Pseudozyma saprotrophic yeasts have retained a large effector arsenal, including functional Pep1 orthologs

Rahul Sharma[1,2], Bilal Ökmen[3], Gunther Doehlemann[3], Marco Thines[1,2,4]*

[1]Senckenberg BiK-F, Frankfurt am Main, Germany

[2]Goethe University, Biosciences, Institute of Ecology, Evolution and Diversity, Frankfurt am Main, Germany

[3]Botanical Institute and CEPLAS, University of Cologne, BioCenter, Cologne, Germany.

[4]Integrative Fungal Research Cluster (IPF), Frankfurt am Main, Germany.

* m.thines@thines-lab.eu

Summary

The basidiomycete smut fungi are predominantly plant parasitic, causing severe losses in some crops. Most species feature a saprotrophic haploid yeast stage, and several smut fungi are only known from this stage, with some isolated from habitats without suitable hosts, e.g. from Antarctica. Thus, these species are generally believed to be apathogenic, but recent findings that some of these might have a plant pathogenic sexual counterpart, casts doubts on the validity of this hypothesis. Here, four *Pseudozyma* genomes were re-annotated and compared to published smut pathogens and the well-characterised effector gene *Pep1* from these species was checked for its ability to complement a *Pep1* deletion strain of *Ustilago maydis*. It was found that 113 high-confidence putative effector proteins were conserved among smut and *Pseudozyma* genomes. Among these were several validated effector proteins, including Pep1. By genetic complementation we show that *Pep1* homologs from the supposedly apathogenic yeasts restore virulence in *Pep1*-deficient mutants *Ustilago maydis*. Thus, it is concluded that *Pseudozyma* species have retained a suite of effectors. This hints at the possibility that *Pseudozyma* species have kept an unknown plant pathogenic stage for sexual recombination or that these effectors have positive effects when colonising plant surfaces.

Key words: core effectors; effector complementation; plant pathogens; *Pseudozyma*; *Ustilago*; yeast

Introduction

Smut fungi in a broad sense are one of the three major lineages of Basidiomycota. Some species threaten crop plants, e.g. *Ustilago maydis* (maize) and *Ustilago scitaminea* (sugarcane). Smut fungi are dimorphic, with a filamentous and a yeast-like morph in different life stages. The yeast stage is haploid and saprotrophic, the hyphal stage is dicaryotic and obligate plant parasitic. Yeasts of matching sex form fusing conjugation hyphae, producing an infective dikaryon penetrating the host or colonisation. However, some smuts are known only from their yeast stage. Since the ending of dual naming for anamorphic and teleomorphic fungi (Hawksworth *et al.*, 2011), the yeast genus *Pseudozyma* became obsolete, being scattered throughout the smut phylogeny (Wang *et al.*, 2015). It was recently established that *Pseudozyma prolifica* is conspecific with *U. maydis* (Wang *et al.*, 2015), but it is believed that other *Pseudozyma* species have lost pathogenicity and exist only as yeasts (Lefebvre *et al.*, 2013; Morita *et al.*, 2013; Morita *et al.*, 2014). A well-known species of these is the biocontrol agent *Pseudozyma aphidis* (Gafni *et al.*, 2015), closely related or probably conspecific with *Moesziomyces bullatus* (Kruse *et al.*, 2017).

A key to investigating the apathogenic or pathogenic nature of *Pseudozyma*-like yeasts are putative secreted effector proteins (PSEPs). For successful colonization of hosts, pathogens secrete hundreds of effector proteins (Latijnhouwers *et al.*, 2003). Effectors often have limited sequence conservation (Schirawski *et al.*, 2010; Kemen *et al.*, 2011; Sharma *et al.*, 2014), but some are conserved in genomes of related pathogens (Sharma *et al.* 2014, 2015; Quinn *et al.* 2013; Hemetsberger *et al.* 2015; Lanver et al. (2017). Previous definitions of putative effector proteins were often associated with several restrictions, such as size cut-off (e.g. below 300aa), certain amino acid composition (e.g. Cys-rich), or the absence of sequence motifs predicting an enzymatic function. However, such formal restrictions artificially exclude a large fraction of proteins that contribute to virulence and interfere with the plant immune system. Effectors with enzymatic functions have actually been identified in

various filamentous plant pathogens (Franceschetti *et al.*, 2017). Recently, the *U. maydis* metalloprotease Fly1 was found to contribute to virulence by targeting maize chitinase. Here, we investigated conservation of PSEPs, including effectors with predicted functional domains, among *Pseudozyma* genomes (Morita *et al.*, 2013; Konishi *et al.*, 2013; Oliveira *et al.*, 2013; Lorenz *et al.*, 2014) to evaluate the potential pathogenicity of *Pseudozyma*-like yeasts. We were focusing on well-studied effectors, in particular the core-effector *Pep1*, which previously was identified as an essential virulence factor of *U. maydis* which is functionally conserved amongst pathogenic smuts (Hemetsberger *et al.*, 2012, 2015).

Results

Of 211 core PSEPs, a total of 178, 199, 182 and 171 candidates were conserved in the non-pathogenic yeasts *P. antarctica, P. hubeiensis, P. brasiliensis,* and *P. aphidis,* respectively (Tables 1, 2). Of these, 158, 182, 166, and 151, respectively, were predicted to have a secretion signal (Table 2). In total, 113 PSEPs were found to be conserved among all eight genomes. Functional annotation revealed features associated with pathogenicity, such as glycoside hydrolase and aspartic peptidase domains (Supplementary Table 2).

Almost all of the well-studied *Ustilago maydis* virulence factors conserved among Ustilaginales were found conserved in the four *Pseudozyma* genomes (Table 1). The conservation of the *Pit* cluster (Doehlemann *et al.*, 2011) *Cmu1* (Djamei *et al.*, 2011), *Cwh41* (Martínez-Soto *et al.*, 2013), and *Hum3* (Muller *et al.*, 2008) was found in all four *Pseudozyma* genomes (Table 1). *Tin2* (Tanaka *et al.*, 2014) was also found conserved, but the *P. hubeiensis* ortholog was lacking a strong secretion signal (Table 1). The *Pep1* effector was found to be conserved among the eight genomes, with all structural features intact (Figure 1). Besides these effectors with already known virulence function, also the membrane proteins Msb2 (Lanver *et al.*, 2010) and Pit1 (Doehlemann et al., 2011), which hold crucial virulence

functions, but are also known from non-pathogenic species were conserved amongst the yeast genomes.

