Parallel Rate Heterogeneity in Chloroplast and Mitochondrial Genomes of Brazil Nut Trees (Lecythidaceae) Is Consistent with Lineage Effects

David F. Soria-Hernanz,[1] John M. Braverman,* and Matthew B. Hamilton*†*

*Department of Biology, Georgetown University; and †Biological Dynamics of Forest Fragments Project, National Institute for Research in the Amazon, Manaus, Brazil

We investigated whether relative rates of divergence were correlated between the mitochondrial and chloroplast genomes as expected under lineage effects or were genome specific as expected with locus-specific effects. Five mitochondrial noncoding regions (*nad1B_C*, *nad4exon1_2*, *nad7exon2_3*, *nad7exon3_4*, and *rps14-cob*) for 21 samples from Lecythidaceae were sequenced. Three chloroplast regions (*rpl*20-5′*rps*12, *trn*S-*trn*G, and *psb*A-*trn*H) were sequenced to expand the taxa in an existing data set. Absolute rates of nucleotide and insertion and deletion (indel) changes were 13 times faster in the chloroplast genome than in the mitochondrial genome. Similar indel length frequency distributions for both organelles suggested that common mechanisms were responsible for generating indels. Molecular clock tests applied to phylogenetic trees estimated from mitochondrial and chloroplast sequences revealed global rate heterogeneity of nucleotide substitution. Maximum likelihood and Tajima's 1*D* relative rate tests show that *Lecythis zabucajo* exhibited a rate acceleration for both the mitochondrial and chloroplast sequences. Whereas *Eschweilera romeu-cardosoi* showed a significant rate slowdown for chloroplast sequences, the mitochondrial sequences for 3 *Eschweilera* taxa showed evidence for a rate slowdown only when compared with *L. zabucajo*. Significant rate heterogeneity was also observed for indel changes in the mitochondrial genome but not for the chloroplast. The lack of mitochondrial nucleotide changes for some taxa as well as chloroplast indel homoplasy may have limited the power of relative rate tests to detect rate variation. Relative ratio tests consistently indicated rate proportionality among branch lengths between the mitochondrial and chloroplast phylogenetic trees. The relative ratio tests showed that taxa possessing rate heterogeneity had parallel relative divergence rates in both mitochondrial and chloroplast sequences as expected under lineage effects. A neutral replication–dependent model of rate heterogeneity for both nucleotide and indel changes provides a simple explanation for common patterns of rate heterogeneity across the 2 organelle genomes in Lecythidaceae. The lineage effects observed here were uncoupled from annual/perennial habit because all the species from this study are perennial.

Introduction

The molecular clock hypothesis first proposed by Zuckerkandl and Pauling (1965) generated a debate about whether molecular divergence among species follows a constant rate through time. This debate has implications in 2 broad areas of molecular evolution (reviewed by Bromham and Penny 2003). First, rates of molecular evolution could be used to estimate time of divergence among species if there is a molecular clock that is constant over time for at least some loci or for some groups of species (reviewed by Arbogast et al. 2002). Second, testing for divergence rate constancy is a tool to distinguish among the evolutionary processes responsible for the fixation of new mutations, called substitutions, that are the basis of genetic divergence. The neutral theory (Kimura 1968, 1983) provides a null model of substitution by genetic drift as well as potential explanations for rate heterogeneity. Under neutrality, the rate of molecular divergence is expected to be equal to the mutation rate and independent of the effective population size. Rate heterogeneity under neutrality can be explained by differences in the underlying mutation rates among loci (manifested as locus effects) as well as by mutation rates at the same locus that are variable among lineages (manifested as lineage effects). When considering multiple independent loci, a lineage effect would be characterized by a proportional and correlated slowdown (or acceleration) in rates of divergence at all loci. Natural selection is also a potential explanation for variation in rates of molecular evolution because the fitness of new mutations affects their probability of fixation. Purifying or positive selection pressures on new mutations alter the probability of substitution and thereby impact the rate of molecular evolution (Gillespie 1991; Muse and Gaut 1997). However, it is considered very unlikely that natural selection would act in such a way as to produce correlated effects on rates of divergence among numerous independent loci or lineages. Therefore, a principal method to distinguish between neutral and selective causes of rate heterogeneity is to compare rates of divergence at multiple independent loci and in multiple independent lineages.

Lineage-specific rates of molecular evolution have been observed in both animals (e.g., Sarich and Wilson 1967; Wu and Li 1985; Britten 1986; Martin and Palumbi 1993; Bromham et al. 1996; Gissi et al. 2000; Bromham 2002) and plants (e.g., Gaut et al. 1992, 1997; Eyre-Walker and Gaut 1997; Muse and Gaut 1997; Andreasen and Baldwin 2001; Young and dePamphilis 2005; Kay et al. 2006). In plants, rates of molecular evolution have been studied primarily for chloroplast sequences (e.g., Gaut et al. 1992; Muse and Gaut 1997; Hamilton et al. 2003) because numerous universal primers are available for the chloroplast genome and less often for mitochondrial (Cho et al. 2004; Parkinson et al. 2005) and nuclear genomes (Gaut 1998; Muse 2000; Small et al. 2004). However, rates of divergence estimated with DNA sequences from a single organelle genome are not able to distinguish if lineage effects encompass multiple genomes or are limited to multiple regions within a single genome.

A limited number of studies have compared rates of divergence across multiple independent loci in plants. A very widely cited comparison of grass and palm taxa (all monocots) demonstrated a faster rate of synonymous

[1] Present address: The Genographic Project, National Geographic Society, Washington, DC.

Key words: chloroplast DNA, lineage effect, mitochondrial DNA, rate heterogeneity, relative rate test, relative ratio test.

E-mail: dsoria@ngs.org.

Mol. Biol. Evol. 25(7):1282–1296. 2008
doi:10.1093/molbev/msn074
Advance Access publication April 2, 2008

© The Author 2008. Published by Oxford University Press on behalf of the Society for Molecular Biology and Evolution. All rights reserved.
For permissions, please e-mail: journals.permissions@oxfordjournals.org

Downloaded from https://academic.oup.com/mbe/article-abstract/25/7/1282/1041900 by guest on 03 June 2020

substitutions in grasses that was correlated across loci in the chloroplast, mitochondrial, and nuclear genomes as expected for lineage effects (Eyre-Walker and Gaut 1997). Notably, rates of molecular evolution at nonsynonymous sites were not correlated across the independent loci, consistent instead with locus effects. Two more recent studies have identified heterogeneity in absolute rates of synonymous substitutions in the mitochondrial genome but not in the nuclear and chloroplast genomes among recently diverged species within the angiosperm genera *Pelargonium* (Parkinson et al. 2005) and *Plantago* (Cho et al. 2004). Both these studies showed uncorrelated rate heterogeneity between the mitochondrial genome and the other plant genomes consistent with locus effects. This mitochondrial rate variation was hypothesized to be caused by differences among taxa in the efficiency of DNA repair and replication fidelity.

Few studies have distinguished lineage effects across multiple genomes and lineage effects at the level of multiple loci within single genomes. To do this for plant organelle genomes, it is possible to build on previous studies that detected chloroplast genome rate heterogeneity with the addition of mitochondrial genome sequence data. However, identification and amplification of homologous mitochondrial regions among multiple taxa are difficult because plant mitochondrial genomes generally exhibit a high degree of structural variation (Wolfe et al. 1987; Palmer and Herbon 1988; Laroche et al. 1997; Gaut 1998; Muse 2000; Palmer et al. 2000; Lynch et al. 2006). An additional limitation is that the mitochondrial genome usually has the lowest rate of nucleotide change among the 3 genomes in plants, requiring a large sample of base pairs be collected in order to have power for comparisons such as relative rate tests (Bromham et al. 2000; Muse 2000). Sequencing noncoding mitochondrial regions is one strategy to increase the number of observed nucleotide changes because such regions are expected to be less constrained by function than coding regions and should thereby exhibit a higher proportion of DNA changes per sequence length (Laroche et al. 1997). Comparing divergence rates in chloroplast and mitochondrial genomes can also help to elucidate causes of molecular evolution in the plant mitochondrial genome, a plant genome that currently has poorly described patterns of molecular evolution (Palmer et al. 2000).

Nucleotide substitutions and insertion and deletion (indel) variation are the 2 types of DNA changes widely observed in plant organelle genomes. Indels are a major component in the evolution of noncoding DNA sequences (Britten et al. 2003) but relatively little is known about the mechanisms generating indels (Graham et al. 2000; Palmer et al. 2000). Similar indel length frequency distributions would be expected between chloroplast and mitochondrial genomes if similar mechanisms generate these sequence changes (Laroche et al. 1997). Comparing rates of nucleotide and indel sequence changes can also indirectly test whether the same evolutionary processes affect the origination and fixation of these 2 types of mutations. Proportional rates of nucleotide and indel changes within genomes are expected if both types of mutations are neutral and share a proportional origination (mutation) rate at the molecular level. Proportional divergence rates for nucleotides and indels between genomes are consistent with a replication rate–dependent model of neutral divergence where the mutation rate varies with the frequency of DNA replication (Ohta and Kimura 1971; Wu and Li 1985; Weinreich 2001). If nucleotide and indel mutations share an approximately equal origination rate but their probability of fixation varies due to different selection pressures, rates of nucleotide and indel changes would be uncoupled. Comparing rates of divergence for nucleotide and indel changes within and between organelle genomes is therefore a means to distinguish neutral evolution versus the action of natural selection as well as whether nucleotide and indel mutations evolve under common processes. This same logic underlies rate comparisons of synonymous and nonsynonymous nucleotide divergence within and among plant genomes (reviewed by Muse 2000).

Several processes have been hypothesized to cause lineage-specific mutation rates that result in lineage effects. Rate heterogeneity is frequently correlated with differences in life history features such as generation time, body size, and metabolic rate (e.g., Gaut et al. 1992; Martin and Palumbi 1993; Mooers and Harvey 1994; Bromham et al. 1996; Eyre-Walker and Gaut 1997; Muse and Gaut 1997; Andreasen and Baldwin 2001; Bromham 2002; Kay et al. 2006); differences in biochemical features such as efficiency of DNA repair machinery (e.g., Britten 1986; Cho et al. 2004; Parkinson et al. 2005); as well as differences in environmental variables such as energy and temperature (e.g., Rand 1994; Gillooly et al. 2001, 2005; Wright et al. 2003; Davies et al. 2004; but see Bromham and Cardillo 2003; Brown and Pauly 2005). However, establishing cause and effect relationships between organismal, biochemical, or environmental variables and rate heterogeneity is difficult because such features are often intercorrelated themselves and can equally well explain cases of rate heterogeneity (Bromham et al. 1996; Bromham 2002). Moreover, in some cases, highly diverged species were compared and numerous variables that potentially influence mutation rates may have varied during the evolutionary history of the species. Deep phylogenetic comparisons make it very difficult to unambiguously identify specific causes of rate heterogeneity (Thorne et al. 1998; Whittle and Johnston 2003).

