

ICS 07.080

B 50

SZDB/Z

深圳市标准化指导性技术文件

SZDB/Z 315—2018

基于基因条形码的鱼类物种鉴定技术要求

Technical specification for identification of fish species based on DNA barcoding

2018-08-02 发布

2018-09-01 实施

深圳市市场和质量监督管理委员会 发布

目 次

前 言	II
1 范围	1
2 规范性引用文件	1
3 术语与定义	1
4 缩略语	2
5 原理	2
6 仪器与试剂	2
7 取样	3
8 实验步骤	3
9 结果判断	5
附录 A（规范性附录） 试剂配制	6
附录 B（规范性附录） 序列的有效性判定	7
附录 C（资料性附录） 基因条形码在 BOLD 系统中物种鉴定流程示意图	9
附录 D（资料性附录） 基因条形码在 NCBI 系统中物种鉴定流程示意图	11

前 言

本文件根据GB/T 1.1-2009给出的规则起草。

本文件由深圳市检验检疫科学研究院提出。

本文件由中华人民共和国深圳海关归口。

本文件起草单位：深圳市检验检疫科学研究院、深圳出入境检验检疫局动植物检验检疫技术中心、深圳出入境检验检疫局食品检验检疫技术中心、深圳华大海洋研究院、镇江华大检测有限公司。

本文件主要起草人：郑晓聪、刘荭、石琼、王敏、何俊强、黄玉、阮周曦、于力、王津津、贾鹏、刘莹、温志清、史秀杰、兰文升、秦智锋。

基于基因条形码的鱼类物种鉴定技术要求

1 范围

本文件规定了鱼类基因条形码鉴定中的样品处理、基因扩增、序列分析、比对鉴定及结果判定等操作程序和规范。

本文件适用于对自然环境种化产生的野生型鱼类或其残存的组织进行物种鉴定, 不适用于杂交品种。

2 规范性引用文件

下列文件对于本文件的应用是必不可少的。凡是注日期的引用文件, 仅所注日期的版本适用于本文件。凡是不注日期的引用文件, 其最新版本(包括所有的修改单)适用于本文件。

GB/T 6682 分析实验室用水规格和试验方法

FDA SOP for Generating DNA Barcodes Suitable for Species Identification of an Unknown Fish Tissue Sample

3 术语与定义

下列术语与定义适合于本文件。

3.1

基因条形码 (DNA barcoding)

是指可用于物种鉴定的具有生物物种序列特征的DNA片段, 本文件中指鱼类线粒体基因组中的 *COI* 基因片段。

3.2

基因条形码鉴定技术 (DNA barcoding identification technique)

一种利用基因片段进行生物物种鉴定和系统进化关系研究的技术, 建立在基因扩增和序列比对的基础之上, 它具有不受物种发育状态、雌雄性个体及形态特征完整性与否的限制、减少对分类专家的依赖、不受主观因素影响、准确率高、借助互联网可以实现数据共享等特点。

3.3

空白对照 (Blank control)

从基因组DNA提取步骤开始至PCR扩增和PCR产物检测整个过程, 没有加入任何鱼源性成分的空白的灭菌洁净离心管, 与其他受试的样品平行进行操作以观察实验是否处于正常状态。空白对照没有PCR产物条带产生。

3.4

阳性对照 (Positive control)

SZDB/Z 315—2018

在PCR扩增过程中,与受试样品平行进行的对照反应,以在正常反应情况下一定能产生PCR产物的DNA为反应模板,用于观察整个反应体系和反应过程是否正常,阳性对照应有PCR产物带产生。

3.5

阴性对照 (Negative control)

非鱼类样品的其他物种样品,与受试样品平行进行的对照反应,用于观察整个反应体系是否正常,确认没有受到污染,阴性对照没有PCR产物条带产生。

3.6

序列相似度 (Similarity of the sequences)

反映序列间相似程度的数值,一般以百分数表示。

3.7

FASTA格式 (FASTA format)

生物信息学术语,又称Pearson格式,是一种用文本表示核苷酸序列或者氨基酸序列的格式。核苷酸或者氨基酸碱基用单个字母表示,序列的第一行一般用“>”起始,随后可以添加序列名或者注释,第二行为序列本身,只允许使用既定的核苷酸或者氨基酸编码符号,如核苷酸使用ATCG表示,序列中不能出现回车等字符。

示例 1:

> 1

```
GTGTGTGGGTATGATTCTAAACTGACAAATCGCGCCTATTTATACTTTTTCCCCGGGGTTTAAAT  
TCCTTGGCCCCCATGTGGTTTCTTTAATTTAGATTGTGGAACCCATTTTCTTAATCCGTAATCG
```

