Syst. Biol. 51(4):638–647, 2002
DOI: 10.1080/10635150290102348

Phylogeny Reconstruction and Functional Constraints in Organellar Genomes: Plastid *atpB* and *rbcL* Sequences Versus Animal Mitochondrion

VINCENT SAVOLAINEN,[1] MARK W. CHASE,[1] NICOLAS SALAMIN,[2] DOUGLAS E. SOLTIS,[3] PAMELA S. SOLTIS,[4] ANDRÉS J. LÓPEZ,[5] OLIVIER FÉDRIGO,[5] AND GAVIN J. P. NAYLOR[5]

[1]*Molecular Systematics Section, Jodrell Laboratory, Royal Botanic Gardens, Kew, Richmond, TW9 3DS, U.K.; E-mail: v.savolainen@kew.org (V.S.), m.chase@kew.org (M.W.C.)*
[2]*Department of Botany, Trinity College, Dublin 2, Ireland; E-mail: salaminn@tcd.ie*
[3]*Department of Botany and the Genetics Institute, University of Florida, Gainesville, Florida 32611, USA; E-mail: dsoltis@botany.ufl.edu*
[4]*Florida Museum of Natural History and the Genetics Institute, University of Florida, Gainesville, Florida 32611, USA; E-mail: psoltis@flmnh.ufl.edu*
[5]*Department of Zoology and Genetics, Iowa State University, Ames, Iowa 50011, USA; E-mail: andresl@iastate.edu (A.L.), ofedrigo@iastate.edu (O.F.), gnaylor@iastate.edu (G.J.P.N.)*

Interest in phylogeny reconstruction has increased so rapidly during the past decade that now roughly 4,000 articles that include a phylogenetic tree are published each year (Pagel, 1999). With the increasing numbers of simultaneously analyzed taxa (Chase et al., 1993; Källersjö et al., 1998; Savolainen et al., 2000a, 2000b; Soltis et al., 1999, 2000), accuracy of such large molecular trees has been questioned (Graybeal, 1998; Hillis et al., 1996; Hillis, 1998), and one of the key problems has been to discriminate between phylogenetic signal and "noise." Because constraints acting on DNA can strongly bias the assessment of homology (viz. homoplasy due to convergence and parallelism preserving function and structure of proteins), characterization of these evolutionary forces may shed light on the robustness of phylogenetic inference (Naylor et al., 1995; Naylor and Brown, 1997, 1998).

Because hundreds of organelles are present in each cell, thereby making haploid organellar DNA easier to amplify and sequence than single-copy nuclear (diploid) genes, most phylogenetic trees have been based on these data: Mitochondrial DNA is generally used to infer mammalian phylogenies (Waddell et al., 1999), including hominids (Kings et al., 1997; Ovchinnikov et al., 2000), whereas plant studies are mostly performed using plastid DNA (e.g., Chase et al., 1993; Källersjö et al., 1998; Savolainen et al., 2000a, 2000b; but many studies at lower levels are also based on nuclear ITS sequences, as reviewed in Soltis and Soltis, 1998). Using sequences of the entire coding mitochondrial genome in a wide range of animals, Naylor and Brown (1998) suggested that recognizing misleading signals required integrating knowledge of structure and function of encoded products. This is important since functional requirements (e.g., chemical properties, charge, and hydrophobicity) in the animal mitochondrial genome seriously compromised estimation of relationships, with incorrect trees sometimes receiving high bootstrap percentages (Naylor and Brown, 1997, 1998).

To determine whether these issues were a concern in studies of seed plant relationships, we investigated the substitution properties of plastid *rbcL* and *atpB* sequences. This was done assuming a phylogenetic tree previously reconstructed for concatenated plastid (*rbcL* and *atpB*) and nuclear 18S rDNA sequences (Soltis et al., 1999, 2000), which has led to better resolution and higher internal support than analyses based on single genes (Chase and Cox, 1998; Soltis et al., 1997, 1998). We have looked for an association of functional constraints and phylogenetic informativeness in plastid DNA by calculating the retention index (RI, formally "ri" in Farris, 1989) of each site according to codon position, amino acid, chemical properties, charge, and hydrophobicity (*atpB* and *rbcL* were optimized on the combined three-gene tree published by Soltis et al., 1999 and Soltis et al., 2000; sites were classified according to coded

amino acids). We have also compared the RI for *rbcL*, *atpB*, and 18S rDNA; they exhibited 1072, 1018, and 857 variable sites, respectively.

If the consistency index (CI) is the simplest measure of the amount of homoplasy, the RI relates to the pattern of homoplasy (Farris, 1989). The CI was first introduced by Kluge and Farris (1969) as a measure of fit of a character to a tree. For any character, CI is defined as $CI = m/s$ where m is the minimum amount of change the character may show on any tree, and s denotes the observed quantity of change (steps) in the cladogram (perfect fit $CI = 1$). However, "it might be desirable to use a measure that reaches zero when a character fits the tree as poorly as possible" (Farris, 1989, p. 418); the CI does not have this property, being no less than m/g where g is the greatest number of steps that the character may require on any tree. Therefore, Farris (1989) defined the RI as, for any character, $RI = (g - s)/(g - m)$. RI measures how many times a character changes over a tree versus how many times it would have changed on an unresolved bush; it is a measure of phylogenetic signal, and we expect low RIs for variable sites constrained by function.

567-Taxon Plant Phylogenetic Analysis

In contrast to mitochondrial genes for which all comparisons were highly significant (RIs varied relative to functional constraints, $P < 0.0005$; Naylor et al., 1995; Naylor and Brown, 1997, 1998), we found phylogenetic information in plastid DNA in seed plants to be evenly distributed among sites (see Fig. 1). RIs were not significantly different for genes, charge, or hydrophobicity. Differences were detected for codon positions, chemical properties, and amino acids. However, in the case of chemical properties, statistical significance was due to sites coding for aromatic and sulphur-containing amino acids, which had higher RIs (when these sites

