Extensive Horizontal Transfer and Homologous Recombination Generate Highly Chimeric Mitochondrial Genomes in Yeast

Baojun Wu,1 Adnan Buljic,1 and Weilong Hao*1
1Department of Biological Sciences, Wayne State University
*Corresponding author: E-mail: haow@wayne.edu.
Associate editor: Andrew Roger

Abstract

The frequency of horizontal gene transfer (HGT) in mitochondrial DNA varies substantially. In plants, HGT is relatively common, whereas in animals it appears to be quite rare. It is of considerable importance to understand mitochondrial HGT across the major groups of eukaryotes at a genome-wide level, but so far this has been well studied only in plants. In this study, we generated ten new mitochondrial genome sequences and analyzed 40 mitochondrial genomes from the Saccharomycetales to assess the magnitude and nature of mitochondrial HGT in yeasts. We provide evidence for extensive, homologous-recombination-mediated, mitochondrial-to-mitochondrial HGT occurring throughout yeast mitochondrial genomes, leading to genomes that are highly chimeric evolutionarily. This HGT has led to substantial intraspecific polymorphism in both sequence content and sequence divergence, which to our knowledge has not been previously documented in any mitochondrial genome. The unexpectedly high frequency of mitochondrial HGT in yeast may be driven by frequent mitochondrial fusion, relatively low mitochondrial substitution rates and pseudohyphal fusion to produce heterokaryons. These findings suggest that mitochondrial HGT may play an important role in genome evolution of a much broader spectrum of eukaryotes than previously appreciated and that there is a critical need to systematically study the frequency, extent, and importance of mitochondrial HGT across eukaryotes.

Key words: horizontal gene transfer, gene conversion, mitochondrial fusion.

Introduction

Horizontal gene transfer (HGT) in mitochondrial DNA (mtDNA) has been increasingly recognized as an important evolutionary process during eukaryotic evolution, and the frequency of mitochondrial HGT has been documented to vary substantially among eukaryotic groups (Keeling and Palmer 2008; Rice et al. 2013). Thousands of animal mitochondrial genomes have been sequenced, yet mitochondrial HGT (i.e., mitochondrial-to-mitochondrial HGT) has been credibly reported in only three studies, all involving cox1 introns and nonbilaterian animals (Goddard et al. 2006; Fukami et al. 2007; Szitenberg et al. 2010). In sharp contrast, HGT is relatively common in plant mitochondrial genomes (Won and Renner 2003; Berghthorsson et al. 2004; Davis and Wurdack 2004; Rice et al. 2013; Xi et al. 2013), with foreign mitochondrial sequences far outnumbering native ones in the truly exceptional mitochondrial genome of the angiosperm Amborella trichopoda (Rice et al. 2013). With the surprising discovery of frequent and massive mitochondrial HGT in plants, it is important to determine the prevalence and significance of mitochondrial HGT in other eukaryotes.

Fungi represent another highly successful group of eukaryotes, with over 100,000 described species (Blackwell 2011). The yeast family Saccharomycetales is arguably the most well-studied fungal group with abundant genomic resources. Mitochondrial genome in the Saccharomycetales encodes eight protein genes and two rRNA genes, with the cox1, cob, and 21S-rRNA genes often containing intron(s). Yeast mitochondrial introns belong to group I (often encoding homing endonuclease genes, HEGs [Cech 1990; Haugen et al. 2005]) or group II (often encoding reverse-transcriptase genes [Lambowitz and Belfort 1993]) classes, some of which can catalyze their own splicing. Mitochondrial introns can invade intronless alleles via intron homing or retrohoming (Lambowitz and Belfort 1993; Eskes et al. 2000; Chevalier and Stoddard 2001) and have been used as markers for mtDNA exchange during intraspecific genetic crosses (Kleese et al. 1972; Shannon et al. 1972; Dujon et al. 1974; Williamson and Fennell 1974). The frequent occurrence of mtDNA exchange among conspecific yeast strains under laboratory conditions (Fritsch et al. 2014) inspired us to systematically examine the frequency of mtDNA exchange between yeast species, that is, mitochondrial HGT.

