encyclopedia of mycology. Elsevier, Oxford, United*

6

Kingdom.

https://doi.org/10.1016/B978-0-12-809633-8.21072-8

09 . **Singh G. (2023).** Linking Lichen Metabolites to Genes: Emerging Concepts and Lessons from Molecular Biology and Metagenomics. *Journal of fungi (Basel, Switzerland), 9(2), 160*.

https://doi.org/10.3390/jof9020160

10 . **Jiang, L., Lv, K., Zhu, G., Lin, Z., Zhang, X., Xing, C., Yang, H., Zhang, W., Wang, Z., Liu, C., Qu, X., Hsiang, T., Zhang, L., & Liu, X. (2022).** Norditerpenoids biosynthesized by variediene synthase-associated P450 machinery along with modifications by the host cell Aspergillus oryzae. *Synthetic and systems biotechnology, 7(4), 1142–1147*.

https://doi.org/10.1016/j.synbio.2022.08.002

11 . **Tagami, K., Liu, C., Minami, A., Noike, M., Isaka, T., Fueki, S., Shichijo, Y., Toshima, H., Gomi, K., Dairi, T., & Oikawa, H. (2013).** Reconstitution of biosynthetic machinery for indole-diterpene paxilline in *Aspergillus oryzae. Journal of the American Chemical Society, 135(4), 1260–1263.* https://doi.org/10.1021/ja3116636

12 . **Han, H., Yu, C., Qi, J., Wang, P., Zhao, P., Gong, W., Xie, C., Xia, X., & Liu, C. (2023).** High-efficient production of mushroom polyketide compounds in a platform host *Aspergillus oryzae. Microbial cell factories, 22(1), 60*. https://doi.org/10.1186/s12934-023-02071-9

13 . **Pahirulzaman, K. A., Williams, K., & Lazarus, C. M. (2012).** A toolkit for heterologous expression of metabolic pathways in *Aspergillus oryzae. Methods in enzymology, 517, 241–260*.

https://doi.org/10.1016/B978-0-12-404634-4.00012-7

14 . **de Mattos-Shipley, K. M. J., Lazarus, C. M., & Williams, K. (2022).** Investigating Fungal Biosynthetic Pathways Using Heterologous Gene Expression: *Aspergillus oryzae* as a Heterologous Host. *Methods in molecular biology (Clifton, N.J.), 2489, 23–39*.

https://doi.org/10.1007/978-1-0716-2273-5_2

15 . **Bertrand RL, Sorensen JL. (2019).** Lost in translation: challenges with heterologous expression of lichen polyketide synthases. *ChemistrySelect 4:6473–6483*.

https://doi.org/10.1002/slct.201901762

16 . **Liang, Y., Xiong, W., Steinkellner, S., & Feng, J. (2018).** Deficiency of the melanin biosynthesis genes SCD1 and THR1 affects sclerotial development and vegetative growth, but not pathogenicity, in *Sclerotinia sclerotiorum. Molecular plant pathology, 19(6), 1444–1453*.

https://doi.org/10.1111/mpp.12627

17 . **Griffiths S, Mesarich CH, Saccomanno B, Vaisberg A, De Wit PJGM, Cox R, Collemare J. (2016).** Elucidation of cladofulvin biosynthesis reveals a cytochrome P450 monooxygenase required for anthraquinone dimerization. *Proc Natl Acad Sci USA 113:6851–6856*.

https://doi.org/10.1073/pnas.1603528113

18 . **Mosunova, O. V., Navarro-Muñoz, J. C., Haksar, D., van Neer, J., Hoeksma, J., den Hertog, J., & Collemare, J. (2022).** Evolution-Informed Discovery of the Naphthalenone Biosynthetic Pathway in Fungi. *mBio, 13(3), e0022322*.

https://doi.org/10.1128/mbio.00223-22

19 . **Salvatore, M.M.; Alves, A.; Andolfi, A. (2021).** Secondary Metabolites Produced by *Neofusicoccum* Species Associated with Plants: A Review. *Agriculture 2021, 11, 149*.

https://doi.org/10.3390/agriculture11020149

20 . **Paguirigan, J. A. G., Kim, J. A., Hur, J. S., & Kim, W. (2023).** Identification of a biosynthetic gene cluster for a red pigment cristazarin produced by a lichen-forming fungus *Cladonia metacorallifera. PloS one, 18(6), e0287559*.

