

水龙骨科附生蕨类 Rubisco 大亚基的适应性进化: 正向选择位点的鉴定*

苏应娟¹, 王 艇²

(1. 中山大学生命科学学院, 广东 广州 510275; 2. 中国科学院武汉植物园, 湖北 武汉 430074)

摘 要: 核酮糖-1, 5-二磷酸羧化酶/加氧酶 (Rubisco, EC 4. 1. 1. 39) 在植物光合作用中起重要作用, 既负责碳同化又引发光呼吸。催化固定 CO₂ 的活性位点位于 Rubisco 大亚基, 由叶绿体 *rbcL* 基因编码。水龙骨科是现存蕨类中最为衍生的类群, 多为附生植物。为验证蕨类植物在白垩纪适应被子植物兴起而发生分化的观点, 本研究以水龙骨科附生蕨类为对象, 利用位点间可变 ω (非同义替换率 d_N 和同义替换率 d_S 的比值) 模型对其 *rbcL* 基因的适应性进化进行分析。通过比较模型 M1a/M2a 和模型 M7/M8, 在氨基酸水平上, 共鉴定出 7 个正向选择位点: 133L, 251M, 262A, 265V, 282Y, 326I 和 362I。其中位点 262A, 265V, 282Y 及 326I 对维持 Rubisco 功能的重要性也得到已发表实验数据的支持。这些结果一方面显示了基于 ω 比值检验 DNA 编码序列分子适应的有效性, 另一方面也提示水龙骨科类可借助 *rbcL* 等功能基因的适应性进化, 应对白垩纪被子植物引发的陆地生态系统改变。

关键词: 水龙骨科; 核酮糖-1, 5-二磷酸羧化酶/加氧酶; *rbcL* 基因; 正向选择位点

中图分类号: Q949; Q349 **文献标识码:** A **文章编号:** 0529-6579 (2008) 05-0074-07

核酮糖-1, 5-二磷酸羧化酶/加氧酶 (Rubisco, EC 4. 1. 1. 39) 可能是地球上最丰富的蛋白质^[1]。该酶具双重功能, 能催化两种反应: 一是 CO₂ 与核酮糖-1, 5-二磷酸 (RUBP) 共价结合生成不稳定的六碳化合物, 分裂出 2 分子 3-磷酸甘油酸; 二是 O₂ 同样和 RUBP 反应, 生成 1 分子 3-磷酸甘油酸和 1 分子 2-磷酸乙醇酸^[2-3]。因此, 在光合作用中, Rubisco 既是固定 CO₂ 的限速酶, 又是决定 CO₂ 和 O₂ 相互竞争的关键。陆生植物的 Rubisco 全酶为一包含 8 个大亚基和 8 个小亚基的 L₈S₈ 复合体, 大、小亚基分别由来自叶绿体 DNA 的 *rbcL* 基因和核 DNA 的 *rbcS* 基因编码^[4]。现已知道, Rubisco 催化固定 CO₂ 的活性位点位于大亚基, 尤其是大亚基的 C 末端在功能上起重要作用^[5]。如果能小幅度地增加 CO₂ 和 Rubisco 大亚基的亲合力, 那就可显著提高光合作用的生产力。

分析蛋白质序列的适应性进化能为识别重要的氨基酸结构和功能位点提供信息。在蛋白质水平上, 可以用非同义替换率 (d_N) 和同义替换率 (d_S) 的比值 ω ($= d_N/d_S$) 来测度选择压力, 进而判断自然选择对非同义突变的固定是起了助推还是阻碍作用^[6]。如果选择对适合度没有影响, 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, 如果非同义突变有害, 那么净化选择就会减低这些突变的固定速率, 从而有 $d_N < d_S$, $\omega < 1$ 。与之相反, 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, 可被当作蛋白质发生适应性进化的证据。Nielsen 和 Yang^[7], Yang 等^[8] 和 Wong 等^[9] 针对编码蛋白质的 DNA 序列, 先后建立、改进了允许不同氨基酸位点取不同 ω 值的一系列似然进化模型。经由这些模型验证存在正向选择位点后, 可进一步通过贝叶斯途径将发生正向选择的氨基酸位点加以鉴定。值得注意的是, Suzuki 和 Nei^[10-11] 曾指出上述似然法会给出假阳性结果。但是, Wong 等^[9] 基于模拟研究对似然法的可信度和统计效力给出有力回应, 并特别提到 Suzuki 和 Nei^[10] 的分析有误。Zhang^[12] 也质疑过 Yang 和 Nielsen^[13] 的分支-位点模型易产生假阳性, 但随后 Zhang 等^[14] 同 Yang 合作对该模型予以了改进, 很好地解决了假阳性问题。

