Genome Sequence, Comparative Analysis, and Evolutionary Insights into Chitinases of Entomopathogenic Fungus *Hirsutella thompsonii*

Yamini Agrawal[1], Indu Khatri[1], Srikrishna Subramanian[1,*], and Belle Damodara Shenoy[1,2,*]

[1]CSIR-Institute of Microbial Technology, Chandigarh, India

[2]Present address: CSIR-National Institute of Oceanography, Dona Paula, Goa, India

*Corresponding author: E-mail: krishna@imtech.res.in; shenoynio@gmail.com.

Accepted: February 20, 2015

Data deposition: KM220955–KM221027, KM504392. (Chitinase sequences) Genome deposition: APKB00000000, APKU00000000.

Abstract

Hirsutella thompsonii (*Ht*) is a fungal pathogen of acarines and the primary cause of epizootics among mites. The draft genomes of two isolates of *Ht* (MTCC 3556: *Ht3*, 34.6 Mb and MTCC 6686: *Ht6,* 34.7 Mb) are presented and compared with the genomes of *Beauveria bassiana* (*Bb*) ARSEF 2860 and *Ophiocordyceps sinensis* (*Os*) CO18. Comparative analysis of carbohydrate active enzymes, pathogen–host interaction genes, metabolism-associated genes, and genes involved in biosynthesis of secondary metabolites in the four genomes was carried out. Reduction in gene family sizes in *Ht3* and *Os* as compared with *Ht6* and *Bb* is observed. Analysis of the mating type genes in *Ht* reveals the presence of MAT idiomorphs which is suggestive of cryptic sexual traits in *Ht*. We further identify and classify putative chitinases that may function as virulence factors in fungal entomopathogens due to their role in degradation of arthropod cuticle.

Key words: glycoside hydrolase family 18, mycoinsecticide, pathogen–host interactions, phylogenetic classification, subgroups A, B, and C.

Introduction

Hirsutella thompsonii (*Ht*) (*Ophiocordycipitaceae*, *Hypocreales,* and *Sordariomycetes*) is a pathogenic fungal species that infects arthropod hosts and causes natural epizootics in populations of blueberry bud mites (*Acalitus vaccinii*), citrus rust mites (*Phyllocoptruta oleivora*), and coconut fruit mites (*Aceria guerreronis*) (Samson et al. 1980; Weibelzahl and Liburd 2009). It produces an insecticidal protein Hirsutellin A, which is ribotoxic to a variety of pest populations (Herrero-Galán et al. 2008, 2013). The insecticidal property of *Ht* has been exploited, for example, spore suspension of *Ht* is used against *Varroa* mites and promoted under the trade name *Mycar*, but this product has limited commercial success due to efficacy issues (Haragsim and Ruzicka 2001; Weinzierl et al. 2005). The genus *Hirsutella* includes some other well-known insect-pathogenic fungi, such as *Hirsutella rhossiliensis* and *Hirsutella sinensis* (Sung et al. 2007). A range of bioactive secondary metabolites have been reported in *Hirsutella* species, such as linear and cyclic nonribosomal peptides: Hirsutellic acid A, hirsutide, hirsutatins A and B; polyketide synthase–nonribosomal peptide synthase (PKS–NRPS) hybrid metabolites: Cytostatic cytochalasin Q (Molnar et al. 2010). Some strains of *Hirsutella* are also used in traditional Chinese medicines and neutraceutical preparations (Paterson 2008; Molnar et al. 2010).

The dynamics of insect–fungal interactions is complex and has critical implications in biocontrol of insect pests. Recent genome-based studies are providing interesting insights into genetic and pathogenic mechanisms involved in insect–fungal interactions (Gao et al. 2011; Kubicek et al. 2011; Zheng et al. 2011; Xiao et al. 2012; Bushley, Raja, et al. 2013; Wang et al. 2014). For example, phylogenomic analyses of insect-pathogenic fungi *Beauveria bassiana* (*Bb*) (Xiao et al. 2012), *Cordyceps militaris* (Zheng et al. 2011), and *Metarhizium anisopliae* (Gao et al. 2011) indicate similarity in evolutionary expansion of gene families such as chitinases, lipases, proteases, etc. within insect-pathogenic fungi that could be attributed to their insect killing strategies and host ranges. Such

genome-wide studies are the leading steps for comprehensive understanding of insect–fungal associations in the context of fungal evolution and ecology. This study was thus initiated to decipher the genome sequence of *Ht* and compare it with two phylogenetically closely related species within the fungal order *Hypocreales*: *Ophiocordyceps sinensis* (*Os*) (*Ophiocordycipitaceae*) (Liu et al. 2002; Hu et al. 2013), and *Bb* (*Cordycipitaceae*), an effective mycoinsecticide (Xiao et al. 2012), so as to reveal similarities and differences in their genetic make-up and its possible implications in insect–fungal interactions and pathogenicity.

Insect-pathogenic fungi disintegrate the chitin-rich insect cuticle by the action of a variety of chitinases (Charnley 2003). Fungal chitinases (EC 3.2.1.14) have been exploited in the development of biocontrol strategies against insects, fungi, or nematodes (Seidl 2008; St. Leger and Wang 2010). Pathogenic fungi contain various chitin-degrading enzymes with an average of 10–20 different chitinases, where chitinases play diverse biological roles, such as cell-wall remodeling, nutrient scavenging and aggressive functions (Seidl 2008; Gao et al. 2011; Hartl et al. 2012; Hamid et al. 2013). Genome surveys of chitinases from pathogenic fungi suggest high variability in gene size and modular structure of chitinases that could be related to their diverse lifestyle pattern (Seidl et al. 2005; Karlsson and Stenlid 2009; Hartl et al. 2012). However, there has been no report on diversity and phylogeny of chitinases from *Hirsutella* species. Our study focuses on the evolutionary relationships of *Bb, Ht,* and *Os* chitinases.

Materials and Methods

Fungal Isolates and Maintenance

Hirsutella thompsonii isolates MTCC 3556 (*Ht3*) and MTCC 6686 (*Ht6*) were accessed from the Microbial Type Culture Collection (MTCC), located at CSIR-Institute of Microbial Technology, Chandigarh, India. The *Ht3* strain, isolated from coconut perianth mite (*Aceria guerreronis*), was deposited in MTCC by S. Ramarethinam from J. Stanes and Co. Ltd., Coimbatore, Tamil Nadu. The *Ht6* strain, isolated from red spider mite, was deposited in MTCC by S. Debnath of the Tea Research Association, P.O. Cinnamara, Jorhat, Assam.

The fungal cultures were grown on potato-dextrose agar medium and incubated at 25 °C for 6–8 days. Glycerol stocks were prepared from fresh mycelia in 10% glycerol and stored at −70 °C for further use. Genomic DNA was isolated from fresh mycelia using ZR Fungal/Bacterial DNA kit (Zymo Research, catalog number D6005) following manufacturer's protocol. The concentration and quality of DNA extracted were assessed by Nanodrop Spectrophotometer ND-1000 (Thermo Scientific, Wilmington) and 0.8% agarose gel electrophoresis, respectively.

Morphological Characterization of *Ht3* and *Ht6*

Fungal morphological features were observed under Olympus BX51 microscope. The microscopic images were captured using Olympus XC10 digital camera and analyzed using CellB image analysis software (Olympus, Tokyo, Japan).

Phylogenetic Position of *Ht3* and *Ht6*

The nuclear ribosomal small and large subunits (*nrSSU* and *nrLSU*), elongation factor 1α (*tef1*), the largest subunit of RNA polymerase II (*rpb1*), and β-tubulin (*tub*) genes were retrieved from the *Ht3* and *Ht6* genomes. The gene sequences were included in the concatenated five gene data set of hypocrealean fungi (Sung et al. 2007). Maximum parsimony analysis was performed in PAUP v4.0 (Swofford 2003). Ambiguously aligned regions were excluded from the analysis. Gaps were treated as missing data. Trees were inferred using the heuristic search option with 1,000 random sequence additions.

Genome Sequencing, Assembly, and Annotation

Paired-end (PE) (insert size of 350 bp and length 101 bp) shotgun sequencing of *Ht* genomes was performed using Illumina HiSeq 1000 technology at the Centre for Cellular and Molecular Platforms (C-CAMP) Bangalore, India. Mate-pair (MP) (insert size of 2 kb and length 101 bp) information was also obtained for *Ht3*. De novo assembly was performed using CLCbio wb6.0 (http://www.clcbio.com/, last accessed March 10, 2015). The data sets were scanned and filtered for the presence of any reads from bacterial DNA that may be present as contaminants. The DNA of *Ht6* was found to likely be contaminated by some of the bacterial strains, mainly belonging to the genera, *Bordetella*, *Burkholderia*, *Escherichia*, and *Phyllobacterium*. The genomes of all available species of these bacteria were retrieved from National Centre for Biotechnological Information (NCBI) and the reads of *Ht6* were aligned to these genomes. The contigs obtained were also subjected to BLAST against nucleotide collection database at NCBI, and the reads that showed significant matches to only bacterial sequences but no hits to fungal sequences were discarded in the final assembly. The refined contigs obtained were scaffolded using SSPACE v2.0 (Boetzer et al. 2011) and gaps were filled using GapFiller v1.10 (Boetzer and Pirovano 2012).

Genome annotation was performed using the MAKER pipeline v2.0 with Augustus as the gene predictor (Stanke and Morgenstern 2005; Cantarel et al. 2008). *Bb* ARSEF 2860 (PRJNA38719) and *Os* CO18 (PRJNA59569) genomes were reannotated using the same methodology. Combinations of both the original and reannotated set of genes/proteins from all genomes were used for further computational analyses. The predicted proteins were subjected to BLASTp (Altschul et al. 1990) against the nonredundant protein sequences database at NCBI. The tRNAs in *Ht3* and *Ht6* were identified using the tRNAscan-SE v1.3 (Lowe and Eddy

1997). The draft genomes of *Ht3* and *Ht6* were deposited at NCBI-GenBank under the accession number APKB00000000 and APKU00000000, respectively.

