

doi:10.3969/j.issn.1672-5565.2013.02.01

类群特异性 PCR 引物设计与评估

董声^{1,2}, 俞志明^{1*}, 宋秀贤¹, 韩笑天¹, 曹西华¹

(1. 中国科学院海洋研究所 海洋生态与环境科学重点实验室, 山东 青岛 266071; 2. 中国科学院大学, 北京 100049)

摘要: 本文以真核藻类 18S rDNA 类群特异性 PCR 引物的设计与评估为例, 详细介绍了如何利用 Primrose 等一系列程序设计类群特异性 PCR 引物并评估其敏感性与特异性, 并对这一方法的优势与应用类群特异性 PCR 引物进行群落多样性分析需要注意的问题进行了讨论。

关键词: 类群特异性 PCR 引物; 真核藻类; Primrose; 18S rDNA 通用引物

中图分类号: Q518.2 **文献标识码:** A **文章编号:** 1672-5565(2013)-02-079-07

Design and evaluation of group – targeted PCR primers

DONG Sheng^{1,2}, YU Zhi-ming^{1*}, SONG Xiu-xian¹, HAN Xiao-tian¹, CAO Xi-hua¹

(1. Key Laboratory of Marine Ecology and Environmental Sciences, Institute of Oceanology, Chinese Academy of Sciences, Qingdao 266071, China; 2. University of Chinese Academy of Sciences, Beijing 100049, China)

Abstract: We take the design and evaluation of eukaryotic algae-targeted 18S rDNA PCR primers as an example, to give a step-by-step illustration of how to design and assess group-targeted PCR primers using Primrose and other programs. Then we summarize the advantages of the design & evaluation pipeline by comparing the amplification efficiency of newly designed and previously reported 18S rDNA universal primers. Cautions are also emphasized on using PCR-based strategies to assess community diversity.

Key words: Group-targeted Primers; Eukaryotic Algae; Primrose; 18S rDNA Universal Primers

类群特异性 PCR 引物 (Group-targeted PCR primers) 指的是以一个类群 (group) 生物共有的某一进化保守基因作为扩增产物的 PCR 引物。通过应用类群特异性 PCR 产物扩增环境宏基因组中该保守基因, 并进行宏基因组学分析 (DGGE-测序、T-RFLP、454 测序), 可以获得关于该类群遗传多样性的信息。为了保证结果全面而特异地反映该类群的多样性水平, 类群特异性 PCR 引物需要具备区别于普通 PCR 引物的一些特性, 以系统发育学研究中常用的核糖体 RNA 编码基因为例:

(1) 敏感性: PCR 反应是以多来源 rDNA 而非单一来源 rDNA 作为模板, 需要通过多序列全局比对寻找序列保守区间设计引物, 以尽可能地保证引

物能够扩增所有的目的来源 rDNA;

(2) 特异性: 多来源 rDNA 中存在大量非目标扩增的 rDNA, 需要将引物与这些非目标扩增 rDNA 逐个进行两序列局部比对并汇总, 排除非特异性扩增严重的引物。

类群特异性 PCR 引物的两大特性决定了其设计与评估方法具有一定的特殊性, 利用普通 PCR 引物的设计程序很难独立完成^[1]。本文通过设计并评估一对真核藻类特异性 18S rDNA 引物, 说明如何利用 Primrose 等一系列程序 (见表 1) 方便快捷地设计类群特异性 PCR 引物并评估其敏感性与特异性。通过比较一对新设计的与一对已报道的真核生物 18S rDNA 通用引物的评估结果, 说明基于 Prim-

收稿日期: 2012-12-12; 修回日期: 2013-03-08.

资助项目: 国家基金委创新研究群体 (项目批准号: 41121064); 国家 973 项目 (项目批准号: 2010CB428706); 国家自然科学基金 (项目批准号: 41276115)。

作者简介: 董声, 男, 硕士研究生, 研究方向: 海洋生态安全。E-mail: 03dstmac@163.com.

