Title

Mapping mitonuclear epistasis using a novel recombinant yeast population

Authors and Affiliations

Tuc H.M. Nguyen[1,3], Austen Tinz-Burdick[1], Meghan Lenhardt[1], Margaret Geertz[1], Franchesca Ramirez[1], Mark Schwartz[1], Michael Toledano[1], Brooke Bonney[1], Benjamin Gaebler[1], Weiwei Liu[1], John F. Wolters[1], Ken Chiu[2], Anthony C. Fiumera[1], Heather L. Fiumera[1*]

[1]Department of Biological Sciences, Binghamton University, Binghamton, NY USA
[2]Department of Computer Sciences, Binghamton University, Binghamton, NY USA
[3] Current address: Department of Biological Sciences, New York University, New York, NY USA

* Corresponding Author
Heather L. Fiumera (hfiumera@binghamton.edu)
Department of Biological Sciences
Binghamton University
Binghamton, NY 13902

Abstract

Natural genetic variation in mitochondrial and nuclear genomes can influence phenotypes by perturbing coadapted mitonuclear interactions. Mitonuclear epistasis, i.e. non-additive phenotype effects of interacting mitochondrial and nuclear alleles, is emerging as a general feature in eukaryotes, yet very few mitonuclear loci have been identified. Here, we present a novel advanced intercrossed population of *S. cerevisiae* yeasts, called the Mitonuclear Recombinant Collection (MNRC), designed explicitly for detecting mitonuclear loci contributing to complex traits, and use this population to map the genetic basis to mtDNA loss. In yeast, spontaneous deletions within mtDNAs lead to the *petite* phenotype that heralded mitochondrial research. We show that in natural populations, rates of *petite* formation are variable and influenced by genetic variation in

nuclear, mtDNAs and mitonuclear interactions. We then mapped nuclear and mitonuclear alleles contributing mtDNA stability using the MNRC by integrating mitonuclear epistasis into a genome-wide association model. We found that associated mitonuclear loci play roles in mitotic growth most likely responding to retrograde signals from mitochondria, while associated nuclear loci with main effects are involved in genome replication. We observed a positive correlation between growth rates and *petite* frequencies, suggesting a fitness tradeoff between mitotic growth and mtDNA stability. We also found that mtDNA stability was influenced by a mobile mitochondrial GC-cluster that is expanding in certain populations of yeast and that selection for nuclear alleles that stabilize mtDNA may be rapidly occurring. The MNRC provides a powerful tool for identifying mitonuclear interacting loci that will help us to better understand genotype-phenotype relationships and coevolutionary trajectories.

Author Summary

Genetic variation in mitochondrial and nuclear genomes can perturb mitonuclear interactions and lead to phenotypic differences between individuals and populations. Despite their importance to most complex traits, it has been difficult to identify the interacting loci. Here, we created a novel population of yeast designed explicitly for mapping mitonuclear loci contributing to complex traits and used this population to map genes influencing the stability of mitochondrial DNA (mtDNA). We found that mitonuclear interacting loci were involved in mitotic growth while non-interacting loci were involved in genome replication. We also found evidence that selection for mitonuclear loci that stabilize mtDNAs occurs rapidly. This work provides insight into mechanisms underlying maintenance of mtDNAs. The mapping population presented here is an important new resource that will help to understand genotype/phenotype relationships and coevolutionary trajectories.

Background

Interactions between mitochondrial and nuclear genomes are essential for the mitochondrial functions that power eukaryotic life. Mitonuclear interactions can be direct, as physical contacts between mitochondrial and nuclear genes and their products are

needed for mitochondrial DNA (mtDNA) replication and maintenance, transcription, translation, assembly and function of mitochondrially encoded components of oxidative phosphorylation (OXPHOS) complexes [1]. Mitonuclear interactions can also be indirect, though anterograde (nucleus-to-mitochondria) and retrograde (mitochondria-to-nucleus) signaling where metabolites, biochemicals or RNAs direct gene expression in response to metabolic needs or environmental stressors [2, 3]. Genetic variation in natural systems can alter the efficiencies of these interactions leading to phenotypic differences between individuals and populations. Mitonuclear epistasis, defined as the non-additive phenotypic effects of interacting mitochondrial and nuclear allele pairs, has been demonstrated across Eukarya, including vertebrates [4-6], invertebrates [7-12], plants [13] and fungi [14-20]. In humans, allelic variation in mitonuclear interactions contributes to human diseases [21-25]. Identification of important mitonuclear loci will provide insight into evolutionary and coevolutionary processes and facilitate mitochondrial disease prediction and treatments.

Selection for mitonuclear interactions could contribute to speciation [26, 27]. Because of the large interest in uncovering speciation loci, strategies to identify mitonuclear epistasis often focus on analyzing mitonuclear incompatibilities in interspecific and inter-population hybrids or in mitonuclear hybrids where nuclear genomes from one population are paired with the mtDNAs from another. Sometimes candidate genes for these incompatibilities can be revealed through deductive reasoning. For example, in crosses between populations of *Tigriopus californicus* where mtDNA inheritance was controlled, F2 hybrids showed mitonuclear-specific OXPHOS enzyme activities and mtDNA copy number differences, prompting investigations into mitochondrially encoded electron transport proteins and the mtRNA polymerase [28, 29]. In *Drosophila*, a mtDNA from *D. simulans* paired with a nuclear genome from *D. melanogaster* resulted in mitonuclear genotype with impaired development and reproductive fitness [30]. Fortuitously for mapping purposes, the mtDNA sequences in this mitonuclear panel had very few polymorphisms, enabling the causative alleles behind this incompatibility (a mitochondrially encoded tRNA and a nuclear encoded tRNA synthetase) to be identified [31].

Genetic approaches can also be used to reveal mitonuclear epistatic loci. Chromosomal replacements in interspecific *Saccharomyces* yeast hybrids, followed by plasmid library screening, allowed the identification of mitonuclear incompatibilities between nuclear genes encoding intron splicing factors from one species and their mitochondrially-encoded targets in the other [17, 32, 33]. Quantitative trait loci (QTL) mapping approaches using genotype/phenotype associations of recombinant progeny containing different mtDNAs have also been used to seek intraspecific mitonuclear incompatibilities [34-37]. Analysis of meiotic segregants following forced polyploidy enabled the detection of QTLs for interspecific mitonuclear incompatibilities contributing to sterility barriers in *Saccharomyces* [38]. Due to the low resolution of QTL mapping, specific loci were not identified, but in some cases, regions of mitonuclear genomic interest were implicated. Other approaches to uncovering mitonuclear epistatic loci include differential expression analysis [5, 39-41] and experimental evolution [42, 43]. Because relatively few genetic backgrounds are used for these mapping approaches, even if single gene pairs are identified, it is not clear if these approaches will reveal a general picture of mitonuclear epistasis or lineage specific idiosyncrasies.

Mitonuclear epistasis could explain over 30% of the phenotypic variances observed in a panel of *S. cerevisiae* yeasts consisting of 225 unique mitonuclear genotypes [15]. In the current study, our goal was to map naturally occurring alleles leading to mitonuclear epistasis in this yeast population. Detecting mitonuclear epistatic loci (or any g x g interaction) using association approaches is challenging due to low allele frequencies, the large number of tests required to detect pairwise epistasis and the potential for the environment to affect genetic interactions [44-46]. To overcome some of these challenges, we created a recombinant population of *S. cerevisiae* designed explicitly to detect mitonuclear epistasis through association testing. We used controlled and random matings with 25 natural yeast isolates to create a multiparent advanced intercross population in order to reduce effects of the known population structure in yeast [47] and allow for finer mapping resolution. We controlled the mtDNA inheritance in the recombinant collection such that strains shared a single mitotype, and then replaced this mitotype with two different mtDNAs resulting in a collection of nuclear genotypes paired with three different mitotypes. This should increase power to detect

mitonuclear interactions across the full mapping population and to detect nuclear effects within a given mitotype. Mitonuclear interactions could thus be integrated into phenotype/genotype association models to detect loci contributing to complex traits.

We focused on mapping alleles contributing the stability of the mitochondrial genome, a complex trait controlled by many nuclear loci [48-52]. Large scale deletions within mtDNAs are a common form of mtDNA instability, leading to mitochondrial heteroplasmies (defined as having more than one mitotype present) and deterioration of organismal health [53-57]. *Saccharomyces* yeasts do not tolerate mtDNA heteroplasmies and will fix for a single mitotype after just a few mitotic divisions [58]. Yeast cells containing these spontaneous mtDNA deletions can grow via fermentation and will form small, *petite*, colonies in comparison to the larger *grande* colonies formed by respiring cells, making it possible to quantify rates of mtDNA deletions [59]. Laboratory strains have accumulated multiple genetic variants leading to increased *petite* frequencies [51]. We reasoned that natural genetic variation would lead to differences in *petite* frequencies among wild isolates of *S. cerevisiae* through mitonuclear epistasis and that our recombinant population would enable the identification mitonuclear epistatic loci.

Here, we show that mtDNA stability in *S. cerevisiae* populations is influenced by mitonuclear epistasis and by independent effects of both nuclear and mitochondrial genomes. We describe the construction of the Mitonuclear Recombinant Collection, and how we were able to use this to detect nuclear loci that participate in the mtDNA stability through main (independent) effects and mitonuclear interactions. The epistatic loci identified revealed that growth rates influence the stability of mtDNAs. The mitonuclear recombinant collection provides a new tool for identifying mitonuclear loci that are important in nature.

Results

mtDNA stability is a quantitative trait influenced by mitochondrial GC clusters and mitonuclear interactions.

Given the mechanistic complexities of mtDNA replication and maintenance, we hypothesized that standing genetic variation would contribute to differences in mtDNA stability. To test this, we examined the frequencies of spontaneous *petite* colony formation in 21 isolates of *S. cerevisiae* (**Figure 1**, **Table S1**). Rates of mtDNA deletions in these isolates ranged from 0.3 to 9.9%. All 21 isolates had lower *petite* frequencies than S288c, a widely used laboratory strain known to harbor nuclear variants that promote exceptionally high rates of *petite* formations [51]. Within the wild isolates, genetic variation within and between broadly-defined ancestral populations contributed to differences in the rates of mtDNA loss (P<0.001 for strain and population, **Table S2**). Strains with Wine/European or mosaic origins had the lowest rates of *petite* formation, while strains with West African or Sake/Asian origins had the highest.

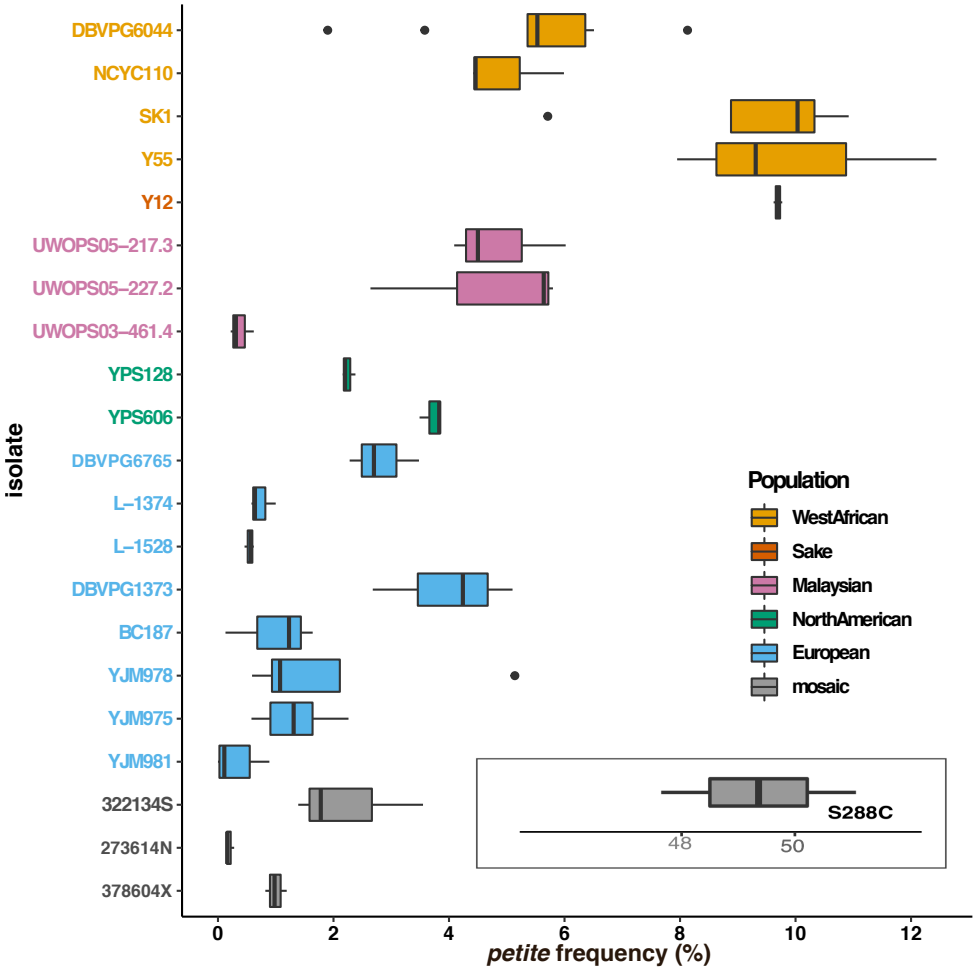

Figure 1. Genetic variation contributes to *petite* frequency. Boxplots showing *petite* frequencies (total *petite* colonies/total CFU *100) are presented for haploid derivatives of 21 wild *S. cerevisiae* isolates and the reference strain, S288c. Strains are colored by

broad population identities from [47, 60, 61] as indicated in the key. The *petite* frequency for the reference strain is offset for scaling purposes. *Petite* frequencies differ between populations and between strains within populations (ANOVA, **Table S2**).

The differences in *petite* formation rate between wild isolates could be due to genetic variation in mtDNAs, nuclear genomes and/or mitonuclear interactions. To examine the independent effects of different mitotypes on mtDNA stability, we introduced 18 different mtDNAs into a common nuclear background and assayed rates of mtDNA deletions. In these iso-nuclear strains, *petite* frequencies ranged from 5.5 to 31.0 due to differences in mitotypes (ANOVA, $P<0.001$, **Table S3**). To determine if *petite* formation correlated with any particular mitochondrial feature, we analyzed the mtDNAs for their length, intron content, GC% and GC-cluster content (**Table S4**). The GC-rich regions of yeast mtDNAs are predominantly due to short (~30-40 bp) mobile GC-rich palindromic sequences called GC-clusters that interrupt long intergenic AT-rich sequences [62, 63]. GC content and GC cluster numbers were tightly correlated ($r = 0.95$, $P<0.001$), and both were correlated with the total length of mtDNAs ($r = 0.59$, $P=0.001$ and $r=0.70$, $P<0.001$, respectively). *Petite* frequencies, however, did not correlate with the total length of mtDNAs ($r=0.12$, $P=0.63$) nor the total lengths of intron sequences ($r=-0.09$, $P=0.72$). Thus, it is likely that GC rich regions play an important role in mtDNA stability. In fact, we found that *petite* frequencies correlated with overall GC% of mtDNAs (Pearson's $r=0.59$, $P=0.013$, **Fig 2A**) but not the total numbers of GC clusters ($r=0.44$, $P=0.07$), mtDNA length ($r=0.12$, $P=0.63$) nor total length of intron sequences ($r=-0.09$, $P=0.72$). A presumed mechanism for *petite* formation is illegitimate recombination in mitochondrial GC clusters [59, 64, 65]. To see if any particular type of GC cluster could explain mtDNA instabilities, we sorted the GC clusters into 9 classes described by sequence homologies [66]. We found that the M4 GC-cluster class, but no other, correlated with *petite* frequencies ($r=0.67$, $P=0.004$, **Fig 2B**). This particular cluster appears to be expanding in the mtDNAs of strains with West African lineages [67] and explain the observed correlation.

