Bioinformatics Approaches in Food Industry: An Overview

Kamran Manzoor Waidha[1*], Neetu Jabalia[2], Divyam Singh[3‡], Adya Jha[4] and Rupsneet Kaur[5‡]

[1,2,3,4,5]*Amity Institute Of Biotechnology Amity University Sector-125 Noida-201313*
E-mail: [1]kamranwaidha1@gmail.com [2]njabalia@amity.edu [3]divyamsingh711@gmail.com
[4]adya.j11@gmail.com [5]rupsneetk@yahoo.com

Abstract—*Bioinformatics is use of modern computer sciences in management of biological data so that gathered biological information can be applied extensively in biotechnology. The knowledge of Bioinformatics can be used for ease of access to all genomic and proteomic data (sequence) discovered so far and uploading this data to make this accessible to every individual group industry or company so as to increase quality and nutritional value of food that is to be produced. Being the important part of food research with its help we can predict the behavior of normal organisms or GMOs and products that they can produce when kept in particular condition in food industry (e.g. yeast) using molecular dynamic simulation In addition to this prediction of the amino acid sequence produced by specific genes can be studied so that its effect or side-effect can be predicted and genome of crops can be altered accordingly by genetic engineering to produce desired product. Nutrient data has been created by various organizations where information about nutritional values of various food items has been uploaded and is accessible through internet. Using bioinformatics method is economically more effective than using lab kits. In recent days the algorithms for data integration and analysis have been developed to decrease the level of error.*

Keyword:- *Bioinformatics Genomic Proteomic GMO Molecular dynamic simulation Genetic engineering food.*

1. INTRODUCTION

Bioinformatics can be explained as the utilization of computer science and its resources to the biological data mainly concerned with data related to sequences of DNA and proteins. With various success stories accredited to the application of bioinformatics approaches It prospects to be beneficial in food sciences. Bioinformatics field is expanding at an alarming rate and becoming a very essential tool for research in life sciences by integrating the traditional techniques with massive amount of data generated from the fields of comparative genomics genomics and proteomics in food research [1]. It has made easier to evaluate allergenic potential of normal proteins by playing an important role in safety assessment of genetically modified crops because it is important to have safety from food allergy [2]. This technique help in the prediction of functionality and allergenicity of food products by analyzing the protein sequence of their ingredients [3]. In practical context comparative genomics technique of bioinformatics has been applied to many food related pathogens significantly to food and sources related to its production have been the center of interest of many sequencing and comparative genomic research projects [4]. It is believed that such studies can majorly play important role in prevention of crop related disease[5] and poisoning of food [6]. With this increasing the information related to the various processes and properties like crop improvement flavor formation and ripening [7 8 9 10].

Crops are important part of food industry so crop must be of good quality (i.e. high yielding and disease resistant). With the help of bioinformatics we can identify the genes in the commercially important crops using genomics and functional genomics it can also help in the development of transgenic crops and in the identification of new genes that can increase quality and quantity of food products produced by these crops. This technique can be also beneficial in development of agro-chemicals based on the concepts of signal perception transduction pathways for specific targets and identification of compounds that can be potentially used as pesticide herbicide or insecticide [11]. This review will primarily be focusing on the advances future possibilities and role of bioinformatic approaches in food industry.

2. WHY WE NEED BIOINFORMATICS FOR FOOD?

Food plays an important role in regulating the various processes within the body like metabolic mental etc. However there is always growing admiration for the role nutrition plays in halting the growth of various chronic disease. Significant efforts are made to encourage and increase the nutritional potential and quality of food sources. Food science has developed outstandingly accepting various emanate techniques like " omics" suite. Due to these diversification it has started to face difficulties in interpretation and amalgamation of various varied data types so it is important to develop and establish an database which can store and keep updating the vast amount of biological data and resources which are important to food sciences as well as for the nutritional

sciences. It is believed that the development of bioinformatics in food will help in providing the simple and convenient ways for researchers to accept a new field or discipline[12].

2.1. Detection of Allergen

Due to very distinct origin of allergens they have very large sequence and similarity in the structure causing equivalent responses of IgE. It wont be unexpected that bioinformatics is the most necessary resource for food allergen related studies. Various databases which are primarily dedicated to the food allergen focusing on food exist such as AllerMatch[13] Informall FARRP Allergen database and SDAP [14] show the use of bioinformatics in this area.

Homology of prevalent bioinformatics with inclusion of structural bioinformatics are systematically performed to build up the possible allergenicity and cross reactivity of proteins [15]. In fact the utility of such methodologies has incited the WHO to incorporate sequences similarity search as a feature of the rules for evaluating allergenicity for food thats genetically modified [16]. Recently various techniques of bioinformatics have been utilized for allergen diagnostic development to predict the peanut allergy with the help of machine learning [17].


