Nep1-like proteins (NLP) are best known for their cytotoxic activity in dicot plants. NLP are taxonomically widespread among microbes with very different lifestyles. To learn more about this enigmatic protein family, we analyzed more than 500 available NLP protein sequences from fungi, oomycetes, and bacteria. Phylogenetic clustering showed that, besides the previously documented two types, an additional, more divergent, third NLP type could be distinguished. By closely examining the three NLP types, we identified a noncytotoxic subgroup of type 1 NLP (designated type 1a), which have substitutions in amino acids making up a cation-binding pocket that is required for cytotoxicity. Type 2 NLP were found to contain a putative calcium-binding motif, which was shown to be required for cytotoxicity. Members of both type 1 and type 2 NLP were found to possess additional cysteine residues that, based on their predicted proximity, make up potential disulfide bridges that could provide additional stability to these secreted proteins. Type 1 and type 2 NLP, although both cytotoxic to plant cells, differ in their ability to induce necrosis when artificially targeted to different cellular compartments in planta, suggesting they have different mechanisms of cytotoxicity.

The 24-kDa necrosis-and-ethylene inducing peptide 1 (Nep1) of *Fusarium oxysporum*, which was isolated from culture filtrates of this plant-pathogenic fungus (Bailey 1995), is the founding member of a widely occurring protein family. The Nep1 protein induces necrosis and ethylene production when infiltrated into the extracellular space in leaves of dicot, but not monocot, plant species. Since the discovery of Nep1, many Nep1-like proteins (NLP) have been identified, mostly in plant-related microorganisms of both prokaryotic (gram-negative and gram-positive bacteria) and eukaryotic (fungi and oomycetes) origin (Gijzen and Nürnberger 2006; Pemberton and Salmond 2004). As the NLP constituted a new family of proteins, a Pfam domain was defined based on NPP1, an NLP of *Phytophthora parasitica* (Pfam domain PF05630) (Fellbrich et al. 2002). A prominent feature of the NPP1 domain is a seven–amino-acid (heptapeptide) motif, GHRHDWE, which is most strongly conserved in NLP of different species. Based on initial phylogenetic clustering, two types of NLP could be distinguished by the presence of two conserved cysteine residues in type 1 NLP and four in type 2 NLP (Gijzen and Nürnberger 2006). The vast majority of NLP have an N-terminal signal peptide (SP), suggesting they are secreted and function extracellularly. In planta expression of *PsjoNIP*, a type 1 NLP of *Phytophthora sojae*, in soybean hypocotyls (Qutob et al. 2002) and *Arabidopsis thaliana* leaves (Qutob et al. 2006) showed that the protein only caused necrosis, or cell death, when containing its SP but not when expressed without it. This suggested that the NLP acts extracellular of the plant cell to confer its cytotoxic activity.

The mechanism by which NLP induce necrosis is poorly understood. The main question is whether NLP-induced necrosis is caused by stimulation of the plant’s immune system and associated programmed cell death or by direct toxicity of the NLP through disruption of the plant membrane (Gijzen and Nürnberger 2006). Application of recombinant NLP to *A. thaliana* leads to rapid activation of genes associated with defense and cell death (Ba et al. 2006; Qutob et al. 2006), suggesting an active role of the plant in necrosis induction. This was emphasized by the finding that necrosis induction by NPP1 in tobacco requires light and an active host metabolism (Qutob et al. 2006). If NLP would induce necrosis because of their recognition by a receptor, it would likely require only a small conserved peptide fragment (Boller and Felix 2009). However, induction of necrosis by NLP requires the proteins to be mostly intact (Veit et al. 2001), as minimal truncations on either of the N- or C- termini of, for instance, NPP1 of *Phytophthora parasitica* prevent it from inducing necrosis in plants (Fellbrich et al. 2002). Ottmann and associates (2009) showed that NLP permeabilize dicot-derived membrane vesicles in vitro, suggesting that cytotoxicity is the result of membrane leakage, which in turn activates defense responses and cell death. Monocot-derived membrane vesicles were not permeabilized, suggesting NLP cytotoxicity requires a dicot-specific target protein or membrane architecture. The cytolytic activity suggested the protein could form a pore in the plant membrane. However, evidence for a pore-forming activity is lacking so that the exact molecular mechanism of NLP cytotoxicity remains unknown.

The protein structures of two type 1 NLP have been published to date: NLP<sub>ps</sub> of the oomycete *Pythium aphaniderma-tum* (PDB:3GNU) (Öttmann et al. 2009) and MpNLP2 of the Basidiomycete fungus *Moniliophthora perniciosa* (PDB:3ST1) (Zaparoli et al. 2011). These structures revealed that NLP adopt a fold similar to that of pore-forming Actinoporins (Öttmann et al. 2009). The two cysteines that are conserved and essential for induction of necrosis in type 1 NLP (Fellbrich et al. 2002) were confirmed to form a disulfide bridge.
more, three residues (underlined) of the highly conserved GHRHDWE heptapeptide motif were found to be part of an acidic cation-binding pocket. Substitution of these conserved histidine (H), aspartic acid (D), and glutamic acid (E) residues in NLP$_{Py}$ by alanine disabled the induction of necrosis and membrane permeabilization by the protein, suggesting cation-binding is essential for these activities (Ottmann et al. 2009). The acidic pocket was proposed to interact with polar head groups of membrane lipids, thereby damaging or interacting with the plant membrane (Küfner et al. 2009).

NLP could contribute to the virulence of pathogenic microorganisms, e.g., NLP$_{Pc}$ of the soft-rot bacterium Erwinia (Pectobacterium) carotovora subsp. carotovora, which was shown to be essential for virulence on potato tubers (Mattinen et al. 2004). In contrast, virulence of the wheat pathogen Mycosphaerella graminicola was not affected by deletion of its sole NLP gene (Motteram et al. 2009). Also, in several (fungal) species containing multiple NLP, the loss of a single NLP gene did not affect virulence (Arenas et al. 2010; Santhanam et al. 2013; Staats et al. 2007; Zhou et al. 2012), suggesting NLP do not play an essential role in virulence of these organisms. Alternatively, the loss of a single NLP gene in these species may not result in reduced virulence due to genetic redundancy.

