

Impact of ancient oxygen levels on mitochondrial genome evolution inferred by supertrees, supermatrices, and relaxed molecular clocks

L.R. Lopes^{1,2}, R.C. Ferreira³, F. Antoneli¹, P.B. Paiva¹ and M.R.S. Briones^{1,2}

¹ Departamento de Informática em Saúde, Universidade Federal de São Paulo, São Paulo, SP, Brasil

² Departamento de Microbiologia, Imunologia e Parasitologia, Universidade Federal de São Paulo, São Paulo, SP, Brasil

³ Departamento de Neurologia e Neurocirurgia, Universidade Federal de São Paulo, São Paulo, SP, Brasil

Corresponding author: M.R.S. Briones
E-mail: briones.marcelo.rs@gmail.com

Genet. Mol. Res. 20 (2): gmr18485
Received September 06, 2019
Accepted April 22, 2021
Published May 31, 2021
DOI <http://dx.doi.org/10.4238/gmr18485>

ABSTRACT. Evolution of mitochondrial genomes is essential for the adaptation of yeasts to changes in environmental oxygen levels. Although *Saccharomyces cerevisiae* mitochondrial DNA lacks all complex I genes, respiration is possible because alternative NADH dehydrogenases are encoded by *NDE1* and *NDI1* nuclear genes. The apparent whole genome duplication (WGD) in the yeast ancestor 100-150 million years ago caused nuclear gene duplications and secondary losses, although its relation to the loss of mitochondrial complex I is unknown. We produced phylogenomic supertrees and a supermatrix tree of 46 mitochondrial genomes, showing that the loss of complex I predates WGD and occurred independently in the *S. cerevisiae* group and the fission yeast *Schizosaccharomyces pombe*. The branching patterns did not differ substantially in supertrees and supermatrix phylogenies. We found consistent relations between conserved mitochondrial chromosomal gene order (synteny) in closely related yeasts. Correlation of mitochondrial molecular clock estimates and atmospheric oxygen variation in the Phanerozoic suggests that the *Saccharomyces* lineage might have lost complex I during hypoxic periods near Permian-Triassic or Triassic-Jurassic mass extinction

events, while the *Schizosaccharomyces* lineage possibly lost complex I during hypoxic environment periods during the Middle Cambrian until the Lower Devonian. The loss of mitochondrial complex I, as a result of low oxygen levels, might not affect yeast metabolism due to a fermentative switch. The return to increased oxygen periods could have favored adaptations to aerobic metabolism. Additionally, we also show that *NDE1* and *NDII* phylogenies indicate evolutionary convergence in yeasts in which mitochondrial complex I is absent.

Key words: Yeast evolution; Mitochondrial genome; *Saccharomyces cerevisiae*; *Candida albicans*; Mitochondrial complex I

INTRODUCTION

The Ascomycota yeast group includes *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*, both fundamental models for eukaryotic genetics and biotechnology (Botstein et al., 1997; Hoffman et al., 2015). Despite the phylogenetic distance, *Saccharomyces* and *Schizosaccharomyces* spp. share similar adaptations. In both species groups the ability to adapt to varying oxygen concentrations is essential for survival. *Saccharomyces cerevisiae* and *Sz. pombe* readily ferment sugars and produce ethanol in the presence of oxygen, a process known as the Crabtree effect (Crabtree, 1928; Snoek and Steensma, 2006; Dashko et al., 2014), and are also promptly able to sustain anaerobic growth despite higher energetic costs (Lai et al., 2006; Dashko et al., 2014). Therefore, these yeast species are adapted to current oxygen atmospheric levels conditions (21%) and hypoxic environments, in which oxygen concentrations are lower than current atmospheric levels (Postmus et al., 2011a).

During a series of geological periods, the oxygen levels significantly varied and might have exerted strong pressure on various species for the evolution of respiratory mechanisms and adaptations to cope with hyperoxic periods (Carboniferous, Early Permian and Late Cretaceous-Early Tertiary), hypoxic periods (Ordovician, Devonian, Late-Permian and Triassic-Jurassic boundary) and normoxic periods (present) (Berner, 1999; Ward, 2006; Glasspool and Scott, 2010a).

The evolution of respiratory mechanisms in fungi produced patterns with different combinations of fermentation types (Li et al., 2011). For example, *Candida albicans*, the causative agent of candidiasis, the most common fungal infection worldwide, is dependent on respiration, but is able to adapt to hypoxic environments inside its host (Grahl et al., 2012). *Candida albicans* has a complete gene set of the mitochondrial respiratory chain (Cavalheiro et al., 2004). On the other hand, *S. cerevisiae* and *Sz. pombe* lack all seven mitochondrial genes encoding respiratory complex I subunits, but they are adapted to oxygenic conditions (Bullerwell et al., 2003; Skoneczna et al., 2015).

In *S. cerevisiae*, the absence of mitochondrial respiratory complex I is compensated by *NDE1*, *NDE2* and *NDII* nuclear gene products, although these proteins are not evolutionarily related to any classical complex I subunits (Small and McAlister-Henn, 1998; Luttik et al., 1998; Overkamp et al., 2000; Li et al., 2006). *NDE1*, *NDE2* and *NDII* proteins act by transferring electrons to ubiquinone, as an alternative NADH dehydrogenase (Luttik et al., 1998; Postmus et al., 2011b). Therefore, *NDE1*, *NDE2* and *NDII* are essential for *S. cerevisiae* viability (Melo et al., 2004).

The loss of complex I genes promotes gene order variability from *Saccharomycetes* and *Schizosaccharomyces* spp. mitochondrial genomes (Aguileta et al., 2014). Fungal mitochondrial gene order is relatively free to vary, especially within the *Saccharomycetes* (Aguileta et al., 2014). As expected, conservation of gene order in mitochondrial and nuclear genomes is higher in closely related species (Keogh et al., 1998). While the loss of complex I genes is associated with a lack of conserved mitochondrial gene order, the WGD is associated with a lack of gene order in the yeast nuclear genome (Wolfe, 2015). There is evidence that a WGD event happened in the yeast ancestor about 100-150 million years ago (Wolfe and Shields, 1997). This genomic event doubled the dose of each nuclear gene per cell, causing gene duplications, secondary losses, neofunctionalization and the origin of ohnologs (Kellis et al., 2004; Sémon and Wolfe, 2007).

In our study we investigated the mitochondrial phylogenomics of yeasts and whether their ability to adapt to varying oxygen concentrations could be correlated with the evolutionary history of Saccharomycetales mitochondrial genomes during the Phanerozoic (from 541 Myr ago to present). The great diversification of these fungi, which occurred in this geologic eon, paralleled major fluctuations in atmospheric O₂ and CO₂ concentrations (Berner, 1999; Ward, 2006; Mentel and Martin, 2008; Glasspool and Scott, 2010a). For this we combined phylogenomics with quantitative synteny analysis and molecular clock methods to estimate the dates of the loss of mitochondrial complex I in the *Saccharomyces* and *Schizosaccharomyces* lineages.

MATERIAL AND METHODS

Sequence Data

The 39 complete yeast mitochondrial genome sequences, and all other individual gene sequences, were downloaded from GenBank (<http://ncbi.nlm.nih.gov/Genbank>) and the Candida Genome Database (<http://www.candidagenome.org>), with accession numbers available in [Supplementary Table 1](#). Individual mitochondrial gene sequences were analyzed separately. The set of mitochondrial genes including apocytochrome b (*COB*), ATP synthetase subunits (*ATP6*, *ATP8*, and *ATP9*) and cytochrome oxidase subunits (*COX1*, *COX2*, and *COX3*), conserved in genomes of all yeast species, were included in this study. Another set of (variable) genes, not present in all mitochondrial genomes, is composed by NADH dehydrogenase (ubiquinone) complex I subunits (*NAD1*, *NAD2*, *NAD3*, *NAD4*, *NAD4L*, *NAD5* and *NAD6*). Additionally, large (LSU) and small (SSU) ribosomal RNA subunits genes were also analyzed. Sequences were aligned, individually, using Clustal W (Larkin et al., 2007), with default settings, and manually adjusted using Seaview software (Galtier et al., 1996).

Inference of Source Trees

To infer phylogenetic Bayesian trees, substitution models were determined for each dataset based on the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), using MEGA 6 software (“find best DNA model” option) (Tamura et al., 2013). Depending on the gene, we selected either general time-reversible (GTR) model (Rodríguez et al., 1990) with Γ distribution to correct for rate variation among sites

(GTR+G) or GTR with Γ distribution and proportion of invariant sites, (GTR+G+I) (Yang, 1994) (the complete parameter list is available in [Supplementary Table 2](#)).

All the mitochondrial source trees and the nuclear phylogenies based on 18S rRNA, *NDE1* and *NDII* genes were inferred using MrBayes 3.2 software (Ronquist et al., 2012). The gene sequences were downloaded from GenBank according to the availability. Only 37 nucleotide sequences of 18S rRNA gene from the species set here considered were available, while few species had *NDE1* and *NDII* sequences available (see [Supplementary Table 1](#)). Bayesian trees were searched for 10^7 generations with sampling every 100 generations until the standard deviation from split frequencies were under 0.01. The parameters and the trees were summarized by wasting 25% of the samples obtained (burn-in). The consensus trees were then used to determine the posterior probabilities. All phylogenetic trees were then formatted with the FigTree v1.3.1 software (<http://tree.bio.ed.ac.uk/software/figtree/>).

Inference of Supertrees and a Supermatrix

The 16 mitochondrial source trees ([Supplementary Figures 1-16](#)) were combined as a comprehensive Robinson-Foulds (RF) supertree and as a Matrix Representation Parsimony (MRP) supertree. The mitochondrial RF supertree was inferred using RF-supertree software (<http://genome.cs.iastate.edu/CBL/RFsupertrees>) and CompPHY software (<http://www.atgc-montpellier.fr/compiphy/>) was used to infer MRP supertree (Fiorini et al., 2014). Both supertree methods consider even the phylogenetic source trees with low frequency species overlap to generate a supertree containing all species. The analysis included genomes with varying gene contents and three species (*Lachancea waltii*, *Saccharomyces bayanus* and *Zygosaccharomyces rouxii*) represented only by 15S rRNA and *COX2* genes, because their mitochondrial genomes were not available.