To test the functionality as virulence factors on the example of the well-studied effector Pep1, coding regions of the *Pseudozyma Pep1* orthologs were fused to the *U. maydis pep1* promoter and stably integrated in the *ip*-locus of *U. maydis* strain SG200Δpep1, which is unable to infect maize plants due to the deletion of *Pep1* (Doehlemann *et al.*, 2009). The resulting *U. maydis* strains were verified by Southern Blot (Supplementary Figure 1) and subsequently inoculated to maize seedlings. Strikingly, scoring of *U. maydis* tumour formation revealed that all strains expressing Pep1 homologs from *Pseudozyma* species were fully virulent, i.e. produced plant tumours were indistinguishable from the progenitor strain SG200 (Figure 2). This result demonstrates that *Pseudozyma* species encode functional orthologs of the Pep1 effector, which restore virulence of *U. maydis* in maize.

Discussion

To avoid recognition by R-proteins, many secreted effectors adapt quickly and show limited conservation (Schirawski *et al*., 2010; Laurie *et al*., 2012). However, conservation of putative secreted effector protein (PSEP)-encoding genes among related pathogens has been reported (Hemetsberger *et al*., 2015; Sharma et al., 2015). These effectors have been termed "core" effectors (Sharma et al. 2014). The well-studied Pep1 effector, required for successful host colonisation (Doehlemann *et al.*, 2009), is an archetypal core effector, functionally conserved among monocot and dicot infecting smut pathogens (Hemetsberger *et al*., 2015). Our finding that also Cmu1, Tin2, and Hum3 remained conserved with secretion signal peptides in the genomes of smuts fungi only known from the yeast stage suggests that they either feature an unknown pathogenic stage, or that effectors also have a positive effect when settling on plant surfaces. Also in *P. antarctica* isolated from sediments in Antarctica, these

111 effectors remained conserved, with secretion signal peptides intact. If these effectors were not
112 needed anymore, it can be expected that they would either get lost quickly, as in *U.*
113 *pennsylvanica* after a host jump from monocots to dicots (Sharma *et al*., 2014), or acquire
114 new functions (Sharma *et al*., 2015). For *Pep1*, the finding that orthologs from all yeast-only
115 species were able to fully restore pathogenicity in *U. maydis*, demonstrates that its virulence
116 function remained conserved amongst species. Moreover, our finding highlights that more
117 research is needed to investigate, if fully saprotrophic species of the Ustilaginales exist (Kruse
118 *et al*., 2017) and if so, how different lifestyles of smut yeasts evolved . Regarding the
119 functional conservation of Pep1 one could speculate that its function to suppress PAMP-
120 triggered ROS generation also benefits epiphytic yeasts on plant surfaces. However, in *U.*
121 *maydis* transcription of Pep1, as well as that of other known effectors, is only activated upon
122 mating when compatible heterodimers of the b-transcription factor are present in the cell.
123 Therefore, a putative role of effectors in non-biotrophic stages would imply a fundamental
124 change in the transcriptional regulatory cascade of anamorphic smut yeasts. Even though the
125 alternative explanation that all the conserved proteins would also benefit a saprotrophic yeasts
126 seems highly unlikely, as several are known to interact only with targets inside the plant
127 cytoplasm. Given the conservation of more than 100 PSEPs and almost all core effectors with
128 known virulence activity, it seems more likely that *Pseudozyma* species have a plant
129 pathogenic stage, rather than that they have lost it only recently, simultaneously in four
130 species. It is conceivable that those few species frequently encountered as yeasts, are
131 competitive saprotrophs and that the plant pathogenic stage is only maintained to allow for
132 infrequent sexual recombination.

133

134 # Materials and Methods

135 ## Bioinformatics

Out of 248 PSEPs conserved among the four smut genomes i.e. *Ustilago maydis* (Kämper *et al.*, 2006), *U. hordei* (Laurie *et al.*, 2012), *U. reiliana* (Schirawski *et al.*, 2010), and *U. pennsylvanica* (Sharma *et al.*, 2014), a high-confidence (secretion strongly supported) core set of 211 PSEPs was inferred (Sharma *et al.*, 2015). The genomes of four *Pseudozyma* species, *P. antarctica, P. hubeiensis, P. brasiliensis*, and *P. aphidis*, were scanned for the presence of the 211 PSEPs, with *U. maydis* proteins as query. To investigate the conservation, also *ab initio* prediction was done using GeneMark (Ter-Hovhannisyan *et al.*, 2008), trained on the other Ustilaginales. The resulting protein sequences were aligned to the 211 PSEPs of *U. maydis* using Blastp. A PSEP was considered present if exceeding 45% identity, an e-value of e-5, and alignment coverage of 60%. Candidate orthologs were scanned for secretion signals as described before (Sharma *et al.*, 2015). Start-codon positions of candidate orthologs were manually checked and corrected, if necessary, using the well-annotated *U. maydis* proteins as reference.

The PSEPs conserved among smuts and yeasts were annotated based on *U. maydis* proteins (ftp://ftpmips.gsf.de/fungi/Ustilaginaceae/Ustilago_maydis_521/) and InterProScan (Quevillon *et al.*, 2005) using Blast2GO (Conesa *et al.*, 2005). Particular attention was paid to validated effector proteins of *U. maydis*.

Δ*umpep1* complementation and disease assay

In order to show functional conservation between the *Pep1* orthologs, those from *P. antarctica, P. hubeiensis, P. brasiliensis*, and *P. aphidis* were amplified by PCR with the primers given in Supplementary Table 1 and subsequently expressed in the *U. maydis* SG200Δpep1 strain, which is deleted for *pep1* (Doehlemann *et al.*, 2009). For proper expression, *pep1* orthologs were expressed controlled by the endogenous *U. maydis pep1*-promoter, integrated in single copy (checked by Southern blots) into the *ip*-locus of *U. maydis*, as described previously (Hemetsberger *et al.*, 2015).

The *U. maydis* disease assays were performed according to Hemetsberger *et al.* (2015). Briefly, the cell culture control and complementation strains were inoculated onto maize seedlings (variety Early Golden Bantam) with a syringe and needle into the leaf whirl. All assays were performed in biological triplicates (≥30 plants each). The quantification of disease symptoms was performed at 12 dpi as described previously (Kämper *et al.*, 2006).