* 通讯作者: 俞志明, 研究员, E-mail: zyu@qdio.ac.cn.

rose 引物设计方法的优势与实际应用类群特异性

PCR 引物进行群落多样性分析需要注意的问题。

表 1 类群特异性 PCR 引物设计与评估程序

Table 1 Programs for design & evaluation of group-specific PCR primers

程序名	来源	用途	操作平台
ARB	http://www.arb-home.de/downloads.html	序列比对、引物搜索等	Linux Mac OS
Primrose	http://www.bioinformatics-toolkit.org/Primrose/index.html	引物搜索	Java
BEsTRF	http://lie.fe.uni-lj.si/bestrf/	扩增片段长度确定	MS-DOS
Taxman	http://www.ibi.vu.nl/programs/taxmanwww/	引物敏感性与特异性评估	HTML
Oligo	http://www.oligo.net/downloads.html	引物热力学参数分析	Java

1 类群特异性 PCR 引物的设计方法

目前常用的类群特异性引物设计软件包括基于 Linux 和 Mac OS 的 ARB、基于 Java 语言的 Primrose^[2-3]。由于软件安装上的复杂性以及无法进行本地数据库分析,ARB 并不适合初学者使用,因此选择可在本地分析并在 windows 操作系统下运行的 Primrose 来实现类群特异性 PCR 引物的搜索。下面以真核藻类特异性 18S rDNA 引物设计为例,具体说明 Primrose 的使用方法和注意事项。

1.1 构建数据库,确定目标扩增类群

随着分子生物学技术的不断发展与普及,世界三大核苷酸序列一级数据库 GenBank、EMBL 和 DDBJ 条目数量不断增大,在为人们提供大量及时数据的同时,不可避免地存在一些序列质量上的问题:序列长度较短、未经比对的 DNA 序列大量存在。一些规模较小、数据质量较高的核苷酸序列二级数据库应运而生,系统发育学常用二级数据库包括 RDP (Ribosomal Database Project)、greengenes、SILVA rRNA database project 以及 ITS2 Ribosomal RNA Database (见表 2)。

表 2 系统发育学常用核苷酸序列二级数据库的特性

Table 2 Features of nucleotide secondary databases widely-used in phylogenetic research

	RDP	Greengenes	SILVA	ITS2
基因	16S rDNA	16S rDNA	16S/18S&23S/28S rDNA	ITS
筛选器(Filter) ^a	√ ^b	- ^c	√	-
筛除嵌合体	-	√	-	-
二级结构	-	-	-	√

注:a:筛选器(Filter)指的是可通过条件筛选数据库进行浏览、下载;

b:“√”代表存在相关功能;

c:“-”代表暂无相关功能。

Notes:a: Filter is referred to as the availability to browse and download sequences under requisition;

b:“√” implies related service exists.

c:“-” denoted tool mentioned is invalid.

设计类群特异性引物的第一步是构建数据库并确定目标扩增类群,通过从 SILVA 数据库下载序列

并导入 Primrose 来实现,主要包括以下几步:

(1) 根据实验目的,从以上数据库中下载目的扩增数据库与非扩增数据库的 DNA 序列,而后导入 Primrose。(需要注意的是 Primrose 在构建数据库时需要将目的扩增数据库与非扩增数据库分别置于不同的文件中,否则会出现警告而终止程序)。

(2) 点击 Construct Database - Load Sequence File(s),选择数据库文件(见表 3)^[4],包括扩增数据文件(命名为 Algae.fas)与非扩增数据文件(命名为 Non_Algae.fas),加载完成后点击 Accept;若需重新加载点击 Cancel 并重新选择文件。

(3) 目的扩增数据库的选择:点击 Specify target Taxon,弹出新对话框,左半栏是已加载的各个序列文件,左键单击某个文件夹,右半栏将弹出该序列文件中所有的序列条目信息。这时可以通过右键单击某一条目信息或整个文件来确定需扩增的序列,完成后点击 Accept;若需重新选择点击 Reset,若对选择结果不满意可点击 Cancel 恢复上一次选择结果。

表 3 引物设计数据库构成

Table 3 Content of database for primer design

文件名	包含类群(taxa)	序列数
Algae.fas	绿藻门、裸藻门、轮藻门、金藻门、黄藻门、硅藻门、甲藻门、隐藻门、褐藻门、红藻门、定鞭藻门	10 513
Non_Algae.fas	除以上门类的真核生物、细菌、古菌	138 816

注:所有序列从 SILVA (Version 111) 下载,序列长度 > 1 200bp。

Notes: All sequences are downloaded from SILVA (Version 111), whose length > 1 200bp.