https://doi.org/10.1371/journal.pone.0287559

21 . **Zaehle, C., Gressler, M., Shelest, E., Geib, E., Hertweck, C., & Brock, M. (2014).** Terrein biosynthesis in *Aspergillus terreus* and its impact on phytotoxicity. *Chemistry & biology, 21(6), 719–731*. https://doi.org/10.1016/j.chembiol.2014.03.010

22 . **Watanabe, A., & Ebizuka, Y. (2004).** Unprecedented mechanism of chain length determination in fungal aromatic polyketide synthases. *Chemistry & biology, 11(8), 1101–1106*.

https://doi.org/10.1016/j.chembiol.2004.05.015

23 . **Vagstad, A. L., Hill, E. A., Labonte, J. W., & Townsend, C. A. (2012).** Characterization of a fungal thioesterase having Claisen cyclase and deacetylase activities in melanin biosynthesis. *Chemistry & biology, 19(12), 1525–1534*.

https://doi.org/10.1016/j.chembiol.2012.10.002

6

24 . Crawford, J. M., & Townsend, C. A. (2010). New insights into the formation of fungal aromatic polyketides. *Nature reviews. Microbiology, 8(12), 879–889*. https://doi.org/10.1038/nrmicro2465

25 . **Wattanachaisaereekul, S., Lantz, A. E., Nielsen, M. L., & Nielsen, J. (2008).** Production of the polyketide 6-MSA in yeast engineered for increased malonyl-CoA supply. *Metabolic engineering, 10(5), 246–254.* https://doi.org/10.1016/j.ymben.2008.04.005

26 . **Jørgensen, S. H., Frandsen, R. J., Nielsen, K. F., Lysøe, E., Sondergaard, T. E., Wimmer, R., Giese, H., & Sørensen, J. L. (2014).** *Fusarium graminearum* PKS14 is involved in orsellinic acid and orcinol synthesis. *Fungal genetics and biology : FG & B, 70, 24–31.* https://doi.org/10.1016/j.fgb.2014.06.008

27 . **Packter, N. M., & Steward, M. W. (1967).** Studies on the biosynthesis of phenols in fungi. Biosynthesis of 3,4-dimethoxy-6-methyltoluquinol and gliorosein in *Gliocladium roseum I.M.I. 93 065. The Biochemical journal, 102(1), 122–132.* https://doi.org/10.1042/bj1020122

28 . **Cragg, G. M., Schepartz, S. A., Suffness, M., & Grever, M. R. (1993).** The taxol supply crisis. New NCI policies for handling the large-scale production of novel natural product anticancer and anti-HIV agents. *Journal of natural products, 56(10), 1657–1668.* https://doi.org/10.1021/np50100a001

29 . **Danner, C., Mach, R. L., & Mach-Aigner, A. R. (2023).** The phenomenon of strain degeneration in biotechnologically relevant fungi. *Applied microbiology and biotechnology, 107(15), 4745–4758.* https://doi.org/10.1007/s00253-023-12615-z

30 . **Harvey, C. J. B., Tang, M., Schlecht, U., Horecka, J., Fischer, C. R., Lin, H. C., Li, J., Naughton, B., Cherry, J., Miranda, M., Li, Y. F., Chu, A. M., Hennessy, J. R., Vandova, G. A., Inglis, D., Aiyar, R. S., Steinmetz, L. M., Davis, R. W., Medema, M. H., Sattely, E., Hillenmeyer, M. E. (2018).** HEx: A heterologous expression platform for the discovery of fungal natural products. *Science advances, 4(4), eaar5459.* https://doi.org/10.1126/sciadv.aar5459

31 . **Qian, Z., Liu, Q., & Cai, M. (2022).** Investigating Fungal Biosynthetic Pathways Using *Pichia pastoris* as a Heterologous Host. *Methods in molecular biology (Clifton, N.J.), 2489, 115–127.*

https://doi.org/10.1007/978-1-0716-2273-5_7

32 . **Geib, E., & Brock, M. (2017).** ATNT: an enhanced system for expression of polycistronic secondary metabolite gene clusters in *Aspergillus niger. Fungal biology and biotechnology, 4, 13.* https://doi.org/10.1186/s40694-017-0042-1

