# Mycorrhiza-induced mycocypins of Laccaria bicolor are potent protease inhibitors with nematotoxic and collembola antifeedant activity

Jonathan Plett, Jerica Sabotič, Eva Vogt, Fridtjof Snijders, Annegret Kohler, Uffe Nielsen, Markus Künzler, Francis Martin, Claire Veneault-Fourrey

Received: 11 June 2020 | Accepted: 21 June 2022

DOI: 10.1111/1462-2920.16115

RESEARCH ARTICLE

environmental microbiology sfam

# Mycorrhiza-induced mycocypins of *Laccaria bicolor* are potent protease inhibitors with nematotoxic and collembola antifeedant activity

**Jonathan M. Plett**[1,2] | **Jerica Sabotič**[3] | **Eva Vogt**[4] | **Fridtjof Snijders**[2] | **Annegret Kohler**[1] | **Uffe N. Nielsen**[2] | **Markus Künzler**[4] | **Francis Martin**[1] | **Claire Veneault-Fourrey**[1]

[1]Université de Lorraine, INRAE, UMR 1136 Interactions Arbres-Microorganismes, Centre INRAE Grand Est-Nancy, Champenoux, France

[2]Hawkesbury Institute for the Environment, Western Sydney University, Richmond, New South Wales, Australia

[3]Department of Biotechnology, Jožef Stefan Institute, Ljubljana, Slovenia

[4]Department of Biology, Institute of Microbiology, ETH Zürich, Zürich, Switzerland

**Correspondence**
Claire Veneault-Fourrey, Université de Lorraine INRAE, UMR 1136 Interactions Arbres-Microorganismes, Centre INRAE Grand Est-Nancy, 54280 Champenoux, France.
Email: claire.veneault-fourrey@inrae.fr

**Funding information**
Centre of Excellence in Plant Energy Biology, Australian Research Council, Grant/Award Number: DE150100408; ENERGYPOPLAR, Grant/Award Number: FP7-211917; Laboratory of Excellence ARBRE, Grant/Award Number: ANR-11-LABX-0002-01; Plant-Microbes Interfaces, US DOE, Grant/Award Number: DE-AC05-00OR22725; Schweizerischer Nationalfonds zur Förderung der Wissenschaftlichen Forschung, Grant/Award Number: 31003A_173097; Slovenian Research Agency, Grant/Award Numbers: P4-0127, P4-0432

## Abstract

Fungivory of mycorrhizal hyphae has a significant impact on fungal fitness and, by extension, on nutrient transfer between fungi and host plants in natural ecosystems. Mycorrhizal fungi have therefore evolved an arsenal of chemical compounds that are hypothesized to protect the hyphal tissues from being eaten, such as the protease inhibitors mycocypins. The genome of the ectomycorrhizal fungus *Laccaria bicolor* has an unusually high number of mycocypin-encoding genes. We have characterized the evolution of this class of proteins, identified those induced by symbiosis with a host plant and characterized the biochemical properties of two upregulated *L. bicolor* mycocypins. More than half of *L. bicolor* mycocypin-encoding genes are differentially expressed during symbiosis or fruiting body formation. We show that two *L. bicolor* mycocypins that are strongly induced during symbiosis are cysteine protease inhibitors and exhibit similar but distinct localization in fungal tissues at different developmental stages and during interaction with a host plant. Moreover, we show that these *L. bicolor* mycocypins have toxic and feeding deterrent effect on nematodes and collembolans, respectively. Therefore, *L. bicolor* mycocypins may be part of a mechanism by which this species deters grazing by different members of the soil food web.

## INTRODUCTION

Tree roots establish mutualistic interactions with ectomycorrhizal (ECM) fungi that can aid the host tree growth and health by supplying nutrients, particularly in nutrient-limited forest soils. During this symbiotic interaction, ECM hyphae aggregate on the surface of roots to form a sheath-like mantle and then grow into the root apoplastic space to form a structure known as the Hartig net, where the exchange of nutrients occurs (Martin et al., 2016; Pellegrin et al., 2019). The colonization process significantly alters the physiology and structure of roots, with tree hosts typically displaying an immune response entailing the induction of secondary metabolites, localized hormone signalling and the secretion of a range of proteases (Basso et al., 2020). How ECM

fungi respond to these physiological changes and avoid the host immune system is still an area of actively developing research. While ECM fungi have been found to secrete effector-like proteins that manipulate host secondary metabolism (Plett et al., 2020) and host hormone signalling (Daguerre et al., 2020; Plett et al., 2014), there are currently no known defences against host proteases characterized in the effectome of ECM fungi. In addition to the structures formed at the host root, ECM fungi produce extensive hyphae into the surrounding soil for nutrient acquisition (e.g. nitrogen and phosphorus) and to form fruiting bodies during sexual propagation (Pickles & Simard, 2017; Simard et al., 2012). The establishment and maintenance of these hyphae is necessary for the maintenance of symbiotic interactions but can be impeded by hyphal predation by soil fauna (Giannakis & Sanders, 1990). Within this context, the mechanisms induced in ECM fungi during colonization of a host dealing with predation also need to be studied.

Following the genome sequencing of the model ECM fungus *L. bicolor* (Martin et al., 2008), a series of transciptomic experiments have identified hundreds of host-induced fungal genes (Kohler et al., 2015; Martin et al., 2008; Plett et al., 2015). The majority of these genes have no known function or have a function specifically attributed to supporting the establishment of mutualistic symbiosis (e.g. nutrient transport and metabolism). Among the genes induced by host colonization identified in the first transcriptomic studies were several genes encoding mycocypin-like proteins (Martin et al., 2008). Characterized mycocypins from the saprotrophic fungi *Clitocybe nebularis* (clitocypin) and *Macrolepiota procera* (macrocypin) act against a wide variety of cysteine proteases such as papain, cathepsin L, legumain and bromelain (Brzin, 2000; Renko et al., 2010; Sabotic et al., 2006). Several plant cysteine proteases of the papain and legumain families are involved in plant defence and are inhibited during colonization by fungal pathogens (Misas Villamil et al., 2019; Passarge et al., 2021; Schulze Hüynck et al., 2019; Van der Linde et al., 2012). It is hence likely that ECM fungi use the induced mycocypins to inactivate defensive plant proteases during colonization of their host. Such a function would necessitate a secretion of these protease inhibitors.

Alternatively and also based on their role in saprotrophic fungi, the mycocypin-like proteins of ECM fungi may inhibit the proteases of predators and parasites of ECM hyphae. As soil-borne hyphae and fruiting bodies represent a major reservoir of nutrients, they are targeted by a number of predatory nematodes, insects and micro-arthropods (Asplund et al., 2013; Bokhorst et al., 2012; Bokhorst & Wardle, 2014; Klironomos & Kendrick, 1996). In order to defend against this predation, fungi have evolved a set of defensive molecules to deter or kill grazers, including secondary metabolites, lectins, biotin-binding proteins, ribotoxins and protease inhibitors (Künzler, 2018; Rohlfs, 2015). Many of these defensive molecules are not secreted but accumulate in the fungal hyphae and are only released upon attack of the fungivore. Mycocypins, members of the I48 and I85 protease inhibitor families of the MEROPS database, negatively affect foraging herbivores. They counteract the activity of proteases in the predator's gut, thereby reducing the digestibility of ingested material (Goulet et al., 2008; Michaud et al., 1995). This process gradually starves the predator or encourages its relocation towards a more acceptable food source (Šmid et al., 2013; Zhao et al., 1996). While these defensive molecules directly benefit all fungi, for those fungi that enter into symbiosis with a host plant they may also indirectly benefit its host. Namely, their presence in the mantle surrounding host roots may alter fungivore feeding, as has been found for other fungal toxins (Kaneda & Kaneko, 2004; Wohlschlager et al., 2014). It is still unknown whether mycocypins of *L. bicolor* function as feeding deterrents to soil fauna.

Given the induction of mycocypins during symbiosis with the host plant, and the different roles that they may play in ECM fungi to support symbiosis, we sought to improve our understanding of this protease inhibitor family in the mutualistic ECM fungus *L. bicolor*. Their representation and inheritance throughout the fungal tree of life were determined using currently available genomes. Two recently published transcriptomic data sets were used to identify which of this manually curated list of mycocypin-like encoding genes were induced during symbiosis. Of these, two *L. bicolor* mycocypins upregulated during symbiotic establishment and/or in fruiting bodies were chosen for further characterization. As these two characterized *L. bicolor* mycocypin-like genes were non-secreted, we hypothesized that these mycorrhiza-induced genes may have a role as feeding deterrents towards soil fauna as opposed to plant defence modification.

## EXPERIMENTAL PROCEDURES

### BLAST search for mycocypin-like genes and phylogeny

In order to find the mycocypin-like gene sequences within the fungal tree of life, a BLAST search and functional domain search approach was taken. Using the Joint Genome Institute fungal website (http://genome.jgi.doe.gov/programs/fungi/index.jsf; accessed June 2019) with 1379 fungal genomes publicly available, we first searched for functional annotations related to mycocypins (i.e. MEROPS:I48; IPR019508; PF10467) and performed tBLASTn using the reference mycocypin sequences from *Clitocybe* (accession number AAZ78483), *S. stellatus* (protein ID 267444),

*Rhizoctonia solani* (protein ID 1087) and *L. bicolor* (protein IDs 311135 and 293818) to find significantly similar gene sequences. A cut-off of $e^{-10}$ was used. From this search, the total number of mycocypin-like genes and the number of genomes are reported in Figure 1 (inset). Mycocypin-like sequences from 14 published fungal genomes (Table S1) were used in the creation of a phylogenetic tree. Predicted protein sequences were analysed using MEGA11 (Tamura et al., 2021) with the MUSCLE 3.8.31 alignment program using the default settings (Dereeper et al., 2008). The evolutionary history was inferred by using the Phylogeny using Maximum Likelihood method (PhyML) using the WAG substitution model. The tree was rooted using a mid-point rooting method and was visualized using FigTree, with branch support values of >70% replicates annotated in the figure.

