

# Genome-Wide Analysis of Biosynthetic Gene Cluster Reveals Correlated Gene Loss with Absence of Usnic Acid in Lichen-Forming Fungi

David Pizarro <sup>1,\*</sup>, Pradeep K. Divakar <sup>1,\*</sup>, Felix Grewe<sup>2</sup>, Ana Crespo<sup>1</sup>, Francesco Dal Grande<sup>3,4</sup>, and Helge Thorsten Lumbsch<sup>2</sup>

<sup>1</sup>Departamento de Farmacología, Farmacognosia y Botánica, Facultad de Farmacia, Universidad Complutense de Madrid, Madrid 28040, Spain

<sup>2</sup>Department of Science & Education, The Field Museum, Chicago, Illinois

<sup>3</sup>Senckenberg Biodiversity and Climate Research Centre, Frankfurt am Main D-60325, Germany

<sup>4</sup>LOEWE Center for Translational Biodiversity Genomics, Frankfurt am Main D-60325, Germany

\*Corresponding authors: E-mails: dapizarr@ucm.es; pdivakar@farm.ucm.es

Accepted: 27 August 2020

## Abstract

Lichen-forming fungi are known to produce a large number of secondary metabolites. Some metabolites are deposited in the cortical layer of the lichen thallus where they exert important ecological functions, such as UV filtering. The fact that closely related lineages of lichen-forming fungi can differ in cortical chemistry suggests that natural product biosynthesis in lichens can evolve independent from phylogenetic constraints. Usnic acid is one of the major cortical pigments in lichens. Here we used a comparative genomic approach on 46 lichen-forming fungal species of the Lecanoromycetes to elucidate the biosynthetic gene content and evolution of the gene cluster putatively responsible for the biosynthesis of usnic acid. Whole-genome sequences were gathered from taxa belonging to different orders and families of Lecanoromycetes, where Parmeliaceae is the most well-represented taxon, and analyzed with a variety of genomic tools. The highest number of biosynthetic gene clusters was found in *Evernia prunastri*, *Pannoparmelia angustata*, and *Parmotrema austrosinense*, respectively, and lowest in *Canoparmelia nairobiensis*, *Bulbothrix sensibilis*, and *Hypotrachyna scytodes*. We found that all studied species producing usnic acid contain the putative usnic acid biosynthetic gene cluster, whereas the cluster was absent in all genomes of species lacking usnic acid. The absence of the gene cluster was supported by an additional unsuccessful search for  $\beta$ -ketoacylsynthase, the most conserved domain of the gene cluster, in the genomes of species lacking usnic acid. The domain architecture of this PKS cluster—homologous to the already known usnic acid PKS cluster (MPAS) and CYT450 (MPO)—varies within the studied species, whereas the gene arrangement is highly similar in closely related taxa. We hypothesize that the ancestor of these lichen-forming fungi contained the putative usnic acid producing PKS cluster and that the gene cluster was lost repeatedly during the evolution of these groups. Our study provides insight into the genomic adaptations to the evolutionary success of these lichen-forming fungal species and sets a baseline for further exploration of biosynthetic gene content and its evolutionary significance.

**Key words:** ascomycota, evolution, comparative-genomics, metabolic gene cluster, usnic acid, Parmeliaceae.

## Significance

The fact that closely related lineages of lichen-forming fungi can differ in cortical chemistry suggests that natural product biosynthesis in lichens can evolve independent from phylogenetic constraints. Here we used a comparative genomic approach on 46 lichen-forming fungal species of the Lecanoromycetes to elucidate the biosynthetic gene content and evolution of the gene cluster putatively responsible for the biosynthesis of usnic acid. Whole-genome sequences were gathered and analyzed with a variety of genomic tools. All species lacking usnic acid production also

© The Author(s) 2020. Published by Oxford University Press on behalf of the Society for Molecular Biology and Evolution.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

lacked the putative biosynthetic gene cluster for usnic acid biosynthesis. This is the first study to report that the presence/absence of an entire biosynthetic gene cluster is putatively responsible for the biosynthesis of the corresponding secondary metabolite in lichen-forming fungi. Our results suggest that gene loss is an evolutionary mechanism underlying secondary metabolite diversity in lichens.

## Introduction

Lichens are fungi that form stable symbiotic relationships with algae and/or cyanobacteria (Nash 2008; Crespo et al. 2014). Lichen-forming fungi house in vegetative structures (lichen thalli) the photosynthetically active partner. Lichens are known to produce a large number of secondary metabolites (many of them extrolites) with almost 1000 known substances, the large majority of which exclusively found in lichen-forming fungi (Hawksworth 1976; Elix et al. 1984; Huneck and Yoshimura 1996; Lumbsch 1998; Stocker-Wörgötter 2008; Crespo et al. 2010; Calcott et al. 2018). Extrolites are deposited extracellularly, mostly in the medullary layer of the lichen thallus or in the cortical layer. Only a small number of substances occur in the cortex. Among those, coupled phenolics, which originate from polyketide pathways, such as depsides, depsidones, and usnic acids are found almost exclusively in lichens. In macrolichens, the most common cortical substances are the depsides atranorin or usnic acid. Usnic acids and cortical substances in general protect the photobiont from solar radiations as UV-B (reviewed in Solhaug and Gauslaa [2012]). The occurrence of these cortical substances is usually constant within clades (e.g., genera) and has been used to circumscribe genera (Lumbsch 1998) especially in Parmeliaceae, which is the largest family of lichen-forming fungi (Elix 1993; Crespo et al. 2010; Thell et al. 2012). This said cortical chemistry can be highly variable even among closely related genera, making the occurrence of these substances often scattered over the phylogenetic tree. This pattern suggests that natural product biosynthesis in lichens may evolve independently from phylogenetic constraints.

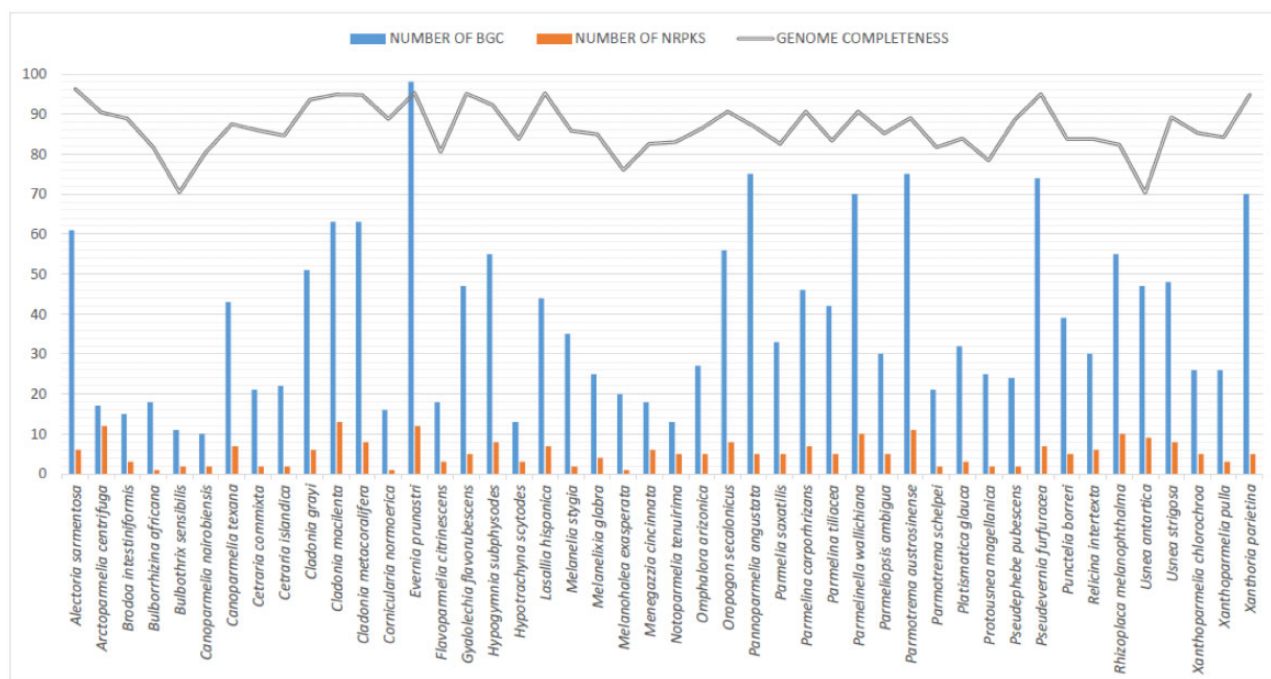
A previous ancestral character state reconstruction analysis suggested that the ancestor of parmelioid lichens, which is the largest group within Parmeliaceae, contained usnic acid and that this substance has been lost and replaced by atranorin several times independently (Divakar et al. 2013). However, the genetic mechanisms behind these repeated losses remained elusive.