Acknowledgements

RS carried out bioinformatics work and participated in data analysis; BO carried out the molecular lab work, participated in data analysis, and edited the manuscript; GD designed the study, participated in data analysis and drafted the manuscript; MT conceived of the study, designed the study, participated in data analysis, coordinated the study and drafted the manuscript. This study was funded by the LOEWE initiative of the government of Hessen, in the framework of the cluster for Integrative Fungal Research (IPF) and the Center of Excellence on Plant Science (CEPLAS). We thank Melanie Kastl and Raphael Wemhöner for assistance in generation of *U. maydis* strains and plant infection assays.

References

Conesa A, Götz S, García-Gómez JM, Terol J, Talón M, Robles M. 2005. Blast2GO: A universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*. 21: 3674-3676.

Djamei A, Schipper K, Rabe F, Ghosh A, Vincon V, Kahnt J, et al. 2011. Metabolic priming by a secreted fungal effector. *Nature* 478: 395-398.

Doehlemann G, Van Der Linde K, Aßmann D, Schwammbach D, Hof A, Mohanty A, et al. 2009. *Pep1*, a secreted effector protein of *Ustilago maydis*, is required for successful invasion of plant cells. *PLoS Pathogens* 5: e1000290.

Doehlemann G, Reissmann S, Aßmann D, Fleckenstein M, Kahmann R. 2011. Two linked genes encoding a secreted effector and a membrane protein are essential for *Ustilago maydis*-induced tumour formation. *Molecular Microbiology* 81: 751-66.

Franceschetti M, Maqbool A, Jiménez-Dalmaroni M J, Pennington HG, Kamoun S, Banfield, MJ. 2017. Effectors of filamentous plant pathogens: commonalities amid diversity. *Microbiology and Molecular Biology Reviews* 81: e00066-16.

Gafni A, Claudia C, Raviv H, Buxdorf K, Avis D, Berger ML. 2015. Biological control of the cucurbit powdery mildew pathogen *Podosphaera xanthii* by means of the epiphytic fungus *Pseudozyma aphidis. Fronttiers in Plant Sciences* 6: 1-11.

Hawksworth DL, Crous PW, Redhead S, Reynolds DR, Samson R, Seifert K, et al. 2011. The Amsterdam declaration on fungal nomenclature. *IMA Fungus* 2: 105-112.

Hemetsberger C, Mueller AN, Matei A, Herrberger C, Hensel G, Kumlehn J, et al. 2015. The fungal core effector *Pep1* is conserved across smuts of dicots and monocots. *New Phytologist* 206: 1116-1126.

Kämper J, Kahmann R, Bölker M, Ma L-J, Brefort T, Saville BJ, et al. 2006. Insights from the genome of the biotrophic fungal plant pathogen *Ustilago maydis. Nature* 444: 97-101.

205 Kemen E, Gardiner A, Schultz-Larsen T, Kemen AC, Balmuth AL, Robert-Seilaniantz A, et
206 al. 2011. Gene gain and loss during evolution of obligate parasitism in the white rust
207 pathogen of *Arabidopsis thaliana*. *PLoS Biology* 9: e1001094.

208 Konishi M, Hatada Y, Horiuchi J. 2013. Draft Genome Sequence of the Basidiomycetous
209 Yeast-Like Fungus *Pseudozyma hubeiensis SY62*, Which Produces an Abundant
210 Amount of the Biosurfactant Mannosylerythritol Lipids. *Genome Announcements* 1:
211 13-14.

212 Kruse J, Doehlemann G, Kemen E, Thines M. 2017. Asexual and sexual morphs of
213 *Moesziomyces* revisited. *IMA Fungus* 8: 117-129.

214 Latijnhouwers M, De Wit PJGM, Govers F. 2003. Oomycetes and fungi: Similar weaponry to
215 attack plants. *Trends in Microbiology* 11: 462-469.

216 Lanver D, Mendoza-Mendoza A, Brachmann A, Kahmann R. 2010. Sho1 and Msb2-Related
217 Proteins Regulate Appressorium Development in the Smut Fungus *Ustilago maydis*.
218 *Plant Cell* 22: 2085-2101.

219 Lanver D, Tollot M, Schweizer G, Presti LL, Reissmann S, Ma LS, *et al*. 2017. *Ustilago*
220 *maydis* effectors and their impact on virulence. *Nature Reviews Microbiology* 15: 409.

221 Laurie JD, Ali S, Linning R, Mannhaupt G, Wong P, Guldener U, et al. 2012. Genome
222 Comparison of Barley and Maize Smut Fungi Reveals Targeted Loss of RNA
223 Silencing Components and Species-Specific Presence of Transposable Elements. Pla*nt*
224 *Cell* 24: 1733-1745.

225 Lefebvre F, Joly DL, Labbé C, Teichmann B, Linning R, Belzile F, et al. 2013. The transition
226 from a phytopathogenic smut ancestor to an anamorphic biocontrol agent deciphered
227 by comparative whole-genome analysis. *Plant Cell* 25: 1946-1959.

Lorenz S, Guenther M, Grumaz C, Rupp S, Zibek S, Sohn K. 2014. Genome Sequence of the Basidiomycetous Fungus *Pseudozyma aphidis DSM70725*, an Efficient Producer of Biosurfactant Mannosylerythritol Lipids. *Genome Announcements* 2: e00053

Martínez-Soto D, Robledo-Briones AM, Estrada-Luna A a, Ruiz-Herrera J. 2013. Transcriptomic analysis of *Ustilago maydis* infecting *Arabidopsis* reveals important aspects of the fungus pathogenic mechanisms. *Plant Signalling and Behaviour* 8: 1-13.

Morita T, Koike H, Koyama Y, Hagiwara H, Ito E, Fukuoka T, et al. 2013. Genome Sequence of the Basidiomycetous Yeast *Pseudozyma antarctica* T-34, a Producer of the Glycolipid Biosurfactants Mannosylerythritol Lipids. *Genome Announcements* 1: e00064-13.

Morita T, Koike H, Hagiwara H, Ito E, Machida M, Sato S, et al. 2014. Genome and transcriptome analysis of the basidiomycetous yeast *Pseudozyma antarctica* producing extracellular glycolipids, mannosylerythritol lipids. *PLoS One* 9: e86490.

