

Article

Biosynthetic Gene Content of the ‘Perfume Lichens’ *Evernia prunastri* and *Pseudevernia furfuracea*

Anjuli Calchera ^{1,2}, Francesco Dal Grande ² , Helge B. Bode ³ and Imke Schmitt ^{1,2,*}

¹ Institute of Ecology, Evolution and Diversity, Goethe University Frankfurt, 60438 Frankfurt am Main, Germany; Anjuli.Calchera@senckenberg.de

² Senckenberg Biodiversity and Climate Research Centre (S-BiKF), Senckenberg Gesellschaft für Naturforschung, 60325 Frankfurt am Main, Germany; Francesco.DalGrande@senckenberg.de

³ Molecular Biotechnology, Department for Biosciences, Goethe University Frankfurt, 60438 Frankfurt am Main, Germany; h.bode@bio.uni-frankfurt.de

* Correspondence: Imke.Schmitt@senckenberg.de; Tel.: +49-(0)69-7542-1855

Received: 14 December 2018; Accepted: 4 January 2019; Published: 8 January 2019



Abstract: Lichen-forming fungi produce a vast number of unique natural products with a wide variety of biological activities and human uses. Although lichens have remarkable potential in natural product research and industry, the molecular mechanisms underlying the biosynthesis of lichen metabolites are poorly understood. Here we use genome mining and comparative genomics to assess biosynthetic gene clusters and their putative regulators in the genomes of two lichen-forming fungi, which have substantial commercial value in the perfume industry, *Evernia prunastri* and *Pseudevernia furfuracea*. We report a total of 80 biosynthetic gene clusters (polyketide synthases (PKS), non-ribosomal peptide synthetases and terpene synthases) in *E. prunastri* and 51 in *P. furfuracea*. We present an in-depth comparison of 11 clusters, which show high homology between the two species. A ketosynthase (KS) phylogeny shows that biosynthetic gene clusters from *E. prunastri* and *P. furfuracea* are widespread across the Fungi. The phylogeny includes 15 genomes of lichenized fungi and all fungal PKSs with known functions from the MIBiG database. Phylogenetically closely related KS domains predict not only similar PKS architecture but also similar cluster architecture. Our study highlights the untapped biosynthetic richness of lichen-forming fungi, provides new insights into lichen biosynthetic pathways and facilitates heterologous expression of lichen biosynthetic gene clusters.

Keywords: lichen secondary metabolites; tree moss; oakmoss; biosynthetic gene clusters; polyketide synthases; non-ribosomal peptide synthetases; terpene synthases; transcription factor; phylogeny; comparative genomics

1. Introduction

Lichens are symbioses, composed of a fungal partner (mycobiont) and one or more photosynthetic partners (photobiont) [1]. Approximately one fifth of all described fungi form lichens and more than 19,000 lichen species are described [2]. Lichens can be found in most terrestrial ecosystems. They impact various community processes, also because of their rich and diverse secondary chemistry [3,4]. Recent studies have shown that lichens represent complex multi-species symbiotic assemblages, forming microhabitats and harbouring a high diversity of other eukaryotic and prokaryotic microorganisms [5–7]. High-throughput sequencing technologies have revealed the presence of multiple fungal and algal species, along with hyper-diverse bacterial communities, within individual lichen thalli [7–13].

Lichens produce a great variety of natural compounds, and more than 1000 secondary metabolites have been identified to date, many of which are exclusively found in lichens [14–17]. The vast

majority of these characteristic lichen secondary metabolites is of fungal origin. Many lichen secondary metabolites have important ecological roles including light-screening, chemical weathering, allelopathic and anti-herbivore defence [18–20]. Furthermore, lichens are a promising source for pharmaceutically interesting natural products, because of the manifold biological activities of lichen compounds—including antiviral, antibiotic, antitumor, allergenic, plant growth inhibitory, antiherbivore and enzyme inhibitory activities [15,19,21,22]. For example, the secondary metabolites atranorin, evernic acid, physodic acid and usnic acid found in the lichens *Evernia prunastri* and *Pseudevernia prunastri* are strong metabolic enzyme inhibitors, and atranorin may inhibit lung cancer cell mobility and tumorigenesis [22,23]. A wide range of lichens have been used in traditional medicines all around the world [24]. Today, one of the more significant economic uses of lichens is in the perfume industry. Combined lichen material is used for extracts added to perfumes for a sweet and mossy smell and to ensure persistence on the skin. Several hundred to thousands of tons of *E. prunastri* and *P. furfuracea* are harvested every year in France, Morocco and South-eastern Europe for this purpose [16,19,25,26].

The most abundant classes of lichen secondary metabolites are phenolic compounds, dibenzofurans, depsides and depsidones built by multienzyme polyketide synthases (PKS). Derived monoaromatic subunits (orcinol, β -orcinol type or methylphloroacetophenone) are then linked by ester, ether or carbon-carbon bonds [1,22,27]. Fungal PKSs consist of a set of active site domains used in an iterative fashion for multiple catalytic cycles and are subdivided into non-reducing (NR-PKS), or reducing (R-PKS) according to the level of reductive processing [28]. Both types of polyketides are found in lichen-forming fungi but fully oxidized polyketides account for most of the well-known and common lichen substance classes such as depsides and depsidones whose formation is controlled by NR-PKS genes [16,29]. Fungal secondary metabolic pathway genes are often grouped in gene clusters including genes encoding tailoring functions, transporters, and pathway-specific regulatory genes along with the core genes [30–32].

Recent advances in genome sequencing and bioinformatic mining algorithms have enabled the identification and characterization of biosynthetic gene clusters from genome sequences and led to the discovery of thousands of biosynthetic gene clusters [33–36]. However, the identity, structure and function of the metabolites encoded by these gene clusters remain mostly unknown. The clusters are therefore referred to as cryptic or orphan clusters [37]. In particular, fungal PKS gene clusters are challenging to functionally characterize, mainly due to the unpredictable intrinsic cryptic programming concealed within iterative PKS genes themselves and missing knowledge on precise selectivity and activity of tailoring enzymes [30,31,38]. From the vast biosynthetic potential found in genomes only a small fraction has been investigated experimentally to connect genes to metabolite production—for example through gene knockout or heterologous expression. But the presence of cryptic or orphan gene clusters in many genomes, and the fact that a vast number of biosynthetic genes is not reflected by the metabolic profile, hint at a plethora of yet undiscovered chemical compounds [32,39,40].

Despite the rich chemical diversity found in lichens, only few biosynthetic studies are available. In lichen-forming fungi the slow growth rates and difficulties in aposymbiotic mycobiont cultivation further hinder the characterization of the molecular mechanisms underlying the biosynthesis of characteristic lichen metabolites [14,41]. The biosynthesis of substances may depend on abiotic factors such as geographic, altitudinal or microhabitat conditions or the response to microclimatic fluctuations, seasonality, chemical signals, hydration, or UV radiation [14,22]. Biosynthesis of natural products may also be influenced by biological factors, such as the presence of competing plant or lichen species, predation by insects, or contact to other symbionts that are part of the lichen [27]. Furthermore, lichen mycobionts synthesize significant quantities of secondary metabolites only under permissive conditions and therefore the production in axenic cultures can differ substantially from that in nature [21,22,42].

In lichen-forming fungi no secondary metabolite has been directly linked with experimental proof (i.e., gene knockout, heterologous expression) to a biosynthetic gene cluster, although several assignments

have been proposed. The first identification of a lichen secondary metabolic gene cluster—grayanic acid in *Cladonia grayi*—was accomplished by Armaleo et al. [43] combining phylogenetic analysis, biosynthetic prediction from PKS domain arrangement and correlation of mRNA levels with metabolic production. Later, the putative gene clusters encoding usnic acid and 6-hydroxymellein in *Cladonia uncials* were identified by Abdel-Hameed et al. [41,44] on the basis of PKS domain arrangement and prediction of the functions of tailoring enzymes encoded in the same gene cluster.

Several methods are applied and combined to discover PKS genes in lichenized fungi: amplification of PKS gene fragments from genomic DNA using degenerate primers and design of probes for library screening; the use of cDNA-based templates generated from reverse-transcribed mRNA, or gaining biosynthetic insights directly from transcriptomes or genomes sequenced from mycobiont cultures or reconstructed from metagenomic thalli [45–47]. The latter became possible through the advances of sequencing technologies and taxonomic assignment tools and led to the publication of several draft genomes of lichen-forming fungi over the last few years (see below). For the prediction of biosynthetic function of PKS genes there are diverse strategies, apart from experimental evidence through gene knockout, heterologous expression, radiolabelling or oxidation experiments [48–51]. These include phylogenetic approaches to infer the functions through a close phylogenetic relatedness to characterized genes [52,53], transcriptional profiling [43,47], or comparative homology mapping of entire gene clusters [31,36,54]. These approaches are often combined to provide putative assignments [47]. A review on reported PKS genes from lichenized fungi including an overview on approaches for PKS identification and characterization has been published recently and highlights the prospects of genomics-driven natural product discovery in lichens [45].

Here, we combine genome mining with comparative genetic mapping and phylogenetics of the two lichenized fungi *Evernia prunastri* and *Pseudevernia furfuracea*. Both lichens are of economic importance [1,19] and their fungal genomes contain a high number of biosynthetic gene clusters [46]. Moreover, a comprehensive review on the multitude of extracts identified from both lichens is available [25,26] and includes, for example, the bioactive metabolite atranorin detected in both lichens. The secondary metabolite richness and partially overlapping chemical profiles make these species particularly interesting study systems for combining genome mining and comparative genomics to investigate biosynthetic enzymes that may be responsible for the production of characteristic lichen substance classes. We further compare the biosynthetic gene richness found in these species with that of 13 other lichenized fungi and 57 representative fungal species and thereby provide the most complete comparison of biosynthetic gene families in lichen genomes based on entire genomes. We present mapped biosynthetic gene clusters of *E. prunastri* and *P. furfuracea* with functionally annotated accessory genes and putative regulatory genes. These carefully annotated clusters provide baseline information for heterologous expression of lichen biosynthetic gene clusters. Furthermore, we present all PKS genes mined from 15 genomes of lichenized fungi in a phylogenetic framework of characterized biosynthetic genes. Information on phylogenetic relatedness to previously characterized genes broadens our understanding of lichen biosynthetic pathways and may help us to identify promising clusters for the production of characteristic substance classes.

Specifically, we address the following questions:

- (I) What is the diversity of biosynthetic gene clusters in *Evernia prunastri* and *Pseudevernia furfuracea* and how does it compare to other lichenized fungi and non-lichenized fungi?
- (II) What is the architecture and gene content of those clusters with high homology between *E. prunastri* and *P. furfuracea*?
- (III) Where do PKSs from *E. prunastri* and *P. furfuracea* group phylogenetically in a phylogeny of PKSs with known functions?

2. Results & Discussion

2.1. Biosynthetic Gene Richness in Fifteen Annotated Genomes of Lichen-Forming Fungi

We investigated the biosynthetic gene richness in a total of fifteen genomes of lichen-forming fungi. The species belong to different phylogenetic groups and synthesize a variety of lichen substances (Table 1).

Table 1. Overview of genomes of lichen-forming fungi used in this study.

