Diversity of MSDIN family members in amanitin-producing mushrooms and the phylogeny of the MSDIN and prolyl oligopeptidase genes

CURRENT STATUS: UNDER REVISION

Zhengmi He
Hunan Normal University

Pan Long
Hunan Normal University

Fang Fang
Hunan Normal University

Sainan Li
Hunan Normal University

Ping Zhang
Hunan Normal University

Zuohong Chen
Hunan Normal University

Email: chenzuohong@263.net

DOI: 10.21203/rs.2.22199/v1

SUBJECT AREAS
Epigenetics & Genomics

KEYWORDS
Amanita, Galerina, Lepiota, cyclopeptide toxin, prolyl oligopeptidase, horizontal gene transfer
Abstract

**Background**

Amanitin-producing mushrooms, mainly distributed in the genera Amanita, Galerina and Lepiota, possess MSDIN gene family for the biosynthesis of many cyclopeptides catalyzed by prolyl oligopeptidase (POP). Recently, transcriptome sequencing has proven to be an efficient way to mine MSDIN and POP genes in these lethal mushrooms. Until now, only A. palloidoides and A. bisporigera from North America and A. exitialis from Asia have been studied based on transcriptome analysis. However, MSDIN and POP genes of many amanitin-producing mushrooms in China remain unstudied, and hence the transcriptomes of these species deserve to be analyzed. Results

In this study, the MSDIN and POP genes from ten Amanita species, two Galerina species and Lepiota venenata were studied and the phylogenetic relationships of their MSDIN and POP genes were analyzed. Through transcriptome sequencing and PCR cloning, 19 POP genes and 151 MSDIN genes predicted to encode 98 non-duplicated cyclopeptides, including α-amanitin, β-amanitin, phallacidin, phalloidin and 94 unknown peptides, were found in these species. Phylogenetic analysis showed that toxin peptide genes were clustered depending on the chemical substance within genus while depending on the taxonomy between genus and that the POPA genes of Amanita, Galerina and Lepiota were clustered and separated in three different groups, but the POPB genes of the three distinct genera were clustered in a highly monophyletic group. Conclusions

These results above indicate that lethal Amanita species have the genetic capacity to produce numerous cyclopeptides, most of which are unknown, while lethal Galerina and Lepiota species seem to only have the genetic capacity to produce α-amanitin. Additionally, the POPB phylogeny of Amanita, Galerina and Lepiota conflicts with the taxonomic status of the three genera, suggesting that horizontal gene transfer might occur among the three genera.

**Background**

Amatoxins, which are lethal substances found in mushrooms, have mainly been reported to be present in species from three distinct genera: Amanita, Galerina and Lepiota [1-4]. Among these amanitin-producing mushrooms, lethal Amanita species are the most well-known and typical mushrooms that produce three primary groups of cyclopeptide toxins, namely, amatoxins,
phallotoxins and virotoxins, which are bicyclic octapeptides, bicyclic heptapeptides and monocyclic heptapeptides, respectively [1-4].

The precursor peptide genes of α-amanitin (α-AMA) and phallacidin (PHD) along with multiple related sequences encoding unknown cyclic peptides were first found and predicted in *Amanita bisporigera* by genome shotgun sequencing, indicating that amatoxins and phallotoxins are encoded by a gene family and biosynthesized on ribosomes [5]. This gene family is called “MSDIN” in reference to the first five conserved amino acids, and as precursor peptides, its members contain 33-37 amino acids: two conserved regions, 10 amino acids upstream and 17 amino acids downstream, a highly variable core region, and a 6-10 amino acid sequence that ultimately forms the corresponding cyclopeptide [6]. GmAMA, which is responsible for producing α-AMA, is also found in the genome of *Galerina marginata* [7]. *G. marginata* is a specific amanitin-containing species in the genus *Galerina*. Unlike lethal amanitas, *G. marginata* does not contain MSDIN-like family genes besides two copies of α-AMA genes. Additionally, α-AMA of *Lepiota brunneoincarnata*, which is a amanitin-containing mushroom of the genus *Lepiota*, has been successfully cloned [8]. Genome sequencing of *L. venenata*, another newly published amanitin-containing species, has been completed, and it has been shown to contain α-AMA genes [9]. Precursor peptide sequence alignment of α-AMA from *Amanita*, *Galerina* and *Lepiota* shows high divergence, except in the toxin region.

It has been strongly indicated that a prolyl oligopeptidase (POP) plays an important role in the initial processing of MSDIN precursor peptides. Since the toxin core regions are flanked by two highly conserved proline (Pro) residues, this enzyme can cleave the C-terminus of Pro residues and release the peptide chain of the toxin to form a cyclopeptide [10]. It has been reported that there are two types of POP in amatoxin-producing mushrooms: POPA, which behaves like a conventional housekeeping protein that is present in all species, and POPB, which is the enzyme that actually catalyzes the cutting and cyclization of precursor peptides [7, 11, 12].

Increasing numbers of MSDIN family members have been published since the first 15 MSDIN genes were found in the *A. bisporigera* genome, and four were amplified by degenerate primers in *A. phalloides* and *A. ocreata* [5]. Li et al. (2014) obtained 24 MSDIN members from 6 *Amanita* species.
using degenerate primers [13]. Recently, Pulman et al. (2016) showed by generating draft genome sequences of *A. palloides* and *A. bisporigera* that both possessed approximately 30 members of the MSDIN family, but only three of these genes were common to both [6]. Luo et al. (2018) found 18 and 22 MSDIN genes in the *A. subjunquillea* and *A. pallidorosea* genomes through PacBio and Illumina sequencing, respectively [8]. However, the MSDIN genes of many *Amanita*, *Galerina* and *Lepiota* amanitin-containing mushrooms have not been investigated in depth to date. Lethal *Amanita* species are classified in the genus *Amanita* section *Phalloideae* (Fr.) Quël [14, 15]. There have been approximately 50 lethal *Amanita* species reported worldwide, and the species diversity of lethal amanitas is strongly underestimated under the current taxonomy [15, 16]. Many new lethal *Amanita* and *Lepiota* species, including *A. rimosas*, *A. subfuliginea*, *A. subpallidorosea*, and *L. venenata*, etc., have been discovered over the past 10 years [17-20]. Meanwhile, in addition to the known twenty-two cyclopeptide toxins, some other new cyclopeptide substances, such as cycloamanide E and cycloamanide F in *A. phalloides* and amanexitide in *A. exitialis*, respectively, have been extracted and identified [6, 21, 22]. It has been reported that *A. bisporigera* and *A. phalloides* have large but essentially nonoverlapping potential for the biosynthesis of a variety of cyclopeptides, most of which are unknown according to predictions. Hence, considering the species diversity of amanitin-containing mushrooms and the broad genetic capacity of lethal amanitas to produce unknown cyclopeptides, there are still many new cyclopeptide genes and corresponding cyclopeptides to be discovered.

According to previous research, whole-genome sequencing has proven to be the most comprehensive and in-depth method to identify MSDIN genes or genes related to the cyclopeptide biosynthetic pathway in amanitin-producing mushrooms [6, 8]. Nevertheless, compared with genome sequencing, transcriptome sequencing provides an alternative efficient and low-cost method to obtain functional gene data. To the best of knowledge, until now, only *A. palloides* and *A. bisporigera* from North America and *A. exitialis* from Asia have been studied using transcriptome sequencing [6, 23]. In this study, the transcriptomes of seven amanitin-producing mushrooms (*A. exitialis*, *A. fuliginea*, *A. molliuscula*, *A. pallidorosea*, *A. rimosas*, *A. subpallidorosea* and *L. venenata*) and an *Amanita* species producing no amanitin (*A. oberwinklerana*) were sequenced. MSDIN and POP genes were searched
and predicted from the transcriptome data. Genomic and coding sequences of toxin MSDIN and POP genes were then cloned and verified. Similarly, MSDIN and POP sequences were also cloned from two Galerina strains (G. marginata and G. sulciceps). In addition to the Amanita species mentioned above, MSDIN genes from A. subfuliginea, A. subjunquillea and A. virosa were cloned using specific and degenerate primers. Furthermore, phylogenetic analysis was performed on the obtained toxin and POP genes obtained. Our study intends to (a) identify MSDIN genes from amanitin-producing mushrooms to guide the isolation and identification of related new unknown cyclopeptides and to (b) determine the evolutionary relationship of toxin MSDIN and POP genes in amanitin-producing mushrooms.

Results

Data filtering and assembly of transcriptomes

Transcriptome sequencing of seven amanitin-producing mushrooms was performed on the BGISEQ-500 platform using the combinational probe-anchor synthesis sequencing method. After the removal of ambiguous, adaptor-containing and low-quality sequences, clean data were obtained and de novo assembled using Trinity software. The main transcriptomic features and the NCBI accession numbers of the transcriptome data obtained in our study are presented in Table 1.

MSDIN and POP genes

Through transcriptome sequencing, 110 MSDIN genes (Shown in Table 2) were manually identified in 7 lethal Amanita and Lepiota species using known MSDIN members of A. bisporigera as TBLASTN queries. Additionally, 70 MSDIN genes (shown in Table 3) were obtained from 12 lethal Amanita, Galerina and Lepiota species by PCR cloning using degenerate and specific primers. In general, a total of 151 nonrepetitive MSDIN genes were identified at the genomic and transcriptomic levels by these methods. All the obtained MSDIN genes were predicted to encode 98 cyclopeptides, including α-amanitin (IWGIGCNP), β-amanitin (IWGIGCDP), phallacidin (AWLVDCP), phalloidin (AWLATCP) and 94 unknown peptides. These predicted cyclopeptides were composed of 6-11 amino acids, including 5 hexapeptides, 30 heptapeptides, 73 octapeptides, 22 nanopeptides, 19 decapeptides, and 2
undecapeptides.

