REPORT

Detecting evolutionary rate heterogeneity among mangroves and their close terrestrial relatives

Yang Zhong,[1]* Qiong Zhao,[2] Suhua Shi,[3] Yelin Huang[3] and Masami Hasegawa[4]
[1]*Ministry of Education Key Laboratory for Biodiversity Science and Ecological Engineering, and Institute of Biodiversity Science, Fudan University, Shanghai 200433, People's Republic of China.*
[2]*Department of Genetics, University of Wisconsin-Madison, WI 53706, USA.*
[3]*Key Laboratory of Gene Engineering of Ministry of Education, School of Life Sciences, Zhongshan University, Guangzhou 510275, People's Republic of China.*
[4]*The Institute of Statistical Mathematics, 4-6-7 Minami-Azabu, Minato-ku, Tokyo 106, Japan*
Correspondence: Institute of Biodiversity Science, Fudan University, Shanghai 200433, People's Republic of China. E-mail: yangzhong@fudan.edu.cn

Abstract

Mangroves form the dominant intertidal ecosystems and differ morphologically and physiologically from their close terrestrial relatives. We investigate the molecular evolutionary pattern of the typical mangrove family, i.e. Rhizophoraceae, and rate heterogeneity for the plastid *mat*K and *rbc*L genes in different species of the family, as revealed by phylogenetic analyses and relative-rate tests. Our study documents evolutionary rate heterogeneity in the Rhizophoraceae for the two genes: the mangrove genus *Bruguiera* has relatively slow substitution rates compared to the terrestrial genus *Carallia* at both synonymous and non-synonymous sites in the *mat*K sequences, and the synonymous and non-synonymous substitution matrices are correlated. However, the *rbc*L non-synonymous sites exhibit a high degree of rate heterogeneity among mangroves and related terrestrial groups, and uncoupling of rates with the synonymous sites. Selection is probably an important influence on the rate variation, suggesting further investigation for better understanding of various forces contributing to the rate heterogeneity and molecular adaptation in mangroves.

INTRODUCTION

Phylogenetic analyses have been a standard tool in systematic and evolutionary biology, with goals ranging from reconstructing evolutionary histories of organisms or genes (Li 1997) to measuring biodiversity (Purvis & Hector 2000), and have also become a comparative approach to ecological studies, e.g. exploring the phylogenetic structure of ecological communities (Webb 2000) and detecting adaptation to stress in various environments (Yang & Bielawski 2000; Schulte 2001). In particular, molecular phylogenetics has begun to attract more attention of ecologists to some difficult problems regarding adaptive evolution of plants and animals in different ecosystems, such as aquatic and terrestrial ecosystems.

Mangroves form the dominant intertidal ecosystems throughout the tropical regions of the world and enter the subtropics, with a total area of about 182 000 km^2 worldwide (Spalding *et al.* 1997). Current knowledge of the evolution of mangroves is still limited in terms of range-wide distributions and is mostly restricted to morphological and physiological analyses, which have indicated a high degree of homogeneity (Dodd *et al.* 1998). For example, mangroves differ from their terrestrial related plants in that the former are characterized by their peculiar adaptive viviparous fruits and their pneumatophores and knee roots,

etc. (Geh & Keng 1974). However, less attention has been paid to molecular evolution of mangroves and their terrestrial related groups (Zhong *et al.* 2000).

Recently, molecular phylogenetic analyses of the mangrove family Rhizophoraceae, which comprise 15 genera and approximately 130 species including eumangrove ("true" mangrove) and terrestrial plants, have been conducted using several kinds of DNA sequences (Huang *et al.* 1999; Setoguchi *et al.* 1999; Schwarzbach & Ricklefs 2000; Zhong *et al.* 2000; Shi *et al.* 2002). Here, we present detailed analyses of relative evolutionary rates of two chloroplast genes, *mat*K (maturase K) and *rbc*L (large subunit of the enzyme ribulose-1,5-bisphosphate carboxylase), among the mangroves and their close terrestrial relatives in Rhizophoraceae. The results have implications both for evolutionary

biology and ecology of mangroves and related groups, especially for better understanding of molecular adaptation in mangroves.

