DNA Research, 2018, 25(3), 245–256
doi: 10.1093/dnares/dsx053
Advance Access Publication Date: 5 January 2018
Full Paper

OXFORD

Full Paper

Comparative genomic analyses reveal the features for adaptation to nematodes in fungi

Ruizhen Wang[1,2], Leiming Dong[3], Ran He[1,2], Qinghua Wang[1], Yuequ Chen[1,4], Liangjian Qu[1,*], and Yong-An Zhang[1,*]

[1]The Key Laboratory of Forest Protection, State Forestry Administration of China, Research Institute of Forest Ecology, Environment and Protection, Chinese Academy of Forestry, Beijing 100091, China, [2]Institute of Botany, Beijing Botanical Garden, Beijing 100093, China, [3]State Key Laboratory of Tree Genetics and Breeding, Research Institute of Forestry, Chinese Academy of Forestry, Beijing 100091, China, and [4]Forestry Resources Protection Institute, Jilin Provincial Academy of Forestry Sciences, Changchun 130031, China

*To whom correspondence should be addressed. Tel. +86-10-62888719. Fax. +86-10-62884972. Email: qulj2001@caf.ac.cn (L.Q.); zhangyab@caf.ac.cn (Y.-A.Z.)

Edited by Prof. Takashi Ito

Received 22 May 2017; Editorial decision 3 December 2017; Accepted 6 December 2017

Abstract

Nematophagous (NP) fungi are ecologically important components of the soil microbiome in natural ecosystems. *Esteya vermicola* (*Ev*) has been reported as a NP fungus with a poorly understood evolutionary history and mechanism of adaptation to parasitism. Furthermore, NP fungal genomic basis of lifestyle was still unclear. We sequenced and annotated the *Ev* genome (34.2 Mbp) and integrated genetic makeup and evolution of pathogenic genes to investigate NP fungi. The results revealed that NP fungi had some abundant pathogenic genes corresponding to their niche. A number of gene families involved in pathogenicity were expanded, and some pathogenic orthologous genes underwent positive selection. NP fungi with diverse morphological features exhibit similarities of evolutionary convergence in attacking nematodes, but their genetic makeup and microscopic mechanism are different. Endoparasitic NP fungi showed similarity in large number of transporters and secondary metabolite coding genes. Noteworthy, expanded families of transporters and endo-beta-glucanase implied great genetic potential of *Ev* in quickly perturbing nematode metabolism and parasitic behavior. These results facilitate our understanding of NP fungal genomic features for adaptation to nematodes and lay a solid theoretical foundation for further research and application.

Key words: nematophagous fungi, *Esteya vermicola*, comparative genome, genome feature, adaptation and evolution

1. Introduction

The pine wood nematode is responsible for an epidemic of pine wilt disease, which causes severe ecological and economic losses in Asia and Europe and more recently spreading to Mexico.[1,2] *Esteya vermicola* (*Ev*) is the first recorded endoparasitic nematophagous (NP) fungus that attacks the pine wilt nematode (PWN; *Bursaphelenchus xylophilus*), which exhibits high infectivity.[3] Certain volatile organic compounds (VOCs), α-pinene, β-pinene, and camphor for PWN attraction were produced by *Ev*.[4] Then lunate adhesive spores produced by *Ev*, whose concave surface has large surface area and thick adhesive layer, adhered to the cuticle of attracted nematode, and penetrated the cuticle of PWN. Afterwards, *Ev* colonized in the body cavity of the

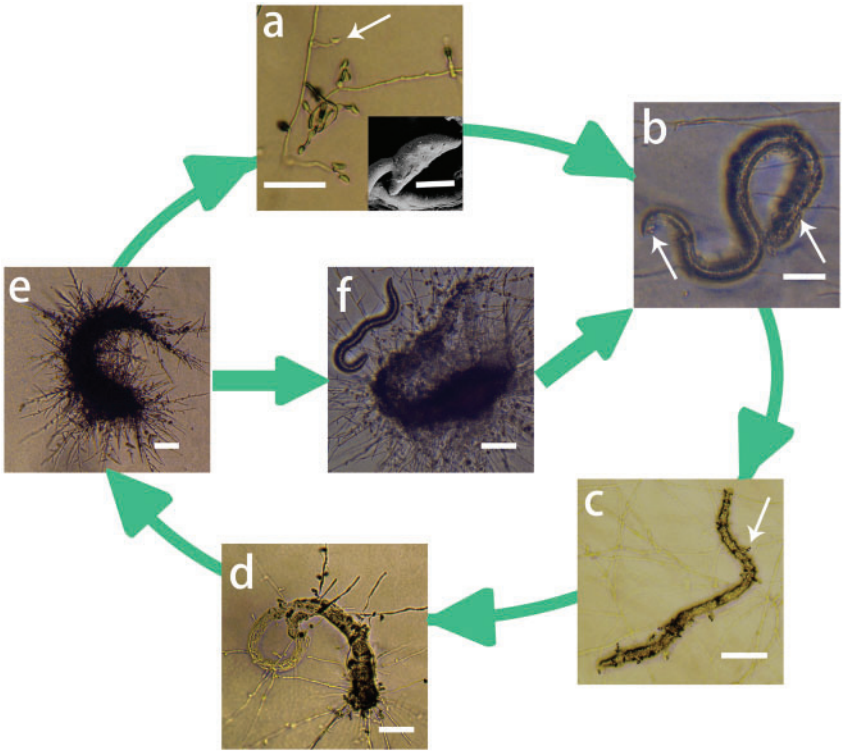


Figure 1. Life cycle of the *Ev*. (a) hyphae and lunate spore (arrow); (b) nematodes were attracted and adhered by lunate spores (arrow); (c) absorption of nutrition from nematode corpse, and producing new lunate spores (arrow); (d) emergence of new hyphae from the nematode; (e), mass reproduction of hyphae and spores on the nematode; f, attracting and infecting nematodes again. The whole infection process about 48 h. Scale bar, 50 μm; inset in a, 5 μm. The infected nematode shown is *B. xylophilus*.

nematode and produced a great deal of mycelia and lunate spores on cadaver of nematode to attract new nematodes for next infection cycle[5] (Fig. 1). Once lunate spore adheres to nematode cuticle, spores fall-off from the conidiophores and move with nematode. Ninety percent of inoculated PWN are infected within 24 h.[3,6] The high infectivity suggests that *Ev* has great potential as biocontrol agent against PWN, but the evolution and mechanisms underlying infection are not clear.

NP fungi are important microorganisms that can suppress the populations of plant-parasitic nematodes and maintain ecological balance and can be used as potential biological control factors.[7] The genomes of NP fungi which have been sequenced include *Arthrobotrys oligospora* (*Ao*) with adhesive networks,[8] *Dactylellina haptotyla* (*Dh*; synonyms *Monacrosporium haptotylum*) of knob-forming species,[9] *Drechmeria coniospora* (*Dc*),[10,11] *Pochonia chlamydosporia* (*Pc*)[12,13] and *Hirsutella minnesotensis* (*Hm*) of endoparasitic fungus,[14] and *Drechslerella stenobrocha* (*Ds*) with constricting ring.[15] Despite the important roles of NP fungi in ecology, comprehensive and detailed analysis of similarity and difference based on the multiple genomic level was still lacking. Due to very different hosts categories and lifestyles of insect pathogenic (IP), plant pathogenic (PP) and NP fungi, it has been expected that their capacity to utilize the host's nutrition and to degrade carbohydrates and proteins would also differ, and that might be reflected in their genetic makeup and gene copy numbers at first. The availability of multiple whole genome data of NP, IP, and PP fungi has provided good opportunity to employ genome-wide surveys to investigate how natural selection shaped the evolution of NP fungi towards nematode host.

The study aimed to decipher the putative genetic components involved in pathogenesis and evolution and to identify the molecular underpinnings underlying the establishment of predation relationship in NP fungi by performing comparative genomic and evolutionary analysis. Here, we reported the complete genome sequence of *Ev* based on the PacBio sequencing technology, comprehensively analyzed the differences among NP, PP, and IP fungi, and uncovered the genetic and evolutionary bases of NP fungi (*Ao*, *Dh*, *Dc*, *Hm*, *Pc* and *Ev*) by the numbers of pathogenic genes, secreted protein, secondary metabolism encoding genes, expanded gene families, and genes undergoing positive selection.

2. Materials and methods

2.1. Fungal strain and DNA preparation

Ev (CBS115803) was purchased from CBS and maintained on cornmeal agar. This fungus was originally isolated from *Scolytus intricatus* and its galleries in oak trees in Czech Republic.[16] Mycelium was grown in liquid medium (potato dextrose broth) and incubated at 25 °C using a shaker at 150 rpm for 7 days. The mycelium (2g) was harvested by aseptic filtering and then grounded in liquid nitrogen. Genomic DNA was extracted using Qiagen Genomic-tip Kit 500 G (Cat No./ID: 10262, Germany) according to the manufacturer's instructions. DNA quality was assessed by spectrophotometry and gel electrophoresis before library construction.

2.2. SMRT PacBio sequencing and HGAP genome assembly

Twenty-five micrograms of high-molecular-weight *Ev* gDNA was used for the PacBio libraries constructing. A 20-kb insert SMRTbell

library was generated using a 15 kb lower-end size selection protocol on the BluePippin (Sage Science). The genome was sequenced using 4 SMRT Cells and P6-C4 chemistry on the PacBio RS II platform (Pacific Biosciences). The *Ev* genome was assembled using the HGAP (version 2.3.0).[17] First, the genome was preassembled with improved consensus accuracy. Then it was assembled through overlap consensus accuracy using Celera assembler and polished with Quiver to improve the site-specific consensus accuracy of the assembly. This Whole Genome project has been deposited at SRA database in NCBI under the accession SRP097009 (BioProject: SUB2318437; BioSample: SUB2318464).

