bioRxiv preprint doi: https://doi.org/10.1101/049924; this version posted February 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

TITLE:

Radical amino acid mutations persist longer in the absence of sex

AUTHORS:

Joel Sharbrough[1,2], Meagan Luse[1], Jeffrey L. Boore[3,4], John M. Logsdon, Jr.[1], Maurine Neiman[1]

[1] Department of Biology, University of Iowa, Iowa City, IA 52242, USA
[2] Department of Biology, Colorado State University, Fort Collins, CO 80523, USA
[3] Department of Integrative Biology, University of California, Berkeley, Berkeley CA 94720, USA
[4] Providence St. Joseph Health and Institute for Systems Biology, 401 Terry Avenue N, Seattle, WA 98109, USA

CORRESPONDING AUTHOR:

Joel Sharbrough
jsharbro@rams.colostate.edu
(970) 491-22256
440 Biology Building
Colorado State University
Fort Collins, CO 80523

AUTHOR CONTRIBUTIONS:
JS – Concept, Data Collection, Data Analysis, Mitochondrial Genome Assembly, Writing
ML – Data Collection, Data Analysis
JLB – Data Collection, Mitochondrial Genome Assembly, Writing
JML – Concept, Writing
MN – Concept, Data Collection, Writing

KEYWORDS:

asexual reproduction, selective interference, mitochondrial genome, *Potamopyrgus antipodarum,* purifying selection, sexual reproduction

bioRxiv preprint doi: https://doi.org/10.1101/049924; this version posted February 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

ABSTRACT

Harmful mutations are ubiquitous and inevitable, and the rate at which these mutations are removed from populations is a critical determinant of evolutionary fate. Closely related sexual and asexual taxa provide a particularly powerful setting to study deleterious mutation elimination because sexual reproduction should facilitate mutational clearance by reducing selective interference between sites and by allowing the production of offspring with different mutational complements than their parents. Here, we compared the rate of removal of conservative (*i.e.*, similar biochemical properties) and radical (*i.e*., distinct biochemical properties) nonsynonymous mutations from mitochondrial genomes of sexual *vs*. asexual *Potamopyrgus antipodarum*, a New Zealand freshwater snail characterized by coexisting and ecologically similar sexual and asexual lineages. Our analyses revealed that radical nonsynonymous mutations are cleared at higher rates than conservative changes and that sexual lineages eliminate radical changes more rapidly than asexual counterparts. These results are consistent with reduced efficacy of purifying selection in asexual lineages allowing harmful mutations to remain polymorphic longer than in sexual lineages. Together, these data illuminate some of the population-level processes contributing to mitochondrial mutation accumulation and suggest that mutation accumulation could influence the outcome of competition between sexual and asexual lineages.

bioRxiv preprint doi: https://doi.org/10.1101/049924; this version posted February 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

INTRODUCTION

One of the primary hypothesized advantages of sexual reproduction is the clearance of harmful mutations, which is expected to be much more effective when linkage disequilibria (LD) are disrupted by sex (Hill and Robertson 1966). Although this LD disruption should be most apparent in the nuclear genome, biparental inheritance and meiotic recombination in the nuclear genome should also decrease LD between the nuclear genome and the mitochondrial genome (*i.e.*, mitonuclear LD) because nuclear and mitochondrial genomes can be transmitted separately. One important consequence of the potential for separate transmission of nuclear and mitochondrial alleles is a substantial reduction in selective inference between the mitochondrial and nuclear genome relative to a scenario (*e.g.*, most forms of asexual reproduction) where these genomes are always or nearly always co-transmitted (Normark and Moran 2000; Neiman and Taylor 2009). The evolutionary consequences of sexual reproduction (here, increased efficacy of natural selection) should therefore extend beyond the nuclear genome to cytoplasmic genomes, which are typically uniparentally transmitted and often lack recombination (Barr et al. 2005). Elevated mitonuclear LD resulting from co-transmission of nuclear and mitochondrial genomes, combined with simultaneous selection on nuclear variants, should impede selection in the cytoplasmic genomes of asexual lineages (Normark and Moran 2000).

Reduced efficacy of selection in the cytoplasmic genomes of asexual lineages is most obviously linked to accumulation of *slightly* deleterious mutations over time (Gabriel et al. 1993; Neiman and Taylor 2009); however, the Hill-Robertson effect (*i.e.,* smaller N_e*s) is expected to reduce the efficacy of selection with respect to *all*

bioRxiv preprint doi: https://doi.org/10.1101/049924; this version posted February 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

mutational changes. The implications are that although most strongly deleterious mutations should eventually be purged from asexual populations, these mutations will tend to remain polymorphic for a longer period of time than similar mutations in sexual populations. Thus even severely deleterious changes will have a higher probability of fixation by random drift in asexual *vs.* sexual lineages.

The prediction that asexual lineages should experience a higher rate of accumulation of harmful mutations has found support from animal (Neiman et al. 2010; Henry et al. 2012) and plant (Voigt-Zielinski et al. 2012; Hollister et al. 2015; Lovell et al. 2017) taxa. While these results represent important steps towards understanding the genomic consequences of asexuality, the evolutionary mechanisms (*e.g.,* deleterious mutation accumulation *vs.* mitonuclear coevolution) underlying these observations remain unclear, in large part because the extent to which accumulated mutations are actually deleterious in asexuals has not been evaluated. This type of information is especially important in light of substantial evidence that nonsynonymous mutations are likely to vary widely in fitness effects (Eyre-Walker and Keightley 2007; Boyko et al. 2008). Predicting and characterizing the fitness effects of these nonsynonymous mutations remains a central area of focus in evolutionary biology (Keightley and Charlesworth 2005; Xue et al. 2008; Eyre-Walker and Keightley 2009; Andolfatto et al. 2011; Halligan et al. 2011). Here, we take advantage of the fact that the biochemical properties of amino acids represent useful heuristics (*e.g*., conservative *vs.* radical amino acid changes) for inferring effects of mutations on protein phenotype (Zhang 2000; Hanada et al. 2007; Popadin et al. 2007) to take critical steps towards assessing

bioRxiv preprint doi: https://doi.org/10.1101/049924; this version posted February 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

whether mutations likely to influence fitness accumulate differently in asexual than in sexual lineages.

The removal of deleterious mutations (and the fixation of beneficial mutations) depends upon the efficacy of selection as well as the fitness effect of mutations, such that elevated N_e and reduced LD in sexual *vs.* asexual populations should result in more rapid removal of deleterious mutations for the former (Birky and Walsh 1988). As such, comparing rates and patterns of evolution across reproductive modes while incorporating information about mutational severity will provide a powerful glimpse into the evolutionary dynamics governing the removal of deleterious mutations. The New Zealand freshwater snail *Potamopyrgus antipodarum* is ideally suited to evaluate this critically important evolutionary process because otherwise similar obligately sexual and obligately asexual *P. antipodarum* frequently coexist within New Zealand lake populations (Lively 1987; Jokela et al. 1997), enabling direct comparisons across reproductive modes, and, thereby, across genomes that experience predictable variation in the efficacy of selection.

Asexual lineages of *P. antipodarum* are the product of multiple distinct transitions from sexual *P. antipodarum* (Neiman and Lively 2004; Neiman et al. 2011), meaning that these asexual lineages represent separate natural experiments into the consequences of the absence of sex. Neiman *et al.* (2010) showed that asexual lineages of *P. antipodarum* experience a higher rate of nonsynonymous substitution in their mitochondrial genomes than sexual lineages. Here, we extend these studies to a relatively short evolutionary time scale by using whole mitochondrial genomes to evaluate whether sexual lineages distinguish between radical (*i.e.,* ancestral and

bioRxiv preprint doi: https://doi.org/10.1101/049924; this version posted February 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

derived amino acids have distinct biochemical properties) and conservative (*i.e.*, ancestral and derived amino acids have similar biochemical properties) amino acid changes more effectively than asexual lineages. This approach allowed us to evaluate whether harmful mutation accumulation is detectable within species and, if so, whether this phenomenon is driven by ineffective selection in asexual lineages. The outcome of these analyses emphasizes fundamental differences in the rate of accumulation of conservative *vs.* radical nonsynonymous mutations and the critical importance of intraspecific sampling. Our results also demonstrate that radical mutations persist longer in mitochondrial genomes of asexual lineages of *P. antipodarum* compared to sexual counterparts, likely as a consequence of reduced efficacy of purifying selection.

MATERIALS & METHODS

Sequencing. We analyzed 31 whole mitochondrial genomes from eight sexual lineages and 23 asexual lineages of *P. antipodarum* and one whole mitochondrial genome from the closely related *Potamopyrgus estuarinus*, representing the wide-ranging mitochondrial genetic diversity of this species in New Zealand (see Figure 1, Table S1; Neiman and Lively 2004; Neiman et al. 2010; Paczesniak et al. 2013). We obtained 19 publicly available mitochondrial genomes (4 sexual, 14 asexual, 1 *P. estuarinus,* Accession Nos.: GQ996415 – GQ996433, Neiman et al. 2010), sequenced eight mitochondrial genomes (2 sexual, 6 asexual) using bi-directional Sanger sequencing on an ABI 3730 (Applied Biosystems, Foster City, CA), and sequenced and assembled five mitochondrial genomes (2 sexual, 3 asexual) using 2x100 bp paired-end sequencing on an Illumina HiSeq 2500 (Illumina, Inc., San Diego, CA) as part of the ongoing *P.*

bioRxiv preprint doi: https://doi.org/10.1101/049924; this version posted February 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

antipodarum nuclear genome project. For all newly sequenced lineages, we determined ploidy and, thus reproductive mode (2x – sexual; ≥ 3x – asexual), using flow cytometry following the protocol outlined in (Neiman et al. 2011; Neiman et al. 2012; Paczesniak et al. 2013; Krist et al. 2014) and extracted total genomic DNA following a mollusk-adapted phenol-chloroform extraction protocol (Fukami et al. 2004). We cleaned DNA extractions using the Zymo Clean and Concentrate Kit (Zymo Research, Irvine, CA), re-suspended DNA in 30-100 µL of T-low-E buffer (10mM Tris pH8.0, 0.1mM EDTA), and determined DNA concentration and purity for each sample on a NanoDrop 1000 (Thermo Fisher Scientific, Waltham, MA). For the eight Sanger-sequenced samples, we amplified mitochondrial genomes in four overlapping fragments using primers and programs designed in Neiman et al (2010). PCR products were cleaned with Shrimp Exo shrimp alkaline phosphatase (Werle et al. 1994) and directly sequenced with internal sequencing primers (Table S2). Sanger-sequenced mitochondrial genomes were assembled and manually edited in Sequencher 5.0 (Gene Codes Corporation, Ann Arbor, MI). Only unambiguous sites supported by ≥ 2 reads were included in the assemblies.

For the remaining five DNA samples, we constructed sequencing libraries using the Illumina Nextera DNA Library Prep Kit and sequenced 2x100 bp paired-end reads on a single lane of an Illumina HiSeq 2500. Following read trimming and quality control (≥ 30 Phred quality score) using the Fastx-toolkit (Gordon and Hannon 2010), we assembled mitochondrial genomes *de novo* using the CLC Genomics Work Bench (Qiagen, Hilden, Germany). Consensus base calls were then determined for sites with >10x coverage and Phred-scaled quality scores >30. Sites with >1 read supporting a

bioRxiv preprint doi: https://doi.org/10.1101/049924; this version posted February 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

minor allele (resulting from *e.g.*, heteroplasmies or *numts*) were not used in this analysis. All newly sequenced mitochondrial genomes have been deposited in GenBank (Accession Nos.: *sequences will be uploaded to GenBank upon acceptance*).

