



## ORIGINAL ARTICLE

# Allergenicity Study of Genetically Modified Herbicide Resistant Crops (Bioinformatics Assessment)

Najaf Allahyari Fard<sup>1</sup>, Zarrin Minucheher<sup>1</sup>, Amir Mousavi<sup>2</sup>

National Institute of Genetic Engineering and Biotechnology (NIGEB),

<sup>1</sup>Bioinformatics Group,

<sup>2</sup> Head of Plant Biotechnology Department.

allahyar@nigeb.ac.ir

### ABSTRACT

Genetically modified herbicide resistant (GMHR) crops have seriously affected the increase of food production. GMHR events are used on >80% 160 million estimated hectares of transgenic crops grown annually across 29 countries. GMHR traits encompass five events, including two transgenes code for glyphosate resistance and insensitive 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS), the cp4 epsps gene from *Agrobacterium tumefaciens* strain CP4 that causes shikimate pathway in another manner as the process mediate by phosphoenol pyruvate (PEP) and the mutated zm-2mepsps from corn (*Zea mays* L.), and three transgenes code for metabolic inactivation. One gene from *Ochrobactrum anthropi* strain LBAA encodes for glyphosate oxidoreductase (GOX), and two genes pat and bar from *Streptomyces viridochromogenes* and *Streptomyces hygroscopicus*, respectively, encode N-acetyltransferases that inactivate the glufosinate. Bioinformatics allergenicity of CP4 epsps, zm-2mepsps, GOX, pat, and bar proteins was studied in this research. Proteins encoded by CP4 epsps, zm-2mepsps, GOX, pat, and bar genes contain 427, 455, 431, 183, 183 amino acids respectively. These sequences were aligned using the FASTA35.04 program in six allergen databases of FARRP, SDAP, ADFS, PSD, Allergome and Allpred. The sequence alignment was implemented with the allergen proteins in three matches including: the full sequence matching sequence, matching the 80 amino acids and eight amino acids. The results showed no similarity among CP4 epsps, zm-2mepsps, GOX, pat, and bar proteins in addition to allergen proteins in the full sequence matching. Matching the 80 amino acid (Domain) in the SDAP database showed three little similarities (35, 36 and 37.25) for CP4 epsps that was not confirmed in the Allpred for mapping of IgE epitopes search. Matching of 8 amino acids showed no similarity to determine the epitope potential. Therefore, we conclude that CP4 epsps, zm-2mepsps, GOX, pat and bar proteins have non-allergenic potential at bioinformatics levels.

**Key words:** Bioinformatics Assessment, Allergenicity, Genetically Modified, Herbicide Resistant Crops.

Received 22/02/2013 Accepted 16/03/2013

© 2013 AEELS, INDIA

### INTRODUCTION

#### Genetically Modified Plants

The world's population will increase from 6 billion to 9.1 billion at the start of this century by 2050. The majority of this increase will occur in the developing countries [1]. The continuing world's population will mean that the global demand for food will increase and the world has an important challenge about food security in the future [2]. Food and Agriculture Organization (FAO) has predicted that food preparation for 9.1 billion world's population in 2050 requires 70 percent increased food production in world and 100 percent increase of food production in developing countries. The report of International Service for the Acquisition of Agri-biotech Applications (ISAAA) indicates that the genetically modified plants (GMPs) can produce genetically modified (GM) foods and is a key solution to growing demand for food in the world [3].

Creation of GMPs via genetic engineering methods to introduce a foreign gene to create a new distinct has been one of the most rapidly adopted technologies in the history of agriculture [4]. Traditional breeding requires compatible gene parents (donor and recipient) and transfers thousands of undesirable genes, whereas genetic engineering allows to introduce one or more benefit genes from living organism into the genome of the another organism such as recipient plant, animal, fungi, bacteria, etc [4].

GMPs have many characteristics such as resistance to herbicides, virus, pests, water stress and salinity, delayed ripening and altered oil contents [1]. GMPs have developed from 1996. In recent

years the GMPs have been developed with the aim to achieve food security, health and industrial applications. Total areas under cultivated GMPs increase from 1/7 million hectares to 160 million hectares (MHs) during 1996 to 2011 (fifteen years) in the world (Figure1) [3]. The main GMPs were soya, followed by maize, cotton and canola. It is estimated that 16.7 million farmers, in 29 countries planted GMPs in 2011. The largest producer of GMPs was the USA (69 MHs) in 2011, followed by Brazil (30.3 MHs), Argentina (23.7 MHs), India (10.6 MHs), Canada (10.4 MHs), China (3.9 MHs), Paraguay (2.8 MHs), Pakistan (2.6 MHs), South Africa (2.3 MHs), Uruguay (1.3 MHs) and other 19 countries producing less than 1 MHs (Bolivia, Australia, Philippines, Myanmar, Burkina Faso, Mexico, Spain, Colombia, Chile, Honduras, Portugal, Czech Republic, Poland, Egypt, Slovakia, Romania, Sweden, Costa Rica, Germany) [3].