Species	Taxonomic Group	Data Repository ¹	Gene Set Previously Available	Genome Size	Number of Scaffolds	Scaffold N50	Number of Genes	Abbreviation	Metabolites Reported [55–59]
<i>Cladonia grayi</i>	Lecanoromycetes, Lecanorales	JGI Clagr3 v2.0	yes	34.6 Mb	414	243,412	11,389	CGRA	grayanic acid, fumarprotocetraric acid complex
<i>Cladonia macilentata</i>	Lecanoromycetes, Lecanorales	NCBI AUPP000000000.1	no	37.1 Mb	240	1,469,036	10,559	CMAC	thamnolic acid, barbatic acid, didymic acid, squamatic acid, usnic acid, rhodocladonic acid
<i>Cladonia metacorallifera</i>	Lecanoromycetes, Lecanorales	NCBI AXCT000000000.2	no	36.7 Mb	30	1,591,850	10,497	CMET	usnic acid, didymic acid, squamatic acid, rhodocladonic acid
<i>Cladonia uncialis</i>	Lecanoromycetes, Lecanorales	NCBI NAPT000000000.1	no	32.9 Mb	2124	34,871	10,902	CUNC	usnic acid, squamatic acid
<i>Endocarpon pusillum</i> (Park et al.) [60]	Eurotiomycetes, Verrucariales	NCBI JFDM000000000.1	no	37.2 Mb	40	1,340,794	11,756	EPUP	(none reported)
<i>Endocarpon pusillum</i> (Wang et al.) [61]	Eurotiomycetes, Verrucariales	NCBI APWS000000000.1	yes	37.1 Mb	908	178,225	9238	EPUW	(none reported)
<i>Evernia prunastri</i>	Lecanoromycetes, Lecanorales	NCBI NKYR000000000.1	yes	40.3 Mb	277	264,454	10,992	EPRU	evernic acid, atranorin, usnic acid
<i>Gyalolechia flavorubescens</i>	Lecanoromycetes, Teloschistales	NCBI AUPK000000000.1	no	34.5 Mb	36	1,693,300	10,460	GFLA	parietin, emodin, fallacinal, fragilin
<i>Lobaria pulmonaria</i>	Lecanoromycetes, Peltigerales	JGI Lobpu1 v1.0	yes	56.1 Mb	1911	50,541	15,607	LPUL	stictic acid, norstictic acid, constictic acid
<i>Pseudevernia furfuracea</i>	Lecanoromycetes, Lecanorales	NCBI NKYQ000000000.1	yes	37.8 Mb	46	1,178,799	8842	PFUR	atranorin, olivetoric acid, physodic acid
<i>Ramalina peruviana</i>	Lecanoromycetes, Lecanorales	NCBI MSTJ000000000.1	no	27.0 Mb	1657	40,431	9338	RPER	sekikaic acid complex
<i>Umbilicaria hispanica</i>	Lecanoromycetes, Umbilicariales	NCBI PKMA000000000.1	yes	41.2 Mb	1619	145,035	8488	LHIS	gyrophoric acid, lecanoric acid, umbilicatic acid, skyrin
<i>Umbilicaria muehlenbergii</i>	Lecanoromycetes, Umbilicariales	NCBI JFDN000000000.1	no	34.8 Mb	7	7,009,248	8968	UMUE	gyrophoric acid
<i>Umbilicaria pustulata</i>	Lecanoromycetes, Umbilicariales	NCBI FWEW000000000.1	yes	39.2 Mb	3891	104,297	8268	LPUS	gyrophoric acid, lecanoric acid, hiascinic acid, skyrin
<i>Xanthoria parietina</i>	Lecanoromycetes, Teloschistales	JGI Xanpa2 v1.1	yes	31.9 Mb	39	1,731,186	11,065	XPAR	physcion, parietinic acid, teloschistin, emodin

¹ NCBI: National Center for Biotechnology Information; JGI: DOE Joint Genome Institute.

The genomes differ substantially in assembly status—from 7 to 3891 scaffolds—but most are 90–96% complete according to BUSCO (Figure 1a). Only the genomes of *Ramalina peruviana* (86%) and *Cladonia uncialis* (89%) score lower in completeness and also have the lowest scaffold N50 (see genome statistics given in Table 1 and Table S1). Since seven genomes did not have available gene sets, we predicted gene models. The limitations of gene prediction without transcriptome data are evident in the assessment of gene set completeness in Figure 1b. The gene sets for which we performed gene annotation are less complete (72–80% compared to 82–97%) and show more fragmented BUSCO marker genes. Nevertheless, the predicted number of genes is in the range of most of the gene sets available for lichen-forming fungi (~8200 to 11,800). The only exception here is the genome of *Lobaria pulmonaria* which has a considerable larger genome size with 56.1 Mb and 15,607 genes but also the highest rate of duplicated marker genes (Table 1, Table S1). Even though the gene prediction done here is not optimal, it improves the downstream biosynthetic gene cluster detection over the less specific annotation that can be done within antiSMASH [62].

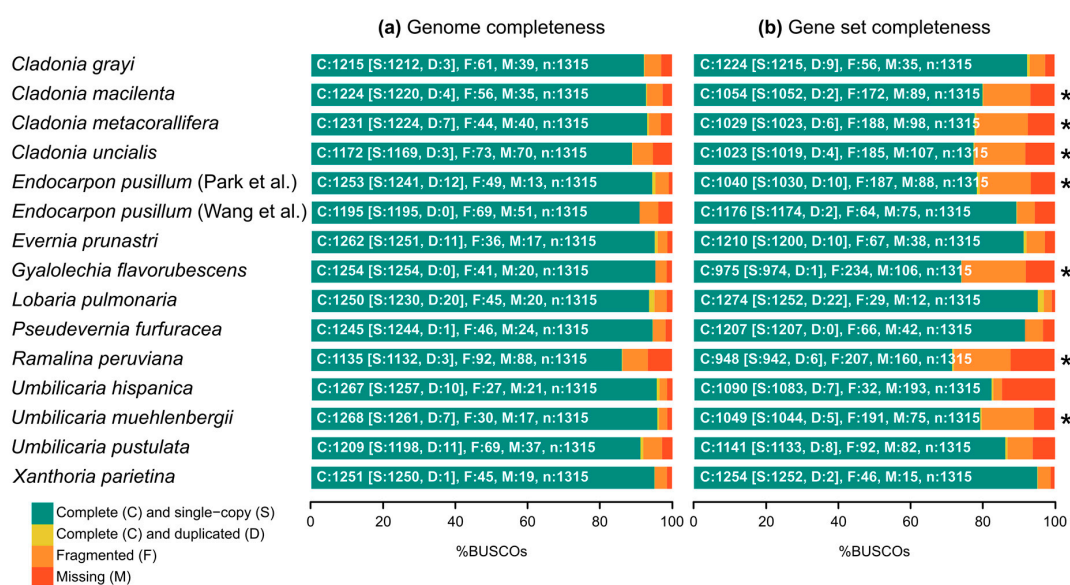


Figure 1. (a) Genome completeness for assemblies and (b) gene set completeness for gene annotations of lichen-forming fungi. The completeness is compared against 1315 orthologous BUSCO marker genes for Ascomycota. Exact percentages can be found together with basic genome statistics in Supplementary Table S1. Asterisk indicates genomes annotated in the present study (without RNA evidence).

We identified a high number of biosynthetic gene clusters in all genomes of lichen-forming fungi (Table 2). The fungal version of antiSMASH annotated on average ~47 clusters per genome, ranging from 27 clusters for *Umbilicaria pustulata* up to as many as 80 gene clusters in *Evernia prunastri*. The most abundant family of secondary metabolite enzymes identified in all genomes are reducing type I polyketides (R-PKS) followed by terpene synthases and non-reducing type I polyketides (NR-PKS). We also detected non-ribosomal peptide synthetases (NRPS), hybrid PKS-NRPS and type III PKSs in most genomes.

Polyketides represent the most abundant class of lichen secondary metabolites and can be reduced or fully oxidized [16,22,27]. It has been reported that most polyketides found in lichens are fully oxidized [63]. Nevertheless, we find that the number of R-PKS genes exceeds by far the number of NR-PKS genes. The genomes of the lichen-forming fungi show—through their large number of secondary metabolic genes and gene clusters—the potential for a much greater number of natural products than have been reported to occur in the respective lichen species (Table 1).

Table 2. Number of biosynthetic gene clusters and main families of secondary metabolite genes found in the genomes of lichen-forming fungi.

Species	Abbreviation	Number of Clusters	Type I NR-PKS	Type I R-PKS	Type I PKS	Type III PKS	Hybrid PKS-NRPS	NRPS	Terpene Synthases	Total KS Sequences for Phylogeny	Complete PKS (KS + AT + ACP)
<i>Cladonia grayi</i>	CGRA	51	8	17	1	3	1	2	5	27	15
<i>Cladonia macilentata</i>	CMAC	52	11	23	-	2	4	2	5	38	25
<i>Cladonia metacorallifera</i>	CMET	59	13	26	1	2	2	2	8	42	29
<i>Cladonia uncialis</i>	CUNC	59	10	25	1	2	3	1	8	39	25
<i>Endocarpon pusillum</i> [8] (Park et al.)	EPUP	31	4	9	-	1	2	2	6	15	9
<i>Endocarpon pusillum</i> [9] (Wang et al.)	EPUP	31	5	12	-	1	2	1	6	19	12
<i>Evernia prunastri</i>	EPRU	80	9	29	1	2	4	4	13	43	30
<i>Gyalolechia flavorubescens</i>	GFLA	37	7	12	-	1	1	3	6	20	10
<i>Lobaria pulmonaria</i>	LPUL	77	8	28	-	-	4	9	16	40	22
<i>Pseudevernia furfuracea</i>	PFUR	51	5	23	-	2	3	4	5	31	17
<i>Ramalina peruviana</i>	RPER	47	9	18	3	1	1	3	2	31	15
<i>Umbilicaria hispanica</i>	LHIS	28	7	10	1	1	-	-	6	18	14
<i>Umbilicaria muehlenbergii</i>	UMUE	31	5	15	-	1	-	-	4	20	17
<i>Umbilicaria pustulata</i>	LPUS	27	6	9	-	1	1	-	6	16	13
<i>Xanthoria parietina</i>	XPAR	51	5	18	1	1	2	5	7	26	13
Sum		712	112	274	9	25	30	38	125	425	266

NR-PKS = Non-reducing type I polyketide synthase; R-PKS = Reducing type I polyketide synthase; NRPS = Non-ribosomal peptide synthetase. KS = Ketosynthase; AT = Acyltransferase; ACP = Acyl carrier protein.

The genomes of lichen-forming fungi in the Lecanoromycetes show a remarkable richness of secondary metabolic gene clusters in comparison with genomes of species from all major fungal classes. We present the number of predicted secondary metabolite gene clusters in 15 genomes of lichenized fungi (this study) and 57 ascomycete and basidiomycete genomes of non-lichenized fungi analysed in a previous study [32] (Figure 2). The average number of predicted secondary gene clusters is ~31 clusters and the average number of predicted PKS clusters is ~11. Except for two *Umbilicaria* species the number of lichen gene clusters is above average and all lichen genomes harbour more than the average amount of PKS gene cluster (~24 clusters on average). Indeed, the highest number of PKS gene clusters are found in *Evernia prunastri* (38 PKS clusters) together with six other lichens including *Pseudevernia furfuracea* in the top ten PKS cluster count. We used a non-representative selection of fungal genomes for the comparison and thus did not include all fungal species that are reported to be rich in secondary metabolite gene clusters, for example, *Pestalotiopsis fici* (Sordariomycetes) [64]. The comparison is meant for the sole purpose of placing the richness of secondary metabolite gene clusters found in lichen-forming fungi in a broader context (Figure 2).

