

- [8] 丁衬衬,周艳红,林凡云,等.转基因小麦田土壤细菌 16S rDNA V3 片段的 DGGE 分析[J]. 江苏农业学报,2011,27(1):66-70.
- [9] 邓先余,王智学,孙成波,等.3种水产病原弧菌(*Vibrio*)的 16S-23S rDNA 间区(IGSs)的克隆、测序与分析[J]. 2006,37(2):162-170.
- [10] Lee S K, WANG H Z, Law S H, et al. Analysis of the 16S-23S rDNA intergenic spacers (IGSs) of marine vibrios for species-specific signature DNA sequences[J]. Mar Pollut Bull, 2002, 44(5):412-420.
- [11] Buchan A, Alber M, Hodson R E. Strain-specific differentiation of environmental *Escherichia coli* isolates via denaturing gradient gel electrophoresis (DGGE) analysis of the 16S-23S intergenic spacer region[J]. FEMS Microbiol Ecol, 2001, 35(3):313-321.
- [12] 谢珍玉,胡超群,周永灿.溶藻弧菌 16S rDNA 与 23S rDNA 间的 ISR 多态性分析[A]//第六届世界华人虾蟹类养殖研讨会论文摘要集[C],广州,2008.北京:海洋出版社,2008.

论著

重组人 α -乳清白蛋白潜在致敏性的生物信息学预测

刘珊,汪会玲,冯晓莲,李晨汐,徐海滨

(国家食品安全风险评估中心 卫生部食品安全风险评估重点实验室,北京 100021)

摘要:目的 通过生物信息学分析预测牛乳腺生物反应器表达的重组人 α -乳清白蛋白(rhLA)的致敏性。方法 使用在线致敏原数据库(The Allergen Online Database)、致敏蛋白结构数据库(Structural Database of Allergenic Proteins)和食品安全致敏原数据库(Allergen Database for Food Safety),分析目标蛋白与已知致敏原的序列或结构相似性。结果 rhLA 致敏性的生物信息学分析结果显示,rhLA 与已知致敏原牛 α -乳清白蛋白、牛乳糖合酶 B 蛋白(致敏原 Bos d 4)、鸡溶菌酶 C(1,4- β -N 乙酰胞壁质酶 C、致敏原 Gal d IV、致敏原 Gal d 4)存在较高的序列或结构同源性。结论 rhLA 具有潜在的致敏可能性。

关键词:重组蛋白;人 α -乳清白蛋白;转基因;致敏性

中图分类号:Q512.1 文献标识码:A 文章编号:1004-8456(2013)03-0210-05

Preliminary assessment of recombinant human α -lactalbumin for allergenic potential

Liu Shan, Wang Huiling, Feng Xiaolian, Li Chenxi, Xu Haibin

(Key Laboratory of Food Safety Risk Assessment of Ministry of Health, China National Center for Food Safety Risk Assessment, Beijing 100021, China)

Abstract: Objective To predict the allergenic potential of recombinant human α -lactalbumin (rhLA) expressed in the milk of transgenic cloned cattle using bioinformatics analysis. **Methods** The Allergen Online Database, Allergen Database for Food Safety and Structural Database of Allergenic Proteins were used to determine the similarity of amino acid sequences and structures between the expressed protein and known allergens. **Results** There was a high level of similarity between rhLA and some known allergens, including bevine alpha lactalbumin, bovine lactose synthas B proteina (Bos d 4) and chicken lysozyme C (1, 4-beta-N-acetylmuramidase C, Gal d 4, Gal d IV). **Conclusion** RhLA has possible allergenic potential.

Key words: Recombinant protein; human α -lactalbumin; transgenic; allergenicity

α -乳清白蛋白(α -lactalbumin, α -LA)由乳腺上皮

细胞合成,分泌到所有的哺乳动物乳汁中,分子量为 14 kDa 左右。作为半乳糖苷转移酶的调节亚基,促进乳中乳糖的生成从而调控乳的渗透压。人乳清白蛋白是母乳中的主要成分, α -LA 的氨基酸组成中包含大量的人体营养所必需的氨基酸,对婴幼儿体格和神经系统的正常发育具有十分重要的作用。此外,据报道 α -LA 还具有抑菌和诱导肿瘤细胞凋亡的功能^[1-3]。通过转基因技术,在牛乳腺中特异表达人 α -LA,既可以提高牛乳中 α -LA 的含量,增强其营养

