DOI: 10.3724/SP.J.1035.2010.00922

非洲爪蟾短型肽聚糖识别蛋白基因的克隆与鉴定

齐志涛[1,2,3] 高 谦[1] 黄 贝[1,3] 昌鸣先[1] 聂 品[1]

(1. 中国科学院水生生物研究所, 淡水生态与生物技术国家重点实验室, 武汉 430072; 2. 盐城工学院海洋技术系, 江苏省沿海池塘养殖生态重点实验室, 盐城 224051; 3.中国科学院研究生院, 北京 100049)

摘要: 采用生物信息学方法首次对非洲爪蟾短型肽聚糖识别蛋白(xePGRP-S)基因进行了克隆, 并对其在胚胎发育和成年爪蟾各组织中的表达状况进行了分析。*xePGRP-S* cDNA 全长 720 bp, 开放阅读框为 549 bp, 编码 182 个氨基酸。序列比对显示 xePGRP-S 与其他物种 PGRP-S 的序列相似性在 42.4%—50.5% 之间。RT-PCR 显示在非洲爪蟾胚胎发育至 3d 时可以明显检测到 *xePGRP-S* 的表达, 之后呈持续性表达, 且在所检测的心、肝、脾、肺、肾、肠和胃这 7 种组织器官中呈组成型表达。

关键词: 非洲爪蟾; 肽聚糖识别蛋白; 克隆; 同源性; 组织表达

中图分类号: S917.4 文献标识码: A 文章编号: 1000-3207(2010)05-0922-05

肽聚糖识别蛋白(Peptidoglycan recognition proteins, PGRPs)是一类重要的模式识别分子, 在宿主对细菌类病原体的固有免疫反应中起着关键性的作用[1]。目前, PGRP 基因已经在果蝇(*Drosophila melanogaster*)、硬骨鱼类、哺乳类中被克隆鉴定。根据氨基酸长度的不同, PGRP 可以分为长型(PGRP-L)、短型(PGRP-S)和中间型(PGRP-I), PGRP-L 和 PGRP-S 在昆虫和脊椎动物中广泛存在, 而 PGRP-I 仅在哺乳类中被报道[1]。果蝇中含有 13 个 PGRP 基因, 分布在 3 条染色体的 8 个不同的基因座上, 至少编码 17 种 PGRP 蛋白[1]。非洲按蚊(*Anopheles gambiae*)中已报道 7 种 PGRP 基因, 其中有 3 个 PGRP 基因编码 3 种 PGRP-S, 4 个 PGRP 基因编码 7 种 PGRP-L[2]。目前在斑马鱼(*Danio rerio*)中报道了 6 种 PGRP, 即 PGRP-SC2、-SC1a、-L、-2A、-2B 和-6, 其中 *PGRP-SC1a* 和 *PGRP-SC*2 基因含有 4 个外显子, 其余的则有 5 个外显子[3]。在哺乳类中只发现了 4 种 PGRP 蛋白, 分别是 PGLYRP-1、-2、-3 和-4, PGLYRP-1 为短型 PGRP, PGLYRP-2 为长型 PGRP, 二者分布在人的 19 号染色体上, 而属中间型的 PGLYRP-3 和-4 位于人 1 号染色体上[4]。

两栖类作为水生脊椎动物向陆生脊椎动物、变温动物向恒温动物演化的过渡类群, 在生物进化过程中具有重要的地位[5,6]。迄今为止, 尚未见有关两栖类 PGRP 的研究报道。本文采用生物信息学方法, 首次对两栖类模式生物非洲爪蟾(*Xenopus tropicalis*)的一种短型 PGRP 基因(*xePGRP-S*)进行了克隆及序列分析, 为进一步研究其在两栖类先天免疫反应中的作用奠定基础。

1 材料与方法

1.1 实验动物

本研究所用非洲爪蟾购自中国科学院遗传与发育生物学研究所。运回本实验室后, 定期投喂猪肝饲养, 水温保持在 23°C 左右。在实验室饲养两周后, 取心、肝、脾、肺、肾、肠等组织提取总 RNA, 按 SMART cDNA Synthesis Kit 操作手册合成 cDNA。

1.2 *xePGRP*-S cDNA 全长的获得

通过对非洲爪蟾基因组数据库进行分析, 在 Scaffold_31 序列上发现了一个短型 PGRP 基因。根据该序列设计 xePGRP-SF1/SR1 引物扩增 *xePGRP-S* cDNA 中间片段。PCR 反应条件为; 94 °C 预变性