1.2 生成备选引物

在构建数据库并确定目标扩增类群之后,将通过 Primrose 获得备选引物。为了平衡引物的敏感性和特异性,以寻找到更合适的类群特异性引物,选择以下 2 种方式生成备选引物:(1) 选取不同真核藻类门类的序列作为备选引物设计来源,以保证引物的高敏感性;(2) 选取某一真核藻类门类的序列作为备选引物设计来源,以保证引物的高特异性(见表 4)。

表4 备选引物设计来源

Table 4 Sequences employed to generate raw primers

类群 (taxa)	方式1 序列来源 (accession number)	方式2 序列来源 (accession number)
绿藻门(Chlorophyta)	X68484	
裸藻门(Euglenida)	AF283306	
轮藻门(Charophyceae)	X70704	
金藻门(Chrysophyceae)	EF432525	
黄藻门(Xanthophyceae)	AM490835	
硅藻门(Bacillariophyta)	AY684941	
甲藻门(Dinophyceae)	EU418969	
隐藻门(Cryptophyta)	AB240956	
褐藻门(Phaeophyceae)	AF350231	
红藻门(Rhodophyta)	AB053489	AB053489
定鞭藻门(Haptophyceae)	AB183595	

具体步骤为:

(1) 点击 Find Oligonucleotides 进入引物搜索参数设置界面。在“Records to target”一栏显示的是上一步选为目标扩增数据库的序列信息,最右边有一个“Design from?”勾选栏,根据表4中2种方式分别勾选11条和1条序列。

(2) 在“Oligonucleotide parameters”一栏中“Required length”指的是引物长度,对于目标扩增序列间亲缘关系较远的引物设计,应尽量缩短引物长度,提高其敏感性;反之,可以适当增长引物长度。由于真核藻类11门之间18S rDNA亲缘关系较远,故选择较短的引物(15bp)。“Maximum degenerate bases”指的是引物简并位点数,为了提高引物敏感性这里设置为3bp。

(3) 若对以上参数设置满意,可以点击“Find oligonucleotides”栏开始备选引物的搜索。程序会首先加载目标扩增数据库,并分析该数据库是否经过比对,弹出对话框询问,选择“Yes”即可,否则程序会开始序列全局比对过程,消耗较长时间,这也是选择经过序列比对的rDNA数据库作为数据来源的优势。比对完成后,程序将进行序列局部比对寻找备选引物,完成后在“Results summary”栏会列出发现的潜在引物数量,以及不同简并度的引物数量。若对结果满意,点击 Accept 进行下一步分析,或者点击 Reset 重置参数进行分析。

1.3 筛选备选引物

生成的备选引物数量较大,还需通过筛选以减少目标,方便后续分析。备选引物的筛选主要包括以下步骤:

(1) 点击 Check Oligonucleotides Against Database 进入数据库评估备选引物界面。界面上半部分分别列出目标与非扩增数据库、引物设计来源序列的信息以及备选引物数量信息。

(2) 对备选引物进行评估前需对引物的特异性水平进行界定,在“Maximum non-targets”栏中输入允许的非扩增序列数量最大值。对于目标扩增序列亲缘较远的引物设计,应尽量设置一个较大的值以保证尽可能多的敏感度较高的备选引物通过筛选;反之,可以设置一个较小的值提高特异性。这里选择一个相对较大的值20 000,然后点击“Search Database Records”进行备选引物评估。

程序首先加载非扩增数据库文件,对备选引物进行筛选,排除非扩增序列数量大于20 000的引物;而后加载目标扩增数据库文件,对筛选后的引物在目标扩增数据库扩增数量进行分析,并计算引物热力学参数,在“Results of search”一栏输出(表5)。

同理,选择 accession number 为 AB053489 的序列进行方式2的引物设计。产生的结果与方式1的结果存在显著不同(表6)。

比较表5与表6可以看出,方式1得到的引物存在较多的简并位点,保证了引物较高的敏感性,全部22条引物中有19条引物其目标扩增数据库扩增比例大于80%,其余3条也大于77%;然而较多的简并位点也降低了引物的特异性,对非扩增数据库的扩增比例均达到20%~40%。方式2因为由1条序列产生,所以不存在简并位点,特异性较好,所有1 351条引物中797条引物其非扩增数据库的扩增比例小于1%;然而低简并性导致了低敏感性,只有序号为897的引物(与表5中11号引物来源于同一段序列)其对目标扩增数据库扩增比例大于80%。由此可见,对于亲缘关系较远的目标扩增序列,为了保证引物的敏感性,需要牺牲一定的特异性,以保证目标扩增序列大部分得到扩增;反之,当目标扩增序列间亲缘关系较近时,利用1条序列就可以较为容易地获得具有合适敏感性和特异性的引物。

