Diversification of fungal chitinases and their functional differentiation in *Histoplasma capsulatum*

Kristie D. Goughenour[1*], Janice Whalin[1],
Jason C. Slot[2], Chad A. Rappleye[1#]

[1] Department of Microbiology, Ohio State University

[2] Department of Plant Pathology, Ohio State University

[#]corresponding author:
rappleye.1@osu.edu
614-247-2718

*current affiliation:
Division of Pulmonary and Critical Care Medicine
University of Michigan
VA Ann Arbor Healthcare System, Research Service
Ann Arbor, Michigan, USA

running title: Fungal chitinases

keywords: chitinase, GH18, fungi, *Histoplasma*

ABSTRACT

Chitinases enzymatically hydrolyze chitin, a highly abundant biomolecule with many potential industrial and medical uses in addition to their natural biological roles. Fungi are a rich source of chitinases, however the phylogenetic and functional diversity of fungal chitinases are not well understood. We surveyed fungal chitinases from 373 publicly available genomes, characterized domain architecture, and conducted phylogenetic analyses of the glycoside hydrolase family 18 (GH18) domain. This large-scale analysis does not support the previous division of fungal chitinases into three major clades (A, B, C). The chitinases previously assigned to the "C" clade are not resolved as distinct from the "A" clade in this larger phylogenetic analysis. Fungal chitinase diversity was partly shaped by horizontal gene transfer, and at least one clade of bacterial origin occurs among chitinases previously assigned to the "B" clade. Furthermore, chitin binding domains (CBD) including the LysM domain do not define specific clades but instead are found more broadly across clades of chitinase enzymes. To gain insight into biological function diversity, we characterized all eight chitinases (Cts) from the thermally dimorphic fungus, *Histoplasma capsulatum:* six A clade (3 A-V, 1 A-IV, and two A-II), one B clade (B-I), and one formerly classified C clade (C-I) chitinases. Expression analyses showed variable induction of chitinase genes in the presence of chitin but preferential expression of *CTS3* in the mycelial stage. Activity assays demonstrated that Cts1 (B-I), Cts2 (A-V), Cts3 (A-V), Cts4 (A-V) have endochitinase activities with varying degrees of chitobiosidase function. Cts6 (C-I) has activity consistent with N-acetyl-glucosaminidase exochitinase function and Cts8 (A-II) has chitobiase activity. This suggests chitinase activity is variable even within sub-clades and that predictions of functionality require more sophisticated models.

INTRODUCTION

Chitin is a (1,4)-β-linked N-acetyl-D-glucosamine (GlcNAc) polymer. As the second most abundant biopolymer after cellulose (Tharanathan and Kittur 2003), chitin and its deacetylated derivative chitosan are abundant sources of organic carbon and nitrogen that have many potential industrial uses, from biomedical to agricultural to water engineering (Ravi Kumar 2000; Zargar et al. 2015). Consequently, there is a great interest in identifying enzymes that efficiently hydrolyze chitin into more soluble mono- and oligomers of GlcNAc. Chitin degrading enzymes also have potential applications in the breakdown of the chitinous structures of fungal and arthropod agricultural pests (Hamid et al. 2013).

Chitinases (E.C 3.2.1.14) are glycosyl hydrolases that are found in a wide range of plants, bacteria, and fungi. For plants and bacteria, which lack chitin, chitinases play roles primarily in defense against fungi and/or arthropods. Chitin is an important structural component of the fungal cell wall, ranging from 0.5%-5% in yeasts to ≥20% in some filamentous fungi (Hartl et al. 2012). Yeast-form fungi possess relatively few chitinases (e.g., two in *Saccharomyces cerevisiae*) compared to the 10-20 chitinases encoded in some filamentous fungal genomes (Seidl 2008). The expansion of chitinases in filamentous fungi has prompted investigation of chitinase roles in formation of hyphal structures (Seidl 2008). While some studies have shown select chitinases are specifically induced during hyphal formation (Gruber, Kubicek, and Seidl-Seiboth 2011; Takaya et al. 1998), many other chitinases are not (Duo-Chuan 2006; Gruber, Vaaje-Kolstad, et al. 2011; Seidl 2008) suggesting alternative roles for diverse chitinases. For example, there is evidence that specific chitinases facilitate mycoparasitism of other fungi by *Trichoderma* (Boer et al. 2007; Cruz et al. 1992; Seidl et al. 2005).

Fungal chitinases are characterized by the presence of the gycoside hydrolase 18 family (GH18) (Seidl 2008) domain. Additional motifs often found in chitinases include: an N-terminal signal peptide region that serves for secretion of the enzyme, a serine/theronine-rich region, one or more chitin-binding domains (CBDs), and LysM domains. The LysM domain enables binding to polysaccharides such as peptidoglycan and chitin (i.e. a CBD). However, none of these motifs beyond the GH18 domain are conserved across all fungal chitinases or required for a protein to be considered a chitinase (Duo-Chuan 2006).

A three-clade classification system for fungal chitinases (clades A, B, and C) emerged from several investigations into the diversity of GH18-domain containing enzymes among fungi (Karlsson and Stenlid 2008, 2009; Seidl 2008; Seidl et al. 2005). While some studies have further divided the B clade into D and E clades, this appears to be a result of their specific datasets rather than a widely applicable feature (Junges et al. 2014). Clade A chitinases are between 40-50 kDa and were previously reported to have no CBD. These are the most well-studied of the three clades (Seidl 2008). Clade B chitinases are variable in size and also in the presence and number of CBDs. Clade C chitinases are distinguished by a significantly larger size (140-170 kDa) due to extension at the C-terminus. Clade C chitinases were also proposed to be defined by the presence of multiple CBDs, and especially LysM motifs (Seidl et al. 2005; Karlsson and Stenlid 2008; Karlsson and Stenlid 2009). Clade C chitinases are the least well characterized of all the clades; few members have been included in previous phylogenetic

89 studies, and none have been functionally characterized (Duo-Chuan 2006; Seidl 2008).
90 However, this three clade classification of fungal chitinases is based on a limited number of
91 fungal genomes, and it remains to be determined if it is robust to more in-depth taxon sampling
92 (Seidl et al. 2005; Karlsson and Stenlid 2008; Karlsson and Stenlid 2009).

93 While many fungal chitinases have been transcriptionally profiled, the characterization of fungal
94 chitinase enzymatic specificity is particularly limited, leaving broad assumptions about clade-
95 specific functions untested. For example, Clades A and C are class V (fungal/bacterial)
96 chitinases, which are generally assumed to be exochitinases based on modeling of the
97 conserved binding groove as deep and tunnel shaped (Duo-Chuan 2006; Seidl 2008; Hartl et al.
98 2012). Clade B corresponds to class III (fungal/plant) chitinases and are predicted to be
99 endochitinases due to the modeling of their binding grooves as shallow and open (Duo-Chuan
100 2006; Seidl 2008; Hartl et al. 2012). While some of these assumptions are supported by the
101 activities of specific chitinases (i.e., the CHIT33 and CHIT42 chitinases of *Trichoderma*
102 *harzianum* (Boer et al. 2007; Lienemann et al. 2009)), these chitinases also have multiple
103 activities and further examples need to be studied to establish clade-defning characteristics
104 (Hartl et al. 2012). The inconsistent use of diverse chitin substrates (e.g., crustacean chitin or
105 fungal chitin) further confounds the reliability of assumptions about functional diversity within
106 and among clades. While there are reports of complete chitinase or clade/sub-clade specific
107 chitinase transcriptional or deletion studies (Alcazar-Fuoli et al. 2011; Dünkler et al. 2005;
108 Gruber, Vaaje-Kolstad, et al. 2011; Gruber, Kubicek, et al. 2011), enzymatic studies have not
109 been as comprehensive.

110 In this study, we identified the GH18 domain proteins encoded in 373 published fungal genomes
111 to conduct a more complete analysis of their distribution and evolution. In addition, we provide
112 initial expression analysis and enzymatic characterization of the chitinase enzymes produced by
113 the thermally dimorphic fungus, *Histoplasma capsulatum.* This organism has distinct and tightly
114 controlled morphologies (mycelia or yeasts), each of which has a specific lifestyle
115 (environmental saprobe or pathogen of mammals, respectively) potentially allowing for
116 identification of specific roles for different chitinases. These chitinases provide new enzymatic
117 examples from each of the major clades, including the enzymatic activity of the first
118 characterized clade C chitinase.

RESULTS

Diversification of fungal GH18 domains

We identified 3,888 GH18 domain-containing proteins in 373 publicly available fungal genomes (Supplemental Data 1) using an HMM search (Eddy 2009). 494 of these proteins contain the LysM chitin binding domain (Supplemental Data 2) considered to be characteristic of the C clade chitinases in fungi (Seidl et al. 2005; Seidl 2008). 1,250 CBDs (i.e., ChtBD1, ChtBD3, COG3979, CBM_1 and CBM_19) were also identified in the GH18 domain containing proteins (Marchler-Bauer et al. 2017).

Alignment of the GH18 domains of these proteins was 1,840-characters long; however, the C-terminus of the alignment is not well-conserved and occasionally absent particularly among members of the B clade (including functionally validated chitinases (Dünkler et al. 2005; Hurtado-Guerrero and van Aalten 2007)). Maximum likelihood phylogenetic analysis using IQ-Tree was used to establish evolutionary relationships among the fungal chitinases and to infer support for clades and subclades defined by the common ancestor of previously classified chitinases, with some slight expansion for closely related but previously unstudied taxa (Figure 1, Table 1, Supplemental Data 3). This identified the previously reported A+C clade, although this was not well-supported (70% of rapid bootstraps (RB)). The B clade was not supported. Constituent sub-clades, determined by identifying the most recent common ancestor of previously categorized chitinase sub-clades (Seidl et al. 2005; Karlsson and Stenlid 2008; Karlsson and Stenlid 2009), were generally recovered with variable support. Further expansion of sub-clades recovers more strongly supported nodes without encroaching on other sub-clades (Figure 1).

In order to retain more alignable characters and thereby reduce error introduced by homoplasy for finer resolution within clades and sub-clades, we separately analyzed sequences on either half of a bifurcation between the A+C clades and an assemblage containing all the B sub-clades in the complete tree (Figure 2). In the complete GH18 tree (Figure 1), we did not recover a B-I sub-clade and only the B-II and B-IV sub-clades received strong rapid bootstrap (RB) support (Table 1). However, in the IQ-tree focused on B sub-clades, B-I (98% RB), B-II (100% RB), and B-V (97% RB) clades were supported, while clades B-III and B-IV were not supported with the current strict definition of sub-clades, and a B clade remained unsupported (Figure 2A and Table 1). However, slight expansions of the sub-clades identified supported nodes consistent with the previous analyses. A RAxML analysis of this alignment also recovered these B sub-clades but without RB support, suggesting the finer scale topology is not always robust to differences in methodology (Table 1 and Supplemental Data 3).

For the branch in the entire GH18 tree containing the A and C clades (hereafter referred to as the “AC” clade), we recovered distinct, yet unsupported A-IV (58% RB) and A-V (50% RB) sub-clades after reclassifying one *Trichoderma reesei* gene (Chi18-5) from an A-V to an A-IV (Figure 1 and Table 1). In the AC-only IQ-Tree (Figure 2B), the A-V sub-clade was not monophyletic. The A-II (100% RB, IQ-Tree) and A-III (100% RB, IQ-Tree) sub-clades were supported in all analyses (Figure 1, Figure 2B, and Table 1). The A-IV and A-V sub-clades were separated

159 (98%RB, IQ-Tree) from the A-II and C clades (Figure 1). While the C group is monophyletic
160 within the AC clade (Figure 1), separation into sub-clade C-I was supported (99% RB) while
161 sub-clade C-II was not (89% RB) in the complete tree (Figure 1, Table 1, Supplemental Data 3).
162 In the separate AC tree C-II is not supported but would be supported by including sister
163 sequences (Figure 2B). Interestingly, in our analysis clade C (both C-I and C-II) groups closely,
164 with strong support, with A-II (97% RB in the all GH-18 IQ-Tree) to the exclusion of other A sub-
165 clades (Figure 1 and Supplemental Data 3).

166 **Taxonomic distribution of GH18 chitinases**

167 Each sub-clade hosts a distinct assortment of fungal taxa. The Ascomycota, particularly
168 Leotiomycetes, Sordariomycetes, Eurotiomycetes, and Dothideomycetes, have the largest
169 number of chitinases and highest diversity of chitinase sub-clades (Figure 3, Supplemental Data
170 4). B sub-clades-I, -II, and -IV are composed almost entirely of Ascomycota. The individual
171 orders Malasseziales, Glomerales, Monoblepharidales, Neocallimastigales, and Rozella have
172 the fewest chitinases per genome (e.g. 1 or 0.5 (Rozella)) and Auriculariales, Basidiobolales,
173 Geastrales, and Xylariales have the most (greater than 17). The B-IV sub-clade consists entirely
174 of Saccharomycetales, while the B-I and B-II sub-clades are composed of non-
175 Saccharomycetales, Ascomycota. B-I consists of Leotiomycetes, Agaricomycetes,
176 Dothideomycetes, Eurotiomycetes, Leotiomycetes, Sordariomycetes, and Xylonomycetes, with
177 multiple orders of Dothideomycetes, Eurotiomycetes and Sordariomycetes represented. B-II
178 contains multiple orders of Eurotiomycetes, Dothideomycetes, Leotiomycetes, and
179 Sordariomycetes. Sub-clade B-V contains a large number of Basidiomycota, however several
180 species of Ascomycota and a single early diverging fungus (*Basidiobolus meristosporus*) are
181 also represented. B-III is composed almost entirely of Basidiomycota with the exception of
182 sequences from the early diverging Mucorales.