## Gene expression analysis

Mycocypin gene expression data were extracted from two previously published experiments (Ruytinx et al., 2021). Briefly, for the first experiment, early

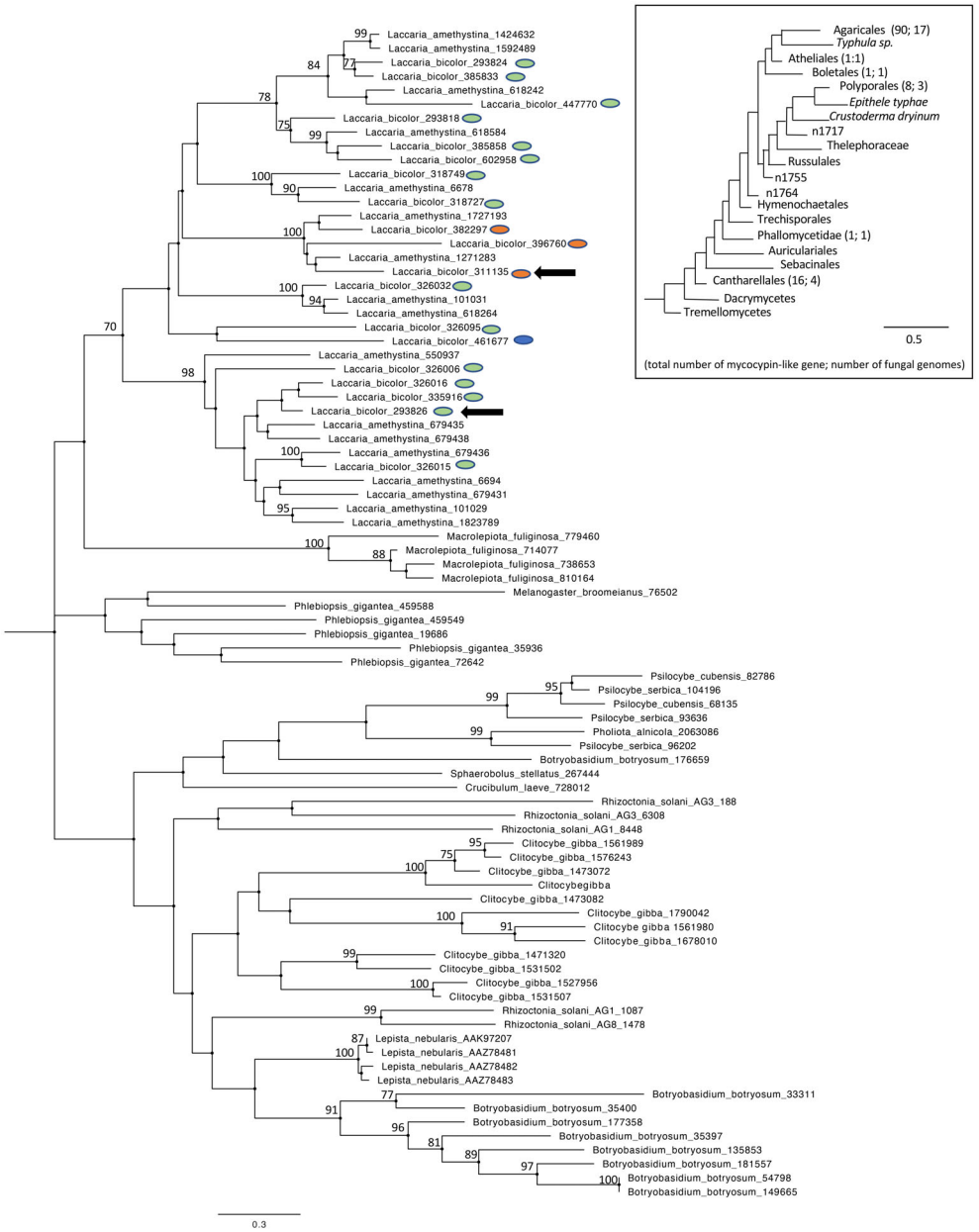


**FIGURE 1** Phylogenetic relationship using maximum likelihood method of all mycocypin-like sequences obtained. Each mycocypin is identified by a JGI number or by protein ID. Within the *L. bicolor* mycocypin-like proteins, the linkage group within which they are found is indicated by a coloured ellipse (blue—LG1; Green—LG8; Orange—Scaffold 17). Black arrows indicate the two mycocypins further characterized in this study. Inset: Total number of mycocypin-like genes and number of genomes within clades of Agaricomycotina

(primordia), medium (non-fully developed cap) and late (fully developed cap) stages of *L. bicolor* fruiting bodies were harvested in a nursery under Douglas fir seedlings inoculated with *L. bicolor* S238N-93.12 mycelium (=*L. bicolor* S238N strain frozen in liquid nitrogen in 1993 as back-up). Basidiocarps were flash-frozen in liquid nitrogen before RNA extraction. Total RNAs of basidiocarp stipes and caps were extracted using a cetyltrimethylammonium bromide–based protocol including a LiCl precipitation. For the second experiment, total RNA from mycelium and mycorrhizal root tips [harvested after 14 d of in vitro contact between *Populus tremula* × *Populus alba* and *L. bicolor* S238N (established as per Felten et al., 2009)] was extracted using the RNeasy Plant Mini Kit (Qiagen, Hilden, Germany).

RNA concentration and integrity were assessed using the Biorad Experion and RNA StdSens kit and chips. Three replicates from each stage and tissue were used for RNA-seq sequencing. The preparation of libraries from total RNA (TruSeq Stranded mRNA Kit Illumina), 2 × 100 bp Illumina HiSeq2000 sequencing for the basidiocarp samples and 2 × 150 bp Illumina HiSeq3000 sequencing for all other samples were performed on the GET platform (Génopole Toulouse Midi-Pyrénées, Auzeville, France) following their standard protocol. Raw reads were trimmed for low quality (quality score, 0.05). Illumina adapters and sequences shorter than 15 nucleotides were removed and remaining reads were aligned to the *L. bicolor* reference transcripts (https://mycocosm.jgi.doe.gov/Lacbi2/Lacbi2.home.html) using CLC Genomics Workbench 7.5.1 and 8 (Qiagen). Reads per kilobase and per Million (RPKM) and mean RPKM from replicates were calculated from unique mapped reads (Table S2). The complete data sets were submitted to NCBI-GEO as GSE190444 and GSE190443.

## Microscopy

Indirect immuno-localization of two mycocypins was performed as described by Plett et al. (2011). Anti-293826 and anti-311135 antibodies were created against the VHGKESAYTISLHTN and KDDSVVPKGP IDTSR regions, respectively, of these proteins (Eurogentec, Seraing, Belgium). Immunolocalization of the mycocypin genes was performed on transverse sections free-living mycelium (FLM) of *L. bicolor* S238N, ECM root tips and the stipe and gills of a mature fruiting body with fully developed stipe and open cap (fruiting body harvested in a nursery under Douglas fir seedlings inoculated with *L. bicolor* S238N-93.12 mycelium) as described in Martin et al. (2008) with a 1:1000 dilution of purified anti-mycocypin rabbit antibodies, followed by incubation with a 1:100 dilution of goat anti-rabbit IgG-Alexa Fluor 488 conjugate (Molecular Probes, Invitrogen™, Carlsbad, CA, USA). As a control, immunolocalization was also performed with only the secondary antibody and no detectable fluorescence was found [Figure S2(B)]. Sections were mounted in propidium iodide and viewed by a Bio-Rad Radiance 2100 AGR3Q-BLD Rainbow confocal microscope with laser levels and emission parameters adjusted to avoid auto-fluorescent signal.

## Secretion analysis

Using cDNA libraries from ECM tissues, full-length mycocypin 293826 and 311135 coding sequences were cloned (lacking the stop codon) into pSUC-GW as described previously (Plett et al., 2011). These constructs were transformed into the *Saccharomyces cerevisiae* strain YTK12. They were plated on selective SD-W medium and cultured for 4 days at 30°C. After this period, positive colonies were selected using colony PCR and replica-plated on SD-W and on YPR-A medium (10 g $L^{-1}$ yeast extract, 20 g $L^{-1}$ peptone, 20 g $L^{-1}$ agar amended with 2 g $L^{-1}$ sucrose and 60 μg $L^{-1}$ antimycin A, pH 6.5). If cultures grew on the latter medium, then the protein of interest contained a secretion motif. Cells were cultured on these media for 7 days at 30°C after which growth was scored. Four separate colonies per construct were tested for their growth on these media.

## Production and purification of recombinant protein

cDNA sequences of 293826 and 311135 *L. bicolor* genes were optimized (Genecust; Luxembourg) for *E. coli* production and cloned into pET28a between *Nde*I and *Sal*I restriction sites. The resulting plasmid, or an empty vector control, was used to transform the Rosetta2 (DE3) pLysS strain of *E. coli* (Novagen). The expression of the recombinant proteins or empty vector control was performed at 37°C in lysogeny broth medium supplemented with ethanol (5 ml $L^{-1}$) until cell culture reached an $OD_{600}$ of 0.7. Recombinant protein expression was induced by the addition of 0.1 mM, and the cells were grown for a further 24 h at 20°C. Cells were then harvested by centrifugation and resuspended in 40 ml chilled His Tag-A lysis buffer (54 mM $Na_2HPO_4$, 4 mM $NaH_2PO_4$). Thereafter, they were sonicated for 2 min at 30-s intervals. After sonication, the samples were centrifuged at 15,000*g* for 20 min at 4°C and then loaded onto a Nickle-NTA column (2 ml bed volume; Qiagen) and allowed to pass through by gravity flow. After adding the cell protein extract, the column was washed using His Tag buffer A until $OD_{280}$ registered no additional protein loss from the column. Protein elution was performed in a step-wise manner first

with the loading buffer supplemented with 25 mM imidazole. Twenty-eight 0.5 ml elution fractions were collected after which the elution concentration of imidazole was changed to 150 mM. After taking additional 18 × 0.5 ml fractions, aliquots of each fraction were run on a denaturing SDS-PAGE gel to verify the production of the proper product and product purity. Fractions exhibiting pure products of the expected size were pooled and loaded into a Slide-A-Lyzer 10 kDa dialysis cassette (Thermo Fisher Scientific) and left to dialyze overnight at 4°C in 1× PBS. As the empty vector control did not have the same product as the target mycocypins, the same fractions were taken during elution as the target proteins eluted and used as negative control. After dialysis, the samples were removed from the Slide-A-Lyzer, the protein concentration was determined and normalized to 10 mg/ml for the two target proteins and aliquots were stored at −20°C. As the negative control had no detectable protein, it was diluted in a similar ratio as for the target proteins.

## Protease inhibition assays

Papain (2× crystallized) (EC 3.4.22.2), ficain (EC 3.4.22.3) and trypsin (EC 3.4.21.4) were obtained from Sigma-Aldrich. Legumain (EC 3.4.22.34) was isolated from germinated bean seeds (Sabotic et al., 2011). Recombinant human cathepsin V and L (EC 3.4.22.43), cathepsin H (EC 3.4.22.16) and cathepsin B (EC 3.4.22.1) were prepared as described by Sabotic et al. (2009) and cathepsin X as described by Mitrović et al. (2017). Substrates benzyloxycarbonyl(Z)-Phe-Arg-MCA [7-(4-methyl)-coumarylamide], Z-Arg-Arg-MCA, Z-Ala-Ala-Asn-MCA, Arg-MCA and Abz-Gly-Ile-Val-Arg-Ala-Lys(Dnp)-OH were obtained from Bachem and substrate bz-FEK(Dnp)-OH was used for cathepsin X (Mitrović et al., 2017). *Caenorhabditis elegans* Bristol N2 crude protein extract was prepared by homogenization of adult nematodes in 0.1 M Tris–HCl, pH 8.5 with 0.5 M NaCl followed by centrifugation to remove insoluble material. Adult nematodes grown in liquid culture (Stiernagle, 2006) and cleaned by sucrose flotation (Portman, 2006) were used.

