

种子植物 *matK* 引物通用性分析研究*

□石林春 (中国医学科学院
北京协和医学院 药用植物研究所 北京 100193)

刘金欣 (北京师范大学资源学院资源生态与中药资源研究所 北京 100875)

姚 辉 徐晓兰 辛天怡 谢彩香 宋经元**

(中国医学科学院
北京协和医学院 药用植物研究所 北京 100193)

摘 要 *matK* 是植物 DNA 条形码鉴定的核心序列之一,其引物通用性一直存在争议。然而,*matK* 不同引物通用性的差异及其在不同植物类群中通用性的差异尚缺乏系统研究。本研究收集种子植物 36 个目,239 个科,3 292 个属,11 429 个种的 14 563 条全长 *matK* 序列。利用生物信息学方法,分析 13 对 *matK* 引物及其 78 种两两引物组合的通用性。结果表明引物 *xf/5r*、*1F/8R*、*390F/1326R* 和 *3F_KIM/1R_KIM* 是通用性最高的 4 对引物,在所有种子植物中分别为 91.18%、84.65%、79.81% 和 80.94%;不同种子植物类群(目),通用性最高的引物存在差异;*xf/5r* 是引物组合的基础序列,*1F/8R*、*1F/1R*、*M3/M4* 和 *3F_KIM/1R_KIM* 可作为引物组合的补充引物。本研究为种子植物不同类群(目)DNA 条形码鉴定 *matK* 序列引物的选择提供有价值的参考。

关键词 :DNA 条形码 种子植物 *matK* 引物通用性

doi: 10.11842/wst.2013.03.008

中图分类号 :R282.5

文献标识码 :A

一、概 述

matK 基因位于叶绿体赖氨酸 tRNA 基因(*trnK*)高度保守的 2 个外显子之间的内含子中,在几乎所有的种子植物中都有分布,为序列长度约 1500 bp 的单拷贝编码基因,编码参与 RNA 转录本中 II 型内含子剪切的成熟酶(Maturase)^[1]。*matK* 基因是叶

绿体基因组中进化最快的编码基因之一^[2]。在科、属水平,*matK* 序列分析为研究类群内部的系统重建提供了较多的信息和较高的支持率^[3, 4]。在物种、种下水平,*matK* 常和 *rbcL*、ITS 联合使用,用于密切相关物种的系统发育关系研究^[5, 6]。*matK* 基因也被用于药用植物的分子鉴定研究,刘静等分析 12 种药用石斛及密花石豆兰的 *matK* 序列,认为 *matK* 可以在分子水平对石斛不同种及种内不同居群(品种)进行鉴别^[7]。2009 年,生物条形码联盟(CBOL)植物工作

收稿日期:2013-05-24

修回日期:2013-06-04

* 科学技术部国家高技术研究发展计划(“863 计划”)项目(2012AA021602)“基于 DNA 条形码的珍稀药用、野生等资源快速检测技术及产品开发”,负责人:陈士林;教育部长江学者和创新团队发展计划项目(IRT1150)“中药资源学”,负责人:陈士林。

** 通讯作者:宋经元,研究员,博士生导师,主要研究方向:中药 DNA 条形码鉴定。

组根据引物通用性、序列测序质量及物种鉴定效率,推荐 *matK* 为植物 DNA 条形码研究的核心 DNA 条形码序列之一^[8]。

然而, *matK* 引物的通用性一直存在争议,不同植物类群往往需要采用不同引物进行扩增^[8, 9]。Kress 等^[10]研究表明,在陆生植物 48 个属中, *matK* 的 PCR 扩增效率为 39.3%。生物条形码联盟植物工作组对 397 份陆生植物样品进行研究, *matK* 单引物在被子植物中的通用性为 90%、在裸子植物中<50%、在隐花植物中为 10%左右^[11]。陈士林等研究表明,在 992 份植物样品中, 390F/1326R 引物的 PCR 扩增效率为 54%^[12]。李德铎等研究表明,使用单一引物 *matK* 在 6 286 样品中的 PCR 扩增效率为 79.5%^[9]。此外,在不同植物类群中,不同 *matK* 引物通用性存在较大差别。Lahaye 等^[13]研究表明,在兰科植物中,引物 390F/1326R 几乎可以对所有的样品进行成功扩增。Gao 等^[14]的研究表明,在豆科植物中,引物 3F_KIM/1R_KIM 和 390F/1326R 的 PCR 扩增效率分别为 91%和 25%。

