



Genome-Wide Analysis and Evolutionary History of the Necrosis- and Ethylene-Inducing Peptide 1-Like Protein (NLP) Superfamily Across the *Dothideomycetes* Class of Fungi

Thaís C. S. Dal'Sasso¹ · Hugo V. S. Rody² · Luiz O. Oliveira¹

Received: 7 February 2022 / Accepted: 15 November 2022

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract

Necrosis and Ethylene-inducing peptide 1-like proteins (NLPs) are broadly distributed across bacteria, fungi, and oomycetes. Cytotoxic NLPs are usually secreted into the host apoplast where they can induce cell death and trigger plant immune responses in eudicots. To investigate the evolutionary history of the NLPs, we accessed the genomic resources of 79 species from 15 orders of *Dothideomycetes*. Phylogenetic approaches searched for biased patterns of NLP gene evolution and aimed to provide a phylogenetic framework for the cytotoxic activities of NLPs. Among *Dothideomycetes*, the NLP superfamily sizes varied, but usually contained from one to six members. Superfamily sizes were higher among pathogenic fungi, with family members that were mostly putative-effector NLPs. Across species, members of the NLP1 family (Type I NLPs) were predominant (84%) over members of the NLP2 family (Type II NLPs). The NLP1 family split into two subfamilies (NLP1.1 and NLP1.2). The NLP1.1 subfamily was broadly distributed across *Dothideomycetes*. There was strong agreement between the phylogenomics of *Dothideomycetes* and the phylogenetic tree based on members of the NLP1 subfamilies. To a lesser extent, phylogenomics also agreed with the phylogeny based on members of the NLP2 family. While gene losses seem to have shaped the evolutionary history of NLP2 family, ancient gene duplications followed by descent with modification characterized the NLP1 family. The strongest cytotoxic activities were recorded on NLPs of the NLP1.1 subfamily, suggesting that biased NLP gene retention in this subfamily favored the cytotoxic paralogs.

Introduction

Necrosis- and Ethylene-inducing peptide 1-like proteins (NLPs) are a superclass of proteins present in bacteria, fungi, and oomycetes [1]. In general, NLPs are small proteins of about 24 kDa [2] that exhibit cytotoxic activity with cell death-inducing properties and triggers of plant immune response in eudicots [1–3]. The first purified NLP was the Necrosis- and Ethylene-inducing peptide 1 (Nep1) from *Fusarium oxysporum* f. sp. *erythroxylae*, a protein capable of inducing both necrosis and production of ethylene when applied to *Erythroxylum coca* (coca) [4].

A number of evidence-based studies indicated that NLPs of plant-associated fungi may play a role as virulence factors during the early stages of infection and disease development. The removal of a cytotoxic NLP-encoding gene decreased the virulence of fungal mutants compared to the wild-type strain [5]. Conversely, the overexpression of a cytotoxic NLP-encoding gene increased the virulence of the mutant [6]. The role of NLPs in the infection process may not be equally important for all phytopathogens. For example, the removal of NLP-encoding genes did not change the virulence of *Magnaporthe oryzae* [7] and *Botrytis elliptica* [8] mutants. Through functional analyses, some members of the prolific NLP superfamilies of both *Diplodia seriata* [9] and *Neofusicoccum parvum* [10] were shown to exhibit varying levels of cytotoxicity. Contrary to their cytolytic counterpart, the noncytolytic NLPs cannot permeabilize the plant membrane but retain the capability of triggering plant immune responses; the biological role of the noncytolytic NLPs is yet to be characterized [11–14].

The defining molecular characteristic of an NLP is its NPP1 domain (Pfam PF05630) [15]. This domain contains

✉ Luiz O. Oliveira
lorlando@ufv.br

¹ Departamento de Bioquímica e Biologia Molecular, Universidade Federal de Viçosa, Viçosa, Brazil

² Departamento de Genética, Universidade de São Paulo/Escola Superior de Agricultura “Luiz de Queiroz”, Piracicaba, Brazil

a highly conserved, seven amino acid long motif (GHRH-DWE) that are involved in cation binding [11, 13, 16]. In the N-terminal half of NLPs, there are conserved cysteine residues that are able to form disulfide bridges, which seem to be essential for protein stabilization and necrosis induction [15, 16]. Consistent with their role as proteins of the secretory pathway, the vast majority of NLPs have an N-terminal signal peptide [3, 17]. These remarkable characteristics, along with their role as toxin-like virulence factor and broad distribution across taxa, usually distinguish NLPs as conserved effectors [1–3, 5, 16].

Based on the number of the conserved cysteines, NLPs have been classified into four main types: (a) Type I NLPs have two conserved cysteine residues and are the most widespread type, occurring in fungi, bacteria, and oomycetes [1, 3, 17]; (b) A variant called Type Ia also has two cysteines, but differs from Type I by the occurrence of distinct amino acid substitutions; Type Ia is found among oomycetes [17]; (c) Type II NLPs have four conserved cysteines, which are responsible for the formation of two disulfide bridges; occurrence of Type II NLPs is scarce, mostly in bacteria and fungi [1]; and (d) Type III NLPs exhibit the least conserved amino acid sequence, most of them carry six cysteine residues that are able to form three disulfide bridges; they occur in ascomycetes and some bacteria [3]. The occurrence of Types I, II, and III NLPs in Ascomycota suggests that the NLPs originated within this phylum; horizontal gene transfer (HGT) is a plausible mechanism that may have allowed NLPs to reach distantly related taxa [17].

A recent study surveyed the taxonomic distribution of NLPs in reference proteomes of about 10 thousand species of bacteria, fungi, and oomycetes; about 500 species contained at least one NLP [3]. The most striking features that bring together most of the species that do contain NLPs are their taxonomic placement within the phylum Ascomycota (60%, 211 of 360 proteomes), their plant pathogenicity (some of the most obnoxious phytopathogens do contain NLPs), and their trophic modes (especially pathogenic and saprophytic microbes). In *Dothideomycetes*, the overall number of NLPs copies is much smaller than the extremely large copy size number found in Oomycetes, in which up to 100 copies have been documented in some genomes [3]. Undoubtedly, the diversification of the NLPs across the *Dothideomycetes* is a complex process and far from being understood.

The *Dothideomycetes* comprises the largest and phylogenetically most-diverse class within the phylum Ascomycota, with an estimated number of members that may reach up to 19 thousand species [18, 19]. They occur across diverse habitats, including extreme environmental conditions, and their lifestyles are very diverse [18, 20, 21]. *Dothideomycetes* diverged from other sister Ascomycota classes around 366 million years ago (mya) [22]. Currently, the class comprises 23 orders [20] that are believed to have evolved in the

range between 100 and 220 mya [22]. The Pleosporales, the Capnodiales, and the Botryosphaerales are three of the largest orders within the *Dothideomycetes*; each of those orders contains a large number of families, which in turn hold some of the most destructive genera of plant pathogens to cereal crops, trees, and dicots [20, 21, 23].

With incredible ecological and morphological diversities and economical importance as plant pathogens, the *Dothideomycetes* have raised the interest of genomic research in the recent decade. Currently, genomes of about 90 genera of *Dothideomycetes* are available in public repositories, most of which are from plant pathogens and plant-associated species [20, 23]. Those genomic resources allow for the investigation of the evolutionary history of NLPs across that class.

We began our study by building up a robust molecular phylogenetic framework based on a set of 1,851 single-copy ortholog (SCO) proteins from 79 species of *Dothideomycetes*. To account for ecological diversity within the class, we included species of three trophic modes (pathotrophs, saprotrophs, and symbiotrophs) sorted out among plant pathogens, plant-associated species, and non-plant-associated species. Built independently from the NLP evolutionary history, our framework revealed the phylogenomic relationships among the studied species and should shed light into the evolutionary history of NLPs across *Dothideomycetes*.

Next, we reconstructed the phylogenetic relationships among NLPs (Types I, II, or III) based on the sequences we had retrieved from those genomic resources and questioned how ubiquitous and diverse the NLP superfamily became across the *Dothideomycetes*. If NLP gene evolution took place in an NLP-type-independent manner, the three types of NLPs should show similar distribution across the phylogenies. Otherwise, if biased patterns drove NLP gene evolution, the NLP-based phylogenies will be imbalanced owing to unequal gene losses, gene duplications, and gene retention that took place over time. As a consequence, NLP types that experienced adverse selection will be rare, likely vanishing from certain sub-clades.

Subsequently, we explored how the processes that drove NLP gene evolution took place at the lower ranks of the NLP phylogenetic hierarchy. To answer our questions, we build independent phylogenies, one for each of the major NLP sub-clades we had recovered previously. Then, we contrasted the topology of those trees with the phylogenetic framework of *Dothideomycetes*.

