

文章编号:1004-7271(2004)03-0193-05

# 从福尔马林保存的牙鲆中克隆的 18S rDNA 进行系统发育分析的可靠性

鲍宝龙<sup>1</sup>, 杨桂梅<sup>1</sup>, 龚小玲<sup>1</sup>, 黄连安<sup>1</sup>, 任大明<sup>2</sup>

(1. 上海水产大学生命科学与技术学院, 上海 200090;

2. 复旦大学生命科学学院, 上海 200433)

**摘要:**从福尔马林保存的牙鲆和活体牙鲆肌肉组织中, 利用基因拼接的方法首次克隆到 1205bp 和 1295bp 的长片段序列。福尔马林保存标本与活体标本 DNA 序列的同源率为 82%。和活体标本的 18S rDNA 序列相比, 标本 DNA 的碱基缺失率为 7%, 碱基替换比例为 11%。利用福尔马林标本的 18S rDNA 同源性高的区域进行系统发育分析, 表明其结果是可信的。

**关键词:**牙鲆; 18S rDNA; 福尔马林保存; 系统发育分析

中图分类号: S917 文献标识码: A

## On the rationality of phylogenetic analyses of 18S rDNA cloned from formalin-fixed *Paralichthys olivaceus*

BAO Bao-long<sup>1</sup>, YANG Gui-mei<sup>1</sup>, GONG Xiao-ling<sup>1</sup>, HUANG Lian-an<sup>1</sup>, REN Da-ming<sup>2</sup>

(1. College of Aqua-life Science and Technology, Shanghai Fisheries University, Shanghai 200090, China;

2. Life Science College, Fudan University, Shanghai 200020, China)

**Abstract:** Because some species of fishes have been becoming more and more difficult to be collected, many phylogenetic researchers have no other choice but adopt formalin-fixed specimens. There are many reports on molecular phylogenetic analyses based on DNA sequences from formalin-fixed specimens, but seldom researches have discussed the reliability about the DNA sequences from formalin-fixed specimens in applying phylogenetic analysis. Here the partial sequences, 1205bp and 1295bp 18S rDNA we reported, were first cloned from muscle of formalin-fixed and live *Paralichthys olivaceus* by using gene overlap strategy. Compared with live specimen, 80% bases of 18S rDNA sequence are identical, 7% are deleted, and 11% are replaced in formalin-fixed specimen. The DNA sequences from formalin-fixed and live specimens were used simultaneously in phylogenetic analysis based on high homology area. We think it is reliable in high-class molecular phylogenetic analysis.

**Key words:** *Paralichthys olivaceus*; 18S rDNA; formalin-fixed; phylogenetic analyses

利用从福尔马林长期保存的鱼类中克隆的 DNA 序列进行分子系统发育研究的可靠性一直受到怀疑, 主要认为福尔马林会造成标本中的 DNA 片段化, 以及碱基的缺失、转换等<sup>[1, 2]</sup>。但由于某些鱼类标本已很难采集, 有些研究不得不借助馆藏标本。国内鱼类标本大多保存在福尔马林溶液中, 利用福尔马

林长期保存的标本进行系统发育研究也有一些报道<sup>[3,4]</sup>,但在利用其 DNA 序列进行系统发育分析的可靠性方面却没有研究报道。

在鱼类分子系统发育研究中,高阶元(属及属以上)的系统发育分析主要采用 18S rDNA<sup>[5,6]</sup>。本实验将以福尔马林长期保存的牙鲮(*Paralichthys olivaceus*)标本和活体牙鲮的 18S rDNA 序列为对象,探讨如何克隆较长的基因片段和分析福尔马林对 DNA 序列变化的影响以及利用其序列进行系统发育研究的可靠性。

