Advances in Bioresearch
Adv. Biores., Vol 11 (1) January 2020: 113-122

Print ISSN 0976-4585; Online ISSN 2277-1573
Journal's URL:http://www.soeagra.com/abr.html
CODEN: ABRDC3
DOI: 10.15515/abr.0976-4585.11.1.113122

Advances in Bioresearch

ORIGINAL ARTICLE

Prediction of Novel Bacteriocin from Human Intestinal Microbiome and Their Growth Modeling

Shazeen Yousaf[1$*], Humaira Parvaiz[1$], Sajid Khan[1], M Rizwan[1], Azhar Mehmood[1], Anum Munir[1,2]
[1]Department of Bioinformatics, Govt Postgraduate College of Mandian Abbottabad, Pakistan
[2]Department of Bioinformatics and Biosciences, Faculty of Health and Life Sciences, Capital University of Science and Technology, Islamabad.
*Correspondence: Email address:wanianoor38@gmail.com
[$ equal contribution in research work and shares first authorship.]

ABSTRACT
The microbiome is the genetic material of all the microbes - bacteria, fungi, protozoa and viruses. It is estimated that the human microbiota is composed of 10^{14} bacterial cells, which is 10 times more than the total number of human cells. The most complex and largest is the one comprised by intestinal bacteria that includes as many as 10^{12} cells per 1 g of feces in the average human individual. The human intestinal microbiome is among the most complecate of the body sites: it includes 500–1000 species and several million genes. The function of Human intestinal microbiota is responsible for both human health and disease in correspondence with its own genetic diversity and in joint with human genetic variation. Intestinal microbiome placed an important role in health and immunity and their produced bacteriocin are used for many purposes therefore here the aim of this study is to Predict the novel bacteriocin from human intestinal microbiome and modeling of their growth. To understand mechanism of bacteriocin production and their growth level the bactriocin production model is designed and several bioinformatics approaches are used to modulate the intestinal microbiota and novel bacteriocin as probioticsfor improving health outcomes. The identified bacteriocins can e used as probiotics, in food industry, as an antibiotics and for several other purposes. This study can be further validated in future to confirm its efficacy in vtro.
Keywords: *microbiome, bacteriocin, metagenome,*

Received 02.01.2020 Revised 13.01.2020 Accepted 30.01.2020

How to cite this article:
S Yousaf, H Parvaiz, S Khan, M Rizwan, A Mehmood, A Munir. Prediction of Novel Bacteriocin from Human Intestinal Microbiome and Their Growth Modeling. Adv. Biores., Vol 11 [1] January 2020.113-122.

INTRODUCTION

Human microbiome**,** the full system of microorganisms (the microbiota) that live on and in humans and, more precisely, the collection of microbial genomes that pays to the deeper genetic portrait, or metagenome, of a human. Different type of human microbiota is skin microbiota, intestinal microbiota, gut microbiota etc.The human intestinal microbiome is among the most complecate of the body sites: it includes 500–1000 species and several million genes.The human intestinal microbiota plays a vital role in human nutrition and health by developing the supply of nutrients, preventing pathogen colonization and shaping and framing normal mucosal immunity [1].The small intestine contains alot of bacterias: include *streptococci, staphylococci* and *lactobacilli*. *Veillonellae* and *Actinomycos .Bacteroides, Bifidobacterium,* and *Eubacterium* (1, 2, 4, 6, 8) are predominant isolates .Most common species include anaerobic gram-positive *cocci, Clostridia, Entero cocci*, and various species of *Enterobacteriaceae.* The intestinal microbiota considered as an important modulator of the systemic immune system and also supports local mucosal immunity. Intestinal microbiota protects the host during pneumococcal pneumonia, as reflected by increased bacterial dissemination, inflammation, organ damage and mortality in microbiota-depleted mice compared with controls [2].

Bacteriocins are the subset of antimicrobial peptides (AMPs) produced by bacteria. They are small amphipathic peptides that interact with bacterial membranes leading to cell death, they are protein in nature. Most of the best known are produced by lactic acid bacteria used as food fermentation starters,

because of their potential use as food preservatives [3]. Bacteriocins are used as food preservatives, either after total or partial purification or as extracts of producing bacteria. In situ production is also used, with the advantage of producing early lysis of the starter bacteria and ripening acceleration of the fermented product [4]. They may also form part of hurdle technologies and be incorporated into packaging systems to allow extended liberation. Medical and veterinary applications are in their infancy but good results have been obtained against infection by Gram-positive bacteria and Helicobacter pylori [5]. Bacteriocin growth is totally dependent on environmental factors such as temperature, pH, water activity, redox potential, and the presence of inhibitory compounds determine bacterial growth and the best conditions for bacteriocin growth modeling are required. Up till now several bacteriocins are available in market such as nisin A; lacticins A164, BH5, JW3, and NK24; pediocin PO2; and leucocin K.These molecules exhibit significant potency against other bacteria (including antibiotic-resistant strains), and are stable and can have narrow or broad activity spectra. These Bacteriocins can even be produced in situ in the gut by probiotic bacteria to combat intestinal infections [6].