183 Within the AC clade, sub-clades A-II and A-IV are composed mostly of Ascomycota with
184 isolated Basidiomycota (Tremellales in A-II and Pucciniales in A-IV) (Figure 3). The A-II sub-
185 clade consists of Eurotiomycetes, Sordariomycetes, Leotiomycetes, and Dothideomycetes,
186 although within the Dothideomycetes the order Capnodiales and an unnamed order in
187 Pleosporomycetidae were not represented. Only the Dothideales, Orbiliales, and Lecanorales
188 were not represented in A-IV. A-III is composed of both Basidiomycota and early diverging
189 fungal chitinases with a few Ascomycota (Sordariomycetes). Malasseziales was the only order
190 of Basidiomycota with no A-III chitinases while in the early diverging fungi Glomerales,
191 Blastocladiales, and Rozella also lacked these chitinases. The A-V sub-clade contains
192 Ascomycota, Basidiomycota and early diverging fungi. All Ascomycota and Basidiomycota were
193 represented in this clade except one member of an unnamed class in Taphrinomycotina, and
194 among early diverging fungi, they are only found in Rozella, Mucorales, Glomerales, and
195 Basidiobolales. The C-I sub-clade mostly contains Ascomycota with a small assortment of
196 sequences from 6 orders of Agaricomycetes while the C-II sub-clade contains both Ascomycota
197 and a few early diverging fungi (Kickxellales, Entomophthorales, and Blastocladiales). Relatively
198 few chitinases are found in Botryosphaeriales.

199 **Patterns of GH18 chitinase diversification**

200 Much of the GH18 chitinase phylogeny is consistent with vertical inheritance. More derived
201 branches of the chitinase phylogeny frequently track the species phylogeny, particularly in the A
202 sub-clades (A-V, A-IV, A-III, and A-II) and the B-V sub-clade. In keeping with the fungal
203 phylogeny, the B-V sub-clade contains distinct Agaricomycetes, Eurotiomycetes, and
204 Dothideomycetes groups, and a single species of Leotiomycetes within a Sordariomycetes
205 group. Similarly, the A-IV sub-clade contains distinct clades of Sordariomycetes,
206 Eurotiomycetes, Leotiomycetes, and Dothideomycetes. The A-II sub-clade has large separate
207 Sordariomycetes and Dothideomycetes groups. The A-V sub-clade contains a distinct
208 Saccharomycetes clade, but there are also multiple Agaricomycetes, Sordariomycetes, and
209 Eurotiomycetes clades. There are large-scale divergences from the fungal tree, including a
210 clade of Sordariomycetes within a clade otherwise composed of Agaricomycetes in the A-III
211 sub-clade. Despite such grouping of related species, the branching order within GH18
212 phylogenies is overall complex, suggesting domain and gene duplications are common. Some
213 sub-clades in particular (e.g. the C sub-clades) show few taxon-specific clades (Figure 3).

214 Part of the contrast with the species tree comes from specific supported instances of horizontal
215 gene transfer (HGT) (Supplemental Figure S1). For example, two clades of mainly Hypocreales
216 (Pezizomycotina) fungi were initially poorly-placed among the B sub-clades in the GH18
217 phylogenies. A re-analysis including fungal and non-fungal sequences at NCBI more confidently
218 placed these sequences among bacterial chitinases (Figure 1), suggesting the clades originated
219 through HGT, further reducing support for a B clade. One of these chitinases appears to have
220 been acquired specifically by insect-pathogenic Hypocreales fungi (e.g. XP_006673222.1
221 *Cordyceps militaris*) from bacteria of an unknown lineage (Supplemental Figure S1A). This *C.*
222 *militaris* chitinase is an endo-N-acetylhexosaminidase active on fucose-containing N-glycans
223 (Huang et al. 2018). The other chitinase, found in a much larger group of fungi in Hypocreales
224 and additional fungi in Pezizomycotina is supported to have been transferred from
225 Actinobacteria to Hypocreales (e.g. XP_006670951.1 *Cordyceps militaris*, Supplemental Figure
226 S1B) and later to other fungi. This single-domain chitinase is most similar to chitinase D
227 (cd02871), which has hydrolytic and transglycosylation activity in bacteria (Vaikuntapu et al.
228 2016). Two other HGTs of chitinase D were supported (Supplementary Figure S1C), one from
229 Streptomyces (Actinobacteria) to Arthrodermataceae (Eurotiomycetes), and the other from
230 Micromonadaceae (Actinobacteria) to Auriculariales (Agaricomycetes). (Vaikuntapu et al. 2016).
231 Still another group of insect-associated Hypocreales chitinases (e.g. KOM21536.1
232 *Ophiocordyceps unilateralis*, Supplemental Figure S1D) is dispersed in a clade of fungal
233 chitinases otherwise composed of multiple paralogs from the arthropod-associated
234 Kickxellomycotina. These chitinases are part of the expanded A-III sub-clade, and contain a
235 GH18 domain and a signal secretion peptide. Outside Hypocreales, the amphibian gut symbiont
236 *Basidiobolus meristosporus* appears to have acquired a B-V chitinase with only a GH18 domain
237 (Basme2_176417, Supplemental Figure S1E) from Agaricomycetes, possibly in Auriculariales,
238 but a specific donor is not supported. Finally, another inter-phylum transfer appears to have
239 occurred from Pezizomycotina to Panaeolus (Pancy2_12872, Agaricales, Supplemental Figure
240 S1F). The nearest sequence, in *Uncinocarpus reesei* (Uncre1_6593), associates this last event
241 with the dung decay niche. This is a C-II chitinase with 2 LysM and 1 CBD domain. Analyses

242 comparing topological constraints were able to reject monophyly of the hypothetical donor
243 groups for most, but not all HGTs, due to the complexity of the overall phylogeny and an
244 absence of a reliable outgroup in some cases (Supplementary Data 5).

245 **Evolution of domain architecture**

246 Chitinases with multiple GH18 domains were rare, observed in only 11 cases. Of these, all
247 contained two GH18 domains except *Psilocybe cyanescens* [Psicy2_12260], which contained
248 three. All multi-GH18 domain proteins were in Basidiomycota; most were in Agaricales (9 of 11
249 cases) and there was one each in Polyporales and Pucciniales. Multi-GH18 domain proteins are
250 not constrained to a specific sub-clade, and are found in the A-III (6 cases), B-III (3 cases) and
251 B-V (2 cases) sub-clades. Interestingly only 5 GH18 domains in multi-GH18-containing proteins,
252 all in B sub-clades, were predicted to be active. In addition, both GH18 domains in multi-domain
253 A-III chitinases lacked conserved active site residues.

254 LysM domains occur in 4 chitinase sub-clades. Most are restricted to the C-II sub-clade (Figure
255 3 and Supplemental Data 2), which also includes the previously described LysM containing
256 class III Hce2 (Homologs of *C. fulvum* Ecp2) effector proteins (Stergiopoulos et al. 2012). Some
257 LysM domain-containing proteins formed a monophyletic group within the A-V sub-clade, and
258 their LysM domain is similar to bacterial spore assembly proteins. Almost all other remaining
259 LysM-containing chitinases, which include the C-I, one A-V member (Talma12_7383), and a B-I
260 (Spoth2_113450) protein, share a different recent LysM common ancestor. There is evidence of
261 additional LysM duplications of various ages that have resulted in phylogenetic diversity among
262 LysM domains in the same protein (Supplemental Data 2), however the amino acid alignment is
263 lacking sufficient characters to infer robust relationships among most LysM domains.

264 Chitin binding domains (CBDs) are dispersed across chitinases in early diverging fungi,
265 Ascomycota and Basidiomycota (Figure 3 and Supplemental Data 2). Zoopagomycota
266 (particularly Basidiobolomycetes) have comparatively large numbers of CBDs, and other early
267 diverging fungi (Mucoromycotina and Kickxellomycotina) have somewhat fewer. CBDs are
268 widely dispersed through the classes and orders of Ascomycota, but in Basidiomycota, they are
269 mostly restricted to Agaricomycetes (Figure 3). The earliest diverging B sub-clade, B-V, has a
270 low percentage of chitinases with CBDs, while there are more of these domains in other sub-
271 clades, particularly the B-III clade (Figure 3). In the AC super clade, most CBDs are found in the
272 C-I and C-II subclades as well as in a low percentage of A-V class proteins (Figure 3).

273 ***H. capsulatum* chitinases**

274 <u>Diversification and taxonomic distribution</u>

275 The *H. capsulatum* genome encodes eight chitinase (Cts) enzymes, which are widely distributed
276 in the GH18 phylogeny (Figure 4 and Supplemental Data 6). Six genomes from diverse *H.*
277 *capsulatum* strains were used for analysis (*H. capsulatum* G186AR, *H. capsulatum* G217B, *H.*
278 *capsulatum H143*, *H. capsulatum* H88, *H. capsulatum* NAM1, and *H. capsulatum* TMU). While
279 the genomes generally contain the same chitinases, there are slight differences in clade
280 distribution. *H. capsulatum* H143 and H88 (two representatives of the African strains) do not

281 contain a C-I chitinase while G217B has two that are 100% identical (likely the result of an
282 assembly error). *H. capsulatum* TMU lacks the otherwise conserved A-IV chitinase, while *H.*
283 *capsulatum* H143 is also missing the B-I chitinase. We cannot at this time rule out the absence
284 of individual chitinases being due to errors in genome assembly (particularly with the low
285 coverage of H143 at only 3-fold).

286 Cts1 is the only *H. capsulatum* sequence in the B-I sub-clade. Cts6 is the sole C sub-clade
287 protein (C-I) and also has the larger size characteristic of this sub-clade. The remaining *H.*
288 *capsulatum* chitinases (Cts2, Cts3, Cts4, Cts5, Cts7, and Cts8) are all members of clade A
289 (Figure 4). Cts2, Cts3, and Cts4 are members of the A-V sub-clade, Cts5 and Cts8 both belong
290 to sub-clade A-II, and Cts7 is a member of sub-clade A-IV. Cts4 appears to be a very recent
291 duplication of Cts2 as the closely related fungus *Blastomyces dermatitidis* has only a single Cts
292 protein that is orthologous to both *H. capsulatum*'s Cts2 and Cts4 proteins. *Histoplasma* is the
293 only genus in the Onygenales that contains chitinases in the A-II sub-clade, while other
294 Onygenales tend to contain B-II sub-clade members, which *Histoplasma* lacks. There are no A-
295 III chitinases in *H. capsulatum,* which is consistent with other Onygenales surveyed (Figure 3).

296 *H. capsulatum* has only a single C-clade chitinase enzyme (C-I) even though the C clade (C-I
297 and C–II) is expanded in multiple other Onygenales members. Among Eurotiomycetes,
298 Onygenales and Eurotiales have similar distributions of chitinases among the sub-clades except
299 as noted above. Chaetothyriales differ in that they contain A-III sub-clade chitinases but lack the
300 B-II and C clade chitinases found in the other Eurotiomycetes (Figure 3).

301 <u>Domain architecture and finer motifs</u>

302 The architectural diversity of *H. capsulatum* chitinases is consistent with the variability observed
303 across the fungal kingdom. In particular, *H. capsulatum* chitinases look similar to those of other
304 Onygenales. For example, other Onygenales have Cts3-like proteins with matching architecture
305 (i.e, containing a LysM domain but no signal peptide). Most chitinases only have 1-2 CBDs even
306 in C clade proteins. *Microsporum canis* has expanded numbers of CBDs and LysM domains
307 corresponding to their expanded C clade, which is unusual for the Onygenales chitinase
308 architecture. At the individual protein level, domain architecture is variable among the A clade
309 chitinases. Cts2, Cts3, and Cts4 are all in the A-V sub-clade. Of these, Cts2 and Cts4 have
310 similar domain architecture consistent with a recent duplication, with secretion signals at the N-
311 terminus and serine/threonine rich regions (Figure 4), which are common sites for O-linked
312 glycosylation (Loibl and Strahl 2013). Cts3 contrasts with Cts2 and Cts4 by lacking the secretion
313 signal and containing the only LysM domain among *H. capsulatum* chitinases. In sub-clade A-
314 IV, Cts7 lacks a secretion signal and contains a serine/threonine rich region. This is also the
315 only *H. capsulatum* chitinase that lacks conserved aspartate residues predicted to comprise the
316 active site (Hartl et al. 2012). In sub-clade A-II, Cts5 and Cts8 both contain a secretion signal,
317 but only Cts5 contains a recognizable CBD. The C-1 sub-clade chitinase, Cts6, possesses the
318 N-terminal secretion signal, serine/threonine rich regions, and two CBDs. The B clade chitinase
319 in *H. capsulatum,* Cts1 (B-I sub-clade), has a GPI anchor region in addition to a secretion signal
320 suggesting this enzyme is exported and anchored to the cell surface of *H. capsulatum* cells. It
321 also contains an extended serine/threonine rich region common among extracellular proteins.