## 1 材料和方法

### 1.1 样品

福尔马林长期保存的牙鲮为上海水产大学鱼类室 1976 年采集的馆藏标本,活体牙鲮于 2002 年购于上海市铜川水产品批发市场。

### 1.2 基因组 DNA 提取

#### 1.2.1 预处理

分别取福尔马林保存的牙鲮标本和活体标本的背部肌肉组织 0.3g。预处理主要采用 TE9.0 置换,具体步骤参考[3]。

#### 1.2.2 DNA 抽提

活体标本的 DNA 抽提采用酚和蛋白酶 K 法具体步骤参考[7]。福尔马林长期保存标本的 DNA 抽提方法略作调整,具体步骤参考文献[3]。

### 1.3 基因克隆

采用的两对脊椎动物保守的引物:一对为 VP1(5'-CTGGTTGATCCTGCCAG)和 VP2(5'-GAA TTACCGCGGCTGCTG),另一对为 VP3(5'-GTGCCAGCAGCCGCGG)和 VP4(5'-CGGCCATGCACC ACC)<sup>[8]</sup>,两对引物扩增区域有重叠序列,希望通过拼接来扩增 1200bp 左右的 18S rDNA 序列。上述两对引物均为上海生工公司合成。

PCR 体系的反应体系为 25 $\mu$ L,包括引物(6.25 $\mu$ mol/L)各 1 $\mu$ L,dNTPs(2.5 $\mu$ mol/L)1 $\mu$ L,10 $\times$  PCR Buffer 2.5 $\mu$ L,Taq 酶 2.5U,模板 DNA 1 $\mu$ L。PCR 反应过程为 94 $^{\circ}$ C 4';35 循环中 94 $^{\circ}$ C 45",53 $^{\circ}$ C(VP1、VP2)或 60 $^{\circ}$ C(VP3、VP4)45",72 $^{\circ}$ C 1';72 $^{\circ}$ C 6'。PCR 产物用 0.8% 琼脂糖凝胶电泳检测。

PCR 产物胶回收、TA 法克隆等按常规方法进行。克隆的质粒送上海亚博生物工程公司进行 DNA 测序。

### 1.4 序列比对分析和进化树分析

应用 ClustX1.81 分析软件进行完全比对后,再用 Bioedit5.0.9 软件进行手工比对。分子进化分析采用 phylip3.5 软件包。

## 2 结果

### 2.1 18S rDNA 的扩增

从图 1 可以看到,从福尔马林长期保存的牙鲮标本中能扩增出 700bp 左右的基因片段。

### 2.2 18S rDNA 序列比较

通过拼接两对引物 VP1、VP2 和 VP3、VP4 扩增的序列,得到了活体牙鲮 18S rDNA 长 1295bp 的序列,福尔马林保存的 1205bp 序列。

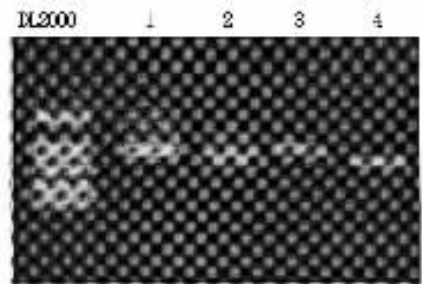


图 1 PCR 扩增产物检测电泳图

Fig.1 Eletrophoretic pattern of the PCR products  
注:1.福尔马林固定牙鲮 VP3、VP4 引物的扩增长度 2.福尔马林固定牙鲮 VP1、VP2 对引物的扩增长度 3.活体牙鲮 VP3、VP4 引物的扩增长度 4.活体牙鲮 VP1、VP2 对引物的扩增长度

福尔马林保存标本与活体标本 DNA 序列相同的碱基有 1063bp,占活体标本 18S rDNA 1295bp 长的 82%,尤其是序列中含有高度同源的区域,可能可以加以利用。和活体标本的 18S rDNA 序列相比,标本 DNA 的碱基缺失 91bp,其中缺失最大的片段有 31bp,占总缺失的 34%;碱基替换 142bp,替换比例为 11%,转换/颠换比例为 1.5(见图 2)。

