

# 南极冰鱼和伯氏肩孔南极鱼 EPO 基因 预测及生物信息学分析\*

张 驰<sup>1</sup> 姜华鹏<sup>1</sup> 王丛丛<sup>1,2</sup> 许强华<sup>1,2,3,4,5</sup>

(1. 上海海洋大学海洋科学学院 上海 201306; 2. 大洋渔业资源可持续开发省部共建教育部重点实验室 上海 201306;  
3. 农业部大洋渔业资源环境科学观测试验站 上海 201306; 4. 国家远洋渔业工程技术研究中心 上海 201306;  
5. 远洋渔业协同创新中心 上海 201306)

**摘要** 运用 PCR、TA 克隆及二代测序技术获得了独角雪冰鱼(*Chionodraco hamatus*)、伯氏肩孔南极鱼(*Trematomus bernacchii*)的 EPO 区域基因组 DNA 序列,并结合生物信息学方法进行了基因结构与保守功能域等生物学特性研究。结果表明:克隆获得的独角雪冰鱼、伯氏肩孔南极鱼 EPO 基因序列分别为 2727 bp 和 2820 bp,包括 5 个外显子和 4 个内含子结构,编码区全长皆为 558 bp,推定编码 185 个氨基酸。理化性质分析得到两者的 EPO 蛋白皆为相对稳定的酸性可溶性蛋白。其三级结构预测结果都显示出典型的 4 个反相平行的  $\alpha$  螺旋结构特征,保持了其在生物体内发挥的原有作用。多重序列比对显示,独角雪冰鱼 EPO 与伯氏肩孔南极鱼的氨基酸同源性为 95.2%,而与其它物种的同源性在 53.9%—85.5%之间。系统发育树显示,独角雪冰鱼、伯氏肩孔南极鱼的 EPO 基因与鱼类的同源基因聚类,系统进化情况同亲缘关系一致,表明 EPO 基因在进化过程中具有较强保守性。本研究为揭示 EPO 基因在南极冰鱼内的遗传特性提供理论依据,并为进一步探讨 EPO 基因在极端低温高氧环境中的鱼类适应性进化研究奠定了基础。

**关键词** 南极冰鱼; EPO 基因; 基因克隆; 适应性进化; 生物信息学

中图分类号 Q71 doi: 10.11693/hyhz20150100036

南极位于地球最南端,又称“第七大陆”,为全球气候变化敏感区。南极水域低温高氧,生存着许多鲜为人知的极地水生生物。南极所具有的独特地理及气候特征,造就了极地生物特殊的生理特征,使其能够适应南极低温高氧极地环境,在生物进化中极具研究价值。南极圈内生活的南极鱼已成为研究鱼类适应性进化和耐寒机制的优良模型。

作者在本文中以独角雪冰鱼(*Chionodraco hamatus*, CH)和伯氏肩孔南极鱼(*Trematomus bernacchii*, TB)为主要研究对象。独角雪冰鱼属于鲈形目(Perciformes)、南极鱼亚目(Notothenoioidei)、鳕冰鱼科(Channichthyidae),生活在南极水深 200—800 m 的水域中,体表无鳞片,

体内不含血红蛋白,可能通过鳃和极薄的皮肤吸收海水中的氧(Bargelloni *et al.*, 1994; Eastman *et al.*, 2004)。伯氏肩孔南极鱼为鲈形目(Perciformes)、南极鱼亚目(Notothenoioidei)、南极鱼科(Notothenoioidei),生活在南极水深 700m 以上的水域中,其生存环境与独角雪冰鱼相似且与独角雪冰鱼有较近的亲缘关系,但体内表达有正常的血红蛋白。本文以探究独角雪冰鱼的进化机制为大背景,从对血红蛋白生成有促进作用关键因子之一 EPO 入手,来探究血红蛋白丢失的潜在机制以及 EPO 在其中起的作用。

促红细胞生成素(erythropoietin, EPO)是哺乳动物调节红细胞生成的主要调控因子,主要由肾脏和

\* 国家自然科学基金重大研究计划培育项目, 91131006 号; 上海市教育发展基金会和上海市教育委员会曙光计划, 13SG51 号; 教育部科学技术研究项目, 213013A 号; 上海市教委水产学一流学科项目。张驰, 硕士研究生, E-mail: wkzhangchi@yeah.net

通讯作者: 许强华, 教授, 博士生导师, E-mail: qhxu@shou.edu.cn

收稿日期: 2015-01-29, 收修改稿日期: 2015-05-17

胚胎肝脏分泌。研究表明 EPO 为氧依赖型基因, 缺氧可以诱导 EPO 基因的转录, 从而达到调控效果(葛全兴等, 2009)。尤其是低氧诱导因子-1(hypoxia inducing factor-1, HIF-1), 在缺氧条件下能够进入核内形成异二聚体 HIF-1 复合物, 并结合在 EPO 的 3'端约 120 bp 的区域内的顺式调控元件上(桂长云等, 1999)。该顺势调控元件为(5'-RCGTG-3')保守序列构成, 是典型的低氧反应元件(hypoxia response element, HRE)(桂长云等, 1999)。EPO 作为一种糖蛋白激素, 能够刺激骨髓造血功能, 及时有效地增加红细胞的数量, 从而提高血液的携氧能力, 增加血红蛋白含量与红细胞比积(血液中红细胞百分比)(Katakura *et al.*, 2013)。高海拔哺乳动物低氧促进红细胞生成是通过提高血浆 EPO 水平来实现的, 以改善缺氧引起的生理反应; 红细胞增多又可反馈性抑制 EPO 的生成, 以使红细胞量精确地适应机体的需要。在非哺乳动物中, EPO 的保守性以及刺激肾脏中红色祖先细胞分裂和分化作用同哺乳动物是一致的(Stachura *et al.*, 2009, 2011)。在非洲爪蟾体外实验中也同样证明 EPO 系统在正常状态下刺激幼红血细胞的增殖, 并使其在急性严重贫血情况中恢复过来(Nogawa-Kosaka *et al.*, 2010)。这表明 EPO 系统在红系器官发育中具有关键作用。近年来, 体外和动物实验报道都暗示了 EPO 的作用不仅仅局限于红系细胞生成。在非造血系统中, EPO 在低氧状态下具有神经营养和神经保护作用(Kulkarni *et al.*, 2010; Ostrowski *et al.*, 2011)。

值得一提的是, 鱼类 EPO 功能的研究尚处于起步阶段。人们先后在河豚、斑马鱼中发现了 EPO 基因的结构、HRE(低氧反应元件)所在位点, 对 EPO 在鱼类缺氧应答机制中所起的作用有了初步认识(Chou *et al.*, 2004; Chu *et al.*, 2007; Paffett-Lugassy *et al.*, 2007)。对鱼类 EPO 结构和功能的研究仍将继续。本研究首次获得了独角雪冰鱼、伯氏肩孔南极鱼的 EPO 基因序列, 对其分子进化水平进行了初步探索, 以期在 EPO 对鱼类的功能研究方面做出贡献, 为极端环境下的生物长期适应机制提供帮助。本研究为揭示 EPO 基因在南极冰鱼内的遗传特性提供理论依据, 并为进一步探讨 EPO 基因在极端低温高氧环境中的鱼类适应性进化研究奠定了基础。