To date, mitochondrial HGT has been reported in six yeast studies (Lang 1984; Lang et al. 1985; Goddard and Burt 1999; Schafer and Wolf 1999; Peris et al. 2014; Wu and Hao 2014). Three of these involve isolated cases of horizontal transfer of an intron in the cox1 gene (Lang 1984; Schafer and Wolf 1999) or in the cob gene (Lang et al. 1985). Two studies documented remarkably rapid turnover and recurrent HGT of an intron (known as the omega intron) in the 21S-rRNA gene of yeast (Goddard and Burt 1999; Wu and Hao 2014). For example, the omega intron in Torulaspora delbrueckii, a ubiquitous yeast associated with winemaking and other bioprocesses, is strikingly different from that of all other Torulaspora species but highly similar to that of Saccharomyces cerevisiae (Goddard and Burt 1999; Wu and Hao 2014), whereas conspecific strains of Torulaspora pretoriensis contain distinctly
different omega intron sequences (Wu and Hao 2014). The highly mosaic nature of both the T. delbrueckii omega intron and HEG suggests that homologous recombination, in addition to intron homing, is an important driving force in intron transfer (Wu and Hao 2014). A recent analysis of the cox2 genes in the Saccharomyces clade showed recombinant cox2 alleles in four beer-brewing hybrids and a related wild isolate in Saccharomyces eubayanus (Peris et al. 2014).

Is the omega intron exceptionally prone to horizontal transfer or is it a harbinger of perhaps common but largely overlooked HGT throughout the mitochondrial genome of yeasts? Is the discovery of HGT in the cox2 gene just a rare incidence related to hybrids and hybrid-forming species, or is it the tip of the iceberg of mitochondrial HGT in protein coding genes? There is, therefore, a crucial need to thoroughly examine mitochondrial HGT across yeast species at a genome-wide level. To answer these questions, we generated mitochondrial genome sequences for one Torulaspora delbrueckii strain and nine strains from four Torulaspora species, in which horizontal transfer of the omega intron has been previously documented (Goddard and Burt 1999; Wu and Hao 2014). We reconstructed phylogenetic relationships (fig. 1) and conducted comparative genomic analysis of these and 30 additional mitochondrial genomes from the yeast family Saccharomycetaceae (supplementary table S1, Supplementary Material online). Our results provide evidence of HGT in introns, exons, and intron-lacking protein genes across the mitochondrial genome. We discuss the underlying mechanism of extensive mitochondrial HGT and predict that extensive mitochondrial HGT might be a propensity common in the entire fungi kingdom and other eukaryotic lineages.

**Results**

**Intron Gain and Loss**

Introns are found in three mitochondrial genes, cox1, cob, and 21S-rRNA in the Saccharomycetaceae (henceforth referred to as “yeast” or “yeasts”). There are ten intron positions in cox1, six intron positions in cob, and one (the omega intron) in 21S-rRNA (fig. 2). These 17 introns have strikingly different distribution patterns. Each intron has a unique distribution pattern, and intron content often varied substantially among conspecific strains despite the small numbers (2–5) of these

---

**FIG. 1.** Phylogeny of yeast strains in the Saccharomycetaceae family based on the concatenated sequences of seven mitochondrial protein genes (atp6, atp8, atp9, cox1, cox2, cox3, and cob). Node support values (likelihood bootstrap/Bayesian posterior probability) are shown when maximum likelihood bootstrap values are $>50\%$. The tree was rooted using Pichia sorbitiphila, Pichia farinosa, Komagataella pastoris, and Hanseniaspora uvarum as the outgroups. The three taxon-rich clades are colored with Lachancea strains in red, Saccharomyces in green, and Torulaspora in blue. The red star represents a whole-genome duplication event in the nuclear genome.
examined (Fig. 2). For example, only one of the seven introns (14.2%) in *T. delbrueckii* (the omega intron) is shared by all three *T. delbrueckii* strains examined. The fraction of introns shared by conspecific strains is 54.5% (6/11) in *Torulaspora globosa* (two strains examined), 60% (9/15) in *S. cerevisiae* (four strains), 60% (6/10) in *S. paradoxus* (two strains), 66.7% (6/9) in *Lachancea kluyveri* (five strains), and 66.7% (6/9) in *T. pretoriensis* (two strains). Identical intron patterns were found for four sets of conspecific strains (three pairs and one trio), but in three of these four cases, the entire mitochondrial genomes were extremely similar (at the extreme, in *Torulaspora franciscae*, differing by only a single nucleotide polymorphism).

