J Mol Evol (1992) 35:292–303

Journal of Molecular Evolution

Relative Rates of Nucleotide Substitution at the *rbc*L Locus of Monocotyledonous Plants

Brandon S. Gaut,[1] Spencer V. Muse,[2] W. Dennis Clark,[3] and Michael T. Clegg[1]

[1] Department of Botany and Plant Sciences, University of California, Riverside, Riverside, CA 92521, USA
[2] Program in Statistical Genetics, North Carolina State University, Raleigh, NC 27695, USA
[3] Department of Botany, Arizona State University, Tempe, AZ 85287, USA

Summary. We subjected 35 *rbc*L nucleotide sequences from monocotyledonous taxa to maximum likelihood relative rate tests and estimated relative differences in rates of nucleotide substitution between groups of sequences without relying on knowledge of divergence times between taxa. Rate tests revealed that there is a hierarchy of substitution rate at the *rbc*L locus within the monocots. Among the taxa analyzed the grasses have the most rapid substitution rate; they are followed in rate by the Orchidales, the Liliales, the Bromeliales, and the Arecales. The overall substitution rate for the *rbc*L locus of grasses is over 5 times the substitution rate in the *rbc*L of the palms. The substitution rate at the third codon positions in the *rbc*L of the grasses is over 8 times the third position rate in the palms. The pattern of rate variation is consistent with the generation-time-effect hypothesis. Heterogenous rates of substitution have important implications for phylogenetic reconstruction.

Key words: *rbc*L — Relative rates of nucleotide substitution — Generation time — Phylogeny construction

Introduction

The molecular clock hypothesis (Zuckerkandl and Pauling 1965) has been the subject of controversy. Early protein sequence data suggested that amino acid substitution rates are constant between different evolutionary lineages (Wilson et al. 1977; Kimura 1983, 1989), while more recent studies of nucleotide sequences have suggested that the rate of the molecular clock varies between evolutionary lineages (Li et al. 1985, 1987a; Wu and Li 1985; Bulmer et al. 1991). A number of factors have been hypothesized to account for heterogeneous substitution rates between lineages, including differences in evolutionary history, selection, generation time, and polymerase fidelity (Li et al. 1985, 1987a; Wu and Li 1985; Britten 1986; Gillespie 1986). A thorough characterization of rate variation is an essential prerequisite to distinguishing among these various hypotheses. Knowledge of rate variation is also important for the study of molecular phylogenies, since rate constancy between lineages is sometimes assumed in the process of phylogenetic reconstruction.

The chloroplast gene encoding ribulose-1,5 bisphosphate-carboxylase (*rbc*L) has been used as a tool in the phylogenetic analysis of angiosperms (see Doebley et al. 1990; Soltis et al. 1990; Clark et al. 1993; Duvall et al. 1993; Giannasi et al. 1992) and has been shown to have heterogeneous rates of nucleotide substitution between some plant lineages (Smith and Doyle 1986; Doebley et al. 1990; Wilson et al. 1990). Rate variation at the *rbc*L locus may be most conspicuous within monocotyledonous plants. For example, the *rbc*L of maize has been found to evolve more rapidly than the *rbc*L of other mem-

Offprint requests to: M.T. Clegg

bers of the grass family (Doebley et al. 1990; Gaut and Clegg 1991). The *rbc*L locus in palms (family Arecaceae) has been found to have an eightfold-lower overall substitution rate and a 36-fold-lower synonomous substitution rate relative to the *rbc*L of some annual plant species (Wilson et al. 1990). The determination of substitution rates by Wilson et al. (1990) relied on fossil evidence to estimate the time of divergence between plant lineages.

The extent of rate variation at the *rbc*L locus over a broad range of monocot taxa is uncertain. In order to examine the heterogeneity of substitution rate at the *rbc*L locus, we have applied maximum likelihood relative rate tests (Muse and Weir 1992) to 35 *rbc*L nucleotide sequences from various monocot taxa. We discuss the extent of rate heterogeneity at the *rbc*L locus within the monocots and estimate relative differences in rates of nucleotide substitution independently of both phylogenetic assumptions and knowledge of divergence times between species. We also discuss the factors which may contribute to variation in substitution rates and the implications of rate heterogeneity for the methodology of phylogenetic reconstruction.

Material and Methods

Sequence Data. DNA was extracted from palm leaves (Wilson et al. 1990) and from leaves of *Zea mays* ssp. *mays* (Doyle and Doyle 1987). Template for sequencing was produced by symmetric polymerase chain reaction (PCR) amplification of the *rbc*L locus. Symmetric PCR amplification was followed by asymmetric amplification in order to generate single-stranded DNA. Asymmetric amplification products were sequenced directly using the di-deoxy method (Sanger et al. 1977). Primers internal to the *rbc*L gene were employed for sequencing reactions.

Relative Rate Tests. Relative rate tests were performed according to the two-parameter maximum likelihood method of Muse and Weir (1992). Each relative rate test requires three nucleotide sequences. Two sequences (A and B) are examined for departures from rate equivalence; the third sequence (D) functions as an outgroup. The three sequences comprise a star phylogeny which includes a node (C) (Fig. 1). The relative rate test examines departures from the null hypothesis (H_0) of rate equivalence; H_0 constrains the rate of transition substitution in the lineage leading from node C to sequence A (α_A) to equal the rate of transition substitution in the lineage leading from node C to sequence B (α_B). H_0 also constrains the rate of transversion substitution in the lineage leading to sequence A (β_A) to be equal to the rate of transversion substitution in the lineage leading to sequence B (β_B). (That is, H_0: $\alpha_A = \alpha_B$, $\beta_A = \beta_B$; H_A: $\alpha_A \neq \alpha_B$ and/or $\beta_A \neq \beta_B$.) The likelihood ratio test statistic is χ^2 distributed with two degrees of freedom; a significant result ($p < 0.01$) indicates that the maximum likelihood estimate of transition and/or transversion substitutions between the lineages leading to sequences A and B are sufficiently unequal to reject H_0.

The *rbc*L sequences from 35 monocot taxa were included in this analysis (Table 1); 1400 base pairs of sequence were used for every taxon except *Hordeum vulgare* (1279 bp), *Nypa fruticans* (1365 bp), *Pontederia sagittaria* (1320 bp), *Aechmea chantinii* (1370 bp), Vellozia sp. (1350 bp), *Anomatheca laxa* (1370 bp), and *Hechtia montana* (1350 bp). Every possible pair of monocot *rbc*L sequences (of which there are $(34 \times 35/2) = 595$ possible pairs) was tested for departures from H_0. The outgroup (D) for every test was *rbc*L sequence from *Magnolia macrophylla* (Table 1). *Magnolia* was chosen as the outgroup because (1) it is a dicot and hence an outgroup to any pairwise comparison of monocots and (2) the Magnoliidae may be basal to monocotoyledonous plants (Cronquist 1988, p. 453; Dahlgren et al. 1985, p. 48) and thus more closely related to monocots than other dicots. Each pair of monocot sequences was subjected to four maximum likelihood relative rate tests (resulting in a total of $595 \cdot 4 = 2380$ tests): one test in which all the nucleotide data were examined and three additional tests in which the data were limited to only first-, second-, or third-position nucleotide data, respectively.

