Factors enforcing the species boundary between the human pathogens *Cryptococcus neoformans* and *Cryptococcus deneoformans*

Shelby J. Priest[1], Marco A. Coelho[1], Verónica Mixão[2,3], Shelly Clancey[1], Yitong Xu[4], Sheng Sun[1], Toni Gabaldón[2,3,5], and Joseph Heitman[1]*

[1] Department of Molecular Genetics and Microbiology, Duke University Medical Center, Durham, NC, USA

[2] Life Sciences Department, Barcelona Supercomputing Center, Barcelona, Spain

[3] Institute for Research in Biomedicine, Barcelona Institute of Science and Technology, Barcelona, Spain

[4] Program in Cell and Molecular Biology, Duke University Medical Center, Durham, NC, USA

[5] Catalan Institution for Research and Advanced Studies (ICREA), Barcelona, Spain

Short title: Factors enforcing the *Cryptococcus* species boundary

* Corresponding author

Email: heitm001@duke.edu (JH)

Abstract

Hybridization has resulted in the origin and variation in extant species, and hybrids continue to arise despite pre- and post-zygotic barriers that limit their formation and evolutionary success. One important system that maintains species boundaries in prokaryotes and eukaryotes is the mismatch repair pathway, which blocks homeologous recombination between divergent DNA sequences. Previous studies illuminated the role of the mismatch repair component Msh2 in blocking genetic recombination between divergent DNA during meiosis in yeast and during conjugational and transductional crosses in bacteria. Loss of Msh2 resulted in increased interspecific genetic recombination in bacterial and yeast models, and increased viability of progeny derived from yeast hybrid crosses. Hybrid isolates of two pathogenic fungal *Cryptococcus* species, *Cryptococcus neoformans* and *Cryptococcus deneoformans*, are isolated regularly from both clinical and environmental sources. In the present study, we sought to determine if loss of Msh2 would relax the species boundary between *C. neoformans* and *C. deneoformans*. We found that crosses between these two species in which both parents lack Msh2 produced hybrid progeny with increased viability and high levels of aneuploidy. Whole-genome sequencing revealed few instances of recombination among hybrid progeny and did not identify increased levels of recombination in progeny derived from parents lacking Msh2. Several hybrid progeny from bilateral *msh2*Δ crosses produced structures associated with sexual reproduction when incubated alone on nutrient-rich medium in light, a novel transgressive phenotype in *Cryptococcus*. These findings represent a unique, unexpected case where rendering the mismatch repair system defective did not result in increased meiotic recombination across a species boundary. This suggests that alternative pathways limit meiotic recombination between

53 homeologous DNA and enforce species boundaries in the basidiomycete *Cryptococcus* species
54 and possibly in others as well.
55

56 ## Author summary

57 Several mechanisms enforce species boundaries by either preventing the formation of
58 hybrids, known as pre-zygotic barriers, or preventing the viability and fecundity of hybrids,
59 known as post-zygotic barriers. Despite these barriers, interspecies hybrids form at an
60 appreciable frequency, such as hybrid isolates of the human fungal pathogenic species,
61 *Cryptococcus neoformans* and *Cryptococcus deneoformans*, which are regularly isolated from
62 both clinical and environmental sources. *C. neoformans* x *C. deneoformans* hybrids are typically
63 highly aneuploid, sterile, and display phenotypes intermediate to those of either parent, although
64 self-fertile isolates and transgressive phenotypes have been observed. One important mechanism
65 known to enforce species boundaries or lead to incipient speciation is the DNA mismatch repair
66 system, which blocks recombination between divergent DNA sequences during meiosis. The aim
67 of this study was to determine if genetically deleting the DNA mismatch repair component Msh2
68 would relax the species boundary between *C. neoformans* and *C. deneoformans*. Progeny derived
69 from *C. neoformans* x *C. deneoformans* crosses in which both parental strains lacked Msh2 had
70 higher viability, several progeny displayed a novel transgressive phenotype, and unlike previous
71 studies in *Saccharomyces*, these *Cryptococcus* hybrid progeny had higher levels of aneuploidy
72 and no observable increase in meiotic recombination at the whole-genome level.
73

74 ## Introduction

75 The mixing of species through sexual reproduction can result in the production of hybrid
76 offspring. While hybridization can have beneficial consequences in some cases (e.g. hybrid vigor
77 and the emergence of novel hybrid species), sexual reproduction between diverging lineages or
78 different species is typically deleterious and results in hybrid progeny with reduced fitness or
79 sterility [1-3]. Thus, mechanisms preventing such events, such as pre- and post-zygotic
80 reproductive barriers, tend to be favored by natural selection. Pre-zygotic species barriers block
81 the formation of a hybrid zygote and include spatial, temporal, behavioral, mechanical, and
82 gametic isolation [4]. In the event that a hybrid zygote forms, several post-zygotic barriers exist
83 that inhibit the viability or fecundity of the hybrid, including 1) gross chromosomal
84 rearrangements that prevent effective meiotic recombination, 2) Bateson-Dobzhansky-Muller
85 incompatibilities in which interactions between nuclear elements or between mitochondrial and
86 nuclear factors are detrimental or lethal, and finally, 3) the mismatch repair (MMR) pathway,
87 which has an important role in blocking meiotic recombination between diverged DNA
88 sequences.
89 The MMR pathway was first identified in prokaryotes as a mechanism to repair
90 replication errors or damage-induced mismatches in DNA [5]. The MMR pathway is highly
91 conserved, playing similar roles in unicellular eukaryotes, such as *Saccharomyces* species, and in
92 multicellular eukaryotes, including humans [6]. Prokaryotic and eukaryotic cells lacking
93 functional MMR components typically have increased mutation rates and therefore display a
94 hypermutator phenotype [7]. The MMR pathway plays an additional role in maintaining species
95 boundaries by inhibiting homeologous chromosome pairing and subsequent recombination
96 during meiosis. The inability of chromosomes to properly and stably pair, and subsequently
97 undergo recombination, can lead to chromosome nondisjunction during meiosis, resulting in high

98 frequencies of aneuploidy and ultimately a decrease in the viability of progeny. Rayssiguier et al.
99 demonstrated that mutation of the MutL, MutS, or MutH MMR components relaxed the species
100 boundary between *Escherichia coli* and *Salmonella typhimurium*, two bacterial species whose
101 genomes are 20% divergent, such that recombination during conjugational and transductional
102 crosses increased up to 1,000-fold [8]. The involvement of MMR in recombination is conserved
103 and has also been shown to enforce the species boundary between *Saccharomyces cerevisiae* and
104 the closely related species *Saccharomyces paradoxus*, whose genomes are ~11% divergent in
105 coding regions and ~20% divergent in non-coding regions. Mutants lacking the eukaryotic MutS
106 homolog, Msh2, or the MutL homolog, Pms1, produce hybrid *S. cerevisiae* x *S. paradoxus*
107 progeny with higher rates of viability as well as approximately ten-fold increased frequencies of
108 meiotic recombination [9,10].

109 Despite the numerous pre- and post-zygotic species barriers and the robust fitness defects
110 associated with hybrids, hybridization occurs at an appreciable frequency. For example, the ursid
111 hybrids produced by matings between grizzly and polar bears are frequently observed in nature
112 and genomic analyses suggest ancient hybridization events between brown and polar bears, as
113 the mitochondrial genomes of brown bears are derived from polar bears [11]. Ligers, the hybrid
114 progeny of a male lion and female tiger, and mules, the hybrid progeny of a male donkey and
115 female horse, are examples of hybrids that normally only occur through human intervention.
116 Ligers are bred for their size, as they are larger than either parent (an instance of hybrid vigor),
117 while mules are bred for their endurance, docile demeanor, and intelligence [12]. In an exciting
118 recent study, researchers have even witnessed the origin and monitored the evolution of a novel
119 finch species that arose through a hybridization event in the Galapagos Islands [13].

120 Hybridization has impacted recent human evolution too, with several modern human populations
121 having introgressed genomic regions from Neanderthals or Denisovans [14].
122 Hybridization is also important in many microbial pathogens and model organisms. For
123 instance, the diploid model organism *S. cerevisiae* is thought to have arisen following a whole-
124 genome duplication that was a direct consequence of interspecies hybridization [15-17]. An
125 instance of relatively recent hybridization is also thought to have led to the emergence of the
126 novel widespread fungal plant pathogen species *Zymoseptoria pseudotritici*, originating from
127 fusion between two diverged haploids followed by mitosis and meiosis to generate a
128 recombinant haploid F_1 hybrid [18]. In recent years, the hybrid nature of several emerging
129 human opportunistic pathogens has been uncovered [19,20], suggesting that hybridization might
130 be a mechanism underlying the emergence of novel pathogens [21].
131 Several *Cryptococcus* species are microbial human fungal pathogens and are responsible
132 for over 200,000 infections in both immunocompromised and immunocompetent individuals
133 annually [22]. Cryptococcal infections are associated with high mortality rates and occur
134 globally. The pathogenic *Cryptococcus* species complex was recently divided into seven distinct
135 species based on phylogenetic analysis of multiple genomic loci [23]. The pathogenic
136 *Cryptococcus* species complex is comprised of two phylogenetically-based subgroups, the
137 *Cryptococcus neoformans* species complex and the *Cryptococcus gattii* species complex, which
138 consist of two and six species, respectively. The present study focuses on the two members of the
139 *C. neoformans* species complex: *C. neoformans* and *C. deneoformans* [23,24].
140 Previously, *C. neoformans* and *C. deneoformans* were considered to be a single species
141 with two varieties and two serotypes: *C. neoformans* var. *grubii* (serotype A) and *C. neoformans*
142 var. *neoformans* (serotype D) [23]. These two species are the most commonly isolated

143 *Cryptococcus* species from clinical and environmental settings and both species serve as model
144 pathogenic eukaryotic organisms [25]. Delineating *C. neoformans* and *C. deneoformans* as two
145 distinct species is in relative agreement with each of the three most common species delimitation
146 approaches: the phylogenetic, morphological, and biological species concepts [26-29]. In line
147 with the phylogenetic definition, there is clear genetic evidence separating these two groups and
148 molecular phylogenetics along with whole-genome sequencing suggests they diverged ~18
149 million years ago [30-33]. There are also several phenotypes that differentiate *C. neoformans* and
150 *C. deneoformans* as distinct species, such as differences in thermotolerance, capsular
151 agglutination reactions, morphology during murine infection, and human disease manifestations
152 and outcomes [34-38].
153 Despite the genetic and phenotypic differences between *C. neoformans* and *C.*
154 *deneoformans*, these two groups produce hybrids in the laboratory and in nature. *C. neoformans*
155 x *C. deneoformans* hybrids, also known as AD hybrids, make up ~7.5% of environmental
156 isolates and up to 30% of clinical isolates in Europe and North America [39-42]. Spores
157 produced by genetic crosses between *C. neoformans* and *C. deneoformans* isolates are known to
158 have poor germination frequencies (~5%), and are typically highly aneuploid or heterozygous
159 diploids [43]. This poor viability is likely due to a combination of gross chromosomal
160 rearrangements between the two parental genomes along with ~15% sequence divergence
161 between the parental species [31,33], leading to a compromised meiosis that produces genetically
162 imbalanced meiotic progeny. *C. neoformans* x *C. deneoformans* hybrids also have unstable
163 karyotypes, can be self-fertile, and display phenotypes intermediate between either parent,
164 although instances of transgressive phenotypes and hybrid vigor have been observed [43-46].
165 Several studies of *C. neoformans* x *C. deneoformans* hybrid genomes have utilized restriction

166 fragment length polymorphism analysis or PCR with sequence-specific primers to assess their
167 genomes. Through these methods it has been demonstrated that although *C. neoformans* x *C.*
168 *deneoformans* hybrids are heterozygous at most loci, some chromosomes seem to be
169 recombinant, indicative of potential mitotic or meiotic recombination [45,47-49].

170 In the present study, we generated *C. neoformans* and *C. deneoformans* strains lacking
171 Msh2 to determine if loss of MMR relaxed the boundary between these two species. As
172 expected, *C. neoformans* and *C. deneoformans msh2*Δ mutants displayed hypermutator
173 phenotypes. Hybrid progeny derived from genetic crosses in which both parental strains lacked
174 Msh2 displayed increased germination frequencies compared to wild-type crosses and also
175 exhibited phenotypes and genotypes in accordance with previous findings [43]. Several instances
176 of genetic recombination were observed in hybrid progeny derived from both wild-type *C.*
177 *neoformans* x *C. deneoformans* crosses and *msh2*Δ mutant crosses, although interestingly,
178 increased frequencies of meiotic recombination were not observed in hybrid progeny derived
179 from crosses involving *msh2*Δ mutants. Additionally, lower rates of loss of heterozygosity
180 (LOH) and higher rates of aneuploidy were observed in hybrid progeny from *msh2*Δ mutant
181 crosses. These results suggest other pathways and mechanisms besides MMR may enforce
182 species boundaries in *Cryptococcus*.

183

184 ## Results

185 ### *C. deneoformans msh2*Δ mutants are hypermutators

186 To assess the role of the MMR pathway in maintaining species boundaries in
187 *Cryptococcus*, we first determined if Msh2 plays a similar role in DNA MMR in *C.*
188 *deneoformans* as in other fungi. Deletion mutants lacking *MSH2* were generated via biolistic

189 transformation and homologous recombination in the *C. deneoformans* JEC20**a** genetic
190 background. A schematic depicting the deletion is shown in Figure S1A. After transformation,
191 mutants were selected on medium containing nourseothricin, and PCR was employed to confirm
192 that the deletion allele had replaced the wild-type *MSH2* allele at its endogenous locus (Figure
193 S1B).

194 Following isolation and confirmation of the desired *msh2*Δ mutants, we determined if the
195 mutants displayed hypermutator phenotypes similar to those observed in *msh2* mutants of other
196 fungi [50,51]. To assess mutation rate, fluctuation assays were performed on either YPD medium
197 supplemented with a combination of rapamycin and FK506 at 37°C (where calcineurin, the
198 target of FKBP12-FK506, is essential) or YNB medium supplemented with 5-fluoroorotic acid
199 (5-FOA) at 30°C. In this analysis, the parental strain of the genetic deletion mutants, JEC20**a**,
200 served as the negative control, and a *C. neoformans msh2*Δ mutant in the KN99α genetic
201 background from the *Cryptococcus* deletion mutant collection [52] served as the positive control.
202 Although fluctuation assays do not provide genome-wide mutation rates, the assays allow us to
203 compare the mutation rates at two unique coding loci of the JEC20**a** *msh2*Δ mutants to those of
204 the JEC20**a** parental strain and the KN99α *msh2*Δ positive control. On the rapamycin and
205 FK506 antifungal drug combination, the *msh2*Δ-*1*, *msh2*Δ-*2*, *msh2*Δ-*3,* and *msh2*Δ-*4* mutants
206 exhibited hypermutator phenotypes with significantly higher mutation rates (1.02×10^{-6}, $1.30 \times$
207 10^{-6}, 1.16×10^{-6}, and 1.16×10^{-6} mutations per cell per generation, respectively) than the parental
208 JEC20**a** strain (8.59×10^{-8} mutations per cell per generation) (Figure 1A).