Phylogenetic analysis. We aligned whole mitochondrial genome sequences from the 31 *P. antipodarum* lineages and one *P. estuarinus* sample using MAFFTv v7.305b (Katoh and Standley 2013) and manually edited the alignments in MEGA v5.2.2 (Kumar et al. 2008, Supplementary File S2). We used jModelTest 2, which uses a corrected Akaike Information Criterion (AICc), Bayesian Information Criterion (BIC), and performance-based decision theory (DT) in a Maximum Likelihood (ML) framework, (Darriba et al. 2012), to select appropriate models of molecular evolution. Of the 1,624 possible models evaluated by jModelTest 2, two models produced the highest log-likelihood scores across all three model estimators: TIM3 (Posada 2003) and GTR (Tavaré 1986), both with invariant sites (I) and a gamma distribution of site classes (G). We therefore used both models (TIM3+I+G, GTR+I+G) to infer ML-based phylogenies using PAUP*4.0 (Swofford and Sullivan 2009) with 1,000 bootstrap replicates each and assuming the following priors: TIM3+I+G: lset – base = (0.2587, 0.1736, 0.1748), nst = 6, rmat = (2.7926, 49.1045, 1.0000, 2.7926, 37.8686), rates = gamma, shape = 0.8150, ncat = 4, pinvar = 0.4550; GTR+I+G: lset – base = (0.2597, 0.1736, 0.1738), nst = 6, rmat = (1.5661, 37.1103, 0.5913, 2.9024, 28.5161), rates = gamma, shape = 0.8380, ncat = 4, pinvar = 0.4610. We also inferred the phylogeny for these mitochondrial genomes under a Bayesian statistical framework using MrBayes v3.2 (Ronquist et al. 2012) assuming the following priors: nucmodel = 4by4, nst = mixed, ploidy = haploid,

bioRxiv preprint doi: https://doi.org/10.1101/049924; this version posted February 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

rates = invgamma, ngammacat = 4, number of generations = 10^6, relative burn-in = 25%, number of chains = 4, heating parameter = 0.5, revmat = all GTR submodels have equal probabilities, statefreq = dirichlet, shape = exponential (1.0), pinvar = uniform (0.0, 1.0), ratemultiplier = fixed (1.0), topology = all topologies equally probable, brlens = unconstrained: gammaDir (1.0, 0.1, 1.0, 1.0). After ensuring that the number of generations (10^6 generations) and burn-in (2.5×10^5 generations) were sufficient for the log probability of trees to plateau, trees were sampled every 500 generations, resulting in 24,016 trees of which 2,045 contributed to the 95% credible tree set. We visualized ML majority rule trees and the consensus Bayesian tree in FigTree v1.4 (Rambaut 2012); only nodes with bootstrap values ≥ 60 and posterior probabilities ≥ 75 were used in subsequent tests of molecular evolution (Figure 1). Although animal mitochondrial genomes are small (~15 kbp), there can still be substantial among-gene heterogeneity in evolutionary rate (Xia 1998; Ballard 2000; Marshall et al. 2009), which can affect phylogenetic inferences (Maddison 1997; Mallo et al. 2016). We therefore used MrBayes v3.2 to infer gene trees for each of the 13 protein-coding genes, each with their own empirically determined substitution model, in the *P. antipodarum* mt genome. We then evaluated whether nodes not present in the genome-wide tree had significant statistical support (posterior probabilities ≥ 75) in individual gene trees.

Recombination can make phylogenetic analysis complex or even impossible (Schierup and Hein 2000; Arenas and Posada 2010a, b; Arenas and Posada 2014; Mallo et al. 2016), even in mitochondrial genomes (Rokas et al. 2003; Barr et al. 2005; Tsaousis et al. 2005). Recombination in mitochondrial genomes is particularly problematic in phylogenetic analyses of sexual lineages *vs.* asexual lineages because

bioRxiv preprint doi: https://doi.org/10.1101/049924; this version posted February 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

only sexual lineages have the potential to experience paternal leakage and recombination between mitochondrial genomes from different lineages. To evaluate evidence for recombination among *P. antipodarum* mitochondrial genomes, we estimated population recombination rate using LDHat (implemented by RDP4 – Martin et al. 2017). To test whether recombination had any effect on our phylogenetic analyses, we also inferred a genome-wide phylogenetic tree using MrBayes excluding those regions with evidence of recombination.

Next, we extracted and concatenated the 13 protein-coding regions from the alignment (~11.2 kbp, 3,738 amino acids) for all 32 lineages, and we mapped all mutational changes in protein-coding regions to the fixed tree topology (described above) according to the rules of parsimony. Only non-homoplasious changes were considered in subsequent phylogenetic analyses of molecular evolution. In our population genetic analyses, we considered all sites that were variable within *P. antipodarum* (including homoplasious sites) to be polymorphic and all sites that were distinct between *P. antipodarum* and *P. estuarinus* to be substitutions.

Identifying conservative and radical amino acid changes. We used seven different amino acid classification schemes, three drawn from Zhang (2000), three drawn from Hanada *et al.* (2007), and a modified Grantham scheme based on charge and hydrophobicity (Grantham 1974), to evaluate rates and patterns of radical and conservative amino acid evolution (Table 1). While there is some overlap between different classification schemes, each scheme highlights different amino acid properties that are likely to shape protein evolution. For example, amino acid charge is a major

bioRxiv preprint doi: https://doi.org/10.1101/049924; this version posted February 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

determinant of protein folding (Perutz et al. 1965; Anfinsen 1973; Nakashima et al. 1986; Bashford et al. 1987; Wright et al. 2005) and three-dimensional structure (Lesk and Chothia 1980; Geisler and Weber 1982; Doms et al. 1988; Rumbley et al. 2001), changes between polar and non-polar amino acids can expose or bury key interaction residues in membrane-associated proteins (von Heijne 1992), and volume and aromaticity can both affect protein folding and play a role in protein-protein interactions (Burley and Petsko 1985). Classification schemes 4 and 7 (Table 1) are unique in that they are based on evolutionary information (although classification scheme 7 largely fits with charge and polarity classifications), meaning that these schemes incorporate aspects of other amino acid characteristics into their classifications. Using these seven amino acid classification schemes, we developed an overall index of the degree of amino acid change that we termed the Conservative-Radical Index (CRI). For each possible amino acid change, we calculated CRI by averaging across the seven amino acid classification schemes (radical changes assigned a value = 1.0, conservative changes assigned a value = 0.0), such that CRI = 1.0 indicates that the amino acid change in question was radical in all amino acid classification schemes (*e.g.*, V ←→ E) whereas CRI = 0.0 indicates that the amino acid change in question was conservative in all amino acid classification schemes (*e.g.*, V ←→ I). Amino acid changes with CRI ≤ 0.5 were treated as conservative changes and amino acid changes with CRI > 0.5 were treated as radical changes (Table S3). All analyses of molecular evolution and population genetics were repeated for each amino acid scheme individually as well as for the classification scheme-averaged classifier.

bioRxiv preprint doi: https://doi.org/10.1101/049924; this version posted February 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

We then calculated the number of mutational target sites per codon for each different type of mutational change (*i.e.*, synonymous, nonsynonymous, conservative nonsynonymous, and radical nonsynonymous) relative to the invertebrate mitochondrial genetic code (http://www.ncbi.nlm.nih.gov/Taxonomy/Utils/wprintgc.cgi#SG5) to ensure that we properly accounted for the different probabilities of different types of mutational changes (Table S4). To confirm that the number of each of type of site was properly calculated, we checked that the number of nonsynonymous sites and the number of synonymous sites per codon summed to three and that the number of conservative nonsynonymous sites and the number of radical nonsynonymous sites per codon summed to the number of nonsynonymous sites per codon. Of particular note is that the "GTG", "TTG", "ATT", "ATC", and "ATA" codons can all be used as alternative start codons in invertebrate mitochondrial genomes (only the GTG alternative start codon was observed in the present dataset), which we accounted for in our site calculations.

Molecular evolution and population genetic analyses. To evaluate the relative intensity of selection acting on different types of mutational changes, we estimated rates of substitution between *P. antipodarum* and *P. estuarinus* (Li et al. 1985), ratios of polymorphism to divergence (McDonald and Kreitman 1991), nucleotide diversity (Watterson 1975; Nei and Li 1979; Fu and Li 1993), and site frequency spectra (SFS, Tajima 1989) in protein-coding regions of the *P. antipodarum* mitochondrial genome for each mutational type using all seven amino acid classification schemes and the scheme-averaged classifier with a custom python tool (available at https://github.com/jsharbrough/conservativeRadicalPolymorphismsSubstitutions). For

bioRxiv preprint doi: https://doi.org/10.1101/049924; this version posted February 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

codons with multiple hits, we assumed the fewest number of nonsynonymous changes necessary to explain the data.

To compare rates of molecular evolution in *P. antipodarum* relative to *P. estuarinus*, we first calculated substitution rates for each mutational type in *P. antipodarum* as in Li *et al.* (1985), corrected these estimates for multiple hits using the Jukes-Cantor model of molecular evolution (Jukes and Cantor 1969), and estimated variance as in Jukes and Cantor (1969). We then compared 95% confidence intervals (CIs) of substitution rates between *P. antipodarum* and *P. estuarinus*, estimated as $95\%\ \mathrm{CI} = 2\sqrt{(\mathrm{variance})}$ for each substitution type, such that non-overlapping CIs indicate statistically different rates of substitution. We also took advantage of the non-homoplasious changes mapped to our fixed tree topology (see above) to compare synonymous substitution rate (K_S), nonsynonymous substitution rate (K_A), conservative substitution rate (K_C), and radical substitution rate (K_R) relative to *P. estuarinus*. We estimated branch lengths by summing the mapped changes for each lineage and mutational type and dividing the number of changes by the number of mutational target sites (Table S4). We then used pairwise Wilcoxon Signed-Rank (WSR) tests and the Holm procedure for correcting for multiple comparisons to determine whether branch lengths differed across mutational types (Holm 1979). We next used McDonald-Kreitman (MK) tests to compare molecular evolution in conservative *vs.* radical changes using a series of Fisher's Exact Tests (FET): (1) synonymous changes to conservative changes, (2) synonymous changes to radical changes, and (3) conservative changes to radical changes. All statistical tests were performed in R v3.2.4 (R Core Team 2016).

bioRxiv preprint doi: https://doi.org/10.1101/049924; this version posted February 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

To compare species-wide patterns of polymorphism, we estimated nucleotide diversity using θ_π – hereafter π – (Nei and Li 1979) and θ_W – hereafter θ – (Watterson 1975), and their respective variances (Durrett 2008) within *P. antipodarum*. We compared levels of nucleotide diversity across mutational types by constructing 95% CIs for each mutational type using 10,000 bootstrap replicates. For each bootstrap replicate, we randomly selected 3,738 codons with replacement and recalculated all population genetic statistics for each replicate using a custom Python tool (available at https://github.com/jsharbrough/popGenBootstrapping). We also performed an estimate of θ using only private polymorphisms – hereafter θ_U. Because private polymorphisms should often represent relatively young mutations that are relatively independent of phylogenetic constraint (Fu and Li 1993), these polymorphisms can be used to estimate rates of mutation accumulation since the transition to asexuality. We compared levels of these private polymorphisms across mutational types with the same bootstrapping approach as with π and θ, using a custom Python tool to compare CIs for each mutational type and compute *p* values (available at https://github.com/jsharbrough/distributionPvalues). This tool rank-orders bootstrap replicates and tests the degree to which bootstrap distributions overlap (*i.e.*, perfect overlap of bootstrap distributions: $p > 0.9998$, no overlap of bootstrap distributions: $p < 0.0002$, for a two-tailed test). Because deleterious changes should generally be found at lower frequencies than relatively neutral changes, we also compared the SFS across mutational types using a series of Goodness-of-Fit (GoF) analyses assuming a χ^2 distribution, first for nonsynonymous, conservative, and radical polymorphisms *vs.* synonymous polymorphisms, and second for conservative *vs.* radical polymorphisms.

bioRxiv preprint doi: https://doi.org/10.1101/049924; this version posted February 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

We performed these GoF analyses with the chisq function in R v3.2.4 (R Core Team 2016) and visualized species-wide polymorphism levels using the R package ggplot2 (Wickham 2016).

After establishing that the substitution rate for radical amino acid changes is lower than for conservative amino acid changes and that radical changes exhibit lower nucleotide diversity within species than conservative changes (see Results, Figure 2), we compared rates of mutation accumulation of conservative and radical amino acid changes across reproductive modes. We used the mapped changes to estimate synonymous-corrected branch lengths for conservative and radical sites and compared rates of molecular evolution for each mutational type separately in sexual *vs*. asexual lineages using Mann-Whitney U (MWU) tests. We also performed these same comparisons on internal branches only and on branch tips only to test whether patterns of mutation accumulation across reproductive modes were robust across evolutionary time.