收稿日期:2013-02-21

基金项目:转基因生物新品种培育科技重大专项(2011ZX08011-005)

作者简介:刘珊 女 副研究员 研究方向为食品安全与毒理,转基因生物及其产品的食用安全
E-mail:liushan78@hotmail.com

通信作者:徐海滨 男 研究员 研究方向为食品安全风险评估
E-mail:hbxiu1231602@vip.sina.com

价值,又可为获得大量人 α -LA 提供原料来源。本研究评价的转人乳清白蛋白是由中国农业大学开发的转人乳清白蛋白转基因牛在牛奶中高效表达的外源蛋白。对于转基因生物表达的外源蛋白,首先必须考虑其安全性,在之前的研究中本研究室已经对该转基因牛生产的牛奶的食用安全性进行了初步评价,未发现该牛奶饲养的动物出现任何有生物学意义的改变^[4]。本文主要讨论该重组蛋白的致敏性,致敏性也是目前评价转基因生物表达的外源蛋白安全性的一个重要方面。2001年FAO/WHO转基因食品致敏性评价导则中将生物信息学分析作为评价蛋白质致敏性的一个重要手段^[5],本研究也将以此初步预测重组人 α -乳清白蛋白的致敏性。

1 材料与方法

1.1 致敏原数据库及其使用^[6]

致敏原数据库主要收录自然界中已知致敏蛋白的信息,目前国外已建立的主要变应原数据库以3个蛋白质数据库(GenPept、TrEMBL和DAD)和3个核酸数据库(GenBank、EMBL和DDBJ)为依托,对已知变应原相关信息,如物种来源、地域分布、蛋白质和核酸编码等进行收录、整理、分析。本研究选用了3个比较常用的致敏原数据库。

1.1.1 在线致敏原数据库

在线致敏原数据库是美国内布拉斯加大学建立的食物变应原相关数据库,可以进行序列搜索和识别蛋白质的交叉反应性,具备了对不同蛋白质的序列进行比较的功能。输入网址 <http://www.allergenonline.com/databasefasta.asp> 并打开链接,在序列输入框输入外源蛋白质英文名称(在外源蛋白质英文名称前用“>”引导),换行后输入外源蛋白质氨基酸全序列(氨基酸序列用大写单字母表示)。点击“Simple FASTA”进行全长比对;点击“Sliding 80mer”进行80个氨基酸片段比对;点击“8mer Exact Match”进行8个连续氨基酸比对。

1.1.2 致敏蛋白结构数据库

致敏蛋白结构数据库是美国于2001年发布且不断更新的、具有丰富计算工具(FSAST、BLAST)的变应原数据库,可对整条变应原序列或者多肽进行比较、计算,最终得出其差异或相似程度,用于蛋白质致敏性预测,变应原之间交叉反应的研究。输入网址 http://fermi.utmb.edu/SDAP/sdap_src.html 并打开链接,点击FAO/WHO Allergenicity Test,分

别输入外源蛋白质英文名称和氨基酸序列。点击“Full FASTA alignment”进行全长比对;点击“FASTA alignments for an 80 amino acids sliding window”进行80个氨基酸序列比对;比对条件设定为8,点击“Exact match for contiguous amino acids”进行8个连续氨基酸比对。

1.1.3 食品安全致敏原数据库

食品安全致敏原数据库是由日本国家健康科学研究生院新型食物和免疫化学部门联合推出的一个项目,包括食物致敏原和致敏性预测数据库,其中收录了所有的已知变应原和B细胞表位。输入网址 <http://allergen.nihs.go.jp/ADFS/> 并打开链接,选择Allergenicity Prediction:FAO/WHO method菜单,输入外源蛋白质氨基酸全序列(氨基酸序列用大写单字母表示),点击“Do a full fasta alignment”进行80个氨基酸片段比对;选择Allergenicity Prediction:Motif-based method菜单,设定比对条件为8,点击“Look for a small exact wordmatch”进行8个连续氨基酸比对。