The aim of this research study is to predict novel Bacteriocin from Human Intestinal Microbiome and modeling of their growth. For modeling purposes, he mathematical model are being used in Bioinformatics and Systems Biology. Mathematical model is a description of a system using mathematical concepts and language. The process of developing a mathematical model is termed mathematical modeling. Mathematical models are used in the natural sciences and engineering disciplines, as well as in the social sciences.

MATERIAL AND METHODS

This section provide general overview of steps performed for prediction of novel bacteriocin from human intestinal microbiome and their growth modeling. The detail of methodology is shown in fig 1.

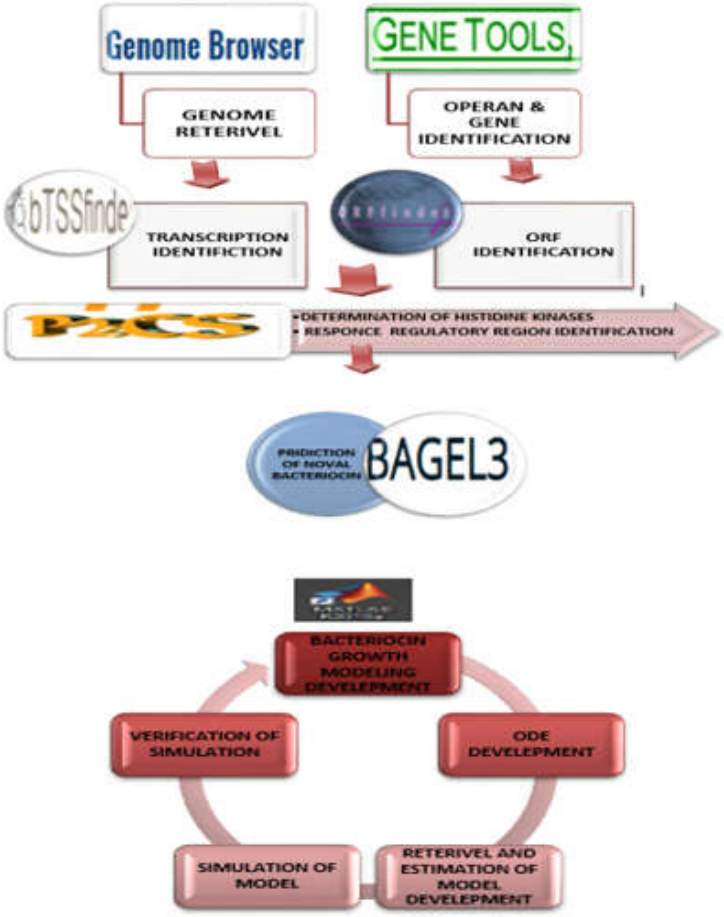


Figure 1. Whole methodology followed in this research work

Genome retrieval

For the prediction of novel bacteriocin from human intestine, the whole genome of bacteria which are present in intestine were retrieved one by one through genome browser. Genome browser is genome retrieval tool which gives the detail about genome sequence its annotation and retrieval [7].

Operon and gene identification

Operons were identified through FGENESB tool. FGENESB proposes an assortment of algorithms which can assign operon as well as to determine promoters, terminators or/and protein coding gene. It can be applied to scaffolds of bacterial genomes or to short sequencing reads. Besides, the program can be used to determine both tRNA and rRNA genes. https://omictools.com/fgenesb-tool

Transcription factor and ORF identification

For the identification of transcription factor **BTSS tool** was used which is a novel tool for the prediction of promoters in *cyanobacteria*, having a higher accuracy of MCC=0.59. and covers multiple classes of promoter. https://omictools.com/btssfinder-tool. ORF identification was done by ORF finder.https://www.ncbi.nlm.nih.gov/orffinder/.ORF were identified on the basis of sequences of genome and parameters were set on the basis of factors effecting on them.

Determination of histidine kinases

Histidine kinases, response regulatory regions were determined through P2CS tool which is very accurate tool. The P2CS database (http://www.p2cs.org/) is a comprehensive resource for the analysis of Prokaryotic Two-Component Systems (TCSs). TCSs are comprised of a receptor histidine kinase (HK) and a partner response regulator (RR) and control important prokaryotic behaviors [8].

Prediction of novel Bacteriocins

Novel Bacteriocins were identified through BAGEL3 tool. BAGEL3 uses DNA nucleotide sequences in FASTA format as input; multiple sequence entries per file are allowed. These DNA sequences are analyzed in parallel using two different approaches, one based on finding genes commonly found in the context of bacteriocin or Rip genes, the other based on finding the gene itself [9].