Cluster of Orthologous Groups

In order to classify the proteins into orthologous groups of various cellular and metabolic functions, clusters of orthologous groups (COG) were generated by subjecting the predicted proteomes to BLASTp against the COG database (Tatusov et al. 2000) (http://www.ncbi.nlm.nih.gov/COG/, last accessed March 10, 2015) with an *E* value cutoff of $1e^{-5}$. For each COG mapping, functional assignment was provided based on COG general category letter associations. Orthologous proteins among all the four genomes were characterized by reciprocal best BLAST hits (*E* value cutoff of $1e^{-10}$) obtained using a PERL script of Proteinortho v2.0 (Lechner et al. 2011). Gene family expansions and contractions were estimated using the Computational Analysis of gene Family Evolution (CAFE) v3.0 (Han et al. 2013). Molecular Evolutionary Genetics Analysis (MEGA) v6.0 (Tamura et al. 2013) was used to prepare Maximum Likelihood Newick tree that was used as an input for CAFE. Significant changes in gene numbers were analyzed at *P* value cutoff of 0.01.

Carbohydrate Active Enzymes

Structurally related catalytic enzymes belonging to classes, glycoside hydrolases (GHs), glycosyltransferases (GTs), polysaccharide lyases (PLs), carbohydrate esterases (CEs) and auxiliary activities (AAs), and associated carbohydrate-binding modules (CBMs) were searched in the predicted proteome of all the four genomes. The HMM profiles of CAZy database (Lombard et al. 2014) (http://www.cazy.org/, last accessed March 10, 2015) on dbCAN web server (Yin et al. 2012) were used to assign the CAZy families to the predicted protein sequences of *Bb, Ht3*, *Ht6,* and *Os*.

Protein Family Classification

The gene clusters and the proteins responsible for the biosynthesis of secondary metabolites were identified by subjecting the predicted proteome from all the four genomes to antiSMASH server (Medema et al. 2011). Pathogen–host interacting (PHI) partners were identified using PHI database (Winnenburg et al. 2006) with an *E* value cutoff of $1e^{-5}$. In order to identify ribonuclease (RNase) sequences from all the four genomes, "Ribonucleases fungi" was used as keyword search in NCBI and 817 RNase sequences were retrieved. These retrieved RNase sequences were searched against protein databases for all the four genomes using BLASTp at *E* value cutoff of $1e^{-5}$. Transporter Classification Database (Saier et al. 2014) was used to identify membrane transport proteins. The cytochrome P450 (CYP) superfamily was classified according to P450 classification database (Nelson 2009). The mating type genes for fruiting body development, karyogamy and meiosis, mating process and mating signaling in *Aspergillus nidulans*, *Metarhizium acridum, M. anisopliae*, and *Neurospora crassa* (Gao et al. 2011) were retrieved from NCBI and the predicted protein sequences of all the four genomes were subjected to BLASTp against these proteins with an *E* value cutoff of $1e^{-5}$.

Identification of MAT Idiomorphs

Sexual development in fungi is regulated through the mating type genes (Lee et al. 2010). The presence of mating type loci in the four genomes was assessed by sequence similarities to the corresponding MAT genes in other filamentous fungi (Gao et al. 2011). MAT1-1-1 idiomorphic region encodes for alpha domain, MAT1-1-2 for conserved Proline–Proline–Phenylalanine domain whereas MAT1-1-3 for High Mobility Group (HMG) domain with a DNA-binding site (Bushley, Li, et al. 2013; Klix et al. 2010). Likewise, MAT1-2-1 idiomorph encodes for MatA HMG box with a DNA-binding site (Bushley, Li, et al. 2013). The presence of these conserved domains was used to identify the mating type loci. Synteny among all the four genomes was analyzed for the scaffolds containing the mating type locus by locating the position of MAT locus and their flanking genes. The absence of any gene was rechecked through mapping back the reads on these genes.

Transposable Elements and Repeat-Induced Point Mutation Analysis

The transposable elements (TEs) were classified by subjecting genome sequences of *Bb, Ht3*, *Ht6,* and *Os* to BLASTn against the Repbase (http://girinst.org, last accessed March 10, 2015) libraries of RepeatMasker (http://www.repeatmasker.org, last accessed March 10, 2015). Repeat-induced point mutations (RIPs) were estimated using the RIPCAL v1.0 (Hane and Oliver 2008).

Phylogenetic Analysis of GH Family 18 Chitinases

Chitinase sequences were retrieved from the predicted proteomes of all the four genomes using ProfileScan (http://www.csd.hku.hk/bruhk/gcgdoc/profilescan.html, last accessed March 10, 2015) by identifying the conserved active site signature motif [LIVMF]-[DN]-G-[LIVMF]-[DN]-[LIVMF]-[DN]-x-E (Jablonowski et al. 2001) (Prosite No. PS01095). The retrieved chitinase sequences were aligned using PCMA (Pei et al. 2003) and a hidden Markov model profile was built using HMMER v3.1b1 (Finn et al. 2011). This HMMER profile hidden Markov model was further used to scan the four predicted proteomes to identify additional homologs of GH18 chitinases that may have a mutated active site motif. This list of GH18 chitinase sequences was found to be identical to what we retrieved through CAZy GH searches (supplementary table S1, Supplementary Material online) with the exception of one chitinase sequence (*Bb GH18-3*) that was observed to be shorter in length.

A total of 140 fungal GH18 chitinase sequences were selected from the chitinase data set reported previously (Karlsson and Stenlid 2009). The 74 chitinase sequences from the four genomes were classified and aligned with these fungal GH18 family chitinase sequences using MEGA. Three alignment files for subgroups A, B, and C fungal chitinases were prepared as previously described (Seidl et al. 2005; Karlsson and Stenlid 2009) (supplementary sets S1–S3, Supplementary Material online). Only the sequences of the catalytic domains were included for the phylogenetic classification (Saito et al. 2003). Each alignment was manually adjusted to allow maximum alignment of residues and minimize gaps. Maximum parsimony analysis was performed using PAUP.

Elucidation of Domains Architecture of GH Family 18 Chitinases

The domain architecture of the putative chitinases was obtained using the InterProScan, Pfam, and SMART databases (Schultz et al. 1998; Quevillon et al. 2005; Letunic et al. 2012; Punta et al. 2012). Also these sequences were analyzed to evaluate their theoretical pI and molecular weight using the Protparam tool (Gasteiger et al. 2005).

Results

General Features

Ht3 and *Ht6* were observed to have typical *Hirsutella*-morphologies (Kirk et al. 2008) (supplementary fig. S1, Supplementary Material online). In the multigene phylogenetic analysis, *Ht3* and *Ht6* were found to be clustered within *Ophiocordycipitaceae* (*Hypocreales*) (supplementary fig. S2, Supplementary Material online). The genome sequencing, assembly, and annotated features for *Ht3* and *Ht6* genomes are presented in table 1. About 75% of predicted proteins in *Ht3* and *Ht6* were found to have homologs in *Bb* ARSEF 2860 (*Cordycipitaceae*, *Hypocreales*), whereas about 44% of predicted proteins in *Ht3* and 38% in *Ht6* were found to have homologs with *M. anisopliae* ARSEF 23 (*Clavicipitaceae, Hypocreales*). Table 2 presents the highlights of the comparative analyses for the four genomes. A total of 10,714, 9,756, 10,259, and 9,130 predicted proteins were identified in *Bb*, *Ht3*, *Ht6,* and *Os* genomes, respectively. A comparative distribution of functional attributes of the predicted proteins is presented in supplementary figure S3*A*, Supplementary Material online. Reciprocal BLAST analysis identified 4,989 proteins as the core proteome of the four genomes contributing almost 50% of their distinct proteomic content (supplementary fig. S3*B*, Supplementary Material online). *Ht3* and *Ht6* were found to share 8,182 proteins, exhibiting maximum orthology as expected. The unique protein content of each genome varied in the order *Ht3* (15%) < *Ht6* (19%) < *Os* (36%) < *Bb* (42%) (supplementary fig. S3*B*, Supplementary Material online).

Table 1

Sequencing, Assembly, and Annotated Features for the *Ht* Genomes

General genome Features	*Ht3*	*Ht6*
Genome size	34.6 Mb	34.7 Mb
PE reads	58,288,788	63,296,052
High-quality reads	57,988,713	56,897,996
High-quality mapped reads	44,357,832	48,273,170
MP reads	192,572	—
N50 contigs	180,407	214,019
Scaffolds	474	981
Contigs	621	1,003
Coverage (mapped)	169× (121×)	165× (116×)
ORFs	9,756	10,259
Annotated ORFs	8,699	8,705
NCBI-GenBank accession no.	APKB00000000	APKU00000000

Survey of Carbohydrate Active Enzymes

A total of six classes: GHs (supplementary table S1, Supplementary Material online), GTs (supplementary table S2, Supplementary Material online), PLs (supplementary table S3, Supplementary Material online), CEs (supplementary table S4, Supplementary Material online), AAs (supplementary table S5, Supplementary Material online), and CBMs (supplementary table S6, Supplementary Material online) as provided in CAZy database were compared in this study. The significantly expanded/contracted gene families are listed in table 3. All the four genomes were found to contain genes putatively encoding enzymes of the GH16 (xyloglucan: xyloglucosyltransferase), GH18 (chitinase), and GH76 (α-1,6-mannanase) families. GT4 family enzyme identified as *N*-acetylglucosaminyltransferase (CAFE *P* value, 0.001) was observed in all four species, showing its expansion in *Ht3*, *Ht6,* and *Bb* genomes. GT41 was observed in *Ht6* and *Bb*. Further, *Ht6* and *Bb* were found to contain CBM66 (CAFE *P* value is 0.01) that is implicated in β-fructosidase activity, usually in bacteria (Cuskin et al. 2012). Both GT41 and CBM66 are expanded in *Ht6* and *Bb*. CBM18 and CBM50 (LysM) family modules implicated in chitin-binding function in association with chitinase catalytic domains were found to be predominantly present in all the four genomes. Moreover, CBM19 and CBM55 also implicated in chitin-binding functions were observed in *Bb* and *Os*, respectively.

