

凤尾蕨科植物 *rbcL* 基因的适应性进化分析

森 林^{1,3}, 苏应娟², 张 冰², 王 艇^{1*}

(1. 中国科学院武汉植物园, 武汉 430074; 2. 中山大学生命科学学院, 广州 510275; 3. 中国科学院研究生院, 北京 100049)

摘要: 为深入理解蕨类植物辐射式物种分化的分子适应机制, 在时间框架下, 采用位点模型和分支-位点模型对凤尾蕨科植物 *rbcL* 基因的进化式样进行了分析。通过比较模型 M1a/M2a 和 M7/M8, 在氨基酸水平上共鉴定出 6 个正选择位点: 149I、251M、255V、282F、359S 和 375F, 其中位点 282F 对维持 Rubisco 功能有重要作用。分别检验凤尾蕨科的附生分支和水蕨类分支发现, 前者不具适应性进化位点, 而后者有两个位点(230A 和 247C)经历正选择。相对于荫蔽的光条件, 水生生境可能对 RbcL 亚基的选择作用更强。另外, 基于 UCLD 分子钟模型估算出的凤尾蕨科各分支分化时间表明, 该科物种丰富度的辐射式增长发生在新生代渐新世, 推测古、始新世最热事件可能对物种分化的形成也产生一定作用。这对认识薄囊蕨类如何应对被子植物兴起导致的陆地生态系统改变具有重要意义。

关键词: 凤尾蕨科; *rbcL* 基因; 核酮糖-1,5-二磷酸羧化酶/加氧酶; 放松分子钟模型; 正选择位点

中图分类号: Q349

文献标识码: A

文章编号: 1005-3395(2010)01-0001-08

Adaptive Evolution of the *rbcL* Gene in Pteridaceous Ferns

SEN Lin^{1,3}, SU Ying-juan², ZHANG Bing², WANG Ting^{1*}

(1. Wuhan Botanical Garden, Chinese Academy of Sciences, Wuhan 430074, China; 2. School of Life Sciences,

Sun Yat-Sen University, Guangzhou 510275, China; 3. Graduate University of Chinese Academy of Sciences, Beijing 100049, China)

Abstract: To further understand the molecular adaptation linked to fern radiation, evolutionary patterns of the *rbcL* sequences in the family Pteridaceae were examined using random-site and branch-site models along with the estimated time-scale of pteridaceous phylogeny. By comparing Models M1a/M2a and M7/M8 under random-site models, six amino acid sites (149I, 251M, 255V, 282F, 359S and 375F) were found to be positively selected. The site 282F plays an important role in maintaining Rubisco function. Under branch-site models, no adaptive site was identified in the epiphytic clade, whereas two sites (230A and 247C) were detected in the ceratopteridoid clade. In comparison with shade conditions, aquatic habitats may have imposed stronger selection on the pteridaceous RbcL subunits. Moreover, the phylogenetic tree reconstructed with uncorrelated lognormal distributed relaxed clock model shows that the pteridaceous fern radiations occurred during the Oligocene, suggesting possible roles played by the Paleocene-Eocene Thermal Maximum. These results provide new insights into how leptosporangiate ferns response to the terrestrial ecosystem changes caused by the rise of angiosperms.

Key words: Pteridaceae; *rbcL* gene; Ribulose-1,5-bisphosphate carboxylase/oxygenase; Relaxed molecular clock model; Positive selection

现存的蕨类植物约有 12 000 种, 其中薄囊蕨类约 9 000 种, 占有所有蕨类的四分之三^[1]。研究表

明, 蕨类植物起源于泥盆纪(距今约 3.59 亿年前), 曾经历过 3 次“物种辐射”^[2-3]。第一次辐射发生在

收稿日期: 2009-03-16 接受日期: 2009-08-09

基金项目: 中国科学院“百人计划”引进国外杰出人才项目(0729281F02); 湖北省自然科学基金青年杰出人才项目(0631061H01); 国家自然科学基金项目(30011597)资助

* 通讯作者 Corresponding author, email: tingwang@wbcas.cn

泥盆纪至石炭纪,产生了 6 个现已灭绝的科;第二次见于古生代晚期至中生代早期(距今约 2.51 亿年前);而第三次辐射则出现在白垩纪晚期(距今约 6500 万年前)。蕨类的第三次辐射使我们对其进化潜力有了新的认识:即在白垩纪被子植物兴起后,蕨类植物非但没有衰退反而通过适应环境使物种丰富度显著增加。对水龙骨科附生蕨类的分析表明,它们的确可借助功能基因的适应性进化对被子植物引发的陆地生态系统变化作出响应^[4]。但是,其它薄囊蕨类是否也会如此,尚需更多验证。

凤尾蕨科(Pteridaceae)含 50 多属,1 000 余种,约占薄囊蕨类的 10%^[5-6],分布于潮湿的热带或干旱地区,占据陆生、岩生、附生和泽生等多种生态位^[6]。考虑到这些生境的环境因子(包括光条件)有明显差异,推测凤尾蕨类植物中在光合作用中起关键作用的一些功能基因,极有可能伴随物种的辐射式分化而发生适应性进化。