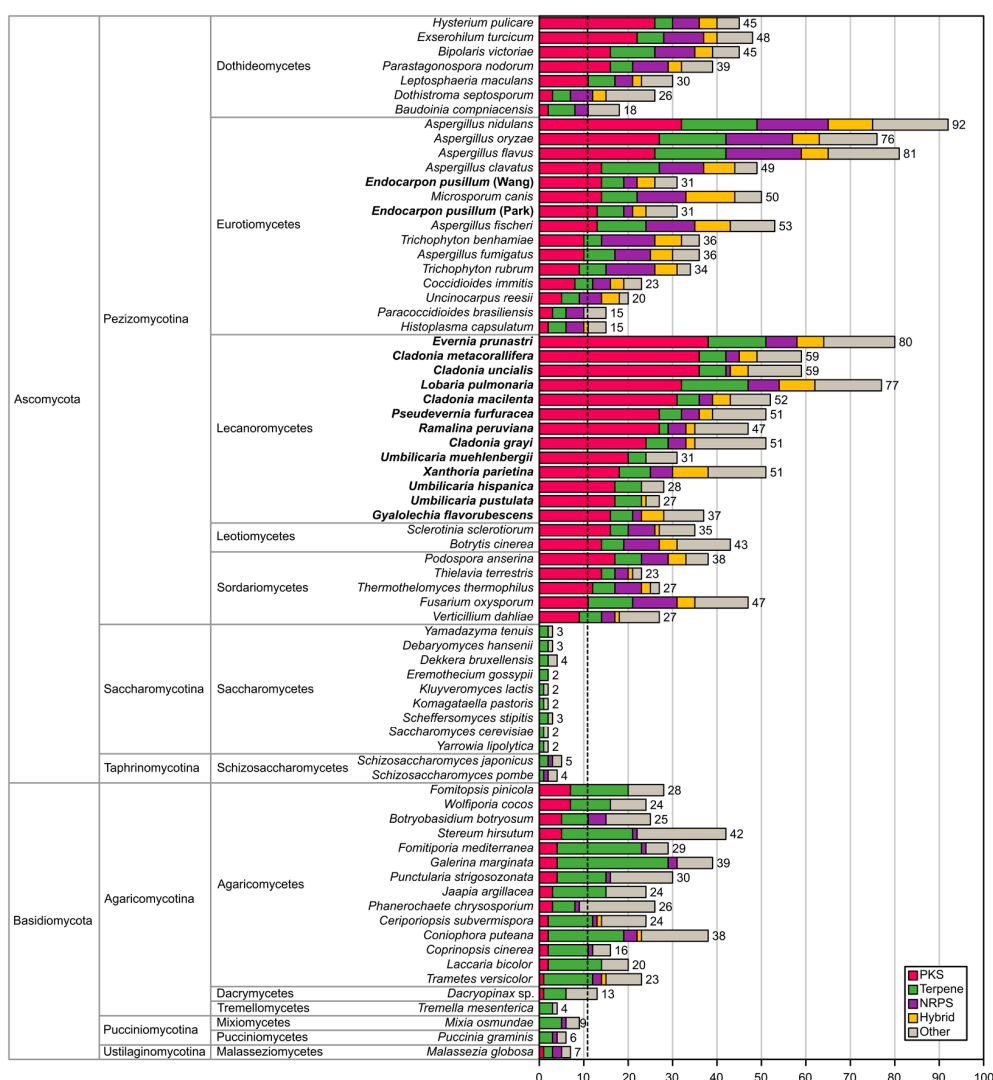


Figure 2. Overview of predicted secondary metabolic gene clusters across the genomes of representative fungal species adopted from [32]. Bold font indicates genomes of lichen-forming fungi included from this study. The dashed line shows the average number of PKS gene clusters found in a genome. “Hybrid” refers to clusters with multiple core genes belonging to different secondary metabolite families. PKS = polyketide synthase; NRPS = non-ribosomal peptide synthetase.

2.2. Gene Cluster Comparison

Pseudevernia furfuracea and especially *E. prunastri* show a high number of biosynthetic gene clusters (Table 2), and a high number of natural products has been reported for both species [25,26]. This richness in secondary metabolites and partially overlapping chemical profiles (e.g., atranorin and chloroatranorin) make these species particularly interesting study systems for combining genome mining and comparative genomics to investigate biosynthetic enzymes that may be responsible for the production of characteristic lichen substance classes. Moreover, both species are harvested in large quantities for the perfume industry and are of economic importance [1,19]. A detailed list of the biosynthetic gene clusters detected in both species can be found in Tables S2 and S3.

For the gene cluster comparison, we functionally annotated 887 cluster genes of *E. prunastri* and 548 cluster genes of *P. furfuracea* with gene ontology terms and protein names. The full Blast2GO annotation is presented in Tables S4 and S5. We then identified 126 orthologous pairs as Reciprocal Best Blast Hits (RBH) between the *E. prunastri* and *P. furfuracea* cluster genes. The full list of RBHs is given in Table S6. From these results we chose all biosynthetic gene clusters that contained an orthologous core PKS, NRPS or hybrid PKS-NRPS for our comparative genetic mapping approach presented in the following synteny plots of Figures 3–5. The phylogenetic placements of these clusters containing a ketosynthase (KS) domain are presented below.

Four NR-PKS clusters show a core RBH gene (Figure 3). Three of these (Figure 3A,B,D) show a high cluster homology between *E. prunastri* and *P. furfuracea* with several genes carrying similar functional annotation and RBHs of each other. The homologous genes include for example cytochrome P450, monooxygenases, serine/threonine kinases, and regulatory genes. Most of the homologous regions between the clusters are confined within genes but there are a few exceptions (Figure 3D). These exceptions with high homology of coding sequences to non-coding regions may be an artefact of gene prediction, and may indicate genes missed in the annotation. The clusters in Figure 3C on the other hand display only similarity between the two core PKS genes—one of which is a R-PKSs—and not between any accessory genes. Closer inspection of the domain architecture of the NR-PKSs in Figure 3 using BLASTp indicates the presence of acyl carrier protein (ACP) transacylase starter units (SAT) and product templates (PT). Both conserved domains were shown to be typical for NR-PKSs in addition to the minimal PKS domain structure of a KS, AT (acyltransferase) and ACP [65–68].

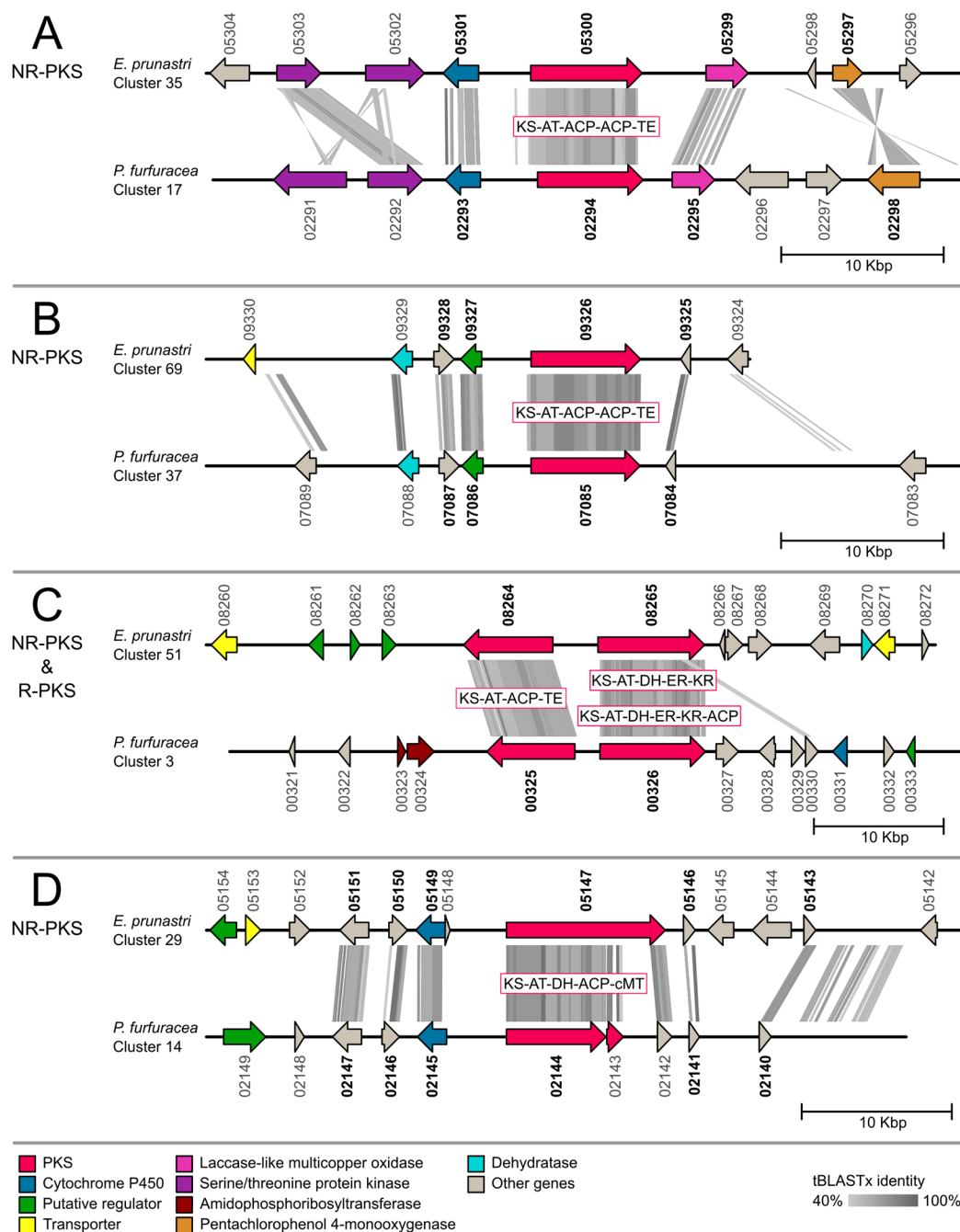


Figure 3. Synteny plots of biosynthetic gene clusters with an orthologous non-reducing polyketide synthase (NR-PKS) core gene in *Evernia prunastri* and *Pseudevernia furfuracea*. Orthologous genes identified with the reciprocal best hit (RBH) approach are highlighted in bold. ACP = Acyl carrier protein; AT = Acyltransferase; cMT = C-Methyltransferase; DH = Dehydratase; ER = Enoylreductase; KR = Ketoreductase; KS = Ketosynthase; TE = Thioesterase.

We further identified four R-PKS clusters with orthologous core genes (Figure 4). Two cluster comparisons (Figure 4A,B) show a high homology between both species. The entire R-PKS in Figure 4A is split over two to three genes—possibly an artefact of gene prediction—but we detected nine RBH accessory genes in the clusters including oxidases and calcium-binding domain proteins. The homologous cluster genes in Figure 4B include cytochrome P450, dehydrogenases, transporter genes and chalcone synthases (type III PKS). The other two comparative mapped clusters (Figure 4C,D)

share only the core gene as a RBH. In Figure 4C we detected a homologous region with annotated genes in *P. furfuracea* while *E. prunastri* lacked any predicted gene models in the concurrent region.

