广西植物 Guihaia Jan. 2016, **36**(1):101-106
http://journal.gxzw.gxib.cn
http://www.guihaia-journal.com

DOI: 10.11931/guihaia.gxzw201508017
李景剑,刘合霞,毛世忠,等. 大黄属(蓼科)植物 *ndh*F 基因的适应性进化[J]. 广西植物, 2016, 36(1):101-106
LI JJ,LIU HX,MAO SZ,et al. Adaptive evolution of the *ndh*F gene in the genus *Rheum* (Polygonaceae)[J]. Guihaia, 2016, 36(1):101-106

Adaptive evolution of the *ndh*F gene in the genus *Rheum* (Polygonaceae)

LI Jing-Jian[1,2], LIU He-Xia[2], MAO Shi-Zhong[2],
ZHAO Bo[2,3], HUANG Shi-Xun[2*]
(1. *College of Forestry and Landscape Architecture*, *South China Agricultural University*, Guangzhou 510642, China; 2. *Guangxi Institute of Botany*, *Guangxi Zhuang Autonomous Region and the Chinese Academy of Sciences*, Guilin 541006, China; 3. *Institute of Chinese Materia Medica*, *China Academy of Chinese Medical Sciences*, Beijing 100193, China)

Abstract: *Rheum*, a highly diversified genus of Polygonaceae, comprising about 60 species, is mainly distributed in the mountainous and desert regions in Asia and Europe. Its high diversification in the Qinghai-Tibetan Plateau and adjacent areas,where approximately 40 species have been described,was suggested to result from the adaptive radiation triggered by the recent uplifts of the Qinghai-Tibetan Plateau and the quaternary climate oscillations. To further understand the molecular adaptation linked to *Rheum* radiation, the adaptive evolution of chloroplast *ndh*F gene belonging to 34 morphologically diversified species of the genus *Rheum* were tested by using the Phylogenetic Analysis Program in this study. The results showed that the pattern of phylogenic tree was a typical "paralleling" phylogeny, suggesting a radiative diversification. Three amino acid sites (188H,465H,551L) of NDHF subunit were identified under positive selection ($\omega > 1$),and the secondary structures of NDHF subunit showed that the 188th amino acid was located in the α-helix. The changes of these adaptive sites may be associated with the evolution of NDH to adapt the extreme habitats of *Rheum* species caused by the extensive uplifts of the QTP and the quaternary climate oscillations.

Key words: *Rheum* (Polygonaceae), *ndh*F gene, branch-specific model, site-specific model, positive selection
CLC number: Q941.2 **Document code**: A **Article ID**: 1000-3142(2016)01-0101-06

大黄属(蓼科)植物 *ndh*F 基因的适应性进化

李景剑[1,2], 刘合霞[2], 毛世忠[2], 赵博[2,3], 黄仕训[2*]
(1. 华南农业大学 林学与风景园林学院, 广州 510642; 2. 广西壮族自治区中国科学院广西植物研究所, 广西 桂林 541006; 3. 中国中医科学院 中药研究所, 北京 100700)

摘　要: 大黄属(*Rheum* L.)是蓼科(Polygonaceae)中一个高度分化的大属,广泛分布在亚洲和欧洲的高山和沙漠地区,全世界约 60 种,其中在青藏高原及其邻近地区发现了约 40 种。该属种的高度分化曾被推测是第三纪末青藏高原的快速隆升以及第四纪气候的反复变化所引发的适应性辐射导致。为进一步了解大黄属植物辐射式物种分化的分子适应机制,该研究选取 34 个形态上多样化的大黄属物种,利用系统发育分析软件,在时间框架下采用位点模型和分支模型对大黄属的叶绿体 *ndh*F 基因进行了适应性进化分析。结果表明:大黄属植物的分

