ORIGINAL RESEARCH PAPER

Food Science

APPLICATIONS OF BIOINFORMATICS IN FOOD ALLERGY: A BRIEF REVIEW

KEY WORDS: Types of Food Allergy, Bioinformatics, Food Science, Immune Response

Anamika Bagchi*

B.Tech Food Technology, Techno India, Salt Lake, West Bengal University of Technology *Corresponding Author

ABSTRACT

This literature review mainly deals with the concepts of food science, food allergy, allergenomics, bioinformatics and its application in solving challenges pertaining to food allergy. When an after eating a certain food, the immune system gives out a reaction, it is defined as Food Allergy. It is a complex disease that has many challenging area related to its cure and is a worldwide problem. Bioinformatics help in providing solutions to these challenges. As allergens are mainly proteins. There are many allergy databases and tools available in the market that can be used to identify the differences between novel proteins and food allergens that could be the epicentre of multiple allergy syndromes. This review article talks about food allergenicity, allergy and computational approaches employed to solves issues related to it.

1. INTRODUCTION

Food science is a field, which is the combination of engineering, chemistry, processing, preservation and microbiological aspects of food products, to study the nature of foods. Food science mainly comprises of Food Chemistry, Food Microbiology, Food Engineering and Food Processing as concentration areas.

In this field of food science, there are many challenges present. One such challenging area is Food allergy or Food allergenomics. It is mainly a complex disease. Food allergy is mainly induced by food ingredients(Wang et al. 2020). It has been estimated that it affects around 240-550 million people worldwide, especially children. The symptoms of food allergy include coughing, itching, vomiting, abdominal pains, swelling of pharynx and anaphylaxis causing death of the individual. It is a severe public health burden affecting almost 10% adults and 8% children in the United States (Albuhairi and Rachid 2021). Death due to food allergy has an incidence rate of 1.81 per million person-years, amongst food allergenic people(Umasunthar et al. 2013). Hence, the treatment for food allergy associated diseases is promptly needed.

An allergen is a substance mainly a protein moiety which is an ingredient of allergenic source that induces allergenic reactions or immune specific reactions in the individuals irrespective of being adult or children(Wang et al. 2020).Since there is a lack in allergen therapies, oral immunotherapy and allergen avoidance are the most effective method currently available, which is also based on identification of allergens from databases. An individual can be allergenic to multiple allergens while a single source may or may not contain various allergens. Therefore, there lies a need to determine the allergen. Hence, it serves two goals: one, it helps to completely identify all the allergenic groups to reduce the risk of allergy and second, it helps to identify and exclude any non-allergenic compounds that do not associate with the quality of life.

With the advancement of science, newer techniques have been introduced. Bioinformatics is one such area (Kumar and Chordia 2017). Bioinformatics is an interdisciplinary field defined as a branch of biology where computation or computer science and statistics or statistical analysis is used in amalgamation with biological sciences. It involves various disciplines such as comparative genomics, molecular medicine, drug discovery, microbial genome applications and biotechnology. The latest technological advances in genomics, transcriptomics, proteomics and metabolomics along with the use of different software and tools produces the opportunity to explore and decipher composition of food , their micro and macro nutrients, chemistry, biology and the nutritive value of food.

Common food items that cause food allergy are milk, peanuts, eggs, shellfish, tree-nuts, wheat, rice and fruit (Kotra, Chaudhary, and Neetu 2015). Food allergy happens as an adverse reaction resulting from an immunological response from a food antigen. Immunoglobulin E(IgE) as it binds to the food produces an allergenic response. IgE molecules towards allergens via B-cells induce this response. After this, degranulation of mast cells occurs. Receptors present on these cells are binded by the IgE and the body starts to experience itchiness, sneezing and coughing when the mast cells release its mediators such as histamine, leukotriene and prostaglandins(Wang et al. 2020).This literature review focuses on the advances on research related to food allergy or allergenomics. Future intensive allergenomics research can improve the efficiency of promoting, preventing and treating of food allergies. This can help in guiding or promoting the production, management and consumption of potentially allergenic foods with the help from bioinformatic tools and softwares.

2. Types of Food Allergy

According to FARE or Food Allergy Research and Education , there are eight foods responsible for food allergy namely: milk, peanut, egg, tree-nuts, soy, fish, wheat and shell-fish (Kotra, Chaudhary, and Neetu 2015).