Yeast mitochondrial introns are all group I or group II introns, and 7 out of the 17 introns have demonstrated mobility in experimental genetic crosses (Dujon 1980; Delahodde et al. 1989; Lazowska et al. 1992; Moran et al. 1992; Lazowska et al. 1994; Szczepanek et al. 1994; Moran et al. 1995). If we assume that the evidence of demonstrated mobility in experimental genetic crosses is an approximation of intron mobility in natural yeast populations, we would expect these mobile introns to have higher rates of turnover than inactive introns. Our results (Table 1), however, show that the seven mobile introns do not have significantly different turnover rates from the remaining introns (*P* = 0.49, Wilcoxon rank test). Although the interpretation can be confounded by the fact that the mobility might not have been exhaustively examined for all 17 introns, it is important to mention that the *cob* i4 intron (a group I intron also known as b13) was demonstrated not to be mobile (Zinn and Butow 1984; Moran et al. 1992) but has a turnover rate higher than five out of the seven mobile introns (Table 1). Furthermore, high-mobility introns do not necessarily have faster turnover rates than low-mobility introns. For instance, the *cox1* i8 intron (a group I intron also known as a15c) has been shown to be less efficient as a mobile element than the omega intron (Moran et al. 1992), but the

Fig. 2. Distribution of mitochondrial introns in the Saccharomycetaceae. Introns with demonstrated mobility in experimental genetic crosses are in black, whereas the remaining introns are in gray. Four introns labeled with asterisks are group II introns, whereas the remaining 13 belong to group I introns. The homologous intron positions in the *Kluyveromyces lactis* could not all be confidently determined; as a result *K. lactis* and *Eremothecium gossypii* were excluded from the intron analyses.
turnover rate of the cox1 i8 intron is about 2.7 times (58.4 vs. 21.7) as that of the omega intron (table 1). These suggest that the intron turnover rate is not necessarily correlated with intron mobility.

Intron homing has long been recognized as an important mechanism driving mitochondrial intron turnover (Goddard and Burt 1999). If intron homing is the sole predominant mechanism of intron turnover, one would expect a combination of fast intron gain, slow intron degeneration and sequential loss. However, when the rates of intron gain and loss are estimated separately, four introns show a significantly higher loss rate than the gain rate, but no introns show a significantly higher gain rate than the loss rate (table 1). These data suggest that the intron turnover process is also governed by factors other than intron homing or retrohoming. The results in table 1 were estimated using the mitochondrial protein–gene tree as the "organismal" tree (fig. 1), which is not completely identical with previously published relationships in yeast, but the analysis on the most reasonable alternative phylogeny (supplementary fig. S1, Supplementary Material online) yielded highly similar results (supplementary table S2, Supplementary Material online).

Phylogenetic Incongruence in Introns
All natural yeast hybrids documented to date involve closely related species within the same genus, for example, within Saccharomyces sensu stricto (Morales and Dujon 2012). To minimize the effect of natural hybridization and identify bona fide HGT events, we took a conservative approach and focused on three well-separated, taxon-rich clades (Lachancea, Saccharomyces sensu stricto, and Torulaspora), so that HGT between these clades can be detected with confidence. We restricted our attention to the nine most abundant introns (i.e., those present in ten or more strains, in aggregate, of the three clades). Given the highly sporadic distribution and frequent gain and loss of these introns, it is intrinsically problematic to root most intron trees in the conventional manner, that is, using outgroups. Therefore, all intron trees are shown midpoint rooted. Five of these nine introns (cox1 i3, cox1 i5, cox1 i6, cox1 i8, and cob i4) show intermixing between clades, with all of these incongruences strongly supported at $P < 0.0001$ (fig. 3 and supplementary tables S3 and S4, Supplementary Material online). Putative HGT events are also evident among taxa within a clade, including cases of conspecific strains that do not form a monophyletic group in seven of the intron trees (fig. 3). Here, we cannot rule out the possibility that these inferred HGT events within a clade might result from natural hybridization, that is, successful interspecific mating. The topologies of all nine examined introns are significantly incongruent, with all $P$ values < 0.05 (supplementary tables S3 and S4, Supplementary Material online) according to the approximately unbiased (AU) test. Furthermore, the intron trees show pronounced branch-length disparities (fig. 3). For example, the Torulaspora and Lachancea clades group closely with cox1 i7 (fig. 3E) cob i1 (fig. 3H), and the protein–gene concatenate (fig. 1), whereas Torulaspora and Saccharomyces group closely with cox1 i9 (fig. 3G). This suggests the occurrence of recent HGT for the cox1 i9 intron.