4 缩略语

下列缩略语适用于本文件。

BOLD: 生物条形码数据系统 (barcoding of life data system)

NCBI: 美国生物技术信息中心 (National center for biotechnology information)

5 原理

线粒体COI基因是位于线粒体DNA的蛋白质编码基因,其长度为700 bp左右,易于扩增,进化速度较快又相对保守,既有足够的变异又便于通用引物扩增,DNA序列很少有插入与缺失现象,便于序列比对分析。利用通用引物对鱼类物种线粒体COI进行PCR扩增,得到目的片段进行纯化、测序,然后将所有序列进行两两比较并计算其差异值,根据差异值来确定物种之间的关系。

6 仪器与试剂

6.1 仪器

水浴锅、恒温箱、纯水器、普通冰箱和超低温冰箱、台式高速冷冻离心机、PCR扩增仪、水平电泳仪、微量移液器、微量核酸蛋白测定仪、紫外观察灯或凝胶成像仪、电子天平、测序仪。

6.2 试剂

6.2.1 试剂级别

除另有规定，试验用水应为符合GB/T 6682规定的一级水，所用化学试剂均为分析纯。

6.2.2 引物

用去离子水将每条引物配制成 100 $\mu\text{mol/L}$ 储存液，置 -20°C 以下冻存；使用时取适量配制成 10 $\mu\text{mol/L}$ 工作液，避免多次冻融。序列如下，

引物	名称	5' -3'	备注
上游引物	FishF2t1	<u>TG</u> TAAAACGACGGCCAGT <u>CG</u> ACTAATCATAAAGATATCGGCAC	其中下划线部分为通用测序引物 M13F(-20)
	VF2t1	<u>TG</u> TAAAACGACGGCCAGTCAACCAACCACAAAGACATTGGCAC	
下游引物	FishR2t1	CAGGAAACAGCTATGAC <u>ACT</u> TGACCGGATGACCGAAGAATCAGAA	其中下划线部分为通用测序引物 M13R(-27)
	FR1dt1	CAGGAAACAGCTATGAC <u>AC</u> CTCAGGGTGTCCGAARAAYCARAA	

6.2.3 其他试剂

醋酸铵、乙醇、异丙醇、蛋白酶 K、高保真 DNA 聚合酶、dNTPs (含 dATP、dTTP、dGTP、dCTP)、 MgSO_4 、PCR buffer (不含 Mg^{2+})、溴酚蓝、标准分子量 (推荐使用 100 bp DNA Marker)、琼脂糖。

7 取样

7.1 取样对象

活鱼、鱼卵，鲜冻的肉块，单一鱼类制作的肉糜等。

7.2 取样部位

宜取鲜活的组织，成鱼宜取但不限于肌肉组织 (心、脑、脾、肾、肝、肠、胃等器官也适用)，鱼卵取卵膜。

如需非致死性取样，剪取部分鳍条或者刮去部分皮肤后取皮下一小块 $3\text{ mm}^3 \sim 5\text{ mm}^3$ 的肌肉块即可。

7.3 样品处理

采集的组织置于冰上低温保存备用，如需长期保存应置于 $\leq -20^\circ\text{C}$ 冷藏或置于 75% 乙醇中固定保存。

8 实验步骤

8.1 设立对照

设置空白对照、阴性对照和阳性对照。

8.2 DNA 提取

取 20 mg 样品组织剪碎置于洁净 1.5 mL 离心管中，加入 600 μL 裂解液 (见附录 A.1) 和 20 μL 蛋白酶 K 溶液 (见附录 A.2)，混匀，于 56°C 水浴 1-3h，期间不时轻微振动混匀；待组织消化完全，冷却至室温，加入 200 μL 醋酸铵溶液 (见附录 A.3)，混匀，冰上放置或 4°C 冷却 5 min~10 min，使蛋白质盐析出来； 4°C ，12 000 r/min 离心 10 min，将上清液转移至另一洁净的 1.5 mL 离心管中， 4°C ，12 000 r/min 离心 10 min，取上清液至另一个洁净 1.5 mL 离心管中，加等体积异丙醇，轻轻摇匀， 4°C 放置 15 min； 4°C ，12 000 r/min 离心 10 min 沉淀 DNA；弃上清，加 1 mL 75% 乙醇，轻轻混匀，于 4°C ，12 000 r/min 离心 5 min，