## 1 材料与方法

### 1.1 材料

1.1.1 实验动物 独角雪冰鱼和伯氏肩孔南极鱼样本分别采集于南极埃默里冰架和南极熊猫码头。利

用冰凿打冰洞, 再进行渔钓的方式, 获得活体南极鱼样本数十条。活体迅速冻存于 $-80^{\circ}\text{C}$ 超低温冰箱。由雪龙号带回的冻存样本用锡箔纸包裹后液氮带回上海海洋大学海洋科学学院保护遗传学实验室,  $-80^{\circ}\text{C}$ 保存备用。

1.1.2 主要试剂 Taq DNA Polymerase (艾德莱); Agarose N、氯仿、异丙醇、 $5\times\text{TAE}$  缓冲液、琼脂糖、氯化钠、酵母提取物(上海生工);  $6\times\text{loading buffer}$ 、 $0.1\%$  DEPC 及  $75\%$ 乙醇(天根); X-gal、IPTG (Amersco); DNA 胶回收试剂盒(Axygen Biosciences); TRIzol<sup>®</sup>试剂 (Invitrogen Corp); 胰蛋白酶 (OXOID); pMD19-T 载体、DH5 $\alpha$  感受态细胞及 DL5000 DNA 标记(TaKaRa); Trans 2K Plus DNA 标记(TransGen)。

### 1.2 方法

1.2.1 南极鱼基因组 DNA(gDNA)提取及 PCR 扩增 取独角雪冰鱼和伯氏肩孔南极鱼的肌肉组织, 参照 Trizol 法(Invitrogen Corp, USA)分别提取其基因组 DNA。用 NanoDrop<sup>®</sup> 2000 (Thermo)与  $1.5\%$ 琼脂糖凝胶电泳检测提取质量。根据 GenBank 数据库中 EPO 相应的同源序列(表 1)(Katakura *et al.*, 2013), 利用 Primer Premier 5 软件设计特异性引物, 引物由上海生工合成(表 2)。

表 1 EPO 相应的同源序列信息  
Tab.1 The information of the respective counterparts of EPO homologous sequences

物种	登录号
斜带石斑鱼 <i>Epinephelus coioides</i>	AAW29029.1
红鳍东方鲀 <i>Takifugu rubripes</i>	AAQ72467.1
罗非鱼 <i>Oreochromis niloticus</i>	XP_003457688.1
斑马鱼 <i>Danio rerio</i>	NP_001108600.1
鲫鱼 <i>Carassius auratus</i>	AGH20610.1
鲤鱼 <i>Cyprinus carpio</i>	ABB83930.1
清斑河豚 <i>Tetraodon nigroviridis</i>	AAR25698.1
青鳉 <i>Oryzias latipes</i>	XP_004079700.1
热带爪蟾 <i>Xenopus tropicalis</i>	NP_001233194.1
非洲爪蟾 <i>Xenopus laevis</i>	BAI82351.1
马 <i>Equus caballus</i>	NP_001075294.1
猫 <i>Felis catus</i>	NP_001009269.1
狗 <i>Canis lupus</i>	NP_001006647.1
猪 <i>Sus scrofa</i>	NP_999299.1
兔 <i>Oryctolagus cuniculus</i>	NP_001075559.1
牛 <i>Bos taurus</i>	NP_776334.1
人 <i>Homo sapiens</i>	AAI43226.1
猕猴 <i>Macaca mulatta</i>	NP_001036201.1
黑猩猩 <i>Pan troglodytes</i>	XP_519268.2
大鼠 <i>Rattus norvegicus</i>	NP_058697.1
高原鼯鼠 <i>Eospalax baileyi</i>	EU183314.1
小鼠 <i>Mus musculus</i>	AAK28825.1

表 2 EPO 基因 PCR 引物信息  
Tab.2 The information of PCR primers of EPO gene

引物名称	引物序列(5'—3')	扩增长度(bp)
EPO-1-F	ATGTTGCAGAAAACGGGTAGA	1164—1193
EPO-1-R	CCCTCAGGTCACAGATTGGC	
EPO-2-F	GACTGCTCACCTTGCTGTTGT	623—727
EPO-2-R	CTGGATGTGGAGGCTGTGCA	
EPO-3-F	GCACAGCCACTTAGATAACTCCA	891—977
EPO-3-R	GCCCCGAGGAAGTTGA	
EPO-4-F	GAAGGAACATGGAGGTTGCTCGG	120
EPO-4-R	TCAGCTGACGTCTTGCTGGC	

PCR 总反应体系为 25  $\mu$ L, 含 0.5  $\mu$ L 10  $\mu$ mol/L 正向和反向引物、2.5  $\mu$ L 10 $\times$ Buffer (含  $MgCl_2$ )、2  $\mu$ L dNTP Mixture (各 2.5 mmol/L)、0.5  $\mu$ L Taq DNA 聚合酶(5U/ $\mu$ L)。反应程序为 94 $^{\circ}C$  预变性 5 min; 94 $^{\circ}C$  变性 30 s, 58—64 $^{\circ}C$  退火 30 s (退火温度依引物而异), 72 $^{\circ}C$  延伸 1—2 min (延伸温度依扩增片段长短而异), 32 个循环; 72 $^{\circ}C$  终延伸 10 min。PCR 产物用 1% 的琼脂糖凝胶电泳检测。

**1.2.2 PCR 扩增产物的克隆、转化与测序** 琼脂糖凝胶回收试剂盒回收目的片段, 与 pMD19-T 克隆载体于 16 $^{\circ}C$  下连接过夜; 以连接产物转化大肠杆菌 DH5 $\alpha$  菌株感受态细胞, 蓝白斑筛选阳性克隆。菌落 PCR (以阳性菌落为模板)鉴定重组转化菌(莎姆布鲁克, 2002; 陈书霞等, 2006)。经菌落 PCR 鉴定为阳性克隆的菌落重新挑入 500  $\mu$ L 含 Amp 的 LB 培养基中 37 $^{\circ}C$  210 r/min 培养 14 h 后, 送上海生工进行双向测序, 测序引物为载体通用引物。

**1.2.3 生物信息学分析** 使用 Seqman、Editseq 等软件进行测序序列片段拼接; 通过 NCBI blast two sequences 完成 mRNA 序列与基因组 DNA 比对, 利用相关软件处理结果, 并仔细核对序列信息; EPO 基因开放阅读框(ORF)利用 NCBI 的 ORF Finder 程序分析(罗轶, 2008); 使用 Smart 在线分析网站(<http://smart.embl-heidelberg.de/>)以及 NCBI 中的 CDD 库 (Conserved Domain Database)进行结构域的保守性分析(Marchler-Bauer *et al*, 2005); 使用 Scanprosite 在线软件 (<http://prosite.expasy.org/cgi-bin/prosite/>)分析其功能位点差异; 使用 SignalP 4.1 server 在线处理网站 (<http://www.cbs.dtu.dk/services/SignalP/>)分析蛋白潜在信号肽; 利用 ProtParam (<http://web.expasy.org/protparam/>)在线分析氨基酸序列的理化特性(Wilkins

