

热带假丝酵母 1230 *POX4*、*POX5* 基因及其编码蛋白的研究

秦文彦 任 刚 荣 东 陈远童

(中国科学院微生物研究所 微生物资源前期开发国家重点实验室 北京 100080)

摘 要 根据 GenBank 中热带假丝酵母菌株 pk233(ATCC 20336) *POX4* 和 *POX5* 基因序列,设计了两对引物。通过 PCR 扩增,用一种简易的方法克隆了 1230 菌株中这两个基因。对 *POX4*、*POX5* 的测序表明,1230 菌株的 *POX5* 基因与 pk233 菌株的 *POX5* 基因存在菌株间的差异。这种差异对蛋白功能的影响有待进一步研究。对 *POX4* 和 *POX5* 编码的蛋白 PXP4、PXP5 进行 Pfam 分析和二级结构预测后,推测 PXP4 和 PXP5 的 N 端可能是 FAD 结合部位。

关键词: 热带假丝酵母, PCR, Pfam, 二级结构预测

中图分类号: Q812 文献标识码: A 文章编号: 1001-6209(2002)02-0193-07

热带假丝酵母是一种具有重要工业用途的微生物。它因为能够利用正烷烃生产二元酸而受到人们的广泛关注^[1]。正烷烃在热带假丝酵母中的降解主要通过 β -氧化和 ω -氧化两条途径^[2]。 β -氧化途径完全在过氧化物酶体中进行。催化该途径第一步反应的酰基辅酶 A 氧化酶是限速酶,该酶由八个亚基组成^[3]。目前发现有三个基因(*POX2*、*POX4* 和 *POX5*)编码该酶的亚基。这三个基因编码的蛋白 PXP2、PXP4 和 PXP5 在一级结构上具有较高的同源性^[4~5]。它们都以 FAD 作为辅基,但与 FAD 的结合方式尚不清楚。

本实验室在七十年代从土壤中分离到一株热带假丝酵母 1230^[6]。根据 GenBank 中热带假丝酵母 pk233 菌株(ATCC 20336) *POX4* 和 *POX5* 基因序列,设计两对引物,通过 PCR 扩增克隆了这两个基因并测序。对 *POX4*、*POX5* 编码的蛋白 PXP4、PXP5 在 Pfam(Protein families database of alignments and HMMs)^[7]中分析并进行二级结构预测,探讨它们与 FAD 的结合方式。

1 材料和方法

1.1 菌株

热带假丝酵母(*Candida tropicalis*)1230,本实验室保存菌株。

1.2 DNA 提取

根据《酵母遗传学方法实验指南》^[8]中 DNA 分离方法提取。

1.3 *POX4* 和 *POX5* 的克隆

根据 GenBank 中 *Candida tropicalis* pk233 *POX4* 和 *POX5* 序列,设计两对引物(见表 1),扩增片段的长度分别为 2814bp 和 2013bp。PCR 条件如下:100 μ L 反应体系中,dNTP 0.2mmol/L,引物各 20pmol,模板 DNA 0.5 μ g,Taq plus(Taq DNA polymerase + Pfu DNA polymerase)5U,94 $^{\circ}$ C 预变性 2min 后,进行如下 25 个循环:94 $^{\circ}$ C 变性 35s,50 $^{\circ}$ C 复性 35s,72 $^{\circ}$ C 延伸 *POX4* 180s,*POX5* 120s。最后延伸 600s。取 1~2 μ L 反应液直接与 T-载体连接(T-载体

购自 Promega)。4℃放置过夜。第二天用 2μL 连接液转化大肠杆菌 XL₁-Blue。涂平板后, 37℃培养过夜。挑取白斑再涂平板。37℃培养过夜。挑取薄而透明的菌落检测质粒。

1.4 POX4 与 POX5 的测序

挑选出的阳性克隆送上海基康生物技术公司测序,测序引物为 M₁₃通用引物。

1.5 POX4、POX5 的 Pfam 分析

在英国 Sanger 中心网站(www.sanger.ac.uk)上进行。

1.6 POX4、POX5 二级结构的预测

在瑞士 www.expasy.ch 网站上用 GOR4 方法进行。

表 1 本实验中使用的引物序列
Table 1 Sequences of the deoxy-oligonucleotides used in this study