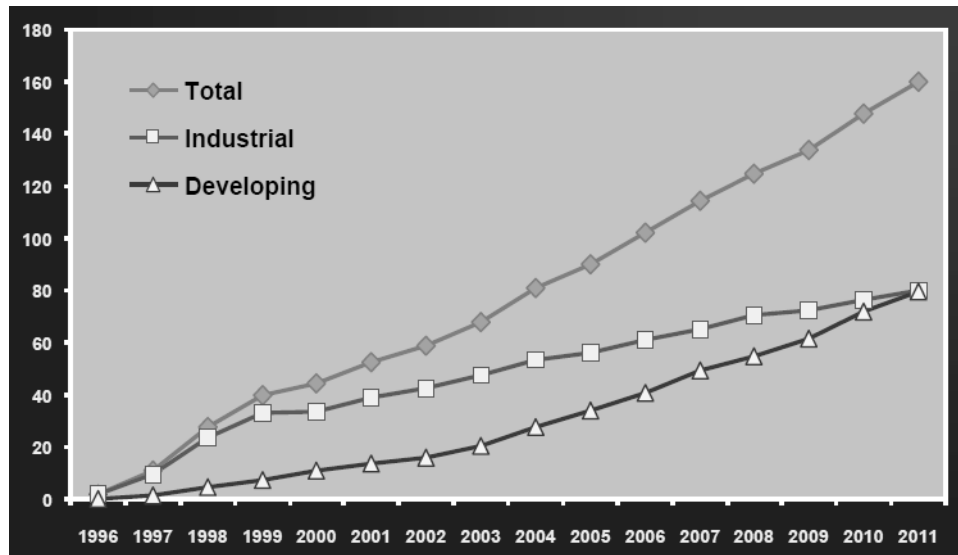


Figure1.Total Area of GMPs, 1996 to 2011: Industrial and Developing Countries (MHs) [3]

### Allergenicity

Allergy is an important health issue, which is caused by certain food proteins, including proteins derived from GMPs [5]. The prevalence of allergic diseases has grown in many industrialized and urbanized countries during the last 50 years and outbreak of allergy is increasing in a daily manner [6]. Almost 30 to 40 percent of the world's populations suffer from allergic diseases [7]. Allergens are proteins or glycoproteins that are recognized by Immunoglobulin E (IgE) which is produced by the immune system of individuals [8].

Allergen is an antigen capable of stimulating a type-I hypersensitivity reaction in atopic persons through IgE responses [9]. Allergenic proteins must crosslink specific IgE molecules, bound to the surface of mast cells and basophiles to stimulate an immune response [10]. Common allergens include animal proteins and animal dander, bacteria and viruses, chemicals, dust, drugs (such as antibiotics or medications you put on your skin), foods (such as milk, chocolate, strawberries, and wheat), perfumes, plants, pollen (Figure2), and smoke.

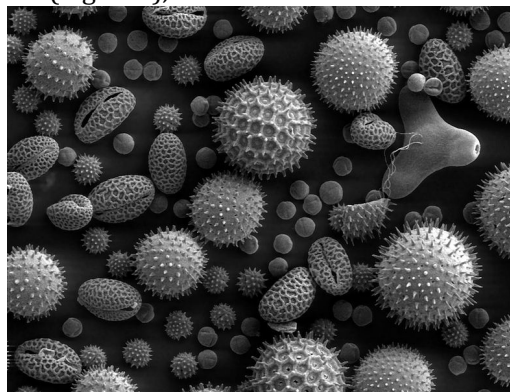


Figure2. A scanning electron microscope (SEM) of miscellaneous plant pollens.

Until now more than 1600 allergens were indentified and many allergen databases have been approved for identification of allergens and bioinformatics allergenicity assessment of novel proteins [8].

### ALLERGENICITY ASSESSMENT OF GMPS NOVEL PROTEINS

One of the most important issues in the consumption of GMPs products is to ensure safety, health and non-allergenicity of them. Biosafety of GMPs and their derivatives is a major topic in the agenda of governments and societies worldwide [11]. An increasing interest in the development of GMPs, has generated considerable debate about allergenicity [12]. There has been a growing research to assess the potential allergenicity of novel proteins with the development of GMPs [13]. Based on the increases of different allergies, allergy status has become more difficult in the world. According to the international regulations (FAO, WHO) evaluation of the allergenicity of novel proteins is necessary in genetically modified organisms (GMOs) [14]. Eight major allergenic foods are peanuts, tree nuts, cow's milk, hen's eggs, fish, crustacea (e.g., shrimp), wheat, and soybeans. Approximately 90 percent of food allergies are associated with a small number of specific proteins represented by these foods. Geography and place of life can affect on the prevalence of food allergy (e.g., walnut and pecan in the US, buckwheat in Asia, celery in Europe) [15].