209 Insertion/deletion mutations (INDELs) in homopolymeric nucleotide runs are a hallmark
210 mutation pattern of *msh2*Δ mutants [50,53]. To determine if the four independent JEC20**a** *msh2*Δ
211 mutants also displayed similar mutation patterns, *FRR1*, the gene encoding FKBP12 (the

212 common target of FK506 and rapamycin), was PCR amplified and mutations were identified
213 through Sanger sequencing. This analysis revealed single base pair INDELs in homopolymeric
214 nucleotide runs within *FRR1* in ten of ten independent colonies analyzed for of the *msh2*Δ-*1*,
215 *msh2*Δ-*2*, and *msh2*Δ-*4* strains and nine of ten independent colonies analyzed for the *msh2*Δ-*3*
216 strain (Figure 1B). This is significantly different from the frequency of 1-bp INDEL mutations
217 observed in the wild-type JEC20**a** parental strain, in which only one of nine colonies had a 1-bp
218 INDEL mutation in a homopolymer run ($p<0.001$, Fisher's exact test). As anticipated, the
219 KN99α *msh2*Δ mutant produced resistant colonies with 1-bp INDEL mutations in homopolymer
220 runs in *FRR1* in all (10/10) colonies analyzed (Figure 1B).
221 Mutation rates of three of the four independent *msh2*Δ mutants (*msh2*Δ-*2*, *msh2*Δ-*3*, and
222 *msh2*Δ-*4*) were also significantly higher than the mutation rate of the parental strain, JEC20**a**, on
223 5-FOA (Figure 1C). Interestingly, the *msh2*Δ-*1* mutant failed to produce any 5-FOA resistant
224 isolates during the fluctuation assay, which is unusual as even JEC20**a** yielded 5-FOA resistant
225 colonies in this assay despite its relatively low mutation rate. The inability of the *msh2*Δ-*1*
226 mutant to produce 5-FOA-resistant colonies, which is typically mediated by loss-of-function
227 mutations in *URA5* or *URA3*, was explained by an INDEL in a homopolymeric nucleotide run in
228 *FUR1*, which encodes a uracil phosphoribosyl transferase involved in the pyrimidine salvage
229 pathway (deletion of a thymine at position 570229 on Chr5 (AE017345.1), GeneID: CNE02100
230 (AE017345.1:569,042-570,826)). Mutations in components of the pyrimidine *de novo*
231 biosynthesis pathway, such as *URA3*, have been demonstrated to be synthetically lethal with
232 *FUR1* mutations in yeast [54,55]. Independent overnight cultures of each of the independent
233 JEC20**a** *msh2*Δ mutants, a KN99α *fur1*Δ mutant, the JEC20**a** parental strain (negative control),
234 and a KN99α *msh2*Δ mutant (positive control) were swabbed on YNB plates supplemented with

235 5-FOA to illustrate that both JEC20**a** *msh2*Δ-*1* and KN99α *fur1*Δ mutants are incapable of
236 producing 5-FOA resistant colonies (Figure 1D). Furthermore, *fur1*Δ mutants are also known to
237 be resistant to the antifungal drug 5-fluorouridine (5FU) and the clinically relevant antifungal
238 drug 5-fluorocytosine (5FC) [56]. Papillation assays with the same strains used on 5-FOA were
239 also conducted on YNB plates supplemented with 5FC or 5FU. As anticipated, the JEC20**a**
240 *msh2*Δ-*1* and the KN99α *fur1*Δ mutant displayed congruent growth phenotypes and were
241 resistant to both 5FC and 5FU, unlike the parental JEC20**a**, the KN99α *msh2*Δ strain, or any of
242 the other JEC20**a** *msh2*Δ mutants (Figure 1E and Figure 1F).
243
244 **Progeny from *msh2*Δ hybrid genetic crosses display increased viability**
245 Interspecific crosses involving MMR-deficient *Saccharomyces* strains produce progeny
246 with increased viability [9,10]. To determine if a similar increase in viability would be observed
247 in *Cryptococcus*, genetic crosses were conducted between *C. neoformans* and *C. deneoformans*
248 wild-type strains as well as corresponding *msh2*Δ mutants. A cross between H99α (a laboratory
249 standard *C. neoformans* reference strain) and JEC20**a** served as a control for the *msh2*Δ hybrid
250 crosses. H99α x JEC20**a** interspecific crosses produced robust mating structures, including
251 hyphae, basidia, and basidiospores (Figure 2A). Interestingly, bilateral crosses in which both
252 parent strains lacked *MSH2* produced similarly abundant hyphae as those produced by the H99α
253 x JEC20**a** cross, but exhibited more bald basidia and therefore fewer basidiospore chains (Figure
254 2A, 2B). In these bilateral crosses, 66%, 64%, 63%, and 57% of the basidia produced by each of
255 the four independent JEC20**a** *msh2*Δ mutants, respectively, lacked basidiospores compared to the
256 34% of bald basidia produced by wild-type H99α x JEC20**a** crosses (Figure 2A, 2B).

257 Basidiospores from wild-type, unilateral *msh2*Δ x wild type, and bilateral *msh2*Δ x
258 *msh2*Δ *C. neoformans* x *C. deneoformans* genetic crosses involving each of the four independent
259 JEC20**a** *msh2*Δ mutants and the KN99α *msh2*Δ strain were randomly dissected via
260 micromanipulation onto nutrient-rich YPD medium. The spores were allowed to germinate for
261 up to three weeks on YPD medium at room temperature, and total germination frequencies were
262 calculated (Figure 2B and Table S1). Dissected basidiospores from wild-type H99α x JEC20**a**
263 crosses germinated at an average frequency of 5.8%, which is similar to the previously published
264 frequency of 5% [43]. Basidiospores from unilateral crosses in which only one parent lacked
265 *MSH2* germinated at highly variable frequencies, ranging from an average of 2% to 11% (Figure
266 2B and Table S1). Progeny from bilateral crosses involving *msh2*Δ mutants in both parents
267 displayed increased germination frequencies compared to the progeny from the respective
268 unilateral crosses. Progeny germination frequencies were as high as 41% for some bilateral
269 *msh2*Δ x *msh2*Δ sexual crosses (Table S1), nearing the expected upper germination frequency
270 limit for *C. neoformans* x *C. deneoformans* crosses based on the presence of one large reciprocal
271 chromosomal translocation between *C. neoformans* H99α and *C. deneoformans* JEC21α, a strain
272 congenic to JEC20**a** with the exception of the *MAT* locus [33,57].
273

274 **Hybrid progeny are capable of hyphal growth and sporulation on nutrient-rich medium**
275 **and display phenotypes and genotypes typically associated with *C. neoformans* x *C.***
276 ***deneoformans* hybrids**

277 Due to the potentially unstable nature of hybrid progeny from *C. neoformans* x *C.*
278 *deneoformans* matings, progeny from crosses were allowed to germinate for up to three weeks
279 following dissection on YPD plates, compared to the standard 5-day incubation time [44,58].

280 During this prolonged incubation period, YPD agar plates were kept sealed with parafilm on the
281 benchtop. Surprisingly, after 14 days of incubation, aerial filaments began to emerge from the
282 periphery of several germinated progeny dissected from bilateral *msh2*Δ *C. neoformans* x *msh2*Δ
283 *C. deneoformans* crosses. Through microscopic analysis we found that these progeny were
284 producing structures resembling those observed during sexual reproduction under mating-
285 inducing conditions. Continued incubation and additional microscopic analysis revealed that
286 after approximately three weeks of incubation, these progeny produced basidia and sparse
287 basidiospores under conditions previously not known to support sexual reproduction of
288 *Cryptococcus* (rich medium, incubated in the light in sealed plates) (Figure 3).
289 The hybrid progeny from both wild-type *C. neoformans* x *C. deneoformans* crosses, as
290 well as those from unilateral and bilateral *msh2*Δ crosses, exhibited typical *Cryptococcus* hybrid
291 genotypes and phenotypes. With sequence-specific primers for the gene *STE20*, which exists as
292 two mating-type specific alleles and encodes a kinase involved in the pheromone response
293 signaling cascade, the mating types of the hybrid progeny were determined. Nearly all progeny
294 inherited and maintained both the *STE20*α allele from the *C. neoformans* parent and the *STE20***a**
295 allele from the *C. deneoformans* parent, which was expected due to the diploid hybrid
296 characteristic of *C. neoformans* x *C. deneoformans* hybrids (Figure S2) [43,46]. Based on flow
297 cytometry, the ploidy of the majority of the hybrid progeny was diploid or aneuploid (Figure S3).
298 Three out of the 27 progeny for which whole genome sequencing was obtained (progeny YX1,
299 YX3, and YX6) were, however, estimated to be close to haploid, and all three were from a wild-
300 type *C. neoformans* x *C. deneoformans* cross (Figure S3).
301 Although the majority of naturally isolated *C. neoformans* x *C. deneoformans* hybrid
302 strains are not self-fertile, *C. neoformans* x *C. deneoformans* hybrids produced under laboratory

303 conditions, similar to those presented in this study, are often self-fertile [43,44]. This fertility is
304 characterized by the ability to produce hyphae when incubated alone on mating-inducing media,
305 such as MS agar plates, at room temperature in the dark. All but one of the hybrid progeny
306 assessed produced hyphae on MS medium and many progeny were also capable of producing
307 basidia and basidiospores (Figure S4). The one progeny that was not self-fertile (YX3) had lost
308 the *C. deneoformans MAT***a** locus (Figure S2 and S4).
309

310 **Hybrid progeny from *msh2*Δ genetic crosses are highly aneuploid or diploid and do not**
311 **exhibit increased meiotic recombination frequencies**

312 Following isolation and characterization of hybrid progeny from the wild-type and
313 unilateral and bilateral *msh2*Δ mutant crosses, high molecular weight DNA was isolated from 27
314 hybrid progeny derived from wild-type (H99α x JEC20**a**), unilateral (KN99α *msh2*Δ x JEC20**a**
315 and H99α x JEC20**a** *msh2*Δ-*1*), and bilateral (KN99α *msh2*Δ x JEC20**a** *msh2*Δ-*1*) crosses for
316 whole-genome sequencing. The chromosomal composition of the hybrid progeny, along with the
317 identification of potential sites of meiotic recombination, LOH events, and heterozygosity across
318 the genome, was assessed by employing the analytic pipeline described in Figure S5 (see
319 methods for details). To determine ploidy, read depth was assessed in conjunction with flow
320 cytometry data (Figure 4, S2, and S5).

321 Hybrid progeny from crosses between wild-type *C. neoformans* and *C. deneoformans*
322 parental strains, as well as those from unilateral and bilateral *msh2*Δ hybrid crosses, displayed
323 high rates of aneuploidy (Figure 4, Figure S6). Whole-genome sequencing revealed five
324 instances of potential meiotic recombination in the hybrid progeny, in which a single copy of a
325 chromosome was inherited and was chimeric, meaning that various loci along the same

326 chromosome were homologous to either the *C. neoformans* or the *C. deneoformans* parent, but
327 not both. For example, in progeny YX3, a portion of *C. neoformans* Chr7 (homologous to *C.*
328 *deneoformans* chromosome 6) was inherited solely from the H99α parent while the remaining
329 portion of the chromosome was inherited from the JEC20**a** parent (Figure 4A). Other examples
330 of these instances of recombination are seen in progeny YX1 on Chr2 and Chr10, in progeny
331 YX6 on *C. neoformans* Chr5/*C. deneoformans* Chr4, and also again in progeny YX3 on *C.*
332 *neoformans* Chr14/*C. deneoformans* Chr8. All five of these instances of potential meiotic
333 recombination identified were in progeny derived from the wild-type cross, rather than from
334 crosses involving *msh2*Δ mutants. Interestingly, hybrid progeny from bilateral *msh2*Δ mutant
335 crosses had ploidies closer to 2N, were nearly completely heterozygous across a majority of the
336 genome, and displayed fewer LOH events (Figure 4, S3, S6, S7). Quantifying this
337 heterozygosity, hybrid progeny derived from bilateral *msh2*Δ x *msh2*Δ mutant crosses had
338 significantly more heterozygosity ($p<0.05$, Kruskal-Wallis test, Dunn's test) across the entirety
339 of their genomes with an average of 96% heterozygosity relative to hybrid progeny derived from
340 wild-type H99α x JEC20**a** crosses, which had an average of 67% heterozygosity across their
341 genomes (Figure S7 and Table S2). Another unexpected finding was a phenomenon associated
342 only with hybrid progeny derived from unilateral or bilateral *msh2*Δ crosses in which
343 chromosome breaks occurring at centromeres, transposable elements, and at other regions in the
344 genome appeared to have been repaired by *de novo* addition of telomeric repeats at breakpoints
345 (Figure 4C, Figure S6, and Figure S8).

