

DOI: 10.11913/PSJ.2095-0837.2020.10010

熊哲铭, 高一波, 任慧莹, 徐波, 吴思婉, 彭圆, 森林. 40 种蕨类植物 *matK* 基因的系统分类及分子进化研究 [J]. 植物科学学报, 2020, 38(1): 10~22Xiong ZM, Gao YB, Ren HY, Xu B, Wu SW, Peng Y, Sen L. Analysis on the phylogenetic classification and molecular evolution of the *matK* gene in 40 fern species [J]. *Plant Science Journal*, 2020, 38(1): 10~22

40 种蕨类植物 *matK* 基因的系统分类及分子进化研究

熊哲铭¹, 高一波¹, 任慧莹¹, 徐波², 吴思婉¹, 彭圆^{2*}, 森林^{1*}

(1. 湖北中医药大学药学院, 武汉 430065; 2. 湖北中医药大学老年医学研究所, 武汉 430065)

摘要: 采用“放松分子钟”模型、氨基酸位点正选择模型和分子内共进化网络估算方法, 对蕨类植物Ⅱ型内含子成熟酶蛋白 K(Maturase K, MATK)编码基因 *matK* 的进化趋势进行研究。结果显示: *matK* 基因在蕨类植物系统学研究中具有一定的应用价值, 与 *rbcL* 基因和 *psaA* 基因联合后能显著提升系统发育树的可信度; 蕨类植物 MATK 蛋白中存在少数曾经历正选择的位点; MATK 蛋白内部有多对氨基酸位点共同构成共进化网络。在被子植物兴起环境改变后, MATK 蛋白部分位点发生适应性进化, 通过位点间共进化网络协同作用方式提升蕨类植物对新光合环境的适应能力。

关键词: 蕨类植物; 阴地蕨属; *matK* 基因; 适应性进化; 共进化

中图分类号: Q949.36

文献标识码: A

文章编号: 2095-0837(2020)01-0010-13

Analysis on the phylogenetic classification and molecular evolution of the *matK* gene in 40 fern species

Xiong Zhe-Ming¹, Gao Yi-Bo¹, Ren Hui-Ying¹, Xu Bo², Wu Si-Wan¹, Peng Yuan^{2*}, Sen Lin^{1*}

(1. College of Pharmacy, Hubei University of Chinese Medicine, Wuhan 430065, China;

2. Institute of Gerontology, Hubei University of Chinese Medicine, Wuhan 430065, China)

Abstract: The *matK* gene is the only group II intron maturase encoded in the chloroplast genome. It is the potentiality left untested that applying the *matK* gene as the molecular marker reconstructs credible fern phylogenetic trees under estimated timescale. Using the relaxed molecular clock model, positive selection model, and co-evolutionary analysis, we attempted to unravel the evolutionary pattern of the *matK* gene. Results showed that *matK* had certain application value in the phylogenetic study of ferns, and the reliability of the phylogenetic tree was significantly enhanced when combined with *rbcL* and *psaA*. The study also indicated a few positively selected amino-acid sites in the MATK protein. Multiple pairs of amino-acid sites in the MATK protein evolved to form a co-evolutionary network. The modification of the MATK protein by site adaptation and the protein intra-network may be important factors that benefited the adaptive evolution of ferns under the change in photosynthetic environment after the rise of angiosperms.

Key words: Monilophytes; *Botrychium*; *matK* gene; Adaptive evolution; Co-evolution

收稿日期: 2019-06-10, 退修日期: 2019-07-26。

基金项目: 国家自然科学基金项目(31500260, 81574037); 湖北中医药大学“青苗计划”项目(2016ZZX015)。

This work was supported by grants from the National Natural Science Foundation of China (31500260, 81574037) and Youth Talent Project of Hubei University of Chinese Medicine (2016ZZX015).

作者简介: 熊哲铭(1995-), 男, 硕士研究生, 研究方向为中药资源及品质(E-mail: xiongheming_hbtcm@126.com)。

* 通讯作者(Author for correspondence. E-mail: senlin_hbtcm@126.com, happyppy05@126.com)。

传统分类学观点认为蕨类植物一般分为松叶蕨类(Whisk ferns)、木贼类(Horsetails)、石松类(Lycopods)和真蕨类(Ferns)4类,其中前3类合称拟蕨类(fern-allies),真蕨类又称为蕨类(ferns)^[1]。随着分子生物学技术在蕨类植物系统学研究中的广泛应用,传统分类系统受到了分子数据的冲击和检验。基于形态学和分子序列证据,植物分类学家对原归于蕨类植物门(Pteridophyta)的物种进行了系统发育重排,并以蕨类植物(Moniliophytes)和石松类植物(Lycopophytes)为名将物种进行重新区分^[2, 3]。新分类系统将蕨类植物分为4大类群:薄囊蕨(Leptosporangiate)、厚囊蕨(Eusporangiate)、木贼和松叶蕨^[4]。一般认为,在被子植物出现后,蕨类植物逐渐失去生态上的优势地位,其物种多样性和丰度都急剧下降^[5]。然而,研究发现,随着被子植物的兴起,部分蕨类植物因适应阴生环境而出现大规模的物种辐射^[6]。目前,蕨类植物适应新光合环境以及物种辐射现象所蕴藏的分子机理尚未完全厘清。

植物适应新的光合环境时,其在光合作用中起关键作用的一些基因可能伴随着物种辐射而发生适应性进化。许可等^[7]研究发现,树蕨类和其他蕨类的 *psbD* 基因进化趋势各异;森林等^[8]和周媛等^[9]研究发现,凤尾蕨科(Pteridaceae)植物旱生分支、水生分支和附生分支的 *rbcL* 基因受到不同程度的选择压力;Sen 等^[10]和吴筱婷等^[11]对蕨类植物光合系统核心蛋白编码基因的研究结果表明,蕨类植物光合系统Ⅱ核心蛋白编码 *psbA* 基因和光合系统Ⅰ核心蛋白编码基因 *psaA* 呈现出不同的进化趋势;Sen 等^[12, 13]对裸子植物的研究发现,其 *rbcL* 和 *psbA* 基因也曾在正选择作用下发生了改变。物种的演化历程即是不断适应生存环境的过程,对其适应性进化的探究有利于理解物种进化潜力及进化机理。

matK 基因位于叶绿体 *trnK* 基因的内含子中,长约 1500 bp,编码一种成熟酶 K,该成熟酶参与叶绿体Ⅱ型内含子的剪切,涉及对象包括 *trnK*、*trnA*、*trnl*、*atpF* 等基因的转录本,对维持叶绿体正常功能有重要意义^[14]。有研究报道 *matK* 基因在蕨类植物系统学中能解决一些复杂类群的系统关系问题^[15, 16],还有人提出将 *matK* 和 *rbcL* 基因共同作为蕨类植物的核心 DNA 条形码^[17]。在植物系

统学研究中,基因全长序列往往包含更多的演化信息,能反映物种更真实的演化历史。虽然目前 *matK* 基因在蕨类植物中应用广泛,但大多以 *matK* 部分序列作为分子标记^[18~20], *matK* 全长序列的应用价值还有待研究。另外,最近的研究显示 *psaA* 基因的全序列在蕨类植物系统学研究中具有应用潜力^[11],而 *matK*、*rbcL* 及 *psaA* 基因全序列在系统发育重建中的效率仍需要进一步研究。

因此,本研究对 25 科 40 种具有代表性的蕨类植物 *matK* 基因的编码区进行系统发育重建、适应性进化分析及其进化关系研究,探究在“放松分子钟”模型下, *matK* 基因序列能否正确刻画蕨类植物 4 大类群的演化关系,并比较该基因与传统分子标记 *rbcL* 基因、新分子标记 *psaA* 基因及联合数据集间所构建的系统发育树的可信度及效率;探讨随着部分蕨类植物的物种辐射, *matK* 基因是否受到选择压力而发生适应性进化;估算 MATK 蛋白氨基酸位点间的共进化关系,以期为明晰蕨类植物适应环境变化的分子机制研究奠定基础。