Among the MSDIN members found in the nine lethal species of *Amanita* sect. *Phalloideae* used in our study, in addition to the common α-amanitin, β-amanitin, phallacidin and phalloidin (PHA), a few unnamed predicted peptides overlapped among different *Amanita* species, including “FNFFRFYP” in *A. exitialis* and *A. rimosae*; “FPWTGPFPVP” in *A. fuliginea* and *A. pallidorosea*; “IIIVGLIIP” in *A. fuliginea* and *A. rimosae*; “YFLPPIFSPPP” in *A. mollioasca* and *A. subpallidorosea*; “ISDPTAYP” in *A. pallidorosea* and *A. rimosae*; “IFWFIYFP” in *A. exitialis*, *A. fuliginea*, *A. rimosae* and *A. subpallidorosea*; and “ISDPTAYP” in *A. pallidorosea*, *A. rimosae*, *A. subfuliginea* and *A. subpallidorosea*. The remaining 87 core regions were unique to their corresponding species. The MSDIN genes encoding “AWLTDCP” in *A. exitialis*; “AWLMTTCP” in *A. pallidorosea*; “AWLECP” in *A. rimosae*; “AWLVTCP” in *A. fuliginea*, *A. subpallidorosea* and *A. virosa*; and “AWITDCP” and “AWLITCP” in *A. subpallidorosea* probably expressed new unknown phallotoxins because their core regions were similar to those of phallacidin (AWLVDCP) and phalloidin (AWLATCP). Additionally, no MSDIN genes were found, as expected, in *A. oberwinklerana*, a species of *Amanita* sect. *Lepidella* sensu Bas [16] or sect. *Roanokenses*, that does not contain cyclopeptide toxins [15].

In *G. marginata*, *G. sulciceps* and *L. venenata*, only MSDIN genes encoding α-amanitin were found, and this gene was the only gene common in the amanitin-producing genera *Amanita*, *Galerina* and *Lepiota*. Unlike lethal *Amanita* species, no other extra MSDIN genes were discovered except the α-amanitin gene. Interestingly, an MSDIN gene with the full amino acid sequence

```
MFDTNSTRPL*GIGCNWPFAHIDQTLVSGNDTC* (with the core region in bold and underlined)
```

was found in *G. sulciceps*. Due to its identity with the α-amanitin gene *Gs_α-AMA1* (MFDTNATRLP*WGIGCNWPFAHVDQTLALSGNDIC*) in *G. sulciceps*, it was named *Gs_α-AMA2*.

Similarly, 19 POP genes were identified from the transcriptomes of nine *Amanita*, two *Galerina* and a *Lepiota* species using known POPA and POPB genes of *A. bisporigera* and *G. marginata*, respectively, as the TBLASTN queries and these sequences were further verified by PCR amplification. Eleven lethal *Amanita*, *Galerina* and *Lepiota* species contained both POPA and POPB genes, but *A. oberwinklerana*, an *Amanita* species containing no cyclopeptide toxins, only contained the POPA gene. All of the
obtained POP sequences and their accession numbers are listed in Table 4.

**Comparison of MSDIN precursor peptide sequences**

WebLogo alignment was carried out for 145 MSDIN sequences obtained from nine *Amanita* species (Fig. 1A). The composition and structure of these sequences and the relative degree of conservation of amino acids at each point were analyzed. As shown in Fig. 1A, MSDIN precursor peptides in *Amanita* species were 31-38 amino acids in length and could be divided into three regions: a highly conserved upstream leader peptide, a relatively conserved downstream recognition sequence and a highly variable core peptide. The core peptide is located between P\textsuperscript{10} and P\textsuperscript{21} and contains the latter proline, and its two ends are leader peptide and recognition sequence of MSDIN. The leader peptide contains 10 amino acids, and M\textsuperscript{1}S\textsuperscript{2}D\textsuperscript{3}I\textsuperscript{4}N\textsuperscript{5}R\textsuperscript{8}L\textsuperscript{9}P\textsuperscript{10} are highly conserved, with conservation rates of 100% (145/145), 91.7% (133/145), 97.2% (141/145), 93.1% (135/145), 99.3% (144/145), 99.3% (144/145), 94.5% (137/145) and 99.3% (144/145), respectively. In the leader peptide, P\textsuperscript{10} is the first cleavage site for prolyl oligopeptidase (POPB) (Luo et al., 2014). The recognition sequence usually contains 17 amino acids, beginning with C\textsuperscript{22}V\textsuperscript{23}G\textsuperscript{24}D\textsuperscript{25}D\textsuperscript{26}, with conservation rates of 76.5% (111/145), 79.3% (115/145), 66.2% (96/145), 93.1% (135/145) and 83.4% (121/145), respectively, and ending with L\textsuperscript{31}T\textsuperscript{32}R\textsuperscript{33}G\textsuperscript{34}E\textsuperscript{35}L\textsuperscript{37}C\textsuperscript{38}, with conservation rates of 86.9% (126/145), 73.8% (107/145), 82.1% (119/145), 96.6% (140/145), 93.1% (135/145), 97.9% (142/145) and 91.7% (133/145), respectively. In the recognition sequence, L\textsuperscript{31} and L\textsuperscript{37} are conducive to the formation of an alpha helix and substrate recognition by the POPB enzyme [24]; meanwhile, the C-terminal Cys\textsuperscript{38} (sometimes replaced with Ser) is indispensable for performing the function of POPB [12]. The core peptide encodes cyclopeptides, and its last amino acid, P\textsuperscript{21}, is highly conserved, with a conservation rate of 94.5%, which is the second cleavage site for POPB [12].

The \(\alpha\)-AMA precursor peptide sequences of the genera *Amanita*, *Galerina* and *Lepiota* were compared, as shown in Fig. 1B. The \(\alpha\)-AMA sequences show few differences within the same genera but more
differences between the different genera. The $\alpha$-AMA leader peptides of the three genera show few differences and more conserved than the other sequences. The leader peptides of Amanita and Galerina contain 10 amino acids, while that of the genus Lepiota contains 9, and the three genera start with “MSDIN”, “MFDTN”, and “MDAN”, respectively. For the recognition sequences of the three genera, except for several highly conserved amino acids, specifically V, L, G, and LC or LS at the end, the sequences showed many differences. Overall, there were large differences among the $\alpha$-AMA sequences of Amanita, Galerina and Lepiota, but the Galerina and Lepiota $\alpha$-AMA sequences were closer to each other than Amanita.

The Amanita toxin MSDIN genes encoding amatoxins ($\alpha$-, $\beta$-amanitin) and phallotoxins (phallacidin and phalloidin), the major cyclopeptides in these mushrooms, were aligned with highlighted variations compared to the represented consensus sequences (Fig. 1C). In general, the sequences are more identical when with same core toxin peptide. More variations occur in the recognition sequences than those in leader peptides and phallotoxins sequences present higher amount of variations than amatoxins sequences.

**Structures of MSDIN and POP genes**

The genomic sequences and coding sequences of the toxin MSDIN and POP genes were pairwise aligned, and the gene compositions of the exons and introns were analyzed. As shown in Fig. 2, the POPA genes comprised 19 exons and 18 introns, while the POPB genes comprised 18 exons and 17 introns, which are very similar to other known POP genes. The $\alpha$-AMA genes of Amanita and Galerina contained three introns, while the $\alpha$-AMA genes of Lepiota contained two or three introns. In addition to $\alpha$-AMA, other toxin MSDIN genes in Amanita species, such as $\beta$-AMA, PHA and PHD, were also composed of three introns.

**Phylogenetic analysis of MSDIN and POP genes**

From the phylogenetic analysis, two maximum likelihood (ML) trees based on 46 toxin MSDIN genes from 14 amanitin-producing mushrooms and 58 POP genes from 46 agaric species were constructed.
In the toxin MSDIN gene tree (Fig. 3), all toxin MSDIN gene sequences were distributed in four clades. Clade I contained 10 α-amanitin gene sequences and 10 β-amanitin gene sequences from 10 lethal *Amanita* species forming a cluster with a 95% bootstrap and a 1.0 Bayesian posterior probabilities. Clade II contained MSDIN genes encoding AWLVDCP (phallacidin, PHA) and the unknown related variants AWLAEC, AWITDCP and AWLTDCCP forming a cluster with a 95% bootstrap and a 1.0 Bayesian posterior probabilities. Clade III contained MSDIN genes encoding AWLATCP (phalloidin, PHD) and the unknown related variants AWLMTCP and AWLVTCP forming a cluster with a 100% bootstrap and a 1.0 Bayesian posterior probabilities. Clade IV contained α-amanitin genes from *Galerina* and *Lepiota* species, including *G. marginata*, *G. sulciceps*, *L. subincarnata* and *L. venenata*, forming a cluster with a 100% bootstrap and a 1.0 Bayesian posterior probabilities. In the POP gene tree (Fig. 4), POPA sequences from *Amanita*, *Galerina* and *Lepiota* were separated from each other in different groups. *Amanita* POPA sequences (12) were clustered together as a single group, while *Galerina* POPA sequences (2) were clustered in a group containing *Gymnopilus dilepis* and *Gymnopilus chrysophilus*, and *Lepiota* POPA sequences (2) were clustered in a group containing *Agaricus bisporus* var. *bisporus*, *Agaricus bisporus* var. *burnettii*, *Leucoagaricus* sp. and *Macrolepiota fuliginosa*. However, POPB sequences (13) belonging to three disjunct genera (*Amanita*, *Galerina* and *Lepiota*) were clustered together forming a monophyletic group.

