

利用早园竹叶绿体基因序列分析其在禾本科植物中的分类地位

李潞滨¹, 唐 征², 庄彩云^{1,2}, 胡 陶^{1,2}, 姚 娜^{1,2}, 杨 凯^{2*}

(¹中国林业科学研究院林业研究所, 国家林业局林木培育重点实验室, 北京 100091; ²北京农学院农业应用新技术北京市重点实验室, 北京 102206)

摘 要: 从构建的早园竹 (*Phyllostachys propinqua* McClure) 叶绿体 DNA 基因组文库中克隆了早园竹叶绿体 *atpA* (EU64861)、*psaB* (EU64863)、*rpoA* (EU64862) 和 *rpoC1* (EU64864) 4 个基因, 并将其序列分别与 GenBank 中部分禾本科植物的基因序列进行了比对分析, 构建核苷酸序列的 Neighbor-joining 系统进化树, 分析探讨其在禾本科中的分类地位。结果表明, 禾本科植物在系统分类上呈现出两大类群。第一大类中包括两个亚类, 水稻与竹类植物聚为一亚类; 玉米、高粱和甘蔗聚为另一亚类。另一大类为小麦、大麦、匍荦剪股颖和黑麦草。表明早园竹与水稻分类地位最为接近; 比水稻和玉米、高粱与甘蔗更为接近, 与禾本科中麦类及其近缘植物关系较远。

关键词: 禾本科; 早园竹; 叶绿体; DNA 序列; 系统分类

中图分类号: S 68; Q 78 **文献标识码:** A **文章编号:** 0513-353X (2009) 10-1498-06

Taxonomic Status of *Phyllostachys propinqua* McClure in Poaceae Basing on Sequence Analysis of Chloroplast Genes

LI Lu-bin¹, TANG Zheng², ZHUANG Cai-yun^{1,2}, HU Tao^{1,2}, YAO Na^{1,2}, and YANG Kai^{2*}

(¹ Key Laboratory of Tree Breeding and Cultivation, State Forestry Administration, Research Institute of Forestry, Chinese Academy of Forestry, Beijing 100091, China; ² Key Laboratory of Agricultural New Technology and Application in Beijing, Beijing Agricultural College, Beijing 102206, China)

Abstract: Following of *Phyllostachys propinqua* McClure chloroplast genomic DNA library constructed, the *atpA* (EU64861), *psaB* (EU64863), *rpoA* (EU64862) and *rpoC1* (EU64864) genes were cloned and analyzed basing on sequence alignment with other chloroplast genes of Poaceae species in GenBank. In order to analyze the taxonomic status of *Phyllostachys propinqua* in Poaceae, four phylogenetic trees were constructed by neighbor-joining method. The results indicated that the gramineous plants can be classified as two classes in the phylogenetic systematics, the first class contained two subclasses, with *Phyllostachys propinqua* and rice together as one of the subclass and maize, sorghum, saccharum composing another; The second class included wheat, barley, ryegrass and creeping bentgrass. It showed that *Phyllostachys propinqua* had the closest taxonomic status with rice, and it was much closer than that between rice and maize or sorghum and saccharum and the relation between *Phyllostachys propinqua* and wheat as well as its kindred plant was far.

Key words: Poaceae; *Phyllostachys propinqua* McClure; chloroplast; DNA sequence; phylogenetic systematics

竹为禾本科 (Poaceae) 竹亚科 (Bambusoideae) 多年生常绿单子叶植物, 是禾本科最原始的亚科,

收稿日期: 2008 - 12 - 24; 修回日期: 2009 - 09 - 07

基金项目: 国家林业局林业科学技术研究项目 (2007-01); 国家林业局 '948' 项目 (2004-4-60)

* 通讯作者 Author for correspondence (E-mail: yangkai8978@126.com)