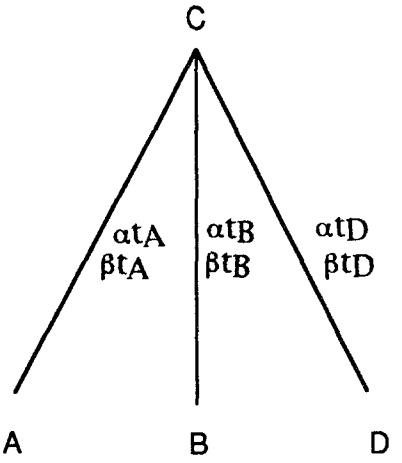

Fig. 1. Star phylogeny for three sequences used in maximum likelihood relative rates tests. Sequences A and B share the ancestral sequence C. αt_A and βt_A are the estimable parameters from the lineage leading from sequence C to sequence A. αt_B, βt_B, αt_D, and βt_D are the estimable parameters from the lineages leading to sequences B and D, respectively.

Estimation of Relative Rates of Nucleotide Substitution. According to the two-parameter model of Muse and Weir (1992), the total substitution rate (μ) within a lineage is $\alpha + 2\beta$, where α is the rate of transitions and β is that of transversions. These parameters are not estimable themselves. Instead, they are confounded with divergence times, so αt and βt are the estimable parameters, with t the length of time along an evolutionary pathway (Fig. 1). Using this framework, the ratio of evolutionary rates between two groups of sequences can be estimated in some cases. Consider two groups of homologous sequences, $A_1 \ldots A_m$ and $B_1 \ldots B_n$, with homogeneous rates of substitution within each group. The parameter of interest is

$$u = \frac{\alpha_A t_A + 2\beta_A t_A}{\alpha_B t_B + 2\beta_B t_B} = \frac{\mu_A}{\mu_B} \qquad (1)$$

the ratio of total substitution rates. (Note that $t_A = t_B$ by design.) Given a pair of sequences, A_i and B_j, and an appropriate outgroup D, maximum likelihood estimates $\mu_{A_{ij}}$ and $\mu_{B_{ij}}$ may be found. ($\mu_{A_{ij}}$ is the maximum likelihood estimate of μ_A found by using sequences A_i and B_j; $\mu_{B_{ij}}$ is the maximum likelihood estimate of μ_B using sequence A_i and B_j.) By using all pairs of sequences in groups A and B, a combined estimate of u may be computed as

Table 1. Taxa for relative rate analysis[a]

	Species	Source	Abv
Order Cyperales			
family Poaceae	*Zea mays*	This paper	zea
	Avena sativa	Garcia and Clegg, '91	aven
	Puccinellia distans	Doebley et al., '90	pucc
	Pennisetum glaucum	Doebley et al., '90	penn
	Neurachne munroi	Hudson et al., '90	neum
	Neurachne tenuifolia	Hudson et al., '90	neut
	Oryza sativa	Moon et al., '87	oryz
	Cenchrus setigerus	Doebley et al., '90	cenc
	Triticum aestivum	Terachi et al., '87	trit
	Aegilops crassa	Terachi et al., '87	aegi
	Hordeum vulgare	Zurawski et al., '84	hord
Order Liliales			
family Liliaceae	*Colchicum speciosum*	Chase & Hills, unpub.	colc
	Danae racemosa	Chase & Hills, unpub.	dana
	Hypoxis leptocarpa	Chase & Hills, unpub.	hypo
	Kniphofia uvaria	Chase & Hills, unpub.	knip
	Lilium superbum	Chase & Hills, unpub.	lili
family Amaryllidaceae	*Aletris farinacea*	Chase & Hills, unpub.	alet
family Iridaceae	*Anomatheca laxa*	Chase & Hills, unpub.	anom
family Pontederiaceae	*Pontederia sagittaria*	Clark et al., '92	pont
family Smilaceae	*Smilax glauca*	Chase & Hills, unpub.	smil
family Velloziaceae	*Vellozia* sp.	Clark et al., '92	vell
Order Orchidales			
family Orchidaceae	*Oncidium excavatum*	Chase & Hills, unpub.	onci
family Burmmaniaceae	*Burmannia biflora*	Chase & Hills, unpub.	burm
Order Commelinales			
family Rappateaceae	*Stegolepis allenii*	Clark et al., '92	steg
Order Bromeliales			
family Bromeliaceae	*Tillandsia elizabethae*	Clark et al., '92	till
	Puya dyckioides	Clark et al., '92	puya
	Hechtia montana	Clark et al., '92	hech
	Ananas comosus	Clark et al., '92	anan
	Aechmea chantinii	Clark et al., '92	aech
Order Arecales			
family Arecaceae	*Phoenix reclinata*	This paper	phoe
	Serenoa repens	Wilson et al., '90	sere
	Calamus usitatus	Wilson et al., '90	cala
	Caryota mitis	This paper	cary
	Nypa frucitans	This paper	nypa
Order Magnoliales	*Drymophloeus subdisticha*	This paper	drym
family Magnoliaceae	*Magnolia macrophylla*	Golenberg et al., '90	—

[a] All species are monocotyledonous except *Magnolia*. Abv refers to the abbreviation for the species used in tables and the appendix. Classification as per Cronquist (1988)

$$\hat{u} = \frac{\sum_i \sum_j \mu_{A_{ij}}}{\sum_i \sum_j \mu_{B_{ij}}} \qquad (2)$$

The standard error of this estimate may be found using the jackknife, omitting each pair of sequences in turn. The estimate of the standard error may be used to form a confidence interval for u. As nothing is known about the small sample distribution of $\hat{u}$, a simple method is to use Chebychev's Inequality, which states that an estimate is within k standard deviations of its mean with probability of at least $1 - 1/k^2$, no matter what the distribution. For 95% confidence intervals, $k = \sqrt{20}$. This interval will most certainly be overly conservative, but it will be sufficient for our purposes.

Results

Sequence Data

Sequences of the *rbc*L locus were generated for six palm taxa representing four of the six subfamilies of the Arecaceae (Uhl and Dransfield 1987). Two sequences, *Calamus usitatus* (subfamily Calamoideae) and *Serenoa repens* (subfamily Coryphoideae), have been published previously in truncated form (Wilson et al. 1990). These two sequences have been augmented to include 460 base pairs which were lacking from the previous analysis. Four other palm taxa have also been sequenced:

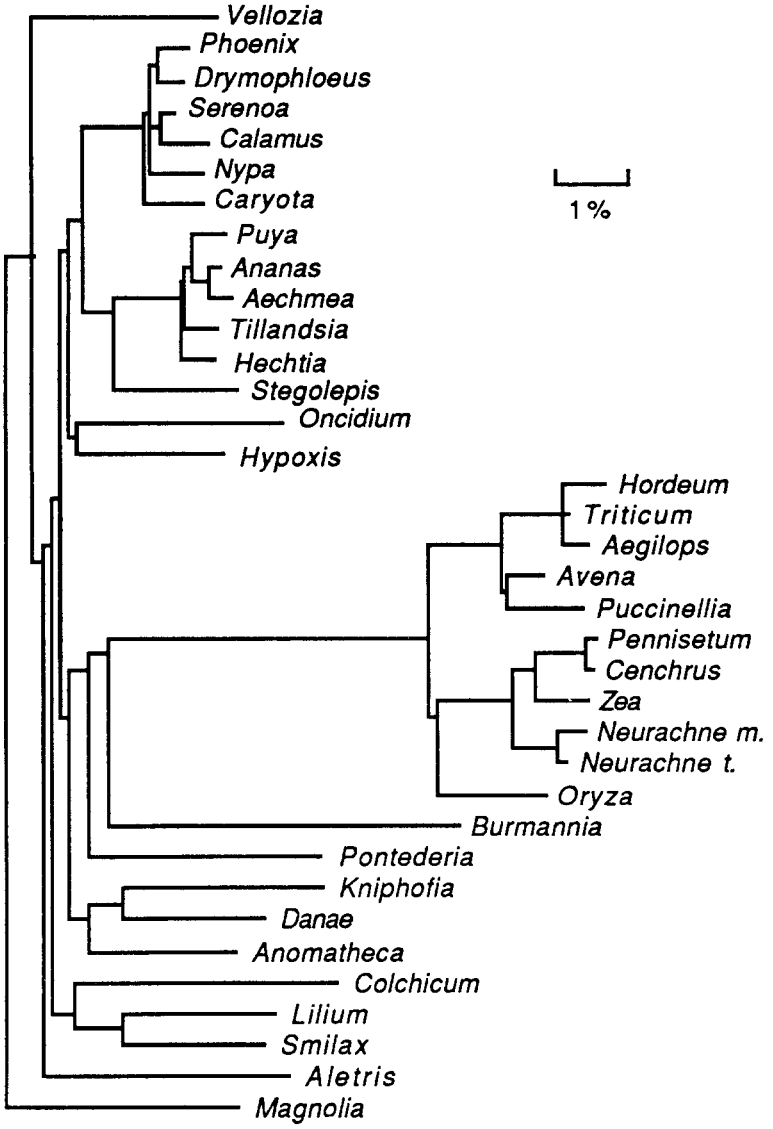

Fig. 2. A phylogeny based on the 35 monocotyledonous taxa used in this study and one dicotyledonous outgroup (*Magnolia*); 1% bar provides a rough indication of sequence divergence. The phylogeny was produced by the neighbor-joining method (Saitou and Nei 1987).

Phoenix reclinata (subfamily Coryphoideae), *N. frucitans* (subfamily Nypoideae), *Caryota mitis* (subfamily Arecoideae), and *Drymophloeus subdisticha* (subfamily Arecoideae); 1400 base pairs of *rbc*L sequence are reported for each taxon except *Nypa* (1369 bp), for which the first 31 bases of coding sequence were not determined. Sequence data for the *rbc*L of these taxa have been deposited in Genbank under accession numbers M81810 through M81815. The *rbc*L from *Zea mays* ssp. *mays* has been deposited in Genbank under accession number Z11973.

Relative Rate Tests

There are 595 possible pairs of monocot taxa, and each pair was subjected to four tests. For the tests in which all nucleotide positions were included in the data, 307 of 595 tests (51.6%) reject H_0 (see Appendix for test results). The tests which partitioned the sequence data into first-, second-, and third-codon position yielded 18 (3.0%) (data not shown), 0 (0%) (data not shown), and 304 (51.1%) (see Appendix) rejections of H_0, respectively, out of 595 tests for each category. The low percentage of rejection of H_0 in tests which used nucleotides from the first- and second-codon positions may be indicative of the low number of substitution events at these positions, probably due to selective constraint on missense substitutions. It should also be noted that individual tests are not statistically independent and thus the percentage of significant tests at all positions and at third positions are somewhat inflated. The results below pertain only to those tests in which all the nucleotide data were examined, and the results are presented by taxonomic groups (Cronquist 1988). A phylogeny is presented to indicate both rough distances between sequences and relationships between groups of sequences (Fig. 2). (For a thorough discussion of the system-

atic relationships of these species see Duvall et al. 1993; Clark et al. 1993).

Cyperales

The order Cyperales is represented by 11 members of the grass family (family Poaceae) (Table 1). All pairwise comparisons of *rbc*L from grass taxa accept H_0, indicating homogeneous rates of nucleotide substitution among *rbc*L from grass taxa. All relative rate tests which pair a grass taxon with a nongrass taxon, with the exception of *Burmannia biflora,* reject H_0.

Orchidales

Two members of the order Orchidales were analyzed. Tests of *B. biflora* (Burmanniaceae) with the *rbc*L of all grass taxa accept H_0. Rate tests which pair sequences of *B. biflora* to that of *Oncidium excavatum* (Orchidaceae), *P. sagittaria* (Pontederiaceae), and *Colchicum speciosum* (Liliaceae) also accept H_0. All remaining tests of the *rbc*L sequence of *B. biflora* to those of other taxa reject H_0.

Every comparison of *O. excavatum rbc*L to the *rbc*L of the grasses rejects H_0. Tests of the *rbc*L of *O. excavatum* with the *rbc*L of *Hypoxis leptocarpa* (Liliaceae) and the Arecales taxa also reject H_0.

Liliales

The order Liliales is represented in this analysis by ten taxa from six families (Table 1). Of these six families the family Liliaceae is the best represented (five species) in this analysis. Relative rate tests within the Liliaceae indicate some rate heterogeneity; tests of *Colchicum speciosum rbc*L to *H. leptocarpa rbc*L reject H_0. All other pairwise comparisons within the family Liliaceae accept H_0.

Rate heterogeneity is apparent when *rbc*L of *C. speciosum* is tested against the *rbc*L of *A. laxa* (Iridaceae). All other comparisons of *rbc*L among members of the order Liliales lead to the acceptance of H_0. As previously mentioned, all tests pairing *rbc*L from taxa in the Liliales to *rbc*L from taxa in the Cyperales reject H_0. Tests which pair *rbc*L from taxa in the Liliales to *rbc*L from the taxa in the Bromeliales and the Commelinales are not significant at the 1% level.

Pairings of *rbc*L of members of the Liliales with *rbc*L from members of the order Arecales separate the Liliales into two clear groups. The *rbc*L of six members of the Liliales (*Aletris farinacea, Anomatheca laxa, H. leptocarpa, Vellozia, Smilax glauca,* and *Lilium superbum*) accept H_0 in every test relative to the *rbc*L of the Arecales. The *rbc*L of four members of the Liliales (*C. speciosum, Kniphofia uvaria, P. sagittaria,* and *Danae racemosa*) reject H_0 in the majority of comparisons to the Arecales taxa.

Bromeliales

Five taxa from the Bromeliaceae are analyzed. Every test which pairs *rbc*L from bromeliads accept H_0, indicating homogeneity of substitution rate at the *rbc*L locus within the family. Tests which pair *rbc*L from a bromeliad to that of the Cyperales or *B. biflora* are significant.

Commelinales

Stegolepis allenii (Rapateaceae) is the sole representative of the Commelinales in this study. Relative rate tests of *rbc*L from *S. allenii* with those of the Cyperales and *B. biflora* reject H_0. Comparison of *S. allenii rbc*L with the *rbc*L of *O. excavatum,* the Liliales, the Bromeliales, and the Arecales are not significant.

Arecales

The palms, like the bromeliads, have a homogenous rate of *rbc*L evolution; no rate tests pairing the *rbc*L of palm taxa reject H_0. Tests of the *rbc*L of palms to the *rbc*L of the Cyperales, the Orchidales, and some of the Liliales reject H_0. All other tests involving *rbc*L of the Arecales accept H_0.

The results of the rate tests using all the nucleotide data establish that there is heterogeneity of substitution rate within taxonomic orders. The *rbc*L of *B. biflora,* for example, behaves quite unlike the *rbc*L of the other member of the Orchidales; comparison of the *rbc*L of *B. biflora* to *rbc*L from grasses leads to acceptance of H_0 while comparison of the *rbc*L of *O. excavatum* to *rbc*L of the grasses rejects H_0 in every case. Although there is little heterogeneity within the Liliales, Liliales taxa perform differentially with respect to Arecales *rbc*L. This result suggests that there are two distinct groups of Liliales with respect to rate of nucleotide substitution in *rbc*L. Conversely, the Cyperales, the Bromeliales, and the Arecales appear to represent three major *rbc*L lineages which have homogenous substitution rates.

The results of rate tests also clearly establish heterogeneity of substitution rate between major monocot lineages. The grasses are heterogenous for *rbc*L substitution rate relative to all other monocot lineages except the lineage leading to *B. biflora,* while some taxa within the Liliales and the Orchidales are heterogenous in *rbc*L substitution rate relative to the Arecales.