SZDB/Z 315—2018

弃上清，晾干；干燥后加入50 μL 去离子水溶解；DNA提取液应置于冰上备用，若需长期保存应置于 $\leq -20^\circ\text{C}$ 保存。

也可采用经验证可靠的DNA提取纯化试剂盒进行核酸提取，参照说明书使用。

8.3 DNA浓度和纯度的测定

吸取DNA提取液 2 μL ，以去离子水作为空白对照，使用 NanoDrop微量核酸蛋白测定仪测定 DNA 模板的质量浓度及 A260/A280比值。DNA 的质量浓度根据以下公式计算：DNA 的质量浓度= A260 \times 50 \times 稀释倍数($\text{ng}/\mu\text{L}$)

当A260/A280 比值在1.7 ~ 1.9之间时，适合于PCR扩增。

8.4 PCR 扩增

在0.2 mL PCR反应管中，按表1配制反应体系，加样宜按空白对照、阴性对照、待检样品、阳性对照的次序分别加入模板，盖上管盖，充分混匀。

表1 PCR反应体系(50 μL)

组分	加样量(μL)	含量/浓度
10 \times PCR Buffer (不含 Mg^{2+})	5	1 \times
MgCl_2 (25 mmol/L)	5	2.5 mmol/L
dNTPs (10 mmol/L each)	1	0.2 mmol/L
FishF2t1 (10 $\mu\text{mol/L}$)	0.5	0.1 $\mu\text{mol/L}$
VF2t1 (10 $\mu\text{mol/L}$)	0.5	0.1 $\mu\text{mol/L}$
FishR2t1 (10 $\mu\text{mol/L}$)	0.5	0.1 $\mu\text{mol/L}$
FR1dt1 (10 $\mu\text{mol/L}$)	0.5	0.1 $\mu\text{mol/L}$
高保真 DNA 聚合酶 (5 U/ μL)	0.5	0.05 U/ μL
样品核酸/对照	1	50 ng ~500 ng
去离子水	35.5	补充至 50 μL

8.5 PCR 反应程序

将已加样的PCR反应管短暂离心后放入PCR仪，扩增反应条件设定：94 $^\circ\text{C}$ 预变性2 min；94 $^\circ\text{C}$ 变性30 s，50 $^\circ\text{C}$ 退火30 s，72 $^\circ\text{C}$ 延伸45 s，30个循环；72 $^\circ\text{C}$ 7 min；4 $^\circ\text{C}$ 保存。

8.6 琼脂糖电泳

用新鲜配制的1 \times TAE电泳缓冲液（见附录A.4）配制含1 $\mu\text{g/mL}$ 溴化乙锭（见附录A.5）的2%琼脂糖凝胶。取适量扩增产物和溴酚蓝指示剂按比例混匀后进行电泳检测，设DNA Marker同步电泳；采用5 V/cm \sim 8 V/cm，电泳1 h，用凝胶成像仪或紫外透射仪观察，拍照记录结果。

8.7 电泳结果

当阳性对照扩增片段大小约700bp，阴性对照和空白对照没有扩增片段时，结果成立。此时，若待测样品出现约700bp扩增片段则判断该样品PCR扩增结果阳性，无扩增片段或者片段大小不符则判断为PCR扩增结果阴性。

8.8 序列的测定

将PCR扩增结果阳性产物进行序列测定，测序引物使用通用引物M13F (-20)和M13R (-27)。

8.9 序列的有效性分析

通过查看序列的原始峰图，有效序列段峰高、基部杂峰少且低的序列（遵照附录B中的B.1执行）且长度大于500 bp为有效序列。将符合要求的序列进行拼接，去除引物片段后应能得到一个具有完整开放阅读框的拼接序列，并可翻译成一条完整的蛋白质序列；否则序列仅供参考（遵照附录B中的B.2执行）。

8.10 序列比对鉴定

与BOLD系统或NCBI的序列进行比对。具体操作遵照附录C和附录D进行。

9 结果判断

- 9.1 根据序列的相似度判定鱼类的种类，相似度最高且 $>98.0\%$ 者，可判定为同一种类。
- 9.2 若相似度 $>98.0\%$ 时系统给出两个或以上的物种，需参考形态学特征进行综合判断。
- 9.3 序列相似度 $\leq 98.0\%$ 者，不能进行种类的判断。
- 9.4 当 BOLD 与 NCBI 给出不同结论时，以 BOLD 给出的结论为准，NCBI 给出的结论供参考。

