Divergent and Convergent Evolution of Fungal Pathogenicity

Yanfang Shang[†], Guohua Xiao[1,†], Peng Zheng, Kai Cen, Shuai Zhan, and Chengshu Wang[*]

Key Laboratory of Insect Developmental and Evolutionary Biology, Institute of Plant Physiology and Ecology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai, China

*Corresponding author: E-mail: cswang@sibs.ac.cn.

[†]These authors contributed equally to this work.

[1]Present address: School of Computer Science, Fudan University, Shanghai, China

Abstract

Fungal pathogens of plants and animals have multifarious effects; they cause devastating damages to agricultures, lead to life-threatening diseases in humans, or induce beneficial effects by reducing insect pest populations. Many virulence factors have been determined in different fungal pathogens; however, the molecular determinants contributing to fungal host selection and adaptation are largely unknown. In this study, we sequenced the genomes of seven ascomycete insect pathogens and performed the genome-wide analyses of 33 species of filamentous ascomycete pathogenic fungi that infect insects (12 species), plants (12), and humans (9). Our results revealed that the genomes of plant pathogens encode more proteins and protein families than the insect and human pathogens. Unexpectedly, more common orthologous protein groups are shared between the insect and plant pathogens than between the two animal group pathogens. We also found that the pathogenicity of host-adapted fungi evolved multiple times, and that both divergent and convergent evolutions occurred during pathogen–host cospeciation thus resulting in protein families with similar features in each fungal group. However, the role of phylogenetic relatedness on the evolution of protein families and therefore pathotype formation could not be ruled out due to the effect of common ancestry. The evolutionary correlation analyses led to the identification of different protein families that correlated with alternate pathotypes. Particularly, the effector-like proteins identified in plant and animal pathogens were strongly linked to fungal host adaptation, suggesting the existence of similar gene-for-gene relationships in fungus–animal interactions that has not been established before. These results well advance our understanding of the evolution of fungal pathogenicity and the factors that contribute to fungal pathotype formation.

Key words: fungal pathogen, comparative genomics, convergent evolution, pathotype, host adaptation, effector.

Introduction

Evolution and speciation of organisms are primarily achieved by genetic divergence that eventually become distinct biological traits suited to adapt to different environments (Clark et al. 2007; Hu et al. 2014). Fungal species have been widely used as models to study the patterns of divergent speciation as well as convergent evolution for environmental adaptations including host selection and infection. Various fungal pathotypes (pathogenic variants for different hosts) have been observed in different pathogens such as the seven pathotypes of plant pathogenic fungus *Alternaria alternata* against different plant hosts (Takaoka et al. 2014). However, infection across kingdom or phylum by a single fungal species rarely occurs in nature. Therefore, it is uncommon that the plant pathogen *Fusarium oxysporum* can infect immunocompromised humans as well as insects (Coleman et al. 2011; Salah et al. 2015). Aside from this exception, most single species or phylogenetically related or unrelated fungi are largely pathogenic against hosts belonging to the same phylum or even the same species. Although fungal virulence factors have been widely studied (van de Wouw and Howlett 2011), the evolution of fungal pathotypes against alternative hosts is still not clearly understood.

Fungal pathogens of plants and humans cause devastating damage to agricultures or cause life-threatening diseases in

humans. Insect pathogenic fungi, however, play a pivotal role in maintaining insect populations in ecosystems. A number of factors are known to contribute to fungal virulence during their interactions with different hosts (van de Wouw and Howlett 2011; Urban et al. 2015). For example, alkaline proteases and chitinases are crucial for insect fungal pathogens to degrade insect cuticles (Ortiz-Urquiza and Keyhani 2013; Wang and Feng 2014). Similarly, enzymes digesting carbohydrates are the main virulence determinants for plant pathogens to digest plant cell walls (Perez-Nadales et al. 2014). Differently, fungal pathogens infecting mammals have to survive a high temperature (37 °C), and evade both innate and adaptive immune systems in hosts (Casadevall 2007). The protein families of adhesins and aspartic proteinases are particularly important for human pathogens to mediate cell adherence to host tissues, formation of biofilms, and hydrolysis of host tissues (Schaller et al. 2005; de Groot et al. 2013). However, it is unclear which determinants enable fungi to recognize and infect alternative hosts particularly across kingdom or phylum.

It is known that approximately 1,000 known fungal species can kill insects (Shang et al. 2015). Despite their prevalence, genome sequences are available for only a few insect fungal pathogens such as *Metarhizium* spp. (Gao et al. 2011; Hu et al. 2014), *Beauveria bassiana* (BBA) (Xiao et al. 2012), *Cordyceps militaris* (CCM) (Zheng et al. 2011), and *Ophiocordyceps* spp. (Bushley et al. 2013; Hu et al. 2013; de Bekker et al. 2015; Wichadakul et al. 2015). Therefore, in this study, seven filamentous ascomycete insect pathogens were sequenced for de novo assemblies and comparative analyses with the plant and human pathogens to explore the evolution of fungal pathotype-associated determinants. The seven species include the honeybee pathogen (Chalkbrood disease) *Ascosphaera apis* (locus-tagged and abbreviated as AAP), *Nomuraea rileyi* (NOR), *Aschersonia aleyrodis* (AAL), *Lecanicillium lecanii* (LEL), *B. brongniartii* (BBO), *Isaria fumosorosea* (ISO), and *Sporothrix insectorum* (SPI). Evolutionary and comparative analyses were performed with the published genome sequences from other filamentous ascomycetes. In total, we analyzed 33 fungal pathogens that infect insects (12 species, abbreviated as IFs), plants (12 species, PFs), and mammals (9 species, MFs). The results puzzle out the evolutionary patterns of fungal pathogenicity and the protein families that correlate to fungal pathotype formations.

Materials and Methods

Fungal Species

Seven insect pathogens were selected for de novo genome sequencing (supplementary table S1, Supplementary Material online). These include NOR (strain RCEF 4871), *Isaria fumosorosea* (ARSEF 2679), AAL (RCEF 2490), *Lecanicillium lecanii* (RCEF 1005), *Beauveria brongiartii* (RCEF 3172), SPI (RCEF

264), and the bee pathogen, AAP (ARSEF 7405). Single spore cultures were prepared for each species and used for DNA isolation and genome sequencing. Besides, five genomes of previously sequenced IF species belonging to *Cordyceps* sensu lato were also included in the analyses (Gao et al. 2011; Zheng et al. 2011; Xiao et al. 2012; Bushley et al. 2013).

For phylogenetic and evolutionary correlation analyses, 12 filamentous ascomycete plant pathogens (*Fusarium graminearum* [FGR], *Verticillium dahliae*, *V. albo-atrum*, *Magnaporthe oryzae*, *Botrytis cinerea*, *Sclerotinia sclerotiorum*, *Mycosphaerella graminicola* [MGR], *Mycosphaerella fijiensis*, *Leptosphaeria maculans*, *Stagonospora nodorum*, *Cochliobolus heterostrophus*, *Pyrenophora tritici-repentis*) and nine ascomycete human pathogens (*Aspergillus fumigatus*, *Coccidioides immitis*, *C. posadasii*, *Histoplasma capsulatum*, *Paracoccidioides lutzii*, *P. brasiliensis*, *Trichophyton rubrum*, *Microsporum gypseum*, *Arthroderma benhamiae*) with published genome data were selected in this study (supplementary table S2, Supplementary Material online). Thus, in total, we analyzed 33 fungal pathogens including 12 insect pathogenic fungi (labeled as I1–I12 for different species for simplicity in analyses), 12 plant pathogenic fungi (P1–P12), and 9 mammalian pathogenic fungi (M1–M9).

Genome Sequencing, Assembly, and Annotation

Three DNA libraries with 170-, 500-, and 5,000-bp fragment sizes from each species were constructed and sequencing was performed with the Illumina Hiseq2000 system (BGI, Shenzhen, China). The obtained data were first processed using FastQC (ver. 0.10.1, http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) to filter out low-quality reads before de novo assembly with the program SOAPdenovo (ver. 1.05, http://soap.genomics.org.cn/) using a *K*-mer value of 51. Four programs were used for gene prediction as we previously described (Hu et al. 2014), and the best gene model was filtered for each species using the EvidenceModeler algorithm (Haas et al. 2008). To assess the completeness of each genome, a mapping approach for core eukaryotic genes was used (Parra et al. 2007).

Phylogenetic Analyses of Fungal Pathogens

By referring to a core set of proteins used for building robust fungal phylogeny (Floudas et al. 2012), 47 protein sequences were selected from 33 fungal species. In addition, the genome of the basidiomycete plant pathogenic fungus *Ustilago maydis* was included as an outgroup. Protein sequence alignments were performed using MUSCLE (ver. 3.8.31) (Edgar 2004). Concatenated sequences were then used to generate a maximum-likelihood tree using the program TREE-PUZZLE (ver. 5.2) (Schmidt et al. 2002), which was run with the parameter settings of a Dayhoff substitution model, a partial deletion for

gaps or missing data, and 100 bootstrap replicates for phylogeny test.

Analysis of Genetic Similarity among Three Groups of Fungi

To investigate the genetic similarity among different fungal species, we calculated the global frequencies of consensus clustering at each gene locus. We first identified 2,754 universal genes from 33 fungal species (allowing up to one gene missing in a given species) as well as the outgroup species *U. maydis* based on the OrthoMCL analysis (http://www.orthomcl.org/). After individual protein sequence alignment, the conserved blocks of these alignments were extracted using Gblocks (ver. 0.91b) with the default parameter settings (Talavera and Castresana 2007). Multidimensional scaling (MDS) was subsequently performed based on the concatenated sequences extracted from the conserved blocks, and the *K*-means ($K = 3$) clustering was conducted to group each fungal species into three clusters: IF, PF, and MF. Pairwise genetic distance was calculated by the Protdist function in Phylip using the JTT (Jones, Taylor, and Thorton) model (Plotree and Plotgram 1989). Average values of the distance to all other members within the groups were then calculated for each species. Scatter plots of the resulting MDS were generated by an R package in Bios2mds (Pele et al. 2012).

