

# An *in silico* Investigation of Anticancer Peptide Candidates in Fermented Food Microbiomes

Muzaffer Arikan <sup>1,2</sup> 

<sup>1</sup>Research Institute for Health Science and Technologies (SABITA), Istanbul Medipol University, Istanbul, Turkiye

<sup>2</sup>Department of Medical Biology, Faculty of Medicine, Istanbul Medipol University, Istanbul, Turkiye

ORCID ID: M.A.0000-0001-5162-2000

**Cite this article as:** Arikan M. An *in silico* investigation of anticancer peptide candidates in fermented food microbiomes. Experimed 2023; 13(1): 64-72.

## ABSTRACT

**Objective:** Cancer is a leading cause of death worldwide, requires development of new effective, specific, and safe strategies that do not carry the disadvantages of traditional cancer treatment approaches. Hence, this study aimed to identify anticancer peptide candidates in fermented food microbiomes through an *in silico* investigation.

**Materials and Methods:** One hundred eight shotgun metagenomic sequencing samples from six studies on fermented food microbiomes were downloaded from the NCBI and ENA databases and included in the study. Bioinformatic analyses including quality control of raw data, *de novo* assembly, prediction of protein sequences, anticancer peptide predictions by an integrated use of four different prediction tools, toxicity predictions and database comparisons were performed.

**Results:** One hundred forty-two novel anticancer peptide candidates were identified. Liquor, coffee, kefir fermentation samples contained the greatest numbers of anticancer peptide candidates while sugar, dairy, coconut kefir and brine-type fermentations were dominant sources according to the substrate type.

**Conclusion:** This study indicates the potential of fermented food microbiomes as a useful source for candidate anticancer peptide detection. *In vitro* and *in vivo* validations of detected peptides may lead to development of new candidate molecules for cancer therapy in the future.

**Keywords:** Fermented food, microbiome, bioinformatics, metagenomics, anticancer peptide

## INTRODUCTION

Cancer is a leading cause of death worldwide and responsible for an estimated 9.6 million deaths each year (1). According to the Global Cancer Observatory, the most common types of cancer are lung, breast and colorectal cancer (2). Traditional approaches to cancer management, such as radiotherapy, chemotherapy, surgery, can have significant negative effects (3,4). Considering the disadvantages of current treatment approaches such as medical complexities, adverse effects and high treatment

costs, the development of effective, specific and safe new strategies is of paramount importance in terms of public health and economic benefits (5).

The focus on peptides as potential anticancer agents gained momentum after studies reporting antimicrobial peptides with varying levels of activity against tumor cells (6). Anticancer peptides show their cytotoxic activities in a similar way to antimicrobial peptides. Many anticancer peptides destroy cancer cells via apoptosis and necrosis by membrane lysis or pore formation (7). Anticancer peptides

**Corresponding Author:** Muzaffer Arikan **E-mail:** muzafferarikan@gmail.com

**Submitted:** 09.03.2023 **Revision Requested:** 27.03.2023 **Last Revision Received:** 27.03.2023 **Accepted:** 04.04.2023 **Published Online:** 07.04.2023



Content of this journal is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

have many advantages such as crossing biological barriers more easily, broad target range, less side effects, less accumulation in tissues and lower toxicity which make them stand out as potential anticancer agents when compared to currently used chemotherapy drugs (8). Peptides with anticancer properties and desirable properties such as specificity, solubility, tumor penetration and safety are among the new alternatives suitable for cancer therapy (9).

Fermented foods have been part of the human diet for several millennia and are consumed in a variety of ways around the world. Fermented foods have been associated with alleviating various health problems in humans (10), and these properties have been attributed to bioactive compounds formed as a result of microbial fermentation (11). In recent years, there has been a growing number of reports on the anticancer effects of fermented foods (12). Anticancer peptides derived from fermented foods have the potential of being used as suitable alternatives to traditional cancer management approaches (5), many computational methods for anti-cancer peptide identification have been developed in the last decade (13). However, considering the massive amount of data produced by high throughput genomics technologies for fermented foods, there is a strong need for efficient strategies for the prediction and testing of anticancer peptide candidates in fermented food microbiomes.

