Hybridization drives mitochondrial DNA degeneration and metabolic shift in a species with biparental mitochondrial inheritance

Mathieu Hénault[1245]*, Souhir Marsit[12345#a], Guillaume Charron[1345#b] and Christian R. Landry[12345]

[1]Institut de Biologie Intégrative et des Systèmes (IBIS), Université Laval, Québec, QC, Canada

[2]Département de biochimie, microbiologie et bioinformatique, Université Laval, Québec, QC, Canada

[3]Département de biologie, Université Laval, Québec, QC, Canada

[4]Quebec Network for Research on Protein Function, Engineering, and Applications (PROTEO), Université Laval, Québec, QC, Canada

[5]Université Laval Big Data Research Center (BDRC_UL), Québec, QC, Canada

[#a]Current address: Département de biologie, chimie et géographie, Université du Québec à Rimouski, Rimouski, QC, Canada

[#b]Current address: Centre de foresterie des Laurentides, Ressources naturelles Canada, Québec, QC, Canada

*Corresponding author

E-mail: mathieu.henault.1@ulaval.ca

ABSTRACT

Mitochondrial DNA (mtDNA) is a cytoplasmic genome that is essential for respiratory metabolism. While uniparental maternal mtDNA inheritance is most common in animals and plants, distinct mtDNA haplotypes can coexist in a state of heteroplasmy, either because of paternal leakage or de novo mutations. MtDNA integrity and the resolution of heteroplasmy have important implications, notably for mitochondrial genetic disorders, genome evolution and speciation. However, knowledge on the factors driving heteroplasmy resolution is sparse. Here, we use *Saccharomyces* yeasts, fungi with constitutive biparental mtDNA inheritance, to investigate the resolution of mtDNA heteroplasmy in a variety of hybrid genotypes. We analyzed mtDNA sequence evolution and growth phenotypes in a collection of 864 hybrid yeast lines that were experimentally evolved under relaxed selection for mitochondrial function, allowing us to explore the near-neutral evolution of mtDNAs. The lines were subdivided into 11 different hybrid crosses among natural lineages of *Saccharomyces paradoxus* and its sister species *Saccharomyces cerevisiae*, yielding a variety of hybrid genotypes along a gradient of parental divergence. We observed extensive mtDNA recombination, but recombination rates were not predicted by parental divergence and were genotype-specific. However, we found a strong positive relationship between parental divergence and the rate of large-scale mtDNA deletions, which lead to the loss of respiratory metabolism. We also uncovered associations between mtDNA recombination, deletion, and genome instability that were genotype-specific. Our results show that hybridization in yeast induces mtDNA degeneration, with deep consequences for mtDNA evolution, metabolism and the emergence of reproductive isolation between species.

INTRODUCTION

Mitochondrial DNAs (mtDNAs) are genomes that are maintained and expressed in mitochondria, which are ATP-producing organelles shared by virtually all Eukaryotes (Müller et al. 2012). Although mtDNA content varies greatly across Eukaryotes, all encode genes involved in respiration and protein synthesis (Adams and Palmer 2003). Most animals and plants have uniparental maternal mtDNAs inheritance. Many exceptions exist to this rule, either because of constitutive uniparental or biparental inheritance (Breton et al. 2015; Wilson and Xu 2012; Wang et al. 2015), or leakage of paternal mtDNA (Parakatselaki and Ladoukakis 2021). Leakage and biparental mtDNA inheritance lead to a state of heteroplasmy, which is the intracellular coexistence of multiple mtDNA haplotypes. Evidence for paternal leakage is growing, even in species long considered to have strict uniparental maternal inheritance (Ladoukakis and Zouros 2017; Kondo et al. 1990; Payne et al. 2013; Gyllensten et al. 1991). Heteroplasmy may have important implications for mtDNA evolution, notably by enabling recombination between otherwise clonal mtDNAs (Rokas et al. 2003; Tsaousis et al. 2005; Ladoukakis et al. 2011). In humans, heteroplasmy is determinant in the onset of mitochondrial genetic disease, which are linked to inherited or de novo mtDNA variants (Stewart and Chinnery 2015; Wallace 2015; Stewart and Chinnery 2021). Additionally, when diverged populations or species hybridize, heteroplasmy can lead to the coexistence of more or less diverged mtDNAs which can then interact.

Genes encoded on mtDNAs are necessary, but not sufficient for carrying their essential metabolic function. Interactions between gene products encoded by mtDNAs and nuclear genomes imply the requirement for tight coevolution (Burton and Barreto 2012; Piccinini et al. 2021). Mitonuclear coevolution has important implications in the context of heteroplasmy in hybrids. Because mtDNA replication and segregation are non-mendelian, heteroplasmy is generally transient and a single mtDNA haplotype is eventually fixed by vegetative segregation

3

66 (Birky 2001). The return to homoplasmy fixes a set of mtDNA alleles belonging to either parental
67 haplotype, or a mix of parental alleles in the case of recombination (Rokas et al. 2003; Tsaousis
68 et al. 2005). Heteroplasmy resolution can lead to the fixation of alleles involved in mitonuclear
69 incompatibilities, thus compromising mitochondrial function and hybrid fitness. In addition,
70 mitochondrial alleles can interact epistatically within recombinant mtDNAs (mito-mito epistasis)
71 (Wolters et al. 2018).

72 Fundamental properties of hybrids may determine the outcome of heteroplasmy resolution. One
73 general prediction is that the potential for epistatic interactions will scale with the level of
74 evolutionary divergence between the parents (Orr 1995). As such, mitonuclear incompatibilities
75 are expected to be more frequent between distantly related genomes compared to closely
76 related ones (Burton 2022). Under the same rationale, mito-mito incompatibilities are expected
77 to show increased prevalence between more diverged mtDNAs. Nucleotide sequence
78 divergence between mtDNAs is also expected to reduce the opportunities for homologous
79 recombination. Little is known about the path to heteroplasmy resolution in hybrids and its
80 consequences for hybrid evolution. Notably, how parental divergence shapes the landscape of
81 mtDNA recombination is unknown. Additionally, whether parental divergence favors the loss of
82 mitochondrial function in hybrids and, if so, by which mechanism, remains unexplored.
83 Answering these questions requires a model system in which heteroplasmy is dominant and
84 that can easily generate a large diversity of F1 hybrid genotypes.

85 Like many other fungi (Wilson and Xu 2012), yeasts of the *Saccharomyces* genus have
86 biparental mtDNA inheritance (Solieri 2010). This enables the inheritance of either parental
87 mtDNA in F1 hybrids, as well as recombination between parental haplotypes (Dujon et al.
88 1974). Because yeast F1 hybrids are initially heteroplasmic and quickly transition to
89 homoplasmy by mitotic vegetative segregation (Solieri 2010), they are a powerful system to
90 investigate the factors driving heteroplasmy resolution. In addition, respiratory metabolism in

91 yeast is facultative, enabling the study of neutral mtDNA evolution without the selective pressure
92 to maintain respiration. The undomesticated species *Saccharomyces paradoxus* has natural
93 lineages with wide variation in genetic diversity and evolutionary divergence (Kuehne et al.
94 2007; Charron et al. 2014a; Xia et al. 2017; Leducq et al. 2016), allowing laboratory crosses that
95 span various levels of parental divergence.

96 Mutation accumulation (MA) is a type of experimental evolution that minimizes the power of
97 natural selection. Combined with genome sequencing, MA experiments produce near-unbiased
98 estimates of the spectrum of genomic changes (Lynch et al. 2016). We previously performed a
99 large-scale MA experiment on yeast hybrids (Charron et al. 2019; Hénault et al. 2020). We
100 generated 864 hybrid MA lines from independent mating events, which were subdivided into
101 eleven crosses (with 48 to 96 independent lines each). Parental strains comprised 11 wild
102 isolates of *S. paradoxus* and two wild isolates of its sister species *Saccharomyces cerevisiae*,
103 spanning a range of evolutionary divergence from intra-lineage to interspecific (Fig. 1A). We
104 submitted each line to periodical extreme bottlenecks by streaking for single colonies on solid
105 medium (Joseph and Hall 2004; Lynch et al. 2008) (Fig. 1B) until they reached ~770 mitotic
106 generations (Charron et al. 2019). Here, we use this collection and the associated whole-
107 genome sequencing datasets (Hénault et al. 2020; Marsit et al. 2021) to investigate the
108 resolution of heteroplasmy in a variety of hybrid genotypes.

109 **RESULTS**

110 *Parental strains of the MA experiment display extensive variation in mtDNA content*

111 We sequenced the genome of the 13 parental strains (Supplemental Table S1) using Oxford
112 Nanopore long reads to produce high-quality reference assemblies of their mtDNAs
113 (Supplemental Fig. S1). The annotation of these assemblies revealed variation in mtDNA
114 content, mostly in the presence of introns in the *COB* and *COX1* genes (Fig. 1C). MtDNA

content was consistent within lineages, although important diversity was found in *SpB*. To enable comparative analyses based on a common reference sequence, the most complete genome (LL2012_028) was used as a template to build an artificial reference genome by adding missing introns at the orthologous splice sites.

