From Genome Variation to Molecular Mechanisms: What we Have Learned From Yeast Mitochondrial Genomes?

Weilong Hao*

Department of Biological Sciences, Wayne State University, Detroit, MI, United States

Analysis of genome variation provides insights into mechanisms in genome evolution. This is increasingly appreciated with the rapid growth of genomic data. Mitochondrial genomes (mitogenomes) are well known to vary substantially in many genomic aspects, such as genome size, sequence context, nucleotide base composition and substitution rate. Such substantial variation makes mitogenomes an excellent model system to study the mechanisms dictating mitogenome variation. Recent sequencing efforts have not only covered a rich number of yeast species but also generated genomes from abundant strains within the same species. The rich yeast genomic data have enabled detailed investigation from genome variation into molecular mechanisms in genome evolution. This mini-review highlights some recent progresses in yeast mitogenome studies.

Keywords: mutation, recombination, repeat, mobile introns, gene conversion, GC-content

INTRODUCTION

Mitochondrial genomes (mitogenomes) originated from an alpha-proteobacterium via endosymbiosis (Lang et al., 1999), and have adopted radically different shapes, sizes, and organizations (Burger et al., 2003; Shao et al., 2009; Sloan et al., 2012; Smith et al., 2012). The great variation of mitogenome diversity and complexity has revolutionized our view of genome evolution and facilitated development of new evolutionary theories (Bazin et al., 2006; Lynch et al., 2006; Whitney and Garland, 2010; Sloan et al., 2012; Christensen, 2013). The excitement of mitogenomes inspired more sequencing projects, and perhaps more importantly, many mindful and in-depth comparative genomics studies (Smith, 2015). These efforts continue to push the boundaries of our understanding in genome evolution.

Mitogenomes are highly variable among yeast species. Mitogenomes show substantial differences in genome size and organization, GC-content, mutation rates, and recombination frequencies. As mutation is the fundamental source of genetic variation, mitogenome differences provide important insights into the underlying mechanisms in the mutation processes. The budding yeast Saccharomyces cerevisiae is among the best-studied model organisms, an abundant number of mitogenomes have been sequenced for S. cerevisiae (Strope et al., 2015; De Chiara et al., 2020) and related species (Freel et al., 2014; Nguyen et al., 2020b). The abundant genomic data have allowed comparative analyses of mitogenomes among closely related yeast species and also intraspecific strains within a species. Detailed variations among closely related mitogenomes become uniquely informative to help identify molecular mechanisms driving the genomic changes.

MITOCHONDRIAL DNA DELETION

The mitochondrial DNAs (mtDNAs) of related yeast strains is well known to show hypervariability in sequence [reviewed in Borst and Grivell (1978)]. Studies of hypervariable yeast mitogenomes
have led to discoveries of new molecular and cellular mechanisms (Dujon, 2020). Yeast spontaneously loses mtDNAs and develops respiratory-incompetent petite colonies (rho−) (Whittaker et al., 1972). Petite colonies arise naturally under normal growth conditions, but the frequency of petite formation varies among strains. In S. cerevisiae, the frequency of developing spontaneous petite colonies varies by about 100-fold between laboratory strains versus natural isolates (Dimitrov et al., 2009). The formation of petite colonies, or petite-positive trait, has been associated with whole-genome duplication of the nuclear genome (Piskur, 2001), which provides the basis for neofunctionalization. A survey in more than one hundred strains, however, shows that the petite positive trait is throughout the Saccharomycetaceae family, much beyond the whole genome duplication species (Fekete et al., 2007). Some deletions in yeast mtDNAs are mediated by gene conversion between GC cluster repeats (Weiller et al., 1991), which are characterized by high GC content and palindromic structure (Yin et al., 1981) (also described below). Petite formation, as a genetic trait, has been studied in genome-wide association studies to identify the associated genes (Dimitrov et al., 2009). The different frequencies in petite formation among strains have been associated with genetic variation in at least four nuclear-encoded genes, the mtDNA polymerase (MIP1) (Baruffini et al., 2007a) and other less-well studied SAL1, CAT5, and MKT1 genes (Dimitrov et al., 2009).

**MOVEMENT OF MITOCHONDRIAL INTRONS**

Genetic crosses of different mitochondrial genotypes have been conducted in S. cerevisiae to understand yeast mitochondrial inheritance and genetics (Wilkie and Thomas, 1973). Mitochondrial markers are inherited in a non-mendelian manner (Jacquier and Dujon, 1985). A well-studied example is the omega (ω) locus in 21S rRNA (Butow, 1985), in which the ω+ strains can transfer the whole intron unidirectionally to an intron-less locus (Dujon et al., 1974). The ω locus belongs to group I intron, a class of self-catalytic ribozymes that often encode a homing endonuclease gene (HEG) (Perlman and Butow, 1989). The self-splicing intron and intron-encoded endonuclease play an important role in driving intron invasion and mobility (Dujon, 1989). The mobility and invasion of the ω intron can go beyond the species boundary. The phylogeny of ω-HEG is significantly different from that of the host 21S rRNA (Goddard and Burt, 1999), suggesting cross-species horizontal transfer of ω-HEG. Given the highly invasive nature of the ω intron, one would expect most yeast strains to be ω+. Yet, many yeast strains remain ω−. Goddard and Burt (1999) provided a framework to explain the sporadic distribution of the ω intron; after invasion, the omega-HEGs undergo rapid degeneration and loss, followed by reinvansion. This process is coined as the Goddard–Burt life cycle by Mukhopadhyay and Hausner (2021). The actual life cycle of introns is likely much more complex than cyclical invasion, degeneration, loss and then reinvansion.

