

# 线粒体 *nad1* 基因内含子在稻族系统学研究中的价值 ——兼论 *Porteresia* 的系统位置

郭亚龙 葛颂\*

(中国科学院植物研究所系统与进化植物学重点实验室 北京 100093)

## The utility of mitochondrial *nad1* intron in phylogenetic study of Oryzeae with reference to the systematic position of *Porteresia*

GUO Ya-Long GE Song

(Laboratory of Systematic and Evolutionary Botany, Institute of Botany, the Chinese Academy of Sciences, Beijing 100093, China)

**Abstract** In the study of plant systematics and evolution, the gene sequences of mitochondrial genome are less used for their low informative sites and complex structure compared to those of the nuclear and chloroplast genomes. In this study, we sequenced the intron 2 of mitochondrial *nad1* gene to study the phylogeny of Oryzeae, in particular emphasizing the systematic position of *Porteresia*, a monotypic genus of Oryzeae with the only species *P. coarctata*. The results indicate that the tribe Oryzeae should be divided into two subtribes (Oryzinae and Zizaniinae) and that *Leersia* is the most closely related genus to *Oryza*. The present data clearly demonstrate that *Porteresia coarctata* should be treated as a member in *Oryza* rather than as a separate genus. The sequences of *nad1* intron 2 are informative enough in Oryzeae, and thus appear appropriate for the phylogenetic study at higher taxonomic levels. The gaps, when treated appropriately, can provide additional information in phylogenetic study.

**Key words** *nad1* intron, gap, mitochondrial DNA, *Porteresia coarctata*, Oryzeae, phylogeny.

**摘要** 在植物的系统与进化研究中,相对于叶绿体和核基因片段来说,线粒体基因片段由于其分辨率较低、结构复杂而较少应用到系统学研究中。本研究选用 *nad1* 基因的第 2 内含子序列探讨了稻族 Oryzeae 的系统发育关系,并着重分析了单型属 *Porteresia* 的系统学位置,并在此基础上探讨了 *nad1* 基因内含子序列的系统学意义以及空位性状在系统发育重建中的价值。本研究结果支持将稻族分为两个亚族 (Oryzinae 和 Zizaniinae) 的处理,稻族中假稻属 *Leersia* 与稻属 *Oryza* 最近缘, *Porteresia coarctata* 不应该单独成属而应处理为稻属的成员。研究表明, *nad1* 第 2 内含子具有一定的系统发育信息,尤其是在较高分类等级的系统发育重建中具有较大价值,同时表明空位的合理利用将在已有序列信息基础上提供更多的系统发育信息。

**关键词** *nad1* 内含子;空位;线粒体 DNA; *Porteresia coarctata*;稻族;系统发育

由于分子证据的广泛应用,植物系统学在过去 20 多年中取得了巨大的进步。迄今,

2004-02-11 收稿,2004-06-10 收修改稿。

基金项目:中国科学院“十五”重要方向项目(KSCXZ-SW-101A)、国际科技合作重点项目计划(2001CB711103)、国家杰出青年科学基金(30025005) Supported by the Chinese Academy of Sciences (Grant No. KSCXZ-SW-101A), the Program for Key International Science and Technique Cooperation Project of China (Grant No. 2001CB711103) and the National Natural Science Foundation of China (Grant No. 30070057)。

\* 通讯作者 (Author for correspondence. E-mail: gesong@ibcas.ac.cn)

在植物系统发育重建中,来自叶绿体基因组的 DNA 片段和来自核基因编码核糖体的 DNA 片段(nuclear ribosomal DNA)是两个最主要的分子证据来源(Olmstead & Palmer, 1994; Soltis & Soltis, 1998; Alvarez & Wendel, 2003)。近年来,单拷贝或低拷贝的核基因片段被越来越多地用于不同等级的系统发育重建研究中,为系统发育重建研究提供了更为重要的信息来源(Clegg et al., 1997; Ge et al., 1999; Sang, 2002)。相比之下,利用来自植物线粒体基因组的 DNA 片段进行系统发育研究仍十分有限,一方面是因为植物线粒体基因组的进化速率比叶绿体基因组更慢从而提供的进化信息有限(Wolfe et al., 1987),另一方面由于植物线粒体基因组在大小和结构上十分不稳定(高频重排、重复或缺失以及外源 DNA 的转入等),这些都制约了其在植物系统发育研究中的广泛应用(Palmer, 1992; Adams & Palmer, 2003)。但最新的研究表明,植物线粒体基因组中有些片段是相对比较稳定的,如与呼吸代谢有关的基因(包括 *cob*、*cox1*、*nad1*、*nad4* 和 *nad5*) (Adams & Palmer, 2003),而且植物线粒体中许多基因存在进化较快的内含子,加上被子植物线粒体是母系遗传,因此某些线粒体 DNA 片段信息同样提供了重要的进化信息(Demesure et al., 1995),在植物系统发育重建、居群生物学和生物地理学研究中许多成功的报道(Qiu & Palmer, 1999; Mitton et al., 2000; Gugerli et al., 2001; Sanjur et al., 2002)。

稻族 *Oryzae* 是禾本科 *Poaceae* 中与人类生活密切相关的一个族,包括水稻所在的稻属 *Oryza* L. 和茭白 *Zizania latifolia* (Griseb.) Turcz. ex Stapf 所在的菰属 *Zizania* L. 等大约 12 个属 70 余种,广泛分布在全球热带和温带地区(Clayton & Renvoize, 1986; Vaughan, 1994)。尽管迄今对稻族采用不同手段进行了一系列研究(Hitchcock & Chase, 1951; Hutchinson, 1959; Terrell & Robinson, 1974; Zhang & Second, 1989; Duvall et al., 1993; Ge et al., 2002),其作为一个单系类群已不存在疑问(Ge et al., 2002),但稻族中一些基本的分类和系统发育问题仍有待澄清,包括族下的分类处理、属的界限和属间关系等(Terrell et al., 2001; Ge et al., 2002)。例如, *Porteresia coarctata* (Roxb.) Tateoka 曾经被认为是稻属的 1 个种(*O. coarctata* Roxb.), 后来由于形态上的某些差异被处理为单型属(Tateoka, 1965; Vaughan, 1989),但近来的分子证据表明这种处理是有疑问的, *P. coarctata* 不应该处理为单型属而应重新置于稻属中(Ge et al., 1999, 2002),当然这仍需要来自不同方面证据的验证。