Estimates of Relative Rates of Nucleotide Substitution

The monocot taxa can be partitioned into eight groups which are internally homogenous in their

Table 2. Estimated differences in overall rate of *rbc*L nucleotide substitution between various groups of monocots[a]

	grasses	burm	onci	lilies I	lilies II	steg	broms	palms
grasses		1.34 (11)	2.21 (11)	2.40 (44)	3.18 (66)	4.26 (11)	4.10 (55)	5.12 (66)
burm	(1.25, 1.43)		1.73 (1)	1.81 (4)	2.76 (6)	2.76 (1)	2.86 (5)	4.10 (6)
onci	(2.00, 2.42)	NA		1.04 (4)	1.61 (6)	1.90 (1)	1.81 (5)	3.01 (6)
lilies I	(2.19, 2.61)	(1.16, 2.46)	(0.76, 1.32)		1.51 (24)	1.64 (4)	1.60 (20)	2.37 (24)
lilies II	(2.93, 3.43)	(1.86, 3.66)	(1.20, 2.02)	(1.30, 1.72)		1.02 (6)	1.00 (20)	1.54 (24)
steg	(3.74, 4.78)	NA	NA	(1.27, 2.01)	(0.97, 1.07)		1.05 (5)	1.58 (6)
broms	(3.78, 4.42)	(2.52, 3.20)	(1.50, 2.12)	(1.42, 1.78)	(0.89, 1.11)	(0.89, 1.21)		1.78 (30)
palms	(4.73, 5.51)	(3.31, 4.89)	(2.11, 3.91)	(2.09, 2.65)	(1.38, 1.70)	(1.27, 1.89)	(1.59, 1.97)	

[a] $\hat{u}$ and the number of comparisons leading to the estimation of $\hat{u}$ (in parentheses) are given above the diagonal. In each case, $\hat{u}$ represents a ratio of total substitution rates for which the group or taxon on the vertical axis is the numerator and the group or taxon on the horizontal axis is the denominator. For example, the bromeliads are estimated to have an *rbcL* substitution rate, which is 1.78 times greater than that of the palms; 95% confidence intervals are given (below diagonal). Groups are as follows (abbreviations as in Table 1): grasses (aven, pucc, neum, neut, oryz, cenc, trit, aegi, zea, hord), lilies I (colc, knip, pont, dana), lilies II (alet, lili, hypo, anom, vell, smil), steg, broms (till, puya, hech, anan, aech), palms (phoe, sere, cala, cary, nypa, drym). NA = not available

rate of nucleotide substitution at the *rbc*L locus. The groups are (1) the grass taxa (*Z. mays, Avena sativa, Puccinellia distans, Pennisetum glaucum, Neurachne munroi, Neurachne tenuifolia, Oryza sativa, Cenchrus setigerus, Triticum aestivum, Aegilops crassa,* and *Hordeum vulgare*), (2) *B. biflora,* (3) *O. excavatum,* (4) the Liliales taxa which reject H_0 relative to the *rbc*L of Arecales taxa (*C. speciosum, K. uvaria, P. sagittaria,* and *D. racemosa*), (5) the Liliales taxa which accept H_0 relative to the *rbc*L of Arecales taxa (*A. farinacea, L. superbum, H. leptocarpa, A. laxa, Vellozia,* and *S. glauca*), (6) the Bromeliales, (7) the Commeliniales, and (8) the Arecales. This partition was used to estimate relative rates of nucleotide substitution between groups.

Relative rates of nucleotide substitution between groups were estimated for both total substitution rates (Table 2) and for substitution rates at the third-codon position (Table 3). Estimates of relative rate differences suggest that the grasses have the most rapid overall nucleotide substitution rate among taxa included in this analysis. The grasses are followed in substitution rate by the *rbc*L of *B. biflora, O. excavatum,* the two groups of the Liliales, the Bromeliles, *S. allenii,* and the Arecales.

The largest differences in substitution rate occur between members of the Cyperales and members of the Arecales. *rbc*L from the grasses are estimated to have an overall substitution rate which is 5.12 times faster than that of *rbc*L from palms (Table 2). Cyperales *rbc*L are also found to have third-position substitution rates which are 8.14 times faster than the third-position rates of the Arecales (Table 3).

The *rbc*L from the grass taxa are found to have overall substitution rates which are at least four times faster than the rates of *rbc*L substitution of the Bromeliales and *S. alenii;* Cyperales *rbc*L also have an overall substitution rate which is three times faster than one group of Lilies (Table 2). The palms are remarkable for their relatively slow rates of evolution. The Arecales have twofold-slower rates of overall nucleotide substitution relative to four of seven groups (Table 2) and a twofold-slower rate of third-position substitution relative to six of seven groups (Table 3).

Discussion

The majority of relative rate tests between total *rbc*L sequences from monocot taxa lead to a rejec-

Table 3. Estimated differences in *rbc*L nucleotide substitution at the third-codon position[a]

	grasses	burm	onci	lilies I	lilies II	steg	broms	palms
grasses		1.33 (11)	2.53 (11)	3.07 (44)	4.76 (66)	5.55 (11)	6.14 (55)	8.14 (66)
burm	(1.27, 1.39)		1.92 (1)	2.27 (4)	4.18 (6)	3.05 (1)	3.65 (5)	5.91 (6)
onci	(2.15, 2.91)	NA		1.11 (4)	1.98 (6)	1.86 (1)	1.94 (5)	3.73 (6)
lilies I	(2.57, 3.57)	(0.53, 4.01)	(0.50, 1.72)		1.72 (24)	1.51 (4)	1.58 (20)	2.72 (24)
lilies II	(4.10, 5.42)	(1.81, 6.55)	(0.72, 3.24)	(1.26, 2.18)		0.83 (6)	0.87 (30)	1.55 (36)
steg	(4.87, 6.23)	NA	NA	(0.65, 2.37)	(0.04, 1.23)		1.13 (30)	2.04 (36)
broms	(4.64, 7.64)	(3.16, 4.14)	(1.63, 2.25)	(1.24, 1.92)	(0.69, 1.05)	(0.92, 1.34)		2.12 (30)
palms	(7.44, 8.84)	(5.06, 6.76)	(3.34, 4.12)	(2.14, 3.30)	(1.21, 1.80)	(1.84, 2.24)	(1.94, 2.30)	

[a] Estimates and the number of comparisons (in parentheses) are given above the diagonal; 95% confidence intervals are given below the diagonal. Groups are as defined in Table 2. NA = not available

tion of the null hypothesis, indicating that the *rbc*L gene has heterogeneous rates of nucleotide substitution among most major monocot lineages. Tests which utilize only nucleotides from the third-codon position produce results similar to the results produced by testing total sequence data. In contrast, tests which utilize data from only the first- or the second-codon position yield fewer significant results. Substitutions at first- and second-codon positions are primarily missense substitutions, and hence there are fewer substitutions at these codon positions relative to synonomous substitutions at the third-codon position. It is possible that the lack of significant results at first- and second-codon positions can be attributed to a reduction in statistical power associated with a low number of substitution events. Nonetheless, the tests which partitioned the nucleotide data into codon positions suggest, as Wu and Li (1985) and Li et al. (1987a) found, that missense substitutions demonstrate less rate variation than synonomous substitutions.