In this study, shotgun metagenomics based fermented food microbiome samples reported by different studies were analyzed to conduct an *in silico* investigation of anticancer peptide candidates encoded within the genomes of the microbiome members using a bioinformatics workflow including metagenome assembly, combined use of anticancer peptide prediction tools, toxicity and novelty analyses.

## MATERIALS AND METHODS

One hundred eight shotgun metagenomic sequencing samples from six studies focusing on fermented food microbiomes were downloaded from the National Center for Biotechnology Information (NCBI), and the European Nucleotide Archive

(ENA) databases and included in the study. The overall analysis strategy of the study is presented in Figure 1. Quality control of the data was first performed to clean raw sequences. Using trimmomatic, the adapter sequences were removed, the low-quality ends of the sequences were trimmed from the position with the quality score value below 20, and the sequences with a length shorter than 30 bases were eliminated. Quality controlled and filtered clean sequences were used for metagenome assembly analysis which was performed using the default parameters of the metaSPAdes (v3.15.0) (14) and contig sequences were obtained. Protein-coding sequences (CDS) for each of these contigs were predicted and annotated using Prokka (v1.4.0) with option `-metagenome` (15). Thereafter, CDS with the length between 15 amino acids and 50 amino acids were determined using SeqKit (v2.3.0) (16) and used for downstream analyses. The anticancer activity potentials of these peptides were estimated by four different anticancer peptide prediction tools, AntiCP2 (v2.0) (17), mACPred (v1.0) (18), ACP-MHCNN (v1.0) (19), ACPred (v1.0) (20) using default options for each tool. By combining the results obtained from these tools, the anticancer peptides predicted commonly by all four prediction tools were determined and considered as a reliable set. Next, ToxinPred (v1.0) (21) was used to carry out an *in silico* toxicity analysis. Finally, CancerPPD (v1.0) (22), a manually curated database of experimentally validated anticancer peptides, was used to check the novelty of anticancer peptide, candidates with minimum similarity threshold of 100% and the peptides with the highest anticancer property score and the lowest toxicity were determined.

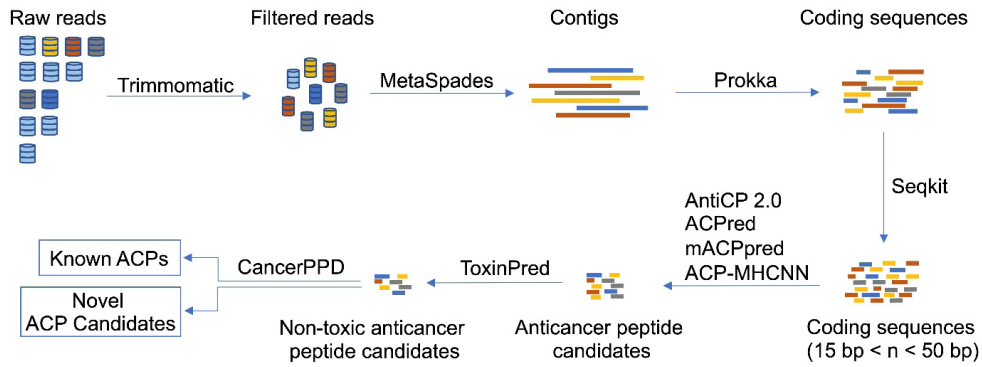
## RESULTS

One hundred eight shotgun metagenomics samples from six studies focusing on fermented food microbiomes were included in this study. Fermented food samples represented five main substrate types, namely brine, coconut kefir, dairy, soy, and sugar (Figure 2). Anticancer peptides candidates were recovered in samples from a total of 13 countries (Cote d'Ivoire, Saudi Arabia, Ecuador, Egypt, Benin, Russia, United Kingdom, Germany, Mexico, China, Ireland, Japan, and Turkiye) across five continents (Asia, Africa, Europe, North America, South America)