核酮糖-1,5-二磷酸羧化酶/加氧酶(Rubisco, EC 4.1.1.39)是植物参与光合作用的关键酶,其大亚基由叶绿体 *rbcL* 基因编码^[7]。它具有双重功能,能催化两种反应:一是羧化反应;二是加氧反应^[8-9]。Rubisco 既是光合作用固定 CO₂ 的限速酶,又是 CO₂ 和 O₂ 相互竞争的场所。Rubisco 固定 CO₂ 的活性位点就位于大亚基,尤其是其 C 末端的氨基酸有十分重要的功能^[10]。若增加 CO₂ 和 Rubisco 大亚基的亲合力,就能显著提高光合作用的生产力。对日本眼子菜属(*Potamogeton*)水生植物的研究表明,环境因素可直接影响 *rbcL* 基因的适应性进化^[11]。

非同义替换率(d_N)和同义替换率(d_S)的比值 ω 可用来测度蛋白质编码序列所受的选择压力,并判断选择压力对非同义突变的固定是起助推还是阻碍作用^[12]。如果非同义突变有害,那么净化选择就会降低这些突变的固定速率,从而有 $d_N < d_S$, $\omega < 1$ 。如果选择对适合度无影响,那么非同义突变和同义突变的固定速率相同,因而 $d_N = d_S$, $\omega = 1$ 。但是,如果发生的非同义突变有利于选择,那么它们会以比同义突变更快的速率被固定,致使 $d_N > d_S$, $\omega > 1$ 。所以,非同义替换率比同义替换率统计上显著偏高,能被当作蛋白质发生适应性进化的证据。此外,“分子钟”的研究也取得重要进展,新近建立的模型已无需假设核苷酸替换率在不同时间

段下保持恒定^[13],这为在时间框架下探讨分子适应性进化提供了契机。

本研究针对凤尾蕨科植物的 *rbcL* 基因展开进化分析。希望探讨凤尾蕨科 *rbcL* 基因是否经受正选择以及正选择发生的时间和蕨类的第几次物种辐射相关联;并鉴别凤尾蕨科 *rbcL* 基因的哪些位点在适应不同生境时起作用,为了解蕨类的分子适应机制提供重要参考。

1 材料和方法

1.1 序列数据

由 GenBank 获得凤尾蕨科 47 种植物以及外类群 8 种植物的 *rbcL* 基因序列(表 1)。使用 ClustalW^[14] 软件进行序列比对并进行手工校正。比对得到的数据矩阵中,每条序列包括 436 个密码子。为便于比较,密码子(氨基酸)位点的编号参照台湾水龙骨(*Polypodiodes formosana*) *rbcL* 基因序列进行^[15]。

1.2 统计分析

运行 Modeltest 3.7^[16] 软件选取核苷酸进化模型。采用 MrBayes 3.1.2^[17] 软件经贝叶斯途径推测系统发育关系。后验概率依据 Metropolis-Hastings-Green 算法通过 4 条链(Markov Chain Monte Carlo, MCMC)运行 15 000 000 代估计。MCMC 分析以随机树起始,每 100 代保存 1 棵树。使用 Tracer v1.4.1 软件^[18] 检验模型运行程度;当所有参数的有效取样大小(ESS, efficient sampling size)值均大于 500 时,认为迭代运算达到收敛。摒弃老化样本(burnin = 37 500)后,利用剩余样本构建一致树。一致树的绘制采用 Figtree v1.2.2 软件^[19]。

利用 BEAST 软件包中的 UCLD (Uncorrelated Lognormal Distributed Relaxed Clock Model)分子钟模型估算系统树上每个节点的分歧时间^[20]。依获得的贝叶斯法树将所研究植物分别归入 5 个类群集。基于最近共同祖先时间(tMRCA)值校正主要分支的分歧时间。采用 Modeltest 3.7 选取的 GTR + G + I 模型,迭代计算 30 000 000 代,每 1 000 代保存 1 株样本;Tracer v1.4.1^[21] 检测收敛程度。最初的 7 500 株被当作老化样本去除,用剩余样本重建时间尺度下一致树。

表 1 植物材料及其 *rbcL* 基因 GenBank 登录号Table 1 Plant materials and their Genbank accession numbers of *rbcL* genes