Protein Family Classification and Comparative Analyses

The total sets of proteins encoded in the genomes of the 33 fungal species were scanned for protein families by InterProScan analysis (Jones et al. 2014). Further search for proteases, cytochrome P450s (CYPs), carbohydrate-active enzymes (CAZy), transporters, G-protein coupled receptors (GPCRs), dehydrogenases, and putative virulence factors were performed according to our previous protocols (Xiao et al. 2012; Hu et al. 2013). To identify gene clusters involved in the biosynthesis of secondary metabolites, the whole-genome data set of each species was analyzed with the program antiSMASH (ver. 2.02) (Blin et al. 2013). Orthologous relationship of each gene cluster identified in different species was determined based on the OrthoMCL data and reciprocal BLAST (Basic Local Alignment Search Tool) analysis. The effector-like SSCPs (small secreted cysteine-rich proteins, less than 300 amino acid residues) from each selected species were classified based on a previously described algorithm (Kohler et al. 2015). The average size of each protein family was estimated for each fungal group and compared for the level of difference between the groups by *t*-test. The significance of the *P* values was also checked by calculation of the false discovery rate when it was applicable based on a mixture model (Efron 2007). The evolution of protein family size variation (loss or gain) was analyzed using the program CAFE (de Bie et al. 2006). To evaluate the significance of protein family size variation, one-tail *t*-test was conducted for each protein family

among IF, PF, and MF. Principal component analysis (PCA) was also carried out based on protein family sizes.

Phylogenetic Independent Contrasts Tests

To determine the protein families' associations with fungal pathotypes against different hosts (i.e., insects, plants, or mammals) by removing the effect of shared evolutionary history, we conducted the analysis of phylogenetically independent contrasts (PIC) using the APE package (ver. 2.2-2) in R (Felsenstein 1985). Briefly, the established fungal phylogeny (fig. 1) was parsed into APE to infer the phylogenetic branch lengths, and the protein family sizes of each species were transformed into a data matrix, accompanying with their pathotype variables of 0, 1, and 2 for the pathogenic fungi of plants, mammals, and insects based on their overall phylogenetic relationships (fig. 1). A linear regression was subsequently run to calculate the correlation coefficients between the paired pathotype and protein family features (Hu et al. 2014).

Growth Test of Insect Pathogens on Different Nitrogen and Carbon Sources

To further verify the phenotypic features associated with protein family traits, 12 insect pathogens and the plant pathogen, FGR (strain PH-1, used as a reference), were tested for the growth on different carbon and nitrogen sources. The fungi were grown on potato dextrose agar for 14 days, and conidia of each species were harvested and adjusted to a final concentration of 10^7 cells/ml. Conidial suspension (2.5 μl) from each species was then inoculated onto a minimal medium agar (NaNO$_3$ 6 g/l, KCl 0.52 g/l, MgSO$_4$.7H$_2$O 0.52 g/l, KH$_2$PO$_4$ 0.25 g/l) supplemented with 1% casein acid hydrolysate, cellulose, glucose, pectin, sucrose or xylose and incubated for 14 days to monitor growth.

Results

Genome Features and Phylogenetic Analysis

Genome sequencing of seven insect fungal pathogens obtained greater than 97% completeness for each genome. Genome assembly and annotation results revealed the typical genomic features of filamentous ascomycete fungi in these pathogens with the genome sizes ranging from 30 to 40 Mb that encode 9,000–12,000 genes except for the bee-specific pathogen, AAP (genome size 20.4 Mb with 6,442 predicted genes). In terms of the GC content, the genome of bee pathogen is also unique with a GC ratio (47.67%) lower than those of other fungi (49.94–54%) (supplementary table S1, Supplementary Material online). Different from the highly repetitive genome of the caterpillar fungus *O. sinensis* (37.98%, Hu et al. 2013), the repeat rates of these sequenced fungal genomes are typically low (0.41–6.15%). Our genome survey of mating-type genes indicated that a single mating-type
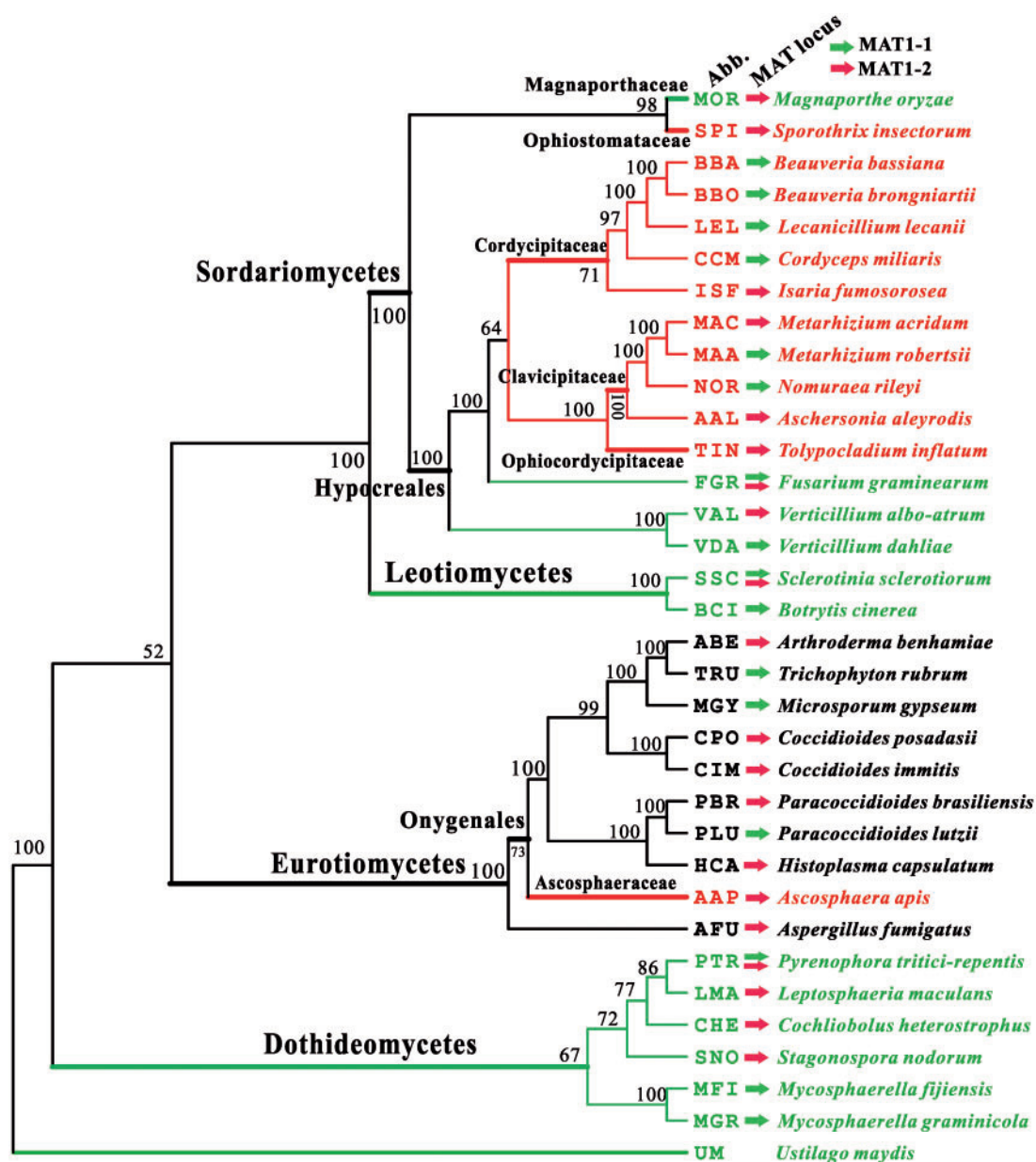
Fig. 1.—Phylogenetic analysis of insect, plant, and mammalian pathogenic fungi. The maximum-likelihood tree was generated using the concatenated sequences of 47 universal proteins with the program TREE-PUZZLE using a Dayhoff substitution model, a partial deletion for gaps or missing data, and 100 bootstrap replicates. The tree is rooted by the basidiomycete plant pathogen *Ustilago maydis* (UM). Green (MAT 1-1) and red (MAT 1-2) arrows indicate the mating-type of each sequenced strain. Species with both arrows are homothallic fungi. Insect fungi are labeled in red, plant fungi in green, and mammalian fungi in black.

(MAT) locus, that is, either MAT1-1 (NOR, BBO, LEL) or MAT1-2 (SPI, ISF, AAL, AAP), is present in the haploid genome of each sequenced strain (fig. 1). Thus, in contrast to the homothallic (self-fertile) pathogen *O. sinensis* (Hu et al. 2013), the heterothallic (self-sterile) nature of these species is similar to other species of *Cordyceps* sensu lato (Zheng et al. 2013). Overall, most of the selected pathogenic fungi are self-sterile (30/33 only containing a single MAT locus), and no clear

pattern could be evident between fungal reproductive modes (i.e., self-fertile or self-sterile) and pathotypes for different hosts.