We detected significant evidence for recombination in most (eight out of nine) examined introns and, in all but one case, by six or all seven recombination-detection methods used (supplementary table S5, Supplementary Material online). One of these cases is illustrated phylogenetically (fig. 4): The five L. kluyveri strains form the expected
monophyletic clade using the last 670 alignment positions of Cox1 i5, but break into two utterly different groups using the first 2,491 alignment positions. Moreover, multiple recombination events, that is, events with different break points were detected within six of the nine introns (data not shown). Such intron chimerism could not be explained by intron homing, but instead should be considered as evidence of homologous-recombination-mediated horizontal transfer.

Phylogenetic Incongruence in Protein Genes
Because our intron analyses revealed extensive intronic horizontal transfer mediated by homologous recombination at different periods of evolutionary time and among strains between clades, which of course can also occur within coding sequences, we searched for evidence of HGT in mitochondrial exons. We examined seven of the eight protein genes found in all yeast mtDNAs (the var1 gene was
excluded due to its extremely high level of sequence variation). Full coding-sequence analyses produced significantly incongruent topologies relative to the organismal tree for at least four of these seven protein genes (supplementary tables S3 and S4, Supplementary Material online). Unlike the whole-intron analyses, no compelling evidence of intermixing between the three color-coded clades (Lachancea, Saccharomyces sensu stricto, and Torulaspora) was found in the full-coding-sequence analyses (data not shown). At this level, one would expect that the incongruence must only occur within one or more of these clades. However, statistically significant recombination events involving between-clade intermixing were detected for four of seven examined protein genes (figs. 5 and 6 and table 2). For example, positions 1051–1142 of the cob gene alignment group Lachancea thermodora with T. globosa and Lachancea mirantina with the Saccharomyces strains (fig. 5A), whereas the rest of this gene supports a monophyletic relationship of each of the three clades analyzed (fig. 5B). One might argue that the region showing phylogenetic incongruence could be a coconversion tract resulted from intron homing of a nearby cob intron. It is important to note here that region 1051–1142 starts 243 nucleotides to the downstream of the last cob intron but before alignment position 1051 were also obtained (fig. 5C).

Fig. 5. Phylogenetic evidence for chimerism of the cob gene. Node support values (likelihood bootstrap/Bayesian posterior probability) are shown when maximum likelihood bootstrap values are > 50%. Different relationships were obtained for alignment positions 1051–1142 (A) and the remaining sequence alignment (B). The relationships of the downstream region of the last cob intron but before alignment position 1051 were also obtained (C). Lachancea strains are in red, Saccharomyces in green, and Torulaspora in blue.

Fig. 6. Alignment of positions 412–459 of the cob gene. The Torulaspora-specific informative sites are highlighted in red.
and instead supports homologous recombination between coding sequences of relatively distantly yeast strains.

Exonic HGT can introduce amino acid changes and potentially affect the function of mitochondrial proteins. For example, a short region of the *L. mirantina* *cox3* gene is identical with those of a three-species *Torulaspora* clade but contains at least four informative nucleotides not present in any other *Lachancea* strains, with two of these nucleotide differences producing a valine (GTA) to isoleucine (ATT) change in *L. mirantina* (fig. 6). Similarly, the HGT event involving positions 1051–1142 of the *cob* gene alignment has also resulted in amino acid alteration (fig. 5A). We suspect that many amino acid changes induced by mitochondrial HGT are functionally neutral or slightly deleterious, and some amino acid changes may have a small, positive effect on fitness and serve as the raw materials for adaptive evolution.