Finally, we provide a context to explore whether the cytotoxic activities reported previously through functional analyses are associated with NLP gene evolution and phylogenetic relationships among species. If preferential gene retention patterns maintained functional paralogs over time, cytotoxic activities will tend to exhibit some phylogenetic signals. This investigation shed light on the likely mechanisms that contributed to the evolution of NLPs in *Dothideomycetes*.

Material and Methods

Data Assembly and Trophic Mode Prediction

From MycoCosm [24], we retrieved protein sequences of 79 species (representing 15 orders) of *Dothideomycetes*. There was one species per genus, and two additional species of *Eurotiomycetes* (*Aspergillus fumigatus* and *Aspergillus nidulans*). The 81 species (79 *Dothideomycetes* and two *Aspergillus* spp.) used in this study are listed (Supplementary Table S1).

To infer the trophic modes, we run the FUNGuild annotation tool locally [25] using the database for fungi and a list of species (Supplementary Table S1) as input. FUNGuild assigns trophic modes based on assessments given in primary research literature and it uses species and genus as taxon levels for inferences.

Orthogroups and Phylogenomic Analysis

OrthoFinder v2.3.3 [26] established orthologous relationships among members of the *Dothideomycetes* (Supplementary Table S1). OrthoFinder calculated length and phylogenetic distance-normalized scores from an all-versus-all alignment using DIAMOND v0.9.24 [27] and identified reciprocal best normalized hits. Normalized scores above default thresholds were assigned to the orthogroup graph and subjected to the Markov clustering analysis in order to assume ortholog groups (orthogroups).

We used a Maximum-likelihood phylogeny approach to uncover the phylogenetic relationships among species of the *Dothideomycetes*. We firstly prepared a dataset with the SCO proteins as identified by OrthoFinder. Proteins were aligned using the L-INS-I method as implemented in MAFFT v7.453 [28]. TrimAL v1.4.rev22 [29] trimmed the alignments, with the parameters “-gt 0.95” and “-cons 60.” After trimming and concatenation, we obtained dataset 1.

We subjected dataset 1 as input to IQ-TREE v1.6.11 [30] to estimate both the best evolutionary model and the phylogenomic tree. According to the Bayesian information criteria (BIC), IQ-TREE standard model selection indicated LG + F + I + G4 as the best fit model. Subsequently, we carried out a Maximum-likelihood phylogenomic tree using 1,000 ultrafast bootstrap replicates and ten independent runs. *Aspergillus fumigatus* and *Aspergillus nidulans* were set as outgroups. The tree was visualized in FigTree v1.4.4 (<http://tree.bio.ed.ac.uk/software/figtree/>).

Protein Annotation and Assemble of NLP Homologues

The obtained proteomes of the set of 81 species were annotated using PfamScan with Pfam v32.0 [31] and InterProScan v5.30.69 [32] with the following eight parameters: SMART-7.1, SUPERFAMILY-1.75, ProDom-2006.1, CDD-3.16, TIGRFAM-15.0, Pfam v31.0, Coils-2.2.1, and Gene3D-4.2.0.

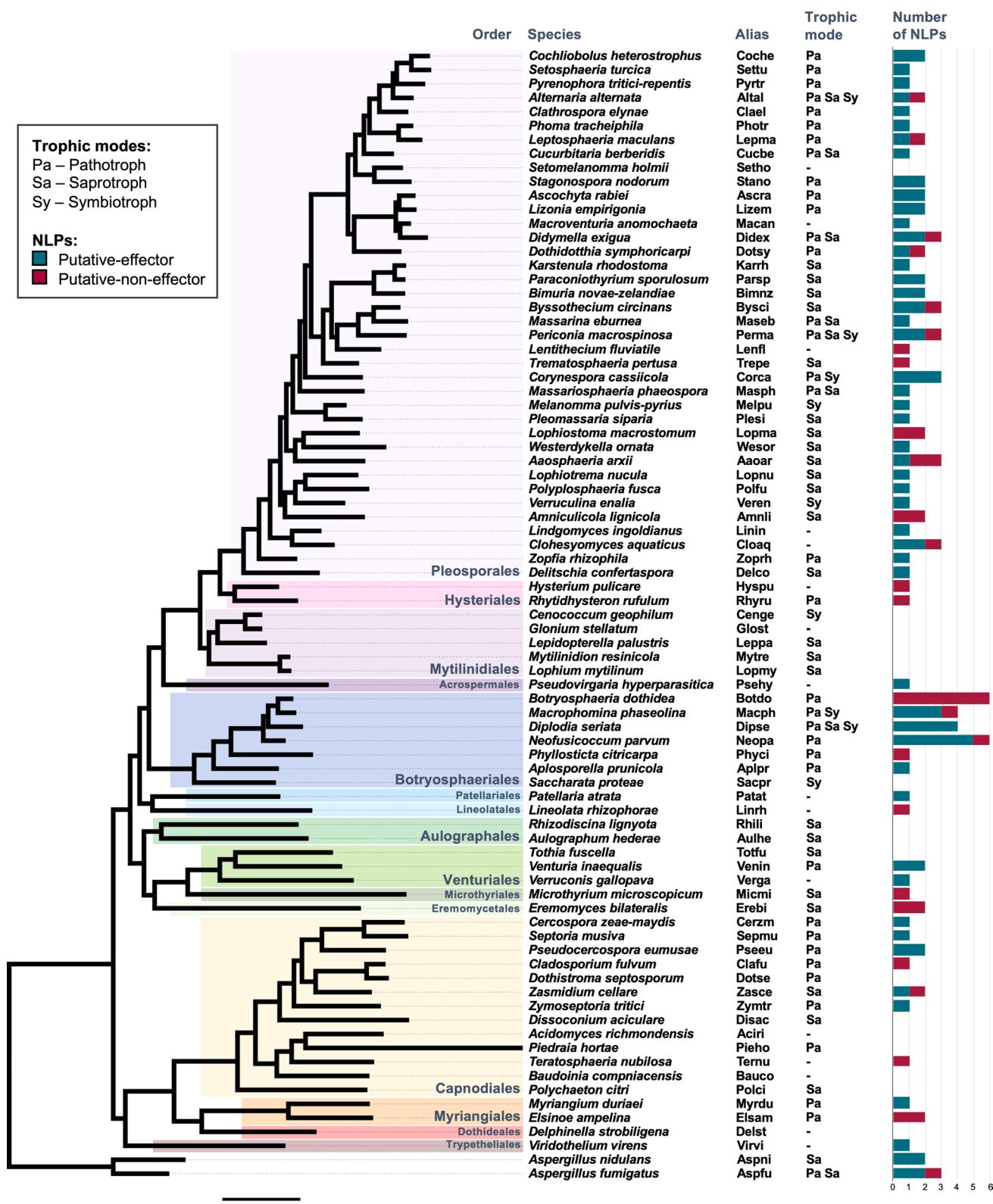
The presence of a signal peptide (according to SignalP v4.1; [33]) and the absence of transmembrane domains (according to TMHMM v2.0; [34]) defined the proteins that we predicted to be part of the secretome. Subsequently, TargetP [35] predicted the subcellular localization of the predicted secretome.

Finally, HMMER v3.2.1 (<http://hmmer.org>) predicted the NLP homologues (E-value < 0.001) using profile hidden Markov models (HMMs) for NPP1 domain (PF05630) from the proteomes. A protein was declared to be an NLP when the NPP1 domain was annotated by at least two out of three softwares: PfamScan, InterproScan, and HMMER. Within the pool of NLPs, a given protein was considered to be an ‘putative-effector NLP’ when it passed the following three tests: (a) SignalP predicted it harbors a signal peptide, (b) TMHMM predicted it has no transmembrane domain, and (c) TargetP predicted it to be part of a secretory pathway. When a given NLP failed any of these three tests, it was regarded as a ‘putative-non-effector NLP.’ Finally, we assembled dataset 2, which contained the predicted NLPs (protein data) that were present in the set of 81 species.

NLP Maximum-Likelihood Phylogenies

Maximum-likelihood phylogenetic analysis based on dataset 2 allowed us to reconstruct the phylogenetic relationships of NLPs among the members of the *Dothideomycetes*.

Alignments were obtained using the L-INS-I method, as implemented in MAFFT v7.453. MAFFT L-INS-I allows for aligning a set of sequences containing sequences flanking around one alignable domain with high accuracy [28]. To find the best fit model, we used dataset 2 and its partitions as input to IQ-TREE. According to BIC, IQ-TREE selected WAG + F + I + G4 for dataset 2; for its partitions, the best evolutionary models were WAG + F + I + G4, WAG + I + G4, or VT + G4. Maximum-likelihood phylogenetic trees for NLPs were also performed in IQ-TREE as described previously. Phylogenetic analysis of the full dataset 2 was performed without outgroup. However, phylogenetic analyses on partitions of dataset 2 (see results section) were carried out using *A. fumigatus* as outgroup.