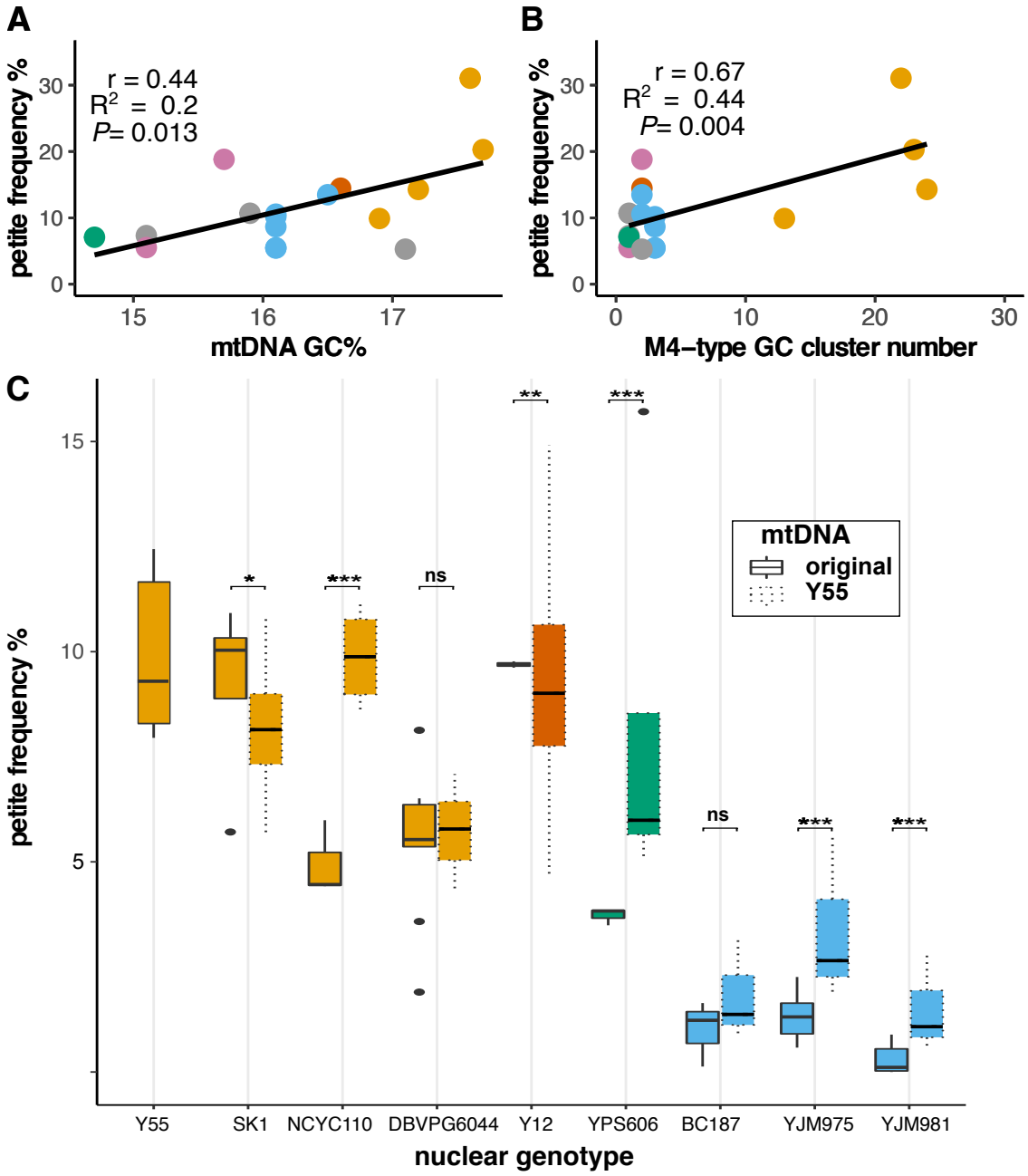

Figure 2. Mitochondrial GC-content influences stability of mtDNAs. *Petite* frequencies of iso-nuclear strains (containing the nuclear genome from strain Y55) correlate with **A.** total GC% of mtDNAs and **B.** the numbers of M4-type GC clusters. Pearson's correlation coefficient (r), coefficient of determination (R^2) and *P*-value significance are shown. **C.** *Petite* frequencies for strains containing original (solid outlines) or the GC-cluster rich mtDNA from Y55 (dotted outlines) are shown as box and whisker plots. Colors indicate nuclear genotype population as described in **Fig 1**. Significance of individual ANOVAs comparing the original and synthetic mitonuclear genotypes are shown. * *P<0.05,* ** $P \leq 0.005$, *** $P \leq 0.001$.

GC clusters, however, do not solely control rates of *petite* formation. When a mtDNA from the West African strain with the highest *petite* frequency was introduced into different nuclear backgrounds, *petite* frequencies were increased in 4 of 8 nuclear backgrounds tested and decreased in 2 of them (**Fig 2C**). This is consistent with both nuclear and mitonuclear effects**.** Nuclear backgrounds from Wine/European origins maintained relatively low *petite* frequencies even when harboring this GC-cluster rich mtDNA, indicating that nuclear genotypes largely control mtDNA stability, at least for these strains.

Some of the synthetic mitonuclear combinations led to even higher *petite* frequencies than observed in the original isolates suggesting that mitonuclear interactions also play a role in mtDNA stability. To formally test for mitonuclear effects on mtDNA stability, we examined *petite* frequencies in 16 mitonuclear cybrids created by exchanging mtDNAs between 4 strains with West African lineages (4 nuclear x 4 mtDNAs). A two-way ANOVA showed that mtDNA stability was influenced by nuclear genotypes, mitotypes, and mitonuclear interactions ($P<0.001$ for each term, **Table S5**), where certain mitonuclear combinations showed very large increases in the rates of *petite* formation (**Fig 3**). Within these strains, the original mitonuclear genotypes had lower *petite* frequencies than any of their synthetic mitonuclear combination (**Fig S1**). Given the recent expansion of the destabilizing M4 clusters within the West African strains, this observation suggests that selection for mitonuclear interactions that stabilize mtDNAs has occurred quickly in ways that are strain-specific.

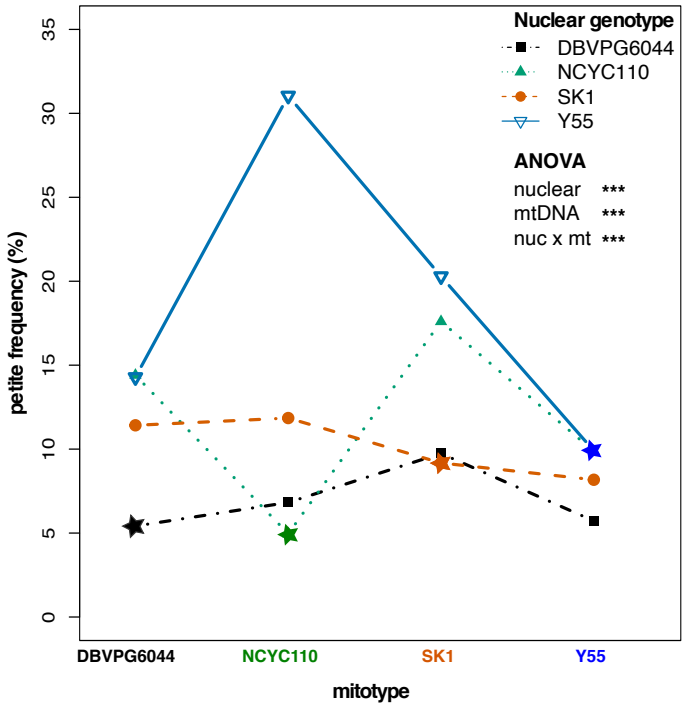

Fig 3. Coadapted mitonuclear interactions stabilize mtDNAs. *Petite* frequencies for strains containing 4 nuclear and 4 mtDNA backgrounds are shown as an interaction plot. Colored lines connect a nuclear genotype paired with different mitotypes. The original mitonuclear genotype combination is starred. The nonparallel lines indicate mitonuclear epistasis. ANOVA revealed highly significant nuclear, mtDNA, and mitonuclear genotypes contributions (**Table S5**). *** $P \leq 0.001$.

Creation of a Mitonuclear Recombinant Collection for association studies

Mitonuclear interactions explain a significant proportion of phenotypic variances in *S. cerevisiae* yeasts and involve numerous, as yet unmapped, loci [14, 15, 68]. We sought a general genome-wide mapping approach that would facilitate the mapping of both the nuclear and mitonuclear loci underlying complex traits such as mtDNA stability. We created a multiparent recombinant collection of *S. cerevisiae* strains specifically designed for association mapping of nuclear and mitonuclear loci (called the Mitonuclear Recombinant Collection, or MNRC) (**Fig 4**). To do this, we first replaced the mtDNAs in 25 wild divergent yeast isolates such that each contained an identical mitotype, and then mated them to create each possible heterozygous diploid. The diploids were sporulated and ~10,000 haploid F1 haploid progeny were isolated and then randomly mated. F1 diploids were then isolated and sporulated. The process was repeated for a total of 7 rounds of meiosis. A collection of 181 F7 haploids (named

Recombinant Collection 1 (RC1)) were isolated and fully sequenced. The mtDNAs from RC1 strains were removed (creating RCρ^0) and replaced with two additional mtDNAs via karyogamy-deficient matings, creating populations RC2 and RC3. The GC cluster content of the mtDNAs in the MNRC are classified as low (117 clusters in RC2), medium (137 in RC1), or high (203 in RC3).

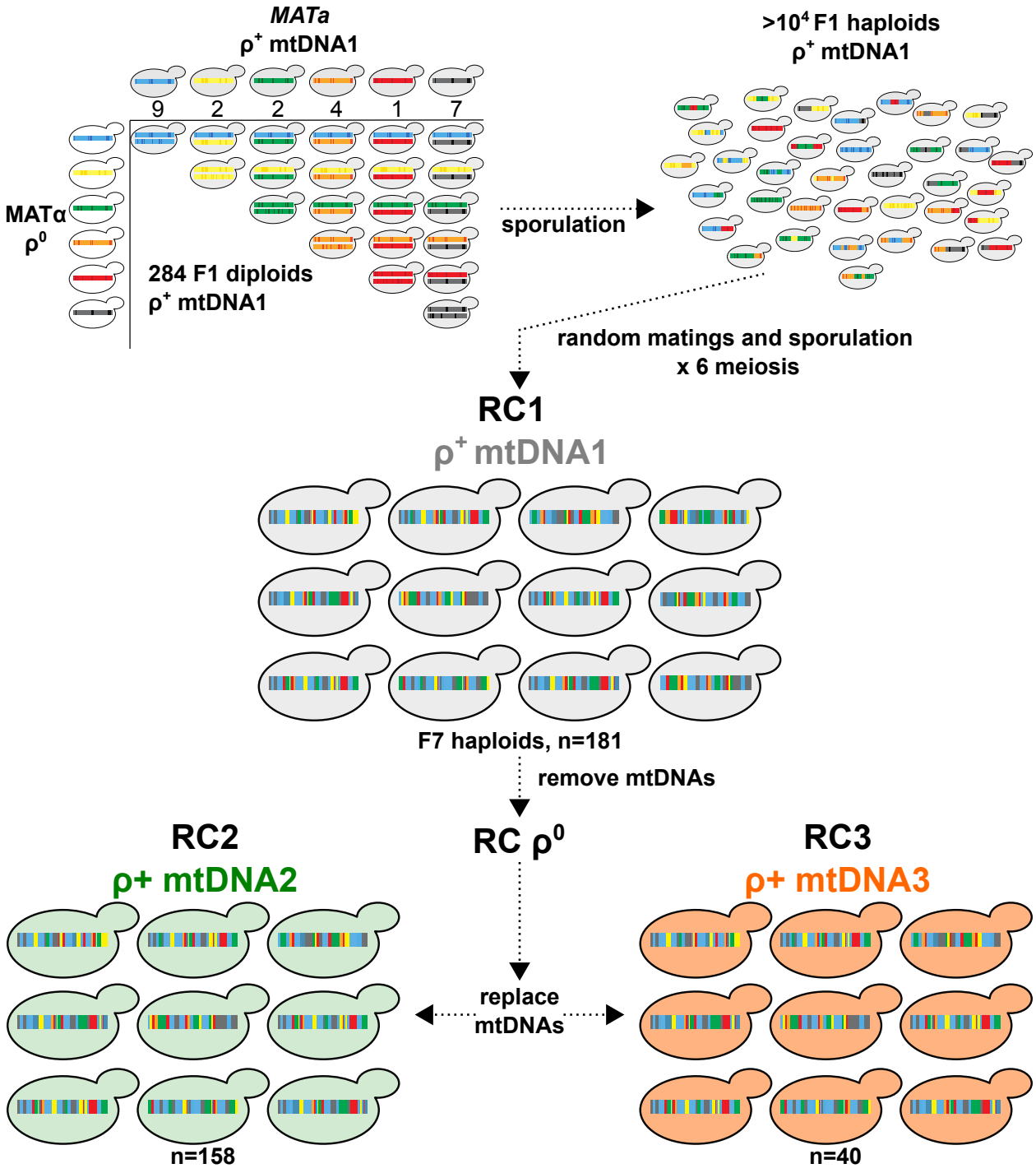

Figure 4. A Mitonuclear Recombinant Collection designed for association mapping. 25 unique genetic backgrounds of *S. cerevisiae* fixed for a single mitotype were systematically crossed to create each possible F1 heterozygous diploid with identical mtDNAs (284 total). The numbers of unique genotypes from each parental ancestral background are indicated; blue (Wine/European); yellow (Malaysian); green (North American); orange (West African); black (mosaic). The F1 diploids were sporulated and F1 recombinant haploid progeny isolated. Following 6 rounds of random matings, haploid F7 progeny were isolated to create RC1. The mtDNAs from RC1 were removed (RC1ρ^0) and replaced with 2 different mtDNAs, creating RC2 and RC3. The mtDNAs is RC1,2, and 3 are from the wild isolates 273614N, YPS606, and NCYC110, respectively.

To generate SNPs tables for association testing, RC1 strains were sequenced to ~40x coverage. Paired-end reads were aligned to the *S. cerevisiae* reference genome and the locations of nuclear SNPs and small indels were extracted from each alignment. Polymorphic sites were filtered by removing telomeric regions and SNPs/indels with low allele frequencies (MAF <5%). Following filtering, 24,955 biallelic sites across the 16 yeast chromosomes with an average of ~2200 SNPs/chromosome were available for association testing. Chromosomal polymorphic data are summarized in **Table S6**. Our read alignments and subsequent analyses did not account for chromosomal rearrangements, such as copy number variants, translocations and inversions, that would map to similar locations of the reference genome nor genomic regions absent in the reference strain.

We validated that this novel recombinant population could be used for simple association studies. Strains from RC1 were phenotyped for growth on copper sulfate and an association test was performed to identify SNP variants associated with copper tolerance. A single peak on Chr. 8 coincided with a region containing *CUP1*, the copper binding metallothionein (**Fig S2**). Variation at this locus is known to lead to high copper tolerances found in Wine/European isolates [69] and has previously been identified through association studies using wild isolates [47, 70]. Thus, the recombinant collection produced here is successful for association studies despite a relatively low number of parental strains.

Nuclear and mitonuclear associations for mtDNA stability

To map nuclear and mitonuclear associations, *petite* frequencies were collected for each strain in RC1, RC2 and RC3. In RC1, the *petite* frequencies ranged from 0.0% to 27.7% forming a continuum, as would be expected for a complex trait involving numerous loci (**Fig 5A**). The same rank orderings were not observed in RC2 or RC3, revealing the influences of mitonuclear interactions. RC3, containing the GC cluster rich mtDNA, had a higher average *petite* frequency than RC1 or RC2 (**Fig S3**).

In theory, the recombinant genomes and fixed mitotypes of the RC strains should reduce effects of population structure, improve statistical power while using a smaller number of samples, limit false positives, and control for mitonuclear interactions. We performed association testing to identify nuclear loci that had both a main effect and interacted with the mtDNA to influence mtDNA stability. Mating types, auxotrophic markers and residual population structure as determined by principal component analyses were included as covariates (see METHODS). The significance profiles of associations for nuclear variants that were independent (*nuclear SNP*, **Fig 5B**) and dependent on mitotype (*nuclear SNP× mtDNA*, **Fig 5C**) were different, providing confidence that the association model is able to detect nuclear features that are unique to either main or epistatic effects.

Nuclear SNPs whose effects were independent of mitotype resulted in stronger associations than mtDNA-dependent alleles. This is not surprising given that independent contributions of nuclear genotypes influence growth phenotypes to a greater extent than mitonuclear interactions [14, 15, 68]. At a false discovery rate (FDR) < 0.1% (Q-value = 0.001), we observed 130 mtDNA-independent associated SNPs located within or 250 bp upstream of coding sequences (**Table S7**). In comparison, at FDR <5% (Q-value = 0.05), there were 3 mtDNA-dependent associated SNPs (**Table S8**). Strong nuclear effects could mask mitonuclear interactions. We identified the alleles with the strongest effects by calculating the effect size of each SNP with mitotype-independent associations as the difference between the average *petite* frequencies of each allele weighted by its frequency in the recombinant collections (**Table S9**). This revealed that the highest effect size was attributed to a SNP on Chr. 15 predicting a G50D missense mutation in *MIP1*, the mitochondrial DNA polymerase required for replication and maintenance of mtDNA. To improve power of detecting

mitonuclear associations, we repeated the analysis including the *MIP1* SNPs as covariates. This removed the mitotype-independent associations on Chr. 15 and increased the numbers of significant mitonuclear associations from 3 to 27 without changing the overall association profiles (**Fig S4, Table S8**).

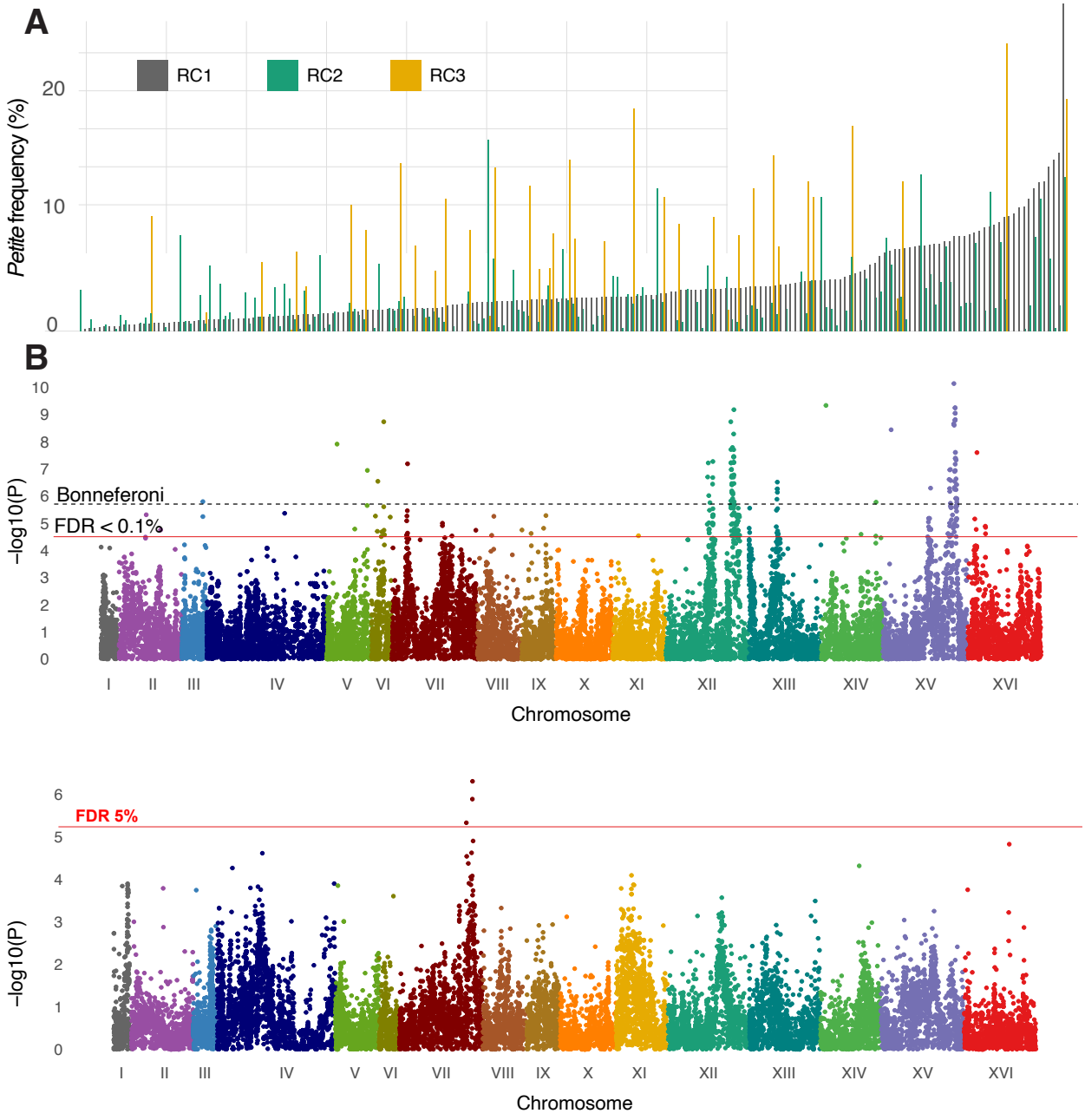

Figure 5. Nuclear SNPs associate with mtDNA stability through main effects and mitonuclear interactions. A. *Petite* frequency is a quantitative trait influenced by natural genetic variation. *Petite* frequencies for strains from RC1 (gray) were rank ordered from lowest to highest. The *petite* frequencies for strains from RC2 (green) and RC3 (gold) are plotted next to isonuclear strains from RC1. **B.** Manhattan plots show

main effect nuclear SNPs associated with *petite* frequencies in ways that were independent of mtDNA and **C.** nuclear SNPs whose association is dependent on mitotype (ie. mitonuclear)**.** The different plot profiles indicate that the main effect nuclear SNPs are different that those involved with mitonuclear interactions. FDR thresholds at 0.1% ($P < 4.1 \times 10^{-5}$) for nuclear associations and 5.0% ($P < 1.2 \times 10^{-5}$) for mitonuclear associations and a conservative Bonferroni threshold ($P < 2.0 \times 10^{-6}$) are shown.

