

• 药材与资源 •

冬虫夏草指标蛋白 IP4 的 cDNA 全长克隆及抗原决定簇预测研究

张 涵, 童芯铎[#], 王 芳, 王艺璇, 彭 成, 国锦琳*

成都中医药大学 中药材标准化教育部重点实验室 中药资源系统研究与开发利用国家重点实验室培育基地, 四川 成都 611137

摘要: 目的 克隆获取冬虫夏草指标蛋白氰酸酶(IP4)的cDNA全长及对其抗原表位预测。方法 结合课题组前期采用蛋白组学方法筛选得到的冬虫夏草指标性蛋白IP4和冬虫夏草转录组数据,挖掘IP4 mRNA序列,并克隆IP4 cDNA全长,通过NCBI结构分析服务系统对IP4进行保守区和功能区分析,利用DNAMAN 6.0和DNASar软件分别进行序列比对和抗原表位分析。结果 从转录组数据中挖掘到IP4的2个可变剪接体,属于内含子保留的可变剪接。RT-PCR结果显示,IP4 cDNA全长465 bp,编码154 aa(氨基酸)。IP4蛋白的保守区和主要催化活性位点位于87~154 aa,该区域为氰酸酶家族保守核心区域和二聚体结合区域,DNAMAN 6.0分析结果表明1~39 aa和55~81 aa区具有较高种属特异性,DNASar预测IP4蛋白的抗原表位区主要分布于25~90 aa。结论 最终确定IP4的25~39 aa和55~81 aa为最佳抗原特异区,为后续规模化制备IP4蛋白特异抗原区及免疫法鉴定冬虫夏草奠定基础。

关键词: 冬虫夏草; 指标蛋白; cDNA克隆; DNASar; 抗原区分析**中图分类号:** R282.12 **文献标志码:** A **文章编号:** 0253-2670(2018)20-4864-06**DOI:** 10.7501/j.issn.0253-2670.2018.20.024cDNA cloning and prediction of antigenic determinants of protein marker IP4 in *Ophiocordyceps sinensis*

ZHANG Han, TONG Xin-xin, WANG Fang, WANG Yi-xuan, PENG Cheng, GUO Jin-lin

State Key Laboratory Breeding Base of TCM Resources Research and Development, Key Laboratory of Standardization of Chinese Herbal Medicine, Chengdu University of Traditional Chinese Medicine of Ministry of Education, Chengdu 611137, China

Abstract: **Objective** To clone the full-length of cDNA protein marker cyanase (IP4) of *Ophiocordyceps sinensis* and predict the antigenic sites. **Methods** The information of protein marker of *O. sinensis* was obtained by proteomics technology. The transcriptome database of *O. sinensis* and mycelium constructed in our lab were used to analyze IP4 unigenes and appropriate primers designed to amplify IP4, which was then cloned and sequenced. The conserved sequence of IP4, homology comparison, and the antigenic site were analyzed by DNAMAN6.0 and DNASar software. **Results** Two alternative splice variants of IP4 were discovered from the transcriptome data and belonged to the invariant alternative splicing. The protein marker of *O. sinensis* was identified as cyanate hydratase, 465 bp, encoding 154 aa. The analysis of the conserved and functional regions showed that catalytic site and binding site mainly lied in C-terminal of IP4. Analysis of DNAMAN 6.0 showed that species 1—39 aa and 55—81 aa were of higher species specificity, and DNASar results showed that the epitope region of IP4 protein is distributed from 25 to 90 aa. **Conclusion** The optimal antigen region of IP4 lied in the region 25—39 and 55—81 aa, which lays a foundation for the subsequent large-scale preparation of IP4 protein and utilizing ELISA to identify authenticity of *O. sinensis*.