33 . **Richter, L., Wanka, F., Boecker, S., Storm, D., Kurt, T., Vural, Ö., Süßmuth, R., & Meyer, V. (2014).** Engineering of *Aspergillus niger* for the production of secondary metabolites. *Fungal biology and biotechnology, 1, 4.* https://doi.org/10.1186/s40694-014-0004-9

34 . **Sakai, K., Kinoshita, H., & Nihira, T. (2012).** Heterologous expression system in *Aspergillus oryzae* for fungal biosynthetic gene clusters of secondary metabolites. *Applied microbiology and biotechnology, 93(5), 2011–2022.* https://doi.org/10.1007/s00253-011-3657-9

35 . **Katsuya Gomi, Yuzuru Iimura & Shodo Hara (1987).** Integrative Transformation of *Aspergillus oryzae* with a Plasmid Containing the *Aspergillus nidulans* argB Gene, Agricultural and Biological Chemistry, 51:9, 2549-2555 https://doi.org/10.1080/00021369.1987.10868429

36 . **Jin, F. J., Maruyama, J., Juvvadi, P. R., Arioka, M., & Kitamoto, K. (2004).** Development of a novel quadruple auxotrophic host transformation system by argB gene disruption using *adeA* gene and exploiting adenine auxotrophy in *Aspergillus oryzae. FEMS microbiology letters, 239(1), 79–85.* https://doi.org/10.1016/j.femsle.2004.08.025

37 . **Roux, I., & Chooi, Y. H. (2022).** Heterologous Expression of Fungal Biosynthetic Pathways in *Aspergillus nidulans* Using Episomal Vectors. *Methods in molecular biology (Clifton, N.J.), 2489, 75–92.* https://doi.org/10.1007/978-1-0716-2273-5_5

38 . **Frandsen, R. J. N., Khorsand-Jamal, P., Kongstad, K. T., Nafisi, M., Kannangara, R. M., Staerk, D., Okkels, F. T., Binderup, K., Madsen, B., Møller, B. L., Thrane, U., & Mortensen, U. H. (2018).** Heterologous production of the widely used natural food colorant carminic acid in *Aspergillus nidulans. Scientific reports, 8(1), 12853.* https://doi.org/10.1038/s41598-018-30816-9

39 . **Yang, Y., Lin, Y., Li, L., Linhardt, R. J., & Yan, Y. (2015).** Regulating malonyl-CoA

6

metabolism via synthetic antisense RNAs for enhanced biosynthesis of natural products. *Metabolic engineering, 29, 217–226.* https://doi.org/10.1016/j.ymben.2015.03.018

40. **van Santen, J. A., Poynton, E. F., Iskakova, D., McMann, E., Alsup, T. A., Clark, T. N., Fergusson, C. H., Fewer, D. P., Hughes, A. H., McCadden, C. A., Parra, J., Soldatou, S., Rudolf, J. D., Janssen, E. M., Duncan, K. R., & Linington, R. G. (2022).** The Natural Products Atlas 2.0: a database of microbially-derived natural products. *Nucleic acids research, 50(D1), D1317–D1323.* https://doi.org/10.1093/nar/gkab941

41. **Buckingham, J. (2014).** Dictionary of natural products. https://doi.org/10.1007/978-1-4899-3314-0

42. **Liu, G., Catacutan, D. B., Rathod, K., Swanson, K., Jin, W., Mohammed, J. C., Chiappino-Pepe, A., Syed, S. A., Fragis, M., Rachwalski, K., Magolan, J., Surette, M. G., Coombes, B. K., Jaakkola, T., Barzilay, R., Collins, J. J., & Stokes, J. M. (2023).** Deep learning-guided discovery of an antibiotic targeting *Acinetobacter baumannii. Nature chemical biology, 19(11), 1342–1350.* https://doi.org/10.1038/s41589-023-01349-8

43. **Wong, F., Zheng, E. J., Valeri, J. A., Donghia, N. M., Anahtar, M. N., Omori, S., Li, A., Cubillos-Ruiz, A., Krishnan, A., Jin, W., Manson, A. L., Friedrichs, J., Helbig, R., Hajian, B., Fiejtek, D. K., Wagner, F. F., Soutter, H. H., Earl, A. M., Stokes, J. M., Renner, L. D., Collins, J. J. (2024).** Discovery of a structural class of antibiotics with explainable deep learning. *Nature, 626(7997), 177–185.* https://doi.org/10.1038/s41586-023-06887-8