Experimental work on lichen secondary metabolism is hampered by the fact that lichen-forming fungi are extremely slow growing. However, the recent applications of metagenomic approaches to the study of the lichen symbiosis can fortunately obviate the limitations imposed by the slow-growing nature of lichen-forming fungi and shed light on the underlying genetics of their secondary metabolism (Bertrand et al. 2018a). For example, a comparative genomic study identified a polyketide synthase (PKS) gene cluster putatively encoding for the biosynthesis of usnic acid (Abdel-

Hameed et al. 2016). Genomic data sets are also increasing our knowledge of the presence and domain architecture of PKSs, which are key for the biosynthesis of fungal phenolics, at an unprecedented rate (Calchera et al. 2019).

The secondary metabolite biosynthetic pathways in filamentous fungi—including lichen-forming fungi—are typically organized into contiguous gene clusters in the genome, that is biosynthetic gene clusters (BGCs). These gene clusters contain the chemical backbone synthesis genes, such as nonribosomal peptide synthetases (NRPSs) and PKSs, tailoring enzymes, transporters and, often, transcription factors that control the expression of the clustered genes (Keller 2019). PKS catalyze repetitive condensations of an acetyl-coenzyme A (CoA) starter with malonyl-CoA units. Most of the fungal PKSs are type I PKSs, which consist of a set of  $\beta$ -ketoacyl synthase (KS), acyl transferase (AT), and acyl carrier protein (ACP) domains. The PKSs are classified into non-reducing (NR-), partially reducing (PR-), and highly reducing (HR-PKS) types depending on the extent of chemical reduction in their polyketide structure (Miao et al. 1998; Bingle et al. 1999; Nicholson et al. 2001). Based on the phylogenetic relationships and domain architecture, the PKSs are further divided into subclades (Kroken et al. 2003). Within the NR-PKS eight groups (I to VIII) with known functions have been described in fungi so far (Liu et al. 2015). The genomes of ascomycetes fungi usually contain dozens of BGCs. These are either species specific or broadly taxonomically distributed and are often very different between species (Khaldi et al. 2010; Keller 2015). The total numbers of BGCs can differ widely even between very closely related species (Lind et al. 2015, 2017). Although the number and functions of BGCs have been largely studied in non-lichenized fungal model organisms, they are poorly known in lichen-forming fungal taxa (Calchera et al. 2019).

Here, we used a genomic approach to elucidate the biosynthetic gene content and the evolution of the PKS gene that putatively is central for the biosynthesis of usnic acid. We tested whether the presence and absence of usnic acid in a lichen is caused by differential expression of genes or reflects the presence of a gene cluster in the genome of a lichen-forming fungus. We specifically focused on the following research questions: 1) do all usnic acid-containing lichen-forming fungi share a homologous gene cluster?, 2) is this gene cluster present in lichen-forming fungi that do not produce usnic acid?, 3) is the gene cluster architecture conserved across Parmeliaceae? and 4) how can the presence or absence



**Fig. 1.**—Genome completeness and BGC assessment of 46 lichen-forming fungi genomes used for this study. Bars represent the number of BGC (blue), number of non-reducing PKS (orange), and percentage of genome completeness (grey). The Y-axis represents the values of percentage and the number of BGCs.

of the putatively usnic acid producing gene cluster be explained?

## Results and Discussion

### Genome Completeness to Detect Putative Usnic Acid BGC

Metagenomic methods recovered the large parts of the lichen-fungal genomes with a genome completeness varying from 70.51% to 98.4%. The lowest values of completeness were found for *Bulbothrix sensibilis* (70.5%) and *Usnea antarctica* (70.4%), and the highest in *Alectoria sarmentosa* (96.4%) and *Evernia prunastri* (95.3%). This heterogeneity allowed us to test whether the ability to detect the presence of the putative usnic acid gene cluster was dependent on the level of genome completeness. The fact that we detected the putative usnic acid cluster in *U. antarctica* demonstrates that the detectability of the targeted gene cluster and the level of genome completeness were independent in our study. The statistics of genome assemblies of the studied species are depicted in [supplementary table S1, Supplementary Material](#) online.

### BGC Content

The species with a higher number of BGCs were *E. prunastri* with 98 BGCs followed by *Pannoparmelia angustata* with 78 BGC and *Parmotrema austrosinense* with 75 BGCs. The species with lowest number were *Canoparmelia nairobiensis* with

10 BGCs, *B. sensibilis* with 11 BGCs, and *Hypotrachyna scytodes* with 13 BGCs (fig. 1). Although the BGCs have been studied for individual species in lichen-forming fungi (see e.g. Armaleo et al. 2011; Bertrand et al. 2018b; Dal Grande et al. 2018; Calchera et al. 2019), this has not yet been compared at large scale for example family level. To our knowledge, this is the first study comparing BGC contents in species belonging to different major clades of the family Parmeliaceae (fig. 1). Within Parmeliaceae, 80 BGCs have been reported in *E. prunastri* and 51 in *Pseudevernia furfuracea* (Calchera et al. 2019); in Cladoniaceae, *Cladonia uncialis* contained 48 BGCs (Bertrand et al. 2018b); in Teloschistaceae, *Caloplaca flavorubescens* had 13 BGCs (Park et al. 2013); and in Umbilicariaceae, *Lasallia hispanica* contained 18 BGCs (Dal Grande et al. 2018). Although the highest number of BGCs has been found in some Parmeliaceae taxa (fig. 1), no correlations between BGCs and secondary metabolism or ecology were found. For example, *E. prunastri* with a high number of BGCs (98) contained three main secondary metabolites (atranorin, usnic, and evernic acids) whereas *P. furfuracea* with a relatively low number of BGCs (51) also contained a similar number of main substances (atranorin, physodic, and olive-toric acids) ([supplementary table S2, Supplementary Material](#) online). Both species are found in same ecological conditions and widespread in temperate habitats. *Parmotrema austrosinense* with a high number of BGCs (75) contained two main substances (atranorin and lecanoric acid) ([supplementary table S2, Supplementary Material](#) online) and are widely

distributed in the tropical habitats. Nonetheless, a genome-wide examination of BGC contents provides a road map to the genomic changes essential for the extensive diversity of secondary metabolites in this group of lichen-forming fungi. Parmeliaceae is one of the most diverse families both in species diversity and in containing a large number of different secondary metabolites. Usnic acid and atranorin are the most common cortical substances in the family. Usnic acid protects the photobiont from solar radiations as UV-B and act as defense against herbivores (Solhaug and Gauslaa 2012) and has been attributed to adaptive radiation in Parmeliaceae occupying different climatic regions, for example *Xanthoparmelia* in semi-arid regions (Lumbsch et al. 2008; Divakar et al. 2013). This is the most diverse genus of lichen-forming with more than 800 described species; of these over 600 species contained usnic acid.

