

旋毛虫 *plancitoxin-1-like* 基因的序列分析与编码蛋白结构预测

廖成水^{1,4}, 王晓利², 田文静¹, 杨亚东¹, 张梦珂¹, 王安琪⁴, 张春杰¹, 刘明远⁴,
Boireau Pascal^{4,5}, 程相朝^{1,3*}

(1. 河南科技大学, 动物科技学院/洛阳市活载体生物材料与动物疫病防控重点实验室, 洛阳 471023;

2. 河南科技大学 医学院, 洛阳 471023; 3. 洛阳职业技术学院, 洛阳 471003;

4. 吉林大学 人兽共患病研究所/人兽共患病研究教育部重点实验室, 长春 130062;

5. 法国国家食品、环境及劳动卫生署动物卫生实验室, 迈松阿尔福 94706)

摘要: 基因组预测旋毛虫拥有 125 种庞大的 DNase II 家族蛋白, 但只有 *plancitoxin-1* 唯一在 C 端含有一个保守的 DNase II 活性位点(HKD 基序)。为研究旋毛虫 *plancitoxin-1* 基因的序列结构特征, 以旋毛虫总 RNA 为模板进行 RT-PCR, 获得 *plancitoxin-1* 基因的序列, 并利用生物信息学方法进行系统性分析。结果显示, 扩增得到的核苷酸序列比 GenBank 中相关序列短 210 bp, 编码 294 个氨基酸, 由 20 种氨基酸组成, 重新命名为 *plancitoxin-1-like*, 且氨基酸序列中在 N 端和 C 端都具有 HKD 基序。蛋白质理论相对分子质量为 33.19 ku, 等电点为 9.17, 是稳定存在的亲水性蛋白。*Plancitoxin-1-like* 的氨基酸序列与其他物种 DNase II 的相似性为 35% 左右。*Plancitoxin-1-like* 含有 4 个 N-糖基化位点、3 个 O-糖基化位点、31 个磷酸化位点、17 个 B 细胞线性结合位点和 9 个 T 细胞结合位点以及 3 个二硫键。二级结构中, α -螺旋占 14.29% (42 个), 延伸链占 28.23% (83 个), 无规卷曲占 55.10% (162 个)。本研究成功克隆旋毛虫 *plancitoxin-1-like* 基因, 对其序列结构进行分析, 为进一步研究庞大的 DNase II 家族蛋白在旋毛虫发育和感染中的作用奠定基础。

关键词: 旋毛虫; DNase II; 克隆; 生物信息学

中图分类号: S852.731

文献标志码: A

文章编号: 0366-6964(2018)01-0218-05

Cloning and Sequence Analysis of *Plancitoxin-1-like* Gene in *Trichinella spiralis*

LIAO Cheng-shui^{1,4}, WANG Xiao-li², TIAN Wen-jing¹, YANG Ya-dong¹, ZHANG Meng-ke¹,
WANG An-qi⁴, ZHANG Chun-jie¹, LIU Ming-yuan⁴, BOIREAU Pascal^{4,5}, CHENG Xiang-chao^{1,3*}

(1. College of Animal Science and Technology / Luoyang Key Laboratory of Live Carrier
Biomaterial and Animal Disease Prevention and Control, Henan University of
Science and Technology, Luoyang 471023, China;

2. Medical College, Henan University of Science and Technology, Luoyang 471023, China;

3. Luoyang Vocational and Technical College, Luoyang 471003, China;

4. Key Laboratory of Zoonosis of Ministry of Education, Institute of Zoonosis,
Jilin University, Changchun 130062, China;

5. ANSES's Laboratory for Animal Health, Maisons-Alfort 94706, France)

Abstract: It is noteworthy that more extensive expansion of the DNase II-like protein family (estimated 125 genes) in the genome of the parasitic nematode *Trichinella spiralis* (*T. spiralis*).

