

<b>Manuscript Number:</b>	GIGA-D-22-00073	
<b>Full Title:</b>	Clonality, inbreeding, and hybridization in two extremotolerant black yeasts	
<b>Article Type:</b>	Research	
<b>Funding Information:</b>	Javna Agencija za Raziskovalno Dejavnost RS (MRIC UL I0-0022)	Not applicable
	Javna Agencija za Raziskovalno Dejavnost RS (P4-0432)	Assoc Prof Cene Gostinčar
	Javna Agencija za Raziskovalno Dejavnost RS (P1-0198)	Not applicable
	Javna Agencija za Raziskovalno Dejavnost RS (J4-2549)	Assoc Prof Cene Gostinčar
<b>Abstract:</b>	<p>The great diversity of lifestyles and survival strategies observed in fungi is reflected in the many ways in which they reproduce and recombine. Although truly asexual fungi are rare, population genomic data support the clonality of two extremotolerant black yeasts from Dothideomycetes: <i>Hortaea werneckii</i> and <i>Aureobasidium melanogenum</i>. Thus, the discovery of a number of diploid strains of these species could not be explained as the product of conventional sexual reproduction.</p> <p>Genome sequencing revealed that the ratio of diploid to haploid strains in both <i>H. werneckii</i> and <i>A. melanogenum</i> is approximately 2:1. Linkage disequilibrium between pairs of polymorphic loci and a high degree of concordance between the phylogenies of different genomic regions confirmed that both species are clonal. Heterozygosity of diploid strains is high, with several hybridizing genome pairs reaching the intergenomic distances typically seen between different fungal species. The origin of diploid strains collected worldwide can be traced to a handful of hybridization events that produced diploids, which were stable over long periods of time and distributed over large geographic areas.</p> <p>Our results, based on the genomes of over 100 strains of two black yeasts, show that although they are asexual, they occasionally form stable and highly heterozygous diploid hybrids. The mechanism of these apparently rare hybridization events, which are not followed by meiosis or haploidisation, remains unknown. Both extremotolerant yeasts, <i>H. werneckii</i> and even more so <i>A. melanogenum</i>, a close relative of the intensely recombining and biotechnologically relevant <i>Aureobasidium pullulans</i>, provide an attractive model for studying the role of clonality and ploidy in extremotolerant fungi.</p>	
<b>Corresponding Author:</b>	Cene Gostinčar, Ph.D. University of Ljubljana Biotechnical faculty: Univerza v Ljubljani Biotehniška fakulteta Ljubljana, SLOVENIA	
<b>Corresponding Author Secondary Information:</b>		
<b>Corresponding Author's Institution:</b>	University of Ljubljana Biotechnical faculty: Univerza v Ljubljani Biotehniška fakulteta	
<b>Corresponding Author's Secondary Institution:</b>		
<b>First Author:</b>	Cene Gostinčar, Ph.D.	
<b>First Author Secondary Information:</b>		
<b>Order of Authors:</b>	Cene Gostinčar, Ph.D.	
	Xiaohuan Sun	
	Anja Černoša	

	Chao Fang
	Nina Gunde-Cimerman
	Zewei Song
<b>Order of Authors Secondary Information:</b>	
<b>Additional Information:</b>	
<b>Question</b>	<b>Response</b>
Are you submitting this manuscript to a special series or article collection?	No
<p><b>Experimental design and statistics</b></p> <p>Full details of the experimental design and statistical methods used should be given in the Methods section, as detailed in our <a href="#">Minimum Standards Reporting Checklist</a>. Information essential to interpreting the data presented should be made available in the figure legends.</p> <p>Have you included all the information requested in your manuscript?</p>	Yes
<p><b>Resources</b></p> <p>A description of all resources used, including antibodies, cell lines, animals and software tools, with enough information to allow them to be uniquely identified, should be included in the Methods section. Authors are strongly encouraged to cite <a href="#">Research Resource Identifiers</a> (RRIDs) for antibodies, model organisms and tools, where possible.</p> <p>Have you included the information requested as detailed in our <a href="#">Minimum Standards Reporting Checklist</a>?</p>	Yes
<p><b>Availability of data and materials</b></p> <p>All datasets and code on which the conclusions of the paper rely must be either included in your submission or deposited in <a href="#">publicly available repositories</a> (where available and ethically</p>	Yes

appropriate), referencing such data using a unique identifier in the references and in the “Availability of Data and Materials” section of your manuscript.

Have you have met the above requirement as detailed in our [Minimum Standards Reporting Checklist?](#)

1 **Clonality, inbreeding, and hybridization in two extremotolerant black yeasts**

2

3 Cene Gostinčar <sup>a,b,¶,\*</sup>, Xiaohuan Sun <sup>c,¶</sup>, Anja Černoša <sup>a</sup>, Chao Fang <sup>c</sup>, Nina Gunde-  
4 Cimerman <sup>a,+</sup>, Zewei Song <sup>c,+</sup>

5

6 <sup>a</sup> Department of Biology, Biotechnical Faculty, University of Ljubljana, 1000 Ljubljana,  
7 Slovenia

8 <sup>b</sup> Lars Bolund Institute of Regenerative Medicine, BGI-Qingdao, Qingdao 266555,  
9 China

10 <sup>c</sup> BGI-Shenzhen, Beishan Industrial Zone, Shenzhen 518083, China

11

12 ¶¶ *These authors contributed equally to this work as first authors.*

13 + *These authors contributed equally to this work as leading authors.*

14

15 \* **Corresponding author: Cene Gostinčar**

16 Department of Biology, Biotechnical Faculty

17 University of Ljubljana

18 Jamnikarjeva 101

19 SI-1000 Ljubljana

20 Slovenia

21 Tel.: +386-1-320 3392

22 Fax: +386-1-2573390

23 E-mail: [cene.gostincar@bf.uni-lj.si](mailto:cene.gostincar@bf.uni-lj.si), [cgostincar@gmail.com](mailto:cgostincar@gmail.com)

24

25 Xiaohuan Sun: [sunxiaohuan@genomics.cn](mailto:sunxiaohuan@genomics.cn)

26 Anja Černoša: [anja.cernosa@bf.uni-lj.si](mailto:anja.cernosa@bf.uni-lj.si)

27 Chao Fang: [fangchao@genomics.cn](mailto:fangchao@genomics.cn)

28 Nina Gunde-Cimerman: [nina.Gunde-Cimerman@bf.uni-lj.si](mailto:nina.Gunde-Cimerman@bf.uni-lj.si)

29 Zewei Song: [songzewei@genomics.cn](mailto:songzewei@genomics.cn)

30

31 **Running Title**

32 *Hybridization in extremotolerant yeasts*

33

34

35 **Abstract**

36 The great diversity of lifestyles and survival strategies observed in fungi is reflected in  
37 the many ways in which they reproduce and recombine. Although truly asexual fungi  
38 are rare, population genomic data support the clonality of two extremotolerant black  
39 yeasts from Dothideomycetes: *Hortaea werneckii* and *Aureobasidium melanogenum*.  
40 Thus, the discovery of a number of diploid strains of these species could not be  
41 explained as the product of conventional sexual reproduction.

42 Genome sequencing revealed that the ratio of diploid to haploid strains in both *H.*  
43 *werneckii* and *A. melanogenum* is approximately 2:1. Linkage disequilibrium between  
44 pairs of polymorphic loci and a high degree of concordance between the phylogenies  
45 of different genomic regions confirmed that both species are clonal. Heterozygosity of  
46 diploid strains is high, with several hybridizing genome pairs reaching the intergenomic  
47 distances typically seen between different fungal species. The origin of diploid strains  
48 collected worldwide can be traced to a handful of hybridization events that produced  
49 diploids, which were stable over long periods of time and distributed over large  
50 geographic areas.

51 Our results, based on the genomes of over 100 strains of two black yeasts, show that  
52 although they are asexual, they occasionally form stable and highly heterozygous  
53 diploid hybrids. The mechanism of these apparently rare hybridization events, which  
54 are not followed by meiosis or haploidisation, remains unknown. Both extremotolerant  
55 yeasts, *H. werneckii* and even more so *A. melanogenum*, a close relative of the  
56 intensely recombining and biotechnologically relevant *Aureobasidium pullulans*,  
57 provide an attractive model for studying the role of clonality and ploidy in  
58 extremotolerant fungi.

59

60 Keywords: population genomics, halotolerance, extremotolerance, halophilic fungus,  
61 *Hortaea werneckii*, *Aureobasidium melanogenum*, hybridization

62

## 63 **Introduction**

64 No single reproductive strategy is optimal for all species and all conditions in which  
65 they live. This results in the coexistence of a wide variety of ways in which organisms  
66 reproduce and recombine their genetic material. Among the most diverse are fungi,  
67 which exhibit a wide range of strategies, from strictly clonal species [1] to species with  
68 thousands of mating types [2]. Sexual, parasexual, and clonal reproduction are broad  
69 categories that encompass a variety of different phenomena, processes, and  
70 mechanisms, some of which are typical of larger groups (e.g., the dikaryon of  
71 basidiomycetes), while others differ even among closely related species [3].

72 Traditionally, up to one-fifth of fungi were thought to be asexual [4]. Subsequent genetic  
73 and genomic analyses found at least a rudimentary mating-type locus in nearly every  
74 species studied. Population genetics showed that most species previously thought to  
75 be asexual are actually recombining [5–7]. However, in most species, asexual  
76 reproduction dominates over occasional recombination. To account for the observation  
77 that some apparently clonal species can nevertheless recombine at levels low enough  
78 not to break the pattern of population structure typical of clonal species, Tibayrenc and  
79 Ayala [5] introduced the concept of “restricted recombination”. Nevertheless, some  
80 species appear to be strictly clonal even by highly sensitive measures of recombination  
81 used by population genomics, such as linkage disequilibrium [1].

82 Several reasons for the pervasive clonality in fungi have been proposed, such as hybrid  
83 incompatibility or limited opportunities to meet strains of the opposite mating type [6].  
84 Severe bottlenecks (e.g., following the introduction of a small number of strains to a  
85 new site) can lead to a skewed balance of mating types. Strains of pathogens with  
86 host-to-host transmission may encounter other strains of the species very rarely [6,8].  
87 Similar may be true for fungi with poor dispersal abilities that are restricted to rare and  
88 isolated ecological islands, such as certain extreme environments [1,9]. However, even  
89 without these constraints, sexually meiotic reproductive events are often majorly  
90 outweighed by asexual mitotic events, as shown in *Saccharomyces paradoxus* [10].

91 The absence of recombination carries the risk of accumulation of deleterious  
92 mutations, a process known as Muller’s ratchet. Sexual reproduction remedies this and  
93 efficiently generates diversity, which is a substrate for selection and adaptation to novel  
94 conditions. But asexual reproduction has its own advantages. For example, it can

95 conserve energy by eliminating the need to maintain the mating system and form  
96 sexual structures [11,12]. It also allows the organism to faithfully reproduce successful  
97 genomic configurations and thus avoid the recombination load, a loss of fitness due to  
98 a break up of advantageous combinations of interacting alleles [12–14]. This may be  
99 particularly beneficial in specialists, such as those inhabiting extreme environments  
100 [14,15].

101 In addition to sexual reproduction, fungi can employ another tool to recombine their  
102 genomes: parasexuality [16]. Two cells can fuse to combine their genetic material,  
103 providing the opportunity for mitotic recombination. Haploid parents thus produce  
104 diploid offspring, but this ploidy change is generally considered unstable. However, it  
105 does not revert to the original ploidy of the parental strains through the tightly controlled  
106 and high-fidelity process of meiosis. Instead, chromosomes are lost randomly through  
107 a series of aneuploid generations. The importance of parasexuality outside of  
108 laboratory settings has been questioned [17], but in at least some cases the process  
109 appears to drive adaptation and facilitate survival, e.g., in *Aspergillus fumigatus* in the  
110 lungs of patients with cystic fibrosis [18].

111 Changes in ploidy itself may be a form of adaptation, either through parasexuality or  
112 other processes such as abnormal cell division [19]. Both polyploidy and aneuploidy  
113 can be a response to adverse or novel environmental conditions [20–22]. They  
114 influence fitness through changes in cell size and shape, changes in the transcriptome  
115 (by altering gene dosage) and in the rate of adaptation, but also by providing new  
116 options for repairing DNA damage and temporarily masking deleterious mutations [19–  
117 21,23]. A growing body of evidence shows that variation in ploidy is a widespread  
118 transient adaptation of fungi to novel conditions (reviewed by Naranjo–Ortiz and  
119 Gabaldón [21]). Aneuploidies are more common and more easily reversed than  
120 tandem gene duplications, which are an alternative way to increase gene dosage.  
121 During cultivation under optimal conditions, such altered ploidies tend to revert to a  
122 baseline ploidy of the species, often without clear increases in fitness [20,24].

123 Sexual and parasexual recombination can lead to recombinant lineages and  
124 interspecies hybrids [25], another process increasingly recognised as an important  
125 generator of fungal diversity, including in industrial and clinical settings (reviewed by  
126 Naranjo–Ortiz and Gabaldón [21]). With the increasing accessibility of genome  
127 sequencing, research on this topic is rapidly expanding to non-model species. Some

128 hybrids exhibit higher fitness than their parental strains [26], making hybridization an  
129 important driver of adaptation to novel environments [21,27]. Divergent hybrid  
130 genomes can be stabilised by chromosomal aberrations [28,29] and the outcome of  
131 hybridization is often similar to parasexuality. Hybrids of *Cryptococcus neoformans* and  
132 *Cryptococcus gattii*, for example, rapidly lose chromosomes and rearrange them  
133 [28,29].

134 Indications of hybridization were also reported by Gostinčar et al. [1] in the extremely  
135 halotolerant black yeast *Hortaea werneckii* (Capnodiales, Dothideomycetidae,  
136 Dothideomycetes, Pezizomycotina, Ascomycota), a globally distributed species  
137 specialised for survival in saline environments and able to grow in nearly salt-saturated  
138 solutions [30,31]. Whole genome sequencing of twelve strains indicated that the  
139 species is clonal, but also that a majority of the strains are highly heterozygous diploids.  
140 These diploids appeared to be stable enough to spread over considerable distances,  
141 with little evidence of haploidisation [1,32]. This explained the ploidy of the reference  
142 genome, which was originally interpreted as the result of endoreduplication [33,34].  
143 Subsequent genome sequencing of two additional *H. werneckii* strains provided  
144 additional support for the hybridization hypothesis [35]. However, the total number of  
145 sequenced *H. werneckii* genomes remained low, limiting the power of the analyses  
146 and the interpretation of the results.

147 A similar pattern of haploid and diploid strains coexisting within an apparently clonal  
148 species has since been discovered in *Aureobasidium melanogenum* (Dothideales,  
149 Dothideomycetidae, Dothideomycetes, Pezizomycotina, Ascomycota), another black  
150 yeast only distantly related to *H. werneckii* [36]. While *A. melanogenum* tolerates less  
151 extreme conditions than *H. werneckii*, it is tolerant of a wider range of types of stress  
152 and occurs in a variety of environments, from hypersaline waters to various indoor  
153 habitats (reviewed by Černoša et al. [36]).