Fig. 1: FARRP Allergen database homepage.

2.2. Bioinformatics in ensuring Quality taste and safety of food

Bioinformatics is playing an important role in ensuring the quality taste flavour of the food while maintaining its safety. Working in relation with the molecular evolution bioinformatics played an essential role in study of evolution of receptors for taste [18].With various studies being conducted primarily focusing on the taste receptors with the link between the glucose regulation and bitter taste receptors established[19]. Recently electronic database was established which include the chemical properties of various compounds related to their taste and flavour [20]. Moreover study of genetic sequence of lactic acid bacteria played an important role in uncovering the formation of specific flavouring potential helping in giving flavour to many fermented foodstuffs[21].

Bioinformatics is growing while playing crucial importance in food safety and quality. Due to its potential many genomic sequencing projects are targeting on the food-borne pathogens[22]. With the development of genomic sequencing technologies bioinformatics has come up with an innovative way which will help in determining the source of the food-borne illness[23]. For instance recently developed tool by the FDA (Food and Drug Administration) helps in detection of the bacterial food pathogen where as other computer based tools are being developed focusing on prediction of growth of microbes within a given food source[24]. In regards to ensuring food quality progress is made with the help of bioinformatics that allows in detection of various properties of food automatically. [26].
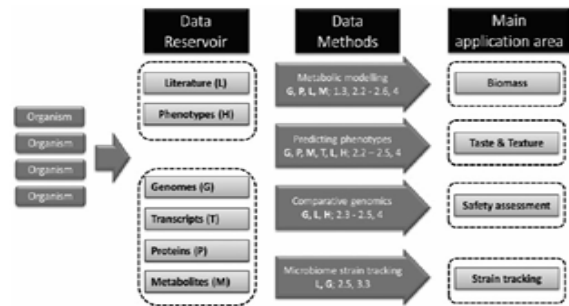
Fig. 2: Applied Bioinformatics In Food[25]

3. UNCOVER BIOACTIVE PEPTIDES

Bioinformatics recently got to divulge in the discovery of bioactive peptides present within the food proteins[27 28]. Due to its simultaneous evaluation of different food proteins and the enzymes which help in the digestion of proteins i.e proteolytic enzymes bioinformatics can help in bringing the major change in bioactive peptide research.The approach includes the utilization of data present in various databases like BIOPEP[29]. The protein sequences from UniProtKB SwissProt and TreMBL database can be obtained for determining the occurring frequency of bioactive peptide crypt in primary structures of food proteins. The occurance frequency can be calculated by a/N where "a" stands for the no. of peptides exhibiting bioactivity of particular type present in sequence and "N" stands for the total no. of Amino acids residues present within the proteins. Bioinformatics developed softwares are used to generate profiles for *in silico* peptides resulted from the stimulation of proteolytic specificities of various enzymes[30]. Various *in silico* tools for proteolysis are:- BIOPEP ExPASY and PoPS. Ideally this process is meticulous and reduces the time which was required earlier to screen for bioactive peptides which are present in distinctive protein sources by utilizing a few proteases and which can prompt the discovery of new and credible forerunners of known bioactive peptides

4. BIOINFORMATICAL APPROACH IN CROP PRODUCTION

Genetic research in plants have provided aid in increasing the success rate of production of genetically modified crops of general regard for food industry. Main objective of production of Genetically Modified Crops is to improve quality of raw materials of food supply for better processability lowering cost and safer food [31]. Genome sequencing projects help in identification of plant biosynthetic genes of health importance. Thus genome research can be directly implemented to increase the efficiency and efficacy in breeding of plants for improvement [32].

Theobroma cacao i.e. Cocoa is used as a raw material for almost every food product containing chocolate. It has been difficult in past to recognise a seed of higher quality and good flavor. Trees are supposed to be maintained for 3-5 years before seeds can be harvested. With implementation of DNA fingerprinting in screening of plant collection RFPL markers for detection of genotypic connection between breeds and detection and identification of more than 300 markers breeding programs has been extensively enhanced. The further availability of EST (Expressed Sequence Tags) sequences and genetic comparisons to other identified plants which depends on bioinformatics techniques will further increase the possibility to select desired traits in early stage of plant's development on basis of genotype and phenotype [33].