◀Fig. 1 Phylogenomic relationships among 79 species of *Dothideomycetes*. The Maximum-likelihood consensus tree is based on a dataset of 1851 single-copy ortholog proteins, with *Aspergillus* spp. (*Eurotiomycetes*) as outgroups. Clades are color coded according to the orders of *Dothideomycetes*, as indicated. Five-letter codes indicate alias of each species name. Trophic modes (according to FUNGuild classification) are coded as indicated. Bars indicate the number of NLPs (putative effectors and putative non-effectors) per species. Branch lengths are drawn to scale; nodal support values are equal to 100 local bootstraps for all nodes. Scale bar corresponds to the expected number of substitutions per site

Survey on Functional Activities of NLPs

We surveyed the literature for functional analyses of NLPs across Ascomycota. We recorded the taxonomic placement of the study species, the NLP names attributed to the proteins, and the original outcome of the functional analyses (usually each study ranked the activity as either strong, weak, or absent). We included in our phylogenetic approach the *Dothideomycetes* species for which a functional analysis had been carried out. Species of other classes of fungi were not included in our phylogenies, but we carried out protein sequence alignments individually in order to assign the placement of each of their NLPs into the phylogenetic treatments we developed herein.

Statistical Analysis

To determine the statistical differences in the number of NLPs and putative-effector NLPs among the predicted trophic modes, we applied a one-way ANOVA followed by the Sidak test (P value < 0.05) for the mean number of NLPs and the mean number of putative effectors NLPs, independently. Statistical analyses were implemented in R v4.2.1 (<http://www.R-project.org/>) using the *emmeans* and *multicomp* packages.

Results

Genome-Wide Phylogeny of *Dothideomycetes*

The set of 79 species (79 genera, 15 orders) of *Dothideomycetes* allowed OrthoFinder to uncover 1,851 SCO proteins (a concatenated set of 910,900 amino acids). The whole-genome data provided support for the phylogenetic reconstruction within *Dothideomycetes* (Fig. 1). The phylogenomic tree showed well-supported nodes (bootstrap values = 100, for all nodes). Overall, there were two major clades. The first major clade held 17 species of four orders (Capnodiales, Myriangiales, Dothideales, and Trypetheliales). The second major clade held 62 species within 11 orders of *Dothideomycetes*; it was split further into two sister sub-clades. The first of these sub-clades encompassed seven species of four orders: Aulographiales (two species),

Venturiales (three), Mycrothyriales (one), and Eremomycetales (one). The second sub-clade comprised 55 species of seven orders: Pleosporales (38 species), Hysteriales (two), Mytilinidiales (five), Acrospermales (one), Botryosphaerales (seven), Patellariales (one), and Lineolatales (one).

Genome-Wide Identification of Homologues of NLPs in *Dothideomycetes*

Apart from the NPP1 domain (PF05630), our pipeline predicted no other domain as part of the NLPs. Each NLP had a single copy of the NPP1 domain. Among the NLPs, our pipeline distinguished between two sets of proteins: (a) putative-effector NLPs and (b) putative-non-effector NLPs. In the first set, the protein harbored a signal peptide, had no transmembrane domain, and was part of the secretory pathway. In the second set, despite the presence of the NPP1 domain, the protein failed to comply with any of the previous three requirements.

In general, NLP superfamily sizes (putative-effector NLPs + putative-non-effector NLPs) among species of *Dothideomycetes* varied (Fig. 1, Supplementary Table S1). There were 17 species without NLPs: six species of Capnodiales (*Acidomyces richmondensis*, *Baudoinia compniacensis*, *Dissoconium aciculare*, *Dothistroma septosporum*, *Piedraia hortae*, and *Polychaeton citri*), all five species of Mytilinidiales (*Cenococcum geophilum*, *Glonium stellatum*, *Lepidothecula palustris*, *Lophium mytilinum*, and *Mytilinidion resinicola*), the two species of Aulographales (*Aulographum hederae* and *Rhizodiscina lignyota*), one species of Pleosporales (*Setomelanomma holmii*), one of Botryosphaerales (*Saccharata proteae*), one of Dothideales (*Delphinella strobiligena*), and one of Venturiales (*Tothia fuscella*).

Among the species that do harbor NLPs, the copy number varied from one (62 species) to six (*Botryosphaeria dothidea* and *Neofusicoccum parvum*; Botryosphaerales) (Supplementary Table S1). In a given species, the general trend was that the number of putative-effector NLPs was larger than the number of putative-non-effector NLPs. The copy number of putative-effector NLPs ranged from one (32 species) to five (*Neofusicoccum parvum*). In 14 species, however, the NLP family contained exclusively putative-non-effector NLPs (from one to six members).

Broad Sense Phylogeny of NLPs

After carrying out a phylogenomic reconstruction within the set of 79 species of *Dothideomycetes* (Fig. 1), we re-examined the phylogenetic relationships among species with protein data from the NLPs as the only source of information. Therefore, 17 species that lacked NLPs did not take part in this new analysis. The Maximum-likelihood phylogenetic tree contained a total of 111 NLPs, 106 of them came from a total of 62 species

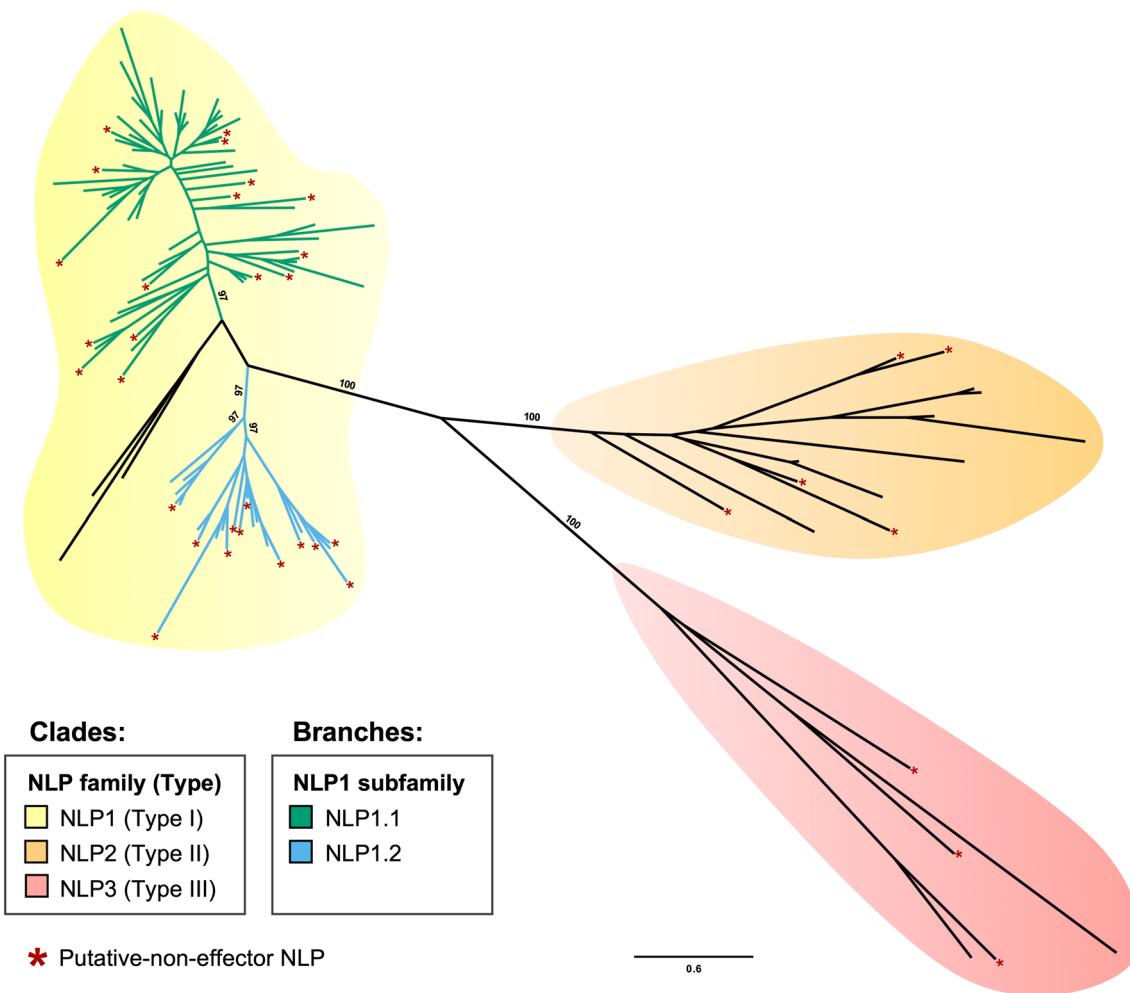


Fig. 2 Phylogenetic relationships among 111 predicted Necrosis- and Ethylene-inducing like proteins (NLPs). The dataset for the unrooted Maximum-likelihood phylogeny (consensus tree) comprised proteins that were 572 amino acids long. A total of 106 out of 111 NLPs were from the *Dothideomycetes*. Clades are color coded according to the number of conserved cysteine residues in the protein (Type I, two residues; Type II, four residues; Type III, variable number of residues),