形态构造较独特。竹秆寿命短, 开花周期长, 物种传播和繁殖更新主要依靠营养体分生, 表现出木质的秆、复合分枝、发达的根系和地下茎, 既是养分的贮存和输导的主要器官, 也具有强大的分生繁殖能力。我国是世界上竹类的分布中心之一, 有着极为丰富的竹种资源, 自然分布约有 40 余属, 500 余种 (赵奇僧和汤庚国, 1993)。竹类具有良好的经济价值和生态价值, 早园竹在我国江、浙、苏、皖一带广泛栽培。作为观赏竹, 四季常青; 作为笋用竹, 上市早、笋质鲜嫩、甘美。

竹类植物集原始和进化特征于一体, 在进化中具有与其它亚科不同的独特路线, 其系统分类研究对于整个禾本科的进化关系研究具有重要意义。长期以来, 完全针对竹类植物在禾本科内的系统研究较少, 竹亚科在禾本科的分类地位和进化关系一直没有达到共识, 依据形态学、不同来源 (核、叶绿体和线粒体) 的基因序列或者同一来源但不同基因获得的系统分类都有所差异 (Grosser & Liese, 1971; Clark et al, 1995; Mathews et al, 2000)。

叶绿体 DNA (chloroplast DNA, cpDNA) 属母性遗传, 相对于核基因组分子量小, 结构简单, 序列保守, 突变率较低, 遗传稳定, cpDNA 分析被认为是研究系统发生和群体遗传学的理想方法, 并被广泛应用 (Palmer, 1987; Badenes & Parfitt, 1995; Xu et al, 2001)。

早园竹 (*Phyllostachys propinqua* McClure) 是竹亚科刚竹属中比较高级的竹种之一。本研究中尝试以均属保守编码基因序列、适合进行高等级分类地位研究的早园竹叶绿体 4 个基因 *atpA*、*psaB*、*rpoA* 和 *rpoC1* 为研究对象, 对早园竹和水稻、玉米、大麦和小麦等重要禾本科植物的相关叶绿体基因进行了分析, 旨在对存在争议的竹类植物在禾本科中的分类地位, 提供较为适宜和确切的分子生物学层面的研究手段和结论。

1 材料与方法

1.1 材料

早园竹 (*Phyllostachys propinqua* McClure) 幼嫩叶片 2008 年 8 月采集自中国林业科学研究院北京良乡竹藤资源圃。

1.2 叶绿体基因组 DNA 的提取

依照龚小松等 (1994) 的方法。1% 琼脂糖凝胶电泳检测提取 DNA 质量, 用紫外分光光度计测定 DNA 样品的浓度, 4℃ 保存备用。

1.3 叶绿体基因组 DNA 文库的构建

用 10 U 内切酶 *EcoR* 部分酶切 10 μg 早园竹叶绿体基因组 DNA, 反应体积 100 μL, 37℃ 酶切 15 min, 终止反应后通过 0.8% 琼脂糖凝胶电泳分离、回收 1~2 kb 长度酶切 DNA, 将回收的部分酶切 DNA 片段与以 *EcoR* 酶切并经牛肠碱性磷酸脂酶处理的 pUC18 载体用 T4 DNA Ligase 进行连接; 连接产物转化感受态 *E. coli* Top 10; 铺平板培养进行文库的滴度测定。获得 3.04×10^5 个重组克隆。

同法构建早园竹叶绿体基因组内切酶 *Hind* 部分酶切叶绿体基因组 DNA 文库, 获得约 2.9×10^5 个重组克隆。

1.4 测序和序列分析

早园竹叶绿体基因组 DNA 文库克隆送北京金螺旋生物科技中心测序, 获得早园竹叶绿体基因组基因序列。利用 BLAST 软件与 GenBank 数据库中的序列进行比对分析, 获得早园竹和其它禾本科植物相应基因的序列。

利用 MEGA 4.0 软件对早园竹叶绿体基因组基因序列和相应禾本科植物基因序列按照最大同源性的原则进行排序, 采用 Kimura 2 (Kimura, 1980) 计算核苷酸差异值, 并用 Neighbor-joining 法 (Saitou & Nei, 1987) 构建系统进化树, 自展数 (bootstrap) 为 1 000。