植物 Species	GenBank 登录号 GenBank accession No.	植物 Species	GenBank 登录号 GenBank accession No.
<i>Monogramma graminea</i>	EF452157	<i>Ochropteris pallens</i>	EF452160
<i>Monogramma</i> sp.	EF452156	欧洲凤尾蕨 <i>Pteris cretica</i>	EF452170
<i>Haplopteris elongata</i>	EF452153	井栏边草 <i>Pteris multifida</i>	EF452171
<i>Antrophyum latifolium</i>	EF452138	<i>Pteris tremula</i>	EF452174
<i>Ananthacorus angustifolius</i>	U20932	<i>Pteris argyraea</i>	EF452169
假鞭叶铁线蕨 <i>Adiantum malesianum</i>	EF452132	<i>Pteris quadriaurita</i>	EF452173
细叶铁线蕨 <i>Adiantum venustum</i>	EF452136	<i>Neurocallis praestantissima</i>	EF452158
<i>Adiantum peruvianum</i>	EF452133	<i>Eriosorus cheilanthoides</i>	EF452152
<i>Argyrosma limitanea</i>	EF452139	<i>Jamesonia blepharum</i>	EF452154
<i>Pellaea intermedia</i>	EF452163	<i>Jamesonia verticalis</i>	EF452155
<i>Pellaea truncata</i>	EF452164	<i>Pterozonium brevifrons</i>	EF452175
<i>Pellaea andromedifolia</i>	EF452162	<i>Anogramma chaerophylla</i>	AY168712
欧洲金毛裸蕨 <i>Paragymnopteris marantae</i>	EF452161	<i>Pityrogramma calomelanos</i>	EF452166
<i>Astrolepis sinuata</i>	EF452141	<i>Pityrogramma jamesonii</i>	EF452167
<i>Cheilanthes viridis</i>	EF452147	<i>Cosentinia vellea</i>	AY168720
银粉背蕨 <i>Aleuriopteris argentea</i>	EF452137	<i>Acrostichum danaeifolium</i>	EF452129
旱蕨 <i>Cheilanthes nitidula</i>	EF452146	<i>Ceratopteris richardii</i>	EUB52297
<i>Cheilanthes micropteris</i>	EF452145	全缘凤丫蕨 <i>Coniogramme fraxinea</i>	AM177359
<i>Cheilanthes alabamensis</i>	EF452143	<i>Cryptogramma crista</i>	EF452148
<i>Cheilanthes eatonii</i>	EF452144	沼泽蕨 <i>Thelypteris palustris</i>	U05947
<i>Adiantopsis radiata</i>	EF452131	<i>Cystopteris reevesiana</i>	EF452149
箭叶黑心蕨 <i>Doryopteris sagittifolia</i>	EF452151	<i>Blechnum occidentale</i>	U05910
<i>Pentagramma triangularis</i>	EF452165	半边铁角蕨 <i>Asplenium unilaterale</i>	EF452140
<i>Notholaena aschenborniana</i>	EF452159	<i>Nephrolepis cordifolia</i>	U05933
<i>Bommeria hispida</i>	EF452142	稀子蕨 <i>Monachosorum henryi</i>	U05932
戟叶黑心蕨 <i>Doryopteris ludens</i>	EF452150	食蕨 <i>Pteridium esculentum</i>	U05940
<i>Actinopteris dimorpha</i>	EF452130	阔叶鳞盖蕨 <i>Microlepia platyphylla</i>	U18642
<i>Platyozoma microphyllum</i>	AY168721	台湾水龙骨 <i>Polypodiodes formosana</i>	AB043100

使用位点模型^[21-22]和分支-位点模型^[23-24]进行分子进化分析。位点模型假设系统树的不同分支所受选择压力相同,但不同的氨基酸位点经历的选择压力各异。模型 M0(单一比率)假定所有位点具有相同的 ω 值。模型 M1a(近中性)假设蛋白质序列仅有保守位点($0 < \omega < 1$)和中性位点($\omega = 1$),而没有正选择位点($\omega > 1$)存在,并估算这两类位点的比率(p_0 及 p_1)和 ω 值(ω_0 及 ω_1)。与模型 M1a 不同,模型 M2a(正选择)假设序列不仅存在保守位点与中性位点,还存在处于正选择压力下的位点($\omega > 1$),估算获得这 3 类位点的比例(p_0 、 p_1 及 p_2)和 ω 值(ω_0 、 ω_1 及 ω_2)。模型 M3(离散)假定不同位点间的 ω 值不同,根据估测的 ω 值将所有位点分为保守位点($0 < \omega < 1$)、中性位点($\omega = 1$)和自由参数位点

($\omega > 1$) 3 类,它们所占比率分别为 p_0 、 p_1 及 p_2 ,对应的 ω 值为 ω_0 、 ω_1 及 ω_2 。模型 M7(beta)假设位点的 ω 值在 0 至 1 之间,满足 beta 分布。模型 M8(beta 和 ω) 在 M7 基础上增加一类位点,其比例由数据估计得出,位点的 ω 取值可大于 1。模型 M8a (beta 和 $\omega = 1$) 同 M8 类似,只是将 ω_1 值固定为 1。两两比较模型 M1a 和 M2a、模型 M7 和 M8 以及模型 M8 和 M8a 即可检验是否存在正选择位点。

分支-位点模型将位点分为 4 类:I 类位点在所有分支中高度保守,其 d_N/d_S (ω_0) 值较小;II 类位点经受中性或近中性选择,它们的 d_N/d_S (ω_1) 值接近 1 或者略小于 1;III 类位点在背景分支中高度保守 (ω_0),但前景分支比率(ω_2)显著大于 1;IV 类位点在背景分支经历中性进化或所受选择压力较小(ω_1),

但前景分支比率 ω_2 显著大于 1。采用分支-位点模型 A 的检验 $2^{[25]}$ 鉴别目标前景分支的正选择位点。以上进化分析均使用 PAML4 软件^[25] 完成。

将推测的氨基酸序列提交瑞士生物信息研究所(European Bioinformatics Institute: <http://www.isb-sib.ch/>) Swiss-model, 基于同源建模原理预测全缘凤丫蕨 RbcL 大亚基的三维结构。

2 结果

2.1 凤尾蕨科的系统发育和分化时间

利用贝叶斯法重建了凤尾蕨科的系统发育关系(图 1)。凤尾蕨科明显可分为两大分支(后验概率为 1.00): 第 1 支包括铁线蕨类(AD)和碎米蕨类

(CH), 互为姐妹群(后验概率为 1.00); 第 2 支由珠蕨类(CR)、凤尾蕨类(PT)和水蕨类(CE)组成(后验概率为 1.00), 其中珠蕨类首先分歧出来(后验概率为 0.84)。本研究所得贝叶斯树(图 1)的分支结构同以前报道的基本一致^[6]; 但以前认为珠蕨类是凤尾蕨科最基部类群。