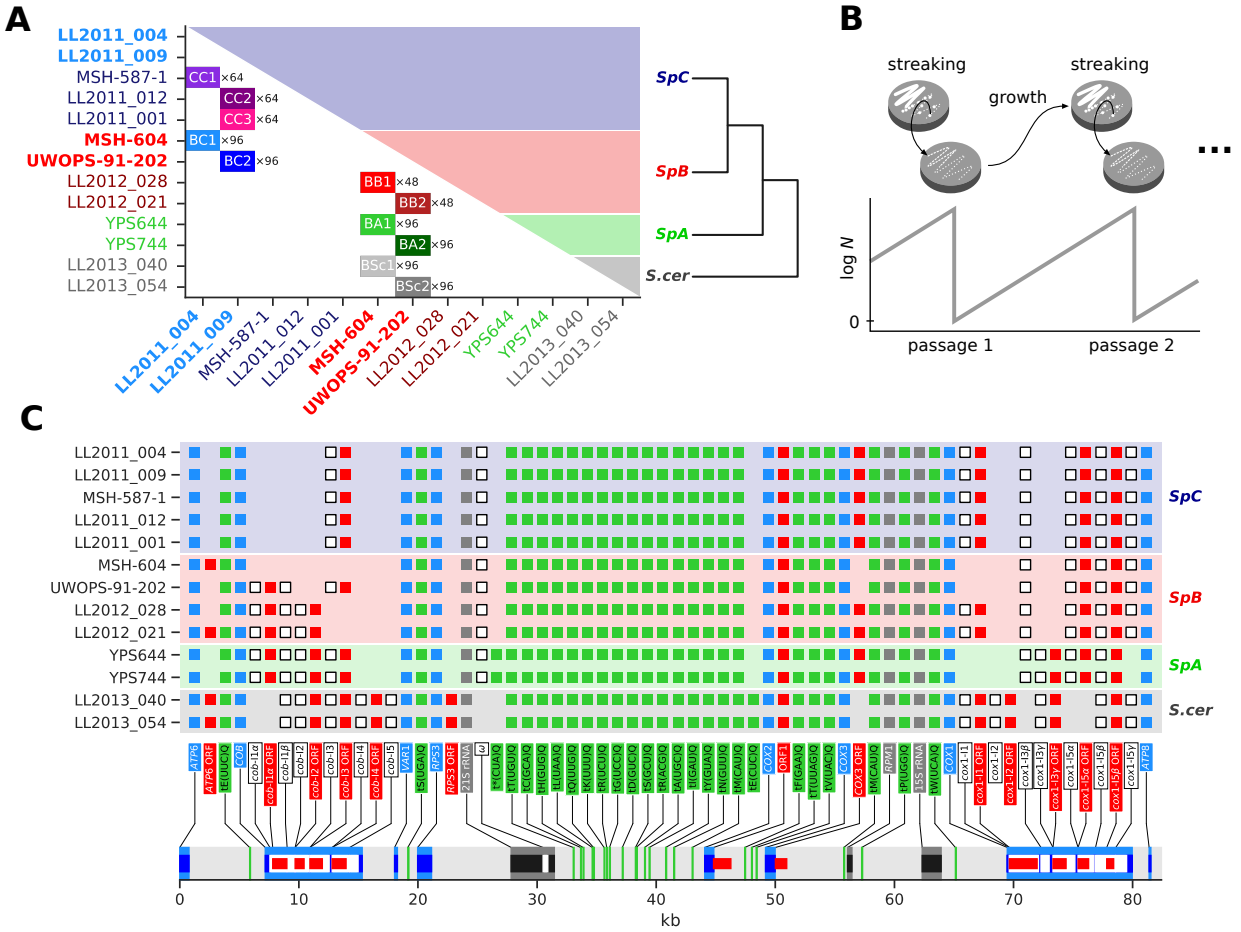


Figure 1. Experimental design of the MA experiment and mtDNA contents of the parental strains. (*A*) Cross scheme for the generation of 864 independent MA lines subdivided into 11 crosses spanning intra-lineage (*SpC*×*SpC*, CC; *SpB*×*SpB*, BB), intraspecific (*SpB*×*SpC*, BC; *SpB*×*SpA*, BA) and interspecific (*SpB*×*S. cerevisiae*, BSc) parental divergence. Each cross was replicated to found 48 to 96 independent evolution lines, all arising from distinct mating events. (*B*) Microbial MA experiments minimize the efficiency of natural selection. Periodical extreme bottlenecks are achieved through streaking for single colonies on solid medium, which amplifies genetic drift. (*C*) Annotation summary of high-quality reference mtDNA assemblies for each parental strain. Squares denote the presence of each feature in the corresponding genome. Light blue, protein-coding genes; dark blue, protein-coding exons; white, introns; grey, RNA-coding genes; black, RNA-coding exons; red, intronic or free-standing ORFs; green, tRNAs.

130 *The landscape of mtDNA recombination is not predicted by parental divergence*

131 For each cross, 36 to 47 lines (average 41) were previously selected at random for short-read
132 whole-genome sequencing at the initial and final timepoints of the MA experiment (Hénault et al.
133 2020; Marsit et al. 2021) (Supplemental Table S2). For each cross, we identified a set of
134 confident mtDNA variants and used them as markers to discriminate parental haplotypes. We
135 then used alleles at marker positions to identify the recombination tracts for individual MA lines,
136 yielding 482 independent mtDNA haplotypes. For all lines at both timepoints, minor allele
137 frequencies were consistent with homoplasmy, with the vast majority of minor alleles below
138 0.5% (Supplemental Fig. S2). MtDNA haplotypes revealed that lines from all cross types
139 experienced some extent of recombination (Fig. 2A; Supplemental Fig. S3-13). Only BA1
140 exhibited no recombination event across all the sequenced lines. Some crosses exhibited
141 recombination hotspots, yielding high local similarity among haplotypes (Fig. 2B).

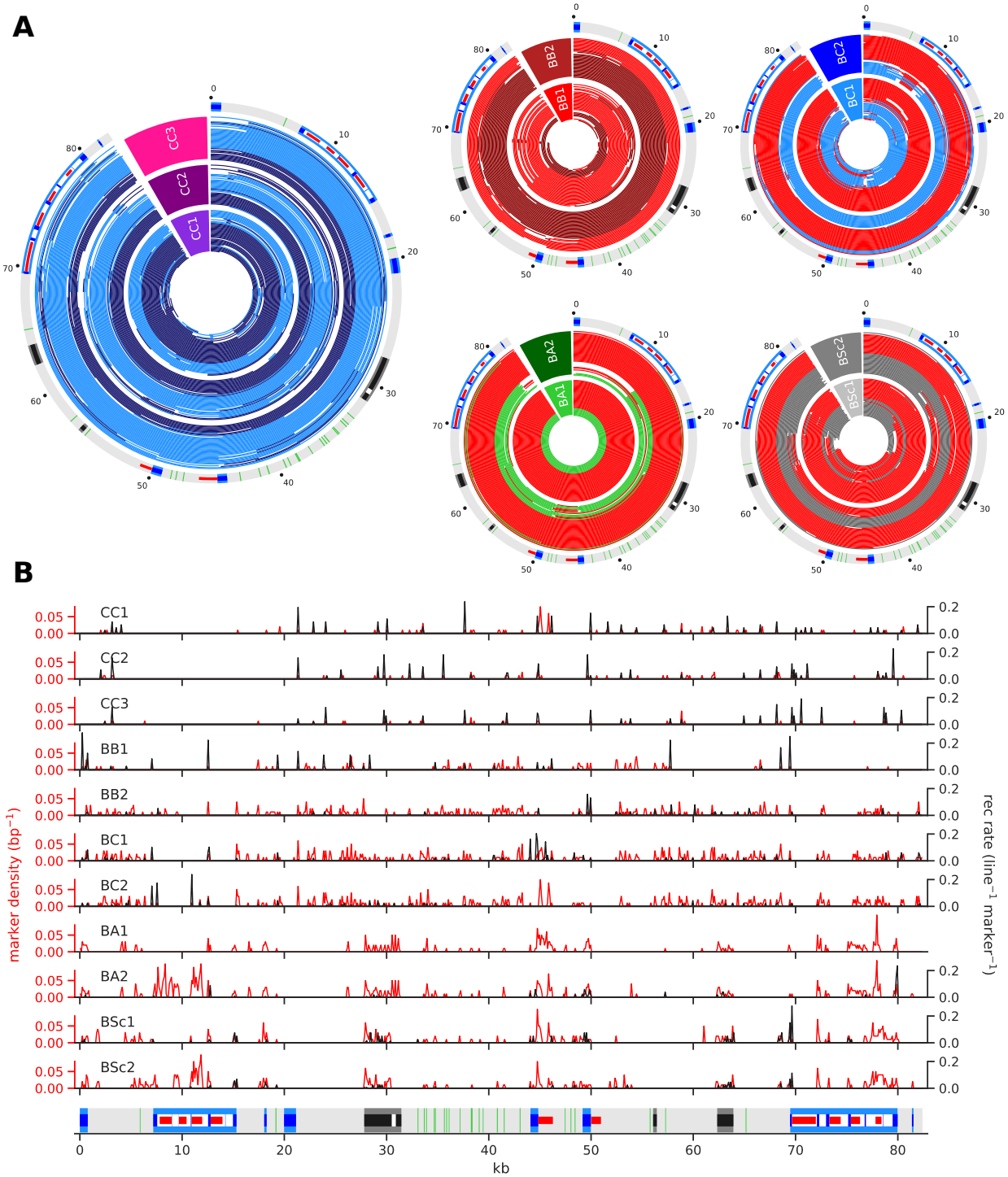


Figure 2. Recombination is pervasive between parental mtDNA haplotypes of hybrid MA lines. (*A*) Stacks show recombination tracts colored by parental haplotype (Fig 1A). Outer circles show annotation summary as in Fig 1C, with coordinates in kb. (*B*) Density of markers (red) and recombination junctions (black) in 100 bp-wide windows. Recombination junction density is normalized by marker density. Annotation summary corresponds to Fig. 1C.

8

147 CC2 and BC1 crosses harbored the mtDNA haplotypes with the highest count of recombination
148 tracts (Fig. 3A). Recombination rate displayed no significant trend with evolutionary divergence
149 (based on nuclear variants) between the parents of a cross (Fig. 3B). Notably, the interspecific
150 BSc crosses had recombination rates similar to intraspecific crosses. Excluding BSc crosses did
151 not yield better correlations, suggesting an absence of relationship with evolutionary
152 divergence, even at the intraspecific level. The average length of recombination tracts was
153 higher for BSc crosses and had a significant positive relationship with parental divergence (Fig.
154 3C). Neither recombination rate nor tract length correlated with marker counts, suggesting that
155 these estimates are largely unaffected by parental marker density. This indicates that mtDNA
156 recombination rate is independent of the global evolutionary divergence level between the
157 parents of a hybrid, contrary to the expectation that nucleotide sequence similarity should
158 enable recombination.