附录 A
(规范性附录)
试剂配制

A.1 裂解液

1 mol/L Tris-HCl (pH8.0)	1 mL
0.5 mol/L EDTA	1 mL
0.5 mol/L NaCl	20 mL
10% SDS	10 mL
定容至100 mL	

A.2 蛋白酶K溶液 (20 mg/mL)

将200 mg的蛋白酶K加入到9.5 ml水中，轻轻摇动，直至蛋白酶K完全溶解。不要涡旋混合。加水定容到10 ml，然后分装成小份贮存于-20℃

A.3 醋酸铵溶液 (7.5 mol/L)

称取578.12 g醋酸铵，加去离子水定容到1 L。

A.4 1×TAE电泳缓冲液

称取242 g Tris和37.2 g $\text{Na}_2\text{EDTA}\cdot 2\text{H}_2\text{O}$ 加800 mL去离子水充分搅拌溶解后，加入57.1 mL的醋酸，用去离子水定容至1L，即为50×TAE电泳缓冲液，灭菌后室温保存。使用时用去离子水稀释50倍即得1×TAE电泳缓冲液。

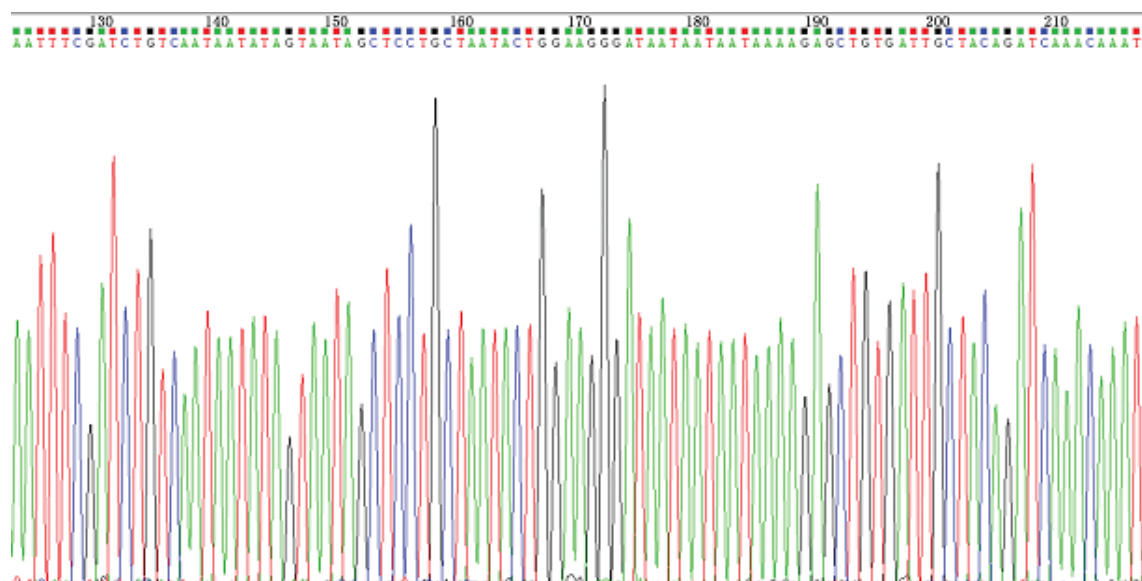
A.5 溴化乙锭 (Ethidium Bromide, EB)

用去离子水配制成10 mg/mL的浓缩液。用时每100 mL 琼脂糖凝胶中加10 μL 。

附录 B (规范性附录) 序列的有效性判定

B.1 通过原始峰图判断

如下图所示，有效序列段峰高、基部杂峰少且低的序列，且长度大于500 bp为有效序列。



图B.1 原始峰图

B.2 通过翻译成氨基酸判断

B.2.1 登陆网址：http://www.ebi.ac.uk/Tools/st/emboss_transeq，序列输入窗口截图见图B.2。

B.2.2 将序列复制至上图Step 1中的空格处，Step 2中的Frame栏中选择6 (All Frame)，Condon Table中选择“Vertebrate mitochondrial”，点击Submit，得到如图B.3所示画面。

B.2.3 给出的六条氨基酸序列中，只要有一条没有星号，表明可以得到一条具有完整开放阅读框的拼接序列，并能翻译成完整的蛋白质，即可判定为有效序列；否则序列仅供参考。

The screenshot shows the EMBOSS Transeq web interface. At the top, there is a navigation bar with 'Input form', 'Web services', and 'Help & Documentation'. Below this, the page title 'EMBOSS Transeq' is displayed. The main content area is divided into three steps:

- STEP 1 - Enter your input sequence:** A text input field with a dropdown menu set to 'DNA/RNA'. Below the field is a large empty text area for pasting a sequence. An option to 'upload a file' is also present.
- STEP 2 - Select Parameters:** Two dropdown menus are visible. The first is labeled 'FRAME' and is set to '6 (All six frames)'. The second is labeled 'CODON TABLE' and is set to 'Vertebrate Mitochondrial'. A note below states: 'The default settings will fulfill the needs of most users. More options... (Click here, if you want to view or change the default settings.)'
- STEP 3 - Submit your job:** A checkbox for 'Be notified by email' and a 'Submit' button.