In the Saccharomycetaceae family, mitochondrial introns are found in three genes, cox1, cob, and 21S rRNA, in total 17 intron positions (Wu et al., 2015). Each intron has a unique distribution pattern, and intron content often varies substantially even among small numbers of conspecific strains within species. Except the cox1 i1 intron, all other 16 introns are sporadically distributed. The evolutionary turnover rates of gain and loss among these introns were quantitatively measured (Wu et al., 2015). The high-mobility introns documented in genetic crosses do not necessarily have faster turnover rates than low-mobility introns. The cox1 i1 intron is currently only found in S. cerevisiae strains, with a high rate of intron insertion, it will not be surprising if the cox1 i1 intron is present in some upcoming non-S. cerevisiae mitogenomes. Furthermore, phylogenetically mosaic sequences are evident in both introns and HEGs (Wu and Hao, 2014). Thus, intron and its encoded HEG do not always transmit together as a unit. These findings support that gene conversion between the donor and recipient sequences can take place at both the gene and intragenic levels and lead to insertion or deletion of the adjacent HEG/intron content (Wu and Hao, 2014).

**ELEVATED SEQUENCE EVOLUTION NEAR MOBILE INTRONS**

The distribution of mutation along DNA sequences is not uniform, and yeast mitogenomes are no exception. The recent availability of abundant yeast population genomic data makes it possible to examine genetic diversity along mitochondrial genes. One striking finding is the increased density of single nucleotide polymorphisms (SNP) in exon regions approaching intron boundaries (Repar and Warnecke, 2017). Although intron mobility is recognized to play a critical role in driving the sequence diversity of host genes, the underlying mechanisms cannot be easily identified. There are two possible mechanisms that can increase SNP density in exons (Repar and Warnecke, 2017). First, diverse exonic sequences are gene conversion tracts, or known as co-conversion tracts, acquired from distantly related species. Horizontal transfer of introns and co-conversion tracts has been well documented in plant mitogenomes (Cho and Palmer, 1999). Since most plant mitogenomes have exceptionally low mutation rates, shared long co-conversion tracts among distantly related intron-containing sequence can be easily and convincingly identified (Sanchez-Puerta et al., 2011). Similar to plant mitochondrial introns, yeast mitochondrial introns also undergo horizontal transfer (Wu et al., 2015) and gene conversion at intragenic level (Wu and Hao, 2014). The relative high sequence divergence among yeast mitogenomes makes it challenging to accurately identify the donor species of the diverse co-conversion tracts. Alternatively, the flanking regions of each intron insert site are mutation hotspots because of endonuclease activity and subsequent error-prone repair. Yeast mitogenomes contain stand-alone HEGs, which are not associate with mitochondrial introns. A strong association is evident between the presence of a stand-alone endonuclease gene and high sequence diversity at the end of the endonuclease-adjacent gene (Wu and Hao, 2019). This finding is consistent with the notion that the recognition sites of endonuclease are mutation hotspots.
THE REBIRTH OF GC-CLUSTERS

The var1 gene (or called rps3 in non-yeast fungi) is another well-studied example of unidirectional inheritance in *S. cerevisiae* (Strausberg and Butow, 1981). The var1 gene is polymorphic, and different forms of the var1 gene differ by in-frame insertions of short GC-rich palindromic cluster (GC-cluster) in the coding region (Hudspeth et al., 1984). The terminal sequences of most GC clusters are repeats and different GC clusters can share the same terminal repeats. For instance, AG dinucleotide and TAG trinucleotide repeats are common among GC-clusters in *S. cerevisiae* (Weiller et al., 1989). The terminal repeats are regarded as target-site duplication, and GC-clusters have been suggested to bear ribozyme activity, which catalyzes self-cleavage and ligation reactions (Weiller et al., 1989; Lang et al., 2014). GC-cluster sequences rapidly accumulate substitutions especially in the loop regions, and also undergo dynamic merger and shuffling to form new GC-clusters (Wu and Hao, 2015). Changes at nucleotide bases as well as sequence structure result in highly variable GC cluster sequences among different yeast mitogenomes. GC-clusters are most often found in intergenic regions, but many of them are transcribed into RNAs (Wu and Hao, 2015). All these support the notion that GC clusters are transposable elements. GC-clusters can also be found in protein-coding regions. In yeast *Magnusiomyces capitatus* (in the Dipodascaceae family), GC-clusters inserted in protein-coding regions are transcribed in mRNAs, but the GC-cluster region in mRNA gets bypassed (or ignored) during translation (Lang et al., 2014). GC clusters have been suggested as recombination hotspots in mitogenomes (Dieckmann and Gandy, 1987). GC cluster-mediated gene conversion can insert or delete large genomic fragments (Weiller et al., 1991), which ultimately lead to alteration of genome size. GC clusters have been suggested to induce long AT-rich sequences into the *Nakaseomyces bacillisporus* mitogenome (Bouchier et al., 2009).