### Phylogeny of Non-Reducing PKS

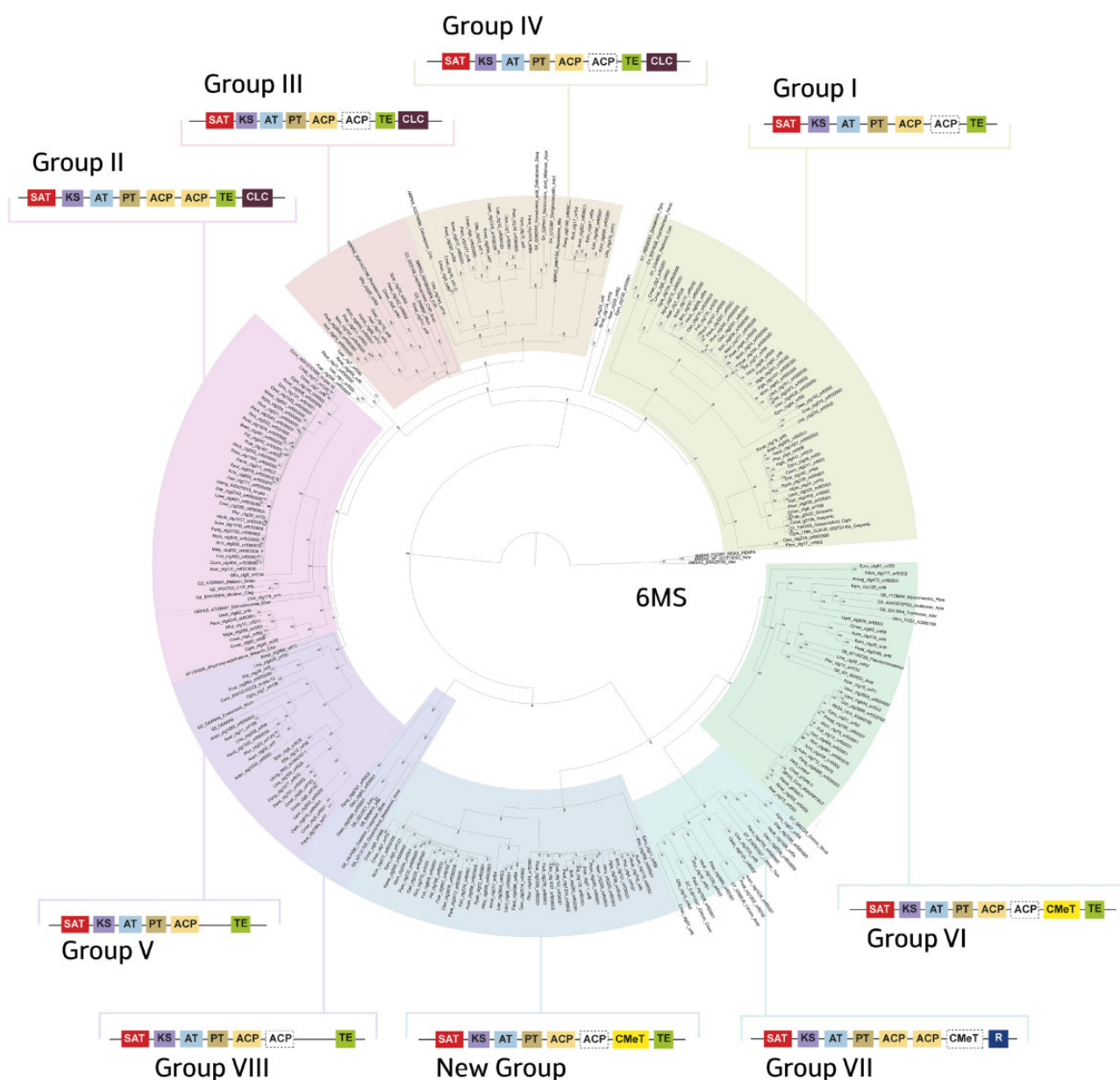
A NJ phylogeny of the KS domain and full PKS sequence revealed the evolutionary relationship of biosynthetic gene content and the origins of the putative usnic acid PKS genes in lichen-forming fungi. The KS domain and PKS full-length sequence showed highly similar tree topology (supplementary fig. S2, Supplementary Material online) and therefore only PKS phylogeny is discussed in detail. The data matrix of amino acid sequences of PKS domain consists of 356 columns (residues) and 167 taxa (of these 45 were from curated reference data). All sequences included in the analysis passed the composition  $\chi^2$ -test ( $P < 5\%$ ) and the best substitution model was LG+I+G4 according to BIC. The phylogenetic tree inferred from the PKS data matrix was rooted using polyketide 6-methylsalicylic acid sequences. In the resulted tree, two main groups were found among NR-PKSes (fig. 2). The well-supported topology is concordant with previous reported KS phylogenies (Liu et al. 2015; see Supplementary fig. S1, Supplementary Material online).

The first monophyletic cluster comprised the groups I–V of NR-PKS. The group I included PKS sequences for the biosynthesis of aromatic compound derived from orsellinic acid, such as grayanic acid in *Cladonia grayi* (Armaleo et al. 2011) or mycotoxins, such as zearalenone in *Fusarium graminearum* (Gaffoor and Trail 2006). The PKS of this group shows a domain arrangement of SAT, KS, AT, PT, one or two ACP domains and TE. Group V comprised PKS without the TE domain and is implicated in the production of different mycotoxin, such as desertorin (*Aspergillus nidulans*) or atrochryson (*Aspergillus fumigatus*) (Galagan et al. 2005; Lim et al. 2012). Groups II–III–IV formed a supported sister-relationship to group V, comprising PKS with TE/CLC domain located on the N-terminal with a similarity in the domain arrangement. Group II contained PKS involved in melanin biosynthesis as PKS1 of *Exophiala dermatitidis* (Feng et al. 2001) and it is characterized by having two ACP domain between PT and

TE; in this group, many homologous genes to melanin biosynthesis of some Pezizomycotina fungi were found in lichen-forming fungi. This could indicate that the NR-PKS group II could be responsible for the biosynthesis of melanin in lichen-forming fungi. Moreover, homology search using a known gene clusters have been applied for the identification of related gene clusters in other fungal genomes (Gardiner and Howlett 2005; Khaldi et al. 2010). However, an additional study will be necessary in order to confirm its function in lichens. Groups III and IV NR-PKS contain proteins that synthesize large polyketides chains as conidial yellow pigment Alb1 (group III) or aflatoxin/sterigmatocystin of *A. nidulans* (Yu and Leonard 1995) (group IV) (supplementary table S3, Supplementary Material online).