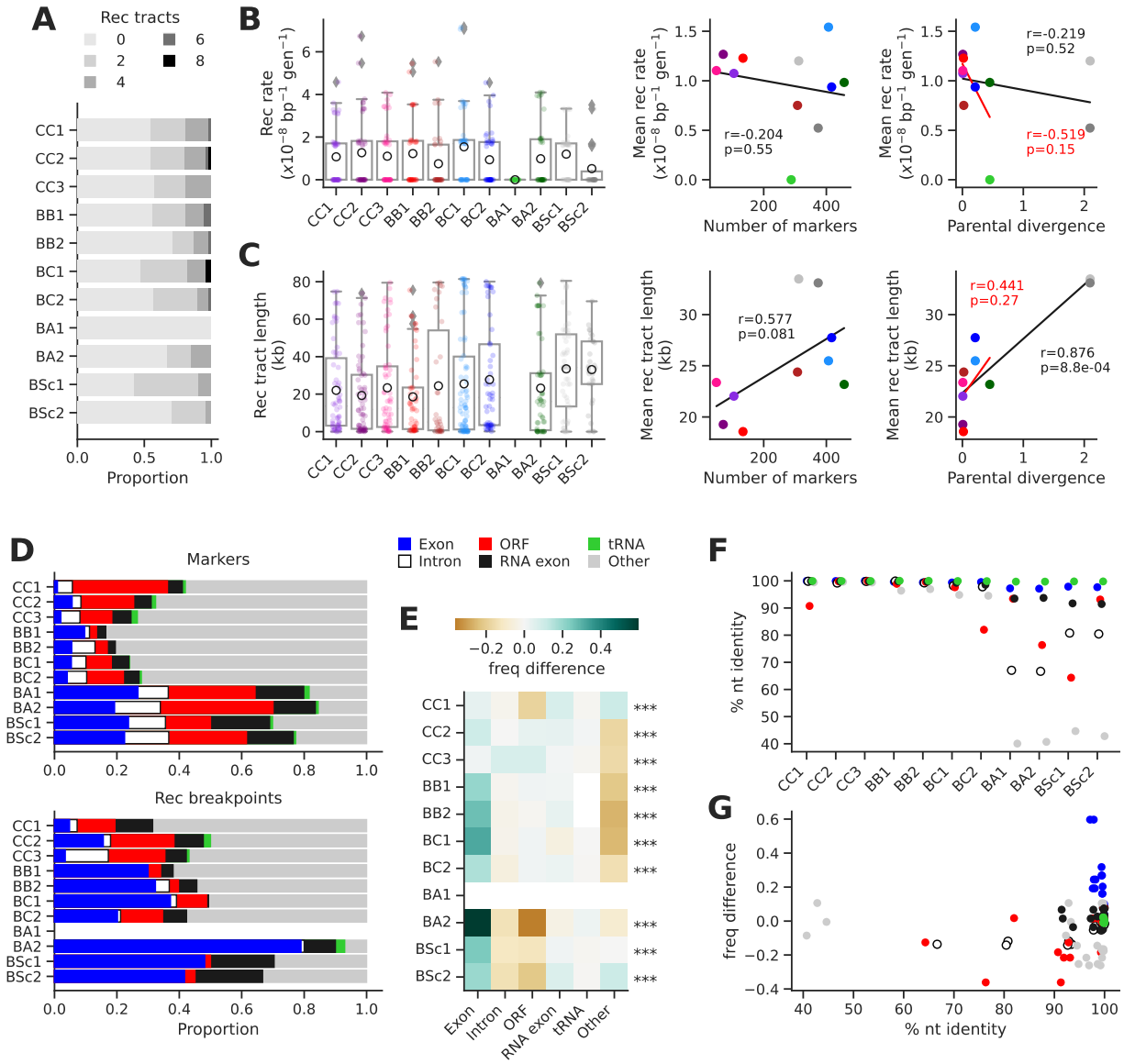


Figure 3. MtDNA recombination breakpoints are enriched for loci corresponding to protein-coding exons. (*A*) Counts of recombination tracts per line. (*B*) Left: recombination rate distributions for individual lines. Black circles show average recombination rates. Right: Pearson's correlations between mean recombination rate and number of markers or parental evolutionary divergence. The correlation in red excludes BSc crosses. (*C*) Left: length distributions for individual recombination tracts. Black circles show average recombination tract lengths. Right: Pearson's correlations between mean recombination tract length and number of markers or parental evolutionary divergence. The correlation in red excludes BSc crosses. (*D*) Distributions of annotation feature types for markers (top) and recombination breakpoints (bottom). (*E*) Frequency difference between breakpoints and markers for each feature type. FDR-corrected p-values of chi-squared tests for frequency deviations within each cross are shown (***: $p \leq 0.001$). (*F*) Percentage of nucleotide identity between parental mtDNAs for each annotation category. (*G*) Frequency difference against nucleotide identity.

10

172 We classified recombination breakpoints by annotation feature type and contrasted the resulting
173 distributions with parental markers (Fig. 3D). We found a global overrepresentation of protein-
174 coding exons in recombination breakpoints through all crosses (Fig. 3E). While protein-coding
175 exons retain high nucleotide sequence identity in all crosses (Fig. 3F), it does not solely explain
176 their overrepresentation, as other annotation types displayed equally high sequence identity
177 (Fig. 3G). In summary, recombination between mtDNAs occurs preferentially near protein-
178 coding exon sequences, regardless of the parental divergence of the crosses.

179 We asked whether the overrepresentation of exons can be explained by adjacent intron
180 mobilization. We examined the depth of coverage at intron-exon junctions, which show
181 distinctive patterns in the presence or absence of an intron. We identified seven putative
182 mobility events of *cox1*-I1 in BB1 and one putative mobility event of *cob*-I3 in BC1
183 (Supplemental Fig. S14). A total of 15 recombination junctions were consistent with
184 mobilization-mediated recombination (Supplemental Fig. S15). Notably, *cox1*-I1 accounts for six
185 of the eight putative recombination-associated mobilization events in BB crosses. Although
186 many intron presence/absence polymorphisms were consistent with mobilization-associated
187 recombination, mobilization potentially explained only a minor fraction of the total recombination
188 junctions (7/50 in BB1, 1/34 in BB2, 3/78 in BC1, 3/54 in BSc1 and 1/30 in BSc2).

189 MtDNA recombination may yield biases in the inheritance of parental alleles. The average
190 mtDNA-wide inheritance ratio for each cross did not correlate with parental mtDNA copy number
191 ratios estimated from depth of coverage of long-read sequencing libraries (Supplemental Fig.
192 S16). However, copy number estimates may have low accuracy since mtDNA abundance
193 depends on the growth phase (Galeota-Sprung et al. 2021) and we did not optimize the DNA
194 extraction cultures for synchronicity. They nevertheless suggest that parental mtDNA copy
195 numbers do not determine inheritance probability. We found significant deviations from mtDNA-
196 wide inheritance ratios for some genes (Supplemental Fig. S17). Similar patterns were found in

197 the two BSc crosses, with segments centered on *COX2* being mostly inherited from the *SpB*
198 parent.

199 *Parental divergence predicts the frequency of loss of respiratory function and large-scale*
200 *mtDNA deletions*

201 We next characterized the loss of respiratory function in MA lines. The loss of respiration in
202 yeast is most often explained by the loss of mtDNA integrity through large-scale deletions
203 (Bernardi 1979). We measured the growth of all MA lines on media containing fermentable
204 (glucose, YPD) and non-fermentable (glycerol and ethanol, YPEG) carbon sources using high-
205 density colony array imaging. Since respiratory effects of mtDNA haplotypes or mitonuclear
206 interactions are often revealed with exposure to high temperatures (Li et al. 2019; Paliwal et al.
207 2014), growth was measured at both 25°C and 37°C. Growth on YPEG revealed differences
208 between cross types. Many CC lines became thermosensitive, while many BC lines exhibited
209 growth advantages at either 25°C or 37°C (Supplemental Fig. S18). We classified lines as non-
210 respiring if no growth was scored on YPEG at both temperatures (Supplemental Fig. S19). The
211 frequency of respiration loss was higher for more divergent crosses, with BSc being one order
212 of magnitude higher than CC crosses (Fig. 4A). We quantified the depth of sequencing
213 coverage along the mtDNA sequence and classified lines as harboring complete mtDNAs or
214 mtDNAs with deletions (Supplemental Fig. S20-30). All parental strains had complete mtDNAs.
215 The frequencies of mtDNAs with deletions closely matched the frequencies of respiratory loss
216 (Fig. 4B) and had a significant positive relationship with parental divergence (Pearson's
217 correlation, r=0.88, p-value=3×10^{-4}, Fig. 4C). For most crosses, deletions in mtDNAs had
218 significantly negative effects on growth in YPD (Fig. 4D), consistent with loss of respiratory
219 ability in *S. cerevisiae* (Ephrussi et al. 1949; Vowinckel et al. 2021). In contrast, deletions were
220 associated with complete loss of growth in YPEG (Fig. 4D).

221 Focusing on respiring lines, we looked at the effect of culture medium and evolution time (final
222 vs initial timepoint) on growth. A linear mixed effect model with random intercepts for crosses
223 showed significant global negative effect of YPEG versus YPD, as well as significant pairwise
224 interactions between carbon source, temperature and time (Supplemental Table S3). Linear
225 models fitted individually for each cross recapitulated a globally negative effect of YPEG versus
226 YPD on growth (Supplemental Fig. S31). Many terms exhibited patterns consistent with cross
227 types. For instance, the growth of less divergent CC crosses generally decreased with time,
228 while the more divergent crosses BC, BA and BSc increased (Supplemental Fig. S31,32).

229 We asked if mtDNA haplotypes impacted the growth phenotypes of respiring lines. The only
230 difference was found between non-recombinant *SpB* and *SpA* mtDNAs in the BA1 cross on
231 YPEG at 37°C (Supplemental Fig. S33,34), with *SpB* having a significant growth advantage
232 (Mann-Whitney U test p-value=1.56×10^{-6}, FDR-corrected). Lines with *SpB* mtDNAs also tended
233 to grow better on YPEG at 37°C in the BC2 cross, which has many of the largest effect sizes
234 (Supplemental Fig. S34), hinting towards a similar thermotolerance phenotype against *SpC* and
235 recombinant mtDNAs. These results suggest that mitochondrial genetic variation is linked to
236 thermal and respiratory adaptations in wild *S. paradoxus*.

13

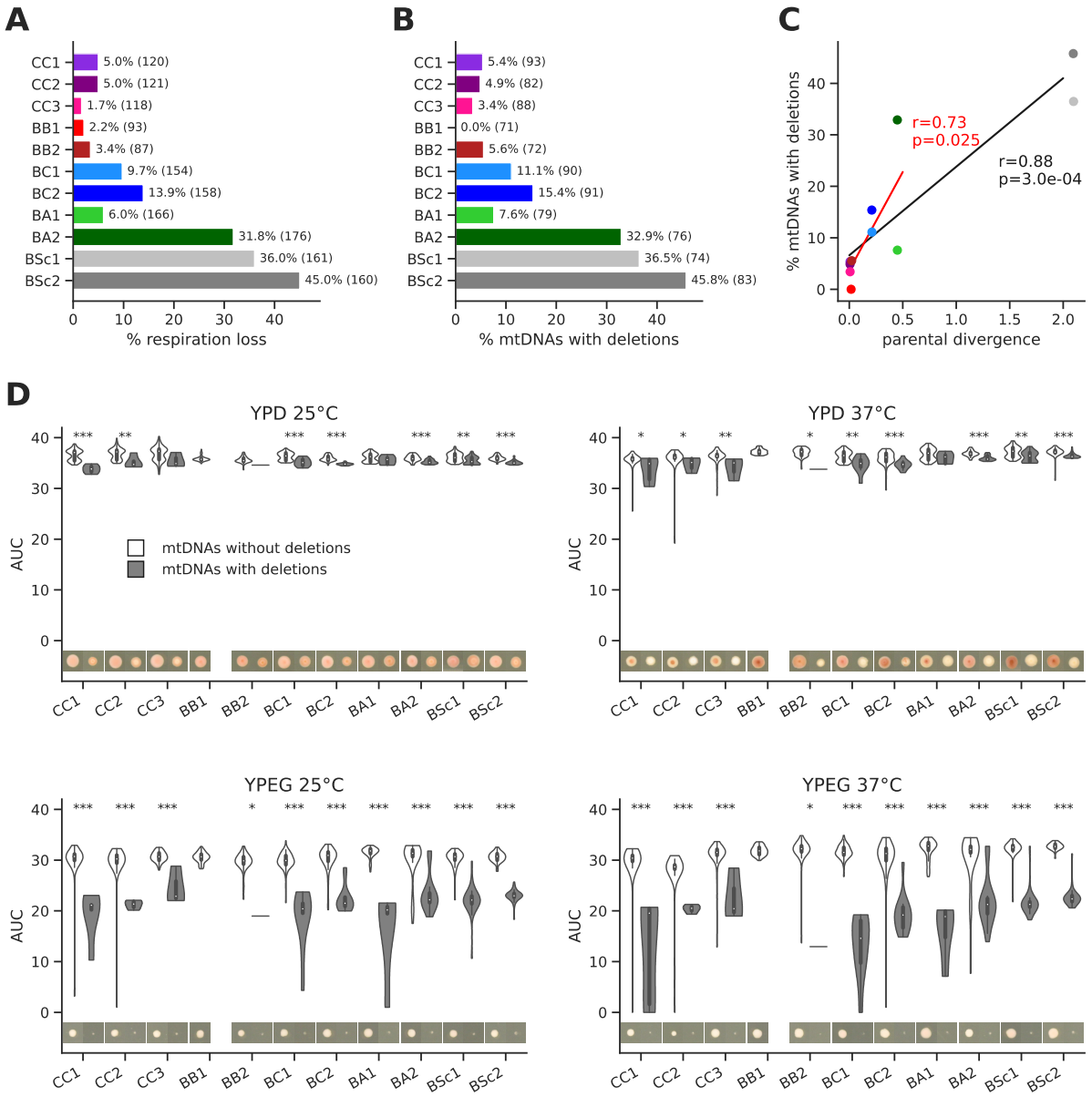


Figure 4. Frequent deletions in mtDNAs lead to loss of respiratory function and occur preferentially in crosses with higher parental divergence. (*A*) Percentage of all MA lines with loss of respiration. Counts of lines (at initial and final timepoints combined) are shown. (*B*) Percentage of sequenced MA lines per cross with mtDNA deletions. Counts of lines (at initial and final timepoints combined) are shown. (*C*) Pearson's correlation between percentage of mtDNAs with deletions and parental divergence for each cross. The correlation shown in red excludes BSc crosses. (*D*) Area under growth curves (AUC) for lines without and with mtDNA deletions in the four tested conditions. FDR-corrected p-values for Mann-Whitney U tests between mtDNAs without and with deletions are shown. Images show the growth of representative lines (the closest to the median of each distribution) after four days.

