Genetics: Early Online, published on August 17, 2020 as 10.1534/genetics.120.303549

1 **Variation among biosynthetic gene clusters, secondary metabolite profiles, and cards of**
2 **virulence across *Aspergillus* species**
3
4 Jacob L. Steenwyk[1], Matthew E. Mead[1,+], Sonja L. Knowles[2,+], Huzefa A. Raja[2], Christopher D.
5 Roberts[2], Oliver Bader[3], Jos Houbraken[4], Gustavo H. Goldman[5], Nicholas H. Oberlies[2], Antonis
6 Rokas[1,*]
7
8 [1] Department of Biological Sciences, Vanderbilt University, Nashville, TN, 37235, USA
9 [2] Department of Chemistry and Biochemistry, University of North Carolina at Greensboro,
10 Greensboro, NC, USA
11 [3] Institute for Medical Microbiology, University Medical Center Göttingen, Göttingen, Germany
12 [4] Westerdijk Fungal Biodiversity Institute, Uppsalalaan 8, 3584 CT Utrecht, the Netherlands
13 [5] Faculdade de Ciências Farmacêuticas de Ribeirão Preto, Universidade de São Paulo, São Paulo,
14 Brazil
15 [+] M. E. M. and S. L. K. contributed equally to this work.
16
17 [*] Correspondence: antonis.rokas@vanderbilt.edu
18
19 **Running title:** Secondary metabolism in *A. fumigatus* and its relatives
20
21 **Keywords:** secondary metabolites, specialized metabolism, gliotoxin, chemodiversity,
22 mycotoxin, biosynthetic gene cluster, cards of virulence, pathogenicity,

Abstract

Aspergillus fumigatus is a major human pathogen. In contrast, *Aspergillus fischeri* and the recently described *Aspergillus oerlinghausenensis*, the two species most closely related to *A. fumigatus*, are not known to be pathogenic. Some of the genetic determinants of virulence (or "cards of virulence") that A. *fumigatus* possesses are secondary metabolites that impair the host immune system, protect from host immune cell attacks, or acquire key nutrients. To examine whether secondary metabolism-associated cards of virulence vary between these species, we conducted extensive genomic and secondary metabolite profiling analyses of multiple *A. fumigatus*, one *A. oerlinghausenensis*, and multiple *A. fischeri* strains. We identified two cards of virulence (gliotoxin and fumitremorgin) shared by all three species and three cards of virulence (trypacidin, pseurotin, and fumagillin) that are variable. For example, we found that all species and strains examined biosynthesized gliotoxin, which is known to contribute to virulence, consistent with the conservation of the gliotoxin biosynthetic gene cluster (BGC) across genomes. For other secondary metabolites, such as fumitremorgin, a modulator of host biology, we found that all species produced the metabolite but that there was strain heterogeneity in its production within species. Finally, species differed in their biosynthesis of fumagillin and pseurotin, both contributors to host tissue damage during invasive aspergillosis. *A. fumigatus* biosynthesized fumagillin and pseurotin, while *A. oerlinghausenensis* biosynthesized fumagillin and *A. fischeri* biosynthesized neither. These biochemical differences were reflected in sequence divergence of the intertwined fumagillin/pseurotin BGCs across genomes. These results delineate the similarities and differences in secondary metabolism-associated cards of virulence between a major fungal pathogen and its nonpathogenic closest relatives, shedding light onto the genetic and phenotypic changes associated with the evolution of fungal pathogenicity.

46 **Introduction**

47 Fungal diseases impose a clinical, economic, and social burden on humans (Drgona *et al.* 2014;
48 Vallabhaneni *et al.* 2016; Benedict *et al.* 2019). Fungi from the genus *Aspergillus* are responsible
49 for a considerable fraction of this burden, accounting for more than 250,000 infections annually
50 with high mortality rates (Bongomin *et al.* 2017). *Aspergillus* infections often result in
51 pulmonary and invasive diseases that are collectively termed aspergillosis. Among *Aspergillus*
52 species, *Aspergillus fumigatus* is the primary etiological agent of aspergillosis (Latgé and
53 Chamilos 2019).

54

55 Even though *A. fumigatus* is a major pathogen, its closest relatives are not considered pathogenic
56 (Mead *et al.* 2019a; Steenwyk *et al.* 2019; Rokas *et al.* 2020b). Numerous studies have identified
57 genetic determinants that contribute to *A. fumigatus* pathogenicity, such as the organism's ability
58 to grow well at higher temperatures and in hypoxic conditions (Kamei and Watanabe 2005;
59 Tekaia and Latgé 2005; Abad *et al.* 2010; Grahl *et al.* 2012). Genetic determinants that
60 contribute to pathogenicity could be conceived as analogous to individual "cards" of a "hand"
61 (set of cards) in a card game – that is, individual determinants are typically insufficient to cause
62 disease but can collectively do so (Casadevall 2007).

63

64 *Aspergillus fumigatus* biosynthesizes a cadre of secondary metabolites and several metabolites
65 could be conceived as "cards" of virulence because of their involvement in impairing the host
66 immune system, protecting the fungus from host immune cell attacks, or acquiring key nutrients
67 (Shwab *et al.* 2007; Losada *et al.* 2009; Yin *et al.* 2013; Wiemann *et al.* 2014; Knox *et al.* 2016;
68 Bignell *et al.* 2016; Raffa and Keller 2019; Blachowicz *et al.* 2020). For example, the secondary

69 metabolite gliotoxin has been shown in *A. fumigatus* to inhibit the host immune response (Sugui
70 *et al.* 2007; Spikes *et al.* 2008). Other secondary metabolites implicated in virulence include:
71 fumitremorgin, which inhibits the activity of the breast cancer resistance protein (González-
72 Lobato *et al.* 2010); verruculogen, which modulates the electrophysical properties of human
73 nasal epithelial cells (Khoufache *et al.* 2007); trypacidin, which is cytotoxic to lung cells and
74 inhibits phagocytosis (Gauthier *et al.* 2012; Mattern *et al.* 2015); pseurotin, which inhibits
75 immunoglobulin E (Ishikawa *et al.* 2009); and fumagillin which causes epithelial cell damage
76 (Guruceaga *et al.* 2018) and impairs the function of neutrophils (Fallon *et al.* 2010, 2011).
77
78 By extension, the metabolic pathways responsible for the biosynthesis of secondary metabolites
79 could also be conceived as components of these secondary metabolism-associated "cards" of
80 virulence. Genes in these pathways are typically organized in contiguous sets termed
81 biosynthetic gene clusters (BGCs) (Keller 2019). BGCs are known to evolve rapidly, and their
82 composition can differ substantially across species and strains (Lind *et al.* 2015, 2017; de Vries
83 *et al.* 2017; Kjærbølling *et al.* 2018, 2020; Rokas *et al.* 2018, 2020a; Vesth *et al.* 2018). For
84 example, even though *A. fumigatus* contains 33 BGCs and *A. fischeri* contains 48 BGCs, only 10
85 of those BGCs appear to be shared between the two species (Mead *et al.* 2019a). Interestingly,
86 one of the BGCs that is conserved between *A. fumigatus* and *A. fischeri* is the gliotoxin BGC and
87 both species have been shown to biosynthesize the secondary metabolite, albeit at different
88 amounts (Knowles *et al.* 2020). These results suggest that the gliotoxin "card" is part of a
89 winning "hand" that facilitates virulence only in the background of the major pathogen *A.*
90 *fumigatus* and not in that of the nonpathogen *A. fischeri* (Knowles *et al.* 2020).
91

92 To date, such comparisons of BGCs and secondary metabolite profiles among *A. fumigatus* and
93 closely related nonpathogenic species have been few and restricted to single strains (Mead *et al.*
94 2019a; Knowles *et al.* 2020). However, genetic and phenotypic heterogeneity among strains of a
95 single species is an important consideration when studying *Aspergillus* pathogenicity (Kowalski
96 *et al.* 2016, 2019; Keller 2017; Ries *et al.* 2019; Blachowicz *et al.* 2020; Bastos *et al.* 2020; Drott
97 *et al.* 2020; dos Santos *et al.* 2020; Steenwyk *et al.* 2020). Examination of multiple strains of *A.*
98 *fumigatus* and close relatives—including the recently described closest known relative of *A.*
99 *fumigatus*, *A. oerlinghausenensis*, whose virulence has yet to be examined but which is not
100 thought to be a human pathogen (Houbraken *et al.* 2016) and has never been associated with
101 human infections—will increase our understanding of the *A. fumigatus* secondary metabolism-
102 associated "cards" of virulence.
103
104 To gain insight into the genomic and chemical similarities and differences in secondary
105 metabolism among *A. fumigatus* and nonpathogenic close relatives, we characterized variation in
106 BGCs and secondary metabolites produced by *A. fumigatus* and nonpathogenic close relatives.
107 To do so, we first sequenced and assembled *A. oerlinghausenensis* CBS 139183^{T} as well as *A.*
108 *fischeri* strains NRRL 4585 and NRRL 4161 and analyzed them together with four *A. fumigatus*
109 and three additional *A. fischeri* publicly available genomes. We also characterized the secondary
110 metabolite profiles of three *A. fumigatus*, one *A. oerlinghausenensis*, and three *A. fischeri* strains.
111 We observed both variation and conservation among species- and strain-level BGCs and
112 secondary metabolites. We found that the biosynthesis of the secondary metabolites gliotoxin
113 and fumitremorgin, which are both known to interact with mammalian cells (Yamada *et al.* 2000;
114 González-Lobato *et al.* 2010; Li *et al.* 2012; Raffa and Keller 2019), as well as their BGCs, were

115 conserved among pathogenic and nonpathogenic strains. Interestingly, we found only *A. fischeri*
116 strains, but not *A. fumigatus* strains, biosynthesized verruculogen, which changes the
117 electrophysical properties of human nasal epithelial cells (Khoufache *et al.* 2007). Similarly, we
118 found that both *A. fumigatus* and *A. oerlinghausenensis* biosynthesized fumagillin and
119 trypacidin, whose effects include broad suppression of the immune response system and lung cell
120 damage (Ishikawa *et al.* 2009; Fallon *et al.* 2010, 2011; Gauthier *et al.* 2012), but *A. fischeri* did
121 not. Taken together, these results reveal that nonpathogenic close relatives of *A. fumigatus* also
122 produce some, but not all, of the secondary metabolism-associated cards of virulence known in
123 *A. fumigatus*. Further investigation of the similarities and differences among *A. fumigatus* and
124 close nonpathogenic relatives may provide additional insight into the "hand of cards" that
125 enabled *A. fumigatus* to evolve into a deadly pathogen.
126

127 **Materials and Methods**

128 **Strain acquisition, DNA extraction, and sequencing**

129 Two strains of *Aspergillus fischeri* (NRRL 4161 and NRRL 4585) were acquired from the
130 Northern Regional Research Laboratory (NRRL) at the National Center for Agricultural
131 Utilization Research in Peoria, Illinois, while one strain of *Aspergillus oerlinghausenensis* (CBS
132 139183^{T}) was acquired from the Westerdijk Fungal Biodiversity Institute, Utrecht, The
133 Netherlands. These strains were grown in 50 ml of liquid yeast extract soy peptone dextrose
134 (YESD) medium. After approximately seven days of growth on an orbital shaker (100 rpm) at
135 room temperature, the mycelium was harvested by filtering the liquid media through a
136 Corning®, 150 ml bottle top, 0.22µm sterile filter and washed with autoclaved distilled water.
137 All subsequent steps of DNA extraction from the mycelium were performed following protocols

138 outlined previously (Mead *et al.* 2019b). The genomic DNA from these three strains was
139 sequenced using a NovaSeq S4 at the Vanderbilt Technologies for Advanced Genomes facility
140 (Nashville, Tennessee, US) using paired-end sequencing (150 bp) strategy with the Illumina
141 TruSeq library kit.
142
143 **Genome assembly, quality assessment, and annotation**
144 To assemble and annotate the three newly sequenced genomes, we first quality-trimmed raw
145 sequence reads using Trimmomatic, v0.36 (Bolger *et al.* 2014) using parameters described
146 elsewhere (ILLUMINACLIP:TruSeq3-PE.fa:2:30:10, leading:10, trailing:10,
147 slidingwindow:4:20, minlen:50) (Steenwyk and Rokas 2017). The resulting paired and unpaired
148 quality-trimmed reads were used as input to the SPAdes, v3.11.1 (Bankevich *et al.* 2012),
149 genome assembly algorithm with the 'careful' parameter and the 'cov-cutoff' set to 'auto'.
150
151 We evaluated the quality of our newly assembled genomes, using metrics based on continuity of
152 assembly and gene-content completeness. To evaluate genome assemblies by scaffold size, we
153 calculated the N50 of each assembly (or the shortest contig among the longest contigs that
154 account for 50% of the genome assembly's length) (Yandell and Ence 2012). To determine gene-
155 content completeness, we implemented the BUSCO, v2.0.1 (Waterhouse *et al.* 2018), pipeline
156 using the 'genome' mode. In this mode, the BUSCO pipeline examines assembly contigs for the
157 presence of near-universally single copy orthologous genes (hereafter referred to as BUSCO
158 genes) using a predetermined database of orthologous genes from the OrthoDB, v9 (Waterhouse
159 *et al.* 2013). We used the OrthoDB database for Pezizomycotina (3,156 BUSCO genes). Each
160 BUSCO gene is determined to be present in a single copy, as duplicate sequences, fragmented, or

161 missing. Our analyses indicate the newly sequenced and assembled genomes have high gene-
162 content completeness and assembly continuity (average percent presence of BUSCO genes:
163 98.80 ± 0.10%; average N50: 451,294.67 ± 9,696.11; Fig. S1). These metrics suggest these
164 genomes are suitable for comparative genomic analyses.
165
166 To predict gene boundaries in the three newly sequenced genomes, we used the MAKER,
167 v2.31.10, pipeline (Holt and Yandell 2011) which, creates consensus predictions from the
168 collective evidence of multiple *ab initio* gene prediction software. Specifically, we created
169 consensus predictions from SNAP, v2006-07-28 (Korf 2004), and AUGUSTUS, v3.3.2 (Stanke
170 and Waack 2003), after training each algorithm individually on each genome. To do so, we first
171 ran MAKER using protein evidence clues from five different publicly available annotations of
172 *Aspergillus* fungi from section *Fumigati*. Specifically, we used protein homology clues from *A.*
173 *fischeri* NRRL 181 (GenBank accession: GCA_000149645.2), *A. fumigatus* Af293 (GenBank
174 accession: GCA_000002655.1), *Aspergillus lentulus* IFM 54703 (GenBank accession:
175 GCA_001445615.1), *Aspergillus novofumigatus* IBT 16806 (GenBank accession:
176 GCA_002847465.1), and *Aspergillus udagawae* IFM 46973 (GenBank accession:
177 GCA_001078395.1). The resulting gene predictions were used to train SNAP. MAKER was then
178 rerun using the resulting training results. Using the SNAP trained gene predictions, we trained
179 AUGUSTUS. A final set of gene boundary predictions were obtained by rerunning MAKER
180 with the training results from both SNAP and AUGUSTUS.
181
182 To supplement our data set of newly sequenced genomes, we obtained publicly available ones.
183 Specifically, we obtained genomes and annotations for *A. fumigatus* Af293 (GenBank accession:

184 GCA_000002655.1), *A. fumigatus* CEA10 (strain synonym: CBS 144.89 / FGSC A1163;
185 GenBank accession: GCA_000150145.1), *A. fumigatus* HMR AF 270 GenBank accession:
186 GCA_002234955.1), *A. fumigatus* Z5 (GenBank accession: GCA_001029325.1), *A. fischeri*
187 NRRL 181 (GenBank accession: GCA_000149645.2). We also obtained assemblies of the
188 recently published *A. fischeri* genomes for strains IBT 3003 and IBT 3007 (Zhao *et al.* 2019)
189 which, lacked annotations. We annotated the genome of each strain individually using MAKER
190 with the SNAP and AUGUSTUS training results from a close relative of both strains, *A. fischeri*
191 NRRL 4161. Altogether, our final data set contained a total of ten genome from three species:
192 four *A. fumigatus* strains, one *A. oerlinghausenensis* strain, and five *A. fischeri* strains (Table 1).
193

194 **Maximum likelihood phylogenetics and Bayesian estimation of divergence times**

195 To reconstruct the evolutionary history among the ten *Aspergillus* genomes, we implemented a
196 recently developed pipeline (Steenwyk *et al.* 2019), which relies on the concatenation-approach
197 to phylogenomics (Rokas *et al.* 2003) and has been successfully used in reconstructing species-
198 level relationships among *Aspergillus* and *Penicillium* fungi (Steenwyk *et al.* 2019; Bodinaku *et*
199 *al.* 2019). The first step in the pipeline is to identify single copy orthologous genes in the
200 genomes of interest which, are ultimately concatenated into a larger phylogenomic data matrix.
201 To identify single copy BUSCO genes across all ten *Aspergillus* genomes, we used the BUSCO
202 pipeline with the Pezizomycotina database as described above. We identified 3,041 BUSCO
203 genes present at a single copy in all ten *Aspergillus* genomes and created multi-FASTA files for
204 each BUSCO gene that contained the protein sequences for all ten taxa. The protein sequences of
205 each BUSCO gene were individually aligned using Mafft, v7.4.02 (Katoh and Standley 2013),
206 with the same parameters as described elsewhere (Steenwyk *et al.* 2019). Nucleotide sequences

207 were then mapped onto the protein sequence alignments using a custom Python, v3.5.2
208 (https://www.python.org/), script with BioPython, v1.7 (Cock *et al.* 2009). The resulting codon-
209 based alignments were trimmed using trimAl, v1.2.rev59 (Capella-Gutierrez *et al.* 2009), with
210 the 'gappyout' parameter. The resulting trimmed nucleotide alignments were concatenated into a
211 single matrix of 5,602,272 sites and was used as input into IQ-TREE, v1.6.11 (Nguyen *et al.*
212 2015). The best-fitting model of substitutions for the entire matrix was determined using
213 Bayesian information criterion values (Kalyaanamoorthy *et al.* 2017). The best-fitting model was
214 a general time-reversible model with empirical base frequencies that allowed for a proportion of
215 invariable sites and a discrete Gamma model with four rate categories (GTR+I+F+G4) (Tavaré
216 1986; Yang 1994, 1996; Vinet and Zhedanov 2011). To evaluate bipartition support, we used
217 5,000 ultrafast bootstrap approximations (Hoang *et al.* 2018).
218
219 To estimate divergence times among the ten *Aspergillus* genomes, we used the concatenated data
220 matrix and the resulting maximum likelihood phylogeny from the previous steps as input to
221 Bayesian approach implemented in MCMCTree from the PAML package, v4.9d (Yang 2007).
222 First, we estimated the substitution rate across the data matrix using a "GTR+G" model of
223 substitutions (model = 7), a strict clock model, and the maximum likelihood phylogeny rooted on
224 the clade of *A. fischeri* strains. We imposed a root age of 3.69 million years ago according to
225 results from recent divergence time estimates of the split between *A. fischeri* and *A. fumigatus*
226 (Steenwyk *et al.* 2019). We estimated the substitution rate to be 0.005 substitutions per one
227 million years. Next, the likelihood of the alignment was approximated using a gradient and
228 Hessian matrix. To do so, we used previously established time constraints for the split between
229 *A. fischeri* and *A. fumigatus* (1.85 to 6.74 million years ago) (Steenwyk *et al.* 2019). Lastly, we

230 used the resulting gradient and Hessian matrix, the rooted maximum likelihood phylogeny, and
231 the concatenated data matrix to estimate divergence times using a relaxed molecular clock
232 (model = 2). We specified the substitution rate prior based on the estimated substitution rate
233 (rgene_gamma = 1 186.63). The 'sigma2_gamma' and 'finetune' parameters were set to '1 4.5'
234 and '1', respectively. To collect a high-quality posterior probability distribution, we ran a total of
235 5.1 million iterations during MCMC analysis which, is 510 times greater than the minimum
236 recommendations (Raftery and Lewis 1995). Our sampling strategy across the 5.1 million
237 iterations was to discard the first 100,000 results followed by collecting a sample every 500th
238 iteration until a total of 10,000 samples were collected.
239
240 **Identification of gene families and analyses of putative biosynthetic gene clusters**
241 To identify gene families across the ten *Aspergillus* genomes, we used a Markov clustering
242 approach. Specifically, we used OrthoFinder, v2.3.8 (Emms and Kelly 2019). OrthoFinder first
243 conducts a blast all-vs-all using the protein sequences of all ten *Aspergillus* genomes and NCBI's
244 Blast+, v2.3.0 (Camacho *et al.* 2009), software. After normalizing blast bit scores, genes are
245 clustered into discrete orthogroups using a Markov clustering approach (van Dongen 2000). We
246 clustered genes using an inflation parameter of 1.5. The resulting orthogroups were used proxies
247 for gene families.
248
249 To identify putative biosynthetic gene clusters (BGCs), we used the gene boundaries predictions
250 from the MAKER software as input into antiSMASH, v4.1.0 (Weber *et al.* 2015). To identify
251 homologous BGCs across the ten *Aspergillus* genomes, we used the software BiG-SCAPE,
252 v20181005 (Navarro-Muñoz *et al.* 2020). Based on the Jaccard Index of domain types, sequence

253 similarity among domains, and domain adjacency, BiG-SCAPE calculates a similarity metric
254 between pairwise combinations of clusters where smaller values indicate greater BGC similarity.
255 BiG-SCAPE's similarity metric can then be used as an edge-length in network analyses of
256 cluster similarity. We evaluated networks using an edge-length cutoff from 0.1-0.9 with a step of
257 0.1 (Fig. S3). We found networks with an edge-length cutoff of 0.4-0.6 to be similar and based
258 further analyses on a cutoff of 0.5. Because BiG-SCAPE inexplicably split the gliotoxin BGC of
259 the *A. fumigatus* Af293 strain into two cluster families even though the BGC was highly similar
260 to the gliotoxin BGCs of all other strains, we supplemented BiG-SCAPE's approach to
261 identifying homologous BGCs with visualize inspection of microsyteny and blast-based analyses
262 using NCBI's BLAST+, v2.3.0 (Camacho *et al.* 2009) for BGCs of interest. Similar sequences in
263 microsynteny analyses were defined as at least 100 bp in length, at least 30 percent similarity,
264 and an expectation value threshold of 0.01. Lastly, to determine if any BGCs have been
265 previously linked to secondary metabolites, we cross referenced BGCs and BGC families with
266 those found in the MIBiG database (Kautsar *et al.* 2019) as well as previously published *A.*
267 *fumigatus* BGCs (Table S2). BGCs not associated with secondary metabolites were considered to
268 likely encode for unknown compounds.
269
270 **Identification and characterization of secondary metabolite production**
271 **General experimental procedures**
272 The ^{1}H NMR data were collected using a JOEL ECS-400 spectrometer, which was equipped
273 with a JOEL normal geometry broadband Royal probe, and a 24-slot autosampler, and operated
274 at 400 MHz. HRESIMS experiments utilized either a Thermo LTQ Orbitrap XL mass
275 spectrometer or a Thermo Q Exactive Plus (Thermo Fisher Scientific); both were equipped with

276 an electrospray ionization source. A Waters Acquity UPLC (Waters Corp.) was utilized for both
277 mass spectrometers, using a BEH C_{18} column (1.7 μm; 50 mm x 2.1 mm) set to a temperature of
278 40°C and a flow rate of 0.3 ml/min. The mobile phase consisted of a linear gradient of CH_3CN-
279 H_2O (both acidified with 0.1% formic acid), starting at 15% CH_3CN and increasing linearly to
280 100% CH_3CN over 8 min, with a 1.5 min hold before returning to the starting condition. The
281 HPLC separations were performed with Atlantis T3 C_{18} semi-preparative (5 μm; 10 x 250 mm)
282 and preparative (5 μm; 19 x 250 mm) columns, at a flow rate of 4.6 ml/min and 16.9 ml/min,
283 respectively, with a Varian Prostar HPLC system equipped with a Prostar 210 pumps and a
284 Prostar 335 photodiode array detector (PDA), with the collection and analysis of data using
285 Galaxie Chromatography Workstation software. Flash chromatography was performed on a
286 Teledyne ISCO Combiflash Rf 200 and monitored by both ELSD and PDA detectors.
287
288 **Chemical characterization**
289 To identify the secondary metabolites that were biosynthesized by *A. fumigatus*, *A.*
290 *oerlinghausenensis*, and *A. fischeri*, these strains were grown as large-scale fermentations to
291 isolate and characterize the secondary metabolites. To inoculate oatmeal cereal media (Old
292 fashioned breakfast Quaker oats), agar plugs from fungal stains grown on potato dextrose agar;
293 difco (PDA) were excised from the edge of the Petri dish culture and transferred to separate
294 liquid seed media that contained 10 ml YESD broth (2% soy peptone, 2% dextrose, and 1% yeast
295 extract; 5 g of yeast extract, 10 g of soy peptone, and 10 g of D-glucose in 500 ml of deionized
296 H_2O) and allowed to grow at 23˚C with agitation at 100 rpm for three days. The YESD seed
297 cultures of the fungi were subsequently used to inoculate solid-state oatmeal fermentation
298 cultures, which were either grown at room temperature (approximately 23˚C under 12h

299 light/dark cycles for 14 days), 30˚C, or 37˚C; all growths at the latter two temperatures were
300 carried out in an incubator (VWR International) in the dark over four days. The oatmeal cultures
301 were prepared in 250 ml Erlenmeyer flasks that contained 10 g of autoclaved oatmeal (10 g of
302 oatmeal with 17 ml of deionized H_2O and sterilized for 15–20 minutes at 121°C). For all fungal
303 strains three flasks of oatmeal cultures were grown at all three temperatures, except for *A.*
304 *oerlinghausenensis* (CBS 139183^T) at room temperature and *A. fumigatus* (Af293) at 37˚C. For
305 CBS 139183^T, the fungal cultures were grown in four flasks, while for Af293 eight flasks were
306 grown in total. The growths of these two strains were performed differently from the rest because
307 larger amounts of extract were required in order to perform detailed chemical characterization.
308 The cultures were extracted by adding 60 ml of (1:1) MeOH-$CHCl_3$ to each 250 ml flask,
309 chopping thoroughly with a spatula, and shaking overnight (~ 16 h) at ~ 100 rpm at room
310 temperature. The culture was filtered *in vacuo*, and 90 ml $CHCl_3$ and 150 ml H_2O were added to
311 the filtrate. The mixture was stirred for 30 min and then transferred to a separatory funnel. The
312 organic layer ($CHCl_3$) was drawn off and evaporated to dryness *in vacuo*. The dried organic layer
313 was reconstituted in 100 ml of (1:1) MeOH–CH_3CN and 100 ml of hexanes, transferred to a
314 separatory funnel, and shaken vigorously. The defatted organic layer (MeOH–CH_3CN) was
315 evaporated to dryness *in vacuo*.
316
317 To isolate compounds, the defatted extract was dissolved in $CHCl_3$, absorbed onto Celite 545
318 (Acros Organics), and fractioned by normal phase flash chromatography using a gradient of
319 hexane-$CHCl_3$-MeOH. *Aspergillus fischeri* strain NRRL 181 was chemically characterized
320 previously (Mead *et al.* 2019a; Knowles *et al.* 2019). *A. fumigatus* strain Af293, grown at 37°C,
321 was subjected to a 12g column at a flow rate of 30 ml/min and 61.0 column volumes, which

322 yielded four fractions. Fraction 2 was further purified via preparative HPLC using a gradient
323 system of 30:70 to 100:0 of CH_3CN-H_2O with 0.1% formic acid over 40 min at a flow rate of
324 16.9 ml/min to yield six subfractions. Subfractions 1, 2 and 5, yielded cyclo(L-Pro-L-Leu) (Li *et*
325 *al.* 2008) (0.89 mg), cyclo(L-Pro-L-Phe) (Campbell *et al.* 2009) (0.71 mg), and
326 monomethylsulochrin (Ma *et al.* 2004) (2.04 mg), which eluted at approximately 5.7, 6.3, and
327 10.7 min, respectively. Fraction 3 was further purified via preparative HPLC using a gradient
328 system of 40:60 to 65:35 of CH_3CN-H_2O with 0.1% formic acid over 30 min at a flow rate of
329 16.9 ml/min to yield four subfractions. Subfractions 1 and 2 yielded pseurotin A (Wang *et al.*
330 2011) (12.50 mg) and bisdethiobis(methylthio)gliotoxin (Afiyatullov *et al.* 2005) (13.99 mg),
331 which eluted at approximately 7.5 and 8.0 min, respectively.
332
333 *A. fumigatus* strain CEA10, grown at 37°C, was subjected to a 4g column at a flow rate of 18
334 ml/min and 90.0 column volumes, which yielded five fractions. Fraction 1 was purified via
335 preparative HPLC using a gradient system of 50:50 to 100:0 of CH_3CN-H_2O with 0.1% formic
336 acid over 45 min at a flow rate of 16.9 ml/min to yield eight subfractions. Subfraction 1, yielded
337 fumagillin (Halász *et al.* 2000) (1.69 mg), which eluted at approximately 18.5 min. Fraction 2
338 was purified via semi-preparative HPLC using a gradient system of 35:65 to 80:20 of CH_3CN-
339 H_2O with 0.1% formic acid over 30 min at a flow rate of 4.6 ml/min to yield 10 subfractions.
340 Subfraction 5 yielded fumitremorgin C (Kato *et al.* 2009) (0.25 mg), which eluted at
341 approximately 15.5 min. Fraction 3 was purified via preparative HPLC using a gradient system
342 of 40:60 to 100:0 of CH_3CN-H_2O with 0.1% formic acid over 30 min at a flow rate of 16.9
343 ml/min to yield nine subfractions. Subfraction 2 yielded pseurotin A (1.64 mg), which eluted at
344 approximately 7.3 min.