Key words: *Ophiocordyceps sinensis* (Berk.) Sacc.; protein marker; cDNA clone; DNASar; antigenic site analysis

冬虫夏草为麦角菌科冬虫夏草菌 *Ophiocordyceps sinensis* (Berk.) Sacc. 寄生在鳞翅目蝙蝠蛾科昆虫幼虫上形成的幼虫尸体及子座的复合体,属国家二级保护的名贵药材,与人参、鹿茸并列为中国的三大补品。冬虫夏草始载于《本草从新》,具有补肺益肾、止血、化痰的作用,用于久咳虚喘、

收稿日期: 2018-04-30

基金项目: 国家自然科学基金资助项目(30801522, 81373920); 四川省中医药管理局川产道地药材大品种项目; 成都市科技局科技惠民项目资助**作者简介:** 张 涵(1995—),女,在读硕士,研究方向为中药资源与鉴定。Tel: 18408271679 E-mail: 1340757919@qq.com***通信作者** 国锦琳(1976—),男,博士,教授,研究方向为名贵中药资源与鉴定。Tel: 13981938406 E-mail: guo596@163.com**#并列第一作者** 童芯铎,博士,讲师,研究方向为中药资源与鉴定。

劳嗽咯血、阳痿遗精、腰膝酸痛^[1-3]。其药材需求量大,由于生长区域少,目前无成熟的人工培植技术,市场供求严重失衡。加之由于全球变暖、雪线上移、蝙蝠蛾大量死亡,其产量逐年减少,资源缺乏,故极为珍贵,价格连年上涨^[4];同时因为全世界已报道的虫草属真菌有400多种,我国有60多种,可寄生9个目昆虫的幼虫、蛹和成虫,但《中国药典》收载的冬虫夏草的寄主是产于我国青藏高原特殊高寒生境的蝠蛾幼虫^[5],因此市场掺伪、假冒现象日益严重,严重影响了冬虫夏草的流通安全和临床安全,目前无有效的方法进行鉴别。

课题组前期利用蛋白组学技术筛选得到冬虫夏草指标性蛋白IP4,经质谱鉴定为氰酸酶(cyanate hydratase, CH)^[6-8],并初步开发了ELISA的鉴定新方法,可进行定性、定量分析。但冬虫夏草中IP4含量低,从冬虫夏草中直接提取制备抗体难以大量获得,限制了试剂盒的规模化开发。基于此本实验克隆了IP4的cDNA全长并预测了其抗原决定簇,为下一步大量转化获得IP4蛋白奠定了基础。

1 材料与试剂

1.1 材料

新鲜冬虫夏草于2017年5月采集自四川省阿坝州小金县,由成都中医药大学国锦琳教授鉴定为麦角菌科冬虫夏草菌*Ophiocordyceps sinensis* (Berk.) Sacc.,留样在成都中医药大学中药材标准化教育部重点实验室。大肠杆菌*E. coli* DH5 α 菌株由实验室保存;克隆的载体购于北京全式金生物技术有限公司的pEasy-Blunt Simple Cloning Vector。

1.2 试剂

Trizol Reagent (Ambion 公司, 美国); RT reagent Kit with gDNA Eraser 逆转录试剂盒 (Takara 公司, 日本); TransStart FastPfu DNA Polymerase、高纯度质粒小提试剂盒、Trans2K Plus DNA Marker (北京全式金生物技术有限公司); 2 $\times$ Taq MasterMix (Dye)、琼脂糖凝胶DNA回收试剂盒(康为世纪生物科技有限公司); 其他实验试剂均为国产分析纯。

2 方法

2.1 冬虫夏草指标蛋白信息获得

前期课题组通过蛋白质组学方法筛选获得冬虫夏草指标性蛋白IP4,此蛋白表达丰度大,相对分子质量较小(15 000~25 000),相对稳定。进一步经MALDI-TOF-MS鉴定,测得肽段氨基酸序列

为“K. ALSFEQIAQHIGR”,覆盖率16%,经NCBI数据搜索确定为冬虫夏草氰酸酶(CH OS)同源蛋白,gi|751836131,相对分子质量为17 780,等电点(pI)是5.93^[6]。