We next investigated which mtDNA segments were mostly involved in deletions. We compared depth of coverage profiles along mtDNAs and found recurring patterns among mtDNAs of non-

250 respiring lines (Fig. 5). Notably, many BSc haplotypes comprised only *COX2* and *COX3*. This
251 recurrent pattern emerged at two levels, among independently evolved lines and in the
252 independent crosses BSc1 and BSc2. These results suggest that mtDNA deletions may be
253 predictable and repeatable in hybrids, and that high frequencies of deletions are partly
254 explained by deletion hotspots.

255 Depth of coverage profiles from short read mappings are highly heterogeneous along the
256 mtDNA sequence (Supplemental Fig. S20-30), making it difficult to identify deletion breakpoints
257 with precision. We randomly selected 122 lines from the final timepoint for long-read sequencing
258 and assembled their mtDNAs. 15 haplotypes harbored large-scale deletions, and most deletion
259 breakpoints were in unannotated regions (Supplemental Fig. S35, Supplemental Table S4). This
260 result indicates that the types of sequences involved in mtDNA deletions are different from
261 mtDNA homologous recombination, suggesting distinct mechanisms.

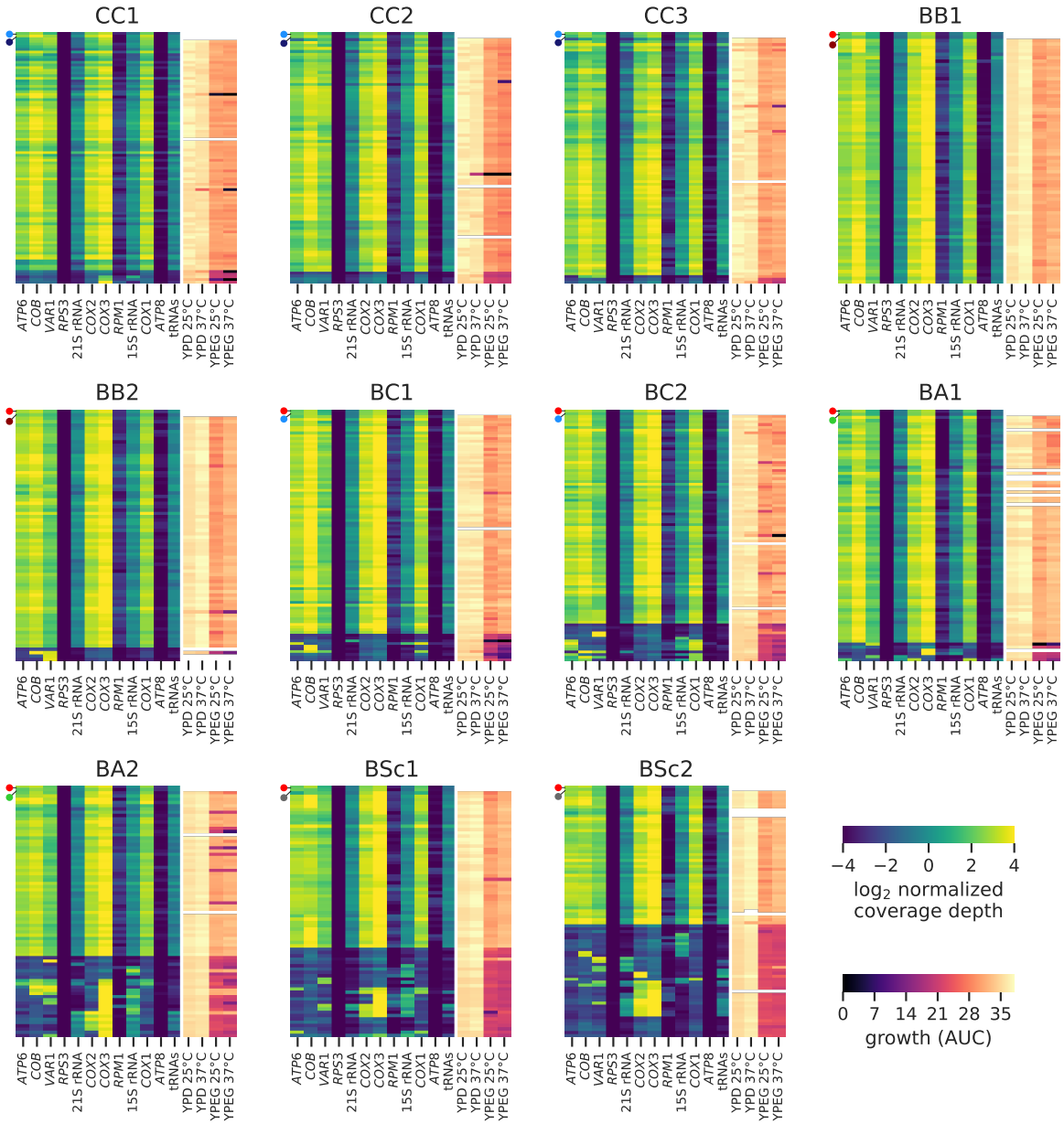


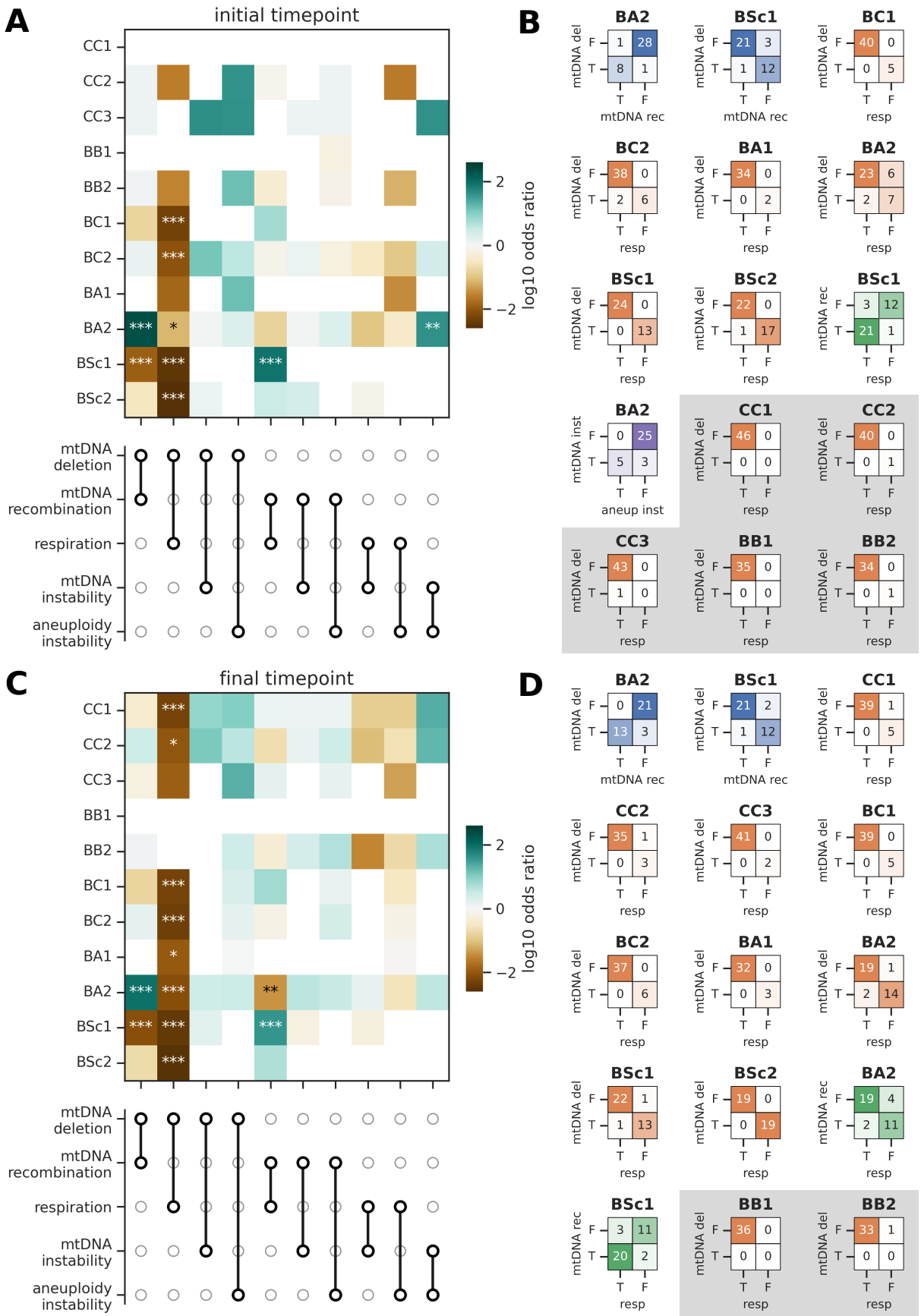
Figure 5. Recurrent mtDNA deletion patterns emerge in non-respiring lines. The left part of each heatmap shows normalized depth of coverage grouped by mtDNA features. The right part of each heatmap shows the growth AUC in the four conditions tested. Rows represent individual MA lines (including both initial and final timepoints) grouped by cross. Lines are hierarchically clustered by depth of coverage profile similarity. Top lines represent the parental strains.