收稿日期: 2015-08-24 **修回日期**: 2015-12-02
基金项目: 国家科技基础性工作专项(2009FY120200);广西自然科学基金(2012GXNSFBA053075);广西植物研究所基本业务费(桂植业 14003) [Supported by the Special Program for Basic Research of Science and Technology of China(2009FY120200); the Natural Science Foundation of Guangxi (2012GXNSFBA053075); the Science Research Foundation of Guangxi Institute of Botany (Guizhiye14003)]。

作者简介: 李景剑(1983-),男,广东廉江人,博士研究生,主要从事植物分子生物学研究,(E-mail)calljone@163.com。
[*] **通讯作者**: 黄仕训,研究员,主要从事濒危植物保护研究,(E-mail)hsx@gxib.cn。

子进化系统树呈现短而平行的辐射式分支式样，显示出典型的物种快速辐射多样化特征；用位点模型检验 *ndh*F 基因是否存在经受正向选择（$\omega>1$）时，在氨基酸水平上共鉴定出 3 个 NDHF 亚基的正选择位点（188H，465H，551L），对 NDHF 亚基的二级结构进行分析后发现编码的 188H 氨基酸位于 α 螺旋上。大黄属植物可能通过这些结构域的适应性进化，适应青藏高原的快速隆升以及第四纪气候的反复变化而引发的陆地生态系统改变。该研究结果可为今后对该属植物的实验分析提供首选位点。

关键词：大黄属（蓼科），*ndh*F 基因，分支模型，位点模型，正选择位点

The genus *Rheum* (Polygonaceae) with about 60 species, primarily distributed in mountainous and desert regions of the Qinghai-Tibetan Plateau and adjacent areas (Kao & Cheng, 1975; Li, 1998). The distribution and the ancestral area reconstruction analyses consistently suggested that rapid radiations of *Rheum* have occurred, and may have been caused by the extensive uplifts of the Qinghai–Tibetan Plateau (Wang et al, 2005; Wan et al, 2011; Sun et al, 2012). To adapt to the new alterations of habitat, morphological traits of this genus are highly diversified. Some species have evolved into dwarf plants with coriaceous basal leaves or drooping bracts to defense them from freeze injury and could distribute up to snow line at altitude of 5 400 m, for example *R. Nobile* (Xie, 2000). For other species, stem leaves are have degenerated and basal leaves are covered with verruca or indumentum to reduce water transpiration, to avoid the burning from high temperature and to avoid damage by strong winds, therefore, these species can grow in the Gobi Desert at altitude 700 m. *R. palaestinum* in particular, has broad, rigid leaves, with a waxy surface, and channels cut into them that funnel any water that drops onto them toward its root, with enough force to cause deep soil penetration (Lev-Yadun et al, 2009). These changes in morphology and physiology might be resulted from the adaptive evolution of some genes which encode functional proteins, such as chloroplast *ndh*F gene that related to photosynthesis and photorespiration (Zapata et al, 2005).

Chloroplast is thought to be a very conservative part of plant genome but little is known about the evolution of this plastome promoters. Previous study showed that the alignment of sequences upstream *ndh*F suggested that promoters of this gene underwent comparatively rapid evolution in flowering plants (Seliverstov et al, 2009). The *ndh*F gene is located in a small single-copy region of the chloroplast genome that rarely underwent substantial rearrangements in terrestrial plants (Hiratsuka et al, 1989). Its nucleotide sequence predicts a hydrophobic protein of 664 amino acids with a calculated mass of 72.9 kDa (Schluchter et al, 1993). The *ndh*F gene encodes NADH dehydrogenase F subunit of the plastid NDH complex which regulated the activity of NDH complexes by its phosphorylation. The plastid NDH complex in chloroplast thylakoid membranes is involved in photosystem I cyclic and chlororespiratory electron transport in photosynthetic regulation of higher plants (Lascano et al, 2003).

Considering the adaptability of *Rheum* to extreme habitats, the sequences of *ndh*F gene from 34 species of *Rheum* and 2 species of *Oxyria* in Polygonaceae were retrieved from the National Center for Biotechnology Information (NCBI) for adaptive evolution testing in this study. Our finding may provide new molecular evidence for the rapid putative radiations of *Rheum* triggered by the recent uplifts of the Qinghai-Tibetan Plateau.