1 材料与方法

1.1 样品情况

本研究选取 4 大类群 25 科 40 种蕨类植物的 *matK*、*rbcL* 及 *psaA* 基因编码区全长序列(表 1)。其中,阴地蕨(*Botrychium ternatum* (Thunb.) Sw.)、华东阴地蕨(*Botrychium japonicum* (Prantl) Underw.)和蕨萁(*Botrychium virginianum* (L.) Sw.)共 9 条基因序列由测序方法获得,测序标本具体信息见表 2;其他序列从 GenBank 网站(<https://www.ncbi.nlm.nih.gov/>)下载得到。

1.2 DNA 提取、PCR 扩增及测序

取新鲜叶片,采用改良的 CTAB 法^[21]提取总 DNA。使用 Primer Premier 5.0 软件分别设计 *matK*、*rbcL* 和 *psaA* 基因的特异性引物。PCR 反应体系为 25 μL 体系,其中包括: 2 × Taq PCR Master Mix 12.5 μL, 正、反向引物(2.5 μmol/L)各 1 μL, 总 DNA 1 μL。扩增程序为: 94℃预变性 5 min, 94℃变性 30 s, 52℃退火 30 s, 72℃延伸 60 s, 共 35 个循环;最后 72℃延伸 5 min。对阳性扩增产物进行测序,序列采用 MEGA 7.0 软件进行拼接以获得相应的基因编码区全序列^[22]。

表1 本研究所选取的植物及其GenBank登录号
Table 1 Sampled plants and GenBank accession numbers

类别 Classification	科名 Family	种名 Species	matK		rbcL		psaA	
			GenBank 登录号 GenBank acc. no.	长度 Length (bp)	GenBank 登录号 GenBank acc. no.	长度 Length (bp)	GenBank 登录号 GenBank acc. no.	长度 Length (bp)
薄囊蕨 Leptosporangiatae	Aspleniaceae	长叶铁角蕨 <i>Asplenium prolongatum</i> Hook.	NC_035838	1506	NC_035838	1428	NC_035838	2253
		半边铁角蕨 <i>Hymenophyllum unilaterale</i> (Lam.) Hayata	NC_035856	1497	NC_035856	1428	NC_035856	2253
	Athyriaceae	<i>Athyrium opacum</i> (D. Don) Copel.	NC_035841	1506	NC_035841	1428	NC_035841	2253
		单叶对囊蕨 <i>Diparia lancea</i> (Thunb.) Fraser-Jenk.	NC_035844	1506	NC_035844	1428	NC_035844	2253
	Blechnaceae	顶芽狗脊 <i>Woodwardia unigemmata</i> (Makino) Nakai	NC_028543	1509	NC_028543	1428	NC_028543	2253
		<i>Austroblechnum melanocaulon</i> (Brack.) Gasper & V. A. O. Dittrich	NC_035840	1503	NC_035840	1428	NC_035840	2253
	Cibotiaceae	金毛狗 <i>Cibotium barometz</i> (L.) J. Sm.	NC_037893	1509	NC_037893	1428	NC_037893	2253
	Cyatheaceae	桫椤 <i>Alsophila spinulosa</i> (Wall. ex Hook.) R. M. Tryon	NC_012818	1503	NC_012818	1428	NC_012818	2253
	Cystopteridaceae	<i>Cystopteris chinensis</i> Ching	NC_035843	1503	NC_035843	1428	NC_035843	2253
	Dennstaedtiaceae	欧洲蕨 <i>Pteridium aquilinum</i> (L.) Kuhn	NC_014348	1494	NC_014348	1428	NC_014348	2253
球子蕨 Onocleaceae	Diplaziopsidaceae	川黔肠蕨 <i>Diplaziopsis cavalieriana</i> (H. Christ) C. Chr.	NC_035847	1512	NC_035847	1428	NC_035847	2253
		<i>Homalosorus pycnocarpus</i> (Spreng.) Pic. Serm.	NC_035855	1524	NC_035855	1428	NC_035855	2253
	Dryopteridaceae	披针贯众 <i>Cyrtomium devexiscapulae</i> (Koidz.) Koidz. & Ching	NC_028542	1503	NC_028542	1428	NC_028542	2253
		迷人鳞毛蕨 <i>Dryopteris decipiens</i> (Hook.) Kuntze	NC_035854	1506	NC_035854	1428	NC_035854	2253
	Dicksoniaceae	<i>Dicksonia squarrosa</i> (Forst.) Sw.	KJ569698	1506	KJ569698	1428	KJ569698	2253
	Gleicheniaceae	<i>Diplopterygium glaucum</i> (Thunb. ex Houtt.) Nakai	NC_024158	1518	NC_024158	1428	NC_024158	2253
	Lygodiaceae	海金沙 <i>Lygodium japonicum</i> (Thunb.) Sw.	NC_022136	1518	NC_022136	1428	NC_022136	2253
		<i>Lygodium microphyllum</i> (Cav.) R. Br.	NC_039378	1518	NC_039378	1428	NC_039378	2253
	Marsileaceae	南国田字草 <i>Marsilea crenata</i> C. Presl	NC_022137	1527	NC_022137	1428	NC_022137	2253
	Onocleaceae	菱果蕨 <i>Matteuccia struthiopteris</i> (L.) Tod.	NC_035859	1506	NC_035859	1428	NC_035859	2253
		球子蕨 <i>Onoclea sensibilis</i> L.	NC_035860	1500	NC_035860	1428	NC_035860	2253

续表 1

类别 Classification	科名 Family	种名 Species	matK		rbcL		psaA	
			GenBank 登录号 GenBank acc. no.	长度 (bp)	GenBank 登录号 GenBank acc. no.	长度 (bp)	GenBank 登录号 GenBank acc. no.	长度 (bp)
	Osmundaceae	<i>Osmundastrum cinnamomeum</i> (L.) C. Presl	NC_024157	1548	NC_024157	1428	NC_024157	2253
	Polypodiaceae	<i>Polypodium glycyrrhiza</i> D. C. Eaton	KP136832	1503	KP136832	1428	KP136832	2253
	Pteridaceae	碎米蕨 <i>Cheilanthes lindheimeri</i> Hook.	NC_014592	1503	NC_014592	1428	NC_014592	2253
		<i>Ceratopteris richardii</i> Brongn.	KM052729	1503	KM052729	1428	KM052729	2253
薄囊蕨 Leptosporangiate	Rhachidosoraceae	喜钙轴果蕨 <i>Rhachidosorus consimilis</i> Ching	NC_035862	1488	NC_035862	1428	NC_035862	2253
	Schizaeaceae	<i>Schizaea pectinata</i> (L.) Sw.	NC_035808	1473	NC_035808	1428	NC_035808	2253
	Thelypteridaceae	普通针毛蕨 <i>Macrothelypteris torresiana</i> (Gaudich.) Ching	NC_035858	1506	NC_035858	1428	NC_035858	2253
		耳状紫柄蕨 <i>Thelypteris aurita</i> (Hook.) Ching	NC_035861	1506	NC_035861	1428	NC_035861	2253
	Woodsiaceae	耳羽岩蕨 <i>Woodsia polystichoides</i> D. C. Eaton	NC_035865	1503	NC_035865	1428	NC_035865	2253
厚囊蕨 Eusporangiate	Ophioglossaceae	七指蕨 <i>Helminthostachys zeylanica</i> (L.) Hook.	KM817788	1503	KM817788	1431	KM817788	2253
		<i>Mankyua chejuensis</i> B. Y. Sun, M.H.Kim & C. H. Kim	NC_017006	1500	NC_017006	1434	NC_017006	2253
		阴地蕨 <i>Botrychium ternatum</i> (Thunb.) Sw.	MN030170	1500	MN030172	1428	MN030171	2253
		华东阴地蕨 <i>Botrychium japonicum</i> (Prantl) Underw.	MN030167	1500	MN030169	1428	MN030168	2253
		蕨萁 <i>Botrychium virginianum</i> (L.) Sw.	MN030173	1500	MN030175	1428	MN030174	2253
	Marattiaceae	<i>Angiopteris angustifolia</i> C. Presl	NC_026300	1521	NC_026300	1428	NC_026300	2253
木贼类 Horsetails	Equisetaceae	木贼 <i>Equisetum hyemale</i> L.	NC_020146	1467	NC_020146	1428	NC_020146	2253
		问荆 <i>Equisetum arvense</i> L.	NC_014699	1473	NC_014699	1428	NC_014699	2253
松叶蕨类 Whisk ferns	Psilotaceae	松叶蕨 <i>Psilotum nudum</i> (L.) P. Beauv.	NC_003386	1512	NC_003386	1428	NC_003386	2253
		<i>Tmesipteris elongata</i> Danguy	KJ569699	1518	KJ569699	1428	KJ569699	2253

