

DISS. ETH NO. 28555

**Combining reverse genetics and peptidomics to characterize  
ribosomal peptides in fungi**

A thesis submitted to attain the degree of

DOCTOR OF SCIENCES of ETH ZÜRICH

(Dr. sc. ETH Zürich)

presented by

EVA VOGT

MSc, ETH Zürich

born on 06.06.1990

citizen of Switzerland

accepted on the recommendation of

Prof. Dr. Markus Künzler

Prof. Dr. Christine Beemelmans

Prof. Dr. Jörn Piel

2022



## Acknowledgements

First, I thank Prof. Dr. Markus Künzler for the opportunity to work in his group and for his continuous support, generosity, patience and guidance during these four years. I was fortunate to be part of his group and to have the chance to work in this fascinating field of research.

I am also thankful to my Ph.D. committee, Prof. Dr. Christine Beemelmans, Prof. Dr. Jörn Piel and Prof. Dr. Cyril Zipfel, for investing their precious time in our annual meetings and providing valuable input.

During my Ph.D., I supervised three talented Master students. I would like to thank Lukas Sonderegger for his work on the heterologous expression of KEPs and establishment of the *C. cinerea* knockout protocol during his semester project and his Master thesis, Marc Flachsmann for his attempts to test a new knockout protocol during his semester project, and Ying-Yu Chen for his work on peptide extractions from fruiting bodies during his Master thesis.

I would also like to thank our collaborators, Dr. Antje Dittmann from the Functional Genomic Center Zürich, Dr. Stephanie Herzog and Prof. Dr. André Fleißner from the Technical University of Braunschweig, Germany, and Dr. Jerica Sabotič and Tanja Zupan from the Jožef Stefan Institute, Slovenia.

I am also grateful to past and present members of the Künzler and Aebi lab, especially Emmanuel Matabaro for his scientific support and friendship during my time in the lab.

Finally, my deepest gratitude goes to my parents Christine Vogt-Jaeger and Hans-Peter Vogt for everything they have done for me, including their kind and generous support during my studies at ETH, to my siblings Katharina and Viktor, and to Ben, who has accompanied and supported me since my Bachelor studies.

During my time at ETH, I've shared my home with two cats, Merlin and Spock. I thank them, too, for their companionship and love.



## Contents

<b>Summary</b> .....	<b>5</b>
<b>Zusammenfassung</b> .....	<b>7</b>
<b>Chapter 1</b> .....	<b>10</b>
Introduction: Discovery of novel fungal RiPP biosynthetic pathways and their application for the development of peptide therapeutics.....	10
Appendix 1: Current state of research on fungal RiPPs .....	37
<b>Chapter 2</b> .....	<b>45</b>
Structural and functional analysis of peptides derived from KEX2-processed repeat proteins in agaricomycetes using reverse genetics and peptidomics.....	45
Supplementary information .....	75
Appendix 2: Functional assays on KEP-derived peptides .....	135
<b>Chapter 3</b> .....	<b>143</b>
Genome sequences of <i>Rhizopogon roseolus</i> , <i>Mariannaea elegans</i> , <i>Myrothecium verrucaria</i> and <i>Sphaerostilbella broomeana</i> and the identification of biosynthetic gene clusters for fungal peptide natural products.....	143
Supplementary information .....	156
<b>Chapter 4</b> .....	<b>158</b>
Concluding remarks and future perspectives.....	158
<b>Curriculum Vitae</b> .....	<b>170</b>

## Summary

Fungi are competent producers of a large variety of natural products, including peptides. Peptides can be non-ribosomally encoded, referred to as non-ribosomal peptides (NRPs), or ribosomally encoded, referred to as ribosomally synthesized and posttranslationally modified peptides (RiPPs). While many fungal NRPs have been studied to date, including prominent representatives such as penicillin and cyclosporin, the first fungal RiPP was not discovered until 2007. RiPPs have several properties that make them interesting targets for research and biotechnological applications: Their precursors are genetically encoded and can be genetically modified, they are processed by promiscuous enzymes, and they can adapt a high degree of complexity through various posttranslational modifications. In recent years, another chapter was added to research on RiPPs when it was discovered that a class of peptide-encoding precursor proteins, called KEX2-processed repeat proteins (KEPs), are widely distributed in the fungal kingdom and their functions are largely unknown. The aim of this thesis was to expand our understanding of ribosomally produced fungal peptides, with a special focus on peptides derived from KEPs.

Chapter 1 provides an introduction to fungal RiPPs and their biosynthesis. First, the RiPP family of amatoxins/phallotoxins is discussed, which includes famous lethal toxins present in, e.g., the death cap, followed by the most recently discovered RiPP family, the borosins, which carry backbone N-methylations. Finally, the dikaritins are examined. These are functionally diverse cyclic peptides processed from precursor proteins with a KEP-like structure. In addition, the potential of RiPP biosynthetic pathways for the development of peptide therapeutics is discussed, both in terms of the identification of new fungal RiPPs and in terms of a synthetic biology approach using biosynthetic enzymes to generate new-to-nature peptide libraries that could be screened for bioactivity.

Chapter 2 describes the structural and functional characterization of KEP-derived peptides in the agaricomycetes *Coprinopsis cinerea*, *Lentinula edodes*, *Pleurotus ostreatus* and *Pleurotus eryngii*. Genome mining identified several KEPs in *C. cinerea*. A selection of these proteins was successfully expressed in *Pichia pastoris* and processed by the yeast's enzymatic machinery, resulting in an accumulation of KEP-derived peptides in the culture supernatant. This experimental setup was used to establish a protocol for the extraction and detection of KEP-derived peptides in fungal tissues by mass spectrometry. Using this KEP detection protocol, multiple KEP-derived peptides were confirmed in culture supernatant and tissue samples of *C. cinerea*, *L. edodes*, *P. ostreatus* and *P. eryngii*. In addition, we established a CRISPR-mediated knockout protocol for *C. cinerea* and generated knockout strains of six different KEPs, three KEX2 homologs and one KEX1 homolog. We found that the *kep* knockouts resulted in no discernible phenotype, while the *kex* knockouts revealed defects in mycelial growth and fruiting body formation. These results suggest that the investigated KEP-derived peptides do not have an organism-intrinsic role but rather function in the interaction with the biotic environment. The KEX proteases, on the other hand, apparently target a wide variety of substrates, including some with roles in mycelial growth and differentiation.

Chapter 3 presents the new draft genome sequences of four fungal species: *Rhizopogon roseolus*, *Mariannaea elegans*, *Myrothecium verrucaria* and *Sphaerostilbella broomeana*. The genomes were screened for precursor proteins or gene clusters involved in the synthesis of the peptide macrocycles mariannamides and the backbone N-methylated macrocycles verrucamides and broomeanamides. We performed BLAST searches to find potential RiPP precursor proteins of these peptides and included a general screen for borosin-producing precursor proteins. While this search was unsuccessful, we found non-ribosomal peptide synthase (NRPS) gene clusters in an antiSMASH screening that may be responsible for the biosynthesis of mariannamides, verrucamides and broomeanamides. In addition, we discovered several gene clusters encoding putative peptaibols, an antimicrobial class of peptides.

Finally, Chapter 4 summarizes the findings of this thesis and gives an outlook on future experiments

## Zusammenfassung

Pilze sind kompetente Produzenten einer großen Vielfalt von Naturprodukten, einschließlich Peptiden. Peptide können nicht-ribosomal kodiert sein, bezeichnet als «non-ribosomal peptides» (NRPs), oder ribosomal, bezeichnet als «ribosomal synthesized and posttranslationally modified peptides» (RiPPs). Während bisher viele pilzliche NRPs untersucht wurden, darunter prominente Vertreter wie Penicillin und Cyclosporin, wurde das erste pilzliche RiPP erst 2007 entdeckt. RiPPs haben mehrere Eigenschaften, die sie zu interessanten Zielen für Forschung und biotechnologische Anwendungen machen: Ihre Präkursor-Proteine sind genetisch kodiert und können gentechnisch verändert werden, sie werden von promiskuitiven Enzymen prozessiert und sie können durch verschiedene posttranslationale Modifikationen ein hohes Maß an Komplexität erreichen. In den letzten Jahren wurde der Forschung zu RiPPs ein weiteres Kapitel hinzugefügt, als entdeckt wurde, dass eine Klasse von Peptid-kodierenden Präkursor-proteinen, sogenannte «KEX2-processed repeat proteins» (KEPs), im Pilzreich weit verbreitet und ihre Funktionen weitgehend unbekannt sind. Das Ziel dieser Doktorarbeit war es, unser Verständnis von ribosomal produzierten Pilzpeptiden zu erweitern, mit besonderem Fokus auf Peptide, die aus KEPs hergestellt werden.

Kapitel 1 gibt eine Einführung in pilzliche RiPPs und ihre Biosynthese. Zunächst wird die RiPP-Familie der Amatoxine/Phallotoxine diskutiert, die berühmte tödliche Toxine umfasst, die z.B. im Grünen Knollenblätterpilz auftreten, gefolgt von der zuletzt entdeckten RiPP-Familie, den Borosinen, die N-Methylierungen im Peptid-Rückgrat tragen. Schließlich werden die Dikaridine untersucht. Dies sind funktionell diverse zyklische Peptide, die aus Präkursor-Proteinen mit einer KEP-ähnlichen Struktur prozessiert werden. Darüber hinaus wird das Potenzial von RiPP-Biosynthesewegen bei der Entwicklung von Peptidtherapeutika diskutiert, sowohl im Hinblick auf die Identifizierung neuer RiPPs als auch im Hinblick auf einen Ansatz der synthetischen Biologie, wo unter Verwendung biosynthetischer Enzyme neuartige Peptidbibliotheken erzeugt werden, die auf Bioaktivität gescreent werden können.

Kapitel 2 beschreibt die strukturelle und funktionelle Charakterisierung von Peptiden, die aus KEPs hergestellt werden, in den Agaricomyceten *Coprinopsis cinerea*, *Lentinula edodes*, *Pleurotus ostreatus* und *Pleurotus eryngii*. Genome Mining identifizierte mehrere KEPs in *C. cinerea*. Eine Auswahl dieser Proteine wurde erfolgreich in *Pichia pastoris* exprimiert und von der enzymatischen Maschinerie der Hefe verarbeitet, was zu einer Akkumulation von Peptiden im Kulturüberstand führte. Dieser experimentelle Aufbau wurde verwendet, um ein Protokoll für die Extraktion und den Nachweis von KEP-Peptiden in Pilzgeweben durch Massenspektrometrie zu erstellen. Unter Verwendung dieses Nachweisprotokolls wurden mehrere KEP-Peptide in Kulturüberständen und Gewebeproben von *C. cinerea*, *L. edodes*, *P. ostreatus* und *P. eryngii* bestätigt. Darüber hinaus haben wir ein CRISPR- Knockout-Protokoll für *C. cinerea* etabliert und Knockout-Stämme von sechs verschiedenen KEPs, drei KEX2-Homologen und einem KEX1-Homologen generiert. Wir fanden heraus, dass die *kep-*

Knockouts zu keinem erkennbaren Phänotyp führten, während die *kex*-Knockouts Defekte im Myzelwachstum und in der Fruchtkörperbildung aufzeigten. Diese Ergebnisse deuten darauf hin, dass die untersuchten KEP-Peptide keine organismus-intrinsische Rolle spielen, sondern vielmehr in der Interaktion mit der biotischen Umgebung funktionieren. Zudem zielen die KEX-Proteasen anscheinend auf eine große Vielzahl von Substraten ab, einschließlich einiger mit Rollen im Wachstum und der Differenzierung des Myzels.

Kapitel 3 stellt die neuen Genomsequenzen von vier Pilzarten vor: *Rhizopogon roseolus*, *Mariannaea elegans*, *Myrothecium verrucaria* und *Sphaerostilbella broomeana*. Die Genome wurden auf Präkursor-Proteine oder Gencluster gescreent, die an der Synthese der Peptid-Makrocyclen Mariannamide und der  $\alpha$ N-methylierten Makrocyclen Verrucamide und Broomeamide beteiligt sind. Wir suchten mit BLAST nach potenziellen RiPP-Präkursor-Proteinen dieser Peptide, und schlossen einen allgemeinen Screen für Borosin-produzierende Präkursor-Proteine ein. Während diese Suche erfolglos war, fanden wir in einem antiSMASH-Screening «non-ribosomal peptide synthase» (NRPS) Gencluster, die für die Biosynthese von Mariannamiden, Verrucamiden und Broomeamidn verantwortlich sein könnten. Darüber hinaus entdeckten wir mehrere Gencluster, die mutmaßliche Peptaibole codieren, eine antimikrobielle Klasse von Peptiden.

Abschließend fasst Kapitel 4 die Ergebnisse dieser Arbeit zusammen und gibt einen Ausblick auf zukünftige Experimente.



## Chapter 1

Introduction: Discovery of novel fungal RiPP biosynthetic pathways and their application for the development of peptide therapeutics

Eva Vogt<sup>1</sup> and Markus Künzler<sup>1</sup>

<sup>1</sup>ETH Zürich, Department of Biology, Institute of Microbiology, Vladimir-Prelog-Weg 4, CH-8093 Zürich, Switzerland

Published in:

Applied Microbiology and Biotechnology, 103, pages 5567–5581 (2019),  
<https://doi.org/10.1007/s00253-019-09893-x>

### Contributions:

- Figure preparation
- Writing, revisions

### **Abstract**

Bioactive peptide natural products are an important source of therapeutics. Prominent examples are the antibiotic penicillin and the immunosuppressant cyclosporin which are both produced by fungi and have revolutionized modern medicine. Peptide biosynthesis can occur either non-ribosomally via large enzymes referred to as non-ribosomal peptide synthetases (NRPS) or ribosomally. Ribosomal peptides are synthesized as part of a larger precursor peptide where they are posttranslationally modified and subsequently proteolytically released. Such peptide natural products are referred to as ribosomally synthesized and posttranslationally modified peptides (RiPPs). Their biosynthetic pathways have recently received a lot of attention, both from a basic and applied research point of view, due to the discoveries of several novel posttranslational modifications of the peptide backbone. Some of these modifications were so far only known from NRPSs and significantly increase the chemical space covered by this class of peptide natural products. Latter feature, in combination with the promiscuity of the modifying enzymes and the genetic encoding of the peptide sequence, makes RiPP biosynthetic pathways attractive for synthetic biology approaches to identify novel peptide therapeutics via screening of *de novo* generated peptide libraries and, thus, exploit bioactive peptide natural products beyond their direct use as therapeutics. This review focuses on the recent discovery and characterization of novel RiPP biosynthetic pathways in fungi and their possible application for the development of novel peptide therapeutics.

### **Introduction**

Bioactive natural products have been of great interest for therapeutic health care throughout human history. While many natural products have been successfully used for thousands of years, only fairly recent advancements have allowed elucidating their structure, biosynthesis and mode of action (Dias et al. 2012). Meanwhile, the search for novel natural products with pharmacological activities continues. High-throughput screens using molecular and cellular bioassays allow the evaluation of large numbers of natural products and synthetically generated compound libraries for bioactivities of interest. Statistical analysis of the complementarity of natural products and synthetic compounds showed that 40% of chemical scaffolds of natural products are absent in synthetic compound libraries (Henkel et al. 1999), demonstrating that the biomedical relevance of natural products lies in their wide range of structural diversity and complexity. Thus, natural products continue to play a significant role in drug development. From the 1940s to 2014, 40% of small molecules approved for cancer treatment were natural products or derivatives thereof (Newman and Cragg 2016). Plants and microorganisms are the main sources of bioactive natural products (Dias et al. 2012). The vast dimension of 'microbial dark matter' suggested by genomic microbiome studies promises an enormous variety of bioactive natural products yet to be discovered (Solden et al. 2016). This is also true for fungi where it is assumed that only about one-twentieth of all existing fungi have been described and an even smaller fraction has been cultured successfully in the lab (Jiang and An 2000; Hawksworth and Lücking 2017). In addition to the direct use of natural products for pharmacological purposes, their chemical structures might serve as an inspiration

## Chapter 1

for structural motifs in synthetic chemistry, and the knowledge about the biosynthetic pathways leading to distinct natural products might allow to biotechnologically produce libraries of new-to-nature compounds for screening (Jiang and An 2000; Boecker et al. 2016; Bozhüyük et al. 2018).

Natural products can be divided into different structural classes including alkaloids, terpenoids, polyketides and peptides (Cragg and Newman 2013). Among these classes, peptides cover the largest chemical space due to their polymeric nature and the large number and variety of building blocks (20 in case of proteinogenic amino acids). Thus, a large fraction of bioactive natural products used as therapeutics are peptides (Lau and Dunn 2018). As prominent examples, the discoveries of two fungal peptide natural products, the antibiotic penicillin in the late 1920s and the immunosuppressant cyclosporin in the early 1970s, have revolutionized modern medicine (Bills and Gloer 2016). Therapeutic peptides are recognized to be efficient and selective while being, at the same time, well tolerated and safe (Fosgerau and Hoffmann 2015). On the other hand, peptides are comparatively large molecules with usually poor chemical and physical properties: They have a short life span due to degradation by proteases and a limited membrane permeability. Their binding specificity is favorable compared to typical small-molecule drugs, but off-target effects e.g. due to adoption of different conformations can still pose a risk (Bruno et al. 2013). The pharmacological applicability of peptide natural products depends on their properties, e.g. their serum stability and oral availability, which are influenced by modifications of the peptide side chains and backbone (Beutler 2009).

The biosynthesis of peptides can occur either non-ribosomally or ribosomally. Non-ribosomal biosynthesis is mediated by so-called non-ribosomal peptide synthetases (NRPSs), large modular enzymes that accept also non-proteinogenic amino acids and often contain amino acid-modifying modules (Süssmuth and Mainz 2017). Non-ribosomally produced peptide natural products include above mentioned examples (penicillin and cyclosporin) from fungi and many examples from bacteria e.g. vancomycin and actinomycin. Conversely, ribosomal biosynthesis involves the translation of a precursor peptide which is then posttranslationally processed to the mature peptide (Fig. 1a). Since the discovery of the first family of such ribosomally synthesized and post-translationally modified peptides (RiPP) in 1988, over 20 additional families from all domains of life have been discovered to date (Arnison et al. 2013). The interest in this class of natural products is increasing both from a basic and applied research point of view due to two main reasons: First, the amino acid sequence of RiPPs is genetically encoded and thus, the peptide-encoding gene (including the usually proximal genes for the modifying enzymes) can be easily identified in genome sequences and also modified in order to change the sequence of the peptide (McIntosh et al. 2009). Second, some of the recently identified posttranslational modifications of RiPPs were previously regarded as hallmarks of non-ribosomally produced peptides, affect the peptide backbone and significantly improve the pharmacological properties of these peptides e.g. their proteolytic stability and membrane permeability (Freeman et al. 2012; Van Der Velden et al. 2017; Müller 2018). These features, together with the promiscuity of the RiPP modifying enzymes, may

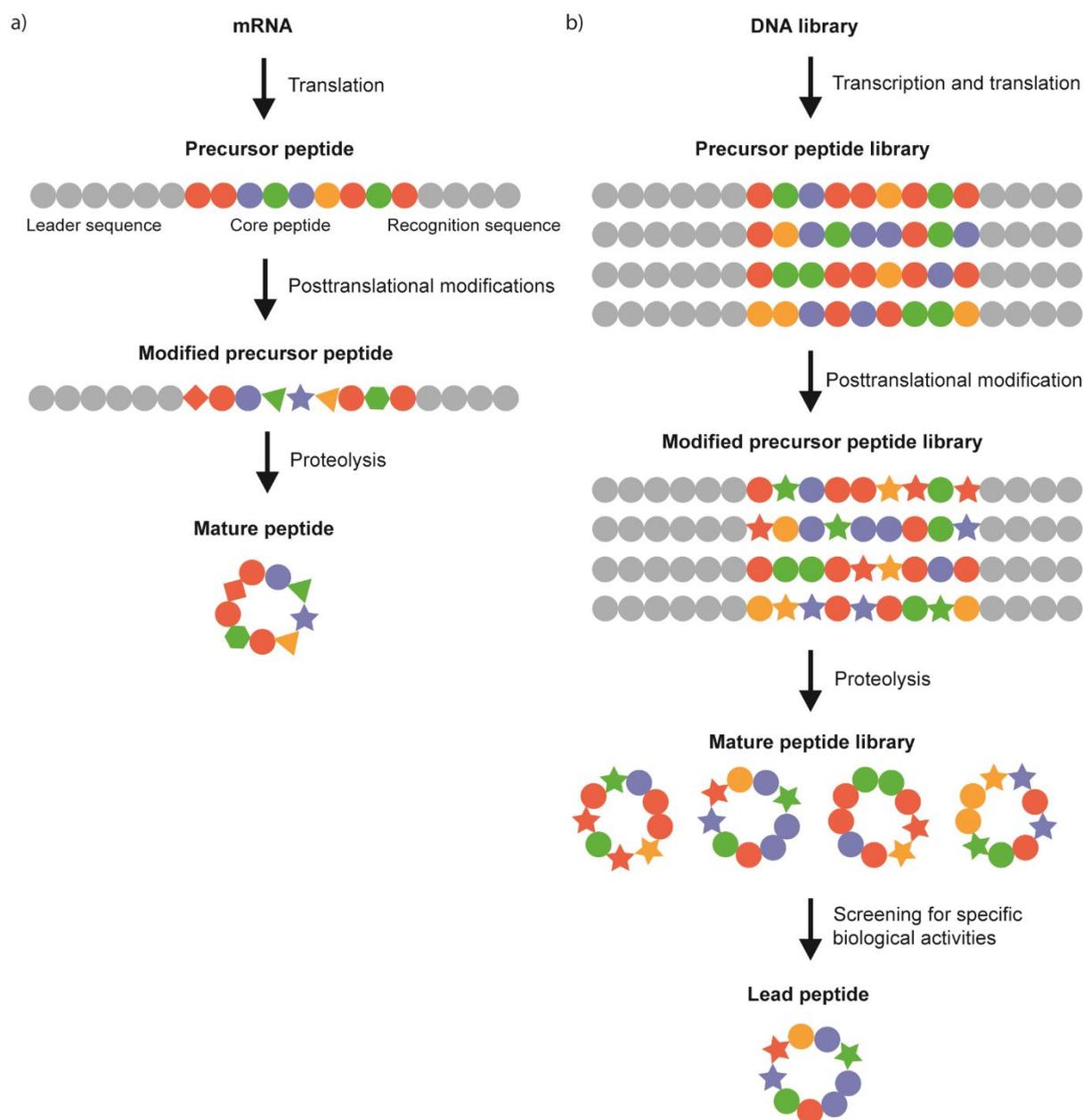
allow to exploit these pathways for the biotechnological production of new-to-nature libraries of peptides with advantageous pharmacological properties.

This history of fungal RiPPs is very recent with the first family discovered only in 2007 (Hallen et al. 2007). In this review, we describe the commonalities and the differences between these RiPP biosynthetic pathways and discuss their potential for the development of novel peptide therapeutics, having the successful history of non-ribosomal fungal peptide natural products in mind.

### **General biosynthetic pathway of RiPPs**

RiPPs share several features with regard to their biosynthetic pathways. Usually, these pathways are encoded in the genome by gene clusters, where one gene encodes the peptide precursor while neighboring genes encode enzymes that will modify the precursor (Yang and Van Der Donk 2013). In RiPPs, common posttranslational modifications are single residue modifications such as methylation, hydroxylation, acetylation, epimerization and glycosylation, as well as crosslinks between residues via disulfide, lanthionine, thiazole and oxazole bridges (Arnison et al. 2013). The peptide precursor is usually 20-110 residues long and composed of a N-terminal leader sequence and the C-terminal core peptide (Yang and Van Der Donk 2013). The core peptide will undergo various posttranslational modifications followed by proteolytic cleavage of the RiPP from the leader sequence. In some cases the core peptide is followed by a recognition sequence at the C-terminus which is relevant in excision and cyclization of the core peptide (Koehnke et al. 2012; Luo et al. 2014). In eukaryotes, a signal sequence preceding the leader peptide can regulate the transport of the precursor to cellular compartments.

The presence of the leader sequence is usually required for the core peptide modifications. Deletion of the leader sequence resulted in a reduction or absence of core peptide modifications (Ortega and Van Der Donk 2016). However, the exact function of the leader sequence is still unclear and might differ between different peptide precursors. It has been suggested that the leader sequence supports the posttranslational modification steps as a recognition motif for the modifying enzymes and assists the export of the core peptide from the producing cell as a non-classical secretory signal. In addition, the leader sequence probably plays a role in the folding, stabilization and activation of the bound modifying enzyme (Ortega and Van Der Donk 2016). For such a role, the leader sequence could either induce a conformational shift in the enzyme and thereby activate enzyme activity, or it could trap enzymes in a specific, active conformation (Ortega and Van Der Donk 2016).



**Figure 1.** Schematic illustration of the general RiPP biosynthesis pathway (a) and its biotechnological application for the identification of leads for novel peptide drugs via generation and screening of new-to-nature peptide libraries (b). In RiPP biosynthesis, the precursor is ribosomally synthesized and consists of a leader sequence and the core peptide. Depending on the molecule, the precursor can also contain an N-terminal signal sequence or a C-terminal recognition sequence. The core peptide within the precursor is modified by enzymes that are encoded by large gene clusters comprising also the precursor gene. Afterwards, the peptide is proteolytically cleaved, optionally followed by additional modifications, resulting in the bioactive, mature form. Figure is based on image by Arnison *et al.* (Arnison *et al.* 2013). RiPP biosynthesis pathways may be applied for the de novo generation and screening of modified peptide libraries. A DNA library would serve as source for the translation of a large variety of precursor peptides with unique core peptides. Each core peptide would be modified individually and cleaved from the precursor, resulting in a library of new-to-nature mature peptides that display a unique combination of different residues and modifications. The mature peptides would be screened for specific bioactivities and hits would serve as lead peptides for the development of novel peptide drugs.

The importance of the leader sequence was demonstrated in an experiment where the core peptide sequences of precursors were exchanged with foreign sequences. The foreign core peptides were still posttranslationally modified suggesting that the modifying enzymes only recognized the leader and not the core peptide sequence (Ortega and Van Der Donk 2016). This promiscuity of the modifying enzymes towards the modified core sequence is regarded as a hallmark of RiPP biosynthetic pathways. In addition, posttranslational modifications of the core peptide usually proceed in a processive manner from the N- to the C-terminus (Kelleher et al. 1999; Lee et al. 2009; Lubelski et al. 2009) with some instances proceeding in the opposite direction (Krawczyk et al. 2012; Melby et al. 2012). The processivity of these modifications indicates that the intermediates are kept bound to the modifying enzyme after every reaction (Ortega and Van Der Donk 2016).

### Fungal RiPPs

Fungi have been shown to be a valuable source of peptide natural products exemplified by (non-ribosomally produced) penicillin and cyclosporin (Bills and Gloer 2016). However, to date, only four families of RiPPs have been found in fungi: amatoxins/phallotoxins (Hallen et al. 2007), borosins (Van Der Velden et al. 2017), dikaritins\* (Umemura et al. 2014; Nagano et al. 2016; Ding et al. 2016) and epichloëcyclins (Fig. 2) (Johnson et al. 2015). See Table 1 for a summary of all known fungal RiPP families, their representatives and their activities.

#### *Amatoxins/Phallotoxins*

Amatoxins and phallotoxins are structurally similar peptides synthesized by several species of the basidiomycete genera *Amanita*, *Galerina*, *Lepiota* and *Conocybe* (Hallen et al. 2007; Luo et al. 2009; Walton et al. 2010; Luo et al. 2012; Sgambelluri et al. 2014). Amatoxins, exemplified by  $\alpha$ -amanitin, are highly toxic to insects, nematodes, and mammals and many of the amatoxin-synthesizing fungi, such as the death cap (*Amanita phalloides*), are infamous for causing fatal mushroom poisonings. Amatoxins selectively inhibit RNA polymerase II which leads to a stop of mRNA transcription, a halt of cell metabolism and ultimately to cell death (Chafin et al. 1995; Rudd and Luse 1996; Bushnell et al. 2002). In spite of their structural similarity, phallotoxins, exemplified by phalloidin, rarely exert toxicity due to their poor absorption in the gut (Hallen et al. 2007). If injected in the bloodstream, they inhibit actin polymerization primarily in liver cells which results in cell membrane dysfunction and severe liver damage (Lengsfeld et al. 1974). The ecological function of these peptides is not clear but based on their toxicity and their induction upon initiation of fruiting body formation it is assumed that they protect the early stage fruiting bodies from fungivores (Spiteller 2015; Zhang et al. 2018).

Amatoxins and phallotoxins consist of eight respectively seven amino acid residues, forming a bicyclic peptide (Fig. 2a) (Walton et al. 2010). The bicyclic structure is established through a crosslink between a cysteine and a tryptophan residue, referred to as tryptathionine (May and Perrin 2007). Nine amatoxins and seven phallotoxins have been identified so far:  $\alpha$ -/ $\beta$ -/ $\gamma$ -/ $\epsilon$ -amanitin, amanullin, amanullinic acid, amaninamide, amanin and proamanullin, respectively phalloidin, phalloidin, prophalloin, phalloin, phallisin, phallacin and phallisacin (Table 1)

## Chapter 1

(Wieland 1987; Baumann et al. 1993). The variants share the same peptide backbone (with variations at one or two positions) and differ in hydroxylations at different positions. Phallotoxins contain one residue in D-configuration, which can be either D-hydroxy-Asp or D-Thr (Hallen et al. 2007).

A number of closely related, monocyclic peptides have been found in *Amanita* that seem to belong to the RiPP family of amatoxins/phallotoxins. The so-called virotoxins, consisting of the peptides alaviroidin, viroisin, deoxoviroisin, viroidin and deoxoviroidin, were isolated from *A. phalloides* (Table 1) (Baumann et al. 1993; Vetter 1998). Additionally, a monocyclic nonapeptide termed amanexitide of unknown function was recently identified in *A. exitialis* from China (Xue et al. 2011). Its structure is related to antamanide, a monocyclic decapeptide with antidote activity against amatoxins and phallotoxins previously identified in *A. phalloides* (Wieland et al. 1978). Antamanide was recently shown to be immunosuppressive and bind, similar to cyclosporin A, the peptidyl prolyl cis-trans isomerase cyclophilin D (Siemion et al. 1992; Azzolin et al. 2011). Immunosuppressive activity has also been reported for several monocyclic hepta- and octapeptides, referred to as cycloamanides A-F, identified from different *Amanita phalloides* specimens (Gauhe and Wieland 1977; Wieczorek et al. 1993; Pulman et al. 2016).

In 2007,  $\alpha$ -amanitin and phallacidin were found to be derived from ribosomally produced peptide precursors encoded in the genome of *A. bisporigera* (destroying angel) and to represent the first fungal RiPP family (Hallen et al. 2007). The peptide precursors consist of 35 respectively 34 amino acid residues and comprise a 10-residue leader sequence, the 8/7-residue core peptide and a 17-residue C-terminal recognition sequence (Fig. 2b) (Hallen et al. 2007). Amatoxin and phalloxin precursors do not contain a signal sequence for classical secretion and are thus synthesized and processed in the cytoplasm, eventually close to the vacuole (Luo et al. 2010). Analysis of the draft genome sequences of *A. bisporigera* and *A. phalloides* suggests that the two species code together for over 50 unique small, cyclic peptides of this family (Pulman et al. 2016). The encoded peptide precursors are very conserved in the leader and C-terminal recognition sequences but vary significantly in the length and the sequence of the core peptide. Due to a conserved MSDIN motif in the leader sequence, these peptides are also referred to as MSDIN family of peptides.



## Chapter 1

with recognition sequences for the kexin Golgi protease in bold. The schematic representation of the Omphalotin A-I precursor is based on image by Van der Velden *et al.* (Van Der Velden et al. 2017). The protein accession numbers for the peptide precursors are A8W7M4 for  $\alpha$ -amanitin, A8W7M7 for phalloidin, XP\_002381318 for ustiloxins, AMR44282 for phomopsins, XP\_002377602 for asperipin-2a and KP797979 for epichloëcyclins. The JGI protein ID of the omphalotin precursor is 2087.

Remarkably, the two different *Amanita* species share only three peptides,  $\alpha$ -amanitin, phalloidin and an uncharacterized MSDIN peptide, demonstrating the sequence and species diversity within this RiPP family. Based on the transcriptomes of the two *Amanita* species, it can be assumed that many of the detected precursor genes are expressed. Accordingly, two novel monocyclic, but otherwise not posttranslationally modified hepta- and nonapeptides, referred to as cycloamanide E and F, could be identified from *A. phalloides* based on the core peptide sequences of the respective peptide precursors. It should be noted, however, that this RiPP family is only present in species of *Amanita* sect. *Phalloideae* and not in other, more distantly related *Amanita* species. Similarly, the genome sequences of the distantly related mushrooms *Galerina marginata* and *Lepiota brunneoincarnata* contain only two precursor genes each, both coding for  $\alpha$ -amanitin and no extended MSDIN-like peptide family (Luo et al. 2012; Luo et al. 2018). The leader and C-terminal recognition sequences of these precursors are of the same length as the ones of the different *Amanita* precursors but the individual residues are less conserved.

Regarding the modifying enzymes of the amatoxin/phallotoxin RiPP family, thus far, a prolyloligopeptidase (POP) of serine protease family S9a responsible for the dual function of core peptide release and macrocyclization has been identified and characterized both biochemically and structurally (Luo et al. 2014; Czekster et al. 2017; Czekster and Naismith 2017). The POP-encoding gene clusters with one of the two  $\alpha$ -amanitin precursor genes of *G. marginata* was shown to be required for  $\alpha$ -amanitin production in this mushroom (Luo et al. 2014). Accordingly, homologs of this gene are found in the genomes of all amatoxin- and phallotoxin-producing basidiomycetes and the gene product colocalizes with  $\alpha$ -amanitin in the same cells of *Amanita* fruiting bodies (Luo et al. 2010). The POP involved in the amatoxin/phallotoxin biosynthetic pathway (POPB) may have evolved from a related, housekeeping POP (POPA) that is found in all basidiomycetes (Zhang et al. 2005; Luo et al. 2010). *In vitro* studies with recombinant *G. marginata* POPB demonstrated that the enzyme processes the peptide precursor in two consecutive steps: In a first step, it removes the leader sequence of the 35-residue  $\alpha$ -amanitin precursor by proteolytic cleavage after the terminal and conserved Pro-residue of the leader sequence; in a second step, the enzyme performs a transpeptidation reaction between the generated N-terminus of the shortened peptide precursor and the second conserved Pro-residue at the end of the core peptide, thereby releasing the C-terminal recognition sequence (Luo et al. 2014). Based on kinetic studies and substrate-bound structures of the enzyme, it was proposed that the enzyme recognizes the original peptide precursor via its leader sequence, releases both the leader sequence and the truncated precursor upon the first proteolytic cleavage and rebinds the truncated precursor to catalyze the transpeptidation (Czekster et al. 2017; Czekster and Naismith 2017). In accordance with such a mechanism, the leader sequence of the peptide precursor was shown

to be required for the first proteolytic cleavage and the C-terminal recognition sequence for the second proteolytic cleavage and the transpeptidation (Luo et al. 2014; Czekster et al. 2017). None of the enzymes catalyzing any of the other known modifications of members of this RiPP family, including the formation of the tryptathionine crossbridge and the oxidation, L-to-D-epimerization or hydroxylation of specific residues, has been identified yet. Hence, it is not clear whether these enzymes act at the level of the peptide precursor or on the released and cyclized core peptide. At least *in vitro*, the POP does not seem to depend on any additional modification of the core peptide for its activity which rather argues for latter option.

### *Borosins*

The fungal RiPP family of borosins was recognized only in 2017 (Ramm et al. 2017; Van Der Velden et al. 2017). Omphalotin A, the founding member of this RiPP family, was isolated already in 1997 and 1998 from the basidiomycete *Omphalotus olearius* in a screen for natural products from submerged cultures of asco- and basidiomycetes exerting toxicity against the plant parasitic nematode *Meloidogyne incognita* (Mayer et al. 1997; Sterner et al. 1997; Büchel et al. 1998; Liermann et al. 2009). Further analysis revealed that this toxicity was higher than the one of the bacterial nematotoxin ivermectin and very specific for this nematode species with very low or no toxicity towards mammalian cells, plants, insects, other nematode species, fungi and bacteria (Mayer et al. 1997). Omphalotin A is a cyclic dodecapeptide with backbone N-methylations on 9 of the 12 amino acid residues (Fig. 2a) (Mayer et al. 1997; Sterner et al. 1997). The peptide consists mainly of hydrophobic amino acids and all residues are in L-configuration (Büchel et al. 1998). Eight other derivatives of omphalotin called omphalotin B-I, varying in additional modifications i.e. hydroxylations, acetylations and an unusual tryptophan modification, have been discovered (Table 1) (Büchel et al. 1998; Liermann et al. 2009). Similar to the amatoxins/phallotoxins, omphalotins are not secreted and have to be extracted from *O. olearius* cells (Mayer et al. 1997). In contrast to the amatoxins/phallotoxins, however, omphalotins are exclusively produced in the vegetative mycelium of *O. olearius* and not in its fruiting bodies (Anke and Schüffler 2018). Accordingly, submerged cultures of monokaryotic strains of *O. olearius* are efficient producers of omphalotins (Liermann et al. 2009).

Backbone N-methylation was long thought to be an exclusive modification of non-ribosomally produced peptides exemplified by the immunosuppressant cyclosporin A (Dreyfuss et al. 1976), the antibiotic actinomycin S (Hollstein 1974) and several depsipeptides (Maharani et al. 2015). In all these cases, the methylation of the nitrogen occurs on the level of the primary amine before formation of the peptide bond which makes sense given the low reactivity of a secondary amine in particular in the context of the delocalization of its free electron pair in a peptide bond (Scherer et al. 1998). Hence, it came as a surprise when omphalotin A was shown to be genetically encoded and derived from the ribosomally produced precursor OphMA\*\* (Ramm et al. 2017; Van Der Velden et al. 2017). This peptide precursor is, with 417 residues, unusually long and its N-terminus (residues 12 to 219) shows sequence homology to SAM-dependent methyltransferases (Fig. 2b) (Ramm et al. 2017; Van Der Velden et al. 2017). This domain was shown to be responsible for the backbone N-methylation of the omphalotin core

## Chapter 1

peptide located at the C-terminus of the OphMA precursor (Ramm et al. 2017; Van Der Velden et al. 2017). The modified residues of the OphMA core peptide were identical to the ones of the natural product omphalotin A demonstrating that omphalotin A was a RiPP and that enzymatic backbone N-methylation of the nitrogen in preformed peptide bonds is possible (Van Der Velden et al. 2017). The omphalotin precursor is so far the only example where a RiPP modifying enzyme is covalently coupled to the modified core peptide. An enzymatic domain was reported for the precursors of a bacterial RiPP family but this domain appears to function rather in export than in modification of the core peptide and is not catalytically active (Haft et al. 2010). In analogy to the automethylating activity of the OphMA precursor, the name "borosins" was proposed for this novel fungal RiPP family, after the ancient mythological symbol Ouroboros depicting a serpent biting its own tail (Van Der Velden et al. 2017).

Structural analysis of OphMA has revealed that the precursor forms a dimer with an unusual catenane-like arrangement in which the methyltransferase domain of one protomer methylates the core peptide of the other protomer (Ongpipattanakul and Nair 2018; Song et al. 2018). Analysis of the methylation states of reaction intermediates and the structure of the enzyme suggest an overall mechanism in which the core peptide is coordinated into and processively pulled, from the N- to the C-terminus, through a substrate tunnel which is formed at the interface between the two protomers and contains the active site (Van Der Velden et al. 2017; Song et al. 2018). The structure of the active site and computational reaction modelling suggest a base-catalyzed reaction mechanism that involves a protein-mediated clamping of the amide nitrogen towards the methyl group of SAM and stabilization of the reaction intermediate (Ongpipattanakul and Nair 2018; Song et al. 2018). Methylation analysis of hybrid precursors where the core peptide of OphMA was replaced by unrelated sequences revealed a considerable promiscuity of the enzyme with regard to its substrate (Van Der Velden et al. 2017). These results also suggested a minor role of the six-residue C-terminal recognition sequence in methylation as methylation also occurred when this sequence was deleted - although the last two positions of the core peptide remain unmethylated in this case (Van Der Velden et al. 2017). In accordance with this result, at least in case of the *D. bispora* OphMA homolog, no binding of synthetic peptides corresponding to the C-terminal recognition sequence to the OphMA methyltransferase domain could be detected (Ongpipattanakul and Nair 2018). In addition, recombinant full length OphMA methylates 1-2 residues in the C-terminal recognition sequence beyond the core peptide (Ramm et al. 2017; Van Der Velden et al. 2017). All these results suggest that the C-terminal recognition sequence has a role in the completion of the methylation of the core peptide but it is not clear whether this effect is sequence-specific.

The *ophMA* gene is part of a gene cluster of *O. olearius* where neighbouring genes are likely responsible for the release, macrocyclization and additional modifications of the core peptide (Ramm et al. 2017; Van Der Velden et al. 2017). Thus far, only the role of one of these genes has been elucidated: *ophP* codes for a prolyl oligopeptidase of the same serine protease family as POPB, that is involved in the amatoxin/phallotoxin biosynthetic pathway, and its coexpression with *ophMA* in the yeast *Pichia pastoris* lead to formation of both linear and

cyclic forms of omphalotin A (Ramm et al. 2017). These results suggest that OphP is sufficient for the release and the macrocyclization of the methylated core peptide similar to the role of POPB in amatoxin/phallotoxin biosynthesis. Accordingly, the residue preceding the core peptide is a proline but the terminal residue is a glycine. So far, the specificity of OphP with regard to the sequence, methylation state or flanking residues of the core peptide has not been addressed experimentally.

BLAST searches revealed homologous gene clusters in the basidiomycetes *Dendrothele bispora* and *Lentinula edodes* (Ramm et al. 2017; Van Der Velden et al. 2017; Ongpipattanakul and Nair 2018). The coding region of the *D. bispora* OphMA precursor was expressed in *Escherichia coli* and the core peptide of the recombinant protein was shown to be methylated in the same pattern as OphMA from *O. olearius* despite three sequence variations each therein (Van Der Velden et al. 2017; Ongpipattanakul and Nair 2018). Additional homologs of the *ophMA* precursor gene were found in genomes of many other basidiomycetes suggesting that borosins are widespread in this fungal phylum (Van Der Velden et al. 2017). The genome coding for the highest number of predicted borosin precursors is thus far the one of the ectomycorrhizal fungus *Rhizopogon vinicolor* (unpublished results) The C-termini of all these predicted borosin precursors differ considerably from the one of *O. olearius*, *D. bispora* and *L. edodes* but are generally also rich in hydrophobic residues. In none of these fungi, backbone N-methylated peptides have been described so far. On the other hand, such peptides have been described for several fungi for which no genome sequences are available at the moment. Examples are the dictyonamides, linear dodecapeptides from an unknown marine-derived fungus (Komatsu et al. 2001), the RHM family of linear octapeptides from a sponge-derived ascomycota (Boot et al. 2006), the verrucamides from the plant-pathogenic ascomycete *Myrothecium verrucaria* (Zou et al. 2011) and the gymnopeptides, cyclic octadecapeptides from the wood-degrading basidiomycete *Gymnopus fusipes* (Ványolós et al. 2016). It can be assumed that at least some of these peptides will turn out to be new members of the borosin RiPP family.

### *Dikaritins*

In contrast to the cytoplasmic localization of the amatoxin/phallotoxin and borosin precursors and their processing proteases, the precursors of the other known fungal RiPP families are secreted and proteolytically processed by the Golgi-localized endopeptidase kexin that is present in most fungi (Yoshimi et al. 2016; Li et al. 2017). The name of the dikaritin RiPP family was proposed based on the production of these cyclopeptides by the fungal subkingdom Dikarya comprising the phyla Ascomycota and Basidiomycota (Ding et al. 2016). Hitherto characterized examples include the ustiloxins, phomopsins and asperipins originally identified in the ascomycetes *Ustilaginoidea virens*, *Phomopsis leptostromiformis* and *Aspergillus flavus*, respectively (Tsukui et al. 2014; Umemura et al. 2014; Nagano et al. 2016; Ding et al. 2016). A common feature of all dikaritins is their cyclization via ether bridges between the hydroxyl group of tyrosine and the C $\beta$  of isoleucine (ustiloxins and phomopsins) or phenylalanine/tyrosine residues (asperipins) (Fig. 2a). This special type of peptide cyclization is catalyzed by DUF3328-containing oxidases encoded by all dikaritin-encoding gene clusters

(Ye et al. 2016; Ye et al. 2019). In case of ustiloxins and phomopsins, the cyclizing hydroxyl group of the tyrosine residue is provided by a tyrosinase. Other modifications besides cyclization include methylation of the N-terminus (ustiloxin and phomopsin), addition of the non-proteinogenic amino acid norvaline through a sulfoxide bond (ustiloxin), hydroxylation (ustiloxin) and dehydrogenation of side chains (phomopsin).

Ustiloxins are mycotoxins produced by the plant pathogen *U. virens*. The fungus infects the spikelets of rice, leading to the so-called "rice false smut disease" which is known as one of the most destructive fungal diseases in rice (Koiso et al. 1992; Koiso et al. 1994; Koiso et al. 1998; Wang et al. 2017). It was shown that ustiloxins cause phytotoxic effects in rice, wheat and maize, such as growth reduction of radicles and plumules (Koiso et al. 1998; Abbas et al. 2014). Ustiloxins inhibit the polymerization of tubulin and suppress mitosis, causing a range of symptoms in domestic animals that were fed with contaminated rice. Observed symptoms were diarrhea, hemorrhage, poor growth, abortion and organ damage (Wang et al. 2017). Until today, contamination of food crops with ustiloxins remains a serious safety risk and financial burden (Wang et al. 2017). Detection and analysis of ustiloxin production in the mold *A. flavus* revealed that these peptide natural products are RiPPs (Umemura et al. 2014; Tsukui et al. 2014). The cyclic tetrapeptides of the sequence "Y-V/A-I-G" are modified with methyl groups, hydroxyl groups and norvaline (Fig. 2a) (Umemura et al. 2014). Seven different ustiloxins have been discovered, called ustiloxin A-G (Table 1) (Koiso et al. 1992; Koiso et al. 1998; Ding et al. 2016). All *A. flavus* ustiloxin variants are derived from precursor *ustA*, consisting of a signal peptide for classical secretion and a 16-fold repeat of the core peptide "YAIG" (Fig. 2b). Each core peptide is flanked by a leader peptide at the N-terminus and a recognition sequence at the C-terminus. These peptide units are interspersed with dibasic residues "KR" which are recognized by the Golgi endoprotease kexin (Umemura et al. 2014). The *ustA*-encoding gene is part of a gene cluster that encodes most of the enzymes that are needed for biosynthesis of ustiloxin B (Umemura et al. 2014).

As a result of a detailed genetic and biochemical analysis, the biosynthetic pathway could be established as follows: Upon kexin-mediated cleavage of the precursor protein UstA, the tetrapeptides YVIG and YAIG are cyclized by the tyrosinase-homolog UstQ and the DUF3328-containing proteins UstYa/Yb. Upon N-methylation of the terminal amino group by the methyltransferase UstM, UstC (a cytochrome P450) probably installs a cysteine residue by oxidative substitution. After the substitution, UstF1 (a Flavin-containing monooxygenase [FMO]) catalyzes oxygenation of the sulfur atom to a sulfoxide, and the subsequent N-hydroxylation by another FMO, UstF2, results in an aldoxime. Finally, UstD, a PLP-dependent enzyme homologous to cysteine desulfurase, catalyzes condensation of a C3 unit derived from aspartate (Ye et al. 2016). Although Kexin plays a vital role in the ustiloxin biosynthesis by cleaving the repeating units after the second basic residue, it is not encoded by the cluster but elsewhere in the genome (Umemura et al. 2014; Yoshimi et al. 2016). Analysis of biosynthetic intermediates accumulating in gene deletion mutants allowed the establishment of the biosynthetic pathway (Ye et al. 2016). However, the subcellular localization of the various

reactions remains, with the exception of the kexin-mediated endoproteolytic cleavage of the precursor in the Golgi, unclear.

Phomopsins are mycotoxins produced by a number of legume-infesting ascomycetes including *Phomopsis leptostromiformis* which causes stem blight in lupins (Battilani et al. 2011). Similar to ustiloxins, phomopsins bind tubulin and prevent it from polymerization, leading to inhibition of spindle formation during mitosis and mitotic arrest (Battilani et al. 2011). Affected are mainly hepatocytes, as seen in the occurrence of liver damage in livestock fed with phomopsin-contaminated crops (Battilani et al. 2011). Inspired by the related structure (Fig. 2a) and the ribosomal biosynthesis of ustiloxins, Ding *et al.* recently established that phomopsins are derived from the ribosomally produced precursor *phomA* in *P. leptostromiformis* (Ding et al. 2016). This precursor contains, similar to the ustiloxin precursor *ustA*, a signal peptide followed by eight repeats of the phomopsin core peptide "Y-V-I-P-I/F-D" (Fig. 2b). The core peptide in each repeating sequence is flanked by an N-terminal leader sequence and the C-terminal dibasic residue "K-K/R". The gene coding for the precursor PhomA is part of the gene cluster *phom* encoding 18 modifying enzymes (Ding et al. 2016). These enzymes include peptidases that might cleave the leader and the dibasic residues from the core peptide, a tyrosinase that catalyzes the cyclization of the peptide scaffold possibly together with several DUF3328-containing oxidases, a SAM-dependent methyltransferase that methylates the amino-terminus and several enzymes with unknown functions (Ding et al. 2016). Gene clusters that generate phomopsin-like cyclic peptides are widespread in fungal genomes, including plant-pathogenic, animal parasitic and saprobic ascomycetes but also basidiomycetes (Ding et al. 2016). Based on these biocomputational findings, the name dikaritins was suggested for this RiPP family.

At the time of its discovery, the structure of the precursor peptide *UstA* with its 16-fold repeat of the core tetrapeptide sequence seemed rather unique, since RiPP precursors usually contain only one single core peptide sequence (Umemura et al. 2014). However, the previous discovery of cyanobactins in bacteria (Gu et al. 2018) and the subsequent discoveries of the phomopsins and the epichloëcyclin RiPP family demonstrate that RiPP precursors with repetitive peptide sequences are more common than previously thought (Fig. 2b). The acquisition of several repeats of the core sequence was suggested to increase the production of the peptide natural product and allow diversification of the peptide sequence in terms of its bioactivity. The presence of highly repetitive core peptides also simplifies the biocomputational detection of new RiPP precursor genes with different core peptide sequences. An algorithmic approach for the identification of secreted, kexin-cleavable and highly repetitive precursor genes clustering with DUF3328-encoding genes in the *Aspergilli* genomic database (AspGD) (Cerqueira et al. 2014) suggested an average of 70 candidate precursors per *Aspergillus* genome and 94 candidate gene clusters coding for novel dikaritins\* (Nagano et al. 2016). Gene deletions in one of these clusters in *A. flavus* lead to the identification of asperipin-2a, a bicyclic hexapeptide of the sequence "FYITGY" (Fig. 2). In contrast to ustiloxins and phomopsins, the cyclization of asperipin occurs via the para-hydroxyl group of two tyrosine residues and is mediated by a single DUF3328-containing

oxidase encoded in the respective gene cluster without the involvement of a tyrosinase (Ye et al. 2019). Although the biological activity of asperipin-2a has not been characterized yet, the results of these studies demonstrate that in particular ascomycete genomes encode a large variety of dikaritins some of which are likely to have other targets than ustiloxins and phomopsins.

### *Epichloëcyclins*

Epichloëcyclins are cyclic nonapeptides produced by ascomycetes of the genus *Epichloë* (Johnson et al. 2015). These fungi form a symbiotic relationship with certain grasses by growing in the intercellular space of stems, leaves, inflorescences and seeds. The infected grasses benefit in several ways from the fungal endophyte: Fungal alkaloids contribute to the plant defense against herbivores, while other products secreted by the fungus increase the plants stress resistance and growth (Johnson et al. 2003; Khan et al. 2010; Ambrose and Belanger 2012). One of the most abundant fungal transcripts in infected plants is *gigA* (Johnson et al. 2015). *gigA* encodes a precursor that consists of a signal sequence, followed by four repeats of 27 amino acids. Each repeat contains a dibasic "KR" motif which is highly conserved between species and probably cleaved by the two kexin proteases of *Epichloë*, *kexA* and *kexB*. The repeating units in *gigA* are not identical but slightly variable. Accordingly, metabolic analysis of *gigA*-expressing *Epichloë* showed that six different Epichloëcyclins (A-F) are produced, corresponding to a slightly variable core nonapeptide within the repeating units of the *gigA* gene (Table 1) (Johnson et al. 2015). These core peptides are cyclized by a bridge between the conserved tyrosine residue at position 7 and the aminoterminal proline or isoleucine residue and are in addition modified by a dimethylation of the conserved lysine residue at position 4. Since the chemical nature of the cyclizing bridge has not been characterized yet and no information about the composition of the respective gene cluster is available, epichloëcyclins are at the moment categorized as separate RiPP family and not as additional representatives of dikaritins. The high expression of *gigA* during infection suggests a role of the epichloëcyclins in the symbiosis with grass. However, the biological function of epichloëcyclins remains unknown (Johnson et al. 2015).

### *Beyond RiPPs: Peptides derived from Kex2-processed repeat proteins (KEPs)*

A recent biocomputational analysis of 250 fungal genomes for the occurrence of secreted, kexin-processed and highly repetitive proteins suggests that such proteins, referred to as KEX2-processed repeat proteins (KEPs), are very common in all fungal phyla and give rise to a large diversity of yet uncharacterized, secreted fungal peptides (Le Marquer et al. 2019). Some of the peptides derived from such precursors do not need additional modifications for bioactivity, and do therefore not meet the definition of a RiPP. Characterized examples are  $\alpha$ -factor, one of the two types of the *Saccharomyces cerevisiae* mating pheromones (Brake et al. 2015) and candidalysin, a recently identified virulence factor of the mucosal pathogen *Candida albicans* (Moyes et al. 2016). These ribosomally synthesized peptides are not further discussed here.

### **Application of RiPP biosynthetic pathways for the development of peptide therapeutics**

In the light of the blockbuster therapeutics penicillin and cyclosporin A, there is a strong interest in exploiting fungi for the development of novel peptide therapeutics. In this regard, RiPP biosynthetic pathways have several advantages over the non-ribosomal biosynthesis of peptides which can be exploited using two main approaches:

*Discovery approach:* The examples of penicillin and cyclosporin A show that fungal peptide natural products can sometimes be used directly as therapeutics without significant alterations. Thus, it is worthwhile to search for novel peptide natural products in fungi as they may be used, with minor alterations, as peptide therapeutics. In this regard, precursors of new members of already characterized RiPP families can be easily identified by biocomputational analyses of the rapidly increasing number of available fungal genome sequences. Based on the available knowledge about the processing of the precursor in the respective RiPP family, a putative sequence of the core peptide can be derived. This information, together with data on the expression of the precursor gene, helps in the subsequent identification of the respective RiPP natural product and the elucidation of its structure. Latter analysis is further aided by the fact that the various modifications of a core peptide are usually mediated by individual enzymes which are encoded by genes clustering around the precursor-encoding gene. The biosynthetic pathway of a novel fungal RiPP can thus be dissected by deletion or heterologous expression of the individual biosynthetic genes and biochemical characterization of the encoded enzymes. The new compounds can then be tested for their biological and/or pharmaceutical activities. A nice example for this approach is the recent identification and characterization of asperipin-2a whose biological activity is not known yet (Nagano et al. 2016; Ye et al. 2019). Once a novel RiPP natural product with an interesting biological activity is identified, variants of the compound could be easily generated by genetic alteration of the core peptide sequence in the precursor and/or modification of the biosynthetic pathway by removing one or more modifying enzymes.

*Synthetic biology approach:* The promiscuity of the modifying enzymes for the core peptide sequence may be exploited for the development of peptide therapeutics beyond the direct use of the natural products. The idea of this synthetic biology approach is to use these enzymes, along with engineered precursors, to generate, *in vitro* or *in vivo*, libraries of new-to-nature peptides with the respective posttranslational modifications. These libraries can then be screened, *in vitro* or *in vivo*, for desired biological activities and the hits can serve as leads for novel peptide therapeutics. A scheme of this strategy is shown in Fig. 1b. As an example, Sgambelluri *et al.* managed to create *de novo* libraries of macrocyclic peptides using chemically synthesized or *E. coli*-produced linear precursor peptides. The peptides were cyclized *in vitro* using *S. cerevisiae*-produced POPB from *G. marginata* (Sgambelluri et al. 2018). Such libraries are of interest as macrocyclization is, together with backbone N-methylation, one of the favourite strategies to improve the pharmacological properties of peptides. Both modifications increase the metabolic stability, structural rigidity, target specificity, cell permeability and oral availability of peptides (Pattabiraman and Bode 2011; Chatterjee et al. 2013). Since retroactive modification of peptides often destroys their

## Chapter 1

biological activity, it would be preferable if the peptide library to be screened for a biological activity would already contain these modifications. Both of these modifications are difficult to achieve by chemical synthesis (Pattabiraman and Bode 2011) and engineering of respective NRPSs, such as the one for cyclosporin A, is a difficult task (Winn et al. 2016), in particular for the production of a peptide library. Thus, alternative approaches are wanted. One such approach is based on *in vitro* translation and reprogramming of codons from proteinogenic to non-proteinogenic amino acids (Passioura and Suga 2017). A key element of this system are promiscuous, ribozyme-based aminoacyl-tRNA-synthetases that accept a large variety of chemically synthesized, non-proteinogenic amino acids, including  $\alpha$ -N-methylated ones, and thereby massively increase the chemical space covered by ribosomally synthesized peptides. Coupling of this reprogrammed ribosomal synthesis with ribosomal display and screening technologies allowed the recent identification of a natural product-like macrocyclic N-methyl peptide inhibitor of a ubiquitin ligase (Yamagishi et al. 2011). A possibility for another biotechnological approach is based on the promiscuity of the methyltransferase and prolyl oligopeptidase involved in omphalotin A biosynthesis. In this regard, Van der Velden et al. demonstrated that the methyltransferase domain of the omphalotin A precursor OphMA functions in the cytoplasm of *E. coli* and accepts omphalotin A-unrelated sequences in its core peptide region (Van Der Velden et al. 2017). Ramm et al. demonstrated that OphMA also functions in the cytoplasm of *P. pastoris* and that coexpression of the prolyl oligopeptidase OphP results in formation of the peptide natural product in this organism (Ramm et al. 2017). If OphP also shows promiscuity for the core peptide region of OphMA, libraries of macrocyclic, backbone N-methylated peptides could be generated *in vivo* by coexpression of OphP and hybrid OphMAs containing a randomized core peptide region. This *in vivo* approach would, in comparison to above mentioned *in vitro* approach, have the advantages that it would allow a higher number of backbone N-methylations per peptide and a coupling of the production and the screening of the peptide in the same microbial cell. Such systems were constructed for bacterial RiPP pathways and recently lead to the identification of natural-product-like peptides (Yang and Van der Donk 2016). Further characterization of the promiscuity of OphMA and OphP is needed to judge the feasibility of such an approach.

## Chapter 1

RiPP family	Representatives	Activity	References
Amatoxins/ phallotoxins	$\alpha$ -/ $\beta$ -/ $\gamma$ -/ $\epsilon$ -amanitin, amanullin, amanullic acid, amaninamide, amanin, proamanullin	Inhibition of RNA polymerase II. Highly toxic to insects, nematodes and mammals.	(Baumann et al. 1993; Garcia et al. 2015)
	Phalloidin, phalloidin, prophalloin, phalloin, phallisin, phallacin, phallisin	Inhibition of actin polymerization. Low toxicity due to poor absorption by the gut of fungivores.	(Wieland et al. 1978; Garcia et al. 2015)
	Antamanide	Antidote activity against amatoxins and phallotoxins. Immunosuppressive activity by binding to cyclophilin D.	(Wieland et al. 1978; Siemion et al. 1992; Azzolin et al. 2011)
	Cycloamanides A-F	Immunosuppressive activity (cycloamanides A-D)	(Gauhe and Wieland 1977; Wieczorek et al. 1993; Pulman et al. 2016)
	Amanexitide	Unknown activity	(Xue et al. 2011)
	Virotoxins***	Alaviroidin, viroisin, deoxoviroisin, viroidin, deoxoviroidin	Similarly to phallotoxins: Inhibition of actin polymerization. Low toxicity due to poor absorption by the gut of fungivores.
Borosins	Omphalotin A-I	Toxicity towards nematodes (omphalotin A-I).	(Büchel et al. 1998; Liermann et al. 2009)
Dikaritins	Ustiloxin A-G	Suppression of mitosis by inhibition of tubulin polymerization (ustiloxin A-F).	(Koiso et al. 1994; Koiso et al. 1998; Ding et al. 2016)
	Phomopsins A-E, P	Suppression of mitosis by inhibition of tubulin polymerization (phomopsin A, B, C, D).	(Allen and Hancock 1989; Battilani et al. 2011; Ding et al. 2016)
	Asperipin-2a	Unknown activity	(Nagano et al. 2016; Ye et al. 2019)
Epichloëcyclins	Epichloëcyclin A-F	Unknown activity. Only produced in endosymbiosis with grass.	(Johnson et al. 2015)

\*\*\*Likely also MSDIN peptides but precursor genes are not known yet.

**Table 1:** Compilation of all currently known fungal RiPP classes, their representatives and their activities.

### **Concluding remarks**

In the light of the rapidly increasing number of available fungal genome sequences, the recent discovery of fungal RiPPs has revived the interest in fungi as sources of peptide natural products. There are two main reasons for this interest: First, the biocomputational identification of structural or sequence variants of characterized RiPP precursors is comparably easy and can serve as basis for the identification of novel peptide natural products with interesting biological activities. Second, the discovery of modifying enzymes that introduce unprecedented peptide modifications that may allow to produce peptide libraries. These libraries would contain new-to-nature peptides with favorable pharmacological properties which can be screened for interesting biological activities. Both of these approaches will speed up the identification and development of novel peptide therapeutics, some of which may be of similar clinical significance as the non-ribosomal peptides penicillin and cyclosporin A.

### **Notes**

\*The authors prefer 'dikaritins' over 'ust-RiPS' suggested by Nagano et al 2016 as designation for the RiPP family comprising ustiloxins, phomopsins and asperipin-2a although the original definition of 'dikaritins' by Ding et al 2016 was restricted to peptides with sequence similarity to ustiloxins and would, thus, not include asperipin-2a.

\*\*The authors suggest to use the designation 'OphMA' (as used in Ramm et al 2017 and Ongpipattanakul and Nair 2018) instead of 'OphA' (as used in van der Velden et al 2017 and Song et al 2018) for the omphalotin precursor to indicate the presence of the methyltransferase domain and to avoid dual designations in future publications.