44. **Mullowney, M. W., Duncan, K. R., Elsayed, S. S., Garg, N., van der Hooft, J. J. J., Martin, N. I., Meijer, D., Terlouw, B. R., Biermann, F., Blin, K., Durairaj, J., Gorostiola González, M., Helfrich, E. J. N., Huber, F., Leopold-Messer, S., Rajan, K., de Rond, T., van Santen, J. A., Sorokina, M., Balunas, M. J., Medema, M. H. (2023).** Artificial intelligence for natural product drug discovery. *Nature reviews. Drug discovery, 22(11), 895–916.* https://doi.org/10.1038/s41573-023-00774-7

Summary

Secondary metabolites (SMs) are biosynthesized by plants and microbes and often possess biological activities that can be applied in medicine and biotechnology. Increasing amount of sequenced fungal genomes and genomic studies highlighted that biosynthetic potential of fungi is not fully exploited. In **Chapters 1 and 2**, we reviewed the encoded potential of Lecanoromycetes to produce secondary metabolites (SMs), with particular expansion of encoded pathways for polyketides in their genomes.

In the current thesis, we explored the chemical space of non-reduced polyketides produced by BGCs encoded in the genomes of Lecanoromycetes. We used heterologous expression as a strategy to access the encoded potential of lichenizing fungi to produce potentially bioactive molecules. Because high number of encoded pathways per genome, we had to prioritize pathways for functional characterization using heterolgous expression. We utilized evolution-guided pathway dereplication, when we analysed phylogenetic positioning of studies nrPKSs in relation to already characterized enzymes.

Using this approach, we have identified two phylogenetic clades that are distant from the clades that contained already characterized nrPKS(s). Functional investigation of the ortholog of nrPKSs from lichenizing fungi in **Chapter 3** provided assignment of the core nrPKS Apr1 to the production of AT4HN, and the whole pathway was assigned to the production of naphthalenones. In **Chapter 4** we elucidated individual steps of the naphthalenone pathway, and found a previously unreported new derivative. We confirmed that FAD-binding oxidoreductase Apr2 is responsible for the introduction of p-hydroxygroup to the A ring of the molecule, and hypothesized that this reaction could also occur spontaneously. We observed stereoselectivity of the Apr2, which leads to the formation of a diastereomer

6

that differs to one that is hypothetically occurring due to spontaneous oxidation. We made more detailed predictions about the assembly of naphthalenones based on the obtained data.

In **Chapter 5** we applied phylogenetic dereplication to another subset of genomes of lichenizing fungi, and identified a set of three phylogenetically distant nrPKSs. We performed heterologous expression of pathway that contains close homologue nrPKS to the identified nrPKSs from lichen forming fungi. We linked nrPKS AmoA to the production of orcinol and orsellinic acid. By performing heterologous expression of the entire pathway, we obtained a molecule with antibiotic and demelanizing activity, and proposed a putative pathway.

In **Chapter 6**, we reviewed findings from the experimental chapters in a broader context and discussed future horizons for the fungal natural product research aiming at finding novel bioactive chemical moieties. We propose a focus shift towards creating a catalogue of functionally characterized tailoring genes, and suggest using tailoring genes for derivatization of backbones to generate novel potentially bioactive molecules.

In this work we have demonstrated the power of the evolution-guided investigation of pathways in order to find new chemistries, or evolutionary distinct pathways. This not only leads to the discovery of new molecules or linking pathways to already known chemistries, but also generates valuable knowledge for future development. Using this approach, we functionally assigned two pathways that belong to two distinct GCFs to their respective chemical backbone, and proposed a shift towards focusing on entire GCFs rather than individual pathways for further research. This approach streamlines the process of massive delimiting core genes to their respective backbones, saving time by avoiding exhaustive investigations of each biosynthetic pathway individually.