**Discussion**

It has been proven that lethal *Amanita* species are classified in the *Amanita* section *Phalloideae* and that these species contain the MSDIN gene family, allowing them to produce many small cyclopeptides, such as α-amanitin, on ribosomes [6, 16]. In the present study, nine lethal *Amanita* species from China, including *A. exitialis*, *A. fuliginea*, *A. mulliussula*, *A. pallidorosea*, *A. rimos*, *A. subfuliginea*, *A. subjunquillea*, *A. subpallidorosea* and *A. virosa*, were proven to contain MSDIN genes, the same as other lethal *Amanita* species, such as *A. bisporigera*, described previously. These results further suggest that species of the *Amanita* section *Phalloideae* are genetically similar and are able to biosynthesize amatoxins, phallotoxins and some other unknown peptides. Based on the MSDIN gene data from nine lethal *Amanita* species in China obtained in our study and
some other European and North American species, such as *A. bisporigera*, *A. phalloides* and *A. ocreata* [5, 6], most MSDIN genes were found to not be common and were even unique among these lethal *Amanita* species. In species of the *Amanita* section *Phalloideae*, the MSDIN gene encoding α-amanitin was present in all species, and MSDIN genes encoding β-amanitin, phallacidin and phalloidin were widely distributed but not common. This evidence suggested that each lethal *Amanita* species had its own independent MSDIN family and that the MSDIN genes demonstrated little crossing among lethal *Amanita* species.

In addition to species of the *Amanita* section *Phalloideae*, some *Galerina* and *Lepiota* species, such as *G. marginata* and *L. brunneoincarnata*, produce amatoxins as well [2, 25]. In our study, the MSDIN gene mining results showed that *G. marginata* and *L. venenata* only had two copies of the α-amanitin gene, and no additional MSDIN genes were found, consistent with the results of Luo et al. (2012) and Lüli et al. (2019) [7, 9]. Lethal *Galerina* and *Lepiota* species are supposed to have two copies of the α-amanitin gene. However, when analyzing the MSDIN genes of another *Galerina* species, *G. sulciceps*, the results showed that *G. sulciceps* only had a single copy of the α-amanitin gene, although it also contained an MSDIN gene that was extremely similar to the α-amanitin gene with an I*GIGCN* core region. This MSDIN gene seemed to be an α-amanitin gene mutation, and we speculated that its tryptophan (W) codon, TGG, in the core region was mutated to a termination codon, TGA, via a single-base substitution, which inhibited proper expression of the gene. In general, only the α-amanitin gene is found in *Amanita*, *Galerina* and *Lepiota*, which indicates that the α-amanitin genes in the three genera might share a common origin or originate from one genus. Additionally, MSDIN genes including β-AMA, PHA, PHD, etc. were only found in *Amanita*, which indicated that these MSDIN genes, except for α-AMA, were derived from lethal *Amanita* species.

Lethal *Amanita* species contain three primary kinds of peptide toxins, namely, amatoxins, phallotoxins and virotoxins [21]. MSDIN genes encoding amatoxin and phallotoxin were discovered in 2007 [5], but there has been no related evidence of MSDIN genes encoding virotoxins to date. It has been reported that *A. subpallidorosea* and *A. virosa* contain virotoxins [3, 26]. In this study, toxin genes of the two lethal *Amanita* species were also identified, and no virotoxin genes were found. Nevertheless, the two
species both contain MSDIN genes encoding AWLATCP (PHD) and AWLVTCP, which only had an amino acid difference in the composition of virotoxins (AWLATSP or AWLVTSP). Therefore, we speculated that virotoxins might be encoded by the PHD gene or the phallotoxin-like gene AWLVTCP and that cysteine (C) was transformed to serine (S) during post-translational modification.

Phylogenetic analysis showed that the Galerina α-AMA genes and Lepiota α-AMA genes were homologous but distant from the Amanita α-AMA gene. In the genus Amanita, α-AMA and β-AMA were mixed and clustered in a clade, which indicated that β-AMA might be derived from α-AMA. PHA genes (AWLVDCP) were clustered with MSDIN genes encoding AWLAECM, AWITDCP and AWLTDCP, and PHD genes (AWLATCP) were clustered with MSDIN genes encoding AWLMTCP and AWLVTCP, which indicated that the encoding products of these MSDIN genes were very likely to correspond to new unknown phallotoxins in consideration of their amino acid composition similarity with PHA and PHD and their capacity to contain tryptathione (Trp-Cys). These phallotoxin-like genes might be variants derived from PHA and PHD. For example, we found that the PHA gene (AWLVDCP) sequence in A. exitialis was almost the same as the sequence of the MSDIN gene encoding AWLTDCP, with only a two-nucleotide difference in the core region, and the valine (V) codon GTA likely mutated into the threonine (T) codon ACA. According to this evidence, it can be inferred that the MSDIN genes in Amanita evolved faster than those in Galerina and Lepiota, which induced the generation of a variety of new peptide genes and might also be the reason why the Galerina and Lepiota α-AMA genes differed from the Amanita α-AMA genes.

Horizontal gene transfer (HGT), also known as lateral gene transfer, refers to the transmission of genetic material between distinct organisms, specifically across species boundaries [27, 28]. It has been reported that HGT is very common in prokaryotes and can be an important source of their biological evolution, and HGT also occurs in eukaryotes at a lower frequency than that in prokaryotes [29-32]. The latest reports suggest that HGT may be responsible for the α-amanitin biosynthetic pathway found in the three distantly-related genera Amanita, Galerina and Lepiota [8, 9]. It has been reported that in amanitin-producing mushrooms, the POPB gene catalyzes the cleavage and cyclization of the toxin precursor peptide, while the POPA gene is a housekeeping gene unrelated to
toxin biosynthesis [7, 12]. In our study, phylogenetic analysis based on the POP gene showed that the POPA genes of Amanita, Galerina and Lepiota were distributed in three separated groups, but the POPB genes of the three genera were highly homologous forming a highly monophyletic group, which apparently conflicted with the species taxonomic status and could not be explained by conserved gene inheritance. Additionally, the MSDIN and POP genes were proven to have the same exon and intron structures. Regardless, these results can be considered evidence for HGT events among Amanita, Galerina and Lepiota. For a complete validation of HGT among amanitin-producing mushrooms in the future, more related species and their genomic data are required to perform a phylogenetic analysis with appropriate taxon sampling and tree building methodologies.

Conclusions

In conclusion, the MSDIN gene family is abundant and diverse. In addition to the peptide toxins α-amanitin, β-amanitin, phallacidin, phalloidin, etc., the MSDIN family encodes a variety of unknown small cyclopeptides. The amanitin-producing species Amanita, Galerina and Lepiota have a common toxin biosynthetic pathway, and their α-amanitin genes and POPB genes may have a common origin that includes HGT among the three distant genera.

Methods

Sample collection and preparation

Seven Amanita species and Lepiota venenata samples were collected from the wild for RNA extraction and sequencing, and their fresh basidiocarps were cleaned and placed on dry ice and transported back to the lab and stored at -80 °C. The mushroom samples intended to undergo DNA extraction were dried with silica gel and then stored at 4 °C. The mycelium of two Galerina strains were cultivated to grow for DNA and RNA extraction. Detailed information of the mushroom materials used in this study is given in Table 5.

Nucleic acid extraction and cDNA preparation

Total genomic DNA was extracted using the Fungal DNA Mini Kit (Omega Bio-tek, Norcross, USA).

Total RNA was isolated using TRIzol Reagent (Invitrogen, Carlsbad, USA) and following the TRIzol User
Guide. cDNA was synthesized using TransScript® One-Step gDNA Removal and cDNA Synthesis SuperMIX (Transgen Biotech, Beijing, China). The DNA and RNA quality and yield were detected using a SmartSpec Plus (Bio-Rad, Hercules, USA).

**Transcriptome sequencing and de novo assembly**

The concentration, purity and integrity of the RNA samples used for next-generation sequencing were further examined using an Agilent 2100 bioanalyzer (Agilent, Santa Clara, USA). Qualified RNA samples were used to construct circular single-stranded cDNA libraries, and then, the libraries were sequenced on a BGISEQ-500 sequencer (BGI, Shenzhen, China). Clean reads were obtained using the filtering software SOAPnuke to remove reads containing adaptors, reads with unknown bases accounting for more than 5%, and low-quality reads (bases with a quality value < 15 accounting for more than 20% in a read) from raw reads. These clean reads were de novo assembled using Trinity software. Finally, nonredundant unigenes were obtained using Tgicl software. All of these steps were completed by Beijing Genomic Institute (BGI)-Wuhan in China.

**Retrieval and annotation of MSDIN and POP genes**

The unigene data obtained above were searched for MSDIN and POP genes (*Galerina* unigenes were provided by Professor Zhang Ping at Hunan Normal University) by using the known amino acid sequences of the MSDIN family and POP genes from *A. bisporigera* and *G. marginata* [5, 7] as NCBI online TBLASTN tool queries. Then, unigenes similar to the queries were manually annotated, and the coding sequences were predicted and translated into protein sequences using DNAMAN 7.0 software.