PHI Genes

The virulence-associated genes that are likely to be involved in a variety of fungal–host interactions were identified in *Bb*, *Ht3*, *Ht6,* and *Os* (table 4 and supplementary table S7, Supplementary Material online). We identified 10.3%, 10.1%, 10.4%, and 8.2% of the 19,162 experimentally verified PHI genes (Winnenburg et al. 2006) involved in virulence from *Bb*, *Ht3*, *Ht6,* and *Os*, respectively. List of significantly expanded/contracted PHI families is provided in table 3.

Table 2
Major Highlights of the Comparative Analyses for the Four Genomes: *Ht3*, *Ht6*, *Os,* and *Bb*

Category	*Ht3*	*Ht6*	*Os*	*Bb*	Range
Total no. of predicted proteins	9,756	10,259	9,130	10,714	9,130–10,714
Total no. of GHs	149	149	108	186	108–186
Total no. of GTs	99	103	79	107	79–107
Total no. of PLs	1	2	2	3	1–3
Total no. of CEs	78	82	38	90	38–90
Total no. of AAs	58	69	28	66	28–69
Total no. of CBMs	25	31	33	60	25–60
No. of CBM1 proteins	0	0	0	4	0–4
No. of CBM18 proteins	5	7	6	10	5–10
No. of CBM50 (LysM) proteins	10	13	15	19	10–19
No. of PHI proteins	1,931	1,995	1,572	1,979	1,572–1,995
No. of membrane transporters	568	610	743	764	568–764
No. of G-protein coupled receptors	15	14	15	9	9–15
No. of RNase genes	24	25	10	18	10–25
No. of genes involved in biosynthesis of sec. metabolites	56	59	32	43	32–59
No. of CYPs	137	140	92	132	92–140
No. of chitinase sequences	19	15	17	23	15–23
Subgroup-A chitinases	12	9	9	5	5–12
Subgroup-B chitinases	4	3	3	11	3–11
Subgroup-C chitinases	3	3	5	7	3–7

Table 3
List of Statistically Significant Expanded/Contracted Gene Families

Family	*Ht3*	*Ht6*	*Os*	*Bb*	Gene Family Expansion
GH15	1	3	1	2	*Ht6*, *Bb*
GH18	19	15	17	22	*Ht3*, *Bb*
GH74	2	4	5	1	*Os*, *Ht6*
GT4	5	8	2	4	*Ht3*, *Ht6*, *Bb*
GT26	1	0	0	0	*Ht3*
GT41	0	2	0	1	*Ht6*, *Bb*
AA7	17	23	11	28	*Ht3*, *Ht6*, *Bb*
CBM66	0	2	0	5	*Ht6*, *Bb*
CoA_Ligase	22	29	9	16	*Ht3*, *Ht6*, *Bb*
Intermediate_filament_protein_MDM1	1	0	1	1	*Ht3*, *Os*, *Bb*
Oxidoreductase FAD/NAD(P)-binding	28	35	4	27	*Ht3*, *Ht6*, *Bb*

Genes involved in signal transduction were identified in all the four genomes (supplementary table S8, Supplementary Material online). These included several multidrug resistance transporters, responsible for protection against a variety of toxic compounds (Morschhauser 2010). We also identified RNases (supplementary table S9, Supplementary Material online) in the four genomes that catalyze the cleavage of phosphodiester bonds in RNA and play important role in RNA processing and nonspecific RNA degradation (Mishra and Nawin 1995).

Gene Clusters Involved in Biosynthesis of Secondary Metabolites

In the present genome analysis, genes involved in the biosynthesis of secondary metabolites from the four genomes were found to mainly consist of NRPS, PKS, and terpene gene clusters (supplementary table S10, Supplementary Material online). The contraction of these gene clusters was observed in *Os*, except for unusual heterocyst glycolipid synthase (HglD/E)-like PKS known for production of unsaturated fatty acids and lipids (Oßwald et al. 2014) that were expanded in *Os*, and is perhaps related to their psychrophilic adaptation.

Metabolism-Associated Genes

CYP family enzymes, CYP51 (lanosterol 14 alpha-demethylase) which is involved in ergosterol biosynthesis and in membrane permeabilization (Ferreira et al. 2005) and CYP53 involved in detoxification of toxic benzoate molecules (Jawallapersand et al. 2014), were observed in all the four genomes. In addition to these, CYP504, involved in

Table 4
Pathogen–Host Interacting Genes and Their Distribution in *Ht3*, *Ht6*, *Os,* and *Bb* Genomes

Crucial Steps in Pathogen–Host Interactions	Interacting Virulence-Associated Genes	*Ht3*	*Ht6*	*Os*	*Bb*	References
1. Adhesion	Hydrophobins	3	3	1	4	St Leger et al. (1992)
	Cell-wall protein genes (Mad1, Mad2)	1	1	2	0	Wang and St. Leger (2007)
2. Appresorium formation	Protein kinases (MAPK/hog1)	123	129	128	139	Zhang et al. (2010); Zhang et al. (2009)
3. Epicuticle Penetration	CYP	69	72	42	71	Pedrini et al. (2010)
	Vitamin H: biotin transporter	2	1	1	1	Fan et al. (2011)
4. Cuticle Penetration	Glycosyl transferases (carboxylate transporter genes: increase pH)	37	37	31	35	Jin et al. (2009)
	Subtilisin protease: protein lysis in insect epidermis	12	13	9	15	Donatti et al. (2008)
	Glutaminase and NADH-specific glutamate dehydogenase: catabolism of aminoacids	4	4	7	2	Freimoser et al. (2005)
	GHs	22	20	15	26	Fang et al. (2005)

detoxification of phenylacetate derivatives (Ferrer-Sevillano and Fernandez-Canon 2007), was observed in maximum copy numbers as compared with the other CYPs in the four genomes. CYP52 family proteins, known to participate in host–cuticle penetration by metabolizing host epicuticle lipids (Zhang et al. 2012), were also observed in all the four genomes, with comparatively more proteins in *Bb*. *Ht3*, *Bb,* and *Os* showed contraction (negative CAFE average expansion values, *Ht3*: −0.03, *Bb*: −0.32, *Os*: −0.44) of CYPs whereas expansion of CYPs was observed in *Ht6* (CAFE average expansion is 0.02) (supplementary table S11, Supplementary Material online). Interestingly, a protein belonging to the CYP55/PNOR family of nitric oxide reductases was detected in the *Ht* genomes. This protein was observed to share synteny in *Ht3* and *Ht6* with conserved flanking regions (supplementary fig. S4, Supplementary Material online).

RIP and Diversity of TEs

Repeat-induced point mutation (RIP) is a genome defense mechanism in fungi that acts on repetitive DNA to control the progression of TEs through hypermutation (Hood et al. 2005). The RIP frequency index in all the four genomes was found to be almost similar whereas the distribution of TEs differed (fig. 1). Overall, an irregular pattern of TEs was observed in all the four genomes.

Mating Type

A number of mating type genes common in fungi were observed in the *Ht3*, *Ht6*, *Bb*, and *Os* genomes (supplementary table S12, Supplementary Material online). Based on the presence of conserved domains and sequence similarity with mating type genes of hypocrealean entomopathogenic and plant pathogenic fungi, six mating type genes, that is, MAT1-1-1, MAT1-1-2, MAT1-1-3 (MAT1-1 locus), MAT1-2-1 (MAT1-2 locus), matA/MAT A3 (MAT2 locus: Two copies) were identified in the *Ht3* and *Os* genomes, whereas only four mating type genes were identified in the *Ht6* and *Bb* genomes, including MAT1-2-1 (MAT1-2 locus: Two copies) and matA/MAT A3 (MAT2 locus: Two copies) in *Ht6* genome; and MAT1-1-1 (MAT1-1 locus), MAT1-2-1 (MAT1-2 locus), and matA/MAT A3 (MAT2 locus: Two copies) in *Bb* genome. Interestingly, syntenic analysis revealed the presence of MAT1-2 locus in place of MAT1-1 locus and the conservation of genes flanking the MAT1-1 locus in *Ht6* (fig. 2). *Bb* showed the presence of only MAT1-1-1 gene (fig. 2). Synteny was also analyzed for other MAT loci, MAT1-2 and MAT2, and the genes flanking these MAT loci were found to be mostly conserved in the four genomes (supplementary fig. S5, Supplementary Material online). The genomic location of these mating type genes is provided in supplementary table S13, Supplementary Material online.

Phylogenetic Analysis of GH Family 18 Chitinase Sequences

Fungal chitinases have previously been classified into three subgroups A, B, and C on the basis of amino acid (aa) sequence similarity of their GH18 catalytic domains (Seidl et al. 2005; Seidl 2008). In our analysis, the three data sets, through, subgroups A, B, and C were individually analyzed to reconstruct the phylogeny of fungal chitinases. Uniformity in the sequence profile of *Ht3* and *Ht6* chitinases was observed. These *Ht3* and *Ht6* chitinase sequences have been named based on their increasing molecular weight values (supplementary table S14, Supplementary Material online). *Bb* and *Os* chitinases have been named on the basis of sequence similarity with *Ht* chitinases. These chitinases are represented as