322 Thus, most *H. capsulatum* chitinases are predicted to be secreted enzymes and active on chitin
323 substrates.

324 Expression of *H. capsulatum* chitinases

325 To provide insight into the physiological roles of diverse chitinases in *H. capsulatum*, the
326 expression of each was determined under different environmental and nutritional conditions. As
327 a thermally-controlled dimorphic fungus, *H. capsulatum* provides an opportunity to ascribe
328 morphology-specific (yeast- or mycelial-) roles to individual chitinases. Surprisingly, most
329 chitinase-encoding genes (*CTS*) were expressed at constant, but low levels across most
330 conditions including in the presence of exogenous chitin (Figure 5). Only *CTS3* (one of the A-V
331 chitinases) was specifically up-regulated in *H. capsulatum* filamentous cells (on average roughly
332 11-fold higher in mycelia compared to yeasts; Figure 5) suggestive of mycelia-specific functions.
333 *CTS2* (encoding another A-V-class chitinase) was expressed at higher levels overall suggesting
334 that the Cts2 protein could be important for general growth or the major functional chitinase
335 under the tested conditions. *CTS4* (A-V) and *CTS8* (A-II) were expressed at low but consistent
336 levels. *CTS1* (the only B-I chitinase) showed variable expression. *CTS6* (the C-I chitinase) also
337 showed highly variable levels of expression that somewhat mirrored that of *CTS1* (Figure 5).
338 *CTS7* (the only A-IV chitinase), which lacks the active site D-X-X-D-X-D-X-E residues (Figure
339 4), showed very low levels of expression. *CTS5* (A-II) expression was not detectable under any
340 condition tested. Interestingly, the presence of exogenous chitin in the media did not
341 consistently induce chitinase gene expression, with the exception of *CTS6* which was induced in
342 the presence of chitin in most conditions except yeast-phase growth on rich media (Figure 5).
343 Regardless, neither *H. capsulatum* yeast nor mycelia were able to grow on chitin as the carbon
344 source of the growth media (data not shown). Thus, with the exception of *CTS3* and *CTS6*,
345 expression studies did not reveal specific environmental conditions for *H. capsulatum* chitinase
346 expression or a yeast-phase specific chitinase.

347 Enzymatic activities of *H. capsulatum* chitinases

348 To determine if the clades represented by the *H. capsulatum* chitinases (Cts proteins)
349 correspond to different enzyme activities as suggested by previous phylogenetic studies, all
350 eight *H. capsulatum* chitinases were purified and tested for chitin degradation profiles. Three
351 artificial substrate mimics were used to determine the specificity of each purified chitinase
352 enzyme (Figure 6): N-acetyl-glucosaminidase exochitinase activity (liberation of 4-
353 methylumbelliferone (4-MU) from all substrates including sequential hydrolysis of GlcNAc units
354 from the reducing end of oligosaccharides), endochitinase activity (liberation of 4-MU from
355 oligosaccharides with at least 2 GlcNAc units), chitobiosidase activity (hydrolysis of the
356 disaccharide chitobiose from the reducing end thus liberating 4-MU from 4MU-$(GlcNAc)_2$), and
357 chitobiase, which hydrolyzes only the disaccharide chitobiose (e.g., hydrolysis only of the
358 glucosidic bond in 4-MU-GlcNAc). The B-clade chitinase Cts1 exhibited endochitinase activity
359 (Figure 6) consistent with the proposed activity suggested by the more open-channel B-clade
360 enzyme structure (Hartl et al. 2012). The A-V chitinases (Cts2, Cts3, and Cts4) also showed
361 endochitinase activity and chitobiosidase function consistent with hydrolysis of internal
362 glucosidic linkages. Cts4 also had a low level of exochitinase activity similar to N-acetyl-
363 glucosaminidases suggesting neofunctionalization after duplication from Cts2. Of the other A-

364 clade proteins, only Cts8 (an A-II sub-clade protein) exhibited significant chitinase activity, which
365 was consistent with chitobiase activity (Figure 6). The other A-II sub-clade protein, Cts5 had
366 detectable but extremely poor activity on any substrate. Cts7 (sub-clade A-IV), lacked any
367 detectable activity (Figure 6), consistent with the lack of active site residues in the Cts7 protein
368 (Figure 4). Cts6, *H. capsulatum*'s only C-clade protein (C-1 sub-clade) hydrolyzed all substrates
369 consistent with N-acetyl-glucosaminidase exochitinase activity (Figure 6).

370 **DISCUSSION**

371 **Previous chitinase ontologies are largely robust to increased sampling, but the A clade is**
372 **polyphyletic.**

373 Our analyses support the existence of two major classes of chitinases defined by an ancient
374 divergence between B and AC seen in previous phylogenetic analyses. We also recapitulate the
375 previous sub-clades with varying degrees of support (Figure 1 and Table 1). However, we find
376 that there is not a single A clade, because C clade chitinases are more closely related to A-II
377 chitinases. The divergence of the A-IIIs and the close clustering of the A-II sub-clade with the C
378 clade suggest the definition of A and C clades could use revision. Since A clade chitinases are
379 not monophyletic, either C clade chitinases should be subsumed into a larger A clade, or
380 alternately, the A clade chitinases could be split into multiple new clades informed by additional
381 gene architecture and enzymatic activity investigations. Previous work has supported both an
382 independent C clade (Seidl et al. 2005; Karlsson and Stenlid 2008; Alcazar-Fuoli et al. 2011)
383 and a C + A-II clade (Karlsson and Stenlid 2009) that is consistent with our analyses. However,
384 only the present analysis strongly supports the divergence of the A-III clade from the rest of the
385 A sub-clades and the C clade, perhaps because the limited sample of A-III chitinases in
386 previous analyses (< 10 members) did not reflect overall A-III diversity. Our analysis suggests
387 additional fungal chitinases need to be characterized in order to accurately describe functional
388 diversity among chitinases and thereby more accurately inform the designation of clades and
389 sub-clades. Additionally the location of the LysM domain has been used to distinguish the C
390 clade from the A clade (Gruber, Vaaje-Kolstad, et al. 2011; Gruber and Seidl-Seiboth 2012);
391 however, in our analyses, LysM domains are found in the C-II sub-clade as well as some A-V
392 chitinases, including the *H. capsulatum* Cts3 (Figure 4). Furthermore, C-I sub-clade chitinases
393 often lack a detectable LysM domain (e.g., H*. capsulatum*'s Cts6). Thus, the presence of LysM
394 is not a C-clade-defining feature as originally proposed. Additionally, the A clade was generally
395 thought to be lacking in CBDs (Hartl et al. 2012), however CBDs are widespread in the A-II sub-
396 clade (Figure 3), further supporting the need to redefine the A and C clades.

397 **Species representation within chitinase sub-clades are explained by phylogeny and**
398 **ecology.**

399 The B vs AC split appears to be an early divergence in chitinases, preceding the origin of fungi.
400 For the B clade (GH18 class III) chitinases, previous analyses suggest that B-Vs were the first
401 to diverge after the fungal B chitinases split from bacterial chitinase (Karlsson and Stenlid 2009).
402 Of the AC (Class V) chitinases, sub-clade A-V is the most widespread throughout the taxa, with
403 members found in early diverging fungi, Basidiomycota, and Ascomycota, suggesting it
404 maintains a core function in fungal physiology. In contrast, the A-III sub-clade is greatly reduced
405 in Ascomycota, as compared to early diverging fungi and Basidiomycota, suggesting A-III
406 functions may be conditionally dispensable in Ascomycota. In general, the Pezizomycotina
407 (Ascomycota), particularly Leotiomycetes, Sordariomycetes, Eurotiomycetes, and
408 Dothideomycetes, have chitinases from the greatest diversity of sub-clades. Genomes of early
409 diverging fungi generally contain few chitinases, and these are mostly from the C-II, A-V, and
410 particularly the A-III sub-clades. *Rozella allomycis* is an early diverging species that lacks chitin,

411 but maintains a single A-V chitinase that may play a role in its obligate endoparasitism of chytrid
412 fungi (Jones et al. 2011).

413 Clearly restricted taxonomic distributions of certain chitinase sub-clades suggest they may be
414 ancestrally orthologous. For example, the B-IV sub-clade is strictly limited to the
415 Saccharomycetales. However, most B-IV characterized functions (Table 2) relate to fungal cell
416 division and morphology (Colussi, Specht, and Taron 2005; Hurtado-Guerrero and van Aalten
417 2007; Kuranda and Robbins 1991; Selvaggini et al. 2004). As cell division is fundamental, it
418 would be unusual that this function itself is performed only by chitinases in Saccharomycetales.
419 The constitutive expression of *H. capsulatum's* B-clade chitinase, and the GPI-attachment motif
420 suggestive of anchoring at the cell surface, supports a generalized function for B clade
421 chitinases in fungal cell division and growth. The B-III sub-clade may represent the
422 Basidiomycota-specific B chitinases, but as delimited it also includes Mucorales chitinases,
423 which may reflect uncertainty in deep branching order, or loss of an Ascomycota paralog. B-I
424 and B-II chitinases are limited to the Pezizomycotina and may represent a taxon-specific
425 radiation of B chitinases.

426 The fungal ecological diversity represented among chitinase clades is complex, precluding
427 simple ecological explanations for differences among chitinase repertoires; however the
428 variability in the number of chitinases in species is consistent with some expectations. For
429 example, the genomes with the highest number of chitinases include fungi with predominantly
430 insect and fungus-derived nutrition like Basidiobolus (Tabima et al. 2020) Cordyceps, and
431 Trichoderma (Supplemental Data 7), but also wood-decay fungi like Ganoderma and
432 Auricularia, which may use additional chitinases in competition or secondary decomposition. In
433 contrast we see the fewest chitinases (≤1 per sequenced genome) in fungi that are sequestered
434 from the open environment or obligately associated with plants or animals. For example, few
435 chitinases are produced by Rozella (endoparasites), Neocallimastigomycota (obligate gut rumen
436 symbionts), Glomerales (arbuscular mycorrhizae), and Malassezia (skin commensals). We
437 identified no GH18 chitinases in the obligate intracellular animal parasites Microsporidia, and
438 the GH18 function may have been subsumed by their recently described GH19 chitinases (Han
439 et al. 2016).

440 Although it was not feasible to perform a comprehensive analysis of individual molecular
441 evolution events among fungal chitinases in this dataset, there is evidence for a significant role
442 of HGT in their diversification, and these events suggest ecological roles for the transferred
443 genes (Supplemental Figure 1 and Supplemental Data 5). For example, we observed multiple
444 HGT events and gene family expansions among insect-associated fungi. As insect exoskeletons
445 are a major source of chitin (Tharanathan and Kittur 2003), the acquisition and expansion of
446 chitinases might facilitate greater chitinase production or enable degradation of diverse forms of
447 arthropod chitin under diverse environmental conditions or pathogen relationships (Karlsson and
448 Stenlid 2009). LysM domain-containing chitinases are distributed among plant pathogenic,
449 insect parasitic and saprotrophic fungi.

450 LsyM domains have been shown to bind chitin although carbohydrate-binding specificities are
451 not well-described and their role in most fungi is not well understood (Gruber, Vaaje-Kolstad, et

452 al. 2011). In our analysis, the LysM domain, although largely corresponding to the C-II clade, is
453 not a clade-defining characteristic, making prediction of the functional role of such chitinases
454 difficult (Figure 3). Some LysM-containing proteins have also been shown to bind fungal chitin,
455 contributing to fungal evasion of host plant immunity (Bolton et al. 2008; de Jonge and Thomma
456 2009). One chitinase in the single Chytrid genome analyzed, *Spizellomyces punctatus*, which
457 may be a decomposer of arbuscular mycorrhizal fungi (Paulitz and Menge 1984) contains a
458 large number of LysM domains. However, although there are many plant and animal pathogens
459 with LysM-containing chitinases, there does not appear to be a strong ecological association
460 (Stergiopoulos et al. 2012).

461 **Evolution of the *H. capsulatum* chitinases indicates a degree of differentiation and**
462 **expansion that is reflected in fungal chitinases in general**

463 Placing *H. capsulatum* in context of the other Onygenales analyzed, *H. capsulatum* was the only
464 genus that contained A-II sub-clade members. One of the most interesting features of *H.*
465 *capsulatum* chitinase evolution is the presence of two chitinases in the A-II sub-clade. A-II sub-
466 clade members are not widespread in fungal taxonomy; they are almost exclusively found in the
467 Leotiomyceta (sub-division of the filamentous Pezizomycotina). A-II chitinases are found in
468 Eurotiomycetes, Sordariomycetes, Dothideomycetes, and Leotiomycetes. In the
469 Eurotiomycetes, Eurotiales, Onygenales, and Chaetothyriales contain A-II sub-clade members.
470 However, *H. capsulatum* is the only Onygenales to contain this sub-clade. That indicates these
471 chitinases are ancestral to the Leotiomyceta, however they have been largely lost in the
472 Onygenales with the exception of *H. capsulatum*. Therefore, these may have a specific function
473 that is necessary for *H. capsulatum* biology that is lacking in other Onygenales.

474 The B-clade chitinase in *H. capsulatum* (Cts1) belongs to the B-I clade. The other Onygenales
475 tend to contain B-II sub-clade members that *H. capsulatum* lacks. In addition, the C clade
476 chitinases are expanded in multiple other Onygenales members but is restricted to one
477 chitinase in *H. capsulatum* (Cts6). This is most obvious in *Microsporum canis* with 9 C-I and 4
478 C-II members. However, upon closer examination, only 3 of each sub-clade are predicted to be
479 active (contain the active site residues) indicating that this larger clade representation may not
480 be as functionally extreme as otherwise indicated.

481 In terms of other members of Eurotiomycetes, the Onygenales and Eurotiales show matching
482 patterns of chitinases distribution in the sub-clades. Chaetothyriales differ in that they contain A-
483 III sub-clade chitinases members but are lacking B-II or C clade members seen in the other
484 Eurotiomycetes. When comparing the Eurotiomycetes to other Pezizomycotina,
485 Saccharomycetes is rather divergent. It is the only one to contain a B-IV domain, and is the only
486 order from the Pezizomycotina to be missing a B-V or C-I sub-clade chitinase from the
487 Pezizomycotina. Saccharomycetes are limited to the B-IV, C-II, A-IV, and A-V sub-clades, which
488 is less diversity seen in other Pezizomycotina. Sordariomycetes and Eurotiomycetes had
489 identical patterns of sub-clade distribution, missing the A-III proteins. Leotiomyceta members
490 (excluding Xylonomycetes) have A-II, B-I, and C-II sub-clade proteins. The A-IV clade is missing
491 in Orbiliomycetes while C-I is missing in Xylonomycetes and Pezizomycetes. The A-V sub-clade
492 is conserved in all Pezizomycotina.