为准确估计隶属凤尾蕨科的附生蕨类的分歧时间, 对图 1 中凤尾蕨科、铁线蕨类、碎米蕨类及凤尾蕨类等 4 个节点的分歧时间^[1] 进行多点校对。水蕨类卤蕨属(*Acrostichum*)的祖先化石证据也被用于时间校正^[26]。结果除戟叶黑心蕨(*Doryopteris ludens*)和秘鲁铁线蕨(*Adiantum peruvianum*)外, 其它植物的分化均发生在新生代以后。

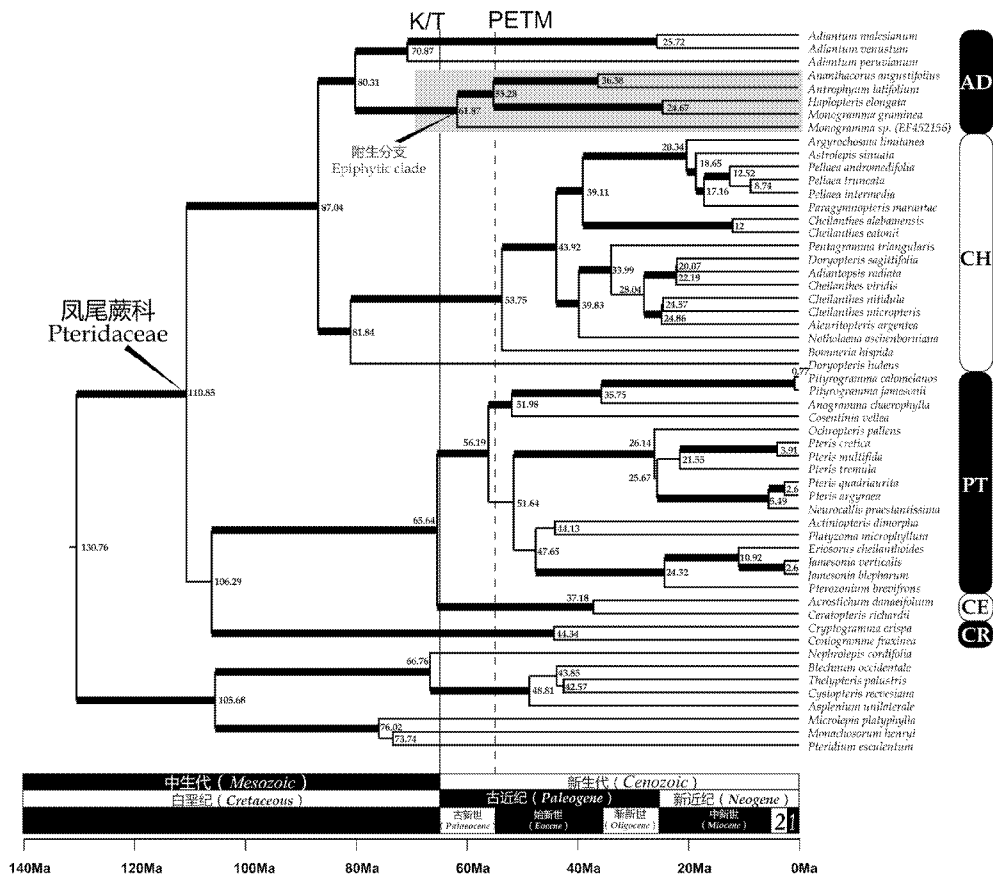


图 1 基于凤尾蕨科 *rbcL* 基因序列数据和 UCLD 分子钟模型构建的系统发育树

Fig. 1 The phylogenetic tree of *rbcL* gene sequences of Pteridaceae using uncorrelated lognormal distributed relaxed clock model

凤尾蕨科的 5 个主要分支, AD 为铁线蕨类, CH 为碎米蕨类, PT 为凤尾蕨类, CE 为水蕨类, CR 为珠蕨类。粗线表示该分支后验概率大于 0.9。灰色方框表示附生分支所处位置。系统树下端为地质时间表, 其中 1 为更新世, 2 为上新世。竖实线标出白垩纪与第三纪分界线, 竖虚线标出古、始新世最热事件。分支节点处数字为估测的分歧时间(单位: 百万年)。Five major clades in Pteridaceae family are indicated (black and white boxes): Adiantoid clade (AD); Cheilanthe clade (CH); Pteridoid clade (PT); Ceratopteridoid clade (CE); and Cryptogrammoid clade (CR). Heavily thickened lines indicate posterior probability $\geq 90\%$. Epiphytic clade is demonstrated in grey box. Geologic timescale is labelled below the phylogenetic tree. The epoch 1 and 2 stand for Pleistocene and Pliocene, respectively. The Cretaceous/Tertiary boundary (K/T) and Paleocene/Eocene thermal maximum (PETM) are showed with solid and dashed vertical line, respectively. Estimated divergence time of each node under UCLD are illustrated beside them (Unit: Million years).

2.2 凤尾蕨科 *rbcL* 基因正选择位点的鉴定

针对凤尾蕨科植物 *rbcL* 基因数据集,求算了不同模型对应的对数似然值和参数估计值。模型 M2a(选择)、M3(离散)及 M8(beta 和 ω)的假定条件中均允许 ω 比值大于 1,它们均提示 *rbcL* 基因存在正选择位点(表 2)。同它们各自对应的原假设模型 M1a(近中性)、M0(单一比率)及 M7(beta)相比,这些模型均得到显著改进,更加符合所分析数据(表 3)。例如,模型 M2a 比 M1a 多两个参数,它们的 LRT(likelihood ratio test)统计量 $2\Delta\ell = 16.23$,明显大于当 d.f.=2、 $P=5\%$ 时 χ^2 分布临界值 5.99。模型 M8a 的似然值 $\ell = -9550.67$,和模型 M8 相比 $2\Delta\ell = 2 \times (-9536.48 + 9550.67) = 28.38$;因为 χ^2 分布(d.f. = 1、 $P=5\%$)临界值为 3.84,所以模型 M8a 被极显著拒绝。