表2 测序标本信息
Table 2 Information on sequencing samples

种名 Species	凭证标本号 Voucher specimen no.	标本采集地 Sampling position	NCBI 数据库信息		凭证标本保存处 Specimen storage
			基因名 Gene	GenBank 登录号 Accession no.	
阴地蕨 <i>Botrychium ternatum</i>	YJD01	湖北恩施	<i>matK</i>	MN030170	湖北中医药大学
			<i>rbcL</i>	MN030172	
			<i>psaA</i>	MN030171	
华东阴地蕨 <i>Botrychium japonicum</i>	HD201	湖北恩施	<i>matK</i>	MN030167	湖北中医药大学
			<i>rbcL</i>	MN030169	
			<i>psaA</i>	MN030168	
蕨萁 <i>Botrychium virginianum</i>	JQ01	湖北荆门	<i>matK</i>	MN030173	湖北中医药大学
			<i>rbcL</i>	MN030175	
			<i>psaA</i>	MN030174	

1.3 序列数据

使用 MEGA 7.0 软件包中 Clustal W(密码子)方式进行序列比对, 比对后 *matK* 基因数据集长度为 1620 bp, *rbcL* 数据集长度为 1434 bp, *psaA* 数据集长度为 2253 bp。为了方便比较, 密码子位点的编号参照欧洲蕨 (*Pteridium aquilinum* (L.) Kuhn., NC_014348) 的基因序列。构建以下 7 个数据集: (1) 仅包括 *matK* 编码区; (2) 仅包括 *rbcL* 编码区; (3) 仅包括 *psaA* 编码区; (4) *matK* 及 *rbcL* 编码区全长序列; (5) *matK* 及 *psaA* 编码区全长序列; (6) *rbcL* 及 *psaA* 编码区全长序列。 (7) 联合 *matK*、*rbcL* 及 *psaA* 编码区全长序列。多序列联合数据集通过将多基因编码区全序列首尾相连得到。

1.4 统计分析

采用 jModeltest v2.1.10 软件分别选取上述 7 个数据集的最适核苷酸进化模型^[23-25]。分别利用 BEAST v1.8.4 软件的 UCLD(uncorrelated lognormal distributed relaxed clock model) 分子钟模型估算系统树每个节点的分歧时间^[26]。用最近共同祖先时间(t_{MRCA})值校正主要分支的分歧时间: 薄囊蕨类植物总体类群 t_{MRCA} 值为 299.9 Mya(百万年前); 松叶蕨 (Psilotaceae) 和瓶尔小草 (Ophioglossaceae) 的 t_{MRCA} 值为 135.7 Mya^[10]。根据 jModeltest v2.1.10 软件所选取的 GTR + I +G 模型, 在 BEAUti v1.8.4 软件中设置贝叶斯迭代模型, *matK*、*rbcL* 及 *psaA* 编码区数据集迭代计算 3×10^8 代; *matK* + *rbcL*、*matK* + *psaA* 及 *rbcL* + *psaA* 数据集迭代计算 4×10^8 代; *matK* + *rbcL* + *psaA* 联合数据集迭代计算 5×10^8 代, 所有迭代运

算中每 1000 代保存 1 株样本; 将输出结果导入 Tracer v1.6.0 软件检测各参数后验概率的收敛程度, 所有参数的有效取样大小 (efficient sampling size, ESS) 值均大于 200 时, 确定运行达到收敛终点。将树结果导入 TreeAnnotator v1.8.4 软件去除最初 10% 的样本后, 用剩余样本中后验概率大于 95% 的拓扑结构重建时间尺度下的一致树。利用 Figtree v1.4.3 软件展开 7 个数据集各自一致树的拓扑结构。

利用 PAMLx v1.3.1 软件提供的正选择模型 (M2a、M5、M8) 与对照模型 (M1a、M7、M8a) 进行位点适应性进化分析^[27]。通过 6 个不同模型估算 *matK* 基因各位点的选择压力 (ω 值) 和后验概率 (P 值), 检测各模型下 *matK* 基因的正选择位点 ($\omega > 1$)、中性位点 ($\omega = 1$) 和负选择位点 ($\omega < 1$), 再将正选择模型与相应的对照模型进行似然率检测 (likelihood ratio test, LRT), 通过比较模型间差异的显著性来确定更合适的模型, 从而确定是否存在正选择位点。为了保证适应性进化分析的准确性, 将 *matK* 编码数据集中比对模糊及无法分析的区域移除后, 得到包含 502 个氨基酸位点的 *matK* 数据集, 再进行适应性进化分析。

利用 CAPS 软件分析蕨类植物蛋白质 MATK 内部的共进化氨基酸位点。

2 结果与分析

2.1 序列长度

本研究中 *rbcL* 和 *psaA* 基因较为保守, 而 *matK* 基因长度差异较大。除七指蕨 (*Helminthostachys zeylanica* (L.) Hook.) 和 *Mankyua chejuensis*

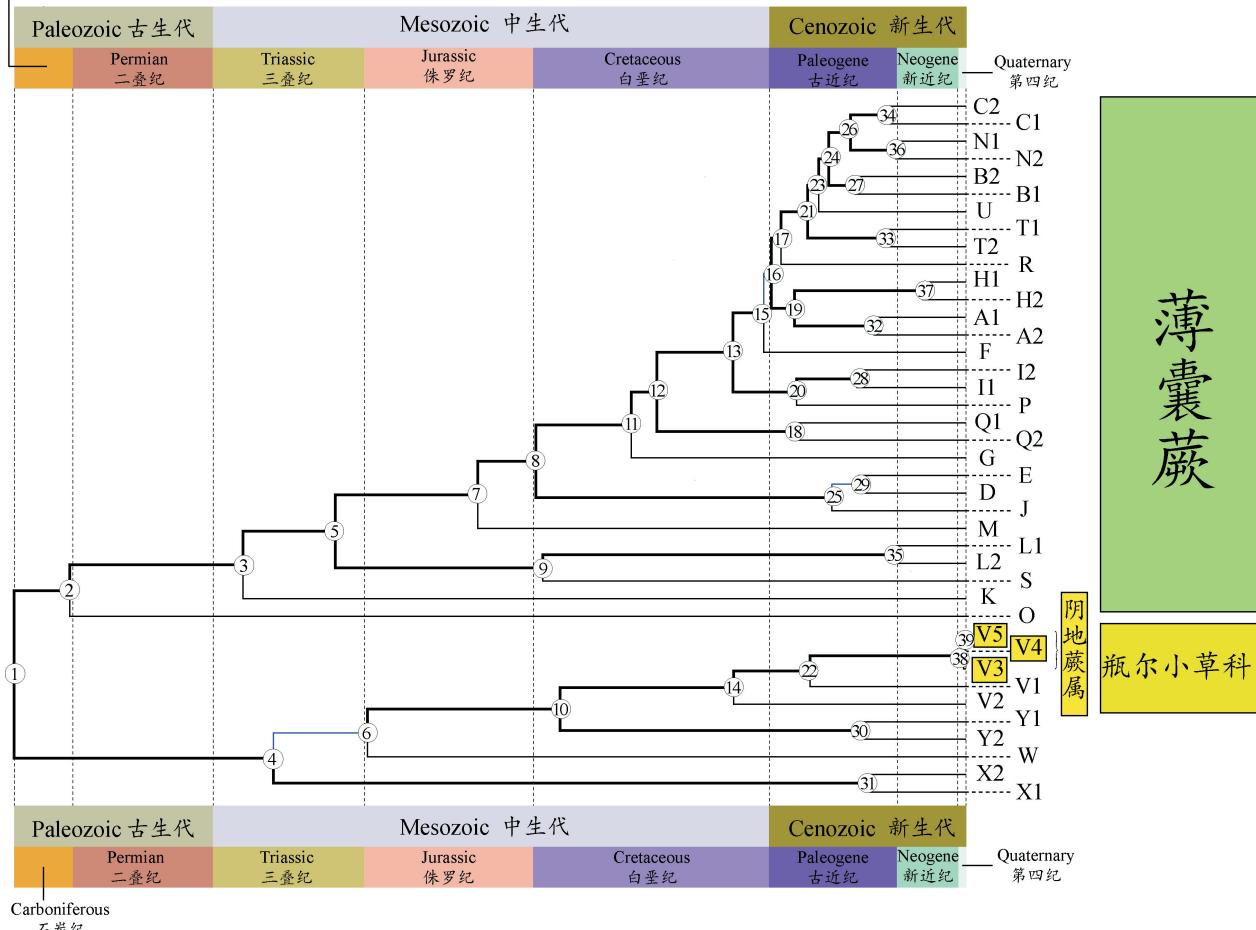
nsis B. Y. Sun, M. H. Kim & C. H. Kim以外, 所有物种的 *rbcL* 基因长均为 1428 bp; 所有物种的 *psaA* 基因长为 2253 bp; 半数 *matK* 基因的全长序列为 1506 bp, 木贼 (*Equisetum hyemale* L.) 的 *matK* 基因长度最短(1467 bp), *Osmundastrum cinnamomeum* (L.) C. Presl 的 *matK* 基因长度最长(1548 bp), 说明蕨类植物 *matK* 基因可能在进