345

346 *Aspergillus oerlinghausenensis* strain CBS 139183^{T}, grown at RT, was subjected to a 4g column
347 at a flow rate of 18 ml/min and 90 column volumes, which yielded 4 fractions. Fraction 3 was
348 further purified via preparative HPLC using a gradient system of 35:65 to 70:30 of CH_3CN-H_2O
349 with 0.1% formic acid over 40 min at a flow rate of 16.9 ml/min to yield 11 subfractions.
350 Subfractions 3 and 10 yielded spiro [5H,10H-dipyrrolo[1,2-a:1′,2′-d]pyrazine-2-(3H),2′-
351 [2H]indole]-3′,5,10(1′H)-trione (Wang *et al.* 2008) (0.64 mg) and helvolic acid (Zhao *et al.*
352 2010) (1.03 mg), which eluted at approximately 11.5 and 39.3 min, respectively. (see NMR
353 supporting information; figshare: https://doi.org/10.6084/m9.figshare.12055503).
354

355 **Metabolite profiling by mass spectrometry**

356 The metabolite profiling by mass spectrometry, also known as dereplication, was performed as
357 stated previously (El-Elimat *et al.* 2013). Briefly, ultraperformance liquid chromatography-
358 photodiode array-electrospray ionization high resolution tandem mass spectrometry (UPLC-
359 PDA-HRMS-MS/MS) was utilized to monitor for secondary metabolites across all strains
360 (Af293, CEA10, CEA17, CBS 139183^{T}, NRRL 181, NRRL 4161, and NRRL 4585). Utilizing
361 positive-ionization mode, ACD MS Manager with add-in software IntelliXtract (Advanced
362 Chemistry Development, Inc.; Toronto, Canada) was used for the primary analysis of the UPLC-
363 MS chromatograms. The data from 19 secondary metabolites are provided in the Supporting
364 Information (see Dereplication table; figshare: https://doi.org/10.6084/m9.figshare.12055503),
365 which for each secondary metabolite lists: molecular formula, retention time, UV-absorption
366 maxima, high-resolution full-scan mass spectra, and MS-MS data (top 10 most intense peaks).
367

368 **Metabolomics analyses**

369 Principal component analysis (PCA) analysis was performed on the UPLC-MS data. Untargeted
370 UPLC-MS datasets for each sample were individually aligned, filtered, and analyzed using
371 MZmine 2.20 software (https://sourceforge.net/projects/mzmine/) (Pluskal *et al.* 2010). Peak
372 detection was achieved using the following parameters, *A. fumigatus* at (Af293, CEA10, and
373 CEA17): noise level (absolute value), 1×10^6; minimum peak duration, 0.05 min; *m/z* variation
374 tolerance, 0.05; and *m/z* intensity variation, 20%; *A. fischeri* (NRRL 181, NRRL 4161, and
375 NRRL 4585): noise level (absolute value), 1×10^6; minimum peak duration, 0.05 min; *m/z*
376 variation tolerance, 0.05; and *m/z* intensity variation, 20%; and all strains (Af293, CEA10,
377 CEA17, CBS 139183^T, NRRL 181, NRRL 4161, and NRRL 4585): noise level (absolute value),
378 7×10^5; minimum peak duration, 0.05 min; *m/z* variation tolerance, 0.05; and *m/z* intensity
379 variation, 20%. Peak list filtering and retention time alignment algorithms were used to refine
380 peak detection. The join algorithm integrated all sample profiles into a data matrix using the
381 following parameters: *m/z* and retention time balance set at 10.0 each, *m/z* tolerance set at 0.001,
382 and RT tolerance set at 0.5 mins. The resulting data matrix was exported to Excel (Microsoft) for
383 analysis as a set of *m/z* – retention time pairs with individual peak areas detected in triplicate
384 analyses. Samples that did not possess detectable quantities of a given marker ion were assigned
385 a peak area of zero to maintain the same number of variables for all sample sets. Ions that did not
386 elute between 2 and 8 minutes and/or had an *m/z* ratio less than 200 or greater than 800 Da were
387 removed from analysis. Relative standard deviation was used to understand the quantity of
388 variance between the technical replicate injections, which may differ slightly based on
389 instrument variance. A cutoff of 1.0 was used at any given *m/z* – retention time pair across the
390 technical replicate injections of one biological replicate, and if the variance was greater than the

391 cutoff, it was assigned a peak area of zero. Final chemometric analysis, data filtering (Caesar *et*
392 *al.* 2018) and PCA was conducted using Sirius, v10.0 (Pattern Recognition Systems AS)
393 (Kvalheim *et al.* 2011), and dendrograms were created with Python. The PCA scores plots were
394 generated using data from either the three individual biological replicates or the averaged
395 biological replicates of the fermentations. Each biological replicate was plotted using averaged
396 peak areas obtained across four replicate injections (technical replicates).
397

398 **Data Availability**

399 Sequence reads and associated genome assemblies generated in this project are available in
400 NCBI's GenBank database under the BioProject PRJNA577646. Additional descriptions of the
401 genomes including predicted gene boundaries are available through figshare
402 (https://doi.org/10.6084/m9.figshare.12055503). The figshare repository is also populated with
403 other data generated from genomic and natural products analysis. Among genomic analyses, we
404 provide information about predicted BGCs, results associated with network-based clustering of
405 BGCs into cluster families, phylogenomic data matrices, and trees. Among natural products
406 analysis, we provide information that supports methods and results, including NMR spectra.
407

408 **Results**

409 **Conservation and diversity of biosynthetic gene clusters within and between species**

410 We sequenced and assembled *A. oerlinghausenensis* CBS 139183^{T} and *A. fischeri* strains NRRL
411 4585 and NRRL 4161. Together with publicly available genomes, we analyzed 10 *Aspergillus*
412 genomes (five *A. fischeri* strains; four *A. fumigatus* strains; one *A. oerlinghausenensis* strain; see
413 Methods). We found that the newly added genomes were of similar quality to other publicly

414 available draft genomes (average percent presence of BUSCO genes: 98.80 ± 0.10%; average
415 N50: 451,294.67 ± 9,696.11; Fig. S1). We predicted that *A. oerlinghausenensis* CBS 139183^{T}, *A.*
416 *fischeri* NRRL 4585, and *A. fischeri* NRRL 4161 have 10,044, 11,152 and 10,940 genes,
417 respectively, numbers similar to publicly available genomes. Lastly, we inferred the evolutionary
418 history of the 10 *Aspergillus* genomes using a concatenated matrix of 3,041 genes (5,602,272
419 sites) and recapitulated species-level relationships as previously reported (Houbraken *et al.*
420 2016). Relaxed molecular clock analyses suggested that *A. oerlinghausenensis* CBS 139183^{T}
421 diverged from *A. fumigatus* approximately 3.9 (6.4 – 1.3) million years ago and that *A.*
422 *oerlinghausenensis* and *A. fumigatus* split from *A. fischeri* approximately 4.5 (6.8 – 1.7) million
423 years ago (Fig. 1A; Fig. S2).
424
425 Examination of the total number of predicted BGCs revealed that *A. fischeri* has the largest BGC
426 count. Among *A. fumigatus*, *A. oerlinghausenensis*, and *A. fischeri*, we predicted an average of
427 35.75 ± 2.22, 40, 50.80 ± 2.17 BGCs, respectively, and found they spanned diverse biosynthetic
428 classes (e.g., polyketides, non-ribosomal peptides, terpenes, etc.) (Fig. 1B). Network-based
429 clustering of BGCs into cluster families (or groups of homologous BGCs) resulted in
430 qualitatively similar networks when we used moderate similarity thresholds (or edge cut-off
431 values; Fig. S3A). Using a (moderate) similarity threshold of 0.5, we inferred 88 cluster families
432 of putatively homologous BGCs (Fig. 1C).
433
434 Examination of BGCs revealed extensive presence and absence polymorphisms within and
435 between species. We identified 17 BGCs that were present in all 10 *Aspergillus* genomes
436 including the hexadehydroastechrome (HAS) BGC (cluster family 311 or CF311), the

437 neosartoricin BGC (CF61), and other putative BGCs likely encoding unknown products (Fig.
438 S3B; Table S1; data available from figshare, https://doi.org/10.6084/m9.figshare.12055503). In
439 contrast, we identified 18 BGCs found in single strains, which likely encode unknown products.
440 Between species, similar patterns of broadly present and species-specific BGCs were observed.
441 For example, we identified 18 BGCs that were present in at least one strain across all species; in
442 contrast, *A. fumigatus*, *A. oerlinghausenensis*, and *A. fischeri* had 16, 8, and 27 BGCs present in
443 at least one strain but absent from the other species, respectively. These results suggest each
444 species has a largely distinct repertoire of BGCs.
445
446 Examination of shared BGCs across species revealed *A. oerlinghausenensis* CBS139183^{T} and *A.*
447 *fischeri* shared more BGCs with each other than either did with *A. fumigatus*. Surprisingly, we
448 found ten homologous BGCs between *A. oerlinghausenensis* CBS 139183^{T} and *A. fischeri* but
449 only three homologous BGCs shared between *A. fumigatus* and *A. oerlinghausenensis* CBS
450 139183^{T} (Fig. 2A; Fig. S3C) even though *A. oerlinghausenensis* is more closely related to *A.*
451 *fumigatus* than to *A. fischeri* (Fig. 1A). BGCs shared by *A. oerlinghausenensis* CBS 139183^{T} and
452 *A. fischeri* were uncharacterized while BGCs present in both *A. fumigatus* and *A.*
453 *oerlinghausenensis* CBS 139183^{T} included those that encode fumigaclavine and
454 fumagillin/pseurotin. Lastly, to associate each BGC with a secondary metabolite in *A. fumigatus*
455 Af293, we cross referenced our list with a publicly available one (Table S2) (Lind *et al.* 2017).
456 Importantly, all known *A. fumigatus* Af293 BGCs were represented in our analyses.
457
458 At the level of gene families, there were few species-specific gene families in *A.*
459 *oerlinghausenensis* (Fig. 2B). *A. oerlinghausenensis* CBS 139183^{T} has only eight species-

460 specific gene families, whereas *A. fischeri* and *A. fumigatus* have 1,487 and 548 species-specific
461 gene families, respectively. Examination of the best BLAST hits of the eight species-specific
462 gene families suggest that most are hypothetical or uncharacterized fungal genes. To determine if
463 the eight *A. oerlinghausenensis* CBS 139183^{T} specific gene families were an artifact of using a
464 single representative strain, we conducted and additional ortholog clustering analysis using a
465 single strain of *A. fischeri* (NRRL 181), a single strain of *A. fumigatus* (Af293), or a single strain
466 of each species (CBS 139183, NRRL 181, Af293). When using a single strain of *A. fischeri* or *A.*
467 *fumigatus*, there were 23 or six gene families unique to each species, respectively. Therefore, the
468 low number of *A. oerlinghausenensis*-specific gene families likely stems from our use of the
469 genome of a single strain.
470
471 Despite a closer evolutionary relationship between *A. oerlinghausenensis* and *A. fumigatus*, we
472 found *A. oerlinghausenensis* shares more gene families with *A. fischeri* than with *A. fumigatus*
473 (685 and 109, respectively) suggestive of extensive gene loss in the *A. fumigatus* stem lineage.
474 Lastly, we observed strain heterogeneity in gene family presence and absence within both *A.*
475 *fumigatus* and *A. fischeri* (Fig. S4). For example, the largest intersection that does not include all
476 *A. fischeri* strains is 493 gene families, which were found in all but one strain, NRRL 181. For *A.*
477 *fumigatus*, the largest intersection that does not include all strains is 233 gene families, which
478 were shared by strains Af293 and CEA10.
479
480 **Within and between species variation in secondary metabolite profiles of *A. fumigatus* and**
481 **its closest relatives**

482 To gain insight into variation in secondary metabolite profiles within and between species, we
483 profiled *A. fumigatus* strains Af293, CEA10, and CEA17 (a *pyrG1/URA3* derivative of CEA10),
484 *A. fischeri* strains NRRL 181, NRRL 4585, and NRRL 4161, and *A. oerlinghausenensis* CBS
485 139183^{T} for secondary metabolites. Specifically, we used three different procedures, including
486 the isolation and structure elucidation of metabolites, where possible, followed by two different
487 metabolite profiling procedures that use mass spectrometry techniques. Altogether, we isolated
488 and characterized 19 secondary metabolites; seven from *A. fumigatus*, two from *A.*
489 *oerlinghausenensis*, and ten from *A. fischeri* (Fig. S5). These products encompassed a wide
490 diversity of secondary metabolite classes, such as those derived from polyketide synthases, non-
491 ribosomal peptide-synthetases, terpene synthases and mixed biosynthesis enzymes.
492
493 To characterize the secondary metabolites biosynthesized that were not produced in high enough
494 quantity for structural identification through traditional isolation methods, we employed
495 "dereplication" mass spectrometry protocols specific to natural products research on all tested
496 strains at both 30°C and 37°C (see supporting information, dereplication example; figshare:
497 https://doi.org/10.6084/m9.figshare.12055503) (El-Elimat *et al.* 2013; Ito and Masubuchi 2014;
498 Gaudêncio and Pereira 2015; Hubert *et al.* 2017). We found that most secondary metabolites
499 were present across strains of the same species (Table S3); for example, monomethylsulochrin
500 was isolated from *A. fumigatus* Af293, but through metabolite profiling, its spectral features were
501 noted also in *A. fumigatus* strains CEA10 and CEA17. We identified metabolites that were
502 biosynthesized by only one species; for example, pseurotin A was solely present in *A. fumigatus*
503 strains. Finally, we found several secondary metabolites that were biosynthesized across species,
504 such as fumagillin, which was biosynthesized by *A. fumigatus* and *A. oerlinghausenensis*, and

505 fumitremorgin B, which was biosynthesized by strains of both *A. oerlinghausenensis* and *A.*
506 *fischeri.* Together, these analyses suggest that closely related *Aspergillus* species and strains
507 exhibit variation both within as well as between species in the secondary metabolites produced.
508
509 To further facilitate comparisons of secondary metabolite profiles within and between species,
510 we used the 1,920 features (i.e., unique *m/z* – retention time pairs) that were identified from all
511 strains at all temperatures (Fig. 3A), to perform hierarchical clustering (Fig. 3B) and Principal
512 Components Analysis (PCA) (Fig. S6). Hierarchical clustering at 37°C and 30°C indicated the
513 chromatogram of *A. oerlinghausenensis* CBS 139183^{T} is more similar to the chromatogram of *A.*
514 *fischeri* than to that of *A. fumigatus*. PCA results were broadly consistent with the clustering
515 results, but suggested that *A. oerlinghausenensis* was just as similar to *A. fischeri* strains as it was
516 to *A. fumigatus* strains. This difference likely stems from the fact that hierarchical clustering is a
517 total-evidence approach whereas PCA captures most but not all variance in the data (e.g., the two
518 principal components in Fig. S6B and S6C capture 84.6% of the total variance). PCA analysis
519 revealed greater variation in secondary metabolite production at 30°C compared to 37°C (Fig.
520 S6), suggesting there is a more varied response in how BGCs are being utilized at 30°C. PCA at
521 both 37°C and 30°C showed that variation between *A. oerlinghausenensis* CBS 139183^{T} and *A.*
522 *fischeri* strains was largely captured along the second principal component; in contrast, the
523 differences between *A. oerlinghausenensis* CBS 139183^{T} and *A. fumigatus* strains are captured
524 along the first principal component (Fig. S6D-E). Taken together, these results suggest that the
525 three *A. fischeri* strains and *A. oerlinghausenensis* were the most chemically similar to each
526 other.
527

528 In summary, even though *A. oerlinghausenensis* is phylogenetically more closely related to *A.*
529 *fumigatus* than to *A. fischeri* (Fig. 1A), our chemical analyses suggest that the secondary
530 metabolite profile of *A. oerlinghausenensis* is more similar to the profile of *A. fischeri* than it is
531 to the profile of *A. fumigatus* (Fig. 3B and S6B-E). The similarity of secondary metabolite
532 profiles of *A. oerlinghausenensis* and *A. fischeri* is consistent with our finding that the genome of
533 *A. oerlinghausenensis* shares higher numbers of BGCs and gene families with *A. fischeri* than
534 with *A. fumigatus* (Fig. 2). The broad clustering patterns in secondary metabolite-based plots
535 (Fig. S6B-E) are less robust than, but consistent with, those of BGC-based plots (Fig. S6A),
536 suggesting that the observed similarities in the secondary metabolism-associated genotypes of *A.*
537 *oerlinghausenensis* and *A. fischeri* are likely reflected in their chemotypes.
538

539 **Conservation and divergence among biosynthetic gene clusters implicated in *A. fumigatus***
540 **pathogenicity**

541 Secondary metabolites are known to play a role in *A. fumigatus* virulence (Raffa and Keller
542 2019). We therefore conducted a focused examination of specific *A. fumigatus* BGCs and
543 secondary metabolites that have been previously implicated in the organism's ability to cause
544 human disease (Table 2). We found varying degrees of conservation and divergence that were
545 associated with the absence or presence of a secondary metabolite. Among conserved BGCs that
546 were also associated with conserved secondary metabolite production, we highlight the
547 mycotoxins gliotoxin and fumitremorgin. Interestingly, we note that only *A. fischeri* strains
548 synthesized verruculogen, a secondary metabolite that is implicated in human disease and is
549 encoded by the fumitremorgin BGC (Khoufache *et al.* 2007; Kautsar *et al.* 2019). Among BGCs
550 that exhibited varying degrees of sequence divergence and divergence in their production of the

551 corresponding secondary metabolites, we highlight those associated with the production of the
552 trypacidin and fumagillin/pseurotin secondary metabolites. We found that nonpathogenic close
553 relatives of *A. fumigatus* produced some but not all mycotoxins, which provides novel insight
554 into the unique cocktail of secondary metabolites biosynthesized by *A. fumigatus*.
555
556 ***Gliotoxin.*** Gliotoxin is a highly toxic compound and known virulence factor in *A. fumigatus*
557 (Sugui *et al.* 2007). Nearly identical BGCs encoding gliotoxin are present in all pathogenic (*A.*
558 *fumigatus*) and nonpathogenic (*A. oerlinghausenensis* and *A. fischeri*) strains examined (Fig. 4).
559 Additionally, we found that all examined strains synthesized bisdethiobis(methylthio)gliotoxin a
560 derivative from dithiogliotoxin, involved in the down-regulation of gliotoxin biosynthesis (Dolan
561 *et al.* 2014), one of the main mechanisms of gliotoxin resistance in *A. fumigatus* (Kautsar *et al.*
562 2019).
563
564 ***Fumitremorgin and Verruculogen.*** Similarly, there is a high degree of conservation in the
565 BGC that encodes fumitremorgin across all strains (Fig. 5). Fumitremorgins have known
566 antifungal activity, are lethal to brine shrimp, and are implicated in inhibiting mammalian
567 proteins responsible for resistance to anticancer drugs in mammalian cells (Raffa and Keller
568 2019). We found that conservation in the fumitremorgin BGC is associated with the production
569 of fumitremorgins in all isolates examined. The fumitremorgin BGC is also responsible for the
570 production of verruculogen, which is implicated to aid in *A. fumigatus* pathogenicity by changing
571 the electrophysical properties of human nasal epithelial cells (Khoufache *et al.* 2007).
572 Interestingly, we found that only *A. fischeri* strains produced verruculogen under the conditions
573 we analyzed.