Aspects of mtDNA evolution show contrasted associations between MA crosses

We next characterized the associations between mtDNA recombination, mtDNA deletion and loss of respiratory ability. All three phenotypes were observed in lines sampled at the initial timepoint, indicating that they can occur shortly after F1 zygotes are formed. We also observed

276 distinct haplotypes sampled through time from the same line, indicating that heteroplasmy still
277 existed in the lines at the initial timepoint. We thus analyzed the associations at initial and final
278 timepoints separately. As expected, there was a strong negative relationship between mtDNA
279 deletions and respiration (Fig. 6A,C). There were also significant associations between mtDNA
280 deletion and recombination in BA2 and BSc1. However, the sign of the association was different
281 in the two cases. While recombinant mtDNA haplotypes were more likely to harbor deletions in
282 BA2, only one BSc1 haplotype showed both recombination and deletion (Fig. 6B,D).

17


A
initial timepoint
CC1
CC2
CC3
BB1
BB2
BC1
BC2
BA1
BA2
BSc1
BSc2
mtDNA deletion
mtDNA recombination
respiration
mtDNA instability
aneuploidy instability
log10 odds ratio
B
BA2
BSc1
BC1
BC2
BA1
BA2
BSc1
BSc2
BSc1
BA2
CC1
CC2
CC3
BB1
BB2
mtDNA del
mtDNA rec
mtDNA inst
resp
aneup inst
C
final timepoint
D

18

Figure 6. Associations between aspects of mtDNA evolution in MA lines. (*A*) Associations for lines at the initial timepoint, grouped by cross. The heatmap shows the odds ratio for each pair of variables among mtDNA deletion, mtDNA recombination, respiration, mtDNA instability and aneuploidy instability. FDR-corrected p-values of Fisher's exact test, *: p≤0.05, **: p≤0.01, ***: p≤0.001. (*B*) Counts of MA lines for each comparison with a statistically significant association, classified as true (T) or false (F) for the given binary variable. The shaded background comprises associations between mtDNA deletion and respiration that were not statistically significant. (*C*,*D*) Same as A and B, but for lines at the final timepoint.

While most MA lines had the expected negative association between mtDNA deletions and respiration, we observed many exceptions. Lines with non-deleted mtDNAs could harbor other mutations (mitochondrial or nuclear) leading to respiratory loss. However, mtDNAs with large deletions (which are incompatible with respiration) were found in six and three respiring lines at the initial and final timepoints, respectively (Fig. 6B,D). Moreover, loss of respiration without deletion was found in six lines at the initial timepoint, all in BA2, which are unlikely explained by de novo mutations given the short evolution time. These inconsistencies may be explained by differences between the cryopreserved archive stocks used for genome sequencing, and the copies of those stocks used for phenotyping. Spot assays in the four conditions revealed that both stocks exhibited the same phenotypes (Supplemental Fig. S36), thus ruling out an effect of stock propagation.

We looked for genomic changes other than mtDNA deletions that could have occurred rapidly in MA line evolution and explain respiratory loss. We searched for de novo nucleotide variants in mtDNAs of the MA lines by looking for confident variants that are private to single lines. We found a handful of candidate de novo mutations at the initial timepoint but none were in coding regions (Supplemental Table S5), making them poor candidates to explain respiratory loss of function. We next focused on aneuploidies in the nuclear genome, which were previously shown to arise early in MA (Fijarczyk et al. 2021; Marsit et al. 2021). Several inconsistent lines had aneuploidies, including five BA2 lines with monosomies for chromosomes I, III, VI and IX and/or trisomies for chromosomes IV and XII (Supplemental Fig. S37). However, our reanalysis of

313 parental allele frequencies revealed that all closely matched the diploid expectation of 50%. We
314 quantified the deviation in depth of coverage from the value expected given the copy number
315 call for each chromosome. We validated that these deviations were not artifacts from variation in
316 library size (Supplemental Fig. S38). Many lines with aneuploidy calls showed abnormally high
317 deviations from the expectation of true aneuploidies (Supplemental Fig. S37A). Continuous
318 variation in depth of coverage and symmetrical parental allele frequencies are consistent with
319 unbiased, segregating aneuploidies rather than fixed aneuploidies. These aneuploidies arose
320 from single euploid clones in the population expansion leading to DNA extraction. We thus
321 named this phenotype aneuploidy instability.

322 Focusing on BA2 at the initial timepoint, we assessed mtDNA presence by PCR at two loci in
323 the *ATP6* and 21S rRNA genes for all lines (Supplemental Table S6). While globally consistent
324 with sequencing libraries, both PCR amplicons were detected in the two respiring lines
325 genotyped as having mtDNA deletions (Supplemental Fig. S37). Furthermore, at least one
326 amplicon was missing for five out of six non-respiring lines with complete mtDNAs. This
327 suggests that minor mtDNA variants segregating in the archive stocks were selected by chance
328 when single clones were isolated for genome sequencing. We named the inconsistency
329 between mitochondrial deletion and respiration mtDNA instability. BA2 had the highest
330 proportion of lines with mtDNA instability, and those were primarily found at the initial timepoint
331 (Supplemental Fig. S39A). We found a statistically significant association of mtDNA instability
332 with aneuploidy instability in BA2 at the initial timepoint (Fig. 6B). Out of eight lines displaying
333 mtDNA instability, five also were among the lines with the highest aneuploidy instability. There
334 was a similar trend in other crosses, although only BA2 was statistically significant
335 (Supplemental Fig. S39B). This suggests that a proportion of hybrid lines, specifically in the BA2
336 cross, suffered from high genomic instability, manifested by both frequent aneuploidy and
337 mtDNA deletion.

338 **DISCUSSION**

339 How diverged mtDNA haplotypes interact to shape the resolution of heteroplasmy and the
340 evolution of mitochondrial functions remains vastly unexplored. In this study, we used
341 *Saccharomyces* yeasts as a model to probe the neutral evolution of mtDNAs in heteroplasmic
342 hybrids through experimental evolution by MA. We uncovered that recombination rates are
343 independent of parental divergence, and recombination occurs preferentially near protein-
344 coding exons. On the other hand, parental divergence is a strong predictor of the rate of large-
345 scale mtDNA deletions, which is the primary mechanism for loss of respiratory metabolism. We
346 also uncover that some crosses have contrasted associations between mtDNA recombination
347 and deletion, as well as shared genomic instability at the level of mtDNAs and nuclear
348 aneuploidies.

349 Recombination between mtDNAs in heteroplasmic *S. cerevisiae* has been known for decades
350 (Dujon et al. 1974). Recombination is likely central in the replication of mtDNAs, making it an
351 inherent part of mitochondrial genetics rather than an accidental byproduct of heteroplasmy
352 (Dujon 2020). Genomic studies brought further support for the idea that mtDNA recombination is
353 pervasive (Fritsch et al. 2014; Bágeľová Poláková et al. 2021). Fritsch and collaborators
354 characterized recombination in pools of diploid F1 hybrids from intraspecific crosses of *S.*
355 *cerevisiae*, and found that recombination hotspots were concentrated in tRNA clusters and
356 rRNA genes (Fritsch et al. 2014). Bágeľová Poláková and collaborators investigated three
357 recombinant mtDNA haplotypes in F1 hybrids between *S. cerevisiae* and *S. paradoxus*, and
358 attributed most of the observed recombination events to the mobilization of introns and free-
359 standing ORFs (Bágeľová Poláková et al. 2021). Thus, our study is the first large-scale report of
360 mtDNA recombination at the intra- and inter-specific levels, allowing the resolution of 482
361 independent mtDNA haplotypes. Our data confirms that mtDNA recombination is pervasive in a
362 variety of hybrid genotypes. In contrast to previous studies, we found an association of

363 recombination junctions with protein-coding exons and only a minor potential contribution for
364 intron mobilization.

365 We observed associations between mtDNA recombination and deletion that suggest a
366 mechanistic interaction, although this remains to be investigated. The most contrasting patterns
367 were between BA2 and BSc1, two crosses in which mtDNA deletions occurred frequently. In
368 BA2, haplotypes harboring deletions were almost always recombinant, and non-deleted
369 haplotypes rarely showed evidence for recombination. In contrast, in BSc1, mtDNA
370 recombination and deletion were almost mutually exclusive. This is consistent with differences
371 in mtDNA architecture between *S. paradoxus* lineages *SpB* and *SpA*. While *SpB* mtDNAs are
372 collinear with other North American lineages and with *S. cerevisiae*, *SpA* harbors rearranged
373 mtDNAs (Leducq et al. 2017; Yue et al. 2017). As such, homologous recombination between
374 rearranged mtDNAs could lead to frequent deletions. However, deletion breakpoints were often
375 located in unannotated regions, contrasting with recombination breakpoints and suggesting a
376 distinct mechanism. The association between mtDNA instability and aneuploidy instability in
377 BA2 is more challenging to interpret. Nevertheless, it raises the possibility that both phenotypes
378 could be linked or share a common underlying cause. It appears unlikely that a genetic basis is
379 involved, as these phenotypes appeared among isogenic lines after a very short evolution time,
380 and show little evidence for persistence through time (Supplemental Fig. S40)

381 Experimental evolution by MA minimizes the efficiency of natural selection, enabling a
382 selectively unbiased characterization of mtDNA evolution. The characterization of de novo
383 mutation in nuclear genomes of our MA lines revealed signatures that were entirely consistent
384 with neutral evolution (Fijarczyk et al. 2021). Our observation of up to 45% MA lines with a
385 complete loss of respiratory function further confirms the absence of selection. While the MA
386 design minimized the efficiency of natural selection at the level of individual cells, we highlight
387 that it did not alleviate it at the intracellular level. MtDNA haplotypes can outcompete others if

388 they have a replication advantage, regardless of their effect on organismal fitness (Ma and
389 O'Farrell 2016; Taylor et al. 2002). As such, some mtDNA haplotypes could have been driven to
390 homoplasmy by replication advantages.

391 Our results show that respiratory function is primarily lost through large-scale deletions within
392 mtDNAs. The level of replication of our MA experiment yielded a high number of recombinant
393 haplotypes. Nevertheless, cases of respiratory loss of function in our experiment are likely
394 entirely explained by mtDNA deletions. Additionally, we found no significant negative effect of
395 recombinant mtDNA haplotypes on respiration. In BSc1, the non-significant growth difference
396 between lines with *SpB* versus recombinant haplotypes on YPEG at 25°C was the second
397 largest effect size, suggesting that some recombinant haplotypes might be disadvantaged.
398 Despite this hypothetical case, we found no support for negative mito-mito epistasis, contrary to
399 previous reports (Wolters et al. 2018; Leducq et al. 2017).

400 Mitonuclear incompatibilities are known to contribute to reproductive isolation between species
401 of the *Saccharomyces* genus (Chou et al. 2010; Lee et al. 2008; Jhuang et al. 2017), along with
402 other types of reproductive barriers (Hunter et al. 1996). Even at the intraspecific level in *S.*
403 *cerevisiae* and *S. paradoxus*, crosses between diverged isolates often have low fertility (Greig et
404 al. 2003; Charron et al. 2014b; Leducq et al. 2016; Eberlein et al. 2019), including most of our
405 MA crosses (Charron et al. 2019; Marsit et al. 2021). Our results show that loss of respiration
406 adds another reproductive barrier, since meiosis (sporulation) is dependent on functioning
407 respiratory metabolism (Jambhekar and Amon 2008). Yeast can divide strictly mitotically in
408 permissive laboratory conditions, and mitosis is dominant in the wild (Tsai et al. 2008).
409 However, sporulation is triggered in starvation conditions (Bautz Freese et al. 1982) and is likely
410 to be essential in nature, thus selecting for mtDNA integrity. Surprisingly, natural isolates of *S.*
411 *cerevisiae* were reported to completely lack mtDNA (De Chiara et al. 2020), suggesting that loss

412 of sporulation ability can arise in nature. As such, mtDNA degeneration in hybrids acts as an
413 additional mechanism of reproductive isolation, both within and between species.