154 The role of hybridization and ploidy changes are among the overlooked dimensions of  
155 fungal genetics [21]. Here we analyse 66 genomes of *H. werneckii* and 49 genomes of  
156 *A. melanogenum* to provide insight into the reproductive strategy of these two  
157 extremotolerant fungi, characterised by coexistence of haploid and highly  
158 heterozygous diploid strains that are stable over large geographic and temporal  
159 distances.



160

## 161 **Results**

162 Whole genomes of 54 strains of the extremely halotolerant black yeast *H. werneckii*  
163 were sequenced. Combined with previously sequenced strains [1,34], this resulted in  
164 a data set of 66 whole-genome sequences (Table 1, Supplemental Table S1). A  
165 majority of strains (26) were isolated from brine, evaporation-concentrated seawater  
166 during salt extraction; followed by strains isolated from bittern (7), a saturated,  
167 magnesium-rich solution that remains after precipitation of halite during salt extraction.  
168 Nine strains were isolated from a seacoast cave in Atacama, where some of the strains  
169 grew on spider webs along with the alga *Dunaliella atacamensis* [37]. Twelve strains  
170 were isolated from marine habitats and four were clinical isolates. All genomes of *A.*  
171 *melanogenum* (Table 2) were sequenced and described in a previous study [36]. The  
172 largest number of strains (19) were isolated from bathroom and kitchen surfaces  
173 (including from kitchen appliances), followed by 16 strains from tap water or springs of  
174 tap water. In the case of both *H. werneckii* and *A. melanogenum*, the majority of strains  
175 were isolated in Slovenia.

176 Based on previous studies [1,33,34], the haploid genomes of both *H. werneckii* and *A.*  
177 *melanogenum* are approximately 25 Mbp in size. Comparing the sizes of genome  
178 assembly and the number of predicted genes in each genome, 20 (30%) *H. werneckii*  
179 genomes were recognised as haploid, 45 (68%) as diploid, and one genome as  
180 tetraploid. This was similar to the *A. melanogenum* genomes where 16 (33%) genomes  
181 were recognised as haploid, 30 (61%) as diploid, and the ploidy of three genomes (2,  
182 18, 38) was unclear [36].

183 The assembly size, number of predicted genes and other genomic characteristics of  
184 *H. werneckii* were fairly consistent within groups of specific ploidy (Table 3). The  
185 average genome assembly size was 26.52 Mbp ( $\pm 1.47$  SD) for haploid and 49.30 Mbp  
186 ( $\pm 1.74$  SD) for diploid genomes. The average number of predicted genes was 9519  
187 ( $\pm 665$  SD) for haploids and 20417 ( $\pm 1709$  SD) for diploids. As expected, the quality of  
188 assembly was much lower in diploid strains, as evidenced by a higher number of  
189 contigs and a smaller N50 value, presumably due to regions of high similarity between  
190 the two haploid subgenomes, a long-standing challenge in sequencing *H. werneckii*  
191 genomes [34]. Nevertheless, the assembly and annotation of all strains were of

192 reasonable quality, with only about 3.19% ( $\pm 0.30$  SD) Benchmarking Universal Single-  
193 Copy Orthologs missing completely in the haploid genomes and 6.09% ( $\pm 4.24$  SD) in  
194 the diploid genomes.

195 Single-nucleotide polymorphism (SNP) analysis was performed on all strains except  
196 *H. werneckii* strain 36 due to its tetraploid genome. The average density of single  
197 nucleotide polymorphisms (SNPs) in haploid strains of *H. werneckii* was high: 4.04%  
198 ( $\pm 1.11$  SD) (Table 3). For diploids, the SNP density was 3.44% ( $\pm 1.12$  SD), of which  
199 71% of the loci were heterozygous. In *A. melanogenum* the average SNP density was  
200 4.41% ( $\pm 1.87$  SD) in haploids and 3.79% ( $\pm 0.21$  SD) in diploids, with 44% of the latter  
201 heterozygous. Based on the SNP data, the genomes of both *H. werneckii* and *A.*  
202 *melanogenum* were clustered into 5 clusters in principal component analysis, with the  
203 first two principal components explaining 57.2% of the SNP diversity of *H. werneckii*  
204 and 59.3% of *A. melanogenum* (Fig. 1). SNP-based phylogenetic analyses of both  
205 species revealed considerable reticulation (Fig. 1). The largest cluster of strains  
206 identified by both network analysis and PCA contained 18 strains in *H. werneckii* and  
207 20 strains in *A. melanogenum*.

208 The squared correlation coefficient ( $r^2$ ) was calculated for all pairs of biallelic SNP loci  
209 present in at least two genomes analysed and within 10 kbp of each other. Plotting  $r^2$   
210 as a function of the distance between pairs of loci showed very little decay of linkage  
211 disequilibrium in either species. Linkage disequilibrium remained high above half of the  
212 maximum value even for alleles that were 10 kbp apart (Fig. 2). Such strong linkage  
213 between loci can be explained by a lack of recombination that would break the linkage,  
214 confirming previous reports that *H. werneckii* and *A. melanogenum* are strictly clonal  
215 [1,36].

216 The phylogenies of the fifty longest alignable genomic regions were also consistent  
217 with the presumed lack of recombination within *H. werneckii* and *A. melanogenum*.  
218 The phylogenetic trees showed a high degree of concordance (Fig. 3), meeting the  
219 “strong phylogenetic signal” criterion for clonality [5]. Sequences representing different  
220 haploid subgenomes of diploid strains were positioned in different parts of the  
221 phylogeny, corresponding to the high heterozygosity of the strains. An extreme case  
222 of this was the tetraploid *H. werneckii* strain 36, which was positioned in four different  
223 parts of the phylogeny. When all 50 multi-labelled trees for each species were

224 collapsed into a consensus supernetwork (Fig. 3), the result was similar to the  
225 phylogenetic network estimated from SNP data (Fig. 1).

226 The topology of the consensus phylogenies and supernetworks was best explained by  
227 a number of hybridization events in each species: 9 or 10 events in the case of *H.*  
228 *werneckii* (with an additional event leading to tetraploid strain 36) and 7 events in the  
229 case of *A. melanogenum* (Fig. 4). Several phylogenetic lineages resulting from these  
230 events appeared to be relatively widespread – more than one representative strain  
231 was found for most lineages, often in different habitats and geographic locations.  
232 However, lineage composition was skewed in favour of specific localities or habitats.  
233 This was confirmed by Fisher's Exact Test, which found significant differences  
234 between groups in both the isolation habitat and the geographic location of origin for  
235 both species ("*H. werneckii* – habitats":  $p < 0.01$ ; all other:  $p < 0.001$ ). For example, for  
236 both species, the two largest groups (groups 1 in Fig. 4) contained isolates from  
237 Europe, with only one exception. In terms of habitats, *H. werneckii* group 1 was isolated  
238 mainly from hypersaline habitats and group 9 from seawater; groups 4 and 5 were  
239 found mainly in a cave on a desert coast. The largest group of haploid strains was also  
240 found mainly in hypersaline habitats. Clinical isolates of *H. werneckii* belonged to  
241 different phylogenetic lineages. The tetraploid *H. werneckii* strain 36 was isolated from  
242 the deep sea (Figs. 3, 4) and most likely arose by hybridization between diploid hybrids  
243 of groups 1 and 9. In the case of *A. melanogenum*, most strains originated in Europe.  
244 Some groups showed preference for particular habitats: *A. melanogenum* group 1 was  
245 mostly isolated from tap water and its sources, while the majority of isolates from  
246 household surfaces were classified into other groups.

247 Aneuploidy in the genomes of *H. werneckii* and *A. melanogenum* was investigated by  
248 searching for large genome segments with different sequencing coverage than the rest  
249 of the genome. Evidence of aneuploidy was found in 23 genomes of *H. werneckii* (35%)  
250 and 8 genomes of *A. melanogenum* (16%) (Fig. 5, Supplemental Figs. S1, S2). The  
251 majority of these genomes were diploid: 18 (78%) in the case of *H. werneckii* and 8  
252 (100%) in the case of *A. melanogenum*. Some parts of the genome were aneuploid in  
253 several strains, with most aneuploid parts representing an increase in ploidy rather  
254 than a decrease. The aneuploid strains included three of four clinical *H. werneckii*  
255 isolates and the only clinical isolate of *A. melanogenum* in the study. In some diploid  
256 genomes loss of heterozygosity was observed over large regions (Supplemental Fig.

257 S3). Some of these could be explained by aneuploidy (loss of one copy of a  
258 chromosome or part of chromosome), while others appeared to be copy-neutral,  
259 possibly caused by mitotic recombination.

260 A putative mating-type locus was found in the majority of the genomes, although in  
261 many cases it was poorly assembled, especially in diploid genomes and even more so  
262 in the tetraploid genome of *H. werneckii*, precluding a conclusive analysis in all strains.  
263 However, where entire genes MAT1-1 and MAT1-2 were assembled, they showed a  
264 substantial diversity, clustering into distinct phylogenetic groups (Supplemental Figs.  
265 S4, S5). In case of *H. werneckii* 'blastx' searches against the non-redundant GenBank  
266 protein database showed that one large phylogenetic group of each locus was highly  
267 similar to homologues from other fungi ("true" MAT1-1 and MAT1-2), while the other  
268 groups only produced matches with these "true" *H. werneckii* homologues, but not with  
269 homologues from other fungi, especially in case of MAT1-2, a result of an intense  
270 diversification. In the case of *A. melanogenum*, all putative MAT1-1 and MAT1-2 could  
271 be matched to homologues from other fungi. Phylogenetic groups of MAT1-1 and  
272 MAT1-2 generally corresponded to hybrid and haploid groups of strains, but with  
273 numerous exceptions (Supplemental Figs. S4, S5). For example, in case of *H.*  
274 *werneckii* hybrid groups generally contained two similar copies of MAT1-1, but groups  
275 1a, 3 and 4 contained very different copies and group 9 consistently lacked one copy  
276 of MAT1-1 altogether (possibly due to poor assembly of the locus). In contrast, two  
277 different copies of MAT1-2 were found in group 1, but two similar copies in group 9.  
278 Three distinct copies of the mating-type locus were found in strain 10. In *A.*  
279 *melanogenum* hybrid group 1 contained only a single well-assembled homologue of  
280 MAT1-1 and two homologues of MAT1-2, but otherwise the diversity of the mating-type  
281 locus in this species was generally lower and mostly corresponded to groups of hybrids  
282 and haploid strains.

283 The search for proteins of the Pfam families HET and Het-C typical of heterokaryon  
284 incompatibility proteins, identified a large number of such proteins in the predicted  
285 proteomes of *A. melanogenum* (on average 4.87 HET and 3.74 Het-C proteins per  
286 strain) and even more in *H. werneckii* (on average 27.82 HET and 3.5 Het-C proteins  
287 per strain) (Supplemental Table S2). Phylogenetic analysis of these proteins showed  
288 that they formed several clusters, some of which contained representatives only from  
289 specific groups of hybrid strains (as identified in Fig. 4). For example, in *H. werneckii*,

290 the hybrid genomes of groups 9 and 10 (and in one case genome 23) were the only  
291 ones to contain HET genes belonging to phylogenetic clusters 6 and 24 (Supplemental  
292 Table S2). HET proteins from cluster 12 were found only in hybrid groups 1-3 and their  
293 closely related haploid strains (and tetraploid strain 36). In *A. melanogenum*, hybrid  
294 groups 5 and 6 were the only ones to contain HET proteins belonging to a small  
295 phylogenetic cluster 5 (Supplemental Table S2) and similarly hybrid group 2 was the  
296 only one to contain HET proteins from phylogenetic group 6.

297

## 298 **Discussion**

299 Genome sequencing of 66 strains of the black yeast *Hortaea werneckii* and 49 strains  
300 of the black yeast *Aureobasidium melanogenum* revealed some unexpected  
301 similarities between these species, which belong to different orders of the subclass  
302 Dothideomycetidae. Approximately one-third of the sequenced strains of each species  
303 were haploid and approximately two-thirds were diploid. Principal component analysis  
304 of single-nucleotide polymorphisms identified several clusters of strains in each  
305 species. In both cases, the first two principal components explained nearly 60% of the  
306 observed diversity – much more than, for example, in the homogenous and  
307 recombining species *Aureobasidium pullulans*, where the first two principal  
308 components together explained less than 15% of the diversity [38]. The clustering of  
309 strains was consistent with the previous reports that both *H. werneckii* and *A.*  
310 *melanogenum* are strictly clonal [1,36]. Despite the presence of a mating-type locus in  
311 the reference genomes of both species [1,39], the clonality of the species was  
312 confirmed here by a high degree of concordance between the phylogenetic histories  
313 of different genomic regions and by a lack of decay in linkage disequilibrium, an  
314 established measure of recombination often expressed as the distance over which  
315 linkage disequilibrium falls to half its maximum value [6].

316 The existence of highly heterozygous hybrids, first observed in *H. werneckii* [1], is  
317 confirmed here on a much larger genomic data set of *H. werneckii* and also *A.*  
318 *melanogenum*. The mechanism of hybridization is unknown and could range from  
319 vegetative hyphal fusion between different haploid strains to plasmogamy and  
320 karyogamy of gametes. Regardless of the mechanism, at least some diploid hybrids  
321 appear to be stable over long periods of time, allowing them to disperse over long

322 distances and constitute a large proportion of the species in some habitats. Hybrid  
323 strains have previously been reported in many other fungal species, including  
324 *Saccharomyces cerevisiae* [40], *Candida tropicalis* [41], and *Cryptococcus*  
325 *neoformans* [42], but none of these species are strictly clonal. However, diploid hybrids  
326 that cannot reproduce sexually have also been reported in some species [43]. The  
327 majority of *Candida orthopsilosis* are hybrids with 5% divergence between their haploid  
328 genomes at the nucleotide level, arising from at least four independent hybridization  
329 events [44]. *Candida metapsilosis* is also a species originating from hybridization, with  
330 a similar divergence between haploid genomes [45]. It is possible that *H. werneckii* and  
331 *A. melanogenum* follow a similar reproductive strategy as the *Candida parapsilosis* /  
332 *C. orthopsilosis* / *C. metapsilosis* group, but in the case of both black yeasts the hybrids  
333 are not recognised as separate species (for reasons discussed below).

334 In the reticulate history of *H. werneckii* and *A. melanogenum*, the unit of genetic  
335 exchange are whole haploid genomes. This allowed us to trace phylogenies of haploid  
336 genomes, for example, in the case of *H. werneckii* (Fig. 4A) of the “green haploid”  
337 genome (strains 2 and 3) and the “red haploid” genome (strains 4 and C), as well as  
338 their “red and green diploid” hybrids (group 1). The absence of haploid strains with  
339 genomes combined from two or more genomic lineages implies that – as is often seen  
340 in hybrids – diploids do not revert to haploid state but are stuck at the F1 stage. They  
341 are either unable to undergo meiosis, or the progeny of such meiosis has sufficiently  
342 low viability or fitness to evade sampling. The presence of a putative homothallic  
343 mating-type locus in the majority of sequenced strains and the distribution of diversified  
344 lineages of MAT genes in groups of hybrids and haploids raise the possibility that these  
345 loci play a role in the formation of hybrid lineages. At the same time a substantial  
346 number of strains containing unexpected combinations of MAT lineages, the fast  
347 diversification of the MAT genes (particularly in *H. werneckii*), and the methodological  
348 limitations of the analysis (suboptimal assembly of the putative mating-type locus),  
349 demand further research before providing conclusive evidence on the presence, type  
350 and functionality of mating-type genes, and their possible role in the formation of  
351 hybrids in *H. werneckii* and *A. melanogenum*.