574

575 ***Trypacidin.*** Examination of the trypacidin BGC, which encodes a spore-borne and cytotoxic
576 secondary metabolite, revealed a conserved cluster found in four pathogenic and nonpathogenic
577 strains: *A. fumigatus* Af293, *A. fumigatus* CEA10, *A. oerlinghausenensis* CBS 139183^{T}, and *A.*
578 *fischeri* NRRL 181 (Fig. S7). Furthermore, we found that three of these four isolates (except *A.*
579 *fischeri* NRRL 181) biosynthesized a trypacidin analog, monomethylsulochrin. Examination of
580 the microsynteny of the trypacidin BGC revealed that it was conserved across all four genomes
581 with the exception *A. fischeri* NRRL l81, which lacked a RING (Really Interesting New Gene)
582 finger gene. Interestingly, RING finger proteins can mediate gene transcription (Poukka *et al.*
583 2000). We confirmed the absence of the RING finger protein by performing a sequence
584 similarity search with the *A. fumigatus* Af293 RING finger protein (AFUA_4G14620;
585 EAL89333.1) against the *A. fischeri* NRRL 181 genome. In the homologous locus in *A. fischeri*,
586 we found no significant BLAST hit for the first 23 nucleotides of the RING finger gene
587 suggestive of pseudogenization. Taken together, we hypothesize that presence/absence
588 polymorphisms or a small degree of sequence divergence between otherwise homologous BGCs
589 may be responsible for the presence or absence of a toxic secondary metabolite in *A. fischeri*
590 NRRL 181. Furthermore, inter- and intra-species patterns of trypacidin presence and absence
591 highlight the importance of strain heterogeneity when examining BGCs.

592

593 ***Fumagillin/pseurotin.*** Examination of the intertwined fumagillin/pseurotin BGCs
594 revealed that fumagillin has undergone substantial sequence divergence and that pseurotin is
595 absent from strains of *A. fischeri*. The fumagillin/pseurotin BGCs are under the same regulatory
596 control (Wiemann *et al.* 2013) and biosynthesize secondary metabolites that cause cellular

damage during host infection (fumagillin (Guruceaga *et al.* 2019)) and inhibit immunoglobulin E production (pseurotin (Ishikawa *et al.* 2009)). Microsynteny of the fumagillin BGC reveals high sequence conservation between *A. fumigatus* and *A. oerlinghausenensis*; however, sequence divergence was observed between *A. oerlinghausenensis* and *A. fischeri* (Fig. 5). Accordingly, fumagillin production was only observed in *A. fumigatus* and *A. oerlinghausenensis* and not in *A. fischeri*. Similarly, the pseurotin BGC is conserved between *A. fumigatus* and *A. oerlinghausenensis*. Rather than sequence divergence, no sequence similarity was observed in the region of the pseurotin cluster in *A. fischeri*, which may be due to an indel event. Accordingly, no pseurotin production was observed among *A. fischeri* strains. Despite sequence conservation between *A. fumigatus* and *A. oerlinghausenensis*, no evidence of pseurotin biosynthesis was observed in *A. oerlinghausenensis*, which suggests regulatory decoupling of the intertwined fumagillin/pseurotin BGC. Alternatively, the genes downstream of the *A. fumigatus* pseurotin BGC, which are absent from the *A. oerlinghausenensis* locus, may contribute to BGC production and could explain the lack of pseurotin production in *A. oerlinghausenensis*. Altogether, these results show a striking correlation between sequence divergence and the production (or absence) of secondary metabolites implicated in human disease among *A. fumigatus* and nonpathogenic closest relatives.

Discussion

Aspergillus fumigatus is a major fungal pathogen nested within a clade (known as section *Fumigati*) of at least 60 other species, the vast majority of which are nonpathogenic (Steenwyk *et al.* 2019; Rokas *et al.* 2020b). Currently, it is thought that the ability to cause human disease evolved multiple times among species in section *Fumigati* (Rokas *et al.* 2020b). Secondary

620 metabolites contribute to the success of the major human pathogen *A. fumigatus* in the host
621 environment (Raffa and Keller 2019) and can therefore be thought of as "cards" of virulence
622 (Casadevall 2007; Knowles *et al.* 2020). However, whether the closest relatives of *A. fumigatus*,
623 *A. oerlinghausenensis* and *A. fischeri*, both of which are nonpathogenic, biosynthesize secondary
624 metabolites implicated in the ability of *A. fumigatus* to cause human disease remained largely
625 unknown. By examining genomic and chemical variation between and within *A. fumigatus* and
626 its closest nonpathogenic relatives, we identified both conservation and divergence (including
627 within species heterogeneity) in BGCs and secondary metabolite profiles (Fig. 1-5, S3, S5-8;
628 Table 2, S1, S3). Examples of conserved BGCs and secondary metabolites include the major
629 virulence factor, gliotoxin (Fig. 4), as well as several others (Fig. 5, S7; Table 2, S1, S3);
630 examples of BGC and secondary metabolite heterogeneity or divergence include pseurotin,
631 fumagillin, and several others (Fig. 5; Table 2, S1, S3). Lastly, we found that the fumitremorgin
632 BGC, which biosynthesizes fumitremorgin in all three species, is also associated with
633 verruculogen biosynthesis in *A. fischeri* strains (Fig. 5).
634
635 One of the surprising findings of our study was that although *A. oerlinghausenensis* and *A.*
636 *fumigatus* are evolutionarily more closely related to each other than to *A. fischeri* (Fig. 1), *A.*
637 *oerlinghausenensis* and *A. fischeri* appear to be more similar to each other than to *A. fumigatus* in
638 BGC composition, gene family content, and secondary metabolite profiles. The power of
639 pathogen-nonpathogen comparative genomics is best utilized when examining closely related
640 species (Fedorova *et al.* 2008; Jackson *et al.* 2011; Moran *et al.* 2011; Mead *et al.* 2019a; Rokas
641 *et al.* 2020b). Genomes from additional strains from the closest known nonpathogenic relatives
642 of *A. fumigatus*, including from the closest species relative *A. oerlinghausenensis*, *A. fischeri*,

643 and other nonpathogenic species in section *Fumigati* will be key for understanding the evolution
644 of *A. fumigatus* pathogenicity.
645
646 Our finding that *A. oerlinghausenensis* and *A. fischeri* shares more gene families and BGCs with
647 each other than they do with *A. fumigatus* (Fig. 1C, 2, S3, S4, S8) suggests that the evolutionary
648 trajectory of the *A. fumigatus* ancestor was marked by gene loss. We hypothesize that there were
649 two rounds of gene family and BGC loss in the *A. fumigatus* stem lineage: (1) gene families and
650 BGCs were lost in the common ancestor of *A. fumigatus* and *A. oerlinghausenensis* and (2)
651 additional losses occurred in the *A. fumigatus* ancestor. In addition to losses, we note that 548
652 gene families and 16 BGCs are unique to *A. fumigatus*, which may have resulted from genetic
653 innovation (e.g., *de novo* gene formation) or unique gene family and BGC retention (Fig. 2, S8).
654 In line with the larger number of shared BGCs between *A. oerlinghausenensis* and *A. fischeri*, we
655 found their secondary metabolite profiles were also more similar (Fig. 3, S6). Notably, the
656 evolutionary rate of the internal branch leading to the *A. fumigatus* common ancestor is much
657 higher than those in the rest of the branches in our genome-scale phylogeny (Fig. S2B),
658 suggesting that the observed gene loss and gene gain / retention events specific to *A. fumigatus*
659 may be part of a wider set of evolutionary changes in the *A. fumigatus* genome. Analyses with a
660 greater number of strains and species will help further test the validity of this hypothesis. More
661 broadly, these results suggest that comparisons of the pathogen *A. fumigatus* against either the
662 non-pathogen *A. oerlinghausenensis* (this manuscript) or the non-pathogen *A. fischeri* ((Mead *et*
663 *al.* 2019a; Knowles *et al.* 2020) and this manuscript) will both be instructive in understanding the
664 evolution of *A. fumigatus* pathogenicity.
665

666 When studying *Aspergillus* pathogenicity, it is important to consider any genetic and phenotypic
667 heterogeneity between strains of a single species (Knox *et al.* 2016; Kowalski *et al.* 2016, 2019;
668 Keller 2017; Ries *et al.* 2019; Blachowicz *et al.* 2020; Bastos *et al.* 2020; Drott *et al.* 2020; dos
669 Santos *et al.* 2020; Steenwyk *et al.* 2020). Our finding of strain heterogeneity among gene
670 families, BGCs, and secondary metabolites in *A. fumigatus* and *A. fischeri* (Fig. 1-3, S3, S4, S6,
671 S8) suggests considerable strain-level diversity in each species. For example, we found
672 secondary metabolite profile strain heterogeneity was greater in *A. fumigatus* than *A. fischeri*
673 (Fig. S6B-E). These results suggest that strain-specific secondary metabolite profiles may play a
674 role in variation of pathogenicity among *A. fumigatus* strains. In support of this hypothesis,
675 differential secondary metabolite production has been associated with differences in virulence
676 among isolates of *A. fumigatus* (Blachowicz *et al.* 2020). More broadly, our finding supports the
677 hypothesis that strain-level diversity is an important parameter when studying pathogenicity
678 (Kowalski *et al.* 2016, 2019; Keller 2017; Ries *et al.* 2019; Blachowicz *et al.* 2020; Bastos *et al.*
679 2020; Drott *et al.* 2020; dos Santos *et al.* 2020; Steenwyk *et al.* 2020).
680
681 Secondary metabolites contribute to *A. fumigatus* virulence through diverse processes including
682 suppressing the human immune system and damaging tissues (Table 2). Interestingly, we found
683 that the nonpathogens *A. oerlinghausenensis* and *A. fischeri* produced several secondary
684 metabolites implicated in the ability of *A. fumigatus* human disease, such gliotoxin, trypacidin,
685 verruculogen, and others (Fig. 4, 5, S7; Table 2, S3). Importantly, our work positively identified
686 secondary metabolites for many structural classes implicated in a previous taxonomic study
687 (Samson *et al.* 2007). These results suggest that several of the secondary metabolism-associated
688 cards of virulence present in *A. fumigatus* are conserved in closely related nonpathogens

689 (summarized in Fig. 6) as well as in closely related pathogenic species, such as *A. novofumigatus*
690 (Kjærbølling *et al.* 2018). Interestingly, disrupting the ability of *A. fumigatus* to biosynthesize
691 gliotoxin attenuates but does not abolish virulence (Sugui *et al.* 2007; Dagenais and Keller 2009;
692 Keller 2017), whereas disruption of the ability of *A. fischeri* NRRL 181 to biosynthesize
693 secondary metabolites, including gliotoxin, does not appear to influence virulence (Knowles *et*
694 *al.* 2020). Our findings, together with previous studies, support the hypothesis that individual
695 secondary metabolites are "cards" of virulence in a larger "hand" that *A. fumigatus* possesses.
696

697 **Funding**

698 JLS and AR are supported by the Howard Hughes Medical Institute through the James H.
699 Gilliam Fellowships for Advanced Study program. AR has additional support from a Discovery
700 Grant from Vanderbilt University and the National Science Foundation ((DEB-1442113). GHG
701 is supported by the Brazilian funding agencies Fundacão de Amparo a Pesquisa do Estado de
702 São Paulo (FAPESP 2016/07870-9) and Conselho Nacional de Desenvolvimento Cientıfico e
703 Tecnologico (CNPq). NHO is supported by the National Cancer Institute (P01 CA125066). SLK
704 and CDR were supported in part by the National Institutes of Health via the National Center for
705 Complementary and Integrative Health (F31 AT010558) and the National Institute of General
706 Medical Sciences (T34 GM113860), respectively.
707

708 **Acknowledgements**

709 We thank the labs of Rokas, Oberlies, and Goldman for helpful discussion and support of this
710 work.

711 **Table 1. Species and strains used in the present study**

Genus and species	Strain	Environmental/Clinical	Genomic analysis	Secondary metabolite profiling	Reference
Aspergillus oerlinghausenensis	CBS 139183^{T}	Environmental	+	+	This study
Aspergillus fischeri	NRRL 4585	Environmental	+	+	This study
Aspergillus fischeri	NRRL 4161	Unknown	+	+	This study
Aspergillus fischeri	NRRL 181	Environmental	+	+	(Fedorova *et al.* 2008)
Aspergillus fischeri	IBT 3007	Environmental	+	-	(Zhao *et al.* 2019)
Aspergillus fischeri	IBT 3003	Environmental	+	-	(Zhao *et al.* 2019)
Aspergillus fumigatus	Af293	Clinical	+	+	(Nierman *et al.* 2005)
Aspergillus fumigatus	CEA10 / CEA17	Clinical	+	+	(Fedorova *et al.* 2008)
Aspergillus fumigatus	HMR AF 270	Clinical	+	-	BioSample: SAMN07177964
Aspergillus fumigatus	Z5	Environmental	+	-	(Miao *et al.* 2015)

712 **'+' and '-' indicate if BGCs and secondary metabolite profiling was conducted on a particular strain. More specifically '+'**
713 **indicates the strain was analyzed whereas '-' indicates that the strain was not analyzed.**

714 **Table 2. Select *A. fumigatus* secondary metabolites implicated in modulating host biology**