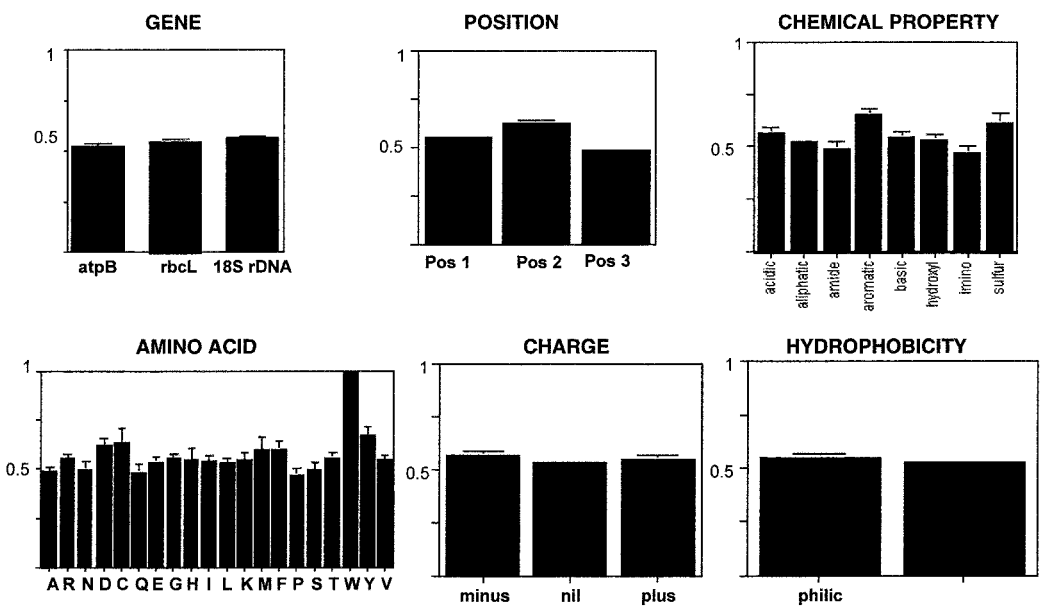

FIGURE 1. Relationship between functional characteristics and phylogenetic signal (as measured by the RI; y-axis, poorest fit = 0, strongest signal = 1) in *rbcL* and *atpB* sequences data from 567 seed plants (see text for details). No significant differences were found for genes (Kruskal-Wallis $H = 5.15$, $P = 0.076$), charge (Kruskal-Wallis $H = 2.29$, $P = 0.318$), and hydrophobicity (Mann-Whitney $U = 2236349.5$, $P = 0.171$; e.g., see Zar, 1999, for Mann-Whitney tests). Differences were detected in codon positions (Kruskal-Wallis $H = 55.750$, $P = 0.000$), chemical properties (Kruskal-Wallis $H = 26.03$, $P = 0.000$), and amino acids (Kruskal-Wallis $H = 43.82$, $P = 0.001$), but see text. Nonparametric tests were performed because RIs did not follow a normal distribution. Compare this result for plastid DNA with results from animal mitochondria in Table 2 and in Figure 5 of Naylor and Brown (1998, p. 68): analysis of variance indicated that all factors in animal mitochondrial genes had highly significant effects on RI; $P = 0.0005$ in each case.

were excluded from the analysis, Kruskal-Wallis $H = 8.739$, $P = 0.1200$; e.g., see Zar, 1999, for statistical tests); for amino acids, significance was due to the sites coding for tryptophan, which fitted the tree perfectly (with tryptophan excluded from the analysis, Kruskal-Wallis $H = 26.168$, $P = 0.096$, Fig. 1). These latter sites represent, however, a minority of all characters involved in each category (aromatic, 9%; sulphur, 3.7%; tryptophan, 8.3%); the high RIs that caused the rejection of the null hypothesis may be due to the small number of characters present in these categories. The general trend is for no significant differences.

Unlike animal mitochondrial DNA, phylogenetic informativeness in seed plant plastid DNA appears evenly dispersed across functional categories of sites. This would imply that protein-coding regions of plastid genomes do not suffer from the same differential constraints seen in the mitochondrial genomes of animals. If this is truly the case, then plastid sequences are less likely to be plagued by problems of among-site and among-lineage rate heterogeneity, which has often compromised the accuracy of phylogenetic inference in animals (Naylor and Brown, 1997, 1998). The alpha shape parameter for the gamma distribution (Yang, 1996) is slightly higher in the plastid genes (*atpB* + *rbcL* for the plant tree) compared to the animal tree: 0.5824 versus 0.4586. To explain the differences we propose several explanations: (1) the rate of evolution in plastid genomes might be so slow that it has not proceeded to the point at which constraints cause signal saturation. We found that the average pair-wise divergence for the plant data set is roughly 1/7 (0.0639 ± 0.0140 versus 0.4313 ± 0.0983; maximum divergence within plants is 0.1873; see Table 1) that of the animal data set of Naylor and Brown (1997, 1998); thus, the animal mitochondrial genome indeed evolves much faster than plastid DNA, and this might, in theory, result in higher levels of homoplasy that would be detected as "noise" in phylogenetic studies. However, low average pair-wise distances could result from either low rates of change and/or a relatively younger age of divergence. The animals examined by Naylor and Brown are early Paleozoic in age (Benton, 1993; Naylor and Brown, 1998), and, despite the fact that the plant tree also included all major lineages of seed plants (except cycads), many angiosperms were much younger. Seed plants are mid-Paleozoic (Soltis et al., 2002), whereas angiosperms are mid-Mesozoic (Wikström et al., 2001); a fairer comparison might be to use mammals since they are of similar age to the angiosperms, predating the K/T boundary around 100 million years ago (Penny et al., 1999). We have repeated all analyses but looking at an association between functional characteristics and phylogenetic signal (RI) in 13 mammal mitochondrial genes from the rat, mouse, cow, opossum, and two species of whales (viz. all mammals in Naylor and Brown's data set, 1998; accounting for ca. 1/3 of the total tree length). For all categories except for hydrophobicity, RIs were significantly different among sites, thereby showing that phylogenetic signal was also compromised by functional constraints in mammal mitochondrial DNA (Table 2). However, note that despite the fact that they are of similar ages, average pair-wise divergence in mammal mitochondrial sequences is higher than plastid *rbcL* and *atpB* sequences (0.2580 ± 0.0600 versus 0.0639 ± 0.0140; Table 1) consistent with the suggestion that the former genome evolves faster (Palmer and Delwiche, 1998).

Comparing Similar Taxon Sampling

Two other hypotheses may explain the differences between Naylor and Brown's results (1997, 1998) and ours (this paper, Fig. 1): (2) the sampling in the three-gene plant analysis was more thorough than was that of animals investigated by Naylor and Brown, which

TABLE 1. Pair-wise divergence in data sets (see text for details).

Matrices	Average distances ± S.D.	Minimum	Maximum
Animals (full matrix)	0.4313 ± 0.0983	0.0856	0.6095
Mammals	0.2580 ± 0.0600	0.0856	0.3038
Plants (full matrix)	0.0639 ± 0.0140	0.0020	0.1873
Plants (random 19-taxon matrices)	0.0852 ± 0.0207	0.0055	0.1781
Plants (lineage specific 19-taxon matrices)	0.0535 ± 0.0199	0.0020	0.1246

TABLE 2. Relationship between functional characteristics and phylogenetic signal (as measured by the RI) in a subset of Naylor and Brown's 1998 data comprising 13 mitochondrial coding genes but only for the mammal species (*Mus musculus*, *Rattus norvegicus*, *Bos taurus*, *Balaneopterus physalus*, *Balaneopterus musculus*, *Didelphis virginiana*).

