

一种苦荞主要过敏原基因 cDNA 的克隆及序列分析

侯晓军¹⁾, 畅文军¹⁾, 陈立钊²⁾, 张政¹⁾, 刘知学³⁾, 王转花¹⁾*

(¹⁾山西大学生命科学与技术学院生物技术研究所,太原 030006; ²⁾燕山大学环境与化学工程系,秦皇岛 066004;

³⁾中国科学院上海生命科学院生物化学与细胞生物学研究所,上海 200031)

摘要 为了获得苦荞中主要过敏原的 cDNA 和由此推导的蛋白质序列,分析其结构特点,以苦荞幼根根尖为材料,提取总 RNA 并反转录 mRNA 为 cDNA 第一链.通过 RT-PCR、3 RACE、基因克隆及序列测定,获得一种苦荞主要过敏蛋白基因的 cDNA 片段(GenBank 登录号为 AY044918).该 cDNA 片段由 768 bp 组成,包括 3 端非编码区 180 个 bp,开放阅读框 588 bp.可编码一个由 195 个氨基酸残基组成的功能蛋白及一个终止密码.苦荞主要过敏原基因与甜荞 22 kD 过敏蛋白、豆球蛋白的核苷酸序列分别有 95%和 93%的同源性.其推导的氨基酸序列与甜荞球蛋白、刀豆蛋白、甜橙柠檬素分别有 93%、83%和 57%的同源性.该过敏蛋白 183~188 位氨基酸残基 KEEKE 在多数不同过敏原中均存在,推测可能为其中的抗原决定簇序列.

关键词 苦荞, 过敏蛋白, 3-RACE, 克隆与序列分析

中图分类号 Q78,R392

Cloning and Sequencing of Major Allergenic Gene cDNA from Tartary Buckwheat

HOU Xiaojun¹⁾, CHANG Wenjun¹⁾, CHEN Lichao²⁾, ZHANG Zheng¹⁾, LIU Zhixue³⁾, WANG Zhuanghua¹⁾*

(¹⁾ College of Life Science and Technology, Shanxi University, Taiyuan 030006, China; ²⁾ Department of

Environmental and Chemical Engineering, Yanshan University, Qinhuangdao 066004, China; ³⁾ Institute of Biochemistry and Cell Biology,

Shanghai Institute for Biological Sciences, Chinese Academy of Sciences, Shanghai 200031, China)

Abstract The cDNA sequence of allergenic gene expresses a kind of 22 kD protein in tartary buckwheat (TB), was obtained using RT-PCR and 3-rapid amplification of cDNA ends (3 RACE). The cDNA sequence coding for this allergenic protein, consisted of 768 nucleotides, including 588 nucleotides in the open reading frame (ORF) and 180 nucleotides in the 3' untranslated region. The ORF encoded a functional protein of 195 amino acids. Comparing this allergy from TB with those of common buckwheat (CB), it showed that this allergy in TB shared 95% and 93% nucleotide sequence homology with allergenic protein and legumin-like protein in CB, respectively. The protein sequence of TB allergenic protein shared 93% and 83% homology with those of globulin protein in CB and sword bean protein, respectively. The conserved sequence KEEKE at position 183-188 was probably considered to be the antigen determinant sequence similar to that of other allergenic proteins found in food and pollen.

Key words *Fagopyrum tataricum* (tartary buckwheat), allergenic protein, 3-RACE, cloning and sequencing

荞麦属于蓼科植物,其栽培品种为甜荞麦(*Fagopyrum esculentum*, Common buckwheat, 简称 CB)和苦荞麦(*Fagopyrum tataricum*, Tartary buckwheat, 简称 TB).前者在亚洲、欧洲和拉美等国家栽培较广,后者在我国西南和一些北方高寒地区种植,食用较多^[1].荞麦种子中的蛋白质含量可达 15%,富含多种必需氨基酸,具有很高的营养价值和药用价值^[2,3].特别是其中的类黄酮等生物活性物质能有

效预防糖尿病、高血压和风湿性关节炎等疾病^[4-6],

收稿日期:2002-09-23,接受日期:2002-12-16
山西省科技厅攻关项目(011001)