	Function	Reference(s)	Evidence of biosynthetic gene cluster / secondary metabolite						
			A. fumigatus			*A. oerlinghausenensis*	*A. fischeri*		
			Af293	CEA10	CEA17	CBS 139183[T]	NRRL 181	NRRL 4585	NRRL 4161
Gliotoxin	Inhibits host immune response	(Sugui *et al.* 2007)	+/+	+/+	+/+	+/+	+/+	+/+	+/+
Fumitremorgin	Inhibits the breast cancer resistance protein	(González-Lobato *et al.* 2010)	+/-	+/+	+/-	+/+	+/+	+/+	+/+
Verruculogen	Changes electrophysical properties of human nasal epithelial cells	(Khoufache *et al.* 2007)	+/-	+/+	+/-	+/+	+/+	+/+	+/+
Trypacidin	Damages lung cell tissues	(Gauthier *et al.* 2012)	+/+	+/+	+/-	+/+	+/-	-/-	-/-
Pseurotin	Inhibits immunoglobulin E	(Ishikawa *et al.* 2009)	+/+	+/+	+/+	+/+	-/-	-/-	-/-
Fumagillin	Inhibits neutrophil function	(Fallon *et al.* 2010, 2011)	+/+	+/+	+/+	+/+	-/-	-/-	-/-

715 **A list of select secondary metabolites implicated in human disease and their functional role are described here. All secondary**
716 **metabolites listed or analogs thereof were identified during secondary metabolite profiling. Plus (+) and minus (-) signs**
717 **indicate the presence or absence of the BGC and secondary metabolite, respectively. For example, +/+ indicates both BGC**
718 **presence and evidence of secondary metabolite production, whereas +/- indicates BGC presence but no evidence of secondary**
719 **metabolite production.**

720 **References**

721 Abad A., J. Victoria Fernández-Molina, J. Bikandi, A. Ramírez, J. Margareto, *et al.*, 2010 What
722 makes Aspergillus fumigatus a successful pathogen? Genes and molecules involved in
723 invasive aspergillosis. Rev. Iberoam. Micol. 27: 155–182.
724 https://doi.org/10.1016/j.riam.2010.10.003

725 Afiyatullov S. S., A. I. Kalinovskii, M. V. Pivkin, P. S. Dmitrenok, and T. A. Kuznetsova, 2005
726 Alkaloids from the Marine Isolate of the Fungus Aspergillus fumigatus. Chem. Nat. Compd.
727 41: 236–238. https://doi.org/10.1007/s10600-005-0122-y

728 Bankevich A., S. Nurk, D. Antipov, A. A. Gurevich, M. Dvorkin, *et al.*, 2012 SPAdes: A New
729 Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. J. Comput.
730 Biol. 19: 455–477. https://doi.org/10.1089/cmb.2012.0021

731 Bastos R. W., C. Valero, L. P. Silva, T. Schoen, M. Drott, *et al.*, 2020 Functional
732 Characterization of Clinical Isolates of the Opportunistic Fungal Pathogen Aspergillus
733 nidulans, (A. P. Mitchell, Ed.). mSphere 5. https://doi.org/10.1128/mSphere.00153-20

734 Benedict K., B. R. Jackson, T. Chiller, and K. D. Beer, 2019 Estimation of Direct Healthcare
735 Costs of Fungal Diseases in the United States. Clin. Infect. Dis. 68: 1791–1797.
736 https://doi.org/10.1093/cid/ciy776

737 Bignell E., T. C. Cairns, K. Throckmorton, W. C. Nierman, and N. P. Keller, 2016 Secondary
738 metabolite arsenal of an opportunistic pathogenic fungus. Philos. Trans. R. Soc. B Biol. Sci.
739 371: 20160023. https://doi.org/10.1098/rstb.2016.0023

740 Blachowicz A., N. Raffa, J. W. Bok, T. Choera, B. Knox, *et al.*, 2020 Contributions of Spore
741 Secondary Metabolites to UV-C Protection and Virulence Vary in Different Aspergillus
742 fumigatus Strains, (J. Andrew Alspaugh, Ed.). MBio 11.

743 https://doi.org/10.1128/mBio.03415-19
744 Bodinaku I., J. Shaffer, A. B. Connors, J. L. Steenwyk, M. N. Biango-Daniels, *et al.*, 2019 Rapid
745 Phenotypic and Metabolomic Domestication of Wild Penicillium Molds on Cheese, (J. W.
746 Taylor, Ed.). MBio 10. https://doi.org/10.1128/mBio.02445-19
747 Bolger A. M., M. Lohse, and B. Usadel, 2014 Trimmomatic: A flexible trimmer for Illumina
748 sequence data. Bioinformatics 30: 2114–2120.
749 https://doi.org/10.1093/bioinformatics/btu170
750 Bongomin F., S. Gago, R. Oladele, and D. Denning, 2017 Global and Multi-National Prevalence
751 of Fungal Diseases—Estimate Precision. J. Fungi 3: 57. https://doi.org/10.3390/jof3040057
752 Caesar L. K., O. M. Kvalheim, and N. B. Cech, 2018 Hierarchical cluster analysis of technical
753 replicates to identify interferents in untargeted mass spectrometry metabolomics. Anal.
754 Chim. Acta 1021: 69–77. https://doi.org/10.1016/j.aca.2018.03.013
755 Camacho C., G. Coulouris, V. Avagyan, N. Ma, J. Papadopoulos, *et al.*, 2009 BLAST+:
756 architecture and applications. BMC Bioinformatics 10: 421. https://doi.org/10.1186/1471-
757 2105-10-421
758 Campbell J., Q. Lin, G. D. Geske, and H. E. Blackwell, 2009 New and Unexpected Insights into
759 the Modulation of LuxR-Type Quorum Sensing by Cyclic Dipeptides. ACS Chem. Biol. 4:
760 1051–1059. https://doi.org/10.1021/cb900165y
761 Capella-Gutierrez S., J. M. Silla-Martinez, and T. Gabaldon, 2009 trimAl: a tool for automated
762 alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25: 1972–1973.
763 https://doi.org/10.1093/bioinformatics/btp348
764 Casadevall A., 2007 Determinants of virulence in the pathogenic fungi. Fungal Biol. Rev. 21:
765 130–132. https://doi.org/10.1016/j.fbr.2007.02.007

766 Cock P. J. A., T. Antao, J. T. Chang, B. A. Chapman, C. J. Cox, *et al.*, 2009 Biopython: freely
767 available Python tools for computational molecular biology and bioinformatics.
768 Bioinformatics 25: 1422–1423. https://doi.org/10.1093/bioinformatics/btp163
769 Dagenais T. R. T., and N. P. Keller, 2009 Pathogenesis of Aspergillus fumigatus in Invasive
770 Aspergillosis. Clin. Microbiol. Rev. 22: 447–465. https://doi.org/10.1128/CMR.00055-08
771 Dolan S. K., R. A. Owens, G. O'Keeffe, S. Hammel, D. A. Fitzpatrick, *et al.*, 2014 Regulation of
772 Nonribosomal Peptide Synthesis: bis-Thiomethylation Attenuates Gliotoxin Biosynthesis in
773 Aspergillus fumigatus. Chem. Biol. 21: 999–1012.
774 https://doi.org/10.1016/j.chembiol.2014.07.006
775 Dongen S. van, 2000 Graph clustering by flow simulation. Graph Stimul. by flow Clust. PhD
776 thesis: University of Utrecht. https://doi.org/10.1016/j.cosrev.2007.05.001
777 Drgona L., A. Khachatryan, J. Stephens, C. Charbonneau, M. Kantecki, *et al.*, 2014 Clinical and
778 economic burden of invasive fungal diseases in Europe: focus on pre-emptive and empirical
779 treatment of Aspergillus and Candida species. Eur. J. Clin. Microbiol. Infect. Dis. 33: 7–21.
780 https://doi.org/10.1007/s10096-013-1944-3
781 Drott M. T., R. W. Bastos, A. Rokas, L. N. A. Ries, T. Gabaldón, *et al.*, 2020 Diversity of
782 Secondary Metabolism in Aspergillus nidulans Clinical Isolates, (A. P. Mitchell, Ed.).
783 mSphere 5. https://doi.org/10.1128/mSphere.00156-20
784 El-Elimat T., M. Figueroa, B. M. Ehrmann, N. B. Cech, C. J. Pearce, *et al.*, 2013 High-
785 Resolution MS, MS/MS, and UV Database of Fungal Secondary Metabolites as a
786 Dereplication Protocol for Bioactive Natural Products. J. Nat. Prod. 76: 1709–1716.
787 https://doi.org/10.1021/np4004307
788 Emms D. M., and S. Kelly, 2019 OrthoFinder: phylogenetic orthology inference for comparative

789 genomics. Genome Biol. 20: 238. https://doi.org/10.1186/s13059-019-1832-y
790 Fallon J. P., E. P. Reeves, and K. Kavanagh, 2010 Inhibition of neutrophil function following
791 exposure to the Aspergillus fumigatus toxin fumagillin. J. Med. Microbiol. 59: 625–633.
792 https://doi.org/10.1099/jmm.0.018192-0
793 Fallon J. P., E. P. Reeves, and K. Kavanagh, 2011 The Aspergillus fumigatus toxin fumagillin
794 suppresses the immune response of Galleria mellonella larvae by inhibiting the action of
795 haemocytes. Microbiology 157: 1481–1488. https://doi.org/10.1099/mic.0.043786-0
796 Fedorova N. D., N. Khaldi, V. S. Joardar, R. Maiti, P. Amedeo, *et al.*, 2008 Genomic islands in
797 the pathogenic filamentous fungus Aspergillus fumigatus. PLoS Genet. 4.
798 https://doi.org/10.1371/journal.pgen.1000046
799 Gaudêncio S. P., and F. Pereira, 2015 Dereplication: racing to speed up the natural products
800 discovery process. Nat. Prod. Rep. 32: 779–810. https://doi.org/10.1039/C4NP00134F
801 Gauthier T., X. Wang, J. Sifuentes Dos Santos, A. Fysikopoulos, S. Tadrist, *et al.*, 2012
802 Trypacidin, a Spore-Borne Toxin from Aspergillus fumigatus, Is Cytotoxic to Lung Cells,
803 (S. G. Filler, Ed.). PLoS One 7: e29906. https://doi.org/10.1371/journal.pone.0029906
804 González-Lobato L., R. Real, J. G. Prieto, A. I. Álvarez, and G. Merino, 2010 Differential
805 inhibition of murine Bcrp1/Abcg2 and human BCRP/ABCG2 by the mycotoxin
806 fumitremorgin C. Eur. J. Pharmacol. 644: 41–48.
807 https://doi.org/10.1016/j.ejphar.2010.07.016
808 Grahl N., K. M. Shepardson, D. Chung, and R. A. Cramer, 2012 Hypoxia and Fungal
809 Pathogenesis: To Air or Not To Air? Eukaryot. Cell 11: 560–570.
810 https://doi.org/10.1128/EC.00031-12
811 Guruceaga X., G. Ezpeleta, E. Mayayo, M. Sueiro-Olivares, A. Abad-Diaz-De-Cerio, *et al.*, 2018

812 A possible role for fumagillin in cellular damage during host infection by Aspergillus
813 fumigatus. Virulence 9: 1548–1561. https://doi.org/10.1080/21505594.2018.1526528
814 Guruceaga X., U. Perez-Cuesta, A. Abad-Diaz de Cerio, O. Gonzalez, R. M. Alonso, *et al.*, 2019
815 Fumagillin, a Mycotoxin of Aspergillus fumigatus: Biosynthesis, Biological Activities,
816 Detection, and Applications. Toxins (Basel). 12: 7. https://doi.org/10.3390/toxins12010007
817 Halász J., B. Podányi, L. Vasvári-Debreczy, A. Szabó, F. Hajdú, *et al.*, 2000 Structure
818 Elucidation of Fumagillin-Related Natural Products. Tetrahedron 56: 10081–10085.
819 https://doi.org/10.1016/S0040-4020(00)00979-0
820 Hoang D. T., O. Chernomor, A. von Haeseler, B. Q. Minh, and L. S. Vinh, 2018 UFBoot2:
821 Improving the Ultrafast Bootstrap Approximation. Mol. Biol. Evol. 35: 518–522.
822 https://doi.org/10.1093/molbev/msx281
823 Holt C., and M. Yandell, 2011 MAKER2: an annotation pipeline and genome-database
824 management tool for second-generation genome projects. BMC Bioinformatics 12: 491.
825 https://doi.org/10.1186/1471-2105-12-491
826 Houbraken J., M. Weig, U. Groß, M. Meijer, and O. Bader, 2016 Aspergillus oerlinghausenensis
827 , a new mould species closely related to A. fumigatus, (S. Poeggeler, Ed.). FEMS
828 Microbiol. Lett. 363: fnv236. https://doi.org/10.1093/femsle/fnv236
829 Hubert J., J.-M. Nuzillard, and J.-H. Renault, 2017 Dereplication strategies in natural product
830 research: How many tools and methodologies behind the same concept? Phytochem. Rev.
831 16: 55–95. https://doi.org/10.1007/s11101-015-9448-7
832 Ishikawa M., T. Ninomiya, H. Akabane, N. Kushida, G. Tsujiuchi, *et al.*, 2009 Pseurotin A and
833 its analogues as inhibitors of immunoglobuline E production. Bioorg. Med. Chem. Lett. 19:
834 1457–1460. https://doi.org/10.1016/j.bmcl.2009.01.029

835 Ito T., and M. Masubuchi, 2014 Dereplication of microbial extracts and related analytical
836 technologies. J. Antibiot. (Tokyo). 67: 353–360. https://doi.org/10.1038/ja.2014.12
837 Jackson R. W., L. J. Johnson, S. R. Clarke, and D. L. Arnold, 2011 Bacterial pathogen evolution:
838 breaking news. Trends Genet. 27: 32–40. https://doi.org/10.1016/j.tig.2010.10.001
839 Kalyaanamoorthy S., B. Q. Minh, T. K. F. Wong, A. von Haeseler, and L. S. Jermiin, 2017
840 ModelFinder: fast model selection for accurate phylogenetic estimates. Nat. Methods 14:
841 587–589. https://doi.org/10.1038/nmeth.4285
842 Kamei K., and A. Watanabe, 2005 Aspergillus mycotoxins and their effect on the host. Med.
843 Mycol. 43: 95–99. https://doi.org/10.1080/13693780500051547
844 Kato N., H. Suzuki, H. Takagi, Y. Asami, H. Kakeya, *et al.*, 2009 Identification of Cytochrome
845 P450s Required for Fumitremorgin Biosynthesis in Aspergillus fumigatus. ChemBioChem
846 10: 920–928. https://doi.org/10.1002/cbic.200800787
847 Katoh K., and D. M. Standley, 2013 MAFFT Multiple Sequence Alignment Software Version 7:
848 Improvements in Performance and Usability. Mol. Biol. Evol. 30: 772–780.
849 https://doi.org/10.1093/molbev/mst010
850 Kautsar S. A., K. Blin, S. Shaw, J. C. Navarro-Muñoz, B. R. Terlouw, *et al.*, 2019 MIBiG 2.0: a
851 repository for biosynthetic gene clusters of known function. Nucleic Acids Res.
852 https://doi.org/10.1093/nar/gkz882
853 Keller N. P., 2017 Heterogeneity confounds establishment of “a” model microbial strain. MBio
854 8.
855 Keller N. P., 2019 Fungal secondary metabolism: regulation, function and drug discovery. Nat.
856 Rev. Microbiol. 17: 167–180. https://doi.org/10.1038/s41579-018-0121-1
857 Khoufache K., O. Puel, N. Loiseau, M. Delaforge, D. Rivollet, *et al.*, 2007 Verruculogen