Categories	Statistics	*P*-values	Sites	RI means ± S.D.
Codon positions	Kruskal-Wallis $H = 209.3799$	$P = 0.000$	Pos1	0.6840 ± 0.0152
			Pos2	0.7757 ± 0.0196
			Pos3	0.4753 ± 0.0104
Genes	Kruskal-Wallis $H = 43.8888$	$P = 0.000$	atp6	0.5934 ± 0.0328
			atp8	0.6538 ± 0.0568
			CO1	0.4677 ± 0.2485
			CO2	0.5102 ± 0.0396
			cytb	0.5364 ± 0.0281
			ND1	0.5042 ± 0.0310
			ND2	0.6150 ± 0.0241
			ND3	0.5727 ± 0.0441
			ND4	0.5611 ± 0.0230
			ND4L	0.6609 ± 0.0500
			ND5	0.5934 ± 0.0189
			ND6	0.6627 ± 0.0344
Charges	Kruskal-Wallis $H = 12.419$	$P = 0.020$	Minus	0.5205 ± 0.0538
			Nil	0.5629 ± 0.0084
			Plus	0.7467 ± 0.0458
Amino acids	Kruskal-Wallis $H = 51.7902$	$P = 0.000$	Ala (A)	0.6314 ± 0.0300
			Arg (R)	0.6842 ± 0.0664
			Asn (N)	0.5457 ± 0.0391
			Asp (D)	0.4891 ± 0.0668
			Cys (C)	0.6176 ± 0.1008
			Gln (Q)	0.5740 ± 0.0951
			Glu (E)	0.5740 ± 0.0913
			Gly (G)	0.5766 ± 0.0386
			His (H)	0.4666 ± 0.0544
			Ile (I)	0.5039 ± 0.0239
			Leu (L)	0.5749 ± 0.0165
			Lys (K)	0.8076 ± 0.0626
			Met (M)	0.7232 ± 0.0568
			Phe (F)	0.5420 ± 0.0305
			Pro (P)	0.5621 ± 0.0356
			Ser (S)	0.5988 ± 0.0291
			Thr (T)	0.5748 ± 0.0297
			Trp (W)	0.6111 ± 0.0937
			Tyr (Y)	0.3869 ± 0.0496
			Val (V)	0.5671 ± 0.0373
Chemical properties	Kruskal-Wallis $H = 14.4586$	$P = 0.043$	Acidic	0.5205 ± 0.0538
			Aliphatic	0.5662 ± 0.0112
			Amide	0.5502 ± 0.0362
			Aromatic	0.5089 ± 0.0252
			Basic	0.6085 ± 0.0373
			Hydroxyl	0.5871 ± 0.0208
			Imino	0.5621 ± 0.0356
			Sulphur	0.6986 ± 0.0494
Hydrophobicity	Mann-Whitney $U = 0.554$	$P = 0.456$	Philic	0.5518 ± 0.0158
			Phobic	0.5667 ± 0.0103

made historical signal easier to detect in plant studies; (3) different clades of plants have different constraints that, when pooled and averaged over the entire data set, make them appear homogeneous.

Taxon density is known to affect both tree building and site informativeness (Graybeal, 1998; Hillis et al., 1996; Hillis, 1996, 1998), and Naylor and Brown's results could have been biased by limited sampling of highly divergent taxa. In an effort to parallel the taxon density and divergences explored in Naylor and Brown's study, we sampled 19 taxa for the plant phylogeny (i.e., the same number of taxa in the animal tree) from divergent clades. We randomly chose these 19 taxa in 100 replicates (using a program written for this purpose, RandomTaxa, available at http://www.tcd.ie/Botany/NS/software.html). For the 19-taxon plant data subsets,

TABLE 3. Proportion of significant *P*-values in lineage-specific and random replicates of 19 taxa (Kruskal-Wallis tests; lineage-specific n = 20; random n = 100; see text for details).

	Lineage-specific				Random			
	Codon positions		Amino acids		Codon positions		Amino acids	
	RI	CI	RI	CI	RI	CI	RI	CI
$P < 0.05$	20%	40%	15%	30%	17%	all	5%	35%
$P < 0.01$	10%	20%	none	10%	4%	99%	1%	17%
$P < 0.001$	5%	15%	none	none	none	none	none	5%

we computed the RI for codon positions and amino acids (triplets), both types of sites reflecting most of the evolutionary constraints that might have affected the plastid genes in question. We have also evaluated whether constraints in plastid DNA, if any, might have been lineage-specific but masked by our sampling of taxa. To do this we examined subsets of 19 taxa that formed monophyletic (or at least paraphyletic) groups; 20 subsets met these criteria (Appendix 1). We compared the results from these groups with those for similar sized groups in which taxa were randomly chosen. In addition to RI, we also recorded the CI for each of these sites in these two resampling procedures.

Kruskal-Wallis tests were performed in each case, and all *P*-values are listed in Appendix 2; for clarity, Table 3 presents the significant sets of *P*-values. RIs for codon positions were significantly different in approximately 20% of the comparisons, both when 19 taxa were chosen randomly and from closely related lineages. All comparisons for CI/codon positions were significant with randomly chosen subsets, whereas only 40% were siginificant when taxa formed mono- or paraphyletic groups: as expected, the latter simply indicates that homoplasy is higher in comparisons of distantly related taxa (parallelism and convergence are more likely to occur), especially at third codon positions due to degeneracy of the genetic code. Regarding the triplets coding for amino acids, only a few of the subsets exhibited significantly different RIs ($\leq$15%), and this was three times more frequent in lineage-specific comparisons (15% versus 5% for $P < 0.05$). CIs were similar in random and lineage-specific comparisons with approximately 30% of the comparisons significant. We also compared the proportion of invariant sites and the alpha shape parameters of the gamma distribution between the random and the lineage-specific subsets. They were significantly different: the proportion of invariant sites was lower in distantly related taxa (0.4663 ± 0.0357 versus 0.5353 ± 0.0536, t-test $P < 0.0010$), whereas rate heterogeneity among sites was lower within closely related taxa (α 0.8544 ± 0.1489 versus 0.7738 ± 0.0952, t-test $P = 0.0296$). Therefore, although we cannot exclude some type of differential lineage-specific constraints (especially in amino acids in which more replicates exhibited significant differences in mono/paraphyletic subsets; Table 3), it is unlikely that the homogeneity seen in the complete plant data set is due to either averaged differences among various subsets (Table 3 and Appendix 2) or improved hierarchical structure only. It is noteworthy that if the minimum pair-wise divergence of the random plant subsets was approximately twice that of the entire plant data set, its maximum was only 1/3 the average divergence observed for the animal data set (Table 1). Therefore, the lower sequence divergence in *atpB* and *rbcL* sequences (hypothesis 1) is likely to be important in explaining different observations on the substitution properties of plastid and animal mitochondrial sequences (Fig. 1).