414 In conclusion, our study provides the first large-scale investigation of neutral mtDNA evolution in
415 hybrids. Our results show that distinct aspects of neutral mtDNA evolution can be either
416 genotype-specific or, in contrast, highly predictable given the genomic composition of hybrids.
417 We show that hybridization can trigger mtDNA degeneration with profound metabolic
418 consequences and the emergence of reproductive isolation.

419 **METHODS**

420 *Long read library preparation and sequencing*

421 Stocks of the 13 haploid parental strains (Supplemental Table S1) and of 122 randomly selected
422 MA lines from the final timepoint (Charron et al. 2019; Hénault et al. 2020) were thawed,
423 streaked on YPD agar medium (1% yeast extract, 2% bio-tryptone, 2% glucose, 2% agar) and
424 incubated at room temperature for three days. Cultures in YPD medium (1% yeast extract, 2%
425 bio-tryptone, 2% glucose) were inoculated with single colonies and incubated overnight at room
426 temperature without shaking. DNA was extracted from cultures following a standard phenol-
427 chloroform protocol. Oxford Nanopore native (PCR-free) genomic DNA libraries were prepared
428 in multiplex with kits SQK-LSK109 and EXP-NBD104 (Oxford Nanopore Technologies, Oxford,
429 UK). Libraries were sequenced with FLO-MIN106 (revC) flowcells on a MinION sequencer
430 (MIN-101B). The sequencing and basecalling were run on a MinIT computer (MNT-001) with
431 MinKNOW v3.3.2 and guppy v3.0.3. Reads were demultiplexed using guppy_basecaller v3.1.5.

432 *Long read mtDNA assemblies*

433 Long reads matching mtDNA sequences were extracted from whole-genome libraries. For each
434 strain, we made a compound reference genome comprising the nuclear genome of the

435 corresponding *S. paradoxus* lineage(s) (*SpA*: LL2012_001, *SpB*: MSH-604, *SpC*: LL2011_012)
436 (Eberlein et al. 2019), or *S. cerevisiae* (YPS128) (Yue et al. 2017), supplemented with a
437 concatenated mitochondrial contig comprising *S. paradoxus SpB* (YPS138), *SpA* (CBS432) and
438 (YPS138) and *S. cerevisiae* (S288c) (Yue et al. 2017) mtDNAs joined by 100 Ns. Nanopore
439 libraries were mapped on the compound references using minimap2 v2.20-r1061 (Li 2018a)
440 with preset map-ont. Reads mapping to the mtDNA contig were extracted using samtools v1.15
441 (Li et al. 2009) and seqtk v1.3-r106 (Li 2018b). Assembly of the extracted reads was performed
442 using wtdbg2 v2.5 (Ruan and Li 2019) with options –AS 2 –g 72k –X 50 and a grid search of the
443 following parameters: –k {23, 21, 19, 17, 15, 13, 11}, –p {8, 5, 2, 1, 0}, –l {4096 2048 1024}.

444 For parental strains, draft assemblies were aligned against the reference mtDNA assembly (Yue
445 et al. 2017) of the closest species/lineage (*SpB* and *SpC*: YPS138, *SpA*: CBS432, *S. cerevisiae*:
446 S288c) using MUMmer v3.23 (Kurtz et al. 2004). For each strain, the best assembly was
447 selected on the basis of a quality score computed from the alignment of each query assembly
448 using custom Python v3.9.10 (Van Rossum and Drake 2009) scripts. A genome coverage
449 metric was first computed as the sum of absolute deviations from 1 of both query and reference
450 coverage. A contiguity metric was computed as the inverse of the largest contig size in the
451 assembly. Both coverage and continuity metrics were normalized to z-scores and summed to
452 yield the final score. Assemblies with the lowest score were circularized at the start of the *ATP6*
453 gene and resulting duplications were manually trimmed. Circularized assemblies were polished
454 by first aligning the raw reads to draft assemblies using minimap2 v2.20-r1061 with preset map-
455 ont. Alignments were polished using the raw nanopore signal using nanopolish v0.13.3 (Loman
456 et al. 2015) with parameter —min-candidate-frequency 0.1. Short read libraries for each
457 parental strain (Charron et al. 2019; Hénault et al. 2020) were aligned on draft assemblies using
458 bwa-mem v0.7.17 and polishing was performed using Pilon v1.22 (Walker et al. 2014).

459 For MA lines, draft assemblies were polished with basecalled reads using medaka v1.4.4
460 (https://github.com/nanoporetech/medaka) with model r941_min_high_g303. Assemblies were
461 circularized at the start of the *ATP6* gene (if present) and aligned to the artificial mtDNA
462 reference (see below) using MUMmer v3.23.

463 *MtDNA assembly annotation and generation of mtDNA artificial reference*

464 Genomes of the 13 parental strains were annotated using the MFannot server
465 (https://megasun.bch.umontreal.ca/cgi-bin/mfannot/mfannotInterface.pl) and annotations were
466 manually curated. From the comparative analysis of genome annotations, we built an artificial
467 mtDNA reference comprising the most complete set of features. We used the LL2012_028
468 genome as a template, and added introns *cob*-I3, *cob*-I4, *cob*-I5 and *cox1*-I3γ from YPS744 or
469 LL2013_054 at the orthologous splice sites using custom Python v3.9.10 scripts. Genomes
470 were aligned to the artificial mtDNA contig using mugsy v1r2.3 (Angiuoli and Salzberg 2011).
471 Liftover tables were built using a custom Python v3.9.10 script.

472 *Short read alignment and variant calling*

473 Short read libraries corresponding to the 13 parental haploid strains and 447 hybrid yeast MA
474 lines, sampled at both the initial (i.e. shortly after zygote formation) and final (i.e. after ~770
475 mitotic generations) timepoints, were downloaded from NCBI (BioProject PRJNA515073). Short
476 read preprocessing was performed as described in (Hénault et al. 2020). We built a
477 concatenated reference genome comprising nuclear contigs of *S. paradoxus SpB* strain MSH-
478 604 (Eberlein et al. 2019) and *S. cerevisiae* strain YPS128 (Yue et al. 2017), to which we
479 appended the artificial mtDNA contig. Reads were aligned to this reference using bwa mem
480 v0.7.17 and mappings to the mtDNA contig were extracted using samtools v1.8. Variant calling
481 was performed separately for each cross using freebayes v1.3.5 (Garrison and Marth 2012).
482 Multiallelic variants were decomposed using vcfbreakmulti from vcflib v1.0.2 (Garrison 2018).

483 *Definition of mtDNA recombination tracts*

484 The characterization of recombination in mtDNA haplotypes was performed using custom
485 Python v3.9.10 scripts. For each cross, marker variants were identified for the parental
486 genomes based on variant calls from short-read alignments on the artificial mtDNA reference.
487 Candidate markers were filtered as follows. Each variant had to be supported by at least two
488 read alignments. Parental alleles at each marker locus were validated using whole-genome
489 alignments of the high-quality genome assemblies of each parental mtDNA and the associated
490 liftover tables. Markers were required to have nucleotide calls and assembly alignments for both
491 parents, thus excluding insertion and deletions.

492 For MA lines of each cross, variants corresponding to parental markers were filtered to exclude
493 loci with low depth of coverage. For each line we defined a threshold as the highest value
494 between the 10th percentile of depth of coverage values for the nuclear genome of each line, or
495 20 reads. Variants passed the filter if supporting reads for the major allele were higher than the
496 threshold. Since each MA line was sampled for sequencing at two timepoints, mtDNA
497 haplotypes were compared across timepoints of individual lines to determine whether they were
498 clonal or independent haplotypes that segregated from initially heteroplasmic lines. For each
499 line, we computed the pairwise identity between mtDNA haplotypes as the proportion of
500 identical calls at marker variants (excluding missing positions in either mtDNA). If the identity
501 was 95% or higher, we kept only one mtDNA (preferentially the latest timepoint) as
502 representative of the line. Independent haplotypes within each cross were clustered by identity
503 using the nearest point algorithm from scipy v1.8.0 (Virtanen et al. 2020).

504 Recombination tracts were defined by scanning and grouping adjacent marker variants of the
505 same genotype along the genome, with a tolerance of one marker of the opposite genotype (i.e.
506 two marker variants of the opposite genotype were required to break a tract). Tracts were

507 circularized at the ends of the mtDNA contig. Recombination breakpoints were defined as the
508 marker variants at the ends of each tract. Recombination breakpoint and marker variant
509 densities were computed for 100 bp windows along the genome. For haplotypes sampled at the
510 final timepoint, recombination rates were computed as the number of recombination tracts per
511 haplotype divided by the combined length of parental mtDNAs, and the number of generations
512 of the MA experiment corrected for growth yield estimated by (Fijarczyk et al. 2021). Average
513 recombination rate and tract length for each cross were correlated with marker variant counts
514 and parental evolutionary divergence values computed by (Hénault et al. 2020).

515 *MtDNA inheritance biases*

516 Recombination tracts were used to compute the proportion of each mtDNA haplotype that was
517 inherited from either parent. The inheritance ratios were averaged over all independent
518 haplotypes of a cross to yield genome-wide average inheritance ratios. The depth of coverage
519 of long reads on each parental mtDNA assembly compared to nuclear chromosomes was used
520 to compute a relative copy number. For individual genes within haplotypes, discrete inheritance
521 calls were generated by attributing either parental genotype to annotations that had ratios of
522 10% or less, or 90% or more. Genes with ratios between those values were considered as
523 recombinant and excluded. Binomial tests were used to test the deviation of non-recombinant
524 genes from genome-wide inheritance ratios.

525 *Growth measurements of the MA lines collection*

526 During the MA experiment, MA lines were sampled at regular intervals and archived in
527 cryopreserved stocks in 96 well microplates. Archives from the stocks sampled at the initial and
528 final timepoints were replicated in YPD medium and cryopreserved as copies of the original
529 archives. For high-throughput phenotyping, stock copies were thawed and spotted on YPD agar
530 medium in OmniTray plates (Thermo Fisher Scientific, Waltham, USA) using a BM5-SC1 colony

531 processing robot (S&P Robotics Inc., Toronto, Canada) and incubated at room temperature for
532 four days. Colonies were replicated on YPD agar by condensing them into six 384-position
533 source arrays and incubated at room temperature for three days. Source arrays were replicated
534 twice on YPD agar medium and twice on YPEG agar medium (1% yeast extract, 2% bio-
535 tryptone, 3% glycerol, 3% ethanol, 2% agar). Plates were incubated at 25°C or 37°C for the first
536 round of selection during two days. Plates were replicated for the second round of selection on
537 the same media and photographed manually after 24, 48 and 96 hours of incubation with the
538 robotic platform. Colony sizes were quantified from the plate photos using gitter v1.1 (Wagih
539 and Parts 2014) in R v3.3.2 (R Core Team 2016). Colony sizes in pixels were transformed in log
540 and summed across imaging timepoints to estimate the area under growth curves (AUC). One
541 colony (BSc2, line F53, initial timepoint, YPD 37°C) was removed due to visible contamination.
542 Calls of growth absence based on pixel quantifications were manually curated with raw images.