346 Due to the high levels of heterozygosity across the genomes of many of the *C.*
347 *neoformans* x *C. deneoformans* hybrid progeny, it is possible that some recombination events
348 might not be detected by merely assessing the depth of reads aligning to each parental reference

349 strain. For instance, if a hybrid progeny inherited both homeologous chromosomes involved in a
350 meiotic reciprocal recombination event, read depth coverage would be equal across both parental
351 chromosomes. To ensure that additional recombination events like these were not being missed,
352 Illumina paired-end reads were aligned to a combined reference that included both parental
353 genomes to identify additional possible recombination sites. This analysis was based on the
354 assumption that if a recombination event occurred, the two reads of a pair would align to
355 different chromosomes. Different filtering thresholds were applied based on mapping quality and
356 the number of read-pairs supporting the event. The results of these analyses, using more or less
357 stringent filtering thresholds to detect recombination events, show similar frequencies of
358 recombination events for each of the different types of crosses: wild-type *C. neoformans* x *C.*
359 *deneoformans* crosses, unilateral crosses in which one parent lacks Msh2, and bilateral crosses in
360 which both of the parental strains lack Msh2 (Figure 5 and Table S3). With the strictest filtering
361 thresholds (at least 15 read-pairs supporting the event and a mapping quality of 60), a median of
362 1 recombination event is detected across the whole genomes of progeny derived wild-type,
363 unilateral KN99α *msh2*Δ x JEC20**a**, and bilateral crosses, while a median of 1.5 events are
364 detected in progeny from unilateral H99α x JEC20**a** *msh2*Δ-*1* crosses. Furthermore, with the
365 least strict thresholds (at least 5 read-pairs supporting and mapping quality greater than or equal
366 to 0), a median of 10 recombination events are detected in progeny derived from a wild-type
367 cross, a median of 8.5 and 6.5 events are detected from both unilateral crosses (H99α x JEC20**a**
368 *msh2*Δ-*1*and KN99α *msh2*Δ x JEC20**a**, respectively), and a median of 5.5 events are detected in
369 progeny from bilateral crosses (Figure 5, Table S3).
370

371 **Discussion**

372 In this study, we investigated the role of MMR in maintaining the species boundary
373 between two closely related and prevalent human fungal pathogens, *C. neoformans* and *C.*
374 *deneoformans*. In addition to its role in DNA repair, the MMR pathway plays a highly conserved
375 role in blocking recombination between homeologous DNA sequences during meiosis, serving as
376 a post-zygotic barrier. Findings from previous studies in prokaryotic and eukaryotic models
377 indicated that rendering the MMR pathway nonfunctional by genetically deleting pathway
378 components, particularly Msh2, allows increased homoleogous recombination during meiosis [8-
379 10]. Therefore, we hypothesized that lack of Msh2 in *Cryptococcus* would allow increased
380 pairing and recombination between homeologous chromosomes during meiosis, leading to
381 decreased rates of chromosome nondisjunction, and ultimately the production of more viable
382 progeny with fewer instances of aneuploidy.
383 We genetically engineered *C. deneoformans* strains lacking *MSH2* and confirmed
384 through fluctuation analysis and characterization of mutational spectra that Msh2 plays a similar
385 role in DNA MMR in *C. deneoformans* as has been previously observed across eukaryotes and in
386 other pathogenic *Cryptococcus* species [50,51]. Although all of the independent *C.*
387 *deneoformans msh2*Δ mutants displayed increased rates of mutation that conferred resistance to
388 FK506 and rapamycin, the *msh2*Δ-*1* mutant failed to produce 5-FOA resistant colonies, which
389 was attributable to a frameshift mutation identified in the *FUR1* that is synthetically lethal with
390 mutations that can confer 5-FOA resistance. This conclusion was further supported by the
391 inability of a *C. neoformans fur1*Δ mutant to produce 5-FOA resistant colonies as well as the
392 inherent 5FC and 5FU resistance displayed by both the *msh2*Δ-*1* and *fur1*Δ mutants [56]. This
393 finding illustrates that, in some instances, *Cryptococcus* is unable to acquire concurrent
394 resistance to both 5-FOA and the clinically relevant drug 5-flucytosine (5FC) as well as the

395 antimetabolite 5-fluorouracil (5FU). However, there are additional mutations that confer 5FC
396 and 5FU resistance, such as those in the*UXS1* gene, which encodes a UDP-glucuronate
397 decarboxylase, which are not presumed to be synthetically lethal with mutations conferring 5-
398 FOA resistance [56].

399 Similar to results from previous studies in prokaryotic and eukaryotic microorganisms,
400 germination frequencies of *C. neoformans* x *C. deneoformans* hybrid progeny derived from
401 bilateral *msh2*Δ crosses increased substantially relative to wild-type hybrid germination
402 frequencies [8-10]. Although loss of *MSH2* could lead to increased frequencies of homeologous
403 recombination in *Cryptococcus* and subsequently increased viability of progeny, it is also
404 possible that loss of *MSH2* contributes to spore inviability due to the hypermutator phenotype of
405 *msh2*Δ mutants, as has been observed in previous studies in *S. cerevisiae* [59]. Despite the low
406 germination frequencies of some progeny isolated from unilateral *msh2*Δ crosses, progeny
407 derived from crosses that involved the same independent JEC20**a** *msh2*Δ mutant (i.e. H99α x
408 JEC20**a** *msh2*Δ-*1* and KN99α *msh2*Δ x JEC20**a** *msh2*Δ-*1*) all had increased viability in bilateral
409 crosses compared to the corresponding unilateral crosses.

410 Hybrid progeny derived from genetic crosses between *C. neoformans* and *C.*
411 *deneoformans* isolates are known to display unique phenotypes. For example, hybrids, especially
412 those generated under laboratory conditions, can exhibit self-fertility, producing hyphae and in
413 some cases basidia and spores under mating-inducing conditions [43,46]. The majority of the
414 hybrid progeny derived from the crosses described here displayed phenotypes in accord with
415 previously published findings for *C. neoformans* x *C. deneoformans* hybrids [43,44]. However, a
416 novel transgressive phenotype was exhibited by multiple progeny isolated from bilateral *msh2*Δ
417 mutant crosses: the ability to produce sexual reproduction structures on nutrient-rich YPD

418 medium in sealed plates in the light. This was surprising, because environmental cues known to
419 suppress mating in *Cryptococcus* include nutrient-rich environments, such as YPD medium, as
420 well as light and high levels of humidity and carbon dioxide [37,60-63]. The production of
421 sexual reproduction structures by these hybrid progeny under nutrient-rich, light, sealed
422 conditions represents sexual reproduction of *Cryptococcus* in a novel environment. The isolation
423 of self-fertile progeny provides an important example of how hybridization can enable
424 *Cryptococcus* to access a dimorphic state under conditions that are not conducive for the parental
425 isolates to sexually reproduce. The ability to produce spores, the infectious propagules of
426 *Cryptococcus* human pathogens, in a previously prohibitive environment also provides the
427 opportunity to produce more infectious propagules, which could potentially result in a higher rate
428 of infection. This is in line with previous studies which found that up to 30% of clinical isolates
429 in Europe were *C. neoformans* x *C. deneoformans* hybrids [40-42]. Future studies using
430 quantitative trait loci mapping as well as other molecular genetic techniques could provide
431 insight into the genetic determinants contributing to this novel self-fertility phenotype.
432 Unlike their haploid parents, *C. neoformans* x *C. deneoformans* hybrid progeny typically
433 have relatively unstable aneuploid or diploid karyotypes [44-46]. All hybrid progeny derived
434 from *C. neoformans* x *C. deneoformans* crosses in this study were aneuploid based on both
435 FACS analysis and whole-genome sequencing data. We observed instances of low read-depth
436 associated with several aneuploid chromosomes, which are thought to represent loss of
437 chromosomes in fractions of populations of cells that were sequenced at the whole-genome level,
438 reflecting the karyotypic instability of these progeny. The majority (21/27) of the hybrid progeny
439 were diploid (11/27 progeny) or nearly diploid (2N+1 or 2N-1) (10/27 progeny), and all progeny
440 derived from unilateral and bilateral *msh2*Δ crosses were close to diploid. The higher ploidy

441 levels in progeny derived from crosses involving *msh2*Δ mutants were unexpected because
442 MMR mutants in yeast produce progeny with lower levels of aneuploidy [9]. It is possible that
443 the diploids or near diploid *C. neoformans* x *C. deneoformans* progeny were the meiotic products
444 of a tetraploid, generated by the fusion of two diploids that formed within the mating patch.
445 However, interspecific progeny derived from divergent *Saccharomyces* tetraploids display much
446 higher germination frequencies (>90%), which is very different from the frequencies associated
447 with these *C. neoformans* x *C. deneoformans* hybrid progeny, making it unlikely that they are
448 derived from tetraploids [64]. It is also possible that the high frequency of diploid progeny could
449 be indicative of failed meiosis. However, *Cryptococcus* mutants lacking known meiotic genes,
450 such as those encoding the meiosis-specific endonuclease Spo11 or the key meiotic regulator
451 Dmc1, are unable to produce basidiospores at an appreciable frequency and their basidia lack the
452 typical four chains of basidiospores [65,66]. In contrast to the abnormal sexual reproduction
453 structures produced by these meiosis-deficient mutants, the basidia produced by hybrid crosses
454 in this study generated four lengthy spore chains, albeit at a lower frequency than intra-species
455 crosses in some cases. The higher ploidy in hybrid progeny from bilateral *msh2*Δ x *msh2*Δ
456 crosses was also associated with significantly more heterozygosity across their genomes than
457 their counterparts from wild-type *C. neoformans* x *C. deneoformans* crosses. It is possible that
458 this higher level of heterozygosity contributes to the increased viability by masking deleterious
459 alleles or overcoming Bateson-Dobzhansky-Muller genetic incompatibilities, as has been
460 observed in previous studies of *C. neoformans* x *C. deneoformans* hybrids [49,67].
461 An interesting phenomenon associated with progeny derived from unilateral and bilateral
462 *msh2*Δ crosses was the *de novo* addition of telomeric repeat sequences to various locations in the
463 genome including centromeric repeat sequences, non-centromeric transposable elements, and the

464 rDNA repeat locus. Although the addition of telomeric repeats at non-telomeric sites can
465 promote the stabilization of a broken chromosome, it also leads to the loss of a large portion of a
466 chromosome, which would normally threaten cell viability. However, the presence of both
467 homeologous chromosomes from each parental species in these hybrid progeny alleviates this
468 problem. *De novo* telomeric repeat addition has also been observed in *Cryptococcus* following
469 CRISPR-mediated double-stranded breaks at centromeres [68]. Overall, our results suggest that
470 in *C. neoformans* x *C. deneoformans* crosses involving *msh2*Δ mutants, chromosomes may be
471 more prone to double-stranded breaks, or that the normally occurring double-stranded breaks are
472 unable to be properly repaired.

473 Many instances of loss of heterozygosity and potential recombination were observed in
474 the hybrid progeny assessed in this study. Three hybrid progeny isolated from wild-type *C.*
475 *neoformans* x *C. deneoformans* crosses were close to haploid with five or fewer aneuploid
476 chromosomes. Conversely, hybrid progeny derived from unilateral or bilateral *msh2*Δ crosses
477 were more highly aneuploid or diploid. Although instances of recombination were detected in
478 progeny derived from each of the different types of crosses (wild-type, unilateral, and bilateral
479 *msh2*Δ), it is difficult to determine if these recombination events occurred during meiosis or
480 during subsequent mitotic growth during the culturing required for whole-genome sequencing.
481 Despite this inherent difficulty, we expected to see significantly higher rates of potential
482 instances of meiotic recombination based on previous studies in both prokaryotes and eukaryotes
483 [8-10]. Surprisingly, based on read-depth analysis as well as detecting recombination events by
484 aligning whole-genome sequencing from the progeny to a combined reference genome with both
485 parental species, no increase in recombination was observed in crosses involving either a single
486 parent lacking Msh2 or both parents lacking Msh2. One potential explanation for this

487 observation could be that the level of sequence divergence between *C. deneoformans* and *C.*
488 *neoformans* is large enough that meiotic recombination will be inefficient, even in the absence of
489 MMR. However, previous studies on meiotic recombination in *S. cerevisiae* indicate that
490 although recombination is less efficient at high levels of sequence divergence, loss of MMR
491 leads to approximately a 24-fold increase in meiotic recombination between 85% divergent
492 sequences, the same level of divergence as between *C. neoformans* and *C. deneoformans*
493 [31,33,69]. Furthermore, studies on mitotic recombination in *S. cerevisiae* also similarly indicate
494 that even at 26% sequence divergence, loss of MMR leads to a 55-fold increase in mitotic
495 recombination [70,71].

496 In summary, this study illustrates several key findings on the roles of MMR in
497 *Cryptococcus*. In *C. deneoformans*, *msh2*Δ mutants behave as hypermutators on various selective
498 media and acquire INDELs in homopolymeric nucleotide runs. Hybrid *C. neoformans* x *C.*
499 *deneoformans* progeny dissected from genetic crosses involving *msh2*Δ mutants generally
500 displayed increased germination frequencies compared to those from wild-type crosses. Hybrid
501 progeny derived from *msh2*Δ crosses displayed phenotypes and karyotypes characteristic of *C.*
502 *neoformans* x *C. deneoformans* hybrid strains, such as diploidy/aneuploidy and self-fertility,
503 although some progeny derived from bilateral *msh2*Δ crosses displayed a novel, transgressive
504 phenotype in which they were capable of producing hyphae, basidia, and basidiospores on a
505 glucose- and nutrient-rich medium in the light. The increased viability of hybrid progeny derived
506 from bilateral *msh2*Δ crosses suggests that loss of Msh2 in *Cryptococcus* may allow
507 homeologous chromosomes to pair more efficiently during meiosis. However, it is interesting to
508 note that loss of Msh2 did not seem to increase the frequency of meiotic recombination between
509 *C. neoformans* and *C. deneoformans* homeologous chromosomes. These results suggest that

510 alternative pathways or additional MMR components may play different or more important roles
511 in maintaining species boundaries in *Cryptococcus* than in other previously studied organisms,
512 an intriguing avenue for future studies. Unlike bacterial and budding yeast studies, loss of Msh2
513 did not relax the species boundary between *C. neoformans* and *C. deneoformans,* illustrating
514 robust post-zygotic mechanisms that ultimately maintain integrity between these two closely
515 related species.
516

517 ## Materials and methods

518 **Strains and growth**

519 The *C. neoformans* and *C. deneoformans* strains described in this study are listed in Table
520 S4. Strains were stored at -80°C in liquid yeast extract peptone dextrose (YPD) supplemented
521 with 15% glycerol. Strains were inoculated on YPD agar plates, initially grown at 30°C for 3
522 days, and then maintained at 4°C. Due to the hypermutator phenotypes associated with *msh2*Δ
523 strains and the genomic instability associated with *C. neoformans* x *C. deneoformans* hybrids,
524 strains used in the experiments of this study were not maintained for more than two weeks at 4°C
525 on YPD agar plates and fresh cells from frozen glycerol stocks were inoculated to new YPD agar
526 plates at the end of each two-week period.
527

528 **Generation of *msh2*Δ deletion mutants**

529 The open-reading frame of the *MSH2* gene (gene ID: CNA07480) in the *C. deneoformans*
530 JEC20**a** genetic background was replaced with the gene encoding the dominant drug-resistance
531 marker for nourseothricin resistance, *NAT*, by homologous recombination via biolistic
532 transformation as previously described [72]. Following transformation and selection on YPD +