Inhibition kinetics were determined under pseudo-first-order conditions and analysed according to Henderson (1972). Kinetic assays were performed for the cysteine proteases papain, ficain and cathepsin V using Z-Phe-Arg-MCA as the substrate in 0.1 M MES buffer, pH 6.5 with 5 mM DTT (dithiothreitol); for cathepsin L using the same substrate in 0.1 M acetate buffer, pH 5.5 with 5 mM DTT and 1.5 mM EDTA; for cathepsin B (endopeptidase activity) Z-Arg-Arg-MCA in phosphate buffer, pH 6 with 5 mM DTT and 1.5 mM EDTA; for cathepsin B (exopeptidase activity) using Abz-Gly-Ile-Val-Arg-Ala-Lys(Dnp)-OH in 60 mM acetate buffer, pH 5 with 5 mM DTT and 1.5 mM EDTA; for cathepsin H using Arg-MCA in 0.1 M MES buffer with 1.5 mM EDTA and 2 mM DTT and for legumain using Z-Ala-Ala-Asn-AMC substrate in 0.1 M Na-acetate buffer, pH 5.5 with 5 mM DTT. Trypsin was assayed using the substrate Z-Phe-Arg-AMC in 0.1 M Tris–HCl buffer, pH 8.0, containing 20 mM $CaCl_2$. The released MCA was measured using an Infinite®M1000 microplate reader (Tecan).

## Nematotoxicity tests

To study the nematotoxicity of the *L. bicolor* mycocypins 293826 and 311135, we transformed *E. coli* BL21 with plasmids directing the heterologous expression of both mycocypins and of the lectin CGL2 (as positive control) and with the empty vector pET28b (as negative control). The strains were cultivated in LB medium containing 50 μg/ml kanamycin or 100 μg/ml ampicillin at 37°C until an OD of 0.5 was reached. The cells were then chilled on ice for 20 min and expression of proteins was induced by incubating the cells with 0.5 mM IPTG for 2 h at 16°C. Then, 300 μl of cell solution was transferred to nematode growth medium (NGM) plates (2.5 g/L peptone, 50 mM NaCl, 1.7% agar, 1 mM $CaCl_2$, 25 mM $KH_2PO_4$, 1 mM $MgSO_4$ and 13 mM cholesterol) with 1 mM IPTG and 50 μg/ml kanamycin or 100 μg/ml ampicillin, and incubated overnight at 23°C. For verification of protein expression, the bacterial lawns of separate plates were washed off and collected. After incubation, 20 synchronized L1 larvae of *C. elegans* Bristol N2, *C. briggsae* AF16 or *C. tropicalis* JU1373 were added to each plate. After 72 h at 23°C, the fraction of nematodes that reached the L4 developmental stage was determined by counting under a stereo microscope. All nematode strains were obtained from the Caenorhabditis Genetics Center (CGC) at the University of Minnesota. The nematodes were maintained on NGM plates using the *E. coli* strain OP50 and synchronized L1 larvae were isolated according to guidelines (Stiernagle, 2006). The isolated bacterial lawns were used for the preparation of whole-cell extracts and high spin fractions (S), which were centrifuged for 20 min at 14,000*g* to test solubility of the expressed proteins. These samples were loaded on an SDS-PAGE gel and stained with Coomassie (Figure S4). Statistical analysis was conducted using one-way ANOVA, followed by Dunnett's multiple-comparison test to compare the mean of the negative control group to the means of the other groups (Dunnett, 1964). Each group consists of five biological replicates. The ANOVA results are shown in the Supplemental Table S3.

## Collembola isolation and identification

Collembola (springtails) were isolated from soil sampled from Winmalee, NSW, Australia (S33.7°, E150.6°). Multiple specimens were handpicked from a wetted soil using an eyelash and transferred to 90 mm Petri dishes with a substrate of plaster of Paris and activated charcoal. Plates were made by mixing activated charcoal, Paris plaster, and water, in a ratio of 1:8:6 by weight and pouring them into 90 mm Petri dishes. Thereafter, the plates were left to dry for 1 week. Prior to use, to provide enough humidity for the collembola, the charcoal plaster plates were flooded with Milli-Q water and set to soak for 3 min, after which they were drained of excess water. Collembola were reared on dried baker's yeast and identified as *Orthonychiurus* sp. (closest match to *O. folsomi* with 94% similarity) based on the mitochondrial DNA cytochrome C oxidase I gene (Schäffer, 1900). This genus and species have a global distribution (Environment Canada, 2007; Greenslade, 2018; Ospina-Sánchez et al., 2020; Yao et al., 2020) and, therefore, is a proper model organism to test toxicity and feeding deterrence. When sufficient juvenile numbers were obtained, 15 collembola from a variety of stock plates were added and transferred to new plates to establish breeding colonies. The stock colonies were fed baker's yeast as required and plates were kept at high relative humidity using RO water and sealed with parafilm.

## Juvenile collembola feeding trials

To test the impact of the *L. bicolor* mycocypins on *Orthonychiurus* sp. feeding, their food source was treated with purified protein and feeding was measured. To do this, 15 mg of dry baker yeast was mixed with 20 μl of either *L. bicolor* protein (5 mg/ml), Mcp1 at the same concentration (positive control), or 20 μl of the negative control. The negative control was generated by extracting protein from IPTG-induced *E. coli* transformed with an empty protein expression vector and then sent through the same purification protocol as for the mycocypins. As no detectable protein was recovered in this manner, we pooled the same elution fractions as for the mycocypins, dialysed them, and then used this as a negative control. The treated yeasts were added to new plates and 10 juvenile collembola at second to third stage of development were placed onto each plate. In total three replicates were completed at staged intervals to reduce the impact on the breeding colonies. After 1 month, yeast pellets were removed and weighed to determine the amount consumed.

## Adult collembola feeding choice trials

To determine if the *L. bicolor* mycocypins may act as feeding deterrents, a feeding choice trial was performed using adult *Orthonychiurus* sp. To establish these, two pellets of yeast were produced per plate (15 mg yeast + 20 μl protein at 10 mg/ml) whereby one pellet was always treated with the negative control protein purification and the other pellet was treated with Mcp1 (positive control), 293826, or 311135. Following this, five adult collembola were transferred into the plates and left to feed for 1 month. Across this incubation period, regular observations in the late evening (when the most activity in the plates was typically observed) were recorded as to the number of collembola actively feeding at either of the pellets or that were not feeding. Following the incubation period, yeast pellets were removed and weighed to determine the amount of each pellet consumed. A total of six replicates per treatment were performed.

# RESULTS

## Mycocypin-coding genes appear in multiple clades of the Agaricomycotina

Within the 1379 fungal genomes publicly available through the Joint Genome Institute MycoCosm database (Grigoriev et al., 2014), a total of 177 mycocypin-like gene sequences were found in 27 published genomes (including *L. bicolor*). More genes were identified based on annotated functional domains, such as the beta-trefoil fold, rather than using a tBLASTn alignment, suggesting that these are highly divergent genes. However, the beta-trefoil fold is common with fungal lectins and only functional tests could confirm whether the protein is a lectin or a protease inhibitor. Phylogenetically, the very high sequence divergence observed among mycocypin genes was supported by low confidence values on a number of the predicted ancestral branches (Figure 1). Lacking in Ascomycota, mycocypin-like sequences were only found within five different clades of the Agaricomycotina: Cantharellales, Geastrales (within Phallomycetidae), Polyporales, Boletales and Agaricales (Figure 1, inset; Table S1), with the Cantharellales as the most basal clade. Mycocypin-like sequences were not found in all genomes of intervening clades between Cantharellales and Agaricales. An increased number of mycocypin-encoding genes were apparent within the genomes of more recently evolved clades of the Agaricomycotina (e.g. Agaricales) as opposed to more basal clades (e.g. Cantharellales; Figure 1). The phylogenetic tree suggests that mycocypin gene duplications took place

in the symbiotic genus *Laccaria*. Within the genome of *L. bicolor*, mycocypins were clustered on scaffold 17 (i.e. 396760, 382297, 311135). Linkage group 8 exhibited clusters of phylogenetically related mycocypin sequences (e.g. 326006, 326016) together with other clusters of phylogenetically distinct mycocypins (e.g. 385833). An analysis of the orientation of clustered genes found that 25% displayed divergent alignment, 8% showed a convergent alignment and 67% had a co-linear alignment (Table S1).

## Transcripts of *L. bicolor* mycocypin-encoding genes accumulate during mutualistic symbiosis and in fruiting bodies

Using the newly identified and refined list of *L. bicolor* mycocypins (Figure 1), we wished to understand which genes were expressed during symbiosis, fruiting-body development (early as primordia, stipe and cap at medium and late stage; see details in Experimental Procedures) and in axenic FLM (Table S2). Hierarchical clustering based on normalized expression levels (RPKM) revealed five expression profiles (Figure 2). Eight mycocypins genes (335916, 44770, 326095, 461677, 326006, 396760, 385858, 326032) were not expressed or expressed at very low levels in all conditions. Three mycocypins genes (318727, 318749, 293818) were highly, and almost specifically, expressed in stipes and caps of fruiting bodies (Figure 2), while two other genes (326016, 293826) were expressed in both stipes and caps of fruiting bodies as well as in ECM roots. Five mycocypin-encoding genes (602958, 293824, 385833, 382297, 311135) were expressed in ECM roots and, to a lower extent, in FLM, while only one (326015) displayed the highest expression level in FLM (Figure 2). Some of the paralogues (Figure 1) did not share the same expression profile (Figure 2). The *L. bicolor* mycocypin genes