of *Dothideomycetes* and five NLPs came from *Aspergillus* spp. (Fig. 2, Supplementary Table S2). The phylogenetic tree showed well-supported nodes (bootstrap values > 95 , for the main nodes). There were three major clades. A shared trait of all NLPs within a given major clade was the number of conserved cysteine residues in the protein (which separates NLPs into Types: I, II, or III). Notably, the placement of putative-non-effector NLPs was spread across the phylogeny.

The first major clade harbored 93 NLPs, all of which were considered as Type I NLPs (with two cysteine residues; see [3]). Among the 93 NLPs, 75 (81%) were predicted to be putative-effector NLPs. The first major clade was split further into three sub-clades. One of the three sub-clades contained three highly differentiated putative-effector NLPs: one from *Aaospaeria arxii* (Aaoar_355692), one

as indicated. Branches are color coded according to the NLP1 subfamily, as indicated. Along the phylogeny, asterisks indicate the phylogenetic placements of putative-non-effector NLPs. Branch lengths are drawn to scale; nodal support values are given as local bootstraps (when ≥ 95 , for the main nodes) above the branches. Scale bar corresponds to the expected number of substitutions per site (Color figure online)

from *Paraconiothyrium sporulosum* (Parsp_876375), and one from *Periconia macrospinosa* (Perma_691703). Species composition among the remaining two sub-clades was large and very diverse. The second major clade brought together 13 NLPs (Type II NLPs; with four cysteine residues). Among the 13 NLPs, eight (62%) were predicted to be putative-effector NLPs. Meanwhile, the third major clade encompassed only five NLPs (Type III NLPs; with variable number of cysteine residues). Among them, only two (40%) were predicted to be putative-effector NLPs.

Narrow Sense Phylogeny of NLPs

We carried out four subsequent studies based on the phylogenetic relationships we recovered previously (Fig. 2,

Supplementary Table S2). The data from the first major clade gave rise to two independent phylogenies (one for each of the two best-resolved sub-clades (Fig. 3, Supplementary Table S2). Hereafter, these two narrow sense phylogenies are referred to as 'NLP1.1' and 'NLP1.2' phylogenies, respectively. We also built independent phylogenies for the remaining two major clades (Fig. 4, Supplementary Table S2). Hereafter, these two narrow sense phylogenies are referred to as 'NLP2' and 'NLP3' phylogenies, respectively.

Unequivocally, NLP1.1 was the most member-rich phylogeny; it contained 66 Type I NLPs (50 putative effectors, 76%) from 60 species of *Dothideomycetes* and *Aspergillus* spp. (Fig. 3A). Among the 37 species that contained a single NLP, 34 (92%) of them were present in NLP1.1. Among the 64 NLP1.1 proteins from *Dothideomycetes*, there were a total of 48 putative-effector NLPs. The richness toward putative-effector NLPs was higher among Pleosporales, in which 32 out of 38 NLPs were predicted as putative-effector proteins.

The NLP1.1 phylogeny split further into two major clades (Fig. 3A). The major clade 1 contains six NLPs from Botryosphaerales (Botdo_289801, Macph_6483, Neopa_7612, Dipse_9327, Aplpr_237706, Phyci_582607). The major clade 2 was member rich and split further into two sister sub-clades; the smallest of which contained 12 NLPs from seven orders (Botryosphaerales, Pleosporales, Lineolatales, Venturiales, Capnodiales, Myriangiales, and Trypetheliales). Finally, the large sub-clade 2 contained NLPs from Capnodiales, Eremomycetales, Acrospermales, Hysteriales, and a large second-order sub-clade of NLPs from Pleosporales. Overall, NLP1.1 paralleled quite well the molecular phylogenetic framework of *Dothideomycetes*, especially across Pleosporales. Botryosphaerales retained a major gene duplication of NLP1.1, with sets of corresponding paralogs in distinct genera being retained over time. This major gene duplication is absent from either Capnodiales or Pleosporales.

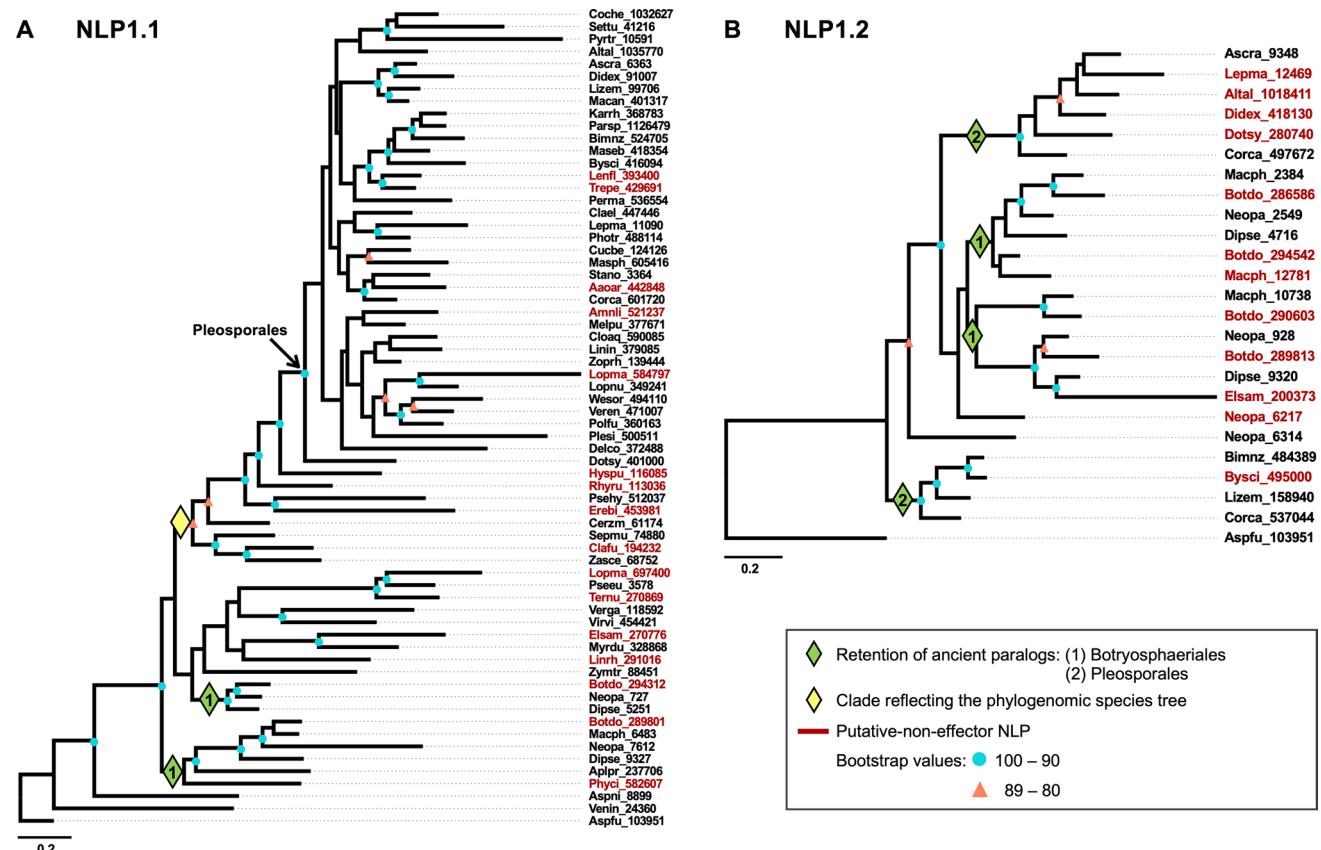


Fig. 3 Phylogenetic relationships among members in two subfamilies of the Necrosis- and Ethylene-inducing like protein family 1 (NLP1) in *Dothideomycetes*. Each phylogeny represents a sub-clade of NLP1 family depicted in Fig. 2: **A** Members of NLP1.1 (66 proteins that were 307 amino acids long) and **B** Members of NLP1.2 (24 proteins that were 270 amino acids long). Protein IDs in red indicate putative-non-effector NLPs. Yellow diamond indicates the major clade in NLP1.1 that reflects the phylogenomic species tree (see Fig. 1); the

clade with NLPs from Pleosporales as indicated. Green diamonds indicate retentions of NLP paralogs that took place over time in Botryosphaerales and Pleosporales. For each phylogeny, Aspfu_103951 (NLP1.1) from *Aspergillus fumigatus* (Eurotiomycetes) was used as outgroup. Branch lengths are drawn to scale; nodal support values are given as local bootstraps, as indicated. Scale bar corresponds to the expected number of substitutions per site (Color figure online)