2.3. Gene prediction and functional annotation

Gene annotation was conducted by combining *de novo* and homology-based methods. Augustus[18] and GeneMark-HMM[19] were employed to *de novo* predict gene structures in the *Ev* genome, respectively. For homology-based prediction, protein sequences from three species [*Grosmannia clavigera* (*Gc*), *Ophiostoma piceae* (*Op*), *Sporothrix schenckii* (*Ssch*)] were initially mapped onto the *Ev* genome using tBlastn (E-value $\leq 1e^{-5}$). The homologous genome sequences were aligned against the matching protein using GeneWise[20] for accurate spliced alignments. All the predictions were combined by EVidenceModeler (EVM)[21] to produce a consensus gene sets. Gene functions were assigned according to the best matches derived from the alignments to proteins annotated in KOG, NR, SwissProt, and TrEMBL databases using Blastp[22] (E-value $\leq 1e^{-5}$). Then the pathway in which the gene might be involved was annotated by KAAS[23] according the KEGG database. Motifs and domains were annotated using InterProScan[24] by searching against GO databases. Finally, the result annotated from the KOG, GO, KEGG, NR, Swissprot, and TrEMBL databases were combined to obtain the final annotation of *Ev* genome.

2.4. Protein family classifications and prediction of pathogenicity-related genes

The carbohydrate active enzymes (CAZy) genes were identified using Blastp according to the CAZy database.[25] Whole genome protein families were classified by Pfam analysis.[26] The families of proteases were identified by Blastp searching against the MEROPS peptidase database release 9.4.[27] Transporters were retrieved from the Transporter Classification Database (TCDB).[28] Putative virulence factors were identified by searching against the pathogen–host interaction database (PHI),[29] database of virulence factors in fungal pathogens (DFVF),[30] and virulence factors of bacterial pathogens (VFDB). Peroxibase-encoding genes were predicted searching against Peroxidases database. Antibiotic resistance genes were identified according to Antibiotic Resistance Genes Database (ARDB).[31] An E-value cut-off of 1.0E-5 was adopted to filter the BLAST results of the above databases. Twenty-one strains of fungi were used for the above BLAST analysis (Supplementary Table S1). SignalP 4.1 and TargetP were used to predict potential secreted proteins with default parameters. Secondary metabolite genes of six NP fungi, six IP, and six PP fungi were predicted by SMURF (Supplementary Table S1). The antiSMASH pipeline with HMM signatures was used to identify and annotate putative polyketide synthase (PKS), nonribosomal peptide synthetase (NRPS), and terpene synthase (TPS) genes and other gene clusters about secondary metabolites and to predict the PKS and NRPS domain architecture. The whole genome dataset was subjected to analysis with default settings.[32] The correlations between these 18 fungal species (IP, PP, and NP) and gene variables (PHI, TCDB, DFVF, and VFDB) were evaluated by correspondence analysis. All statistical analyses were performed with R project V.3.4.2.[33]

2.5. Phylogeny construction and estimation of divergence time

Twenty-three fungi species [*Ev*, *Dc*, *Ao*, *Dh*, *Hm*, *Pc*, *Gc*, *Op*, *Ss*, *Sporothrix brasiliensis* (*Sb*), *Fusarium graminearum* (*Fg*), *Magnaporthe oryzae* (*Mo*), *Colletotrichum higginsianum* (*Ch*), *Verticillium dahliae* (*Vd*), *Zymoseptoria tritici* (*Zt*), *Cordyceps militaris* (*Cm*), *Beauveria bassiana* (*Bb*), *Metarhizium anisopliae* (*Ma*), *Metarhizium robertsii* (*Mr*), *Moelleriella libera* (*Ml*), *Sporothrix insectorum* (*Si*), *Neurospora crassa* (*Nc*), *Saccharomyces cerevisiae* (*Sc*)] were used for gene family clustering and phylogeny tree construction. Blastp was used to generate the pairwise protein sequence with similarity (E-value $\leq 1e^{-5}$). OrthoMCL[34] was used to cluster similar genes by setting main inflation value 1.5 and other default parameters; 1203 single-copy gene families were extracted for the phylogenomic analysis. The protein sequences were aligned by Mafft[35] and back translated into CDS alignments. Then, we eliminated poorly aligned positions and divergent regions of alignment of CDS sequences using Gblocks.[36] Four-fold degenerate sites of all these single-copy genes in each species were extracted and concatenated them to be one supergene for phylogeny construction. Phylogenetic analyses using maximum likelihood (ML) method were conducted in RaxML version 8.2.4[37] and set *Sc* as out-group. The divergence time was estimated based on established phylogeny tree. Markov chain Monte Carlo algorithm for Bayes estimation was adopted to estimate the neutral evolutionary rate and species divergence time using the program MCMCTree of the PAML package.[38] The calibration times for divergence between *Sc* and *Dh* (460-726My), *Dh* and *Zt* (358-468Mya), *Pc* and *Ch* (194-888Mya) were obtained from the TimeTree database.[39]

2.6. Expansion and contraction of gene families

Gene families that have undergone expansions or contractions were identified using CAFE version 3.0[40] program with default parameters. The algorithm CAFE takes a matrix of gene family sizes in extant species as input and uses a probabilistic graphical model to ascertain the rate and direction of changes in gene family size across the given phylogenetic tree. The tree and expansion/contraction data were displayed using the iTOL web tool.[41]

2.7. Identification of orthologous genes and positively selected genes (PSGs)

Orthologous genes in 23 species were identified. Genes with CDS length less than 450 bp were filtered. All-versus-all Blastp comparisons were used to identify clusters of homologous genes. For avoiding noise information in the Blastp output, orthologous clusters that contained more than 10 species were retained. The values of Ks and Ka substitution rates and the Ka/Ks ratio were estimated using the Codeml program in the PAML package.[38] Codeml was used to calculate the selection pressure in the phylogenetic tree ($P \leq 0.05$).

3. Results

3.1. General features of *Ev* genome

Ev genome was sequenced at approximately 100 × coverage, generating a total of 4.44 Gbp sequences. *De novo* assembly yielded *Ev* genome of 34.2 Mb in length with average 68.59 × coverage depth.

The N50 sizes of 50 contigs and 42 scaffolds were 4.26 Mbp and 4.41 Mbp, respectively. Combined *de novo* and homology prediction produced 8,427 protein-coding genes. The detailed annotation results are given in the supporting information.

3.2. Comparative analysis of gene involved in pathogenicity and virulence

To identify functional gene categories characteristic of adaptation in killing nematode, we compared 21 genomes, including 6 IP fungi, 6 PP fungi, and 6 NP fungi and 3 genetically close species with *Ev*, *Gc*, and *Si* (Supplementary Table S1). The genome integrity and completeness of gene prediction assessed by BUSCO are shown in Supplementary Table S2. The genome integrities of the selected species were more than 95%, and the completeness of protein prediction was greater than 92%. The corresponding datasets based on different databases were shown in supporting information (Supplementary Table S3–S11). Extracellular adhesive proteins comprise GLEYA proteins, CFEM proteins, WSC-containing proteins and lectin, each of which was averagely the highest in NP genomes except CFEM (Supplementary Fig. S1), as well as total adhesion (Supplementary Fig. S2). *Dh* encoded the highest number of adhesive proteins GLEYA and WSC, whereas *Ao* encoded most abundant lectin among 21 analyzed fungi genomes (Supplementary Fig. S1a). However, our comparative analysis found that endoparasitic fungi (*Ev*, *Dc*, *Pc* and *Hm*) contain fewer extracellular adhesive proteins (Lectin and GLEYA) than adhesive networks producer *Ao* and adhesive knobs producer *Dh* (Supplementary Figs S1 and S2). It was found that subtilases/peptidase S8 and subtilisin-like participating in lethal activity and the infection process of nematodes[42–44] were averagely abundant in NP (Supplementary Fig. S3 and Supplementary Table S6). More subtilisin S8 and subtilisin-like were found in *Ao* and *Dh* than that in other NP fungi (Supplementary Table S6). NP fungi encoded averagely richer ankyrin repeat protein, while PP fungi encoded more leucine-rich repeat, tyrosinase, xylanase (Supplementary Fig. S4). IP fungi encoded the least cellulose, compared with PP, NP and CP fungi (Supplementary Fig. S6). Interestingly, all analyzed fungi except *Z.tritici* possessed appressorial penetration protein (PHI: 256), homologs of proteins in plant-pathogenic fungi *Mo*, especially in *Ao* (25) and *Dh* (28) most abundant (Supplementary Table S4).

Chitinase is important for IP and egg-parasitic fungi to infect insect and nematode eggshell.[45,46] IP fungi were most abundant averagely in enzymes catalyzing the decomposition of chitin (Supplementary Fig. S7). *Pc*, an egg-parasitic NP fungus, encoded the highest number of chitinase (159, CAZy) among NP fungi (Supplementary Fig. S7). Chitosanases (including GH5, GH7, GH8, GH46, GH75 and GH80) are another enzymes involved in chitin degradation and nematode parasitism.[13] The analysis suggested that *Ev* had the most chitosanases, especially GH7; *Ao* encoded the richest GH75 (Supplementary Table S3 and Supplementary Fig. S7). IP, NP, and PP fungi also contained averagely more peptidase than CP fungi (Supplementary Fig. S8). IP fungi possessed most pkinase and enterotoxin-a, furthermore, and both pkinase and enterotoxin-a in IP fungi was significantly higher than that in PP fungi (Supplementary Figs S9 and S11). Based on Pfam database, IP had significantly more toxin genes than PP and CP fungi (Supplementary Fig. S10). The number of aspartic peptidase in NP and IP fungi was averagely higher than PP fungi (Supplementary Fig. S5). PP fungi genome encoded the richest CAZy enzymes (Supplementary Table S3), peroxidase (Supplementary Table S9), in particular, hemicellulose (Supplementary Fig. S12), G-protein receptor (Supplementary Fig. S12), pectate lyase (Supplementary Fig. S13) and cutinase (Supplementary Fig. S14), P450 (Supplementary Fig. S15). The important claim here was that the number of hemicellulose, G-protein receptor, and pectate lyase in PP fungi was significantly higher than that in IP fungi (Supplementary Figs S12b and S13b). Moreover, the content of cutinase in PP group was significantly greater than that in IP and CP group (Supplementary Fig. S14b). It suggested that PP fungi evolved to adapt to higher carbohydrate content in plants.