The RF supertree provides a binary supertree, by heuristic search, that minimizes the sum of the Robinson-Foulds distances (defined as a normalized count of the symmetric difference between the sets of clusters of the two trees) between every rooted source tree and the supertree (Bansal et al., 2010). Forty-four optimal tree outputs were generated by RF-supertree software with a default setting. All the optima supertrees were combined into a single consensus tree using Geneious v6.1.6 software. MRP has been the method of choice for most of supertree analyses. The MRP method, independently developed by Baum (1992) and Ragan (1992), consists in constructing a matrix where each taxon is coded according to the presence in a clade (“1” for presence, “0” for absence in the clade and “?” when absent in the source tree). Based on the constructed matrix, the most parsimonious tree was built. CompPHY software provided an MRP supertree that was edited in Figtree v1.3.1.

The software based on RF and MRP can estimate supertrees that are consistent with the largest number of clusters (or clades) from the source trees. The approach can consider even the phylogenetic source trees with low frequency species overlap to generate a supertree containing all species. Species included in this study, such as *Sz. pombe*, for example, which have a low number of genes in the mitochondrial genome, were not included in all source trees, but was included in the supertrees. Even those species that were represented only by 15S rRNA and *COX2* gene were included in supertrees.

The concatenated align sequences to construct the supermatrix phylogenetic tree included a 16 mitochondrial gene dataset (22.6 kilobases). Alignments from mitochondrial genes were concatenated using Seaview software. Best-fit models of molecular evolution for concatenated alignment were selected based on AIC and BIC using MEGA 6 software (Lanfear et al., 2012) ([Supplementary Table 2](#)). A Bayesian tree with partitioning schemes was inferred based on concatenated alignment using MrBayes 3.2 software with 10^7 generations, sampling every 100 generations, until the split frequencies standard deviations were under 0.01. The selected trees were summarized by wasting 25% of the samples obtained (burn-in) and consensus trees were then used to determine the posterior probability values.

Robinson-Foulds distances

To infer congruency between two phylogenetic trees, we used RF distance. The symmetric difference between the sets of clusters of the two phylogenetic trees can be defined as a normalized count by RF distance, as score value. The most congruent tree comparison has a low RF-distance, low score. Using CompPHY software (Fiorini et al., 2014) to provide RF distance scores, we compared the congruency between supertrees, supermatrix phylogeny and nuclear 18S rRNA.

Mitochondrial Genome Synteny Analysis

The synteny (or gene order) map was constructed using Geneious v6.1.6 to verify gene positions and transcriptional directions. The graphic with gene order and gene presence/absence was constructed using Microsoft Excel and Libre Office Draw. All yeast mitochondrial genomes were delineated with all genes linearly depicted as arrow boxes indicating forward (right) or reverse (left) sense. Genomes were manually aligned, using 15S rRNA as starting point, and progressively assembled blocks of genes that exhibit conserved features across the genome. Mitochondrial genomes were pairwise compared using the progressive Mauve software (Darling et al., 2010) to search for regions of local synteny (colinear blocks), counted as units of conserved blocks. The nonparametric Mann-Whitney test, using GraphPad InStat software, was used to infer that the synteny mean values are distinct among the clades. NADH gene remnants were searched, using BLAST (Altschul et al., 1990), within the nuclear genome of the yeasts *S. cerevisiae*, *Sz. pombe* and *C. glabrata*. The search was optimized using the megablast option, to find highly similar sequences, and the blastn option, for sequences that were somewhat similar.

Molecular Clock Tests and Divergence Date Estimates

For divergence date estimates, we used BEAST v2.3.1 software (Drummond et al., 2012). The null hypothesis of a strict global clock was rejected for all mitochondrial genes, by the likelihood ratio test for $\alpha=0.05$ (Felsenstein, 1981). The likelihood ratio for molecular clocks was calculated based on 2Δ , the difference between the likelihood scores ($\text{Ln}L_0 - \text{Ln}L_1$) of the trees, considering that $\text{Ln}L_0$ represents the likelihood from the molecular clock tree and $\text{Ln}L_1$ represents the likelihood from the non-clock tree. Molecular clocks only allow a null hypothesis model. A non-molecular clock allows alternative

hypotheses. In testing the molecular clock hypothesis, the degrees of freedom work out to be $s-2$, where “ s ” is the number of taxa in the phylogeny (Felsenstein, 1981). Substitution models were selected based on AIC and BIC using MEGA 6 software for each mitochondrial gene. We implemented a Yule speciation process prior to this. Bayesian coestimation of topology and divergence time were obtained using a uniform prior analysis for calibration ages. The divergence dates for the mrca priors of Ascomycota (811.57Myr and 690.55Myr) and Saccharomycetales (473.33Myr and 648.32Myr), obtained from the Padovan et al. (2005) study, were used as calibration points. BEAST analyses were run for 10 million generations, logging parameters and trees every 1000 generations.

RESULTS AND DISCUSSION

After the alignment of 39 mitochondrial complete yeast genomes ([Supplementary Table 1](#)) the coding domains sequences of 16 genes were extracted and used to infer unrooted Bayesian phylogenies for each separate gene. These trees (or source trees), available as [Supplementary Figures 1-16](#), were used as inputs for the inferences of the Robinson-Foulds (RF) supertree (Bansal et al., 2010) and the Matrix Representation Parsimony (MRP) supertree (Baum, 1992; Ragan, 1992), in Figure 1 (and [Supplementary Figures 17 and 18](#)). The RF supertree is the tree that minimizes the RF distances between each individual source tree and considers the presence-absence of species in different source trees. This supertree is estimated by the heuristic tree-bisection reconnection, where the RF distances are the number of operations necessary to convert one topology to another (Bansal et al., 2010). On the other hand, the MRP supertree is based on the constructed matrix, where each taxon is coded according to the presence in a clade (“1”), absence in a clade (“0”), or absent in the source trees (“?”). Thus, the most parsimonious tree built represents a supertree (Sanderson et al., 1998; Bininda-Emonds, 2004).

Supertree methods have been broadly applied in phylogenetic studies, aimed at the combination of phylogenetic trees based on different types of data, joining molecular and morphology information (Liu et al., 2001; Creevey and McInerney, 2005; Beck et al., 2006; Bininda-Emonds et al., 2007; Ruta et al., 2007; Nyakatura and Bininda-Emonds, 2012; Chang et al., 2013; Davis and Page, 2014; Sigwart and Lindberg, 2015; Duda and Jan Zrzavý, 2016), or exclusively molecular data (Daubin et al., 2001; Fitzpatrick et al., 2006; Higdon et al., 2007; Campbell and Lapointe, 2010; Swenson et al., 2011; Modica et al., 2011). More specifically, mitochondrial gene datasets were used to construct supertrees (Fitzpatrick et al., 2006; Higdon et al., 2007; Campbell and Lapointe, 2010).

In Figure 1, we observe that MRP and RF supertrees included two groups, which correspond to the four nuclear (Nc) and mitochondrial (Mt) genetic code patterns observed: Group 1 (blue group) has Nc=1 (standard) (Abramczyk et al., 2003) and Mt=3 (yeast mitochondrial) (Clark-Walker and Weiller, 1994) and Group 2 (yellow group) has Nc=1 and 12 (alternative yeast nuclear genetic code) (Ohama et al., 1993) and Mt=4. All species from the blue group are included in Saccharomycetacea family plus *Wickerhamomyces canadensis*, which belongs to the Wickerhamomycetaceae family. This group also includes the species that were described as WGD positive species (in purple). The two supertree methods supported similar topologies, but there were few differences in the blue group, in the phylogenetic position of pre-WGD and post-WGD species.

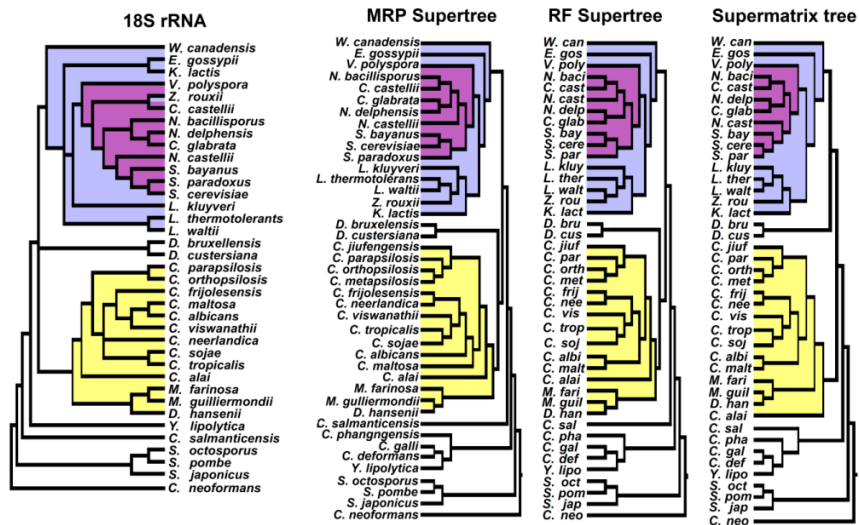


Figure 1. Multigene phylogenetic analysis, supermatrix phylogenetic tree and supertrees. Supertrees were inferred from 16 source trees, each corresponding to all Saccharomycota mitochondrial genes excepting the tRNA genes. The phylogenetic tree based on supermatrix was constructed using concatenated alignment of mitochondrial gene sequences. Only three node labels had posterior probabilities values inferior to 1. RF Supertree was inferred using the Robinson-Foulds algorithm as implemented in the RF program (Bansal et al. 2010). Supertree based on Matrix Representation Parsimony algorithm (Baum 1992; Ragan 1992), MRP Supertree, was constructed using CompPHY software (Fiorini et al. 2014). Phylogenetic tree based on 18S rRNA was the only tree inferred using single gene sequence. The blue group includes pre and pos-WGD species. WGD was highlighted in purple. The yellow group includes most *Candida* species, the CTG clade.