2.2 指标蛋白基因的挖掘及差异表达分析

根据前期采用Illumina/Solexa HiSeq 2500高通量测序平台对冬虫夏草菌丝体(HH)和子实体(CS)转录组进行测序,每个样品设定2个重复。数据组装分析的结果采用2种方法挖掘指标蛋白基因序列:(1)根据搜索功能注释为CH的同源序列,将序列在NCBI中进行Blast X比对,分析序列的编码区及注释结果的可靠性;(2)从NCBI下载CH序列与转录组组装结果进行比对,搜索相似度高的序列^[9]。经前期数据分析得到unigenes,并且以FPKM值表示unigene的表达丰度。差异表达基因(DEGs)的筛选标准为FDR(false discovery rate)<0.05且 $\log_2|FC|\geq 1$,其中FC(fold change)表示2样品间表达量的比值。

2.3 冬虫夏草总RNA的提取与指标蛋白cDNA的制备

用Trizol试剂法^[10]提取新鲜冬虫夏草总RNA,采用微量分光光度法Thermo Nanodrop 2000及1%琼脂糖凝胶电泳检测总RNA质量,保证RNA样品质量浓度 ≥ 400 ng/ μ L, $A_{260}/A_{280}>1.8$ 方可进行下列实验。取2 μ g冬虫夏草总RNA进行反转录合成cDNA,反应条件为42 $^{\circ}$ C水浴2 min,转至冰上上放置2 min,37 $^{\circ}$ C孵育15 min,85 $^{\circ}$ C热激5 s,再将cDNA放置于4 $^{\circ}$ C短期保存(合成的cDNA需要长期保存时,于-20 $^{\circ}$ C或更低温度保存)。

2.4 IP4蛋白基因cDNA克隆及鉴定

以cDNA为模板,按照TransStart FastPfu DNA Polymerase使用说明书进行实验。根据转录组中提取的IP4 mRNA序列,利用Primer 6.0引物设计软件,设计一对特异性引物,由成都擎科梓熙生物技术有限公司合成。上游引物P1:5'-ATGTCT-CTCGTAACTCTTCAAACCAGCC-3',下游引物P2:5'-TCACCACTTCCCCCGCAGG G-3'。扩增条件:95 $^{\circ}$ C预变性5 min,95 $^{\circ}$ C变性10 s,60退火10 s,70 $^{\circ}$ C延伸25 s,32个循环,1%琼脂糖凝胶电泳检测PCR产物,置于自动凝胶成像仪上检测结果。而后用琼脂糖凝胶DNA回收试剂盒回收目的片段,操作方法按DNA回收试剂盒说明书进行。将回收的目的片段连接到

pEasy-Blunt Simple Cloning 载体上, 连接产物转化到大肠感受态细胞 *E. coli* DH5 α 中, 再将转化后的菌液涂板于含有 50 mg/mL Kanamycin 的平板上, 37 °C 过夜倒置培养。挑取单个菌落于含有 Kanamycin 的 LB 液体培养基, 200 r/min, 振荡培养 6 h, 重组质粒经菌落 PCR、酶切及重组子测序的方法鉴定后, 送至成都擎科梓熙生物技术有限公司。

2.5 IP4 序列分析

通过 NCBI 结构分析服务系统 (<http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>) 对氨基酸序列进行保守序列分析; 利用 DNAMAN 6.0 软件进行对核苷酸序列和氨基酸序列进行多重比对; 根据 SMART 结构分析服务系统^[11]和 EMBL-EBI^[12]的 ProScan 序列搜索 (<http://www.ebi.ac.uk/Tools/pfa/ipscan/>) 对氨基酸序列进行保守区域的分析和活性中心预测。

2.6 IP4 抗原区分析

利用 DNASTar 中 Protean 软件对指标蛋白的二级结构和抗原决定簇进行分析和预测, 用 Gamier-Robson^[13]、Chou-Fasman^[14]方法对 IP4 蛋白二级结构中存在的 α 螺旋、 β 折叠、 β 转角和无规卷曲进行预测。采用 Kyte-Doolittle^[15]方法分析氨基酸序列预测蛋白的亲水与疏水区域; 用 Emini 等^[16]方法对特定区域位于蛋白质的表面的可能性进行预测; 利用 Jameson-wolf^[17]方法分析抗原指数, 选