**Table 1.** General characteristics of sequencing data at different steps of bioinformatics analysis.

Study ID	Raw	QC Filtered*	Contigs	CDS*	CDS (15<n<50)	Reference**
PRJEB19968	2,287,242	1,824,414	48,044	36,241	3,348	-
PRJEB21086	26,002,221	22,511,121	86,532	34,884	3,550	-
PRJEB22200	2,275,543,745	1,991,566,237	608,723	308,949	43,830	(29)
PRJEB24129	56,774,402	46,729,715	3,669,361	1,349,589	183,550	(30)
PRJEB35321	347,841,507	263,627,918	7,102,933	2,076,255	305,492	(23)
PRJNA260163	1,683,868,773	1,425,472,508	12,142,357	7,586,687	444,668	-
<b>Total</b>	<b>4,392,317,890</b>	<b>3,751,731,913</b>	<b>23,657,950</b>	<b>11,392,605</b>	<b>984,438</b>	

\*QC: Quality Control Filtered, CDS: Protein Coding Sequence, \*\*3 studies did not have any associated publication in the NCBI and ENA databases



**Figure 1.** Overall analysis strategy.

(Figure 2). The majority of samples were collected in China (n=24), Ireland (n=24) and Turkiye (12).

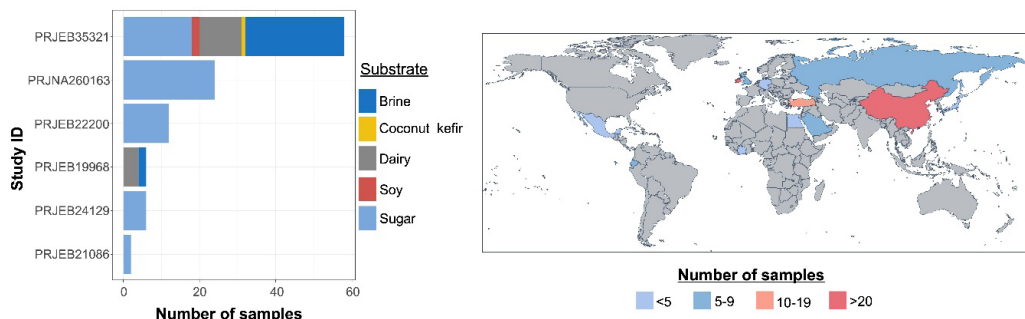
To predict anticancer peptide candidates in fermented food microbiomes, a total of 4.4 billion reads were analyzed (Table 1). Metagenome assemblies yielded 23.6 million contigs. Then, CDS for each of these contigs were predicted and a combined set of 11.4 million CDS were obtained among which 984,438 short CDS (between 15 amino acids and 50 amino acids in length) were used for prediction of anticancer peptide candidates.

Anticancer peptide predictions were performed by combining four different anticancer peptide prediction tools, namely AntiCP2, mACPred, ACP-MHCNN and ACPred. ACPred predicted the highest number of anticancer peptide candidates (n=80,646) while AntiCP2 predicted the lowest number of anticancer peptide candidates (n=7,570; Figure 3). Since the correct identification of anticancer peptide candidates vary between different methods, only 168 anticancer peptide candidates predicted by all tools were considered as reliable and used for downstream analyses. The overlap between all four prediction tools was approximately 0.001% (168 anticancer peptide candidates in 133,263 anticancer peptide candidates predicted in total).