最近的分类系统认为水龙骨科属现存蕨类中最为衍生的类群, 包括约 56 属、1 200 种^[15]。这些

* 收稿日期: 2008-05-28

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

作者简介: 苏应娟 (1965 生), 女, 副教授; 通讯联系人: 王艇; E-mail: tingwang@wbcas.cn

植物主要分布在泛热带,不仅形态性状变异极其多样,而且大多为附生植物。关于水龙骨科的物种分化,近来有学者通过将分子系统发育分析和利用化石资料对估计的分歧时间进行校正这两者相结合的方法,得出同以往截然不同的看法。他们指出水龙骨科的物种多样性和丰富度是在白垩纪被子植物兴起后发生的;也即被子植物在白垩纪的繁盛不仅没有导致水龙骨科的衰退,而且还为附生型水龙骨科的分化提供了更为复杂多样的生态条件,从而引发其丰富度呈“辐射”式增长^[16]。若接受这种观点,那就可以预期:水龙骨科中,一些同适应被子植物所形成的陆地生态系统相关的基因在分子水平也会发生适应性进化。鉴于 Rubisco 酶大亚基的重要催化活性以及 GenBank 中登录有大量的 *rbcL* 基因序列数据,本研究对水龙骨科附生蕨类的 *rbcL* 基因进行了适应性进化分析。目的是:①鉴定水龙骨科 Rubisco 大亚基受正向选择而发生适应性进化的氨基酸位点;②为水龙骨科在白垩纪发生“辐射”式物种分化的理论提供新的佐证。

1 材料和方法

1.1 序列数据

由 GenBank 获得 51 种蕨类植物的 *rbcL* 基因序列(表1)。使用 Clustal W 软件^[17]进行序列比对,并随后对结果施行手工校正。比对后得到的数据矩阵中,每条序列包含 406 个密码子。密码子(氨基酸)位点的编号参照 Hirohara 等^[18]报道的台湾水龙骨科(*Polypodioides formosana*) *rbcL* 基因序列。运行 Modeltest 3.7 软件^[19]选取核苷酸进化模型。采用 MrBayes 3.1.2 软件^[20]经贝叶斯途径推测系统发育关系。系统发育模型的后验概率依 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。摒弃老化样本(burnin = 37500)后由剩余样本构建一致树。

1.2 统计分析

采用位点间可变 ω 比值模型检验是否存在经受正向选择($\omega > 1$)的位点^[7]。模型 M0(单一比值)假定所有位点和分支具相同 ω 比值。模型 M1a(近中性)假定蛋白质具两类位点:一类为保守位点, $0 < \omega < 1$; 另一类为中性位点,其 $\omega = 1$ 。

模型 M2a(选择)增加第三类位点,其 ω 为自由参数,取值可大于 1。模型 M3(离散)利用三类位点的离散分布;三类位点对应的比例 p_0 , p_1 及 p_2 和 ω 比值 ω_1 , ω_2 及 ω_3 分别由数据估计得出。模型 M7(beta)利用 beta 分布 $B(p, q)$; 随参数 p 和 q 的改变,该分布可取不同形状。模型 M8(beta 和 ω) 在模型 M7 基础上增加一类位点,其比例和 ω 由数据估计得出,这样位点的 ω 取值就可大于 1。模型 M8a(beta 和 $\omega = 1$) 同 M8 类似,只是将 ω 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向选择位点^[7-8]。利用极大似然法估计出的模型参数值,再通过贝叶斯定律计算出每个氨基酸所属位点类别的后验概率,依据后验概率鉴定出正向选择位点。以上分析采用软件 PAML 4 完成^[21]。