### 2.3 系统发育分析

把以上福尔马林保存的牙鲆 18S rDNA 序列(共 1205bp)通过 Blastn 在 genebank 进行同源比较,同源性最高的是莫桑比克罗非鱼(*Oreochromis mossambicus*),其 18S rDNA(序列号:AF497908.1)的 381~895bp 之间序列和福尔马林保存的牙鲆 18S rDNA 的 688~1205bp 间有很高的同源性(score = 892 bits)。进一步以福尔马林保存的牙鲆 18S rDNA 的 688~1205bp 间的序列重新在 genebank 用 blastn 进行同源搜索,把搜索结果中包含的 22 种硬骨鱼类的 18S rDNA 相对应的序列进行进化分析,同时包含本实验新鲜牙鲆的 18S rDNA 相对应序列,采用最大简约性法构建系统树,来探讨应用福尔马林浸泡标本的 18S rDNA 应用分子系统发育分析的可靠性。这 22 种硬骨鱼类及它们 18S rDNA 序列号分别是非洲肺鱼一种(*Protopterus* sp.) (AF188370.1),弓鳍鱼(*Amia calva*) (AF188368.1),雀鳝属一种(*Lepisosteus osseus*) (AF188369.1),花鲈(*Lateolabrax japonicus*) (AB089346.1),鲀属一种(*Tetraodon nigroviridis*) (AJ270031.1),长鳍魮属一种(*Sebastolobus altivelis*) (M91182.1),莫桑比克罗非鱼(*Oreochromis mossambicus*) (AF497908.1),罗非鱼属一种(*Oreochromis esculentus*) (AF337051),真鲷(*Chrysophrys major*) (AB028214.1),黄鳍棘海鲷(*Acanthopagrus latus*) (AB089345.1),牙鲷属一种(*Dentex dentex*) (AB089347.1),鲷形目一种(*Fundulus heteroclitus*) (M91180.1),河鲈(*Salmo trutta*) (X98843.1),青鳉(*Oryzias latipes*) (AB105163.1),大麻哈鱼属一种(*Oncorhynchus mykiss*) (AF308735.1),鲑(*Salmo salar*) (AJ427629.1),蛇鳗属一种(*Ophichthus rex*) (X98843.1),海鲢属一种(*Elops hawaiiensis*) (X98841.1),北梭鱼(*Albula vulpes*) (X98842.1),海鲢目一种(*Megalops atlanticus*) (X98846.1),月目鱼(*Hiodon alosoides*) (X98840.1),鳗鲡目一种(*Echiophis punctifer*) (X98844.1)。

图 3 系统树中月目鱼、海鲢、北梭鱼、蛇鳗属的一种、鳗鲡距离较近,非洲肺鱼、弓鳍鱼及雀鳝距离较近但距其它种类较远,其余种类基本上分为距离较近的两大类,一大类基本上是以大麻哈鱼属的一种、鲑、河鲈为主,另一类则是以花鲈、鲀属一种、长鳍魮属一种、莫桑比克罗非鱼、罗非鱼属一种、真鲷、黄鳍棘海鲷、牙鲷属一种等构成,这个系统树揭示的系统关系基本符合目前国际上普遍接受的硬骨鱼类分类系统<sup>[9]</sup>的框架,系统树中福尔马林保存的和活体牙鲆的亲缘关系最为接近。

## 3 讨论

福尔马林会导致蛋白质变性,样品处理过程中残留的福尔马林会导致蛋白酶 K 的消化效果大大降低。在随后的酚氯仿抽提中 DNA 会随蛋白质而一同丢失<sup>[1]</sup>。通过饱和酚重复抽提,能提高 DNA 得率。否则 DNA 浓度太低而无法进行 PCR 扩增。另外,不能用饱和酚反复抽提, DNA 浓度高,其含福尔马林浓度相对就高,从而抑制 Taq 酶活性导致 PCR 失败。可以通过减少饱和酚的抽提次数或稀释 DNA 模板,以降低福尔马林的浓度。