Supertree methodology has been considered a robust approach to analyze broad phylogenies (Daubin et al., 2001; Qian and Zhang, 2016), but the difficulty to carry out statistical analysis is considered challenging in supertree methodology (von Haeseler, 2012). A phylogenetic approach based on Bayesian trees could overcome this disadvantage in supertree methodology. Bayesian inference allows analysis of large datasets and incorporates full models of sequence evolution and is not restricted to a unique best tree (Padovan et al., 2005). We tried to compensate the lack of statistical support by inferring the Bayesian source trees with the best gene substitution models (Supplementary Table 2), including a large number of tree generations and considering significant statistical support measures to combine source trees into RF and MRP supertrees. Each source tree, obtained by MrBayes, reached estimated sample size (ESS) values above 200 (they should be at least 100), and reached potential scale reduction factor (PSRF) values equal or above to 1.0 (the recommendation is close to 1.0). Furthermore, 87.5% of posterior probabilities values of source trees were equal or greater than 0.75 (Supplementary Figures 1-16). Posterior probability represents the probability that a clade is true given the data, the model, and the priors (Larget and Simon, 1999).

In addition to supertrees, we made a supermatrix phylogenetic tree (Figure 1). We applied the yeast mitochondrial genome sequences dataset, the same dataset used in supertrees, to infer a phylogenomic tree based on concatenated alignment (also called a supermatrix). Different from supertree methodologies, the supermatrix phylogenetic tree was not constructed by source trees combinations. Instead, this phylogenetic tree was

constructed based on concatenated alignments (head to tail linkage) from 16 mitochondrial genes sequences. By supermatrix phylogenetic methodology, we constructed a Bayesian phylogenetic tree that allowed us to infer satisfactory statistical support values (ESS and PSRF). Mitochondrial supermatrix phylogenetic tree also presented consistent posterior probability values. Only three nodes had support values inferior to 1 (see [Supplementary Figure 19](#)).

The supermatrix phylogenetic tree also conserved clades presented in RF and MRP supertrees. Differently from the supertrees, the supermatrix phylogenetic tree placed *Millerozyma farinosa*, *Meyerozyma guilliermondii* and *Debaryomyces hansenii* nearer the other CTG clade species than *C. alai*, similar to what was found in a previous study (Valach et al., 2011). The *Dekkera* genus is the first branch out from the blue group, while in the supertrees, the *Dekkera* genus is the first branch out from the yellow group.

Despite some minor topology divergences between supertrees and the supermatrix tree, the pattern of branching trees did not differ dramatically (Figure 1). Most clades were conserved even in supertrees as in the supermatrix. As in the study by Nyakatura and Bininda-Emonds (2012), we did find significant differences between supertree and supermatrix topologies (Nyakatura and Bininda-Emonds, 2012), even considering higher accuracy from the supermatrix method comparing with supertrees (Kupczok et al., 2010).

Besides multi-gene mitochondrial phylogeny, we also included 18S rRNA phylogenetic tree (Figure 1 and Figure S20), representing nuclear evolution hypothesis. Mitochondrial genome evolves considerably faster than nuclear genome due strong selective pressure under mitochondria (Brown et al., 1979; Saccone et al., 2000). Consequently, mitochondrial genomes from different yeast species present variable size and structure and accumulate higher mutation frequencies than nuclear genomes (Saccone et al., 2000; Freel et al., 2015). Nuclear 18S rRNA phylogeny has the same three groups identified in the mitochondrial phylogenies, but with additional differences along the tree (inconclusive place of *Dekkera*, inclusion of *Z. rouxii* into WGD group and polytomy into CTG clade, for example).

To compare differences into topologies we applied Robinson-Foulds distance method. This method quantifies topological differences between two phylogenetic trees. Comparison between more congruent trees results in low RF distance. Based on RF distance score, supertrees (MRP and RF) phylogeny are the more congruent trees (scores = 8, Table 1). When nuclear 18S rRNA based tree is compared with supertrees or supermatrix the scores of RF distances are higher than comparisons between mitochondrial trees.

Table 1. Comparison of Robinson Foulds distance scores of supertrees inferred by MRP, RF, Supermatrix tree and the 18S rRNA tree.

	Robinson-Foulds distance scores			18S rRNA
	MRP	RF	Supermatrix	
MRP Supertree	NA	8	12	37
RF Supertree	NA	NA	16	31
Supermatrix tree	NA	NA	NA	31
18S rRNA	NA	NA	NA	NA

To ratify the phylogenetic results and support the topology of conserved groups placed on the mitochondrial phylogenetic trees, we investigated the gene order of mitochondrial genomes from yeast species comparing with phylogeny (Figure 2). MRP supertree (most congruent with other trees - Table 1) was compared with mitochondrial gene order map which revealed that the phylogeny correspond to synteny patterns. Fungal mitochondrial genomes have

highly variable size between their species, with mtDNA ranging from 19kbp in *Sz. pombe* to 100 kbp *S. cerevisiae*. Therefore, co-localization of the block of the genes into genome was quantified, using Mauve software, and is remarkably similar between closely related yeast species (Table 2 and Figure 3). The pairwise matrix comparison confirmed higher synteny between species within the same clade than synteny between species from different clades, supported by nonparametric Mann-Whitney (Table 3).

Comparison between gene order map and phylogeny allowed us to identify ancestral point from the mitochondrial complex I loss (Figure 2). This comparison also showed that the absence of mitochondrial complex I occurred independently in *Saccharomyces* lineage and the *Schizosaccharomyces* genus, also presented by previous studies (de Zamaroczy and Bernardi, 1986; Nosek and Fukuhara, 1994; Bullerwell et al., 2003; Schäfer, 2003; Procházka et al., 2010). After the divergence between *W. canadensis* and the common ancestral of yeast species that composed blue group (see Figure 2) occurred an event (or a sequence of events) that promoted the loss of complex I. The loss of mitochondrial complex I in the *Saccharomyces* lineage leads us to suggest that complex I might, at least in part, affected genome stability. Nuclear genome instability may be resulted from a failure of various mitochondrial functions, such as an electron transport chain activity breakdown, for example, which changed ATP dynamic production (Kaniak-Golik and Skoneczna, 2015). Functional mitochondria are crucial to maintaining the stability of cellular genetic information (Skoneczna et al., 2015). This requirement is not limited to the proteins and RNAs encoded by the mitochondrial genome or proteins whose expression needs tight feedback control. Mitochondrial dysfunction can results in increased reactive oxygen species (ROS) production that, consequently, increases frequencies of mutations and lesions in the mitochondrial genome and may lead the loss of parts of mitochondrial DNA or all mitochondrial genome (Rasmussen et al., 2003; Yazgan and Krebs, 2012). So, mutations and deletions may have affected the mitochondrial genome and, consequently, impaired the stability of the nuclear genome (Veatch et al., 2009).

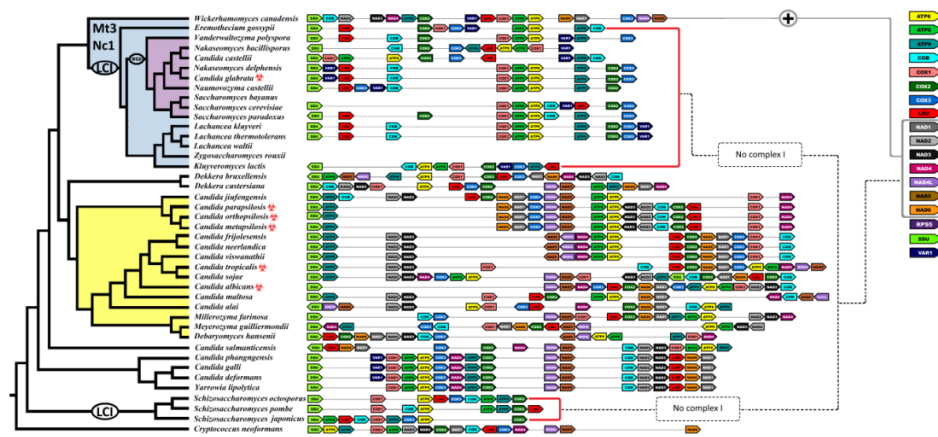


Figure 2. MRP supertree, most congruent tree, represents mitochondrial phylogeny. Mitochondrial gene order map of every yeast specie is represented by colored boxes with direction according to the coding sense. The blue group, (*Saccharomyces* lineage), has nuclear code 1 and mitochondrial code 3. WGD occurred in ancestors from the blue group, highlighted as purple clade. The gene order map indicates that the blue group has no genes encoding mitochondrial complex I (NADH dehydrogenase subunits), except for *W. canadensis*. This is evidence that the loss of complex I occurred after the divergence between *W. canadensis* and *Saccharomyces* lineage. The yellow group contains only representatives of genus *Candida*, including *C. albicans*, with nuclear code 1 and 12 and mitochondrial 4. Biohazard symbols indicate pathogenic yeasts. LCI = Loss of complex I, Mt3 = mitochondrial translational code 3 and Nc1 = nuclear translational code 1.

Table 2. Quantitative synteny of yeast mitochondrial genomes. The co-localization of the block of the genes into genome was quantified, using Mauve software, revealing strong similarities between closely related yeast species. The correspondence of abbreviations and complete species names are in [Supplementary Table S1](#).