Milk Allergy: Generally, it has been found out that children are more prone to this allergy. The major allergens identified are casein, β- lactoglobulin, and α-lactalbumin have been well associated(Cox, Eigenmann, and Sicherer 2021). These proteins share a significant sequence homology and immunological cross- reactivity between milks from various mammals.

Peanut Allergy: Ranging between 30-50% in some study populations, peanut and tree-nut allergies have been found to co-exist in same patients. Although the peanut and tree-nut protein allergenic proteins share homology, they are from different plant taxonomic groups. The relationship between peanut ad tree-nut is patient specific and requires further exploration.

Egg Allergy: Ovumucoid has been associated with egg allergy as major hen's egg food allergen. Hen's egg is by far the most prevalent food allergies associated worldwide. Studies have shown quail and duck egg allergy in the absence of hen's egg allergy but, however clinically relevant avian egg cross-reactivity has not been systematically investigated.

Tree-nut Allergy: Tree-nuts mainly comprises of almond, walnut, pistachios, pecans, hazelnut, cashew, macadamia, Brazil nut and pine nuts. They affect 0.5-1.2% of the population and can easily become severe. Studies have shown various prevalence rates in different regions for different age groups. In the United States, the most common tree nut allergies are

from walnut and cashew, whereas in Europe, hazelnut predominates and the most frequently seen tree nut allergies in the United Kingdom are Brazil nut, almond, and walnut.

Soy Allergy: It is one of the most common food allergies present among children's("The Natural History of Soy Allergy ScienceDirect" n.d.). It has been estimated that approximately 0.4% of children are allergic to soy. Soy protein is sought to be a popular vegetarian alternative to meat and milk allergenic groups. Hence soy avoidance has become challenging with people with soy allergy. It's been difficult to determine an exact relationship between the soy-specific IgE concentrations with clinical soy allergy, both being predominately different in nature. Studies suggest that soy-specific IgE levels of between 20 and 30 kU/L shows a 50% chance of passing the soy challenge.

Fish Allergy: The major fish allergens are parvalbumins, aldolase and enolase. There are various amounts of homologous proteins present among seafoods and particularly fishes which can cause various degree of food allergy among group of individuals(Cox, Eigenmann, and Sicherer 2021). There are many variables affecting the observed clinical cross-reactivity which are related to patient's immune response and the fish being itself. Unit operations such as canning has been seen to reduce fish protein cross-reactivity as different parts of fish contain different concentration of major proteins that could be allergenic to the consumer. There is additional amount of complexity that pervades among various types of fishes procured from different geographical areas.

Wheat Allergy: The major allergens present in wheat belong to the gliadin or glutenin family(Cox, Eigenmann, and Sicherer 2021). Majority of the people having wheat allergy would test positive to skin prick testing or SPT or wheat specific IgE, while some may have clinical reactions towards barley. Studies show that only 20% of wheat allergenic individuals show allergy towards rye or barley. Another alternative for those having wheat allergy are millet, corn , sorghum , teff and pseudocereals such as quinoa are gluten free and generally recognized as safe for consumptions. *Shell-fish Allergy*: Major causes of anaphylaxis determined by the European community are fish and shell-fish("Prevalence of Fish and Shellfish Allergy: A Systematic Review - ScienceDirect" n.d.). Allergenic reactions are usually found to be immediate and can be caused by direct ingestion or skin contact or through cooking vapors. The allergy induced by shellfish does not resolve with age and subsequent lifelong avoidance of consumption is necessary. Some allergens that include in crustacean and mollusk species are tropomyosin, arginine kinase, myosin light chain (MLC), sarcoplasmic calcium binding protein, troponin triose phosphate isomerase, paramyosin,and others. Due to the high amount of homology present among crustacean species tropomyosin a high amount of cross-reactivity has been observed.

Table 1: Showing Different Proteins Present In Different Food Items.