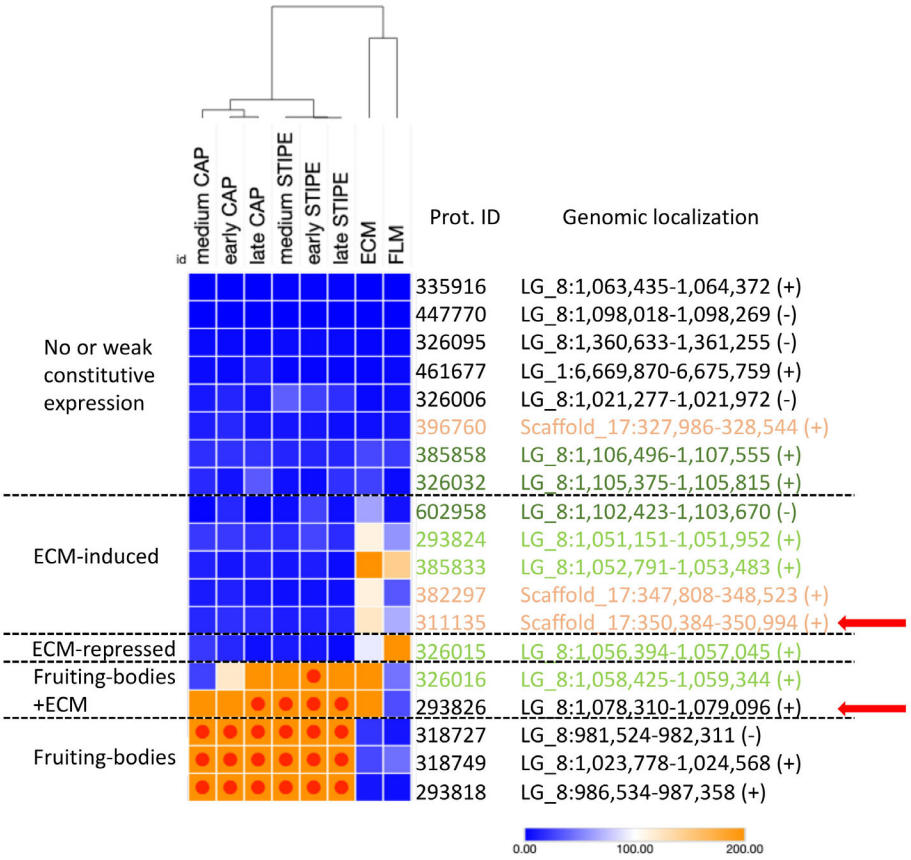


**FIGURE 2** Expression profiles of 19 *Laccaria bicolor* mycocypin genes, during fruiting body development (early, medium and late stages of both stipe and cap), in axenically grown free-living mycelium (FLM) and in vitro ectomycorrhiza (ECM) formed between *L. bicolor* and *Populus tremula* × *alba* roots. Mean expression levels (RPKM) are shown with their hierarchical clustering (1-r, Pearson correlation with average linkage method). Dendrogram produced using Morpheus (https://software.broadinstitute.org/morpheus/). Each gene is represented by a row of coloured boxes (RPKM values) with a single column representing one condition. Expression levels range from pale to saturated colours (orange for high expression; blue for low expression). Red dots indicate the highest gene expression levels (>1000 RPKM). Raw data were extracted from Ruytinx et al. (2021) and are available in Table S2. Protein IDs are given for each gene. Colours in genomic localization indicate genes belonging to the same cluster (orange for scaffold 17 and pale green and dark green for two clusters in LG_8). Red arrows indicate the two mycocypins further characterized

belonging to clusters did not always share the same expression patterns, indicating that they were not co-regulated. For example, mycocypins-311135 and 382297 were strongly upregulated during host colonization while mycocypin-396760 was not significantly expressed or regulated under the tested conditions, although they belong to the same gene cluster. None of the host-induced mycocypin-like genes was predicted with high confidence to be secreted (i.e. SignalP >4).

## Mycocypins are localized in distinct *L. bicolor* tissues

Two *L. bicolor* mycocypin-encoding genes were chosen for further characterization. The choice was made to follow-up on two host-induced sequences that were divergent based on (i) sequence (i.e. phylogenetic relationship; Figure 1), (ii) location within the genome (i.e. different linkage groups/scaffolds) and (iii) expression profiles. Based on these criteria, we chose mycocypin 293826 and mycocypin 311135. Neither mycocypin 293826 nor 311135 were found to be secreted in accordance with *in silico* predictions using SignalP (Petersen et al., 2011) or expression in yeast (Figure S1). Recombinant proteins were produced in, and purified from, *E. coli*. Antibodies against the recombinant proteins were obtained and tested for binding to the proteins of interest [Figure S2(A)]. Indirect immunolocalization in *L. bicolor* tissues was unable to detect either of these mycocypins within the FLM [Figure 3(A,B)] suggesting that protein levels of these two mycocypins were below

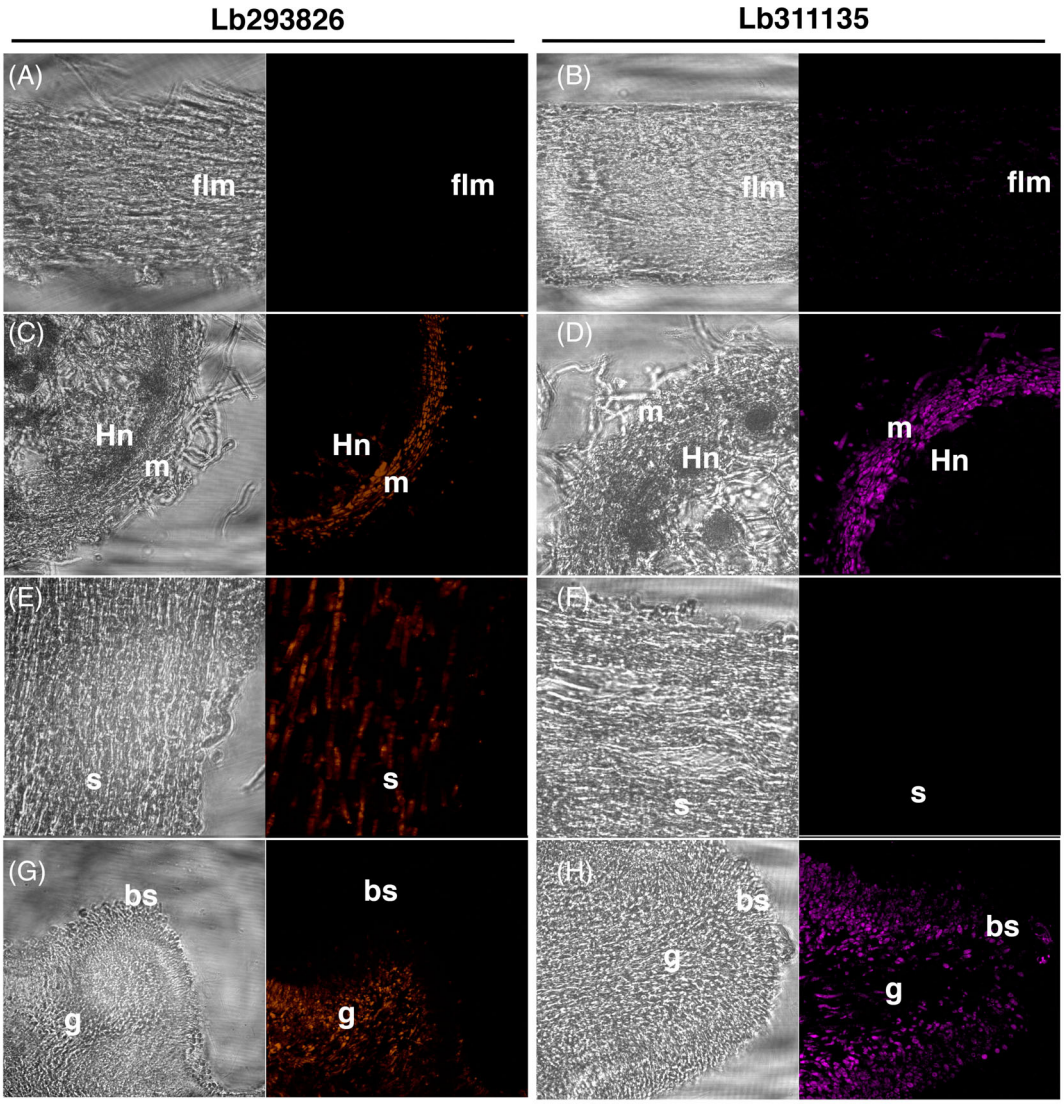


**FIGURE 3** Mycocypins 293826 and 311135 of *L. bicolor* show overlapping protein localization patterns. Indirect immunolocalization of mycocypins 293826 (orange false-coloured signal) and 311135 (purple false-coloured signal) in free-living mycelium (A, B), in transverse cross-sections of a fully colonized mycorrhizal root tip (C, D), in a longitudinal cross-section of the stipe (E, F) and in the gills of a mature fruiting body (G, H). Flm = free-living mycelium; m = mantle; Hn = Hartig net; s = stipe; g = gill; bs = basidium/basidiospores. Photographs were taken with ×60 magnification

detectable levels, despite a detectable expression at the transcript level (Figure 2). Within ECM root tips, mycocypin 293826 exhibited strong expression in the mantle and weaker expression in the hyphae of the Hartig net [Figure 3(C)], while mycocypin 311135 was only localized to the mantle [Figure 3(D)]. Within the mature fruiting body, we found mycocypin 293826 within the stipe [Figure 3(E,F)], consistent with gene expression profiles (Figure 2), while both mycocypins were found in the gills and cap of the fruiting body [Figure 3G,H)]. In the latter tissues, however, mycocypin 293826 was more localized to interior hyphae while mycocypin 311135 was localized to the outer layers of the cap and the basidiospores, despite the latter gene being expressed at very low levels in these tissues.

## *Laccaria bicolor* mycocypins inhibit different cysteine proteases

The high sequence variability and high plasticity of function within mycocypins make it impossible to predict their substrate specificity as protease inhibitors based on sequence alone. To test if mycocypins 293826 and 311135 could inhibit cysteine proteases, we performed in vitro protease inhibition assays using model proteases of either plant or animal origin (Table 1). Mycocypin 293826 inhibited the plant C1-family protease papain, with an inhibitory constant (or $K_i$) in the μM range. The inhibition of papain by mycocypin 311135 was weaker. Furthermore, mycocypin 293826 exhibited weak inhibition of ficain, showing approximately 50% inhibition at 50 μM, whereas mycocypin 311135 did not inhibit this protease at this concentration. Both mycocypins inhibited legumain, a C13-family cysteine protease, albeit weakly, showing approximately 40% inhibition at 50 μM, whereas no inhibition was observed for trypsin, a serine protease. Sequence comparison (Figure 4) explained the differences in inhibition of plant C1-family proteases between mycocypin 293826 and 311135 (Sabotic et al., 2007, 2009). The predicted inhibitory loops in 293826 have high similarity (44%–57%) to those of clitocypin and macrocypin, which are strong inhibitors of papain-like proteases. The protein loops predicted to mediate inhibition of the C1-protease family are shorter (eight and nine residues in 293826 and five and nine residues in 311135) compared to the same loops in either clitocypin (11 and 10 residues) or macrocypin (15 and 9 residues). These changes likely affect the level of inhibition by the *L. bicolor* proteins. Furthermore, the predicted legumain inhibitory loop in mycocypin 293826 contains an asparagine that presumably interacts with the enzyme's active site, not present in the loop of mycocypin 311135. However, the measured strength of inhibition was similarly weak for both. The latter indicates that a different loop could be involved in legumain inhibition in mycocypin 311135 compared to other mycocypins, as observed for trypsin inhibition by the different fungal beta-trefoil inhibitors cnispin and cospin (Caglic et al., 2014; Renko et al., 2012).