通过基因拼接的思路,本实验首次报道了从福尔马林长期保存牙鲆标本中扩增出 1205bp 长的 18S rDNA 大片段基因,这为解决福尔马林造成标本基因的片段化而使扩增长片段基因的困难,提供了新的思路。因为,目前报道的利用福尔马林保存的鱼类标本扩增出的基因长度在 300~700bp 之间<sup>[2,4]</sup>,没有超过 1kb 的长片段报道。同时,由于标本 DNA 扩增时易发生的 Jumping PCR 现象<sup>[10]</sup>会产生片段缺失(本实验结果也反映了片段和碱基缺失比较普遍存在的现象)短的序列很难直接应用分子系统发育分析。

通过和活体牙鲆 18S rDNA 序列比对分析,发现序列中在某一区域内有很大的缺失和碱基差异,而在其他区域相对要低得多,而长片段基因序列为利用其作为系统发育研究提供了可能。

本研究的系统分析表明,利用从福尔马林长期保存鱼类中克隆出的 18S rDNA 同源性相对较高的区

		VP1 →
Live	1	<u>CTGGTTGATCCTGCCAGTAG</u> -CATATGCTTGCTCAAAGATTAAGCCATG
Formalin-fixed	1	*****T*****C*****
Live	50	CAAGTCTAAGTACACACGGCCGGTACAGTGAAACTGCGAATGGCTCATTA
Formalin-fixed	51	**T*****T*A**AA*TA-*****C*****
Live	100	AATCAGTTATGGTTCTTTGATCGCTCTCACGTTACTTGGATAACTGTGG
Formalin-fixed	100	*****C**TA*****A*T*T*-----A*****C**A
Live	150	CAATTCCAGAGCTAATACATGCCAACGGGCGCTGACCTTCGGGGACGCGT
Formalin-fixed	146	T*C*T***A*****A*G*-----*T*T*-----C
Live	200	GCATTTATCAGACCCAAAACCCATGCGGGGTGCCACCACCGGGGTGCCCC
Formalin-fixed	180	*T*A**T***TA***C*-----
Live	250	GGCCGCTTTGGTGACTCTAGATAACCTCGAGCTGATCGCTGCCCTCGTG
Formalin-fixed	199	-AA*****T**---**TA**T*TC**--A*C*AT*T*---
Live	300	GCGCGACGTCTCATTGCAATGCTGCCCTATCACTTTCGATGGTACTT
Formalin-fixed	239	---*G*T*A*****A**T*****G*****GGA
Live	350	TTTGTGCCTACCATGGTGACCACGGGTAACGGGAATCAGGGTTCGATTC
Formalin-fixed	286	*CGTG*****A**A*****A*****
Live	400	CGGAGAGGGAGCCTGAGAAACGGCTACCACATCCAAGGAAGCAGCAGGC
Formalin-fixed	336	*****
Live	450	GCGCAAATTACCCACTCCGACTCGGGGAGGTAGTGACGAAAAATAACAA
Formalin-fixed	386	*****A**T**A*A*****A*T*T**G*
Live	500	TACAGGACTCTTCGAGGCCCTGAATTGGAATGAGTACACTTAAATCC
Formalin-fixed	436	*C*G*G*****-GA*T*TC*G*****A*****CA*
		VP3 →
Live	550	TTAACGAGGATCCATTGGAGGGCAAAGTCTGGT <u>GTGCCAGCAGCCGCGG</u>
Formalin-fixed	484	C*****A*A*****-*****-*****
		← VP2
Live	600	<u>TAATTC</u> CAGCTCCAATAGCGTATCTTAAAGTTGCTGCAGTTAAAAGCTC
Formalin-fixed	531	*****A*τ*****T*****
Live	650	GTAGTTGGACCTCGGGATCGAGCTGACGGTCCGCCGCGAGGCGACTACC
Formalin-fixed	581	*****T*T**TCGT*AGG*T*****T*AC**T*---*T
Live	700	GTCTGTCCCAGCCCTGCCTCTCGGCCCCCTCGATGCTCTTAGCTGAG
Formalin-fixed	630	***-G**G*G**T*T**T**T*AT*TG*T*---*T*CG*A*CA
Live	750	TGTCCCGCGGGTCCGAAGCGTTACTTTGAAAAATTAGAGTGTCAA
Formalin-fixed	674	-----**AA**AGGA*C*****
Live	800	GCAGCCCGTCCGCTGAATACCGCAGCTAGGAATAATGGAATAGGACTC
Formalin-fixed	717	*****
Live	850	CGGTTCTATTTGTGGTTTCTCTCTGAAGTGGGCCATGATTAAGAG
Formalin-fixed	767	*****
Live	900	GGACGGCCGGGGCATTGATTTGTGCCGCTAGAGTGAAATCTTGGAC
Formalin-fixed	817	*****
Live	950	CGGCGCAAGACGGACGAAAGCGAAAGCATTTGCCAAGATGTTTCATTA
Formalin-fixed	867	*****
Live	1000	ATCAAGAACGAAAGTCGGAGGTTCGAAGACGATCAGATACCGTCGTAGTT
Formalin-fixed	917	*****
Live	1050	CCGACCGTAAACGATGCCAACTAGCGATCCGGCCGCTTATCCCATGAC
Formalin-fixed	967	*****
Live	1100	CCGCCGGGCGCTCCGGAAACCAAAGTCTTTGGGTTCCGGGGGAAGT
Formalin-fixed	1017	****T*****_**
Live	1150	ATGGTTGCAAAAGCTTGAAACTTAAAGGAATTGACGGAAGGCACCACCA
Formalin-fixed	1066	*****_*****
Live	1200	GGAGTGGGAGCCTGCGGCTTAATTTGACTCAACACGGGAAACCTCACCC
Formalin-fixed	1115	*****_*****G**A*****-
Live	1250	GGCCCGGACCCGAAAGGATTGACAGATTGGATAGCTTTTTTCT
Formalin-fixed	1163	A**T**A***-T*ATG*****_*****C***-*