1 Materials and Methods

Thirty-four species of *Rheum* and two species of *Oxyria* in Polygonaceae used in this study were listed in Table 1. Sequences of *ndh*F gene were downloaded from NCBI (http://www.ncbi.nlm.nih.gov/guide/). Sequence alignments were conducted using the software CLUSTAL W ver. 1.83 (Thompson et al, 1994) and adjusted manually in BioEdit 5.0.9.1 (Hall, 1999). *Oxyria digyna* and *O. sinensis* were used as outgroup. Maximum Parsimony (MP) analysis was conducted using PAUP 4.0b10 (Swofford, 2003), Heuristic searches were conducted 1 000 times with random taxon-addition sequences, with tree-bisection-reconnection (TBR) branch swapping, and with the options MULPARS in effect and STEEPEST DE-

SCENT off. Support for internal nodes was estimated with bootstrap values (Felsenstein,1985).

Table 1 Species names and accession numbers of *ndh*F sequences

No.	Species	GenBank No.	No.	Species	GenBank No.
1	*Rheum acuminatum*	FJ872086	19	*R. palaestinum*	FJ872084
2	*R. alexandrae*	EU840401	20	*R. palmatum*	EU840398
3	*R. alpinum*	EU840385	21	*R. przewalskyi*	EU840379
4	*R. altaicum*	EU840404	22	*R. pumilum*	EU840389
5	*R. australe*	EU840393	23	*R. mreticulatum*	EU840383
6	*R. compactum*	EU840403	24	*R. rhaponticum*	EU840402
7	*R. delavayi*	FJ872085	25	*R. rhizostachyum*	EU840380
8	*R. forrestii*	EU840388	26	*R. rhomboideum*	EU840378
9	*R. franzenbachii*	EU840405	27	*Rheum* sp.	EU840392
10	*R. globulosum*	EU840386	28	*R. spiciforme*	EU840377
11	*R. hotaoense*	EU840406	29	*R. sublanceolatum*	EU840390
12	*R. kialense*	EU840387	30	*R. tanguticum*	EU840399
13	*R. lhasaense*	EU840395	31	*R. tataricum*	EU840407
14	*R. likiangense*	EU840396	32	*R. tibeticum*	EU840391
15	*R. moorcroftianum*	EU840384	33	*R. webbianum*	EU840381
16	*R. nanum*	EU840397	34	*R. wittrockii*	EU840400
17	*R. nobile*	EU840382	35	*Oxyria digyna*	EU840375
18	*R. officinale*	EU840394	36	*O. sinensis*	JQ342161

Based on the MP tree, the analysis of adaptive evolution of *ndh*F gene was implemented in the program of CODEML from PAML package version 4 (Yang,2007). The lnL values under one-ratio model as well as free ratio model were calculated, and the Likelihood Ratio Test (LRT) was conducted to test whether there were different ratios for each lineage. Site-specific models, which allowed the ω ratio to vary among sites but fixed a single ω ratio in all branches, were used to detect positive selection and to identify positively selected sites. Three pairs of site-specific models were calculated to test for recurrent, diversifying, selection: M0 (one ratio) and M3 (Discrete), M1 (Neutral) and M2 (Selection), and M7 (Beta) and M8 (Beta & ω). (Yang & Nielsen, 2002; Yang et al,2005). Log likelihoods of models (M1 vs. M2; M0 vs. M3; M7 vs. M8) were compared using LRT.

For the spatial analysis of the codon site under positive selection, the PSIPRED server (http://bioinf.cs.ucl.ac.uk/psipred/) was used to analysis the secondary structure of NDHF subunit for *Rheum palaestinum*.