*联系人 Tel:(0351)7010599;E-mail:zhwang@sxu.edu.cn
侯晓军,女,1977年8月生,硕士生

Received:September 23,2002;Accepted:December 16,2002

Supported by the Key Program of Science and Technology Committee(011001) of Shanxi Province

* Corresponding author Tel:(0351)7010599;E-mail:zhwang@sxu.edu.cn

受到医学界的普遍关注. 荞麦中也含有许多过敏性成份, 使一些接触或食用它的人产生过敏性症状. 1909年 Smith^[7]首次报道了甜荞麦的过敏性. 近年来, 国外对甜荞麦过敏原的研究取得了长足的进展^[8~11]. Nakamura 和 Urisu 的研究揭示其病理机制为 IgE 介导的 I 型速发性过敏反应^[12,13]. 目前国内对苦荞过敏性的研究尚属少见. 我们曾首次从中国黑苦荞中纯化出一种分子量为 22 kD 的球蛋白(称作 TB22)^[14], 采用 Western 印迹及 ELISA 分析显示, 苦荞球蛋白中分子量在 19~70 kD 的部分均具有程度不同的 IgE 结合活性, 其中又以分子量约 22 kD 的蛋白质结合能力较强.

为揭示苦荞麦中过敏蛋白的致敏机理, 并克隆出一种植物过敏原家族的新基因, 本研究以纯化的苦荞 22 kD 蛋白的部分氨基酸序列设计简并性引物, 采用 RT-PCR 及 3 RACE 技术扩增, 克隆苦荞过敏原基因片段并分析其核苷酸序列以及氨基酸序列的同源性. 以为深入研究苦荞过敏蛋白等功能蛋白奠定基础.

1 材料和方法

1.1 材料

苦荞 (*Fagopyrum tataricum*) 为黔威苦荞, 当年收获. 大肠杆菌 DH5 为本室保存. AMV-Rtase, pGEM-T Easy 载体, T4 DNA 连接酶, RNase-Free DNase 为 Promega 公司产品, *Taq* DNA 聚合酶, DNA 分子量标准购自 Sangon 公司. 小量胶回收试剂盒及质粒纯化试剂盒为上海华舜生物公司产品. 3 RACE 试剂盒为 GibcoBRL 公司产品. 其他为国产分析纯试剂.

1.2 方法

1.2.1 总 RNA 的提取及 cDNA 第一链的合成 将苦荞种子洗净, 浸润 4 h, 室温培养 2 d 后用液氮研磨新鲜根尖, 参考植物材料总 RNA 的提取方法, 采用含 50 mmol/L EDTA, 500 mmol/L NaCl, 10 mmol/L 巯基乙醇及 1% SDS 的 Tris-HCl 缓冲液 (pH 8.0) 提取苦荞总 RNA. 以苦荞总 RNA 为模板, 以 Oligo (dT) 为引物, 由 AMV-Rtase 催化合成 cDNA 第一链用于苦荞过敏原基因的克隆.

1.2.2 RT-PCR 及 3-RACE 根据我们先前纯化的 TB22 蛋白的 N 端氨基酸序列并参考 Nair 等^[15]报道的甜荞过敏蛋白的氨基酸序列设计兼并性引物. 以总 cDNA 为模板, RT-PCR 扩增后得到该蛋白的编码区片段, 测序后设计特异性引物, 利用 3-RACE 方法结合巢式 PCR 扩增基因的 3 末端. 反应条件: 反

应体系为 25 μ l, 采用热启动方式, 两步循环. 第一循环 5 个周期, 第二循环 30 个周期. 退火温度根据各对引物组成而定. 1.5% 琼脂糖凝胶电泳检测扩增结果. 实验所用引物分别为:

引物 : 5'-GGATTGGARCAA GCVTTCTGYAACCT-3'

引物 : 5'-AACWATRGA GAAACGYTCCCTCTCCT-3'

其中, R 为 (A + G), V 为 (A + G + C), Y 为 (C + T), W 为 (A + T)