543 *Characterization of mtDNA deletions*

544 Depth of coverage values along mtDNAs were extracted from short read alignments using
545 samtools v1.15. Median depth values were computed for 100 bp windows and normalized by
546 the depth of coverage on nuclear chromosomes. Depth values were grouped by annotation type
547 and the fraction of positions with normalized depth values higher than 0.1 was extracted. For
548 each cross, these fractions were normalized into z-scores, and an mtDNA completion score was
549 computed as the sum of z-scores across annotation types for individual mtDNA haplotypes.
550 Haplotypes were sorted by mtDNA completion score to classify them as complete or deleted,
551 and the classification was manually curated by examination of the depth of coverage profiles.
552 Haplotypes were clustered by Pearson's correlation coefficient between normalized depth
553 profiles using the nearest point algorithm from scipy v1.8.0.

554 *Detection of mobile introns presence/absence polymorphisms*

555 Presence or absence of mobile introns in mtDNA haplotypes was called by analyzing depth of
556 coverage profiles at intron-exon junctions. For each polymorphic intron, the marker variants
557 closest to both junctions (upstream or downstream) were identified. 100 bp windows centered
558 around each junction were used to analyze depth of coverage profiles. Depth of coverage
559 values were normalized by the 80th percentile of depth values within a window. Euclidean
560 distance and Pearson's correlation coefficient were used to compare profiles of each line with
561 both parental profiles. Evidence for intron presence was called at a junction if the euclidean
562 distance and Pearson's correlation coefficient were respectively lower and higher for
563 comparisons to the parent harboring the intron. An intron was called present if its two junctions
564 showed consistent evidence. Putative intron mobilization events were defined as intron
565 presence with two flanking markers of the opposite ancestry. Putative recombination-associated
566 mobilization events were defined as intron presence with one flanking marker of the opposite
567 ancestry, consistent with neighboring recombination tracts.

568 *Confirmation of growth phenotypes by spot assays*

569 A subset of lines classified as having mtDNA instability at the initial timepoint were selected for
570 growth phenotypes validation by spot assays. Original archive stocks and copies were thawed,
571 spotted on YPD agar medium and incubated at room temperature. Spots were used to inoculate
572 precultures in YPD medium that were incubated overnight at 30°C. Precultures were diluted at 1
573 OD ml^{-1} in sterile water and five 5X serial dilutions were prepared. Aliquots of 5 µl of each
574 dilution were spotted twice on YPD agar and YPEG agar. Plates were incubated at 25°C or
575 37°C for three days and photographed.

576 *De novo mutations identification*

577 Candidate de novo single nucleotide mutations were identified from variant calls based on short
578 read alignments. First, variant calls supported by at least five reads and which had an
579 alternative allele ratio of at least 0.8 were extracted. Variants which were only observed in a
580 single line were further checked for the absence of matching alleles in the parental variant calls
581 from all crosses, and from lines of all the other crosses. For each remaining candidate, we
582 extracted the proportion of lines in the cross with missing data at that locus. We defined a more
583 stringent alignment criteria for supporting reads by requiring the length of the aligned reference
584 segment to be between 141 and 161 bp (for 151 bp reads), and counted those high confidence
585 alignments for each candidate variant. Finally, we counted the number of supporting reads for
586 matching alleles in the whole dataset, excluding the line of interest. The final set of candidate de
587 novo mutations consisted in variants for which the total number of supporting reads was at least
588 twice the number of supporting reads in the whole dataset excluding the line, and had at least
589 two supporting reads with high confidence alignments.

590 *Depth of coverage deviations on nuclear chromosomes*

591 Variant calls from short-read alignments and aneuploidy calls were previously done by Fijarczyk
592 and collaborators (Fijarczyk et al. 2021). BSc crosses were excluded from this analysis as the
593 mapping strategy was different (using concatenated *SpB* and *S. cerevisiae* nuclear reference
594 genomes). For each cross, 10000 nuclear variants were chosen at random. Allele ratios were
595 examined along nuclear chromosomes. For individual lines, the median depth of coverage of
596 each chromosome was normalized by the genome-wide median to yield the depth ratio. The
597 copy number ratio was defined by dividing chromosome copy number calls by the global ploidy
598 level of each line. Depth of coverage deviation scores were computed as (depth ratio – copy
599 number ratio). Absolute deviation scores were averaged over all chromosomes and lines with

600 an average absolute deviation score of at least 0.08 were classified as having the aneuploidy
601 instability phenotype.

602 *MtDNA genotyping by PCR*

603 Two loci were genotyped for presence or absence by PCR for all lines of the BA2 cross at the
604 initial and final timepoints. DNA was extracted by resuspending colonies in 40 µl NaOH 20 mM
605 and incubating at 95°C for 20 min. Two mtDNA amplicons in the *ATP6* and 21S rRNA were
606 amplified with primers and cycles detailed in Supplemental Table S6 (Leducq et al. 2017).
607 Amplicons of the nuclear mating type loci were amplified as positive controls with primers and
608 cycles detailed in Supplemental Table S6 (Huxley et al. 1990). Presence and absence were
609 scored by agarose gel electrophoresis.

610 **DATA ACCESS**

611 Long read datasets are available at NCBI under accession PRJNA828354. Scripts, genome
612 assemblies and annotations are available at https://github.com/Landrylab/mito_ma.

613 **COMPETING INTEREST STATEMENT**

614 We declare no competing interest.

615 **ACKNOWLEDGEMENTS**

616 We thank Anna Fijarczyk for comments on the manuscript and contributions to sequencing data
617 preprocessing and analysis. We thank Hélène Martin for contributions to sequencing data
618 preprocessing. We thank the members of the Landry laboratory for discussions on the project.
619 This project was supported by funding to C.R.L. from a NSERC discovery grant and FRQNT
620 team grant, and a NSERC Alexander Graham Bell doctoral scholarship to M.H. CRL holds the
621 Canada Research Chair in Cellular Systems and Synthetic Biology.

622 **AUTHOR CONTRIBUTIONS**

623 CRL and MH designed the project. MH performed the analyses and experiments. MH, SM and
624 GC performed the MA evolution experiment and short read sequencing as cited. MH wrote the
625 manuscript with inputs from CRL and SM.

626 **REFERENCES**

627 Adams KL, Palmer JD. 2003. Evolution of mitochondrial gene content: Gene loss and transfer to
628 the nucleus. *Mol Phylogenet Evol* **29**: 380–395.

629 Angiuoli SV, Salzberg SL. 2011. Mugsy: Fast multiple alignment of closely related whole
630 genomes. *Bioinformatics* **27**: 334–342.

631 Bágeľová Poláková S, Lichtner Ž, Szemes T, Smolejová M, Sulo P. 2021. Mitochondrial DNA
632 duplication, recombination, and introgression during interspecific hybridization. *Sci Rep* **11**:
633 12726.

634 Bautz Freese E, Chu MI, Freese E. 1982. Initiation of Yeast Sporulation by Partial Carbon,
635 Nitrogen, or Phosphate Deprivation. *J Bacteriol* **149**: 840–851.

636 Bernardi G. 1979. The petite mutation in yeast. *Trends Biochem Sci* **4**: 197–201.

637 Birky CW. 2001. T He R Ole of Knox G Enes in. *Annu Rev Genet* **35**: 125/148.

638 Breton S, Stewart DT, Bonen L. 2015. Atypical mitochondrial inheritance patterns in eukaryotes.
639 *Genome* **58**: 423–431.

640 Burton RS. 2022. The role of mitonuclear incompatibilities in allopatric speciation. *Cell Mol Life*
641 *Sci* **79**: 103.

642 Burton RS, Barreto FS. 2012. A disproportionate role for mtDNA in Dobzhansky-Muller

643 incompatibilities? *Mol Ecol* **21**: 4942–4957.

644 Charron G, Leducq JB, Bertin C, Dubé AK, Landry CR. 2014a. Exploring the northern limit of
645 the distribution of Saccharomyces cerevisiae and Saccharomyces paradoxus in North
646 America. *FEMS Yeast Res* **14**: 281–288.

647 Charron G, Leducq JB, Landry CR. 2014b. Chromosomal variation segregates within incipient
648 species and correlates with reproductive isolation. *Mol Ecol* **23**: 4362–4372.

649 Charron G, Marsit S, Hénault M, Martin H, Landry CR. 2019. Spontaneous whole-genome
650 duplication restores fertility in interspecific hybrids. *Nat Commun* **10**.
651 http://dx.doi.org/10.1038/s41467-019-12041-8.

652 Chou J-Y, Hung Y-S, Lin K-H, Lee H-Y, Leu J-Y. 2010. Multiple Molecular Mechanisms Cause
653 Reproductive Isolation between Three Yeast Species ed. M.A.F. Noor. *PLoS Biol* **8**:
654 e1000432.

655 De Chiara M, Friedrich A, Barré B, Breitenbach M, Schacherer J, Liti G. 2020. Discordant
656 evolution of mitochondrial and nuclear yeast genomes at population level. *BMC Biol* **18**: 49.

657 Dujon B. 2020. Mitochondrial genetics revisited. *Yeast* **37**: 191–205.

658 Dujon B, Slonimski PP, Weill L. 1974. Mitochondrial genetics IX: A model for recombination and
659 segregation of mitochondrial genomes in saccharomyces cerevisiae. *Genetics* **78**: 415–
660 437.

661 Eberlein C, Hénault M, Fijarczyk A, Charron G, Bouvier M, Kohn LM, Anderson JB, Landry CR.
662 2019. Hybridization is a recurrent evolutionary stimulus in wild yeast speciation. *Nat*
663 *Commun* **10**: 923.

664 Ephrussi B, Hottinguer H, Tavlitzki J. 1949. Action de l'acriflavine sur les levures. I. La mutation"

665 petite colonie." *Ann Inst Pasteur* **76**: 351–367.

666 Fijarczyk A, Hénault M, Marsit S, Charron G, Landry CR. 2021. Heterogeneous Mutation Rates
667 and Spectra in Yeast Hybrids. *Genome Biol Evol* **13**.
668 http://dx.doi.org/10.1093/gbe/evab282.

669 Fritsch ES, Chabbert CD, Klaus B, Steinmetz LM. 2014. A genome-wide map of mitochondrial
670 DNA recombination in yeast. *Genetics* **198**: 755–771.

671 Galeota-Sprung B, Fernandez A, Sniegowski P. 2021. Changes to the mtDNA copy number
672 during yeast culture growth. *bioRxiv* 2021.09.02.458779.
673 https://www.biorxiv.org/content/10.1101/2021.09.02.458779v4.full (Accessed March 1,
674 2022).