533 100 µg/mL nourseothricin agar medium, genomic DNA was isolated from candidate mutants
534 with the MasterPure DNA purification kit (Epicentre) and PCR followed by gel electrophoresis
535 confirmed correct integration of the *NAT* dominant resistance marker. The locations of the
536 primers used to generate the deletion allele and confirm deletion mutants are depicted in Figure
537 S1A and their sequences are given in Table S5.
538
539 **Fluctuation analysis to quantify mutation rates**
540 Fluctuation analysis was utilized to quantify the mutation rates, or the number of
541 mutations per cell per generation, of the JEC20**a** *msh2*Δ mutants. The wild-type strain JEC20**a**
542 served as a negative control and the *msh2*Δ mutant from the 2015 Madhani deletion collection in
543 the KN99α genetic background served as a positive control [52]. For each strain, including
544 controls, ten 5 mL YPD liquid cultures were each inoculated with a single colony from a YPD
545 agar stock plate. Cultures were incubated overnight at 30°C. After incubation, cultures were
546 pelleted at 3,000 x *g*, washed twice with 5 mL dH_2O, and resuspended in 4 mL dH_2O. 100 µL of
547 undiluted, washed cells were plated directly to YPD + 100 ng/mL rapamycin + 1 µg/mL FK506
548 or yeast nitrogen base (YNB) + 1 mg/mL 5-fluoroorotic acid (5-FOA) solid agar medium.
549 Washed cells were diluted 1:100,000 and 100 µL of the dilution was plated to YPD solid agar
550 medium. Inoculated YPD and YNB+5-FOA plates were incubated at 30°C for 4 or 14 days,
551 respectively. Inoculated YPD+rapamycin+FK506 plates were incubated at 37°C for 14 days.
552 Following incubation, colonies on each of the media were counted and mutation frequencies
553 were calculated with the FluCalc program, which utilizes the Ma-Sandri-Sarkar maximum-
554 likelihood estimation (MSS-MLE) equation for calculations [73].
555

556 **Mutation spectra analysis**

557 Single resistant colonies from fluctuation analyses were streak purified to YPD +
558 rapamycin + FK506 medium and grown for 3 days at 37°C. Genomic DNA was isolated using
559 the MasterPure Yeast DNA Purification Kit (Epicenter Biotechnologies, Madison, WI), and the
560 *FRR1* gene was PCR-amplified with Phusion High-Fidelity DNA Polymerase (NEB, Ipswich
561 MA, USA). PCR products were subjected to gel electrophoresis, extracted using a QIAgen gel
562 extraction kit, and mutations were identified through classical Sanger Sequencing at Genewiz.
563 Fisher's exact probability test was used to calculate statistically significant differences between
564 the frequencies of 1-bp INDEL mutations compared to other types of mutations in
565 YPD+rapamycin+FK506-resistant colonies from strains lacking *MSH2* compared to the wild-
566 type JEC20**a** strain using the VassarStats online software (http://vassarstats.net).

567

568 **Papillation assays**

569 Papillation assays were conducted on YNB solid agar medium supplemented with 1
570 mg/mL 5-FOA, 100 µg/mL 5-fluorocytosine (5FC), or 100 µg/mL 5-fluorouridine (5FU). For
571 this assay, ten independent YPD liquid cultures were inoculated with ten single colonies from
572 YPD agar plates and incubated overnight at standard laboratory conditions. Following overnight
573 culture, cells were pelleted at 3,000 x *g* and resuspended in 2 mL dH_2O. Sterile cotton swabs
574 were then used to inoculate quadrants of YNB + 5-FOA agar plates with each independent
575 overnight culture. The inoculated agar plates were incubated at 30°C for 6 days to allow
576 sufficient growth to visualize resistant colonies.

577

578 **Genetic crosses and progeny dissection**

579 All genetic crosses were conducted on Murashige and Skoog (MS) agar medium
580 following Basic Protocol 1 for mating assays as previously described [58]. The frequency of bald
581 basidia was calculated by imaging random areas surrounding two mating patches per cross,
582 counting the number of bald basidia and basidia producing basidiospores to determine the
583 frequency of bald basidia, and the frequencies from the two mating patches were averaged. For
584 each genetic cross, two independent mating patches were assessed; for each mating patch, over
585 130 total basidia were assessed across at least 11 images. Images used to quantify bald basidia
586 were taken on a Zeiss Axio Scope.A1 with camera. Random basidiospore dissection was
587 performed as described in Basic Protocol 2 [58]. Following dissection, the micromanipulated
588 basidiospores were germinated for up to 3 weeks on YPD agar plates sealed with parafilm and
589 incubated on the laboratory benchtop. Images of self-fertile hybrid progeny derived from
590 bilateral *msh2*Δ crosses producing sexual structures on YPD agar plates following dissection
591 were taken with an Accu-Scope EXC-500 microscope with an attached Nikon DXM1200F
592 microscope camera.
593
594 **Phenotyping and genotyping of hybrid progeny**
595 Primers designed to specifically amplify only the *C. neoformans STE20***a**, *C. neoformans*
596 *STE20*α, *C. deneoformans STE20***a**, or *C. deneoformans STE20*α alleles aided in identifying the
597 mating types of the hybrid progeny (primers listed in Table S5). To assess self-filamentation,
598 hybrid progeny were spotted onto MS agar plates and incubated for 14 days at room temperature
599 (approximately 24°C) in the dark as described in Basic Protocol 1 [58]. Filamentation was
600 assessed via microscopy after 14 days of incubation with an Accu-Scope EXC-500 microscope
601 with an attached Nikon DXM1200F microscope camera.

602

603 **Flow Cytometry**

604 Cells were patched onto YPD agar medium and incubated at 30°C overnight; strains that
605 exhibited slow growth were incubated at 30°C for three days. Cells were harvested by scraping a
606 2 mm sized colony with a toothpick and resuspending in 1 mL PBS buffer. Cells were washed
607 once with 1 mL PBS and then fixed in 1 mL 70% ethanol overnight at 4°C. After fixation, cells
608 were washed once with 1 mL NS buffer (10 mM Tris-HCl pH=7.6, 150 mM sucrose, 1 mM
609 EDTA pH=8.0, 1 mM $MgCl_2$, 0.1 mM $ZnCl_2$, 0.4 mM phenylmethylsulfonyl fluoride, 1 mM β-
610 mercaptoethanol) and resuspended in 180 µl NS buffer supplemented with 5 µl propidium iodide
611 (0.5 mg/mL) and 20 µl RNase A (10 mg/mL). Cells were incubated covered, overnight at 4°C
612 with shaking. Prior to analysis, 50 µl of cells were diluted in 2 mL of 50 mM Tris-HCl (pH=8.0)
613 in a 5 mL falcon tube. Flow cytometry was performed on 10,000 cells and analyzed on the FL1
614 channel on a Becton-Dickinson FACScan.

615

616 **Whole-genome sequencing**

617 Single colonies of strains for whole-genome sequencing were inoculated into 50 mL of
618 liquid YPD and grown overnight at 30°C in standard laboratory conditions. Overnight cultures
619 were then pelleted at 3,000 x *g* in a tabletop centrifuge and subsequently lyophilized. High
620 molecular weight DNA was extracted from lyophilized cells following the CTAB protocol as
621 previously described [74]. Genomic DNA libraries were prepared for sequencing with a Kapa
622 Hyper plus library kit for 300 bp inserts and sequenced at the Duke Sequencing and Genomic
623 Technologies Shared Research core facility. Libraries were sequenced using paired-end, 2 x 150
624 bp reads on an Illumina HiSeq 4000 platform.

625

626 **Whole-genome and chromosome composition analyses of the hybrid progeny**

627 *C. deneoformans* strain JEC21α is the congenic mating partner of JEC20**a**, which was
628 obtained in an earlier study by selecting *MAT*α progeny after ten rounds of backcrossing to
629 JEC20**a** [75]. The genomes of *C. deneoformans* JEC21α and JEC20**a** strains are therefore nearly
630 identical, with the exception of 5,322 SNPs mainly distributed over three genomic regions
631 (including the mating-type locus) [76]. Because a highly contiguous *de novo* genome assembly
632 of JEC20**a** is not currently available, the genome assembly of strain JEC21α
633 (GCA_000091045.1) was used for all comparisons. The nuclear genomes of two parental strains,
634 *C. neoformans* H99α and *C. deneoformans* JEC21α were compared by performing whole-
635 genome alignments with Satsuma (https://github.com/bioinfologics/satsuma2) [77], using default
636 parameters. The output of Satsuma was input to the visualization tools "BlockDisplaySatsuma"
637 and "ChromosomePaint", included in the same package to generate a postscript file. For
638 representation purposes, chromosome color codes were modified in Adobe Illustrator, and
639 centromeres and other genomic features (rDNA and *MAT* loci) were superimposed at scale based
640 on their respective genomic coordinates. Characterization of the chromosome composition of the
641 hybrid progeny followed the procedure summarized in Figure S5A. The combined nuclear
642 reference genome used in this study was built with the genome assemblies of the two parental
643 strains after reordering and reorienting the JEC21α contigs to maximize collinearity with the
644 H99α assembly (Figure S5B), using the Mauve Contig Mover tool [78]. A dot plot analysis
645 comparing the H99α assembly with the JEC21α rearranged assembly (Figure S5B) was
646 performed with D-Genies application [79], which uses minimap2 [80] for aligning the two
647 genomes. Raw Illumina paired-end reads of the selected progeny and the combined reference

genomes were input into the sppIDer pipeline [81], which is a wrapper that sequentially maps the Illumina short reads to the combined reference, performs quality filtering (MQ > 3), and generates depth of coverage plots (Figure 4 and S5). For each progeny, the number of chromosomes and the ploidy were estimated from the sppIDer plots in conjunction with the flow-cytometry data. Chromosomal aberrations, e.g. due to recombination, or chromosome breakage followed by *de novo* telomere addition, were inferred from the sppIDer plots, and further validated by visual inspection of the mapped reads in IGV [82].

Heterozygosity of the hybrid progeny

To inspect the heterozygosity levels of the hybrid progeny, the same set of Illumina paired-end reads were mapped to the *C. neoformans* H99α reference genome, using the BWA-MEM short-read aligner (v0.7.17-r1188) with default settings [83]. SNP discovery, variant evaluation, and further refinements were carried out with the Genome Analysis Toolkit (GATK) best-practices pipeline [84,85] (v4.0.1.2), including the use of Picard tools to convert SAM to sorted BAM files, fix read groups (module: 'AddOrReplaceReadGroups'; SORT_ORDER=coordinate), and mark duplicates. Variant sites were identified with HaplotypeCaller from GATK and only high-confidence variants that passed filtration were retained (the "VariantFiltration" module used the following criteria: DP < 20 || QD < 2.0 || FS > 60.0 || MQ < 40.0 || SOR > 4.0). The genome-wide level of heterozygosity was defined as the ratio of the number of heterozygous SNPs divided by the total number of SNPs (i.e. heterozygous and non-reference homozygous SNPs), and was calculated from the resulting VCF files on a per-individual basis after extracting the corresponding sites using the module VariantsToTable from GATK. Due to the high variance in heterogeneity across the genomes of

671 progeny from the different genetic crosses and unequal numbers of progeny analyzed from each
672 different genetic cross, the non-parametric Kruskal-Wallis test followed by Dunn's test were
673 performed using JMP v15 (SAS Institute).
674
675 **Detection of variants in the genome of the JEC20a *msh2*Δ-*1* mutant**
676 To predict variants in the JEC20**a** *msh2*Δ-*1* mutant, we employed the variant calling
677 procedure described above, using the JEC21 genome (GCA_000091045.1) as reference. The
678 effect of a variant (SNPs and INDELs) was predicted and annotated using SnpEff [86]. Only
679 variants of moderate and high impact were considered (i.e. excluding synonymous and non-
680 coding variants), and variants in common between the mutant and the JEC20**a** parent were
681 flagged background mutations and excluded. Given the inability of the *msh2*Δ-*1* mutant to
682 produce 5-FOA resistant colonies, we specifically focused on mutations in genes involved in the
683 *de novo* and salvage pathways of pyrimidine biosynthesis.
684
685 **Detection of recombination events based on read-pair alignment**
686 To detect possible recombination events in *C. neoformans* x *C. deneoformans* hybrids,
687 the respective reference genomes were retrieved from NCBI (accession: ASM301198v1 for H99
688 and ASM9104v1 for JEC21 as no genome assembly for JEC20 was available). As the H99 strain
689 was not sequenced in this project, the respective Illumina paired-end reads were retrieved from
690 SRA (SRR7042283). All Illumina paired-end libraries were filtered with the default parameters
691 of Trimmomatic v0.36 [87]. Filtered reads of each strain (including the parental strains) were
692 aligned on the combined reference (containing both H99 and JEC21 genome assemblies) with

693 HaploTypo pipeline [88], using BWA-MEM v0.7015 [83], samtools v1.9 (Li 2011) and GATK
694 v4.0.2.1 [85].
695 To detect recombination events, we assumed that in recombinant sites it would be
696 possible to detect read-pairs with the two reads aligning to different chromosomes. Therefore, for
697 each strain (including the parental strains), we obtained the coordinates of all the pairs with each
698 read aligning to a chromosome of a different parental species. Only reads with at least 75 bp and
699 a maximum of 3 mismatches were considered. As there were some differences between the
700 parental JEC20 and the reference JEC21, variant calling was performed with HaploTypo pipeline
701 [88], using GATK v4.0.2.1 [85]. Bedtools v2.26 [90] intersect was used to determine how many
702 SNPs overlapped each read, and in cases in which more than 3 SNPs would overlap, the filter of
703 mismatches was adjusted to that value. The final results were filtered based on mapping quality
704 (MQ) and number of pairs supporting a given recombination event (Bedtools v2.26 [90] merge
705 with a distance of 100bp). Four different analyses were performed using more or less stringent
706 filters: i) MQ = 60 and at least 15 reads supporting the event (stricter); ii) MQ = 60 and at least
707 five reads supporting the event; iii) MQ >= 0 and at least 15 reads supporting the event; iv) MQ
708 >= 0 and at least five reads supporting the event (less strict). The script that automatizes this
709 pipeline is available on GitHub (https://github.com/Gabaldonlab/detect_recombination/). As it
710 was possible to detect recombination in the parental strains (false positives), Bedtools v2.26 [90]
711 subtract was used to remove all the regions detected in hybrid progeny that were also detected in
712 the parental strains.
713
714 **Data availability**

715 Raw sequence data has been deposited into the National Center for Biotechnology
716 Information Sequence Read Archive under BioProject accession no. PRJNA626676. The
717 accession numbers for each sample are provided in Table S6.
718

Acknowledgements

We thank Sue Jinks-Robertson, Tom Petes, and Paul Magwene for critical reading and comments on the manuscript. We thank Jay Jawahar for his assistance in the initial phases of constructing the JEC20**a** *msh2*Δ mutants. We also thank the Madhani Laboratory and NIH grant R01 AI100272 for the KN99α *msh2*Δ and *fur1*Δ deletion strains.