2 Results and Analysis

The *ndh*F dataset had an aligned length of 1 944 characters in the dataset, of which, 1 751 characters were constant, 111 were variable and parsimony-uninformative, and 82 were parsimony-informative. Maximum Parsimony analysis yielded 84 equally parsimonious trees, and a strict consensus tree of these trees was shown in Fig.1. The topology of MP tree was consistent with the molecular phylogenies published to date (Wang et al, 2005; Sun et al, 2012). The consensus tree revealed three major clades (A, B, C) within the genus *Rheum*, all species of *Rheum* comprised a well-supported lineage, with a sister relationship to *Oxyria*.

To analyze the possibility that positive selection acts on *ndh*F genes, we used the maximum-likelihood codon model from the CODEML program in the PAML4 package. The topology of the MP tree mentioned above was modified for all CODEML analyses. All calculations and tests are listed in Tables 2 and 3. Under the one-ratio model which allowed for only a single ω ratio across all sites of the gene phylogeny and the same ω ratio for all branches in the phylogenic tree (Fig. 1), the log-likelihood value was $\omega = 0.296\ 5$, lower than 1 (Table 2). In the branch-specific analysis, the LRT statistic for the comparison of the one-ratio model vs. the free-ratio model was $2\Delta\ell = 77.467\ 0$ with $P<0.05$ and $df = 54$, suggesting that there had different ratios for each lineage of *Rheum* species (Table 3). But no sites with a Bayesian posterior probability of positive selection larger than 0.95 in one or more cases were found when analyzed by Bayes.

In site-specific models, models M2, M3 and M8 allowed sites with $\omega>1$. The LRT statistic of M0-M3, M1-M2 and M7-M8 comparison all with $P<0.05$, so models M3, M2 and M8 was significantly better than M0, M1 and M7. Under both M2 and M8 models, three sites were under positive selection with $\omega>1$ and identified three NDHF residues (188H, 465H, 551L) with a Bayesian posterior probability of positive selection larger than 0.95 in one or more cases when analyzed by Empirical Bayes (Table 2).