1.2 比对结果判定

将rhLA的氨基酸序列(不包括前导序列)与3个致敏原数据库进行比对。通过全长比对(Full Fasta)获得E值,E值 $<3.9E-07$ 时,认为rhLA与已知致敏原具有较高的序列同源性^[7]。通过80个氨基酸片段序列相似性比对获得序列一致性,若序列一致性 $>35\%$,则认为rhLA与已知致敏原具有较高的序列同源性^[5]。通过连续8个氨基酸相同的精确比对获得匹配记录,若rhLA与已知致敏原具有完全匹配的连续8个氨基酸,则认为rhLA与已知致敏原具有较高的序列同源性^[8]。结构比对E值 <0.001 时,认为rhLA与已知致敏原具有较高的结构相似性^[9]。

2 结果

rhLA含有123个氨基酸残基(不包括前导序列),详细信息如下:

```
EQLTHCEV FQHLHDLHDYGGVSLPEWVCTAFH  
TSGYDTQAIVQNNST EYGLFQINHIWCHDDQNP  
SRNICNISCDHFLDDDLTDDIVCAHHILDHVGINYWL  
AHHALCSEHLDQWLCEHL.
```

2.1 在线致敏原数据库比对结果

在该数据库中通过全长比对、80个氨基酸片段序列相似性比对和连续8个氨基酸相同的精确比对结果见表1。

表1 在线致敏原数据库比对结果

Table 1 Search results against the Allergen Online Database

过敏原	物种	序列编号	序列长度	E 值	序列一致性	匹配记录
α-乳清白蛋白	牛	GB:CAA29664.1 GI:295774	142	5.8E-055	90.0%	39
α-乳清白蛋白;乳糖合酶 B 蛋白; 致敏原 Bos d 4(前体)	牛	SP:P00711.2 GI:125996	142	1.4E-054	88.7%	31
溶菌酶 C;1,4-β-N-乙酰胞壁质酶 C; 致敏原 Gal d IV;致敏原 Gal d 4(前体)	鸡	SP:P00698.1 GI:126608	147	7.6E-02	48.1%	—

注:—为未见匹配记录。

2.2 致敏蛋白结构数据库比对结果

在该数据库中通过全长比对、80 个氨基酸片段序列相似性比对和连续 8 个氨基酸相同的精确比对结果见表 2。

表2 致敏蛋白结构数据库比对结果

Table 2 Search results against the Structural Database of Allergenic Proteins

过敏原	物种	序列编号	序列长度	E 值	序列一致性	匹配记录
Bos d 4	牛	SP:Q28049 GI:75039460	123	2.8E-53	85.37%	39
Bos d 4	牛	GB:CAA29664 GI:295774	142	3.0E-53	85.37%	39
Bos 4	牛	SP:P00711 GI:125996	142	7.2E-53	84.55%	39
Gal d 4	鸡	SP:P00698 GI:126608	147	3.4E-20	39.84%	—

注:“—”为未见匹配记录。

2.3 食品安全致敏原数据库比对结果

在该数据库中通过全长比对、80 个氨基酸片段序列相似性比对、连续 8 个氨基酸相同的精确比对结果见表 3,结构相似性比对结果见表 4。

与食品安全致敏原数据库进行 Motif-based 三维结构比对获得 E 值,E 值 < 0.001 时,认为 rhLA 与已知致敏原具有较高的结构同源性。结果显示 rhLA 与 Motif ADFS 0065 具有相似的三维结构(E 值 = 7.3E-004)。含有 Motif ADFS 0065 的致敏原有以下 3 种(见表 4)。

2.4 同源性比对结果汇总

综合对上述 3 个致敏蛋白数据库的搜索结果,与 rhLA 具有较高的序列或结构同源性的致敏原主要为以下 4 种(见表 5)。

rhLA 致敏性的生物信息学分析结果显示:rhLA

表3 食品安全致敏原数据库比对结果

Table 3 Search results against the Allergen for Food Safety

过敏原	序列编号	序列长度	E 值	序列一致性	匹配记录
α-乳清白蛋白;乳糖合酶 B 蛋白;致敏原 Bos d 4(前体)	SP:P00711 GI:125996	123	3.9E-59	84.6%	5
溶菌酶 C;1,4-β-N-乙酰胞壁质酶 C;致敏原 Gal d IV;致敏原 Gal d 4(前体)	SP:P00698 GI:126608	101	6.8E-22	45.5%	—