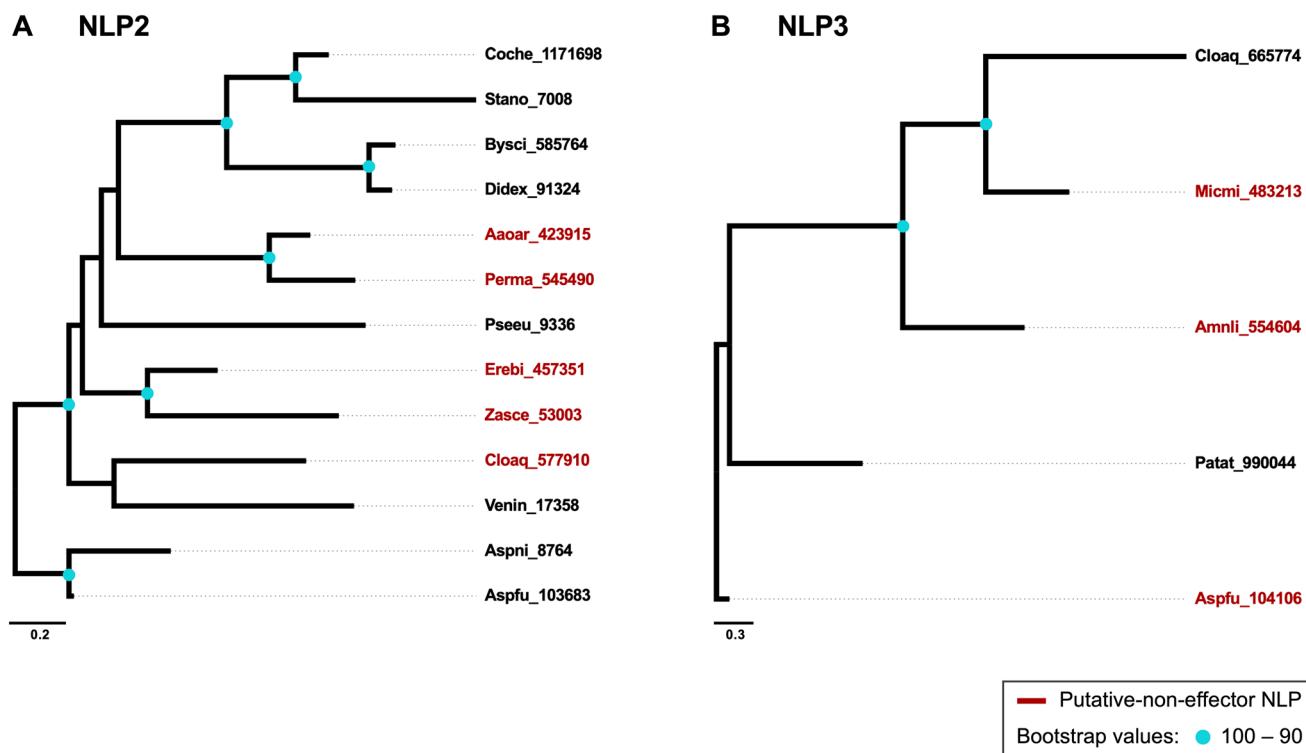


Fig. 4 Phylogenetic relationships among members of two Necrosis- and Ethylene-inducing like protein families—NLP2 and NLP3—in *Dothideomycetes*. Each phylogeny represents a sub-clade depicted in Fig. 2: **A** Members of NLP2 (13 proteins that were 491 amino acids long) and **B** Members of NLP3 (five proteins that were 390 amino acids long). Protein IDs in red indicate putative-non-effector NLPs. For each phylogeny, an NLP from *Aspergillus fumigatus* (*Eurotiomycetes*) was used as outgroup. Branch lengths are drawn to scale; nodal support values are given as local bootstraps, as indicated. Scale bar corresponds to the expected number of substitutions per site

The remaining 24 Type I NLPs from fourteen species of *Dothideomycetes* were analyzed in the NLP1.2 phylogeny; half of the members (12 out of 24) were putative-non-effector NLPs (Fig. 3B). This phylogeny revealed two additional duplication events; one of them was shared among Botryosphaerales only. The second duplication gave rise to a sub-clade that has a sister relationship to Pleosporales. Although not as member rich as NLP1.1, NLP1.2 also resembled well the molecular phylogenetic framework of *Dothideomycetes*.

The Type II NLPs clustered together in the NLP2 phylogeny; they were of rare occurrence (nine out of 79 species) among the *Dothideomycetes* and *Aspergillus* spp.; five species were Pleosporales (Fig. 4A). Finally, there was the NLP3 phylogeny, which consisted of only five Type III NLPs, from four species among the *Dothideomycetes* and *Aspergillus fumigatus* (Fig. 4B).

Cytotoxic NLPs Within a Phylogenetic Context

Our survey recovered 27 functional analyses of NLPs (Table 1). There were 11 instances in which the authors attributed to the NLP under investigation a strong activity, all of the proteins were Type I NLPs. Among those 11 proteins

acids long). Protein IDs in red indicate putative-non-effector NLPs. For each phylogeny, an NLP from *Aspergillus fumigatus* (*Eurotiomycetes*) was used as outgroup. Branch lengths are drawn to scale; nodal support values are given as local bootstraps, as indicated. Scale bar corresponds to the expected number of substitutions per site

with strong cytotoxic activity, there were seven NLP1.1 and four NLP1.2. There were nine instances in which a Type 1 NLP showed weak/absent activity; those instances took place in species with multiple copies of the Type I NLPs and a strong activity already had been attributed to one of the NLP1 paralogs. Finally, all six functional analyses of Type II NLPs (NLP2) and a single analysis of Type III NLP (NLP3) pointed to weak/absent activity. In *Neofusicoccum parvum* (Botryosphaerales), the NLP1.2 paralogs Neopa_928 and Neopa_6217 had been both removed from functional analysis experiment [10]. While the overexpression of Neopa_928 as a recombinant protein was not possible, Neopa_6217 coded for a truncated protein that lacked signal peptide. Meanwhile, our analyses predicted that Neopa_928 was a putative-effector NLP and Neopa_6217 was a putative-non-effector NLP.

NLPs Within a Trophic Mode Context

Using the FUNGuild annotation tool, we inferred the trophic mode for each of the *Dothideomycetes* species into three groups: pathotrophs, saprotrophs, and symbiotrophs (Fig. 5, Supplementary Table S1). Sixty-three (80%) out of 79