*et al*, 1999); 利用 ProtScale 网站(<http://web.expasy.org/protscale/>)分析蛋白质亲疏水性; 使用 PredictProtein 网站(<http://www.predictprotein.org>)和 Psipred 分析网站 (<http://bioinf.cs.ucl.ac.uk/psipred/>)预测蛋白质二级结构(李盛杰等, 2013); 根据 PSD 数据库和 SWISS - MODEL 在线建模网站(<http://swissmodel.expasy.org/>)进行同源建模, 构建蛋白质的 3D 模型(陶士珩, 2007; 刘虎岐等, 2010); 利用 PyMOL 和 Spdbv 软件修饰并输出其 3D 结构(李盛杰等, 2013)。最后使用 DNAMAN、MEGA 软件进行同源性分析并构建其系统进化树(季舒涵等, 2010; 张宝等, 2013)。

## 2 结果与分析

### 2.1 南极鱼 EPO 基因结构及序列比较

通过 PCR 分段扩增, 分别获得了独角雪冰鱼和伯氏肩孔南极鱼 DNA 片段(图 1)。阳性克隆菌经过测序后获得测序结果, 利用 Editseq 和 Seqman 将测序结果进行序列拼接后, 获得了该基因扩增的核心区及侧翼区部分序列。其中, 得到独角雪冰鱼基因全长 2727 bp, 基因编码区 558 bp, 而伯氏肩孔南极鱼基因全长 2820 bp, 基因编码区 558 bp。为保证分析序列结果的可靠性, 拼装 EPO 序列与红旗东方鲀、罗非鱼等同源性高的相关物种进行序列特征性比对, 结果证明其拼接序列位于基因组 EPO 区域。

对 EPO 基因的结构分析表明, 独角雪冰鱼和伯氏肩孔南极鱼的 EPO 皆由 5 个外显子组成, 中间穿插有 4 个内含子。其序列差异比对结果表明, 这两个基因 gDNA 序列相似性为 80.1%, 其主要的序列差异位于内含子非编码区, 共有 6 个 Gap, 平均 Gap 长度为 19 bp, 最大一致序列长度为 640 bp。独角雪冰鱼同红旗东方鲀序列相似性仅 48.9%, 表明 EPO 基因的种属差异性(图 2)。

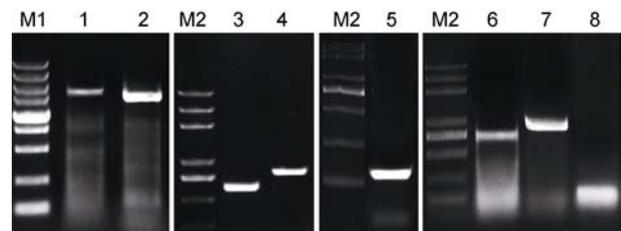


图 1 南极鱼扩增产物电泳图

Fig.1 Electrophoresis of Antarctic ice fish EPO amplified fragments  
M1: DL5000 DNA marker; M2: Trans 2K Plus DNA marker; 独角雪冰鱼 EPO 扩增片段 1, 2, 3, 4 (1, 3, 4, 5); 伯氏肩孔南极鱼 EPO 扩增片段 1, 2, 3, 4 (2, 6, 7, 8)

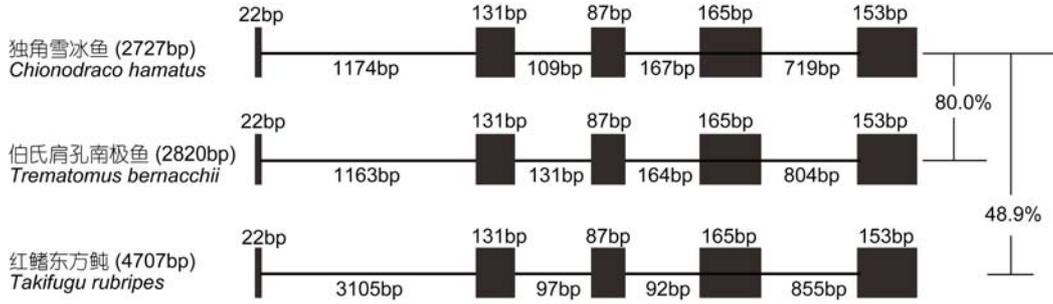


图 2 南极鱼基因结构示意图

Fig.2 Diagram of EPO genetic structures of Antarctic icefish

黑色方框表示为外显子区域, 直线表示为内含子区域。红旗东方鲀 EPO 基因为外源参考基因

### 2.2 南极鱼 EPO 基因 CDS 区核苷酸序列及编码氨基酸序列分析

为了探讨 EPO 基因的功能域在独角雪冰鱼、伯氏肩孔南极鱼中的差异, 对克隆的基因序列进行对比(图 3)。两者之间的 ORF 序列之间同源性为 95.2%, 表明南极鱼中的相对保守性。EPO 蛋白前体为一种跨膜糖蛋白分子。根据 SignalP 4.1 server、SMART 软件和 NCBI 检测得到两者的 EPO 都包含一个信号肽区和一个 EPO\_TPO 保守结构域(图 4)。独角雪冰

鱼和伯氏肩孔南极鱼 EPO 基因 ORF 推导氨基酸序列之间存在 16 处位点差异, 其中 EPO\_TPO 区域有 15 个氨基酸的差异。经过 Scanprosite 的分析, 结果表明伯氏肩孔南极鱼有两个苏氨酸(Thr)和一个丝氨酸(Ser)可以成为蛋白激酶 C 磷酸化位点, 有一个苏氨酸和三个丝氨酸可以成为酪蛋白激酶 II 的磷酸化位点; 而相比较而言, 独角雪冰鱼分别增加了一个丝氨酸位点可以被蛋白酶 C 和酪蛋白激酶 II 磷酸化(图 4)。