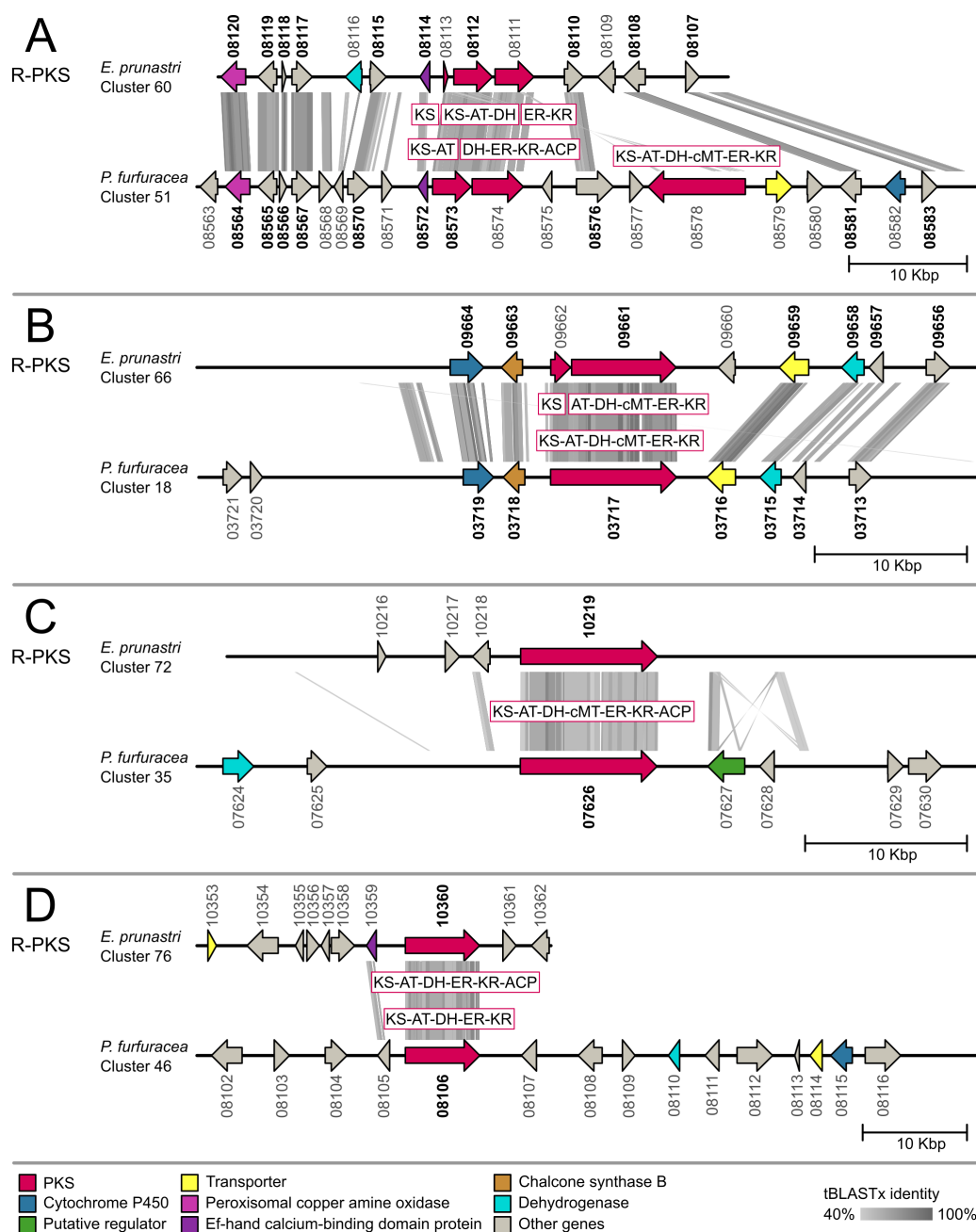


Figure 4. Synteny plots of biosynthetic gene clusters with an orthologous reducing polyketide synthase (R-PKS) core gene in *Evernia prunastri* and *Pseudevernia furfuracea*. Orthologous genes identified with the reciprocal best hit (RBH) approach are highlighted in bold. ACP = Acyl carrier protein; AT = Acyltransferase; cMT = C-Methyltransferase; DH = Dehydratase; ER = Enoylreductase; KR = Ketoreductase; KS = Ketosynthase; TE = Thioesterase.

Last, we included NRPS and hybrid PKS-NRPS clusters with orthologous core genes in the comparative analyses (Figure 5). We detected two hybrid PKS-NRPS clusters (Figure 5A,C) and two NRPS clusters (Figure 5B,C) all showing a high similarity in functionally annotated genes and through RBHs. Accessory genes include transporter genes, hydrolases and putative regulatory genes. In Figure 5C a hybrid PKS-NRPS and a NRPS are encoded closely together. This might suggest that

both enzymes participate in the biosynthesis of one natural product or that the two natural products might function together.

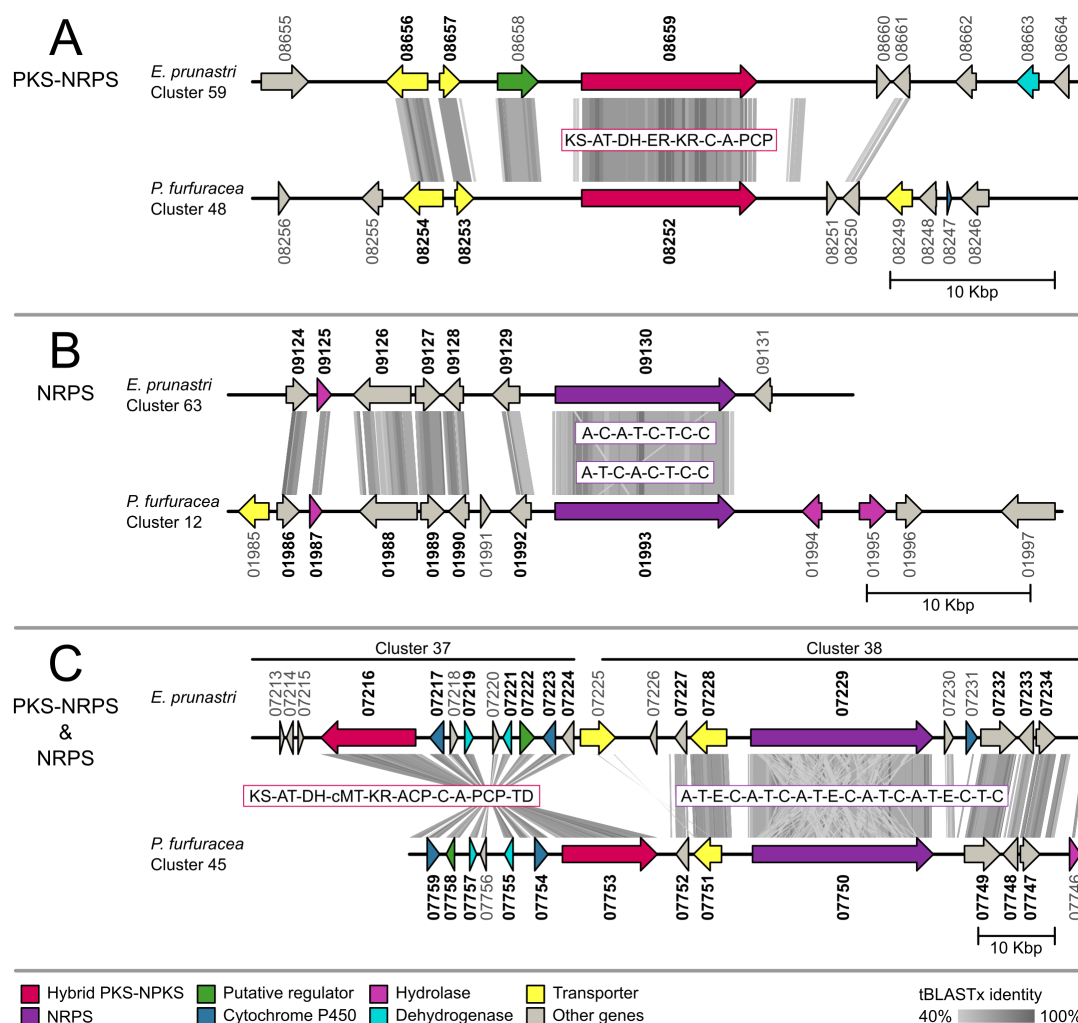


Figure 5. Synteny plots of biosynthetic gene clusters with an ortholog non-ribosomal peptide synthetase (NRPS) core gene or hybrid with polyketide synthase (PKS-NRPS) in *Evernia prunastri* and *Pseudevernia furfuracea*. Orthologous genes identified with the reciprocal best hit (RBH) approach are highlighted in bold. A = Adenylation; ACP = Acyl carrier protein; AT = Acyltransferase; C = Condensation; cMT = C-Methyltransferase; DH = Dehydratase; E = Epimerization; ER = Enoylreductase; KR = Ketoreductase; KS = Ketosynthase; PCP = peptide carrier protein; T = Thiolation; TD = Terminal domain.

2.3. Putative Regulators of Biosynthetic Gene Clusters

We identified a total of 60 putative intracluster regulators in the biosynthetic gene clusters of *E. prunastri* and *P. furfuracea*. Most of these were present in the gene clusters of *E. prunastri* (44 genes). Two pairs are orthologs between both species based on RBH analysis and can be found among the clusters investigated with comparative mapping (see clusters in Figures 3B and 5C). We detected ten C6 zinc finger domain proteins, six Zn₂Cys₆ DNA-binding proteins and two ankyrin repeat proteins among other fungal specific transcription factors. These are typical regulators of secondary metabolites [37,69]. A detailed list of these putative regulatory genes can be found in Table S7.

2.4. Phylogenetic Analysis with Characterized Fungal Polyketides

We reconstructed a maximum likelihood tree of the conserved KS domain that allows inference of domain architecture and pathway association of PKSs [52,70,71]. We included a total of 413

KS sequences from fifteen genomes of lichen-forming fungi in relation to 131 MIBiG entries of characterized fungal PKSs and six animal fatty acid synthases (FAS) as an outgroup. We also included partial PKSs from lichen-forming fungi that lack one of the required domains AT or ACP for a minimal PKS organization. Out of 413 lichen KS sequences five originate from a PKS without an AT domain and 136 from a PKS without an ACP domain. We included these since such partial PKSs are also present in the MIBiG repository of characterized enzymes [72] and may nevertheless represent functional and interesting genes. In our manually curated MIBiG dataset we found one entry without an AT domain and 32 entries without an ACP domain.

The entire phylogenetic tree is shown in full length in Supplementary Figure S1. Overall the phylogeny shows support for a clade with NR-PKSs, a clade with R-PKS and a clade containing 6-methylsalicylic acid synthases (6-MSAS). Hybrid PKS-NRPS genes contain reducing domains and fall within the clade of R-PKS. We provide here a few examples of interesting clades found in the phylogenetic analysis with a focus on NR-PKS since the basic building subunits for lichen-specific depsides and depsidones are believed to be encoded by this gene family [1,14,27,73].

Within the NR-PKS clade we find one supported clade of PKSs containing C-methyltransferase (cMT) domains including a subclade with the PKS proposed to produce the lichen substance usnic acid [41]. *Evernia prunastri* is also known to produce usnic acid (Table 1) and one PKS (gene 02873) falls phylogenetically close to the potential usnic acid producer (Figure 6).

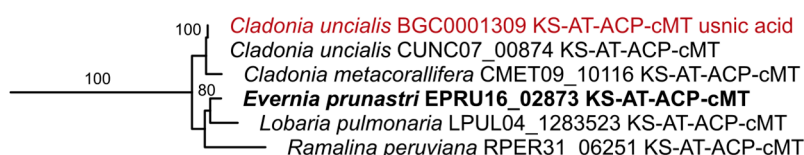


Figure 6. Supported clade containing the gene *MPAS* (methylphloroacetophenone synthase) of *Cladonia uncialis* (MIBiG-ID BGC0001309) putatively associated with usnic acid biosynthesis. For the complete KS tree see Figure S1.

Another clade contains the NR-PKSs of *E. prunastri* and *P. furfuracea* (Figure 3A) which are phylogenetically close to the PKS gene that is putatively associated with grayanic acid biosynthesis [43]. All *Cladonia* species included in this study (*C. grayi*, *C. macilenta*, *C. metacorallifera* and *C. uncialis*) have members within this clade (Figure 7).



Figure 7. Supported clade containing the gene *PKS16* of *Cladonia grayi* (MIBiG-ID BGC0001266), putatively involved in grayanic acid biosynthesis. The cluster 35 of *Evernia prunastri* (EPRU35) and cluster 17 of *Pseudevernia prunastri* (PFUR17) are phylogenetically close and are presented in detail in synteny plot Figure 3A. For the complete KS tree see Figure S1.