本研究收集种子植物 36 个目, 239 个科, 3 292 个属 11 429 个种的 14 563 条全长 *matK* 序列,利用生物信息学方法,分析 13 对 *matK* 引物及其 78 对引物组合的通用性,为种子植物 DNA 条形码鉴定中 *matK* 引物的选择提供有价值参考。

二、材料与方法

1. 序列下载和提取

将植物 GenBank (V195) 文件下载到本地服务器,共 61 个文件。对每个 GB 格式 Record 的 DEFINITION 字段采用 *matK* 关键词检索;对

ORGANISM 字段采用 Spermatophyta 关键词检索。对检索到 Record,从 FEATURES 字段中解析 *matK* 基因位置信息,并从 ORIGIN 中提取相对应的 *matK* 基因序列。仅保留 *matK* 基因位置信息无“>”和“<”(即 *matK* 序列被注释为全长序列),且序列长度大于 100 bp 的序列。结合 Taxonomy 库,进一步去除“门、目、科、属、种”分类信息不完整的序列,以及序列中碱基为“N”的数目大于序列全长 1%的序列,共得到 58 个目, 274 个科, 3 339 个属, 11 480 个种的 14 644 条 *matK* 序列。进一步去除物种数小于 10

表 1 种子植物分类类群(目)及其全长 *matK* 序列数目

Class	Order	Family	Genus	Species	Samples
Angiosperms	Apiales	2	6	22	60
Angiosperms	Asparagales	14	514	1 888	2 424
Angiosperms	Asterales	5	243	552	759
Angiosperms	Brassicales	11	99	230	302
Angiosperms	Bruniales	1	12	63	63
Angiosperms	Caryophyllales	30	259	1 022	1 272
Angiosperms	Cornales	4	34	133	161
Angiosperms	Crossosomatales	5	11	14	16
Angiosperms	Cucurbitales	2	3	32	32
Angiosperms	Dioscoreales	2	8	41	48
Angiosperms	Dipsacales	2	29	31	31
Angiosperms	Ericales	16	119	793	963
Angiosperms	Fabales	4	472	1 510	1 933
Angiosperms	Fagales	2	13	70	119
Angiosperms	Gentianales	5	61	144	166
Angiosperms	Geraniales	2	6	11	11
Angiosperms	Lamiales	22	204	677	833
Angiosperms	Laurales	5	41	108	116
Angiosperms	Liliales	5	19	98	150
Angiosperms	Magnoliales	5	29	112	154
Angiosperms	Malpighiales	14	124	583	835
Angiosperms	Malvales	4	68	153	190
Angiosperms	Myrtales	7	83	154	199
Angiosperms	Nymphaeales	2	8	37	54
Angiosperms	Oxalidales	2	2	14	15
Angiosperms	Piperales	4	15	206	226
Angiosperms	Poales	13	374	1 008	1 215
Angiosperms	Proteales	3	65	172	188
Angiosperms	Ranunculales	7	39	334	392
Angiosperms	Rosales	7	83	301	406
Angiosperms	Sapindales	9	104	240	384
Angiosperms	Saxifragales	10	49	190	218
Angiosperms	Solanales	2	17	147	209
Angiosperms	Zingiberales	7	68	263	315
Gymnosperms	Cycadales	3	10	63	90
Gymnosperms	Ephedrales	1	1	13	14

个的种子植物类群(目)后,共获得 36 个目,239 个科,3 292 个属,11 429 个种的 14 563 条 *matK* 全长序列(见表 1)。

2. 引物通用性分析

根据文献报道^[8, 9, 12],种子植物可用引物有 13 对,见表 2。使用 EMBOSS 中的 Primersearch 和 Fuzznuc 程序,碱基不匹配率设为 20%,分别分析 13 对引物及其 78 种两两引物组合在种子植物 36 个目中的通用性。

三、结 果

1. 单引物通用性

同一种子植物类群(目),不同引物的通用性存在差异。不同种子植物类群(目),同一引物的通用性也存在差异(见表 3)。在 Apiales 等 25 个目中,引物 *xf/5r* 具有最大的引物通用性。除 *Malpighiales* 目(68.5%)外,*xf/5r* 的引物通用性均大于 90%。在 *Zingiberales*、*Poales* 和 *Nymphaeales* 目,引物 390F/1326R 具有最大的引物通用性,分别为 99.68%、99.84%和 100%。在 *Fabales*、*Asterales* 和 *Magnoliales* 目,3F_KIM/1R_KIM 具有最大的引物通用性,分别为 90.02%、96.97%和 100%。在 *Caryophyllales* 目中,