Bacteriocin growth modeling development

Bacteriocin production model was develop through matlab Simulink software. Simulink is a graphical extension to MATLAB for modeling and simulation of systems. Modeling and Simulation of Systems. Using MATLAB and Simulink provides comprehensive, state-of-the-art coverage of all the important aspects of modeling and simulating both physical and conceptual systems [10].

Parameter Estimation

Bacterial cell growth development strongly dependent on environmental conditions. Factors such as temperature, PH, PI, elimination rate and degradation rate etc [11, 12]. Estimation of parameters were done by ODE solver 23 of matlab.

Simulation of Model

Simulations of model were performed to determine the expression of bacteriocin through matlab simbiology toolbox. SimBiology is a graphical environment for (PK/PD) modeling and analysis. The SimBiology environment provides point-and-click tools to make PK/PD modeling and analysis accessible, even if you have little to no programming experience [13].

RESULTS AND DISCUSSION

Genome retrieval of selected bacterial species.

For the prediction of novel Bacteriocins, the genome size of 7 bacterial species: *E-coli, Lactobacillus, Clostridium, Bacteroidesvulgatus, Methanosphaerastadtmanae, Bifidobacterium longum,* and *Lactobacillus plantarum*was retrieved one by one through genome browser. In the selected bacterial species, some have chromosomal genome and others have plasmid genome. Theselected bacterial species and their size of genomes are shown in (table 1).

Operons and gene identification.

We identified operons through FGENESB tool. Operons are groups of genes that are transcribed in a single mRNA.When an operon is transcribed, all of the genes on the operon are on the same mRNA. Operons occur in prokaryotes [14].The identified operons and genes in the bacterial genomes are shown in (table 2).*Escherichia coli* contains maximum number ofoperons, whereas no operon was found in *Lactobacillus*

Transcriptional units ORF calculations.

Transcriptional units were identified through BTSS tool and ORF identifications were done by ORF finder. Transcriptional units and ORF calculations are shown in (table 3). In *Lactobacillus plantarum*, the highest number of transcription units were identified and largest number of ORFs were obtained in *Escherichia coli.* The *Lacobacillus* againfound to be constitute minimum number of transcription units.

Determination of Histidine kinesis and Response regulatory regions.

Histidine kinases are found in most bacteria, archaea, and lower eukaryotic species where they function as two component signal transduction pathway.The canonical two-component system (TCS) utilizes two distinctive signaling proteins, a membrane-bound sensor histidine kinase (HK) and a cytoplasmic

response regulator (RR) protein. In response to extracellular stimuli, the HK catalyzes ATP-dependent autophosphorylation of its conserved histidine residue (within the HK dimerization domain) and then transfers phosphoryl groups to a conserved aspartate residue within the receiver domain of the RR protein [15].Histidine kinesis and response regulatory regions were determined through P2CS tool. The different values of histidine kinesis and response regulatory region for selected species are shown in (table 4).

Predicted Bacteriocin from different bacteria.