2 结果与分析

2.1 早园竹叶绿体基因的克隆和序列比对分析

经过检测,多数克隆插入片段长度在 300 ~ 2 000 bp 之间,随机选取两个叶绿体基因组文库各 100 个克隆进行测序,测序结果拼接后,共计长度 104 339 bp,所测定的早园竹叶绿体 DNA 片段通过与水稻日本晴 Nipponbare 叶绿体 DNA 序列 (AY522330) 的比对,其长度相当于水稻叶绿体基因组长度的 71%。

早园竹叶绿体基因组测序片段与 GenBank 中水稻、甘蔗、高粱、玉米、匍茎剪股颖和小麦等禾本科植物叶绿体基因比对,获得早园竹叶绿体基因序列。

2.2 基于早园竹叶绿体基因组基因序列的系统分析

为研究竹类植物的系统分类,本研究从克隆的早园竹叶绿体基因组基因中,依据 Olmstead 和 Palmer (1994) 对叶绿体基因组中哪些基因适于进行系统分类研究的结论,选取早园竹叶绿体基因中 *atpA* (EU64861)、*psaB* (EU64863)、*rpoA* (EU64862) 和 *rpoC1* (EU64864) 4 个基因用于本研究,通过搜索 GenBank 数据库获得其它禾本科植物相应基因的序列,用于系统进化分析。

叶绿体基因 *atpA* 编码叶绿体 ATP 合成酶 CF₁- σ 亚基蛋白。基于禾本科植物 *atpA* 基因序列构建的系统树显示,禾本科植物 *atpA* 基因在聚类关系上分为两大类:水稻和早园竹一起与甘蔗、杂交甘蔗、高粱、玉米聚为一大类;小麦、大麦、匍茎剪股颖和黑麦草聚为另一大类 (图 1)。

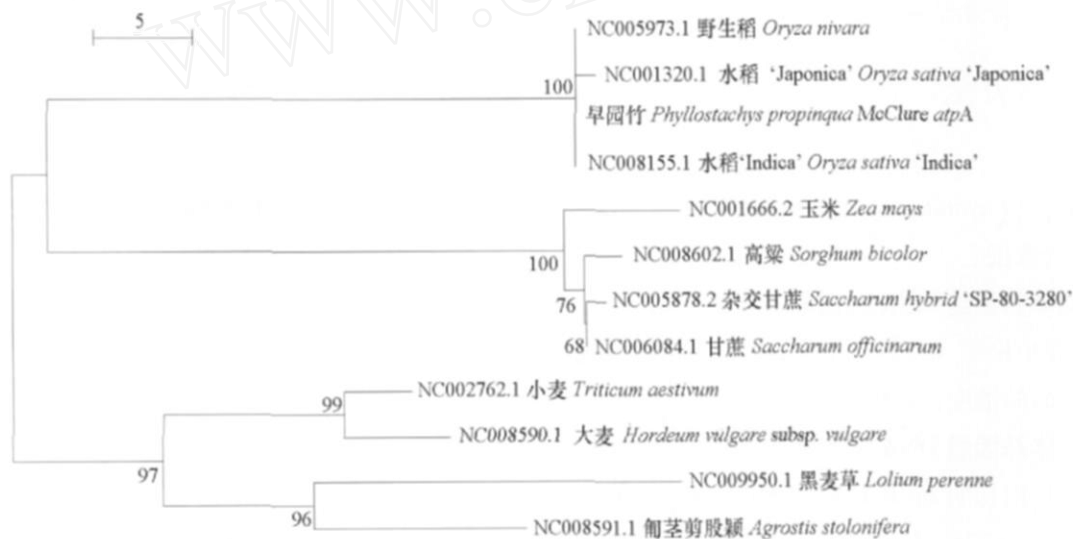


图 1 基于早园竹和其它禾本科植物 *atpA* 基因序列所构建的系统进化树
分支节点数值为置信度值,下同。

Fig 1 Phylogenetic relationships of the *Phyllostachys propinqua* McClure together with grasses (Poaceae) inferred from *atpA* gene sequence
Bootstrap supporting values are shown at branch nodes. The same below.