POX4	primer 1	5'-TCGGGAGAAATATCGTTGGG-3'
	primer 2	5'-TCGCAACTATCAACTAAGGG-3'
POX5	primer 1'	5'-ACCACGCTATCATGCCTACCG-3'
	primer 2'	5'-ACCATGTGCTCTATTAACTGGAC-3'

2 实验结果

2.1 POX4 和 POX5 的克隆

挑取薄菌落接种含抗生素的 LB 培养液,过夜培养。离心收集细胞后,苯酚抽提直接破裂大肠杆菌细胞,上清液电泳。结果见图 1。从图 1 可以看出有三个转化子含有较大的质粒。进一步纯化后,酶切分析证明它们都是目的转化子。PCR 扩增的 POX4、POX5 以及与 T-载体连接后的重组质粒(T-POX4、T-POX5)酶切电泳见图 2。

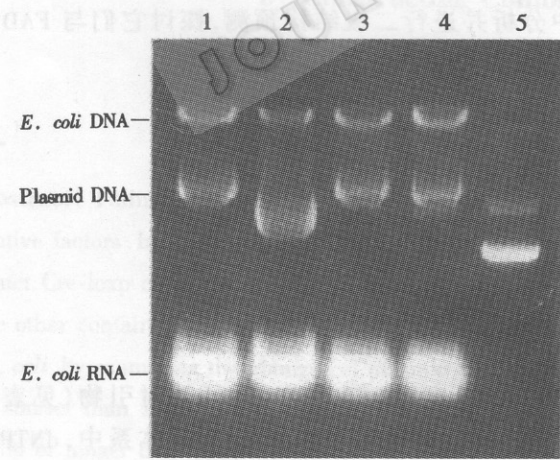


图 1 上清液电泳图
Fig.1 Electrophoresis of supernatant treated by phenol extraction
1 ~ 4. Four transformants; 5. pUC19.

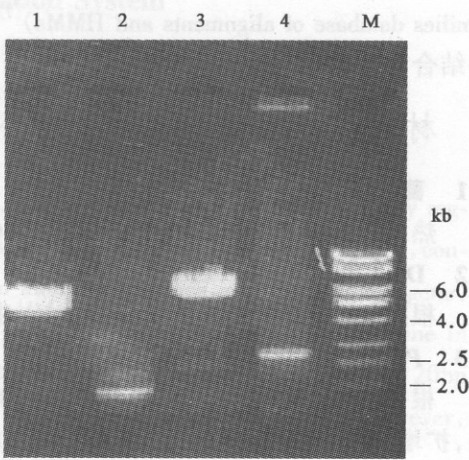


图 2 酶切质粒电泳图
Fig.2 Electrophoresis of digested plasmids
1. BamHI digested T-POX5; 2. Purified POX5;
3. BamHI digested T-POX4; 4. Purified POX4;
M. Marker.

2.2 *POX4*、*POX5* 的测序

POX4 和 *POX5* 都分别测了三个克隆。总结测序结果 ,*POX4* 在我们扩增的区域(*ATC* 前 513 个碱基 ,*TAA* 后 171 个碱基)内与 GenBank 中 pk233 菌株的 *POX4* 完全相同。但 *POX5* 与 pk233 菌株 *POX5* 在编码区存在 4 处差别 ,编码的蛋白质有 3 处氨基酸不同。如下 :

位置		1	382	403	529	1456	1984
DNA 序列	1230	ATG	aaG	ggT	Acc	TCg	AGT TAA
	pk233	ATG	ccG	ccT	Agg	TCc	AGT TAA
氨基酸序列	1230	Met	Lys	Gly	Thr	Ser	Ser *
	pk233	Met	Pro	Pro	Arg	Ser	Ser *

这种差异对蛋白功能的影响有待进一步研究。热带假丝酵母中 ,*CUG* 密码子编码的氨基酸不是亮氨酸 ,而是丝氨酸^[1]。在 *POX5* 编码区的 793-795 位为 *CTG* ,编码的氨基酸应该是丝氨酸。GenBank 中给出的 *POX5* 基因(accession number M12161)编码的氨基酸序列在该位置的氨基酸是亮氨酸 ,应该是不对的。*POX4* 基因的编码区没有此密码子。

