

药用/工业用转基因植物数据库的构建

盛耀¹, 罗云波¹, 谢子鑫¹, 贺晓云^{1,2}, 黄昆仑^{1,2}, 许文涛^{1,2*}

(1. 中国农业大学食品科学与营养工程学院, 北京 100083;

2. 农业部转基因生物食用安全监督检验测试中心, 北京 100083)

摘要: **目的** 收集以植物作为生物反应器生产药用/工业用产品的数据信息, 构建药用/工业用转基因植物数据库。**方法** 收集药用/工业用转基因植物相关资料, 每个转化事件收集包括外源蛋白表达量、毒性、致敏性、临床数据、安全等级在内的 20 多种信息。利用 Linux Apache Mysql PHP(LAMP)架构首次设计药用/工业用转基因植物数据库, 面向用户不同需求设计初级检索和高级检索两种检索方式。**结果** 本研究成功构建了药用/工业用转基因植物数据库。目前, 收集了 300 多份数据信息, 整理了转化事件近 200 条, 完成信息上传 108 条, 并将定期更新。**结论** 该数据库具有综合性、专业性、实用性和首创性的特点, 为药用/工业用转基因植物新品种的研发提供信息帮助, 为政府机构审批和监管以及制定其安全评价标准提供技术支撑。

关键词: 药用/工业用; 转基因植物; 数据库; 外源蛋白; 毒性; 致敏性

Construction of database of pharmaceutical and industrial transgenic plants

SHENG Yao¹, LUO Yun-Bo¹, XIE Zi-Xin¹, HE Xiao-Yun^{1,2},
HUANG Kun-Lun^{1,2}, XU Wen-Tao^{1,2*}

(1. College of Food Science and Nutritional Engineering, China Agricultural University, Beijing 100083, China;

2. The Supervision, Inspection and Testing Center of Genetically Modified Organisms, Ministry of Agriculture, Beijing 100083, China)

ABSTRACT: Objective To construct the database of pharmaceutical and industrial transgenic plants, the research group collected published data related with genetically modified (GM) plants as bioreactors. **Methods** Collect information associated with GM plants. Every GM event contains more than 20 parameters including the expression of exogenous protein, toxicity, allergenicity, clinical data and security level, etc. The database was designed with Linux Apache Mysql PHP (LAMP) and two kinds of retrieval modes, primary retrieval method and advanced retrieval method, were realized for different users. **Results** In this study, the database of pharmaceutical and industrial transgenic plants was constructed successfully. Presently, we collected more than 300 published data related with GM plants as bioreactors, compiled 200 GM events approximately and uploaded 108 GM events of the information. **Conclusion** The database is characterized by comprehensiveness, specialty, practicability and innovation, providing information resources for new GM plant varieties development and technological support for government to formulate regulations, approve and supervise pharmaceutical and industrial GM plants.

基金项目: 转基因生物新品种培育重大专项(2012ZX08011-003)

Fund: Supported by Major Project of the Development of New Genetically Modified Organisms (2012ZX08011-003)

*通讯作者: 许文涛, 副教授, 博士生导师, 主要研究方向为食品安全。E-mail: xuwentao1111@sina.com.

*Corresponding author: XU Wen-Tao, Associate Professor, College of Food Science and Nutritional Engineering, China Agricultural University, Beijing, China. E-mail: xuwentao1111@sina.com.

KEY WORDS: pharmaceutical and industrial; transgenic plant; database; exogenous protein; toxicity; allergenicity

1 引言

2013 年, 全球转基因植物的种植面积达到 1.752 亿公顷, 比 1996 年增加了 100 多倍^[1], 累计经济效益超过 1200 亿美元^[2], 转基因植物已经成为现代农业史上发展最为迅速的领域。药用/工业用转基因植物是以植物为生物反应器生产药用/工业用原料的新型转基因生物, 被称为第三代转基因产品^[3], 例如利用转基因植物生产疫苗^[4, 5]、抗体^[6]、蛋白酶^[7]等。美国孟山都公司培育出一种生产人类抗体的转基因玉米, 每公顷可产生 37 kg 抗体蛋白^[8]; 先正达种子研发的转基因玉米 Event 3272 含有微生物 α -淀粉酶基因, 该酶能够迅速地将淀粉分解为糖, 从而生产乙醇^[9]。在美国和加拿大, 利用转基因植物生产药用蛋白的田间实验超过 370 次, 有 16 个植物源药品进入临床试验阶段^[10]。