Some new functions were identified along the newly predicted Bacteriocins. The novel Bacteriocins were identified through BAGEL3 tool. The predicted Bacteriocins are shown in (figure 2a -2c).The bacteriocin produced by *Bifidiobacterium* is of type head to tail cyclized peptide, is found in small ORF_7 and contain a protein sequence
NPGQRSPVAISHSNPVAARMSSMSCTASMCRSSPTARASLIRPHIGSNAMFALCCMVCGVGMMVHPFRVA. It lies in the area where gene start at 60 and ends at 179 with a Length of sequence is 71 and having an isoelectric point 10.5. Along with the bacteriocin 3 transport regions, 1 modification and 2 undefined associates were identified. The Bacteriocin produced by *Clostridium* is of type class II, is found in ORF 012 Linocin_M18 and contain a protein sequence
MENEVIKLASSQGIWSVLSIVLIFYIIKTQEKRDLRQEEREKNYQSIISELTSKFYIVEEIKKDVEEIKESVIV. It lies in the area where gene start at 5157 and ends at 5219 with a length of sequence 74 and having Isoelectric point 4.7. The bacteriocin of *Eubacterium* is of type head to tail cyclized peptide, is found in ORF_ 038 and contain a protein sequence
MDHALLIPPHNTVGIDLYIPLTSAKNVDIKCSFKNVLGEKTAGFLHLDKADDVYHGIYNLRIVEA.The bacteriocin lie in area where gene start at 19756 and ends at 19953 with a Length of sequence 65 and having Isoelectric point 6.5. Along with bacteriocin 1 modification region, 2 regulation region and 1 immunity region were newly identified. The *E-coli* bacteriocin is of type bottromycin, is found in small ORF_1 and contain a protein sequence MKYDDLHGHKQKCDSQLQMDEGNNCTQYRRQLPLRTGKPLIPKYQCRDRSGREP. The bacteriocin lie in area where gene start at 273 and ends at 437 with a length of sequence 54 and having isoelectric point 9.3.Along with bacteriocin 1 transport region and 1 modification were newly identified.6 Bacteriocins are predicted by Lactobacillus Plantarum, are of type Class II and are found in ORF 016 with a protein sequence of length 57
1). MKIKLTVLNEFEELTADAEKNISGNRRSRKNGIGYAIGYAFGAVERAVLGGSRDYNK The bacteriocin lie in area where gene start at 8061 and ends at 8234 and having an isoelectric point 10.0. 2) small ORF_30comc with a protein sequence of length 55
MTVNKMIKDLDVVDAFAPISNNKLNGVVGGGAWKNFWSSLRKGFYDGEAGRAIR. The bacteriocin lie in area where gene starts at 8265 and ends at 8432 and having isoelectric point 10.7. 3) small ORF_13 bacteriocin IIc with a protein sequence of length 55
MKSLDKIAGLGIEMAEKDLTTVEGGKNYSKTWWYKSLTLLGKVAEGTSSAWHGLG. The bacteriocin lie in area where gene starts at 9622 and ends at 9789 and having isoelectric point 9.0. 4) small ORF_17 DUF1858 with a protein sequence of length MKIQIKGMKQLSNKEMQKIVGGKSSAYSLQMGATAIKQVKKLFKKWGW.
The bacteriocin lie in area where gene starts at 12761 and ends at 12907 and having isoelectric point 11.2. 5) small ORF_24 Lactococcin with a protein sequence of length 52 MKKFLVLRDRELNAISGGVFHAYSARGVRNNYKSAVGPADWVISAVRGFIHG. The bacteriocin lie in area where gene starts at 17209 and ends at 17367 and having Isoelectric point 11.1. 6) ORF 033 bacteriocin _IIc with a protein sequence of length 56
MTVNKMIKDLDVVDAFAPISNNKLNGVVGGGAWKNFWSSLRKGFYDGEAGRAIRR. The bacteriocin lie in area where gene starts at 17392 and ends at 17562 and having isoelectric point 11.7. Along with these bacteriocins 1 transport and leader cleavage and 4 undefined associates were identified.

Estimation of parameters.

The estimated value of parameters like temperature, PH, elimination rate, degradation rate and isoelectric point are required for the production of bacteriocin. So the estimation of parameters was done through the ODE solver 23 in matlab to predict the growth of bacteriocins. The estimated value of parameters are shown in (table 5).

Bacteriocin growth modeling:

Bacteriocin growth model was developed in Simulink matlab comprises of four different phases the central phase, biophase, growing phase and resting phase. In central phase and biophase, the bacterial culturing was shown. If bactrias are provided by proper culturing then bacterial growth occurs and if proper culturing is not provided, then bactrias goes towards resting phase [16]. In the model, the bacterial cell growth was strongly dependent on environmental conditions like temperature, pH, water activity, redox potential, elimination rate, degradation rate, isoelectric point and the presence of

inhibitory compounds. Any increase or decrease in the value of parameter highly effects the growth of bacteriocin

Simulation results

Simulations were performed to check the expression of bacteriocin production through matlab simbiology toolbox. The expression level of different bacteriocin produced by different bacteria are shown in (table 6) where x axis shows the time rate and y-axis shows the production rate of bacteriocin.

TABLE 1: Various types of bacterial species and their size of genomes are shown

Bacterial species	Types of Genome		Size of Genome
	Chromosomal	plasmid	
Escherichia coli	Chromosomal	___	5132068
Clostridium	Chromosomal	___	3886916
Lactobacillus	Chromosomal	___	1917
Bacteroides vulgatus	Chromosomal	___	5163189
Methanosphaerastadtmanae	Chromosomal	___	3091
Bifidobacterium longum	Chromosomal	___	2256640
Bifidobacterium longum	___	plasmid	2705
Lactobacillus plantarum	Chromosomal	___	5163189

TABLE 2:genes, operons values are shown

Bacterial species	Genome size	Genes	Operon
Escherichia coli	5132068	5253	1060
Clostridium	3886916	4420	896
Lactobacillus	1917	3	0
Bacteroidesvulgatus	5163189	4507	979
Bifidobacterium longum	2256640	2119	425
Lactobacillus plantarum	5163189	4822	747

Table 3. transcriptionalunits ORF calculations are shown

Bacterial species	Genome size	Transcriptional Unit	ORF
Escherichia coli	5132068	3224	143
Clostridium	3886916	2536	104
Lactobacillus	1917	3	90
Bacteroidesvulgatus	5163189	2768	73
Bifidobacterium longum	2256640	1370	103
Lactobacillus plantarum	5163189	3796	13