MITOCHONDRIAL DNA RECOMBINATION

Pioneer studies on mtDNA recombination through mating between *S. cerevisiae* strains can be traced back to the early 1970s (Kleese et al., 1972; Shannon et al., 1972). Recent large-scale genomic survey confirms frequent mtDNA recombination in natural *S. cerevisiae* populations (De Chiara et al., 2020). Surprisingly, yeasts were not among the organisms in early discoveries of mtDNA recombination between different species, as shown in plants (Hao et al., 2010; Mower et al., 2010). Subsequent analysis on closely related yeast mitogenomes found extensive recombination throughout the mitogenome between yeast species (Wu et al., 2015). There are two important identified issues in assessing the extent of mtDNA recombination events. (1) Accurate detection of mtDNA recombination relies on the abundance of closely related mitogenomes, preferably, an abundant number of intraspecific mitogenomes from several related species. (2) Many mtDNA recombinant events are fine-scale and often overlooked when using whole genes as the unit of phylogenetic analysis.

Fritsch et al. (2014) have constructed the genome-wide map of mtDNA recombination events in *S. cerevisiae* and found that recombinant hotspots are preferentially localized in intergenic and intronic regions. They further investigated the impact of individual deletion of four genes [namely *NgI* (Ling et al., 2007), *MgtI* (Lockshon et al., 1995), *MHR1*, and *Din7* (Ling et al., 2013)] previously associated with mtDNA recombination. The deletion of *NgI*, *MgtI*, and *MHR1* had little influence on mtDNA recombination hotspots, and the deletion of *Din7* resulted primarily in DNA degradation. These findings hint that the study of only nuclear-encoded genes is unlikely to achieve a complete understanding of the localization of recombination hotspots along the yeast mitogenome. Given the dynamic nature of mitochondrial -encoded sequences, such as introns and GC-clusters, a plausible alternative could be that mtDNA recombination hotspots are primarily driven by the sequence features of the mitogenome. Future studies are needed to test this hypothesis.

Mitochondrial DNA recombination can impact mitochondrial function in different ways. As described above (on GC-clusters), Non-allelic mtDNA recombination can lead to deletion and insertion of sequences. Allelic mtDNA recombination, on the other hand, can prevent mtDNA deletion and repair mitogenome damage (Ling et al., 2019). Mitochondrial recombination introduces rapid sequence changes, and could have a significant functional impact on the host. There are direct functional impacts of the introduced mito-genotype, effects of the altered mito-nuclear interaction, and effects of the interaction between the introduced mito-loci and native mito-loci (mito-mito interaction) (Wolters et al., 2018). Mitochondrial recombination can enhance phenotypic variation among diploid hybrids, and facilitate the phenotypic differentiation of hybrid species (Leduq et al., 2017). Given the prevalence of yeast mtDNA recombination and even across different species, much phenotypic diversity in yeast could have resulted from mitochondrial recombination.

YEAST MITOGENOMES ARE HIGHLY VARIABLE IN SIZE

In the Saccharomycetaceae family, the smallest and largest mitogenomes belong in the same genus *Nakaseomyces* (*Figure 1*). The mitogenome in *Candida glabrata* is 20.1 kb (Koszul et al., 2003), while the mitogenome in *Nakaseomyces bacillisporus* is over five-time larger, at 107.1 Kb (Bouchier et al., 2009). In contrast, the sizes of nuclear genomes among the *Nakaseomyces* species are remarkably similar, ranging from 10.2 to 12.3 Mb (Gabaldon et al., 2013). In the sister family Saccharomycodaceae, *Hanseniaspora uvarum* has a linear mitogenome at 11.1 kb (Pramateftaki et al., 2006), and *Saccharomyces ludwigii* has a circular mitogenome at 69.0 kb (Nguyen et al., 2020b). Mitogenome sizes also vary among strains within the same yeast species. Among the 109 *S. cerevisiae* mitogenomes examined in
STABLE GENE CONTENT IN YEAST MITOGENOMES

Despite the various sizes, genes encoded in yeast mitogenomes are remarkably stable. Unlike mitogenomes in many other fungal species, none of the mitogenomes in the Saccharomycetaceae and Saccharomycodaceae encode the respiratory-chain NADH dehydrogenase (complex I) (Dujon, 2010; Freel et al., 2014). The loss of complex I in yeasts is generally believed as a result of adaptation to fermentative lifestyles, where complex I is not essential (Schikora-Tamarit et al., 2021). All complete mitogenomes in Saccharomycetaceae encode eight protein genes, seven respiratory-chain protein genes (atp6, atp8, atp9, cob, cox1, cox2, and cox3) and one ribosomal protein gene var1, SSU and LSU rRNAs, and 22–24 tRNAs. This is in contrast with a recent report that “no gene is universally conserved in fungal mitogenomes” (Fonseca et al., 2021). Although it is inevitable that fewer mitochondrial genes are shared when more diverse lineages are included (Roger et al., 2017), the variation of mitogenome sequence quality could have been an important reason for the discrepancy on gene conservation among yeast mitogenomes. For instance, the two S. cerevisiae mitogenomes (accessions: CM002421 and CP046458) reportedly missing atp6 by Fonseca et al. (2021) are 22,149 bp (including 13,247 Ns for gaps) and 49,451 bp (517 Ns) in length, respectively, much shorter than the average 82.5 kb [ranging from 74.2 to 92.2 kb in Xiao et al. (2017)] of S. cerevisiae mitogenomes (Figure 1). The Saccharomycodaceae family has a sole case of gene loss involving a protein gene. That is the var1 gene absent from Hanseniaspora uvarum (Pramateftaki et al., 2006). This is likely a lineage-specific gene loss, as var1 is present in Saccharomycodes ludwigii (Nguyen et al., 2020b), a related species within the Saccharomycodaceae family.