TAXON SAMPLING AND CODON POSITION INFORMATIVENESS

We further examined informativeness of codon positions, an issue receiving by far the most attention in *a priori* weighting schemes: third positions are usually down weighted (or eliminated altogether by translation to amino acid sequences) because of their expected high level of homoplasy (e.g., Allard and Carpenter, 1996; Waddell et al., 1999). Using simulations of mitochondrial coding genomes (total 12,234 base pairs), we increased the number of taxa in the animal tree from 19 to 600 (see Fig. 2). We used a HKY85 + Γ model of DNA evolution that accounts for different base frequencies,

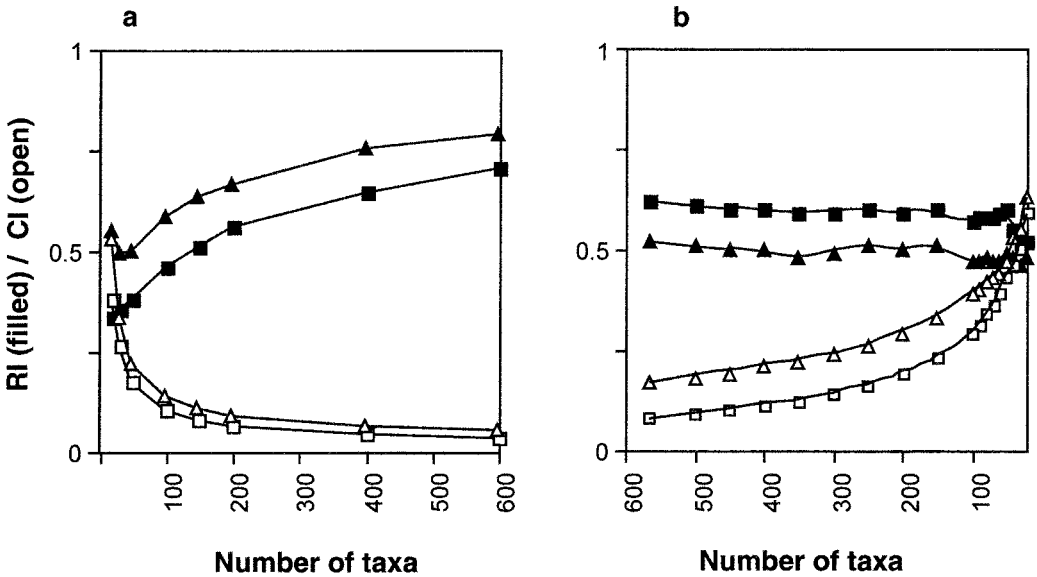

FIGURE 2. Relationship among codon position, number of taxa, and phylogenetic signal. RIs are shown with filled figures, whereas CIs are shown with open figures. (a) Using the animal mitochondrial phylogeny from Naylor and Brown (1998), we simulated first and third codon positions (respectively triangles and squares) for an additional 600 taxa. (b) When plant taxa were deleted randomly in the angiosperm phylogeny from Soltis et al. (2000), the RI remained roughly constant, and third codon positions performed the best throughout.

transition/transversion bias, and rate heterogeneity among sites (Hasegawa et al., 1985; Yang, 1996) with all parameters inferred from the real data but separately for first and third codon positions (parameters first positions: A = 0.3262, C = 0.1736, G = 0.1716, T = 0.3284, $\kappa = 2.3634$, % invariant sites = 0.1133, $\alpha = 0.8959$; parameters third positions: A = 0.3813, C = 0.2053, G = 0.0854, T = 0.3278, $\kappa = 6.1542$; % invariant sites = 0.0001, $\alpha = 1.6093$). Sequences were simulated using Seq-Gen (Rambault and Grassly, 1997; available at http://evolve.zoo.ox.ac.uk/software/Seq-Gen/Seq-Gen.html). In the growing tree, each new taxon was joined at the midpoint of the longest branches, but the length of these new branches was set such that the total tree length did not change (Graybeal, 1998); this was performed using the program BranchCut (available at http://www.tcd.ie/Botany/NS/software.html). For comparison, we randomly decreased the number of plant taxa from 567 to 20 using the tree from Soltis et al. (2000; Fig. 2). Plastid third-codon positions always exhibited the greatest informativeness (Fig. 2, see also Chase et al., 1995; Källersjö et al., 1998; Olmstead et al., 1998), whereas in the animal mitochondrial genome these positions performed most poorly (Naylor and Brown, 1997, 1998). The rate of change of the third codon position in the plastid genome (data set from Soltis et al., 2000) is roughly equivalent to that of the first codon positions in animal mitochondrial DNA (Naylor and Brown, 1998). The common view that saturation, and consequently lack of phylogenetic utility of third codon positions is a source of confusion because this does not hold for plastid DNA under these circumstances.

CONCLUSIONS

In plants, phylogenetic analyses of large DNA data sets have been shown to be more tractable than would have been predicted (Hillis, 1996, 1998; Källersjö et al., 1998; Savolainen et al., 2000a, 2000b; Soltis et al., 1998, 1999, 2000). Molecular phylogenetic studies in angiosperms involving multiple genes are among the largest ever conducted, and they have produced highly congruent results (Hoot et al., 1999; Qiu et al., 1999, 2000; Savolainen et al., 2000a, 2000b; Soltis et al., 1999, 2000). Large phylogenetic analyses of plastid genes have also been conducted involving all vascular plants (Pryer et al., 2001) and green plants (Källersjö et al., 1998; Karol et al., 2001).

One suggestion for the reasons behind this success has been that by adding numerous taxa, "noise" has been dispersed, leading to more accurate reconstruction (Hillis, 1996, 1998; Purvis and Quicke, 1997). However, true "noise" (randomness) cannot be dispersed simply by adding taxa. This is only expected if sampling helps to emphasize the underlying hierarchical nature of a phylogeny. This may not be the case, either because the underlying phylogeny is not bifurcating or because some lineages in the phylogenetic analysis are poorly sampled. Here, we report another reason why building plant phylogenies with some plastid genes has turned out to a more straightforward task: low substitution rates and spatial substitution patterns not limited by detectable functional constraints have lead to high signal content.

If correct, this means that botanists have made enormous strides in inferring seed plant phylogeny from *atpB* and *rbcL* sequences due simply to the substitution properties of these genes. Unfortunately, our findings do not mean that all plastid sequences will necessarily share these properties in seed plants. Further, at deeper levels, *atpB* and *rbcL* sequences may also show evidence of lineage-specific constraints, as might be expected given the observations reported for other plastid genes in anciently diverged taxa (Lockhart et al., 1999, 2000). The evolution of *rbcL* for example is clearly under structural and functional constraint (Kellogg and Juliano, 1997) and at deeper levels it may well be important to determine whether differential structural constraints are likely to influence phylogeny reconstruction. Some specific questions, such as the rooting of angiosperms or deep branching pattern in embryophytes, may particularly benefit from examining in more detail homoplasy in various categories of sites (Qiu et al., 2000; Barkman et al., 2000; Sanderson et al., 2000; Zanis et al., 2002).