We next compared polymorphism levels for conservative and radical changes across reproductive modes by dividing our sample into sexual (n = 8) and asexual (n = 23) groups and separately calculating π, θ, and θ_U for each mutational type. Asexual *P. antipodarum* have already been shown to accumulate nonsynonymous changes more rapidly than their sexual counterparts (Neiman et al. 2010), leading to the *a priori* expectation that asexual lineages should exhibit higher levels of nucleotide diversity at nonsynonymous sites than sexual lineages. To address whether this prediction was met, we calculated a test statistic, D_{AS}, in which $D_{AS} = \bar{x}_{Asex} - \bar{x}_{Sex}$, where $\bar{x}$ represents the diversity-related population genetic statistic under consideration. We then constructed a

bioRxiv preprint doi: https://doi.org/10.1101/049924; this version posted February 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

null distribution for each of these statistics using a custom-built python program (available at https://github.com/jsharbrough/conRadNullDistribution), which randomly assigned lineages without replacement and without consideration of reproductive mode to one of two groups: group 1 (n = 23, representing the asexual lineage sample size) and group 2 (n = 8, representing the sexual lineage sample size), calculated the difference between the two groups, D_{12}, where $D_{12} = \bar{x}_1 - \bar{x}_2$, for each population genetic statistic, and repeated this process 10,000 times. Because we expected asexual lineages to exhibit higher levels of nonsynonymous polymorphism than sexual lineages, we conducted one-tailed tests of polymorphism across reproductive modes (H_O: $D_{AS} = 0$; H_A: $D_{AS} > 0$) by comparing D_{AS} to the null distribution for each population genetic statistic (*i.e.*, π, θ, and θ_U) for synonymous sites, synonymous-corrected conservative sites, and synonymous-corrected radical sites. We inferred *p* values directly from these comparisons using a custom python tool (available at https://github.com/jsharbrough/distributionPValues) and visualized null distributions and test statistics using the ggplot2 package for R (Wickham 2016).

RESULTS

Phylogenetic analysis of mitochondrial genetic diversity in *P. antipodarum*. We used 32 whole mitochondrial genomes to infer a phylogeny of sexual and asexual *P. antipodarum* relative to a *P. estuarinus* outgroup using Maximum Likelihood and Bayesian methods (Figure 1). Because we did not detect any major differences from our genome-wide tree in individual gene trees (Figure S1) or in a tree that excluded regions

bioRxiv preprint doi: https://doi.org/10.1101/049924; this version posted February 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

suggestive of recombination (Figure S2), we performed our molecular evolutionary analyses assuming the genome-wide tree topology.

These phylogenetic analyses revealed substantial intraspecific genetic variation within *P. antipodarum* (mean pairwise genetic distance ≈ 0.025 changes/site) as well as relatively deep divergence for one entirely asexual clade (clade A; mean pairwise distance between clade A *vs.* clades B/C ≈ 0.035 changes/site, *vs.* mean pairwise distance within B/C ≈ 0.013). Clade A is composed mostly of lineages from the North Island of New Zealand (Lakes Waikaremoana and Tarawera) but also includes one invasive lineage from Wales and two lineages collected from lakes on the South Island of New Zealand (Lakes Brunner and Kaniere). Clades B/C are predominantly composed of lineages collected from the South Island of New Zealand (Lakes Alexandrina, Brunner, Grasmere, Gunn, Heron, Ianthe, Kaniere, Lady, McGregor, Poerua, Rotoiti, Rotoroa); however, two invasive lineages (collected from Lake Superior in Duluth, MN, USA and Denmark), one North Island lineage (collected from Lake Waikaremoana), and a lineage collected from Lake Alexandrina form a monophyletic, all-asexual clade within clade C.

We next used parsimony to map the 1,524 non-homoplasious changes (out of 1,711 total variable sites) present in the ~11 kbp protein-coding region of the mitochondrial genome onto this tree topology (Figure 1). In all, there were 814 synonymous substitutions and 35 nonsynonymous substitutions between *P. antipodarum* and *P. estuarinus*. Of these 35 nonsynonymous substitutions, 30 were considered conservative amino acid changes and 5 were considered radical amino acid changes by our scheme-averaged classifier index, CRI (see Materials and Methods).

bioRxiv preprint doi: https://doi.org/10.1101/049924; this version posted February 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Within *P. antipodarum*, we observed 651 synonymous polymorphisms and 211 nonsynonymous polymorphisms; CRI classified 151 of the latter as conservative amino acid polymorphisms and the other 60 nonsynonymous polymorphisms as radical amino acid changes (Figure 2b).

Substitution rates and polymorphism levels for conservative vs. radical amino acid changes. Using the mapped changes to calculate branch-specific estimates of Jukes-Cantor-corrected substitution rate (K), we found that mean branch-specific synonymous substitution rate (K_S) was significantly higher than mean branch-specific nonsynonymous substitution rate (K_A; WSR: $V = 496$, $p < 0.0001$), conservative substitution rate (K_C; WSR: $V = 496$, $p < 0.0001$), and radical substitution rate (K_R; WSR: $V = 496$, $p < 0.0001$), revealing that nonsynonymous changes evolve at lower rates than do synonymous changes (Table 2). Branch-specific estimates of K_C were also significantly higher than branch-specific estimates of K_R (WSR: $V = 496$, $p < 0.0001$), indicating that among amino acid-changing mutations, radical amino acid changes evolve particularly slowly (Figure 2a, Table 2). In accordance with branch-specific estimates, species-wide estimates of substitution rates that only use fixed differences between *P. antipodarum* and *P. estuarinus* revealed that K_S was significantly greater than K_A, K_C, and K_R, and that K_C was significantly greater than K_R (Table S5). In summary, branch-specific and species-wide estimates of substitution rate indicate that nonsynonymous sites have lower rates of substitution than synonymous sites and that radical nonsynonymous sites have lower rates of substitution than conservative nonsynonymous sites.

bioRxiv preprint doi: https://doi.org/10.1101/049924; this version posted February 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

To determine the probability of fixation of conservative *vs.* radical nonsynonymous polymorphisms, we performed pairwise MK tests of selection across all mutational types. We found that nonsynonymous (FET: $p < 0.0001$), conservative (FET: $p < 0.0001$), and radical (FET: $p < 0.0001$) amino acid polymorphisms were significantly less likely to reach fixation than synonymous polymorphisms (Figure 2b, Table 3). There was a trend for conservative polymorphisms to become fixed within *P. antipodarum* at higher rates than radical polymorphisms, but this difference was not significant (FET: $p = 0.098$). These data indicate that nonsynonymous changes have a lower probability of fixation than synonymous changes, but that there is no difference in fixation probabilities across different types of nonsynonymous changes.

To investigate how conservative and radical changes evolve within species, we estimated nucleotide diversity for all mutational types using the frequency-dependent measure π, the frequency-independent measure θ, and a measure of private nucleotide diversity θ_U, treating all lineages as a single population. Comparisons of π revealed that nonsynonymous, conservative, and radical sites exhibited lower levels of nucleotide diversity than synonymous sites ($p < 0.0002$ for all comparisons, Figure S3a, Table 2). Similarly, nonsynonymous, conservative, and radical sites all exhibited lower levels of nucleotide diversity than synonymous sites, as measured by θ ($p < 0.0002$ for all comparisons, Figure S3b, Table 2), and θ_U ($p < 0.0002$ for all comparisons, Table 2, Figure S3c, Table 2). Together, these results indicate that all types of nonsynonymous changes are eliminated from *P. antipodarum* mitochondrial genomes more rapidly than synonymous changes (Figure 2c, Table 2). We also found that π_C and θ_C were significantly greater than π_R ($p = 0.0004$) and θ_R ($p < 0.0002$), respectively (Table 2,

bioRxiv preprint doi: https://doi.org/10.1101/049924; this version posted February 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 2c); however, $\theta_{U\text{-}C}$ and $\theta_{U\text{-}R}$ are statistically indistinguishable within *P. antipodarum* (p = 0.072, Figure S3c, Table 2). Because asexual lineages likely experience distinct demographic forces that sexual lineages do not (Charlesworth and Wright 2001; Kaiser and Charlesworth 2009), we also compared these population genetic statistics using only sexual lineages and found evidence of similarly elevated intensity of selection against radical *vs.* conservative polymorphisms as observed in analyses incorporating all *P. antipodarum* mitochondrial genome sequences (Figure S4).

Finally, we compared allele frequencies of different mutational types by comparing their SFS (Figure 2d). We found that nonsynonymous (χ^2 = 56.14, df = 7, p < 0.0001), conservative (χ^2 = 28.57, df = 7, p < 0.0001), and radical SFS (χ^2 = 35.04, df = 7, p < 0.0001) are significantly left-skewed (*i.e.*, excess of rare variants) relative to synonymous changes. We did not detect any differences in conservative *vs.* radical SFS (χ^2 = 6.36, df = 4, p = 0.17). In sum, all types of nonsynonymous changes are found in lower numbers and at lower frequencies than synonymous changes and that radical amino acid polymorphisms tend to be found in lower numbers and at lower frequencies than conservative amino acid polymorphisms.

Efficacy of purifying selection in sexual *vs.* asexual lineages. To evaluate the degree to which the efficacy of selection is reduced in asexual *vs.* sexual lineages of *P. antipodarum*, we used the same set of 1,524 synapomorphic and autapomorphic changes (Figure 1) to estimate sexual and asexual lineage branch lengths relative to outgroup *P. estuarinus* and compare relative rates of conservative and radical amino acid molecular evolution. After accounting for the number of mutational target sites

bioRxiv preprint doi: https://doi.org/10.1101/049924; this version posted February 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

(Table S4) and correcting for multiple hits (Jukes and Cantor 1969), we found that asexual lineages exhibited higher rates of substitution at synonymous sites than sexual lineages (MWU: $W = 142$, $p = 0.025$, Figure S5a, Table 4), potentially indicating that the mitochondrial genomes of asexual lineages of *P. antipodarum* experience higher mutation rates than their sexual counterparts. We therefore corrected nonsynonymous rates of molecular evolution with synonymous rates to ensure that we could adequately compare efficacy of selection acting on conservative and radical changes across reproductive modes. We did not detect a difference in the rate of molecular evolution at conservative sites in sexual *vs.* asexual lineages (MWU: $W = 95$, $p = 0.91$, Figure 3a, Table 4). By contrast, we did find that asexual lineages exhibit significantly higher rates of molecular evolution at radical sites than sexual lineages (MWU: $W = 168$, $p < 0.0001$, Figure 3b, Table 4), indicating that asexual lineages accumulate radical amino acid changes more rapidly than sexual lineages. We obtained a similar pattern of accelerated accumulation of radical amino acid changes in asexual *vs.* sexual lineages when excluding branch tips from the analysis (K_C/K_s – MWU: $W = 84$, $p = 0.74$; K_R/K_S – MWU: $W = 168$, $p < 0.0001$), indicating that this pattern is robust across evolutionary time. Results from all amino acid classification schemes are detailed in Table S6.

We next compared patterns of polymorphism across reproductive modes and mutational types using π *and* θ. For each mutational type, we calculated an estimate of nucleotide diversity for asexuals and sexuals separately, using the difference D as a test statistic to compare against a null distribution (see Materials and Methods). Although sexual and asexual lineages do not appear to differ in terms of synonymous polymorphism levels (Figure S5b, Table 4), K_S, was elevated in asexual compared to

bioRxiv preprint doi: https://doi.org/10.1101/049924; this version posted February 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

sexual lineages (Figure S5a, Table 4). Because this result could reflect a scenario where differential effects of neutral processes change in response to the transition to asexuality (Charlesworth and Wright 2001; Kaiser and Charlesworth 2009), we accounted for levels of neutral polymorphism in both sexuals and asexuals by correcting nonsynonymous estimates of nucleotide diversity with the corresponding synonymous rate. For both π and θ, there was no evidence that sexual lineages differed from asexual lineages in the level of conservative amino acid polymorphism (D_π = -0.0086, p = 0.91, Figure S6a; D_θ = 0.014, p = 0.70, Figure 4a, Figure S6c, Table S6). The same analyses applied to radical amino acid polymorphisms revealed significantly higher levels of radical polymorphisms in asexual *vs.* sexual lineages (D_π = 0.012, p = 0.049, Figure 4a Figure S6b; D_θ = 0.035, p = 0.0015, Figure S6d, Table S6).