### Discussion

**Mechanistic Factors Promoting Intron Mobility**

The hyperdynamic process of mitochondrial intron gain and loss is in striking contrast to nuclear intron dynamics. The analysis of a much larger dataset of 292 nuclear-encoded splicing introns failed to detect any intron presence–absence polymorphism among 38 *S. cerevisiae* strains (Skelly et al. 2009). One explanation for this marked contrast is the presence of intron-encoded HEGs, which could promote high rates of mitochondrial intron gain and loss. However, introns with demonstrated mobility in experimental genetic crosses (Dujon 1980; Delahodde et al. 1989; Moran et al. 1992) do not all have high rates of intron gain and loss, whereas introns demonstrated to have inactive mobility (Zinn and Butow 1984; Moran et al. 1992) can have high rates of intron gain and loss (table 1). In the intron turnover rate analysis, we used the tree branch length as a relative time scale. One caveat is that the yeast mitochondrial-gene tree is not clock-like and the mtDNA genes in post (nuclear) whole-genome duplication species have undergone accelerated evolution (fig. 1), probably due to relaxed mitochondrial function (Jiang et al. 2008). In the rate estimation, all transition rate matrices are calibrated (please see [Hao and Golding 2010; Kim and Hao 2014] for details), and the estimated intron turnover rate is the average number of transition events per intron site per nucleotide substitution. The strikingly different turnover rates among the introns (ranging from 0.5 to 180 in table 1) can be attributed to the fact that certain introns have different turnover rates in different species. Nevertheless, the overall inference of hyperdynamic mitochondrial intron gain and loss should be considered robust.

Given the fact that a number of the introns are chimeric, we believe that homologous recombination rather than intron mobility (e.g., intron homing or retrohoming) is likely the major mechanism for high rates of mitochondrial intron turnover. Homologous recombination and intron homing/retrohoming, however, are not mutually exclusive. Homologous recombination can take place, perhaps recurrently, after an intron homing or retrohoming event. It is possible that the substantially divergent sequences were initially seeded by intron homing or retrohoming from distantly related strains and then became widely spread by sequential homologous recombination among close relatives.

**Extensive Mitochondrial HGT in Yeast**

To date, widespread mitochondrial-to-mitochondrial HGT has been reported only in certain plants (Berghthorson et al. 2004; Rice et al. 2013; Xi et al. 2013). Our findings provide evidence for frequent HGT in yeast mitochondrial introns and protein genes, including many transfers that affect only part of a gene or intron sequence, creating evolutionarily chimeric genes and introns. We posit that standard phylogenetic analyses using entire genes, much less concatenates of multiple genes, as the unit of analysis will fail to detect many HGT events in yeast mtDNAs. The amount of HGT found in this study should be considered as a conservative estimate, because existing recombination detection programs have limited sensitivity, especially when recombination takes place in short patches, frequently, recurrently, and/or between very similar mtDNA sequences (Hao and Palmer 2009; Hao 2010; Hao et al. 2010; Hao and Palmer 2011). Consistent with the fact that increased sequence similarity promotes higher rates of homologous recombination (Majewski and Cohan 1999; Hao and Palmer 2009), we found relatively many intermixing events between the *Lachancea* and *Torulaspora* clades, the two most closely related clades sequence-wise (but probably not genealogically) among the three clade–clade combinations.

We could not completely rule out the possibility of some false positives for mitochondrial HGT. Recombination could be mistakenly inferred by molecular evolutionary process such as mutation hot spots (Galtier et al. 2006) and evolutionary rate heterogeneity (Sun et al. 2011). In addition, yeast mitochondrial genomes have strong base compositional bias and sometimes different codon assignments (Miranda et al. 2006; Ling et al. 2014), which can potentially confound phylogenetic analysis (Jermyn et al. 2004). In our analysis, most HGT sequences share more similarity with homologs from different clades than those from their native clade, which is
very unlikely to be simply due to standard mutation processes alone. Furthermore, our inferred HGT cases in the protein-coding regions (e.g., the identical cox3 fragment between L. mirantina and Torulaspora) cannot be explained solely by codon usage bias. The codon usage biases were measured and compared among the Lachancea genomes using the CodonO program (Angellotti et al. 2007), but the codon usage in L. mirantina was not significantly different from that in other Lachancea genomes.