We identified 231 transporter superfamily in all 21 fungi species. *Ev*, *Hm* and *Pc* genome encodes a large number of transporters (1,482, 1,546, and 1,986, respectively; Supplementary Table S10). Remarkably, *Ev* and *Pc* had more carbon (72, 90), iron (23, 19) and vitamin transporters (81, 55) than do other NP fungi (Fig. 2 and Supplementary Table S10). The numbers of genes encoding allantoate transporters, and vitamin or iron in *Ev* were the greatest and second highest among all 21 fungi, respectively (Fig. 2). Ammonium and urea import capacity of *Ev* was higher than other NP, IP, and PP fungi. Therefore, *Ev* shows an increased genetic potential in terms of vitamin, iron, and nitrogen uptake. This could be an important mechanism for perturbing nematode metabolism and ultimately causing nematode death. It indicated that the transport capacity of ammonium, nucleobase, and nucleoside in CP group was very strong (Supplementary Fig. S16). The number of oligopeptide transporter in IP group was averagely higher than other groups and significantly higher than that of in NP group (Supplementary Fig. S16). This implied the great potential of absorbing oligopeptide of nitrogen sources in IP fungi.

Strikingly, all 21 fungi contained a wide array of antibiotic resistance genes. The total number of antibiotic resistance genes in IP fungi (207) was the richest, followed by PP fungi (201), and NP fungi (174) (Supplementary Table S8 and Supplementary Fig. S17). The substances encoded by these genes were likely to play roles in competition with bacteria in insect, nematode, and plant, then pathogenic fungi can colonize in their host. Among antibiotic synthesizing genes, macrolide transporter-encoding genes and multidrug-encoding genes were most abundant in each group (Supplementary Table S8 and Supplementary Fig. S17).

3.3. Secreted proteins and small secreted proteins

Secretion signals were predicted in 21 fungi. The results showed that PP had the highest number of secreted proteins (SPs) and small secreted proteins (SSPs; < 300 amino acids), followed by NP and IP fungi (Supplementary Table S12). Among the 21 analyzed fungi, *Mo* had amazing ability to produce SSP (about one-tenth of the total protein). Among NP fungi, *Pc* had the most SPs and SSPs, followed by *Dh*, *Ao* and *Hm* (Supplementary Table S12). Out of the 1,318 proteins predicted to be secreted by *Ev*, 46.7% can be ascribed known function. Of these 397 SSPs of *Ev*, 302 was annotated into NR database, but only 131 (33.0%) were annotated with known function; 10 of 397 SSPs was related with transporters; 48 of 397 SSPs (12.1%) was related with pathogen-host interaction based on PHI database; and 180 of 397 SSPs (45.3%) was found listed in Pfam database.

The possible pathogenic-related SPs are shown in Fig. 3. NP fungi were rich in SPs of subtilisin or subtilisin-like protease, lectin, WSC, and iron transporter. Only *Ev* was predicted to produce secretory cellobiose dehydrogenase in NP fungi. *Ev* was remarkable in SPs of putative aspartic protease, glucanase and endoglucanase, laccase and cellobiose dehydrogenase. It was worth noting that the number of secreted oligopeptide transporter and aspartic proteinase was remarkably high in endoparasitic fungi *Ev*, *Dc*, *Pc*, and *Hm*. The important

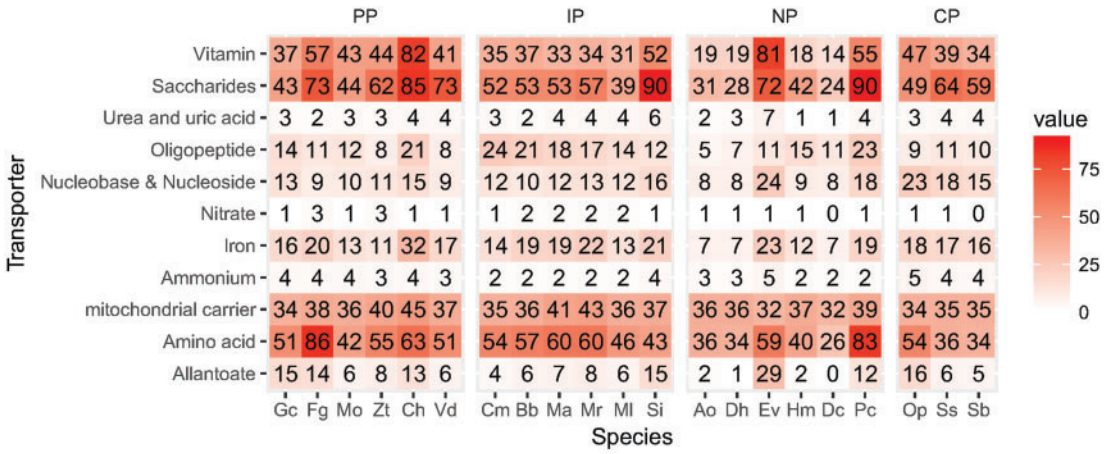


Figure 2. Heatmap of vitamin, iron, saccharides and nitrogen transporter-encoded genes in 21 fungi.

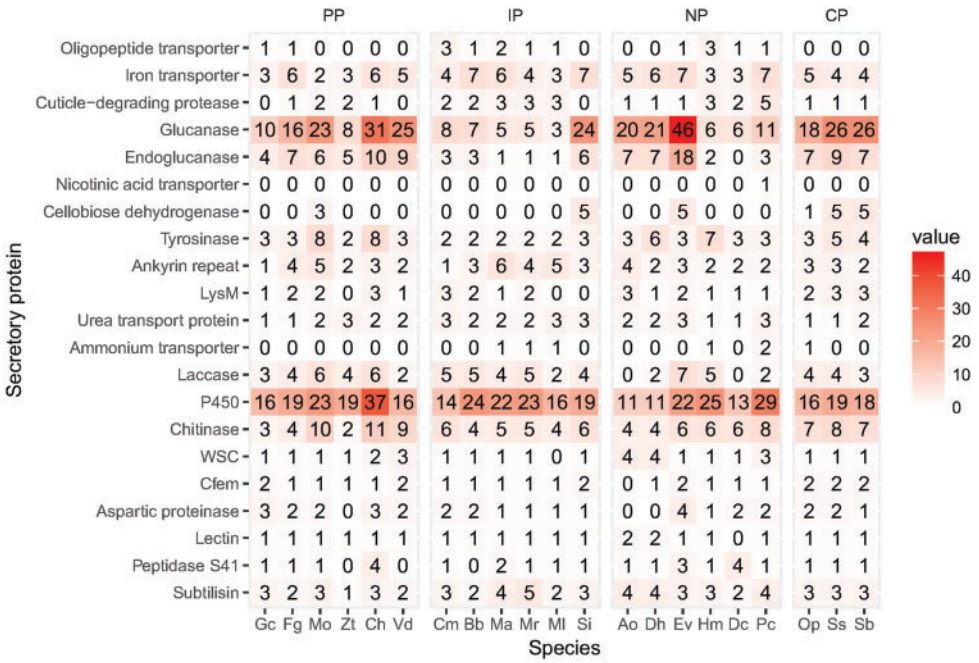


Figure 3. Heatmap of possible pathogenic related SPs in 21 fungi.

SPs involving pathogenicity in *Ev* were comparable to those in *Mo* and *Ch*, such as P450, laccase, and cellobiose dehydrogenase.

3.4. Secondary metabolite analysis

PP fungi can produce diverse secondary metabolites (SMs) that aid in pathogenicity.[47–49] However, the extent and distribution of the capacity for secondary metabolite synthesis have not been investigated in NP group. Furthermove, there was no comparative analysis about the capacity for secondary metabolite among groups (IP, NP, and PP) before. SMURF predicted that endoparasitic fungi (*Hm*, *Pc*, *Dc*, and *Ev*) had more genes encoding secondary metabolites than species of adhesive networks and knob-forming species (*Ao* and *Dh*) (Supplementary Table S13). *Hm* had the highest number of gene clusters and core genes involved in biosynthesis of secondary metabolites, followed by *Pc* among NP group. PKS and NRPS synthesis genes in *Pc* (195, 156) and *Hm* (299, 238) were more abundant than that of other analyzed NP fungi. Especially endoparasitic NP fungi had tremendous potential for biosynthesis of secondary metabolites that act as toxins or signals in the interactions between fungus and host. The rich unknown secondary metabolism in NP fungi is worthy of further research and exploration. Furethermore, *Mr* and *Ma* of IP fungi, and *Ch* and *Mo* of PP fungi also had great capacity for biosynthesis of secondary metabolites, which possibly greatly enhanced their disease-causing abilities.