Contrary to previous studies (Doebley et al. 1990; Gaut and Clegg 1991), we do not find that the *rbc*L of *Z. mays* is accelerated in substitution rate relative to the *rbc*L of other members of the grasses. Previous studies have used an existing nucleotide sequence (McIntosh et al. 1980; Poulsen 1981; Kreppers et al. 1982), while this study employs a newly generated sequence. Maximum-likelihood relative rate tests using the old *Zea* sequence indicate acceleration of this sequence relative to other grasses (Muse and Weir 1992). Thus, the discrepancy in results lies in the differences in the two reported *rbc*L sequences, not in different methods of the relative rate test.

The discrepancy between the two maize *rbc*L sequences may be attributed to one of two factors: (1) wide chloroplast DNA variation within *Z. mays* and (2) sequence error. However, the genus *Zea* appears to have relatively little chloroplast DNA variation (Doebley et al. 1987). Owing to the cumbersome and error-prone sequencing strategies of the late 1970s, we believe all of the differences between the new *Zea rbc*L sequence and the original *Zea rbc*L sequence are the result of technical errors and are not genuine polymorphisms.

Estimates of overall rates of nucleotide substitution indicate less rate heterogeneity between the grasses and the palms than found by Wilson et al. (1990), who found an eightfold difference. Our estimate of rate differences between grasses and palms at the third position (an eightfold difference) is also lower than the Wilson et al. (1990) estimate for differences in silent site substitution rate (a 36-fold difference). The estimates of Wilson et al. (1990) relied on fossil-based divergence times and a relatively small sample of nucleotide changes, so one might expect their estimates to have large standard errors. Our estimates do not rely on fossil-based divergence times, and our sample of nucleotide changes is much greater.

Table 4. Substitution rates vs. minimum generation time (MGT)[a]

	overall	third	MGT (yrs.)	refs.
grasses	1.00	1.00	<1 to 2	Hitchcock, '35
burm	0.75	0.75	>1	Cowley, '88
onci	0.45	0.40	2 to 7	Goh et al., '82
lilies I	0.42	0.33	<1 to 15	Ivashchenko, '79
broms	0.24	0.16	>1 to 36	Augsburger, '85 Rauh, '79
palms	0.20	0.12	8 to 40	Ash, '88 D. DeMason[b]

[a] *Overall* and *third* refer to overall and third-position substitution rates, respectively, relative to the estimated substitution rate of the grasses. Groups are defined in Table 2

[b] Personal communication

Substitution rates are hypothesized to be a function of many factors including G/C content (Bulmer et al. 1991), selection (Gillespie 1986; Ohta 1987), molecular effects (such as DNA polymerase fidelity) (Wu and Li 1985; Britten 1986), and generation time (Li et al. 1985, 1987a; Wu and Li 1985). Of these, it is reasonable to dismiss G/C content as the primary causative factor of *rbc*L rate heterogeneity between lineages due to the similar G/C content of the genes examined (range: 40–45% G/C, unpublished data).

Differences in substitution rates between lineages may depend on variable generation times (T) between lineages; this is a view consistent with the "generation-time-effect" hypothesis of the neutral theory (Li et al. 1985, 1987a; Wu and Li 1985). The generation-time-effect hypothesis predicts that the rate of nucleotide substitution at the *rbc*L locus should be proportional to the inverse of the generation time (1/T).

The determination of generation time for plant species is fraught with difficulties. We do have some estimates of the time to first flowering, or minimum generation time (MGT), for some of the species in this analysis (Table 4). Table 4 outlines the relation of MGT to the rate of *rbc*L nucleotide substitution. Table 4 demonstrates that substitution rate decreases with increasing MGT, as would be predicted by the generation-time-effect hypothesis.

Thus there is a correlation between rates of nucleotide substitution at the *rbc*L locus and the MGT. It is clear, however, that a perfect correlation between MGT and rate of nucleotide substitution does not exist. The *rbc*L from perennial grass species (e.g., *P. distans* and *Neurachne* sp.) appear to have homogeneous substitution rates relative to

many annual grass species. Further, all the Liliales are perennials but members of some genera (e.g., *Pontederia*) flower within the first year of growth while members of other genera (e.g., *Colchicum*) may require up to 15 years of growth before flowering (Ivashchenko 1979). Yet, relative rate tests on *rbc*L from these two genera do not reject H_0.

The evolutionary history of generation times may be important in determining substitution rates at the *rbc*L locus. For example, while *B. biflora* is perennial, *Burmannia* is the only nongrass genus in this analysis which includes annual species (Cowley 1988). *B. biflora* is also the only nongrass taxon which accepts H_0 when tested for rate heterogeneity against the *rbc*L from grass species. Perhaps rapid substitution rates at the *rbc*L locus in *B. biflora* reflect an evolutionary history which includes short generation times. A similar argument can be made about perennial grass species; rapid substitution rates may reflect the recent acquisition of perennial generation times. It appears, then, that generation times (and an understanding of the evolutionary history of generation times) may account for heterogeneous rates of *rbc*L evolution among the analyzed monocotyledonous taxa.

Well-defined groups such as the grasses, the bromeliads, and the palms have similar rates of *rbc*L evolution, suggesting that substitution rate is a trait which reflects phylogeny in a broad sense. Relative rates of nucleotide substitution at the *rbc*L locus among the Orchidales and the Liliales appear to be heterogenous. Assuming that rates of evolution are an indication of relatedness, this analysis would tend to imply that these orders (as defined by Cronquist 1988) are not monophyletic in origin. This view is supported by both molecular (Duvall et al. 1992) and morphological (Dahlgren et al. 1985) analyses of the Orchidales and Liliales.

Many phylogenetic studies have used *rbc*L sequences. At higher taxonomic levels (e.g., above the family level) *rbc*L clearly demonstrates heterogenous rates of nucleotide substitution. Phylogenetic methods which generate ultrametric trees should be avoided with *rbc*L data sets of wide taxonomic range. At lower taxonomic levels (e.g., family and below) certain *rbc*L data sets may not violate assumptions of rate constancy. The Poaceae, the Bromeliaceae, and the Arecaceae, for example, appear to have homogeneous rates of *rbc*L evolution within families. In such cases use of ultrametric trees may be proper.

Parsimony methods of phylogenetic inference should also be used with caution on *rbc*L data sets. Despite arguments that parsimony methods make no a priori assumptions about the equality of rates between lineages (Farris 1983; Sober 1983), numerous theoretical and simulation studies have shown maximum parsimony methods to be inconsistent estimators of tree topology under conditions of unequal rates of substitution between lineages (Felsenstein 1978, 1983; Li et al. 1987b; Saitou and Imanishi 1989; Hasegawa et al. 1991). One simulation which examined the ability of maximum parsimony to reconstruct the correct topology of a tree given four operational taxonomic units (OTUs) and fivefold differences in branch lengths (which is equivalent to the rate variation found in this study) showed that maximum parsimony never selected the correct tree topology (Hasegawa et al. 1991). In another study with four OTUs (represented by 1000 base pairs of nucleotide sequence) with different branch lengths leading to each OTU (and a maximum rate difference of four between branches), maximum parsimony selected the correct topology only 35% of the time (Li et al. 1987b). With six OTUs represented by 600 base pairs of sequence and a maximum rate difference of 8 between branches, maximum parsimony chose the correct topology for the phylogeny an average of 33% of the time (Saitou and Imanishi 1989). Simulations have clearly shown that maximum likelihood methods (Felsenstein 1981)—and frequently distance methods—outperform parsimony methods of phylogeny reconstruction under conditions of unequal rates of nucleotide substitution.

It has been shown that additional taxa increase the reliability of the parsimony method (Penny et al. 1987), so studies which use parsimony methods and a great many taxa may be more reliable than the simulation studies suggest. For example, the application of parsimony and maximum likelihood algorithms to a data set of 79 sequences produces topologies which cannot be distinguished as statistically different by the criterion of Kishino and Hasegawa (1989; Duvall et al. 1992).