psaB 基因编码光合系统中 PS 复合体中的 *psaB* 蛋白。基于禾本科植物 *psaB* 基因序列构建的系统树显示,禾本科植物 *psaB* 基因在聚类关系上分为两类,水稻 Indica、Japonica 和野生稻聚合后与早园竹聚合,再与甘蔗、杂交甘蔗与高粱、玉米聚合为一大类;小麦和大麦、匍茎剪股颖和黑麦草聚为一大类 (图 2)。

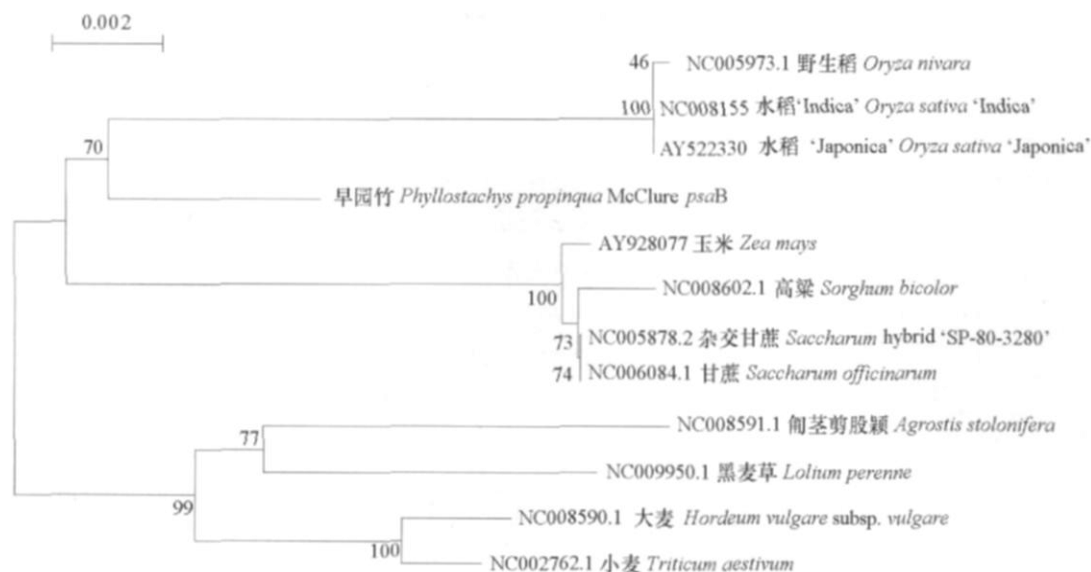


图 2 基于早园竹和其它禾本科植物 *psbB* 基因序列所构建的系统进化树

Fig 2 Phylogenetic relationships of the *Phyllostachys propinqua* McClure together with grasses (Poaceae) inferred from *psbB* gene sequence

叶绿体基因 *rpoA* 编码叶绿体可溶性 RNA 聚合酶的 α 亚基。基于早园竹和其它禾本科植物 *rpoA* 基因序列构建的系统树显示，禾本科植物甘蔗、高粱、玉米聚为一类，早园竹与野生稻和水稻 *Indica* 聚为一类，这两类聚合后与匍茎剪股颖、小麦、大麦聚合（图 3）。