6
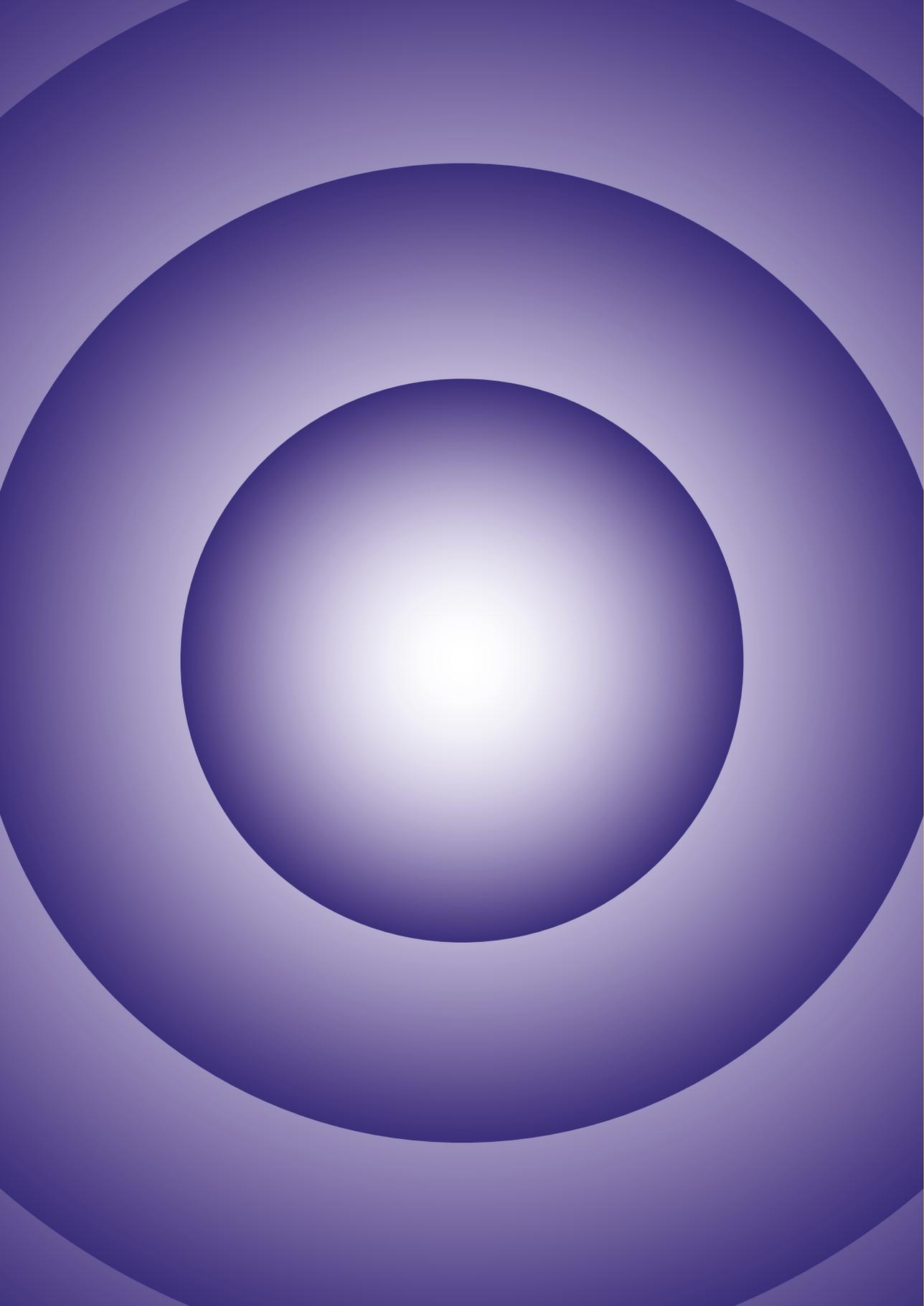
Appendix

Acknowledgements

This thesis and a PhD itself was certainly a life changing journey for me, a huge professional and personal growth. I've got to meet and work with so many incredible people, each being a gem in their very own way. I've got to work on this thesis during pandemic, and wrap up with my journey during an ongoing war. After those 6 years all these experiences distill to one single thing: gratitude to all the people I have met or kept by me these years. This work would not have happened without your contribution, support and inspiration.

First of all, I would like to thank my co-promotor **Jérôme Collemare** for giving me this breathtaking opportunity to work in the group on Fungal Natural Products. I appreciate all the guidance and support that you gave me, even if I was resistant to getting an advice at first. It was a lot of fun (and difficult sometimes) to start a first PhD project is a new group. I've learnt a lot, and enjoyed being part of the team that was so tiny at the beginning, and developed into a multicultural welcoming group with a vision on how things should be done. I especially enjoyed the philosophical turns in our discussions about the biosynthetic pathways, nature, science and the fate of scientists.