Mitonuclear associations: mitotic growth signaling pathways

The 27 mitonuclear SNP associations corresponded to 21 unique genes, including 11 in a QTL on Chr. 7 (**Table S8**). Within this QTL, the three strongest associations corresponded to SNPs near the coding start sites of the genes *HGH1* and *SMI1* and an in-frame deletion within *BNS1*. Interestingly, mitochondrial activities have not been shown for these genes. A second QTL on Chr. 1 contained a SNP upstream of *YAT1*. To verify the involvement of these genes in mitonuclear interactions affecting mtDNA stability, we first looked at how removing each gene influenced *petite* frequencies. In the parental background of the *S. cerevisiae* knockout collection (BY4741) and with the BY4741 mitotype, the null mutant *hgh1*Δ lowered rates of *petite* formation while null mutants *bns1*Δ*, smi1*Δ and *yat1*Δ showed no significant effect (**Fig 6A**). We next introduced two different mitotypes (the GC cluster-rich mitotype NCYC110 used in RC3 and the GC cluster-poor mitotype YPS606 used in RC2) into the parental and deletion strains and measured *petite* frequencies. Mitonuclear interactions were readily observed; the GC-rich mitotype led to higher *petite* frequencies in the *hgh1*Δ and *yat1*Δ deletion strains, whereas the same mitotype led to a lower *petite* frequency in the parental background (**Fig 6B**). Two-way ANOVAs showed significant mitonuclear interactions when comparing the parental strain to each deletion strain with each mtDNA comparison (**Table S10**). Similar mitonuclear effects were also observed for *bns1*Δ and *smi1*Δ when compared to the parental background (**Fig S5, Table S11**). While the high *petite* frequency of the parental strain complicates the interpretation of these assays, the differential responses provide strong evidence that the association model was successful in identifying genes that influence mtDNA stability through mitonuclear epistatic interactions.

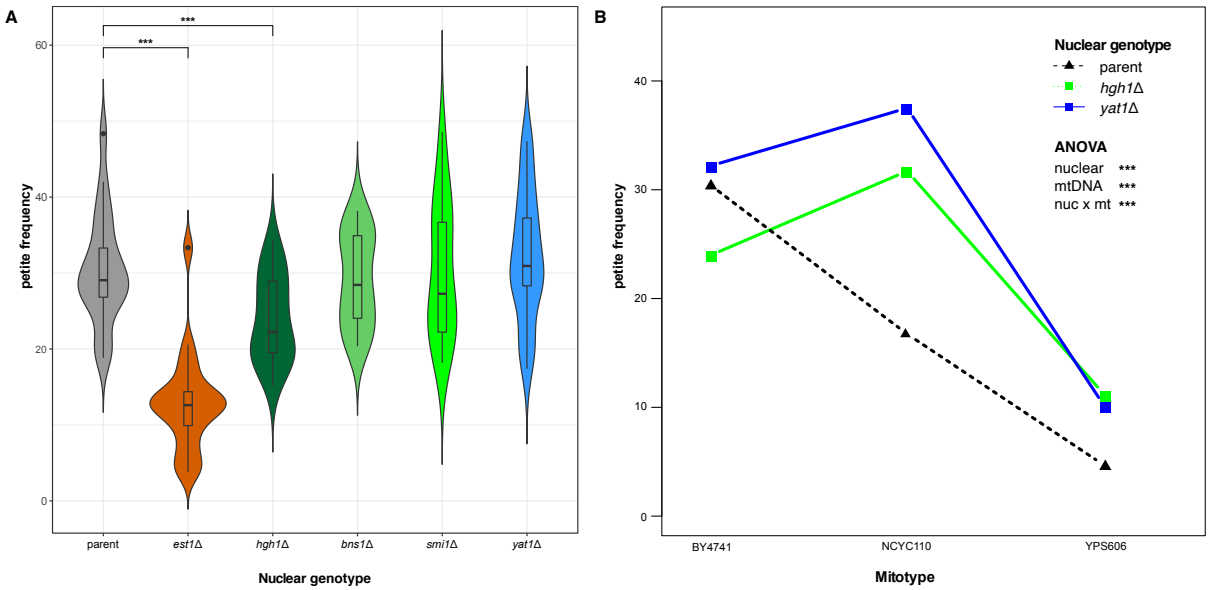

Figure 6. Mitonuclear interactions influencing mtDNA stability include *HGH1* and *YAT1*. A. *Petite* frequencies for strains lacking genes with mitotype-independent (*est1*Δ) and mitonuclear associations (*hgh1*Δ, *bns1*Δ, *smi1*Δ and *yat1*Δ) are shown as boxes in violin plots. Significant differences between the parental and each deletion strain are shown. * $P<0.05$, ** $P \leq 0.005$, *** $P \leq 0.001$. **B.** *Petite* frequencies of *hgh1*Δ and *yat1*Δ strains depend on mitotype. Colored lines connect each nuclear genotype paired with different mtDNAs. Non-parallel lines indicate that mitotypes influence mtDNA stability through mitonuclear interactions. Comparisons of the parental and each deletion strain with any 2 different mtDNAs showed highly significant mitonuclear interactions ($P<0.001$ for each comparison, **Table S10**). Mitotype BY4741 is the parental mtDNA in the yeast deletion collection. Mitotypes NCYC110 and YPS606 were used in RC3 and RC2, respectively. A minimum of 20 replicates for all *petite* assays were performed.

We then tested if the mitonuclear associations could be explained by differences in gene expression. mRNAs were extracted from a subset of 18 strains from the RC2 and RC3 collections containing GC-cluster poor and rich mitotypes, respectively. We attempted to control for *MIP1* alleles and the 2 most common SNP haplotypes for mitonuclear candidate genes on Chr. 12 (containing *HGH1, BNS1, and SMI1)* and Chr. 1 (containing *YAT1*). Across the subset of RC2 and RC3 strains containing alternate mitonuclear haplotypes with each *MIP1* allele, *YAT1* SNPs (Chr. 1) were in complete LD with the associated SNPs on Chr. 12 and are thus considered together. Expression of *HGH1* showed a modest positive correlation with *petite* frequencies (r=0.43, P=0.08)

(**Fig S6**). Expression of *MIP1, YAT1*, *BNS1* and *SMI1* showed slight but non-significant positive correlations (**Fig S6**). In these strains, mitotypes did not influence expression of the mitonuclear candidate genes nor on the expression of the mtDNA polymerase, *MIP1* (**Table S12**). The expression of *MIP1* was, however, influenced by the haplotype of the Chr. 12/1 QTLs (**Fig S7A** while expression of *BNS1* was influenced by *MIP1* alleles (**Fig S7B**, **Table S12**). *BNS1* and *YAT1* expression showed *MIP1* haplotype x mitonuclear haplotype effects. Interestingly, the higher expression in these genes were observed in strains containing the *MIP1* alleles that led to a high *petite* frequency. These data suggest that the genotypes of *MIP1* and the mitonuclear candidate genes may influence the expression of each other to affect mtDNA stability.

The mitonuclear candidate genes, *YAT1, HGH1*, *BNS1*, and *SMI1* are broadly connected to mitotic growth. To investigate a potential relationship between mitotic growth and mtDNA stability, we compared the growth rates of wild yeast isolates and their *petite* frequencies and found that the faster growing strains were more likely to lose their mtDNAs (**Fig 7A**). Increasing growth rates by changing glucose availability or temperature also led to an increase in *petite* frequencies (**Fig 7B)**. This suggests a fitness tradeoff between rapid cell growth and deletions of mtDNA. We also noted that increase in *petite* frequencies of the West African strains with synthetic mitonuclear genotypes (**Fig S1**) coincided with an increase in growth rates (**Fig S8**). Selection for mitonuclear interactions that stabilize mtDNAs may act by lowering overall growth.

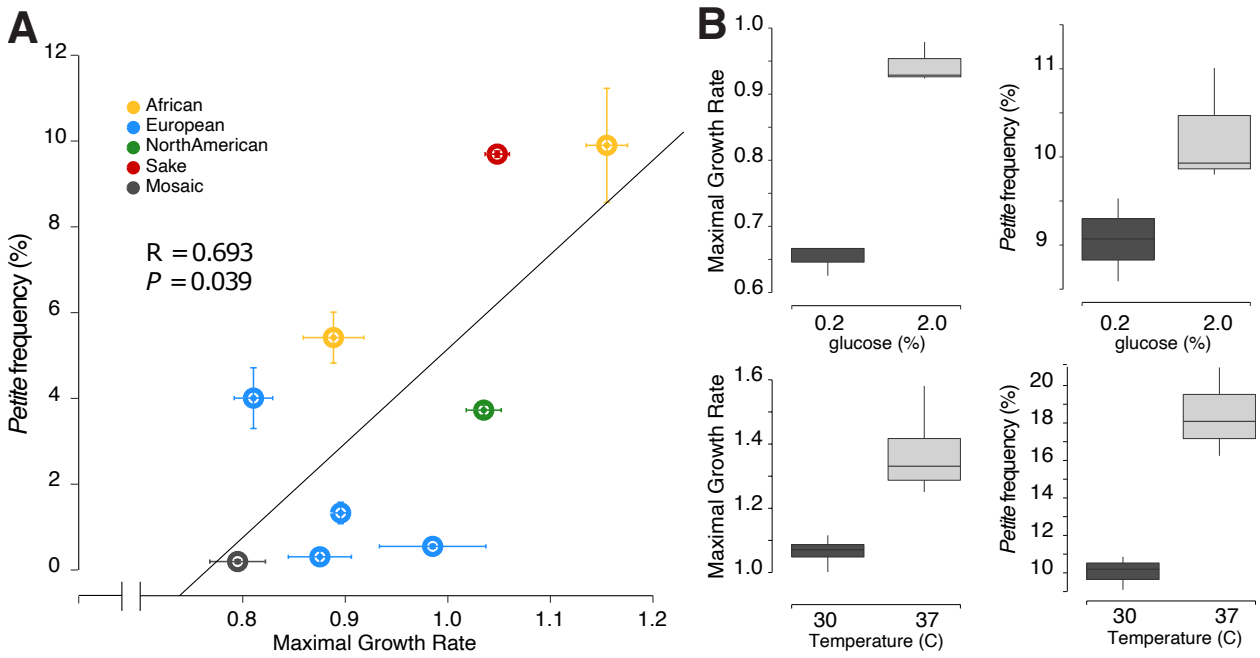

Figure 7. Rapid mitotic growth increases rates of mtDNA loss. A. Growth rates of wild isolates and spontaneous *petite* formation are positively correlated. Maximal growth rates (from [14]) of strains grown in conditions used for *petite* assays were plotted against rates of spontaneous *petite* formation. **B.** Growth rates and *petite* frequencies increased as a result of increased glucose concentration or temperature. Assays were performed on Sake strain Y12 (NCYC3605) a minimum of three times. Growth rates are reported as mOD/min.

Mitotype-independent associations: mtDNA polymerase and a G-quadruplex stabilizing protein

In addition to identifying mitonuclear loci, our model identified 301 SNP whose associations were independent of mitotype. Of these, 130 SNPs were located within 250 bp upstream of coding sequences for 86 unique genes (**Table S7).** Gene Ontology (GO) analyses revealed enrichments in mitochondrially-targeted genes (27 genes), genes involved in mitochondrial organization (9 genes) and homeostatic processes (10 genes) (**Table S13**). Sixteen of the associated genes generate non-respiratory phenotypes when deleted from the genome [49, 71-73]. In all, 33 (38.4%) of the mitotype-independent associations have known mitochondrial activities.

A large number of the associated SNPs occurred within QTLs on chromosomes 12, 13 and 15 (**Fig 5B**) and were likely associated through linkage. To help identify causative loci, effect sizes for each SNP were calculated as the difference between the

441 average *petite* frequencies of each allele weighted by its frequency in the recombinant
442 collections (**Table S9**). The highest effect size was attributed to a SNP on Chr. 15
443 predicting a G50D missense mutation in *MIP1*, the mitochondrial DNA polymerase
444 required for replication and maintenance of mtDNA. Associations for two additional
445 *MIP1* missense variants (T540M and H541N) were also significant. High effect sizes
446 were also seen in associated SNPs in genes flanking *MIP1*, including *FRT1* and
447 *PDR10*, and essential genes *KRE5* and *ALA1*. Deletions of *FRT1* or *PDR10* did not
448 significantly alter *petite* frequencies suggesting that these associations are the result of
449 linkage to *MIP1* (**Fig S9**).
450 Thirty-two associated genes were located within two QTLs on Chr. 12. Here, the
451 largest effect size was for a putative missense variant (L13I) in *EST1* (**Table S9**),
452 encoding a telomerase accessory protein that helps to stabilize of G-quadruplex
453 structures in telomeres during certain stages of the cell cycle [74]. We found that a
454 deletion of *EST1* significantly reduced *petite* frequencies as compared to the parental
455 strain in the yeast knockout collection (**Fig 6A**) while a deletion of *TOP3,* a candidate
456 gene tightly linked to *EST1*, did not (**Fig S9**). This suggests that *EST1* variants likely
457 drive the associations in this QTL. A second QTL on Chr. 12 contained associations for
458 5 genes with known mitochondrial functions. *ILV5* is a nonspecific mtDNA binding
459 protein involved in the packaging of mtDNA into nucleoids and is required for mtDNA
460 stability [75]. *NAM2,* a mitochondrial tRNA synthetase [76], and *SSQ1*, a mitochondrial
461 chaperone [77], are required for growth via mitochondrial respiration. We found that a
462 deletion of *MDM30,* a component of the ubiquitin ligase complex involved in
463 mitochondrial morphology and turnover [72], lowered *petite* frequencies while a deletion
464 of *ATG33,* a protein involved in mitochondrial mitophagy (KANKI *et al.* 2009), had no
465 effect (**Fig S9**). A single QTL on Chr. 13 contained associations for 6 genes, four of
466 which encode mitochondrially targeted proteins. *COX14* and *SAM37,* and *SOV1* are
467 required for mitochondrial respiration [49, 71, 72] and *RNA14* is essential for growth.
468 Deletions of a 5th gene, *CSM3*, did not significantly alter *petite* frequencies (**Fig S9**).
469

470 **Discussion**

471 **Mitonuclear Recombinant Collection**

We created a multiparent advanced intercross population (Mitonuclear Recombinant Collection or MNRC) to identify mitonuclear interactions driven by allelic variation in *S. cerevisiae* yeasts through association testing. The MNRC incorporates the genetic variation from 25 wild isolates into ~200 fully sequenced recombinant strains. These strains were paired with three different mitotypes, enabling the detection of mitonuclear interactions via genome-wide association tests. Using the MNRC, we were able to identify loci contributing to mtDNA stability through independent (main) effects and mitonuclear interactions.

The MNRC was designed to help overcome many of the challenges of detecting epistatic interactions. Multiple rounds of random mating should lower existing LD created by the strong population structure in *S. cerevisiae* [47, 61] and reduce the numbers of strains needed for association testing. The MNRC strains are haploid, eliminating dominance effects. Because novel mtDNAs created through mitochondrial recombination may produce unexplained phenotypic variances [19, 68], mtDNA inheritance was controlled. By specifically pairing each mitotype with the nuclear recombinants, the power to detect specific mitonuclear loci was increased. Environments can easily be controlled in yeast, eliminating unexplained variances due to mtDNA x nuclear x environment interactions. We showed that the MNRC can successfully detect nuclear alleles that contribute to phenotypes through independent (main) or epistatic (mitonuclear) effects.