化过程中发生了某种程度的插入或缺失事件, 可能与部分蕨类植物 *trnK* 基因的丢失有关。

2.2 基于7种数据集构建的蕨类植物系统发育关系

本研究利用贝叶斯法重建了蕨类植物系统发育关系, 结果显示(图 1、图 2), 不同数据集构建的系统发育树的拓扑结构基本一致, 薄囊蕨、木贼类和松叶蕨类分别聚为一支, 而厚囊蕨出现了多歧分

Carboniferous
石炭纪



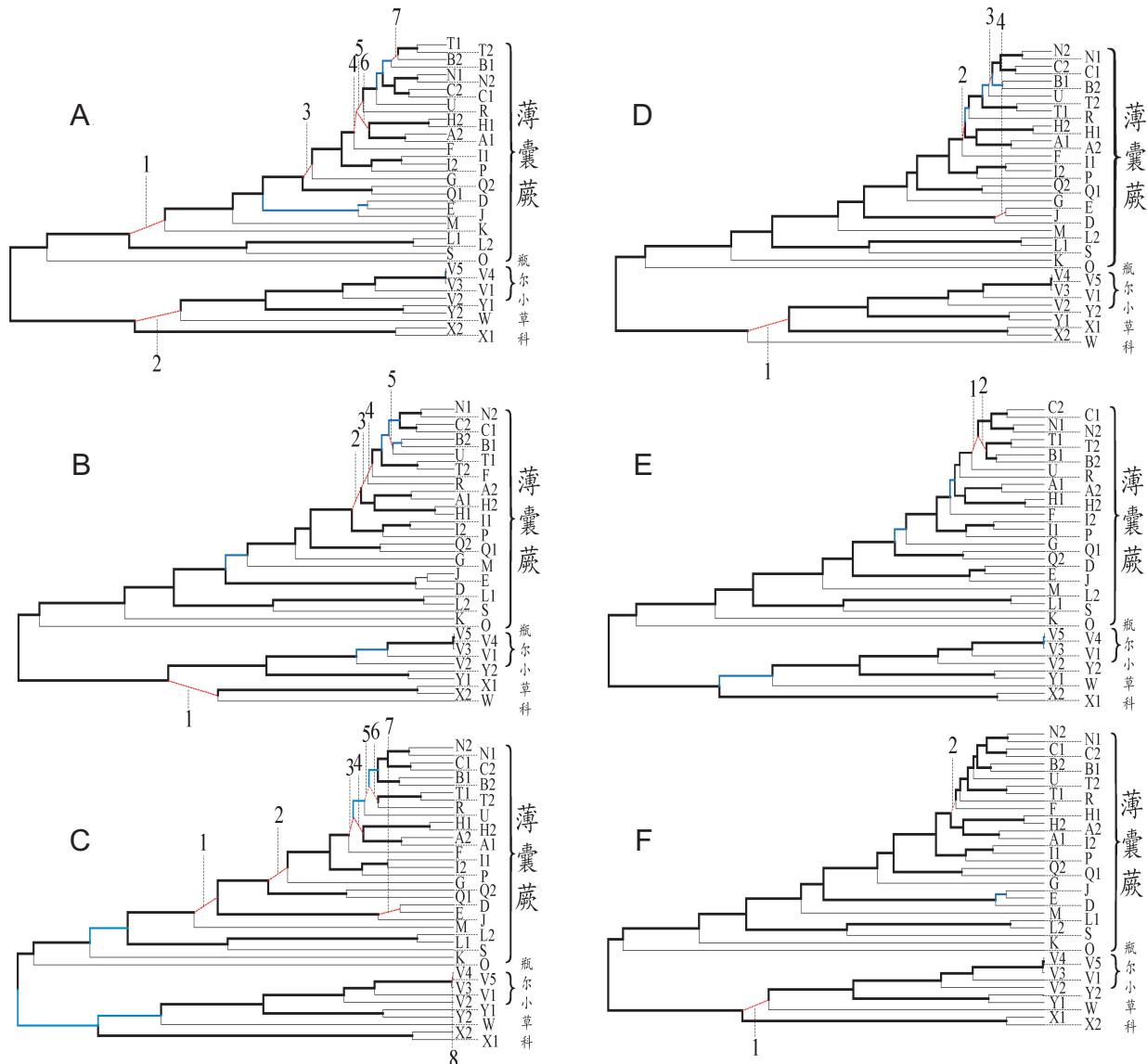
T1: 普通针毛蕨 (*Macrothelypteris torresiana*)；T2: 耳状紫柄蕨 (*Thelypteris aurita*)；B2: 单叶对囊蕨 (*Deparia lancea*)；B1: *Athyrium opacum*；N1: 荚果蕨 (*Matteuccia struthiopteris*)；N2: 球子蕨 (*Onclea sensibilis*)；C2: *Austroblechnum melanocaulon*；C1: 顶芽狗脊 (*Woodwardia unigemmata*)；U: 耳羽岩蕨 (*Woodsia polystichoides*)；R: 喜钙轴果蕨 (*Rhachidosorus consimilis*)；H2: *Homalosorus pycnocarpos*；H1: 川黔肠蕨 (*Diplaziopsis cavaleriana*)；A2: 半边铁角蕨 (*Hymenasplenium unilaterale*)；A1: 长叶铁角蕨 (*Asplenium prolongatum*)；F: *Cystopteris chinensis*；I1: 披针贯众 (*Cyrtomium devexiscapulae*)；I2: 迷人鳞毛蕨 (*Dryopteris decipiens*)；P: *Polypodium glycyrrhiza*；G: 欧洲蕨 (*Pteridium aquilinum*)；Q2: *Ceratopteris richardii*；Q1: 碎米蕨 (*Cheilanthes lindheimeri*)；D: 金毛狗 (*Cibotium barometz*)；E: 桑椤 (*Alsophila spinulosa*)；J: *Dicksonia squarrosa*；M: 南国田字草 (*Marsilea crenata*)；S: *Schizaea pectinata*；O: *Osmundastrum cinnamomeum*；V5: 蕨萁 (*Botrychium virginianum*)；V4: 华东阴地蕨 (*Botrychium japonicum*)；V3: 阴地蕨 (*Botrychium ternatum*)；V1: 七指蕨 (*Helminthostachys zeylanica*)；V2: *Mankyua chejuensis*；Y1: 松叶蕨 (*Psilotum nudum*)；Y2: *Tmesipteris elongata*；W: *Angiopteris angustifolia*；X2: 问荆 (*Equisetum arvense*)；X1: 木贼 (*Equisetum hyemale*)。

①~⑩代表通过贝叶斯算法得到的节点序号, 详见表 3。粗实线表示该分支后验概率大于 90%, 蓝色实线表示该分支后验概率在 70%~90% 之间。

The numbers in the circles (i. e. ①~⑩) represent the nodes estimated from Bayesian approaches respectively, as shown in Table 3. Thick lines indicate posterior probability greater than 90% and blue lines indicate posterior probability between 70% and 90%.