724 ## References

725 1. Arnold ML, Bulger MR, Burke JM, Hempel AL, Williams JH. Natural hybridization: how
726 low can you go and still be important? *Ecology* 1999;80(2): 371-381. doi: 10.1890/0012-
727 9658(1999)080[0371:NHHLCY]2.0.CO;2
728 2. Goulet BE, Roda F, Hopkins R. Hybridization in plants: old ideas, new techniques. *Plant*
729 *Physiol*. 2017;173(1):65-78. PMID: 27895205.
730 3. Vallejo-Marin M, Hiscock SJ. Hybridization and hybrid speciation under global change. *New*
731 *Phytol*. 2016;211(4):1170-87. PMID: 27214560.
732 4. Fu C, Coelho MA, David-Palma M, Priest SJ, Heitman J. Genetic and genomic evolution of
733 sexual reproduction: echoes from LECA to the fungal kingdom. *Curr Opin Genet Dev*.
734 2019;58-59:70-75. PMID: 31473482.
735 5. Lu AL, Clark S, Modrich P. Methyl-directed repair of DNA base-pair mismatches in vivo.
736 *Proc Natl Acad Sci U S A*. 1983;80(15):4639-43. PMID: 6308634.
737 6. Grilley M, Holmes J, Yashar B, Modrich P. Mechanisms of DNA-mismatch correction.
738 *Mutat Res*. 1990;236(2-3):253-67. PMID: 2144613.
739 7. Harfe BD, Jinks-Robertson S. Mismatch repair proteins and mitotic genome stability. *Mutat*
740 *Res*. 2000;451(1-2):151-67. PMID: 10915870.
741 8. Rayssiguier C, Thaler DS, Radman M. The barrier to recombination between *Escherichia*
742 *coli* and *Salmonella typhimurium* is disrupted in mismatch-repair mutants. *Nature*.
743 1989;342(6248):396-401. PMID: 2555716.
744 9. Hunter N, Chambers SR, Louis EJ, Borts RH. The mismatch repair system contributes to
745 meiotic sterility in an interspecific yeast hybrid. *EMBO J.* 1996;15(7):1726-33. PMID:
746 8612597.

747 10. Chambers SR, Hunter N, Louis EJ, Borts RH. The mismatch repair system reduces meiotic
748 homeologous recombination and stimulates recombination-dependent chromosome loss. *Mol*
749 *Cell Biol*. 1996;16(11):6110-20. PMID: 8887641.
750 11. Pongracz JD, Paetkau D, Branigan M, Richardson E. Recent hybridization between a polar
751 bear and grizzly bears in the Canadian Arctic. *Arctic*. 2017;70(2):151-60. doi:
752 10.14430/arctic4643
753 12. Darwin C. *What Mr. Darwin saw in his voyage round the world in the ship "Beagle."* New
754 York: Harper & Bros; 1879.
755 13. Lamichhaney S, Han F, Webster MT, Andersson L, Grant BR, Grant PR. Rapid hybrid
756 speciation in Darwin's finches. *Science*. 2018;359(6372):224-8. PMID: 29170277.
757 14. Kelso J, Prufer K. Ancient humans and the origin of modern humans. *Curr Opin Genet Dev*.
758 2014;29:133-8. PMID: 25286439.
759 15. Dietrich FS, Voegeli S, Brachat S, Lerch A, Gates K, Steiner S, et al. The *Ashbya gossypii*
760 genome as a tool for mapping the ancient *Saccharomyces cerevisiae* genome. *Science*.
761 2004;304(5668):304-7. PMID: 15001715.
762 16. Wolfe KH. Origin of the yeast whole-genome duplication. *PLOS Biol*. 2015;13(8):e1002221.
763 PMID: 26252643.
764 17. Marcet-Houben M, Gabaldón T. Beyond the whole-genome duplication: phylogenetic
765 evidence for an ancient interspecies hybridization in the baker's yeast lineage. *PLOS Biol*.
766 2015;13(8):e1002220. PMID: 26252497.
767 18. Stukenbrock EH, Christiansen FB, Hansen TT, Dutheil JY, Schierup MH. Fusion of two
768 divergent fungal individuals led to the recent emergence of a unique widespread pathogen
769 species. *Proc Natl Acad Sci U S A*. 2012;109(27):10954-9. PMID: 22711811.

770 19. Pryszcz LP, Nemeth T, Saus E, Ksiezopolska E, Hegedusova E, Nosek J, et al. The genomic
771 aftermath of hybridization in the opportunistic pathogen *Candida metapsilosis*. *PLOS Genet*.
772 2015;11(10):e1005626. PMID:26517373.
773 20. Mixão V, Hansen AP, Saus E, Boekhout T, Lass-Florl C, Gabaldón T. Whole-genome
774 sequencing of the opportunistic yeast pathogen *Candida inconspicua* uncovers its hybrid
775 origin. *Front Genet*. 2019;10:383. PMID: 31105748.
776 21. Mixão V, Gabaldón T. Hybridization and emergence of virulence in opportunistic human
777 yeast pathogens. *Yeast*. 2018;35(1):5-20. PMID: 28681409.
778 22. Rajasingham R, Smith RM, Park BJ, Jarvis JN, Govender NP, Chiller TM, et al. Global
779 burden of disease of HIV-associated cryptococcal meningitis: an updated analysis. *The*
780 *Lancet Infect Dis*. 2017;17(8):873-81. PMID: 28483415.
781 23. Hagen F, Khayhan K, Theelen B, Kolecka A, Polacheck I, Sionov E, et al. Recognition of
782 seven species in the *Cryptococcus gattii*/*Cryptococcus neoformans* species complex. *Fungal*
783 *Genet Biol*. 2015;78:16-48. PMID: 25721988.
784 24. Farrer RA, Chang M, Davis MJ, van Dorp L, Yang D, Shea T, et al. A new lineage of
785 *Cryptococcus gattii* (VGV) discovered in the Central Zambezian Miombo Woodlands. *mBio*.
786 2019;10(6):e02306-19. PMID: 31719178.
787 25. Chayakulkeeree M, Perfect JR. Cryptococcosis. *Infect Dis Clin North Am*. 2006;20(3):507-
788 44. PMID: 16984867.
789 26. Cracraft J. Species concepts and speciation analysis. In: Johnston RF, editors. *Current*
790 *Ornithology*. New York, NY: Springer; 1983. p. 159-187.
791 27. Cracraft J. Species concepts and the ontology of evolution. *Biol Philos*. 1987;2:329-346.
792 doi:10.1007/BF00128837

793 28. Cronquist A. Once again, what is a species? In: Knutson LV, editor. *Biosystems in*
794 *Agriculture*. Montclair, NJ: Allanheld Osmun; 1978. p. 3-20.
795 29. Mayr E, Ashlock PD. *Principles of Systematic Zoology*. New York: McGraw-Hill; 1991.
796 30. Xu J, Vilgalys R, Mitchell TG. Multiple gene genealogies reveal recent dispersion and
797 hybridization in the human pathogenic fungus *Cryptococcus neoformans*. *Mol Ecol*.
798 2000;9(10):1471-81. PMID: 11050543.
799 31. Kavanaugh LA, Fraser JA, Dietrich FS. Recent evolution of the human pathogen
800 *Cryptococcus neoformans* by intervarietal transfer of a 14-gene fragment. *Mol Biol Evol*.
801 2004;23(10):1879-90. PMID: 16870684.
802 32. Loftus BJ, Fung E, Roncaglia P, Rowley D, Amedeo P, Bruno D, et al., The genome of the
803 basidiomycetous yeast and human pathogen *Cryptococcus neoformans*. *Science*.
804 2005;307(5713):1321-4. PMID: 15653466.
805 33. Janbon G, Ormerod KL, Paulet D, Byrnes EJ 3rd, Yadav V, Chatterjee G, et al. Analysis of
806 the genome and transcriptome of *Cryptococcus neoformans* var. *grubii* reveals complex RNA
807 expression and microevolution leading to virulence attenuation. *PLOS Genet*.
808 2014;10(4):e1004261. PMID: 24743168.
809 34. Martinez LR, Garcia-Rivera J, Casadevall A. *Cryptococcus neoformans* var. *neoformans*
810 (serotype D) strains are more susceptible to heat than *C. neoformans* var. *grubii* (serotype A)
811 strains. *J Clin Microbiol*. 2001;39(9):3365-7. PMID: 11526180.
812 35. Feldmesser M, Kress Y, Casadevall A. Dynamic changes in the morphology of *Cryptococcus*
813 *neoformans* during murine pulmonary infection. *Microbiology*. 2001;147(Pt 8):2355-65.
814 PMID: 11496012.

815 36. Desnos-Ollivier M, Patel S, Raoux-Barbot D, Heitman J, Dromer F, The French
816 Cryptococcosis Study Group. Cryptococcosis serotypes impact outcome and provide
817 evidence of *Cryptococcus neoformans* speciation. *mBio*. 2015;6(3):e00311-15. PMID:
818 26060271.
819 37. Casadevall A, Perfect JR. *Cryptococcus neoformans*. Washington, DC: ASM Press; 1998.
820 38. Dromer F, Mathoulin-Pelissier S, Launay O, Lortholary O, The French Cryptococcosis Study
821 Group. Determinants of disease presentation and outcome during cryptococcosis: the Crypto
822 A/D Study. *PLOS Med*. 2007;4(2):e21. PMID: 17284154.
823 39. Litvintseva AP, Kestenbaum L, Vilgalys R, Mitchell TG. Comparative analysis of
824 environmental and clinical populations of *Cryptococcus neoformans*. *J Clin Microbiol*.
825 2005;43(2):556-64. PMID: 15695645.
826 40. Cogliati M, Esposto MC, Clarke DL, Wickes BL, Viviani MA. Origin of *Cryptococcus*
827 *neoformans* var. *neoformans* diploid strains. *J Clin Microbiol*. 2001;39(11): 3889-94. PMID:
828 11682503.
829 41. Frases S, Ferrer C, Sanchez M, Colom-Valiente MF. Molecular epidemiology of isolates of
830 the *Cryptococcus neoformans* species complex from Spain. *Rev Iberoam Micol*.
831 2009;26(2):112-7. PMID: 19631160.
832 42. Maduro AP, Mansinho K, Teles F, Silva I, Meyer W, Martins ML, et al. Insights on the
833 genotype distribution among *Cryptococcus neoformans* and *C. gattii* Portuguese clinical
834 isolates. *Curr Microbiol*. 2014;68(2):199-203. PMID: 24077953.
835 43. Lengeler KB, Cox GM, Heitman J. Serotype AD strains of *Cryptococcus neoformans* are
836 diploid or aneuploid and are heterozygous at the mating-type locus. *Infect Immun*.
837 2001;69(1):115-22. PMID: 11119496.

838 44. Lin X, Litvintseva AP, Nielsen K, Patel S, Floyd A, Mitchell TG, et al. αADα hybrids of
839 *Cryptococcus neoformans*: evidence of same-sex mating in nature and hybrid fitness. *PLOS*
840 *Genet*. 2007;3(10):1975-90. PMID: 17953489.
841 45. Kwon-Chung KJ, Varma A. Do major species concepts support one, two or more species
842 within *Cryptococcus neoformans*. *FEMS Yeast Res*. 2006;6(4):574-87. PMID: 16696653.
843 46. Li W, Averette AF, Desnos-Ollivier M, Ni M, Dromer F, Heitman J. Genetic diversity and
844 genomic plasticity of *Cryptococcus neoformans* AD hybrid strains. *G3 (Bethesda)*.
845 2012;2(1):83-97. PMID: 22384385.
846 47. Sun S, Xu J. Genetic analyses of a hybrid cross between serotypes A and D strains of the
847 human pathogenic fungus *Cryptococcus neoformans*. *Genetics*. 2007;177(3):1475-1486.
848 PMID: 17947421.
849 48. Vogan A, Khankhet J, Xu J. Evidence for mitotic recombination within the basidia of a
850 hybrid cross of *Cryptococcus neoformans*. *PLOS One*. 2013;8(5):e62790. PMID: 23690954.
851 49. Samarasinghe H, Vogan A, Pum N, Xu J. Patterns of allele distribution in a hybrid
852 population of the *Cryptococcus neoformans* species complex. *Mycoses*. 2019;63(3):275-83.
853 PMID: 31774582.
854 50. Billmyre RB, Clancey SA, Heitman J. Natural mismatch repair mutations mediate
855 phenotypic diversity and drug resistance in *Cryptococcus deuterogattii*. *eLife* 2017;6:e28802.
856 PMID: 28948913.
857 51. Boyce KJ, Wang Y, Verma S, Shakya VPS, Xue C, Idnurm A. Mismatch repair of DNA
858 replication errors contributes to microevolution in the pathogenic fungus *Cryptococcus*
859 *neoformans*. *mBio*. 2017;8(3):e00595-17. PMID: 28559486.