The second monophyletic cluster included the groups VI–VII–VIII and an undescribed group (named hereafter “New Group”) sister to VI and VII group, which does not contain PKS of known function. As eight NR-PKS groups with known functions have been described in fungi (Liu et al. 2015), our results suggest that this group with unknown function may belong to a new NR-PKS group. This result can serve as valuable entry point to search for functional gene clusters in fungi. However, we refrain to describe this group formally as this is not the main focus of our study and an additional study will be needed. Sequences located on the groups VI, VII and the New Group are largely characterized by having a methylation domain (CMet) between the ACP domain and the N-terminal domain. Group VII, which is responsible for the biosynthesis of citrinin in many fungal species (*A. nidulans*, *Coccidioides immitis*, supplementary table S3, Supplementary Material online) (Chiang et al. 2009; Gallo et al. 2013), is characterized by having a reductase (R) domain placed on the N-terminal. The domain arrangement of this probable new group is similar to group VI. Although the new group included only lichen-forming fungi, the blast analysis found homologous sequences belonging to other groups of fungi with unknown functions. The PKS putatively responsible for usnic acid biosynthesis in lichens was found in NR-PKS group VI (fig. 2). We found that all producer species contained PKS sequences homologous to the putative PKS of usnic acid of *C. uncialis* (MPAS) (Abdel-Hameed et al. 2016). This orthologous gene cluster was absent in all non-producer species.

Within group VI, a strong phylogenetic structure was observed. Three main strongly supported (>95 bootstrap) monophyletic groups were found, in accordance with the domain arrangement of the cluster. The short phylogenetic branches of usnic acid PKS gene cluster and the concordance in phylogenetic relationships between PKS phylogeny and species phylogeny (see, e.g., *Usnea*, fig. 3) (Pizarro et al. 2018) indicate that the scattered occurrence of usnic acid across the phylogenetic tree of these lichen-forming fungi may not be the result of horizontal gene transfer (Kroken et al. 2003; Bushley and Turgeon 2010; Campbell et al. 2013; Lind et al. 2017).

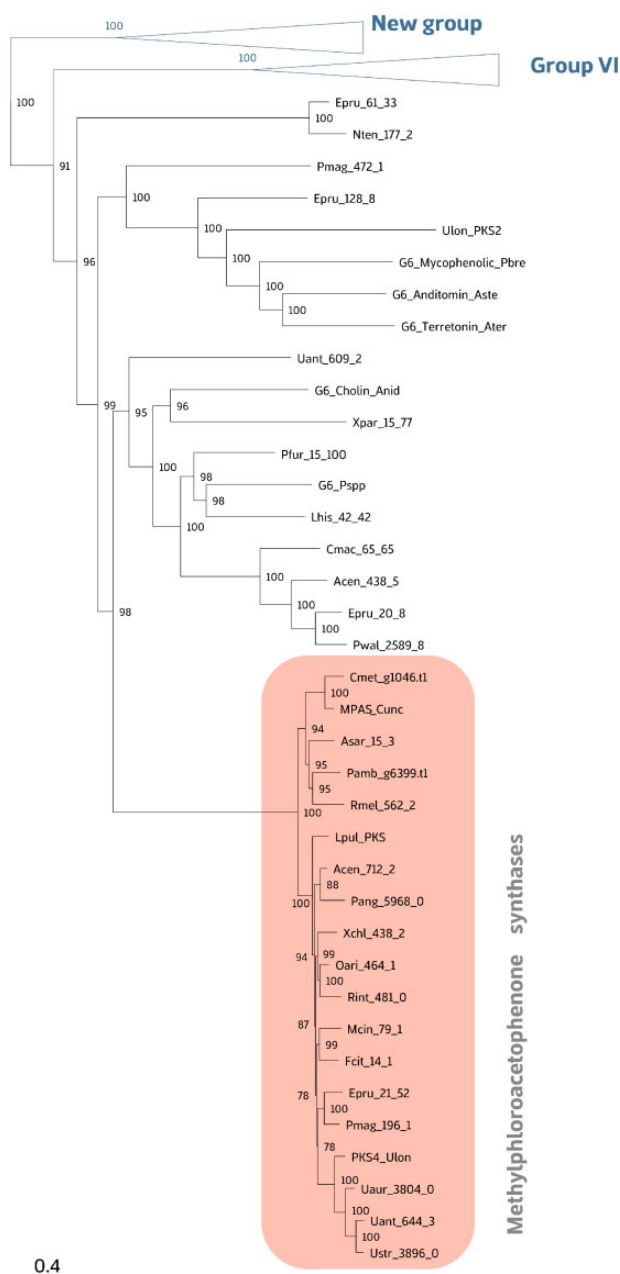


**FIG. 2.**—Gene tree of NR-PKS dataset inferred by ML analysis in IQ-Tree using 6 Methylsalicylic PKS protein sequences as outgroup. The distinct colors represent the different groups corresponding to the domain arrangement. In each group the domain arrangement of NR-PKS is highlighted with distinct colors. SAT = starter unit-ACP transacylase, KS = ketosynthase, AT = acyl transferase, PT = product template, ACP = acyl carrier protein, TE = thioesterase, TE/CLC = thioesterase/Claisen cyclase, CMET = C-methyltransferase, R = reductase.

### Gene Arrangement of Putative Usnic Acid BGC

We then compared gene cluster architecture of the putative usnic acid BGC among six species of Lecanoromycetes. All genomes of producer species shared homologous genes, that is a core NR-PKS biosynthetic gene (corresponding to *MPAS* in *C. uncialis*), a cytochrome P450 gene (*CYT450*) (corresponding to *MPAO* in *C. uncialis*) (Abdel-Hameed et al. 2016) and one or two putative transcription factors (fig. 4). The putative transcription factors (in red in fig. 4) contain

Cys6-Zn domains, as often found in regulators of the secondary metabolism in fungi (Shelest 2017). The gene arrangement of the BGC in *Usnea florida* is very similar to those of the other Parmeliaceae species, *E. prunastri* and *A. sarmentosa*. They share homologous genes that could be likely involved in the usnic acid biosynthesis like the Lacase (Tsai et al. 1999) and the putative oxidoreductase gene, which could be implicated in the oxidation of precursors molecules of usnic acid.



**FIG. 3.**—Fragment of NR-PKS tree showing the relationship of putative PKS responsible for the biosynthesis of usnic acid.