Both mycocypins strongly inhibited human cathepsin L with $K_i$ in the pM range. Mycocypin 293826 also strongly inhibited human cathepsin V with $K_i$ in the nM range, whereas mycocypin 311135 was in the μM range. Mycocypin 311135 inhibited both endopeptidase

**TABLE 1** Inhibition of various proteases from plant or animal origin by recombinant mycocypins

| | Inhibition ($K_i$) | | $K_i$ (nM) | |
|---|---|---|---|---|
| **Protease (family, source organism)** | **293826** | **311135** | **Mcp1** | **Clt** |
| Papain (C1, *Carica papaya*) | 7.96 ± 0.38 μM | >100 μM | 0.95 ± 0.33 | 6.2 ± 0.55 |
| Ficain (C1, *Ficus glabrata*) | Yes (weak) | No | ND | ND |
| Legumain (C13, *Phaseolus vulgaris*) | Yes (weak) | Yes (weak) | 3.38 ± 1.44 | 21.5 ± 2.81 |
| Cathepsin V (C1, *Homo sapiens*) | 0.52 ± 0.09 nM | 1.16 ± 0.10 μM | 0.69 ± 0.06 | 0.08 ± 0.03 |
| Cathepsin L (C1, *Homo sapiens*) | 12.8 ± 1.5 pM | 12.4 ± 0.5 pM | 0.64 ± 0.22 | 0.02 ± 0.001 |
| Cathepsin B—endo (C1, *Homo sapiens*) | >100 μM | 1.80 ± 0.30 μM | ND | 480 ± 90 |
| Cathepsin B—exo (C1, *Homo sapiens*) | No | 39.3 ± 5.0 nM | ND | ND |
| Cathepsin H (C1, *Homo sapiens*) | No | No | 100 ± 10 | No |
| Z-Arg-Arg-activity (C1, *C. elegans*)—cathepsin B endopeptidase-like activity | No | Yes (strong) | ND | ND |
| Z-Phe-Arg-activity (C1, *C. elegans*)—cathepsin L/V-like activity | No | Yes (strong) | ND | ND |
| Trypsin (S1, *Bos taurus*) | No | No | No | No |

Equilibrium constants for the inhibition of different proteases were determined under pseudo-first-order conditions and analysed according to Henderson (1972). Kinetic data for the interaction of clitocypin (Clt) from *Clitocybe nebularis* and macrocypin 1 (Mcp1) from *Macrolepiota procera* were reported previously (Sabotic et al., 2009). Standard deviation is given where appropriate; ND, not determined.

and exopeptidase activities of human cathepsin B, whereas 293826 showed weak or no inhibition. No inhibition was found for the human exopeptidase cathepsin H. Interestingly, only mycocypin 311135 showed inhibition of cysteine protease activity in *C. elegans* extracts, and the inhibition of both cathepsin B-like and cathepsin L/V-like activity was strong (approximately 50% inhibition at 1 μM). Therefore, *L. bicolor* mycocypins 293826 and 311135 both inhibit cysteine proteases of different origins. However, the strength of inhibition differs by several orders of magnitude, with mycocypin 311135 being more effective against nematode proteases.

## *Laccaria bicolor* mycocypins inhibit the development of bacterivorous *Caenorhabditis* sp. juveniles

To test whether mycocypin 293826 and 311135 had an impact on nematode development, *E. coli* BL21 strains expressing the recombinant mycocypins were supplied as food to juvenile L1 larvae of the bacterivorous nematode strains *C. elegans* Bristol N2, *C. briggsae* AF16 and *C. tropicalis* JU1373. A previously characterized nematotoxic lectin (CGL2) from the fungus *Coprinopsis cinerea* was used as positive control (Bleuler-Martínez et al., 2011). Protein expression by *E. coli* BL21 was verified for all experiments using a Coomassie-stained SDS-PAGE gel showing heterologous expression and solubility of the proteins [Figure S3(a,b)]. When fed with the *E. coli* strain expressing CGL2, none of the *C. elegans* juveniles reached L4 stage [Figure 5(A)]. Significantly fewer nematodes developed into L4 stage compared to the negative control when the juveniles were fed on a diet of *E. coli* expressing either *L. bicolor* mycocypin 293826 or 311135 [Figure 5(A)]. Most of the mycocypin-fed nematodes died during the L1 developmental stage, but the larvae that survived the first 24 h of the experiment continued to grow until L4 (data not shown). This indicates that L1 stages of *C. elegans*


**FIGURE 4** Alignment of *L. bicolor* mycocypin sequences with those of Clitocypin (Clt) from *Clitocybe nebularis* and Macrocypin 1 (Mcp1) from *Macrolepiota procera*. Identical residues are highlighted in dark grey and similar residues shown in light grey. The two loops involved in inhibition of papain-like proteases (family C1) are underlined with a solid line; the loop involved in inhibition of family C13 or S1 proteases is underlined with a dotted line

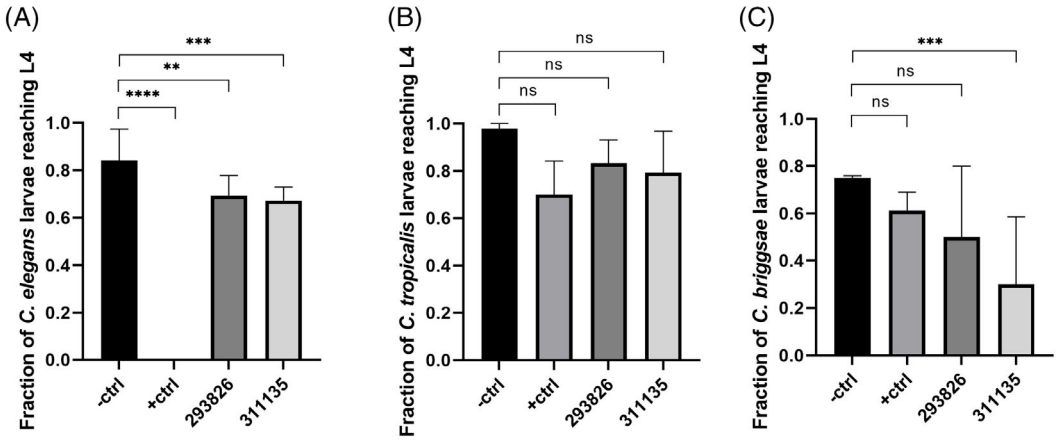


**FIGURE 5** Mycocypins 293826 and 311135 inhibit larval survival of three different *Caenorhabditis* spp. Percentage of *C. elegans* Bristol N2 (A), *C. tropicalis* JU1373C (B) and *C. briggsae* AF16 (C) that survive from larval stage L1 through to L4 after feeding on *E. coli BL21* expressing empty vector pET28b as a negative control (−ctrl), the lectin CGL2 as a positive control (+ctrl) and mycocypins 293826 and 311135. Asterisk denotes significant difference of the mean of a group compared to the mean of control condition (Dunnett's multiple-comparison test); the bars represent the median value of five biological replicates, the error bars represent the 95% confidence interval. ns, not significant; *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; ****$p < 0.0001$ (Table S3)

juveniles are most susceptible to *Laccaria* mycocypins. *Caenorhabditis briggsae* and *C. tropicalis* juveniles were not affected by the lectin CGL2, nor by the mycocypin 293826 [Figure 5(B,C)]. However, a significant toxic effect of mycocypin 311135 on *C. briggsae* was detected [Figure 5(C)]. These results suggest that mycocypins are highly species-specific for bacterivorous nematodes. To assess whether nematotoxicity is specific to *L. bicolor* mycocypins, previously characterized mycocypins from saprotrophic fungi, namely, macrocypins (Mcp) 1, 3 and 4 from *M. procera* and clitocypin (Clt) from *C. nebularis*, were tested for a nematotoxic effect (Figure S4). Mcp1 and Mcp3 showed weak nematotoxicity, confirming this effect is not limited to *L. bicolor* mycocypins and may be a common trait of several mycocypin family members.

## *Laccaria bicolor* mycocypin 311135 acts as a feeding deterrent to collembola

To understand the putative impact of the *L. bicolor* mycocypins on the feeding habits of larger soil fauna, these proteins were also produced in vitro and purified [Figure S3(c)]. They were tested to determine how they impact the grazing of yeast by the springtail *Orthonychiurus* sp. (Collembola). This forager was chosen because it is an eudaphic species (i.e. foragers in the soil horizons associated with ECM hyphae), found in both the Northern and Southern hemisphere locations where *Laccaria* sp. are known to co-occur (Environment Canada, 2007; Greenslade, 2018; Ospina-Sánchez et al., 2020; Yao et al., 2020). The characterized mycocypin Mcp1 from *M. procera* was used as positive control. We found that juvenile collembola consumed an average of 25 $\pm$ 5 μg/day when their yeast food source was treated with the negative control solution [Figure 6(A)]. This feeding was significantly reduced if their food source was treated with either Mcp1 or with *L. bicolor* mycocypin 311135, but not when the yeast was treated with *L. bicolor* mycocypin 293826. We further tested if adult *Orthonychiurus* sp. would, if given a choice of food source, avoid yeast treated with the mycocypins. Similar to juvenile feeding assays, we found that adults were significantly more likely to feed off of yeast treated with the negative control protein source rather than Mcp1 or 311135 [Figure 6(B)]. Visitation of feeding was also recorded during the evening when feeding was most consistently observed. The collembola were more likely to be found actively eating yeast without mycocypin treatment, when the other option was either Mcp1 or 311135-treated yeasts, while they were found to be more equally split between control and 293826-treated yeasts [Figure 6(C)]. In this latter measure, no active feeding was observed on 311135-treated yeasts [Figure 6(C)] despite a quantifiable consumption of these yeasts [Figure 6(B)].

# DISCUSSION

Mutualistic ECM fungal hyphae are exposed to a range of protease defences when growing within plant tissues. Furthermore, their ability to support plant health is dependent upon an intact hyphal network within the soil. Therefore, ECM fungi must have evolved protective mechanisms against both plant defences and predation. In this study, we determined that a suite of uncharacterized mycocypin genes encoded of *L. bicolor* were induced during the colonization process. Work hypothesis was that their products may play a role in either subverting plant protease defences or acting as feeding deterrents for soil fauna. It is worth noting that none of the *L. bicolor* mycocypins induced during symbiosis were predicted to be secreted. This would suggest that, rather than being produced to counter plant immune responses, these proteins are more likely active against fungivores, as data from controlled assays support this latter claim.