Table 1 The cytotoxicity of NLPs found in species of the phylum Ascomycota

Reference	Class/order/species	NLP name	Cytotoxicity to dicots	Protein ID (this work)	NLP context (this work)
[9]	<i>Dothideomycetes/Botryosphaerales/Diplodia seriata</i>	DserNEP1	Strong	Dipse_5251	NLP1.1
[9]	<i>Dothideomycetes/Botryosphaerales/Diplodia seriata</i>	DserNEP2	Weak	Dipse_9327	NLP1.1
[10]	<i>Dothideomycetes/Botryosphaerales/Neofusicoccum parvum</i>	NprvNep1	Strong	Neopa_727	NLP1.1
[10]	<i>Dothideomycetes/Botryosphaerales/Neofusicoccum parvum</i>	NprvNep2	Strong	Neopa_7612	NLP1.1
[10]	<i>Dothideomycetes/Botryosphaerales/Neofusicoccum parvum</i>	NprvNep3	Weak	Neopa_6314	NLP1.2
[10]	<i>Dothideomycetes/Botryosphaerales/Neofusicoccum parvum</i>	NprvNep4	Weak	Neopa_2549	NLP1.2
[42]	<i>Dothideomycetes/Capnodiales/Zymoseptoria tritici</i>	MgNLP	Strong	Zymtr_88451	NLP1.1
[5]	<i>Eurotiomycetes/Eurotiales/Penicillium expansum</i>	PeNLP1	Strong	-	NLP1.1 ^a
[5]	<i>Eurotiomycetes/Eurotiales/Penicillium expansum</i>	PeNLP2	Weak	-	NLP3 ^a
[37]	<i>Sordariomycetes/Hypocreales/Verticillium dahliae</i>	NLP1	Strong	-	NLP1.2 ^a
[37]	<i>Sordariomycetes/Hypocreales/Verticillium dahliae</i>	NLP2	Weak	-	NLP1.1 ^a
[37]	<i>Sordariomycetes/Hypocreales/Verticillium dahliae</i>	NLP3	Absent	-	NLP1.2 ^a
[37]	<i>Sordariomycetes/Hypocreales/Verticillium dahliae</i>	NLP4	Absent	-	NLP2 ^a
[37]	<i>Sordariomycetes/Hypocreales/Verticillium dahliae</i>	NLP5	Absent	-	NLP2 ^a
[37]	<i>Sordariomycetes/Hypocreales/Verticillium dahliae</i>	NLP6	Absent	-	NLP1.2 ^a
[37]	<i>Sordariomycetes/Hypocreales/Verticillium dahliae</i>	NLP7	Absent	-	NLP2 ^a
[37]	<i>Sordariomycetes/Hypocreales/Verticillium dahliae</i>	NLP9	Absent	-	NLP2 ^a
[7]	<i>Sordariomycetes/Magnaportheales/Magnaporthe oryzae</i>	MoNLP1	Strong	-	NLP1.1 ^a
[7]	<i>Sordariomycetes/Magnaportheales/Magnaporthe oryzae</i>	MoNLP2	Weak	-	NLP2 ^a
[7]	<i>Sordariomycetes/Magnaportheales/Magnaporthe oryzae</i>	MoNLP3	Absent	-	NLP2 ^a
[7]	<i>Sordariomycetes/Magnaportheales/Magnaporthe oryzae</i>	MoNLP4	Strong	-	NLP1.2 ^a
[41]	<i>Leotiomycetes/Helotiales/Sclerotinia sclerotiorum</i>	SsNLP1	Weak	-	NLP1.2 ^a
[41]	<i>Leotiomycetes/Helotiales/Sclerotinia sclerotiorum</i>	SsNLP2	Strong	-	NLP1.1 ^a
[40]	<i>Leotiomycetes/Helotiales/Botrytis cinerea</i>	BcNEP1	Strong	-	NLP1.2 ^a
[40]	<i>Leotiomycetes/Helotiales/Botrytis cinerea</i>	BcNEP2	Weak	-	NLP1.1 ^a
[8]	<i>Leotiomycetes/Helotiales/Botrytis elliptica</i>	BeNEP1	Strong	-	NLP1.2 ^a
[8]	<i>Leotiomycetes/Helotiales/Botrytis elliptica</i>	BeNEP2	Weak	-	NLP1.1 ^a

^aProteins that were individually aligned against our sequence database, but data not included in Fig. 2

species of *Dothideomycetes* were classified into at least one of the three groups (Fig. 5A). There were 26 pathotrophs, 25 saprotrophs, and four symbiotrophs. Eight species were classified as part of more than one group: pathotrophs–symbiotrophs (two species), pathotrophs–saprotrophs–symbiotrophs (three), and pathotrophs–saprotrophs (three). In general, the average number of NLPs per species was higher among the species classified as pathotrophs (exclusively or not) and saprotrophs (Fig. 5B). The average number of putative-effector NLPs per species was higher among the species classified as pathotrophs (exclusively or not) (Fig. 5C).

Discussion

The current type-based classification of NLPs (Types I, II, and III) relies exclusively on the number of conserved cysteine residues that form disulfide bridges along the

N-terminal half of the proteins [1, 17]. Previously, a broad phylogenetic analysis of NLPs from bacteria, fungi, and oomycetes [3] provided hints that suggested NLPs of a given type might share a common ancestor. Herein, we used phylogenomics and NLP phylogenetic data to explore the evolutionary history of NLPs in the *Dothideomycetes* class of fungi.

Our studies support a phylogeny-based classification of NLPs into families. Accordingly, our phylogenetic framework uncovered at least two subordinate families that corresponded to existing NLP types. Hereafter, we will refer to Type I NLPs collectively as the ‘NLP1 family’ and Type II NLPs as the ‘NLP2 family.’ Furthermore, the phylogenetic relationships within the member-rich NLP1 family of *Dothideomycetes* supported two further natural subdivisions: hereafter referred to as NLP1.1 and NLP1.2 subfamilies.

Likely, Type III NLPs may also merit their own phylogenetic-based category into a NLP3 family; however, several

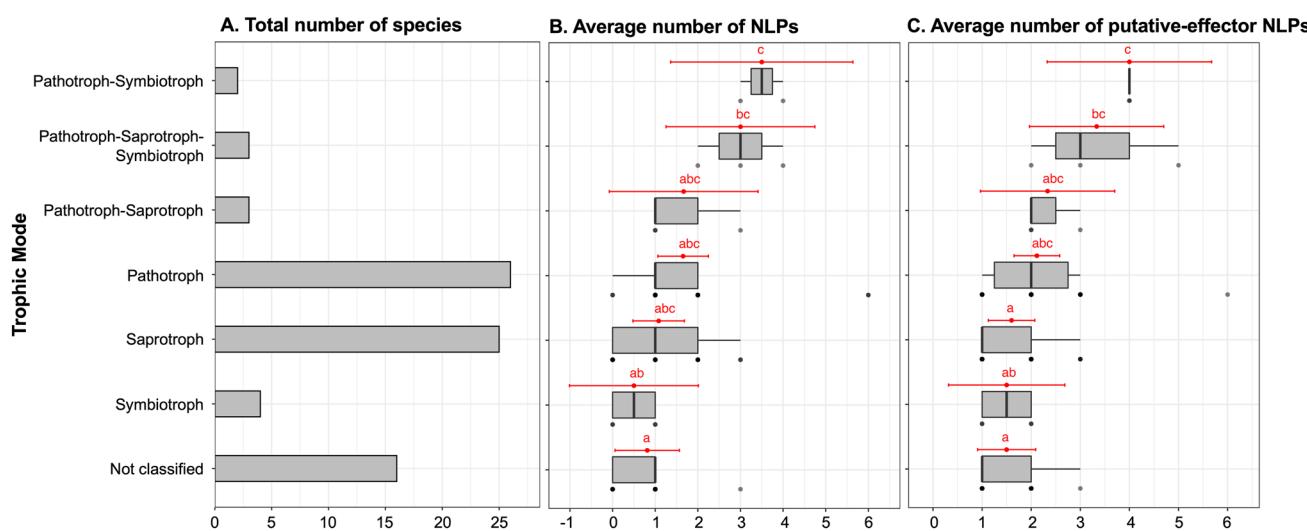


Fig. 5 Trophic modes predicted among 79 species of *Dothideomycetes* and its relationship with NLP copy number. **A** Total number of species per trophic mode. **B** Average numbers of NLPs per species per trophic mode. **C** Average numbers of putative-effector NLPs per species per trophic mode. Black dots represent raw data. Red dots and

error bars represent (estimated marginal) means \pm 95% confidence interval per group. Means not sharing any letter are significantly different by the Sidak test at the 5% level of significance. Classification of *Dothideomycetes* into trophic modes was performed using the FUNGuild annotation tool

aspects of their evolutionary history are still unclear, such as their low sequence conservation and likely association with horizontal gene transfer [17]. Moreover, the rare occurrence of Type III NLPs in *Dothideomycetes* may bias our studies. Thus, we refrained from making further attempts to explore the phylogenetics of NLP3 family with the current dataset available to this investigation.

The NLP superfamily exhibited a small size across *Dothideomycetes*, with one (62 species) to six members (two species in Botryosphaeraiales). Among *Dothideomycetes*, NLP family size is likely associated with trophic mode. Fungi inferred to be pathotrophs detained the highest number of putative-effector NLP copies; meanwhile, fungi inferred to be pathotrophs and saprotrophs detained more NLP copies than those inferred to be symbiotrophs. These findings agree with a recent study that describe some plant pathogens (necrotrophs and hemibiotrophs) detaining more NLP copies and a broad distribution of NLPs across saprotrophic species [3]. Cell death-inducing proteins, such as NLPs, which act at the plant apoplast level, are essential for host colonization [9, 10, 36] and may contribute to the decay of plant material [3]. Moreover, the role of the larger NLP family in necrotrophic plant pathogens, such as that of the Botryosphaeraiales, can go beyond differential cytotoxicity and detain levels of functional diversification at different life stages [37, 38] or, to some extent, contribute to infection of wider host ranges by those pathogens [37, 39].