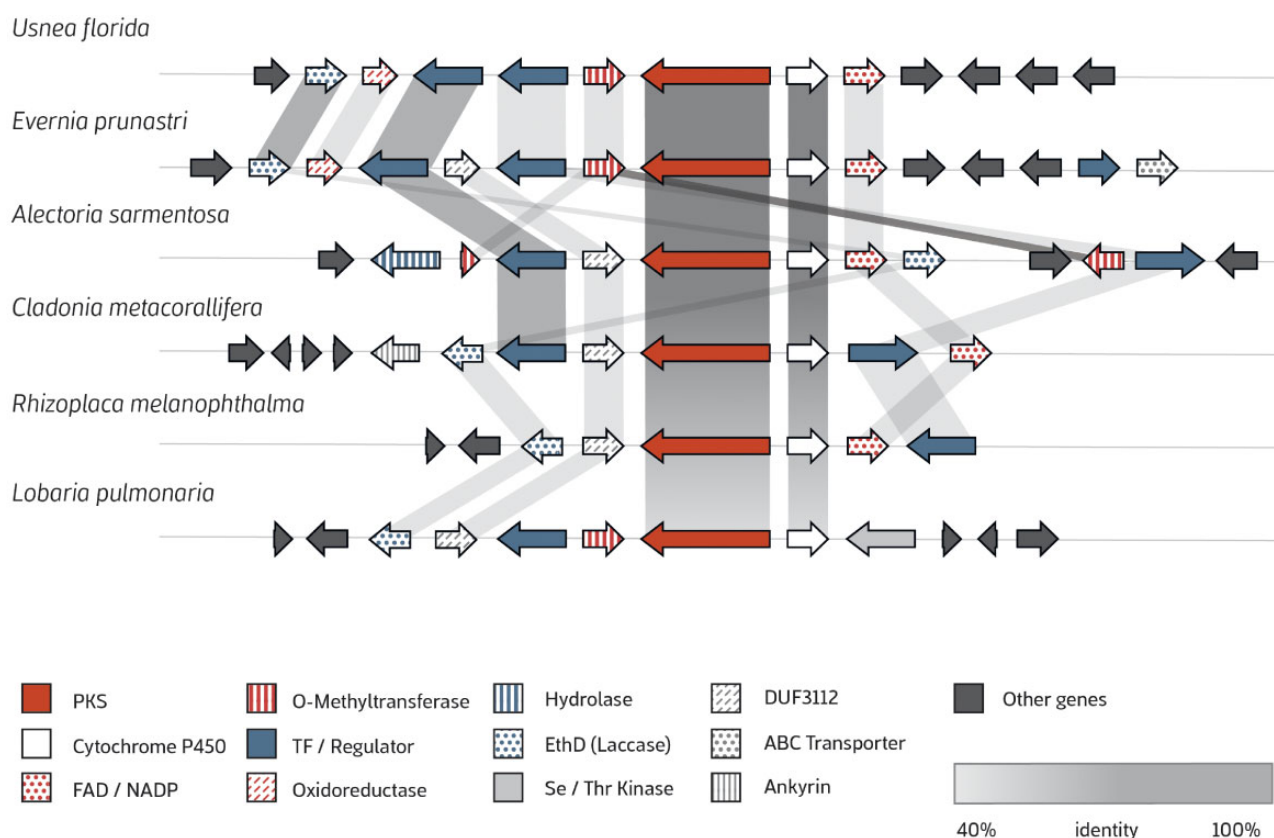
We found an O-methyl transferase gene flanking the core biosynthetic PKS genes in all species, except *Cladonia metacorrallifera* and *Rhizoplaca melanophthalma*. This O-methyl-transferase gene is common to many BGC and is homologous to the coactivator *AflJ*, implicated in the biosynthesis of aflatoxin in *Aspergillus parasiticus* (Chang 2003) and to *mdpA* in *A. nidulans* (Chiang et al. 2009). In the 5' flanking upstream region of the *CYT450*, we found a FAD/NADP-containing protein in all species included with exception of *Lobaria pulmonaria*, which contained a serin/threonin kinase.

*Evernia prunastri*, *A. sarmentosa*, *C. metacorrallifera*, and *R. melanophthalma* contained a gene DUF3112 with unknown function, which was in different locations depending on the species. For example, in *E. prunastri* it is located between the regulators, whereas in *A. sarmentosa* it was between the PKS gene and a regulator. An ABC transporter that may be implicated in the secretion of usnic acid was only found in the *E. prunastri*.

### Repeated Independent Usnic Acid BGC Losses in Non-Producers

The phylogenetic tree inferred from the concatenated dataset of 2556 BUSCO genes showed similar results as in our previous study (fig. 5; Pizarro et al. 2018). All nodes in the tree were maximally supported (100%). Usnic acid-producing species were widely spread across the Lecanorales. All families included in this study had representatives of usnic acid-producing species, such as *R. melanophthalma* (Lecanoraceae), *C. metacorrallifera* (Cladoniaceae), and 13 usnic acid-producing species representing the 6 major clades of the Parmeliaceae (see fig. 5). The orthologous putative usnic acid BGC was absent in all studied species lacking usnic acid whereas it was present in all studied usnic acid producer species (fig. 5). Interestingly, in some cases producer and non-producer species formed strongly supported sister-group relationship, for example *Cladonia macilenta* (non-producer) and *C. metacorrallifera* (producer), *Xanthoparmelia pulla* (non-producer) and *Xanthoparmelia chlorochroa* (producer). This strongly suggests that the entire usnic acid BGC was lost independently several times during the evolution of this fungal group. Similar results have been found in other groups of BGC and fungi, for example the bikaverin BGC was entirely lost in different *Botrytis* species (Campbell et al. 2013).

The discontinuous distribution of PKS genes among fungal groups has been often explained by gene loss and convergent evolution (Stayton 2015) rather than horizontal gene acquisition (see, e.g. Kroken et al. 2003; Bushley and Turgeon 2010; Campbell et al. 2012; Lind et al. 2017). Corroborating this scenario of repeated gene cluster losses, the amino acid sequences of the putative usnic acid biosynthetic genes were highly similar and showed similar tree topology as the species tree (see figs. 3 and 5). Furthermore, within the putative usnic acid PKS gene phylogeny, the branches were homogenous and short, and no long branches were detected (see fig. 3). In the 2,556-gene phylogeny, the species containing the usnic acid BGC (i.e. producer) do not form one monophyletic cluster. Instead, producers and non-producers were part of a large monophyletic group, suggesting an ancestral state of studied gene cluster in this fungal group (Lind et al. 2017). This is in accordance with the previous finding supporting an ancestral state of usnic acid production in the parmelioid lichens, the major clade within Parmeliaceae (Divakar et al. 2013). Due to limited taxon sampling, a more conclusive



**FIG. 4.**—Putative usnic acid compound cluster (UA) conservation and synteny. The UA cluster from other lichen-forming fungi resembled to the characterized usnic acid core genes PKS (MPAS) and Cytochrome P450 (MPAO) from *C. uncialis* (high percentage of identity in protein-by-protein comparisons of both genes in all species included).

ancestral state reconstruction analysis was not possible, and thus remains a hypothesis to test in future studies. On the other hand, the secondary acquisition or horizontal gene transfer may be possible but it is less probable in our case as. The evolutionary relationships among Lecanorales taxa were highly similar to those published previously (Pizarro et al. 2018) and therefore are not discussed in detail here.