In addition to cytotoxic NLP, many noncytotoxic NLP have been described that do not induce necrosis. Two of three tested NLP of Phytophthora infestans (Kanneganti et al. 2006), 11 of the 19 tested NLP of Phytophthora sojae (Dong et al. 2012), all 12 NLP of Hyaloperonospora arabidopsidis (Calbral et al. 2012), five of the seven full-length NLP of Verticillium dahliae (Zhou et al. 2012), and one of two tested NLP of Colletotrichum higginsianum (Kleemann et al. 2012) did not cause necrosis when transiently expressed in plants. The genes encoding noncytotoxic NLP were found to be expressed during the biotrophic and early stages of infection by these (hemibiotrophic) pathogens (Calbral et al. 2012; Dong et al. 2012; Kanneganti et al. 2006; Kleemann et al. 2012), suggesting they play a role during penetration or establishment of infection.

As many new microbial genome sequences became available in the last years, we bioinformatically analyzed the family of NLP by combining phylogenetic analyses and available structural information. Our results suggest that the NLP family consists of three distinct types, each with specific characteristics. We identified an important putative calcium-binding site in type 2 NLP that clearly distinguishes them from type 1 and type 3 NLP. Our study highlights the diversity in phylogenetic distribution, protein sequence, and function of NLP.

**RESULTS AND DISCUSSION**

NLP group in three distinct subfamilies.

Two proteins that represent both described types of NLP, the type 1 NLP$_{Py}$ of the oomycete Pythium aphanidermatum and the type 2 NLP$_{Pc}$ of the bacterium Pectobacterium carotovorum, formed the starting point of our search for NLP homologs. Exhaustive screening of DNA and protein databases for NLP family members using Blast (detailed below) resulted in a collection of 533 NLP homologs from more than 150 different species of bacteria, fungi, and oomycetes (Supplementary Data; Supplementary Table S1), yet not from any other taxonomic group. Phylogenetic analysis of all collected NLP resulted in a neighbor-joining tree (Fig. 1A) that showed a clearly distinguishable third group of NLP proteins (designated type...
3) in addition to the previously described types 1 and 2 NLP (Gijzen and Nünberger 2006). The type 3 NLP are more divergent from types 1 and 2, sharing only a central stretch of 50 amino acids with them, including the highly conserved heptapeptide motif. Sequences N- and C-terminal of these 50 amino acids in the type 3 NLP are very distinct from those of type 1 and type 2 NLP and are, therefore, not comparable. When performing the phylogenetic analysis using the conserved 50 amino-acid central domain of all NLP proteins, the three NLP types still clustered in the same three branches, indicating the type 3 NLP are not a subgroup of the other NLP types (Supplementary Fig. S1).

Of all our 533 collected NLP, 95% carry a SP for secretion, all of them the heptapeptide motif (or variant thereof), and more than 98% of the type 1 and type 2 NLP contain the two conserved cysteine residues that form the disulfide bond in the published NLP structures. Most type 2 NLP have two additional cysteine residues (total of four), while only the type 2 NLP identified in members of genus Bacillus lack these. Type 3 NLP are characterized by a total of six cysteine residues, but these are in the N- and C-terminal part of the proteins that share no similarity to type 1 and type 2 NLP and can, thus, not be aligned with the conserved cysteine in type 1 and type 2 NLP. All type 1 and type 2 NLP match the NLP1 family Pfam domain (PF05630), whereas this is only the case with 66% of the type 3 NLP. The overall architecture of the three NLP types is shown for a representative member of each type, the type 1 NLP Pya, the type 2 NLP Pcc, and the type 3 NLP. The overall architecture of the three NLP types still clustered in the same three branches, indicating the location of the conserved cysteines and heptapeptide motif for all three NLP types.

The distribution of NLP across taxa differs between the three types (Fig. 1C). Whereas type 1 NLP are found in bacte- ria, fungi, and oomycetes, type 2 NLP are found in bacteria and fungi but not in oomycetes. Type 3 NLP have the narrow- est distribution and have only been identified in a restricted number of fungal species and do not occur in bacteria or oomy- cetes. Sixty bacterial species contain a type 2 NLP, only ten contain a type 1 NLP, and, contrary to fungi, none contain both a type 1 and type 2 NLP (Fig. 1C). It is also important to note that the number of bacterial species having one or more NLP is rather limited, knowing that several thousands of bacterial genomes have been sequenced. In oomycetes, only type 1 NLP are found, which form a large clade as a result of expansion of the number of NLP genes within oomycete species belonging to the plant-infecting order Peronosporales.

In fungi, species with all possible combinations of the three types of NLP are found. The genomes of five fungal species encode all three NLP types. Phylogenetic clustering of these fungal NLP resulted in three major clades (Supplementary Fig. S2) fitting the three NLP types shown in Figure 1A. The fact that the three types only occur together in fungal species belonging to the Ascomycetes suggests that the evolutionary origin of the NLP is in this phylum. Duplication of an ancestor in this phylum. Duplication of an ancestral NLP gene and diversification events then led to the three types that expanded even further in a selection of fungal spe- cies, mostly plant-associated ones. Oomycete NLP all cluster together, suggesting a single horizontal gene transfer (HGT) event of an Ascomycete fungus type 1 NLP to a common ancestor of members of the order Peronosporales, which then further expanded within this lineage. The latter hypothesis was also proposed by Richards and associates (2011), as NLP are one of the 20 gene families predicted to be acquired by oomy- cetes from fungi through HGT. HGT is also the most plausible explanation for the occurrence of type 2 NLP in a limited number of bacterial species belonging to the Gammaproteobacteria and Actinobacteria.

Type 3 NLP are specific for ascomycetes.

Type 3 NLP are only found in orders of the Ascomycete fungi (Supplementary Fig. S3). With 33 members identified thus far, they are not as widespread as types 1 and 2 NLP. The fungi containing type 3 NLP have many different lifestyles; in addition to plant pathogens, type 3 NLP are, among others, found in the plant endosymbiont Oidiodendron maius, the nematode pathogen Arthrobotrys oligospora, the bat pathogen Geomyces destructans, and the coprophilous Ascosbolus immersus. In the latter species, which lacks type 1 and type 2 NLP, we find the highest number (three) of type 3 NLP of all examined species. Little is known about these proteins, and thus far, experimental data is only available for the type 3 NLP (Afu5g02100) of the opportunistic human pathogen Aspergillus fumigatus, which was identified in a proteomics analysis of secreted fibrinogen-binding proteins (Upadhyay et al. 2012). Preliminary data suggest that expression of the Verticillium dahliae type 3 NLP VDAC 07972T0 is induced during infec- tion of tomato (P. Santhanam and B. Thomma personal communi- cation). Also, expression of a type 3 NLP was observed during infection of insects by Metarhizium anisopliae (EFY97649) and M. acridum (EFY91472) (Gao et al. 2011; C. Wang personal communication).