注:“—”为未见匹配记录。

表4 含有 Motif ADFS 0065 的致敏原

Table 4 Allergens containing Motif ADFS 0065

过敏原	过敏原通用名称	物种	编号
Ani s 7	UA3 识别抗原(片段)	鲱鱼蠕虫	A9XBJ8;GI:119524036
Bos d 4	α-LA(乳糖合酶 B 蛋白)(Bos d 4)(前体)	牛	SP:P00711;SP:Q3T111; SP:Q95NE4;GI:162644; GI:295774;GI:162644
Gal d 4	溶菌酶 C(EC3.2.1.17)(1,4β-N 乙酰胞壁质酶 C)(致敏原 Gal d IV)(Gal d 4)(前体)	鸡	SP:P00698;SP:Q90884; GI:126608;GI:212279

表5 3 个致敏原数据库比对结果汇总

Table 5 Summary results against three Allergen Databases

过敏原	在线致敏原数据库			食品安全致敏原数据库			致敏蛋白结构数据库		
	全长	80 个氨基酸	连续 8 个氨基酸	80 个氨基酸	连续 8 个氨基酸	三维结构	全长	80 个氨基酸	连续 8 个氨基酸
牛 α-乳清白蛋白(GB:CAA29664.1 GI:295774)	+	+	+	-	-	+	+	+	+
牛 α-乳清白蛋白(乳糖合酶 B 蛋白、致敏原 Bos d 4)(SP:P00711 GI:125996)	+	+	+	+	+	+	+	+	+
鸡溶菌酶 C(1,4β-N 乙酰胞壁质酶 C、致敏原 Gal d IV、致敏原 Gal d 4)(SP:P00698 GI:126608)	-	+	-	+	+	+	+	+	-
牛 α-乳清白蛋白(SP:Q28049 GI:75039460 GB:CAA44927.1)	-	-	-	-	-	-	+	+	+

注: + 为阳性; - 为阴性。

与已知致敏原牛 α -乳清白蛋白、牛乳糖合酶 B 蛋白(致敏原 Bos d 4)、鸡溶菌酶 C(1,4- β -N 乙酰溶菌酶、致敏原 Gal d IV、致敏原 Gal d 4)存在较高的序列同源性或结构相似性。

3 讨论

食物过敏是全球关注的公共卫生问题之一,主要由食物中的蛋白质引起,轻者出现皮肤或肠道的刺激症状,重者可发生过敏性休克甚至死亡。常见的致敏食物有花生、大豆、牛奶、鸡蛋、鱼类、贝类、小麦和坚果 8 种,其中又以水产品、牛奶和鸡蛋最为常见^[10-11]。尽管目前评价致敏性的方法很多,但由于致敏性在生物体内的表现差异很大,许多方法的灵敏度和特异性尚有所欠缺,目前比较公认的有生物信息学分析、消化稳定性和血清学筛选试验等方法^[12-13]。

本研究选择的 3 个致敏原数据库是国际上比较公认、应用较多且信息比较全面的数据库。链接上述三个数据库,用全长比对、80 个氨基酸片段序列相似性比对、连续 8 个氨基酸相同的精确比对以及结构相似性比对 4 种方法与数据库中的已知致敏原进行比对。由于一直以来用全长比对分析 80 个以上氨基酸序列的相似性在结果判断时存在争议,2001 年 FAO/WHO 的转基因食物的致敏性评价导则中只将全长比对应用于少于 80 个氨基酸的序列。然而近年随着生物信息学研究和深入,以及蛋白质数据库和已知过敏原数据库的不断扩大,研究者通过分析大量的已知过敏原和非过敏原,提出若以 E 值 $<3.9E-07$ 作为判断阈值,其效率与 80 个氨基酸片段序列相似性比对应序列一致性 $>35\%$ 的效率相同^[7],因此在本研究中全长比对以 E 值 $<3.9E-07$ 作为判断阈值。三个数据库的比对结果提示,重组人 α -乳清白蛋白具有潜在的致敏性。