352 While hybridization to diploids and the absence of meiosis is reminiscent of  
353 parasexuality, *H. werneckii* and *A. melanogenum* do not conform to this mode of  
354 reproduction either. In parasexual reproduction diploids typically revert to haploids

355 through haploidisation – a random loss of chromosomes with the end result similar to  
356 meiosis [17]. In addition to the absence of recombinant haploid strains there is also  
357 little evidence for the existence of intermediate aneuploid states characteristic of  
358 haploidisation. Either such haploidisation does not occur, or its products are not viable  
359 due to incompatibility of parental chromosomes. The aneuploidy observed in both *H.*  
360 *werneckii* and *A. melanogenum* mostly involves increases in ploidy above the diploid  
361 state rather than decreases below it, as would be expected with haploidisation. Thus,  
362 this aneuploidy is almost certainly not part of a parasexual cycle, but might be an  
363 adaptive evolutionary response to adverse or novel conditions – a common adaptive  
364 response of fungi [20–22]. This explanation is also supported by the observations of  
365 large-scale duplications of specific genomic regions in a diploid strain of *H. werneckii*  
366 subjected to long-term experimental evolution at extreme salinity [46]. Interestingly, the  
367 high level of heterozygosity of diploid strains of *H. werneckii* and *A. melanogenum*  
368 resulting from hybridisation is mostly preserved not only by the rarity of large-scale  
369 deletions, but also by the relative scarcity of large-scale copy-neutral loss of  
370 heterozygosity. In other hybrids loss of heterozygosity has been recognised as a  
371 common tool for genome shaping and stabilisation after hybridization [47,48], but  
372 appears to be largely avoided by both *H. werneckii* and *A. melanogenum*.

373 An integral part of fungal parasexuality is the heterokaryon – a cell with two genetically  
374 distinct nuclei that sometimes undergo karyogamy [49]. In *H. werneckii* both haploid  
375 and diploid cells contain a single nucleus per cell [50]. The viability and stability of  
376 heterokaryons are controlled by heterokaryon incompatibility loci. At least three genes  
377 for proteins with domains characteristic of such loci were found in all *A. melanogenum*  
378 genomes and at least 17 in *H. werneckii* (with up to 43 in diploid genomes, although  
379 this number may be an overestimate due to fragmented genome assembly). While it  
380 has been shown in some species that even differences in heterokaryon incompatibility  
381 loci as small as a single amino acid can be sufficient to trigger incompatibility [51], the  
382 diversity of these loci in *H. werneckii* and *A. melanogenum* was much higher than that.  
383 The distribution of certain types of these loci is consistent with the hybrid groups  
384 described above. On the one hand, this could simply reflect the phylogenetic distance  
385 between the strains. On the other hand, it opens the possibility that heterokaryon  
386 incompatibility loci might be involved in the successful formation of diploid hybrids in  
387 *H. werneckii* and *A. melanogenum*.

388

389 Hybrid fungal genomes have so far been described mostly in pathogens of animals or  
390 plants. This is the first time we document the formation of stable and highly  
391 heterozygous diploids in wild populations of two extremotolerant clonal species. Five  
392 aspects of this phenomenon are discussed below.

393 1. Is clonality related to the extremotolerance of *H. werneckii* and *A. melanogenum*? It  
394 has long been speculated that asexuality may be advantageous in extreme  
395 environments, conserving the energy required for sexual reproduction and allowing the  
396 fixation of beneficial alleles and genomic configurations in small populations that have  
397 managed to adapt to extreme conditions at the ecological edge of the species [14,15].  
398 One of the mechanisms that can promote adaptation at the margin of species  
399 distribution is hybridization [52].

400 Of course, sexual reproduction does not only have shortcomings, but also considerable  
401 advantages in adapting to stress [6,12]. The same is true for parasexuality, which not  
402 only alters ploidy but also increases diversity through cycles of regular and double  
403 ploidy. A high frequency of diploids in *Aspergillus fumigatus* has been reported in cystic  
404 fibrosis, presumably due to local stress (e.g., nitrogen deficiency or the presence of  
405 certain drugs) that promotes parasexual recombination [18]. In *Candida albicans*,  
406 which can undergo regular sexual recombination, stress additionally promotes the  
407 parasexual cycle, which generates a high degree of diversity, including aneuploidy  
408 [53].

409 There are also several examples of at least occasionally recombining extremotolerant  
410 and extremophilic species. For example, the polyextremotolerant yeast *Aureobasidium*  
411 *pullulans*, a close relative of *A. melanogenum*, has one of the highest rates of  
412 recombination demonstrated in fungi by population genomics [38]. Two salt-adapted  
413 basidiomycetes, the halotolerant *Wallemia mellicola* and the halophilic *Wallemia*  
414 *ichthyophaga*, also appear to recombine, albeit much less frequently than *A. pullulans*,  
415 even though *W. ichthyophaga* is exceptionally rare and limited to highly fragmented  
416 environments [9,54]. If clonality or hybridization are indeed beneficial for adaptation to  
417 extreme conditions – and more data are needed to test this hypothesis – recombination  
418 appears to be compatible with extremotolerant lifestyle as well.



419 2. Do clonality and hybridization allow for greater specialisation? While recombination  
420 generates potentially useful diversity and thus provides substrate for natural selection,  
421 it can also break successful genomic configurations – a shortcoming known as  
422 recombination load [13–15]. In well-adapted subpopulations clonality prevents  
423 beneficial adaptations from being diluted by gene flow from other environments, which  
424 promote different adaptations. Clonal lineages may thus be more successful in the  
425 short term, but may collapse due to reduced adaptability or Muller’s ratchet and are  
426 replaced by the next successful clone, which can be generated by sexual or parasexual  
427 recombination – or hybridization [55].

428 Some species are able to thrive in a wide range of different environments without  
429 adapting to any of them at the genomic level – the ubiquitous and polyextremotolerant  
430 *A. pullulans* is one such example [38]. But while *A. pullulans* is an exceptionally  
431 generalistic species, the two species analysed here are less so: *A. melanogenum* is  
432 mostly restricted to aquatic and indoor environments, while *H. werneckii* is mostly  
433 found in marine and hypersaline environments and has a much higher upper salinity  
434 growth limit than *A. pullulans*. The preference of some *H. werneckii* and *A.*  
435 *melanogenum* strain groups for specific habitats (Fig. 4) possibly indicates an ongoing  
436 clonality-driven specialisation of these groups. This would be in line with the suggestion  
437 of Romeo et al. [35]. However, based on the here studied dataset the observed habitat  
438 preferences might be an artefact of skewed geographic distribution due to limited  
439 dispersal and unequal habitat sampling in different locations.

440 Interestingly, the clinical isolates of *H. werneckii* belong to different strain groups within  
441 the species (Fig. 4). Although the number of clinical isolates analysed here was too  
442 small to draw reliable conclusions, this could mean that no lineage within the species  
443 is better able to cause infections in humans than others. Aneuploidy was observed in  
444 four out of five analysed clinical strains of both species. As discussed above,  
445 aneuploidy can be a signature of adaptation to novel environments [20–22]. Both *H.*  
446 *werneckii* and *A. melanogenum* are opportunistic pathogens that rarely cause  
447 infections. The conditions they encounter in the human body are almost certainly  
448 suboptimal for their survival [56,57], resulting in high selection pressure and possibly  
449 in aneuploidies. However, due to their rarity, such infections most likely do not  
450 contribute meaningfully to the evolution and specialisation of either of the two species  
451 [56].

452 3. If the formation of diploids in *H. werneckii* and *A. melanogenum* is irreversible, what  
453 drives the co-existence of diploid and haploid strains? A study of 12 *H. werneckii*  
454 genomes reported 7 successful hybridization events, and expansion of the data set to  
455 66 genomes uncovered only 2 or 3 additional hybridizations in the history of the  
456 species. While isolation of strains from novel environments or geographic locations  
457 might lead to the discovery of new hybrid lineages, their number is likely to remain  
458 limited. This might indicate that hybridization events are relatively uncommon or that  
459 only a small number of them result in offspring with sufficient fitness to persist in the  
460 environment. The co-existence of haploid and diploid strains may be supported by their  
461 divergent performance in different conditions. A preliminary comparison of  
462 halotolerance between diploid and haploid strains shows slightly higher halotolerance  
463 of diploid *A. melanogenum*, but no such difference in *H. werneckii* (our unpublished  
464 data), but this comparison was limited to only one parameter tested in a laboratory  
465 setting. The possibility of different adaptation value of haploids vs. diploids should be  
466 more carefully addressed in the future, for example by competition experiments.

467 Whatever the mechanism of hybridization, it appears to operate almost exclusively  
468 between haploid strains, and compared to many other fungal species [20] the range of  
469 observable ploidies in *H. werneckii* and *A. melanogenum* is modest. Although a  
470 randomly selected environmental strain is about twice as likely to be diploid as haploid,  
471 a single tetraploid strain of *H. werneckii* is the only evidence that these diploids can  
472 hybridize further. No triploid strains of either species have been found. Either the fusion  
473 of diploid cells is prevented by some as yet unknown mechanism, or the resulting  
474 strains do not persist in the environment long enough to be detected.

475 4. How should the hybrids of clonal species be treated in taxonomy? In such situation  
476 even the definition of otherwise well-established terminology is not trivial. For example,  
477 Naranjo–Ortiz and Gabaldón [21] defined hybrids as lineages emerging from  
478 ancestors, which differ from each other more than the most distant strains of well-  
479 recognised species. According to Boekhout et al. [47], lineages of interspecies hybrids  
480 can be recognised as separate species, while intraspecies hybrids better fit in the  
481 concept of varieties. This returns us to the problem of species delineation in clonal  
482 taxa, which may involve arbitrary decisions. On the one hand, diversity in *H. werneckii*  
483 is high, and distances between genomes of some strains are substantially greater than  
484 what is typical for fungal species [58]. On the other hand, Fig. 4 clearly illustrates why

485 a more fragmented taxonomy of *Hortaea* would result in unpractical taxonomic  
486 inflation. If hybrids and the remaining monophyletic groups of haploids were treated as  
487 species, *H. werneckii* could easily be split into 15 or more new species. In several of  
488 these new species two different sequences of standard taxonomic markers carried by  
489 a single diploid strain [32] would in many cases belong to different species – a  
490 decidedly untenable situation. Similarly, if all haploid strains were treated as one  
491 species and all diploids as another, such species would be polyphyletic. Since clonality  
492 precludes the application of the biological species concept to *H. werneckii*, we suggest  
493 the dense reticulation of its phylogeny can be pragmatically interpreted as an analogue  
494 of interbreeding. Thus, a single *H. werneckii* species is maintained despite its high  
495 diversity, as suggested also by a recent in-depth phylogenetic study of the taxon [32].  
496 Although a comparably detailed taxonomic revision of *Aureobasidium melanogenum*  
497 is still pending, it is expected to lead to a similar outcome.

498 5. How common is hybridisation in clonal fungi? Clonality itself appears to be much  
499 rarer in fungi than once believed, but population genomic studies of *Neurospora* spp.  
500 showed that even closely related species can differ substantially in their reproductive  
501 strategies [3]. This is also the case here: while *A. melanogenum* is clonal, the closely  
502 related species *A. pullulans* has exceptionally high recombination rates [38]. Genome  
503 sequencing of another species of the genus, *Aureobasidium subglaciale*, revealed a  
504 small number of apparently clonal diploid strains that may belong to a new species  
505 [36], the reproductive strategy of which should be investigated if more such strains can  
506 be isolated and sequenced.

507 Of five fungal species from extreme environments that we have previously studied  
508 using population genomics, two were strictly clonal and both contained stable diploid  
509 hybrids. This situation is at least similar to the one described in the *Candida*  
510 *parapsilosis* / *C. orthopsilosis* / *C. metapsilosis* group [48]. Other examples may include  
511 the plant pathogen *Verticillium longisporum* [59], and some clinically relevant hybrids  
512 of *Cryptococcus* spp. [60] and *Aspergillus* spp. [61]. Such reproductive strategy may  
513 thus be more common than currently known, especially since it can be easily  
514 overlooked without performing careful population genomic studies. Polyploid strains  
515 often produce highly fragmented but otherwise inconspicuous assemblies, and even  
516 after genome sequencing, hybrids, polyploids, and aneuploids can easily go

517 undetected [21]. Any study that discovers genomes of different ploidy in clonal fungal  
518 species should investigate hybridization as a possible explanation for the data.

519

## 520 **Conclusions**

521 Genome sequencing of two black yeasts from extreme environments, *H. werneckii* and  
522 *A. melanogenum*, revealed that both species are strictly clonal. Their populations  
523 consist of both haploid and diploid strains, and diploid strains were formed by a handful  
524 of hybridization events between haploids. These hybridizations were not followed by  
525 meiosis as part of sexual reproduction, nor by haploidisation through random  
526 chromosome loss, as is typical of parasexuality. Hybrid lineages avoid the loss of  
527 heterozygosity even over timeframes allowing them to disperse over large geographic  
528 distances. Such “stable parasexuality”, a process of forming stable and highly  
529 heterozygous diploids in a clonal species without evidence of subsequent meiosis or  
530 haploidisation, is an unusual reproductive strategy, which merits further study. This is  
531 the first time it has been documented in wild populations of extremotolerant fungi. The  
532 increasing use of population genomics in fungi will show whether this reproductive  
533 strategy is more widespread than is currently known and careful comparative studies  
534 should investigate its potential role in adaptation to extreme (and other) environments.

535

## 536 **Materials and methods**

### 537 ***Cultivation and DNA isolation***

538 Strains of the extremely halotolerant *Hortaea werneckii* (Table 1) were obtained from  
539 the Ex Culture Collection of the Department of Biology, Biotechnical Faculty, University  
540 of Ljubljana (IC Mycosmo, MRIC UL, Slovenia). The cultivation and DNA isolation were  
541 performed as described previously [1], using the standard chemically defined Yeast  
542 Nitrogen Base medium (Qbiogene). Biomass harvested from liquid cultures was frozen  
543 in liquid nitrogen and kept at -80 °C until DNA isolation, performed as described  
544 previously [1], using the UltraClean Microbial DNA isolation kit (MO BIO Laboratories,  
545 USA), preceded by homogenisation with a pestle and mortar in liquid nitrogen and 1  
546 min in Retsch Mixer Mill 301 (ThermoFisher Scientific, USA) at 20 Hz. After the RNase  
547 A treatment (ThermoFisher Scientific, USA), the isolated DNA was evaluated using  
548 agarose electrophoresis and by fluorometry (Qubit; ThermoFisher Scientific, USA).

549

550 **Genome sequencing**

551 The genome sequencing was performed using the platform BGISEQ-500, with 2×150-  
552 bp libraries, prepared as described previously [62], with multiplexed sequencing  
553 barcodes. The resulting output was demultiplexed, the quality was checked with  
554 FastQC, and the reads were trimmed for adaptors and quality (removal of bases with  
555 Q <20) using the 'bbduk' script (<https://jgi.doe.gov/data-and-tools/bbtools/>).