*nad1* 基因编码线粒体呼吸传递链的复合体 I——泛醌氧化还原酶(NADH:ubiquinone oxidoreductase, 也可称 NADH 脱氢酶)的 *nad1* 亚基,是一个结构较为稳定的线粒体基因,在迄今已研究过的物种中由 5 个外显子构成(*nad1/A-E*) (Gutierrez et al., 1999)。本文通过对稻族 11 属 19 个代表种线粒体 *nad1* 第 2 内含子序列的测定和系统发育分析,从线粒体基因组角度探讨了稻族的族下分类和属间关系,尤其是对 *P. coarctata* 的系统位置进行了分析,同时探讨了 *nad1* 第 2 内含子在稻族系统发育研究中的价值,为该片段在系统学研究中的进一步应用提供资料。

## 1 材料和方法

### 1.1 实验材料

本研究所选取类群包括 19 个种代表稻族,其中包括与稻属关系较为密切且有争议的

*Porteresia* Tateoka 这一单型属。根据 GPWG (2001) 对禾本科的最新处理, 稻族 Oryzaceae 是稻亚科 Ehrhartoideae 三个族之一, 而稻亚科和竹亚科 Bambusoideae 是姊妹群关系, 因此我们选取竹亚科刚竹属 *Phyllostachys* Sieb & Zucc. 的 *P. aurea* Riviere & C. Riviere 作为外类群。本研究所涉及的材料均为硅胶干燥的叶片材料(表 1), 除 *O. granulata* Nees & Arn. ex Watt、*Zizania latifolia* 和外类群为野外采集外, 其余均由菲律宾的国际水稻研究所(IRRI) 种质资源库提供。

表 1 植物样品材料及其基因库序列号

Table 1 The species used in the study and their GenBank accession number

物种 Species	样品编号 Accession number	来源 Provenance	染色体数 2n	基因库序列号 GenBank number
亚洲栽培稻 <i>Oryza sativa</i> L.	30416	Philippines	24	AY507930
斑点野稻 <i>O. punctata</i> Kotschy ex Steud.	103887	Cameroon	24	AY507931
药用野稻 <i>O. officinalis</i> Wall ex Watt	105085	Philippines	24	AY507932
长颖野稻 <i>O. longiglumis</i> Jansen	106525	Papua New Guinea	48	AY507933
短花药野稻 <i>O. brachyantha</i> A. Chev. & Roehr.	105151	Sierra Leone	24	AY507934
<i>Porteresia coarctata</i> (Roxb.) Tateoka	104502	Bangladesh	48	AY507935
西来特野稻 <i>O. schlechteri</i> Pilger	82047	Papua New Guinea	48	AY507936
澳洲野稻 <i>O. australiensis</i> Domin.	101144	Australia	24	AY507937
疣粒野稻 <i>O. granulata</i> Nees & Arn. ex Watt	2422	Hainan, China	24	AY507938
<i>Leersia tisserantii</i> (A. Chev.) Launert	105610	Cameroon	24	AY507939
<i>L. hexandra</i> SW.	105252	Philippines	48	AY507940
茭白 <i>Zizania latifolia</i> (Griseb.) Turcz. ex Stapf	GS 0202	Beijing, China	34	AY507941
<i>Zizaniopsis villanensis</i> Quarin	85425	Argentina	24	AY507942
<i>Prospyrtchloa prehnensis</i> Schweick	-	S. Africa	24	AY507943
<i>Potamophila parviflora</i> R. Br.	85424	Australia	24	AY507944
<i>Rhynchoryza subulata</i> (Nees) Baillon	100913	Argentina	24	AY507945
<i>Luziola leiocarpa</i> Lindm.	82043	Argentina	24	AY507946
<i>Hygroryza aristata</i> (Retz.) Nees	105460	Sri Lanka	24	AY507947
<i>Chikusichloa aquatica</i> Koidzumi	106186	Japan	24	AY507948
早园竹 <i>Phyllostachys aurea</i> Riviere & C. Riviere	-	Beijing, China	48	AY507949

## 1.2 DNA 提取、PCR 扩增、纯化和测序

总 DNA 的提取参照 Ge et al. (1999) 的方法。在小麦 *Triticum aestivum* L. 中 *nad1* 的外显子 2 (exon 2) 为 82 bp, 内含子 2 (intron 2) 为 1423 bp, 外显子 3 为 192 bp (Chapdelaine & Bonen, 1991)。本研究选取 *nad1* 基因的第 2 内含子进行序列测定。扩增 *nad1* 的第 2 内含子的引物见表 2。*nad1* 的扩增体系均为 25  $\mu$ L, 反应体系包括 10–20 ng 的总 DNA 模板, 10 mmol/L Tris-HCl (pH 8.3) 2.0 mmol/L MgCl<sub>2</sub> 4 种 dNTP 各 200  $\mu$ mol/L 2 个引物各 5 pmol/L 0.75 单位的 *Taq* 聚合酶(宝生物)。扩增反应在 PTC-200(PE) PCR 扩增仪上进行。扩增反应程序如下: 70  $^{\circ}$ C 4 min  $\rightarrow$  94  $^{\circ}$ C 1 min, 50  $^{\circ}$ C 20 s, 72  $^{\circ}$ C 1.5 min, 2 个循环  $\rightarrow$  94  $^{\circ}$ C 20 s, 50  $^{\circ}$ C 20 s, 72  $^{\circ}$ C 1.5 min, 35 个循环  $\rightarrow$  72  $^{\circ}$ C 10 min。PCR 的产物在 1.5% 的琼脂糖胶上电泳分离, 目的片段割胶后使用 DNA 纯化试剂盒(Promega, Madison, WI) 纯化。

测序的反应体系(10  $\mu$ L): 参考 ABI Prism BigDye<sup>TM</sup> Terminator Cycle Sequencing Ready