表 5 方式 1 生成的引物及其特性

Table 5 Raw primers generated by method 1 and their features

序号	引物(5' - 3')	命中数	目标库内 命中数	占库内 总数%	目标库外 命中数	占库外 总数%	简并碱 基数
1	TTACCAGGTCMRGAC	18 778	9 317	89.115 26	9 461	20.526 339	2
2	ARYTTCGMCCTATC	26 175	8 129	77.752 27	18 046	39.152 13	3
3	RYYTTCGMCCTATCA	26 387	8 269	79.091 35	18 118	39.308 34	3
4	YTTCTGMCCTATGAR	26 340	8 399	80.334 77	17 941	38.924 324	3
5	TTCTGMCCTATCARY	26 274	8 362	79.980 865	17 912	38.861 41	3
6	CGGGRAAWSTTACCA	19 979	9 176	87.766 62	10 803	23.437 906	3
7	GGGRAAWSTTACCAG	19 833	9 183	87.833 57	10650	23.105 963	3
8	GGRAAWSTTACCAGG	19 813	9 216	88.149 216	10 597	22.990 974	3
9	GRAAWSTTACCAGGT	19 390	9 176	87.766 62	10 214	22.160 028	3
10	RAAWSTTACCAGGTC	19 376	9 188	87.881 4	10 188	22.103 619	3
11	AAWSTTACCAGGTCM	19 352	9 185	87.852 71	10 167	22.058 058	3
12	STTACCAGGTCMRGA	18 763	9 322	89.163 086	94 41	20.482 946	3
13	TACCAGGTCMRGACR	18 667	9 219	88.177 91	9 448	20.498 135	3
14	CACCGCCGTCGYWS	21 801	9 469	90.569 1	12 332	26.755 184	3
15	ACCGCCGTCGYWSC	20 943	9 452	90.406 5	11 491	24.930 574	3
16	CCGCCGTCGYWSCT	19 689	9 239	88.369 2	10 450	22.672 047	3
17	CGCCGTCGYWSCTA	19 627	9 165	87.6614	10462	22.698 082	3
18	GCCCGTCGYWSCTAC	19 611	9 142	87.441 414	10 469	22.713 27	3
19	CCCGTCGYWSCTACC	19 504	9 056	86.618 84	10 448	22.667 707	3
20	CCGTCGYWSCTACCG	19 393	9 077	86.819 7	10 316	22.381 323	3
21	CGTCGYWSCTACCGA	19 392	9 076	86.810 135	10 316	22.381 323	3
22	GTCGYWSCTACCGAT	19 328	9 060	86.657 104	10 268	22.277 185	3

注:表中引物首先按简并碱基数由小到大分块,每块中的引物再按照该引物在基因序列中的位置由小到大排列。

Notes: Primers are sorted into blocks according to the number of degenerate bases, within each block arranged on the basis of their position in template sequence.

表 6 方式 2 生成的引物及其特性

Table 6 Raw primers generated by method 2 and their features

序号	引物(5' - 3')	命中数	目标库内 命中数	占库内 总数%	目标库外 命中数	占库外 总数%	简并碱 基数
1	GTGAAACTGCCAATG	19 390	3 651	34.921 09	15 739	34.146 92	0
.....							
368	CCCTTATCGAGGATC	1 711	1 670	15.973 22	41	0.0889 53	0
.....							
402	CGCTCGTAGTCGGAC	372	372	3.558 106	0	0	0
.....							
897	AAACTTACCAGGTCC	18 278	8 509	81.386 89	9 769	21.194 57	0
.....							

注:原表中大部分引物的敏感性较差,故未予列出。

Notes: Primers are partially listed to omit the low sensitive counterparts in the original list.

2 类群特异性 PCR 引物对 (primer pairs) 的综合评估

通过 Primrose 分析数据库得到足够多的引物之后,还需要对引物进行配对,作为 PCR 反应的上游引物与下游引物。这里涉及到引物对评估中需要解决的三个问题:

(1) 该引物对用于 PCR 反应产生的扩增片段 (Amplicon) 长度是否适合下游实验;

(2) 该引物对组合后其敏感性和特异性相对单一引物是否变化明显;

(3) 该引物对的热力学参数是否适合用于 PCR 反应。

由于表 6 中引物敏感性较低,因而选择表 5 中引物进行配对。根据 Primrose 分析结果中给出的 T_m 值或其他计算 T_m 值的方法可知,表 5 中任意引物之间 T_m 值相差较小 (小于 $10\text{ }^\circ\text{C}$), 适合组成引物对。下面将对以上三个问题进行说明。