Requirements for the allergenicity assessment of novel proteins and how these requirements might best be addressed have been widely discussed [16]. Allergenicity assessment is one stage in process of biosafety assessment of GMPs. Utilization of bioinformatics tools to evaluate the allergenicity potential of a novel protein has seen significant progress over the last decade [17]. Bioinformatics analysis has been considered a main part of the safety assessment for GMPs because it is the first step in supporting an assessment of the biosafety with food allergy [18]. Bioinformatics analysis for allergenicity assessment of proteins expressed from induced genes is a primary step for determination of allergenicity of novel proteins. The use of bioinformatics screening of allergenicity assessment for the novel proteins in allergen databases is recommended by the World Health Organization [19], the European Food Safety Authority (EFSA) and the US Environmental Protection Agency [20]. Bioinformatics allows two questions to be asked: a) Is the novel protein an existing allergen? b) Is the novel protein likely to cross-react with an existing allergen? [17]. A 'weight of evidence' approach is recommended by the Codex Alimentarius Guidelines to assess the risk of allergenicity of GMPs, including bioinformatics analysis, digestibility and animal models [21] (Tabel1).

Table1. Codex Alimentarius Guidlines (2003) [22].

Codex recommended allergy assessment
If introduced protein from a non-allergenic source
Assess amino acid sequence similarity to known allergens*
Assess in vitro pepsin resistance
If introduced protein from an allergenic source
Assess amino acid sequence similarity to known allergens*
Assess in vitro pepsin resistance
Assess specific IgE binding
Assess skin-prick testing on appropriate individuals
Other considerations as scientific knowledge and technology evolves
Exposure level of the introduced protein
As science and technology evolves other methods may be considered
Targeted sera screens
Animal models
Examination of newly expressed proteins for T cell epitopes and structural motifs associated with allergens

Bioinformatics analysis for allergenicity assessment of proteins is performed via allergen databases [23]. Bioinformatics, databases, and computational tools are increasingly important in the study of allergen proteins, particularly in the assessment of allergenicity [24]. Until now many allergen

databases have been approved for identification of allergens and bioinformatics allergenicity assessment of novel proteins. Table 2 has showed 14 allergen databases. These databases are general and specialist either contain different datasets of allergen or have different tools for assessing potential allergenicity of proteins [25] (Table 2).

Table2. Allergen databases

Database Name	URL
IUIS Allergen Nomenclature (IUIS)	<a href="http://www.allergen.org">www.allergen.org</a>
Structural Database of Allergenic Proteins (SDAP)	<a href="http://fermi.utmb.edu/SDAP">fermi.utmb.edu/SDAP</a>
Food Allergy Research and Resource program (Farrp)	<a href="http://www.allergenonline.com">www.allergenonline.com</a>
Food Allergens of Plant Origin (Protall)	<a href="http://www.ifr.bbsrc.ac.uk/Protall">www.ifr.bbsrc.ac.uk/Protall</a>
Biotechnology Information for Food Safety (BIFS)	<a href="http://www.iit.edu/~sgendel/fa.htm">www.iit.edu/~sgendel/fa.htm</a>
Central Science Laboratory Allergen Database (ADB)	<a href="http://www.csl.gov.uk/allergen">www.csl.gov.uk/allergen</a>
Allergen Families (AllFam)	<a href="http://www.meduniwien.ac.at/allergens/allfam/">http://www.meduniwien.ac.at/allergens/allfam/</a>
Gateway to All Allergy on the Web(All Allergy)	<a href="http://allallergy.net/">http://allallergy.net/</a>
Food Allergy Database (InformAll)	<a href="http://foodallergens.ifr.ac.uk/">http://foodallergens.ifr.ac.uk/</a>
The Platform for Allergen Knowledge (Allergome)	<a href="http://www.allergome.org">http://www.allergome.org</a>
Allermatch™	<a href="http://www.allermatch.org">http://www.allermatch.org</a>
Protein Structure Discovery (PSD)	<a href="http://www.bionet.ssc.ru/psd/">http://www.bionet.ssc.ru/psd/</a>
Algpred	<a href="http://www.imtech.res.in/raghava/algpred/index.html">http://www.imtech.res.in/raghava/algpred/index.html</a>
Allergen Database for Food Safety (ADFS)	<a href="http://allergen.nihs.go.jp/ADFS/">http://allergen.nihs.go.jp/ADFS/</a>

Allergens have variety of biologic functions, including proteases, pathogenesis-related proteins, structural proteins, ligand-binding proteins, lipid transfer proteins, profilins, and calcium-binding proteins [26] that are growing increasingly [27] (Figure 3).