Reaction Kit (Perkin Elmer): BigdyeMix 1.5  $\mu\text{L}$ , 引物 5 pmol, 模板 25–50 ng。测序反应的条件如下: 94  $^{\circ}\text{C}$  15 s, 52  $^{\circ}\text{C}$  8 s, 60  $^{\circ}\text{C}$  4.5 min, 29 个循环。测序反应产物的纯化如下: 将测序反应的产物转入含有 25  $\mu\text{L}$  无水乙醇: 3 mol/L NaAC = 24:1 的溶液中, 冰浴 10 min。4  $^{\circ}\text{C}$  12000 r/min 离心 20 min, 弃上清液, 然后加入 150  $\mu\text{L}$  70% 的乙醇洗涤沉淀。4  $^{\circ}\text{C}$  12000 r/min 离心 10 min, 弃上清液, 避光吹干, 加入 3  $\mu\text{L}$  Loading buffer, 用 ABI-377 DNA 测序仪收集数据。

表 2 所用扩增及测序引物(引物名有下划线表示该引物仅用于测序)

Table 2 The primers used in this study (The primer underlined is designed for sequencing only)

引物名 Primer name	方向 Direction	引物序列(5'–3') Primer sequence(5'–3')	参考文献 Reference
<i>nad1</i> exon B	正向 Forward	GCATTACGATCTGCAGCTCA	Demesure et al. (1995)
<i>nad1</i> exon C	反向 Reverse	GGAGCTCGATTAGTTTCTGC	Demesure et al. (1995)
<u><i>nad1</i> exon B-FF</u>	正向 Forward	GGATATACACCAGGCAAC	本文 This paper

### 1.3 数据处理

所得序列用 Clustal W (Thompson et al., 1994) 进行对位排列 (alignment)。碱基组成和 GC 含量、序列间成对距离 (pairwise distances) 等均用 MEGA 2.1 (Kumar et al., 2001) 计算。为了解位点变异在序列上的分布, 利用 DAMBE 4.1.19 (Xia & Xie, 2001) 计算作图。系统发育分析用 PAUP\* 4.0 (Swofford, 1998) 中的最大简约 (MP) 法和最大似然 (ML) 法进行。对排好的序列中所有的碱基同等加权并作为无序 (unordered) 性状。MP 法分析为 MULRARS 启发式搜寻 (Heuristic search), random stepwise addition 序列添加方式重复 1000 次, tree-bisection-reconnection (TBR) 做 branch swapping。ML 法依据 HKY 模式分析 (Hasegawa et al., 1985), 选择分支限界法 (branch and bound search) 搜寻 1000 次。拓扑结构的可靠性用 1000 次重复的自展检验来完成 (Felsenstein, 1985)。

由于序列数据中的空位 (gap) 往往含有系统发育信息, 曾被作为性状分别处理为不同的状态 (Simmons & Ochoterena, 2000)。但由于 ML 法尚没有合适的模型来处理空位性状 (Kawakita et al., 2003), 本研究仅用 MP 法分析空位性状。除按常规将其处理为缺失 (missing) 以外, 还根据 Simmons 和 Ochoterena (2000) 的方法将空位编码为有/无 (presence/absence) 性状, 即依据相邻空位更可能是 1 次进化事件的结果, 把单碱基 (占 1 个位点) 空位与多碱基 (占多个相邻位点) 空位均编码为 1, 无空位编码为 0, 并同序列数据一并计算。

## 2 实验结果

### 2.1 *nad1* 序列的特点

对稻族 19 种及 1 个外类群 20 个物种测序得到 1.44 kb 左右的 DNA 片段, 其中 5' 端从 *nad1* 基因第 2 内含子 33 bp 处起始, 不包括第 2 外显子, 这是因为测序起始部分可读性差, 3' 端则包括外显子 3 部分 60 bp。最短片段来自 *Luziola leiocarpa* (1371 bp), 最长片段来自 *Hygroryza aristata*、*Chikusichloa aquatica* 和外类群早园竹 (1451 bp)。稻族物种的 GC 含量都在 53.0%–54.4%, 平均值为 54.1%。

变异在序列位点上的分布情况见图 1,图中纵坐标表示某一位点的变异性( variability )用熵( entropy )或信息( information )来表示,单位是字节( byte ),其值越大表示变异性越高( Xia & Xie, 2001 )。横坐标表示序列的位置。由图 1 可见, *nad1* 序列的变异基本上均匀地分布在整個内含子上,但序列的 3' 端相对较为保守一些。

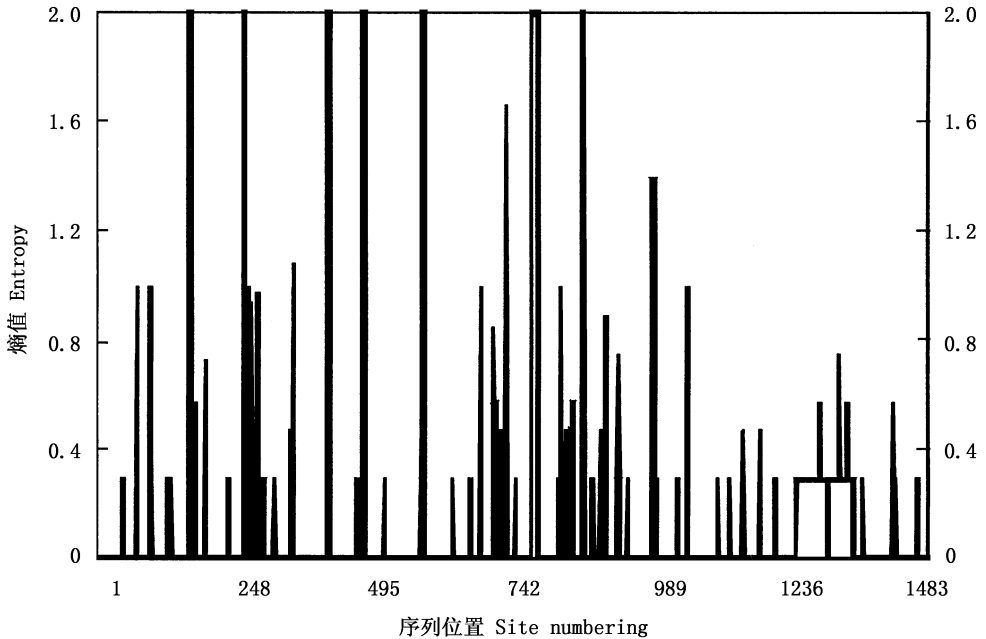


图 1 变异位点在序列上的分布  
Fig. 1. Variable substitution rate over sites.