MATERIALS AND METHODS

The sequences of chloroplast DNA (cpDNA) *mat*K and *rbc*L genes used in this study are listed in Table 1. According to the viewpoint of the APG (Angiosperm Phylogeny Group 1998), the family Rhizophoraceae belongs to the order Malpighiales, so that *Byrsonima crassifolia* (Malpighiaceae) can be selected as an outgroup for our phylogenetic analyses.

Each sequence data set was aligned using the program CLUSTAL-X (Thompson *et al.* 1997). The phylogenetic analy-

Table 1 Sequences of *mat*K and *rbc*L genes used in this study

| Genus | Abbreviation | Species | Habitat and distribution[2] | GenBank Accession No. | |
				*mat*K	*rbc*L
Blepharistemma	BL	*B. membranifolia*	T; AS		AF006761
Bruguiera	BR	*B. cylindrica*	M; AS, AU		AF127694
		B. exaristata	M; AU		AF127695
		B. gymnorhiza	M; AS, AF, AU	AF105088	AF127693
		B. parviflora	M; AS, AU		AF127692
		B. sexangula	M; AS, AU	AF105091	AF127691
Carallia	CR	*C. brachiata*	T; AS	AF105086	AF006757
		C. graciniifolia	T; AS	AF126370	AF127373
		C. pectinifolia	T; AS	AF105087	AF127372
Cassipourea	CS	*C. ceylanica*	T; AF		AF127674
		C. guianensis	T; AM		AF127673
		C. elliptica	T; AM		AF127672
		C. rotundifolia	T; AF		AF006762
Ceriops	CE	*C. australis*	M; AS, AF, AU		AF127683
		C. tagal	M; AS, AF, AU	AF105089	AF127684
Crossostylis	CO	*C. biflora*	T; SPI		AF127679
		C. grandiflora	T; SPI		AF006760
Dactylopetalum	DA	*D. verticillatum*	T; AF		AF127676
Gynotroches	GY	*G. axillaris*	T; AS, SPI		AF127678
Kandelia	KA	*K. candel*	M; AS	AF105090	AF127682
Macarisia	MA	*M. emarginata*	T; AF		AF129130
Pellacalyx	PE	*P. axillaris*	T; AS		AF127681
		P. saccardians	T; AS		AF127680
		P. yunnanensis	T; AS	AF126371	
Rhizophora	RH	*R. apiculata*	M; AS, AF, AU		AF127685
		R. mangle	M; AM, AF		AF127689
		R. mucronata	M; AS, AF, AU		AF127687
		R. racemosa	M; AM		AF127690
		R. stylosa	M; AS	AF105092	AF127686
Sterigmapetalum	ST	*S. guianense*	T; AM		AF127671
Byrsonima[1]	BY	*B. crassifolia*	T; AM	Unpublished[3] L01892	

[1]Outgroup.
[2]M, mangrove; T, terrestrial; AS, Asia; AF, Africa; AM, America; AU, Australia; SPI, South Pacific Islands.
[3]Provided by Charles C. Davis at Harvard University.

ses of the nucleotide sequences of the *rbc*L and *mat*K genes were performed using the NucML program in the MOLPHY Package version 2.3 (Adachi & Hasegawa 1996) with the HKY 85 model (Hasegawa *et al.* 1985). Bootstrap proportions (BPs) were estimated by the RELL (Resampling of Estimated Log-Likelihoods) method (Kishino *et al.* 1990; Hasegawa & Kishino 1994) with 10 000 bootstrap resamplings.