2003 年 3 月 6 日, 美国颁布了生物安全管理法规修订案, 加强了对药用/工业用转基因植物田间试验、运输和进口的管理, 由宽松的通知管理改为严格的许可管理。目前, 药用/工业用转基因植物多数处于研发、田间试验或临床试验阶段, 上市品种较少^[11]。但是药用/工业用转基因植物具有低投入、低风险、易于产业化生产、耐储存、高回报等优点^[12, 13], 已经成为转基因技术的研究热点, 市场前景广阔。

转基因产品的安全性是制约其产业化发展的重要因素。尽管其安全性受到广泛质疑, 但是自转基因产品上市近 20 年来有数亿人食用了转基因食品, 迄今确实没有发生过具有科学实证的转基因食用和环境安全事件^[14]。人们对转基因产品的质疑更多的是出于对其潜在安全性问题的一种担忧。目前, 国际上基于“实质等同性”原则^[15, 16], 对传统转基因植物(例如抗虫、耐除草剂性状)制定了比较全面的、系统的食用安全评价策略^[17, 18], 但是药用/工业用转基因植物外源基因种类和功能与传统转基因植物存在较大差异性, 其食用安全评价技术体系也不能照搬传统转基因植物食用安全评价系统。对于药用/工业用转基因植物的安全性评价一般是按照临床医学评价体系进行评价。国内外还没有建立对其食用安全性

的评价体系^[10]。因此, 构建药用/工业用转基因植物数据库将有助于政府机构制定其安全评价标准, 同时为科研机构 and 生物技术公司研发新植物反应器提供参考。

目前, 国内外关于转基因植物构建的数据库有很多, 各具优点和特色。例如, 国际农业生物技术应用服务组织 (ISAAA) 建立的转基因审批数据库 (<http://www.isaaa.org/gmaprovaldatabase/default.asp>), 国际生命科学学会 (ILSI) 的转基因农作物数据库 (CERA) (<http://cera-gmc.org/index.php/GMCropDatabase>), 加拿大转基因生物风险评估数据库 (<http://www.agbios.com/dbase.php>), 欧盟转基因生物检测数据库 (http://www.gmo_compass.org), 中国转基因作物检测与监测网转基因作物数据平台 (http://www.gmcrop.cn/crop_gene.php), 上海交通大学的转基因数据库 (GMDD) (<http://gmdd.shgmo.org/>) 等。这些数据库基本涵盖了各种已经审批的商业化种植的转基因植物详细遗传信息、安全评价数据、分子技术检测和种植信息等。但是, 目前国内外并没有关于药用/工业用转基因植物信息的专业数据库。

药用/工业用转基因植物不同于传统转基因植物, 其目的是获取药用蛋白或工业原料, 而非食用目的, 但是仍存在意外泄漏扩散到食物链而产生潜在的食用安全的风险。药用/工业用转基因植物表达产物的毒性、致敏性以及其临床应用等信息都是考察其食用安全性的重要依据。我们收集了全球以转基因植物作为生物反应器生产药用/工业用原料的数据信息, 为了确定每个转化事件基因操作过程、基因表达谱、表达产物的安全性和意外扩散后的潜在风险, 整理了包括外源基因、克隆载体、表达量、毒性、致敏性、临床数据、动物实验数据、用途、安全等级等 20 多种信息。并做成开放式数据库, 供研究人员、管理人员及公众进行查询。

2 方法学

2.1 数据库信息收集和整理

2.1.1 信息收集

收集转基因植物作为生物反应器生产药用/工业

用产品相关资料。例如，转基因植物生产乙肝疫苗、狂犬病疫苗、霍乱疫苗、免疫球蛋白 G 抗体、人白介素抗体、人血清白蛋白、干扰素、生长因子、蛋白酶类等。资料主要来源于谷歌学术、美国国立生物技术信息中心(NCBI)、中国知网(CNKI)、美国食品药品监督管理局(FDA)、美国农业部(USDA)、欧洲食品安全局(EFSA)以及各个生物技术公司资源。目前本数据库收集资料 300 份左右，并将定期更新。