## Conclusions

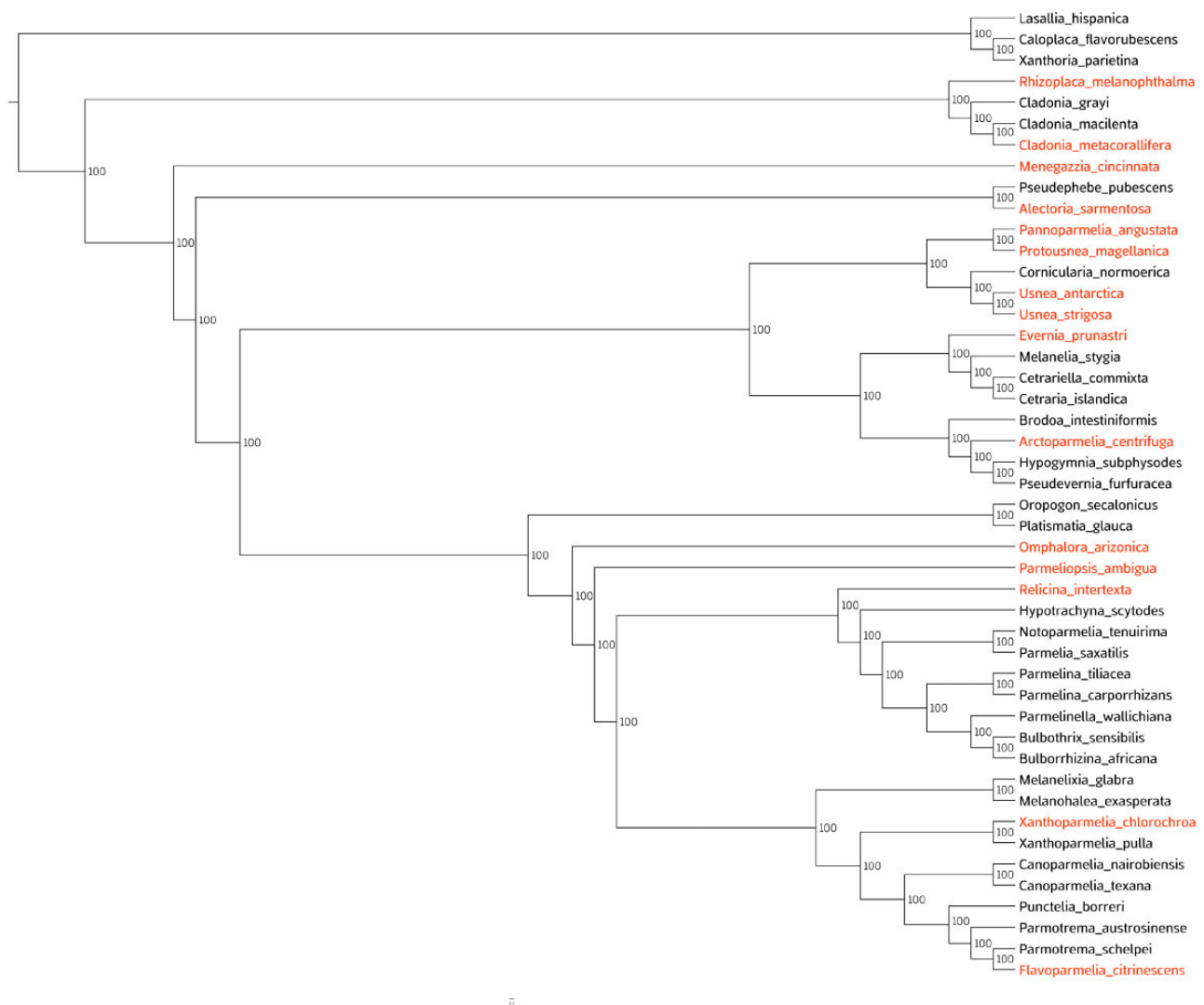
Here we present a first comprehensive study investigating biosynthetic gene contents in the genomes of the largest clade of lichen-forming fungi. A total of eight NR-PKS groups (I–VIII) with known functions and, in addition, one novel NR-PKS group with unknown function in fungi were identified. In this study, we showed that, all species lacking usnic acid production also lacked the putative BGC for usnic acid biosynthesis. To our best knowledge, this is the first study to report that the presence/absence of an entire BGC is putatively responsible for the biosynthesis of the corresponding secondary metabolite in lichen-forming fungi. The short phylogenetic branches of usnic acid PKS gene cluster, the concordance in phylogenetic relationships between PKS phylogeny and

species phylogeny and high similarity in the amino acid sequences of this BGC among all usnic acid producer species, suggest that the discontinuous distribution of usnic acid across the phylogenetic tree of these lichen-forming fungi may not be explained by horizontal gene transfer. Rather, we hypothesized that the usnic acid BGC might have been entirely lost in the non-producer species during the evolution of these fungi. Our results suggest that gene loss is an evolutionary mechanism underlying secondary metabolite diversity in lichens. Furthermore, our study sets the path for future research on the molecular detection of producer species and the metabolic engineering of lichen secondary metabolite biosynthesis. Investigating gene losses has great potential to reveal the association of BGC with natural compounds and to understand the genomic basis underlying ecological and metabolic changes.

## Materials and Methods

### Taxon Sampling

A total of 46 lichen-forming fungal species of producers and non-producers of usnic acid were included in this study. We included genomes of species belonging to different orders of



**Fig. 5.**—Phylogenetic tree from the IQ-Tree analysis based on a concatenated dataset of 2556 BUSCO genes (Pizarro et al. 2018). Numbers at the nodes represent ML bootstrap support values based on 1,000 bootstrap pseudoreplicates. Species in orange and black indicate the usnic acid producer species and non-producer species, respectively.

Lecanoromycetes including Teloschistales (*Xanthoria parietina* and *Gyalolechia flavorubescens*), Umbilicariales (*Umbilicaria hispanica*; Dal Grande et al. 2018), and Lecanorales (family Cladoniaceae: *C. grayi*, *C. macilenta*, *C. metacorallifera*; Lecanoraceae: *R. melanophthalma*; Parmeliaceae: 39 species representing six of its seven major clades; see [supplementary table S2, Supplementary Material](#) online).

### Genome Sequencing and Assembly

Total genomic DNA of 38 specimens of Parmeliaceae was extracted from apothecia or thalli using the Quick-DNA™ Fungal/Bacterial Miniprep Kit (Zymo Research, Irvine, CA) following the manufacturers' instructions. Genome sequencing, assembly, and taxonomic assignment were carried out as described in Pizarro et al. (2018).

### Genome Completeness Assessment and Phylogenomic Analysis

BUSCO analysis was used to evaluate the genome completeness (Simão et al. 2015). Here, we used the Pezizomycotina data set containing 3,156 single-copy genes to assess the completeness of the 46 genomes of this study. To infer the phylogenetic relationships between producers and non-producers of usnic acid, every genome was explored following a methodology as described before (Pizarro et al. 2018). The complete BUSCO single-copy genes predicted in each genome were extracted and aligned using MAFFT L-INS-i (Standley 2013). A supermatrix was created by concatenating all alignments using FASconCAT.pl (Kück and Longo 2014). Evolutionary relationships were inferred from this subset using maximum likelihood (ML) analysis implemented in IQTree



v1.5.5 (Nguyen et al. 2015; Kalyaanamoorthy et al. 2017) with standard model selection and 1,000 bootstrap replicates. The resulting tree was visualized using FigTree 1.3.1 (Rambaut 2009).