Rate heterogeneity was previously observed in noncoding regions of the chloroplast genome among recently diverged species of tropical trees of the Brazil nut family or Lecythidaceae (Hamilton et al. 2003). In the chloroplast study, 2 taxa (*Lecythis zabucajo* and *Eschweilera romeu-cardosoi*) showed consistent rate heterogeneity for nucleotide substitutions at 6 chloroplast intergenic regions. *Lecythis zabucajo* showed a relatively recent rate acceleration because rate heterogeneity was only apparent using a less diverged outgroup but not using a more diverged outgroup. In contrast, *E. romeu-cardosoi* exhibited a relative slowdown of nucleotide substitutions when compared with the more diverged outgroup. Rate heterogeneity was consistent across the 6 chloroplast regions examined, suggesting that rate heterogeneity was caused by differences in the underlying mutation rate consistent with a lineage effect within the chloroplast genome. However, the generation time hypothesis predicts that lineage effects should be common to both the chloroplast and mitochondria genomes. In contrast, different organelles in the same lineage

Downloaded from https://academic.oup.com/mbe/article-abstract/25/7/1282/1041900 by guest on 03 June 2020

Table 1
List of Species, Population and Individual Identification Numbers, and Taxonomic Abbreviation for the mtDNA Lineages Sampled

Family	Species	Population[a]	Individual[a]	Abbreviation
Lecythidaceae	*Bertholletia excelsa*	n/a	n/a	Bex
	Cariniana micrantha	2107	122	Cmi
	Corythophora alta	1202	2058	Cal
		1202	2829	
		1202	4133	
		1202	5022	
		1501	4133	
		3114	216	
		3209	1752	
	Corythophora rimosa	1202	5614	Cri
	Couratari guianensis	1501	15	Cgu
	Couratari multiflora	1501	23	Cmu
		1501	26	
	Eschweilera amazoniciformis	1202	756	Eam
	Eschweilera romeu-cardosoi	1501	4514	Ero
		2206	842	
	Lecythis zabucajo	1202	1267	Lza
		2108	86	
		2206	1977	
	Lecythis poiteaui	2108	664	Lpo
Sapotaceae	*Chrysophylum cainito*	n/a	n/a	Cca

[a] Populations are numbered according to BDFFP nomenclature, and individual numbers are tree tag numbers if present.

not showing rate heterogeneity in common would indicate a narrower lineage effect that could be explained by a mechanism such as differences in organelle-specific polymerase accuracies (Muse 2000).

To assess whether relative rates of molecular evolution are correlated between organelle genomes as expected under lineage effects or are genome specific as expected with locus effects, we sequenced 5 mitochondrial noncoding regions for a sample of 21 taxa from Lecythidaceae that could be compared with the existing chloroplast data set (Hamilton et al. 2003). In addition, we expanded the chloroplast data of Hamilton et al. (2003) by the addition of new taxa for 3 chloroplast regions found to have the highest number of DNA changes per sequence length. Patterns of nucleotide and indel changes were compared between organelle genomes for the same taxa. We employed 2 methods to test for rate constancy of nucleotide substitutions among lineages for the mitochondrial and chloroplast genomes separately. First, we compared molecular clock constrained and unconstrained maximum likelihood (ML) phylogenies. Second, we used both ML and nonparametric relative rate tests to compare divergence between pairs of sequences relative to an outgroup. Relative ratio tests were used to determine whether relative rates of nucleotide divergence for each lineage were proportional between the chloroplast and mitochondrial genomes. Evidence for rate heterogeneity among sequences for one or the other locus but failure to reject the null in the relative ratio test is consistent with a lineage effect and a common cause of rate heterogeneity in the 2 genomes.

Materials and Methods

Sampling

A total of 22 individuals were sampled in this study, including 21 individuals from 9 Lecythidaceae species and the outgroup *Chrysophylum cainito* in the Sapotaceae (table 1). Leaf and bark tissues of the Lecythidaceae were collected at the Biological Dynamics of Forest Fragments Project (BDFFP) in Manaus, Brazil (Gascon and Bierregaard 2001). Consistent with BDFFP botanical inventory plot nomenclature, individual trees were identified by a 4-digit population number and a 1- to 4-digit individual number (e.g., 1202–2829). Trees sampled in population 1501(camp 41), which have no BDFFP botanical inventory plot, were from the Lecythidaceae plot established by S. Mori and P. Becker (Mori S, personal communication). DNA from *Bertholletia excelsa* was extracted from seeds purchased from AmazonHerbs (Paramaribo, Suriname), and leaves from *C. cainito* were collected in Panama. A member of the Sapotaceae was chosen as outgroup because the family is closely related to Lecythidaceae (Morton et al. 1997; Savolainen et al. 2000) and therefore provided an outgroup in another family but with minimal divergence.

Amplification and Sequencing

A total of 5 intergenic mitochondrial regions (*nad1B_C* intron between exons B and C of *nad1*, *nad4exon1_2* intron between exons 1 and 2 of *nad4*, *nad7exon2_3* intron between exons 2 and 3 of *nad7*, *nad7exon3_4* intron between exons 3 and 4 of *nad7*, and the *rps14-cob* intergenic region) were sequenced for this study. The mitochondrial *nad*1 and *rps14-cob* regions were amplified using primers described by Demesure et al. (1995). Novel primers (table 2) were designed for the other 3 mitochondrial intron regions using complete mitochondrial genomes from *Arabidopsis thaliana* (NC 001284), *Beta vulgaris* (NC 002511), *Zea mays* (NC 008332), and *Oryza sativa* (NC 007886). Multiple sequence alignments from these 4 genome sequences were used to identify conserved mitochondrial regions to design

Downloaded from https://academic.oup.com/mbe/article-abstract/25/7/1282/1041900 by guest on 03 June 2020

Table 2
Oligonucleotide Primers and Annealing Temperature Used to Amplify 5 Mitochondrial Intergenic Regions

Primers Pairs (5′–3′)	Primer Name	Annealing Temperature (°C)
GCGTGCCAATCCCTATGTT	3′exon1_nad4	60
CTTTCTTTGTCTCGAACCCCATA	5′exon2_nad4	60
ATCTTCAAGCTTTACCTTATT	3′exon2_nad7	52
CGTGTTCTTGGGCCATCATAG	5′exon3_nad7	62
GGAACCGGTGAACTACACTTGCT	nad7intron2F	64
CCCAGGCCATGCTATTCAATAATA	nad7intron2R	61
TTCAGTGGTGTAATGTTAAG	3′exon3_nad7	54
AACATCGTAAGGTGCTGCTC	5′exon4_nad7	60
GCATTACGATCTGCAGCTCA	nad1 exon B	57.5
GGAGCTCGATTAGTTTCTGC	nad1 exon C	57.5
CACGGGTCGCCCTCGTTCCG	rpS14	57.5
GTGTGGAGGATATAGGTTGT	cob	57.5

primers that would amplify intron and intergenic regions. The software Amplify 1.2 (B. Engels, University of Wisconsin) was used to test potential primers. Primers were synthesized (Operon Biotechnologies, Germantown, MD) and tested in polymerase chain reactions (PCRs) using DNA from the Lecythidaceae and Sapotaceae species.

The chloroplast sequence data of Hamilton et al. (2003) were expanded with the addition of the Lecythidaceae *Eschweilera amazoniciformis*, *E. romeu-cardosoi* (population 2206), and *Lecythis poiteaui* sequenced for the regions *rpl*20-5′*rps*12, *trn*S-*trn*G, and *psb*A-*trn*H. These taxa were added because they expanded the sample of lineages from the 2 genera that demonstrated cpDNA rate heterogeneity in the Hamilton et al. (2003) study. These 3 cpDNA regions were sequenced because they were the most variable among 6 intergenic regions reported in the original cpDNA data set of Hamilton et al. (2003). Chloroplast sequences were amplified using previously described primers and PCR conditions (Hamilton 1999b; Hamilton et al. 2003).

Genomic DNA was extracted by grinding frozen leaf, inner bark, or seed embryo tissue in liquid nitrogen and using a DNeasy plant kit (QiaGen, Valencia, CA) according to the manufacturer's instructions. PCRs to amplify mtDNA contained 2.5 μl of genomic DNA template (DNA concentration was not determined), 5 μl of 10× Thermopol buffer (containing 20 mM $MgSO_4$), 0.2 mM each deoxynucleoside triphosphate, 0.4 μM of each primer, 5 μl of Betaine, and 0.4 units of Vent exo- polymerase (New England Biolabs, Cambridge, MA) in a total volume of 50 μl. The thermal cycling profiles were 5 min at 96 °C followed by 30–40 cycles of 96 °C for 45 s, annealing temperature for 1 min, and 72 °C extension for 2 min. The specific annealing temperatures used for each mitochondrial primer set are shown in table 2. For a few taxa, PCRs repeatedly yielded weak DNA products. In order to increase the final concentration of DNA, in these cases, the initial PCR product was purified with a StrataPrep column and amplified again using the same conditions and primers. PCR products were purified with StrataPrep PCR Purification Kit spin columns (STRATAGENE, La Jolla, CA) and then sequenced in reactions containing 5 μl DNA template, 2 μl primer (1 μM), 1 μl of big dye 5× sequencing buffer, and 2 μl Big Dye v.3.1 terminator reaction ready mix (Applied Biosystems, Foster City, CA). Sequence reactions were purified by ethanol precipitation and electrophoresed on a model 3100 sequencer (Applied Biosystems).

Alignment of DNA Sequences

Sequences were aligned into contigs for each individual using Sequencher 4.5 (GeneCodes, Ann Arbor, MI) and trimmed to exclude terminal coding regions. Gaps in multiple sequence alignments were positioned to minimize the number of nucleotide differences among sequences. Indel characters, character states, and step matrices were coded using Seq State (Muller 2005) following the complex coding method of Simmons and Ochoterena (2000). After multiple sequence alignment, single-base indels and indels associated with mononucleotide runs were all subject to an additional check against original sequence chromatograms to verify that they were not a result of base-calling error. To test whether reamplifying PCR products changed the number of perceived nucleotide substitutions, a sample of sequences from 1-round PCRs were compared with sequences from the same taxa generated from a 2-round PCR. No additional base pair changes were observed in the 2-round PCR sequences. Sequences were deposited in GenBank accession numbers (EU049614–EU049697).