Clade A asexuals appear to be genetically distinct relative to the rest of the *P. antipodarum* dataset (Figure 1), raising the question of whether this group might be contributing disproportionately to our observations of elevated radical amino acid polymorphism levels in asexual *P. antipodarum*. We therefore repeated the comparisons of π and θ with clade A excluded, using the clade B sexuals as an outgroup to clade C (Figure 4b, Table 4). In this more limited analysis, we again found that asexuals harbored higher levels of radical (D_π = 0.023, p = 0.0038; D_θ = 0.050, p < 0.0001) but not conservative amino acid polymorphism (D_π = -0.00057, p = 0.60,; D_θ = 0.0040, p = 0.74) (Figure 4b, Figure S6e-h), indicating that the clade A asexuals were not entirely responsible for the patterns of elevated radical amino acid polymorphism in *P. antipodarum* asexuals.

bioRxiv preprint doi: https://doi.org/10.1101/049924; this version posted February 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

To account for the phylogenetic non-independence of asexual lineages in *P. antipodarum* in our population genetic analyses, we also compared levels of private polymorphism across reproductive modes and mutational types (Figure 4c, Table 4). There was a trend towards higher levels of synonymous private polymorphism in sexual compared to asexual lineages ($D_{\theta U}$ = -0.0017, p = 0.057, Figure S5b), so we corrected for private synonymous polymorphism levels before comparing private nonsynonymous sites between sexuals and asexuals. These analyses revealed that asexual lineages exhibited significantly higher levels of both private conservative amino acid polymorphism ($D_{\theta U}$ = 0.28, p = 0.0031) and private radical amino acid polymorphism than sexual lineages ($D_{\theta U}$ = 0.27, p = 0.0059, Figure 4c, Table S6). The lower levels of private polymorphism in sexual compared to asexual lineages for both conservative and radical mutational types indicates that selection is less effective in asexual lineages and that this effect is detectable even over relatively short time scales.

DISCUSSION

Radical amino acid changes accumulate in the absence of sex

Here, we provide evidence from a non-model system featuring reproductive mode polymorphism (a primary determinant of N_e) that there exists more stringent purifying selection on radical *vs.* conservative mutational types. This pattern is pronounced enough to be evident even among intraspecific comparisons. These results are consistent with the expectation that radical mutations are usually more harmful than conservative mutations (Freudenberg-Hua et al. 2003; Smith 2003), and demonstrate that radical and conservative mutations appear to experience very different histories of

bioRxiv preprint doi: https://doi.org/10.1101/049924; this version posted February 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

selection in natural populations. The implication is that radical amino acid changes usually impart substantially more severe fitness effects than conservative changes and should therefore return to mutation-selection-drift equilibrium more rapidly than conservative amino acid changes (Fisher 1930).

We also found that asexual *P. antipodarum* lineages experience higher rates of radical amino acid substitution and exhibit higher incidences (θ) and frequencies (π) of radical amino acid polymorphisms than sexual counterparts. These results suggest that asexual *P. antipodarum* experience a relatively high rate of accumulation of likely harmful mutations in their mitochondrial genomes. Although using biochemical properties to predict likely phenotypic consequences of amino acid changes is ultimately limited by the reality that amino acid position can also play a role in protein properties (see *e.g.,* Zhai et al. 2009), compensatory coevolution between nuclear and mitochondrial genomes has been documented on a background of conservative, but not radical, structural evolution (Havird et al. 2015). The implications of this latter finding are that radical amino acid changes are rarely, if ever, tolerated in the multi-subunit complexes of OXPHOS, making adaptive evolution an unlikely explanation underlying the elevated numbers of radical mutations we have observed in mitochondrial genomes of asexual *P. antipodarum*. We therefore interpret this result as evidence that mutations in asexual lineages experience lower N_e*s than in sexual lineages, with the implication that the elimination of even strongly deleterious mutations should occur more slowly in asexual than in sexual lineages. Theoretical (Ohta 1987; Charlesworth et al. 1993; Charlesworth and Wright 2001) and empirical (Eyre-Walker and Keightley 2007; Wright et al. 2008; Katju et al. 2015) work support this conclusion, suggesting that populations

bioRxiv preprint doi: https://doi.org/10.1101/049924; this version posted February 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

with low N_e should harbor a larger proportion of "effectively neutral" mutations than populations with large N_e. The somewhat surprising upshot of this prediction is that even severely deleterious mutations become more effectively neutral and thus more likely to be subject to fixation by drift when N_e is low. The absence of any detectable differences in sexual and asexual lineages in their rate of removal of conservative changes suggest that the conservative changes observed in this dataset are evolving *via* neutral processes in *both* sexual and asexual lineages. Only radical amino acid changes exhibit severe enough effects on fitness to be differentially maintained in asexual *vs.* sexual lineages, highlighting the utility of distinguishing between mutations with different effects on fitness when comparing the efficacy of selection across lineages.

The observation that asexual *P. antipodarum* accumulate radical nonsynonymous mutations more rapidly than sexual *P. antipodarum* raises the intriguing possibility that asexual *P. antipodarum* might exhibit decreased mitochondrial function compared to sexual counterparts. The presumed severity of some mutations in these genomes (*e.g.,* a nonsense mutation in *nad2* of one asexual lineage that would truncate the NAD2 protein by three amino acids) suggests that either mitochondrial function is decreased in at least some of these asexual lineages or that asexuals possess one or more mechanisms to compensate for deleterious mutation load (*e.g.,* RNA editing). RNA editing of mitochondrially encoded transcripts has been observed in a variety of plant taxa (Covello and Gray 1989; Gualberto et al. 1989) and in the mitochondrially encoded tRNAs of land snails (Yokobori and Paabo 1995); whether *P. antipodarum* employs similar strategies has not been addressed. Asexual lineages of *P. antipodarum* do exhibit substantial phenotypic variation for mitochondrial function, which

bioRxiv preprint doi: https://doi.org/10.1101/049924; this version posted February 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

has in turn been linked to mitonuclear genotype (Sharbrough et al. 2017). Future work evaluating mitochondrial function at the organelle and organismal levels in *P. antipodarum* will be essential to understanding how the efficacy of selection influences the relative success of sexual *vs.* asexual lineages and the maintenance and distribution of sexual *P. antipodarum*.

Mitonuclear coevolution in the absence of sex

We interpret these results as a consequence of reduced efficacy of selection in asexual lineages; however, another possible (and non-mutually exclusive) explanation is that the co-transmission (and thus, effective linkage) between the nuclear and mitochondrial genomes in asexuals has facilitated the persistence and spread of beneficial nonsynonymous mutations *via* selection imposed by cooperation with nuclear-encoded genes (Blier et al. 2001; Meiklejohn et al. 2007). Because apomictic asexuals co-transmit their nuclear and mitochondrial genomes, mutations in either genome may cause decreases in mitochondrial function. Therefore, long-term co-transmission of the nuclear and mitochondrial genomes may provide a scenario in which asexuals experience relatively strong selection favoring compensatory mutation(s). This expectation aligns well with theory suggesting that recombination can be selected against in a scenario where the genes contributing to fitness interact non-additively (see *e.g.,* Maynard Smith 1978; Kondrashov 1994; Neiman and Linksvayer 2006). Epistatic selection in which recombination is disfavored has found support from the multiple examples of relatively rapid responses to selection and high heritability of polygenic traits in asexual taxa (*e.g.*, Stratton 1991; Browne 1992; Christensen et al. 1992;

bioRxiv preprint doi: https://doi.org/10.1101/049924; this version posted February 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Stratton 1992; Weeks and Hoffmann 1998). More recently, Lenski and colleagues have shown extensive and rapid adaptive evolution in asexual bacteria lineages in the course of their long-term experimental evolution studies (Lenski 2017). The strong body of work suggesting that mitonuclear interactions are epistatic in nature (*e.g.,* Osada and Akashi 2012; Barreto and Burton 2013; Beck et al. 2015; Havird et al. 2015) indicates that the multi-subunit enzymes of the oxidative phosphorylation pathway might be subject to relatively effective evolution by natural selection in an asexual context.

Whether these epistatic interactions contribute to coevolution between the nuclear and mitochondrial genomes of asexual lineages remains an open question. We have not detected any evidence of positive selection acting in the mitochondrial genome of *P. antipodarum* (*e.g.,* codon-by-codon $d_N/d_S < 1$, Neutrality Index > 1 for all 13 protein-coding genes, sliding window $\pi_A/\pi_S < 1$ at all sites; Sharbrough et al. in preparation), though indirect evidence that a particular mitochondrial haplotype is spreading among asexual lineages hints that selection favoring particular mitochondrial haplotypes or mitonuclear combinations might be involved (Paczesniak et al. 2013). Evaluation of rates and patterns of evolution in the nuclear-encoded mitochondrial genes that make up ≥95% of the genes that influence mitochondrial function (Sardiello et al. 2003), coupled with the functional analyses mentioned above, is thus a critical next step towards determining whether mitonuclear linkage in asexuals is at least in part responsible for elevated retention of apparently harmful mutations in mitochondrial genomes.

Ultimately, inefficient removal of seemingly deleterious mutations in asexual lineages is only important to understanding the maintenance of sex and/or the

bioRxiv preprint doi: https://doi.org/10.1101/049924; this version posted February 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

persistence of asexual lineages if those mutations, in fact, negatively affect fitness. Indeed, consistent with the potential that mitonuclear coevolution might be especially effective in asexuals, we cannot exclude the possibility that radical mitochondrial mutations are not as harmful as they appear from our analyses. In particular, we speculate that the selective regime operating on mitochondrial genes may be affected by asexuality itself (*e.g.,* elevated heritability of mitonuclear genotypes, allowing for effective compensation of deleterious mutations), such that apparently deleterious mutations do not, in fact, impart reduced function and fitness. Functional validation of deleterious effects of mutations therefore represents an especially central component of determining the evolutionary implications of accelerated accumulation of radical mutations.

Mitonuclear LD and the efficacy of purifying selection in mitochondrial genomes

For decades, deleterious mutation accumulation has been thought to be an unavoidable hallmark of mitochondrial genome evolution (Gabriel et al. 1993; Neiman and Taylor 2009). In the last few years, several examples of relatively effective natural selection operating in mitochondrial genomes compared to nuclear genomes (Cooper et al. 2015; Konrad et al. 2017) have called this widely held assumption into question. Indeed, the extent to which mitochondrial genomes actually undergo mutational meltdown has profound implications for a number of downstream evolutionary hypotheses including the compensatory model of mitonuclear coevolution (Rand et al. 2004; Dowling et al. 2008; Niehuis et al. 2008; Osada and Akashi 2012; Havird et al. 2015), the evolution of mitonuclear reproductive incompatibilities that could contribute to speciation (Harrison

bioRxiv preprint doi: https://doi.org/10.1101/049924; this version posted February 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

and Burton 2006; Niehuis et al. 2008; Burton and Barreto 2012; Hill 2016), introgressive mitochondrial replacement (Toews and Brelsford 2012; Sloan et al. 2017), and rates of extinction (Lynch et al. 1993; Havird et al. 2016). If natural selection is indeed effective in mitochondrial genomes (especially purifying selection), selection to alleviate deleterious epistatic interactions between nuclear and mitochondrial genomes may not play as central a role in the generation of eukaryotic diversity as previously thought (Montooth et al. 2010; Adrion et al. 2016).

While the results presented here are restricted to mitochondrial genomes, they do provide empirical evidence regarding the efficacy of selection in mitochondrial genomes. In particular, we show that in low N_e conditions (*i.e.,* asexual lineages), elevated selective interference appears to increase rates of deleterious mutation accumulation in mitochondrial genomes. That is, as expected under the Hill-Robertson effect, the higher mitonuclear LD experienced by mitochondrial genomes “trapped” in asexual lineages is associated with mitochondrial mutation accumulation. It follows that mitochondrial genomes that appear to experience relatively effective purifying selection (*e.g.,* the mitochondrial genomes in sexual lineages of *P. antipodarum*), there must be some mechanism (likely sexual reproduction) whereby the Hill-Robertson effect is ameliorated. Based on this logic, we predict that species harboring relatively few deleterious polymorphisms in mitochondrial genomes (*e.g., Caenorhabditis*, *Drosophila*) will also exhibit low levels of mitonuclear LD compared to those harboring many deleterious polymorphisms in mitochondrial genomes (*e.g.*, *Nasonia*, *Potamopyrgus*).. Characterizing the genetic mechanisms that promote and restrict mitochondrial mutation accumulation in some lineages but not others therefore represents a critical component

bioRxiv preprint doi: https://doi.org/10.1101/049924; this version posted February 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

of understanding the role that mitonuclear coevolution has played in eukaryotic evolution.

bioRxiv preprint doi: https://doi.org/10.1101/049924; this version posted February 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

DATA ARCHIVAL LOCATION:

Sequence data – GenBank

Python scripts – https://github.com/jsharbrough

CONFLICT OF INTEREST

The authors do not declare a conflict of interest.