675 Garrison E. 2018. *vcflib*. https://github.com/vcflib/vcflib.

676 Garrison E, Marth G. 2012. Haplotype-based variant detection from short-read sequencing. 1–9.

677 Greig D, Travisano M, Louis EJ, Borts RH. 2003. A role for the mismatch repair system during
678 incipient speciation in Saccharomyces. *J Evol Biol* **16**: 429–437.

679 Gyllensten U, Wharton D, Josefsson A, Wilson AC. 1991. Paternal inheritance of mitochondrial
680 DNA in mice. *Nature* **352**: 255–257. http://dx.doi.org/10.1038/352255a0.

681 Hénault M, Marsit S, Charron G, Landry CR. 2020. The effect of hybridization on transposable
682 element accumulation in an undomesticated fungal species. *Elife* **9**.
683 http://dx.doi.org/10.7554/eLife.60474.

684 Hunter N, Chambers SR, Louis EJ, Borts RH. 1996. The mismatch repair system contributes to
685 meiotic sterility in an interspecific yeast hybrid. *EMBO J* **15**: 1726–1733.

686 Huxley C, Green ED, Dunham I. 1990. Rapid assessment of S. cerevisiae mating type by PCR.
687 *Trends Genet* **6**: 236.

688 Jambhekar A, Amon A. 2008. Control of meiosis by respiration. *Curr Biol* **18**: 969–975.

689 Jhuang HY, Lee HY, Leu JY. 2017. Mitochondrial-nuclear co-evolution leads to hybrid
690 incompatibility through pentatricopeptide repeat proteins. *EMBO Rep* **18**: 87–101.

691 Joseph SB, Hall DW. 2004. Spontaneous mutations in diploid Saccharomyces cerevisiae: more
692 beneficial than expected. *Genetics* **168**: 1817–1825.

693 Kondo R, Satta Y, Matsuura ET, Ishiwa H, Takahata N, Chigusa SI. 1990. Incomplete maternal
694 transmission of mitochondrial DNA in Drosophila. *Genetics* **126**: 657–663.

695 Kuehne HA, Murphy HA, Francis CAA, Sniegowski PD. 2007. Allopatric Divergence, Secondary
696 Contact, and Genetic Isolation in Wild Yeast Populations. *Curr Biol* **17**: 407–411.

697 Kurtz S, Phillippy A, Delcher AL, Smoot M, Shumway M, Antonescu C, Salzberg SL. 2004.
698 Versatile and open software for comparing large genomes. *Genome Biol* **5**: R12.

699 Ladoukakis ED, Theologidis I, Rodakis GC, Zouros E. 2011. Homologous recombination
700 between highly diverged mitochondrial sequences: examples from maternally and
701 paternally transmitted genomes. *Mol Biol Evol* **28**: 1847–1859.

702 Ladoukakis ED, Zouros E. 2017. Evolution and inheritance of animal mitochondrial DNA: rules
703 and exceptions. *J Biol Res* **24**: 2.

704 Leducq J-B, Henault M, Charron G, Nielly-Thibault L, Terrat Y, Fiumera HL, Shapiro BJ, Landry
705 CR. 2017. Mitochondrial Recombination and Introgression during Speciation by
706 Hybridization. *Mol Biol Evol* **34**: 1947–1959.

707 Leducq JB, Nielly-Thibault L, Charron G, Eberlein C, Verta JP, Samani P, Sylvester K, Hittinger
708 CT, Bell G, Landry CR. 2016. Speciation driven by hybridization and chromosomal plasticity
709 in a wild yeast. *Nature Microbiology* **1**: 1–58.

710 Lee H-Y, Chou J-Y, Cheong L, Chang N-H, Yang S-Y, Leu J-Y. 2008. Incompatibility of Nuclear
711 and Mitochondrial Genomes Causes Hybrid Sterility between Two Yeast Species. *Cell* **135**:
712 1065–1073.

713 Li H. 2018a. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**: 3094–
714 3100.

715 Li H. 2018b. Seqtk: a fast and lightweight tool for processing FASTA or FASTQ sequences.
716 https://github.com/lh3/seqtk.

717 Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R.
718 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**: 2078–2079.

719 Li XC, Peris D, Hittinger CT, Sia EA, Fay JC. 2019. Mitochondria-encoded genes contribute to
720 evolution of heat and cold tolerance in yeast. *Sci Adv* **5**: eaav1848.

721 Loman NJ, Quick J, Simpson JT. 2015. A complete bacterial genome assembled de novo using
722 only nanopore sequencing data. *Nat Methods* **12**: 733–735.

723 Lynch M, Ackerman MS, Gout J-F, Long H, Sung W, Thomas WK, Foster PL. 2016. Genetic
724 drift, selection and the evolution of the mutation rate. *Nat Rev Genet* **17**: 704–714.

725 Lynch M, Sung W, Morris K, Coffey N, Landry CR, Dopman EB, Dickinson WJ, Okamoto K,
726 Kulkarni S, Hartl DL, et al. 2008. A genome-wide view of the spectrum of spontaneous
727 mutations in yeast. *Proc Natl Acad Sci U S A* **105**: 9272–9277.

728 Ma H, O'Farrell PH. 2016. Selfish drive can trump function when animal mitochondrial genomes

729 compete. *Nat Genet* **48**: 798–802.

730 Marsit S, Hénault M, Charron G, Fijarczyk A, Landry CR. 2021. The neutral rate of whole-
731 genome duplication varies among yeast species and their hybrids. *Nat Commun* **12**: 3126.

732 Müller M, Mentel M, van Hellemond JJ, Henze K, Woehle C, Gould SB, Yu R-Y, van der Giezen
733 M, Tielens AGM, Martin WF. 2012. Biochemistry and evolution of anaerobic energy
734 metabolism in eukaryotes. *Microbiol Mol Biol Rev* **76**: 444–495.

735 Orr HA. 1995. The population genetics of speciation: The evolution of hybrid incompatibilities.
736 *Genetics* **139**: 1805–1813.

737 Paliwal S, Fiumera AC, Fiumera HL. 2014. Mitochondrial-Nuclear Epistasis Contributes to
738 Phenotypic Variation and Coadaptation in Natural Isolates of Saccharomyces cerevisiae.
739 *Genetics* **198**: 1251–1265.

740 Parakatselaki M-E, Ladoukakis ED. 2021. mtDNA Heteroplasmy: Origin, Detection,
741 Significance, and Evolutionary Consequences. *Life* **11**.
742 http://dx.doi.org/10.3390/life11070633.

743 Payne BAI, Wilson IJ, Yu-Wai-Man P, Coxhead J, Deehan D, Horvath R, Taylor RW, Samuels
744 DC, Santibanez-Koref M, Chinnery PF. 2013. Universal heteroplasmy of human
745 mitochondrial DNA. *Hum Mol Genet* **22**: 384–390.

746 Piccinini G, Iannello M, Puccio G, Plazzi F, Havird JC, Ghiselli F. 2021. Mitonuclear
747 Coevolution, but not Nuclear Compensation, Drives Evolution of OXPHOS Complexes in
748 Bivalves. *Mol Biol Evol* **38**: 2597–2614.

749 R Core Team. 2016. R: A Language and Environment for Statistical Computing. https://www.R-
750 project.org/.

751 Rokas A, Ladoukakis E, Zouros E. 2003. Animal mitochondrial DNA recombination revisited.
752 *Trends Ecol Evol* **18**: 411–417.

753 Ruan J, Li H. 2019. Fast and accurate long-read assembly with wtdbg2. *Nat Methods*.
754 http://dx.doi.org/10.1038/s41592-019-0669-3.

755 Solieri L. 2010. Mitochondrial inheritance in budding yeasts: towards an integrated
756 understanding. *Trends Microbiol* **18**: 521–530.

757 Stewart JB, Chinnery PF. 2021. Extreme heterogeneity of human mitochondrial DNA from
758 organelles to populations. *Nat Rev Genet* **22**: 106–118.

759 Stewart JB, Chinnery PF. 2015. The dynamics of mitochondrial DNA heteroplasmy: implications
760 for human health and disease. *Nat Rev Genet* **16**: 530–542.

761 Taylor DR, Zeyl C, Cooke E. 2002. Conflicting levels of selection in the accumulation of
762 mitochondrial defects in Saccharomyces cerevisiae. *Proc Natl Acad Sci U S A* **99**: 3690–
763 3694.

764 Tsai IJ, Bensasson D, Burt A, Koufopanou V. 2008. Population genomics of the wild yeast
765 Saccharomyces paradoxus: Quantifying the life cycle. *Proc Natl Acad Sci U S A* **105**: 4957–
766 4962.

767 Tsaousis AD, Martin DP, Ladoukakis ED, Posada D, Zouros E. 2005. Widespread
768 recombination in published animal mtDNA sequences. *Mol Biol Evol* **22**: 925–933.

769 Van Rossum G, Drake FL. 2009. *Python 3 Reference Manual: (Python Documentation Manual*
770 *Part 2)*. CreateSpace Independent Publishing Platform.

771 Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, Burovski E,
772 Peterson P, Weckesser W, Bright J, et al. 2020. SciPy 1.0: fundamental algorithms for

773 scientific computing in Python. *Nat Methods* **17**: 261–272.

774 Vowinckel J, Hartl J, Marx H, Kerick M, Runggatscher K, Keller MA, Mülleder M, Day J, Weber
775 M, Rinnerthaler M, et al. 2021. The metabolic growth limitations of petite cells lacking the
776 mitochondrial genome. *Nat Metab* **3**: 1521–1535.

777 Wagih O, Parts L. 2014. Gitter: A robust and accurate method for quantification of colony sizes
778 from plate images. *G3: Genes, Genomes, Genetics* **4**: 547–552.

779 Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo CA, Zeng Q,
780 Wortman J, Young SK, et al. 2014. Pilon: An Integrated Tool for Comprehensive Microbial
781 Variant Detection and Genome Assembly Improvement ed. J. Wang. *PLoS One* **9**:
782 e112963.

783 Wallace DC. 2015. Mitochondrial DNA variation in human radiation and disease. *Cell* **163**: 33–
784 38.

785 Wang Z, Wilson A, Xu J. 2015. Mitochondrial DNA inheritance in the human fungal pathogen
786 Cryptococcus gattii. *Fungal Genet Biol* **75**: 1–10.

787 Wilson AJ, Xu J. 2012. Mitochondrial inheritance: Diverse patterns and mechanisms with an
788 emphasis on fungi. *Mycology* **3**: 158–166.

789 Wolters JF, Charron G, Gaspary A, Landry CR, Fiumera AC, Fiumera HL. 2018. Mitochondrial
790 recombination reveals mito–mito epistasis in yeast. *Genetics* **209**: 307–319.

791 Xia W, Nielly-Thibault L, Charron G, Landry CR, Kasimer D, Anderson JB, Kohn LM. 2017.
792 Population genomics reveals structure at the individual, host-tree scale and persistence of
793 genotypic variants of the undomesticated yeast Saccharomyces paradoxus in a natural
794 woodland. *Mol Ecol* **26**: 995–1007.

795 Yue JX, Li J, Aigrain L, Hallin J, Persson K, Oliver K, Bergström A, Coupland P, Warringer J,
796 Lagomarsino MC, et al. 2017. Contrasting evolutionary genome dynamics between
797 domesticated and wild yeasts. *Nat Genet* **49**: 913–924.

798

41