Outgroups with different levels of divergence were utilized to test for changes in relative rates over time (e.g., basal rate homogeneity that evolves to rate heterogeneity in derived lineages). For *L. zabucajo* in the previous study of Lecythidaceae chloroplast sequences (Hamilton et al. 2003), a faster relative rate of nucleotide substitution was observed with the less diverged outgroup *B. excelsa* but rate homogeneity was observed using the more diverged outgroup *C. cainito*. An outgroup difference in relative rate was also seen for *E. romeu-cardosoi*, where a slower relative rate was observed using *C. cainito* but rate homogeneity was observed with *B. excelsa*. To permit comparisons to these previous chloroplast results, 2 data sets of mitochondrial sequences with different outgroup taxa were constructed to test for deviations from the molecular clock. Both mitochondrial data sets comprised concatenated multiple sequence alignments for all 5 noncoding regions. The first data set included the entire sample of 21 sequences from 11 species and the more diverged *C. cainito* as the outgroup (see table 1), hereafter referred to as the 21 sequence mtDNA data set. The second mitochondrial sequence data set included a subset of 15 sequences from 7 species (*B. excelsa*, *Corythophora alta*, *Corythophora rimosa*, *E. amazoniciformis*, *E. romeu-cardosoi*, *L. zabucajo*, and *L poiteaui*) and the less diverged *B. excelsa* as the outgroup, hereafter referred to as the 15 sequence mtDNA data set. In addition, a single chloroplast sequence data set was constructed using the 3 intergenic regions *rpl*20-5′*rps*12, *trn*S-*trn*G, and *psb*A-*trn*H for the same 7 taxa represented in the 15 sequence mtDNA data set. This alignment of chloroplast sequences is hereafter referred to as the 15 sequence cpDNA data set. In the mtDNA NAD7 exons 2–3 sequences, there were 5 clustered nucleotide changes flanking a 1

Downloaded from https://academic.oup.com/mbe/article-abstract/25/7/1282/1041900 by guest on 03 June 2020

bp gap in *C. cainito*. These 5 nt changes were treated as a single change because they appeared to be an inversion associated with an indel event rather than the product of 5 independent nucleotide changes and no other clustered nucleotide changes were observed.

Sequence Analyses

Sequence motifs of indels were classified into 6 types as in Hamilton et al. (2003): 1) mononucleotide repeats, 2) perfect repeats of $\geq$2 bp motif, 3) imperfect repeats of $\geq$2 bp motif, 4) palindromic sequences, 5) reverse-compliment indels or indels abutting such repeats, and 6) other sequences that were not apparently repetitive, structured, or related to abutting sequence. Indel length frequencies in the 21 sequence mtDNA data set were compared with a previous cpDNA indel data set reported in Hamilton et al. (2003). Sequence variation in both genomes was quantified by substitution model–corrected total number of nucleotide substitutions (K_o) and indels (I_o) per site, as well as the K_o/I_o ratio for both 15 sequence mtDNA and 15 sequence cpDNA data sets for each sequence region. Numbers of indels per site, I_o was calculated by dividing the total number of indels by the total number of nucleotide sites excluding all sites in gaps and adding the total number of indels (Laroche and Bousquet 1999). The ratio of the total number of nucleotide sites within indels to the total number of variable nucleotide sites (R_u) in the multiple sequence alignments was calculated according to Britten et al. (2003).

Modeltest 3.5 (Posada and Crandall 1998) was used to determine the best supported model of nucleotide substitution. Tests for the most appropriate nucleotide substitution model used data sets that excluded all sites with gaps in any sequence and all missing and ambiguous sites in the multiple sequence alignments of the 21 and 15 sequence mtDNA and 15 sequence cpDNA data sets. To estimate the phylogenetic relationships within Lecythidaceae for each data set, ML trees were constructed with PAUP* (4.0b10; Swofford 2002) using the nucleotide substitution models from Modeltest.

To assess whether indel and nucleotide rates of change within each organelle genome as well as indel or nucleotide rates of change between organelle genomes were correlated, the numbers of indel and nucleotide changes on each branch of a phylogenetic tree were compared using the combined nucleotide and indel data of the 21 and 15 sequence mtDNA and 15 sequence cpDNA data sets. Because indel characters cannot be used in a ML framework, a maximum parsimony (MP) tree was estimated for combined nucleotide and indel data under the topological constraints of the ML tree for nucleotide data. The number of indel changes on each branch of the phylogenetic tree was estimated by subtracting the number of nucleotide changes on each branch estimated by the nucleotide data MP tree from the total number of changes on the same branch of the combined nucleotide and indel data MP tree. The correlation between the number of nucleotide changes and the number of indel changes could then be estimated with independent comparisons using the number of changes estimated on each branch of the MP tree (e.g., Golenberg et al. 1993; Saitou and Ueda 1994; Graham et al. 2000; Hamilton et al. 2003). Correlation analyses were carried out with JMP (version 5.0.1a; SAS Institute, Cary, NC).

Relative Rate Tests

To test for a constant rate of nucleotide substitution over an entire phylogeny, the molecular clock option in HyPhy (version 0.99 beta; Kosakovsky Pond et al. 2005) was used for each of 3 data sets in conjunction with the substitution models estimated in Modeltest. The molecular clock option computes the likelihood ratio statistic delta = $2(\ln L_A - \ln L_0)$ to test whether the null hypothesis (likelihood L_0) of rate constancy among lineages fits the data better than the alternative hypothesis (likelihood L_A) of rate heterogeneity. Delta has a χ^2 distribution with 2 degrees of freedom (df) in this case, and a significant delta value rejects the null hypothesis of proportional rates between sequences compared.

To investigate whether nucleotide and indel changes exhibited rate heterogeneity between pairs of lineages within mitochondria or chloroplast genomes, Tajima's 1*D* and 2*D* relative rate tests were applied separately to indel characters and nucleotide changes following Hamilton et al. (2003). Tajima's 1*D* test was carried out on either nucleotide or indel changes using a χ^2 distribution with 1 df to test for deviations from the constant rate expectation that the number of unique changes along one lineage equals that along its sister lineage (Tajima 1993). Calculations for Tajima's relative rate tests were conducted with the program T1Dand2D (version 6.4) in conjunction with an Excel spreadsheet (Hamilton et al. 2003). In order to apply the χ^2 approximation for testing the null hypothesis of equal numbers of changes, each lineage must have at least 6 changes (Tajima 1993; Nei and Kumar 2000). For 1*D* tests where at least 1 lineage had 5 or fewer changes, an exact binomial test was applied in lieu of the χ^2 approximation as described in Hamilton et al. (2003). All pairs of taxa were also tested for rate heterogeneity of nucleotide substitution using ML relative rate test (Muse and Weir 1992) conducted with HyPhy and using the nucleotide substitution model parameters from Modeltest.

The relative ratio test was employed to determine whether nucleotide substitution rates for each lineage were proportional between the mitochondrial and chloroplast genomes. This test computes a likelihood ratio statistic to test whether the null hypothesis of proportional relative rates between loci fits the data better than the alternative hypothesis that relative rates lack proportionality between loci (Muse and Gaut 1997). An important feature of the relative ratio test is that it allows comparison of relative rates even if loci evolve at different absolute rates. The relative ratio test requires an explicit nucleotide substitution model and the lineages compared have identical tree topologies for both loci (Muse and Gaut 1997; Gaut 1998; Muse 2000). The relative ratio test was conducted in HyPhy using identical phylogenetic trees for the mitochondrial and chloroplast data sets and nucleotide substitution models estimated in Modeltest. The default options of HyPhy do not permit relative ratio tests using different nucleotide substitution models for the 2 loci. A batch file for HyPhy that calculates the relative ratio test using different nucleotide substitution

Downloaded from https://academic.oup.com/mbe/article-abstract/25/7/1282/1041900 by guest on 03 June 2020

Table 3
Numbers and Types of Indels Observed in 5 Mitochondrial Intergenic Regions Compared with a Previous Study Based on 5 Chloroplast Intergenic Regions in Lecythidaceae

Region	Length in Base Pairs[a] (length gaps excluded)	Mononucleotide Repeats	Perfect Repeats	Imperfect Repeats	Palindromic or Inversion	Other[b]
nad1exonB-C	1,362 (1,315)	0	5	1	1	3
nad4 exon1-2	1,419 (1,358)	1 A/T	3	1	1	7
nad7 exon2-3	1,487 (1,152)	0	2	0	2	2
nad7 exon3-4	1,054 (1,028)	1 A/T, 1 G/C	2	0	0	2
rps14-cob	1,079 (917)	1 A/T	1	1	0	5
Total mtDNA	6,401 (5,770)	3 A/T, 1 G/C	13	3	4	19
Total[c] cpDNA	4,209 (3,317)	16 A/T, 5 G/C	27	11	4	34

NOTE.—A total of 39 indels were observed in the mtDNA sequences compared with 76 total indels observed in cpDNA by Hamilton et al. (2003).
[a] Total length of aligned intergenic sequence including or excluding sites with gaps.
[b] Insertion or deletion of sequence that is not obviously repetitive, structured, or similar to nearby sequence.
[c] Results from Hamilton et al. (2003).

Downloaded from https://academic.oup.com/mbe/article-abstract/25/7/1282/1041900 by guest on 03 June 2020

models for each locus kindly provided by S. L. Kosakovsky was used to compare relative ratios for the chloroplast and mitochondrial sequences.

Results

Patterns of Molecular Evolution in the Organelle Genomes

The 2 *Couratari multiflora* individuals did not show any sequence variation for the mitochondrial and chloroplast sequences. Thus, the mitochondrial and chloroplast data sets represented *C. multiflora* with a single sequence. The 7 *C. alta* sequences formed 2 distinct chloroplast haplotypes, with individuals from the 1202 and 1501 populations sharing identical sequences and individuals from the 3114 and 3209 populations sharing identical sequences. In the chloroplast data sets, *C. alta* was represented by these 2 sequences. However, there was no variation among the *C. alta* mitochondrial sequences. In order to allow phylogenies for the chloroplast and mitochondrial genomes to be compared in the mitochondrial data sets, *C. alta* was represented by 2 identical DNA sequences.

The numbers and types of indels observed in the 21 sequence mtDNA data set (the data set with the more diverged *C. cainito* as outgroup) are summarized in table 3. The 21 sequence mtDNA data set exhibited a total of 6,368 nt sites with 43 indels between 1 and 329 bp in length. Indels $\leq$10 bp were most frequent (40 of 43 or 93.27%). The lengths of the 3 remaining indels were 12, 127, and 329 bp. The total number of indels observed in the 21 sequence mtDNA data set was lower when compared with the number of indels observed in a previous study focused on 5 chloroplast intergenic regions for the same taxa (Hamilton et al. 2003) in spite of mtDNA data set having more nucleotide sites. The frequencies of indel classes were similar in both organelle genomes with nonrepetitive indels being the most common. The total GC content in the 5 mitochondrial regions was almost 50%, whereas the chloroplast sequences were strongly AT rich (AT content = 72.8%).