493 ***H. capsulatum* chitinase expression and functional differences suggest complementary**
494 **functional roles**

495 *H. capsulatum* expression data allowed for an initial look at how multiple chitinases could
496 potentially play different roles. Cts1 is secreted and contains a GPI-attachment motif indicating
497 that it is likely anchored at the fungal cell wall for cell wall remodeling or modification during cell
498 division and growth. Cts1 had a low, but generally constant level of expression in all growth
499 conditions consistent with an essential function in fungal growth. As transcriptional data was not
500 collected on synchronized stages of division or morphology, it is possible the low expression is a
501 result of a mix of cells at different stages of replication with high and low expression. The finding
502 that Cts1 is an endochitinase active only on larger oligosaccharides (Figure 6) rather than a
503 processive exochitinase fits with a putative role in remodeling, but not complete dismantling, of
504 chitin in the fungal cell wall.

505 Cts3 was the only chitinase that was induced in the mycelial phase of *Histoplasma,* supporting a
506 functional role in the mycelial state of this organism, which is an environmental saprotroph in
507 contrast to the yeast phase that is parasitic for animals. As Cts3 lacks a secretion signal, this
508 may suggest that Cts3 plays an intracellular role in cell division, septa formation, or in the
509 preparation of the cell wall for formation of mycelial-specific structures (e.g., conidiophores). The
510 chitobiosidase activity of Cts3 suggests the existing cell wall must be extensively altered for
511 such structures. Chitinases are generally expanded in fungi characterized by filamentous
512 growth and are few in yeast-form fungi like *S. cerevisiae*, supporting the hypothesis that hyphal
513 growth or reproductive morphologies requires multiple different chitinases (Karlsson and Stenlid
514 2008). As Cts3 is the only *H. capsulatum* chitinase with a LysM domain, this may further
515 indicate that Cts3 activity is directed at “self” chitin (e.g., during formation of hyphal wall
516 structures) similar to the role of LysM-proteins in binding fungal pathogen chitin to prevent plant
517 immune responses. *H. capsulatum* dimorphism is strictly regulated (Edwards et al. 2013),
518 primarily controlled by the Ryp transcription factors (Shen and Rappleye 2017; Sil 2019) while
519 other common fungal transcription factors, such as the APSES family, appear to have very
520 limited function in dimorphism (Longo et al. 2018). Therefore, any chitinases necessary for
521 hyphal growth or mycelial structures should be closely linked to transcriptional regulation of the
522 mycelial state.

523 *CTS2*, which encodes another A-V subclade chitinase, has consistently higher expression
524 among all *H. capsulatum* chitinase genes. This suggests that under the conditions tested Cts2
525 may be the general functional chitinase while the other related chitinase, Cts4, is either simply
526 less used or used under specific, unknown conditions. Other *H. capsulatum* chitinases show low
527 or highly variable expression among diverse conditions precluding specific hypotheses as to
528 their biological roles based on gene expression.

529 As *H. capsulatum* is the only Onygenales member with A-II chitinases, the secreted Cts5 and
530 Cts8 chitinases are of particular interest. Cts8 hydrolysis activity is consistent with that of a
531 chitobiase, an enzyme that specifically hydrolyzes the disaccharide chitobiose which is
532 produced by the activity of chitobiosidases. This may indicate that the primary role of Cts8 is to
533 further hydrolyze disaccharides in the extracellular environment (such as those produced by

534 other chitinases) into GlcNAc monomers. We hypothesize that Cts8 thus plays a nutritional role
535 by liberating hexosamines from other chitin hydrolysis products. Alternatively, generation of
536 GlcNAc monomers from environmental chitin degradation products (i.e., chitobiose) may serve
537 a signaling function since adoption of the mycelial state is potentiated by free GlcNAc but not
538 glucosamine (Gilmore et al. 2013). The other A-II *Histoplasma* chitinase, Cts5, has very minimal
539 activity, which along with the lack of any detectable expression suggests that this chitinase may
540 no longer have a functional role.

541 *CTS7* expression is very low to undetectable, and Cts7 lacks any chitinase activity suggesting
542 that it also may no longer serve a functional role in *H. capsulatum* biology. Consistent with this,
543 Cts7 also lacks a signal sequence for secretion unlike the majority of the other chitinases.
544 However, at this time a non-chitin-hydrolysis role for Cts7 cannot be ruled out.

545 Cts6 is the most unique of *H. capsulatum*'s chitinases having multiple domains and an
546 increased size. Cts6 is the only C-clade chitinase produced by *H. capsulatum*, and the
547 expression pattern of *CTS6* suggests that its transcription increases in the presence of
548 exogenous chitin. Together these characteristics suggest that Cts6 may be suited for hydrolysis
549 of diverse chitin substrates in the environment. Consistent with a role in the degradation of
550 environmental chitin, Cts6 is secreted and has N-acetyl-glucosaminidase activity (Figure 6), an
551 exochitinase function which would enable it to processively hydrolyze chitin polysaccharides.

552 ***H. capsulatum* chitinase specificity demonstrates the difficulty in using phylogeny for**
553 **activity prediction**

554 Determination of the chitinase activities of the eight *H. capsulatum* chitinases highlights the
555 limitations of previous phylogenetic analyses to predict enzymatic activities. Few fungal
556 chitinases have been enzymatically characterized (Table 2) and the characterization is
557 sometimes not complete. The multiple A-V *H. capsulatum* chitinases illustrate how functional
558 variation can be present even within sub-clades, such as the exochitinase activity of Cts4 in
559 addition to the endochitinase and chitobiosidase activity of its closest paralog Cts2. When
560 compared to some other characterized A-V clade members we see a similar pattern of multiple
561 and complex activities (Table 2). The demonstrated endochitinase activity with Cts2 and Cts3 is
562 inconsistent with the previous prediction that A clade members are characterized by
563 exochitinase activity (Duo-Chuan 2006; Seidl 2008; Hartl et al. 2012). However, this sub-clade
564 also shows a strong presence of chitobiosidase activity (Table 2). The A sub-clades might have
565 functionally diverged, which may be reflected in the potential polyphyletic phylogeny of the A
566 clade chitinases. For example, the characterized members of the A-II sub-clade have varied
567 activities, not the previously suggested restriction to exochitinase activity (Table 2). In addition,
568 the A-III sub-clade is lacking in enzymatically-characterized members, and the A-IV only has
569 one characterized member (Cts7 from this study), which does not have any enzymatic activity at
570 all. As the A sub-clades are more variable in enzymatic activity than expected and the A-III sub-
571 clade is so divergent from the other AC sub-clades, these chitinases need further study to
572 uncover how the A sub-clades have diverged and how that might influence their activities.

573 This study is the first report of the enzymatic activity of a C clade member (Table 2). The
574 hydrolysis profile of *H. capsulatum* Cts6 is consistent with an N-acetyl-glucosaminidase as the
575 successive hydrolysis of GlcNAc from oligosaccharides enables this enzyme to hydrolyze all
576 substrates tested (Figure 6).

577 In contrast to the AC clade, the B clade functional data is much more consistent with
578 predictions, with characterized members of multiple clades having endochitinase or
579 chitobiosidase activity (Duo-Chuan 2006; Seidl 2008; Hartl et al. 2012). However, sub-clades B-
580 III and B-V have not been extensively investigated to determine the universality of the prediction
581 for B clade enzymes (Table 2). In addition, some of these functionally characterized enzymes
582 have not been extensively studied for substrate preference and thus a strict categorization is
583 premature.

584 The large increase in available fungal genomes is improving our understanding of the
585 evolutionary relationships among fungal chitinases. As demonstrated in this study, combination
586 of phylogenetics with empirical characterizations, including determination of enzymatic activities,
587 further refines the predictive power of fungal chitinase phylogenetic classifications. Using this
588 fungal chitinase phylogenetic framework, future studies to define the functional roles of enzymes
589 using the increased feasibility of molecular genetic tools for fungi (Wang et al. 2017;
590 Raschmanová et al. 2018; Song et al. 2019) will connect evolution of chitinases to their
591 individual and specific roles in fungal biology.

592 **MATERIALS AND METHODS**

593 **Phylogenetic Analyses**

594 Putative chitinases were retrieved from 373 published fungal genomes, including 6 *Histoplasma*
595 genomes (Supplemental Data 1) obtained from the Joint Genome Institute (JGI) website and
596 NCBI. Proteomes were independently mined for glycosyl hydrolase 18 (GH18) containing
597 proteins (El-Gebali et al. 2019) using hidden Markov models (HMM search) (Eddy 2009). A
598 CDD search was then performed on proteins containing GH18 domains to identify additional
599 carbohydrate binding domains (Marchler-Bauer et al. 2017) and LysM domains. Alignments of
600 the GH18 domains and LysM domains were performed by HMMalign which uses a Viterbi
601 algorithm to align each sequence to the given HMM (Eddy 2009). Sequences were removed if
602 they were missing alignment between positions 87-238. Poorly aligned characters were
603 removed using trimAl (v. 1.4) (Capella-Gutiérrez et al. 2009) with a gap threshold of 0.01.

604 Phylogenetic analysis of the GH18 domains was performed using maximum-likelihood methods
605 using VT+F+G4 (GH18) or WAG+R5 (LysM) model (automatically determined) in IQ-TREE
606 (Nguyen et al. 2015). This tree was used to determine the sequences in the AC clades and the
607 B clade. These sequences were then split into separate phylogenetic analyses again using
608 maximum-likelihood methods (model WAG+G4 for AC tree and PMB+F+I+G4 for 4B tree as
609 automatically determined) implemented in IQ-TREE (Nguyen et al. 2015). Statistical support for
610 all IQ-TREEs was assessed by Ultrafast bootstrap analysis using 1,000 replicates (Minh et al.
611 2013). An ultrafast bootstrap value of ≥ 95 was considered a strongly supported branch.
612 Parameters of additional phylogenetic analyses in IQ-TREE to assess support for HGT events,
613 including topology tests using the Approximately Unbiased test (Shimodaira 2002) implemented
614 in IQ-TREE are in Supplemental Data 5. In addition, maximum-likelihood methods were
615 implemented in RAxML v.(SSE3) (Stamatakis 2014), 100 alternative runs on 100 distinct trees,
616 bootstrapping (100) and best scoring ML in 1 run and GAMMA models autodetermined (WAG
617 for AC and PMB for B). Statistical support for RAxML trees was assessed by rapid
618 bootstrapping, where nodes receiving ≥60 percent of bootstraps were considered supported.

619 **Chitinase gene expression**

620 *H. capsulatum CTS* gene transcription was determined by quantitative RT-PCR (qRT-PCR).
621 The WU15 strain of *H. capsulatum*, a uracil auxotroph of the North America 2 clade was grown
622 on HMM (Worsham and Goldman 1988) or Sabouraud's Dextrose Agar (SDA) media
623 supplemented with 100 µg/mL uracil and solidified with 0.6% agarose. For tests of chitin
624 induction of *CTS* gene transcription, colloidal chitin (prepared from shrimp chitin (Rodriguez-
625 Kabana et al. 1983)) was added to media (1.2% final concentration). Media were inoculated with
626 *H. capsulatum* cells and incubated at 25 degrees (mycelial culture) for 4 weeks or 37 degrees
627 (yeasts culture) for 5-7 days. *H. capsulatum* cells were scraped from the solid media and
628 collected by centrifugation (5 minutes at 2000 x g). RNA was isolated from fungal cells by
629 mechanical disruption with 0.5 mm glass beads, extraction with RiboZol (Amresco), and
630 purification using an affinity column (Direct-zol RNA MiniPrep Plus; Research Products
631 International). Following DNA removal with DNase Invitrogen, RNA was reversed transcribed

632 (Maxima reverse transcriptase; Thermo Scientific) primed with random pentadecamers.
633 Quantitative PCR was carried out using *CTS* gene specific primer pairs (Supplemental Data 8)
634 with SYBR green-based visualization of product amplification (Bioline). Changes in *CTS* gene
635 transcript levels relative to Actin (*ACT1*) and Ribosomal Protein S15 (*RPS15*) were determined
636 using the ΔΔCt method (Schmittgen and Livak 2008) after normalization of cycle thresholds to
637 *ACT1* and *RPS15* mRNA levels. RNAs were prepared from 3 biological replicates for each
638 media/condition. Significant differences in expression ($P<0.05$) were determined by Student's t-
639 test between paired samples differing by a single parameter. Samples with transcripts below the
640 detectable limit were set to -12.00 ΔΔCt for analysis purposes.

641 **Purification of *H. capsulatum* chitinases**

642 Chitinases were purified from transgenic *H. capsulatum* yeasts overexpressing each protein.
643 Chitinase-encoding genes were amplifed by PCR from wild-type *H. capsulatum* G217B genomic
644 DNA and cloned into the *H. capsulatum* expression vector pAG38 placing the gene under
645 transcriptional control of the *H2B* constitutive promoter and fusing the chitinase to a C-terminal
646 hexahistidine tag. For Cts1, sequences encoding the putative GPI-attachment site (nucleotides
647 2141-2215) were removed to permit recovery of soluble protein. Overexpression plasmids were
648 transformed into *H. capsulatum* WU15 yeasts by *Agrobacterium*-mediated transformation
649 (Zemska and Rappleye 2012) and transformants selected by uracil prototrophy. Transformants
650 were screened by immunoblotting of culture filtrates and cellular lysates for the hexahistidine tag
651 (GnScrpt antibody A00186 to 6xhistidine). Chitinase-expressing transformants were grown in
652 liquid HMM until stationary phase and the yeasts separated from the culture supernatant by
653 centrifugation (5 minutes at 2000 x g). For Cts1, Cts2, Cts4, Cts5, Cts6, and Cts8, culture
654 filtrates were prepared by filtration of the supernatant (0.45 um pore; Millipore), concentrated
655 100-fold by ultrafiltration (10 kDa MWCO; Whatman), and the proteins exchanged into
656 phosphate-buffered saline (PBS). For Cts3 and Cts7, which lack secretion signals, lysates were
657 prepared from the yeast cells by suspension of yeasts in PBS and mechanical breakage with
658 0.5 um diameter glass beads. Debris was removed from the cellular lysate by centifiguation (10
659 minutes at 12000 x g). Hexahistidine-tagged chitinase proteins were purified from the
660 concentrated culture filtrates or cellular lysates by metal affinity chromatography (HisPur Co^{2+}
661 Resin, Thermo Fisher Scientific) and the chitinase-containing elution fractions exchanged into
662 PBS by ultrafiltration. Resultant protein concentrations were determined using a Bradford assay
663 (Sigma-Aldrich) and purity of the protein preparation determined by SDS-PAGE followed by
664 silver staining. Purified proteins were stored at −20 °C in 50% glycerol.

