AENSI Journals

Advances in Environmental Biology

ISSN-1995-0756 EISSN-1998-1066

Journal home page: http://www.aensiweb.com/AEB/

Anticancer Peptides Investigation in Silico Using TRAINER Tool

[1]Akram Songol and [2]Sayed Mohammad Hoseini Sedeh

[1]MSc student in Biotechnology, Isfahan University, Isfahan, Iran.
[2]MSc student in Agriculture Biotechnology, Faculty of Agriculture, Islamic Azad University of sabzevar, Mashhad, Iran.

ARTICLE INFO

Article history:
Received 11 October 2014
Received in revised form 21 November 2014
Accepted 25 December 2014
Available online 16 January 2015

Keywords:
Anticancer peptides, Cancer, Naive Bayes, TRAIER, Radial Basis

ABSTRACT

One of the causes of death in the world is cancer. Radiotherapy and chemotherapy are among the treatment methods existing against cancer cells. Identification and developing a new method for treating cancer is extremely essential because traditional methods have side effects on normal cells and are very costly. Recently discovered agents are Antimicrobial peptides which are present in a wide variety of organisms including plants, amphibians and mammals. Developing a computational method for predicting these anticancer peptides is helpful since these peptides have a variety of structures, sizes and molecular compositions. In the current research, primarily, 2 databases with 138 and 206 anticancer and non-anticancer peptides were introduced which were classified by TRAINER. TRAINER (http://www.baskent.edu.tr /~hogul/TRAINER/) is a new online tool designed for classification of any alphabet of sequences. With TRAINER users are capable of selecting from among several feature representation schemes and supervised machine learning methods with pertinent parameters. Moreover, naive Bayes and radial basis were used in a support vector machine in the present study. The accuracy and specificity in combination of features by Naive Bayes were 83% and by radial basis 87% and 92%correspondingly. Based on the findings two methods are very helpful for categorization of these peptides; yet, the precision of Radial Basis is superior to that of Naive Bayes.

To Cite This Article: Akram Songol and Sayed Mohammad Hoseini Sedeh., Anticancer peptides investigation in silico using TRAINER tool. ***Adv. Environ. Biol., 9(2),*** *224-228, 2014*

INTRODUCTION

Throughout the world, in spite of current advances in treatment modalities, cancer remains a major source of morbidity and death. Cancer is the most important cause of death for individuals younger than 85 years of age [1] in the United States. Furthermore, the occurrence of various cancers including cancers of the skin, prostate, breast, and kidney, are on the raise [2]. Actually, cancer is a general term that refers to more 100 different diseases affecting many different tissues and cell types. Nonetheless, all cancers are distinguished by abnormal cell growth resulting from a relatively small number of inherited or environmentally-induced genetic mutations [3]. According to Hanahan and Weinberg [4], in order for a cell to become cancerous, it must obtain six distinctive traits as a result of altered cell physiology. The distinctive traits of cancer cells are as follows: 1) capability of generating their own growth signals or reacting to weak growth signals ignored by healthy cells; 2) insensitivity to anti proliferative signals; 3) resistance to cellular suicide mechanisms that normally cause abnormal cells to die by apoptosis; 4) capability of infinite replication; 5) being able to stimulate new blood vessel development allowing for tumour growth; and 6) the ability to attack tissues, at first locally, and afterward to spread or metastasize all over the body. Chemotherapy continues to be the usual treatment of choice for advanced or metastatic disease [5] though localized cancers can often be effectively treated by surgery and/or radiation therapy. Nevertheless, the application of usual chemotherapeutic agents that normally target rapidly dividing cancer cells is often associated with harmful side-effects caused by unintentional drug-induced damage to healthy cells and tissues [6, 7]. In addition, cancer cells that are inactive or gradually proliferating are refractory to the cytotoxic effect of chemotherapeutic drugs that act at the level of DNA synthesis [8]. Moreover, as a result of cellular changes that include increased expression of drug-detoxifying enzymes and drug transporters, altered interactions between the drug and its target, an increased ability to repair

Corresponding Author: Akram Songol, MSc student in Biotechnology, Isfahan University, Isfahan, Iran.

Advances in Environmental Biology, 9(2) January 2015, Pages: 224-228

DNA damage, and defects in the cellular machinery that mediate apoptosis [9] cancer cells often become resistant to chemotherapy. A key advance in cancer treatment would be the development of a new class of anticancer drugs that lack the toxicity of common chemotherapeutic agents and are unchanged by common mechanisms of chemo resistance. In the present article, we studies on naturally occurring, in addition to some selected hybrid and synthetic cationic antimicrobial peptides (AMPs) that display anticancer activities.