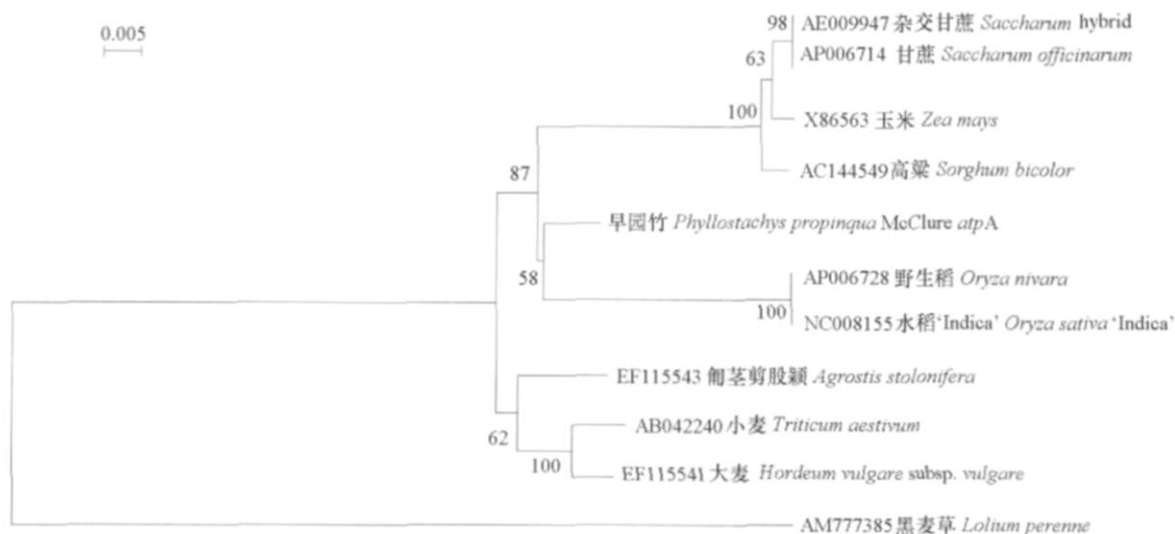


图 3 基于早园竹和其它禾本科植物 *rpoA* 基因序列所构建的系统进化树

Fig 3 Phylogenetic relationships of the *Phyllostachys propinqua* McClure together with grasses (Poaceae) inferred from *rpoA* gene sequence

叶绿体基因 *rpoC1* 编码叶绿体可溶性 RNA 聚合酶的 β 亚基。基于禾本科植物 *rpoC1* 基因序列构建的系统树显示，水稻 *Indica*、*Japonica* 和野生稻先聚为一类，然后与早园竹聚为一大类；另一大类为甘蔗、玉米、高粱先聚合，然后和小麦聚为一类（图 4）。

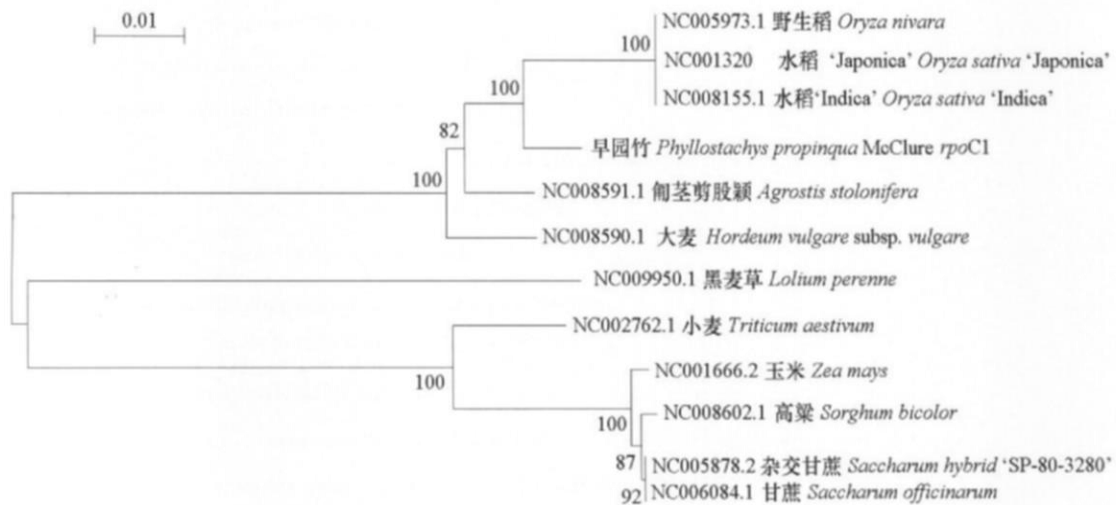


图 4 基于早园竹和其它禾本科植物 *rpoC1* 基因序列的系统进化树

Fig. 4 Phylogenetic tree of the *Phyllostachys propinqua* McClure and Poaceae based on *rpoC1* gene sequence