### BGC Prediction and Usnic Acid Genes Cluster Identification

In order to identify the putative usnic acid PKS cluster, gene prediction and annotation were performed in every genome using AUGUSTUS (Stanke et al. 2004) and MAKER2 (Holt and Yandell 2011). In parallel, we used every genome and its respective gene prediction as input for antiSMASH 4.0 (Blin et al. 2017) for secondary metabolites gene cluster prediction. Afterwards, MPAS of *C. uncialis* (A0A0R8YWJ7) and the orthologous gene PKS4 of *Usnea longissima* (AGI60156) sequences were downloaded from Uniprot database and HMM profile was created using MAFFT and HMMER hmmbuild (Mistry et al. 2013). These profiles and hmmsearch tool were used to search the most similar sequences present in every protein data set predicted from each genome. Only non-reducing PKSES were considered for further analyses.

### Full-Length NR-PKS and KS Domain Phylogenetic Analysis

We used Hidden Markov Model profiles for the identification of a specific type of domain. The recovered NR-PKS proteins were scanned using HMMER against the Pfam (Mitchell et al. 2016) protein domain collection. Then using a bioinformatic approach with BEDtool (getfasta) (Quinlan and Hall 2010), the different KS domains from PKS were identified in the corresponding amino acid sequences. An identical procedure was carried out over an additional curated NR-PKS data set (see [supplementary table S3, Supplementary Material](#) online) of characterized fungal NR-PKSs. Multiple alignments were performed using MAFFT L-INS-I over the amino acid sequence of full-length PKS dataset and KS domain dataset. Evolutionary relationships were inferred using ML analysis implemented in IQ-Tree v1.5.5 with standard model selection. For each analysis, a Neighbour Joining and 1,000 bootstrap replicates were calculated using fast bootstrapping option. The resulting phylogenetic trees were rooted with 6 Methylsalicylic PKS (6MS) sequences and drawn using FigTree v1.3.

### Synteny Comparison of Usnic Acid BGC

Genomes that contain the full putative gene clusters of usnic acid biosynthesis were selected for synteny comparison. Since some of the usnic acid producer species had low gene cluster completeness, only *C. macilenta*, *A. sarmentosa*, *E. prunastri*, and *R. melanophthalma* were considered for the analysis. In addition, *U. florida* and *L. pulmonaria*, downloaded from JGI, were also included in the analysis. Using Orthofinder (Emms and Kelly 2015), we initially compared the homology among different genomes and subsequently drew their relationship.

## Supplementary Material

Supplementary material is available at *Genome Biology and Evolution* online (<http://www.gbe.oxfordjournals.org/>).

## Acknowledgements

We thank Daniele Armaleo for granting permission to use the draft genome of *C. grayi*, via the JGI portal. We thankfully acknowledge the valuable feedback from two anonymous reviewers and the editor that greatly improved this study. This study was supported by the Spanish Ministerio de Ciencia e Innovacion (CGL2013-42498-P and PID2019-105312GB-I00), the Santander-Universidad Complutense de Madrid (PR75/18-21605, PR87/19-22637, and G/6400100/3000), the LOEWE-Centre TBG funded by the Hessen State Ministry of Higher Education, Research and the Arts (HMWK) and the Grainger Bioinformatics Center at the Field Museum (Chicago, IL).

**Data and materials availability:** All figures and tables generated during the study are included in the paper and/or the [supplementary materials, Supplementary Material](#) online; and all data used have been deposited in FigShare (10.6084/m9.figshare.12942110). Additional data related to this paper may be requested from the authors.

## References

- Abdel-Hameed M, Bertrand RL, Piercey-Normore MD, Sorensen JL. 2016. Putative identification of the usnic acid biosynthetic gene cluster by de novo whole-genome sequencing of a lichen-forming fungus. *Fungal Biol.* 120:306–316.
- Armaleo D, Sun X, Culbertson C. 2011. Insights from the first putative biosynthetic gene cluster for a lichen depside and depsidone. *Mycologia* 103(4):741–754.
- Bertrand RL, Abdel-Hameed M, Sorensen JL. 2018a. Lichen biosynthetic gene clusters part II: homology mapping suggests a functional diversity. *J Nat Prod* 81(4):723–731.
- Bertrand RL, Abdel-Hameed M, Sorensen JL. 2018b. Lichen biosynthetic gene clusters. Part I. Genome sequencing reveals a rich biosynthetic potential. *J Nat Prod* 81(4):732–748.
- Bingle LEH, Simpson TJ, Lazarus CM. 1999. Ketosynthase domain probes identify two subclasses of fungal polyketide synthase genes. *Fungal Genet Biol* 26(3):209–223.
- Blin K, et al. 2017. AntiSMASH 4.0—improvements in chemistry prediction and gene cluster boundary identification. *Nucleic Acids Res* 45(1):36–41.
- Bushley KE, Turgeon BG. 2010. Phylogenomics reveals subfamilies of fungal nonribosomal peptide synthetases and their evolutionary relationships. *BMC Evol Biol.* 10(1):26.
- Calchera A, Dal Grande F, Bode HB, Schmitt I. 2019. Biosynthetic gene content of the ‘perfume lichens’ *Evernia prunastri* and *Pseudevernia furfuracea*. *Molecules* 24(1):203.
- Calcott MJ, Ackerley DF, Knight A, Keyzers RA, Owen JG. 2018. Secondary metabolism in the lichen symbiosis. *Chem Soc Rev.* 47(5):1730–1760.
- Campbell MA, Rokas A, Slot JC. 2012. Horizontal transfer and death of a fungal secondary metabolic gene cluster. *Genome Biol Evol.* 4(3):289–293.