Loci affecting mtDNA stability

Mitonuclear interactions

Our model found epistatic associations for SNPs in four genes (*HGH1, SMI1, BNS1,* and *YAT1*) that are broadly related to mitotic growth. We confirmed these associations by showing that mitotypes influence *petite* frequencies differently in null mutations of each gene as compared to a wild type genetic background. *YAT1* encodes a mitochondrial outer membrane protein that participates in mitochondrial metabolism by importing the acyl groups that enter the Krebs cycle for energy production [78, 79]. Industrial yeast strains have likely adapted to high levels of toxic aldehydes produced during ethanol production by increasing *YAT1* copy numbers [80]. The Yat1 protein is

phosphorylated [81], suggesting that its activity is regulated through cell signaling. *HGH1* encodes a translation factor chaperone involved in protein synthesis that is activated in response to DNA damage [82] [83]. The molecular function of *SMI1* is unknown, but it is thought to be a member of the cell signaling pathway that regulates cell wall biosynthesis during mitotic growth [84]. Null mutations of *SMI1* have reduced respiratory growth [85], suggesting its activity influences mitochondrial function. During mitosis, *BNS1* participates in the signaling network directing the exit from anaphase [86]. A high-quality mitochondrial proteomics screen found the Bns1 protein in the mitochondrial matrix [87]. Possibly, these genes are involved in mitonuclear epistasis by changing growth parameters in response to retrograde signals produced by mitochondria with damaged mtDNAs. We expect this to be a general trend. Dominant mutations that suppressed the low growth phenotype of strains with mitochondrial ribosomal protein defects also led to higher *petite* frequencies [88].

Mitotic growth also linked mitonuclear epistasis and mtDNA stability when exchanging mtDNAs between strains from West African lineages. This strain grouping had higher *petite* frequencies (**Fig 1**) and growth rates (**Fig 7**) than most other strains. When mtDNAs were exchanged between West African isolates, strains with original, coadapted mitonuclear genome combinations had lower *petite* frequencies (**Fig 3**) and mitotic growth rates (**Fig S8**) than strains harboring introduced, non-coadapted, mtDNAs. This suggests that selection for mtDNA-stabilizing mitonuclear alleles is rapid and may come at the cost of lowering overall growth rates and may even dictate an upper limit for optimal growth. It is interesting to note that mtDNA instability is a hallmark of fast growing human cancer cells [89]. Experimental evolutions aimed at altering growth rates may be one way to demonstrate this potential fitness tradeoff. In yeast cells exposed to oxidative stress agents, partial deletions within mtDNAs may initially be under genetic control, as a way to quickly reduce endogenous ROS levels by preventing OXPHOS [52]. We found that growth rates and *petite* frequencies were increased by altering environmental conditions. It is possible the increased OXPHOS requirements of rapidly dividing cells may also stimulate this retrograde signaling activity, resulting in higher rates of mtDNA deletions.

In constant environments, there should be rapid selection for variants providing optimal growth rates. This selection may occur more often in yeast where cells are likely to experience singular environments over a generation time. There is some evidence for this. When grown in media attempting to emulate natural habitats, we previously showed that coadapted mitonuclear genome combinations had higher growth rates [15]. Because there is a wide landscape of potential alleles involved in growth, different coadapted mitonuclear loci may evolve rapidly in different populations. In yeast, this could be important in maintaining allelic variation across the species. It will be interesting to see whether growth is a general feature of mitonuclear loci contributing to phenotypes other than mtDNA stability. Given that mitochondria are central to maintaining cellular homeostasis, this may be likely.

Main effects

Loci with main effects for mtDNA stability detected by our model interact with mtDNA through non-specific binding. In the MNRC, missense alleles of the mtDNA polymerase, *MIP1*, had the strongest effects on *petite* frequencies. It is well documented that the *MIP1* allele in the yeast reference strain, and other experimentally derived variants, lead to a decrease in mtDNA stability [51, 90, 91]. Our work suggests that naturally occurring variants contribute to basal differences in mtDNA stability in populations. Low expression of the mammalian homolog, PolG, leads to the accumulation of mtDNA deletions [92, 93]. Consistent with this, we found that expression of *MIP1* was lower in strains with the low *petite* frequency alleles and positively correlated with *petite* frequencies (**Fig S6, S7**), although not to statistical significance. Interestingly, *MIP1* expression was dependent on the genotypes of the mitonuclear associated genes, and *BNS1* expression was dependent on the genotype of *MIP1* (**Table S12**). From this data, it can't be deduced whether cells alter *MIP1* expression in response to growth differences caused by genotypes of mitonuclear loci or if expression of mitonuclear genes is regulated by Mip1 levels. Significant mitonuclear genotype x *MIP1* genotype interactions influenced the expression of the mitonuclear associated genes *BNS1* and *YAT1* (**Table S12**) and suggest that higher order interactions are involved in mtDNA stabilities.

We also found that *EST1* variants associated with mtDNA stability independently of mitotype and that its deletion altered *petite* frequencies. Est1 maintains telomerase at the linear ends of nuclear chromosomes [94, 95] by stabilizing G-quadruplex secondary structures in the DNA [96]. G-quadruplexes play a role in stabilizing human mtDNAs and can cause mtDNA polymerase to stall [97]. The GC-clusters in yeast mtDNAs likely form secondary structures [63] though it is not known whether these form typical G-quad structures or whether they act to stabilize mtDNAs. While Est1 is normally targeted to the nucleus, a proteomics experiment found that Est1 coimmunoprecipitated with mitochondrial ribosomes just prior to cell division (in G2 phase) [98]. It is plausible that Est1 plays dual roles in nuclear and mitochondrial genome maintenance. Another non-specific mtDNA binding protein identified by our association model (but not verified here) was *ILV5*, a mitochondrial nucleoid associated protein that was previously shown to influence mtDNA stability [99].

While higher GC content typically stabilizes DNAs, we found that mtDNAs with the highest GC-content were the least stable (**Fig 2A**). This is consistent with the observation that *petite* frequency differences between different yeast species correlates with their GC-cluster content [100]. Within *S. cerevisiae,* the M4 family of GC-clusters appears to be expanding in the mtDNAs with West African lineages by targeting another family of clusters [67] and likely explains the high *petite* frequencies in these strains (**Fig 2B**). Not much is known about functional differences between different categories of these mobile elements [63, 66]. Previously, we carefully aligned 9 mtDNAs, including their long intergenic regions, and showed that only a small number of GC clusters were in conserved positions [67]. Interestingly, the M4 clusters in the West African mtDNA in that alignment interrupted 15% (2 of 13) of the conserved clusters. The M4 clusters could destabilize mtDNAs directly or perhaps their expansion has interrupted genome stabilizing functions provided by the positions of the conserved clusters.

Significance

By focusing on naturally occurring genetic variation in *S. cerevisiae*, the alleles mapped using the MNRC provide insight into evolutionary potential. We found that mitonuclear interactions affecting mitotic growth contribute to mtDNA stability although the

595 significance of these specific interactions in yeast populations is not yet known and will
596 depend on allele frequencies and relative effect sizes in context with environments in
597 mating populations. Mitochondrial and nuclear sequences follow different evolutionary
598 trajectories in yeast populations [101] and much of the genetic variation is unique to a
599 single population [47]. It is likely that the MNRC contains combinations of interacting loci
600 that do not exist in nature. Large scale intra- and interpopulation surveys will be
601 required to determine how selection on mitonuclear interactions has shaped patterns of
602 diversity across geographic ranges and ecological niches. Thus, it is not clear that the
603 alleles identified here contribute to population specific phenotypic differences but it is
604 clear that mitonuclear interactions have shaped phenotypic diversity. In addition, the
605 mitonuclear loci identified using the MNRC provide insight into pathways and
606 mechanisms that generate phenotypic differences. The MNRC offers a powerful tool to
607 identify mitonuclear interactions and helps us better understand and predict the complex
608 genotype/phenotype relationships that shape life.
609
610 **Methods**
611 *Yeast strains*
612 All strains are listed in **Table S1**. Wild yeast isolates, described in [102], were obtained
613 from the National Collection of Yeast Cultures General. Deletion strains [103] were
614 obtained from the Yeast Knockout Collection (Horizon Discovery). To replace mtDNAs,
615 karyogamy deficient matings were performed as previously described [15].
616
617 *Media*
618 Media recipes include: SD (6.7 g/L yeast nitrogen base without amino acids, 20 g/L
619 glucose); CSM with or without amino acids as specified (SD + 800mg/L CSM premix or
620 as recommended by the manufacturer (Sunrise Science)); (CSM containing 30mL/L
621 ethanol and 30mL/L glycerol instead of glucose); sporulation media (1% IOAc, 0,1%
622 yeast extract, 0.05% dextrose); YPD (10g/L yeast extract, 20g/L peptone, 20g/L
623 glucose) supplemented with 10 mM CuSO4 when indicated; YPEG (YPD + 30mL/L
624 ethanol, 30mL/L glycerol instead of glucose); YPDG (1% yeast extract, 2% peptone,

0.1% glucose, 3% glycerol). Sugar concentrations in CSM media were altered as indicated. For solid media, agar was added to 2% prior to autoclaving.

Petite *frequency assays*

To assay rates of mtDNA deletions, 5 mL YPD cultures were inoculated with freshly grown colonies, grown in roller tubes at 30°C for exactly 15.0 hrs, diluted and plated onto YPDG solid media. After 2-3 days at 30°C, large (*grande*) and small (*petite*) colonies were counted manually or photographed and counted using an imaging system (sp-Imager-SA, S&P Robotics, Inc). To quantify *petite* frequencies in the deletion strains, 20 freshly grown colonies were scooped from solid media, diluted and plated directly onto YPDG solid media. Single assays were performed for each strain in the recombinant collections. All other assays were performed in 3-20 replicates. *Petite* frequencies are reported as the ratio of *petite* to *grande* ([*#petite colonies/ total colonies*] * *100*).

Growth phenotyping

To test to the effects of sugar and temperature on growth, strains were cultivated in 96 well microtiter plates using Biotek Eon photospectrometers using double orbital shaking. Optical densities (600nm) were recorded at 15-minute intervals until cells reached stationary phase. Maximal growth rates (V_{max}) were determined as the highest slope of regression lines modeled over sliding windows of 5 data points from growth curves and normalized to the mean of a reference strain included on each plate.

Multiparent recombinant strain collection

An F7 multiparent recombinant collection of *S. cerevisiae* yeast derived from 25 wild isolates and paired with 3 different mtDNAs was created for genome-wide association studies. We first created a set of parental strains containing a single mitotype with selectable markers to facilitate matings. We began with haploid derivatives (*MAT*a and *MATα ura3::KanMX)* representing 25 wild yeast isolates (**Table S1**). To facilitate the selection of diploids between these haploid strains, an *arg8::URA3* disruption cassette from *BamHI* linearized plasmid, pSS1, was introduced into each *MATα* strain through

chemical competence (EZ Yeast Transformation Kit (Zymo Research) or [104]) or electroporation [105]. Transformants (strains MLx2x1UA-MLx28x1UA) were selected on CSM-ura solid media and arginine auxotrophies and respiratory competencies were verified by replica plating to CMS-arg and YPEG, respectively. Correct integration of the *arg8::URA3* disruption cassette was verified through tetrad analysis: each *MATα arg8::URA3* strain was mated to their *MATa ura3::KanMX* isogenic counterpart, diploids were selected on SD media and sporulated on SPO media. Spores were dissected from ≥ 12 tetrads and printed to CSM-ura, CSM-arg, and YPEG media, verifying a 2:2 segregation of Arg- Ura+: Arg+ Ura- phenotypes. All crosses had >90% spore viabilities. The mtDNAs in the *MAT*a strains were replaced with the mtDNA from 273614N using karyogamy-deficient matings as previously described [15]. The mtDNAs from the *MATα arg8::URA3* strains were removed through ethidium bromide treatment.

To create the multiparent recombinant collection, the MATa *ura3::KanMX* ρ^+ and MAT*α ura3::KanMX arg8::URA3* ρ^0 strains were mated to create each possible heterozygous F1 diploid. To do this, aliquots (50 µL) of haploid strains were mixed with 200 µL fresh YPD media and incubated without mixing for 2 days at 30°C in 96-well plates. Mating mixtures were harvested by centrifugation, washed (200 µL ddH_20), re-suspended in 250 µL CSM-URA-ARG liquid media and incubated for 2 days at 30°C with continuous shaking, and then spotted onto solid CSM-URA-ARG media. This yielded 279 F1 heterozygous diploids (of 284 attempted crosses). To create F1 haploids, 10 µL aliquots of diploid-enriched mixtures were spotted to sporulation media and incubated at room temperature until tetrads were visible via compound microscopy (7-11 days). The sporulated cells were collected and pooled by washing the plates with 5mL ddH_20. To isolate spores from asci and vegetative cells, a random spore analysis protocol was followed [106], with modifications. Cell walls were digested by incubating 1 mL aliquots of the sporulated cells with zymolyase 20T (1 mg/ml)) at room temperature for 1 hr with gentle rocking. The cells were centrifuged (12000 rpm, 4°C, 1min), washed (1 mL ddH_20) and resuspended in 100 µL ddH_20. The treated cell mixture was vortexed vigorously for 3 minutes to adhere spores to tube walls. The supernatant was carefully aspirated, and the spores were gently washed (1mL ddH_20) to remove remaining

vegetative cells. To release spores from the plastic tube walls, cells were sonicated for 10-20 sec at 110 V (Stamina XP s50.0.7L, SharperTek) in 1mL of 0.02% Triton-X. The released spores were centrifuged (12000 rpm, 4°C, 1min), washed (1mL ddH_20), and re-pelleted. The resulting freed spore mixtures were combined into a single tube with 1mL ddH_20. Spore density was determined with a hemocytomer and plated for single F1 haploid colonies (~500 CFU/plate) onto 10 petri plates containing CSM-URA media (selecting for *MATa* or *α ura3 arg8::URA3*) and 10 petri plates containing CMS-ARG media (selecting for *MATa* or *α ura3 ARG8*). The F1 haploid cells were pooled by washing the colonies from each plate using ~5 mL ddH_20. The pooled cells were washed in 5 mL ddH_20, resuspended in 5mL YPD media and incubated at 30°C without shaking. The random mating mixtures were washed in 5 mL ddH_20, resuspended in 5 mL CSM-URA-ARG and incubated for 8-12 hours to enrich for F2 diploids. The cell mixtures were pelleted, resuspended in 2.5 mL ddH_20, aliquoted (250 ul) onto solid sporulation media and incubated at room temperature for 7 days. Spores were released and F2 haploids were isolated as described above. Random matings and haploid selection were repeated for a total of 7 meioses. Haploid F7 progeny were isolated as single colonies on YPD media, and replica plated to CSM-ARG, CSM-URA, and YPEG to determine auxotrophies and respiratory growth. Mating types were tested using mating type testers. Approximately 200 verified haploid recombinants were selected as strains for Recombinant Collection 1 (RC1) and include ~50 isolates of each selectable genotype. RC001-RC049: *MATa ura3 ARG8;* RC102-146: MATα *ura3 arg8::URA3;* RC201-247: *MAT*a *ura3 arg8::URA3;* RC301-347: MATα *ura3 ARG8.*

The mtDNAs from each strain in RC1 were removed using an ethidium bromide treatment (creating RC001-347 ρ^0) and replaced using karyogamy deficient matings to create RC2 and RC3. RC2 strains contain the mtDNA from YPS606 and were created using mitochondrial donor strains SPK27 (for RC2:001-048), CK520E1 (for RC2:102-146 and RC2:301-347) and MKG109 (for RC2:301-307). RC3 strains contain the mtDNA from NCYC110 and were created using mtDNA donor strain TUC131.

Genome Sequencing and analysis

Genomic DNAs from 192 F7 recombinants from RC1 were isolated following Hoffman-Winston genomic DNA protocol (Hoffman & Winston 1987). DNA samples were concentrated using ZR-96 Genomic DNA Clean & Concentrator^{TM} – 5 (Zymo Research Corp). DNA concentrations were measured using Qubit and diluted to the final concentration of 0.2 ng/µl. DNA sequencing libraries were prepared using an Illumina Nextera XT DNA Library Prep Kit according to manufacture instructions. The DNA concentration of each library was determined using Qubit, normalized, and pooled. The pooled libraries were sequenced using a single run of paired end 2 x 150 bp on Nexseq500 (Illumina) at the Institute for Biotechnology and Life Sciences Technology at Cornell University. The reads for each RC strain were individually mapped to the annotated reference genome S288C (R64-2-1_20110203) using Bowtie2. Single nucleotide variants (SNPs) and regions of low and high coverage were identified using the Find Variations/SNPs function in Geneious v.8, filtering the results to regions with a minimum coverage $\geq$5, minimum variant frequencies within the reads $\geq$ 0.8 and the maximum variant *P* values (the probability of a sequencing error) $\leq 10^{-6}$. To create a SNP table for association mapping, adjacent polymorphisms were merged, telomeric regions were removed and polymorphisms from each strain were combined into a single file. Low coverage areas were converted to deletions and the polymorphisms were filtered for biallelic variants with MAF > 0.5%. This resulted in in 24,955 polymorphic sites.

The mtDNA sequences were analyzed for GC cluster and intron content as previously described [67]. Accession numbers are provided in **Table S4**.