5. BIOINFORMATICAL APPROACH IN FOOD PROCESSING

Implementation of bioinformatics tool kit and intellect in food processing will optimise the quantitative and compositional parameters of traditional operations. Food products are processed to increase storage stability and safety. But due to lack of knowledge of composition and structural complexity of biological materials the variability of living organisms (generally microbes) as food process input material and response of these materials on processing parameters brings out large array of errors. With the use knowledge of organisms for food processing and acquired aid from bioinformatics the processing will become optimized with lesser chances of error

The food processing involves integration of biological intellect with the knowledge of biomaterials. The traditional techniques of food processing involved high energy input in reconstruction of biomaterials of used organisms to simpler macrostructures forms of stable food. But in most cases inherent biological properties are lost in final product while elimination of hazardous properties of some molecules. The intellect build on bioinformatics illustrates characteristic features for complexity of biomolecules and their functions such knowledge paves way to circumstances that yield to precise and apt assessment of biomaterial properties of each of the molecules in complex mixtures. For example. The unabridged protein-protein synergy map of yeast i.e. all possible interactions between the 6000 proteins of yeast has been completed [34]. The constructional and utilitarian features of living organisms that are being researched and examined with the aid of bioinformatics will progressively affect the new food processes and food-products.

6. CONCLUSION

Bioinformatics shows the potential for diverse applications in food industry and offers almost unlimited scope for the development of new methods to implement them. The huge amounts of data that can be generated from the outlined projects which are to take place will make bioinformatics an important area in the food industry. Attempts are being made via using bioinformatics tool kit to genetically engineer transgenic plants to obtain healthy nutritious and beneficial foodstuffs and crops for public to rage for an ultimate breakthrough for the food industry. All the fields and arenas of food industry will benefit from the efficacy efficiency accuracy and speed that the breakthrough will provide. The beauty and future potential of the bioinformatics approach to food industry depends on the design of multiple models based on our increasing knowledge of genes and their functions. However the full commitment of this promise is yet to be established.

REFERENCES

[1] Frank Desiere Bruce German Heribert Watzke Andrea Pfeifer Sam Saguy Bioinformatics and data knowledge: the new frontiers for nutrition and foods Trends in Food Science & Technology Volume 12 Issue 7 July 2001 Pages 215-229 ISSN0922244 http://dx.doi.org/10.1016/S09242244(01)00089-9.

[2] Gregory S. Ladics Robert F. Cressman Corinne Herouet-Guicheney Rod A. Herman Laura Privalle Ping Song Jason M. Ward Scott McClain Bioinformatics and the allergy assessment of agricultural biotechnology products: Industry practices and recommendations Regulatory Toxicology and Pharmacology Volume 60 Issue 1 1 June 2011 Pages 46-53 ISSN 0273-2300 http://dx.doi.org/10.1016/j.yrtph.2011.02.004.

[3] Bioinformatics may isolate food allergy proteins By Chris Mercer 05-Aug-2005.

[4] Mayer K. F. X. Martis M. Hedley P. E. Simkova H. Liu H. Morris J. A. et al. (2011). Unlocking the barley genome by chromosomal and comparative genomics. The Plant Cell Online 23 1249-1263.

[5] Qian W. Jia Y. Ren S.X. He Y.Q. Feng J.-X. Lu L.-F. et al.(2005). Comparative and functional genomic analyses of the pathogenicity of phytopathogen Xanthomonas campestris pv.campestris. Genome Research 15 757-767.

[6] Abee T. Van Schaik W. & Siezen R. J. (2004). Impact of genomics on microbial food safety. Trends in Biotechnology 22 653-660.

[7] Fei Z. Tang X. Alba R. M. White J. A. Ronning C. M. Martin G. B. et al. (2004). Comprehensive EST analysis of tomato and comparative genomics of fruit ripening. The Plant Journal 40 47-59.

[8] Liu M. Nauta A. Francke C. & Siezen R. J. (2008). Comparative genomics of enzymes in flavor-forming pathways from amino acids in lactic acid bacteria. Applied and Environmental Microbiology 74 4590-4600.

[9] Mochida K. & Shinozaki K. (2010). Genomics and bioinformatics resources for crop improvement. Plant and Cell Physiology 51 497-523.

[10] Varshney R. K. Graner A. & Sorrells M. E. (2005). Genomics assisted breeding for crop improvement. Trends in Plant Science 10 621-630.

[11] Bennetzen J. L. Kellogg E. A. Lee M. & Messing J. (1998). A plant genome initiative. The Plant Cell 10(4) 488-493.

[12] Lemay D. G. Zivkovic A. M. & German J. B. (2007). Building the bridges to bioinformatics in nutrition research. The American Journal of Clinical Nutrition 86 1261-1269.

[13] Fiers M. Kleter G. Nijland H. Peijnenburg A. Nap J. & Van Ham R. (2004). AllermatchTM a webtool for the prediction of potential allergenicity according to current FAO/WHO Codex alimentarius guidelines. BMC Bioinformatics 5 133.

[14] Mari A. Scala E. Palazzo P. Ridolfi S. Zennaro D. & Carabella G. (2006). Bioinformatics applied to allergy: allergen databases from collecting sequence information to data integration. The allergome platform as a model. Cellular Immunology 244 97-100.