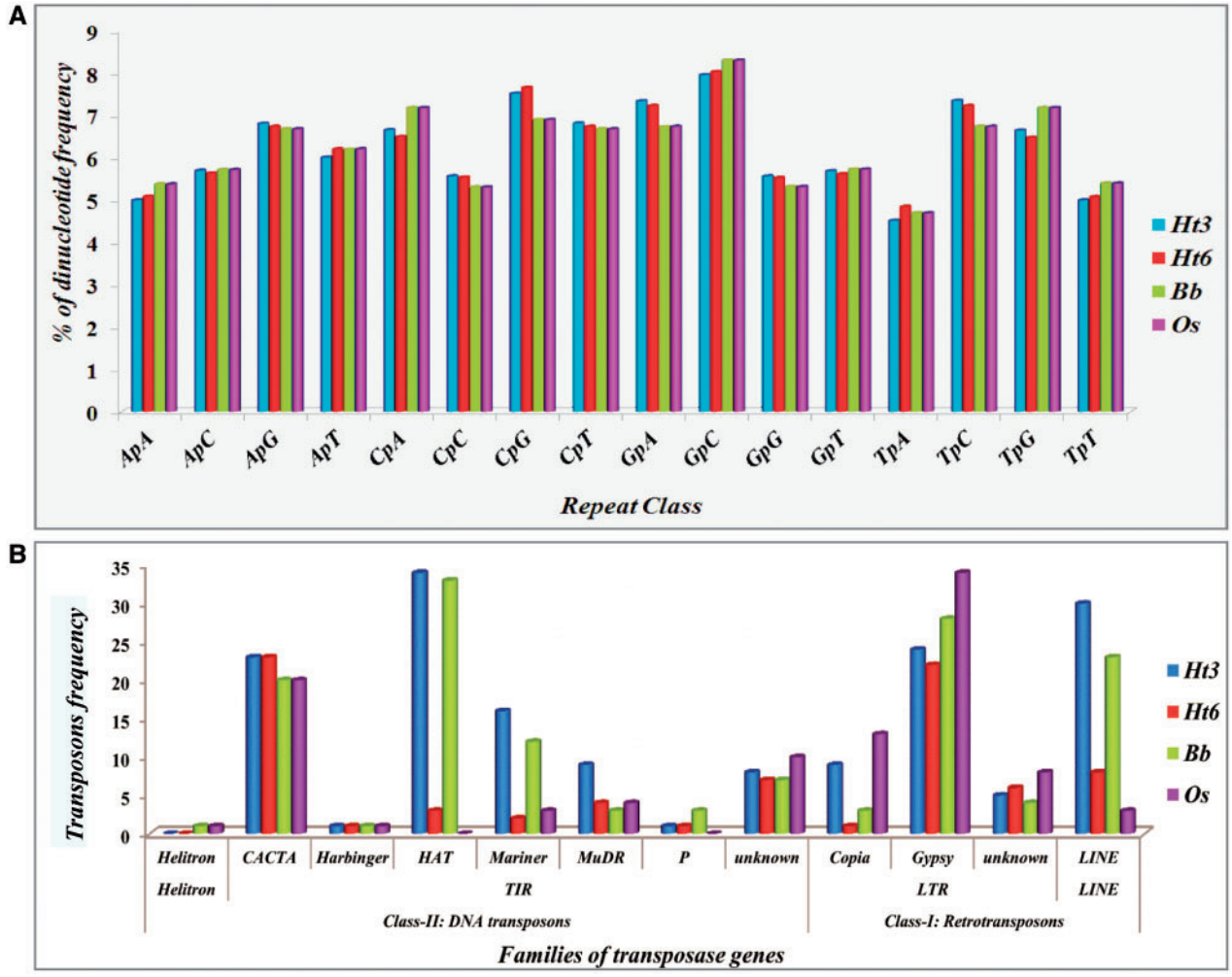

Fig. 1.—(A) Estimation of RIPs and (*B*) Diversity of transposase genes in *Ht3*, *Ht6*, *Bb,* and *Os* genomes. RIPs were estimated using the RIPCAL. RepeatMasker was used to search for the TEs.

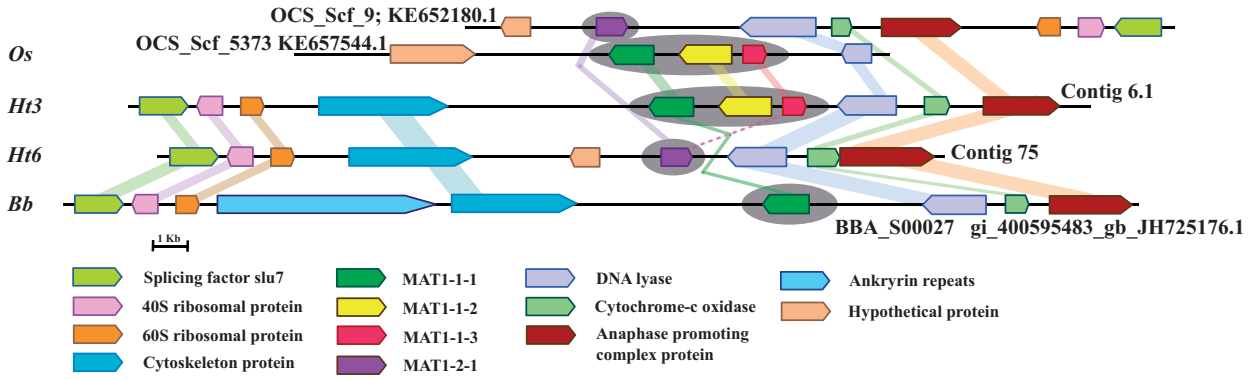

Fig. 2.—Synteny of mating-type loci (MAT1-1-1, MAT1-1-2, and MAT1-1-3) and their flanking regions in *Os*, *Ht3*, *Ht6,* and *Bb* genomes. MAT loci and their flaking gene regions showing orthologous relationships are labeled in the same color.

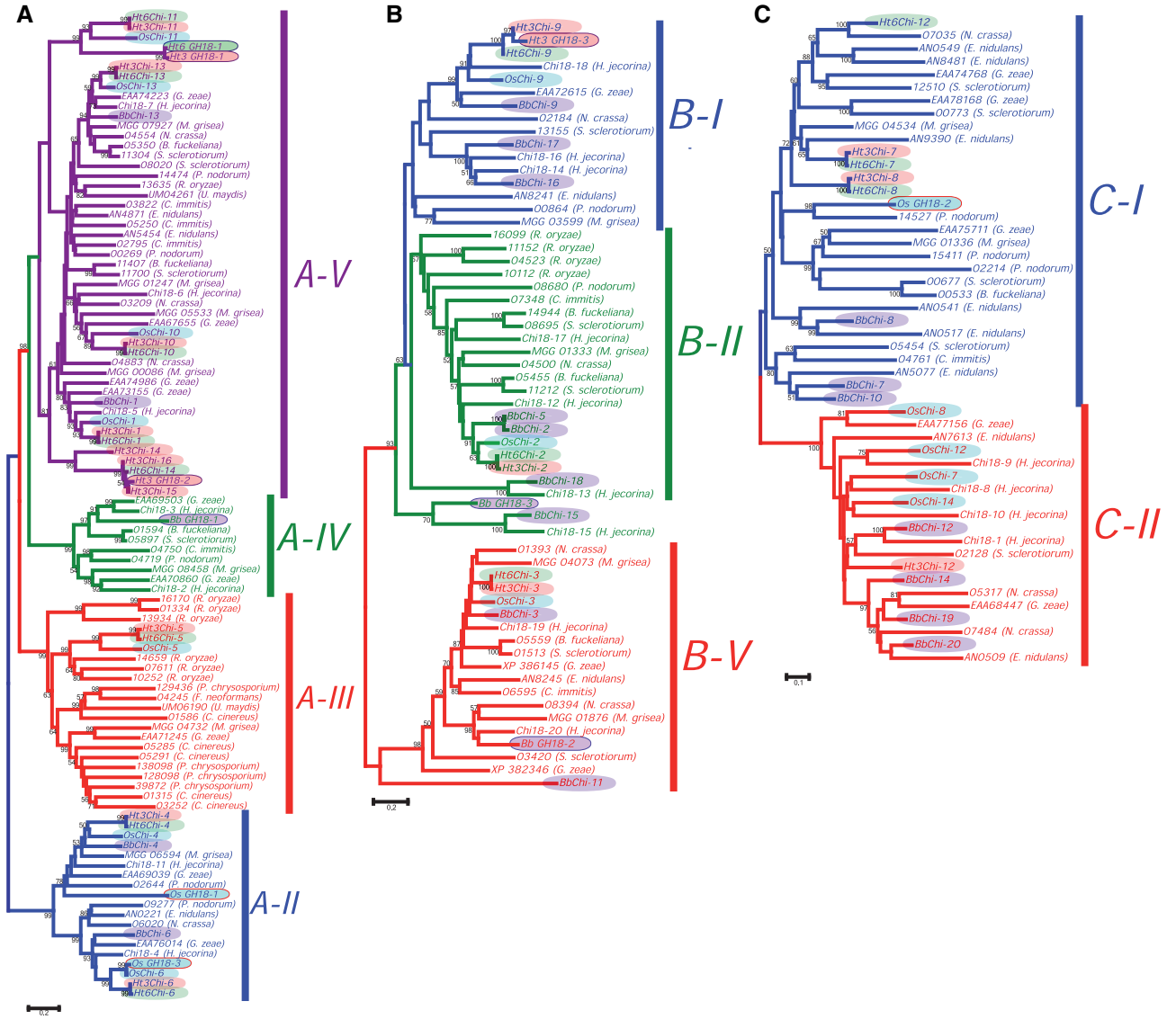

Fig. 3.—Phylogenetic relationships of catalytic domains of GH family 18 chitinases subgroups A, B, and C (*A*–*C*). MP tree showing evolutionary history (PAUP* 4.0 b win). Bootstrap values greater than 50% are shown. Shaded region shows the chitinase sequences obtained from *Ht3*, *Ht6*, *Bb*, and *Os* genomes.

XChi-z, where *X* is the organism label (*Ht3*, *Ht6*, *Bb,* or *Os*) and *z* is the designated number. GH18 chitinase sequences with mutations in the catalytic site are referred to with a different nomenclature, where *X GH18-z* is used for demarcating it from the putative chitinolytically active proteins. Phylogenetic trees are shown in figure 3*A*–*C* and the PAUP statistical information on data sets is presented in supplementary table S15, Supplementary Material online.

In this study, 23, 19, 15, and 17 chitinase sequences from *Bb*, *Ht3*, *Ht6,* and *Os* genomes, respectively, were identified. Among four additional chitinase sequences identified in *Ht3* as compared with *Ht6*, three belong to subgroup-A and one to subgroup-B (fig. 3). The three chitinase sequences (*Ht3Chi-15*, *Ht3Chi-16*, and *Ht3 GH18-2*) belonging to subgroup-A are observed to cluster with *Ht6Chi-14*. Further, all these chitinase sequences show close similarity with *Ht3Chi-14* (fig. 3*A*). *Ht3Chi-15* and *Ht3Chi-16* are identical in length (278 aa) and with a predicted molecular weight of 31 kDa. Their sequences are 97% identical, differing only at 7 aa positions. *Ht3 GH18-2* sequence is 208 aa long and 23 kDa in weight. *Ht3Chi-14* sequence differs from *Ht3Chi-15*, *Ht3Chi-16*, and *Ht3GH18-2* in having a higher predicted molecular weight (38 kDa) and length (338 aa).