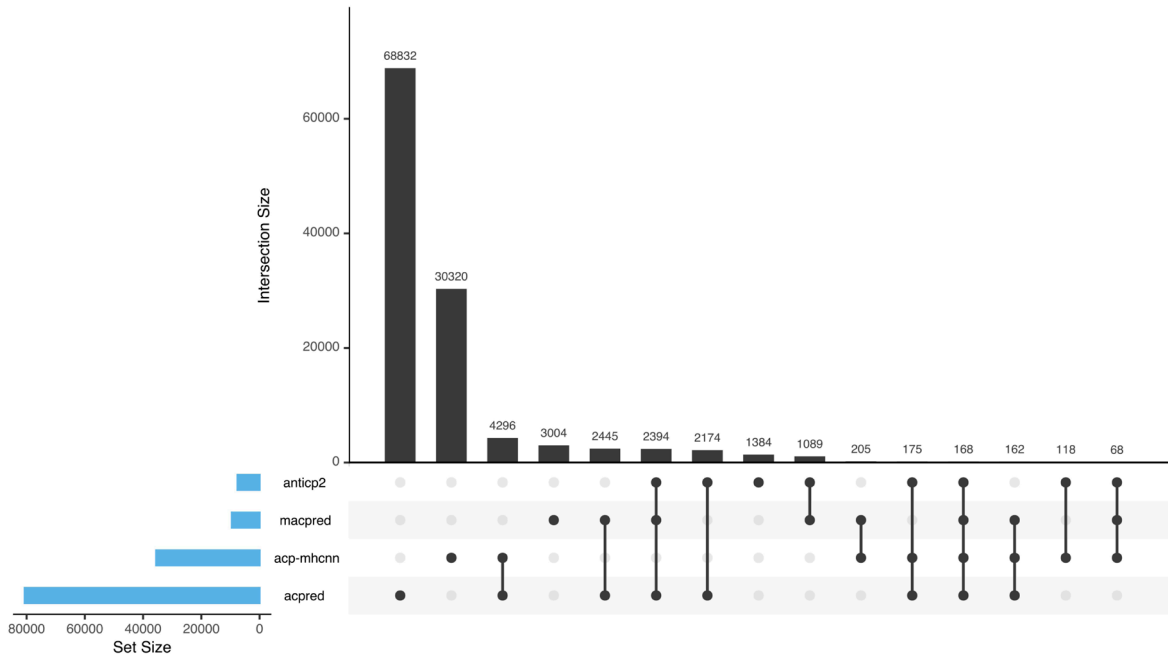
One hundred sixty-eight anticancer peptide candidates commonly predicted by all prediction tools were further analyzed to assess their anticancer activity potential. These anticancer peptides were first tracked back to the studies and samples they originated from. The results showed liquor, coffee, and kefir fermentation samples as main sources of these anticancer peptides while sugar and brine type fermentations were main substrate types (Figure 4). Next, the potential toxicity and novelty of 168 anticancer peptide candidates were examined which resulted in 142 anticancer peptide candidates predicted as non-toxin and have not been reported to date (Supplementary File 1).