Type 2 NLP are widely distributed in fungi and bacteria with various lifestyles.

Fungal type 2 NLP are found in many orders of the Asco- mycota (Fig. 2), and in one species of the primitive chytrid fungi (Gonapodya prolifera, phylum Monoblepharidomy- cetes). Based on its phylogenetic clustering, the latter likely obtained a type 2 NLP from a Gammaproteobacterium (Fig. 2). Type 2 NLP are widely spread in two groups of bacteria; they were identified in species from several orders of Gamma- proteobacteria and in species of the phylogenetically distant region of Actinomycetales. The only exception to this is a type 2 NLP homolog identified in a small number of species from the genus Bacillus (phylum Firmicutes). These are also the only type 2 NLP lacking the characteristic second disulfide bridge. In addition, these Bacillus NLP are fused to a lectin domain that shares homology with ricin B (Schouten et al. 2008). Their clustering with the NLP of the distantly related Actinobacteria and their G/C content (Supplementary Fig. S4) suggest HGT from a Streptomyces species to these Bacillus species. Another possible HGT is likely responsible for the occurrence of a type 2 NLP in a Streptomyces species (Strepto- myces sp. strain Mg1_2) that clusters within the Ascomycete NLP.

Type 2 NLP are found in plant-associated microbes, e.g., the necrotrophic bacterium Pectobacterium carotovorum, the dicot plant pathogen Necricia haematococca, and the monocot symbiont Epichloë festucae, but also in non-plant-associated microbes, e.g., in the coprophilous Podospora anserina. Interest- ingly, NLP were also found in insect pathogenic fungi of the genera Beauveria, Cordiceps, and Metarhizium, which have also been isolated as plant endophytes (Vega et al. 2009). During infection of insects by the fungi M. anisopliae and M. acridum, expression of type 2 NLP was detected at 24 h post- inoculation (Gao et al. 2011; C. Wang personal communication), suggesting these proteins may play a role in insect patho- genesis. Animal cell cytoly sis caused by NLP has been tested on sheep erythrocytes. However, no hemolysis was observed in response to the type 1 NLP PpNPP1 of Phytophthora parasitica, which is cytolytic in dicot plants (Qutob et al. 2006). Also, other animal-related microorganisms harbor type 2 NLP, e.g., the coral pathogen Vibrio coralliilyticus (O de Santos et al. 2011) and the bivalve endosymbiont Teredinibacter turnerae (Yang et al. 2009). Strikingly, type 2 NLP are also found in the
Fig. 2. Phylogenetic distribution of type 2 Nep1-like proteins (NLP). Aspergillus includes Neosartoria spp., Glomerella includes Colletotrichum spp., Fusarium includes Gibberella spp., and Erwinia includes Pectobacterium spp. NLP shown: ScoeNIP (Qutob et al. 2002); NLP_Pcc (Müllin et al. 2004, Pinnab et al. 2009); NLP_Eca (Pemberton et al. 2005); ChNL4 (Kleemann et al. 2012); VdNLP4, VdNLP5, VdNLP7, VdNLP9 (Zhou et al. 2012). Additional disulfide bridges found in the indicated branches are discussed below. Unrooted maximum likelihood tree of 61 fungal and 61 bacterial type 2 NLP. Only bootstrap values $>60$ are shown.
animal-associated *Bacillus cereus* and *B. thuringiensis* (Jensen et al. 2003), while type 1 NLP are only found in this genus in the plant-associated *B. subtilis* (Earl et al. 2008), its close relative *B. licheniformis*, and the alkalophilic plant litter–degrading *B. halodurans* (Takami et al. 2000). Only two of the six tested type 2 NLP have been shown to induce necrosis, NLP$_{Ecc}$ (NLP$_{Ecc}$) from *Pectobacterium (Erwinia) carotovorum* subsp. *carotovorum* (Mattinen et al. 2004) and NLP$_{Eca}$ from *Pectobacterium carotovorum* subsp. *atrosepticum* (Pemberton et al. 2005). None of the four type 2 NLP of *Verticillium dahliae* induce necrosis when transiently expressed in plants (Zhou et al. 2012).

**Fig. 3.** Phylogenetic relationship between fungal and bacterial type 1 Nep1-like proteins (NLP). *Aspergillus* includes *Neosartoria* spp., *Glomerella* includes *Colletotrichum* spp., *Botrytis* includes *Botryotinia* spp., and *Fusarium* includes *Gibberella* spp. NLP shown: NEP1 (Bailey 1995); BhalNIP (Qutob et al. 2002); BeNEP1, BeNEP2 (Staats et al. 2007); BcNEP1, BcNEP2 (Schouten et al. 2008); MgNLP (Motteram et al. 2009); SsNEP1, SsNEP2 (Bashi et al. 2010); VdNEP1, VdNEP2, VdNEP3 (Zhou et al. 2012); ChNLP1, ChNLP2, ChNLP3, ChNLP5 (Kleemann et al. 2012); Nep1$_{Mo}$ (Zhang et al. 2012). Additional disulfide bridges found in the indicated branches are discussed below. Unrooted maximum likelihood tree of 135 fungal and 12 bacterial type 1 NLP. Only bootstrap values >60 are shown.
Fig. 4. Phylogenetic distribution of Nep1-like proteins (NLP) in oomycetes. A, Unrooted maximum likelihood tree of 231 type 1 oomycete NLP. The tree can be divided in three groups: type 1 NLP, type 1 NLP with an added region consisting of one or both a Q-rich region and a TPAP repeat, and type 1a NLP. *Phytophthora* contains *P. infestans*, *P. capsici*, *P. sojae*, *P. ramorum*, *P. parasitica*, *P. megakarya*, and *P. cinnamomi*. NLP shown: PmegNEP1-6 (Bae et al. 2005); PiNPP1.1, PiNPP1.2, PiNPP1.3 (Kanegami et al. 2006); MpNEP1, MpNEP2 (García et al. 2007); PpNPP1, NLP Pya (Ottmann et al. 2009); PsojNIP, HaNLP1-10 (Cabral et al. 2012); PsNLPI, 2, 3, 5, 7, 8, 9, 11, 14, 19, 21, 23, 24, 25, 27, 29, 30, 32, 35, 37, 38, 39, 40, 41, 42, 44, 51, 54, 55, 57, 58, 59, and 60 (Dong et al. 2012). Additional disulfide bridges found in the indicated branches are discussed below. Only bootstrap values >60 are shown. B, Schematic representation of an NLP with a Q-rich region and TPAP-repeat, indicating the relative positions of the Q-rich region and TPAP repeat.
Type 1 NLP are found mostly in plant-associated species.