	<i>Wc</i>	<i>Eg</i>	<i>Vp</i>	<i>Nb</i>	<i>Cc</i>	<i>Nd</i>	<i>Cg</i>	<i>Nc</i>	<i>Sb</i>	<i>Sp</i>	<i>Sc</i>	<i>Lk</i>	<i>Lt</i>	<i>Lw</i>	<i>Zr</i>	<i>Kl</i>	<i>Cj</i>	<i>Cp</i>	<i>Co</i>	<i>Cm</i>	<i>Cf</i>	<i>Cn</i>	<i>Cv</i>	<i>Ct</i>	<i>Cs</i>	<i>Ca</i>	<i>Cm</i>	<i>Ca</i>	<i>Mf</i>	<i>Cg</i>	<i>Dh</i>	
<i>W. canadensis</i>	0	5	8	8	5	6	6	7	NA	3	4	5	8	NA	NA	10	7	6	8	8	3	3	3	5	3	7	7	7	9	8	10	
<i>E. gossypii</i>		0	6	10	11	8	8	8	NA	8	9	5	6	NA	NA	7	3	6	6	6	1	1	1	3	3	1	5	5	5	5	4	
<i>V. polyspora</i>			0	10	9	10	9	8	NA	8	11	7	5	NA	NA	10	1	5	4	5	3	3	4	1	4	3	4	4	5	6	5	
<i>N. bacillisporus</i>				0	7	8	12	12	NA	11	10	9	13	NA	NA	6	4	3	4	3	4	4	4	4	1	2	3	3	3	4	3	
<i>C. castellii</i>					0	9	8	9	NA	6	6	6	12	NA	NA	9	3	2	4	4	4	4	4	2	4	2	2	2	2	5	5	
<i>N. delphensis</i>						0	4	7	NA	12	6	4	5	NA	NA	9	4	3	3	3	4	4	4	5	4	5	5	3	4	6	2	
<i>C. glabrata</i>							0	6	NA	12	11	4	4	NA	NA	9	4	3	3	4	5	5	2	4	4	5	5	5	4	6	2	
<i>N. castellii</i>								0	0	7	6	5	5	NA	NA	9	1	3	5	1	3	3	4	4	3	4	4	3	2	6	1	
<i>S. paradoxus</i>									0	4	10	7	NA	NA	10	3	3	3	5	3	3	2	3	3	2	5	1	4	2	2		
<i>S. cerevisiae</i>										0	10	9	NA	NA	11	3	4	5	4	3	3	3	3	3	3	3	3	4	1	4	3	
<i>L. kluyveri</i>											0	4	NA	NA	8	4	3	4	4	4	4	4	4	3	3	3	5	4	4	1		
<i>L. thermotolerans</i>												0	0	0	11	4	3	5	3	2	2	2	4	3	3	3	4	5	4	6	5	
<i>K. lactis</i>															0	3	4	4	6	4	4	4	3	4	2	6	4	5	4	4		
<i>C. juufengensis</i>																0	11	12	12	10	11	10	8	7	10	11	8	5	8	11		
<i>C. parapsilosis</i>																	0	12	12	12	11	13	8	8	10	13	10	8	9	8		
<i>C. orthopsilosis</i>																		0	12	11	11	10	7	8	10	13	11	11	9	4		
<i>C. metapsilosis</i>																			0	14	11	10	11	8	10	14	8	10	9	8		
<i>C. frijolesensis</i>																			0	9	8	4	5	11	6	8	8	8	11			
<i>C. neerlandica</i>																					0	8	4	5	9	12	8	6	12	11		
<i>C. viswanathii</i>																						0	4	5	9	10	10	4	12	7		
<i>C. tropicalis</i>																							0	6	9	13	5	6	10	7		
<i>C. sojae</i>																								0	5	9	3	4	4	4		
<i>C. albicans</i>																									0	13	8	10	10	8		
<i>C. maltosa</i>																									0	5	8	8	11			
<i>C. alai</i>																									0	8	9	9				
<i>M. farinosa</i>																										0	10	10				
<i>M. guilliermondii</i>																										0	9					
<i>D. hansenii</i>																											0					

Table 3. Statistics of intra-clade and inter-clade synteny. The pairwise matrix comparison confirmed the inter-species and inter-clade synteny as supported by the nonparametric Mann-Whitney test.

Synteny means (standard deviation)		Difference means	Mann-Whitney test p-value
Clade 1	Clade 1 and 2		
7.88 (2.38)	3.75 (1.53)	4.13	>0.0001
Clade 2	Clade 2 and 3		
8.77 (2.63)	3.75 (1.53)	5.02	>0.0001

The absences of mitochondrial complex I from *S. cerevisiae* and *Sz. pombe* were documented by previous studies (Lang and Wolf, 1984; de Zamaroczy and Bernardi, 1986; Nosek and Fukuhara, 1994; Bullerwell et al., 2003; Procházka et al., 2010). However, when the loss of complex I happened remained unknown. Moreover, whether *Schizosaccharomyces* and *Saccharomyces* lineages lose the complex I in the same episode (or not) was not explored yet. For while, estimated divergence of *S. cerevisiae* and *Sz. pombe* is widely ranged between 200 Myr until 1000 Myr, and the divergence between *Candida albicans* and *S. cerevisiae* is estimated to be 871 Myr, by multiprotein analysis and 526.95 Myr depending on the calibration of 18S rRNA phylogeny (Heckman et al., 2001; Padovan et al., 2005). The mitochondrial supertrees and phylogenetic tree based on supermatrix indicate that the loss of mitochondrial complex I occurred after the divergence of *W. canadensis* from the *Saccharomyces* lineage, certainly older than 150-100 Myr ago, the estimate for the WGD (Wolfe and Shields, 1997). To estimate the date of complex I loss and WGD we used molecular clock analysis based on mitochondrial genes *COB*, *ATP*

synthetase subunits (*ATP6*, *ATP8*, and *ATP9*), cytochrome oxidase subunits (*COX1*, *COX2* and *COX3*) and LSU and SSU ribosomal RNA genes. We also included nuclear 18S rRNA to estimate loss of mitochondrial complex I.

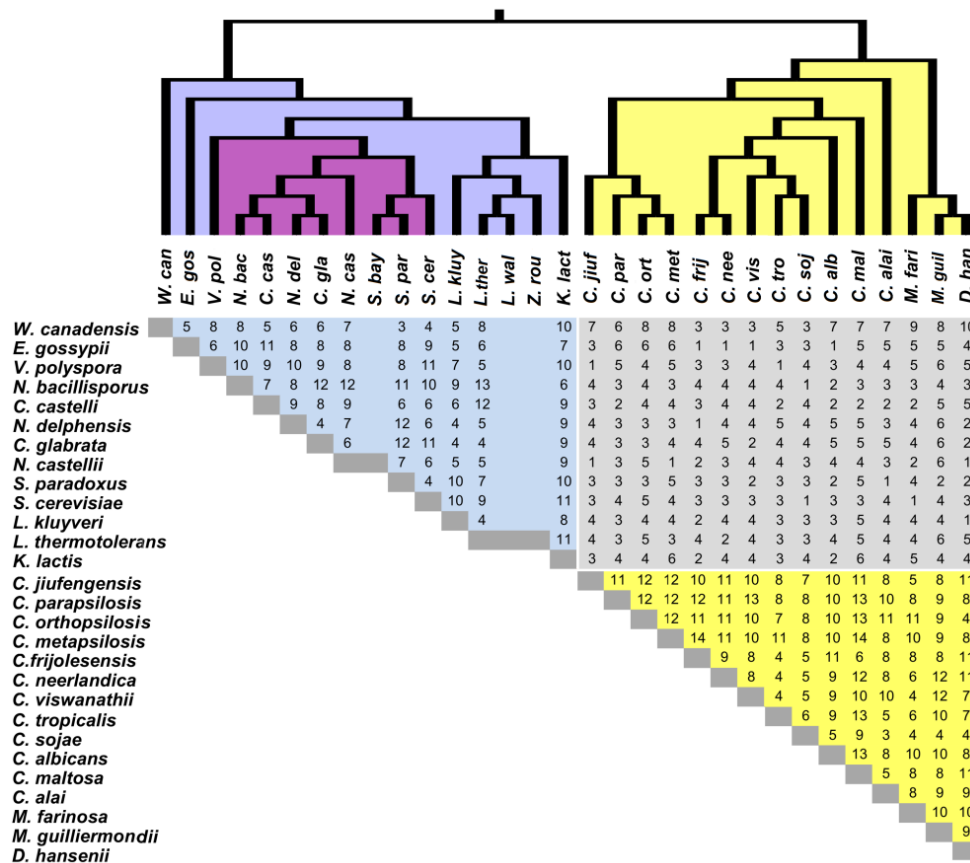


Figure 3. Pairwise similarity of colinear blocks in mitochondrial genomes aligned by Mauve progressive algorithm. Regions of local colinearity colinear blocks were counted as one conserved block unit. Colored areas (yellow and blue) indicate comparisons within the same clade whereas gray area indicate comparisons between clades. Means of syntenic block numbers of each clade are higher than the mean number of conserved blocks between different clades. The nonparametric Mann-Whitney test was used to estimate the significance levels between comparisons. The correspondence of species names abbreviations and complete names is on [Supplementary Table S1](#).

Based in our dataset genes, the likelihood ratio test for molecular clock rejected the null hypothesis of a strict global clock by comparison of the likelihoods of the clock constrained tree versus the non-clock unconstrained tree (see [Supplementary Table 3](#) for details). This approach is considered more realistic and focused on estimate periods with variable rates (Wilke et al., 2009).

Defined the model of molecular clock, we aimed to estimate when the loss of mitochondrial complex I occurred. As indicated in the phylogenetic analysis (Figure 2), the loss of mitochondrial complex I occurred after the divergence between *W. canadensis* with

other species from blue group. Additionally, we also estimated loss of complex I from *Schizosaccharomyces* genus. Concomitantly, WGD, previously dated between 150 Myr to 100 Myr (Wolfe and Shields, 1997), was reanalyzed with our dataset.

For molecular clock analysis, we adopted the confidence intervals (95% HPDI - highest posterior density interval) as main result, considering the low reliability of mean ages. Confidence intervals may provide more likelihood dates, even when intervals are broad (Warnock et al., 2017).

In agreement with phylogenetic and synteny results, the molecular clock indicated that the loss of mitochondrial complex I by *Schizosaccharomyces* and *Saccharomyces* occurred in distinct events (Table 4 and Figure 4).

Table 4. Divergence date estimates with 95% highest posterior density interval (HPDI) of different mitochondrial genes using relaxed clocks. WGD = Whole Genome Duplication and Myr = Million years.