图 3 南极鱼核苷酸序列间比对

Fig.3 Alignment of nucleotide sequences of Antarctic ice fish EPO

灰色部分表示物种间氨基酸核苷酸序列相同; 白色表示物种间核苷酸序列不同

### 2.3 南极鱼 EPO 蛋白质结构分析

**2.3.1 EPO 蛋白理化性质分析** 用 ProtParam 软件预测南极鱼 EPO 的理化性质, 结果显示独角雪冰鱼和伯氏肩孔南极鱼皆由 185 个氨基酸组成, 其分子式分别为  $C_{904}H_{1472}N_{262}O_{271}S_8$  和  $C_{907}H_{1469}N_{257}O_{267}S_9$ , 相应分子量分别为 20.60 kDa 和 20.53 kDa, 理论等电点(pI)分别为 6.65 和 6.28, 蛋白呈酸性。独角雪冰鱼 EPO 氨基酸残基中 Leu (16.8%)、Ser (9.2%)及 Val (8.1%) 的频率较高, 疏水性氨基酸占 43.8%, 极性不带电荷氨基酸占 32.4%, 极性带电荷氨基酸占 23.8%, 其中包括 10.3%的酸性氨基酸(D, E)及 13.5%的碱性氨基酸(H, R, K); 伯氏肩孔南极鱼中 Leu (18.4%)、Ser

(8.1%)、Ala (7.0%)的频率较高, 疏水性氨基酸占 45.4%, 极性不带电荷氨基酸占 30.8%, 极性带电荷氨基酸占 23.8%, 其中包括 10.3%的酸性氨基酸及 13.5%的碱性氨基酸(表 3)。独角雪冰鱼和伯氏肩孔南极鱼 EPO 蛋白消光系数( $M^{-1} cm^{-1} \gamma=280nm$ )同样为 23740, 其不稳定系数分别为 36.00 和 36.87(通常当不稳定系数小于 40 时, 可将蛋白质确定为稳定类蛋白质), 皆属于稳定类蛋白质(Guruprasad *et al*, 1990), 且预计在体外哺乳动物网织红细胞内的半衰期为 30 h。独角雪冰鱼疏水指数为 106.43, 伯氏肩孔南极鱼疏水指数为 110.6, 平均亲水性分别为 -0.071 和 0.022, 均属可溶性蛋白范围(Kyte *et al*, 1982)。

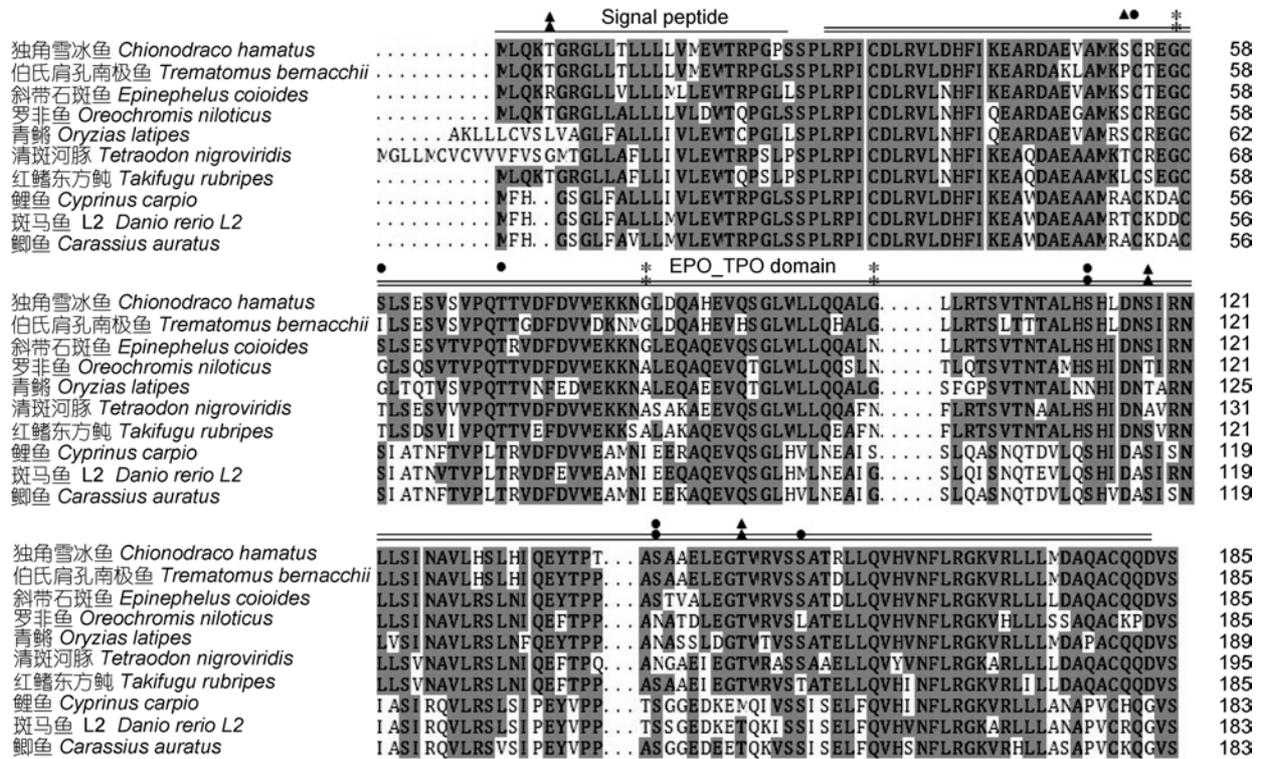


图 4 EPO 氨基酸序列多重比对及保守结构域和功能位点分析

Fig. 4 Alignment of amino acid sequences and conservative domain structure and functional sites of Antarctic icefish EPO. 灰色部分表示有一半以上物种间氨基酸序列相同; 白色表示有一个以上物种间氨基酸序列的不同。单划线表示南极冰鱼信号肽(第 1—25 个氨基酸残基), 双划线表示南极冰鱼 EPO\_TPO 结构域(第 28—183 个氨基酸残基), 黑三角标注蛋白激酶 C 磷酸化位点, 黑点标注酪蛋白激酶 II 磷酸化位点, 星号标注酰基化位点

表 3 南极鱼 EPO 的氨基酸组成  
Tab.3 The composition of amino acids in EPO

独角雪冰鱼 <i>Chionodraco hamatus</i>			伯氏肩孔南极鱼 <i>Trematomus bernacchii</i>		
氨基酸	数量	频率(%)	氨基酸	数量	频率(%)
Ala (A)	13	7.0	Ala (A)	13	7.0
Arg (R)	12	6.5	Arg (R)	10	5.4
Asn (N)	6	3.2	Asn (N)	5	2.7
Asp (D)	9	4.9	Asp (D)	11	5.9
Cys (C)	4	2.2	Cys (C)	4	2.2
Gln (Q)	11	5.9	Gln (Q)	9	4.9
Glu (E)	10	5.4	Glu (E)	8	4.3
Gly (G)	9	4.9	Gly (G)	10	5.4
His (H)	7	3.8	His (H)	9	4.9
Ile (I)	5	2.7	Ile (I)	6	3.2
Leu (L)	31	16.8	Leu (L)	34	18.4
Lys (K)	6	3.2	Lys (K)	6	3.2
Met (M)	4	2.2	Met (M)	5	2.7
Phe (F)	3	1.6	Phe (F)	3	1.6
Pro (P)	6	3.2	Pro (P)	7	3.8
Ser (S)	17	9.2	Ser (S)	15	8.1
Thr (T)	12	6.5	Thr (T)	13	7.0
Trp (W)	4	2.2	Trp (W)	4	2.2
Tyr (Y)	1	0.5	Tyr (Y)	1	0.5
Val (V)	15	8.1	Val (V)	12	6.5