2.1 获取引物对的扩增片段长度

微生物群落多样性常用技术(DGGE/TGGE、T-RFLP、454 测序)要求 DNA 片段长度大致在 300 ~ 700bp 左右,因此需首先分析不同引物组合的扩增片段长度以选择长度合适的引物对。通过观察表 5 序列可以发现序号 2 ~ 5, 6 ~ 11, 12 ~ 13, 14 ~ 22 来自 3 段序列保守区,但具体在基因中的位置 Primrose 并未给出,因而无法分析扩增片段长度。这里可以采用 bl2seq 分别获得两条引物在某一特定序列所在位置,然后相减得到扩增片段长度;为了获得更全面的扩增片段长度分布信息,可以使用 T-RFLP 分析软件 BEsTRF 获得扩增片段长度的频数分布图(图 1)^[5]。从图中可以看出,表 5 中 1 号与 22 号引物组合扩增 18S rDNA 的片段长度主要分布在 440 ~ 480bp 之间(众数 445bp),适合微生物群落多样性分析。

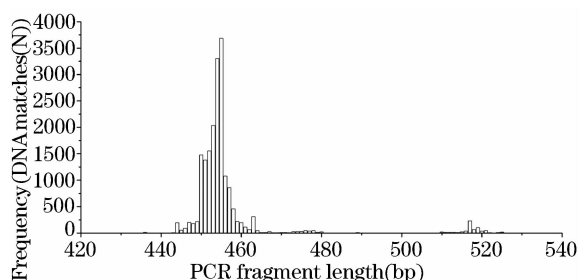


图1 表5中1号与22号引物扩增片段长度的频数分布

Fig. 1 Length frequency distribution of amplicons primed by No. 1&22 primers in Table 5

2.2 评估引物对的敏感性与特异性

通过表 5 可以获得某条引物的敏感性与特异性粗略信息,然而对于两条引物搭配成为引物对后敏感性与特异性的改变,以及引物对在目标扩增数据库中各个类群的敏感性与特异性等具体信息,Primrose 并未给出。这里采用线上程序 Taxman 对引物对的敏感性与特异性进行评估^[6],主要步骤如下:

分别在 Forward primer 与 Reverse primer 空格中填写上、下游引物序列,Database 栏选择“SILVA SSU NR rRNA v. 108_2”,勾选“You are an academic user”。为了获得更为贴近实际 PCR 反应结果的理论模拟结果,Mismatch 空格处填 10 percent,3' mismatch window 处填 5,然后点击 submit 运行。

根据结果 Taxonomic tree with sequence-based counts 栏中的数据,可以获得引物对敏感性与特异性的信息(图 2),从图中可以看出表 5 中 1 号与 22 号引物组合在 11 个真核藻类类群中扩增效果较好(charophyceae 除外),而对非目标扩增类群(untar-

geted groups) 扩增较少,表明该引物对具有较高的敏感性与特异性。

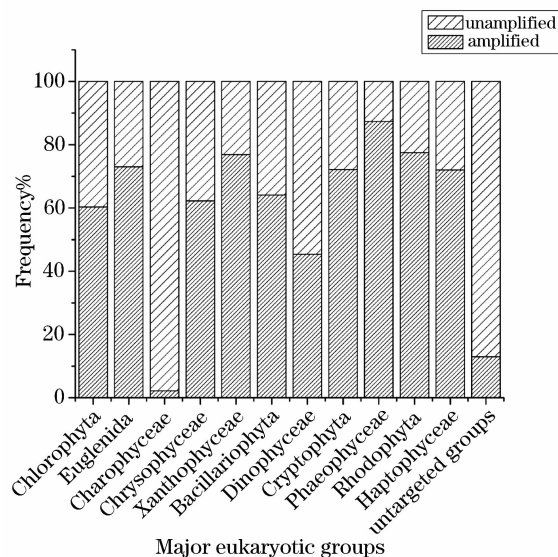


图2 表5中1与22号引物对扩增真核藻类类群的频率分布

Fig. 2 Eukaryotic algae taxa Frequency distribution of amplicons primed by No. 1&22 primers in Table 5

2.3 分析引物对的热力学参数

在满足了扩增片段长度、敏感性与特异性之后,需要对引物对的 T_m 值、 ΔG 、错配等热力学参数进行分析,以获得最优的 PCR 条件。这里采用 Oligo 7 对引物对的热力学参数进行分析,分析步骤如下:

(1) 点击菜单栏 File 中的 Open 按钮,导入目标扩增数据库中某一选定的序列文件,点击菜单栏 Select 中的 New Current Oligo Position 按钮,输入根据 bl2seq 得到的上游(下游)引物位置信息,点击 Sequence 活动栏中部 Forward(Reverse) Primer 前小方框,确定上、下游引物。