Muller O, Schreier PH, Uhrig JF. 2008. Identification and characterization of secreted and pathogenesis-related proteins in *Ustilago maydis*. *Molecular Genetics and Genomics* 279: 27–39.

Oliveira JV de C, Dos Santos RAC, Borges TA, Riaño-Pachón DM, Goldman GH. 2013. Draft Genome Sequence of *Pseudozyma brasiliensis sp.* nov. Strain GHG001, a High Producer of Endo-1,4-Xylanase Isolated from an Insect Pest of Sugarcane. *Genome Announcements* 1: e00920-13.

Quevillon E, Silventoinen V, Pillai S, Harte N, Mulder N, Apweiler R, et al. 2005. InterProScan: Protein domains identifier. *Nucleic Acids Research* 33: XY.

Quinn L, O'Neill PA, Harrison J, Paskiewicz KH, Mccracken AR, Cooke LR, et al. 2013. Genome-wide sequencing of *Phytophthora lateralis* reveals genetic variation among isolates from Lawson cypress (*Chamaecyparis lawsoniana*) in Northern Ireland. *FEMS Microbiology Letters* 344: 179-185.

Redkar A, Villajuana-Bonequi M, Doehlemann G. 2015. Conservation of the *Ustilago maydis* effector *See1* in related smuts. *Plant Signalling and Behaviour* 10: e1086855.

Schirawski J, Mannhaupt G, Munch K, Brefort T, Schipper K, Doehlemann G, et al. 2010. Pathogenicity determinants in smut fungi revealed by genome comparison. *Science* 330: 1546-1548.

Sharma R, Mishra B, Runge F, Thines M. 2014. Gene loss rather than gene gain is associated with a host jump from monocots to dicots in the smut fungus *Melanopsichium pennsylvanicum*. *Genome Biology and Evolution* 6: 2034-2049.

Sharma R, Xia X, Riess K, Bauer R, Thines M. 2015. Comparative genomics including the early-diverging smut fungus *Ceraceosorus bombacis* reveals signatures of parallel evolution within plant and animal pathogens of fungi and oomycetes. *Genome Biology and Evolution* 7: 2781-2798.

Tanaka S, Brefort T, Neidig N, Djamei A, Kahnt J, Vermerris W, et al. 2014. A secreted *Ustilago maydis* effector promotes virulence by targeting anthocyanin biosynthesis in maize. *Elife* 3: e01355.

Ter-Hovhannisyan V, Lomsadze A, Chernoff YO, Borodovsky M. 2008. Gene prediction in novel fungal genomes using an ab initio algorithm with unsupervised training. *Genome Research* 18: 1979-1990.

Wang QM, Begerow D, Groenewald M, Liu XZ, Theelen B, Bai FY, et al. 2015. Multigene phylogeny and taxonomic revision of yeasts and related fungi in the Ustilaginomycotina. *Studies in Mycology* 81: 55-83.

Tables

Table 1. Conservation of functionally characterized pathogenicity related proteins of *U. maydis* among four *Pseudozyma* genomes.

U. maydis gene	Gene name	*P. antarctica*	*P. aphidis*	*P. brasiliensis*	*P. hubeiensis*
UMAG_04433	*Hum3*	Pant_3371_t	Paph_3632_t.1*	Pbra_985_t	Phub_3033_t
UMAG_01987	*Pep1*	Pant_5912_t	Paph_3137_t	Pbra_3010_t	Phub_108_t
UMAG_01375	*Pit2*	Pant_3484_t***	Paph_3237_t***	Pbra_3853_t***	Phub_3768_t
UMAG_05302	*Tin2*	Pant_4343_t.2***	Paph_1304_t.2***	Pbra_3373_t	Phub_4666_t.1**
UMAG_05731	*Cmu1*	Pant_6209_t	Paph_4714_t	Pbra_2221_t	Phub_1879_t

* Sequence contains a stretch of Ns, could be resolved by experimental methods. ** No signal peptide predicted by using SignalP4.1, but by TargetP1 and Protcomp9. Genes with secretion signal and some conservation of some domains, but not matching the strict orthology criteria over the entire length.

Table 2. Number of orthologs of putative effectors found among the four *Pseudozyma* genomes.

Species	All Orthologs	Orthologs with SP in all species
P. antarctica	178	158
P. hubeiensis	199	182
P. brasiliensis	182	166
P. aphidis	171	151

Figure captions

Figure 1. Multiple sequence alignment of eight candidate *Pep1* proteins from pathogens and *Pseudozyma* yeasts. Multiple sequence alignment (MSA) shows high sequence conservation of the candidate *Pep1* effector proteins. The conserved four cysteine residues needed for the funtion of *Pep1* are highlighted in yellow.

Figure 2. Disease assay on EGB maize lines. For testing biological conservation, Pep1 orthologs from *Pseudozyma* yeasts were expressed in a *Pep1* deletion background (SG200Δumpep1). The restoration of pathogenicity in the complemented lines is indicative of functional conservation.

Supporting information figure caption

Figure S1. Southern blot analysis to confirm single integration events. All complementation events were performed in the *ip* locus in SG200Δumpep1 background. Restriction enzyme, DNA probe that were used and the expected fragments sizes for each southern blot analysis are indicated below each picture. Red arrows indicate single integration event in the correct genomic locus.

Tables

Table S1. Primers for delta-umpep1 complementation with *Pep1* homologs

Primer name	Primer sequence	Description
Pant_Pep1_fr	CTCTCGTCTAGAATGAGGTTCATGCTTGCCACTG	XbaI site
Pant_Pep1_rw	GCAGCGCCCGGGTTAAAAGCCAAGCAGATTAC	XmaI site
Paph_Pep1_fr	CTCTCGTCTAGAATGAGGTTCATGCTTGCCACTG	XbaI site
Paph_Pep1_rw	GCAGCGCCCGGGTTAAAAGCCAAGCAGATTAC	XmaI site
Pbra_Pep1_fr	CTCTCGTCTAGAATGAAGACGACACTCCTCAC	XbaI site
Pbra_Pep1_rw	GCAGCGCCCGGGCTACATCCCGAACATGCTTC	XmaI site
Phub_Pep1_fr	CTCTCGTCTAGAATGAAGACGACTCCTTCTC	XbaI site
Phub_Pep1_rw	GCAGCGCCCGGGTCACATGCCGAACATGCTG	XmaI site

318 **Table S2.** Annotation of putative secreted effector proteins conserved in all eight genomes.