**2.3.2 南极鱼 EPO 二级、三级结构预测** 使用 PredictProtein 和 PSIPRED 软件预测南极鱼的二级结构, 结果表明: 独角雪冰鱼和伯氏肩孔南极鱼 EPO\_TPO 域都主要由 4 个反向平行的  $\alpha$  螺旋和两长两短的环状结构(Loops)区域构成, 为 EPO 蛋白的典型结构(Boissel *et al.*, 1993), 仅伯氏肩孔南极鱼预测出置信度较低的两个小的  $\beta$  折叠区域(图 5)。

利用 SWISS-MODEL 同源建模网站和 PyMol 软件对南极冰鱼 EPO 建模获得三级结构模型, 该结果表明独角雪冰鱼和伯氏肩孔南极鱼都呈 4 个反向平行的  $\alpha$  螺旋结构, 由  $\alpha$  螺旋与无规卷曲构成。即使同源构建模型所用模板不同, 其主要的 4 个螺旋结构也是保守的, 这与二级结构预测结果基本一致(图 6)。

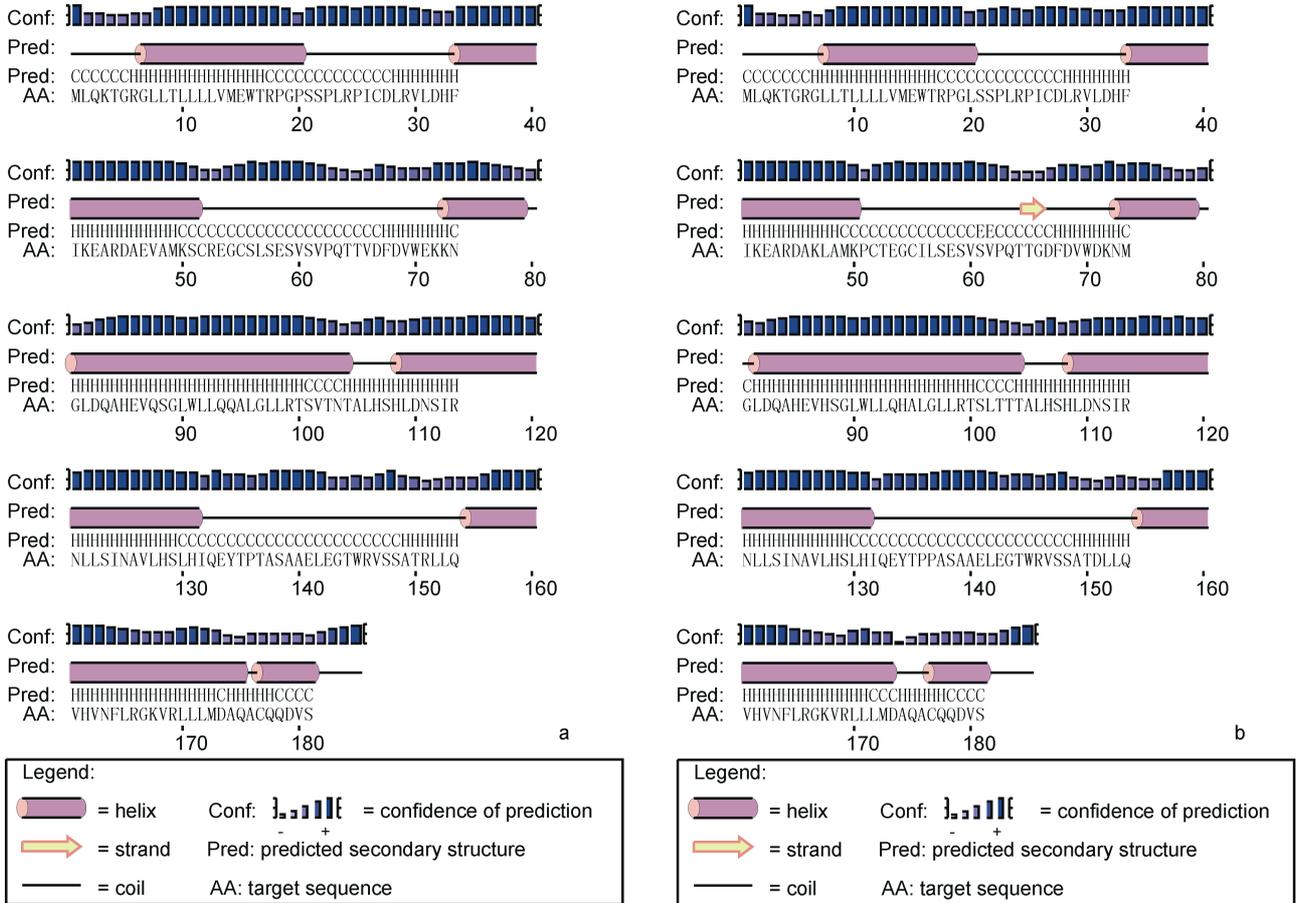


图 5 南极鱼 EPO 的二级结构预测

Fig.5 Predictions of the secondary structure of the Antarctic icefish EPO  
a. 独角雪冰鱼二级结构预测结果; b. 伯氏肩孔南极鱼二级结构预测结果

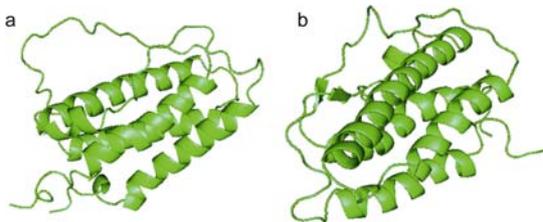


图 6 南极鱼 EPO 的三级结构预测

Fig.6 Predictions of the tertiary structure of the Antarctic icefish EPO

a. 独角雪冰鱼 EPO 同源建模结果(以 PDB 数据库中 1eer.1.A 为模板); b. 伯氏肩孔南极鱼 EPO 同源建模结果(以 PDB 数据库中 1cn4.1.C 为模板)

**2.4 南极鱼 EPO 同源性及系统发育分析**

将独角雪冰鱼和伯氏肩孔南极鱼 EPO 基因的氨基酸序列与其它物种的相应序列进行同源性分析(图 7)。其中, 独角雪冰鱼 EPO 与其它鱼类的同源性在 70%—90%之间, 与两栖类的同源性约 85.5%, 与哺乳动物同源性在 53.9%—85.5%之间。数据表明, EPO 种内差异较小, 同源性总体较高, 进化比较保守; 但 EPO 基因的种间差异比较明显。用 MEGA5 软件分析南极冰鱼及其它鱼类、两栖类、哺乳类 EPO 基因编码区氨基酸系统发生树, 其中独角雪冰鱼和伯