The rates of synonymous and non-synonymous substitutions were estimated by using the method of Li (1993) and Pamilo & Bianchi (1993). The relative-rate tests of between the lineages at generic level in the phylogenetic trees were performed with the method proposed by Li & Bousquet 1992) and using the RRTree program (Robinson *et al.* 1998; Robinson-Rechavi & Huchon 2000). The closest sister group of each genus pair to be compared in a phylogenetic tree was selected as the reference group for a relative-rate test. The binary transferred matrices were constructed with the results of the relative-rate tests, and Mantel tests (Sokal & Rohlf 1995; Liedloff 1999) were conducted to detect the associations among these matrices. Significance values of the Mantel tests were based on 100 random iterations.
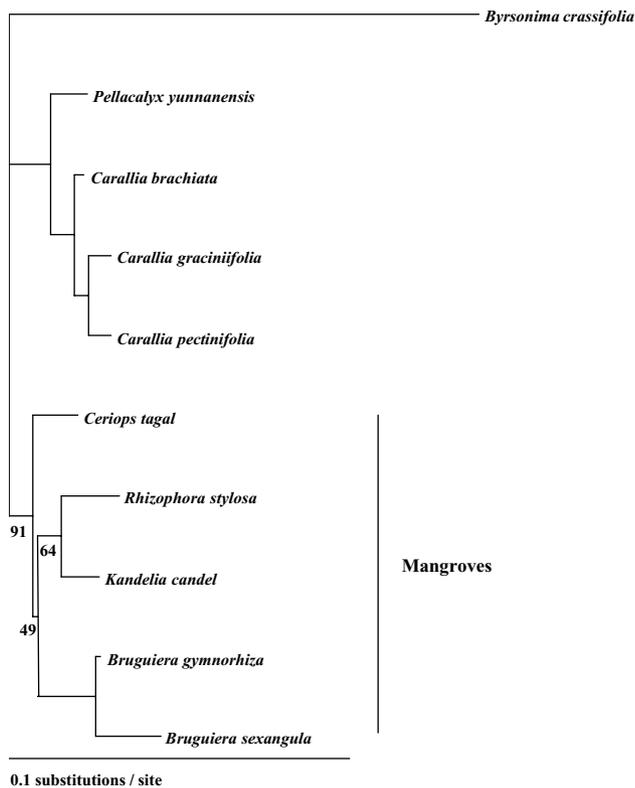
RESULTS AND DISCUSSION

The phylogenetic trees of the Rhizophoraceae based on the sequences of *mat*K and *rbc*L genes are shown in Figs 1 and 2, respectively. The *mat*K tree strongly supports the idea that the mangroves form a monophyletic group with a relatively high BP value (91%). The BP value for the monophyly of mangroves is 81% in the highest likelihood tree among 945 trees, and the BP value for the monophyly based on the total *rbc*L and *mat*K sequences is 97%. However, the terrestrial plants form different phylogenetic patterns: monophyly in the *mat*K tree (Fig. 1) but paraphyly in the *rbc*L tree (Fig. 2).

The results obtained from the relative-rate tests at synonymous and non-synonymous sites in the sequences of *mat*K and *rbc*L genes are given in Table 2 and Table 3, respectively. Our exploratory analyses detect a number of significant rate differences at synonymous and non-synonymous sites in the two plastid sequences. First, two significant contrasts are that the mangrove *Bruguiera* has relatively slower substitution rates than the terrestrial *Carallia* at both
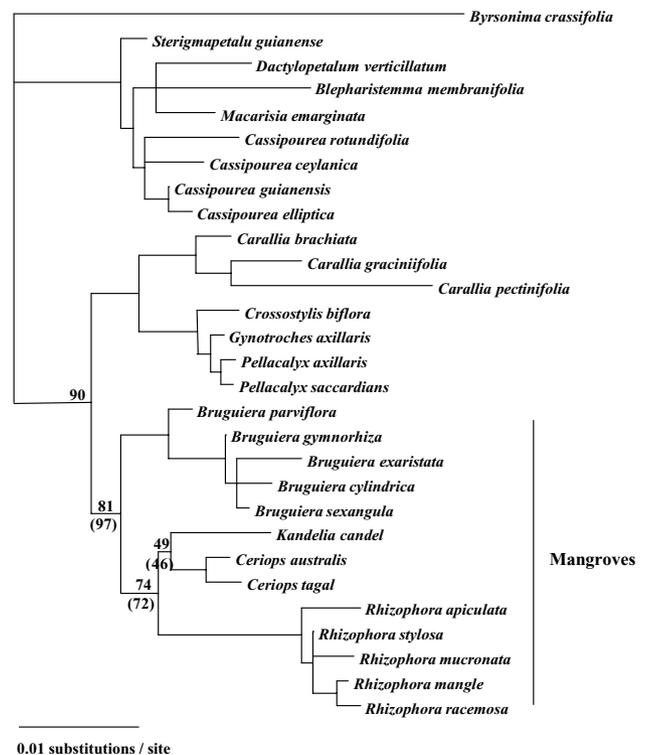