3 讨论

植物叶绿体 DNA (cpDNA) 为闭环双链 DNA, 原核生物型, 大小为 120 ~ 210 kb。叶绿体基因相当保守, 进化速率平均每年每个位点约为 $(0.2 \sim 1.0) \times 10^{-9}$, 这仅为核基因的 1/5 (Badenes & Parfitt, 1995); 此外, cpDNA 属母体遗传, 一般不会在杂交中出现胞质混合, 在物种进化中形成相对独立稳定的体系, 有独立的进化路线, 不依赖于其它任何数据即可构建分子系统树。故对 cpDNA 进行分析, 可为从历史和系统发育的角度解释生物多样性提供可靠和准确的信息。迄今为止的植物分子系统学研究大多基于对 cpDNA 比较分析。研究结果表明, 在分类群中亲缘关系越近, 其 cpDNA 的同源性越高, 符合协同进化的理论。另外, cpDNA 不同区域的 DNA 序列的进化速率存在着差异, 这就使 cpDNA 序列的分析可以适合于任何一级分类群的系统学研究。在一些保守区域的序列比较可以应用于研究大类群的系统与进化。而在特定的非编码区, 由于序列结构特征而出现突变热点区, 这对于较近分类群的研究有重要价值。本研究选取的早园竹叶绿体基因中 *rpoA*、*rpoC1* 分别编码 RNA 聚合酶的 α 和 γ 亚基。RNA 聚合酶在进化过程中高度保守, 其许多结构特征在从细菌到人类的不同物种间也具有保守性 (Chapman, 2007)。

目前对竹类植物的分子生物学研究薄弱, 公布的竹类植物 DNA 序列也多集中在对核、叶绿体和线粒体个别基因的研究上。多数有关竹类植物系统学方面的研究是以单个叶绿体基因组基因或核基因为研究对象, 分别得出竹类植物与水稻, 竹类植物与玉米, 竹类植物与小麦有较近亲缘关系的结论 (Mason-Gamer et al, 1998; Gaut et al, 1999; Mathews et al, 2000)。由于不同叶绿体基因可能存在进化速度的不同, 而核基因还存在进化过程中的重组、重排、加倍等情况, 利用单个基因研究系统进化关系可能导致一些偏差。本研究以 *atpA*、*psaB*、*rpoA* 和 *rpoC1* 为研究对象, 其中 *atpA*、*psaB* 和 *rpoA* 的结果显示出高度的一致性; 而 *rpoC1* 的结果与其它 3 个基因的分析结果有部分出入。所以在进行系统学亲缘关系分析时, 采用多个基因为研究对象进行分析可以避免单个基因分析过程中的偏差, 有助于正确反映物种的分类地位和系统进化关系。

本研究结果表明: 选用的 4 个基因 *atpA*、*psaB*、*rpoA* 和 *rpoC1* 在系统进化关系上都显示出早园竹

与水稻亲缘关系最为接近; 其次与甘蔗、玉米、高粱进化关系较近, 与禾本科中麦类和麦类近缘植物的进化关系较远。樊龙江等 (2006) 依据 GenBank 中已测序的绿竹 (*Dendrocalamopsis oldhami* Munro)、毛竹 (*Phyllostachys heterocycla* Carr.) 和麻竹 (*Dendrocalamus latiflorus* Munro) 的 8 个全长 mRNA 进行的系统学分析得出的竹类植物与水稻亲缘关系最近的结果。Zhang (2000) 基于叶绿体 *psl16* 基因获得的系统进化关系为竹类与水稻先聚合然后与玉米再聚在一起; Nadot 等 (1995) 同样基于叶绿体基因 *ps4* 获得竹类与水稻先聚合然后与玉米、小麦再聚在一起。

研究认为禾本科植物在系统进化上大体分为两个进化类群, 第一个类群中包括两个亚类, 水稻和竹类为一亚类; 高粱、甘蔗和玉米为另一亚类; 第二类群为麦类植物和其近缘植物。水稻与竹类植物有最近的亲缘关系, 这种较近的亲缘关系比水稻与玉米、高粱与甘蔗的亲缘关系更接近。水稻是目前禾本科植物中研究最为详细的物种, 已经完成基因组测序; 由于水稻和竹类植物具有很近的亲缘关系, 提示我们可充分利用水稻的研究背景加速竹类植物的分子生物学研究。