氏肩孔南极鱼与其它鱼类的遗传关系很近, 与两栖类、哺乳类的遗传距离较远(图 8)。

### 3 讨论

南极作为地球上特殊的极端环境之一, 其特殊

物种南极冰鱼在长期进化选择过程中所占据的重要位置值得去做深入研究。EPO 是与低氧适应、组织保护和血系生成相关的重要基因, 但目前还未获得其在极端环境下尤其是低温下的表达调控的机制和功能。尤其是南极冰鱼的特殊存在以及其基因组研究价

P-distance

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27		
1. Chionodraco hamatus		0.03	0.03	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04		
2. Trematomus bernacchi	0.11		0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04		
3. Epinephelus coioides	0.13	0.19		0.04	0.04	0.03	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.05	0.05	0.05	0.05	0.04	0.04	0.04	0.05		
4. Takifugu rubripes	0.26	0.29	0.20		0.03	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.05	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04		
5. Tetraodon nigroviridis	0.23	0.29	0.19	0.10		0.04	0.04	0.04	0.04	0.04	0.04	0.05	0.04	0.04	0.04	0.05	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04		
6. Oreochromis niloticus	0.21	0.27	0.16	0.22	0.22		0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.05	0.05	0.04	0.04	0.04	0.04	0.04		
7. Oryzias latipes	0.28	0.34	0.24	0.31	0.29	0.22		0.04	0.04	0.04	0.04	0.05	0.04	0.04	0.04	0.05	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04		
8. Danio rerio isoform L1	0.37	0.40	0.32	0.40	0.40	0.39	0.40		0.00	0.00	0.02	0.02	0.02	0.02	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04		
9. Danio rerio isoform S	0.37	0.40	0.32	0.40	0.40	0.39	0.40	0.00		0.00	0.02	0.02	0.02	0.02	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04		
10. Danio rerio isoform L2	0.37	0.40	0.32	0.40	0.40	0.39	0.40	0.00	0.00		0.02	0.02	0.02	0.02	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04		
11. Carassius auratus	0.39	0.43	0.35	0.42	0.43	0.42	0.45	0.08	0.08	0.08		0.02	0.02	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04		
12. Cyprinus carpio II	0.40	0.44	0.34	0.43	0.44	0.41	0.45	0.08	0.08	0.08	0.06		0.02	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04		
13. Cyprinus carpio	0.39	0.43	0.34	0.40	0.41	0.40	0.44	0.08	0.08	0.08	0.04	0.04		0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04		
14. Xenopus laevis	0.66	0.68	0.66	0.70	0.68	0.70	0.68	0.71	0.71	0.71	0.71	0.72	0.71		0.03	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04		
15. Xenopus tropicalis	0.66	0.68	0.67	0.71	0.70	0.70	0.68	0.71	0.71	0.71	0.71	0.72	0.71	0.11		0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04		
16. Oryctolagus cuniculus	0.61	0.62	0.59	0.59	0.58	0.61	0.60	0.53	0.63	0.63	0.62	0.64	0.62	0.64	0.62	0.64		0.03	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.03	0.04	
17. Homo sapiens	0.63	0.65	0.61	0.63	0.62	0.64	0.63	0.63	0.63	0.63	0.63	0.64	0.63	0.61	0.63	0.18	0.01		0.03	0.04	0.04	0.03	0.03	0.03	0.03	0.03	0.03		
18. Pan troglodytes	0.63	0.65	0.61	0.63	0.62	0.64	0.63	0.63	0.63	0.63	0.63	0.64	0.63	0.61	0.63	0.18	0.01	0.03		0.04	0.04	0.03	0.03	0.03	0.03	0.03	0.03		
19. Macaca mulatta	0.62	0.66	0.61	0.65	0.65	0.65	0.64	0.64	0.64	0.64	0.65	0.65	0.64	0.62	0.64	0.21	0.10	0.11	0.03		0.04	0.04	0.03	0.04	0.04	0.04	0.03	0.04	
20. Mus musculus	0.61	0.62	0.57	0.62	0.61	0.61	0.61	0.61	0.61	0.61	0.61	0.62	0.61	0.61	0.63	0.64	0.22	0.18	0.18	0.20		0.01	0.03	0.03	0.03	0.03	0.03	0.03	
21. Rattus norvegicus	0.62	0.63	0.58	0.62	0.61	0.61	0.61	0.61	0.61	0.61	0.62	0.61	0.61	0.63	0.64	0.23	0.19	0.19	0.21	0.03		0.03	0.03	0.04	0.03	0.03	0.03	0.03	
22. Sus scrofa	0.62	0.63	0.58	0.61	0.61	0.62	0.62	0.61	0.61	0.61	0.61	0.61	0.61	0.61	0.61	0.59	0.62	0.18	0.12	0.12	0.16	0.15	0.17		0.03	0.03	0.03	0.03	
23. Eospalax baileyi	0.62	0.63	0.58	0.63	0.62	0.61	0.62	0.61	0.61	0.61	0.60	0.61	0.60	0.61	0.60	0.61	0.63	0.21	0.18	0.18	0.21	0.10	0.11	0.13		0.03	0.03	0.03	
24. Bos taurus	0.61	0.62	0.58	0.62	0.61	0.62	0.61	0.62	0.62	0.62	0.61	0.62	0.61	0.58	0.61	0.19	0.14	0.14	0.19	0.18	0.19	0.09	0.17		0.03	0.03	0.03	0.03	
25. Equus caballus	0.61	0.62	0.58	0.62	0.61	0.62	0.61	0.61	0.61	0.61	0.62	0.63	0.62	0.61	0.63	0.18	0.15	0.15	0.19	0.17	0.18	0.12	0.15	0.10		0.03	0.03	0.03	
26. Felis catus	0.61	0.63	0.58	0.62	0.61	0.62	0.62	0.62	0.62	0.62	0.62	0.63	0.62	0.60	0.62	0.17	0.13	0.14	0.18	0.16	0.16	0.10	0.15	0.12	0.08		0.02	0.02	
27. Canis lupus familiaris	0.62	0.64	0.59	0.63	0.62	0.62	0.63	0.62	0.62	0.62	0.61	0.62	0.61	0.59	0.61	0.19	0.16	0.17	0.20	0.17	0.18	0.12	0.18	0.11	0.11	0.05		0.02	0.02

图 7 各物种 EPO 基因推定的氨基酸序列间的遗传距离比较  
Fig.7 The sequence identity of inferred EPO amino acid sequences