3.5. Insights into pathogenic differentiation leading to different hosts categories

To assess pathogenic differentiation leading to different host classification, we performed correspondence analysis based on multiple large datasets (Supplementary Tables S3–S11). It was hypothesized

that pathogenic fungi belonging to the same host categories with similar nutritional mode, would be clustered together. Figure 4 shows separation result of three clusters (PP, IP, and NP) according to gene variables of PHI genes, membrane transport proteins (TCDB), fungal virulence factors (DFVF), virulence factors of bacterial pathogens (VFDB). The species of IP fungi were clustered except *Si* species, and so did the six PP species. However, the distribution of NP species was dispersive, and *Ev*, *Dc* and *Pc* were always clustered with PP, IP fungi, respectively. CAZy, peroxidises, and antibiotic resistance genes failed to separate IP and PP fungi (figures not shown). Also, the analysis suggestsed that the difference in pathogenicity genes or membrane transport proteins was possibly the important factors to drive pathogenic differentiation and adaptation to different host categories. Endoparasitic fungal *Ev*, *Dc*, *Pc* and *Hm*, which were very different from *Ao* and *Dh*, were likely to be the intermediate type between the IP and PP fungi.

3.6. Phylogeny and divergence time estimates

OrthoMCL was used to identify gene orthologs. Totally 253,363 genes from 23 species were used for gene family clustering analysis. Finally, 20,524 gene families containing 206,300 total genes from 23 species were generated, out of which 1,203 possessed a strict single-copy orthologous relationship. The phylogenomic tree was constructed using concatenated nucleotide sequences of 1,203 single-copy orthologous genes to infer NP fungal phylogenetic relationships in the context of other fungal taxa. It indicated that *Ev* diverged from plant-associated *Gc* about 135 million years ago (Mya), and *Ev* and *Gc* had closer genetic relationship (Fig. 5). The phylogenetic analysis showed that the relationship of *Ao* and *Dh* in NP fungi were closely related, and the divergence of the two species took place around 212 Mya. *Dc*, *Pc* and *Hm* had a close genetic relationship with IP fungal *Metarhizium*, which was consistent with previous phylogenetic analysis results.[10–14]

3.7. Gene family and gene family expansion

Gene duplication and protein family expansion are important genomic mechanisms that shape the evolution of pathogenic fungi.[50] CAFE analysis was performed to assess the gene family variations; 276 gene families were expanded in *Ev*, 1,170 families had undergone contraction (Fig. 6). We found a marked expansion of transporter genes in *Ev*. Among the 276 expanded gene families in *Ev*, 10.7% (37) was related to transporters or permeases, which was the highest among all the fungi analyzed (Supplementary Table S14), while the proportions were less than 2.4% and 6.8% in other five NP fungi and other analyzed species, respectively (Supplementary Table S14). It was surprising that the proportion of expanded gene families of transporters in *Ev* was much higher than that of the other analyzed species. The transporters of expanded gene families in *Ev* occurred mainly for the carbon source, nitrogen source, iron and vitamin transport, such as ABC transporter, sugar transporter, iron permease, urea transporter, allantoate/allantoale permease, vitamin H transporter and amino acid permease (Supplementary Table S15), which provide the basis for effectively absorbing nutrition of nematodes. Gene families involving in aflatoxin efflux pump had undergone expansion in *Ev* (Supplementary Table S16). Gene families about endo-beta-glucanase, which catalyse hydrolyzing O-glycosyl compounds, were substantially expanded in *Ev* compared with other fungi species. Important pathogenicity-related gene families were

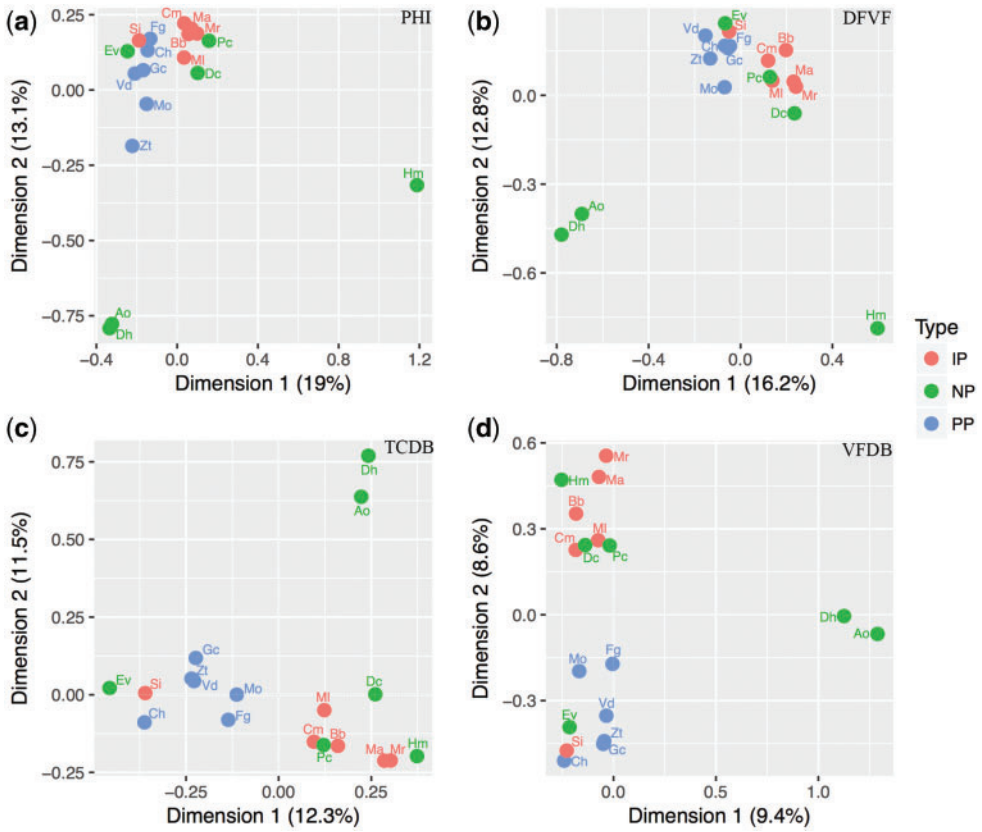


Figure 4. Correspondence analysis grouping pathogenic fungi by host class based on the set of predicted (a) PHI, (b) DFVF, (c) TCDB, (d) VFDB.

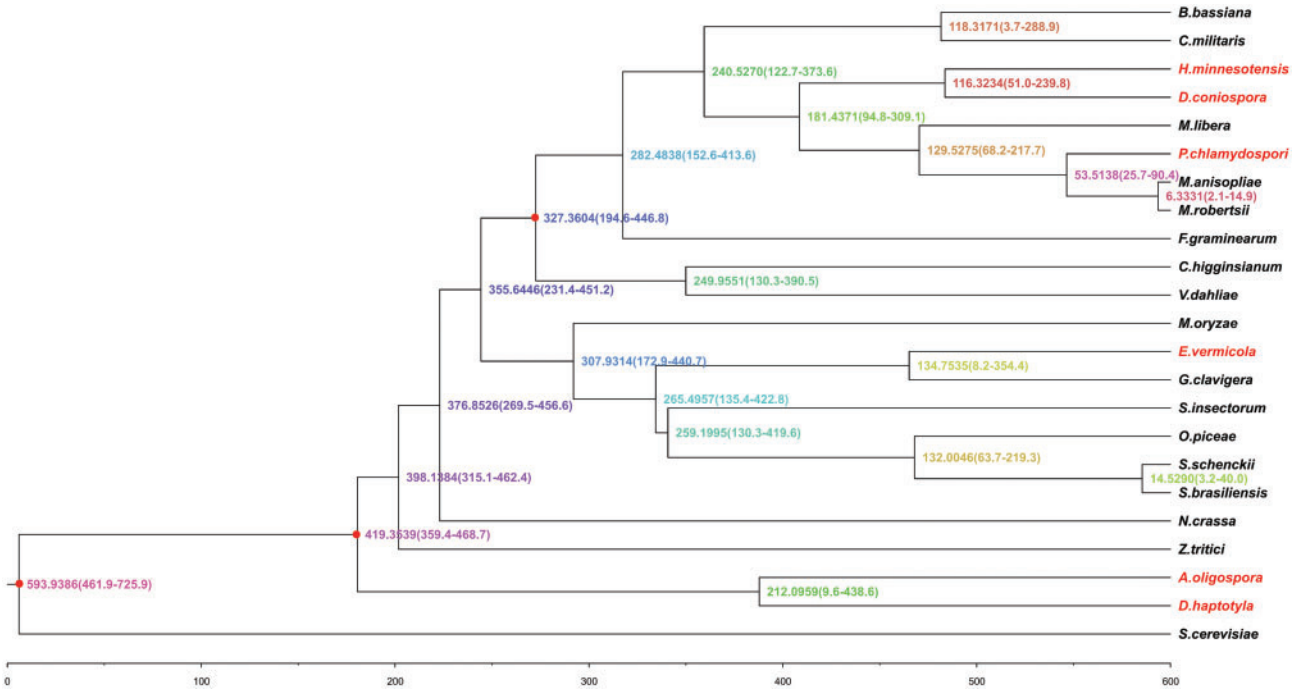


Figure 5. Phylogenetic tree and divergence time estimates. The numbers at nodes in the tree are the divergence times from the present (Million years ago, Mya). The nodes with red dot represent the reliable calibration time of the divergence. Bootstrap values in each node are 100 (not shown in the diagram). Six NP fungi *H. minnesotensis, D. coniospora, P. chlamydosporia, E. vermicola, A. oligospora, and D. haptotyla*

