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

3

4 Rahul Sharma^{1,2}, Bilal Ökmen³, Gunther Doehlemann³, Marco Thines^{1,2,4*}

5

6 ¹Senckenberg BiK-F, Frankfurt am Main, Germany

7 ²Goethe University, Biosciences, Institute of Ecology, Evolution and Diversity, Frankfurt am
8 Main, Germany

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

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

11

12

13 * m.thines@thines-lab.eu

14   **Summary**

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

31

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

34

35 Introduction

36 Smut fungi in a broad sense are one of the three major lineages of Basidiomycota.
 37 Some species threaten crop plants, e.g. *Ustilago maydis* (maize) and *Ustilago scitaminea*
 38 (sugarcane). Smut fungi are dimorphic, with a filamentous and a yeast-like morph in different
 39 life stages. The yeast stage is haploid and saprotrophic, the hyphal stage is dikaryotic and
 40 obligate plant parasitic. Yeasts of matching sex form fusing conjugation hyphae, producing an
 41 infective dikaryon penetrating the host or colonisation. However, some smuts are known only
 42 from their yeast stage. Since the ending of dual naming for anamorphic and teleomorphic
 43 fungi (Hawksworth *et al.*, 2011), the yeast genus *Pseudozyma* became obsolete, being
 44 scattered throughout the smut phylogeny (Wang *et al.*, 2015). It was recently established that
 45 *Pseudozyma prolifica* is conspecific with *U. maydis* (Wang *et al.*, 2015), but it is believed that
 46 other *Pseudozyma* species have lost pathogenicity and exist only as yeasts (Lefebvre *et al.*,
 47 2013; Morita *et al.*, 2013; Morita *et al.*, 2014). A well-known species of these is the biocontrol
 48 agent *Pseudozyma aphidis* (Gafni *et al.*, 2015), closely related or probably conspecific with
 49 *Moesziomyces bullatus* (Kruse *et al.*, 2017).

50 A key to investigating the apathogenic or pathogenic nature of *Pseudozyma*-like
 51 yeasts are putative secreted effector proteins (PSEPs). For successful colonization of hosts,
 52 pathogens secrete hundreds of effector proteins (Latijnhouwers *et al.*, 2003). Effectors often
 53 have limited sequence conservation (Schirawski *et al.*, 2010; Kemen *et al.*, 2011; Sharma *et al.*,
 54 2014), but some are conserved in genomes of related pathogens (Sharma *et al.* 2014,
 55 2015; Quinn *et al.* 2013; Hemetsberger *et al.* 2015; Lanver *et al.* (2017). Previous definitions
 56 of putative effector proteins were often associated with several restrictions, such as size cut-
 57 off (e.g. below 300aa), certain amino acid composition (e.g. Cys-rich), or the absence of
 58 sequence motifs predicting an enzymatic function. However, such formal restrictions
 59 artificially exclude a large fraction of proteins that contribute to virulence and interfere with
 60 the plant immune system. Effectors with enzymatic functions have actually been identified in

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

69

70 Results

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

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

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

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

98

99 Discussion

100 To avoid recognition by R-proteins, many secreted effectors adapt quickly and show
101 limited conservation (Schirawski *et al.*, 2010; Laurie *et al.*, 2012). However, conservation of
102 putative secreted effector protein (PSEP)-encoding genes among related pathogens has been
103 reported (Hemetsberger *et al.*, 2015; Sharma *et al.*, 2015). These effectors have been termed
104 “core” effectors (Sharma *et al.* 2014). The well-studied Pep1 effector, required for successful
105 host colonisation (Doehlemann *et al.*, 2009), is an archetypal core effector, functionally
106 conserved among monocot and dicot infecting smut pathogens (Hemetsberger *et al.*, 2015).
107 Our finding that also Cmu1, Tin2, and Hum3 remained conserved with secretion signal
108 peptides in the genomes of smuts fungi only known from the yeast stage suggests that they
109 either feature an unknown pathogenic stage, or that effectors also have a positive effect when
110 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**

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

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

153

154 ***Δumpep1* complementation and disease assay**

155 In order to show functional conservation between the *Pep1* orthologs, those from *P.*
 156 *antarctica*, *P. hubeiensis*, *P. brasiliensis*, and *P. aphidis* were amplified by PCR with the
 157 primers given in Supplementary Table 1 and subsequently expressed in the *U. maydis*
 158 SG200 Δ pep1 strain, which is deleted for *pep1* (Doehlemann *et al.*, 2009). For proper
 159 expression, *pep1* orthologs were expressed controlled by the endogenous *U. maydis pep1*-
 160 promoter, integrated in single copy (checked by Southern blots) into the *ip*-locus of *U.*
 161 *maydis*, as described previously (Hemetsberger *et al.*, 2015).