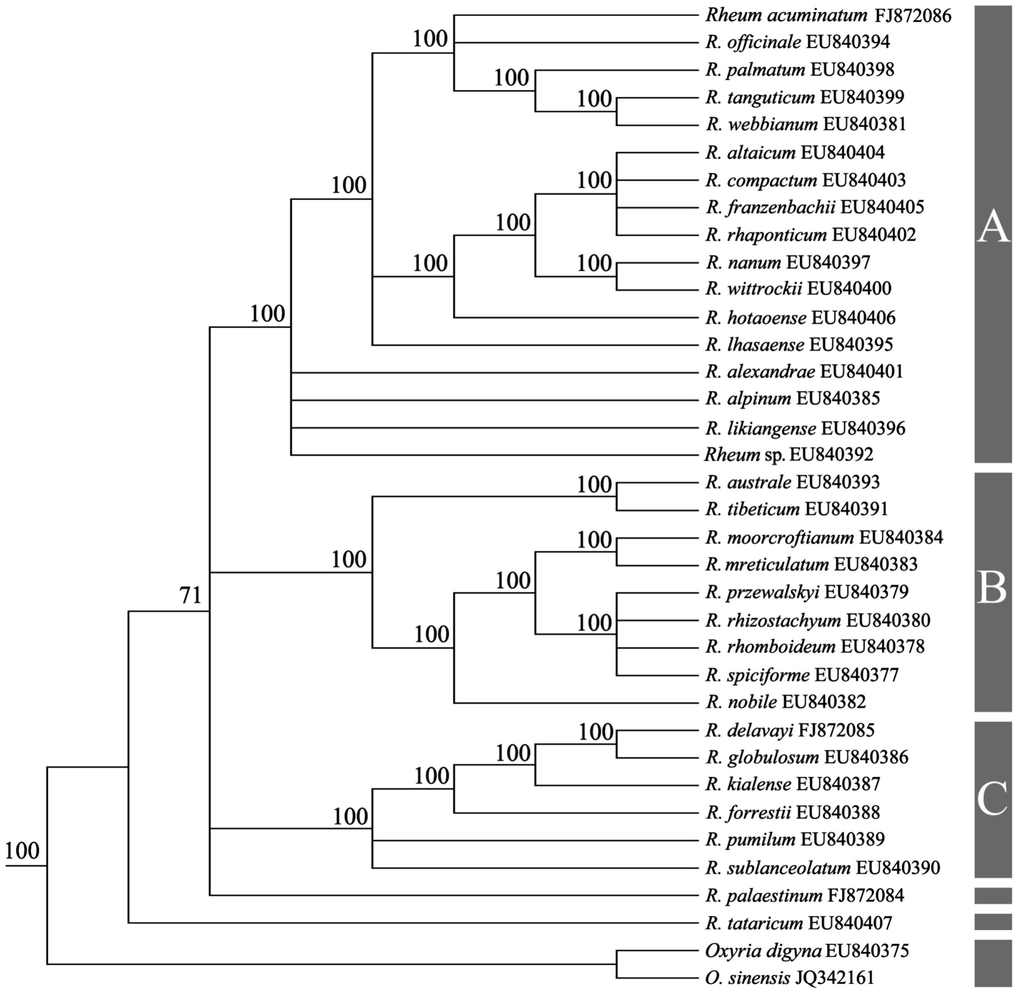

Fig. 1 Strict consensus tree from maximum-parsimony (MP) analysis based on *ndh*F sequences of *Rheum* Numbers above the branches indicate bootstrap values by MP analysis. Numbers with bootstrap values > 50% are shown. Shadows on the right indicate five clades of *Rheum* displaying rapid radiation.

With no obvious sequence similarity to structures present in PDB, the secondary structure of NDHF subunit for *Rheum palaestinum* was predicted by PSIPRED server (Fig. 2). The test results showed the 188th amino acid which located in the α-helix was histidine (H) in *R. palaestinum*, while asparagine (N) was also found in other species of *Rheum*. The amino acid encoded by the 465th codon was histidine (H) in *R. palaestinum*, while asparagine (N) and tyrosine (Y) was found in other species of *Rheum*. In addition, the 551th codon encoded was leucine (L) in *R. palaestinum*, while phenylalanine (F) was found in other species of *Rheum*.

3 Discussion

The distribution and the ancestral area reconstruction analyses suggests that rapid putative radiations of *Rheum* might have been triggered by the recent uplifts of the Qinghai-Tibetan Plateau and the Quaternary climate oscillations. Geological evidence indicates that at least four different periods at the early Miocene (i.e., 22, 15–13, 8–7, and 3.5–1.6 Ma) occurred during recent extensive uplifting of the Qinghai-Tibetan Plateau (Shi et al, 1998; Sun et al, 2012), and new habitats may have been created while old ones became fragmented within each period. The new alterations of habitat of *Rheum* species are various, from snow line at altitude 5 400 m to Gobi Desert at altitude 700 m. Through the adaptive evolution of *ndh*F gene involved in photosynthesis pathways, some species of *Rheum* could adapt the extreme habitats.

By comparing Models M1a/M2a and M7/M8, three

Table 2　Maximum likelihood parameter estimates for *ndh*F gene

Model	n_p	ℓ	Estimates of parameter	Positively selected site
M0：one-ratio	57	-4205.9904	$\omega=0.2965$	None
Branch-specific models	—	—	—	—
F：free-ratio	111	-4167.2569	—	Not allowed
Site-specific models	—	—	—	—
M1：neutral	58	-4162.9674	$P_0=0.7513$；$P_1=0.2487$	Not allowed
M2：selection	60	-4154.5725	$P_0=0.8588$；$P_1=0.0900$； $P_2=0.0512$；$\omega=3.6640$	22M, 24L, 42T, 110F, 169F, 178F, 188H＊, 243V, 461P, 462I, 465H＊, 470K, 472P, 476R, 487N, 495S, 551L＊＊, 555S, 602G, 611V
M3：discrete（K=3）	61	-4154.4436	$P_0=0.9010$；$P_1=0.0890$； $P_2=0.0100$；$\omega_0=0.1081$； $\omega_1=2.2318$；$\omega_2=6.1884$	Not allowed
M7：β, neutral	58	-4163.5194	a=0.0050；b=0.0114	Not allowed
M8：β and ω selection	60	-4154.6300	$P_0=0.9360$；$P_1=0.0640$； a=0.4887；b=2.7044； $\omega=3.3811$	22M, 24L, 42T, 110F, 169F, 178F, 188H＊＊, 243V, 461P, 462I, 465H＊, 470K, 472P, 476R, 487N, 495S, 551L＊＊, 555S, 590V, 602G, 611V, 628I