由于致敏原十分复杂,不仅仅取决于氨基酸序列,因此基于序列相似性的生物信息学预测方法不可避免地存在一些缺陷。有研究者对 2 290 个已知过敏原和 234 760 个已知非过敏原进行分析,发现 FAO/WHO 推荐的以 80 个氨基酸序列一致性 $>35\%$ 作为判断标准的灵敏度(已知致敏原被判定为致敏原的比例)为 99.2%,特异度(已知非致敏原被判定为非致敏原的比例)仅为 9.6%,即非致敏原中有 90.4% 被误判为致敏原^[14]。2012 年广州医学院在创建 AFFP(allergen family featured peptide)算法基础上自主开发了过敏原判别软件 SORTALLER,用该软件分析上述已知过敏原及非致敏原的灵敏度为 98.6%、特异度为 98.4%。笔者用 SORTALLER 分析

重组人 α -乳清白蛋白(<http://sortaller.gzhmc.edu.cn/>),根据 0.348 的得分将该蛋白判断为非致敏原。

本研究组在此之前还进行了重组人 α -乳清白蛋白的消化稳定性试验,结果显示,重组人 α -乳清白蛋白在模拟胃肠液和小型猪胃肠液中均容易被消化,这与天然的人 α -乳清白蛋白是一致的^[15]。由于蛋白质经口摄入后一般经消化后才吸收,而消化后的小片段通常不具有抗原性,因此根据该蛋白不抵抗胃肠液消化的结果判断其致敏可能性不大。

综合分析上述结果,由于与已知致敏原的相似性不能排除重组人 α -乳清白蛋白具有潜在致敏性,但 SORTER 分析和消化稳定性结果提示其产生致敏性的可能性不大。在后续研究中将以过敏患者血清进行血清学筛选试验,为该蛋白的致敏性评价提供更充足的依据。

参考文献

- [1] Lönnerdal B, Lien E L. Nutritional and physiologic significance of alpha-lactalbumin in infants [J]. *Nutr Rev*, 2003, 61(9): 295-305.
- [2] Montagne P, Cuillière M L, Molé C, et al. Immunological and nutritional composition of human milk in relation to prematurity and mother's parity during the first 2 weeks of lactation [J]. *J Pediatr Gastroenterol Nutr*, 1999, 29(1): 75-80.
- [3] 张英, 刘珊, 徐海波, 等. 转基因牛的人 α -乳清白蛋白的生理作用与安全性 [J]. *上海预防医学杂志*, 2010(22): 484-486.
- [4] 支媛, 刘海波, 耿桂英, 等. 转人 α 乳清白蛋白基因奶粉亚慢性毒性试验 [J]. *卫生研究*, 2011, 40(4): 426-430.
- [5] Food and Agriculture Organization, World Health Organization. Evaluation of Allergenicity of Genetically Modified Foods [R]. Rome, Italy: FAO/WHO, January 2001.
- [6] 程芳, 孙月眉, 唐宁波, 等. 变应原数据库在变态反应学领域的应用 [J]. *中华临床免疫和变态反应杂志*, 2010, 4(2): 140-144.
- [7] Andre S, Gary B, Scott M. The use of E-scores to determine the quality of protein alignments [J]. *Regul Toxicol Pharmacol*, 2009, 54: S26-S31.
- [8] Silvanovich A, Nemeth M A, Song P, et al. The Value of Short Amino Acid Sequence Matches for Prediction of Protein Allergenicity [J]. *Toxicol Sci*, 2006, 90(1): 252-258.
- [9] Stadler M B, Stadler B M. Allergenicity prediction by protein sequence [J]. *FASEB J*, 2003, 17(9): 1141-1143.
- [10] 甄宇江. 食物致敏原与食品安全 [M]. 北京: 中国标准出版社, 2011: 224-235.
- [11] 王蒙, 张余, 贾小丽, 等. 食物中常见过敏原及其过敏特征 [J]. *中国食物与营养*, 2008, 11: 62-64.
- [12] Bruce G H. Food Safety of Proteins in Agricultural Biotechnology [M]. CRC Taylor & Francis, 2007: 209-222.
- [13] Goodman R E, Tetteh A O. Suggested Improvements for the Allergenicity Assessment of Genetically Modified Plants Used in Foods [J]. *Curr Allergy Asthma Rep*, 2011, 11: 317-324.