Our phylogenetic analysis demonstrated that the examined pathogenic fungi could be divided into four groups of Dothideomycetes, Eurotiomycetes, Leotiomycetes, and Sordariomycetes in accordance with the overall evolutionary relationships of ascomycetes (Schoch et al. 2009). The result

suggested that fungal entomopathogenicity and phytopathogenicity evolved multiple times in the phylum Ascomycota (fig. 1). Although we did not include the yeast-like pathogens (e.g., *Candida* spp. and *Pneumocystis* spp.) in this study due to the concerns of the immense evolutionary gaps existing between the yeasts and filamentous fungi, it appears that the pathogenicity of ascomycetes infecting mammals may also have arisen repeatedly during fungal evolution (James et al. 2006; Schoch et al. 2009). The phylogeny analysis confirmed again that AAP is a unique insect pathogen because it clustered with mammalian pathogens. Consistent with a previous analysis of human Sporotrichosis pathogens *Sporothrix schenckii* (SCH) and *S. brasiliensis* (SBR) (Teixeira et al. 2014), the insect pathogen SPI appeared to be closely related to the plant pathogens *Grosmannia clavigera* and *Magnaporthe oryzae* (fig. 1 and supplementary fig. S1A, Supplementary Material online). Otherwise, ascomycete insect fungal pathogens have diversified in three families of Cordycipitaceae, Clavicipitaceae, and Ophiocordycipitaceae in the order Hypocreales (fig. 1). Further analyses of the three *Sporothrix* species indicated that the two human pathogens (SBR, SCH) share more orthologs with each other than with the insect pathogen SPI (supplementary fig. S1B, Supplementary Material online). In addition, comparison of the amino acid sequences of 5,925 universal proteins indicated that the two human pathogens share a higher identity (97.0%) than

between the human and insect pathogens (SBR vs. SPI, 65.5% and SCH vs. SPI, 65.1%), an indicative of host adaptive evolution opposed to phylogenetic relatedness.

For the three groups of fungal pathogens, the average protein-coding capacity of the IF genomes (9,527) was similar to those of MF (average 9,918, $P = 0.2232$). However, both numbers were fewer than the average number of the proteins encoded in the PF genomes (12,308, $P < 0.001$) (table 1). To further explore the relationships among the three fungal groups, we performed OrthoMCL analysis to identify orthologous protein groups. The results revealed that more common groups were shared between IF and PF (2,455) than between IF and MF (2,399), and between PF and MF (1,786) (fig. 2A). We also found that MF (2,293) had more lineage-specific orthologous groups followed by the IF (2,161) and PF (548), an indication of divergent evolution occurring at various levels. Based on the similarities among universal genes present in each species, an MDS analysis was performed. The results demonstrated again that IF and PF were more closely related to each other than each to MF, especially for those species within the class Sordariomycetes (fig. 2B). Except for AAP, our PCA based on the OrthoMCL data divided the three fungal groups into three clusters (fig. 3A), which indicated a convergent evolution of fungal pathotypes in association with

Table 1
Comparison of the Protein Family Sizes among the Three Groups of Fungi

Protein Family	Average Size			P value (*t*-test)			PIC Test	
	IF	PF	MF	IF versus PF	IF versus MF	PF versus MF	*F* value	*P*
Total proteins	9526.83	12308.25	9198.33	0.0002	0.2232	0.0001	18.2300	0.0002
Conserved protein families	2791.67	2849.42	2715.44	0.1048	0.1114	0.0118	4.2420	0.0479
Conserved proteins	6907.00	7487.00	5928.56	0.0300	0.0017	1.7694E-06	10.5200	0.0028
Secreted proteins	1130.17	1373.08	701.78	0.0142	1.2240E-04	3.5398E-08	16.6600	0.0003
PHI[a]	1550.75	1724.08	1249.22	0.0268	0.0033	3.2948E-05	14.4000	0.0006
SSCP[b]	262.25	333.33	158.56	0.0233	0.0009	1.80242E-06	16.0500	0.0004
GPCR	37.75	58.00	27.22	0.0060	0.0234	0.0002	13.3700	0.0002
TFs	144.83	164.50	132.11	0.0612	0.1078	0.0146	2.7620	0.1066
Secreted proteases	109.08	88.17	53.89	0.0769	0.0006	0.0018	3.0450	0.0909
GHs	128.33	207.67	89.44	3.8215E-06	0.0576	1.6147E-04	15.6000	0.0004
CEs	8.92	24.17	9.56	6.2655E-04	0.3761	1.0176E-03	14.4400	0.0006
PLs	0.17	9.67	1.44	8.1521E-05	0.2016	7.9088E-04	17.3700	0.0002
GTs	67.83	79.08	68.11	0.0123	0.4790	0.0470	8.4460	0.0067
Cutinases	4.08	11.08	2.00	2.4531E-04	0.0757	1.0263E-06	23.0900	3.745e-5
Lipases	24.67	58.58	30.56	0.0002	0.1814	2.5270E-03	17.1800	0.0002
Dehydrogenases	241.50	302.50	190.44	0.0012	0.0126	2.2845E-05	5.7940	0.0222
CYPs	80.92	116.58	62.89	0.0039	0.0470	4.8784E-05	13.9600	0.0008
Transporters	517.00	512.25	393.89	0.4315	0.0009	8.4126E-04	0.9541	0.3362
SMs	34.83	35.50	26.33	0.4419	0.0481	0.0390	4.2080	0.0488
Horizontally transferred genes	43.92	45.00	14.56	0.4560	0.0001	0.0017	2.2090	0.1473

[a]Proteins involved in PHI.
[b]Putative effectors of SSCP.
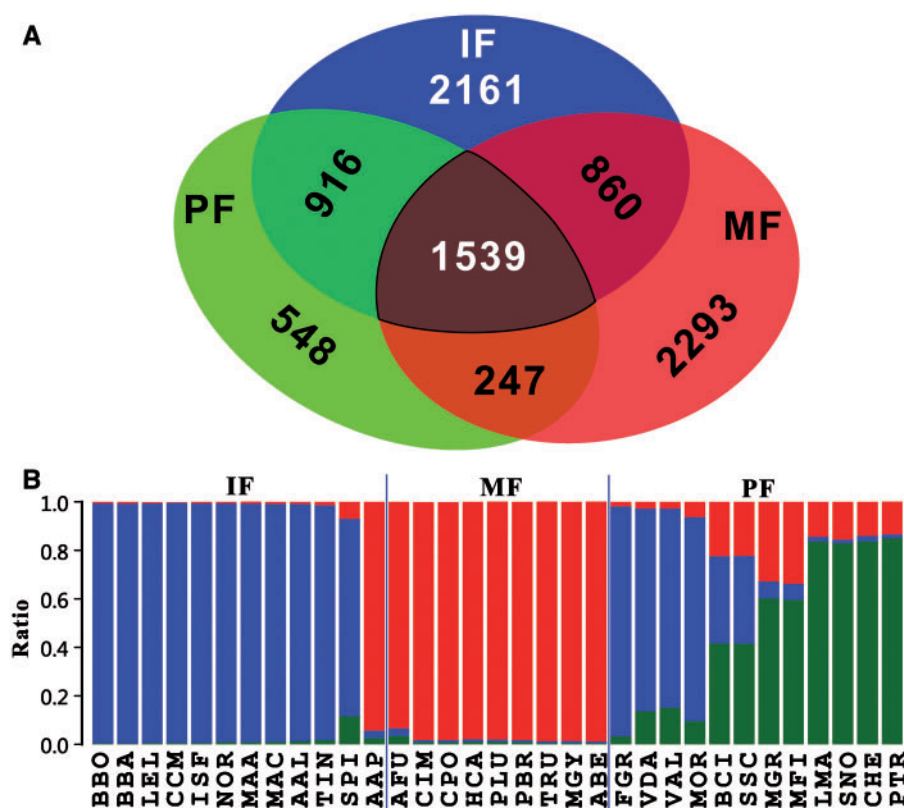
FIG. 2.—Analyses of relationships between three groups of fungi. (A) Venn diagram of IF, PF, and MF pathogenic fungi based on the orthologous protein groups identified by OrthoMCL analysis. (B) Genetic similarity among three groups of fungi. MDS analysis was performed based on the extracted conserved blocks of individual orthologous protein sequence ($n = 2,754$) alignments, and the K-means clustering ($K = 3$) was used to group each fungal species into three clusters for IF (blue), MF (red), and PF (green). Pairwise genetic distance was calculated, and the length of each colored segment in each species represents the proportion of its orthologous proteins that represent the minimum average genetic distance to the corresponding fungal group.

different hosts. The finding that AAP is more like a mammal pathogen would suggest a role of phylogenetic relatedness being opposed to host adaptation.

Protein Family Variations and Pathotype Associations

InterProScan analysis identified 5,337 conserved protein families from 33 examined fungal species. For these protein families, the differences were significant only between MF (2,715) and PF (2,849) ($P = 0.0118$). It was found that insect (1,130) and plant (1,373) pathogens have higher numbers of secreted proteins than MF (702) ($P < 0.001$) (table 1; fig. 4A and B). By homology analysis with the previously reported factors involved in pathogen–host interactions (PHI) (Urban et al. 2015), plant pathogens were found to encode more PHI factors (1,724) than IF (1,551) and MF (1,249) ($P < 0.05$). Based on the size of PHI factors in each species, PCA could generally separate the three groups of fungi into different clusters (fig. 3B).