收稿日期: 2017-06-26

基金项目: 国家自然科学基金(31030064); 国家自然科学基金面上项目(31572489); 河南科技大学博士启动基金(13480071); 河南省高等学校重点科研项目计划(17A230009)

作者简介: 廖成水(1986-), 男, 博士, 主要从事微生物感染与先天性免疫细胞作用研究, E-mail: liaochengshui33@163.com

* 通信作者: 程相朝, 教授, 主要从事动物疫病分子与免疫学机制研究, E-mail: chengxch@126.com

However, only *plancitoxin-1* possesses one predicted active site of DNase II (HKD motif) in its C-terminus domain in 125 DNase II-like protein family. To explore the molecular characteristics of *plancitoxin-1* gene in *T. spiralis*, *plancitoxin-1* gene was cloned by RT-PCR from total RNA of *T. spiralis* larvae, and analyzed with bioinformatics software. The results showed that the sequences of *plancitoxin-1* cloned from *T. spiralis* were shorter (210 bp less) than the XM_003370715.1 sequence in GenBank, and named after *plancitoxin-1-like*. The gene encoded a protein of 294 amino acids. Intriguingly, there were two HKD motifs in the N- and C-termini in the cloned sequences. *plancitoxin-1-like* is a stable protein with a predicted molecular mass (Mr) of 33.19 kDa and theoretical pI of 9.17. By sequence alignment, *plancitoxin-1-like* protein revealed 35% sequence identity to previously reported DNase II family members from other species. A total of 4 N-linked glycosylation sites, 3 O-linked glycosylation sites, 31 phosphorylation site 17 linear B cell epitope, 9 T cell binding sites, and 3 disulfide bond were predicted in *plancitoxin-1-like*. Analysis of secondary structure revealed that α -helix, extended strand, and random coil were 14.29% (42), 28.23% (83), and 55.10% (162), respectively. In this study, the *plancitoxin-1-like* gene was cloned and analyzed. The results laid the foundation for the further study on the role of development infection in *T. spiralis*.

Key words: *Trichinella spiralis*; DNase II; clone; bioinformatics

脱氧核糖核酸酶(deoxyribonuclease, DNase)具有 DNase I(EC 3.1.21.1)和 DNase II(EC 3.1.22.1)两大类^[1]。1988 年首次报道人类 DNase II 基因后^[2], DNase II 在脊椎动物、无脊椎动物和单细胞原虫中被陆续报道^[3]。旋毛虫是一种寄生于骨骼肌细胞的重要人兽共患胞内寄生虫,主要分布于中国、阿根廷和部分东欧国家等^[4]。旋毛虫基因组显示旋毛虫具有多达 125 个 DNase II 蛋白家族基因^[5]。在这之前含有最多 DNase II 的物种是秀丽杆线虫(仅 10 种)^[6],旋毛虫如此庞大的蛋白家族在其他物种中实属罕见,而关于旋毛虫 DNase II 的研究仍未见有明显进展。经典 DNase II 的氨基酸序列在 N 端和 C 端都具有高度保守区域,即 HKD 基序^[7]。在线软件分析显示旋毛虫 125 个基因中仅仅 *plancitoxin-1* (1 095 bp, GenBank accession No. XM_003370715.1)含有一个 HKD 基序,且位于 C 端。因此,本研究克隆得到旋毛虫 *plancitoxin-1* 基因,并分析其序列特征,以期为庞大的 DNase II 蛋白家族在旋毛虫发育和感染中的作用研究提供科学资料。

1 材料与方 法

1.1 主要材料

Wistar 大鼠(200.0 g \pm 20.0 g)购自吉林大学白求恩医学院动物实验中心;中国河南猪旋毛虫分离株 *Trichinella spiralis* (genotype T1),国际标准

虫种编号为 ISS534,由本实验室小鼠传代保种;pMD18-T 购自宝生物工程(大连)有限公司;大肠杆菌 DH5 α 购自德国 Novagen 公司,由本实验室常规冻存;胰蛋白胨和酵母浸出物购自英国 OXOID 公司;胃蛋白酶购自中国医药集团总公司;RPMI-1640 培养液购自 Hyclone 公司;青霉素、硫酸链霉素均购自美国 Sigma-Aldrich 公司;Taq DNA 聚合酶、dNTPs、DNA Marker、BamH I、Hind III 和 RNA PCR Kit (AMV) Ver. 3.0 均购自宝生物工程(大连)有限公司;其他化学试剂均为国产分析纯。

1.2 *Plancitoxin-1* 基因的克隆

根据 GenBank 上的旋毛虫 *plancitoxin-1* 基因序列(GenBank No.: XM_003370715.1)设计引物(上游引物:5'-GGATCCATGGACGCACGTCG-GCCGGTAT-3', BamH I;下游引物:5'-AAGCTT TCAATATGGTGAATAGGACAAAGT-3', Hind III)。常规方法收集旋毛虫肌幼虫、新生幼虫和成虫虫体,根据 Trizol 试剂盒的方法得到三个时期总 RNA。根据 RNA PCR Kit (AMV) Ver. 3.0 试剂盒的方法得到 cDNA。以旋毛虫 cDNA 为模板,常规 PCR 扩增获得 *plancitoxin-1* 基因。