Type 1 NLP form a monophyletic group and can be divided into fungal/bacterial NLP and oomycete NLP (Fig. 1A). A closer look at the fungal/bacterial NLP (Fig. 3) shows most of the fungal type 1 NLP are identified in many orders of the phylum Ascomycota. In this phylum, type 1 NLP are found mostly in genera with plant-associated lifestyles, ranging from the symbiotic (endomycorrhizal) genus *Oidiodendron* to the necrotrophic *Sclerotinia* and *Botrytis* pathogens. The only included fungal species not directly related to plants is the saprobic wine cellar fungus *Zasmidium cellare* (yet it belongs to the family *Mycosphaerellaceae*, which contains many plant pathogens [Crous et al. 2009]). Surprisingly, as monocots are insensitive to NLP (Bailey 1995), NLP genes have been found in fungi that solely interact with monocot plants, e.g., the grass pathogens belonging to the genera *Pyrenophora*, *Cochliobolus*, *Setosphaeria*, and *Magnaporthe*. This suggests that, in these species, cytotoxic type 1 NLP may have a function other than killing host cells. Only two fungal NLP in this group are found in species of phylum Basidiomycota, i.e., *Polyporus arcularius* and *Ganoderma* sp., which are both wood-degrading species of the order of Polyporales.

Bacterial species containing type 1 NLP are not very numerous. Phylogenetic distribution suggests that, in the evolutionary past, a type 1 NLP gene was acquired by a bacterial species, from which it was subsequently transferred to other bacterial species. Type 1 NLP are found in a few distantly related species of Proteobacteria (Rhizobiaceae and a single *Pseudomonas* species) and Firmicutes (*Bacillus* species). Of these bacteria, only *Pseudomonas psychrotolerans* (Hauser et al. 2004) may not be directly plant-associated.

Of 14 tested fungal type 1 NLP described in literature, 12 were found to induce necrosis. The heptapeptide is highly conserved in most bacterial and fungal NLP, except for the NLP found in the Rhizobiaceae family and a branch of type 1 NLP found in the genus *Glomerella/Colletotrichum*. The only tested NLP of this last group (ChNLP3 from *Colletotrichum higginsianum*) did not induce necrosis, as tested by transient expression in *Nicotiana benthamiana* leaves (Kleemann et al. 2012).

**Diversification of oomycete type 1 NLP.**

In oomycetes, NLP are only found in the Pythiaceae and Peronosporaceae families (Fig. 4A) of the Peronosporales order. They are, so far, not found in the third family of this order, the...
Albuginaceae, as the genomes of both Albugo candida (Links et al. 2011) and Albugo laibachii (Kemen et al. 2011) completely lack NLP genes. All oomycete NLP form a separate branch with a common origin (Figs. 1A and 4A), and are not found together with fungal or bacterial NLP, suggesting that oomycetes have acquired these genes only once. The only exceptions are the NLP of the Basidiomycete fungi Moniliophthora perniciosa and Moniliophthora roseri, which group together with the oomycete proteins (Fig. 4A), suggesting they are the result of HGT. Based on an aberrant G/C content, this gene transfer has likely occurred from an oomycete to these fungi (Tiburcio et al. 2010). The oomycete NLP are thus far exclusively found in plant-pathogens, e.g., the necrotrophic Pythium, the hemibiotrophic Phytophthora, and the obligate biotrophic Peronospora, Hyaloperonospora, Bremia, and Plasmopara genera. Many oomycete NLP do not induce necrosis (Fig. 4A); of 32 tested NLP, only 10 are cytotoxic, Phytophthora infestans (Kanneganti et al. 2006), Phytophthora sojae (Dong et al. 2012), and H. arabidopsis (Cabral et al. 2012) all express NLP during infections that have been shown not to induce necrosis.

In the Peronosporaceae, diversification of type 1 NLP can be observed (Fig. 4A). A subgroup of proteins has an additional hydrophilic domain between the SP and NPP1 domain. Within this “glutamine and proline-rich hydrophilic domain” (Gijzen and Nürnberger 2006), two regions are distinguished: i) a glutamine (Q)-rich region and ii) a TPAP repeat (Dong et al. 2012). These regions can be present independently or together; in the latter case, the Q-rich region is always more N-terminal than the TPAP repeat (Fig. 4B). Adding this domain rich in glutamine or proline, or both, to a cytotoxic NLP or removing it from a cytotoxic NLP that already contains one does not abolish necrosis-inducing activity in transient expression in Nicotiana sp. leaves (Cabral et al. 2012; Dong et al. 2012). This suggests that the necrosis-inducing activity of NLP is independent of the presence or absence of these additional domains.

The proline-rich hydrophilic region has also been described for the noncytotoxic HaNLP9 of H. arabidopsis and was predicted to be O-glycosylated (Cabral et al. 2012). In 27 of the identified Phytophthora NLP, this proline-rich region forms TPAP repeats (Supplementary Fig. S5) that fit the maximum consensus sequence for O-glycosylation (Nishikawa et al. 2010; Wilson et al. 1991). This is also predicted by NetOGlyc3.1 (Julienius et al. 2004), suggesting a high degree of O-glycosylation in these parts of the proteins. The number of TPAP repeats varies from five (Ps_138848, PPTG_08005) up to 14 (Ps_127887). Similar potentially O-glycosylated regions have been described for Phytophthora sojae elicitors (Qutob et al. 2003). The O-glycosylated TPAP-repeat domain has been proposed by Jentoft (1990) to adopt a rigid and extended conformation, like a rod or stick that can penetrate the protective extracellular layer surrounding cells.

Many oomycete NLP have substitutions in the GHRHDWE heptapeptide motif; only 38% of oomycete NLP contain the fully conserved motif. Examination of the phylogenetic distribution of the oomycete NLP with alterations in the heptapeptide motif revealed that most of these cluster together (Figs. 1A and 4A) in one group, which was designated type 1a.