图 2 福尔马林保存和活体牙鲆 18S rDNA 部分序列的差异比较

Fig.2 18S rDNA partial sequence variation comparison between live and formalin-fixed *Paralichthys ovivaceus*  
注 : \* 表示相同碱基 , - 表示缺失。VP3 和 VP4 扩增的片段测序时未能完全测通 故图中未标出 VP4。

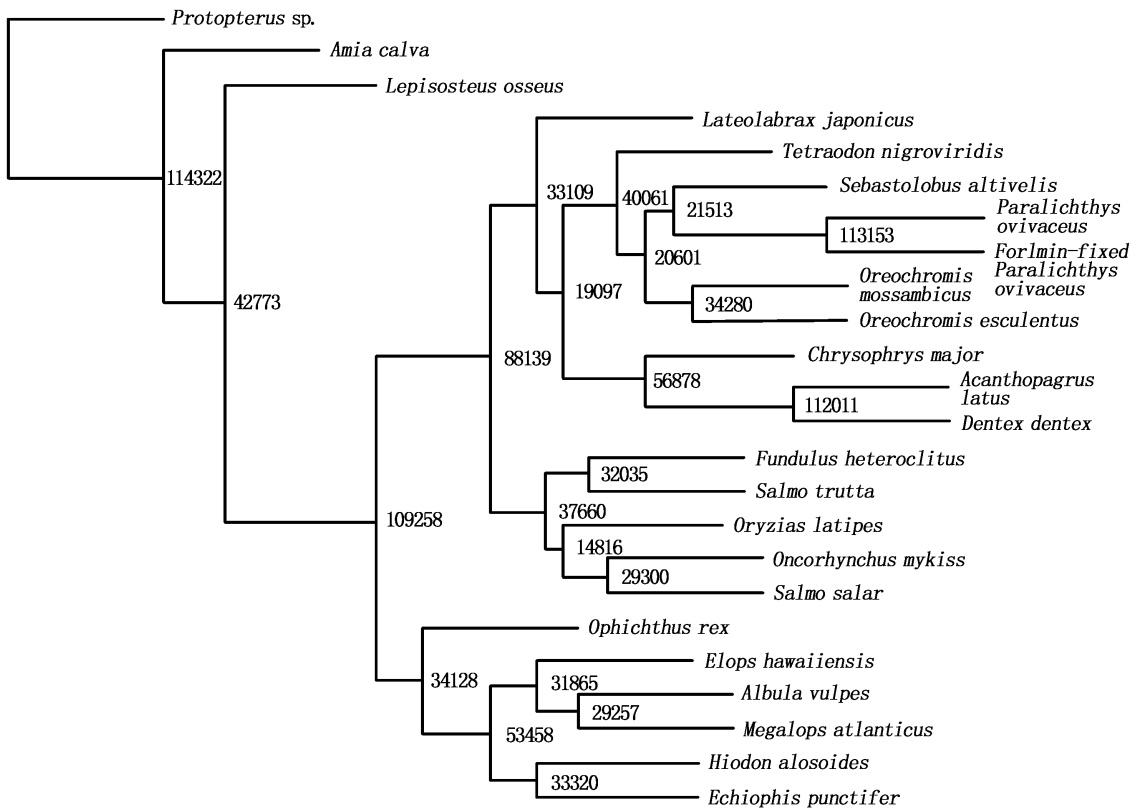


图 3 根据 18S rDNA 部分序列特征而采用最大简约性法构建的硬骨鱼类系统树

Fig.3 Phylogram of Osteichthyes reconstructed with parsimony based on partial sequences of 18S rDNA

域序列,作为高阶元系统发育分析的特征数据是可靠的。

由于从福尔马林长期保存的鱼类中克隆的 DNA 序列存在较为明显的缺失和碱基转换,不如同源分析而直接利用克隆的序列进行分子系统分析是不可靠的。本实验结果提供了从福尔马林长期保存的牙鲆标本中克隆的 18S rDNA 序列同源区域应用高阶元分析的可靠性例证,由于鲽形目鱼类目前没有 18S rDNA 序列公布,有关低阶元系统发育可靠性的分析尚有待今后进一步研究。

#### 参考文献:

- [1] Kaufman B A, Newman S M, Hallberg R L, et al. In organello formaldehyde crosslinking of proteins to mtDNA: identification of bifunctional proteins [J]. Proc Natl Sci USA, 2000, 97(144): 7772 - 7777.
- [2] 庞峻峰, 张亚平. 标本 DNA 研究 [J]. 动物学研究, 2001, 22(6): 490 - 496.
- [3] 肖武汉, 吴春花. 福尔马林固定云南鲷的 DNA 提取及其细胞色素 b 基因序列分析 [J]. 动物学研究, 1997, 18(3): 242, 252, 258, 284.