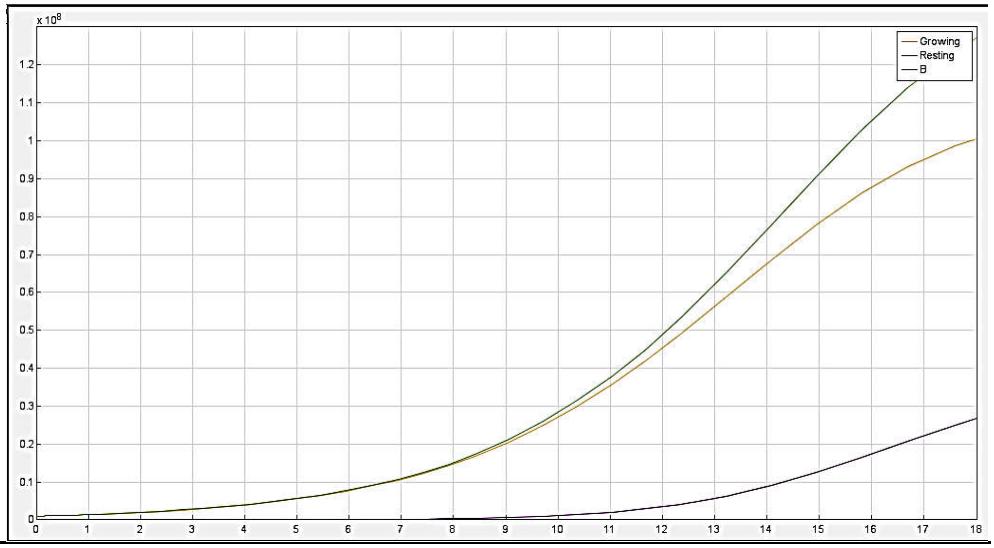
Table 4. histidine kinases and response regulatory regions are shown

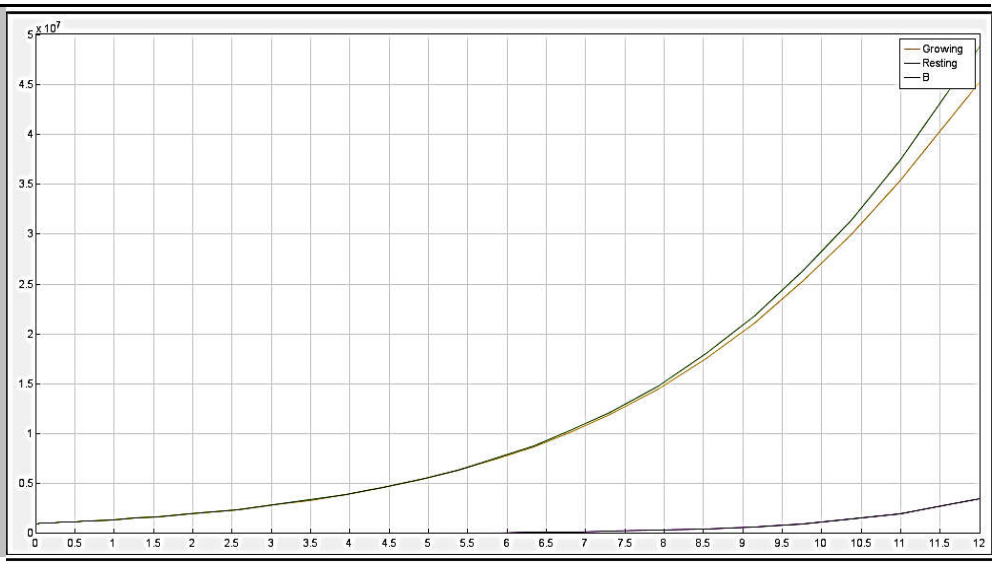
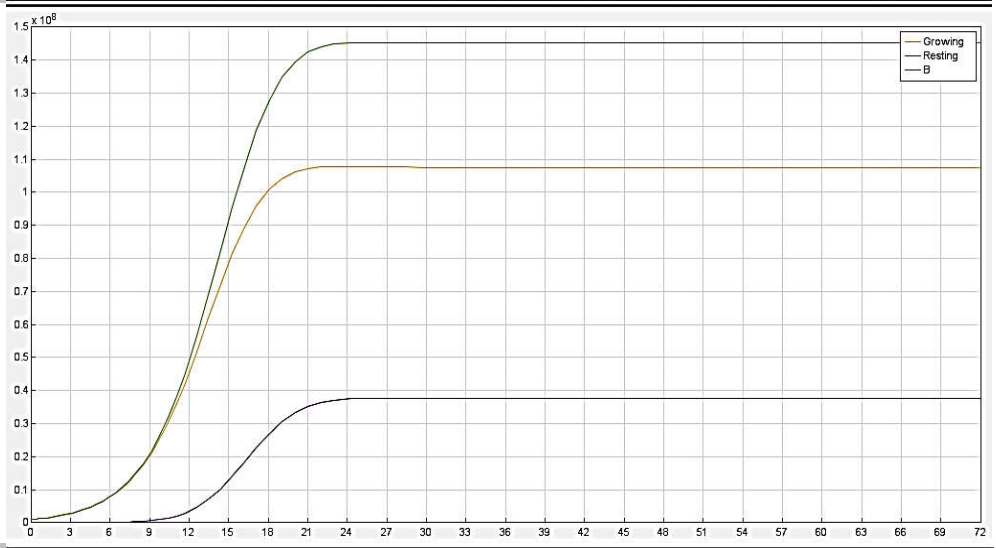
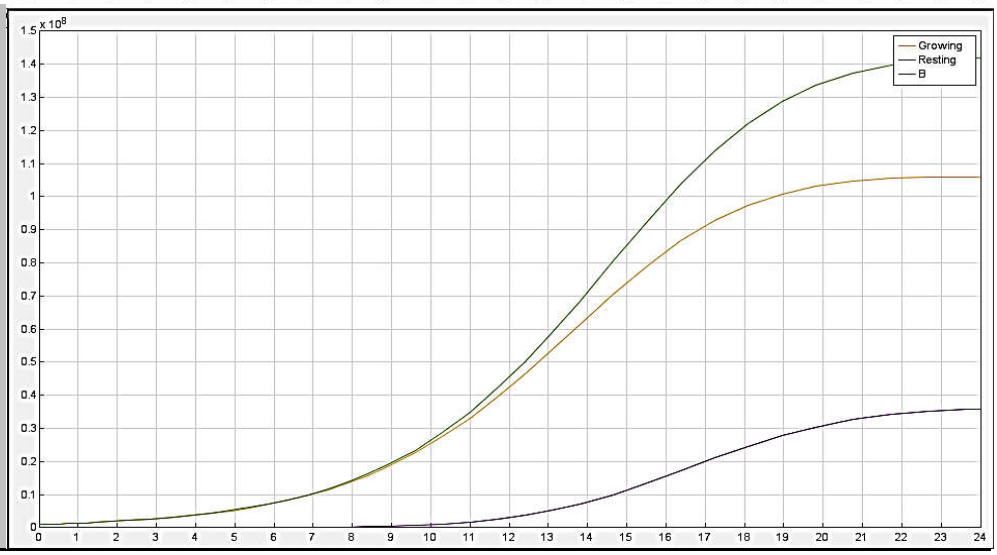
Bacterial species	Histidine kinases	ResponseRegulatory Region
Escherichia coli	30	34
Clostridium	10	9
Lactobacillus	7	8
Bacteroidesvulgates	0	0
Bifidobacterium longum	13	13
Lactobacillus plantarum	8	8