## References

- Abbas HK, Shier WT, Cartwright RD, Sciumbato GL. 2014. *Ustilaginoidea virens* Infection of Rice in Arkansas: Toxicity of False Smut Galls, Their Extracts and the Ustiloxin Fraction. *Am J Plant Sci*. 5(October):3166–3176. <https://doi.org/10.4236/ajps.2014.521333>
- Allen JG, Hancock GR. 1989. Evidence that phomopsins A and B are not the only toxic metabolites produced by *Phomopsis leptostromiformis*. *J Appl Toxicol*. 9(2):83–89. <https://doi.org/10.1002/jat.2550090203>
- Ambrose K V., Belanger FC. 2012. SOLiD-SAGE of Endophyte-Infected Red Fescue Reveals Numerous Effects on Host Transcriptome and an Abundance of Highly Expressed Fungal Secreted Proteins. *PLoS One*. 7(12):1–17. <https://doi.org/10.1371/journal.pone.0053214>
- Anke T, Schüffler A. 2018. *Physiology and Genetics: Selected Basic and Applied Aspects (The Mycota)*. XV. Heidelberg: Springer International Publishing.
- Arnison PG, Bibb MJ, Bierbaum G, Bowers AA, Bugni TS, Bulaj G, Camarero JA, Campopiano DJ, Challis GL, Clardy J, et al. 2013. Ribosomally synthesized and post-translationally modified peptide natural products: Overview and recommendations for a universal nomenclature. *Nat Prod Rep*. 30(1):108–160. <https://doi.org/10.1039/c2np20085f>
- Azzolin L, Antolini N, Calderan A, Ruzza P, Sciacovelli M, Marin O, Mammi S, Bernardi P, Rasola A. 2011. Antamanide, a Derivative of *Amanita phalloides*, is a Novel Inhibitor of the Mitochondrial Permeability Transition Pore. *PLoS One*. 6(1):26–29. <https://doi.org/10.1371/journal.pone.0016280>
- Battilani P, Gualla A, Dall'Asta C, Pellacani C, Galaverna G, Giorni P, Caglieri A, Tagliaferri S, Pietri A, Dossena A, et al. 2011. Phomopsins: An overview of phytopathological and chemical aspects, toxicity, analysis and occurrence. *World Mycotoxin J [Internet]*. 4(4):345–359. <https://doi.org/10.3920/WMJ2011.1302>
- Baumann K, Muentner K, Faulstich H. 1993. Identification of Structural Features Involved in Binding of alpha-Amanitin to a Monoclonal Antibody. *Biochemistry [Internet]*. 32(15):4043–4050. <https://doi.org/10.1021/bi00066a027>
- Beutler JA. 2009. Natural products as a foundation for drug discovery. *Curr Protoc Pharmacol.(SUPPL. 46):1–30*. <https://doi.org/10.1002/0471141755.ph0911s46>
- Bills GF, Gloer JB. 2016. Biologically Active Secondary Metabolites from the Fungi. *Microbiol Spectr [Internet]*. 4(6):1–32. <https://doi.org/10.1128/microbiolspec.FUNK-0009-2016>
- Boecker S, Zobel S, Meyer V, Süßmuth RD. 2016. Rational biosynthetic approaches for the production of new-to-nature compounds in fungi. *Fungal Genet Biol [Internet]*. 89:89–101. <https://doi.org/10.1016/j.fgb.2016.02.003>
- Boot CM, Tenney K, Valeriote FA, Crews P. 2006. Highly N-methylated linear peptides produced by an atypical sponge-derived *Acremonium sp.* *J Nat Prod*. 69(1):83–92. <https://doi.org/10.1021/np0503653>
- Bozhüyük KAJ, Fleischhacker F, Linck A, Wesche F, Tietze A, Niesert CP, Bode HB. 2018. De novo design and engineering of non-ribosomal peptide synthetases. *Nat Chem [Internet]*. 10(3):275–281. <https://doi.org/10.1038/NCHEM.2890>

## Chapter 1

- Brake AJ, Julius DJ, Thorner J. 2015. A functional prepro-alpha-factor gene in *Saccharomyces* yeasts can contain three, four, or five repeats of the mature pheromone sequence. *Mol Cell Biol.* 3(8):1440–1450. <https://doi.org/10.1128/mcb.3.8.1440>
- Bruno BJ, Miller GD, Lim CS. 2013. Basics and recent advances in peptide and protein drug delivery. *Ther Deliv.* 4(11):1443–1467. <https://doi.org/10.4155/tde.13.104>
- Büchel E, Martini U, Mayer A, Anke H, Sterner O. 1998. Omphalotins B, C and D, nematocidal cyclopeptides from *Omphalotus olearius*. Absolute configuration of omphalotin A. *Tetrahedron.* 54(20):5345–5352. [https://doi.org/10.1016/S0040-4020\(98\)00209-9](https://doi.org/10.1016/S0040-4020(98)00209-9)
- Bushnell DA, Cramer P, Kornberg RD. 2002. Structural basis of transcription: alpha-Amanitin-RNA polymerase II cocystal at 2.8 Å resolution. *Proc Natl Acad Sci [Internet].* 99(3):1218–1222. <https://doi.org/10.1073/pnas.251664698>
- Cerqueira GC, Arnaud MB, Inglis DO, Skrzypek MS, Binkley G, Simison M, Miyasato SR, Binkley J, Orvis J, Shah P, et al. 2014. The *Aspergillus* Genome Database: Multispecies curation and incorporation of RNA-Seq data to improve structural gene annotations. *Nucleic Acids Res.* 42(D1):705–710. <https://doi.org/10.1093/nar/gkt1029>
- Chafin DR, Guo H, Price DH. 1995. Action of alpha-amanitin during pyrophosphorolysis and elongation by RNA polymerase II. *J Biol Chem [Internet].* 270(32):19114–9. <https://doi.org/10.1074/jbc.270.32.19114>
- Chatterjee J, Rechenmacher F, Kessler H. 2013. N-Methylation of peptides and proteins: An important element for modulating biological functions. *Angew Chemie - Int Ed.* 52(1):254–269. <https://doi.org/10.1002/anie.201205674>
- Cragg GM, Newman DJ. 2013. Natural products: A continuing source of novel drug leads. *Biochim Biophys Acta - Gen Subj [Internet].* 1830(6):3670–3695. <https://doi.org/10.1016/j.bbagen.2013.02.008>
- Czekster CM, Ludewig H, McMahon SA, Naismith JH. 2017. Characterization of a dual function macrocyclase enables design and use of efficient macrocyclization substrates. *Nat Commun [Internet].* 8(1):1–10. <https://doi.org/10.1038/s41467-017-00862-4>
- Czekster CM, Naismith JH. 2017. Kinetic landscape of a peptide bond-forming prolyl oligopeptidase. *Biochemistry.* 56(15):2086–2095. <https://doi.org/10.1021/acs.biochem.7b00012>
- Dias DA, Urban S, Roessner U. 2012. A Historical overview of natural products in drug discovery. *Metabolites.* 2(2):303–336. <https://doi.org/10.3390/metabo2020303>
- Ding W, Liu W-Q, Jia Y, Li Y, van der Donk WA, Zhang Q. 2016. Biosynthetic investigation of phomopsins reveals a widespread pathway for ribosomal natural products in Ascomycetes. *Proc Natl Acad Sci [Internet].* 113(13):3521–3526. <https://doi.org/10.1073/pnas.1522907113>
- Dreyfuss M, Härrä E, Hofmann H, Kobel H, Pache W, Tschertter H. 1976. Cyclosporin A and C New Metabolites from *Trichoderma polysporum*. *Eur J Appi Microbiol.* 3(1):125–133. <https://doi.org/10.1007/BF00928431>
- Faulstich H, Buku A, Bodenmüller H, Wieland T. 1980. Virotoxins: Actin-Binding Cyclic Peptides of *Amanita Virosa* Mushrooms. *Biochemistry.* 19(14):3334–3343. <https://doi.org/10.1021/bi00555a036>

## Chapter 1

- Fosgerau K, Hoffmann T. 2015. Peptide therapeutics: Current status and future directions. *Drug Discov Today* [Internet]. 20(1):122–128. <https://doi.org/10.1016/j.drudis.2014.10.003>
- Freeman MF, Gurgui C, Helf MJ, Morinaka BI, Uria AR, Oldham NJ, Sahl HG, Matsunaga S, Piel J. 2012. Metagenome mining reveals polytheonamides as posttranslationally modified ribosomal peptides. *Science* (80- ). 338(6105):387–390. <https://doi.org/10.1126/science.1226121>
- Garcia J, Costa VM, Carvalho A, Baptista P, de Pinho PG, de Lourdes Bastos M, Carvalho F. 2015. *Amanita phalloides* poisoning: Mechanisms of toxicity and treatment. *Food Chem Toxicol*. 86:41–55. <https://doi.org/10.1016/j.fct.2015.09.008>
- Gauhe A, Wieland T. 1977. Über die Inhaltsstoffe des grünen Knollenblätterpilzes, LI. Die Cycloamanide, monocyclische Peptide; Isolierung und Strukturaufklärung eines cyclischen Heptapeptids (CyA B) und zweier cyclischer Oktapeptide (CyA C und CyA D). *Justus Liebigs Ann Chem*. 1977(5):859–868. <https://doi.org/10.1002/jlac.197719770515>
- Gu W, Sardar D, Pierce E, Schmidt EW. 2018. Roads to Rome: Role of Multiple Cassettes in Cyanobactin RiPP Biosynthesis. *J Am Chem Soc* [Internet]. 140(47):16213–16221. <https://doi.org/10.1021/jacs.8b09328>
- Haft DH, Basu MK, Mitchell DA. 2010. Expansion of ribosomally produced natural products: A nitrile hydratase- and Nif11-related precursor family. *BMC Biol*. 8(70):1–15. <https://doi.org/10.1186/1741-7007-8-70>
- Hallen HE, Luo H, Scott-Craig JS, Walton JD. 2007. Gene family encoding the major toxins of lethal *Amanita* mushrooms. *Proc Natl Acad Sci* [Internet]. 104(48):19097–19101. <https://doi.org/10.1073/pnas.0707340104>
- Hawksworth DL, Lücking R. 2017. Fungal Diversity Revisited: 2.2 to 3.8 Million Species. *Microbiol Spectr* [Internet]. 5(4). <https://doi.org/10.1128/microbiolspec.FUNK-0052-2016.Correspondence>
- Henkel T, Brunne RM, Müller H, Reichel F. 1999. Statistical investigation into the structural complementarity of natural products and synthetic compounds. *Angew Chemie - Int Ed*. 38(5):643–647. [https://doi.org/10.1002/\(SICI\)1521-3773\(19990301\)38:5<643::AID-ANIE643>3.0.CO;2-G](https://doi.org/10.1002/(SICI)1521-3773(19990301)38:5<643::AID-ANIE643>3.0.CO;2-G)
- Hollstein U. 1974. Actinomycin. Chemistry and mechanism of action. *Chem Rev*. 74(6):625–652. <https://doi.org/10.1021/cr60292a002>
- Jiang ZD, An Z. 2000. Bioactive fungal natural products through classic and biocombinatorial approaches. *Stud Nat Prod Chem* [Internet]. 22(PART C):245–272. [https://doi.org/10.1016/S1572-5995\(00\)80027-7](https://doi.org/10.1016/S1572-5995(00)80027-7)
- Johnson LJ, Johnson RD, Schardl CL, Panaccione DG. 2003. Identification of differentially expressed genes in the mutualistic association of tall fescue with *Neotyphodium coenophialum*. *Physiol Mol Plant Pathol*. 63(6):305–317. <https://doi.org/10.1016/j.pmpp.2004.04.001>
- Johnson RD, Lane GA, Koulman A, Cao M, Fraser K, Fleetwood DJ, Voisey CR, Dyer JM, Pratt J, Christensen M, et al. 2015. A novel family of cyclic oligopeptides derived from ribosomal peptide synthesis of an in planta-induced gene, *gigA*, in *Epichloë* endophytes of grasses. *Fungal Genet Biol* [Internet]. 85:14–24. <https://doi.org/10.1016/j.fgb.2015.10.005>

- Kelleher NL, Hendrickson CL, Walsh CT. 1999. Posttranslational heterocyclization of cysteine and serine residues in the antibiotic Microcin B17: Distributivity and directionality. *Biochemistry*. 38(47):15623–15630. <https://doi.org/10.1021/bi9913698>
- Khan A, Bassett S, Voisey C, Gaborit C, Johnson L, Christensen M, McCulloch A, Bryan G, Johnson R. 2010. Gene expression profiling of the endophytic fungus *Neotyphodium lolii* in association with its host plant perennial ryegrass. *Australas Plant Pathol*. 39(5):467–476. <https://doi.org/10.1071/AP09084>
- Koehnke J, Bent A, Houssen WE, Zollman D, Morawitz F, Shirran S, Vendome J, Nneoyiegbe AF, Trembleau L, Botting CH, et al. 2012. The mechanism of patellamide macrocyclization revealed by the characterization of the PatG macrocyclase domain. *Nat Struct Mol Biol*. 19(8):767–772. <https://doi.org/10.1038/nsmb.2340>
- Koiso Y, Li Y, Iwasaki S. 1994. Ustiloxins, antimitotic cyclic peptides from false smut balls on rice panicles caused by *Ustilaginoidea virens*. *J Antibiot (Tokyo)* [Internet]. 47(7):765–773. <https://doi.org/10.7164/antibiotics.47.765>
- Koiso Y, Morisaki N, Yamashita Y, Mitsui Y, Shirai R, Hashimoto Y, Iwasaki S. 1998. Isolation and structure of an antimitotic cyclic peptide, ustiloxin F: Chemical interrelation with a homologous peptide, ustiloxin B. *J Antibiot (Tokyo)* [Internet]. 51(4):418–422. <https://doi.org/10.7164/antibiotics.51.418>
- Koiso Y, Natori M, Iwasaki S, Sato S, Sonoda R, Fujita Y, Yaegashi H, Sato Z. 1992. Ustiloxin: a phytotoxin and a mycotoxin from false smut balls on rice panicles. *Tetrahedron Lett*. 33(29):4157–4160. [https://doi.org/10.1016/S0040-4039\(00\)74677-6](https://doi.org/10.1016/S0040-4039(00)74677-6)
- Komatsu K, Shigemori H, Kobayashi J. 2001. Dictyonamides A and B, new peptides from marine-derived fungus. *J Org Chem*. 66(18):6189–6192. <https://doi.org/10.1021/jo0156767>
- Krawczyk B, Ensle P, Müller WM, Süßmuth RD. 2012. Deuterium labeled peptides give insights into the directionality of class III lantibiotic synthetase LabKC. *J Am Chem Soc*. 134(24):9922–9925. <https://doi.org/10.1021/ja3040224>
- Lau JL, Dunn MK. 2018. Therapeutic peptides: Historical perspectives, current development trends, and future directions. *Bioorganic Med Chem* [Internet]. 26(10):2700–2707. <https://doi.org/10.1016/j.bmc.2017.06.052>
- Lee MV, Ihnken LAF, Young OY, McClerren AL, Van Der Donk WA, Kelleher NL. 2009. Distributive and directional behavior of lantibiotic synthetases revealed by high-resolution tandem mass spectrometry. *J Am Chem Soc*. 131(34):12258–12264. <https://doi.org/10.1021/ja9033507>
- Lengsfeld AM, Löw I, Wieland T, Dancker P, Hasselbach W. 1974. Interaction of phalloidin with actin. *Proc Natl Acad Sci U S A*. 71(7):2803–2807. <https://doi.org/10.1073/pnas.71.7.2803>
- Li J, Gu F, Wu R, Yang JK, Zhang KQ. 2017. Phylogenomic evolutionary surveys of subtilase superfamily genes in fungi. *Sci Rep* [Internet]. 7(March):1–15. <https://doi.org/10.1038/srep45456>
- Liermann JC, Opatz T, Kolshorn H, Antelo L, Hof C, Anke H. 2009. Omphalotins E-I, five oxidatively modified nematocidal cyclopeptides from *Omphalotus olearius*. *European J Org Chem*. 2009(8):1256–1262. <https://doi.org/10.1002/ejoc.200801068>

## Chapter 1

- Lubelski J, Khusainov R, Kuipers OP. 2009. Directionality and coordination of dehydration and ring formation during biosynthesis of the lantibiotic nisin. *J Biol Chem.* 284(38):25962–25972. <https://doi.org/10.1074/jbc.M109.026690>
- Luo H, Cai Q, Lüli Y, Li X, Sinha R, Hallen-Adams HE, Yang ZL. 2018. The MSDIN family in amanitin-producing mushrooms and evolution of the prolyl oligopeptidase genes. *IMA Fungus* [Internet]. 9(2):225–242. <https://doi.org/10.5598/imafungus.2018.09.02.01>
- Luo H, Hallen-Adams HE, Scott-Craig JS, Walton JD. 2010. Colocalization of amanitin and a candidate toxin-processing prolyl oligopeptidase in *Amanita basidiocarps*. *Eukaryot Cell.* 9(12):1891–1900. <https://doi.org/10.1128/EC.00161-10>
- Luo H, Hallen-Adams HE, Scott-Craig JS, Walton JD. 2012. Ribosomal biosynthesis of  $\alpha$ -amanitin in *Galerina marginata*. *Fungal Genet Biol* [Internet]. 49(2):123–129. <https://doi.org/10.1016/j.fgb.2011.12.005>
- Luo H, Hallen-Adams HE, Walton JD. 2009. Processing of the phalloidin proprotein by prolyl oligopeptidase from the mushroom *Conocybe albipes*. *J Biol Chem.* 284(27):18070–18077. <https://doi.org/10.1074/jbc.M109.006460>
- Luo H, Hong SY, Sgambelluri RM, Angelos E, Li X, Walton JD. 2014. Peptide Macrocyclization Catalyzed by a Prolyl Oligopeptidase Involved in  $\alpha$ -Amanitin Biosynthesis. *Chem Biol* [Internet]. 21(12):1610–1617. <https://doi.org/10.1016/j.chembiol.2014.10.015>. Peptide
- Maharani R, Sleebs BE, Hughes AB. 2015. Macrocyclic N-methylated cyclic peptides and depsipeptides. *Stud Nat Prod Chem* [Internet]. 44:113–249. <https://doi.org/10.1016/B978-0-444-63460-3.00004-3>
- Le Marquer M, San Clemente H, Roux C, Savelli B, Frei dit Frey N. 2019. Identification of new signalling peptides through a genome-wide survey of 250 fungal secretomes. *BMC Genomics* [Internet]. 20(1):64. <https://doi.org/10.1186/s12864-018-5414-2>
- May JP, Perrin DM. 2007. Tryptathionine Bridges in Peptide Synthesis. *Biopolymers.* 88(5):714–724. <https://doi.org/10.1002/bip>
- Mayer A, Anke H, Sterner O. 1997. Omphalotin, a new cyclic peptide with potent nematocidal activity from *Omphalotus olearius*. I. Fermentation and biological activity. *Nat Prod Lett.* 10(1):25–32. <https://doi.org/10.1080/10575639708043691>
- McIntosh JA, Donia MS, Schmidt EW. 2009. Ribosomal peptide natural products: Bridging the ribosomal and nonribosomal worlds. *Nat Prod Rep.* 26(4):537–559. <https://doi.org/10.1039/b714132g>
- Melby JO, Dunbar KL, Trinh NQ, Mitchell DA. 2012. Selectivity, directionality, and promiscuity in peptide processing from a *bacillus* sp. Al Hakam cyclodehydratase. *J Am Chem Soc.* 134(11):5309–5316. <https://doi.org/10.1021/ja211675n>
- Moyes DL, Wilson D, Richardson JP, Mogavero S, Tang SX, Wernecke J, Höfs S, Gratacap RL, Robbins J, Runglall M, et al. 2016. Candidalysin is a fungal peptide toxin critical for mucosal infection. *Nature* [Internet]. 532(7597):64–68. <https://doi.org/10.1038/nature17625>
- Müller MM. 2018. Post-Translational Modifications of Protein Backbones: Unique Functions, Mechanisms, and Challenges. *Biochemistry.* 57(2):177–185. <https://doi.org/10.1021/acs.biochem.7b00861>

## Chapter 1

- Nagano N, Umemura M, Izumikawa M, Kawano J, Ishii T, Kikuchi M, Tomii K, Kumagai T, Yoshimi A, Machida M, et al. 2016. Class of cyclic ribosomal peptide synthetic genes in filamentous fungi. *Fungal Genet Biol* [Internet]. 86(January):58–70. <https://doi.org/10.1016/j.fgb.2015.12.010>
- Newman DJ, Cragg GM. 2016. Natural Products as Sources of New Drugs from 1981 to 2014. *J Nat Prod*. 79(3):629–661. <https://doi.org/10.1021/acs.jnatprod.5b01055>
- Ongpipattanakul C, Nair SK. 2018. Molecular Basis for Autocatalytic Backbone N-Methylation in RiPP Natural Product Biosynthesis. *ACS Chem Biol* [Internet]. 13(10):2989–2999. <https://doi.org/10.1021/acschembio.8b00668>
- Ortega MA, Van Der Donk WA. 2016. New Insights into the Biosynthetic Logic of Ribosomally Synthesized and Post-translationally Modified Peptide Natural Products. *Cell Chem Biol* [Internet]. 23(1):31–44. <https://doi.org/10.1016/j.chembiol.2015.11.012>
- Passioura T, Suga H. 2017. A RaPID way to discover nonstandard macrocyclic peptide modulators of drug targets. *Chem Commun*. 53(12):1931–1940. <https://doi.org/10.1039/c6cc06951g>
- Pattabiraman VR, Bode JW. 2011. Rethinking amide bond synthesis. *Nature* [Internet]. 480(7378):471–479. <https://doi.org/10.1038/nature10702>
- Pulman JA, Childs KL, Sgambelluri RM, Walton JD. 2016. Expansion and diversification of the MSDIN family of cyclic peptide genes in the poisonous agarics *Amanita phalloides* and *A. bisporigera*. *BMC Genomics* [Internet]. 17(1):1038. <https://doi.org/10.1186/s12864-016-3378-7>
- Ramm S, Krawczyk B, Mühlenweg A, Poch A, Mösker E, Süßmuth RD. 2017. A Self-Sacrificing N-Methyltransferase Is the Precursor of the Fungal Natural Product Omphalotin. *Angew Chemie - Int Ed*. 56(33):9994–9997. <https://doi.org/10.1002/anie.201703488>
- Rudd MD, Luse DS. 1996. Amanitin greatly reduces the rate of transcription by RNA polymerase II ternary complexes but fails to inhibit some transcript cleavage modes. *J Biol Chem*. 271(35):21549–21558. <https://doi.org/10.1074/jbc.271.35.21549>
- Scherer G, Kramer ML, Schutkowski M, Reimer U, Fischer G. 1998. Barriers to rotation of secondary amide peptide bonds. *J Am Chem Soc*. 120(22):5568–5574. <https://doi.org/10.1021/ja980181t>
- Sgambelluri MR, Epis S, Sasser D, Luo H, Angelos ER, Walton JD. 2014. Profiling of amatoxins and phallotoxins in the genus *Lepiota* by liquid chromatography combined with UV absorbance and mass spectrometry. *Toxins (Basel)*. 6(8):2336–2347. <https://doi.org/10.3390/toxins6082336>
- Sgambelluri RM, Smith MO, Walton JD. 2018. Versatility of Prolyl Oligopeptidase B in Peptide Macrocyclization. *ACS Synth Biol*. 7(1):145–152. <https://doi.org/10.1021/acssynbio.7b00264>
- Siemion IZ, Pędyczak A, Trojnar J, Zimecki M, Wieczorek Z. 1992. Immunosuppressive activity of antamanide and some of its analogues. *Peptides*. 13(6):1233–1237. [https://doi.org/10.1016/0196-9781\(92\)90034-Z](https://doi.org/10.1016/0196-9781(92)90034-Z)
- Solden L, Lloyd K, Wrighton K. 2016. The bright side of microbial dark matter: Lessons learned from the uncultivated majority. *Curr Opin Microbiol* [Internet]. 31:217–226.

## Chapter 1

<https://doi.org/10.1016/j.mib.2016.04.020>

Song H, Van Der Velden NS, Shiran SL, Bleiziffer P, Zach C, Sieber R, Imani AS, Krausbeck F, Aebi M, Freeman MF, et al. 2018. A molecular mechanism for the enzymatic methylation of nitrogen atoms within peptide bonds. *Sci Adv.* 4(8). <https://doi.org/10.1126/sciadv.aat2720>

Spiteller P. 2015. Chemical ecology of fungi. *Nat Prod Rep.* 32(7):971–993. <https://doi.org/10.1039/c4np00166d>

Sterner O, Etzel W, Mayer A, Anke H. 1997. Omphalotin, a new cyclic peptide with potent nematocidal activity from *Omphalotus olearius*. II. Isolation and structure determination. *Nat Prod Lett.* 10(1):33–38. <https://doi.org/10.1080/10575639708043692>

Süssmuth RD, Mainz A. 2017. Nonribosomal Peptide Synthesis—Principles and Prospects. *Angew Chemie - Int Ed.* 56(14):3770–3821. <https://doi.org/10.1002/anie.201609079>

Tsukui T, Nagano N, Terai G, Kumagai T, Machida M, Umemura M, Asai K. 2014. Ustiloxins, fungal cyclic peptides, are ribosomally synthesized in *Ustilaginoidea virens*. *Bioinformatics.* 31(7):981–985. <https://doi.org/10.1093/bioinformatics/btu753>

Umemura M, Nagano N, Koike H, Kawano J, Ishii T, Miyamura Y, Kikuchi M, Tamano K, Yu J, Shin-ya K, Machida M. 2014. Characterization of the biosynthetic gene cluster for the ribosomally synthesized cyclic peptide ustiloxin B in *Aspergillus flavus*. *Fungal Genet Biol [Internet].* 68:23–30. <https://doi.org/10.1016/j.fgb.2014.04.011>

Ványolós A, Dékány M, Kovács B, Krámos B, Bérdi P, Zupkó I, Hohmann J, Béni Z. 2016. Gymnopeptides A and B, Cyclic Octadecapeptides from the Mushroom *Gymnopus fusipes*. *Org Lett.* 18(11):2688–2691. <https://doi.org/10.1021/acs.orglett.6b01158>

Van Der Velden NS, Kälin N, Helf MJ, Piel J, Freeman MF, Künzler M. 2017. Autocatalytic backbone N-methylation in a family of ribosomal peptide natural products. *Nat Chem Biol.* 13(8):833–835. <https://doi.org/10.1038/nchembio.2393>

Vetter J. 1998. Toxins of *Amanita phalloides*. *Toxicon.* 36(1):13–24. [https://doi.org/10.1016/S0041-0101\(97\)00074-3](https://doi.org/10.1016/S0041-0101(97)00074-3)

Walton JD, Hallen-Adams HE, Luo H. 2010. Ribosomal biosynthesis of the cyclic peptide toxins of *Amanita* mushrooms. *Biopolymers.* 94(5):659–664. <https://doi.org/10.1002/bip.21416>

Wang X, Wang J, Lai D, Wang W, Dai J, Zhou L, Liu Y. 2017. Ustiloxin G, a new cyclopeptide mycotoxin from rice false smut balls. *Toxins (Basel).* 9(2):1–9. <https://doi.org/10.3390/toxins9020054>

Wieczorek Z, Siemion IZ, Zimecki M /k k., Bolewska-Pedyczak E, Wieland T. 1993. Immunosuppressive activity in the series of cycloamanide peptides from mushrooms. *Peptides.* 14(1):1–5. [https://doi.org/10.1016/0196-9781\(93\)90003-Y](https://doi.org/10.1016/0196-9781(93)90003-Y)

Wieland T. 1987. 50 Jahre Phalloidin. *Naturwissenschaften [Internet].* 74(8):367–373. <https://doi.org/10.1007/BF00405464>

Wieland T, Faulstich H, Fiume L. 1978. Amatoxins, phallotoxins, phallolysin, and antamanide: The biologically active components of poisonous amanita mushroom. *Crit Rev Biochem Mol Biol.* 5(3):185–260. <https://doi.org/10.3109/10409237809149870>

Winn M, Fyans JK, Zhuo Y, Micklefield J. 2016. Recent advances in engineering nonribosomal

## Chapter 1

- peptide assembly lines. *Nat Prod Rep.* 33(2):317–347. <https://doi.org/10.1039/c5np00099h>
- Xue J-H, Wu P, Chi Y-L, Xu L-X, Wei X-Y. 2011. Cyclopeptides from *Amanita exitialis*. *Nat Products Bioprospect.* 1(1):52–56. <https://doi.org/10.1007/s13659-011-0013-9>
- Yamagishi Y, Shoji I, Miyagawa S, Kawakami T, Katoh T, Goto Y, Suga H. 2011. Natural product-like macrocyclic N-methyl-peptide inhibitors against a ubiquitin ligase uncovered from a ribosome-expressed de novo library. *Chem Biol [Internet].* 18(12):1562–1570. <https://doi.org/10.1016/j.chembiol.2011.09.013>
- Yang X, Van der Donk W. 2016. A lanthipeptide library used to identify a protein–protein interaction inhibitor. *35(14):1252–1260.* <https://doi.org/10.1177/0333102415576222>.ls
- Yang X, Van Der Donk WA. 2013. Ribosomally synthesized and post-translationally modified peptide natural products: New insights into the role of leader and core peptides during biosynthesis. *Chem - A Eur J.* 19(24):7662–7677. <https://doi.org/10.1002/chem.201300401>
- Ye Y, Minami A, Igarashi Y, Izumikawa M, Umemura M, Nagano N, Machida M, Kawahara T, Shin-ya K, Gomi K, Oikawa H. 2016. Unveiling the Biosynthetic Pathway of the Ribosomally Synthesized and Post-translationally Modified Peptide Ustiloxin B in Filamentous Fungi. *Angew Chemie - Int Ed.* 55(28):8072–8075. <https://doi.org/10.1002/anie.201602611>
- Ye Y, Ozaki T, Umemura M, Liu C, Minami A, Oikawa H. 2019. Heterologous production of asperipin-2a: Proposal for sequential oxidative macrocyclization by a fungi-specific DUF3328 oxidase. *Org Biomol Chem.* 17(1):39–43. <https://doi.org/10.1039/c8ob02824a>
- Yoshimi A, Umemura M, Nagano N, Koike H, Machida M, Abe K. 2016. Expression of ustR and the Golgi protease KexB are required for ustiloxin B biosynthesis in *Aspergillus oryzae*. *AMB Express.* 6(1):1–8. <https://doi.org/10.1186/s13568-016-0181-4>
- Zhang C hua, Zou J ping, Deng W qiu, Li T hui, Jiang Z de. 2018. Molecular cloning and the expression pattern of AePOPB involved in the  $\alpha$ -amanitin biosynthesis in *Amanita exitialis* fruiting bodies. *Gene [Internet].* 662(April):123–130. <https://doi.org/10.1016/j.gene.2018.04.006>
- Zhang P, Chen ZH, Hu JS, Wei BY, Zhang ZG, Hu WQ. 2005. Production and characterization of Amanitin toxins from a pure culture of *Amanita exitialis*. *FEMS Microbiol Lett.* 252(2):223–228. <https://doi.org/10.1016/j.femsle.2005.08.049>
- Zou X, Niu S, Ren J, Li E, Liu X, Che Y. 2011. Verrucamides A-D, antibacterial cyclopeptides from *Myrothecium verrucaria*. *J Nat Prod.* 74(5):1111–1116. <https://doi.org/10.1021/np200050r>

## Appendix 1: Current state of research on fungal RiPPs

**A 1.1 The fungal peptide victorin is a ribosomally synthesized and posttranslationally modified peptide**

The necrotrophic plant pathogen *Cochliobolus victoriae* caused an epidemic blight in oats in the 1940s (Meehan and Murphy 1946) and the host-specific toxin victorin was isolated from the fungus in the 1980s (Macko et al. 1985; Wolpert et al. 1986; Kono et al. 1989). Victorin is required for pathogenicity of the fungus and a single gene is responsible for host plant susceptibility to the toxin (Wolpert and Lorang 2016). The same susceptibility gene was originally introduced into oats during breeding programs to confer resistance to another plant pathogen, *Puccinia coronata* (Mayama et al. 1995). Victorin consists of a mixture of the peptide derivatives victorin B, C, D, E, victoricine and HV-toxin M (Macko et al. 1985; Wolpert et al. 1986; Kono et al. 1989). All derivatives contain a 12-membered ring with an ether bond, five peptide bonds and most derivatives are chlorinated (Appendix Fig. 1a). Victorin binds to and activates the immunoregulatory thioredoxin TRX-h5 in the model host *Arabidopsis thaliana* (Lorang et al. 2012). As a result, a hypersensitivity response is triggered in the plant, leading to cell death, a strategy normally used to halt biotrophic infection. Thus, *C. victoriae* exploits a plant defense mechanism for disease susceptibility. The fungus produces a virulence factor to trigger innate plant defense signalling and cell death, thus gaining access to dead cells for necrotrophic growth (Lorang et al. 2012).

Despite knowledge of the role of victorin in infection, the biosynthesis of the peptide remained unknown until recently. As with other complex cyclic peptides, extensive posttranslational modifications can mask the ribosomal origin of a peptide, and thus victorin was assumed to be synthesized by a non-ribosomal peptide synthase (NRPS). After unsuccessful attempts to delete candidate NRPSs (Condon et al. 2013; Zainudin et al. 2015), a retrosynthetic analysis suggested that the peptide backbone of victorin might correspond to the peptide sequence GLKLAF (Kessler et al. 2020). A subsequent BLAST search led to a KEX2-processed repeat protein with a secretion signal sequence and multiple “GLKLAF” core sequences flanked by KR dibasic residues, which proved to be the precursor protein of victorin (Appendix Fig. 1b, Kessler et al. 2020).

Three victorin precursors were identified in two different loci with low gene density and a high number of transposons. These loci contained several biosynthesis-related genes such as copper amine oxidases, an NADPH oxidase, DUF3328 proteins, peptidases, a transporter and a transcription factor (Kessler et al. 2020). As with the other members of the RiPP family of dikaritins, the synthesis of victorin relies on the activity of a DUF3328 protein. Deletion of the DUF3328 gene abolished the production of victorin. Additionally, a copper amine oxidase was shown to be required for oxidation of the N-terminal glycyl residue to a glyoxylate, a modification required for victorin toxicity (Kessler et al. 2020).



the latter two modifications were elucidated by the discovery of two new proteins encoded by the biosynthetic genes *GmFMO1* and *GmP450-29* (Luo et al. 2022). *GmFMO1* is predicted to encode a Flavin-containing mono-oxygenase, and disruption of the gene resulted in immature amanitin compounds, suggesting that oxygenation via *GmFMO1* is required for other posttranslational modifications. *GmP450-29* catalyzes the hydroxylation of several residues of  $\alpha$ -amanitin that are required for its affinity for target RNA polymerase. Disruption of *GmFMO1* or *GmP450-29* resulted in immature amanitin that lacks sulfoxidation and hydroxylations important for the peptide's toxic effect in inhibiting RNA polymerase (Luo et al. 2022).

The genes of these two new biosynthetic proteins, as well as the prolyl oligopeptidase and the amanitin precursor, were analyzed for their phylogenetic distance, revealing that the original gene cluster was likely spread by an ancestral donor among the distantly related genera *Amanita*, *Galerina* and *Lepiota* (Luo et al. 2022). The horizontally transferred genes underwent distinct evolutionary pathways in each of these genera, including gene expansion and genomic rearrangements caused by transposable elements, resulting in cluster disorganization and distribution of biosynthetic genes over a large portion of the genome. In *Amanita*, the biosynthetic machinery increased in complexity and versatility, allowing the generation of a broader range of different toxic peptides. For example, three additional cytochrome P450 enzymes are found in *Amanita* but not in *Galerina* or *Lepiota*, indicating that these and other genes may expand the repertoire of possible cyclic peptides and thus turn the recruitment of new biosynthetic genes into an evolutionary advantage (Luo et al. 2022).

### **A 1.3 Updates on classifications and labels of fungal RiPP families**

In our publication on fungal RiPPs in Chapter 1, we classified the known fungal RiPPs into four different families: the amatoxins/phallotoxins, the borosins (e.g., omphalotin), dikaritins (including ustiloxins, phomopsins and asperipin-2a) and the epichloëcyclins. The amatoxins/phallotoxins possess a conserved N-terminal sequence and are cleaved by a prolyl oligopeptidase, whereas the backbone N-methylated borosins arise from cleavage of a self-methylating precursor protein. The dikaritins and epichloëcyclins derive from KEX2-processed repeat proteins that have an N-terminal signal sequence for secretion and repetitive peptide cores flanked by KEX2 cleavage sites. The dikaritins are currently defined as cyclic peptides with an ether bond bridge (Vogt and Künzler 2019; Kessler and Chooi 2022). Structural elucidation of epichloëcyclins thus far relies on tandem mass spectrometry data, leaving the exact structure of epichloëcyclins unknown, including the nature of the cyclic bond (Johnson et al. 2015). Therefore, we designated epichloëcyclins as a separate RiPP family. However, a recent review on fungal RiPPs assigns the epichloëcyclins to the dikaritin family, arguing that the presence of a DUF3328 protein in the epichloëcyclin gene cluster is strong evidence that this peptide is cyclized via an ether bond like the other members of the dikaritins, whose cyclization depends on the presence of a DUF3328 protein (Kessler and Chooi 2022). This reasoning is plausible and our compilation of dikaritin members, as shown in Appendix 1 Table 1, is thus updated to include the epichloëcyclins and the newly discovered

## Chapter 1

victorins (Kessler et al. 2020). The function of the epichloëcyclins, like that of asperipin-2a, remains unknown.

Epichloëcyclins are structurally similar to dikaritins and their precursor proteins all share the same KEP-like organization. Therefore, even if the epichloëcyclins were shown to be cyclized without forming an ether bridge, it could be argued that they should still be considered dikaritins. The family of dikaritins could be redefined to include all cyclic KEP-derived peptides, regardless of how they are cyclized.

RiPP family	Representatives	Activity	References
Dikaritins	Ustiloxin A-G	Suppression of mitosis by inhibition of tubulin polymerization (ustiloxin A-F).	(Koiso et al. 1994; Koiso et al. 1998; Ding et al. 2016)
	Phomopsins A-E, P	Suppression of mitosis by inhibition of tubulin polymerization (phomopsin A, B, C, D).	(Allen and Hancock 1989; Battilani et al. 2011; Ding et al. 2016)
	Asperipin-2a	Unknown activity	(Nagano et al. 2016; Ye et al. 2019)
	Epichloëcyclin A-F	Unknown activity. Only produced in endosymbiosis with grass.	(Johnson et al. 2015)
	Victorin B, C, D, E, victoricine and HV-toxin M	Induction of programmed cell death in plant host.	(Macko et al. 1985; Wolpert et al. 1986; Kono et al. 1989; Lorang et al. 2012; Kessler et al. 2020)

**Appendix 1 Table 1:** Updated compilation of the fungal RiPP family dikaritins, adapted from Vogt and Künzler 2019.

The classification of fungal RiPPs into different families and their terminology has already undergone several changes in recent years (e.g., the replacement of the term “ust-RiPS” originally used by Nagano et al. 2016) and will likely need to be updated again soon, especially in light of the newly discovered abundance of KEPs in fungal genomes. On a similar note, the RiPP family of amatoxins/phallotoxins (sometimes referred to as MSDIN family of peptides) are labeled as “cycloamanides” in Kessler and Chooi 2022 and Walton 2018 for simplicity and consistency.

### A 1.4 New insights into KEX2-processed repeat proteins and their position within the RiPP family

Following the publication of Le Marquer et al. 2019, in which a genomic screen was performed to find KEX2-processed repeat proteins in fungi, another paper with the same goal and a larger genomic sample size was published in 2020 (Umemura 2020). This work used a different screening method but came to the same conclusion that almost all fungi display KEPs and found that these KEPs are highly diverse. A total of 1461 fungal strains were screened and 7878 KEPs were detected, of which 2560 had no homologs in other fungi. The rest were

## Chapter 1

classified into 838 KEP types based on core peptide homology, e.g., the KEP core peptide type P-6, which contains N-terminal STE13 recognition motifs and several tryptophan residues, including one at the C-terminus. 22% of all detected KEPs were accompanied by genes encoding putative DUF3328-domain proteins, indicating that the processed peptide products may be cyclic and dikaritin-like. The author speculates that KEPs were present in the common ancestors of fungi and animals, and that some KEPs, like the precursors of dikaritins, accumulated other biosynthetic genes that allowed them to cyclize and posttranslationally modify their products, making them potent toxins (Umemura 2020).

Dikaritins, which are produced by precursors with a KEP-like architecture, are cyclic and highly modified and therefore count as RiPPs (Ding et al. 2016). However, several characterized examples of KEP-derived peptides are linear and, apart from proteolytic cleavage, not further modified (e.g., the  $\alpha$ -pheromone of ascomycetes or candidalysin of *Candida albicans* (Julius et al. 1983; Moyes et al. 2016)). According to the current consensus, these peptides do not count as RiPPs (Vogt and Künzler 2019; Vignolle et al. 2020; Kessler and Chooi 2022). However, it may be useful to reconsider this classification once more KEP-derived peptides have been structurally investigated. The question will arise whether linear peptides with minimal modifications such as pyroglutamations or amidations meet the requirements to be classified as RiPPs. And if not, it might be appropriate to define RiPPs more narrowly, for example as ribosomally synthesized and *extensively* posttranslationally modified peptides, where proteolytic cleavage and minimal modifications implicitly do not qualify a peptide as a RiPP.

## Appendix references

- Allen JG, Hancock GR. 1989. Evidence that phomopsins A and B are not the only toxic metabolites produced by *Phomopsis leptostromiformis*. *J Appl Toxicol*. 9(2):83–89. <https://doi.org/10.1002/jat.2550090203>
- Battilani P, Gualla A, Dall'Asta C, Pellacani C, Galaverna G, Giorni P, Caglieri A, Tagliaferri S, Pietri A, Dossena A, et al. 2011. Phomopsins: An overview of phytopathological and chemical aspects, toxicity, analysis and occurrence. *World Mycotoxin J* [Internet]. 4(4):345–359. <https://doi.org/10.3920/WMJ2011.1302>
- Chafin DR, Guo H, Price DH. 1995. Action of  $\alpha$ -amanitin during pyrophosphorolysis and elongation by RNA polymerase II. *J Biol Chem*. 270(32):19114–19119. <https://doi.org/10.1074/jbc.270.32.19114>
- Condon BJ, Leng Y, Wu D, Bushley KE, Ohm RA, Otilar R, Martin J, Schackwitz W, Grimwood J, MohdZainudin NAI, et al. 2013. Comparative Genome Structure, Secondary Metabolite, and Effector Coding Capacity across Cochliobolus Pathogens. *PLoS Genet*. 9(1). <https://doi.org/10.1371/journal.pgen.1003233>
- Ding W, Liu W-Q, Jia Y, Li Y, van der Donk WA, Zhang Q. 2016. Biosynthetic investigation of phomopsins reveals a widespread pathway for ribosomal natural products in Ascomycetes. *Proc Natl Acad Sci* [Internet]. 113(13):3521–3526. <https://doi.org/10.1073/pnas.1522907113>
- Jiang Y, Ozaki T, Liu C, Igarashi Y, Ye Y, Tang S, Ye T, Maruyama JI, Minami A, Oikawa H. 2021. Biosynthesis of Cyclochlorotine: Identification of the Genes Involved in Oxidative Transformations and Intramolecular O, N-Transacylation. *Org Lett*. 23(7):2616–2620. <https://doi.org/10.1021/acs.orglett.1c00525>
- Johnson RD, Lane GA, Koulman A, Cao M, Fraser K, Fleetwood DJ, Voisey CR, Dyer JM, Pratt J, Christensen M, et al. 2015. A novel family of cyclic oligopeptides derived from ribosomal peptide synthesis of an in planta-induced gene, *gigA*, in *Epichloë* endophytes of grasses. *Fungal Genet Biol* [Internet]. 85:14–24. <https://doi.org/10.1016/j.fgb.2015.10.005>
- Julius D, Blair L, Brake A, Sprague G, Thorner J. 1983. Yeast  $\alpha$  factor is processed from a larger precursor polypeptide: The essential role of a membrane-bound dipeptidyl aminopeptidase. *Cell*. 32(3):839–852. [https://doi.org/10.1016/0092-8674\(83\)90070-3](https://doi.org/10.1016/0092-8674(83)90070-3)
- Kessler SC, Chooi Y-H. 2022. Out for a RiPP: challenges and advances in genome mining of ribosomal peptides from fungi. *Nat Prod Rep*. 39(2):222–230. <https://doi.org/10.1039/d1np00048a>
- Kessler SC, Zhang X, McDonald MC, Gilchrist CLM, Lin Z, Rightmyer A, Solomon PS, Gillian Turgeon B, Chooi YH. 2020. Victorin, the host-selective cyclic peptide toxin from the oat pathogen *Cochliobolus victoriae*, is ribosomally encoded. *Proc Natl Acad Sci U S A*. 117(39):24243–24250. <https://doi.org/10.1073/pnas.2010573117>
- Koiso Y, Li Y, Iwasaki S. 1994. Ustiloxins, antimitotic cyclic peptides from false smut balls on rice panicles caused by *Ustilaginoidea virens*. *J Antibiot (Tokyo)* [Internet]. 47(7):765–773. <https://doi.org/10.7164/antibiotics.47.765>
- Koiso Y, Morisaki N, Yamashita Y, Mitsui Y, Shirai R, Hashimoto Y, Iwasaki S. 1998. Isolation and structure of an antimitotic cyclic peptide, ustiloxin F: Chemical interrelation with a homologous peptide, ustiloxin B. *J Antibiot (Tokyo)* [Internet]. 51(4):418–422.

## Chapter 1

<https://doi.org/10.7164/antibiotics.51.418>

Kono Y, Kinoshita T, Takeuchi S, Daly JM. 1989. Structure of amino acids isolated from hydrolyzed hv-toxin m, a host-specific toxin-related compound produced by *helminthosporium victoriae*. *Agric Biol Chem.* 53(2):505–511. <https://doi.org/10.1080/00021369.1989.10869289>

Lorang J, Kidarsa T, Bradford CS, Gilbert B, Curtis M, Tzeng S-C, Maier CS, Wolpert TJ. 2012. Tricking the Guard : Exploiting Plant Defense for Disease Susceptibility. *Science* (80- ). 338(November):659–662. <https://doi.org/10.1093/plphys/kiaa088>

Luo H, Hallen-Adams HE, Lüli Y, Sgambelluri RM, Li X, Smith M, Yang ZL, Martin FM. 2022. Genes and evolutionary fates of the amanitin biosynthesis pathway in poisonous mushrooms. *Proc Natl Acad Sci U S A.* 119(20). <https://doi.org/10.1073/pnas.2201113119>

Macko V, Wolpert TJ, Acklin W, Jaun B, Seibl J, Meili J. 1985. Characterization of victorin C, the major host-selective toxin from *Cochliobolus victoriae*: structure of degradation products. *Experientia.* 41:1366–1370. <https://doi.org/10.1007/BF01949993>

Le Marquer M, San Clemente H, Roux C, Savelli B, Frei dit Frey N. 2019. Identification of new signalling peptides through a genome-wide survey of 250 fungal secretomes. *BMC Genomics* [Internet]. 20(1):64. <https://doi.org/10.1186/s12864-018-5414-2>

Mayama S, Bordin AP, Morikawa T, Tanpo H, Kato H. 1995. Association of avenalumin accumulation with co-segregation of victorin sensitivity and crown rust resistance in oat lines carrying the Pc-2 gene. *Physiol Mol Plant Pathol.* 46(4):263–274. <https://doi.org/10.1006/pmpp.1995.1021>

Meehan F, Murphy HC. 1946. A New *Helminthosporium* Blight of Oats. *Science* (80- ). 104(2705):413–414. <https://doi.org/10.1126/science.104.2705.413>

Moyes DL, Wilson D, Richardson JP, Mogavero S, Tang SX, Wernecke J, Höfs S, Gratacap RL, Robbins J, Runglall M, et al. 2016. Candidalysin is a fungal peptide toxin critical for mucosal infection. *Nature* [Internet]. 532(7597):64–68. <https://doi.org/10.1038/nature17625>

Nagano N, Umemura M, Izumikawa M, Kawano J, Ishii T, Kikuchi M, Tomii K, Kumagai T, Yoshimi A, Machida M, et al. 2016. Class of cyclic ribosomal peptide synthetic genes in filamentous fungi. *Fungal Genet Biol* [Internet]. 86(January):58–70. <https://doi.org/10.1016/j.fgb.2015.12.010>

Umemura M. 2020. Peptides derived from Kex2-processed repeat proteins are widely distributed and highly diverse in the Fungi kingdom. *Fungal Biol Biotechnol* [Internet]. 7(1). <https://doi.org/10.1186/s40694-020-00100-5>

Van Der Velden NS, Kälin N, Helf MJ, Piel J, Freeman MF, Künzler M. 2017. Autocatalytic backbone N-methylation in a family of ribosomal peptide natural products. *Nat Chem Biol.* 13(8):833–835. <https://doi.org/10.1038/nchembio.2393>

Vignolle GA, Mach RL, Mach-Aigner AR, Derntl C. 2020. Novel approach in whole genome mining and transcriptome analysis reveal conserved RiPPs in *Trichoderma* spp. *BMC Genomics.* 21(1):1–12. <https://doi.org/10.1186/s12864-020-6653-6>

Vogt E, Künzler M. 2019. Discovery of novel fungal RiPP biosynthetic pathways and their application for the development of peptide therapeutics. *Appl Microbiol Biotechnol.*

## Chapter 1

103(14):5567–5581. <https://doi.org/10.1007/s00253-019-09893-x>

Walton J. 2018. The cyclic peptide toxins of amanita and other poisonous mushrooms. New York: Springer US. <https://doi.org/10.1007/978-3-319-76822-9>

Wolpert TJ, Lorang JM. 2016. Victoria Blight, defense turned upside down. *Physiol Mol Plant Pathol* [Internet]. 95:8–13. <https://doi.org/10.1016/j.pmpp.2016.03.006>

Wolpert TJ, Macko V, Acklin W, Jaun B, Arigoni D. 1986. Structure of minor host-selective toxins from *Cochliobolus victoriae*. *Experientia*. 42(11–12):1296–1299. <https://doi.org/10.1007/BF01946431>

Ye Y, Ozaki T, Umemura M, Liu C, Minami A, Oikawa H. 2019. Heterologous production of asperipin-2a: Proposal for sequential oxidative macrocyclization by a fungi-specific DUF3328 oxidase. *Org Biomol Chem*. 17(1):39–43. <https://doi.org/10.1039/c8ob02824a>

Zainudin NAIM, Condon B, De Bruyne L, Van Poucke C, Bi Q, Li W, Hofte M, Gillian Turgeon B. 2015. Virulence, host-selective toxin production, and development of three *cochliobolus* phytopathogens lacking the sfp-type 49-phosphopantetheinyl transferase Ppt1. *Mol Plant-Microbe Interact*. 28(10):1130–1141. <https://doi.org/10.1094/MPMI-03-15-0068-R>

## Chapter 2

Structural and functional analysis of peptides derived from KEX2-processed repeat proteins in agaricomycetes using reverse genetics and peptidomics

Eva Vogt<sup>1</sup>, Lukas Sonderegger<sup>1</sup>, Ying-Yu Chen<sup>1</sup>, Tina Segesseemann<sup>1</sup>, Markus Künzler<sup>1</sup>

<sup>1</sup> ETH Zürich, Department of Biology, Institute of Microbiology, Vladimir-Prelog-Weg 4, CH-8093 Zürich, Switzerland

Manuscript in revision:

American Society for Microbiology (ASM): Microbiology Spectrum

### Contributions:

- Planning and performing main experiments
- Cloning
- Heterologous expression in *P. pastoris*
- Peptide extraction and measurement of fungal samples
- Data analysis
- Knockout strain establishment
- Functional assays
- Writing, revisions

### Abstract

Bioactivities of fungal peptides are of interest for basic research and therapeutic drug development. Some of these peptides are derived from “KEX2-processed repeat proteins” (KEPs), a recently defined class of precursor proteins that contain multiple peptide cores flanked by KEX2 protease cleavage sites. Genome mining has revealed that KEPs are widespread in the fungal kingdom. Their functions are largely unknown. Here, we present the first in-depth structural and functional analysis of KEPs in a basidiomycete. We bioinformatically identified KEP-encoding genes in the genome of the model agaricomycete *Coprinopsis cinerea* and established a detection protocol for the derived peptides by overexpressing the *C. cinerea* KEPs in the yeast *Pichia pastoris*. Using this protocol, which includes peptide extraction and mass spectrometry with data analysis using the search engine Mascot, we confirmed the presence of several KEP-derived peptides in *C. cinerea* as well as in the edible mushrooms *Lentinula edodes*, *Pleurotus ostreatus*, and *Pleurotus eryngii*. While CRISPR-mediated knockout of *C. cinerea* *kep* genes did not result in any detectable phenotype, knockout of *kex* genes caused defects in mycelial growth and fruiting body formation. These results suggest that KEP-derived peptides may play a role in the interaction of *C. cinerea* with the biotic environment and that the KEP-processing KEX proteases target a variety of substrates in agaricomycetes, including some important for mycelial growth and differentiation.

### Importance

Two recent bioinformatics studies demonstrated that KEX2-processed repeat proteins are widespread in the fungal kingdom. However, despite the prevalence of KEPs in fungal genomes, only few KEP-derived peptides have been detected and studied so far. Here, we present a protocol for the extraction and structural characterization of KEP-derived peptides from fungal culture supernatants and tissues. The protocol was successfully used to detect several linear and minimally modified KEP-derived peptides in the agaricomycetes *C. cinerea*, *L. edodes*, *P. ostreatus*, and *P. eryngii*. Our study establishes a new protocol for the targeted search of KEP-derived peptides in fungi, which will hopefully lead to the discovery of more of these interesting fungal peptides and allow a further characterization of KEPs.

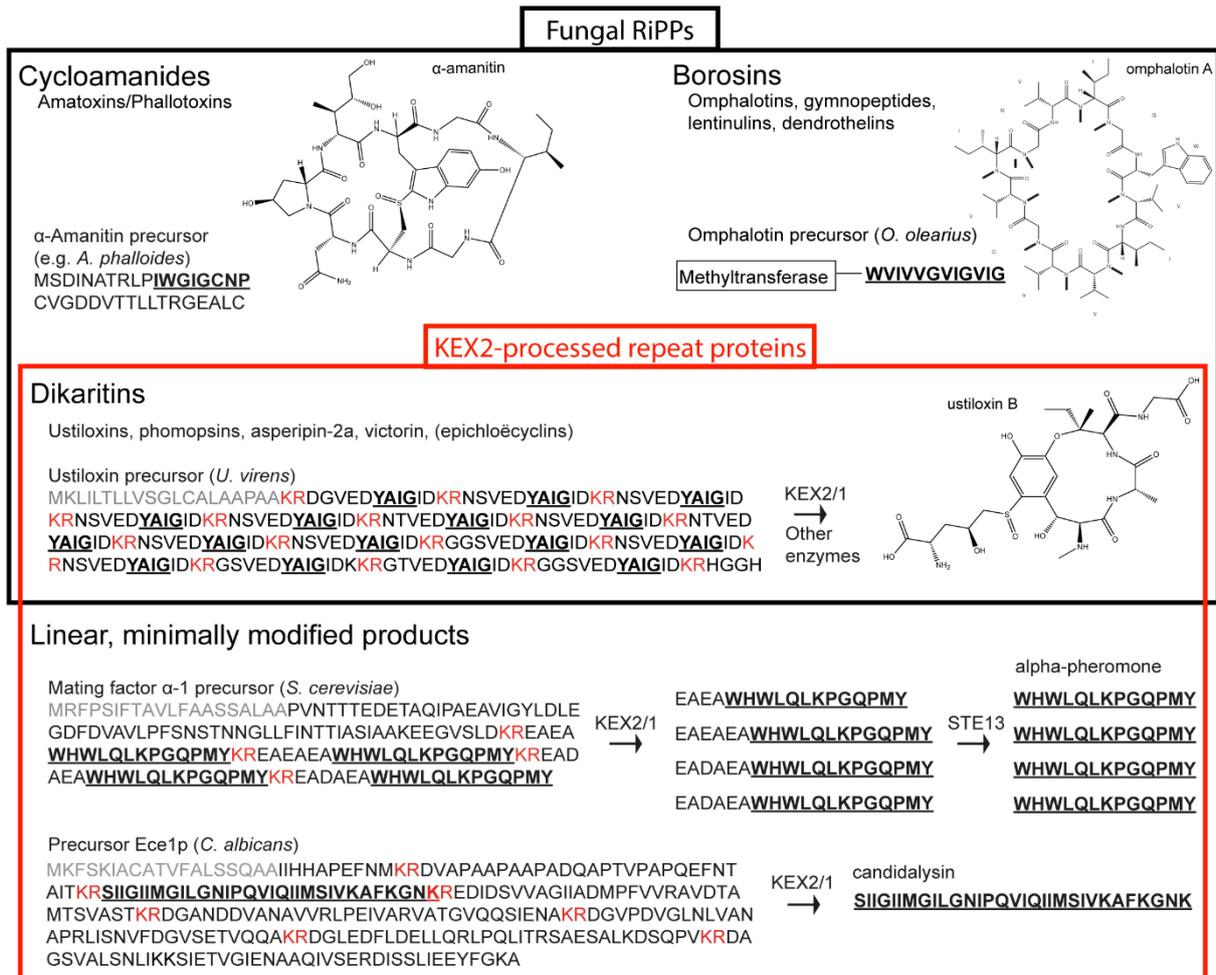
### Introduction

Fungi produce a wide range of bioactive natural products, including peptides. Some peptides have been shown to be of immense value as therapeutics. For example, the use of the non-ribosomal peptides penicillin and cyclosporine as antibiotics and immunosuppressants, respectively, has revolutionized modern medicine (Bills and Gloer 2016). While the synthesis of non-ribosomal peptides (NRPs) relies on large modular enzymes called non-ribosomal peptide synthases (NRPSs) (Süssmuth and Mainz 2017), the sequences of “ribosomally synthesized and posttranslationally modified peptides” (RiPPs) are genetically encoded as parts of precursor proteins. Residues in the core peptide region undergo posttranslational modifications, e.g., cyclization, acetylation, glycosylation, epimerization, and methylation,

followed by release of the core peptide from the precursor protein by proteolytic cleavage (Arnison et al. 2013).

To date, only a few RiPP classes from fungi have been discovered (Fig. 1) (Vogt and Künzler 2019; Kessler and Chooi 2022). The precursors of the two RiPP classes cycloamanides and borosins are processed in the cytoplasm by oligopeptidases of the S9 protease family (Luo et al. 2014; Ramm et al. 2017; Van Der Velden et al. 2017), whereas processing of dikaritins, including ustiloxins, phomopsins, asperipin-2a, victorin, and possibly epichloëcyclins, is mediated by the Golgi-localized kexin endoprotease KEX2 (Umemura et al. 2014; Johnson et al. 2015; Tsukui et al. 2015; Ding et al. 2016; Nagano et al. 2016; Kessler et al. 2020). The dikaritin precursors contain an N-terminal signal peptide for secretion and multiple repeats of core peptides separated by dibasic residues “KR”, “RR”, or “KK” that serve as recognition and cleavage sites for the KEX2 endoprotease. The signal peptide directs the protein to the secretory pathway and is removed by the signal peptidase upon translocation into the lumen of the endoplasmic reticulum (ER). KEX2 is localized in the late Golgi and cleaves the precursor protein at the C-terminal end of the dibasic residue motifs. Subsequently, the C-terminal basic residues are removed from the peptides by the Golgi-localized exopeptidase KEX1 and the peptides are secreted after additional modification in the Golgi (Fig. 1) (Vogt and Künzler 2019).

Recently, two bioinformatic studies showed that precursor proteins with a dikaritin-like architecture are widely distributed in all fungal phyla (Le Marquer et al. 2019; Umemura 2020). These proteins, which contain an N-terminal signal sequence for secretion and repeats of short peptide sequences flanked by KEX2 cleavage sites, were termed “KEX2-processed repeat proteins” (KEPs). Well-characterized representatives of KEP-derived peptides are the aforementioned dikaritins. However, while the dikaritins are extensively posttranslationally modified and therefore considered RiPPs, this does not seem to be the case for many other KEP-derived peptides. The following KEP-derived peptides are linear with no or at most one additional processing step after KEX2/1 cleavage: The cytolytic candidalysin is produced by the opportunistic pathogen *Candida albicans* and is crucial for mucosal infection (Fig. 1) (Moyes et al. 2016), while peptides from the precursor protein Rep1 in *Ustilago maydis* form amyloid-like fibrils (Wösten et al. 1996; Teertstra et al. 2009). The best studied KEP-derived peptides are the  $\alpha$ -pheromones of ascomycetes, e.g. *Saccharomyces cerevisiae*, as shown in Fig. 1. The  $\alpha$ -pheromone precursor contains an STE13 recognition motif, a repetitive dipeptide sequence of XP or XA, where X is often aspartic acid or glutamic acid, located after the KEX2 cleavage site and removed by the dipeptidyl aminopeptidase STE13 after KEX2/1 cleavage (Julius et al. 1983; Jones and Bennett 2011).



**Figure 1:** Illustration of the three major classes of fungal “ribosomally synthesized and post-translationally modified peptides” (RiPPs, black box) and their overlap with peptides derived from KEX2-processed repeat proteins (KEPs, red box). Fungal RiPPs include cycloamanides (Luo et al. 2014) and borosins (Van Der Velden et al. 2017; Quijano et al. 2019; Matabaro et al. 2021), as well as dikaritins (Johnson et al. 2015; Ding et al. 2016; Nagano et al. 2016; Kessler et al. 2020). The dikaritin precursors possess a KEP-like organization with a signal peptide for secretion (labeled in grey) and repetitive units of core peptides that are flanked by dibasic KEX2 cleavage sites (labeled in red). The dibasic residues are cleaved by endoproteinase KEX2 and carboxypeptidase KEX1 and the resulting peptides are further modified and cyclized. According to the current consensus, the cyclic, modified peptides of the dikaritins are considered RiPPs (e.g., ustiloxin B (Tsukui et al. 2014; Umemura et al. 2014)), whereas the linear, minimally modified peptides of other KEPs are not. These KEP-derived peptides include, for example, the  $\alpha$ -pheromones of ascomycetes, in which the dipeptides XP/XA are removed by the dipeptidyl aminopeptidase STE13 (Jones and Bennett 2011), and candidalysin of *C. albicans* (Moyes et al. 2016).

Most KEPs encode peptides of unknown function. Many ascomycetous KEPs likely represent precursors of  $\alpha$ -type mating pheromones, based on the presence of STE13 recognition motifs. Interestingly, the same precursor architecture is also found in many KEPs from basidiomycetes. This is surprising because basidiomycetous mating pheromones are of the  $\alpha$ -type which are synthesized in the cytoplasm, prenylated, and secreted via ABC transporters (Raudaskoski and Kothe 2010; Freiherst et al. 2016). Thus, the identification of  $\alpha$ -type pheromone precursors in basidiomycetes was unexpected and raises the question for the function of these peptides.

## Chapter 2

In the past, fungal RiPPs and KEP-derived peptides were mostly studied using forward genetics, where the peptides were first isolated and characterized and only then the corresponding precursor genes were identified (e.g., ustiloxins, phomopsins, candidalysin, Rep1, omphalotin). In this paper, we established a protocol for the identification of KEP-derived peptides in fungal samples by a combination of reverse genetics and peptidomics. The protocol was developed using the model agaricomycete *Coprinopsis cinerea*, but is applicable to other fungi. In a first step, we bioinformatically screened the predicted proteome of *C. cinerea* for KEP-encoding genes. Second, we expressed six different KEP genes from *C. cinerea* in the yeast *Pichia pastoris* to (1) establish a protocol for extraction and detection of KEP-derived peptides by mass spectrometry in *P. pastoris* culture supernatants known for their low complexity in terms of proteins and peptides (Higgins 1995); (2) obtain an indication that the detected KEPs were indeed processed to peptides, and (3) use the structures of the detected peptides as a reference as to which peptides to expect in *C. cinerea* in terms of peptide length and peptide modifications. In a third step, we applied the established peptide extraction and detection protocols to detect KEP-derived peptides in culture supernatants and tissue samples of the agaricomycetes *C. cinerea*, *Lentinula edodes*, *Pleurotus ostreatus*, and *Pleurotus eryngii*. Finally, we examined the phenotypes of *C. cinerea* *kep* and *kex* knockout strains regarding mycelial growth and fruiting body formation, and tested synthetic KEP-derived peptides for growth inhibition of bacteria. The results of these analyses suggest that individual KEX proteases are redundant in the processing of KEPs but are required for normal mycelial growth and fruiting body formation. Our results did not reveal any specific function for any of the KEP-derived peptides examined.

### Results

#### **Bioinformatic analysis predicts 22 KEX2-processed repeat proteins (KEPs) in *Coprinopsis cinerea***

We screened the predicted proteome of *Coprinopsis cinerea* AmutBmut available on JGI MycoCosm (Nordberg et al. 2014; Muraguchi et al. 2015) for proteins with a KEP-like organization. We used two different methods: In method 1, all proteins were cleaved *in silico* at putative KEX2 cleavage sites, and the resulting fragments of each protein were aligned with BLAST+ (Altschul et al. 1990; Camacho et al. 2009). Proteins lacking a signal sequence for secretion were excluded. Method 2 used the tool RADAR (Heger and Holm 2000) to highlight repeats in signal-peptide containing proteins shorter than 300 amino acids, which were then visually inspected.

The first method using BLAST+ yielded 19 putative *C. cinerea* KEPs. The results of method 2 largely overlapped with the ones of method 1, with two additional hits. Additionally, one protein that had been missed by both methods was manually detected by chance and added to the list of putative KEPs. Of these total 22 hits, two were previously detected by both Le Marquer et al. 2019 and Umemura 2020, and eight were detected only by Umemura 2020. Additionally, Le Marquer et al. and Umemura had found two and seven proteins, respectively, that were not detected in this study, resulting in a total number of 31 putative *C. cinerea* KEPs

(Table S1). In comparison, the KEP screening method of Le Marquer et al. used a similar approach to ours, i.e. proteins with signal sequences were cut *in silico* at KEX2 cleavage sites and a sequence comparison was performed with the resulting peptide fragments (Le Marquer et al. 2019). The main difference is the use of the tool FIMO (Grant et al. 2011) to determine sequence similarity instead of BLAST+. Umemura's approach differed slightly in that this author first screened the genome for proteins with signal sequences and sequence repeats using an in-house script, before checking for the presence of KEX2 cleavage sites.

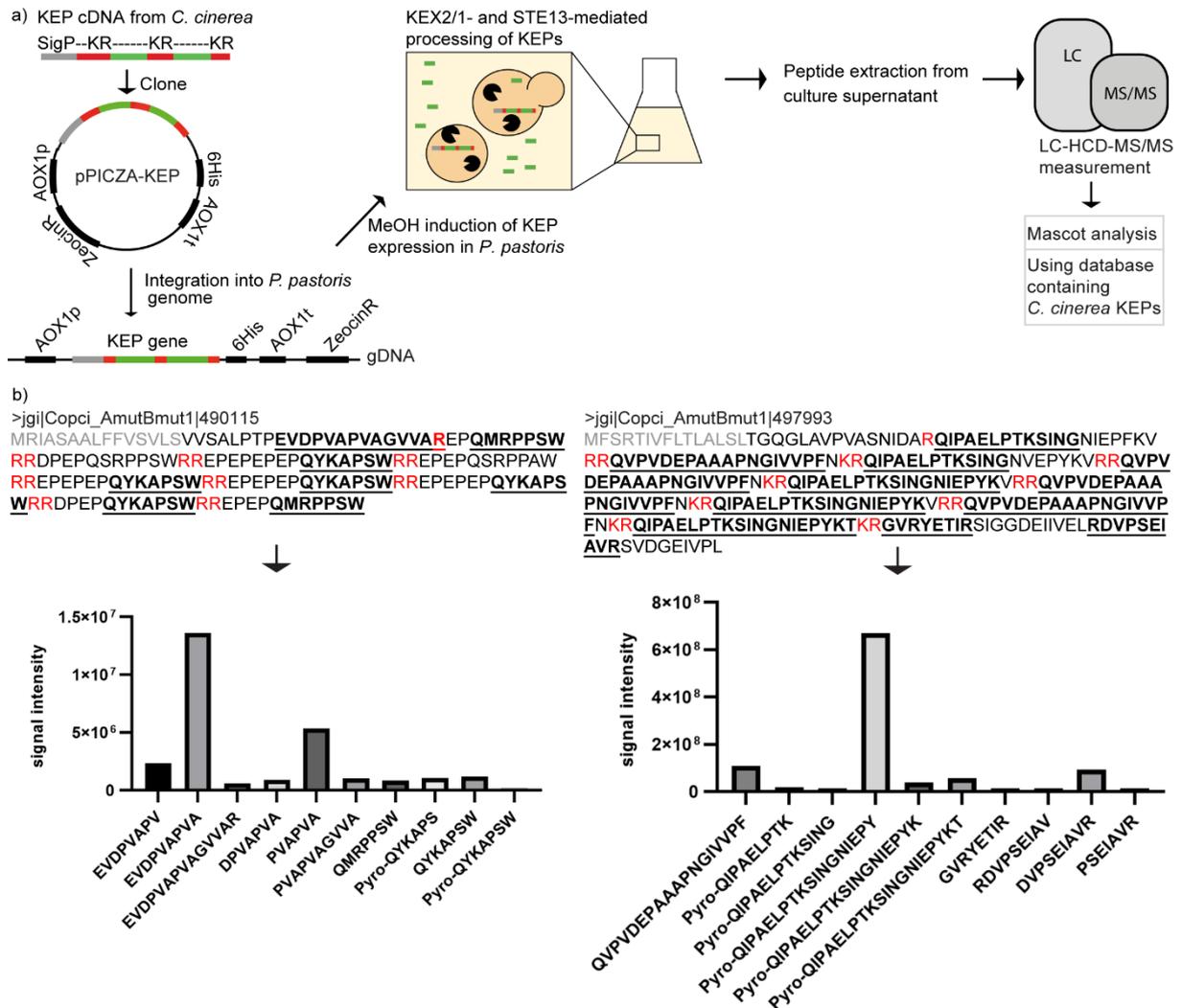
### **Heterologous expression of KEPs from *C. cinerea* in *Pichia pastoris* and detection of the derived peptides in the culture supernatant**

We selected six KEPs from *C. cinerea* for further investigation. All of these proteins are short and contain a high number of KEX2 cleavage sites separating peptide repeats. Previously acquired transcriptome data indicated that some of the corresponding genes were transcribed in the vegetative mycelium and fruiting bodies of *C. cinerea* (Fig. S1) (Muraguchi et al. 2015; Kombrink et al. 2019; Tayyrov et al. 2019). We expressed the *C. cinerea* KEPs in the yeast *P. pastoris*, which produces functional homologs of the endoproteinase KEX2 (Fig. S11a, Yang et al. 2013), the carboxypeptidase KEX1 (Fig. S11c, Dmochowska et al. 1987), and the dipeptidyl aminopeptidase STE13 (Fig. S12a, Hopkins et al. 2014). The function of *P. pastoris* KEX2 and KEX1 on *C. cinerea* proteins and peptides was demonstrated by the successful production of the KEX2-cleaved antimicrobial peptide copsin (CPP1) and its paralog CPP2 from *C. cinerea* in *P. pastoris* (Essig et al. 2014; Kombrink et al. 2019). Thus, we expected that the KEPs from *C. cinerea* would be properly processed in this host and that the derived peptides would accumulate in the culture supernatants of the respective transformants (Fig. 2a).

Expression of the C-terminally His6-tagged *C. cinerea* KEPs in *P. pastoris* was confirmed by immunoblotting whole cell extracts of the respective transformants using anti-His6 antibodies (Fig. S2a). In some of the extracts, a distinct pattern of several smaller proteins was detectable below the intact precursor protein, possibly corresponding to KEX2 cleavage products (Fig. S2b). These bands were of equal intensity, indicating that KEX2 cleaves KEPs stochastically rather than in an N- to C- or C- to N-terminus-directed manner. Peptides were extracted from *P. pastoris* culture supernatant by solid phase extraction (SPE) and the extracts were analyzed by Liquid Chromatography Higher-energy Collisional Dissociation tandem Mass Spectrometry (LC-HCD-MS/MS) and the MS/MS database search tool Mascot (Fig. 2a).

We found a variety of *C. cinerea* KEP-derived peptides in the *P. pastoris* supernatants, as shown in Figure 2b for the expression of the two *C. cinerea* KEPs 490115 and 497993 or in supplemental Figure S3 for KEPs 405832, 434504, 503649, and 426342. Listed are the top 10 peptides detected in the *P. pastoris* supernatant with the highest signal intensity. Some of these peptides were modified at their N-terminal glutamine to pyroglutamate, a modification known to occur spontaneously or enzymatically and to confer increased thermal and proteolytic stability to peptides (Wu et al. 2017). Here, this modification must have occurred spontaneously as the *P. pastoris* genome does not encode a glutaminyl cyclase homolog. The ratio between unmodified peptides and spontaneously formed pyroglutamated species was

shown to be around 9:1 in previous experiments (Wu et al. 2017). Comparison of the signal intensity of pyroglutamated and unmodified peptides in our samples demonstrated that the peptide length and sequence as well as the extraction method influenced the ratio of measured signal intensity between pyroglutamated and unmodified peptides (Fig. S1e-f). As a negative control, we confirmed that no *C. cinerea* KEP-derived peptides could be detected in an induced *P. pastoris* strain transformed with an empty PICZA vector.



**Figure 2:** KEP-derived peptides detected in culture supernatant of *Pichia pastoris* after heterologous expression of KEX2-processed repeat proteins (KEPs) from *Coprinopsis cinerea*. a) Workflow of the experimental procedure. KEP cDNAs from *C. cinerea* were cloned into *P. pastoris* expression vectors, integrated into the alcohol oxidase 1 (AOX1) locus of the *P. pastoris* genome and KEP production was induced by methanol. The supernatant of the liquid cultures was harvested, peptides were extracted and the extracts measured by mass spectrometry (MS). MS/MS spectra analysis was performed using Mascot and a database of *C. cinerea* KEP sequences. b) Peptides derived from expression of *C. cinerea* KEPs 490115 and 497993 in *P. pastoris*. In the respective protein sequences, the KEX2 cleavage sites are shown in red and the signal peptides in grey. The 10 peptides that were detected with the highest signal intensities in the supernatant are listed in the graph below. Their sequences are labeled in bold and underlined in the protein sequence. Many of the detected peptides contain a modification of their N-terminal glutamine residue to pyroglutamate (labeled as “pyro-Q”).

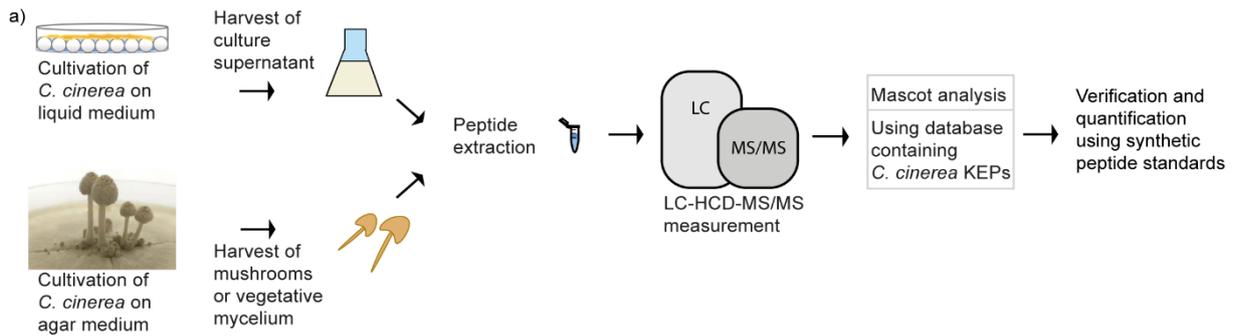
The data for KEP 490115 demonstrates that *P. pastoris* not only successfully cleaved the KEX2 cleavage site RR, but also removed the STE13 recognition motif XA/XP from subsequent peptide products, resulting in the peptides QMRPPSW, pyro-QYKAPS, pyro-QYKAPSW, and QYKAPSW (Fig. 2b). Interestingly, the other detected peptides seem to be derived from the non-repetitive sequence between the N-terminal signal peptide and a single arginine residue. BLAST searches using KEP 490115 reveal many other basidiomycetous KEPs with STE13 recognition motifs and core peptides with a C-terminal tryptophan (Fig. S10a).

Heterologous expression of KEP 497993 resulted in the detection of the peptide pyro-QIPAEELPTKSINGNIEPYKV and four C-terminally truncated cleavage products thereof (Fig. 2b). These results are consistent with the N-terminal pyroglutamate protecting the N-terminus from further degradation, whereas the C-terminus appears to be further trimmed under these conditions. Four other peptides among the top 10 peptide hits were rather short peptides from the C-terminal end of the protein without obvious sequence repetitiveness. In the heterologous expressions of KEPs 490115, 497993, and 434504 (Fig. 2b, Fig. S3), several peptide products appear to result from cleavage after single arginine residues, possibly indicating that not only dibasic but also monobasic residues are cleaved by KEX2 *in vivo*, as previously shown in *S. cerevisiae* (Bevan et al. 1998). Similarly, many products of KEP 503649 end with a lysine residue (Fig. S3). Taken together, these results confirm that our peptide extraction protocol combined with the search engine Mascot is suitable for detecting heterologously expressed KEP-derived peptides from culture supernatants of the ascomycetous yeast *P. pastoris*.