Statistical analysis and association testing

Statistics and plotting were performed using R 4.1.2 [107] and association models were performed using R 4.0.4 in RStudio [108]. Generalized linear models of the family binomial accounting for the different numbers of *petite* and *grande* colonies were performed using *glm* with the *cbind* function. ANOVAs for the analyses of growth or expression were performed using linear models with the lm function. Correlations were run using the *corr.test* function. Gene ontology classes for associated genes were identified using GO::TermFinder (Boyle, 2004

748 https://doi.org/10.1093/bioinformatics/bth456) and tested for over-representation using
749 Fisher's exact tests.
750
751 *Genome-wide association analysis*
752 A biallelic SNP table containing 24955 unique SNPs was used for association testing.
753 To address residual population structure in the recombinant strains, a principal
754 component analysis was performed using PLINK v1.9 [109]. A simple association test
755 was performed on the growth parameters from RC1 strains grown on YPD + 10 mM
756 CuSO4 using the linear model: *colony size ~ covariates + nuclear SNP + error* where
757 colony size is the strain mean derived from the linear model, nuclear SNP represented
758 SNP variation at a given locus, and covariates included auxotrophies, mating type, and
759 the first ten principal components from the PCA analysis. Association tests to detect
760 mitonuclear interactions were performed using the strain means for *petite* frequency
761 from RC1, RC2 and RC3 in the following model:: *phenotype ~ covariates + nuclear SNP*
762 *+ mtDNA + nuclear SNP*mtDNA+ error*, where phenotype is either the petite
763 frequencies. Covariates included MIP1 alleles when indicated. False discovery rates
764 were calculated using the "*qvalue"* package version 2.22.0 [110].
765
766 Associated SNPs that crossed FDR threshold were further filtered to non-synonymous
767 SNPs within coding sequence (CDS) regions, and upstream regions within 250bp
768 upstream of the CDS. Effect sizes for the associated SNPs with main effects were
769 determined as the absolute differences in *petite* frequencies for strains with each variant
770 weighted by their allele frequencies (effect size =Δ= |(freq $_{allele1} \times$ *petite* $freq_{allele1}$) – (f
771 $_{allele2} \times$ *petite* $freq_{allele2}$)|).
772
773 *RT-qPCR*
774 Cells from overnight CSM cultures were lysed according to [111] and total RNA was
775 isolated using Qiagen RNeasy mini spin columns. RNA was quantified (Invitrogen Qubit
776 Fluorometer) and diluted to 92 ng/µL. BioRad iScript cDNA Synthesis Kit was used to
777 create cDNA, according to manufacturer's instructions. RT-qPCR was performed on a
778 BioRad CFX Connect Real-Time PCR Detection System by adding 2 µl (~2.3 ng/µl)

cDNA to 10 ul of BioRad SsoAdvanced Univerisal SYBR Green Supermix, 1 µL each of forward and reverse primers (500 nM) and 6 uL nuclease-free H_20 in BioRad Hard-Shell 96-Well PCR Plates according to the following protocol: [95C-30s, (95C-15s, 60C 15-40s)$_{40}$, 65C-95C-0.5C/5s]. Primer sequences are provided in **Table S14.** Relative expression was determined as the residuals from a linear regression of 1/Ct for each candidate gene against 1/Ct for the control gene, UBC6 as previously described [112]. This approach controls for differences in RNA extraction and cDNA manufacturing efficiencies and results in more normally distributed data where higher residual values correspond to higher starting mRNA levels.

Data Availability

Reads for the MNRC strains can be retrieved from NCBI under the BioProject ID PRJNA871925. Scripts, SNP table and data available at https://github.com/mito32/Mitonuclear-Recombinant-Collection.

Author Contributions

HLF, ACF, and KC conceived the project. ML, MG, and FR created the recombinant collection. Library preps were performed by THMN, JFW, and WL. Phenotyping was performed by THMN, MG, FR, WL, ATB, MS, BB, MT, BG. Data analysis was performed by THMN, ATB, JFW, ACF and HLF. The manuscript is written by THMN, ACF and HLF. All authors read and improved the manuscript.

Acknowledgements

We gratefully acknowledge T.D. Fox for the gift of plasmid, pSS1 and L. P. Musselman for assistance with RT-qPCR. We are also grateful to Binghamton University undergraduate students K. Dave, B. DeJesus, J. DeJesus, A. Federico, G. Lal, S. Sondhi, S. Oh, B.A. Wong, and A. Ziesel for technical assistance with *petite* frequencies in RC1.

Funding

This research was supported by an NIH award (GM101320) to HLF, ACF and KC. MS was supported by an NSF REU (EEC 1757846). ATB, BB, MT, and BG received Binghamton University Undergraduate Research awards.

REFERENCES

1. Wolff JN, Ladoukakis ED, Enriquez JA, Dowling DK. Mitonuclear interactions: evolutionary consequences over multiple biological scales. Philos Trans R Soc Lond B Biol Sci. 2014;369(1646). Epub 2014/05/28. doi: rstb.2013.0443 [pii] 10.1098/rstb.2013.0443. PubMed PMID: 24864313; PubMed Central PMCID: PMC4032519.
2. Quiros PM, Mottis A, Auwerx J. Mitonuclear communication in homeostasis and stress. Nat Rev Mol Cell Biol. 2016;17(4):213-26. Epub 20160309. doi: 10.1038/nrm.2016.23. PubMed PMID: 26956194.
3. Pozzi A, Dowling DK. New Insights into Mitochondrial-Nuclear Interactions Revealed through Analysis of Small RNAs. Genome Biol Evol. 2022;14(2). doi: 10.1093/gbe/evac023. PubMed PMID: 35143645; PubMed Central PMCID: PMCPMC8883506.
4. Baris TZ, Wagner DN, Dayan DI, Du X, Blier PU, Pichaud N, et al. Evolved genetic and phenotypic differences due to mitochondrial-nuclear interactions. PLoS Genet. 2017;13(3):e1006517. doi: 10.1371/journal.pgen.1006517. PubMed PMID: 28362806; PubMed Central PMCID: PMCPMC5375140.
5. Ding Y, Chen W, Li Q, Rossiter SJ, Mao X. Mitonuclear mismatch alters nuclear gene expression in naturally introgressed Rhinolophus bats. Front Zool. 2021;18(1):42. Epub 20210906. doi: 10.1186/s12983-021-00424-x. PubMed PMID: 34488775; PubMed Central PMCID: PMCPMC8419968.
6. Ma H, Marti Gutierrez N, Morey R, Van Dyken C, Kang E, Hayama T, et al. Incompatibility between Nuclear and Mitochondrial Genomes Contributes to an Interspecies Reproductive Barrier. Cell Metab. 2016;24(2):283-94. Epub 20160714. doi: 10.1016/j.cmet.2016.06.012. PubMed PMID: 27425585; PubMed Central PMCID: PMCPMC4981548.
7. Mossman JA, Ge JY, Navarro F, Rand DM. Mitochondrial DNA Fitness Depends on Nuclear Genetic Background in Drosophila. G3 (Bethesda). 2019;9(4):1175-88. Epub 2019/02/13. doi: 10.1534/g3.119.400067. PubMed PMID: 30745378; PubMed Central PMCID: PMCPMC6469417.
8. Immonen E, Berger D, Sayadi A, Liljestrand-Ronn J, Arnqvist G. An experimental test of temperature-dependent selection on mitochondrial haplotypes in Callosobruchus maculatus seed beetles. Ecol Evol. 2020;10(20):11387-98. Epub 20200917. doi: 10.1002/ece3.6775. PubMed PMID: 33144972; PubMed Central PMCID: PMCPMC7593184.
9. Bever BW, Dietz ZP, Sullins JA, Montoya AM, Bergthorsson U, Katju V, et al. Mitonuclear Mismatch is Associated With Increased Male Frequency, Outcrossing, and Male Sperm Size in Experimentally-Evolved C. elegans. Front Genet. 2022;13:742272. Epub 20220311. doi: 10.3389/fgene.2022.742272. PubMed PMID: 35360860; PubMed Central PMCID: PMCPMC8961728.

10. Lima TG, Burton RS, Willett CS. Genomic scans reveal multiple mito-nuclear incompatibilities in population crosses of the copepod Tigriopus californicus. Evolution. 2019;73(3):609-20. Epub 20190208. doi: 10.1111/evo.13690. PubMed PMID: 30675905.
11. Carnegie L, Reuter M, Fowler K, Lane N, Camus MF. Mother's curse is pervasive across a large mitonuclear Drosophila panel. Evol Lett. 2021;5(3):230-9. Epub 20210313. doi: 10.1002/evl3.221. PubMed PMID: 34136271; PubMed Central PMCID: PMCPMC8190446.
12. Plytycz B, Bigaj J, Rysiewska A, Osikowski A, Hofman S, Podolak A, et al. Impairment of reproductive capabilities in three subsequent generations of asymmetric hybrids between Eisenia andrei and E. fetida from French, Hungarian and Polish laboratory colonies. PLoS One. 2020;15(7):e0235789. Epub 2020/07/10. doi: 10.1371/journal.pone.0235789. PubMed PMID: 32645117; PubMed Central PMCID: PMCPMC7347191.
13. Chase CD. Cytoplasmic male sterility: a window to the world of plant mitochondrial-nuclear interactions. Trends Genet. 2007;23(2):81-90. Epub 20061222. doi: 10.1016/j.tig.2006.12.004. PubMed PMID: 17188396.
14. Paliwal S, Fiumera AC, Fiumera HL. Mitochondrial-nuclear epistasis contributes to phenotypic variation and coadaptation in natural isolates of Saccharomyces cerevisiae. Genetics. 2014;198(3):1251-65. Epub 2014/08/29. doi: 10.1534/genetics.114.168575. PubMed PMID: 25164882; PubMed Central PMCID: PMC4224164.
15. Nguyen THM, Sondhi S, Ziesel A, Paliwal S, Fiumera HL. Mitochondrial-nuclear coadaptation revealed through mtDNA replacements in Saccharomyces cerevisiae. BMC Evol Biol. 2020;20(1):128. Epub 2020/09/27. doi: 10.1186/s12862-020-01685-6. PubMed PMID: 32977769; PubMed Central PMCID: PMCPMC7517635.
16. Giordano L, Sillo F, Garbelotto M, Gonthier P. Mitonuclear interactions may contribute to fitness of fungal hybrids. Sci Rep. 2018;8(1):1706. Epub 2018/01/28. doi: 10.1038/s41598-018-19922-w. PubMed PMID: 29374209; PubMed Central PMCID: PMCPMC5786003.
17. Lee HY, Chou JY, Cheong L, Chang NH, Yang SY, Leu JY. Incompatibility of nuclear and mitochondrial genomes causes hybrid sterility between two yeast species. Cell. 2008;135(6):1065-73. Epub 2008/12/17. doi: S0092-8674(08)01385-8 [pii] 10.1016/j.cell.2008.10.047. PubMed PMID: 19070577.
18. Zeyl C, Andreson B, Weninck E. Nuclear-mitochondrial epistasis for fitness in Saccharomyces cerevisiae. Evolution. 2005;59(4):910-4. Epub 2005/06/02. PubMed PMID: 15926700.
19. Leducq JB, Henault M, Charron G, Nielly-Thibault L, Terrat Y, Fiumera HL, et al. Mitochondrial Recombination and Introgression during Speciation by Hybridization. Mol Biol Evol. 2017;34(8):1947-59. doi: 10.1093/molbev/msx139. PubMed PMID: 28444332.
20. Vijayraghavan S, Kozmin SG, Strope PK, Skelly DA, Lin ZG, Kennell J, et al. Mitochondrial Genome Variation Affects Multiple Respiration and Nonrespiration Phenotypes in Saccharomyces cerevisiae. Genetics. 2019;211(2):773-86. doi: 10.1534/genetics.118.301546. PubMed PMID: WOS:000458574800026.
21. Pickett SJ, Deen D, Pyle A, Santibanez-Koref M, Hudson G. Interactions between nuclear and mitochondrial SNPs and Parkinson's disease risk. Mitochondrion. 2022;63:85-8. Epub 20220212. doi: 10.1016/j.mito.2022.02.002. PubMed PMID: 35167983.
22. Gonzalez S. The Role of Mitonuclear Incompatibility in Bipolar Disorder Susceptibility and Resilience Against Environmental Stressors. Front Genet. 2021;12:636294. Epub 20210316. doi: 10.3389/fgene.2021.636294. PubMed PMID: 33815470; PubMed Central PMCID: PMCPMC8010675.

23. Son JM, Lee C. Aging: All roads lead to mitochondria. Semin Cell Dev Biol. 2021;116:160-8. Epub 20210323. doi: 10.1016/j.semcdb.2021.02.006. PubMed PMID: 33741252.
24. Kozin M, Kulakova O, Kiselev I, Baulina N, Boyko A, Favorova O. Mitonuclear interactions influence multiple sclerosis risk. Gene. 2020;758:144962. Epub 20200717. doi: 10.1016/j.gene.2020.144962. PubMed PMID: 32687946.
25. Andrews SJ, Fulton-Howard B, Patterson C, McFall GP, Gross A, Michaelis EK, et al. Mitonuclear interactions influence Alzheimer's disease risk. Neurobiol Aging. 2020;87:138 e7-e14. Epub 20190924. doi: 10.1016/j.neurobiolaging.2019.09.007. PubMed PMID: 31784277; PubMed Central PMCID: PMCPMC7205324.
26. Hill GE, Havird JC, Sloan DB, Burton RS, Greening C, Dowling DK. Assessing the fitness consequences of mitonuclear interactions in natural populations. Biol Rev Camb Philos Soc. 2019;94(3):1089-104. Epub 2018/12/28. doi: 10.1111/brv.12493. PubMed PMID: 30588726; PubMed Central PMCID: PMCPMC6613652.
27. Burton RS. The role of mitonuclear incompatibilities in allopatric speciation. Cell Mol Life Sci. 2022;79(2):103. Epub 20220129. doi: 10.1007/s00018-021-04059-3. PubMed PMID: 35091831.
28. Ellison CK, Burton RS. Interpopulation hybrid breakdown maps to the mitochondrial genome. Evolution. 2008;62(3):631-8. Epub 2007/12/18. doi: EVO305 [pii] 10.1111/j.1558-5646.2007.00305.x. PubMed PMID: 18081717.
29. Ellison CK, Burton RS. Disruption of mitochondrial function in interpopulation hybrids of Tigriopus californicus. Evolution. 2006;60(7):1382-91. Epub 2006/08/26. PubMed PMID: 16929655.
30. Montooth KL, Meiklejohn CD, Abt DN, Rand DM. Mitochondrial-nuclear epistasis affects fitness within species but does not contribute to fixed incompatibilities between species of Drosophila. Evolution. 2010;64(12):3364-79. Epub 2010/07/14. doi: EVO1077 [pii] 10.1111/j.1558-5646.2010.01077.x. PubMed PMID: 20624176; PubMed Central PMCID: PMC2997886.
31. Meiklejohn CD, Holmbeck MA, Siddiq MA, Abt DN, Rand DM, Montooth KL. An Incompatibility between a mitochondrial tRNA and its nuclear-encoded tRNA synthetase compromises development and fitness in Drosophila. PLoS Genet. 2013;9(1):e1003238. Epub 2013/02/06. doi: 10.1371/journal.pgen.1003238 PGENETICS-D-12-02557 [pii]. PubMed PMID: 23382693; PubMed Central PMCID: PMC3561102.
32. Jhuang HY, Lee HY, Leu JY. Mitochondrial-nuclear co-evolution leads to hybrid incompatibility through pentatricopeptide repeat proteins. Embo Reports. 2017;18(1):87-101. doi: 10.15252/embr.201643311. PubMed PMID: WOS:000392430400014.
33. Chou JY, Hung YS, Lin KH, Lee HY, Leu JY. Multiple molecular mechanisms cause reproductive isolation between three yeast species. PLoS Biol. 2010;8(7):e1000432. Epub 20100720. doi: 10.1371/journal.pbio.1000432. PubMed PMID: 20652018; PubMed Central PMCID: PMCPMC2907292.
34. Gibson JD, Niehuis O, Peirson BR, Cash EI, Gadau J. Genetic and developmental basis of F2 hybrid breakdown in Nasonia parasitoid wasps. Evolution. 2013;67(7):2124-32. Epub 2013/07/03. doi: 10.1111/evo.12080. PubMed PMID: 23815665.

35. Riles L, Fay JC. Genetic Basis of Variation in Heat and Ethanol Tolerance in Saccharomyces cerevisiae. G3-Genes Genom Genet. 2019;9(1):179-88. doi: 10.1534/g3.118.200566. PubMed PMID: WOS:000455206600017.

36. Healy TM, Burton RS. Strong selective effects of mitochondrial DNA on the nuclear genome. Proc Natl Acad Sci U S A. 2020;117(12):6616-21. Epub 2020/03/12. doi: 10.1073/pnas.1910141117. PubMed PMID: 32156736; PubMed Central PMCID: PMCPMC7104403.

37. Haddad R, Meter B, Ross JA. The Genetic Architecture of Intra-Species Hybrid Mito-Nuclear Epistasis. Front Genet. 2018;9:481. Epub 2018/12/07. doi: 10.3389/fgene.2018.00481. PubMed PMID: 30505316; PubMed Central PMCID: PMCPMC6250786.

38. Naseeb S, Visinoni F, Hu Y, Hinks Roberts AJ, Maslowska A, Walsh T, et al. Restoring fertility in yeast hybrids: Breeding and quantitative genetics of beneficial traits. Proc Natl Acad Sci U S A. 2021;118(38). doi: 10.1073/pnas.2101242118. PubMed PMID: 34518218; PubMed Central PMCID: PMCPMC8463882.

39. Santiago JC, Boylan JM, Lemieux FA, Gruppuso PA, Sanders JA, Rand DM. Mitochondrial genotype alters the impact of rapamycin on the transcriptional response to nutrients in Drosophila. BMC Genomics. 2021;22(1):213. Epub 2021/03/26. doi: 10.1186/s12864-021-07516-2. PubMed PMID: 33761878; PubMed Central PMCID: PMCPMC7992956.

40. Mossman JA, Biancani LM, Rand DM. Mitochondrial genomic variation drives differential nuclear gene expression in discrete regions of Drosophila gene and protein interaction networks. BMC Genomics. 2019;20(1):691. Epub 2019/09/04. doi: 10.1186/s12864-019-6061-y. PubMed PMID: 31477008; PubMed Central PMCID: PMCPMC6719383.

41. Rand DM, Mossman JA, Spierer AN, Santiago JA. Mitochondria as environments for the nuclear genome in Drosophila: mitonuclear GxGxE. J Hered. 2022;113(1):37-47. doi: 10.1093/jhered/esab066. PubMed PMID: 34964900; PubMed Central PMCID: PMCPMC8851671.

42. Pereira RJ, Lima TG, Pierce-Ward NT, Chao L, Burton RS. Recovery from hybrid breakdown reveals a complex genetic architecture of mitonuclear incompatibilities. Mol Ecol. 2021. Epub 2021/05/19. doi: 10.1111/mec.15985. PubMed PMID: 34003535.