FACTORS DRIVING MITOGENOME SIZE VARIATION

Mobile introns and variable intergenic regions are known factors driving variation of yeast mitogenome architecture (Bouchier et al., 2009; Freel et al., 2014). Ironically, the mitogenome in Nakaseomyces bacillipsorus at 107 kb contains no intron, while

FIGURE 1 | Variation in size, sequence divergence, and GC-content of mtDNAs of 36 yeast species. The phylogenetic tree is based on nucleotide sequences of seven core mitochondrial genes (cob, cox1, cox2, cox3, atp6, atp8, and atp9). The very long branch leading to Hanseniaspora uvarum is shown by a thickened line at fivefold reduced scale. If two or more strains are analyzed in a species, the averages are shown for genome size and GC% (strain number in parentheses). Otherwise, the strain names are shown. Details of all strain names in the Saccharomycetaceae family are in Xiao et al. (2017), and the Saccharomycodes ludwigii strains are in Nguyen et al. (2020b).

Xiao et al. (2017), and their genome sizes range from 74.2 to 92.2 kb. The fast increasing number of yeast mitogenomes will only expand the range in mitogenome size difference.
mitogenomes in *Candida glabrata* at around 20 kb contain 3–4 introns in at least four intron-distribution patterns [(Koszul et al., 2003), and unpublished observation]. This irony can be solved by separate analyses at different time scales (Xiao et al., 2017). When intraspecific mitogenomes (within the same species) are compared, intron sequences show the highest variance in length and significantly overrepresented in large mitogenomes. When interspecific mitogenomes (among different species) are compared, tandem repeats show the highest variance in sequence length and significantly overrepresented in large mitogenomes. In other words, the rapid turnover of mobile introns can significantly impact genome size, but the number of available introns insertion sites are limited; while expansion and contraction of repeats may cause only subtle change per event, but they take place persistently with little space limit.

**ORIGIN OF MITOGENOME SIZE**

The question on whether genome size is under selective constraint has been a subject of debate (Lynch and Conery, 2003; Whitney and Garland, 2010; Shtolz and Mishmar, 2019). The mutational burden hypothesis (MBH) was postulated to explain the origin of organellar genome size (Lynch et al., 2006). Introns and intergenic DNAs in mitogenomes are genetic liability, as they are targets for deleterious and potentially lethal mutations (Lynch et al., 2006). Following MBH, introns and intergenic DNAs tend to accumulate when natural selection is less efficient at purging hazardous non-coding DNA. Using the ratio of non-synonymous over synonymous Ka/Ks rates as a proxy for the level of genetic drift, a significant positive correlation was evident between the genome-wide Ka/Ks ratios and mitogenome sizes among seven yeast species with sufficient intraspecific diversity (Xiao et al., 2017). This finding is consistent with the notion that introns, GC-clusters, and repeats in yeast mitogenomes are mostly deleterious (Bernardi, 2005; Wu and Hao, 2014, 2015). *Hanseniaspora uvarum* has accelerated sequence evolution compared with related species (Figure 1) and has a small genome size, which is consistent with the MBH hypothesis. Genetic drift still faces challenges to explain mitogenome size variation in many yeast species. For example, significant relaxation of mitochondrial functions has been documented after whole-genome duplication (Jiang et al., 2008), yet many post-whole-genome duplication species (including *Candida glabrata*) have compact mitogenomes. Unfortunately, many yeast species either lack intraspecific genomic data or suffer insufficient sequence diversity, making it impossible to estimate the degree of genetic drift. To address this issue, extended efforts are needed to sample and sequence an abundant number of intraspecific strains in a broad range of yeast species.

**EXTREME GENOME-WIDE G + C CONTENT**

Even though yeast mitogenomes are overall AT-rich, their G + C-contents vary greatly. GC-content of the *Hanseniaspora uvarum* mitogenome is 29.3% (Pramateftaki et al., 2006), while the most AT-rich mitogenome in *Saccharomyces ludwigii* is at 7.6% G + C (Nguyen et al., 2020b). Generally speaking, mutation is nearly universally biased from C/G to T/A (Hershberg and Petrov, 2010), and mutation rates are often higher at C/G nucleotides than at A/T nucleotides (Zhu et al., 2014). Alterations of GC content have been shown to impact mutation and recombination rates (Kiktev et al., 2018). The mitogenomes with extreme base composition offer important and unique insights into the mechanisms governing mutation processes (Gardner et al., 2002; McCutcheon and Moran, 2010; Smith et al., 2011; Su et al., 2019). Comparative genomics of the 10 extreme AT-rich mitogenomes in *Saccharomyces ludwigii* (Nguyen et al., 2020b) found a strong mutation bias toward A/T, but the expected equilibrium G + C content under mutation pressure alone is still higher than observed G + C content. Interestingly, mitogenomes in *Saccharomyces ludwigii* undergo frequent recombination, a genetic process that normally increases G + C content by GC-biased gene conversion (Pessia et al., 2012). These findings suggest other mechanisms alongside with AT-biased mutation operating to increase A/T in *Saccharomyces ludwigii*. Another important, but perhaps underappreciated finding is the prevalence of indel mutations in yeast mitogenomes (Xiao et al., 2017). Indel mutations contribute much more to genomic variation among closely related mitogenomes than nucleotide substitutions. Further studies are needed to investigate the molecular mechanisms driving indel mutations and quantitatively model the evolutionary process of indel mutations.