2.3 *PXP4* 与 *PXP5* 的 Pfam 分析

2.3.1 *PXP4* 分析结果 :见表 2、表 3。

2.3.2 *PXP5* 分析结果 :见表 4、表 5。

表 2 确信匹配

Table 2 Trusted matches

Domain	Start	End	Bits	Evalue
ACOX	183	688	974.70	2.3e-289

表 3 Pfam-B 中的匹配

Table 3 Matches to Pfam-B

Domain	Start	End	Evalue
Pfam-B_28106	1	30	5.5e-11
Pfam-B_369	33	114	2e-38
Pfam-B_382	115	182	1.2e-32

表 4 确信匹配

Table 4 Trusted matches

Domain	Start	End	Bits	Evalue
ACOX	168	641	937.50	3.5e-278

表 5 Pfam-B 中的匹配

Table 5 Matches to Pfam-B

Domain	Start	End	Evalue
Pfam-B_369	19	99	1.8e-38
Pfam-B_382	100	167	1.7e-33

从以上的结果看出 ,*PXP4* 蛋白中酰基辅酶 A 氧化酶(*ACOX*)区域在 183 ~ 688 位氨基酸残基之间(*PXP4* 共 708 个氨基酸残基)。*PXP5* 的 *ACOX* 区域在 168 ~ 641 位氨基酸残基之间(*PXP5* 共 661 个氨基酸残基)。*PXP4* 与 *PXP5* 的 N 端都包含有 Pfam-B.369 和 Pfam-B.382。此外 ,*PXP4* 的 N 末端还包含一个 Pfam-B.28106。对于这些区域的功能尚不清楚。我们在 Pfam 数据库中以 Pfam-B.369 和 Pfam-B.382 为关键字进行了检索 ,发现任何一个 Pfam-B.369 或 Pfam-B.382 都与 *ACOX* 有关 ,它们的序列都非常保守 ,但对于它们的功能没有说明。

2.4 *PXP4*、*PXP5* 二级结构的预测

用 GOR^[4]方法对 *PXP4* 和 *PXP5* 进行了二级结构的预测。结果见图 3、图 4。

TFTKKNVSVSQGPDPRSSIQKERDSSKWNPQOMNYFLEGSVERSELMKALAQQMERDPILFTDGSYYDLT 70
ccccccccccccchhhcccccccccccccccccccccccccccccccccccccccccccccccccccccccc 140
KDQGRELTAVKINRIARYREQESIDTFNKRLSLIGIFDPQVGTRIGVNLGLFLSCIRGNGTTSQLNYYWAN
hhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhh 210
EKETADVKGITYGCFGMTELAHGSNVAGLETTATFDKESDEFVINTPHIGATKWWIGGAHSATHCSVYAR
hhccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccchhhhce 280
LIVDQGDYGVKTFVVPLRDSNHDLMPGVTVGDIGAKMGRDGDNGWIGFSNVRIPRFFMLQKFKCKVSAEG
ccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccchhhhhhhhhc 350
EVTLPPELQLSYALLGGRVMMVLDSYRMLARMSTIALRYAIGRRQFKGDNVDPKDPNALETQLIDYPLH
ccccccccchhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhh 420
QKRLFPYLAAYVIGSAGALKVEDTIHNTLAELDAAVEKNDTKAIFKSIDDMKSLFVDSGSLKSTATWLGA
hhhhchhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhh 490
EADQCRQACGGHYSSYNGFGKAYNDWVQCTWEGDNNVLAMSVGKPIVKQVISIEDAGKTVRGSTAFI
hhhhhhhhhhccccccccccccccccccccccccccccccccccccccccccccccccccccccccchhhhhhh 560
NQLKDYTGSSSKVVLNTVADLDDIKTVIKAIEVAIIRLSQEAASIVKESFDYVGAELVQLSKLKAHHY
hhccccccccccccccccccccccccchhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhh 630
LLTEYIRRIDTFDQKDLVPYLITLGKLYAATIVLDRFAGVFLTFNVASTEAITALASVQIPKLCAEVRPN
hhhhhhhhccccccccccccchhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhh 700
VVAYTDSFQQSDMIVNSAIGRYDGDYENYFDLVKLQNPPSKTKAPYSDALEAMLNRPLDERERFEKSD
ccccccccccccchhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhh 708
ETAAILSK
hhheeeec