在 95% 水平上,模型 M2a 鉴定出 3 个氨基酸位点(149I, 251M 和 282F)受到正选择;而模型 M8 鉴定出有 6 个氨基酸位点(149I, 251M, 255V, 282F, 359S 和 375F)受正选择(表 2)。

对凤尾蕨科的附生分支,利用分支-位点模型

进行检验(图 1),结果显示在后验概率大于 80% 的水平,虽有 4 个位点(226V、243T、348S 和 367G) ω 值大于 1,但在后验概率大于 95% 水平却没有发现正选择位点(表 2)。与之不同,对水蕨类分支进行检验,结果在后验概率大于 95% 的水平上,发现有两个位点(230A 和 247C)经历正选择,其中 230A 的后验概率超过 99% (表 2)。

2.3 凤尾蕨科 RbcL 亚基正选择位点的空间位置

根据 Swiss-Model 数据库的比对结果,绘制 RbcL 大亚基各位点的空间位置。Rubisco 大亚基的羧基末端拥有一个 α/β 桶状结构域,其活性中心由一个大亚基的 α/β 桶状结构域同另一个大亚基的氨基末端结构域组成^[27-28]。由凤尾蕨科 *rbcL* 基因鉴定出的 6 个正选择位点,除位点 149I, 359S 和 375F 外,其余 3 个位点 251M, 255V 和 282F 均位于 α/β 桶状结构域的底部(图 3)。149I 位于 α 螺旋一侧,359S 位于 loop 环上,375F 则位于大亚基内侧的 α 螺旋上。对于水蕨类分支 *rbcL* 基因鉴定出的 2 个正选择位点,230A 位于 α 螺旋外侧,247C 也位于 α/β 桶状结构域底部。

表 2 各类位点间可变 ω 值模型下的参数估计值和对数似然值

Table 2 Parameter estimates and log-likelihood values under models of variable ω ratios among sites

模型 Models	p^*	ℓ	参数估计值 Estimated value of parameters	正选择位点 Positive selection sites
M0 单一比值 One ratio	1	-9930.37	$\omega = 0.056$	无 None
M1a 近中性 Near neutral	2	-9577.32	$p_0 = 0.934, \omega_0 = 0.011; p_1 = 0.066, \omega_1 = 1$	不允许 Not allowed
M2a 选择 Positive selection	4	-9569.21	$p_0 = 0.934, \omega_0 = 0.011; p_1 = 0.056, \omega_1 = 1;$ $p_2 = 0.009, \omega_2 = 2.376$	$P > 99\% : 282F; P > 95\% : 149I, 251M$
M3: 离散 Discrete	5	-9536.75	$p_0 = 0.895, \omega_0 = 0.004; p_1 = 0.086,$ $\omega_1 = 0.348; p_2 = 0.019, \omega_2 = 1.858$	$P > 99\% : 149I, 251M, 255V, 282F, 359S, 375F$
M7: beta	2	-9552.35	$p = 0.018, q = 0.125$	不允许 Not allowed
M8: beta 和 ω beta & ω	4	-9536.48	$p_0 = 0.982, p = 0.052, q = 0.992$ $p_1 = 0.018, \omega = 1.919$	$P > 99\% : 149I, 251M, 255V, 282F, 359S, 375F$
分支-位点模型 Branch site models				
附生分支 Epiphytic clade				
模型 A (ω_2 固定为 1)	119	-12175.22	$p_0 = 0.918, \omega_0 = 0; p_1 = 0.061, \omega_1 = 1;$ $p_2 + p_3 = 0.03, \omega_2 = 1$	不允许 Not allowed
模型 A ($\omega_2 = 1$, fixed)				
模型 A (ω_2 为估计值)	120	-12175.42	$p_0 = 0.909, \omega_0 = 0; p_1 = 0.055, \omega_1 = 1;$ $p_2 + p_3 = 0.03, \omega_2 = 1$	$P > 80\% : 226V, 243T, 348S, 367G; P > 50\% : 32L, 219L, 226V, 232V, 243T, 276S, 348S, 358Y, 367G$
模型 A (ω_2 estimated)				
水蕨类分支 Ceratopteridoid clade				
模型 A (ω_2 固定为 1)	119	-12185.06	$p_0 = 0.921, \omega_0 = 0.015; p_1 = 0.056, \omega_1 = 1;$ $p_2 + p_3 = 0.022, \omega_2 = 1$	不允许 Not allowed
模型 A ($\omega_2 = 1$, fixed)				
模型 A (ω_2 为估计值)	120	-12183.02	$p_0 = 0.931, \omega_0 = 0.015; p_1 = 0.057, \omega_1 = 1;$ $p_2 + p_3 = 0.012, \omega_2 = 4.03$	$P > 99\% : 230A; P > 95\% : 230A, 247C; P > 50\% : 31T, 226N, 230A, 247C, 359K$
模型 A (ω_2 estimated)				

p^* 表示 ω 分布所用的参数数目。氨基酸位点编号参照台湾水龙骨 *rbcL* 基因序列^[15]。 p^* stands for the number of parameters in different models. The site No. is associated with the *rbcL* sequence of *Polypodiodes formosana*^[15].