**VARIABLE EVOLUTIONARY RATES AMONG YEAST MITOGENOMES**

Yeast mitogenomes show variable evolutionary rates. The branch leading to *Hanseniaspora uvarum* is at least five times longer than the branch leading to its sister species. Similarly, the branch leading to a plant pathogen *Eremothecium gossypii* is at least three times longer than the branch leading to the related *Kluyveromyces* species (Figure 1). Spontaneous mitochondrial mutation rates have been measured in several yeast species. The mitochondrial base-substitution mutation (BSM) rates are all higher than their corresponding nuclear BSM rates. The mitochondrial BSM rates in *S. cerevisiae* range from $4.47 \times 10^{-10}$ to $122.3 \times 10^{-10}$ per site per cell division (Lynch et al., 2008). Please also note that the *S. cerevisiae* FY10 strain (isogenic to S288c) used in the Lynch et al. (2008) study contains a single non-synonymous mutation in the mtDNA polymerase (*MIP1*) linked to reduced fidelity of mtDNA replication (Baruffini et al., 2007b). *Hanseniaspora uvarum* does have a higher mitochondrial BSM rate than other two *Hanseniaspora* species, at $13.1 \times 10^{-10}$ mutations per site per cell division, compared with $5.94 \times 10^{-10}$ mutations per site per cell division in *Hanseniaspora valbyensis*, and $3.65 \times 10^{-10}$ mutations per site per cell division in *Hanseniaspora osmophila* (Nguyen et al., 2020a). The measured spontaneous mutation rates will also help us estimate the effective population size for each species following equation...
\[ \pi_{silent} = 2N_e \mu, \text{ where } \pi_{silent} \text{ is nucleotide diversity at silent sites, } N_e \text{ is effective population size, } \mu \text{ is mutation rate (Lynch, 2006). Precise estimation of effective population size for mtDNA holds the key to understanding the significance of mtDNA recombination at the population level. To achieve this, it is critical to obtain both intraspecific diversity and spontaneous mutation rate for a variety of yeast species. Extra attention must also be paid to accumulate sufficient number of mitochondrial mutations due to the small mitogenome size relative to the nuclear genome size.}

**MOVING FORWARD**

Yeast mitogenomes are highly diverse ranging from fast-evolving compact mitogenomes (similar to animal mitogenomes) to slow-evolving mitogenomes inflated by large introns, repeats and non-coding sequences (similar to plant mitogenomes). The fast-growing yeast mitogenome data have allowed us to begin to identify mechanisms driving genome diversity. Specific efforts are needed to sequence and study an abundant number of intraspecific strains from closely related species for contrast genomic features. The future of yeast mitogenome studies is bright, and the generated knowledge will no doubt benefit our understanding of mitogenomes much beyond the fungal kingdom.

**AUTHOR CONTRIBUTIONS**

WH wrote the manuscript.

**FUNDING**

Previous work in the lab was supported by funds from Wayne State University and the National Science Foundation (ISO1838291) to WH.

**ACKNOWLEDGMENTS**

I am grateful to GH and three reviewers for helpful comments on a previous version of this manuscript.

**REFERENCES**

Baruffini, E., Ferrero, I., and Foury, F. (2007a). Mitochondrial DNA defects in *Saccharomyces cerevisiae* caused by functional interactions between DNA polymerase gamma mutations associated with disease in human. *Biochim. Biophys. Acta* 1772, 1223–1235. doi: 10.1016/j.bbadis.2007.10.002

Baruffini, E., Lodi, T., Dallabona, C., and Foury, F. (2007b). A single nucleotide polymorphism in the DNA polymerase gamma gene of *Saccharomyces cerevisiae* laboratory strains is responsible for increased mitochondrial DNA mutability. *Genetics* 177, 1227–1231. doi: 10.1534/genetics.107.079293

Bazin, E., Glennin, S., and Galtier, N. (2006). Population size does not influence mitochondrial genetic diversity in animals. *Science* 312, 570–572. doi: 10.1126/science.1122033

Bernardi, G. (2005). Lessons from a small, dispensable genome: the mitochondrial genome of yeast. *Gene* 354, 189–200. doi: 10.1016/j.gene.2005.03.024

Borst, P., and Grivell, L. A. (1978). The mitochondrial genome of yeast. *Cell* 15, 705–723. doi: 10.1016/0092-8674(78)90257-x

Boucher, C., Ma, L., Creno, S., Dujon, B., and Fairhead, C. (2009). Complete mitochondrial genome sequences of three *Nakaseomyces* species reveal invasion by palindromic GC clusters and considerable size expansion. *FEBS Yeast Res.* 9, 1283–1292. doi: 10.1111/j.1567-1364.2009.00551.x

Burger, G., Gray, M. W., and Lang, B. F. (2003). Mitochondrial genomes: anything goes. *Trends Genet.* 19, 709–716.