GOR4 :

Alpha helix	(Hh) :	307 is	43.36%
Beta sheet	(Ee) :	116 is	16.38%
Random coil	(Cc) :	285 is	40.25%

图3 PXP4 二级结构预测结果

Fig.3 GOR4 result for PXP4

从二级结构预测结果可以看出,PXP4 和 PXP5 的二级结构非常相似。对 N 端 Pfam-B_369 和 Pfam-B_382 区域二级结构的统计如下(表 6,表 7):

表 6 Pfam-B_369(81 个氨基酸残基)区二级结构

Table 6 Secondary structure in Pfam-B_369(81 amino acid residues)

Peptide	Secondary structure type	Number of segment	Number of residue	Percentage/%
PXP4	α-helix	3	43	53.1
	β-sheet	3	12	14.8
	Random coil	6	26	32.1
PXP5	α-helix	3	40	49.4
	β-sheet	3	12	14.8
	Random coil	5	30	35.8

PTELQKERELTKFNPKELNYFLEGSQERSEIIISMVSEQMKDPIKVDASYNNLTkdQGREVTAKK|ARL 70
ccccchhhhhccccceeeccccchhhhhhhhhhhhhccccceeeccccccccchhhhhhhhhhhhhhhhh
SRYFEHEYPDQQAQRLSILGVFDPQVFTRI|GVNLGLFVSCVRNGTNSQFFYWT|NKGI|DKLRGI|YGCFG 140
hhhhccccchhhhhhhhhccccccccceeeccccceeeeeeccccccccceeeeeeccccccccceeecccc
MTELAHGSNVQGIETATFDEDTDEFVINTPHI|GATKWW|GGAAHSATHCSVYARLKVKGKDYGVKTFVV 210
ceeeccccccccceeeeeeccccccccceeeccccceeeeeeccccccccchhhhhhhheeeccccccccceeee
PLRDSNHDLPEPGTVGDI|GAKMGRDGI|DNGW|IQFSNVR|PRFFMLQKYCKVSRLGEVTMPPEQLSYSAL 280
ccccccccccccceeeeeeccccccccccccceeeccccccccceeeeeeccccceeeeeeccccccccchhhhhhe
IGGRVTMMDSYRMTSRFI|T|ALRYA|IHRRQFKKKD|TDT|ETKL|DYPLHQKRLFPFLAAAYLFSQGALY 350
ecchhhhhhhhhhhchchhhhhhhhhhhhhhhhhhhccccccccchhhhhccccchhhhhhhchhhhhhhhhhhhh
LEQTMNATNDKLDEAVSAGEKEA|DAA|VESKKLFVASGCLKSTCTWLTAEA|DEARQACGGHGYSSYNG 420
hhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhheeeccccccccccccchhhhhhhhhhhccccceeecccc
FGKAYSOWVVOCTWEGDNNI|LAMNVAKPMVRDLLKEPEQKGLVLSSVADLDDPAKLKAFDHALSGLARD 490
ceeeeeeceeececcccchhhhhhhchchhhhhhhchchhhhhceeececcccchhhhhhhhhhhhhhhhhhhhh
|GAVAEDKGFDI|TGPSLVLVSKLNAHRFL|DGGFFKR|TPEWSEVLRPLGFLYADWILTNFGATFLQYGI| 560
hhhhhhccccccccccccchhhhhhhhhchhhhhccccceeeccccchhhhhccccccccchhhhhccccceeecee
TPDYSRK|SSEHFPALCAKVRPNVVGLTDGFNLTDMMTNAA|GRYDGNVYEHYFETVKALNPPENTKAPY 630
ccccccccccccccccccccccccceeeccccchhhhhhhhhccccceeechhhhhccccccccccccch
SKALEDMLNRPDLEVRERGEKSEAAEILSS 661
Hhhhhhhccccchhhhhhhhhhhhhhhheee

GOR4 :

Alpha helix (Hh) : 258 is 39.03%

Extended strand (Ee) : 122 is 18.46%

Random coil (Cc) : 281 is 42.51%

图 4 PXP5 二级结构预测结果

Fig.4 GOR4 result for PXP5

表 7 Pfam-B_382(68 个氨基酸残基)区二级结构

Table 7 Secondary structure in Pfam-B_382(68 amino acid residues)