收稿日期: 2009-06-04; 修订日期: 2010-03-18

基金项目: 国家自然科学基金资助项目(30830083); 中国科学院创新基金(KSCX2-YW-N-021)资助

作者简介: 齐志涛(1979—), 男, 内蒙古赤峰人; 博士; 研究方向为鱼类免疫与分子生物学。E-mail: qizhitao@ycit.edu.cn

通讯作者: 聂品, E-mail: pinnie@ihb.ac.cn

5min; 94°C 30s, 56°C 40s, 72°C 90s, 35 个循环; 72°C 延伸 10min。根据获得的中间片段设计 RACE-PCR 引物, 分别扩增 *xePGRP-S* cDNA 的 5′和 3′ 端序列。所用引物(表 1)。

1.3 RT-PCR 检测 *xePGRP-S* 在胚胎发育和成年阶段的组织表达状况

根据 Dale 和 Slack[7]报道的方法, 在 23°C 条件下分别收取非洲爪蟾受精卵和出膜后幼体, 取样时间点为受精后 3、6、15、48、72、111、216 和 360h。用 Trizol 分别提取上述时间点胚胎发育样品和非洲爪蟾正常成年个体的心、肝、脾、肺、肾、肠、胃的总 RNA, 取 2 μg 逆转录成 cDNA, 使用 RT-PCR 方法检测在胚胎发育不同时期和成年个体各组织器官中 *xePGRP-S* 的表达水平。所用引物见表 1。

表 1　本研究所用引物及其用途
Tab. 1　Primers used in the study

引物名称 Primer	序列 Sequence from 5′ to 3′	引物用途 Usage
xePGRP-SF1	TGCCCTACAATCCTATC	PGRP-S 中间序列扩增
xePGRP-SR1	GCCAAGTTTTTACGGTG	
UPM	CTAATACGACTCACTATAGGGC	RACE-PCR 通用引物
xePGRP-S5out	GCGTAGCCATTGCTGTCC	5′-RACE 特异性引物
xePGRP-S5in	ATAGGTGACAGGCGTAGCC	
xePGRP-S3out	CACCGTAAAAACTTGGCC	3′-RACE 特异性引物
xePGRP-S3in	CTTGGCCCCGTTTCCAAGC	
xePGRP-SF2	CTACGCCTGTCACCTAT	RT-PCR
xePGRP-SR2	TTCCATCTTCACCAACT	
xeActin-F	GGTCGCCCAAGACATCAG	RT-PCR
xeActin-R	GCATACAGGGACAACACA	

1.4 序列分析与系统发育树的构建

利用 ExPASy 网站(http: //www.expasy.org)的 Translate 程序进行氨基酸序列推导。采用 SignalP3.0 预测信号肽[8]。使用 NCBI 上的 BLASTp 软件进行同源基因搜索; 用 PROSITE 预测氨基酸序列中可能的结构域及 N-糖基化位点。氨基酸序列比对使用 Clustalw 1.83[9], 同源基因的序列相似性用 DNASTAR 软件包中的 MEGALIGN 程序分析[10]。采用 Mega4.1 软件中的 N-J 法构建系统发育树, 并进行 1000 次 Bootstrap 分析以评估各分支的可信度[11]。采用 Cn3D (http://www.ncbi.nlm.nih.gov/structure/CN3D/ cn3d.html)软件预测 xePGRP-S 的 3D 结构。

2 结　果

2.1 *xePGRP-S* cDNA 全长序列及其编码的氨基酸序列

xePGRP-S cDNA(GenBank 登录号为 HM012720)全长 720 bp, 包括 5′-UTR 19 bp, 3′-UTR 152 bp, 开放阅读框为 549 bp, 编码 182 个氨基酸, 分子量为 19.6 kD, 等电点为 9.22。*xePGRP-S* 基因定位于 Scaffold_31 序列上, 其基因全长为 4.6 kb, 含有 4 个外显子和 3 个内含子, 外显子长度分别为 49、202、119 和 179 bp。内含子剪切模式符合 GT-AG 模式。SignalP 预测结果显示 xePGRP-S 含有一个 17 氨基酸的信号肽(MLRLLAFFATLCAVANS)。xePGRP-S 与其他物种 PGRP-S 的序列相似性在 42.4%—50.5%之间, 与果蝇 PGRP-SC2 的相似性最高, 与大肠杆菌 *Enterobacteria phage* 的 T7 酰胺酶相似性为 28.3%。T7 酰胺酶所必需的一些氨基酸位点, 如 His-17、His-122 及 Cys-130 在 xePGRP-S 中同样保守(图 1)。