择编码亲水性较强、抗原指数较高且表面可及性好的区域为最佳抗原区。

3 结果与分析

3.1 冬虫夏草转录组数据中蛋白核酸序列挖掘及差异表达分析

前期采用 Illumina/Solexa HiSeq 2500 高通量测序平台对冬虫夏草 HH 和 CS 转录组进行测序分析, NCBI 注册编号为 MH100886, 从转录组数据中挖掘 Cyanate hydratase (IP4) 核酸序列, 结果显示 IP4 具有 2 个可变剪接转录本, 序列结构如图 1 所示, 分别为 IP4 和 IP4-intron, IP4 基因 mRNA 全长为 465 bp, 编码 155 aa, IP4-intron mRNA 全长 517 bp, 于 21 bp 处连续插入 52 bp 片段, 该插入片段为内含子, 由于可变剪接体为内含子保留引起的可变剪接, 会影响其编码蛋白的结构域, 继而影响其功能。差异表达量分析结果如表 1 所示, 子实体中 IP4 表达量高于菌丝体, 推测子实体抗逆性增强, 基础代谢更活跃。

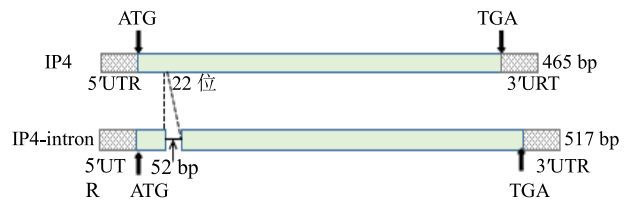


图 1 IP4 基因 2 个可变剪接体结构示意图

Fig. 1 Diagram of IP4 gene structure and variants

表 1 IP4 转录子差异表达量分析

Table 1 Transcript expression pattern of IP4 gene

基因名称	Transcript ID	CS-1_RPKM	CS-2_RPKM	HH-1_FPKM	HH-1_FPKM	log ₂ (CS_FPKM/HH_FPKM)
IP4 (CH)	CL1526 Contig1	3.25	4.36	2.08	2.49	0.88

HH-冬虫夏草菌丝体 CS-冬虫夏草子实体

HH- *O. sinensis* mycelium CS-*O. sinensis* fruiting body

3.2 IP4 蛋白基因 cDNA 克隆及序列分析

根据冬虫夏草转录组数据中挖掘的 IP4 序列, 并在序列头尾设计引物, 以 cDNA 为模板, 对 IP4 蛋白基因进行 PCR 扩增, 扩增产物经琼脂糖凝胶电泳检测, 如图 2 所示, 回收后, 将目的片段克隆到 pEASY 载体上, 鉴定、测序, 经 Blast 2 进行氨基酸序列同源比对, 结果显示克隆序列与转录组挖掘序列完全一致, GenBank 注册编号为 MH100886。利用 NCBI 数据库 (<http://www.ncbi.nlm.nih.gov>) 进行 blast 核酸比对, 显示 IP4 含有氰酸酶家族核心保守区。

3.3 IP4 序列分析

依据 NCBI 结构分析系统、EMBL-EBI 的 ProScan 和 SMART 服务序列检索。结果如图 3 所示, IP4 含有保守序列, 位于 87~154 aa, 序列中含二聚体结合区域, 还含有催化活性位点 (100 aa, 126~127 aa), 该区域是氰酸酶家族保守的核心结构域。IP4 蛋白参与冬虫夏草的氰酸盐代谢过程, 与 GO 分析的功能注释结果一致, 具有克服环境氰酸盐毒性功能, 同时为冬虫夏草生长提供氮源。