Secondly, I would like to thank **Pedro Crous** for keeping the Westerdijk Institute what it is and facilitating the Golden Mycological Triangle project and giving me an opportunity to visit Biotec, Thailand. I would also like to express my gratitude to **Noppol Kobmoo** who supervised me during my stay in Thailand.

I would also like to thank **Jeroen den Hertog** for providing a ground for a collaboration between Hubrecht and Westerdijk institutes, for providing the personnel and equipment to do the work on compounds.

I am very grateful for all the help from my supervisors **Jorge Navarro-Muñoz** and **Scott Griffiths**. **Jorge**, you gave me an indispensable experience with bioinformatics, and guided me through many critical moments in the

A

project that I won't be able to deal with on my own. I enjoyed our (sometimes long) talks about science, then coffee, then music, then something else again and back to science. I cannot express how thankful I am to have had you as a supervisor. I remember the panic I had when your contract was coming closer to its end, luckily the institute building did not fall in ruins after that, and neither did our connection. Thank you and **Pily** for feeding me delicious Mexican food. Dr **Scott**, you saved me from having an experience with local medicine by driving me to Ikea to get that neck pillow, which was the only way to get some sleep after I moved to the Netherlands. I was so inspired by how passionate you were about the natural products that I had no experience with, and the vision that you had on scientific and industrial aspects of the field. Thank you for being a supporting shoulder, I wish we had more time to do some cool fungal stuff. I really appreciate you left me all your furniture when you moved back to the UK. I still use that pillow and your protocols have been very helpful when we started heterolgous expression, and other things too. Your thesis is not too bad.

I would also like to thank the students that were doing their projects with me: **Charlotte Borrias**, **Elske Dwars**, **Ella Schunselaar**, **Janieke Klusener**. Ladies, it was a pleasure to work with all of you, and I've learnt a lot from our interactions. If I could, I'd CRISPR myself and use your genes as a repair template: persistence and ability to structure of **Charlotte**, passion and endurance of **Elske**, commitment and dedication of **Ella**, and ownership and independence of **Janieke**. I'm proud of what and how we did, especially at how you managed to do such great job at hard times like corona lockdown or scattered supervision when I got a new job. Three of you continued their journeys in science, and I'm excited to see the results of your future work. Please invite me to your PhD defences! Special thanks to **Jacque van Neer** who was a student of FNP group with whom we worked closely together, for the Apr1 PKS cloning. It is my favourite pathway now :)

The Fungal Natural Products Team: **Caroline**, **Diksha**, **Yanfang**, **Xin**, **Trung**, **Sandriele**, **Roya**, **Lhais**, **Khyati**, **Lotte**, **Thomas**, **Debbie**, and many more who were or are a part of this wonderful research group. The multicultural crowd with such different and somewhat similar visions, with our unique twists. I savoured the distance between our cultures that somehow made every connection so special. It was a pleasure to work with you all! **Caro**, thank you for your soft power that was the axis of the lab operations for a long time. I think many people appreciated you for what you are, especially during the covid times. **Diksha**, you gave me a chemical dimension to reality, from looking at these 'beautiful hydrogens' in our compounds NMR spectra to talking about formulations of skin treatments. It was nice to be silly and serious with you, and I'm happy we stay connected. **Yanfang**, I remember you just came to the Netherlands like yesterday and - puf!- you are about to finish your PhD. What a journey! I really like the work you did and I know it was not easy to get where you are now. Thank you for being unapologetically yourself! **Sandriele**, **Roya**, **Lhais**, **Khyati** - it's soo nice to see that you liked the institute and the group and you rejoined the group after finishing your projects, or their parts, or by whatever reason brought you back. **Sandri**, I promise to come to Brazil so we can make pelmeni and drink caipirinha again! **Xin**, I loved climbing with you, thank you for those evenings and talks afterwards. Maybe we can pick new sport to do together? Congratulations with your new job! **Roya**, you are so incredibly talented, I would definitely use this feature of you as a repair template for the above mentioned CRISPR fantasy :) **Trung**, it was never enough of those talks in the lab about learning russian language, your adventures in Ukraine and Chernobyl, and Vietnamese cuisine.