Based on the established phylogeny (fig. 1), we performed PIC tests to determine the evolutionary correlation between protein families across taxa (Felsenstein 1985). It was found

that the features of total proteins ($F = 18.23$; $P = 0.0002$), protein families ($F = 4.24$; $P = 0.0479$), conserved proteins ($F = 10.52$; $P = 0.0028$), and secreted proteins ($F = 16.66$; $P = 0.0003$) significantly correlated to the fungal pathotypes against different hosts (table 1). Not surprisingly, the PIC test revealed that the PHI factors significantly ($F = 14.4$; $P = 0.0006$) determine fungal pathogenicity against different hosts. Secreted proteins, especially the SSCP-type effectors, are the key determinants of fungus–host interactions (Schmidt and Panstruga 2011). On average, PF (333) had more SSCPs than IF (262; $P = 0.0233$) and MF (159; $P = 1.802e-6$) (fig. 4C). That these factors are the essential determinants of fungal pathogenicity against different hosts was well supported by the PIC test ($F = 16.05$; $P = 0.0004$). Based on the SSCP orthologous groups, the PCA also split the three fungal groups into three clusters, in particular, the honeybee pathogen AAP could be well separated from the same clade of mammalian fungi (fig. 3C). In total, 467 of more than 2,700 conserved protein families were identified to have significant ($P < 0.05$) correlations to fungal pathotypes (supplementary tables S3–S17, Supplementary Material online), which could
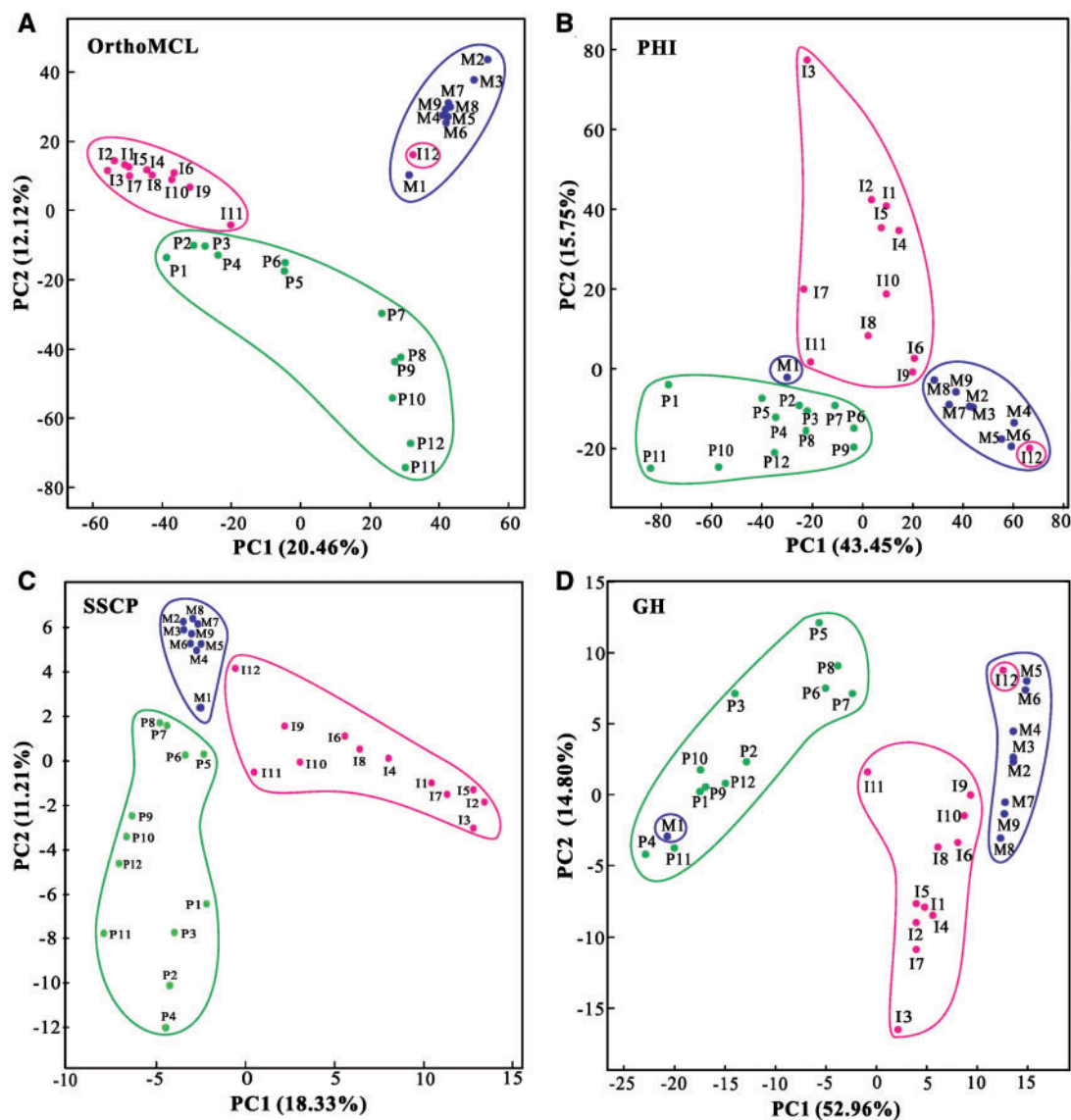
Fig. 3.—PCA of fungal species. PCA of examined fungal species based on (A) the orthologous protein groups identified in OrthoMCL analysis, (B) the number of proteins putatively involved in PHI, (C) the number of SSCP effectors, and (D) the number of GHs. Species I1–I12 are simplified for 12 insect pathogenic fungi, P1–P12 for 12 plant pathogenic fungi, and M1–M9 for nine mammalian pathogenic fungi that are detailed in supplementary table S2, Supplementary Material online.

be the important factors to drive fungal evolution and adaptation to different hosts.

Correlation between Host Selection and Signal Sensing in Fungi

Similar to other organisms, GPCRs mediate recognition of environmental signals (including host signals) to activate downstream transcription factors (TFs) in fungi (Xue et al. 2008; Gao et al. 2011). We surveyed the putative GPCRs encoded in the 33 examined fungal species, and found that the number of GPCR classes varied substantially among individual species

(supplementary table S3, Supplementary Material online), that is, the example of divergent evolution of fungal host recognition mechanisms. On average, PF (58) had more ($P < 0.05$) GPCRs than IF (38) and MF (27) (table 1). Consistent with the significant correlation between the GPCR numbers and host range in *Metarhizium* spp. (Hu et al. 2014), the PIC test indicated that fungal GPCR family sizes essentially correlated to pathogen–host selection and infection ($F = 13.37$; $P = 0.0002$). Among the different classes of GPCRs, the Pth11-like GPCRs were significantly linked to fungal pathotypes ($F = 14.4$; $P = 0.0006$), that is, the essential
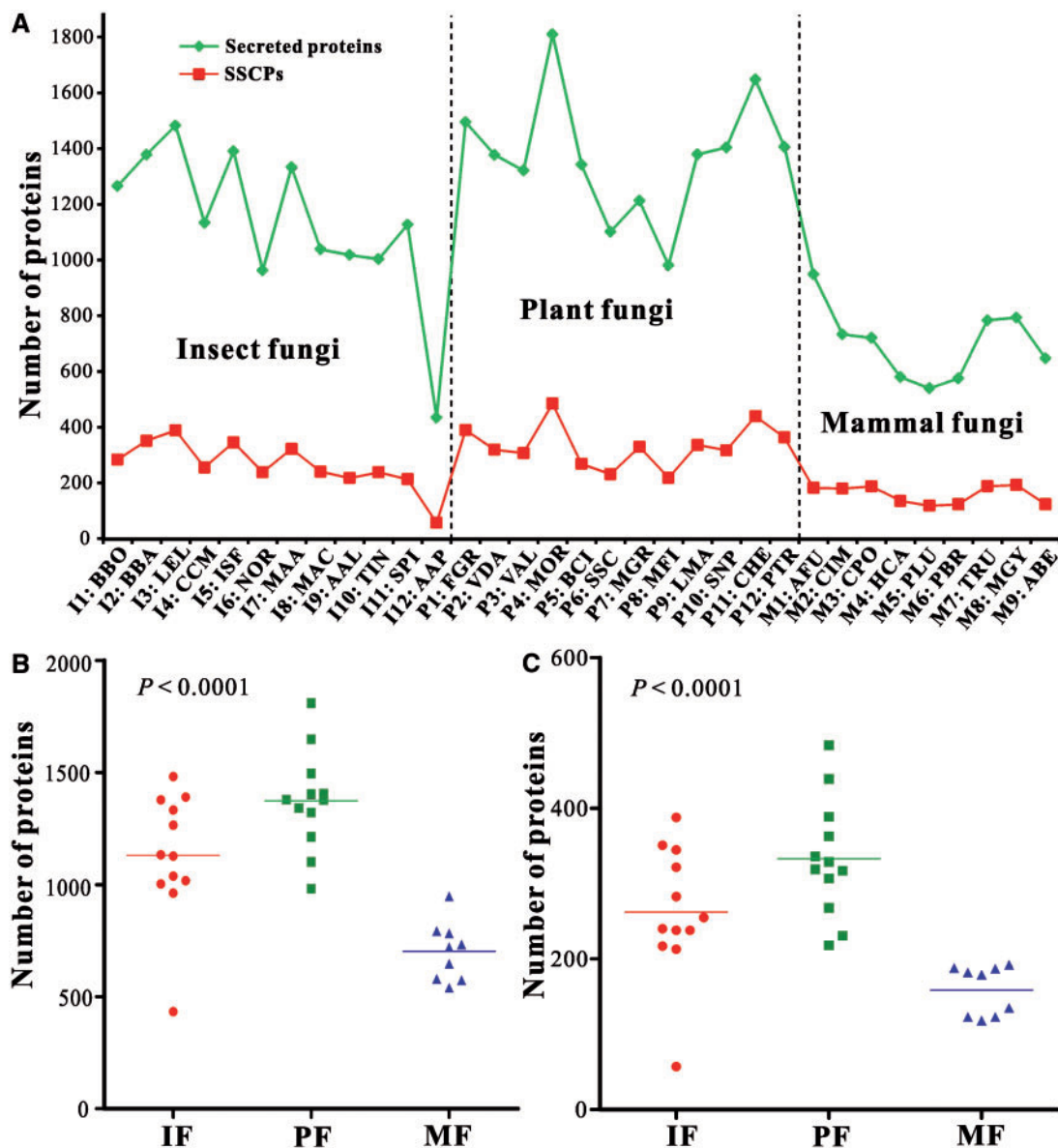
FIG. 4.—Characterization of secreted proteins and putative effectors of SSCPs. (A) Kinetic profiles of the secreted proteins and SSCPs encoded by each fungal species. Species I1–I12 are simplified for 12 insect pathogenic fungi (IF), P1–P12 for 12 plant pathogenic fungi (PF), and M1–M9 for nine mammalian pathogenic fungi (MF) that are detailed in supplementary table S2, Supplementary Material online. (B) Average number of secreted proteins encoded by IF, PF, and MF, respectively. (C) Average number of SSCPs encoded by IF, PF, and MF, respectively.

factors in association with fungal pathotype formations. However, the Class B type GPCRs involved in stress responses ($F = 0.1004$; $P = 0.7534$), the Class D GPCRs functioning as fungal mating pheromone receptors ($F = 0.0076$; $P = 0.931$), and Class E type of cAMP receptors ($F = 1.394$; $P = 0.2468$) (Xue et al. 2008) are not linked to fungal pathotypes against different hosts (supplementary table S3, Supplementary Material online). Nevertheless, functional characterization of a class B-type GPCR3 (BBA_00868) in the insect pathogen BBA indicated that the protein mediates stress responses and contributes to fungal virulence (Ying et al. 2013).