**Cloning of MSDIN and POP genes**

Partial MSDIN gene sequences were amplified from *Amanita* genomic DNA using the following degenerate primers: forward (5’-ATGTCNGAYATYAAYGCNACNCG-3’) and reverse (5’-CCAAGCCTRAYAWRGTCMACAACC-3’), according to the method of Li et al. (2014) [13]. The PCR mixtures contained 1× PCR buffer, 1.5 mM MgCl₂, 0.2 mM dNTPs, 0.4 μM each primer, 1.25 U of Taq
polymerase (Comwin Biotech, Beijing, China), and 1 μL of DNA template in a total volume of 25 μL. PCR was performed with the following program: initial denaturation at 94°C for 4 min, 35 cycles at 94°C for 30 s, 52°C for 30 s, and 72°C for 30 s, and the reaction batches were incubated at 72°C for 2 min for terminal elongation.

Using the known MSDIN and POP genes from *A. bisporigera* and *G. marginata* as reference models [5, 7], specific primers (shown in Table S1) were designed to obtain target products close to the full lengths of the genes according to the flanking sequences of the CDS. The genomic DNA and cDNA of *Amanita, Galerina* and *Lepiota* species were used as templates, and PCR was performed as follows: initial denaturation at 94°C for 4 min, followed by 32 cycles of denaturation at 94°C for 30 s, 55-60 °C for 30 s (annealing temperature for each target shown in Table S1), extension at 72°C (MSDIN gene for 30 s, POP for 2 min), and a final extension at 72°C for 5 min.

All PCR products were detected by agarose gel electrophoresis and purified using an EasyPure Quick Gel Extraction Kit (Transgen Biotech, Beijing, China). The purified products were ligated into the pEASY®-Blunt Zero Cloning Vector (Transgen Biotech, Beijing, China) and transformed into competent cells. Positive clones to be sequenced were selected using Amp-resistance LB medium and further verified by colony PCR. Finally, all of the obtained gene genomic and coding sequences were used to manually predict the corresponding functions and structures by DNAMAN 7.0 software.

**Phylogenetic tree construction of MSDIN and POP genes**

Forty-six coding sequences (CDSs) of toxin MSDIN genes and fifty-eight CDSs of POP genes were used for phylogenetic analysis, and their source and GenBank accession numbers are presented in Table 4. These sequences were aligned by MAFFT v7.374 [33] and then manually adjusted by BioEdit [34]. HKY+I+G and GTR+I+G were inferred as the best-fit models for the CDSs of MSDIN and POP genes selected by AIC in MrModeltest v2.3 [35]. Maximum likelihood (ML) trees with 1000 bootstrap replicates and Bayesian inferences were carried out in RAxML v7 [36] and MrBayes v3.1.2 [37], respectively.

**Abbreviations**
Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

The mushroom species and related transcriptome data, MSDIN and POP genes used in this study has been deposited at GenBank and their accession numbers can be found in Table 1, 3, 4 and 5.

Competing interests

The authors declare that they have no competing interests.

Funding

This study was supported by the National Natural Science Foundation of China (Grant No. 31872616), and the Scientific Research Foundation of Hunan Provincial Education Department, China (Grant no. 18A026).

Author contributions

Zuohong Chen and Zhengmi He conceived and designed the experiments. Zhengmi He and Pan Long carried out the MSDIN and POP genes cloning, Fang Fang and Sainan Li carried out the phylogenetic analysis. Ping Zhang provided some Amanita materials and identified the species. Zhengmi He and Zuohong Chen wrote the manuscript.

Acknowledgements

The authors are very grateful to Dr. Yang ZL and Luo H (Kunming Institute of Botany, Chinese Academy of Sciences, China) for critically reviewing the manuscript.

References

1. Wieland T. Peptides of poisonous Amanita, New York: Springer; 1986.

2. Enjalbert F, Cassanas G, Rapior S, Chaumont JP, 2004. Amatoxins in wood-rotting
Galerina marginata. Mycologia. 2004;96:720-729.

3. Tang SS, Zhou Q, He ZM, Luo T, Zhang P, Cai Q, et al. Cyclopeptide toxins of lethal amanitas: compositions, distribution and phylogenetic implication. Toxicon. 2016;120:78-88.

4. Walton JD. The cyclic peptide toxins of Amanita and other poisonous mushrooms. New York: Springer; 2018.

5. Hallen HE, Luo H, Scott-Craig JS, Walton JD. 2007. Gene family encoding the major toxins of lethal Amanita Proc Nati Acad Sci USA. 2007;104:19097-19101.

6. Pulman JA, Childs KL, Sgambelluri RM, Walton JD. Expansion and diversification of the MSDIN family of cyclic peptide genes in the poisonous agarics Amanita phalloides and A. bisporigera. BMC Genomics. 2016;17:1038.

7. Luo H, Hallen HE, Scott-Craig JS, Walton JD. Ribosomal biosynthesis of α-amanitin in Galerina marginata. Fungal Genet Biol. 2012;49:123-129.

8. Luo H, Cai Q, Lüli YJ, Li X, Sinha R, Hallen HE, et al. The MSDIN family in amanitin-producing mushrooms and evolution of the prolyl oligopeptidase genes. IMA Fungus. 2018;9:225-242.

9. Lüli YJ, Cai Q, Chen ZH, Sun H, Zhu XT, Li X, et al. Genome of lethal Lepiota venenata and insights into the evolution of toxin-biosynthetic genes. BMC Genomics. 2019;20:198.

10. Walton JD, Hallen HE, Luo H. Ribosomal biosynthesis of the cyclic peptide toxins of Amanita Biopolymers. 2010;94:659-664.

11. Luo H, Hallen HE, Scott-Craig JS, Walton JD. Colocalization of amanitin and a candidate toxin-processing prolyl oligopeptidase in Amanita Eukaryot Cell. 2010;9:1891-1900.

12. Luo H, Hong SY, Sgambelluri RM, Angelos E, Li X, Walton JD. Peptide macrocyclization
catalyzed by a prolyl oligopeptidase involved in α-amanitin biosynthesis. Chem Biol. 2014;21:1610-1617.

13. Li P, Deng WQ, Li TH. The molecular diversity of toxin gene families in lethal Amanita Toxicon. 2014;83:59-68.

14. Yang ZL. Atlas of the Chinese species of Amanitaceae. Beijing: Science; 2015.

15. Cui YY, Cai Q, Tang LP, Liu JW, Yang ZL. The family Amanitaceae: molecular phylogeny, higher-rank taxonomy and the species in China. Fungal Divers. 2018;91:5-230.

16. Cai Q, Tulloss RE, Tang LP, Tolgor B, Zhang P, Chen ZH, et al. Multi-locus phylogeny of lethal amanitas: implications for species diversity and historical biogeography. BMC Evol Biol. 2014;14:143.

17. Zhang P, Chen ZH, Xiao B, Tolgor B, Bao HY, Yang ZL. Lethal amanitas of East Asia characterized by morphological and molecular data. Fungal Divers. 2010;42:119-133.

18. Li Hj, Xie JW, Zhang S, Zhou Yj, Ma PB, Zhou J, et al. Amanita subpallidorosea, a new lethal fungus from China. Mycol Prog. 2015;14:43.

19. Cai Q, Cui YY, Yang ZL. Lethal Amanita species in China. Mycologia. 2016;108:993-1009.

20. Cai Q, Chen ZH, He ZM, Luo H, Yang ZL. Lepiota venenata, a new species related to toxic mushroom in China. J Fungal Res. 2018;16:63-69.

21. Clarke DB, Lloyd AS, Robb P. Application of liquid chromatography coupled to time-of-flight mass spectrometry separation for rapid assessment of toxins in Amanita Anal Methods. 2012;4:1298.

22. Xue JH, Wu P, Chi YL, Xu LX, Wei XY. Cyclopeptides from Amanita exitialis. Nat Prod Bioprospect. 2011;1:52-56.

23. Li P, Deng WQ, Li TH, Song B, Shen YH. Illumina-based de novo transcriptome
sequencing and analysis of *Amanita exitialis* Gene. 2013;532:63-71.

24. Sgambelluri RM, Smith MO, Walton JD. Versatility of prolyl oligopeptidase B in peptide macrocyclization. ACS Synth Biol. 2018;7:145-152.

25. Yilmaz I, Bakirci S, Akata I, Bayram R, Kaya E. Toxin content and toxicological significance in different tissues and development stages of *Lepiota brunneoincarnata* Toxin Rev. 2015;34:109-114.

26. Wei JH, Wu JF, Chen J, Wu BD, He ZM, Zhang P, et al. Determination of cyclopeptide toxins in *Amanita sub pallidorosea* and *Amanita virosa* by high-performance liquid chromatography coupled with high-resolution mass spectrometry. Toxicon. 2017;133:26-32.

27. Keeling PJ, Palmer JD. Horizontal gene transfer in eukaryotic evolution. Nat Rev Genet. 2008;9:605-618.

28. Andersson JO. Gene transfer and diversification of microbial eukaryotes. Annu Rev Microbiol. 2009;63:177-193.

29. Richards TA, Leonard G, Soanes DM, Talbot NJ. Gene transfer into the fungi. Fungal Biol Rev. 2011;25:98-110.

30. Richards TA, Soanes DM, Jones MDM, Vasieva O, Leonard G, Paszkiewica K, et al. Horizontal gene transfer facilitated the evolution of plant parasitic mechanisms in the oomycetes. Proc Nati Acad Sci USA. 2011;108:15258-15263.