对位排列后的序列长 1485 bp,不包括外类群时,共有 58 个核苷酸位点( 3.9% )为多态位点,其中 27 个位点( 1.8% )具有系统发育信息。当把空位作为有/无性状处理后可增加 18 个变异位点,其中 8 个具系统发育信息,分别位于序列 175 ( gap-1 )、474 - 478 ( gap-2 )、581 - 584 ( gap-3 )、709 - 715 ( gap-4 )、733 ( gap-5 )、852 - 854 ( gap-6 )、869 - 872 ( gap-7 )、996 - 1000 ( gap-8 )等位置,使变异位点和信息位点分别增加到 76 个( 5.1% )和 35 个( 2.3% )。这些空位性状中包括 2 个单碱基空位和 6 个多碱基空位。

## 2.2 系统发育分析

将空位作为缺失处理时,MP 法分析得到 18 棵相同的 MP 树,树长 79,除去非信息位点后的一致性指数( *CI* )为 0.8354,保持性指数( *RI* )为 0.8571。18 棵最大简约树的严格一致树见图 2。ML 法分析所得树的拓扑结构与 MP 法完全一致,但各分支的自展支持率略有不同(图 2)。由图 2 可见,稻族可分为两大支,第一分支具有很高的支持率( 98% ),包括稻属、*Porteresia* 和假稻属 3 个属,其中 2 个假稻属物种形成很高支持率的一支( 93% ),稻属和 *Porteresia coarctata* 则形成另一支( 58% )。第二分支包括稻族其他 8 个属,但自展支持率低于 50%( 48% )。在该分支中, *Luziola leiocarpa* 和 *Zizaniopsis villanensis* 聚为一支(支持率 98%), *Potamophila parviflora* 和 *Prospyrtochloa prehensilis* 聚为一支(支持率为 86%)。

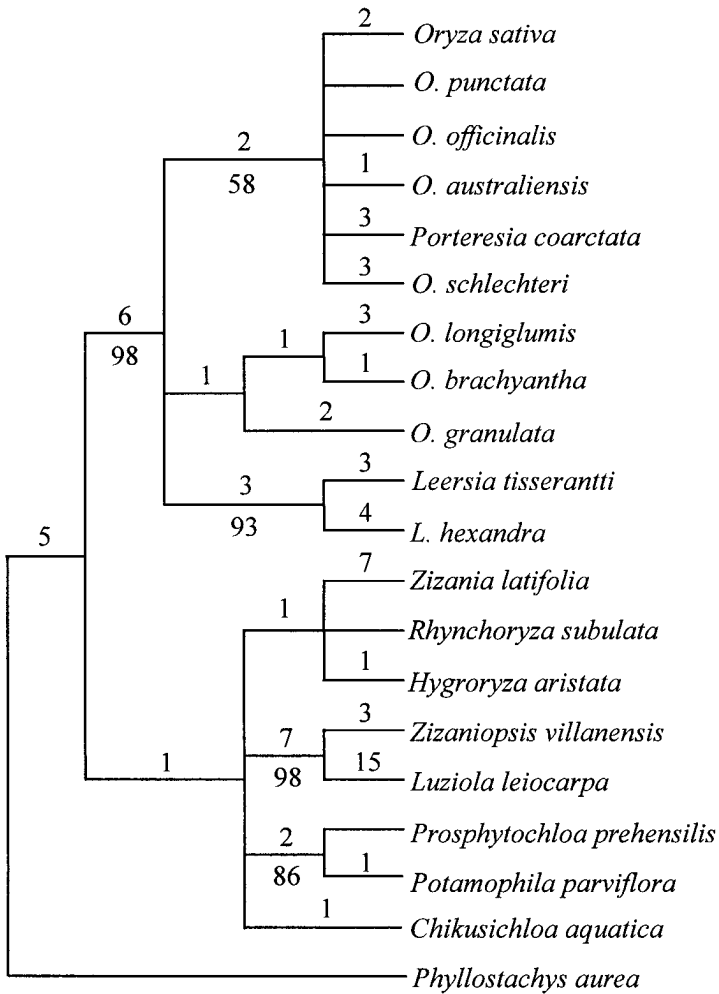


图2 *nad1* 序列最大简约性分析所得到的 18 棵树的严格一致树(空位处理为缺失) 树长 = 79,  $CI = 0.8354$ ,  $RI = 0.8571$ 。分支上下的数值分别表示置换数目和自展(bootstrap)支持率。

Fig. 2. The strict consensus tree of the 18 most parsimonious trees of *nad1* with the gaps treated as missing (Tree length = 79;  $CI = 0.8354$ ;  $RI = 0.8571$ ). Numbers above and below the branches represent nucleotide substitutions and bootstrap values, respectively.