The repositories of NLPs are the product of dynamic events that took place during the evolutionary history of extant *Dothideomycetes*. The NLP1 family and the NLP2

family clearly underwent distinct pathways during their evolution. Notably, gene retention and gene loss over time are major drivers that distinguish the NLP1 family apart from NLP2 family. The NLP1 family likely underwent rounds of gene duplication followed by descent with modification, which allowed for its diversification into at least two subfamilies (NLP1.1 and NLP1.2). However, these subfamilies experienced distinct evolutionary processes over time. Preferential gene retention contributed to preserve the NLP1.1 paralogs that emerged after gene duplication and explains the ubiquitous presence of the NLP1.1 subfamily members across most of the extant *Dothideomycetes* (Fig. 3A). On the contrary, gene loss is a realistic scenario that explains the evolutionary history of the NLP1.2 subfamily. Over time, preferential gene loss eliminated NLP1.2 members from the genomes of most extant *Dothideomycetes* (Fig. 3B). The fact that the phylogenies of each subfamily resembled the phylogenomics of their species supports the scenario of early gene duplications and subsequent biased gene losses in some lineages as a likely explanation to account for the extant pattern of distribution of the NLP1 family across *Dothideomycetes*.

The signatures of past lineage-specific gene duplication and gene retention are visible in the phylogenies of subfamilies NLP1.1 and NLP1.2, especially in the order Botryosphaeraiales, which showed above-average number of the NLP1 family (Fig. 3). We speculate that the otherwise redundancy of extant copies in Botryosphaeraiales may contribute to shape a strategy of functional diversification to overcome host-related defense mechanisms, with distinct NLPs playing

a functional role in a time-dependent manner to allow successful colonization of plant tissues.

In sharp contrast to the ubiquitous presence of NLP1 family members, the occurrence of the NLP2 family members among *Dothideomycetes* was rare. Similarly, to the NLP1.2 subfamily, biased gene losses also seem to have played a major role in driving the trajectory of the NLP2 family. Yet, few members of the NLP2 family survived and are distributed across phylogeny of the extant *Dothideomycetes* (Fig. 4A).

Putative-non-effector NLPs were those NLPs that do harbor NPP1 domain (PF05630) but either do not contain a signal peptide, may contain transmembrane domains, or are not part of the secretory pathway. Apparently, a putative-non-effector NLP lacks some of the key traits that define a protein as a functional NLP. Their presence across the phylogeny implies that putative-non-effector NLPs originated multiple times over time. Along their evolution, likely, the proteins we predicted to be putative-non-effector NLPs underwent independent rounds of sequence rearrangements, with losses and gain of traits, including their secretion signals. High frequency of putative-non-effectors among NLP1.2 paralogs also suggests that ongoing pseudogenization may be helping to avoid functional redundancy of NLPs in *Dothideomycetes*. The retention of those genes over time is puzzling; whether this is a strategy to escape from host recognition remains to be demonstrated experimentally.

Through functional analyses, a number of NLPs have been demonstrated to exhibit cytotoxic activity in eudicot plants [5, 7–10, 37, 40–42]. These analyses were carried out at the species level; the authors originally summarized the final outcomes of the cytotoxicity levels of a given NLP as either ‘strong,’ ‘weak,’ or ‘absent.’ In the present investigation, we demonstrated that the cytotoxicity of NLPs contains a strong phylogenetic signal (Table 1).

We applied our phylogenetic framework to those NLPs that had undergone functional analyses in previous studies. The vast majority of the NLPs that had exhibited ‘strong’ activity were phylogenetically related; they belong to the member-rich family NLP1 and grouped mostly within subfamily NLP1.1. Our findings suggested that members of the NLP1 family may have been favored through biased gene retention over time. Likely, the NLP1 paralogs that encode for proteins with ‘strong’ cytotoxicity are actually encoding putative-effector NLPs with a functional role in promoting virulence. Therefore, the maintenance of these putative-effector-encoding genes in the genome would be advantageous at evolutionary time scale and could explain their maintenance across extant *Dothideomycetes*.

Experiment-based evidence supports NLP1 members as genes that encode for essential virulence factors in plant pathogenic fungi. In the postharvest pathogen *Penicillium expansum* (*Eurotiomycetes*), the in vitro deletion of a

member of the NLP1 family led to the reduction of virulence; meanwhile the virulence remained unchanged when a Type III NLP was the deleted copy [5]. In *Neofusicoccum parvum* (*Dothideomycetes*), NLP1.1 paralogs encoded proteins that were strongly cytotoxic to both tomato plants and mammalian cells, while NLP1.2 paralogs encoded for proteins that exhibited weak cytotoxicity [10]. The only NLP from *Zymoseptoria tritici* (*Dothideomycetes*), which induced defense responses and cell death in dicots but not in monocots [42] and shared homology to members of the NLP1.1 subfamily. Paralogs of the NLP1.1 subfamily of *Diplodia seriata* (*Dothideomycetes*) encoded proteins that showed distinct levels of cytotoxicity to grapevine leaves [9].

Our phylogenetic framework grouped most of the paralogs that encode for NLP proteins with ‘weak’ cytotoxicity—as well as proteins with activity classified as ‘absent’—together with members of either the NLP2 or NLP3 families. The scarcity of either NLP2 or NLP3 members in the genome of most species of *Dothideomycetes* suggests that the proteins these genes encode for likely play a secondary role in virulence. In some instances, NLP cytotoxicity and immunity induction are uncoupled, and non-cytotoxic NLPs may retain the ability to trigger host immune responses [11, 13]. An additional evidence for a secondary role of NLP2 and NLP3 family members in virulence comes from the fact that fungal genomes that do possess NLP2 or NLP3 genes frequently harbored at least one NLP1 gene that encoded for a protein with ‘strong’ activity in the functional analyses. Therefore, these functional studies support an evolutionary scenario in which members of the NLP1 family likely play a critical role as effector genes, therefore, contributing actively as a tool to overcome host defense systems and facilitate the infection process.

Conclusion

The phylogenies of each NLP subfamilies followed closely the phylogenomic relationships of the genera in *Dothideomycetes*. Nevertheless, the imbalanced phylogeny of the NLP superfamily across the *Dothideomycetes* revealed that subfamilies underwent independent evolutionary paths, each of which showing different signatures of ancient gene duplications and biased successive gene losses that may be associated with changes in cytotoxic activity.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00284-022-03125-8>.

Author Contributions All authors contributed to the study conception. TCSD performed data assembly and analyses. TCSD and LOO wrote the manuscript. HVSR commented and edited the manuscript. LOO and HVSR supervised the research. All authors read and approved the final manuscript.

Funding This work was supported by The Minas Gerais State Foundation of Research Aid – FAPEMIG (Grant No. APQ-00150-17) and by The National Council of Scientific and Technological Development – CNPq (Fellowship Number PQ 302336/2019-2) to LOO. TCSD received student fellowships from the CAPES Foundation (PROEX – 0487 No. 1684083) and CNPq (GM/GD 142400/2018-1). HVSR was supported by a PD fellowship from the São Paulo Research Foundation – FAPESP (2018/04555-0).

Data Availability This study used genomic data available from public repositories and their information is in the Electronic Supplementary Material.

Code Availability Not applicable.

Declarations

Conflict of interest The authors have declared that no competing interests exist.

Ethics Approval Not applicable.

Consent to Participate Not applicable.

Consent for Publication Not applicable.