Factors Driving Mitochondrial HGT in Yeast

Many strains of Saccharomyces, Lachancea, and Torulaspora share the same habitat, living on the bark of oak trees (Sampaio and Goncalves 2008; Zhang et al. 2010). Diverse yeast strains have long been cocultured intentionally and unintentionally during winemaking and other bioprocesses (Ciani et al. 2010). The shared niche of these diverse yeast strains provides opportunities for physical contact and interspecific genetic interactions, including nonsexual cellular fusion. Indeed, several horizontal transfer cases in nuclear genes have been reported between distantly related winemaking yeasts (Novo et al. 2009; Marsi et al. 2015). Yeast cells also undergo filamentous (pseudohyphal) growth in response to starvation for nitrogen (Gimeno et al. 1992). Hyphal or pseudohyphal fusion can lead to the production of heterokaryons, that is, cytoplasmic fusion with maintenance of separate nuclei (Bever and Wang 2005; Mehrabi et al. 2011). Heterokaryon formation in conjunction with very frequent (i.e., occurring on an organismal time scale) fusion and fission of mitochondria (Shaw and Nunnari 2002) can effectively promote recombination between different mitochondrial haplotypes. Mitochondrial fusion is already proposed to be a crucial force driving mitochondrial HGT in plants (Rice et al. 2013).

In yeast, the mitochondrial genome has been shown to have lower substitution rates than the nuclear genome (Clark-Walker 1991; Freal et al. 2014), with the notable exceptions of some yeast mitochondrial mutator strains (Dimitrov et al. 2009). Low substitution rates can facilitate mitochondrial HGT via both intron invasion and homologous recombination. Low substitution rates can also facilitate intron invasion by preserving the specific recognition sites required for intron homing. As noted above, the frequency of homologous-recombination-mediated HGT is expected to correlate positively with DNA sequence similarity between donor and recipient (Majewski and Cohan 1999). In combination, heterokaryon production, mitochondrial fusion, and low mitochondrial substitution rates are predicted to drive high levels of HGT between mitochondrial genomes in yeasts relative to their nuclear genomes. Comparative studies on rates of mitochondrial and nuclear HGT in yeast are needed to test this prediction.

These yeast results are part of an emerging pattern in mitochondrial evolution, namely, a correlation between low substitution rates and relatively high rates of HGT. Mitochondrial HGT is best known in land plants, whose mtDNAs generally have very low substitution rates (Wolfe et al. 1987). Conversely, higher animal mtDNAs have high substitution rates and lack any evidence for HGT (Shearer et al. 2002; Hellberg 2006; Keeling and Palmer 2008), whereas the several mitochondrial HGTs described in basal animals, for example, corals and sponges, correlate with the unusually low substitution rates in their mtDNAs (Shearer et al. 2002; Schroder et al. 2003; Goddard et al. 2006; Hellberg 2006; Fukami et al. 2007; Szitzenberg et al. 2010).

HGT among Nonyeast Fungal Mitochondrial Genomes

Although the demonstration of extensive mitochondrial–mitochondrial HGT in this study is in yeast, the implication could be extended to the entire fungi kingdom and perhaps beyond. In fact, recent genomic studies have observed a highly sporadic distribution of mitochondrial introns in nonyeast fungal species (Ferandon et al. 2010; Joardar et al. 2012; Mardanov et al. 2014; Zhang et al. 2015), which is remarkably similar to what has been observed in yeast (e.g., in fig. 2). Chimeric introns were also observed in the cox1 gene of the fungus Sclerotinia borealis suggesting mitochondrial–mitochondrial HGT independent of the intron-encoded endonuclease activity, which is compatible with the homologous-recombination-mediated intron exchange observed in yeast (Wu and Hao 2014). Furthermore, protein-coding genes in several fungal species, for example, Allomyces arbusculus, Mortierella verticillata, Rhizopus oryzae, Smittium culisetae, and Glomus sp., have undergone HGT due to the activity of their adjacent endonucleases (Paquin et al. 1994; Seif et al. 2005; Beaudet et al. 2013). Therefore, it may be that fungal mitochondrial genomes are in general subject to extensive mitochondrial–mitochondrial HGT. There are currently at least two obstacles to the discovery of extensive mitochondrial–mitochondrial HGT in nonyeast fungi. 1) There is a lack of complete mitochondrial genome sequences from an abundant number of conspecific strains or related species in nonyeast fungi. 2) Many mitochondrial–mitochondrial HGT events appear to involve fine scale recombination events, which are difficult to detect using existing recombination detection programs and most likely overlooked when using whole genes as the unit of phylogenetic analysis. Our prediction could soon been tested in the rapidly increasing fungal mitochondrial genome sequences from conspecific strains using more sophisticated phylogenetic methodologies.