665 **Chitinase activity and specificity determination**

666 Chitinase enzymatic activities were determined via a fluorimetric chitinase assay kit (Sigma-
667 Aldrich-CS1030). Three artificial substrate mimics, 4-Methylumbelliferyl N-acetyl-β-D-
668 glucosaminide (4MU-(GlcNAc)), 4-Methylumbelliferyl N,N'-diacetyl-β-D-chitobioside (4MU-
669 $(GlcNAc)_2$), and 4-Methylumbelliferyl β-D-N,N',N''-triacetylchitotriose (4MU-$(GlcNAc)_3$) were
670 used to determine the activity of each chitinase. Hydrolysis of each substrate releases 4-
671 methylbelliferone, the fluorescence of which was measured using a plate reader (360 nm
672 excitation and 450 nm emission; BioTek Synergy 2). For chitinase activity reactions, varying

673 amounts of purified chitinases were added to 0.5mg/mL substrate in reaction buffer (see kit
674 protocol) and incubated at 37°C. Reactions were monitored by endpoint fluorescence at 60
675 minutes. Data is reported as nanograms of 4-methylumbelliferone released per minute per
676 nanogram of purified chitinase. The equation of the best-fit line to the data was computed and
677 compared to a standard curve of 4-methylumbelliferone. Activity assays were performed in
678 triplicate.

ACKNOWLEDGEMENTS AND FUNDING

We thank Stephanie Ray and Qian Shen for technical reviewing and editing of the manuscript. This work was supported in part by grant AI117122 to C.A.R. from the National Institutes of Health and grant DEB-1638999 to J.C.S. from the National Science Foundation. K.D.G was supported in part through a predoctoral fellowship sponsored by the NIH/NIAID award T32, 5T32AI112542-05 (MPI), a NRSA training grant administered by the Department of Microbial Infection and Immunity at Ohio State University and a postdoctoral fellowship sponsored by the NIH award T32, 5T32HL007749-27 at the University of Michigan. The funders had no role in study design, data collection and interpretation, or the submission of the work for publication.

688 **REFERENCES**

689 Alcazar-Fuoli L, Clavaud C, Lamarre C, Aimanianda V, Seidl-Seiboth V, Mellado E,
690 Latgé J-P. 2011. Functional analysis of the fungal/plant class chitinase family in
691 Aspergillus fumigatus. *Fungal Genetics and Biology* 48:418–429.

692 Baratto CM, Vanusa da Silva M, Santi L, Passaglia L, Schrank IS, Vainstein MH,
693 Schrank A. 2003. Expression and characterization of the 42 kDa chitinase of the
694 biocontrol fungus Metarhizium anisopliae in Escherichia coli. *Can. J. Microbiol.*
695 49:723–726.

696 Boer H, Simolin H, Cottaz S, Söderlund H, Koivula A. 2007. Heterologous expression
697 and site-directed mutagenesis studies of two Trichoderma harzianum chitinases,
698 Chit33 and Chit42, in Escherichia coli. *Protein Expr. Purif.* 51:216–226.

699 Bolton MD, van Esse HP, Vossen JH, de Jonge R, Stergiopoulos I, Stulemeijer IJE, van
700 den Berg GCM, Borrás-Hidalgo O, Dekker HL, de Koster CG, et al. 2008. The
701 novel Cladosporium fulvum lysin motif effector Ecp6 is a virulence factor with
702 orthologues in other fungal species. *Mol. Microbiol.* 69:119–136.

703 Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. 2009. trimAl: a tool for automated
704 alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*
705 25:1972–1973.

706 Colussi PA, Specht CA, Taron CH. 2005. Characterization of a Nucleus-Encoded
707 Chitinase from the Yeast Kluyveromyces lactis. *Appl. Environ. Microbiol.*
708 71:2862–2869.

709 Cruz J, Hidalgo-Gallego A, Lora JM, Benitez T, Pintor-Toro JA, Llobell A. 1992. Isolation
710 and characterization of three chitinases from Trichoderma harzianum. *European*
711 *Journal of Biochemistry* 206:859–867.

712 Dünkler A, Walther A, Specht CA, Wendland J. 2005. Candida albicans CHT3 encodes
713 the functional homolog of the Cts1 chitinase of Saccharomyces cerevisiae.
714 *Fungal Genetics and Biology* 42:935–947.

715 Duo-Chuan L. 2006. Review of Fungal Chitinases. *Mycopathologia* 161:345–360.

716 Eddy SR. 2009. A new generation of homology search tools based on probabilistic
717 inference. *Genome Inform* 23:205–211.

718 Edwards JA, Chen C, Kemski MM, Hu J, Mitchell TK, Rappleye CA. 2013. Histoplasma
719 yeast and mycelial transcriptomes reveal pathogenic-phase and lineage-specific
720 gene expression profiles. *BMC Genomics* 14:695.

721 El-Gebali S, Mistry J, Bateman A, Eddy SR, Luciani A, Potter SC, Qureshi M,
722 Richardson LJ, Salazar GA, Smart A, et al. 2019. The Pfam protein families
723 database in 2019. *Nucleic Acids Res* 47:D427–D432.

724 Fan Y, Fang W, Guo S, Pei X, Zhang Y, Xiao Y, Li D, Jin K, Bidochka MJ, Pei Y. 2007.
725 Increased insect virulence in Beauveria bassiana strains overexpressing an
726 engineered chitinase. *Appl. Environ. Microbiol.* 73:295–302.

727 Fang W, Leng B, Xiao Y, Jin K, Ma J, Fan Y, Feng J, Yang X, Zhang Y, Pei Y. 2005.
728 Cloning of Beauveria bassiana Chitinase Gene Bbchit1 and Its Application To
729 Improve Fungal Strain Virulence. *Appl Environ Microbiol* 71:363–370.

730 Fukamizo T, Sasaki C, Schelp E, Bortone K, Robertus JD. 2001. Kinetic Properties of
731 Chitinase-1 from the Fungal Pathogen Coccidioides immitis. *Biochemistry*
732 40:2448–2454.

733 Gilmore SA, Naseem S, Konopka JB, Sil A. 2013. N-acetylglucosamine (GlcNAc)
734 Triggers a Rapid, Temperature-Responsive Morphogenetic Program in Thermally
735 Dimorphic Fungi. *PLoS Genet* [Internet] 9. Available from:
736 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3778022/

737 Gruber S, Kubicek CP, Seidl-Seiboth V. 2011. Differential regulation of orthologous
738 chitinase genes in mycoparasitic Trichoderma species. *Appl. Environ. Microbiol.*
739 77:7217–7226.

740 Gruber S, Seidl-Seiboth V. 2012. Self versus non-self: fungal cell wall degradation in
741 Trichoderma. *Microbiology* 158:26–34.

742 Gruber S, Vaaje-Kolstad G, Matarese F, López-Mondéjar R, Kubicek CP, Seidl-Seiboth
743 V. 2011. Analysis of subgroup C of fungal chitinases containing chitin-binding
744 and LysM modules in the mycoparasite Trichoderma atroviride. *Glycobiology*
745 21:122–133.

746 Hamid R, Khan MA, Ahmad M, Ahmad MM, Abdin MZ, Musarrat J, Javed S. 2013.
747 Chitinases: An update. *J Pharm Bioallied Sci* 5:21–29.

748 Han B, Zhou K, Li Z, Sun B, Ni Q, Meng X, Pan G, Li C, Long M, Li T, et al. 2016.
749 Characterization of the First Fungal Glycosyl Hydrolase Family 19 Chitinase
750 (NbchiA) from Nosema bombycis (Nb). *J. Eukaryot. Microbiol.* 63:37–45.

751 Haran S, Schickler H, Oppenheim A, Chet I. 1995. New components of the chitinolytic
752 system of Trichoderma harzianum. *Mycological Research* 99:441–446.

753 Hartl L, Zach S, Seidl-Seiboth V. 2012. Fungal chitinases: diversity, mechanistic
754 properties and biotechnological potential. *Appl Microbiol Biotechnol* 93:533–543.

755 Hoell IA, Vaaje-Kolstad G, Eijsink VGH. 2010. Structure and function of enzymes acting
756 on chitin and chitosan. *Biotechnology and Genetic Engineering Reviews* 27:331–
757 366.

758 Huang Y, Higuchi Y, Kinoshita T, Mitani A, Eshima Y, Takegawa K. 2018.
759 Characterization of novel endo-β- N -acetylglucosaminidases from

760 Sphingobacterium species, Beauveria bassiana and Cordyceps militaris that
761 specifically hydrolyze fucose-containing oligosaccharides and human IgG.
762 *Scientific Reports* 8:1–10.

763 Hurtado-Guerrero R, van Aalten DMF. 2007. Structure of Saccharomyces cerevisiae
764 Chitinase 1 and Screening-Based Discovery of Potent Inhibitors. *Chemistry &*
765 *Biology* 14:589–599.

766 Jaques AK, Fukamizo T, Hall D, Barton RC, Escott GM, Parkinson T, Hitchcock CA,
767 Adams DJ. 2003. Disruption of the gene encoding the ChiB1 chitinase of
768 Aspergillus fumigatus and characterization of a recombinant gene product.
769 *Microbiology,* 149:2931–2939.

770 Jones MDM, Richards TA, Hawksworth DL, Bass D. 2011. Validation and justification of
771 the phylum name Cryptomycota phyl. nov. *IMA Fungus* 2:173–175.

772 de Jonge R, Thomma BPHJ. 2009. Fungal LysM effectors: extinguishers of host
773 immunity? *Trends in Microbiology* 17:151–157.

774 Junges Â, Boldo JT, Souza BK, Guedes RLM, Sbaraini N, Kmetzsch L, Thompson CE,
775 Staats CC, Almeida LGP de, Vasconcelos ATR de, et al. 2014. Genomic
776 Analyses and Transcriptional Profiles of the Glycoside Hydrolase Family 18
777 Genes of the Entomopathogenic Fungus Metarhizium anisopliae. *PLOS ONE*
778 9:e107864.

779 Karlsson M, Stenlid J. 2008. Comparative Evolutionary Histories of the Fungal Chitinase
780 Gene Family Reveal Non-Random Size Expansions and Contractions due to
781 Adaptive Natural Selection. *Evol Bioinform Online* 4:47–60.

782 Karlsson M, Stenlid J. 2009. Evolution of family 18 glycoside hydrolases: diversity,
783 domain structures and phylogenetic relationships. *J. Mol. Microbiol. Biotechnol.*
784 16:208–223.

785 Kuranda MJ, Robbins PW. 1991. Chitinase is required for cell separation during growth
786 of Saccharomyces cerevisiae. *J. Biol. Chem.* 266:19758–19767.

787 Lienemann M, Boer H, Paananen A, Cottaz S, Koivula A. 2009. Toward understanding
788 of carbohydrate binding and substrate specificity of a glycosyl hydrolase 18
789 family (GH-18) chitinase from Trichoderma harzianum. *Glycobiology* 19:1116–
790 1126.

791 Loibl M, Strahl S. 2013. Protein O-mannosylation: What we have learned from baker's
792 yeast. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*
793 1833:2438–2446.

794 Longo LVG, Ray SC, Puccia R, Rappleye CA. 2018. Characterization of the APSES-
795 family transcriptional regulators of Histoplasma capsulatum. *FEMS Yeast Res.*
796 18.

797 Marchler-Bauer A, Bo Y, Han L, He J, Lanczycki CJ, Lu S, Chitsaz F, Derbyshire MK,
798 Geer RC, Gonzales NR, et al. 2017. CDD/SPARCLE: functional classification of
799 proteins via subfamily domain architectures. *Nucleic Acids Res.* 45:D200–D203.

800 Minh BQ, Nguyen MAT, von Haeseler A. 2013. Ultrafast Approximation for Phylogenetic
801 Bootstrap. *Mol Biol Evol* 30:1188–1195.

802 van Munster JM, van der Kaaij RM, Dijkhuizen L, van der Maarel MJEC. 2012.
803 Biochemical characterization of Aspergillus niger CfcI, a glycoside hydrolase
804 family 18 chitinase that releases monomers during substrate hydrolysis.
805 *Microbiology* 158:2168–2179.

806 Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. 2015. IQ-TREE: A Fast and
807 Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies.
808 *Mol Biol Evol* 32:268–274.

809 Paulitz TC, Menge JA. 1984. Is Spizellomyces Punctatum a Parasite or Saprophyte of
810 Vesicular-Arbuscular Mycorrhizal Fungi? *Mycologia* 76:99–107.

811 Pinto A de S, Barreto CC, Vainstein MH, Schrank A, Ulhoa CJ. 1997. Purification and
812 characterization of an extracellular chitinase from the entomopathogen
813 Metarhizium anisopliae. *Can. J. Microbiol.* 43:322–327.