2.1.2 信息整理

根据药用/工业用转基因植物的特点以及其安全评价实验，对每个转化事件进行信息整理，统计了包括外源基因、开发团队、国家、研发或上市时间、商业化进程、克隆载体、表达量、糖基化情况、毒性、致敏性、临床数据、动物实验结果、用途、安全等级、参考资料等 20 多种信息。目前共整理了近 200 个转化事件，完成数据上传 108 条，并将定期更新。

2.2 数据库设计和编程

2.2.1 架构设计

该系统采用互联网主流的 LAMP 架构，即 CentOS 5.5 32 位操作系统, Apache, php 以及 mysql 数据库搭建。该架构易于通过横向扩展而达到性能优化, 通过配备资源来达到高可用性, 同时易于开发和维护。全文检索使用 sphinx 服务器进行响应, 由于索引结果会缓存到内存, 查询 I/O 直接访问内存, 可容纳 1000 人同时进行 10 万次请求的情况, 且不影响网站其它页面的正常浏览。

2.2.2 网站设计

前端页面设计了一套采用 CSS+DIV 排版的皮肤, 兼容 IE、火狐、谷歌等主流浏览器, 达到各页面统一规范的效果。后端部分使用 CI 框架进行开发, 该框架支持 MVC 开发模式, 有效地解决了前端页面和后台数据的分离, 实现代码可重用性, 并简化了代码设计的耦合度。

2.3 部署实施

分别安装 LAMP 架构中的 Apache、php、mysql, 并配置相对应的配置文件, 开启 httpd 和 mysql 服务。在此之中, 需要通过安装过程中提示缺失的软件包(例如 windows 下的 ftp server 等服务软件), 通过 yum 进行在线安装。安装 sphinx, 生成需要检索的表字段的全文索引, 并开启服务等待检索请求。

3 结 果

3.1 数据库网站

药用/工业用转基因植物数据库(http://175.102.8.19:8000/)以子网站的形式搭载在中国农业大学食品安全门户网站(http://www.food12315.com.cn/)。

3.2 信息检索系统

3.2.1 检索方式

本数据库针对用户不同需求和 ISAAA 及 CERA 转基因数据库的检索方法设计了两种检索方式: 一种是初级检索(图 1), 输入需要查询的外源蛋白、转基因植物、国家、商业化进程和时间的任意组合信息, 系统将给出符合条件的所有药用/工业用转基因植物事件, 例如“乙型肝炎疫苗 美国”(图 1); 另一种为高级检索(图 2), 依次选择系统给定的外源蛋白、转基因植物、国家、商业化进程信息, 系统将给出特定的药用/工业用转基因植物事件。



图 1 初级检索方式及一级查询结果

Fig. 1 Primary retrieval method and level one query results



图 2 高级检索方式

Fig. 2 Advanced retrieval method

3.2.2 数据展示

输入查询信息, 系统将给出一级查询结果(图 1), 即符合要求的所有转化事件。点击“代码号”超链接, 进入二级查询结果, 即特定转化事件的详细信息, 包

括基本信息(图 3-1)、致敏性生物信息学比对结果(图 3-2)、其他相关信息和参考文献目录(图 3-3)。数据库还提供了参考文献下载链接, 供用户进行深入研究。

4 讨 论

转基因技术自上世纪 80 年代以来发展非常迅速, 经过 30 多年的发展, 转基因技术已经被广泛应用到

农业、医药、化工、食品、环境保护和能源等领域^[14]。近年来, 药用/工业用转基因植物已经成为转基因领域的研究热点。目前, 国际上并没有全面的、系统的评价药用/工业用转基因植物安全性的方法^[19]。本研究构建的药用/工业用转基因植物数据库可以为政府机构制定药用/工业用转基因植物安全评价标准提供技术支撑, 为科研机构 and 生物技术公司研发药用/工