Horizontal Transfer as a Population Genetic Process

Our results do not show faster intron gain over intron loss, which would be expected in the cyclical model of rapid intron invasion, slow intron degeneration, and loss, followed by reinvansion (Goddard and Burt 1999). Four introns have significantly different rates between intron gain and intron loss, but all four introns show faster intron loss over intron gain (table 1). The observed bias toward intron loss could be explained by the slightly deleterious nature of intron sequences (Lynch 2002). There are fewer mitochondrial HGT events in the protein genes than in the introns, and the detected
interclade mitochondrial HGT events in the protein genes are all at subgenic levels, which could be explained by the fact that the mitochondrial protein genes are under very strong functional constraint (Mamirova et al. 2007; Stewart et al. 2008). The introduction of substantially divergent sequences in the protein-coding region may happen, but many are deleterious and selected against during evolution.

The nuclear and mitochondrial genomes have a tight co-evolutionary relationship (Barreto and Burton 2013), which helps to prevent mito-nuclear incompatibility. For instance, the *Mrs1* gene in *S. cerevisiae* was found to be incompatible with the *cox1* i4 intron (also know as I3f) in *S. paradoxus* (Chou et al. 2010). Horizontal transfer of any mitochondrial sequences incompatible with the recipient nuclear genome would be selected against. Therefore, the observed yeast mtDNA dynamics should be viewed as a snapshot of a population genetic process with evolutionary forces including functional constraint and mito-nuclear interaction acting on the highly recombinogenic and dynamic mtDNAs in yeast.

Materials and Methods

Mitochondrial Genome Sequencing, Assembly, and Annotation

Eight strains of *Torulaspora* from four species were kindly provided by Dr. Matthew Goddard (The University of Auckland) and the National Center of Agricultural Utilization Research (IL). *Saccharomyces paradoxus* YPS138 was purchased from National Collection of Yeast Cultures (NCYC, UK). mtDNA of each strain was extracted from a 2-day culture inoculated from a single colony as in (Defontaine et al. 1991). Eight strains were sequenced at an average coverage of 238–1,194 × by the Illumina MiSeq platform (150 bp PE); assembly used software programs SOAPdenovo (Luo et al. 2012), SPAdes (Bankevich et al. 2012), Velvet (Zerbino and Birney 2008), and Consed (Gordon and Green 2013) with K-mers from 21 to 127. *Torulaspora globosa* CBS764 was sequenced at an average coverage of 258 × by the Ion Torrent PGSM assembly used gsAssembler (Roche, CT). Gaps between contigs were checked and filled using SSPACE (Boetzer et al. 2011) and GapFiller (Boetzer and Pirovano 2012). mtDNAs from nine strains were assembled into single, circularized genomes of length 28–45 kb (supplementary table S1, Supplementary Material online), whereas *T. globosa* CBS764 was assembled into four contigs totaling 37 kb. The genomes were annotated using MFannot (Nadimi et al. 2012), followed by manual correction of intron boundaries. Raw sequence data from *T. delbrueckii* CBS1146 were kindly provided by Dr. Kenneth Wolfe (University College Dublin). These ten mitochondrial genomes have been submitted to GenBank (accession numbers KM595067–KM595076).