(2) 引物内与引物间的错配分析:以表 5 序号 1 号与 22 号引物为例,分别点击菜单栏 Analyze 中 Duplex Formation-Forward Primer、Reverse Primer 和 Mixed Oligos,分析上游引物内部、下游引物内部和上、下游引物间错配物情况,确保潜在的错配物 ΔG 满足要求(表 7)。

表7 引物 ΔG 的参考值

Table 7 reference value of primer ΔG

	3'端 ΔG	全长 ΔG
发夹结构(Hairpin)	≥ -2 kcal/mol	≥ -3 kcal/mol
同种引物间	≥ -5 kcal/mol	≥ -6 kcal/mol
上、下游引物间	≥ -5 kcal/mol	≥ -6 kcal/mol

结果(表 8)与 ΔG 的参考值(表 7)比较,可以看出表 5 中 1 号与 22 号引物组合满足要求,从热力

学参数的角度看适合用于 PCR 反应。

表 8 表 4 中 1 与 22 号引物对的 ΔG

Table 8 primer ΔG of No. 1&22 primer pairs in Table 4

	3' 端 ΔG		全长 ΔG	
	1 号上游引物	22 号下游引物	1 号上游引物	22 号下游引物
发夹结构 (Hairpin)	-	0.5 kcal/mol	-	0.5 kcal/mol
同种引物间	-0.4 kcal/mol	-0.3 kcal/mol	-4.7 kcal/mol	-4.1 kcal/mol
上、下游引物间	-2.1 kcal/mol		-2.6 kcal/mol	

至此,我们得到了扩增片段长度适宜、敏感性与特异性较好、满足热力学条件的真核藻类 18S rDNA 类群特异性引物。

3 基于 Primrose 的类群特异性 PCR 引物设计方法的优势

相对于以单一序列作为引物设计来源的引物设计方法,Primrose 允许多达 15 条序列作为引物设计的来源,在目前分子系统学占主导地位的时代,15 条序列基本可以包含某一分类单元(如真核生物域)的所有下一级分类单元(如变形虫界、囊泡藻界等),这就基本保证了所得引物的高敏感性。利用这 15 条序列得到的 1 条高简并性共有序列(consensus sequence)进行引物搜索,排除在热力学参数上不能满足的、在序列内部存在错配等不符合条件的引物,而后对获得的备选引物在目标扩增数据库和非扩增数据库中进行两序列局部比对,以评估引物的敏感性与特异性,并给出可以进行后续数据分析的引物列表。在类群特异性引物设计过程中,这种批量分析引物并在数据库中评估引物的敏感性与特异性的引物设计方式,相对其他引物设计程序具有显著优势。

由于缺少目标扩增类群为真核藻类 18S rDNA 引物对的报道,这里通过比较一对被大量引用的真核生物 18S rDNA 引物对与通过 Primrose 设计得到的引物对来说明后者设计所得引物的高敏感性(表 9)。Euk1A-Euk516r 是一对以真核生物群落为目标扩增对象的 18S rDNA 引物对,已被大量应用于真核生物中某类群及全部类群的多样性分析^[10-11]。该引物对是通过人工比对一定数量的 18S rDNA 序列,而后根据保守序列获得引物。鉴于当时较小的 rDNA 数据库,评估该引物对的敏感性与特异性的结果并不全面,随着 18S rDNA 数据库的不断扩大(图 3),评估敏感性与特异性结果变得更为可靠。通过 Primrose 设计得到的一对引物 1143F 与

1637R,与 Euk1A-Euk516r 的敏感性与特异性在 Taxman 中进行评估(图 4、图 5)。

表 9 真核生物 18S rDNA 通用引物

Table 9 eukaryote-targeted 18S rDNA universal primers

引物命名	序列(5' - 3')	扩增长度	文献来源
Euk1A	CTGGTTGATCCTGCCAG	约 560bp	[7]
Euk516r	ACCAGACTTGCCCTCC		[8]
1143F	AATTGACGGAAGGCA	约 510bp	本文设计
1637R	CGACGGCGGTGTGTA		[9]