Ortholog	*U. maydis* ids			
1	UMAG_00013	hypothetical protein	noIPR	
2	UMAG_00054	hypothetical protein	noIPR	
3	UMAG_00064	related to pepsin precursor (aspartate protease)	IPR001461; IPR021109	Aspartic peptidase; Aspartic peptidase domain
4	UMAG_00102	hypothetical protein	IPR017853	Glycoside hydrolase, superfamily
5	UMAG_00144	hypothetical protein	IPR000782	FAS1 domain
6	UMAG_00159	hypothetical protein	noIPR	
7	UMAG_00235	probable EXG1 - exo-beta-1,3-glucanase (I/II), major isoform	IPR017853	Glycoside hydrolase, superfamily
8	UMAG_00309	hypothetical protein	noIPR	
9	UMAG_00330	related to UTR2 - cell wall protein	IPR008985	Concanavalin A-like lectin/glucanases superfamily
10	UMAG_00598	hypothetical protein	noIPR	
11	UMAG_00655	probable Family 9 glycosyl hydrolase	IPR001701; IPR008928	Glycoside hydrolase, family 9; Six-hairpin glycosidase-like
12	UMAG_00688	related to bacterial leucyl aminopeptidase precursor	IPR007484	Peptidase M28
13	UMAG_00692	hypothetical protein	IPR018466	Ser-Thr-rich glycosyl-phosphatidyl-inositol-

				anchored membrane family
14	UMAG_00695	putative exochitinase	IPR017853; IPR025705	Glycoside hydrolase, superfamily; Beta-hexosaminidase subunit alpha/beta
15	UMAG_00715	hypothetical protein	noIPR	
16	UMAG_00723	related to stress response protein rds1p	IPR009078	Ferritin-like superfamily
17	UMAG_00882	related to PPN1 - vacuolar endopolyphosphatase	IPR004843	Phosphoesterase domain
18	UMAG_00961	hypothetical protein	noIPR	
19	UMAG_01022	hypothetical protein	IPR009009	RlpA-like double-psi beta-barrel domain
20	UMAG_01082	hypothetical protein	noIPR	
21	UMAG_01512	related to UDP-glucose:glycoprotein glucosyltransferase precursor	IPR002495; IPR009448	Glycosyl transferase, family 8; UDP-glucose:Glycoprotein Glucosyltransferase
22	UMAG_01705	hypothetical protein	IPR027706	Mitochondrial PGP phosphatase
23	UMAG_01772	related to Aldose 1-epimerase precursor	IPR008183; IPR011013; IPR015443	Aldose 1-/Glucose-6-phosphate 1-epimerase; Galactose mutarotase-like domain; Aldose 1-epimerase
24	UMAG_01774	hypothetical protein	IPR005887; IPR008928	Alpha-1,2-mannosidase, putative; Six-hairpin glycosidase-like

25	UMAG_01788	related to deacetylase	IPR011330	Glycoside hydrolase/deacetylase, beta/alpha-barrel
26	UMAG_01802	probable protein disulfide-isomerase precursor	IPR005746; IPR011679; IPR012336	Thioredoxin; Endoplasmic reticulum, protein ERp29, C-terminal; Thioredoxin-like fold
27	UMAG_01820	hypothetical protein	noIPR	
28	UMAG_01851	hypothetical protein	IPR018466	Ser-Thr-rich glycosyl-phosphatidyl-inositol-anchored membrane family
29	UMAG_01855	hypothetical protein	IPR027372	Phytase-like domain
30	UMAG_01886	related to carboxypeptidase	IPR001563	Peptidase S10, serine carboxypeptidase
31	UMAG_01888	probable serine-type carboxypeptidase f precursor	IPR001563	Peptidase S10, serine carboxypeptidase
32	UMAG_01937	related to acid sphingomyelinase	IPR004843	Phosphoesterase domain
33	UMAG_01940	hypothetical protein	IPR003892	Ubiquitin system component Cue
34	UMAG_01945	probable SUC2 - invertase (sucrose hydrolyzing enzyme)	IPR001362; IPR008985; IPR023296	Glycoside hydrolase, family 32; Concanavalin A-like lectin/glucanases superfamily; Glycosyl hydrolase, five-bladed beta-propellor domain
35	UMAG_01957	related to Mannosyl-oligosaccharide alpha-1,2-mannosidase precursor	IPR001382	Glycoside hydrolase, family 47

36	UMAG_01977	hypothetical protein	noIPR	
37	UMAG_01987	fungal core effector Pep1	noIPR	
38	UMAG_02006	hypothetical protein	IPR009003	Trypsin-like cysteine/serine peptidase domain
39	UMAG_02019	probable Chitin deacetylase	IPR011330	Glycoside hydrolase/deacetylase, beta/alpha-barrel
40	UMAG_02035	related to yellowish-green 1 (ayg1)	IPR010520	Protein of unknown function DUF1100, hydrolase-like
41	UMAG_02111	related to triacylglycerol lipase precursor	IPR002921	Lipase, class 3
42	UMAG_02178	related to aspartic protease	IPR001461; IPR021109	Aspartic peptidase; Aspartic peptidase domain
43	UMAG_02204	related to beta-galactosidase	IPR017853	Glycoside hydrolase, superfamily
44	UMAG_02381	hypothetical protein	IPR011330	Glycoside hydrolase/deacetylase, beta/alpha-barrel
45	UMAG_02465	related to Calcium influx promoting protein ehs1	IPR024338	Stretch-activated cation channel Mid1
46	UMAG_02523	related to Endoglucanase 1 precursor	IPR009009	RlpA-like double-psi beta-barrel domain
47	UMAG_02611	related to ribonuclease M	IPR001568	Ribonuclease T2-like
48	UMAG_02620	hypothetical protein	noIPR	
49	UMAG_02865	hypothetical protein	IPR000782	FAS1 domain