Phylogenetic Analyses and Tests for Incongruence

Gene sequences were extracted from the ten newly sequenced mtDNAs and the 30 yeast mtDNAs available in GenBank (supplementary table S1, Supplementary Material online). These were aligned using MUSCLE (Edgar 2004), with the intron sequence alignments further refined using SATe-II (Liu et al. 2012), and all alignments were manually inspected. Phylogenetic trees were constructed using PhyML (Guindon and Gascuel 2003) under a GTR + I + I nucleotide substitution model. Bayesian analysis was performed using MrBayes version 3.2 (Ronquist et al. 2012) under a GTR + I + I nucleotide substitution model for 100,000 generations with trees sampled every 100 generations after a burn-in of 25,000 generations. Since the corresponding nuclear genomes are not all available for the strains used in this study, it is currently impossible to obtain a nuclear-genome-based phylogenomic tree. The concatenated mitochondrial protein–gene tree (fig. 1) is in good agreement with relationships inferred in nuclear-gene-based phylogenetic studies (Kurtzman 2003; Kurtzman and Robnett 2013; Salichos and Rokas 2013) and was used as the organismal tree in this study. We used four outgroup taxa, *Pichia sorbitophila* (Debaryomycetaceae), *Pichia farinosa* (Debaryomycetaceae), *Komagataella pastoris* (Phaffomycetaceae), and *Hanseniaspora uvarum* (Saccharomycodaceae) to root the tree. *H. uvarum*, a basal lineage in the Saccharomycodaceae (Kurtzman 2003), has undergone accelerated evolution (Pramatetaki et al. 2006), which might introduce phylogenetic artifacts (Bergsten 2005; Su et al. 2011). To reduce such a concern, we have performed phylogenetic analyses both including *H. uvarum* and excluding it from the outgroups and obtained the same phylogenetic relationship for the in-group taxa (fig. 1). To be conservative, we have compiled the phylogenetic information of the Saccharomycetaceae from previous studies (Kurtzman 2003; Friedrich et al. 2012; Gaillardin et al. 2012; Jung et al. 2012; Kurtzman and Robnett 2013; Salichos and Rokas 2013; Lang et al. 2014; Ling et al. 2014; Kurtzman and Sugiyama 2015) and generated a best alternative “organismal” tree (supplementary fig. S1, Supplementary Material online). Phylogenetic incongruence of each gene tree relative to the organismal tree (or the alternative organismal tree) was examined by the AU test (Shimodaira 2002) implemented in the CONSEL program (Shimodaira and Hasegawa 2001).

Detection of Homologous Recombination

Recombination events were detected using seven recombination-detection methods (RDP, GENECONV, CHIMAERA, SISCAN, MaxChi, Bootscan, and 3Seq), all implemented in the RDP4 (recombination detection program) program (Martin et al. 2010). Significant recombinant events were further examined for phylogenetic incongruence between sequence regions using the AU test (Shimodaira 2002).

Quantification of Intron Turnover Rates

Gain and loss of introns located at homologous sites were modeled as a two-state continuous-time Markov process, with states 0 (absence) and 1 (presence) mapped on a phylogeny. The rate of intron gain and loss were estimated in the R package DiscML (Kim and Hao 2014) using the tree branch length as a relative time scale. The turnover rate is expressed as the number of gains/losses per site per nucleotide substitution (Hao and Golding 2006, 2010). Intron turnover rates
were estimated using both a simplistic one-parameter model (by constraining gain and loss rates to be the same) and a two-parameter model (gain and loss rates calculated separately).

**Supplementary Material**

Supplementary figure S1 and tables S1–S5 are available at Molecular Biology and Evolution online (http://www.mbe.oxfordjournals.org/).

**Acknowledgments**

The authors thank three anonymous reviewers for their helpful comments; Jeffrey Palmer and Brian Golding for critically reading the manuscript; the National Center of Agricultural Utilization Research and Matthew Goddard for providing the Torulaspora yeast strains; and Kenneth Wolfe for providing sequence data from T. delbrueckii CBS1146. They are grateful for the grid computing service from Computing & Information Technology of Wayne State University. This work was supported by funds from Wayne State University to W.H.

**References**


Lang BF. 1984. The mitochondrial genome of the fission yeast Schizosaccharomyces pombe: highly homologous introns are inserted at the same position of the otherwise less conserved cox1 genes in Schizosaccharomyces pombe and Aspergillus nidulans. EMBO J. 3:2129–2136.


Lazowska J, Szczepanek T, Macadre C, Dokova M. 1992. Two homologous mitochondrial introns from closely related Saccharomyces species differ by only a few amino acid replacements in their Open Reading Frames: one is mobile, the other is not. C R Acad Sci. Ill 315:37–41.


Schafer B, Wolf K. 1999. A novel group-II intron in the *cox1* gene of the fission yeast *Schizosaccharomyces pombe* is inserted in the same codon as the mobile group-I intron al2 in the *Saccharomyces cerevisiae* *cox1* homologue. *Curr Genet.* 35:602–608.


Zinn AR, Butow RA. 1984. Kinetics and intermediates of yeast mitochondrial DNA recombin-