2.2 系统发育分析

N-J 系统进化树(图 2)显示, xePGRP-S 与其他脊柱动物的 PGRP-S 聚为一枝。相对于 *PGRP-L* 基因而言, *PGRP-S* 基因较为古老, *PGRP-L* 可能是 *PGRP-S* 通过基因重复而产生的。

2.3 xePGRP-S 3D 结构预测

对 xePGRP-S 进行 3D 结构预测显示, 与人 PGRP-S 结构相似, xePGRP-S 核心结构包括 3 个 α 螺旋和 5 个 β-折叠(图 3)。

2.4 xePGRP-S 在早期发育阶段的表达分析

采用 RT-PCR 检测了 xePGRP-S 在非洲爪蟾胚

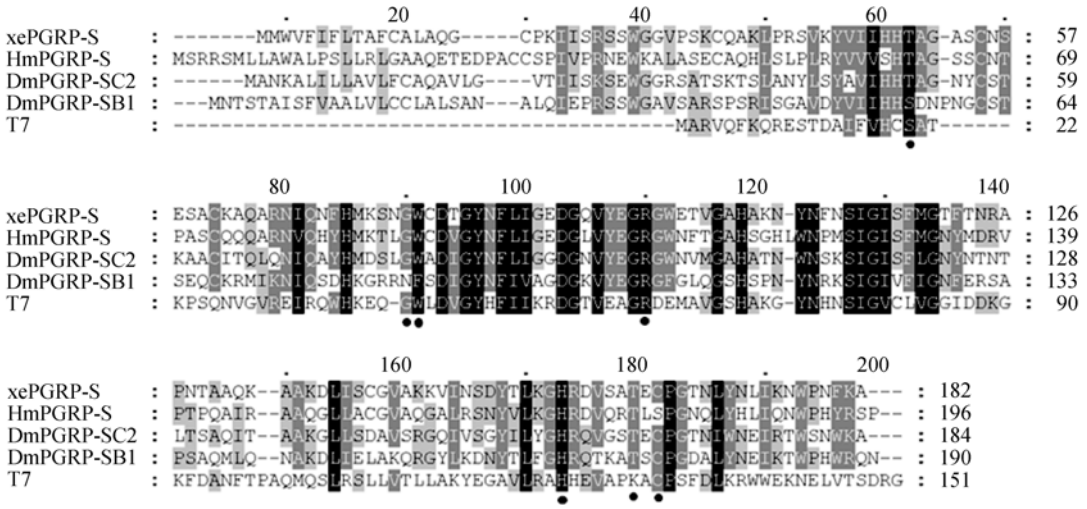


图 1 xePGRP-S 与其他几种 PGRP-S 及 T7 酰胺酶的氨基酸序列比对

Fig. 1 Multiple alignment of xePGRP-S with other PGRP-Ss and T7 amidase by using CLUSTL W

用 CLUSTL W 软件进行序列比较并用 GeneDoc 软件进行修饰; 保守的 T7 酰胺酶氨基酸位点用圆点标注; 所用序列见表 2

The three levels of shading indicate 100%, 80% or 60% similar amino acids decorated with the GeneDoc program. The amino acids marked with dots are necessary for T7 amidase. The GenBank accession numbers of all sequences used are listed in Tab. 2

表 2 用于系统进化树构建的各物种 **PGRP** 分子及其 **GenBank** 登录号

Tab. 2 Peptidoglycan recognition protein sequences used for the multiple alignment and the construction of phylogenetic tree with their GenBank accession numbers