Genes	WGD		Loss of Complex I Saccharomycetales		Loss of Complex I Schizosaccharomyces		Substitution rates (per site/million years)	
	Mean	HPDI Myr	Mean	HPDI Myr	Mean	HPDI Myr	Mean	HPDI
<i>ATP6</i>	200.757	220.214-181.220	244.378	263.399-222.732	370.376	409.057-330.160	7.691x10 ⁻⁴	7.029x10 ⁻⁴ -8.360x10 ⁻⁴
<i>ATP8</i>	124.916	134.899-108.413	182.256	220.351-156.003	198.341	265.832-136.831	6.851x10 ⁻⁴	5.947x10 ⁻⁴ -7.798x10 ⁻⁴
<i>ATP9</i>	207.527	246.297-166.663	232.788	266.790-197.769	176.280	219.947-129.551	4.117x10 ⁻⁴	3.493x10 ⁻⁴ -4.769x10 ⁻⁴
<i>COB</i>	153.429	168.732-139.203	214.812	225.690-197.227	440.386	479.850-402.595	6.561x10 ⁻⁴	5.871x10 ⁻⁴ -7.276x10 ⁻⁴
<i>COX1</i>	187.299	203.042-172.470	267.252	286.792-248.078	348.480	381.190-314.141	4.310x10 ⁻⁴	3.899x10 ⁻⁴ -4.735x10 ⁻⁴
<i>COX2</i>	163.946	205.566-121.007	220.777	239.120-190.512	381.309	424.094-335.067	4.526x10 ⁻⁴	3.998x10 ⁻⁴ -5.062x10 ⁻⁴
<i>COX3</i>	181.173	194.882-167.538	205.087	207.997-199.635	457.189	496.991-419.014	5.903x10 ⁻⁴	5.582x10 ⁻⁴ -6.218x10 ⁻⁴
<i>SSU</i>	154.761	164.799-142.544	210.812	226.494-192.894	470.171	507.809-420.295	1.014x10 ⁻³	9.143x10 ⁻⁴ -1.113x10 ⁻³
<i>LSU</i>	198.511	210.636-189.250	250.513	260.160-235.935	427.513	459.401-403.040	1.106x10 ⁻³	1.026x10 ⁻³ -1.189x10 ⁻³
<i>18S</i>	105.014	124.952-085.653	167.333	196.996-138.736	352.561	418.701-281.318	1.448x10 ⁻⁴	1.223x10 ⁻⁴ -1.687x10 ⁻⁴

*18S = 18S rRNA

Considering the confidence intervals based on nine mitochondrial genes, most of them (seven genes – *ATP6*, *COB*, *COX1-3*, *SSU* and *LSU* rRNA, see Table 4 and Figure 4A), dated the loss of complex I by *Schizosaccharomyces* in a period which comprehends the middle of Cambrian until the middle of Carboniferous (Figure 4B). The estimates for the loss of complex I by *Schizosaccharomyces* based on *COB* (479.850 - 402.595 Myr), *COX3* (496.991 - 419.014 Myr), *LSU* (507.809 - 420.295 Myr) and *SSU* (459.401 - 403.040 Myr) encompass the Late Ordovician mass extinction event (Figure 4B). This event, dated between 458 and 443 Myr, probably occurred due variation of climatic conditions, characterized by a high CO₂ concentration, hypoxia and glaciation (Raup and Sepkoski, 1982; Berner, 1999; Sheehan, 2001; Finnegan et al., 2011; Pohl et al., 2014). The Late Ordovician extinction event eliminated around 25% of the families and nearly 60% of the genera of marine organisms (Wake and Vredenburg, 2008). This scenario seems to provide conditions for the loss of a mitochondrial complex I due low atmospheric oxygen availability. Alternatively, the Late Devonian mass extinction event (around 364 Myr), was encompassed in the estimate for loss of complex I by *Schizosaccharomyces* lineage, based on *ATP6* (409.057 - 330.160 Myr), *COX1* (381.190 - 314.141 Myr) and *COX2* (424.094 - 335.067 Myr). Late Devonian mass extinction event might have offered significant pressure on microbial community, as discussed above, due hypoxic ocean and atmospheric environments (Berner, 2006; McGhee, 2012). This event, which a bolide impact as a cause controversial (Racki, 2012), shifted the temperature down and decreased oxygen levels, causing major mass extinction of marine life, around 57% of marine genera (Ward, 2006; Wake and Vredenburg, 2008).

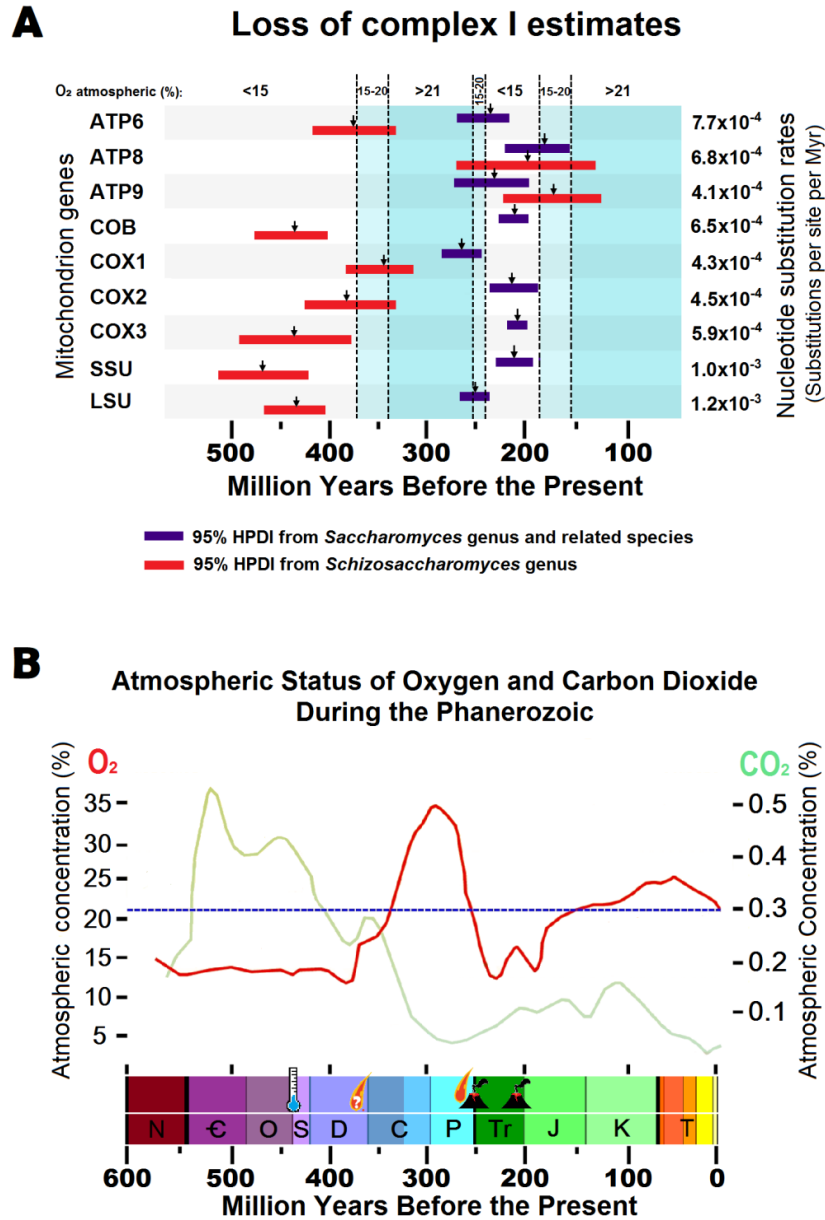


Figure 4. In (A) loss of mitochondrial complex I during the Phanerozoic as estimated by lognormal relaxed molecular clocks and (B) oxygen and carbon dioxide levels and mass extinction events from the Phanerozoic. In (A), red bars indicate 95% HPDI from estimate of the loss of mitochondrial complex I in *Schizosaccharomyces* lineage and blue bars in *Saccharomyces* lineage. Vertical divisions in blue represent oxygenic atmospheric level. In (B) data were compiled from (Berner 1999; Huey and Ward 2005; Ward 2006; Glasspool and Scott 2010b). The O₂ levels are indicated by red lines and CO₂ levels by green lines. The five major events extinction events are depicted by symbols representing their putative triggers. The geological periods depicted at the bottom are: Cambrian (C), Ordovician (O), Silurian (S), Devonian (D), Carboniferous (C), Permian (P), Triassic (Tr), Jurassic (J), Cretaceous (K) and Tertiary (T).

According to estimated ages based on *ATP8* (265.832 - 136.831 Myr) and *ATP9* (265.832 - 136.831 Myr) genes, mitochondrial complex I was lost by *Schizosaccharomyces* more recently, in a period that comprehends the end of Permian up to Jurassic, the last great hypoxic period (Table 4 and Figure 4). This period estimated includes two mass extinction events: Permian–Triassic extinction (252 Myr) and End Triassic mass extinction (199 – 214 Myr). The Permian–Triassic mass extinction episode, occurred around 252 Myr, was considered the worst of the five mass extinctions, due 95% of all species, including marine and terrestrial, were eradicated (Huey and Ward, 2005; Ward, 2006; Benton, 2015). The event was, possibly, caused by a bolide impact that initiated a volcanic flood emanating from the Siberian Traps (Kaiho et al., 2001; Wake and Vredenburg, 2008; Benton, 2015). Consequently, there was a severe decline of oxygen (Huey and Ward, 2005). The hypoxic environment persisted along million years until the next mass extinction event, in the End Triassic. During Triassic–Jurassic (Tr–J) boundary (Figure 4B), around 200 Myr ago, occurred the Central Atlantic Magmatic Province (CAMP) volcanic eruptions, one of the strongest volcanic episodes in Earth's history (Pálffy et al., 2000; Hesselbo et al., 2002). Emission of voluminous volcanic gases lead to increased CO₂ and decreased O₂ and coincided with climatic crisis and abrupt decline of marine and terrestrial flora and fauna (Hesselbo et al., 2002), leading to the loss of the whole class of conodonts (Stanley, 2009), and the extinction of 48% invertebrate species (Keller, 2005).

While most of our estimates indicate that *Schizosaccharomyces* may have lost the complex I in the first half of Phanerozoic, the loss of mitochondrial complex I by *Saccharomyces* was dated exclusively in second half, between Middle Permian until Middle Jurassic periods (Table 4 and Figure 4). As well as discussed above, this hypoxic period coincides with End Permian mass extinction and End Triassic mass extinction events. The persistent oxygen low levels associated with events that provide abrupt changes in climate and environment may have favored the loss of respiratory complex I. During hypoxia, our results allow us to suppose that the complex I genes were lost as an economical molecular strategy. Declined oxygen levels may have been consumed even when mitochondrial respiratory chain from yeasts became defective. Without complex I to transfer electrons using NAD cofactor, succinate dehydrogenase, the mitochondrial complex II, transfers electrons reducing FAD cofactor, promoting the execution of respiratory chain (Rutter et al., 2010). However, decreased oxygen levels induces production of ROS (Pitkanen and Robinson, 1996) and, consequently, contribute to damages and instability on mitochondrial and nuclear genomes (Rasmussen et al., 2003; Stuart et al., 2006; Kaniak-Golik and Skoneczna, 2015).