31. Savory F, Leonard G, Richards TA. The role of horizontal gene transfer in the evolution of the oomycetes. PLoS Pathog. 2015;11:e1004805.

32. Dhillon B, Feau N, Aerts AL, Beauseigle S, Bernier L, Foster A, et al. Horizontal gene transfer and gene dosage drives adaptation to wood colonization in a tree pathogen. Proc Nati Acad Sci USA. 2015;112:3451-3456.

33. Katoh K, Standley DM. A simple method to control over-alignment in the MAFFT
multiple sequence alignment program. Bioinformatics. 2016;32:1933-1942.

34. Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for windows 95/98/NT. Nucleic Acids Sym Ser. 1999;41:95-98.

35. Nylander JAA. MrModeltest v2.2 Uppsala: evolutionary biology Centre, Uppsala University. 2004.

36. Stamatakis A. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics. 2006;22:2688-2690.

37. Ronquist F, Huelsenbeck JP. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics. 2003;19:1572-1574.

Tables

Table 1 Features and accession numbers of transcriptomes.

| Species          | Total clean bases (Gb) | Q30 (%) | Total number of unigene | Total length of unigene (nt) | Mean length of unigene (nt) | N50  |
|------------------|------------------------|---------|-------------------------|-------------------------------|----------------------------|------|
| A. excitialis    | 6.67                   | 86.25   | 24,578                  | 48,383,999                    | 1,968                      | 2,891|
| A. fuliginea     | 6.55                   | 88.20   | 21,624                  | 36,817,429                    | 1,702                      | 2,599|
| A. molliuscula   | 6.32                   | 90.65   | 46,471                  | 79,566,952                    | 1,712                      | 3,007|
| A. oberwinklerana| 6.59                   | 86.87   | 24,326                  | 61,864,918                    | 2,543                      | 3,993|
| A. pallidorosea  | 6.25                   | 89.78   | 36,846                  | 79,216,743                    | 2,149                      | 3,375|
| A. rmosa         | 6.57                   | 87.93   | 22,532                  | 36,712,344                    | 1,629                      | 2,648|
| A. subpallidorose| 10.24                  | 91.21   | 42,803                  | 110,323,057                   | 2,577                      | 3,630|
| L. venenata      | 8.47                   | 93.83   | 13,859                  | 21,738,818                    | 1,569                      | 2,994|

Table 2 MSDIN family members searched from the transcriptomes of seven amanitin-producing mushrooms.

| Name             | No. | Leader peptide | Core peptide | Recognition sequence | Prod |
|------------------|-----|----------------|--------------|----------------------|------|

19
### A. exitialis

|   | Amino Acid Sequence |   |   |   |
|---|---------------------|---|---|---|
| Ae1 | MTDINDTRLP | FIWLLWIWL | SVGDDNNILNMRGEDLC* |
| Ae2 | MSDINATRLP | LFPFPDFRPP | CVGDADNFTLTRGENLC* |
| Ae3 | MSDVNAVTRLP | FNFFRFYP | CIGDDGSALRLGESLC* |
| Ae4 | MSDINTARLP | IPvPPFFIP | FVGDDIDVVLRRGENLC* |
| Ae5 | MSDINVTRLP | VFIFFIIP | CVGDGTADIRKGQENLC* |
| Ae6 | MSDINTARLP | VFSLPVF | FVSDDIQAVLTRGLAESLC* |
| Ae7 | MSDINTTRLP | FVFVASPP | CVGDDIAMVLTRGLAESLC* |
| Ae8 | MSDINPTRLP | IFWFIYFP | CVSDVDSTLTLGISLS* |
| Ae9 | MSDINTARLP | IIWIIGNP | CVSDVERILTRGLAESLC* |
| Ae10 | MSDINATRLP | IIWAPVVP | CISDDNDSTLTRGLAESLC* |
| Ae11 | MSDINATRLP | IGRPOPLL | CVGGDVNYILSGLAESLC* |
| Ae12 | MSDINATRLP | IWGIGCDP | CVGDDVTALLTRGLAESLC* |
| Ae13 | MSDINATRLP | IWGIGCNP | CVGDDVTSVTRGLAESLC* |
| Ae14 | MSDINVIRLP | SMLTILPP | CVSSDAORTLTRGLAESLC* |
| Ae15 | MSDINATRLP | AWLTDCP | CVGDDVNRLTRGLAESLC* |
| Ae16 | MSDINATRLP | AWLVDCLP | CVGDDVNRLTRGLAESLC* |
| Ae17 | MSDINLTRLP | GIIAIIP | CVGDDVNSTLTRGLAESLC* |
| Ae18 | MSDINATRLP | VWIGYSP | CVGDDCIALTRGLAESLC* |
| Ae19 | MSDINATRLP | GFLFWA | YVGDDVDYITRGLAESLC* |

### A. fuliginea

|   | Amino Acid Sequence |   |   |   |
|---|---------------------|---|---|---|
| Af1 | MSDINATRLP | IIIVLGLIIP | LCVSDIEMILTRGLAESLC* |
| Af2 | MSDLNASRLP | ILSVLGLPVP | HVGEETNSTLARGVSLCGE |
| Af3 | MSDINSARLP | LFPFPDFRPP | CVSDVQVONGLAESLC* |
| Af4 | MSDINAARLP | FPFPFF | CIGDDTSIVQRAGNLGEG |
| Af5 | MSDINTIRIP | FPWTGFVP | CVSDVGSVMRTRGLAESLC* |
|     | Sequence 1     | Sequence 2     | Sequence 3     |
|-----|----------------|----------------|----------------|
| Af6 | MSDTNATRLP     | IWFQIQIP       | CAGDDVNSSLTREGESLC* |
| Af7 | MSDINVTRLP     | VLVFQFP       | YISDDAVNILKQGENLC* |
| Af8 | MFDINGSRLP     | AFRILPPP      | CVGDDVDSTLTSGESLC* |
| Af9 | MSDINATRLP     | GILIVFPP      | CVGDDVNSTLTRGESLC* |
| Af10| MSDINATRLP     | HLFTWIP       | CISDDSTLTRGESFC*  |
| Af11| MFDINSSRLP     | HLYPNRSP      | CVCCDACSTLSTAESLC* |
| Af12| MSDINATRLP     | IFWFIYFP      | CVGDDVDNLTRGESLS*  |
| Af13| MSDINATRLP     | IWGIGCDP      | CVGDDVAALITRGEALC* |
| Af14| MSCIATRLP      | LPSVPVF       | FVSDAIEVVLGRGELDC* |
| Af15| MSDINSLRLP     | VVNSRFNP      | CVGDDVSPSLTRGEGLC* |
| Af16| MSDINASRLP     | AWLACCP       | CIGDDVPNPTTRGESLC* |
| Af17| MSDINATRLP     | AWLDCP        | CVGDDVNLRLARGENLC* |
| Af18| MSDINATRLP     | AWLVTCR       | CVGDDINRLTRGERNC*  |
|     |                |                | β-amanitin      |
|     |                |                | phallidin       |
|     |                |                | phalla          |
|     |                |                | phallo          |

**A. molliscula**

|     | Sequence 1     | Sequence 2     | Sequence 3     |
|-----|----------------|----------------|----------------|
| Am1 | MSDINTARLP     | YFLPPIFSPP     | CVSDIEMVLTREGENLC* |
| Am2 | MTDINATRLP     | ILFGFLLP       | CVGVDNLTHSGENLC*  |
| Am3 | MSNIASRLP      | IWAAFFRP       | CVGDEVGDILRSGESLC* |
| Am4 | MSDINATRLA     | IWGICDP        | CVGDDVTALLTRGEALC* |
| Am5 | MSDINASRLP     | RLLVPRYP       | CIDEDAEGATYLC*   |
| Am6 | MSNIARLIP      | GFFAVVP        | YLATSITFSLLGRGESC*  |
| Am7 | MTDINATRLP     | WIFFFPP        | CVGVDNLTHSGENLC*  |
| Am8 | MSNIALRLP      | GFFFP          | YASGDVYTLTREGSL*  |
| Am9 | MSDINATRFP     | GKVNNP         | YVGDDVDIIIIRGELC*  |