The relative rate test is not without its limitations. Fitch (1976) pointed out that the relative rate test cannot detect changes in evolutionary rates if substitution rates between lineages change proportionally. The relative rate test also cannot detect stochastic changes in rate within a lineage (such as described by Gillespie 1984), because the relative rate test compares average substitution rates between two lineages. Nonetheless, unequal rates clearly exist at the *rbc*L locus in monocotyledonous plants, and it appears that this variation loosely conforms to the predictions of the generation-time-effect hypothesis.

Acknowledgments. The authors would like to thank M.W. Chase and H.G. Hills for the use of unpublished sequence data. The authors are grateful to B.S. Weir and L.E. Eguiarte for helpful discussion, G.H. Learn for phylogenetic analysis, and M.R. Duvall, B. Morton, G. Huttley, and two anonymous re-

viewers for helpful comments on earlier versions of the manuscript. This research was supported by NIH grant GM45144 to M.T.C., NIH grant GM45344 to North Carolina State University, an NSF Graduate Fellowship to S.V.M., and NSF grant BSR-8904637 to W.D.C.

References

Ash J (1988) Demography and production of *Balaka microcarpa* Burret (Arecaceae), a tropical understory palm in Fiji. Aust J Bot 6:67–68

Augsburger CK (1985) Demography and life history variation of *Puya dasylirioides*, a long-lived rosette in tropical subalpine bogs. Oikos 45(3):341–352

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

Bulmer R, Wolfe KH, Sharp PM (1991) Synonomous nucleotide substitution rates in mammalian genes: implication for the molecular clock and the relationship of mammalian orders. Proc Natl Acad Sci USA 88:5974–5978

Clark WD, Gaut BS, Duvall MR, Clegg, MT (1993) Phylogenetic relationships of the Bromeliaceae based on *rbc*L sequence comparisons. Ann Miss Bot Gard, submitted

Cowley EJ (1988) Burmmaniaceae. In: Polhill RM (ed) Flora of tropical East Africa. AA Balkema, Rotterdam, pp 1–9

Cronquist A (1988) The evolution and classification of flowering plants, 2nd edition. The New York Botanical Garden, Bronx, NY

Dahlgren RMT, Clifford HT, Yeo PF (1985) The families of the Monocotyledons: structure, evolution and taxonomy. Springer-Verlag, Berlin

Doebley J, Renfroe W, Blanton A (1987) Restriction site variation in the Zea chloroplast genome. Genetics 117:139–147

Doebley J, Durbin M, Golenberg EM, Clegg MT, Ma D-P (1990) Evolutionary analysis of the large subunit of carboxylase (*rbc*L) nucleotide sequence among the Grasses (Gramineae). Evolution 44:1097–1108

Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure from small quantities of fresh leaf tissue. Phytochem Bull 19:11–15

Duvall MR, Clegg MT, Chase MW, Clark WD, Kress WJ, Zimmer EA, Hills HG, Eguiarte LE, Smith JF, Gaut BS, Learn GH (1993) Rapid radiation of ancestral monocotyledons into seven primary lineages indicated by analysis of DNA sequence of a plastid locus obtained from 104 species. Ann Miss Bot Gard, in press

Farris JS (1983) The logical basis of phylogenetic analysis. In: Platnick NI, Funk VA (eds) Advances in cladistics vol. 2. Columbia University Press, New York, pp 7–36

Felsenstein JF (1978) Cases in which parsimony or compatibility methods will be positively misleading. Syst Zoo 27:401–410

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

Felsenstein JF (1983) Parsimony in systematics: biological and statistical issues. Ann Rev Ecol Syst 14:313–333

Fitch WM (1976) Molecular evolutionary clocks. In: Ayala FJ (ed) Molecular evolution. Sinauer Associates, Sunderland, MA, pp 160–178

Garcia P, Clegg MT (1991) The gene *rbc*L for *Avena sativa* L. 26th Portugese-Spanish meetings of Genetics. Coimbra

Gaut BS, Clegg MT (1991) Molecular evolution of alcohol dehydrogenase 1 in members of the grass family. Proc Natl Acad Sci USA 88:2060–2064

Giannasi DE, Zurawski G, Learn GH, Clegg MT (1992) Evolutionary relationships of the Caryophillidae based on comparative *rbc*L sequences. Syst Bot 17:1–5

Gillespie JH (1984) The molecular clock may be an episodic clock. Proc Natl Acad Sci USA 81:8009–8013

Gillespie JH (1986) Natural selection and the molecular clock. Mol Biol Evol 3:138–155

Goh C-J, Strauss MS, Arditti J (1982) Flower induction and physiology in orchids. In: Arditti J (ed) Orchid biology, reviews and perspectives, vol 2. Cornell University Press, Ithaca, pp 213–241

Golenberg EM, Giannasi DE, Clegg MT, Smiley CJ, Durbin M, Henderson D, Zurawski G (1990) Chloroplast DNA sequence from a Miocene *Magnolia* species. Nature 344(6267):656–658

Hasegawa M, Kishino H, Saitou N (1991) On the maximum likelihood method in molecular phylogenetics. J Mol Evol 32: 443–445

Hitchcock AS (1935) Manual of the grasses of the United States. US Government Printing Office, Washington

Hudson GS, Mahon JD, Anderson PA, Gibbs MJ, Badger MR, Andrew TJ, Whitfield PR (1990) Comparisons of the *rbc*L genes for the large subunit of ribulose bisphosphate carboxyase from closely related C3 and C4 plant species. J Biol Chem 265:808–814

Ivashchenko AA (1979) Characteristics of the greater life cycle of the Yellow Autumn Crocus at the northern boundary of its range. Sov J Ecol 10:431–432

Kimura M (1983) The neutral theory of molecular evolution. Cambridge University Press, Cambridge, p 70

Kimura M (1989) The neutral theory of evolution and the world view of the neutralists. Genome 31:24–31

Kishino H, Hasegawa M (1989) Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in Hominoidea. J Mol Evol 16:111–120

Kreppers ET, Larrinua IM, McIntosh L, Bogorad L (1982) The maize chloroplast genes for the β and ϵ subunits for the photosynthetic coupling factor CF_1 are fused. Nucleics Acid Res 10:4985–5002

Li W-H, Luo C-C, Wu C-I (1985) Evolution of DNA sequences. In: Macintyre RJ (ed) Molecular evolutionary genetics. Plenum, New York, pp 1–94

Li W-H, Tanimura M, Sharp PM (1987a) An evaluation of the molecular clock hypothesis using mammalian DNA sequences. J Mol Evol 25:330–342

Li W-H, Wolfe KH, Sourdis J, Sharp PM (1987b) Reconstruction of phylogenetic trees and estimation of divergence times under nonconstant rates of evolution. Cold Spring Harbor Symp Quant Biol 52:874–856

McIntosh L, Poulsen C, Bogorad L (1980) Chloroplast gene sequences for the large subunit of ribulose bisphosphate carboxylase of maize. Nature 288:556–560

Moon E, Kao TH, Wu R (1987) Rice chloroplast DNA molecules are heterogenous as revealed by DNA sequences of a cluster of genes. Nucl Acids Res 15:611–630

Muse SV, and Weir BS (1992) Testing for equality of evolutionary rates. Genetics, in press

Ohta T (1987) Very slightly deleterious mutations and the molecular clock. J Mol Evol 26:1–6

Penny D, Hendy MD, Henderson IM (1987) Reliability of evolutionary trees. Cold Spring Harbor Symp Quant Biol 52:857–862

Poulsen C (1981) Comments on the structure and function of the large subunit of the enzyme ribulose bisphosphate carboxylase-oxygenase. Carsberg Res Commun 46:259–278