ACKNOWLEDGEMENTS

We thank Cindy Toll and Gery Hehman for their assistance with DNA sequencing. We also thank Stephen I. Wright and Aneil F. Agrawal for helpful discussions regarding interpretation of intraspecific data. We thank Samuel J. Fahrner for helpful discussions regarding bootstrapping. We thank Jeremiah W. Busch and several anonymous reviewers who saw previous versions of this manuscript for their helpful comments. Some of the data presented herein were obtained at the Flow Cytometry Facility, which is a Carver College of Medicine / Holden Comprehensive Cancer Center core research facility at the University of Iowa. The Facility is funded through user fees and the generous financial support of the Carver College of Medicine, Holden Comprehensive Cancer Center, and Iowa City Veteran's Administration Medical Center. The National Science Foundation (NSF: MCB – 1122176; DEB – 1310825) and the Iowa Academy of Sciences (ISF #13-10) funded this research.

bioRxiv preprint doi: https://doi.org/10.1101/049924; this version posted February 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

656 **TABLES**

Table 1. Amino acid classification schemes.

1 Zhang (2000) **Charge**	2 Zhang (2000) **Polarity**	3 Zhang (2000) **Polarity & Volume**	4 Hanada et al. (2007) **Correlation with K_A/K_S**	5 Hanada et al. (2007) **Charge & Aromaticity**	6 Hanada et al. (2007) **Charge & Polarity**	7 Grantham (1974) **Charge & Hydrophobicity**
Neutral A, C, F, G, I, L, M, N, P, Q, S, T, V, W, Y	*Polar* C, D, E, G, H, K, N, Q, R, S, T, Y	*Neutral, small* A, G, P, S, T	*Neutral, small* A, C, G, N, P, S, T	*Neutral, non-aromatic* A, C, G, I, L, M, N, P, Q, S, T, V	*No polarity* A, F, G, I, L, M, P, V, W	*Neutral, hydrophobic,* A, F, I, L, M, P, V, W
Positively charged H, K, R		*Polar, small* D, E, N, Q *Nonpolar, small* I, L, M, V	*Basic, acid, aromatic, small* F, H, K, Q, R, W, Y	*Neutral, aromatic* F, W, Y	*Neutral, polar* C, N, Q, S, T, Y	*Neutral, hydrophilic* C, G, N, Q, S, T, Y
Negatively charged D, E	*Non-polar* A, F, I, L, M, P, V, W	*Polar, large* H, K, R *Nonpolar, large* F, W, Y *Special* *C*	*Neutral, large* I, L, M, V *Acidic, large* D, E	*Basic* H, K, R *Acidic* D, E	*Basic, polar* H, K, R *Acidic, polar* D, E	*Positive, hydrophilic* H, K, R *Negative, hydrophilic* D, E

657

bioRxiv preprint doi: https://doi.org/10.1101/049924; this version posted February 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table 2. Mean branch-specific substitution rates and nucleotide diversity across mutational types in *P. antipodarum* mitochondrial genomes.

Mutational Type	*K* (Variance[a])	π^b (95% CI)	θ^b (95% CI)	$\theta_U{}^b$ (95% CI)
Syn	0.48 * (0.46 – 0.51)	0.064 * (0.058 – 0.069)	0.063 * (0.058 – 0.067)	0.0087 * (0.0069 – 0.011)
Nsyn	0.0077 (0.0060 – 0.0093)	0.0046 (0.0038 – 0.0053)	0.0061 (0.0052 – 0.0069)	0.0025 (0.0020 – 0.0031)
Con[c]	0.013 (0.010 – 0.015)	0.0065 (0.0050 – 0.0075)	0.0083 (0.0067 – 0.0094)	0.0034 (0.0022 – 0.0036)
Rad[c]	0.0059 ** (0.0034 – 0.0085)	0.0025 *** (0.0018 – 0.0034)	0.0037 *** (0.0028 – 0.0047)	0.0017 (0.00076 – 0.0024)
Con-1	0.0087 (0.0072 – 0.010)	0.0058 (0.0047 – 0.0066)	0.0074 (0.0062 – 0.0084)	0.0029 (0.0019 – 0.0034)
Rad-1	0.0049 ** (0.0017 – 0.0081)	0.0015 *** (0.00078 – 0.0024)	0.0026 *** (0.0016 – 0.0037)	0.0015 (0.0016 – 0.0033)
Con-2	0.0094 (0.0078 – 0.011)	0.0055 (0.0040 – 0.0061)	0.0065 (0.0053 – 0.0075)	0.0026 (0.0022 – 0.0042)
Rad-2	0.0052 ** (0.0029 – 0.0076)	0.0037 (0.0028 – 0.0048)	0.0055 (0.0044 – 0.0068)	0.0024 (0.0016 – 0.0028)
Con-3	0.015 (0.012 – 0.017)	0.0077 (0.0060 – 0.0094)	0.0094 (0.0076 – 0.011)	0.0032 (0.0018 – 0.0036)
Rad-3	0.0042 ** (0.0026 – 0.0057)	0.0030 *** (0.0022 – 0.0036)	0.0045 *** (0.0035 – 0.0052)	0.0022 (0.0017 – 0.0031)
Con-4	0.0145 (0.011 – 0.018)	0.0069 (0.0054 – 0.0084)	0.0084 (0.0069 – 0.010)	0.0027 (0.0025 – 0.0042)
Rad-4	0.0032 ** (0.0022 – 0.0043)	0.0031 *** (0.0022 – 0.0037)	0.0046 *** (0.0036 – 0.0054)	0.0024 [e] (0.0011 – 0.0024)
Con-5	0.011 (0.0095 – 0.013)	0.0075 (0.0061 – 0.0088)	0.0093 (0.0078 – 0.011)	0.0033 (0.0020 – 0.0037)
Rad-5	0.0041 ** (0.0017 – 0.0064)	0.0017 *** (0.00095 – 0.0022)	0.0028 *** (0.0020 – 0.0035)	0.0017 (0.0016 – 0.0029)
Con-6	0.012 (0.0094 – 0.014)	0.0055 (0.0040 – 0.0064)	0.0067 (0.0053 – 0.0078)	0.0029 (0.0020 – 0.0038)
Rad-6	0.0041 ** (0.0025 – 0.0058)	0.0039 (0.0030 – 0.0048)	0.0056 (0.0046 – 0.0067)	0.0022 (0.0016 – 0.0030)
Con-7	0.014 (0.011 – 0.016)	0.0062 (0.0046 – 0.0073)	0.0075 (0.0059 – 0.0087)	0.0029 (0.0024 – 0.0040)
Rad-7	0.0029 ** (0.0017 – 0.0041)	0.0034 *** (0.0026 – 0.0043)	0.0050 (0.0041 – 0.0060)	0.0023 (0.0012 – 0.0025)

[a] – Variance calculated as standard deviation multiplied by a factor of two
[b] – 95% CIs estimated from 10,000 bootstrap replicates
[c] – Mutational type defined by CRI score
* – All nonsynonymous mutational types had significantly lower values ($p < 0.05$ after correcting for multiple comparisons) than synonymous sites for all estimates of polymorphism and substitution
** – Radical substitution rate was significantly lower ($p < 0.0001$) than conservative substitution rate
*** – Nucleotide diversity was significantly lower ($p < 0.05$ after correcting for multiple comparisons) at radical sites than at conservative sites

bioRxiv preprint doi: https://doi.org/10.1101/049924; this version posted February 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table 2. Substitution rate and nucleotide diversity across mutational types in *P. antipodarum* mitochondrial genomes. For each mutational type, we estimated branch-specific substitution rates (K), nucleotide diversity (π and θ), and private allele diversity (θ_U) using seven distinct amino acid classification schemes plus a model-averaged classifier (CRI).

bioRxiv preprint doi: https://doi.org/10.1101/049924; this version posted February 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table 3. McDonald-Kreitman tests of selection across mutational types in *P. antipodarum* mitochondrial genomes.

	Syn	Nsyn (Nsyn/Syn)	Con[a] (Con/Syn)	Rad[a] (Rad/Syn)
Polymorphism	651	211 (0.32)	151 (0.23)	60 (0.092)
Divergence	814	35 (0.043)	30 (0.037)	5 (0.0061)
FET *vs.* Syn	-	$p < 0.0001$	$p < 0.0001$	$p < 0.0001$
FET *vs.* Con	-	$p = 0.59$	-	$p = 0.098$

[a] – Mutational type defined by CRI score

bioRxiv preprint doi: https://doi.org/10.1101/049924; this version posted February 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table 4. Population genetic comparisons of conservative and radical mutations in sexual vs. asexual lineages of *P. antipodarum.*

	Con[a]			Rad[a]		
	Sex	Asex	p^b	Sex	Asex	p^b
K/K_S (variance)	0.026 (0.024 - 0.029)	0.026 (0.022 - 0.030)	0.95	0.0048 (0.0036 – 0.0059)	0.0063 (0.0039 – 0.0087)	0.00065

	Con[a]				Rad[a]			
	Sex	Asex	D	p^c	Sex	Asex	D	p^c
Polymorphisms	53	113	-	-	13	54	-	-
π/π_S (variance)	0.11 (0.058 - 0.16)	0.099 (0.036 - 0.061)	-0.0075	0.91	0.030 (0.017 - 0.043)	0.043 (0.027 - 0.059)	0.013	0.049
θ/θ_S (variance)	0.12 (0.070 - 0.17)	0.13 (0.061 - 0.14)	0.0070	0.70	0.029 (0.018 - 0.041)	0.068 (0.049 - 0.087)	0.039	0.0015
$\theta_U/\theta_{U\text{-}S}$ (variance)	0.25 (0.14 - 0.36)	0.53 (0.35 - 0.70)	0.28	0.0031	0.055 (0.033 - 0.077)	0.33 (0.23 - 0.43)	0.27	0.0059

[a] - Mutational type defined by CRI score
[b] - Mann-Whitney U test
[c] - One-tailed test of significance ($H_{A:}$ $D > 0$)

bioRxiv preprint doi: https://doi.org/10.1101/049924; this version posted February 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE LEGENDS

Figure 1. Whole-mitochondrial genome phylogeny for *P. antipodarum* and an outgroup, *P. estuarinus*. Consensus phylogenetic tree depicting evolutionary relationships of 23 asexual (red) and eight sexual (blue) *P. antipodarum* mitochondrial genomes. Tree topology was inferred using Maximum Likelihood (ML, performed in PAUP*) and Bayesian (performed in MrBayes) methods. Two models of molecular evolution identified by jModelTest v 2.0 (GTR + I + G, TIM3 + I + G) were used to infer ML-based trees, while the model of molecular evolution was directly inferred from the alignment over 1.0 x 10^6 generations for the Bayesian (MrBayes) tree. Branch support from the GTR + I + G ML approach and from the Bayesian consensus tree is listed below branches (bootstrap support/posterior probability). Asterisks indicate nodes for which both ML models exhibited ≥ 90 bootstrap support. Numbers above each line represent the numbers of non-homoplasious mutational changes, with the ratio of synonymous: nonsynonymous changes in plain font and the ratio of conservative: radical changes in bold font and parentheses. Clade A represents a particularly distinct group of asexual lineages.

Figure 2. Molecular evolution and population genetics of synonymous, nonsynonymous, conservative, and radical mutations in *P. antipodarum* mitochondrial genomes. a) Branch-specific estimates of synonymous-corrected substitution rates for all nonsynonymous changes (black), conservative nonsynonymous changes (blue), and radical nonsynonymous changes (orange), as identified by the model-averaged amino acid classification scheme (CRI). Error bars represent standard

bioRxiv preprint doi: https://doi.org/10.1101/049924; this version posted February 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

deviations of branch-specific substitution rates. Lower-case letters reflect statistical groupings based on pairwise Mann-Whitney U tests ($p < 0.017$). b) Ratios of polymorphism to divergence for synonymous, all nonsynonymous, conservative nonsynonymous, and radical nonsynonymous changes. Lower-case letters reflect statistical groupings based on pairwise Fisher's Exact Tests ($p < 0.013$). c) Nucleotide diversity estimated using π (circles) and θ (triangles) for all nonsynonymous (black), conservative nonsynonymous (blue), and radical nonsynonymous (orange) sites. Error bars reflect variance calculated as in Durrett (2008). Radical nonsynonymous sites exhibit significantly lower levels of nucleotide diversity than conservative sites ($p < 0.0002$) for both measures of polymorphism. d) Site frequency spectra of synonymous (white), nonsynonymous (black), conservative nonsynonymous (blue), and radical nonsynonymous (orange) polymorphisms within *P. antipodarum*.