Note：ℓ. Log likelihood；np. Number of parameters；ω. Ratio of synonymous/non-synonymous substitutions；P. Proportion of sites estimated to be under positive selection with $\omega>1$；a, b. parameters of the beta distribution. The positively selected sites were in ＊ and ＊＊ under 95% and 99% posterior probability values, respectively.

Fig. 2　The secondary structures of NDHF protein for *Rheum palaestinum*　The sites（188H, 465H, 551L）under adaptive evolution are marked with red boxes.

amino acid sites（188H, 465H, 551L）were identified under positive selection. The secondary structure of NDHF subunit showed thatthe positively selected sites（465H and 551L）were on the loops. The 188th amino acid which located in the α-helix was histidine（H）in *R. palaestinum*（Fig. 2）, while asparagine（N）was found in other species. The *ndh*F gene encoded a subunit of the plastid NDH complex, and this complex assembly might be regulated on the pos-transcriptional level in a way that the quantity of whole NDH complexes

Table 3 Likelihood ratio test (LRTs) of the variable ω ratios under different models for *ndh*F gene

Comparison	2⊿ℓ	Degree of freedom	*P*
M0 vs. F	77.4670*	54	$P<0.05$
M0 vs. M3	103.0936***	4	$P<0.001$
M1 vs. M2	16.7898**	2	$P<0.01$
M7 vs. M8	17.7788**	2	$P<0.01$

Note: *Significant difference at the $P<0.05$ level; ** Significant difference at the $P<0.01$ level; *** Significant difference at the $P<0.001$ level.

could be determined by the quantity of one of its subunits, e.g. NDHF. The activity of NDH complexes was also regulated by phosphorylation of the NDHF polypeptide. Our study found that the *ndh*F gene was at high expression level under stress conditions, and those stress factors from environment might be the selective pressure to lead the adaptive evolution of *ndh*F gene. Our results indicated that the change of spatial structure may have a relationship with the adaptation of *Rheum* to the environment. For example, *R. palaestinum* with histidine mutation on the 188th codon site is the rare *Rheum* plant growing in mountainous desert areas (receiving an average annual rainfall of ca. 75 mm) in the world, and it has a single deep main vertical root (Zohary, 1966). Previous studies released that histidine is one of the essential amino acids for plant growth and survival, especially for root meristem maintenance, the higher histidine content in plant, the faster root growth and better adaptability to the environments (Mo, 2006; Malki & Jaeobs, 2001). So the adaptive evolution of *ndh*F subunit might be important for *Rheum* species to adapt various habitats.

Acknowledgements We are grateful to LIU Lei for his help with data analysis and Qian Guo for language editing support.

References:

FELSENSTEIN J, 1985. Confidence limits on phylogenies: an approach using the bootstrap [J]. Evolution, 39: 783-791.

HALL TA, 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT [M]. Nucl Acids Symp Ser, 41: 95-98.

HIRATSUKA J, SHIMADA H, WHITTIER R, et al, 1989. The complete sequence of the rice (*Oryza sativa*) chloroplast genome: Intermolecular recombination between distinct tRNA genes accounts for a major plastid DNA inversion during the evolution of the cereals [J]. Mol Gen Genet, 217: 185-194.