Rauh W (1979) Bromeliads for home, garden and greenhouse. Blandford Press, Poole, p 19

Saitou N, Nei M (1987) The neighbor-joining method for molecular phylogeny. J Mol Evol 27:261–273

Saitou N, Imanishi T (1989) Relative efficiencies of the Fitch-Margoliash, Maximum-Parsimony, Maximum-Likelihood,

Minimum-Evolution, and Neighbor-joining Methods of phylogenetic tree construction in obtaining the correct tree. Mol Biol Evol 6:514–525

Sanger FS, Nicklen S, Colson AR (1977) DNA sequencing with chain terminating inhibitors. Proc Natl Acad Sci USA 74: 5463–5467

Smith JF, Doyle JJ (1986) Chloroplast DNA variation and evolution in the Juglandaceae. Am J Bot 78:730

Sober E (1983) Parsimony in systematics: philosophical issues. Ann Rev Ecol Syst 14:335–357

Soltis DE, Soltis PS, Clegg MT, Durbin M (1990) *rbc*L sequence divergence and phylogenetic relationships in Saxifragaceae *sensu lato*. Proc Natl Acad Sci USA 87:4640–4644

Terachi T, Ofihara Y, Tsunewaki K (1987) The molecular basis of genetic diversity among cytoplasms of *Triticum* and *Aegilops*. VI. Complete nucleotide sequences of the *rbc*L genes encoding H and L type rubisco large subunits in common wheat and *Aegilops crassa*. 4X. Jpn J Genet 62:375–387

Uhl NW, Dransfield J (1987) Genera Plamarum, LH Bailey Hortorium and the International Palm Society. Allen Press, Lawrence, Kansas

Wilson AC, Carlson SS, White T (1977) Biochemical evolution. Ann Rev Biochem 46:573–639

Wilson MA, Gaut B, Clegg MT (1990) Chloroplast DNA evolves slowly in the Palm family (Arecaceae). Mol Biol Evol 7:303–314

Wu C-I, Li W-H (1985) Evidence for higher rates of nucleotide substitution in rodents than in man. Proc Natl Acad Sci USA 82:1741–1745

Zuckerkandl E, Pauling L (1965) In: Bryson V, Vogel HJ (eds) Evolving genes and proteins. Academic Press, New York, pp 97–166

Zurawski G, Clegg MT, Brown AHD (1984) The nature of nucleotide sequence divergence between barley and maize chloroplast DNA. Genetics 106:735–749

Received November 30, 1991/Revised and Accepted May 1, 1992

Appendix. Results of relative rates tests. Results using all sequence data are given above the diagonal. Results using only third-codon-position data are given below the diagonal. Species A and B are found on the horizontal and vertical axes; *Magnolia rbc*L was used as the outgroup sequence in every test. Abbreviations found in Table 1. The test statistic is distributed χ^2 with 2 *df* ($p(\chi^2 > 9.21) < 0.01$)

	zea	hord	neum	aegi	oryz	cen	neut	penn	pucc	aven	trit	burm	onci	colc	pont	knip	dana
zea		2.10	0.47	0.47	0.53	1.77	0.11	0.90	0.54	2.52	0.76	3.90	19.66	17.28	21.32	26.86	31.02
hord	2.81		0.24	0.00	0.85	0.82	0.87	1.75	3.54	5.23	0.00	4.32	22.76	17.88	24.65	26.57	36.66
neum	0.54	0.67		0.30	0.32	0.22	3.09	0.16	1.66	2.47	0.68	4.45	22.00	21.14	20.82	28.42	37.60
aegi	0.76	0.00	0.06		0.01	0.07	0.35	0.18	1.37	3.10	0.00	3.38	21.18	19.82	23.38	25.29	34.49
oryz	0.83	3.38	0.00	1.80		0.13	0.28	0.14	0.52	1.26	0.09	3.25	23.16	19.46	18.38	24.19	33.87
cenc	1.56	1.56	0.00	0.06	0.00		0.64	0.00	1.29	1.98	0.31	3.74	20.86	19.52	20.66	25.48	35.58
neut	0.30	1.28	0.00	0.43	0.11	0.00		0.23	0.41	1.14	0.44	3.32	19.89	18.30	18.21	24.66	32.67
penn	0.59	3.22	0.00	1.02	0.00	0.00	0.00		0.91	1.71	0.42	3.61	20.41	18.83	18.94	24.55	33.63
pucc	0.57	3.06	0.00	1.03	0.00	0.00	0.00	0.00		3.94	1.12	1.46	17.76	16.78	18.18	21.72	29.66
aven	0.28	0.00	0.00	0.00	0.35	0.00	0.00	0.06	0.00		1.47	1.46	16.59	15.26	16.77	19.98	29.23
trit	0.76	0.00	0.06	0.00	1.79	0.06	0.43	1.02	1.03	0.00		2.82	19.66	18.34	22.16	23.57	32.41
burm	1.94	2.45	2.77	1.83	3.04	2.44	2.00	2.47	2.45	2.06	1.83		8.62	6.23	7.53	12.03	20.11
onci	14.81	19.86	17.39	17.34	20.72	17.07	15.40	16.49	17.29	18.30	17.34	7.59		0.20	0.54	0.52	8.20
colc	14.28	17.16	17.74	16.90	18.19	16.83	15.15	16.10	17.38	17.57	16.91	6.08	0.04		1.06	0.54	5.98
pont	15.21	20.90	17.58	20.18	17.96	17.39	15.20	15.76	18.70	18.02	20.18	7.09	1.27	0.88		1.44	3.22
knip	24.54	22.73	29.03	24.14	27.42	26.44	26.10	25.87	25.79	25.21	24.14	14.39	2.20	1.44	3.84		4.91
dana	29.77	35.19	36.98	33.48	34.87	36.12	33.28	34.60	34.00	34.30	33.48	21.30	3.46	3.07	3.29	6.51	
alet	40.97	49.29	48.20	50.92	50.51	47.27	44.33	45.47	47.32	49.12	50.92	32.00	10.21	10.00	10.18	8.07	2.37
lili	29.37	32.25	33.30	31.28	34.38	31.90	30.27	31.36	30.70	30.71	31.30	19.48	4.07	3.54	6.09	0.51	2.96
smil	34.77	36.01	37.82	35.91	36.27	36.26	34.68	35.02	35.10	35.97	35.91	23.25	4.91	4.90	7.25	1.48	2.09
vell	34.35	37.84	39.30	39.64	40.16	38.69	36.60	36.36	38.25	38.56	39.65	23.41	3.37	5.33	5.81	4.38	0.77
hypo	47.27	45.68	51.73	50.02	50.04	52.19	48.89	49.99	48.92	50.01	50.02	38.79	14.86	11.53	9.92	8.54	3.99
anom	38.10	42.92	44.64	42.25	44.54	42.65	41.92	41.67	42.50	42.12	42.25	28.42	7.66	8.05	7.89	12.29	1.44
steg	33.57	43.65	37.96	39.41	41.68	38.00	35.07	36.55	38.02	39.58	39.42	19.06	4.07	3.26	5.24	2.63	0.40
aech	35.72	42.61	38.32	40.44	42.95	39.89	35.49	38.61	38.41	40.68	40.46	20.99	3.26	3.43	5.39	3.05	0.48
puya	34.89	39.65	37.94	38.48	42.52	38.50	35.18	37.45	36.07	38.00	38.51	20.90	4.04	3.38	5.75	2.11	1.19
till	36.64	45.96	39.84	43.08	45.19	41.24	37.05	39.73	40.74	43.30	43.11	22.91	4.91	4.62	7.15	4.35	0.30
anan	38.85	46.64	45.61	44.03	46.30	43.06	38.76	41.63	41.57	44.25	44.03	23.91	5.59	4.89	7.41	4.42	0.63
hech	34.98	43.96	38.41	41.09	42.83	38.51	35.81	36.74	38.22	40.99	41.10	23.58	4.81	6.05	7.00	6.41	0.31
nypa	54.21	53.88	60.77	58.27	61.34	59.04	57.69	57.34	57.15	58.22	58.27	41.54	16.67	15.42	17.15	9.46	7.14
cary	50.53	52.54	55.21	54.40	56.72	54.91	52.19	53.05	53.38	54.28	54.52	36.93	14.40	13.12	16.40	9.25	4.67
cala	51.15	55.60	57.36	54.67	59.38	55.53	54.25	53.34	53.59	57.35	54.67	37.35	15.52	14.19	17.13	9.94	5.19
drym	54.06	57.47	60.70	60.55	61.99	58.92	57.56	57.25	59.38	60.49	60.55	40.31	18.00	14.48	16.62	9.17	6.97
sere	55.87	57.48	61.09	59.27	63.08	60.39	57.90	58.26	58.05	59.12	59.30	39.41	17.15	14.84	16.92	10.56	6.02
phoe	55.79	55.73	62.25	58.92	62.11	60.09	59.10	57.81	59.19	60.21	58.92	41.63	17.17	15.69	17.20	11.46	5.86