- [4] He S P, Chen Y J, Zhang Y P. Preliminary study on Mitochondrial Cytochrome b DNA sequences and phylogeny of formalin fixed Sisorid fishes [J]. 动物学研究, 1999, 20(2): 81 - 87.
- [5] Gardiner B G, Greenwood P H, Miles R S, et al. Interrelationships of Basia Neopterygians. Interrelationships of Fishes [C]. San Diego: Academic Press, 1996. 117 - 146.
- [6] Lecointer G, Nelson G. Clupeomorpha, sister-group of Ostariophys [A]. Interrelationships of Fishes [C]. San Diego: Academic Press, 1996. 193 - 207.
- [7] Joseph S, David R. Molecular Cloning: A Laboratory Manual 3rd ed. Cold Spring Harbor Lab [M]. NY: Cold Spring Harbor Press, 1989.
- [8] Peterson K J, Eernisse D J. Animal phylogeny and the ancestry of bilaterians: inferences from morphology and 18S rDNA gene sequences [J]. Evo & Dev, 2001, 3(3): 170 - 205.
- [9] Nelson J S. "Fishes of the world", 3rd ed [M]. New York: Wiley, 1994.
- [10] Paabo S. Amplifying ancient DNA [A]. PCR Protocols: A Guide to Methods and Applications [C]. San Diego: Academic Press, 1990. 159 - 166.