43. Wernick RI, Christy SF, Howe DK, Sullins JA, Ramirez JF, Sare M, et al. Sex and Mitonuclear Adaptation in Experimental Caenorhabditis elegans Populations. Genetics. 2019;211(3):1045-58. Epub 2019/01/24. doi: 10.1534/genetics.119.301935. PubMed PMID: 30670540; PubMed Central PMCID: PMCPMC6404257.

44. Rand DM, Mossman JA. Mitonuclear conflict and cooperation govern the integration of genotypes, phenotypes and environments. Philos Trans R Soc Lond B Biol Sci. 2020;375(1790):20190188. Epub 2019/12/04. doi: 10.1098/rstb.2019.0188. PubMed PMID: 31787039; PubMed Central PMCID: PMCPMC6939372.

45. Campbell RF, McGrath PT, Paaby AB. Analysis of Epistasis in Natural Traits Using Model Organisms. Trends Genet. 2018;34(11):883-98. Epub 20180827. doi: 10.1016/j.tig.2018.08.002. PubMed PMID: 30166071; PubMed Central PMCID: PMCPMC6541385.

46. Mackay TF. Epistasis and quantitative traits: using model organisms to study gene-gene interactions. Nat Rev Genet. 2014;15(1):22-33. doi: 10.1038/nrg3627. PubMed PMID: 24296533; PubMed Central PMCID: PMCPMC3918431.

47. Peter J, De Chiara M, Friedrich A, Yue JX, Pflieger D, Bergstrom A, et al. Genome evolution across 1,011 Saccharomyces cerevisiae isolates. Nature. 2018;556(7701):339-44. doi: 10.1038/s41586-018-0030-5. PubMed PMID: 29643504.
48. Gilea AI, Ceccatelli Berti C, Magistrati M, Di Punzio G, Goffrini P, Baruffini E, et al. Saccharomyces cerevisiae as a Tool for Studying Mutations in Nuclear Genes Involved in Diseases Caused by Mitochondrial DNA Instability. Genes. 2021;12(12):1866. doi: 10.3390/genes12121866.
49. Stenger M, Le DT, Klecker T, Westermann B. Systematic analysis of nuclear gene function in respiratory growth and expression of the mitochondrial genome in S. cerevisiae. Microb Cell. 2020;7(9):234-49. Epub 2020/09/10. doi: 10.15698/mic2020.09.729. PubMed PMID: 32904421; PubMed Central PMCID: PMCPMC7453639.
50. Contamine V, Picard M. Maintenance and integrity of the mitochondrial genome: a plethora of nuclear genes in the budding yeast. Microbiol Mol Biol Rev. 2000;64(2):281-315. PubMed PMID: 10839818.
51. Dimitrov LN, Brem RB, Kruglyak L, Gottschling DE. Polymorphisms in multiple genes contribute to the spontaneous mitochondrial genome instability of Saccharomyces cerevisiae S288C strains. Genetics. 2009;183(1):365-83. Epub 2009/07/08. doi: genetics.109.104497 [pii] 10.1534/genetics.109.104497. PubMed PMID: 19581448; PubMed Central PMCID: PMC2746160.
52. Stenberg S, Li J, Gjuvsland AB, Persson K, Demitz-Helin E, Gonzalez Pena C, et al. Genetically controlled mtDNA deletions prevent ROS damage by arresting oxidative phosphorylation. Elife. 2022;11. Epub 20220708. doi: 10.7554/eLife.76095. PubMed PMID: 35801695.
53. Wallace DC, Chalkia D. Mitochondrial DNA genetics and the heteroplasmy conundrum in evolution and disease. Cold Spring Harb Perspect Biol. 2013;5(11):a021220. Epub 2013/11/05. doi: 5/11/a021220 [pii] 10.1101/cshperspect.a021220. PubMed PMID: 24186072.
54. Payne BA, Wilson IJ, Yu-Wai-Man P, Coxhead J, Deehan D, Horvath R, et al. Universal heteroplasmy of human mitochondrial DNA. Hum Mol Genet. 2013;22(2):384-90. Epub 20121016. doi: 10.1093/hmg/dds435. PubMed PMID: 23077218; PubMed Central PMCID: PMCPMC3526165.
55. Yusoff A, Radzak S, Mohd Khair S. Research highlights on contributions of mitochondrial DNA microsatellite instability in solid cancers – an overview. Współczesna Onkologia. 2022;26(1):8-26. doi: 10.5114/wo.2022.115674.
56. El-Hattab AW, Scaglia F. Mitochondrial DNA depletion syndromes: review and updates of genetic basis, manifestations, and therapeutic options. Neurotherapeutics. 2013;10(2):186-98. doi: 10.1007/s13311-013-0177-6. PubMed PMID: 23385875; PubMed Central PMCID: PMCPMC3625391.
57. Nogueira C, Almeida LS, Nesti C, Pezzini I, Videira A, Vilarinho L, et al. Syndromes associated with mitochondrial DNA depletion. Ital J Pediatr. 2014;40:34. Epub 20140403. doi: 10.1186/1824-7288-40-34. PubMed PMID: 24708634; PubMed Central PMCID: PMCPMC3985578.
58. Dujon B. Mitochondrial genetics revisited. Yeast. 2020;37(2):191-205. Epub 2019/11/07. doi: 10.1002/yea.3445. PubMed PMID: 31691343.
59. Faugeron-Fonty G, Culard F, Baldacci G, Goursot R, Prunell A, Bernardi G. The mitochondrial genome of wild-type yeast cells. VIII. The spontaneous cytoplasmic "petite"

1032 mutation. J Mol Biol. 1979;134(3):493-57. Epub 1979/11/05. doi: 0022-2836(79)90365-6 [pii].
1033 PubMed PMID: 231670.
1034 60. Liti G, Carter DM, Moses AM, Warringer J, Parts L, James SA, et al. Population
1035 genomics of domestic and wild yeasts. Nature. 2009;458(7236):337-41. Epub 2009/02/13. doi:
1036 nature07743 [pii]
1037 10.1038/nature07743. PubMed PMID: 19212322; PubMed Central PMCID: PMC2659681.
1038 61. Strope PK, Skelly DA, Kozmin SG, Mahadevan G, Stone EA, Magwene PM, et al. The
1039 100-genomes strains, an S. cerevisiae resource that illuminates its natural phenotypic and
1040 genotypic variation and emergence as an opportunistic pathogen. Genome Res. 2015;25(5):762-
1041 74. doi: 10.1101/gr.185538.114. PubMed PMID: 25840857; PubMed Central PMCID:
1042 PMC4417123.
1043 62. de Zamaroczy M, Bernardi G. The GC clusters of the mitochondrial genome of yeast and
1044 their evolutionary origin. Gene. 1986;41(1):1-22. Epub 1986/01/01. doi: 0378-1119(86)90262-3
1045 [pii]. PubMed PMID: 3009268.
1046 63. Wu B, Hao W. A Dynamic Mobile DNA Family in the Yeast Mitochondrial Genome. G3
1047 (Bethesda). 2015;5(6):1273-82. doi: 10.1534/g3.115.017822. PubMed PMID: 25897011;
1048 PubMed Central PMCID: PMCPMC4478555.
1049 64. Weiller GF, Bruckner H, Kim SH, Pratje E, Schweyen RJ. A GC cluster repeat is a
1050 hotspot for mit− macro-deletions in yeast mitochondrial DNA. Molecular and General Genetics
1051 MGG. 1991;226(1):233-40.
1052 65. Dieckmann CL, Gandy B. Preferential Recombination between Gc Clusters in Yeast
1053 Mitochondrial-DNA. Embo J. 1987;6(13):4197-203. PubMed PMID: WOS:A1987L435200041.
1054 66. Weiller G, Schueller CM, Schweyen RJ. Putative target sites for mobile G + C rich
1055 clusters in yeast mitochondrial DNA: single elements and tandem arrays. Mol Gen Genet.
1056 1989;218(2):272-83. PubMed PMID: 2674655.
1057 67. Wolters JF, Chiu K, Fiumera HL. Population structure of mitochondrial genomes in
1058 Saccharomyces cerevisiae. BMC Genomics. 2015;16:451. doi: 10.1186/s12864-015-1664-4.
1059 PubMed PMID: 26062918; PubMed Central PMCID: PMC4464245.
1060 68. Wolters JF, Charron G, Gaspary A, Landry CR, Fiumera AC, Fiumera HL. Mitochondrial
1061 Recombination Reveals Mito-Mito Epistasis in Yeast. Genetics. 2018;209(1):307-19. doi:
1062 10.1534/genetics.117.300660. PubMed PMID: 29531011; PubMed Central PMCID:
1063 PMCPMC5937185.
1064 69. Warringer J, Zorgo E, Cubillos FA, Zia A, Gjuvsland A, Simpson JT, et al. Trait
1065 variation in yeast is defined by population history. PLoS Genet. 2011;7(6):e1002111. Epub
1066 2011/06/24. doi: 10.1371/journal.pgen.1002111. PubMed PMID: 21698134; PubMed Central
1067 PMCID: PMCPMC3116910.
1068 70. Strope PK, Skelly DA, Kozmin SG, Mahadevan G, Stone EA, Magwene PM, et al. The
1069 100-genomes strains, anS. cerevisiaeresource that illuminates its natural phenotypic and
1070 genotypic variation and emergence as an opportunistic pathogen. Genome Res. 2015;25(5):762-
1071 74. doi: 10.1101/gr.185538.114. PubMed PMID: WOS:000353712200014.
1072 71. Luban C, Beutel M, Stahl U, Schmidt U. Systematic screening of nuclear encoded
1073 proteins involved in the splicing metabolism of group II introns in yeast mitochondria. Gene.
1074 2005;354:72-9.
1075 72. Dimmer KS, Fritz S, Fuchs F, Messerschmitt M, Weinbach N, Neupert W, et al. Genetic
1076 basis of mitochondrial function and morphology in Saccharomyces cerevisiae. Mol Biol Cell.

2002;13(3):847-53. Epub 2002/03/22. doi: 10.1091/mbc.01-12-0588. PubMed PMID: 11907266; PubMed Central PMCID: PMCPMC99603.

73. Merz S, Westermann B. Genome-wide deletion mutant analysis reveals genes required for respiratory growth, mitochondrial genome maintenance and mitochondrial protein synthesis in Saccharomyces cerevisiae. Genome Biol. 2009;10(9):R95. Epub 2009/09/16. doi: gb-2009-10-9-r95 [pii]
10.1186/gb-2009-10-9-r95. PubMed PMID: 19751518.

74. Zhang Z, Dai J, Veliath E, Jones RA, Yang D. Structure of a two-G-tetrad intramolecular G-quadruplex formed by a variant human telomeric sequence in K+ solution: insights into the interconversion of human telomeric G-quadruplex structures. Nucleic Acids Res. 2010;38(3):1009-21. Epub 2009/12/01. doi: 10.1093/nar/gkp1029. PubMed PMID: 19946019; PubMed Central PMCID: PMCPMC2817458.

75. Bateman JM, Perlman PS, Butow RA. Mutational bisection of the mitochondrial DNA stability and amino acid biosynthetic functions of ilv5p of budding yeast. Genetics. 2002;161(3):1043-52. doi: 10.1093/genetics/161.3.1043. PubMed PMID: 12136009; PubMed Central PMCID: PMCPMC1462179.

76. Herbert CJ, Labouesse M, Dujardin G, Slonimski PP. The NAM2 proteins from S. cerevisiae and S. douglasii are mitochondrial leucyl-tRNA synthetases, and are involved in mRNA splicing. Embo J. 1988;7(2):473-83. Epub 1988/02/01. PubMed PMID: 3284745; PubMed Central PMCID: PMCPMC454344.

77. Schilke B, Forster J, Davis J, James P, Walter W, Laloraya S, et al. The cold sensitivity of a mutant of Saccharomyces cerevisiae lacking a mitochondrial heat shock protein 70 is suppressed by loss of mitochondrial DNA. J Cell Biol. 1996;134(3):603-13. Epub 1996/08/01. doi: 10.1083/jcb.134.3.603. PubMed PMID: 8707841; PubMed Central PMCID: PMCPMC2120932.

78. Schmalix W, Bandlow W. The ethanol-inducible YAT1 gene from yeast encodes a presumptive mitochondrial outer carnitine acetyltransferase. J Biol Chem. 1993;268(36):27428-39. PubMed PMID: 8262985.

79. van Rossum HM, Kozak BU, Niemeijer MS, Dykstra JC, Luttik MA, Daran JM, et al. Requirements for Carnitine Shuttle-Mediated Translocation of Mitochondrial Acetyl Moieties to the Yeast Cytosol. mBio. 2016;7(3). Epub 2016/05/05. doi: 10.1128/mBio.00520-16. PubMed PMID: 27143389; PubMed Central PMCID: PMCPMC4959659.

80. Liu ZL, Huang X. A glimpse of potential transposable element impact on adaptation of the industrial yeast Saccharomyces cerevisiae. FEMS Yeast Res. 2020;20(6). doi: 10.1093/femsyr/foaa043. PubMed PMID: 32780789.

81. Reinders J, Wagner K, Zahedi RP, Stojanovski D, Eyrich B, van der Laan M, et al. Profiling phosphoproteins of yeast mitochondria reveals a role of phosphorylation in assembly of the ATP synthase. Mol Cell Proteomics. 2007;6(11):1896-906. Epub 2007/09/01. doi: 10.1074/mcp.M700098-MCP200. PubMed PMID: 17761666.

82. Tkach JM, Yimit A, Lee AY, Riffle M, Costanzo M, Jaschob D, et al. Dissecting DNA damage response pathways by analysing protein localization and abundance changes during DNA replication stress. Nat Cell Biol. 2012;14(9):966-76. Epub 20120729. doi: 10.1038/ncb2549. PubMed PMID: 22842922; PubMed Central PMCID: PMCPMC3434236.

83. Mönkemeyer L, Klaips CL, Balchin D, Körner R, Hartl FU, Bracher A. Chaperone Function of Hgh1 in the Biogenesis of Eukaryotic Elongation Factor 2. Mol Cell. 2019;74(1):88-100.e9. Epub 2019/03/17. doi: 10.1016/j.molcel.2019.01.034. PubMed PMID: 30876804.

84. Martin-Yken H, Dagkessamanskaia A, Talibi D, Francois J. KNR4 is a member of the PKC1 signalling pathway and genetically interacts with BCK2, a gene involved in cell cycle progression in Saccharomyces cerevisiae. Curr Genet. 2002;41(5):323-32. Epub 2002/08/20. doi: 10.1007/s00294-002-0299-6. PubMed PMID: 12185498.
85. Brown JA, Sherlock G, Myers CL, Burrows NM, Deng C, Wu HI, et al. Global analysis of gene function in yeast by quantitative phenotypic profiling. Mol Syst Biol. 2006;2:2006 0001. Epub 20060117. doi: 10.1038/msb4100043. PubMed PMID: 16738548; PubMed Central PMCID: PMCPMC1681475.
86. Visintin R, Stegmeier F, Amon A. The role of the polo kinase Cdc5 in controlling Cdc14 localization. Mol Biol Cell. 2003;14(11):4486-98. Epub 2003/10/11. doi: 10.1091/mbc.e03-02-0095. PubMed PMID: 14551257; PubMed Central PMCID: PMCPMC266767.
87. Morgenstern M, Stiller SB, Lubbert P, Peikert CD, Dannenmaier S, Drepper F, et al. Definition of a High-Confidence Mitochondrial Proteome at Quantitative Scale. Cell Rep. 2017;19(13):2836-52. Epub 2017/06/29. doi: 10.1016/j.celrep.2017.06.014. PubMed PMID: 28658629; PubMed Central PMCID: PMCPMC5494306.
88. Li J, Rinnerthaler M, Hartl J, Weber M, Karl T, Breitenbach-Koller H, et al. Slow Growth and Increased Spontaneous Mutation Frequency in Respiratory Deficient afo1-Yeast Suppressed by a Dominant Mutation in ATP3. G3 Genes|Genomes|Genetics. 2020;10(12):4637-48. doi: 10.1534/g3.120.401537.
89. Lee H-C, Wei Y-H. Mitochondrial DNA Instability and Metabolic Shift in Human Cancers. International journal of molecular sciences. 2009;10(2):674-701. doi: 10.3390/ijms10020674.
90. Stuart GR, Santos JH, Strand MK, Van Houten B, Copeland WC. Mitochondrial and nuclear DNA defects in Saccharomyces cerevisiae with mutations in DNA polymerase gamma associated with progressive external ophthalmoplegia. Hum Mol Genet. 2006;15(2):363-74. Epub 20051220. doi: 10.1093/hmg/ddi454. PubMed PMID: 16368709.
91. Baruffini E, Horvath R, Dallabona C, Czermin B, Lamantea E, Bindoff L, et al. Predicting the contribution of novel POLG mutations to human disease through analysis in yeast model. Mitochondrion. 2011;11(1):182-90. Epub 20100929. doi: 10.1016/j.mito.2010.09.007. PubMed PMID: 20883824.
92. Lee Y, Kim T, Lee M, So S, Karagozlu MZ, Seo GH, et al. De Novo Development of mtDNA Deletion Due to Decreased POLG and SSBP1 Expression in Humans. Genes (Basel). 2021;12(2). Epub 20210217. doi: 10.3390/genes12020284. PubMed PMID: 33671400; PubMed Central PMCID: PMCPMC7922481.
93. Vermulst M, Wanagat J, Kujoth GC, Bielas JH, Rabinovitch PS, Prolla TA, et al. DNA deletions and clonal mutations drive premature aging in mitochondrial mutator mice. Nat Genet. 2008;40(4):392-4. Epub 20080302. doi: 10.1038/ng.95. PubMed PMID: 18311139.
94. Chen H, Xue J, Churikov D, Hass EP, Shi S, Lemon LD, et al. Structural Insights into Yeast Telomerase Recruitment to Telomeres. Cell. 2018;172(1-2):331-43 e13. Epub 20171228. doi: 10.1016/j.cell.2017.12.008. PubMed PMID: 29290466; PubMed Central PMCID: PMCPMC5839504.
95. Lubin JW, Tucey TM, Lundblad V. Using Separation-of-Function Mutagenesis To Define the Full Spectrum of Activities Performed by the Est1 Telomerase Subunit in Vivo. Genetics. 2018;208(1):97-110. Epub 20171129. doi: 10.1534/genetics.117.300145. PubMed PMID: 29187507; PubMed Central PMCID: PMCPMC5753878.