556 The raw sequencing reads have been deposited in China National GeneBank  
557 Sequence Archive (CNSA) of China National GeneBank DataBase (CNGBdb) with  
558 accession number CNP0001993. Sequencing reads together with assembly and  
559 annotation data have been deposited in Genbank under BioProject PRJNA428320.  
560 Genome sequences of previously sequenced *H. werneckii* strains [1] are deposited in  
561 Genbank under the same BioProject (PRJNA428320). Genome sequences of  
562 previously sequenced *A. melanogenum* strains [36] are deposited in Genbank under  
563 the BioProject PRJNA721240 and listed in Table 2. Genome 7 from the study by  
564 Černoša et al. [36] was excluded from this study due to the large phylogenetic distance  
565 from other *A. melanogenum* strains, while strains 2, 18, and 38 were excluded from  
566 phylogenetic analyses based on alignments produced by SibeliaZ (described below)  
567 due to their unclear ploidy.

568

569 **Variant calling**

570 Sequencing reads of *H. werneckii* genomes were mapped to the reference genome of  
571 *H. werneckii* EXF-2000 (GenBank MUNK00000000.1 [34]), which was first haploidised  
572 with HaploMerger2 [63]. Mapping was performed by 'bwa mem', using the default  
573 parameters. The reads were sorted with Samtools 1.6 [64], deduplicating with Picard  
574 2.10.2 and then used for variant calling with Genome Analysis Toolkit 4.1 [65], following  
575 the 'Genome Analysis Toolkit (GATK) Best Practices' workflow in diploid mode, but  
576 using 'hard filtering' with the expression 'QD < 2.0 || FS > 20.0 || SOR > 3.0 || MQ <  
577 50.0'. Strain 36 was excluded from the analysis due to its tetraploid genome. Variants  
578 of *A. melanogenum* genomes were determined by Černoša et al. [36].

579

580 **Variant-based analysis**

581 Variant-based analyses for both *H. werneckii* and *A. melanogenum* were performed in  
582 R [66], except the calculation of the linkage disequilibrium squared correlation  
583 coefficient ( $r^2$ ; described below). Genomes were clustered based on the single-  
584 nucleotide polymorphism data using the principal component analysis with the 'glPca'  
585 function of the 'adgenet' package in R [67]. The phylogenetic networks estimated from  
586 SNP data were reconstructed with the Neighbor-Net algorithm of the package  
587 'phangorn' [68] based on a dissimilarity distance matrix calculated with the package  
588 'poppr' [69].

589 Linkage disequilibrium squared correlation coefficient ( $r^2$ ) was calculated for all pairs  
590 of biallelic SNP loci within 10 000 nucleotides of each other with 'vcftools' [70]. Then  $r^2$   
591 was plotted as a function of distance between pairs of loci using 'ggplot2' [71]. A  
592 generalized additive model ("gam") curve was fitted to the data.

593

594 **Assembly and annotation**

595 Reference-guided genome assembly was performed for all here sequenced *H.*  
596 *werneckii* genomes with IDBA-Hybrid 1.1.3 [72] using the same reference as for variant  
597 calling. The maximum k value selected was 120, the minimum support in each iteration  
598 was 2, the similarity threshold for alignment was 0.95, seed kmer was 20, maximum  
599 allowed gap in the reference was 100, and the minimum size of contigs included in the  
600 final assembly was 500. Genomes were annotated with Augustus 3.4 [73]. Augustus  
601 parameters were optimised with training using the scripts provided with the program  
602 with (i) the RNAseq data from Sinha et al. [34] deposited at GenBank Sequence Read  
603 Archive with the accession number SRP094740 and (ii) all predicted proteins of *H.*  
604 *werneckii* EXF-2000 (GenBank MUNK00000000.1). These hints were also used for  
605 the final annotation.

606 Predicted proteomes were benchmarked with the BUSCO 4.1.1 [74] using the default  
607 parameter values and the data set for Dothideomycetes [75].

608 The files for submission to GenBank were prepared with the Genome Annotation  
609 Generator (GAG) 2.0.1 [76]. Gene models with short coding regions (<150 bp) and/or  
610 introns (<10 bp) were removed before the submission.

611

## 612 ***Phylogenetic analyses***

613 Sibeliaz 1.2.2 [77] was used to align parts of the genomes of both *H. werneckii* and *A.*  
614 *melanogenum* into multiple sequence alignments. The parameters used were  $k=21$ ,  
615  $a=150$ ,  $b=15000$ . Alignments were then filtered to keep only those in which the number  
616 of sequences from each genome exactly matched the ploidy of the genome.  
617 Alignments were optimised with Gblocks 0.91 [78], using the options '-b3=10 -b4=3 -  
618 b5=n' and then inspected manually to trim the ends to the shortest sequence in the  
619 alignment and remove any alignments with long gaps in any of the sequences. Fifty  
620 longest alignments were selected for each species and each alignment was used for  
621 the estimation of the phylogenetic tree with IQ-TREE 2.0.3 using standard model  
622 selection and 1000 replicates for the SH approximate likelihood ratio test [79]. The  
623 resulting collection of 50 phylogenetic trees for each species was visualized as an  
624 overlay using 'densiTree()' from the 'phangorn' package in R [68] and as a consensus  
625 supernetwork using SplitsTree 4.16.2 [80]. These visualisations and a majority rule  
626 consensus tree calculated with the 'consensus.edges' from the package 'phytools' in  
627 R [81] were used to draw a schematic representation of phylogenetic histories of  
628 genomes in the open-source vector graphics software Inkscape 1.1  
629 (<http://inkscape.org>). Enrichment of phylogenetic clusters of strains for certain  
630 geographic origin or habitat was analysed in R using Fisher's Exact Test with simulated  
631 p-value [66].

632 Genes with HET (PF06085) and Het-C (PF07217) domains were identified in predicted  
633 proteomes of all strains investigated in this study (Table 1,2) with 'hmmsearch' 3.3.1  
634 and HMM profiles with default parameters from the Pfam-A.hmm database version  
635 34.0 [82]. The identified proteins were aligned with Mafft 7.475 [83] and the alignments  
636 were used for reconstruction of phylogenies with IQ-TREE 2.0.3 using standard model  
637 selection and 1000 replicates for the SH approximate likelihood ratio test [79].

638 Putative mating-type loci were identified in the genomes with stand-alone BLAST  
639 2.9.0+ [84], aligned with Mafft 7.475 [83], and annotated based on previously published  
640 annotations of mating-type loci in *H. werneckii* [1] and *A. melanogenum* [39].  
641 Phylogeny of putative MAT1-1 and MAT1-2 homologues was estimated with IQ-TREE  
642 2.0.3 using standard model selection and 1000 replicates for the SH approximate

643 likelihood ratio test [79] after first excluding all putative homologues truncated to less  
644 than 80% of expected length due to suboptimal genome assembly.

645

646 ***Detection of aneuploidies and loss of heterozygosity***

647 Per-nucleotide sequencing depth of reads mapped to the reference genome as  
648 described above was calculated with Samtools 1.6 [64]. For each sequenced genome  
649 the median values of per-nucleotide depths in 30 kbp windows were plotted as  
650 proportion of the median depth of the whole genome. These values were calculated in  
651 R and visualised with 'ggplot2' [66,71] for 50 longest reference genome contigs in the  
652 case of *H. werneckii* and for 35 longest reference genome contigs in the case of *A.*  
653 *melanogenum*.

654 Evidence for loss of heterozygosity in diploid genomes was searched for by counting  
655 the number of heterozygous SNPs in 25 kbp windows along the longest reference  
656 genome contigs (50 in case of *H. werneckii*, 35 in case of *A. melanogenum*) and plotted  
657 as proportion of the median heterozygosity of each genome with 'ggplot2' [66,71].  
658 Depth of sequencing was plotted as described above, but in 25 kbp windows, to  
659 distinguish between copy-neutral loss of heterozygosity and loss of heterozygosity  
660 caused by aneuploidy.

661



## 662 **Acknowledgements**

663 The authors would like to thank Yonglun Luo (Lars Bolund Institute of Regenerative  
664 Medicine, Qingdao-Europe Advanced Institute for Life Sciences, BGI-Qingdao; BGI-  
665 Shenzhen; Department of Biomedicine, Aarhus University) for supporting the work on  
666 genomics of fungi from extreme environments; Yuchong Tang (China National  
667 GeneBank, BGI-Shenzhen) for his invaluable help in project management,  
668 organization of scientific visits and facilitation of collaboration between the project  
669 partners; and Toni Gabaldón (Biomedical Research Institute (IRB), Barcelona  
670 Supercomputing Centre (BSC)) for inspiring and constructive discussions on  
671 hybridization in fungi. The authors thank Rafael R. Montalvo-Rodriguez (University of  
672 Puerto Rico) for the strains EXF-2515 and EXF-2516; Filomena de Leo (University of  
673 Messina, Italy) for the strains EXF-10508 to EXF-10512; Armando Azua Bustos  
674 (Pontificia Universidad Católica de Chile) for the strains EXF-11540, EXF-11650 and  
675 EXF-11651; Zhu-Hua Luo (Third Institute of Oceanography, China) for the strains EXF-  
676 12619 and EXF-12620; Amy Gladfelter (University of North Carolina) for the strains  
677 EXF-14591 and EXF-14592; and Christine Beemelmans (Hans Knöll Institute,  
678 Germany) for the strain EXF-14590. We acknowledge the China National Gene Bank  
679 for the support of sequencing library preparation and shotgun sequencing. This study  
680 was supported by funding from the Slovenian Research Agency to Infrastructural  
681 Centre Mycosmo (MRIC UL I0-0022), programmes P4-0432, P1-0198, and project J4-  
682 2549.

683

## 684 **Authors contributions**

685 Conceptualization of the study: NGC, ZS and CG; experimental work: XS, AČ, and CF;  
686 bioinformatic analyses: CG; data curation: CG; resources: ZS, NGC; preparation of the  
687 manuscript and visualisations: CG; review and editing of the manuscript: CG, NGC,  
688 AČ, XS, CF and ZS; supervision: ZS and NGC; funding acquisition: NGC, ZS.

689

## 690 **Declarations**

## 691 **Funding**

692 The authors acknowledge the China National Gene Bank for the support of sequencing  
693 library preparation and shotgun sequencing. This study was supported by funding from  
694 the Slovenian Research Agency to Infrastructural Centre Mycosmo (MRIC UL),  
695 programmes P1-0170, P1-0198, project J4-2549, and young researcher grant to Anja  
696 Černoša.

697 **Conflicts of interest/Competing interests**

698 Not applicable.

699 **Ethics approval**

700 Not applicable.

701 **Consent to participate**

702 Not applicable.

703 **Consent for publication**

704 Not applicable.

705 **Availability of data and material**

706 All data used in the study are available in GenBank.

707 **Code availability**

708 Not applicable.

709

710 **References**

- 711 1. Gostinčar C, Stajich JE, Zupančič J, Zalar P, Gunde-Cimerman N. Genomic  
712 evidence for intraspecific hybridization in a clonal and extremely halotolerant yeast.  
713 *BMC Genomics*. 2018; doi: 10.1186/s12864-018-4751-5.
- 714 2. Brown A. Mating in mushrooms: increasing the chances but prolonging the affair.  
715 *Trends Genet*. 2001; doi: 10.1016/S0168-9525(01)02343-5.
- 716 3. Gladieux P, De Bellis F, Hann-Soden C, Svedberg J, Johannesson H, Taylor JW.  
717 *Neurospora* from natural populations: Population genomics insights into the life history  
718 of a model microbial eukaryote. In: Y. DJ, editor. *Stat Popul Genomics*. New York:  
719 Humana;
- 720 4. Hawksworth DL, Kirk PM, Sutton BC, Pegler DN. Ainsworth & Bisby's Dictionary of  
721 the Fungi. 8th ed. Oxford University Press;
- 722 5. Tibayrenc M, Ayala FJ. Reproductive clonality of pathogens: A perspective on  
723 pathogenic viruses, bacteria, fungi, and parasitic protozoa. *Proc Natl Acad Sci*. 2012;  
724 doi: 10.1073/pnas.1212452109.
- 725 6. Taylor JW, Hann-Soden C, Branco S, Sylvain I, Ellison CE. Clonal reproduction in  
726 fungi. *Proc Natl Acad Sci*. 2015; doi: 10.1073/pnas.1503159112.
- 727 7. Carreté L, Ksiezopolska E, Pegueroles C, Gómez-Molero E, Saus E, Iraola-Guzmán  
728 S, et al.. Patterns of genomic variation in the opportunistic pathogen *Candida glabrata*  
729 suggest the existence of mating and a secondary association with humans. *Curr Biol*.  
730 2018; doi: 10.1016/j.cub.2017.11.027.
- 731 8. Kasuga T, White TJ, Koenig G, Mcewen J, Restrepo A, Castañeda E, et al..  
732 Phylogeography of the fungal pathogen *Histoplasma capsulatum*. *Mol Ecol*. 2003; doi:  
733 10.1046/j.1365-294X.2003.01995.x.
- 734 9. Gostinčar C, Sun X, Zajc J, Fang C, Hou Y, Luo Y, et al.. Population genomics of  
735 an obligately halophilic basidiomycete *Wallemia ichthyophaga*. *Front Microbiol*. 2019;  
736 doi: 10.3389/fmicb.2019.02019.
- 737 10. Tsai IJ, Bensasson D, Burt A, Koufopanou V. Population genomics of the wild yeast  
738 *Saccharomyces paradoxus*: Quantifying the life cycle. *Proc Natl Acad Sci*. 2008; doi:  
739 10.1073/pnas.0707314105.
- 740 11. Xu J. Cost of interacting with sexual partners in a facultative sexual microbe.  
741 *Genetics*. 2005; doi: 10.1534/genetics.105.045302.
- 742 12. Sun S, Heitman J. Is sex necessary? *BMC Biol*. 2011; doi: Artn 56 Doi

743 10.1186/1741-7007-9-56.

744 13. Otto SP. The evolutionary enigma of sex. *Am Nat.* 2009; doi: 10.1086/599084.

745 14. Gostinčar C, Grube M, De Hoog S, Zalar P, Gunde-Cimerman N. Extremotolerance  
746 in fungi: evolution on the edge. *FEMS Microbiol Ecol.* 2010; doi: 10.1111/j.1574-  
747 6941.2009.00794.x.

748 15. Gostinčar C, Gunde-Cimerman N, Grube M. Polyextremotolerance as the fungal  
749 answer to changing environments. In: Bakermans C, editor. *Microb Evol under Extrem*  
750 *Cond.* Berlin, München, Boston: DE GRUYTER;

751 16. Pontecorvo G. The parasexual cycle in fungi. *Annu Rev Microbiol.* 1956; doi:  
752 10.1146/annurev.mi.10.100156.002141.

753 17. Clutterbuck AJ. Parasexual recombination in fungi. *J Genet.* 1996; doi:  
754 10.1007/BF02966308.

755 18. Engel T, Verweij PE, van den Heuvel J, Wangmo D, Zhang J, Debets AJM, et al..  
756 Parasexual recombination enables *Aspergillus fumigatus* to persist in cystic fibrosis.  
757 *ERJ Open Res.* 2020; doi: 10.1183/23120541.00020-2020.