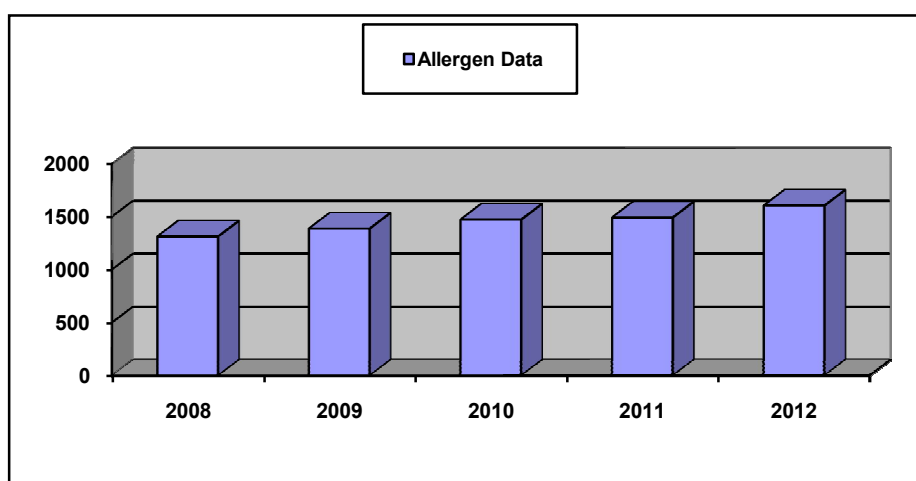


Figure3. Yearly allergen updated in the FARPP database ([www.allergenonline.org](http://www.allergenonline.org)) from 2008–2012. [18]

FAO/WHO guidelines about evaluation of allergenicity of foods derived from GMPs have stated that a query protein is potentially allergenic if it either has an equal or bigger 35% than sequence similarity over a window of 80 amino acids or identity of at least six contiguous amino acids (identity length  $n=6-8$ ) when compared with known allergens [28].

### GMHR crops

GMHR crops are of five events, including two transgenes codes for glyphosate resistance and insensitive 5- enolpyruvylshikimate-3-phosphate synthase (EPSPS), the *cp4 epsps* gene from *Agrobacterium tumefaciens* strain CP4 that causes shikimate pathway in another manner as the process mediate by phosphoenol pyruvate (PEP) and the mutated *zm-2mepsps* from corn (*Zea mays* L.), and three transgenes codes for metabolic inactivation. One gene from *Ochrobactrum anthropi*

strain LBAA is encoded for glyphosate oxidoreductase (GOX), and two genes *pat* and *bar* from *Streptomyces viridochromogenes* and *Streptomyces hygroscopicus*, respectively, encode N-acetyltransferases that inactivate glufosinate [29].

In the present investigation, CP4 epsps, zm-2mepsps, GOX, *pat* and *bar* proteins expressed in GMHR crops have been evaluated for allergenicity by sequence alignment searches using six allergen databases including FARRP, SDAP, ADFS, PSD, Allergome, and Algpred and bioinformatics tools. The sequence alignment searches were conducted using FASTA algorithm. The CP4 epsps, zm-2mepsps, GOX, *pat* and *bar* proteins were also assessed for presence of domains similar to allergen proteins to test for possibility of allergenic cross reactivity.

## MATERIALS AND METHODS

### 1. Query protein sequence

The amino acid sequence of the cp4 epsps, zm-2mepsps, GOX, *pat*, and *bar* was taken up for the study from NCBI in FASTA format and used for sequence alignment searches.

### 2. Bioinformatics analysis to search allergenicity

Six allergen database systems, i.e., FARRP, SDAP, ADFS, Allergome, Algpred, and PSD, using FASTA algorithm available at <<http://farrp.unl.edu/>>, <<http://fermi.utmb.edu/SDAP/>>, <<http://allergen.nihs.go.jp/ADFS/>>, <<http://www.allergome.org/>>, <<http://www.imtech.res.in/raghava/algpred/index.html>>, <<http://www-bionet.ssc.ru/psd/cgi-bin/programs/Allergen/allergen.cgi/>> respectively were used to identify any potential sequence matches to allergen proteins that might indicate allergenic cross-reactivity with our above mentioned sequence.