### **KEP-derived peptides are detected in *Coprinopsis cinerea* culture supernatant, vegetative mycelium and fruiting bodies**

We tested our established KEP-peptide detection protocol on culture supernatants and tissues of the agaricomycete *Coprinopsis cinerea*. For this purpose, we cultivated *C. cinerea* on glass beads immersed in minimal medium, a setup that allows the fungus to form a mycelial lawn on the solid surface of the beads while secreting products into the minimal medium (Van Schöll et al. 2006; Essig et al. 2014). We harvested the culture supernatant and extracted the peptides. In parallel, we prepared peptide extracts from vegetative mycelium, premature and mature fruiting body caps, fruiting body stems, and basidiospores from *C. cinerea* cultures on agar-solidified complete medium (YMG). All samples were measured using LC-HCD-MS/MS and the resulting spectra were analyzed using the MS/MS database search tool Mascot on the basis of a database containing *C. cinerea* KEPs. The predicted peptides were confirmed and quantified using synthetic peptide standards (Fig. 3a).

## Chapter 2



b)

```
>gij|Copci_AmutBmut1|497993
MFSRTIVFLTLALSLTGQGLAVPVASNIDARQIPAELPTKSINGNIEPFKVRQVPVDEPAAAPNGIVVPFNKRQIPAELPTKSINGNVEPYKVR
QVPVDEPAAAPNGIVVPFNKRQIPAELPTKSINGNIEPYKVRQVPVDEPAAAPNGIVVPFNKRQIPAELPTKSINGNIEPYKVRQVPVDEPA
AAPNGIVVPFNKRQIPAELPTKSINGNIEPYKKRGVRYETIRSIGGDEIIVELRDVPSEIAVRSVDGEIVPL
```

```
>gij|Copci_AmutBmut1|447393
```

```
MRFSNSVALLVFAFLSTLASPVGALVNRGALDDLERLPARRELVDGRITPPPWRDHVEPTAGITPPPWRDNLPEPTVPVPPRREDSRPPPAWR
PIDSETDSELRPPPAWRPVEAESRPPPAWRPVKREADARPPPAWRPIDISDDDEGPPAPSWRRRETVADAAGVIAPDW
```

	Culture supernatant	Vegetative mycelium	Fruiting body			Basidiospores
			Premature cap	Mature cap	Stem	
Pyro-QVPVDEPA	1.650	0.490	-	-	-	-
Pyro-QIPAE LPTKSINGNIEPY	-	0.010	0.399	-	-	-
Pyro-QIPAE LPTKSINGNIEPYK	-	-	0.006	-	-	-
GPPAPSW	0.003	-	-	-	-	-
DDDPEGPPAPSW	0.062	-	-	-	-	-

c)

*C. cinerea*.497993  
*Lentinula edodes*  
*Pleurotus ostreatus*  
*Pleurotus eryngii*  
*Dendrothele bispora*  
*Dendrothele bispora*  
*Gymnopus luxurians*  
*Termitomyces*.sp.  
*Panaeolus cyanescens*  
*Laccaria bicolor*  
*Hypholoma sublateritium*  
*Hypholoma sublateritium*  
*Crucibulum laeve*  
*Hypsizygus marmoreus*

**Figure 3:** KEP-derived peptides detected in culture supernatants and tissue samples of *Coprinopsis cinerea*. a) Workflow of the experimental procedure. *C. cinerea* was cultivated on glass beads in liquid minimal medium or on agar-solidified complete medium. Peptides were extracted from the culture supernatant, vegetative mycelium, fruiting bodies, and basidiospores and analyzed using LC-HCD-MS/MS and the search engine Mascot. Potential peptide hits were confirmed and quantified using synthetic peptide standards. b) Detected KEP-derived peptides. Three peptides, pyro-QVPVDEPA, pyro-QIPAE LPTKSINGNIEPY, and pyro-QIPAE LPTKSINGNIEPYK, derived from KEP 497993, were detected in the culture supernatant, vegetative mycelium, and fruiting body caps of *C. cinerea*. The two peptides GPPAPSW and DDDPEGPPAPSW, derived from KEP 447393, were detected in the culture supernatant. In the sequence of the precursor protein, the detected peptide sequences are shown in bold and underlined, the KEX2 cleavage sites in red and the signal peptide in grey. In the table below, the abundance of each peptide in the respective fungal tissues is given as peptide weight (ng) per fungal dry weight (mg). c) Conservation of *C. cinerea* KEP 497993 among agaricomycetes. The alignment displays core peptides of the homologs, starting with the first KEX2 cleavage site (indicated by black boxes). The full protein length is indicated on the right.

In *C. cinerea*, we confirmed the presence of three peptides derived from KEP 497993 (pyro-QVPVDEPA, pyro-QIPAE LPTKSINGNIEPY, and pyro-QIPAE LPTKSINGNIEPYK) (Fig. 3b). We

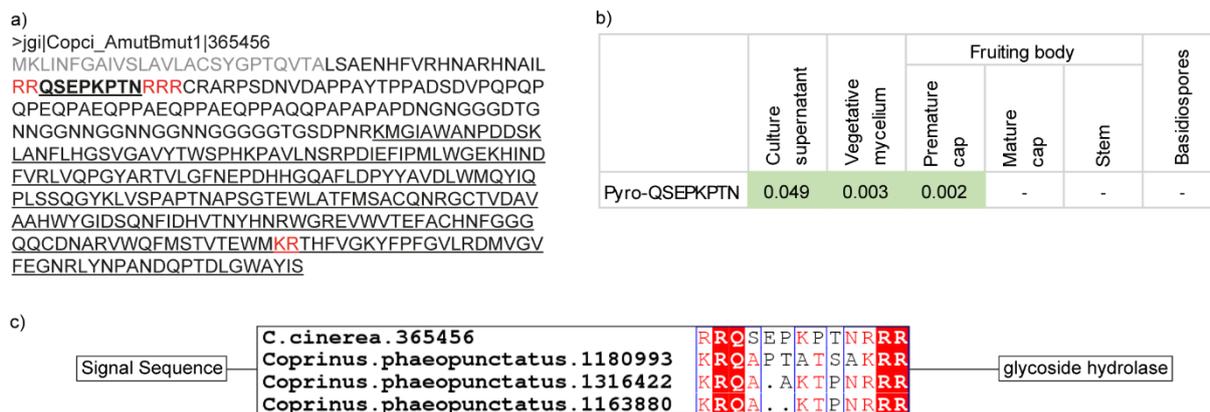
analyzed the presence of the KEP-derived peptides across all culture supernatant and tissue samples to obtain an overview of their expression profile. We found that the presence of each peptide was tissue-specific. The peptide pyro-QVPVDEPA was present in the supernatant of bead assay cultures and in vegetative mycelium, while being absent from fruiting body samples (Fig. 3b). In the culture supernatant, the peptide accumulated even after the fungus had fully covered the plate (4 days) and reached a maximum concentration after 14 days (Fig. S5).

The peptides pyro-QIPAELPTKSINGNIEPY and pyro-QIPAELPTKSINGNIEPYK were primarily detected in the premature fruiting body cap, but were not detectable in the culture supernatant or other fruiting body samples. The precursor KEP 497993, from which all these peptides are derived, showed high transcription in the vegetative mycelium, the fruiting body, and the stem (Fig. S1). This KEP is of special interest due to its high conservation among agaricomycetes. A BLAST search revealed at least 13 homologs in other species, including the edible mushrooms *Lentinula edodes*, *Pleurotus ostreatus*, *Pleurotus eryngii*, and *Hypsizygos marmoreus*, as well as the hallucinogenic mushroom *Panaeolus cyanescens* (Fig. 3c, S10b), indicating a conserved function.

In addition to the peptides derived from KEP 497993, the two peptides GPPAPSW and DDDPEGPPAPSW from the non-repetitive C-terminus of KEP 447393 were detected in the supernatant of *C. cinerea* cultures (Fig. 3b). This precursor is transcribed at very low levels (Fig. S1) and the peptides were detected only in extractions of large volumes of supernatant. A BLAST search in fungal genomes for this KEP reveals two similar KEPs in *Crassisporium funariophilum* and *Coprinellus micaceus* (Fig. S10d)

We performed additional Mascot data analyses by expanding our peptide search to the entire *C. cinerea* proteome instead of just the KEPs and then screened the results for peptides derived from potential KEX2 cleavage sites. Using this approach, we detected the peptide pyro-QSEPKPTN, derived from the KEX2 cleavage of a protein with a predicted C-terminal glycoside hydrolase 128 (GH128) domain (Fig. 4), which belongs to the same family as a  $\beta$ -1,3-glucanase in *L. edodes* and could have a cell-wall remodelling activity (Sakamoto et al. 2011). The peptide was detected in supernatants of bead assay cultures, vegetative mycelium, and fruiting bodies of *C. cinerea* cultivated on agar-solidified medium, in agreement with the protein transcription profile (Fig. S1). Although this protein cannot be classified as a KEP due to the lack of peptide repeats, the unique protein organization with a signal-sequence, a single short peptide flanked by KEX2 cleavage sites and a C-terminal glycoside hydrolase makes this protein and its derived peptide an interesting candidate for further investigation. A BLAST search with the protein reveals three similar proteins in *Coprinus phaeopunctatus* that also contain a signal peptide for secretion, a glycosyl hydrolase domain, and KEX2 cleavage sites flanking peptides with the sequences QAPTATSA, QAAKTPN, and QAKTPN (Fig. 4c). Five other potential homologs are found in *Cyathus striatus*, *Crucibulum laeve*, *Laccaria amethystina*, *Lyophyllum atratum*, and *Asterophora parasitica* with the sequences ASTS, ASTS, GTTA, AANP, and ASTG, respectively, flanked by KEX2 cleavage sites (Fig. S10c).

## Chapter 2



**Figure 4:** The KEX2-processed peptide pyro-QSEPKPTN is detected in culture supernatants and tissue samples of *C. cinerea*. The peptide is produced by KEX2-cleavage of the glycoside hydrolase 128 (GH128) domain-containing protein 365456. a) Sequence of the precursor protein. The detected core peptide is shown in bold and underlined, KEX2 cleavage sites in red, the signal peptide in grey and the glycoside hydrolase domain in underlined. b) Detection pattern of pyro-QSEPKPTN. The peptide abundance in the respective fungal tissues is given as peptide weight (ng) per fungal dry weight (mg). The peptide was confirmed and quantified using synthetic peptide standards. c) Schematic illustration of *C. cinerea* 365456 and homolog proteins in *Coprinus phaeopunctatus*. Each protein contains a signal sequence, KEX2 cleavage sites flanking a core peptide and a predicted glycoside hydrolase domain.

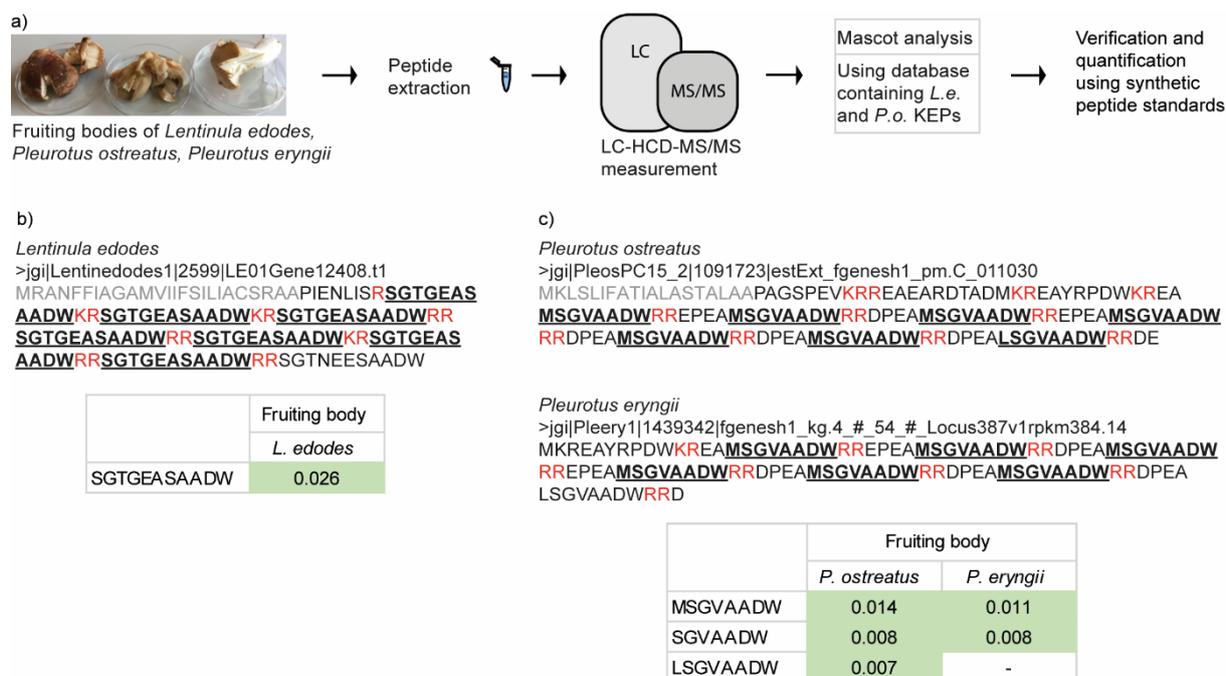
Many of the confirmed peptides in *C. cinerea* are pyroglutamated. While we detected both the unmodified and pyroglutamated peptide species in the culture supernatant of KEP expressions in *P. pastoris* (Fig. S3e-f), this does not seem to be the case in *C. cinerea* where no unmodified peptide species could be confirmed (Fig. S4g). Since the genome of *C. cinerea*, unlike *P. pastoris*, encodes a putative glutaminyl cyclase (QC) (Fig. S13), it seems likely that the formation of these pyroglutamated peptides in *C. cinerea* is an enzyme-mediated process. QCs are widely spread in the fungal kingdom (Wu et al. 2017). So far, only the two QCs of *Neurospora crassa* were studied in depth and found to be localized in the ER (Wu et al. 2017). A BLAST alignment of the suspected *C. cinerea* QC with fungal and human homologs interestingly demonstrates that its sequence is more similar to the human QCs than to the ones of *N. crassa* (Fig. S13b). The two human QCs contain either a signal peptide and are secreted into the extracellular space, or a signal anchor and remain localized in the Golgi (Xu et al. 2021). The similarity in sequence and the presence of a signal sequence implies that the *C. cinerea* QC is secreted into the extracellular space like the corresponding human homolog (Fig. S13b-c).

### KEP-derived peptides are detected in the fruiting bodies of shiitake, oyster mushroom and king oyster mushroom

Having demonstrated the presence of KEP-derived peptides in *C. cinerea*, we applied our KEP-peptide detection protocol to other agaricomycetes. Fruiting bodies of *L. edodes*, *P. ostreatus*, and *P. eryngii* were purchased from a local mushroom farm and peptides were extracted and analyzed by LC-HCD-MS/MS. For the identification of the KEP-derived peptides, the Mascot search engine in conjunction with the list of KEPs of *L. edodes* and *P. ostreatus* according to Le Marquer et al. 2019 and Umemura 2020 was used (Fig. 5a).

## Chapter 2

The KEP 2599 from *L. edodes* encodes a sevenfold repeat of the peptide SGTGEASAADW, the presence of which was confirmed in fruiting body extracts (Fig. 5b). This KEP has four homologs encoding nearly identical peptides in *Auriculariales sp.*, *Gymnopus luxurians*, *Gymnopus confluens*, and *Rhodocollybia butyracea* (Fig. S10e). In *P. ostreatus*, we confirmed the presence of three peptide products, MSGVAADW, SGVAADW, and LSGVAADW, derived from KEP 1091723. In *P. eryngii*, we detected the peptides MSGVAADW and SGVAADW from the homolog KEP 1439342 (Fig. 5c). Remarkably, the latter KEPs contain not only the classical KEX2 cleavage sites flanking repetitive peptide sequences, but STE13 recognition motifs, repetitive units of dipeptides XA or XP, which are removed by dipeptidyl aminopeptidases STE13 in ascomycetes (Jones and Bennett 2011). A BLAST search demonstrates that STE13 homologs are present not only in *C. cinerea*, *L. edodes*, *P. ostreatus*, and *P. eryngii*, but in many more basidiomycetes (Fig. S12). Interestingly, all detected KEP-derived peptides from *L. edodes* and *P. ostreatus/eryngii* contain an identical “-AADW” sequence at their C-termini. A BLAST search of these KEPs reveals some interesting homologs with STE13 recognition motifs and peptide cores with a similar C-terminus (Fig. S10f).



**Figure 5:** KEP-derived peptides detected in fruiting bodies of *Lentinula edodes* (shiitake), *Pleurotus ostreatus* (oyster mushroom), and *Pleurotus eryngii* (king oyster mushroom). a) Workflow of the experimental procedure. Peptides were extracted from purchased fruiting bodies followed by LC-HCD-MS/MS measurements and analysis using Mascot. Peptides were confirmed and quantified using synthetic peptide standards. b) Detected KEP-derived peptide from *L. edodes* and from c) *P. ostreatus* and *P. eryngii*. The respective *Pleurotus* precursor proteins contain putative STE13 recognition motifs, repetitive units of XP/XA dipeptides following dibasic residues. The absence of a signal peptide in KEP 1439342 of *P. eryngii* is likely due to misannotation of the protein at the N-terminus (Fig. S8). The peptide abundance in the respective strains is given as peptide weight (ng) per fungal dry weight (mg).

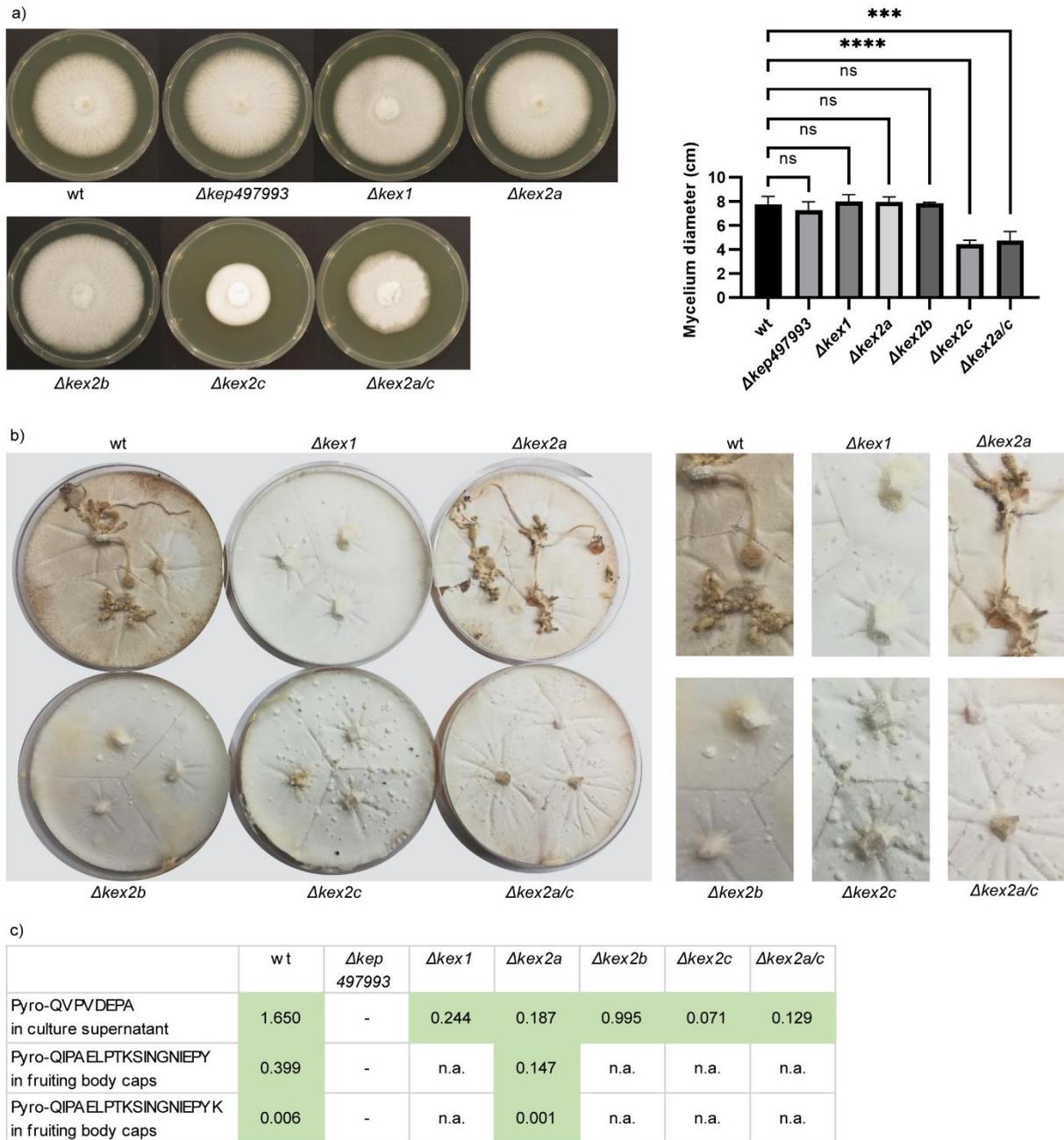
**Knockout of *kex2* and *kex1* genes in *C. cinerea* affects mycelial growth and fruiting body formation**

We established a protocol to increase the efficiency of gene knockouts in *C. cinerea* using cotransformation of preassembled gRNA/Cas9 complexes, a strategy previously established in the fruiting body forming basidiomycete *Schizophyllum commune* (Jan Vonk et al. 2019). In the course of establishing the protocol, we compared the knockout efficiency between the original knockout procedure, in which protoplasts derived from the vegetative mycelium of a *C. cinerea* AmutBmut  $\Delta ku70$  strain are transformed with a plasmid carrying a gene-specific knockout cassette (Nakazawa et al. 2011; Stöckli et al. 2017; Stöckli et al. 2019), and the newly established protocol in which protoplasts are transformed with both the plasmid and preassembled gRNA/Cas9 complexes. The use of CRISPR increased knockout efficiency in three different genes severalfold (Fig. S7a).

Using the CRISPR-assisted knockout protocol, we generated *C. cinerea* knockout strains for the six KEPS 405832, 434504, 490115, 497993, 503649, and 426342. We confirmed that the peptides pyro-QVPVDEPA, pyro-QIPAELPTKSINGNIEPY, and pyro-QIPAELPTKSINGNIEPYK, which were detected in samples from the *C. cinerea* wildtype strain, were no longer detectable in the  $\Delta kep 497993$  mutant strain (Fig. 6c, S9b), confirming that the detected peptides were derived from this protein. All *kep* knockout strains showed normal mycelial growth when compared with the original wildtype strain macroscopically (Fig. S7n). No differences in quantity or morphology of formed fruiting bodies could be detected. In order to test a possible function of the peptides in the defense of *C. cinerea* against bacterial competitors, we performed bacterial growth inhibition tests using synthetic peptides and *B. subtilis* 168, *B. subtilis* NCBI 3610, *Micrococcus luteus*, *Staphylococcus aureus*, and *E. coli* BL21. No antibacterial activity was observed (Fig. S14).

Since the *kep* knockouts showed no apparent phenotype, we generated knockout strains for the three KEX2 homologs of *C. cinerea*: KEX2a (502579), KEX2b (406374), and KEX2c (448165). A sequence alignment between the KEX2 homologs of *C. cinerea*, *P. pastoris*, and *S. cerevisiae* demonstrates that of the three *C. cinerea* KEX2 paralogs, KEX2c has the highest sequence similarity to *S. cerevisiae* and *P. pastoris* KEX2 (Fig. S11b). There are at least two potential homologs for KEX1 in the *C. cinerea* genome, and we knocked out the one with higher sequence similarity to *S. cerevisiae* KEX1 (437675) (Fig. S11c). We also generated a double knockout strain for KEX2a and KEX2c. These *kex* knockout strains were analysed for the presence of previously confirmed KEP-derived peptides and for growth and developmental phenotypes.

## Chapter 2



**Figure 6:** Phenotypic analysis of *C. cinerea* *kep* and *kex* knockout strains. a) Mycelial growth of knockout strains on solid agar medium after four days with statistical analysis. \*Denotes a significant difference of the mean of a group compared to the mean of a control condition (Dunnett's multiple-comparison test); bars represent the median value of three biological replicates, error bars represent the 95% confidence interval; ns, not significant; \*\*\*\*,  $P < 0.0001$  (ANOVA Table S11). b) Fruiting body formation by *kex* knockout strains. The mycelia were exposed to a night-day cycle for 40 days. During this period, the wt and  $\Delta kex2a$  strains formed multiple fruiting bodies, whereas the other strains did not. The cropped pictures on the right show the mycelium in double magnification. c) Detection pattern of KEP 497993-derived peptides. The peptide abundance in the respective strains is given as peptide weight (in ng) per fungal dry weight (mg). The presence of the peptide pyro-QVPVDEPA was analyzed in the bead assay supernatant of wildtype (wt) and knockout strains. The same analysis was performed for the peptides pyro-QIPAE LPTKSINGNIEPY and pyro-QIPAE LPTKSINGNIEPYK in premature fruiting body caps of wildtype,  $\Delta kex2a$ , and  $\Delta kep$  497993 strains. The other *kex* knockout strains were deficient in fruiting body formation (labeled n.a.).

The KEP-derived peptides pyro-QVPVDEPA, pyro-QIPAELOPTKSINGNIEPY, and pyro-QIPAELOPTKSINGNIEPYK, as well as the KEX-processed pyro-QSEPKPTN, were still present in all the *kex* knockout strains including the  $\Delta kex2a/c$  double knockout strain (Fig. 6c, S9). The strain  $\Delta kex2c$  and the double knockout strain  $\Delta kex2a/c$  grew more slowly on agar-solidified complete medium compared to the wildtype strain, while the other *kex* knockout strains grew normally (Fig. 6a). These results suggest that the function of the three *C. cinerea* KEX2 proteases is redundant for some substrates, e.g. the KEPs, but not for all substrates. Consistent with this conclusion, fruiting body formation was abolished in the  $\Delta kex1$ ,  $\Delta kex2b$ ,  $\Delta kex2c$ , and  $\Delta kex2a/c$  strains, whereas the  $\Delta kex2a$  strain and all *kep* knockout strains fruited normally (Fig. 6b). The fruiting-deficient mutant strains formed many hyphal knots with a diameter of 1-3 mm, but they did not develop further even after weeks of incubation. Taken together, these results suggest a role for KEX2c in mycelial growth and for KEX1, KEX2b, and KEX2c in fruiting body formation. KEX2a and the six investigated KEPs do not appear to be involved in these processes.

### Discussion

In this work, we have established a protocol for the detection of KEP-derived peptides in a variety of fungal tissues. These secreted, ribosomally synthesized fungal peptides represent a very interesting class of natural products, although they are only minimally modified compared to more complex fungal RiPPs. KEPs have been shown to be widely distributed in fungi (Le Marquer et al. 2019; Umemura 2020) and the sequence diversity of predicted KEP-derived peptides is huge. Therefore, the analysis of KEPs has a high potential for the discovery of novel peptides with interesting bioactivities. The already characterized examples function as mating factors (Julius et al. 1983), virulence factors (Moyes et al. 2016) and mutualism modulators (Pellegrin et al. 2019) and are, thus, mostly involved in biotic interactions of the respective fungi.

Using our detection protocol, we confirmed the presence of several KEP-derived peptides in *C. cinerea*, *L. edodes*, and *P. ostreatus/eryngii* (Fig. 3, 5). Interestingly, we observed that certain peptides, although derived from the same precursor protein, were exclusively detected in specific tissues, indicating that there is a tissue-specific processing of the precursor protein or of the derived peptides. This tissue-specificity seems unrelated to KEX protease expression levels, as all of these processing enzymes are expressed at approximately the same levels in the respective tissues (Fig. S1). Some of the detected KEP-derived peptides likely originate from STE13-mediated processing, just as the  $\alpha$ -pheromones of ascomycetes (Fig. 5). To our knowledge, this is the first indication of STE13 dipeptidyl aminopeptidase activity in a basidiomycete, and the first demonstration that these  $\alpha$ -pheromone-like peptide precursor proteins are not only present in the predicted proteome of basidiomycetes but are also actively processed to peptides. Further studies are needed to elucidate the function of these peptides in basidiomycetes.

Our Mascot-based approach to finding KEP-derived peptides seems to be best suited for the discovery of minimally modified linear peptides. For complex products, a more targeted

approach might be useful, in which the genomic neighborhood of KEP-candidates is examined for the presence of modifying enzymes such as the DUF3328-domain-containing enzymes, which would then be indicative of cyclized peptides (Umemura 2020). Furthermore, the detected peptides might not be the actual bioactive products. In this context, it is interesting to note that once a KEP-derived peptide is detected, there is a high probability that cleaved versions of the same peptide will also be detected: For example, in *P. ostreatus* and *P. eryngii* we found both MSGVAADW and SGVAADW (Fig. 5). In addition, we were unable to detect KEP-derived peptides of certain KEPs, although they should have been highly transcribed in the respective tissue of *C. cinerea*. Possible explanations include low peptide abundances and unexpected posttranslational modifications.

The established knockout strains for six different KEPs in *C. cinerea* did not show any apparent phenotype, indicating that the function of these peptides is not critical for axenic growth and development under the used growth conditions or is rather relevant in the communication with the biotic environment. KEP-derived peptides were present in all *kex* knockouts (Fig. 6), suggesting that there is redundancy among the three proteases regarding KEP processing. The apparent redundancy regarding KEPs does not seem to apply to all targets of KEX proteases, as the individual *kex2b*, *kex2c*, and *kex1* knockout strains displayed aberrant growth and development phenotypes.

We observed defects in mycelial growth in *C. cinerea kex2c* knockout strains. Mycelial growth defects or abnormal hyphal morphology are typical phenotypes of *kex2* knockouts in ascomycetes. Examples are hyperbranching, thickened cell walls, and upregulation of cell wall stress response genes in *Aspergillus niger* (Jalving et al. 2000; van Leeuwe et al. 2020), hyperbranching and disordered cell integrity signalling in *A. oryzae* (Mizutani et al. 2004), hypersensitivity to fungicides targeting the cell membrane in *C. glabrata* (Bader et al. 2001), and lack of hyphae formation in *C. albicans* (Newport and Agabian 1997). The ascomycetous strains in which *kex2* knockouts were established possess only one *kex2* gene in contrast to the three *kex2* genes of *C. cinerea* and other agaricomycetes (Li et al. 2017), implying a diversification of KEX2 function in agaricomycetes. Consistent with the high similarity of *C. cinerea* KEX2c with *S. cerevisiae* and *P. pastoris* KEX2 (Fig. S11b), only knockouts of the *kex2c* gene reduced mycelial growth in *C. cinerea*. Inhibition of fruiting body formation was observed in knockout strains of *kex1*, *kex2b*, and *kex2c* (Fig. 6c). Since no *kex* knockout strains have been established in mushroom-forming fungi before, this is a previously undescribed phenotype. Both  $\Delta kex2b$  and  $\Delta kex2c$  exhibited fruiting deficiencies, implying that they cannot substitute for each other's function and are both required for efficient fruiting body formation via processing of one or more targets. The fruiting deficiency of  $\Delta kex1$  suggests that the same target(s) may be involved and that they require processing by KEX2b, KEX2c, and KEX1. In contrast, defects in mycelial growth were observed only in the  $\Delta kex2c$  strain, suggesting that a different target is affected that does not require processing by KEX1.

KEX2 and KEX1 proteases process not only KEPs but also several cell wall modulating enzymes and structural proteins that pass the secretory pathway, e.g., by removing N-terminal or C-terminal propeptides (Cooper and Bussey 1989; Conesa et al. 2001; Bader et al. 2008). In

## Chapter 2

previous experiments in ascomycetes, knockout of *kex2* resulted in a phenotype similar to knockouts of KEX2-processed cell wall remodelling enzymes or components, e.g., increased sensitivity to fungicides targeting the cell wall (Martínez et al. 2004; Firon et al. 2007). Therefore, we hypothesise that the observed phenotypes of *C. cinerea kex* knockout strains are due to an impairment in protein processing rather than to an impairment in KEP processing. This hypothesis is also supported by the observation that KEP-derived peptides are still present in *kex* knockouts.

Further work is needed to decipher the function of the KEP-derived peptides and the molecular basis of the phenotypes of *kex* knockout mutations in *C. cinerea*. The specificity of basidiomycetous KEX proteases should be further investigated, and a bioinformatic screen could be performed to find potential KEX2 substrates with putative carbohydrate-active enzymatic domains. According to reports in ascomycetes, such proteins could be responsible for the observed  $\Delta kex2$  phenotypes. Protein 365456 that contains KEX2 cleavage sites and a C-terminal glycoside hydrolase domain may fall in this class.

### Conclusion

Our results suggest that fungi, including the agaricomycetes *Coprinopsis cinerea*, *Pleurotus ostreatus*, *Pleurotus eryngii*, and *Lentinula edodes*, secrete a variety of linear and minimally modified ribosomal peptides. These peptides derive from secreted precursor proteins (KEPs) that are processed in the Golgi apparatus by the endoproteinase KEX2 and the carboxypeptidase KEX1. While knockouts of *C. cinerea kex* genes resulted in defects in mycelial growth and differentiation, knockouts of *kep* genes showed no phenotype under axenic conditions. These results suggest that the KEP-processing enzymes target a wide variety of substrates, some of which play a role in mycelial growth and differentiation, but that the investigated KEP-derived peptides are likely to rather play a role in the interaction of the mushroom with the biotic environment.

### Contributions

EV performed the main experiments, analyzed and interpreted the data and wrote the manuscript draft. LS performed the *Pichia* transformation for heterologous expression of the KEPs 405832, 434504, 497993, 503649, 426342, established the described knockout procedure for *C. cinerea* and performed the knockout of the *kex2a* gene. YC performed the peptide extractions from fruiting bodies of *C. cinerea* AmutBmut and performed the knockout of the *kex2b* gene. TS performed the KEP screening of the *C. cinerea* proteome. MK proposed and supervised the project. All authors have read and approved the manuscript.

### Acknowledgements

We thank Antje Dittmann from the Functional Genomic Center (Zürich, Switzerland) for her support in the mass spectrometry measurements, Marc Flachsmann for cloning the plasmid containing the *kex1* knockout cassette, and Annageldi Tayyrov for his supervision of the bioinformatic KEP screen, preparing the *C. cinerea* cDNA and the manual annotation of the *C. cinerea* proteome.

### Funding

This work was supported by the Swiss National Science Foundation (Grant No. 31003A-173097) and ETH Zürich (Grant No. ETH-25 20-1).

### Conflict of interest

The authors declare that they have no conflict of interest.

### Methods

#### *Strains*

A full list of all used strains can be found in the supplementary Table S3. *C. cinerea* samples were extracted from strains AmutBmut or AmutBmut  $\Delta$ ku70. The same  $\Delta$ ku70 strain was used to establish all knockout cell lines and is therefore labelled as wildtype (wt) throughout the paper. Peptides derived from the KEP 497993 were confirmed in the strain AmutBmut  $\Delta$ ku70 for comparison to the  $\Delta$ ku70-based knockout strains, all other peptides were confirmed in AmutBmut. Fresh fruiting bodies of *Lentinula edodes* (strain 4312, Sylvan, USA), *Pleurotus ostreatus* (strain P24/HK35, Sylvan, USA), and *Pleurotus eryngii* (strain 3066, Sylvan, USA) were purchased from a grocery store and processed immediately. *Pichia pastoris* strain GS115 was ordered from Invitrogen. Molecular cloning was performed using *E. coli* DH5 $\alpha$ , protein expression of Cas9 was performed using *E. coli* BL21.

#### *Genomic screening for KEX2-processed repeat proteins in C. cinerea*

The predicted proteome of *Coprinopsis cinerea* AmutBmut pab1.2 was downloaded from JGI in early 2018 («*Copci\_AmutBmut1\_GeneModels\_FrozenGeneCatalog\_20160912\_aa.fasta*») (Nordberg et al. 2014, Muraguchi et al. 2015) (<https://mycocosm.jgi.doe.gov>). Two different methods were tested to screen for KEPS. For the first method a Java script was written that cleaved the proteome *in silico* at the dibasic residues KR, RR, KK, and RK. These four dibasic cleavage sites were chosen based on the processing of KEP-homolog neuropeptide precursors. After *in silico* cleavage, the cleavage products of individual proteins were compared using the protein BLAST+ command line application (version 2.7.1). (Altschul et al. 1990; Camacho et al. 2009). In a first screen, sequences were considered to be similar if they reached a percentage identity of 70%. Similarity was assessed only for cleaved fragments with a length difference of fewer than 30 amino acid residues. To account for short repetitive sequences for which a percentage identity threshold of 70% was too strict, a second screen was performed in which short cleaved fragments (<20 residues) with a length difference of fewer than 8 residues were assessed for their percentage identity and considered similar if they scored at least 40%. Results from both screens were combined and filtered: Excluded were proteins with no signal sequence (assessed using SignalP 4.0, Petersen et al. 2011), fewer than two cleavage sites and protein lengths greater than 800 amino acids, as 96% of known neuropeptide precursors were below that threshold (Kang et al. 2019). KEPS had to contain at least one repetitive partial sequence that was similar to at least two other partial sequences. The java script is deposited on zenodo (<https://doi.org/10.5281/zenodo.7034114>). In the second screening method, only

proteins with a signal sequence and shorter than 300 amino acids were analyzed. Repetitive sequences in these proteins were highlighted using the rapid automated detection and alignment of repeats tool (RADAR) (Heger and Holm 2000). The output was visually inspected and proteins with a KEP-like architecture were manually picked.

#### *Heterologous expression of KEX2-processed repeat proteins in Pichia pastoris*

All gene and primer sequences are listed in the supplementary Table S4 and S5. RNA was isolated from *C. cinerea* mycelium using the Norgen RNA extraction kit (Norgen Biotek Corporation, Canada) according to the manufacturer's protocol. cDNA was prepared using the Transcriptor first strand cDNA synthesis kit (Roche) using oligo-dT primers. The cDNA was used for amplifying intron-less KEP gene sequences via PCRs, which were then ligated into pGEM T-easy vectors (Promega, USA). This was done for the KEPs 405832, 434504, 497993, 503649, 426342. The coding region of KEP 490115 could not be amplified from cDNA, possibly due to highly repetitive nucleotide regions in the peptide cores and was instead ordered from Twist Bioscience (USA) in a codon optimized form. All plasmids were transformed into *E. coli* DH5 $\alpha$  and isolated using miniprep kits (Qiagen, Germany). The coding regions were subsequently amplified from the plasmids using primers with a 5' restriction site overhang. These amplified products were digested and ligated into pPICZA vectors (ThermoFisher Scientific, USA). The stop codons of the *kep* genes were removed in the PCR primers to allow for C-terminal fusion of the proteins to the myc and 6xHis tags on the PICZA vector. The PICZA vectors were amplified in *E. coli* DH5 $\alpha$ , purified using miniprep, the sequence confirmed via Sanger sequencing, the plasmids linearized with the restriction enzyme PmeI and then used for transformation of *Pichia pastoris* GS115 by electroporation with 1.2 kV of charging voltage using a BioRad MicroPulser electroporator. Positive clones were selected for on YPD plates (1% (w/v) yeast extract, 2% (w/v), peptone, 2% (w/v) glucose, 2% (w/v) agar) containing 0.2 mg/mL zeocin. Successful integration of pPICZA at the AOX1-locus was tested via colony PCR. For heterologous expression of *C. cinerea* KEP genes, *P. pastoris* transformants were cultivated until they reached OD 2 in buffered minimal glycerol medium (BMGH) (100 mM potassium phosphate (pH 6.0), 1.34% yeast nitrogen base (YNB, with ammonium sulfate and without amino acids) (w/v),  $4 \times 10^{-5}$ % biotin (w/v), 0.004% histidine (w/v), 1% glycerol (v/v)). The culture was spun down at 2000 rcf for 10 min, the pellet resuspended in buffered minimal methanol medium (BMMH) (100 mM potassium phosphate (pH 6.0), 1.34%, YNB (w/v),  $4 \times 10^{-5}$ % (w/v), biotin (w/v), 0.004% histidine (w/v), 0.5% methanol (v/v)), spun down again, and then resuspended to an OD of 1. The culture was then incubated for three days at 30°C under constant shaking, with an addition of 0.5% (v/v) methanol to the medium every 24 hours. Afterwards, the culture was spun down with 3000 rcf for 10 min. The cell pellets were washed with 2x PBS, lysed using glass beads using a Thermo FastPrep FP120 cell disruptor and used for an immunoblot to confirm successful expression of the tagged KEP. His-tagged proteins on the blots were detected using anti-His antibodies (Qiagen, Germany). The supernatant was harvested, filtered through a 10 kDa Amicon filter, and 10 ml were subsequently used for extraction of peptides.

## Chapter 2

### *Cultivation of C. cinerea and harvest of fungal samples*

Cryostocks (oidia) of *Coprinopsis cinerea* were revived on YMG 1.5% agar plates (0.4% (w/v) yeast extract, 1% (w/v) malt extract, 0.4% (w/v) glucose) and allowed to grow for four days at 37°C in a dark aerated box with wet tissue paper. A mycelial plug was cut from the edge of the mycelium and transferred to a YMG plate to be incubated again for four days. Bead assays with liquid minimal medium (per liter of medium: 5 g glucose, 2 g asparagine, 50 mg adenine sulfate, 1 g KH<sub>2</sub>PO<sub>4</sub>, 2.3 g Na<sub>2</sub>HPO<sub>4</sub> (anhydrous), 0.3 g Na<sub>2</sub>SO<sub>4</sub>, 0.5 g C<sub>4</sub>H<sub>12</sub>N<sub>2</sub>O<sub>6</sub>, 40 g thiamine-HCl, 0.25 g MgSO<sub>4</sub>, and 5 mg p-aminobenzoic acid) were then set up as previously reported (Van Schöll et al. 2006; Essig et al. 2014): Three agar plugs were cut from the edge of a *C. cinerea* mycelial colony and transferred to a petri dish filled with approx. one layer of sterile borosilicate glass beads (5 mm) and 15 ml of minimal medium. The fungus was allowed to grow for four days under the usual growth conditions (37°C, darkness) until harvest of the liquid medium. For the comparison between different knockout strains, 125 µl of media were harvested. For the time course samples, 125 µl of media were taken continuously over 28 days from the same culture. For the rest of the samples, 1 ml of supernatant were harvested. The liquid medium was filtered through a 10 kDa filter (Amicon, Merck, Germany), snap-frozen and then used for peptide extraction. For the harvest of dry vegetative mycelium, the fungus was cultivated by transferring three agar plugs to an YMG agar plate covered with one sheet of sterile cellophane. The fungus was then allowed to grow on top of the cellophane for four days at 37°C in darkness, before the mycelium was scratched from the cellophane and snap-frozen. For the harvest of fruiting bodies, thick YMG agar plates (containing 35 ml of medium) were inoculated with three agar plugs of *C. cinerea* and incubated for four days at 37°C in darkness (6 days for the growth deficient  $\Delta$ kex2c strains). The plates were then moved to a fruiting body room with 25°C, 85% humidity and a day-night cycle of 12 h of light and 12 h of darkness. After approximately 10 days (AmutBmut) or 20 days (AmutBmut  $\Delta$ ku70) under these conditions, fruiting bodies started to form on the plates. The fruiting bodies were harvested when their stipes were fully expanded and their cap was still half-closed in the middle of the “night” phase (premature caps and stipes) or four hours later when the caps had fully opened and started to darken (mature caps). In addition, basidiospores were isolated by harvesting multiple mature dark caps and their “ink”, the black liquid within the area. The samples were resuspended in water, filtered through glass wool funnels and the basidiospores were pelleted by centrifuging. All tissue samples were flash-frozen in liquid nitrogen and dried in a Speedvac (SPD111V, ThermoFisher Scientific, USA) before peptide extraction.

### *Peptide extraction*

For peptide extractions from liquid medium, including *C. cinerea* bead assay samples and *P. pastoris* supernatant samples, samples were either purified using Oasis HLB 1 cc vac cartridges (10 mg sorbent, 30 µm, Waters, US), Sep-Pak 1 cc cartridges (10 mg sorbent, 30 µm, Waters, US) or first dried using speedvac and then purified using SP3 (Single-pot solid-phase-enhanced sample preparation, Hughes et al. 2019). For initial peptide extractions of *P. pastoris* samples, extraction of KEP 497993 heterologous expression samples was tested using Sep-Pak C18 and Oasis HLB cartridges. The concentration of extracted peptides was highest with HLB cartridges,

## Chapter 2

so they were used for extraction of all *Pichia* samples. For peptide extractions from fungal tissues, 30 mg of fruiting body caps, stipes and mycelium or 10 mg of basidiospores were soaked in 1.5 ml of methanol for 1 hour at 37°C while gently shaking. The methanol was then moved to a new tube and evaporated by speedvac. Other tested parameters included different solvents like water and acetonitrile and different soaking temperatures of room temperature, 37°C and 99°C (for water). After drying, the samples were reconstituted in ddH<sub>2</sub>O and passed through a 10 kDa Amicon filter. The peptides were then either purified using HLB columns, or dried using a speedvac for SP3. Solid phase extraction using HLB and Sep-Pak cartridges were done according to the manufacturer's protocol. In brief, the cartridges were conditioned with 1 ml of methanol, equilibrated with 1 ml of ddH<sub>2</sub>O and loaded with the aqueous samples. The cartridges were washed with 1 ml of 5% methanol and then eluted with 1 ml of 100% methanol. The samples were dried in a speedvac and resuspended in MS buffer (3% acetonitrile, 0.1% formic acid). For SP3 peptide extraction, dried samples were resuspended in 25 µl of 50 mM HEPES buffer and processed following the official guidelines with modifications (Hughes et al. 2019). Magnetic hydrophilic and hydrophobic beads (Sera-Mag SpeedBead Carboxylate-Modified Magnetic Particles, GE Life, Germany) were mixed 1:1 and washed three times with ddH<sub>2</sub>O at a concentration of 10 µg/µl and then resuspended at a concentration of 5 µg/µl. The samples were then gently mixed with 15 µl of bead suspension. 800 µl of acetone were added and the samples carefully mixed. Samples were incubated at room temperature for 8 min, then moved to a magnetic rack for 2 min. The supernatants were removed and the beads washed twice with 200 µl of 95% acetonitrile. Then, the beads were resuspended in 25 µl of 2% DMSO for peptide elution. The samples were sonicated for 1 min and incubated at room temperature for 5 min. The tubes were moved to the magnetic rack and the supernatant transferred to a new tube. The samples were dried by speedvac and redissolved in MS buffer. The approximate peptide concentration was determined using nanodrop (Witec, Switzerland) and samples were diluted to concentrations of 2 mg/ml.

### *LC-HCD-MS/MS measurements and analysis by Mascot*

Samples were measured using liquid chromatography higher-energy collisional dissociation tandem mass spectrometry (LC-HCD-MS/MS) on a Q Exactive HF (Thermo Scientific) coupled to an ACQUITY UPLC M-Class system. Chromatographic separation was performed using trapped elution using a nanoEASE Symmetry C18 100 A column (5 µm, 1/PK 180 µm x 20 mm), followed by analytical elution using a nanoEase HSS C18 T3 100 A column (1.8 µm, 1 P/K 75 µm x 250 mm). Solvent A was water with 0.1% formic acid (v/v) and solvent B was acetonitrile with 0.1% formic acid. For the initial measurement of *Pichia pastoris* supernatant samples as well as *C. cinerea* supernatant and tissue samples, the samples were measured using an untargeted approach. Peptides were eluted with a linear gradient from 5% to 30% solvent B in 90 min at a flow rate of 0.3 µl/min, followed by a linear gradient from 30% to 95% in 5 min and a return to 5% in 10 min. Full MS was measured using data-dependent acquisition with a resolution of 120 000, automated gain control (AGC) target  $3 \times 10^6$ , scan range 350 to 1500 m/z and maximum IT 50 ms. MS/MS spectra acquisition was performed using a resolution of

## Chapter 2

30 000, AGC target  $10^5$ , maximum IT 50 ms. The minimum AGC was  $4.5 \times 10^3$  and the dynamic exclusion time 30 s.

In the subsequent Mascot search, the mgf file containing all MS/MS spectra of the measurement was searched against databases containing the *C. cinerea* KEPs, the *C. cinerea* proteome (*Coprinopsis cinerea* AmutBmut *pab1.2*, from <https://mycocosm.jgi.doe.gov>, manually annotated) or *L. edodes* and *P. ostreatus* KEPs (Le Marquer et al. 2019; Umemura 2020). The search settings were: no enzyme cleavage, error windows of 10 ppm for peptide tolerance and 0.05 Da for MS/MS tolerance, measured on a Q-exactive instrument and with an enabled decoy database. Variable modifications that were tested included N-terminal pyroglutamation, C-terminal amidation, hydroxylation, oxidation and methylation. To avoid false-positives, a false-discovery-rate was calculated, based on matches from a “decoy” database containing only randomized and inverted sequences. No true matches are expected from the decoy database, so the number of false matches is a good estimate of the number of false positives from the target database. In the analysis of *P. pastoris* samples, we set the false-discovery rate in Mascot to 1%, meaning that the risk of a false positive being called significant is lower than 1%. For all the potential KEP-derived peptide candidates from *C. cinerea*, *L. edodes*, *P. ostreatus*, and *P. eryngii*, the Mascot hits were manually screened for interesting peptide candidates. Peptides with fewer than six residues were excluded, while modified peptides or peptides with a promising sequence (e.g., a repetitive core region flanked by KEX2 cleavage sites) were chosen for a verification using synthetic peptides. We ordered synthetic peptides at Genscript Biotech (USA, pyro-QVPVDEPA) and EZBiolab (USA, all other peptides) at 85% purity for verification. We measured the synthetic peptides with a concentration of 100 ng/ml. An initial comparison of the spectra of the synthetic peptides to the fungal samples was done using the software Skyline 20.2.0.343 (MacCoss Lab Software, USA) on the basis of the (isotope) dot product ((i)dotp). For a final confirmation of peptide presence and for figures, the retention times and MS/MS spectra were compared manually in the software Xcalibur (Thermo Scientific, USA). Once the peptides were confirmed, samples were remeasured using a targeted approach with an inclusion list containing the masses of the peptides of interest. For the time course samples (Fig. S5) and the knockout comparison samples (Fig. 6), a shorter linear gradient was used from 5% to 30% acetonitrile in 50 min, followed by a linear gradient from 30% to 95% in 5 min and a return to 5% in 10 min. We performed BLAST searches in the respective fungal species using the confirmed peptide sequences to ensure that they do not arise from cleavage of another protein. The abundance of peptides in fungal samples was quantified by comparing the peak area values of the extracted ion chromatogram to the respective synthetic peptides.

### *Generation of C. cinerea knockout strains*

For the establishment of knockout strains in *C. cinerea*, we used the para-aminobenzoic acid synthase-encoding gene (*pab1*) that allows growth of *pab1*-mutant strains AmutBmut and AmutBmut  $\Delta ku70$  on para-aminobenzoic acid-deficient medium as selection marker. In order to avoid recombination of the marker with the *C. cinerea pab1* locus, we used a previously constructed, heterologous *pab1* cassette (Pcpab1) consisting of the *Phanerochaete*

## Chapter 2

*chryso sporium pab1* gene under control of the *Agaricus bisporus gpdII* promoter and *P. chryso sporium mnp* terminator regions (Stöckli et al. 2017). For each knockout, we cloned a template plasmid with the *Pcpab1* cassette flanked by 750-1200 nt homology arms of the respective gene, as described by Stöckli et al., 2017. The homology arms were first amplified from *C. cinerea* gDNA using primer pairs containing a short *Pcpab1* overhang, while the *Pcpab1* cassette was amplified from the template plasmid using primer pairs containing a short overhang of the homology arms. All three PCR products, meaning the 5' homology arm, the *Pcpab1* cassette, and the 3' homology arm, were then assembled in a one-step overlap extension PCR and ligated into a pGEM T-easy vector (Promega, USA). crRNAs were designed to cut proximal to the 5' and 3' regions of the genes. The design tool CRISPR RGEN Tools (<http://www.rgenome.net/cas-designer>) was used to design and test the crRNAs for potential off-target effects and the crRNAs were ordered at Integrated DNA Technologies (IDT, USA) and hybridized with the tracrRNA (IDT, USA) to form the gRNA. The *S. pyogenes* Cas9 enzyme (Zuris et al. 2015) was heterologously produced in *E. coli* BL21 via the plasmid pET-NLS-Cas9-6xHis from Addgene and the enzyme was isolated and purified using Ni-NTA beads. For all *C. cinerea* transformations we used the strain AmutBmut *pab1.2 Δku70* that contains a knockout of the *ku70* gene which reduces the frequency of non-homologous end-joining (Nakazawa and Honda 2015; Jan Vonk et al. 2019). The protoplasting of *C. cinerea* mycelia was done as described in Nakazawa et al. 2010). In short, *C. cinerea* was grown on cellophane-covered YMG agar for three days at 37°C in darkness, until the mycelium was scraped off, mixed with YMG medium and blended using a Waring laboratory blender (US). The blended mycelium was then grown in shaking flasks in liquid YMG medium for two days. The mycelium was harvested and incubated for 1 h with protoplasting enzyme mix containing cellulase ("Onozuka R-10", Serva, Germany) and chitinase ("C-6137", Sigma, USA). The protoplasts were then filtered through funnels with Miracloth, mixed with the template DNA and pre-assembled gRNA/Cas9 complexes as described by Jan Vonk et al. 2019 and transformed in a heat shock incubation at room temperature. The transformation mixtures were then plated on para-aminobenzoic acid-deficient agar plates and growing colonies were picked and transferred to new plates. These colonies were then screened in PCR reactions and a final confirmation of knockout was done via southern blotting. The KO strains of *kex2b* and *kex1* were produced using oidial protoplasts, meaning that plates covered with *C. cinerea* mycelia were moved to constant light for three days and the oidia were rinsed and filtered from the plates through glass-wool funnels. They were treated with the same enzyme mix as the mycelium for generation of protoplasts. The *kex2c* KO and the double KO strain of *kex2a* and *kex2c* (labelled  $\Delta kex2a/c$ ) were both produced in a transformation reaction where the gRNAs and template plasmids of KEX2a, KEX2b and KEX2c were all added to the same reaction and the subsequent colonies were tested for KO of all three KEX proteases.

### *Southern blots*

As a final confirmation that the constructed knockout strains no longer contained the wildtype gene, we performed southern blots on all knockout strains according to the protocol by Wälti et al., 2006. In short, DIG-labelled hybridization probes (labeled dNTPs from Roche,

## Chapter 2

Switzerland) for one homology arm of the respective genes were amplified in a PCR reaction from the respective template plasmid that was used for transformation of the wildtype strain. These PCR samples were run on an agarose gel and the band of the correct size was purified. 5 µg of gDNA isolated from the respective knockout strains were digested with the chosen restriction enzyme and the samples were run on an agarose gel, from which they were transferred to a Hybond-N Nylon membrane (Cytiva, USA). The membranes were hybridized with the DIG probes, stained with an anti-DIG-antibody (Roche, Switzerland), and imaged using CDP-Star substrate (Roche, Switzerland).

### *Phenotypic comparisons of knockout strains*

For a comparison of mycelial growth speed, knockout strains were cultivated on YMG plates with a single inoculum and imaged after four days of incubation at 37°C in darkness. For a comparison of peptide presence between *kex* and *kep* knockout strains, the strains were cultivated in bead assays as described earlier. Samples were taken after four days. For the strains  $\Delta kex2c$  and  $\Delta kex2a/c$ , an additional set of plates was prepared whose supernatant was harvested after six days to account for the slower growth speed of these strains. For a comparison of fruiting body formation, thick YMG agar plates were inoculated with three agar plugs. After four days of growth, plates were almost covered with mycelium (with the exception of strains  $\Delta kex2c$  and  $\Delta kex2a/c$  which were left to grow for total six days), were sliced with a scalpel to induce the formation of more fruiting bodies and moved to a fruiting room with 25°C, 85% humidity and a day-night cycle of 12 h of light and 12 h of darkness. Six replicates of plates were left in the fruiting room for continued observation for several weeks, without any fruiting bodies forming for  $\Delta kex1$ ,  $\Delta kex2b$ ,  $\Delta kex2c$ , and  $\Delta kex2a/c$ .

### *Data availability*

The mass spectrometry data have been deposited to the ProteomeXchange Consortium via the PRIDE (Perez-Riverol et al. 2019) partner repository with the dataset identifier PXD036934. A list of the individual datasets is provided in Table S11.

## References

- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J Mol Biol.* 215(3):403–410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2)
- Arnison PG, Bibb MJ, Bierbaum G, Bowers AA, Bugni TS, Bulaj G, Camarero JA, Campopiano DJ, Challis GL, Clardy J, et al. 2013. Ribosomally synthesized and post-translationally modified peptide natural products: Overview and recommendations for a universal nomenclature. *Nat Prod Rep.* 30(1):108–160. <https://doi.org/10.1039/c2np20085f>
- Bader O, Krauke Y, Hube B. 2008. Processing of predicted substrates of fungal Kex2 proteinases from *Candida albicans*, *C. glabrata*, *Saccharomyces cerevisiae* and *Pichia pastoris*. *BMC Microbiol.* 8:1–16. <https://doi.org/10.1186/1471-2180-8-116>
- Bader O, Schaller M, Klein S, Kukula J, Haack K, Mühlischlegel F, Korting HC, Schäfer W, Hube B. 2001. The KEX2 gene of *Candida glabrata* is required for cell surface integrity. *Mol Microbiol.* 41(6):1431–1444. <https://doi.org/10.1046/j.1365-2958.2001.02614.x>
- Bevan A, Brenner C, Fuller RS. 1998. Quantitative assessment of enzyme specificity in vivo: P2 recognition by Kex2 protease defined in a genetic system. *Proc Natl Acad Sci U S A.* 95(18):10384–10389. <https://doi.org/10.1073/pnas.95.18.10384>
- Bills GF, Gloer JB. 2016. Biologically Active Secondary Metabolites from the Fungi. *Microbiol Spectr* [Internet]. 4(6):1–32. <https://doi.org/10.1128/microbiolspec.FUNK-0009-2016>
- Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. BLAST+: Architecture and applications. *BMC Bioinformatics.* 10:1–9. <https://doi.org/10.1186/1471-2105-10-421>
- Conesa A, Weelink G, Van den Hondel CAMJJ, Punt PJ. 2001. C-terminal propeptide of the *Caldariomyces fumago* chloroperoxidase: An intramolecular chaperone? *FEBS Lett.* 503(2–3):117–120. [https://doi.org/10.1016/S0014-5793\(01\)02698-9](https://doi.org/10.1016/S0014-5793(01)02698-9)
- Cooper A, Bussey H. 1989. Characterization of the yeast KEX1 gene product: a carboxypeptidase involved in processing secreted precursor proteins. *Mol Cell Biol.* 9(6):2706–2714. <https://doi.org/10.1128/mcb.9.6.2706-2714.1989>
- Ding W, Liu W-QQ, Jia Y, Li Y, Van Der Donk WA, Zhang Q. 2016. Biosynthetic investigation of phomopsins reveals a widespread pathway for ribosomal natural products in Ascomycetes. *Proc Natl Acad Sci U S A* [Internet]. 113(13):3521–3526. <https://doi.org/10.1073/pnas.1522907113>
- Dmochowska A, Dignard D, Henning D, Thomas DY, Bussey H. 1987. Yeast KEX1 gene encodes a putative protease with a carboxypeptidase B-like function involved in killer toxin and  $\alpha$ -factor precursor processing. *Cell.* 50(4):573–584. [https://doi.org/10.1016/0092-8674\(87\)90030-4](https://doi.org/10.1016/0092-8674(87)90030-4)
- Essig A, Hofmann D, Münch D, Gayathri S, Künzler M, Kallio PT, Sahl H, Wider G, Schneider T, Aebi M. 2014. Copsin, a Novel Peptide-based Fungal Antibiotic Interfering with the Peptidoglycan Synthesis \*. 289(50):34953–34964. <https://doi.org/10.1074/jbc.M114.599878>
- Firon A, Aubert S, Iraqui I, Guadagnini S, Goyard S, Prévost MC, Janbon G, D’Enfert C. 2007. The SUN41 and SUN42 genes are essential for cell separation in *Candida albicans*. *Mol Microbiol.* 66(5):1256–1275. <https://doi.org/10.1111/j.1365-2958.2007.06011.x>

## Chapter 2

- Freiherst D, Fowler TJ, Bartholomew K, Raudaskoski M, Horton JS, Kothe E. 2016. 13 The Mating-Type Genes of the Basidiomycetes. *Growth, Differ Sex.*:329–349. [https://doi.org/10.1007/978-3-319-25844-7\\_13](https://doi.org/10.1007/978-3-319-25844-7_13)
- Grant CE, Bailey TL, Noble WS. 2011. FIMO: Scanning for occurrences of a given motif. *Bioinformatics.* 27(7):1017–1018. <https://doi.org/10.1093/bioinformatics/btr064>
- Heger A, Holm L. 2000. Rapid automatic detection and alignment of repeats in protein sequences. *Proteins Struct Funct Genet.* 41(2):224–237. [https://doi.org/10.1002/1097-0134\(20001101\)41:2<224::AID-PROT70>3.0.CO;2-Z](https://doi.org/10.1002/1097-0134(20001101)41:2<224::AID-PROT70>3.0.CO;2-Z)
- Higgins DR. 1995. Overview of Protein Expression in *Pichia pastoris*. *Curr Protoc Protein Sci.* 2(1):1–16. <https://doi.org/10.1002/0471140864.ps0507s02>
- Hopkins D, Gomathinayagam S, Lynaugh H, Stadheim TA, Hamilton SR. 2014. Elimination of diaminopeptidase activity in *Pichia pastoris* for therapeutic protein production. *Appl Microbiol Biotechnol.* 98(6):2573–2583. <https://doi.org/10.1007/s00253-013-5468-7>
- Hughes CS, Moggridge S, Müller T, Sorensen PH, Morin GB, Krijgsveld J. 2019. Single-pot, solid-phase-enhanced sample preparation for proteomics experiments. *Nat Protoc.* 14(1):68–85. <https://doi.org/10.1038/s41596-018-0082-x>
- Jalving R, Van De Vondervoort PJI, Visser J, Schaap PJ. 2000. Characterization of the kexin-like maturase of *Aspergillus niger*. *Appl Environ Microbiol.* 66(1):363–368. <https://doi.org/10.1128/AEM.66.1.363-368.2000>
- Jan Vonk P, Escobar N, Wösten HAB, Lugones LG, Ohm RA. 2019. High-throughput targeted gene deletion in the model mushroom *Schizophyllum commune* using pre-assembled Cas9 ribonucleoproteins. *Sci Rep.* 9(1):1–8. <https://doi.org/10.1038/s41598-019-44133-2>
- Johnson RD, Lane GA, Koulman A, Cao M, Fraser K, Fleetwood DJ, Voisey CR, Dyer JM, Pratt J, Christensen M, et al. 2015. A novel family of cyclic oligopeptides derived from ribosomal peptide synthesis of an in planta-induced gene, *gigA*, in *Epichloë* endophytes of grasses. *Fungal Genet Biol [Internet].* 85:14–24. <https://doi.org/10.1016/j.fgb.2015.10.005>
- Jones SK, Bennett RJ. 2011. Fungal mating pheromones: Choreographing the dating game. *Fungal Genet Biol [Internet].* 48(7):668–676. <https://doi.org/10.1016/j.fgb.2011.04.001>
- Julius D, Blair L, Brake A, Sprague G, Thorner J. 1983. Yeast  $\alpha$  factor is processed from a larger precursor polypeptide: The essential role of a membrane-bound dipeptidyl aminopeptidase. *Cell.* 32(3):839–852. [https://doi.org/10.1016/0092-8674\(83\)90070-3](https://doi.org/10.1016/0092-8674(83)90070-3)
- Kang J, Fang Y, Yao P, Li N, Tang Q, Huang J. 2019. NeuroPP: A Tool for the Prediction of Neuropeptide Precursors Based on Optimal Sequence Composition. *Interdiscip Sci Comput Life Sci [Internet].* 11(1):108–114. <https://doi.org/10.1007/s12539-018-0287-2>
- Kessler SC, Chooi Y-H. 2022. Out for a RiPP: challenges and advances in genome mining of ribosomal peptides from fungi. *Nat Prod Rep.* 39(2):222–230. <https://doi.org/10.1039/d1np00048a>
- Kessler SC, Zhang X, McDonald MC, Gilchrist CLM, Lin Z, Rightmyer A, Solomon PS, Gillian Turgeon B, Chooi YH. 2020. Victorin, the host-selective cyclic peptide toxin from the oat pathogen *Cochliobolus victoriae*, is ribosomally encoded. *Proc Natl Acad Sci U S A.* 117(39):24243–24250. <https://doi.org/10.1073/pnas.2010573117>

- Kombrink A, Tayyrov A, Essig A, Stöckli M, Micheller S, Hintze J, van Heuvel Y, Dürig N, Lin C wei, Kallio PT, et al. 2019. Induction of antibacterial proteins and peptides in the coprophilous mushroom *Coprinopsis cinerea* in response to bacteria. *ISME J* [Internet]. 13(3):588–602. <https://doi.org/10.1038/s41396-018-0293-8>
- van Leeuwe TM, Arentshorst M, Forn-Cuní G, Geoffrion N, Tsang A, Delvigne F, Meijer AH, Ram AFJ, Punt PJ. 2020. Deletion of the *aspergillus niger* pro-protein processing protease gene *kexB* results in a pH-dependent morphological transition during submerged cultivations and increases cell wall chitin content. *Microorganisms*. 8(12):1–19. <https://doi.org/10.3390/microorganisms8121918>
- Li J, Gu F, Wu R, Yang JK, Zhang KQ. 2017. Phylogenomic evolutionary surveys of subtilase superfamily genes in fungi. *Sci Rep*. 7(September 2016):1–15. <https://doi.org/10.1038/srep45456>
- Luo H, Hong SY, Sgambelluri RM, Angelos E, Li X, Walton JD. 2014. Peptide macrocyclization catalyzed by a prolyl oligopeptidase involved in  $\alpha$ -amanitin biosynthesis. *Chem Biol* [Internet]. 21(12):1610–1617. <https://doi.org/10.1016/j.chembiol.2014.10.015>
- Le Marquer M, San Clemente H, Roux C, Savelli B, Frei dit Frey N. 2019. Identification of new signalling peptides through a genome-wide survey of 250 fungal secretomes. *BMC Genomics* [Internet]. 20(1):64. <https://doi.org/10.1186/s12864-018-5414-2>
- Martínez AI, Castillo L, Garcerá A, Elorza M V., Valentín E, Sentandreu R. 2004. Role of Pir1 in the construction of the *Candida albicans* cell wall. *Microbiology*. 150(10):3151–3161. <https://doi.org/10.1099/mic.0.27220-0>
- Matabaro E, Kaspar H, Dahlin P, Bader DLV, Murar CE, Staubli F, Field CM, Bode JW, Künzler M. 2021. Identification, heterologous production and bioactivity of lentinulin A and dendrothelin A, two natural variants of backbone N-methylated peptide macrocycle omphalotin A. *Sci Rep*. 11(1):1–12. <https://doi.org/10.1038/s41598-021-83106-2>
- Mizutani O, Nojima A, Yamamoto M, Furukawa K, Fujioka T, Yamagata Y, Abe K, Nakajima T. 2004. Disordered cell integrity signaling caused by disruption of the *kexB* gene in *Aspergillus oryzae*. *Eukaryot Cell*. 3(4):1036–1048. <https://doi.org/10.1128/EC.3.4.1036-1048.2004>
- Moyes DL, Wilson D, Richardson JP, Mogavero S, Tang SX, Wernecke J, Höfs S, Gratacap RL, Robbins J, Runglall M, et al. 2016. Candidalysin is a fungal peptide toxin critical for mucosal infection. *Nature* [Internet]. 532(7597):64–68. <https://doi.org/10.1038/nature17625>
- Muraguchi H, Umezawa K, Niikura M, Yoshida M, Kozaki T, Ishii K, Sakai K, Shimizu M, Nakahori K, Sakamoto Y, et al. 2015. Strand-specific RNA-seq analyses of fruiting body development in *Coprinopsis cinerea*. *PLoS One*. 10(10):1–23. <https://doi.org/10.1371/journal.pone.0141586>
- Nagano N, Umemura M, Izumikawa M, Kawano J, Ishii T, Kikuchi M, Tomii K, Kumagai T, Yoshimi A, Machida M, et al. 2016. Class of cyclic ribosomal peptide synthetic genes in filamentous fungi. *Fungal Genet Biol* [Internet]. 86(January):58–70. <https://doi.org/10.1016/j.fgb.2015.12.010>
- Nakazawa T, Ando Y, Kitaaki K, Nakahori K, Kamada T. 2011. Efficient gene targeting in  $\delta$ Cc.ku70 or  $\delta$ Cc.lig4 mutants of the agaricomycete *coprinopsis cinerea*. *Fungal Genet Biol* [Internet]. 48(10):939–946. <https://doi.org/10.1016/j.fgb.2011.06.003>
- Nakazawa T, Honda Y. 2015. Absence of a gene encoding cytosine deaminase in the genome

## Chapter 2

of the agaricomycete *Coprinopsis cinerea* enables simple marker recycling through 5-fluorocytosine counterselection. *FEMS Microbiol Lett.* 362(15):1–7. <https://doi.org/10.1093/femsle/fnv123>

Nakazawa T, Tatsuta Y, Fujita T, Nakahori K, Kamada T. 2010. Mutations in the *Cc.rmt1* gene encoding a putative protein arginine methyltransferase alter developmental programs in the basidiomycete *Coprinopsis cinerea*. *Curr Genet.* 56(4):361–367. <https://doi.org/10.1007/s00294-010-0307-1>

Newport G, Agabian N. 1997. KEX2 influences *Candida albicans* proteinase secretion and hyphal formation. *J Biol Chem [Internet].* 272(46):28954–28961. <https://doi.org/10.1074/jbc.272.46.28954>

Nordberg H, Cantor M, Dusheyko S, Hua S, Poliakov A, Shabalov I, Smirnova T, Grigoriev I V., Dubchak I. 2014. The genome portal of the Department of Energy Joint Genome Institute: 2014 updates. *Nucleic Acids Res.* 42(D1):26–31. <https://doi.org/10.1093/nar/gkt1069>