引物 (AP) : 5'-GGCCACGC GTC GACTA GTACTTTTTTTTTTTTTTTTTTTF3

引物 : 5'-AACGCCATAACCA GTCCGATTF3

引物 (AUAP) : 5'-GGCCACGC GTC GACTA GTAC3

1.2.3 PCR 产物的克隆及 DNA 序列分析 用 1.5% 琼脂糖凝胶电泳鉴定 PCR 产物, 将目的片段回收, 与质粒载体 pGEM -T Easy Vector 连接并转化大肠杆菌 DH5 感受态细胞, 在含有氨苄青霉素, IPTG 和 X-Gal 的 LB 平板上挑取白色单克隆, 筛选重组质粒. 通过酶切和 DNA 测序鉴定. 核苷酸及氨基酸序列同源性分析使用 BLAST (Ver 2.00) 软件在 GenBank 数据库中进行. 氨基酸序列与已知过敏原比较在 Database of Food Allergens in AgMoBiol 中进行.

2 结果

2.1 苦荞总 RNA 的获得

苦荞总 RNA 如图 1 所示, 泳道 1、2 均能够清晰地看到 18 S rRNA 和 28 S rRNA 及 mRNA, 且 18 S rRNA 和 28 S rRNA 两者比例符合 2:1 的特点, 说明提取过程中无 RNA 降解现象, 提取方法效果较好.

2.2 目的基因编码区的扩增与克隆

用 Oligo (dT) 为引物, 合成的 cDNA 为模板, 以简并性引物 和 进行 RT-PCR 扩增. 扩增结果如图 2 所示, 在 600 bp 处的条带与预期的目的片段大小吻合, 回收目的片段, 并连入 pGEM -T Easy 载体, 得到重组质粒. RT-PCR 扩增产物的序列测定结果显示, 该基因片段长度为 585 bp. Internet 搜寻比较表明, 所克隆的目的基因片段与 Nair^[15]报道的过敏性蛋白的结构基因序列有 95% 的同源性, 说明此克隆产物是苦荞中主要过敏蛋白基因的编码区.

2.3 目的基因 3 端的克隆及序列分析

1 2

— 28S rRNA
— 18S rRNA
— mRNA

Fig. 1 Electrophoresis of total RNA from tartary buckwheat
1 and 2: Total RNA

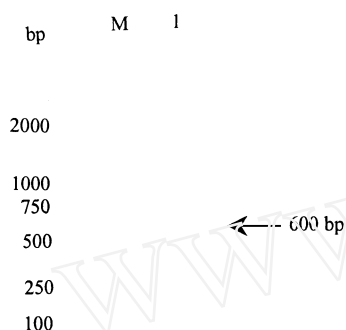


Fig. 2 The PCR fragment of allergenic protein gene coding
M:DNA marker, 1: PCR fragment

用 3 RACE 试剂盒中的 AP (adapter primer) 为引物,合成的 cDNA 为模板并根据测得的编码区序列设计特异性引物,采用巢式 PCR,先利用引物和 3 RACE 试剂盒中的另一引物 AUAP (Abridged Universal Amplification Primer) 进行 PCR,再以 PCR 产物为模板,用引物和 AUAP 进行扩增,得到了包括部分编码区及 3 端非编码区的基因片段,将产物纯化、连接并转化,筛选出阳性重组质粒, *Eco*R 酶切分析,得到大小约 400 bp 的片段. 测序结果显示,该片段确切长度为 404 bp,其中包括部分编码区碱基序列和一段 180 bp 的 3 末端非编码区. 将获得的 2 个核苷酸序列片段进行叠加分析,得到一长度为 768 bp 的基因序列 (Fig. 3). 包括编码区的 588 个碱基及 3 端的非编码区的 180 个碱基. 该基因已在 GenBank 中公布,登录号为 AY044918. 来源于苦荞中的过敏蛋白基因序列在国内外尚属首次报道.

在 GenBank 中用 BLAST 软件进行分析,表明此核苷酸序列与甜荞过敏性贮藏蛋白及豆球类蛋白的

同源性分别达到 95 %和 93 %.