物种 Species	蛋白名称 Protein	GenBank 登录号 Accession No.	物种 Species	蛋白名称 Protein	GenBank 登录号 Accession No.
Homo sapiens (H.m)	PGRP1	O75594	*Drosophila simulans* (D.s)	PGRP-SA	AAQ64766
	PGRP-L	AAL05629			
	PGRP-Iα	NP_443123		PGRP-SC2	XP_699934
	PGRP-Iβ	Q96LB8	*Danio rerio* (D.r)	PGLYRP6	ABE01406
Macaca mulatto (Ma.m)	PGRP-L	XP_001111895		PGRP2-A	ABE01404
Mus musculus (M.m)	PGRP1	NP_033428		PGRP-SB1	CAD89129
	PGRP-L	NP_067294		PGRP-SC1B	CAD89167
	PGRP-Iβ	NP_997130		PRGRP-SC2	CAD89177
Sus scrofa (S.s)	PGRP-L	NP_998903	*Drosophila melanogaster* (D.m)	PGRP-LCx	AAM18530
Canis familiaris (C.f)	PGRP-L	XP_852999		PGRP-LCy	AAQ16306
Rattus norvegicus (R.n)	PGRP1	AAF73252		PGRP-LD	AAG36063
Camelus dromedaries (C.d)	PGRP	CAC19553		PGRP-LB	NP_731576
Bos taurus (B.t)	PGRP2	XP_588006		PGRP-LE	NP_573073
Trichoplusia ni (T.n)	PGRP	AAC31820			
Enterobacteria phage	T7 lysozyme	NP_041973			

胎早期发育阶段的表达情况。图 4 显示, 在胚胎发育至 3d 时可以明显检测到 xePGRP-S 的表达, 之后 xePGRP-S 呈持续性表达。

2.5 xePGRP-S 基因的组织表达特征

采用 RT-PCR 方法对 xePGRP-S 基因在非洲爪蟾各组织中的表达状况进行分析, 结果显示, 在不同组织中均能检测到作为对照的 *β-actin* 基因的表达, 说明所制备的 cDNA 模板完整(图 5)。xePGRP-S 在心、肝、脾、肺、肾、肠和胃中均能检测到, 呈组成型表达(图 5)。

3 讨 论

随着人类及其他物种基因组计划的完成, 从果蝇、鱼类到哺乳类, 越来越多的 PGRPs 被报道, 但

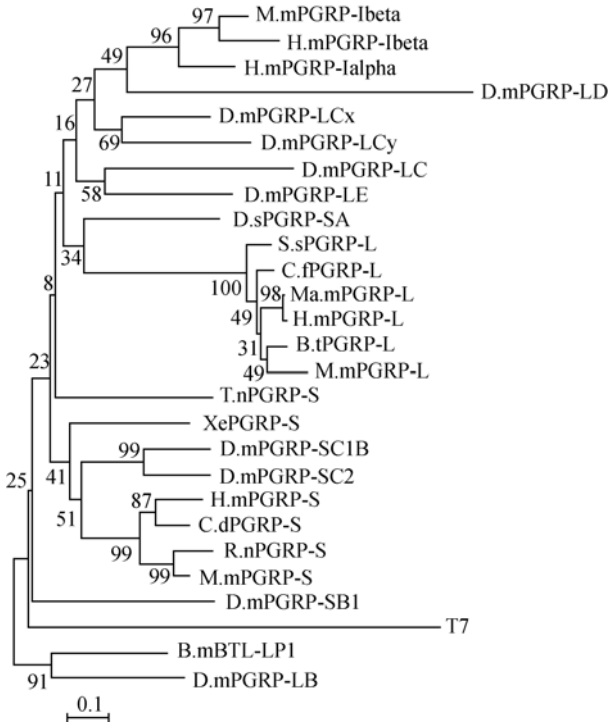


图 2　依据 PGRP 分子氨基酸序列采用邻接法构建的系统进化树(所用序列见表 2)

Fig. 2　An unrooted phylogenetic tree constructed by the neighbor joining method from amino acid sequences of PGRPs

The number indicates the bootstrap confidence values obtained for each node after 1000 replications. The GenBank accession numbers of all sequence used are listed in Tab. 2

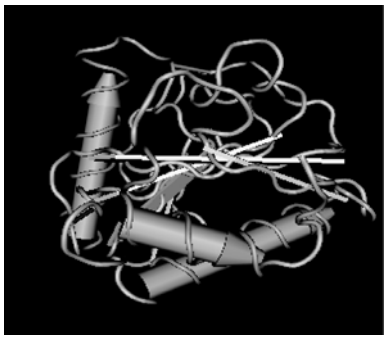


图 3　预测的 xePGRP-S 3-D 结构

Fig. 3　The 3-D structure of xePGRP-S predicted using Cn3D software

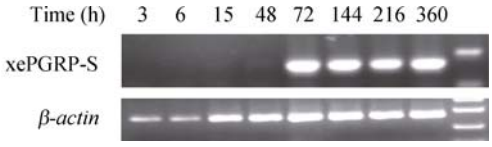


图 4　RT-PCR 分析非洲爪蟾胚胎发育阶段 xePGRP-S 的表达情况

Fig. 4　The expression analysis of xePGRP-S in early developing stage of xenopus by using RT-PCR