图B.2 序列输入窗口示意图

The screenshot shows the EMBOSS Transeq web interface displaying the results of a translation job. The page title is 'EMBOSS Transeq'. Below the navigation bar, the job ID is 'emboss_transeq-l20170919-091228-0485-87689905-es'. There are tabs for 'Tool Output' and 'Submission Details'. Below the tabs are buttons for 'Download' and 'Show Colors'. The main content area displays the translated protein sequences for three different frames:

```

>EMBOSS_001_1
TFL**FSL**CQLWLGAEITD*PHL*LVPLMWFSPWTT*ASDYSFHLSPFFSLPGLKP
GLALGEQFIHWRVILPMQEHQST*PFSPLT*RVFHQFLVQLTLLPSSMWNFLPFSI*
PPCSFGPFLPLSPFFYPCCQLLGLQCS*PIVI*TLFSLTQLEETQFFINTCFDPLVT
X
>EMBOSS_001_2
RICKRFLYSNAHYD*GLWKLINSTYNWCFWYGLFPNEQHKLITPFFSSTRFLRWSR
GWHVNSLSTTGG*SCPC*SISRPNHLFSSLSGYFINS*CH*LIYYHQYETPCYFFVSD
PLVRLGRSYRCPSS*TPASSCC*QWAPRNS*GHWLL*PSK***PNSLSTLVLIFWSP
>EMBOSS_001_3
AFVMIFFWVPMI*GGPGRWLIPLMIGAFDMAFPMMNMSFWLLPFSFLLASSGVEAG
AGTGWVYVPLAGLHAHAGASVDLTI*FSLHLAGISSILGAINFITIINMKPFAISQIQT
PLVWAVLITAVLILLSLPVLAAGITMLLDRMLNTFFPDPAGGGDPIYQLHFFFGHP
>EMBOSS_001_4
GQKIKTSYDKELGLLQIGQ***CLYDRLGAL*SLQQLAGMKKEEQR**ERPKRT*G
SDFGKQVSTVW***S*LQELMKPLSEK*WLGRLMLLHQDFFVVDKLTQCQP
RLAFR*RVVEEGEMGGV*SLCCSPGG*PIQ*HQL*VELISTQSFSQ*LALL**KLLQMR
>EMBOSS_001_5
GFFKQKQK**IGSFPFAGSKVVF*LESV*SIVIPAA*TG*DE***TAVM*TAQVHG
VYVEMAGGMMLMVMKLIAP*IDEMAPK*EDM*STDAFA*LEASGG*TVHVEA
PASTPEAS***KDGSQKMLFIRKAMS*AFIMSGINGPFPKPPIMI*GITMKLITNA
X
>EMBOSS_001_6
VTKRSQVLMKRWVSSSVKESV*ITIG*EHCNFCSKWQ*IKKDS*GKNGFNEQG
LM/LGNS*GFHDDGSKVCTKQWNTA*VK*KD*VDWCSMGKITBQ*WNCSPASP
GFNPGSE*IKK*GE*SEAYVHSGEHI*GTNYKN*SVSKAPRHWYIKENYIKV

```

图B.3 输出的翻译蛋白质序列窗口示意图

附录 C (资料性附录)