Butow, R. A. (1985). Nonreciprocal exchanges in the yeast mitochondrial genome. *Trends Genet.* 1, 81–84.

Cho, Y., and Palmer, J. D. (1999). Multiple acquisitions via horizontal transfer of a group I intron in the mitochondrial cox1 gene during evolution of the Araceae family. *Mol. Biol. Evol.* 16, 1153–1165.

Christensen, A. C. (2013). Plant mitochondrial genome evolution can be explained by DNA repair mechanisms. *Genome Biol. Evol.* 5, 1079–1086. doi: 10.1093/gbe/evt069

De Chiara, M., Friedrich, A., Barré, B., Breitenbach, M., Schacherer, J., and Liti, G. (2020). Discordant evolution of mitochondrial and nuclear yeast genomes at population level. * BMC Biol.* 18:49. doi: 10.1186/s12915-020-00786-4

Dieckmann, C. L., and Gandy, B. (1987). Preferential recombination between GC clusters in yeast mitochondrial DNA. *EMBO J.* 6, 419–4203.

Dimitrov, L. N., Brem, R. B., Kruglyak, L., and Gottschling, D. E. (2009). Polymorphisms in multiple genes contribute to the spontaneous mitochondrial genome instability of *Saccharomyces cerevisiae* S288C strains. *Genetics* 183, 365–383. doi: 10.1534/genetics.109.104497

Dujon, B. (1989). Group I introns as mobile genetic elements: facts and mechanistic speculations—a review. *Gene* 82, 91–114.

Dujon, B. (2010). Yeast evolutionary genomics. *Nat. Rev. Genet.* 11, 512–524. doi: 10.1038/nrg2811

Dujon, B. (2020). Mitochondrial genetics revisited. *Yeast* 37, 191–205. doi: 10.1002/yea.3445

Dujon, B., Slonimski, P. P., and Weill, L. (1974). Mitochondrial genetics IX: a model for recombination and segregation of mitochondrial genomes in *Sacharomyces cerevisiae*. *Genetics* 78, 415–437. doi: 10.1093/genetics/78.1.415

Fekete, V., Cierna, M., Poláková, S., Piskur, J., and Sulo, P. (2007). Transition of the ability to generate petites in the *Saccharomyces/Klugeyeromyces* complex. *FEBS Yeast Res.* 7, 1237–1247. doi: 10.1111/j.1567-1364.2007.00287.x

Fonseca, P. L. C., De-Paula, R. B., AraiJho, D. S., Tomé, L. M. R., Mendes-Pereira, T., Rodrigues, W. F. C., et al. (2021). Global characterization of fungal mitogenomes: new insights on genomic diversity and dynamism of coding genes and accessory elements. *Front. Microbiol.* 12:787283. doi: 10.3389/fmicb.2021.787283

Freed, K. C., Friedrich, A., Hou, J., and Schacherer, J. (2014). Population genomic analysis reveals highly conserved mitochondrial genomes in the yeast species *Lachancea thermotolerans*. *Genome Biol. Evol.* 6, 2586–2594. doi: 10.1093/gbe/evu203

Fritsch, E. S., Chabbert, C. D., Klaus, B., and Steinmetz, L. M. (2014). A genome-wide map of mitochondrial DNA recombination in yeast. *Genetics* 198, 755–771. doi: 10.1534 genetics.114.166637

Gabaldon, T., Martin, T., Marceœ-Hœben, M., Durrens, P., Bolotin-Fukuhara, M., Lespinet, O., et al. (2013). Comparative genomics of emerging pathogens in the *Candida glabrata* clade. * BMC Genomics* 14:623. doi: 10.1186/1471-2164-14-623

Gardner, M. J., Hall, N., Fung, E., White, O., Berriman, M., Hyman, R. W., et al. (2002). Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature* 419, 498–511. doi: 10.1038/nature01097

Goddard, M. R., and Burt, A. (1999). Recurrent invasion and extinction of a selfish gene. *Proc. Natl. Acad. Sci. U.S.A.* 96, 13880–13885.