Food	Protein
Milk	Casein, casein, β- lactoglobulin, and α-lactalbumin
Peanut	Cupin,Prolamin and Profilin protein family
Egg	Ovalbumin, Ovomucoidand ovotransferrin
Tree- nut	2S albumins, vicilins, legumins, and nsLTPs. Bet v 1-homologues and profilins are involved in pollen associated tree nut allergy.
Soy	glycinin and beta-conglycinin
Fish	Myosin ,Actin, Tropomyosin, Troponin, Actinin, parvalbumins, aldolase and enolase
Wheat	Glutenin and Gliadin
Shell- fish	Tropomyosin

3. Mechanism of Food Allergy:

There are ubiquitous homologous proteins available across many plant and animal species. This is projected as a challenge. Many a time it is thought that a person who is allergenic to cherries may stop eating fruits or not, or whether a person allergenic to peanuts has concerns about other nuts and tree-nuts for consumption as a whole. Thus the diagnosis of food allergens has become complicated due to cross-reactivity among homologous proteins in edible foods and aero-allergens(Cox, Eigenmann, and Sicherer 2021). Patients are sensitised or give positive tests to may biological foods while many other are sensitized without any clinical reactivity. The reasons behind these positive tests may relate to sensitisation to conserved homologous protein moieties or cross-reactivity. It has become highly essential to understand the distinction between allergenic sensitization and clinical allergy and to consider for both interpretation and decision-making. Sensitization can be indicated by the presence of IgE binding to an allergen. It can be demonstrated by serum IgE testing or skin- prick testing or assays and not necessarily associated with clinical or symptomatic allergies.

Food allergy or food hypersensitivity is described as a harmful trigger response generated by antigens (Wang et al. 2020). Hypersensitivity is caused due to an inappropriate immunological response to an antigen contained in the food or food additive. It can be classified into four types mainly: Type I/II/III/IV. Food allergies are mainly type I hypersensitivities mediated by immunoglobulin E. Food allergies occur in two stages, the sensitization stage and the effector stage. The first stage of food allergy is an asymptomatic primary immune response stage. Here the mucosal immune system is exposed to the allergen and a series of immune responses are triggered. With the help of specific immune cells such as the dendritic cells,T cells, and B cells an allergen specific antibody IgE is produced. This IgE then binds onto the specific receptors on the surface of the mast cells and basophils leading to positive test or sensitized status.

The second stage is the effector stage, which is a harmful secondary response. When the body is re-exposed to the same allergen, the allergen binds to the IGE on the surface of the sensitised mast cells and basophils. Then these cells are degranulated to release inflammatory mediators such as histamine, leukotriene and prostaglandin, which finally stimulate the effector organs leading to triggered immune or allergenic responses such as vomiting, vascular collapse and even life threatening anaphylaxis(Wang et al. 2020). With thedevelopment of allergenic knowledge due to the presence of databases, it has now become possible to identify allergens by DNA or protein sequence based prediction. This in- silico approach is of relatively low cost and is extremely high throughput and therefore is attracting increasing interest for allergenomics studies.

4. Diagnosis of Food Allergy

Food allergy diagnosis is done with the help of immunoglobulins. There are several types of immunoglobulin's that are present, of which four types are commonly used for measurement. They are: Food-specific IgE, Allergen-specific IgE, Epitope-specific IgE and Food-specific IgG and IgA(Patil, Bunyavanich, and Berin 2020).

Based on World Health Organization standards, Food-specific IgE is measured in serums, while values are reported in kilounits of antibody per liter (kU/L). The most commonly used IgE laboratory test used is ImmunoCAP. In order to obtain accurate measurement of Food-specific IgE range of ImmunoCAP is 0.1 to 100 kU/L and, sera above 100 kU/L can be diluted. It is a useful tool in estimating and understanding the probable clinical reactivity to certain foods. However the results vary depending on geography and age. Allergen-specific IgE is the tool used by obtaining values from blood biomarkers which can be used as a guide to make clinical

decision in allergen-specific IgE. The measurement of IgE binding to food allergens provides the improved predictive performance compared with IgE levels to whole allergens. Epitope-specific IgE are of more clinical relevance. Epitopes are areas or regions on the allergens where there is high specificity involved in which IgE binding to sequential 15-20 amino acids peptides along the allergen is tested. It has been found out that epitope-specific IgE binding to outperform food-specific IgE binding and component- specific IgE binding in predicting allergies from milk and peanut. They have equally performed well in differentiating shellfish epitopes from cross-reactivity to inhaled mite and cockroach allergens.

The measurement of food specific IgG, IgG4 or IgA is limited for the diagnosis of food allergy and do not predict clinical reactivity well.

5. Application of Bioinformatics in Food Allergy

Bioinformatics is the science that deals with biology and computer science, and for the analysis part statistical analytical tools are employed. The information gathered is then represented in a statistical format(Kotra, Chaudhary, and Neetu 2015). Some conducted studies show that common molecular features of protein with same homology or in the same or different protein family account for clinically significant cross-reactivity and sensitivity. Computational approach is therefore necessary to understand and compare these known with unknown allergens.