By contrast, systematic zoology will greatly benefit from an increased understanding of the pattern exhibited by the rapidly evolving mitochondrial genome and potentially strong functional constraints inherent in this small genome (Brown et al., 1979; Brown, 1983, 1985; Naylor and Brown, 1997, 1998). Pollock et al. (2000) also examined 69 vertebrate mitochondrial genomes and made a strong case in favor of collecting genome-scale data to examine functional and structural constraints and to improve phylogenetic inference using large data sets.

ACKNOWLEDGMENTS

We thank Chris Simon, Pete Lockhart, and an anonymous reviewer for their comments on the manuscript. This study was partly funded by the Swiss National Science Foundation (grant 3152378.97 to V.S.), the Roche Research Foundation, the U.S. National Science Foundation (DEB-9707868 to PSS and DES), the Irish Higher Education Authority, and a joint U.S.–U.K. Fulbright Distinguished Professorship to D.E.S. and P.S.S.

REFERENCES

ALLARD, M. W., AND J. M. CARPENTER. 1996. On weighting and congruence. Cladistics 12:183–198.

BARKMAN, T. J., G. CHENERY, J. R. MCNEAL, J. LYONS-WEILER, W. J. ELLISENS, M. MOORE, A. D. WOLFE, AND C. W. DEPAMPHILIS. 2000. Independent and combined analyses of sequences from all three genome compartments converge on the root of flowering plant phylogeny. Proc. Natl Acad. Sci. USA 97:13166–13171.

BENTON, M. J. 1993. The fossil record 2. Chapman and Hall, London.

BROWN, W. M. 1983. Evolution of animal mitochondrial DNA. Pages 62–88 *in* Evolution of genes and proteins (M. Nei and R. K. Koehn, eds.). Sinauer Associates, Sunderland, Massachusetts.

BROWN, W. M. 1985. The mitochondrial genome of animals. Pages 95–130 *in* Molecular evolutionary genetics (R. J. MacIntyre, ed.). Plenum, New York.

BROWN, W. M., M. GEORGE JR., AND A. C. WILSON. 1979. Rapid evolution of animal mitochondrial DNA. Proc. Natl Acad. Sci. USA 76:1967–1971.

CHASE, M. W., D. E. SOLTIS, R. G. OLMSTEAD, D. MORGAN, D. H. LES, B. D. MISHLER, M. R. DUVALL, R. A. PRICE, H. G. HILLS, Y.-L. QIU, K. A. KRON, J. H. RETTIG, E. CONTI, J. D. PALMER, J. R. MANHART, K. J. SYTSMA, H. J. MICHAELS, W. J. KRESS, K. G. KAROL, W. D. CLARK, M. HEDRÉN, B. S. GAUT, R. K. JANSEN, K.-J. KIM, C. F. WIMPEE, J. F. SMITH, G. R. FURNIER, S. H. STRAUSS, Q.-Y. XIANG, G. M. PLUNKETT, P. S. SOLTIS, S. M. SWENSEN, S. E. WILLIAMS, P. A. GADEK, C. J. QUINN, L. E. EGUIARTE, E. GOLENBERG, G. H. LEARN, JR., S. W. GRAHAM, S. C. BARRETT, S. DAYANANDAN, AND V. A. ALBERT. 1993. Phylogenetics of seed plants: An analysis of nucleotide sequences from the plastid gene *rbcL*. Ann. Missouri Bot. Gard. 80:528–580.

CHASE, M. W., M. R. DUVALL, H. G. HILLS, J. G. CONRAN, A. V. COX, L. E. EGUIARTE, J. HARTWELL, M. F. FAY, L. R. CADDICK, K. M. CAMERON, AND S. B. HOOT. 1995. Molecular phylogenetics of Lilianae. Pages 109–137 *in* Monocotyledons: Systematics and evolution (P. Rudall, P. J. Cribb, D. F. Cutler, and C. J. Humphries, eds.). Royal Botanic Gardens, Kew, London.

CHASE, M. W., AND A. V. COX. 1998. Gene sequences, collaboration, and analysis of large data sets. Austr. Syst. Bot. 11:215–229.

FARRIS, J. S. 1989. The retention index and the rescaled consistency index. Cladistics 5:417–419.

GRAYBEAL, A. 1998. Is it better to add taxa or characters to a difficult phylogenetic problem? Syst. Biol. 47: 9–17.

HASEGAWA, M., H. KISHINO, AND T. YANO. 1985. Dating the human-ape splitting by a molecular clock of mitochondrial DNA. J. Mol. Evol. 21:160–174.

HILLIS, D., K. MABLE, AND C. MORITZ. 1996. Applications of molecular systematics: The state of the field and a look to the future. Pages 515–543 *in* Molecular systematics (D. Hillis, C. Moritz, and B. Mable, eds.). Sinauer Associates, Sunderland, Massachusetts.

HILLIS, D. M. 1996. Inferring complex phylogenies. Nature 383:130–131.

HILLIS, D. M. 1998. Taxonomic sampling, phylogenetic accuracy, and investigator bias. Syst. Biol. 47:3–8.

HOOT, S. B., S. MAGALLÓN, AND P. R. CRANE. 1999. Phylogeny of basal eudicots based on three molecular data sets: *atpB*, *rbcL*, and 18S nuclear ribosomal DNA sequences. Ann. Missouri Bot. Gard. 86:1–32.

KÄLLERSJÖ, M., J. S. FARRIS, M. W. CHASE, B. BREMER, M. F. FAY, C. J. HUMPHRIES, G. PETERSEN, O. SEBERG, AND K. BREMER. 1998. Simultaneous parsimony jackknife analysis of 2538 *rbcL* DNA sequences reveals support for major clades of green plants, land plants, seed plants, and flowering plants. Pl. Syst. Evol. 213:259–287.

KAROL, K. G., R. G. MCCOURT, M. T. CIMINO, AND C. H. DELWICHE. 2001. The closest living relatives of land plants. Science 294:2351–2353.

KELLOGG, E. A., AND N. D. JULIANO. 1997. The structure and function of RuBisCO and their implications for systematic studies. Amer. J. Bot. 84:413–428.

KINGS, M., A. STONE, R. W. SCHMITZ, H. KRAINITZKI, M. STONEKING, AND S. PÄÄBO. 1997. Neandertal DNA sequences and the origin of modern humans. Cell 90:32–42.

KLUGE, A. G., AND J. S. FARRIS. 1969. Quantitative phyletics and the evolution of anurans. Syst. Zool. 18:1–32.

LOCKHART, P. J., C. J. HOWE, A. C. BARBROOK, A. W. D. LARKUM, AND D. PENNY. 1999. Spectral analysis, systematic bias, and the evolution of chloroplast. Mol. Biol. Evol. 16:573–576.

LOCKHART, P. J., D. HUSON, U. MAIER, M. J. FRAUNHOLZ, Y. VAN DE PEER, A. C. BARBROOK, C. J. HOWE, AND M. A. STEEL. 2000. How molecules evolve in eubacteria. Mol. Biol. Evol. 17:835–838.