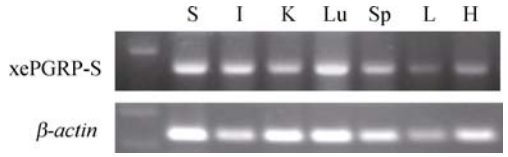


图 5　RT-PCR 分析非洲爪蟾成体各组织器官中 xePGRP-S 的表达状况

Fig.5　The expression analysis of *xePGRP-S* in the tissues of normal xenopus by using RT-PCR

H. 心; L. 肝; Sp. 脾; Lu. 肺; K. 肾; I. 肠; S. 胃

H. Heart; L. Liver; Sp. Spleen; Lu. Lung; K. Kidney; I. Intestine; S. Stomach

有关两栖类 PGRPs 的报道较少。本研究通过数据库挖掘和基因克隆与测序获得了两栖类模式生物非洲爪蟾的一种可能的 PGRP 的 cDNA 全长, 与其他物种 PGRP-S 的序列相似性在 42.4%—50.5%之间, 其中与果蝇 PGRP-SC2 的相似性达 50.5%。3D 结构预测显示 xePGRP-S 为 α-螺旋/β-折叠混合型蛋白, 含有 3 个 α-螺旋和 5 个 β-折叠结构, 此种构象在短型 PGRP 中具有保守性。因此推测本研究所克隆的基因应该是 PGRP 家族中短型亚族的成员。

昆虫和哺乳类的 *PGRP* 表达模式差异较大, 昆虫的 *PGRP* 主要在具有免疫活性的组织中表达[12]。哺乳类不同类型 *PGRP* 在各组织器官中表达水平存在较大差异, *PGLYRP*-1 主要在骨髓中表达[4], *PGLYRP*-3 主要在食管中表达, *PGLYRP*-4 仅在食管、扁桃体和胸腺中表达[13], 而 *PGLYRP*-2 则在各种组织器官中呈组成型表达[4]。本文所克隆的 xePGRP-S 在非洲爪蟾各组织器官中呈组成型表达, 同时, 在非洲爪蟾早期胚胎发育阶段也具有较高的表达水平。这提示 xePGRP-S 在非洲爪蟾胚胎及成体阶段的固有性免疫反应方面均起作用。

目前, 在海胆、文昌鱼及软骨鱼类中均只发现 PGRP-S, 推测在肽聚糖识别蛋白家族中 PGRP-S 可能是比较古老的类型[14]。PGRP-S 的主要功能为酰胺酶(如 *Drosophila melanogaster* PGRP-SC1B、-SC2)[15]和杀菌活性(如鼠和牛的 PGRP-S)[13,16]。本文通过序列分析显示, 酰胺酶活性所需的氨基酸位点在 xePGRP-S 中同样具有保守性, 但 xePGRP-S 是否也具有酰胺酶活性有待进一步研究确定。

参考文献:

[1]　Werner T, Liu G, Kang D, *et al*. A family of peptidoglycan

recognition proteins in the fruit fly *Drosophila melanogaster* [J]. *Proceeding of National Academy of Science of United states of America*, 2000, **97**: 13772—13777

[2] Chang M X, Nie P, Wei L L. Short and long peptidoglycan recognition proteins (PGRPs) in zebrafish, with findings of multiple PGRP homologs in teleost fish [J]. *Molecular Immunology*, 2007, **44**: 3005—3023

[3] Liu C, Gelius E, Liu G, *et al*. Mammalian peptidoglycan recognition protein binds peptidoglycan with high affinity, is expressed in neutrophils, and inhibits bacterial growth [J]. *Journal of Biology Chemistry*, 2000, **275**: 24490—24499

[4] Liu C, Xu Z, Gupta D, *et al*. Peptidoglycan recognition proteins: A novel family of four human innate immunity pattern recognition molecules [J]. *Journal of Biology Chemistry*, 2001, **276**: 34686—34694