2.4 TB22 过敏蛋白的氨基酸序列分析

由苦荞过敏蛋白基因核苷酸序列导出的氨基酸序列分析显示:该蛋白由 195 个氨基酸组成. 分子量约为 22 kD. 酸性氨基酸 (Asp + Glu) 23 个,碱性氨基酸 (Arg + Lys) 25 个. 与来自甜荞中的球蛋白,刀豆蛋白有 93 % ~ 83 %同源性,与甜橙柠檬素有 57 %同源性,与不同来源 (如:蓖麻、咖啡、木兰、葫芦、水稻) 的种子贮藏蛋白,11S 球蛋白 亚基前体,豆球蛋白前体均有 59 % ~ 50 %同源性 (Table 1). 在食物过敏原数据库中进行 TB22 蛋白与已知过敏原氨基酸序列的比较,发现连续相同的氨基酸最多有 5 ~ 6 个,分别为 EEAYK, EEEKEH 和 KEEEEE. 其中氨基酸序列中 183 ~ 188 位的多肽 KEEEEE 在不同过敏原中都有存在,推测可能为潜在的抗原决定簇.

Table 1 Homologous comparison of amino acid sequence of TB 22 kD allergenic protein to 11S storage globulin and legumin-like protein

Protein source	Homology (%)
Allergenic protein, storage protein, legumin-like protein (<i>Fagopyrum esculentum</i>)	83 ~ 93
Citrus-sweet orange (<i>Citrus sinensis</i>)	57
Beta-globulin, legumin, storage protein upland cotton (<i>Gossypium hirsutum</i>)	49 ~ 58
Seed storage protein, legumin (<i>Ricinus communis</i>)	51 ~ 56
11S storage globulin, 11S storage protein-coffee (<i>Coffea arabica</i>)	58 ~ 59
Gobulin, legumin (<i>Magnolia salicifolia</i>)	55 ~ 56
11S globulin (<i>Sesamum indicum</i>)	51
11S globulin oat, 12S seed storage protein (<i>Avena sativa</i>)	48 ~ 52
11S globulin beta subunit- cucurbit (<i>Cucurbita pepo</i>)	51
Gutelin-rice (<i>Oryza sativa</i>)	51

3 讨论

威胁人类健康的许多过敏性疾病 (包括药物过敏、食物过敏、环境过敏等) 的病理机制是一个非常复杂的过程. 对于一般的食物过敏许多研究认为与其中的过敏蛋白有关,因此,寻找可能引起过敏反应的过敏蛋白及阐明其抗原决定簇氨基酸序列等,对于深入了解过敏机制及定向修饰或改造过敏原是非常必要的. 近年来,荞麦食品在欧美和日本等国越来越受到消费者的欢迎,但一些人群食用荞麦食品后引起的皮肤搔痒、头晕、哮喘等过敏症状引起了国外医学和营养学界的普遍关注,成为食物过敏原研