The absence of complex I could have affected yeasts mitochondrial-nuclear genomic stability. Could WGD event have occurred as result of nuclear instability? Two estimates for WGD, based on *ATP9* (246.297 - 166.663 Myr) and *COX2* (205.566 - 121.007 Myr) genes, overlapped with the estimates of the loss of complex I by *Saccharomyces*, *ATP9* (266.790 - 197.769 Myr) and *COX2* (239.12 - 190.512 Myr). From these results we hypothesize that WGD was a consequence of mitochondrial instability caused by the loss of complex I. On the other hand, the most of WGD estimates are not overlapped with loss of complex I event, although some estimate pointed both events as subsequent episodes (*ATP6* and *COX3*). Our WGD estimate based on nuclear 18S rRNA (134.899 - 108.413 Myr) is compatible with the WGD estimate of 150 to 100 Myr (Wolfe and Shields, 1997). *ATP8*, *COB*, *COX2* and SSU mitochondrial genes also comprehends part of the classical estimate from Wolfe and Shields. Therefore, based on molecular clock estimate, WGD occurred after the loss of complex I by *Saccharomyces*.

Hyperoxic or normoxic conditions might have alleviated the negative selection on respiratory systems of yeasts, which by this time had a fermentative metabolism well developed

since the availability of land plant fruits (Dashko et al., 2014). *S. cerevisiae* obtain energy preferentially by fermentation, and they do not exhibit the Pasteur effect, where fermentation is inhibited by oxygen. Therefore, complex I defective mutants in the *S. cerevisiae* lineage under hypoxia, where fermentation is preferred, would have lost the complex I. With the subsequent increase on oxygen levels, the adaptations for respiration might have pressured these organisms to use the nuclear *NDE1* and *NDI1* genes to replace the absent to classical complex I to start the respiratory chain (Luttik et al., 1998; Postmus et al., 2011b). Despite the evolutive distance between complex I negative species *Schizosaccharomyces pombe* and *S. cerevisiae* (Melo et al., 2004), the *NDE1* and *NDI1* genes have a pattern of convergent evolution, as observed by the comparison of *NDE1* and *NDI1* genealogies (Figures 5A and 5C). Differently than 18S rRNA phylogeny (Figures 5B and 5D), *Sz. pombe* and *S. cerevisiae* were placed closest on phylogenetic trees based on *NDE1* and *NDI1*. *NDE1* and *NDI1* genes are not related to the original mitochondrion encoded NADH genes and are present in complex I positive and complex I negative lineages. This can be explained because in complex I negative lineages (*Schizosaccharomyces* and *Saccharomyces*) the *NDI* and *NDE1* genes must complement the function of missing mitochondrial encoded complex I genes. This functional complementation is however not necessary in complex I positive lineages. Other Saccharomycetales subgroups that don't have strong fermentative metabolism might have devised other strategies to cope with hypoxia that could not allow for the loss of the complex I, as is the case of *Candida albicans* (Rozpedowska et al., 2011). These might include increased respiratory rates and uncoupling (Cavalheiro et al., 2004). In the *S. cerevisiae* pre-WGD clade it is important to notice that *Kluyveromyces lactis* does not have the Crabtree effect and cannot use ethanol (Liti et al., 2001). NADH gene remnants were not observed in complex I negative mitochondrial DNA (data not shown) which is consistent with a scenario of ancient loss of complex I, which occurred most likely by deletions of inactive genes.

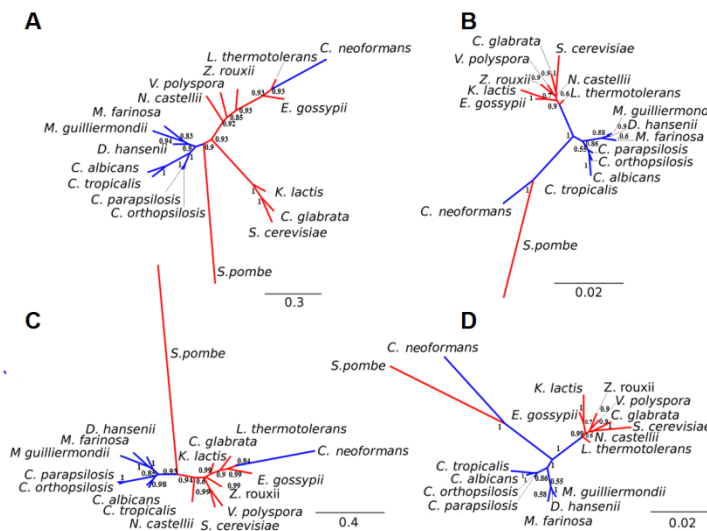


Figure 5. Gene tree of *NDI1* (A) and corresponding 18S rRNA 1 phylogeny (B) and gene tree of *NDE1* (C) and the corresponding 18S rRNA phylogeny (D) in *Saccharomycetales*. In red are the complex I negative taxa and blue the complex I positive taxa. In both *NDI1* and *NDE1* phylogenies the *Schizosaccharomyces* branch is more closely related to other complex I negative clades as compared to the 18S rRNA trees where *Schizosaccharomyces* is distantly related to *S. cerevisiae*. This can be explained because in complex I negative lineages (*Schizosaccharomyces* and *Saccharomyces*) the *NDI* and *NDE1* genes must substitute the function of missing mitochondrial encoded complex I genes. This functional complementation is however not necessary in complex I positive lineages (e.g., *C. albicans*).

CONCLUSIONS

The mitochondrial phylogenomics, synteny analysis and molecular clock estimates here presented confirm that the loss of complex I occurred before, or maybe subsequently to the WGD, and independently in *S. cerevisiae* and *Schizosaccharomyces*. Correlations of estimated dates and variations of Phanerozoic atmospheric oxygen concentrations suggest that the loss of complex I by yeast may have occurred during hypoxic periods, possibly associated with mass extinction events. The *Saccharomyces* lineage might have lost the mitochondrial complex I thanks to hypoxia caused by the End of Permian or the End of Triassic mass extinction events, while the *Schizosaccharomyces* lineage might have lost the mitochondrial complex I associated with Late Ordovician or Late Devonian mass extinction events. Both periods had low oxygen environments, in which negative selection was probably, alleviated on complex I defective mutants under hypoxia as fermentation would have been preferred. The loss of complex I and its pressure for necessary substitution by nuclear gene products during the increased oxygen periods might have, at least in part, contributed to genome instabilities associated with WGD as in the case of *S. cerevisiae* and its relatives. We conclude that supertrees and relaxed molecular clocks suggest losses of yeast mitochondrial complex I in Late Devonian and Triassic-Jurassic boundary.

ACKNOWLEDGMENTS

This work was supported by grants to M.R.S.B from Fundação de Amparo à Pesquisa do Estado de São Paulo – FAPESP, Brazil (2013/07838-0) and Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq, Brazil (303905/2013-1).

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

REFERENCES

- Abramczyk D, Tchorzewski M and Grankowski N (2003). Non-AUG translation initiation of mRNA encoding acidic ribosomal P2A protein in *Candida albicans*. *Yeast* 20: 1045-1052. doi: 10.1002/yea.1020.
- Aguileta G, Vienne DMD, Ross ON, Hood ME, et al. (2014). High Variability of Mitochondrial Gene Order among Fungi. *Genome Biol Evol* 6:451–465. doi: 10.1093/gbe/evu028.
- Altschul SF, Gish W, Miller W, Myers EW, et al. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215: 403-410. doi: 10.1016/S0022-2836(05)80360-2.
- Bansal MS, Burleigh JG, Eulenstein O and Fernández-Baca D (2010). Robinson-Foulds supertrees. *Algorithms Mol. Biol.* 5: 18. doi: 10.1186/1748-7188-5-18.
- Baum BR (1992). Combining Trees as a Way of Combining Data Sets for Phylogenetic Inference, and the Desirability of Combining Gene Trees. *Taxon.* 41: 3-10. doi: 10.2307/1222480.
- Beck RM, Bininda-Emonds OR, Cardillo M, Liu F-GR, et al. (2006). A higher-level MRP supertree of placental mammals. *BMC Evol. Biol.* 6: 93. doi: 10.1186/1471-2148-6-93.
- Benton MJ (2015). When Life Nearly Died: The Greatest Mass Extinction of All Time. Thames & Hudson.
- Berner RA (1999). Atmospheric oxygen over Phanerozoic time. *Proc. Natl. Acad. Sci.* 96: 10955-10957. doi: 10.1073/pnas.96.20.10955.
- Berner RA (2006). GEOCARBSULF: A combined model for Phanerozoic atmospheric O₂ and CO₂. *Geochim Cosmochim Acta.* 70: 5653-5664. doi: 10.1016/j.gca.2005.11.032.
- Bininda-Emonds ORP (2004). The evolution of supertrees. *Trends Ecol Evol.* 19: 315-322. doi: 10.1016/j.tree.2004.03.015.