Pellegrin C, Daguerre Y, Ruytinx J, Guinet F, Kemppainen M, Frey NF dit, Puech-Pagès V, Hecker A, Pardo AG, Martin FM, Veneault-Fourrey C. 2019. *Laccaria bicolor* MiSSP8 is a small-secreted protein decisive for the establishment of the ectomycorrhizal symbiosis. *Environ Microbiol.* 21(10):3765–3779. <https://doi.org/10.1111/1462-2920.14727>

Perez-Riverol Y, Csordas A, Bai J, Bernal-Llinares M, Hewapathirana S, Kundu DJ, Inuganti A, Griss J, Mayer G, Eisenacher M, et al. 2019. The PRIDE database and related tools and resources in 2019: Improving support for quantification data. *Nucleic Acids Res.* 47(D1):D442–D450. <https://doi.org/10.1093/nar/gky1106>

Petersen TN, Brunak S, Von Heijne G, Nielsen H. 2011. SignalP 4.0: Discriminating signal peptides from transmembrane regions. *Nat Methods.* 8(10):785–786. <https://doi.org/10.1038/nmeth.1701>

Quijano MR, Zach C, Miller FS, Lee AR, Imani AS, Künzler M, Freeman MF. 2019. Distinct Autocatalytic  $\alpha$ -N-Methylating Precursors Expand the Borosin RiPP Family of Peptide Natural Products. *J Am Chem Soc.* 141(24):9637–9644. <https://doi.org/10.1021/jacs.9b03690>

Ramm S, Krawczyk B, Mühlenweg A, Poch A, Mösker E, Süßmuth RD. 2017. A Self-Sacrificing N-Methyltransferase Is the Precursor of the Fungal Natural Product Omphalotin. *Angew Chemie - Int Ed.* 56(33):9994–9997. <https://doi.org/10.1002/anie.201703488>

Raudaskoski M, Kothe E. 2010. Basidiomycete mating type genes and pheromone signaling. *Eukaryot Cell.* 9(6):847–859. <https://doi.org/10.1128/EC.00319-09>

Sakamoto Y, Nakade K, Konno N. 2011. Endo- $\beta$ -1,3-Glucanase GLU1, from the fruiting body of *Lentinula edodes*, belongs to a new glycoside hydrolase family. *Appl Environ Microbiol.* 77(23):8350–8354. <https://doi.org/10.1128/AEM.05581-11>

Van Schöll L, Hoffland E, Van Breemen N. 2006. Organic anion exudation by ectomycorrhizal fungi and *Pinus sylvestris* in response to nutrient deficiencies. *New Phytol.* 170(1):153–163. <https://doi.org/10.1111/j.1469-8137.2006.01649.x>

Stöckli M, Lin C wei, Sieber R, Plaza DF, Ohm RA, Künzler M. 2017. *Coprinopsis cinerea* intracellular lactonases hydrolyze quorum sensing molecules of Gram-negative bacteria. *Fungal Genet Biol.* 102:49–62. <https://doi.org/10.1016/j.fgb.2016.07.009>

- Stöckli M, Morinaka BI, Lackner G, Kombrink A, Sieber R, Margot C, Stanley CE, DeMello AJ, Piel J, Künzler M. 2019. Bacteria-induced production of the antibacterial sesquiterpene lagopodin B in *Coprinopsis cinerea*. *Mol Microbiol.* 112(2):605–619. <https://doi.org/10.1111/mmi.14277>
- Süssmuth RD, Mainz A. 2017. Nonribosomal Peptide Synthesis—Principles and Prospects. *Angew Chemie - Int Ed.* 56(14):3770–3821. <https://doi.org/10.1002/anie.201609079>
- Tayyrov A, Stanley CE, Azevedo S, Künzler M. 2019. Combining microfluidics and RNA-sequencing to assess the inducible defensome of a mushroom against nematodes. *BMC Genomics [Internet]*. 20(1):243. <https://doi.org/10.1186/s12864-019-5607-3>
- Teertstra WR, van der Velden GJ, De Jong JF, Kruijtzter JAW, Liskamp RMJ, Kroon-Batenburg LMJ, Müller WH, Gebbink MFBG, Wösten HAB. 2009. The filament-specific Rep1-1 repellent of the phytopathogen *Ustilago maydis* forms functional surface-active amyloid-like fibrils. *J Biol Chem.* 284(14):9153–9159. <https://doi.org/10.1074/jbc.M900095200>
- Tsukui T, Nagano N, Terai G, Kumagai T, Machida M, Umemura M, Asai K. 2014. Ustiloxins, fungal cyclic peptides, are ribosomally synthesized in *Ustilaginoidea virens*. *Bioinformatics.* 31(7):981–985. <https://doi.org/10.1093/bioinformatics/btu753>
- Tsukui T, Nagano N, Umemura M, Kumagai T, Terai G, Machida M, Asai K. 2015. Ustiloxins, fungal cyclic peptides, are ribosomally synthesized in *Ustilaginoidea virens*. *Bioinformatics.* 31(7):981–985. <https://doi.org/10.1093/bioinformatics/btu753>
- Umemura M. 2020. Peptides derived from Kex2-processed repeat proteins are widely distributed and highly diverse in the Fungi kingdom. *Fungal Biol Biotechnol [Internet]*. 7(1):1–24. <https://doi.org/10.1186/s40694-020-00100-5>
- Umemura M, Nagano N, Koike H, Kawano J, Ishii T, Miyamura Y, Kikuchi M, Tamano K, Yu J, Shin-ya K, Machida M. 2014. Characterization of the biosynthetic gene cluster for the ribosomally synthesized cyclic peptide ustiloxin B in *Aspergillus flavus*. *Fungal Genet Biol [Internet]*. 68:23–30. <https://doi.org/10.1016/j.fgb.2014.04.011>
- Van Der Velden NS, Kälin N, Helf MJ, Piel J, Freeman MF, Künzler M. 2017. Autocatalytic backbone N-methylation in a family of ribosomal peptide natural products. *Nat Chem Biol.* 13(8):833–835. <https://doi.org/10.1038/nchembio.2393>
- Vogt E, Künzler M. 2019. Discovery of novel fungal RiPP biosynthetic pathways and their application for the development of peptide therapeutics. *Appl Microbiol Biotechnol.* 103(14):5567–5581. <https://doi.org/10.1007/s00253-019-09893-x>
- Wälti MA, Villalba C, Buser RM, Grünler A, Aebi M, Künzler M. 2006. Targeted gene silencing in the model mushroom *Coprinopsis cinerea* (*Coprinus cinereus*) by expression of homologous hairpin RNAs. *Eukaryot Cell.* 5(4):732–744. <https://doi.org/10.1128/EC.5.4.732-744.2006>
- Wösten HAB, Bohlmann R, Eckerskorn C, Lottspeich F, Bölker M, Kahmann R, Bolker M, Kahmann R, Bölker M, Kahmann R. 1996. A novel class of small amphipathic peptides affect aerial hyphal growth and surface hydrophobicity in *Ustilago maydis*. *EMBO J [Internet]*. 15(16):4274–4281. <https://doi.org/10.1002/j.1460-2075.1996.tb00802.x>
- Yang S, Kuang Y, Li H, Liu Y, Hui X, Li P, Jiang Z, Zhou Y, Wang Y, Xu A, et al. 2013. Enhanced Production of Recombinant Secretory Proteins in *Pichia pastoris* by Optimizing Kex2 P1' site. *PLoS One.* 8(9):1–11. <https://doi.org/10.1371/journal.pone.0075347>

## Chapter 2

Zuris JA, Thompson DB, Shu Y, Guilinger JP, Bessen JL, Hu JH, Maeder ML, Joung JK, Chen ZY, Liu DR. 2015. Cationic lipid-mediated delivery of proteins enables efficient protein-based genome editing in vitro and in vivo. *Nat Biotechnol.* 33(1):73–80. <https://doi.org/10.1038/nbt.3081>

Supplementary information

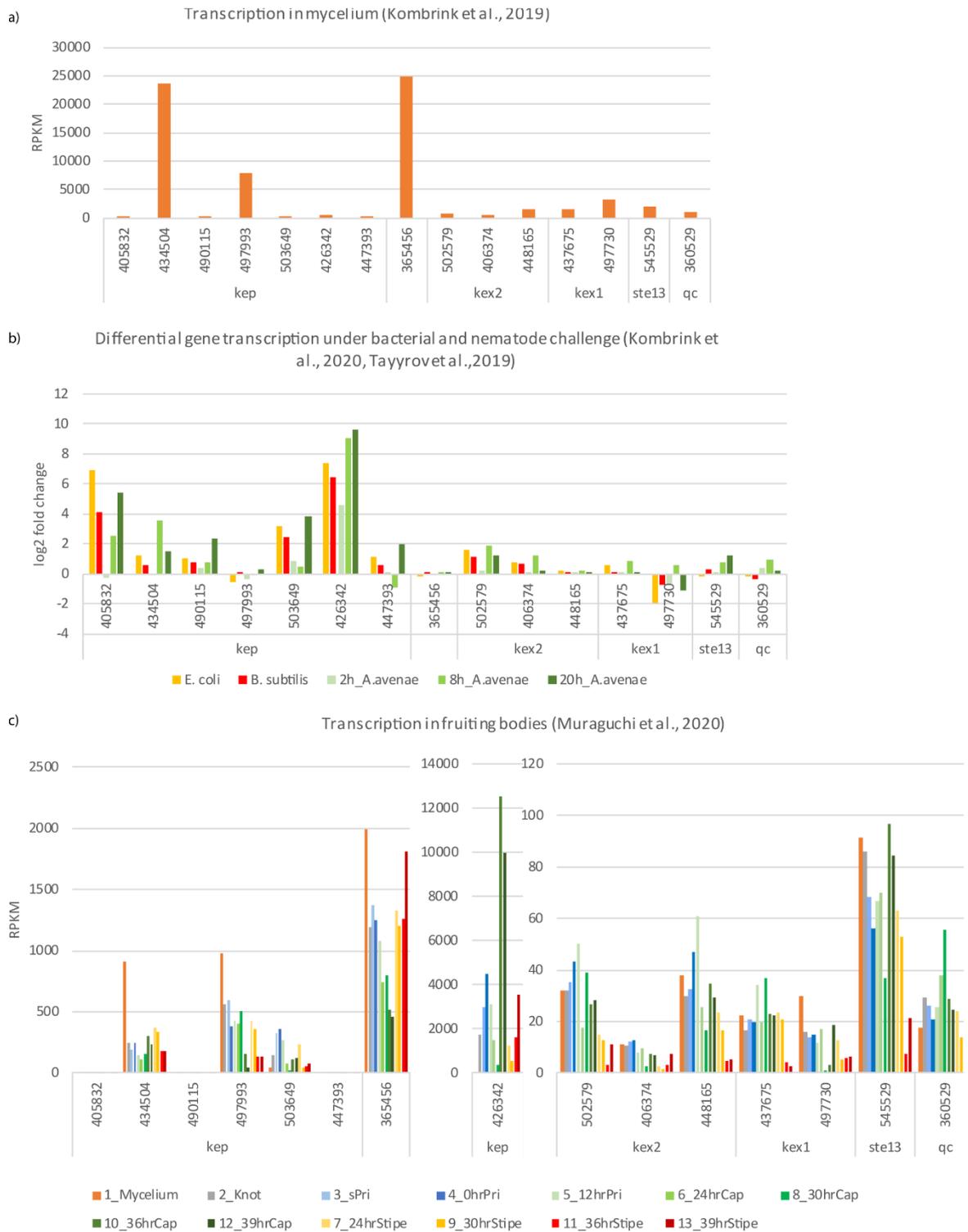
Structural and functional analysis of peptides derived from KEX2-processed repeat proteins in agaricomycetes using reverse genetics and peptidomics

Eva Vogt<sup>1</sup>, Lukas Sonderegger<sup>1</sup>, Ying-Yu Chen<sup>1</sup>, Tina Segesseemann<sup>1</sup>, Markus Künzler<sup>1</sup>

<sup>1</sup> ETH Zürich, Department of Biology, Institute of Microbiology, Vladimir-Prelog-Weg 4, CH-8093 Zürich, Switzerland

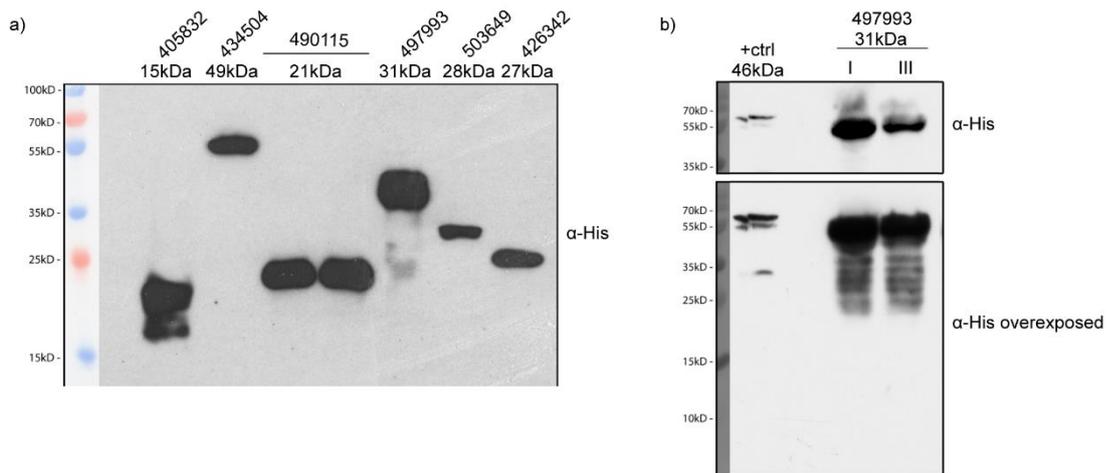
**Supplemental Materials and Methods:** Antimicrobial assays using synthetic peptides. For the disk diffusion assay, the bacterial strains *B. subtilis* 168, *B. subtilis* NCBI 3610, *Micrococcus luteus*, *Staphylococcus aureus*, and *E. coli* BL21 were streaked on agar-solidified LB plates and incubated overnight at 37°C. Freshly grown colonies were resuspended in 1 ml liquid LB medium to an OD<sub>595</sub> of 0.1 and 150 µl of culture suspensions were spread evenly on agar-solidified LB plates. Peptides were not tested individually. Instead, three or four peptides were combined in peptide mixtures and tested together. Filter paper disks were loaded with 70 nmol of each individual peptide or 5 µg of vancomycin hydrochloride (from *Streptomyces orientalis*, Merck, Germany), or ampicillin sodium salt (BioChemica, AppliChem GmbH, Germany) as positive controls. Bacterial growth was checked after incubation at 37°C for 24 h. For bacterial assays using an optical density (OD) readout, resuspended bacteria were allowed to grow to an OD<sub>595</sub> of 0.5 in liquid LB medium at 37°C while shaking, were then diluted to an OD<sub>595</sub> of 0.1, and incubated in triplicates in 96-well microtiter plates at a concentration of 0.5 mM for each peptide mixture and 50 µg/ml of vancomycin or ampicillin. The OD<sub>595</sub> was measured after 24 hours and the change of optical density was calculated for each microtiter well.

Supplemental Figures



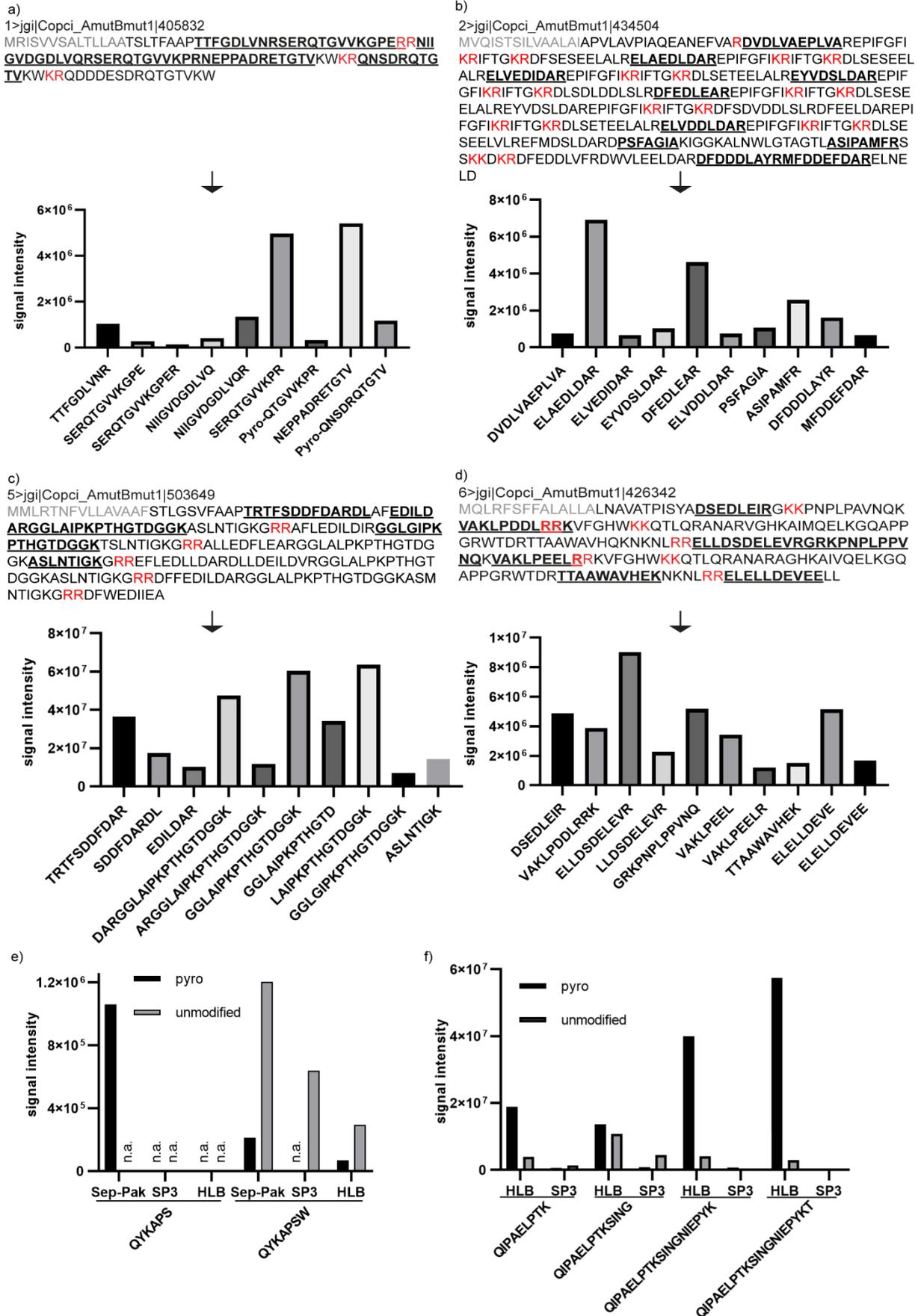
**Figure S1:** Gene transcription profiles of the *C. cinerea* *kep* genes, the gene 365456, *kex* genes, *ste13*, and a glutaminyl cyclase gene. a) Gene transcription in the vegetative mycelium. The y-axis describes reads per kilobase of transcript per million reads mapped (RPKM). b) Differential gene transcription of *C. cinerea* challenged with the bacteria *E. coli* or *B. subtilis* or the nematode *A. avenae*. c) Gene transcription in fruiting body development. For visual clarity, *kep426342* and *kex2*, *kex1*, *ste13*, and glutaminyl cyclase (*qc*) genes are shown on different scales than the other genes. Samples harvested at 39 h are premature fruiting bodies. Raw data from Muraguchi et al. 2015; Kombrink et al. 2019; Tayyrov et al. 2019.

## Chapter 2



**Figure S2:** Verification of expression of *C. cinerea* KEPs in *P. pastoris*. a) Immunoblot of whole cell extracts of *P. pastoris* expressing *C. cinerea* KEPs 405832, 434504, 490115, 497993, 503649, and 426342. Samples were run on a 12% SDS-PAGE, blotted on a nitrocellulose membrane and hybridized with anti-His antibodies. Detection of bands of the correct size confirmed the successful expression of the six proteins. b) Immunoblot of whole cell extracts of *P. pastoris* expressing *C. cinerea* KEP497993. Solubility of the protein was tested by preparing samples before (pellet I) and after (pellet III) high speed centrifugation at 16000 rcf for 30 min at 4°C.

## Chapter 2



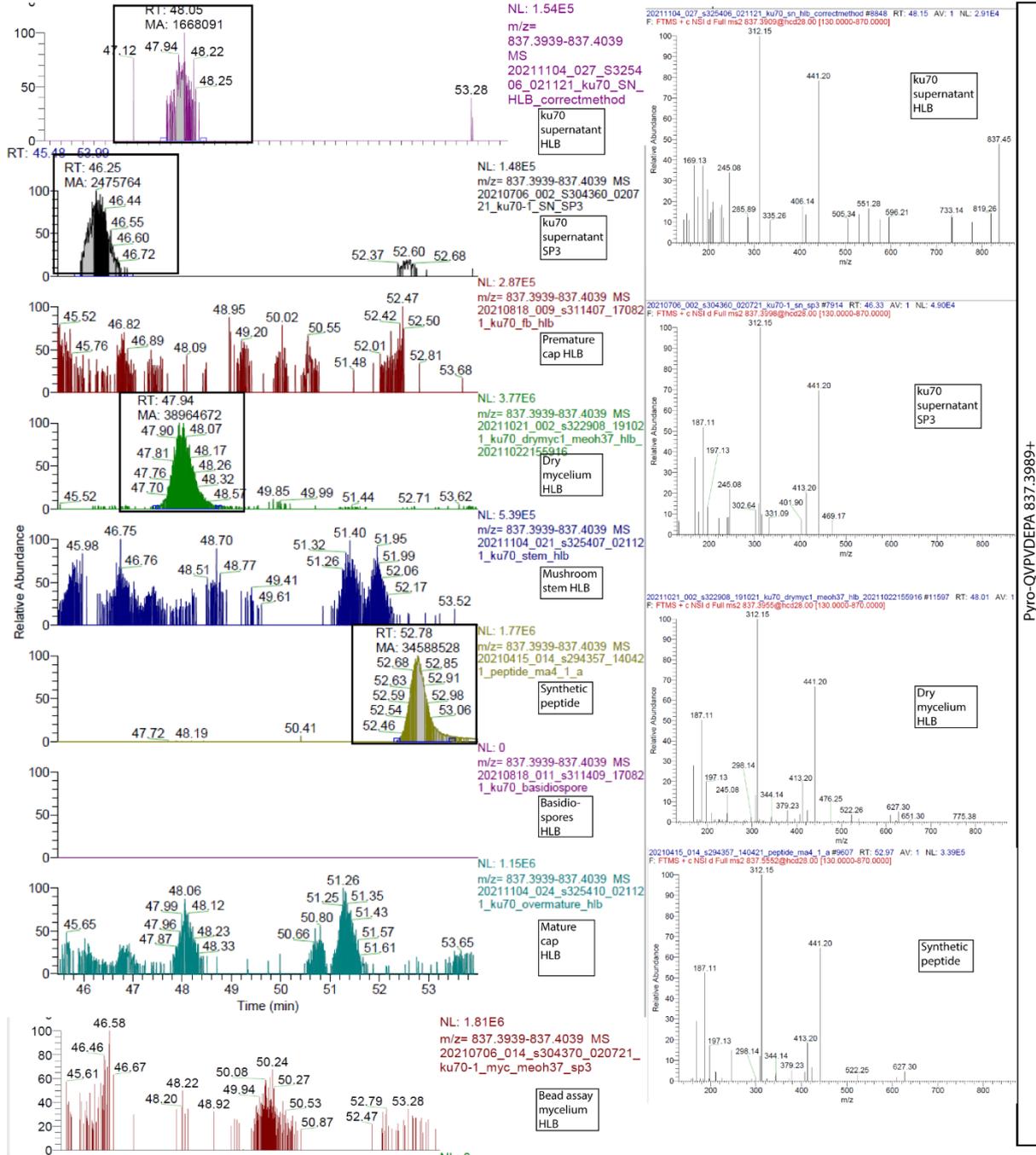
**Figure S3:** KEP-derived peptides can be detected in the culture supernatant of *Pichia pastoris* upon expression of *C. cinerea* KEPs. a-d) Peptides derived from the expression of *C. cinerea* KEPs 405832,

## Chapter 2

434504, 503649, and 426342 in *P. pastoris*. Given are the protein sequences with KEX2 cleavage sites labeled red and signal sequences labeled grey. The 10 peptides that were measured with the highest mass spectrometry signal intensities in the supernatant are indicated in the graphs. Their sequences are indicated in bold and underlined in the protein sequence. e-f) Comparison of pyroglutamated and unmodified KEP-derived peptides across different extraction methods in *P. pastoris* heterologous expressions. Signal intensities of a selection of peptides that were extracted using the solid phase extraction methods Sep-Pak, HLB, and SP3 from the *P. pastoris* expression of *C. cinerea* KEP 490115 (left) and KEP 497993 (right) are compared.

# Chapter 2

S4a)

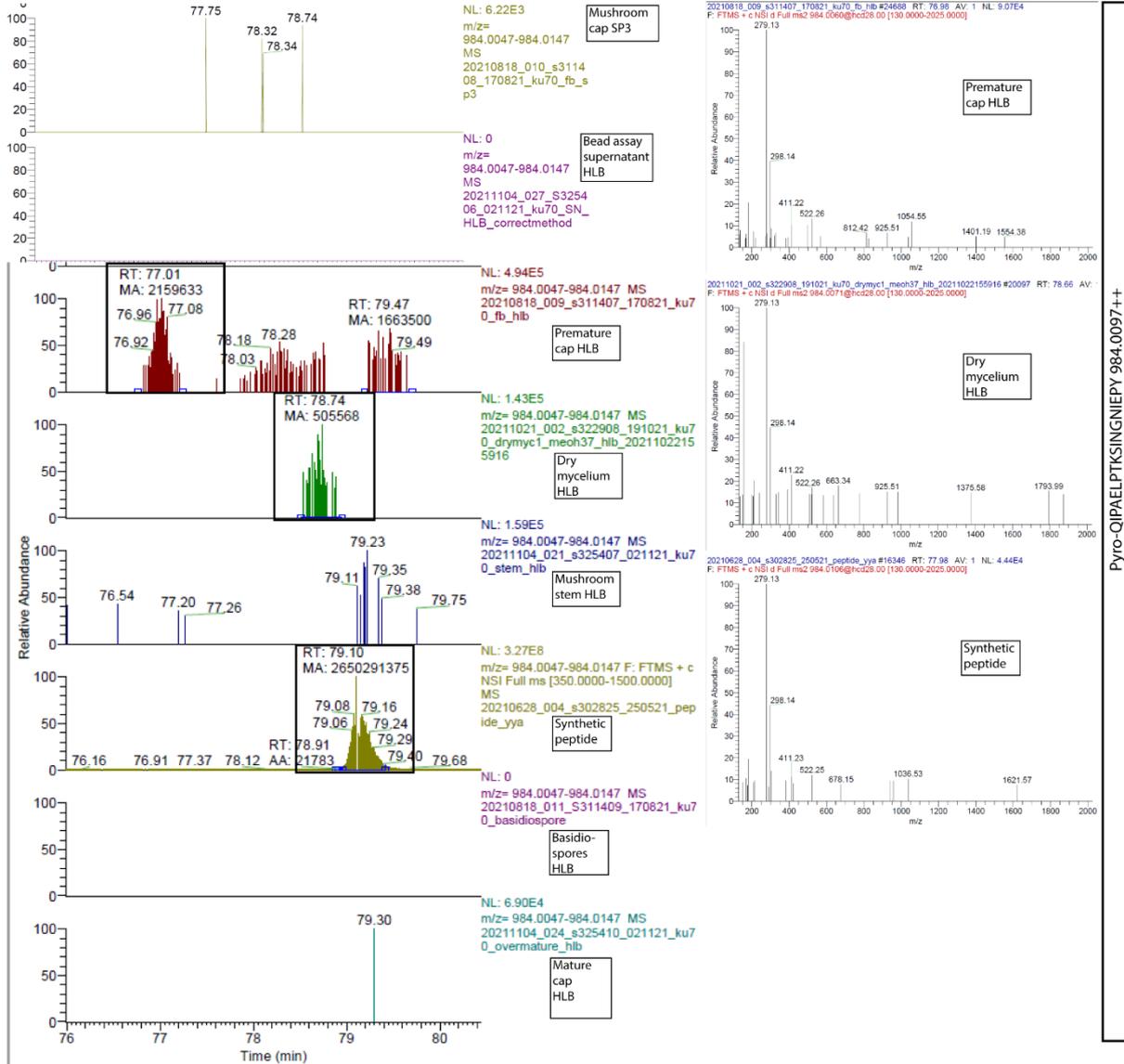


Pyro-QVPVDEPA 837.3989+

Pyro-QVPVDEPA	idotp	dotp
Supernatant	0.99	0.79
Dry mycelium	0.92	0.83

# Chapter 2

S4b)

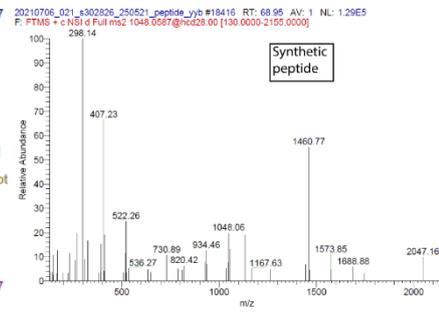
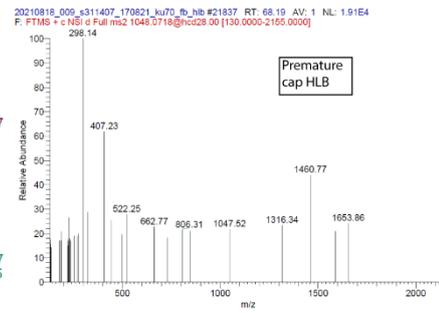
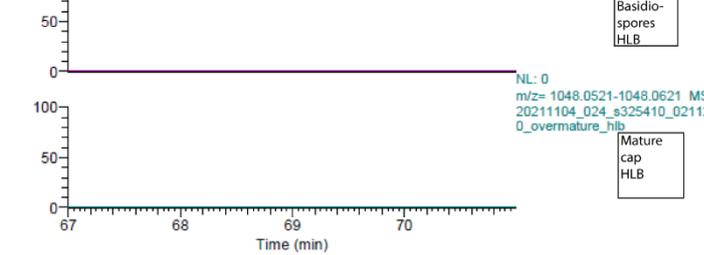
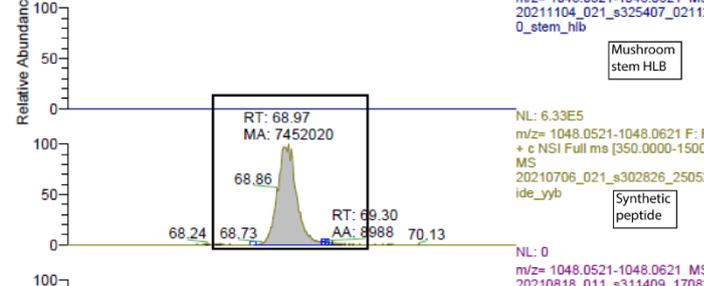
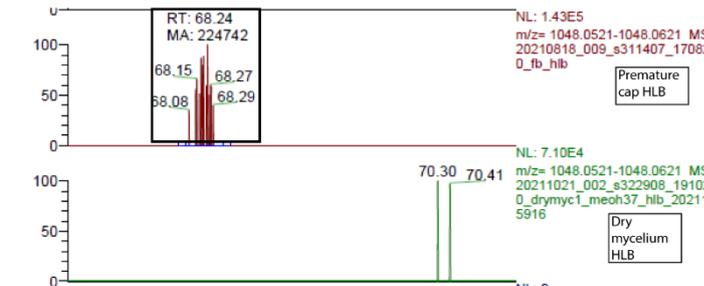


Pyro-QIPALPTKSINGNIEPY 984.0097++

Pyro-QIPALPTKSINGNIEPY	idotp	dotp
Dry mycelium	0.9	0.82
Mushroom cap	0.94	0.91

# Chapter 2

S4c)

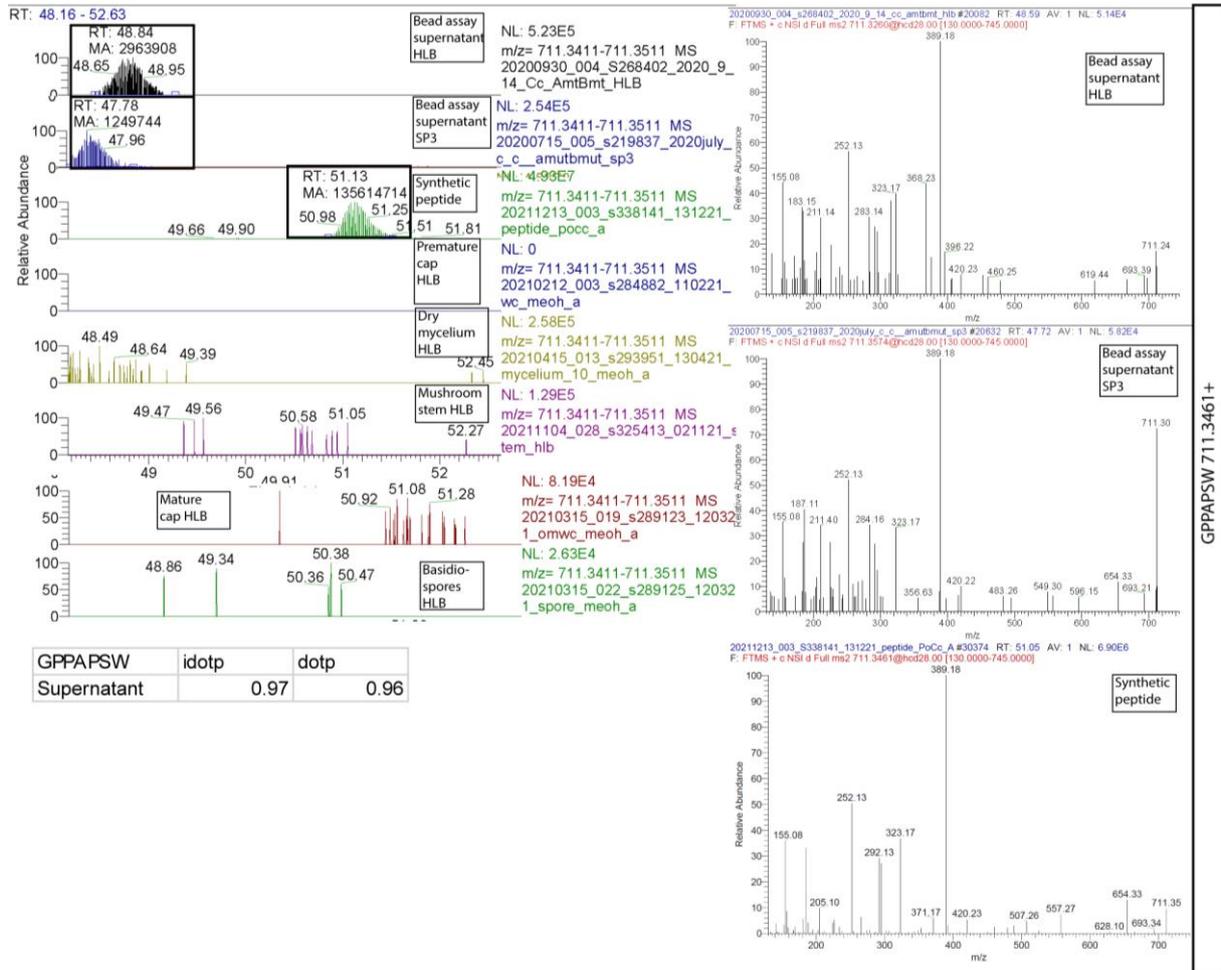


Pyro-QIPAEIPTSINGNIEPYK 1048.0571++

Pyro-QIPAEIPTSINGNIEPYK	idotp	dotp
Mushroom cap	0.86	0.91

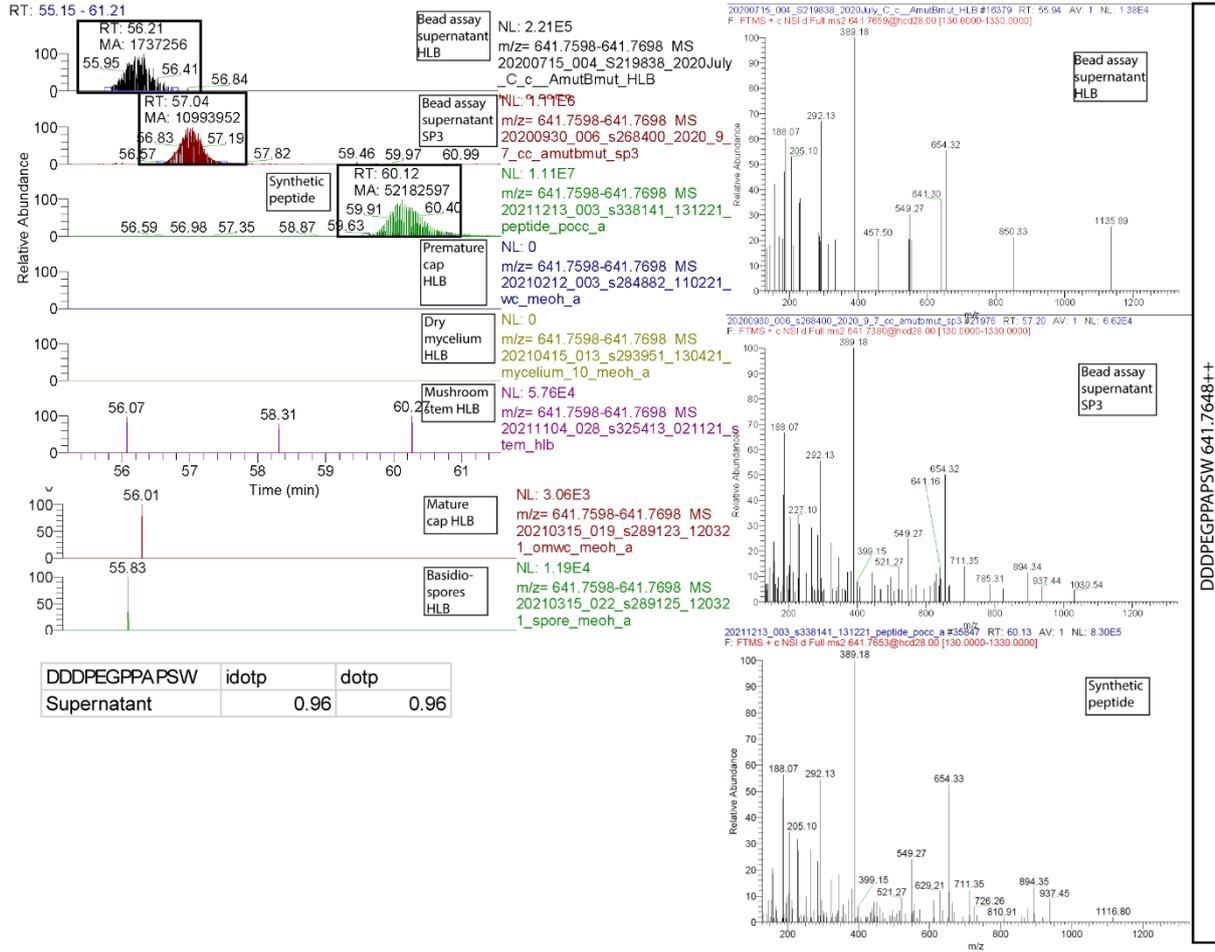
# Chapter 2

S4d)



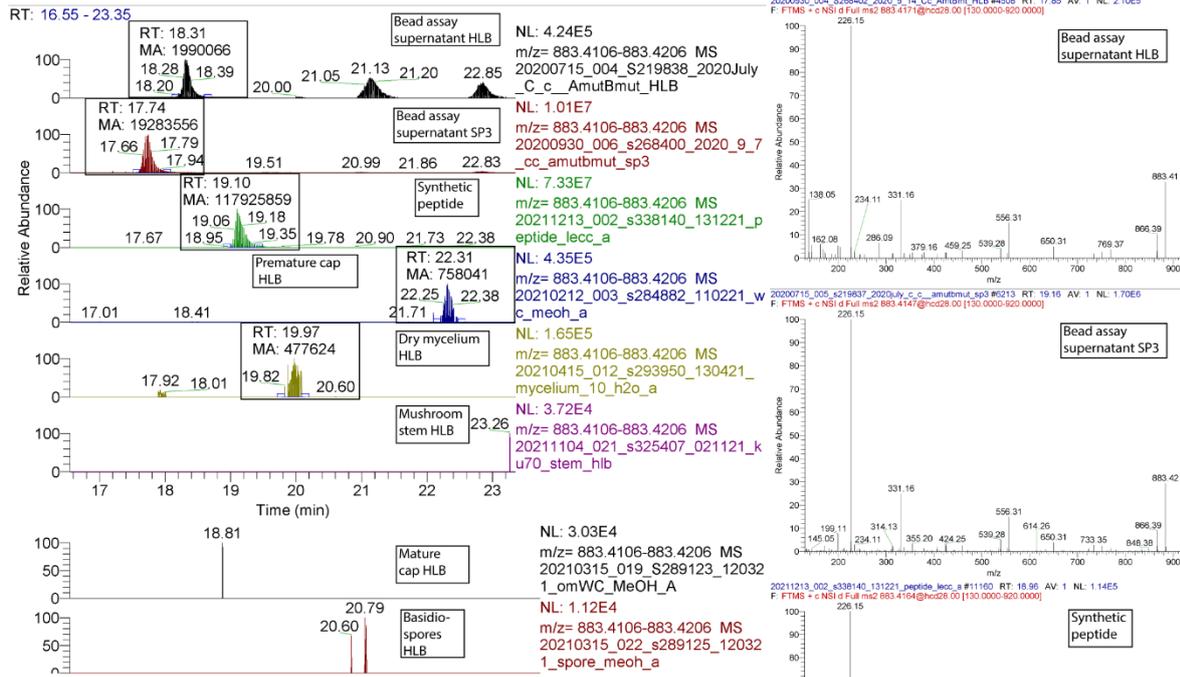
# Chapter 2

S4e)

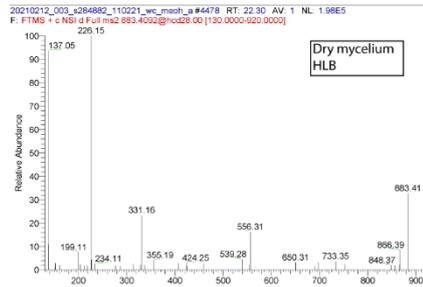
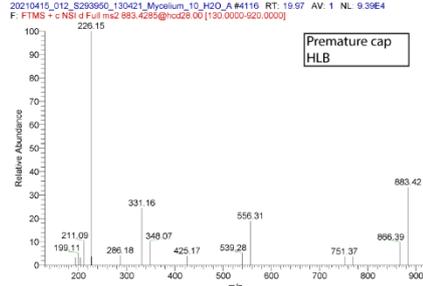
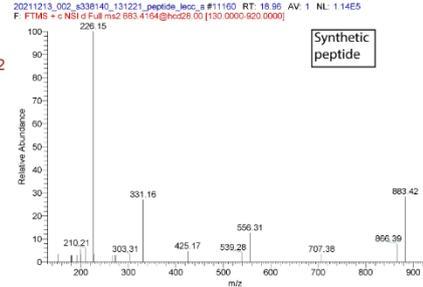
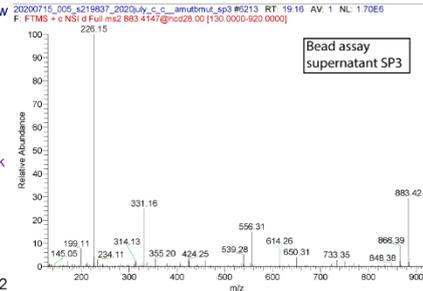
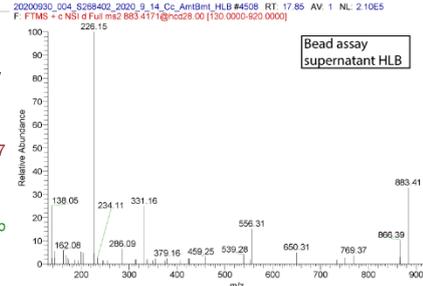


# Chapter 2

S4f)



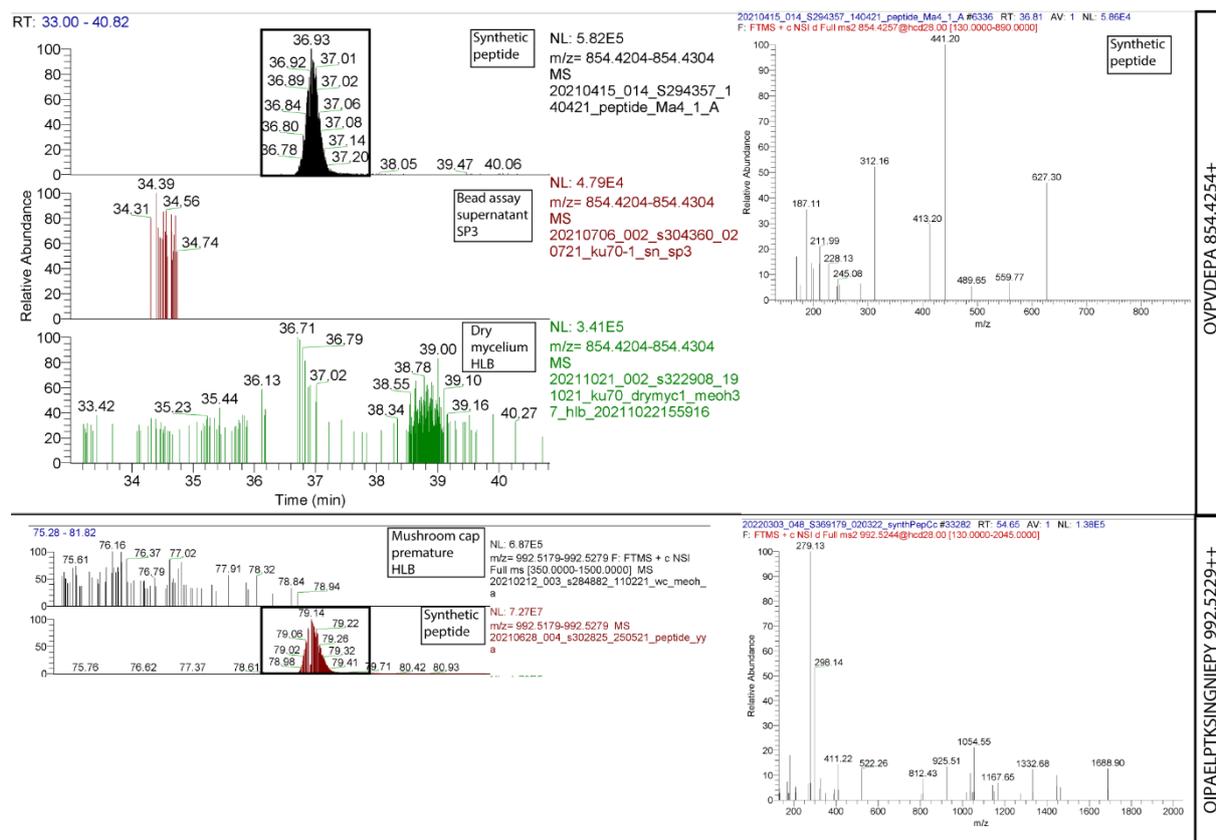
Pyro-QSEPKPTN	idotp	dotp
Supernatant	0.96	0.94
Dry mycelium	0.88	0.84
Mushroom cap	0.86	0.78



Pyro-QSEPKPTN 883.4156+

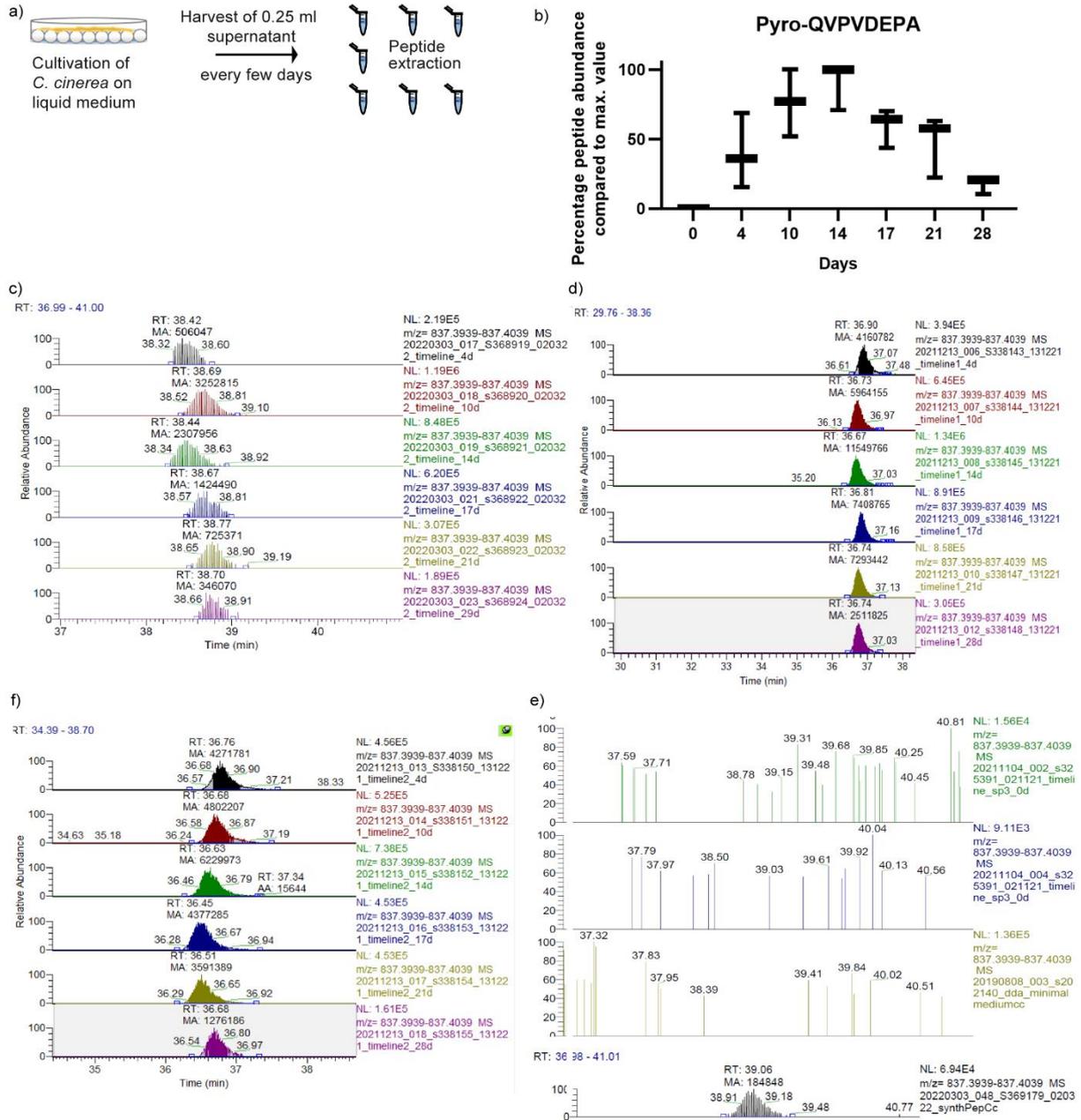
## Chapter 2

S4g)



**Figure S4:** Extracted ion chromatograms and MS/MS spectra of peptides in *C. cinerea* tissues. Panels a) to f) depict the extracted ion chromatograms (EIC, left) and the MS/MS spectra (right) with the retention times (RT) and the peak areas (MA) of endogenous peptides from *C. cinerea* samples or synthetic peptides that are used as gold standards. The isotope dot product (idot) and dot product (dotp) values, determined using the software Skyline 20.2.0.343 (MacCoss Lab Software, USA), are indicated as an additional means to compare similarity between the spectra. Samples were extracted using the solid phase extraction cartridge HLB or SP3 magnetic beads. Panels a) to f) each show the presence of one peptide in a selection of *C. cinerea* tissue samples. The peptides are a) pyro-QVPVDEPA, b) pyro-QIPAEIPTSINGNIEPY, c) pyro-QIPAEIPTSINGNIEPYK, d) GPPAPSW, e) DDDPEGPPAPSW, f) pyro-QSEPKPTN, and g) unmodified QVPVDEPA and QIPAEIPTSINGNIEPY.

## Chapter 2

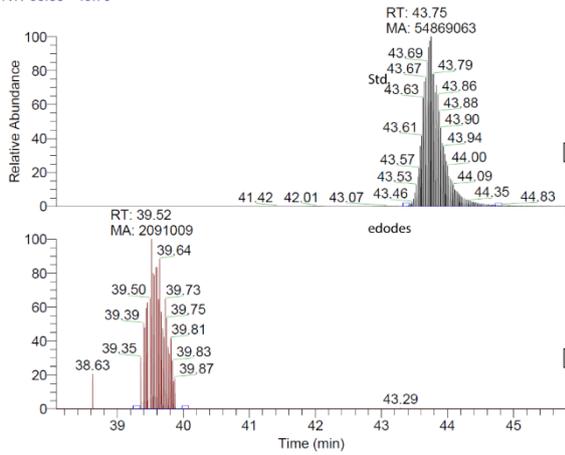


**Figure S5:** Presence of the KEP-derived peptide pyro-QVPVDEPA in time-course sampling of *C. cinerea* culture supernatant. a) Workflow of the experiment. A triplicate of *C. cinerea*  $\Delta ku70$  bead assays were prepared where the fungus grows on glass beads covered with liquid minimal medium. The supernatant was sampled every few days. b) Time-course of peptide abundance over 28 days. The peptide peak areas of the extracted ion chromatograms of each time point were determined, and the maximal value of each of three data sets was set as 100%. The rest of the values were then compared to this maximum. The single bands represent the median value of three biological replicates, the whiskers represent the 10-90% percentile range. The subpanels c) to f) show the extracted ion chromatograms of three independent data sets, e) shows the extracted ion chromatograms of the 0 d minimal medium samples and the synthetic peptide.

# Chapter 2

S6a)

RT: 38.08 - 45.79



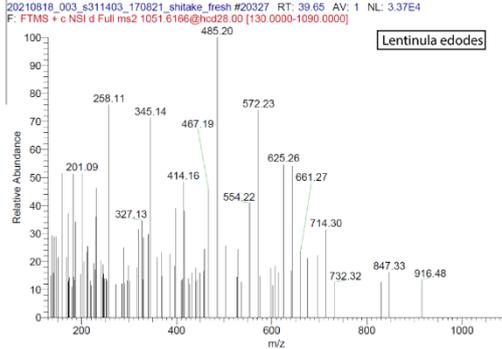
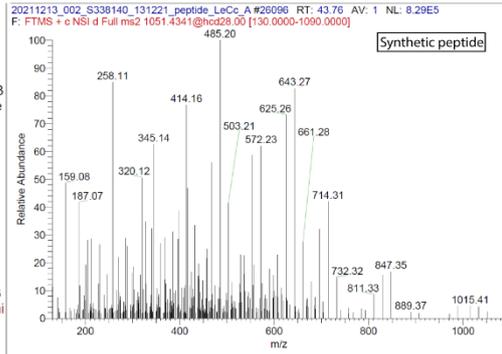
SGTGEASAADW	idotp	dotp
<i>L. edodes</i>	0.96	0.92

NL: 1.88E7  
m/z= 1051.4277-1051.4377 MS  
20211213\_002\_S338140\_131221\_peptide\_LeCc\_A #26096  
20211213\_002\_S338140\_131221\_peptide\_LeCc\_A

Synthetic peptide

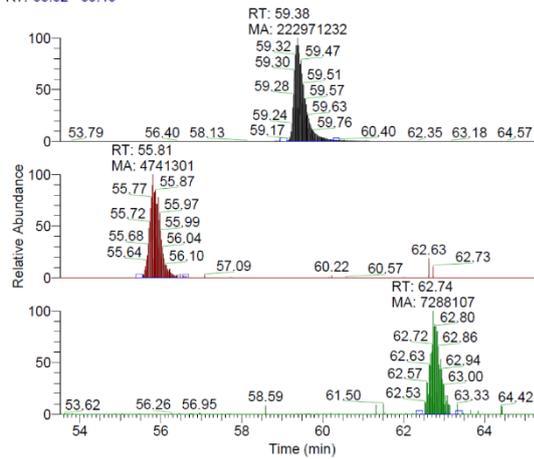
NL: 8.40E5  
m/z= 1051.4277-1051.4377 MS  
20210818\_003\_s311403\_170821\_shitake\_fresh

Lentinula edodes



SGTGEASAADW 1051.4327+

RT: 53.52 - 65.46



MSGVAADW	idotp	dotp
<i>P. ostreatus</i>	0.97	0.97
<i>P. eryngii</i>	0.91	0.97

NL: 8.24E7

m/z= 836.3557-836.3657 MS  
20211213\_003\_S338141\_131221\_peptide\_PoC c\_A

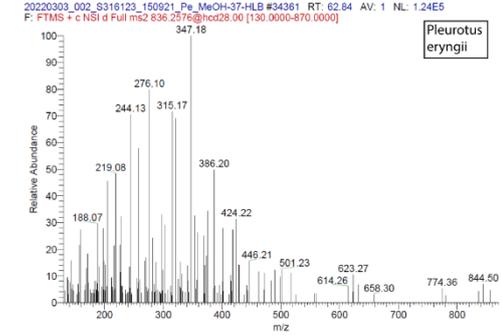
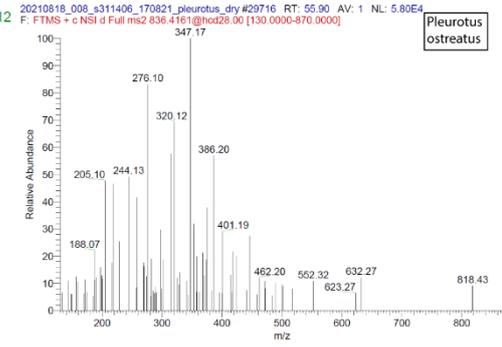
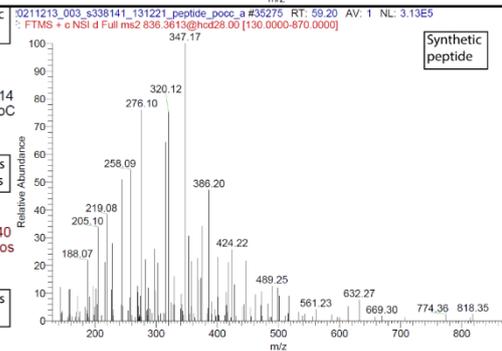
Synthetic peptide

NL: 1.84E6  
m/z= 836.3557-836.3657 MS  
20210818\_008\_s311406\_170821\_pleurotus\_ostreatus\_dry

Pleurotus ostreatus

NL: 2.92E6  
m/z= 836.3557-836.3657 MS  
20220303\_002\_s316123\_150921\_pe\_meh-37-hlb

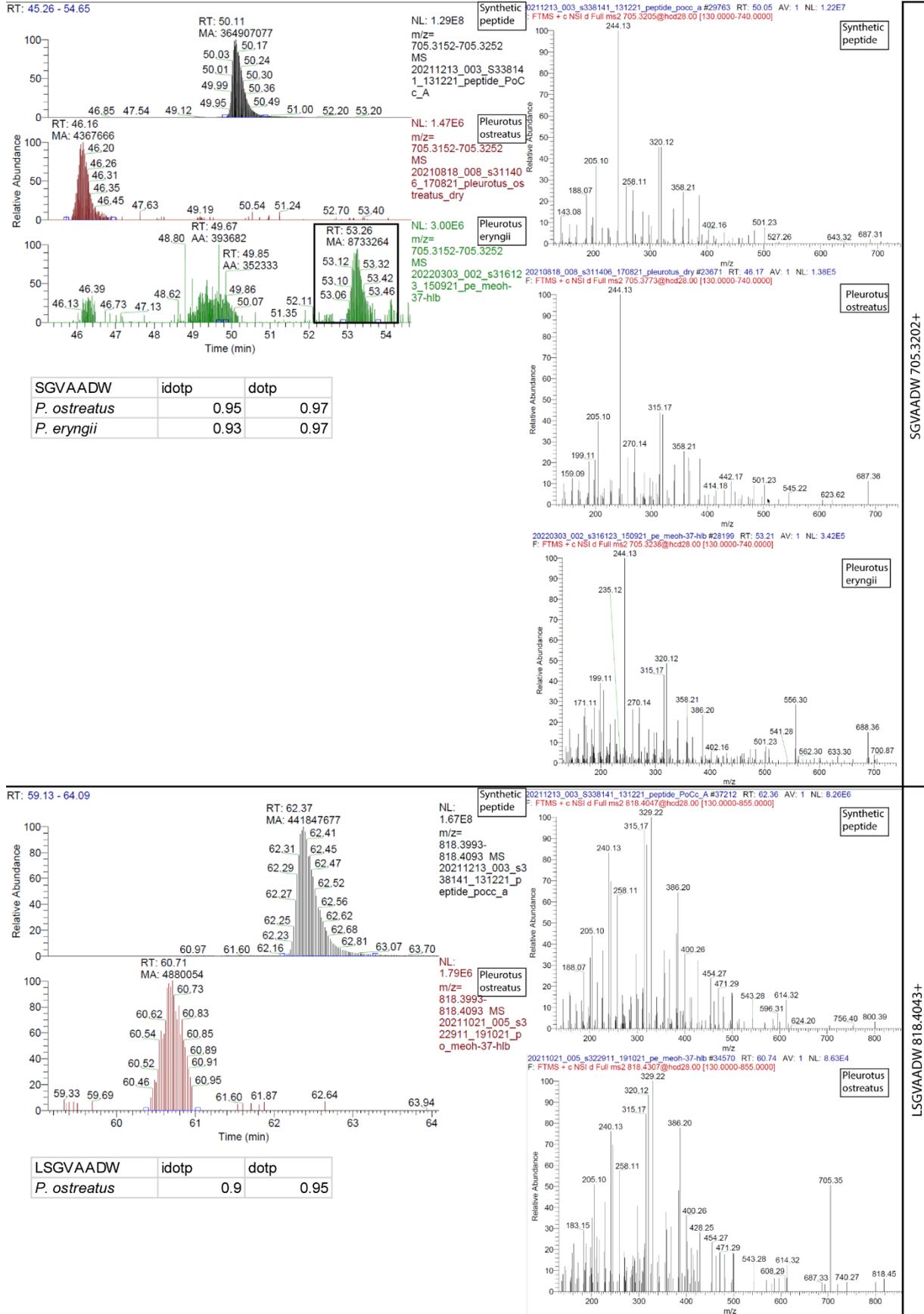
Pleurotus eryngii



MSGVAADW 836.3607+

## Chapter 2

S6b)

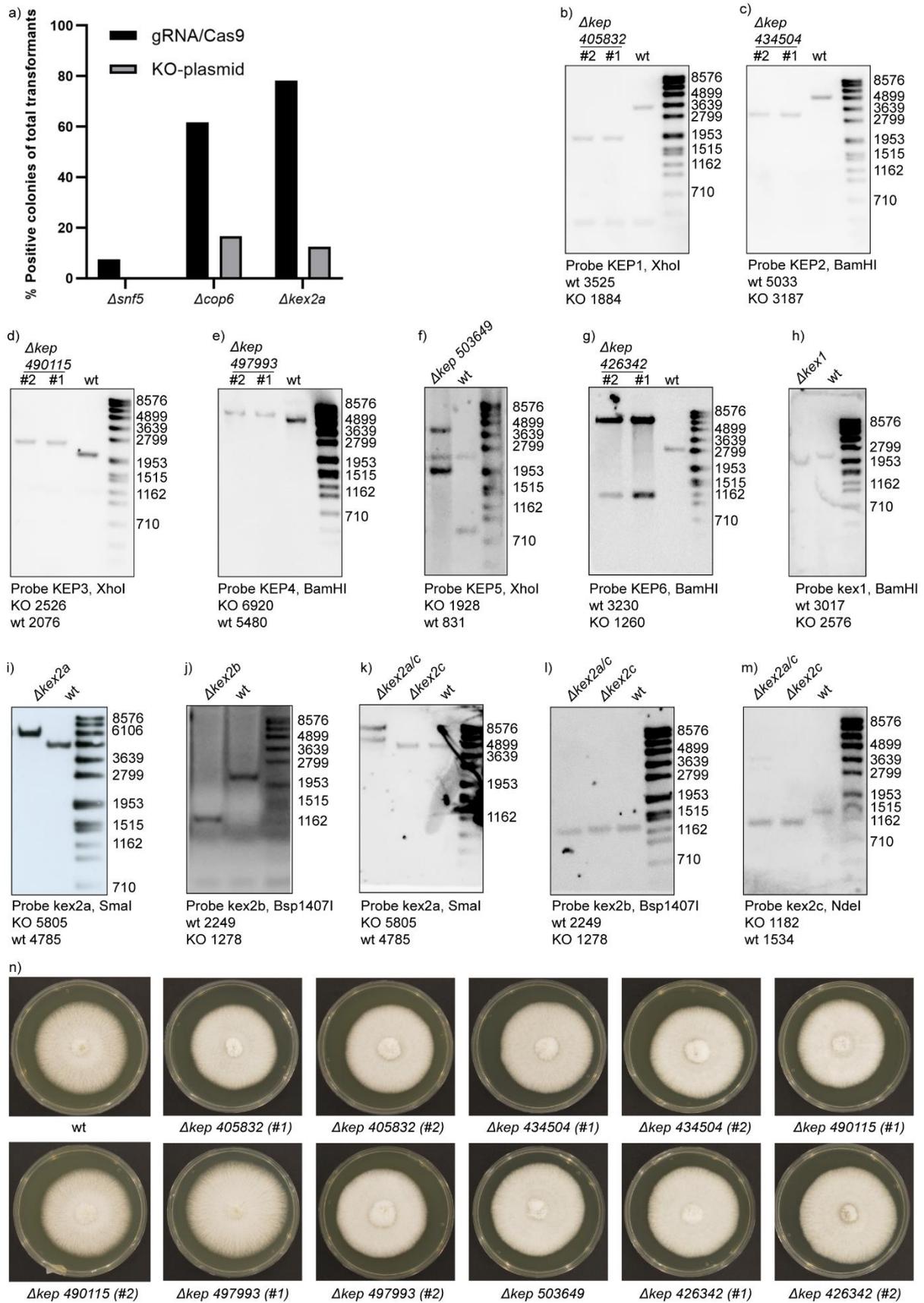


**Figure S6:** Extracted ion chromatograms and MS/MS spectra of peptides in *L. edodes*, *P. ostreatus*, and *P. eryngii* fruiting bodies. Panels a) and b) depict the extracted ion chromatograms (left) and the

## Chapter 2

MS/MS spectra (right) with the retention times and the peak areas of endogenous peptides from *L. edodes*, *P. ostreatus*, and *P. eryngii* fruiting body samples or synthetic peptides that are used as gold standards. The isotope dot product (idot) and dot product (dotp) values, determined using the software Skyline, are given as an additional means to compare similarity between the spectra. The peptides are a) SGTGEASAADW and MSGVAADW, and b) SGVAADW and LSGVAADW.

## Chapter 2



**Figure S7:** Establishment of *kex* and *kep* knockout strains in *C. cinerea*. All knockouts are based on the strain AmutBmut  $\Delta ku70$  that is labeled as wildtype. a) Comparison of the knockout efficiency for three genes. The efficiency of a gRNA/Cas9-guided knockout protocol was assessed in comparison to a traditional non-CRISPR knockout protocol for the genes *snf5* (365798), *cop6* (394772), and *kex2a*

## Chapter 2

(502579). After transformation, picked mycelial colonies were tested in PCR reactions for the presence of a band indicating successful genomic mutation. The numbers of tested colonies were 67 and 56 for  $\Delta snf5$ , 60 and 24 for  $\Delta cop6$ , and 55 and 16 for  $\Delta kex2$ , for gRNA/Cas9 and KO-plasmid alone, respectively. Given is the percentage of colonies whose tests were positive. The genes for *cop6* and *kex2a* were fully replaced with a *pab1* selection marker (Stöckli et al. 2017) using two different gRNAs that cleaved at the 5' and 3' end of the gene, while the *snf5* gene was simply disrupted with the *pab1* gene using a single gRNA according to the knockout strategy by Ando et al. 2013. Subpanels b) to m) depict the confirmation of gene deletions by Southern blot analysis. b) to g) show blots of *kep* knockouts, h) to m) *kex* knockouts. The bands in h) displayed a mass shift due to the use of GelRed. n) Mycelial growth of *kep* knockout strains on solid YMG agar medium. Shown are plates after three days of growth at 37°C.