758 19. Todd RT, Forche A, Selmecki A. Ploidy variation in fungi: Polyploidy, aneuploidy,  
759 and genome evolution. *Microbiol Spectr.* 2017; doi: 10.1128/microbiolspec.FUNK-  
760 0051-2016.

761 20. Gerstein AC, Sharp NP. The population genetics of ploidy change in unicellular  
762 fungi. *FEMS Microbiol Rev.* 2021; doi: 10.1093/femsre/fuab006.

763 21. Naranjo–Ortiz MA, Gabaldón T. Fungal evolution: cellular, genomic and metabolic  
764 complexity. *Biol Rev.* 2020; doi: 10.1111/brv.12605.

765 22. Hill R, Leitch IJ, Gaya E. Targeting Ascomycota genomes: what and how big?  
766 *Fungal Biol Rev.* 2021; doi: 10.1016/j.fbr.2021.03.003.

767 23. Chen G, Rubinstein B, Li R. Whole chromosome aneuploidy: Big mutations drive  
768 adaptation by phenotypic leap. *BioEssays.* 2012; doi: 10.1002/bies.201200069.

769 24. Gerstein AC, Lim H, Berman J, Hickman MA. Ploidy tug-of-war: Evolutionary and  
770 genetic environments influence the rate of ploidy drive in a human fungal pathogen.  
771 *Evolution (N Y).* 2017; doi: 10.1111/evo.13205.

772 25. Peter J, De Chiara M, Friedrich A, Yue J-X, Pflieger D, Bergström A, et al.. Genome  
773 evolution across 1,011 *Saccharomyces cerevisiae* isolates. *Nature.* 2018; doi:  
774 10.1038/s41586-018-0030-5.

775 26. Depotter JR, Seidl MF, Wood TA, Thomma BP. Interspecific hybridization impacts  
776 host range and pathogenicity of filamentous microbes. *Curr Opin Microbiol.* 2016; doi:

777 10.1016/j.mib.2016.04.005.

778 27. Samarasinghe H, You M, Jenkinson TS, Xu J, James TY. Hybridization facilitates  
779 adaptive evolution in two major fungal pathogens. *Genes (Basel)*. 2020; doi:  
780 10.3390/genes11010101.

781 28. Morrow CA, Fraser JA. Ploidy variation as an adaptive mechanism in human  
782 pathogenic fungi. *Semin Cell Dev Biol*. 2013; doi: 10.1016/j.semcdb.2013.01.008.

783 29. Forche A. Large-scale chromosomal changes and associated fitness  
784 consequences in pathogenic fungi. *Curr Fungal Infect Rep*. 2014; doi:  
785 10.1007/s12281-014-0181-2.

786 30. Gunde-Cimerman N, Zalar P, Hoog S, Plemenitaš A. Hypersaline waters in  
787 salterns - natural ecological niches for halophilic black yeasts. *FEMS Microbiol Ecol*.  
788 2000; doi: 10.1111/j.1574-6941.2000.tb00716.x.

789 31. Plemenitaš A, Lenassi M, Konte T, Kejžar A, Zajc J, Gostinčar C, et al.. Adaptation  
790 to high salt concentrations in halotolerant/halophilic fungi: a molecular perspective.  
791 *Front Microbiol*. 2014; doi: 10.3389/fmicb.2014.00199.

792 32. Zalar P, Zupančič J, Gostinčar C, Zajc J, de Hoog GSS, De Leo F, et al.. The  
793 extremely halotolerant black yeast *Hortaea werneckii* - a model for intraspecific  
794 hybridization in clonal fungi. *IMA Fungus*. 2019; doi: 10.1186/s43008-019-0007-5.

795 33. Lenassi M, Gostinčar C, Jackman S, Turk M, Sadowski I, Nislow C, et al.. Whole  
796 genome duplication and enrichment of metal cation transporters revealed by *de novo*  
797 genome sequencing of extremely halotolerant black yeast *Hortaea werneckii*. *PLoS*  
798 *One*. 2013; doi: 10.1371/journal.pone.0071328.

799 34. Sinha S, Flibotte S, Neira M, Formby S, Plemenitaš A, Gunde-Cimerman N, et al..  
800 Insight into the recent genome duplication of the halophilic yeast *Hortaea werneckii*:  
801 Combining an improved genome with gene expression and chromatin structure. *G3-*  
802 *Genes Genomes Genet*. 2017; doi: 10.1534/g3.117.040691.

803 35. Romeo O, Marchetta A, Giosa D, Giuffrè L, Urzì C, De Leo F. Whole genome  
804 sequencing and comparative genome analysis of the halotolerant deep sea black yeast  
805 *Hortaea werneckii*. *Life*. 2020; doi: 10.3390/life10100229.

806 36. Černoša A, Sun X, Gostinčar C, Fang C, Gunde-Cimerman N, Song Z. Virulence  
807 traits and population genomics of the black yeast *Aureobasidium melanogenum*. *J*  
808 *Fungi*. 2021; doi: 10.3390/jof7080665.

809 37. Azua-Bustos A, Gonzalez-Silva C, Salas L, Palma RE, Vicuna R. A novel subaerial  
810 *Dunaliella* species growing on cave spiderwebs in the Atacama Desert. *Extremophiles*.

811 2010; doi: 10.1007/s00792-010-0322-7.

812 38. Gostinčar C, Turk M, Zajc J, Gunde-Cimerman N. Fifty *Aureobasidium pullulans*  
813 genomes reveal a recombining polyextremotolerant generalist. *Environ Microbiol.*  
814 2019; doi: 10.1111/1462-2920.14693.

815 39. Gostinčar C, Ohm RA, Kogej T, Sonjak S, Turk M, Zajc J, et al.. Genome  
816 sequencing of four *Aureobasidium pullulans* varieties: biotechnological potential,  
817 stress tolerance, and description of new species. *BMC Genomics.* 2014; doi:  
818 10.1186/1471-2164-15-549.

819 40. Zhu YO, Sherlock G, Petrov DA. Whole genome analysis of 132 clinical  
820 *Saccharomyces cerevisiae* strains reveals extensive ploidy variation. *G3*  
821 *Genes/Genomes/Genetics.* 2016; doi: 10.1534/g3.116.029397.

822 41. O'Brien CE, Oliveira-Pacheco J, Ó Cinnéide E, Haase MAB, Hittinger CT, Rogers  
823 TR, et al.. Population genomics of the pathogenic yeast *Candida tropicalis* identifies  
824 hybrid isolates in environmental samples. Vylkova S, editor. *PLOS Pathog.* 2021; doi:  
825 10.1371/journal.ppat.1009138.

826 42. Rhodes J, Desjardins CA, Sykes SM, Beale MA, Vanhove M, Sakthikumar S, et  
827 al.. Tracing genetic exchange and biogeography of *Cryptococcus neoformans* var.  
828 *grubii* at the global population level. *Genetics.* 2017; doi:  
829 10.1534/genetics.117.203836.

830 43. Gabaldón T. Hybridization and the origin of new yeast lineages. *FEMS Yeast Res.*  
831 2020; doi: 10.1093/femsyr/foaa040.

832 44. Prysycz LP, Németh T, Gácsér A, Gabaldón T. Genome comparison of *Candida*  
833 *orthopsilosis* clinical strains reveals the existence of hybrids between two distinct  
834 subspecies. *Genome Biol Evol.* 2014; doi: 10.1093/gbe/evu082.

835 45. Prysycz LP, Németh T, Saus E, Ksiezopolska E, Hegedúsová E, Nosek J, et al..  
836 The genomic aftermath of hybridization in the opportunistic pathogen *Candida*  
837 *metapsilosis*. Zhang J, editor. *PLOS Genet.* 2015; doi: 10.1371/journal.pgen.1005626.

838 46. Gostinčar C, Stajich JE, Kejžar A, Sinha S, Nislow C, Lenassi M, et al.. Seven  
839 years at high salinity — experimental evolution of the extremely halotolerant black  
840 yeast *Hortaea werneckii*. *J Fungi.* 2021; doi: 10.3390/jof7090723.

841 47. Boekhout T, Aime MC, Begerow D, Gabaldón T, Heitman J, Kemler M, et al.. The  
842 evolving species concepts used for yeasts: from phenotypes and genomes to  
843 speciation networks. *Fungal Divers.* 2021; doi: 10.1007/s13225-021-00475-9.

844 48. Mixão V, Gabaldón T. Hybridization and emergence of virulence in opportunistic

845 human yeast pathogens. *Yeast*. 2018; doi: 10.1002/yea.3242.

846 49. Strom NB, Bushley KE. Two genomes are better than one: history, genetics, and  
847 biotechnological applications of fungal heterokaryons. *Fungal Biol Biotechnol*. 2016;  
848 doi: 10.1186/s40694-016-0022-x.

849 50. Mitchison-Field LMY, Vargas-Muñiz JM, Stormo BM, Vogt EJD, Van Dierdonck S,  
850 Pelletier JF, et al.. Unconventional cell division cycles from marine-derived yeasts. *Curr*  
851 *Biol*. 2019; doi: 10.1016/j.cub.2019.08.050.

852 51. Saupe SJ. Molecular genetics of heterokaryon incompatibility in filamentous  
853 Ascomycetes. *Microbiol Mol Biol Rev*. 2000; doi: 10.1128/MMBR.64.3.489-502.2000.

854 52. Gladieux P, Ropars J, Badouin H, Branca A, Aguilera G, Vienne DM, et al.. Fungal  
855 evolutionary genomics provides insight into the mechanisms of adaptive divergence in  
856 eukaryotes. *Mol Ecol*. 2014; doi: 10.1111/mec.12631.

857 53. Berman J, Hadany L. Does stress induce (para)sex? Implications for *Candida*  
858 *albicans* evolution. *Trends Genet*. 2012; doi: 10.1016/j.tig.2012.01.004.

859 54. Sun X, Gostinčar C, Fang C, Zajc J, Hou Y, Song Z, et al.. Genomic Evidence of  
860 Recombination in the Basidiomycete *Wallemia mellicola*. *Genes (Basel)*. 2019; doi:  
861 10.3390/genes10060427.

862 55. Drenth A, McTaggart AR, Wingfield BD. Fungal clones win the battle, but  
863 recombination wins the war. *IMA Fungus*. 2019; doi: 10.1186/s43008-019-0020-8.

864 56. Gostinčar C, Zajc J, Lenassi M, Plemenitaš A, de Hoog S, Al-Hatmi AMS, et al..  
865 Fungi between extremotolerance and opportunistic pathogenicity on humans. *Fungal*  
866 *Divers*. 2018; doi: 10.1007/s13225-018-0414-8.

867 57. Zajc J, Gostinčar C, Černoša A, Gunde-Cimerman N. Stress-tolerant yeasts:  
868 Opportunistic pathogenicity versus biocontrol potential. *Genes (Basel)*. 2019; doi:  
869 10.3390/genes10010042.

870 58. Gostinčar C. Towards genomic criteria for delineating fungal species. *J Fungi*.  
871 2020; doi: 10.3390/jof6040246.

872 59. Fogelqvist J, Tzelepis G, Bejai S, Ilbäck J, Schwelm A, Dixelius C. Analysis of the  
873 hybrid genomes of two field isolates of the soil-borne fungal species *Verticillium*  
874 *longisporum*. *BMC Genomics*. 2018; doi: 10.1186/s12864-017-4407-x.

875 60. Michelotti LA, Sun S, Heitman J, James TY. Clonal evolution in serially passaged  
876 *Cryptococcus neoformans* x *deneoformans* hybrids reveals a heterogenous landscape  
877 of genomic change. Stajich J, editor. *Genetics*. 2022; doi: 10.1093/genetics/iyab142.

878 61. Steenwyk JL, Lind AL, Ries LNA, dos Reis TF, Silva LP, Almeida F, et al..

879 Pathogenic allodiploid hybrids of *Aspergillus* fungi. *Curr Biol.* 2020; doi:  
880 10.1016/j.cub.2020.04.071.

881 62. Fang C, Zhong H, Lin Y, Chen B, Han M, Ren H, et al.. Assessment of the cPAS-  
882 based BGISEQ-500 platform for metagenomic sequencing. *Gigascience.* 2018; doi:  
883 10.1093/gigascience/gix133.

884 63. Huang S, Kang M, Xu A. HaploMerger2: rebuilding both haploid sub-assemblies  
885 from high-heterozygosity diploid genome assembly. Berger B, editor. *Bioinformatics.*  
886 2017; doi: 10.1093/bioinformatics/btx220.

887 64. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al.. The Sequence  
888 Alignment/Map format and SAMtools. *Bioinformatics.* 2009; doi:  
889 10.1093/bioinformatics/btp352.

890 65. Alkan C, Coe BP, Eichler EE. GATK toolkit. *Nat Rev Genet.* 2011; doi:  
891 10.1038/nrg2958.

892 66. R Development Core Team. R: A language and environment for statistical  
893 computing. Vienna, Austria: R Foundation for Statistical Computing;

894 67. Jombart T, Ahmed I. adegenet 1.3-1: new tools for the analysis of genome-wide  
895 SNP data. *Bioinformatics.* 2011; doi: 10.1093/bioinformatics/btr521.

896 68. Schliep K, Potts AJ, Morrison DA, Grimm GW. Intertwining phylogenetic trees and  
897 networks. *Methods Ecol Evol.* 2017; doi: 10.1111/2041-210X.12760.

898 69. Kamvar ZN, Brooks JC, Grünwald NJ. Novel R tools for analysis of genome-wide  
899 population genetic data with emphasis on clonality. *Front Genet.* 2015; doi:  
900 10.3389/fgene.2015.00208.

901 70. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, et al.. The  
902 variant call format and VCFtools. *Bioinformatics.* 2011; doi:  
903 10.1093/bioinformatics/btr330.

904 71. Wickham H. ggplot2: Elegant graphics for data analysis. Springer-Verlag New  
905 York;

906 72. Peng Y, Leung HCM, Yiu SM, Chin FYL. IDBA-UD: A *de novo* assembler for single-  
907 cell and metagenomic sequencing data with highly uneven depth. *Bioinformatics.*  
908 2012; doi: 10.1093/bioinformatics/bts174.

909 73. Stanke M, Morgenstern B. AUGUSTUS: a web server for gene prediction in  
910 eukaryotes that allows user-defined constraints. *Nucleic Acids Res.* 2005; doi:  
911 10.1093/nar/gki458.

912 74. Simão FA, Waterhouse RM, Ioannidis P, Kriventseva E V., Zdobnov EM. BUSCO:



913 Assessing genome assembly and annotation completeness with single-copy  
914 orthologs. *Bioinformatics*. 2015; doi: 10.1093/bioinformatics/btv351.

915 75. Kriventseva E V., Tegenfeldt F, Petty TJ, Waterhouse RM, Simão FA, Pozdnyakov  
916 IA, et al.. OrthoDB v8: update of the hierarchical catalog of orthologs and the underlying  
917 free software. *Nucleic Acids Res*. 2015; doi: 10.1093/nar/gku1220.

918 76. Geib SM, Hall B, Derego T, Bremer FT, Cannoles K, Sim SB. Genome Annotation  
919 Generator: a simple tool for generating and correcting WGS annotation tables for NCBI  
920 submission. *Gigascience*. 2018; doi: 10.1093/gigascience/giy018.