We further inspected the phylogenetic placement of the comparative mapped NR-PKSs of the *E. prunastri* and *P. furfuracea* gene clusters. The gene clusters mapped in Figure 3B fall into a group with tandem ACP domains and in the proximity of characterized PKSs involved in the production of hydroxy naphthalenes [74,75] (Figure 8). Naphthalene is found in the extracts of *E. prunastri* and *P. furfuracea* used for the perfume industry [1] (pp. 132–133).

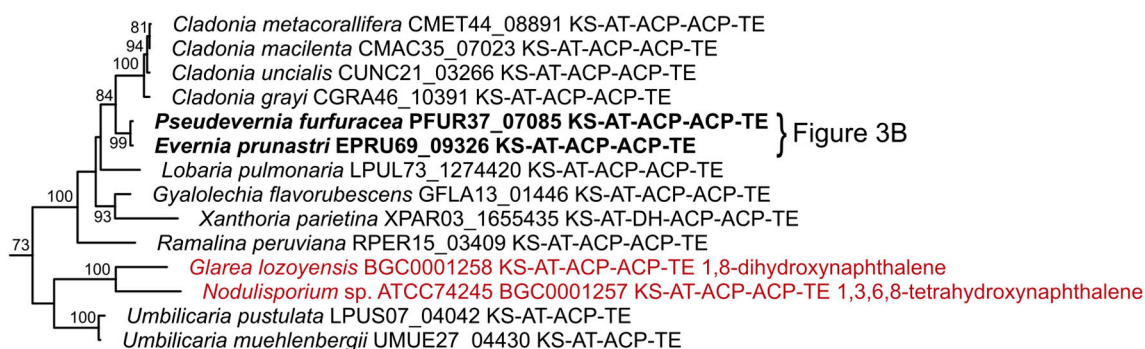


Figure 8. Supported clade containing PKSs from dihydroxy and tetrahydroxy naphthalene producers (MIBiG-ID BGC0001257, BGC0001258). For the complete KS tree see Figure S1.

Two closely related clusters of *E. prunastri* and *P. furfuracea* (Figure 3C) fall into a clade with the experimentally characterized PKS gene *orsA* known to be involved in orsellinic acid biosynthesis in *Aspergillus nidulans* [76]. This clade (Figure 9) provides an especially interesting starting point for further investigations as the most abundant class of lichen metabolites are composed of orcinol or β -orcinol monomers [27].

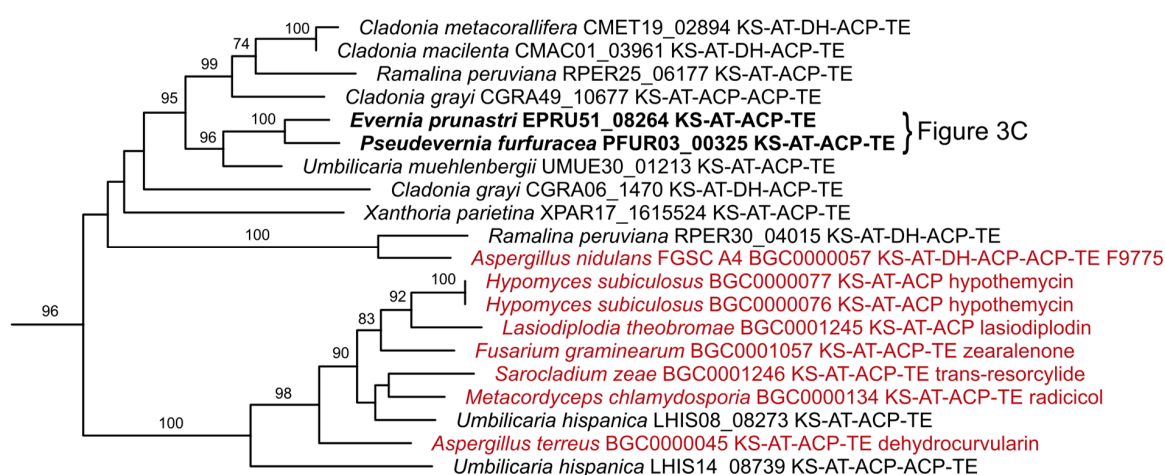


Figure 9. Supported clade containing the PKS *orsA* for orsellinic acid in *Aspergillus nidulans* (MIBiG-ID BGC0000057). For the complete KS tree see Figure S1.

We also want to underline that combining comparative mapping of entire gene clusters with a phylogenetic approach based only on the KS sequence of the core PKS gene shows that entire cluster similarities correspond to KS topology. This is in line with findings that emphasize the predictive power of the conserved KS domains for the investigation of enzyme architecture and pathway association [52,70,71].

Our study combines genome mining and comparative genomics and highlights the high diversity of biosynthetic gene clusters that can be found in fifteen genomes of lichen-forming fungi. This number exceeds by far the number of lichen metabolites that are reported for these species. The secondary metabolite gene cluster richness found in the genomes of lichen-forming fungi is above the average richness found in other representative fungal species, especially polyketide synthase gene clusters. The comparative mapping of interesting biosynthetic gene clusters, functional annotation of accessory genes together with the identification of putative regulatory genes presented here will aid in providing new insights into lichen biosynthetic pathways and serve as a valuable resource for developing heterologous expression of lichen biosynthetic gene clusters.

3. Materials and Methods

3.1. Identification and Annotations of Biosynthetic Gene Clusters

We used the genomes of the two lichen-forming fungi of *Evernia prunastri* and *Pseudevernia furfuracea* for biosynthetic gene cluster identification (accession numbers in Table 1). Sequencing of cultures, genome assemblies, gene prediction and genome mining for biosynthetic gene clusters in both species were done as described in Meiser et al. [46]. We provide here a short description on the identification of biosynthetic gene clusters to give a full depiction on our focused genome mining approach. We identified typical gene families of secondary metabolism and annotated biosynthetic gene clusters with the fungal version of antiSMASH v.4.0.2 (fungiSMASH) [62,77], including polyketide synthases (PKS), non-ribosomal peptide synthetases (NRPS) and terpene synthases. These gene families are typical targets in lichen-metabolite genome mining, because they encode the structural frame of most secondary metabolites [19,37,41,44]. As input for antiSMASH we used an annotated nucleotide file (EMBL format) constructed from the genome FASTA file and the GFF file from gene prediction.

Both investigated species show a high number of biosynthetic gene clusters with reducing and non-reducing PKSs, NRPSs, and terpene synthases [46] and represent a promising source of bioactive secondary metabolites. To further investigate these biosynthetic gene clusters for accessory genes involved in the lichen metabolite synthesis, we annotated all predicted protein sequences belonging to an antiSMASH cluster with gene ontology (GO) terms and protein names. The functional annotation was done using Blast2GO [78] v.5.0.22 and by running BLASTp [79] v.2.2.29 against the NCBI GenBank protein database *nr* (downloaded May 2017) [80] and InterProScan [81] v.5.28–67.0 with a matching lookup service and the Pfam (protein family) database v.31.0 [82].

3.2. Gene Cluster Comparison in *Evernia prunastri* and *Pseudevernia furfuracea*

We applied a two-step approach to compare the biosynthetic genes cluster of *E. prunastri* and *P. furfuracea*. First, as a simple and fast way of detecting orthologous pairs [83], we identified Reciprocal Best Blast Hits (RBH) between all cluster genes of both species. For better detecting orthologs we used BLASTp v.2.2.31 + with final Smith-Waterman alignment and soft filtering (BLAST flags *use_sw_tback*, *soft_masking true*, *seq yes* and *evaluate 1e-6*) [83,84]. The BLAST hits were then filtered for a minimum query coverage of 50% (*qcovhsp*) and a minimum identity of 70% over the alignment length (*pident*) and sorted for highest bit-score and lowest e-value [85].

Second, we used comparative genetic mapping to analyse the homology and arrangement of entire clusters instead of looking at core gene orthologs only. For the comparison we chose all biosynthetic gene clusters that contained a RBH core PKS, NRPS, or hybrid PKS-NRPS. Synteny plots were generated with Easyfig python script v.2.2.3 [86] and the tBLASTx option with a minimum identity value of 40% and a minimum alignment length of 50 [87]. Gene clusters were reversed when necessary to have a matching orientation. We generated GBK files with seqkit [88] v.0.7.2 and the seqret tool in the EMBOSS package [89] v6.6.0.0 as input for Easyfig.

3.3. Putative Regulators of Biosynthetic Gene Clusters

The transcriptional control of biosynthetic genes by transcription factors is complex and can occur on several levels from pathway-specific to a broader global regulation [37]. Specific acting regulatory genes are usually found in the cluster that the factors regulate. Examples include Zn(II)₂Cys₆, Cys₂His₂, basic leucine zipper (bZip), winged helix, zinc-finger, or ankyrin repeat proteins [30,69]. The identification of these regulators involved in the synthesis of secondary metabolites may be crucial for successful heterologous expression experiments or for activating silent gene clusters.

For the detection of putative regulators within the biosynthetic gene clusters of *E. prunastri* and *P. furfuracea* we made use of the Blast2GO annotation. We searched for the following key terms in the GO names: ‘signal transducer activity’, ‘transcription factor activity’ and ‘transcription factor binding’.

3.4. Biosynthetic Gene Richness in Other Genomes of Lichenized and Non-Lichenized Fungi

We assessed the chemical diversity and biosynthetic gene richness of *E. prunastri* and *P. furfuracea* in comparison to the genomes of other lichen-forming fungi. We analysed all genomes of lichen-forming fungi from fungal culture available in NCBI (National Center for Biotechnology Information) and for authorized genomes in JGI (DOE Joint Genome Institute) (Table 1). The only genome not derived from fungal culture is the genome of *Umbilicaria pustulata* where the fungal genome was reconstructed with metagenomic sequencing of the lichen thallus instead [90]. We ran antiSMASH as described above on all thirteen additional genomes.

For seven of these genomes no gene set was available and gene models had to be predicted before running antiSMASH to improve the biosynthetic gene cluster detection. The de novo gene prediction and annotation of these genomes was done with MAKER [91] v2.31.8 using an iterative approach as recommended in the protocols of Campbell et al. [92]. In the first round of MAKER we used Hidden Markov Models (HMMs) generated with GeneMark-ES [93] v4.33 and SNAP [94] with hints from CEGMA [95] v2.4 (performed on iPlant [96]) as training data for gene finding. The first-round results were then converted to new SNAP and Augustus [97] v3.0.2 HMMs for the second round of MAKER resulting in the final set of gene models.

The genome and gene set completeness of all 15 lichen genomes was assessed based on evolutionarily-informed expectations of gene content with BUSCO (Benchmarking Universal Single-Copy Orthologs) [98] v3.0 and a lineage-specific set of 1315 Ascomycota single-copy orthologs from OrthoDB [99] v.9. Completeness estimates were visualized with the package ggplot2 [100] v.2.2.1 in R [101] v.3.4.4.

Additionally, we provide an overview on how richness of biosynthetic gene clusters found in lichenized fungi compare to the cluster richness found in non-lichenized fungi. Therefore, we combined the secondary metabolic gene cluster counts for the 15 genomes of lichenized fungi obtained from antiSMASH in this study with a recently published analysis from 57 genomes of representative fungal ascomycete and basidiomycete species [32].

3.5. Phylogenetic Analysis with Characterized Fungal Polyketides

We used phylogenetics to analyse the core enzymes of PKS and hybrid PKS-NRPS gene clusters in relation to characterized fungal biosynthetic gene clusters. Combining genome mining with phylogenetic analysis may provide clues for identifying promising clusters for the production of characteristic substance classes through phylogenetic relatedness to previously characterized genes [36,62,102].