858 associated with Aspergillus fumigatus hyphae and conidia modifies the electrophysiological
859 properties of human nasal epithelial cells. BMC Microbiol. 7: 5.
860 https://doi.org/10.1186/1471-2180-7-5
861 Kjærbølling I., T. C. Vesth, J. C. Frisvad, J. L. Nybo, S. Theobald, *et al.*, 2018 Linking
862 secondary metabolites to gene clusters through genome sequencing of six diverse
863 Aspergillus species. Proc. Natl. Acad. Sci. 115: E753–E761.
864 https://doi.org/10.1073/pnas.1715954115
865 Kjærbølling I., T. Vesth, J. C. Frisvad, J. L. Nybo, S. Theobald, *et al.*, 2020 A comparative
866 genomics study of 23 Aspergillus species from section Flavi. Nat. Commun. 11: 1106.
867 https://doi.org/10.1038/s41467-019-14051-y
868 Knowles S. L., N. Vu, D. A. Todd, H. A. Raja, A. Rokas, *et al.*, 2019 Orthogonal Method for
869 Double-Bond Placement via Ozone-Induced Dissociation Mass Spectrometry (OzID-MS).
870 J. Nat. Prod. 82: 3421–3431. https://doi.org/10.1021/acs.jnatprod.9b00787
871 Knowles S. L., M. E. Mead, L. P. Silva, H. A. Raja, J. L. Steenwyk, *et al.*, 2020 Gliotoxin, a
872 Known Virulence Factor in the Major Human Pathogen Aspergillus fumigatus , Is Also
873 Biosynthesized by Its Nonpathogenic Relative Aspergillus fischeri, (Y.-S. Bahn, Ed.). MBio
874 11. https://doi.org/10.1128/mBio.03361-19
875 Knox B. P., A. Blachowicz, J. M. Palmer, J. Romsdahl, A. Huttenlocher, *et al.*, 2016
876 Characterization of Aspergillus fumigatus Isolates from Air and Surfaces of the
877 International Space Station, (Y.-S. Bahn, Ed.). mSphere 1.
878 https://doi.org/10.1128/mSphere.00227-16
879 Korf I., 2004 Gene finding in novel genomes. BMC Bioinformatics 5: 59.
880 https://doi.org/10.1186/1471-2105-5-59

881 Kowalski C. H., S. R. Beattie, K. K. Fuller, E. A. McGurk, Y.-W. Tang, *et al.*, 2016
882 Heterogeneity among Isolates Reveals that Fitness in Low Oxygen Correlates with
883 Aspergillus fumigatus Virulence. MBio 7. https://doi.org/10.1128/mBio.01515-16
884 Kowalski C. H., J. D. Kerkaert, K.-W. Liu, M. C. Bond, R. Hartmann, *et al.*, 2019 Fungal
885 biofilm morphology impacts hypoxia fitness and disease progression. Nat. Microbiol. 4:
886 2430–2441. https://doi.org/10.1038/s41564-019-0558-7
887 Kvalheim O. M., H. Chan, I. F. F. Benzie, Y. Szeto, A. H. Tzang, *et al.*, 2011 Chromatographic
888 profiling and multivariate analysis for screening and quantifying the contributions from
889 individual components to the bioactive signature in natural products. Chemom. Intell. Lab.
890 Syst. 107: 98–105. https://doi.org/10.1016/j.chemolab.2011.02.002
891 Latgé J.-P., and G. Chamilos, 2019 Aspergillus fumigatus and Aspergillosis in 2019. Clin.
892 Microbiol. Rev. 33. https://doi.org/10.1128/CMR.00140-18
893 Li Z., C. Peng, Y. Shen, X. Miao, H. Zhang, *et al.*, 2008 l,l-Diketopiperazines from Alcaligenes
894 faecalis A72 associated with South China Sea sponge Stelletta tenuis. Biochem. Syst. Ecol.
895 36: 230–234. https://doi.org/10.1016/j.bse.2007.08.007
896 Li X.-J., Q. Zhang, A.-L. Zhang, and J.-M. Gao, 2012 Metabolites from Aspergillus fumigatus,
897 an endophytic fungus associated with Melia azedarach, and their antifungal, antifeedant,
898 and toxic activities. J. Agric. Food Chem. 60: 3424–31. https://doi.org/10.1021/jf300146n
899 Lind A. L., J. H. Wisecaver, T. D. Smith, X. Feng, A. M. Calvo, *et al.*, 2015 Examining the
900 evolution of the regulatory circuit controlling secondary metabolism and development in the
901 fungal genus Aspergillus. PLoS Genet. 11: e1005096.
902 https://doi.org/10.1371/journal.pgen.1005096
903 Lind A. L., J. H. Wisecaver, C. Lameiras, P. Wiemann, J. M. Palmer, *et al.*, 2017 Drivers of

904 genetic diversity in secondary metabolic gene clusters within a fungal species. PLoS Biol.
905 15. https://doi.org/10.1371/journal.pbio.2003583
906 Losada L., O. Ajayi, J. C. Frisvad, J. Yu, and W. C. Nierman, 2009 Effect of competition on the
907 production and activity of secondary metabolites in Aspergillus species. Med. Mycol. 47:
908 S88–S96. https://doi.org/10.1080/13693780802409542
909 Ma Y. ., Y. Li, J. . Liu, Y. . Song, and R. . Tan, 2004 Anti-Helicobacter pylori metabolites from
910 Rhizoctonia sp. Cy064, an endophytic fungus in Cynodon dactylon. Fitoterapia 75: 451–
911 456. https://doi.org/10.1016/j.fitote.2004.03.007
912 Mattern D. J., H. Schoeler, J. Weber, S. Novohradská, K. Kraibooj, *et al.*, 2015 Identification of
913 the antiphagocytic trypacidin gene cluster in the human-pathogenic fungus Aspergillus
914 fumigatus. Appl. Microbiol. Biotechnol. 99: 10151–10161. https://doi.org/10.1007/s00253-
915 015-6898-1
916 Mead M. E., S. L. Knowles, H. A. Raja, S. R. Beattie, C. H. Kowalski, *et al.*, 2019a
917 Characterizing the Pathogenic, Genomic, and Chemical Traits of Aspergillus fischeri , a
918 Close Relative of the Major Human Fungal Pathogen Aspergillus fumigatus, (A. P.
919 Mitchell, Ed.). mSphere 4. https://doi.org/10.1128/mSphere.00018-19
920 Mead M. E., H. A. Raja, J. L. Steenwyk, S. L. Knowles, N. H. Oberlies, *et al.*, 2019b Draft
921 Genome Sequence of the Griseofulvin-Producing Fungus Xylaria flabelliformis Strain
922 G536, (J. E. Stajich, Ed.). Microbiol. Resour. Announc. 8.
923 https://doi.org/10.1128/MRA.00890-19
924 Miao Y., D. Liu, G. Li, P. Li, Y. Xu, *et al.*, 2015 Genome-wide transcriptomic analysis of a
925 superior biomass-degrading strain of A. fumigatus revealed active lignocellulose-degrading
926 genes. BMC Genomics 16: 459. https://doi.org/10.1186/s12864-015-1658-2

927 Moran G. P., D. C. Coleman, and D. J. Sullivan, 2011 Comparative Genomics and the Evolution
928 of Pathogenicity in Human Pathogenic Fungi. Eukaryot. Cell 10: 34–42.
929 https://doi.org/10.1128/EC.00242-10
930 Navarro-Muñoz J. C., N. Selem-Mojica, M. W. Mullowney, S. A. Kautsar, J. H. Tryon, *et al.*,
931 2020 A computational framework to explore large-scale biosynthetic diversity. Nat. Chem.
932 Biol. 16: 60–68. https://doi.org/10.1038/s41589-019-0400-9
933 Nguyen L.-T., H. A. Schmidt, A. von Haeseler, and B. Q. Minh, 2015 IQ-TREE: A Fast and
934 Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. Mol.
935 Biol. Evol. 32: 268–274. https://doi.org/10.1093/molbev/msu300
936 Nierman W. C., A. Pain, M. J. Anderson, J. R. Wortman, H. S. Kim, *et al.*, 2005 Genomic
937 sequence of the pathogenic and allergenic filamentous fungus Aspergillus fumigatus.
938 Nature 438: 1151–1156. https://doi.org/10.1038/nature04332
939 Pluskal T., S. Castillo, A. Villar-Briones, and M. Orešič, 2010 MZmine 2: Modular framework
940 for processing, visualizing, and analyzing mass spectrometry-based molecular profile data.
941 BMC Bioinformatics 11: 395. https://doi.org/10.1186/1471-2105-11-395
942 Poukka H., P. Aarnisalo, H. Santti, O. A. Jänne, and J. J. Palvimo, 2000 Coregulator Small
943 Nuclear RING Finger Protein (SNURF) Enhances Sp1- and Steroid Receptor-mediated
944 Transcription by Different Mechanisms. J. Biol. Chem. 275: 571–579.
945 https://doi.org/10.1074/jbc.275.1.571
946 Raffa N., and N. P. Keller, 2019 A call to arms: Mustering secondary metabolites for success and
947 survival of an opportunistic pathogen, (D. C. Sheppard, Ed.). PLOS Pathog. 15: e1007606.
948 https://doi.org/10.1371/journal.ppat.1007606
949 Raftery A. E., and S. M. Lewis, 1995 The number of iterations, convergence diagnostics and

950 generic Metropolis algorithms. Pract. Markov Chain Monte Carlo 7: 763–773.
951 https://doi.org/10.1.1.41.6352
952 Ries L. N. A., J. L. Steenwyk, P. A. de Castro, P. B. A. de Lima, F. Almeida, *et al.*, 2019
953 Nutritional Heterogeneity Among Aspergillus fumigatus Strains Has Consequences for
954 Virulence in a Strain- and Host-Dependent Manner. Front. Microbiol. 10.
955 https://doi.org/10.3389/fmicb.2019.00854
956 Rokas A., B. L. Williams, N. King, and S. B. Carroll, 2003 Genome-scale approaches to
957 resolving incongruence in molecular phylogenies. Nature 425: 798–804.
958 https://doi.org/10.1038/nature02053
959 Rokas A., J. H. Wisecaver, and A. L. Lind, 2018 The birth, evolution and death of metabolic
960 gene clusters in fungi. Nat. Rev. Microbiol. https://doi.org/10.1038/s41579-018-0075-3
961 Rokas A., M. E. Mead, J. L. Steenwyk, H. A. Raja, and N. H. Oberlies, 2020a Biosynthetic gene
962 clusters and the evolution of fungal chemodiversity. Nat. Prod. Rep.
963 https://doi.org/10.1039/C9NP00045C
964 Rokas A., M. E. Mead, J. L. Steenwyk, N. H. Oberlies, and G. H. Goldman, 2020b Evolving
965 moldy murderers: Aspergillus section Fumigati as a model for studying the repeated
966 evolution of fungal pathogenicity, (D. C. Sheppard, Ed.). PLOS Pathog. 16: e1008315.
967 https://doi.org/10.1371/journal.ppat.1008315
968 Samson R. A., S. Hong, S. W. Peterson, J. C. Frisvad, and J. Varga, 2007 Polyphasic taxonomy
969 of Aspergillus section Fumigati and its teleomorph Neosartorya. Stud. Mycol. 59: 147–203.
970 https://doi.org/10.3114/sim.2007.59.14
971 Santos R. A. C. dos, J. L. Steenwyk, O. Rivero-Menendez, M. E. Mead, L. P. Silva, *et al.*, 2020
972 Genomic and Phenotypic Heterogeneity of Clinical Isolates of the Human Pathogens

973 Aspergillus fumigatus, Aspergillus lentulus, and Aspergillus fumigatiaffinis. Front. Genet.
974 11. https://doi.org/10.3389/fgene.2020.00459
975 Shwab E. K., J. W. Bok, M. Tribus, J. Galehr, S. Graessle, *et al.*, 2007 Histone Deacetylase
976 Activity Regulates Chemical Diversity in Aspergillus. Eukaryot. Cell 6: 1656–1664.
977 https://doi.org/10.1128/EC.00186-07
978 Spikes S., R. Xu, C. K. Nguyen, G. Chamilos, D. P. Kontoyiannis, *et al.*, 2008 Gliotoxin
979 Production in Aspergillus fumigatus Contributes to Host-Specific Differences in Virulence.
980 J. Infect. Dis. 197: 479–486. https://doi.org/10.1086/525044
981 Stanke M., and S. Waack, 2003 Gene prediction with a hidden Markov model and a new intron
982 submodel. Bioinformatics 19: ii215–ii225. https://doi.org/10.1093/bioinformatics/btg1080
983 Steenwyk J., and A. Rokas, 2017 Extensive Copy Number Variation in Fermentation-Related
984 Genes Among Saccharomyces cerevisiae Wine Strains. G3 Genes, Genomes, Genet. 7.
985 Steenwyk J. L., X.-X. Shen, A. L. Lind, G. H. Goldman, and A. Rokas, 2019 A Robust
986 Phylogenomic Time Tree for Biotechnologically and Medically Important Fungi in the
987 Genera Aspergillus and Penicillium, (J. P. Boyle, Ed.). MBio 10.
988 https://doi.org/10.1128/mBio.00925-19
989 Steenwyk J. L., A. L. Lind, L. N. A. Ries, T. F. dos Reis, L. P. Silva, *et al.*, 2020 Pathogenic
990 Allodiploid Hybrids of Aspergillus Fungi. Curr. Biol. 30: 2495-2507.e7.
991 https://doi.org/10.1016/j.cub.2020.04.071
992 Sugui J. A., J. Pardo, Y. C. Chang, K. A. Zarember, G. Nardone, *et al.*, 2007 Gliotoxin Is a
993 Virulence Factor of Aspergillus fumigatus : gliP Deletion Attenuates Virulence in Mice
994 Immunosuppressed with Hydrocortisone. Eukaryot. Cell 6: 1562–1569.
995 https://doi.org/10.1128/EC.00141-07

996 Tavaré S., 1986 Some probabilistic and statistical problems in the analysis of DNA sequences.
997 Lect. Math. life Sci. 17: 57–86.
998 Tekaia F., and J.-P. Latgé, 2005 Aspergillus fumigatus: saprophyte or pathogen? Curr. Opin.
999 Microbiol. 8: 385–392. https://doi.org/10.1016/j.mib.2005.06.017
1000 Vallabhaneni S., R. K. Mody, T. Walker, and T. Chiller, 2016 The Global Burden of Fungal
1001 Diseases. Infect. Dis. Clin. North Am. 30: 1–11. https://doi.org/10.1016/j.idc.2015.10.004
1002 Vesth T. C., J. L. Nybo, S. Theobald, J. C. Frisvad, T. O. Larsen, *et al.*, 2018 Investigation of
1003 inter- and intraspecies variation through genome sequencing of Aspergillus section Nigri.
1004 Nat. Genet. https://doi.org/10.1038/s41588-018-0246-1
1005 Vinet L., and A. Zhedanov, 2011 A 'missing' family of classical orthogonal polynomials. J.
1006 Phys. A Math. Theor. 44: 085201. https://doi.org/10.1088/1751-8113/44/8/085201
1007 Vries R. P. de, R. Riley, A. Wiebenga, G. Aguilar-Osorio, S. Amillis, *et al.*, 2017 Comparative
1008 genomics reveals high biological diversity and specific adaptations in the industrially and
1009 medically important fungal genus Aspergillus. Genome Biol. 18: 28.
1010 https://doi.org/10.1186/s13059-017-1151-0
1011 Wang F., Y. Fang, T. Zhu, M. Zhang, A. Lin, *et al.*, 2008 Seven new prenylated indole
1012 diketopiperazine alkaloids from holothurian-derived fungus Aspergillus fumigatus.
1013 Tetrahedron 64: 7986–7991. https://doi.org/10.1016/j.tet.2008.06.013
1014 Wang F.-Z., D.-H. Li, T.-J. Zhu, M. Zhang, and Q.-Q. Gu, 2011 Pseurotin A 1 and A 2 , two new
1015 1-oxa-7-azaspiro[4.4]non-2-ene-4,6-diones from the holothurian-derived fungus Aspergillus
1016 fumigatus WFZ-25. Can. J. Chem. 89: 72–76. https://doi.org/10.1139/V10-157
1017 Waterhouse R. M., F. Tegenfeldt, J. Li, E. M. Zdobnov, and E. V. Kriventseva, 2013 OrthoDB: a
1018 hierarchical catalog of animal, fungal and bacterial orthologs. Nucleic Acids Res. 41:

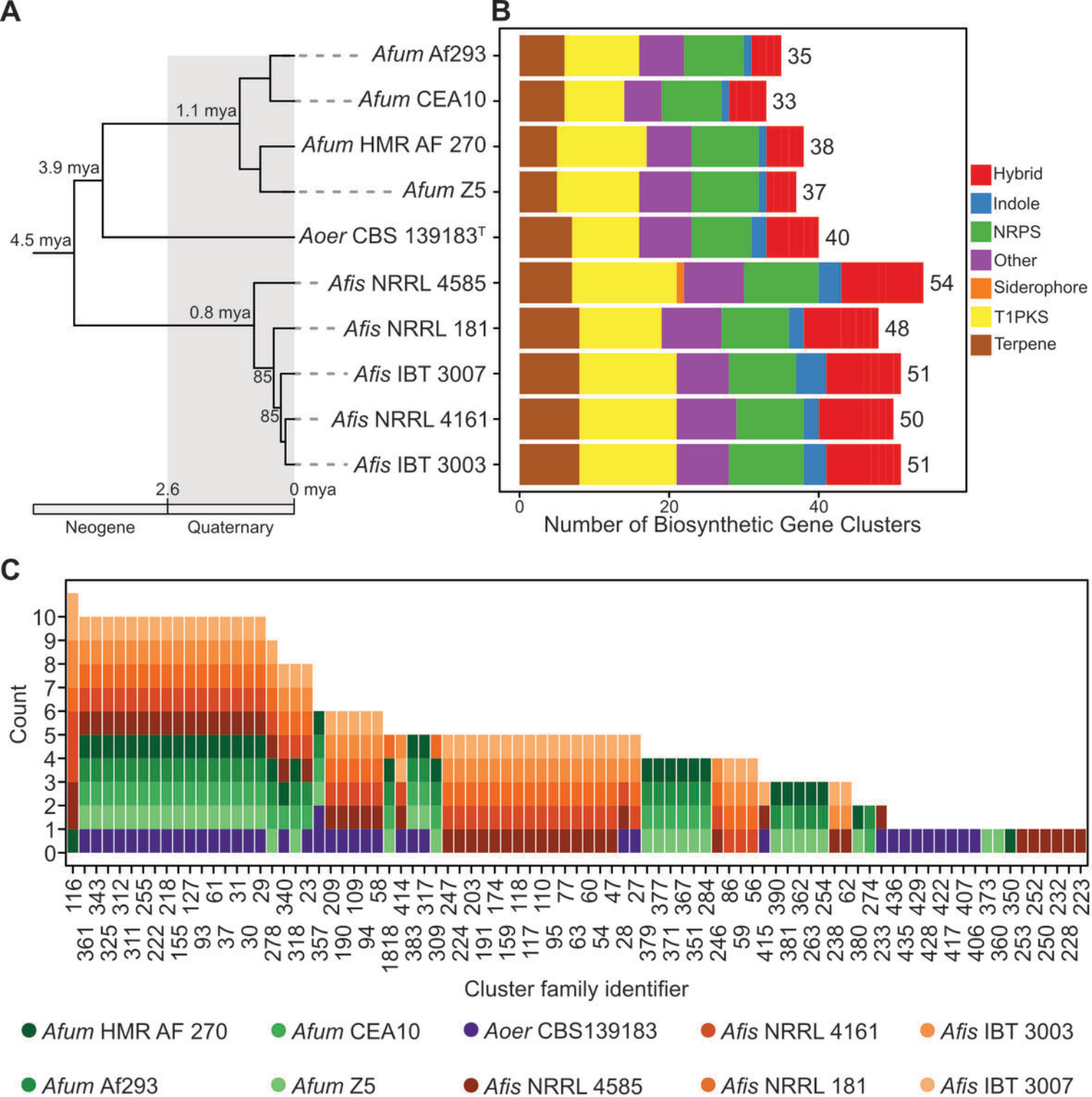
1019 D358–D365. https://doi.org/10.1093/nar/gks1116
1020 Waterhouse R. M., M. Seppey, F. A. Simão, M. Manni, P. Ioannidis, *et al.*, 2018 BUSCO
1021 Applications from Quality Assessments to Gene Prediction and Phylogenomics. Mol. Biol.
1022 Evol. 35: 543–548. https://doi.org/10.1093/molbev/msx319
1023 Weber T., K. Blin, S. Duddela, D. Krug, H. U. Kim, *et al.*, 2015 antiSMASH 3.0—a
1024 comprehensive resource for the genome mining of biosynthetic gene clusters. Nucleic Acids
1025 Res. 43: W237–W243. https://doi.org/10.1093/nar/gkv437
1026 Wiemann P., C.-J. Guo, J. M. Palmer, R. Sekonyela, C. C. C. Wang, *et al.*, 2013 Prototype of an
1027 intertwined secondary-metabolite supercluster. Proc. Natl. Acad. Sci. 110: 17065–17070.
1028 https://doi.org/10.1073/pnas.1313258110
1029 Wiemann P., B. E. Lechner, J. A. Baccile, T. A. Velk, W.-B. Yin, *et al.*, 2014 Perturbations in
1030 small molecule synthesis uncovers an iron-responsive secondary metabolite network in
1031 Aspergillus fumigatus. Front. Microbiol. 5. https://doi.org/10.3389/fmicb.2014.00530
1032 Yamada A., T. Kataoka, and K. Nagai, 2000 The fungal metabolite gliotoxin:
1033 immunosuppressive activity on CTL-mediated cytotoxicity. Immunol. Lett. 71: 27–32.
1034 https://doi.org/10.1016/s0165-2478(99)00155-8
1035 Yandell M., and D. Ence, 2012 A beginner's guide to eukaryotic genome annotation. Nat. Rev.
1036 Genet. 13: 329–42. https://doi.org/10.1038/nrg3174
1037 Yang Z., 1994 Maximum likelihood phylogenetic estimation from DNA sequences with variable
1038 rates over sites: Approximate methods. J. Mol. Evol. 39: 306–314.
1039 https://doi.org/10.1007/BF00160154
1040 Yang Z., 1996 Among-site rate variation and its impact on phylogenetic analyses. Trends Ecol.
1041 Evol. 11: 367–372. https://doi.org/10.1016/0169-5347(96)10041-0

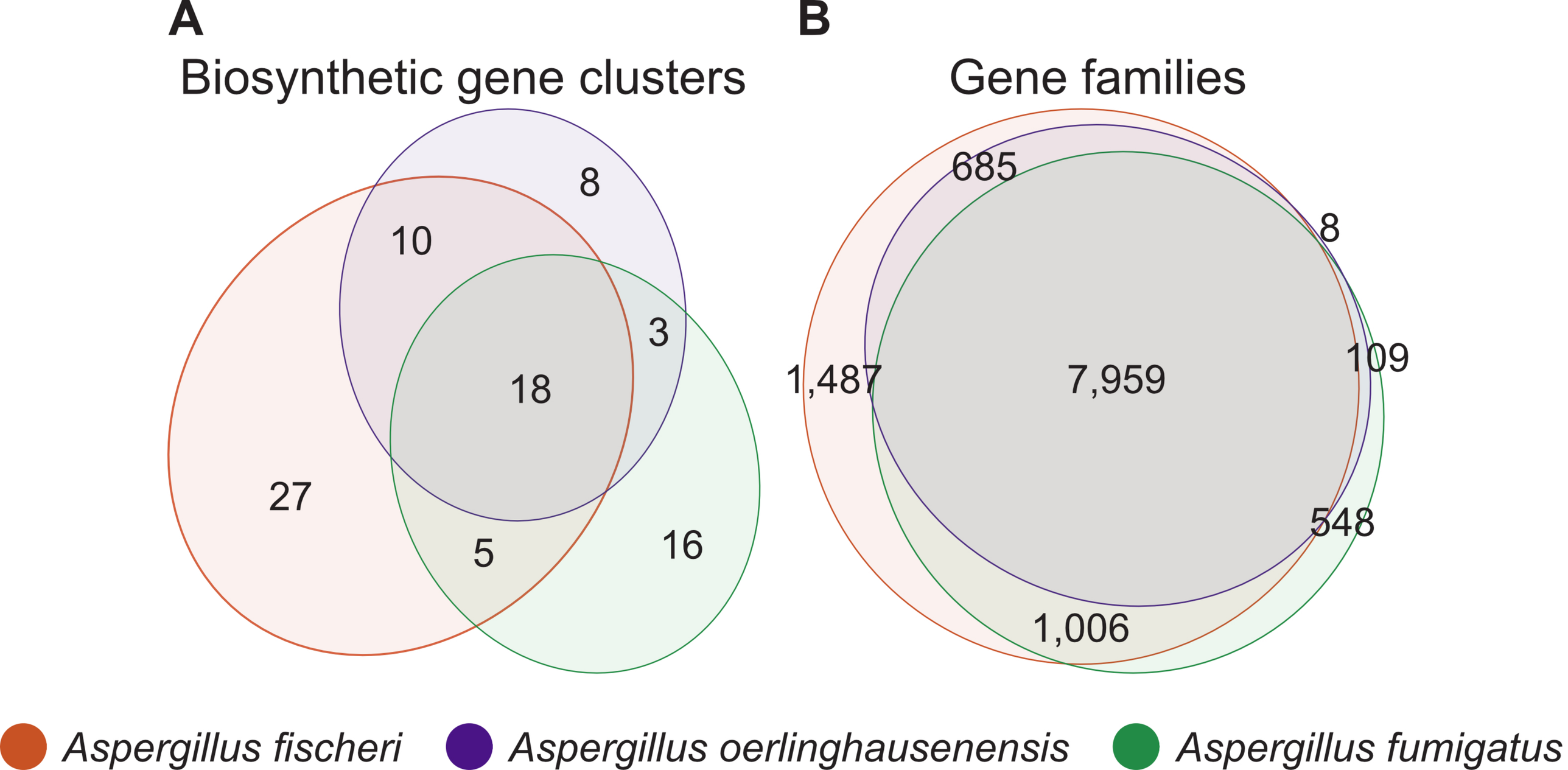
1042 Yang Z., 2007 PAML 4: Phylogenetic Analysis by Maximum Likelihood. Mol. Biol. Evol. 24:
1043 1586–1591. https://doi.org/10.1093/molbev/msm088
1044 Yin W.-B., J. A. Baccile, J. W. Bok, Y. Chen, N. P. Keller, *et al.*, 2013 A Nonribosomal Peptide
1045 Synthetase-Derived Iron(III) Complex from the Pathogenic Fungus Aspergillus fumigatus.
1046 J. Am. Chem. Soc. 135: 2064–2067. https://doi.org/10.1021/ja311145n
1047 Zhao J., Y. Mou, T. Shan, Y. Li, L. Zhou, *et al.*, 2010 Antimicrobial Metabolites from the
1048 Endophytic Fungus Pichia guilliermondii Isolated from Paris polyphylla var. yunnanensis.
1049 Molecules 15: 7961–7970. https://doi.org/10.3390/molecules15117961
1050 Zhao S., J.-P. Latgé, and J. G. Gibbons, 2019 Genome Sequences of Two Strains of the Food
1051 Spoilage Mold Aspergillus fischeri, (A. Rokas, Ed.). Microbiol. Resour. Announc. 8.
1052 https://doi.org/10.1128/MRA.01328-19
1053
1054

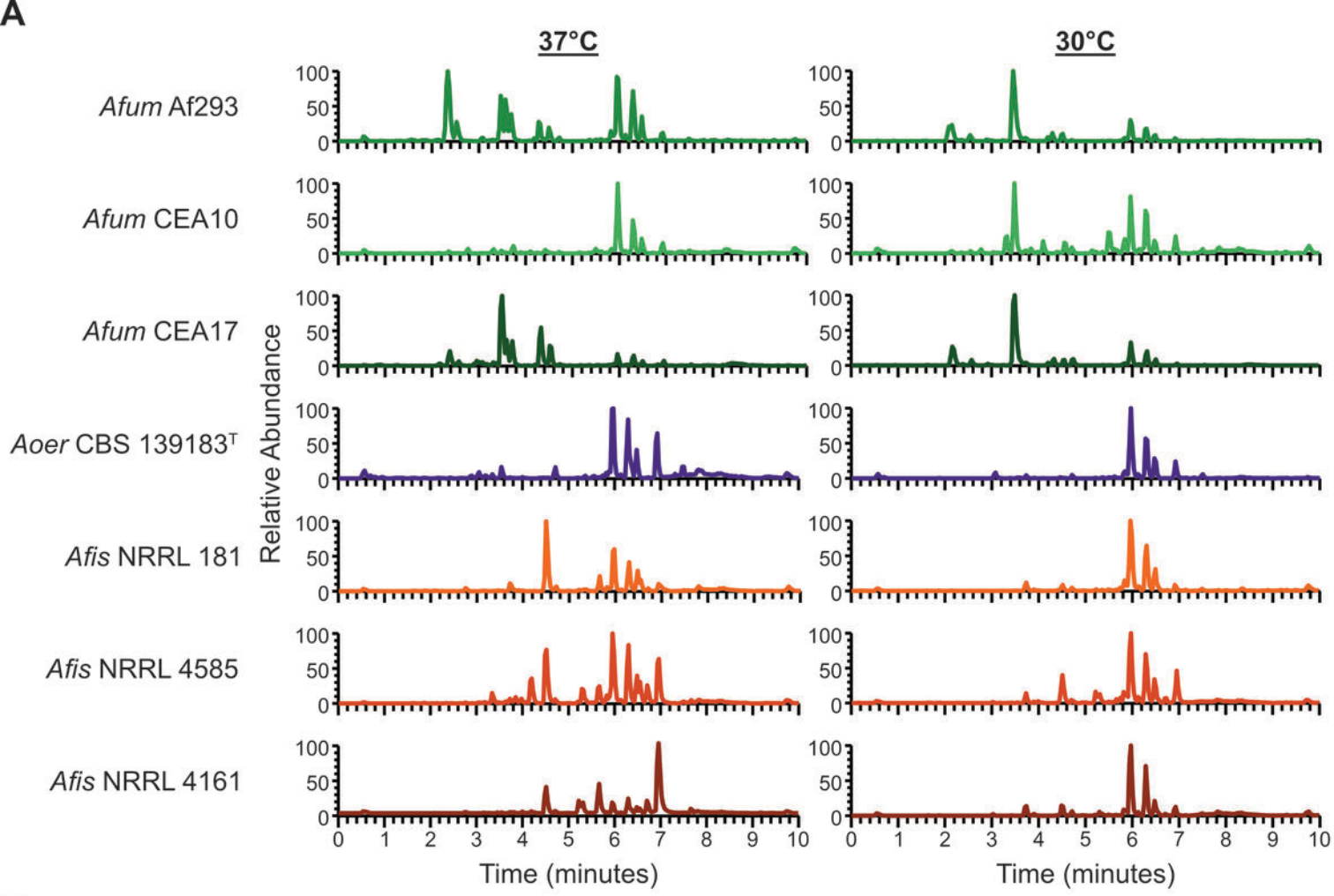
1055 **Figure 1. Diverse genetic repertoire of biosynthetic gene clusters and extensive presence**
1056 **and absence polymorphisms between and within species.** (A) Genome-scale phylogenomic
1057 analysis confirms *A. oerlinghausenensis* is the closest relative to *A. fumigatus*. Relaxed
1058 molecular clock analyses suggest *A. fumigatus, A. oerlinghausenensis,* and *A. fischeri* diverged
1059 from one another during the Neogene geologic period. Bipartition support is depicted for
1060 internodes that did not have full support. (B) *A. fumigatus* harbors the lowest number of BGCs
1061 compared to its two closest relatives. (C) Network-based clustering of BGCs into cluster families
1062 reveal extensive cluster presence and absence polymorphisms between species and strains.
1063 Cluster family identifiers are depicted on the x-axis; the number of strains represented in a
1064 cluster family are shown on the y-axis; the colors refer to a single strain from each species.
1065 Genus and species names are written using the following abbreviations: *Afum*: *A. fumigatus*;
1066 *Aoer*: *A. oerlinghausenensis*; *Afis*: *A. fischeri*. Classes of BGCs are written using the following
1067 abbreviations: NRPS: nonribosomal peptide synthetase; T1PKS: type I polyketide synthase;
1068 Hybrid: a combination of multiple BGC classes.
1069
1070 **Figure 2. *Aspergillus oerlinghausenensis* shares more gene families and BGCs with *A.***
1071 ***fischeri* than *A. fumigatus*.** (A) Euler diagram showing species-level shared BGCs. (B) Euler
1072 diagram showing species-level shared gene families. In both diagrams, *A. oerlinghausenensis*
1073 shares more gene families or BGCs with *A. fischeri* than *A. fumigatus* despite a closer
1074 evolutionary relationship. The Euler diagrams show the results for the species-level comparisons,
1075 which may be influenced by the unequal numbers of strains used for the three species; strain-
1076 level comparisons of BGCs and gene families can be found in Figures 1C and S4, respectively.
1077