Figure 1 A phylogenetic tree of Rhizophoraceae based on the *mat*K sequences constructed by the maximum likelihood method with MOLPHY. Numbers indicate the bootstrap proportions (percentage over 10 000 replicates).

Figure 2 A phylogenetic tree of Rhizophoraceae based on the *rbc*L sequences constructed by the maximum likelihood method with MOLPHY. Numbers without the parentheses indicate the bootstrap proportions (percentage over 10 000 replicates). Numbers in the parentheses indicate the bootstrap proportions for the total of *rbc*L and *mat*K sequences.

	PE	CR	RH	KA	CE	BR
PE	–	−0.62	0.03	0.09	−0.56	1.08
CR	0.48	–	−0.04	0.56	−0.63	−4.41**
RH	−0.62	−1.75	–	0.36	−0.45	1.11
KA	0.92	0.82	1.60	–	−0.99	0.77
CE	0.19	−0.57	1.60	−0.59	–	1.16
BR	0.12	2.04*	1.23	−0.31	−0.72	–

Table 2 Results of relative-rate tests for *mat*K sequences[1]

K_{ij}/SE is given as $(K_i - K_j)$/SE so that a K_{ij}/SE > 0 means that the *i*th column is estimated to have faster substitution rates than the row. Each lower-triangular matrix represents the results at synonymous sites, and each upper-triangular matrix represents the results at non-synonymous sites.
[1]Abbreviations of genera see Table 1.
*$P < 0.05$, **$P < 0.01$.

Table 3 Results of relative-rate tests for *rbc*L sequences[1]

	BL	ST	DA	CS	MA	CO	GY	PE	CR	RH	KA	CE	BR
BL	–	0	−0.49	−1.29	−1.42	−1.76	−2.01*	−1.61	1.15	0.30	−0.71	−0.51	−0.82
ST	1.36	–	1.97*	2.57*	0.51	0.22	−0.23	0.21	1.63	1.95*	1.24	1.32	1.19
DA	0.20	−0.78	–	−0.78	−2.00*	−1.37	−1.81	−1.41	−0.39	0.35	−0.29	−0.33	−0.33
CS	1.55	0.42	1.15	–	−1.87	−1.12	−1.26	−0.78	0.76	−0.16	0.44	0.44	0.36
MA	1.22	−0.54	0.20	−1.03	–	−0.13	−0.27	0.24	1.30	2.25*	1.08	1.38	1.19
CO	−0.03	−0.62	−0.18	−0.83	−0.33	–	−0.44	0.45	2.44*	2.39*	1.29	1.61	1.31
GY	0.13	−0.57	−0.02	−0.77	−1.16	0.48	–	1.44	2.06*	2.70**	2.98**	1.77	1.42
PE	0.03	−0.66	−0.10	−0.86	−0.26	0.19	−0.35	–	1.37	2.26*	1.22	1.26	1.02
CR	0.00	−1.36	−0.70	−1.82	0.00	−1.41	−1.50	−1.32	–	−0.40	−2.11*	−0.77	−0.96
RH	0.03	−0.47[2]	−0.10[2]	−0.42[2]	0.19[2]	0.87[2]	−1.18[2]	−1.19[2]	1.08[2]	–	−0.92	−1.06	−0.94
KA	−0.55	−1.11	−0.55	−1.31	−0.87	−0.68	−1.01	−0.59	0.92	−0.13[2]	–	−0.04	−0.06
CE	−0.30	−0.91	−0.30	−1.13	−0.62	−0.35	−0.39	−0.26	1.36	0.75[2]	0.50	–	−0.03
BR	−0.13[2]	−0.41[2]	0.04[2]	−0.59[2]	−0.35[2]	−1.10[2]	−1.10[2]	−1.11[2]	0.68[2]	0.19[2]	0.30[2]	−0.62[2]	–