1.3 同源性分析

利用 CLASTALX 2.1 进行二十余种物种 DNase II 蛋白多重序列分析。采用 PHYLIP 3.695-NEIGHBOR 重复 1 000 次的 bootstrap 法进行氨基酸序列的遗传进化树分析。

1.4 理化性质分析

在线网站 ProtParam (<http://web.expasy.org/protparam/>)分析蛋白质相对分子质量大小、理论等电点和氨基酸组成。在线网站 <http://people.mbi.ucla.edu/sumchan/caltor.html> 和 <http://bio-tech.ou.edu/>分别用于分析氨基酸的稀有密码子和重组蛋白质的可溶性。在线网站 ProtScale (<http://web.expasy.org/protscale/>)分析蛋白质的亲水性与疏水性。

1.5 结构分析

在线网站 <http://www.cbs.dtu.dk/services/>的 NetNGlyc、NetOGlyc 和 NetPhos 分别分析 N-糖基化位点、O-糖基化位点和磷酸化位点。在线网站 <http://www.embl-heidelberg.de/predictprotein/predictprotein.html> 分析二硫键数量。<http://tools.immuneepitope.org/bcell/>和 NetCTL (<http://www.cbs.dtu.dk/services/>)预测蛋白可能存在的线性 B 细胞抗原表位和 T 细胞结合位点。在线软件 https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_server.html 预测二级结构。SWISS-MODEL Workspace 分析三级结构。

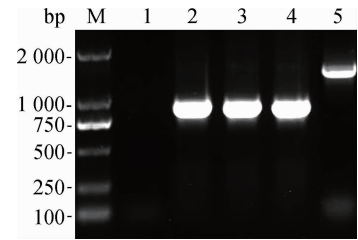
2 结果

2.1 *Plancitoxin-1-like* 基因的克隆

从旋毛虫中扩增得到一条 885 bp 的序列(图 1),与 *plancitoxin-1* (Tsp_09974, GenBank No. XM_003370715.1)相比,核苷酸序列少 210 bp,缺少部分位于 184—393 位,将其命名为 *plancitoxin-1-like*。从肌幼虫、成虫和新生幼虫扩增得到的 *plancitoxin-1-like* 序列均一致,而从基因组扩增的序列与 NCBI 公布的 1 471 bp 一致。*Plancitoxin-1-like* 编码 294 个氨基酸,比 *plancitoxin-1* 少 70 个氨基酸,缺少部分位于 62—131 位。保守区域预测显示 *plancitoxin-1* 只有一个 HKD 基序,且位于 C 端,但 *plancitoxin-1-like* 在 N 端和 C 端均具有 HKD 基序。

2.2 同源性分析

Plancitoxin-1-like 的核苷酸序列与其他物种的 DNase II 的相似性较低,最高的是广杆属线虫 CBR-CRN-6 基因(10.07%),与人类、牛、马、猪、小鼠、大鼠、鸡、斑马鱼、非洲蟾蜍、果蝇、海星、河豚、类鼻疽杆菌、盘基网柄菌、鸡痘病毒、金丝雀痘病毒、冈比亚按蚊、广杆属线虫和秀丽杆线虫等 DNase II 成



M. DNA 相对分子质量标准 (DL2000);1. 水对照;2. 肌幼虫;3. 新生幼虫;4. 成虫;5. 基因组

M. DNA marker ladder (DL2000); 1. ddH₂O control; 2. ML; 3. NBL; 4. Ad; 5. *T. spiralis* genome

图 1 旋毛虫 *plancitoxin-1-like* 基因的 PCR 扩增

Fig. 1 Amplification of *T. spiralis plancitoxin-1-like* gene by PCR

员的相似性集中在 34%~38%,与海星 *plancitoxin-1* 具有 40.18%的相似性,遗传进化关系上被分在一个进化支(图 2),但保守区域基本相同。*Plancitoxin-1-like* 与低等生物具有较近的亲缘关系,且更靠近于 DNase II β 。