162   The *U. maydis* disease assays were performed according to Hemetsberger *et al.*
163   (2015). Briefly, the cell culture control and complementation strains were inoculated onto
164   maize seedlings (variety Early Golden Bantam) with a syringe and needle into the leaf whirl.
165   All assays were performed in biological triplicates (≥ 30 plants each). The quantification of
166   disease symptoms was performed at 12 dpi as described previously (Kämper *et al.*, 2006).
167
168

169

170 **Acknowledgements**

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

179

References

- 180
- 181 Conesa A, Götz S, García-Gómez JM, Terol J, Talón M, Robles M. 2005. Blast2GO: A
- 182 universal tool for annotation, visualization and analysis in functional genomics
- 183 research. *Bioinformatics*. 21: 3674-3676.
- 184 Djamei A, Schipper K, Rabe F, Ghosh A, Vincon V, Kahnt J, et al. 2011. Metabolic priming
- 185 by a secreted fungal effector. *Nature* 478: 395-398.
- 186 Doehlemann G, Van Der Linde K, Aßmann D, Schwammbach D, Hof A, Mohanty A, et al.
- 187 2009. *Pep1*, a secreted effector protein of *Ustilago maydis*, is required for successful
- 188 invasion of plant cells. *PLoS Pathogens* 5: e1000290.
- 189 Doehlemann G, Reissmann S, Aßmann D, Fleckenstein M, Kahmann R. 2011. Two linked
- 190 genes encoding a secreted effector and a membrane protein are essential for *Ustilago*
- 191 *maydis*-induced tumour formation. *Molecular Microbiology* 81: 751-66.
- 192 Franceschetti M, Maqbool A, Jiménez-Dalmaroni M J, Pennington HG, Kamoun S, Banfield,
- 193 MJ. 2017. Effectors of filamentous plant pathogens: commonalities amid diversity.
- 194 *Microbiology and Molecular Biology Reviews* 81: e00066-16.
- 195 Gafni A, Claudia C, Raviv H, Buxdorf K, Avis D, Berger ML. 2015. Biological control of the
- 196 cucurbit powdery mildew pathogen *Podosphaera xanthii* by means of the epiphytic
- 197 fungus *Pseudozyma aphidis*. *Frontiers in Plant Sciences* 6: 1-11.
- 198 Hawksworth DL, Crous PW, Redhead S, Reynolds DR, Samson R, Seifert K, et al. 2011. The
- 199 Amsterdam declaration on fungal nomenclature. *IMA Fungus* 2: 105-112.
- 200 Hemetsberger C, Mueller AN, Matei A, Herrberger C, Hensel G, Kumlehn J, et al. 2015. The
- 201 fungal core effector *Pep1* is conserved across smuts of dicots and monocots. *New*
- 202 *Phytologist* 206: 1116-1126.
- 203 Kämper J, Kahmann R, Bölker M, Ma L-J, Brefort T, Saville BJ, et al. 2006. Insights from the
- 204 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-Seilanianantz 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. *Plant*
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.

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

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

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

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

241 Muller O, Schreier PH, Uhrig JF. 2008. Identification and characterization of secreted and
242 pathogenesis-related proteins in *Ustilago maydis*. *Molecular Genetics and Genomics*
243 279: 27–39.

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

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

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

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

256 Schirawski J, Mannhaupt G, Munch K, Brefort T, Schipper K, Doehlemann G, et al. 2010.
257 Pathogenicity determinants in smut fungi revealed by genome comparison. *Science*
258 330: 1546-1548.

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

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

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

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

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

275
276
277
278

279 Tables

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

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

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

285

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

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

288

289   **Figure captions**

290

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

295

296   **Figure 2. Disease assay on EGB maize lines.** For testing biological conservation, *Pep1*
 297   orthologs from *Pseudozyma* yeasts were expressed in a *Pep1* deletion background
 298   (SG200Δ*umpep1*). The restoration of pathogenicity in the complemented lines is indicative of
 299   functional conservation.