D wing guys: **Roland**, **Sara**, **Sandra**, **Astrid**, **Tania**, **Agata**, **Xinxin**, **Jiali**, **Mao**, **Chendo**, **Mar**, **Ola**, **Dujuan**, **Adiphol**, **Kelli**, **Gert-Jan**, **Abigail**, **Bas**, **Kimberly**, **Vicky**, **Tiziano**, and everyone I unintentionally forgot. Thank you for

A

bringing the D wing people together, we had a good time and I enjoyed how we had a good laugh together time to time, and we definitely had the most sweets in the entire institute. Never low blood sugar! **Vicky**, thank you for helping me with bioinformatics when Jorge was out of reach. **Sara**, thank you for helping me having a good start in the Westerdijk. **Roland**, so many thanks for your company, hospitality and talks about random topics. You're such a history nerd! And thank you for introducing me to you wife **Dora** :) I wish I didn't miss most of your wedding to stomach flu, it was truly awesome to be there and share the moment with you. Hope we'll travel more, like we did in India, but in good health. **Bas Viergever**, thank you for being a good friend for me and Ronnie, and for feeding us a lot of comfort food.

Our beloved DTO group, Bartin -oh sorry!- **Bart** and **Martin**. The siamese twin of the C wing, thank you for your help and for making me laugh a lot, and for organising borrels. **Martin**, thank you for not throwing me out of the window when I hugged you, I know you hate it but we both got used to each other at the end. **Tom**, thank you for the fun at the borrels, and your brutal honesty. **Alex**, thank you for the small talks, было здорово знать, что кто-то в институте поймет тебя на культурном уровне, и можно шутить родные шутки.

Phytopathology group: **Roy**, **Marcelo**, **Lin**, **Hazal**, **Margarita**, **Arien**, **Ewald**. Thank you for the talks, your enthusiasm and fun at institute outings. **Roy**, thank you so much for helping us with klussen and moving two times!

Medical mycology: **Auke**, **Ferry**, **Bert**. Thank you for allowing me to work in your lab with Candida species. **Auke**, thank you for the fun at the borrels, and talks about science. **Ferry**, thank you for being so open for a talk with beginner researchers, I think your perspective and advice on career development helped some of us with their choices.

Manon, all thankful words of all the languages are not enough to express how grateful I am for everything you did for us PhD students, for the institute, and for me personally. You are amazing! I cannot imagine the institute without you.

Magazijn, **Collection** and **Support**. Thank you for keeping us running, I appreciate how helpful you guys always were.

Kitchen: **Ralph**, **Remco**, **Kevin**, **Sander**, **Walter**, **Ronald**. Thank you for making colossal amounts of media to order in such timely manner! **Ralph**, thank you for a good laugh we always had.

Hubrecht institute people: **Jelmer**, **Wouter**. It was super fun to work together! You both were very-very helpful and I appreciate how ready you were to show me the instruments, do experiments and just talk. **Jelmer**, thank you for explaining me how to look at the NMR data, it did get better at the end :) I wish you success with the continuation of the fungal project in your lab, and all your new beginnings.

IT: Thank you guys for keeping us safe and working smoothly, and your swift help with any IT-related problem we had. **Robin**, thank you for setting up the computing server, it helped to run my phylogenetic trees in 30 min instead of three long loooooooong days before I realize I forgot to add yet another sequence to the dataset.

My colleagues from Ginkgo Bioworks Netherlands, thank you for being supportive and understanding during times I was trying to finish writing while working full time. Special thanks to **Daphne** and **Michiel** for the moral support with finishing the last writing bits. **Niki**, thank you for your helping hand when (and not only when) I was exhausted. **Brandon** and **Tim**, thank you for checking on me, I appreciate your care. Also thanks for the jokes on the suboptimal planning for finalising this thesis :) **Marian**, thank you for your cheering.

Kelli and **Bas**, thank you for such warm welcoming in Belgium in your very own botanical garden of a flat, and for introducing me to the one and only Cheddar. **Kelli**, thank you for being you and for keeping us connected!

A

Valerie and Ronnie. For meeting you two I'm grateful the most. Valerie, we met and connected when both of us were taking a steep dive into depression. Thankfully, we rented an apartment together, which was a safe haven where I knew I would be completely understood, and never judged. You are one of my most intense and happy friendship experiences. I'm surprised of how we seem to resonate in so many levels, it makes it almost surreal. I deeply value our connection, the additional perspective on things that you give me. Thank you for all the scientific talks, the figure you made and I did not use here at the end although it was so cool, but I could not finish the chapter on time. One day we'll be finally writing that paper on SM pathways in phytopathogens, or if not, we'd have a good time trying. I hope we gonna be friends for many many years to come. .