当把空位作为有/无性状同序列数据一并分析时,MP 树从 18 棵降到 2 棵(图 3),树长 92,  $CI$  为 0.8152,  $RI$  为 0.8381。由图 3 可见,利用空位信息所得严格一致树的拓扑结构与将空位作为缺失处理时基本相同,但图 3 中的一些分支得到了更高的支持率,而且 *Porteresia coarctata* 和 *O. schlechteri* 聚为一支(支持率为 51%)。根据图 3 中空位信息的分布可见, *gap-5* 和 *gap-7* 是平行演化性状(homoplasy),而其余 6 个空位均为共有衍征(synapomorphy)。

### 3 讨论

#### 3.1 稻族的系统发育关系和 *Porteresia coarctata* 的系统位置

本文结果支持稻族分为两大支,第一支由稻属、假稻属和 *Porteresia* 构成,得到了很高

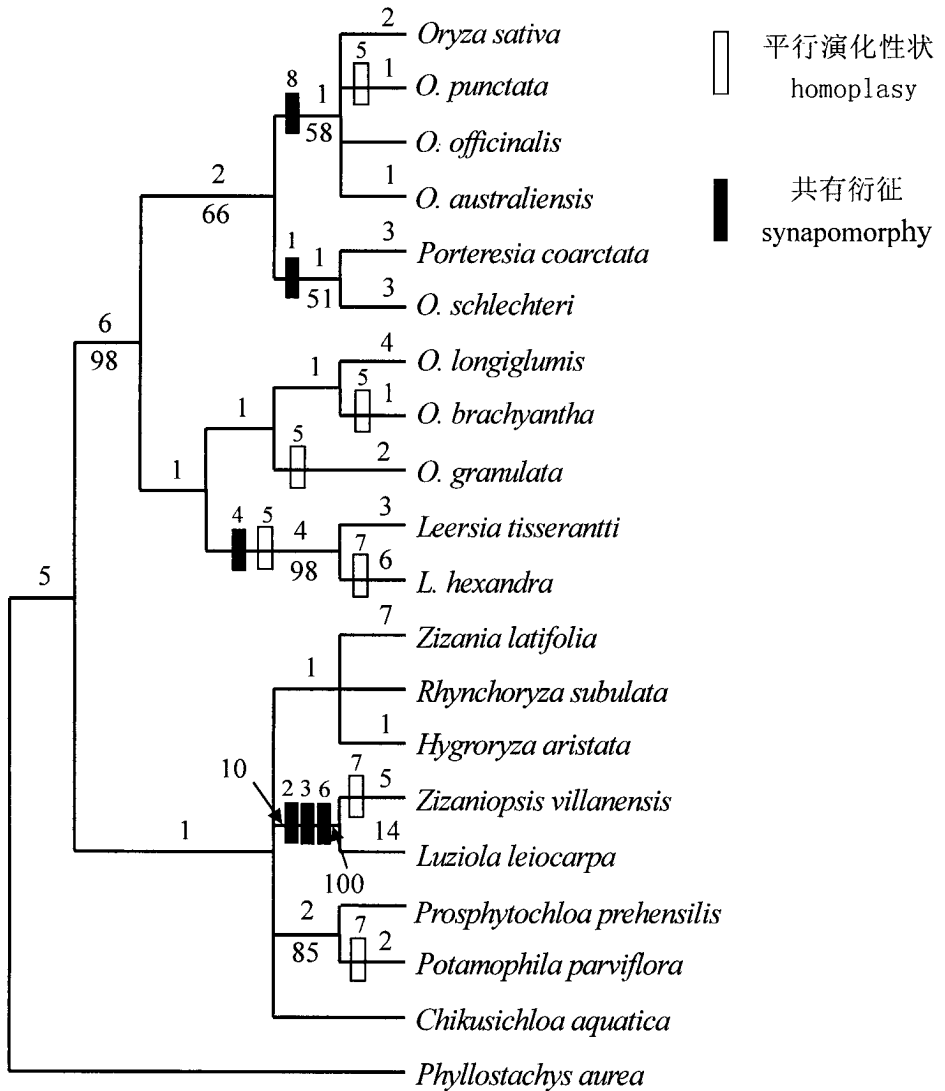


图3 *nad1* 序列最大简约性分析所得 2 棵树的严格一致树(空位处理为有/无性状) 树长 = 92,  $CI = 0.8152$ ,  $RI = 0.8381$ 。分支上下的数值分别表示间碱基置换数目和自展支持率。空白的和黑色的矩形框分别表示空位是平行演化性状还是共有衍征, 矩形框上的数字表示相应的空位。

Fig. 3. The strict consensus tree of the two most parsimonious trees of *nad1* with the gaps treated as present/absent characters (Tree length = 92;  $CI = 0.8152$ ;  $RI = 0.8381$ ). Numbers above and below the branches represent nucleotide substitutions and bootstrap values, respectively. The numbers above rectangles represent corresponding gap.

的支持(95%)。第二支由稻族其他 8 个属构成, 但统计支持率较低(48%), 这可能反映了这些属在进化历史上的快速分化。Ge 等(2002)利用叶绿体 *matK* 基因构建的系统发育树也体现了这一特点, 并推测稻族是一个快速分化(rapid radiation)的类群。*Luziola* Juss. 和 *Zizaniopsis* Doell. & Asch. 两属以及 *Potamophila* R. Br. 和 *Prophytochloa* Schweick. 两属关系非常密切, 均得到很高的统计支持(图 2 3)。本研究结果与以往叶绿体 *matK* 基因片段所得结果十分吻合(Ge et al., 2002), 说明稻族被分为两个亚族(Oryzinae 和 Zizaniinae)的处

理是可取的,而 Terrell 和 Robinson (1974) 把 *Luziola* 和 *Zizaniopsis* 单列为一个新亚族 (*Luziolinae*) 的建议得不到支持, 尽管这两个属关系十分密切(图 2.3)。如果不考虑 *Porteresia* 属 稻族中假稻属与稻属最近缘, 这一结果也得到其他方面证据的支持 (Terrell & Robinson, 1974; Duvall et al., 1993; Ge et al., 2002)。

*Porteresia coarctata* 曾是稻属中的成员 (*Oryza coarctata*), 后来依据某些形态上的差异被作为一个单型属处理 (Tateoka, 1965)。Zhang 和 Second (1989) 用叶绿体 DNA 限制性片段分析对稻族系统发育关系进行了初步研究, 发现 *P. coarctata* 位于稻属分支中并与 *O. meyeriana* (Zoll. & Mor. ex Steud.) Baill. 关系密切, 而其他 7 个属则形成两个截然不同的分支。通过 2 个核基因 (*Adh1* 和 *Adh2*) 及 *matK* 的测序分析, Ge 等 (1999) 发现 *P. coarctata* 无论是在 2 个核基因还是 *matK* 基因树上均嵌入稻属分支中, 进一步对稻族 11 个属叶绿体 *matK* 序列的分析表明, *P. coarctata* 以很高的支持率 (95%) 嵌入稻属中 (Ge et al., 2002)。尽管本研究 *nad1* 基因第 2 内含子在稻族中的系统发育信息有限 (2.3%), 但也很清楚地表明 *Porteresia coarctata* 和 *O. schlechteri* 关系较近且嵌入稻属其他类群中。因此, 线粒体片段明显不支持将 *Porteresia coarctata* 作为单型属的处理。