1078 **Figure 3. *A. oerlinghausenensis* and *A. fischeri* have more similar secondary metabolite**
1079 **profiles than *A. fumigatus*.** (A) UPLC-MS chromatograms of secondary metabolite profiles of
1080 *A. fumigatus* and its closest relatives, *A. oerlinghausenensis* and *A. fischeri* at 37°C and 30°C
1081 (left and right, respectively). (B) Hierarchical clustering of chromatograms (1,920 total features)
1082 reveals *A. oerlinghausenensis* clusters with *A. fischeri* and not its closest relative, *A. fumigatus* at
1083 37°C and 30°C (left and right, respectively).
1084
1085 **Figure 4. Conservation in the gliotoxin BGC correlates with conserved production of**
1086 **gliotoxin analogs in *A. fumigatus* and nonpathogenic close relatives.** Microsynteny analysis
1087 reveals a high degree of conservation in the BGC encoding gliotoxin across all isolates. The
1088 known gliotoxin gene cluster boundary is indicated above the *A. fumigatus* Af293 BGC. Black
1089 and white squares correspond to evidence or absence of evidence of secondary metabolite
1090 production, respectively. Genes are drawn as arrows with orientation indicated by the direction
1091 of the arrow. Gene function is indicated by gene color. Grey boxes between gene clusters
1092 indicate BLAST-based similarity of nucleotide sequences defined as being at least 100 bp in
1093 length, share at least 30% sequence similarity, and have an expectation value threshold of 0.01.
1094 Genus and species names are written using the following abbreviations: *Afum*: *A. fumigatus*;
1095 *Aoer*: *A. oerlinghausenensis*; *Afis*: *A. fischeri*. Below each genus and species abbreviation is the
1096 cluster family each BGC belongs to and their cluster number.
1097
1098 **Figure 5. Conservation and divergence in the locus encoding the fumitremorgin and**
1099 **intertwined fumagillin/pseurotin BGCs.** Microsynteny analysis reveals conservation in the
1100 fumitremorgin BGC across all isolates. Interestingly, only *A. fischeri* strains synthesize

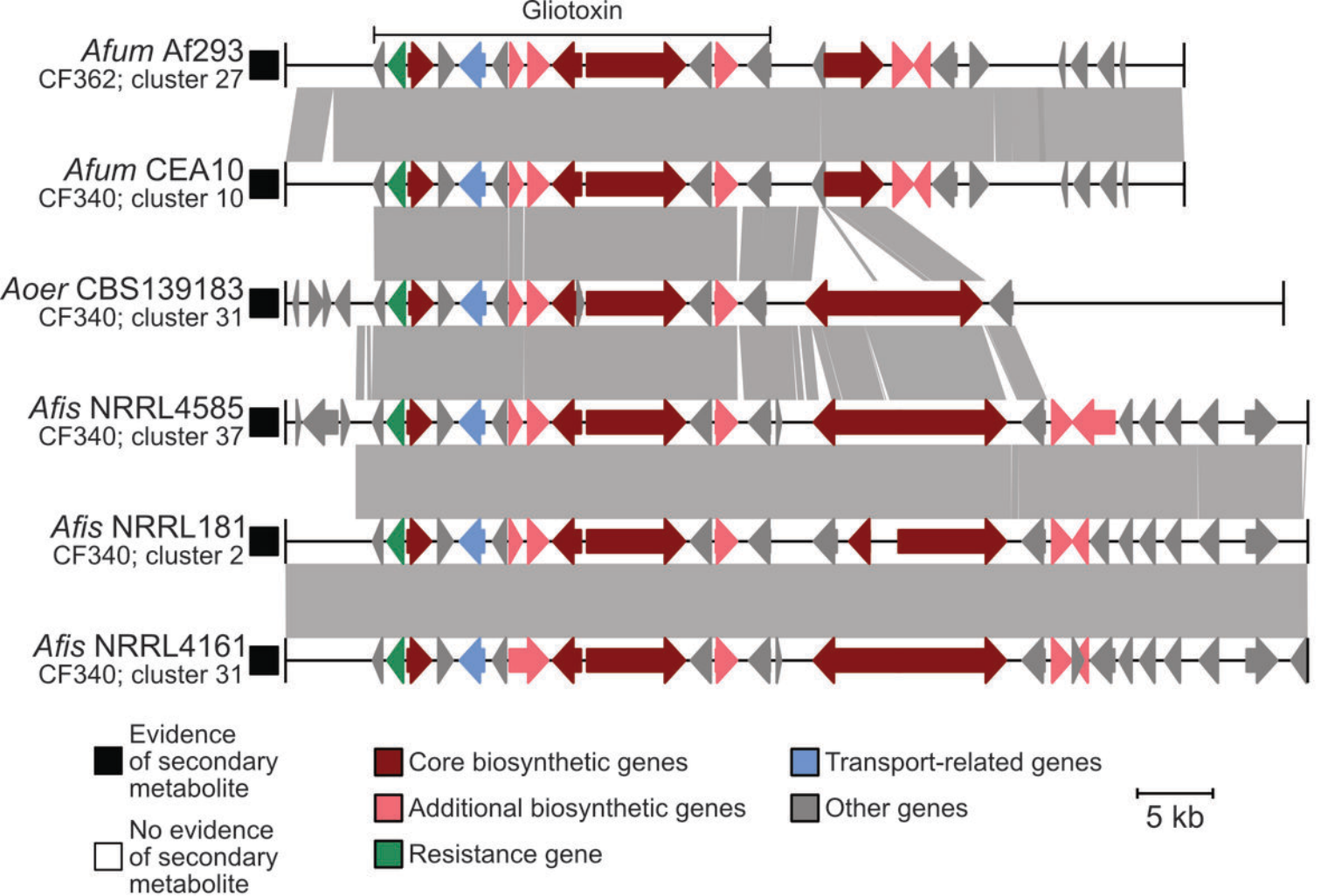
1101 verruculogen, a secondary metabolite also biosynthesized by the fumitremorgin BGC. In
1102 contrast, the intertwined fumagillin/pseurotin BGCs are conserved between *A. fumigatus* and *A.*
1103 *oerlinghausenensis* but divergent in *A. fischeri*. BGC conservation and divergence is associated
1104 with the presence and absence of a secondary metabolite, respectively. The same convention
1105 used in Fig. 4 is used to depict evidence of a secondary metabolite, represent genes and broad
1106 gene function, BGC sequence similarity, genus and species abbreviations, and BGC cluster
1107 families and cluster numbers.
1108
1109 **Figure 6. Secondary metabolism-associated "cards" of virulence among *A. fumigatus* and**
1110 **close relatives.** Secondary metabolites contribute to the "hand of cards"' that enable *A.*
1111 *fumigatus* to cause disease. Here, we show that the nonpathogenic closest relatives of *A.*
1112 *fumigatus* possess a subset of the *A. fumigatus* secondary metabolism-associated cards of
1113 virulence. We hypothesize that the unique combination of cards of *A. fumigatus* contributes to its
1114 pathogenicity and that the cards in *A. oerlinghausenensis* and *A. fischeri* (perhaps in combination
1115 with other non-secondary-metabolism-associated cards, such as thermotolerance) are insufficient
1116 to cause disease. Pathogenic and nonpathogenic species are shown in red and black, respectively.
1117 Cartoons of *Aspergillus* species were obtained from WikiMedia Commons (source: M.
1118 Piepenbring) and modified in accordance with the Creative Commons Attribution-Share Alike
1119 3.0 Unported license (https://creativecommons.org/licenses/by-sa/3.0/deed.en).

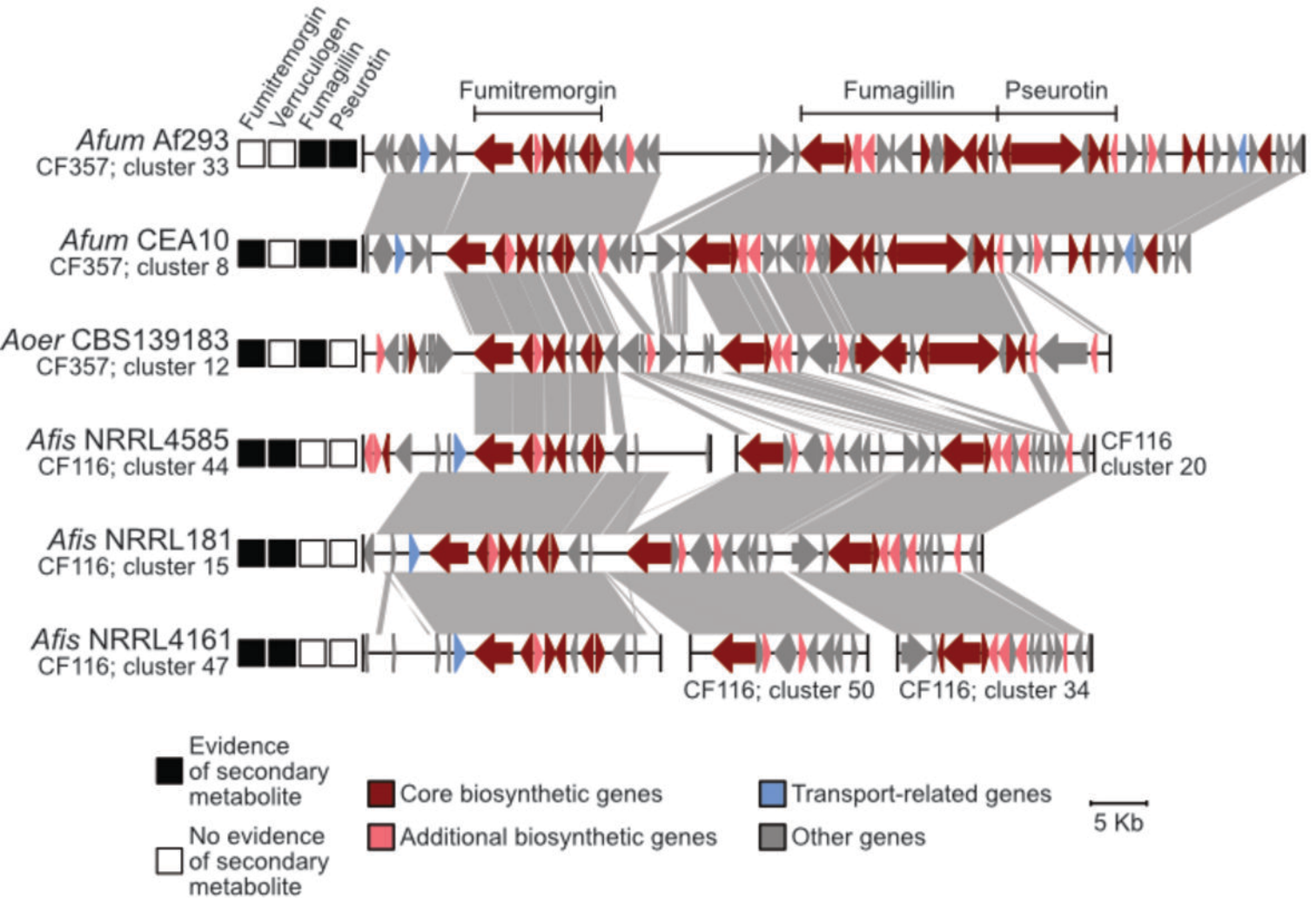

A
B
C
Afum Af293
Afum CEA10
Afum HMR AF 270
Afum Z5
Aoer CBS 139183T
Afis NRRL 4585
Afis NRRL 181
Afis IBT 3007
Afis NRRL 4161
Afis IBT 3003
1.1 mya
3.9 mya
4.5 mya
0.8 mya
85
85
2.6
0 mya
Neogene
Quaternary
35
33
38
37
40
54
48
51
50
51
0
20
40
Number of Biosynthetic Gene Clusters
Hybrid
Indole
NRPS
Other
Siderophore
T1PKS
Terpene
Count
Cluster family identifier
Afum HMR AF 270
Afum CEA10
Aoer CBS139183
Afis NRRL 4161
Afis IBT 3003
Afum Af293
Afum Z5
Afis NRRL 4585
Afis NRRL 181
Afis IBT 3007


A
Biosynthetic gene clusters
8
10
3
18
27
5
16
B
Gene families
685
8
1,487
7,959
109
548
1,006
Aspergillus fischeri
Aspergillus oerlinghausenensis
Aspergillus fumigatus


A
37°C
30°C
Afum Af293
Afum CEA10
Afum CEA17
Aoer CBS 139183T
Afis NRRL 181
Afis NRRL 4585
Afis NRRL 4161
Relative Abundance
100
50
0
0 1 2 3 4 5 6 7 8 9 10
Time (minutes)


B
37°C
30°C
Aoer CBS 139183T
Afis NRRL 181
Afis NRRL 4585
Afis NRRL 4161
Afum CEA17
Afum Af293
Afum CEA10


Gliotoxin
Afum Af293
CF362; cluster 27
Afum CEA10
CF340; cluster 10
Aoer CBS139183
CF340; cluster 31
Afis NRRL4585
CF340; cluster 37
Afis NRRL181
CF340; cluster 2
Afis NRRL4161
CF340; cluster 31
Evidence of secondary metabolite
No evidence of secondary metabolite
Core biosynthetic genes
Additional biosynthetic genes
Resistance gene
Transport-related genes
Other genes
5 kb


Fumitremorgin
Verruculogen
Fumagillin
Pseurotin
Fumitremorgin
Fumagillin
Pseurotin
Afum Af293
CF357; cluster 33
Afum CEA10
CF357; cluster 8
Aoer CBS139183
CF357; cluster 12
Afis NRRL4585
CF116; cluster 44
CF116
cluster 20
Afis NRRL181
CF116; cluster 15
Afis NRRL4161
CF116; cluster 47
CF116; cluster 50
CF116; cluster 34
Evidence of secondary metabolite
No evidence of secondary metabolite
Core biosynthetic genes
Additional biosynthetic genes
Transport-related genes
Other genes
5 Kb

Secondary metabolism-associated "cards" of virulence