**A. pallidorosea**

|     | Sequence 1     | Sequence 2     | Sequence 3     |
|-----|----------------|----------------|----------------|
| Ap1 | MADINAARLP     | FHGLFPPLPPP    | FVDDDATSTLTRGESLC* |
| Ap1  | MADINASRLP | LNILPFHLPP | CVSDDATSTLTRGLSLC* |
| Ap2  | MSDINATRLP | NWHAGPRTRP | CVADDVSSLTRGLSLC* |
| Ap3  | MSDINATRLP | VFFMPFPFPP | CVSDDIQMVTRGLSLC* |
| Ap4  | MSDINATRLP | EFIIVGIFFP | CVGDDQTGTRGLSLC* |
| Ap5  | MSDINATRLP | FFPEVGFPP | CVGDDTNPTTRGLSLC* |
| Ap6  | MSDINATRLP | FNLFRFPPP | CVGDDGSVLTRGLSLC* |
| Ap7  | MSDINATRLP | FPWTGFPFVP | CVGDDGNSVLTRGLSLC* |
| Ap8  | MSDINATRLP | HPFPLGQP | CVGDDQSVGSLTRGLSLC* |
| Ap9  | MSDINATRLP | DPPRLILP | CVGDDVSDQLTRGLSLC* |
| Ap10 | MSDINATRLP | HFFNLMPR | CVGDDIETVTRGLSLC* |
| Ap11 | MSDINATRLP | HQQHPFVP | CVGDDVGSTTRGLSLC* |
| Ap12 | MSDINATRLP | IWGIGCDP | CVGDDVTAVTRGLSLC* |
| Ap13 | MSDINATRLP | IWGIGCNP | CVGDEVAALLTRGLSLC* |
| Ap14 | MSDINATRLP | IWGIGCNP | CVGDEVTALTRGLSLC* |
| Ap15 | MSDINATRLP | IWGIGCNP | CVGDDVASLTRGLSLC* |
| Ap16 | MSDINATRLP | IWGIGCNP | CVGDDVSDLTRGLSLC* |
| Ap17 | MSDINATRLP | LGRPESLP | CVGDDVNYILVSGlobal |
| Ap18 | MSDINATRLP | LVYMILFP | CVGDDVYLMVGSGlobal |
| Ap19 | MSDINATRLP | MAFPEFLA | CVGDDVHGLTRGLSLC* |
| Ap20 | MSDINATRLP | MHIAPP | CVGDDMVATTRGLSLC* |
| Ap21 | MSDINATRLP | NLFWVIPP | CVGDDMVATTRGLSLC* |
| Ap22 | MSDINATRLP | YMWDHHL | CVGDDMVATTRGLSLC* |
| Ap23 | MSDINATRLP | AWLATCP | CVGDDNPTTRGLSLC* |
| Ap24 | MSDINATRLP | AWLMTCP | CVGDDNPTTRGLSLC* |
| Ap25 | MSDINATRLP | AWLVDIP | CVGDDINRLTRGLSLC* |

β-amînótoxin, α-amînótoxin, phalloidin, phallotoxin, phallacidin.
| Protein   | Amino Acid Sequence | Peptide Sequence | N-Terminal Peptide | C-Terminal Peptide |
|-----------|---------------------|------------------|--------------------|------------------|
| Ar1       | MSDINTSRPLP         | FIPLGIIILTP      | CVSDDVNTTITRGESLC*|                  |
| Ar2       | MTDINSDTRLP         | FVWILWLWLA       | CVGDDTSILNRGEDLC* |                  |
| Ar3       | MSDINTATRILP        | IIIVGLIIP        | LCVSDIEMILTRGESLC*|                  |
| Ar4       | MSDVNTTRILP         | FNFFRFPYP        | CICDDSEKVLGENLC*  |                  |
| Ar5       | MSDINTATRILP        | HPFPLGLQP        | CAGDVNFTVSCHSLC*  |                  |
| Ar6       | MLDINTARFLP         | LGVPTTPLP        | CVGDDVNYILIGNGENLC*|                  |
| Ar7       | MSDINASCLP          | LILVANGMA        | YVSDDVSPTELTRGENLC*|                  |
| Ar8       | MPDINTVTRLP         | LLIVLLTP         | CISDDNINLRGKDL*   |                  |
| Ar9       | MSDIHAARLP          | FPTRPVFP         | SAGDDMIEVLGRGDEL* |                  |
| Ar10      | MSDINNARLP          | FYFYLGIN         | SDDAHPILTRGESLC*  |                  |
| Ar11      | MSDINJARLP          | IFWFIDFP         | CVGDDVDNTLSRGESLC*|                  |
| Ar12      | MSDINASRLP          | ILKKPWAP         | SVCDDVNSLTRGLC*   |                  |
| Ar13      | MSDINVARLP          | ISDPTAYP         | CVGDDIQAVKRGESLC* |                  |
| Ar14      | MSDINTARLP          | IWGIGCDP         | CVGDDVALTRGRGELC* | β-aman*         |
| Ar15      | MSDINSTRLP          | IWGIGCNP         | SVGDEVTALLTRGRGELC*| α-aman*         |
| Ar16      | MSDINTARLP          | AWDSKHP          | CVGDDVSRLLTRGLC*  |                  |
| Ar17      | MSDINATRVP          | AOWLAECP         | CVGDDISHLLTRGLC*  | phallo           |
| Ar18      | MSDINATRVP          | AOWLVDCP         | CVGDDISRLTRGLC*   | phalla           |

| Protein   | Amino Acid Sequence | Peptide Sequence | N-Terminal Peptide | C-Terminal Peptide |
|-----------|---------------------|------------------|--------------------|------------------|
| Asp1      | MTDVNDTRLP          | FIWLIWLWLP       | SVGDDINILNGGEDLC* |                  |
| Asp2      | MTDINYARLP          | ITLFLFFIP        | CLSDDNNINLRGKDL*  |                  |
| Asp3      | MSDINTARLP          | YFLPPIFSPP       | CVSDDIEMVTRGLC*   |                  |
| Asp4      | MSDINTARLP          | HPFPLGLQP        | CAGDVNFTLTGRGELC* |                  |
| Asp5      | MSDINTARLP          | GILIVWP          | CVGDDVNTTRGLC*    |                  |
| Asp6      | MSDINTTRLRPLP       | IAFPEFIA         | RVGDDIHRTLTRGLC*  |                  |
Asp7  MSDINVTRLP  IFWFIYFP  CVGDDVNDTLTRGESLS*
Asp8  MSDINARLP  IGRPENKP  CVGDDNVYILISGEKLC*
Asp9  MSDINATRLP  IVFLEFYS  CVGDDVNSTLTRGESLC*
Asp10 MSDINATRLP  IWGIGCDP  CVGDDVAFLTRGEALC*  β-amanitin
Asp11 MSDINATRLP  IWGIGCNP  SVGDEVTALLTRGEALC*  α-amanitin
Asp12 MSDINASRLP  VIGLFGLP  YVSDDVPILTRGDSLC*
Asp13 MSDINASRLP  VIPFLLPP  CVSDDVNFTLTRGESLC*
Asp14 MSDINATRLP  YFRPAPPP  CVSDDINPILTCGESLC*
Asp15 MSDINAARLP  AWITDCP  CVGDDINRILTRGENIC*  "phallo"
Asp16 MSDINASRFP  AWLATCP  CVGDDVNPITARGESLC*  phallotoxin
Asp17 MSDINATRLP  AWLVTCP  CVGDDVNFTLTRGESLC*  "phallo"
Asp18 MSDINATRLP  AWLVTCP  CVGDDVNPITTTRGESLC*  "phallo"
Asp19 MSDINTIRIP  GPFGFA  YVGDEVENLLKRGESLS*

L. venenata
Lv1  MDANATRLP  IWGIGCNP  WTPESVNDLTLDLS*  α-amanitin
Lv2  MDANSTRLP  IWGIGCNP  WAPESVNDLTGKDLSC*  α-amanitin

The MSDIN members with underlined numbers were verified at the genomic level. “Phallotoxin” means a novel heptapeptide similar to the phallotoxin cyclopeptide and capable of containing tryptathione (Trp-Cys).

Table 3 MSDIN family members cloned from genomic DNA of twelve amanitin-producing mushrooms.

| Name  | No. | Leader peptide | Core peptide | Recognition sequence | Product       |
|-------|-----|----------------|--------------|----------------------|---------------|
| A. exitialis |     |                |              |                      |               |
| Ae1a  |     | MSDINATRLP     | FIWVFGIP     | GDIGTVLTRGENLC*      |               |
| Ae2a  |     | MSDINATRLP     | IIWIIGNP     | CVSDDVERILTRGESLC*   |               |
| Ae3ab |     | MSDINATRLP     | IWGIGCDP     | CVGDDVTALLTRGEALC*   | β-amanitin    |
| Ae4ab |     | MSDINATRLP     | IWGIGCNP     | CVGDDVTSVLTRGEALC*   | α-amanitin    |
| Ae5<sup>b</sup> | MSDINATRLP | AWLTDCP | CVGDDVNRLTRGESLC* | "phallotoxins" |
| Ae6<sup>b</sup> | MSDINATRLP | AWLVDCP | CVGDDVNRLTRGESLC* | phallacidin |
| Ae7<sup>a</sup> | MSDINATRLP | VVIGYSP | CVGDDCIALLTRGEGLC* | |
| Ae8<sup>a</sup> | MSDINATRLP | GFLFA | YVGDDVYILTRGESLA* | |
| Ae9<sup>a</sup> | MSDINATRLP | GFLLA | YVGDDVYILTRGESLA* | |

A. fuliginea

| Af1<sup>a</sup> | MSDINATRLP | FPFPYNNPP | CVSDDIHMVLTRGENLC* | |
| Af2<sup>a</sup> | MSDINATRLP | YLLLLLPP | CVSDDLQTVLTRGENLC* | |
| Af3<sup>a</sup> | MSDINATRLP | IFWIYFP | CVGDDVDNLARGEGLC* | |
| Af4<sup>b</sup> | MSDINATRLP | IWGIGCDP | CVGDDVAALLTRGEALC* | β-amanitin |
| Af5<sup>a</sup> | MSDINATRLP | IWGIGCDP | CVGDDVAALLTRGEALC* | β-amanitin |
| Af6<sup>a</sup> | MSDINATRLP | IWGIGCNP | SVGDEVAALLTRGENLC* | α-amanitin |
| Af7<sup>a</sup> | MSDINATRLP | LPSRPVFP | FVSDAIYVLGRGEDI* | |
| Af8<sup>b</sup> | MSDINASRLP | AWLATCP | CIGDDVNPTLTRGESLC* | phalloidin |
| Af9<sup>ab</sup> | MSDINATRLP | AWLVDCP | CIGDDVNLLARGEGLC* | phallacidin |