Hao, W., Richardson, A. O., Zheng, Y., and Palmer, J. D. (2010). Gorgeous mosaic of mitochondrial genes created by horizontal transfer and gene conversion. *Proc. Natl. Acad. Sci. U.S.A.* 107, 21576–21581. doi: 10.1073/pnas.1016295107

Hershberg, R., and Petrov, D. A. (2010). Evidence that mutation is universally biased towards AT in bacteria. *PLoS Genet.* 6:e1001115. doi: 10.1371/journal.pgen.1001115

Hudspeth, M. E., Vincent, R. D.,Perlman, P. S., Shumard, D. S., Treisman, L. O., and Grossman, L. I. (1984). Expandable var1 gene of yeast mitochondrial DNA:
in-frame insertions can explain the strain-specific protein size polymorphisms. Proc. Natl. Acad. Sci. U.S.A. 81, 3148–3152. doi: 10.1073/pnas.81.10.3148
Jacquier, A., and Dujon, B. (1985). An intron-encoded protein is active in a gene conversion process that spreads an intron into a mitochondrial gene. Cell 41, 383–391. doi: 10.1016/0092-8674(85)90011-8
Jiang, H., Guan, W., Pinney, D., Wang, W., and Gu, Z. (2008). Relaxation of yeast mitochondrial functions after whole-genome duplication. Genome Res. 18, 1466–1471. doi: 10.1101/gr.74674.107
Kiytek, D. A., Sheng, Z., Lobachev, K. S., and Petes, T. D. (2018). GC content elevates mutation and recombination rates in the yeast Saccharomyces cerevisiae. Proc. Natl. Acad. Sci. U.S.A. 115, E7109–E7118. doi: 10.1073/pnas.1807341115
Kleece, R. A., Grotheck, R. C., and Snyder, J. R. (1972). Recombination among three mitochondrial genes in yeast (Saccharomyces cerevisiae). J. Bacteriol. 112, 1023–1025.
Koszul, R., Malpertuy, A., Frangeul, L., Bouchier, C., Wincker, P., Thierry, A., et al. (2003). The complete mitochondrial genome sequence of the pathogenic yeast Candida (Torulopsis) glabrata. FEBS Lett. 534, 39–48.
Lang, B. F., Gray, M. W., and Burger, G. (1999). Mitochondrial genome evolution and the origin of eukaryotes. Annu. Rev. Genet. 33, 351–397. doi: 10.1146/annurev.genet.33.1.351
Lang, B. F., Jakubkova, M., Hageduova, E., Daoud, R., Forget, L., Brejova, B., et al. (2014). Massive programmed translational jumping in mitochondria. Proc. Natl. Acad. Sci U.S.A 111, 5926–5931. doi: 10.1073/pnas.1322190111
Leduq, J. B., Henault, M., Charron, G., Nielly-Thibault, L., Terrat, Y., Fiumera, H. L., et al. (2017). Mitochondrial recombination and introgression during speciation by hybridization. Mol. Biol. Evol. 34, 1947–1959. doi: 10.1093/molbev/msx139
Ling, F., Bradshaw, E., and Yoshida, M. (2019). Prevention of mitochondrial genomic instability in yeast by the mitochondrial recombainase Mhr1. Sci. Rep. 9:5433. doi: 10.1038/s41598-019-41699-9
Ling, F., Hori, A., and Shibata, T. (2007). DNA recombination-initiation plays a role in the extremely biased inheritance of yeast [rho−] mitochondrial DNA that contains the replication origin ori5. Mol. Cell. Biol. 27, 1133–1145. doi: 10.1128/MCB.00770-06
Ling, F., Hori, A., Yoshihata, N., Niu, R., Yoshida, M., and Shibata, T. (2013). Din7 and Mhr1 expression levels regulate double-strand-break-induced replication and recombination of mtDNA at ori5 in yeast. Nucleic Acids Res. 41, 5799–5816. doi: 10.1093/nar/gkt273
Lockshon, D., Weifel, S. G., Freeman-Cook, L. L., Lorimer, H. E., Brewer, B. J., and Fangman, W. L. (1995). A role for recombination junctions in the segregation of mitochondrial DNA in yeast. Cell 81, 947–955.
Lynch, M. (2006). The origins of eukaryotic gene structure. Trends Genet. 22, 23, 947–955.
Lynch, M., Koskella, B., and Schaack, S. (2006). Mutation pressure and the origin of genome complexity. Curr. Biol. 16, 23, 947–955.
Lynch, M., and Conery, J. S. (2003). The origins of genome complexity. Proc. Natl. Acad. Sci. U.S.A. 100, 13,977–13,982. doi: 10.1073/pnas.2132190111
Mhr1, J. P., and Warnecke, T. (2017). Mobile introns shape the genetic diversity of the human body louse, Pediculus humanus. Genome Res. 19, 904–911. doi: 10.1101/gr.534010.108
Sharp, N. P., Sandell, L., James, C. G., and Otto, S. P. (2018). The genome-wide rate and spectrum of spontaneous mutations differ between haploid and diploid yeast. Proc. Natl. Acad. Sci. U.S.A. 115, E5046–E5055. doi: 10.1073/pnas.1801040115
Shao, R., Kirkness, E. F., and Barker, S. C. (2009). The single mitochondrial chromosome typical of animals has evolved into 18 minichromosomes in the human body louse, Pediculus humanus. Genome Res. 19, 904–911. doi: 10.1101/gr.534010.108
Shulz, N., and Mishmar, D. (2019). The mitochondrial genome–on selective constraints and signatures at the organism, cell, and single mitochondrial levels. Front. Ecol. Evol. 7:342. doi: 10.3389/felev.2019.00342
Smith, D. R. (2015). The past, present and future of mitochondrial genomics: have we sequenced enough mtDNAs? Brief. Funct. Genomics 15, 47–54. doi: 10.1093/bfgp/elv027
Smith, D. R., Burki, F., Christians, R., and Otto, S. P. (2018). The genome-wide rate and spectrum of spontaneous mutations differ between haploid and diploid yeast. Proc. Natl. Acad. Sci. U.S.A. 115, E5046–E5055. doi: 10.1073/pnas.1801040115
Smith, D. R. (2015). The past, present and future of mitochondrial genomics: have we sequenced enough mtDNAs? Brief. Funct. Genomics 15, 47–54. doi: 10.1093/bfgp/elv027
Smith, D. R., Burki, F., Christians, R., and Otto, S. P. (2018). The genome-wide rate and spectrum of spontaneous mutations differ between haploid and diploid yeast. Proc. Natl. Acad. Sci. U.S.A. 115, E5046–E5055. doi: 10.1073/pnas.1801040115
Smith, D. R. (2015). The past, present and future of mitochondrial genomics: have we sequenced enough mtDNAs? Brief. Funct. Genomics 15, 47–54. doi: 10.1093/bfgp/elv027
Smith, D. R., Burki, F., Christians, R., and Otto, S. P. (2018). The genome-wide rate and spectrum of spontaneous mutations differ between haploid and diploid yeast. Proc. Natl. Acad. Sci. U.S.A. 115, E5046–E5055. doi: 10.1073/pnas.1801040115
Smith, D. R. (2015). The past, present and future of mitochondrial genomics: have we sequenced enough mtDNAs? Brief. Funct. Genomics 15, 47–54. doi: 10.1093/bfgp/elv027
Smith, D. R., Burki, F., Christians, R., and Otto, S. P. (2018). The genome-wide rate and spectrum of spontaneous mutations differ between haploid and diploid yeast. Proc. Natl. Acad. Sci. U.S.A. 115, E5046–E5055. doi: 10.1073/pnas.1801040115
Smith, D. R. (2015). The past, present and future of mitochondrial genomics: have we sequenced enough mtDNAs? Brief. Funct. Genomics 15, 47–54. doi: 10.1093/bfgp/elv027
Smith, D. R., Burki, F., Christians, R., and Otto, S. P. (2018). The genome-wide rate and spectrum of spontaneous mutations differ between haploid and diploid yeast. Proc. Natl. Acad. Sci. U.S.A. 115, E5046–E5055. doi: 10.1073/pnas.1801040115
Smith, D. R. (2015). The past, present and future of mitochondrial genomics: have we sequenced enough mtDNAs? Brief. Funct. Genomics 15, 47–54. doi: 10.1093/bfgp/elv027
Weiller, G., Schueller, C. M., and Schweyen, R. J. (1989). Putative target sites for mobile G + C rich clusters in yeast mitochondrial DNA: single elements and tandem arrays. Mol. Gen. Genet. 218, 272–283.