**Comparing NLP proteins.**

To gain insight into the similarities and differences among and between the three types of NLP proteins, sequences were aligned and consensus sequences were generated. Based on phylogenetic distribution, we separated the NLP into five distinct groups: fungal/bacterial type 1, oomycete type 1, oomycete type 1a, type 2, and type 3. For each type, the conservation of amino-acid residues was determined (details below). Each of these groups is represented by a consensus Weblogo, displaying the conserved residues per group. The Weblogos of the first four groups were aligned to a characterized member of their respective type (NLP<sub>Py<Psub>a</sub> for all type 1 consensus Weblogos and NLP<sub>Pv</sub> for the type 2 consensus Weblogo) (Fig. 5A). All numbering of type 1 residues is based on the NLP<sub>Py<Psub>a</sub> structure 3GNU (which lacks the SP), and numbering of type 2 NLP is based on NLP<sub>Pv</sub> starting directly after its predicted SP of 22 amino acids. The type 3 consensus Weblogo was found to be too divergent for proper alignment with the other four groups, and is therefore discussed separately below. It was found that the first four groups of NLP differ substantially from each other.

---

**Fig. 6.** Effect of alanine substitutions of the putative calcium-binding site and proximate conserved residues of NLP<sub>Pv</sub> (coordinates based on protein without a signal peptide [SP]). A, Model of NLP<sub>Pv</sub> indicating the location of the mutated residues (black) and the conserved cysteins (white). B, 3D model of NLP<sub>Pv</sub>. C24, D26, and D28 are indicated in red, L31 and R57 in blue, E166 and E169 in yellow. NLP<sub>Pv</sub> was modeled on 3GNU using I-Tasser and made with Polyview 3D. C, Agroinfiltrations of the various mutants in tobacco leaves. Changing the three conserved aspartic acids into alanines completely abolishes necrosis-inducing activity, while L31A and R57A do not affect the cytotoxic activity. Changing the two conserved glutamic acids to alanines results in a reduced activity. Plants were followed to day 10; the last 5 days, no more changes in necrosis were observed. Standard deviation represent 3 × 15 plants. D, Changing the two conserved glutamic acids to alanines results in an incomplete, patchy necrosis. Photograph taken at day 5.
Oomycete type 1a NLP lack the acidic cation-binding pocket.

The N-termini, between the SP and the first conserved cysteines (shown in orange in Figure 5) of type 1 and type 2 NLP, differ too much for proper alignment (Fig. 5A). In type 1 NLP, this part contains four highly conserved residues (I3, H5, D6, and V8), which, together with Y40, D44, T49, and S50, are in close proximity to each other in the 3D structure of NLP_pya (Supplementary Fig. S6A), forming a region not present or not conserved in type 2 NLP.

When comparing type 1 NLP of fungi and bacteria with those of oomycetes, 12 amino-acid positions were found to be highly conserved in oomycete NLP but less in fungal/bacterial NLP: P10, T18, K22, A24, K26, K28, Q30, S39, S69, S124, D172, Q178, and R184. Except for S124, all these residues are on the surface and are located on one side of the protein (Supplementary Fig. S7). Notably, a number of positively charged amino acids is highly conserved, suggesting these oomycete proteins have a local surface charge different from fungal type 1 NLP. When comparing the cytotoxic fungal/bacterial and oomycete type 1 NLP to the noncytotoxic oomycete type 1a NLP, several differences are observed. Of the seven amino acids making up the heptapeptide motif, type 1a NLP only show conservation of three (R102, H103, and W105), whereas four others do not (G100, H101, D104, and E106). Of these latter four, three residues (H101, D104, and E106) have been shown to be crucial for necrosis induction in NLP_pya and are part of its acidic cation-binding pocket (Ottmann et al. 2009). In each type 1a NLP, one or more of the eight residues making up this pocket (D93, H101, D104, E106, H128, Y151, D/N158, and H159) (Fig. 5B) are substituted. Also, three residues near this pocket (G100, G193, and A195) that are highly conserved in both fungal type 1 NLP, fungal/bacterial type 1 NLP, and type 2 NLP are not conserved in type 1a NLP, suggesting they play a role in cytotoxicity. The same 11 residues are substituted in most Rhizobiaceae NLP and in the noncytotoxic branch of Glomerella/Colletotrichum NLP in which the heptapeptide was found to have substitutions (Fig. 3). The only residue that is highly conserved in type 1a NLP but not in the other type 1 NLP is D200 (corresponding to K200 in NLP_pya).

Type 3 NLP contain all cation-binding residues.

As type 3 NLP share no clear homology with the other two types of NLP, except for the 50 residues surrounding the heptapeptide motif, the rest of the sequence cannot be compared with type 1 and type 2 NLP. Most type 3 NLP (22 of 33) are characterized by having six conserved cysteines corresponding to positions 39, 67, 74, 76, 99, and 270 in Afu5g02100 (Supplementary Fig. S9A). Eleven type 3 NLP do not have cysteine residues corresponding to positions 67 and 74, suggesting that in the other 22 type 3 NLP C67 and C74 form a disulfide bridge.

In a single protein (ELR02678 from Geomyces destructans), cysteines 39 and 76 have been substituted for a threonine and a serine, respectively, while the other four cysteines are present, suggesting C39 and C76 form a disulfide bridge as well in the other proteins. The 50-amino-acid region that contains the heptapeptide motif can be aligned with the homologous region from type 1 and type 2 NLP. Strikingly, all residues in this region of NLP_pya, which make up the acidic cation-binding pocket (D93, G100, H101, D104, E106, and H128) and which are also found in the cytotoxic type 2 NLP_pcc, are highly conserved in type 3 NLP, suggesting that type 3 NLP also bind a cation. The amino-acid sequences N- and C-terminal of the conserved 50-residue-stretch are very different from those of types 1 and 2 NLP. The arginine and histidine residues (underlined) of the heptapeptide motif (GHRHDWE), which are not conserved in type 3 NLP, are not surface exposed but, rather, seem to play a structural role in the NLP_pya protein.

Type 2 NLP contain a putative calcium-binding motif.

Based on their amino-acid sequence, type 2 NLP share an overall structure with type 1 NLP, yet there are specific differences. In type 2 NLP, we find a strongly conserved KxYxHK DxxxDxxRXA motif (residues 146 to 162 in NLP_pcc), of which the five underlined residues show no conservation in type 1 NLP (Fig. 5A). The corresponding five positions in NLP_pya (S150, K152, S153, T154, and G162) are in close proximity to each other and locate to two antiparallel beta sheets, suggesting they form a positively charged patch.