### 3.2 线粒体 *nad1* 基因第 2 内含子的系统学价值

在被子植物中, 线粒体基因的进化速率要比核基因慢 5 倍, 与哺乳动物线粒体基因比核基因进化至少快 5 倍形成巨大的反差 (Wolfe et al., 1987)。加上在基因组大小和结构上的不稳定, 植物线粒体基因片段在系统学中的应用远不如动物线粒体。尽管如此, 植物线粒体中仍然存在一些结构及功能上保守的基因, 并被用于大尺度的系统发育研究 (Wahleithner & Wolstenholme, 1988; Schuster et al., 1990; Qiu & Palmer, 1999; Wang et al., 2000; Meng et al., 2002; Adams & Palmer, 2003)。由于植物线粒体基因有内含子, 如 *nad1* 和 *nad4* 基因在所有已研究过的物种中都具有内含子, 因此 Demesure et al. (1995) 设计了一系列扩增线粒体片段的引物, 并在许多研究中得以应用, 其中包括扩增 *nad1* 基因外显子 B 和外显子 C 之间第 2 内含子的引物对 (*nad1* exon B 和 *nad1* exon C)。

本研究发现在稻族中线粒体 *nad1* 序列变异在整个第 2 内含子区是均匀分布的, 适合进行系统发育分析。为了探讨这一片段在不同类群水平上的替代速率, 我们比较了水稻及其近缘类群在 *nad1* 第 2 内含子序列上的差异 (表 3), 其中稻族以外的序列从 GenBank

表 3 水稻及其近缘类群 *nad1* 第 2 内含子序列的长度、GC 含量和序列差异

Table 3 Sizes, GC contents of and pairwise distance between *Oryza sativa* and its related taxa

物种 Species	序列间差异 Pairwise distance (%)	序列长度 Length of sequences	GC 含量 GC content (%)	基因库序列号 GenBank No.
<i>O. sativa</i> L.	—	1446	54.2	AY507930
<i>O. granulata</i> Nees & Arn. ex Watt	0.005	1442	54.1	AY507938
<i>Leersia tisserantii</i> (A. Chev.) Launert	0.008	1443	54.1	AY507939
<i>Phyllostachys aurea</i> Riviere & C. Riviere	0.011	1451	54.5	AY507949
<i>Triticum aestivum</i> L.	0.019	1449	54.4	X57967
<i>Cephalanthera damasonium</i> (Mill.) Druce	0.057	1456	54.6	AF314858
<i>Alocasia cucullata</i> (Lour.) G. Don	0.071	1228	54.0	AY243116



下载。结果表明,该片段在种间和属间序列差异很小,分别为 0.004% 和 0.008%;亚科和科间差异明显较大,亚科类群间序列差异从 1.1% 到 1.9%,科间差异从 5.7% 到 7.1%;而且分析表明稻族和表 3 中的植物类群的线粒体 *nad1* 序列变异在整个第 2 内含子区是均匀分布的,没有碱基偏好(分析结果图未列出),说明该片段在较高等级类群的系统发育重建中有更高的信息量,而且变异水平也更适于进行较高等级类群的系统发育重建。

本研究结果表明,相对于核基因和叶绿体基因来说,虽然线粒体由于变异速度低及其结构的不稳定在植物系统学研究中应用很少,但只要筛选合适的序列片段,线粒体基因组仍能为系统学研究提供新的证据。尤其是线粒体基因组在被子植物和裸子植物中均为母系遗传,同双亲遗传的核基因组序列进行对比分析,可以为进一步揭示系统发育关系和探讨物种形成机制提供重要信息。

### 3.3 空位性状的利用及其系统学价值

在进行系统发育分析时,不同物种同一 DNA 片段序列必须首先进行序列的对位排列,从而不可避免地出现或长或短的空位。一般情况下,在系统发育分析中空位常常被排除在系统发育分析之外(均作缺失状态)(Terry et al., 1997; Denton et al., 1998; Mason-Gamer et al., 1998),主要有以下三方面原因。首先,在关系较远的物种中产生空位的插入/缺失可能是多次起源的,因此空位性状不同源(Johnson & Soltis, 1995)。其次,如何处理空位性状仍存在争议,有人将其作为第 5 状态来处理(对氨基酸来说作为第 21 状态来处理),而另一些研究者则把它们作为不同于 4 个碱基的有/无(presence/absence)性状来处理(Simmons & Ochoterena, 2000)。第三,对多碱基空位如何处理仍缺乏一致意见,一些研究者把多碱基空位分别处理为多个独立性状(Wheeler & Hayashi, 1998; Giribet & Wheeler, 1999),而另一些研究者则把它们处理为单个插入/缺失性状(如 Lloyd & Calder, 1991; Barriol, 1994; Van Dijk et al., 1999)。已有的一些研究表明,不管空位出现在编码区还是非编码区都是很好的性状(以平行演化性状的水平来衡量),即具有很低的平行演化现象,可用于系统发育分析(Baldwin & Markos, 1998; Lloyd & Calder, 1991; Van Ham et al., 1994; Johnson & Soltis, 1995; Simmons et al., 2001)。