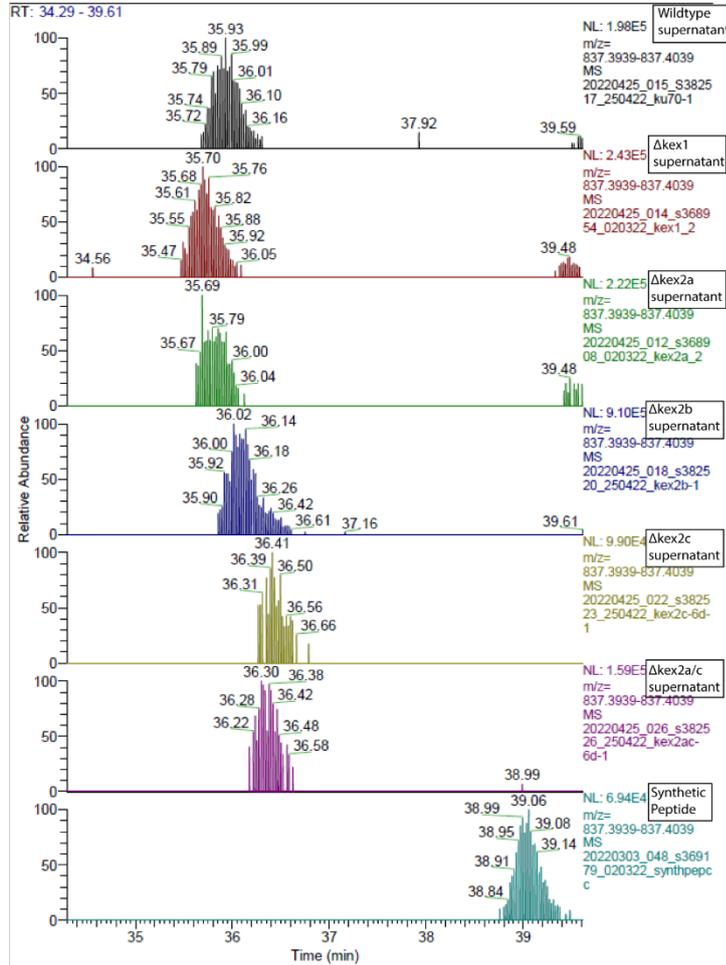


**Figure S8:** Analysis of gene annotation for the gene 439342 from *P. eryngii*. a) Genomic location and RNA coverage of the gene 439342. The gene is located on scaffold\_4:168642-169504 on the negative-sense strand, as shown in the JGI MycoCosm genome viewer (<https://mycoCosm.jgi.doe.gov>). b) Alignment of the nucleotide sequence of the two KEP homologs 1091723 and 439342 from *P. ostreatus* and *P. eryngii*, respectively. Shown is the gene coding for KEP 1091723, including introns, and the

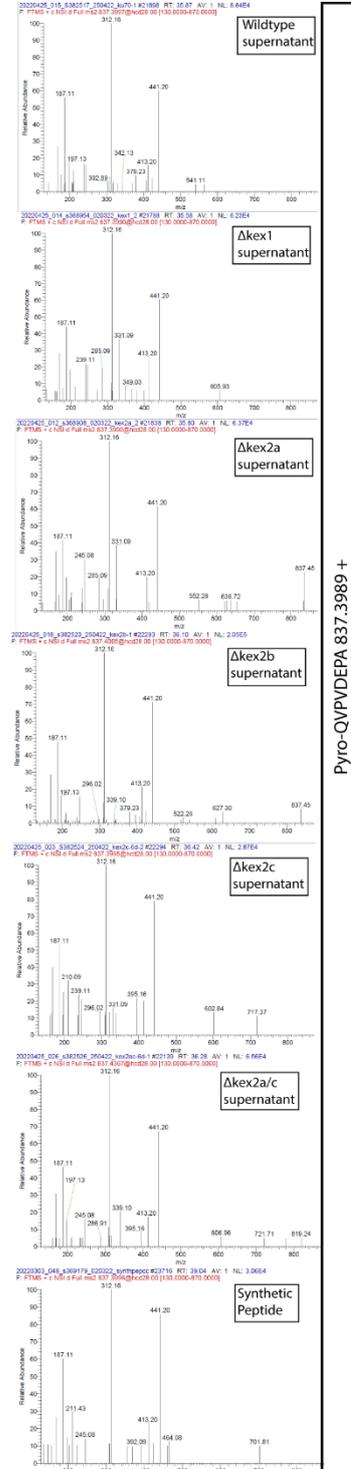
## Chapter 2

genomic locus 168642-169701, scaffold 4, of *P. eryngii*. The annotated gene of *P. eryngii* starts at the Start codon ATG at residue 226.

S9a)

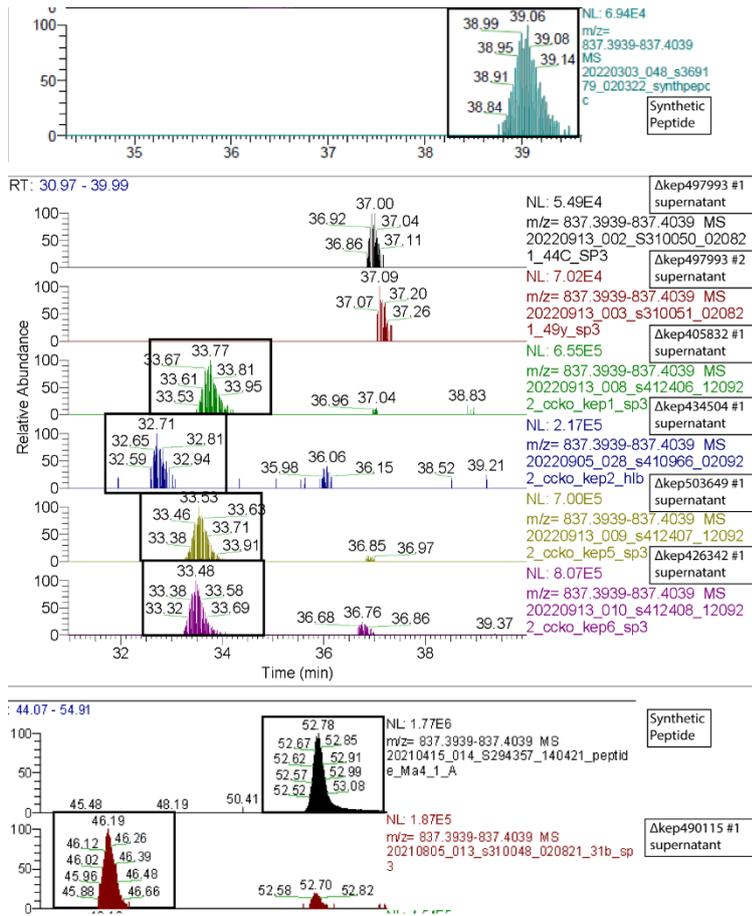


Pyro-QVPV/DEPA	idotp	dotp
Wildtype	0.99	0.79
Δkex1	0.93	0.78
Δkex2a	0.99	0.85
Δkex2b	0.94	0.87
Δkex2c	0.95	0.78
Δkex2a/c	0.82	0.85

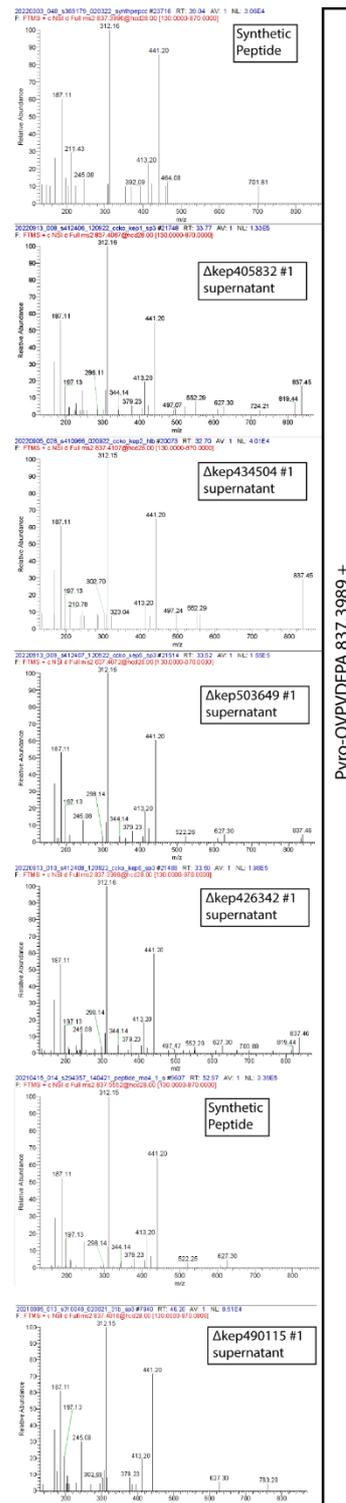


# Chapter 2

S9b)



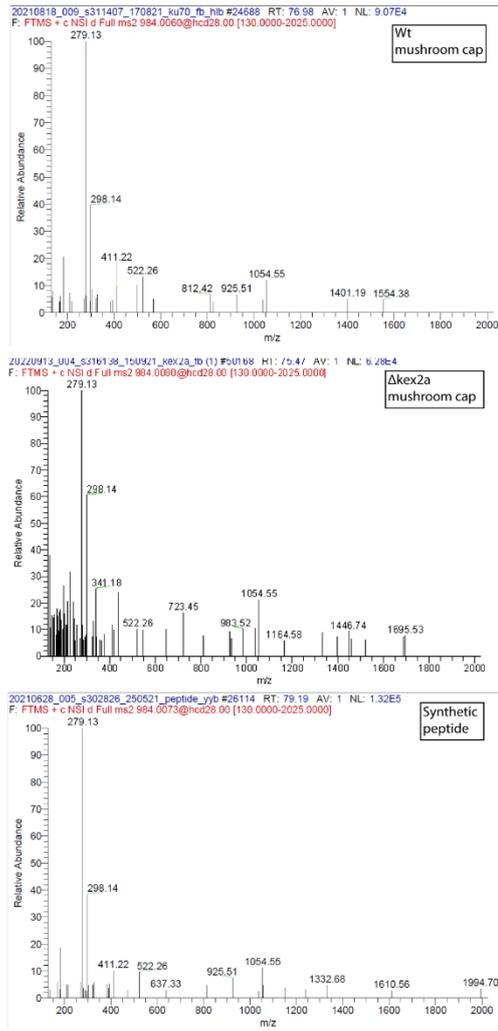
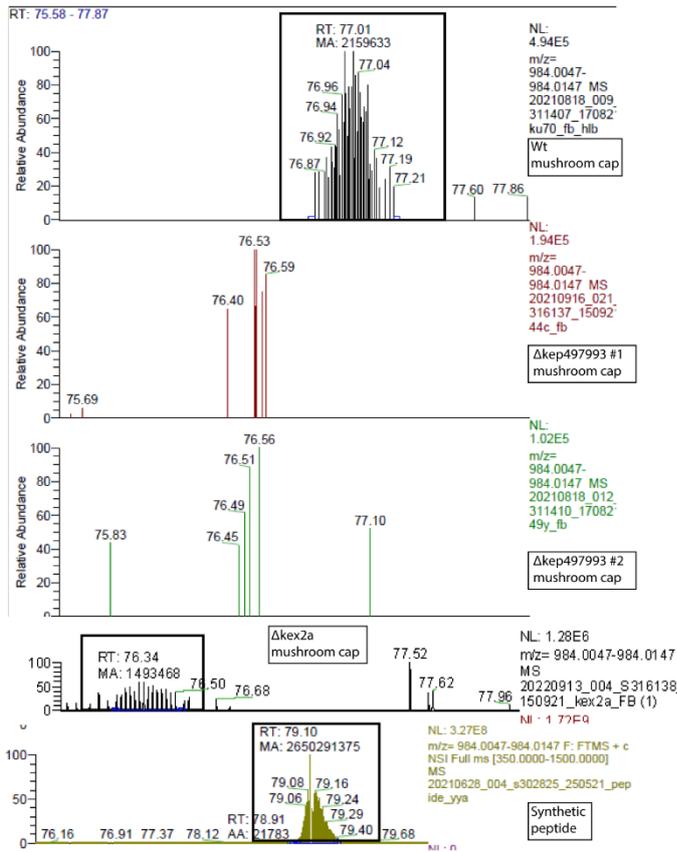
Pyro-QVPVDEPA	idotp	dotp
$\Delta$ kep405832	0.95	0.85
$\Delta$ kep434504	0.8	0.72
$\Delta$ kep490115	0.96	0.83
$\Delta$ kep503649	0.94	0.84
$\Delta$ kep426342	0.97	0.84



Pyro-QVPVDEPA 837.3989 +

# Chapter 2

S9c)

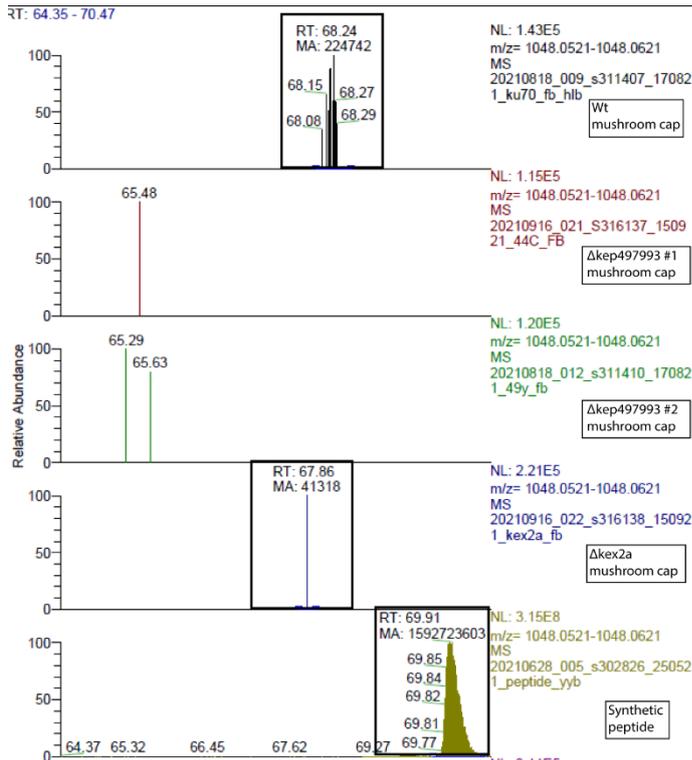


Pyro-QIPAEIPTSINGNIEPY 984.0097++

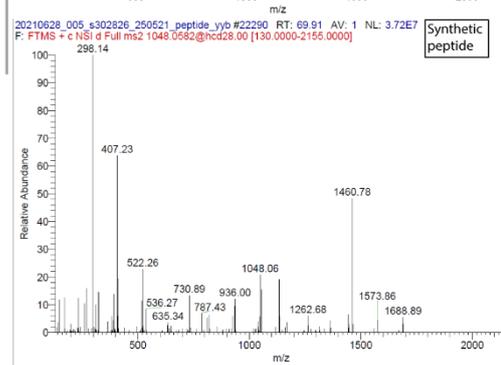
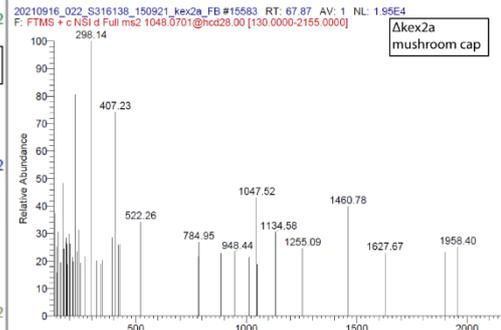
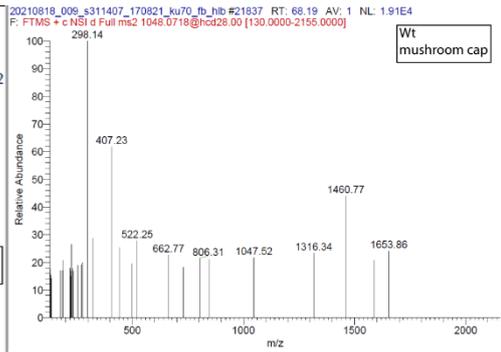
Pyro-QIPAEIPTSINGNIEPY	idotp	dotp
Δkex2a	0.94	0.91

# Chapter 2

S9d)



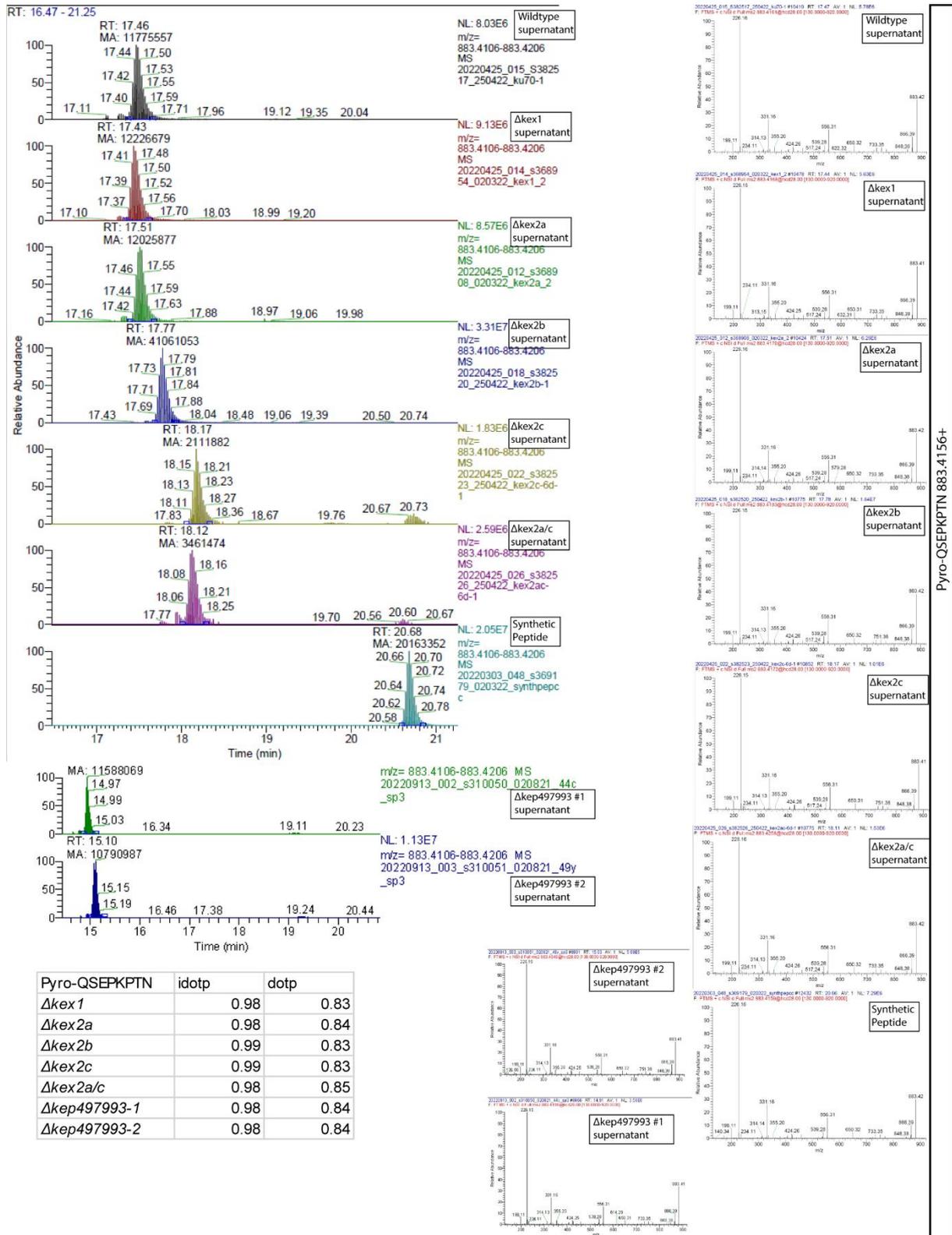
Pyro-QIPAE LPTKSINGNIEPYK	idotp	dotp
Δkex2a	0.87	0.76



Pyro-QIPAE LPTKSINGNIEPYK 1048.0571++

## Chapter 2

S9e)



**Figure S9:** Extracted ion chromatograms and MS/MS spectra of peptides of all constructed *C. cinerea* knockout strains. Panels a) to d) depict the extracted ion chromatograms (EIC, left) and the MS/MS spectra (right) with the retention times and the peak areas of endogenous peptides from *C. cinerea* samples or synthetic peptides that are used as gold standards. The isotope dot product (idot) and dot product (dotp) values, determined using the software Skyline, are given as an additional means to compare similarity between the spectra. The peptides are a) pyro-QVPVDEPA in *kex* knockout strains,

Chapter 2

b) pyro-QVPVDEPA in *kep* knockout strains, c) pyro-QIPAEIPTSINGNIEPY, d) pyro-QIPAEIPTSINGNIEPYK, and e) pyro-QSEPKPTN.

S10a)

	1	10	20	30								
Coprinopsis.cinerea.490115	MR	IASAA	LV	VSVLSV	SLPT	PEVDF	APVAG	.....	VVA			
Pleurotus.ostreatus.KDQ23127.1	MR	VQNI	LV	SSVLSLA	AAV	QPDSSV	AARDGFE	.....	A.RDM.LLS			
Auriculariopsis.ampla.TRM58449.1	MR	LQIF	LA	AALASAA	AGSPV	QPDSDV	DARGYRTS	.....	.DWRKDEEARGY			
Armillaria.gallica.PBK88411.1	MR	FHNL	VV	ASAMAFV	GAFA	..GELSE	ELREAEFANGGGGQGGQ	WR	DAEIIHVA			
Amanita.thiersii.PFH51582.1	MR	SIPFA	LA	LSFLSLA	TAAP	..LNDD	LALRS	.....	.FEA			
Gloeophyllum.trabeum.XP_00787023	MR	VQALL	AV	ASALAVARA	APV	..ETRD	LVARSPVVE	.....	.A.RAD.VVA			
Heliocybe.sulcata.TFK50419.1	MR	FGSLL	VV	ASAVAVIA	APV	..ESGE	LVRSPVVE	.....	.A.RAD.VVA			
Coprinopsis.marcescibilis.TFK229	MR	FSTT	VML	LTSLLAVS	ALPI	..SGEG	VVARDGA	.....	.LVE			
Exidia.glandulosa.KZV96953.1	MR	LTSFV	LLS	IFS	AVIVIA	APV	..AAP	IDP	PAPN.....W.K.....			
Psilocybe.cyanescens.PPQ90175.1	MR	VSV	..	LSAFLV	LAAP	..NDGV	ALRSGL	.....	.VAA			
Termitomyces.sp.KNZ76543.1	MR	IVGF	LL	SSVFL	LVANV	..SLLY	QDANYEMA	.....	.A.RHV.DKA			
Psathyrella.aberdarensis.RXW2148	MR	FCNFS	VIA	LTLT	VTSASP	..PSEI	RAESEFAGP	.....	.W.KE..TPQ			
Hebeloma.cylindrosporium.KIM43538	MR	LSSTA	LV	LSLL	VALSAP	..NDAM	EARGGEEV	.....	.A.REPATLVN			
Hypoholoma.sublateritium.KJA22546	MR	VVSFT	LV	VV	SALS	AVFAAP	..NDA	..	.L			
	40			50		60						
Coprinopsis.cinerea.490115	REPQMR	.....	.....	PF	WR	RP	BPQS	RPPS	WR	RP	PEPE	
Pleurotus.ostreatus.KDQ23127.1	REASPEANPRFR	.....	.....	.....	SP	WR	APGR	WPPS	WR	DAE	AEF	
Auriculariopsis.ampla.TRM58449.1	STSDWRREEAR	YGG	.....	.....	SD	WR	EAR	GYGR	WR	DEAR	DEAR	
Armillaria.gallica.PBK88411.1	REPANQASAR	RRRGPANQ	.....	.....	SA	WR	PANQA	SA	WR	PPANQAS		
Amanita.thiersii.PFH51582.1	REP	GVVA	.....	.....	VP	IL	REP	PELP	TRN	WRA	ATP	
Gloeophyllum.trabeum.XP_00787023	RASDWKREPEPEPE	PASDWK	RASDWK	RASDWK	RASDWK	RASDWK	RASDWK	RASDWK	RASDWK	RASDWK	RASDWK	
Heliocybe.sulcata.TFK50419.1	RASDWKREPE	.....	.....	ASD	WR	REP	PKR	..GSD	WR	REP	PASD	
Coprinopsis.marcescibilis.TFK229	RDPIFK	.....	.....	SSS	WR	REP	ADPQG	GRPN	QR	.....	.....	
Exidia.glandulosa.KZV96953.1	REP	EVH	.....	.....	PD	WR	REP	AVH	..NPD	WR	REP	VHP
Psilocybe.cyanescens.PPQ90175.1	RDAAPV	.....	.....	.....	NPD	WR	REP	AVH	..NPD	WR	REP	AVH
Termitomyces.sp.KNZ76543.1	RRVTAKG	.....	.....	.....	SP	WR	REP	AVH	..SRG	WR	REP	AVH
Psathyrella.aberdarensis.RXW2148	CAPGWK	.....	.....	TAHS	REP	REP	REP	REP	REP	REP	REP	REP
Hebeloma.cylindrosporium.KIM43538	REAEAG	.....	.....	TF	WR	REP	REP	REP	REP	REP	REP	REP
Hypoholoma.sublateritium.KJA22546	REAEAG	.....	.....	TF	WR	REP	REP	REP	REP	REP	REP	REP
	70			80		90		100				
Coprinopsis.cinerea.490115	PQYKA	.....	.....	PS	WR	REP	REP	REP	REP	REP	REP	REP
Pleurotus.ostreatus.KDQ23127.1	GYGT	.....	.....	EA	WR	.....	AD	WR	REP	REP	REP	REP
Auriculariopsis.ampla.TRM58449.1	AAWRRRDP	.....	NQAS	.....	AA	WR	REP	REP	REP	REP	REP	REP
Armillaria.gallica.PBK88411.1	.....	.....	.....	.....	PD	WR	REP	REP	REP	REP	REP	REP
Amanita.thiersii.PFH51582.1	.....	.....	.....	.....	SD	WR	REP	REP	REP	REP	REP	REP
Gloeophyllum.trabeum.XP_00787023	.....	.....	.....	.....	SD	WR	REP	REP	REP	REP	REP	REP
Heliocybe.sulcata.TFK50419.1	.....	.....	.....	.....	SD	WR	REP	REP	REP	REP	REP	REP
Coprinopsis.marcescibilis.TFK229	.....	.....	.....	.....	SD	WR	REP	REP	REP	REP	REP	REP
Exidia.glandulosa.KZV96953.1	.....	.....	.....	.....	PD	WR	REP	REP	REP	REP	REP	REP
Psilocybe.cyanescens.PPQ90175.1	PIWKREAE	.....	AQRN	.....	PI	WR	REP	REP	REP	REP	REP	REP
Termitomyces.sp.KNZ76543.1	EVAEAKTV	.....	.....	.....	PE	WR	REP	REP	REP	REP	REP	REP
Psathyrella.aberdarensis.RXW2148	PGWRKDAEP	GGPPGWR	READAD	PAPPG	WR	REP	REP	REP	REP	REP	REP	REP
Hebeloma.cylindrosporium.KIM43538	PGWRKDAEP	GGPPGWR	READAD	PAPPG	WR	REP	REP	REP	REP	REP	REP	REP
Hypoholoma.sublateritium.KJA22546	PGWRKDAEP	GGPPGWR	READAD	PAPPG	WR	REP	REP	REP	REP	REP	REP	REP
	110			120		130						
Coprinopsis.cinerea.490115	PS	.....	WR	REP	.....	PE	REP	REP	REP	REP	REP	REP
Pleurotus.ostreatus.KDQ23127.1	AD	.....	WR	REP	.....	EA	REP	REP	REP	REP	REP	REP
Auriculariopsis.ampla.TRM58449.1	SD	.....	WR	REP	.....	EA	REP	REP	REP	REP	REP	REP
Armillaria.gallica.PBK88411.1	AA	.....	WR	REP	.....	EA	REP	REP	REP	REP	REP	REP
Amanita.thiersii.PFH51582.1	PVQS	.....	WR	REP	.....	EA	REP	REP	REP	REP	REP	REP
Gloeophyllum.trabeum.XP_00787023	SD	.....	WR	REP	.....	EA	REP	REP	REP	REP	REP	REP
Heliocybe.sulcata.TFK50419.1	SD	.....	WR	REP	.....	EA	REP	REP	REP	REP	REP	REP
Coprinopsis.marcescibilis.TFK229	PNQGR	.....	WR	REP	.....	EA	REP	REP	REP	REP	REP	REP
Exidia.glandulosa.KZV96953.1	PD	.....	WR	REP	.....	EA	REP	REP	REP	REP	REP	REP
Psilocybe.cyanescens.PPQ90175.1	PI	.....	WR	REP	.....	EA	REP	REP	REP	REP	REP	REP
Termitomyces.sp.KNZ76543.1	PG	.....	WR	REP	.....	EA	REP	REP	REP	REP	REP	REP
Psathyrella.aberdarensis.RXW2148	PN	.....	WR	REP	.....	EA	REP	REP	REP	REP	REP	REP
Hebeloma.cylindrosporium.KIM43538	PG	.....	WR	REP	.....	EA	REP	REP	REP	REP	REP	REP
Hypoholoma.sublateritium.KJA22546	PG	.....	WR	REP	.....	EA	REP	REP	REP	REP	REP	REP
	140			150		160						
Coprinopsis.cinerea.490115	DP	.....	PE	REP	REP	REP	REP	REP	REP	REP	REP	REP
Pleurotus.ostreatus.KDQ23127.1	DW	.....	KR	REP	REP	REP	REP	REP	REP	REP	REP	REP
Auriculariopsis.ampla.TRM58449.1	DAQGYRTSD	WR	REP	REP	REP	REP	REP	REP	REP	REP	REP	REP
Armillaria.gallica.PBK88411.1	EA	.....	FP	MN	QPAAG	WR	REP	REP	REP	REP	REP	REP
Amanita.thiersii.PFH51582.1	EP	.....	F	AG	P	TRS	WR	REP	REP	REP	REP	REP
Gloeophyllum.trabeum.XP_00787023	AS	.....	DW	K	RASD	WR	REP	REP	REP	REP	REP	REP
Heliocybe.sulcata.TFK50419.1	GG	.....	PD	WR	KRG	PD	WR	REP	REP	REP	REP	REP
Coprinopsis.marcescibilis.TFK229	EP	.....	EA	REP	REP	REP	REP	REP	REP	REP	REP	REP
Exidia.glandulosa.KZV96953.1	EP	.....	EA	REP	REP	REP	REP	REP	REP	REP	REP	REP
Psilocybe.cyanescens.PPQ90175.1	EA	.....	EA	REP	REP	REP	REP	REP	REP	REP	REP	REP
Termitomyces.sp.KNZ76543.1	EA	.....	EV	DG	NTP	WR	KRG	.....	.....	.....	.....	.....
Psathyrella.aberdarensis.RXW2148	DA	.....	EA	REP	REP	REP	REP	REP	REP	REP	REP	REP
Hebeloma.cylindrosporium.KIM43538	EA	.....	EA	REP	REP	REP	REP	REP	REP	REP	REP	REP
Hypoholoma.sublateritium.KJA22546	EA	.....	VP	QP	GPP	WR	REP	REP	REP	REP	REP	REP
Coprinopsis.cinerea.490115	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
Pleurotus.ostreatus.KDQ23127.1	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
Auriculariopsis.ampla.TRM58449.1	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
Armillaria.gallica.PBK88411.1	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
Amanita.thiersii.PFH51582.1	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
Gloeophyllum.trabeum.XP_00787023	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
Heliocybe.sulcata.TFK50419.1	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
Coprinopsis.marcescibilis.TFK229	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
Exidia.glandulosa.KZV96953.1	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
Psilocybe.cyanescens.PPQ90175.1	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
Termitomyces.sp.KNZ76543.1	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
Psathyrella.aberdarensis.RXW2148	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
Hebeloma.cylindrosporium.KIM43538	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
Hypoholoma.sublateritium.KJA22546	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....



Chapter 2

S10c)

C. cinerea. 365456  
 Coprinus. phaeopunctatus. 1316422  
 Coprinus. phaeopunctatus. 1180993  
 Coprinus. phaeopunctatus. 1163880  
 Laccaria. amethystina. KIK08407.1  
 Cyathus. striatus. KAF9008876.1  
 Crucibulum. laeve. TFK43346.1  
 Lyophyllum. atratum. KAF8076205.1  
 Asterophora. parasitica. KAG564866

1 10 20 30 40  
 . . M K L D N F G A L V S L A V L A C S Y G P T Q V R A D S A E N . H H F V R H N N A . R H N A I . . L  
 M A V A K L D N D F A V S L A L A C S Y G P T Q V R A D S A E N . H H F V R H N N A A A H G L V . . A  
 M A V A K L D N D F A V S L A L A C S Y G P T Q V R A D S A E N . H H F V R H N N A A A H G L V . . A  
 M A V A K L D N D F A V S L A L A C S Y G P T Q V R A D S A E N . H H F V R H N N A A A H G L V . . A  
 M A V A K L D N D F A V S L A L A C S Y G P T Q V R A D S A E N . H H F V R H N N A A A H G L V . . A  
 M A V A K L D N D F A V S L A L A C S Y G P T Q V R A D S A E N . H H F V R H N N A A A H G L V . . A  
 M A V A K L D N D F A V S L A L A C S Y G P T Q V R A D S A E N . H H F V R H N N A A A H G L V . . A  
 M A V A K L D N D F A V S L A L A C S Y G P T Q V R A D S A E N . H H F V R H N N A A A H G L V . . A  
 M A V A K L D N D F A V S L A L A C S Y G P T Q V R A D S A E N . H H F V R H N N A A A H G L V . . A

50 60 70 80 90  
 R R Q S E P P T T N R R C R A R P S D N V D . . . . . A P P A Y T P P A D S D V P Q P Q P Q P E Q P A E  
 K R Q A . A K T P N R R R C K A . P T S S G L A S S A A A A P T S P A G E N P G E Q A P T P A P Q . .  
 K R Q A P T A T S A K R R C K V R S T S V A I . . . . . P A P T S P A E Q A P P Q O D P S T P A P S  
 K R Q A . . K T P N R R R C K A . P T S S G L A S S A A A A P T S P A G E N P G E Q A P T P A P Q . .  
 K K K . . . . . R G T I A R R C K P R P S A L . . . . . S S A A K P T T T P A P A . . . . .  
 K K K . . . . . R A S T S K R C K P R P T S S I A A T V V Q T S A A I S K A P A T S E V P V I S K A P V T T  
 K K K . . . . . R A S T S K R C K P R P S S V V . . . . . V A T P K T T S A  
 K K K . . . . . R A A N P K R C R P R V S S T V V . . . . . K P T V A P T T A P A P V T T  
 K K K . . . . . R A S T S K R C R P R P T T S A . . . . . N V P P Q S P P V A K P T T T

100 110 120 130 140  
 Q P P A E Q P P A E Q P P A Q Q P . . . . . A P A P A P D N G N G G D T G N N G G N N G C N N G C N N G G  
 . . . . . Q P A P S P S P E P A F . . . . . A P A P A P P P . . . . . S T P S N G G . . . . .  
 Q D A P A P A P S Q D S P T P A P S Q D A P A P S Q D T P A P A P S Q D T P A P P S N G G A K . . . . .  
 . . . . . Q P A P S P S P E P A F . . . . . A P A P A P P P . . . . . S T P S N G G . . . . .  
 . . . . . P A P T T T P R A M P . . . . . A T S T K A S G . . . . . G S G T P S N T G . . . . . G Y  
 Q S E S Y K S S S S S T A V . . . . . V T T K A S T S . . . . . N S N A T S S S . . . . . S S  
 A S A P A C T P K T T P A P A . . . . . A T T S K R A S . . . . . S G S G S N T G N T N T  
 K P A T T K P A T T K P . . . . . P A P . . . . . T T S S N P P P . . . . . P T N G C G N T G S  
 A V A T T S E P P P P P P Q P A E . . . . . T S T K A P P P A P S P P A P S A P S N N G P T N G C G N N G S

150 160 170 180 190  
 G G G T G S D F N R K M G A W A N P D S K L A N F L H G S V G A Y Y W S P K P P V I N S R P D I  
 . . . . . K V G A W A N P D N E L A N M M G R Y S A Y Y W S P K P P I L N N Y F Q I  
 . . . . . K V G A W A N P D N E L A N M M G R Y S A Y Y W S P K P P I L N N Y F Q I  
 . . . . . K V G A W A N P D N E L A N M M G R Y S A Y Y W S P K P P I L N N Y F Q I  
 N G P P G . . . . . K V G A Y A G P D D G T I S H F I T P K V T A I Y W S P A P A S A A L G L  
 S G S G . . . . . K V G A W N N A D E S S I K N F V N G Q V S A I Y W S P K P P V V D S L G L  
 N P G A . . . . . K V G A W P N G D D S L N F K T N K V S R I Y W S P K P P A L S D S L G L  
 D G G V . . . . . K V G A W P Q D D K A L A N F K T G K V S P I Y W S P I P C A A K K L G F  
 N G G P G . . . . . K V G A W P I N D K A L A N F K T S K V S P I Y W S P I P C A A K E L G E

200 210 220 230 240  
 E F I P M L W G E K H I N D F V L V Q F G W A R T V I G F N E P D H H G O A F D P Y Y A V D I P M Q  
 D F I P M L W G E K Q I E D F Y R W V K P G W A R Y I A G F N E P D G G C S N D L D V G Y A V D I W Q  
 D F I P M L W G E K Q I E D F Y R W V K P G W A R Y I A G F N E P D G G C S N D L D V G Y A V D I W Q  
 D F I P M L W G E K Q I E D F Y R W V K P G W A R Y I A G F N E P D G G C S N D L D V G Y A V D I W Q  
 E F I P M L W G E K Q I G D F T Q L V K P G W A R T V I G F N E P N S G G A E M T P A R A A Q W W Q  
 E F I P M L W G E K Q I S D F T S L V K A G W A R T V I G F N E P N S G G A E M T P A R A A Q W W Q  
 E F I P M L W G E K Q I A D F Q R L V K P G W A R T V I G F N E P N Q G G S D M S P Q R A A Q W W Q  
 E F I P M L W G E K Q A T D F K R L V K G W A R T V I G F N E P N Q G G S D M T P Q R G A Q W W Q  
 E F I P M L W G E K Q A E F K R L V K G W A N T V I G F N E P N Q G G S D M S P Q R G A Q W W Q

250 260 270 280 290  
 Y I O P L K S Q G Y T L I S P A C T N S P A G T K W L Q F I S A C N N R G C N V D I V A H W Y G T D  
 H I O P L K S Q G Y T L I S P A C T N S P A G T K W L Q F I S A C N N R G C N V D I V A H W Y G T S  
 H I O P L K S Q G Y T L I S P A C T N S P A G T K W L Q F I S A C N N R G C N V D I V A H W Y G T D  
 H I O P L K S Q G Y T L I S P A C T N S P A G T K W L Q F I S A C N N R G C N V D I V A H W Y G T S  
 Y I O P L K N Q G Y T L I S P A C T N S P A G T K W L Q F I S A C . . . . . G C T P D I V A H W Y G T S  
 Y I O P L K N Q G Y T L I S P A C T N S P A G T K W L Q F I S A C . . . . . G C T P D I V A H W Y G T D  
 N I O P L K N S G M A L I S P A C T S A P S K T W M K L P F A C N . . . . . S P S D G A H W Y G T D  
 Y I O P L K A Q G Y S L I S P A C T N A P S K K W L Q F F C A C K . . . . . G C T A D G V A L H E Y G T D  
 Y I O P L K A Q G Y S L I S P A C T N A P S K K W L K D F F R A C D . . . . . G C T F D G V A L H E Y G T D

300 310 320 330 340  
 S Q N E I D H V T N Y H N R W G R E V V V T E F A C N R S G G C Q C D N A R V W G F M S T V T E F W M  
 S D A F I Q H V E S H N T F N R E V V V T E F A C N R S G G C Q C D A G Q V W F M Q T V T W M  
 S Q N E I D H V S Q H S V F G R E V V V T E F A C N R S G G C Q C S D Q L W S F M K V T D W M  
 S D A F I Q H V E S H N T F N R E V V V T E F A C N R S G G C Q C D A G Q V W F M Q T V T W M  
 P Q D E I N Y I T D H N T F N L D V F V T E F A C N R S G G C Q C S D V T E F M D T V T F M  
 P Q E E I T Y L O D E H N T F N L P I W F T E F A C N R S G G C Q C T S D Q V W F M Q T V K D F M  
 A Q E E I K Y V T D E H T F N L P I W F T E F A C N R S G G C Q C T K D Q V F S F M D T V T F M  
 P Q A M N Y L I D M H N T F N L P I W F T E F A C N R S G G C Q C S K D D V F F M S K K T F M  
 P Q V E I D V I T D H N T F N L P I W F T E F A C N R S G G C Q C S R D Q V F G F M S K V K N F M

350 360 370 380 390  
 K R T H F V G H Y F F G V L R D M V G V F E G N R L Y N P A N D Q P H D L G W A Y I S  
 D G Q W Y G R Y F F G V L H D M Y N V N Y A N Q L M G G . G G Q N D L G R E Y I Y  
 D N T S F V G R Y F F G V L H D M G N V N Y A N Q L M N S . R G Q N D L G R Y B I Y  
 D G Q W Y G R Y F F G V L H D M Y N V N Y A N Q L M G G . G G Q N D L G R E Y I Y  
 D S T S W V K H Y F F G L Q H D M V G V N P A N Q L L A S . N G Y P T A L G R G Y I N  
 D N T S W V G H Y F F G V L H D M A G V N S L D Q L L G S . D G Q P T A L G R I Y T Q  
 D N S P V A A Y F F G V M H D M V G V N I L D Q L M G P . N G P T D L G Y S Y I N  
 D E T P W L K H Y F F G A M D M G N V N P L N Q L L G N . D G R P T D L G Y M I N  
 D N T S W V A H Y F F G A M Y D M N V N P L N Q L L G S . N G R P E D L G Y I N

S10d)

C. cinerea. 447393  
 Crassisporium. funariophilum. KAF8  
 Coprinellus. micaceus. TEB25085.1

1 10 20 30 40 50 60  
 M R F S N S V A L L V F A F L S I S L A S F V P G E A L V N R G A L D D L E R L P A R R E L V D R G I T P P W R D H V  
 M R S R R A A F L A F T F T I S V H A S I S V E L L N G R S A V D . . . . . V V V R E A D T I A G P S A W R R . .  
 M R P E T A A V V T A L S L F S V V S A R P V D . . . . . L A A R G P P S . . . . . W K R D E L D A R . G P S W R D .

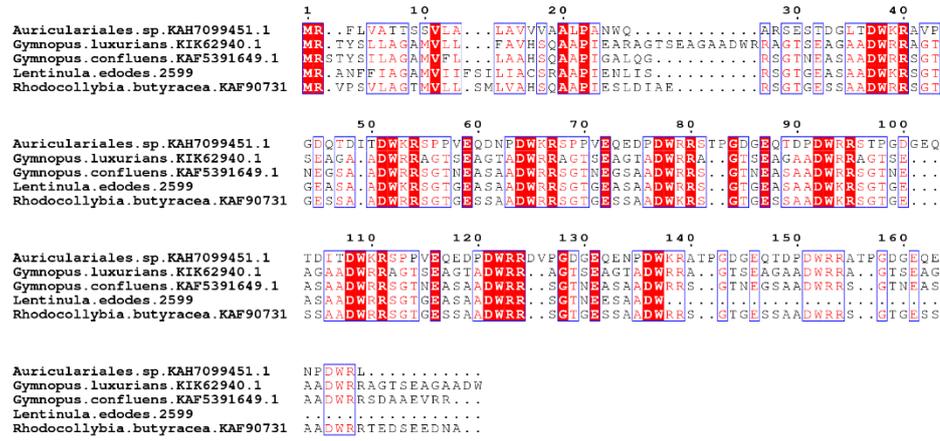
70 80 90 100  
 E P T P A G I T P P W R D N L P P T P V P P V R R R . . . . . D S R P . . . . . P P A W R P I D S E T S D S  
 . . . . . D D A V A L P L K R E S V V S N G P P A W R R . . . . . A A T P G P T . . . . . P P S W R R . . . . .  
 . . . . . E I D A R G P P S W K R D E V D A R G P P G W R R E V D A R G P P S W K R E I D V R G P P S W R R . . . . . D

110 120 130 140 150 160  
 E L E S R P P P A W R P V E A E S R P . P P A W R P V K R E A D A R P P A W R P I D S D D D P E G P P A P S W R R  
 . . . . . A P P S P D W R R A S S G S A P P P A W R . . . . . R G S N E P A G P P A W R R E N V . . . . . G T S G P P . . . . . W R R  
 E V D A R G P P S W K R D E V D A R R . P P S W R . . . . . R E E V D A R G P P S W R D E V . . . . . D A R G P P S . . . . . W R R

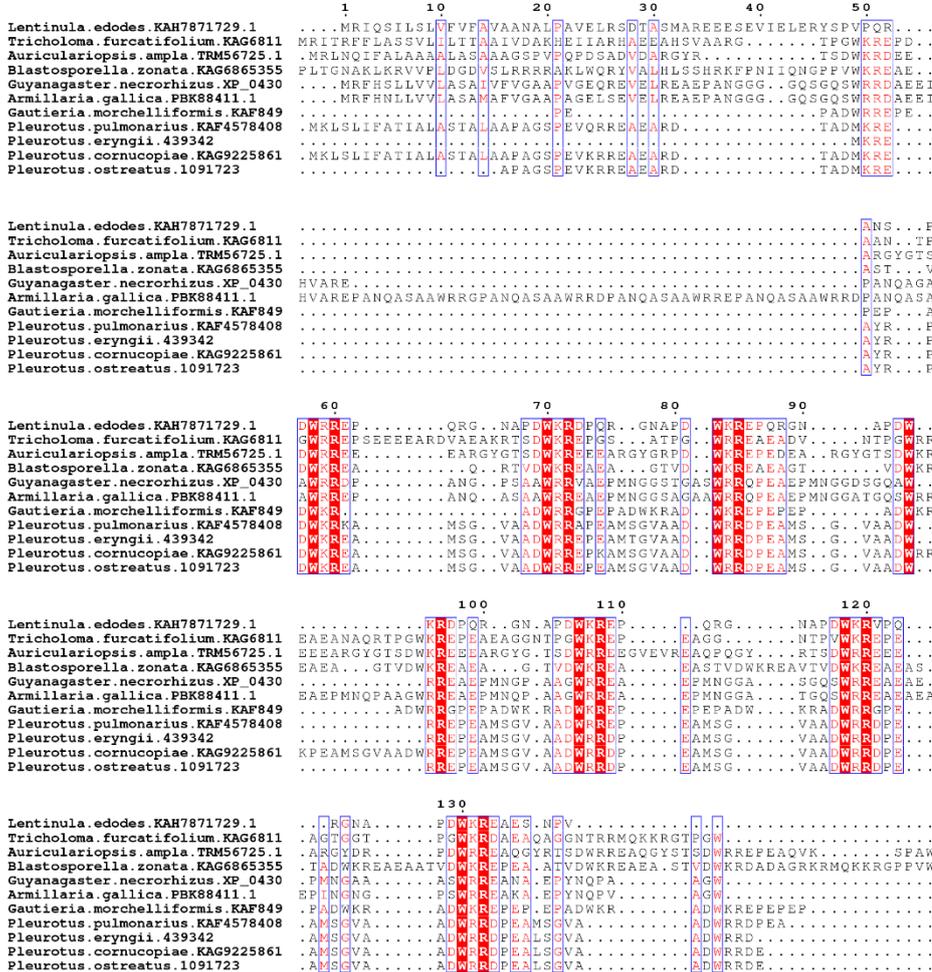
170  
 E T V A I A A G V I A P D W . . . . .  
 S . . . . . D P A T . . . . . P S W . . . . .  
 E . . . . . V A R G . . . . . P S W K R G P P S W

## Chapter 2

S10e)



S10f)

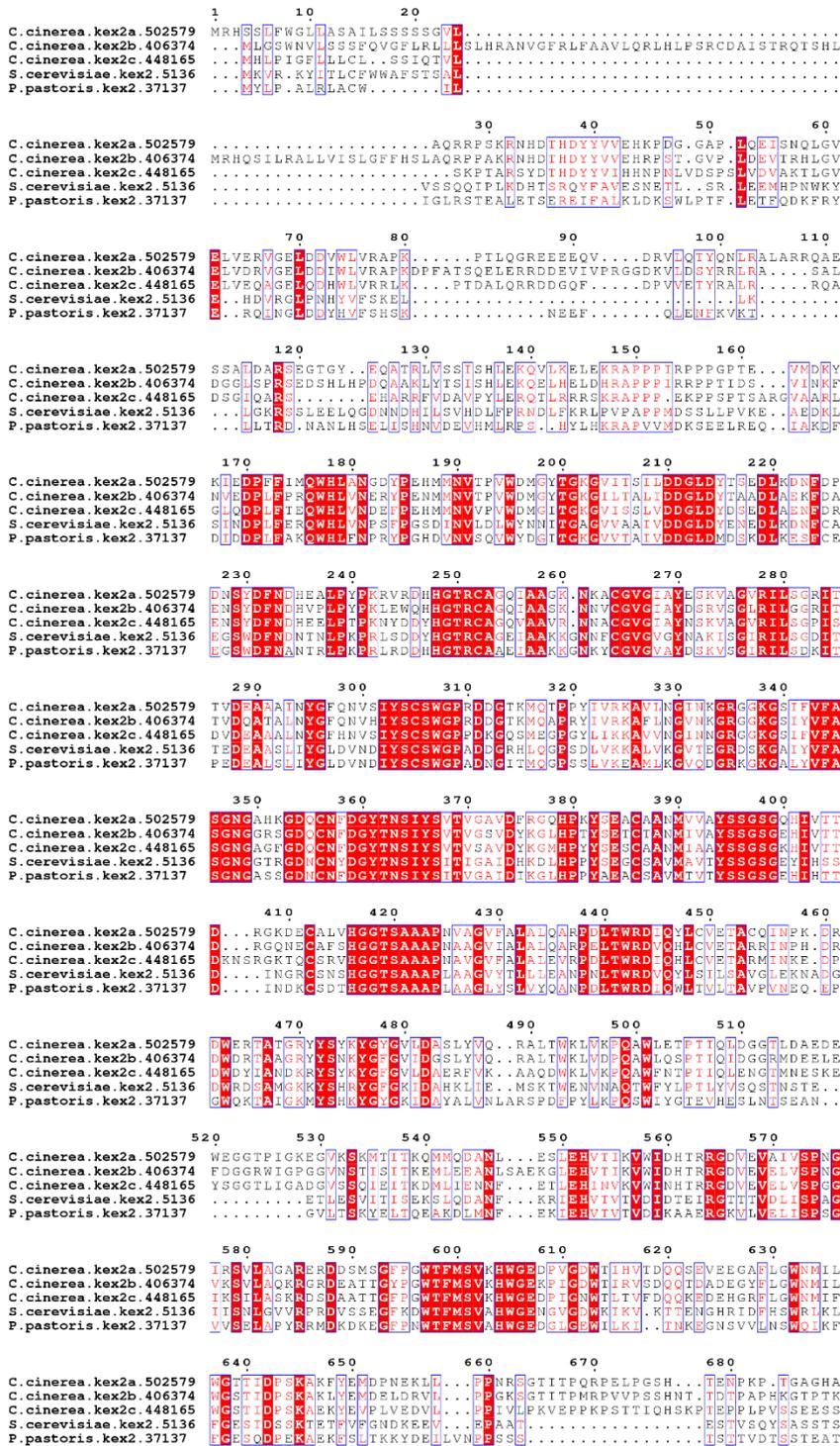


**Figure S10:** Alignments of analyzed KEX2-cleaved proteins with homologs in other fungal species. BLASTs were carried out using NCBI (<https://www.ncbi.nlm.nih.gov>) or JGI (<https://mycocosm.jgi.doe.gov>), alignments were carried out using MUSCLE (MULTiple Sequence Comparison by Log- Expectation) (Pearce et al. 2022) (<https://www.ebi.ac.uk/Tools/msa/muscle/>), and the graphic rendering of the alignment results using ESPrnt (Easy Sequencing in PostScript) (<https://esprnt.ibcp.fr>) (Robert and Gouet 2014). Shown are the alignments of a) *C. cinerea* KEP 490115, b) *C. cinerea* KEP 497993, and c) *C. cinerea* protein 365456 that contains a C-terminal glycoside hydrolase domain (starting at position 153) with two KEX2 cleavage sites (starting at position 46)

# Chapter 2

flanking the sequences QAAKTPN, QAPTATSA, QAKTPN in *Coprinus phaeopunctatus* and GTTA, ASTS, ASTS, AANP, ASTG in five other fungal species. Subpanel d) shows *C. cinerea* KEP 447393, e) *Lentinula edodes* KEP 2599, and f) the homologs *Pleurotus ostreatus* KEP 1091723 and *Pleurotus eryngii* KEP 439342.

S11a)



## Chapter 2

```

690      700      710      720      730      740
C. cinerea. kex2a.502579 IEGKPTADLPHDSTSPDKLGDADAGK..KPENKKGWIGGKDKAAKQTM..AGLIGVLV
C. cinerea. kex2b.406374 FKSPFHEIHKAKPS.PNPVLPATSQKEDGDGEKSTFHKIKDTAIEHSW..SGLVFFLF
C. cinerea. kex2c.448165 VEASTTFLSPTESAPASASASFTQNI..PTIDQGWFFDMSKLIANQKMF..FGAIGAVA
S. cerevisiae. kex2.5136  LSISATSTLSSISIGVETSALPQITITASTDFDSDPNTPFKLSSPRQAMHYFLTIFLIGATF
P. pastoris. kex2.37137  FSSSSVSEHSAT...ETDNRKETSITIGDEDFENEKGSYKHLNSSHITEL..AFLIGLGF

750      760      770      780      790      800
C. cinerea. kex2a.502579 LANIAGAIYFCCW...KKGCGDGEKYNALPGEENVPMSINGGGYDQDGERIALVGGG
C. cinerea. kex2b.406374 VGSIIISIAFCCW..RRRGGAINQYVFLSAVEEESVP.....EDAE
C. cinerea. kex2c.448165 VFGIAGVAFF..W..RRKLARRANTEYSAL.ANDELMSITVSGPLTGHARTTRELVDAGF
S. cerevisiae. kex2.5136  LVLYYFMFFMKSR..RRIRRSRAETYEFDITDTSSEYDSTLDNGTSGITEPEEVEDDFD
P. pastoris. kex2.37137  LTCITLFTNRRKLERQRNRRRDEYEFDLIPADDFD.....TEEDQEANSQFT

810      820      830      840      850
C. cinerea. kex2a.502579 DEEEDGVDTRFAARAQQTGFHPQARSTIGGLGFHSGFD DDDDPMTAASPR..KDGSSVVQSA
C. cinerea. kex2b.406374 DVSSDEDEDETAALRPGQTQPL...GGQLGYHSGFLDDGEPVSGPTATTYRDEPNPEDA
C. cinerea. kex2c.448165 LSDEDH.....LALSSSSSENGDAEHTI..D..SVLTN
S. cerevisiae. kex2.5136  LDSDAE.....LMEFEDTSQREASPHYED..SLGSN
P. pastoris. kex2.37137

860      870      880      890      900      910
C. cinerea. kex2a.502579 VGELIPTGDAPVAAAVSSSGPQPSVQQQRNLVSESEEEDEDEDDDEDEDEDEDEDE
C. cinerea. kex2b.406374 NGQAREGGSQOPANMISEQ.....LEPERSIASKSEATKISE.....
C. cinerea. kex2c.448165 ETPTQPLPRSETATSITST.....TASPSSGGSSWEHASS.....
S. cerevisiae. kex2.5136  ENPFSDPKQKFPNDANAESASNKLEQLQFDVPPSSGRS.....
P. pastoris. kex2.37137  EHPKRAAL.....

920      930      940
C. cinerea. kex2a.502579 SSEEDESGSQQFLQQNQVRRPGEVDSSASLI
C. cinerea. kex2b.406374 .....
C. cinerea. kex2c.448165 .....
S. cerevisiae. kex2.5136  .....
P. pastoris. kex2.37137  .....

```

S11b)

	<i>C. cinerea. kex2c.448165</i>	<i>C. cinerea. kex2a.502579</i>	<i>C. cinerea. kex2b.406374</i>	<i>S. cerevisiae. kex2.5136</i>	<i>P. pastoris. kex2.37137</i>
<i>C. cinerea. kex2c.448165</i>	100	54.08	51.59	37.16	39.71
<i>C. cinerea. kex2a.502579</i>	54.08	100	61.33	34.67	38.04
<i>C. cinerea. kex2b.406374</i>	51.59	61.33	100	36.54	38.47
<i>S. cerevisiae. kex2.5136</i>	37.16	34.67	36.54	100	47.28
<i>P. pastoris. kex2.37137</i>	39.71	38.04	38.47	47.28	100

Chapter 2

S11c)

```

S. cerevisiae.kex1.2361 .....1 10 20 30
P. pastoris.kex1.36999 .....1 10 20 30
C. cinerea.497730 .....1 10 20 30
C. cinerea.437675 .....1 10 20 30
C. cinerea.71140 .....1 10 20 30
C. cinerea.496475 .....1 10 20 30

S. cerevisiae.kex1.2361 .....40 50 60
P. pastoris.kex1.36999 .....40 50 60
C. cinerea.497730 .....40 50 60
C. cinerea.437675 .....40 50 60
C. cinerea.71140 .....40 50 60
C. cinerea.496475 .....40 50 60

S. cerevisiae.kex1.2361 .....70 80 90 100 110 120
P. pastoris.kex1.36999 .....70 80 90 100 110 120
C. cinerea.497730 .....70 80 90 100 110 120
C. cinerea.437675 .....70 80 90 100 110 120
C. cinerea.71140 .....70 80 90 100 110 120
C. cinerea.496475 .....70 80 90 100 110 120

S. cerevisiae.kex1.2361 .....130 140 150 160 170 180
P. pastoris.kex1.36999 .....130 140 150 160 170 180
C. cinerea.497730 .....130 140 150 160 170 180
C. cinerea.437675 .....130 140 150 160 170 180
C. cinerea.71140 .....130 140 150 160 170 180
C. cinerea.496475 .....130 140 150 160 170 180

S. cerevisiae.kex1.2361 .....190 200 210 220 230 240
P. pastoris.kex1.36999 .....190 200 210 220 230 240
C. cinerea.497730 .....190 200 210 220 230 240
C. cinerea.437675 .....190 200 210 220 230 240
C. cinerea.71140 .....190 200 210 220 230 240
C. cinerea.496475 .....190 200 210 220 230 240

S. cerevisiae.kex1.2361 .....250 260 270 280 290
P. pastoris.kex1.36999 .....250 260 270 280 290
C. cinerea.497730 .....250 260 270 280 290
C. cinerea.437675 .....250 260 270 280 290
C. cinerea.71140 .....250 260 270 280 290
C. cinerea.496475 .....250 260 270 280 290

S. cerevisiae.kex1.2361 .....300 310
P. pastoris.kex1.36999 .....300 310
C. cinerea.497730 .....300 310
C. cinerea.437675 .....300 310
C. cinerea.71140 .....300 310
C. cinerea.496475 .....300 310

S. cerevisiae.kex1.2361 .....320 330 340 350
P. pastoris.kex1.36999 .....320 330 340 350
C. cinerea.497730 .....320 330 340 350
C. cinerea.437675 .....320 330 340 350
C. cinerea.71140 .....320 330 340 350
C. cinerea.496475 .....320 330 340 350

S. cerevisiae.kex1.2361 .....360 370 380 390 400
P. pastoris.kex1.36999 .....360 370 380 390 400
C. cinerea.497730 .....360 370 380 390 400
C. cinerea.437675 .....360 370 380 390 400
C. cinerea.71140 .....360 370 380 390 400
C. cinerea.496475 .....360 370 380 390 400

S. cerevisiae.kex1.2361 .....410 420 430 440 450 460
P. pastoris.kex1.36999 .....410 420 430 440 450 460
C. cinerea.497730 .....410 420 430 440 450 460
C. cinerea.437675 .....410 420 430 440 450 460
C. cinerea.71140 .....410 420 430 440 450 460
C. cinerea.496475 .....410 420 430 440 450 460

S. cerevisiae.kex1.2361 .....470 480 490 500 510 520
P. pastoris.kex1.36999 .....470 480 490 500 510 520
C. cinerea.497730 .....470 480 490 500 510 520
C. cinerea.437675 .....470 480 490 500 510 520
C. cinerea.71140 .....470 480 490 500 510 520
C. cinerea.496475 .....470 480 490 500 510 520

```

## Chapter 2

```

          530      540      550      560      570      580
S.cerevisiae.kex1.2361 FKENLEEEQEAQNEEGKEKEGNKDKGDDDDNDNDDDDDDHNSGDDDDDDDDDDDDNN
P.pastoris.kex1.36999 PRL.....QNGPKSSSTDDSAAHGN.....
C.cinerea.497730      ..EGKNLPAGPEIYMGEQATQST.....
C.cinerea.437675      PQP.....STPALPAGKTPAQDKAMWEA.....
C.cinerea.71140      .....
C.cinerea.496475      .....

          590      600      610      620      630      640
S.cerevisiae.kex1.2361 EKQSNQGLSDSRHKSSEYEQEEEEVEEFEEIISMYKHKAVVVIIVTFLIVVLGVYATDRR
P.pastoris.kex1.36999 .....PFFYYVFELFVIVLLCGLVLYQY...
C.cinerea.497730      .....HVMPEA
C.cinerea.437675      .....YYNAGSAALVVLVLMVIGLFIW...
C.cinerea.71140      .....W...
C.cinerea.496475      .....W...

          650      660      670      680      690      700
S.cerevisiae.kex1.2361 VRRKARHTIIVDFNNRQHDSPNKIVSWADDLESGLGAEDDLEQQDEQLEGGAPISSTSNKA
P.pastoris.kex1.36999 YSNSAPHSITADKHKKSKNKSKNVRFLLDLESNLDLDNTDDKDNSVMSKLLSSMGYQA
C.cinerea.497730      TRAAWQSFTITETATTPAAAP.....TSTS
C.cinerea.437675      YRRKSRQLQIPSNQSGELAEESIPLRSEMEDRGGNGVSNNGAWKKGKARASEPVFEVGDSD
C.cinerea.71140      IFDLPLNLPVGAELGSIPLP.....
C.cinerea.496475      FDGRKFT.....

          710      720
S.cerevisiae.kex1.2361 GSKLTKKKKKYISLPNTEIDESFEMTDF
P.pastoris.kex1.36999 QEPYKPLDKGANADL.DIEMD.SHGTSEK
C.cinerea.497730      SSRVRPTRRR.....
C.cinerea.437675      EDEPSPYRRPDGGGDRNV.....
C.cinerea.71140      .....
C.cinerea.496475      .....

```

**Figure S11:** Alignments of *C. cinerea* KEX proteases with homologs in other fungal species. BLASTs were carried out using NCBI (<https://www.ncbi.nlm.nih.gov>) or JGI (<https://mycocosm.jgi.doe.gov>). Shown are a) KEX2 alignments, b) the percent identity matrix calculated by MUSCLE (Pearce et al. 2022) of KEX proteases of *S. cerevisiae*, *P. pastoris* and *C. cinerea*, and c) alignments of KEX1.