Domain Architecture of GH Family 18 Chitinases

The domain organization for all the chitinases identified in this study is presented in figure 4. *Ht3* and *Ht6* were found to have a similar domain pattern with the exception of *Chi-12*. *Ht3Chi-12* was observed to have a single chitin-binding domain belonging to CBM18 family, whereas *Ht6Chi-12* had two CBM18 domains. Additional domains for AMP-binding enzyme, phosphopantetheine attachment site and nicotinamide adenine dinucleotide (NAD)-binding enzyme were observed in *HtChi-13* and *OsChi-13*. Similarly, WD40 domain was observed in *BbChi-1* and Hce2 domain was detected in *BbChi-14* and *BbChi-20*.

Discussion

Comparative Distinctions among *Bb*, *Ht*, and *Os* Genomes

Genome comparisons of *Bb, Ht3, Ht6,* and *Os* suggest overall reduced gene family sizes in *Ht3* and *Os* as compared with *Ht6* and *Bb* genomes. It is known that RIP is a maintenance force of evolutionary diversity and genome organization during sexual cycle (Selker et al. 1987; Galagan et al. 2003). We speculate that the differences in gene family sizes in the four fungi may perhaps be related to the frequency of meiotic/recombination events linked to the active RIP mechanism (Meerupati et al. 2013), in addition to the other selective forces.

Hypocrealean fungi are known to exhibit diversity in their mating behavior. Some of these fungi are self-fertile/homothallic, whereas some are self-sterile/heterothallic (Ni et al. 2011). Genotypically, their mating response has been indicated to be regulated by mating type genes (Lee et al. 2010; Ni et al. 2011). The presence of two compatible MAT idiomorphs, MAT1-1 and MAT1-2, is suggested to decide homothallism and heterothallism in ascomycetes (Lee et al. 2010; Ni et al. 2011). As of now, there has been no report on mating/teleomorph in *Ht*. It is quite possible that *Ht* undergoes recombination/meiosis but it is uncertain yet. Interestingly, in this study, MAT1-1-1, MAT1-1-2, and MAT1-1-3 mating type genes (MAT1-1 idiomorph) are identified in *Ht3*, whereas synteny analysis reveals the presence of MAT1-2-1 gene (MAT1-2 idiomorph) in place of MAT1-1 idiomorph in *Ht6* (fig. 2). This indicates at cryptic sexuality in *Ht*. Similar to this study, isolate-specific differences in mating genes were observed in two isolates of *Metarhizium robertsii* (Previous name: *M. anisopliae*) (Pattemore et al. 2014). However, fungi can switch between homo- and heterothallic mode (Ni et al. 2011); therefore, additional gene expression and experimental studies are required to investigate their homo-/heterothallic mode of mating.

Entomopathogenic fungi have a number of virulence-associated genes in their genome to adapt to the diverse ecological niches (Klosterman et al. 2011; Gardiner et al. 2013; Zhao et al. 2014). Therefore, the expansion of virulence genes is suggested to affect the pathogenicity potential and evolution of host affiliation. *Os* genome (size: 78.5 Mb) is about more than double the size of other three fungal genomes but shows a comparatively reduced set of protein-coding genes and is rich in repeats similar to that observed for many obligatory pathogens such as powdery mildew fungi (Spanu et al. 2010) and rust fungi (Duplessis et al. 2011). It shows peculiar behavioral traits, such as a biphasic pathogenic lifestyle, obligatory mycoparasitism, and psychrophilic adaptation. Unlike the other three fungi in this study, *Os* has a unique lifestyle exhibiting concealed virulence in early larval developmental stages of ghost moths and a fatal virulence in late caterpillar instars (Hu et al. 2013). Interestingly, *Os* contains several unusual PKS that participate in accumulation of lipids and unsaturated fatty acids (Oßwald et al. 2014) that perhaps help it to adapt to extreme cold environments.

CYPs represent one of the largest protein families which are involved in a number of biochemical reactions associated with primary, secondary, and xenobiotic metabolism (Cresnar and Petric 2011). CYPs are suggested to be important for organisms to acclimatize to novel ecological niches through functional diversification (Cresnar and Petric 2011). These proteins play important roles in detoxification, biosynthesis, and other metabolic reactions (Cresnar and Petric 2011). *Ht* genomes contain a CYP55/PNOR family gene that is reported to participate in anaerobic denitrification and detoxification (Takaya et al. 2002; Shoun et al. 2012), possibly enabling adaptation to hypoxic/anoxic conditions in *Ht*. This gene was likely acquired through horizontal transfer from actinobacteria as inferred by sequence identities (about 40%) with bacterial CYP105 (Tomura et al. 1994; Shoun et al. 2012). This gene is suggested to be an important constituent of fungal denitrification system that involves reduction of toxic nitric oxide into nontoxic nitrous oxide (Shoun et al. 2012). The metabolic process of this system is not connected to the mitochondrial respiratory system and redox partners, rather electrons are directly received by CYP55 from NAD to reduce nitric oxide to nitrous oxide (Shoun et al. 2012). Therefore, the presence of CYP55 could be related to the fungal survival strategy under oxygen deficient or oxidative stress conditions in response to the antimicrobial defense of the host (Nittler et al. 2005; Shoun et al. 2012). It is thus suggested as one of the factors in *Ht* that could be responsible for its variable ecological niche adaptations and host affiliations.

Ht3 and *Ht6* Chitinases

This study reports 19 and 15 chitinase sequences, including three and one GH18 sequences with mutations at the active site in *Ht3* and *Ht6* genomes, respectively. This observation is in accordance with the previous reports that there are higher numbers of chitinase genes in entomopathogenic fungi (ranging from 10 to 20) than plant-pathogenic fungi (ranging from

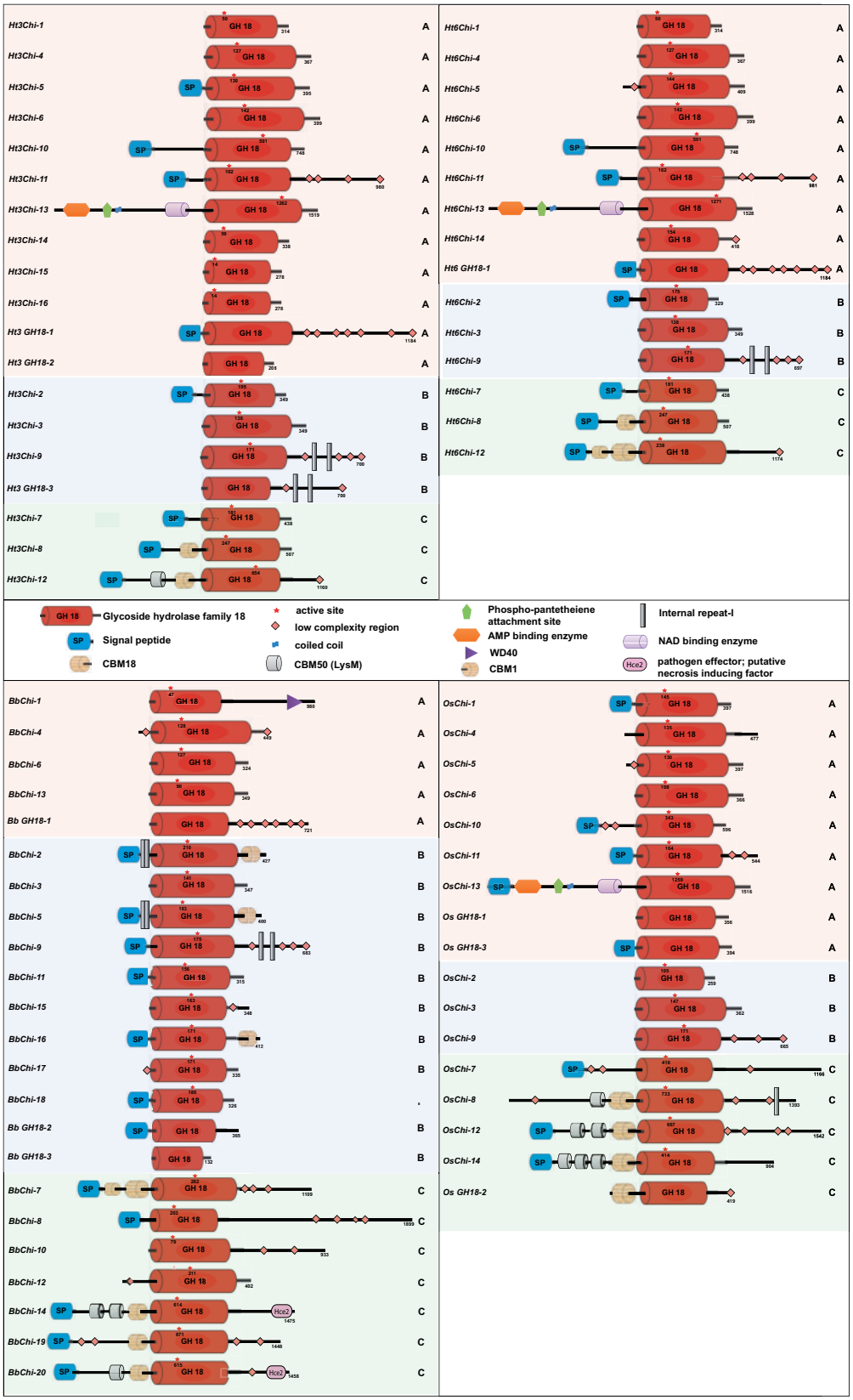

Fig. 4.—Domain organization of *Ht3*, *Ht6*, *Bb*, and *Os* chitinases. Protein domains, as identified with InterProScan, Pfam, and SMART databases are shown.