50	UMAG_03024	related to subtilisin-like serine protease	IPR000209; IPR015500	Peptidase S8/S53 domain; Peptidase S8, subtilisin-related
51	UMAG_03551	related to Glucose oxidase	IPR012132	Glucose-methanol-choline oxidoreductase
52	UMAG_03614	hypothetical protein	noIPR	
53	UMAG_03615	related to Glucose oxidase	IPR012132	Glucose-methanol-choline oxidoreductase
54	UMAG_03807	hypothetical protein	IPR001938	Thaumatin
55	UMAG_03880	hypothetical protein	IPR017946	PLC-like phosphodiesterase, TIM beta/alpha-barrel domain
56	UMAG_03947	probable carboxypeptidase 2	IPR001563	Peptidase S10, serine carboxypeptidase
57	UMAG_04044	related to Glucose oxidase	IPR012132	Glucose-methanol-choline oxidoreductase
58	UMAG_04145	hypothetical protein	noIPR	
59	UMAG_04171	related to ROT1 - molecular chaperone in the endoplasmic reticulum	IPR019623	Chaperone, endoplasmic reticulum protein-folding, fungi
60	UMAG_04282	related to 3-phytase A precursor	IPR000560	Histidine phosphatase superfamily, clade-2
61	UMAG_04318	related to Prenylcysteine oxidase precursor	IPR010795	Prenylcysteine lyase
62	UMAG_04355	related to Sel-1 homolog precursor	noIPR	
63	UMAG_04364;	probable EXG1 - Exo-1,3-beta-glucanase	IPR017853	Glycoside hydrolase,

	UMAG_05550	precursor		superfamily
64	UMAG_04400	probable PRB1 - protease B, vacuolar	IPR000209; IPR009020; IPR015500	Peptidase S8/S53 domain; Proteinase inhibitor, propeptide; Peptidase S8, subtilisin-related
65	UMAG_04405	alpha-glucosidase II precursor	IPR000322; IPR011013; IPR017853	Glycoside hydrolase, family 31; Galactose mutarotase-like domain; Glycoside hydrolase, superfamily
66	UMAG_04531	hypothetical protein	IPR009011	Mannose-6-phosphate receptor binding domain
67	UMAG_04630	related to MUS81 - endonuclease involved in DNA repair and replication fork stability	IPR010996; IPR011335	DNA polymerase beta-like, N-terminal domain; Restriction endonuclease type II-like
68	UMAG_04641.2	related to PRC1 - carboxypeptidase y, serine-type protease	IPR001563	Peptidase S10, serine carboxypeptidase
69	UMAG_04733	related to phosphatidylglycerol/phosphatidylinositol transfer protein	IPR014756	Immunoglobulin E-set
70	UMAG_04926	probable PEP4 - aspartyl protease	IPR001461; IPR021109	Aspartic peptidase; Aspartic peptidase domain
71	UMAG_05036	related to endo-1,3(4)-beta-glucanase	IPR008985	Concanavalin A-like lectin/glucanases superfamily
72	UMAG_05109	related to ECM14 - involved in cell wall biogenesis and architecture	IPR000834	Peptidase M14, carboxypeptidase A

73	UMAG_05222	hypothetical protein	noIPR	
74	UMAG_05227	hypothetical protein	noIPR	
75	UMAG_05352	related to MPD1 - Disulfide isomerase related protein	IPR005746; IPR012336	Thioredoxin; Thioredoxin-like fold
76	UMAG_05366	hypothetical protein	noIPR	
77	UMAG_05528	hypothetical protein	IPR017853	Glycoside hydrolase, superfamily
78	UMAG_05604	hypothetical protein	noIPR	
79	UMAG_05704	hypothetical protein	IPR017853	Glycoside hydrolase, superfamily
80	UMAG_05731	secreted chorismate mutase	IPR008238; IPR020822	Chorismate mutase, AroQ class, eukaryotic type; Chorismate mutase, type II
81	UMAG_05774; UMAG_03630	related to metalloprotease MEP1	IPR008754	Peptidase M43, pregnancy-associated plasma-A
82	UMAG_05988	hypothetical protein	IPR015889	Intradiol ring-cleavage dioxygenase, core
83	UMAG_06071	related to Para-nitrobenzyl esterase	IPR000997	Cholinesterase
84	UMAG_06075	related to beta-glucosidase	IPR002772; IPR017853; IPR026892	Glycoside hydrolase family 3 C-terminal domain; Glycoside hydrolase, superfamily; Glycoside hydrolase family 3
85	UMAG_06118	related to Tripeptidyl-peptidase I precursor	IPR000209; IPR009020	Peptidase S8/S53 domain; Proteinase inhibitor, propeptide