814 Raschmanová H, Weninger A, Glieder A, Kovar K, Vogl T. 2018. Implementing
815 CRISPR-Cas technologies in conventional and non-conventional yeasts: Current
816 state and future prospects. *Biotechnol. Adv.* 36:641–665.

817 Ravi Kumar MNV. 2000. A review of chitin and chitosan applications. *Reactive and*
818 *Functional Polymers* 46:1–27.

819 Rodriguez-Kabana R, Godoy G, Morgan-Jones G, Shelby RA. 1983. The determination
820 of soil chitinase activity: Conditions for assay and ecological studies. *Plant Soil*
821 75:95–106.

822 Rush CL, Schüttelkopf AW, Hurtado-Guerrero R, Blair DE, Ibrahim AFM, Desvergnes S,
823 Eggleston IM, van Aalten DMF. 2010. Natural Product–Guided Discovery of a
824 Fungal Chitinase Inhibitor. *Chemistry & Biology* 17:1275–1281.

825 Schmittgen TD, Livak KJ. 2008. Analyzing real-time PCR data by the comparative C(T)
826 method. *Nat Protoc* 3:1101–1108.

827 Seidl V. 2008. Chitinases of filamentous fungi: a large group of diverse proteins with
828 multiple physiological functions. *Fungal Biology Reviews* 22:36–42.

829 Seidl V, Huemer B, Seiboth B, Kubicek CP. 2005. A complete survey of Trichoderma
830 chitinases reveals three distinct subgroups of family 18 chitinases: Trichoderma
831 chitinases. *FEBS Journal* 272:5923–5939.

832 Selvaggini S, Munro CA, Paschoud S, Sanglard D, Gow NAR. 2004. Independent
833 regulation of chitin synthase and chitinase activity in Candida albicans and
834 Saccharomyces cerevisiae. *Microbiology,* 150:921–928.

835 Shen Q, Rappleye CA. 2017. Differentiation of the fungus Histoplasma capsulatum into
836 a pathogen of phagocytes. *Curr Opin Microbiol* 40:1–7.

837 Shimodaira H. 2002. An approximately unbiased test of phylogenetic tree selection.
838 *Syst. Biol.* 51:492–508.

839 Sil A. 2019. Molecular regulation of Histoplasma dimorphism. *Curr. Opin. Microbiol.*
840 52:151–157.

841 da Silva MV, Santi L, Staats CC, da Costa AM, Colodel EM, Driemeier D, Vainstein MH,
842 Schrank A. 2005. Cuticle-induced endo/exoacting chitinase CHIT30 from
843 Metarhizium anisopliae is encoded by an ortholog of the chi3 gene. *Res.*
844 *Microbiol.* 156:382–392.

845 Song R, Zhai Q, Sun L, Huang E, Zhang Y, Zhu Y, Guo Q, Tian Y, Zhao B, Lu H. 2019.
846 CRISPR/Cas9 genome editing technology in filamentous fungi: progress and
847 perspective. *Appl. Microbiol. Biotechnol.* 103:6919–6932.

848 Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-
849 analysis of large phylogenies. *Bioinformatics* 30:1312–1313.

850 Stergiopoulos I, Kourmpetis YAI, Slot JC, Bakker FT, De Wit PJGM, Rokas A. 2012. In
851 Silico Characterization and Molecular Evolutionary Analysis of a Novel
852 Superfamily of Fungal Effector Proteins. *Mol Biol Evol* 29:3371–3384.

853 Tabima JF, Trautman IA, Chang Y, Wang Y, Mondo S, Kuo A, Salamov A, Grigoriev IV,
854 Stajich JE, Spatafora JW. 2020. Phylogenomic analyses of non-Dikarya fungi
855 supports horizontal gene transfer driving diversification of secondary metabolism
856 in the amphibian gastrointestinal symbiont, Basidiobolus.
857 *bioRxiv*:2020.04.08.030916.

858 Takaya N, Yamazaki D, Horiuchi H, Ohta A, Takagi M. 1998. Cloning and
859 Characterization of a Chitinase-encoding Gene (*chiA*) from *Aspergillus nidulans*,
860 Disruption of Which Decreases Germination Frequency and Hyphal Growth.
861 *Bioscience, Biotechnology, and Biochemistry* 62:60–65.

862 Tharanathan RN, Kittur FS. 2003. Chitin--the undisputed biomolecule of great potential.
863 *Crit Rev Food Sci Nutr* 43:61–87.

864 Vaikuntapu PR, Rambabu S, Madhuprakash J, Podile AR. 2016. A new chitinase-D
865 from a plant growth promoting Serratia marcescens GPS5 for enzymatic
866 conversion of chitin. *Bioresour. Technol.* 220:200–207.

867 Wang S, Chen H, Tang X, Zhang H, Chen W, Chen YQ. 2017. Molecular tools for gene
868 manipulation in filamentous fungi. *Appl. Microbiol. Biotechnol.* 101:8063–8075.

869 Worsham PL, Goldman WE. 1988. Quantitative plating of Histoplasma capsulatum
870 without addition of conditioned medium or siderophores. *J. Med. Vet. Mycol.*
871 26:137–143.

872 Zargar V, Asghari M, Dashti A. 2015. A Review on Chitin and Chitosan Polymers:
873 Structure, Chemistry, Solubility, Derivatives, and Applications. *ChemBioEng*
874 *Reviews* 2:204–226.

875 Zemska O, Rappleye CA. 2012. Agrobacterium-mediated insertional mutagenesis in
876 Histoplasma capsulatum. *Methods Mol. Biol.* 845:51–66.

877

878 **FIGURE LEGENDS**

879 **Figure 1. Phylogenetic analysis of fungal GH18 domains**

880 Phylogenetic tree depicts relationships among 3,888 fungal chitinase proteins based on
881 alignment of their glycosyl hydrolase (GH18) domains. Trees were built using maximum-
882 likelihood methods (IQ-Tree) and branch support assessed by 1000 Ultrafast bootstrap
883 replicates (values indicated by numbers at the root of common ancestor branches). Major
884 groups corresponding to the previously named A, B, and C clades are indicated by color
885 (yellow, blue, red, respectively). Clades comprised of bacterial-like GH18 proteins are indicated
886 (orange). Collapsed sub-clades, indicated as triangles with names to the right, are defined by
887 the most recent common ancestor of previously categorized chitinases, in some cases extended
888 to supported clades consistent with the fungal phylogeny. The tree is rooted along the midpoint,
889 which generates the B vs AC clade split that has been widely accepted in the literature
890 (Karlsson and Stenlid 2008, 2009; Seidl et al. 2005). Pie charts at the far right show the fraction
891 of the constituent chitinase proteins in each clade occurring in early-diverging fungi (black),
892 Basidiomycota (gray), and Ascomycota (white).

893 **Figure 2. Phylogenetic analysis of fungal B and AC clades**

894 Relationships among constituent proteins in the B and AC clades when analyzed as separate
895 groups. Trees were built using maximum-likelihood methods (IQ-Tree) and branch support
896 assessed by 1000 Ultrafast bootstrap replicates (values indicated as percentages at the base of
897 the clade). Trees are rooted as in the complete GH18 IQ-Tree in Figure 1. Collapsed sub-
898 clades, indicated as triangles with names to the right, are defined by the most recent common
899 ancestor of previously categorized chitinases, in some cases extended to supported clades
900 consistent with the fungal phylogeny. (**A**) IQ-Tree based phylogeny of B clade chitinase proteins
901 with sub-clades indicated and collapsed (blue). (**B**) IQ-Tree based phylogeny of chtinases
902 belonging to the A (yellow) and C clades (red) showing the evolution of the C clade as a branch
903 of the A clade. As sub-clade A-V is polyphyletic in the AC tree, the ancestral nodes giving rise to
904 previously identified chitinases are indicated with red stars.

905 **Figure 3. Distribution of chitinases among fungal taxa**

906 Heat map showing the distribution of chitinase sub-clades across fungal orders. Shading
907 indicates the average number of GH18 domain-containing proteins belonging to each sub-clade
908 (columns) per genome with orders with highly expanded numbers (at least 6 per genome) of
909 sub-clade chitinases indicated (red). Taxonomic relationships among fungal orders is presented
910 for reference (left side) and colored for early-diverging fungi (blue), Ascomycota (green), and
911 Basidiomycota (red). The right two columns of the heat map indicate the number of chitinases
912 per genome with LysM or CBD domains, respectively. Lower grayscale heat map indicates the
913 percentage of the subclade proteins with LysM and CBD domains.

914 **Figure 4. Domain and motif structure of *H. capsulatum* chitinases**

915 The eight chitinase proteins (Cts) encoded in the *H. capsulatum* genome are schematically
916 shown and grouped by sub-clade. Glycosyl hydrolase family 18 (GH18) domains (green) are
917 indicated and noted as class III (GH18(III), class V (GH18(V)), or those with shortened length
918 (gh18). Recognized chitin-binding domains (CBD and LysM) are indicated (orange). N-terminal
919 signal peptides (SP; white) directing secretion of the protein, serine/threonine-rich regions (S/T;
920 pink) as potential sites for O-linked glycosylation, and the motif for glycosylphosphatidylinositol
921 attachment (GPI, yellow) are indicated. Presence and location of residues for the conserved
922 DxxDxDxE motif for chitin hydrolysis is also shown (blue stars). Sub-clade to which each protein
923 belongs is presented along the left.

924 **Figure 5. Expression of chitinase-encoding genes by *H. capsulatum* yeasts and mycelia**

925 Expression of each chitinase-encoding gene (*CTS*) by *H. capsulatum* cells was measured using
926 qRT-PCR. Transcription was determined for *H. capsulatum* cells grown on two media, one
927 based on basal cell culture medium (HMM) and one based on glucose and peptone
928 (Sauboraud's dextrose; SDA). For some tests, media was supplemented with chitin (CHIT).
929 Yeast (red bars) and mycelial (blue bars) phases were maintained by growth at 37°C or 25°C,
930 respectively. Data shows expression of each *CTS* gene relative to the expression of actin (*ACT*)
931 and organized by sub-clades (B, cyan; A, magenta; C, green). Data represent the average ±
932 standard deviation among biological replicates (n=3). Significant differences in expression
933 between yeasts and mycelia (*), absence and presence of chitin (†), or HMM compared to SDA
934 (#) are indicated for corresponding conditions differing by the single parameter. Single symbols
935 indicate $P<0.05$ and double symbols indicate $P<0.01$ as determined by pairwise Student's t-
936 tests.

937 **Figure 6. Enzymatic activity of *H. capsulatum* chitinases**

938 The enzymatic activities of purified *H. capsulatum* chitinases were tested using fluorigenic
939 GlcNAc oligomers linked to 4-methylumbelliferone (4MU) by a glycosidic bond. Chitinase
940 activity that hydrolyzes this linkage liberates 4-MU which was detected by fluorescence.
941 Substrates tested were 4MU-(GlcNAc) (black bars), 4MU-$(GlcNAc)_2$ (white bars), 4MU-
942 $(GlcNAc)_3$ (gray bars). The exochitinase activity, N-acetyl-glucosaminidase, processively
943 removes GlcNAc saccharides from the non-reducing end resulting in hydrolysis of all substrates.
944 Chitobiosidase activity (release of chitobiose from the non-reducing end) is detected as
945 hydrolysis of 4MU-$(GlcNAc)_2$. Endochitinase activity generates fluorescence from 4MU-
946 $(GlcNAc)_2$ and 4MU-$(GlcNAc)_3$. Chitobiase activity (hydrolysis of the disaccharide chitobiose)
947 generates fluorescence only from the chitobiose mimic 4MU-(GlcNAc). Activity rates were
948 calculated as nanograms of hydrolysis product (4MU) released per minute per nanogram of *H.*
949 *capsulatum* protein. Data represent the average activity ± standard deviation of replicates (n =
950 3). GH18 sub-clades to which each Cts protein belong are indicated (A, yellow; B , blue; and C,
951 red).

952 **Supplemental Figure 1. Detection of potential horizontal gene transfer occurrances**

953 Maximum Likelihood phylogenies (IQ-tree) to infer origins of potential HGT chitinases (red text).
954 (**A**) The A clade of chitinases from arthropod pathogens in Hypocreales occupy an uncertain
955 position among diverse bacteria. (**B**) The A clade of mostly Hypocreales chitinases are placed
956 within Streptomyces (Actinobacteria). (**C**) Two clades of fungal chitinases occupy separate parts
957 of the phylogeny of D-like chitinases from Actinobacteria. (**D**) The A clade of Kickxellales
958 chitinases also contains sequences from multiple hypocrealean insect pathogens. (**E**) The B-V
959 chitinase from *Basidiobolus meristosporus* (Zoopagomycota) is placed in Agaricomycetes
960 (Basidiomycota). (**F**) The C-II chitinase in *Panaeolus cyanescens* (Agaricales) is sister to a
961 chitinase in *Uncinocarpus reesei* (Onygenales). Support values indicate percentage of rapid
962 bootstraps.

963 **Supplemental Figure 2. Phylogeny of LysM domains**

964 Maximum Likelihood phylogeny (IQ-tree) of LysM domains extracted from all fungal chitinases.
965 Support values indicate percentage of rapid bootstraps.