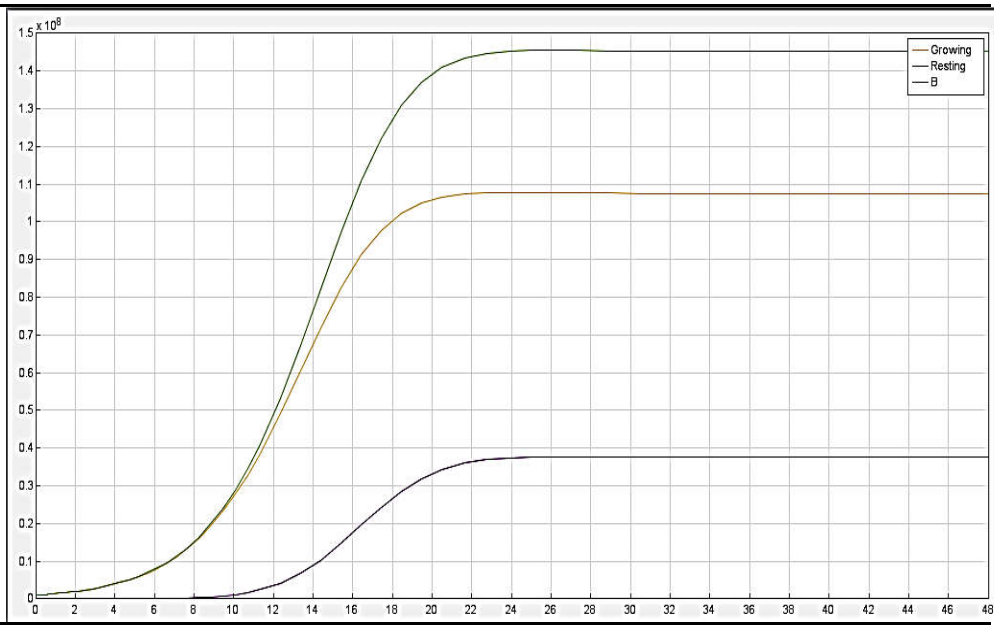
Table 5. Estimated values of parameter required for bacteriocin production