基因条形码在 BOLD 系统中物种鉴定流程示意图

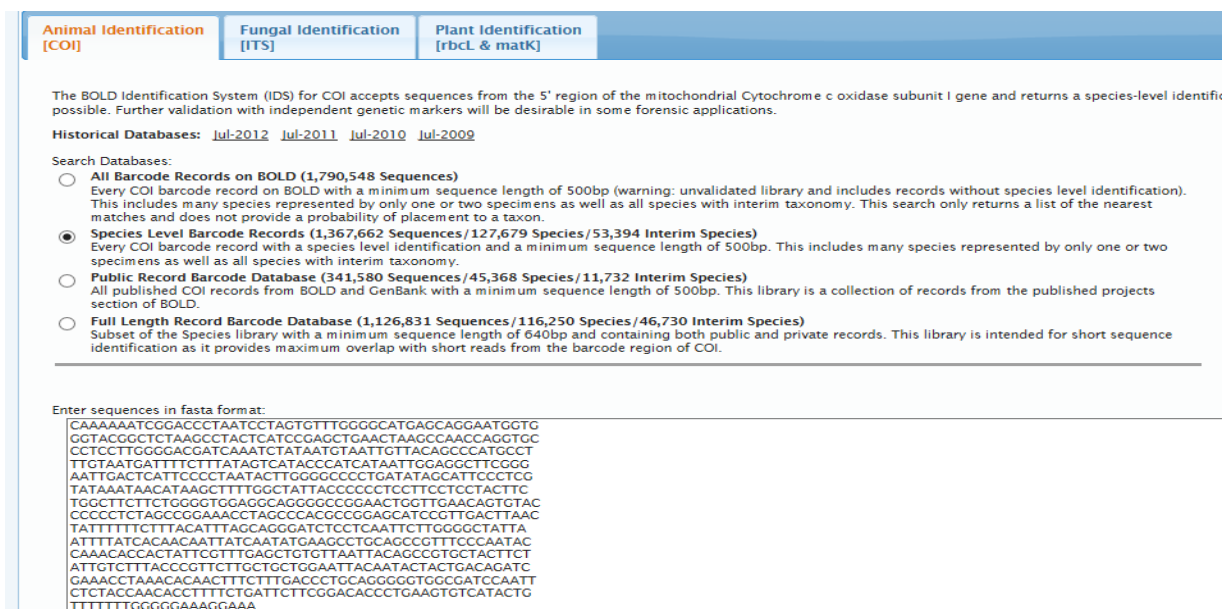
C.1 鉴定流程

C.1.1 登陆网站<http://v4.boldsystems.org>，进入下图所示画面。



图C.1 BOLD 登陆窗口示意图

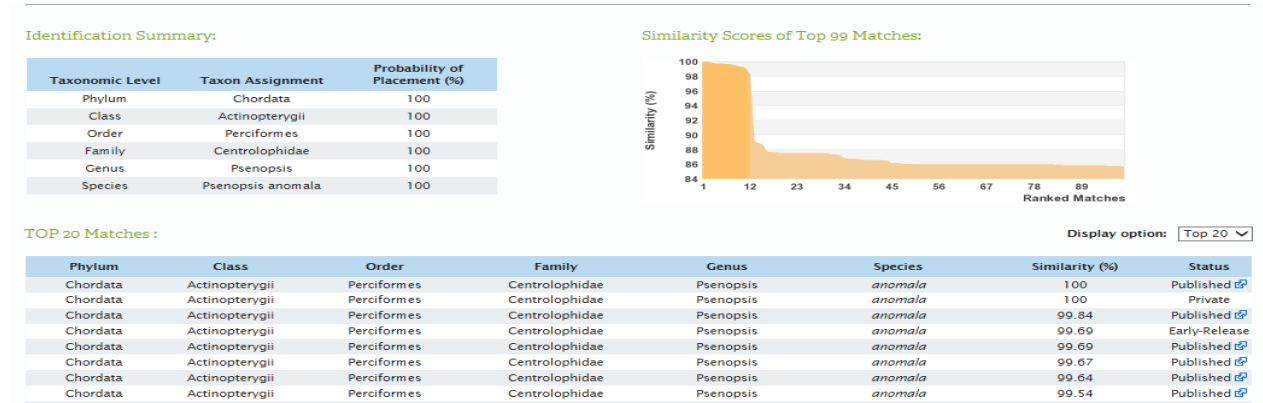
C.1.2 点击Identification进入下图所示画面。



图C.2 序列输入窗口示意图

将有效的序列以Fasta格式黏贴在上方的空格内，点击Submit。

C.1.3 得到如下图所示报告。



图C.3 BOLD 系统 BLAST 比对呈现结果

C.2 结果判定

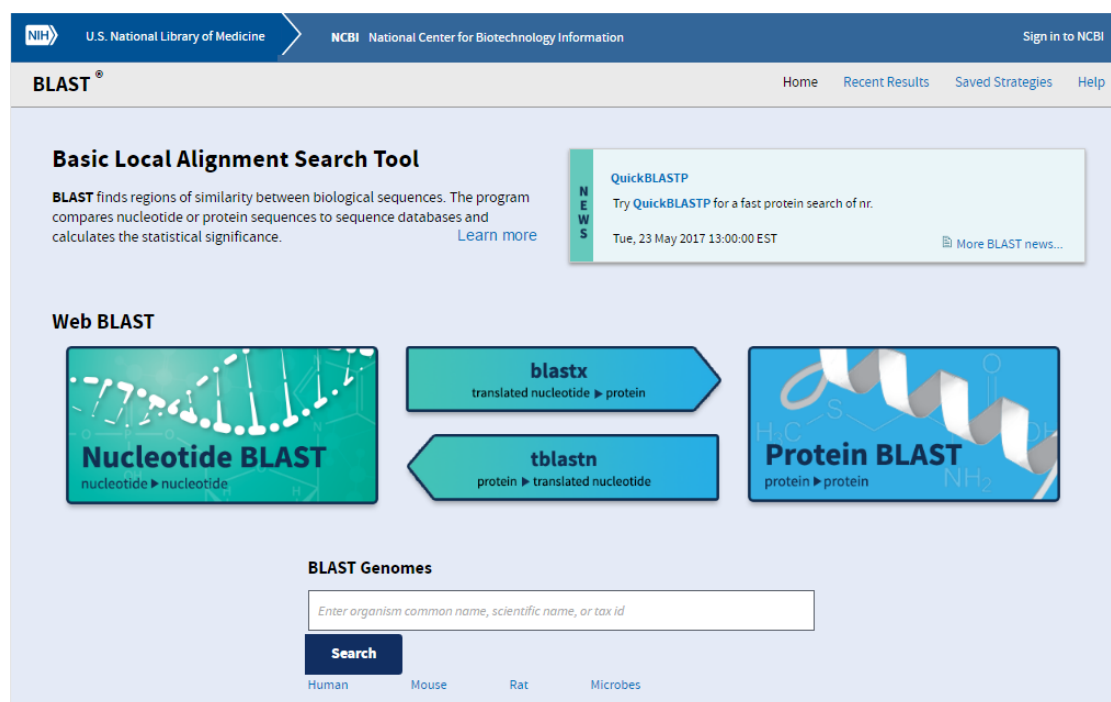
根据序列的相似度判定鱼类的种类，相似度最高且>98.0%者，可判定为同一种类。若相似度>98.0%时系统给出两个或以上的物种，需参考形态学特征进行综合判断。序列相似度≤98.0%者，不能进行种类的判断。

附录 D (资料性附录)

基因条形码在 NCBI 系统中物种鉴定流程示意图

D.1 鉴定流程

D.1.1 登陆网站 <https://www.ncbi.nlm.nih.gov/blast.cgi>，进入如下图示画面。



图D.1 NCBI 登陆窗口示意图

D.1.2 点击nucleotide blast进入如图所示画面。

Standard Nucleotide BLAST

Enter Query Sequence

Enter accession number(s), gi(s), or FASTA sequence(s) [Clear](#) Query subrange From To

Or, upload file 未选择任何文件

Job Title

Enter a descriptive title for your BLAST search

Align two or more sequences

Choose Search Set

Database Human genomic + transcript Mouse genomic + transcript Others (nr etc.):
Nucleotide collection (nr/nt)

Organism Optional Exclude
Enter organism common name, binomial, or tax id. Only 20 top taxa will be shown