During the characterization of mycocypins from saprotrophic fungi, it was hypothesized that their expansion within fungal lineages would allow for novel functional adaptations of the duplicated genes (Sabotic et al., 2006). Expression divergence is considered to occur through post-gene duplication events (Brawand et al., 2011; Farre & Alba, 2010; Ganko et al., 2007; Gu et al., 2002; Huerta-Cepas et al., 2011; Makova & Li, 2003). For example, differential expression of paralogous genes for defence effectors, such as copsin and its paralogues, was previously observed (Kombrink et al., 2019), suggesting that different effector paralogues with different specificities may be active in different fungal tissues. In *L. bicolor* genome, a few of these genes showed expression patterns responsive to the host tissues colonization, the formation of the reproductive tissues, or both. However, no correlation was found between gene expression patterns and gene clustering among *L. bicolor* mycocypin genes. Since the discovery of the first mycocypin sequence (Brzin, 2000), the evolutionary origin of this class of protease inhibitors has been questioned (Sabotic et al., 2006). With the recent increase of available fungal genomes (Kohler et al., 2015; Miyauchi et al., 2020), it is now possible to start tracing their origin. It is worth noting that the high sequence divergence shown by this class of proteins makes their detection difficult, based on sequencing alone. We observed an amplification in the number of mycocypin-coding genes within the Agaricomycotina where they are found in four separate clades (particular in the Agaricales family and *Laccaria* genus) as well as genes clustered within the genome of *L. bicolor*. The evolutionary source of gene clusters has also been debated. Horizontal gene transfer (Walton, 2000) or gene duplication and specialization by inversion during meiosis (Hane et al., 2011) or by the activity of transposable elements (Ehrlich & Yu, 2010; Young

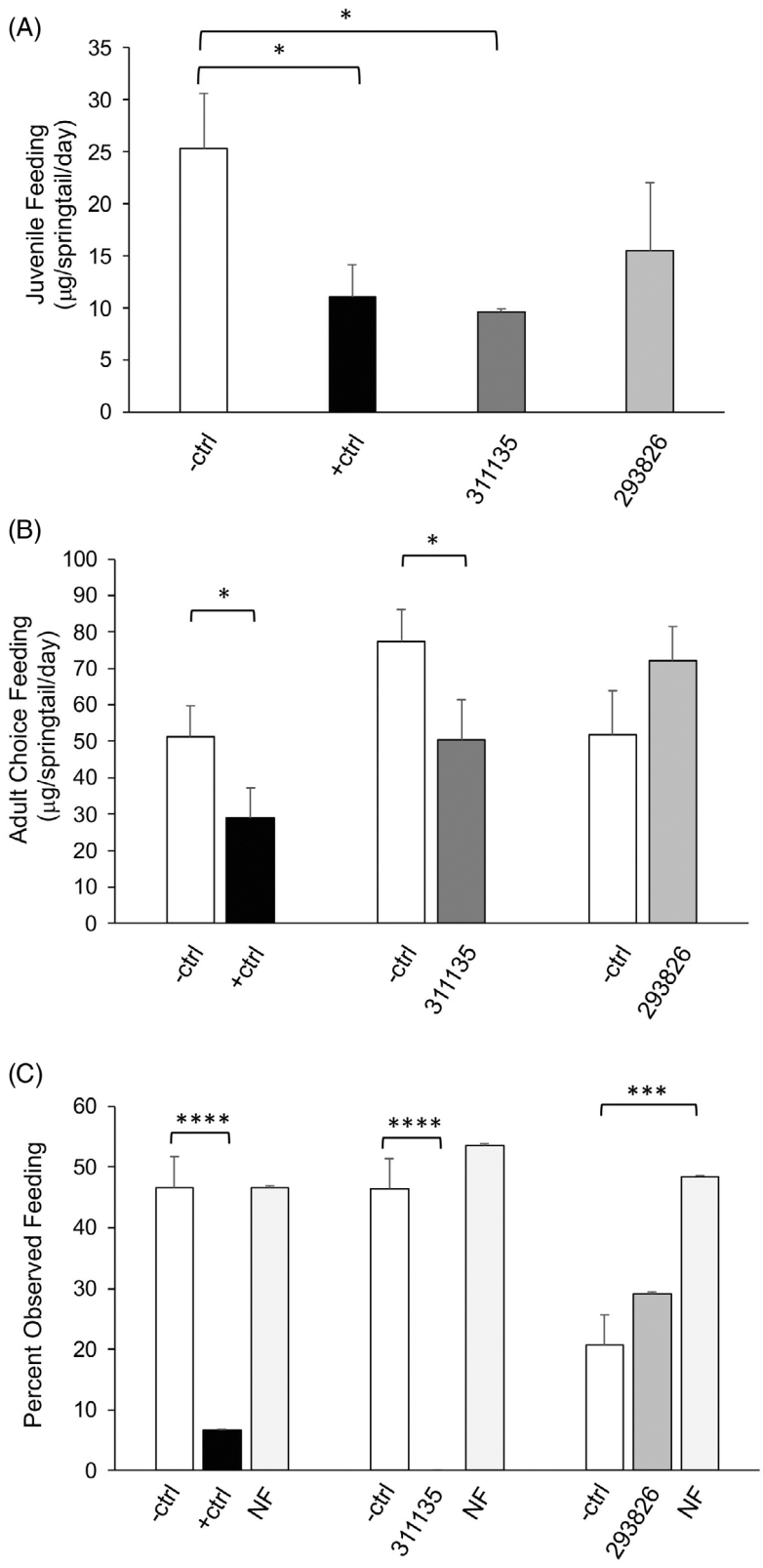


**FIGURE 6** Mycocypin 311135 acts as a feeding deterrent to juvenile and adult collembola. (A) Average daily feeding of juvenile collembola on yeast treated with empty vector protein purification (−ctrl.), Mcp1 (+ctrl), and *L. bicolor* mycocypins 311135 and 293826. (B) Average daily feeding of adult collembola in choice feeding trials where each colony had the choice of yeast treated with empty vector protein purification (−ctrl) paired with either: Mcp1 (+ctrl), or *L. bicolor* mycocypins 311135 or 293826. (C) Percentage of times that adult collembola from (C) were found feeding at either yeast food source or not feeding (NF). Asterisk = significant difference from Neg (Student *t*-test; $*p < 0.05$; $**p < 0.01$,$***p < 0.001$, $****p < 0.0001$), all values ±SE

et al., 2006) have been considered. Most mycocypins from the annotated genomes searched to date, however, do not follow these rules of cluster formation. Altogether, this suggests that mycocypin gene duplication has occurred in *Laccaria* genus, which is an interesting deviation from a typical gene clustering process.

Due to the mycocypins redundancy and the limitation of the molecular tools available in *L. bicolor*, the functional characterization of mycocypins by gene inactivation is challenging. Despite this drawback, we showed that these two proteins were unlikely to be secreted, as has been found for most proteins in this class (Sabotic et al., 2006). However, one secreted mycocypin was previously found in the exudates of axenic mycelium of *L. bicolor* (Vincent et al., 2012). Since we did not characterize the full set of the *L. bicolor* mycocypins, it is likely that the mycocypin detected in *L. bicolor* exudates is either not the same as the ones we studied, derived from non-intact mycelium, or that certain mycocypins of *L. bicolor* are not secreted through a conventional pathway. The non-secretion of the mycorrhiza-induced *L. bicolor* mycocypins is consistent with the hypothesis that they act as toxins against fungivores, taken up during feeding (Künzler, 2018). The presence of the two mycocypins in the hyphal mantle surrounding roots colonized by *L. bicolor* as well as in fruiting bodies (Figure 3), similar to clitocypin and macrocypin from the saprotrophic fungus *M. procera* (Sabotic et al., 2011, 2016), is also indicative of a role as potential fungivory deterrents.

We found that both *L. bicolor* mycocypins inhibit cysteine proteases of different origins, with an inhibition constant in the range of most enzyme-inhibiting drugs (from pM to μM). This is consistent with their sequence divergence and different locations within the genome. Their inhibitory activity is comparable to that of other known mycocypins (Renko et al., 2010; Sabotic et al., 2011), targeting both plant and animal cysteine proteases, with a much stronger activity against the latter. The strong inhibition of the cathepsin-like protease of *C. elegans* by *L. bicolor* mycocypins corroborates the proposed function in defence against nematodes (probably by affecting their development). This is also indicated by the observed nematotoxicity, while an inhibitory role against plant proteases involved in plant defence mechanisms appears less likely.

Interestingly, the strength of inhibition of different cysteine proteases differs greatly between the two *L. bicolor* proteins, indicating different target proteases/protease families. For instance, mycocypin 311135 inhibited *C. elegans* cathepsin L, but not *C. elegans* cathepsin B. In addition, very weak inhibition of human cathepsin X, an orthologue of nematode cathepsin Z, was also observed (data not shown). Comparing the differences between inhibition of human and nematode cathepsin B, it is conceivable that mycocypins show stronger inhibition of the cathepsin Z (nematode-derived cathepsin X). As both cathepsin L and Z play a role in the development of *C. elegans* and are increasingly expressed before the L1/L2 moult (Hashmi et al., 2002, 2004), data suggest that *L. bicolor* mycocypin 311135 impacts the L1/L2 moulting step. Data also showed that only mycocypin 311135 acted as a feeding deterrent for the fungivore Collembola. The same result was observed with the bacterivorous nematode *C. briggsae*, while both *L. bicolor* mycocypins tested inhibited the development of *C. elegans*. Interestingly, neither of the two mycocypins affected the development of *C. tropicalis*. These results suggest that the effect of mycocypins on nematodes is species-specific, as is the case for the lectin CGL2 (this study; Tayyrov et al., 2019). Differences in toxicity and antifeedant properties against fungivorous nematodes have also been observed previously for other saprotrophic fruiting body defence proteins (Tayyrov et al., 2018).

In summary, we established that the genome of the mutualistic fungus *L. bicolor* displays a duplication of protease inhibitor mycocypin-encoding genes that are either expressed in symbiotic tissues, the reproductive tissues, both, or not expressed under the tested conditions. Both toxic and feeding deterrent properties were found for two host-induced mycocypins from *L. bicolor*. It remains to be seen if other mutualistic fungi produce similar mycocypins with sequence divergence too great to be captured in the present study. It would also be of interest to investigate other possible roles of mycocypins in fungal biology (e.g. development and regulation of endogenous proteases). In conclusion, we have identified one pathway by which ECM fungi may prevent feeding on their own hyphae and host roots by soil fauna and, thereby, indirectly promote plant growth. Further research *in situ* on the complex role of mycocypins in fungal physiology is needed to elucidate their greater ecological significance.