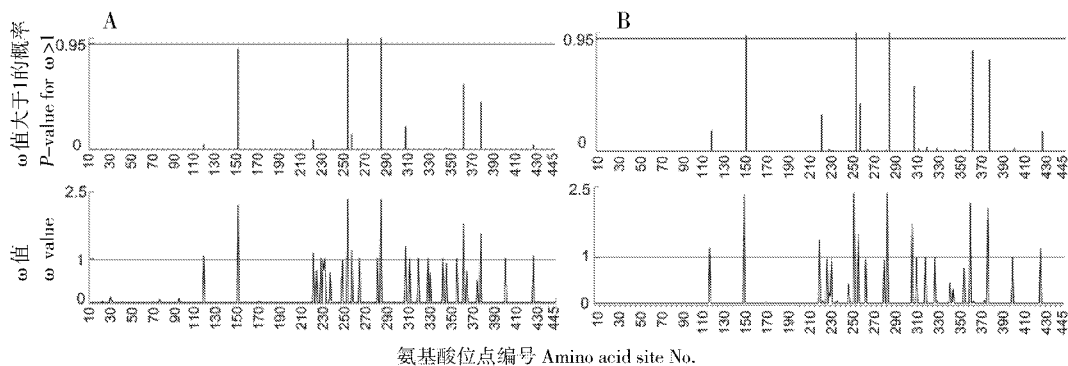


图 2 模型 M2a 所估测的参数值

Fig. 2 Estimated values of the parameters under Model M2a

氨基酸位点编号参照台湾水龙骨 *rbcL* 基因序列^[15]。图 A、B 分别表示基于 NEB 和 BEB 方法测算的结果。The site No. are referenced to the *rbcL* sequence of *Polypodiodes formosana*^[15]. A. Naïve Empirical Bayes (NEB); B. Bayes Empirical Bayes (BEB).

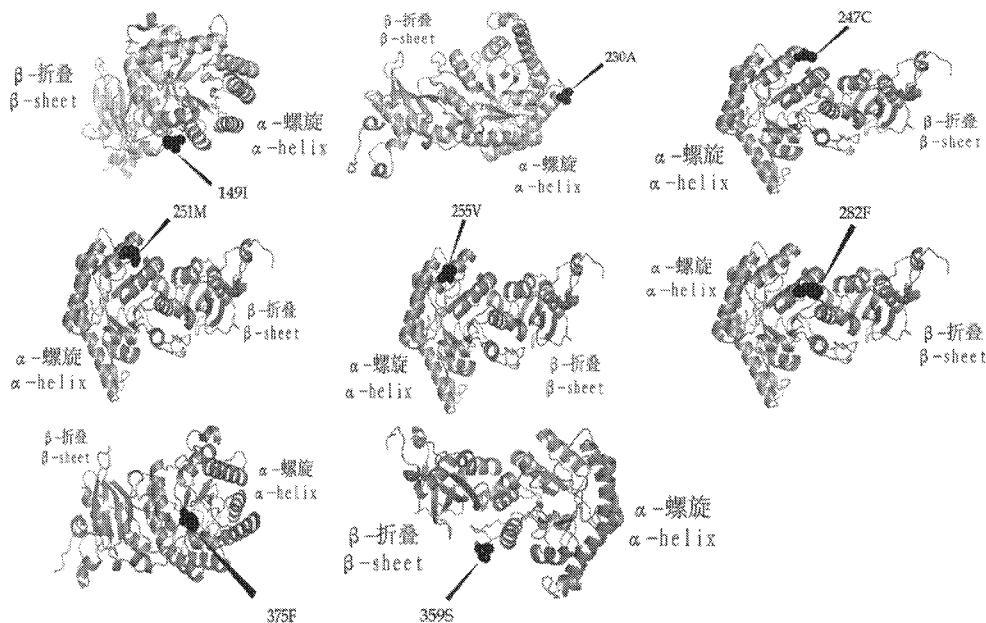


图 3 全缘凤丫蕨(*Coniogramme fraxinea*) RbcL 大亚基正选择位点的空间位置

Fig. 3 Spatial locations of the positively selected sites in the RbcL subunit of *Coniogramme fraxinea*
黑色箭头所指为正选择位点。Amino acid sites under positive selection are labeled with black arrows.

表 3 似然比值检验统计量 ($2\Delta\ell$)

Table 3 Likelihood ratio statistics ($2\Delta\ell$)

模型比较 Comparison of Models	$2\Delta\ell$	d.f.	$\chi^2_{5\%}$
M0-M3	787.25	4	9.49
M1a-M2a	16.23	2	5.99
M7-M8	31.73	2	5.99
M8a-M8	28.38	1	3.84
附生分支 Epiphytic clade			
模型 A Model A	0.2	1	3.84
水蕨类分支 Ceratopteridoid clade			
模型 A Model A	4.08	1	3.84

3 讨论

本研究显示,凤尾蕨科的分化最早发生于距今约 1.1 亿年前的中生代白垩纪,然而直至白垩纪结束(距今约 6 500 万年前)该科的分化仍然很少(图 1)。但是,在随后的新生代渐新世,凤尾蕨科的分化开始明显改变,物种丰富度呈现辐射式增长。这些结果支持蕨类面对被子植物兴盛有能力发生多样化的观点^[29],可是也提示某些类群(例如凤尾蕨科)的适应过程可能长达数千万年,而且古、始新世最热事件所起的作用也不容忽视。此外,我们重建的凤尾蕨科贝叶斯树在珠蕨类的位置上与