KAO TC, CHENG CY, 1975. Synopsis of the Chinese *Rheum* [J]. Acta Phytotax Sin, 13: 69-82.

LASCANO HR, CASANO LM, MARTIN M, et al, 2003. The activity of the chloroplastic ndh complex is regulated by phosphorylation of the NDH-F subunit [J]. Plant Physiol, 13: 2 256-2 262.

LEV-YADUN S, KATZIR G, NEEMAN G, 2009. *Rheum palaestinum* (desert rhubarb), a self-irrigating desert plant [J]. Beijing: Naturwissenschaften, 96: 393-397.

LI AR, 1998. Flora republicae popularis sinicae [M]. Beijing: Beijing Science Press.

MALKI FE, JAEOBS M, 2001. Molecular characterization and expression study of a histidine auxotorphic mutant (hisl-) of Nicotiana plumbaginifolia [J]. Plant Mol Biol, 45: 191-199.

MO XR, 2006. The functional analysis of *Arabidopsis* HPA1 gene [D]. Hangzhou: Zhejiang University.

SELIVERSTOV AV, LYSENKO EA, LYUBETSKY VA, 2009. Rapid evolution of promoters for the plastome Gene *ndh*F in flowering plants [J]. Russ J Plant Physl, 56(6): 838-845.

SCHLUCHTER WM, ZHAO JD, BRYANT DA, 1993. Isolation and characterization of the *ndh*F gene of *Synechococcus* sp. Strain PCC 7002 and initial characterization of an interposon mutant [J]. J Bacteriol, 175(11): 3 343-3 352.

SHI YF, LI JJ, LI BY, 1998. Uplift and environmental changes of Qinghai-Tibetan Plateau in the Late Cenozoic [M]. Guangzhou: Guangdong Science and Technology Press.

SUN Y, WANG A, WAN D, et al, 2012. Rapid radiation of *Rheum* (Polygonaceae) and parallel evolution of morphological traits [J]. Mol Phylogenet Evol, 63: 150-158.

SWOFFORD DL, 2003. PAUP*: Phylogenetic analysis using parsimony (*and other methods) [M]. Sunderlande, MA: Sinauer Associates.

THOMPSON JD, GIBSON TJ, PLEWNIAK F, et al, 1997. The ClustalX Windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools [M]. Nucleic Acids Res, 24: 4 876-4 882.

WAN D, WANG A, ZHANG X, et al, 2011. Gene duplication and adaptive evolution of the CHS-like genes within the genus *Rheum* (Polygonaceae) [J]. Biochem Syst Ecol, 39: 651-659.

WANG AL, YANG MH, LIU JQ, 2005. Molecular phylogeny, recent radiation and evolution of gross morphology of the Rhubarb genus *Rheum* Polygonaceae inferred from chloroplast DNA *trn*L-F sequences [J]. Ann Bot, 96: 489-498.

XIE ZQ, 2000. Eco-geographical distribution of the species from *Rheum* L. Polygonaceae in China. China's Biodiversity conservation toward the 21^{st} Century [M]. Beijing: China Forest Press: 230-236.

YANG Z, NIELSEN R, 2002. Codon-substitution models for detecting molecular adaptation at individual sites along specific lineages [J]. Mol Biol Evol, 19: 908-917.

YANG Z, WENDY SWW, RASMUS N, 2005. Bayes empirical bayes inference of amino acid sites under positive selection [J]. Mol Biol Evol, 22: 1 107-1 118.

YANG Z, 2007. PAML 4: a program package for phylogenetic analysis by maximum likelihood [J]. Mol Biol Evol, 24: 1 586-1 591.

ZAPATA JM, GUERA A, ESTEBAN-CARRASCO A, et al, 2005. Chloroplasts regulate leaf senescence: delayed senescence in transgenic *ndh*F-defective tobacco [J]. Cell Death Diff, 12: 1 277-1 284.

ZOHARY M, 1966. Flora Palaestina. Vol. I [M]. Jerusalem: The Israel Academy of Sciences and Humanities.