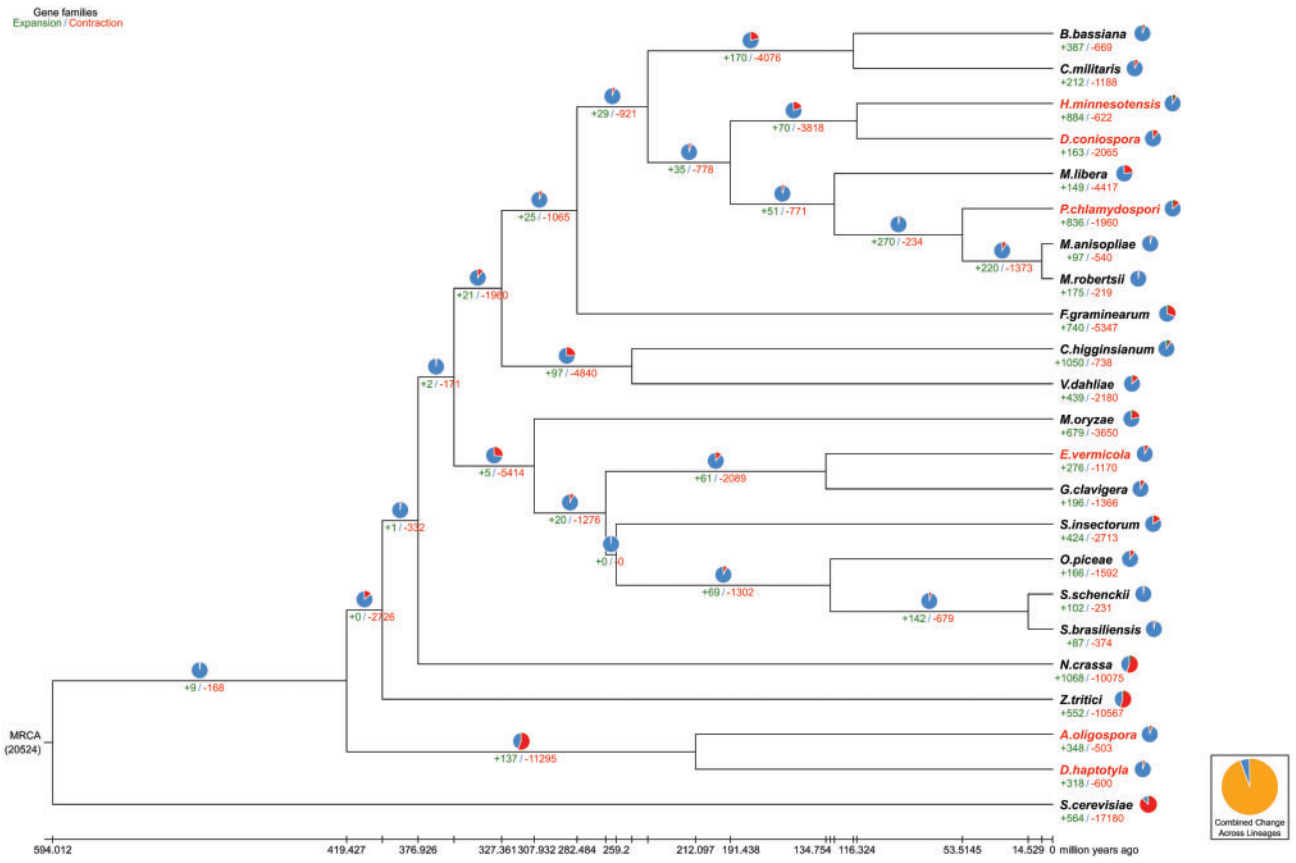


Figure 6. Gene family expansions and contractions. Branch length represents differentiation time. Numbers for expanded (green) and contracted (red) gene families are shown below branches or taxon names with percentages indicated by pie charts. Blue represents the proportion of gene families no changed in the current species or branch. Orange in pie charts represents the proportion of the gene families combined change across lineages. Six NP fungi are marked with red.

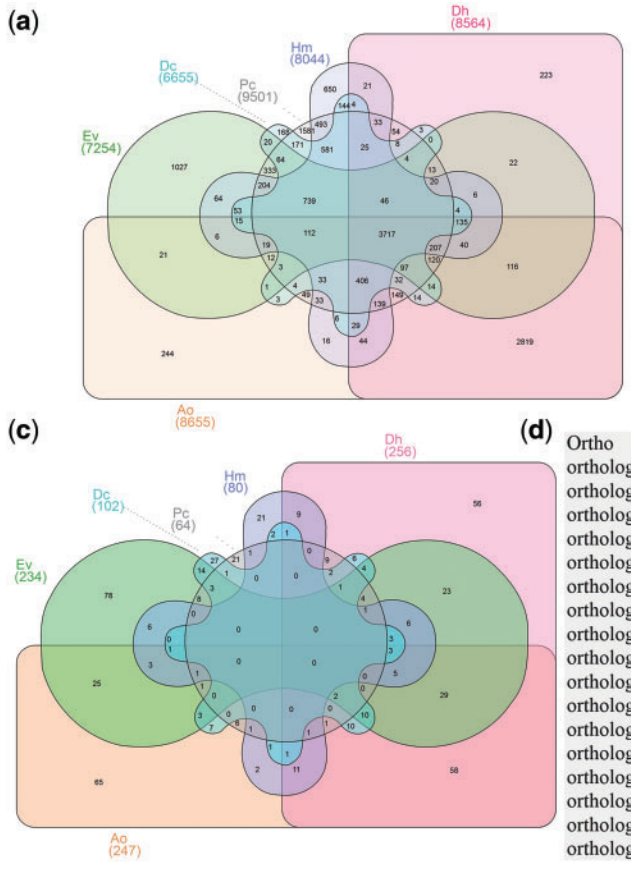
founded expanded, including peptidase S41, cellobiohydrolase and laccase.

The six NP fungi in total shared 3,717 gene families, but no expansion gene families in common were observed (Fig. 7a and b). Surprisingly, we found that *Ao* and *Dh* shared the largest set of gene families (8,078) through pair comparison, followed by *Pc* and *Hm* with 6,807 gene families. Beyond that, the two species (*Pc* and *Hm*) with similar nematicidal mechanism (endoparasitism) shared the most expanded gene families (132). Substantial differences were observed in species-specific gene families among six fungi (*Ev* 1,027, *Hm* 650, *Ao* 244, *Dh* 223, *Dc* 168, *Pc* 1,581) in magnitude. The ratio of species-specific gene family to the total gene family was the highest in *Pc* (16.6%), followed by *Ev* (14.2%) and *Hm* 8.1% (Supplementary Table S17). The ratio of species-specific expanded gene family to the total expanded gene family was 71–86.8% among NP fungi (Supplementary Table S17). NP fungi contained more gene families not found in other pathogenic fungi (Supplementary Fig. S18) indicating their unknown functional diversification.

We further compared expanded gene families in 23 species to detect gene family evolution (Supplementary Table S16). Subtilisin-like gene family in NP fungi (*Dh* and *Hm*) and IP fungi (*Mr*) were expanded, except that in PP species. Genes encoding chitinase, aspartic proteinase and Wsc domain protein had undergone expansion in NP and PP group, but not in IP group. Furthermore, peptidase S41 family protein exhibited expansion in NP fungi (*Ev* and *Dc*) and IP fungi (*Ma*). Intriguingly, gene families about trehalose metabolism expanded greatly in *Hm*. Sabine Fillinger et al. found trehalose rapidly degraded upon induction of conidial germination and was required for the acquisition of tolerance to a variety of stresses in *Aspergillus nidulans*.[51] Compared with other analyzed species, there was a major expansion in gene families related to allantoate and iron transport, laccase and endoglucanase in *Ev* (Supplementary Table S16). The ratio of expansion gene families to total gene families in *Hm* and *Pc* peaked at an incredible 11.0% and 8.8%, respectively, while that in other four NP fungi was no more than 4% (Supplementary Table S17), which partly explained the differences in genomic size of these species.

3.8. Positive selection

Positive selection, acting in the evolution of functionally important gene families, is an important driving force for pathogens in adaptation to a pathogenic lifestyle and host specialization.[52–54] Identifying targets of positive selection will help to annotate the genome functionally and to elucidate evolutionary processes. Signatures of positive selection specify functionally important regions of the genome.[55] Based on the established phylogeny and 10,610 high-confidence orthologues identified among 23 species, positively selected genes (PSGs) of six NP fungi and PP fungal *Gc* was detected. *Dh* and *Ao* had more PSGs (256, 247) than endoparasitic NP (*Ev* 234, *Dc* 102, *Hm* 80, *Pc* 64) (Supplementary Table S18); 97 PSGs in *Ev*, this was the highest count among 7 analyzed species, belonged to putative



(d)

Ortho	*Ao*	*Dh*	*Dc*	*Hm*	*Pc*	*Ev*	*Gc*	Swissprot function
ortholog00083	0	0	0	0	0	1	0	Nitrogen permease regulator
ortholog00552	1	1	0	1	0	0	0	Subtilisin-like proteinase
ortholog04865	1	1	0	1	0	1	0	aspartic endopeptidase
ortholog04778	1	1	1	0	0	1	0	Aminopeptidase
ortholog06321	0	1	1	1	0	1	0	Probable dipeptidyl-aminopeptidase
ortholog08256	0	1	0	0	0	1	0	Probable Xaa-Pro aminopeptidas
ortholog03864	1	1	0	0	0	1	0	zinc metalloprotease
ortholog00176	1	1	0	0	0	0	0	NADPH-cytochrome P450 reductase
ortholog00371	1	1	1	1	0	1	0	Cytochrome P450
ortholog01184	1	0	0	0	0	0	1	cytochrome c oxidase assembly protein
ortholog07190	0	0	0	0	0	1	1	NADH-cytochrome b5 reductase
ortholog09764	0	0	0	0	0	1	0	Sporulation protein
ortholog08144	0	1	0	0	0	1	0	Bud site selection protein
ortholog09010	0	0	0	0	0	1	0	Ubiquinone biosynthesis monooxygenase
ortholog08515	1	0	0	0	0	0	0	Dynamin-related protein DNM1
ortholog03712	0	0	0	0	0	1	0	Dynamin-related protein DNM1
ortholog07048	0	0	0	0	1	1	0	Vacuolar protein sorting-associated protein

Figure 7. Venn diagrams and selected orthologous genes undergoing positive selection in six NP fungi. (a) Venn diagram of gene families; (b) Venn diagram of expanded gene families; (c) Venn diagram of orthologous genes of positive selection; (d) Selected orthologous genes of positive selection. '1'means undergoing positive selection; '0'means no positive selection.