2.3 *Plancitoxin-1-like* 的理化性质分析

Plancitoxin-1-like 的理论相对分子质量为 33.19 ku,理论等电点为 9.17。分子式为 C₁₅₀₇ H₂₃₃₃ N₃₉₉ O₄₂₅ S₁。不稳定系数为 39.25,属于稳定存在蛋白质。氨基酸由 20 种氨基酸组成,不含 Sec 和 Pyl,其中频次较高的氨基酸残基有 Leu、Ser 和 Lys;频次较低的氨基酸残基有 Met、Cys 和 His。正电荷残基数(R+K)为 33 个,负电荷残基数(D+E)为 24 个。含有 26 个稀有密码子,其中有两处连续出现稀有密码子,在大肠杆菌中呈可溶性表达的可能性为 8.8%。在 287 位亲水性分值最高(1.967),有最强的疏水性;在 175 位分值最低(-2.811),有最强的亲水性。

2.4 序列结构分析

在 N19GS、N49NT、N209CS 和 N262RS 共有 4 个 N-糖基化位点。在 L22、L26 和 A108 共有 3 个 O-糖基化位点。存在 31 个磷酸化位点,分别为 14 个 Ser、9 个 Thr 和 8 个 Tyr,潜力值均大于 0.5。有 17 个 B 细胞线性结合位点和 9 个 T 细胞结合位点。蛋白质二级结构主要由 42 个 α -螺旋(14.29%)、83 个延伸链(28.23%)和 162 个无规卷曲(55.10%)构成。存在 3 个二硫键,分别位于 59 和 101 位、210 和 269 位、250 和 289 位(图 3)。*Plancitoxin-1-like* 三维结构同源模型见图 4,三级结构预测发现,蛋白质序列与 PDB 数据库中 5i3e.1.A 模板序列相似性为 31.79%。

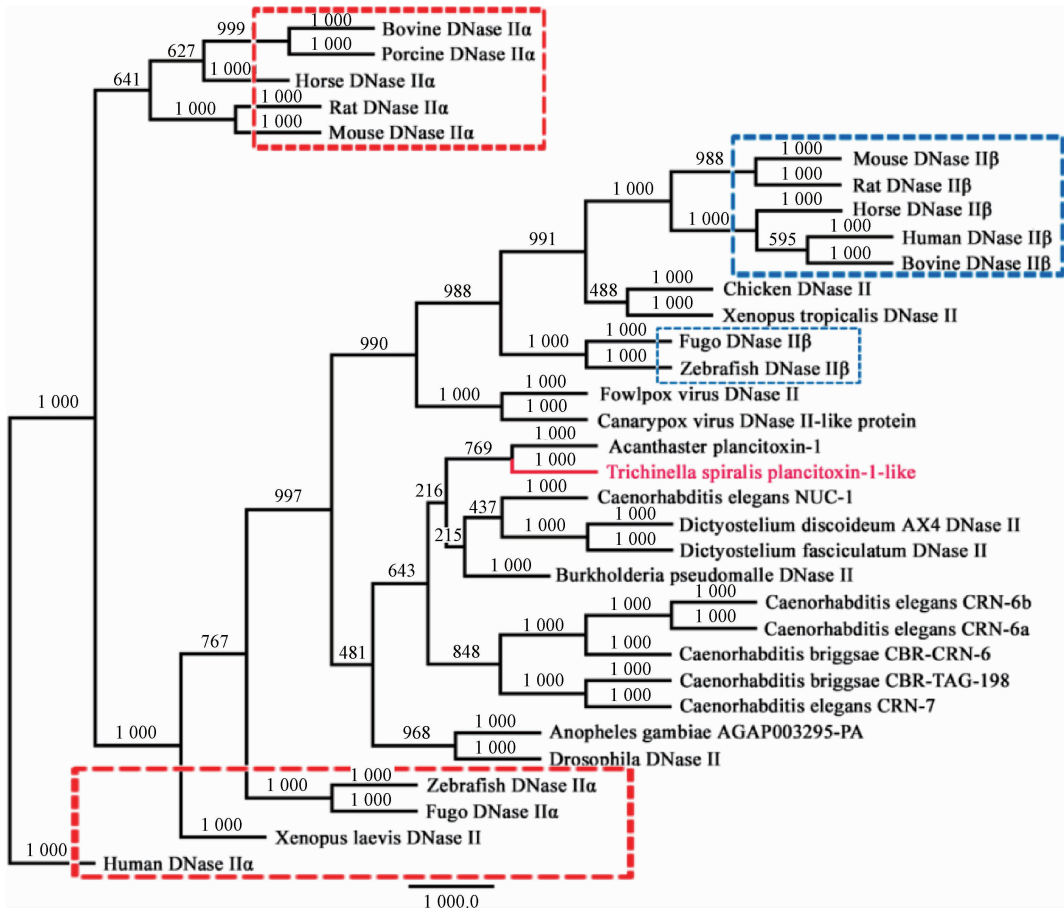


图 2 Plancitoxin-1-like 与其他物种 DNase II 家族蛋白的遗传进化树分析

Fig. 2 Phylogenetic analysis of plancitoxin-1-like and other previously reported DNase II family proteins