基本信息

外源蛋白:	人乳铁蛋白
英文名:	Recombinant human lactoferrin (rhLF)
国家:	美国
商业化进程:	上市 (实验用); 药用
第一次研发成功时间:	2003
最新优化时间或上市时间:	2008 (实验用上市); 2014 (临床第二阶段)
转基因植物:	水稻
开发团队:	Ventria Bioscience
转入基因:	hLF (codon-optimized HLF gene) 人工合成413/629
载体:	pAPI164, ExpressTecTM
启动子:	水稻胚乳特异性谷蛋白 (GT1)
终止子:	NOS
表达部位:	种子
表达量:	25%总溶解蛋白; 0.5%总谷物
是否糖基化:	是 (植物模式的糖基化, 多木糖缺乏唾液酸)
毒性级别:	实际无毒
半数致死量 (LD50):	>1g/kg (大鼠)
无毒性反应浓度 (NOAEL):	1g/kg (大鼠)
致敏性:	生物信息学分析预测与已知过敏原存在较高同源性, 潜在致敏性较高; 但动物实验中大鼠和儿童均未出现明显致敏
是否为已知过敏原:	否
全序列比对:	
80氨基酸比对:	83.8%ID lactotransferrin precursor [Bos taurus]
8氨基酸比对:	3

图 3-1 二级检索结果——基本信息

Fig. 3-1 Level two query results——basic information

Display Settings: ☑ FASTA

lactoferrin [Homo sapiens]

GenBank: AAA59511.1
[GenPept](#) [Graphics](#)

```
>gi|187122|gb|AAA59511.1| lactoferrin [Homo sapiens]
MKLVFLVLLFLGLGCLAGRRRRSQWCAVVSQPEATKCPQQRHMKVRGPPVSPICIKRDPSPICIQAI
ENRADAVILDGGFIYEAGLAPYKLRVAAEVYTERQPRTHYAVAVVKKGGSPQLNELQGLKSCHTGLR
RTAGVNVPIGTLRPFLLNWTPPEPEIAAARFFSASCVPADKGGQFPNLCRLCAGTGENKCAFSSQEPYF
SYGAFKCLRDGAGVAFIRESTVFEDLSDEAERDEYELLPDNTKRPVDFKDLARVPSHAVVARSV
NGKEDAIWNLRAQEKFGKDKSPKFLPGSPSQKDLDFKDSAGIFSRVPPRIDSGLYLGSYFTAIQN
LRKSEBEVAARRRVVAVGQBLKRCNWSGLSBG SVTCSASTEDICIALVLRGADAMSLDEGVY
TAGKGLVPLAENYKQSSDPDPNCVDRPVEGYLAAYVRRSDTSLTWNVSKGKKSCHTAVDRTAGVN
IPMGLLFNQITGCKDFEYFSQSCAPGSDPRSNLCAICIGDEQGENKCVFNSNRYTYGTAFRCLAENAG
DVAFVKDVTYVQNTDGNNEAWARDLKLADFALLCLDGRKRPVTEARSKLAPAMNHAVSRMDKVERLK
QVLLHQARFGKRGSDCPDKFLQSETKNLLFNDNTECLARLHGKTYTEKYLGPQYVAGITNLKCKSTS
PLLEACEFLKX
```

Hit #	Defline	Species	Best %ID	# Hits > 35%	Full Alignment			Links	
					E-val	%ID	length	NCBI	Details
1	gi 30794292 ref NP_851341.1 lactotransferrin prec	Bos taurus	83.80%	632of632	0.69	50%	711	gi 30794292	GO!
2	gi 1351295 sp P02789.2 TRFE_CHICK RecName: Full=Ov	Gallus gallus	67.55%	625of632	1.1e-156	51.40%	718	gi 1351295	GO!
3	gi 757851 emb CAA26040.1 ovotransferrin [Gallus g	Gallus gallus	66.70%	625of632	3.7e-156	51.10%	718	gi 757851	GO!