## DISCUSSION

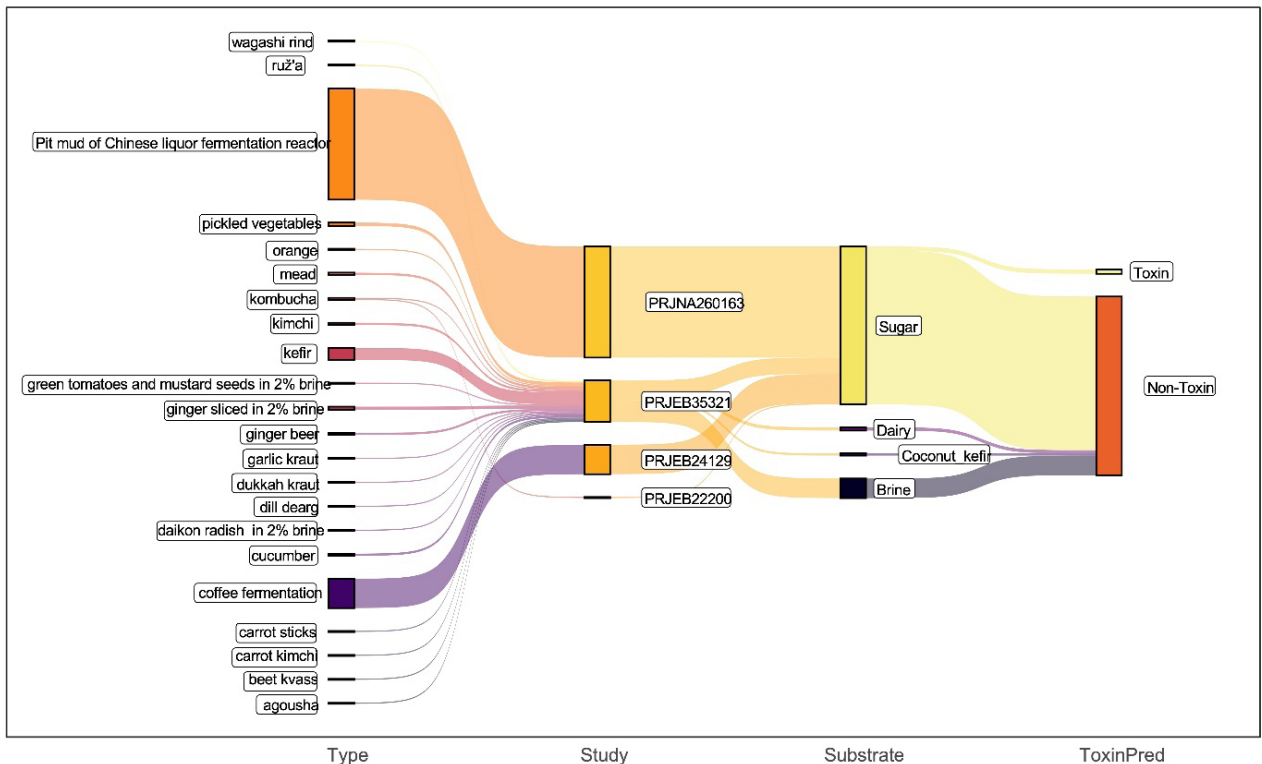
Microorganisms involved in the fermentation process play an important role in the formation of health associated properties of fermented foods (23). It is known that the bioactive molecules produced by these microorganisms could have anti-inflammatory, antifungal, antibiotic, or anticancer properties (12). Thus, examination of microbial genomes can lead to the discovery of new biological agents used in treatment of a variety of diseases, or facilitating biotechnological process. Considering the increasing number of reports on the anticancer effects of fermented food (12), the genomes of microorganisms participating in the fermentation process may be a potential reservoir for novel anticancer peptides. Thanks to the growing



**Figure 2.** Number of samples for each study used to predict anticancer peptide candidates, colored according to fermented-food substrates. Geographic distribution of the number of samples retrieved per country.



**Figure 3.** Number of predicted anticancer peptide candidates across anticancer peptide prediction tools. Vertical bars represent the number of predicted anticancer peptide candidates shared between the specific tools highlighted with connected dots in the lower panel. Horizontal bars in the lower panel indicate the total number of anticancer peptide candidates predicted by each prediction tool.



**Figure 4.** Sankey plot displaying distribution of 168 anticancer peptide candidates across fermented food, study, substrate type and predicted toxicity categories.

use of the next generation sequencing technologies and development of bioinformatics tools, the microbial genomes can be screened for anticancer peptide candidates through much more comprehensive approaches (13).

In this study, microbiomes of many fermented food types originating from different countries were analyzed together to detect anticancer peptide candidates. Starting with billions of DNA sequencing reads, 142 novel anticancer peptide candidates have been predicted and reported. One of the findings of the study was the high-level inconsistency between different anticancer peptide prediction tools which suggests that caution is needed when using a single anticancer peptide prediction tool. Thus, a consensus approach based on multiple anticancer peptide prediction tools -as applied in this study- can be used to alleviate this issue and ensure robust predictions. Moreover, the size distribution after length-based filtering of CDS yielded peptide sequences with lengths ranging from 30 amino acids to 50 amino acids and potentially caused missing shorter anticancer peptide candidates. As this is a known limitation related to software used for CDS prediction and annotations, new tools for small open reading frames would serve as valuable tools to overcome this limitation in the future (24).

The detected anticancer peptide candidates mainly originated from liquor, coffee and kefir fermentations which are classified as sugar, dairy and brine type fermentations. Among these fermented foods, interestingly, kefir has been reported to have anticancer effects by several studies (25–28). This distribution could be partly attributed to the number of raw reads obtained from the samples; however, interestingly, there was very low number of anticancer peptide candidates detected in other samples sequenced with very high coverage such as kombucha samples. It should also be noted that the microbial diversity in the samples potentially have significant effects on the number of the predicted anticancer peptides. The toxicity analysis and comparison with the previously characterized anticancer peptides revealed that most of the predicted anticancer peptide candidates have not been reported before. Considering the high throughput characteristics of the approach applied in this study, comprehensive high throughput wet laboratory testing, and characterization methods will be strongly needed in the future.