After GPCRs' sense of host inductive signals, the downstream mitogen-activated protein kinase and cAMP-dependent protein kinase A pathways are activated to trigger the expression of TFs (Xue et al. 2008; Gao et al. 2011). Our genome survey revealed the presence of various numbers of TF families in each fungal species (supplementary table S4, Supplementary Material online). Overall, the TF numbers did not correlate to fungal pathotype formations ($F = 2.762$; $P = 0.1066$). The PCA also failed to separate the three groups of fungi based on TF families (supplementary fig. S2A, Supplementary Material online), that is, the indicative

of divergent evolution of TF family sizes within and between fungal pathotypes

Correlation between the Hydrolytic Enzymes and Fungal Pathotypes

Pathogenic microbes have evolved multiple enzymes to access nutrients from hosts (Schaible and Kaufmann 2005). To correlate the evolutionary relationship of fungal pathogenicity with these enzymes, we classified and compared the protein families involved in the hydrolysis of host constituents including the proteases, CAZy enzymes, cutinases, and lipases (table 1 and supplementary tables S5–S11, Supplementary Material online). The results showed that IF had the highest number of secreted proteases (average 109) followed by PF (88) and MF (54). PCA separated the three groups of fungi into different clusters (supplementary fig. S2B, Supplementary Material online), a feature of convergent evolution. Unexpectedly, our PIC test indicated that secreted proteases were not related to fungal pathotype formation ($F = 3.045$; $P = 0.909$). Previous studies have established that aspartic proteinases (A01) are the essential virulence factors in mammalian pathogens (Schaller et al. 2005), whereas subtilisin-like proteases (S08) are critical for IF pathogens (St. Leger and Wang 2010). These two subfamilies were highly represented in IF (A01 in average: IF:PF:MF = 16.9:11.4:4.6; S08: IF:PF:MF = 16:2.9:8.1) ($P < 0.01$). PIC tests also indicated that both A01 ($F = 0.0237$; $P = 0.8786$) and S08 ($F = 0.0599$; $P = 0.8083$) subfamilies were not related to determining fungal pathotypes against different hosts (supplementary table S5, Supplementary Material online). However, the metalloprotease subfamilies M14 ($F = 22.78$; $P = 4.115e-5$) and M28 ($F = 19.41$; $P = 0.0001$), serine peptidases S09 ($F = 27.7$; $P = 1.009e-5$) and S26 ($F = 4.458$; $P = 0.0429$), and the serine carboxypeptidase inhibitor I51 ($F = 11.47$; $P = 0.0019$) were significantly correlated to fungal pathogenicity against different hosts (supplementary table S5, Supplementary Material online), that is, the factors to contribute to fungal pathogenicities against insects, plants, and humans.

CAZy enzymes were well classified for each fungal species. Not surprisingly, plant pathogens had higher ($P < 0.05$) numbers of these family members than IF and PF for degradation of plant cell walls. On average, PF had 207.7 glycoside hydrolases (GHs), which was much higher than those in IF (128.3; $P = 3.8125e-6$) and MF (89.4; $P = 1.6147e-4$). Similarly, PF had 24 carbohydrate esterase (CE) enzymes but IF had only 9 whereas MF had 10 in average (table 1). The families of GHs ($F = 15.6$; $P = 0.0004$), glycosyl transferases (GTs, $F = 8.446$; $P = 0.0067$), CEs ($F = 14.44$; $P = 0.0006$), polysaccharide lyases (PL, $F = 17.37$; $P = 0.0002$), and carbohydrate-binding domain proteins ($F = 15.55$; $P = 0.004$) were significantly associated with pathogen lifestyle adaptation (table 1). PCA based on GH family sizes also indicated that, except for AAP (I12) and As. fumigatus (M1), the three fungal groups could be divided into different clusters (fig. 3D). Among the 66 GH families identified in this study, 19 were significantly ($P < 0.05$) correlated to fungal host adaptation, for example, the β-glucosidase GH3, cellobiohydrolase GH6, and β-xylosidases GH43 and GH51 (supplementary table S6, Supplementary Material online). Besides proteins, insect cuticles are rich in chitins, and chitinases (GH18) that assist fungal penetration of the insect cuticles are the virulence factors in insect pathogens (Ortiz-Urquiza and Keyhani 2013). Thus, IF had more GH18 (16.3) than PF (11.1) and MF (12.3). However, this family size had no correlation with fungal pathotypes against insects, plants, and mammals ($F = 0.0052$; $P = 0.943$) (supplementary table S6, Supplementary Material online).

The cutin polymers and pectin polysaccharides are unique plant cell wall components (Serrano et al. 2014). Therefore, PF (11) had more ($P < 0.001$) cutinases than IF (4) and MF (2), and this family size was highly associated with fungal host adaptation ($F = 23.09$; $P = 3.745e-5$) (table 1). Likewise, the pectin lyases PL1 and PL3 as well as the UDP (uridine 5′-diphosphate)-glucuronosyl transferase GT1, and acetyl xylan esterase CE1 were expanded in plant pathogens, and were also found to have correlation with fungal host adaptations ($P < 0.01$) (supplementary tables S7–S10, Supplementary Material online).

Previous studies revealed that lipases contribute to virulence in mammal (Schaller et al. 2005), plant (Blumke et al. 2014), and insect pathogenic fungi (Luo et al. 2014). Our genome survey indicated that plant pathogens had more lipases (59) than IF (25, $P = 0.0002$) and MF (31, $P = 2.527e-3$) (table 1), even though that insects and mammals may have more oily or fatty components than plants (Schaller et al. 2005). The varied numbers of lipases in different fungi had significant ($F = 17.18$; $P = 0.0002$) correlation with fungal pathotype formations (table 1 and supplementary table S11, Supplementary Material online). However, PCA based on the lipase subfamily sizes in each species could not well separate the three groups of fungi into different clusters (supplementary fig. S2C, Supplementary Material online), that is, the lipase family sizes divergently evolved during fungal speciation and host adaptation.

Pathotype Association with Other Protein Families

We also investigated the evolutionary association between fungal lifestyles and protein families involved in detoxification and secondary metabolisms (SMs). Both dehydrogenases and CYPs are known to be involved in multiple physiological processes including the degradation/detoxification of xenobiotics, biotransformations and the biosynthesis of bioactive metabolites in fungi (Christian et al. 2005; Cresnar and Petric 2011). Among the three fungal groups, PF had more ($P < 0.05$) dehydrogenases and CYPs than IF and MF (supplementary tables S12–S14, Supplementary Material online), that is, the

divergent evolution of these protein families among fungal pathotypes. In particular, alcohol dehydrogenases (EC 1.1.1.1; EC 1.1.1.2) were more highly ($P < 0.05$) represented in IF (average 44) and PF (59) than in MF (33). PIC tests indicated that the enzymes contribute to fungal host adaptation ($P < 0.05$) (supplementary table S12, Supplementary Material online). However, PCA failed to separate the three groups of fungi from each other based on the family sizes of dehydrogenases (supplementary fig. S2D, Supplementary Material online) and CYPs (supplementary fig. S2E, Supplementary Material online). Among the 295 subfamilies of CYPs identified in our study, 54 were found ($P < 0.05$) to have correlation with fungal lifestyle adaptation (supplementary table S13, Supplementary Material online). For example, the CYP505 clan, involved in the hydrolysis of fatty acids (Kitazume et al. 2000), was present in all three groups of fungi, and was linked to fungal pathogenicity against different hosts ($F = 9.016$; $P = 0.0053$). The CYP6001 subfamily is known to be involved in the production of oxylipins for signaling and SMs in fungi (Van Bogaert et al. 2011). On average, MF (2.1) had more CYP6001 than PF (1.5; $P = 0.0072$) and IF (0.9; $P = 19.9443e-6$), and the enzyme had significant contribution to fungal host adaptation ($F = 6.895$; $P = 0.0133$) (supplementary table S13, Supplementary Material online). The transporters were not linked to fungal host selection ($F = 0.9541$; $P = 0.3362$) (table 1). However, based on the PIC tests, 14 of 114 transporter subfamilies were found to be associated with fungal lifestyle adaptation (supplementary table S14, Supplementary Material online). For example, consistent with the protease expansion in IF (table 1 and supplementary table S5, Supplementary Material online), the oligopeptide transporters were overrepresented in IF (average 15.7) than in PF (6.8; $P = 1.9923e-5$) and MF (2.6; $P = 1.5058e-7$), and this subfamily had significant correlation with fungal lifestyles ($F = 10.41$; $P = 0.003$) (supplementary table S14, Supplementary Material online).