- Bininda-Emonds ORP, Cardillo M, Jones KE, MacPhee RDE, Beck et al. (2007). The delayed rise of present-day mammals. *Nature*. 446: 507-512. doi: 10.1038/nature05634.
- Botstein D, Chervitz SA and Cherry JM (1997). Yeast as a Model Organism. *Science*. 277: 1259-1260.
- Brown WM, George M and Wilson AC (1979). Rapid evolution of animal mitochondrial DNA. *Proc. Natl. Acad. Sci. USA*. 76: 1967-1971.
- Bullerwell CE, Leigh J, Forget L and Lang BF (2003). A comparison of three fission yeast mitochondrial genomes. *Nucleic Acids Res*. 31: 759-768.
- Campbell V and Lapointe F-J (2010). An Application of Supertree Methods to Mammalian Mitogenomic Sequences. *Evol. Bioinforma Online*. 6: 57-71.
- Cavalheiro RA, Fortes F, Borecký J, Faustinoni VC, et al. (2004). Respiration, oxidative phosphorylation, and uncoupling protein in *Candida albicans*. *Braz. J. Med. Biol. Res*. 37: 1455-1461. doi: 10.1590/S0100-879X2004001000003.
- Chang W-C, Górecki P and Eulenstein O (2013). Exact solutions for species tree inference from discordant gene trees. *J. Bioinform. Comput. Biol*. 11: 1342005. doi: 10.1142/S0219720013420055.
- Clark-Walker GD and Weiller GF (1994). The structure of the small mitochondrial DNA of *Kluyveromyces thermotolerans* is likely to reflect the ancestral gene order in fungi. *J. Mol. Evol*. 38: 593-601. doi: 10.1007/BF00175879.
- Crabtree HG (1928). The carbohydrate metabolism of certain pathological overgrowths. *Biochem. J*. 22: 1289-1298. doi: 10.1042/bj0221289.
- Creevey CJ and McInerney JO (2005). Clann: investigating phylogenetic information through supertree analyses. *Bioinforma Oxf. Engl*. 21: 390-392. doi: 10.1093/bioinformatics/bti020.
- Darling AE, Mau B and Perna NT (2010). progressiveMauve: Multiple Genome Alignment with Gene Gain, Loss and Rearrangement. *PLoS one*. 5: e11147. doi: 10.1371/journal.pone.0011147.
- Dashko S, Zhou N, Compagno C and Piškur J (2014). Why, when, and how did yeast evolve alcoholic fermentation? *FEMS Yeast Res*. 14: 826-832. doi: 10.1111/1567-1364.12161.
- Daubin V, Gouy M and Perrière G (2001). Bacterial molecular phylogeny using supertree approach. *Genome Inform. Int. Conf. Genome Inform*. 12: 155-164.
- Davis KE and Page RDM (2014). Reweaving the Tapestry: a Supertree of Birds. *PLoS Curr*. doi: 10.1371/currents.tol.c1af68dda7c999ed9f1e4b2d2df7a08e.
- de Zamaroczy M and Bernardi G (1986). The primary structure of the mitochondrial genome of *Saccharomyces cerevisiae* — a review. *Gene*. 47: 155-177. doi: 10.1016/0378-1119(86)90060-0.
- Drummond AJ, Suchard MA, Xie D and Rambaut A (2012). Bayesian Phylogenetics with BEAUti and the BEAST 1.7. *Mol. Biol. Evol*. 29: 1969-1973. doi: 10.1093/molbev/mss075.
- Duda P and Jan Zrzavý (2016). Human population history revealed by a supertree approach. *Sci. Rep*. doi: 10.1038/srep29890.
- Felsenstein J (1981). Evolutionary trees from DNA sequences: a maximum likelihood approach. *J. Mol. Evol*. 17: 368-376.
- Finnegan S, Bergmann K, Eiler JM, Jones DS, et al. (2011). The magnitude and duration of Late Ordovician-Early Silurian glaciation. *Science*. 331: 903-906. doi: 10.1126/science.1200803.
- Fiorini N, Lefort V, Chevenet F, Berry V and Chifolleau A-MA (2014). CompPhy: a web-based collaborative platform for comparing phylogenies. *BMC Evol. Biol*. 14: 253. doi: 10.1186/s12862-014-0253-5.
- Fitzpatrick DA, Logue ME, Stajich JE and Butler G (2006). A fungal phylogeny based on 42 complete genomes derived from supertree and combined gene analysis. *BMC Evol. Biol*. 6: 99. doi: 10.1186/1471-2148-6-99.
- Freel KC, Friedrich A and Schacherer J (2015). Mitochondrial genome evolution in yeasts: an all-encompassing view. *FEMS Yeast Res*. doi: 10.1093/femsyr/fov023.
- Galtier N, Gouy M and Gautier C (1996). SEAVIEW and PHYLO_WIN: two graphic tools for sequence alignment and molecular phylogeny. *Comput. Appl. Biosci*. 12: 543-548.
- Glasspool IJ and Scott AC (2010a). Phanerozoic concentrations of atmospheric oxygen reconstructed from sedimentary charcoal. *Nat Geosci* 3: 627-630. doi: 10.1038/ngeo923.
- Glasspool IJ and Scott AC (2010b). Phanerozoic concentrations of atmospheric oxygen reconstructed from sedimentary charcoal. *Nat. Geosci*. 3: 627-630. doi: 10.1038/ngeo923.
- Grahl N, Shepardson KM, Chung D and Cramer RA (2012). Hypoxia and Fungal Pathogenesis: To Air or Not To Air? *Eukaryot Cell*. 11: 560-570. doi: 10.1128/EC.00031-12.
- Heckman DS, Geiser DM, Eidell BR, Stauffer RL, et al. (2001). Molecular Evidence for the Early Colonization of Land by Fungi and Plants. *Science*. 293: 1129-1133. doi: 10.1126/science.1061457.
- Hesselbo SP, Robinson SA, Surlyk F and Piasecki S (2002). Terrestrial and marine extinction at the Triassic-Jurassic boundary synchronized with major carbon-cycle perturbation: A link to initiation of massive volcanism? *Geology*. 30: 251-254. doi: 10.1130/0091-7613(2002)030<0251:TAMEAT>2.0.CO;2.
- Higdon JW, Bininda-Emonds OR, Beck RM and Ferguson SH (2007). Phylogeny and divergence of the pinnipeds (Carnivora: Mammalia) assessed using a multigene dataset. *BMC Evol. Biol*. 7: 216. doi: 10.1186/1471-2148-7-216.

- Hoffman CS, Wood V and Fantes PA (2015). An Ancient Yeast for Young Geneticists: A Primer on the *Schizosaccharomyces pombe* Model System. *Genetics*. 201: 403-423. doi: 10.1534/genetics.115.181503.
- Huey RB and Ward PD (2005). Hypoxia, global warming, and terrestrial late Permian extinctions. *Science*. 308: 398-401. doi: 10.1126/science.1108019.
- Kaiho K, Kajiwara Y, Nakano T, Miura Y, et al. (2001). End-Permian catastrophe by a bolide impact: Evidence of a gigantic release of sulfur from the mantle. *Geology*. 29: 815-818. doi: 10.1130/0091-7613(2001)029<0815:EPCBAB>2.0.CO;2.
- Kaniak-Golik A and Skoneczna A (2015). Mitochondria-nucleus network for genome stability. *Free Radic. Biol. Med.* 82: 73-104. doi: 10.1016/j.freeradbiomed.2015.01.013.
- Keller G (2005). Impacts, volcanism and mass extinction: random coincidence or cause and effect? *Aust. J. Earth Sci.* 52: 725-757. doi: 10.1080/08120090500170393.
- Kellis M, Birren BW and Lander ES (2004). Proof and evolutionary analysis of ancient genome duplication in the yeast *Saccharomyces cerevisiae*. *Nature*. 428: 617-624. doi: 10.1038/nature02424.
- Keogh RS, Seoghe C and Wolfe KH (1998). Evolution of gene order and chromosome number in *Saccharomyces*, *Kluyveromyces* and related fungi. *Yeast Chichester Engl.* 14: 443-457. doi: 10.1002/(SICI)1097-0061(19980330)14:5<443::AID-YEA243>3.0.CO;2-L.
- Kupczok A, Schmidt HA and von Haeseler A (2010). Accuracy of phylogeny reconstruction methods combining overlapping gene data sets. *Algorithms Mol. Biol.* 5: 37. doi: 10.1186/1748-7188-5-37.
- Lai L-C, Kosorukoff AL, Burke PV and Kwast KE (2006). Metabolic-State-Dependent Remodeling of the Transcriptome in Response to Anoxia and Subsequent Reoxygenation in *Saccharomyces cerevisiae*. *Eukaryot Cell* 5: 1468-1489. doi: 10.1128/EC.00107-06.
- Lanfear R, Calcott B, Ho SYW and Guindon S (2012). Partitionfinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Mol. Biol. Evol.* 29: 1695-1701. doi: 10.1093/molbev/mss020.
- Lang BF and Wolf K (1984). The mitochondrial genome of the fission yeast *Schizosaccharomyces pombe*. 2. Localization of genes by interspecific hybridization in strain ade7-50h- and cloning of the genome in small fragments. *Mol. Gen. Genet.* 196: 465-472.
- Larget B and Simon DL (1999) Markov Chasin Monte Carlo Algorithms for the Bayesian Analysis of Phylogenetic Trees. *Mol. Biol. Evol.* 16: 750-750. doi: 10.1093/oxfordjournals.molbev.a026160.
- Larkin MA, Blackshields G, Brown NP, Chenna R, et al. (2007). Clustal W and Clustal X version 2.0. *Bioinforma Oxf. Engl.* 23: 2947-2948. doi: 10.1093/bioinformatics/btm404.
- Li D, Chen H, Florentino A, Alex D, et al. (2011). Enzymatic dysfunction of mitochondrial complex I of the *Candida albicans* goal mutant is associated with increased reactive oxidants and cell death. *Eukaryot Cell*. 10: 672-682. doi: 10.1128/EC.00303-10.
- Liti G, Wardrop FR, Cardinali G, Martini A, et al. (2001). Differential responses to antimycin A and expressions of the Crabtree effect in selected *Kluyveromyces* spp. *Ann. Microbiol.* 51: 235-243.
- Liu FG, Miyamoto MM, Freire NP, Ong PQ, et al (2001). Molecular and morphological supertrees for eutherian (placental) mammals. *Science*. 291: 1786-1789. doi: 10.1126/science.1056346.
- Luttik MA, Overkamp KM, Kötter P, de Vries S, et al (1998). The *Saccharomyces cerevisiae* NDE1 and NDE2 genes encode separate mitochondrial NADH dehydrogenases catalyzing the oxidation of cytosolic NADH. *J. Biol. Chem.* 273: 24529-24534.
- McGhee GR (2012). Extinction: Late Devonian Mass Extinction. In eLS. (Ed.) doi: 10.1002/9780470015902.a0001653.pub3.
- Melo AMP, Bandejas TM and Teixeira M (2004). New Insights into Type II NAD(P)H:Quinone Oxidoreductases. *Microbiol. Mol. Biol. Rev.* 68: 603-616. doi: 10.1128/MMBR.68.4.603-616.2004.
- Mentel M and Martin W (2008). Energy metabolism among eukaryotic anaerobes in light of Proterozoic ocean chemistry. *Philos Trans. R. Soc. B. Biol. Sci.* 363: 2717-2729. doi: 10.1098/rstb.2008.0031.
- Modica MV, Bouchet P, Cruaud C, Utge J et al (2011). Molecular phylogeny of the nutmeg shells (Neogastropoda, Cancellariidae). *Mol. Phylogenet. Evol.* 59: 685-697. doi: 10.1016/j.ympev.2011.03.022.
- Nosek J and Fukuhara H (1994). NADH dehydrogenase subunit genes in the mitochondrial DNA of yeasts. *J. Bacteriol.* 176: 5622-5630.
- Nyakatura K and Bininda-Emonds OR (2012). Updating the evolutionary history of Carnivora (Mammalia): a new species-level supertree complete with divergence time estimates. *BMC Biol.* 10: 12. doi: 10.1186/1741-7007-10-12.
- Ohama T, Suzuki T, Mori M, Osawa S, et al (1993). Non-universal decoding of the leucine codon CUG in several *Candida* species. *Nucleic Acids Res.* 21: 4039-4045.
- Padovan ACB, Sanson GFO, Brunstein A and Briones MRS (2005). Fungi evolution revisited: application of the penalized likelihood method to a Bayesian fungal phylogeny provides a new perspective on phylogenetic relationships and divergence dates of Ascomycota groups. *J. Mol. Evol.* 60: 726-735. doi: 10.1007/s00239-004-0164-y.
- Pálffy J, Mortensen JK, Carter ES, Smith PL, et al. (2000). Timing the end-Triassic mass extinction: First on land, then in the sea? *Geology*. 28: 39-42. doi: 10.1130/0091-7613(2000)28<39:TTEMEF>2.0.CO;2.