Peptide	Secondary structure type	Number of segment	Number of residue	Percentage/%
PXP4	α-helix	1	9	13.2
	β-sheet	6	24	35.3
	Random coil	7	35	51.5
PXP5	α-helix	0	0	0
	β-sheet	7	29	42.6
	Random coil	6	39	57.4

从上面的统计可以看出，Pfam-B_369 区域规则的二级结构主要是 α-螺旋。而 Pfam-B_382 区域规则的二级结构主要是 β-折叠，尤其是 PXP5 的该区域没有 α-螺旋。Jung-J P K 等^[10]测定了猪肝中等碳链长度酰基辅酶 A 脱氢酶的晶体结构。该酶以 FAD 为辅基催化脂酰辅酶 A 脱氢形成反式-2,3-烯酰辅酶 A，其性质与 PXP4 和 PXP5 催化反应的性质相

似。猪肝中等碳链长度酰基辅酶 A 脱氢酶的二级结构分为三个区域,从 N 端到 C 端依次是 α -螺旋区、 β -折叠区和 α -螺旋区。晶体结构中 FAD 核黄素环位于 N 端的 α -螺旋区与 β -折叠区之间的缝隙中。与之相比较, PXP4 和 PXP5 的 N 端也有一个 α -螺旋区(Pfam-B_369) 和一个 β -折叠区(Pfam-B_382)。二者的氨基酸序列在 ACOX 中非常保守,推测这两个 Pfam 可能就是 FAD 结合区。

3 讨论

一般 PCR 产物的克隆,先要进行纯化,以去除产物中短的引物二聚体^[11]。我们在实验中总结出一种简易的方法,直接用 PCR 反应液与 T-载体连接。然后挑选白斑转移到另一平板,观察转化子形成菌落的形态。一般含有大质粒的转化子生长慢,形成的菌落薄。而含有小质粒的转化子生长快,形成的菌落厚实。通过这一步的筛选,可大大消除引物二聚体带来的干扰。然后用苯酚直接破裂培养的大肠杆菌细胞,检测质粒大小。用这种方法进行 1.5kb 和 847bp 片段克隆试验发现,对于 1.5kb 片段,该方法同样简便有效。但对于 847bp 的片段,转化子之间形成菌落的差别不是特别明显,需要仔细观察和扩大筛选范围。

对得到的 POX4 和 POX5 基因测序后发现, POX4 与 GenBank 中 *Candida tropicalis* pk233(ATCC 20336) 的 POX4 完全相同。而 POX5 与 pk233 菌株的 POX5 存在菌株之间的差别。这种差别对蛋白功能的影响有待进一步研究。热带假丝酵母中, CUG 密码子编码的氨基酸不是亮氨酸,而是丝氨酸^[1]。在 POX5 编码区的 793~795 位为 CTG, 编码的氨基酸应该是丝氨酸。POX4 基因的编码区无该密码。

Pfam^[7]由两倍分组成: Pfam-A 由包含确定 HMM(hidden Markov model) 的蛋白质家族组成, Pfam-B 则由不包含在 Pfam-A 中的保守序列片段组成。对 POX4、POX5 编码的蛋白 PXP4、PXP5 在 Pfam 中分析后,得到了蛋白质的不同功能结构域。PXP4 和 PXP5 的 N 端都包含有 Pfam-B_369 和 Pfam-B_382, 二者的具体功能尚不清楚,但它们的序列非常保守且都与 ACOX 相关。对 PXP4、PXP5 的 Pfam 分析并没有直接给出 FAD 结合区域,然而用“FAD binding”关键词在 Pfam 数据库中进行查询,检索到了十二种 FAD 结合方式,说明 PXP4 和 PXP5 结合 FAD 的方式与这十二种 FAD 结合方式不同。