[1]Abbreviations of genera see Table 1.
K_{ij}/SE is given as $(K_i - K_j)$/SE so that a K_{ij}/SE > 0 means the *i*th column is estimated to have faster substitution rates than the row. Each lower-triangular matrix represents the results at synonymous sites, and each upper-triangular matrix represents the results at non-synonymous sites.
[2]Synonymous transversion substitution when K_s is saturated.
*$P < 0.05$, **$P < 0.01$.

synonymous and non-synonymous sites in the *mat*K sequences (Table 2). The Mantel tests show that the synonymous and non-synonymous relative-rate matrices are correlated at the *mat*K ($P = 0.01$). This observation suggests that selective constraint at non-synonymous sites is fairly constant among evolutionary lineages of the *mat*K locus.

Second, there are 13 significant contrasts at non-synonymous sites in the *rbc*L sequences (Table 3). Among them, six indicate that the mangrove genera have relatively faster non-synonymous substitution rates than the related terrestrial groups. However, the terrestrial *Carallia* still shows a relatively faster non-synonymous rate than the mangrove *Kandelia*. In our previous paper (Zhong *et al.* 2000), we noted that there were significant total substitution rate differences between *Carallia brachiata* and other *Carallia*

species in the *mat*K sequences. Obviously, these *Carallia* species need to be further studied as important reference taxa relative to mangroves.

Moreover, the *rbc*L non-synonymous sites also exhibit rate heterogeneity among the terrestrial groups, regardless of their geographical distributions (Table 3). The Mantel tests show that the *rbc*L rates at synonymous and non-synonymous sites are uncorrelated ($P = 1.00$). This can be also seen to some extent by visual examination of Table 3, e.g. there are independently increased non-synonymous rates against synonymous rates in the two genera *Carallia* and *Rhizophora*.

Our analyses of the two chloroplast genes not only clearly indicate that the rates of synonymous and non-synonymous substitution vary among evolutionary lineages of mangroves

and their terrestrial related groups, but also increase the understanding of possible forces contributing to the patterns. It is well known that synonymous substitutions primarily reflect the process of neutral evolution but non-synonymous substitutions reflect either positive selection or purifying selection over evolutionary time (Gaut *et al.* 1992; Gaut *et al.* 1997; Li 1997; Yang & Bielawski 2000). Recently, Gaut *et al.* (1997) reported the uncoupling of *rbc*L non-synonymous and synonymous rates in the grass family, i.e. Poaceae, suggesting that locus-specific selection coefficients can change significantly among lineages. Our study reveals molecular evolutionary patterns of mangroves and their terrestrial relatives in which non-synonymous and synony-mous substitution rates are uncoupled, and therefore selection is probably an important influence on the rate variation.

The speciation rate hypothesis has previously also been invoked to explain rate heterogeneity at *rbc*L (Gaut *et al.* 1997). Mayr (1954) postulated that rates of genetic change should influence rates of speciation, and Bousquet *et al.* (1992) and Barraclough *et al.* (1996) revealed a correlation between speciation rates and nucleotide substitution rates. Gaut *et al.* (1997) analysed the *rbc*L sequences in the grass family and found that the basal grass lineage (subfamily Anomochlooideae) had slower nucleotide substitution rates than other grass subfamilies, that is, their prediction based on the speciation rate hypothesis was consistent with the results of synonymous relative-rate tests. However, our relative-rate tests reveal that no significant contrast at the synonymous sites in the *rbc*L sequences is detected (Table 3), yet there is no basal lineage of mangroves in the *rbc*L tree (Fig. 2).