Exclude Optional Models (XM/XP) Uncultured/environmental sample sequences

Limit to Optional Sequences from type material

Entrez Query Optional [YouTube](#) [Create custom database](#)
Enter an Entrez query to limit search

Program Selection

Optimize for Highly similar sequences (megablast)
 More dissimilar sequences (discontiguous megablast)
 Somewhat similar sequences (blastn)
Choose a BLAST algorithm

Search database Nucleotide collection (nr/nt) using Megablast (Optimize for highly similar sequences)
 Show results in a new window

[Algorithm parameters](#)

图D.2 序列输入窗口示意图

将所需查询的序列以Fasta格式输入空格内，在“Choose Search set”栏中选择圈注的Others，在Optimize for栏中选择圈注的Highly similar sequence (Megablast)

D.1.3 点击Blast进入如图所示画面。

U.S. National Library of Medicine | NCBI National Center for Biotechnology Information | Sign in to NCBI

BLAST » blastn suite » RID-W1N9NJVF016 | Home | Recent Results | Saved Strategies | Help

BLAST Results

Job title: Nucleotide Sequence (541 letters)

RID [W1N9NJVF016](#) (Expires on 09-20 16:28 pm)

Query ID: Id|Query_7159
 Description: None
 Molecule type: nucleic acid
 Query Length: 541