Simmons 和 Ochoterena(2000)主张,将空位编码为有/无性状是比较理想的,因为空位性状与碱基性状的性质完全不同。无论在编码区还是非编码区,多碱基空位通常是一次插入/缺失产生的(Pascarella & Argos, 1992; Gu & Li, 1995),故将多碱基空位编码为多个单独性状,意味着把每个空位均作为独立事件来处理,从而过度加权了一次插入/缺失事件。Simmons 等(2001)通过对 38 个片段(5 个结构 rDNA、5 个 ITS、6 个内含子、22 个外显子)的矩阵进行空位处理的比较研究认为,空位性状中有相当大的系统发育信息,空位性状加入到序列分析中能改变拓扑结构和提高严格一致树的分辨率和分支的支持率。本研究中利用 Simmons 和 Ochoterena(2000)提出的简单插入缺失编码方法,把单碱基空位与多碱基空位均编码为 1,无空位则编码为 0,结果表明空位性状确实能够提供一些系统发育信息,本研究中 16 个空位性状中,8 个具系统发育信息,只有 2 个出现平行演化现象。其中,*Porteresia coarctata* 和 *O. schlechteri* 在空位性状加入序列分析后才形成具有统计支持率的一支,而且空位性状的利用减少最大简约性系统发育树的数目并提高了大多数分支的支持率。因此,我们认为空位的合理利用将在已有序列信息基础上提供更多的系统发育

信息。

致谢 感谢菲律宾国际水稻研究所 (IRRI) 种质资源库为本研究提供材料 ;感谢张富民博士和包颖博士在数据处理和实验工作中的热忱帮助和指导 ;感谢中国科学院植物研究所李振宇研究员在样品采集和鉴定时的热忱指导。

### 参 考 文 献

- Adams K L, Palmer J D. 2003. Evolution of mitochondrial gene content: gene loss and transfer to the nucleus. *Molecular Phylogenetics and Evolution* 29: 380 – 395.
- Alvarez I, Wendel J F. 2003. Ribosomal ITS sequences and plant phylogenetic inference. *Molecular Phylogenetics and Evolution* 29: 417 – 434.
- Baldwin B G, Markos S. 1998. Phylogenetic utility of the external transcribed spacer (ETS) of 18S-26S rDNA: congruence of ETS and ITS trees of *Calycadenia* (Compositae). *Molecular Phylogenetics and Evolution* 10: 449 – 463.
- Barriel V. 1994. Molecular phylogenies and how to code insertion/deletion events. *Life Science* 317: 693 – 701.
- Chapdelaine Y, Bonen L. 1991. The wheat mitochondrial gene for subunit I of the NADH dehydrogenase complex: a trans-splicing model for this gene-in-pieces. *Cell* 65: 465 – 472.
- Clayton W D, Renvoize S A. 1986. Genera Graminum. *Kew Bulletin Additional Series XIII*: 1 – 389.
- Clegg M T, Cummings M P, Durbin M L. 1997. The evolution of plant nuclear genes. *Proceedings of the National Academy of Sciences, USA* 92: 7791 – 7798.
- Demesure B, Sodzi N, Petit R J. 1995. A set of universal primers for amplification of polymorphic non-coding regions of mitochondrial and chloroplast DNA in plants. *Molecular Ecology* 4: 129 – 131.
- Denton A L, McConaughy B L, Hall B D. 1998. Usefulness of RNA polymerase II coding sequences for estimation of green plant phylogeny. *Molecular Biology and Evolution* 15: 1082 – 1085.
- Duvall M R, Peterson P M, Terrell E E, Christensen A H. 1993. Phylogeny of North American Oryzoid grasses as construed from maps of plastid DNA restriction sites. *American Journal of Botany* 80: 83 – 88.
- Felsenstein J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39: 783 – 791.
- Ge S, Li A, Lu B-R, Zhang S-Z, Hong D-Y. 2002. A phylogeny of the rice tribe Oryzaceae (Poaceae) based on *matK* sequence data. *American Journal of Botany* 89: 1967 – 1972.
- Ge S, Sang T, Lu B-R, Hong D-Y. 1999. Phylogeny of rice genomes with emphasis on origins of allotetraploid species. *Proceedings of the National Academy of Sciences, USA* 96: 14400 – 14405.
- Giribet G, Wheeler W C. 1999. On gaps. *Molecular Phylogenetics and Evolution* 13: 132 – 143.
- GPWG (Grass Phylogeny Working Group). 2001. Phylogeny and subfamilial classification of the grasses (Poaceae). *Annals of the Missouri Botanical Garden* 88: 373 – 457.
- Gu X, Li W H. 1995. The size distribution of insertions and deletions in human and rodent pseudogenes suggests the logarithmic gap penalty for sequence alignment. *Journal of Molecular Evolution* 40: 464 – 473.
- Gugerli F, Senn J, Anzidei M, Madaghiele A, Büchler U, Sperisen C, Vendramin G G. 2001. Chloroplast microsatellites and mitochondrial *nad1* intron 2 sequences indicate congruent phylogenetic relationships among Swiss stone pine (*Pinus cembra*), Siberian stone pine (*Pinus sibirica*), and Siberian dwarf pine (*Pinus pumila*). *Molecular Ecology* 10: 1489 – 1497.
- Gutierrez S, Combettes B, Paepé R D, Mirande M, Lelandais C, Vedel F, Chétrit P. 1999. In the *Nicotiana sylvestris* CMSII mutant, a recombination-mediated change 5' to the first exon of the mitochondrial *nad1* gene is associated with lack of the NADH: ubiquinone oxidoreductase (complex I) NAD1 subunit. *European Journal of Biochemistry* 261: 361 – 370.
- Hasegawa M, Kishino H, Yano T. 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution* 22: 160 – 174.
- Hitchcock C, Chase A. 1951. *Manual of the Grasses of the United States*. 2nd ed. Washington, DC.: USDA Miscellaneous Publication.