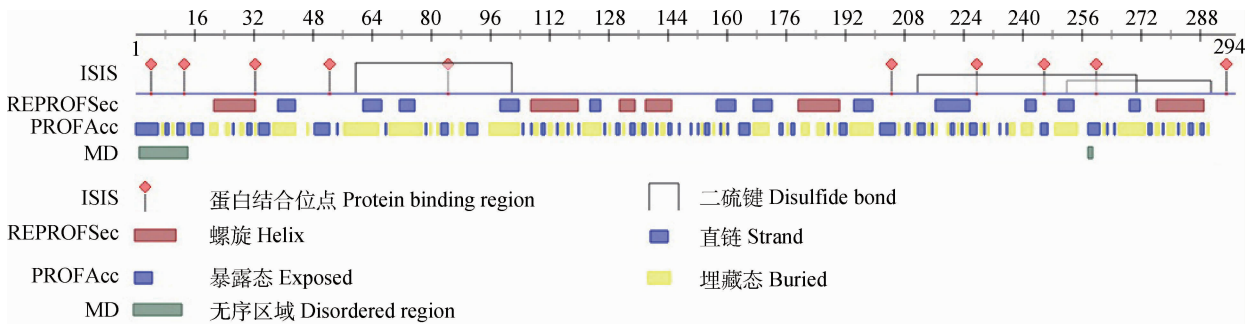


图 3 Plancitoxin-1-like 蛋白二级结构分析

Fig. 3 Predicted secondary structure of plancitoxin-1-like protein

3 讨论

1999年, C. H. Mak等首次发现旋毛虫排泄分泌物具有降解DNA的活性^[8]。来源于排泄分泌物的43 ku蛋白质与哺乳动物DNase II序列具有一定的同源性^[9],但由于缺乏DNase II保守的活性位点,该43 ku蛋白质是否具有DNase II活性仍然存在争议。2011年公布的旋毛虫基因组显示,旋毛虫可能存在一个高达125种DNase II蛋白家族。

人类、牛、马、猪、小鼠、大鼠、鸡、斑马鱼、非洲蟾蜍、果蝇、海星、河豚、类鼻疽杆菌、盘基网柄菌、鸡痘病毒、金丝雀痘病毒、冈比亚按蚊、广杆属线虫和秀丽隐杆线虫等物种只有一种或几种DNase II^[10],而秀丽隐杆线虫具有最多DNase II。但旋毛虫拥有如此多的DNase II家族成员有何作用仍不清楚。43 ku蛋白具有一个螺旋环螺旋基序分子结构,具有调节肌细胞分化的功能^[11],因此与旋毛虫包囊形成有关^[12]。125种旋毛虫DNase II蛋白50%以上的基