There are various kinds of allergens database and bioinformatics tools present or available. Some of the popular public databases include Genbank, EMBL, DDBJ, PIR, SWISS-PROT and PDB or Protein Data Bank. The main aim of such databases is to collect, annotate and provide access to entries of sequence. Some of the manually annotated databases such as SWISS-PROT and PIR have higher detailed annotation as compared to databases such as GenPept, EMBL and DAD. Some of other include allergome that comprises of IUIS(International Union of Immunological Societies) and non-IUIS based allergens. The IFBS/ILSI (International Food Biotechnology Council and International Life Science Institute) is also a public database where list of allergen sequences are submitted. FARRP or the Food Allergy Research and Resource Program is yet another database that consists of species of origin, common and nomenclature names and accession number linked to ENTREZ. Alongside ALLALLERGY, SDAP or Structural Database of Allergenic Protein is another database that provides rapid, cross-referenced access to the sequences, structures and IgE epitopes of allergenic proteins. SDAP contains information of over eight hundred allergens that are freely available over the web to clinicians and patients. It is a relational database that contains links to other publicly available databases which is a state-of-the-art bioinformatics tool used to distinguish between allergens from non-allergens("Bioinformatics Approaches to Classifying Allergens and Predicting Cross-Reactivity"n.d.).

Since 1980's a lot of data has been collected about allergen proteins in order to establish various types of allergen databases with different types and functions(Zhou et al. 2021). The National Centre for Biotechnology Information or NCBI along with SDAP contains data about allergen sequence, epitopes, structure and homology model that can be used to further predict food allergy among individuals. Bioinformatics can be used to predict the epitopes of allergens, along with that it can help in predicting cross-reactivity between similar allergens through the known epitopes knowledge from the database. Bioinformatics is a low cost and high throughput method employed to predict the potential epitopes of a protein allergen. For example, in one of the studies ("Mapping IgE Binding Epitopes of Major Shrimp (Penaeus Monodon) Allergen with Immunoinformatics Tools - ScienceDirect" n.d.) Showed that using DNAStar, a bioinformatics tool and BepiPred 1.0 server to predict the linear IgE epitopes of tropomyosin present in crustaceans. By using three chemoinformatics tool they determined their results, DNAStar, BPAP (Bioinformatics Predicted Antigenic Peptides) and BediPred1.0 server. The results showed that BediPred 1.0 server website predicted 11 epitopes; DNAStar predicted 9 epitopes while the BPAP website showed 7 epitopes. Comparing the results from these three tools, 10 potential epitope regions were determined and solid phase synthetic peptide method and dot-blot inhibition experiment was concluded to obtain 8 tropomyosin linear epitopes. Through the deduction it was observed that aromatic amino acids easily formed epitopes when compared with aliphatic amino acids.

BLAST and FASTA databases are used to determine the amino acid sequences more promptly, likewise SDAP provides the protein family grouping for most of the proteinaceous allergens(Kotra, Chaudhary, and Neetu 2015). FASTA algorithm is usually used to find the local high scoring alignments between a nucleotide pair and a protein. BLASTp is the basic local search alignment tools for proteins and amino acids. Mutations of amino acid sequences can alter epitope of the antigen which can cause and increase or decrease in the binding efficiency of epitope and immunoglobulin(Zhou et al. 2021). Hence, bioinformatics can be combined with mutagenesis techniques to understand and predict the key amino acids on epitopes that have a binding effect on immunoglobulin. Mutagenesis techniques are mainly: shotgun mutagenesis, site-directed masking and site-directed mutagenesis.

FASTA and BLASTp can be used to for the following purposes manly: comparison of sequence of allergenic proteins, prediction of the food allergenicity and sequence combination and structural information(Kotra, Chaudhary, and Neetu 2015). Another power bioinformatics tool is molecular dynamics simulation tool(Zhou et al. 2021). It can be used to determine the binding affinity between receptors and epitopes. It includes T-cell epitopes to alleles and B-cells epitopes to antibodies. Molecular dynamics method can help us analyze food processing methods on the structural changes occurring on allergenic proteins, amino acids, and epitopes.