- Hutchinson J. 1959. The Families of Flowering Plants. II. Monocotyledons. 2nd ed. London: Clarendon Press.
- Johnson L A, Soltis D E. 1995. Phylogenetic inference in Saxifragaceae sensu stricto and *Gilia* (Polemoniaceae) using *matK* sequences. *Annals of the Missouri Botanical Garden* 82: 149 – 175.
- Kawakita A, Sota T, Ascher J S, Ito M, Tanaka H, Kato M. 2003. Evolution and phylogenetic utility of alignment gaps within intron sequences of three nuclear genes in bumble bees (*Bombus*). *Molecular Biology and Evolution* 20: 87 – 92.
- Kumar S, Tamura K, Jakobsen I B, Nei M. 2001. MEGA2: Molecular Evolutionary Genetics Analysis software. *Bioinformatics* 17: 1244 – 1245.
- Lloyd D G, Calder V L. 1991. Multi-residue gaps, a class of molecular characters with exceptional reliability for phylogenetic analyses. *Journal of Evolutionary Biology* 4: 9 – 21.
- Mason-Gamer R J, Weil C F, Kellogg E A. 1998. Granule-bound starch synthase: structure, function, and phylogenetic utility. *Molecular Biology and Evolution* 15: 1658 – 1673.
- Meng S W, Chen Z D, Li D Z, Liang H X. 2002. Phylogeny of Saururaceae based on mitochondrial *matR* gene sequence data. *Journal of Plant Research* 115: 71 – 76.
- Mitton J B, Kreiser B R, Latta R G. 2000. Glacial refugia of limber pine (*Pinus flexilis* James) inferred from the population structure of mitochondrial DNA. *Molecular Ecology* 9: 91 – 97.
- Olmstead R G, Palmer J D. 1994. Chloroplast DNA systematics: a review of methods and data analysis. *American Journal of Botany* 81: 1205 – 1224.
- Palmer J D. 1992. Mitochondrial DNA in plant systematics: applications and limitations. In: Soltis P S, Soltis D E, Doyle J J eds. *Molecular Systematics of Plants*. New York: Chapman and Hall. 36 – 49.
- Pascarella S, Argos P. 1992. Analysis of insertions/deletions in protein structures. *Journal of Molecular Biology* 224: 461 – 471.
- Qiu Y-L, Palmer J D. 1999. Phylogeny of early land plants: insights from genes and genomes. *Trends in Plant Science* 4: 26 – 30.
- Sang T. 2002. Utility of low-copy nuclear gene sequences in plant phylogenetics. *Critical Reviews in Biochemistry and Molecular Biology* 37: 121 – 147.
- Sanjur O I, Piperno D R, Andres T C, Wessel-Beaver L. 2002. Phylogenetic relationships among domesticated and wild species of *Cucurbita* (Cucurbitaceae) inferred from a mitochondrial gene: implications for crop plant evolution and areas of origin. *Proceedings of the National Academy of Sciences, USA* 99: 535 – 540.
- Schuster W, Unsel M, Wissinger B, Brennicke A. 1990. Ribosomal protein S14 transcripts are edited in *Oenothera* mitochondria. *Nucleic Acids Research* 18: 229 – 233.
- Simmons M P, Ochoterena H. 2000. Gaps as characters in sequence-based phylogenetic analysis. *Systematic Biology* 49: 369 – 381.
- Simmons M P, Ochoterena H, Carr T G. 2001. Incorporation, relative homoplasy, and effect of gap characters in sequence-based phylogenetic analyses. *Systematic Biology* 50: 454 – 462.
- Soltis D E, Soltis P S. 1998. Choosing an approach and an appropriate gene for phylogenetic analysis. In: Soltis D E, Soltis P S, Doyle J J eds. *Molecular Systematics of Plants. II. DNA sequencing*. Boston: Kluwer Academic Publications.
- Swofford D L. 1998. PAUP\* 4.0: Phylogenetic Analysis Using Parsimony (\* and other methods). Beta version 4.0b4A. Sunderland: Sinauer Associates.
- Tateoka T. 1965. *Porteresia*, a new genus of Graminae. *Bulletin of National Science Museum, Tokyo* 8: 405 – 406.
- Terrell E E, Robinson H. 1974. *Luziolinae*, a new subtribe of oryzoid grass. *Bulletin of the Torrey Botanical Club* 101: 235 – 245.
- Terrell E E, Peterson P M, Wergin W P. 2001. Epidermal features and spikelet micromorphology in *Oryza* and related genera (Poaceae: Oryzaceae). *Smithsonian Contributions to Botany* 91: 1 – 50.
- Terry R G, Brown G K, Olmstead R G. 1997. Phylogenetic relationships in subfamily Tillandsioideae (Bromeliaceae) using *ndhF* sequences. *Systematic Biology* 22: 333 – 346.
- Thompson J D, Higgins D G, Gibson T J. 1994. Clustal W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. *Nucleic*

- Acids Research 22: 4673 – 4680.
- Van Dijk M A M, Paradis E, Catzefflis F, De Jong W W. 1999. The virtues of gaps: Xenarthran (Edentate) monophyly supported by a unique deletion in alphaA-crystallin. *Systematic Biology* 48: 94 – 106.
- Van Ham R C, Hart H, Mes T H, Sandbrink J M. 1994. Molecular evolution of noncoding regions of the chloroplast genome in the Crassulaceae and related species. *Current Genetics* 25: 558 – 566.
- Vaughan D A. 1989. The genus *Oryza* L. Current status of taxonomy. Philippines: International Rice Research Institute.
- Vaughan D A. 1994. The Wild Relatives of Rice: A Genetic Resources Handbook. Philippines: International Rice Research Institute.
- Wahleithner J A, Wolstenholme D R. 1988. Ribosomal protein S14 genes in broad bean mtDNA. *Nucleic Acids Research* 16: 6897 – 6913.
- Wang X-Q, Tank D C, Sang T. 2000. Phylogeny and divergence times in Pinaceae: evidence from three genomes. *Molecular Biology and Evolution* 17: 773 – 781.
- Wheeler W C, Hayashi C Y. 1998. The phylogeny of the extant chelicerate orders. *Cladistics* 14: 173 – 192.
- Wolfe K H, Li W-H, Sharp P M. 1987. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *Proceedings of the National Academy of Sciences, USA* 84: 9054 – 9058.
- Xia X, Xie Z. 2001. DAMBE: Data analysis in molecular biology and evolution. *Journal of Heredity* 92: 371 – 373.
- Zhang S-H, Second G. 1989. Phylogenetic analysis of the tribe Oryzeae: Total chloroplast DNA restriction fragment analysis (A preliminary report). *Rice Genetics Newsletter* 6: 76 – 80.