The contribution of secondary metabolites to fungal virulence has been evident in different fungal pathogens; for example, destruxins and oosporein in insect pathogens (Wang et al. 2012; Feng et al. 2015); deoxynivalenol in plant pathogens, such as the *Fusarium* spp. (Audenaert et al. 2013); and gliotoxin and siderophores in the human pathogen *As. fumigatus* (Abad et al. 2010). Both IF (35) and PF (36) had more ($P < 0.05$) SM core enzymes than MF (26) (table 1). It is not surprising that SM enzymes correlated well ($F = 4.208$; $P = 0.0488$) with fungal host adaptations (table 1 and supplementary table S15, Supplementary Material online). However, the PCA could not separate the three groups of fungi into different clusters based on the SM enzymes encoded in the examined fungi (supplementary fig. S2F, Supplementary Material online). Overall, both divergent and convergent evolutions of SM gene clusters were observed in different fungi. Among 725 identified gene clusters, 608 (84%) clusters were species-specific, that is, only present in any single species among the 33 examined fungi; 54 clusters were shared among two fungal species whereas only one nonribosomal peptide synthetase (NRPS) gene cluster, putatively involved in the biosynthesis of siderophores (Schrettl et al. 2007), was conserved in 25 species (supplementary table S16, Supplementary Material online). Interestingly, a phylogenetic tree generated based on the adenylation-domain sequences in these orthologous NRPSs was congruent with fungal speciation tree, for example, AAP_04092 of AAP was within the MF clade (fig. 5A). We also found the lineage-specific clusters. For example, there were nine clusters common to eight fungal species (supplementary table S16, Supplementary Material online). One NRPS cluster among these was highly conserved in AAP and seven mammalian pathogens (fig. 5B), which suggests a phylogenetically related origin of this cluster, for example, the inheritance from a common ancestor.

Substrate Utilization Tests for Insect Pathogens

To examine the association between the hydrolytic enzymes and ability to use different substrates, 12 insect pathogens and the plant pathogen, FGR were chosen for culturing on a minimal medium supplemented with casein hydrolysate, cellulose, pectin, xylose, sucrose, or glucose for 2 weeks. The insect pathogen, SPI grew slowly whereas FGR grew well on all examined nutrients (supplementary fig. S3, Supplementary Material online). Otherwise, consistent with the feature of protease expansion in IF (table 1), insect pathogens grew well on casein-supplemented medium. Unlike FGR with expanded CAZy enzymes, insect pathogens failed to utilize cellulose. Except for SPI, NOR, and AAL, the remaining insect fungi could grew well on pectin (supplementary fig. S3, Supplementary Material online). It was also found that NOR and CCM had lost the ability to grow on xylose. This might be due to the absence of a bacterial-like XpkA gene (Duan et al. 2009) in CCM (supplementary table S17, Supplementary Material online). It is unclear why NOR grew poorly on xylose. Unexpectedly, the AAP with a highly reduced genome and protein family sizes could grow similarly well on the examined nutrients except for cellulose. Genome-wide analysis using the program CAFE indicated that substantial gene losses occurred during the speciation of SPI and AAP, and both species also had higher numbers of orphan genes (supplementary fig. S4, Supplementary Material online). Overall, the data indicated again that both convergent and divergent evolutions of protein families occurred in insect pathogens, which resulted in the similar or altered abilities in these fungi to utilize different substrates.

Discussion

Fungal pathogens have a great impact on agriculture, human health, and invertebrate population density. Large efforts have been put forth to understand the determinants or virulence
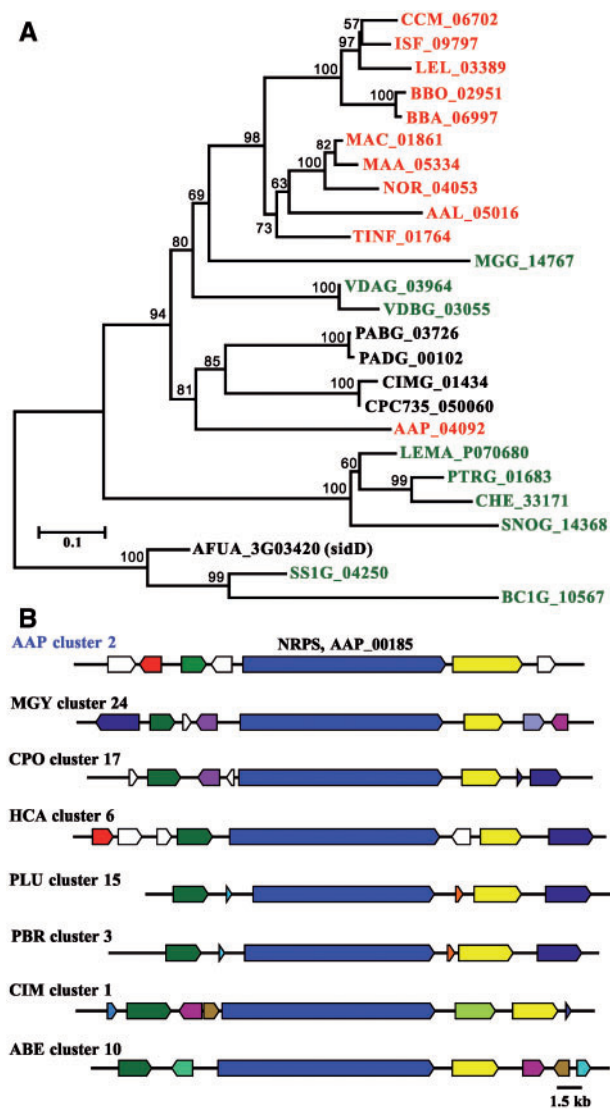
Fig. 5.—Conserved and lineage-specific evolution of SMs in different fungi. (*A*) Phylogeny of a highly conserved NRPS putatively involved in siderophore biosynthesis in 25 fungal species. The adenylation domain in the NRPS protein of each species was retrieved and used to generate the maximum-likelihood tree, which revealed a relatively congruent relationship with the fungal speciation tree. The proteins from IF are labeled in red, from PF in green, and from MF in black. (*B*) A highly conserved NRPS gene cluster is present in the bee pathogen, *A. apis* and seven mammalian pathogens. Orthologous genes are indicated by the same color. Genes with blank labels are hypothetical.

factors in fungal pathogens (Casadevall 2007). It is unknown whether the virulence factors identified thus far contribute to the diversity of fungal pathotypes against different hosts. In this study, we sequenced the genomes of seven insect pathogenic fungi, and, in total, analyzed three groups of 33 filamentous ascomycetous fungi that are the representative pathogens of insects, plants, and humans. We found that

the pathogenicity of either pathotype evolved multiple times (fig. 1; supplementary fig. S1, Supplementary Material online). Insect and plant pathogens are more closely related to each other than to human pathogens, which is consistent with our previous observations (Xiao et al. 2012; Hu et al. 2014). An exception to this norm was the honeybee pathogen AAP, which clustered within the MF clade. Overall, the pathogens that infect the same group of hosts have similar features of different protein family sizes, a clear indication of convergent evolution. Consistently, the analysis of 18 Dothideomycetes plant pathogenic fungi also found a similar size of core genes in different species, which, however, differ dramatically from each other in genome size due to the repetitive content variations (Ohm et al. 2012). Likewise, the diversified mycorrhizal fungi have a similarly reduced set of plant cell wall-degrading enzymes compared with plant pathogens (Kohler et al. 2015).

Despite the observations of the clear pattern of host-adaptive convergent evolution in different pathogens, divergent evolution also occurred within fungal pathotypes. For example, in contrast to other insect pathogens, AAP that causes the chalkbrood disease has a highly reduced genome size and gene-coding capacity (supplementary table 1, Supplementary Material online). This insect pathogen infects honeybee larvae after spore ingestion and germination in larval gut instead of cuticle penetration (Vojvodic et al. 2011), which is the preferred mode of host entry by other insect pathogens (St Leger and Wang 2010; Wang and Feng 2014). Such adaptation likely led to the highly reduced coding capacity and much smaller protein family sizes in AAP when compared with other fungi. Consistent with previous analyses (Goodwin et al. 2011; Stukenbrock et al. 2011), the wheat pathogen MGR also has highly reduced protein family sizes compared with other plant pathogens. For example, MGR has 147 GH proteins but other pathogens have 213 in average (supplementary table S6, Supplementary Material online). Comparative genomic analyses of three plant pathogens *Fusarium* spp. indicated that the lineage-specific acquisition of genomic regions, including four entire chromosomes, in *F. oxysporum* enabled the fungus to infect a broader range of hosts than FGR and *F. verticillioides* (Ma et al. 2010). The dermatophytes *Arthroderma benhamiae*, *Trichophyton rubrum*, and *Microsporum gypseum* also have specifically expanded or contracted protein families compared with their counterparts (Burmester et al. 2011; Martinez et al. 2012). For example, on average, these dermatophytes had 74 secreted proteases versus 44 in other fungi such as *As. fumigatus* and *Coccidioides immitis*, and so forth that can cause systematic infections (supplementary table S5, Supplementary Material online). Substrate utilization tests revealed the growth feature differences among insect pathogens (supplementary fig. S3, Supplementary Material online). In particular, although NOR has been reclassified as *Metarhizium rileyi* (Kepler et al. 2014), the growth of NOR was poor on pectin, xylose and even sucrose, which was in

contrast to *M. acridum* and *M. robertsii*. We also found that the virulence factors of serine proteases (S08 family) and chitinases (GH18) in insect pathogens and aspartic proteases (A01) in mammalian fungi were unrelated to fungal pathotype formation (supplementary table S5, Supplementary Material online), which suggests that the virulence gene of a pathogen may not necessarily contribute to fungal pathotype formation. Overall, host-adaptive convergent evolution is therefore coupled with divergent evolutions, which results in the altered protein family features within or between fungal pathotypic groups. Nevertheless, the role of phylogenetic relatedness on fungal pathotype formation could not be ruled out as the gene inventory and presence or absence of virulence factors could be affected by common ancestry (Sanchez-Rodriguez et al. 2010).