图 1 基于 *matK* + *rbcL* + *psaA* 联合数据集构建的蕨类植物系统发育树

Fig. 1 Phylogenetic tree constructed from combined *matK* + *rbcL*+ *psaA* dataset based on Bayesian analysis



A: *matK*; B: *psaA*; C: *rbcL*; D: *matK* + *psaA*; E: *matK* + *rbcL*; F: *rbcL* + *psaA*。物种编号同图 1。数字(例如, 图 2: A 中的 1 ~ 7)代表各系统发育树中后验概率低于 70% 的分支。粗实线表示该分支后验概率大于 90%, 蓝色实线表示该分支后验概率在 70% ~ 90% 之间, 红色虚线表示该分支后验概率低于 70%。

Species number is consistent with Fig. 1. The numbers (e. g. 1-7 in Fig. 2: A) represent the branches in each phylogenetic tree whose posterior probability is less than 70%, respectively. Thick lines, blue lines, and red dotted lines indicate posterior probability greater than 90%, between 70% and 90%, and less than 70%, respectively.

图 2 基于贝叶斯算法根据 6 个数据集分别构建的蕨类植物系统发育树

Fig. 2 Six different phylogenetic trees constructed from distinct datasets based on Bayesian analysis

化, 提示厚囊蕨可能不是单系类群。在蕨类植物系统发育关系重建过程中, 低后验概率分支可能会产生错误的分类结果, 如同属于蹄盖蕨科 (*Athyriaceae*) 的单叶对囊蕨 (*Deparia lancea* (Thunb.) Fraser-Jenk.) 和 *Athyrium opacum* (D. Don) Copel. 并未聚为一支(图 2: A)。7 个数据集中, *matK* + *rbcL* + *psaA* 构建的系统发育关系具有最高的可信度(图 1, 所有分支后验概率均大于

70%), 其余 6 个数据集包含后验概率较低分支(后验概率低于 70%)的个数由多到少依次为 *rbcL*(图 2: C, 8 个分支)、*matK*(图 2: A, 7 个分支)、*psaA*(图 2: B, 5 个分支)、*matK*+*psaA*(图 2: D, 4 个分支)、*matK*+*rbcL*(图 2: E, 2 个分支)和 *rbcL*+*psaA* 数据集(图 2: F, 2 个分支)。结果说明: (1) *matK* 在蕨类植物系统发育关系构建中潜力可能不如新分子标记基因 *psaA*, 但

相对传统的 *rbcL* 基因更有开发价值; (2) 多基因位点联合 (*matK* + *rbcL* + *psaA*) 能显著提高系统发育关系构建的可信度。

研究发现, 瓶尔小草科阴地蕨属 (*Botrychium*) 的 3 个成员 (*B. ternatum*、*B. japonicum*、*B. virginianum*) 在各个数据集构建的系统发育关系中

均以较高的后验概率聚为一支, 其分歧时间估算值分别为: 53.46 Mya (*matK*)、56.39 Mya (*rbcL*)、48.18 Mya (*psaA*)、50.64 Mya (*matK* + *psaA*)、53.06 Mya (*matK* + *rbcL*)、50.84 Mya (*rbcL* + *psaA*)、51.97 Mya (*matK* + *rbcL* + *psaA*) (表 3)。分歧时间估算值均指向阴地蕨属植物在古近

表 3 依据“放松分子钟”UCLD 模型估算各节点的参数值

Table 3 Estimated values at each node based on UCLED relaxed molecular clock model

节点 Node	最近共同祖先时间 t_{MRCA}	地质年代 Geological timescale			后验概率 Posterior (%)
		纪 Period	世 Epoch	期 Age	
1	318.59	石炭纪 Carboniferous	宾夕法尼亚世 Pennsylvanian	巴什基尔期 Bashkirian	100
2	299.89	石炭纪 Carboniferous	宾夕法尼亚世 Pennsylvanian	格舍尔期 Gzhelian	100
3	241.90	三叠纪 Triassic	中三叠纪 Middle	拉丁期 Ladinian	100
4	231.73	三叠纪 Triassic	晚三叠纪 Late	卡尼期 Carnian	96.73
5	211.06	三叠纪 Triassic	晚三叠纪 Late	诺利期 Norian	100
6	200.16	侏罗纪 Jurassic	早侏罗纪 Early	赫塘期 Hettangian	73.72
7	163.34	侏罗纪 Jurassic	晚侏罗纪 Late	牛津期 Oxfordian	100
8	143.86	白垩纪 Cretaceous	早白垩纪 Early	贝利阿斯期 Berriasian	100
9	141.49	白垩纪 Cretaceous	早白垩纪 Early	贝利阿斯期 Berriasian	100
10	135.71	白垩纪 Cretaceous	早白垩纪 Early	凡兰吟期 Valanginian	100
11	111.84	白垩纪 Cretaceous	早白垩纪 Early	阿尔布期 Albian	100
12	103.26	白垩纪 Cretaceous	早白垩纪 Early	阿尔布期 Albian	92.29
13	77.81	白垩纪 Cretaceous	晚白垩纪 Late	坎帕期 Campanian	100
14	77.47	白垩纪 Cretaceous	晚白垩纪 Late	坎帕期 Campanian	100
15	67.51	白垩纪 Cretaceous	晚白垩纪 Late	马斯特里赫特期 Maastrichtian	100
16	64.92	古近纪 Paleogene	古新世 Paleocene	丹尼期 Danian	80.85
17	61.62	古近纪 Paleogene	古新世 Paleocene	丹尼期 Danian	99.39
18	57.56	古近纪 Paleogene	古新世 Paleocene	坦尼特期 Thanetian	100
19	57.11	古近纪 Paleogene	古新世 Paleocene	坦尼特期 Thanetian	99.99
20	56.45	古近纪 Paleogene	古新世 Paleocene	坦尼特期 Thanetian	100
21	52.69	古近纪 Paleogene	始新世 Eocene	伊普雷期 Ypresian	100
22	51.97	古近纪 Paleogene	始新世 Eocene	伊普雷期 Ypresian	100
23	49.03	古近纪 Paleogene	始新世 Eocene	伊普雷期 Ypresian	98.41
24	45.58	古近纪 Paleogene	始新世 Eocene	鲁帝特期 Lutetian	98.81
25	44.75	古近纪 Paleogene	始新世 Eocene	鲁帝特期 Lutetian	100
26	38.37	古近纪 Paleogene	始新世 Eocene	巴尔通期 Bartonian	100
27	36.76	古近纪 Paleogene	始新世 Eocene	普利亚本期 Priabonian	99.99
28	36.25	古近纪 Paleogene	始新世 Eocene	普利亚本期 Priabonian	100
29	35.76	古近纪 Paleogene	始新世 Eocene	普利亚本期 Priabonian	86.67
30	35.57	古近纪 Paleogene	始新世 Eocene	普利亚本期 Priabonian	100
31	32.44	古近纪 Paleogene	渐新世 Oligocene	鲁培尔期 Rupelian	100
32	30.53	古近纪 Paleogene	渐新世 Oligocene	鲁培尔期 Rupelian	100
33	27.05	古近纪 Paleogene	渐新世 Oligocene	夏特期 Chattian	100
34	26.72	古近纪 Paleogene	渐新世 Oligocene	夏特期 Chattian	100
35	24.21	古近纪 Paleogene	渐新世 Oligocene	夏特期 Chattian	100
36	23.16	古近纪 Paleogene	渐新世 Oligocene	夏特期 Chattian	100
37	14.19	近新纪 Neogene	中新世 Miocene	兰哥期 Langhian	100
38	0.46	第四纪 Quaternary	更新世 Pleistocene	卡拉布里期 Calabrian	100
39	0.06	第四纪 Quaternary	更新世 Pleistocene	卡拉布里期 Calabrian	99.15

注: 节点序号如图 1。

Note: Node number see Fig. 1.