The patterns of nucleotide and indel changes within Lecythidaceae were compared between organelle genomes using the 15 sequence mtDNA and cpDNA data sets (table 4). The number of indel changes per site (I_o) for all 3 chloroplast regions was always lower than the number of substitutions per site (K_o), whereas in 2 out of 5 mitochondrial regions, I_o was greater than K_o. Over most regions in both organelles, nucleotide substitutions occurred about 2–10 times more frequently than indel changes as measured by K_o/I_o. Indels encompassed on average twice as many nucleotide sites compared with the number of variable nucleotide sites (average R_u = 2.06) in the mitochondrial regions. There were on average 5 times more nucleotide

Table 4
The Ratio of the Total Number of Nucleotide Sites within Indels to the Total Number of Variable Nucleotide Sites (R_u), Substitution Model–Corrected Total Numbers of Nucleotide Substitutions (K_o), and Number of Indels (I_o) Per Site for 5 Mitochondrial and 3 Chloroplast Sequence Regions for Both the 15 Taxa mtDNA Data Set and the 15 Taxa cpDNA Data Set

Genome	Region (length gaps excluded)	Length in Base Pairs	R_u	K_o	I_o	K_o/I_o	GC (%)
Mitochondria	*nad1exonB-C*	1,346 (1,324)	5.5	0.003	0.0037	0.8	53
	nad4 exon1-2	1,417 (1,389)	4.6	0.0043	0.0036	1.19	53
	nad7 exon2-3	1,486 (1,485)	0.09	0.0074	0.0007	10.5	52
	nad7 exon3-4	1,045 (1,032)	4.34	0.0029	0.0048	0.6	54
	rps14-cob	1,074 (1,074)	—	0.0065	0	—	46
	Total	6,368 (6,304)	2.06	0.0049	0.0025	1.96	51.6
Chloroplast	*trnS-trnG*	928 (664)	8.25	0.048	0.0376	1.23	25
	rpl20-5′rps12	828 (772)	1.36	0.053	0.0127	4.1	32.8
	trnH-psbA	490 (231)	7.61	0.147	0.0905	1.47	22
	Total	2,246 (1,667)	5.41	0.0641	0.0342	1.88	27.2

NOTE.—*Bertholletia excelsa* was employed as the outgroup.

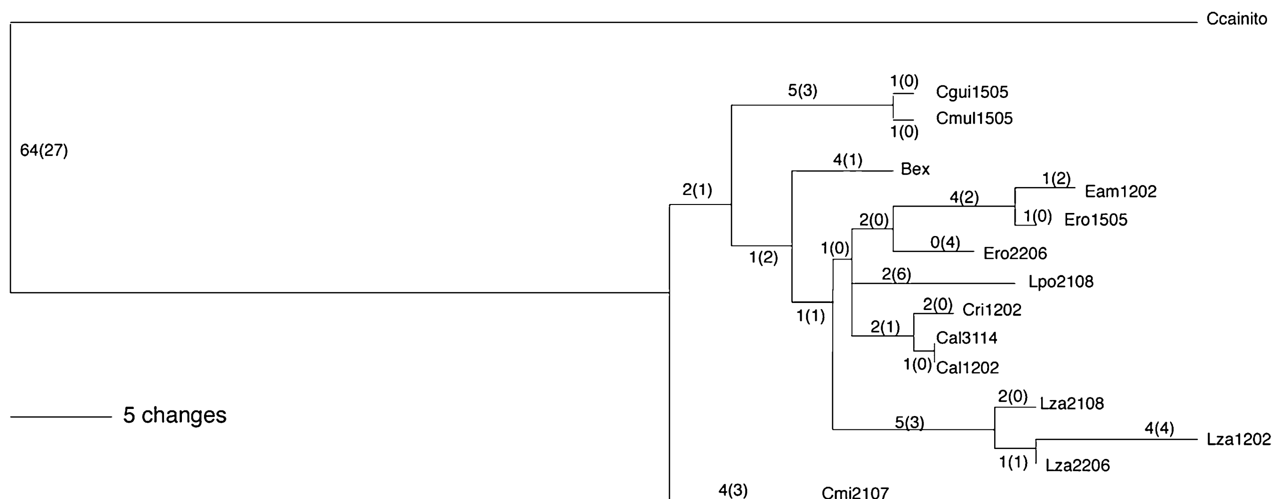

FIG. 1—The ML tree for the 21 sequence mtDNA data set with *Chrysophylum cainito* as the outgroup. The numbers of inferred nucleotide and indel changes (in parenthesis) estimated on ML topology–constrained MP trees are indicated on each branch. Note that the branch length scale is based on total changes (nucleotide plus indel changes).

sites within indels than variable nucleotide sites (R_u = 5.41) for all chloroplast regions. Absolute numbers of nucleotide and indel changes were 13 times greater in the chloroplast genome than in the mitochondrial genome. Because chloroplast and mitochondrial sequences were taken from the same lineages, this translates into a 13-fold higher absolute rate of substitution in the chloroplast genome

Modeltest indicated the F81 + I + G model of nucleotide substitution as most likely for both the 21 and 15 sequence mtDNA data sets. This nucleotide substitution model indicated that base frequencies were variable, all substitutions were equally probable, rates varied among sites with a gamma distribution, and a proportion of sites were invariant (Felsenstein 1981). The substitution models for the 21 and 15 sequence mtDNA data sets differed in the gamma parameters (0.895 and 0.011, respectively) and proportion of invariant sites (0.917 and 0.936, respectively). On the other hand, Modeltest indicated the K81uf (also called K3P, Kimura 1981) model of nucleotide substitution for the 15 sequence cpDNA data set. Under this model, nucleotide frequencies varied (estimated frequencies: A = 0.3118, C = 0.1554, and G = 0.1375) and rates of transition and transversion were unequal (rate matrices: 1.0000,1.2484, 0.3778, 0.3778, and 1.2484).

Figure 1 illustrates the topology of the ML tree for the 21 sequence mtDNA data set with *C. cainito* as outgroup along with the MP tree–inferred indel and nucleotide changes on each branch. The phylogenies in figure 2 show the ML trees for the 15 sequence cpDNA and mtDNA data sets with *B. excelsa* as the outgroup, which exhibited identical topologies. The trees in figure 2 also show the MP-inferred indel and nucleotide changes on each branch. The 3 *Eschweilera* sequences (Ero1501, Ero2206, and Eam) grouped in an unexpected way on all the trees (figs. 1 and 2) because the 2 Ero did not cluster together. It is possible that 2 of the samples were switched or mislabeled during collection or that because these 3 taxa do not resolve properly by chance as they are closely related and show relatively few nucleotide changes. Note that this discrepancy does not impact the utility of the phylogenies to examine relative rates of change. The phylogenies using more or less diverged outgroups (figs. 1 and 2, respectively) differed slightly in topology with the Ero/Eam group being slightly more ancestral to the Lpo/Lza group in the 15 taxa data set due to a relatively small number of changes. This minor difference could have been caused by fewer informative sites in the 15 taxa data set.

Numbers of indel and nucleotide changes were positively correlated for the 21 sequence mtDNA data set (t = 16.01, $P < 0.0001$, R^2 = 0.92), as well as for the 15 sequence cpDNA data set (fig. 3*B*) but not for the 15 sequence mtDNA data set (fig. 3*A*). The numbers of indel (fig. 3*C*) and nucleotide changes (fig. 3*D*) were both positively correlated when compared between the 2 organelle genomes.

Relative Rates within Organelle Genomes

Global rate heterogeneity of nucleotide substitution was observed among lineages in both the chloroplast and mitochondrial genomes. For the chloroplast genome, a global molecular clock among all lineages was always rejected using the 21 sequence cpDNA data set with the more diverged *C. cainito* as outgroup ($P < 0.001$) and the 15 sequence cpDNA data set with the less diverged *B. excelsa* as outgroup ($P < 0.001$). Additionally, a global molecular clock was rejected ($P < 0.0001$) for the 3 regions of cpDNA using those same 16 lineages used in Hamilton et al. (2003).

For the mitochondrial genome, a global molecular clock was rejected when the recently diverged outgroup taxa *Cariniana micrantha* ($P < 0.001$) or *B. excelsa* ($P < 0.001$) were employed. In contrast, with the more diverged *C. cainito* as outgroup, a global molecular clock could not be rejected for the 21 sequence mtDNA data set (P = 0.156) or for the same subset of 16 sequences (P = 0.054) used by Hamilton et al. (2003). In the 21 sequence mtDNA data set, *E. romeu-cardosoi* (the lineage

Downloaded from https://academic.oup.com/mbe/article-abstract/25/7/1282/1041900 by guest on 03 June 2020

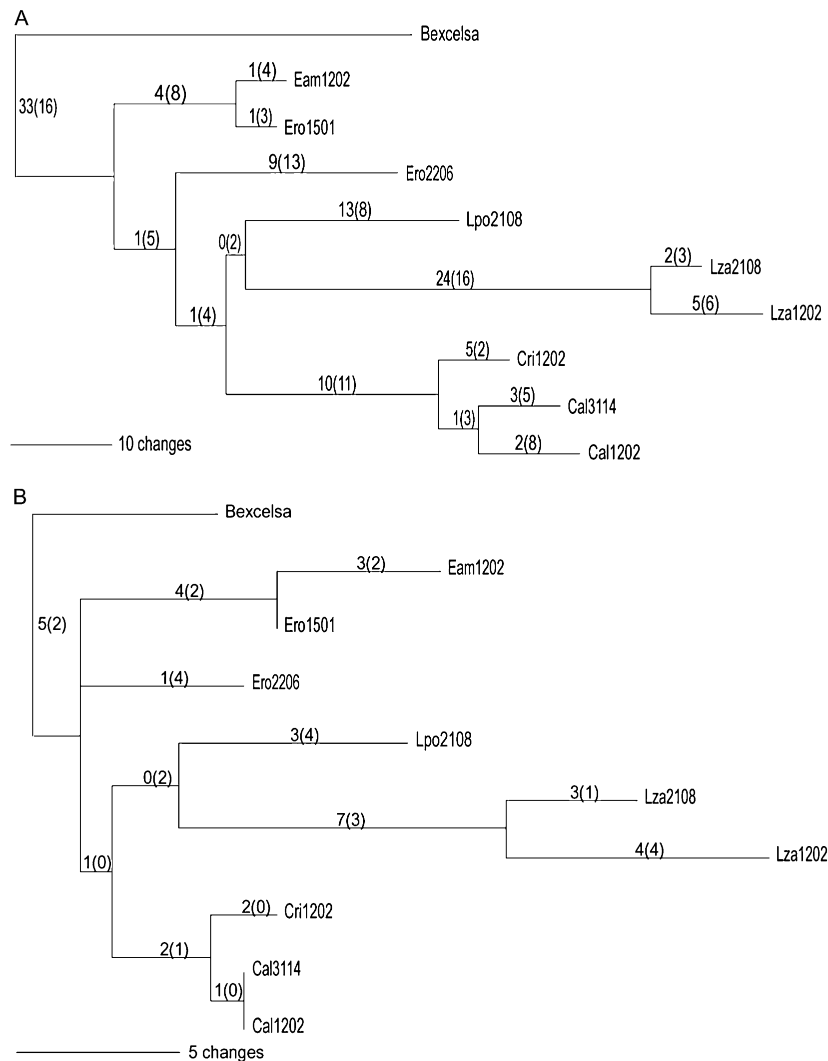

FIG. 2—The ML tree for the 15 sequence cpDNA (*A*) and 15 sequence mtDNA (*B*) data sets with *Bertholletia excelsa* as the outgroup. The numbers of inferred nucleotide and indel changes (in parenthesis) estimated on ML topology–constrained MP trees are indicated on each branch. Note that the branch length scale is based on total changes (nucleotide plus indel changes).