300

301

302   **Supporting information figure caption**

303

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

309

310

311

312 Tables

313

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

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

315

316

317

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

Ortholog	<i>U. maydis</i> 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	Thaumatococcus
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	M	R	F	M	L	A	T	G	A	L	F	A	L	A	S	M	L	V	S	A	D	D	A	-	-	-	-	-	-	M	P	L	P	N	Y	T	P	K	D	V	P	I	A	S	T	F	Y	W	F	S	S	V	E	V	G	V	C	Y	S	P	V	A	R	V	A	G	I	K	G	A	I	H	C	T	H	Q	G	N	Y			
Paph_3137	M	R	F	M	L	A	T	G	A	L	F	A	L	A	S	M	L	V	S	A	D	D	A	-	-	-	-	-	-	M	P	L	P	N	Y	T	P	R	D	E	P	I	A	S	T	F	Y	W	F	S	S	V	E	V	G	V	C	Y	S	P	V	A	R	V	A	G	I	K	G	A	I	H	C	T	H	Q	G	N	Y			
mp03314	M	R	I	T	L	-	T	T	V	L	A	T	F	A	I	L	L	F	A	S	V	D	Q	A	V	A	V	D	T	S	T	G	V	G	I	P	D	Y	T	A	N	A	P	V	I	A	S	S	F	I	W	F	S	S	V	E	V	G	V	C	Y	S	P	A	A	R	L	G	S	I	K	G	T	I	H	C	T	H	Q	E	K	Y
sr12947	M	R	T	T	L	V	Q	T	L	I	L	T	L	T	L	L	T	T	P	C	I	R	-	A	D	-	D	A	N	G	A	M	Q	L	P	D	F	T	P	A	K	F	P	I	A	S	T	F	Y	W	Y	S	S	V	E	V	G	V	C	Y	A	P	Q	A	R	V	A	S	I	K	G	A	I	H	C	T	H	Q	E	N	Y	
UHOR_02965	M	K	L	T	L	N	T	A	F	L	L	T	L	A	S	L	L	V	I	S	I	D	S	-	V	D	A	D	-	-	-	V	P	L	P	D	Y	T	A	H	D	Y	P	L	A	A	T	F	Y	W	F	S	S	V	E	V	G	V	C	Y	S	P	K	A	R	V	G	S	I	K	G	A	L	H	C	T	H	Q	E	N	Y	
UM01987	M	M	T	T	L	V	Q	T	T	L	L	S	L	A	L	V	L	L	G	S	T	V	P	-	V	H	A	D	A	A	G	A	V	P	L	P	N	F	K	V	D	P	Q	P	L	A	S	T	F	Y	W	F	S	S	V	E	V	G	V	C	Y	N	P	Q	A	R	V	G	S	I	K	G	A	L	H	C	T	H	Q	E	N	Y
Pbra_3010	M	K	T	T	L	L	T	A	S	F	V	S	L	A	S	L	L	F	T	S	L	T	Q	-	V	S	A	-	-	-	G	A	V	P	V	P	D	Y	T	V	H	P	L	P	L	A	S	T	F	Y	W	F	S	S	V	E	V	G	V	C	Y	S	P	T	A	R	V	A	S	I	P	G	A	L	H	C	T	H	Q	E	K	Y
Phub_108	M	K	T	T	P	S	Q	A	A	F	L	V	L	-	S	F	L	F	T	A	P	L	R	-	V	S	A	D	S	S	G	A	I	T	L	P	D	F	S	T	K	P	L	P	L	A	S	T	F	Y	W	F	S	S	V	E	V	G	V	C	Y	N	P	Q	A	R	I	G	S	I	A	G	A	L	H	C	T	H	Q	E	K	Y