96. Zhang ML, Tong XJ, Fu XH, Zhou BO, Wang J, Liao XH, et al. Yeast telomerase subunit Est1p has guanine quadruplex-promoting activity that is required for telomere elongation. Nature structural & molecular biology. 2010;17(2):202-9. Epub 2010/01/26. doi: 10.1038/nsmb.1760. PubMed PMID: 20098422.

97. Butler TJ, Estep KN, Sommers JA, Maul RW, Moore AZ, Bandinelli S, et al. Mitochondrial genetic variation is enriched in G-quadruplex regions that stall DNA synthesis in vitro. Hum Mol Genet. 2020. Epub 2020/03/20. doi: 10.1093/hmg/ddaa043. PubMed PMID: 32191790.

98. Lin KW, McDonald KR, Guise AJ, Chan A, Cristea IM, Zakian VA. Proteomics of yeast telomerase identified Cdc48-Npl4-Ufd1 and Ufd4 as regulators of Est1 and telomere length. Nat Commun. 2015;6:8290. Epub 2015/09/15. doi: 10.1038/ncomms9290. PubMed PMID: 26365526; PubMed Central PMCID: PMCPMC4579843.

99. Macierzanka M, Plotka M, Pryputniewicz-Drobinska D, Lewandowska A, Lightowlers R, Marszalek J. Maintenance and stabilization of mtDNA can be facilitated by the DNA-binding activity of Ilv5p. Biochim Biophys Acta. 2008;1783(1):107-17. Epub 20071030. doi: 10.1016/j.bbamcr.2007.09.009. PubMed PMID: 18023287.

100. Špírek M, Šoltésová A, Horváth A, Sláviková E, Sulo P. GC clusters and the stability of mitochondrial genomes ofSaccharomyces cerevisiae and related yeasts. Folia Microbiologica. 2002;47(3):263-70. doi: 10.1007/bf02817649.

101. De Chiara M, Friedrich A, Barre B, Breitenbach M, Schacherer J, Liti G. Discordant evolution of mitochondrial and nuclear yeast genomes at population level. BMC Biol. 2020;18(1):49. Epub 2020/05/13. doi: 10.1186/s12915-020-00786-4. PubMed PMID: 32393264; PubMed Central PMCID: PMCPMC7216626.

102. Cubillos FA, Louis EJ, Liti G. Generation of a large set of genetically tractable haploid and diploid Saccharomyces strains. FEMS Yeast Res. 2009;9(8):1217-25. Epub 2009/10/21. doi: FYR583 [pii]
10.1111/j.1567-1364.2009.00583.x. PubMed PMID: 19840116.

103. Winzeler EA, Shoemaker DD, Astromoff A, Liang H, Anderson K, Andre B, et al. Functional characterization of the S. cerevisiae genome by gene deletion and parallel analysis. Science. 1999;285(5429):901-6. Epub 1999/08/07. doi: 7737 [pii]. PubMed PMID: 10436161.

104. Gietz RD, Schiestl RH. High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. Nat Protoc. 2007;2(1):31-4. Epub 2007/04/03. doi: nprot.2007.13 [pii]
10.1038/nprot.2007.13. PubMed PMID: 17401334.

105. Becker DM, Guarente L. High-efficiency transformation of yeast by electroporation. Methods Enzymol. 1991;194:182-7. Epub 1991/01/01. doi: 10.1016/0076-6879(91)94015-5. PubMed PMID: 2005786.

106. Duveau F, Metzger BP, Gruber JD, Mack K, Sood N, Brooks TE, et al. Mapping small effect mutations in Saccharomyces cerevisiae: impacts of experimental design and mutational properties. G3 (Bethesda). 2014;4(7):1205-16. doi: 10.1534/g3.114.011783. PubMed PMID: 24789747.

107. Team RC. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2020.

108. Team R. RStudio: Integrated Development Environment for R. RStudio, PBC, Boston, MA. 2021.

109. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007;81(3):559-75. Epub 2007/08/19. doi: S0002-9297(07)61352-4 [pii] 10.1086/519795. PubMed PMID: 17701901; PubMed Central PMCID: PMC1950838.

110. Storey JD, Bass AJ, Dabney A, Robinson D. qvalue: Q-value estimation for false discovery rate control. http://github.com/jdstorey/qvalue2022.

111. Hoffman CS, Winston F. A ten-minute DNA preparation from yeast efficiently releases autonomous plasmids for transformation of Escherichia coli. Gene. 1987;57(2-3):267-72. PubMed PMID: 3319781.

112. Fiumera AC, Dumont BL, Clark AG. Sperm competitive ability in Drosophila melanogaster associated with variation in male reproductive proteins. Genetics. 2005;169(1):243-57. Epub 2004/10/07. doi: genetics.104.032870 [pii] 10.1534/genetics.104.032870. PubMed PMID: 15466425; PubMed Central PMCID: PMC1448872.

Supplemental Tables

Table S1. Strain table

Table S2. ANOVA *Petite* frequencies in wild isolates vary by strain and population

Table S3. ANOVA mtDNAs influence *petite* frequencies in the Y55 nuclear background

Table S4. mtDNA GC clusters and introns

Table S5. ANOVA *Petite* frequencies influenced by mitonuclear interactions in a 4x4 mitonuclear genotype panel

Table S6. Recombinant Collection SNP descriptions

Table S7. Associated SNPs that influence *petite* frequencies independent of mitotype (main effect loci)

Table S8. Associated SNPs that influence *petite* frequencies dependent on mitotype (mitonuclear loci)

Table S9. Effect sizes of main effect loci

Table S10. Effect sizes of mitonuclear loci

Table S11. ANOVAs Mitonuclear effects of *HGH1* and *YAT1*

Table S12. ANOVA Mitonuclear effects of *HGH1*, *BSN1*, *SMI1*

Table S13. ANOVAs Expression differences of mitonuclear candidate loci

Table S14. Primer sequences for qRT-PCR

Supplemental Figures

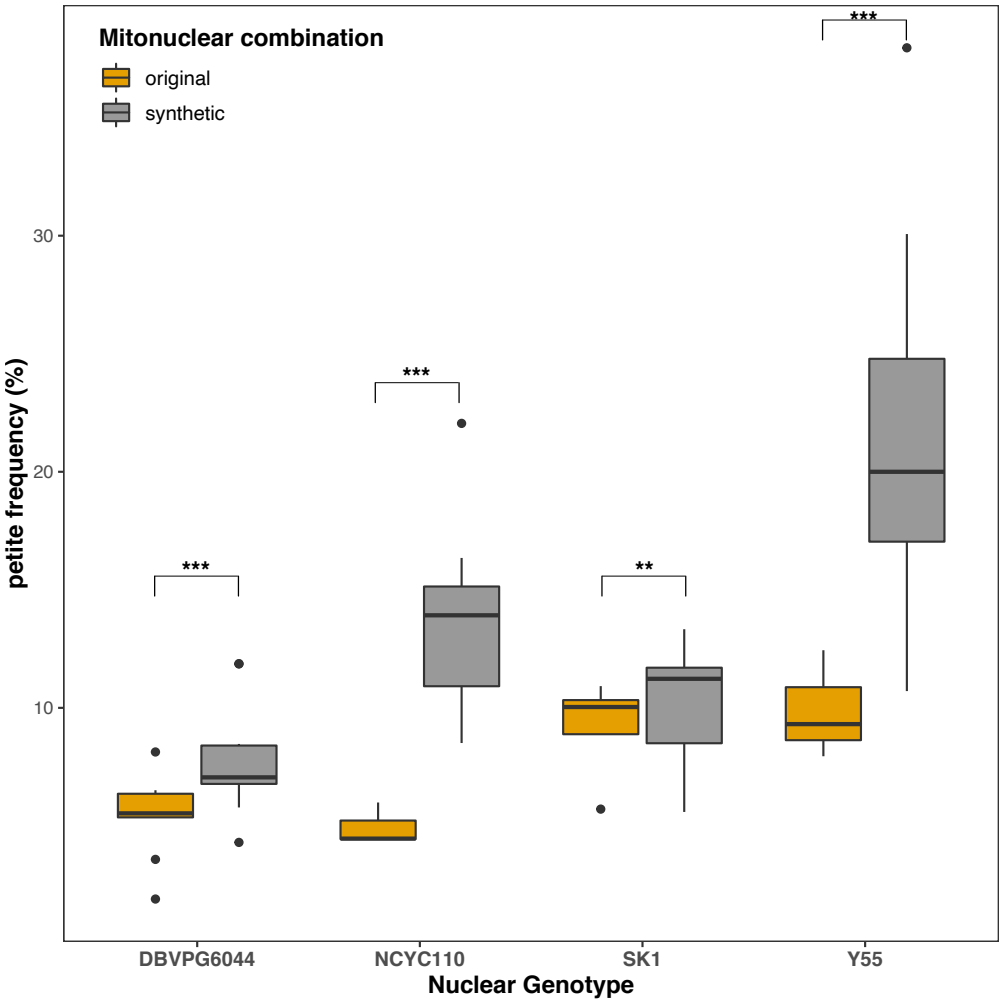

Fig S1. Strains harboring original mtDNAs have lower *petite* frequencies than those with synthetic mitonuclear combinations. *Petite* frequencies of strains with the original (gold) vs. synthetic (grey) mitonuclear genotypes from **Fig 3** are replotted as box plots with the synthetic combinations combined. All nuclear and mtDNAs are from strains with West African lineages. ANOVA significances are shown. * $P<0.05$, ** $P \leq 0.005$, *** $P \leq 0.001$.

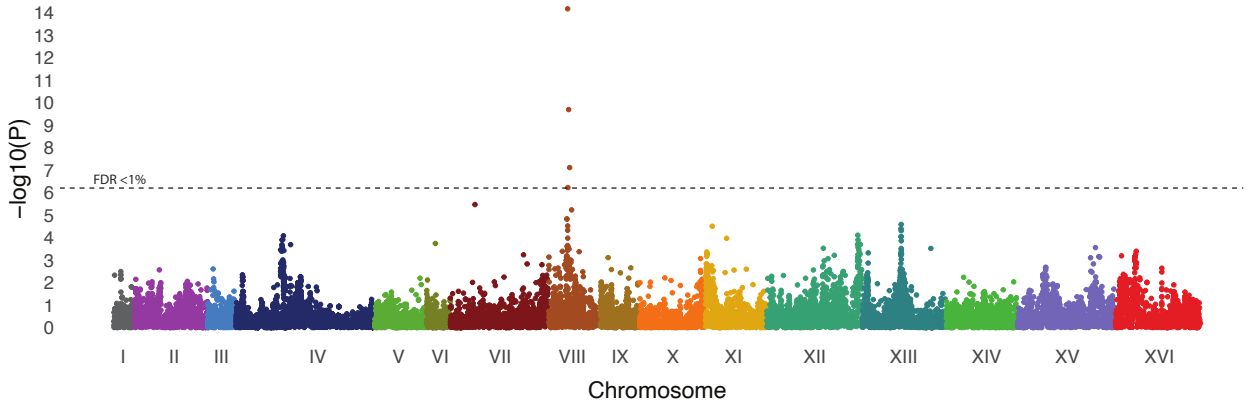

Fig S2. Association Mapping using the Recombinant Collection A Manhattan plot of -log10 of *P* values plotted against chromosomal position shows associations for maximal colony sizes for RC1 strains grown on copper sulfate. A single peak, corresponding to the location of *CUP1* on Chr. 8, is the only significant association at FDR<1%.

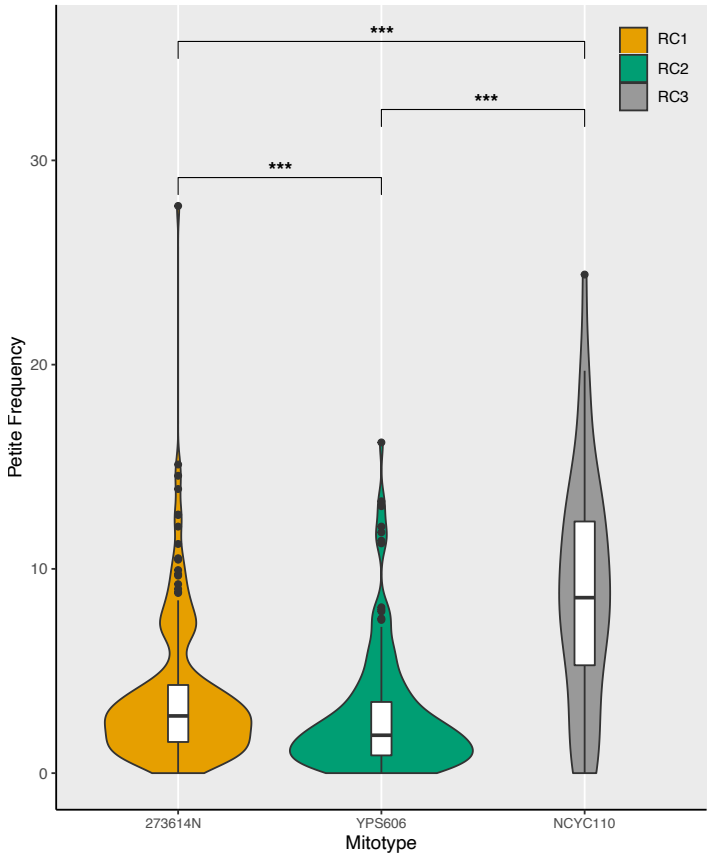

Fig S3. Strains with the GC-rich mtDNA in RC3 strains have higher *petite* frequencies. *Petite* frequencies of strains from RC1, RC2, and RC3 are presented as violin plots. RC1 strains harbor mtDNA from 273614N containing a medium number (137) of GC clusters. RC2 strains harbor mtDNA from YPS606 containing a low number (117) of GC clusters. RC3 harbor mtDNA from NCYC110 containing a high number (210) of GC clusters.

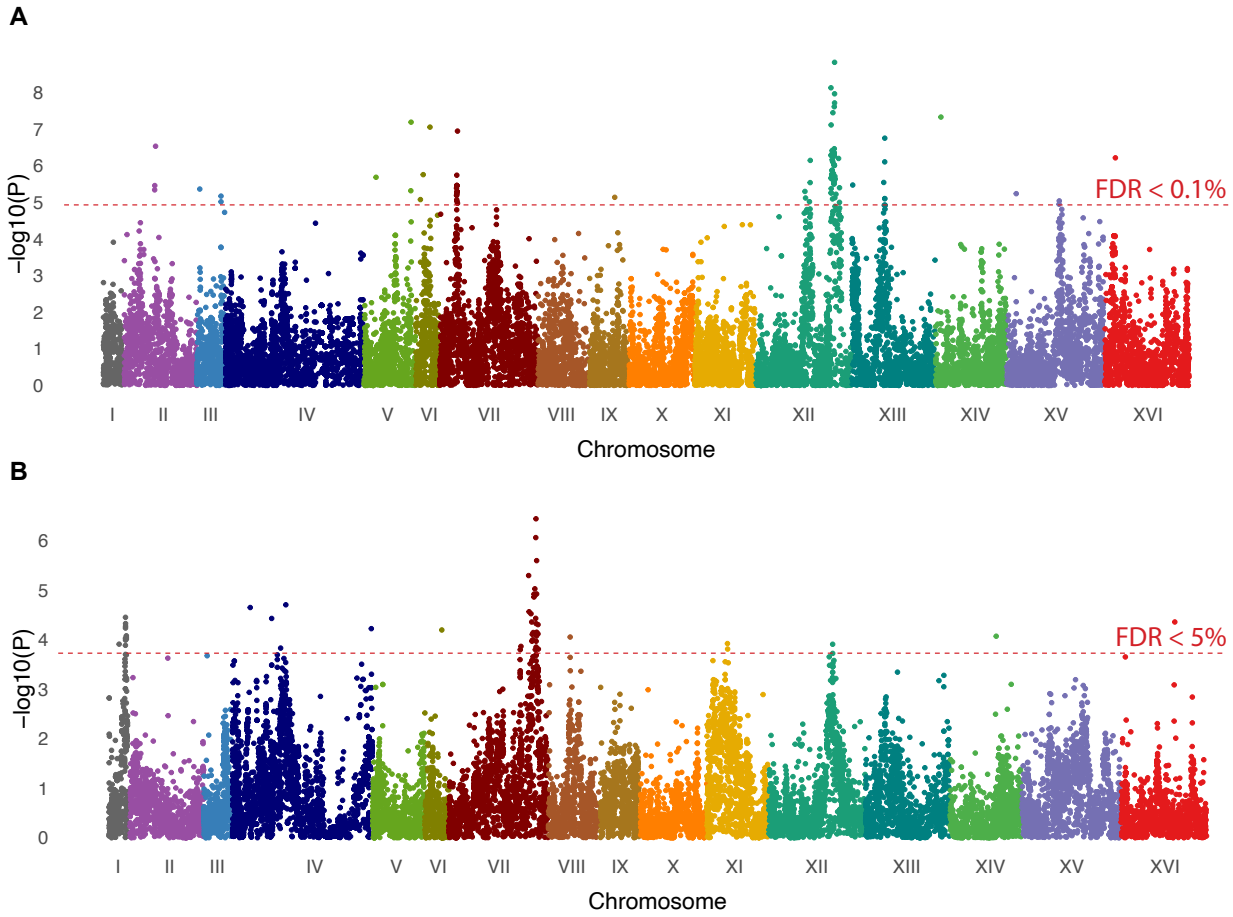

Fig S4. GWAS using MIP1 variants as covariates increases power to detect mitonuclear associations. Manhattan plots show mitotype **A.** independent and **B.** mitotype dependent associations, when *MIP1* variants were included as covariates. Red lines indicate FDR thresholds at 0.1% for nuclear associations and 5.0% for mitonuclear associations.