2 结果

2.1 分子系统发育分析

利用贝叶斯法重建出蕨类植物水龙骨科的系统发育关系(图1)。奶囊蕨(*Thylacopteris papillosa*)独自构成一支,占据最基部位置;其次是由显脉蕨属(*Goniophlebium*)代表植物组成的一个单系分支(后验概率 1.00),它同由星蕨类聚成的核心分支(后验概率 0.82)彼此为姊妹群。星蕨类分支进一步分出两个得到很好支持的分支:一支包括所有的蚁蕨属(*Lecanopteris*)代表植物和部分星蕨属(*Microsorium*)植物(后验概率 0.80),另一支由线蕨属(*Colysis*)、伏石蕨属(*Lemmaphyllum*)、扇蕨属(*Neocheiropteris*)、瓦韦属(*Lepisorus*)、尖嘴蕨属(*Belvisia*)、丝带蕨属(*Drymotaenium*)、薄唇蕨属(*Leptochilus*)及星蕨属(*Microsorium*)植物组成(后验概率 0.81)。后者又分为两个亚分支,其中之一包含薄唇蕨属和星蕨属植物(后验概率 0.78);而另一亚分支包括线蕨属、星蕨属的显脉星蕨(*Microsorium zippelii*)、伏石蕨属、扇蕨属、瓦韦属、尖嘴蕨属及丝带蕨属的代表植物(后验概率 1.00),其中瓦韦属—尖嘴蕨属—丝带蕨属组成的分支得到较好支持(后验概率 0.73)。

2.2 正向选择氨基酸位点的鉴定

针对水龙骨科植物的 *rbcL* 数据集,求算出了不同模型相应的对数似然值和参数估计值。模型

表 1 植物材料及其 *rbcL* 基因 GenBank 登录号Tab. 1 Plant materials and their *rbcL* gene GenBank accession numbers

中文名	学名	GenBank 登录号
南海尖嘴蕨	<i>Belvisia mucronata</i> (Fée) Copel.	AY362562
	<i>Belvisia platyrhynchos</i> (Kunze) Copel.	DQ642152
新店线蕨	<i>Colysis shintenensis</i> (Hayata) H. Ito	CSU05612
丝带蕨	<i>Dryotaenium miyoshianum</i> (Makino) Makino	AY362563
	<i>Goniophlebium amoenum</i> (Mett.) Bedd. var. <i>amoenum</i>	DQ078626
	<i>Goniophlebium amoenum</i> var. <i>chinense</i> Ching	DQ078630
箭叶水龙骨	<i>Goniophlebium argutum</i> (Wall. ex Hook.) J. Sm. ex Hook.	DQ164442
台湾水龙骨	<i>Goniophlebium formosanum</i> (Baker) Rödl-Linder	AB043100
拟水龙骨	<i>Goniophlebium mengtzeense</i> (H. Christ) Rödl-Linder	AY362560
	<i>Goniophlebium percussum</i> (Cav.) Wagner & Grether	AY362561
	<i>Goniophlebium subauriculatum</i> (Blume) C. Presl	AF470342
	<i>Lecanopteris balgoyii</i> Hennipman	AF470328
	<i>Lecanopteris carnosus</i> Blume	AF470322
	<i>Lecanopteris celebica</i> Hennipman	AF470323
	<i>Lecanopteris crustacea</i> Copel.	AF470329
	<i>Lecanopteris deparioides</i> (Ces.) Baker	AF470324
	<i>Lecanopteris lomarioides</i> (Kunze ex Mett.) Copel.	AF470326
	<i>Lecanopteris luzonensis</i> Hennipman	AF470325