Our phylogenetic approach is based on protein sequences of the ketosynthase (KS) domain of PKSs and hybrid PKS-NRPSs. The KS domains are the most conserved and essential part of a PKS cluster and are highly predictive of enzyme architecture and pathway association [52,70,71]. We included all KS sequences identified in the fifteen genomes of lichen-forming fungi (Table 2) that stem from genes with at least three PKS domains and with KS sequences longer than 200 bases. This resulted in a dataset of 413 KS sequences originating from lichen-forming fungi. Additionally, we made use of the MIBiG repository (Minimum Information about Biosynthetic Gene cluster [103]) to include experimentally characterized biosynthetic gene clusters [72]. We compiled a custom database with 131 entries of all fungal PKS records in MIBiG (downloaded January 2018). Further we included six animal fatty acid synthase (FAS) protein sequences as outgroup for the phylogenetic inference (NCBI Reference Sequence accessions: *Bos taurus* NP_001012687.1; *Gallus gallus* NP_990486.2; *Homo sapiens* NP_004095.4; *Mus musculus* NP_032014.3; *Rattus norvegicus* NP_059028.1; *Sus scrofa* NP_001093400.1).

We ran MAFFT [104,105] v7.309 in Geneious v9.1.8 (<https://www.geneious.com>) to generate a multiple sequence alignment of all 550 KS amino acid sequences resulting in an alignment with a length of 1043 characters. We chose the iterative refinement algorithm G-INS-i with a BLOSUM62 scoring matrix, a gap open penalty of 1.53 and an offset value of 0.123. We performed ProtTest [106] v.3.4.2 on our alignment to select the best-fitting substitution model of amino acid replacement for

subsequent tree reconstruction. Model-testing indicated LG + I + G as the best model for both the Akaike and the Bayesian Information Criterion. The phylogenetic tree was then generated with RAxML-HPC Black Box v.8.2.10 (Randomized Accelerated Maximum Likelihood) [107] on CIPRES v3.3 (Cyberinfrastructure for Phylogenetic Research) [108,109] with automated bootstrapping, protein sequence type, FAS sequences set as outgroup, estimating proportion of invariable sites (GTRGAMMA + I), protein substitution matrix LG, no empirical base frequencies and printing branch lengths. The tree was visualized with Figtree v.1.4.3 (<http://tree.bio.ed.ac.uk/software/figtree>) rooting the tree with the help of the FAS outgroup and ordering nodes increasingly.

Supplementary Materials: The following are available online, Table S1: Genome Statistics and Completeness of Lichen-Forming Fungi; Table S2: Overview of Biosynthetic Gene Clusters in *Evernia prunastri*; Table S3: Overview of Biosynthetic Gene Clusters in *Pseudevernia prunastri*; Table S4: Functional Annotation of *Evernia prunastri* Cluster Genes; Table S5: Functional Annotation of *Pseudevernia furfuracea* Cluster Genes; Table S6: Reciprocal Best Blast Hits of *Evernia prunastri* and *Pseudevernia furfuracea* Cluster Genes; Table S7: List of Putative Regulatory Genes in *Evernia prunastri* and *Pseudevernia furfuracea*; Figure S1: Phylogenetic Tree Based on the Ketosynthase (KS) Domain; Supplementary File S1: Ketosynthase (KS) Alignment for Phylogenetic Tree Reconstruction (FASTA); Supplementary Files S2: Predicted Gene Models for Seven Genomes of Lichen-Forming Fungi (GFF3); All output of antiSMASH analyses are available upon request from the corresponding author.

Author Contributions: Conceptualization, A.C., F.D.G. and I.S.; formal analysis, A.C.; investigation A.C.; methodology A.C.; validation, F.D.G., I.S. and H.B.B.; writing—original draft preparation, A.C.; visualization, A.C.; writing—review and editing, F.D.G., I.S. and H.B.B.; supervision, F.D.G. and I.S.; project administration, I.S.

Funding: This research was supported by the research funding program Landes-Offensive zur Entwicklung Wissenschaftlich-Ökonomischer Exzellenz (LOEWE) of Hesse's Ministry of Higher Education, Research and the Arts through the Priority Programme Integrative Pilzforschung (IPF) and the Translational Biodiversity Genomics (TBG) Center. Additional funding came from the RMU Initiativfond "Biotic interactions". Some analyses were performed on the FUCHS cluster of the Center for Scientific Computing (CSC) of the Goethe University in Frankfurt am Main.

Acknowledgments: We thank the following colleagues for permission to use genomes prior to full publication: Daniele Armaleo (Duke University)—*Cladonia grayi*, Ólafur S. Andr sson (University of Iceland)—*Lobaria pulmonaria* and Paul Dyer (University of Nottingham)—*Xanthoria parietina*.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Nash, T.H. *Lichen Biology*, 2nd ed.; Nash, T.H., Ed.; Cambridge University Press: Cambridge, UK, 2008; ISBN 9780521871624.
- L cking, R.; Hodkinson, B.P.; Leavitt, S.D. The 2016 classification of lichenized fungi in the Ascomycota and Basidiomycota—Approaching one thousand genera. *Bryologist* **2016**, *119*, 361–416. [[CrossRef](#)]
- Asplund, J.; Wardle, D.A. How lichens impact on terrestrial community and ecosystem properties. *Biol. Rev.* **2017**, *92*, 1720–1738. [[CrossRef](#)] [[PubMed](#)]
- Seaward, M.R.D. Environmental role of lichens. In *Lichen Biology*; Nash, T.H., Ed.; Cambridge University Press: Cambridge, UK, 2008; pp. 274–298. ISBN 9780521871624.
- Grube, M.; Wedin, M. Lichenized fungi and the evolution of symbiotic organization. *Microbiol. Spectr.* **2016**, *4*, 1–17. [[CrossRef](#)]
- Cernava, T.; Berg, G.; Grube, M. High life expectancy of bacteria on lichens. *Microb. Ecol.* **2016**, *72*, 510–513. [[CrossRef](#)] [[PubMed](#)]
- Aschenbrenner, I.A.; Cernava, T.; Berg, G.; Grube, M. Understanding microbial multi-species symbioses. *Front. Microbiol.* **2016**, *7*, 180. [[CrossRef](#)] [[PubMed](#)]
- Spribille, T.; Tuovinen, V.; Resl, P.; Vanderpool, D.; Wolinski, H.; Aime, M.C.; Schneider, K.; Stabentheiner, E.; Toome-Heller, M.; Thor, G.; et al. Basidiomycete yeasts in the cortex of ascomycete macrolichens. *Science* **2016**, *353*, 488–492. [[CrossRef](#)] [[PubMed](#)]
- Fern ndez-Mendoza, F.; Fleischhacker, A.; Kopun, T.; Grube, M.; Muggia, L. ITS1 metabarcoding highlights low specificity of lichen mycobiomes at a local scale. *Mol. Ecol.* **2017**, *26*, 4811–4830. [[CrossRef](#)] [[PubMed](#)]

10. Dal Grande, F.; Rolshausen, G.; Divakar, P.K.; Crespo, A.; Otte, J.; Schleuning, M.; Schmitt, I. Environment and host identity structure communities of green algal symbionts in lichens. *New Phytol.* **2018**, *217*, 277–289. [[CrossRef](#)]
11. Machado, H.; Tuttle, R.N.; Jensen, P.R. Omics-based natural product discovery and the lexicon of genome mining. *Curr. Opin. Microbiol.* **2017**, *39*, 136–142. [[CrossRef](#)] [[PubMed](#)]
12. Cernava, T.; Erlacher, A.; Aschenbrenner, I.A.; Krug, L.; Lassek, C.; Riedel, K.; Grube, M.; Berg, G. Deciphering functional diversification within the lichen microbiota by meta-omics. *Microbiome* **2017**, *5*, 82. [[CrossRef](#)]
13. Moya, P.; Molins, A.; Martínez-Alberola, F.; Muggia, L.; Barreno, E. Unexpected associated microalgal diversity in the lichen *Ramalina farinacea* is uncovered by pyrosequencing analyses. *PLoS ONE* **2017**, *12*, e0175091. [[CrossRef](#)] [[PubMed](#)]
14. Stocker-Wörgötter, E. Metabolic diversity of lichen-forming ascomycetous fungi: Culturing, polyketide and shikimate metabolite production, and PKS genes. *Nat. Prod. Rep. Prod. Rep.* **2008**, *25*, 188–200. [[CrossRef](#)] [[PubMed](#)]
15. Molnár, K.; Farkas, E. Current results on biological activities of lichen secondary metabolites: A review. *Z. Naturforsch. C* **2010**, *65*, 157–173. [[CrossRef](#)] [[PubMed](#)]
16. Elix, J.A.; Stocker-Wörgötter, E. Biochemistry and secondary metabolites. In *Lichen Biology*; Nash, T.H., Ed.; Cambridge University Press: Cambridge, UK, 2008; pp. 104–133. ISBN 9780521871624.
17. Huneck, S.; Yoshimura, I. *Identification of Lichen Substances*; Springer: Berlin/Heidelberg, Germany, 1996; ISBN 978-3-642-85245-9.
18. Rundel, P.W. The ecological role of secondary lichen substances. *Biochem. Syst. Ecol.* **1978**, *6*, 157–170. [[CrossRef](#)]
19. Huneck, S. The significance of lichens and their metabolites. *Naturwissenschaften* **1999**, *86*, 559–570. [[CrossRef](#)] [[PubMed](#)]
20. Lawrey, J.D. Biological role of lichen substances. *Bryologist* **1986**, *89*, 111. [[CrossRef](#)]
21. Ranković, B.; Kosanić, M. *Lichen Secondary Metabolites*; Ranković, B., Ed.; Springer: Cham, Switzerland, 2015; ISBN 978-3-319-13373-7.
22. Boustie, J.; Grube, M. Lichens—A promising source of bioactive secondary metabolites. *Plant Genet. Resour. Charact. Util.* **2005**, *3*, 273–287. [[CrossRef](#)]
23. Zhou, R.; Yang, Y.; Park, S.-Y.; Nguyen, T.T.; Seo, Y.-W.; Lee, K.H.; Lee, J.H.; Kim, K.K.; Hur, J.-S.; Kim, H. The lichen secondary metabolite atranorin suppresses lung cancer cell motility and tumorigenesis. *Sci. Rep.* **2017**, *7*, 8136. [[CrossRef](#)]
24. Crawford, S.D. Lichens used in traditional medicine. In *Lichen Secondary Metabolites*; Springer: Cham, Switzerland, 2015; pp. 27–80. ISBN 9783319133744.
25. Joulain, D.; Tabacchi, R. Lichen extracts as raw materials in perfumery. Part 1: Oakmoss. *Flavour Fragr. J.* **2009**, *24*, 49–61. [[CrossRef](#)]
26. Joulain, D.; Tabacchi, R. Lichen extracts as raw materials in perfumery. Part 2: Treemoss. *Flavour Fragr. J.* **2009**, *24*, 105–116. [[CrossRef](#)]
27. Calcott, M.J.; Ackerley, D.F.; Knight, A.; Keyzers, R.A.; Owen, J.G. Secondary metabolism in the lichen symbiosis. *Chem. Soc. Rev.* **2018**, *47*, 1730–1760. [[CrossRef](#)] [[PubMed](#)]
28. Crawford, J.M.; Townsend, C.A. New insights into the formation of fungal aromatic polyketides. *Nat. Rev. Microbiol.* **2010**, *8*, 879–889. [[CrossRef](#)] [[PubMed](#)]
29. Schmitt, I.; Martín, M.P.; Kautz, S.; Lumbsch, H.T. Diversity of non-reducing polyketide synthase genes in the Pertusariales (lichenized Ascomycota): A phylogenetic perspective. *Phytochemistry* **2005**, *66*, 1241–1253. [[CrossRef](#)] [[PubMed](#)]
30. Keller, N.P.; Turner, G.; Bennett, J.W. Fungal secondary metabolism—From biochemistry to genomics. *Nat. Rev. Microbiol.* **2005**, *3*, 937–947. [[CrossRef](#)] [[PubMed](#)]
31. Cox, R.J.; Skellam, E.; Williams, K. Biosynthesis of fungal polyketides. In *Physiology and Genetics*; Springer: Cham, Switzerland, 2018; pp. 385–412.
32. Rokas, A.; Wisecaver, J.H.; Lind, A.L. The birth, evolution and death of metabolic gene clusters in fungi. *Nat. Rev. Microbiol.* **2018**, *16*, 731–744. [[CrossRef](#)] [[PubMed](#)]
33. Cimermancic, P.; Medema, M.H.; Claesen, J.; Kurita, K.; Wieland Brown, L.C.; Mavrommatis, K.; Pati, A.; Godfrey, P.A.; Koehrsen, M.; Clardy, J.; et al. Insights into secondary metabolism from a global analysis of prokaryotic biosynthetic gene clusters. *Cell* **2014**, *158*, 412–421. [[CrossRef](#)]