引物 1F/1R 具有最大的引物通用性,为 96.93%。在 *Laurales* 目,1F/8R 具有最大的引物通用性,为 100%。在 *Cycadales* 中,GYM_R2A/GYM_F2A 具有最大的引物通用性,为 100%。在 *Sapindales* 目中,M3/M4 具有最大的引物通用性,为 77.6%。在 *Ephedrales* 目,13 对引物的通用性均为 0。

2. 两两引物组合通用性

除 13 对引物通用性均为 0 的 *Ephedrales* 目,单引物通用性为 100%的 *Apiales* 等 15 个目,对其余 20 个目进行两两引物组合通用性分析,分析不同引物之间的互补性。在 *Asterales* 等 6 个目中,*xf/5r*+1F/1R 具有最大的引物通用性,除 *Malpighiales* (69.82%)外,*xf/5r*+1F/1R 通用性为 99.47%~100%。在 *Asparagales* 等 5 个目中,*xf/5r*+M3/M4 具有最大的引物通用性,为 96.77%~100%。在 *Dioscoreales* 等 4 个目中,*xf/5r*+1F/8R 具有最大的引物通用性,为 97.92%~100%。在 *Fabales*、*Ericales* 和 *Poales* 目,*xf/5r*+3F_KIM/1R_KIM 具有最大的引物通用性,分别为 97.26%、99.58%和 99.92%。在 *Sapindales* 目中,390F/1326R+M3/M4 具有最大的引物通用性,为 92.97%。在 *Brassicales* 目中,1F/8R+3F_KIM/1R_KIM 具有最大的引物通用性,为 100%。引物组合通用

表 2 种子植物 13 对引物及其序列

Primer	F	R
F_Abies/R_Abies	ATTCTTCCGAACCTACAC	AACCTTTTCGTCGCTGGAT
3F_KIM/1R_KIM	CGTACAGTACTTTTGTGTTTACGAG	ACCCAGTCCATCTGGAAATCTTGCTTC
390F/1326R	CGATCTATTCAATCAATATTTTC	TCTAGCACACGAAAGTCGAAGT
GYM_R2A/GYM_F2A	ACYTTTCGYRCTGGATCCAAG	GTTTTAGCRCATGRRAGTCGAAG
xf/5r	TAATTTACGATCAATTCATTC	GTTCTAGCACAAGAAAGTCG
matK-445mF/matK-1209mR	AAGGAATGGATGGATGGAATAGTCT	CACCCTGAGATGTCACAAAA
matK_5F/matK_C369F	GTATCGCACTATGTATCATTTGA	GTGGAAATTCATTTTC
MF/MR	ATTTGCGATCTATTCAATCAAT	TGAGATTCCGCAGGTCATT
M3/M4	GCAACAATACTTCTATATCCGCTT	GAACCTCTCTAATAATCCCGAACCT
1F/1R	ATGTATCAACAGAATCRT	AGAGTTTGAACCAAGACTTCCA
1F/8R	ATGTATCAACAGAATCRT	AAAGTTCTAGCACAAGAAAGTCGA
W/9R	TACCCTATCCTATCCAT	TACGAGCTAAAGTTCTAGC
trnK_685F/trnK_2R	GTATCGCACTATGTATGATTTGA	AAGTCTAGTGGATGGAGTAG
matK/matK-R	CCTATCCATCTGGAAATCTTAG	GTTCTAGCACAAGAAAGTCG
matK-F/matK-R	GAACTCGTCGGATGGAGTG	TAAACGATCCTCTCATTCACGA
trnK-710f/matK-iR	GTATCGCACTATGT[TTA]TCATTTGA	AAATGCAAAGAGGAAGCATCT

表 3 不同种子植物类群(目)中 *matK* 单引物通用性(%)