蚁蕨	<i>Lecanopteris mirabilis</i> (C. Chr.) Copel.	AF470330
	<i>Lecanopteris pumila</i> Blume	AF470331
	<i>Lecanopteris sinuosa</i> (Hook.) Copel.	AF470321
	<i>Lecanopteris spinosa</i> Jermy & Walker	AF470327
肉质伏石蕨	<i>Lemmaphyllum carnosum</i> (J. Sm. ex Hook.) C. Presl	AF470332
网眼瓦韦	<i>Lepisorus clathratus</i> (C. B. Clarke) Ching	DQ642154
	<i>Lepisorus excavatus</i> (Willd.) Ching	DQ642155
	<i>Lepisorus excavatus</i> (Willd.) Ching	DQ642156
长叶瓦韦	<i>Lepisorus longifolius</i> (Bl.) Holtt.	DQ642157
	<i>Lepisorus nudus</i> (Hook.) Ching	AY362564
瓦韦	<i>Lepisorus thunbergianus</i> (Kaulf.) Ching	LTU05629
薄唇蕨	<i>Leptochilus decurrens</i> Blume	AY096203
	<i>Microsorium commutatum</i> (Bl.) Copel.	AY362571
	<i>Microsorium cuspidatum</i> (D. Don) Tagawa	AF470335
	<i>Microsorium grossum</i> (Langsd. & Fisch.) S. B. Andrews	DQ179633
	<i>Microsorium membranifolium</i> (R. Br.) Ching	DQ642161
	<i>Microsorium musifolium</i> (Blume) Copel.	AF470333
	<i>Microsorium novo-zealandiae</i> (Baker) Copel.	DQ401116
	<i>Microsorium papuanum</i> (Baker) Parris	DQ642162
	<i>Microsorium punctatum</i> (L.) Copel.	DQ164444
	<i>Microsorium pustulatum</i> (G. Forst.) Copel.	DQ401117
	<i>Microsorium scandens</i> (G. Forst.) Tindale	DQ401118
	<i>Microsorium scandens</i> (G. Forst.) Tindale	DQ212057
	<i>Microsorium scolopendrium</i> (Burm. f.) Copel.	DQ642163
	<i>Microsorium scolopendrium</i> (Burm. f.) Copel.	DQ642164
	<i>Microsorium varians</i> (Mett.) Hennipman & Hett.	AY362566
	<i>Microsorium viellardii</i> (Mett.) Copel.	DQ179635
显脉星蕨	<i>Microsorium zippelii</i> (Blume) Ching	AB232411
扇蕨	<i>Neochiropteris palmatopedata</i> (Baker) H. Christ	AY362567
奶囊蕨	<i>Thylacopteris papillosa</i> (Blume) Krause ex J. Sm.	AY459174
鹿角蕨	<i>Platyserium stemaria</i> (Beauv.) Desv.	EF463257
槭叶石韦	<i>Pyrrosia polydactyla</i> (Hance) Ching	EF463259
葡萄石韦	<i>Pyrrosia serpens</i> (Forst.) Ching	EF463260