- Campbell MA, Staats M, van Kan JAL, Rokas A, Slot JC. 2013. Repeated loss of an anciently horizontally transferred gene cluster in *Botrytis*. *Mycologia* 105(5):1126–1134.
- Chang P-K. 2003. The *Aspergillus parasiticus* protein AFLJ interacts with the aflatoxin pathway-specific regulator AFLR. *Mol Genet Genomics* 268(6):711–719.
- Chiang YM, et al. 2009. A gene cluster containing two fungal polyketide synthases encodes the biosynthetic pathway for a polyketide, asperfuluranone, in *Aspergillus nidulans*. *J Am Chem Soc* 131(8):2965–2970.
- Crespo A, Divakar PK, Lumbsch HT. 2014. Fungi: Hyperdiversity closer to animals than to plants. *Tree Life*. p. 168–181.
- Crespo A, et al. 2010. Phylogenetic generic classification of parmelioid lichens (Parmeliaceae, Ascomycota) based on molecular, morphological and chemical evidence. *Taxon* 59(6):1735–1753.
- Dal Grande F, et al. 2018. The draft genome of the lichen-forming fungus *Lasallia hispanica* (Frey) Sancho & A. Crespo. *Lichenologist* 50(3):329–340.
- Divakar PK, Kauff F, Crespo A, Leavitt SD, Lumbsch HT. 2013. Understanding phenotypical character evolution in parmelioid lichenized fungi (Parmeliaceae, ascomycota). *PLoS One* 8(11):e83115.
- Elix JA. 1993. Progress in the generic delimitation of *parmelia sensu lato* lichens (Ascomycotina: Parmeliaceae) and a synoptic key to the parmeliaceae. *Bryologist* 96(3):359–383.
- Elix JA, Whitton AA, Sargent MV. 1984. Recent progress in the chemistry of lichen substances. In: Herz W, Grisebach H, Kirby GW, editors. *Fortschritte der Chemie organischer Naturstoffe/Progress in the Chemistry of Organic Natural Products*. Vienna: Springer. p. 103–234.
- Emms DM, Kelly S. 2015. OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biol.* 16(1):157.
- Feng B, et al. 2001. Molecular cloning and characterization of WdPKS1, a gene involved in dihydroxynaphthalene melanin biosynthesis and virulence in *Wangiella* (*Exophiala*) *dermatitidis*. *Infect Immun* 69(3):1781–1794.
- Gaffoor I, Trail F. 2006. Characterization of two polyketide synthase genes involved in zearalenone biosynthesis in *Gibberella zeae*. *Appl Environ Microbiol* 72(3):1793–1799.
- Galagan JE, et al. 2005. Sequencing of *Aspergillus nidulans* and comparative analysis with *A. fumigatus* and *A. oryzae*. *Nature* 438(7071):1105–1115.
- Gallo A, Ferrara M, Perrone G. 2013. Phylogenetic study of polyketide synthases and nonribosomal peptide synthetases involved in the biosynthesis of mycotoxins. *Toxins (Basel)*. 5(4):717–742.
- Gardiner DM, Howlett BJ. 2005. Bioinformatic and expression analysis of the putative gliotoxin biosynthetic gene cluster of *Aspergillus fumigatus*. *FEMS Microbiol Lett* 248(2):241–248.
- Hawksworth DL. 1976. Lichen Chemotaxonomy. In: Brown DH, Hawksworth DL, Bailey RH, editors. *Lichenology: progress and problems*. London: Academic Press. p. 139–184.
- Holt C, Yandell M. 2011. MAKER2: an annotation pipeline and genome-database management tool for second-generation genome projects. *BMC Bioinformatics*. 12(1):491.
- Huneck S, Yoshimura I. 1996. Identification of lichen substances. Berlin, Heidelberg: Springer.
- Kalyaanamoorthy S, Minh BQ, Wong TKF, Von Haeseler A, Jermini LS. 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods*. 14(6):587–589.
- Keller NP. 2015. Translating biosynthetic gene clusters into fungal armor and weaponry. *Nat Chem Biol*. 11(9):671–677.
- Keller NP. 2019. Fungal secondary metabolism: regulation, function and drug discovery. *Nat Rev Microbiol*. 17(3):167–180.
- Khalidi N, et al. 2010. SMURF: genomic mapping of fungal secondary metabolite clusters. *Fungal Genet Biol* 47(9):736–741.
- Kroken S, Glass NL, Taylor JW, Yoder OC, Turgeon BG. 2003. Phylogenomic analysis of type I polyketide synthase genes in pathogenic and saprobic ascomycetes. *Proc Natl Acad Sci U S A* 100(26):15670–15675.
- Kück P, Longo GC. 2014. FASconCAT-G: extensive functions for multiple sequence alignment preparations concerning phylogenetic studies. *Front Zool*. 11(1):81.
- Lim FY, et al. 2012. Genome-based cluster deletion reveals an endocrocin biosynthetic pathway in *Aspergillus fumigatus*. *Appl Environ Microbiol* 78(12):4117–4125.
- Lind AL, et al. 2015. Examining the evolution of the regulatory circuit controlling secondary metabolism and development in the fungal genus *Aspergillus*. *PLoS Genet*. 11(3):e1005096.
- Lind AL, et al. 2017. Drivers of genetic diversity in secondary metabolic gene clusters within a fungal species. *PLoS Biol*. 15(11):e2003583.
- Liu L, et al. 2015. Bioinformatical analysis of the sequences, structures and functions of fungal polyketide synthase product template domains. *Sci Rep*. 5(1):10463.
- Lumbsch HT. 1998. The taxonomic use of metabolic data in lichen-forming fungi. In: Frisvad JC, Bridge PD, Arora DK, editors. *Chemical fungal taxonomy*. New York: CRC Press. p. 345–387.
- Lumbsch HT, Hipp AL, Divakar PK, Blanco O, Crespo A. 2008. Accelerated evolutionary rates in tropical and oceanic parmelioid lichens (Ascomycota). *BMC Evol Biol*. 8(1):257.
- Miao VPW, LeGal-Coeffé MF, Brown DW, Davies JE. 1998. Polyketide synthase genes in lichens. In: Sixth International Mycological Congress (IMC6). Abstracts. Jerusalem. p. 84.
- Mistry J, Finn RD, Eddy SR, Bateman A, Punta M. 2013. Challenges in homology search: HMMER3 and convergent evolution of coiled-coil regions. *Nucleic Acids Res* 41(12):e121.
- Mitchell AL, et al. 2016. The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res*. 44(D1):D279–D285.
- Nash T. 2008. Lichen biology. London: Cambridge University Press.
- Nguyen LT, 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(1):268–274.
- Nicholson TP, et al. 2001. Design and utility of oligonucleotide gene probes for fungal polyketide synthases. *Chem Biol* 8(2):157–178.
- Park S-Y, et al. 2013. Draft genome sequence of lichen-forming fungus *Caloplaca flavorubescens* strain KoLRI002931. *Genome Announc* 1(4):e00678–13.
- Pizarro D, et al. 2018. Phylogenomic analysis of 2556 single-copy protein-coding genes resolves most evolutionary relationships for the major clades in the most diverse group of lichen-forming fungi. *Fungal Divers*. 92(1):31–41.
- Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26(6):841–842.
- Rambaut A. 2009. FigTree v1.3.1: Tree figure drawing tool. Website <http://tree.bio.ed.ac.uk/software/figtree/>. Available from: [https://scholar.google.com/scholar?hl=en&as\\_sdt=0,25&q=rambaut+figtree#4](https://scholar.google.com/scholar?hl=en&as_sdt=0,25&q=rambaut+figtree#4). Accessed 2017.
- Shelest E. 2017. Transcription factors in fungi: TFome dynamics, three major families, and dual-specificity TFs. *Front Genet* 8:53.
- Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31(19):3210–3212.
- Solhaug KA, Gauslaa Y. 2012. Secondary lichen compounds as protection against excess solar radiation and herbivores. In: Lüttge U, Beyschlag W, Büdel B, Francis D. editors. *Progress in Botany* 73. Berlin: Springer.
- Stanke M, Steinkamp R, Waack S, Morgenstern B. 2004. AUGUSTUS: a web server for gene finding in eukaryotes. *Nucleic Acids Res* 32(Web Server issue):W309–W312.

- Stayton CT. 2015. The definition, recognition, and interpretation of convergent evolution, and two new measures for quantifying and assessing the significance of convergence. *Evolution* (N Y). 69(8):2140–2153.
- Stocker-Wörgötter E. 2008. Metabolic diversity of lichen-forming ascomycetous fungi: culturing, polyketide and shikimate metabolite production, and PKS genes. *Nat Prod Rep* 25(1):185–200.
- Thell A, et al. 2012. A review of the lichen family Parmeliaceae—history, phylogeny and current taxonomy. *Nord J Bot* 30(6):641–664.
- Tsai HF, Wheeler MH, Chang YC, Kwon-Chung KJ. 1999. A developmentally regulated gene cluster involved in conidial pigmentbiosynthesis in *Aspergillus fumigatus*. *J Bacteriol.* 181(20):6469–6477.
- Yu JH, Leonard TJ. 1995. Sterigmatocystin biosynthesis in *Aspergillus nidulans* requires a novel type I polyketide synthase. *J Bacteriol* 177(16):4792–4800.

**Associate editor:** Stajich Jason