Order	xf/5r	1F/8R	390F/ 1326R	3F_KIM/ 1R_KIM	1F/1R	matkF/ matkR	W/9R	M3/M4	GYM_R2A/ GYM_F2A	MF/MR	tmK_710f/ matK_iR
Apiales	100	100	100	100	100	100	98.33	100	96.67	0	0
Asparagales	98.6	78.47	98.56	95.13	75.7	95.59	87.95	95.38	2.93	0	0
Asterales	96.84	91.83	96.57	96.97	93.02	57.84	4.74	64.82	13.7	0	0
Brassicales	98.68	97.35	30.46	97.35	98.34	96.03	64.9	21.19	3.97	0	0
Bruniales	100	100	100	100	100	100	96.83	100	4.76	0	0
Caryophyllales	96.7	95.44	96.38	96.38	96.93	83.18	48.74	3.62	73.35	0	0
Cornales	98.76	98.76	97.52	98.14	98.76	96.89	97.52	97.52	82.61	0	0
Crossosomatales	100	100	68.75	100	87.5	100	56.25	56.25	12.5	0	0
Cucurbitales	100	100	100	96.88	96.88	93.75	0	3.13	3.13	0	0
Cycadales	0	6.67	0	0	0	0	0	0	100	0	0
Dioscoreales	97.92	93.75	87.5	85.42	72.92	95.83	95.83	29.17	12.5	56.25	0
Dipsacales	90.32	77.42	64.52	9.68	70.97	3.23	3.23	61.29	45.16	0	0
Ephedrales	0	0	0	0	0	0	0	0	0	0	0
Ericales	98.44	86.19	98.03	83.7	86.19	90.24	31.67	67.5	63.55	0	0
Fabales	68.8	66.37	6.67	90.02	80.44	62.91	0.57	15.83	7.81	0	0
Fagales	100	100	96.64	100	100	99.16	0	26.05	0	0	0
Gentianales	99.4	99.4	99.4	98.8	97.59	74.7	9.64	1.81	0	0	0
Geraniales	100	81.82	90.91	72.73	36.36	81.82	9.09	18.18	18.18	0	0
Lamiales	97.96	96.28	97.6	88.36	95.2	62.06	55.34	56.06	6.12	0	0
Laurales	95.69	100	100	99.14	95.69	98.28	71.55	91.38	12.93	0	0
Liliales	99.33	98.67	99.33	64.67	58.67	97.33	98.67	99.33	7.33	0	0
Magnoliales	99.35	99.35	99.35	100	100	77.92	70.13	71.43	72.08	0	0
Malpighiales	68.5	65.87	62.51	25.63	4.91	10.54	21.56	0.12	20	0	0
Malvales	97.37	95.79	13.68	96.84	64.74	96.84	47.89	0	6.32	0	12.63
Myrtales	95.48	93.47	95.48	45.73	94.97	54.27	61.81	34.67	48.24	0	0.5
Nymphaeales	98.15	98.15	100	98.15	100	98.15	94.44	38.89	0	0	3.7
Oxalidales	100	100	100	73.33	100	80	93.33	13.33	0	0	0
Piperales	98.67	88.94	97.35	45.13	43.81	85.84	92.92	56.19	89.38	0	0
Poales	99.75	97.7	99.84	88.64	58.85	98.52	99.59	49.55	0.33	0	0
Proteales	100	100	100	54.79	1.06	98.94	75	98.94	0	0	0
Ranunculales	99.49	98.21	98.72	25.26	53.83	29.59	96.68	77.55	94.13	0.26	0
Rosales	100	100	99.26	99.01	99.75	99.51	93.1	9.11	2.46	0	0
Sapindales	64.06	53.13	65.89	0.26	15.36	0	0	77.6	1.82	0	0
Saxifragales	100	94.04	96.33	50.46	51.38	52.29	88.99	86.24	22.02	0	0
Solanales	100	99.52	100	100	99.52	96.65	3.35	85.65	0.48	0	0
Zingiberales	98.73	70.16	99.68	81.59	64.76	80	92.38	96.83	0	0	0

注:①引物 F_Abies/R_Abies 在 Ginkgoales 目为 100% ,在其他目为 0 ;②引物 matK_445mF/matK_1209mR 在 Cycadales 目为 82.22% ,在其他目为 0。

表 4 不同种子植物类群(目)中通用性最高的 *matK* 引物组合

Order	Primer	Universality(%)	Order	Primer	Universality(%)
Asparagales	xf/5r+M3/M4	99.71	Lamiales	xf/5r+1F/8R	99.16
Asterales	xf/5r+1F/1R	99.74	Liliales	xf/5r+M3/M4	100
Brassicales	1F/8R+3F_KIM/1R_KIM	100	Malpighiales	xf/5r+1F/1R	69.82
Caryophyllales	xf/5r+1F/1R	99.76	Malvales	xf/5r+1F/1R	99.47
Cornales	xf/5r+1F/1R	100	Myrtales	xf/5r+1F/1R	100
Dioscoreales	xf/5r+1F/8R	97.92	Piperiales	xf/5r+1F/8R	99.12
Dipsacales	xf/5r+M3/M4	96.77	Poales	xf/5r+3F_KIM/1R_KIM	99.92
Ericales	xf/5r+3F_KIM/1R_KIM	99.58	Ranunculales	xf/5r+M3/M4	100
Fabales	xf/5r+3F_KIM/1R_KIM	97.26	Sapindales	390F/1326R+M3/M4	92.97
Gentianales	xf/5r+1F/8R	100	Zingiberales	xf/5r+M3/M4	100