## AUTHOR CONTRIBUTIONS

J.M.P., F.M. and C.V.-F. designed the research. J.M.P. performed phylogenetic analysis, yeast-secretion assays and immunolocalization experiments. A.K. provided transcriptomic analysis. C.V.-F. produced recombinant proteins. J.S. performed proteases-inhibition assays and nematotoxicity assay of Mcps and Clt. E.V. and M.K. performed nematoxocity assays. J.M.P., F.S. and U.N.N. performed the springtail assays. All authors participated in the writing of this manuscript.

## ACKNOWLEDGEMENTS

This work was supported by the European Commission within the Project ENERGYPOPLAR (FP7-211917) and the Laboratory of Excellence ARBRE (ANR-11-LABX-0002-01). This research was also sponsored by the Genomic Science Program (project 'Plant–Microbe Interactions'), U.S. Department of Energy, Office of

Science, Biological and Environmental Research under the contract DE-AC05-00OR22725. J.M.P. would like to acknowledge the Australian Research Council for research funding (DE150100408). We thank the Joint Genome Institute for making their fungal genome data publicly available through the MycoCosm database. We thank Dr. Urša Pečar Fonović for the cathepsin X inhibition assay. We thank Giles Ross for springtail COI sequencing. E.V. and M.K. are supported by the Swiss National Science Foundation (Grant No. 31003A_173097 to M.K.). J.S. is supported by Slovenian Research Agency (Grant Nos. P4-0127 and P4-0432). We thank Krista Plett for the English proofreading.

## CONFLICT OF INTEREST

The authors declare that there is no conflict of interest related to this work.

## DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

## ORCID

*Jonathan M. Plett* https://orcid.org/0000-0003-0514-8146
*Jerica Sabotič* https://orcid.org/0000-0002-2404-0192
*Eva Vogt* https://orcid.org/0000-0002-3761-5266
*Fridtjof Snijders* https://orcid.org/0000-0002-6287-2758
*Annegret Kohler* https://orcid.org/0000-0002-9575-9567
*Uffe N. Nielsen* https://orcid.org/0000-0003-2400-7453
*Markus Künzler* https://orcid.org/0000-0003-1275-0629
*Francis Martin* https://orcid.org/0000-0002-4737-3715
*Claire Veneault-Fourrey* https://orcid.org/0000-0001-9778-2659

## REFERENCES

Asplund, J., Bokhorst, S. & Wardle, D.A. (2013) Secondary compounds can reduce the soil micro-arthropod effect on lichen decomposition. *Soil Biology and Biochemistry*, 66, 10–16.

Basso, V., Kohler, A., Miyauchi, S., Singan, V., Guinet, F., Šimura, J. et al. (2020) An ectomycorrhizal fungus alters sensitivity to jasmonate, salicylate, gibberellin, and ethylene in host roots. *Plant, Cell and Environment*, 43, 1047–1068.

Bleuler-Martínez, S., Butschi, A., Garbani, M., Wälti, M.A., Wohlschlager, T., Potthoff, E. et al. (2011) A lectin-mediated resistance of higher fungi against predators and parasites. *Molecular Ecology*, 20(14), 3056–3070.

Bokhorst, S. & Wardle, D.A. (2014) Snow fungi as a food source for micro-arthropods. *European Journal of Soil Biology*, 60, 77–80.

Bokhorst, S., Phoenix, G.K., Bjerke, J.W., Callaghan, T.V., Huyer-Brugman, F. & Berg, M.P. (2012) Extreme winter warming events more negatively impact small rather than large soil fauna: shift in community composition explained by traits not taxa. *Global Change Biology*, 18(3), 1152–1162.

Brawand, D., Soumillon, M., Necsulea, A., Julien, P., Csardi, G., Harrigan, P. et al. (2011) The evolution of gene expression levels in mammalian organs. *Nature*, 478(7369), 343.

Brzin, J. (2000) Clitocypin, a new type of cysteine proteinase inhibitor from fruit bodies of mushroom clitocybe nebularis. *Journal of Biological Chemistry*, 275(26), 20104–20109.

Caglic, P.A., Renko, M., Turk, D., Kos, J. & Sabotic, J. (2014) Fungal beta-trefoil trypsin inhibitors cnispin and cospin demonstrate the plasticity of the beta-trefoil fold. *Biochimica Et Biophysica Acta-Proteins and Proteomics*, 1844(10), 1749–1756.

Daguerre, Y., Basso, V., Hartmann-Wittulski, S., Schellenberger, R., Meyer, L., Bailly, J. et al. (2020) The mutualism effector MiSSP7 of *Laccaria bicolor* alters the interactions between the poplar JAZ6 protein and its associated proteins. *Scientific Reports*, 10(1), 20362. https://doi.org/10.1038/s41598-020-76832-6

Dereeper, A., Guignon, V., Blanc, G., Audic, S., Buffet, S., Chevenet, F. et al. (2008) Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Research*, 36, W465–W469.

Dunnett, C.W. (1964) New tables for multiple comparisons with a control. *Biometrics*, 20(3), 482–491.

Ehrlich, K.C. & Yu, J.J. (2010) Aflatoxin-like gene clusters and how they evolved. In: Rai, M. & Varma, A. (Eds.) *Mycotoxins in food, feed and bioweapons*. Berlin, Heidelberg: Springer, pp. 65–75.

Environment-Canada. (2007) Test for measuring survival and reproduction of springtails exposed to contaminants in soil. In: Scroggins, R.P. (Ed.) *Environmental protection series. Environment Canada, Environmental Technology Centre, Report EPS 1/RM/47*. Ottawa, ON, Canada: Ottawa, p. 146.

Farre, D. & Alba, M.M. (2010) Heterogeneous patterns of gene-expression diversification in mammalian gene duplicates. *Molecular Biology and Evolution*, 27(2), 325–335.

Felten, J., Kohler, A., Morin, E., Bhalerao, R.P., Palme, K., Martin, F. et al. (2009) The ectomycorrhizal fungus *Laccaria bicolor* stimulates lateral root formation in poplar and Arabidopsis through auxin transport and signaling. *Plant Physiology*, 151(4), 1991–2005.

Ganko, E.W., Meyers, B.C. & Vision, T.J. (2007) Divergence in expression between duplicated genes in Arabidopsis. *Molecular Biology and Evolution*, 24(10), 2298–2309.

Giannakis, N. & Sanders, F.E. (1990) Interactions between mycophagous nematodes, mycorrhizal and other soil fungi. *Agriculture, Ecosystems & Environment*, 29(1–4), 163–167. https://doi.org/10.1016/0167-8809(90)90270-N

Goulet, M.C., Dallaire, C., Vaillancourt, L.P., Khalf, M., Badri, A.M., Preradov, A. et al. (2008) Tailoring the specificity of a plant cystatin toward herbivorous insect digestive cysteine proteases by single mutations at positively selected amino acid sites. *Plant Physiology*, 146(3), 1010–1019.

Greenslade, P. (2018) Why are there so many exotic springtails in Australia? A review. *Soil Organisms*, 90(3), 141–156.

Grigoriev, I.V., Nikitin, R., Haridas, S., Kuo, A., Ohm, R., Otillar, R. et al. (2014) MycoCosm portal: gearing up for 1000 fungal genomes. *Nucleic Acids Research*, 42(D1), D699–D704.

Gu, Z.L., Nicolae, D., Lu, H.H.S. & Li, W.H. (2002) Rapid divergence in expression between duplicate genes inferred from microarray data. *Trends in Genetics*, 18(12), 609–613.

Hane, J.K., Rouxel, T., Howlett, B.J., Kema, G.H.J., Goodwin, S.B. & Oliver, R.P. (2011) A novel mode of chromosomal evolution peculiar to filamentous Ascomycete fungi. *Genome Biology*, 12(5), R45.

Hashmi, S., Britton, C., Liu, J., Guiliano, D.B., Oksov, Y. & Lustigman, S. (2002) Cathepsin L is essential for embryogenesis and development of *Caenorhabditis elegans*. *Journal of Biological Chemistry*, 277(5), 3477–3486.

Hashmi, S., Zhang, J., Oksov, Y. & Lustigman, S. (2004) The *Caenorhabditis elegans* cathepsin Z-like cysteine protease, Ce-CPZ-1,