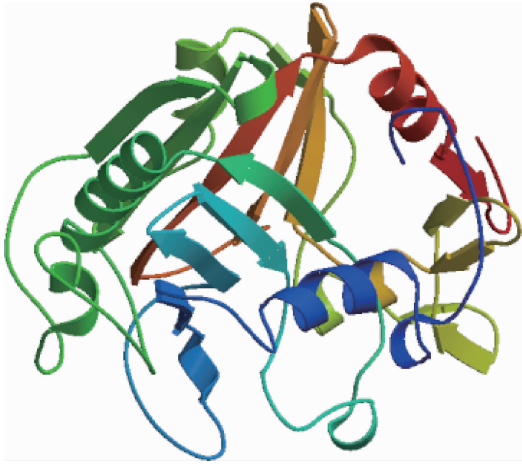


图 4 Plancitoxin-1-like 三维结构同源模型

Fig. 4 Three-dimensional model of plancitoxin-1-like

因编码排泄分泌产物, DNase II 蛋白被认为是控制和预防旋毛虫病的理想疫苗候选蛋白质。

DNase II 在 N 端和 C 端包含有两个保守的 HKD 基序(H-x-K-x(4)-D)^[7]。旋毛虫 p43、SS1 和 AAK85403 三个蛋白与人类 DNase II 具有一定的相似性,但缺乏 DNase II 经典的 HKD 基序^[13]。但除了 plancitoxin-1 具有一个 DNase II 保守活性位点,其他 124 种 DNase II 蛋白都不具有 DNase II 保守活性位点。在本研究中,plancitoxin-1-like 具有两个 HKD 基序,分别位于序列的 N 端和 C 端。由于小鼠、大鼠、河豚和斑马鱼 DNase II β 的 C 端保守区域为 DHSK,因此,根据 DNase II α 和 DNase II β 在 C 端的 HKD 保守基序分别为 DHSK 和 DHAK 的说法很难把 plancitoxin-1-like 归为 DNase II α 还是 DNase II β 。但遗传进化树分析显示,plancitoxin-1-like 的亲缘关系更靠近于 DNase II β (图 2)。

4 结 论

以旋毛虫总 RNA 为模板进行 RT-PCR,获得旋毛虫 *plancitoxin-1-like* 基因,得到的核苷酸序列为 885 bp,编码 294 个氨基酸,由 20 种氨基酸组成。其氨基酸序列在 N 端和 C 端都具有 HKD 基序。蛋白质理论相对分子质量为 33.19 ku,等电点为 9.17,是稳定存在的亲水性蛋白。plancitoxin-1-like 的氨基酸序列与其他物种 DNase II 的相似性为 35%左右。

参考文献 (References):

[1] CUNNINGHAM L, LASKOWSKI M. Presence of two different deoxyribonucleode- polymerases in veal kidney[J]. *Biochim Biophys Acta*, 1953, 11(4):

590-591.

- [2] YASUDA T, TAKESHITA H, IIDA R, et al. Molecular cloning of the cDNA encoding human deoxyribonuclease II[J]. *J Biol Chem*, 1998, 273(5): 2610-2616.
- [3] HOU N, PIAO X, CAI P, et al. A novel Schistosoma japonicum endonuclease homologous to DNase II [J]. *BMC Genomics*, 2015, 16: 126.
- [4] SHIMONI Z, FROOM P. Uncertainties in diagnosis, treatment and prevention of trichinellosis[J]. *Expert Rev Anti Infect Ther*, 2015, 13(10): 1279-1288.
- [5] MITREVA M, JASMER D P, ZARLENGA D S, et al. The draft genome of the parasitic nematode *Trichinella spiralis*[J]. *Nat Genet*, 2011, 43(3): 228-235.
- [6] YU H, LAI H J, LIN T W, et al. Autonomous and non-autonomous roles of DNase II during cell death in *C. elegans* embryos[J]. *Biosci Rep*, 2015, 35(3). pii:e00203.
- [7] SCHÄFER P, CYMERMAN I A, BUJNICKI J M, et al. Human lysosomal DNase IIalpha contains two requisite PLD-signature (HxK) motifs; evidence for a pseudodimeric structure of the active enzyme species [J]. *Protein Sci*, 2007, 16(1): 82-91.
- [8] MAK C H, KO R C. Characterization of endonuclease activity from excretory/secretory products of a parasitic nematode, *Trichinella spiralis*[J]. *Eur J Biochem*, 1999, 260(2): 477-481.
- [9] BIEN J, NÄREAHO A, VARMANEN P, et al. Comparative analysis of excretory-secretory antigens of *Trichinella spiralis* and *Trichinella britovi* muscle larvae by two-dimensional difference gel electrophoresis and immunoblotting[J]. *Proteome Sci*, 2012, 10(1): 10.
- [10] VARELA-RAMIREZ A, ABENDROTH J, MEJIA A A, et al. Structure of acid deoxyribonuclease[J]. *Nucleic Acids Res*, 2017, 45(10): 6217-6227.
- [11] JASMER D P, KWAK D. Fusion and differentiation of murine C2C12 skeletal muscle cells that express *Trichinella spiralis* p43 protein[J]. *Exp Parasitol*, 2006, 112(2): 67-75.
- [12] VASSILATIS D K, DESPOMMIER D, MISEK D E, et al. Analysis of a 43-kDa glycoprotein from the intracellular parasitic nematode *Trichinella spiralis* [J]. *J Biol Chem*, 1992, 267(26): 18459-18465.
- [13] MACLEA K S, KRIESER R J, EASTMAN A. A family history of deoxyribonuclease II: surprises from *Trichinella spiralis* and *Burkholderia pseudomallei* [J]. *Gene*, 2003, 305(1): 1-12.