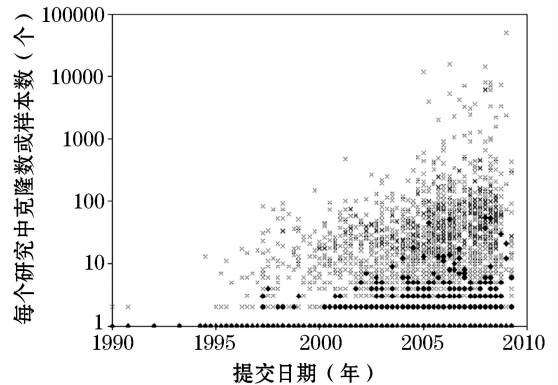


图 3 rDNA 研究数量与深度的年际变化

(“x”表示克隆数,“◆”表示样本数),转引自文献[12]

Fig. 3 Increases in the number of rDNA clones and samples per study since 1990, modified from Reference[12]

比较图 4、图 5 可以看出:新设计的引物对对 10 个主要真核生物类群的扩增结果明显优于已报道引物对的扩增结果,表明新设计引物对相对已报道引物对具有更高的敏感性;同时,新设计引物对与已报道引物对的非特异性扩增数(扩增生物不是真核生物的物种数)分别为 3 和 122,远小于各自的目标扩增数(23 706 和 5 947),表明这两对引物的特异性较好。

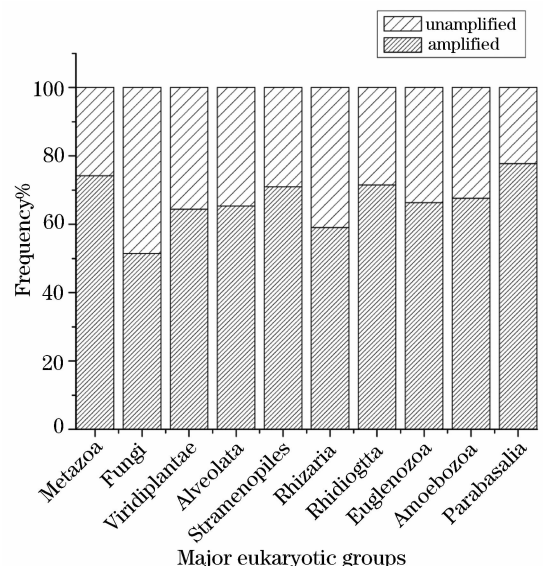


图 4 新设计引物 1143F-1637R 扩增主要真核生物类群的频率分布

Fig. 4 Eukaryotic taxa Frequency distribution of amplicons primed by 1143F-1637R primers

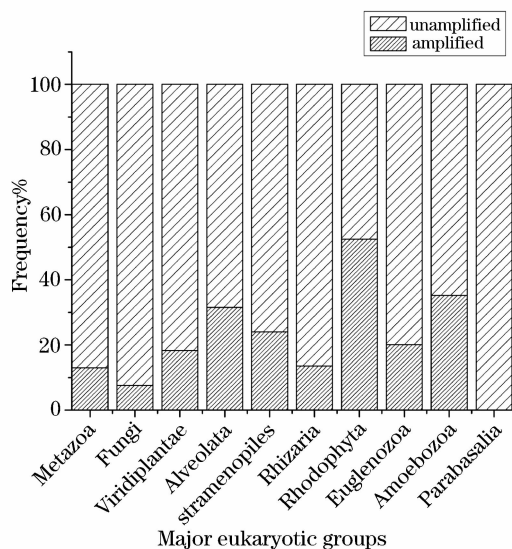


图 5 已报道引物 Euk1A-Euk516r 扩增
主要真核生物类群的频率分布

Fig. 5 Eukaryotic taxa Frequency distribution
of amplicons primed by Euk1A-Euk516r primers

4 应用类群特异性 PCR 引物进行群落多样性分析需注意的问题

具有热稳定性的 DNA 聚合酶 Taq 酶的发现与 PCR 的广泛应用为快速获得某一 DNA 片段大量扩增产物提供了一条捷径。然而,当 PCR 反应的模板由单一来源变成混合来源、当分子生态学家应用 PCR 反应研究群落多样性时,也产生了一些新问题:异源二聚体的形成^[13],引物-模板结合的偏好性^[14],PCR 抑制物的影响^[15]。这其中,PCR 反应的循环数、高简并度的引物,都可能对群落多样性、特别是 species evenness 产生巨大的影响。为了尽可能获得真实的群落多样性信息,应尽量避免 PCR 反应进入平台期对 species evenness 和低丰度物种的影响,选取低简并度的引物、适当减少 PCR 循环数或适当增大引物的浓度。利用 Primrose 等软件设计与评估得到的类群特异性引物尚需在实际 PCR 反应中进行条件的优化,以获得更接近真实群落多样性水平的实验结果。