Results and analysis often based on computation methods have gap, which needs to be filled with need-based experiments. Zhou et al 2021 describes bioinformatics as a low cost and a high throughput method with resourceful data analysis techniques available in the market. Bioinformatics technology can help visualize protein structures with the distribution of linear epitopes and conformational epitopes on the surface of the protein moiety. Additionally bioinformatics tool can be used to identify unknown allergen proteins using *denovo* methods of analysis.

6. Conclusion and Future Perspective

Food allergy is a concerning food safety issue and is a problem worldwide. Investigation is food allergens is the most appropriate way of preventing and treating them, but relevant information is far less available. An allergy is defined as an adverse immune reaction, which is a public health concern. Most of the known allergens are proteins. Hence, finding homologous proteins that share allergenic cross-reactivity and sensitivity among conformational and linear epitopes incomparison to unknown allergenic protein is highly advantageous. Allergenic cross- reactivity caused by homologous protein moieties requires about 170% amino acid similarity. In addition, such high levels can be determined using FASTA or BLASTp analysis.

Bioinformatics plays a defining role in this scenario. Tools such as SDAP can help to find allergenic proteins and help to

determine the structural and functional relationship among proteins or allergens. They can even help in identifying cross-reacting antigens in the system. Allergy related databases are useful in collecting and accessing data. Bioinformatics based tools have furthered research progress in allergenomics and using the sequence analysis tools have created the foundations for proteomics, genomics. As a screening tool, bioinformatics can pick potential allergens out of massive data, which is relatively low in cost and less time consuming. Also, bioinformatics can provide specific information about the protein allergens which is hard to find out using traditional methods. In addition, bioinformatics can also support results of investigations carried out by traditional methods. For future development, improving the available technologies, combining the available methods and using novel strategies, we can move forward through the current limitations of allergenomics in clinical and industrial fields.

REFERENCES:

1. Albuhairi, Sultan, and Rima Rachid. 2021. "Biologics and Novel Therapies for Food Allergy." Immunology and Allergy Clinics of North America, Food Allergy, 41 (2): 271–83. https://doi.org/10.1016/j.iac.2021.01.002.
2. "Bioinformatics Approaches to Classifying Allergens and Predicting Cross-Reactivity." n.d. Accessed May 12, 2021.https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1941676/.
3. Cox, Amanda L., Philippe A. Eigenmann, and Scott H. Sicherer. 2021. "Clinical Relevance of Cross-Reactivity in Food Allergy." The Journal of Allergy and Clinical Immunology: In Practice 9 (1): 82–99. https://doi.org/10.1016/j.jaip.2020.09.030.
4. Kotra, Anoushka, Nidhee Chaudhary, and Jabalia Neetu. 2015. "Role of Bioinformatics in Food Allergens: An Overview." In .
5. "Mapping IgE Binding Epitopes of Major Shrimp (Penaeus Monodon) Allergen with Immunoinformatics Tools - ScienceDirect." n.d. Accessed May 12, 2021.https://www.sciencedirect.com/scien ce/article/pii/ S0278691511003772?via%3Dihub.
6. Patil, Sarita U., Supinda Bunyavanich, and M. Cecilia Berin. 2020. "Emerging Food Allergy Biomarkers."The Journal of Allergy and Clinical Immunology: In Practice 8 (8): 2516–24. https://doi.org/10.1016/j.jaip.2020.04.054.
7. "Prevalence of Fish and Shellfish Allergy: A Systematic Review - ScienceDirect." n.d. Accessed May 12, 2021. https://www.sciencedirect.com/scien ce/article/pii/S1081120616304276.
8. "The Natural History of Soy Allergy - ScienceDirect." n.d. Accessed May 12, 2021.https://www.sciencedirect.com/scien ce/article/pii/ S0091674910000072.
9. Umasunthar, T, J Leonardi-Bee, M Hodes, P J Turner, C Gore, P Habibi, J O Warner, and R J Boyle. 2013. "Incidence of Fatal Food Anaphylaxis in People with Food Allergy: A Systematic Review and Meta-Analysis." Clinical and Experimental Allergy 43 (12): 1333– 41. https://doi.org/10.1111/cea.12211. Wang, Chong, Yanbo Wang, Guangming Liu, and Linglin Fu. 2020. "Food Allergomics Based on High- Throughput and Bioinformatics Technologies." Food Research International 130 (April): 108942. https://doi.org/10.1016/j.foodres.2019.108942.