References

- Gijzen M, Nürnberger T (2006) Nep1-like proteins from plant pathogens: recruitment and diversification of the NPP1 domain across taxa. *Phytochemistry* 67:1800–1807. <https://doi.org/10.1016/j.phytochem.2005.12.008>
- Pemberton CL, Salmond GPC (2004) The Nep1-like proteins—a growing family of microbial elicitors of plant necrosis. *Mol Plant Pathol* 5:353–359. <https://doi.org/10.1111/j.1364-3703.2004.00235.x>
- Seidl MF, Van den Ackerveken G (2019) Activity and phylogenetics of the broadly occurring family of microbial Nep1-Like Proteins. *Annu Rev Phytopathol* 57:367–386. <https://doi.org/10.1146/annurev-phyto-082718-100054>
- Bailey BA (1995) Purification of a protein from culture filtrates of *Fusarium oxysporum* that induces ethylene and necrosis in leaves of *Erythroxylum coca*. *Phytopathology* 85:1035–1039
- Levin E, Raphael G, Ma J et al (2019) Identification and functional analysis of NLP-encoding genes from the postharvest pathogen *Penicillium expansum*. *Microorganisms*. <https://doi.org/10.3390/microorganisms7060175>
- Amsellem Z, Cohen BA, Gressel J (2002) Engineering hypervirulence in a mycoherbicidal fungus for efficient weed control. *Nat Biotechnol* 20:1035–1039. <https://doi.org/10.1038/nbt743>
- Fang YL, Peng YL, Fan J (2017) The Nep1-like protein family of *Magnaporthe oryzae* is dispensable for the infection of rice plants. *Sci Rep* 7:1–10. <https://doi.org/10.1038/s41598-017-04430-0>
- Staats M, van Baarlen P, Schouten A et al (2007) Positive selection in phytotoxic protein-encoding genes of *Botrytis* species. *Fungal Genet Biol* 44:52–63. <https://doi.org/10.1016/j.fgb.2006.07.003>
- Cobos R, Calvo-Peña C, Álvarez-Pérez JM et al (2019) Necrotic and cytolytic activity on grapevine leaves produced by Nep1-like proteins of *Diplodia seriata*. *Front Plant Sci* 10:1–13. <https://doi.org/10.3389/fpls.2019.01282>
- Pour FN, Cobos R, Coque JJR et al (2020) Toxicity of recombinant Necrosis and Ethylene-Inducing Proteins (NLPs) from *Neofusicoccum parvum*. *Toxins* 12:1–17. <https://doi.org/10.3390/toxins12040235>
- Böhm H, Albert I, Oome S et al (2014) A conserved peptide pattern from a widespread microbial virulence factor triggers pattern-induced immunity in *Arabidopsis*. *PLoS Pathog*. <https://doi.org/10.1371/journal.ppat.1004491>
- Cabral A, Oome S, Sander N et al (2012) Nontoxic Nep1-like proteins of the downy mildew pathogen *Hyaloperonospora arabidopsis*: Repression of necrosis-inducing activity by a surface-exposed region. *Mol Plant-Microbe Interact* 25:697–708. <https://doi.org/10.1094/MPMI-10-11-0269>
- Oome S, Raaymakers TM, Cabral A et al (2014) Nep1-like proteins from three kingdoms of life act as a microbe-associated molecular pattern in *Arabidopsis*. *Proc Natl Acad Sci USA* 111:16955–16960. <https://doi.org/10.1073/pnas.1410031111>
- Qutob D, Kemmerling B, Brunner F et al (2006) Phytotoxicity and innate immune responses induced by Nep1-like proteins. *Plant Cell* 18:3721–3744. <https://doi.org/10.1105/tpc.106.044180>
- Fellbrich G, Romanski A, Varet A et al (2002) NPP1, a *Phytophthora*-associated trigger of plant defense in parsley and *Arabidopsis*. *Plant J* 32:375–390. <https://doi.org/10.1046/j.1365-313X.2002.01454.x>
- Ottmann C, Luberacki B, Küfner I et al (2009) A common toxin fold mediates microbial attack and plant defense. *Proc Natl Acad Sci USA* 106:10359–10364. <https://doi.org/10.1073/pnas.0902362106>
- Oome S, Van Den Ackerveken G (2014) Comparative and functional analysis of the widely occurring family of Nep1-like proteins. *Mol Plant-Microbe Interact* 27:1081–1094. <https://doi.org/10.1094/MPMI-04-14-0118-R>
- Hyde KD, Jones EBG, Liu JK et al (2013) Families of Dothideomycetes. *Fungal Divers* 63:1–313. <https://doi.org/10.1007/s13225-013-0263-4>
- Kirk P, Cannon P, Minter D, Stalpers J (2008) Ainsworth and Bisby's dictionary of the Fungi, 10th edn. CAB International, Wallingford
- Haridas S, Albert R, Binder M et al (2020) 101 Dothideomycetes genomes: a test case for predicting lifestyles and emergence of pathogens. *Stud Mycol* 96:141–153. <https://doi.org/10.1016/j.simyco.2020.01.003>
- Schoch CL, Crous PW, Groenewald JZ et al (2009) A class-wide phylogenetic assessment of Dothideomycetes. *Stud Mycol* 64:1–15. <https://doi.org/10.3114/sim.2009.64.01>
- Liu JK, Hyde KD, Jeewon R et al (2017) Ranking higher taxa using divergence times: a case study in Dothideomycetes. *Fungal Divers* 84:75–99. <https://doi.org/10.1007/s13225-017-0385-1>
- Ohm RA, Feau N, Henrissat B et al (2012) Diverse lifestyles and strategies of plant pathogenesis encoded in the genomes of eighteen Dothideomycetes fungi. *PLoS Pathog*. <https://doi.org/10.1371/journal.ppat.1003037>
- Grigoriev IV, Nikitin R, Haridas S et al (2013) MycoCosm portal: gearing up for 1000 fungal genomes. *Nucleic Acids Res* 42:D699. <https://doi.org/10.1093/nar/gkt1183>
- Nguyen NH, Song Z, Bates ST et al (2016) FUNGuild: an open annotation tool for parsing fungal community datasets by ecological guild. *Fungal Ecol* 20:241–248. <https://doi.org/10.1016/j.funeco.2015.06.006>
- Emms DM, Kelly S (2015) OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biol* 16:1–14. <https://doi.org/10.1186/s13059-015-0721-2>
- Buchfink B, Xie C, Huson DH (2014) Fast and sensitive protein alignment using DIAMOND. *Nat Methods* 12:59–60. <https://doi.org/10.1038/nmeth.3176>
- Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and

usability. *Mol Biol Evol* 30:772–780. <https://doi.org/10.1093/molbev/mst010>

29. Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T (2009) trimal: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 25:1972–1973. <https://doi.org/10.1093/bioinformatics/btp348>

30. Nguyen LT, Schmidt HA, Von Haeseler A, Minh BQ (2015) IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* 32:268–274. <https://doi.org/10.1093/molbev/msu300>

31. El-Gebali S, Mistry J, Bateman A et al (2019) The Pfam protein families database in 2019. *Nucleic Acids Res* 47:D427–D432. <https://doi.org/10.1093/nar/gky995>

32. Jones P, Binns D, Chang HY et al (2014) InterProScan 5: genome-scale protein function classification. *Bioinformatics* 30:1236–1240. <https://doi.org/10.1093/bioinformatics/btu031>

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

34. Krogh A, Larsson B, Von Heijne G, Sonnhammer ELL (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol* 305:567–580. <https://doi.org/10.1006/jmbi.2000.4315>

35. Emanuelsson O, Brunak S, von Heijne G, Nielsen H (2007) Locating proteins in the cell using TargetP, SignalP and related tools. *Nat Protoc* 2:953–971. <https://doi.org/10.1038/nprot.2007.131>

36. Li Y, Han Y, Qu M et al (2020) Apoplastic cell death-inducing proteins of filamentous plant pathogens: roles in plant-pathogen interactions. *Front Genet*. <https://doi.org/10.3389/fgene.2020.00661>

37. Santhanam P, Van Esse HP, Albert I et al (2013) Evidence for functional diversification within a fungal Nep1-like protein family. *Mol Plant-Microbe Interact* 26:278–286. <https://doi.org/10.1094/MPMI-09-12-0222-R>

38. Seidl MF, Faino L, Shi-Kunne X et al (2015) The genome of the saprophytic fungus *Verticillium tricorpus* reveals a complex effector repertoire resembling that of its pathogenic relatives. *Mol Plant-Microbe Interact* 28:362–373. <https://doi.org/10.1094/MPMI-06-14-0173-R>

39. Baroncelli R, Amby DB, Zapparata A et al (2016) Gene family expansions and contractions are associated with host range in plant pathogens of the genus *Colletotrichum*. *BMC Genom* 17:555

40. Arenas YC, Kalkman ERIC, Schouten A et al (2010) Functional analysis and mode of action of phytotoxic Nep1-like proteins of *Botrytis cinerea*. *Physiol Mol Plant Pathol* 74:376–386. <https://doi.org/10.1016/j.pmp.2010.06.003>

41. Bashi ZD, Hegedus DD, Buchwaldt L et al (2010) Expression and regulation of *Sclerotinia sclerotiorum* necrosis and ethylene-inducing peptides (NEPs). *Mol Plant Pathol* 11:43–53. <https://doi.org/10.1111/j.1364-3703.2009.00571.x>

42. Motteram J, Küfner I, Deller S et al (2009) Molecular characterization and functional analysis of MgNLP, the sole NPP1 domain-containing protein, from the fungal wheat leaf pathogen *Mycosphaerella graminicola*. *Mol Plant-Microbe Interact* 22:790–799. <https://doi.org/10.1094/MPMI-22-7-0790>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.