921 77. Minkin I, Medvedev P. Scalable multiple whole-genome alignment and locally  
922 collinear block construction with SibeliaZ. *Nat Commun*. 2020; doi: 10.1038/s41467-  
923 020-19777-8.

924 78. Talavera G, Castresana J. Improvement of phylogenies after removing divergent  
925 and ambiguously aligned blocks from protein sequence alignments. *Syst Biol*. 2007;  
926 doi: 10.1080/10635150701472164.

927 79. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler  
928 A, et al.. IQ-TREE 2: New models and efficient methods for phylogenetic inference in  
929 the genomic era. Teeling E, editor. *Mol Biol Evol*. 2020; doi: 10.1093/molbev/msaa015.

930 80. Huson DH, Bryant D. Application of phylogenetic networks in evolutionary studies.  
931 *Mol Biol Evol*. 2006; doi: 10.1093/molbev/msj030.

932 81. Revell LJ. phytools: an R package for phylogenetic comparative biology (and other  
933 things). *Methods Ecol Evol*. 2012; doi: 10.1111/j.2041-210X.2011.00169.x.

934 82. Mistry J, Chuguransky S, Williams L, Qureshi M, Salazar GA, Sonnhammer ELL,  
935 et al.. Pfam: The protein families database in 2021. *Nucleic Acids Res*. 2021; doi:  
936 10.1093/nar/gkaa913.

937 83. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7:  
938 improvements in performance and usability. *Mol Biol Evol*. 2013; doi:  
939 10.1093/molbev/mst010.

940 84. Altschul S. Gapped BLAST and PSI-BLAST: a new generation of protein database  
941 search programs. *Nucleic Acids Res*. 1997; doi: 10.1093/nar/25.17.3389.

942

943

944 **Table 1.** *Hortaea werneckii* strains analysed in this study.

Culture collection strain number	Present study number*	Isolation habitat	Sampling site location
EXF-9	1	brine	Ebre Delta salterns, Spain
EXF-12	2	brine	Santa Pola salterns, Spain
EXF-15	3	brine	Santa Pola salterns, Spain
EXF-20	4	brine	Santa Pola salterns, Spain
EXF-152	5	brine	Sečovlje salterns, Slovenia
EXF-153, EXF-2781	6	brine	Sečovlje salterns, Slovenia
EXF-154	7	brine	Sečovlje salterns, Slovenia
EXF-156, CBS 116.90	8	eye infection of aquarium <i>Spondyliosoma cantharus</i>	unknown
EXF-157, CBS 115.90	9	kidney of <i>Bufo granulosis</i>	Brazil
EXF-161, EXF-2689, CBS 706.76	10	leaf of <i>Rhizophora mangle</i>	Senegal
EXF-166, CBS 100496	11	sea water-sprayed marble	Delos, Greece
EXF-177, CBS 705.76	12	<i>tinea nigra</i>	France
EXF-241	13	brine	Sečovlje salterns, Slovenia
EXF-269, EXF-108	14	brine	Santa Pola salterns, Spain
EXF-291	15	brine	Sečovlje salterns, Slovenia
EXF-561	16	brine	Namibia, salterns at the Atlantic coast
EXF-2515	17	brine	salterns, Puerto Rico
EXF-2516	18	brine	salterns, Puerto Rico
EXF-2683, CBS 117.90	19	salted fish, <i>Osteoglossum bicirrhosum</i>	Brazil
EXF-2685	20	brine	Sečovlje salterns, Slovenia
EXF-2783	21	brine	Sečovlje salterns, Slovenia
EXF-2785	22	brine	Sečovlje salterns, Slovenia
EXF-3845	23	brine	Candelaria salterns, Puerto Rico
EXF-3846	24	brine	Candelaria salterns, Puerto Rico
EXF-4716	25	brine bait	Sečovlje salterns, Slovenia
EXF-6274	26	brine	Sečovlje salterns, Slovenia
EXF-6652	27	spider web in a cave close to the ocean	Atacama, Chile
EXF-6663	28	spider web in a cave close to the ocean	Atacama, Chile
EXF-8170	29	brine	Sečovlje salterns, Slovenia
EXF-8422	30	biofilm from cheese factory brine	Celje, Slovenia
EXF-10304	31	brine	Sečovlje salterns, Slovenia
EXF-10508	32	sea water, depth 25 m	Italy
EXF-10509	33	sea water, depth 200 m	Italy
EXF-10510	34	sea water, depth 94 m	Italy
EXF-10511	35	sea water, depth 25 m	Italy
EXF-10512	36	sea water, depth 25 m	Italy
EXF-10816	37	bittern after halite precipitation	Sečovlje salterns, Slovenia
EXF-10819	38	bittern after halite precipitation	Sečovlje salterns, Slovenia
EXF-10820	39	bittern after halite precipitation	Sečovlje salterns, Slovenia
EXF-10843	40	brine	Sečovlje salterns, Slovenia
EXF-10904	41	bittern after halite precipitation	Sečovlje salterns, Slovenia
EXF-10907	42	bittern after halite precipitation	Sečovlje salterns, Slovenia
EXF-10919	43	bittern after halite precipitation	Sečovlje salterns, Slovenia
EXF-10958	44	bittern after halite precipitation	Sečovlje salterns, Slovenia
EXF-10974	45	brine	Sečovlje salterns, Slovenia
EXF-11540	46	sand in a cave close to the ocean	Atacama, Chile
EXF-11650	47	sand in a cave close to the ocean	Atacama, Chile
EXF-11651	48	sand in a cave close to the ocean	Atacama, Chile
EXF-12619	49	coral or deep sea	China
EXF-12620	50	coral or deep sea	China
EXF-14591, CMF-020	51	plankton tow	Vineyard Sound, USA
EXF-14592, AMF 061	52	plankton tow	Vineyard Sound, USA
EXF-225**	53	malt extract medium, 25% NaCl (w/v)	long-term experimental evolution
EXF-14590, MSW 12-1B	54	marine	List on Sylt, Germany

EXF-2000	A***	brine	Sečovlje salterns, Slovenia
EXF-120	B	brine	Santa Pola saltpans, Spain
EXF-562	C	Soil on the sea coast	Namibia
EXF-2788	D	brine	Sečovlje salterns, Slovenia
EXF-171	E	keratomycosis	Brazil
EXF-2682	F	<i>trichomyces nigra</i>	Italy
EXF-10513	G	deep sea water	Italy
EXF-151	H	<i>tinea nigra</i>	Portugal
EXF-6651	I	spider web in a cave close to the ocean	Atacama, Chile
EXF-6669	J	spider web in a cave close to the ocean	Atacama, Chile
EXF-6654	K	spider web in a cave close to the ocean	Atacama, Chile
EXF-6656	L	rock wall in a cave close to the ocean	Atacama, Chile

---

945 \* strains 1-54 were sequenced in this study; strains A-L were sequenced and named

946 by Gostinčar et al. [1]

947 \*\* strain EXF-225 after 15 years of repeated subcultivation at 25% NaCl (w/v) 5

948 \*\*\* reference *H. werneckii* genome [34]; naming of strains A-L corresponds to names

949 in Gostinčar et al. [1]

950

951

952 **Table 2.** *Aureobasidium melanogenum* strains analysed in this study.

Culture collection strain number	Present study number*	Isolation habitat	Sampling site location
EXF-924	1	ponds on sea ice	Svalbard, Norway
EXF-926	2	surface glacial ice	Svalbard, Norway
EXF-3233	3	deep sea (4500 m b.s.l.)	Japan
EXF-3371	4	soil	Thailand
EXF-3378	5	public fountain	Thailand
EXF-3397	6	endoperitoneal fluid	Greece
EXF-4450	8	Iskra factory	Slovenia
EXF-5590	9	dishwasher rubber seal	Slovenia
EXF-6171	10	glacial ice	Argentina
EXF-7932	11	metal drain on the kitchen sink	Sweden
EXF-7946	12	kitchen metal holder for washed dishes	Sweden
EXF-8016	13	bathroom faucet and sink contact	Sweden
EXF-8022	14	refrigerator inner surface	Sweden
EXF-8044	15	kitchen metal holder for washed dishes	Sweden
EXF-8258	16	well water	Slovenia
EXF-9877	17	tap water	Slovenia
EXF-11403	18	refrigerator inner surface	Sweden
EXF-8492	19	well water	Slovenia
EXF-8678	20	well water	Slovenia
EXF-8689	21	well water	Slovenia
EXF-8695	22	well water	Slovenia
EXF-8702	23	well water	Slovenia
EXF-8986	24	fango mud from Sečovlje salterns	Slovenia
EXF-9262	25	rubber on kitchen drain	Slovenia
EXF-9470	26	kitchen counter above dishwasher	Slovenia
EXF-9272	27	kitchen strainer basket	Slovenia
EXF-9298	28	plastic mesh on kitchen drain	Slovenia
EXF-9304	29	kitchen strainer basket	Slovenia
EXF-9313	30	kitchen sink	Slovenia
EXF-9454	31	tap water	Slovenia
EXF-9484	32	kitchen counter above dishwasher	Slovenia
EXF-9887	33	tap water	Slovenia
EXF-9516	34	kitchen sink drain	Slovenia
EXF-9539	35	kitchen strainer basket	Slovenia
EXF-9540	36	dishwasher door	Slovenia
EXF-10064	37	tap water	Slovenia
EXF-11060	38	ceiling surface	Slovenia
EXF-9875	39	tap water	Slovenia
EXF-9906	40	<i>Arthrocnemum</i> sp. plant surface from Sečovlje saltern	Slovenia
EXF-9911	41	kitchen sink drain	Slovenia
EXF-9937	42	kitchen sink drain	Slovenia
EXF-10061	43	tap water	Slovenia
EXF-10062	44	tap water	Slovenia
EXF-10066	45	tap water	Slovenia
EXF-10333	46	tap water	Slovenia
EXF-10372	47	air in the National Gallery restoration centre	Slovenia
EXF-10726	48	integument of a male alate ant of <i>Atta sexdens rubropilosa</i>	Brazil
EXF-11028	49	water from the aquarium with <i>Proteus anguinus</i>	Slovenia

953 \* same numbering as in Černoša et al. [36]

954

955 **Table 3.** Statistics for the *H. werneckii* genomes sequenced in this study (strains 1-  
 956 54).

	haploid strains			diploid strains			tetraploid strain
	average	median	SD	average	median	SD	/
Coverage	730	619	464	469	485	177	276
Genome assembly size (Mbp)	26.52	26.19	1.47	49.30	49.22	1.74	94.67
Number of contigs	796	638	421	6885	3806	4457	30312
Contig N50 (kbp)	136	138	28	22	26	14	5
GC content	53.22%	53.33%	0.33%	53.40%	53.40%	0.19%	53.40%
CDS total length (Mbp)	14.56	14.39	0.80	27.02	27.87	1.45	49.02
CDS total length (% of genome)	54.94%	55.27%	1.67%	54.80%	55.59%	2.15%	51.78%
Gene models (n)	9519	9344	665	20417	19240	1709	46596
Exons per gene (average)	2.34	2.34	0.06	2.10	2.20	0.14	1.87
Intron average length (bp)	93.17	93.00	2.53	94.11	94.00	4.73	92.00
Complete BUSCOs *	95.99%	96.00%	0.34%	86.86%	93.40%	10.09%	89.60%
Complete and single-copy BUSCOs	95.83%	95.90%	0.35%	21.33%	16.30%	10.50%	33.20%
Complete and duplicated BUSCOs	0.16%	0.20%	0.06%	65.53%	77.30%	19.20%	56.40%
Fragmented BUSCOs	0.82%	0.75%	0.14%	7.05%	3.20%	5.87%	5.10%
Missing BUSCOs	3.19%	3.10%	0.30%	6.09%	3.30%	4.24%	5.30%
Total SNP density (SNPs per total genome size)	4.04%	4.54%	1.11%	3.44%	3.56%	1.12%	/
Heterozygous SNP density (SNPs per total genome size)	0.01%	0.01%	0.01%	2.46%	2.58%	0.74%	/

957 \* BUSCOs, Benchmarking Universal Single-Copy Orthologues.

958

959 **Figure legends**

960

961 **Figure 1.** Single-nucleotide polymorphism (SNP) diversity of *Hortaea werneckii* (**A, B**)  
962 and *Aureobasidium melanogenum* (**C, D**). (**A, C**) Phylogenetic networks reconstructed  
963 with a Neighbor-Net algorithm from a dissimilarity distance matrix calculated from SNP  
964 data. (**B, D**) Principal component analysis of SNPs. The genomes are represented by  
965 circles.

966

967 **Figure 2.** Linkage disequilibrium (LD) decay in *Hortaea werneckii* (**A**) and  
968 *Aureobasidium melanogenum* (**B**). Squared correlation coefficient ( $r^2$ ) was calculated  
969 for all pairs of non-singleton biallelic loci within the distance of 10 kbp or less and  
970 plotted as a function of the distance between the loci (blue line). The maximum  
971 observed value and its half value are marked with red horizontal dashed lines. A  
972 generalized additive model curve was fitted to the data (black line).

973

974 **Figure 3.** Phylogenies of 50 longest alignable genomic regions of *Hortaea werneckii*  
975 (**A, C**) and *Aureobasidium melanogenum* (**B, D**). The alignable regions were extracted  
976 from the genomes and aligned with SibeliaZ, optimized with Gblocks, manually  
977 inspected and used for phylogeny reconstruction with IQ-TREE and standard model  
978 selection. (**A, B**) Overlay of 50 phylogenies for each species. Numbers on leaf nodes  
979 represent genomes, different sequences from the same genomes are distinguished  
980 with letters added to the genome numbers. (**C, D**) Consensus supernetworks  
981 calculated from 50 phylogenies for each species in SplitsTree.

982

983 **Figure 4.** Hypothesis of the genome evolution and hybridization in *Hortaea werneckii*  
984 (**A**) and *Aureobasidium melanogenum* (**B**). The hypothesis is based on the majority  
985 consensus phylogeny of 50 longest alignable regions per species. Each coloured line  
986 in the central tree represents a haploid genome. The distances between the nodes of  
987 the tree correspond to the distances in an ultrametric majority consensus phylogeny.  
988 Haploid genomes are represented by a single coloured line in the outermost edge of  
989 the tree, diploid genomes are represented by a double coloured line, the only tetraploid  
990 genome is represented by four coloured lines. Around the tree, coloured symbols mark  
991 the continent (inner circle) and habitat (outer circle) from which the strains have been

992 originally isolated. Black lines and numbers in the outermost circle mark the genome/  
993 strain groups presumably originating from the same hybridization event.

994

995 **Figure 5.** Aneuploid regions in *Hortaea werneckii* (A) and *Aureobasidium*  
996 *melanogenum* (B) genomes. Per-nucleotide sequencing depth of regions  
997 corresponding to 50 and 35 longest contigs of *H. werneckii* and *A. melanogenum* were  
998 converted into proportion of the median sequencing depth of each individual genome.  
999 Circles represent an average of this depth in 30 kbp windows. The central horizontal  
1000 line marks the median sequencing depth of the genome. Upper and lower horizontal  
1001 lines mark the expected depth for haploid and triploid regions in an otherwise diploid  
1002 genome. Genomes with at least one putatively aneuploid region are plotted in colour.  
1003 Other genomes are plotted in light grey.