Ronnie, you caught me falling and it helped me a lot to stay afloat when I really needed it. You are kind and smart, and you are a very good scientist. I enjoy how we have endless talks about fungi, biology and science in general, how I can complain about everything I usually try not to notice, how it all becomes obvious and falls into places with you. I really enjoyed looking at pathways and having inspiring discussions on how can the study be more elegant, and I'm grateful for all your careful scientific input. Thank you for all your non-scientific support and care, you always find time and place for me and my needs. Thank you for being at peace with my weirdness. It feels like the great adventure just begins.

Many thanks to the Molecular Aspect of Biotransformation lab and mycology department of the Lomonosov MSU in Russia for wonderful filed and lab experience with fungi. Special thanks to Daria Vasina who taught me everything I knew about how to deal with fungi in the lab, and supervised me during my student years. I did not blow the lab, and we had a great time. I miss out travels, hopefully we could see each other regularly again soon. I would

also like to thank my Helicon Company friends (Катя, Тимофей, Стёпа, Егор, Мамед) that helped me to stay positive and keep writing applications for a position. Egor, thank you providing valuable input on how to present myself and having a mock-interview with me. It worked :)

To my russian/ Ukrainian friends in the Netherlands: Юля, Ира, Наташа. Огромное спасибо за вашу поддержку, за весь этот гомерический хохот, за задушевные беседы, ваш юмор и готовность помочь. Люблю и ценю каждую!

To my friends in Russia that are now everywhere: Дашкевич Юля и Ваня, Давыдов Ваня, Влад, Гоша, Артём, Надя и Валера, Лёша и Саша. Спасибо за ваши слова, пусть иногда редкие, но очень теплые. Даже когда вас нет рядом, я чувствую ваши ладони на моих плечах. Безумно скучаю. Наташа Кугушева и Лена Кудинова, спасибо за ваше проявленное и вдохновляющее женское начало, за ваш вкус к жизни и то, как вы смотрите на бытие. Юля Саитгалеева, спасибо за нашу неугасающую связь (со средней школы в Казани!), за татарский юмор, пироги, полночные тусовки и все твое тепло.

To my brother Eugene, my mom Margarita, my dad Vladimir and my family in Russia. Спасибо, что вы так в меня верите и любите! Много раз, из года в год я все крепче убеждаюсь в том, что у меня самая лучшая семья. Каждый из вас по-особенному мне дорог, в вас много доброты, которую вы так щедро дарите мне и всем моим друзьям. С вами безумно интересно говорить, я особенно люблю то, как в вас соседствуют житейская мудрость и юмор. Спасибо самым замечательным крестным Гуленьке и доктору Персику за вашу чуткость и поддержку. Мои любимые Лютики-Гомеры, хочу видеться с вами почаще, и вместе смаковать жизнь.

To my late uncle Dr. Sergey Mosunov, who was an astronomer. Оказалось, звёзды не мигают специально для меня!

A

Curriculum vitae

Olga Vladimirovna Mosunova was born on 14th of August 1992 in Kazan, Republic of Tatarstan, Russia. She obtained specialist degree (= MSc) in botany with specialization in mycology in Lomonosov Moscow State University, Department of Mycology, Algology and Lichenology in 2014. She did her qualification research project at A. N. Bach institute of Biochemistry of Russian Academy of Sciences in the laboratory of Molecular Aspect of Biotransformation of Prof., Dr. O. Koroleva under the supervision of Dr. Daria V. Vasina. She worked on characterization of lignolytic enzyme complex of wood decaying white rot basidiomycete *Phanerochaete avellanea*. In 2014 she enrolled into a PhD program affiliated with A. N. Bach Institute of Biochemistry, which she left in 2016. Her work was focused on regulatory elements in promoters of genes from selected families of wood decaying enzymes from fungi. In 2018 she started a PhD project in Westerdijk Institute of Fungal Biodiversity in the Natural Products laboratory under the supervision of Dr. Jérôme Collemare and Dr. Jorge C. Navarro-Muñoz, and Prof. Dr. Pedro Crous as a promotor.

A

**Всем любопытным,
отчаянным и бесстрашным**