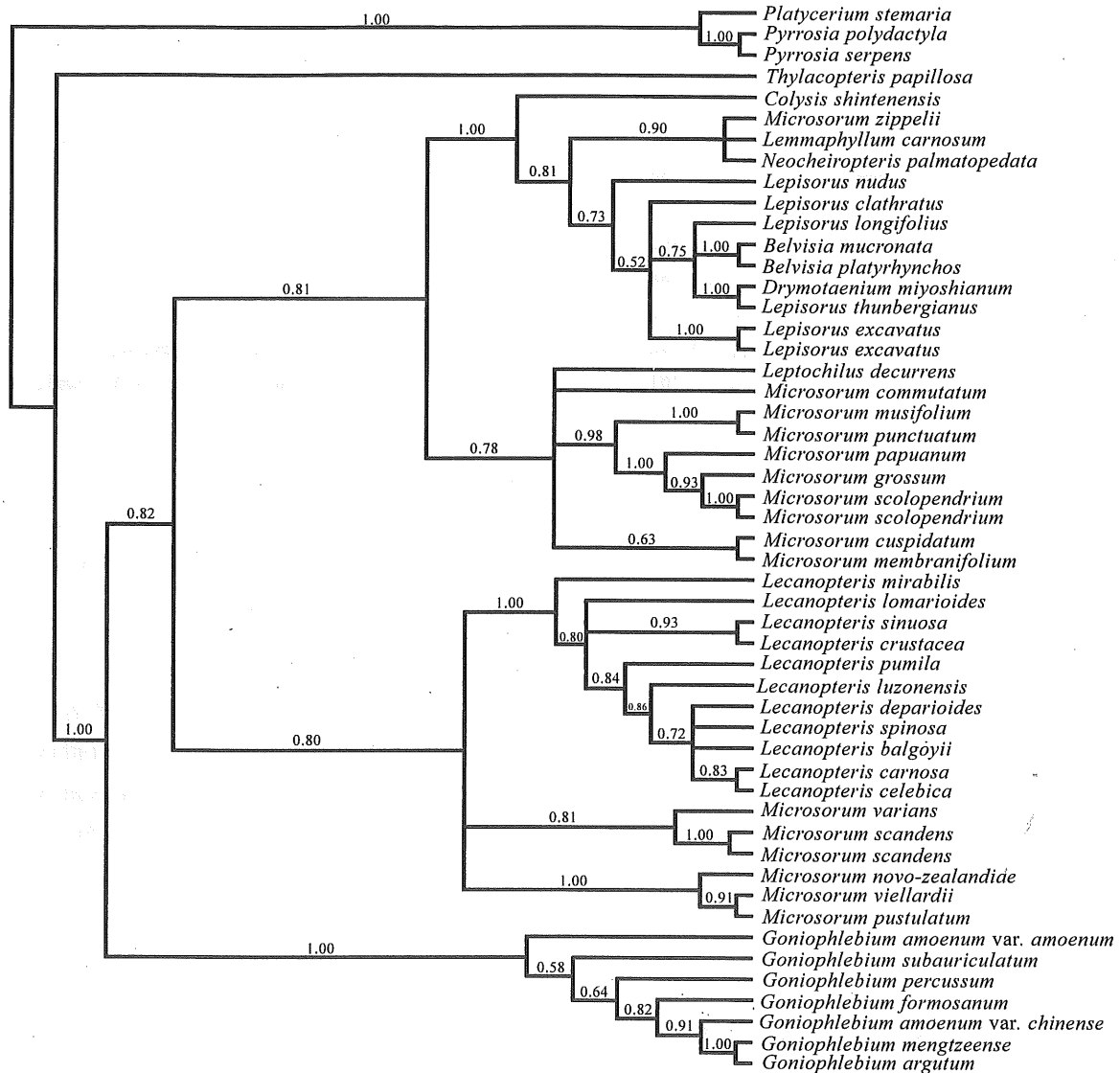


图 1 基于水龙骨科 *rbcL* 基因序列数据利用贝叶斯法得出的多数一致树 (分支上给出的为后验概率值)

Fig. 1 The majority consensus tree resulting from the Bayesian inference of Polypodiaceae *rbcL* gene sequences (Posterior probability values are given above branches)

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, 它们一律提示 *rbcL* 基因中存在受到正向选择的位点 (表 2)。同它们各自对应的空模型 M1a (近中性)、M0 (单一比值) 及 M7 (beta) 相比, 这些模型具显著改进, 明显更符合数据 (表 3)。例如, 模型 M2a 比 M1a 多两个参数, 它们的似然比值检验 (LRT) 统计量 $2\Delta l = 27.34$, 明显大于当 d. f. = 2、 $p = 1\%$ 时的 χ^2 分布临界值 9.21。需要指出的是, 比较模型 M0 和 M3 并不检验是否有位点受到正向选择; 它所检验的其实是各位点是否取不同的 ω 值。另外, 比较

模型 M8 和 M8a 有助于降低假阳性。模型 M8a 的似然值 $l = -5\ 695.34$, 和模型 M8 相比 $2\Delta l = 2 \times (-5\ 679.93 + 5\ 695.34) = 30.82$, 由 χ^2 分布 (d. f. = 1) 得出 $P = 0.28 \times 10^{-7}$, 因此模型 M8a 被极显著地拒绝。

在 95% 水平上, 模型 M2a 鉴定出有 4 个氨基酸位点 (133L, 251M, 282Y 和 326I) 受到正向选择, 而模型 M8 鉴定出有 7 个氨基酸位点 (133L, 251M, 262A, 265V, 282Y, 326I 和 362I) 受到正向选择 (表 2)。