[15] Jenkins J. A. Griffiths-Jones S. Shewry P. R. Breiteneder H. & Mills E. N. C. (2005). Structural relatedness of plant food allergens with specific reference to cross-reactive allergens: an in silico analysis. Journal of Allergy and Clinical Immunology 115 163-170.

[16] FAO/WHO. (2001). Evaluation of allergenicity of genetically modified foods. Report of a joint FAO/WHO expert consultation on 14 T.A. Holton et al. / Trends in Food Science & Technology 34 (2013) 5-17 allergenicity of foods derived from biotechnology. Rome: Food and Agriculture Organization of the United Nations (FAO).

[17] Lin J. Bruni F. M. Fu Z. Maloney J. Bardina L. Boner A. L. et al. (2012). A bioinformatics approach to identify patients with symptomatic peanut allergy using peptide microarray immunoassay. Journal of Allergy and Clinical Immunology 129 1321-1328.e5.

[18] Dong D. Jones G. & Zhang S. (2009). Dynamic evolution of bitter taste receptor genes in vertebrates. BMC Evolutionary Biology 9 12.

[19] Dotson C. D. Zhang L. Xu H. Shin Y.-K. Vigues S. Ott S. H. et al. (2008). Bitter taste receptors influence glucose homeostasis. PLoS ONE 3 e3974.

[20] Wiener A. Shudler M. Levit A. & Niv M. Y. (2012). BitterDB: a database of bitter compounds. Nucleic Acids Research 40 D413-D419.

[21] Liu M. Nauta A. Francke C. & Siezen R. J. (2008). Comparative genomics of enzymes in flavor-forming pathways from amino acids in lactic acid bacteria. Applied and Environmental Microbiology 74 4590-4600.

[22] Brul S. Schuren F. Montijn R. Keijser B. J. F. Van Der Spek H. & Oomes S. J. C. M. (2006). The impact of functional genomics on microbiological food quality and safety. International Journal of Food Microbiology 112 195-199.

[23] Lienau E. K. Strain E. Wang C. Zheng J. Ottesen A. R. Keys C. E. et al. (2011). Identification of a salmonellosis outbreakby means of molecular sequencing. New England Journal of Medicine 364 981-982.

[24] García-Gimeno R. M. A. Hervas-Martinez C. & De Silonizc M. I.(2002). Improving artificial neural networks with a pruning methodology and genetic algorithms for their application in microbial growth prediction in food. International Journal of Food Microbiology 72 19-30.

[25] Alkema W. Boekhorst J. Wels M. & S. A. F. T. Van Hijum. (2015). Microbial bioinformatics for food safety and production. *Briefings in Bioinformatics*.

[26] Mery D. Pedreschi F. & Soto A. (2012). Automated design of a computer vision system for visual food quality evaluation. Food and Bioprocess Technology 1-16.

[27] Carrasco-Castilla J. Hernandez-Alvarez A. J. Jimenez-Martınez C. Gutierrez-Lopez G. F. & Davila-Ortiz G. (2012). Use of proteomics and peptidomics methods in food bioactive peptide science and engineering. Food Engineering Reviews 4 224-243.

[28] Lacroix I. M. & Li-Chan E. C. (2012). Evaluation of the potential of dietary proteins as precursors of dipeptidyl peptidase (DPP)-IV inhibitors by an in silico approach. Journal of Functional Foods 4 403-422.

[29] Dziuba J. Minkiewicz P. Na1e̦cz D. & Iwaniak A. (1999).Database of biologically active peptide sequences. Nahrung 43 190-195.

[30] Chibuike C. Udenigwe Bioinformatics approaches prospects and challenges of food bioactive peptide research Trends in Food Science & Technology Volume 36 Issue 2 April 2014 Pages 137-143 ISSN 0924-2244.

[31] Dellapenna D. (1999). Nutritional Genomics: Manipulating Plant Micronutrients to Improve Human Health. Science 375-379.

[32] Frank Desiere Bruce German Heribert Watzke Andrea Pfeifer Sam Saguy Bioinformatics and data knowledge: the new frontiers for nutrition and foods Trends in Food Science & Technology Volume 12 Issue 7 July 2001 Pages 215-229 ISSN 0924-2244.

[33] R.D Pridmore D Crouzillat C Walker S Foley R Zink M.-C Zwahlen H Brüssow V Pétiard B Mollet Genomics molecular genetics and the food industry Journal of Biotechnology Volume 78 Issue 3 31 March 2000 Pages 251-258 ISSN 0168-1656.

[34] Ito T. Chiba T. Ozawa R. Yoshida M. Hattori M. & Sakaki Y. (2001). A comprehensive two-hybrid analysis to explore the yeast protein interactome. Proceedings of the National Academy of Sciences 4569-4574.