Parameter Estimation	Symbolic representation	Species	Values
Temperature	Temp	*Bifidobacterium*	39.0
		Clostridium	37.0
		Lactobacillus	38.0
		Eubacterium	37.0
		Escherichia Coli	37.0
Elimination rate	Ke	*Bifidobacterium*	23.79
		Clostridium	57.0
		Lactobacillus	26.3
		Eubacterium	0.861
		Escherichia Coli	0.861
Degradation rate	Kdeg	*Bifidobacterium*	61.0
		Clostridium	0.026
		Lactobacillus	50.0
		Eubacterium	42.5
		Escherichia Coli	4.5
Potential Hydrogen	PH	*BIfidobcterium*	6.75
		Clostridium	4.0
		Lactobacillus	7.26
		Eubacterium	7.2
		Escherichia Coli	7.5
Isoelectric point	PI	*BIfidobcterium*	4.6
		Clostridium	6.75
		Lactobacillus	7.17
		Eubacterium	6.84
		Escherichia Coli	5.6

Table 6. Shows the expression level of bacteriocin produced by different bacteria

Bifidobacterium	

Clostridium	
Eubacterium	
Escherichia coli	

Lactobacillus plantarum	

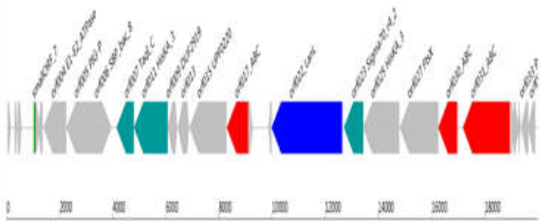
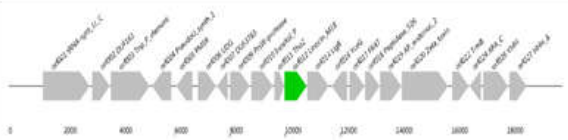
Bifidobacterium	This Bacteriocin is of type head to tail cyclized peptide, is found in small ORF_7. Along with the bacteriocin 3 transport regions, 1 modification and 2 undefined associates were identified.	
Clostridium	This Bacteriocin is of type class II, is found in ORF 012 Linocin_M18.	

Figure 2a. Predicted Bacteriocins identified by different bacteria

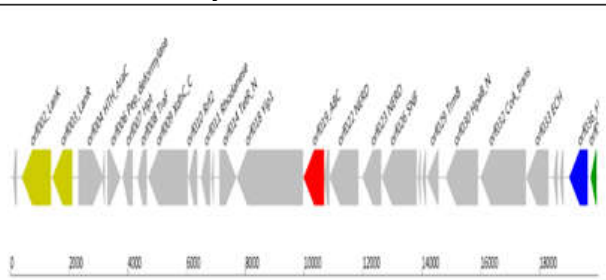
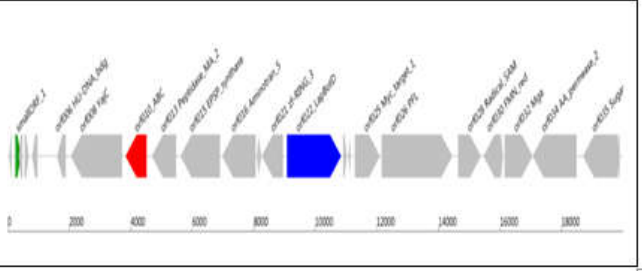
Eu bacterium	This Bacteriocin is of type head to tail cycle peptide, is found in ORF_ 038. Along with Bacteriocin 1 modification region, 2 regulation region and 1 immunity region were newly identified.	
Escherichia coli	This Bacteriocin is of type Bottromycin, is found in small ORF_1. Along with Bacteriocin 1 transport region and 1 modification were newly identified.	

Figure 2b. Predicted Bacteriocins identified by different bacteria:

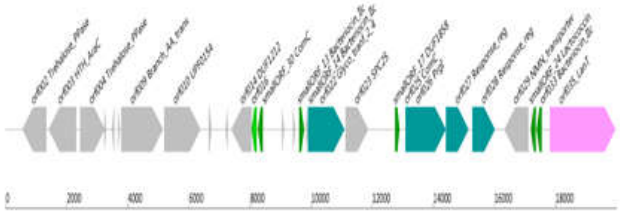
Lactobacillus plantarum	6 Bacteriocins were predicted by Lactobacillus Plantarum, are of type Class II and found in ORF 016, 30,13,17 DUF1858,24 and 033.1 transport and 4 undefined associates were also identified.	