表 2 各类位点间可变 ω 比值模型下的参数估计值和对数似然值

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

模型	p^*	l	参数估计值	正向选择位点
M0: 单一比值	1	-5853.01	$\omega = 0.144$	无
M1a: 近中性	2	-5699.14	$p_0 = 0.888, \omega_0 = 0.035$ $p_1 = 0.112, \omega_1 = 1$	不允许
M2a: 选择	4	-5685.47	$p_0 = 0.892, \omega_0 = 0.039$ $p_1 = 0.088, \omega_1 = 1$ $p_2 = 0.020, \omega_2 = 3.162$	133L, 251M, 262A, 265V, 282Y, 326I, 362I
M3: 离散	5	-5680.51	$p_0 = 0.812, \omega_0 = 0.018$ $p_1 = 0.161, \omega_1 = 0.499$ $p_2 = 0.027, \omega_2 = 2.805$	133L, 221V, 251M, 262A, 265V, 282Y, 326I, 362I
M7: beta	2	-5701.25	$p = 0.073, q = 0.429$	不允许
M8: beta 和 ω	4	-5679.93	$p_0 = 0.974, p = 0.132, q = 1.136$ $p_1 = 0.026, \omega = 2.843$	133L, 221V, 251M, 262A, 265V, 282Y, 326I, 362I

注: * 表示 ω 分布所用的参数数目。在 99% 和 95% 水平上推测出的受选择位点分别用黑体和斜体标出。氨基酸位点编号参照台湾水龙骨 *rbcl* 基因序列^[18]

表 3 似然比值检验统计量 ($2\Delta l$)

Tab. 3 Likelihood ratio statistics ($2\Delta l$)

模型比较	$2\Delta l$	d. f.	$\chi^2_{1\%}$
M0 - M3	345.00	4	13.28
M1a - M2a	27.34	2	9.21
M7 - M8	42.64	2	9.21

3 讨论

为保证数据的一致和完整,我们舍弃了长度较短的 *rbcl* 基因序列,最终只将 49 种水龙骨科植物的 *rbcl* 基因序列用于构树(图 1)。尽管构建此树的目的是为了随后用树文件进行 *rbcl* 基因适应性进化分析而不是探讨系统学问题,但是该树仍能给出一些颇具系统分类意义的信息。首先,瓦韦属为并系类群,尖嘴蕨属和丝带蕨属成员嵌套其间;建议把后两属并入前者以建立自然的瓦韦属。Zink^[22]也曾指出尖嘴蕨属和丝带蕨属可能是瓦韦属的衍生分支。第二,星蕨属是多系类群,应该将不属于该属的成员分离出去。例如,把 *M. varians*、*M. scandens*、*M. novo-zealandiae*、*M. viellardii* 及 *M. pustulatum* 移入蚁蕨属。另外,支持显脉蕨属为自然类群,处在星蕨类基部位置。

Rubisco 大亚基的羧基末端拥有一 α/β 桶结构域,其活性中心即是由一个大亚基的 α/β 桶结构域同另一大亚基的氨基末端结构域组成^[23-24]。由水龙骨科 *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 4 个和 7 个正向选择位点。在这些位点当中,除位点 133L, 326I 和 362I 外,其余 4 个位点 251M, 262A, 265V 和

282Y 均位于 α/β 桶结构域的底部(表 2)。实验生物化学方面,以往为鉴定 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。对于位点 262,突变体回复实验显示它与 Rubisco 维持 CO_2/O_2 专一性和羧化作用的最大反应速度有关^[25]。位点 265,如果单独对其进行定向诱变,Rubisco 的稳定性不变,酶动力学方面也只是使羧化作用的最大反应速度 V_c 、米氏常数 K_c 以及 V_c/K_c 比值稍微降低^[26]。但是,如果再同时替换位点 265 的临近位点 256 和 258,就会导致 Rubisco 变得更易降解、催化效率明显降低^[26]。实验表明, α/β 桶中的环 6 对维持 Rubisco 催化作用和决定羧化作用与氧化作用的比值(Ω)至关重要^[27-28]。本研究鉴定出的位点 326 就位于环 6。Zhu 和 Spreitzer^[29]注意到,替换位点 326 对 Rubisco 的动力学特性和 Ω 值改变不大,不影响催化效率,但却会使全酶的稳定性显著下降。至于位点 282,目前还缺乏碱基突变方面的报道。然而,来自结构生物学的数据显示,通过范德华作用,位点 283 能同时与位点 282 和 258 相作用,而位点 284 又能同时与位点 282 和 265 相作用^[30-31]。这提示位点 282 可以以位点 283 和 284 为中介同位点 258 和 265 发生作用,从而影响 Rubisco 的功能。