Chapter 2

S12a)

```
1          10          20          30
P.pastoris.ANZ77179.1  MRIP...TGRKTA...LPHYPHWSLLLTTFMYF...EHKYRDYK...
Coprinopsis.cinerea.545529  MRIS...EDDVLHRRSSPGTPAE...MRQTSQATVVDPA...LIR...SVYVNDGPPDAPGSSD...EA
L.edodes.GAW08536.1  MRIP...SYERLE...TDDPSE...TSTIVSSVHNSHNSVAK...FATYYGGQPPDAPSSD...EE
L.edodes.XP_046085150.1  MRIP...SYERLE...TDDPSE...TSTIVSSVHNSHNSVAK...FATYYGGQPPDAPSSD...EE
P.ostreatus.KDQ26420.1  ...
P.eryngii.KAF9497814.1  MHVP...SYEAVPNEAVPGGTFDSFGGSKLYEHPNQASS...IRP...FAYYEDGPPDPPSSD...DE
S.cerevisiae.STE13 Dipeptidyl.58  MSAS...TH...SHKRRN...SHLFPQRKSSNSNDK...FFPNNDVANTDQPSN...NG
P.pastoris.ANZ75207.1  MTSR...AENPFD...IELQEN...LSPRSNSNSILEN...LNE...YARRHRNDSLSQDCDNE...EN

P.pastoris.ANZ77179.1  ...
Coprinopsis.cinerea.545529  ESLLD...DDDKDRPDSPGMARLGLGARSSTVDGR
L.edodes.GAW08536.1  DE...RKG...PSTPGIAEAGFNAPY...
L.edodes.XP_046085150.1  DE...RKG...PSTPGIAEAGFNAPY...
P.ostreatus.KDQ26420.1  ...
P.eryngii.KAF9497814.1  TLLEK...DGALEVNEDERRRYVFTG...
S.cerevisiae.STE13 Dipeptidyl.58  HT...EIRPTEATIDVTDVQTFP...EQYGMRRP...RFSFQNDIENQHHTHSFFSVNKFNR
P.pastoris.ANZ75207.1  EN...Y...TDNLTFLSKSGVS...

40          50          60          70
P.pastoris.ANZ77179.1  ...RRVPL...Q...S...L...V...L...V...G...S...L...T...I...N...I...H...Y...N...H...K...N...Y...H...A...H...P...V...Y...T...
Coprinopsis.cinerea.545529  SMRSP...L...K...G...T...S...A...S...L...R...W...L...I...I...L...L...V...L...V...G...L...V...G...I...G...F...I...A...A...V...I...T...D...A...P...L...V...Q...L...R...K...L...T...M...D...
L.edodes.GAW08536.1  ...P...S...L...R...Q...K...R...A...S...N...R...L...V...I...L...L...V...L...V...L...C...G...I...G...I...A...A...Q...I...Y...V...G...T...V...Y...R...
L.edodes.XP_046085150.1  ...P...S...L...R...Q...K...R...A...S...N...R...L...V...I...L...L...V...L...V...L...C...G...I...G...I...A...A...Q...I...Y...V...G...T...V...Y...R...
P.ostreatus.KDQ26420.1  ...
P.eryngii.KAF9497814.1  ...L...R...D...R...L...G...C...R...R...P...T...S...R...L...V...L...A...F...V...L...V...L...S...L...S...A...I...G...I...V...A...A...Y...Y...S...G...A...N...S...Y...H...
S.cerevisiae.STE13 Dipeptidyl.58  WGEWS...L...P...K...R...S...Y...V...L...V...F...T...L...I...L...S...V...L...L...V...L...I...P...S...K...L...L...P...T...K...I...F...R...P...K...T...S...A...G...D...S...S...L...G...K...R...F...S...I...E...
P.pastoris.ANZ75207.1  ...R...K...G...C...M...L...I...F...G...I...C...F...V...I...W...L...L...F...A...L...Y...A...R...D...N...R...F...S...N...L...N...Y...V...P...D...S...N...S...H...G...T...A...S...I...S...

80          90          100          110          120          130
P.pastoris.ANZ77179.1  NTVLNDFA.EEDS...FSSNG...T...L...N...L...E...N...W...R...N...G...I...F...S...P...K...F...H...S...I...O...I...T...E...V...G...R...D...D...G...O...Y...Y...I...L...S...
Coprinopsis.cinerea.545529  HIFNG...T...F...S...A...R...T...I...Q...L...N...W...V...E...G...L...R...K...I...T...M...D...H...I...F...G...I...T...S...P...O...S...T...S...L...H...M...I...F......E...A...G...D...V...H...V...S...V...I...
L.edodes.GAW08536.1  ...L...P...G...N...R...K...I...T...M...D...H...I...F...G...I...T...S...P...O...S...T...S...L...H...M...I...F......E...A...G...D...V...H...V...S...V...I...
L.edodes.XP_046085150.1  ...L...P...G...N...R...K...I...T...M...D...H...I...F...G...I...T...S...P...O...S...T...S...L...H...M...I...F......E...A...G...D...V...H...V...S...V...I...
P.ostreatus.KDQ26420.1  ...M...D...H...V...F...N...G...I...T...M...D...H...I...F...G...I...T...S...P...O...S...T...S...L...H...M...I...F......E...A...G...D...V...H...V...S...V...I...
P.eryngii.KAF9497814.1  ...A...P...G...V...K...I...T...M...H...V...F...N...G...I...T...M...D...H...I...F...G...I...T...S...P...O...S...T...S...L...H...M...I...F......E...A...G...D...V...H...V...S...V...I...
S.cerevisiae.STE13 Dipeptidyl.58  NVLNG...D...F...A...I...P...E...D...T...F...H...I...D...P...Q...R...L...L...G...Q...D...S...D...P...G...L...Y...F...T...T...K...E......I...D...G...H...T...N...I...A...K...Q...
P.pastoris.ANZ75207.1  SVEQK...Q...T...G...L...P...E...G...D...S...N...S...G...F...I...G...S...K...L...S...I...S...G...W...R...G...L...Y...N...V...Y...P...K...I...S...L...G...E...D...D...Y...Y...E...R...H...

140          150          160          170          180
P.pastoris.ANZ77179.1  ...N...S...S...Y...I...V...K...S...I...D...P...D...F...E...S...V...I...F...N...E...S...T...I...T...Y...N...G...E...H...H...V...E...D...V...I...V...S...N...N...I...Q...A...L...V...V...T...K...K...R...R...H...
Coprinopsis.cinerea.545529  ...D...G...S...I...N...L...V...S...L...S...T...R...K...L...V...D...L...A...D...I...R...E...O...C...N...P...L...A...I...A...D...W...K...V...S...P...D...M...H...H...I...L...V...K...A...D...Y...R...K...O...R...H...
L.edodes.GAW08536.1  ...Q...D...G...F...K...L...V...D...L...K...T...N...T...S...R...L...V...K...K...D...I...K...D...E...H...K...P...L...Q...H...W...A...I...W...L...S...P...D...M...K...H...L...L...V...K...G...D...H...K...O...R...W...
L.edodes.XP_046085150.1  ...Q...D...G...F...K...L...V...D...L...K...T...N...T...S...R...L...V...K...K...D...I...K...D...E...H...K...P...L...Q...H...W...A...I...W...L...S...P...D...M...K...H...L...L...V...K...G...D...H...K...O...R...W...
P.ostreatus.KDQ26420.1  ...D...G...S...I...N...L...V...S...L...S...T...R...K...L...V...D...L...A...D...I...R...E...O...C...N...P...L...A...I...A...D...W...K...V...S...P...D...M...H...H...I...L...V...K...A...D...Y...R...K...O...R...H...
P.eryngii.KAF9497814.1  ...D...G...S...I...N...L...V...S...L...S...T...R...K...L...V...D...L...A...D...I...R...E...O...C...N...P...L...A...I...A...D...W...K...V...S...P...D...M...H...H...I...L...V...K...A...D...Y...R...K...O...R...H...
S.cerevisiae.STE13 Dipeptidyl.58  ...L...F...D...E...T...F...E...V...N...L...G...G...N...R...F......L...Y...B...G...V...E...F...I...V...S...I...V...Q...I...N...Y...K...L...D...L...I...F...G...T...N...L...E...S...E...R...H...
P.pastoris.ANZ75207.1  ...F...H...R...I...D...E...K...K...I...D...E...R...H...D...I...F...N...H...E...K...I...E...V...N...G...I...T...Y...V...S...V...I...L...S...F...Y...S...T...K...F...L...V...A...C...N...Y...E...K...H...R...H...

190          200          210          220          230          240
P.pastoris.ANZ77179.1  SFFAN...Y...W...L...Y...K...V...D...N...P...G...E...V...Q...P...I...F...D...S...D...L...S...L...N...A...L...V...S...L...V...H...S...P...D...S...Q...V...A...F...V...L...N...N...I...Y...L...K...
Coprinopsis.cinerea.545529  SFFGN...Y...Y...V...H...N...L...E...T...K...Q...I...W...P...I...I...P...S...S...P...S...K...I......A...Y...A...T...W...S...P...T...G...E...A...I...Y...V...E...N...D...L...Y...I...L...
L.edodes.GAW08536.1  SFFGN...Y...Y...I...H...D...I...E...A...K...V...T...R...P...M...I...P...P...A...N...P...P...T...I......A...Y...A...T...W...S...P...T...G...N...S...L...A...Y...V...E...N...D...L...Y...I...L...
L.edodes.XP_046085150.1  SFFGN...Y...Y...I...H...D...I...E...A...K...V...T...R...P...M...I...P...P...A...N...P...P...T...I......A...Y...A...T...W...S...P...T...G...N...S...L...A...Y...V...E...N...D...L...Y...I...L...
P.ostreatus.KDQ26420.1  SFFGN...Y...Y...V...H...N...L...E...T...R...I...T...R...P...L...I...T...S...N...P...P...T...I......A...F...A...T...W...A...B...I...G...E...A...I...Y...V...E...N...D...L...Y...I...V...V...
P.eryngii.KAF9497814.1  SFFGN...Y...Y...V...H...N...L...E...T...R...I...T...R...P...L...I...T...S...N...P...P...T...I......A...F...A...T...W...A...B...I...G...E...A...I...Y...V...E...N...D...L...Y...I...V...V...
S.cerevisiae.STE13 Dipeptidyl.58  SFFGN...Y...Y...V...H...N...L...E...T...R...I...T...R...P...L...I...T...S...N...P...P...T...I......A...F...A...T...W...A...B...I...G...E...A...I...Y...V...E...N...D...L...Y...I...V...V...
P.pastoris.ANZ75207.1  SFFAR...Y...F...I...Y...D...Q...S...D...Q...E...N...S...T...Q...V...Y...D...D...K...A...L......S...F...Y...S...S...G...D...H...V...V...F...E...N...N...I...Y...L...K...

250          260          270          280          290          300
P.pastoris.ANZ77179.1  HLNN...F...S...D...S...R...V...Q...I...T...Y...D...G...E...N...I...F...Y...G...K...P...D...W...Y...E...E...V...F...E...S...N...S...A...N...W...S...P...N...G...O...F...L...S...I...R...T...N...D...T...Q...V...
Coprinopsis.cinerea.545529  P...I...A...D...P...S...A...E...H...I...R...V...T...I...T...G...N...A...I...F...F...R...G...V...P...D...W...Y...E...E...V...F...S...G...D...H...A...W...F...S...P...D...S...K...I...A...L...I...S...F...D...E...T...E...V...
L.edodes.GAW08536.1  T...S...P...S...P...S...T...A...P...I...R...V...T...S...G...N...A...S...L...F...H...G...V...P...D...W...Y...E...E...I...F...S...S...D...H...A...W...S...P...D...A...Q...K...V...A...F...I...R...F...D...E...T...A...V...
L.edodes.XP_046085150.1  T...S...P...S...P...S...T...A...P...I...R...V...T...S...G...N...A...S...L...F...H...G...V...P...D...W...Y...E...E...I...F...S...S...D...H...A...W...S...P...D...A...Q...K...V...A...F...I...R...F...D...E...T...A...V...
P.ostreatus.KDQ26420.1  A...D...Q...S...S...P......I...R...V...T...S...G...N...A...S...L...F...H...G...V...P...D...W...Y...E...E...V...F...S...A...S...A...W...F...S...P...D...S...K...L...A...F...I...A...F...D...E...T...V...
P.eryngii.KAF9497814.1  Q...V...N...S...G...V...A......K...K...V...E...D...C...S...K...D...I...F...N...A...K...P...D...W...Y...E...E...V...L...S...D...A...W...W...A...P...D...S...K...A...V...F...A...N...D...T...S...V...
S.cerevisiae.STE13 Dipeptidyl.58  Q...L...S...T...L...K...I......T...Q...V...T...F...D...C...D...E...S...I...Y...N...G...K...P...D...W...Y...E...E...V...L...S...D...A...W...W...N...E...D...S...Y...F...T...F...I...R...L...D...D...S...K...V...
P.pastoris.ANZ75207.1  Q...L...S...T...L...K...I......T...Q...V...T...F...D...C...D...E...S...I...Y...N...G...K...P...D...W...Y...E...E...V...L...S...D...A...W...W...N...E...D...S...Y...F...T...F...I...R...L...D...D...S...K...V...

310          320          330          340
P.pastoris.ANZ77179.1  F...Y...Y...F...I...Y...V...O...S...D...A...E...T...A...T...D...E...F...L...L...K...H...K...Y...E...K...K...G...E...P...N...F...V...D...V...I...V...Y...D...V...
Coprinopsis.cinerea.545529  F...E...F...F...I...Y...N...P...T...E...D...A...T...I...V...N...E...T...T...E...V...V...K...Y...P...K...G...Y...N...N...P...L...V...S...V...H...V...F...D...L...E...R...Y...
L.edodes.GAW08536.1  D...E...F...S...F...I...Y...N...P...T...E...N...N...S...A...V...I...P...T...S...E...I...T...K...Y...P...K...G...Y...N...N...P...L...V...S...V...H...V...F...D...L...G...R...Y...
L.edodes.XP_046085150.1  D...E...Y...N...F...F...Y...N...P...T...E...D...S...N...R...V...V...P...Y...T...D...V...S...N...K...Y...P...K...G...Y...N...N...P...L...V...T...H...V...F...D...I...D...K...Y...R...D...R...I...A...H...N...A...T...E...E...
P.ostreatus.KDQ26420.1  D...E...Y...N...F...F...Y...N...P...T...E...D...S...N...R...V...V...P...Y...T...D...V...S...N...K...Y...P...K...G...Y...N...N...P...L...V...T...H...V...F...D...I...D...K...Y...R...D...R...I...A...H...N...A...T...E...E...
P.eryngii.KAF9497814.1  D...D...I...R...L...N...R...Y......T...N...M...N...E...A...Y...L...S...D...T...K...I...K...Y...P...K...G...F...O...N...P...Q...F...D...L...F...L...V...N...L...
S.cerevisiae.STE13 Dipeptidyl.58  P...T...F...N...L...Q...H...E...E...F...I...G...S...V...S...K...Y...P...V...I...D...R...I...K...Y...P...K...G...E...D...N...P...L...A...S...I...F...S...Y...S...
P.pastoris.ANZ75207.1  P...T...F...N...L...Q...H...E...E...F...I...G...S...V...S...K...Y...P...V...I...D...R...I...K...Y...P...K...G...E...D...N...P...L...A...S...I...F...S...Y...S...

350          360          370          380          390
P.pastoris.ANZ77179.1  ...H...R...Q...H...I...C...R...L...P...A...D...D...F...F...Y...N......D...E...N...I...T...E...E...R...L...I...T...I...I...W...G...D...S...R...F...T...I...T...R...E...S...L...
Coprinopsis.cinerea.545529  ...L...N...E...H...E...G......S...G...G...D...G...T...E...V...L...T...D...W...P...N...R...Q...A...A...N...S...V...I...M...E...T...A...W...S...N...S...G...L...I...V...K...E...V...R...N...A...D...D...
L.edodes.GAW08536.1  ...L...D...S...D...S...V...V...N...G...F...P...A...N...A...T...L...E...L...D...W...P...G...R...H...P...I...S...N...I...M...E...V...A...W...D...D...T...Q...L...L...L...K...E...V...R...N...A...D...S...
L.edodes.XP_046085150.1  ...L...D...S...D...S...V...V...N...G...F...P...A...N...A...T...L...E...L...D...W...P...G...R...H...P...I...S...N...I...M...E...V...A...W...D...D...T...Q...L...L...L...K...E...V...R...N...A...D...S...
P.ostreatus.KDQ26420.1  E...G...W...D...D...Q...T...K...R...T...I...S...G...F...P...A...I...N...A...T...L...E...L...D...W...P...Y...R...Q...A...K...E...N...S...I...L...E...V...A...W...G...N...A...S...L...I...M...K...E...V...R...N...A...D...N...
P.eryngii.KAF9497814.1  E...G...W...D...D...Q...T...K...R...T...I...S...G...F...P...A...I...N...A...T...L...E...L...D...W...P...Y...R...Q...A...K...E...N...S...I...L...E...V...A...W...G...N...A...S...L...I...M...K...E...V...R...N...A...D...N...
S.cerevisiae.STE13 Dipeptidyl.58  ...Q...N...G...I...Y...S...I...N...T...G...G......O...K...D...S...I...L...Y...N...G...K...W...S...P...D...I...F...R...F...I...D...R...N...S...K...I...
P.pastoris.ANZ75207.1  ...A...E...E...L...E...K...L...N...I...G...A...A...A...S...L......E...E...D...F...V...L...Y...S...L...K...W...H...K...S...L...F...I...S...T...R...D...I...S...K...I...
```

# Chapter 2

```
400 410 420 430 440
P.pastoris.ANZ77179.1 LAHYIDRAN...SKIVRQDAKSTK...RFIEHNTF...IPKDAVGR
Coprinopsis.cinerea.545529 GHAIFFDSSME...ESVRRARACQVVRKLGKDGEGDDGWDDAQAIYVLYLDG...
L.edodes.GAW08536.1 GSVVFFDSSS...DDKARSRTVVRKLGKGEGERDDGWDDNNQNF...LP...KSMKGS
L.edodes.XP_046085150.1 GSVVFFDSSS...DDKARSRTVVRKLGKGEGERDDGWDDNNQNF...LP...KSMKGS
P.ostreatus.KDQ26420.1 GTVVLFDDSSGAGVSGDGRKVTGKIVRKLKGDGEGDDGWDDNNQNYIHP...IPFRSARRQ
P.eryngii.KAF9497814.1 GTVVLFDDSSGAGVSGDGRKVTGKIVRKLKGDGEGDDGWDDNNQNYIHP...IP...KKSARR
S.cerevisiae.STE13_Dipeptidyl.58 LDVVFYDPS...MLTVRNTNSNLFNG...WEKTKDLS...LFPKPELKR
P.pastoris.ANZ75207.1 MEVTLVDV...SASVVRKHDAAVFKG...WETGDFSVY...VIR

450 460 470 480 490
P.pastoris.ANZ77179.1 ...EQDGYHDTIDV.NGYNHLYVFSFDNPNKIILTSCEWVVD.SFSAFDFKRNVYV
Coprinopsis.cinerea.545529 ...RDAYDIDVPSKDGFNHIALFRPATSSEF.QWLTCDWVVTG.LIRAVDTKGLLVYV
L.edodes.GAW08536.1 ...GSAAVYDIDVPTPEGYNHIALYSPANSNTF.RRLTGGWVWVSS.GIKGVDAKGLLVYV
L.edodes.XP_046085150.1 ...GSAAVYDIDVPTPEGYNHIALYSPANSNTF.RRLTGGWVWVSS.GIKGVDAKGLLVYV
P.ostreatus.KDQ26420.1 ...QAERYDIDVPTPEGYNHIALYSPASSSTF.HRLTGGWVWVSS.GIKGVDAKGLLVYV
P.eryngii.KAF9497814.1 ...QAERYDIDVPTPEGYNHIALYSPASSSTF.HRLTGGWVWVSS.GIKGVDAKGLLVYV
S.cerevisiae.STE13_Dipeptidyl.58 ...MDYGVYDIDVHADSRRGFSRLLFYVVFVAKELIQLTGKMWVWVWVGNIVGYEYEDTIF
P.pastoris.ANZ75207.1 ...DITVGVYDIDVYVYEDYDRLVYVFDCTSDR...VVVLTGKMWVWVWVWVGNIVGYEYEDTIF

500 510 520 530
P.pastoris.ANZ77179.1 TATKK...SSTERHVVSVGT...GKLLNNTDV...SSDGY...YVSTS
Coprinopsis.cinerea.545529 TAAKSERQASIERHDFSVPLP.LDGGSTPV...ATDSTPLTLDI.TKPGY...YVSTS
L.edodes.GAW08536.1 EAAANP...SSIERHDFSVPLP.LDGGSTPV...ATDSTPLTLDI.TKPGY...YVSTS
L.edodes.XP_046085150.1 EAAANP...SSIERHDFSVPLP.LDGGSTPV...ATDSTPLTLDI.TKPGY...YVSTS
P.ostreatus.KDQ26420.1 VAAKA...SSTERHDFSVPLP.SLVRSPAESDAPTYRVEFTLTD...SERAY...YDIN
P.eryngii.KAF9497814.1 VAAKA...SSTERHDFSVPLP.SLVRSPAESDAPTYRVEFTLTD...SERAY...YDIN
S.cerevisiae.STE13_Dipeptidyl.58 TANE...GYMSQHYYSISIT...DSTQNTFQSLQNSD...VDFYDDE
P.pastoris.ANZ75207.1 TGRE...SMEHRYYSISIT...DSTQNTFQSLQNSD...VDFYDDE

540 550 560 570 580
P.pastoris.ANZ77179.1 FSPGARFVLLSHOGPHVYQ...KMLDVLK...GTBEIT...IBSNBDDKD.SVSLFN
Coprinopsis.cinerea.545529 FSPGARFVLLSHOGPHVYQ...KMLDVLK...GTBEIT...IBSNBDDKD.SVSLFN
L.edodes.GAW08536.1 FSPGARFVLLSHOGPHVYQ...KMLDVLK...GTBEIT...IBSNBDDKD.SVSLFN
L.edodes.XP_046085150.1 FSPGARFVLLSHOGPHVYQ...KMLDVLK...GTBEIT...IBSNBDDKD.SVSLFN
P.ostreatus.KDQ26420.1 FSPGARFVLLSHOGPHVYQ...KMLDVLK...GTBEIT...IBSNBDDKD.SVSLFN
P.eryngii.KAF9497814.1 FSPGARFVLLSHOGPHVYQ...KMLDVLK...GTBEIT...IBSNBDDKD.SVSLFN
S.cerevisiae.STE13_Dipeptidyl.58 LSSARVAISKKLQPDTPIKVAGPLRVLNVAAE...IHDSTLQDKRFEK.KTKNVD
P.pastoris.ANZ75207.1 I...KELKVALLSYRGP...KPYQ...KFDLSELATTNLDIT...LSNRRITVEVSLATIS

590 600 610 620 630
P.pastoris.ANZ77179.1 TPDVRYGEVLEHRCVKSNYVITPKNFHS...KYPVLFVYGGGFSQLVTKTFS...SFG
Coprinopsis.cinerea.545529 SBITVYSLIM.SDGYELNVEHTPPKFDGRTKYPVLFVYGGGFSQVDAARWPTANS
L.edodes.GAW08536.1 SBITVYSLIM.SDGYELNVEHTPPKFDGRTKYPVLFVYGGGFSQVDAARWPTANS
L.edodes.XP_046085150.1 SBITVYSLIM.SDGYELNVEHTPPKFDGRTKYPVLFVYGGGFSQVDAARWPTANS
P.ostreatus.KDQ26420.1 ABTILYSITIE.SDGFVYVREVPVPRMDDSGRTKYPVLFVYGGGFSQVDAARWPTANS
P.eryngii.KAF9497814.1 ABTILYSITIE.SDGFVYVREVPVPRMDDSGRTKYPVLFVYGGGFSQVDAARWPTANS
S.cerevisiae.STE13_Dipeptidyl.58 ABTILYSITIE.SDGFVYVREVPVPRMDDSGRTKYPVLFVYGGGFSQVDAARWPTANS
P.pastoris.ANZ75207.1 ABTILYSITIE.SDGFVYVREVPVPRMDDSGRTKYPVLFVYGGGFSQVDAARWPTANS

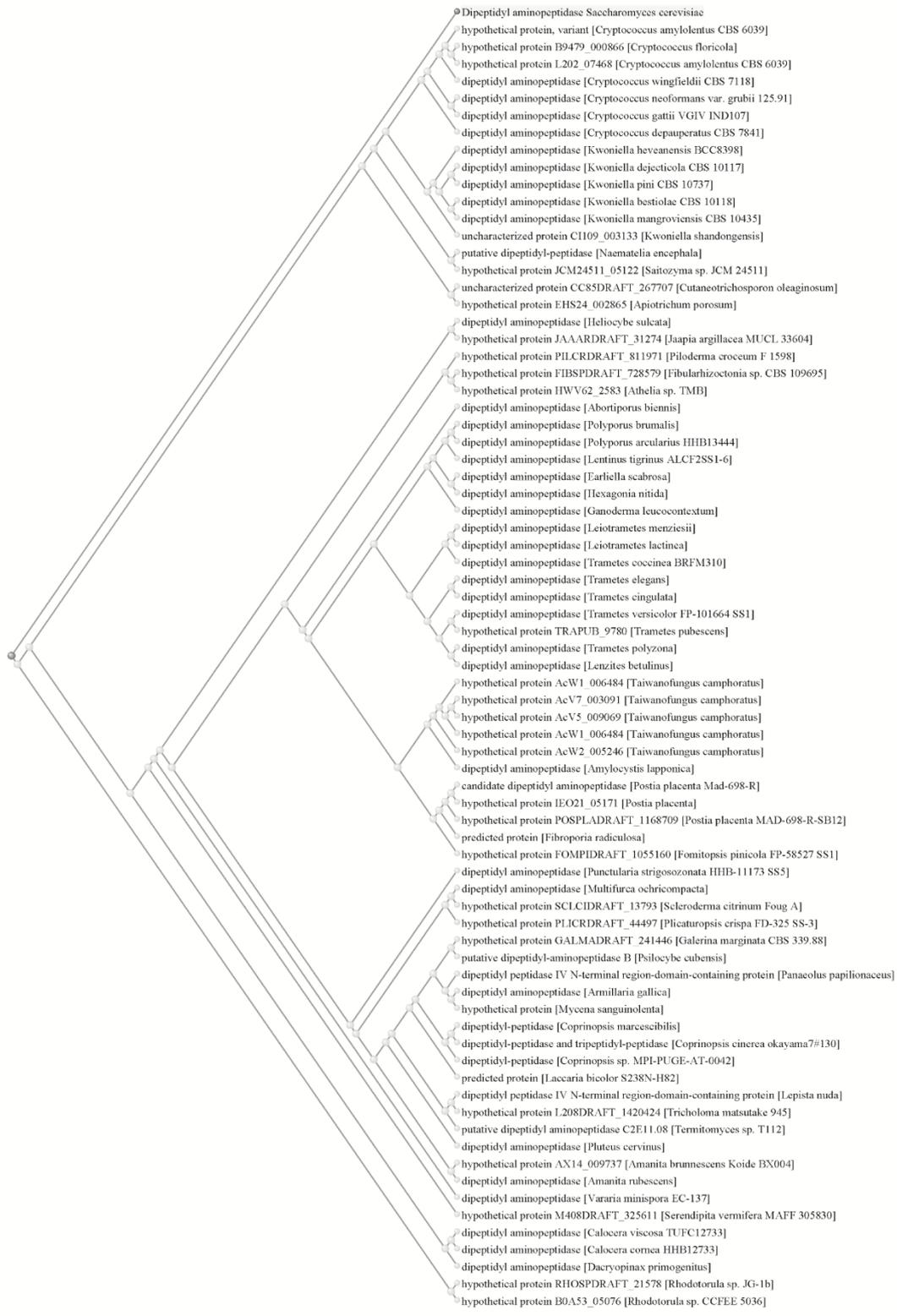
640 650 660 670 680 690
P.pastoris.ANZ77179.1 HVVASELDIVVTVDGRCTGFRGRDYRSLVRRNICHVPSLDQTKAGHIWAS...KPYVDEN
Coprinopsis.cinerea.545529 NYLAAAGLYIVSVVDRGCTGKGRKLRNPKNNLGFETIDQIEAARQYVR...KPYVDSK
L.edodes.GAW08536.1 DYLAAGLQFIVYVVDGRCTGKGRKLRNPKNNLGFETIDQIEAARQYVR...KPYVDSK
L.edodes.XP_046085150.1 DYLAAGLQFIVYVVDGRCTGKGRKLRNPKNNLGFETIDQIEAARQYVR...KPYVDSK
P.ostreatus.KDQ26420.1 DYLAAGLQFIVYVVDGRCTGKGRKLRNPKNNLGFETIDQIEAARQYVR...KPYVDSK
P.eryngii.KAF9497814.1 DYLAAGLQFIVYVVDGRCTGKGRKLRNPKNNLGFETIDQIEAARQYVR...KPYVDSK
S.cerevisiae.STE13_Dipeptidyl.58 QAVVSGLDIVVLDQTEPRGTGKGSWSPKSYATKTKGYWRDRLTAVSKWIDHVPVME
P.pastoris.ANZ75207.1 HTLSSLDIVVLDQTEPRGTGKGSWSPKSYATKTKGYWRDRLTAVSKWIDHVPVME

700 710 720 730 740 750
P.pastoris.ANZ77179.1 RLAIWGSYGGFMILKVLBQDKGEBTPKYGMSVAPVTNKEFYDSYTERYMHHPDNPNY
Coprinopsis.cinerea.545529 RLAIWGSYGGFMSSKVVEANEG.HITLAMVAPVTSWLDYDIYTERYMHHPDNPNY
L.edodes.GAW08536.1 RLAIWGSYGGFMSSKVVEANEG.HITLAMVAPVTSWLDYDIYTERYMHHPDNPNY
L.edodes.XP_046085150.1 RLAIWGSYGGFMSSKVVEANEG.HITLAMVAPVTSWLDYDIYTERYMHHPDNPNY
P.ostreatus.KDQ26420.1 RLAIWGSYGGFMSSKVVEANEG.HITLAMVAPVTSWLDYDIYTERYMHHPDNPNY
P.eryngii.KAF9497814.1 RLAIWGSYGGFMSSKVVEANEG.HITLAMVAPVTSWLDYDIYTERYMHHPDNPNY
S.cerevisiae.STE13_Dipeptidyl.58 RLAIWGSYGGFETLKVLEYSCEVFKYGMVAPVTNKEFYDSYTERYMHHPDNPNY
P.pastoris.ANZ75207.1 RLAIWGSYGGFETLKVLEYSCEVFKYGMVAPVTNKEFYDSYTERYMHHPDNPNY

760 770 780 790 800 810
P.pastoris.ANZ77179.1 YNSG.IHEIDNLDKGVKRFLLHGGDDNVHFQNTLKVLDLFDLGLGENDVYVHVPDSDS
Coprinopsis.cinerea.545529 YNAS.ISNVTGFHNVD.YLLHGGDDNVHFANSAHLDLDMFTQAKVRKFRFRMFDSDS
L.edodes.GAW08536.1 YNAS.ISNVTGFHNVD.YLLHGGDDNVHFANSAHLDLDMFTQAKVRKFRFRMFDSDS
L.edodes.XP_046085150.1 YNAS.ISNVTGFHNVD.YLLHGGDDNVHFANSAHLDLDMFTQAKVRKFRFRMFDSDS
P.ostreatus.KDQ26420.1 YNAS.ISNVTAFDKVD.YLLHGGDDNVHFANSAHLDLDMFTQAKVRKFRFRMFDSDS
P.eryngii.KAF9497814.1 YNAS.ISNVTAFDKVD.YLLHGGDDNVHFANSAHLDLDMFTQAKVRKFRFRMFDSDS
S.cerevisiae.STE13_Dipeptidyl.58 FEVSIQNFKAFESLKRFLVHGTEDDNVHINQNTFRLLDQLNLLGLINVDMHIFDSDS
P.pastoris.ANZ75207.1 SEYS.IKIVSFKKVNRFVLCGHTDDNVHFQNTLKVLDLFDLGLGENDVYVHVPDSDS

820 830
P.pastoris.ANZ77179.1 RRYHNSNVIYVQK...RRAFRAGR...
Coprinopsis.cinerea.545529 RRRGANREYVEYMLGFLEKWKGGRRRGW...
L.edodes.GAW08536.1 RRRGANREYVEYMLGFLEKWKGGRRRGW...
L.edodes.XP_046085150.1 RRRGANREYVEYMLGFLEKWKGGRRRGW...
P.ostreatus.KDQ26420.1 IRRGANREYVEYMLGFVVEKWKGGRRRGW...
P.eryngii.KAF9497814.1 IRRGANREYVEYMLGFVVEKWKGGRRRGW...
S.cerevisiae.STE13_Dipeptidyl.58 RRYHNAQRIVEQKLY...YWLRFDAERFDNTEVHL
P.pastoris.ANZ75207.1 RTHHNAQRIVEYER...KWLRFDAERFDNTEVHL
```

S12b)



**Figure S12:** Homologs of dipeptidyl aminopeptidases STE13 in different fungal species. BLASTs were carried out using NCBI (<https://www.ncbi.nlm.nih.gov>) or JGI (<https://mycocosm.jgi.doe.gov>). Shown are a) the alignments of *S. cerevisiae* STE13 with *P. pastoris*, *C. cinerea*, *L. edodes*, *P. ostreatus*, and *P. eryngii*, b) distance tree of BLAST alignments of STE13-homologs in basidiomycetes, visualized using NCBI.





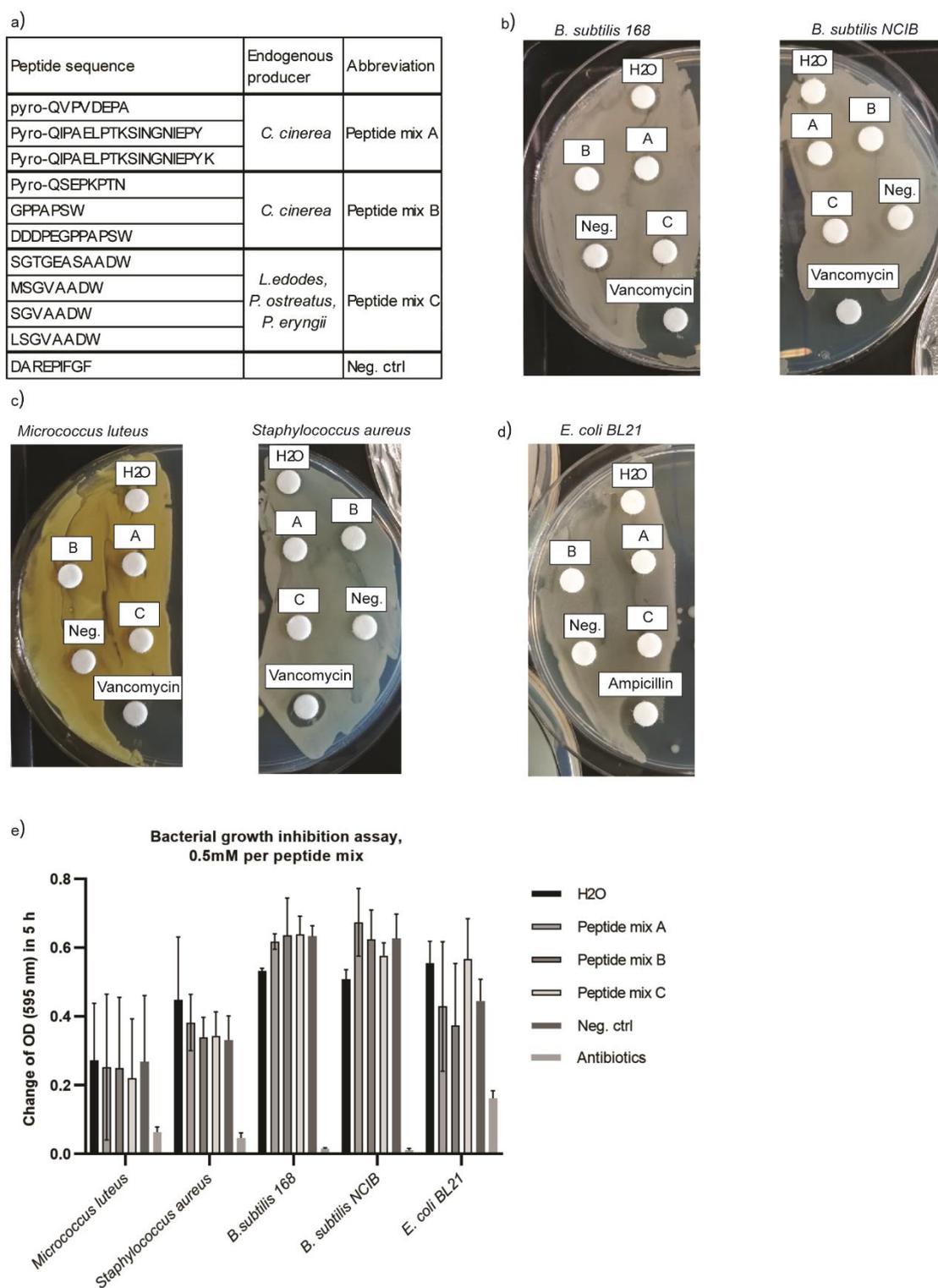
## Chapter 2

S13c)

# ID	Prediction	OTHER	SP(Sec/SPI)
N.crass.glutaminyI.cyclase.4896	NO_SP	0.627403	0.369004
N.crass.glutaminyI.cyclase.1284	SP	0.258902	0.738022
Homo.sapiens.Q16769.QC	SP	0.000294	0.999144
Homo.sapiens.Q9NXS2.isoQC	NO_SP	0.999948	0.000057
Lentinula.edodes.XP_046087353.1	SP	0.000234	0.999205
C.cinerea.360529	SP	0.263988	0.733891
Pleurotus.ostreatus.XP_036636044.1	SP	0.016276	0.982791
Pleurotus.eryngii.KAF9497455.1	SP	0.012853	0.986179

**Figure S13:** Alignments of glutaminyI cyclase (QC) homologs across different species. BLASTs were carried out using NCBI (<https://www.ncbi.nlm.nih.gov>) or JGI (<https://mycosm.jgi.doe.gov>). Shown are a) alignment of glutaminyI cyclases of *C. cinerea* and homologs from *N. crassa*, *Homo sapiens*, *L. edodes*, *P. ostreatus*, and *P. eryngii*, b) distance tree of BLAST alignments of *C. cinerea* glutaminyI cyclase against the homologs from *N. crassa*, *Homo sapiens*, *L. edodes*, *P. ostreatus* and *P. eryngii* and c) the individual signal peptide predictions of the glutaminyI cyclases (assessed using SignalP 4.0, Petersen et al. 2011).

## Chapter 2



**Figure S14:** Bacterial growth inhibition assays of synthetic KEP-derived peptides. Chemically synthesized peptides of each confirmed KEP-derived peptide from *C. cinerea*, *L. edodes*, *P. ostreatus*, and *P. eryngii* were tested for bacterial growth inhibition in disk diffusion assays and in optical density (OD) assays in liquid culture. a) List of each KEP-derived peptide confirmed in this study. The peptides were combined in three different peptide mixtures A, B, and C for the tests. b) Disk diffusion assays on *B. subtilis* 168 and *B. subtilis* NCBI 3610, b) *Micrococcus luteus* and *Staphylococcus aureus*, and c) *E. coli* BL21. 70 nmol of each individual peptide (corresponding to approx. 50  $\mu$ g) and 5  $\mu$ g of either vancomycin or ampicillin were added to the disks. e) Bacterial growth inhibition assay using optical density measurement. The same bacterial strains were cultivated in 96 well plates and the OD change

## Chapter 2

after 24 hours was plotted. Each peptide mix had a concentration of 0.5 mM. 50 µg/ml of ampicillin (for *E. coli* BL21) or vancomycin (for all other strains) were used as control. The bars represent the median value of three biological replicates, the error bars represent the 95% confidence interval.

### Supplemental Tables

ID	Complete sequence	Detected in			
		This study	Method	Le Marquer et al., 2019	Umemura, 2000
139620	MVKAIVVAAAALAAVVVLSIPINVEYDLEHRDVLDFERQYNGDEGSTVKAGEDKGGSAVKGGIVPSGKQRTGRRSFDDYDLFERQDPNSGSGEPDEGTSKGGAIKTNITPKGGRKGRRLSDYYDLFERQDPNSGSGEPDEGKKSAGGAIKTNITPKGGRKGRRSFDDYDLFERQDPNSGSGEPDEGTSKGGAIKTNITPKGGRK MRLSLSNFLSFFLLSSLAHAHAHDSIQRRSRSLKLRQRLSDIPVLDGVLGGDGLPFGFPGSGNDNDDDDDDQPDPAPTQEEVTLPPPMSTSWDPVVPVTPDESSDSVPEPSPTSPDPSIPESSEPVPEPTESVPEPSDPSEPVPEPSDPSEPVPEPSAPSSSEPEPAPTSAPSSSSTPSGLLPTLTSIVGDVTSGLGSIVTEVLIPTPSSSVVPEPSVPEPSQPEPSIPEPSLPLPSDPVPEPSDPVPEPSDPVPEPEPVPVPEPSDPLPSLPLPSVSLPLPSISLPLPSTTEDPGTDPSTSEEPGPGTSEEPGP	Yes	1	No	Yes
197636	SVSVSDPPISEPTPGPSGVSVDGSPSISDPGPSISDPPISEPTPGPSGVSVSISEPPISEPSGVSPTSLEPEVPEPSETLSAPIDTSVGDIPGLSFVLTESQLAINTLPSSTITPTISDPDQATATPEPIVTSQATQSFSAAPLPTGIPNKIYPNKRIGEDELDTGFTLISILFNRELNWPHVIRNENASSQIFAYTPMVIANTLEIGTSSIKTYALQVYIPTSYPGADQTELGTMYLAYIPELVDTLAAAIKTRNSKFTYATPGIAGDLAKYVNSGFALRSVQDPGVGNNGTNGGAPGNTGATSSDGKSRQDAIIGVVSALGAIALFVLIFLVYRSLKRRRELAHRRLSDPPNADAHGYRPEGREFDQDSIGGQRRRSFYFAEDSLRGYAGAAAQEQEYTYNHHQQSQQQMSQRRNINVNPNAISAPILRDNTMNW MVAFRSATNWVWTSCLALLWFYSQYLSQASILGERSLEPLSAEGLVLLKPKDPVKNIDPRDPSSHLSKILIPRAPGTQNNFVRNYLISTLKAIDWHVETDEFVADTPIGKKKMANVIATKDKPKASRRVLVLSAHFDSKYFPNYPENQFLGATDAAACAMMLDVAEALNPFLEKRMKAMKDADYEEDDIGDLTLQLVFFDGEEAFDWDTDTSIYGRHLAEKWASTYIQPHQKRRMLMNDATTEISTIEHFILLDLGAKQLIRSYFLDTAWLFDAMVAVEERLGMGAFAEYEGEEKSMAPGKWTSYFMKRRKGMTNMGYIGDDHVPFLQRGVSVLHIITEPFPRVWHTIKDDASALDIPTMRRWNLILRVFLAEYFSLDSHRPNARSLEDEDESDFISRSVSELVSAARVSIEQGSNRPLYS	Yes	1	No	No
360529	MRFSNSIALLVFAFLSTSLASPVGEALVNRGALDDLERLPARELADRGITPPPWRDHVEPTPVGITPPPWRDNLPAPTPVPPVRRRDRSPPPAWRPIEADSRPPPAWRPVKRAVDSRPPPAWRPISDDDEGPPAPSRRETADAAGVVGPDW MIYPASTLILATLVFNAASVLGAPVRIPASPPVARALADDPVLPKVEIVDVPARMRLDDDDPMDARSAPRIDIRDTASSEIEARDPLPEVEEIVRRYPRKVVYHDFYQKRSQPSAPPEKRAPEPDTPIQERRFPRRSLYERYQKRSEASPPVEARAQAAEPELQQRRYPRKVVYADYIEKRSTPPSPPEQPVAREPAPEPETVQARRFPREVYMKHYERQADPVKHYGRQAETPESAPADATTPDAATPPTDAAAAPQPSGAAVSGGAAVQPPQASTPPVPLNSTSTPLDNALFHNITSNDAGFQPGTTIQETTIVL	Yes	1	No	No
377821	VHKITHISEPQQKVDTPPVVPPSPPTGSDPTAPSPSDPASSASPSASGGVAADPSAPPVAGGEQAPPQPSPSDAATSPPTDSSTTPPGAPTTEDGASEPKPEETNSGSESTEGGAAAPPPGEDGAASRRSIGSDASPEPILRRRNLGYSGPAWASYAKRSVTN	Yes	1	No	No
392677	MRISVVSALTLLAATSLTFAAPTTFGDLVNRSERQTGVVKGPERRNIIGVDGDLVQRSERQTGVVVKPRNEPPADRETGTVKWKRQNSDRQTGTVKWKRQDDDESDRQTGTVKW MQLRFSFFALALLALNAVATPISYADSEDLERIGKKNPLPAVNQKVAKLPPDDLRRKVFGHWKKQTLQRANARVGHKAIMQELKGQAPPGRWTDRTTAAWAVHQKKNLRRRELLDSELEVRGRKNPLPPVNQKVAKLPEELRRKVFGHWKKQTLQRANARAGHKAIVQELKGQAPPGRWTDRTTAAWAVHEKKNLRRLELLDEVEELL MVQISTSVLVAALAIAPVLAIVIAQEADEFVARVDVLDVAEPLVAREPIFGFIKRIFTGKRDFSESEELALRELAEDLDAREPIFGFIKRIFTGKRDLSESEELALRELVEDIDAREPIFGFIKRIFTGKRDLSETEELALREYVDSLAREPIFGFIKRIFTGKRDLSDLLDL	Yes	1	No	No
405832	SLRDFEDLEAREPIFGFIKRIFTGKRDLSESEELALREYVDSLAREPIFGFIKRIFTGKRDFSDVDDLRLRDFEELDAREPIFGFIKRIFTGKRDLSETEELALRELVDLDAREPIFGFIKRIFTGKRDLSESEELVREFMDSLARDPSPFAGIAGKGGKALNWLGTAGTLASIPAMFRSSKKKDRDFEDDLVFRDWVLEELDARDFDDDLAYRMFDFEFDARELNELD	Yes	1	No	Yes
426342	MVNCYTLAFAASSLLSSVGPSLVVQAAPVWGSSDFDILLGRATVANGAANVIPSKRSYDLDQRAPGHKCPGRGGRWKS	Yes	1	No	Yes
434504	TLDPGRENLERGIADTRRALGSASDLLELRFIANLEERNLELEEIGKERNLELEGRAGRCTVCGRWRGQVRGTVHVCPGSTNDPKPEPGVKRCSVCNQWLGTSGGVHTCPGGGGRRKKN	Yes	1	No	No
440153		Yes	1	No	No









## Chapter 2

<i>B. subtilis</i> NCBI 3610	“	SMA #1279	Richard Losick, University of Harvard
<i>Micrococcus luteus</i>	“	Practical course database #72	DSMZ 20030
<i>Staphylococcus aureus</i> 113	“	SMA #1253	DSMZ 4910
<i>Coprinopsis cinerea</i> AmutBmut	A43mut B43mut <i>pab 1.2*</i>	CMK #1	U. Kües, Georg-August-Universität Göttingen
<i>C. cinerea</i> AmutBmut <i>pab1-2 Δku70</i>	A43mut B43mut <i>pab 1.2 Δku70::Flt<sup>R</sup>; Δku70</i> strain used for mutagenesis of <i>kex</i> genes	CMK #100	(Nakazawa and Honda 2015)
<i>C. cinerea</i> AmutBmut <i>Δku70 Δkep 405832 #1</i>	Knockout strain of KEP 405832 based on the parent strain <i>C. cinerea</i> AmutBmut <i>pab1.2 Δku70</i>	SMA #3212	This study
<i>C. cinerea</i> AmutBmut <i>Δku70 Δkep 405832 #2</i>	Knockout strain of KEP 405832 based on the parent strain <i>C. cinerea</i> AmutBmut <i>pab1.2 Δku70</i>	SMA #3213	This study
<i>C. cinerea</i> AmutBmut <i>Δku70 Δkep 434504#1</i>	Knockout strain of KEP 434504 based on the parent strain <i>C. cinerea</i> AmutBmut <i>pab1.2 Δku70</i>	SMA #3214	This study
<i>C. cinerea</i> AmutBmut <i>Δku70 Δkep 434504#2</i>	Knockout strain of KEP 434504 based on the parent strain <i>C. cinerea</i> AmutBmut <i>pab1.2 Δku70</i>	SMA #3215	This study
<i>C. cinerea</i> AmutBmut <i>Δku70 Δkep 490115#1</i>	Knockout strain of KEP 490115 based on the parent strain <i>C. cinerea</i> AmutBmut <i>pab1.2 Δku70</i>	SMA #3216	This study
<i>C. cinerea</i> AmutBmut <i>Δku70 Δkep 490115#2</i>	Knockout strain of KEP 490115 based on the parent strain <i>C. cinerea</i> AmutBmut <i>pab1.2 Δku70</i>	SMA #3217	This study
<i>C. cinerea</i> AmutBmut <i>Δku70 Δkep 497993#1</i>	Knockout strain of KEP 497993 based on the parent strain <i>C. cinerea</i> AmutBmut <i>pab1.2 Δku70</i>	SMA #3218	This study
<i>C. cinerea</i> AmutBmut <i>Δku70 Δkep 497993#2</i>	Knockout strain of KEP 497993 based on the parent strain <i>C. cinerea</i> AmutBmut <i>pab1.2 Δku70</i>	SMA #3219	This study
<i>C. cinerea</i> AmutBmut <i>Δku70 Δkep 503649</i>	Knockout strain of KEP 503649 based on the parent strain <i>C. cinerea</i> AmutBmut <i>pab1.2 Δku70</i>	SMA #3220	This study
<i>C. cinerea</i> AmutBmut <i>Δku70 Δkep 426342#1</i>	Knockout strain of KEP 426342 based on the parent strain <i>C. cinerea</i> AmutBmut <i>pab1.2 Δku70</i>	SMA #3221	This study
<i>C. cinerea</i> AmutBmut <i>Δku70 Δkep 426342#2</i>	Knockout strain of KEP 426342 based on the parent strain <i>C. cinerea</i> AmutBmut <i>pab1.2 Δku70</i>	SMA #3222	This study
<i>C. cinerea</i> AmutBmut <i>Δku70 Δkex1 (437675)</i>	Knockout strain of Kex protease based on the parent strain <i>C. cinerea</i> AmutBmut <i>pab1.2 Δku70</i>	SMA # 3317	This study
<i>C. cinerea</i> AmutBmut <i>Δku70 Δkex2a (502579)</i>	Knockout strain of Kex protease based on the parent strain <i>C. cinerea</i> AmutBmut <i>pab1.2 Δku70</i>	SMA #3095	This study (L. Sonderegger)
<i>C. cinerea</i> AmutBmut <i>Δku70 Δkex2b (406374)</i>	Knockout strain of Kex protease based on the parent strain <i>C. cinerea</i> AmutBmut <i>pab1.2 Δku70</i>	SMA #3211	This study (Y. Chen)
<i>C. cinerea</i> AmutBmut <i>Δku70 Δkex2c (448165)</i>	Knockout strain of Kex protease based on the parent strain <i>C. cinerea</i> AmutBmut <i>pab1.2 Δku70</i>	SMA #3318	This study
<i>C. cinerea</i> AmutBmut <i>Δku70 Δkex2a/c (502579, 448165)</i>	Double knockout strains of Kex proteases based on the parent strain <i>C. cinerea</i> AmutBmut <i>pab1.2 Δku70</i>	SMA #3319	This study

**Table S3:** Strains used in this study. The *C. cinerea* strain AmutBmut *pab 1.2* is labeled according to Wälti et al. 2006. The same strain is labeled as AmutBmut *pab 1-1* in other publications.

PCR/Sequencing primer	Sequence (5'-3')	Description
Primer_f_KEP1(405832)	<u>GTATACGCT</u> ATGCGCATCTC	Amplification of the KEP genes (with a small overhang into UTR (underlined) from the <i>C. cinerea</i> cDNA for heterologous expression in <i>P. pastoris</i> )
Primer_r_KEP1(405832)	CCGTCATTATGCATTACC	“
Primer_f_KEP2(434504)	<u>TCAACATGGTTCAAATCTCC</u>	“
Primer_r_KEP2(434504)	CGTGCACTAGTCGAGCT	“
Primer_f_KEP4(497993)	ATGTTCTCTCGCACTATCG	“
Primer_r_KEP4(497993)	TCACAAAGGTACGATTTACC	“
Primer_f_KEP5(503649)	ATGATGCTCCGCACCAAC	“

## Chapter 2

Primer_r_KEP5(503649)	TCAAGCCTCAATTATATCCTCC	"
Primer_f_KEP6(426342)	TCAAGATGCAACTCCGTTTC	"
Primer_r_KEP6(426342)	CTAITTACAACAACCTCTCAACC	"
F_EcoRI_ApaI_KEP1(405832)	TAAGCAGAATTTCATGCGCATCTCCGTTGTTTCT	Amplification of the KEPs from the pGEM-Teasy vectors for later restriction digest and insertion into pPICZA (EcoRI and ApaI restriction sites), with 5' overhang
R_EcoRI_ApaI_KEP1(405832)	TGCTTAGGGCCCACTTCACAGTCCCAGTCTG	"
F_EcoRI_ApaI_KEP2(434504)	TAAGCAGAATTTCATGGTTCAAATCTCCACCTCC	"
R_EcoRI_ApaI_KEP2(434504)	TGCTTAGGGCCCGTCGAGCTCGTTGAGTTC	"
F_EcoRI_ApaI_KEP4(497993)	TAAGCAGAATTTCATGTTCTCTCGCACTATC	"
R_EcoRI_ApaI_KEP4(497993)	TGCTTAGGGCCCAAAGGTACGATTTACC	"
F_EcoRI_ApaI_KEP5(503649)	TAAGCAGAATTTCATGATGCTCCGCACCAACTTC	"
R_EcoRI_ApaI_KEP5(503649)	TGCTTAGGGCCAGCCTCAATTATATCCTCCCAA	"
F_EcoRI_ApaI_KEP6(426342)	TAAGCAGAATTTCATGCAACTCCGTTTCTCATTT	"
R_EcoRI_ApaI_KEP6(426342)	TGCTTAGGGCCCAACAACCTCTCAACCTC	"
AOX1_f	GACTGGTTCCAATTGACA	Colony PCR to check for integration of genes into Pichia genome
AOX1_r	GCAAATGGCATTCTGACA	"
T7	5' TAATACGACTCACTATAGGG	Sequencing primers to confirm correct pGEM-Teasy plasmid construction
M13r	5' CAGGAAACAGCTATGAC	"
Pcpab1_seq_r	5' TCGGCATAGCGCAGCTAG	Sequencing primers to confirm correct gene mutation
Pcpab1_seq_f	5' TGAGTGGCGAAACAACCTGTG	"
Snf5_seq_f	5' CTGGTTCCTCGACATCTGGC	"
Snf5_seq_r	5' CTCTCCACCAAGACCCAGATC	"
KEP1(405832)_seq_f	TTC AAC CAT CAG CTT CAT C	"
KEP1(405832)_seq_r	GTT ACG GAA TGA ACA CGA G	"
KEP2(434504)_seq_f	GAC ATT TAT TCC TGC TGT TCC	"
KEP2(434504)_seq_r	CAA CGG ATA AGG CAA AGT G	"
KEP3(490115)_seq_f	GTA TCC AAG GTG AAG GTG C	"
KEP3(490115)_seq_r	TAC CGA TAA AGG AGA CAA GGG	"
KEP4(497993)_seq_f	TCA TCT TGT TCT GGG TTT GC	"
KEP4(497993)_seq_r	GAC CTC TGT TTT GTC GAC C	"
KEP5(503649)_seq_f	ATT CAA TTC CGG TCC GAC	"
KEP5(503649)_seq_r	GTT TAC GGT GAC TCC TGG	"
KEP6(426342)_seq_f	CCA GGA GTC GAA GCT ATT ATC	"
KEP6(426342)_seq_r	GAA ATG CGG AGG GTG AAT C	"
KEX1_gene_r	AGTGTCTTGGCCGTTCTCTT	Sequencing primers to confirm correct gene mutation, binding inside gene
KEX2c_gene_r	GACAGGCAGAGGAGCAGAAA	"
KEX2a_gene_r	CGATGAGCTCGAGAGGATG	"
KEX2b_gene_r	TCAGTGACCCGGTATGCAC	"
Cop6_seq_f	5' TTGCGCACCATAGGTCTTG	"
Cop6_seq_r	5' TCCACTAACAGGGTTTCTCC	"
KEX2a_seq_f	5' GGAGTTGCTGGTTGTTTTCAG	"
KEX2a_seq_r	5' GCACTTTCAATCATCGTAGCC	"
Kex1_probe_f	GAAGATAATCGCCAACGCC	"
Kex1_probe_r	GATTGTTGAAATCCTAGAGTC	"
KEX2a_probe_f	5' TAATTGTCGGCTTCATCGTCC	Primers for amplification of labeled DNA probe for southern blots
KEX2a_probe_r	5' TACTGGTAACTGCACATGTCG	"
KEX2b_probe_f	GTCACACTGAGACCTTTGAACGC	"
KEX2b_probe_r	5' GTTCGAGCTTGCACACGTC	"
KEX2c_probe_f	CGACGTTTATAGCAGGGTTGG	"

## Chapter 2

KEK2c_probe_r	GATCCCGTCACAGCCACTCA	“
KEP1(405832)_probe_f	CAAACGAAGGTGAGCAGG	“
KEP1(405832)_probe_r	TTTGGCTTTGGATGTGCG	“
KEP2(434504)_probe_f	AAGTCCAGGTTTCTTCCC	“
KEP2(434504)_probe_r	TTTGGCTTTGGATGTGCG	“
KEP3(490115)_probe_f	TTGTCTCGCACATGATGC	“
KEP3(490115)_probe_r	GGTAAGCGACAGATTTTGG	“
KEP4(497993)_probe_f	AATATCGTCGCCTCCAAG	“
KEP4(497993)_probe_r	CACAACATGATCGCCAGC	“
KEP5(503649)_probe_f	CATGGAGGCGTAATTTCTG	“
KEP5(503649)_probe_r	CATATGCTTGGAGGGTGG	“
KEP6(426342)_probe_f	GTTCTCTTTGATGCTGTG	“
KEP6(426342)_probe_r	AGCTAGGGAATGACAAC	“
KEP1(405832)_A_f	GCCAATGCCCTAAGTGGACTTGCCTG	Primers for cloning of repair templates for knockout establishment
KEP1(405832)_A_r	ggacctctgaattcttcttcAGCGTATACTCAGATCC	“
KEP1(405832)_B_f	GGATCTGAGTATACGCTgaagaagaattcagaggtcc	“
KEP1(405832)_B_r	GTCCGTCCTATTATGCAcagtcacaatgacagc	“
KEP1(405832)_C_f	gctgtcattgtggactgTGCATAATGGACGGAC	“
KEP1(405832)_C_r	GCTCCTTCACTATGTGAGCCGTATTTACTTAGTCC	“
KEP2(434504)_A_f	TGATGCAGCGATAAGTCAACACCC	“
KEP2(434504)_A_r	gcgacctctgaattcttcttcGTTGATAGATAGAAATTG	“
KEP2(434504)_B_f	CAATTTCTATCTATCAACgaagaagaattcagaggtccgc	“
KEP2(434504)_B_r	CTGCATATATCCCGTGCATcagtcacaatgacag	“
KEP2(434504)_C_f	ctgtcattgtggactgATGCACGGGATATATGCAG	“
KEP2(434504)_C_r	GTTGCTGAGCGAGGATCAGTACGG	“
KEP3(490115)_A_f	CACATAAATCGTTCGTGCTTGTCAAGGTCC	“
KEP3(490115)_A_r	cctctgaattcttcttcCGTGAGGACAACAGTGAG	“
KEP3(490115)_B_f	CTCACTGTTGTCTCACGgaagaagaattcagagg	“
KEP3(490115)_B_r	GTTGGCGGTGGAAGGCGAcagtcacaatgacagc	“
KEP3(490115)_C_f	gctgtcattgtggactgTCGCCTTCGACCCCAAC	“
KEP3(490115)_C_r	CATGAATAGACTATGTGGAAGAGAATTCGAAGTC GC	“
KEP4(497993)_A_f	CCTTCATCGTTGTCGCTGTCATCGCC	“
KEP4(497993)_A_r	gacctctgaattcttcttcCGTGGAGGAGACGGG	“
KEP4(497993)_B_f	CCCGTCCTCCTCACGgaagaagaattcagaggtc	“
KEP4(497993)_B_r	GAAAGCTCTCAAGACTAacagtcacaatgacagctctc	“
KEP4(497993)_C_f	gagagctgtcattgtggactgTTAGTCTTGAGAGCTTTC	“
KEP4(497993)_C_r	GCTGAGTCCTCAGGTAGCGTTCGTAGG	“
KEP5(503649)_A_f	AGCCATGTGGTCAGTGTCCCG	“
KEP5(503649)_A_r	gacctctgaattcttcttcCTTGTGAGCAAAAAGT	“
KEP5(503649)_B_f	ACTTTTTGCTCAACAAGgaagaagaattcagaggtc	“
KEP5(503649)_B_r	GTAGTCAAAAAGTCACGCcagtcacaatgacagctc	“
KEP5(503649)_C_f	gagctgtcattgtggactgGCGTGACTTTTACTAC	“
KEP5(503649)_C_r	GAACAGTTGCGGTGAAACCAGAACCC	“
KEP6(426342)_A_f	GATGTACGTGCACTTCGGATGGGTG	“
KEP6(426342)_A_r	GCTTCACCTAAAGAATCAAGgaagaagaattcagaggtc c	“
KEP6(426342)_B_f	GCTTCACCTAAAGAATCAAGgaagaagaattcagaggtc c	“
KEP6(426342)_B_r	GGTCATCATGATCATCTATcagtcacaatgacagctctc	“
KEP6(426342)_C_f	gagagctgtcattgtggactgATAGATGATCATGATGACC	“
KEP6(426342)_C_r	CAAGTCCCCTGGTTCGTAGGGAACG	“
Kex1_A_f	ATGAGGATTTCTCCATGACATTGACGGTATCGTT G	“
Kex1_A_r	GGACCTCTGAATCTTCTTCTCGGAAGCAGATTAC ATTAG	“

## Chapter 2

Kex1_B_f	CTAATGTAATCTGCTCCGAAGAAGAAGAAATTCAG AGGTCC	“
Kex1_B_r	GGCGATTATCTTCAGGTTATCAGTCCACAATGACA GCTC	“
Kex1_C_f	GAGCTGTCATTGTGGACTGATAACCTGAAGATAAT CGCC	“
Kex1_C_r	GTCCGGTGGATCGGTTACTAAACATGATTGTTGG	“
KEX2a_A_f	TTGAACCGTCCTCGAAGGTAAGTGGCTCGCCT	“
KEX2a_A_r	ggacctctgaattcttctcTGTGATGGGGTATGGCGGGT	“
KEX2a_B_f	ACCCGCCATACCCATCACAgagaagaattcagaggtc c	“
KEX2a_B_r	CCCCAAACCAAGAACCAATcagtcacaatgacagctct	“
KEX2a_C_f	agagctgtcattgtggactgATTGGTTCTTGGGTTGGGG	“
KEX2a_C_r	ACCGCTCGTCCCTCGCAATAGCAGCGGAGATAAA	“
KEX2b_A_f	GGACTTGGAGTGTGCGTCTCCG	“
KEX2b_A_r	CCTCTGAATTCTTCTCATGGACAAGAAGTGGTAG AA	“
KEX2b_B_f	TTCTACCACTTCTTGTCCATGAAGAAGAATTCAGA GGTCC	“
KEX2b_B_r	TCTCAGTGTGACTTACGCAGTCCACAATGACAGC	“
KEX2b_C_f	CTGTCAATTGGACTGCGTAAGTCACACTGAGA	“
KEX2b_C_r	AGGGATTCCGGTGTGGAGTATATGCCTC	“
KEX2c_A_f	CGGTCACTCGCGGAGAATAAGGTGAT	“
KEX2c_A_r	CCTCTGAATTCTTCTCAAATCGTCAATCGCCACGA T	“
KEX2c_B_f	GTGGCGATTGACGATTTGAAGAAGAATTCAGAGG	“
KEX2c_B_r	GCTTAGCTCTTGAGCCACAGTCCACAATGACAG	“
KEX2c_C_f	CTGTCAATTGGACTGTGGCTCAAGAGCTAAGC	“
KEX2c_C_r	TGCGGTAAGGGTGGGGAAGTGT	“

**Table S4:** Primers used in this study.

<i>kep1</i> (405832)	ATGCGCATCTCCGTTGTTCTGCCTGACCCTCTTGCGGCGACTAGTCTCACTTTCGCAGCTCCTACTA CCTTTGGAGACCTCGTGAACAGGTCCGAGCGCCAACTGGTGTGCTCAAAGGTCCAGAAAGGAGGAA TATCATCGGGTTGATGGAGATCTAGTTACGCGGTGACAACGGCAGACGGGCGTAAAGCCACGA AATGAGCCACCGGCCGACCCGAAACCGGTACAGTGAAGTGAAGCGACAGAATTCGGATCGCCAAA CCGGGACTGTAAAGTGAAGAGGCAGGACGACGATGAATCAGATCGCCAGACTGGGACTGTGAAGT GGTAA	Coding sequence for heterologous expression in <i>P. pastoris</i> , stop codon in red
<i>kep2</i> (434504)	ATGGTTCAAATCTCCACCTCCATCCTCGTCGCGCCCTTGCCATCGCGCCGTCCTCGCGTCCCATTTG CCCAGGAGGCCAACGAGTTCGTTGCCGGGATGTTGATCTCGTAGCTGAGCCACTGTGCTCGAGAG CCCATCTTGGTTTCATCAAGAGGATTTTACCGGGAAGCGAGACTTCTGAGTCTGAAGAGCTCGCT CTCCGGGAGTTGGCCGAAGACCTTGACGCCCGTGAACCCATCTTCGGTTTCATCAAACGAATCTTACC GGCAAGCGGGACTTGTCCGAGTCTGAGGAGCTTGCCCTCCGTGAGCTTGTGGAAGACATCGACGCCG TGAGCCCATCTTCGGTTTCATCAAACGAATTTTCACTGGCAAGCGTGATCTTTCGGAGACCGAGGAGCT CGCGCTCCGCGAGTACGTCGACTCGCTTACGCCCGTGAACCTATTTTCGGTTCATCAAACGATCTT CACCGGCAAGCGCGACCTTTCGACCTCGACGACCTCTCCCTCCGCGACTTTGAGGACCTTGAAGCTCG CGAGCCCATCTTCGATTTCATCAAGCGAATATTTCACTGGCAAGCGTGATCTTTCGAATCCGAGGAACT CGCCCTCCGCGAGTATGTTGACTCCCTCGACGCCCGTGAACCCATCTTGGCTTCATCAAGAGAATCTTC ACCGGCAAGCGAGACTTCTCCGACGTCGACGACCTTCCCTTCGTGACTTGAAGAGCTCGACGCCGA GAGCCCATCTTCGGTTTCATCAAGAGAATTTCACTGGCAAGCGTGATCTCTGAGACTGAGGAGCTC GCTCTCGTGAAGCTCGTGCAGCAGCTCGATGCCCGTGAACCCATCTTCGGCTTCATCAAGAGGATCTTC ACCGGCAAGCGAGACCTCTCTGAATCCGAAGAGCTCGTCTTCGCGAGTTCATGGACTCTCTCGAGCT CGCGATCTTCTTGTGGCATTGCCAAGATCGGTGCAAGGCCTTGAAGTGGCTCGGTACCGCTGG AACTCTCGTTCAATCCCTGCTATGTTTCAGGAGCAGCAAGAAGGACAAGCGCGACTTCGAGGACGACC TCGTCTTCGGTACTGGTCTCGAGGAGTTGGATGCCCGTACTTTCGAGTACCTCGCCTACCGCA TGTTTCGACGATGAGTTTCGACGCTCGTGAACCAACGAGCTCGACTAG	“

## Chapter 2

<p><i>kep3</i> (490115)</p>	<p>ATGCGGATTGCAAGTGCAGCCCTATTCTTCGTTTCGGTACTGAGTGTGTTAGCGCCTTGCTACACCT GAGGTTCGATCCCGTCGCCCTGTGGCAGGAGTCTGTTGCACGTGAACCTCAAATGAGGCCCTTCCTG GAGAAGAGACCCGGAACCTCAAAGCAGGCCCTTCCTGGCGACGGGAGCCCGAACCCGAGCCCGAA CCTCAATACAAGACCCCTCATGGAGACGCGAACCCGAGCCTCAAAGCAGGCCTCCTGCCTGGCGCCG CGAGCCTGAGCCCGAACCGCAATACAAGGCTCCTTCCTGGCGCCGCGAACCTGAGCCCGAACCGCAAT ACAAGGCTCCTTCCTGGCGCCGCGAACCTGAGCCCGAACCCAGTACAAAGCTCCTTCCTGGCGACGC GATCCGGAACCTCAGTACAAGGCCCTTCCTGGAGGCGCGAACCCGAACCCAGATGAGGCTCCCTC CTGGTGA</p>	<p>“</p>
<p><i>kep4</i> (497993)</p>	<p>ATGTTCTCTCGACTATCGTCTTCTCACCTCGCCCTCTCCCTGACTGGTCAGGGCTAGCCGTCCAG TTGCCTCGAACATTGATGCCCTCAGATCCAGCTGAGCTGCCACCAAGTCCATCAACGGCAACATTG AACCATTCAGGTTCCGCTCAGGTTCCCGTGGACGAGCCTGCTGCTCCCAATGGCATCGTGGTGC CGTTCAACAAGCGTCAAATCCCTGCAGAGCTCCCAACCAAGTCGATCAACGGGAATGTCGAACCTAC AAGGTTCCGCTCAAGTCCCTGTCGATGAGCCTGCTGCCCTCCCAATGGAATCGTAGTGCCTTCAAC AAGCGCAAATTCCTGCTGAGCTCCCGACCAAGTCCATTAACGGCAACATCGAGCCATACAAAGTTCGC CGTCAAGTCCCTGTTGACGAGCCCGCTGCGGCACCCAATGGGATTGTGGTACCGTTCAACAAGCGTCA AATCCCGCTGAACTGCCAACCAAGTCGATCAACGGCAACATTGAGCCTTACAAGTTCGCCGCCAGG TCCCGTGGACGAGCCTGCTGCCGCCCAATGGCATTGTAGTGCCTTCAACAAGCGCAAATCCCG CCGAACTGCCGACCAAGTCTATCAATGGCAACATCGAACCTACAAGACCAAGCGGGGAGTGCCTAC GAACTATCAGGTCCATTGGAGGCGATGAGATCATCGTCAACTACGCGACGTCCCAAGCGAGATTGC CGTCAGGTCCGTTGACGGTCAAATCGTACCTTTGTA</p>	<p>“</p>
<p><i>kep5</i> (503649)</p>	<p>ATGATGCTCCGACCAACTTCGTTCTCCTCGCCGTCGCCCTTCTCCACCCTCGGCTCGGTCTTCGCTG CGCCTACTCGCACTTCTCTGATGACTTTGACGCACGAGACCTCGCTTTTGAGGACATCCTCGACGCAC GGGGTGGTCTCGCCATTCTAAACCCACCCATGGAACCGACGGCGGAAAAGCCTCGTTGAACACCATT GGAAAGGGCAGGCGCGCTTCTTGAAGACATCCTGGACATCCGTGGCGCCTCGGAATCCCAAGCC TACTACGGCACCGACGGTGGCAAACCTCACTCAACACCATCGGCAAGGGAAGCGTGCCTCTTGG AGGACTTCTTGGAGCCAGGGGTGACTCGCGCTTCTAAGCCTACGCACGGCACCGACGGCGGCAA GGCATCGCTGAACACCATCGGCAAGGGCAGGCGGAGTTCTTGGAGGATCTCCTCGATGCTCGAGAC CTGTTGGACGAGATTCTCGATGTGCGGGGTGGTTTGGCTCTCCAAACCTACACATGGCACGGATGG TGGAAGGCTTCCCTCAACACTATTGGCAAAGGCAGGAGGATTTTTTTCGAGGATATCCTTGATGCCA GGGGCGGGCTCGCGCTCCCAAGCCTACCCATGGAAGTACGAGTGGCAAGGCTTCGATGAATACCATT GGAAAGGGCAGGAGGATTTTTGGAGGATATAATTGAGGCTGA</p>	<p>“</p>
<p><i>kep6</i> (426342)</p>	<p>ATGCAACTCCGTTTCTCATTTTTGCTCTCGCCTTGTGGCCCTCAACGACGTGCCACGCCATCTCCTA CGCCGACTCGGAGGACCTCGAAATCCGAGGCAAGAAGCCCAACCCCTCCAGCCGTCAACCAAAGG TCGCCAAACTGCCGACGATCTCCGCGCAAAGTCTTCGGTCACTGGAAGAAGCAGACTTTCGAACGT GCCAACGCTCGGGTTGGACACAAGGCGATTATGCAGGAGTTGAAGGGACAGGCTCTCCCGCCGCT GGACTGATCGCACTACAGCTGCTGGGCTGTACACCAGAAGAACAAGAATCTGAGGAGGGAGTTGCT GGACTCGGACGAGCTCGAAGTCCGAGGTAGGAAGCCCAACCCGCTGCCGCTGTCAACCAAAGGTC GCCAAATGCTGAAGAAGTCCGCCGCAAAGTCTTCGGTCACTGGAAGAAGCAGACTTTGCAGCGTGC AAACGCTCGGGCTGGACACAAGGCGATCGTGCAGGAATTGAAGGGACAGGCTCTCCCGCCGCTGG ACAGATCGAACTACCGCAGCTTGGGCTGTGCATGAGAAGAACAAGAATTTGAGGAGGGAGTTGGAAC TTTTGGATGAGGTTGAGGAGTTGTTGA</p>	<p>“</p>
<p><i>kep3</i> (490115) codon optimized</p>	<p><u>GAATTC</u>ATGCGGATTGCAAGTGCAGCCCTATTCTTCGTTTCGGTACTGAGTGTGTTAGCGCCTTGCT ACACCTGAGGTCGATCCCGTCGCCCTGTGGCAGGAGTCTGTTGCACGTGAACCTCAAATGAGGCCCT TTCCTGGAGAAGAGACCCGGAACCTCAAAGCAGGCCCTTCCTGGCGACGGGAGCCCGAACCCGAG CCCGAACCTCAATACAAGACCCCTCATGGAGACGCGAACCCGAGCCTCAAAGCAGGCCTCCTGCCTG GCGCCGAGCCTGAACCTGAACCCAGTATAAAGCCCAAGCTGGCGTCTGTAACAGAACCTGAGC CACAGTATAAAGCACCTCGTGGCGTCGAGAGCCAGAACCAGGCCACAGTACAAAGCTCCTTCCTGG CGACGATCCGGAACCTCAGTACAAGGCCCTTCCTGGAGGCGCGAACCCGAACCCAGATGAGGCC TCCTCCTGGGGCCC</p>	<p>Coding sequence of <i>kep</i> 490115, codon optimized, with restriction sites (underlined), without STOP codon. The digested piece can immediately be ligated into the pPICZA vector.</p>

**Table S5:** Coding sequences of heterologously expressed *kep* genes.

## Chapter 2

Name	Description	Strain label	Source
pRS426-pAbGPDII-iePcGPD-Pcpab1-tPcMNP	Plasmid containing the <i>P. chrysosporium</i> Pcpab1 cassette selection marker	PMK #476	(Stöckli et al. 2017)
pET-NLS-Cas9-6xHis	Plasmid for <i>S. pyogenes</i> Cas9 expression in bacterial cells	PMA #1582	Addgene (Zuris et al. 2015)
pGEM-T easy_Snf5_Pcpab1	Repair-template plasmid for <i>snf5</i> disruption	PMA #1610	This study
pGEM-03563KO-PcPAB	Repair-template plasmid for <i>cop6</i> deletion	PMA #1184	(Stöckli et al. 2017)
pGEM-T easy_kex2a(502579)_Pcpab1	Repair-template plasmid for <i>kex</i> deletion	PMA #1628	This study
pGEM-Teasy_KO_kex2b(406374)_pab1	"	PMA #1658	This study
pGEM-Teasy_KO_kex2c(448165)_pab1	"	PMA #1657	This study
pGEM-Teasy_KO_kex1(437675)_pab1	"	PMA #1651	This study
pGEM-Teasy_KO_KEP1(405832)_pab1	Repair-template plasmid for <i>kep</i> deletion	PMA #1642	This study
pGEM-Teasy_KO_KEP2(434504)_pab1	"	PMA #1643	This study
pGEM-Teasy_KO_KEP3(490115)_pab1	"	PMA #1644	This study
pGEM-Teasy_KO_KEP4(497993)_pab1	"	PMA #1645	This study
pGEM-Teasy_KO_KEP5(503649)_pab1	"	PMA #1646	This study
pGEM-Teasy_KO_KEP6(426342)_pab1	"	PMA #1647	This study
pPICZA_KEP1(405832)	Plasmid for <i>P. pastoris</i> transformation	PMA #1574	This study
pPICZA_KEP2(434504)	"	PMA #1575	This study
pPICZA_KEP3(490115)	"	PMA #1613	This study
pPICZA_KEP4(497993)	"	PMA #1576	This study
pPICZA_KEP5(503649)	"	PMA #1577	This study
pPICZA_KEP6(426342)	"	PMA #1578	This study

**Table S6:** Plasmids used in this study.

snf5	ACUCAGCAUACCUCUAGAGCAGG
cop6_5'	GUUUACACAAGGACGUCCUACGG
cop6_3'	UCCUCACCACGUUACAGGCUGGG
kex1_5'	UGAGCGGUCUCAGCGGGCUU
kex1_3'	AUCUUCAGGUUUAUCGGAGCG
kex2a_5'	GCGGACGCUUCAAGGAAGGG
kex2a_3'	CUCUGCAACAAAUCAAGUC
kex2b_5'	CGCUCAGGUUGUGGCGUUA
kex2b_3'	GUCUCAGUGGACUUAACGCU
kex2c_5'	AUCGUGGCGAUUGACGAUUU
kex2c_3'	UUAUGUAUCUGUAAGACCGG
kep1_crRNA_5	GAGGGTCAGGGCAGAAACAA
kep1_crRNA_3	GCCAGACTGGGACTGTGAAG
kep2_crRNA_5	TCAATTTCTATCTATCAACA
kep2_crRNA_3	ACGAGCTCGACTAGATGCAC
kep3_crRNA_5	TCACTGTTGTCTCACGATG
kep3_crRNA_3	GTGGAAGGCGATCACCAGGA
kep4_crRNA_5	GATAGTGCAGAGAACATCG
kep4_crRNA_3	CTCTCAAGACTAATCACAA
kep5_crRNA_5	CGAGGAGAACGAAGTTGGTG

## Chapter 2

kep5_crRNA_3	TTGACTACATATGACTACA
kep6_crRNA_5	CGAGAGCGAAAAATGAGAAA
kep6_crRNA_3	TAGATGATCATGATGACCTG

**Table S7:** crRNAs used in this study. These sequences consist of the protospacer sequence without the protospacer adjacent motif (PAM) and without the tracrRNA-complementary region.