- Pitkanen S and Robinson BH (1996). Mitochondrial complex I deficiency leads to increased production of superoxide radicals and induction of superoxide dismutase. *J. Clin. Invest.* 98: 345-351.
- Pohl A, Donnadieu Y, Le Hir G, Buoncristiani J-F, et al. (2014). Effect of the Ordovician paleogeography on the (in)stability of the climate. *Clim. Past.* 10: 2053-2066. doi: 10.5194/cp-10-2053-2014.
- Postmus J, Tuzun I, Bekker M, Müller WH, et al. (2011a). Dynamic regulation of mitochondrial respiratory chain efficiency in *Saccharomyces cerevisiae*. *Microbiol. Read Engl.* 157: 3500-3511. doi: 10.1099/mic.0.050039-0.
- Procházková E, Poláková S, Piškur J and Sulo P (2010). Mitochondrial genome from the facultative anaerobe and petite-positive yeast *Dekkera bruxellensis* contains the NADH dehydrogenase subunit genes. *FEMS Yeast Res.* 10: 545-557. doi: 10.1111/j.1567-1364.2010.00644.x.
- Qian H and Zhang J (2016). Are phylogenies derived from family-level supertrees robust for studies on macroecological patterns along environmental gradients?: Phylogenetic resolution and community ecology. *J. Syst. Evol.* 54: 29-36. doi: 10.1111/jse.12161.
- Racki G (2012). The Alvarez Impact Theory of Mass Extinction; Limits to its Applicability and the “Great Expectations Syndrome.” *Acta Palaeontol Pol.* 57: 681-702. doi: 10.4202/app.2011.0058.
- Ragan MA (1992). Phylogenetic inference based on matrix representation of trees. *Mol. Phylogenet Evol.* 1: 53-58.
- Rasmussen AK, Chatterjee A, Rasmussen LJ and Singh KK (2003). Mitochondria-mediated nuclear mutator phenotype in *Saccharomyces cerevisiae*. *Nucleic Acids Res.* 31: 3909-3917.
- Raup DM and Sepkoski JJ (1982). Mass extinctions in the marine fossil record. *Science.* 215: 1501-1503. doi: 10.1126/science.215.4539.1501.
- Rodríguez F, Oliver JL, Marín A and Medina JR (1990). The general stochastic model of nucleotide substitution. *J. Theor. Biol.* 142: 485-501.
- Ronquist F, Teslenko M, van der Mark P, Ayres DL, et al. (2012). MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* 61: 539-542. doi: 10.1093/sysbio/sys029.
- Rozpędowska E, Galafassi S, Johansson L, Hagman A, et al. (2011). *Candida albicans*—a pre-whole genome duplication yeast — is predominantly aerobic and a poor ethanol producer. *FEMS Yeast Res.* 11: 285-291. doi: 10.1111/j.1567-1364.2010.00715.x.
- Ruta M, Pisani D, Lloyd GT and Benton MJ (2007). A supertree of Temnospondyli: cladogenetic patterns in the most species-rich group of early tetrapods. *Proc. R. Soc. B. Biol. Sci.* 274: 3087-3095. doi: 10.1098/rspb.2007.1250.
- Rutter J, Winge DR and Schiffman JD (2010). Succinate Dehydrogenase—Assembly, Regulation and Role in Human Disease. *Mitochondrion.* 10: 393-401. doi: 10.1016/j.mito.2010.03.001.
- Saccone C, Gissi C, Lanave C, Larizza A, et al (2000). Evolution of the mitochondrial genetic system: an overview. *Gene.* 261: 153-159.
- Sanderson MJ, Purvis A and Henze C (1998). Phylogenetic supertrees: Assembling the trees of life. *Trends Ecol. Evol.* 13: 105-109. doi: 10.1016/S0169-5347(97)01242-1.
- Schäfer B (2003). Genetic conservation versus variability in mitochondria: the architecture of the mitochondrial genome in the petite-negative yeast *Schizosaccharomyces pombe*. *Curr. Genet.* 43: 311-326. doi: 10.1007/s00294-003-0404-5.
- Sémon M and Wolfe KH (2007). Consequences of genome duplication. *Curr. Opin. Genet. Dev.* 17: 505-512. doi: 10.1016/j.gde.2007.09.007.
- Sheehan PM (2001) The Late Ordovician Mass Extinction. *Annu Rev. Earth Planet Sci.* 29: 331-364. doi: 10.1146/annurev.earth.29.1.331.
- Sigwart JD and Lindberg DR (2015). Consensus and confusion in molluscan trees: evaluating morphological and molecular phylogenies. *Syst. Biol.* 64: 384-395. doi: 10.1093/sysbio/syu105.
- Skoneczna A, Kaniak A and Skoneczny M (2015). Genetic instability in budding and fission yeast—sources and mechanisms. *FEMS Microbiol. Rev.* 39: 917-967. doi: 10.1093/femsre/fuv028.
- Snoek ISI and Steensma HY (2006). Why does *Kluyveromyces lactis* not grow under anaerobic conditions? Comparison of essential anaerobic genes of *Saccharomyces cerevisiae* with the *Kluyveromyces lactis* genome. *FEMS Yeast Res.* 6: 393-403. doi: 10.1111/j.1567-1364.2005.00007.x.
- Stanley SM (2009). Evidence from ammonoids and conodonts for multiple Early Triassic mass extinctions. *Proc. Natl. Acad. Sci.* 106: 15264-15267. doi: 10.1073/pnas.0907992106.
- Stuart GR, Santos JH, Strand MK, Van Houten B, et al. (2006). Mitochondrial and nuclear DNA defects in *Saccharomyces cerevisiae* with mutations in DNA polymerase γ associated with progressive external ophthalmoplegia. *Hum. Mol. Genet.* 15: 363-374. doi: 10.1093/hmg/ddi454.
- Swenson MS, Suri R, Linder CR and Warnow T (2011). An experimental study of Quartets MaxCut and other supertree methods. *Algorithms Mol. Biol.* 6: 7. doi: 10.1186/1748-7188-6-7.
- Tamura K, Stecher G, Peterson D, Filipinski A, et al. (2013). MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol. Biol. Evol.* 30: 2725-2729. doi: 10.1093/molbev/mst197.
- Valach M, Farkas Z, Fricova D, Kovac J, et al. (2011). Evolution of linear chromosomes and multipartite genomes in yeast mitochondria. *Nucleic Acids Res.* 39: 4202-4219. doi: 10.1093/nar/gkq1345.
- Veatch JR, McMurray MA, Nelson ZW and Gottschling DE (2009). Mitochondrial dysfunction leads to nuclear genome instability via an iron-sulfur cluster defect. *Cell.* 137: 1247-1258. doi: 10.1016/j.cell.2009.04.014.

- von Haeseler A (2012). Do we still need supertrees? *BMC Biol.* 10: 13. doi: 10.1186/1741-7007-10-13.
- Wake DB and Vredenburg VT (2008). Are we in the midst of the sixth mass extinction? A view from the world of amphibians. *Proc. Natl. Acad. Sci.* 105: 11466-11473. doi: 10.1073/pnas.0801921105.
- Ward P (2006). *Out of Thin Air: Dinosaurs, Birds, and Earth's Ancient Atmosphere*. National Academies Press.
- Warnock RCM, Yang Z and Donoghue PCJ (2017). Testing the molecular clock using mechanistic models of fossil preservation and molecular evolution. *Proc. R. Soc. B. Biol. Sci.* doi: 10.1098/rspb.2017.0227.
- Wilke T, Schultheiß R and Albrecht C (2009). As Time Goes by: A Simple Fool's Guide to Molecular Clock Approaches in Invertebrates. *Am. Malacol. Bull.* 27: 25-45. doi: 10.4003/006.027.0203.
- Wolfe KH (2015) Origin of the Yeast Whole-Genome Duplication. *PLoS Biol.* 13: e1002221. doi: 10.1371/journal.pbio.1002221.
- Wolfe KH and Shields DC (1997). Molecular evidence for an ancient duplication of the entire yeast genome. *Nature.* 387: 708-713. doi: 10.1038/42711.
- Yang Z (1994) Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. *J. Mol. Evol.* 39: 306-314.
- Yazgan O and Krebs JE (2012). Mitochondrial and nuclear genomic integrity after oxidative damage in *Saccharomyces cerevisiae*. *Front Biosci. Landmark Ed.* 17: 1079-1093.