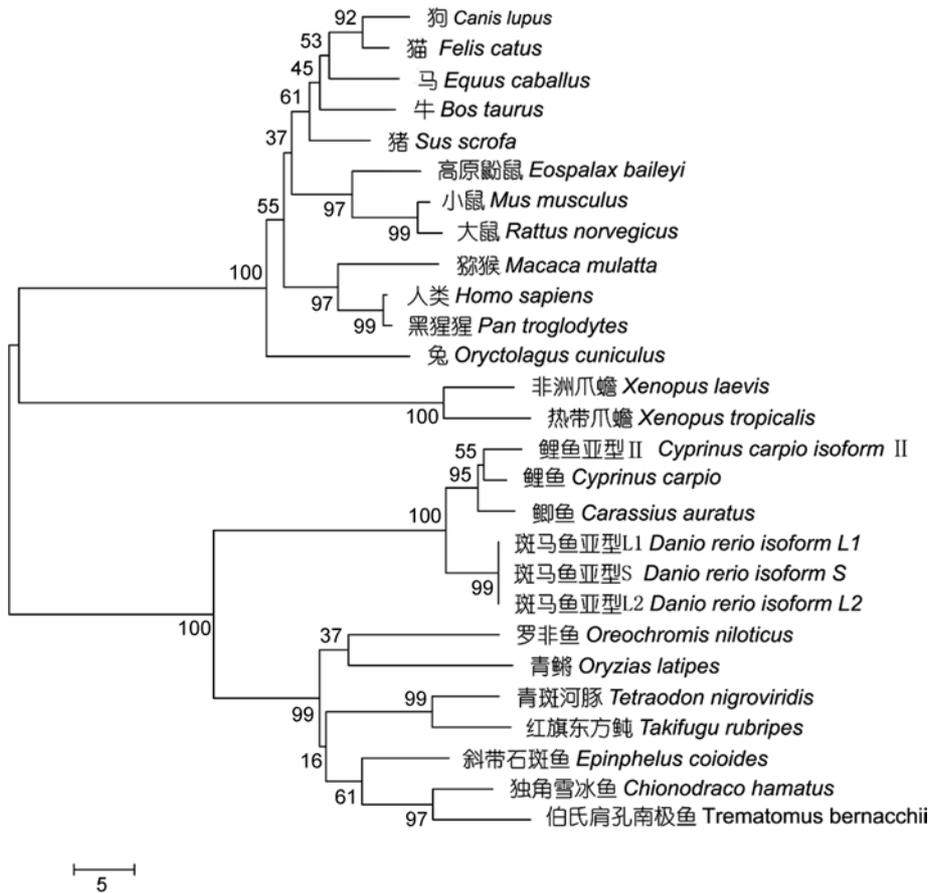


图 8 利用 MEGA 软件中 N-J 法构建的 EPO 系统发育树  
Fig. 8 The Neighbor-Joining tree of EPO constructed with MEGA5 software

值,为人们提供了一个优秀的耐寒基因库(许强华等, 2014)。南极鱼类通过三千万年的快速进化中,在南大洋逐渐变冷直至冰冻的期间演化成了南大洋最多样化的鱼种。作者希望通过比较基因组学、生态学、生理学以及生物统计学等相关专业知识,对 EPO 基因获得更深层的认识,推测它在生物适应性进化过程中所起的作用。尤其是在全球变暖的生态大背景下,南极冰鱼如何适应正在慢慢变化的南极低温环境。目前,针对南极水生动物 EPO 的研究成果在国内外鲜有发布,而对于南极特有的水生动物之一南极冰鱼的研究更是少之甚少。

对独角雪冰鱼和伯氏肩孔南极鱼 EPO 基因结构进行分析,长度分别为 2727 bp 和 2820 bp。南极鱼 EPO 基因结构在鱼类中具有一致性,都包含 5 个外显子和 4 个内含子,种属内部同源性较高,差异性主要体现在非编码区。而针对编码区的研究发现,南极鱼 EPO 推导氨基酸数目与其它鱼类基本一致,但与两栖类、哺乳动物类差异性较大。独角雪冰鱼与伯氏肩孔南极鱼相比,尽管存在有 26 个碱基差异,其中 10 个碱基突变均属同义突变,所编码的氨基酸序列同源率为 91.35%。说明独角雪冰鱼 EPO 与伯氏肩孔南极鱼 EPO 一致,两者之间有生物学功能上的一致性。基于氨基酸序列构建的系统发育树中,独角雪冰鱼和伯氏肩孔南极鱼 EPO 聚类,亲缘关系极近,与鱼类的亲缘关系也在 91.35%—70%之间,符合进化规律。这说明 EPO 在生物长期进化过程中保持着较强的保守性。

在蛋白性质的研究中,蛋白质的稳定性的探究一直都与其半衰期长短相联系。研究表明,半衰期长则蛋白质稳定性高,反之半衰期短则表明蛋白质稳定性低(贾浩等, 2011)。在本研究中,独角雪冰鱼和伯氏肩孔南极鱼的 EPO 都具有较长的半衰期,且属于稳定性蛋白,将其与 EPO 的稳定性分析结果结合起来看,能够确保独角雪冰鱼和伯氏肩孔南极鱼都属于稳定型亲水性蛋白,这也就表明了 EPO 蛋白在体内的生理功能发挥将会维持较稳定的状态。根据以往的研究表明,成熟 EPO 的某些螺旋区域能够影响 EPO 的生物活性,其主要原理是能够增加蛋白质三级结构的灵活度(Boissel *et al*, 1993)。而伯氏肩孔南极鱼在二级结构预测中的两个小  $\beta$  折叠区域能够在确保不会影响 EPO 的功能结构特点的基础上,潜在增加 EPO 的生物活性。两者的三级结构分析结果表明南极冰鱼的 EPO 同哺乳动物研究结果一致,都是

一个典型的 4 个反向平行的  $\alpha$  螺旋折叠结构构成的球形分子(桂长云等, 1999; 葛全兴等, 2009)。虽然独角雪冰鱼和伯氏肩孔南极鱼的 EPO 的理化性质分析结果以及二、三级结构预测都能够确保 EPO 能够正常行使其功能,但是两者之间由碱基突变造成的功能位点的改变对于该蛋白与受体的结合效率、转运、以及参与能量代谢过程的影响仍是一个有待深入研究的过程。该基因的成功克隆及分析为揭示南极冰鱼 EPO 基因的遗传特性以及生物在极端寒冷环境中的长期进化适应提供了理论依据。