#### 2.1. FARRP

FARRP available at <<http://farrp.unl.edu/>> was used for sequence homology of allergens. The Allergen Online database version 12.0 (updated in Feb 2012) is designed by the Food Allergy Research and Resource Program (FARRP) in the Department of Food Science and Technology of Nebraska University provides a comprehensive list of 1603 sequence entries of unique allergen proteins [30]. These proteins have been identified from allergen sources including food, airway, contact, and venom. FARRP includes FASTA algorithm [31] developed by W.R. Pearson at the University of Virginia (1988) which is used for comparing a query protein sequence with allergen proteins in this database. All allergen proteins present in this database are linked to the sequences of the National Centre for Biotechnology Information (NCBI) of the National Institute of Health (NIH). This study was performed on 20<sup>th</sup> December 2012. A full FASTA35.04 search was performed using the default search and scoring criteria of Pearson and default scoring matrix of BLOSUM 50 [30]. The scoring matrix of BLOSUM 50 is supposed to identify highly similar proteins that are likely to have similar overall structure and function, whether of distant evolutionary origin or closely related sequences [32]. The aim of the comparison routine is to assess whether the query protein sequence is identical to, or homologous with known allergens in the database. If a protein aligns greater than 70% identity over its length, relative to allergen, it is likely to be cross-reactive and if it has less than 50% identity, it is not very likely to be cross-reactive [33]. Codex Alimentarius Guidelines (2003) have stated the alignments of >35% identity over segments as short as 80 amino acids for potential cross-reactivity of a protein [22]. So, every possible contiguous 80-amino acid sequence of each query protein was searched for determining the similarity, beginning with amino acids of 1–80, then 2–81, 3–82 and so on until the last amino acid segment of each protein was compared. The expectation value (E value) and the percent identity evaluate the extent of similarity, which reflects the degree of similarity between a pair of sequences based on matches of identical or functionally similar and structurally-related amino acids. The E-value depends on the overall length of joined (gapped) local sequence alignments, the quality (percent identity, similarity) of the overlap, and the size of the database [34]. A low E score indicates a high degree of similarity between the query sequence and the sequence from the database [30].

#### 2.2. SDAP

SDAP (Structural Database of Allergenic Proteins) available at <<http://fermi.utmb.edu/SDAP/>> is a web server that integrates a database of allergenic proteins and provides rapid, cross-referenced access to the sequences, structures and IgE epitopes of allergenic proteins. The “SDAP” consists of various computational tools that can assist structural biology studies related to allergens. SDAP is a tool for investigation of cross-reactivity between known allergens, to test FAO/WHO allergenicity rules for novel proteins, and predicting the IgE-binding potential of GMPs. SDAP database contains

information about the allergen names, sources, sequences, structures, IgE epitopes and literature references and easy links to the major protein (PDB, SWISS-PROT/TrEMBL, PIR-ALN, NCBI Taxonomy Browser) and literature (PubMed, MEDLINE) online servers. Using this Internet service, it is possible to retrieve information related to an allergen from the most common protein sequences and structure databases (SwissProt, PIR, NCBI, PDB), to find sequence neighbors for an allergen, and search the presence of an epitope other than the whole collection of allergens [35].

### 2.3. ADFS

The ADFS (Allergen Database for Food Safety) available at <<http://allergen.nihs.go.jp/ADFS/>> (updated in Feb 2012) provides information of allergenic proteins and low molecular weight (Low Mol Wt) allergenic compounds. The list of allergenic proteins has been collected from literature sources and FARRP database whose allergens have been peer-reviewed by international allergologist. ADFS has the allergens classified into 8 categories (pollen, mite, animal, fungus, insect, food, latex, and others) [36]. ADFS contains supplemental information on each allergen, such as sequence, glycosylation, and structure, by referring public databases like UniProt. In addition, the list of LMW allergenic compounds was extracted from general allergen information provided in 'AllAllergy' (<http://www.allallergy.net/allergenfind.cfm>). Users can search allergens by keywords and amino acid sequences [37].

### 2.4. Allergome

This database available at <<http://www.allergome.org/>> has been designed to supply information on allergens. The Allergome database is based on the published literature. The current base is 5800 selected papers (November 2004) and is continuously updated until last version (Ver4.0, updated in Oct 2012). Before data entry, the experienced reviewers select information from scientific papers. The Allergome is a searchable using the allergenic molecule filter and includes all the identified and characterized allergenic molecules including those in the official International Union of the Immunological Societies (IUIS) present in allergen nomenclature list. Allergenic molecules and sources are included in the Allergome database independently from the sources (animals or plants) and their tissues (dander, fruit, pollen, seed, spores, venoms, whole bodies etc.), or the routes of exposure (contact, ingestion, inhalation, injection etc.) [38].

### 2.5. Algpred

AlgPred is available at <[www.imtech.res.in/raghava/algpred/](http://www.imtech.res.in/raghava/algpred/)>. AlgPred has been developed/ designed for predicting the allergenic proteins and for mapping of IgE epitopes on allergenic proteins. In AlgPred a systematic attempt has been made to integrate various approaches in order to predict allergenic proteins with high accuracy [39].

### 2.6. PSD

PSD (Protein Structure Discovery) is available at <<http://www-bionet.sccc.ru/psd/cgi-bin/programs/Allergen/allergen.cgi/>>. This database includes many parts such as function, structure, polymorphism, evolution, metabolism and immunology. In this database the allergenicity assessment of novel proteins is performed via immunology part and the Query protein sequence must be in FASTA format (length>=8) [40].