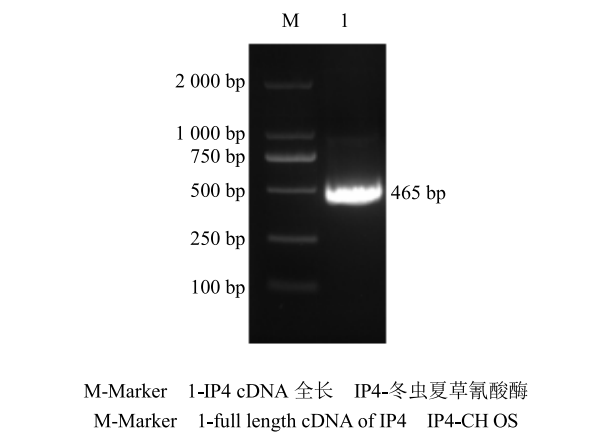


图 2 IP4 基因 PCR 扩增 1%琼脂糖凝胶电泳图
Fig. 2 1% Agar gel electrophoresis map of amplified IP4

从冬虫夏草转录组中挖掘到的 IP4 氨基酸序列，从 NCBI 数据库中提取其他物种的 IP4 同源序列分别为 3 条亲缘关系较近的虫草菌属的 *Ophiocordyceps unilateralis* (XA68_11971)，*Cordyceps militaris* (XM_006673595.1)，*Hirsutella minnesotensis* 3608 (KJZ74399.1) 及 2 条亲缘关系较远的青霉菌属的 *Penicillium italicum* (KGO65543.1) 和曲霉属 *Aspergillus niger* (XP_001397234.1)，利用 DNAMAN 6.0 软件对保守区 (>100 bp) 外的区域进行比对，结果 (图 4) 显示氨基酸序列相似率为 63.19%，为提高抗原特异性，故选择特异性高的区域，分别为肽段 1 (1~39 aa) 和肽段 2 (55~81 aa) 为抗原备选区。

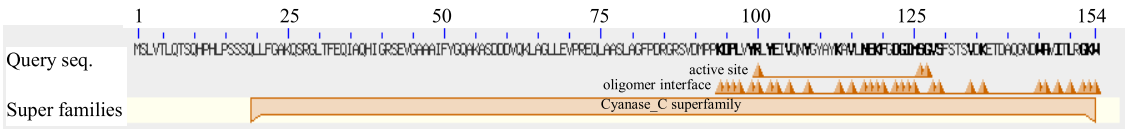


图 3 IP4 保守序列及反应活性位点分析
Fig. 3 Conserved domains and active site analysis of IP4



图 4 IP4 同源氨基酸序列比对结果以及所选肽段在序列上的位置
Fig. 4 Alignment analysis of IP4 amino acid sequence with DNAMAN 6.0 software and sites of specific peptide of *O. sinensis*

3.4 IP4 蛋白抗原区分析

DNAStar Protean 软件分析 IP4 蛋白分析结果如图 5 所示, 经 Gamier-Robson、Chou-Fasman、Kyte-Doolittle、Emini、Jameson-wolf 等方法进行综合分析, 结果显示, α 螺旋、 β 折叠, 分别主要分布在 22~8 aa 和 65 aa 之后, 且分布不均匀; β 转角与无规卷曲结构为凸出结构, 常出现在蛋白质的

抗原表面, 在 IP4 蛋白中, 转角较多且均匀分布, 无规则卷曲结构相对较少, 主要分布于 55 aa 之前和 135 aa 之后; 表面可及性、柔韧性、亲水性及抗原性指数高的区域主要分布于 25~100 aa。结合 IP4 氨基酸序列比对结果^[18-19]。因此, 最终选取 IP4 蛋白 25~39 aa 和 55~81 aa 区域作为最佳抗原区进行下一步克隆和表达。