with a significant cpDNA slowdown in Hamilton et al. [2003]) did not show a significant relative rate slowdown in pairwise comparisons (results not shown). However, there was consistent relative rate acceleration in the 21 sequence mtDNA data set when comparing *L. zabucajo* from population 1202 (the lineage with a significant cpDNA acceleration in Hamilton et al. [2003]) and the 3 *Eschweilera* (Ero1501, Ero2206, and Eam) sequences (results not shown). The speed up of *L. zabucajo* was even more evident when combined indel and nucleotide changes were used in the Tajima's 1*D* tests (results not shown).

Comparing relative rate tests for the 15 sequence mtDNA and cpDNA data sets with *B. excelsa* as outgroup showed a parallel pattern of rate heterogeneity in both genomes (table 5). For cpDNA, the 2 *L. zabucajo* individuals consistently exhibited a faster rate of nucleotide substitution and combined nucleotide and indel changes than the rest of the taxa. The 2 *L. zabucajo* individuals also showed mtDNA rate acceleration for combined nucleotide and indel changes in Tajima's 1*D* tests, but ML relative rate tests consistently rejected rate constancy only for *L. zabucajo* from population 1202. In addition, rate heterogeneity for only indel changes was observed for *L. zabucajo* less often in the chloroplast genome and more often in the mitochondrial genome (results not shown). Tajima's 2*D* tests using indel and nucleotide changes were not reported because significance could not be determined by an exact binomial test and the Tajima's 2*D* test appeared to overstate deviations from rate constancy when numbers of changes were small for nucleotide, indel, or both categories of changes, as was frequently observed in the mtDNA data sets (see Hamilton et al. 2003).

Proportionality of Rates for the 2 Organelle Genomes

Relative ratio tests consistently failed to reject the null hypothesis of proportional relative rates between the mitochondrial and chloroplast genomes for all data sets. The relative ratio test did not reject rate proportionality ($P = 0.45$)

Downloaded from https://academic.oup.com/mbe/article-abstract/25/7/1282/1041900 by guest on 03 June 2020

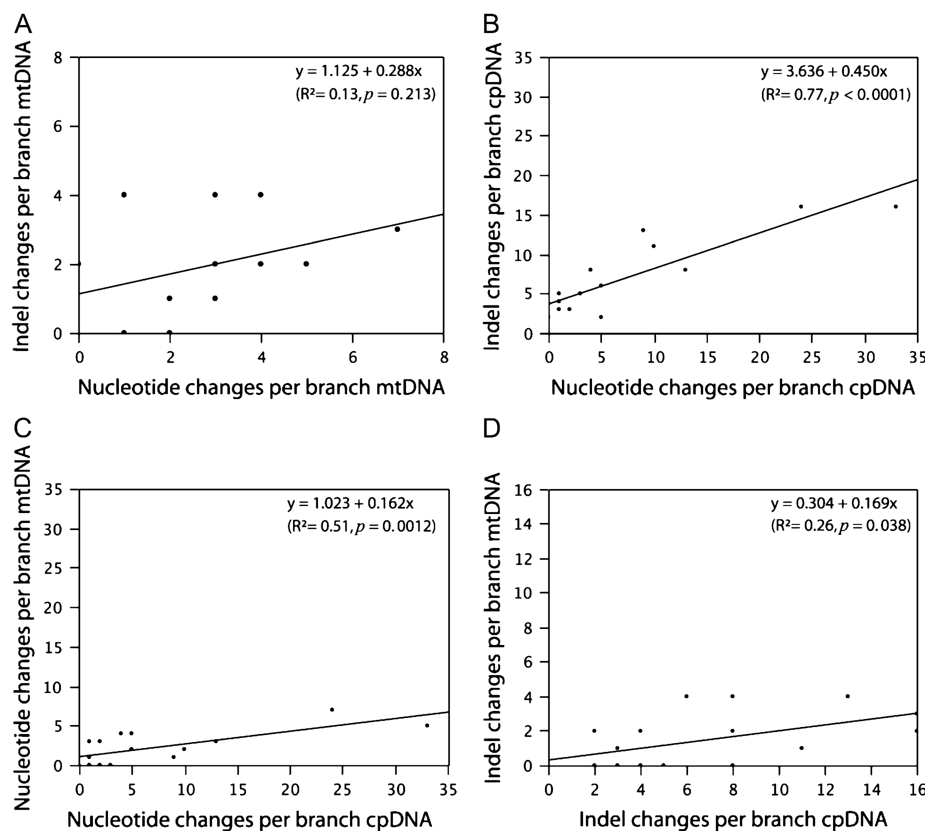

FIG. 3—Joint distributions of nucleotide and indel changes within the same organelle along with joint distributions of nucleotide or indel changes between organelle genomes. Panels (*A*) and (*B*) compare indel and nucleotide changes within the mitochondrial and chloroplast genomes, respectively. Panels (*C*) and (*D*) compare the same type of sequence change, either indel or nucleotide, between the mitochondrial and chloroplast genomes. The numbers of indel and nucleotide changes were estimated from branches of the phylogenetic trees in figure 2. Each panel gives the regression equation, coefficient of determination (R^2), and *t*-test probability that the regression line slope is different than zero.

between the mitochondrial and chloroplast relative rates for the same taxa used in Hamilton et al. (2003). In addition, when the outgroup *B. excelsa* was used the relative ratio test did not reject the null hypothesis of rate proportionality ($P = 0.22$) among branch lengths from the 15 sequence mtDNA and the 15 sequence cpDNA data sets (see phylogenetic trees in fig. 2). Overall, the combination of results from the relative ratio tests and the relative rate tests indicated a parallel pattern for rates of molecular evolution between the mitochondrial and chloroplast sequences.

Discussion

The 13-fold faster absolute rate of substitution in the chloroplast genome compared with the mitochondrial genome is consistent with the generalization that mitochondrial genomes exhibit the lowest overall rate of change among the 3 genomes in plants (Wolfe et al. 1987; Eyre-Walker and Gaut 1997; Laroche et al. 1997; Gaut 1998; Muse 2000; but see exceptions in Cho et al. 2004; Parkinson et al. 2005). The mechanisms causing these differences in absolute rates of molecular evolution among plant genomes remain unclear but are hypothesized to be a consequence of low mitochondrial mutation rates rather than selection (reviewed by Lynch et al. 2006; but see Bazin et al. 2006). Evidence supporting low genome-wide mitochondrial mutation rates comes from plant mitochondrial genes that have transferred and integrated themselves into the nuclear genome, where they have been observed to exhibit higher rates of nucleotide substitution compared with their mitochondrial counterparts (Laroche et al. 1997).

Patterns of Molecular Change in the Organelle Genomes

In spite of the different absolute rates of change, the patterns of nucleotide and indel changes were highly parallel between the mitochondrial and the chloroplast genomes of the Lecythidaceae. The frequency distribution of indel lengths observed in the chloroplast and mitochondrial genomes was very similar, with over 50% short indels ($\leq$10 bp). This result is consistent with previous studies based on nonfunctional DNA sequences in *Drosophila* (Petrov 2002) as well as noncoding chloroplast (Golenberg et al. 1993; Yamane et al. 2006) and mitochondrial (Laroche et al. 1997; Laroche and Bousquet 1999) sequences in angiosperms. Several molecular processes have been proposed to generate different types of indel mutations depending on indel length and specific features of the DNA sequences where indels occur (reviewed by Kelchner 2000). Slipped-strand DNA mispairing events followed by repair of DNA could account for generation of indels shorter than 10 bp (reviewed by Levinson and Gutman

Downloaded from https://academic.oup.com/mbe/article-abstract/25/7/1282/1041900 by guest on 03 June 2020

Table 5
Pairwise Differences for Nucleotide and Indel Changes in the Mitochondrial (below diagonal, 15 sequence mtDNA data set) and Chloroplast (above diagonal, 15 sequence cpDNA data set) Genomes