## RESULTS

Bioinformatics analysis was used to identify proteins that are significantly similar to known allergen proteins to check possible allergenic cross-reactivity. The full-length sequences of *CP4 epsps*, *zm-2mepsps*, *GOX*, *pat*, and *bar* genes as expressed in GMHR crops were subjected to bioinformatics searches. The results of FASTA35.04 search of the GMHR proteins expressed in transgenic plants using FARRP, SDAP, ADFS, Allergome, Algpred, and PSD databases did not identify any significant alignment with any of the known allergens. All the query proteins have less than 50% identity (indicating level for cross-reactivity) to allergen proteins. Aalberse in 2000 described a protein sharing more than 70% identity over its length, relative to an allergen is likely to be cross-reactive, or share IgE binding; and those having less than 50% identity are not very likely to be cross-reactive [33]. International criteria of Codex, 2003 declared it needs for suspecting the potential allergenic cross-reactivity of recombinant protein introduced into a GMPs that have at least 35% identity over an 80 amino acid match which was also demonstrated in Golden Rice 2 by Goodman and Wise, in 2006 for fused chloroplast transit peptidylcarotene destaurase1 (fused CTP-CRT1), phosphomannose isomerase (PMI) and phytoene synthase (Psy) proteins [41].



E score values in all searches were equal or less than 1.0. E score is the probability that the sequence alignment was only due to chance, rather than an alignment due to 'real' similarity. E score much smaller than 1 indicates a highly significant alignment, a high degree of structural similarity, and probable evolutionary relationship. A low E score indicates a high degree of similarity between the query sequence and the sequence from the database [30].

The results showed no significant identity between CP4 epsps, zm-2mepsps, GOX, pat, and bar proteins and allergen proteins in the full sequence matching with E score=1 (Table 3). Matching the 80 amino acid (Domain) in the SDAP database showed three negligible similarities of (35, 36 and 37.25) for CP4 epsps (Table 4) that was not confirmed in the Algpred for Mapping of IgE epitopes search. Matching of 8 amino acids showed no similarity to determine the epitope potential.

Table3. Sequence identity search of CP4 epsps, zm-2mepsps, GOX, pat, and bar proteins in allergen databases, E Score=1

Protein Query	FARRP	SDAP	ADFS	Allergome	Algpred	PSD
CP4 epsps	No	Negligible	No	No	Not Epitop	No
zm-2mepsps	No	No	No	No	Not Epitop	No
GOX	No	No	No	No	Not Epitop	No
pat	No	No	No	No	Not Epitop	No
bar	No	No	No	No	Not Epitop	No

Table4. Sequence identity search for 80 amino acids of CP4 epsps protein in SDAP, E Score=1

No.	Allergen	Organism	Common Name	GenBank	Length (aa)	Identity (%)
1	Allergen Ani s 3	<i>Anisakis simplex</i>	nematode	Y19221.1	284	36.25
2	Allergen Ani s 3	<i>Anisakis simplex</i>	nematode	Y19221.1	284	37.5
3	Allergen Asc l 3.0101	<i>Ascaris lumbricoides</i>	Common roundworm	ACN32322.1	287	35.0

## CONCLUSION

Consumption of GMPs needs to have food products as safe as foods derived from the conventional crops. One of the most important issues in the consumption of GMPs products is to ensure safety, health and non-allergenicity of them. Biosafety of GMPs and their derivatives is a major topic in the agenda of governments and societies worldwide [11]. GMHR crops are of five events contain CP4 epsps, zm-2mepsps, GOX, pat and bar proteins that are used on >80% of the estimated 160 million hectares of transgenic crops grown annually across 29 countries [29]. The Food allergens are generally more permanent than food proteins with no allergenic history in the gastrointestinal model [42]. One of the important considerations in GMPs is the possibility that the transgene(s) may produce an allergen protein(s). The potential transfer of a known allergen from one source to another is most significant potential risk of allergy associated with GMPs. It is also important to determine whether the new protein from transgene is significantly similar to any of the known allergen due to proteins that are structurally very similar may be immunologically cross-reactive. The Sequence similarity search of the novel protein with that of known allergens is an important primary stage in the safety assessment process. Proteins with high similarity in sequence or predicted conformational structure to determined allergen proteins are then further tested with sera from individuals having allergies to the identified allergen to possible assessed IgE binding as an indication of their allergenicity [21].

The present study reports on the assessment of the GMHR proteins, viz., CP4 epsps, zm-2mepsps, GOX, pat and bar expressed in GMHR crops for potential allergenic cross-reactivity using six allergen databases and bioinformatics search tools.

The sequence alignment searched six allergen databases including FARRP, SDAP, ADFS, PSD, Allergome, and Algpred. The sequence alignment searches were carried out using FASTA algorithm, FASTA35.04 software and E score values in all searches were less than 1.0. The CP4 epsps, zm-

2mepsps, GOX, pat and bar proteins were also assessed for presence of domains similar to allergen proteins to test the possibility of allergenic cross reactivity.