Whitney, K. D., and Garland, T. Jr. (2010). Did genetic drift drive increases in genome complexity? PLoS Genet. 6:e1001080. doi: 10.1371/journal.pgen.1001080

Whittaker, P. A., Hammond, R. C., and Luha, A. A. (1972). Mechanism of mitochondrial mutation in yeast. Nat. New Biol. 238, 266–268. doi: 10.1038/newbio238266a0

Wilkie, D., and Thomas, D. Y. (1973). Mitochondrial genetic analysis by zygote cell lineages in Saccharomyces cerevisiae. Genetics 73, 367–377. doi: 10.1093/genetics/73.3.367

Wolters, J. F., Charron, G., Gaspar, A., Landry, C. R., Fiumera, A. C., and Fiumera, H. L. (2018). Mitochondrial recombination reveals mito-mito epistasis in yeast. Genetics 209, 307–319. doi: 10.1534/genetics.117.300660

Wu, B., and Hao, W. (2014). Horizontal transfer and gene conversion as an important driving force in shaping the landscape of mitochondrial introns. G3 4, 605–612.

Wu, B., and Hao, W. (2015). A dynamic mobile DNA family in the yeast mitochondrial genome. G3 5, 1273–1282. doi: 10.1534/g3.115.017822

Wu, B., and Hao, W. (2019). Mitochondrial-encoded endonucleases drive recombination of protein-coding genes in yeast. Environ. Microbiol. 21, 4233–4240. doi: 10.1111/1462-2920.14783

Wu, B., Buljic, A., and Hao, W. (2015). Extensive horizontal transfer and homologous recombination generate highly chimeric mitochondrial genomes in yeast. Mol. Biol. Evol. 32, 2559–2570. doi: 10.1093/molbev/msv127

Xiao, S., Nguyen, D. T., Wu, B., and Hao, W. (2017). Genetic drift and indel mutation in the evolution of yeast mitochondrial genome size. Genome Biol. Evol. 9, 3088–3099. doi: 10.1093/gbe/evx225

Yin, S., Heckman, J., and RajBhandary, U. L. (1981). Highly conserved GC-rich palindromic DNA sequences flank tRNA genes in Neurospora crassa mitochondria. Cell 26, 325–332. doi: 10.1016/0092-8674(81)90201-4

Zhu, Y. O., Siegal, M. L., Hall, D. W., and Petrov, D. A. (2014). Precise estimates of mutation rate and spectrum in yeast. Proc. Natl. Acad. Sci. U.S.A. 111, E2310–E2318. doi: 10.1073/pnas.1323011111

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Hao. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.