Figure 2c. Predicted Bacteriocins identified by different bacteria:

CONCLUSION

Species of microbes in human intestine *are lactobacilli, Clostridium, E Coli, Bifidobacterium, Eubacterium*. Therefore, these species were selected and the novel bacteriocin were identified from the following species through different advance tools. Then after that simulation were performed to check the expression rate of predicted Bacteriocins produced by different bacteria. The maximum numbers of bacteriocin were produced by Lactobacillus Plantarum with a protein sequence of length 57. Whereas *Bifidobacterium*, *Clostridium, Eubacterium* and *Escherichia coli* produced one bacteriocin each. Their growth modeling was done on matlab. These newly predicted Bacteriocins will not only help us in medical and veterinary application but also useful as food preservatives to kill bacteria.

REFERENCES

1. Ursell, L. K., Metcalf, J. L., Parfrey, L. W., & Knight, R. (2012). Defining the human microbiome. *Nutrition reviews*, *70 Suppl 1*(Suppl 1), S38–S44. doi:10.1111/j.1753-4887.2012. 00493.x
2. Martínez, B., Rodríguez, A., & Suárez, E. (2016). Antimicrobial peptides produced by bacteria: the bacteriocins. In *New weapons to control bacterial growth* (pp. 15-38). Springer, Cham. (major group of bacteriocin)
3. Sobrino-López, A., & Martín-Belloso, O. (2008). Use of nisin and other bacteriocins for preservation of dairy products. *International Dairy Journal*, *18*(4), 329-343.
4. Kim, T. S., Hur, J. W., Yu, M. A., Cheigh, C. I., Kim, K. N., Hwang, J. K., &Pyun, Y. R. (2003). Antagonism of Helicobacter pylori by bacteriocins of lactic acid bacteria. *Journal of food protection*, *66*(1), 3-12……h pylori
5. Zhu, B., Wang, X., & Li, L. (2010). Human gut microbiome: the second genome of human body. *Protein & cell*, *1*(8), 718-725.
6. Aasen, I. M., Møretrø, T., Katla, T., Axelsson, L., & Storrø, I. (2000). Influence of complex nutrients, temperature and pH on bacteriocin production by *Lactobacillus sakei* CCUG 42687. *Applied Microbiology and Biotechnology*, *53*(2), 159-166.
7. Fujita, P. A., Rhead, B., Zweig, A. S., Hinrichs, A. S., Karolchik, D., Cline, M. S. & Diekhans, M. (2010). The UCSC genome browser database: update 2011. *Nucleic acids research*, *39*(suppl_1), D876-D882.
8. Ortet, P., Whitworth, D. E., Santaella, C., Achouak, W., & Barakat, M. (2014). P2CS: updates of the prokaryotic two-component systems database. *Nucleic acids research*, *43*(D1), D536-D541.
9. van Heel, A. J., de Jong, A., Montalbán-López, M., Kok, J., & Kuipers, O. P. (2013). BAGEL3: automated identification of genes encoding bacteriocins and (non-) bactericidal post translationally modified peptides. *Nucleic acids research*, *41*(W1), W448-W453.
10. Chaturvedi, D. K. (2017). *Modeling and simulation of systems using MATLAB and Simulink*. CRC press.
11. Leroy, F., & De Vuyst, L. (2001). Growth of the bacteriocin-producing *Lactobacillus sakei* strain CTC 494 in MRS broth is strongly reduced due to nutrient exhaustion: a nutrient depletion model for the growth of lactic acid bacteria. *Appl. Environ. Microbiol.*, *67*(10), 4407-4413.
12. Ilea, M., Turnea, M., &Rotariu, M. (2012). Ordinary differential equations with applications in molecular biology. *Rev Med ChirSoc Med Nat Iasi*, *116*(1), 347-352.

13.]. Saintes, France, U. K. Sandwich, and U. K. Greenford. "S-03 Sven Janssen SimBiology: A Graphical Environment for Population PK/PD."
14. 40 Price, M. N., Arkin, A. P., &Alm, E. J. (2006). The life-cycle of operons. *PLoS genetics*, *2*(6), e96. doi:10.1371/journal.pgen.0020096
15. Yoshida, T., Phadtare, S., & Inouye, M. (2007). Functional and structural characterization of EnvZ, an osmosensing histidine kinase of E. coli. In *Methods in enzymology* (Vol. 423, pp. 184-202). Academic Press.
16. Yi, L., Luo, L., & Lu, X. (2018). Efficient exploitation of multiple novel bacteriocins by combination of complete genome and peptidome. *Frontiers in microbiology*, *9*, 1567.
17. Tuddenham, S., & Sears, C. L. (2015). The intestinal microbiome and health. *Current opinion in infectious diseases*, *28*(5), 464.
18. Cotter, P. D., Ross, R. P., & Hill, C. (2013). Bacteriocins—a viable alternative to antibiotics?. *Nature Reviews Microbiology*, *11*(2), 95.