There are additional filamentous ascomycete pathogens of plants, insects, and humans have not been included in this study. For example, genome analysis of the rice false smut causative pathogen *Ustilaginoidea virens* (Clavicipitaceae) indicated that the fungus is closely related to insect pathogens *Metarhizium* spp. However, its genome encodes fewer numbers of proteases compared with insect pathogens (Zhang et al. 2014), an example of evolutionary host adaptation opposed to phylogenetic relatedness. Genome analyses of the insect pathogens *Hirsutella thompsonii* (Ophiocordycipitaceae) and *Ophiocordyceps unilateralis* indicated that the proteases and chitinases for insect cuticle degradation are expanded in these genomes similar to other insect pathogens (Agrawal et al. 2015; de Bekker et al. 2015; Wichadakul et al. 2015). The *Penicillium* genus (Eurotiomycetes, Ascomycota) contains array of species that are the plant pathogens (e.g., *P. oxalicum*, *P. expansum*, and *P. griseofulvum*) and human pathogen (e.g., *P. marneffei*). Interestingly, the human pathogen *P. marneffei*, close to *As. fumigatus*, diverged ahead the plant pathogens and its genome lacks the secondary metabolic toxin gene clusters that are well conserved in the plant pathogenic *Penicillium* spp. (Banani et al. 2016). The causative life-threatening ascomycete human pathogens *Penumocystis* spp. belong to the class Pneumocystidomycetes (subphylum Taphrinomycotina). The complex surface glycoproteins are particularly expanded in these pathogens as well as the enriched metalloproteases like in other human pathogens (Ma et al. 2016). Thus, these data would support the observations in this study that each pathotype evolved multiple times and both divergent and convergent evolutions occurred during fungal pathotype formation. Further analyses are still required to include the fungal pathogens from other phyla to provide a broader view of the evolution of fungal pathogenicity beyond this study.

To cope with a broad host range, fungal species have evolved more proteins and protein families than the insect pathogens with a narrow host range (Hu et al. 2014). We found that MF in general encoded fewer proteins and protein families followed by IF and then PF (table 1). Theoretically,

mammals have evolved more complicated immune defense systems, which include both innate and adaptive immunities. The latter is absent in insects and plants. In addition, humans as well as the insects employ additional behavioral or social immunities (e.g., social withdrawal strategies and death in isolation) to combat diseases such as fungal infections (Shang et al. 2015). In this regard, plants would be more vulnerable than insects and mammals to fight fungal diseases. Intriguingly, PF instead evolved more proteins and protein families than IF and MF, thereby the sophisticated mechanisms to fight infectious diseases using attack perception, adaptive responses, and building protective cell walls (Dangl and Jones 2001). During the coevolutionary arms race, plant pathogens also evolved highly divergent effectors to counteract host resistances and enzymes to pierce plant cell walls (Schmidt and Panstruga 2011), which is consistent with our observations in this study, for example, the effector and CAzy enzyme expansions in plant pathogens (table 1).

Up to date, huge advances have been made to illustrate the gene-for-gene interactions between the plants and their pathogens including bacteria, fungi, and oomycetes (Giraldo and Valent 2013; Oliveira-Garcia and Valent 2015). A similar evolutionary relationship has also been postulated for plant–insect interactions (Kaloshian 2004). However, the theory has not been well established for fungus–animal interactions. On the one hand, studies on insect and human immune systems have identified receptors, signaling components, and pathways that counteract fungal infections (Gottar et al. 2006; Iwasaki and Medzhitov 2015). On the other hand, proteins or even small molecules have been identified in IF and MF to evade host immunities (Wang and St Leger 2006; Schrettl et al. 2007; Wang et al. 2012; Feng et al. 2015). We identified many effector-like proteins in IF and MF (table 1; fig. 4C), which suggests that the gene-for-gene relationships are highly likely to be present in fungi and their animal hosts. Future functional studies on pathology and host immunology in fungal–animal systems would help establish these associations.

In this study, we identified a number of protein families contributing to fungal pathotype formation. However, it is still unclear how, why, and when a fungal species emerged as a pathogen of alternate hosts. We have found the existence of transitional species with an intermediate host range to bridge the evolutionary trajectory of *Metarhizium* lineages evolving from specialist to generalist (Hu et al. 2014). It is unclear whether the transitional species is present or not for the emergence of a pathogen through cross-kingdom or cross-phylum host-jumping. The closely related *Sporothrix* species either is an insect pathogen (SPI) or are human pathogens (e.g., SCH and SBR). In addition, they are also closely related to plant pathogens. Thus, these species could be good models for future investigation of the patterns and driving force of fungal host switch to form different pathotypes in the future.

In conclusion, we performed de novo genome assemblies, phylogenetic and evolutionary correlation analyses of filamentous ascomycete fungal pathogens. Our results showed that the pathogenicity of host-adapted fungi evolved multiple times. To adapt to different hosts, divergent evolution has coupled with convergent evolution to drive protein family expansions or contractions. Many protein families but not always the virulence factors were identified to contribute to fungal host selection and adaptation. These findings advance our understanding of the evolution of fungal pathogenicity and the determinants that contribute to fungal pathotype formation. Future studies are required to unravel the biological functions of different pathotype determinants and elucidate the forces that drive pathogenic fungal speciation.

Supplementary Material

Supplementary figures S1–S4 and tables S1–S17 are available at *Genome Biology and Evolution* online (http://www.gbe.oxfordjournals.org/).

Acknowledgments

Literature Cited

Abad A, et al. 2010. What makes *Aspergillus fumigatus* a successful pathogen? Genes and molecules involved in invasive aspergillosis. Rev Iberoam Micol. 27:155–182.

Agrawal Y, Khatri I, Subramanian S, Shenoy BD. 2015. Genome sequence, comparative analysis, and evolutionary insights into chitinases of entomopathogenic fungus *Hirsutella thompsonii*. Genome Biol Evol. 7:916–930.

Audenaert K, Vanheule A, Hofte M, Haesaert G. 2013. Deoxynivalenol: a major player in the multifaceted response of fusarium to its environment. Toxins 6:1–19.

Banani H, et al. 2016. Genome sequencing and secondary metabolism of the postharvest pathogen *Penicillium griseofulvum*. BMC Genomics 17:19.

Blin K, et al. 2013. Antismash 2.0–a versatile platform for genome mining of secondary metabolite producers. Nucleic Acids Res. 41:W204–W212.

Blumke A, et al. 2014. Secreted fungal effector lipase releases free fatty acids to inhibit innate immunity-related callose formation during wheat head infection. Plant Physiol. 165:346–358.

Burmester A, et al. 2011. Comparative and functional genomics provide insights into the pathogenicity of dermatophytic fungi. Genome Biol. 12:R7.

Bushley KE, et al. 2013. The genome of *Tolypocladium inflatum*: evolution, organization, and expression of the cyclosporin biosynthetic gene cluster. PLoS Genet. 9:e1003496.

Casadevall A. 2007. Determinants of virulence in the pathogenic fungi. Fungal Biol Rev. 21:130–132.

Christian V, Shrivastava R, Shukla D, Modi HA, Vyas BR. 2005. Degradation of xenobiotic compounds by lignin-degrading white-rot fungi: enzymology and mechanisms involved. Indian J Exp Biol. 43:301–312.

Coleman JJ, Muhammed M, Kasperkovitz PV, Vyas JM, Mylonakis E. 2011. *Fusarium* pathogenesis investigated using *Galleria mellonella* as a heterologous host. Fungal Biol. 115:1279–1289.

Clark AG, et al. 2007. Evolution of genes and genomes on the drosophila phylogeny. Nature 450:203–218.

Cresnar B, Petric S. 2011. Cytochrome P450 enzymes in the fungal kingdom. Biochim Biophys Acta. 1814:29–35.

Dangl JL, Jones JDG. 2001. Plant pathogens and integrated defence responses to infection. Nature 411:826–833.

de Bekker C, et al. 2015. Gene expression during zombie ant biting behavior reflects the complexity underlying fungal parasitic behavioral manipulation. BMC Genomics 16:620.

de Bie T, Cristianini N, Demuth JP, Hahn MW. 2006. CAFE: a computational tool for the study of gene family evolution. Bioinformatics 22:1269–1271.

de Groot PWJ, Bader O, de Boer AD, Weig M, Chauhan N. 2013. Adhesins in human fungal pathogens: glue with plenty of stick. Eukaryot Cell 12:470–481.

Duan ZB, Shang YF, Gao Q, Zheng P, Wang CS. 2009. A phosphoketolase Mpk1 of bacterial origin is adaptively required for full virulence in the insect-pathogenic fungus *Metarhizium anisopliae*. Environ Microbiol. 11:2351–2360.

Efron B. 2007. Size, power and false discovery rates. Ann Stat. 35:1351–1377.

Edgar RC. 2004. Muscle: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32:1792–1797.

Felsenstein J. 1985. Phylogenies and the comparative method. Am Nat. 1:1–15.

Feng P, Shang Y, Cen K, Wang C. 2015. Fungal biosynthesis of the bibenzoquinone oosporein to evade insect immunity. Proc Natl Acad Sci U S A. 112:11365–11370.