Database Name: nr
 Description: Nucleotide collection (nt)
 Program: BLASTN 2.7.0+ » [Citation](#)

Other reports: » [Search Summary](#) | [Taxonomy reports](#) | [Distance tree of results](#) | [MSA viewer](#)

Graphic Summary

Distribution of the top 100 Blast Hits on 100 subject sequences

Mouse over to see the title, click to show alignments

Color key for alignment scores

<40	40-50	50-80	80-200	>=200
-----	-------	-------	--------	-------

Query

Descriptions

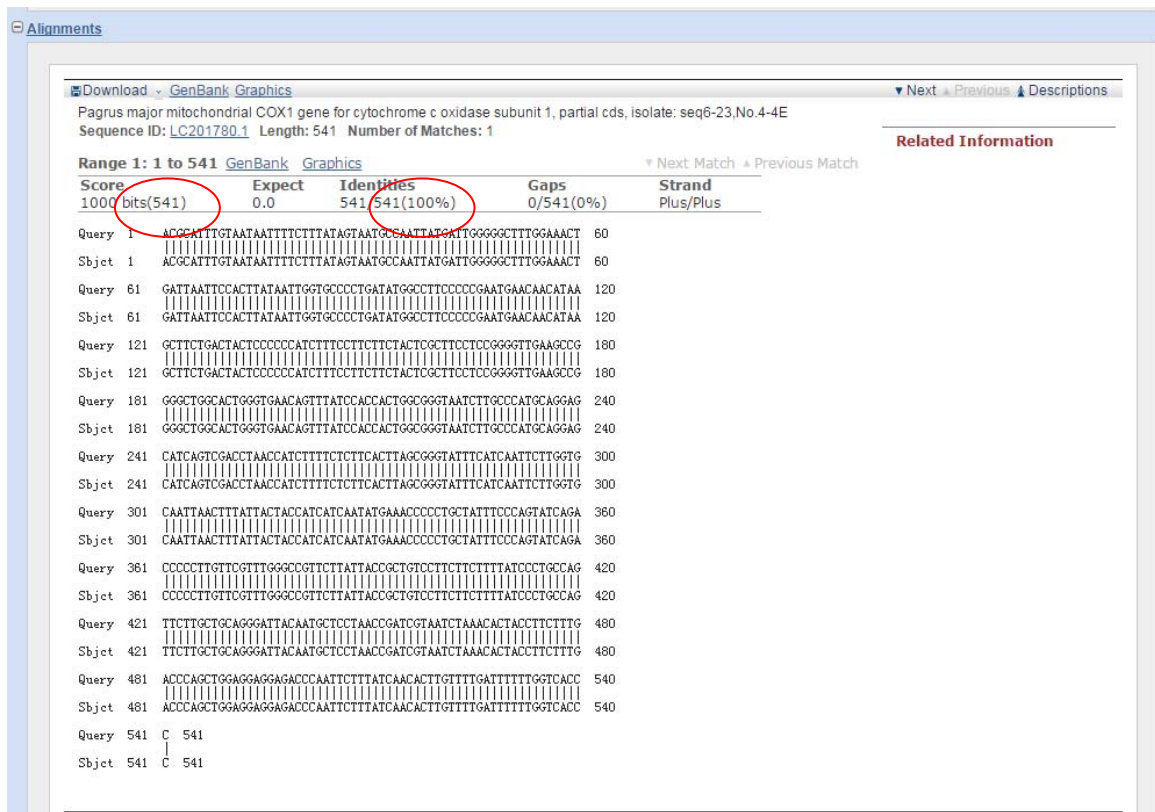
Sequences producing significant alignments:

Select: [All](#) [None](#) Selected: 0

#	Alignments	Description	Max score	Total score	Query cover	E value	Ident	Accession
1	Download GenBank Graphics Distance tree of results	Paqrus major mitochondrial COX1 gene for cytochrome c oxidase subunit 1, partial cds, isolate: seq6-23.No.4-4E	1000	1000	100%	0.0	100%	LC201780.1

图D.3 NCBI 数据库中 COI 序列的比对结果呈现形式

选择相似度最高，排列在第一圈注的点击进入下图所示画面。



图D.4 鉴定结果序列分析详情图

D.2 结果判定

第一个圈标注的序列长度大于500 bp为有效长度，第二个圈标注的序列相似度 $>98.0\%$ 者，可以直接根据相似度最高的结论判断种类。

若相似度 $>98.0\%$ 时系统给出两个或以上的物种，需参考形态学特征进行综合判断。序列相似度 $\leq 98.0\%$ 者，不能进行种类的判断。