Besides this region, type 2 NLP contain two other regions that differ strongly from type 1 NLP. Interestingly, several residues in these regions show a high level of conservation in type 2 NLP. In NLP_pcc, the first region of 28 amino acids located between the SP and the glycine residue preceding the first conserved cysteine contains six conserved residues (Fig. 5A). A second, nine-amino acid region (of which three are conserved), corresponding to residue E166 to K174 in NLP_pcc, is predicted to form an additional (surface-exposed) loop and corresponds to a smaller region in NLP_pya (between T165 and G168). The 28–amino acid N-terminal region contains a highly conserved DxxDxDG motif (resides 24 to 29), which could function as a calcium-binding motif (Rigden and Galperin 2004). The second region (166 to 174 in NLP_pcc) contains three conserved residues: E166, E169, and N170. The two corresponding regions in NLP_pya (L31-N35 and T165-G168) are in close proximity. This suggests the two acidic regions form a binuclear calcium-binding center (Rigden and Galperin 2004). When modeled over the structure of NLP_pya, the putative calcium-binding center of NLP_pcc is visible as a concentrated patch of negatively charged residues (Fig. 6A and B). Other residues conserved in type 2 but not in type 1 NLP correspond to NLP_pcc L31, L38, N43, R57, R64, E91, and F175.

To investigate whether this negatively charged surface region of NLP_pcc is essential for necrosis induction, alanine-substituted versions were created in the conserved DxxDxD motif and in residues E166 and E169 as well as conserved residues L31 and R57, which are predicted to be on the surface of the protein, close to the DxxDxD motif (Fig. 6B) and conserved in type 2 but not in type 1 NLP (Fig. 5A). Transient expression of these alanine-substituted proteins showed that the DxxDxD to AAXA variant of NLP_pcc no longer induces necrosis, while the L31A and R57A variants did not show a marked reduction in the induction of necrosis (Fig. 6C). NLP_pcc, with the substitutions E166xxE169 to A166xxA169 induced necrosis formation, although it was delayed and patchy (Fig. 6D). These data suggest that the conserved acidic residues, predicted to be involved in calcium binding and lacking in type 1 NLP, are important for necrosis induction by type 2 NLP.

NLP have a high number of different potential disulfide bridges.

While comparing the obtained type 1 and type 2 protein sequences, we noticed several conserved cysteine pairs that could form disulfide bridges. Cysteine residues that are part of disulfide bridges are in close proximity in the 3D structures and are orientated such that their sulfur groups are able to form a covalent bond. When, for a given protein, the crystal structure of a homologous protein is available, disulfide bridges can be predicted based on the occurrence of cysteines and their location in the protein. The distance between the α-carbons (Cα) of the two cysteines provides a measure of the proximity in the protein structure, whereas the distance between the β car-
bons (C \( \beta \)) provides a proxy for the distance of the corresponding sulfur groups (S) (Fig. 7A).

The crystal structures of \( \text{NLP}_{\text{Py)} (\text{Ottmann et al. 2009}) \) and \( \text{MpNEP2 (Zaparoli et al. 2011}) \), show that the two strongly conserved cysteines, which are present in both type 1 and type 2 NLP, form a disulfide bridge, which we refer to as “bridge A.” A second pair of conserved cysteines is found in most type 2 NLP. As no type 2 NLP structure is available, we interrogated the positions (outlined below) of the corresponding residues in \( \text{NLP}_{\text{Py}} (T77 \text{ and Y82}) \) and \( \text{MpNEP2 (T92 \text{ and T97})} \). We concluded that these residues are in close proximity, making it likely that the corresponding cysteines in type 2 NLP form a second disulfide bridge (Fig. 7), which we refer to as “bridge B.”

A search in our collected type 1 and type 2 NLP sequences resulted in more than 100 sequences containing two or more extra cysteine residues in addition to the cysteines making up one or both bridge A and B. To predict if the additional cysteine residues are located in sufficient proximity to form a disulfide bridge, we located the corresponding residues for each of them in the structures of \( \text{NLP}_{\text{Py}} (3\text{GNU}) \) and \( \text{MpNEP2 (3ST1}) \) (Fig. 7B). Based on the measured distances of cysteine residues, 12 additional potential disulfide bridges in the NLP family were identified in types 1 and 2 NLP (Fig. 7C). The occurrence of these extra disulfide bridges in NLP is indicated in the phylogenetic protein trees (Figs. 2, 3, and 4). Of these 12 identified additional disulfide bridges, only three (bridges 6, 7, and 9) are occurring in more than 10 NLP. Bridge 6 is found in a branch of oomycete NLP containing both cytotoxic (PsNLP59; Dong et al. 2012) and noncytotoxic (HaNLP3, Cabral et al. 2012) NLP. Bridge 7 is found in a branch of fungal type 1 NLP. In one of the NLP containing this bridge (BcNEP2; Arenas et al. 2010), this bridge was disrupted by amino-acid substitutions without any change in necrosis-inducing activity. Bridge 9 was identified in 16 fungal type 2 NLP and in one unrelated bacterial type 2 NLP, suggesting this bridge has evolved twice independently.

In three additional cases, one of the cysteines potentially making up the disulfide bridge is located in a region just N or C terminal of the part of the proteins that is present in the crystal structures of \( \text{NLP}_{\text{Py}} \) and \( \text{MpNEP2} \), thus not allowing exact distance measurements of C \( \alpha \) and C \( \beta \) of their corresponding residues. Of these, one cysteine pair is present in 14 NLP of species in the genus Glomerella/Colletotrichum, another in four NLP of \( H. \text{arabidopsisid} \) (HaNLP1, 2, 4, and 6) and in a Peronospora tabacina homolog (Supplementary Fig. S10), and a third in HaNLP2. Including these three additional disulfide bridges, 17 potential disulfide bridges are found in the NLP family. Previously, fifteen distinct disulfide bridges were found in the family of \( \beta \)-glucanases that comprises cellulases, glucanases, xylanases, glucuronidases, and mannanases encompassing 14 Pfam domains (Thangudu et al. 2008). In contrast, the NLP family only has a single Pfam domain and has, to our knowledge, the highest number of distinct potential disulfide bridges found in any protein family.