	90	100	110	120	130	140	150	160																																																																										
Pant_5912	D	V	D	N	N	T	W	T	L	P	Q	T	C	V	A	L	K	P	L	G	S	E	L	S	T	A	V	R	D	S	C	K	N	A	K	G	T	F	N	V	I	K	P	A	S	A	N	A	A	G	T	Q	A	A	N	A	I	S	K	N	G	A	G	A	G	A	G	G	D	A	G	A	D	P	S	S	D	G	S	A	P	D
Paph_3137	D	V	D	N	N	T	W	T	L	P	Q	T	C	V	A	L	K	P	L	G	S	E	L	S	T	A	V	R	D	S	C	K	N	A	K	G	T	F	N	V	I	K	P	A	S	A	N	A	A	G	T	Q	A	A	N	A	I	S	K	N	G	A	G	A	G	A	G	G	D	A	G	A	D	P	S	S	D	G	S	A	P	D
mp03314	D	I	D	N	N	T	W	T	L	P	Q	T	C	V	A	L	G	P	L	G	G	P	L	S	S	G	V	R	D	A	C	T	N	A	K	G	T	F	N	V	I	T	P	A	G	A	N	T	D	G	S	Q	A	Y	N	A	I	Q	Q	Q	S	-	G	G	A	G	A	-	-	-	G	A	G	A	G	A	D	D	S	G	D	T
sr12947	D	L	D	N	N	S	W	T	L	P	Q	S	C	V	A	L	K	P	L	G	T	P	L	S	N	A	V	H	Q	S	C	V	N	A	K	G	T	W	N	V	I	K	P	A	A	T	N	A	G	G	G	Q	A	Y	D	T	I	Q	S	K	G	A	G	G	A	G	-	-	-	A	D	A	D	S	A	P	D	S	N	D	Q	
UHOR_02965	D	I	E	N	N	T	W	T	L	P	Q	T	C	V	A	L	K	P	L	G	E	P	L	S	N	A	V	R	D	S	C	T	N	A	K	G	S	F	N	L	I	K	P	A	S	S	N	A	D	G	S	Q	A	Y	N	A	I	S	N	K	G	A	D	G	A	-	-	-	-	A	S	D	P	G	S	Q	T	G	S	D	D	
UM01987	D	R	D	N	N	S	Y	T	L	P	Q	T	C	V	A	L	K	P	L	G	K	A	F	S	S	N	V	R	D	S	C	T	N	A	K	G	I	F	N	V	I	V	P	A	S	S	N	A	L	G	S	Q	A	Y	D	A	V	Q	A	K	G	-	G	T	G	G	T	G	T	-	D	D	D	T	S	A	P	D	S	N	D	Q
Pbra_3010	D	I	D	N	N	S	W	T	L	P	Q	T	C	V	A	L	K	P	L	G	G	P	L	S	S	A	V	R	D	S	C	K	N	A	K	G	T	F	N	V	I	T	P	A	G	S	N	A	E	G	G	Q	A	Y	D	A	I	Q	K	G	G	P	G	G	G	G	A	G	A	G	G	D	G	A	S	T	D	D	S	N	D	A
Phub_108	D	I	D	N	N	T	W	T	L	P	Q	T	C	V	A	L	K	P	L	G	G	P	L	S	S	A	V	R	D	S	C	K	N	A	K	G	T	F	N	V	I	K	P	A	A	G	N	G	D	G	G	Q	A	Y	D	A	I	Q	K	Q	D	G	G	G	A	G	A	G	S	-	G	D	G	D	S	A	P	D	S	N	D	Q

	170	180	185																		
Pant_5912	S	N	G	-	-	K	E	D	K	G	F	L	-	-	-	G	N	L	L	G	F
Paph_3137	S	N	D	-	-	K	E	D	K	G	F	L	-	-	-	G	N	L	L	G	F
mp03314	S	T	-	-	-	Q	S	S	G	G	P	L	S	S	L	G	K	M	V	G	F
sr12947	E	D	A	G	-	K	K	D	G	G	L	L	G	G	I	G	S	A	L	G	M
UHOR_02965	D	E	G	D	K	K	K	D	G	G	M	L	G	G	L	G	K	M	V	G	L
UM01987	E	-	-	-	-	K	K	G	G	G	L	L	G	G	I	G	S	M	F	G	M
Pbra_3010	E	D	S	G	K	K	K	D	G	G	L	L	G	G	L	G	S	M	F	G	M
Phub_108	Q	D	G	D	K	G	K	G	G	G	F	L	S	G	I	G	S	M	F	G	M

Three biological replicates

A

B