The disrupting of cytoplasmic and mitochondrial membranes is the basis of peptides anticancer effect [12-15]. Membranes of cancer cells with negative charge attach to the peptides with positive charge. Peptides'amphipathic structure allows them to insert themselves in the membrane and permeate it through the formation of pores [16]. Still, these peptides have small size and diversity in secondary structure and composition but the following features shared among them including positive charge, high content of hydrophobic residues, amphipathic fold, ease of synthesis and modification and tumor penetrating ability [17-19]. Developing a computational method for the prediction of anticancer peptides is needed in order to decrease costs and time.

A number of databases such as the Antimicrobial Peptide Database (APD2) [20] and the Collection of Antimicrobial Peptides (CAMP) have identified anticancer peptides [21]. APD2 is available at (http://aps.unmc.edu/AP/main.php).

MATERIALS AND METHODS

Datasets:

180 experimentally validated ACPs were extracted from the literature as well as the antimicrobial database (APD2) as positive database. A non-anticancer database was introduced manually from the Universal Protein Resource (UniProt), at http://www.uniprot.org.To introduce the negative database, non-secretary proteins and randomly cut out peptides with the same length range were selected. In the end, negative database containing 215 non-anticancer peptides was used. To eliminate the duplicates[22], the CD-HIT tool (http://cd-hit.org) was used since there are the same sequences in each database. Eventually, positive and negative databases were used, containing 130 and 200peptides respectively.

Radial basis function networks:

Similar to neural networks, radial basis function (RBF)networks [23] have also been demonstrated to be universal approximators [24]. In comparison to neural networks, RBF networks are to a great extent easier to design and train and have strong tolerance to input noise, which enhances the stability of the designed systems. Consequently, it is reasonable to consider an RBF network as a competitive method for nonlinear controller designs. Three steps are included in the basic computations of the RBF network: 1) Input Layer Computation: At the input layer, each input (xp,i) is scaled by the input weights (ui,h) presenting the weight connection between the it h input and RBF unit h.

(1) yp,h,i= xp,iui,h

where vector yp,h= {yp,h,1, yp,h,2 . . . yp,h,i. . . yp,h,I} is the scaled inputs, h is the index of RBF units from 1 to H, i is the index of inputs from 1 to I, and p is the index of training patterns from 1 to P.

2) Hidden Layer Computation: The output of the RBF unit h is calculated by:

(2) ϕh(xp) = exp (|| , – ||)

Where ϕh(•) is the activation function of the RBF unit h,. chand σ hare the center and width, the key properties to describe the RBF unit h,, and || • || represents the computation of Euclidean norm of two vectors.

3) Output Layer Computation: for pattern xp, the network output is calculated as the sum of weighted outputs from RBF units.

(3) op= ( ) +

Where presents the weight value on the connection between the RBF unit h and network output. W0 is the bias weight.

Naive Bayes:

A naive Bayes classifier is a simple probabilistic classifier based on the application of Bayes' theorem with strong (naive) independence assumptions. Considering features as independent makes their learning simple and computationally efficient [25]. This method is generally used due to its advantages. For Naïve Bayes models, parameter estimation is done using maximum likelihood. The method is suited when the dimensionality of the input is high.

In the current study, for classification and evaluating the performance of classifiers Radial Basis and Naive Bayes were employed Performance of the classifier is measured in terms of sensitivity (SEN), specificity (SPEC), accuracy (ACC) and Matthew's correlation coefficients (MCC) given by equations (4-7),

(4) SEN=TP/(TP+FN)

(5) SPEC=TN/(TN+FP)

(6) ACC=(TN+TP)/(TN+FP+TP+FN)

(7) MCC= (TP×TN - FP× FN) / (TP □ □ FP)(TP + FN)(TN + FP)(TN + FN)

Advances in Environmental Biology, 9(2) January 2015, Pages: 224-228

where TP, TN, FP and FN are the numbers of true positives, true negatives, false positives, and false negatives, respectively.
The performances of classifiers are presented in tables (1-4).

Trainable Short Sequence Classifier:

TRAINER is a new online and flexible tool for biosequence analysis. TRAINER interface can provide a user-friendly environment for any basic internet user. This system can respond to thousands of short sequences in seconds and is capable of producing exact results for all types of biological sequences [26]. TRAINER is available at (http://www.baskent.edu.tr/~hogul/TRAINER/). Biologicalsequences are of different lengths, cannot be directly fed in to a classifier and need to be represented by a number of numerical features. There are three different composition representations as 1-mer, 2-mer, 3-mer and their combinations in this tool. In the current research, three different compositions and the combination of all vectors was used.

Results:

In the present study, for the classification and evaluation of the performance of the classifier two machine learning methods with three different compositions of features and their combinations were used. The findings proved that for predicting anticancer peptides these methods were useful. for classifying anticancer and non anticancer peptides both methods are useful but accuracy, sensitivity, specificity and Matthews' correlation coefficients of radial basis were found to be better than Naive Bayes', and in this analysis, the combination of all vectors yielded the best results.