A. molliuscula

| Am1<sup>b</sup> | MSDINATRLA | IWGIGCDP | CVGDDVTALLTRGEALC* | β-amanitin |

A. pallidorosea

| Ap1<sup>a</sup> | MSDINATRLP | LIIFPPFIPP | CVSDDIOMVLTRGENLC* | |
| Ap2<sup>a</sup> | MSDINAPRLP | LIIFPPFIPP | CVSDDIOMVLTRGEGLC* | |
| Ap3<sup>a</sup> | MSDINATRLP | IPFHIPAP | SVGDEVYVLGRGENLC* | |
| Ap4<sup>a</sup> | MSDINATRLP | IWGIGCDP | CVGDDVTALLTRGEALC* | β-amanitin |
| Ap5<sup>b</sup> | MSDINATRLP | IWGIGCDP | CVGDDVTALLTRGEALC* | β-amanitin |
| Ap6<sup>ab</sup> | MSDINATRLP | IWGIGCNP | CVGDEVAALLTRGEALC* | α-amanitin |
|   |   |   |   |   |   |   |
|---|---|---|---|---|---|---|
| Ap7<sup>b</sup> | MSDINATRLP | IWGIGCNP | CVGDEVTLTIRGEALC* | α-amanitin |
| Ap8<sup>a</sup> | MSDINATRLP | AWLATCP | CAGDDVNPTLTRGESLC* | phalloid |
| Ap9<sup>a</sup> | MSDINATRLP | AWLMTCP | CVGDDVNPILTRGESVC* | "phalloto: |
| Ap10<sup>b</sup> | MSDINATRLP | AWLMTCP | CVGDDVNPTLTRGESLC* | "phalloto: |
| Ap11<sup>ab</sup> | MSDVINATRLP | AWLVDCP | CVGDDINRLLTRGENLC* | phallacid |

**A. rimosas**

|   |   |   |   |   |   |   |
|---|---|---|---|---|---|---|
| Ar1<sup>a</sup> | MSDINATRLP | IWGIGCDP | CVGDDVAAALATRGEALC* | β-amanin |
| Ar2<sup>ab</sup> | MSDINATRLP | IWGIGCDP | CVGDDVAAATTRGEALC* | β-amanin |
| Ar3<sup>a</sup> | MSDINATRLP | IWGIGCNP | SVGDEVTALLASGEALC* | α-amanin |
| Ar4<sup>ab</sup> | MSDINATRLP | IWGIGCNP | SVGDEVTALLTRGEALC* | α-amanin |
| Ar5<sup>b</sup> | MSDINATRVP | AWLAECQ | CVGDDISLHLLTRGENLC* | "phalloto: |

**A. subfuliginea**

|   |   |   |   |   |   |   |
|---|---|---|---|---|---|---|
| Asf1<sup>a</sup> | MSDINATRLP | HPFPLGLQP | CAGDVFNFTLTKGEGLC* |
| Asf2<sup>a</sup> | MSDINATRLP | AIIFLAWPP | CVGDNVNSLTRGESLC* |
| Asf3<sup>a</sup> | MSDINATRLP | IWGIGCDP | CVSDDVAALLTRGEALC* | β-amanin |
| Asf4<sup>a</sup> | MSDINATRLP | IWGIGCNP | CVGEVAAALLTRGEALC* | α-amanin |
| Asf5<sup>a</sup> | MSDINATRLP | AWLVDCP | CVGDDVNRLTRGENLC* | phallacid |

**A. subjunquillea**

|   |   |   |   |   |   |   |
|---|---|---|---|---|---|---|
| Asj1<sup>a</sup> | MSDINATRLP | AYLPFLFIPP | CVSSDEIEMVLRGESLC* |
| Asj2<sup>a</sup> | MSDINATRLP | AYLPFLFIPP | CVSDDEIEVVLTRGESLC* |
| Asj3<sup>a</sup> | MSDINATRLP | IWGIGCDP | CIGDDVTALLTRGEALC* | β-amanin |
| Asj4<sup>a</sup> | MSDINATRLP | IWGIGCDP | CVGDEVTAALLTRGEALC* | β-amanin |
| Asj5<sup>a</sup> | MSDINATRLP | IWGIGCNP | CVGDEVAAALLTRGEALC* | α-amanin |
| Asj6<sup>ab</sup> | MSDINATRLP | AWLATCP | CAGDDVNPTLTRGESLC* | phalloid |
| Asj7<sup>a</sup> | MSDINATRLP | AWLATCP | CVGDDVNPTLSRGRESLC* | phalloid |
|        | A. subpallidorosea                                                                 |        | G. marginata                                                                 |        | G. sulciceps                                                                 |
|--------|-----------------------------------------------------------------------------------|--------|----------------------------------------------------------------------------|--------|----------------------------------------------------------------------------|
|        |                                                                                   |        |                                                                            |        |                                                                            |
| Asp1ab | MSDINATRLP  IWGIGCDP  CVGDDVAFLLTRGEALC*                                        | β-aman | Gm1b  MFDTNATRLP  IWGICNPG  WTAEHVDQTLSGNDIC*                                | α-aman | Gs1b  MFDTNATRLP  IWGICNP  WTAEHVDQTLVSGNDIC*                                |
| Asp2ab | MSDINATRLP  IWGIGCNP  SVGDEVTALLLTRGEALC*                                        | α-aman | Gm2b  MFDTNSTRLP  IWGICNP  WTAEHVDQTLVSGNDIC*                                | α-aman | Gs2b  MFDTNSTRLP  GIGICNP  WTAEHIDQTLVSGNDTC*                                |
| Asp3ab | MSDINAARLP  AWITDCP  CVGDDINRILTRGENIC*                                          | “phalloto:” |                                                                            |        |                                                                            |
| Asp4ab | MSDINASRFP  AWLATCP  CVGDDVNPTIARGESLC*                                          | phalloid |                                                                            |        |                                                                            |
| Asp5a  | MSDINATRLP  AWLITCP  CVGDDANPTITRGESLC*                                          | “phalloto:” |                                                                            |        |                                                                            |
| Asp6ab | MSDINATRLP  AWLVTCP  CVGDDVNPTITRGESLC*                                          | “phalloto:” |                                                                            |        |                                                                            |
| Asp7a  | MSDINATRLP  AWLVTCP  CVGDDVNSTITRGESLC*                                          | “phalloto:” |                                                                            |        |                                                                            |
|        |                                                                                   |        |                                                                            |        |                                                                            |
| Av1a   | MSDINATRLP  FLLFIIPP  CVSDDVNSTLTRGESLC*                                         |        |                                                                            |        |                                                                            |
| Av2a   | MSDINATRLP  FYFQPGFP  WSVGDDVNPTLTRGESLC*                                        |        |                                                                            |        |                                                                            |
| Av3b   | MSDINATRLP  IWGIGCNP  SVGDEATALLLTRGEALC*                                        | α-aman |                                                                            |        |                                                                            |
| Av4a   | MSDINATRLP  SILIVWPP  CVGDDVNSTLTRGESLC*                                        |        |                                                                            |        |                                                                            |
| Av5a   | MSDINATRLP  SILVVWPP  CVSDDVNSTLTRGESLC*                                        |        |                                                                            |        |                                                                            |
| Av6a   | MSDINATRLP  AWLATCP  CVGDDVNPTLARGESLC*                                          | phalloid |                                                                            |        |                                                                            |
| Av7a   | MSDINATRLP  AWLVDCP  CVGDDINRLLTRGENLC*                                          | phallacid |                                                                            |        |                                                                            |
| Av8a   | MSDINATRLP  AWLVTCP  CVGDDVNPTLTRGESLC*                                          | “phalloto:” |                                                                            |        |                                                                            |
| Av9a   | MSDINATRLP  GPFLFFP  FVSDDIEVILRRGEDLC*                                          |        |                                                                            |        |                                                                            |
|        |                                                                                   |        |                                                                            |        |                                                                            |
|        |                                                                                   |        |                                                                            |        |                                                                            |
Superscripts a and b are for products cloned with degenerate and specific primers, respectively. “Phallotoxin” means a novel heptapeptide similar to phallotoxin cyclopeptide and capable of containing Tryptathione (Trp-Cys).