1004

1005 **Supplementary Information**

1006

1007 **Supplemental Table S1.** Statistics of *H. werneckii* genomes sequenced in this study.

1008

1009 **Supplemental Table S2.** Putative HET and HET-C proteins in different strains of *H.*  
1010 *werneckii* and *A. melanogenum*.

1011

1012 **Supplemental Figure S1.** Aneuploid regions in *Hortaea werneckii* genomes. Per-  
1013 nucleotide sequencing depth of regions corresponding to 50 longest contigs of *H.*  
1014 *werneckii* and *A. melanogenum* were converted into proportion of the median  
1015 sequencing depth of each individual genome and plotted in 50 kbp rolling median  
1016 windows (black line). Upper and lower horizontal lines mark the expected depth for  
1017 haploid and triploid regions in an otherwise diploid genome. Putatively aneuploid  
1018 region of increased ploidy are marked with red rectangles.

1019

1020 **Supplemental Figure S2.** Aneuploid regions in *Aureobasidium melanogenum*  
1021 genomes. Per-nucleotide sequencing depth of regions corresponding to 35 longest  
1022 contigs of *A. melanogenum* were converted into proportion of the median sequencing  
1023 depth of each individual genome and plotted in 50 kbp rolling median windows (black  
1024 line). Upper and lower horizontal lines mark the expected depth for haploid and triploid  
1025 regions in an otherwise diploid genome. Putatively aneuploid region of increased ploidy  
1026 are marked with red rectangles.

1027

1028 **Supplemental Figure S3.** Heterozygosity in diploid *Hortaea werneckii* and  
1029 *Aureobasidium melanogenum* genomes. Levels of heterozygosity (black lines) and  
1030 sequencing depth (purple lines) were expressed as proportions of median  
1031 heterozygosity and sequencing depth of each individual genome. The values were  
1032 plotted in 25 kbp windows across regions corresponding to 50 and 35 longest contigs  
1033 of *H. werneckii* and *A. melanogenum*, respectively. Diploid regions (i.e. with  
1034 sequencing depth similar to the rest of the diploid genome) with extensive loss of  
1035 heterozygosity are marked with red rectangles.

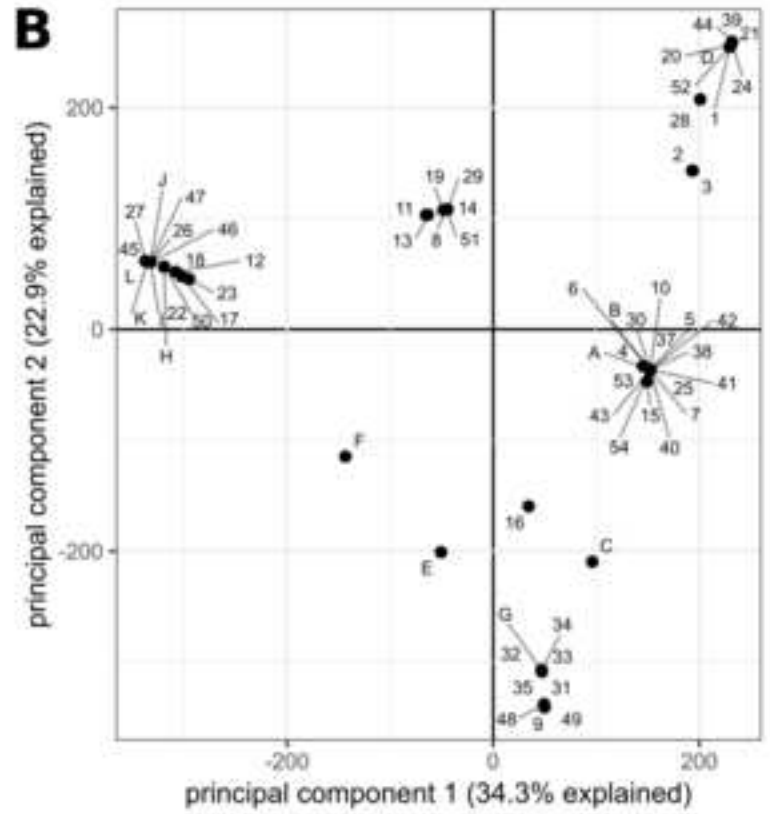
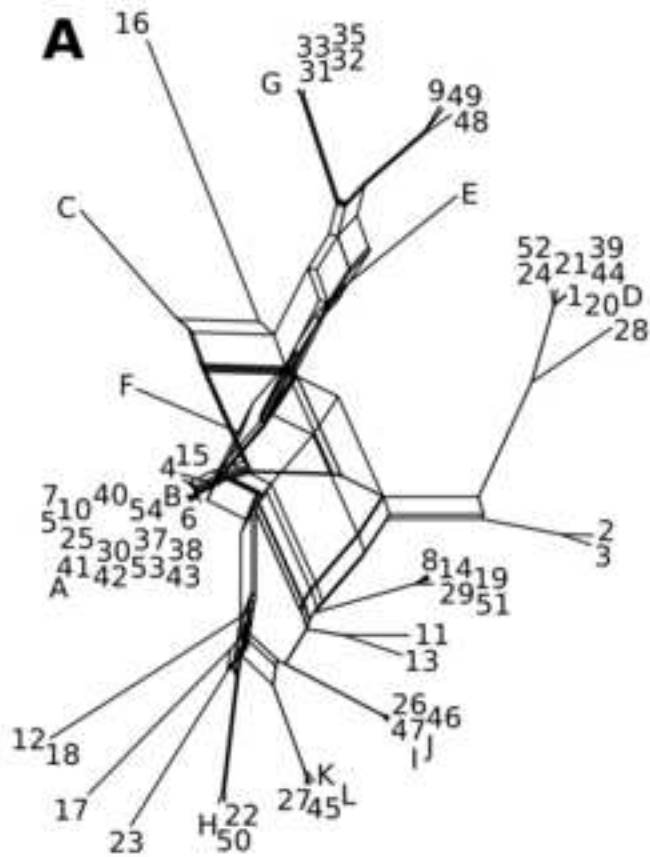
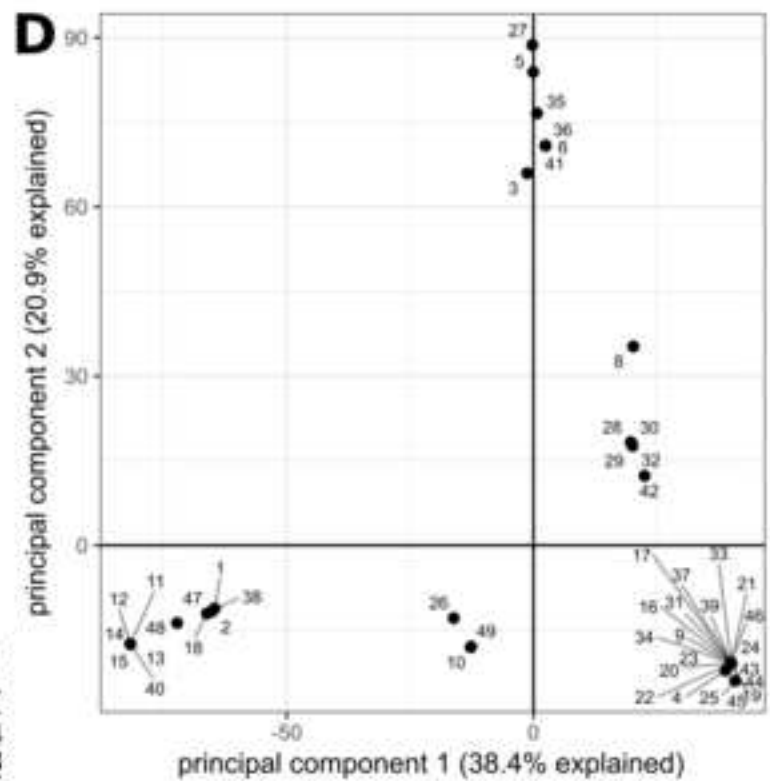
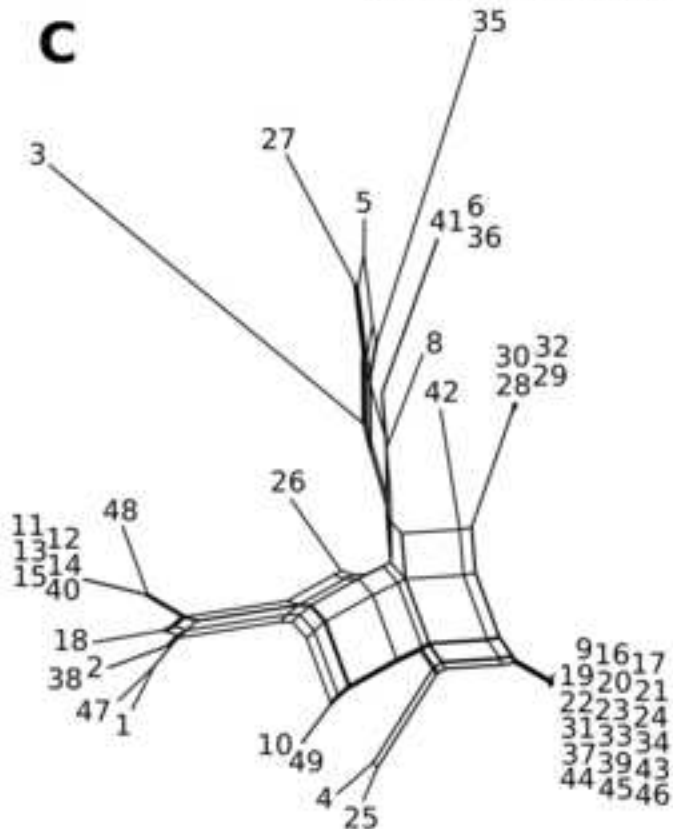
1036

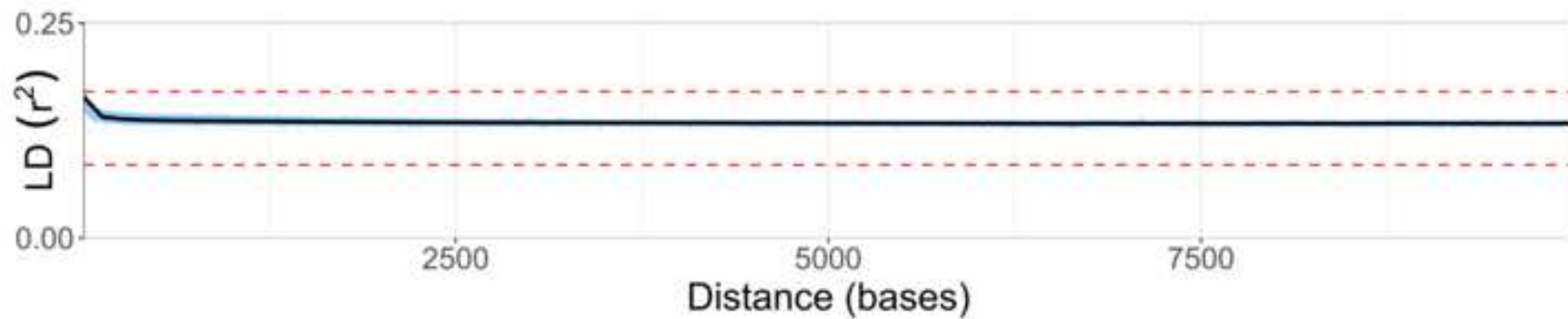
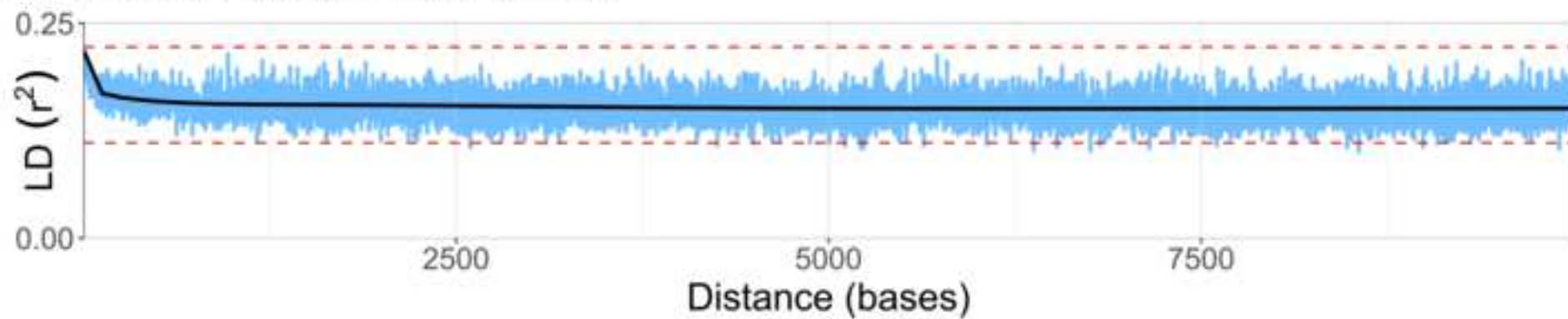
1037 **Supplemental Figure S4.** Putative mating type loci of *Hortaea werneckii*. Phylogenies  
1038 and presence/absence visualisation of MAT1-1 and MAT1-2 homologues, as well as  
1039 an annotated alignment of the whole putative mating locus and its flanking regions.