**Functionality of disulfide bridges in the NLP.**

Formation of bridge A seems essential for necrosis induction, as PpNPP1 of *Phytophthora parasitica* (Fellbrich et al. 2008) and noncytotoxic (HaNLP3, Cabral et al. 2012) NLP. Bridge 7 is found in a branch of fungal type 1 NLP. In one of the NLP containing this bridge (BcNEP2; Arenas et al. 2010), this bridge was disrupted by amino-acid substitutions without any change in necrosis-inducing activity. Bridge 9 was identified in 16 fungal type 2 NLP and in one unrelated bacterial type 2 NLP, suggesting this bridge has evolved twice independently.

**Fig. 7.** Occurrence of possible disulfide bridges in type 1 and type 2 Nep1-like proteins (NLP). A, Disulfide bridges in proteins consist of two cysteine residues, each possessing an \( \alpha \) carbon, a \( \beta \) carbon, and a sulfur. Distances of these atoms between the cysteines are an indication of the disulfide-forming capability of two closely positioned cysteines. B, Distances of the pairs of \( \alpha \) carbons and pairs of \( \beta \) carbon of the predicted disulfide bridges in NLP\(_{\text{Py}}\) (3GNU) and MpNEP2 (3ST1). Residue numbering is based on PDB:3GNU. Asterisks (*) indicates a glycine (lacking a \( \beta \) carbon) located at one of the positions of the possible disulfide bridge. In these cases the glycine was replaced in the structure by an alanine, using I-Tasser, and the position of the alanine \( \beta \) carbon was used. C, Position of the predicted disulfide bridges in different NLP plotted on a line drawing of NLP\(_{\text{Py}}\).
2002), BcNEP1, and BcNEP2 of Botrytis cinerea (Arenas et al. 2010) do not induce necrosis when one of these cysteines is substituted. To test the requirement of the strongly conserved bridge B in type 2 NLP for induction of necrosis, mutant proteins of NLP<sub>PCC</sub> were created in which bridge A, bridge B, or both are disrupted by substituting one of the cysteines by a serine (Fig. 8A). Disruption of bridge A in both PsojNIP and NLP<sub>PCC</sub> led to a complete loss of necrosis induction (Fig. 8B), while disruption of bridge B in NLP<sub>PCC</sub> did not affect its cytotoxicity. Also, simultaneous disruption of both bridges A and B led to a complete loss of necrosis induction. These results suggest that bridge B, although highly conserved in type 2 NLP, is not essential for necrosis induction but has possibly evolved to enhance protein stability.

**Necrosis induction by NLP targeted to different plant cell compartments.**

Experiments by Qutob and associates (2002, 2006) indicated that the induction of cell death by transient expression of PsojNIP in plant cells requires the presence of a SP, suggesting that NLP require an as-yet-unidentified target located on the extracellular side of dicot plasma membranes. However, it is also possible that disulfide bridge A, which is required for necrosis induction by both PsojNIP and NLP<sub>PCC</sub> (Fig. 8B), may not be created when the proteins are produced in the plant cell cytoplasm, an environment that does not support the effective formation of disulfide bonds (Raina and Missiakas 1997). To test if PsojNIP and NLP<sub>PCC</sub> would induce cell death when produced in the endoplasmic reticulum (ER) (in which cysteine bonds are formed [Raina and Missiakas 1997]), we added a C-terminal KDEL sequence to these NLP. Proteins with a C-terminal KDEL sequence that enter the ER are unable to exit it. Leakage from the ER is unlikely, as protein retention in the ER is not significantly impaired by an increased concentration of KDEL-carrying proteins (Denecke et al. 1992). The KDEL sequence is highly specific; the chemically similar sequence KDDL does not function as an ER retention signal (Denecke et al. 1992) and can, therefore, be used as a control. As shown in Figure 9, NLP<sub>PCC</sub> with a C-terminal KDEL sequence does not induce any visible necrosis, while its control, KDDL, gives wild-type levels of necrosis. NLP<sub>PCC</sub> lacking its SP also does not induce any necrosis. In contrast, PsojNIP with a C-terminal KDEL sequence has a one-day delay in onset of necrosis, yet reaches full necrosis at day 4. To our surprise, PsojNIP, lacking a SP, still caused necrosis, although delayed compared with the wild-type protein, while NLP<sub>PCC</sub> without a SP did not induce any necrosis.

Although transient expression in planta is a highly artificial system and may not mimic the natural situation, the difference between type 1 PsojNIP and type 2 NLP<sub>PCC</sub> is remarkable and may suggest these proteins do not have the same mode of action, although they both induce a similar necrotic response and type 1 NLP can restore virulence of the Pectobacterium carotovorum NLP<sub>PCC</sub> mutant (Ottmann et al. 2009). A difference is also observed when the wild-type proteins (with SP)
are tested by transient expression; NLP<sub>Po</sub> requires about one less day than PsojNIP to induce necrosis (Figs. 8 and 9).

**Concluding remarks on the versatile NLP protein family.**

We have defined four different groups of NLP belonging to three phylogenetic types, which are schematically represented in Figure 10. Type 1 NLP are characterized by the presence of (at least) a single conserved disulfide bridge (C-C), an acidic cation-binding pocket, and an exposed region, which are all known to be required for cytotoxicity (Cabral et al. 2012; Fellibrich et al. 2002; Ottmann et al. 2009). Type 1a NLP, which diverged from type 1 NLP and are strongly expanded in oomycetes, are similar to type 1 NLP yet have substitutions in the acidic cation-binding pocket that is conserved in all other NLP types. Type 2 NLP are not only distinguished by a second conserved disulfide bridge but, more importantly, by the presence of a putative calcium-binding domain, equivalent to the exposed region found in types 1 and 1a NLP and that is required for the induction of necrosis. In contrast to type 1 NLP, no type 2 equivalent of the type 1a NLP (i.e., type 2 NLP that lack the acidic cation-binding pocket) were identified. Our knowledge of type 3 NLP is still very limited. Based on the presence of conserved cysteines, we predict that most type 3 NLP contain three disulfide bridges. In addition, the strong conservation of residues that make up the cation-binding pocket in types 1 and 2 NLP suggests that type 3 NLP also contain a similar pocket. The N- and C-terminal parts of type 3 NLP are very different from type 1 and type 2 NLP, precluding any other structure-based predictions.