6 to 15) (Gao et al. 2011; Xiao et al. 2012). *Ht3Chi-14* sequence is predicted to possibly share paralogous relationship with *Ht3Chi-15*, *Ht3Chi-16*, and *Ht3GH18-2* chitinase sequences. We therefore consider that additional subgroup-A chitinase sequences in *Ht3* strain could have been acquired through gene duplication events.

Revision of Classification of *Bb* Chitinases

Our study refines the classification of *Bb* chitinases previously proposed (Xiao et al. 2012), where they compared the occurrence of domains and modules in chitinases and classified nine *Bb* chitinases without CBM in subgroup-A, four *Bb* chitinases with CBM at C-terminal in subgroup-B, and seven *Bb* chitinases with CBM18 and CBM50 in subgroup-C. In our study, we have classified the chitinases based on aa sequence similarity of catalytic domains of subgroups A, B, and C chitinases, similar to previous reports (Seidl et al. 2005; Karlsson and Stenlid 2009). Interestingly, 5, 11, and 7 *Bb* chitinases are grouped in subgroups A, B, and C, respectively. The differences could be observed in domain distribution previously proposed (Xiao et al. 2012) (fig. 4). One of the subgroup-A chitinases (*BbChi-1*) contains a WD domain. Only three of the subgroup-B chitinases classified in our analysis are observed to contain CBM1. The remaining eight subgroup-B chitinases are not found to contain any CBMs. Similarly, only four of seven subgroup-C chitinase sequences are observed to contain CBM18 and CBM50 domains. Subgroup-B and -C chitinases are mainly attributed to participate in nutritional and aggressive functions such as virulence to insects in entomopathogenic fungi (de las Mercedes Dana et al. 2001; da Silva et al. 2005; Baratto et al. 2006); therefore, higher number of subgroup-B and subgroup-C chitinases in *Bb* genome may be related to the virulence of *Bb*.

Evolutionary Peculiarities in *Bb* Chitinases

Previous studies suggest *Hypocrea jecorina Chi18-15* (GenBank: BK006088.1), an active participant in degradation of insect cuticle, to have been horizontally transferred from actinobacteria (Karlsson and Stenlid 2009; Ihrmark et al. 2010). *Hypocrea jecorina Chi18-15* clusters with *BbChi-15* (GenBank: KM220983), sharing 84% aa sequence similarity in the catalytic domain that suggests this chitinase was horizontally transferred prior to the split of *H. jecorina* and *Bb* (fig. 3*B*). The evolution of *H. jecorina Chi18-15* and its fungal and bacterial orthologs have been suggested to occur under different selective forces following horizontal gene transfer where the functional divergence is proposed to be created by loop-associated regions to rapidly evolve selective modifications (Ubhayasekera and Karlsson 2012). This suggests adaptive evolution in *BbChi-15* chitinase and positive selection of adaptable modifications in functional properties (Alsmark et al. 2013) that may affect the entomopathogenicity potential of *Bb*.

Bb GH18-3 represents a paralog of *BbChi-15* and *H. jecorina Chi18-15* (fig. 3*B*). The identified *Bb GH18-3* chitinase sequence (132 aa) is almost half the length of other GH18 chitinase catalytic domain sequences. Sequence and structure-based analysis suggests a half TIM-barrel fold, containing only the N-terminal half region, for *Bb GH18-3*. The TIM-barrel $(\alpha/\beta)_8$ fold has been suggested previously to have evolved by tandem gene duplication and fusion of half-barrels $(\alpha/\beta)_4$ and hence it is possible that *Bb GH8-3* may function as a homodimer (Lang et al. 2000; Nagano et al. 2002). Nevertheless, there is a need to experimentally validate biochemical function of *Bb GH18-3* to comment on its evolutionary implications.

Protein Domain Distribution in Fungal Chitinases

This study provides interesting insights about the domain distribution in chitinases of entomopathogenic fungi (fig. 4). For example, *Ht3Chi-12* is found to have domains CBM18 and CBM50 (LysM), whereas *Ht6Chi-12* has two CBM18 domains. However, both CBM18 and CBM50 are implicated in chitin-binding functions (Seidl 2008). Similarly, presence of AMP+PP+NAD in subgroup-A *Chi-13* of *Ht* and *Os* indicates their role in secondary metabolites (nonribosomal peptide synthetases) production, virulence, and fungal development (Molnar et al. 2010). Likewise, presence of WD40 domain in *Bb-Chi1* belonging to subgroup-A could have implications in molecular recognition events through protein–protein or protein–DNA interaction (Xu and Min 2011). Subgroup-A chitinases are so far known to participate in general physiological functions related to the growth and development of the fungus (Seidl 2008). The occurrences of additional domains AMP+PP+NAD or WD40 in subgroup-A chitinases hints at a possible role of subgroup-A chitinases in pathogenicity in addition to other routine functions. However, further investigations are required to confirm the role of subgroup-A chitinases in virulence.

Though subgroup-B chitinases are expected to have CBM domains (Hartl et al. 2012), our analysis reveals inconsistent scenarios. The CBMs are found to be present in subgroup-B chitinases from *Bb* genome but absent in those from *Ht* and *Os* genomes. *Aspergillus niger*, *A. nidulans*, *A. terreus*, *A. flavus*,*Neosartorya fischeri* (Ascomycota) and *Coprinus cinereus*, *Sporobolomyces roseus* (Basidiomycota) too lack CBMs in their subgroup-B chitinases (Seidl 2008). The CBMs are involved in chitin-binding functions and their absence in *Ht*, *Os* and other ascomycete and basidiomycete genomes may affect the specific recognition events of fungal chitinases. Further, the exclusive association of CBM1 family members with subgroup-B chitinases and the association of CBM18 and CBM50 family members (possessing chitin-binding ability) with subgroup-C chitinases identified in our study are found to be consistent with the CBM distribution of mycoparasitic *Trichoderma* spp. that are important biocontrol agents (Karlsson and Stenlid 2009; Gruber et al. 2011).

The domain architecture of subgroup-C chitinases interestingly reveals the presence of Hce2 domains in *BbChi-14* and *BbChi-20*. Hce2 has been identified as a homolog of the Ecp2 effector protein of plant-pathogenic fungus *Cladosporium fulvus* (Van den Ackerveken et al. 1993; Lauge et al. 1998; Stergiopoulos et al. 2010, 2012). It has been recognized as a necrosis-inducing factor in plants that is involved in pathogenicity against the host (Bolton et al. 2008; Stergiopoulos et al. 2010, 2012). In silico characterization has revealed that in conjugation with GH18 chitinases, Hce2 shows high similarity with killer toxin Zymocin from *Kluyveromyces lactis* with possible antagonistic functions (Stergiopoulos et al. 2012). Similarly, a recent functional study of *A. nidulans* subgroup-C killer toxin-like chitinases has shown the effect of deleted subgroup-C chitinase associated with Hce2 domain in lowering the growth inhibitory activity of culture filtrates, implicating the role of this protein in interspecific fungal–fungal interactions (Tzelepis et al. 2014). Accordingly, the occurrence of Hce2 in *Bb* could be one of the factors for virulence and ecological niche adaptations.

Research Scope

Ht is known to produce the insecticidal ribotoxin hirsutellin A (HtA) (Herrero-Galán et al. 2008). Although about 20 ribotoxins have been described from filamentous fungi, HtA is the only one known from entomopathogenic fungi (Arruda et al. 1992; Lin et al. 1995; Huang et al. 1997; Kao et al. 2001; Herrero-Galán et al. 2008). We have identified a total of 77 putative RNase protein sequences in *Bb*, *Ht3*, *Ht6,* and *Os* genomes (supplementary table S9, Supplementary Material online); however, there is a need to gain more insights about their evolution and function in entomopathogenic fungi that could help in the identification of components responsible for making RNases into ribotoxins. Additionally, it will aid in the exploitation of these ribotoxins in plant protection measures and biomedical applications.

The GHs analysis shows the prevalence of GH families GH2, GH3, GH16, GH18, GH43, GH47, GH72, GH76, and GH109 in the four genomes. Among these characterization of 5 GH2, 30 GH3, 5 GH16, 13 GH18, 6 GH43, and 5 GH47 members of fungal origin have been reported (Murphy et al. 2011), but none from GH72, GH76, and GH109. Moreover, about 25% of annotated proteins in all the four genomes are with unknown/hypothetical functions.

A total of 244 PHI protein families categorized from the four genomes can further help in exploring their role in pathogenicity, such as to aid in the discovery of enzymes with potential insecticidal properties.

Conclusions

The genome sequencing of the insect pathogenic fungi *Ht* and its comparative analysis with *Bb* and *Os* genomes have provided interesting insights into insect–fungal associations, evolutionary insights into chitinases, and offer clues to improve the efficacy of *Ht* as a biological control agent.

Supplementary Material

Supplementary tables S1–S15, figures S1–S5, and data sets S1–S3 are available at *Genome Biology and Evolution* online (http://www.gbe.oxfordjournals.org/).

Acknowledgments

The authors acknowledge the Microbial Type Culture Collection and Gene Bank (MTCC), Chandigarh, India from where the isolates were retrieved and the C-CAMP next-generation genomics facility, Bangalore, India for help in obtaining NGS data. They thank Dr Magnus Karlsson, Department of Forest Mycology and Pathology, Swedish University of Agricultural Sciences, Uppsala, Sweden for providing the chitinase data set of 394 sequences. They also thank Ms Poonam Mual for her help in DNA isolation, Mr Gaurav Sharma, and Ms Gurmeet Kaur for useful discussions on chitinase sequence and structure analysis. They also thank Mr Tarun Narwani for help in statistical analysis. This work was supported by a project "Expansion and modernization of MTCC jointly supported by Council of Scientific and Industrial Research (CSIR) Grant No. BSC0402 and Department of Biotechnology (DBT) Government of India Grant No. BT/PR7368/INF/22/177/2012." Y.A. and I.K. are supported by research fellowships from CSIR and University Grants Commission (UGC), respectively. B.D.S. and S.K.S. conceived the idea. I.K. generated the data. Y.A. analyzed the data and wrote the manuscript. All authors reviewed the manuscript.