Fungal strain	Reference genome
<i>Coprinopsis cinerea</i> AmutBmut pab1.2 v1.0	(Muraguchi et al. 2015)
<i>Coprinopsis cinerea</i> Okayama 7 (#130)	(Stajich et al. 2010)
<i>Lentinula edodes</i> W1-26	(Chen et al. 2016)
<i>Pleurotus ostreatus</i> PC15 v2.0	(Riley et al. 2014)
<i>Pleurotus eryngii</i> ATCC 90797 v1.0	(Ruiz-Dueñas et al. 2021)
<i>Pichia pastoris</i> GS115	(De Schutter et al. 2009)
<i>Saccharomyces cerevisiae</i>	(Goffeau et al. 1996)
<i>Auriculariopsis ampla</i> NL-1724 v1.0	(Almási et al. 2019)
<i>Exidia glandulosa</i> v1.0	(Nagy et al. 2016)
<i>Gloeophyllum trabeum</i> v1.0	(Kerem et al. 1999)
<i>Heliocybe sulcata</i> OMC1185 v1.0	(Varga et al. 2019)
<i>Armillaria gallica</i> 21-2 v1.0	(Sipos et al. 2017)
<i>Psilocybe serbica</i> v1.0	(Fricke et al. 2017)
<i>Coprinopsis marcescibilis</i> CBS121175 v1.0	(Varga et al. 2019)
<i>Amanita thiersii</i> Skay4041 v1.0	(Hess et al. 2014)
<i>Termitomyces</i> sp. J132	(Poulsen et al. 2014)
<i>Psathyrella aberdarensis</i>	(Bau and Yan 2021)
<i>Hebeloma cylindrosporum</i> h7 v2.0	(Kohler et al. 2015)
<i>Hypholoma sublateritium</i> v1.0	(Kohler et al. 2015)
<i>Auricularia subglabra</i> v2.0	(Floudas et al. 2012)
<i>Dendrothele bispora</i> CBS 962.96 v1.0	(Varga et al. 2019)
<i>Gymnopus luxurians</i> v1.0	(Kohler et al. 2015)
<i>Panaeolus cyanescens</i>	GenBank GCA_002938355.1
<i>Laccaria bicolor</i> S238N-H82]	(Martin et al. 2008)
<i>Crucibulum laeve</i> CBS 166.37 v1.0	(Varga et al. 2019)
<i>Hypsizygus marmoreus</i> 51987-8	(Min et al. 2018)
<i>Coprinus phaeopunctatus</i> MPI-PUGE-AT-0042	(Mesny et al. 2021)
<i>Laccaria amethystina</i> LaAM-08-1 v2.0	(Kohler et al. 2015)
<i>Cyathus striatus</i> AH 40144 v1.0	(Ruiz-Dueñas et al. 2021)
<i>Lyophyllum atratum</i> CBS 144462 v1.0	(Steindorff et al. 2021)
<i>Asterophora parasitica</i> isolate AP01	NCBI accession number ASM1828200v1
<i>Crassisporium funariophilum</i> CBS 144457 v1.0	(Steindorff et al. 2021)

## Chapter 2

<i>Coprinellus micaceus</i> FP101781 v2.0	(Varga et al. 2019)
<i>Auriculariales</i> sp. MPI-PUGE-AT-0066	GenBank GCA_020744175.1
<i>Gymnopus luxurians</i> v1.0	(Kohler et al. 2015)
<i>Gymnopus confluens</i>	(Antonín et al. 1997)
<i>Rhodocollybia butyracea</i> AH 40177 v1.0	(Ruiz-Dueñas et al. 2021)
<i>Neurospora crassa</i> OR74A v2.0	(Galagan et al. 2003)
<i>Tricholoma furcatifolium</i>	NCBI accession number ASM1885489v1
<i>Blastosporella zonata</i>	NCBI accession number ASM1885629v1
<i>Guyanagaster necrorhizus</i> MCA 3950 v1.0	(Koch et al. 2021)
<i>Gautieria morchelliformis</i> GMNE.BST v1.0	(Miyachi et al. 2020)
<i>Pleurotus pulmonarius</i>	NCBI accession number ASM1298053v1
<i>Pleurotus cornucopiae</i>	(Zhang et al. 2021)
<i>Cryptococcus amyloletus</i> CBS 6039	(Sun et al. 2017)
<i>Cryptococcus floricola</i>	(Passer et al. 2019)
<i>Cryptococcus wingfieldii</i> CBS 7118	(Passer et al. 2019)
<i>Cryptococcus neoformans</i> var. <i>grubii</i> 125.91	NCBI accession number GCA_002215885
<i>Cryptococcus gattii</i> VGIV IND107	(Farrer et al. 2015)
<i>Cryptococcus depauperatus</i> CBS 7841	NCBI accession number GCA_001720195.1
<i>Kwoniella heveanensis</i> BCC8398	GCA_000507405.3
<i>Kwoniella dejecticola</i> CBS 10117	GCA_000512565.2
<i>Kwoniella pini</i> CBS 10737	GCA_000512605.2
<i>Kwoniella bestiolae</i> CBS 10118	GCA_000512585.2
<i>Kwoniella mangroviensis</i> CBS 10435	GCA_000507885.2
<i>Kwoniella shandongensis</i>	GCA_008629635.1
<i>Naematelia encephala</i>	GCA_002105065.1
<i>Saitozyma</i> sp. JCM 24511	GCA_001600855.1
<i>Cutaneotrichosporon oleaginosum</i>	GCA_008065305.1
<i>Apiotrichum porosum</i>	GCA_003942205.1
<i>Jaapia argillacea</i> MUCL 33604	GCA_000697665.1
<i>Piloderma croceum</i> F 1598	(Kohler et al. 2015)
<i>Fibularhizoctonia</i> sp. CBS 109695	(Nagy et al. 2016)
<i>Athelia</i> sp. TMB	(Konkel et al. 2021)
<i>Abortiporus biennis</i>	GCA_022606235.1
<i>Polyporus brumalis</i>	GCA_003367725.1
<i>Polyporus arcularius</i> HHB13444	GCA_004369055.1
<i>Lentinus tigrinus</i> ALCF2SS1-6	(Wu et al. 2018)
<i>Earliella scabrosa</i>	GCA_022605405.1
<i>Hexagonia nitida</i>	GCA_022606115.1
<i>Ganoderma leucocontextum</i>	GCA_022813035.1

## Chapter 2

<i>Leiotrametes menziesii</i>	GCA_022606275.1
<i>Leiotrametes lactinea</i>	GCA_022376465.1
<i>Trametes coccinea</i> BRFM310	GCA_002092935.1
<i>Trametes elegans</i>	GCA_022606155.1
<i>Trametes cingulata</i>	GCA_022385765.1
<i>Trametes versicolor</i> FP-101664 SS1	GCA_000271585.1
<i>Trametes pubescens</i>	GCA_001895945.1
<i>Trametes polyzona</i>	GCA_022606195.1
<i>Lenzites betulinus</i>	GCA_022264855.1
<i>Taiwanofungus camphoratus</i>	GCA_003999685.1
<i>Amylocystis lapponica</i>	GCA_022376435.1
<i>Postia placenta</i> MAD-698-R-SB12	(Martinez et al. 2009)
<i>Fibroporia radiculosa</i>	GCA_000313525.1
<i>Fomitopsis pinicola</i> FP-58527 SS1	(Floudas et al. 2012)
<i>Punctularia strigosozonata</i> HHB-11173 SS5	GCA_000264995.1
<i>Multifurca ochricompacta</i>	GCA_022496185.1
<i>Scleroderma citrinum</i> Foug A	(Kohler et al. 2015)
<i>Plicaturopsis crispa</i> FD-325 SS-3	GCA_000827205.1
<i>Galerina marginata</i> CBS 339.88	GCA_023014335.1
<i>Psilocybe cubensis</i>	GCA_017499595.2
<i>Panaeolus papilionaceus</i>	GCA_015501605.1
<i>Armillaria gallica</i>	GCA_012064365.1
<i>Mycena sanguinolenta</i>	GCA_014462675.1
<i>Lepista nuda</i>	GCA_015584075.1
<i>Tricholoma matsutake</i> 945	(Miyachi et al. 2020)
<i>Termitomyces</i> sp. T112	GCA_018296085.1
<i>Pluteus cervinus</i>	GCA_004369065.1
<i>Amanita brunnescens</i> Koide BX004	GCA_001691785.2
<i>Amanita rubescens</i>	GCA_015039365.1
<i>Vararia minispora</i> EC-137	(Looney et al. 2022)
<i>Serendipita vermifera</i> MAFF 305830	(Kohler et al. 2015)
<i>Calocera viscosa</i> TUFC12733	GCA_001630345.1
<i>Calocera cornea</i> HHB12733	GCA_001632435.1
<i>Dacryopinax primogenitus</i>	GCA_000292625.1
<i>Rhodotorula</i> sp. JG-1b	(Goordial et al. 2016)
<i>Rhodotorula</i> sp. CCFEE 5036	(Coleine et al. 2020)

**Table S8:** Reference genomes used in this study.

Chapter 2

<b><i>Coprinopsis cinerea</i></b>		
<b>Gene name</b>	<b>JGI protein ID Okayama7 (Stajich et al. 2010)</b>	<b>JGI protein ID AmutBmut (Muraguchi et al. 2015)</b>
<i>kep</i>	CC1G_04905	405832
<i>kep</i>	CC1G_06528	434504
<i>kep</i>	CC1G_11591	490115
<i>kep</i>	CC1G_10545	497993
<i>kep</i>	CC1G_06039	503649
<i>kep</i>	CC1G_09529	426342
<i>kep</i>	CC1G_06036	447393
Gene encoding glycoside hydrolase domain-containing protein	CC1G_01253	365456
<i>kex1</i>	CC1G_01392	437675
<i>kex2a</i>	CC1G_01625	502579
<i>kex2b</i>	CC1G_02379	406374
<i>kex2c</i>	CC1G_12012	448165
<i>Ste13</i>	CC1G_04021	545529
Glutaminy cyclase ( <i>qc</i> )	CC1G_13886	360529
<i>snf5</i>	CC1G_15539	365798
<i>cop6</i>	CC1G_03563	394772
<b><i>Lentinula edodes</i> W1-26</b>		
	<b>JGI protein ID (Chen et al. 2016)</b>	
<i>kep</i>	2599	
<b><i>Pleurotus ostreatus</i> PC15 v2.0</b>		
	<b>JGI protein ID (Riley et al. 2014)</b>	
<i>kep</i>	1091723	
<b><i>Pleurotus eryngii</i> ATCC 90797 v1.0</b>		
	<b>JGI protein ID (Ruiz-Dueñas et al. 2021)</b>	
<i>kep</i>	439342	

**Table S9:** JGI protein IDs of proteins used in this study.

ANOVA results	Sum of Squares	dF	Mean Square	F (DFn, DFd)	P value
Treatment (between columns)	43.27	6	7.211	F (6, 13) = 25.00	P<0.0001
Residual (within columns)	3.75	13	0.2885		
Total	47.02	19			

**Table S10:** ANOVA results of mycelial diameter comparison between *C. cinerea* knockout strains.

## Chapter 2

Strain label	Sample origin	Description	Original file name
<i>P. pastoris</i> GS115-405832	SN	Heterologous expression of <i>C. cinerea</i> KEP in <i>P. pastoris</i>	20200506_006_S216402_2020_Pp_KEP1.raw
<i>P. pastoris</i> GS115-434504	SN	Heterologous expression of <i>C. cinerea</i> KEP in <i>P. pastoris</i>	20200506_002_S216403_2020_Pp_KEP2_rep.raw
<i>P. pastoris</i> GS115-490115	SN	Heterologous expression of <i>C. cinerea</i> KEP in <i>P. pastoris</i>	20200506_003_S216404_2020_Pp_KEP3.raw
<i>P. pastoris</i> GS115-497993	SN	Heterologous expression of <i>C. cinerea</i> KEP in <i>P. pastoris</i>	20200506_007_S216405_2020_Pp_KEP4.raw
<i>P. pastoris</i> GS115-503649	SN	Heterologous expression of <i>C. cinerea</i> KEP in <i>P. pastoris</i>	20200506_004_S216406_2020_Pp_KEP5.raw
<i>P. pastoris</i> GS115-426342	SN	Heterologous expression of <i>C. cinerea</i> KEP in <i>P. pastoris</i>	20200506_008_S216407_2020_Pp_KEP6.raw
<i>P. pastoris</i> GS115-empty	SN	<i>P. pastoris</i> strain transformed with an empty vector, as negative control for heterologous expression	20190808_004_S202141_DDA_Pp_empty.raw
<i>C. cinerea</i> AmutBmut pab1-2	SN	<i>C. cinerea</i> supernatant and tissue samples for screening for KEP-derived peptides	20200715_004_S219838_2020July_C_c_AmutBmut_HLB.raw
<i>C. cinerea</i> AmutBmut pab1-2 $\Delta$ ku70	SN	"	20211104_027_S325406_021121_ku70_SN_HLB_correctmethod.raw
"	Premature cap	"	20210818_009_S311407_170821_ku70_FB_HLB.raw
"	Cap	"	20211104_024_S325410_021121_ku70_overmature_HLB.raw
"	Mycelium	"	20211021_002_S322908_191021_ku70_drymyc1_MeOH37_HLB_20211022155916.raw
"	Basidiospores	"	20210818_011_S311409_170821_ku70_basidiospore.raw
"	Stem	"	20211104_021_S325407_021121_ku70_stem_HLB.raw
<i>C. cinerea</i> AmutBmut $\Delta$ ku70 $\Delta$ kep 405832 #1"	SN	Knockout strain of <i>C. cinerea</i> KEP	20220913_008_S412406_120922_CcKO_KEP1_SP3.raw
<i>C. cinerea</i> AmutBmut $\Delta$ ku70 $\Delta$ kep 434504#1	SN	Knockout strain of <i>C. cinerea</i> KEP	20220905_028_S410966_020922_CcKO_KEP2_HLB.raw
<i>C. cinerea</i> AmutBmut $\Delta$ ku70 $\Delta$ kep 490115#1	SN	Knockout strain of <i>C. cinerea</i> KEP	20210805_013_S310048_020821_31B_SP3.raw
<i>C. cinerea</i> AmutBmut $\Delta$ ku70 $\Delta$ kep 497993#1	SN	Knockout strain of <i>C. cinerea</i> KEP	20220913_002_S310050_020821_44C_SP3.raw
<i>C. cinerea</i> AmutBmut $\Delta$ ku70 $\Delta$ kep 497993#1	Cap	Knockout strain of <i>C. cinerea</i> KEP	20210916_021_S316137_150921_44C_FB.raw
<i>C. cinerea</i> AmutBmut $\Delta$ ku70 $\Delta$ kep 503649"	SN	Knockout strain of <i>C. cinerea</i> KEP	20220913_009_S412407_120922_CcKO_KEP5_SP3.raw
<i>C. cinerea</i> AmutBmut $\Delta$ ku70 $\Delta$ kep 426342#1	SN	Knockout strain of <i>C. cinerea</i> KEP	20220913_010_S412408_120922_CcKO_KEP6_SP3.raw
<i>C. cinerea</i> AmutBmut $\Delta$ ku70 $\Delta$ kex1 (437675)	SN	Knockout strain of <i>C. cinerea</i> KEX protease	20220425_014_S368954_020322_kex1_2.raw
<i>C. cinerea</i> AmutBmut $\Delta$ ku70 $\Delta$ kex2a (502579)	SN	Knockout strain of <i>C. cinerea</i> KEX protease	20220425_012_S368908_020322_kex2a_2.raw
"	Cap	Knockout strain of <i>C. cinerea</i> KEX protease	20220913_004_S316138_150921_kex2a_FB.raw
<i>C. cinerea</i> AmutBmut $\Delta$ ku70 $\Delta$ kex2b (406374)	SN	Knockout strain of <i>C. cinerea</i> KEX protease	20220425_018_S382520_250422_kex2b-1.raw
<i>C. cinerea</i> AmutBmut $\Delta$ ku70 $\Delta$ kex2c (448165)	SN	Knockout strain of <i>C. cinerea</i> KEX protease	20220425_022_S382523_250422_kex2c-6d-1.raw
<i>C. cinerea</i> AmutBmut $\Delta$ ku70 $\Delta$ kex2a/c (502579, 448165)	SN	Double knockout strain of <i>C. cinerea</i> KEX proteases	20220425_026_S382526_250422_kex2ac-6d-1.raw
<i>Lentinula edodes</i> (strain 4312, Sylvan, USA)	Cap	Fruiting bodies purchased from a local mushroom farm	20210818_003_S311403_170821_shitake_fresh.raw

## Chapter 2

<i>Pleurotus ostreatus</i> (strain P24/HK35, Sylvan, USA)	Cap	Fruiting bodies purchased from a local mushroom farm	20210818_008_S311406_170821_pleurotus_dry.raw
<i>Pleurotus ostreatus</i> (strain P24/HK35, Sylvan, USA)	Cap	Fruiting bodies purchased from a local mushroom farm	20211021_005_S322911_191021_Pe_MeOH-37-HLB.raw
<i>Pleurotus eryngii</i> (strain 3066, Sylvan, USA)	Cap	Fruiting bodies purchased from a local mushroom farm	20220303_002_S316123_150921_Pe_MeOH-37-HLB.raw

**Table S11:** List of proteomics data deposited to the ProteomeXchange Consortium. The data can be found online (<http://proteomecentral.proteomexchange.org>) with the dataset identifier PXD036934. The data was uploaded using the PRIDE partner repository (Perez-Riverol et al. 2019).

**Supplemental references**

Almási É, Sahu N, Krizsán K, Bálint B, Kovács GM, Kiss B, Cseklye J, Drula E, Henrissat B, Nagy I, et al. 2019. Comparative genomics reveals unique wood-decay strategies and fruiting body development in the Schizophyllaceae. *New Phytol.* 224(2):902–915. <https://doi.org/10.1111/nph.16032>

Ando Y, Nakazawa T, Oka K, Nakahori K, Kamada T. 2013. Cc.snf5, a gene encoding a putative component of the SWI/SNF chromatin remodeling complex, is essential for sexual development in the agaricomycete *Coprinopsis cinerea*. *Fungal Genet Biol [Internet]*. 50(1):82–89. <https://doi.org/10.1016/j.fgb.2012.09.010>

Antonín V, Halling RE, Noordeloos ME. 1997. Generic concepts within the groups of *Marasmius* and *Collybia sensu lato*. *Mycotaxon*. 63(June):359–368.

Bau T, Yan JQ. 2021. Two new rare species of *Candolleomyces* with pale spores from China. *MycoKeys*. 80:149–161. <https://doi.org/10.3897/MYCOKEYS.80.67166>

Chen L, Gong Y, Cai Y, Liu W, Zhou Y, Xiao Y, Xu Z, Liu Y, Lei X, Wang G, et al. 2016. Genome sequence of the edible cultivated mushroom *Lentinula edodes* (shiitake) reveals insights into lignocellulose degradation. *PLoS One*. 11(8):1–20. <https://doi.org/10.1371/journal.pone.0160336>

Coleine C, Masonjones S, Onofri S, Selbmann L, Stajich JE. 2020. Draft Genome Sequence of the Yeast *Rhodotorula* sp. Strain CCFEE 5036, Isolated from McMurdo Dry Valleys, Antarctica. 9(14):14–16. <https://doi.org/10.1128/MRA.00020-20>

Farrer RA, Desjardins CA, Sakthikumar S, Gujja S, Saif S, Zeng Q, Chen Y, Voelz K, Heitman J, May RC, et al. 2015. Genome evolution and innovation across the four major lineages of *Cryptococcus gattii*. *MBio*. 6(5). <https://doi.org/10.1128/mbio.00868-15>

Floudas D, Binder M, Riley R, Barry K, Blanchette RA, Henrissat B, Martínez AT, Otilar R, Spatafora JW, Yadav JS, et al. 2012. The paleozoic origin of enzymatic lignin decomposition reconstructed from 31 fungal genomes. *Science (80- )*. 336(6089):1715–1719. <https://doi.org/10.1126/science.1221748>

Fricke J, Blei F, Hoffmeister D. 2017. Enzymatic Synthesis of Psilocybin. *Angew Chemie - Int Ed*. 56(40):12352–12355. <https://doi.org/10.1002/anie.201705489>

Galagan JE, Calvo SE, Borkovich KA, Selker EU, Read NO, Jaffe D, FitzHugh W, Ma LJ, Smirnov S, Purcell S, et al. 2003. The genome sequence of the filamentous fungus *Neurospora crassa*. *Nature*. 422(6934):859–868. <https://doi.org/10.1038/nature01554>

Goffeau A, Barrell G, Bussey H, Davis RW, Dujon B, Feldmann H, Galibert F, Hoheisel JD, Jacq C, Johnston M, et al. 1996. Life with 6000 genes. *Science (80- )*. 274(5287):546–567. <https://doi.org/10.1126/science.274.5287.546>

Goordial J, Raymond-Bouchard I, Riley R, Ronholm J, Shapiro N, Woyke T, LaButti KM, Tice H, Amirebrahimi M, Grigoriev I V., et al. 2016. Improved high-quality draft genome sequence of the eurypsychrophile *Rhodotorula* sp. JG1b, isolated from permafrost in the hyperarid upper-elevation McMurdo Dry Valleys, Antarctica. *Genome Announc*. 4(2):9–10. <https://doi.org/10.1128/genomeA.00069-16>

Hess J, Skrede I, Wolfe BE, Butti K La, Ohm RA, Grigoriev I V., Pringle A. 2014. Transposable

element dynamics among asymbiotic and ectomycorrhizal amanita fungi. *Genome Biol Evol.* 6(7):1564–1578. <https://doi.org/10.1093/gbe/evu121>

Kerem Z, Jensen KA, Hammel KE. 1999. Biodegradative mechanism of the brown rot basidiomycete *Gloeophyllum trabeum*: Evidence for an extracellular hydroquinone-driven fenton reaction. *FEBS Lett.* 446(1):49–54. [https://doi.org/10.1016/S0014-5793\(99\)00180-5](https://doi.org/10.1016/S0014-5793(99)00180-5)

Koch RA, Yoon GM, Aryal UK, Lail K, Amirebrahimi M, LaButti K, Lipzen A, Riley R, Barry K, Henrissat B, et al. 2021. Symbiotic nitrogen fixation in the reproductive structures of a basidiomycete fungus. *Curr Biol.* 31(17):3905-3914.e6. <https://doi.org/10.1016/j.cub.2021.06.033>

Kohler A, Kuo A, Nagy LG, Morin E, Barry KW, Buscot F, Canbäck B, Choi C, Cichocki N, Clum A, et al. 2015. Convergent losses of decay mechanisms and rapid turnover of symbiosis genes in mycorrhizal mutualists. *Nat Genet.* 47(4):410–415. <https://doi.org/10.1038/ng.3223>

Kombrink A, Tayyrov A, Essig A, Stöckli M, Micheller S, Hintze J, van Heuvel Y, Dürig N, Lin C wei, Kallio PT, et al. 2019. Induction of antibacterial proteins and peptides in the coprophilous mushroom *Coprinopsis cinerea* in response to bacteria. *ISME J [Internet].* 13(3):588–602. <https://doi.org/10.1038/s41396-018-0293-8>

Konkel Z, Scott K, Slot JC. 2021. Draft Genome Sequence of the Termite-Associated “Cuckoo Fungus,” *Athelia ( Fibularhizoctonia )* sp. TMB Strain TB5. *Microbiol Resour Announc.* 10(1):28–30. <https://doi.org/10.1128/mra.01230-20>

Looney B, Miyauchi S, Morin E, Drula E, Courty PE, Kohler A, Kuo A, LaButti K, Pangilinan J, Lipzen A, et al. 2022. Evolutionary transition to the ectomycorrhizal habit in the genomes of a hyperdiverse lineage of mushroom-forming fungi. *New Phytol.* 233(5):2294–2309. <https://doi.org/10.1111/nph.17892>

Martin F, Aerts A, Ahrén D, Brun A, Danchin EGJ, Duchaussoy F, Gibon J, Kohler A, Lindquist E, Pereda V, et al. 2008. The genome of *Laccaria bicolor* provides insights into mycorrhizal symbiosis. *Nature.* 452(7183):88–92. <https://doi.org/10.1038/nature06556>

Martinez D, Challacombe J, Morgenstern I, Hibbett D, Schmoll M, Kubicek CP, Ferreira P, Ruiz-Duenas FJ, Martinez AT, Kersten P, et al. 2009. Genome, transcriptome, and secretome analysis of wood decay fungus *Postia placenta* supports unique mechanisms of lignocellulose conversion. *Proc Natl Acad Sci U S A.* 106(6):1954–1959. <https://doi.org/10.1073/pnas.0809575106>

Mesny F, Miyauchi S, Thiergart T, Pickel B, Atanasova L, Karlsson M, Hüttel B, Barry KW, Haridas S, Chen C, et al. 2021. Genetic determinants of endophytism in the *Arabidopsis* root mycobiome. *Nat Commun.* 12(1):1–15. <https://doi.org/10.1038/s41467-021-27479-y>

Min B, Kim S, Oh YL, Kong WS, Park H, Cho H, Jang KY, Kim JG, Choi IG. 2018. Genomic discovery of the hypsin gene and biosynthetic pathways for terpenoids in *Hypsizygus marmoreus*. *BMC Genomics.* 19(1):1–12. <https://doi.org/10.1186/s12864-018-5159-y>

Miyauchi S, Kiss E, Kuo A, Drula E, Kohler A, Sánchez-García M, Morin E, Andreopoulos B, Barry KW, Bonito G, et al. 2020. Large-scale genome sequencing of mycorrhizal fungi provides insights into the early evolution of symbiotic traits. *Nat Commun.* 11(1):1–17. <https://doi.org/10.1038/s41467-020-18795-w>

Muraguchi H, Umezawa K, Niikura M, Yoshida M, Kozaki T, Ishii K, Sakai K, Shimizu M, Nakahori

## Chapter 2

- K, Sakamoto Y, et al. 2015. Strand-specific RNA-seq analyses of fruiting body development in *Coprinopsis cinerea*. *PLoS One*. 10(10):1–23. <https://doi.org/10.1371/journal.pone.0141586>
- Nagy LG, Riley R, Tritt A, Adam C, Daum C, Floudas D, Sun H, Yadav JS, Pangilinan J, Larsson KH, et al. 2016. Comparative genomics of early-diverging mushroom-forming fungi provides insights into the origins of lignocellulose decay capabilities. *Mol Biol Evol*. 33(4):959–970. <https://doi.org/10.1093/molbev/msv337>
- Nakazawa T, Honda Y. 2015. Absence of a gene encoding cytosine deaminase in the genome of the agaricomycete *Coprinopsis cinerea* enables simple marker recycling through 5-fluorocytosine counterselection. *FEMS Microbiol Lett*. 362(15):1–7. <https://doi.org/10.1093/femsle/fnv123>
- Passer AR, Coelho MA, Billmyre RB, Nowrousian M, Mittelbach M, Yurkov AM, Averette AF, Cuomo CA, Sun S, Heitman J. 2019. Genetic and genomic analyses reveal boundaries between species closely related to *Cryptococcus* pathogens. *MBio*. 10(3). <https://doi.org/10.1128/mBio.00764-19>
- Pearce M, Tivey ARN, Basutkar P, Lee J, Edbali O, Madhusoodanan N, Kolesnikov A, Lopez R. 2022. Search and sequence analysis tools services from EMBL-EBI in 2022 *F abio*. :1–4.
- Perez-Riverol Y, Csordas A, Bai J, Bernal-Llinares M, Hewapathirana S, Kundu DJ, Inuganti A, Griss J, Mayer G, Eisenacher M, et al. 2019. The PRIDE database and related tools and resources in 2019: Improving support for quantification data. *Nucleic Acids Res*. 47(D1):D442–D450. <https://doi.org/10.1093/nar/gky1106>
- Petersen TN, Brunak S, Von Heijne G, Nielsen H. 2011. SignalP 4.0: Discriminating signal peptides from transmembrane regions. *Nat Methods*. 8(10):785–786. <https://doi.org/10.1038/nmeth.1701>
- Poulsen M, Hu H, Li C, Chen Z, Xu L, Otani S, Nygaard S, Nobre T, Klaubauf S, M. Schindler P, et al. 2014. Complementary symbiont contributions to plant decomposition in a fungus-farming termite. *Proc Natl Acad Sci U S A*. 111(40):14500–14505. <https://doi.org/10.1073/pnas.1319718111>
- Riley R, Salamov AA, Brown DW, Nagy LG, Floudas D, Held BW, Levasseur A, Lombard V, Morin E, Otiillar R, et al. 2014. Extensive sampling of basidiomycete genomes demonstrates inadequacy of the white-rot/brown-rot paradigm for wood decay fungi. *Proc Natl Acad Sci U S A*. 111(27):9923–9928. <https://doi.org/10.1073/pnas.1400592111>
- Robert X, Gouet P. 2014. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res*. 42(W1):320–324. <https://doi.org/10.1093/nar/gku316>
- Ruiz-Deñás FJ, Barrasa JM, Sánchez-García M, Camarero S, Miyauchi S, Serrano A, Linde D, Babiker R, Drula E, Ayuso-Fernández I, et al. 2021. Genomic Analysis Enlightens Agaricales Lifestyle Evolution and Increasing Peroxidase Diversity. *Mol Biol Evol*. 38(4):1428–1446. <https://doi.org/10.1093/molbev/msaa301>
- De Schutter K, Lin YC, Tiels P, Van Hecke A, Glinka S, Weber-Lehmann J, Rouzé P, Van De Peer Y, Callewaert N. 2009. Genome sequence of the recombinant protein production host *Pichia pastoris*. *Nat Biotechnol*. 27(6):561–566. <https://doi.org/10.1038/nbt.1544>
- Sipos G, Prasanna AN, Walter MC, O'Connor E, Bálint B, Krizsán K, Kiss B, Hess J, Varga T, Slot J, et al. 2017. Genome expansion and lineage-specific genetic innovations in the forest

## Chapter 2

pathogenic fungi *Armillaria*. *Nat Ecol Evol.* 1(12):1931–1941. <https://doi.org/10.1038/s41559-017-0347-8>

Stajich JE, Wilke SK, Ahrén D, Au CH, Birren BW, Borodovsky M, Burns C, Canbäck B, Casselton LA, Cheng CK, et al. 2010. Insights into evolution of multicellular fungi from the assembled chromosomes of the mushroom *Coprinopsis cinerea* (*Coprinus cinereus*). *Proc Natl Acad Sci U S A.* 107(26):11889–11894. <https://doi.org/10.1073/pnas.1003391107>

Steindorff AS, Carver A, Calhoun S, Stillman K, Liu H, Lipzen A, He G, Yan M, Pangilinan J, LaButti K, et al. 2021. Comparative genomics of pyrophilous fungi reveals a link between fire events and developmental genes. *Environ Microbiol.* 23(1):99–109. <https://doi.org/10.1111/1462-2920.15273>

Stöckli M, Lin C wei, Sieber R, Plaza DF, Ohm RA, Künzler M. 2017. *Coprinopsis cinerea* intracellular lactonases hydrolyze quorum sensing molecules of Gram-negative bacteria. *Fungal Genet Biol.* 102:49–62. <https://doi.org/10.1016/j.fgb.2016.07.009>

Sun S, Yadav V, Billmyre RB, Cuomo CA, Nowrousian M, Wang L, Souciet JL, Boekhout T, Porcel B, Wincker P, et al. 2017. Fungal genome and mating system transitions facilitated by chromosomal translocations involving intercentromeric recombination. [place unknown]. <https://doi.org/10.1371/journal.pbio.2002527>

Swamy S, Uno I, Ishikawa T. 1984. Morphogenetic effects of mutations at the A and B incompatibility factors in *Coprinus cinereus*. *J Gen Microbiol.* 130(12):3219–3224. <https://doi.org/10.1099/00221287-130-12-3219>

Tayyrov A, Stanley CE, Azevedo S, Künzler M. 2019. Combining microfluidics and RNA-sequencing to assess the inducible defensome of a mushroom against nematodes. *BMC Genomics* [Internet]. 20(1):243. <https://doi.org/10.1186/s12864-019-5607-3>

Varga T, Krizsán K, Földi C, Dima B, Sánchez-García M, Sánchez-Ramírez S, Szöllősi GJ, Szarkándi JG, Papp V, Albert L, et al. 2019. Megaphylogeny resolves global patterns of mushroom evolution. *Nat Ecol Evol.* 3(4):668–678. <https://doi.org/10.1038/s41559-019-0834-1>

Wälti MA, Villalba C, Buser RM, Grünler A, Aebi M, Künzler M. 2006. Targeted gene silencing in the model mushroom *Coprinopsis cinerea* (*Coprinus cinereus*) by expression of homologous hairpin RNAs. *Eukaryot Cell.* 5(4):732–744. <https://doi.org/10.1128/EC.5.4.732-744.2006>

Wu B, Xu Z, Knudson A, Carlson A, Chen N, Kovaka S, LaButti K, Lipzen A, Pennachio C, Riley R, et al. 2018. Genomics and development of *Lentinus tigrinus*: A white-rot wood-decaying mushroom with dimorphic fruiting bodies. *Genome Biol Evol.* 10(12):3250–3261. <https://doi.org/10.1093/gbe/evy246>

Zhang Y, Gao W, Sonnenberg A, Chen Q, Zhang J, Huang C. 2021. Genetic Linkage and Physical Mapping for an Oyster Mushroom (*Pleurotus cornucopiae*) and Quantitative Trait Locus Analysis for Cap Color. *Appl Environ Microbiol.* 87(21):e0095321. <https://doi.org/10.1128/AEM.00953-21>

Zuris JA, Thompson DB, Shu Y, Guilinger JP, Bessen JL, Hu JH, Maeder ML, Joung JK, Chen ZY, Liu DR. 2015. Cationic lipid-mediated delivery of proteins enables efficient protein-based genome editing in vitro and in vivo. *Nat Biotechnol.* 33(1):73–80. <https://doi.org/10.1038/nbt.3081>

## Appendix 2: Functional assays on KEP-derived peptides

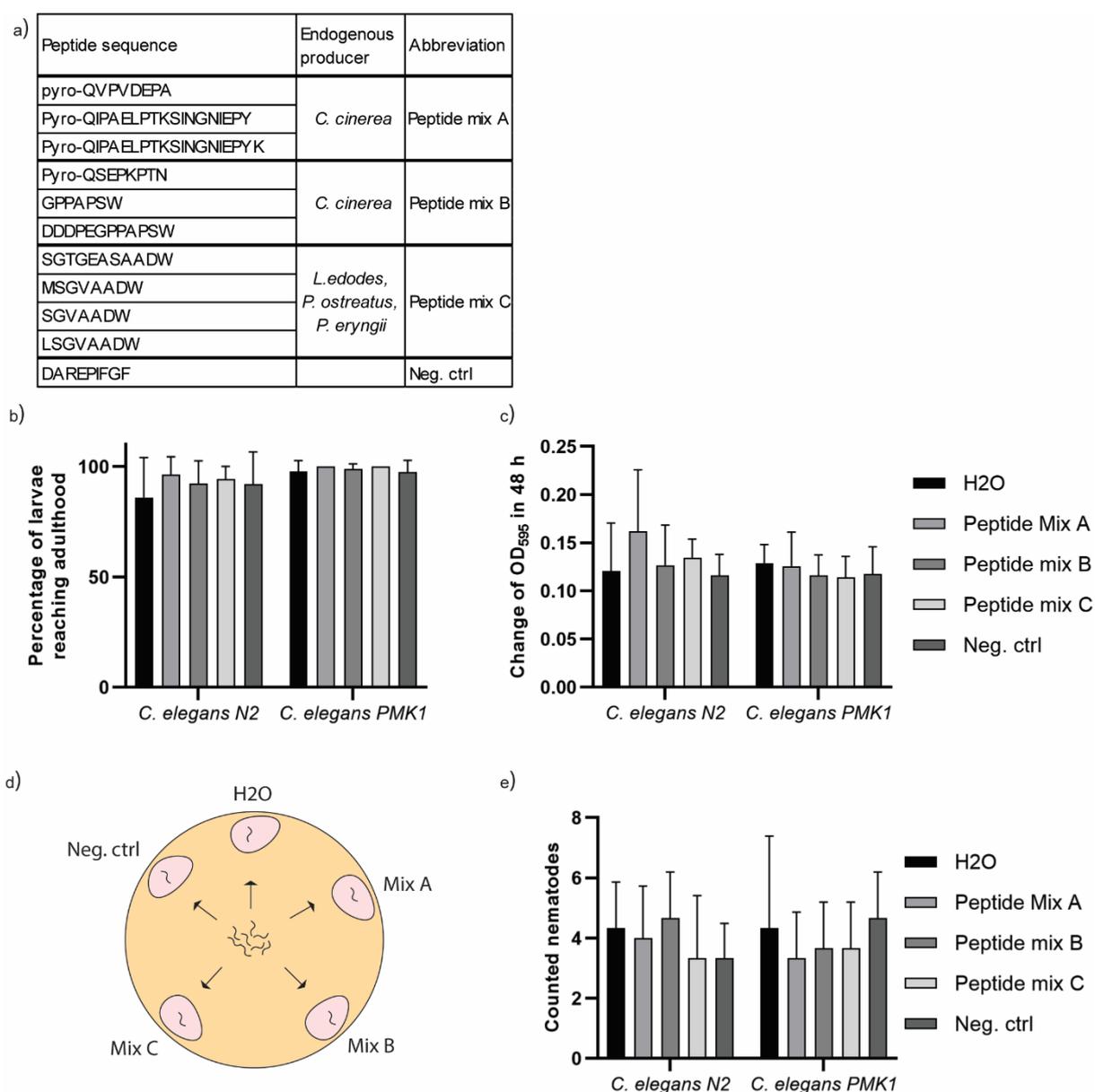
**A 2.1 *C. elegans* toxicity assay with synthetic peptides**

We confirmed several KEP-derived peptides in fruiting bodies: pyro-QIPAE LPTKSINGNIEPY, pyro-QIPAE LPTKSINGNIEPYK and pyro-QSEPKPTN in *C. cinerea*, SGTGEASAADW in *L. edodes* and MSGVAADW, SGVAADW and LSGVAADW in *P. ostreatus* and *P. eryngii*. We speculated that these peptides might play a role in the defense of the fruiting bodies against predators, especially the *C. cinerea* peptides pyro-QIPAE LPTKSINGNIEPY and pyro-QIPAE LPTKSINGNIEPYK, which were detected mainly in the fruiting bodies and were undetectable or barely detectable in the mycelium. Therefore, we tested the peptides in preliminary toxicity and deterrence assays against the nematode strains *Caenorhabditis elegans* N2 (wildtype) and *C. elegans* PMK1 (hypersensitive strain with reduced innate immunity, Bolz et al. 2010), in which the synthetic peptides are dissolved in the bacterial solution that feeds the worms.

Although we did not observe any effect of the synthetic peptides on the growth or deterrence of *C. elegans* N2 or PMK1 (Appendix Fig. 1), we plan to repeat the assay using an adapted experimental procedure. In previous nematode toxicity assays we heterologously expressed the proteins of interest in *E. coli* BL21 and subsequently fed them to *C. elegans* (Künzler et al. 2010), an experimental setup that ensures that the protein of interest reaches the intestinal tract of the nematode. In a heterologous expression of KEP proteins in *E. coli* the precursors would not be processed adequately because of the lack of a secretory system and a dibasic residue-specific protease. The bacterial protease subtilisin shows sequence similarity to fungal KEX2 proteases and mammalian proprotein convertase, but cleaves indiscriminately and shows mainly extracellular activity (Power et al. 1986; Grøn et al. 1992). Therefore, we plan to coexpress a TEV protease with a synthetic KEP construct, consisting of peptide cores flanked by TEV cleavage sites, in *E. coli* BL21. The TEV protease should cleave the KEP construct intracellularly, resulting in the accumulation of peptide products in the bacteria that would then be fed to nematodes in a toxicity assay. The TEV cleavage site is seven residues long with cleavage occurring after the sixth residue. Therefore, a disadvantage of TEV-based cleavage of peptides is the tail of six residues that remains C-terminally attached to the peptide (Parks et al. 1994).

**A2.2 Germination and fusion assays of *C. cinerea* knockout strains**

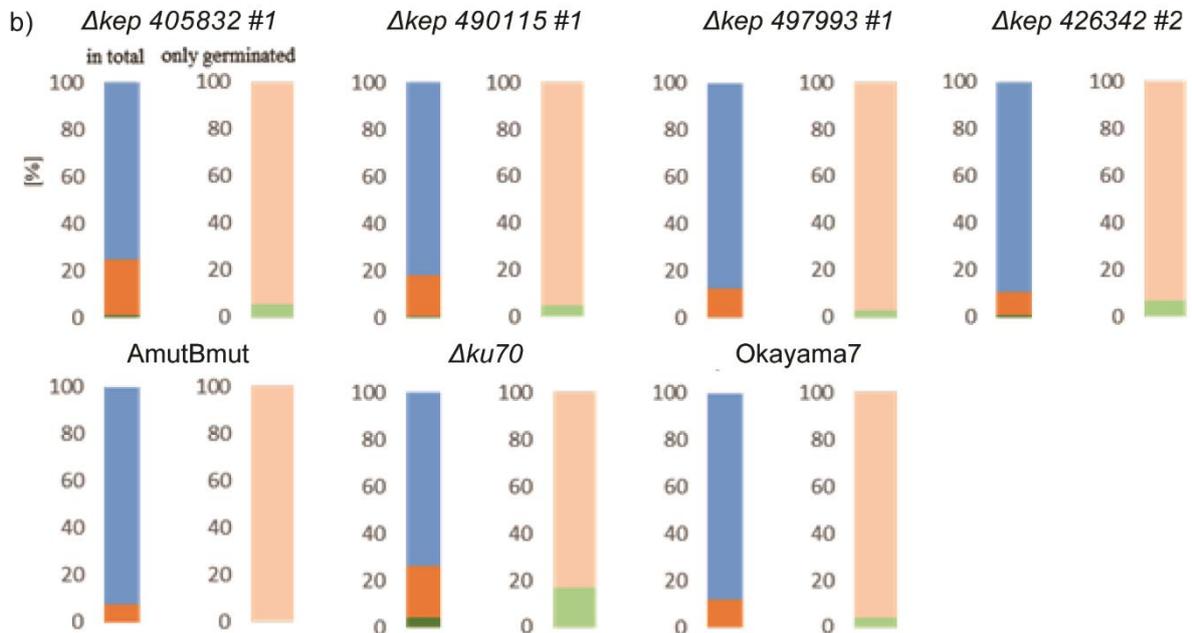
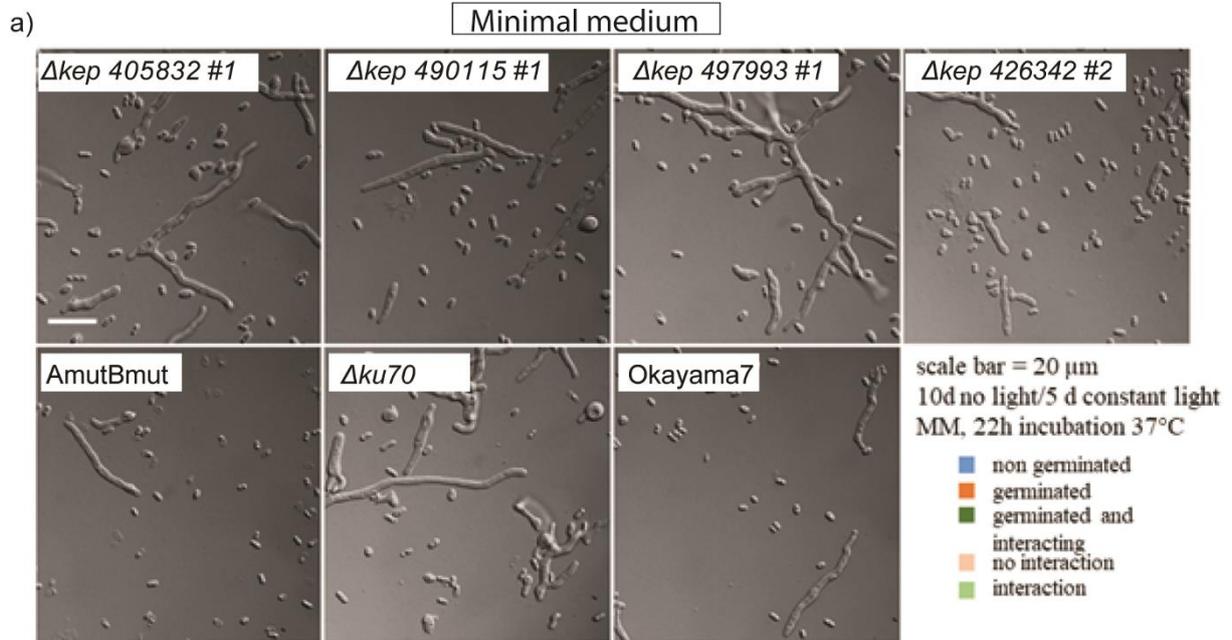
The fungal mycelium is a densely interconnected network formed by the fusion of hyphae, a process called anastomosis in which two hyphae with identical compatibility engage in communication, grow toward each other and fuse (Glass and Fleissner 2006). In *Neurospora crassa*, the two proteins MAP kinase MAK-2 and a protein of unknown molecular function SO are recruited to the plasma membrane of interacting hyphae in an oscillatory manner. This “ping-pong”-like mechanism of oscillatory signalling is thought to ensure that self-stimulation is avoided (Fleissner et al. 2009).

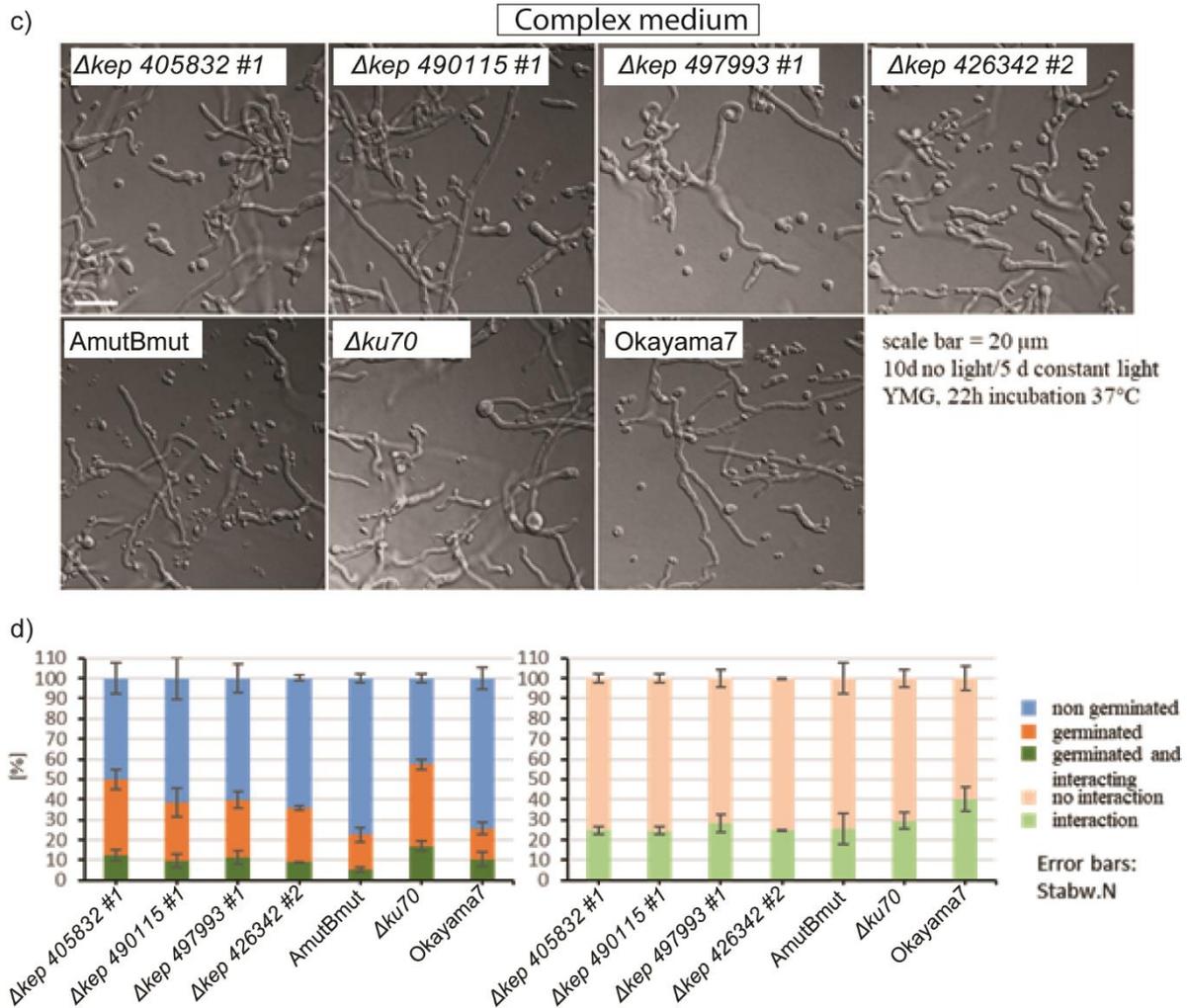


**Appendix 2 Figure 1:** Peptide toxicity and deterrence assays with *C. elegans*. Chemically synthesized peptides previously confirmed in *C. cinerea*, *L. edodes*, *P. ostreatus* and *P. eryngii* were tested for their inhibition of nematode growth and their nematode deterrence. a) List of synthetic peptides tested in this assay. Peptides were not tested individually but were combined in peptide mixtures. b) Development assays of *C. elegans* N2 and PMK1. The percentages of L1 larvae that reached adulthood after 48 hours are shown in the graph. Nematode eggs were bleached according to wormbook guidelines (Stiernagle 2006), and the assay was prepared according to Tayyrov et al. 2019, except that bacteria were diluted to an OD<sub>595</sub> of 1 instead of 2. Briefly, 10-15 L1 larvae were added to an *E. coli* BL21 cell suspension containing synthetic peptides in 96-well microtiter plates. After 48 h, the number of larvae that had reached adulthood was counted. Each peptide was tested at a concentration of 20 µg/ml. c) OD<sub>595</sub> readout of the nematode development assay. Bacterial optical density change was measured to track *C. elegans* feeding. d) Illustration of a *C. elegans* deterrence assay using synthetic peptides. Synthetic peptides were mixed with *E. coli* BL21 with an OD<sub>595</sub> of 1 to a final concentration of 20 µg/ml for each peptide and 15 µl of cell suspension were added to the outer edge of an agar plate and dried. Approximately 50 L1 larvae of *C. elegans* were placed in the center of the plate and the number of nematodes that had crawled onto each bacterial lawn was counted after 24 h. e) Results of the *C. elegans* deterrence assay.

## Chapter 2

Initial cell communication is assumed to rely on the secretion of diffusible molecules that lead to polarized hyphal growth (Fleissner et al. 2009). Based on the discoveries of anastomosis-regulating signalling in *N. crassa*, we speculated that KEP-derived peptides may be involved in this process. Anastomosis between germinating oidia was examined in *C. cinerea* knockout strains in minimal medium (Appendix Fig. 2a-b) and complex medium (Appendix Fig. 2c-d).



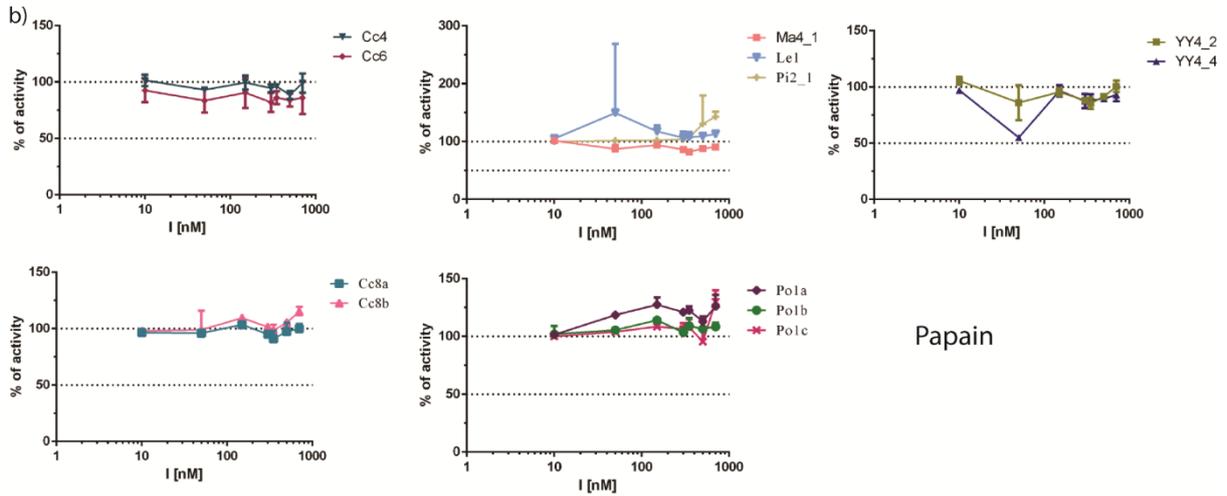


**Appendix 2 Figure 2:** Oidial germination and interaction of *C. cinerea* knockout strains. Freshly harvested oidia were incubated in minimal or complex medium at 37°C for 22 hours and then analyzed under the microscope. a) Microscopic images of germinating oidia in minimal medium. Knockout strains were compared with *C. cinerea* AmutBmut,  $\Delta\text{ku}70$ , and Okayama7. b) Analysis of oidial germination and interaction in minimal medium. The percentage of germinated or germinated and interacting oidia out of the total number of oidia is shown in the left bar for each fungal strain. The right bar shows the percentage of interactions in all germinated oidia. c) Microscopic images of germinating oidia in complex medium. d) Analysis of oidial germination and interaction in complex medium. The experiments were performed by Stephanie Herzog, Technical University of Braunschweig.

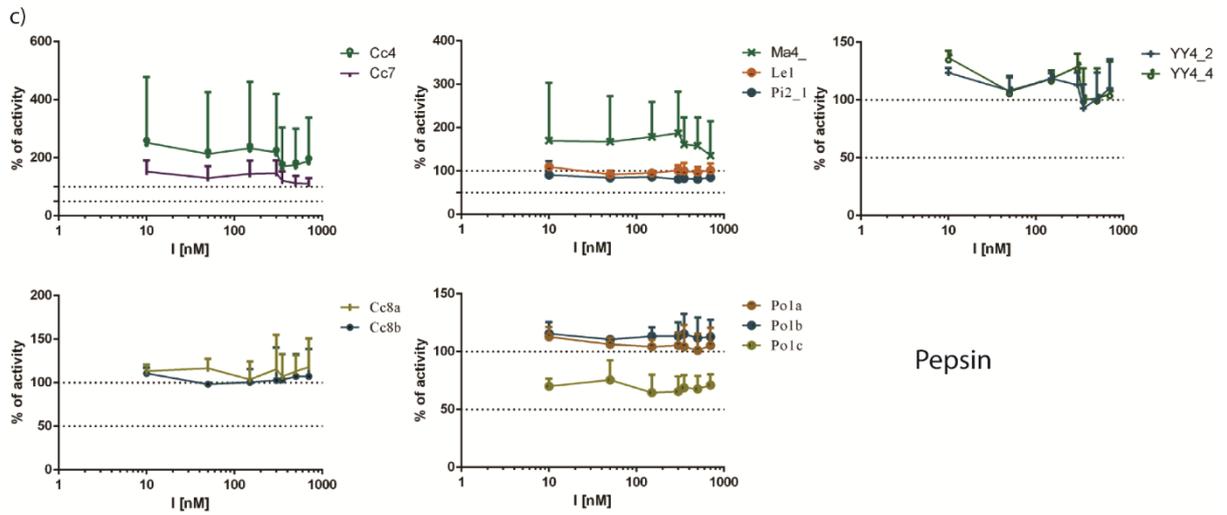
In collaboration with the laboratory of André Fleissner at Technical University of Braunschweig, Germany, the germination rates of the knockout strains compared to the wildtype and the frequency of interactions between germings via the formation of specialized anastomosis tubes were analyzed. There was no significant difference between the *C. cinerea* knockout strains and the wildtype strains in the percentage of germinating oidia and the percentage of interacting oidia (Appendix Fig. 2), suggesting that the respective KEs do not play a role in the chemotropic interactions that mediate polar growth and fusion of compatible hyphae.

## Chapter 2

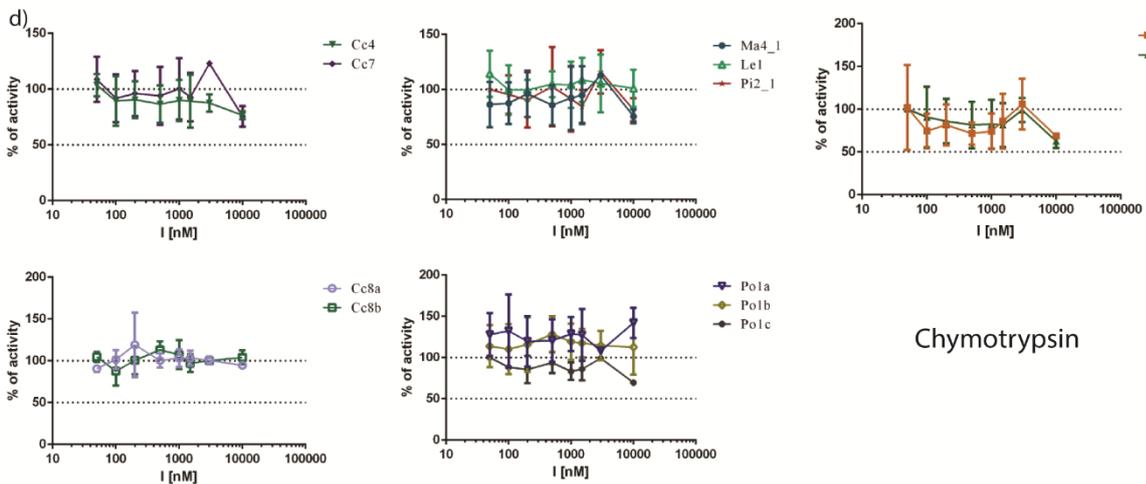
a)	Label	Peptide	Source organism
	Cc Ma4_1	pyro-QVPVDEPA	<i>C. cinerea</i>
	Cc YY4_2	Pyro-QIPAELOPTKSINGNIEPY	<i>C. cinerea</i>
	Cc YY4_4	Pyro-QIPAELOPTKSINGNIEPYK	<i>C. cinerea</i>
	Cc6	Pyro-QSEPKPTN	<i>C. cinerea</i>
	Cc8a	GPPAPSW	<i>C. cinerea</i>
	Cc8b	DDDPEGPPAPSW	<i>C. cinerea</i>
	Le1	SGTGEASAADW	<i>L. edodes</i>
	Po1a	MSGVAADW	<i>P. ostreatus/eryngii</i>
	Po1b	SGVAADW	<i>P. ostreatus/eryngii</i>
	Po1c	LSGVAADW	<i>P. ostreatus</i>
	Pi2_1	DAREPIFGF	Neg. ctrl
	Cc4	Pyro-QIPAELOPTKSINGNIEPYK	Neg. ctrl



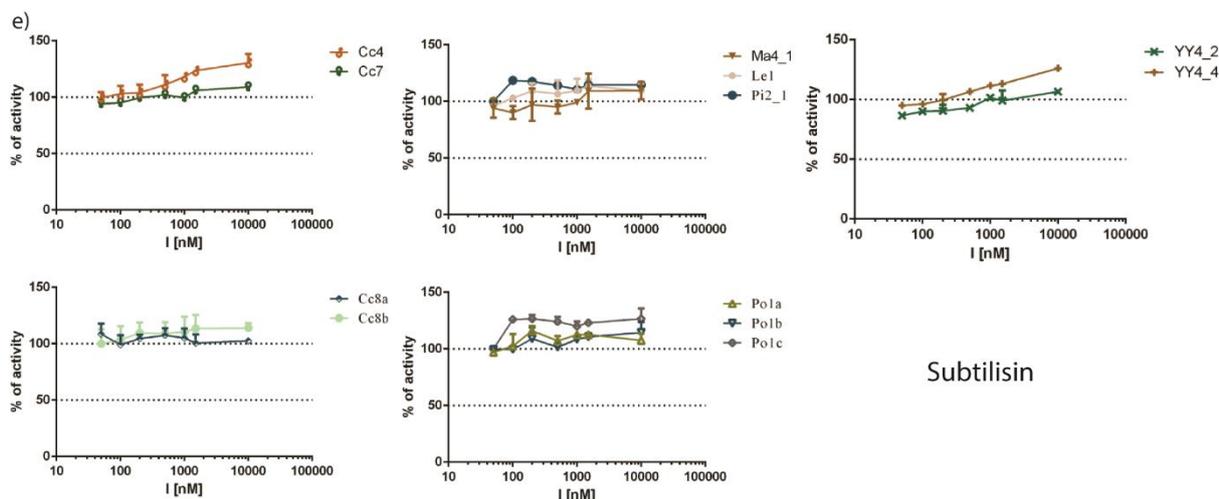
Papain



Pepsin



Chymotrypsin



**Appendix 2 Figure 3:** Protease inhibitor assays with synthetic peptides. a) List of all tested synthetic peptides and their labels. b) Assay of inhibitory activity against the cysteine protease papain, c) the aspartate protease pepsin, d) the serine protease chymotrypsin and e) the serine protease subtilisin. Peptides were tested at concentrations between 10 nM and 1-10  $\mu$ M following the protocol of Brzin et al. 2000 with the fluorogenic substrates Z-Phe-Arg-AMC (in 0.1 M MES buffer, pH 6.5, 5 mM DTT for papain), MOCA DNP (in McIlvaine buffer, pH 3.5 for pepsin) or Suc-AAPF-MCA (in 0.1 M Tris-HCl, pH 8.8 for chymotrypsin and subtilisin). Papain, pepsin, chymotrypsin and subtilisin were tested at concentrations of 6.4, 5.8, 15 and 22 nM, respectively. The experiments were performed by Tanja Zupan, Jožef Stefan Institute.

### A2.3 Protease inhibitor assays with synthetic peptides

Peptides can be potent inhibitors of proteases. They are used as defense effectors in both plants and fungi by inhibiting digestive enzymes in the gastrointestinal tract of herbivores and predators (Ryan 1990; Sabotič and Kos 2017). For example, the fungal protease inhibitors cnispin and cospin are expressed primarily in the fruiting bodies of *Clitocybe nebularis* and *C. cinerea*, respectively, inhibit the protease trypsin and exhibit potent toxicity to *Drosophila melanogaster*. They are thought to play a role in both defense against predators and regulation of endogenous proteases involved in fruiting body development (Sabotič et al. 2007; Avanzo et al. 2009; Sabotič et al. 2012).

We suspected that the KEP-derived peptides that we detected in *C. cinerea*, *L. edodes* and *P. ostreatus/eryngii* might protect the fruiting bodies from predators via a protease inhibitory effect and tested the synthetic peptides in protease inhibition assays in collaboration with the laboratory of Jerica Sabotič at the Jožef Stefan Institute in Ljubljana, Slovenia (Appendix 2 Fig. 3). In these assays, peptide activity was tested against papain, pepsin, chymotrypsin and subtilisin. No effect of the synthetic peptides could be detected.

**Appendix 2 references**

- Avanzo P, Sabotič J, Anžlovar S, Popovič T, Leonardi A, Pain RH, Kos J, Brzin J. 2009. Trypsin-specific inhibitors from the basidiomycete *Clitocybe nebularis* with regulatory and defensive functions. *Microbiology*. 155(12):3971–3981. <https://doi.org/10.1099/mic.0.032805-0>
- Bolz DD, Tenor JL, Aballay A. 2010. A conserved PMK-1/p38 MAPK is required in *Caenorhabditis elegans* tissue-specific immune response to *Yersinia pestis* infection. *J Biol Chem* [Internet]. 285(14):10832–10840. <https://doi.org/10.1074/jbc.M109.091629>
- Brzin J, Rogelj B, Popovič T, Štrukelj B, Ritonja A. 2000. Clitocypin, a new type of cysteine proteinase inhibitor from fruit bodies of mushroom *Clitocybe nebularis*. *J Biol Chem*. 275(26):20104–20109. <https://doi.org/10.1074/jbc.M001392200>
- Fleissner A, Leeder AC, Roca MG, Read ND, Glass NL. 2009. Oscillatory recruitment of signaling proteins to cell tips promotes coordinated behavior during cell fusion. *Proc Natl Acad Sci U S A*. 106(46):19387–19392. <https://doi.org/10.1073/pnas.0907039106>
- Glass NL, Fleissner A. 2006. Re-Wiring the Network: Understanding the Mechanism and Function of Anastomosis in Filamentous Ascomycete Fungi. *Growth, Differ Sex.*:123–139. [https://doi.org/10.1007/3-540-28135-5\\_7](https://doi.org/10.1007/3-540-28135-5_7)
- Grøn H, Meldal M, Breddam K. 1992. Extensive Comparison of the Substrate Preferences of Two Subtilisins As Determined with Peptide Substrates Which Are Based on the Principle of Intramolecular Quenching. *Biochemistry*. 31(26):6011–6018. <https://doi.org/10.1021/bi00141a008>
- Künzler M, Bleuler-Martinez S, Butschi A, Garbani M, Lüthy P, Hengartner MO, Aebi M. 2010. Biotoxicity assays for fruiting body lectins and other cytoplasmic proteins. *Methods Enzymol*. 480(C):141–150. [https://doi.org/10.1016/S0076-6879\(10\)80007-2](https://doi.org/10.1016/S0076-6879(10)80007-2)
- Parks TD, Leuther KK, Howard ED, Johnston SA, Dougherty WG. 1994. Release of proteins and peptides from fusion proteins using a recombinant plant virus proteinase. *Anal Biochem*. 216(2):413–417. <https://doi.org/10.1006/abio.1994.1060>
- Power SD, Adams RM, Wells JA. 1986. Secretion and autoproteolytic maturation of subtilisin. *Proc Natl Acad Sci U S A*. 83(10):3096–3100. <https://doi.org/10.1073/pnas.83.10.3096>
- Ryan CA. 1990. Protease Inhibitors in Plants: Genes for Improving Defenses Against Insects and Pathogens. *Annu Rev Phytopathol*. 28(1):425–449. <https://doi.org/10.1146/annurev.py.28.090190.002233>
- Sabotič J, Bleuler-Martinez S, Renko M, Caglič PA, Kallert S, Štrukelj B, Turk D, Aebi M, Kos J, Künzler M. 2012. Structural basis of trypsin inhibition and entomotoxicity of cospin, serine protease inhibitor involved in defense of *Coprinopsis cinerea* fruiting bodies. *J Biol Chem*. 287(6):3898–3907. <https://doi.org/10.1074/jbc.M111.285304>
- Sabotič J, Kos J. 2017. Fungal Protease Inhibitors. In: Mérillon, JM., Ramawat K, editor. *Fungal Metab*. Cham: Springer International Publishing; p. 853–885. [https://doi.org/10.1007/978-3-319-25001-4\\_10](https://doi.org/10.1007/978-3-319-25001-4_10)
- Sabotič J, Trček T, Popovič T, Brzin J. 2007. Basidiomycetes harbour a hidden treasure of proteolytic diversity. *J Biotechnol*. 128(2):297–307. <https://doi.org/10.1016/j.jbiotec.2006.10.006>

## Chapter 2

Stiernagle T. 2006. Maintenance of *C. elegans*. WormBook.(1999):1–11. <https://doi.org/10.1895/wormbook.1.101.1>

Tayyrov A, Stanley CE, Azevedo S, Künzler M. 2019. Combining microfluidics and RNA-sequencing to assess the inducible defensome of a mushroom against nematodes. BMC Genomics [Internet]. 20(1):243. <https://doi.org/10.1186/s12864-019-5607-3>

## Chapter 3

Genome sequences of *Rhizopogon roseolus*, *Mariannaea elegans*, *Myrothecium verrucaria* and *Sphaerostilbella broomeana* and the identification of biosynthetic gene clusters for fungal peptide natural products

Eva Vogt<sup>1</sup>, Christopher M. Field<sup>1</sup>, Lukas Sonderegger<sup>1</sup>, Markus Künzler<sup>1</sup>

<sup>1</sup> ETH Zürich, Department of Biology, Institute of Microbiology, Vladimir-Prelog-Weg 4, CH-8093 Zürich, Switzerland

Published in:

G3 Genes | Genomes | Genetics, 2022, jkac095, <https://doi.org/10.1093/g3journal/jkac095>

### Contributions:

- gDNA extraction
- antiSMASH genomic screening
- Writing, revisions

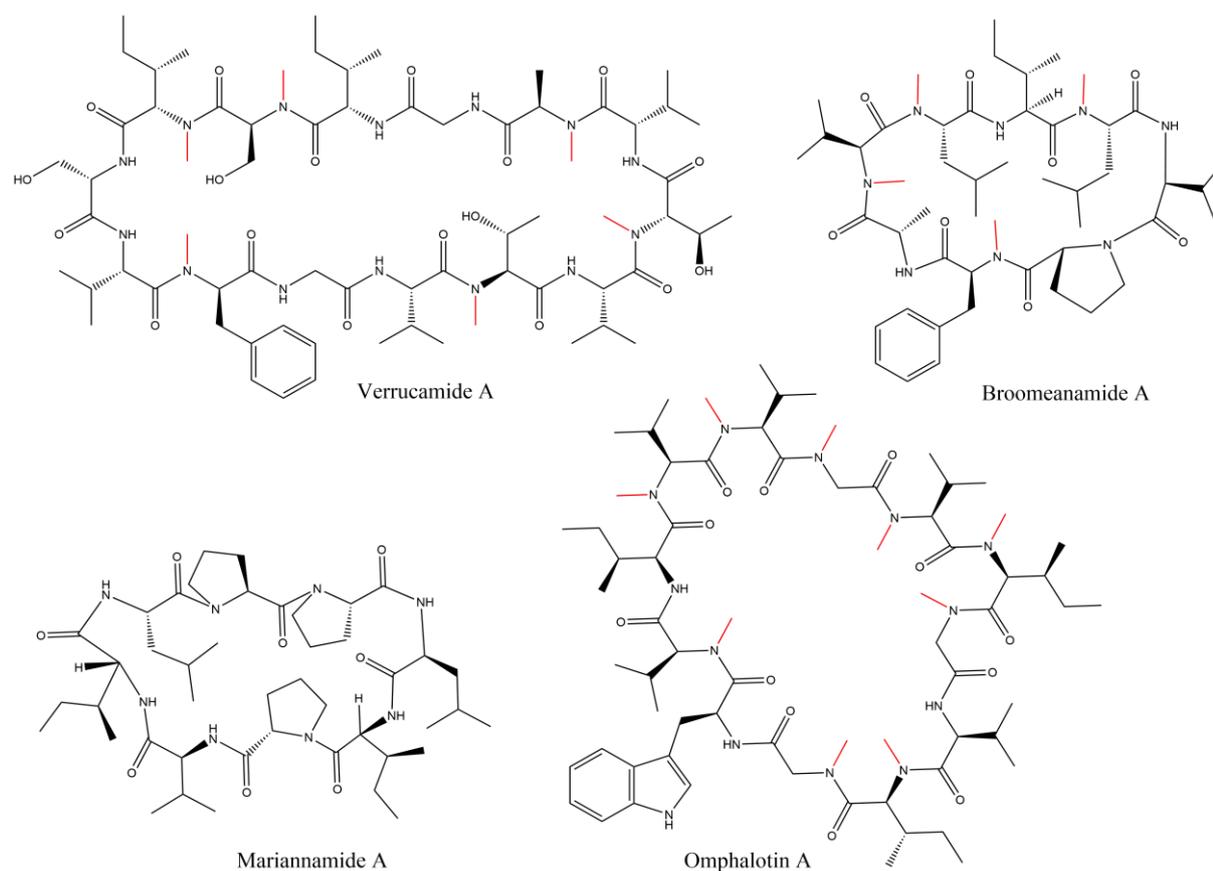
## Abstract

In recent years, a variety of fungal cyclic peptides with interesting bioactivities have been discovered. For many of these peptides, the biosynthetic pathways are unknown and their elucidation often holds surprises. The cyclic and backbone *N*-methylated omphalotins from *Omphalotus olearius* were recently shown to constitute a novel class (borosins) of ribosomally synthesized and posttranslationally modified peptides (RiPP), members of which are produced by many fungi, including species of the genus *Rhizopogon*. Other recently discovered fungal peptide macrocycles include the mariannamides from *Mariannaea elegans* and the backbone *N*-methylated verrucamides and broomeanamides from *Myrothecium verrucaria* and *Sphaerostilbella broomeana*, respectively. Here, we present draft genome sequences of four fungal species *Rhizopogon roseolus*, *Mariannaea elegans*, *Myrothecium verrucaria* and *Sphaerostilbella broomeana*. We screened these genomes for precursor proteins or gene clusters involved in the mariannamide, verrucamide and broomeamide biosynthesis including a general screen for borosin-producing precursor proteins. While our genomic screen for potential RiPP precursor proteins of mariannamides, verrucamides, broomeanamides and borosins remained unsuccessful, antiSMASH predicted non-ribosomal peptide synthase (NRPS) gene clusters that may be responsible for the biosynthesis of mariannamides, verrucamides and broomeanamides. In *M. verrucaria*, our antiSMASH search led to a putative NRPS gene cluster with a predicted peptide product of 20 amino acids, including multiple non-proteinogenic isovalines. This cluster likely encodes a member of the peptaibols, an antimicrobial class of peptides previously isolated primarily from the Genus *Trichoderma*. The NRPS gene clusters discovered in our screenings are promising candidates for future research.