性比单引物通用性平均提高 2.56%, 标准偏差为 $\pm 3.53\%$, 其中 Sapindales 目的提高最为显著, 提高 15.37%。见表 4。

四、讨 论

DNA 条形码鉴定技术是传统形态分类学和中药材鉴定方法强有力的补充^[15, 16], *matK* 是植物 DNA 条形码研究的核心条形码序列之一。*matK* 基因的引物通用性一直是 *matK* 基因作为通用 DNA 条形码序列争议的问题。目前不同引物在不同种子植物类群(目)中的通用性尚缺乏系统研究, 这给种子植物 DNA 条形码研究 *matK* 序列引物的选择, 带来了较大的盲目性和随机性。本研究通过对种子植物 36 个目, 239 个科, 3 292 个属, 11 429 个种的 14 563 条 *matK* 序列的 13 对引物通用性分析, 为种子植物 DNA 条形码研究中, *matK* 序列引物的选择提供重要参考。

在 36 个目中, xf/5r^[17]、390F/1326^[13]、1F/8R^[18]和 3F_KIM/1R_KIM^[19]分别在 31 个、26 个、25 个和 18 个种子植物类群(目)中具有大于 90% 的引物通用性。在种子植物中, xf/5r、390F/1326R、1F/8R 和 3F_KIM/1R_KIM 联合的通用性为 95.7%。因此, 在分类背景不清楚的情况下, 引物 xf/5r、390F/1326R、1F/8R 和 3F_KIM/1R_KIM 可以作为首选尝试引物。除 Ginkgoales 目外, 引物 F_Abies/R_Abies 在其余所有 35 个目中的通用性均为 0。引物 *matK*_445mF/*matK*_1209mR 仅在 Cycadales 目中, 具有较好的引物通用性(82.22%), 在其他目中均为 0。引物 MF/MR 和 tmK_710f/*matK*_iR 在各分类类群中的引物通用性均较差。因此在使用 F_Abies/R_Abies、*matK*_445mF/*matK*_1209mR、MF/MR 和 tmK_710f/*matK*_iR 引物

时, 需要对分类背景有很好的了解, 并且建议先做小规模预实验。

除 Sapindales 目中最好的引物组合为 390F/1326R+M3/M4, Brassicales 目中最好的引物组合为 1F/8R+3F_KIM/1R_KIM。包括 Asparagales 目在内的 18 个目, 最佳的引物组合均为 xf/5r 和其他引物的组合: Asterales 等 6 个目的最佳引物组合为 xf/5r+1F/1R, Asparagales 等 5 个目的最佳引物组合为 xf/5r+M3/M4, Dioscoreales 等 4 个目的最佳引物组合是 xf/5r+1F/8R, Ericales 等 3 个目的最佳引物组合是 xf/5r+3F_KIM/1R_KIM。因此在使用组合引物时, xf/5r 可作为引物组合的基础引物, 1F/8R、1F/1R、M3/M4 和 3F_KIM/1R_KIM 可以作为补充引物。

参考文献

- Hilu K, Liang H. The *matK* gene: sequence variation and application in plant systematics. *Am J Bot*, 1997, 84(6):830~830.
- Wolfe K H. Protein-coding genes in chloroplast DNA: compilation of nucleotide sequences, data base entries and rates of molecular evolution. *Cell culture and somatic cell genetics of plants*, 1991, 7: 467~482.
- Plunkett G, Soltis D, Soltis P. Clarification of the relationship between Apiaceae and Araliaceae based on *matK* and *rbcL* sequence data. *Am J Bot*, 1997, 84(4):565~565.
- Specht C D, Kress W J, Stevenson D W, *et al.* A molecular phylogeny of Costaceae (Zingiberales). *Mol Phylogenet Evol*, 2001, 21 (3):333~345.
- Stanford A M, Harden R, Parks C R. Phylogeny and biogeography of Juglans (Juglandaceae) based on *matK* and ITS sequence data. *Am J Bot*, 2000, 87(6):872~882.
- Kress W J, Wurdack K J, Zimmer E A, *et al.* Use of DNA barcodes to identify flowering plants. *PNAS*, 2005, 102(23):8369~8374.
- 刘静, 何涛, 淳泽. 药用石斛的叶绿体 *matK* 基因序列分析及鉴别. *药学报*, 2009, 44(9):1051~1055.