NAYLOR, G. J. P., AND W. M. BROWN. 1997. Structural biology and phylogenetic estimation. Nature 388:527–528.

NAYLOR, G. J. P., AND W. M. BROWN. 1998. Amphioxus mitochondrial DNA, chordate phylogeny, and the limits of inference based on comparisons of sequences. Syst. Biol. 47:61–76.

NAYLOR, G. J. P., T. M. COLLINS, AND M. B. WESLEY. 1995. Hydrophobicity and phylogeny. Nature 373:565–566.

OLMSTEAD, R. G., P. A. REEVES, AND A. C. YEN. 1998. Patterns of sequence evolution and implication for parsimony analysis of chloroplast DNA. Pages 164–187 *in* Molecular systematics of plants II (D. E. Soltis, P. S. Soltis, and J. J. Doyle, eds.). Kluwer Academic, Norwell, Massachusetts.

OVCHINNIKOV, I. V., A. GÖTHERSTRÖM, G. P. ROMANOVA, V. M. KHARITONOV, K. LIDEN, AND W. GOODWIN. 2000. Molecular analysis of Neanderthal DNA from the northern Caucasus. Nature 404:490–493.

PAGEL, M. 1999. Inferring historical patterns of biological evolution. Nature 401:877–884.

PALMER, J. D., AND C. R. DELWICHE. 1998. The origin and evolution of plastids and their genomes. Pages 375–409 *in* Molecular systematics of plants II (D. E. Soltis, P. S. Soltis, and J. J. Doyle, eds.). Kluwer Academic, Norwell, Massachusetts.

PENNY, D., M. HASEGAWA, P. J. WADDELL, AND M. D. HENDY. 1999. Mammalian evolution: Timing and implications from using the LogDeterminant transform for proteins of differing amino acid composition. Syst. Biol. 48:76–93.

POLLOCK, D. D., J. A. EISEN, N. A. DOGGETT, AND M. P. CUMMINGS. 2000. A case for evolutionary genomics and the comprehensive examination of sequence biodiversity. Mol. Biol. Evol. 17:1776–1788.

PRYER, K. M., H. SCHNEIDER, A. M. SMITH, R. CRANFILL, P. G. WOLF, J. S. HUNT, AND S. D. SIPES. 2001. Horsetails and ferns are a monophyletic group and the closest living relatives to the seed plants. Nature 409:545–648.

PURVIS, A., AND D. L. J. QUICKE. 1997. Building phylogenies—Are the big easy? Trends Ecol. Evol. 12:49–50.

QIU, Y.-L., J. LEE, F. BERNASCONI-QUADRONI, D. E. SOLTIS, P. S. SOLTIS, M. ZANIS, Z. CHEN, V. SAVOLAINEN, AND M. W. CHASE. 1999. The earliest angiosperms: Evidence from mitochondrial, plastid, and nuclear genomes. Nature 402:404–407.

QIU, Y.-L., J. LEE, F. BERNASCONI-QUADRONI, D. E. SOLTIS, P. S. SOLTIS, M. ZANIS, Z. CHEN, V. SAVOLAINEN, AND M. W. CHASE. 2000. Phylogeny of basal angiosperms: Analysis of five genes from three genomes. Int. J. Pl. Sci. 161:S3–S27.

RAMBAUT, A., AND N. C. GRASSLY. 1997. Seq-Gen: An application for the Monte Carlo simulation of DNA sequence evolution along phylogenetic trees. Comput. Applic. Biosci. 13:235–238.

SANDERSON, M., M. F. WOJCIECHOWSKI, J.-M. HU, T. S. KHAN, AND S. G. BRADY. 2000. Error, bias, and long-branch attraction in data for two chloroplast photosystem genes in seed plants. Mol. Biol. Evol. 17:782–797.

SAVOLAINEN, V., M. W. CHASE, C. M. MORTON, S. B. HOOT, D. E. SOLTIS, C. BAYER, M. F. FAY, A. DE BRUIJN, S. SULLIVAN, AND Y.-L. QIU. 2000a. Phylogenetics of flowering plants based upon a combined analysis of plastid *atpB* and *rbcL* gene sequences. Syst. Biol. 49:306–362.

SAVOLAINEN, V., M. F. FAY, D. C. ALBACH, A. BACKLUND, M. VAN DER BANK, K. M. CAMERON, S. A. JOHNSON, M. D. LLEDÓ, J.-C. PINTAUD, M. POWELL, M. C. SHEAHAN, D. E. SOLTIS, P. S. SOLTIS, P. WESTON, W. M. WHITTEN, K. J. WURDACK, AND M. W. CHASE. 2000b. Phylogeny of the eudicots: A nearly complete familial analysis based on *rbcL* gene sequences. Kew Bull. 55:257–309.

SOLTIS, D. E., C. HIBSCH-JETTER, P. S. SOLTIS, M. W. CHASE, AND J. S. FARRIS. 1997. Molecular phylogenetic relationships among angiosperms: An overview based on *rbcL* and 18S rDNA sequences. Pages 157–178 *in* Evolution and diversification of land plants (K. Iwatsuki and P. H. Raven, eds.). Springer Verlag, Tokyo.

SOLTIS, D. E., AND P. S. SOLTIS. 1998. Choosing an approach and an appropriate gene for phylogenetic analysis. Pages 1–42 *in* Molecular systematics of plants II (D. E. Soltis, P. S. Soltis, and J. J. Doyle, eds.). Kluwer Academic, Dordrecht.

SOLTIS, D. E., P. S. SOLTIS, M. E. MORT, M. W. CHASE, V. SAVOLAINEN, S. B. HOOT, AND C. M. MORTON. 1998. Inferring complex phylogenies using parsimony: An

empirical approach using three large DNA data sets for angiosperms. Syst. Biol. 47:32–42.

SOLTIS, D. E., P. S. SOLTIS, M. W. CHASE, M. E. MORT, D. C. ALBACH, M. ZANIS, V. SAVOLAINEN, W. H. HAHN, S. B. HOOT, M. F. FAY, M. AXTELL, S. M. SWENSEN, K. C. NIXON, AND J. S. FARRIS. 2000. Angiosperm phylogeny inferred from a combined data set of 18S rDNA, *rbcL*, and *atpB* sequences. Bot. J. Linn. Soc. 133:381–461.

SOLTIS, P. S., D. E. SOLTIS, AND M. W. CHASE. 1999. Angiosperm phylogeny inferred from multiple genes as a tool for comparative biology. Nature 402:402–404.

SOLTIS, P. S., D. E. SOLTIS, V. SAVOLAINEN, P. R. CRANE, AND T. G. BARRACLOUGH. 2002. Rate heterogeneity among lineages of tracheophytes: Integration of molecular and fossil data and evidence for molecular living fossils. Proc. Natl Acad. Sci. USA. 99:4430–4435.