In the described research, a large number of newly identified NLP sequences from a broad range of diverse organisms was studied to provide insight into the evolution and functions of these remarkable proteins. Phylogenetic analysis revealed three distinct types (and an additional subtype) of NLP, in contrast to the previously described two NLP types (Gijzen and Nünberger 2006). While type 1 NLP are found almost exclusively in plant-associated microbes, type 2 and type 3 NLP occur in microorganisms with different lifestyles. As only Ascomycete fungi have all three NLP types, it is tempting to speculate that NLP find their origin in this group of fungi. In oomycetes, only type 1 NLP occur, which have strongly expanded in this group of plant pathogens. In addition, oomycetes have evolved NLP with extensive N-terminal modifications, of which the role is not clear, yet these additional domains do not interfere with the induction of necrosis. The oomycete type 1a NLP have extensive substitutions in the acidic cation-binding pocket and have lost the ability to induce necrosis. Similar substitutions in the acidic cation-binding pocket were also found in type 1 NLP of members of the plant symbiotic group of Rhizobiaceae and in the hemibiotrophic fungi of the genus *Colletotrichum*, suggesting these have evolved at least three times independently. This suggests that noncytotoxic type 1 NLP lacking the cation-binding pocket serve an as-of-yet unknown function that is not related to the induction of necrosis.

Our multifaceted analysis showed that the NLP domain is highly versatile and has been the basis for at least four different groups of NLP. These proteins are very widespread, both phylogenetically and among microbial lifestyles. In many species, more than one of these groups co-occur, suggesting a different function for each of them. This hypothesis is supported by the finding that necrosis-inducing types 1 and 2 NLP do not only contain different conserved motifs but, also, exhibit different responses when transiently expressed in planta. All these distinctions should be taken into account and may be of great help in deciphering the molecular mechanisms and biological roles of the members of this fascinating protein family.

**MATERIALS AND METHODS**

**Finding NLP genes.**

Databases (National Center for Biotechnology Information, FungiDB, Joint Genome Institute, and Endophyte) were searched for NLP homologs using BlastP with PsojNIP (type 1, NLP<sub>Po</sub> (type 2), and Afu5g02100 (type 3) as queries. Only for searching the Endophyte database, TBlastN was used, because BlastP was not available. Only full-length protein sequences were included in the analysis.

**Construction of phylogenetic trees.**

Neighbor-joining trees were created by aligning the NLP protein sequences with MEGA5 (ClustalW, standard setting, 1,000 bootstraps). Maximum likelihood trees were created by aligning individual NLP types using mafft (LINSi, GINSi, and EINSi) and muscle. Positions that were present in fewer than 80% of the aligned proteins were removed from the alignment, using a customized script. In both methods, all present SP were predicted by SignalP4.1 and were removed before analysis, as these can share similarity not based on ancestry but on physico-chemical constraints (Wong et al. 2010). The best scoring alignment (mafft-LINSi) was chosen based on the quality indications from normd and mumsa. Phylogenetic trees were constructed using phylm (v3.0, WAG model of amino-acid substitution, number of substitution rate categories 4, estimated gamma distribution parameters, and estimated proportion of invariable sites), and the robustness was assessed with 1,008 bootstrap replicates. The best-fitting amino-acid substitution model and parameter settings were chosen based on
ProTest. The created tree files were uploaded in FigTree v1.4.0 to produce the tree image.

Consensus Weblogos.

The NLP amino-acid sequences (without SP) were aligned per type and group using ClustalW. The generated alignment file was used in Consurf without a tree file, using the model NLPP_{cc} (PDB:3GNU) for the type 1 NLP. For the type 2 NLP, NLPP_{cc} (without SP) was modeled on 3GNU, using I-Tasser. For type 3 NLP, Afu5g02100 (without its 20-amino acid SP) was used as a model without PDB file, because Afu5g02100 is too divergent from 3GNU and does not have any related published structural homologs. For types 1 and 2 NLP, all positions scoring a conservation of 8 or 9 in Consurf were selected. Positions scoring below 8 or 9 in the selected type and group yet scoring 8 or 9 in the other group were also selected for the Weblog, to show this homologous position is less conserved. For the type 3 NLP, all positions scoring 7, 8, or 9 in Consurf were used, as fewer sequences are available. Weblogos were created with Weblogo. Colors are set at KHR: green, DE: blue, ST: red, VLAGI: yellow (RGB:FFFF00), NQ: purple, and C: orange. Alignment gaps of the Weblogos were manually corrected using structural information from NLPP_{cc} and the modeled NLPP_{oc}.

Constructs.

Constructs were made by cloning the original genes (PsojNIP, NLPP_{cc}) into pENTRY/D-TOPO vector using Gateway cloning (Invitrogen) and were subcloned into pB7WG2 (Karimi et al. 2002) or pFAST (Shimada et al. 2010), which both have a 35S promoter to express the gene in plants. All modifications (point mutations, removal of SP, adding KDEL and KDDL) were done according to the Phusion site-directed mutagenesis protocol (Finnzymes). All primers used are listed in Supplementary Table S2.

Agroinfiltrations in tobacco.

Five- to six-week old tobacco plants (Nicotiana tabacum cv. Samsun) were grown at short day conditions (10-h light, 14-h darkness) at 21°C and 70% humidity. Agrobacterium tumefaciens (C58C1) (2 ml) was grown overnight with the appropriate antibiotics, was spun down (1 min at full speed), and was resuspended in infiltration medium (10 mM MgCl$_2$, 0.1 mM morpholineethanesulfonic acid, pH 5.7) to an optical density at 600 nm of 0.8. Acetosyringone was added (100 μM final concentration) and bacteria were incubated for 3 to 4 h at room temperature. Infiltration in tobacco leaves was performed by puncturing a small hole in the leaf and infiltrating the bacterial suspension with a needleless 1-ml syringe. The infiltrated areas were scored for necrosis 10 consecutive days after infiltration.

ACKNOWLEDGMENTS

Thanks to M. F. Seidl for assisting in phylogenetic analyses, M. Pirihonen for providing us with the NLP$_{oc}$ gene, and P. Santhanan, B. Thomma, and C. Wang for sharing expression data.

LITERATURE CITED


AUTHOR-RECOMMENDED INTERNET RESOURCES

CaustalW software: www.ebi.ac.uk/Tools/msa/caustalw2
Consurf server: consurf.tau.ac.il
Endophyte Epichloë festucae genome project database: www.endophyte.uky.edu/elF
FugigDB: fugigdb.org
I-Tasser server: zhanglab.ccb.med.umich.edu/I-TASSER
Joint Genome Institute database: genome.jgi-psf.org
National Center for Biotechnology Information BLAST database: blast.ncbi.nlm.nih.gov/BLAST.cgi
SignalP4.1 server: www.cbs.dtu.dk/services/SignalP
Weblog software: weblog.berkeley.edu