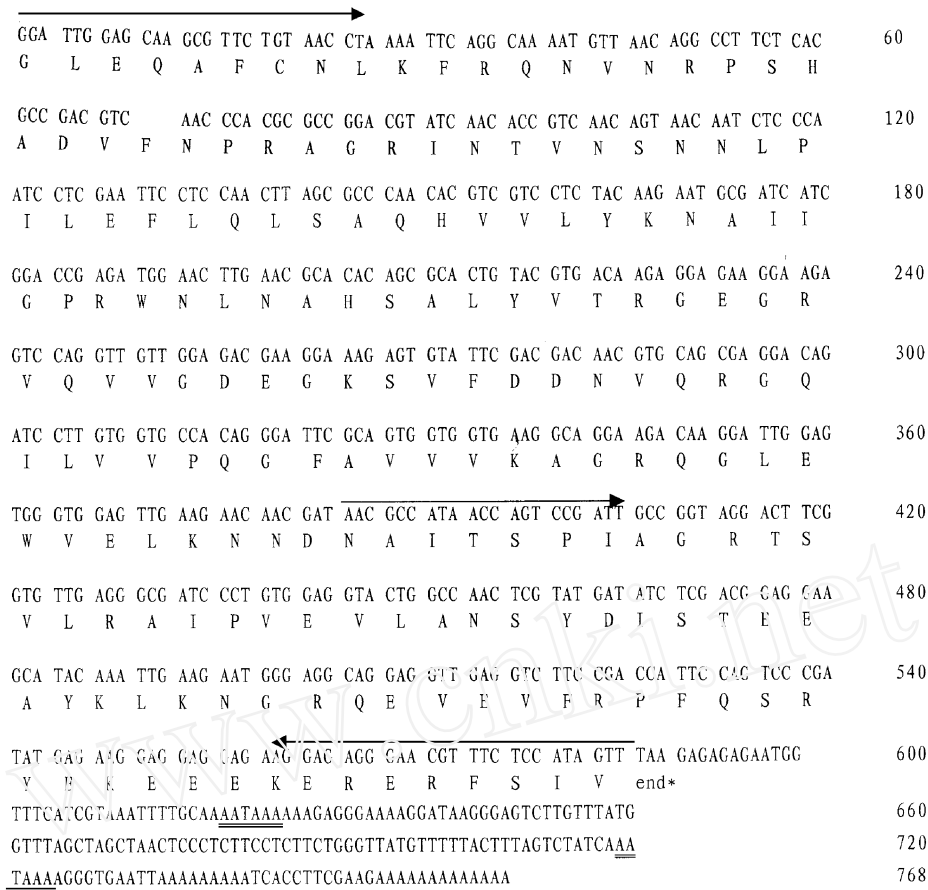


Fig. 3 Nucleotide and deduced amino acid sequences of a cDNA encoding for a tartary buckwheat major allergenic protein. Arrows refer to the four designed primers. The poly adenylation signals are double line underlined (=) and the stop codon is indicated by an asterisk (*)

究的热点课题。1994年, Urisu等^[13]报道甜荞麦中的过敏蛋白氨基酸序列与双子叶物种(花园豆、大豆、旱地棉、卷耳水芹、油菜、南瓜、葫芦、向日葵)11S球蛋白链具有37.9%~48.3%同源性,与单子叶物种(水稻、燕麦)类似蛋白质有51.7%同源性。这些蛋白质与贮藏蛋白的亚基寡聚物相对应。Nair等^[15]1999年测定了甜荞麦过敏蛋白的N端氨基酸序列,分析与其它物种来源的过敏原同源性时发现,甜荞过敏蛋白与单、双子叶物种贮藏蛋白的氨基酸有很大同源性,其中与水稻、豌豆、燕麦、向日葵11S球蛋白有65%~75%同源性。

本研究以我国特有的苦荞麦为材料,采用RT-PCR和3'RACE技术,首次克隆了其中的主要过敏蛋白(TB22)的结构基因部分和3'端序列。其中编码区的588bp可为195个氨基酸编码,3'端180bp包含一段polyA序列(AT约70%),推测的TB22蛋白其氨基酸序列与甜荞主要过敏蛋白相比较,发现有13个氨基酸不同,其中多为碱性氨基酸。另外

160位的Lys在苦荞中被Glu取代,而181位的Asp被Tyr取代。并且增加了185~188四个氨基酸(E-E-K-E),两者氨基酸序列同源性达95%。与甜荞球蛋白、刀豆蛋白、甜橙柠檬素的氨基酸序列分别有93%、83%和57%的同源性。TB22的核苷酸序列与甜荞主要过敏蛋白、豆球蛋白cDNA的核苷酸分别有95%和93%的同源性。按照文献[16]提供的方法在食物过敏原数据库中进行TB22过敏蛋白与目前已知氨基酸序列的198种过敏原蛋白相比较,发现连续相同的氨基酸最多有5~6个,其中183~188位上有在不同过敏原中存在的KEEKE氨基酸序列。我们认为,从食物过敏原数据库分析得到的结果可能是TB22的部分抗原决定簇序列,但也不排除该过敏蛋白中存在不连续抗原决定簇的可能。本实验得到的TB22过敏蛋白基因cDNA可为下一步基因表达及其蛋白质的结构与功能研究奠定基础。核苷酸和氨基酸序列分析显示其cDNA与多种植物过敏原具有较高的同源性,为植物过敏蛋白家