34. Doroghazi, J.R.; Albright, J.C.; Goering, A.W.; Ju, K.-S.; Haines, R.R.; Tchalukov, K.A.; Labeda, D.P.; Kelleher, N.L.; Metcalf, W.W. A roadmap for natural product discovery based on large-scale genomics and metabolomics. *Nat. Chem. Biol.* **2014**, *10*, 963–968. [[CrossRef](#)]
35. Zhao, H.; Medema, M.H. Standardization for natural product synthetic biology. *Nat. Prod. Rep.* **2016**, *33*, 920–924. [[CrossRef](#)]
36. Nielsen, J.C.; Grijsseels, S.; Prigent, S.; Ji, B.; Dainat, J.; Nielsen, K.F.; Frisvad, J.C.; Workman, M.; Nielsen, J. Global analysis of biosynthetic gene clusters reveals vast potential of secondary metabolite production in *Penicillium* species. *Nat. Microbiol.* **2017**, *2*, 17044. [[CrossRef](#)]
37. Brakhage, A.A. Regulation of fungal secondary metabolism. *Nat. Rev. Microbiol.* **2013**, *11*, 21–32. [[CrossRef](#)]
38. Schümann, J.; Hertweck, C. Advances in cloning, functional analysis and heterologous expression of fungal polyketide synthase genes. *J. Biotechnol.* **2006**, *124*, 690–703. [[CrossRef](#)]
39. Hertweck, C. Hidden biosynthetic treasures brought to light. *Nat. Chem. Biol.* **2009**, *5*, 450–452. [[CrossRef](#)] [[PubMed](#)]
40. Bergmann, S.; Schümann, J.; Scherlach, K.; Lange, C.; Brakhage, A.A.; Hertweck, C. Genomics-driven discovery of PKS-NRPS hybrid metabolites from *Aspergillus nidulans*. *Nat. Chem. Biol.* **2007**, *3*, 213–217. [[CrossRef](#)] [[PubMed](#)]
41. Abdel-Hameed, M.; Bertrand, R.L.; Piercey-Normore, M.D.; Sorensen, J.L. Putative identification of the usnic acid biosynthetic gene cluster by de novo whole-genome sequencing of a lichen-forming fungus. *Fungal Biol.* **2016**, *120*, 306–316. [[CrossRef](#)] [[PubMed](#)]
42. Muggia, L.; Schmitt, I.; Grube, M. Lichens as treasure chests of natural products. *Sim News* **2009**, *59*, 85–97.
43. Armaleo, D.; Sun, X.; Culberson, C. Insights from the first putative biosynthetic gene cluster for a lichen depside and depsidone. *Mycologia* **2011**, *103*, 741–754. [[CrossRef](#)]
44. Abdel-Hameed, M.; Bertrand, R.L.; Piercey-Normore, M.D.; Sorensen, J.L. Identification of 6-hydroxymellein synthase and accessory genes in the lichen *Cladonia uncialis*. *J. Nat. Prod.* **2016**, *79*, 1645–1650. [[CrossRef](#)]
45. Bertrand, R.L.; Sorensen, J.L. A comprehensive catalogue of polyketide synthase gene clusters in lichenizing fungi. *J. Ind. Microbiol. Biotechnol.* **2018**, *45*, 1067–1081. [[CrossRef](#)]
46. Meiser, A.; Otte, J.; Schmitt, I.; Dal Grande, F. Sequencing genomes from mixed DNA samples—Evaluating the metagenome skimming approach in lichenized fungi. *Sci. Rep.* **2017**, *7*, 14881. [[CrossRef](#)]
47. Wang, Y.; Geng, C.; Yuan, X.; Hua, M.; Tian, F.; Li, C. Identification of a putative polyketide synthase gene involved in usnic acid biosynthesis in the lichen *Nephromopsis pallescens*. *PLoS ONE* **2018**, *13*, e0199110. [[CrossRef](#)]
48. Taguchi, H.; Sankawa, U.; Shibata, S. Biosynthesis of natural products. VI. Biosynthesis of usnic acid in lichens. (1). A general scheme of biosynthesis of usnic acid. *Chem. Pharm. Bull. (Tokyo)* **1969**, *17*, 2054–2060. [[CrossRef](#)] [[PubMed](#)]
49. Barton, D.H.R.; Deflorin, A.M.; Edwards, O.E. 108. The synthesis of usnic acid. *J. Chem. Soc.* **1956**, *108*, 530–534. [[CrossRef](#)]
50. Gagunashvili, A.N.; Davíðsson, S.P.; Jónsson, Z.O.; Andrésson, Ó.S. Cloning and heterologous transcription of a polyketide synthase gene from the lichen *Solorina crocea*. *Mycol. Res.* **2009**, *113*, 354–363. [[CrossRef](#)] [[PubMed](#)]
51. Chooi, Y.-H.; Stalker, D.M.; Davis, M.A.; Fujii, I.; Elix, J.A.; Louwhoff, S.H.J.J.; Lawrie, A.C. Cloning and sequence characterization of a non-reducing polyketide synthase gene from the lichen *Xanthoparmelia semiviridis*. *Mycol. Res.* **2008**, *112*, 147–161. [[CrossRef](#)] [[PubMed](#)]
52. Ziemert, N.; Jensen, P.R. Phylogenetic approaches to natural product structure prediction. In *Methods in Enzymology*; Elsevier Inc.: San Diego, CA, USA, 2012; Volume 517, pp. 161–182. ISBN 9780124046344.
53. Wang, Y.; Wang, J.; Cheong, Y.H.; Hur, J.-S. Three new non-reducing polyketide synthase genes from the lichen-forming fungus *Usnea longissima*. *Mycobiology* **2014**, *42*, 34–40. [[CrossRef](#)] [[PubMed](#)]
54. Bertrand, R.L.; Abdel-Hameed, M.; Sorensen, J.L. Lichen biosynthetic gene clusters part II: Homology mapping suggests a functional diversity. *J. Nat. Prod.* **2018**, *81*, 732–748. [[CrossRef](#)] [[PubMed](#)]
55. Purvis, O.W.; Coppins, B.J.; Hawksworth, D.L.; James, P.W.; Moore, D.M. (Eds.) *The Lichen Flora of Great Britain and Ireland*; Natural History Museum Publications in association with The British Lichen Society: London, UK, 1992; ISBN 0952304902.
56. Brodo, I.M.; Sharnoff, S.D.; Sharnoff, S. (Eds.) *Lichens of North America*; Yale University Press: New Haven, CO, USA; London, UK, 2001; ISBN 0-300-08249-5.

57. Stenroos, S.; Velmala, S.; Pykälä, J.; Ahti, T. (Eds.) *Lichens of Finland*; Finnish Museum of Natural History LUMOS, University of Helsinki: Helsinki, Finland, 2016; ISBN 978-951-51-2266-7.
58. Posner, B.; Feige, G.B.; Leuckert, C. Beiträge zur Chemie der Flechtengattung *Lasallia* Mérat/On the Chemistry of the lichen genus *Lasallia* Merat. *Z. Naturforsch. C* **1991**, *46*, 19–27. [[CrossRef](#)]
59. Dal Grande, F.; Meiser, A.; Greshake Tzovaras, B.; Otte, J.; Ebersberger, I.; Schmitt, I. The draft genome of the lichen-forming fungus *Lasallia hispanica* (Frey) Sancho & A. Crespo. *Lichenologist* **2018**, *50*, 329–340. [[CrossRef](#)]
60. Park, S.-Y.; Choi, J.; Lee, G.-W.; Park, C.-H.; Kim, J.A.; Oh, S.-O.; Lee, Y.-H.; Hur, J.-S. Draft Genome Sequence of *Endocarpon pusillum* Strain KoLRILF000583. *Genome Announc.* **2014**, *2*, e00452-14. [[CrossRef](#)]
61. Wang, Y.-Y.; Liu, B.; Zhang, X.-Y.; Zhou, Q.-M.; Zhang, T.; Li, H.; Yu, Y.-F.; Zhang, X.-L.; Hao, X.-Y.; Wang, M.; et al. Genome characteristics reveal the impact of lichenization on lichen-forming fungus *Endocarpon pusillum* Hedwig (Verrucariales, Ascomycota). *BMC Genom.* **2014**, *15*, 34. [[CrossRef](#)]
62. Blin, K.; Kim, H.U.; Medema, M.H.; Weber, T. Recent development of antiSMASH and other computational approaches to mine secondary metabolite biosynthetic gene clusters. *Brief. Bioinform.* **2017**, *bbx146*, 1–11. [[CrossRef](#)] [[PubMed](#)]
63. Stocker-Wörgötter, E. Biochemical diversity and ecology of lichen-forming fungi: Lichen substances, chemosyndromic variation and origin of polyketide-type metabolites (biosynthetic pathways). In *Recent Advances in Lichenology*; Springer: New Delhi, India, 2015; pp. 161–179. ISBN 9788132222354.
64. Wang, X.; Zhang, X.; Liu, L.; Xiang, M.; Wang, W.; Sun, X.; Che, Y.; Guo, L.; Liu, G.; Guo, L.; et al. Genomic and transcriptomic analysis of the endophytic fungus *Pestalotiopsis fici* reveals its lifestyle and high potential for synthesis of natural products. *BMC Genom.* **2015**, *16*, 28. [[CrossRef](#)]
65. Crawford, J.M.; Dancy, B.C.R.; Hill, E.A.; Udway, D.W.; Townsend, C.A. Identification of a starter unit acyl-carrier protein transacylase domain in an iterative type I polyketide synthase. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 16728–16733. [[CrossRef](#)]
66. Crawford, J.M.; Thomas, P.M.; Scheerer, J.R.; Vagstad, A.L.; Kelleher, N.L.; Townsend, C.A. Deconstruction of iterative multidomain polyketide synthase function. *Science* **2008**, *320*, 243–246. [[CrossRef](#)]
67. Crawford, J.M.; Korman, T.P.; Labonte, J.W.; Vagstad, A.L.; Hill, E.A.; Kamari-Bidkorpheh, O.; Tsai, S.-C.; Townsend, C.A. Structural basis for biosynthetic programming of fungal aromatic polyketide cyclization. *Nature* **2009**, *461*, 1139–1143. [[CrossRef](#)]
68. Cacho, R.A.; Tang, Y.; Chooi, Y.-H. Next-generation sequencing approach for connecting secondary metabolites to biosynthetic gene clusters in fungi. *Front. Microbiol.* **2015**, *5*, 1–16. [[CrossRef](#)] [[PubMed](#)]
69. Knox, B.P.; Keller, N.P. Key players in the regulation of fungal secondary metabolism. In *Biosynthesis and Molecular Genetics of Fungal Secondary Metabolites, Volume 2*; Zeilinger, S., Martín, J.-F., García-Estrada, C., Eds.; Springer: New York, NY, USA, 2015; Volume 2, pp. 13–28. ISBN 978-1-4939-2530-8.
70. Kroken, S.; Glass, N.L.; Taylor, J.W.; Yoder, O.C.; Turgeon, B.G. Phylogenomic analysis of type I polyketide synthase genes in pathogenic and saprobic ascomycetes. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 15670–15675. [[CrossRef](#)] [[PubMed](#)]
71. Jenke-Kodama, H.; Sandmann, A.; Müller, R.; Dittmann, E. Evolutionary implications of bacterial polyketide synthases. *Mol. Biol. Evol.* **2005**, *22*, 2027–2039. [[CrossRef](#)]
72. Li, Y.F.; Tsai, K.J.S.; Harvey, C.J.B.; Li, J.J.; Ary, B.E.; Berlew, E.E.; Boehman, B.L.; Findley, D.M.; Friant, A.G.; Gardner, C.A.; et al. Comprehensive curation and analysis of fungal biosynthetic gene clusters of published natural products. *Fungal Genet. Biol.* **2016**, *89*, 18–28. [[CrossRef](#)]
73. Nguyen, K.-H.; Chollet-Krugler, M.; Gouault, N.; Tomasi, S. UV-protectant metabolites from lichens and their symbiotic partners. *Nat. Prod. Rep.* **2013**, *30*, 1490. [[CrossRef](#)]
74. Fulton, T.R.; Ibrahim, N.; Losada, M.C.; Grzegorski, D.; Tkacz, J.S. A melanin polyketide synthase (PKS) gene from *Nodulisporium* sp. that shows homology to the *pks1* gene of *Colletotrichum lagenarium*. *Mol. Gen. Genet.* **1999**, *262*, 714–720. [[CrossRef](#)] [[PubMed](#)]
75. Zhang, A.; Lu, P.; Dahl-Roshak, A.M.; Pares, P.S.; Kennedy, S.; Tkacz, J.S.; An, Z. Efficient disruption of a polyketide synthase gene (*pks1*) required for melanin synthesis through *Agrobacterium*-mediated transformation of *Glarea lozoyensis*. *Mol. Genet. Genom.* **2003**, *268*, 645–655. [[CrossRef](#)]
76. Sanchez, J.F.; Chiang, Y.-M.; Szweczyk, E.; Davidson, A.D.; Ahuja, M.; Elizabeth Oakley, C.; Woo Bok, J.; Keller, N.; Oakley, B.R.; Wang, C.C.C. Molecular genetic analysis of the orsellinic acid/F9775 genecluster of *Aspergillus nidulans*. *Mol. Biosyst.* **2010**, *6*, 587–593. [[CrossRef](#)] [[PubMed](#)]