### 参 考 文 献

- 刘虎岐, 刘应保, 宋云鹏等, 2010. 淡色库蚊氯菊酯抗性相关基因 PR-OP 全长 cDNA 的克隆及生物信息学分析. 西北农林科技大学学报(自然科学版), 38(9): 109—117
- 许强华, 吴智超, 陈良标, 2014. 南极鱼类多样性和适应性进化研究进展. 生物多样性, 22(1): 80—87
- 李盛杰, 杜晓华, 罗玉柱等, 2013. 天祝白牦牛 NGB 基因的克隆及生物信息学分析. 畜牧兽医学报, 44(3): 395—398
- 张 宝, 黄克勇, 郭劲松等, 2013. H7N9 病毒的来源和重组模式. 南方医科大学学报, 33(7): 1017—1021
- 陈书霞, 王晓武, 房玉林, 2006. 单菌落 PCR 法直接快速鉴定重组克隆. 微生物学通报, 33(3): 52—56
- 罗 轶, 2008. 鸡 FATP1 基因 cDNA 的克隆、组织表达及其生物信息学分析. 雅安: 四川农业大学硕士学位论文, 22—23
- 季舒涵, 管林森, 王洪宝等, 2010. 秦川牛 A-FABP 基因的生物信息学分析. 西北农林科技大学学报(自然科学版), 38(6): 77—81
- 桂长云, 钱若兰, 1999. EPO, EPOR 及其信号传递研究进展. 细胞生物学杂志, 21(4): 157—163
- 贾 浩, 张小白, 宋晓峰, 2011. 人类胞内蛋白半衰期与其亚细胞定位的相关性研究. 计算机与应用化学, 28(4): 411—414
- 陶士珩, 2007. 生物信息学. 北京: 科学出版社, 177—179
- 葛全兴, 陈 垦, 龙友明等, 2009. EPO/ EPOR 及其信号转导机制研究进展. 国际病理科学与临床杂志, 29(5): 416—420
- 莎姆布鲁克 J, 拉塞尔 D W 编著, 黄培堂译, 2002. 分子克隆实验指南. 北京: 科学出版社, 103—105
- Bargelloni L, Ritchie P A, Patarnello T *et al*, 1994. Molecular evolution at subzero temperatures: mitochondrial and nuclear phylogenies of fishes from Antarctica (suborder Notothenioidei), and the evolution of antifreeze glycopeptides. *Molecular Biology and Evolution*, 11(6): 854—863
- Boissel J P, Lee W R, Presnell S R *et al*, 1993. Erythropoietin structure-function relationships. Mutant proteins that test a model of tertiary structure. *Journal of Biological Chemistry*, 268(21): 15983—15993

- Chou C F, Tohari S, Brenner S *et al*, 2004. Erythropoietin gene from a teleost fish, *Fugu rubripes*. *Blood*, 104(5): 1498—1503
- Chu C Y, Cheng C H, Chen G D *et al*, 2007. The zebrafish erythropoietin: functional identification and biochemical characterization. *FEBS Letters*, 581(22): 4265—4271
- Eastman J T, Lannoo M J, 2004. Brain and sense organ anatomy and histology in hemoglobinless Antarctic icefishes (Perciformes: Notothenioidei: Channichthyidae). *Journal of Morphology*, 260(1): 117—140
- Guruprasad K, Reddy B V B, Pandit M W, 1990. Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting *in vivo* stability of a protein from its primary sequence. *Protein Engineering*, 4(2): 155—161
- Katakura F, Katzenback B A, Belosevic M, 2013. Molecular and functional characterization of erythropoietin receptor of the goldfish (*Carassius auratus* L.). *Developmental & Comparative Immunology*, 45(2): 191—1598
- Kulkarni R P, Tohari S, Ho A *et al*, 2010. Characterization of a hypoxia-response element in the Epo locus of the pufferfish, *Takifugu rubripes*. *Marine Genomics*, 3(2): 63—70
- Kyte J, Doolittle R F, 1982. A simple method for displaying the hydrophobic character of a protein. *Journal of Molecular Biology*, 157(1): 105—132
- Marchler-Bauer A, Anderson J B, Cherukuri P F *et al*, 2005. CDD: A conserved domain database for protein classification. *Nucleic Acids Research*, 33(suppl1): D192—D196
- Nogawa-Kosaka N, Hirose T, Kosaka N *et al*, 2010. Structural and biological properties of erythropoietin in *Xenopus laevis*. *Experimental Hematology*, 38(5): 363—372
- Ostrowski D, Ehrenreich H, Heinrich R, 2011. Erythropoietin promotes survival and regeneration of insect neurons *in vivo* and *in vitro*. *Neuroscience*, 188: 95—108
- Paffett-Lugassy N, Hsia N, Fraenkel P G *et al*, 2007. Functional conservation of erythropoietin signaling in zebrafish. *Blood*, 110(7): 2718—2726
- Stachura D L, Reyes J R, Bartunek P *et al*, 2009. Zebrafish kidney stromal cell lines support multilineage hematopoiesis. *Blood*, 114(2): 279—289
- Stachura D L, Svoboda O, Lau R P *et al*, 2011. Clonal analysis of hematopoietic progenitor cells in the zebrafish. *Blood*, 118(5): 1274—1282
- Wilkins M R, Gasteiger E, Bairoch A *et al*, 1999. Protein identification and analysis tools in the ExPASy Server. *Methods in Molecular Biology*, 112: 531—552

## CLONING AND BIOINFORMATICS ANALYSIS OF EPO GENE IN ANTARCTIC ICEFISH *CHIONODRACO HAMATUS* AND *TREMATOMUS BERNACCHII*

ZHANG Chi<sup>1</sup>, JIANG Hua-Peng<sup>1</sup>, WANG Cong-Cong<sup>1,2</sup>, XU Qiang-Hua<sup>1,2,3,4,5</sup>

(1. College of Marine Sciences, Shanghai Ocean University, Shanghai 201306, China; 2. Key Laboratory of Sustainable Exploitation of Ocean Fisheries Resources, Ministry of Education, Shanghai 201306, China; 3. Scientific Observing and Experimental Station of Oceanic Fishery Resources, Ministry of Agriculture, Shanghai Ocean University, Shanghai 201306, China; 4. National Distant-water Fisheries Engineering Research Center, Shanghai Ocean University, Shanghai 201306, China; 5. Collaborative Innovation Center for Distant-water Fisheries, Shanghai 201306, China)

**Abstract** We used PCR (polymerase chain reaction), TA (based on thymine (T) and adenine (A)) cloning, and next-generation sequencing technology to isolate genomic DNA sequences of EPO (erythropoietin) gene from healthy *Chionodraco hamatus* and *Trematomus bernacchii*, and then we used bioinformatics methods to analyze their genetic characteristics, such as genetic structure and conserved domain. The EPO gene of *C. hamatus* and *T. bernacchii* are 2727 bp and 2820 bp, respectively. Both of them contain five exons and four intron structures, and the coding region is 558 bp in length, encoding 185 amino acids. In addition, EPO of both fish are a relatively soluble and acidic protein. The tertiary structure prediction shows a typical structure characteristic that they fold into a four-parallel- $\alpha$ -helix structure, which assure EPO keep original function in the organism. Amino acids alignment shows 95.2% identity in homology of the EPOs between *C. hamatus* and *T. bernacchii*. Compared with other species, the EPOs shared 53.9%—85.5% identities. The phylogenetic tree of the EPO nucleic acid sequences shows a close relationship among *C. hamatus*, *C. hamatus* and other fishes. The phyletic evolution agrees with their genetics, showing a conservative evolution progress of the EPO gene. This study may provide a theoretical basis for revealing the genetic characteristics of the Antarctic icefish EPO gene, and lay foundation to EPO study on long-term biological adaptation under an extremely low temperature and high oxygen environment.

**Key words** Antarctic icefish; EPO gene; gene cloning; adaptive evolution; bioinformatics