Number of sequences with at least one 8mer match = 3

- ▶ GI:30794292 Hits:75 -- Def: lactotransferrin [Bos taurus]
- ▶ GI:757851 Hits:11 -- Def: ovotransferrin [Gallus gallus]
- ▶ GI:1351295 Hits:14 -- Def: Ovotransferrin precursor (Conalbumin) (Allergen Gal d 3) (Gal d III) (Serum transf

图 3-2 二级检索结果——外源蛋白致敏性生物信息学比对

Fig. 3-2 Level two query results——Allergenic bioinformatics comparison of exogenous protein

其他信息	
动物实验结果:	“大鼠28天重复毒性试验, 100, 500, 1000mg/kg纯蛋白灌胃给药和180, 1800g/kg转基因大米饲养实验, 均未发现显著性毒性及过敏反应。”
临床试验结果:	临床第一阶段儿童口服(ORS)试验未出现明显的毒性及过敏反应。市售乳铁蛋白无不良反应。
蛋白用途:	作为营养保健类成分添加在奶粉中; 用于缺铁性贫血的治疗和腹泻
药用/工业用:	食品/药用/细胞实验
安全等级分类:	一级
参考文献	
参考文献书目:	1. Expression of human lactoferrin in transgenic rice grains for the application in infant formula.2002. 2. Resubmission of Amended GRN 000162 (FDA) 2007. 3. Complete nucleotide sequence of human mammary gland lactoferrin.1990. 4. Recombinant human lactoferrin treatment for global health issues: iron deficiency and acute diarrhea.2004. 5. A rice-derived recombinant human lactoferrin stimulates fibroblast proliferation, migration, and sustains cell survival. 2009. 6. Bioactive recombinant human lactoferrin, derived from rice, stimulates mammalian cell growth.2008.
过敏源查询参考序列	过敏源查询参考序列: http://www.ncbi.nlm.nih.gov/protein/187122?report=fasta

图 3-3 二级检索结果——其他信息及参考文献

Fig. 3-3 Level two query results——other information and reference

业用转基因植物新品种提供全面的、方便的信息资源。同时, 本数据库搭建的信息平台将为政府职能部门对药用/工业用转化事件的审批和监管提供充足的信息。此外, 该数据库的成功建立具有重要的科普宣传作用, 为公众和专业人士了解药用/工业用转基因植物及其产品的研究现状和商业化情况以及安全性提供知情权。

药用/工业用转基因植物数据库总结了近 30 年来全球转基因植物作为生物反应器生产药用/工业用产品的数据信息, 包括处于研发阶段、临床试验阶段和已经上市的转化事件。本数据库具有以下几个特点: (1)综合性, 本数据库信息范围涉及疫苗类、抗体类、干扰素类、生长因子、蛋白酶类和其他药用/工业用产品; (2)专业性, 本数据库针对药用/工业用转基因植物特点, 整理了包括外源基因产物表达量、毒性、致敏性、临床数据(药用产品)在内的 20 多种信息; (3)实用性, 本数据库的两种检索方式满足了不同用户需求, 所有资料免费为任何用户使用, 数据资料将定期进行更新; (4)首创性, 该数据库是全球第一个全面地收录药用/工业用转基因植物的专业数据资源, 具有重要作用。药用/工业用转基因植物数据库已经完成初步构建, 今后将继续对该数据库进行优化和完善: (1)追踪现有数据信息, 即每个转化事件的研发进度或商业化进展; (2)获取最新药用/工业用转基因植物信息, 及时更新数据资源; (3)数据库网站建设优化, 包括高级检索方式的改进、网站页面优化、数据库升级到中英文两种版本。

