How the artificial intelligence tool iHyd-PseAAC is working in predicting the hydroxyproline and hydroxylysine in proteins

Short Communication

Kuo-Chen Chou*
Gordon Life Science Institute, Massachusetts, USA

Received: Mar 06, 2020; Accepted: Mar 20, 2020; Published: Mar 23, 2020

*Corresponding author: Kuo-Chen Chou, Gordon Life Science Institute, Boston, Massachusetts 02478, United States of America

Copyright: © 2020 Kuo-Chen Chou. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

In 2014 a very powerful AI (artificial intelligence) tool has been established for identifying hydroxyproline and hydroxylysine sites in proteins, two of the important post modifications in proteins [1].

To see how the web-server is working, please do the following.

Step 1: Open the web server at the site at http://app.aporc.org/iHyd-PseAAC/ and you will see the top page of the predictor on your computer screen, as shown in Figure 1. Click on the Read Me button to see a brief introduction about iHyd-PseAAC predictor and the caveat when using it.

Step 2: Either type or copy/paste the query protein sequences into the input box at the center of Figure 1. The protein sequences should be in FASTA format. The input examples can be seen by clicking on the Example button right above the input box.

Step 3: Click on the Submit button to see the predicted result. For instance, if you use the protein sequences in the Example window as the input, after a few seconds, you will see the corresponding predicted results, which is fully consistent with experiment observations.

Step 4: Click on the Citation button to find the relevant paper that documents the detailed development and algorithm of iHyd-PseAAC.

Step 5: Click on the Data button to download the benchmark dataset used to train and test the iHyd-PseAAC predictor.

It is anticipated that the Web-Server will be very useful because the vast majority of biological scientists can easily get their desired results without the need to go through the complicated equations in [2] that were presented just for the integrity in developing the predictor.

Also, note that the web-server predictor has been developed by strictly observing the guidelines of “Chou’s 5-steps rule” and hence have the following notable merits (see, e.g., [3,4] and three comprehensive review papers [5-7]): (1) crystal clear in logic development, (2) completely transparent in operation, (3) easily to repeat the reported results by other investigators, (4) with high potential in stimulating other sequence-analyzing methods, and (5) very convenient to be used by the majority of experimental scientists.

It has not escaped our notice that during the development of iDNA6mA-PseKNC web-server, the approach of general pseudo amino acid components [8] or PseAAC [9] had been utilized and hence its accuracy would
be much higher than its counterparts, as concurred by many investigators (see, e.g., [10-13]).

For the marvelous and awesome roles of the "5-steps rule" in driving proteome, genome analyses and drug development, see a series of recent papers [14-35] where the rule and its wide applications have been very impressively presented from various aspects or at different angles.

References

3. O. Barukab, Y.D. Khan, S.A. Khan, K.C. Chou, iSulfoty-PseAAC: Identify tyrosine sulfation sites by incorporating statistical moments via Chou’s 5-steps rule and pseudo components Current Genomics.
7. K.C. Chou, Impacts of pseudo amino acid components and 5-steps rule to proteomics and proteome analysis, Current Topics in Medicinal Chemistry (CTMC) (Special Issue ed. G.P Zhou).
20. K.C. Chou, Showcase to illustrate how the webserver pLoc_bal-mEuk is working, Biomed J Sci & Tech Res.
23. K.-C. Chou, Showcase to illustrate how the web-server iKcrPseEns is working, Journal of Medical Care Research and Review, 3 (2020) 331-347.
24. K.C. Chou, The pLoc_bal-mVirus is a powerful artificial intelligence tool for predicting the subcellular localization of virus proteins according to their sequence information alone, J Gent & Genome, 4 (2020).
27. K.C. Chou, How the Artificial Intelligence Tool iRNA-PseU is Working in Predicting the RNA Pseudouridine Sites, Biomed J Sci & Tech Res.
28. K.C. Chou, Showcase to illustrate how the web-server iSNO-AAPair is working, J Gent & Genome, 4 (2020).
30. K.C. Chou, Showcase to Illustrate How the Web-server iPTM-mLys is working, Infotext Journal of Infectious Diseases and Therapy [IJID], 1 (2020) 1-16.
32. K.C. Chou, Showcase to illustrate how the web-server iPrenyPseAAC is working, Glo J of Com Sci and Infor Tec., 2 (2020) 01-15.

33. K.C. Chou, Some illuminating remarks on molecular genetics and genomics as well as drug development, Molecular Genetics and Genomics, 295 (2020) 261-274.