纪始新世伊普雷斯期发生了较大规模的物种分化(图1、图2)。

2.3 蕨类植物 MATK 蛋白正选择位点的估算

本研究对蕨类植物4大类群25科40个物种的 $matK$ 基因数据集,估算了不同模型对应的对数似然值($\ln L$)和参数估测值。模型M2a(选择)、M5(γ)和M8(β 和 ω)均允许 ω 取值大于1,并通过估算寻找其位置(表4)。经LRT检验后发现,M2a和M8均比其对照模型更符合所分析的数据(表5),其原假设模型在显著性水平 $P=1\%$ 之下被极显著拒绝。

在后验概率大于95%的水平上,模型M2a及M8中均发现了一致的4个正选择位点,即:22I、24S、267Y和465V(表4)。模型M2a、模型M5及模型M8的NEB(Naïve Empirical Bayes)分析

结果如图3、图4所示。模型M2a、模型M5及模型M8均认为89Q存在正选择,模型M2a认为其受到选择压力的概率为 $P=0.959$, $\omega=1.96$;模型M5认为该位点受到选择压力的概率为 $P=0.991$, $\omega=1.65$;模型M8认为该位点受选择压力的概率为 $P=0.946$, $\omega=1.52$ 。

2.4 蕨类植物 MATK 蛋白的重要共进化位点

本研究通过共进化分析鉴定到2个共进化组和19个共进化对(表6)。2个共进化组分别为:43~383~417和43~417~424。所有共进化对中有8个的Bootstrap值大于95%,其中共进化对45~458和145~264的Bootstrap值大于99%。共进化对133~367有较高的相关系数(大于99.9%),表明这两个位点联系非常紧密,一个位点发生改变后,另一个位点可以通过共进化方式进行补偿突

表4 MATK蛋白中的正选择氨基酸位点及其参数估测值
Table 4 Estimated parameters and positively selected sites of MATK protein

模型 Model	参数数量 No. parameters	对数似然值 $\ln L$	参数估计值 Estimated value of parameter	正选择位点 Positively selected site
M1a: 近中性 Near neutral	42	-32 456.56	$p_0 = 0.33742$, $\omega_0 = 0.22678$; $p_1 = 0.66258$, $\omega_1 = 1$	不允许 Not allowed
M2a: 选择 Positive selection	44	-32 436.05	$p_0 = 0.30905$, $\omega_0 = 0.21821$; $p_1 = 0.63053$, $\omega_1 = 1$; $p_2 = 0.06041$, $\omega_2 = 2.00192$	22I、24S、89Q、267Y、465V
M5: γ	42	-32 379.20	$a = 1.67575$, $b = 2.46596$	22I、24S、89Q、116D、 200Q、238V、267Y、465V、 472D、491G、492A
M7: β	42	-32 370.36	$p = 0.73638$, $q = 0.44922$	不允许 Not allowed
M8: β 和 ω	44	-32 357.67	$p_0 = 0.90369$, $p = 0.70929$; $q = 0.53061$, $p_1 = 0.09631$, $\omega = 1.55714$	22I、24S、267Y、465V
M8a: β 和 ω , 且 $\omega = 1$	43	-32 370.65	$p_0 = 0.67148$, $p = 0.85174$; $q = 1.03963$, $p_1 = 0.32852$, $\omega = 1$	不允许 Not allowed

注:仅列入后验概率大于95%的正选择位点,估算结果以Naïve empirical Bayes(NEB)分析为准。

Note: Only sites with probabilities greater than 95% are listed. All results were based on Naïve empirical Bayes (NEB) analysis.

表5 各正选择模型与原假设模型间的似然比值检验统计量($2\Delta\ln L$)

Table 5 Likelihood ratio test statistics between each positive selection model and null hypothesis model ($2\Delta\ln L$)

模型比较 Comparison between models	$2\Delta\ln L$	$d.f.$	χ^2_{f}
M1a vs. M2a	41.02	2	9.2103
M7 vs. M8	25.38	2	9.2103
M8a vs. M8	25.96	1	6.6349

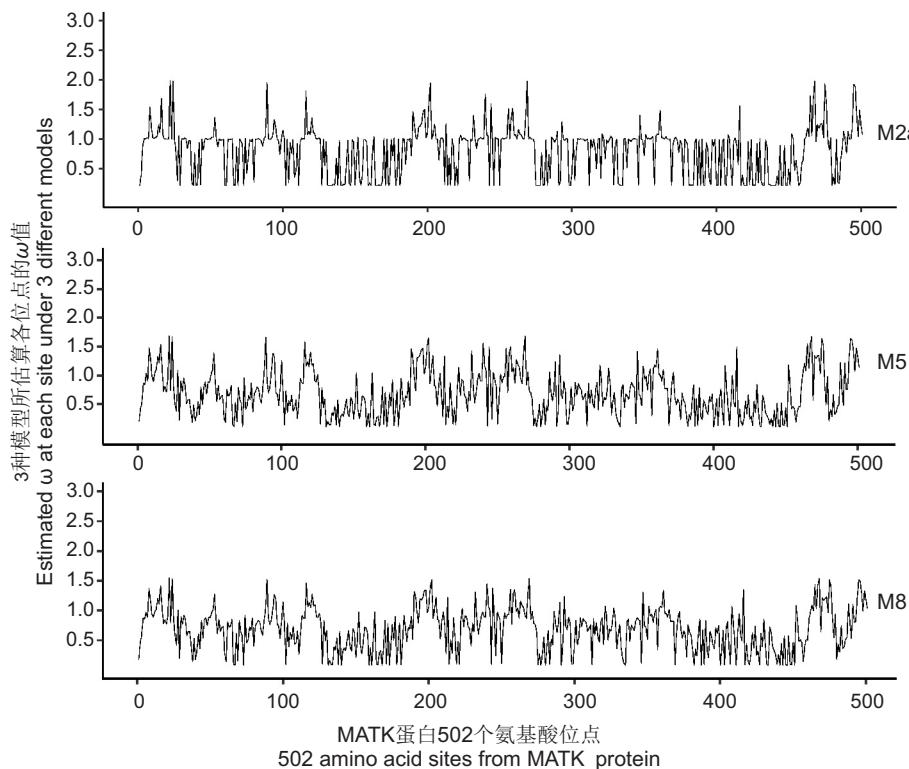
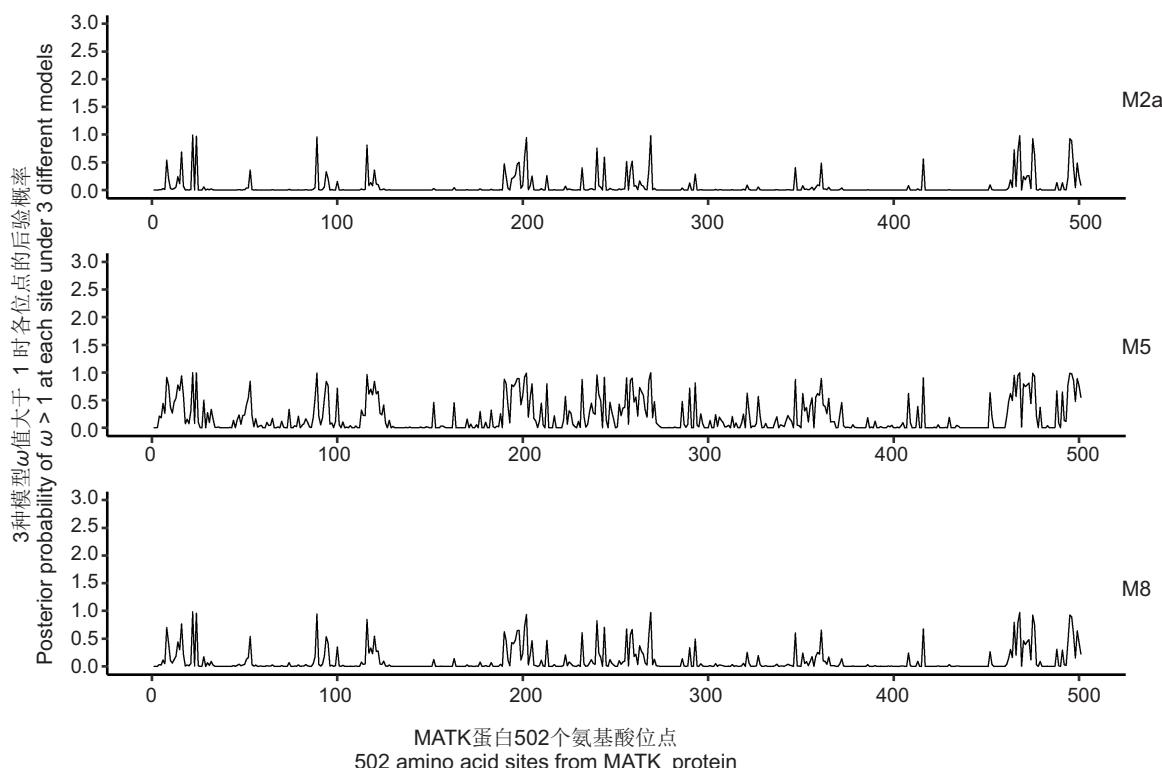
图 3 3 种模型 NEB 算法下估算 MATK 蛋白各位点的 ω 值Fig. 3 Estimated ω at each site in MATK protein based on NEB analysis under three different models图 4 3 种模型 NEB 算法下估算 MATK 蛋白各位点 ω 值大于 1(即正选择)的后验概率Fig. 4 Posterior probabilities of ω values greater than 1 (positively selected) at each site of MATK protein based on NEB analysis under three different models