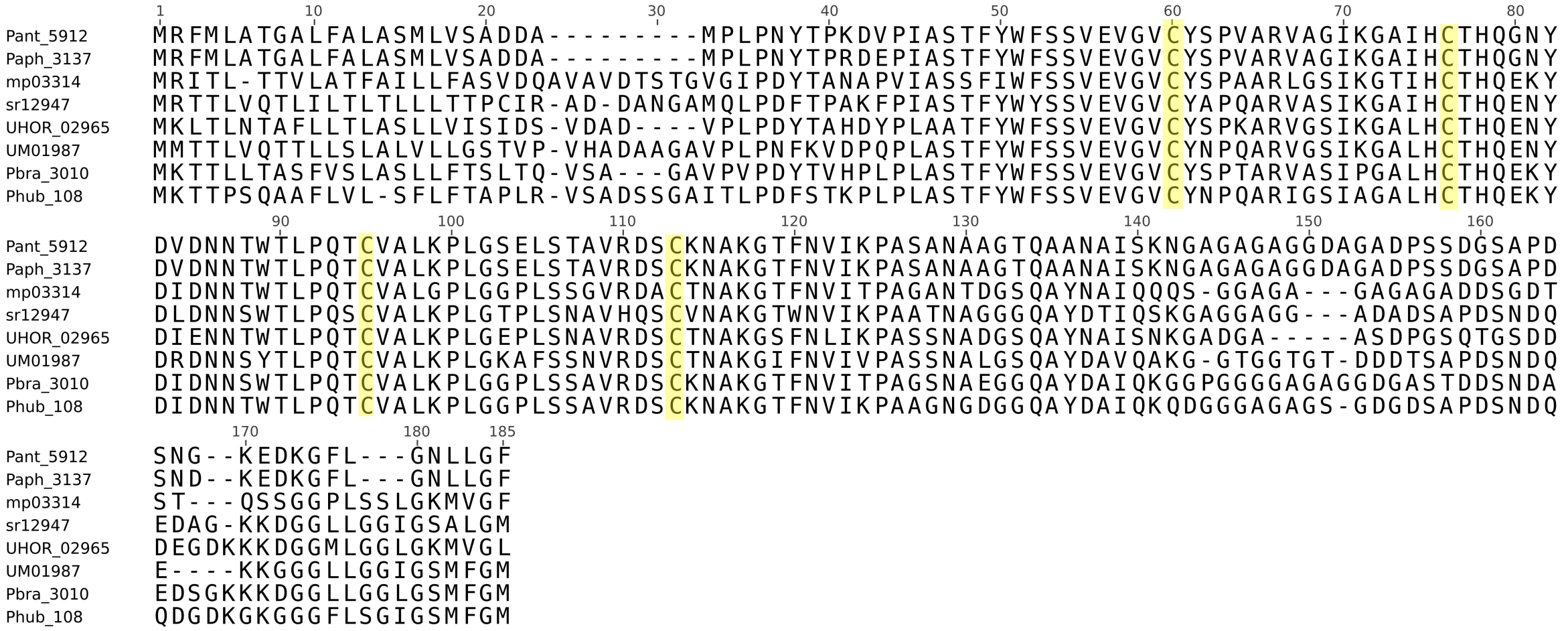
86	UMAG_06120	hypothetical protein	noIPR	
87	UMAG_06162	hypothetical protein	noIPR	
88	UMAG_06190	related to Chitinase	IPR017853	Glycoside hydrolase, superfamily
89	UMAG_06218	hypothetical protein	IPR023346	Lysozyme-like domain
90	UMAG_06332	endoglucanase 1 precursor (egl1)	IPR009009	RlpA-like double-psi beta-barrel domain
91	UMAG_10024	hypothetical protein	noIPR	
92	UMAG_10067	hypothetical protein	noIPR	
93	UMAG_10091	hypothetical protein	IPR008758	Peptidase S28
94	UMAG_10186	hypothetical protein	noIPR	
95	UMAG_10474.2	hypothetical protein	noIPR	
96	UMAG_10536	related to FPR2 - FK506/rapamycin-binding protein of the ER	IPR023566	Peptidyl-prolyl cis-trans isomerase, FKBP-type
97	UMAG_10640	hypothetical protein	IPR009009	RlpA-like double-psi beta-barrel domain
98	UMAG_10676	hypothetical protein	IPR000782	FAS1 domain
99	UMAG_10975.2	hypothetical protein	IPR017853	Glycoside hydrolase, superfamily
100	UMAG_11062	hypothetical protein	noIPR	
101	UMAG_11083	related to p24 protein, involved in membrane trafficking	IPR009038	GOLD DOMAIN
102	UMAG_11187	related to ROT1 - molecular chaperone in the	IPR019623	Chaperone, endoplasmic reticulum protein-folding,

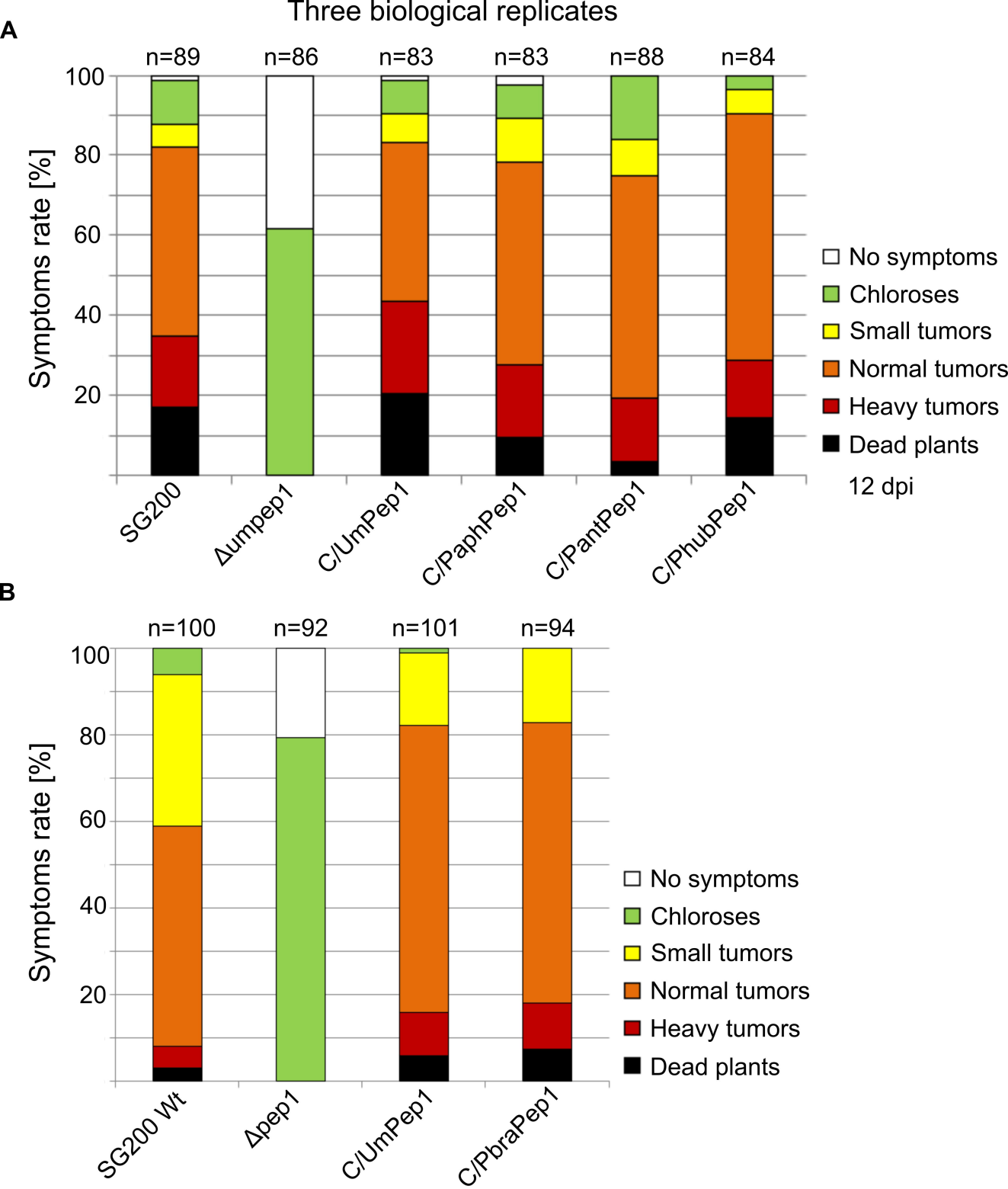
		endoplasmic reticulum		fungi
103	UMAG_11266	probable lysophospholipase (lpl)	IPR016035	Acyl transferase/acyl hydrolase/lysophospholipase
104	UMAG_11403	hypothetical protein	IPR018803	Stress-responsive protein Ish1
105	UMAG_11554	related to Glucose oxidase	IPR012132	Glucose-methanol-choline oxidoreductase
106	UMAG_11562	hydrophobin 2	IPR001338	Hydrophobin
107	UMAG_11886	hypothetical protein	IPR009009	RlpA-like double-psi beta-barrel domain
108	UMAG_11908	related to cathepsin d (lysosomal aspartyl protease)	IPR001461; IPR021109	Aspartic peptidase; Aspartic peptidase domain
109	UMAG_11922	related to Chitin deacetylase precursor	IPR011330	Glycoside hydrolase/deacetylase, beta/alpha-barrel
110	UMAG_11927	related to ROT1 - molecular chaperone in the endoplasmic reticulum	IPR019623	Chaperone, endoplasmic reticulum protein-folding, fungi
111	UMAG_12007	related to cellulase	IPR017853	Glycoside hydrolase, superfamily
112	UMAG_12205	hypothetical protein	noIPR	
113	UMAG_15089	hypothetical protein	IPR017853	Glycoside hydrolase, superfamily