Table 1: Matthew's correlation coefficients of two machine learning methods in three different compositional representations.

Mathew's correlation coefficients				
Methods	1-mer	2-mer	3-mer	all
Naive Bayes	0.60	0.52	0.53	0.58
Radial Basis	0.74	0.82	0.69	0.84

Table 2: Sensitivity and specificity of two machine learning methods in three different compositional representations.

	Sensitivity				Specificity			
Methods	1-mer	2-mer	3-mer	all	1-mer	2-mer	3-mer	all
Naive Bayes	87%	72%	73%	79%	82%	84%	78%	79%
Radial Basis	77%	96%	88%	82%	74%	85%	84%	81%

Table 3: Accuracy of two machine learning methods in three different compositional representations.

Accuracy				
Methods	1-mer	2-mer	3-mer	all
Naive Bayes	76%	77%	86%	86%
Radial Basis	85%	89%	69%	79%

Discussion:

In today's world cancer therapy is an attractive topic. For treating cancer there are many traditional methods, but these methods have their own limitations. Consequently, we concentrated on a more particular method. Anticancer peptides of human, plant, and animal origin are shown to be new agents against cancerous cells. Peptides'oncolytic effect depends on their cationic and amphipathic structure [8]. Asgangliosidesexist in both normal and cancer cells (although by different proportions), the function of the peptides is not specific, however they may be attracted to cancer cells more commonly than normal cells. The positive charge of the peptides is proposed to start electrostatic interaction with the negatively charged membranes of tumor cells. These features could cause the permeation of peptides into the membrane and a succeeding complete membrane disruption. Additionally, by perturbation in the plasma membrane they depolarize the trans membrane potential of cancer cells and kill the cells [11, 27]. Because the examination of peptides in vitro and in vivo is protracted and expensivive, providing a computational method for the prediction of anticancer peptides can be useful. For predicting different aspects of proteins, based on amino acid sequence, template and amino acid composition (AAC) several methods exist. Also, Pse AAC concept has been usually used to predict several aspects of proteins including cyclins [28], risk type of human papillomaviruses [29], GABAA receptors [30], metalloproteinase family [31], antibacterial peptides [32] and allergenic Proteins [33].

As a result, in the present study anticancer and non-anticancer peptides from databases and articles were collected and classified using two machine learning methods in TRAINER tool. A new online tool acceptable for the classification of bio sequences is TRAINER. In the current study, ACC, SEN and SPEC of these methods demonstrated that TRAINER is a useful tool for predicting anticancer and non-anticancer peptides. MCC is a measure for appraising the quality of binary classifications. This measure has a value between -1 and +1.When MCC is higher than 0.7, it is acceptable for predictors. In the current study, the best results for MCC were found to be in combinations of all vectors, MCC of Radial Basis being higher than 0.7, and hence

acceptable. The findings demonstrated that for predicting the above mentioned peptides, the precision of Radial Basis was more than that of Naive Bayes.