参考文献

- [1] James C. Global status of commercialized biotech/GM crops: 2013 [R]. ISAAA Brief, No. 46, ISAAA: Ithaca NY 2013.
- [2] Brookes G, Barfoot P. The global income and production effects of genetically modified (GM) crops 1996–2011 [R]. GM Crops and Food: Biotechnology in Agriculture and the Food Chain, 2013, 4(1): 74–83.
- [3] Spök A, Twyman RM, Fischer R, *et al.* Evolution of a regulatory framework for pharmaceuticals derived from genetically modified plants [J]. Trends Biotechnol, 2008, 26(9): 506–517.
- [4] Sala F, Manuela RM, Barbante A, *et al.* Vaccine antigen production in transgenic plants: strategies, gene constructs and perspectives [J]. Vaccine, 2003, 21(7): 803–808.
- [5] Ma JKC, Drake PMW, Chargelegue D, *et al.* Antibody processing and engineering in plants, and new strategies for vaccine production [J]. Vaccine, 2005, 23(15): 1814–1818.
- [6] Stevens LH, Stoop GM, Elbers IJW, *et al.* Effect of climate conditions and plant developmental stage on the stability of antibodies expressed in transgenic tobacco [J]. Plant Physiol, 2000, 124(1): 173–182.
- [7] Yang DC, Guo FL, Liu B, *et al.* Expression and localization of human lysozyme in the endosperm of transgenic rice [J]. Planta, 2003, 216(4): 597–603.
- [8] 庞俊峰, 黄东光, 吴燕民. 植物生物反应器研究进展[J]. 生物技术通讯, 2011, (1): 21–25.
Pang JF, Huang DG, Wu YM. Research Progress in Corps as Plant Bioreactor [J]. Biotechnol Bull, 2011, (1): 21–25.
- [9] 食品伙伴网. 美国农业部完全解除对转基因乙醇玉米的管制 [EB/OL]. (2011-02-13). <http://www.foodmate.net/news/yujing/>

- 2011/02/175091.html.
Food Mate Website. The United States department of agriculture deregulated the genetically modified maize expressing ethanol [EB/OL]. (2011-02-13). <http://www.foodmate.net/news/yujing/2011/02/175091.html>.
- [10] 梅晓宏, 许文涛, 贺晓云, 等. 新型转基因植物及其食用安全性评价对策研究进展[J]. 食品科学, 2013, 34(5): 308–312.
Mei XH, Xu WT, He XY, *et al.* Research Progress in Countermeasures of Food Safety Assessment to New Type of Genetically Modified Plants [J]. J Food Sci, 2013, 34(5): 308–312.
- [11] 刘培磊, 李宁, 连庆, 等. 利用植物生物反应器生产药用蛋白的研究进展[J]. 生物技术进展, 2013, 3(5): 309–316.
Liu PL, Li N, Lian Q, *et al.* Progress in Using Plant Bioreactor to Produce Pharmaceutical Protein [J]. Cur Biotechnol, 2013, 3(5): 309–316.
- [12] Owen MR, Pen J. Transgenic plants: a production system for industrial and pharmaceutical proteins [M]. John Wiley & Sons, 1996.
- [13] Ma JKC, Chikwamba R, Sparrow P, *et al.* Plant-derived pharmaceuticals—the road forward [J]. Trends Plant Sci, 2005, 10(12): 580–585.
- [14] 盛耀, 许文涛, 罗云波. 转基因生物产业化情况[J]. 农业生物技术学报, 2013, 21(12): 1479–1487.
Sheng Y, Xu WT, Luo YB. Commercialization of Genetically Modified Organisms[J]. J Agr Biotechnol, 2013, 21(12): 1479–1487.
- [15] König A, Cockburn A, Crevel RWR, *et al.* Assessment of the safety of foods derived from genetically modified (GM) crops [J]. Food Chem Toxicol, 2004, 42(7): 1047–1088.
- [16] Domingo JL, Giné BJ. A literature review on the safety assessment of genetically modified plants [J]. Environ Int, 2011, 37(4): 734–742.
- [17] EFSA. Safety and nutritional assessment of GM plants and derived food and feed: the role of animal feeding trials [R]. Food and chemical toxicology: an international journal published for the British Industrial Biological Research Association, 2008, 46: S2.
- [18] Singh OV, Ghai S, Paul D, *et al.* Genetically modified crops: success, safety assessment, and public concern [J]. Appl Microbiol Biot, 2006, 71(5): 598–607.
- [19] Sparrow PAC, Irwin JA, Dale PJ, *et al.* Pharma-Planta: road testing the developing regulatory guidelines for plant-made pharmaceuticals [J]. Transgenic Res, 2007, 16(2): 147–161.

(责任编辑: 白洪健)

作者简介



盛 耀, 硕士研究生, 主要研究方向为转基因食品安全。

E-mail: shengyao314@163.com



许文涛, 副教授, 博士生导师, 主要研究方向为食品安全。

E-mail: xuwentao1111@sina.com