	Eam	Ero1501	Ero2206	Lpo	Lza2108	Lza1202	Cri	Cal3114	Cal1202	Bex
Eam	—	2	15	16	**33*****	**35*****	**22***	**21***	20	38
		0.0012	0.009	0.0096	**0.0198**	**0.021**	**0.0132**	**0.0126**	0.012	0.0228
		9	43	35	**61****	**63*****	**48***	50	**50***	63
		0.0052	0.0249	0.0203	**0.0353**	**0.0365**	**0.0278**	0.0290	**0.0290**	0.0365
Ero1501	3	—	15	16	**33*****	**35*****	**22***	**21***	**20***	38
	0.0005		0.009	0.0096	**0.0198**	**0.021**	**0.0132**	**0.0126**	**0.012**	0.0228
	5*		42	36	**61****	**65*****	**47***	**48***	**49***	63
	0.0008		0.0243	0.0209	**0.0353**	**0.0377**	**0.0272**	**0.0278**	**0.0284**	0.0365
Ero2206	**6****	5	—	21	**36****	**38****	25	24	23	43
	0.0009	0.0008		0.0126	**0.0216**	**0.0228**	0.015	0.0144	0.0138	0.0258
	14*	11		44	67	**70***	52	53	54	76
	0.0022	0.0017		0.0255	0.0388	**0.0406**	0.0301	0.0307	0.0313	0.0440
Lpo	11	8	5	—	**39****	**41****	28	27	26	43
	0.0017	0.0013	0.0008		**0.0234**	**0.0246**	0.0168	0.0162	0.0156	0.0258
	21	16	14		**64****	**68****	50	50	48	67
	0.0033	0.0025	0.0022		**0.0371**	**0.0394**	0.0290	0.0290	0.0278	0.0388
Lza2108	16	13	**12***	13	—	7	40	39	**37***	59
	0.0025	0.0021	**0.0019**	0.0021		0.0042	0.024	0.0234	**0.0222**	0.0354
	23	18	**19***	21		16	69	72	70	88
	0.0036	0.0028	**0.0030**	0.0033		0.0093	0.040	0.0417	0.0406	0.051
Lza1202	17	**14***	**13*****	**14***	7	—	42	**41***	**39***	61
	0.0027	**0.0022**	**0.0021**	**0.0022**	0.0011		0.0252	**0.0246**	**0.0234**	0.0366
	28	**23***	**21****	23	12		72	76	76	92
	0.0044	**0.0036**	**0.0033**	0.0036	0.0019		0.0417	0.0440	0.0440	0.0533
Cri	10	7	6	7	14	15	—	9	8	49
	0.0016	0.0011	0.0009	0.0011	0.0022	0.0024		0.0054	0.0048	0.0294
	15	10	10	14	**18***	**21****		19	19	78
	0.0024	0.0016	0.0016	0.0022	**0.0028**	**0.00332**		0.0110	0.0110	0.0452
Cal3114	9	8	5	6	13	**14***	3	—	5	48
	0.0014	0.0013	0.0008	0.0009	0.0021	**0.00222**	0.00048		0.003	0.0288
	14*	11	9	**13***	**17***	**20****	3		18	77
	0.0022	0.0017	0.0014	**0.0021**	**0.00269**	**0.00316**	0.00047		0.0104	0.0446
Cal1202	9	8	5	6	13	**14***	3	0	—	47
	0.0014	0.0013	0.0008	0.0009	0.0021	**0.0022**	0.0005	0		0.0282
	14*	11	9	**13***	**17***	**20****	3	0		79
	0.0022	0.0017	0.0014	**0.0021**	**0.0027**	**0.0032**	0.0005	0		0.0458
Bex	12	9	6	9	14	17	10	9	9	—
	0.0019	0.0014	0.0009	0.0014	0.0022	0.0027	0.0016	0.0014	0.0014	
	18	13	11	17	19	24	11	10	10	
	0.0028	0.0021	0.0017	0.0027	0.0030	0.0038	0.0017	0.0016	0.0016	

NOTE.—The first 2 values for each pair of species are the total number of nucleotide changes and the nucleotide substitution model–corrected nucleotide changes per site, respectively. The third value is the total number of combined nucleotide and indel changes. The fourth value is the estimated changes per site (uncorrected for multiple hits) after combining nucleotide and indel changes. Corrected nucleotide divergences were estimated using a F81 + I + G model for the mitochondrial sequences and a K81uf + G model for chloroplast sequences. Species pairs with bold values are those that reject the null hypothesis of relative rate constancy with either the ML test using nucleotide changes or Tajima's 1*D* test using combined nucleotide and indel changes. See table 1 for taxon abbreviations. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$.

1987), whereas longer indels could be caused by insertion of mobile elements (Britten et al. 2003) and/or promoted by the presence of long stem-loop structures in DNA sequences (Buroker et al. 1990). The parallel frequency distribution of indel lengths and classes for both organelle genomes suggest that identical mechanisms account for indels in plant mitochondrial and chloroplast genomes, as was proposed previously (Laroche et al. 1997; Yamane et al. 2006).

Overall, the numbers of nucleotide and indel changes were positively correlated within and between the mitochondrial and chloroplast genomes. The only exception to this general pattern was the lack of correlation between the number of nucleotide and indel changes within the mitochondrial genome when estimated with *B. excelsa* as the outgroup. This lack of correlation seems to be due to low divergence for indels because the *nad7 exon2-3*and the *rps14-cob* mitochondrial regions showed 1 and 0 indel changes, respectively. More frequent sequence changes were seen with the more divergent outgroup *C. cainito*, leading to a positive correlation between numbers of nucleotide and indel changes in the mitochondrial genome. Positively correlated rates of nucleotide and indel changes within organelles are consistent with previous studies in the chloroplast (e.g., Golenberg et al. 1993; Saitou and Ueda 1994; Graham et al. 2000; Hamilton et al. 2003; Yamane et al. 2006) and mitochondrial genomes (Laroche et al. 1997; Laroche and Bousquet 1999) of plants. The fact that nucleotide and indel changes were correlated both within and between organelles suggests that both types of sequence changes are evolving neutrally, sharing common origination (random mutation), and fixation (genetic drift) processes. If the fate of indel or nucleotide changes was dictated by natural selection, we would instead expect rates of change to be uncorrelated between organelles.

Downloaded from https://academic.oup.com/mbe/article-abstract/25/7/1282/1041900 by guest on 03 June 2020

Selection pressures would be expected to be unique to each mutation and organelle and would therefore tend to decouple the rates of change within or between the 2 genomes.

Indel Homoplasy in the Chloroplast Genome

Indel events clearly do not occur at random locations and are associated with specific features of DNA sequences (Kelchner 2000). This may lead to recurrent indel mutations along the same stretches of DNA and thereby may result in convergence and/or saturation of indel events among comparisons of multiple sequence alignments from highly diverged taxa. Although indel mutation models that could form the basis of a correction for indel homoplasy have been described (Tajima and Nei 1984; Thorne et al. 1992; McGuire et al. 2001), such corrections are not used widely to adjust indel divergences for multiple indel events. Thus, if indel homoplasy was common, we expect downwardly biased estimates of the total number of indel events per site as time since divergence increases. Estimates of nucleotide change were corrected for saturation and therefore provide more accurate estimates of divergence. Estimates of K_o/I_o increased in both organelle genomes when sequences from increasingly diverged taxa were compared, suggesting some homoplasy for indel changes. This increase was 1.5 times higher for the chloroplast genome, which may indicate that indels exhibit faster saturation in the chloroplast compared with the mitochondrial genome due to a higher absolute mutation rate in the chloroplast. Previous studies of chloroplast and mitochondrial noncoding regions have estimated indel and nucleotide changes per site to be almost equal at intraspecific level (e.g., Hamilton 1999a; Ingvarsson et al. 2003). However, nucleotide changes per site increased more rapidly than indel changes per site with increasing time since divergence (Golenberg et al. 1993; Saitou and Ueda 1994; Laroche et al. 1997; Laroche and Bousquet 1999; Graham et al. 2000; Hamilton et al. 2003; Yamane et al. 2006), which is consistent with indel homoplasy.

Relative Rate Heterogeneity between Organelle Genomes

Phylogeny-based molecular clock tests revealed global rate heterogeneity for nucleotide divergence among the Lecythidaceae in both organelle genomes and at the same time pairwise relative rate tests also showed rate heterogeneity in specific taxa. In contrast, rate heterogeneity was observed for indel changes in the mitochondrial genome but not in the chloroplast genome. This latter finding seems to suggest that indel changes did not exhibit lineage effects because rate heterogeneity was locus specific. However, it can be argued that the higher absolute rate of mutation in the chloroplast genome caused indel saturation that could affect the ability of the Tajima's 1*D* relative rate tests to detect rate variation for indel changes (see Indel Homoplasy in the Chloroplast Genome). This is a plausible explanation because the relative rate tests indicated rate heterogeneity of indel changes in the mitochondrial genome for the same species that also exhibited rate heterogeneity of nucleotide substitutions corrected for saturation. Overall, the rate heterogeneity of nucleotide substitutions observed in this paper can be classified as lineage effects because relative ratio tests indicated that chloroplast and mitochondrial divergences were always proportional.

Comparisons of relative rates of nucleotide divergence between the organelle genomes with *C. cainito* as outgroup are the only instances that might not support lineage effects. Previous relative rate tests in the chloroplast genome indicated that *E. romeu-cardosoi* exhibited a rate slowdown and *L. zabucajo* a rate speed up (Hamilton et al. 2003). Here, *E. romeu-cardosoi* mitochondrial sequences did not show a significant slowdown. However, rate constancy was rejected in comparisons of *L. zabucajo* 1202 with the 3 *Eschweilera* species. The small number of nucleotide changes observed in the mitochondrial genome can explain the lack of a rate slowdown in *E. romeu-cardosoi*. At the shallow divergences studied here, the relative rate tests have low statistical power in the mitochondrial genome. This is particularly true for a lineage that has experienced a rate slowdown because it would have an even lower rate of divergence and therefore very few changes. In addition, because *L. zabucajo* 1202 and the 3 *Eschweilera* taxa exhibited the most extreme rates of nucleotide divergence in the chloroplast genome, it is logical that with low nucleotide divergences the relative rate tests would first detect rate heterogeneity in species pairs with the largest rate differences.

A distant outgroup lineage can also reduce power of relative rate tests (Tajima 1993; Bromham et al. 2000), possibly explaining why rate constancy was not rejected among Lecythidaceae for the mtDNA data set with the more distant *C. cainito* outgroup. The use of more recently diverged outgroup taxa consistently rejected the null hypothesis of overall rate constancy among species, also suggesting low statistical power of the relative rate tests with mitochondrial sequences. Another possibility is that the relative rate acceleration observed in *L. zabucajo* is a recent event because it was more apparent in the faster evolving chloroplast genome and when recently diverged outgroup taxa were used. A recent overall increase in the rates of nucleotide substitution would be reflected in both genomes, but it would be detected soonest in the genome exhibiting the faster absolute rate of molecular evolution. Recent rate acceleration might be lost in the average rate of divergence estimated when more distant ancestors are used in the comparison. Thus, a recent increase in the rates would be more apparent if a more recent ancestor is used.

The failure to reject rate proportionality between the chloroplast and mitochondrial genomes supports the conclusion that the divergent nucleotide sites studied here are not influenced by natural selection. Selection would be expected to increase or decrease rates of divergence compared with rates experienced by neutral mutations, resulting in locus-specific rates of divergence that would lead to rejection of rate proportionality. Proportional rates of divergence, that is, a lack of natural selection, for the 2 organelles might not be surprising because all organelle regions studied are noncoding. Even natural selection occurring at sites linked to the noncoding regions that we studied would not be expected to impact relative divergence rates. If one considered these chloroplast and mitochondrial noncoding

Downloaded from https://academic.oup.com/mbe/article-abstract/25/7/1282/1041900 by guest on 03 June 2020

regions to be neutral, complete linkage to either advantageous or deleterious mutations would not affect their rates of divergence (Birky and Walsh 1988). However, it is possible for noncoding regions to experience direct natural selection (e.g., Casillas et al. 2007).