对 PXP4 和 PXP5 进行二级结构预测,发现 Pfam-B_369 区域内的二级结构主要是 α -螺旋,而 Pfam-B_382 区域内的二级结构主要是 β -折叠。过氧化物酶体中的酰基辅酶 A 氧化酶与线粒体中的酰基辅酶 A 脱氢酶在催化脂酰辅酶 A 转变成 2,3-烯酰辅酶 A 过程中具有很多相似性^[12]。Jung-J P K 等^[10]测定了猪肝中等碳链长度酰基辅酶 A 脱氢酶的晶体结构。与该酶结合 FAD 的方式相比较,推测 Pfam-B_369 和 Pfam-B_382 可能与 ACOX 结合 FAD 有关。这种推测尚需进一步证实。

参 考 文 献

- [1] Gaëlle B, Odile O K, Patrick W, et al. *FEBS Letters* 2000 **487** 91~94.
- [2] Stephen P, Tracy R, Kristine D, et al. *Bio/Technology* 1992 **10** 894~898.
- [3] Atsuo T, Mitsuyoshi U. *Mycol Res* 1993 **97** 1005-1010 中国科学院微生物研究所期刊联合编辑部 <http://journals.im.ac.cn>

[4] Koei O ,Takanori T ,Naoki K ,et al . *Proc Natl Acad Sci USA* ,1986 **83** :1232 ~ 1236 .
[5] Koei O ,Hironobu T ,Sakuzo F ,et al . *Gene* ,1987 **58** 37 ~ 44 .
[6] 中国科学院微生物研究所烃代谢组 . *微生物学报* ,1980 **20** 88 ~ 93 .
[7] Alex B ,Ewan B ,Richard D ,et al . *Nucleic Acids Res* ,1999 **27** 260 ~ 262 .
[8] 亚当斯 A ,戈特施林 D E ,凯泽 C A 等著(刘子铎译) . *酵母遗传学方法实验指南* .北京 : 科学出版社 ,2000 .87 ~ 88 .
[9] Garnier J ,Gibrat J F ,Robson B ,et al . *Methods in Enzymology* ,1996 **266** 540 ~ 543 .
[10] Jung-Ja P K ,Jin W . *Proc Natl Acad Sci USA* ,1988 **85** 6677 ~ 6681 .
[11] 迪芬巴赫 C W ,德维克斯勒 G S 著(黄陪堂 俞炜源 陈添弥等译) *PCR 技术实验指南* .北京 : 科学出版社 ,1998 .380 ~ 381 .
[12] Zhao-yun J ,Colin T . *Biochemistry* ,1983 **22** 3752 ~ 3758 .

**Analysis of *POX4* and *POX5* Gene Encoded Proteins of
Candida tropicalis 1230**

Qin Wenyan Ren Gang Rong Dong Chen Yuantong

(State Key Laboratory of Microbial Resources ,Institute of Microbiology ,CAS ,Beijing 100080 ,China)

Abstract : *POX4* and *POX5* were cloned from *Candida tropicalis* with a simple method .The sequencing results show differences of *POX5* between strain 120 and strain pk233 .The searching of these two sequences against the current collection of Pfam profile HMMS located regions that belong to known domain families .Secondary structure prediction shows that N-terminal secondary structure of PXP4 and PXP5 are homologous to that of pig liver medium-chain acyl-coA dehydrogenase .Thus , the N-terminal of PXP4 and PXP5 may be involved in FAD binding .

Key words : *Candida tropicalis* , PCR , Pfam , Secondary structure prediction



《微生物学报》入选“ 中国科技期刊方阵 ”

近期 ,中国科学技术部公布了我国期刊进入“ 中国期刊方阵 ”的名单 .全国共有 716 种期刊进入了“ 中国期刊方阵 ” ,其中“ 双高 ”期刊 40 种 ;“ 双奖 ”期刊 58 种 ;“ 双百 ”期刊 122 中 ;“ 双效 ”期刊 496 种 .《微生物学报》入选“ 双百 ”期刊方阵 .

“ 中国期刊方阵 ”的建设是现阶段我国期刊出版事业发展的需要 ,是推进新世纪我国期刊发展的战略性举措 ,它将促进我国期刊“ 精品战略 ”的实施 .

《微生物学报》编委会感谢各级领导的关心 ,感谢广大作者、读者对本刊办刊工作的大力支持 .我们将更好地贯彻国家有关办刊工作的各项政策、法规和法令 ,努力适应国家改革开放的需要 ,为我国的经济建设服务 ,为使期刊尽快走向世界做出积极贡献 .