从以上分析可知,对于利用位点间可变 ω 比值模型鉴定出的 7 个正向选择位点,其中已有 4 个已获得实验支持,证明它们是同 Rubisco 的稳定性和动力学特性相关的重要功能位点。这说明,依托 ω 比值建立的 DNA 编码序列进化模型的确是检验分

子适应行之有效的方法。在资源受限的情况下, 未来的定向诱变和遗传操作实验应该优先以剩余的3个正向选择位点为目标, 从而加速阐明 Rubisco 催化机制的进程。另外也提示, 在水龙骨科植物的物种分化过程中, 自然选择不仅对其形态性状变异和生活史特性起作用, 而且还能影响编码重要功能蛋白的基因(例如 *rbcl* 基因)的分子进化。这对理解水龙骨科类发生在白垩纪的“辐射”式物种分化具有特殊意义。在白垩纪, 伴随被子植物的兴盛, 陆地生态系统开始发生明显改变^[16]。水龙骨科类 Rubisco 的适应性进化可被看作是它们对此生态环境变化作出的响应之一。通过对 Rubisco 大亚基上的特定氨基酸位点进行有益替换, 水龙骨科类能够净增 CO₂ 固定量, 显著提升光合作用效率, 进而改善竞争潜力。

参考文献:

- [1] ELLIS R J. The most abundant protein on earth [J]. *Trends Biochem Sci*, 1979, 4: 241 - 244.
- [2] GUTTERIDGE S, GATENBY A A. Rubisco synthesis, assembly, mechanism and regulation [J]. *Plant Cell*, 1995, 7: 808 - 819.
- [3] PORTIS A R J. Regulation of ribulose - 1,5-bisphosphate carboxylase/oxygenase activity [J]. *Annu Rev Plant Physiol Plant Mol Biol*, 1992, 43: 415 - 437.
- [4] SPREITZER R J, SALVUCCI M E. Rubisco: Structure, regulatory interactions, and possibilities for a better enzyme [J]. *Annu Rev Plant Biol*, 2002, 53: 449 - 475.
- [5] 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.
- [6] YANG Z. Computational molecular evolution [M]. Oxford: Oxford University Press, 2006.
- [7] 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.
- [8] YANG Z, SWANSON W J, VACQUIER V D. Maximum-likelihood analysis of molecular adaptation in abalone sperm lysin reveals variable selective pressures among lineages and sites [J]. *Mol Biol Evol*, 2000, 17: 1446 - 1455.
- [9] WONG W S W, YANG Z, GOLDMAN N, et al. Accuracy and power of statistical methods for detecting adaptive evolution in protein coding sequences and for identifying positively selected sites. *Genetics*, 2004, 168: 1041 - 1051.
- [10] SUZUKI Y, NEI M. Reliabilities of parsimony-based and likelihood-based methods for detecting positive selection at single amino acid sites [J]. *Mol Biol Evol*, 2001, 18: 2179 - 2185.
- [11] SUZUKI Y, NEI M. Simulation study of the reliability and robustness of the statistical methods for detecting positive power of the likelihood ratio test in detecting adaptive molecular selection at single amino acid sites [J]. *Mol Biol Evol*, 2002, 19: 1865 - 1869.
- [12] ZHANG J. Frequent false detection of positive selection by the likelihood method with branch-site models [J]. *Mol Biol Evol*, 2004, 21: 1332 - 1339.
- [13] YANG Z, NIELSEN R. Codon-substitution models for detecting molecular adaptation at individual sites along specific lineages [J]. *Mol Biol Evol*, 2002, 19: 908 - 917.
- [14] 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.
- [15] SMITH A R, PRYER K M, SCHUETTPELZ E, et al. A classification for extant ferns [J]. *Taxon*, 2006, 55: 705 - 731.
- [16] SCHNEIDER H, SCHUETTPELZ E, PRYER K M, et al. Ferns diversified in the shadow of angiosperms [J]. *Nature*, 2004, 428: 553 - 557.
- [17] THOMPSON J D, HIGGINS D G, GIBSON T J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice [J]. *Nucl Acids Res*, 1994, 22: 4673 - 4680.
- [18] HIROHARA M, NAKANE T, TERAYAMA Y, et al. Chemotaxonomy of ferns: triterpenoids and *rbcl* gene sequences of *Polypodium*, *Polypodiodes* and *Goniophlebium* [J]. *Natural Med*, 2000, 54: 330 - 333.
- [19] POSADA D, CRANDALL K A. Modeltest: testing the model of DNA substitution [J]. *Bioinformatics*, 1998, 14: 817 - 818.
- [20] HULSENBECK J P, RONQUIST F Mr. Bayes: Bayesian inference of phylogeny [J]. *Bioinformatics*, 2001, 17: 754 - 755.
- [21] YANG Z. PAML 4: a program package for phylogenetic analysis by maximum likelihood [J]. *Mol Biol Evol*, 2007, 24: 1586 - 1591.
- [22] ZINK M J. Affinities and delimitation of the genus *Lepisorus* (J. Sm.) Ching (Polypodiaceae) [M] // SHING K - H, KRAMER K - U. Proceedings of the international symposium on systematic pteridology. Beijing: China Science and Technology Press, 1989: 75 - 81.
- [23] KNIGHT S, ANDERSSON I, BRÄNDÉN C I. Crystallographic analysis of ribulose 1,5 - bisphosphate carboxyl-