Figure 3. Molecular evolution of conservative and radical amino acid substitutions in *P. antipodarum*. Branch-specific estimates of a) synonymous-corrected conservative nonsynonymous substitution rates did not differ across sexual (blue) and asexual (red) lineages ($W = 95$, $p = 0.91$), but b) synonymous-corrected radical nonsynonymous substitution rates were significantly higher in asexual lineages than in sexual lineages ($W = 168$, $p = 0.00065$).

Figure 4. Population genetic comparisons of conservative and radical amino acid polymorphism in sexual (blue) vs. asexual (red) lineages of *P. antipodarum*. a) π_N/π_S (circles) and θ_N/θ_S (triangles) in sexual *vs.* asexual lineages for conservative and

bioRxiv preprint doi: https://doi.org/10.1101/049924; this version posted February 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

radical sites. b) π_N/π_S (circles) and θ_N/θ_S (triangles) in sexual *vs.* asexual lineages for conservative and radical sites. Only lineages from clade C were used in this analysis. c) Private polymorphism ($\theta_{U\text{-}N}/\theta_{U\text{-}S}$) in sexual *vs*. asexual lineages for conservative and radical sites. Error bars reflect variance calculated as in Durrett (2008). For all comparisons, one-tailed *p* values (*i.e.*, Asex > Sex) are indicated by asterisks (* – $p < 0.05$, ** – $p < 0.01$, *** – $p < 0.001$).

bioRxiv preprint doi: https://doi.org/10.1101/049924; this version posted February 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Literature Cited

Adrion, J. R., P. S. White, and K. L. Montooth. 2016. The roles of compensatory evolution and constraint in aminoacyl tRNA synthetase evolution. Mol Biol Evol 33:152-161.

Andolfatto, P., K. M. Wong, and D. Bachtrog. 2011. Effective population size and the efficacy of selection on the X chromosomes of two closely related *Drosophila* species. Genome Biol Evol 3:114-128.

Anfinsen, C. B. 1973. Principles that govern the folding of protein chains. Science 181:223-230.

Arenas, M. and D. Posada. 2010a. Coalescent simulation of intracodon recombination. Genetics 184:429-437.

Arenas, M. and D. Posada. 2010b. The effect of recombination on the reconstruction of ancestral sequences. Genetics 184:1133-1139.

Arenas, M. and D. Posada. 2014. The influence of recombination on the estimation of selection from coding sequence alignments. Natural Selection: Methods and Applications:112-125.

Ballard, J. W. O. 2000. Comparative genomics of mitochondrial DNA in members of the *Drosophila melanogaster* subgroup. J Mol Evol 51:48-63.

Barr, C. M., M. Neiman, and D. R. Taylor. 2005. Inheritance and recombination of mitochondrial genomes in plants, fungi and animals. New Phytol 168:39-50.

Barreto, F. S. and R. S. Burton. 2013. Evidence for compensatory evolution of ribosomal proteins in response to rapid divergence of mitochondrial rRNA. Mol Biol Evol 30:310-314.

bioRxiv preprint doi: https://doi.org/10.1101/049924; this version posted February 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Bashford, D., C. Chothia, and A. M. Lesk. 1987. Determinants of a protein fold - unique features of the globin amino acid sequences. J Mol Biol 196:199-216.

Beck, E. A., A. C. Thompson, J. Sharbrough, E. Brud, and A. Llopart. 2015. Gene flow between *Drosophila yakuba* and *Drosophila santomea* in subunit V of cytochrome c oxidase: A potential case of cytonuclear cointrogression. Evolution 69:1973-1986.

Birky, C. W., Jr. and J. B. Walsh. 1988. Effects of linkage on rates of molecular evolution. Proc Natl Acad Sci USA 85:6414-6418.

Blier, P. U., F. Dufresne, and R. S. Burton. 2001. Natural selection and the evolution of mtDNA-encoded peptides: evidence for intergenomic co-adaptation. Trends Genet 17:400-406.

Boyko, A. R., S. H. Williamson, A. R. Indap, J. D. Degenhardt, R. D. Hernandez, K. E. Lohmueller, M. D. Adams, S. Schmidt, J. J. Sninsky, S. R. Sunyaev, T. J. White, R. Nielsen, A. G. Clark, and C. D. Bustamante. 2008. Assessing the evolutionary impact of amino acid mutations in the human genome. PLoS Genet 4:e1000083.

Browne, R. A. 1992. Population genetics and ecology of *Artemia*: insights into parthenogenetic reproduction. Trends Ecol Evol 7:232-237.

Burley, S. K. and G. A. Petsko. 1985. Aromatic-aromatic interaction: a mechanism of protein structure stabilization. Science 229:23-28.

Burton, R. S. and F. S. Barreto. 2012. A disproportionate role for mtDNA in Dobzhansky-Muller incompatibilities? Mol Ecol 21:4942-4957.

Charlesworth, B., M. T. Morgan, and D. Charlesworth. 1993. The effect of deleterious mutations on neutral molecular variation. Genetics 134:1289-1303.

bioRxiv preprint doi: https://doi.org/10.1101/049924; this version posted February 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Charlesworth, D. and S. I. Wright. 2001. Breeding systems and genome evolution. Curr Opin Genet Dev 11:685-690.

Christensen, B., M. Hvilsom, and B. V. Pedersen. 1992. Genetic variation in coexisting sexual diploid and parthenogenetic triploid forms of *Fridericia galba* (Enchytraeidae, Oligochaeta) in a heterogeneous environment. Hereditas 117:153-162.

Cooper, B. S., C. R. Burrus, C. Ji, M. W. Hahn, and K. L. Montooth. 2015. Similar efficacies of selection shape mitochondrial and nuclear genes in both *Drosophila melanogaster* and *Homo sapiens*. G3 5:2165-2176.

Covello, P. S. and M. W. Gray. 1989. RNA editing in plant mitochondria. Nature 341:662-666.

Darriba, D., G. L. Taboada, R. Doallo, and D. Posada. 2012. jModelTest 2: more models, new heuristics and parallel computing. Nat Methods 9:772.

Doms, R. W., A. Ruusala, C. Machamer, J. Helenius, A. Helenius, and J. K. Rose. 1988. Differential effects of mutations in 3 domains on folding, quaternary structure, and intracellular-transport of vesicular Stomatitis Virus G-Protein. J Cell Biol 107:89-99.

Dowling, D. K., U. Friberg, and J. Lindell. 2008. Evolutionary implications of non-neutral mitochondrial genetic variation. Trends Ecol Evol 23:546-554.

Durrett, R. 2008. Probability models for DNA sequence evolution. Springer, New York ; London.

Eyre-Walker, A. and P. D. Keightley. 2007. The distribution of fitness effects of new mutations. Nat Rev Genet 8:610-618.

bioRxiv preprint doi: https://doi.org/10.1101/049924; this version posted February 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Eyre-Walker, A. and P. D. Keightley. 2009. Estimating the rate of adaptive molecular evolution in the presence of slightly deleterious mutations and population size change. Mol Biol Evol 26:2097-2108.

Fisher, R. A. 1930. The genetical theory of natural selection. The Clarendon Press, Oxford, UK.

Freudenberg-Hua, Y., J. Freudenberg, N. Kluck, S. Cichon, P. Propping, and M. M. Nothen. 2003. Single nucleotide variation analysis in 65 candidate genes for CNS disorders in a representative sample of the European population. Genome Res 13:2271-2276.

Fu, Y. X. and W. H. Li. 1993. Statistical tests of neutrality of mutations. Genetics 133:693-709.

Fukami, H., A. F. Budd, D. R. Levitan, J. Jara, R. Kersanach, and N. Knowlton. 2004. Geographic differences in species boundaries among members of the *Montastraea annularis* complex based on molecular and morphological markers. Evolution 58:324-337.

Gabriel, W., M. Lynch, and R. Burger. 1993. Muller's ratchet and mutational meltdowns. Evolution 47:1744-1757.

Geisler, N. and K. Weber. 1982. The amino acid sequence of chicken muscle desmin provides a common structural model for intermediate filament proteins. Embo J 1:1649-1656.

Gordon, A. and G. Hannon. 2010. Fastx-toolkit. FASTQ/A short-reads preprocessing tools (unpublished) http://hannonlab.cshl.edu/fastx_toolkit.

bioRxiv preprint doi: https://doi.org/10.1101/049924; this version posted February 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Grantham, R. 1974. Amino acid difference formula to help explain protein evolution. Science 185:862-864.

Gualberto, J. M., L. Lamattina, G. Bonnard, J. H. Weil, and J. M. Grienenberger. 1989. RNA editing in wheat mitochondria results in the conservation of protein sequences. Nature 341:660-662.

Halligan, D. L., F. Oliver, J. Guthrie, K. C. Stemshorn, B. Harr, and P. D. Keightley. 2011. Positive and negative selection in murine ultraconserved noncoding elements. Mol Biol Evol 28:2651-2660.

Harrison, J. S. and R. S. Burton. 2006. Tracing hybrid incompatibilities to single amino acid substitutions. Mol Biol Evol 23:559-564.

Havird, J. C., S. W. Fitzpatrick, J. Kronenberger, W. C. Funk, L. M. Angeloni, and D. B. Sloan. 2016. Sex, mitochondria, and genetic rescue. Trends Ecol Evol 31:96-99.

Havird, J. C., N. S. Whitehill, C. D. Snow, and D. B. Sloan. 2015. Conservative and compensatory evolution in oxidative phosphorylation complexes of angiosperms with highly divergent rates of mitochondrial genome evolution. Evolution 69:3069-3081.

Henry, L., T. Schwander, and B. J. Crespi. 2012. Deleterious mutation accumulation in asexual *Timema* stick insects. Mol Biol Evol 29:401-408.

Hill, G. E. 2016. Mitonuclear coevolution as the genesis of speciation and the mitochondrial DNA barcode gap. Ecol Evol 6:5831-5842.

Hill, W. G. and A. Robertson. 1966. The effect of linkage on limits to artificial selection. Genet Res 8:269-294.

bioRxiv preprint doi: https://doi.org/10.1101/049924; this version posted February 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Hollister, J. D., S. Greiner, W. Wang, J. Wang, Y. Zhang, G. K. Wong, S. I. Wright, and M. T. Johnson. 2015. Recurrent loss of sex is associated with accumulation of deleterious mutations in *Oenothera*. Mol Biol Evol 32:896-905.

Holm, S. 1979. A simple sequentially rejective multiple test procedure. Scand J Stat 6:65-70.

Jokela, J., C. M. Lively, M. F. Dybdahl, and J. A. Fox. 1997. Evidence for a cost of sex in the freshwater snail *Potamopyrgus antipodarum*. Ecology 78:452-460.

Jukes, T. H. and C. R. Cantor. 1969. Evolution of protein molecules. Mammalian protein metabolism 3:132.

Kaiser, V. B. and B. Charlesworth. 2009. The effects of deleterious mutations on evolution in non-recombining genomes. Trends Genet 25:9-12.

Katju, V., L. B. Packard, L. Bu, P. D. Keightley, and U. Bergthorsson. 2015. Fitness decline in spontaneous mutation accumulation lines of *Caenorhabditis elegans* with varying effective population sizes. Evolution 69:104-116.

Katoh, K. and D. M. Standley. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol 30:772-780.

Keightley, P. D. and B. Charlesworth. 2005. Genetic instability of *C. elegans* comes naturally. Trends Genet 21:67-70.

Kondrashov, A. S. 1994. Muller's ratchet under epistatic selection. Genetics 136:1469-1473.

Konrad, A., O. Thompson, R. H. Waterston, D. G. Moerman, P. D. Keightley, U. Bergthorsson, and V. Katju. 2017. Mitochondrial mutation rate, spectrum and

bioRxiv preprint doi: https://doi.org/10.1101/049924; this version posted February 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

heteroplasmy in *Caenorhabditis elegans* spontaneous mutation accumulation lines of differing population size. Mol Biol Evol 34:1319-1334.

Krist, A. C., A. D. Kay, K. Larkin, and M. Neiman. 2014. Response to phosphorus limitation varies among lake populations of the freshwater snail *Potamopyrgus antipodarum*. PLoS One 9:e85845.

Kumar, S., M. Nei, J. Dudley, and K. Tamura. 2008. MEGA: a biologist-centric software for evolutionary analysis of DNA and protein sequences. Brief Bioinformat 9:299-306.