Table 4 Gene sequences used in the molecular phylogenetic analyses and their GenBank accession numbers.

| Taxon                        | Gene   | Source             | GenBank accession |
|------------------------------|--------|--------------------|-------------------|
| Agaricus bisporus var. bisporus | POP    | NCBI               | XM0064597         |
| Agaricus bisporus var. burnettii | POP    | NCBI               | JH971409          |
| Amanita bisporigera          | α-AMA1 | Pulman et al, 2016 | -                 |
|                              | α-AMA2 | Pulman et al, 2016 | -                 |
|                              | PHA1   | Pulman et al, 2016 | -                 |
|                              | PHA2   | Pulman et al, 2016 | -                 |
|                              | POPA   | Pulman et al, 2016 | -                 |
|                              | POPB   | Pulman et al, 2016 | -                 |
| Amanita exitialis            | α-AMA  | Our study          | MN26422€          |
|                              | β-AMA  | Our study          | MN26422€          |
|                              | PHA    | Our study          | MN26423€          |
|                              | “AWLTDCP” | Our study   | MN26423€          |
|                              | POPA   | Our study          | MN26423€          |
|                              | POPB   | Our study          | MN26424€          |
| Amanita fuliginea            | β-AMA  | Our study          | MN26422€          |
|                              | PHA    | Our study          | MN26423€          |
|                              | PHD    | Our study          | MN26424€          |
|                              | POPA   | Our study          | MN26423€          |
|                              | POPB   | Our study          | MN26424€          |
| Amanita molliuscula          | β-AMA  | Our study          | MN26422€          |
|                              | POPA   | Our study          | MN26424€          |
|                              | POPB   | Our study          | MN26424€          |
| Amanita muscaria             | POPA   | NCBI               | KN818232          |
| Amanita oberwinklerana       | POPA   | Our study          | MN264241          |
| Amanita pallidorosea         | α-AMA1 | Our study          | MN264221          |
|                              | α-AMA2 | Our study          | MN264222          |
|                              | β-AMA  | Our study          | MN26422€          |
| Genus          | Strain | Source                                    | Accession   |
|---------------|--------|-------------------------------------------|-------------|
| Amanita phalloides | α-AMA  | Pulman. et al., 2016                      | MN264233    |
|                | β-AMA1 | Pulman. et al., 2016                      | MN264236    |
|                | β-AMA2 | Pulman et al., 2016                       | MN264242    |
|                | PHA    | Pulman et al., 2016                       | MN264247    |
|                | PHD1   | Pulman et al., 2016                       | MN264250    |
|                | PHD2   | Pulman et al., 2016                       | MN264251    |
|                | PHD3   | Pulman et al., 2016                       | MN264252    |
|                | POPA   | Pulman et al., 2016                       | MN264253    |
|                | POPB   | Pulman et al., 2016                       | MN264254    |
| Amanita rimosa  | α-AMA  | Our study                                 | MN264222    |
|                | β-AMA  | Our study                                 | MN264223    |
|                | “AWLAECP” | Our study                                | MN264231    |
|                | POPA   | Our study                                 | MN264243    |
|                | POPB   | Our study                                 | MN264244    |
| Amanita subjunquillea | α-AMA | Luo et al., 2018                          | MN264234    |
|                  | β-AMA1 | Luo et al., 2018                          | MN264235    |
|                  | β-AMA2 | Luo et al., 2018                          | MN264236    |
|                  | PHA    | Our study                                 | MN264237    |
|                  | PHD    | Our study                                 | MN264238    |
|                  | POPA   | Luo et al., 2018                          | MN264239    |
|                  | POPB   | Luo et al., 2018                          | MN264240    |
| Amanita subpallidorosea | α-AMA | Our study                                 | MN264241    |
|                   | β-AMA  | Our study                                 | MN264242    |
|                   | PHD    | Our study                                 | MN272401    |
|                   | “AWITDCP” | Our study                               | MN272402    |
|                   | “AWLVTCP” | Our study                              | MN272403    |
|                   | POPA   | Our study                                 | MN272404    |
|                   | POPB   | Our study                                 | MN272405    |
| Amanita thiersii | POPA   | NCBI                                      | KZ302001    |
| Amanita virosa   | α-AMA  | Our study                                 | MN272412    |
| Anomoporia bombycina | POP  | JGI                                       |             |
| Auriculariopsis ampla | POP | NCBI                                      | VDMD010000   |
| Bolbitius vitellinus | POP | JGI                                       |             |
| Conocybe apala   | POP    | NCBI                                      | FJ906819    |
| Coprinellus micaceus | POP | NCBI                                      | QPFP010000   |
| Species                        | Accession | Source     |
|--------------------------------|-----------|------------|
| Coprinopsis cinerea           | POP       | NCBI       | XM0018411 |
| Coprinopsis marcescibilis     | POP       | NCBI       | ML210154  |
| Cortinarius glaucopus         | POP       | JGI        | -         |
| Crucibulum laeve              | POP       | NCBI       | ML213591  |
| Cyathus striatus              | POP       | JGI        | -         |
| Fistulina hepatica            | POP       | NCBI       | KN881639  |
| Galerina marginata            | a-AMA1    | Our study  | MN27241\(1\) |
|                               | a-AMA2    | Our study  | MN27241\(4\) |
|                               | POPA      | Our study  | MN27241\(5\) |
|                               | POPB      | Our study  | MN27241\(6\) |
| Galerina sulciceps            | a-AMA1    | Our study  | MN27241\(7\) |
|                               | “a-AMA2”  | Our study  | MN27241\(8\) |
|                               | POPA      | Our study  | MN27241\(9\) |
|                               | POPB      | Our study  | MN27242\(0\) |
| Gymnopilus chrysopellus       | POP       | JGI        | -         |
| Gymnopilus dilepis            | POP       | NCBI       | NHYE010055 |
| Hebeloma cylindrosporum       | POP       | NCBI       | KN831777  |
| Hypholoma sublateritium       | POP       | NCBI       | KN817688  |
| Hypsizygus marmoreus          | POP       | NCBI       | LUEZ020002 |
| Laccaria amethystina          | POP       | NCBI       | KN838546  |
| Laccaria bicolor              | POP       | NCBI       | DS547115  |
| Lepiota brunneoincarnata      | POPA      | NCBI       | MN91269\(5\) |
| Lepiota subincarnata          | a-AMA1    | Luo et al, 2018 | - |
|                               | a-AMA2    | Luo et al, 2018 | - |
|                               | POPB      | Luo et al, 2018 | - |
| Lepiota venenata              | a-AMA1    | Our study  | MN27242\(1\) |
|                               | a-AMA2    | Our study  | MN27242\(2\) |
|                               | POPA      | Our study  | MN27242\(3\) |
|                               | POPB      | Our study  | MN27242\(4\) |
| Lepista nuda                  | POP       | JGI        | -         |
| Leucoagaricus sp.             | POP       | NCBI       | KQ962668  |
| Macrolepiota fuliginosa       | POP       | JGI        | -         |
| Panaeolus cyanescens          | POP       | NCBI       | NHTK010055 |
| Pleurotus ostreatus           | POP       | NCBI       | KL198007  |
| Plicaturopsis crispa          | POP       | JGI        | -         |
| Pterula gracilis              | POP       | NCBI       | ML178816  |
| Schizophyllum commune          | POP       | NCBI       | GL377318  |
| Termitomyces sp.              | POP       | NCBI       | KQ412502  |

**References:**
- Luo et al., 2018
- Our study
Table 5 Information of the mushroom materials used in this study.

| Species name       | Locality        | Collection time | Specimen no. | GenBank accession no. |
|--------------------|-----------------|-----------------|--------------|-----------------------|
| A. exitialis        | Guangdong, China| 2017-03-27      | MHHNU 30937  | KR996717              |
| A. fuliginea        | Hunan, China    | 2017-06-06      | MHHNU 9047   | MN061271              |
| A. molluscula       | Jilin, China    | 2017-08-07      | MHHNU 9142   | MN061272              |
| A. oberwinklerana   | Hunan, China    | 2017-06-09      | MHHNU 9051   | MN061273              |
| A. pallidorosea     | Shandong, China | 2018-08-13      | MHHNU 31203  | MN061274              |
| A. rimoso           | Hunan, China    | 2017-06-09      | MHHNU 9050   | MN061275              |
| A. subfuliginea     | Chongqing, China| 2015-07-01      | MHHNU 30946  | MN061276              |
| A. subjunquillea    | Hunan, China    | 2012-09-10      | MHHNU 7751   | KR996715              |
| A. subpallidorosea  | Hunan, China    | 2017-09-14      | MHHNU 8617   | KU601411              |
| A. virosa           | Hunan, China    | 2016-09-09      | MHHNU 8621   | KY472227              |
| G. marginata        | -               | -               | MHHNU 8380   | MN061277              |
| G. sulciceps        | -               | -               | MHHNU 7669   | KX214585              |
| L. venenata         | Hubei, China    | 2017-09-10      | MHHNU 31031  | MK095189              |

G. marginata and G. sulcipes samples were cultured mycelia, and the other mushroom samples were wild fruiting bodies.

Figures
Alignment of MSDIN precursor peptide sequences. (A) WebLogo alignment of 145 MSDIN members from 9 Amanita species. The letter height of each amino acid represents its conservation degree, and the higher the letter, the more conserved the site. (B) Alignment of the precursor peptide sequences of α-amanitin from 15 amanitin-producing mushrooms. (C) Alignment of the precursor peptide sequences of α-, β-amanitin, phallacidin and phalloidin from 12 Amanita species. Letters with white background are variations compared with the consensus sequence. The sequences of A. bisporigera, A. phalloides were from Pulman et al. (2016), the sequences of A. fuligineoides were from Li et al. (2014) and the sequences of L. brunneoincarnata were from Lüli et al. (2019).
Structures of α-AMA and POP exemplified by four agaric species.
Figure 3

Phylogenetic trees generated from maximum likelihood analysis based on toxin MSDIN genes. Bootstrap percentages (>50%) based on 1,000 replications and Bayesian posterior probabilities (>0.90) are shown at nodes. Bar, a substitution per 10 nucleotides.
Figure 4

Phylogenetic trees generated from maximum likelihood analysis based on POP genes.

Bootstrap percentages (>50%) based on 1,000 replications and Bayesian posterior probabilities (>0.90) are shown at nodes. Bar, a substitution per 10 nucleotides.

Supplementary Files
This is a list of supplementary files associated with this preprint. Click to download.
Supplementary Material.pdf