860 52. Liu OW, Chun CD, Chow ED, Chen C, Madhani HD, Noble S. Systematic genetic analysis
861 of virulence in the human fungal pathogen *Cryptococcus neoformans*. *Cell*. 2008;135(1):174-
862 88. PMID: 18854164.
863 53. Tran HT, Keen JD, Kricker M, Resnick MA, Gordenin DA. Hypermutability of
864 homonucleotide runs in mismatch repair and DNA polymerase proofreading yeast mutants.
865 *Mol Cell Biol*. 1997;17(5):2859-65. PMID: 9111358.
866 54. Koren A, Ben-Aroya S, Steinlaur R, Kupiec M. Pitfalls of the synthetic lethality screen in
867 *Saccharomyces cerevisiae*: an improved design. *Curr Genet*. 2003;43(1):62-9. PMID:
868 12684846.
869 55. Paluszynski JP, Klassen R, Meinhardt F. Genetic prerequisites for additive or synergistic
870 actions of 5-fluorocytosine and fluconazole in baker's yeast. *Microbiology*. 2008;154(Pt
871 10):3154-64. PMID: 18832321.
872 56. Billmyre RB, Clancey SA, Xi LX, Doering TL, Heitman J. 5-fluorocytosine resistance is
873 associated with hypermutation and alterations in capsule biosynthesis in *Cryptococcus*. *Nat*
874 *Commun*. 2020;11(1):127. PMID: 31913284.
875 57. Sun S, Xu J. Chromosomal rearrangements between serotype A and D strains in
876 *Cryptococcus neoformans*. *PLOS One*. 2009.4(5):e5524. PMID: 19436753.
877 58. Sun S, Priest SJ, Heitman J. *Cryptococcus neoformans* mating and genetic crosses. *Curr*
878 *Protoc Microbiol*. 2019;53(1):e75. PMID: 30661293.
879 59. Alani E, Reenan RAG, Kolodner RD. Interaction between mismatch repair and genetic
880 recombination in *Saccharomyces cerevisiae*. *Genetics*. 1994;137(1):19-39. PMID: 8056309.

881 60. Alspaugh JA, Perfect JR, Heitman J. *Cryptococcus neoformans* mating and virulence are
882 regulated by the G-protein alpha subunit GPA1 and cAMP. *Genes Dev*. 1997;11(23):3206-
883 17. PMID: 9389652.
884 61. Idnurm A, Heitman J. Light controls growth and development via a conserved pathway in the
885 fungal kingdom. *PLOS Biol*. 2005;3(4):e95. PMID: 15760278.
886 62. Heitman J, Kozel TR, Kwon-Chung KJ, Perfect JR, Casadevall A, editors. *Cryptococcus:*
887 *From Human Pathogen to Model Yeast*. Washington, DC: ASM Press; 2011.
888 63. Sun S, Coelho MA, David-Palma MA, Priest SJ, Heitman J. The evolution of sexual
889 reproduction and the mating-type locus: links to pathogenesis of *Cryptococcus* human
890 pathogenic fungi. *Annu Rev Genet*. 2019;53:417-44. PMID: 31537103.
891 64. Greig D, Borts RH, Louis EJ, Travisano M. Epistasis and hybrid sterility in *Saccharomyces*.
892 *Proc Biol Sci*. 2002;269(1496):1167-71. PMID: 12061961.
893 65. Feretzaki M, Heitman J. Genetic circuits that govern bisexual and unisexual reproduction in
894 *Cryptococcus neoformans*. *PLOS Genet*. 2013;9(8):e1003688. PMID: 23966871.
895 66. Lin X, Hull CM, Heitman J. Sexual reproduction between partners of the same mating type
896 in *Cryptococcus neoformans*. *Nature*. 2005;434(7036):1017-21. PMID:15846346.
897 67. Vogan AA, Xu J. Evidence for genetic incompatibilities associated with post-zygotic
898 reproductive isolation in the human fungal pathogen *Cryptococcus neoformans*. *Genome*.
899 2014;57(6):335-44. PMID: 25187972.
900 68. Yadav V, Sun S, Coelho MA, Heitman J. Centromere scission drives chromosome shuffling
901 and reproductive isolation. *PNAS*. 2020;117(14):7917-28. PMID:32193338.

902 69. Chen W, Jinks-Robertson S. The role of the mismatch repair machinery in regulating mitotic
903 and meiotic recombination between diverged sequences in yeast. *Genetics*.
904 1999;151(4):1299-313. PMID:10101158.
905 70. Datta A, Adjiri A, New L, Crouse GF, Jinks-Roberston S. Mitotic crossovers between
906 diverged sequences are regulated by mismatch repair proteins in *Saccharomyces cerevisiae*.
907 *Mol Cell Biol*. 1996;16(3):1085-93. PMID: 8622653.
908 71. Datta A, Hendrix M, Lipsitch M, Jinks-Robertson S. Dual roles for DNA sequence identity
909 and the mismatch repair system in the regulation of mitotic crossing-over in yeast. *Proc Natl*
910 *Acad Sci U S A*. 1997;94(18):9757-62. PMID: 9275197.
911 72. Toffaletti DL, Rude TH, Johnston SA, Durack DT, Perfect JR. Gene transfer in *Cryptococcus*
912 *neoformans* by use of biolistic delivery of DNA. *J Bacteriol*. 1993;175(5):1405-11. PMID:
913 8444802.
914 73. Radchenko EA, McGinty RJ, Aksenova AY, Neil AJ, Mirkin SM. Quantitative analysis of
915 the rates for repeat-mediated genome instability in a yeast experimental system. *Methods Mol*
916 *Biol*. 2018;1672:421-38. PMID: 29043640.
917 74. Pitkin JW, Panaccione DG, Walton JD. A putative cyclic peptide efflux pump encoded by
918 the *TOXA* gene of the plant-pathogenic fungus *Cochliobolus carbonum. Microbiology*.
919 1996;142(Pt 6):1557-65. PMID: 8704997.
920 75. Kwon-Chung KJ, Edman JC, Wickes BL. Genetic association of mating types and virulence
921 in *Cryptococcus neoformans*. *Infect Immun*. 1992;60(2):602-5. PMID: 1730495.
922 76. Hua W, Vogan A, Xu J. Genotypic and phenotypic analyses of two "isogenic" strains of the
923 human fungal pathogen *Cryptococcus neoformans* var. *neoformans*. *Mycopathologia*.
924 2019;184(2):195-212. PMID: 30891668.

925 77. Grabherr MG, Russell P, Meyer M, Mauceli E, Alfoldi J, Di Palma F, et al. Genome-wide
926 synteny through highly sensitive sequence alignment: Satsuma. *Bioinformatics*.
927 2010;26(9):1145-51. PMID: 20208069.
928 78. Darling ACE, Mau B, Blattner FR, Perna NT. Mauve: multiple alignment of conserved
929 genomic sequence with rearrangements. *Genome Res*. 2004;14(7):1394-403. PMID:
930 15231754.
931 79. Cabanettes F, Klopp C. D-GENIES: dot plot large genomes in an interactive, efficient and
932 simple way. *PeerJ*. 2018;6:e4958. PMID: 29888139.
933 80. Li H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics*.
934 2018;34(18):3094-3100. PMID: 29750242.
935 81. Langdon QK, Peris D, Kyle B, Hittinger CT. spplDer: a species identification tool to
936 investigate hybrid genomes with high-throughput sequencing. *Mol Biol Evol*.
937 2018;35(11):2835-49. PMID: 30184140.
938 82. Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G, et al.
939 Integrative genomics viewer. *Nat Biotechnol*. 2011;29(1):24-6. PMID: 21221095.
940 83. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM.
941 *arXiv*:1303.3997v2 [q-bio.GN] [Preprint]. 2013 [cited 2020 May 11]. Available from:
942 arxiv.org/abs/1303.3997
943 84. DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, et al. A framework
944 for variation discovery and genotyping using next-generation DNA sequencing data. *Nat*
945 *Genet.* 2011;43(5):491-8. PMID: 21478889

946 85. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The
947 Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA
948 sequencing data. *Genome Res*. 2010;20(9):1297-303. PMID: 20644199.
949 86. Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, et al. A program for
950 annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. *Fly*
951 *(Austin)*. 2012;6(2):80-92. PMID: 22728672.
952 87. Bolger AM, Lohse M, Usadel B. Trimmomatic: A flexible trimmer for Illumina sequence
953 data. *Bioinformatics*. 2014;30(15):2114-20. PMID: 24695404.
954 88. Pegueroles C, Mixão V, Carreté L, Molina M, Gabaldón T. HaploTypo: A variant-calling
955 pipeline for phased genomes. *Bioinformatics*. 2019;36(8):2569-71. PMID: 31834373.
956 89. Li H. A statistical framework for SNP calling, mutation discovery, association mapping and
957 population genetical parameter estimation from sequencing data. *Bioinformatics*.
958 2011;27(21):2987-93. PMID: 21903627.
959 90. Quinlan AR, Hall IM. BEDTools: A flexible suite of utilities for comparing genomic
960 features. *Bioinformatics*. 2010;26(6):841-2. PMID: 20110278.
961 91. Idnurm A. Tetrad analysis of the basidiomycete fungus *Cryptococcus neoformans*. *Genetics*.
962 2010;185(1):153-63. PMID: 20157004.
963 92. Fraser JA, Huang JC, Pukkila-Worley R, Alspaugh JA, Mitchell TG, Heitman J.
964 Chromosomal translocation and segmental duplication in *Cryptococcus neoformans*.
965 *Eukaryot Cell.* 2005;4(2):401-406. PMID: 15701802.

966 ## Figure Captions

967 **Figure 1. Hypermutator phenotypes of JEC20a *msh2*Δ mutants. (A)** Fluctuation analysis on
968 YPD+rapamycin+FK506 medium was performed to quantify the mutation rates (number of
969 mutations per cell per generation) of four independent JEC20**a** *msh2*Δ mutants: *msh2*Δ-*1*,
970 *msh2*Δ-*2*, *msh2*Δ-*3*, and *msh2*Δ-*4*. The JEC20**a** parental strain in which the *msh2*Δ mutants were
971 constructed served as the negative control and an *msh2*Δ mutant in the KN99α genetic
972 background served as the positive control. Points indicate mean mutation rates and error bars
973 indicate 95% confidence intervals for the mean; 10 independent replicates of each strain were
974 included in mutation rate calculation. **(B)** Spectra of mutations identified through sequencing of
975 the *FRR1* gene in rapamycin+FK506-resistant colonies from fluctuation analysis conducted in
976 panel A. For JEC20**a** n = 9, for all other strains, n = 10. **(C)** Fluctuation analysis on YNB+5-
977 FOA medium was performed in a similar manner to the analysis in panel A. **(D-F)**
978 Representative images of plates used in papillation assays with independent JEC20**a** *msh2*Δ
979 mutants and a KN99α *fur1*Δ mutant on **(D)** YNB+5-FOA medium, **(E)** YNB+5FC medium, and
980 **(F)** YNB+5FU medium with the JEC20**a** parental strain and a KN99α *msh2*Δ mutant as controls.
981 Strains were incubated on YNB+5-FOA medium for 6 days at 30°C before imaging. Strains were
982 incubated on YNB+5FC and YNB+5FU media for 3 days at 30°C before imaging.
983

984 **Figure 2. Mating structure formation in *C. neoformans* x *C. deneoformans* genetic crosses**
985 **and germination frequencies of hybrid progeny. (A)** *C. neoformans* x C. *deneoformans*
986 genetic crosses produced robust hyphal filamentation after 6 days of incubation in dark
987 conditions on MS medium. Scale bars in colony border images (left image of each set) represent
988 200 µm. Scale bars in basidia and basidiospore morphology images (right image of each set)

989 represent 50 µm **(B)** Average frequencies of basidia lacking basidiospores, or bald basidia, in *C.*
990 *neoformans* x *C. deneoformans* genetic crosses. Error bars represent standard deviation. **(C)**
991 Mean germination frequencies of *C. neoformans* x *C. deneoformans* hybrid progeny in wild-
992 type, unilateral, and bilateral *msh2*Δ genetic crosses. Following random basidiospore dissection
993 and micromanipulation onto YPD agar medium, progeny were allowed to germinate for up to
994 three weeks before germination frequencies were assessed. X-axis labels are the same for (B)
995 and (C).
996
997 **Figure 3. Unique phenotypes of *C. neoformans* x *C. deneoformans* hybrid progeny.** Sexual
998 reproduction structures, including hyphae, basidia, and basidiospores, produced by *C.*
999 *neoformans* x *C. deneoformans* hybrid progeny derived from bilateral *msh2*Δ genetic crosses
1000 following dissection and micromanipulation onto YPD agar medium. Sexual structures were
1001 formed on YPD agar plates sealed with parafilm and incubated in light conditions at room
1002 temperature. From left to right, scale bars on microscopy images represent 100 µm, 50 µm, and
1003 10 µm.
1004
1005 **Figure 4. Nuclear genome composition of *C. neoformans* x *C. deneoformans* hybrid progeny**
1006 **reveals substantial aneuploidy and recombination.** Representative sequencing read-depth
1007 coverage and inheritance patterns of progeny derived from **(A)** a wild-type H99α x JEC20**a**
1008 genetic cross, **(B)** a unilateral H99α x JEC20**a** *msh2*Δ genetic cross, and **(C)** a bilateral KN99α
1009 *msh2*Δ x JEC20**a** *msh2*Δ genetic cross. For each progeny, sequencing coverage plots (normalized
1010 to the genome-wide average coverage) are colored according to each parental species
1011 contribution as shown in the key on the top right, and a schematic representation of the inferred

1012 karyotype is depicted on the right. Homeologous chromosomes are color coded based on the H99
1013 reference and asterisks in JEC21 indicate chromosomes in reverse-complement orientation (See
1014 Fig S4 for details). Red arrowheads mark recombination breakpoints between homeologous
1015 chromosomes and/or loss of heterozygosity (also highlighted by red boxes in the karyotype
1016 panels). Circular black labels: (a) indicate changes in ploidy in a subset of the population of cells
1017 that were sequenced; (b) and (c) mark, respectively, chromosome breaks at a transposable
1018 element near the end of the *C. deneoformans* Chr6 and at the centromere of Chr9, which were
1019 both repaired by *de novo* telomere addition (green arrowheads). Note the chromosome order of
1020 each parent; JEC21α contigs have been reordered to maximize collinearity with the H99α
1021 contigs.
1022
1023 **Figure 5. Distribution of the number of recombination events detected in the hybrid**
1024 **progeny based on read-pairs aligning to different chromosomes.** Four different filtering
1025 thresholds were applied to detect potential instances of recombination across the genomes of 27
1026 total *C. neoformans* x *C. deneoformans* hybrid progeny: (A) Mapping quality (MQ) = 60 and at
1027 least 15 read-pairs (Cov) supporting each event; (B) MQ = 60 and at least 5 read-pairs supporting
1028 each event; (C) MQ >= 0 and at least 15 read-pairs supporting each event; and (D) MQ >= 0 and
1029 at least 5 read-pairs supporting each event. Each point represents the number of recombination
1030 events detected across the whole genome of a single hybrid progeny. In the box and whisker
1031 plots, red lines represent the median, shaded boxes represent the interquartile ranges (IQRs),
1032 upper and lower whiskers show the largest or smallest observations, respectively, that lie 1.5 *
1033 IQR beyond the upper and lower quartiles, respectively. Outliers are included.
1034