Fermented food microbiomes were found to be a useful source for candidate anticancer peptide discovery. *In vitro* and *in vivo* validations of the identified peptides may lead to development of new candidate molecules that can be used in cancer therapy.

**Ethics Committee Approval:** Ethics committee approval is not required because of no material or experimental animal that would require permission.

**Peer-review:** Externally peer-reviewed.

**Conflicts of Interest:** The author declare no conflict of interest.

**Financial Disclosure:** The author declare that this study has received no financial support.

## REFERENCES

1. Avilés-Gaxiola S, Gutiérrez-Grijalva EP, León-Felix J, Angulo-Escalante MA, Heredia JB. Peptides in colorectal cancer: Current state of knowledge. *Plant Foods Hum Nutr* 2020; 75(4): 467–76. [\[CrossRef\]](#)
2. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2018; 68(6): 394–424. [\[CrossRef\]](#)
3. Ellis MA, Graboyes EM, Wahlquist AE, Neskey DM, Kaczmar JM, Schopper HK, et al. Primary surgery vs radiotherapy for early stage oral cavity cancer. *Otolaryngol - Head Neck Surg (United States)* 2018; 158(4): 649–59. [\[CrossRef\]](#)
4. Lobb RJ, Jacobson GM, Cursons RT, Jameson MB. The interaction of selenium with chemotherapy and radiation on normal and malignant human mononuclear blood cells. *Int J Mol Sci* 2018; 19(10): 3167. [\[CrossRef\]](#)
5. Sharma P, Kaur H, Kehinde BA, Chhikara N, Sharma D, Panghal A. Food-derived anticancer peptides: a review. *Int J Pept Res Ther* 2021; 27(1): 55–70. [\[CrossRef\]](#)
6. Hoskin DW, Ramamoorthy A. Studies on anticancer activities of antimicrobial peptides. *Biochim Biophys Acta* 2008; 1778(2): 357–75. [\[CrossRef\]](#)
7. Chiangjong W, Chutipongtanate S, Hongeng S. Anticancer peptide: Physicochemical property, functional aspect and trend in clinical application. *Int J Oncol* 2020; 57(3): 678–96. [\[CrossRef\]](#)
8. Pan X, Xu J, Jia X. Research progress evaluating the function and mechanism of anti-tumor peptides. *Cancer Manag Res* 2020; 12: 397–409. [\[CrossRef\]](#)
9. Shoombuatong W, Schaduangrat N, Nantasenamat C. Unraveling the bioactivity of anticancer peptides as deduced from machine learning. *EXCLI J* 2018; 17: 734–52.
10. Baruah R, Ray M, Halami PM. Preventive and therapeutic aspects of fermented foods. *J Appl Microbiol* 2022; 132(5): 3476–89. [\[CrossRef\]](#)
11. Martínez-Villaluenga C, Peñas E, Frias J. Bioactive peptides in fermented foods. In: *Fermented foods in health and disease prevention*. Elsevier; 2017. p. 23–47. [\[CrossRef\]](#)
12. Tasdemir SS, Sanlier N. An insight into the anticancer effects of fermented foods: A review. *J Funct Foods* 2020; 75: 104281. [\[CrossRef\]](#)
13. Liang X, Li F, Chen J, Li J, Wu H, Li S, et al. Large-scale comparative review and assessment of computational methods for anti-cancer peptide identification. *Brief Bioinform*. 2021; 22(4): bbaa312. [\[CrossRef\]](#)
14. Nurk S, Meleshko D, Korobeynikov A, Pevzner PA. MetaSPAdes: A new versatile metagenomic assembler. *Genome Res*. 2017; 27(5): 824–34. [\[CrossRef\]](#)
15. Seemann T. Prokka: Rapid prokaryotic genome annotation. *Bioinformatics* 2014; 30(14): 2068–9. [\[CrossRef\]](#)
16. Shen W, Le S, Li Y, Hu F. SeqKit: A cross-platform and ultrafast toolkit for FASTA/Q file manipulation. *PLoS One* 2016; 11(10): e0163962. [\[CrossRef\]](#)
17. Agrawal P, Bhagat D, Mahalwal M, Sharma N, Raghava GPS. AntiCP 2.0: An updated model for predicting anticancer peptides. *Brief Bioinform* 2021; 22(3): bbaa153. [\[CrossRef\]](#)
18. Boopathi V, Subramaniam S, Malik A, Lee G, Manavalan B, Yang D-C. mACPPred: A support vector machine-based meta-predictor for identification of anticancer peptides. *Int J Mol Sci* 2019; 20(8): 1964. [\[CrossRef\]](#)
19. Ahmed S, Muhammod R, Khan ZH, Adilina S, Sharma A, Shatabda S, et al. ACP-MHCNN: An accurate multi-headed deep-convolutional neural network to predict anticancer peptides. *Sci Rep* 2021; 11(1): 23676. [\[CrossRef\]](#)