Appendix. Extended

alet	lili	smil	vell	hypo	anom	steg	aech	puya	till	anan	hech	nypa	cary	cala	drym	sere	phoe
30.27	36.03	43.74	43.67	47.11	45.32	46.66	45.78	43.00	47.31	49.28	48.75	57.61	58.07	62.13	64.21	67.68	70.00
32.12	36.48	41.53	41.75	44.05	45.75	51.43	50.18	46.54	54.63	57.71	53.79	50.67	54.48	63.38	63.23	64.82	65.27
34.57	39.33	45.69	47.02	52.34	60.05	50.63	45.80	44.56	49.27	51.17	50.97	61.52	61.10	67.05	70.13	71.23	75.52
34.25	36.82	42.82	44.36	49.12	46.74	51.23	46.52	45.23	52.03	53.18	51.13	58.72	60.04	64.04	70.12	69.02	71.75
32.75	37.73	41.69	45.21	46.94	48.25	52.04	47.44	46.85	52.05	53.00	53.27	60.88	60.75	67.38	70.74	71.67	74.36
32.93	36.22	42.89	45.23	48.86	47.05	48.33	45.03	43.25	48.93	50.18	49.60	58.71	59.03	63.02	65.37	68.70	70.78
30.45	35.15	42.16	42.84	48.26	47.07	46.64	42.32	41.48	45.51	47.40	46.93	59.29	57.22	64.00	67.05	66.98	72.54
31.43	35.50	41.68	43.19	46.68	46.02	45.65	43.55	41.99	47.24	48.46	47.65	57.28	57.21	60.59	63.71	66.55	68.37
28.22	31.50	37.15	39.68	43.25	42.55	44.05	39.81	37.48	43.73	44.84	43.88	52.62	53.29	57.14	62.12	61.35	65.77
27.45	29.57	34.89	38.34	41.38	40.53	43.06	39.34	36.67	43.58	44.79	42.27	50.67	51.50	57.93	60.24	59.65	64.09
32.25	34.71	40.53	42.75	46.42	44.24	48.33	43.85	42.70	49.04	50.26	49.06	55.88	57.04	60.91	66.80	65.67	68.28
18.69	20.50	27.41	27.54	32.92	30.33	25.38	23.86	23.13	26.11	27.92	29.84	37.87	37.85	42.31	45.47	43.71	49.32
3.85	3.00	5.40	4.39	12.11	8.89	7.34	4.38	4.91	6.88	8.30	7.78	12.40	14.70	18.83	20.23	20.52	22.64
3.94	3.57	6.21	5.76	10.76	9.60	5.69	0.30	4.30	6.35	6.88	8.24	13.36	14.35	18.08	16.37	17.78	20.75
2.20	4.84	9.07	5.56	6.40	6.60	7.21	5.54	6.02	7.81	9.16	8.64	10.05	12.68	17.52	16.96	17.39	19.16
1.59	1.63	3.72	3.82	6.20	8.24	3.60	2.77	2.55	3.56	4.62	5.11	9.02	9.18	12.00	12.55	12.16	14.62
1.15	6.61	9.16	7.24	4.68	4.17	5.90	4.09	5.61	4.07	4.57	3.84	13.11	9.98	11.45	17.68	14.64	15.16
	1.67	3.61	2.15	1.40	1.28	1.43	0.62	0.96	0.59	1.01	1.13	5.03	3.75	5.50	7.77	6.77	8.47
4.78		1.14	0.56	3.08	4.07	0.48	0.56	0.21	1.22	1.32	2.64	3.47	4.58	6.68	6.01	6.31	7.78
3.03	0.24		0.77	3.80	5.60	0.56	0.62	0.22	1.15	0.94	2.78	2.38	3.70	5.97	3.50	5.00	6.96
0.94	2.84	1.59		2.18	2.36	0.40	0.31	0.83	0.07	0.16	1.41	1.38	1.28	2.72	2.21	2.35	3.57
0.28	4.22	2.34	1.65		0.03	3.02	2.69	2.73	1.91	1.56	0.68	3.28	1.44	1.84	5.72	2.86	4.26
0.27	6.06	4.32	0.99	1.79		1.82	2.08	3.71	1.64	1.38	0.40	3.23	1.71	1.34	5.31	2.77	4.06
2.09	1.23	0.55	0.04	2.42	1.80		0.08	0.20	0.28	0.47	1.28	2.03	2.09	4.10	3.49	3.82	5.85
2.33	1.24	1.26	0.34	1.76	1.53	0.30		1.29	0.18	0.00	1.93	3.14	2.79	4.19	5.53	5.14	6.71
2.35	0.63	0.71	0.25	2.43	2.89	0.21	2.09		2.37	4.15	4.74	3.84	3.73	5.58	6.08	6.34	7.96
1.20	2.26	1.82	0.64	1.00	1.29	0.57	0.25	0.00		0.31	0.60	3.62	2.60	3.56	5.77	4.96	6.51
1.31	1.78	1.94	0.34	0.83	0.70	0.30	0.00	3.74	0.30		0.84	2.18	1.56	2.84	4.17	3.78	5.06
1.04	5.55	2.36	0.42	2.40	0.27	1.08	2.21	0.00	1.22	1.92		3.49	1.99	2.92	5.20	3.85	4.67
1.81	5.52	4.06	4.01	1.05	5.48	5.62	5.47	6.39	5.79	4.47	6.37		0.34	2.25	1.13	1.35	2.66
0.42	5.81	3.68	1.56	0.02	1.85	3.68	3.36	4.32	3.17	2.41	3.09	1.39		1.42	1.47	2.02	3.93
0.59	6.03	3.88	1.94	0.19	2.42	4.03	3.31	4.37	3.23	2.54	3.55	0.00	0.08		0.00	0.43	1.40
1.78	5.49	3.66	3.54	0.00	5.31	4.68	4.77	5.64	5.41	3.84	5.90	0.00	0.96	0.00		2.93	5.03
0.72	6.83	4.55	2.47	0.28	2.60	4.86	4.74	6.00	4.48	3.78	4.58	0.00	1.11	0.66	0.00		0.64
0.97	7.45	5.36	2.74	0.65	2.73	5.39	4.77	6.04	4.78	3.86	4.31	0.00	1.29	0.55	0.00	0.03	