Lenski, R. E. 2017. Experimental evolution and the dynamics of adaptation and genome evolution in microbial populations. ISME J 11:2181-2194.

Lesk, A. M. and C. Chothia. 1980. How different amino acid sequences determine similar protein structures - structure and evolutionary dynamics of the globins. J Mol Biol 136:225-&.

Li, W. H., C. I. Wu, and C. C. Luo. 1985. A new method for estimating synonymous and nonsynonymous rates of nucleotide substitution considering the relative likelihood of nucleotide and codon changes. Mol Biol Evol 2:150-174.

Lively, C. M. 1987. Evidence from a New Zealand snail for the maintenance of sex by parasitism. Nature 328:519-521.

Lovell, J. T., R. J. Williamson, S. I. Wright, J. K. McKay, and T. F. Sharbel. 2017. Mutation accumulation in an asexual relative of *Arabidopsis*. Plos Genet 13:e1006550.

Lynch, M., R. Burger, D. Butcher, and W. Gabriel. 1993. The mutational meltdown in asexual populations. J Hered 84:339-344.

bioRxiv preprint doi: https://doi.org/10.1101/049924; this version posted February 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Maddison, W. P. 1997. Gene trees in species trees. Syst Biol 46:523-536.

Mallo, D., A. Sánchez‐Cobos, and M. Arenas. 2016. Diverse considerations for successful phylogenetic tree reconstruction: impacts from model misspecification, recombination, homoplasy, and pattern recognition. Pattern Recognition in Computational Molecular Biology: Techniques and Approaches:439-456.

Marshall, H. D., M. W. Coulson, and S. M. Carr. 2009. Near neutrality, rate heterogeneity, and linkage govern mitochondrial genome evolution in Atlantic cod (*Gadus morhua*) and other gadine fish. Mol Biol Evol 26:579-589.

Martin, D. P., B. Murrell, A. Khoosal, and B. Muhire. 2017. Detecting and analyzing genetic recombination using RDP4. Methods Mol Biol 1525:433-460.

Maynard Smith, J. 1978. The evolution of sex. Cambridge University Press, Cambridge Eng.; New York.

McDonald, J. H. and M. Kreitman. 1991. Adaptive protein evolution at the *Adh* locus in *Drosophila*. Nature 351:652-654.

Meiklejohn, C. D., K. L. Montooth, and D. M. Rand. 2007. Positive and negative selection on the mitochondrial genome. Trends Genet 23:259-263.

Montooth, K. L., C. D. Meiklejohn, D. N. Abt, and D. M. Rand. 2010. Mitochondrial-nuclear epistasis affects fitness within species but does not contribute to fixed incompatibilities between species of *Drosophila*. Evolution 64:3364-3379.

Nakashima, H., K. Nishikawa, and T. Ooi. 1986. The folding type of a protein is relevant to the amino acid composition. J Biochem 99:153-162.

Nei, M. and W. H. Li. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc Natl Acad Sci USA 76:5269-5273.

bioRxiv preprint doi: https://doi.org/10.1101/049924; this version posted February 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Neiman, M., G. Hehman, J. T. Miller, J. M. Logsdon, Jr., and D. R. Taylor. 2010. Accelerated mutation accumulation in asexual lineages of a freshwater snail. Mol Biol Evol 27:954-963.

Neiman, M., K. Larkin, A. R. Thompson, and P. Wilton. 2012. Male offspring production by asexual *Potamopyrgus antipodarum*, a New Zealand snail. Heredity 109:57-62.

Neiman, M. and T. A. Linksvayer. 2006. The conversion of variance and the evolutionary potential of restricted recombination. Heredity 96:111-121.

Neiman, M. and C. M. Lively. 2004. Pleistocene glaciation is implicated in the phylogeographical structure of *Potamopyrgus antipodarum*, a New Zealand snail. Mol Ecol 13:3085-3098.

Neiman, M., D. Paczesniak, D. M. Soper, A. T. Baldwin, and G. Hehman. 2011. Wide variation in ploidy level and genome size in a New Zealand freshwater snail with coexisting sexual and asexual lineages. Evolution 65:3202-3216.

Neiman, M. and D. R. Taylor. 2009. The causes of mutation accumulation in mitochondrial genomes. Proc Roy Soc Lond B 276:1201-1209.

Niehuis, O., A. K. Judson, and J. Gadau. 2008. Cytonuclear genic incompatibilities cause increased mortality in male F2 hybrids of *Nasonia giraulti* and *N. vitripennis*. Genetics 178:413-426.

Normark, B. B. and N. A. Moran. 2000. Opinion - Testing for the accumulation of deleterious mutations in asexual eukaryote genomes using molecular sequences. J Nat Hist 34:1719-1729.

bioRxiv preprint doi: https://doi.org/10.1101/049924; this version posted February 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Ohta, T. 1987. Very slightly deleterious mutations and the molecular clock. J Mol Evol 26:1-6.

Osada, N. and H. Akashi. 2012. Mitochondrial-nuclear interactions and accelerated compensatory evolution: evidence from the primate cytochrome C oxidase complex. Mol Biol Evol 29:337-346.

Paczesniak, D., J. Jokela, K. Larkin, and M. Neiman. 2013. Discordance between nuclear and mitochondrial genomes in sexual and asexual lineages of the freshwater snail *Potamopyrgus antipodarum*. Mol Ecol 22:4695-4710.

Perutz, M. F., J. C. Kendrew, and H. C. Watson. 1965. Structure and function of haemoglobin II: Some relations between polypeptide chain configuration and amino acid sequence. J Mol Biol 13:669-678.

Posada, D. 2003. Using MODELTEST and PAUP* to select a model of nucleotide substitution. Curr Protoc Bioinformatics Chapter 6:Unit 6 5.

R Core Team. 2016. R: a language and environment for statistical computing. R Development Core Team, Vienna.

Rambaut, A. 2012. FigTree v1. 4. Molecular evolution, phylogenetics and epidemiology. Edinburgh, UK: University of Edinburgh, Institute of Evolutionary Biology.

Rand, D. M., R. A. Haney, and A. J. Fry. 2004. Cytonuclear coevolution: the genomics of cooperation. Trends Ecol Evol 19:645-653.

Rokas, A., E. Ladoukakis, and E. Zouros. 2003. Animal mitochondrial DNA recombination revisited. Trends Ecol Evol 18:411-417.

Ronquist, F., M. Teslenko, P. van der Mark, D. L. Ayres, A. Darling, S. Hohna, B. Larget, L. Liu, M. A. Suchard, and J. P. Huelsenbeck. 2012. MrBayes 3.2: efficient

bioRxiv preprint doi: https://doi.org/10.1101/049924; this version posted February 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Bayesian phylogenetic inference and model choice across a large model space. Syst Biol 61:539-542.

Rumbley, J., L. Hoang, L. Mayne, and S. W. Englander. 2001. An amino acid code for protein folding. Proc Natl Acad Sci USA 98:105-112.

Sardiello, M., F. Licciulli, D. Catalano, M. Attimonelli, and C. Caggese. 2003. MitoDrome: a database of *Drosophila melanogaster* nuclear genes encoding proteins targeted to the mitochondrion. Nucl Acids Res 31:322-324.

Schierup, M. H. and J. Hein. 2000. Consequences of recombination on traditional phylogenetic analysis. Genetics 156:879-891.

Sharbrough, J., J. L. Cruise, M. Beetch, N. M. Enright, and M. Neiman. 2017. Genetic variation for mitochondrial function in the New Zealand freshwater snail *Potamopyrgus antipodarum*. J Hered 108:759-768.

Sloan, D. B., J. C. Havird, and J. Sharbrough. 2017. The on-again, off-again relationship between mitochondrial genomes and species boundaries. Mol Ecol 26:2212-2236.

Smith, N. G. 2003. Are radical and conservative substitution rates useful statistics in molecular evolution? J Mol Evol 57:467-478.

Stratton, D. A. 1991. Life-history variation within populations of an asexual plant, *Erigeron annuus* (Asteraceae). Am J Bot 78:723-728.

Stratton, D. A. 1992. Life cycle components of selection in *Erigeron annuus*: Ii. genetic variation. Evolution 46:107-120.

bioRxiv preprint doi: https://doi.org/10.1101/049924; this version posted February 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Swofford, D. L. and J. Sullivan. 2009. Phylogeny inference based on parsimony and other methods using PAUP*. Phylogenetic Handbook: A Practical Approach to Phylogenetic Analysis and Hypothesis Testing, 2nd Edition:267-312.

Tajima, F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123:585-595.

Tavaré, S. 1986. Some probabilistic and statistical problems in the analysis of DNA sequences. Lectures on Mathematics in the Life Sciences 17:57-86.

Toews, D. P. and A. Brelsford. 2012. The biogeography of mitochondrial and nuclear discordance in animals. Mol Ecol 21:3907-3930.

Tsaousis, A. D., D. P. Martin, E. D. Ladoukakis, D. Posada, and E. Zouros. 2005. Widespread recombination in published animal mtDNA sequences. Mol Biol Evol 22:925-933.

Voigt-Zielinski, M. L., M. Piwczynski, and T. F. Sharbel. 2012. Differential effects of polyploidy and diploidy on fitness of apomictic *Boechera*. Sex Plant Reprod 25:97-109.

von Heijne, G. 1992. Membrane protein structure prediction. Hydrophobicity analysis and the positive-inside rule. J Mol Biol 225:487-494.

Watterson, G. A. 1975. On the number of segregating sites in genetical models without recombination. Theor Popul Biol 7:256-276.

Weeks, A. R. and A. A. Hoffmann. 1998. Intense selection of mite clones in a heterogeneous environment. Evolution 52:1325-1333.

bioRxiv preprint doi: https://doi.org/10.1101/049924; this version posted February 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Werle, E., C. Schneider, M. Renner, M. Volker, and W. Fiehn. 1994. Convenient single-step, one tube purification of PCR products for direct sequencing. Nucl Acids Res 22:4354-4355.

Wickham, H. 2016. ggplot2: elegant graphics for data analysis. Springer.

Wright, C. F., S. A. Teichmann, J. Clarke, and C. M. Dobson. 2005. The importance of sequence diversity in the aggregation and evolution of proteins. Nature 438:878-881.

Wright, S. I., N. Nano, J. P. Foxe, and V. U. Dar. 2008. Effective population size and tests of neutrality at cytoplasmic genes in *Arabidopsis*. Genet Res 90:119-128.

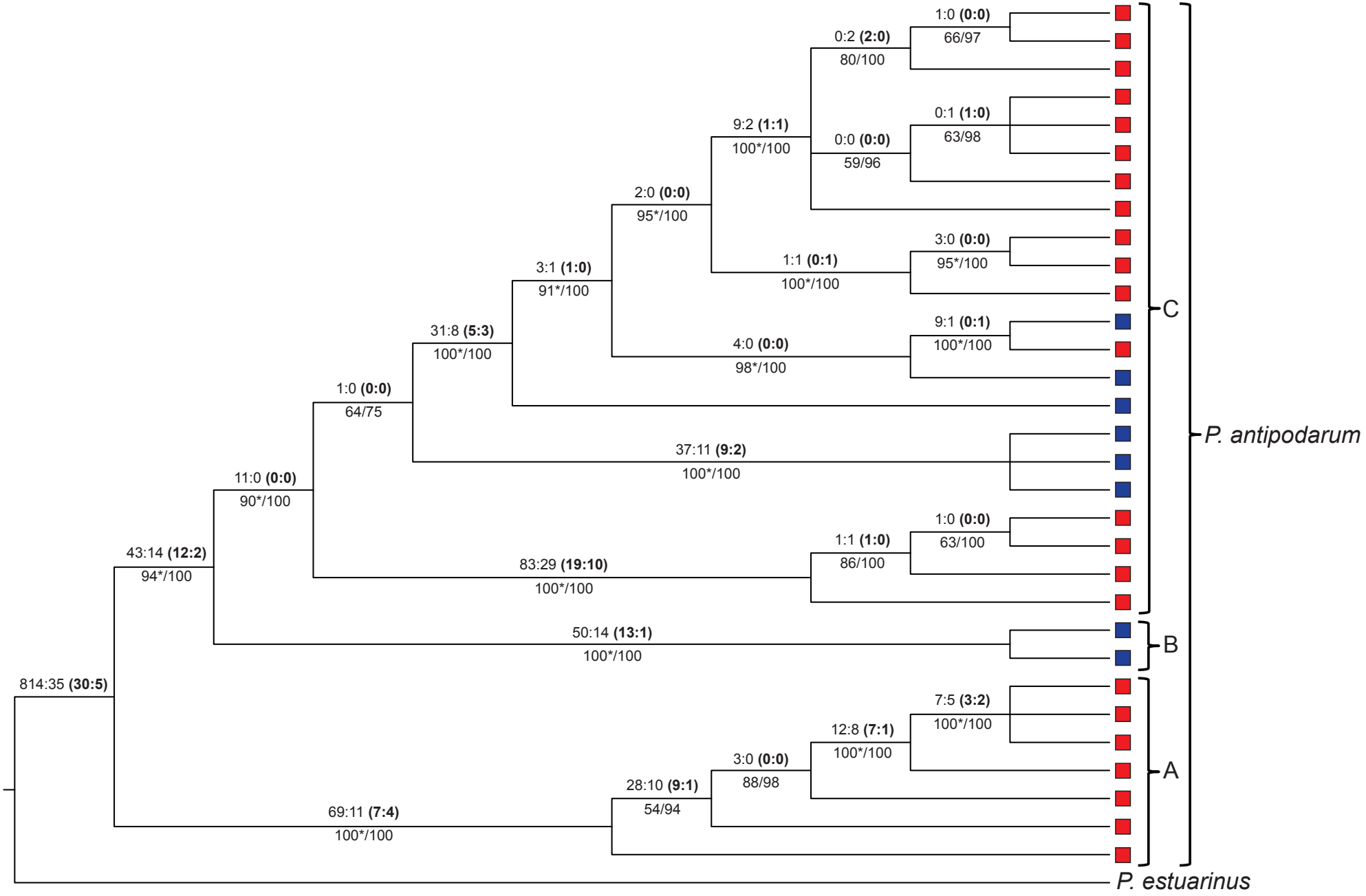
Xia, X. 1998. The rate heterogeneity of nonsynonymous substitutions in mammalian mitochondrial genes. Mol Biol Evol 15:336-344.