表 6 MATK 蛋白内部共进化对位点及参数
Table 6 Estimated parameters and co-evolution pairs of MATK protein

共进化对 Co-evolution pair	1号氨基酸位点 Amino acid site 1	2号氨基酸位点 Amino acid site 2	相关系数 Correlation coefficients (%)	Bootstrap (%)
1	27F	124N	81.42	94.47
2	40Y	285I	92.72	95.73
3	41S	70G	76.79	91.02
4	43A	383S	82.78	95.55
5	43A	417S	81.85	93.70
6	43A	424G	75.67	92.29
7	45K	458E	82.64	99.31
8	84E	383S	83.86	95.19
9	101G	452Q	78.91	92.89
10	121K	330F	72.38	91.86
11	133H	367P	99.98	95.09
12	135L	458E	71.45	94.88
13	145K	264V	77.75	99.07
14	183V	261A	70.26	90.22
15	248K	422R	67.03	92.43
16	282K	309F	75.72	91.79
17	383S	417S	86.07	91.80
18	417S	424G	71.16	95.18
19	422R	423D	80.96	98.16

变，以维持蛋白质特定的功能。

3 讨论

在系统发育关系的构建中，*matK* 虽不如 *psaA* 基因，但要优于传统的 *rbcL* 基因，具有一定的应用价值。另外，联合数据集进行系统发育关系重建的可信度比单基因数据集更高，将来利用多基因联合数据集对蕨类植物系统发育关系进行研究能更加真实的了解蕨类植物的进化动态。

本研究中，阴地蕨属的 3 个成员 (*B. ternatum*、*B. japonicum* 和 *B. virginianum*) 以较高后验概率聚为一支，其中阴地蕨先发生分歧，该分支后验概率为 $P = 100\%$ (图 1，节点 38)，分歧时间为 0.46 Mya(第四纪更新世卡拉布里期)，华东阴地蕨与蕨萁互为姊妹群，后验概率为 $P = 99.15\%$ (图 1，节点 39)，分歧时间为 0.06 Mya(第四纪更新世卡拉布里期)。阴地蕨属 3 个物种的分化事件均发生在第四纪更新世卡拉布里期，说明该时期的环境因素可能导致阴地蕨属植物发生物种分化，今后有望通过对其他阴地蕨属物种深入研究来进行验证。

研究表明，*matK* 基因在被子植物、裸子植物、石松类植物和苔藓类植物中均发生过适应性进

化^[28, 29]，而王博等^[30]未在蕨类植物中发现 *matK* 基因的正选择信号。本研究发现，蕨类植物 *matK* 基因存在具备统计学显著性的正选择信号，出现这种差异的原因可能和早先有关蕨类植物 *psaA* 基因的研究类似：第一，之前关于 *matK* 适应性进化研究的物种数目太少而出现假阴性结果；第二，*matK* 基因中的正选择信号可能来自本研究中新增加的物种，特别是阴地蕨属的 3 个物种。上述情况有望将进一步对阴地蕨属其他成员的叶绿体基因组测序而得到揭示。

研究表明，MATK 蛋白处于强烈的进化限制中，其二级结构非常保守^[31]。与其他光合作用重要基因相比，蕨类植物 *matK* 基因中经历正选择的氨基酸位点较少，这可能与 MATK 蛋白的结构和功能高度保守有关。将来对这些正选择位点(模型 M8，位点 22、24、267 和 465)的研究可望能解释叶绿体 II 型内含子剪切酶结构和功能的变化。

氨基酸位点协同进化是在蛋白质中和蛋白质间普遍存在的现象。本研究发现 MATK 蛋白内存在的 18 对共进化氨基酸位点，其中共进化对 45~458 和 145~264 在蕨类植物 MATK 蛋白的进化历程中可能起到重要作用(Bootstrap 值大于 99%)。值得注意的是，MATK 蛋白并没有像 PSAA 蛋白一

样通过正选择位点形成共进化组来调整蛋白的适合度(fitness)，这可能与 *matK* 基因进化速率有关。随着今后对特定蕨类植物类群，如阴地蕨属的深入研究，将进一步揭示叶绿体基因适应性进化的内在规律。

参考文献:

- [1] 李春香, 王怿, 孙晓燕. 蕨类植物的起源演化: 对“古老”类群的重新审视 [J]. 生命科学, 2007, 19(2): 245–249.
Li CX, Wang Y, Sun XY. Origin and evolution of Pteridophytes: new insights to “ancient” lineage [J]. *Chinese Bulletin of Life Sciences*, 2007, 19(2): 245–249.
- [2] 刘红梅, 王丽, 张宪春, 曾辉. 石松类和蕨类植物研究进展: 兼论国产类群的科级分类系统 [J]. 植物分类学报, 2008, 46(6): 808–829.
Liu HM, Wang L, Zhang XC, Zeng H. Advances in the studies of lycophytes and monilophytes with reference to systematic arrangement of families distributed in China [J]. *Journal of Systematics and Evolution*, 2008, 46(6): 808–829.
- [3] 张宪春, 卫然, 刘红梅, 何丽娟, 王丽, 张钢民. 中国现代石松类和蕨类的系统发育与分类系统 [J]. 植物学报, 2013, 48(2): 119–137.
Zhang XC, Wei R, Liu HM, He LJ, Wang L, Zhang GM. Phylogeny and classification of the extant lycophytes and ferns from China [J]. *Chinese Bulletin of Botany*, 2013, 48(2): 119–137.
- [4] Smith AR, Pryer KM, Schuettpelz E, Korall P, Schneider H, Wolf PG. A classification for extant ferns [J]. *Taxon*, 2006, 55(3): 705–731.
- [5] Crane PR, Friis EM, Pedersen KR. The origin and early diversification of angiosperms [J]. *Nature*, 1995, 374(6517): 27–33.
- [6] Schuettpelz E, Pryer KM. Evidence for a Cenozoic radiation of ferns in an angiosperm-dominated canopy [J]. *Proc Natl Acad Sci USA*, 2009, 106(27): 11200–11205.
- [7] 许可, 王博, 苏应娟, 高磊, 王艇. 蕨类植物 *psbD* 基因的适应性进化和共进化分析 [J]. 植物科学学报, 2013, 31(5): 429–438.
Xu K, Wang B, Su YJ, Gao L, Wang T. Molecular evolution of *psbD* gene in ferns: selection pressure and co-evolutionary analysis [J]. *Plant Science Journal*, 2013, 31(5): 429–438.
- [8] 森林, 苏应娟, 张冰, 王艇. 凤尾蕨科植物 *rbcL* 基因的适应性进化分析 [J]. 热带亚热带植物学报, 2010, 18(1): 1–8.
Sen L, Su YJ, Zhang B, Wang T. Adaptive evolution of the *rbcL* gene in Pteridaceous ferns [J]. *Journal of Tropical and Subtropical Botany*, 2010, 18(1): 1–8.
- [9] 周媛, 王博, 高磊, 王艇. 凤尾蕨科旱生蕨类 *rbcL* 基因的适
应性进化和共进化分析 [J]. 植物科学学报, 2011, 29(4): 409–416.
Zhou Y, Wang B, Gao L, Wang T. Adaptive evolution and coevolution of the *rbcL* gene in xeric Pteridaceae ferns [J]. *Plant Science Journal*, 2011, 29(4): 409–416.
- [10] Sen L, Fares MA, Su Y, Wang T. Molecular evolution of *psbA* gene in ferns: unraveling selective pressure and co-evolutionary pattern [J]. *BMC Evol Biol*, 2012, 12(1): 145–145.
- [11] 吴筱婷, 森林, 陈楠, 张潇, 马朝霞, 张钦宇. 蕨类植物 *psaA* 基因的分子进化研究 [J]. 植物科学学报, 2017, 35(2): 177–185.
Wu XP, Sen L, Chen N, Zhang X, Ma ZX, Zhang QY. Study on the molecular evolution of the *psaA* gene from ferns [J]. *Plant Science Journal*, 2017, 35(2): 177–185.
- [12] Sen L, Fares MA, Liang B, Gao L, Wang B, et al. Molecular evolution of *rbcL* in three gymnosperm families: identifying adaptive and coevolutionary patterns [J]. *Biol Direct*, 2011, 6(1): 1–29.
- [13] 森林, 余坤, 胡志刚, 汪文杰, 徐雷, 等. 裸子植物 *psbA* 基因分子进化式样的研究 [J]. 热带亚热带植物学报, 2016, 24(2): 151–159.
Sen L, Yu K, Hu ZG, Wang WJ, Xu L, et al. Molecular evolutionary patterns of the *psbA* gene in gymnosperms [J]. *Journal of Tropical and Subtropical Botany*, 2016, 24(2): 151–159.
- [14] Vogel J, Borner T, Hess WR. Comparative analysis of splicing of the complete set of chloroplast group II introns in three higher plant mutants [J]. *Nucleic Acids Res*, 1999, 27(19): 3866–3874.
- [15] Kuo LY, Li FW, Chiou WL, Wang CN. First insights into fern *matK* phylogeny [J]. *Mol Phylogenetic Evol*, 2011, 59(3): 556–566.
- [16] Wei R, Yan YH, Harris AJ, Kang J, Shen H, et al. Plastid phylogenomics resolve deep relationships among Eupolypod II ferns with rapid radiation and rate heterogeneity [J]. *Genome Biol Evol*, 2017, 9(6): 1646–1657.
- [17] Li FW, Kuo LY, Rothfels CJ, Ebihara A, Chiou WL, et al. *rbcL* and *matK* earn two thumbs up as the core DNA barcode for ferns [J]. *PLoS One*, 2011, 6(10): e26597.
- [18] Shao YZ, Wei R, Zhang XC, Xiang QP. Molecular phylogeny of the cliff ferns (Woodsiaceae: Polypodiales) with a proposed infrageneric classification [J]. *PLoS One*, 2015, 10(9): e0136318.
- [19] Chao YS, Mustapeng AMA, Chen CW, Chiou WL. *Pteris borneensis* (Pteridaceae), a new species from Borneo, with re-circumscription of *Pteris decrescens* and *Pteris Parviloba* [J]. *Syst Bot*, 2017, 42(4): 724–732.
- [20] Liu ZY, Wei HJ, Shang H, Wei R, Wang P, et al. *Diplazium yinchanianum* (Athyriaceae): a new fern from the

- border between China and Vietnam [J]. *Phytotaxa*, 2018, 343(2): 139–148.
- [21] 李艳, 江玉梅, 鲁顺保, 彭九生, 朱笃. 突托腊梅ISSR引物反应条件的优化与筛选[J]. 植物科学学报, 2008, 26(3): 245–250.
- Li Y, Jiang YM, Lu SB, Peng JS, Zhu D. Optimization of experiment conditions and primer screening with ISSR markers for *Chimonanthus grammatus*[J]. *Plant Science Journal*, 2008, 26(3): 245–250.
- [22] Kumar S, Stecher G, Tamura K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets [J]. *Mol Biol Evol*, 2016, 33(7): 1870–1874.
- [23] Darriba D, Taboada GL, Doallo R, Posada D. jModelTest 2: more models, new heuristics and parallel computing [J]. *Nat Methods*, 2012, 9(8): 772–772.
- [24] Posada D. jModelTest: phylogenetic model averaging[J]. *Mol Biol Evol*, 2008, 25(7): 1253–1256.
- [25] Santorum JM, Darriba D, Taboada GL, Posada D. jModelTest.org: selection of nucleotide substitution models on the cloud [J]. *Bioinformatics*, 2014, 30(9): 1310–1311.
- [26] Drummond AJ, Suchard MA, Xie D, Rambaut A. Bayesian phylogenetics with BEAUti and the BEAST 1.7 [J]. *Mol Biol Evol*, 2012, 29(8): 1969–1973.
- [27] Xu B, Yang Z. PamlX: A graphical user interface for PAML [J]. *Mol Biol Evol*, 2013, 30(12): 2723–2724.
- [28] Hao DC, Chen SL, Xiao PG. Molecular evolution and positive Darwinian selection of the chloroplast maturase matK[J]. *J Plant Res*, 2010, 123(2): 241–247.
- [29] Hao DC, Mu J, Chen SL, Xiao PG. Physicochemical evolution and positive selection of the gymnosperm matK proteins[J]. *J Genet*, 2010, 89(1): 81–89.
- [30] 王博, 高磊, 苏应娟, 王艇. 基于叶绿体基因组全序列分析真叶植物叶绿体基因的适应性进化[J]. 中山大学学报(自然科学版), 2012, 51(3): 108–113.
- Wang B, Gao L, Su YJ, Wang T. Adaptive evolutionary analysis of chloroplast gene in Euphylophytes based on complete chloroplast genome sequences[J]. *Acta Scientiarum Naturalium Universitatis Sunyateni*(Natural Science Edition), 2012, 51(3): 108–113.
- [31] Barthet MM, Hilu KW. Evaluating evolutionary constraint on the rapidly evolving gene matK using protein composition[J]. *J Mol Evol*, 2008, 66(2): 85–97.

(责任编辑:周媛)