Possible Explanations for Lineage Effects in the Absence of Locus Effects

For the plant mitochondrial genome, Cho et al. (2004) and Parkinson et al. (2005) have previously shown a dramatic acceleration in divergence rates that exhibits locus effects (mitochondrial genome–only rate acceleration) and lineage effects (independent acceleration in *Plantago* and in *Pelargonium* lineages only) simultaneously. In those 2 cases, the causal mechanism appears to be evolution of mitochondrial genome mutation rates over time, including both gain and subsequent loss of high mutation rates. Our results show a novel and distinct pattern of rate heterogeneity in plant organelle genomes. In contrast to the *Plantago* and *Pelargonium* cases, heterogeneity in divergence rates in Lecythidaceae is not limited to one organelle genome. The pattern shown in this study is exclusively a lineage effect common between both organelle genomes. Acceleration of divergence rates is parallel in the mitochondrial and chloroplast genomes, so there is no evidence for a locus or genome effect on divergence rates. This finding is more difficult to explain in terms of the evolution of the mutation rate because there is no obvious mechanism that would serve to couple the mutation rates between 2 independent organelle genomes. Our results instead suggest that divergence rates are coupled by processes that alter substitution rates and are common to both organelle genomes but that differ among lineages.

Considering the basic biology of Lecythidaceae, one can evaluate multiple mechanistic hypotheses for the lineage effects observed here. In this study, all taxa were canopy trees collected from the same tropical region sharing similar environmental conditions. Thus, hypotheses that rate heterogeneity is caused by differences in environment among lineages (e.g., solar energy, latitude) seem unlikely. The speciation rate hypothesis postulates that heterogeneity in rates of molecular evolution is related to speciation rates, with more rapidly speciating groups having higher rates of divergence (reviewed by Evans and Gaston 2005). It has been pointed out that the speciation rate hypothesis does not specify a mechanism but rather points to a possible correlation between rates of divergence and species richness (Barraclough et al. 1996). Correlations between species richness and the rate of divergence have been observed in angiosperms (e.g., Bousquet et al. 1992; Savolainen and Goudet 1998; Barraclough and Savolainen 2001). In this study, the genus *Eschweilera* represents one of the most species rich within the Lecythidaceae with 83 described and numerous undescribed species (Mori and Prance 1990). The results here are inconsistent with the speciation rate hypothesis because the *Eschweilera* sequences did not exhibit the fastest rates of divergence. In fact, *Eschweilera* sequences had significantly slower rates of nucleotide substitution. In an attempt to explain the latitudinal gradient in species richness, Davies et al. (2004) suggested that rates of divergence and speciation might be correlated because they are both regulated by “environmental energy.” Our results with *Eschweilera* suggest that, at least for Lecythidaceae, rates of divergence are not related to speciation and that neither of these factors is related in an obvious way to environmental variables.

Lineage-specific divergence rates can be explained by differences in the number of cell divisions per unit of time. Such cell replication rate differences can lead to different numbers of mutations per unit of time in lineages that otherwise share constant mutation rates and probabilities of fixation. Thus, neutral explanations for rate heterogeneity among lineages all rely on some type of replication-dependent mechanism. In angiosperms, the generation time hypothesis has been proposed as an explanation for higher rates of divergence in annual compared with perennial species, previously observed for several nuclear, mitochondrial, and chloroplast loci (e.g., Gaut et al. 1992, 1997; Eyre-Walker and Gaut 1997; Laroche and Bousquet 1999; Andreasen and Baldwin 2001; but see Whittle and Johnston 2003). Our results clearly show lineage effects that are uncoupled from annual/perennial habit because all the species in this study are long-lived perennials.

A neutral replication–dependent model of rate heterogeneity for both nucleotide and indel changes provides a simple explanation for common patterns of rate heterogeneity across the 2 organelle genomes in Lecythidaceae. DNA replication rates may be shared by the organelles because the DNA replication machineries of both mitochondrial and chloroplast genomes are encoded in the nucleus, which places the regulation of organelle DNA replication under common cellular control (Infante and Weissbach 1990; Weinreich 2001). It also seems likely that mitochondrial and chloroplast generation times would be strongly correlated with cell division since the frequency of organelle DNA replication depends on cell generation time. Thus, the lineage effects could be caused by neutral lineage–specific differences in the number of cell divisions per unit time. Alternative hypotheses such as differences in repair fidelity (Britten 1986) and frequency of DNA damage (Martin and Palumbi 1993) could also explain the lineage effects observed here. However, these explanations seem less likely because chloroplast and mitochondrial genomes possess different mechanisms of DNA damage and repair (Prina 1996; Tujeta et al. 2001). Thus, molecular mechanisms impacting the mutation rate would have had to evolve independently in each organelle to produce correlated divergence rates between organelle genomes, an explanation that is not parsimonious.

Data from nuclear genes for these same taxa would distinguish if the mechanism generating the lineage effects is common to all genomes or is limited to only organelle genomes. Because *L. zabucajo* exhibited significant lineage effects for both organelle genomes but only 2 *Lecythis* species were sampled, further taxonomic sampling would help to clarify the extent of lineage effects within the genus *Lecythis* that contains more than 20 species.

Acknowledgments

We thank P. Armbruster, M. Cummings, C. Drummond, C. Lund, and W. Hahn for discussion and comments; S Mori

Downloaded from https://academic.oup.com/mbe/article-abstract/25/7/1282/1041900 by guest on 03 June 2020

and P. Becker for sharing tree locations in the Lecythidaceae plot at camp 41; and S. Kosakovsky for the HyPhy batch file. This work was supported by a doctoral fellowship to D. Soria from the Spanish Ministerio de Educacion y Ciencia, graduate support from Georgetown University and the Department of Biology, the Cosmos Foundation, and a National Science Foundation grant to M.B.H. (DEB9983014). DNA sequencing was supported by an instrumentation award from the National Science Foundation (DBI-0100061) and Georgetown University.

Literature Cited

Andreasen K, Baldwin BG. 2001. Unequal evolutionary rates between annual and perennial lineages of checker mallows (Sidalcea, Malvaceae): evidence from 18S–26S rDNA internal and external transcribed spacers. Mol Biol Evol. 18: 936–944.

Arbogast B, Edwards S, Wakeley J, Beerli P, Slowinski J. 2002. Estimating divergence times from molecular data on phylogenetic and population genetic timescales. Annu Rev Ecol Syst. 33:707–740.

Barraclough TG, Harvey PH, Nee S. 1996. Rate of *rbc*L gene sequence evolution and species diversification in flowering plants (angiosperms). Proc R Soc Lond B Biol sci. 263: 589–591.

Barraclough TG, Savolainen V. 2001. Evolutionary rates and species diversity in flowering plants. Evolution. 55: 677–683.

Bazin E, Glemin S, Galtier N. 2006. Population size does not influence mitochondrial genetic diversity in animals. Science. 312:570–572.

Birky CW, Walsh JB. 1988. Effects of linkage on rates of molecular evolution. Proc Natl Acad Sci USA. 85: 6414–6418.

Bousquet J, Strauss S, Doerksen AH, Price RA. 1992. Extensive variation in evolutionary rate of *rbc*L gene sequences among seed plants. Proc Natl Acad Sci USA. 89:7844–7848.

Britten R. 1986. Rates of DNA sequence evolution differ between taxonomic groups. Science. 231:1393–1398.

Britten RJ, Rowen L, Williams J, Cameron RA. 2003. Majority of divergence between closely related DNA samples is due to indels. Proc Natl Acad Sci USA. 100:4661–4665.

Bromham L. 2002. Molecular clocks in reptiles: life history influences rate of molecular evolution. Mol Biol Evol. 19:302–309.

Bromham L, Cardillo M. 2003. Testing the link between the latitudinal gradient in species richness and rates of molecular evolution. J Evol Biol. 16:200–207.

Bromham L, Penny D. 2003. The modern molecular clock. Nat Rev Genet. 4:216–224.

Bromham L, Penny D, Rambaut A, Hendy MD. 2000. The power of relative rates tests depends on the data. J Mol Evol. 50: 296–301.

Bromham L, Rambaut A, Harvey PH. 1996. Determinants of rate variation in mammalian DNA sequence evolution. J Mol Evol. 43:610–621.

Brown JM, Pauly GB. 2005. Increased rates of molecular evolution in an equatorial plant clade: an effect of environment or phylogenetic nonindependence? Evolution. 59:238–242.

Buroker NE, Brown JR, Gilbert TA, O'Hara PJ, Beckenbach AT, Thomas WK, Smith MJ. 1990. Length heteroplasmy of sturgeon mitochondrial DNA: an illegitimate elongation model. Genetics. 124:157–163.

Casillas S, Barbadilla A, Bergman CM. 2007. Purifying selection maintains highly conserved noncoding sequences in *Drosophila*. Mol Biol Evol. 24:2222–2234.

Cho Y, Mower JP, Qiu YL, Palmer JD. 2004. Mitochondrial substitution rates are extraordinarily elevated and variable in a genus of flowering plants. Proc Natl Acad Sci USA. 101: 17741–17746.

Davies T, Savolainen V, Chase M, Moat J, Barraclough TG. 2004. Environmental energy and evolutionary rates in flowering plants. Proc R Soc Lond B Biol Sci. 271:2195–2200.

Demesure B, Sodzi N, Petit RJ. 1995. A set of universal primers for amplification of polymorphic non-coding regions of mitochondrial and chloroplast DNA in plants. Mol Ecol. 4:129–131.

Evans KL, Gaston KJ. 2005. Can the evolutionary-rates hypothesis explain species-energy relationship? Funct Ecol. 19:899–915.

Eyre-Walker A, Gaut BS. 1997. Correlated rates of synonymous site evolution across plant genomes. Mol Biol Evol. 14: 455–460.

Felsenstein J. 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. J Mol Evol. 17: 368–376.

Gascon C, Bierregaard ROJ. 2001. The biological dynamics of forest fragmentation project: The study site, experimental design, and research activity. In: Bierregaard ROJ, Gascon C, Lovejoy TE, Mesquita R, eds. *Lessons from Amazonia: The Ecology and Conservation of Fragmented Forest*. New Haven (CT): Yale University Press. p. 31–42.

Gaut BS. 1998. Molecular clocks and nucleotide substitution rate in higher plants. Evol Biol. 30:93–120.

Gaut BS, Clark LG, Wendel JF, Muse SV. 1997. Comparisons of the molecular evolutionary/process at *rbc*L and *ndh*F in the grass family (Poaceae). Mol Biol Evol. 14:769–777.