Literature Cited

Alsmark C, et al. 2013. Patterns of prokaryotic lateral gene transfers affecting parasitic microbial eukaryotes. Genome Biol. 14:R19.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol. 215:403–410.

Arruda LK, Mann BJ, Chapman MD. 1992. Selective expression of a major allergen and cytotoxin, Asp f I, in *Aspergillus fumigatus*. Implications for the immunopathogenesis of *Aspergillus*-related diseases. J Immunol. 149:3354–3359.

Baratto CM, et al. 2006. Isolation, characterization, and transcriptional analysis of the chitinase *chi2* Gene (DQ011663) from the biocontrol fungus *Metarhizium anisopliae* var. *anisopliae*. Curr Microbiol. 53: 217–221.

Boetzer M, Henkel CV, Jansen HJ, Butler D, Pirovano W. 2011. Scaffolding pre-assembled contigs using SSPACE. Bioinformatics 27:578–579.

Boetzer M, Pirovano W. 2012. Toward almost closed genomes with GapFiller. Genome Biol. 13:R56.

Bolton MD, et al. 2008. The novel *Cladosporium fulvum* lysin motif effector Ecp6 is a virulence factor with orthologues in other fungal species. Mol Microbiol. 69:119–136.

Bushley KE, Li Y, et al. 2013. Isolation of the MAT1-1 mating type idiomorph and evidence for selfing in the Chinese medicinal fungus *Ophiocordyceps sinensis*. Fungal Biol. 117:599–610.

Bushley KE, Raja R, et al. 2013. The genome of *Tolypocladium inflatum*: evolution, organization, and expression of the cyclosporin biosynthetic gene cluster. PLoS Genet. 9:e1003496.
Cantarel BL, et al. 2008. MAKER: an easy-to-use annotation pipeline designed for emerging model organism genomes. Genome Res. 18: 188–196.
Charnley AK. 2003. Fungal pathogens of insects: cuticle degrading enzymes and toxins. Adv Bot Res. 40:241–321.
Cresnar B, Petric S. 2011. Cytochrome P450 enzymes in the fungal kingdom. Biochim Biophys Acta. 1814:29–35.
Cuskin F, et al. 2012. How nature can exploit nonspecific catalytic and carbohydrate binding modules to create enzymatic specificity. Proc Natl Acad Sci U S A. 109:20889–20894.
da Silva MV, et al. 2005. Cuticle-induced endo/exoacting chitinase CHIT30 from *Metarhizium anisopliae* is encoded by an ortholog of the chi3 gene. Res Microbiol. 156:382–392.
de las Mercedes Dana M, et al. 2001. Regulation of chitinase 33 (*chit33*) gene expression in *Trichoderma harzianum*. Curr Genet. 38: 335–342.
Donatti AC, Furlaneto-Maia L, Fungaro MH, Furlaneto MC. 2008. Production and regulation of cuticle-degrading proteases from *Beauveria bassiana* in the presence of *Rhammatocerus schistocercoides* cuticle. Curr Microbiol. 56:256–260.
Duplessis S, et al. 2011. Obligate biotrophy features unraveled by the genomic analysis of rust fungi. Proc Natl Acad Sci U S A. 108: 9166–9171.
Fan Y, Zhang S, Kruer N, Keyhani NO. 2011. High-throughput insertion mutagenesis and functional screening in the entomopathogenic fungus *Beauveria bassiana*. J Invertebr Pathol. 106:274–279.
Fang W, et al. 2005. Cloning of *Beauveria bassiana* chitinase gene *Bbchit1* and its application to improve fungal strain virulence. Appl Environ Microbiol. 71:363–370.
Ferreira ME, et al. 2005. The ergosterol biosynthesis pathway, transporter genes, and azole resistance in *Aspergillus fumigatus*. Med Mycol. 43: S313–S319.
Ferrer-Sevillano F, Fernandez-Canon JM. 2007. Novel phacB-encoded cytochrome P450 monooxygenase from *Aspergillus nidulans* with 3-hydroxyphenylacetate 6-hydroxylase and 3,4-dihydroxyphenylacetate 6-hydroxylase activities. Eukaryot Cell. 6:514–520.
Finn RD, Clements J, Eddy SR. 2011. HMMER web server: interactive sequence similarity searching. Nucleic Acids Res. 39:W29–W37.
Freimoser FM, Hu G, St, Leger RJ. 2005. Variation in gene expression patterns as the insect pathogen *Metarhizium anisopliae* adapts to different host cuticles or nutrient deprivation in vitro. Microbiology 151:361–371.
Galagan JE, et al. 2003. The genome sequence of the filamentous fungus *Neurospora crassa*. Nature 422:859–868.
Gao Q, et al. 2011. Genome sequencing and comparative transcriptomics of the model entomopathogenic fungi *Metarhizium anisopliae* and *M. acridum*. PLoS Genet. 7:e1001264.
Gardiner DM, Kazan K, Manners JM. 2013. Cross-kingdom gene transfer facilitates the evolution of virulence in fungal pathogens. Plant Sci. 210:151–158.
Gasteiger E, et al.. 2005. Protein identification and analysis tools on the ExPASy server. In: John MW, editor. The proteomics protocols handbook. Totowa (NJ): Humana Press. p. 571–607.
Gruber S, et al. 2011. Analysis of subgroup C of fungal chitinases containing chitin-binding and LysM modules in the mycoparasite *Trichoderma atroviride*. Glycobiology 21:122–133.
Hamid R, et al. 2013. Chitinases: an update. J Pharm Bioallied Sci. 5:21–29.
Han MV, Thomas GW, Lugo-Martinez J, Hahn MW. 2013. Estimating gene gain and loss rates in the presence of error in genome assembly and annotation using CAFE 3. Mol Biol Evol. 30: 1987–1997.
Hane JK, Oliver RP. 2008. RIPCAL: a tool for alignment-based analysis of repeat-induced point mutations in fungal genomic sequences. BMC Bioinformatics 9:478.
Haragsim O, Ruzicka V. 2001. Biological control of *Varroa* mites in honeybee hives with *Hirsutella thompsonii*. US patent (United States). Patent No. 6.277.371 b1. Application No. 09/574,933.
Hartl L, Zach S, Seidl-Seiboth V. 2012. Fungal chitinases: diversity, mechanistic properties and biotechnological potential. Appl Microbiol Biotechnol. 93:533–543.
Herrero-Galán E, et al. 2008. The insecticidal protein hirsutellin A from the mite fungal pathogen *Hirsutella thompsonii* is a ribotoxin. Proteins 72: 217–228.
Herrero-Galán E, et al. 2013. Hirsutellin A: a paradigmatic example of the insecticidal function of fungal ribotoxins. Insects 4:339–356.
Hood ME, Katawczik M, Giraud T. 2005. Repeat-induced point mutation and the population structure of transposable elements in *Microbotryum violaceum*. Genetics 170:1081–1089.
Hu X, et al. 2013. Genome survey uncovers the secrets of sex and lifestyle in caterpillar fungus. Chin Sci Bull. 58:2846–2854.
Huang KC, Hwang YY, Hwu L, Lin A. 1997. Characterization of a new ribotoxin gene (c-sar) from *Aspergillus clavatus*. Toxicon 35:383–392.
Ihrmark K, et al. 2010. Comparative molecular evolution of *Trichoderma* chitinases in response to mycoparasitic interactions. Evol Bioinform Online. 6:1–26.
Jablonowski D, et al. 2001. *Saccharomyces cerevisiae* cell wall chitin, the *Kluyveromyces lactis* zymocin receptor. Yeast 18:1285–1299.
Jawallapersand P, et al. 2014. Cytochrome P450 monooxygenase CYP53 family in fungi: comparative structural and evolutionary analysis and its role as a common alternative anti-fungal drug target. PLoS One 9: e107209.
Jin K, et al. 2009. Carboxylate transporter gene JEN1 from the entomopathogenic fungus *Beauveria bassiana* is involved in conidiation and virulence. Appl Environ Microbiol. 76:254–263.
Kao R, Martinez-Ruiz A, Martinez del Pozo A, Crameri R, Davies J. 2001. Mitogillin and related fungal ribotoxins. Methods Enzymol. 341: 324–335.
Karlsson M, Stenlid J. 2009. Evolution of family 18 glycoside hydrolases: diversity, domain structures and phylogenetic relationships. J Mol Microbiol Biotechnol. 16:208–223.
Kirk PM, Cannon PF, Minter DW, Stalpers JA. 2008. Dictionary of the fungi, 10th ed.. Wallingford (United Kingdom): CAB International.
Klix V, et al. 2010. Functional characterization of MAT1-1-specific mating-type genes in the homothallic ascomycete *Sordaria macrospora* provides new insights into essential and nonessential sexual regulators. Eukaryot Cell. 9:894–905.
Klosterman SJ, et al. 2011. Comparative genomics yields insights into niche adaptation of plant vascular wilt pathogens. PLoS Pathog. 7: e1002137.
Kubicek CP, et al. 2011. Comparative genome sequence analysis underscores mycoparasitism as the ancestral life style of *Trichoderma*. Genome Biol. 12:R40.
Lang D, Thoma R, Henn-Sax M, Sterner R, Wilmanns M. 2000. Structural evidence for evolution of the beta/alpha barrel scaffold by gene duplication and fusion. Science 289:1546–1550.
Lauge R, et al. 1998. Successful search for a resistance gene in tomato targeted against a virulence factor of a fungal pathogen. Proc Natl Acad Sci U S A. 95:9014–9018.
Lechner M, et al. 2011. Proteinortho: detection of (co-)orthologs in large-scale analysis. BMC Bioinformatics 12:124.
Lee SC, Ni M, Li W, Shertz C, Heitman J. 2010. The evolution of sex: a perspective from the fungal kingdom. Microbiol Mol Biol Rev. 74: 298–340.
Letunic I, Doerks T, Bork P. 2012. SMART 7: recent updates to the protein domain annotation resource. Nucleic Acids Res. 40:D302–D305.