[5] Li S S. Cytotaxonomy of Amphibians in China [M]. Beijing: Science Press. 2007, 1—8 [李树深. 中国两栖动物细胞分类学. 北京; 科学出版社. 2007, 1—8]

[6] Zhang J F, Nie L W, Wang Y, *et al*. Sequence of three mitochondrial genes of *L. Kuhlii* and *L. Fragilis* and the phylogenetic relationships among frogs, salamanders and caecilians [J]. *Acta Hydrobiologica Sinica*, 2007, **31**(6): 822—828 [张际峰, 聂刘旺, 王洋, 等. 大头蛙和脆皮大头蛙线粒体 3 个基因的测定及两栖类亲缘关系研究. 水生生物学报, 2007, **31**(6): 822—828]

[7] Dale L, Slack J M. Fate map for the 32-cell stage of *Xenopus laevis* [J]. *Development*, 1987, **99**: 527—551

[8] Bendtsen J D, Nielsen H, Gunnar von H, *et al*. Improved prediction of signal peptides: SignalP 3.0 [J]. *Journal of Molecular Biology*, 2004, **340**: 783—795

[9] Thompson J D, Higgins D G, Gibson T J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice [J]. *Nucleic Acids Research*, 1994, **22**: 4673—4680

[10] Clewley J P, Arnold C. MEGALIGN. The multiple alignment module of LASERGENE [J]. *Methods in Molecular Biology*, 1997, **70**: 119—129

[11] Kumar S, Tamura K, Nei M. MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment [J]. *Brief Bioinformatics*, 2004, **5**: 150—163

[12] Dimopoulos G, Christophides G K, Meister S, *et al*. Genome expression analysis of Anopheles gambiae: responses to injury, bacterial challenge, and malaria infection [J]. *Proceeding of National Academy of Science of United states of America*, 2002, **99**: 8814—8819

[13] Lu X, Wang M, Qi J, *et al*. Peptidoglycan recognition proteins are a new class of human bactericidal proteins [J]. *Journal of Biology Chemistry*, 2006, **281**: 5895—5907

[14] Dziarski R, Gupta D. The peptidoglycan recognition proteins (PGRPs) [J]. *Genome Biology*, 2006, **7**: 232—245

[15] Mellroth P, Karlsson J, Steiner H. A scavenger function for a Drosophila peptidoglycan recognition protein [J]. *Journal of Biology Chemistry*, 2003, **278**: 7059—7064

[16] Tydell C C, Yount N, Tran D, *et al*. Isolation, characterization, and antimicrobial properties of bovine oligosaccharide-binding protein [J]. *Journal of Biology Chemistry*, 2002, **277**: 19658—19664

CLONING AND IDENTIFICATION OF A SHORT TYPE PEPTIDOGLYCAN RECOGNITION PROTEIN IN *XENOPUS TROPICALIS*

QI Zhi-Tao[1,2,3], GAO Qian[1], HUANG Bei[1,3], CHANG Ming-Xian[1] and NIE Pin[1]

(1. *State Key Laboratory of Freshwater Ecology and Biotechnology*, *Institute of Hydrobiology*, *Chinese Academy of Sciences*, *Wuhan* 430072; 2. *Department of Ocean Technology*, *Yancheng Institute of Technology*, *Key Laboratory of Aquaculture and Ecology of Coastal pool of Jiangsu Province*, *Yancheng* 224051; 3. *Graduate University of the Chinese Academy of Sciences*, *Beijing* 100049)

Abstract: A short type peptidoglycan recognition protein (PGRP-S) was first cloned from the amphibian model animal *Xenopus tropicalis* by using the bioinformatics method, and the expression pattern of *xePGRP-S* in early developing stage and normal xenopus was also analyzed. The *xePGRP-S* cDNA was 720 bp in length, containing a 549 bp open reading frame encoding 182 a.a. Multiple sequence alignment showed that the xePGRP-S shared 42.4%—50.5% sequence identities with other PGRP-S. RT-PCR showed that *xePGRP-S* mRNA was not detected in the early development stages (3—48h) until the third day. In the normal adult xenopus, *xePGRP-S* mRNA was constitutively expressed in all the detected tissues including heart, liver, spleen, lung, kidney, intestine and stomach.

Key words: *Xenopus tropicalis*; Peptidoglycan recognition protein; Clone; Identity; Tissue expression