WADDELL, P. J., N. OKADA, AND M. HASEGAWA. 1999. Towards resolving the interordinal relationships of placental mammals. Syst. Biol. 48:1–5.

WIKSTRÖM, N., V. SAVOLAINEN, AND M. W. CHASE. 2001. Evolution of the angiosperms: Calibrating the family tree. Proc. Roy. Soc. B 268:2211–2220.

YANG, Z. H. 1996. Among-site rate variation and its impact on phylogenetic analyses. Trends Ecol. Evol. 11:367–372.

ZANIS, M. J., D. E. SOLTIS, P. S. SOLTIS, S. MATHEWS, AND M. J. DONOGHUE. 2002. The root of the angiosperms revisited. Proc. Natl Acad. Sci. USA. 99:6848–6853.

ZAR, J. H. 1999. Biostatistical analysis. Prentice-Hall, Inc., New Jersey.

First submitted 8 September 2001; reviews returned 13 January 2002; final acceptance 10 May 2002
Associate Editor: Peter Lockhart

APPENDIX 1. LIST OF 20 SUBSETS OF TAXA FORMING MONO/PARAPHYLETIC GROUPS AND USED IN THE LINEAGE-SPECIFIC COMPARISONS

1: *Abelia, Dipsacus, Scabiosa, Symphoricaricarpos, Valeriana, Sambucus, Viburnum, Eremosyne, Escallonia, Apium, Delarbrea, Hedera, Panax, Pittosporum, Sollya, Melanophylla, Griselinia, Berzelia, Helwingia*; 2: *Alseuosmia, Corokia, Phelline, Barnadesia, Gerbera, Helianthus, Tagetes, Tragopogon, Boopis, Scaevola, Menyanthes, Nymphoides, Donatia, Stylidium, Campanula, Codonopsis, Lobelia, Roussea, Gonocaryum*; 3: *Convolvulus, Ipomoea, Duckeodendron, Lycopersicon, Nolana, Nicotiana, Petunia, Schizanthus, Hydrolea, Montinia, Apocynum, Nerium, Gelsemium, Exacum, Spigelia, Cephalanthus, Rogiera, Mitchella, Pentas*; 4: *Idria, Impatiens, Marcgravia, Pellicera, Tetramerista, Roridula, Alangium, Cornus, Camptotheca, Nyssa, Caiophora, Eucnide, Petalonyx, Carpenteria, Philadelphus, Decumaria, Hydrangea, Hydrostachys, Fendlera*; 5: *Ancistrocladus, Triphyophyllum, Nepenthes, Cocoloba, Polygonum, Limonium, Plumbago, Drosera, Frankenia, Tamarix, Asteropeia, Bougainvillaa, Mirabilis, Delosperma, Ercilla, Phytolacca, Mollugo, Pereskia, Portulaca*; 6: *Calathealoes, Maranta, Costus, Dimerocostus, Monocostus, Tapeinochilos, Canna, Globba, Hedychium, Zingiber, Riedelia, Orchidantha, Phenakospermum, Strelitzia, Ravenala, Heliconia, Ensete, Musella, Musa*; 7: *Agave, Behnia, Chlorophytum, Allium, Asparagus, Convallaria, Liriope, Ruscus, Smilacina, Nolina, Luzuriaga, Bowiea, Scilla, Muilla, Clivia, Hippeastrum, Bulbine, Xanthorrhoea, Xeronema*; 8: *Tricyrtis, Smilax, Barbacenia, Stemona, Cyclanthus, Sphaeradenia, Freycinetia, Japonolirion, Petrosavia, Burmannia, Dioscorea, Tacca, Aponogeton, Vallisneria, Zostera, Pleea, Tofieldia, Gymnostachys, Spathiphyllum*; 9: *Annona, Asimina, Eupomatia, Degeneria, Galbulimima, Liriodendron, Magnolia, Manglietia, Knema, Myristica, Belliolium, Drimys, Tasmania, Takhtajania, Canella, Cinnamodenron, Calycanthus, Idiospermum, Cinnamomum*; 10: *Aristolochia, Lactoris, Asarum, Saruma, Houttuyni, Saururus, Peperomia, Piper, Chloranthus, Sarcandra, Hedyosmum, Barclaya, Brasenia, Cabomba, Nuphar, Nymphaea, Austrobaileya, Illicium, Schisandra*; 11: *Akebia, Sinofranchetia, Decaisnea, Sargentodoxa, Circaeaster, Kingdonia, Caulophyllum, Nandina, Coptis, Xanthorhiza, Ranunculus, Glaucidium, Hydrastis, Menispermum, Tinospora, Euptelea, Dicentra, Hypecoum, Pteridophyllum*; 12: *Altingia, Liquidambar, Cercidiphyllum, Crassula, Dudleya, Sedum, Kalenchoe, Haloragis, Myriophyllum, Penthorum, Tetracarpaea, Heuchera, Peltboykinia, Saxifraga_1, Saxifraga_2, Sullivantia, Itea, Pterostemon, Ribes*; 13: *Acer, Aesculus, Cupaniopsis, Koelreuteria, Xanthoceras, Ailanthus, Swietenia, Trichilia, Citrus, Poncirus, Cneorum, Bursera, Schinus, Anisoptera, Sarcolaena, Helianthemum, Muntingia, Aquilaria, Thymelea*; 14: *Akania, Bretschneidera, Tropaeolum, Batis, Koeberlinia, Brassica, Cleome, Capparis, Reseda, Floerkea, Limnanthes, Setchellanthus, Carica, Moringa, Tapiscia, Aphloia, Ixerba, Crossosoma, Stachyurus*; 15: *Albizzia, Pisum, Bauhinia, Polygala, Securidaca, Stylobasium, Alnus, Betula, Casuarina, Carya, Juglans, Myrica, Chrysolepis, Quercus, Fagus, Begonia, Tetrameles, Datisca, Coriaria*; 16: *Cucurbita, Barbeya, Elaeagnus, Shepherdia, Ceanothus, Trevoa, Rhamnus, Boehmeria, Pilea, Celtis, Humulus, Trema, Ficus, Morus, Zelkova, Kerria, Photinia, Prunus, Spiraea*; 17: *Averrhoa, Oxalis, Bauera, Davidsonia, Eucryphia, Crinodendron, Elaeocarpus, Platytheca, Sloanea, Euphorbia, Linum, Reinwardtia, Androstachys, Stachystemon, Balanops, Chrysobalanus, Licania, Dichapetalum, Trigonia*; 18: *Dicella, Malpighia, Galphimia, Hypericum, Marathrum, Mesua, Drypetes, Humiria, Bruguiera, Carallia, Erythroxylon, Caryocar, Irvingia, Medusagyne, Ochna, Quiina, Acharia, Kigellaria, Pangium*; 19: *Barleria, Justicia, Thunbergia, Sesamum, Callicarpa, Clerodendron, Lamium, Pedicularis, Phyla, Paulownia, Euthystachys, Byblis, Proboscidea, Buddleja, Campsis, Veronica, Catalpa, Lantana, Verbena*; 20: *Actinidia, Arbutus, Arctostaphylos, Clethra, Cyrilla, Sarracenia, Camellia, Manilkara, Galax, Halesia, Styrax, Symplococus, Eurya, Ternstroemia, Diospyros, Anagallis, Ardisia, Androsace, Clavija*.