Excluding speciation rate hypothesis, two possible explanations for the rate heterogeneity are about different ecological and biogeographical forces to the mangroves and related terrestrial groups. On the one hand, all mangrove species, regardless of geographical distribution, form a monophyletic group on both the *mat*K and *rbc*L trees. The rate variations do not occur within the mangrove group but between the mangrove and related terrestrial groups, indicating that the marine environment might be a major selection force to mangroves. On the other hand, the related terrestrial groups show different phylogenetic patterns. In particular, there are also relatively complex patterns of rate variation at *rbc*L among the groups, e.g. significant differ-ences between *Carallia* (CR) distributed in Asia and *Crossostylis* (CO) distributed in the South Pacific Islands, and between *Dactylopetalum* (DA) and *Macarisia* (MA), both distributed in Africa. These cases may reflect the effects of selection imposed by different climates and habitats, adding a biogeographical dimension to the problem.

In general, the environmental mechanisms responsible for molecular evolutionary patterns of mangroves and related groups require examination. Positive selection, such as we describe here, is believed to reflect powerful selective forces acting on protein-coding sequences, and may stem from such factors as changed ecological conditions (Ford 2001) or an arms race scenario (Dawkins & Krebs 1979). We have uncovered the traces of such a scenario here, and note that both ecological and molecular studies will be needed to elucidate the causes of the phenomenon in mangroves.

REFERENCES

Adachi, J. & Hasegawa, M. (1996). MOLPHY: Programs for Molecular Phylogenetics, Version 2.3. Computer Science Monographs 28. Institute of Statistical Mathematics, Tokyo.

Angiosperm Phylogeny Group. (1998). An ordinal classification for the families of flowering plants. *Ann. Missouri Bot. Gard.*, 85, 531–553.

Barraclough, T.G., Harvey, P.H. & Nee, S. (1996). Rate of *rbc*L gene sequence evolution and species diversification in flowering plants. *Proc. R. Soc. Lond. B*, 263, 589–591.

Bousquet, J., Strauss, S.H., Doerksen, A.H. & Price, R.A. (1992). Extensive variation in evolutionary rate of *rbc*L gene sequences among seed plants. *Proc. Nat. Acad. Sci. USA*, 89, 7844–7848.

Dawkins, R. & Krebs, J.R. (1979). Arms races between and within species. *Proc. R. Soc. Lond. B*, 205, 489–511.

Dodd, R.S., Rafii, Z.A., Fromard, F. & Blasco, F. (1998). Evolutionary diversity among Atlantic coast mangroves. *Acta Oecologica.*, 19, 323–330.

Ford, M.J. (2001). Molecular evolution of transferrin: evidence for positive selection in salmonids. *Mol. Biol. Evol.*, 18, 639–647.

Gaut, B.S., Clark, L.G., Wendel, J.F. & Muse, S.V. (1997). Comparisons of the molecular evolutionary process at *rbc*L and *ndh*F in the grass family (Poaceae). *Mol. Biol. Evol.*, 14, 769–777.

Gaut, B.S., Morton, B.R., McCaig, B.M. & Clegg, M.T. (1992). Relative rates of nucleotide substitution at the *rbc*L locus of monocotyledonous plants. *J. Mol. Evol.*, 35, 292–303.

Geh, S. & Keng, H. (1974). Morphological studies on some inland Rhizophoraceae. *Gard. Bull. Singapore*, 27, 183–221.

Hasegawa, M. & Kishino, H. (1994). Accuracies of the simple methods for estimating the bootstrap probability of a maximum likelihood tree. *Mol. Biol. Evol.*, 11, 142–145.

Hasegawa, M., Kishino, H. & Yano, T. (1985). Dating of the hu-man–ape splitting by a molecular clock of mitochondrial DNA. *J. Mol. Evol.*, 22, 160–174.