PHI genes. Moreover, *Ev* had the largest number of unique positive selection genes (78) (Fig. 7c). The number of PSGs encoding transcription factors were the most highest in *Dh* (27) all over analyzed seven species (Supplementary Table S18). Six NP fungi did not share orthologous gene undergoing positive selection. Corresponding to the need for nitrogen, orthologous gene (ortholog00083) involving in nitrogen permease regulator have been subject to a positive selection pressure in *Ev* ($P = 0.02$) to adapt to nematode. Subtilisin-like proteinase (ortholog00552) was also under positive selection in *Ao*, *Dh* and *Hm* (Fig. 7d). Orthologous genes related to aspartic endopeptidase, aminopeptidase, and cytochrome P450 reductase revealed positive selection in NP fungi (Fig. 7d).

4. Discussion

4.1. NP fungi encoded rich secondary metabolites and antibiotics

Secondary metabolites involving in host–pathogen interactions may protect the fungus from predation and allow an organism to survive in its ecological niche.[56] Secondary metabolisms of beauvericin,[57] and bassianolide,[58] and destruxin[59] in insect pathogens were insecticidal. Similarly, secondary metabolites in NP fungi can kill nematodes.[60] Endoparasitic fungi had very rich repertories of secondary metabolism genes compared with nematode trapping fungi implying endoparasitic fungal great potential for killing nematodes. One of the most important questions about the ecology of NP fungi is how they defends the nematode cadaver against different microbial competitors. NP fungi, especially *Pc* and *Hm*, encoded a large number of polyketide and nonribosomal peptide synthases which could be part of the biosynthetic pathway of antibiotics. Besides, the phenomenon that NP fungi possess a set of antibiotic resistance genes is a general mechanism responsible for resistance to antibiotics from other microbe. It was suggested that antibiotic and fungal secondary metabolites displayed antagonistic effect against other microbial competitors during colonizing nematode cadaver.

4.2. The abundance and evolution of pathogenic genes in NP fungi

Adhesive proteins accumulating on the outside surface of adhesive traps or spores of NP fungi play vital roles for those fungi to adhere to nematode cuticles.[7] Adhesive materials may harbour many secreted virulence-related proteins.[61] Comparative analysis showed that NP fungi contained abundant pathogenicity genes encoding subtilisins and subtilisin-like protease, of which the nematotoxic activity has been proved. Transcripts encoding subtilisins were highly expressed by *Ao*, *Dh*, and *Hm* during infection.[9,62] Previous studies have proved that subtilisins are important virulence factors involved in the penetration and digestion of host cuticles in IP and NP fungi.[63–65] We found that tyrosinase-encoding gene was the most abundant present in NP fungi. It has been demonstrated that tyrosinases involving in synthesis of melanin in some fungi were highly upregulated in *M. haptotylum* (*Dh*) during infection[9] and participated in pathogenic process.[66] But these genes encoding these proteins were not evenly distributed across NP species, and the number of these genes in different species varied widely. These wide differences in number of pathogenic genes may be a manifestation of diverse pathogenic mechanism of NP fungi.

Subtilisin-like proteases as virulence factors function during fungal pathogenic progress in IP fungi,[67,68] NP fungi,[64,69–71] and other parasite fungi.[72] Our results demonstrated that subtilisin-like serine proteases as SPs underwent expansion and positive selection in NP group. Aspartic protease, which was expressed by many pathogenic ascomycetes during infection,[9] was subjected to positive selection during the evolution in NP fungi. Nitrogen permease regulator, as well as some proteolytic enzymes, was under forward selection, implying the importance of nitrogen sources for NP fungi. Cytochrome P450s play various roles in secondary metabolism and involve in biodegradation of lignin and various xenobiotic compounds.[73] Our results indicated that PP and IP contained richer P450s than NP group. Nevertheless, P450 and P450 reductase underwent positive selection in NP fungi but not in PP fungal *Gc* (Fig. 7d). Therefore, P450s should be functionally important for NP in adaptation to pathogenic lifestyle.

4.3. Abundance of transporter in *Ev* genome

We identified abundant transporters, which underwent expansion in *Ev*. Monosaccharide transporter (Hxt1), also as sensor, is required for virulence of the maize pathogen *Ustilago maydis* and is important for fungal development during the pathogenic stage of the fungus.[74] Pyruvate uptake transporter (BcMctA) is also required for pathogenicity in *Botrytis cinerea*, and disruption of *BcmctA* significantly reduced the virulence of *B. cinerea* on cucumber and tomato leaves.[75] Phytopathogenic bacteria and fungi can use iron uptake systems to multiply in their hosts and to promote infection,[76] and iron transporters have been found to play a critical role in virulence of both animal- and plant-pathogenic fungi,[77] such as siderophores of *Aspergillus fumigatus* and high affinity iron permease of *Rhizopus oryzae*, which play critical roles in the expression of virulence.[78,79] Similarly, these transporters could well take part in pathogenicity during infecting nematodes. In addition, based on a wealth of transporters (Supplementary Table S10), it was possible that NP fungi, especially endoparasitic fungal *Ev*, *Pc*, and *Hm* could be more efficiently utilize nematode nutrients compared with other NP fungi.

4.4. Endo-beta-glucanase in *Ev* genome was highly expanded

Comparative analysis revealed endo-beta-glucanase encoding genes in *Ev* genome were highly expanded. Endo-beta-glucanase likely plays a role in degrading glucans polymers from themselves cell wall or their host nematodes. Endo-beta-glucanases were identified as parasitism genes in plant-parasitic nematode.[80] Extracellular glucanase secreted by *Acremonium persicinum* was considered to participate in the degradation of an extracellular storage glucan.[81] Glucanases produced by some *Trichoderma* species has been recognized as the key enzymes in the lysis of cell walls during their mycoparasitic action against phytopathogenic fungi.[82,83] Highly expanded endo-beta-glucanase in *Ev* genome possibly functions as parasitism, but further research is needed to determine the possibility.

4.5. No α- and β-pinene monoterpenoid synthase-encoding genes in *Ev* genome

Interestingly, the host pine of PWN, vector *Monochamus alternatus* of PWN, and NP fungal *Ev*, belonging to three different kingdoms, all attract nematodes through volatile chemicals: α- and β-pinene.[4,7,84,85] Due to *Ev*' s capacity to attract PWN by volatile chemicals α- and β-pinene[4] and the fact that *Ev* host intracellular bacteria,[86] it remains vague whether the fungus itself produced the compounds. Annotation results of *Ev* genome revealed that there were no α- and β-pinene monoterpenoid synthase-encoding genes.

Here we hypothesize that the endobacteria produce the monoterpene that attracts PWN. The next step for whole genome sequencing of this endobacteria will further verify the authenticity of the endobacteria producing the monoterpenoid that attracts nematodes.

5. Conclusions

We sequenced the genome of *Ev* and embarked upon a comparative genomics analysis. These results indicated that NP fungi was distinguished from other pathogenic fungi in terms of pathogenic genes sets. Despite genetic predispositions for killing nematode, the microscopic mechanism was different in NP fungi. Our analysis elucidated a repertoire of genes in endoparasitic NP fungi distinct from those of other NP fungi. The availability of the first *Ev* genome sequencing, annotation and evolution analysis was an important advance in helping to decipher its pathogenesis. The results would facilitate our comprehension of the genomic features of NP fungi for adaptation to nematodes. Furthermore, the results also set out a solid theoretical foundation for further research and application of these NP fungi as biological control agents.

Acknowledgements

The authors would like to thank Chengshu Wang (Shanghai Institutes for Biological Sciences, CAS) and reviewers for suggestion for the manuscript modification. This work was supported by the Fundamental Research Funds for the Central Non-profit Research Institute of CAF (No. CAFYBB2017SZ003) of China, Institute Special Fund for Basic Research, Institute of Forest Ecology, Environment, and Protection, Chinese Academy of Forestry (Grant no. CAFRIFEEP201403) and State 863 Project funded by Ministry of Science and Technology of the People's Republic of China (Grant no. 2012AA101503).

Accession number

SRP097009

Conflict of interest

The authors declare no conflict of interest.

Supplementary data

Supplementary data are available at *DNARES* online.