Floudas D, et al. 2012. The paleozoic origin of enzymatic lignin decomposition reconstructed from 31 fungal genomes. Science 336:1715–1719.

Gao Q, et al. 2011. Genome sequencing and comparative transcriptomics of the model entomopathogenic fungi *Metarhizium anisopliae* and *M. acridum*. PLoS Genet. 7:e1001264.

Giraldo MC, Valent B. 2013. Filamentous plant pathogen effectors in action. Nat Rev Microbiol. 11:800–814.

Goodwin SB, et al. 2011. Finished genome of the fungal wheat pathogen *Mycosphaerella graminicola* reveals dispensome structure, chromosome plasticity, and stealth pathogenesis. PLoS Genet. 7:e1002070.

Gottar M, et al. 2006. Dual detection of fungal infections in *Drosophila* via recognition of glucans and sensing of virulence factors. Cell 127:1425–1437.

Haas BJ, et al. 2008. Automated eukaryotic gene structure annotation using Evidencemodeler and the program to assemble spliced alignments. Genome Biol. 9:R7.

Hu X, et al. 2013. Genome survey uncovers the secrets of sex and lifestyle in caterpillar fungus. Chin Sci Bull. 58:2846–2854.

Hu X, et al. 2014. Trajectory and genomic determinants of fungal-pathogen speciation and host adaptation. Proc Natl Acad Sci U S A. 111:16796–16801.

Iwasaki A, Medzhitov R. 2015. Control of adaptive immunity by the innate immune system. Nat Immunol. 16:343–353.

James TY, et al. 2006. Reconstructing the early evolution of Fungi using a six-gene phylogeny. Nature 443:818–822.

Jones P, et al. 2014. Interproscan 5: genome-scale protein function classification. Bioinformatics 30:1236–1240.

Kaloshian I. 2004. Gene-for-gene disease resistance: bridging insect pest and pathogen defense. J Chem Ecol. 30:2419–2438.

Kepler RM, Humber RA, Bischoff JF, Rehner SA. 2014. Clarification of generic and species boundaries for *Metarhizium*. and related fungi through multigene phylogenetics. Mycologia 106:811–829.

Kitazume T, Takaya N, Nakayama N, Shoun H. 2000. *Fusarium oxysporum* fatty-acid subterminal hydroxylase (cyp505) is a membrane-bound eukaryotic: counterpart of bacillus megaterium cytochrome P450bm3. J Biol Chem. 275:39734–39740.

Kohler A, et al. 2015. Convergent losses of decay mechanisms and rapid turnover of symbiosis genes in mycorrhizal mutualists. Nat Genet. 47:410–415.

Luo ZB, Qin YQ, Pei Y, Keyhani NO. 2014. Ablation of the crea regulator results in amino acid toxicity, temperature sensitivity, pleiotropic effects on cellular development and loss of virulence in the filamentous fungus *Beauveria bassiana*. Environ Microbiol. 16:1122–1136.

Ma L, et al. 2016. Genome analysis of three *Pneumocystis* species reveals adaptation mechanisms to life exclusively in mammalian hosts. Nat Commun. 7:10740.

Ma LJ, et al. 2010. Comparative genomics reveals mobile pathogenicity chromosomes in *Fusarium*. Nature 464:367–373.

Martinez DA, et al. 2012. Comparative genome analysis of *Trichophyton rubrum* and related dermatophytes reveals candidate genes involved in infection. MBIO 3:e00259–e00253.

Ohm RA, et al. 2012. Diverse lifestyles and strategies of plant pathogenesis encoded in the genomes of eighteen Dothideomycetes fungi. PLoS Pathog. 8:e1003037.

Oliveira-Garcia E, Valent B. 2015. How eukaryotic filamentous pathogens evade plant recognition. Curr Opin Microbiol. 26:92–101.

Ortiz-Urquiza A, Keyhani NO. 2013. Action on the surface: entomopathogenic fungi versus the insect cuticle. Insects 4:357–374.

Parra G, Bradnam K, Korf I. 2007. CEGMA: a pipeline to accurately annotate core genes in eukaryotic genomes. Bioinformatics 23:1061–1067.

Pele J, Becu JM, Abdi H, Chabbert M. 2012. Bios2mds: an R package for comparing orthologous protein families by metric multidimensional scaling. BMC Bioinformatics 13:133.

Perez-Nadales E, et al. 2014. Fungal model systems and the elucidation of pathogenicity determinants. Fungal Genet Biol. 70:42–67.

Plotree D, Plotgram D. 1989. Phylip-phylogeny inference package (version 3.2). Cladistics 5:163–166.

Salah H, et al. 2015. Phylogenetic diversity of human pathogenic *Fusarium* and emergence of uncommon virulent species. J Infect. 71:658–666.

Sanchez-Rodriguez A, Martens C, Engelen K, Van de Peer Y, Marchal K. 2010. The potential for pathogenicity was present in the ancestor of the Ascomycete subphylum Pezizomycotina. BMC Evol Biol. 10:318.

Schaible UE, Kaufmann SHE. 2005. A nutritive view on the host-pathogen interplay. Trends Microbiol. 13:373–380.

Schaller M, Borelli C, Korting HC, Hube B. 2005. Hydrolytic enzymes as virulence factors of *Candida albicans*. Mycoses 48:365–377.

Schmidt HA, Strimmer K, Vingron M, von Haeseler A. 2002. TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. Bioinformatics 18:502–504.

Schmidt SM, Panstruga R. 2011. Pathogenomics of fungal plant parasites: what have we learnt about pathogenesis? Curr Opin Plant Biol. 14:392–399.

Schoch CL, et al. 2009. The Ascomycota tree of life: a phylum-wide phylogeny clarifies the origin and evolution of fundamental reproductive and ecological traits. Syst Biol. 58:224–239.

Schrettl M, et al. 2007. Distinct roles for intra- and extracellular siderophores during *Aspergillus fumigatus* infection. PLoS Pathog. 3:e125.

Serrano M, Coluccia F, Torres M, L'Haridon F, Metraux JP. 2014. The cuticle and plant defense to pathogens. Front Plant Sci. 5:274.

Shang Y, Feng P, Wang C. 2015. Fungi that infect insects: altering host behavior and beyond. PLoS Pathog. 11:e1005037.

St Leger RJ, Wang C. 2010. Genetic engineering of fungal biocontrol agents to achieve greater efficacy against insect pests. Appl Microbiol Biotechnol. 85:901–907.

Stukenbrock EH, et al. 2011. The making of a new pathogen: Insights from comparative population genomics of the domesticated wheat pathogen *Mycosphaerella graminicola* and its wild sister species. Genome Res. 21:2157–2166.

Takaoka S, et al. 2014. Complex regulation of secondary metabolism controlling pathogenicity in the phytopathogenic fungus *Alternaria alternata*. New Phytol. 202:1297–1309.

Talavera G, Castresana J. 2007. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. Syst Biol. 56:564–577.

Teixeira MM, et al. 2014. Comparative genomics of the major fungal agents of human and animal Sporotrichosis: *Sporothrix schenckii* and *Sporothrix brasiliensis*. BMC Genomics 15:943.

Urban M, et al. 2015. The pathogen-host interactions database (PHI-base): additions and future developments. Nucleic Acids Res. 43:D645–D655.

Van Bogaert INA, Groeneboer S, Saerens K, Soetaert W. 2011. The role of cytochrome p450 monooxygenases in microbial fatty acid metabolism. FEBS J. 278:206–221.

van de Wouw AP, Howlett BJ. 2011. Fungal pathogenicity genes in the age of 'omics'. Mol Plant Pathol. 12:507–514.

Vojvodic S, Jensen AB, Markussen B, Eilenberg J, Boomsma JJ. 2011. Genetic variation in virulence among chalkbrood strains infecting honeybees. PLoS One 6:e25035.

Wang C, Feng MG. 2014. Advances in fundamental and applied studies in china of fungal biocontrol agents for use against arthropod pests. Biol Control. 68:129–135.

Wang B, Kang Q, Lu Y, Bai L, Wang C. 2012. Unveiling the biosynthetic puzzle of destruxins in *Metarhizium* species. Proc Natl Acad Sci U S A. 109:1287–1292.

Wang C, St Leger RJ. 2006. A collagenous protective coat enables *Metarhizium anisopliae* to evade insect immune responses. Proc Natl Acad Sci U S A. 103:6647–6652.

Wichadakul D, et al. 2015. Insights from the genome of *Ophiocordyceps polyrhachis-furcata* to pathogenicity and host specificity in insect fungi. BMC Genomics 16:881.

Xiao G, et al. 2012. Genomic perspectives on the evolution of fungal entomopathogenicity in *Beauveria bassiana*. Sci Rep. 2:483.

Xue CY, Hsueh YP, Heitman J. 2008. Magnificent seven: roles of G protein-coupled receptors in extracellular sensing in fungi. FEMS Microbiol Rev. 32:1010–1032.

Ying SH, Feng MG, Keyhani NO. 2013. A carbon responsive g-protein coupled receptor modulates broad developmental and genetic networks in the entomopathogenic fungus, *Beauveria bassiana*. Environ Microbiol. 15:2902–2921.

Zhang Y, et al. 2014. Specific adaptation of *Ustilaginoidea virens* in occupying host florets revealed by comparative and functional genomics. Nat Commun. 5:3849.

Zheng P, et al. 2011. Genome sequence of the insect pathogenic fungus *Cordyceps militaris*, a valued traditional Chinese medicine. Genome Biol. 12:R116.

Zheng P, Xia YL, Zhang SW, Wang CS. 2013. Genetics of *Cordyceps* and related fungi. Appl Microbiol Biotechnol. 97:2797–2804.

Associate editor: Ruth Hershberg