References

- Badenes M L, Parfitt D E. 1995. Phylogenetic relationships of the cultivated *Pinus* species from an analysis of chloroplast DNA variation. *Theor Appl Genet*, 90: 1035 - 1041.
- Chapman D E. 2007. Transcribing RNA Polymerase is phosphorylated at CTD residue serine-7. *Science*, 318 (5857): 1780 - 1782.
- Chao Chison, Tang Geng-guo. 1993. The present status and problems of bamboo classification in China. *Journal of Nanjing Forestry University: Natural Sciences*, 17 (4): 1 - 8. (in Chinese)
- 赵奇僧, 汤庚国. 1993. 中国竹子分类的现状和问题. *南京林业大学学报: 自然科学版*, 17 (4): 1 - 8.
- Clark L G, Zhang W P, Wendel J F. 1995. A phylogeny of the grass family (Poaceae) based on *ndhF* sequence data. *Systematic Botany*, 20 (4): 436 - 460.
- Fan Long-jiang, Guo Xing-yi, Ma Nai-xun. 2006. Comparative study on phylogenetics and sequences composition of bamboos and cereals. *Forest Research*, 19 (2): 165 - 169. (in Chinese)
- 樊龙江, 郭兴益, 马乃训. 2006. 竹类植物与水稻等其它禾本科作物的系统. *林业科学研究*, 19 (2): 165 - 169.
- Gaut B S, Peek A S, Morton B R, Clegg M T. 1999. Patterns of genetic diversification within the *Adh* gene family in the grasses (Poaceae). *Mol Biol Evol*, 16 (8): 1086 - 1097.
- Gong Xiao-song, Zeng Fu-hua, Yan Long-fei. 1994. An effect method for the purification of chloroplast from high plants. *Journal of Wuhan Botanical Research*, 12 (3): 277 - 280. (in Chinese)
- 龚小松, 曾富华, 阎隆飞. 1994. 一种纯化高等植物叶绿体 DNA 的有效方法. *武汉植物研究*, 12 (3): 277 - 280.
- Grosser D, Liese W. 1971. On the anatomy of Asian bamboos with special reference to their vascular bundles. *Wood Science and Technology*, 5 (4): 290 - 312.
- Kimura M. 1980. A simple method for estimating evolutionary rates base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution*, 16: 111 - 120.
- Mason-Gamer R J, Weil C F, Kellogg E A. 1998. Granule-bound starch synthase: Structure, function, and phylogenetic utility. *Mol Biol Evol*, 15 (12): 1658 - 1673.
- Mathews S, Tsai R C, Kellogg E A. 2000. Phylogenetic structure in the grass family (Poaceae): Evidence from the nuclear gene *phytochrome B*. *Am J Bot*, 87 (1): 96 - 107.
- Nadot S, Bittar G, Carter L, Lacroix R, Lejeune B. 1995. A phylogenetic analysis of monocotyledons based on the chloroplast gene *ps4*, using parsimony and a new numerical phenetics method. *Mol Phylogenet Evol*, 4 (3): 257 - 282.
- Olmstead R G, Palmer J D. 1994. Chloroplast DNA systematics: A review of methods and data analysis. *Amer J Bot*, 81: 1205 - 1224.
- Palmer J D. 1987. Chloroplast DNA evolution and biosystematic uses of chloroplast DNA variation. *Am Nat*, 130: S6 - S29.
- Saitou N, Nei M. 1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, 4: 406 - 425.
- Xu D H, Abe J, Kanazawa A, Gai J Y, Shimamoto Y. 2001. Identification of sequence variations by PCR-RFLP and its application to the evaluation of cpDNA diversity in wild and cultivated soybeans. *Theor Appl Genet*, 102: 683 - 688.
- Zhang W. 2000. Phylogeny of the grass family (Poaceae) from *psl16* intron sequence data. *Mol Phylogenet Evol*, 15 (1): 135 - 146.