77. Blin, K.; Wolf, T.; Chevrette, M.G.; Lu, X.; Schwalen, C.J.; Kautsar, S.A.; Suarez Duran, H.G.; de los Santos, E.L.C.; Kim, H.U.; Nave, M.; et al. antiSMASH 4.0—Improvements in chemistry prediction and gene cluster boundary identification. *Nucleic Acids Res.* **2017**, *45*, W36–W41. [[CrossRef](#)] [[PubMed](#)]
78. Conesa, A.; Götz, S. Blast2GO: A comprehensive suite for functional analysis in plant genomics. *Int. J. Plant Genom.* **2008**. [[CrossRef](#)] [[PubMed](#)]
79. Altschul, S. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res.* **1997**, *25*, 3389–3402. [[CrossRef](#)]
80. Benson, D.A.; Cavanaugh, M.; Clark, K.; Karsch-Mizrachi, I.; Lipman, D.J.; Ostell, J.; Sayers, E.W. GenBank. *Nucleic Acids Res.* **2017**, *45*, D37–D42. [[CrossRef](#)]
81. Jones, P.; Binns, D.; Chang, H.-Y.; Fraser, M.; Li, W.; McAnulla, C.; McWilliam, H.; Maslen, J.; Mitchell, A.; Nuka, G.; et al. InterProScan 5: Genome-scale protein function classification. *Bioinformatics* **2014**, *30*, 1236–1240. [[CrossRef](#)]
82. Finn, R.D.; Coghill, P.; Eberhardt, R.Y.; Eddy, S.R.; Mistry, J.; Mitchell, A.L.; Potter, S.C.; Punta, M.; Qureshi, M.; Sangrador-Vegas, A.; et al. The Pfam protein families database: Towards a more sustainable future. *Nucleic Acids Res.* **2016**, *44*, D279–D285. [[CrossRef](#)]
83. Ward, N.; Moreno-Hagelsieb, G. Quickly finding orthologs as reciprocal best hits with BLAT, LAST, and UBLAST: How much do we miss? *PLoS ONE* **2014**, *9*, e101850. [[CrossRef](#)]
84. Moreno-Hagelsieb, G.; Latimer, K. Choosing BLAST options for better detection of orthologs as reciprocal best hits. *Bioinformatics* **2008**, *24*, 319–324. [[CrossRef](#)] [[PubMed](#)]
85. Camacho, C.; Coulouris, G.; Avagyan, V.; Ma, N.; Papadopoulos, J.; Bealer, K.; Madden, T.L. BLAST+: Architecture and applications. *BMC Bioinform.* **2009**, *10*, 421. [[CrossRef](#)] [[PubMed](#)]
86. Sullivan, M.J.; Petty, N.K.; Beatson, S.A. Easyfig: A genome comparison visualizer. *Bioinformatics* **2011**, *27*, 1009–1010. [[CrossRef](#)] [[PubMed](#)]
87. Kjærboelling, I.; Vesth, T.C.; Frisvad, J.C.; Nybo, J.L.; Theobald, S.; Kuo, A.; Bowyer, P.; Matsuda, Y.; Mondo, S.; Lyhne, E.K.; et al. Linking secondary metabolites to gene clusters through genome sequencing of six diverse *Aspergillus* species. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, E753–E761. [[CrossRef](#)]
88. Shen, W.; Le, S.; Li, Y.; Hu, F. SeqKit: A cross-platform and ultrafast toolkit for FASTA/Q file manipulation. *PLoS ONE* **2016**, *11*, e0163962. [[CrossRef](#)]
89. Rice, P.; Longden, I.; Bleasby, A. EMBOSS: The European Molecular Biology Open Software Suite. *Trends Genet.* **2000**, *16*, 276–277. [[CrossRef](#)]
90. Dal Grande, F.; Sharma, R.; Meiser, A.; Rolshausen, G.; Büdel, B.; Mishra, B.; Thines, M.; Otte, J.; Pfenninger, M.; Schmitt, I. Adaptive differentiation coincides with local bioclimatic conditions along an elevational cline in populations of a lichen-forming fungus. *BMC Evol. Biol.* **2017**, *17*, 93. [[CrossRef](#)]
91. Holt, C.; Yandell, M. MAKER2: An annotation pipeline and genome-database management tool for second-generation genome projects. *BMC Bioinform.* **2011**, *12*, 491. [[CrossRef](#)] [[PubMed](#)]
92. Campbell, M.S.; Holt, C.; Moore, B.; Yandell, M. Genome annotation and curation using MAKER and MAKER-P. In *Current Protocols in Bioinformatics*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2014; Volume 48, pp. 4.11.1–4.11.39. ISBN 0471250953.
93. Ter-Hovhannisyan, V.; Lomsadze, A.; Chernoff, Y.O.; Borodovsky, M. Gene prediction in novel fungal genomes using an ab initio algorithm with unsupervised training. *Genome Res.* **2008**, *18*, 1979–1990. [[CrossRef](#)] [[PubMed](#)]
94. Korf, I. Gene finding in novel genomes. *BMC Bioinform.* **2004**, *5*, 59. [[CrossRef](#)] [[PubMed](#)]
95. Parra, G.; Bradnam, K.; Korf, I. CEGMA: A pipeline to accurately annotate core genes in eukaryotic genomes. *Bioinformatics* **2007**, *23*, 1061–1067. [[CrossRef](#)] [[PubMed](#)]
96. Goff, S.A.; Vaughn, M.; McKay, S.; Lyons, E.; Stapleton, A.E.; Gessler, D.; Matasci, N.; Wang, L.; Hanlon, M.; Lenards, A.; et al. The iPlant collaborative: Cyberinfrastructure for plant biology. *Front. Plant Sci.* **2011**, *2*, 34. [[CrossRef](#)] [[PubMed](#)]
97. Stanke, M.; Schöffmann, O.; Morgenstern, B.; Waack, S. Gene prediction in eukaryotes with a generalized hidden Markov model that uses hints from external sources. *BMC Bioinform.* **2006**, *7*, 62. [[CrossRef](#)]
98. Simão, F.A.; Waterhouse, R.M.; Ioannidis, P.; Kriventseva, E.V.; Zdobnov, E.M. BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **2015**, *31*, 3210–3212. [[CrossRef](#)] [[PubMed](#)]

99. Zdobnov, E.M.; Tegenfeldt, F.; Kuznetsov, D.; Waterhouse, R.M.; Simão, F.A.; Ioannidis, P.; Seppey, M.; Loetscher, A.; Kriventseva, E.V. OrthoDB v9.1: Cataloging evolutionary and functional annotations for animal, fungal, plant, archaeal, bacterial and viral orthologs. *Nucleic Acids Res.* **2017**, *45*, D744–D749. [[CrossRef](#)] [[PubMed](#)]
100. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis*; Springer: New York, NY, USA, 2009; ISBN 978-0-387-98140-6.
101. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2015; Available online: <https://www.R-project.org/> (accessed on 29 March 2018).
102. Bahram, M.; Hildebrand, F.; Forslund, S.K.; Anderson, J.L.; Soudzilovskaia, N.A.; Bodegom, P.M.; Bengtsson-Palme, J.; Anslan, S.; Coelho, L.P.; Harend, H.; et al. Structure and function of the global topsoil microbiome. *Nature* **2018**, *560*, 233–237. [[CrossRef](#)] [[PubMed](#)]
103. Medema, M.H.; Kottmann, R.; Yilmaz, P.; Cummings, M.; Biggins, J.B.; Blin, K.; de Bruijn, I.; Chooi, Y.H.; Claesen, J.; Coates, R.C.; et al. Minimum Information about a Biosynthetic Gene cluster. *Nat. Chem. Biol.* **2015**, *11*, 625–631. [[CrossRef](#)]
104. Katoh, K. MAFFT: A novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* **2002**, *30*, 3059–3066. [[CrossRef](#)]
105. Katoh, K.; Standley, D.M. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **2013**, *30*, 772–780. [[CrossRef](#)]
106. Darriba, D.; Taboada, G.L.; Doallo, R.; Posada, D. ProfTest 3: Fast selection of best-fit models of protein evolution. *Bioinformatics* **2011**, *27*, 1164–1165. [[CrossRef](#)]
107. Stamatakis, A. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **2014**, *30*, 1312–1313. [[CrossRef](#)] [[PubMed](#)]
108. Miller, M.A.; Pfeiffer, W.; Schwartz, T. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In *2010 Gateway Computing Environments Workshop (GCE)*; IEEE: New Orleans, LA, USA, 2010; pp. 1–8.
109. Miller, M.A.; Schwartz, T.; Pickett, B.E.; He, S.; Klem, E.B.; Scheuermann, R.H.; Passarotti, M.; Kaufman, S.; O’Leary, M.A. A RESTful API for access to phylogenetic tools via the CIPRES Science Gateway. *Evol. Bioinform.* **2015**, *11*, EBO.S21501. [[CrossRef](#)] [[PubMed](#)]

Sample Availability: Samples of the compounds are not available from the authors.



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).