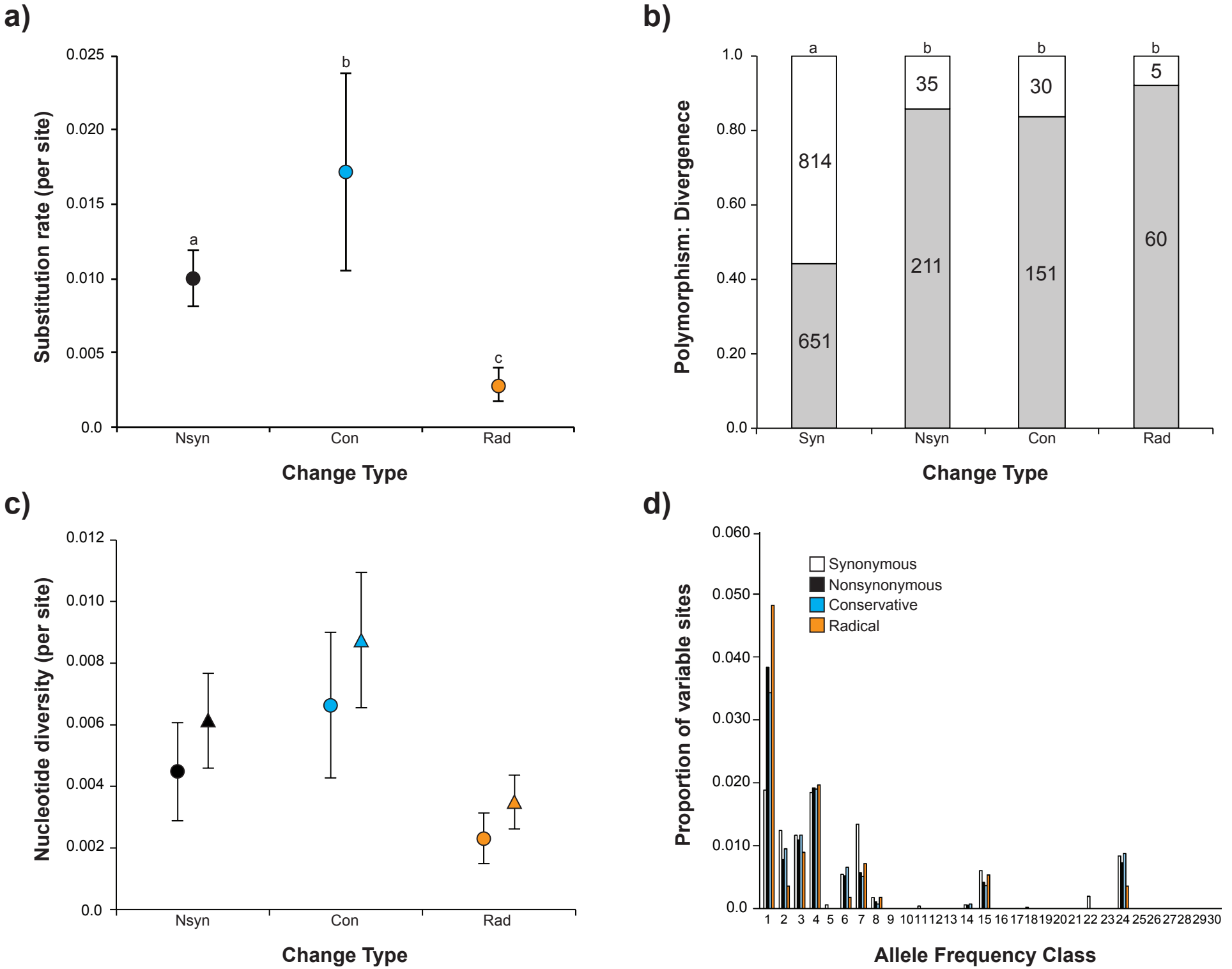
Xue, D., J. Yin, M. Tan, J. Yue, Y. Wang, and L. Liang. 2008. Prediction of functional nonsynonymous single nucleotide polymorphisms in human G-protein-coupled receptors. J Hum Genet 53:379-389.

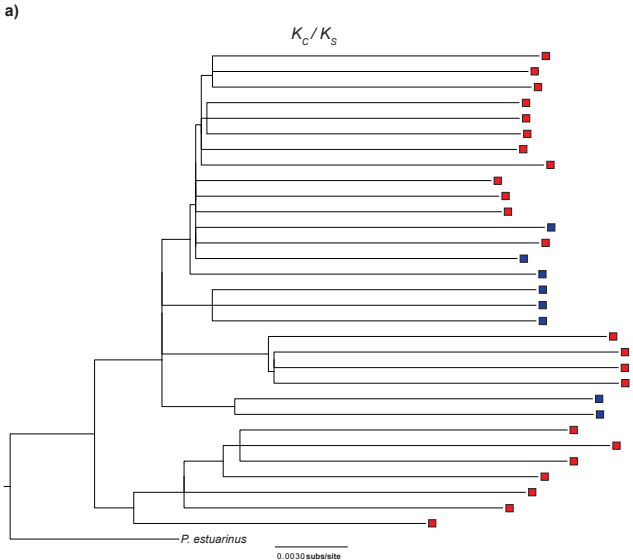
Yokobori, S. and S. Paabo. 1995. Transfer RNA editing in land snail mitochondria. Proc Natl Acad Sci USA 92:10432-10435.

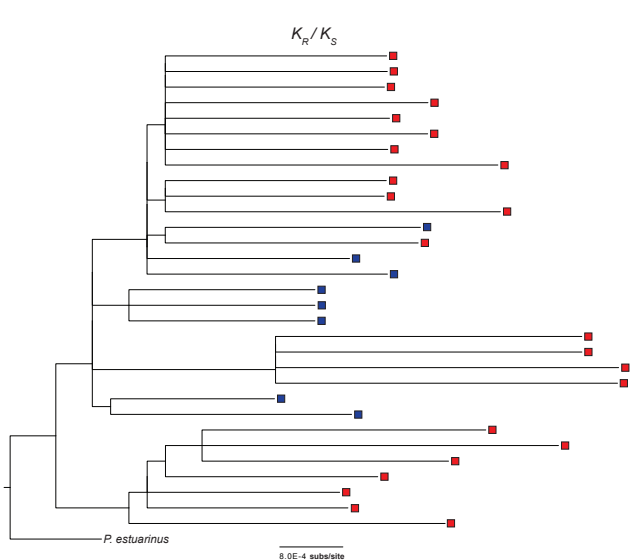
Zhai, W. W., R. Nielsen, and M. Slatkin. 2009. An investigation of the statistical power of neutrality tests based on comparative and population genetic data. Mol Biol Evol 26:273-283.

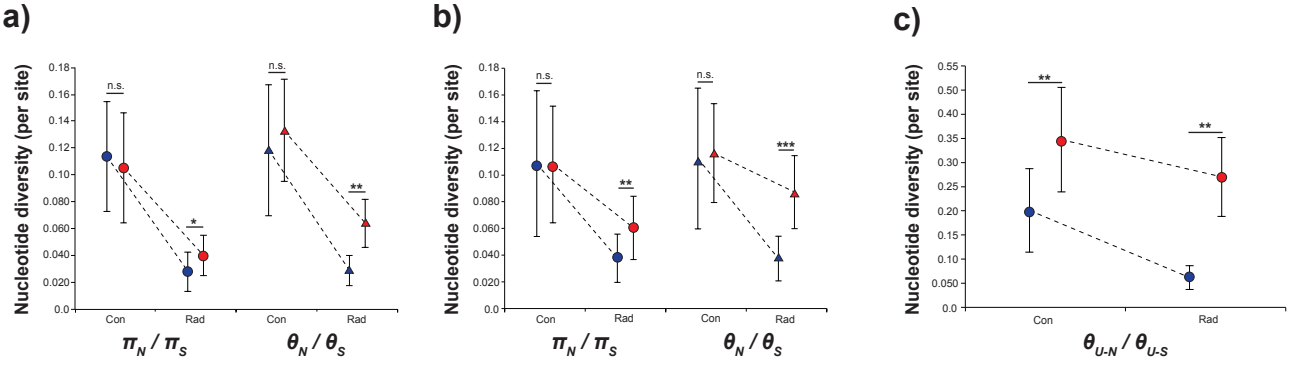
Zhang, J. 2000. Rates of conservative and radical nonsynonymous nucleotide substitutions in mammalian nuclear genes. J Mol Evol 50:56-68.


1:0 (0:0)
66/97
0:2 (2:0)
80/100
0:1 (1:0)
63/98
9:2 (1:1)
100*/100
0:0 (0:0)
59/96
2:0 (0:0)
95*/100
3:0 (0:0)
95*/100
1:1 (0:1)
100*/100
3:1 (1:0)
91*/100
9:1 (0:1)
100*/100
31:8 (5:3)
100*/100
4:0 (0:0)
98*/100
1:0 (0:0)
64/75
37:11 (9:2)
100*/100
11:0 (0:0)
90*/100
1:0 (0:0)
63/100
1:1 (1:0)
86/100
43:14 (12:2)
94*/100
83:29 (19:10)
100*/100
50:14 (13:1)
100*/100
814:35 (30:5)
7:5 (3:2)
100*/100
12:8 (7:1)
100*/100
3:0 (0:0)
88/98
28:10 (9:1)
54/94
69:11 (7:4)
100*/100
C
B
A
P. antipodarum
P. estuarinus


a)
Substitution rate (per site)
0.025
0.020
0.015
0.010
0.005
0.0
a
b
c
Nsyn
Con
Rad
Change Type
b)
Polymorphism: Divergence
1.0
0.80
0.60
0.40
0.20
0.0
a
b
b
b
814
651
35
211
30
151
5
60
Syn
Nsyn
Con
Rad
Change Type
c)
Nucleotide diversity (per site)
0.012
0.010
0.008
0.006
0.004
0.002
0.0
Nsyn
Con
Rad
Change Type
d)
Proportion of variable sites
0.060
0.050
0.040
0.030
0.020
0.010
0.0
Synonymous
Nonsynonymous
Conservative
Radical
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30
Allele Frequency Class


a)
K_C / K_S
P. estuarinus
0.0030 subs/site


b)
K_R / K_S
P. estuarinus
8.0E-4 subs/site


a)
b)
c)
Nucleotide diversity (per site)
0.18
0.16
0.14
0.12
0.10
0.080
0.060
0.040
0.020
0.0
n.s.
*
**
***
Con
Rad
π_N / π_S
θ_N / θ_S
0.55
0.50
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.050
$\theta_{U\text{-}N} / \theta_{U\text{-}S}$