1035 ## Supporting information captions

1036 **Table S1. Germination frequencies of hybrid progeny from individual *C. neoformans* x *C.***
1037 ***deneoformans* genetic crosses.**

1038

1039 **Table S2. (A)Heterozygosity levels of the hybrid progeny, and (B) Kruskal-Wallis and**
1040 **Dunn's statistical tests.**

1041

1042 **Table S3. Recombination events detected in *C. neoformans* x *C. deneoformans* hybrid**
1043 **progeny.**

1044

1045 **Table S4. Strains used in this study.**

1046

1047 **Table S5. Oligonucleotides used in this study.**

1048

1049 **Table S6. NCBI data submissions related to this study.**

1050

1051 **Figure S1. Genetic deletion and confirmation of the *MSH2* gene in JEC20a. (A)** Genetic
1052 deletion mutants lacking *MSH2* were engineered in JEC20**a** by replacing the *MSH2* open reading
1053 frame (ORF) with the dominant drug resistance marker *NEO* which confers resistance to
1054 nourseothricin via biolistic transformation. 5' and 3' UTRs are depicted as yellow boxes. Arrows
1055 depict locations of primers used to generate and verify *MSH2* deletion. The sets of primers
1056 denoted by gray arrows located upstream and downstream of the *MSH2* ORF were used to
1057 amplify flanking sequences homologous to the JEC20**a** *MSH2* endogenous locus to mediate

1058 homologous recombination (JOHE45551, 45552 and JOHE4555, 45556, respectively). Yellow
1059 arrows depict primers (JOHE45553 and JOHE45554) that amplified the *NAT* cassette and share
1060 homology with JOHE45552 and JOHE45556, respectively. The red arrows (JOHE45559,45560)
1061 and blue arrows (JOHE45822,45823) indicate primers that confirmed integration of the deletion
1062 allele and loss of the *MSH2* ORF, respectively. Gray lines indicate syntenic regions shared
1063 between the deletion allele and endogenous locus. **(B)** Gel electrophoresis of PCR products was
1064 used to confirm integration of a single copy of the *NAT* gene at the correct locus and that the
1065 wild-type *MSH2* gene was absent in the JEC20**a** mutant strains. (Spanning: JOHE45559,45560;
1066 WT in-gene: JOHE45822,45823; WT 5' junction: JOHE45559,45823; Δ5' junction:
1067 JOHE45559;45554; WT 3' junction: JOHE45822,45560; Δ3' junction: JOHE45553,45560).
1068
1069 **Figure S2. PCR analysis of mating-type allele inheritance in *C. neoformans* x *C.***
1070 ***deneoformans* hybrid progeny.** Sequence-specific primers for the mating-type (*MAT*) locus
1071 gene *STE20*, which can differentiate between *C. neoformans MAT***a**, *C. neoformans MAT*α, *C.*
1072 *deneoformans MAT***a**, and *C. deneoformans MAT*α, were used to characterize which *MAT* alleles
1073 each of the hybrid progeny inherited.
1074
1075 **Figure S3. FACS analysis of *C. neoformans* x *C. deneoformans* hybrid progeny.** The *C.*
1076 *neoformans* strain H99α was used as a 1N haploid control and the *C. deneoformans* XL143 strain
1077 was used as a diploid 2N control.
1078
1079 **Figure S4. Self-filamentation in *C. neoformans* x *C. deneoformans* hybrid progeny**.
1080 Filamentation of hybrid progeny selected for whole-genome sequencing on MS medium after

1081 incubation for 14 days. AI187 is a self-filamentous, stable diploid *C. neoformans* strain and
1082 served as a positive control. Scale bars represent 100 µm.
1083
1084 **Figure S5. Whole-genome comparisons of *C. neoformans* H99α and *C. deneoformans***
1085 **JEC21α strains. (A)** Workflow to assess the genomic contribution of each parental species in
1086 the hybrid progeny. The sppIDer pipeline uses short-read sequencing data and a combined
1087 genome built from reference genomes of the two parental *Cryptococcus* species. **(B)** Dot-plot
1088 comparing the H99α assembly with the JEC21α reordered and reoriented assembly. Red and blue
1089 lines represent sequence similarities in the same and reverse orientations, respectively. **(C)**
1090 Linear plots showing overall synteny between the H99α and JEC21α genomes. The
1091 chromosomal positions of centromeres and the *MAT* locus are indicated by black and yellow
1092 bars, respectively. Chromosomes of JEC21α are color coded based on their synteny with the
1093 H99α chromosomes. Three major gross chromosomal changes previously documented
1094 distinguishing the two strains correspond to color changes within the same chromosome: TR
1095 indicates a reciprocal chromosomal translocation; INT indicates an introgression of a 14-gene
1096 region from *C. neoformans* to *C. deneoformans* that was mediated by transposable elements
1097 common to both lineages [31]; and SD indicates a segmental duplication following a
1098 nonreciprocal translocation involving the subtelomeric regions of JEC21 chromosomes 8 and 12
1099 that presumably occurred during the construction of the congenic strain pair JEC21α/JEC20**a**
1100 [92]. Chromosomal inversions are not indicated expect for a large inversion on Chr3 of JEC21α
1101 (dashed arrow).
1102

1103 **Figure S6. Nuclear genome composition of additional *C. neoformans* x *C. deneoformans***
1104 **hybrid progeny. (A)** H99α x JEC20**a**. **(B)** H99α x JEC20**a** *msh2*Δ-*1* **(C)** KN99α *msh2*Δ x
1105 JEC20**a (D)** H99α *msh2*Δ x JEC20**a** *msh2*Δ-*1*. For each progeny, read-depth plots (normalized to
1106 the genome-wide average coverage) are colored according to each parental species contribution
1107 as shown in the key on the top right, and a schematic representation of the inferred karyotype is
1108 depicted on the right. Homeologous chromosomes are color coded based on the H99 reference
1109 (see Fig S4 for details) and asterisks in JEC21α indicate chromosomes in reverse-complement
1110 orientation. Red arrowheads mark recombination breakpoints between homeologous
1111 chromosomes and/or loss of heterozygosity (also highlighted by red boxes in the karyotype
1112 panels). Where detected, the breakpoints of additional recombination events within the same
1113 chromosome are indicated by light blue arrowheads. Circular black labels: (a) marks changes in
1114 ploidy in a subset of the population of cells that were sequenced; (b) marks chromosome
1115 breakage events repaired by *de novo* telomere addition (see Fig S6 for details); (c) indicates
1116 recombination events next to the *MAT* locus; (d) marks a break at rDNA locus; (e) marks a
1117 complex chromosomal aberration that cannot be explained by simple rearrangements and
1118 required further investigation.
1119
1120 **Figure S7. Quantification of heterozygosity across genomes of *C. neoformans* x *C.***
1121 ***deneoformans* hybrid progeny.** Plot showing the percentage of heterozygosity for each
1122 individual progeny (represented by different dots) grouped by type of cross (see Table S2 for
1123 details). The horizontal red line depicts the mean heterozygosity values. Genomes of progeny
1124 derived from bilateral *msh2*Δ × *msh2*Δ mutant crosses were significantly more heterozygous

1125 only when compared to hybrid progeny derived from wild-type H99α x JEC20**a** crosses

1126 (Statistical analysis: Kruskal-Wallis test, followed by Dunn's test, $p<0.05$).

1127

1128 **Figure S8. Chromosome breaks in the hybrid progeny can be repaired by *de novo* telomere**

1129 **addition.** Breakpoints were detected in different chromosomal locations, including *CEN14* and

1130 *CEN9* of JEC21α (A and B), a T1 transposable element located at the end of Chr6 of JEC21α

1131 (C), or in other genomic locations (D – F). Each panel shows the result of read mapping for one

1132 or more progeny that underwent chromosome breakage and healing via *de novo* telomere

1133 addition (strains names in boldface type) and a control strain in which no breaks were detected

1134 on the same region (strain names in normal font type). Breakage and *de novo* telomere addition

1135 was inferred, respectively, by abrupt changes in read coverage (depicted as bars on the top and

1136 colored as shown in the key) and by the presence of reads with telomeric repeats at the

1137 breakpoints. When two copies of the same chromosome are present, only a subset of reads are

1138 expected to contain telomeric repeats (as shown e.g. in panel F).

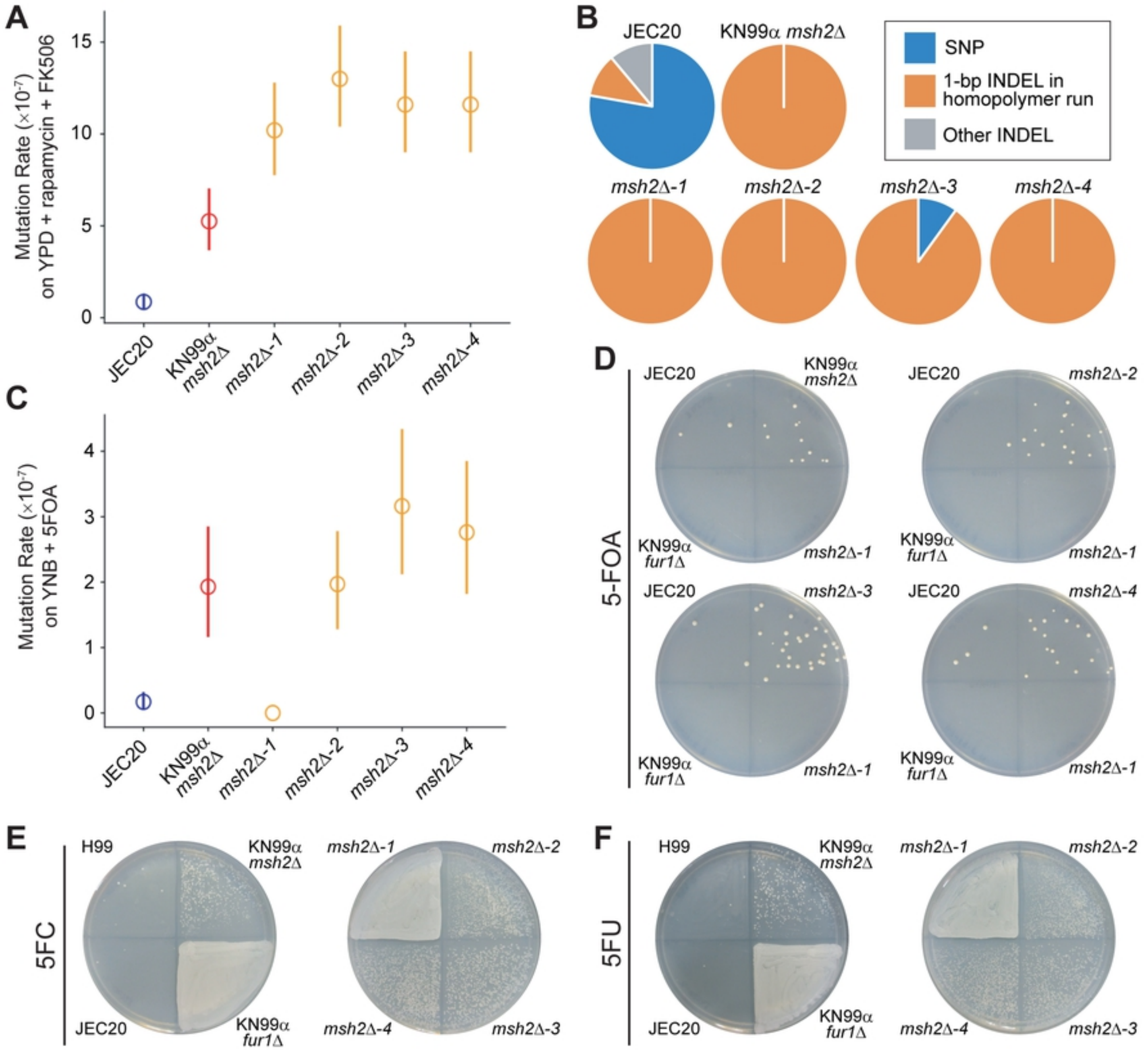
A
Mutation Rate (×10⁻⁷) on YPD + rapamycin + FK506
15
10
5
0
JEC20
KN99α msh2Δ
msh2Δ-1
msh2Δ-2
msh2Δ-3
msh2Δ-4
B
JEC20
KN99α msh2Δ
SNP
1-bp INDEL in homopolymer run
Other INDEL
msh2Δ-1
msh2Δ-2
msh2Δ-3
msh2Δ-4
C
Mutation Rate (×10⁻⁷) on YNB + 5FOA
4
3
2
1
0
JEC20
KN99α msh2Δ
msh2Δ-1
msh2Δ-2
msh2Δ-3
msh2Δ-4
D
5-FOA
JEC20
KN99α msh2Δ
JEC20
msh2Δ-2
KN99α fur1Δ
msh2Δ-1
KN99α fur1Δ
msh2Δ-1
JEC20
msh2Δ-3
JEC20
msh2Δ-4
KN99α fur1Δ
msh2Δ-1
KN99α fur1Δ
msh2Δ-1
E
5FC
H99
KN99α msh2Δ
msh2Δ-1
msh2Δ-2
JEC20
KN99α fur1Δ
msh2Δ-4
msh2Δ-3
F
5FU
H99
KN99α msh2Δ
msh2Δ-1
msh2Δ-2
JEC20
KN99α fur1Δ
msh2Δ-4
msh2Δ-3

