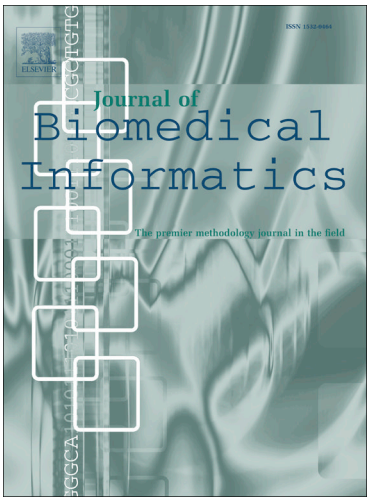

Prediction of therapeutic peptides by incorporating q-Wiener index into Chou's general PseAAC

Chunrui Xu, Li Ge, Yusen Zhang, Matthias Dehmer, Ivan Gutman

PII: S1532-0464(17)30212-5
DOI: https://doi.org/10.1016/j.jbi.2017.09.011
Reference: YJBIN 2861

To appear in: *Journal of Biomedical Informatics*

Received Date: 14 July 2017
Revised Date: 14 September 2017
Accepted Date: 25 September 2017

Please cite this article as: Xu, C., Ge, L., Zhang, Y., Dehmer, M., Gutman, I., Prediction of therapeutic peptides by incorporating q-Wiener index into Chou's general PseAAC, *Journal of Biomedical Informatics* (2017), doi: https://doi.org/10.1016/j.jbi.2017.09.011

Prediction of therapeutic peptides by incorporating q-Wiener index into Chou's general PseAAC

Chunrui Xu[1], Li Ge[1], Yusen Zhang[1,*] , Matthias Dehmer[2,3], Ivan Gutman[4]

[1] *School of Mathematics and Statistics, Shandong University at Weihai, Weihai 264209, China*

[2] *Department of Mechatronics and Biomedical Computer Science, UMIT, Hall in Tyrol, Austria*

[3] *College of Computer and Control Engineering, Nankai University, Tianjin 300071, China*

[4] *Faculty of Science, University of Kragujevac, Kragujevac, Serbia*

Abstract. As therapeutic peptides have been taken into consideration in disease therapy in recent years, many biologists spent time and labor to verify various functional peptides from a large number of peptide sequences. In order to reduce the workload and increase the efficiency of identification of functional proteins, we propose a sequence-based model, q-FP (functional peptide prediction based on the q-Wiener Index), capable of recognizing potentially functional proteins. We extract three types of features by mixing graphic representation and statistical indices based on the q-Wiener index and physicochemical properties of amino acids. Our support–vector–machine–based model achieves an accuracy of 96.71%, 92.52%, 98.40%, and 91.40% for anticancer, virulent, and allergenic proteins datasets, respectively, by using 5-fold cross validation.

Key words: q-Wiener index, Graphic representation, Physicochemical property, Support vector machine, Protein prediction

1 Introduction

Identifying specific peptides from massive proteins is a challenging task for cancer treatment, drug design, and vaccine production.

Anticancer peptides (ACPs), which are small molecules consisting of amino acids, can function against various tumors via disruption of mitochondria or membranolytic

*Corresponding author: zhangys@sdu.edu.cn

†Code is programmed by MATLAB, which can be downloaded from `https://sourceforge.net/projects/application-of-q-wiener-index/files//?upload_just_completed=true`