References

1. Futai, K. 2013, Pine wood nematode, Bursaphelenchus xylophilus, *Annu. Rev. Phytopathol.*, **51**, 61–83.
2. Vicente, C., Espada, M., Vieira, P. and Mota, M. 2012, Pine wilt disease: a threat to European forestry, *Eur. J. Plant Pathol.*, **133**, 89–99.
3. Liou, J., Shih, J. and Tzean, S. 1999, Esteya, a new nematophagous genus from Taiwan, attacking the pinewood nematode (*Bursaphelenchus xylophilus*), *Mycol. Res.*, **103**, 242–8.
4. Lin, F., Ye, J., Wang, H., Zhang, A., Zhao, B. and Hendricks, M. 2013, Host deception: predaceous fungus, *Esteya vermicola*, entices pine wood nematode by mimicking the scent of pine tree for nutrient, *PLoS One*, **8**, e71676.
5. Wang, C.Y., Fang, Z.M. and Wang, Z. 2011, Biological control of the pinewood nematode Bursaphelenchus xylophilus by application of the endoparasitic fungus *Esteya vermicola*, *BioControl*, **56**, 91–100.
6. Wang, C.Y., Fang, Z.M., Sun, B.S., Gu, L. J., Zhang, K.Q. and Sung, C.K. 2008, High infectivity of an endoparasitic fungus strain, *Esteya vermicola*, against nematodes, *J. Microbiol.*, **46**, 380–9.
7. Li, J., Zou, C. and Xu, J. 2015, Molecular mechanisms of nematode-nematophagous microbe interactions: basis for biological control of plant-parasitic nematodes, *Annu. Rev. Phytopathol.*, **53**, 67–95.
8. Yang, J., Wang, L., Ji, X., et al. 2011, Genomic and proteomic analyses of the fungus Arthrobotrys oligospora provide insights into nematode-trap formation, *PLoS Pathogens*, **7**, e1002179.
9. Meerupati, T., Andersson, K.-M., Friman, E., et al. 2013, Genomic mechanisms accounting for the adaptation to parasitism in nematode-trapping fungi, *PLoS Genetics*, **9**, 2005–10.
10. Zhang, L., Zhou, Z. and Guo, Q. 2016, Insights into adaptations to a near-obligate nematode endoparasitic lifestyle from the finished genome of *Drechmeria coniospora*, *Sci. Rep.*, **6**, 23122.
11. Lebrigand, K., He, L.D., Thakur, N., et al. 2016, Comparative genomic analysis of *Drechmeria coniospora* reveals core and specific genetic requirements for fungal endoparasitism of nematodes, *PLoS Genetics*, **12**, 1–41.
12. Larriba, E., Jaime, M.D.L.A.M.D.L.A., Carbonell-Caballero, J.J., et al. 2014, Sequencing and functional analysis of the genome of a nematode egg-parasitic fungus, *Pochonia chlamydosporia*, *Fungal Genet. Biol.*, **65**, 69–80.
13. Aranda-Martinez, A., Lenfant, N., Escudero, N., Zavala-Gonzalez, E.A., Henrissat, B. and Lopez-Llorca, L.V. 2016, CAZyme content of pochonia chlamydosporia reflects that chitin and chitosan modification are involved in nematode parasitism, *Environ. Microbiol.*, **18**, 4200–15.
14. Lai, Y., Liu, K., Zhang, X.X., et al. 2014, Comparative genomics and transcriptomics analyses reveal divergent lifestyle features of nematode endoparasitic fungus Hirsutella minnesotensis, *Genome Biol. Evol.*, **6**, 3077–93.
15. Liu, K., Zhang, W., Lai, Y., et al. 2014, Drechslerella stenobrocha genome illustrates the mechanism of constricting rings and the origin of nematode predation in fungi, *BMC Genomics*, **15**, 114.
16. Chu, W.H., Dou, Q., Chu, H.L., Wang, H.H., Sung, C K. and Wang, C.Y. 2015, Research advance on *Esteya vermicola*, a high potential biocontrol agent of pine wilt disease, *Mycol. Progress*, **14**, 115.
17. Chin, C.-S., Alexander, D.H., Marks, P., et al. 2013, Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data, *Nat. Methods*, **10**, 563–9.
18. Stanke, M., Schöffmann, O., Morgenstern, B. and Waack, S. 2006, Gene prediction in eukaryotes with a generalized hidden Markov model that uses hints from external sources, *BMC Bioinformatics*, **7**, 62.
19. Besemer, J. and Borodovsky, M. 2005, GeneMark: web software for gene finding in prokaryotes, eukaryotes and viruses, *Nucleic Acids Res.*, **33**, W451.
20. Birney, E., Clamp, M. and Durbin, R. 2004, Genewise and Genomewise, *Genome Res.*, **14**, 988–95.
21. Haas, B.J., Salzberg, S.L., Zhu, W., et al. 2008, Automated eukaryotic gene structure annotation using EVidenceModeler and the program to assemble spliced alignments, *Genome Biol.*, **9**, R7.
22. Zhang, J. and Madden, T.L. 1997, PowerBLAST: a new network BLAST application for interactive or automated sequence analysis and annotation, *Genome Res.*, **7**, 649–56.
23. Moriya, Y., Itoh, M., Okuda, S., Yoshizawa, A.C. and Kanehisa, M. 2007, KAAS: an automatic genome annotation and pathway reconstruction server, *Nucleic Acids Res.*, **35**, W182.
24. Quevillon, E., Silventoinen, V., Pillai, S., et al. 2005, InterProScan: protein domains identifier, *Nucleic Acids Res.*, **33**, W116.
25. Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P.M. and Henrissat, B. 2014, The carbohydrate-active enzymes database (CAZy) in 2013, *Nucl. Acids Res.*, **42**, D490.
26. Finn, R.D., Mistry, J., Schuster-Böckler, B., et al. 2006, Pfam: clans, web tools and services, *Nucleic Acids Res.*, **34**, D247–51.
27. Rawlings, N.D., Barrett, A.J. and Finn, R. 2016, Twenty years of the MEROPS database of proteolytic enzymes, their substrates and inhibitors, *Nucleic Acids Res.*, **44**, D343–50.
28. Saier, M.H., Reddy, V.S., Tsu, B.V., Ahmed, M.S., Li, C. and Moreno-Hagelsieb, G. 2016, The Transporter Classification Database (TCDB): recent advances, *Nucleic Acids Res.*, **44**, D372–9.
29. Urban, M., Irvine, A.G., Cuzick, A. and Hammond-Kosack, K.E. 2015, Using the pathogen-host interactions database (PHI-base) to investigate plant pathogen genomes and genes implicated in virulence, *Front. Plant Sci.*, **6**, 605.