REFERENCE

[1] Suarez-Jimenez, G.M., A. Burgos-Hernandez, J.M. Ezquerra-Brauer, 2012. Bioactive peptides and depsipeptides with anticancer potential: Sources from marine animals. Mar Drugs, 10: 963-986.
[2] Dunn, G.P., A.T. Bruce, H. Ikeda, L.J. Old, R.D. Schreiber, 2002. Cancer immunoediting: from immuno surveillance to tumor escape. Nat. Immunol., 3: 991-998.
[3] Mocellin, S., C.R. Rossi, D. Nitti, 2004. Cancer vaccine development: on the way to break immune tolerance to malignant cells. Exp. Cell Res., 299: 267-278.
[4] Massodi, I., S. Moktan, A. Rawat, G.L. Bidwell, D. Raucher, 2010. Inhibition of ovarian cancer cell proliferation by a cell cycle inhibitory peptide fused to a thermally responsive polypeptide carrier. Int. J. Cancer, 126: 533-544.
[5] Shadidi, M., M. Sioud, 2003. Selective targeting of cancer cells using synthetic peptides. Drug Resist Update, 6: 363-371.
[6] Leuschner, C., W. Hansel, 2004. Membrane disrupting lytic peptides for cancer treatments. Curr. Pharm. Design, 10: 2299-2310.
[7] Papo, N., M. Shahar, L. Eisenbach, Y. Shai, 2003. A novel lytic peptide composed of D, L amino acids selectively kills cancer cells in culture and in mice. J. Biol. Chem., 278: 21018-21023.
[8] Tossi, A., L. Sandri, A. Giangaspero, 2000. Amphipathic, α-helical antimicrobial peptides. Pept. Sci., 55: 4-30.
[9] Diamond, G., N. Beckloff, A. Weinberg, K.O. Kisich, 2009. The roles of antimicrobial peptides in innate host defense. Curr. Pharm. Design, 15: 2377.
[10] Bals, R., 2000. Epithelial antimicrobial peptides in host defense against infection. Respir Res., 1: 141-150.
[11] Papo, N., Y. Shai, 2005. Host defense peptides as new weapons in cancer treatment. Cell Mol. Life Sci., 62: 784-790.
[12] Mai, J.C., Z. Mi, S.H. Kim, B. Ng, P.D. Robbins, 2001. A proapoptotic peptide for the treatment of solid tumors. Cancer Res., 61: 7709-7712.
[13] Ellerby, H.M., W. Arap, L.M. Ellerby, R. Kain, R. Andrusiak, G.D. Rio, S. Krajewski, C.R. Lombardo, R. Rao, E. Ruoslahti, D.E. Bredesen, R. Pasqualini, 1999. Anti-cancer activity of targeted pro-apoptotic peptides. Nat. Med., 5: 1032-1038.
[14] Shai, Y., 2002. Mode of action of membrane active antimicrobial peptides. Biopolymers, 66: 236-248.
[15] Hoffmann, J.A., F.C. Kafatos, C.A. Janeway, R.A. Ezekowitz, 1999. Phylogenetic perspectives in innate immunity. Science, 284: 1313-1318.
[16] Shai, Y., 1999. Mechanism of the binding, insertion and destabilization of phospholipid bilayer membranes by α-helical antimicrobial and cell non-selective membrane-lytic peptides. Biochim Biophys Acta, 1462: 55-70.
[17] Nijnik, A., R. Hancock, 2009. Host defense peptides: antimicrobial and immunomodulatory activity and potential applications for tackling antibiotic-resistant infections. Emerg. Health Threats J., 2: e1.
[18] Thayer, A.M., 2011. Improving peptides. Chem. Eng. News, 89: 13-20.
[19] Borghouts, C., C. Kunz, B. Groner, 2005. Current strategies for the development of peptidebased anti-cancer therapeutics. J. Pept. Sci., 11: 713-726.
[20] Wang, G., X. Li, Z. Wang, 2009. APD2: The updated antimicrobial peptide database and its application in peptide design. Nucleic Acids Res., 37: D933-D937.
[21] Thomas, S., S. Karnik, R.S. Barai, V.K. Jayaraman, S. Idicula-Thomas, 2010. CAMP: a useful resource for research on antimicrobial peptides. Nucleic acids Res., 38: D774-D780.
[22] Li, W., A. Godzik, 2006. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics, 22: 1658-1659.
[23] Moody, J., C.J. Darken, 1989. Fast learning in networks of locally-tuned processing units. Neural Comput 1: 281-294.
[24] Park, J., I.W. Sandberg, 1991. Universal approximation using radial basis-function networks. Neural Comput., 3: 246-257.
[25] Duda, R.O., P.E. Hart, D.G. Stork, 2000. Pattern Classification, 2nd edition. John Wiley and Sons.
[26] Oqul, H., A.T. Kalkan, S.U. Umu, M.S. Akkaya, 2013. Trainer: A general-purpose trainable short biosequence classifier. Protein Pept. Lett., [Epub ahead of print].
[27] Sundelacruz, S., M. Levin, D.L. Kaplan, 2009. Role of membrane potential in the regulation of cell proliferation and differentiation. Stem. Cell Rev. Rep., 5: 231-246.
[28] Mohabatkar, H., 2010. Prediction of cyclin proteins using Chou's pseudo amino acid composition. Protein Pept. Lett., 17: 1207-1214.

[29] Esmaeili, M., H. Mohabatkar, S. Mohsenzadeh, 2010. Using the concept of Chou's pseudo amino acid composition for risk type predictionof human papillomaviruses. J. Theor. Biol., 263: 203-209.
[30] Mohabatkar, H., M. Mohammad-Beigi, A. Esmaeili, 2011. Prediction of GABAA receptor proteins using the concept of Chou's pseudoaminoacid composition and support vector machine. J. Theor. Biol., 281: 18-23.
[31] Mohammad-Beigi, M., M. Behjati, H. Mohabatkar, 2011. Prediction of metalloproteinase family based on the concept of Chou's pseudoamino acid composition using a machine learning approach. J, Struct. Funct. Genomics, 12: 191-197.
[32] Khosravian, M., F.K. Faramarzi, M. Mohammad-Beigi, M. Behbahani, H. Mohabatkar, 2013. Predicting antibacterial peptides by the concept of Chou's pseudo-amino acid composition and machine learning methods. Protein Pept. Lett., 20: 1-7.
[33] Mohabatkar, H., M. Mohammad-Beigi, K. Abdolahi, S. Mohsenzadeh, 2013. Prediction of allergenic proteins by means of the concept of Chous pseudo amino acid composition and a machine learning approach. Med. Chem., 9: 133-137.