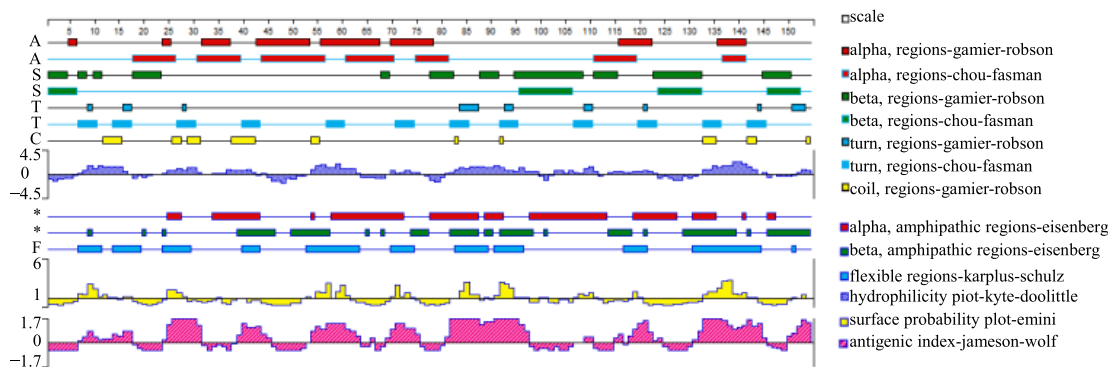


图 5 IP4 蛋白序列及抗原区分析
Fig. 5 Analysis of IP4 protein sequence and antigenic region

4 讨论

本研究前期通过蛋白组法筛选、鉴定得到 22 个冬虫夏草特异蛋白^[7-8], 选择含量高、特异性高、相对分子质量较小(相对稳定)的 IP4 蛋白为指标蛋白, 为建立免疫检测法实现对冬虫夏草进行鉴定打下基础。本实验进而结合冬虫夏草转录组数据, 挖掘 IP4 转录本序列, 有趣的是, 转录组数据中提取到 2 个 IP4 可变转剪接体, 分别为 IP4 和 IP4-intron。IP4 可变剪接属于内含子保留的类型, 已有研究发现, 内含子保留引起的可变剪接广泛存在于冬虫夏草转录组中^[20]。IP4 的功能在大肠杆菌和蓝藻中已较为深入研究, 功能为催化氰酸盐分解为铵和二氧化碳, 除去了有毒的氰酸盐的同时, 也为细菌提供氮源和碳源^[18-19,21]。转录组差异表达分析结果表明冬虫夏草子实体阶段 IP4 表达量高于菌丝体阶段, 推测冬虫夏草子实体阶段抗逆性增强, 基础代谢更为活跃, 与转录组数据分析结果一致。

本实验对 IP4 进行保守区和功能区分析, 发现其含 CH 家族核心保守区分布于 C 端第 87~154 aa (>100 bp 区域), 再通过 DNAStar 软件进行分析, 参考亲水性、抗原性、柔韧性、表面可及性以及二级结构等^[22-25], 对其特异性抗原表位进行预测分析, 选择抗原性强、特异性高的区域为最佳抗原特异区。

综上所述, 本研究基于前期蛋白组学方法确定了冬虫夏草指标蛋白 IP4, 利用转录组数据提取 IP4 蛋白转录本发现该蛋白存在 2 个可变剪接体, 继而首次克隆了冬虫夏草 IP4 cDNA 全长, 并对氨基酸序列功能区、保守区及抗原表位进行分析, 为后续以 IP4 抗原区规模制备以及免疫检测方法在中药质量鉴定领域的应用打下基础。

参考文献

- [1] 中国药典 [S]. 一部. 2015.
- [2] 吴仪洛. 本草从新 [M]. 北京: 中国中医药出版社, 2013.
- [3] 杨 潼. 中国动物志 [M]. 北京: 科学出版社, 2015.
- [4] Dong C, Guo S, Wang W, et al. Cordyceps industry in China [J]. *Mycology*, 2015, 6(2): 1-9.
- [5] 刘 飞, 伍晓丽, 尹定华, 等. 冬虫夏草寄主昆虫的种类和分布研究概况 [J]. *重庆中草药研究*, 2006(1): 47-50.
- [6] 邱 乙. 基于冬虫夏草指标性蛋白质 IP4 的 ELISA 定量评价新方法研究 [D]. 成都: 成都中医药大学, 2015.
- [7] 童芯铤, 王艺璇, 陈 璐, 等. 双向电泳技术筛选鉴定冬虫夏草中指标性蛋白质研究 [J]. *中国现代中药*, 2017, 19(1): 26-30.
- [8] Tong X, Wang Y, Xue Z, et al. Proteomic identification of marker proteins and its application to authenticate *Ophiocordyceps sinensis* [J]. *Biotechnology*, 2018, 8: 246-256.
- [9] 黄羽琪, 童芯铤, 陶 向, 等. 基于转录组测序的冬虫