- ase from spinach at 2.4 Å resolution Subunit interactions and active site[J]. *J Mol Biol*, 1990, 215: 113–160.
- [24] SOPER T S, MURAL R J, LARIMER F W, 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*, 1988, 2: 39–44.
- [25] HONG S, 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.
- [26] 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.
- [27] CHEN Z, SPREITZER R J. Chloroplast intragenic suppression enhances the low CO₂/O₂ specificity of mutant ribulose-bisphosphate carboxylase/oxygenase[J]. *J Biol Chem*, 1989, 264: 3051–3053.
- [28] CHEN Z, YU W, LEE J H, et al. Complementing amino acid substitutions within loop 6 of the .alpha./.beta.-barrel active site influence the carbon dioxide/oxygen specificity of chloroplast ribulose-1,5-bisphosphate carboxylase/oxygenase[J]. *Biochemistry*, 1991, 30: 8846–8850.
- [29] ZHU G, SPREITZER R J. 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.
- [30] 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 biphosphate[J]. *J Mol Biol*, 1996, 259: 160–174.
- [31] TAYLOR T C, BACKLUND A, BJORHALL K, et al. First crystal structure of rubisco from a green alga, *Chlamydomonas reinhardtii* [J]. *J Biol Chem*, 2001, 276: 48159–48164.

Adaptive Evolution in the Rubisco Large Subunits of Epiphytic Ferns (Polypodiaceae): Identification of Positively Selected Sites

SU Ying-juan¹, WANG Ting²

(1. School of Life Sciences, Sun Yat-sen University, Guangzhou 510275, China;

2. Wuhan Botanical Garden, the Chinese Academy of Sciences, Wuhan 430074, China)

Abstract: Ribulose-1, 5-bisphosphate carboxylase/oxygenase (Rubisco, EC 4.1.1.39) plays pivotal roles in plant photosynthesis, initiating both carbon assimilation and photorespiration. Rubisco large subunit, which contains the active sites involved in catalyzing CO₂ fixation, is encoded by the chloroplast *rbcL* gene. The Polypodiaceae are likely the most derived group of extant ferns, and they mainly grow as epiphytes in tropical or subtropical climates. Recently polypod ferns have been demonstrated to undergo diversifications in the Cretaceous after the rise of angiosperms. To further test this proposal, an adaptive evolutionary analysis was performed on the *rbcL* gene sequences of Polypodiaceae ferns under models allowing ω (nonsynonymous/synonymous rate ratio) to vary among sites. By comparing Models M1a/M2a and M7/M8, seven amino acid sites were identified under positive selection: 133L, 251M, 262A, 265V, 282Y, 326I and 362I. Functional significances of four sites 262A, 265V, 282Y and 326I were lent support by published experimental results. This research consolidated that evolutionary models using ω as an indicator of selective pressure provide valuable approaches for analyzing the adaption of protein-encoding sequences. Moreover, it also indicated that Polypodiaceae ferns responded to the terrestrial ecosystem changes initiated by angiosperms through molecular adaption of functional genes such as *rbcL*.

Key words: Polypodiaceae; ribulose-1, 5-bisphosphate carboxylase/oxygenase; *rbcL* gene; positively selected sites