参考文献 (References)

- [1] 张新宇,高燕宁. PCR 引物设计及软件使用技巧[J]. 生物信息学, 2004, (04): 15-18,46.
- [2] Ludwig, W., O. Strunk, R. Westram, L. Richter, H. Meier, Yadhukumar, A. Buchner, T. Lai, S. Steppi, G. Jobb, W. Forster, I. Brettske, S. Gerber, A. W. Ginhart, O. Gross, S. Grumann, S. Hermann, R. Jost, A. König, T. Liss, R. Lussmann, M. May, B. Nonhoff, B. Reichel, R. Strehlow, A. Stamatidis, N. Stuckmann, A. Vilbig, M. Lenke, T. Ludwig, A. Bode, K. H. Schleifer. ARB: a software environment for sequence data[J]. *Nucleic Acids Research*, 2004, 32(4): 1363-1371.
- [3] Ashelford, K. E., A. J. Weightman, J. C. Fry. PRIMROSE: a computer program for generating and estimating the phylogenetic range of 16S rRNA oligonucleotide probes and primers in conjunction with the RDP-II database [J]. *Nucleic Acids Research*, 2002, 30(15): 3481-3489.
- [4] Quast, C., E. Pruesse, P. Yilmaz, J. Gerken, T. Schweer, P. Yarza, J. Peplies, F. O. Glöckner. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools[J]. *Nucleic Acids Research*, 2013, 41(D1): D590-D596.
- [5] Stres, B., J. M. Tiedje, B. Murovec. BEsTRF: a tool for optimal resolution of terminal-restriction fragment length polymorphism analysis based on user-defined primer-enzyme-sequence databases [J]. *Bioinformatics*, 2009, 25(12): 1556-1558.
- [6] Brandt, B. W., M. J. Bonder, S. M. Huse, E. Zaura. TaxMan: a server to trim rRNA reference databases and inspect taxonomic coverage [J]. *Nucleic Acids Research*, 2012, 40(W1): W82-W87.
- [7] Sogin, M. L., J. H. Gunderson. structural diversity of eukaryotic small subunit ribosomal-rnas-evolutionary implications [J]. *Annals of the New York Academy of Sciences*, 1987, 503: 125-139.
- [8] Amann, R. I., B. J. Binder, R. J. Olson, S. W. Chisholm, R. Devereux, D. A. Stahl. combination of 16s ribosomal-rna-targeted oligonucleotide probes with flow-cytometry for analyzing mixed microbial-populations [J]. *Applied and Environmental Microbiology*, 1990, 56(6): 1919-1925.
- [9] Lawley, B., S. Ripley, P. Bridge, P. Convey. Molecular analysis of geographic patterns of eukaryotic diversity in Antarctic soils [J]. *Applied and Environmental Microbiology*, 2004, 70(10): 5963-5972.
- [10] Diez, B., C. Pedros-Alio, T. L. Marsh, R. Massana. Application of denaturing gradient gel electrophoresis (DGGE) to study the diversity of marine picoeukaryotic assemblages and comparison of DGGE with other molecular techniques [J]. *Applied and Environmental Microbiology*, 2001, 67(7): 2942-2951.
- [11] Monchy, S., J. D. Grattepanche, E. Breton, D. Meloni, G. Sancier, M. Chabé, L. Delhaes, E. Viscogliosi, T. Sime-Ngando, U. Christaki. Microplanktonic Community Structure in a Coastal System Relative to a Phaeocystis Bloom Inferred from Morphological and Tag Pyrosequencing Methods [J]. *PLoS one*, 2012, 7(6): e39924.
- [12] Tringe, S. G., P. Hugenholtz. A renaissance for the pioneering 16S rRNA gene [J]. *Current Opinion in Microbiology*, 2008, 11(5): 442-446.
- [13] Ashelford, K. E., N. A. Chuzhanova, J. C. Fry, A. J. Jones, A. J. Weightman. At least 1 in 20 16S rRNA sequence records currently held in public repositories is estimated to contain substantial anomalies [J]. *Applied and Environmental Microbiology*, 2005, 71(12): 7724-7736.
- [14] Polz, M. F., C. M. Cavanaugh. Bias in template-to-product ratios in multitemplate PCR [J]. *Applied and Environmental Microbiology*, 1998, 64(10): 3724-3730.
- [15] Arbeli, Z., C. L. Fuentes. Improved purification and PCR amplification of DNA from environmental samples [J]. *Fems Microbiology Letters*, 2007, 272(2): 269-275.