Gaut BS, Muse SV, Clark WD, Clegg MT. 1992. Relative rates of nucleotide substitution at the *rbcL* locus of the monocotydelonous plants. J Mol Evol. 35:292–303.

Gillespie JH. 1991. The causes of molecular evolution. New York: Oxford University Press.

Gillooly JF, Brown J, West G, Savage V, Charnov E. 2001. Effects of size and temperature on metabolic rate. Science. 293:2248–2251.

Gillooly JF, Allen AP, West GB, Brown JH. 2005. The rate of DNA evolution: effects of body size and temperature on the molecular clock. Proc Natl Acad Sci USA. 102:140–145.

Gissi C, Reyes A, Pesole G, Saccone C. 2000. Lineage-specific evolutionary rates in mammalian mtDNA. Mol Biol Evol. 17:1022–1031.

Golenberg EM, Clegg MT, Durbin ML, Doebley J, Ma DP. 1993. Evolution of a noncoding region of the chloroplast genome. Mol Phylogenet Evol. 2:52–64.

Graham SW, Reeves PA, Burns ACE, Olmstead RG. 2000. Microstructural changes in noncoding chloroplast DNA: interpretation, evolution, and utility of indels and inversions in basal angiosperm phylogenetic inference. Int J Plant Sci. 161:S83–S96.

Hamilton MB. 1999a. Tropical tree gene flow and seed dispersal. Nature. 402:129–130.

Hamilton MB. 1999b. Four primer pairs for the amplification of chloroplast intergenic regions with intraspecific variation. Mol Ecol. 8:521–523.

Hamilton MB, Braverman JM, Soria-Hernanz DF. 2003. Patterns and relative rates of nucleotide and insertion/deletion evolution at six chloroplast intergenic regions in New World species of Lecythidaceae. Mol Biol Evol. 20: 1710–1721.

Downloaded from https://academic.oup.com/mbe/article-abstract/25/7/1282/1041900 by guest on 03 June 2020

Infante D, Weissbach A. 1990. Organellar DNA replication in *Nicotiana tabacum* cultured cells. Plant Mol Biol. 14: 891–897.

Ingvarsson PK, Ribstein S, Taylor DR. 2003. Molecular evolution of insertions and deletion in the chloroplast genome of *Silene*. Mol Biol Evol. 20:1737–1740.

Kay KM, Whittall JB, Hodges SA. 2006. A survey of nr*ITS* substitution rates across angiosperms supports an approximate molecular clock with life history effects. BMC Evol Biol. 6:36.

Kelchner SA. 2000. The evolution of non-coding chloroplast DNA and its application in plant systematics. Ann. MO Bot. Gard. 87:482–498.

Kimura M. 1968. Evolutionary rate at the molecular level. Nature. 217:624–626.

Kimura M. 1981. Estimation of evolutionary distances between homologous nucleotide sequences. Proc Natl Acad Sci USA. 78:454–458.

Kimura M. 1983. The neutral theory of molecular evolution. New York: Cambridge University Press.

Kosakovsky Pond SL, Frost SDW, Muse SV. 2005. HyPhy: hypothesis testing using phylogenies. Bioinformatics. 21: 676–679.

Laroche J, Bousquet J. 1999. Evolution of the mitochondrial *rps*3 intron in perennial and annual angiosperms and homology to *nad*5 intron 1. Mol Biol Evol. 16:441–452.

Laroche J, Li P, Maggia L, Bousquet J. 1997. Molecular evolution of angiosperm mitochondrial introns and exons. Proc Natl Acad Sci USA. 94:5722–5727.

Levinson G, Gutman GA. 1987. Slipped-strand mispairing: a major mechanism for DNA sequence evolution. Mol Biol Evol. 4:203–221.

Lynch M, Koskella B, Schaack S. 2006. Mutation pressure and the evolution of organelle genomic architecture. Science. 311:1727–1730.

Martin AP, Palumbi SR. 1993. Body size, metabolic rate, generation time, and the molecular clock. Proc Natl Acad Sci USA. 90:4087–4091.

McGuire G, Denham MC, Balding DJ. 2001. Models of sequence evolution for DNA sequences containing gaps. Mol Biol Evol. 18:481–490.

Mooers AO, Harvey PH. 1994. Metabolic rate, generation time, and the rate of molecular evolution in birds. Mol Phylogenet Evol. 3:344–350.

Mori SA, Prance GT. 1990. Lecythidaceae—part II: the zygomorphic-flowered New World Genera (*Couroupita, Corythophora, Bertholletia, Couratari, Eschweilera*, & *Lecythis*). With a study of the secondary xylem of Neotropical Lecythidaceae by Carl de Zeeuw. Flora Neotrop Monogr. 21:1–376.

Morton CM, Mori SA, Prance GT, Karol KG, Chase MW. 1997. Phylogenetic relationships of Lecythidaceae: a cladistic analysis using *rbc*L sequence and morphological data. Am J Bot. 84:530–540.

Muller K. 2005. SeqState: primer design and sequence statistics for phylogenetic DNA datasets. Application note. Appl Bioinformatics. 4:65–69.

Muse SV. 2000. Examining rates and patterns of nucleotide substitution in plants. Plant Mol Biol. 42:25–43.

Muse SV, Gaut BS. 1997. Comparing patterns of nucleotide substitution rates among chloroplast loci using the relative ratio test. Genetics. 146:393–399.

Muse SV, Weir BS. 1992. Testing for equality of evolutionary rates. Genetics. 132:269–276.

Nei M, Kumar S. 2000. Molecular evolution and phylogenetic. Oxford (UK): Oxford University Press.

Ohta T, Kimura M. 1971. On the constancy of the evolutionary rate of cistrons. J Mol Evol. 1:18–25.

Palmer JD, Adams KL, Cho Y, Parkinson CL, Qiu Y-L, Song K. 2000. Dynamic evolution of plant mitochondrial genomes: mobile genes and introns, and highly variable mutation rates. Proc Natl Acad Sci USA. 97:6960–6966.

Palmer JD, Herbon L. 1988. Plant mitochondrial DNA evolves rapidly in structure, but slowly in sequence. J Mol Evol. 28:87–97.

Parkinson CL, Mower JP, Qiu YL, Shirk AJ, Song K, Young ND, DePamphilis CW, Palmer JD. 2005. Multiple major increases and decreases in mitochondrial substitution rates in the plant family Geraniaceae. BMC Evol Biol. 5:73.

Petrov DA. 2002. DNA loss and evolution of genome size in *Drosophila*. Genetica. 115:81–91.

Posada D, Crandall KA. 1998. Modeltest: testing the model of DNA substitution. Bioinformatics. 14:817–818.

Prina AR. 1996. Mutator-induced cytoplasmic mutants in barley: genetic evidence of activation of a putative chloroplast transposon. J Hered. 87:385–389.

Rand DM. 1994. Thermal habit, metabolic rate and the evolution of mitochondrial DNA. Trends Ecol Evol. 9:125–131.

Saitou N, Ueda S. 1994. Evolutionary rates of insertion and deletion in noncoding nucleotide sequences of primates. Mol Biol Evol. 11:504–512.

Sarich VM, Wilson AC. 1967. Immunological time scale for hominid evolution. Science. 158:1200–1203.

Savolainen V, Chase M, Hoot S, Morton C, Soltis DE, Bayer C, Fay MF, de Bruijn AY, Sullivan S, Qiu YL. 2000. Phylogenetics of flowering plants based on combined analysis of plastid *atp*B and *rbc*L gene sequences. Syst Biol. 49: 306–362.

Savolainen V, Goudet J. 1998. Rate of gene sequence evolution and species diversification in flowering plants: a re-evaluation. Proc R Soc Lond B Biol Sci. 265:603–607.

Simmons MP, Ochoterena H. 2000. Gaps as characters in sequence-based phylogenetic analyses. Syst Biol. 49:369–381.

Small RL, Cronn RC, Wendel JF. 2004. Use of nuclear genes for phylogeny reconstruction in plants. Aust Syst Bot. 17: 145–170.

Swofford DL. 2002. PAUP*. Phylogenetic analysis using parsimony (*and other methods). Version 4. Sunderland (MA): Sinauer Associates.

Tajima F. 1993. Simple methods for testing the molecular evolutionary clock hypothesis. Genetics. 135:599–607.

Tajima F, Nei M. 1984. Estimation of evolutionary distance between nucleotide sequences. Mol Biol Evol. 1:269–285.

Thorne J, Kishino H, Felsenstein J. 1992. An evolutionary model for maximum likelihood alignment of DNA sequences. J Mol Evol. 33:114–124.

Thorne J, Kishino H, Painter I. 1998. Estimating the rate of evolution of the rate of molecular evolution. Mol Biol Evol. 15:1647–1657.

Tujeta N, Singh MB, Misra MK, Bhalla PL, Tujeta N. 2001. Molecular mechanisms of DNA damage and repair: progress in plants. Crit Rev Biochem Mol Biol. 36:337–397.

Weinreich DM. 2001. The rates of molecular evolution in rodent and primate mitochondrial DNA. J Mol Evol. 52:40–50.

Whittle G, Johnston MO. 2003. Broad-scale analysis contradicts the theory that generation time affects molecular evolutionary rates in plants. J Mol Evol. 56:223–233.

Wolfe KH, Li WH, Sharp P. 1987. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. Proc Natl Acad Sci USA. 84:9054–9058.

Wright SD, Gray RD, Gardner RC. 2003. Energy and the rate of evolution: inferences from plant nrDNA mutation rates in the west Pacific. Evolution. 57:2893–2898.

Downloaded from https://academic.oup.com/mbe/article-abstract/25/7/1282/1041900 by guest on 03 June 2020

Wu ML, Li WH. 1985. Evidence for higher rates of nucleotide substitution in rodents than in man. Proc Natl Acad Sci USA. 82:1741–1745.
Yamane K, Yano K, Kawahara T. 2006. Pattern and rate of indel evolution inferred from whole chloroplast intergenic regions in sugarcane, maize and rice. DNA Res. 13:197–204.
Young N, dePamphilis C. 2005. Rate variation in parasitic plants: correlated and uncorrelated patterns among plastid genes of different function. BMC Evol Biol. 5:16.
Zuckerkandl E, Pauling L. 1965. Evolving genes and proteins. In: Bryson V and Vogel JH, eds. Evolutionary divergence and convergence in proteins. New York: Academic Press. p. 97–166.

Sudhir Kumar, Associate Editor

Accepted March 11, 2008

Downloaded from https://academic.oup.com/mbe/article-abstract/25/7/1282/1041900 by guest on 03 June 2020