has a multifunctional role during the worms' development. *Journal of Biological Chemistry*, 279(7), 6035–6045.
Henderson, P.J. (1972) Linear equation that describes steady-state kinetics of enzymes and subcellular particles interacting with tightly bound inhibitors. *Biochemical Journal*, 127(2), 321.
Huerta-Cepas, J., Dopazo, J., Huynen, M.A. & Gabaldon, T. (2011) Evidence for short-time divergence and long-time conservation of tissue-specific expression after gene duplication. *Briefings in Bioinformatics*, 12(5), 442–448.
Kaneda, S. & Kaneko, N. (2004) The feeding preference of a collembolan (*Folsomia candida* Willem) on ectomycorrhiza (*Pisolithus tinctorius* (Pers.)) varies with mycelial growth condition and vitality. *Applied Soil Ecology*, 27(1), 1–5s.
Klironomos, J.N. & Kendrick, W.B. (1996) Palatability of microfungi to soil arthropods in relation to the functioning of arbuscular mycorrhizae. *Biology and Fertility of Soils*, 21(1–2), 43–52.
Kohler, A., Kuo, A., Nagy, L.G., Morin, E., Barry, K.W., Buscot, F. et al. (2015) Convergent losses of decay mechanisms and rapid turnover of symbiosis genes in mycorrhizal mutualists. *Nature Genetics*, 47(4), 410–U176.
Kombrink, A., Tayyrov, A., Essig, A., Stöckli, M., Micheller, S., Hintze, J. et al. (2019) Induction of antibacterial proteins and peptides in the coprophilous mushroom *Coprinopsis cinerea* in response to bacteria. *The ISME Journal*, 13, 588–602.
Künzler, M. (2018) How fungi defend themselves against microbial competitors and animal predators. *PLoS Pathogens*, 14(9), e1007184.
Makova, K.D. & Li, W.H. (2003) Divergence in the spatial pattern of gene expression between human duplicate genes. *Genome Research*, 13(7), 1638–1645.
Martin, F., Aerts, A., Ahren, D., Brun, A., Danchin, E.G.J., Duchaussoy, F. et al. (2008) The genome of *Laccaria bicolor* provides insights into mycorrhizal symbiosis. *Nature*, 452(7183), 88–U87.
Martin, F., Kohler, A., Murat, C., Veneault-Fourrey, C. & Hibbett, D.S. (2016) Unearthing the roots of ectomycorrhizal symbioses. *Nature Reviews Microbiology*, 14(12), 760–773.
Michaud, D., Berniervadnais, N., Overney, S. & Yelle, S. (1995) Constitutive expression of digestive cysteine proteinase forms during development of the Colorado potato beetle, *Leptinotarsa decemlineata* say (Coleoptera, Chrysomelidae). *Insect Biochemistry and Molecular Biology*, 25(9), 1041–1048.
Misas Villamil, J., Mueller, A., Demir, F., Meyer, U., Ökmen, B., Schulze Hüynck, J. et al. (2019) A fungal substrate mimicking molecule suppresses plant immunity via an inter-kingdom conserved motif. *Nature Communications*, 10(1), 1576.
Mitrović, A., Pečar Fonović, U. & Kos, J. (2017) Cysteine cathepsins B and X promote epithelial-mesenchymal transition of tumor cells. *European Journal of Cell Biology*, 96(6), 622–631.
Miyauchi, S., Kiss, E., Kuo, A., Drula, E., Kohler, A., Sánchez-García, M. et al. (2020) Large-scale genome sequencing of mycorrhizal fungi provides insights into the early evolution of symbiotic traits. *Nature Communications*, 11(1), 5125.
Ospina-Sánchez, C.M., Soto-Adames, F.N. & González, G. (2020) Checklist and distribution of Collembola from greater Puerto Rico. *Biodiversity Data Journal*, 8, e52054.
Passarge, A., Demir, F., Green, J., Depotter, J.R.L., Scott, B., Huesgen, P.F. et al. (2021) Host apoplastic cysteine protease activity is supressed during the mutualistic association of *Lolium perenne* and *Epichloë festucae*. *Journal of Experimental Botany*, 72(9), 3410–3426.
Pellegrin, C., Martin, F. & Veneault-Fourrey, C. (2019) Molecular signalling during the ectomycorrhizal symbiosis. In: Hoffmeister, D. G.M. (Ed.) *Biology of the fungal cell. The Mycota (a comprehensive treatise on fungi as experimental Systems for Basic and Applied Research)*. Cham: Springer, pp. 95–109.
Petersen, T.N., Brunak, S., von Heijne, G. & Nielsen, H. (2011) SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature Methods*, 8(10), 785–786.
Pickles, B.J. & Simard, S.W. (2017) Mycorrhizal networks and forest resilience to drought. In: Johnson, N.C., Gehring, C. & Jansa, J. (Eds.) *Mycorrhizal mediation of soil – fertility, structure, and carbon storage*. Amsterdam: Elsevier, pp. 319–339. https://doi.org/10.1016/B978-0-12-804312-7.00018-8
Plett, J.M., Kemppainen, M., Kale, S.D., Kohler, A., Legue, V., Brun, A. et al. (2011) A secreted effector protein of *Laccaria bicolor* is required for symbiosis development. *Current Biology*, 21(14), 1197–1203.
Plett, J.M., Daguerre, Y., Wittulsky, S., Vayssières, A., Deveau, A., Melton, S.J. et al. (2014) Effector MiSSP7 of the mutualistic fungus *Laccaria bicolor* stabilizes the Populus JAZ6 protein and represses jasmonic acid (JA) responsive genes. *Proceedings of the National Academy of Sciences of the United States of America*, 111(22), 8299–8304.
Plett, J.M., Tisserant, E., Brun, A., Morin, E., Grigoriev, I.V., Kuo, A. et al. (2015) The mutualist *Laccaria bicolor* expresses a core gene regulon during the colonization of diverse host plants and a variable regulon to counteract host-specific defenses. *Molecular Plant-Microbe Interactions*, 28, 261–273.
Plett, J.M., Plett, K.L., Wong-Bajracharya, J., de Freitas, P.M., Dutra Costa, M., Kohler, A. et al. (2020) Mycorrhizal effector PaMiSSP10b alters polyamine biosynthesis in *eucalyptus* root cells and promotes root colonization. *New Phytologist*, 228, 712–727.
Pohleven, J., Renko, M., Magister, Š., Smith, D.F., Künzler, M., Štrukelj, B. et al. (2012) Bivalent carbohydrate binding is required for biological activity of *Clitocybe nebularis* lectin (CNL), the N, N′-diacetyllactosediamine (GalNAcβ1–4GlcNAc, LacdiNAc)-specific lectin from basidiomycete *C. nebularis*. *Journal of Biological Chemistry*, 287, 10602–10612.
Portman, D.S. (2006) Profiling *C. elegans* gene expression with DNA microarrays. *WormBook*, 1–11. https://doi.org/10.1895/wormbook.1.104.1
Renko, M., Sabotic, J., Mihelic, M., Brzin, J., Kos, J. & Turk, D. (2010) Versatile loops in mycocypins inhibit three protease families. *Journal of Biological Chemistry*, 285(1), 308–316.
Renko, M., Sabotic, J. & Turk, D. (2012) beta-Trefoil inhibitors - from the work of Kunitz onward. *Biological Chemistry*, 393(10), 1043–1054.
Rohlfs, M. (2015) Fungal secondary metabolite dynamics in fungus–grazer interactions: novel insights and unanswered questions. *Frontiers in Microbiology*, 5, 788. https://doi.org/10.3389/fmicb.2014.00788
Ruytinx, J., Miyauchi, S., Hartmann-Wittulsky, S., de Freitas Pereira, M., Guinet, F., Churin, J.L. et al. (2021) A transcriptomic atlas of the ectomycorrhizal fungus *Laccaria bicolor*. *Microorganisms*, 9(12), 2612.
Sabotic, J., Gaser, D., Rogelj, B., Gruden, K., Strukelj, B. & Brzin, J. (2006) Heterogeneity in the cysteine protease inhibitor clitocypin gene family. *Biological Chemistry*, 387(12), 1559–1566.
Sabotic, J., Galesa, K., Popovic, T., Leonardi, A. & Brzin, J. (2007) Comparison of natural and recombinant clitocypins, the fungal cysteine protease inhibitors. *Protein Expression and Purification*, 53(1), 104–111.
Sabotic, J., Popovic, T., Puizdar, V. & Brzin, J. (2009) Macrocypins, a family of cysteine protease inhibitors from the basidiomycete *Macrolepiota procera*. *FEBS Journal*, 276(16), 4334–4345.
Sabotic, J., Kilaru, S., Budic, M., Gasparic, M.B., Gruden, K., Bailey, A.M. et al. (2011) Protease inhibitors clitocypin and macrocypin are differentially expressed within basidiomycete fruiting bodies. *Biochimie*, 93(10), 1685–1693.
Sabotic, J., Ohm, R.A. & Kunzler, M. (2016) Entomotoxic and nematotoxic lectins and protease inhibitors from fungal fruiting bodies. *Applied Microbiology and Biotechnology*, 100(1), 91–111.
Schäffer, C. (1900) Ueber württembergische Collembola. In: *Jahreshefte des Vereins für Vaterländische Naturkunde: Württemberg*, Vol. 56, pp. 245–280. Harvard University.

Schulze Hüynck, J., Kaschani, F., van der Linde, K., Ziemann, S., Müller, A.N., Kaiser, M. et al. (2019) Proteases underground: analysis of the maize root apoplast identifies organ specific papain-like cysteine protease activity. *Frontiers in Plant Science*, 10, 473.

Simard, S.W., Beiler, K.J., Bingham, M.A., Deslippe, J.R., Philip, L. J. & Teste, F.P. (2012) Mycorrhizal networks: mechanisms, ecology and modelling. *Fungal Biology Reviews*, 26(1), 39–60.

Šmid, I., Gruden, K., Buh Gašparič, M., Koruza, K., Petek, M., Pohleven, J. et al. (2013) Inhibition of the growth of Colorado potato beetle larvae by macrocypins, protease inhibitors from the parasol mushroom. *Journal of Agricultural and Food Chemistry*, 61(51), 12499–12509.

Stiernagle, T. (2006) *Maintenance of C. elegans*. *WormBook*: *The online review of C. elegans biology*. *The C. elegans Research Community*, *WormBook*. Available from: http://www.wormbook.org.

Tamura, K., Stecher, G. & Kumar, S. (2021) MEGA11: molecular evolutionary genetics analysis version 11. *Molecular Biology and Evolution*, 38, 3022–3027.

Tayyrov, A., Schmieder, S.S., Bleuler-Martinez, S., Plaza, D.F. & Künzler, M. (2018) Toxicity of potential fungal defense proteins towards the Fungivorous nematodes *Aphelenchus avenae* and *Bursaphelenchus okinawaensis*. *Applied and Environmental Microbiology*, 84(23), e02051-18.

Tayyrov, A., Stanley, C.E., Azevedo, S. & Künzler, M. (2019) Combining microfluidics and RNA-sequencing to assess the inducible defensome of a mushroom against nematodes. *BMC Genomics*, 20, 243.

Van der Linde, K., Hemetsberger, C., Kastner, C., Kaschani, F., van der Hoorn, R.A., Kumlehn, J. et al. (2012) A maize cystatin suppresses host immunity by inhibiting apoplastic cysteine proteases. *The Plant Cell*, 24(3), 1285–1300. https://doi.org/10.1105/tpc.111.093732

Vincent, D., Kohler, A., Claverol, S., Solier, E., Joets, J., Gibon, J. et al. (2012) Secretome of the free-living mycelium from the ectomycorrhizal basidiomycete *Laccaria bicolor*. *Journal of Proteome Research*, 11(1), 157–171.

Walton, J.D. (2000) Horizontal gene transfer and the evolution of secondary metabolite gene clusters in fungi: an hypothesis. *Fungal Genetics and Biology*, 30(3), 167–171.

Wohlschlager, T., Butschi, A., Grassi, P., Sutov, G., Gauss, R., Hauck, D. et al. (2014) Methylated glycans as conserved targets of animal and fungal innate defense. *Proceedings of the National Academy of Sciences of the United States of America*, 111(27), E2787–E2796.

Yao, H., Xie, Z., Dong, J. & Sun, X. (2020) The complete mitochondrial genome of Orthonychiurus folsomi (Collembola: Onychiuridae). *Mitochondrial DNA Part B*, 5(1), 340–341.

Young, C.A., Felitti, S., Shields, K., Spangenberg, G., Johnson, R.D., Bryan, G.T. et al. (2006) A complex gene cluster for indole-diterpene biosynthesis in the grass endophyte Neotyphodium lolii. *Fungal Genetics and Biology*, 43(10), 679–693.

Zhao, Y., Botella, M.A., Subramanian, L., Niu, X.M., Nielsen, S.S., Bressan, R.A. et al. (1996) Two wound-inducible soybean cysteine proteinase inhibitors have greater insect digestive proteinase inhibitory activities than a constitutive homolog. *Plant Physiology*, 111(4), 1299–1306.

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Plett, J.M., Sabotič, J., Vogt, E., Snijders, F., Kohler, A., Nielsen, U.N. et al. (2022) Mycorrhiza-induced mycocypins of *Laccaria bicolor* are potent protease inhibitors with nematotoxic and collembola antifeedant activity. *Environmental Microbiology*, 24(10), 4607–4622. Available from: https://doi.org/10.1111/1462-2920.16115