Schuettpelz 等^[6]报道不同,可能是构树用的数据集不同所致。

由凤尾蕨科 *rbcL* 基因推导出的氨基酸序列,通过比较模型 M1a/M2a 和 M7/M8 分别鉴定出 3 个和 6 个正选择位点。在这些位点中,除位点 149I, 359S 和 375F 外,其余 3 个位点 251M, 255V 和 282F 均位于 α/β 桶状结构域的底部(图 3)。251M 含有一硫原子,能够与附近位点形成二硫键,通过范德华力维持 Rubisco 酶催化核心的结构稳定。有研究显示 *rbcL* 基因上甲硫氨酸的突变会显著影响酶的 O₂ 结合量以及同 CO₂/O₂ 结合的特异性,改变催化效率^[30]。另一个正选择位点 255V 的侧链较短,由一个异丙基构成。此较小的侧链,可使桶状结构留有更大内部空间,以保证催化的进行^[31]。结构生物学结果表明,282F 位点能借助范德华作用,以位点 283 和 284 为中介与位点 258 和 265 发生作用,进而影响 Rubisco 的功能^[31-32]。至于其它 3 个鉴定出的正选择位点(149I, 359S 和 375F),它们并不位于酶活性中心区域,推测是在亚基间相互结合时起作用。

分支-位点模型的检验结果显示,凤尾蕨科附生分支的 *rbcL* 基因并无适应性位点,而水蕨类分支则含两个统计上显著的正选择位点(230A 和 247C)。230A 位于 RbcL 和 RbcS 的结合面,247C 处在 Rubisco 活性中心的 α/β 桶状结构域。半胱氨酸(C)的巯基具还原性,可借氢键同其它位点结合以维持酶的催化作用。对眼子菜属的 *rbcL* 基因研究表明,该属植物在 245 与 249 位存在替换位点,这两个位点同样处在 α/β 桶状结构域^[11]。这些结果暗示,位点 247 的有益替换可以提高 Rubisco 的活性。另外,我们的结果还表明,凤尾蕨科植物面对被子植物引发的生境改变,其栖居于不同生态位的分支在特定功能基因的适应性进化上会存在差异。综合水蕨类和眼子菜的分析结果,可以认为相对于荫蔽的光条件,水生生境似乎对 RbcL 亚基的选择作用更强。继续对其它光合作用相关基因进行检测,将有助于加深对这一问题的理解。

以前鉴定 Rubisco 大亚基中影响催化效率的位点,主要采用的是对模式植物进行突变体筛选、突变体回复和定向诱变的技术路线^[33-37]。这种技术路线的优点是能获得直接的经验证据,但也存在实验周期长、成本高以及目标位点欠明确等问题。近来,将实验生物学手段同计算分子进化的分析方

法相结合的研究途径已越来越受到关注。本研究利用位点间可变 ω 比值模型和分支-位点模型分别鉴定出 6 个和 2 个正选择位点,其中一个有明确的实验证据支持它是同 Rubisco 的稳定性和动力学特性相关的重要位点。这再次提示,依托 ω 比值建立的 DNA 编码序列进化模型的确是检验分子适应的有效方法。在资源受限制的情况下,未来的定向诱变和遗传操作实验应该优先以鉴定的正选择位点为目标,从而加速阐明 Rubisco 催化机制的进程。

参考文献

- [1] Schuettpelz E, Pryer K M. Evidence for a Cenozoic radiation of ferns in an angiosperm-dominated canopy [J]. Proc Natl Acad Sci USA, 2009, 106: 11200-11205.
- [2] Lovis J. Evolutionary patterns and processes in ferns [J]. Adv Bot Res, 1977, 4: 229-415.
- [3] Rothwell G. Complex Paleozoic Filicales in the evolutionary radiation of ferns [J]. Amer J Bot, 1987, 74: 458-461.
- [4] Su Y J(苏应娟), Wang T(王艇). Adaptive evolution in the rubisco large subunits of epiphytic ferns (Polypodiaceae): Identification of positively selected sites [J]. Acta Sci Nat Univ Sunyatsen(中山大学学报:自然科学版), 2008, 47: 74-80.(in Chinese)
- [5] Smith A, Pryer K, Schuettpelz E, et al. A classification for extant ferns [J]. Taxon, 2006, 55: 705-732.
- [6] Schuettpelz E, Schneider H, Huiet L, et al. A molecular phylogeny of the fern family Pteridaceae: assessing overall relationships and the affinities of previously unsampled genera [J]. Mol Phylogenet Evol, 2007, 44: 1172-1185.
- [7] Spreitzer R, Salvucci M. Rubisco: Structure, regulatory interactions, and possibilities for a better enzyme [J]. Annu Rev Plant Biol, 2002, 53: 449-475.
- [8] Gutteridge S, Gatenby A. Rubisco synthesis, assembly, mechanism, and regulation [J]. Plant Cell, 1995, 7: 809-819.
- [9] Portis Jr A. Regulation of ribulose 1,5-bisphosphate carboxylase/oxygenase activity [J]. Annu Rev Plant Biol, 1992, 43: 415-437.
- [10] Curmi P M G, Cascio D, Sweet R M, et al. Crystal structure of the unactivated form of ribulose-1,5-bisphosphate carboxylase/oxygenase from tobacco refined at 2.0-Å resolution [J]. J Biol Chem, 1992, 267: 16980-16989.
- [11] Iida S, Miyagi A, Aoki S, et al. Molecular adaptation of *rbcL* in the heterophyllous aquatic plant *Potamogeton* [J]. PLoS ONE, 2009, 4: e4633.
- [12] Yang Z. Computational Molecular Evolution [M]. New York: Oxford University Press, 2006: 259-292.
- [13] Pybus O G, Rambaut A. Evolutionary analysis of the dynamics of viral infectious disease [J]. Nat Rev Genet, 2009, 10: 540-550.
- [14] Thompson J, Higgins D, Gibson T. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight

- matrix choice [J]. *Nucl Acids Res*, 1994, 22: 4673–4680.
- [15] Hirohara M, Nakane T, Terayama Y, et al. Chemotaxonomy of ferns: triterpenoids and *rbcL* gene sequences of *Polypodium*, *Polypodiodes* and *Goniophlebium* [J]. *Nat Med*, 2000, 54: 330–333.
- [16] Posada D, Crandall K. Modeltest: testing the model of DNA substitution [J]. *Bioinformatics*, 1998, 14: 817–818.
- [17] Huelsenbeck J, Ronquist F. MRBAYES: Bayesian inference of phylogenetic trees [J]. *Bioinformatics*, 2001, 17: 754–755.
- [18] Rambaut A, Drummond A J. <http://tree.bio.ed.ac.uk/software/tracer/2008>, Tracer v1.4.1 [CP]. 2008–11–29.
- [19] Rambaut A, Drummond A J. <http://tree.bio.ed.ac.uk/software/figtree/2009>, FigTree v1.2.2 [CP]. 2009–02–24.
- [20] Drummond A J, Rambaut A. BEAST: Bayesian evolutionary analysis by sampling trees [J]. *BMC Evol Biol*, 2007, 7: 214–221.
- [21] Nielsen R, Yang Z. Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene [J]. *Genetics*, 1998, 148: 929–936.
- [22] Yang Z, Bielawski J. Statistical methods for detecting molecular adaptation [J]. *Trends Ecol Evol*, 2000, 15: 496–503.
- [23] Yang Z, Wong W, Nielsen R. Bayes empirical bayes inference of amino acid sites under positive selection [J]. *Mol Biol Evol*, 2005, 22: 1107–1118.
- [24] Zhang J, Nielsen R, Yang Z. Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level [J]. *Mol Biol Evol*, 2005, 22: 2472–2479.
- [25] Yang Z. PAML 4: Phylogenetic analysis by maximum likelihood [J]. *Mol Biol Evol*, 2007, 24: 1586–1591.
- [26] Bonde S, Kumaran K. The oldest macrofossil record of the mangrove fern *Acrostichum* L. from the Late Cretaceous Deccan Intertrappean beds of India [J]. *Cretaceous Res*, 2002, 23: 149–152.
- [27] Knight S, Andersson I, Brändén C I. Crystallographic analysis of ribulose 1,5-bisphosphate carboxylase from spinach at 2.4 Å resolution. Subunit interactions and active site [J]. *J Mol Biol*, 1990, 215: 113–160.
- [28] Soper T, Mural R, Larimer F, et al. Essentiality of Lys-329 of ribulose-1,5-bisphosphate carboxylase/oxygenase from *Rhodospirillum rubrum* as demonstrated by site-directed mutagenesis [J]. *Protein Eng Des Sel*, 1988, 2: 39–44.
- [29] Schneider H, Schuettpelz E, Pryer K, et al. Ferns diversified in the shadow of angiosperms [J]. *Nature*, 2004, 428: 553–557.
- [30] Kreel N, Tabita F. Substitutions at methionine 295 of *Archaeoglobus fulgidus* ribulose-1,5-bisphosphate carboxylase/oxygenase affect oxygen binding and CO₂/O₂ specificity [J]. *J Biol Chem*, 2007, 282: 1341–1351.
- [31] Persson E, Bak H, Olsen O. Substitution of valine for leucine 305 in factor VIIa increases the intrinsic enzymatic activity [J]. *J Biol Chem*, 2001, 276: 29195–29199.
- [32] Andersson I. Large structures at high resolution: The 1.6 Å crystal structure of spinach ribulose-1,5-bisphosphate carboxylase/oxygenase complexed with 2-carboxyarabinitol bisphosphate [J]. *J Mol Biol*, 1996, 259: 160–174.
- [33] Hong S K, Spreitzer R J. Complementing substitutions at the bottom of the barrel influence catalysis and stability of ribulose-bisphosphate carboxylase/oxygenase [J]. *J Biol Chem*, 1997, 272: 11114–11117.
- [34] Du Y C, Peddi S R, Spreitzer R J. Assessment of structural and functional divergence far from the large subunit active site of ribulose-1,5-bisphosphate carboxylase/oxygenase [J]. *J Biol Chem*, 2003, 278: 49401–49405.
- [35] Chen Z, Spreitzer R. Chloroplast intragenic suppression enhances the low CO₂/O₂ specificity of mutant ribulose-bisphosphate carboxylase/oxygenase [J]. *J Biol Chem*, 1989, 264: 3051–3053.
- [36] Chen Z, Yu W, Lee J, et al. Complementing amino acid substitutions within loop 6 of the α/β-barrel active site influence the CO₂/O₂ specificity of chloroplast ribulose-1,5-bisphosphate carboxylase/oxygenase [J]. *Biochemistry*, 1991, 30: 8846–8850.
- [37] Zhu G, Spreitzer R. Directed mutagenesis of chloroplast ribulose-1,5-bisphosphate carboxylase/oxygenase. Loop 6 substitutions complement for structural stability but decrease catalytic efficiency [J]. *J Biol Chem*, 1996, 271: 18494–18498