The results of sequence similarity search using FARRP, SDAP, ADFS, PSD, Allergome, and Algpred databases revealed that none of the proteins met the criteria for suspecting the allergenic cross-reactivity. Results using FASTA35.04 software indicate less than 50% identity so they did not fulfill the criteria for cross-reactivity described by Aalberse (2000) that a protein sharing more than 70% identity over its length, relative to an allergen is likely to be cross-reactive, or share IgE binding. Furthermore, the complementary searches revealed no similarities at the 80 amino acids (domain) level. Because the results showed no similarity among CP4 epsps, zm-2mepsps, GOX, pat, and bar proteins and allergen proteins in the full sequence matching, matching the 80 amino acid (Domain), and matching of 8 amino acids, we conclude that CP4 epsps, zm-2mepsps, GOX, pat, and bar proteins have non-allergenic potential at bioinformatics level. Hence, the bioinformatics search results indicate no need for further in vitro testing such as serum IgE-binding studies. The results indicate that none of the GMHR proteins were found positive for potentially allergenic cross-reactivity.

## ACKNOWLEDGEMENTS

The authors are thankful to National Institute of Genetic Engineering and Biotechnology (NIGEB), Iran for providing the related facilities.

## REFERENCES

1. Purchase IFH. (2005). What determines the acceptability of genetically modified food that can improve human nutrition? *Toxicology and Applied Pharmacology* ;207:19-27.
2. Godfray HC, Beddington JR, Crute IR, Haddad L, Lawrence D, Muir JF, et al. Food security: the challenge of feeding 9 billion people. *Science* 2010;327:812-8.
3. James C. Global status of GM crops for 2011: ISAAA (International Service for the Acquisition of Agric-biotech Applications). 2012.
4. Harlander SK. The evolution of modern agriculture and its future with biotechnology. *Journal of the American College of Nutrition* 2002;21:161S-5S.
5. Konig A, Cockburn A, Crevel RW, Debruyne E, Grafstroem R, Hammerling U, et al. Assessment of the safety of foods derived from genetically modified (GM) crops. *Food and chemical toxicology : an international journal published for the British Industrial Biological Research Association* 2004;42:1047-88.
6. Haahtela T, Valovirta E, Kauppi P, Tommila E, Saarinen K, von Hertzen L, et al. The Finnish Allergy Programme 2008-2018 - scientific rationale and practical implementation. *Asia Pacific allergy* 2012;2:275-9.
7. Pawankar R, Canonica G, Holgate S, Lockey R. WAO white book on allergy. Milwaukee, WI: World Allergy Organization 2011
8. Allahyari Fard N, Minuchehr Z. In Silico Analysis for Allergenicity Assessment of Novel Proteins of GMOs. *Genetics in the 3rd millennium* 2013;10:0-15.
9. Kindt TJ, Osborne BA, Goldsby RA. *Kuby immunology*: WH Freeman; 2006.
10. Power TD, Ivanciuc O, Schein CH, Braun W. Assessment of 3D models for allergen research. *Proteins: Structure, Function, and Bioinformatics* 2013;n/a-n/a.
11. Reis LF, Van Sluys MA, Garratt RC, Pereira HM, Teixeira MM. GMOs: building the future on the basis of past experience. *Anais da Academia Brasileira de Ciencias* 2006;78:667-86.
12. Cao S, He X, Xu W, Luo Y, Ran W, Liang L, et al. Potential allergenicity research of Cry1C protein from genetically modified rice. *Regulatory toxicology and pharmacology : RTP* 2012;63:181-7.
13. Cao S, He X, Xu W, Luo Y, Ran W, Liang L, et al. Potential allergenicity research of Cry1C protein from genetically modified rice. *Regulatory Toxicology and Pharmacology* 2012;63:181-7.
14. Alimentarius C. Foods derived from modern biotechnology. Codex Alimentarius Commission, Joint FAO/WHO Food Standards Programme Rome, Italy 2009.
15. Metcalfe DD, Astwood JD, Townsend R, Sampson HA, Taylor SL, Fuchs RL. Assessment of the allergenic potential of foods derived from genetically engineered crop plants. *Critical reviews in food science and nutrition* 1996;36 Suppl:S165-86.
16. Ladics GS, Knippels LMJ, Penninks AH, Bannon GA, Goodman RE, Herouet-Guicheney C. Review of animal models designed to predict the potential allergenicity of novel proteins in genetically modified crops. *Regulatory Toxicology and Pharmacology* 2010;56:212-24.
17. Ladics GS. Current codex guidelines for assessment of potential protein allergenicity. *Food and chemical toxicology : an international journal published for the British Industrial Biological Research Association* 2008;46 Suppl 10:S20-3.
18. Ladics GS, Cressman RF, Herouet-Guicheney C, Herman RA, Privalle L, Song P, et al. Bioinformatics and the allergy assessment of agricultural biotechnology products: industry practices and recommendations. *Regulatory toxicology and pharmacology : RTP* 2011;60:46-53.
19. Ivanciuc O, Gendel SM, Power TD, Schein CH, Braun W. AllerML: Markup language for allergens. *Regulatory Toxicology and Pharmacology* 2011;60:151-60.