966 **Table 1. Bootstrap support values for sub-clades by various phylogenetic analyses**

Sub-clade	GH18 IQ Tree	B or AC Clade IQ Tree	B or AC Clade RAxML
B-I	NA	98	10
B-II	100	100	75
B-III	60	69	17
B-IV	98	84	17
B-V	82	97	54
A-II	100	100	100
A-III	90	100	68
A-IV	58	67	75
AV	50	NA	11
C-I	99	91	68
C-II	89	80	36

967

968 **Table 2. Enzymatically Characterized Proteins**

Sub-clade	Taxon	Protein	Organism	Activity[1]	Reference
A-II	Eurotiales	Cfu1	*Aspergillus niger*	exochitinase (NAG)	(van Munster et al. 2012)
	Onygenales	Cts5	*Histoplasma capsulatum*	exochitinase (NAG)	(this study)
	Onygenales	Cts8	*Histoplasma capsulatum*	chitobiase	(this study)
A-III					
A-IV	Onygenales	Cts7	*Histoplasma capsulatum*	(no activity)	(this study)
A-V	Eurotiales	ChiB1	*Aspergillus fumigatus*	endochitinase	(Jaques et al. 2003)
	Hypocreales	Ech42/Chit42	*Trichoderma harzianum*	endochitinase/ chitobiosidase	(Haran et al. 1995)
	Onygenales	CiX1/Cts1	*Coccidioides immitis*	endochitinase/ chitobiosidase	(Fukamizo et al. 2001)
	Onygenales	Cts3	*Histoplasma capsulatum*	endochitinase/ chitobiosidase	(this study)
	Onygenales	Cts2	*Histoplasma capsulatum*	endochitinase/ chitobiosidase	(this study)
	Onygenales	Cts4	*Histoplasma capsulatum*	endochitinase/ chitobiosidase/ (weak exochitinase)	(this study)
	Hypocreales	CHIT42	*Metarhizium anisopliae*	endochitinase/ exochitinase (NAG)	(Baratto et al. 2003)
C-I	Onygenales	Cts6	*Histoplasma capsulatum*	exochitinase (NAG)	(this study)
C-II					
B-I	Eurotiales	ChiA1	*Aspergillus fumigatus*	not well defined*	(Rush et al. 2010)

	Hypocreales	Ech30	*Trichoderma atroviride*	endochitinase	(Hoell et al. 2010)
	Onygenales	Cts1	*Histoplasma capsulatum*	endochitinase	(this study)
B-II	Hypocreales	Chit33	*Trichoderma harzianum*	endochitinase/ chitobiosidase	(Haran et al. 1995; Boer et al. 2007)
B-III					
B-IV	Saccharomycetales	Cts1	*Sacharomyces cerevisiae*	not well defined*	(Kuranda and Robbins 1991; Hurtado-Guerrero and van Aalten 2007)
	Saccharomycetales	Chit2	*Candida albicans*	not well defined*	(Selvaggini et al. 2004)
	Saccharomycetales	Chit3	*Candida albicans*	not well defined*	(Selvaggini et al. 2004)
	Saccharomycetales	Cts1	*Kluyveromyces lactis*	endochitinase	(Colussi et al. 2005)
B-V					
“B”	Hypocreales	Chit1	*Beauveria bassiana*	endochitinase	(Fang et al. 2005; Fan et al. 2007)
D-like	Hypocreales	CHIT30	*Metarhizium anisopliae*	endochitinase/ chitobiosidase/ exochitinase (NAG)	(Pinto et al. 1997; da Silva et al. 2005)

[1] Enzymatic activity: hydrolysis of $(GlcNAc)_n$

exochitinase (NAG): N-acetyl-glucosaminidase activity where n ≥2 or greater

endochitinase: n > 2

chitobiosidase: n > 2 with liberation of the disaccharide chitobiose $(GlcNAc)_2$

chitobiase: n = 2

not well defined: limited substrates tested (e.g., only $(GlcNAc)_n$ with n>2) which does not exclude N-acetyl-glucosaminidase (exochitinase) from endochitinase activity

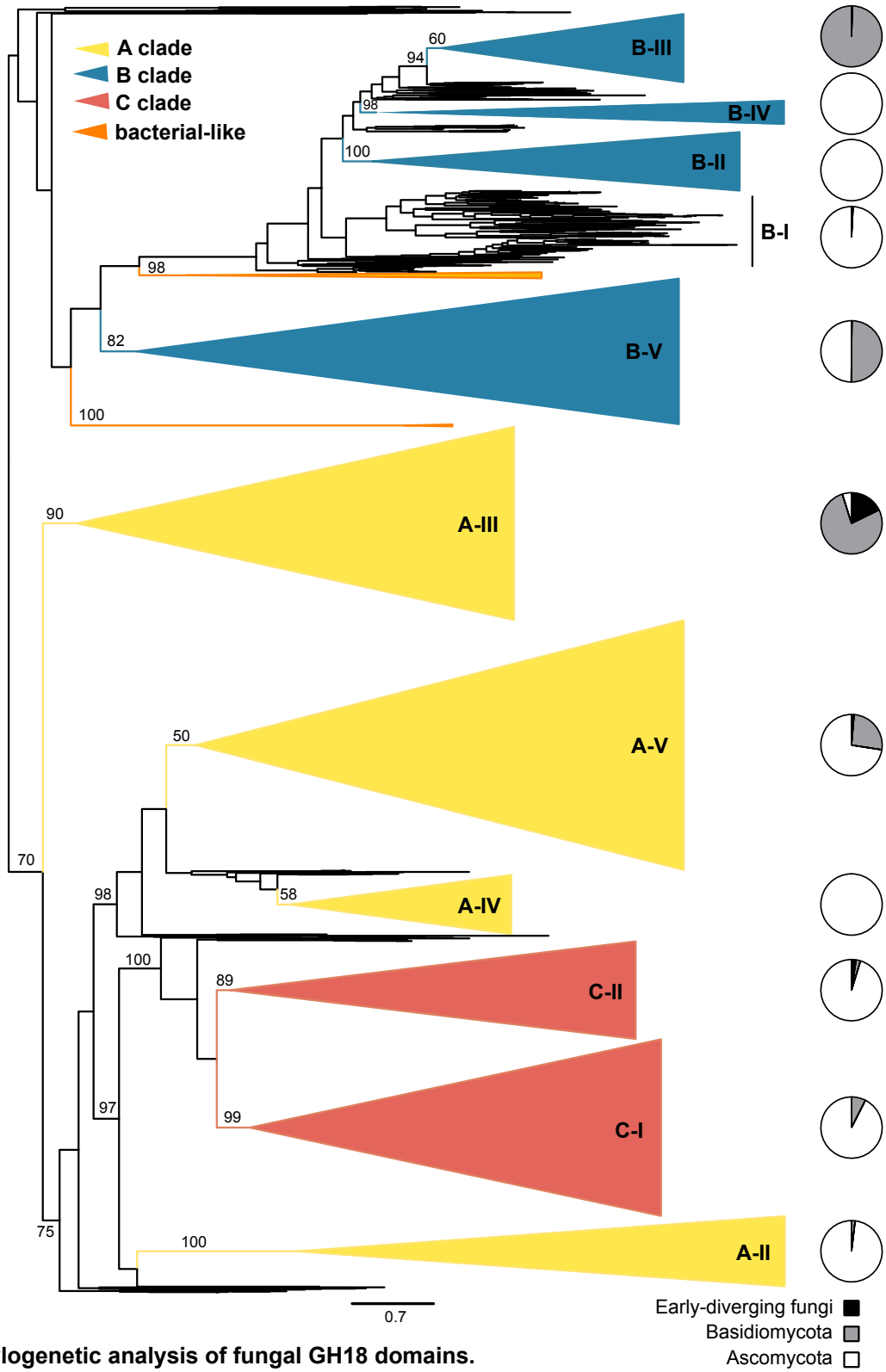

Figure 1. Phylogenetic analysis of fungal GH18 domains.

Phylogenetic tree depicts relationships among 3,888 fungal chitinase proteins based on alignment of their glycosyl hydrolase (GH18) domains. Trees were built using maximum-likelihood methods (IQ-Tree) and branch support assessed by 1000 Ultrafast bootstrap replicates (values indicated by numbers at the root of common ancestor branches). Major groups corresponding to the previously named A, B, and C clades are indicated by color (yellow, blue, red, respectively). Clades comprised of bacterial-like GH18 proteins are indicated (orange). Collapsed sub-clades, indicated as triangles with names to the right, are defined by the most recent common ancestor of previously categorized chitinases, in some cases extended to supported clades consistent with the fungal phylogeny. The tree is rooted along the midpoint, which generates the B vs AC clade split that has been widely accepted in the literature (Karlsson and Stenlid 2008, 2009; Seidl et al. 2005). Pie charts at the far right show the fraction of the constituent chitinase proteins in each clade occurring in early-diverging fungi (black), Basidiomycota (gray), and Ascomycota (white).

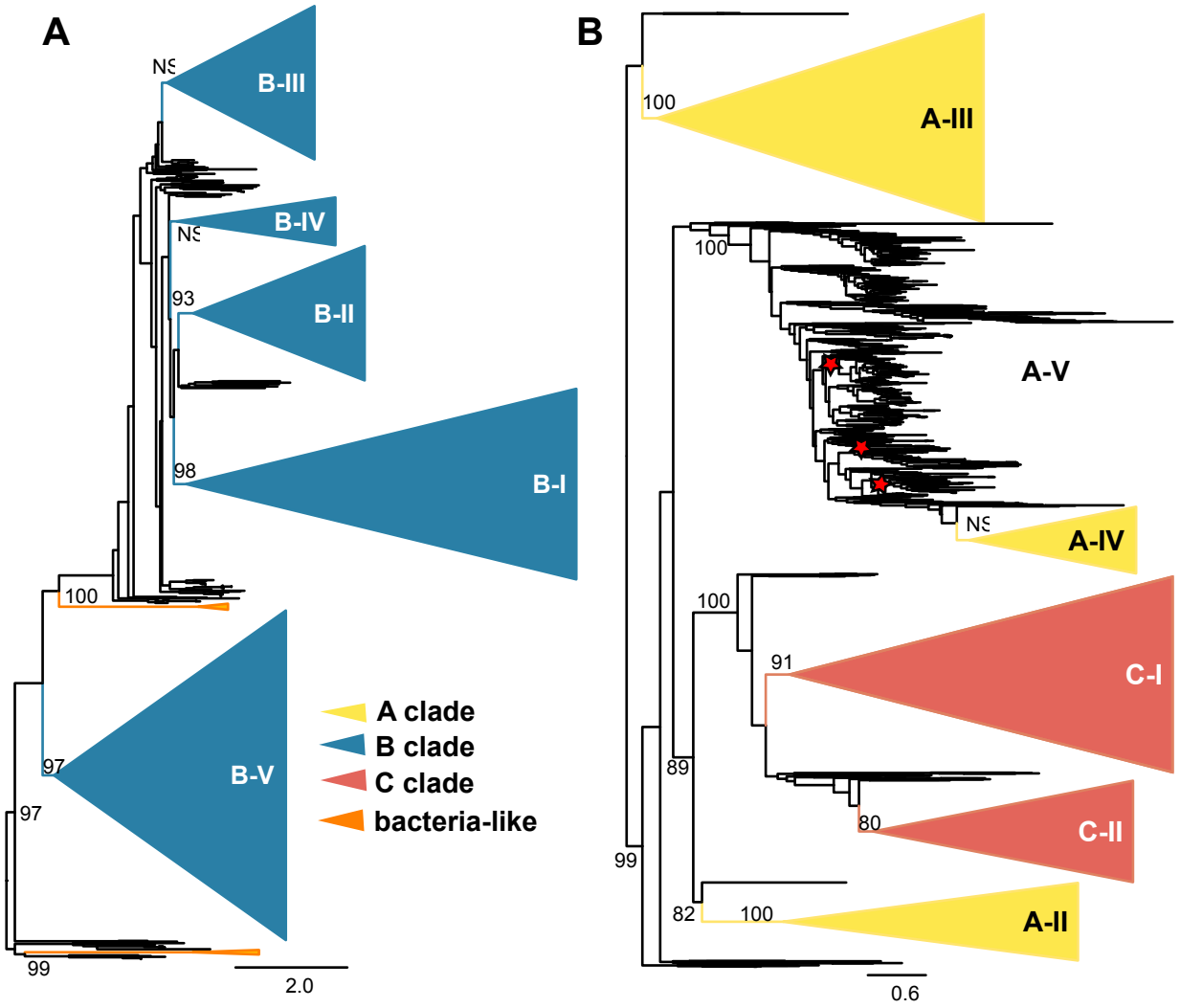

Figure 2. Phylogenetic analysis of fungal B and AC clades

Relationships among constituent proteins in the B and AC clades when analyzed as separate groups. Trees were built using maximum-likelihood methods (IQ-Tree) and branch support assessed by 1000 Ultrafast bootstrap replicates (values indicated as percentages at the base of the clade). Trees are rooted as in the complete GH18 IQ-Tree in Figure 1. Collapsed sub-clades, indicated as triangles with names to the right, are defined by the most recent common ancestor of previously categorized chitinases, in some cases extended to supported clades consistent with the fungal phylogeny. (A) IQ-Tree based phylogeny of B clade chitinase proteins with sub-clades indicated and collapsed (blue). (B) IQ-Tree based phylogeny of chtinases belonging to the A (yellow) and C clades (red) showing the evolution of the C clade as a branch of the A clade. As sub-clade A-V is polyphyletic in the AC tree, the ancestral nodes giving rise to previously identified chitinases are indicated with red stars.