- 夏草虫草素生物合成研究 [J]. 中草药, 2017, 48(19): 4044-4050.
- [10] Chomczynski P, Sacchi N. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol- chloroform extraction [J]. *Anal Biochem*, 1987, 162(1): 156-159.
- [11] Letunic I, Copley R R, Pils B, *et al*. SMART 5: Domains in the context of genomes and networks [J]. *Nucl Acid Res*, 2006, 34: 257-260.
- [12] McWilliam H, Li W, Uludag M, *et al*. Analysis tool web services from the EMBL-EBI [J]. *Nucl Acid Res*, 2013, 41: W597-W600.
- [13] Garnier J, Osguthorpe D J, Robson B. Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins [J]. *J Mol Biol*, 1978, 120(1): 97-120.
- [14] Jr P P, Fasman G D. *Prediction of Protein Structure and the Principles of Protein Conformation* [M]. New York: Springer: 1989.
- [15] Kyte J, Doolittle R F. A simple method for displaying the hydropathic character of a protein [J]. *J Mol Biol*, 1982, 157(1): 105-132.
- [16] Emini E A, Hughes J V, Perlow D S, *et al*. Induction of hepatitis A virus-neutralizing antibody by a virus-specific synthetic peptide [J]. *J Virol*, 1985, 55(3): 836-839.
- [17] Jameson B A, Wolf H. The antigenic index: A novel algorithm for predicting antigenic determinants [J]. *Comput Appl Biosci*, 1988, 4(1): 181-186.
- [18] Ranjan B, Pillai S, Permaul K, *et al*. Expression of a novel recombinant cyanate hydratase (rTl-Cyn) in *Pichia pastoris*, characteristics and applicability in the detoxification of cyanate [J]. *Bioresource Technol*, 2017, 238(4): 582-588.
- [19] 钱 丹. 拟南芥氰酸盐裂解酶 (AtCYN) 生化分析及功能研究 [D]. 北京: 北京大学, 2009.
- [20] Liu Z Q, Lin S, Baker P J, *et al*. Transcriptome sequencing and analysis of the entomopathogenic fungus *Hirsutella sinensis* isolated from *Ophiocordyceps sinensis* [J]. *BMC Genomics*, 2015, 16(1): 106-123
- [21] Sexton A C, Howlett B J. Characterisation of a cyanide hydratase gene in the phytopathogenic fungus *Leptosphaeria maculans* [J]. *Mol General Genetics Mgg*, 2000, 263(3): 463-470.
- [22] 李江英, 白雪娟, 梁 艳, 等. 用 DNASTar 软件预测 Rv1410c 结核分枝杆菌蛋白抗原表位 [J]. 细胞与分子免疫学杂志, 2015, 31(4): 474-477.
- [23] 孙沫逸, 孙文斌, 陈 兴, 等. p12 蛋白抗原决定簇短肽的预测 [J]. 临床口腔医学杂志, 2003, 19(3): 133-135.
- [24] 胡文龙, 唐博恒, 任瑞文, 等. 黄热病毒特异性抗原片段的筛选与鉴定 [J]. 中国预防医学杂志, 2013, 14(2): 114-117.
- [25] 任瑞文, 唐博恒, 骆康杰, 等. 西尼罗病毒特异性抗原片段筛选及 ELISA 方法的研究 [J]. 中国预防医学杂志, 2013, 1(5): 342-345.