Figure 1

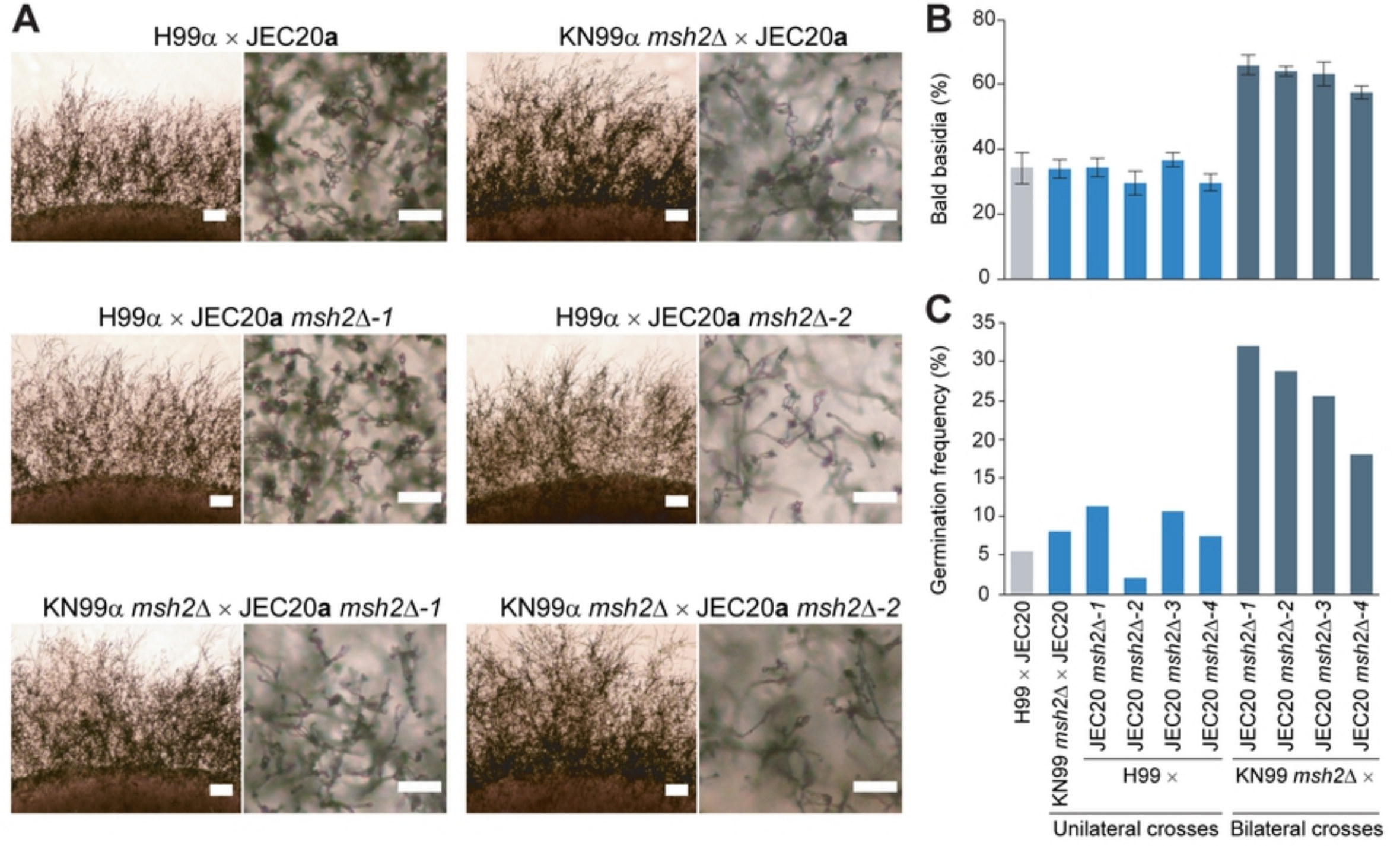
A
H99α × JEC20a
KN99α msh2Δ × JEC20a
H99α × JEC20a msh2Δ-1
H99α × JEC20a msh2Δ-2
KN99α msh2Δ × JEC20a msh2Δ-1
KN99α msh2Δ × JEC20a msh2Δ-2
B
Bald basidia (%)
0
20
40
60
80
C
Germination frequency (%)
0
5
10
15
20
25
30
35
H99 × JEC20
KN99 msh2Δ × JEC20
JEC20 msh2Δ-1
JEC20 msh2Δ-2
JEC20 msh2Δ-3
JEC20 msh2Δ-4
H99 ×
KN99 msh2Δ ×
Unilateral crosses
Bilateral crosses

Figure 2

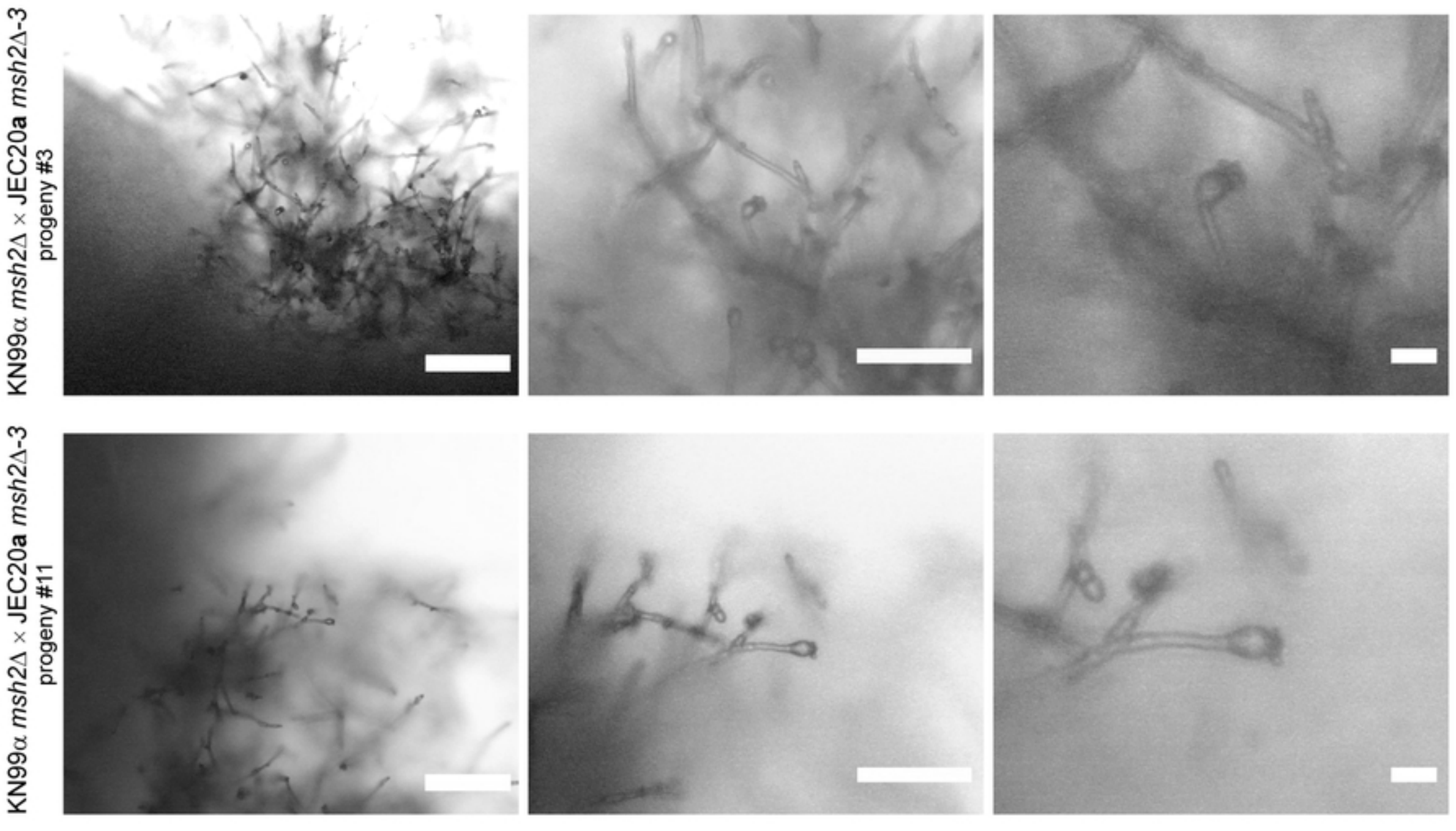
KN99α msh2Δ × JEC20a msh2Δ-3
progeny #3
KN99α msh2Δ × JEC20a msh2Δ-3
progeny #11

Figure 3

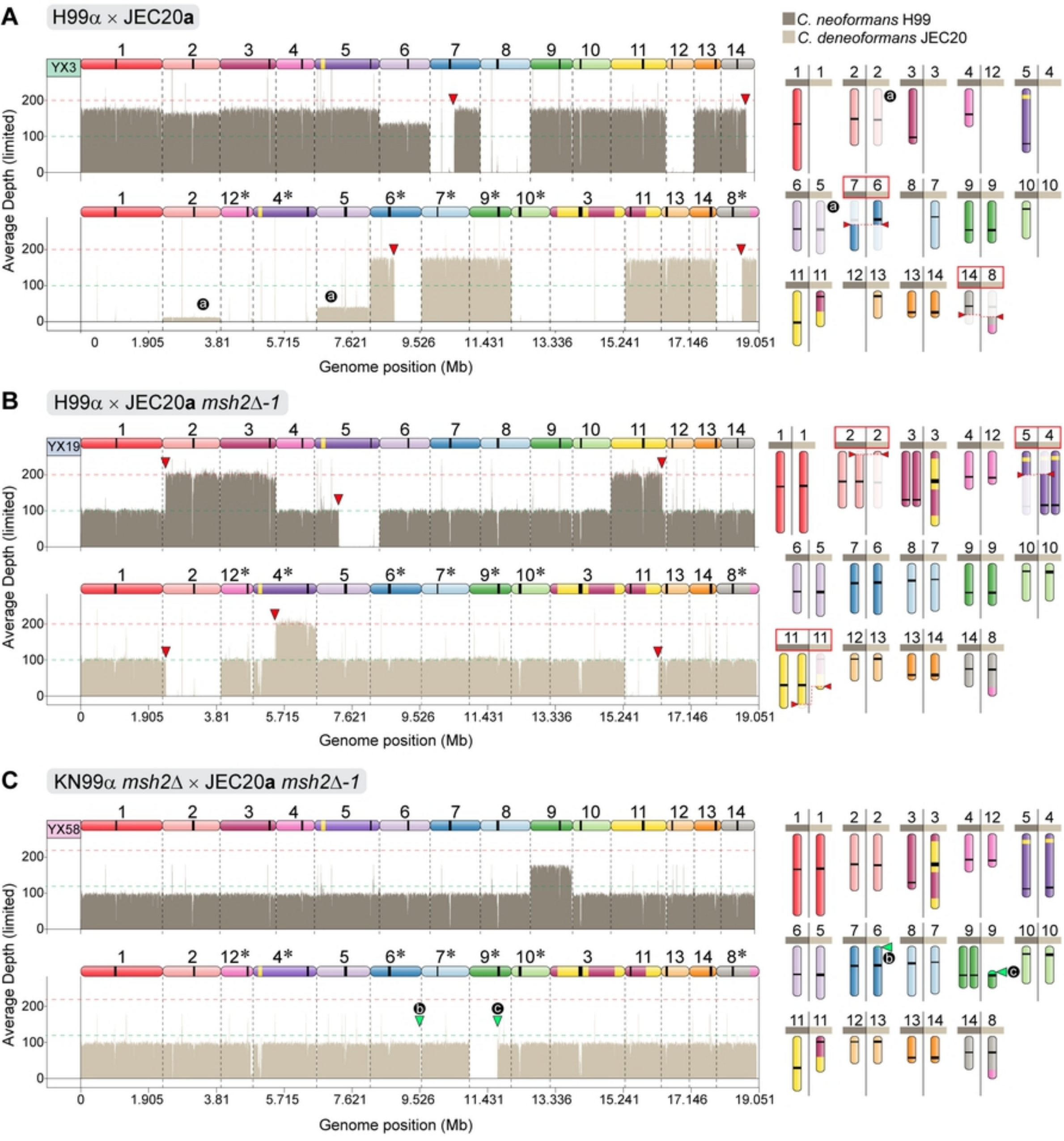
A
H99α × JEC20a
C. neoformans H99
C. deneoformans JEC20
YX3
Average Depth (limited)
Genome position (Mb)
B
H99α × JEC20a msh2Δ-1
YX19
C
KN99α msh2Δ × JEC20a msh2Δ-1
YX58

Figure 4

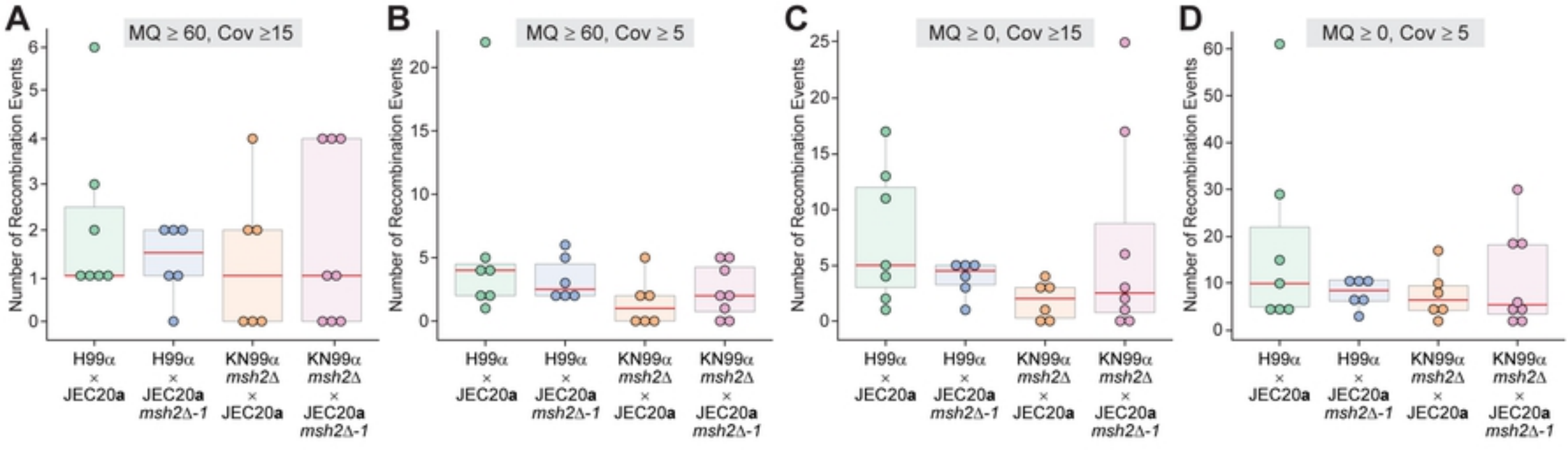
A
MQ ≥ 60, Cov ≥15
B
MQ ≥ 60, Cov ≥ 5
C
MQ ≥ 0, Cov ≥15
D
MQ ≥ 0, Cov ≥ 5
Number of Recombination Events
H99α × JEC20a
H99α × JEC20a msh2Δ-1
KN99α msh2Δ × JEC20a
KN99α msh2Δ × JEC20a msh2Δ-1

Figure 5