20. Kok EJ, Keijer J, Kleter GA, Kuiper HA. Comparative safety assessment of plant-derived foods. *Regulatory Toxicology and Pharmacology* 2008;50:98-113.
21. Randhawa GJ, Singh M, Grover M. Bioinformatic analysis for allergenicity assessment of *Bacillus thuringiensis* Cry proteins expressed in insect-resistant food crops. *Food and chemical toxicology : an international journal published for the British Industrial Biological Research Association* 2011;49:356-62.
22. Ladics GS, Selgrade MK. Identifying food proteins with allergenic potential: evolution of approaches to safety assessment and research to provide additional tools. *Regulatory toxicology and pharmacology : RTP* 2009;54:S2-6.
23. Ivanciuc O, Gendel SM, Power TD, Schein CH, Braun W. AllerML: markup language for allergens. *Regulatory toxicology and pharmacology : RTP* 2011;60:151-60.
24. Zhang ZH, Tan SC, Koh JL, Falus A, Brusic V. ALLERDB database and integrated bioinformatic tools for assessment of allergenicity and allergic cross-reactivity. *Cellular immunology* 2006;244:90-6.
25. Tong JC, Lim SJ, Muh HC, Chew FT, Tammi MT. Allergen Atlas: a comprehensive knowledge center and analysis resource for allergen information. *Bioinformatics* 2009;25:979-80.
26. Chapman MD, Pomes A, Breiteneder H, Ferreira F. Nomenclature and structural biology of allergens. *The Journal of allergy and clinical immunology* 2007;119:414-20.
27. Brusic V, Millot M, Petrovsky N, Gendel SM, Gigonzac O, Stelman SJ. Allergen databases. *Allergy* 2003;58:1093-100.
28. Stadler MB, Stadler BM. Allergenicity prediction by protein sequence. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* 2003;17:1141-3.
29. Green JM, Owen MD. Herbicide-resistant crops: utilities and limitations for herbicide-resistant weed management. *Journal of agricultural and food chemistry* 2011;59:5819-29.
30. FARRP. Allergenonline. 2012.
31. Pearson WR. Rapid and sensitive sequence comparison with FASTP and FASTA. *Methods in enzymology* 1990;183:63-98.
32. Pearson WR. Flexible sequence similarity searching with the FASTA3 program package. *Methods Mol Biol* 2000;132:185-219.
33. Aalberse RC. Structural biology of allergens. *The Journal of allergy and clinical immunology* 2000;106:228-38.
34. Hileman RE, Silvanovich A, Goodman RE, Rice EA, Holleschak G, Astwood JD, et al. Bioinformatic methods for allergenicity assessment using a comprehensive allergen database. *International archives of allergy and immunology* 2002;128:280-91.
35. Ivanciuc O, Schein CH, Braun W. SDAP: database and computational tools for allergenic proteins. *Nucleic acids research* 2003;31:359-62.
36. Nakamura R, Teshima R, Takagi K, Sawada J. [Development of Allergen Database for Food Safety (ADFS): an integrated database to search allergens and predict allergenicity]. *Kokuritsu Iyakuhiin Shokuhin Eisei Kenkyujo hokoku = Bulletin of National Institute of Health Sciences* 2005:32-6.
37. AFDS. Allergen Database for Food Safety. 2012.
38. Allergome. The Platform for Allergen Knowledge. 2012.
39. Saha S, Raghava GP. AlgPred: prediction of allergenic proteins and mapping of IgE epitopes. *Nucleic acids research* 2006;34:W202-9.
40. Bragin AO, Demenkov PS, Kolchanov NA, Ivanisenko VA. Accuracy of protein allergenicity prediction can be improved by taking into account data on allergenic protein discontinuous peptides. *Journal of biomolecular structure & dynamics* 2013;31:59-64.
41. Goodman RE, Wise, J., Bioinformatics analysis of proteins in Golden rice2 to assess potential allergenic cross-reactivity Univ. of Nebraska Food Allergy Research and Resource Program 2006:1-24
42. Astwood J, Fuchs R. Allergenicity of foods derived from transgenic plants. *Monographs in allergy* 1996;32:105.



**HOW TO CITE THIS ARTICLE:** N. A. Fard, Z. Minuchehr, A. Mousavi. Allergenicity Study of Genetically Modified Herbicide Resistant Crops (Bioinformatics Assessment). *Bull. Env. Pharmacol. Life Sci.* 2 [3] February 2013: 24- 32