APPENDIX 2. LIST OF INDIVIDUAL *P*-VALUES (SEE TEXT FOR DETAILS).

Codon positions		Amino acids	
RI	CI	RI	CI
Lineage-specific			
0.8895	0.5156	0.9927	0.1482
0.0533	0.0328	0.4522	0.3598
0.8784	0.5628	0.1915	0.0182
0.6057	0.2503	0.4448	0.1751

APPENDIX 2. Continued.

Codon positions		Amino acids	
RI	CI	RI	CI
0.0664	0.3308	0.7789	0.1131
0.0003	0.2796	0.0364	0.0018
0.0188	0.0945	0.0645	0.2194
0.1186	0.9407	0.1295	0.1483
0.0672	0.0819	0.7899	0.5350
0.0291	0.0000	0.5049	0.5812
0.9034	0.1389	0.0210	0.0590
0.2992	0.0808	0.2478	0.2659
0.3835	0.0035	0.5719	0.0123
0.4026	0.0392	0.0835	0.0345
0.0059	0.0309	0.0194	0.0019
0.2657	0.0123	0.0560	0.1575
0.2998	0.3775	0.7613	0.2766
0.5621	0.1574	0.3772	0.0104
0.0546	0.0000	0.6328	0.2497
0.8386	0.0000	0.3287	0.7853
		Random	
0.1573	0.0002	0.4855	0.0172
0.3287	0.0000	0.9168	0.1930
0.7339	0.0000	0.9926	0.8393
0.5430	0.0000	0.1093	0.1028
0.1172	0.0000	0.7945	0.0945
0.9752	0.0000	0.3469	0.4620
0.0117	0.0000	0.3153	0.0314
0.2573	0.0000	0.8754	0.0725
0.4516	0.0000	0.3613	0.5121
0.7811	0.0000	0.3253	0.0033
0.2178	0.0010	0.9671	0.4322
0.1679	0.0000	0.4785	0.0611
0.1836	0.0000	0.6795	0.0767
0.0690	0.0000	0.5320	0.0099
0.1681	0.0000	0.5369	0.2989
0.9214	0.0000	0.5130	0.0446
0.7144	0.0000	0.4060	0.1196
0.2998	0.0000	0.5996	0.0641
0.1086	0.0000	0.7331	0.0011
0.0811	0.0000	0.8955	0.5736
0.4807	0.0000	0.5832	0.0109
0.4806	0.0000	0.3749	0.4247
0.0049	0.0001	0.0757	0.0626
0.0053	0.0000	0.2039	0.0065
0.4922	0.0000	0.2954	0.1935
0.0860	0.0000	0.5571	0.4173
0.0417	0.0000	0.2928	0.2782
0.1779	0.0000	0.6539	0.0698
0.3288	0.0000	0.1097	0.0007
0.0055	0.0000	0.4076	0.0043
0.4365	0.0000	0.9842	0.0205
0.4296	0.0000	0.5932	0.1763
0.0240	0.0000	0.6456	0.0180
0.0293	0.0000	0.5988	0.2140
0.0475	0.0000	0.4421	0.0103
0.9862	0.0000	0.9663	0.0308
0.7338	0.0000	0.3466	0.1565
0.2685	0.0000	0.7644	0.2751
0.0610	0.0000	0.9398	0.4380
0.1901	0.0000	0.3990	0.0957
0.0133	0.0000	0.6030	0.0090

APPENDIX 2. Continued.

Codon positions		Amino acids	
RI	CI	RI	CI
0.6568	0.0000	0.7327	0.0819
0.6451	0.0000	0.0379	0.4188
0.9320	0.0000	0.8791	0.0002
0.9726	0.0000	0.0246	0.0212
0.2324	0.0000	0.9359	0.1287
0.1237	0.0000	0.5834	0.5753
0.5633	0.0000	0.7373	0.2453
0.2525	0.0000	0.8437	0.2227
0.3493	0.0000	0.2549	0.2817
0.7559	0.0000	0.2890	0.4971
0.1648	0.0000	0.1861	0.1576
0.2368	0.0000	0.5876	0.0344
0.5907	0.0000	0.9320	0.4358
0.3562	0.0000	0.4047	0.0030
0.1066	0.0000	0.3403	0.0436
0.0228	0.0000	0.3931	0.0081
0.4623	0.0000	0.1185	0.2139
0.3699	0.0000	0.6514	0.0701
0.0492	0.0002	0.8597	0.0727
0.5366	0.0000	0.5700	0.5507
0.0906	0.0000	0.1950	0.0043
0.0443	0.0000	0.0127	0.0128
0.5097	0.0000	0.3982	0.2741
0.2598	0.0000	0.0962	0.0002
0.5070	0.0000	0.6535	0.1063
0.4218	0.0000	0.1916	0.0631
0.0429	0.0000	0.6893	0.0915
0.8593	0.0000	0.0846	0.0295
0.8395	0.0000	0.0595	0.0013
0.4488	0.0000	0.0991	0.2067
0.1102	0.0000	0.0412	0.0004
0.0046	0.0000	0.7257	0.2083
0.7284	0.0000	0.0600	0.0599
0.5621	0.0000	0.1914	0.2051
0.1624	0.0001	0.2019	0.1827
0.2747	0.0000	0.2516	0.0127
0.1256	0.0000	0.0075	0.0120
0.6734	0.0000	0.1146	0.0831
0.2240	0.0000	0.1130	0.0278
0.0896	0.0000	0.1804	0.1393
0.2415	0.0000	0.2734	0.8571
0.1022	0.0000	0.7338	0.0744
0.0201	0.0000	0.0613	0.0148
0.2872	0.0000	0.6594	0.0618
0.2195	0.0000	0.7082	0.6737
0.1970	0.0000	0.9113	0.0742
0.0115	0.0000	0.1068	0.2433
0.7533	0.0000	0.0628	0.0379
0.8881	0.0000	0.5652	0.2606
0.4010	0.0000	0.6809	0.2186
0.0524	0.0000	0.3002	0.0069
0.1399	0.0000	0.5368	0.0015
0.0139	0.0000	0.0943	0.0611
0.7238	0.0000	0.8598	0.0512
0.4529	0.0000	0.3967	0.0007
0.4975	0.0000	0.8418	0.1475
0.1412	0.0000	0.4731	0.1076
0.5518	0.0000	0.6534	0.3942
0.2995	0.0000	0.8073	0.0520