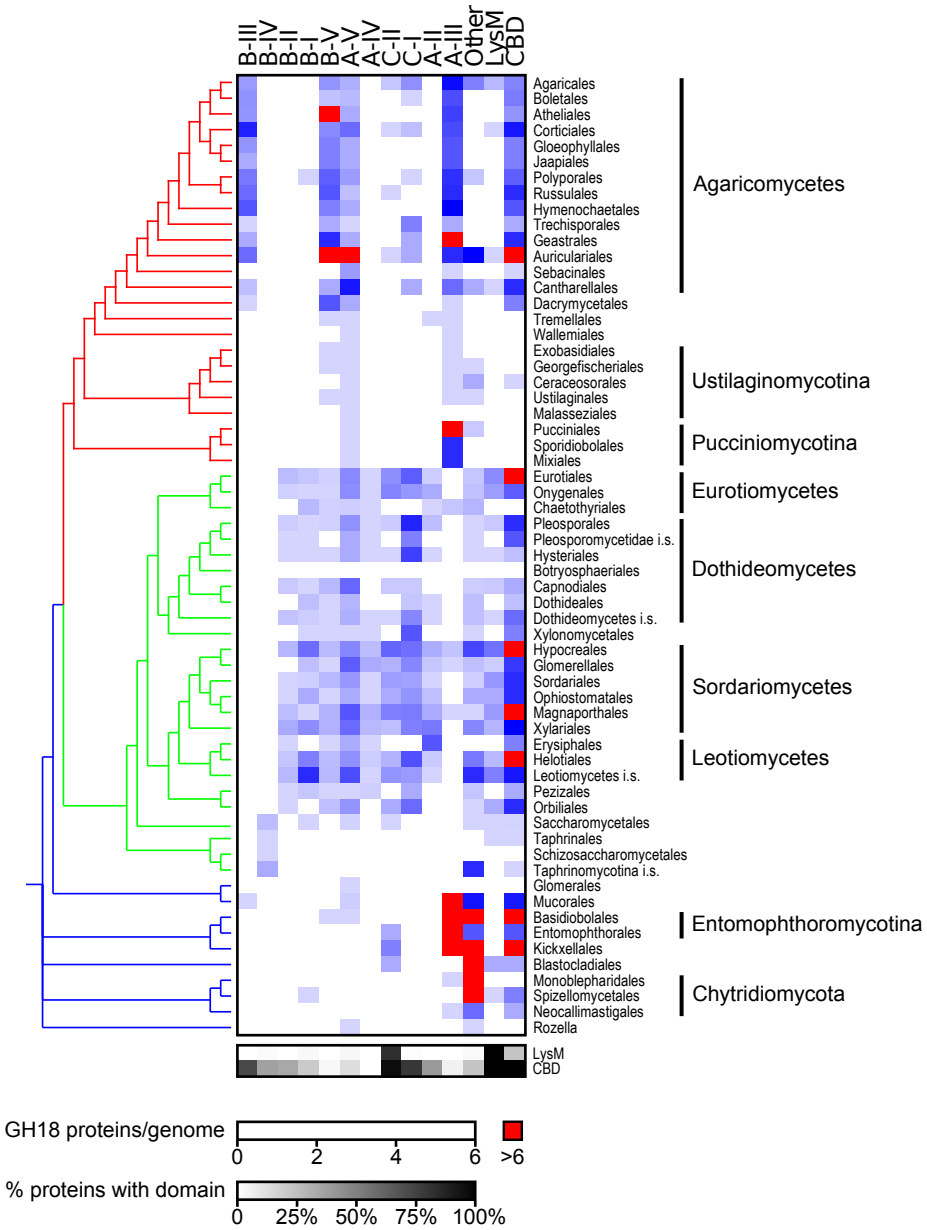

Figure 3. Distribution of chitinases among fungal taxa

Heat map showing the distribution of chitinase sub-clades across fungal orders. Shading indicates the average number of GH18 domain-containing proteins belonging to each sub-clade (columns) per genome with orders with highly expanded numbers (at least 6 per genome) of sub-clade chitinases indicated (red). Taxonomic relationships among fungal orders is presented for reference (left side) and colored for early-diverging fungi (blue), Ascomycota (green), and Basidiomycota (red). The right two columns of the heat map indicate the number of chitinases per genome with LysM or CBD domains, respectively. Lower grayscale heat map indicates the percentage of the subclade proteins with LysM and CBD domains.

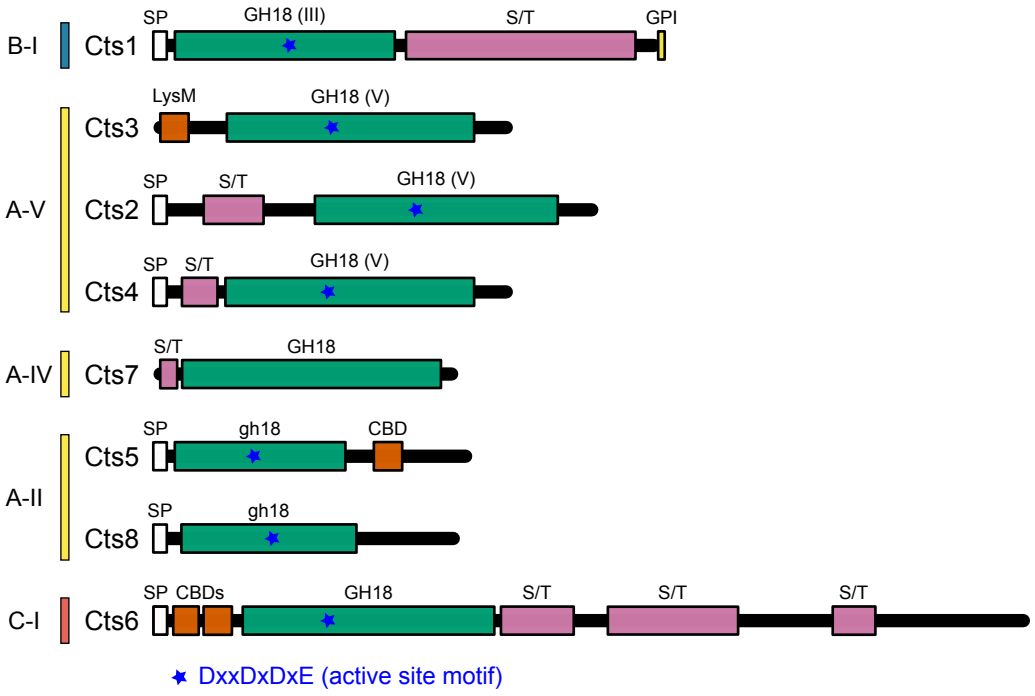

Figure 4. Domain and motif structure of *H. capsulatum* chitinases

The eight chitinase proteins (Cts) encoded in the *H. capsulatum* genome are schematically shown and grouped by sub-clade. Glycosyl hydrolase family 18 (GH18) domains (green) are indicated and noted as class III (GH18(III), class V (GH18(V)), or those with shortened length (gh18). Recognized chitin-binding domains (CBD and LysM) are indicated (orange). N-terminal signal peptides (SP; white) directing secretion of the protein, serine/threonine-rich regions (S/T; pink) as potential sites for O-linked glycosylation, and the motif for glycosylphosphatidylinositol attachment (GPI, yellow) are indicated. Presence and location of residues for the conserved DxxDxDxE motif for chitin hydrolysis is also shown (blue stars). Sub-clade to which each protein belongs is presented along the left.

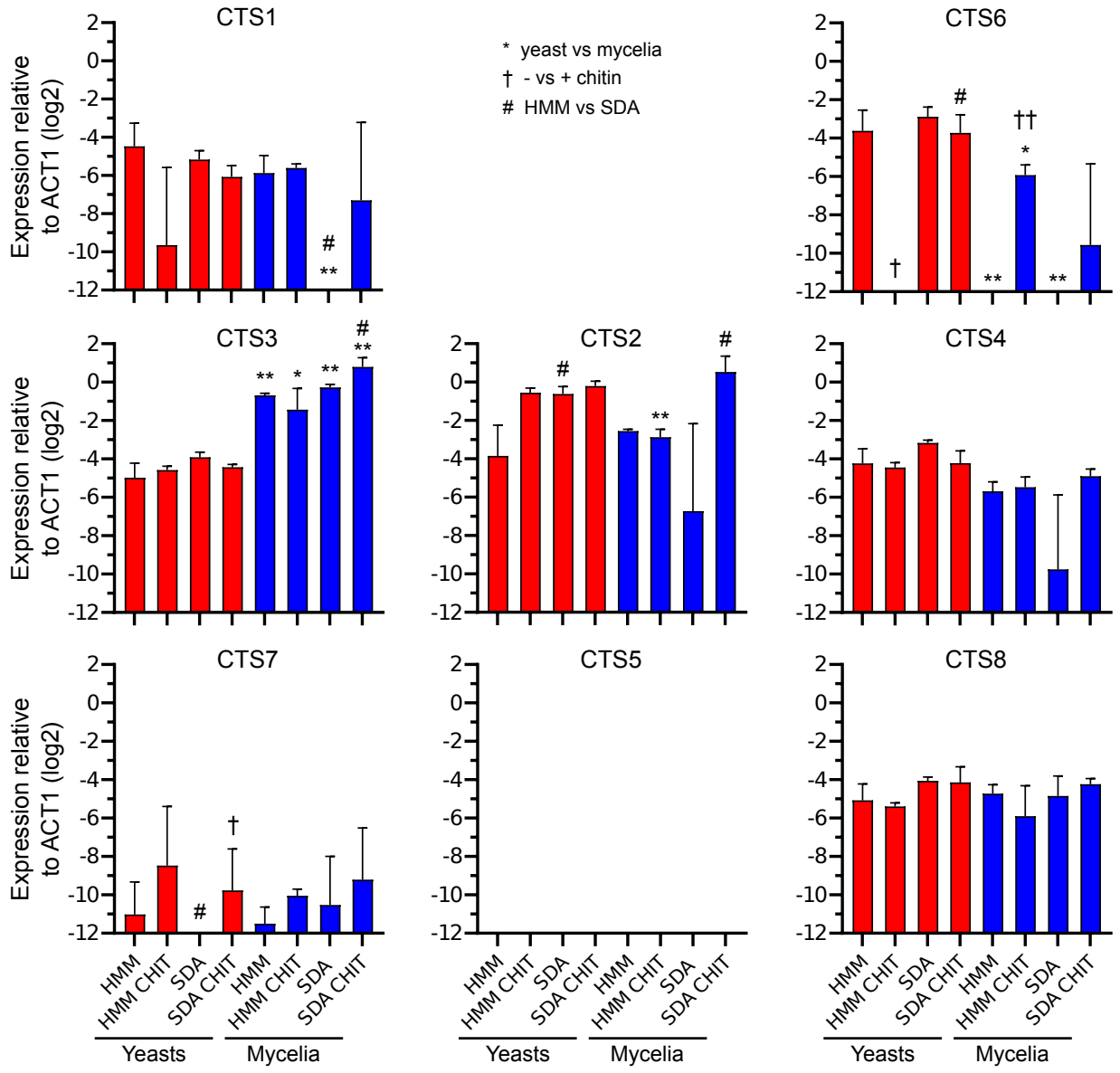

Figure 5. Expression of chitinase-encoding genes by *H. capsulatum* yeasts and mycelia
Expression of each chitinase-encoding gene (CTS) by *H. capsulatum* cells was measured using qRT-PCR. Transcription was determined with *H. capsulatum* grown on two media, one based on basal cell culture medium (HMM) and one based on glucose and peptone (Sauboraud's dextrose; SDA). For some tests, media was supplemented with chitin (CHIT). Yeast (red bars) and mycelial (blue bars) phases were maintained by growth at 37°C or 25°C, respectively. Data shows expression of each CTS gene relative to the expression of actin (ACT) and organized by sub-clades (B, cyan; A, magenta; C, green). Data represent the average ± standard deviation among biological replicates (n=3). Significant differences in expression between yeasts and mycelia (*), absence and presence of chitin (†), or HMM compared to SDA (#) are indicated for corresponding conditions differing by the single parameter. Single symbols indicate P<0.05 and double symbols indicate P<0.01 as determined by pairwise Student's t-tests.

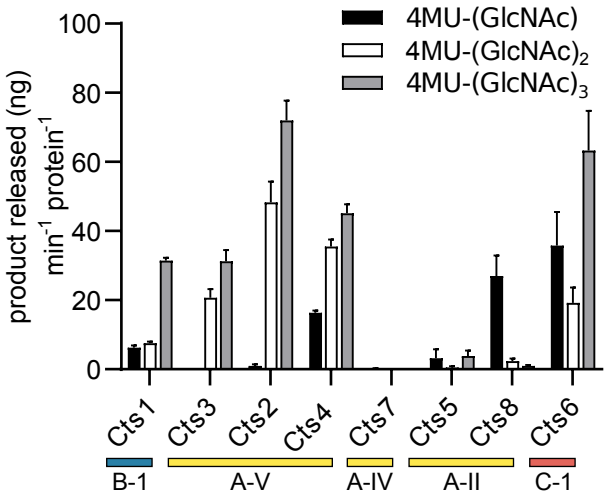

Figure 6. Enzymatic activity of *H. capsulatum* chitinases

The enzymatic activities of purified H. capsulatum chitinases were tested using fluorigenic GlcNAc oligomers linked to 4-methylumbelliferone (4MU) by a glycosidic bond. Chitinase activity that hydrolyzes this linkage liberates 4-MU which was detected by fluorescence. Substrates tested were 4MU-(GlcNAc) (black bars), 4MU-$(GlcNAc)_2$ (white bars), 4MU-$(GlcNAc)_3$ (gray bars). The exochitinase activity, N-acetyl-glucosaminidase, processively removes GlcNAc saccharides from the non-reducing end resulting in hydrolysis of all substrates. Chitobiosidase activity (release of chitobiose from the non-reducing end) is detected as hydrolysis of 4MU-$(GlcNAc)_2$. Endochitinase activity generates fluorescence from 4MU-$(GlcNAc)_2$ and 4MU-$(GlcNAc)_3$. Chitobiase activity (hydrolysis of the disaccharide chitobiose) generates fluorescence only from the chitobiose mimic 4MU-(GlcNAc). Activity rates were calculated as nanograms of hydrolysis product (4MU) released per minute per nanogram of *H. capsulatum* protein. Data represent the average activity ± standard deviation of replicates (n = 3). GH18 sub-clades to which each Cts protein belong are indicated (A, yellow; B , blue; and C, red).