Huang, Y., Qiu, X., Shi, S., Tan, F. & Chang, H.T. (1999). The molecular phylogeny in Rhizophoraceae in China. *Acta Sci. Nat. Univ. Sunyatseni*, 38, 39–45 (in Chinese with English summary).

Kishino, H., Miyata, T. & Hasegawa, M. (1990). Maximum likeli-hood inference of protein phylogeny, and the origin of chlo-roplasts. *J. Mol. Evol.*, 31, 151–160.

Li, W.H. (1993). Unbiased estimation of the rates of synonymous and nonsynonymous substitution. *J. Mol. Evol.*, 36, 96–99.

Li, W.H. (1997). *Molecular Evolution*. Sinauer Associates, Sunderland, MA.

Li, P. & Bousquet, J. (1992). Relative-rate test for nucleotide substitutions between two lineages. *Mol. Biol. Evol.*, 9, 1185–1188.

Liedloff, A. (1999). MANTEL: *Mantel Nonparametric Test Calculator*, Version 2.0. Queensland University of Technology, Brisbane, Australia.

Mayr, E. (1954). Change of genetic environment and evolution. In: *Evolution as a Process* (eds Huxley, J., Hardy, A.C. & Ford, E.B.). George Allen & Unwin, London, pp. 157–180.

Pamilo, P. & Bianchi, N.O. (1993). Evolution of the *Zfx* and *Zfy* genes: rates and interdependence between the genes. *Mol. Biol. Evol.*, 10, 271–281.

Purvis, A. & Hector, A. (2000). Getting the measure of biodiversity. *Nature*, 405, 212–219.

Robinson, M., Gouy, M., Gautier, C. & Mouchiroud, D. (1998). Sensitivity of the relative-rate test to taxonomic sampling. *Mol. Biol. Evol.*, 15, 1091–1098.

Robinson-Rechavi, M. & Huchon, D. (2000). RRTree: relative-rate tests between groups of sequences on a phylogenetic tree. *Bioinformatics*, 16, 296–297.

Schulte, P.M. (2001). Environmental adaptations as windows on molecular evolution. *Comp. Biochem. Physiol. B*, 128, 597–611.

Schwarzbach, A.E. & Ricklefs, R.E. (2000). Systematic affinities of Rhizophoraceae and Anisophylleaceae, and intergeneric relationships within Rhizophoraceae, based on chloroplast DNA, nuclear ribosomal DNA, and morphology. *Am. J. Bot.*, 87, 547–564.

Setoguchi, H., Kosuge, K. & Tobe, H. (1999). Molecular phylogeny of Rhizophoraceae based on *rbc*L gene sequences. *J. Plant Res.*, 112, 443–455.

Shi, S., Zhong, Y., Huang, Y., Qiu, X. & Chang, H.T. (2002). Phylogenetic relationships of the Rhizophoraceae in China based on sequences of the chloroplast gene *mat*K and ITS regions of nuclear ribosomal DNA and combined data set. *Biochem. Syst. Ecol.*, 30, 309–319.

Sokal, R.R. & Rohlf, F.J. (1995). *Biometry*. W.H. Freeman, New York.

Spalding, M., Blasco, F. & Field, C. (1997). *World Mangrove Atlas*. Samara Publishing Co, Cardigan, UK.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F. & Higgins, D.G. (1997). The Clustal–X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucl. Acids Res.*, 25, 4876–4882.

Webb, C.O. (2000). Exploring the phylogenetic structure of ecological communities: an example for rain forest trees. *Am. Nat.*, 156, 145–155.

Yang, Z. & Bielawski, J.P. (2000). Statistical methods for detecting molecular adaptation. *Trends Ecol. Evol.*, 15, 496–503.

Zhong, Y., Shi, S., Tang, X., Huang, Y., Tan, F. & Zhang, X. (2000). Testing relative evolutionary rates and estimating divergence times among six genera of Rhizophoraceae using cpDNA and nrDNA sequences. *Chinese Sci. Bull.*, 45, 1011–1015.