20. Schaduangrat N, Nantasenamat C, Prachayasittikul V, Shoombuatong W. ACPred: A computational tool for the prediction and analysis of anticancer peptides. *Molecules* 2019; 24(10): 1973. [\[CrossRef\]](#)
21. Gupta S, Kapoor P, Chaudhary K, Gautam A, Kumar R, Raghava GPS. In silico approach for predicting toxicity of peptides and proteins. Patterson RL, editor. *PLoS One* 2013; 8(9): e73957. [\[CrossRef\]](#)
22. Tyagi A, Tuknait A, Anand P, Gupta S, Sharma M, Mathur D, et al. CancerPPD: a database of anticancer peptides and proteins. *Nucleic Acids Res* 2015 ; 43(Database Issue): D837–43. [\[CrossRef\]](#)
23. Leech J, Cabrera-Rubio R, Walsh AM, Macori G, Walsh CJ, Barton W, et al. Fermented-food metagenomics reveals substrate-associated differences in taxonomy and health-associated and antibiotic resistance determinants. *mSystems* 2020; 5(6): e00522-20.. [\[CrossRef\]](#)
24. Durrant MG, Bhatt AS. Automated prediction and annotation of small open reading frames in microbial genomes. *Cell Host Microbe* 2021; 29(1): 121-31. [\[CrossRef\]](#)
25. dos Reis SA, da Conceição LL, e Dias MM, Siqueira NP, Rosa DD, de Oliveira LL, et al. Kefir reduces the incidence of pre-neoplastic lesions in an animal model for colorectal cancer. *J Funct Foods* 2019; 53: 1–6. [\[CrossRef\]](#)
26. Jalali F, Sharifi M, Salehi R. Kefir induces apoptosis and inhibits cell proliferation in human acute erythroleukemia. *Med Oncol* 2016; 33(1): 1–9. [\[CrossRef\]](#)
27. Esener OBB, Balkan BM, Armutak EI, Uvez A, Yildiz G, Hafizoglu M, et al. Donkey milk kefir induces apoptosis and suppresses proliferation of Ehrlich ascites carcinoma by decreasing iNOS in mice. *Biotech Histochem* 2018; 93(6): 424–31. [\[CrossRef\]](#)
28. Sharifi M, Moridnia A, Mortazavi D, Salehi M, Bagheri M, Sheikhi A. Kefir: A powerful probiotics with anticancer properties. *Med Oncol* 2017; 34(11):183. [\[CrossRef\]](#)
29. Arikan M, Mitchell AL, Finn RD. GF. Microbial composition of Kombucha determined using amplicon sequencing and shotgun metagenomics. *J Food Sci* 2020; 85(2): 455-64. [\[CrossRef\]](#)
30. Pothakos V, De Vuyst L, Zhang SJ, De Bruyn F, Verce M, Torres J, et al. Temporal shotgun metagenomics of an Ecuadorian coffee fermentation process highlights the predominance of lactic acid bacteria. *Curr Res Biotechnol.* 2020; 2:1–15. [\[CrossRef\]](#)