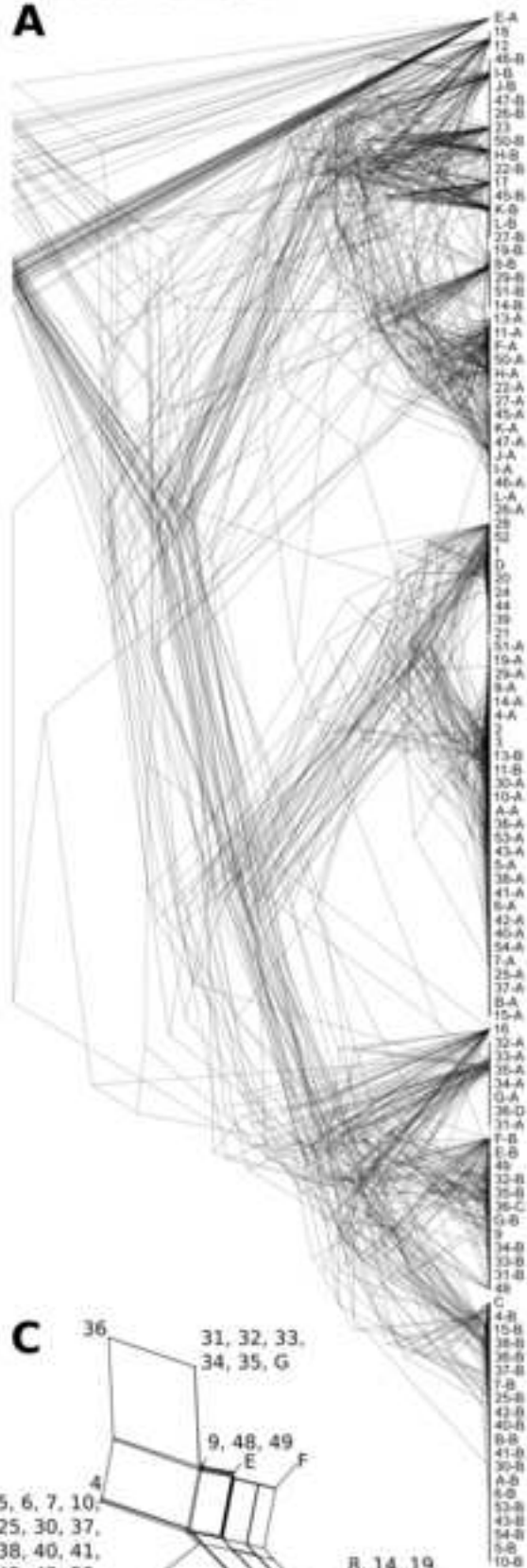
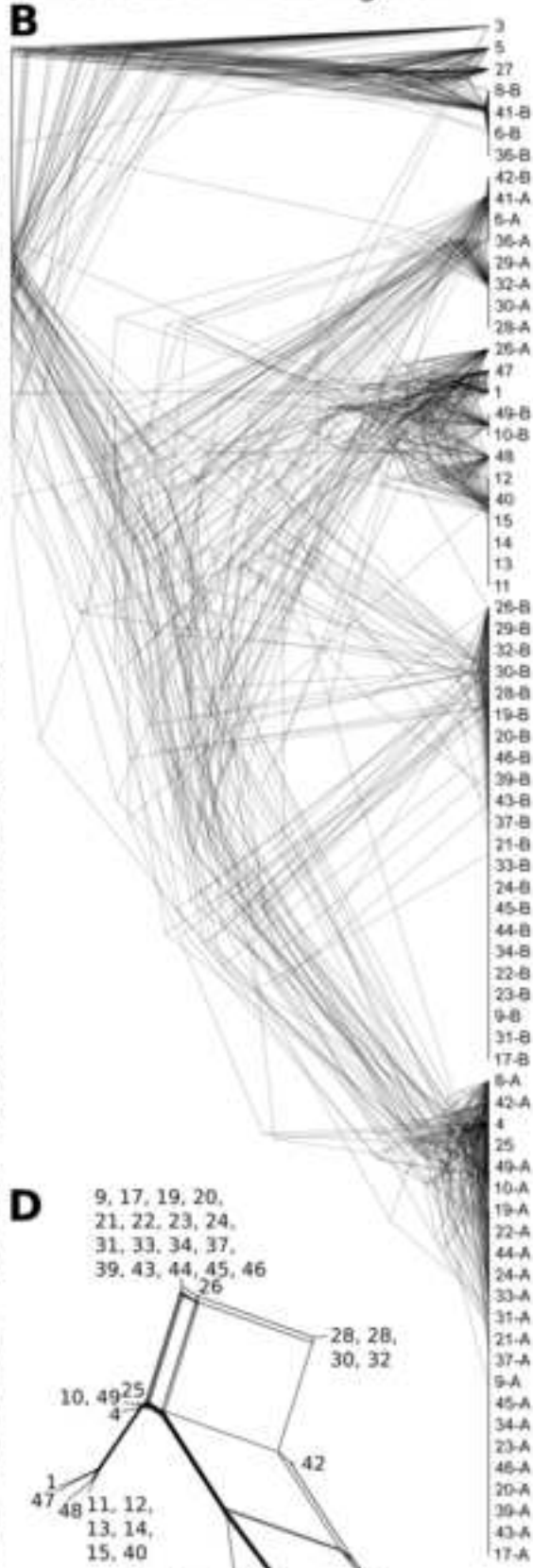
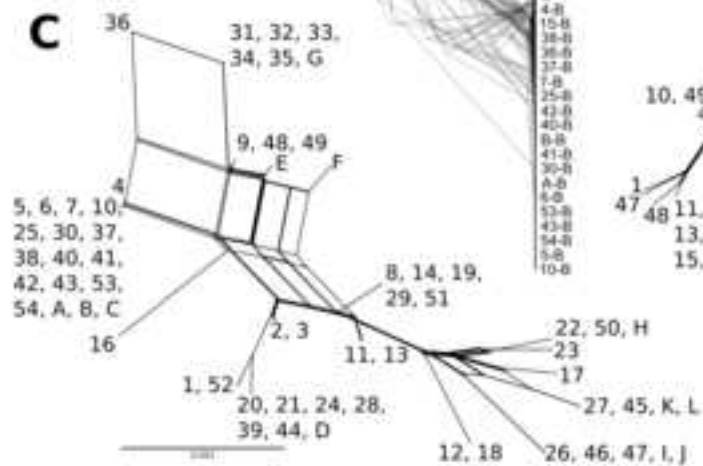
1040

1041 **Supplemental Figure S5.** Putative mating type loci of *Aureobasidium melanogenum*.  
1042 Phylogenies and presence/absence visualisation of MAT1-1 and MAT1-2 homologues,  
1043 as well as an annotated alignment of the whole putative mating locus and its flanking  
1044 regions.

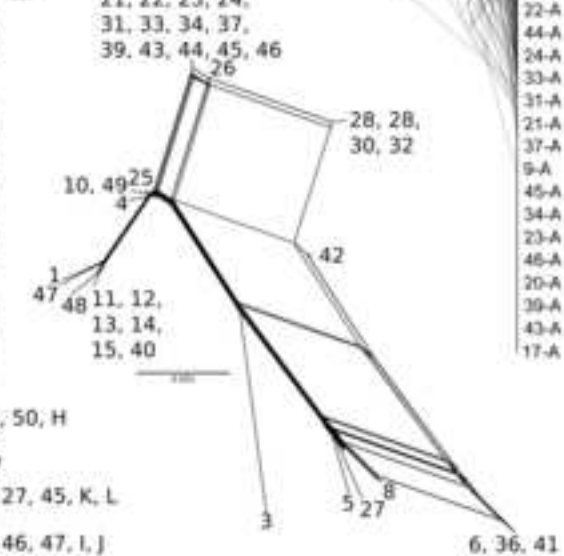


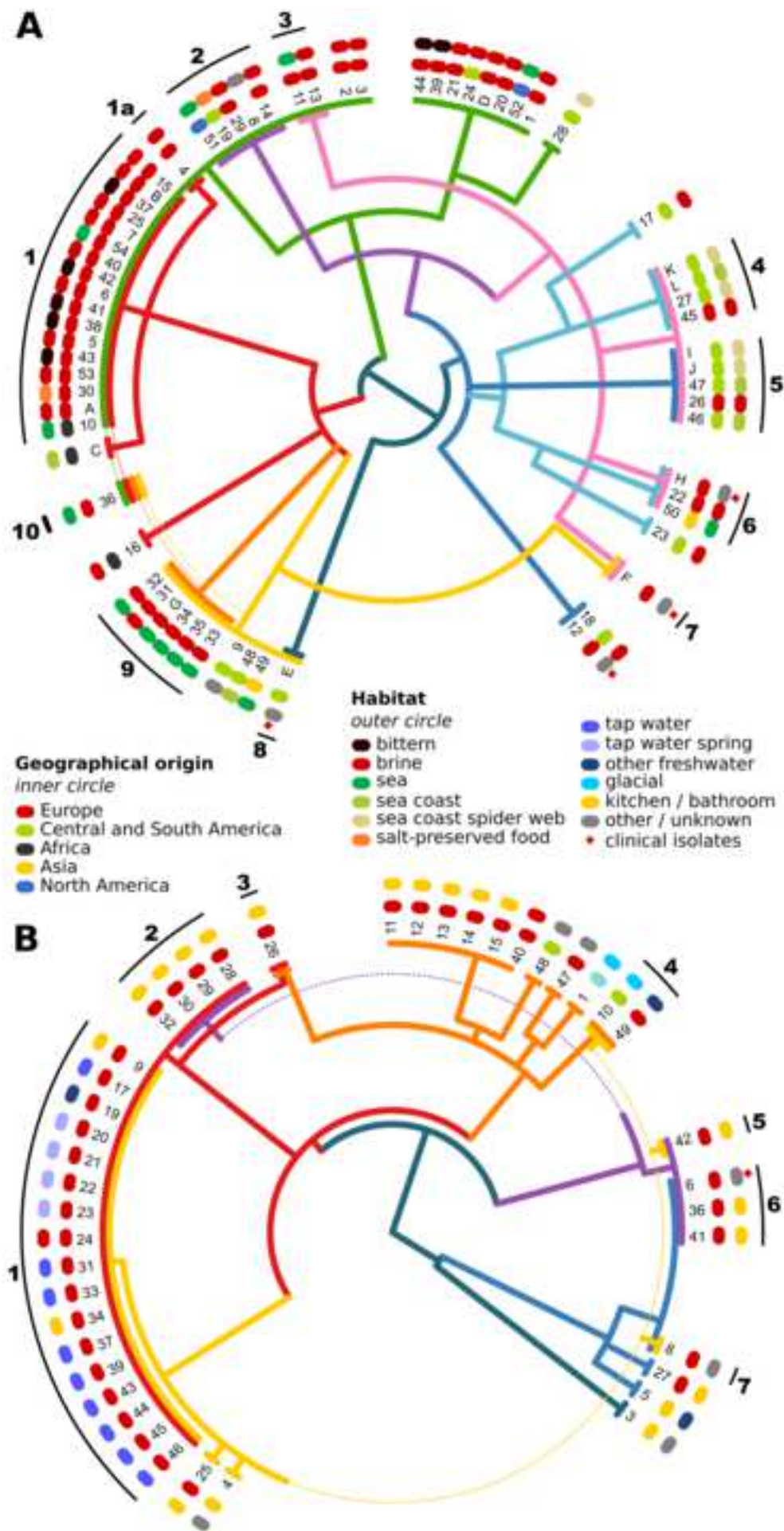
*Hortaea werneckii**Aureobasidium melanogenum*

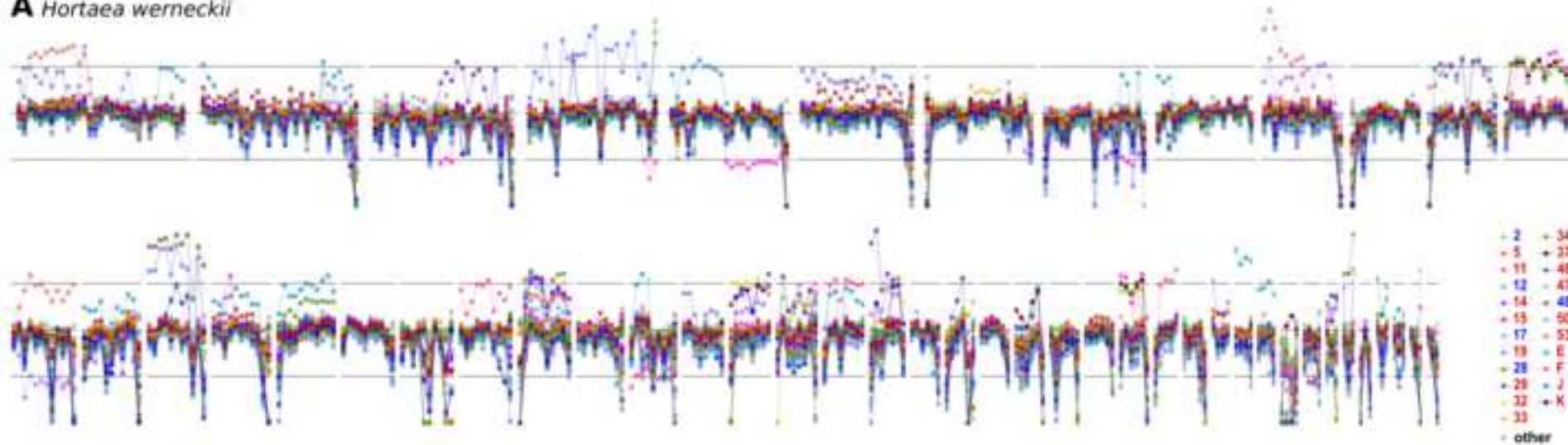
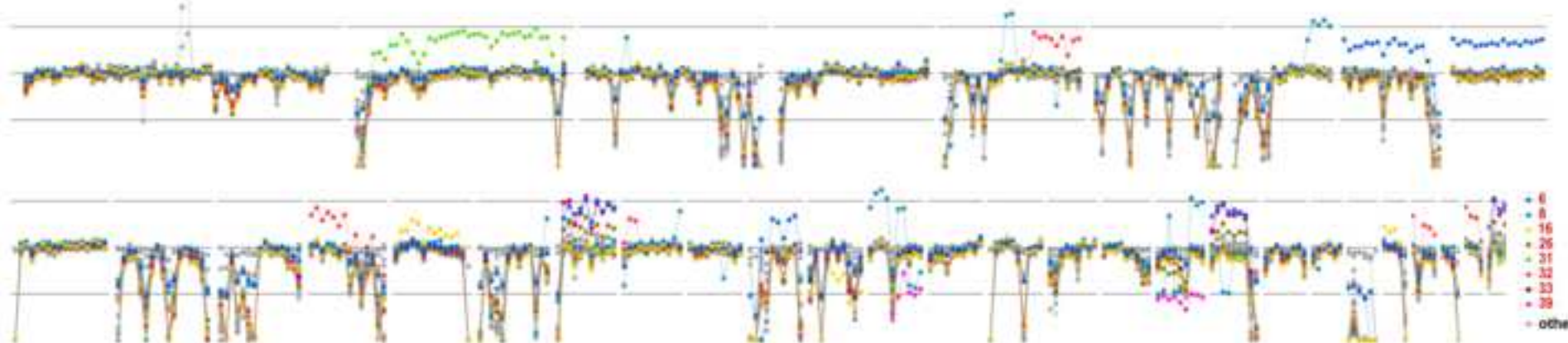
**A** *Hortaea werneckii***B** *Aureobasidium melanogenum*

*Hortaea werneckii***A***Aureobasidium melanogenum***B****C****D**

9, 17, 19, 20,  
21, 22, 23, 24,  
31, 33, 34, 37,  
39, 43, 44, 45, 46

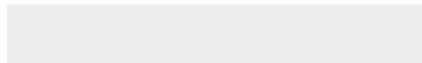




**A** *Hortaea werneckii***B** *Aureobasidium melanogenum*




Click here to access/download  
**Supplementary Material**  
SupplementalTableS1.xlsx



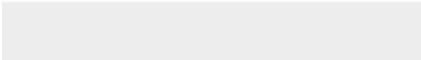



Click here to access/download  
**Supplementary Material**  
SupplementalTableS2.xlsx







Click here to access/download  
**Supplementary Material**  
FigS1.png



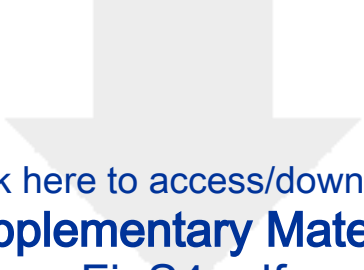




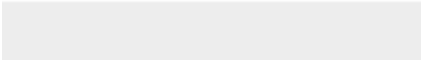

Click here to access/download  
**Supplementary Material**  
FigS2.png




Click here to access/download  
**Supplementary Material**  
FigS3.png



Click here to access/download  
**Supplementary Material**  
FigS4.pdf





Click here to access/download  
**Supplementary Material**  
FigS5.pdf