319

320

321


1 10 20 30 40 50 60 70 80
Pant_5912 MRFMLATGALFALASMLVSADDA---------MPLPNYTPKDVPIASTFYWFSSVEVGVCYSPVARVAGIKGAIHCTHQGNY
Paph_3137 MRFMLATGALFALASMLVSADDA---------MPLPNYTPRDEPIASTFYWFSSVEVGVCYSPVARVAGIKGAIHCTHQGNY
mp03314 MRITL-TTVLATFAILLFASVDQAVAVDTSTGVGIPDYTANAPVIASSFIWFSSVEVGVCYSPAARLGSIKGTIHCTHQEKY
sr12947 MRTTLVQTLILTLTLLLTTPCIR-AD-DANGAMQLPDFTPAKFPIASTFYWYSSVEVGVCYAPQARVASIKGAIHCTHQENY
UHOR_02965 MKLTLNTAFLLTLASLLVISIDS-VDAD----VPLPDYTAHDYPLAATFYWFSSVEVGVCYSPKARVGSIKGALHCTHQENY
UM01987 MMTTLVQTTLLSLALVLLGSTVP-VHADAAGAVPLPNFKVDPQPLASTFYWFSSVEVGVCYNPQARVGSIKGALHCTHQENY
Pbra_3010 MKTTLLTASFVSLASLLFTSLTQ-VSA---GAVPVPDYTVHPLPLASTFYWFSSVEVGVCYSPTARVASIPGALHCTHQEKY
Phub_108 MKTTPSQAAFLVL-SFLFTAPLR-VSADSSGAITLPDFSTKPLPLASTFYWFSSVEVGVCYNPQARIGSIAGALHCTHQEKY
90 100 110 120 130 140 150 160
Pant_5912 DVDNNTWTLPQTCVALKPLGSELSTAVRDSCKNAKGTFNVIKPASANAAGTQAANAISKNGAGAGAGGDAGADPSSDGSAPD
Paph_3137 DVDNNTWTLPQTCVALKPLGSELSTAVRDSCKNAKGTFNVIKPASANAAGTQAANAISKNGAGAGAGGDAGADPSSDGSAPD
mp03314 DIDNNTWTLPQTCVALGPLGGPLSSGVRDACTNAKGTFNVITPAGANTDGSQAYNAIQQQS-GGAGA---GAGAGADDSGDT
sr12947 DLDNNSWTLPQSCVALKPLGTPLSNAVHQSCVNAKGTWNVIKPAATNAGGGQAYDTIQSKGAGGAGG---ADADSAPDSNDQ
UHOR_02965 DIENNTWTLPQTCVALKPLGEPLSNAVRDSCTNAKGSFNLIKPASSNADGSQAYNAISNKGADGA-----ASDPGSQTGSDD
UM01987 DRDNNSYTLPQTCVALKPLGKAFSSNVRDSCTNAKGIFNVIVPASSNALGSQAYDAVQAKG-GTGGTGT-DDDTSAPDSNDQ
Pbra_3010 DIDNNSWTLPQTCVALKPLGGPLSSAVRDSCKNAKGTFNVITPAGSNAEGGQAYDAIQKGGPGGGGAGAGGDGASTDDSNDA
Phub_108 DIDNNTWTLPQTCVALKPLGGPLSSAVRDSCKNAKGTFNVIKPAAGNGDGGQAYDAIQKQDGGGAGAGS-GDGDSAPDSNDQ
170 180 185
Pant_5912 SNG--KEDKGFL---GNLLGF
Paph_3137 SND--KEDKGFL---GNLLGF
mp03314 ST---QSSGGPLSSLGKMVGF
sr12947 EDAG-KKDGGLLGGIGSALGM
UHOR_02965 DEGDKKKDGGMLGGLGKMVGL
UM01987 E----KKGGGLLGGIGSMFGM
Pbra_3010 EDSGKKKDGGLLGGLGSMFGM
Phub_108 QDGDKGKGGGFLSGIGSMFGM


A
Three biological replicates
n=89
n=86
n=83
n=83
n=88
n=84
100
80
60
40
20
Symptoms rate [%]
SG200
Δumpep1
C/UmPep1
C/PaphPep1
C/PantPep1
C/PhubPep1
No symptoms
Chloroses
Small tumors
Normal tumors
Heavy tumors
Dead plants
12 dpi
B
n=100
n=92
n=101
n=94
100
80
60
40
20
Symptoms rate [%]
SG200 Wt
Δpep1
C/UmPep1
C/PbraPep1
No symptoms
Chloroses
Small tumors
Normal tumors
Heavy tumors
Dead plants