## Introduction

Borosins, a class of backbone *N*-methylated ribosomally synthesized and posttranslationally modified peptides (RiPPs), were defined in 2017 following the discovery of the biosynthesis pathway of the founding member omphalotin A (Ramm et al. 2017; Van Der Velden et al. 2017). This nematotoxic peptide macrocycle and its variants are produced by the fungus *Omphalotus olearius* via the self-modifying precursor protein OphMA. OphMA contains an N-terminal  $\alpha$ N-methyltransferase domain that methylates the precursor's C-terminal core peptide, followed by cleavage, cyclization and release of omphalotin (Van Der Velden et al. 2017). Backbone *N*-methylations were previously found exclusively in nonribosomal peptides and were even considered a hallmark of this type of peptides. Therefore, it was a surprise to find them in RiPPs (Vogt and Künzler 2019). Genome mining led to the discovery of many other potential OphMA-like peptide precursors in fungi, including *Dendrothele bispora* and *Lentinula edodes* (Quijano et al. 2019). The genomes of these fungi contain biosynthetic gene clusters with similar composition and organization as the omphalotin cluster. In addition, the encoded OphMA homologs contain core peptide with high sequence similarity to omphalotin A. Analysis of fungal tissue samples confirmed the production of the corresponding peptides, termed dendrothelins and lentinulins (Matabaro et al. 2021). Recent publications demonstrated the presence of borosin clusters with trans-acting  $\alpha$ N-methyltransferases in

bacteria (Cho et al. 2022; Imani et al. 2022). Based on these findings, we were interested in investigating previously discovered, backbone *N*-methylated cyclic peptides that were hypothesized to be of non-ribosomal origin, to represent novel members of the borosin class of RiPPs.



**Figure 1:** Structures of the cyclic fungal peptides verrucamide A (Zou et al. 2011), broomeanamide A (Ekanayake et al. 2021), mariannamide A (Ishiuchi et al. 2020) and omphalotin A (Van Der Velden et al. 2017, the founding member of the peptide class of borosins. Backbone *N*-methylations are colored red.

Recently discovered cyclic, backbone *N*-methylated peptides include verrucamides A-D, tetradecapeptides that are produced by the ascomycete *Myrothecium verrucaria* and contain two D-configured amino acids (Zou et al. 2011), and the octapeptides broomeanamides A-C from the mycoparasitic ascomycete *Sphaerostilbella broomeana* where all eight amino acids are L-configured (Fig. 1) (Ekanayake et al. 2021). Another class of cyclic peptides are the octapeptides mariannamides A and B isolated from the filamentous ascomycete *Mariannaea elegans* that are also composed of all L-amino acids amongst three proline residues but do not contain any backbone *N*-methylations (Fig. 1) (Ishiuchi et al. 2020). Both verrucamides and mariannamides were shown to possess antibacterial properties (Zou et al. 2011; Ishiuchi et al. 2020). The mode of synthesis of all three peptide classes is unknown; the structural similarity of the verrucamides and broomeanamides to the cyclic, backbone *N*-methylated borosins indicated that they may be RiPPs, although the presence of D-amino acids in the verrucamides rather suggested a non-ribosomal origin. Only one fungal RiPP class with a residue in D-configuration has been identified so far (phallotoxins, Hallen et al. 2007).

Here we report the genome sequences of *M. verrucaria*, *M. elegans*, *R. roseolus* and *S. broomeana*. We mined the genomes of *M. verrucaria*, *M. elegans* and *S. broomeana* for potential RiPP precursor proteins of the verrucamides, mariannamides and broomeanamides, respectively. In addition, we performed an antiSMASH search to screen for non-ribosomal peptide (NRP) biosynthetic gene clusters that might encode genes for verrucamide, mariannamide and broomeanamide synthesis. We sequenced the genome of the agaricomycete *Rhizopogon roseolus*, as the genomes of two species of the genus *Rhizopogon* were shown in BLAST searches to encode multiple OphMA homologs each (Quijano et al. 2019). Finally, we performed screens to find new OphMA homologs in *R. roseolus*, *M. verrucaria*, *M. elegans* and *S. broomeana*.

### Methods and Materials

#### *Strains and cultivation*

The sequenced strains of *M. verrucaria*, *M. elegans* and *S. broomeana* are the authentic producers of the verrucamides, mariannamides and broomeanamides as seen in Zou et al. 2011, Ishiuchi et al. 2020 and Ekanayake et al. 2021. *Myrothecium verrucaria* DSM 2087 was received by the Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures, Germany, *Mariannaea elegans* NBRC102301 was ordered from the Biological Resource Center, National Institute of Technology and Evaluation (NITE) in Kisarazu, Japan, heterokaryotic *Rhizopogon roseolus* Mykothek Nr 97.03 (CBS# pending) was received from Martina Peter from the Eidgenössische Forschungsanstalt für Wald, Schnee und Landschaft WSL, Switzerland. *Sphaerostilbella broomeana* TFC201724 is deposited at the Tartu Fungal Culture Collection (TFC) of the University of Tartu, Estonia, and was received by Kadri Põldmaa from the University of Tartu. *M. verrucaria* was cultivated on Corn Meal Agar (CMA) at 30°C, *M. elegans* NBRC102301 on Potato Dextrose Agar (PDA) at room temperature, *R. roseolus* on Yeast Malt Agar (YMG) at room temperature, and *S. broomeana* on PDA at room temperature.

#### *Sample preparation and sequencing*

The fungi were cultivated on cellophane-covered agar plates before their mycelia were harvested. *M. verrucaria*, *M. elegans* and *R. roseolus* mycelia were harvested after 14, 9 and 40 days, respectively, and lysed by grinding with a mortar and pestle in the presence of liquid nitrogen. *S. broomeana* mycelium was harvested after 7 days, mixed in an Eppendorf tube with 0.5 mm glass beads, frozen in liquid nitrogen and then lysed by vigorous shaking in a Fastprep machine for two times 45 seconds at level 6. Genomic DNA was extracted using the QIAGEN DNeasy plant Mini kit, DNA concentration measured using a Qubit dsDNA kit and DNA quality confirmed by running a fraction of the DNA on an agarose gel. The DNA was sent to Novogene, United Kingdom, for shotgun sequencing on an Illumina Novaseq, producing paired-end 150 bp reads, aiming for approximately 100x coverage.

#### *Quality Control*

BBDuk (v38.87, Joint Genome Institute) was first used in right-trimming mode with a kmer length of 23 down to 11 and a hamming distance of 1 to filter out sequencing adapters. A

## Chapter 3

second pass with a kmer length of 31 and a hamming distance of 1 was used to filter out PhiX sequences. A third and final pass performed quality trimming on both read ends with a Phred score cutoff of 14 and an average quality score cutoff of 20, with reads under 45 bp or containing Ns subsequently rejected.

### *Assembly*

The paired-end and singleton reads of each read set were assembled using SPAdes (v3.14.0) (Nurk et al. 2013) in isolate mode, but otherwise default parameters.

### *Gene calling*

GlimmerHMM (v3.0.4) (Majoros et al. 2004) was trained on exon sequences taken from phylogenetically close reference genomes (*Mariannaea sp.* strain PMI\_226 v1.0 for *M. elegans*; *Myrothecium inundatum* CBS 120646 v1.0 for *M. verrucaria*; *Rhizopogon vulgaris* FC72 v1.0 for *R. roseolus*; *Trichoderma reesei* QM6a NW\_006711148.1 for *S. broomeana*) and subsequently used to call genes in the genome assemblies. All reference genomes were obtained from JGI (Nordberg et al. 2014) in August 2020 (*M. elegans*, *M. verrucaria*, *R. roseolus*) and October 2021 (*S. broomeana*). Average nucleotide identity (ANI) between assemblies and reference genomes were performed with FastANI (Jain et al. 2018) that implements a similar calculation to Goris et al. (Goris et al. 2007).

### *Quality assessment*

Completeness of the genome assemblies was assessed using BUSCO (v5.0.0) (Simão et al. 2015) in genome mode with the --auto-lineage-euk parameter to automatically assess the likely lineage of each strain (*M. elegans*, *M. verrucaria*, *S. broomeana*: *Hypocreales*; *R. roseolus*: *Boletales*). To test for bacterial contamination, the tool mOTUs (v3.0.0) (Milanese et al. 2019) was run on the reads for each sample. *M. elegans* had 2 inserts that could not be assigned to a specific mOTU; *M. verrucaria* had 1 insert corresponding to "*Phyllobacterium species incertae sedis*"; *R. roseolus* and *S. broomeana* returned no hits. These very low hit counts indicate that it is very unlikely for there to be any contamination by bacteria in the samples.

### *Taxonomic analysis*

The *ssu\_finder* function of CheckM (v1.0.13) (Parks et al. 2015) was used to extract 16S and 18S rRNA gene sequences from the assemblies. 18S sequences of 1726 bp, 1725 bp and 1726 bp were found for *M. elegans*, *M. verrucaria* and *S. broomeana* respectively, but no such sequence was found for *R. roseolus*, likely because its assembly was highly fragmented. The sequences were aligned with the SILVA taxonomy database (v138) (Quast et al. 2013) using the provided software SINA (v1.6.1) (Pruesse et al. 2012). The internal transcribed spacer (ITS) region of each strain was extracted from its assembly using ITSx (Bengtsson-Palme et al. 2013) and its Fungi profile set. The sequences were then analysed with the UNITE database (Nilsson et al. 2019; Kõljalg et al. 2020).

### *Genome mining*

To search for all possible arrangements of cyclic peptides of interest, a custom Python script generated a fasta file containing all possible variants of linearized peptide sequences for the verrucamides, mariannamides and broomeanamides. For broomeanamide A, for example, these sequences would be VPFVLIL, PFAVLILV, FAVLILVP etc. First the predicted protein sequences were searched for all peptides with *blastp*, then the assemblies were searched for all peptides with *tblastn*, both part of the BLAST+ suite (v2.11.0) (Camacho et al. 2009). As a positive control for the functionality of our mining method, we screened the genome of *Omphalotus olearius* using the peptide sequence of the cyclic RiPP omphalotin (Ramm et al. 2017; Van Der Velden et al. 2017). We found the omphalotin precursor protein OphMA, thus confirming that our method works. The 300 residue long N-terminal methyltransferase domain of the protein OphMA from *Omphalotus olearius* (Quijano et al. 2019) was searched for in the predicted protein sequences with *blastp*. Further, all assemblies were analysed with the fungal version of antiSMASH (v5.1.0) (Blin et al. 2019), which ignores contigs of less than 1 kbp in length by default, to look for biosynthetic gene clusters.

## Results and Discussion

### *Genome assembly and completeness*

The genomes of the fungi *Myrothecium verrucaria*, *Mariannaea elegans*, *Rhizopogon roseolus* and *Sphaerostilbella broomeana* were sequenced using Illumina PE150 and assembled using the reference genomes *Myrothecium inundatum* CBS 120646 v1.0, *Mariannaea sp.* strain PMI\_226 v1.0, *Rhizopogon vulgaris* FC72 v1.0 and *Trichoderma reesei* QM6a. The *M. verrucaria* assembly had the highest quality of the four assemblies with the lowest contig count of 3'197 and the highest N50 value of 1'066'851, while *M. elegans*, *R. roseolus* and *S. broomeana* had contig counts of 3'718, 23'000 and 3'882, respectively, and N50 values of 288'335, 61'999 and 295'130 (Table 1). As an alternative assessment of genome assembly and annotation completeness, the open-source software Benchmarking Universal Single-Copy Orthologs (BUSCO) was used (Simão et al. 2015). *M. verrucaria* was 98.0% complete as a Eukaryote, while *M. elegans*, *R. roseolus* and *S. broomeana* were 99.2%, 95.7% and 98.0% complete, respectively. Completeness for the order *Hypocreales* (for *M. verrucaria*, *M. elegans* and *S. broomeana*) and *Boletales* (for *R. roesolus*) was 96.6%, 97.4%, 97.5% and 95.1%, respectively. Potential 16S rRNA sequences were extracted from the assembly to confirm the absence of bacterial contamination.

*M. verrucaria*, *M. elegans* and *R. roseolus* had an average nucleotide identity (ANI) of 78.8%, 79.9% and 90.2% with their reference fungi *Myrothecium inundatum*, *Mariannaea sp.* strain PMI\_226 and *Rhizopogon vulgaris* (Table 1). No ANI value could be generated for *S. broomeana* and its reference strain *Trichoderma reesei*, meaning that sequence similarity between the two strains is too low (<70%). For a deeper taxonomical analysis of the sequenced species, the 18S rRNA sequences were extracted and ran against the SILVA ribosomal RNA taxonomy database (Quast et al. 2013). *M. verrucaria* was identified as a member of the genus *Myrothecium*, *M. elegans* as a member of the order *Hypocreales*, *R.*

*roseolus* as a member of either the order *Boletales* or *Agaricales* and *S. broomeana* as a member of the genus *Trichoderma*.

		<i>Myrothecium verrucaria</i>	<i>Mariannaea elegans</i>	<i>Rhizopogon roseolus</i>	<i>Sphaerostilbella broomeana</i>
All scaffolds	Count	3'197	3'718	23'000	3'882
	Length	46'297'313	52'632'238	37'675'430	36'266'877
	N50	1'066'851	288'335	61'999	295'130
	N90	2'846'945	1'147'592	169'220	800'567
	Max	4'080'732	2'068'507	547'319	1'217'981
Scaffolds >= 1 kbp	Count	434	1'098	1'543	1'121
	Length	45'599'860	51'857'879	33'959'226	35'139'140
	N50	1'066'851	294'018	70'907	311'573
	N90	2'846'945	1'147'592	173'881	800'567
	Max	4'080'732	2'068'507	547'319	1'217'981
BUSCO	Completeness (Eukaryotes)	98.0%	99.2%	95.7%	98.0% (Eukaryotes)
	Completeness ( <i>Hypocreales</i> )	96.6%	97.4%	95.1% ( <i>Boletales</i> )	97.5% ( <i>Hypocreales</i> )
	Single-copy	95.5%	96.5%	93.8%	97.3%
	Duplicated	1.1%	0.9%	1.3%	0.2%
	Fragmented	0.4%	0.3%	0.6%	0.2%
	Missing	3.0%	2.3%	4.3%	2.3%
	Number of searched genes	4'494	4'494	4'878	4'494
Average nucleotide identity (ANI)	78.8% ( <i>Myrothecium inundatum</i> )	79.9% ( <i>Mariannaea sp.</i> )	90.2% ( <i>Rhizopogon vulgaris</i> )	<70% ( <i>Trichoderma reesei</i> )	

**Table 1:** Summary of the assembled genomes of the four newly sequenced fungal species *Myrothecium verrucaria*, *Mariannaea elegans*, *Rhizopogon roseolus* and *Sphaerostilbella broomeana*. Separate values are given for all scaffolds and scaffolds with a size of 1 kbp or more. Given parameters are the scaffold count, scaffold length, N50 and N90 values, and maximum scaffold length. The N50 and N90 values describe assembly contiguity by giving the minimal contig size that, together with all larger contigs, covers 50% or 90% of the total genome, respectively. BUSCO values describe the assembly completeness compared to Eukaryotes or the orders *Hypocreales* or *Boletales*. Average nucleotide identity describes nucleotide similarity to the reference genomes.

#### Screenings for RiPP precursors and NRP clusters

All four genomes were screened for potential RiPP precursor proteins; *M. verrucaria* for verrucamide precursors, *M. elegans* for mariannamide precursors, *S. broomeana* for broomeanamide precursors and all four genomes, including *R. roseolus*, for OphMA homologs. Screens were performed using all circular permutations of verrucamide, mariannamide and broomeanamide sequences. In addition, all genomes were screened for the N-terminal methyltransferase domain of OphMA. These searches yielded no hits, indicating that the cyclic backbone *N*-methylated verrucamides and broomeanamides and the cyclic mariannamides

### Chapter 3

are not genetically encoded and therefore may indeed be NRPs, and that *R. roseolus*, unlike many of its relatives from the genus *Rhizopogon*, does not contain any OphMA homologs.

<b>Myrothecium verrucaria NRPS 2.4</b>	
Hit region	NODE_2_length_2846945_cov_41.963601 - Region 4 - NRPS
<b>Legend:</b> <span style="color:red">■</span> core biosynthetic genes <span style="color:red">■</span> additional biosynthetic genes <span style="color:blue">■</span> transport-related genes <span style="color:green">■</span> regulatory genes <span style="color:grey">■</span> other genes <span style="border:1px solid grey; padding: 2px;">■</span> resistance	
NRP predicted seq.	A??V?I?????IF
Verrucamide A	<u>A</u> GISISV <u>F</u> GVTVTV
Verrucamide B	<u>A</u> GVSISV <u>F</u> GVTVTV
Verrucamide C	<u>A</u> GISVSV <u>F</u> GVTVTV
Verrucamide D	<u>A</u> GVSVSV <u>F</u> GVTVTV

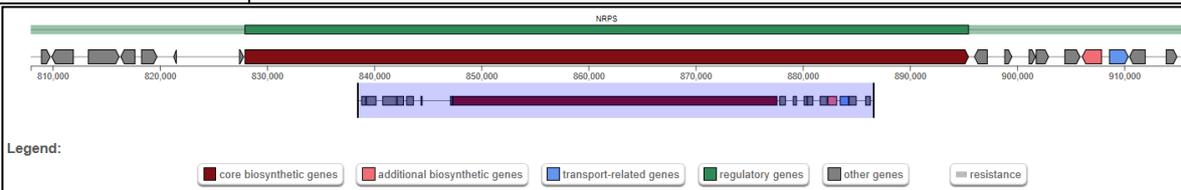
<b>Mariannaea elegans NRPS 180.1</b>	
Hit region	NODE_180_length_54293_cov_43.947949 - Region 1 - NRPS
<b>Legend:</b> <span style="color:red">■</span> core biosynthetic genes <span style="color:red">■</span> additional biosynthetic genes <span style="color:blue">■</span> transport-related genes <span style="color:green">■</span> regulatory genes <span style="color:grey">■</span> other genes <span style="border:1px solid grey; padding: 2px;">■</span> resistance	
NRP predicted seq.	LL??LP?L
Mariannamide A	IPVILPPL
Mariannamide B	IPVVL PPL

<b>Sphaerostilbella broomeana NRPS 77.1</b>	
Hit region	NODE_77_length_123022_cov_43.474692.gene39 - Region 1 - NRPS
<b>Legend:</b> <span style="color:red">■</span> core biosynthetic genes <span style="color:red">■</span> additional biosynthetic genes <span style="color:blue">■</span> transport-related genes <span style="color:green">■</span> regulatory genes <span style="color:grey">■</span> other genes <span style="border:1px solid grey; padding: 2px;">■</span> resistance	
NRP predicted seq.	???I?L
Broomeanamide A	VPF <u>F</u> AVLIL
Broomeanamide B	VPF <u>F</u> AVLVL

**Table 2:** *M. verrucaria*, *M. elegans* and *S. broomeana* possess candidate NRP biosynthetic gene clusters for the synthesis of verrucamide-like, mariannamide-like and broomeanamide-like peptides. A comparison between the predicted product sequence and the known sequences of verrucamide A-D, mariannamide A-B and broomeanamide A-B is given (Zou et al. 2011; Ishiuchi et al. 2020). Red letters indicate *N*-methylated residues, underlined letters represent D-amino acids.

Following the unsuccessful search for RiPP precursors of verrucamides, mariannamides and broomeanamides, an additional search was performed using the fungal version of the “antibiotics and secondary metabolite analysis shell” antiSMASH (Blin et al. 2019) with the goal of finding NRP biosynthetic gene clusters that might direct the biosynthesis of the isolated peptide natural products. antiSMASH currently uses an ensemble prediction method integrating several algorithms to predict the substrate specificity of adenylation domains (Blin et al. 2017). In *M. verrucaria*, one NRP cluster was predicted to produce a verrucamide-like peptide with the correct length and several *N*-methylated residues, whereas in *M. elegans*, one cluster was predicted to produce a peptide of the same length as the mariannamides, containing several leucines and at least one proline (Table 2). In *S. broomeana*, cluster NRPS 77.1 was predicted to produce a six-residue-long peptide containing one isoleucine, one leucine and a total of four *N*-methylations. The broomeanamides are longer (eight residues), but contain four *N*-methylated residues, two leucines and, in the case of Broomeanamide A, one isoleucine (Table 2).

<b>Myrothecium verrucaria NRPS 13.3</b>	
Hit region	NODE_13_length_1066851_cov_42.619387- Region 3 - NRPS. Location: 807,950 - 915 464 nt
	
Predicted sequence	XPXXPXXP?X??XX?XXX??
Trichovirin II 6b ( <i>Trichoderma viride</i> )	AcZGALXQXVZGXZPLZZQLeuol

**Table 3:** Predicted NRP cluster in *M. verrucaria* encoding a putative peptide of the peptaibol class, containing the non-proteinogenic amino acid isovaline (X). The peptaibol Trichovirin from *T. viride* is cited as an example (Jaworski et al. 1999). The letters X and Z stand for isovaline and  $\alpha$ -aminoisobutyric acid, respectively.

Another NRPS gene cluster of *M. verrucaria* was predicted to encode a 20 residue peptide with eleven residues of the non-proteinogenic amino acid isovaline (Table 3). This peptide is likely a peptaibol. Peptaibols are a class of antimicrobial NRPs from fungi that are 5–20 residues long, linear, N- and C-terminally modified with amino alcohol groups, and defined by the presence of the non-proteinogenic amino acids  $\alpha$ -aminoisobutyric acid (Y) and/or isovaline (X) (Fuente-Núñez et al. 2013). The predicted peptide from *M. verrucaria* does not contain  $\alpha$ -aminoisobutyric acid, but eleven residues of isovaline. There are no known peptaibols with such a high content of isovaline, so it is likely that some of these predicted isovalines are rather alpha-aminoisobutyric acids or other residues. To date, over 1000 peptaibols have been characterized in various members of the Order *Hypocreales*, with the vast majority produced by members of the Genus *Trichoderma* (Fuente-Núñez et al. 2013). Our antiSMASH search suggested two additional isovaline-containing NRPs in *Myrothecium verrucaria* (NRPS 3.4, X?X?L?Q?X and NRPS 15.2, XQ?X???) and one in *Sphaerostilbella broomeana* (NRPS 31.1,

XXX?X?QX??). Peptaibols have been previously reported in *Sphaerostilbella toxica* (Perlatti et al. 2020), but to our knowledge no peptaibols have been described in the Genus *Myrothecium*.

In conclusion, we present the complete genome sequences of the fungi *Myrothecium verrucaria*, *Mariannaea elegans*, *Rhizopogon roseolus* and *Sphaerostilbella broomeana*. While our screens of the genomes for genes encoding RiPP precursor proteins of verrucamides, mariannamides, broomeanamides and borosins did not yield any hits, we discovered three NRPS candidate clusters that may control verrucamide, mariannamide and broomeanamide biosynthesis, as well as multiple clusters predicted to produce peptaibol-like peptides. These clusters will be interesting targets for future research, particularly with regard to the antibacterial properties of verrucamides, mariannamides and peptaibols (Zou et al. 2011; Fuente-Núñez et al. 2013; Ishiuchi et al. 2020).

### **Data availability**

All relevant data was submitted to ENA with the study accession number PRJEB50709 (secondary accession number ERP135330). The accession numbers of samples, raw reads and unannotated assemblies are available in the supplementary table 1. The antiSMASH annotations are available on github under [https://microbiologyethz.github.io/vogt\\_fungal\\_genomes\\_2022/](https://microbiologyethz.github.io/vogt_fungal_genomes_2022/).

### **Funding**

This work was supported by the Swiss National Science Foundation (Grant No. 31003A-173097) and ETH Zürich.

### **Acknowledgments**

We thank Alexander Brachmann (ETH Zürich, Switzerland) for the discussion of the NRP biosynthetic gene cluster predictions.

### **Conflict of interest**

The authors declare that they have no conflict of interest.

## References

- Bengtsson-Palme J, Ryberg M, Hartmann M, Branco S, Wang Z, Godhe A, De Wit P, Sánchez-García M, Ebersberger I, de Sousa F, et al. 2013. Improved software detection and extraction of ITS1 and ITS2 from ribosomal ITS sequences of fungi and other eukaryotes for analysis of environmental sequencing data. *Methods Ecol Evol.* 4(10):914–919. doi:10.1111/2041-210X.12073.
- Blin K, Shaw S, Steinke K, Villebro R, Ziemert N, Lee SY, Medema MH, Weber T. 2019. AntiSMASH 5.0: Updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Res.* 47(W1):W81–W87. doi:10.1093/nar/gkz310.
- Blin K, Wolf T, Chevrette MG, Lu X, Schwalen CJ, Kautsar SA, Suarez Duran HG, De Los Santos ELC, Kim HU, Nave M, et al. 2017. AntiSMASH 4.0 - improvements in chemistry prediction and gene cluster boundary identification. *Nucleic Acids Res.* 45(W1):W36–W41. doi:10.1093/nar/gkx319.
- Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. BLAST+: Architecture and applications. *BMC Bioinformatics.* 10:1–9. doi:10.1186/1471-2105-10-421.
- Cho H, Lee H, Hong K, Chung H, Song I, Lee JS, Kim S. 2022. Bioinformatic Expansion of Borosins Uncovers Trans-Acting Peptide Backbone N-Methyltransferases in Bacteria. *Biochemistry.* 61(3):183–194. doi:10.1021/acs.biochem.1c00764.
- Ekanayake DI, Perlatti B, Swenson DC, Põldmaa K, Bills GF, Gloer JB. 2021. Broomeanamides: Cyclic Octapeptides from an Isolate of the Fungicolous Ascomycete *Sphaerostilbella broomeana* from India. *J Nat Prod.* 84(7):2028–2034. doi:10.1021/acs.jnatprod.1c00414.
- Fuente-Núñez C de la, Whitmore L, Wallace BA. 2013. *Peptaibols*. Second Edi. Elsevier Inc. <http://dx.doi.org/10.1016/B978-0-12-385095-9.00022-1>.
- Goris J, Konstantinidis KT, Klappenbach JA, Coenye T, Vandamme P, Tiedje JM. 2007. DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. *Int J Syst Evol Microbiol.* 57(1):81–91. doi:10.1099/ijs.0.64483-0.
- Hallen HE, Luo H, Scott-Craig JS, Walton JD. 2007. Gene family encoding the major toxins of lethal *Amanita* mushrooms. *Proc Natl Acad Sci.* 104(48):19097–19101. doi:10.1073/pnas.0707340104. <http://www.pnas.org/cgi/doi/10.1073/pnas.0707340104>.
- Imani AS, Lee AR, Vishwanathan N, de Waal F, Freeman MF. 2022. Diverse Protein Architectures and  $\alpha$ -N-Methylation Patterns Define Split Borosin RiPP Biosynthetic Gene Clusters. *ACS Chem Biol.* doi:10.1021/acscchembio.1c01002.
- Ishiuchi K, Hirose D, Kondo T, Watanabe K, Terasaka K, Makino T. 2020. Mariannamides A and B, new cyclic octapeptides isolated from *Mariannaea elegans* NBRC102301. *Bioorganic Med Chem Lett.* 30(4):126946. doi:10.1016/j.bmcl.2019.126946.
- Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S. 2018. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat Commun.* 9(1):1–8. doi:10.1038/s41467-018-07641-9.
- Jaworski A, Kirschbaum J, Brückner H. 1999. Structures of trichovirins II, peptaibol antibiotics from the mold *Trichoderma viride* NRRL 5243. *J Pept Sci.* 5(8):341–351.

### Chapter 3

doi:10.1002/(SICI)1099-1387(199908)5:8<341::AID-PSC204>3.0.CO;2-0.

Köljalg U, Nilsson HR, Schigel D, Tedersoo L, Larsson KH, May TW, Taylor AFS, Jeppesen TS, Frøslev TG, Lindahl BD, et al. 2020. The taxon hypothesis paradigm—On the unambiguous detection and communication of taxa. *Microorganisms*. 8(12):1–24. doi:10.3390/microorganisms8121910.

Majoros WH, Pertea M, Salzberg SL. 2004. TigrScan and GlimmerHMM: Two open source ab initio eukaryotic gene-finders. *Bioinformatics*. 20(16):2878–2879. doi:10.1093/bioinformatics/bth315.

Matabaro E, Kaspar H, Dahlin P, Bader DLV, Murar CE, Staubli F, Field CM, Bode JW, Künzler M. 2021. Identification, heterologous production and bioactivity of lentinulin A and dendrothelin A, two natural variants of backbone N-methylated peptide macrocycle omphalotin A. *Sci Rep*. 11(1):1–12. doi:10.1038/s41598-021-83106-2.

Milanese A, Mende DR, Paoli L, Salazar G, Ruscheweyh HJ, Cuenca M, Hingamp P, Alves R, Costea PI, Coelho LP, et al. 2019. Microbial abundance, activity and population genomic profiling with mOTUs2. *Nat Commun*. 10(1). doi:10.1038/s41467-019-08844-4.

Nilsson RH, Larsson KH, Taylor AFS, Bengtsson-Palme J, Jeppesen TS, Schigel D, Kennedy P, Picard K, Glöckner FO, Tedersoo L, et al. 2019. The UNITE database for molecular identification of fungi: Handling dark taxa and parallel taxonomic classifications. *Nucleic Acids Res*. 47(D1):D259–D264. doi:10.1093/nar/gky1022.

Nordberg H, Cantor M, Dusheyko S, Hua S, Poliakov A, Shabalov I, Smirnova T, Grigoriev I V., Dubchak I. 2014. The genome portal of the Department of Energy Joint Genome Institute: 2014 updates. *Nucleic Acids Res*. 42(D1):26–31. doi:10.1093/nar/gkt1069.

Nurk S, Bankevich A, Antipov D, Gurevich AA, Korobeynikov A, Lapidus A, Prjibelski AD, Pyskhin A, Sirotkin A, Sirotkin Y, et al. 2013. Assembling single-cell genomes and mini-metagenomes from chimeric MDA products. *J Comput Biol*. 20(10):714–737. doi:10.1089/cmb.2013.0084.

Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. CheckM: Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res*. 25(7):1043–1055. doi:10.1101/gr.186072.114.

Perlatti B, Nichols CB, Andrew Alspaugh J, Gloer JB, Bills GF. 2020. Sphaerostilbellins, new antimicrobial aminolipopeptide peptaibiotics from *sphaerostilbella toxica*. *Biomolecules*. 10(10):1–15. doi:10.3390/biom10101371.

Pruesse E, Peplies J, Glöckner FO. 2012. SINA: Accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics*. 28(14):1823–1829. doi:10.1093/bioinformatics/bts252.

Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO. 2013. The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res*. 41(D1):590–596. doi:10.1093/nar/gks1219.

Quijano MR, Zach C, Miller FS, Lee AR, Imani AS, Künzler M, Freeman MF. 2019. Distinct Autocatalytic  $\alpha$ -N-Methylating Precursors Expand the Borosin RiPP Family of Peptide Natural Products. *J Am Chem Soc*. 141(24):9637–9644. doi:10.1021/jacs.9b03690.

Ramm S, Krawczyk B, Mühlenweg A, Poch A, Mösker E, Süßmuth RD. 2017. A Self-Sacrificing

### Chapter 3

N-Methyltransferase Is the Precursor of the Fungal Natural Product Omphalotin. *Angew Chemie - Int Ed.* 56(33):9994–9997. doi:10.1002/anie.201703488.

Simão FA, Waterhouse RM, Ioannidis P, Kriventseva E V., Zdobnov EM. 2015. BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics.* 31(19):3210–3212. doi:10.1093/bioinformatics/btv351.

Van Der Velden NS, Kälin N, Helf MJ, Piel J, Freeman MF, Künzler M. 2017. Autocatalytic backbone N-methylation in a family of ribosomal peptide natural products. *Nat Chem Biol.* 13(8):833–835. doi:10.1038/nchembio.2393.

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

Zou X, Niu S, Ren J, Li E, Liu X, Che Y. 2011. Verrucamides A-D, antibacterial cyclopeptides from *Myrothecium verrucaria*. *J Nat Prod.* 74(5):1111–1116. doi:10.1021/np200050r.

Supplementary information

Genome sequences of *Rhizopogon roseolus*, *Mariannaea elegans*, *Myrothecium verrucaria* and *Sphaerostilbella broomeana* and the identification of biosynthetic gene clusters for fungal peptide natural products

Eva Vogt<sup>1</sup>, Christopher M. Field<sup>1</sup>, Lukas Sonderegger<sup>1</sup>, Markus Künzler<sup>1</sup>

<sup>1</sup> ETH Zürich, Department of Biology, Institute of Microbiology, Vladimir-Prelog-Weg 4, CH-8093 Zürich, Switzerland

### Chapter 3

Samples			Reads*		Unannotated assemblies
Name	Accession	Biosample	Accession	Experiment	Accession
<i>Myrothecium verrucaria</i> DSM 2087	ERS10521321	SAMEA12922176	ERR8520425 ERR8520426	ERX8086904 ERX8086903	ERZ5159763
<i>Mariannaea elegans</i> NBRC102301	ERS10521322	SAMEA12922177	ERR8520427	ERX8086902	ERZ5159777
<i>Rhizopogon roseolus</i> Mykothek Nr 97.03	ERS10521323	SAMEA12922178	ERR8520428 ERR8520429	ERX8086901 ERX8086900	ERZ5159795
<i>Sphaerostilbella broomeana</i> TFC201724	ERS10521324	SAMEA12922179	ERR8520430	ERX8086899	ERZ5159804

**Table S1:** The accession numbers of samples, raw reads and unannotated assemblies of all four sequenced fungal genomes (Reads\*: two samples were split over multiple flowcells).

## **Chapter 4**

Concluding remarks and future perspectives

### **Concluding remarks and future perspectives**

The aim of this thesis was to gain a better understanding of the role of ribosomal peptides in fungi, with particular emphasis on peptides derived from KEX2-processed repeat proteins (KEPs) and the unknown biosynthetic pathways of novel peptide macrocycles, the mariannamides, verrucamides and broomeanamides. The following chapter summarizes the main findings of this thesis and discusses implications for future research.

### **Strategies for conducting genomic screenings for KEPs**

While the first KEP-derived peptide was studied in the 1980s ( $\alpha$ -type mating pheromones of *S. cerevisiae*, Julius et al. 1983), KEPs were not formally defined and recognized as widely distributed in almost all fungal species until 2019 (Le Marquer et al. 2019). In 2020, another publication expanded the list of known KEPs, noting that while many KEPs have homologs and can be classified into KEP types, many others are unique (Umemura 2020). We performed our own genomic screening in *C. cinerea* in 2018 using a different strategy than Le Marquer et al. 2019 and Umemura 2020, resulting in the discovery of 22 KEP candidates in *C. cinerea* (Chapter 2, Table S1). Our screening missed 9 potential KEPs found in published KEP screenings, while 14 new KEPs were discovered, increasing the number of KEP candidates from 17 to 31 and suggesting that many more KEP candidates exist in other fungal species. For future in-depth analyses of KEP-derived peptides in target fungi, an in-house KEP screening on the fungus of interest followed by comparison with the results by Le Marquer et al. and Umemura should ensure that all KEPs are found.

For future screenings, it may be interesting to search for secreted proteins with repetitive KEX2-cleavage sites that do not have a high degree of sequence similarity between the interjacent sequences. For example, the precursor of the cytolytic peptide candidalysin from *C. albicans* has a low degree of sequence similarity between KEX2-flanked sequences (Chapter 2, Fig. 1) (Moyes et al. 2016). It is possible that there are other proteins with a similar structure whose derived peptides have interesting bioactivities, although these proteins might not be considered “KEX2-processed repeat proteins.” In our current screening method, at least one KEP with a somewhat unusual sequence slipped through the screening: The *C. cinerea* KEP 447393, whose derived peptides GPPAPSW and DDDPEGPPAPSW were detected in culture supernatant, was found by chance during manual screening of *C. cinerea* proteins with secretory signal sequences. The protein was missed in our screening because we looked exclusively for proteins with at least three sequence repeats separated by cleavage sites. KEP 447393 contains up to six sequence repeats, but these are often separated by single arginines instead of dibasic KEX2 cleavage sites (Chapter 2, Table S1). When multiple repeats are found between two dibasic cleavage sites, the algorithm recognises them only once, so KEP 447393 does not reach the required number of repeats.

### **The expression of KEPs in *Pichia pastoris***

The *P. pastoris* expression system has been successfully used in the past to produce a variety of heterologous proteins and peptides, including the nematotoxic peptide macrocycles

omphalotin, lentinulin and dendrothelin (Matabaro et al. 2021) or the KEX2-processed antimicrobial cysteine-stabilized  $\alpha\beta$ -defensin copsisin (Essig et al. 2014). In Chapter 2, we showed that expression of *C. cinerea* KEPs in *P. pastoris* results in the accumulation of a variety of different KEP-derived peptides in the culture supernatant that appear to have undergone KEX2 and KEX1 cleavage and, in some cases, STE13-mediated cleavage and pyroglutamate formation. In principle, expression in *P. pastoris* could be used as a cost-efficient and convenient approach for large-scale production of KEP-derived peptides with interesting bioactivities. The precursor proteins could be modified in various ways to produce specific peptides. While we used the endogenous *C. cinerea* signal sequences for secretion in our expression, it may be useful to exchange this sequence for the  $\alpha$ -factor signal sequence of *S. cerevisiae*, a strategy commonly used in *P. pastoris* expression to increase secretion (Brake et al. 1984). In addition, the precursor could be modified to contain multiple repeats of the same peptide to increase yield and avoid byproducts. Finally, the coding sequence could be codon optimized for *P. pastoris*.

### **Detection of KEP-derived peptides in fungal tissues and evaluation of their function**

In *C. cinerea*, we confirmed the presence of five peptides derived from two KEPs and detected in different tissues in different concentrations (Chapter 2, Fig. 3). In edible fruiting bodies of *L. edodes* and *P. ostratus/eryngii*, we confirmed the presence of a total of four peptides derived from one KEP of *L. edodes* and one KEP of *P. ostratus/eryngii* (Chapter 2, Fig. 5). We constructed knockout cell lines in *C. cinerea* for six different *kep* genes, but we were unable to detect a phenotype in these strains. We therefore hypothesize that these KEPs and their derived peptides play a redundant role in growth or differentiation, or that they play a role in the interaction of the fungus with the biotic environment. Simple, linear peptides are generally not known to have toxic properties. However, there are exceptions such as the cytolytic peptide candidalysin (Moyes et al. 2016). Therefore, we tested synthetically produced peptides in toxicity assays against the bacteria *E. coli*, *B. subtilis*, *Micrococcus aureus* and *Staphylococcus aureus* (Chapter 2, Fig. S14). Although we did not observe antibacterial activity, it is possible that the peptides are highly specific and therefore their activity is directed against a specific predator or competitor that has not yet been tested. Previous studies have shown the specificity of certain defense molecules against predators, the lectin CGL2 for example, which specifically targets the development of certain nematode species while being harmless to others (Tayrov et al. 2019). Therefore, we will test additional bacteria and fungi, bacterivorous and fungivorous nematodes and fungivorous arthropods such as dipteran larvae and collembolae (springtails). In addition, we will expand the characterization of our *C. cinerea* knockout cell lines, including cultivation on various substrates such as minimal medium and sterilized and unsterilized horse dung. Horse dung is the natural habitat of *C. cinerea*, and cultivation on the unsterilized substrate would allow us to study the individual fitness of the knockout strains in competition with their natural competitors and predators.

### Similarities of KEP-derived peptides with neuropeptides and implications thereof

KEP-derived peptides and animal neuropeptides have structurally similar precursor proteins with N-terminal secretion sequences and peptide cores flanked by dibasic residues, which are cleaved by KEX2 and KEX homologs, respectively (Chen et al. 2019). The putative opisthokont origin of KEPs suggests that a KEP-like precursor organization offers many advantages that have led to its conservation in both fungi and animals: precursor processing by the same set of promiscuous enzymes, parallel production of multiple peptides from the same precursor protein, secretion of peptides via the secretory pathway, and the presence of multiple peptide cores that allow mutation acquisition and peptide diversification. RiPP precursor proteins with multiple encoded peptide cores are also found in bacteria. The cyanobactins are encoded by precursor proteins containing up to four similar peptide cores that are heavily modified before being cleaved from the precursor protein by subtilisin-like proteases (Schmidt et al. 2005). Cyanobactin biosynthesis has already been utilized to generate diversified peptide libraries (Ruffner et al. 2015). Other “multicore RiPPs” in bacteria include microviridin and  $\Omega$ -ester peptides, and cyclotides, orbitides and lyciumins in plants (Rubin and Ding 2020).

In animals, neuropeptides typically bind to G protein-coupled receptors (GPCRs), which are conserved in eukaryotes and involved in many important signal transduction pathways (Hoyer and Bartfai 2012). In fungi,  $\alpha$ -type mating pheromones bind to the GPCR Ste2 (Alvaro and Thorner 2016). It can be speculated that GPCRs may act as receptors for some linear KEP-derived peptides in fungi (Le Marquer et al. 2019). This may not be true for highly modified KEPs such as the dikaritins, some of which are known to target tubulin (Cormier et al. 2008). GPCRs are associated with many signaling pathways related to disease, and it is estimated that approximately 35% of currently available drugs target GPCRs (Sriram and Insel 2018). *C. cinerea* has at least six GPCR homologs (data not shown), some of which are localized in the B mating type locus in *C. cinerea* and act as a-pheromone receptors (Coelho et al. 2017). It would be highly interesting to further investigate whether these receptors, especially those outside the mating locus, might be involved in putative signaling pathways of KEP-derived peptides.

KEP-derived peptides generally do not share sequence similarities with neuropeptides. A BLAST search of *C. cinerea* KEPs does not result in any animal homologs. However, members of a major neuropeptide family in arthropods, the adipokinetic hormone/red pigment-concentrating hormone (AKH/RPCH) family, have interesting sequences reminiscent of some fungal KEP-derived peptides. The neuropeptides are short, pyroglutamated and contain a C-terminal amidated tryptophan (Verlinden et al. 2014). Several peptides contain a C-terminal “PSW” sequence, similar to the *C. cinerea* peptides GPPAPSW and DDDPEGPPAPSW (e.g. Pyro-QLTFTPSW-amide, *Aedes aegypti*), or a C-terminal “DW” sequence, similar to SGTGEASAADW, MSGVAADW, SGVAADW and LSGVAADW (e.g. Pyro-QLTFSPDW-amide, *Drosophila melanogaster*) (Wegener et al. 2006; Predel et al. 2010; Wang et al. 2015). AKH/RPCH peptides, like other neuropeptides, act via GPCR signaling (Verlinden et al. 2014). AKHs regulate energy metabolism in energy-consuming processes in insects, whereas RPCHs are mainly found in the eyestalk ganglia in crustaceans, where they concentrate pigment granules

(Dirksen 2013). In *C. elegans*, some peptides can be found with a C-terminal amidated tryptophan (e.g. SPAQWQRANGLW-amide, Husson et al. 2005).

Neuropeptides are produced by neurons and secreted into the extracellular space of tissues, where they act on other neurons or the periphery (Russo 2017). Unlike classical neurotransmitters, neuropeptides diffuse from their site of release and can act on distant targets. In higher animals, neuropeptides are often produced in more than one tissue and processed in a tissue-specific manner (Russo 2017). In contrast to animals, fungi do not form tissues or organs (with the exception of the fruiting body) and do not have a confined extracellular space that is part of the organism. Each hyphae is in contact with the outside world, and secreted products either remain associated with the hyphal cell wall or diffuse. Thus, unlike neuropeptides, fungal KEP-derived peptides are less likely to have an organism-intrinsic signaling function, although this possibility cannot be ruled out based on the examples of sexual pheromones, quorum sensing factors and putative anastomosis-regulating molecules (Fleissner et al. 2009; Coelho et al. 2017; Mehmood et al. 2019).

### **Phenotypes of *C. cinerea* enzyme knockouts**

We observed defects in mycelial growth in *C. cinerea*  $\Delta kex2c$  strains and defects in fruiting body formation in  $\Delta kex1$ ,  $\Delta kex2b$  and  $\Delta kex2c$  strains (Chapter 2, Fig. 6). Previously established *kex2* knockout strains in ascomycetes frequently displayed growth defects, and *in vitro* digestion assays demonstrated processing of cell wall modulating enzymes and components by KEX2 (Bader et al. 2008). Moreover, knockouts of cell wall modulating enzymes and components showed the same phenotype as *kex2* knockouts (Martínez et al. 2004; Firon et al. 2007). Thus, it seems likely that the  $\Delta kex$  phenotypes in *C. cinerea* are due to defective processing of enzyme substrates rather than a lack of KEP-derived peptide processing.

The three KEX2 proteases show high sequence similarity (Chapter 2, Fig. S11), but the phenotypes of their knockouts differ, suggesting that they have different substrates.  $\Delta kex2c$  strains are the only knockouts with a mycelial growth deficiency, indicating that KEX2c is required for cleavage of a substrate required for growth and the other proteases cannot substitute this function. Similarly,  $\Delta kex2b$  and  $\Delta kex2c$  exhibit fruiting defects, but  $\Delta kex2a$  does not, indicating that KEX2b and KEX2c are required for cleavage of a potential substrate in fruiting, whereas KEX2a is unable to perform the same task. Both  $\Delta kex2b$  and  $\Delta kex2c$  exhibit fruiting deficiency, suggesting that they cannot substitute for each other's function. This could indicate that they either cleave different substrates, each of which is required for fruiting, or that they are involved in cleaving a substrate for which both proteases are required. Interestingly, the  $\Delta kex1$  strain showed the same fruiting deficiency as the  $\Delta kex2b$  and  $\Delta kex2c$  strains, suggesting that the same target(s) are involved in these strains and that they require processing by KEX2b/c and KEX1 to be functional. Thus, cell wall remodeling enzymes or cell wall components with C-terminal propeptides whose removal requires processing by both KEX2 and KEX1 would be good candidates for future studies of the  $\Delta kex$ -mediated defect in fruiting body formation. In contrast, the defect in mycelial growth was observed only in  $\Delta kex2c$

strains, suggesting that a different target is involved here that does not require KEX1 processing.

To further test the specificity of KEX2 proteases, an *in vitro* digestion assay using recombinantly produced KEX2 proteases and synthetic substrates could be performed to test the cleavage specificity of KEX2 proteases. In addition, potential endogenous enzyme substrates of *C. cinerea* could also be tested. While no individual KEX2 protease is required for cleavage of KEPs (Chapter 2, Fig. S9), our experiments did not address whether each of the three KEX2 proteases is capable of cleaving KEPs, which could be tested in additional *in vitro* assays.

The KEX2 protease requires posttranslational processing for maturation. The *S. cerevisiae* KEX2 protease is first cleaved at an N-terminal KR dibasic site in a supposedly autoproteolytic manner in the ER, followed by STE13 processing at the N-terminus and the addition of oligosaccharides during transport from the ER to the Golgi (Wilcox and Fuller 1991). The potential autoproteolysis of KEX2 in the ER is not well studied and it is not known whether the proteolysis occurs intra- or intermolecularly. Therefore, there is a possibility that the three KEX2 proteases of *C. cinerea* cleave each other intermolecularly, which could have implications for knockout strain phenotypes. If the proteases are not equivalent in their ability to perform autoproteolysis, knockout of one protease could affect the maturation of the other two proteases and lead to defects.

We plan to further characterize the *kex* knockout strains, including microscopic analysis of their hyphae, sclerotia and hyphal knots. In addition, we are considering a number of new knockouts. Knockout of the glutaminyl cyclase homolog in *C. cinerea* would allow us to verify whether the pyroglutamated peptides are enzymatically modified and whether they can be detected in their unmodified form. Glutaminyl cyclase-mediated modification of N-terminal glutamine or glutamic acid to pyroglutamate is common not only in peptides but also in proteins (Ball et al. 1998; Liu et al. 2011). Therefore, knockout of *C. cinerea* glutaminyl cyclase could destabilize both peptides and proteins and lead to other phenotypic effects. Screening for proteins with a secretory signal sequence and an N-terminal glutamine or glutamic acid would yield a list of potential target enzymes that could be further investigated. A similar analysis in *Neurospora crassa* led to many potential glutaminyl cyclase substrates, including cell wall-degrading and modifying enzymes and secreted proteases (Wu et al. 2017). Similarly, knockout of *ste13* could lead to new interesting phenotypes and targets. In *S. cerevisiae*, STE13 processes not only  $\alpha$ -pheromones but also enzymatic targets such as the KEX2 protease (Brenner and Fuller 1992). While the KEX2 proteases of *C. cinerea* do not have STE13 recognition motifs, there may be other proteins that require STE13 processing and their incomplete processing could result in a defective phenotype.

KEX2-mediated cleavage of protein 365456 appears to simultaneously generate an enzymatically active C-terminal glycoside hydrolase domain and a small pyroglutamated peptide with the sequence Pyro-QSEPKPTN (Chapter 2, Fig. 4). The presence of the peptide in all *kex* knockout strains suggests that the release of the peptide is not dependent on a single

KEX2 protease. This may indicate that the C-terminal enzymatic domain is also processed and active in all *kex* knockout strains, implying that the *kex* knockout phenotypes are likely unrelated to this protein. Nevertheless, knockout of this protein could lead to an interesting phenotype because the glycoside hydrolase domain belongs to the same family as a  $\beta$ -1,3-glucanase in *L. edodes* (Sakamoto et al. 2011) and thus may have a cell wall-remodeling function. Further studies of the protein may allow functional elucidation of both the enzymatic domain and the simultaneously produced peptide.

### **Elucidating the biosynthesis of novel fungal macrocycles**

In Chapter 3, we investigated a ribosomal origin for the macrocycles mariannamides from *Mariannaea elegans* (Ishiuchi et al. 2020) and the backbone N-methylated macrocycles verrucamides and broomeanamides from *Myrothecium verrucaria* and *Sphaerostilbella broomeana*, respectively (Zou et al. 2011; Ekanayake et al. 2021). Based on the similarity of these peptides to the RiPP family of borosins, which are N-methylated cyclic peptides, we suspected that they might be ribosomally encoded, although the presence of D-amino acids in the verrucamides rather suggested a non-ribosomal origin. However, after genomic sequencing of the three fungi in addition to *Rhizopogon roseolus*, we found no potential precursor proteins for mariannamides, verrucamides and broomeanamides. A general screen for borosin precursor proteins in all four fungi also yielded no results. We then performed an antiSMASH search (Blin et al. 2019) and found potential NRPS biosynthetic gene clusters for mariannamides, verrucamides and broomeanamides, as well as several clusters encoding putative peptaibols, an antimicrobial peptide class (Fuente-Núñez et al. 2013). These NRP clusters will hopefully be further investigated to confirm the biosynthetic origin of these peptides.

Verrucamides, mariannamides and peptaibols possess antibacterial properties (Zou et al. 2011; Fuente-Núñez et al. 2013; Ishiuchi et al. 2020). These peptides are of particular interest with respect to other fungal cyclic peptides with interesting bioactivities, such as the backbone N-methylated macrocycle cyclosporine, which is used as an immunosuppressant (Dreyfuss et al. 1976). While oral delivery of peptide drugs is often poor due to low stability, proteolytic degradation and poor absorption by epithelial membranes (Vorherr et al. 2018), modifications such as cyclizations and N-methylations can improve these pharmacological properties by increasing proteolytic stability, structural rigidity, membrane permeability and oral availability (Marsault and Peterson 2011; Sharma et al. 2018). In addition, N-methylations can modulate selectivity (Chatterjee et al. 2013). Therefore, these novel peptide macrocycles will hopefully be further investigated for their therapeutic potential.

Overall, our work has expanded the knowledge about fungal peptides with a focus on KEP-derived peptides. We demonstrated that the two previous genomic screens for KEPs underestimated the number of KEPs in fungal species and that KEPs can be heterologously expressed in *P. pastoris*. We established a protocol to extract and detect several KEP-derived peptides in different agaricomycetes. Although we did not find a specific function of these

## Chapter 4

peptides, we demonstrate the importance of KEX proteases for normal growth and differentiation in *C. cinerea*.

## References

- Alvaro CG, Thorner J. 2016. Heterotrimeric G protein-coupled receptor signaling in yeast mating pheromone response. *J Biol Chem* [Internet]. 291(15):7785–7798. <https://doi.org/10.1074/jbc.R116.714980>
- Bader O, Krauke Y, Hube B. 2008. Processing of predicted substrates of fungal Kex2 proteinases from *Candida albicans*, *C. glabrata*, *Saccharomyces cerevisiae* and *Pichia pastoris*. *BMC Microbiol.* 8:1–16. <https://doi.org/10.1186/1471-2180-8-116>
- Ball LE, Oatis JE, Dharmasiri K, Busman M, Wang J, Cowden LB, Galijatovic A, Chen N, Crouch RK, Knapp DR. 1998. Mass spectrometric analysis of integral membrane proteins: Application to complete mapping of bacteriorhodopsins and rhodopsin. *Protein Sci.* 7(3):758–764. <https://doi.org/10.1002/pro.5560070325>
- Blin K, Shaw S, Steinke K, Villebro R, Ziemert N, Lee SY, Medema MH, Weber T. 2019. AntiSMASH 5.0: Updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Res.* 47(W1):W81–W87. <https://doi.org/10.1093/nar/gkz310>
- Brake AJ, Merryweather JP, Coit DG, Heberlein UA, Masiarz FR, Mullenbach GT, Urdea MS, Valenzuela P, Barr PJ. 1984.  $\alpha$ -Factor-directed synthesis and secretion of mature foreign proteins in *Saccharomyces cerevisiae*. *Proc Natl Acad Sci U S A.* 81(15 1):4642–4646. <https://doi.org/10.1073/pnas.81.15.4642>
- Brenner C, Fuller RS. 1992. Structural and enzymatic characterization of a purified prohormone- processing enzyme: Secreted, soluble Kex2 protease. *Proc Natl Acad Sci U S A.* 89(3):922–926. <https://doi.org/10.1073/pnas.89.3.922>
- Chatterjee J, Rechenmacher F, Kessler H. 2013. N-Methylation of peptides and proteins: An important element for modulating biological functions. *Angew Chemie - Int Ed.* 52(1):254–269. <https://doi.org/10.1002/anie.201205674>
- Chen M, Talarovicova A, Zheng Y, Storey KB, Elphick MR. 2019. Neuropeptide precursors and neuropeptides in the sea cucumber *Apostichopus japonicus* : a genomic , transcriptomic and proteomic analysis. *Sci Rep* [Internet]. 9(8829). <https://doi.org/10.1038/s41598-019-45271-3>
- Coelho MA, Bakkeren G, Sun S, Hood ME, Giraud T. 2017. Fungal sex: The basidiomycota. *The Fungal Kingdom.* 5(3):147–175. <https://doi.org/10.1128/9781555819583.ch7>
- Cormier A, Marchand M, Ravelli RBG, Knossow M, Gigant B. 2008. Structural insight into the inhibition of tubulin by vinca domain peptide ligands. *EMBO Rep.* 9(11):1101–1106. <https://doi.org/10.1038/embor.2008.171>
- Dirksen H. 2013. Crustacean Bioactive Peptides. Second Edi. Kastin AJ, editor. Cambridge: Elsevier Inc. <https://doi.org/10.1016/B978-0-12-385095-9.00032-4>
- Dreyfuss M, Härrä E, Hofmann H, Kobel H, Pache W, Tschertter H. 1976. Cyclosporin A and C New Metabolites from *Trichoderma polysporum*. *Eur J Appl Microbiol.* 3(1):125–133. <https://doi.org/10.1007/BF00928431>
- Ekanayake DI, Perlatti B, Swenson DC, Pöldmaa K, Bills GF, Gloer JB. 2021. Broomeanamides: Cyclic Octapeptides from an Isolate of the Fungicolous Ascomycete *Sphaerostilbella broomeana* from India. *J Nat Prod.* 84(7):2028–2034. <https://doi.org/10.1021/acs.jnatprod.1c00414>

## Chapter 4

- Essig A, Hofmann D, Münch D, Gayathri S, Künzler M, Kallio PT, Sahl H, Wider G, Schneider T, Aebi M. 2014. Copsin , a Novel Peptide-based Fungal Antibiotic Interfering with the Peptidoglycan Synthesis \*. 289(50):34953–34964. <https://doi.org/10.1074/jbc.M114.599878>
- Firon A, Aubert S, Iraqui I, Guadagnini S, Goyard S, Prévost MC, Janbon G, D'Enfert C. 2007. The SUN41 and SUN42 genes are essential for cell separation in *Candida albicans*. *Mol Microbiol.* 66(5):1256–1275. <https://doi.org/10.1111/j.1365-2958.2007.06011.x>
- Fleissner A, Leeder AC, Roca MG, Read ND, Glass NL. 2009. Oscillatory recruitment of signaling proteins to cell tips promotes coordinated behavior during cell fusion. *Proc Natl Acad Sci U S A.* 106(46):19387–19392. <https://doi.org/10.1073/pnas.0907039106>
- Fuente-Núñez C de la, Whitmore L, Wallace BA. 2013. *Peptaibols*. Second Edi. Cambridge: Elsevier Inc. <https://doi.org/10.1016/B978-0-12-385095-9.00022-1>
- Hoyer D, Bartfai T. 2012. *Neuropeptides and Neuropeptide Receptors: Drug Targets , and Peptide and Non-Peptide Ligands: a Tribute to Prof . Dieter Seebach*. *Chem Biodivers.* 9:2367–2387. <https://doi.org/10.1002/cbdv.201200288>
- Husson SJ, Clynen E, Baggerman G, De Loof A, Schoofs L. 2005. Discovering neuropeptides in *Caenorhabditis elegans* by two dimensional liquid chromatography and mass spectrometry. *Biochem Biophys Res Commun.* 335(1):76–86. <https://doi.org/10.1016/j.bbrc.2005.07.044>
- Ishiuchi K, Hirose D, Kondo T, Watanabe K, Terasaka K, Makino T. 2020. Mariannamides A and B, new cyclic octapeptides isolated from *Mariannaea elegans* NBRC102301. *Bioorganic Med Chem Lett [Internet].* 30(4):126946. <https://doi.org/10.1016/j.bmcl.2019.126946>
- Julius D, Blair L, Brake A, Sprague G, Thorner J. 1983. Yeast  $\alpha$  factor is processed from a larger precursor polypeptide: The essential role of a membrane-bound dipeptidyl aminopeptidase. *Cell.* 32(3):839–852. [https://doi.org/10.1016/0092-8674\(83\)90070-3](https://doi.org/10.1016/0092-8674(83)90070-3)
- Liu YD, Goetze AM, Bass RB, Flynn GC. 2011. N-terminal glutamate to pyroglutamate conversion in vivo for human IgG2 antibodies. *J Biol Chem.* 286(13):11211–11217. <https://doi.org/10.1074/jbc.M110.185041>
- Le Marquer M, San Clemente H, Roux C, Savelli B, Frei dit Frey N. 2019. Identification of new signalling peptides through a genome-wide survey of 250 fungal secretomes. *BMC Genomics [Internet].* 20(1):64. <https://doi.org/10.1186/s12864-018-5414-2>
- Marsault E, Peterson ML. 2011. Macrocycles are great cycles: Applications, opportunities, and challenges of synthetic macrocycles in drug discovery. *J Med Chem.* 54(7):1961–2004. <https://doi.org/10.1021/jm1012374>
- Martínez AI, Castillo L, Garcerá A, Elorza M V., Valentín E, Sentandreu R. 2004. Role of Pir1 in the construction of the *Candida albicans* cell wall. *Microbiology.* 150(10):3151–3161. <https://doi.org/10.1099/mic.0.27220-0>
- Matabaro E, Kaspar H, Dahlin P, Bader DLV, Murar CE, Staubli F, Field CM, Bode JW, Künzler M. 2021. Identification, heterologous production and bioactivity of lentinulin A and dendrothelin A, two natural variants of backbone N-methylated peptide macrocycle omphalotin A. *Sci Rep.* 11(1):1–12. <https://doi.org/10.1038/s41598-021-83106-2>
- Mehmood A, Liu G, Wang X, Meng G, Wang C, Liu Y. 2019. Fungal quorum-sensing molecules and inhibitors with potential antifungal activity: A review. *Molecules.* 24(10):1–18.

<https://doi.org/10.3390/molecules24101950>

Moyes DL, Wilson D, Richardson JP, Mogavero S, Tang SX, Wernecke J, Höfs S, Gratacap RL, Robbins J, Runglall M, et al. 2016. Candidalysin is a fungal peptide toxin critical for mucosal infection. *Nature* [Internet]. 532(7597):64–68. <https://doi.org/10.1038/nature17625>

Predel R, Neupert S, Garczynski SF, Crim JW, Brown MR, Russell WK, Kahnt J, Russell DH, Nachman RJ. 2010. Neuropeptidomics of the mosquito *Aedes aegypti*. GJ S, BW A, RW A, editors. *J Proteome Res.* 9(4):2006–2015. <https://doi.org/10.1021/pr901187p>

Rubin GM, Ding Y. 2020. Recent advances in the biosynthesis of RiPPs from multicore-containing precursor peptides. *J Ind Microbiol Biotechnol* [Internet]. 47(9–10):659–674. <https://doi.org/10.1007/s10295-020-02289-1>

Ruffner DE, Schmidt EW, Heemstra JR. 2015. Assessing the combinatorial potential of the RiPP cyanobactin tru pathway. *ACS Synth Biol.* 4(4):482–492. <https://doi.org/10.1021/sb500267d>

Russo AF. 2017. Overview of Neuropeptides: Awakening the Senses? *Headache.* 57:37–46. <https://doi.org/10.1111/head.13084>

Sakamoto Y, Nakade K, Konno N. 2011. Endo- $\beta$ -1,3-Glucanase GLU1, from the fruiting body of *Lentinula edodes*, belongs to a new glycoside hydrolase family. *Appl Environ Microbiol.* 77(23):8350–8354. <https://doi.org/10.1128/AEM.05581-11>

Schmidt EW, Nelson JT, Rasko DA, Sudek S, Eisen JA, Haygood MG, Ravel J. 2005. Patellamide A and C biosynthesis by a microcin-like pathway in *Prochloron didemni*, the cyanobacterial symbiont of *Lissoclinum patella*. *Proc Natl Acad Sci U S A.* 102(20):7315–7320. <https://doi.org/10.1073/pnas.0501424102>

Sharma A, Kumar A, Abdel Monaim SAH, Jad YE, El-Faham A, de la Torre BG, Albericio F. 2018. N-methylation in amino acids and peptides: Scope and limitations. *Biopolymers.* 109(10). <https://doi.org/10.1002/bip.23110>

Sriram K, Insel PA. 2018. G protein-coupled receptors as targets for approved drugs: How many targets and how many drugs? *Mol Pharmacol.* 93(4):251–258. <https://doi.org/10.1124/mol.117.111062>

Tayyrov A, Azevedo S, Herzog R, Vogt E, Arzt S, Lüthy P, Müller P, Rühl M, Hennicke F, Künzler M. 2019. Heterologous production and functional characterization of ageritin, a novel type of ribotoxin highly expressed during fruiting of the edible mushroom *Agrocybe aegerita*. *Appl Environ Microbiol.* 85(21):1–15. <https://doi.org/10.1128/AEM.01549-19>

Umemura M. 2020. Peptides derived from Kex2-processed repeat proteins are widely distributed and highly diverse in the Fungi kingdom. *Fungal Biol Biotechnol* [Internet]. 7(1):1–24. <https://doi.org/10.1186/s40694-020-00100-5>

Verlinden H, Vleugels R, Zels S, Dillen S, Lenaerts C, Crabbé K, Spit J, Vanden Broeck J. 2014. Receptors for neuronal or endocrine signalling molecules as potential targets for the control of insect pests. Cohen E, editor. Cambridge. <https://doi.org/10.1016/B978-0-12-417010-0.00003-3>

Vorherr T, Lewis I, Berghausen J, Desrayaud S, Schaefer M. 2018. Modulation of Oral Bioavailability and Metabolism for Closely Related Cyclic Hexapeptides. *Int J Pept Res Ther.* 24(1):35–48. <https://doi.org/10.1007/s10989-017-9590-8>

## Chapter 4

Wang Y, Wang M, Yin S, Jang R, Wang J, Xue Z, Xu T. 2015. NeuroPep: A comprehensive resource of neuropeptides. Database. 2015:1–9. <https://doi.org/10.1093/database/bav038>

Wegener C, Reinl T, Jansch L, Predel R. 2006. Direct mass spectrometric peptide profiling and fragmentation of larval peptide hormone release sites in *Drosophila melanogaster* reveals tagma-specific peptide expression and differential processing. *J Neurochem.* 96(5):1362–1374. <https://doi.org/10.1111/j.1471-4159.2005.03634.x>

Wilcox CA, Fuller RS. 1991. Posttranslational Processing of the Prohormone-cleaving Kex2 Protease in the. *J Cell Biol.* 115(2):297–307. <https://doi.org/10.1083/JCB.115.2.297>

Wu VW, Dana CM, Iavarone AT, Clark DS, Glass NL. 2017. Identification of glutaminyl cyclase genes involved in pyroglutamate modification of fungal lignocellulolytic enzymes. *MBio.* 8(1). <https://doi.org/10.1128/mBio.02231-16>

Zou X, Niu S, Ren J, Li E, Liu X, Che Y. 2011. Verrucamides A-D, antibacterial cyclopeptides from *Myrothecium verrucaria*. *J Nat Prod.* 74(5):1111–1116. <https://doi.org/10.1021/np200050r>