Lin A, Huang KC, Hwu L, Tzean SS. 1995. Production of type II ribotoxins by *Aspergillus* species and related fungi in Taiwan. Toxicon 33: 105–110.
Liu Z-Y, et al. 2002. Molecular evidence for teleomorph-anamorph connections in *Cordyceps* based on ITS-5.8S rDNA sequences. Mycol Res. 106:1100–1108.
Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B. 2014. The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res. 42:D490–D495.
Lowe TM, Eddy SR. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. 25: 955–964.
Medema MH, et al. 2011. antiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. Nucleic Acids Res. 39:W339–W346.
Meerupati T, et al. 2013. Genomic mechanisms accounting for the adaptation to parasitism in nematode-trapping fungi. PLoS Genet. 9: e1003909.
Mishra NC. 1995. Molecular biology of nucleases. Boca Raton (FL): CRC Press. p. 46–52.
Molnar I, Gibson DM, Krasnoff SB. 2010. Secondary metabolites from entomopathogenic Hypocrealean fungi. Nat Prod Rep. 27: 1241–1275.
Morschhauser J. 2010. Regulation of multidrug resistance in pathogenic fungi. Fungal Genet Biol. 47: 94–106.
Murphy C, Powlowski J, Wu M, Butler G, Tsang A. 2011. Curation of characterized glycoside hydrolases of fungal origin. Database (Oxford) 2011:bar020.
Nagano N, Orengo CA, Thornton JM. 2002. One fold with many functions: the evolutionary relationships between TIM barrel families based on their sequences, structures and functions. J Mol Biol. 321:741–765.
Nelson DR. 2009. The cytochrome p450 homepage. Hum Genomics. 4: 59–65.
Ni M, Feretzaki M, Sun S, Wang X, Heitman J. 2011. Sex in fungi. Annu Rev Genet. 45:405–430.
Nittler MP, Hocking-Murray D, Foo CK, Sil A. 2005. Identification of *Histoplasma capsulatum* transcripts induced in response to reactive nitrogen species. Mol Biol Cell. 16:4792–4813.
Oßwald C, et al. 2014. A highly unusual polyketide synthase directs dawenol polyene biosynthesis in *Stigmatella aurantiaca*. J Biotechnol. 191:54–63.
Paterson RR. 2008. *Cordyceps*: a traditional Chinese medicine and another fungal therapeutic biofactory? Phytochemistry 69:1469–1495.
Pattemore J, et al. 2014. The genome sequence of the biocontrol fungus *Metarhizium anisopliae* and comparative genomics of *Metarhizium* species. BMC Genomics 15:660.
Pedrini N, Zhang S, Juarez MP, Keyhani NO. 2010. Molecular characterization and expression analysis of a suite of cytochrome P450 enzymes implicated in insect hydrocarbon degradation in the entomopathogenic fungus *Beauveria bassiana*. Microbiology 156:2549–2557.
Pei J, Sadreyev R, Grishin NV. 2003. PCMA: fast and accurate multiple sequence alignment based on profile consistency. Bioinformatics 19: 427–428.
Punta M, et al. 2012. The Pfam protein families database. Nucleic Acids Res. 40:D290–D301.
Quevillon E, et al. 2005. InterProScan: protein domains identifier. Nucleic Acids Res. 33:W116–W120.
Saier MH Jr, Reddy VS, Tamang DG, Vastermark A. 2014. The transporter classification database. Nucleic Acids Res. 42:D251–D258.
Saito A, Fujii T, Miyashita K. 2003. Distribution and evolution of chitinase genes in *Streptomyces* species: involvement of gene-duplication and domain-deletion. Antonie Van Leeuwenhoek 84:7–15.
Samson RA, McCoy CW, O'Donnell KL. 1980. Taxonomy of the acarine parasite *Hirsutella thompsonii*. Mycologia 72:359–377.
Schultz J, Milpetz F, Bork P, Ponting CP. 1998. SMART, a simple modular architecture research tool: identification of signaling domains. Proc Natl Acad Sci U S A. 95:5857–5864.
Seidl V. 2008. Chitinases of filamentous fungi: a large group of diverse proteins with multiple physiological functions. Fungal Biol Rev. 22: 36–42.
Seidl V, Huemer B, Seiboth B, Kubicek CP. 2005. A complete survey of *Trichoderma* chitinases reveals three distinct subgroups of family 18 chitinases. FEBS J. 272:5923–5939.
Selker EU, Cambareri EB, Jensen BC, Haack KR. 1987. Rearrangement of duplicated DNA in specialized cells of *Neurospora*. Cell 51:741–752.
Shoun H, Fushinobu S, Jiang L, Kim SW, Wakagi T. 2012. Fungal denitrification and nitric oxide reductase cytochrome P450nor. Philos Trans R Soc Lond B Biol Sci. 367:1186–1194.
Spanu PD, et al. 2010. Genome expansion and gene loss in powdery mildew fungi reveal tradeoffs in extreme parasitism. Science 330: 1543–1546.
St. Leger RJ, Staples RC, Roberts DW. 1992. Cloning and regulatory analysis of starvation-stress gene, ssgA, encoding a hydrophobin-like protein from the entomopathogenic fungus, *Metarhizium anisopliae*. Gene 120:119–124.
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.
Stanke M, Morgenstern B. 2005. AUGUSTUS: a web server for gene prediction in eukaryotes that allows user-defined constraints. Nucleic Acids Res. 33:W465–W467.
Stergiopoulos I, et al. 2010. Tomato Cf resistance proteins mediate recognition of cognate homologous effectors from fungi pathogenic on dicots and monocots. Proc Natl Acad Sci U S A. 107:7610–7615.
Stergiopoulos I, et al. 2012. *In silico* characterization and molecular evolutionary analysis of a novel superfamily of fungal effector proteins. Mol Biol Evol. 29:3371–3384.
Sung GH, et al. 2007. Phylogenetic classification of *Cordyceps* and the clavicipitaceous fungi. Stud Mycol. 57:5–59.
Swofford DL. 2003. PAUP*: phylogenetic analysis using parsimony (* and other methods). Version 4.0. Sunderland (MA): Sinauer Associates.
Takaya N, Uchimura H, Lai Y, Shoun H. 2002. Transcriptional control of nitric oxide reductase gene (CYP55) in the fungal denitrifier *Fusarium oxysporum*. Biosci Biotechnol Biochem. 66:1039–1045.
Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol. 30:2725–2729.
Tatusov RL, Galperin MY, Natale DA, Koonin EV. 2000. The COG database: a tool for genome-scale analysis of protein functions and evolution. Nucleic Acids Res. 28:33–36.
Tomura D, Obika K, Fukamizu A, Shoun H. 1994. Nitric oxide reductase cytochrome P-450 gene, CYP 55, of the fungus *Fusarium oxysporum* containing a potential binding-site for FNR, the transcription factor involved in the regulation of anaerobic growth of *Escherichia coli*. J Biochem. 116:88–94.
Tzelepis GD, Melin P, Stenlid J, Jensen DF, Karlsson M. 2014. Functional analysis of the C-II subgroup killer toxin-like chitinases in the filamentous ascomycete *Aspergillus nidulans*. Fungal Genet Biol. 64:58–66.
Ubhayasekera W, Karlsson M. 2012. Bacterial and fungal chitinase chiJ orthologs evolve under different selective constraints following horizontal gene transfer. BMC Res Notes. 5:581.
Van den Ackerveken GF, et al. 1993. Characterization of two putative pathogenicity genes of the fungal tomato pathogen *Cladosporium fulvum*. Mol Plant Microbe Interact. 6:210–215.
Wang Z-X, et al. 2014. Comparative transcriptomic analysis of the heat stress response in the filamentous fungus *Metarhizium anisopliae* using RNA-Seq. Appl Microbiol Biotechnol. 98:5589–5597.

Wang C, St. Leger RJ. 2007. The MAD1 adhesin of *Metarhizium anisopliae* links adhesion with blastospore production and virulence to insects, and the MAD2 adhesin enables attachment to plants. Eukaryot Cell. 6: 808–816.
Weibelzahl E, Liburd OE. 2009. Epizootic of *Acalitus vaccinii* (Acari: Eriophyidea) caused by *Hirsutella thompsonii* on southern highbush blueberry in north-central Florida. Fla Entomol. 92:601–607.
Weinzierl R, Henn T, Koehler PG, Tucker CL. 2005. Microbial insecticides. Gainesville (FL): University of Florida Cooperative Extension Service, Institute of Food and Agriculture Sciences, EDIS.
Winnenburg R, et al. 2006. PHI-base: a new database for pathogen host interactions. Nucleic Acids Res. 34:D459–D464.
Xiao G, et al. 2012. Genomic perspectives on the evolution of fungal entomopathogenicity in *Beauveria bassiana*. Sci Rep. 2:483.
Xu C, Min J. 2011. Structure and function of WD40 domain proteins. Protein Cell 2:202–214.
Yin Y, et al. 2012. dbCAN: a web resource for automated carbohydrate-active enzyme annotation. Nucleic Acids Res. 40:W445–W451.
Zhang Y, et al. 2009. Mitogen-activated protein kinase hog1 in the entomopathogenic fungus *Beauveria bassiana* regulates environmental stress responses and virulence to insects. Appl Environ Microbiol. 75: 3787–3795.
Zhang Y, et al. 2010. Requirement of a mitogen-activated protein kinase for appressorium formation and penetration of insect cuticle by the entomopathogenic fungus *Beauveria bassiana*. Appl Environ Microbiol. 76:2262–2270.
Zhang S, et al. 2012. CYP52X1, representing new cytochrome P450 subfamily, displays fatty acid hydroxylase activity and contributes to virulence and growth on insect cuticular substrates in entomopathogenic fungus *Beauveria bassiana*. J Biol Chem. 287:13477–13486.
Zhao H, et al. 2014. Host-to-pathogen gene transfer facilitated infection of insects by a pathogenic fungus. PLoS Pathog. 10:e1004009.
Zheng P, et al. 2011. Genome sequence of the insect pathogenic fungus *Cordyceps militaris*, a valued traditional Chinese medicine. Genome Biol. 12:R116.

Associate editor: Ross Hardison