30. Lu, T., Yao, B. and Zhang, C. 2012, DFVF: database of fungal virulence factors, *Database (Oxford)*, **2012**, bas032.
31. Liu, B. and Pop, M. 2009, ARDB–antibiotic resistance genes database, *Nucleic Acids Res*., **37**, D443–7.
32. Weber, T., Blin, K., Duddela, S., et al. 2015, AntiSMASH 3.0-A comprehensive resource for the genome mining of biosynthetic gene clusters, *Nucleic Acids Res*., **43**, W237–43.
33. R Development Core, T. 2012, R: a language and environment for statistical computing. *R Foundation for Statistical Computing*, Vienna, Austria. ISBN 3-900051-07-0, http://www.R-project.org/ (15 March 2012, date last accessed).
34. Li, L., Stoeckert, C.J. and Roos, D.S. 2003, OrthoMCL: identification of ortholog groups for eukaryotic genomes, *Genome Res*., **13**, 2178–89.
35. Katoh, K. and Standley, D.M. 2013, MAFFT multiple sequence alignment software version 7: Improvements in performance and usability, *Mol. Biol. Evol*., **30**, 772–80.
36. Castresana, J. 2000, Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis, *Mol. Biol. Evol*., **17**, 540–52.
37. Stamatakis, A. 2006, RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models, *Bioinformatics*, **22**, 2688–90.
38. Yang, Z. 2007, PAML 4: phylogenetic analysis by maximum likelihood, *Mol. Biol. Evol*., **24**, 1586–91.
39. Hedges, S.B., Dudley, J. and Kumar, S. 2006, TimeTree: a public knowledge-base of divergence times among organisms, *Bioinformatics*, **22**, 2971–2.
40. De Bie, T., Cristianini, N., Demuth, J.P. and Hahn, M.W. 2006, CAFE: a computational tool for the study of gene family evolution, *Bioinformatics*, **22**, 1269–71.
41. Letunic, I. and Bork, P. 2007, Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation, *Bioinformatics*, **23**, 127–8.
42. Yang, J., Huang, X., Tian, B., Wang, M., Niu, Q. and Zhang, K. 2005, Isolation and characterization of a serine protease from the nematophagous fungus, Lecanicillium psalliotae, displaying nematicidal activity, *Biotechnol. Lett*., **27**, 1123–8.
43. Wang, B., Wu, W. and Liu, X. 2007, Purification and characterization of a neutral serine protease with nematicidal activity from *Hirsutella rhossiliensis*, *Mycopathologia*, **163**, 169–76.
44. Wang, B., Liu, X., Wu, W., Liu, X. and Li, S. 2009, Purification, characterization, and gene cloning of an alkaline serine protease from a highly virulent strain of the nematode-endoparasitic fungus *Hirsutella rhossiliensis*, *Microbiol. Res*., **164**, 665–73.
45. Boldo, J.T., Junges, A., do Amaral, K.B., Staats, C.C., Vainstein, M.H. and Schrank, A. 2009, Endochitinase CHI2 of the biocontrol fungus Metarhizium anisopliae affects its virulence toward the cotton stainer bug Dysdercus peruvianus, *Curr. Genet*., **55**, 551–60.
46. Tikhonov, V.E., Lopez-Llorca, L.V., Salinas, J. and Jansson, H.-B. 2002, Purification and characterization of chitinases from the nematophagous fungi Verticillium chlamydosporium and V. suchlasporium, *Fungal Genet. Biol*., **35**, 67–78.
47. Collemare, J. and Lebrun, M.H. 2011, Fungal secondary metabolites: ancient toxins and novel effectors in plant-microbe interactions, In: Martin, F. and Kamoun, S. (eds), *Effectors in Plant-Microbe Interactions*. John Wiley & Sons, pp. 377–400.
48. Wolpert, T.J., Dunkle, L.D. and Ciuffetti, L.M. 2002, Host-selective toxins and avirulence determinants: what's in a name? *Annu. Rev. Phytopathol*., **40**, 251–85.
49. Dean, R.A., Talbot, N.J., Ebbole, D.J., et al. 2005, The genome sequence of the rice blast fungus *Magnaporthe grisea*, *Nature*, **434**, 980–6.
50. Powell, A.J., Conant, G.C., Brown, D.E., Carbone, I. and Dean, R.A. 2008, Altered patterns of gene duplication and differential gene gain and loss in fungal pathogens, *BMC Genomics*, **9**, 147.
51. Fillinger, S., Chaveroche, M.K., van Dijck, P., et al. 2001, Trehalose is required for the acquisition of tolerance to a variety of stresses in the filamentous fungus Aspergillus nidulans, *Microbiology*, **147**, 1851–62.
52. Lieberman, T.D., Michel, J.-B., Aingaran, M., et al. 2011, Parallel bacterial evolution within multiple patients identifies candidate pathogenicity genes, *Nat. Genet*., **43**, 1275–80.
53. Trivedi, P. and Wang, N. 2014, Host immune responses accelerate pathogen evolution, *ISME J*., **8**, 727–31.
54. Staats, M., van Baarlen, P., Schouten, A., van Kan, J.A.L. and Bakker, F.T. 2007, Positive selection in phytotoxic protein-encoding genes of Botrytis species, *Fungal Genet. Biol*., **44**, 52–63.
55. Biswas, S. and Akey, J.M. 2006, Genomic insights into positive selection, *Trends Genet*., **22**, 437–46.
56. Teichert, I. and Nowrousian, M. 2011, Evolution of genes for secondary metabolism in fungi, *Mycota* ***XVI***, 231–55.
57. Xu, Y., Orozco, R., Wijeratne, E.M.K., Gunatilaka, A.A.L., Stock, S.P. and Molnár, I. 2008, Biosynthesis of the cyclooligomer depsipeptide beauvericin, a virulence factor of the entomopathogenic fungus Beauveria bassiana, *Chem. Biol*., **15**, 898–907.
58. Xu, Y., Orozco, R., Kithsiri Wijeratne, E.M., et al. 2009, Biosynthesis of the cyclooligomer depsipeptide bassianolide, an insecticidal virulence factor of Beauveria bassiana, *Fungal Genet. Biol*., **46**, 353–64.
59. Wang, B., Kang, Q., Lu, Y., Bai, L. and Wang, C. 2012, Unveiling the biosynthetic puzzle of destruxins in *Metarhizium species*, *Proc. Natl. Acad. Sci. USA*, **109**, 1287–92.
60. Li, G., Zhang, K., Xu, J., Dong, J. and Liu, Y. 2007, Nematicidal substances from fungi, *Recent Pat. Biotechnol*., **1**, 212–33.
61. Liang, L., Wu, H., Liu, Z., et al. 2013, Proteomic and transcriptional analyses of Arthrobotrys oligospora cell wall related proteins reveal complexity of fungal virulence against nematodes, *Appl. Microbiol. Biotechnol*., **97**, 8683–92.
62. Andersson, K.-M., Kumar, D., Bentzer, J., Friman, E., Ahrén, D. and Tunlid, A. 2014, Interspecific and host-related gene expression patterns in nematode-trapping fungi, *BMC Genomics,*, **15**, 968.
63. Li, J., Yu, L., Yang, J., et al. 2010, New insights into the evolution of subtilisin-like serine protease genes in Pezizomycotina, *BMC Evol. Biol*., **10**, 68.
64. Zou, C.G., Tao, N., Liu, W.J., et al. 2010, Regulation of subtilisin-like protease prC expression by nematode cuticle in the nematophagous fungus Clonostachys rosea, *Environ. Microbiol*., **12**, 3243–52.
65. Yang, J., Zhao, X., Liang, L., et al. 2011, Overexpression of a cuticle-degrading protease Ver112 increases the nematicidal activity of Paecilomyces lilacinus, *Appl. Microbiol. Biotechnol*., **89**, 1895–903.
66. Langfelder, K., Streibel, M., Jahn, B., Haase, G. and Brakhage, A.A. 2003, Biosynthesis of fungal melanins and their importance for human pathogenic fungi, *Fungal Genet. Biol*., **38**, 143–58.
67. St Leger, R.J., Frank, D.C., Roberts, D.W. and Staples, R. C. 1992, Molecular cloning and regulatory analysis of the cuticle-degrading-protease structural gene from the entomopathogenic fungus Metarhizium anisopliae, *Eur. J. Biochem*., **204**, 991–1001.
68. Małagocka, J., Grell, M.N., Lange, L., Eilenberg, J. and Jensen, A.B. 2015, Transcriptome of an entomophthoralean fungus (Pandora formicae) shows molecular machinery adjusted for successful host exploitation and transmission, *J. Invertebr. Pathol*., **128**, 47–56.
69. Ward, E., Kerry, B.R., Manzanilla-López, R.H., et al. 2012, The *Pochonia chlamydosporia* serine protease gene vcp1 is subject to regulation by carbon, nitrogen and pH: implications for nematode biocontrol, *PLoS One*, **7**, e35657.
70. Yang, J., Tian, B., Liang, L. and Zhang, K.Q. 2007, Extracellular enzymes and the pathogenesis of nematophagous fungi, *Appl. Microbiol. Biotechnol*., **75**, 21–31.
71. Prasad, P., Varshney, D. and Adholeya, A. 2015, Whole genome annotation and comparative genomic analyses of bio-control fungus Purpureocillium lilacinum, *BMC Genomics*, **16**, 1004.
72. Wanyiri, J.W., Techasintana, P., O'Connor, R.M., Blackman, M.J., Kim, K. and Ward, H.D. 2009, Role of CpSUB1, a subtilisin-like protease, in cryptosporidium parvum infection in vitro, *Eukaryotic Cell.*, **8**, 470–7.
73. Martinez, D., Challacombe, J. and Morgenstern, I. 2009, Genome, transcriptome, and secretome analysis of wood decay fungus Postia placenta supports unique mechanisms of lignocellulose conversion, *Proc. Natl. Acad. Sci. USA*, **106**, 1954–9.

74. Schuler, D., Wahl, R., Wippel, K., et al. 2015, Hxt1, a monosaccharide transporter and sensor required for virulence of the maize pathogen Ustilago maydis, *New Phytol*., **206**, 1086–100.
75. Cui, Z., Gao, N., Wang, Q., Ren, Y., Wang, K. and Zhu, T. 2015, BcMctA, a putative monocarboxylate transporter, is required for pathogenicity in *Botrytis cinerea*, *Curr. Genet*., **61**, 545–53.
76. Doherty, C.P. 2007, Host-pathogen interactions: the Role of Iron, *J. Nutr*., **137**, 1341–4.
77. Haas, H., Eisendle, M. and Turgeon, B.G. 2008, Siderophores in fungal physiology and virulence, *Annu. Rev. Phytopathol*., **46**, 149–87.
78. Haas, H. 2012, Iron—a key nexus in the virulence of Aspergillus fumigatus, *Front. Microbiol*., **3**, 1–10.
79. Ibrahim, A.S., Gebremariam, T., Lin, L., et al. 2010, The high affinity iron permease is a key virulence factor required for Rhizopus oryzae pathogenesis, *Mol. Microbiol*., **77**, 587–604.
80. Berg, R.H., Fester, T. and Taylor, C.G. 2009, *Cell Biology of Plant Nematode Parasitism*. In: Berg R.H. and Taylor C.G. (eds), Plant cell monographs, Vol. 15. Springer-Verlag: Berlin, pp. 115–52.
81. Pitson, S., Seviour, R.J., Bott, J. and Stasinopoulos, S.J. 1991, Production and regulation of β-glucanases in Acremonium and Cephalosporium isolates, *Mycol. Res*., **95**, 352–6.
82. de la Cruz, J., Pintor-Toro, J.A., Benítez, T., Llobell, A. and Romero, L.C. 1995, A Novel endo-beta-1,3-glucanase, BGN13.1, involved in the mycoparasitism of Trichoderma harzianum, *J. Bacteriol*., **177**, 6937–45.
83. El-Katatny, M.H., Gudelj, M., Robra, K.H., Elnaghy, M.A. and Gübitz, G.M. 2001, Characterization of a chitinase and an endo-β-1, 3-glucanase from Trichoderma harzianum Rifai T24 involved in control of the phytopathogen Sclerotium rolfsii, *Appl. Microbiol. Biotechnol*., **56**, 137–43.
84. Zhao, L.L., Wei, W., Kang, L. and Sun, J.H. 2007, Chemotaxis of the pinewood nematode, Bursaphelenchus xylophilus, to volatiles associated with host pine, Pinus massoniana, and its vector Monochamus alternatus, *J. Chem. Ecol*., **33**, 1207–16.
85. Niu, H., Zhao, L., Lu, M., Zhang, S., Sun, J. and Yang, C.-H. 2012, The ratio and concentration of two monoterpenes mediate fecundity of the pinewood nematode and growth of its associated fungi, *PLoS One*, **7**, e31716.
86. Wang, R., Dong, L., Chen, Y., Qu, L., Wang, Q. and Zhang, Y. 2017, Esteya Vermicola, a nematophagous fungus attacking the pine wood nematode, harbors a bacterial endosymbiont affiliated with gammaproteobacteria, *Microbes Environ*., **32**, 201–9.