族提供了新内容。

致谢：本研究得到中科院生物化学与细胞生物学研究所戚正武院士的指导，在此致以诚挚的感谢。

参考文献 (References)

- 1 林汝法. 中国荞麦. 北京: 中国农业出版社(Lin R F. *Buckwheat in China*. Beijing: Agriculture Publishing House, China), 1994: 226 ~ 243
- 2 Hong J H, Ikeda K, Kreft I, Yasumoto K. Near-infrared diffuse reflectance spectroscopic analysis of the amounts of moisture, protein, starch, amylose, and tannin in buckwheat flours. *J Nutr Sci Vitaminol (Tokyo)*, 1996, **42** (4): 359 ~ 366
- 3 Wieslander G. Review on buckwheat allergy. *Allergy*, 1996, **51** (10): 661 ~ 665
- 4 Ren W, Qiao Z, Wang H, Zhu L, Zhang L, Lu Y, Cui Y, Zhang Z, Wang Z. Tartary buckwheat flavonoid activates caspase 3 and induces HL-60 cell apoptosis. *Methods Find Exp Clin Pharmacol*, 2001, **23** (8): 427 ~ 432
- 5 张政, 王转花, 林汝法, 金晓弟. 苦荞种子胰蛋白酶抑制剂的分离纯化及部分性质研究. 中国生物化学与分子生物学报 (Zhang Zheng, Wang Zhuar-hua, Lin Rurfa, Jin Xiao-di. Purification and some preperities of the trypsin inhibitor from tartary buckwheat seeds. *Chin J Biochem Mol Biol*), 1999, **15** (2): 247 ~ 251
- 6 Ihme N, Kieseewetter H, Jung F, Hoffmann K H, Birk A, Muller A, Gutzner K I. Leg oedema protection from a buckwheat herb tea in patients with chronic venous insufficiency: a single-centre, randomised, double-blind, placebo-controlled clinical trial. *Eur J Clin Pharmacol*, 1996, **50** (6): 443 ~ 447
- 7 Smith H L. Buckwheat poisoning with report of a case in man. *Arch Intern Med*, 1909, **3**: 350 ~ 359
- 8 Wieslander G, Norback D. Buckwheat allergy. *Allergy*, 2001, **56** (8): 703 ~ 704
- 9 Takahashi Y, Ichikawa S, Aihgara Y, Yokota S. Buckwheat allergy in 90 000 school children in Yokohama. *Averugi*, 1998, **47** (1): 26 ~ 33 (in Japanese)
- 10 Davidson A E, Passero M A, Settipan G A. Buckwheat induced anaphylaxis: a case report. *Ann Allergy*, 1992, **69** (5): 439 ~ 440
- 11 De Maat-Bleeker F, Stapel S O. Cross-reactivity between buckwheat and latex. *Allergy*, 1998, **53** (5): 538 ~ 539
- 12 Nakamura S, Yamaguchi M. Studies on buckwheat allergosis. Report 2: clinical investigation of 169 cases with the buckwheat allergosis gathered from the whole country of Japan. *Allerg Immunol*, 1974/5, **20/21** (4): 457 ~ 465
- 13 Urisu A, Kondo Y, Morita E. Identification of a major allergen of buckwheat seeds by immunoblotting methods. *Allergy Clin Immunol News*, 1994, **6**: 151 ~ 155
- 14 Wang Z H, Zhang Z, Wieslander G, Norback D, LI Y, Yang B, Lin R F. Purification and some properties of the protein with 22 kD in tartary buckwheat. *Fagopyrum*, 2000, **17**: 441 ~ 444
- 15 Nair A, Ohmoto T, Woo SH, Adachi T. A molecular-genetic approach for hypoallergenic buckwheat. *Fagopyrum*. 1999, **16**: 29 ~ 36
- 16 Metcalfe D D, Astwood J D, Townsend R, Townsend R, Sampon H A, Taylor S L, Fuchs R L. Assessment of the allergenic potential of foods derived from genetically engineered crop plants. *Crit Rev Food Sci Nutr*, 1996, **36**: S165 ~ S186