- 8 Hollingsworth P M, Forrest L L, Spouge J L, *et al.* A DNA barcode for land plants. *PNAS*, 2009, 106(31):12794~12797.
- 9 Li D Z, Gao L M, Li H T, *et al.* Comparative analysis of a large dataset indicates that internal transcribed spacer (ITS) should be incorporated into the core barcode for seed plants. *PNAS*, 2011, 108(49):19641~19646.
- 10 Kress W J, Erickson D L. A two-locus global DNA barcode for land plants: the coding *rbcL* gene complements the non-coding *trnH-psbA* spacer region. *PLoS one*, 2007, 2(6):e508.
- 11 Hollingsworth P M. DNA barcoding plants in biodiversity hot spots: progress and outstanding questions. *Heredity*, 2008, 101(1):1~2.
- 12 Chen S, Yao H, Han J, *et al.* Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. *PLoS one*, 2010, 5(1):e8613.
- 13 Lahaye R, van der Bank M, Bogarin D, *et al.* DNA barcoding the floras of biodiversity hotspots. *PNAS*, 2008, 105(8):2923~2928.
- 14 Gao T, Yao H, Song J, *et al.* Identification of medicinal plants in the family Fabaceae using a potential DNA barcode ITS2. *J Ethnopharmacol*, 2010, 130(1):116~121.
- 15 陈士林, 姚辉, 韩建萍, 等. 中药材 DNA 条形码分子鉴定指导原则. *中国中药杂志*, 2013, 38(2):141~148.
- 16 陈士林, 郭宝林, 张贵君, 等. 中药鉴定学新技术新方法研究进展. *中国中药杂志*, 2012, 37(8):1043~1055.
- 17 Ford C S, Ayres K L, Toomey N, *et al.* Selection of candidate coding DNA barcoding regions for use on land plants. *Bot J Linn Soc*, 2009, 159:1~11.
- 18 G B J Grob, B Gravendeel, M C M Eurlings, *et al.* Phylogeny of the Tribe Thomsonieae (Araceae) Based on Chloroplast *matK* and *trnL* Intron Sequences. *Syst Bot*, 2002, 27(3):453~467
- 19 陈士林. 中药 DNA 条形码分子鉴定. 北京: 人民卫生出版社, 2012:14~16

Universality Assessment of *matK* Primer Pairs in Seed Plants

Shi Linchun¹, Liu Jinxin², Yao Hui¹, Xu Xiaolan¹, Xin Tianyi¹, Xie Caixiang¹, Song Jingyuan¹

(1. Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences, Peking Union Medical College, Beijing 100193, China;

2. Beijing Normal University Institute of Resource Ecology and Traditional Chinese Medicine Resources, Beijing 100875, China)

Abstract: *matK* is one of the core DNA barcode markers for plant DNA barcode identification and its universality using single makers has been in controversy. However, the universalities of different *matK* primer pairs in same seed plant group (order) and same *matK* primer pairs in different seed plant groups (order) are lack of systematic research. In this study, we collected 14563 full-length *matK* sequences of 11429 species of 3292 genera in 239 families belonging to 36 orders in seed plants. The universalities of 13 *matK* primer pairs and its 78 primer combinations have been assessed using bioinformatics methods. The results indicated that *xf/5r*, *1F/8R*, *390F/1326R* and *3F_KIM/1R_KIM* were the four most universality primer pairs. The four markers' universalities were 91.18%, 84.65%, 79.81% and 80.94% respectively in all 11429 seed plants. The most universality primer pairs in different orders were different. For each order, the primer pair with maximum universality was different. the *xf/5r* was the basal primer pair for primer combination and *1F/8R*, *1F/1R*, *M3/M4* and *3F_KIM/1R_KIM* could be the complementary primer pairs. This study could be a valuable resource for the primer selection of the research DNA barcoding identification in seed plants.

Keywords: DNA barcoding, seed plants, *matK*, universality

(责任编辑 李沙沙 张志华, 责任译审 陈士林)