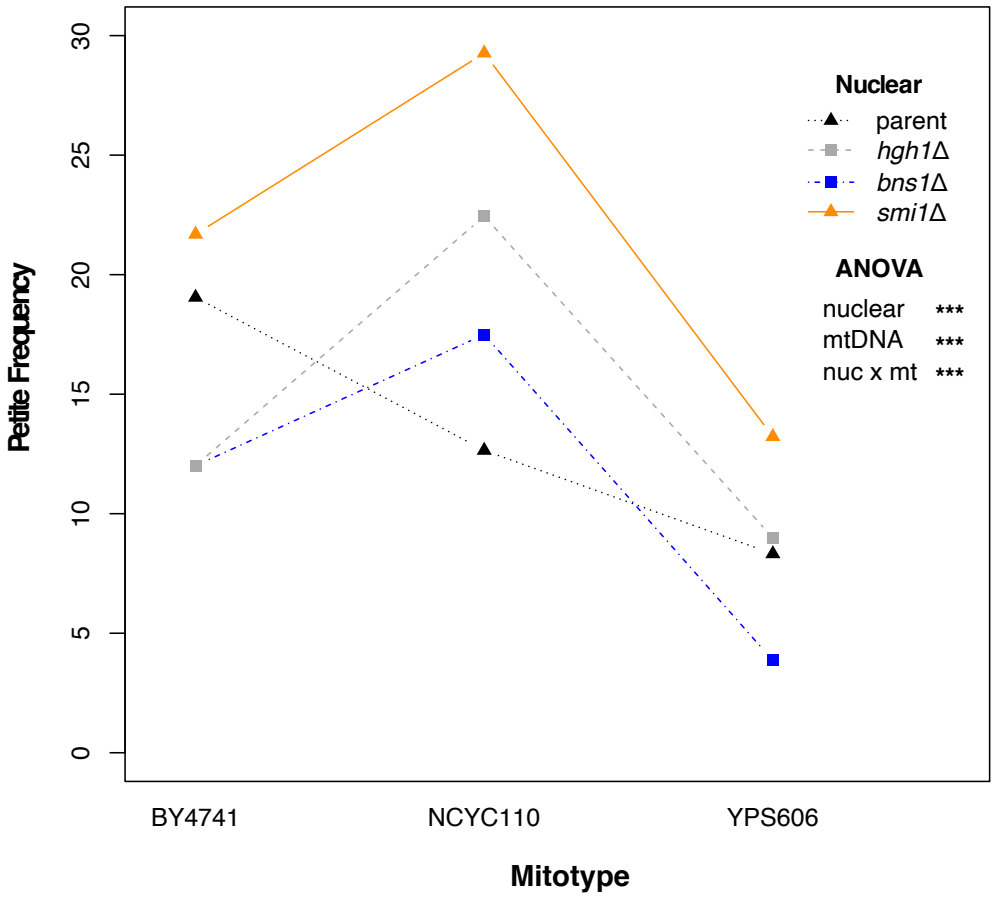

Fig S5. Mitonuclear interactions of BNS1, SMI1 and HGH1 on *petite* frequencies. Interaction plot follows *petite* frequencies for each nuclear genotype paired with different mtDNAs. See **Table S11** for ANOVA. Data was collected using the same assay as performed for phenotyping the RCs (with 8-12 replicates) and cannot be combined with the data shown in **Fig 6**.

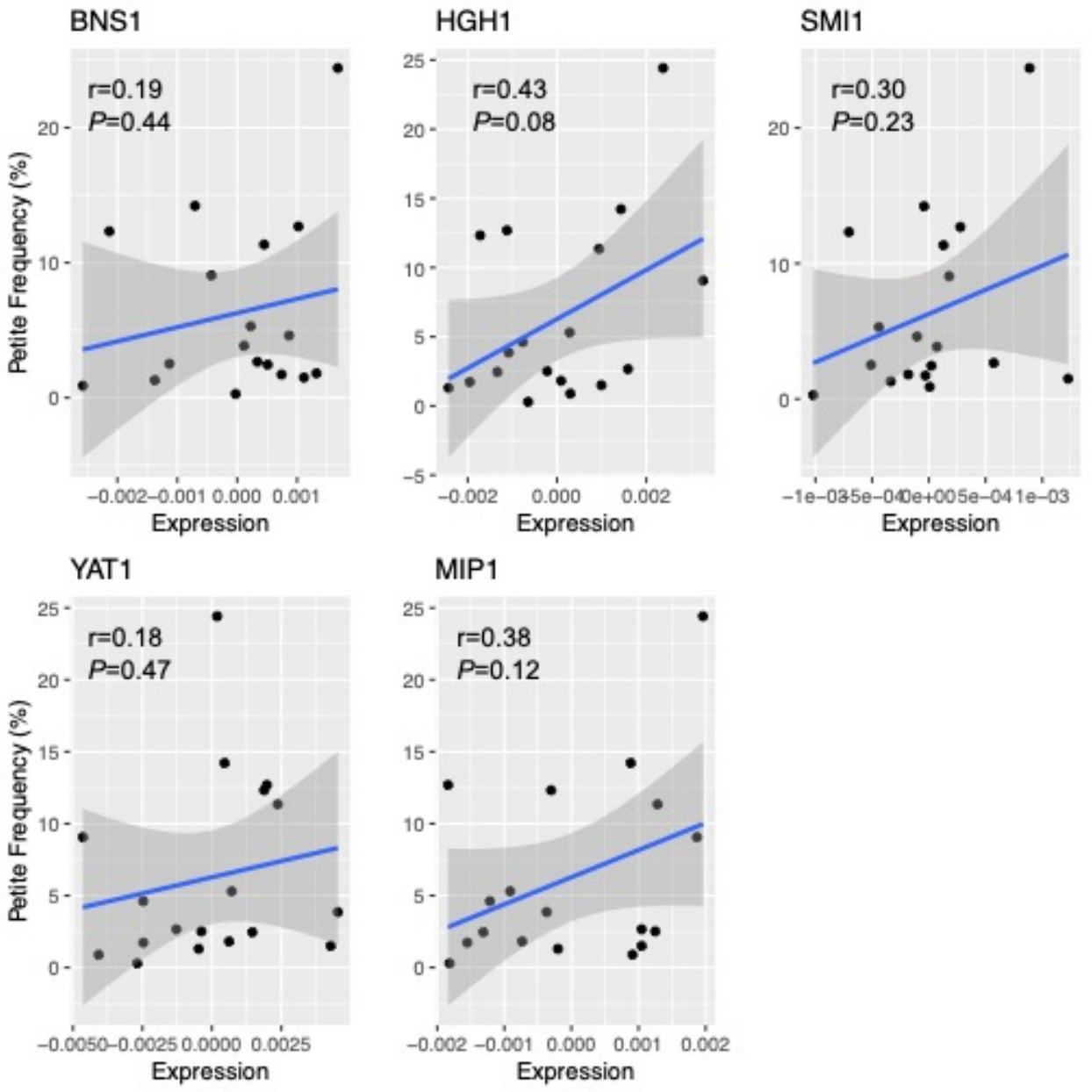

Fig S6. *Petite* frequencies do not correlate with mRNA expression of associated genes. Normalized expression levels (as residuals from regression lines of mRNA levels of each gene compared to a control gene) were plotted against *petite* frequencies. All genes showed positive correlation with *petite* frequencies, though no correlation was statistically significant.

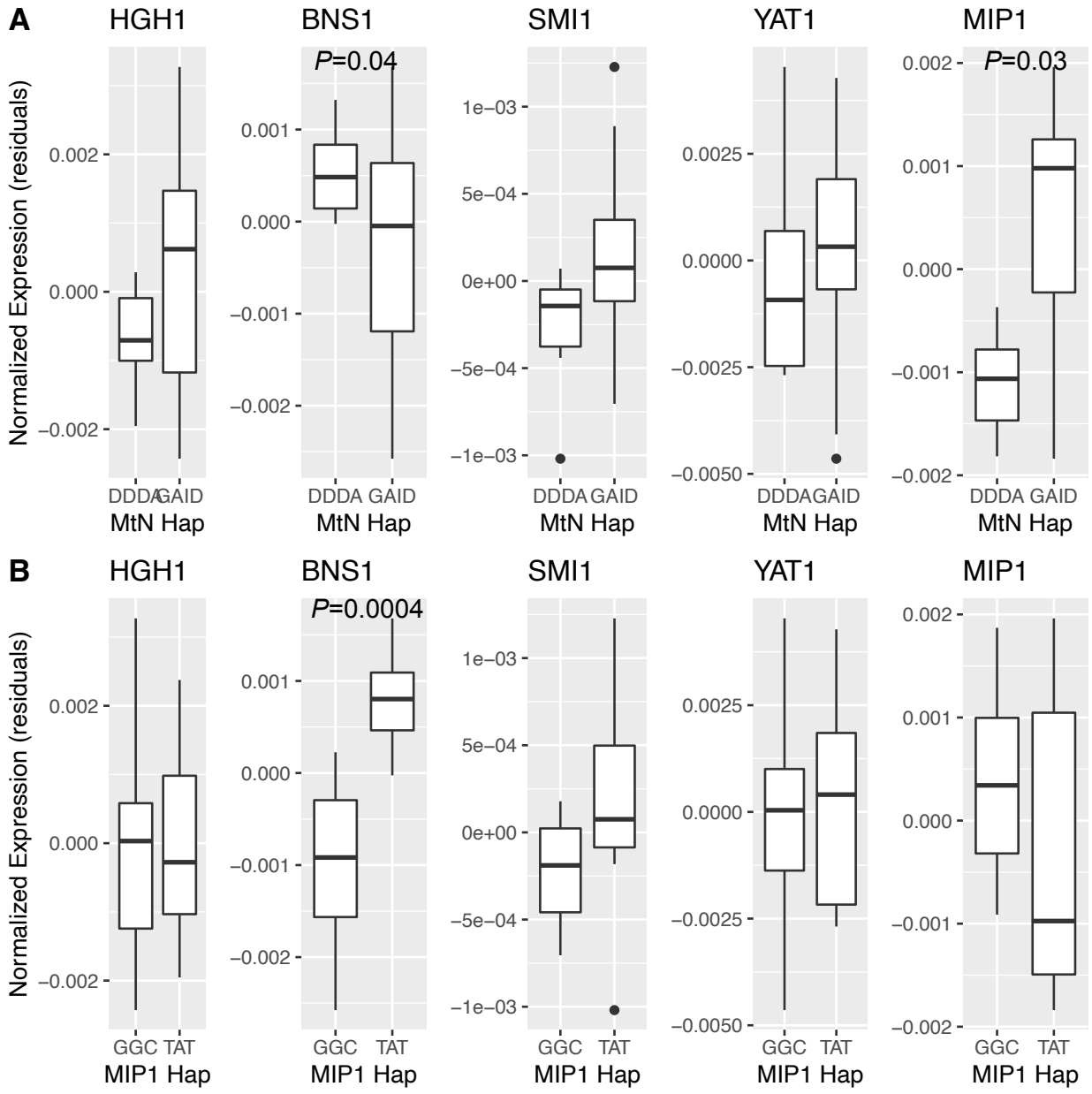

Fig S7. Expression of candidate genes by nuclear haplotype. A. Normalized expression levels of each candidate gene separated by haplotypes of **A.** mitonuclear candidate loci or **B.** *MIP1* loci. The mitonuclear haplotypes represent the SNPs with highest effect sizes for each candidate gene. *P* values for significant differences are shown. All other comparisons were non-significant.

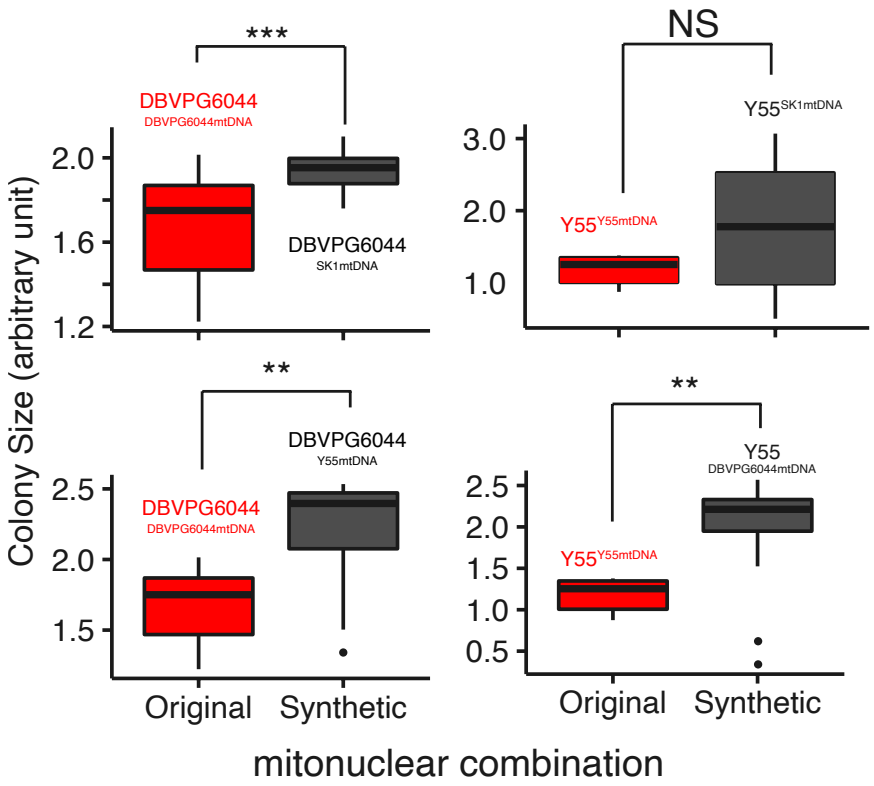

Fig S8. Synthetic mitonuclear genotypes with increased *petite* frequencies have increased growth rates. Maximum colony sizes for strains containing original or synthetic mitonuclear genotypes are presented as boxplots. Each synthetic mitonuclear genotype had higher growth, and higher *petite* frequencies (**Fig S1**), than the original mitonuclear genotype. Growth data were from NGUYEN *et al.* 2020 and collected in the same conditions as the *petite* assays were performed. * $P<0.05$, ** $P \leq 0.005$, *** $P \leq 0.001$

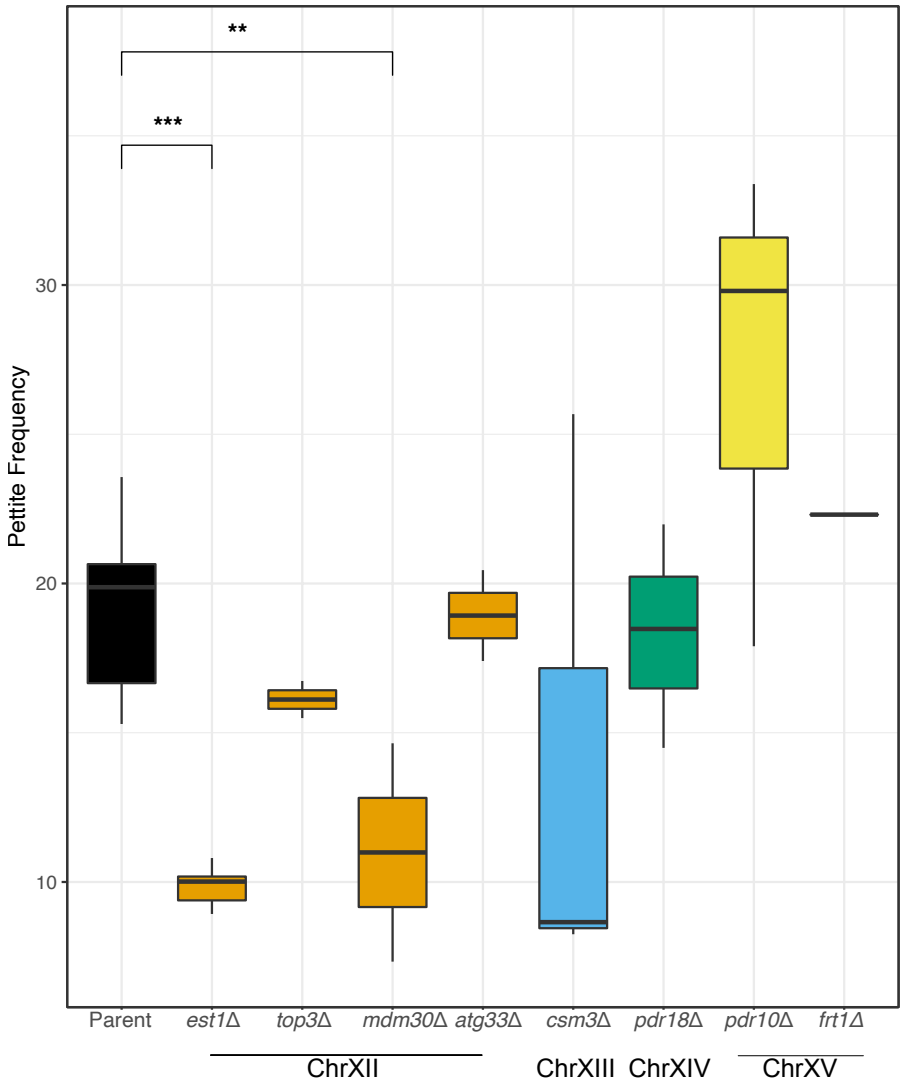

Figure S9. *Petite* frequency of candidate genes associated with mtDNA stability

Petite frequencies of strains containing deletions of candidate genes that did not depend on mitotype are shown as boxplots. Significant differences between the *petite* frequencies of the parental strain and each gene disruption, based on 3 replicates for each strain, are shown. Colors indicate chromosomal location of genes. * $P<0.05$, ** $P \leq 0.005$, *** $P \leq 0.001$