



Showcase to Illustrate how the web-server pLoc_Deep-mGpos is working

By Kuo-Chen Chou

Gordon Life Science Institute

Introduction- In 2020, a very powerful web-server predictor has been established for identifying the subcellular localization of human proteins based on the sequence information alone [1], in which a same protein may occur or move between two or more location sites and hence needs to be marked with the multi-label approach [2].

The web-server predictor is called “pLoc_Deep-mGpos”, where “Deep” means the web-server has been further improved by the “Deep Learning” technique [3-6], and “m” means the capacity able to deal with the multi-label systems. To learn how the web-server is working, please do the following.

Step 1: Click the link at http://www.jci-bioinfo.cn/pLoc_Deep-mGpos/, the top page of the pLoc_bal-mGpos web-server will appear on your computer screen, as shown in **Fig.1**. Click on the Read Me button to see a brief introduction about the predictor.

GJSFR-G Classification: FOR Code: 100499, 080505



Strictly as per the compliance and regulations of:



Showcase to Illustrate how the web-server pLoc_Deep-mGpos is working

Short Title: Showcase for pLoc_Deep-mGpos

Kuo-Chen Chou

INTRODUCTION

In 2020, a very powerful web-server predictor has been established for identifying the subcellular localization of human proteins based on the sequence information alone [1], in which a same protein may occur or move between two or more location sites and hence needs to be marked with the multi-label approach [2].

The web-server predictor is called “pLoc_Deep-mGpos”, where “Deep” means the web-server has been further improved by the “Deep Learning” technique [3-6], and “m” means the capacity able to deal with the multi-label systems. To learn how the web-server is working, please do the following.

Step 1: Click the link at http://www.jci-bioinfo.cn/pLoc_Deep-mGpos/, the top page of the pLoc_bal-mGpos web-server will appear on your computer screen, as shown in Fig.1. Click on the [Read Me](#) button to see a brief introduction about the predictor.

pLoc_Deep-mGpos: predict subcellular localization of gram-positive proteins by deep learning
| [Read Me](#) | [Supporting information](#) | [Citation](#) |

Enter query sequences

Enter the sequences of query proteins in FASTA format ([Example](#)): the number of proteins is limited at 10 or less for each submission.

Or, upload a file for batch prediction

Enter your e-mail address and upload the batch input file ([Batch-example](#)). The predicted result will be sent to you by e-mail once completed; it usually takes 1 minute or so for each protein sequence

Upload file:
Your Email:

Figure 1

Step 2: Either type or copy/paste the sequences of query human proteins into the input box at the center of Fig.1. The input sequence should be in the FASTA format. For the examples of sequences in FASTA format, click 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 four protein sequences in the [Example](#) window as the input, after 10 seconds or so, you will see a new screen (Fig.2) occurring. On its upper

part are listed the names of the subcellular locations numbered from (1) to (8) covered by the current predictor. On its lower part are the predicted results: the query protein P22340 of example-1 corresponds to “2,” meaning it belongs to “Cell outer membrane” only; the query protein P04032 of example-2 corresponds to “8” meaning it belongs to “Periplasm”; the query protein P04825 of example-3 corresponds to “1, 3”, meaning it belongs to “Cell inner membrane” and “Cytoplasm”; the query protein P22251 of example 4 corresponds to “4, 6”, meaning it belongs to “Extracellular” and “Flagellum”. All these results are perfectly consistent with experimental observations.

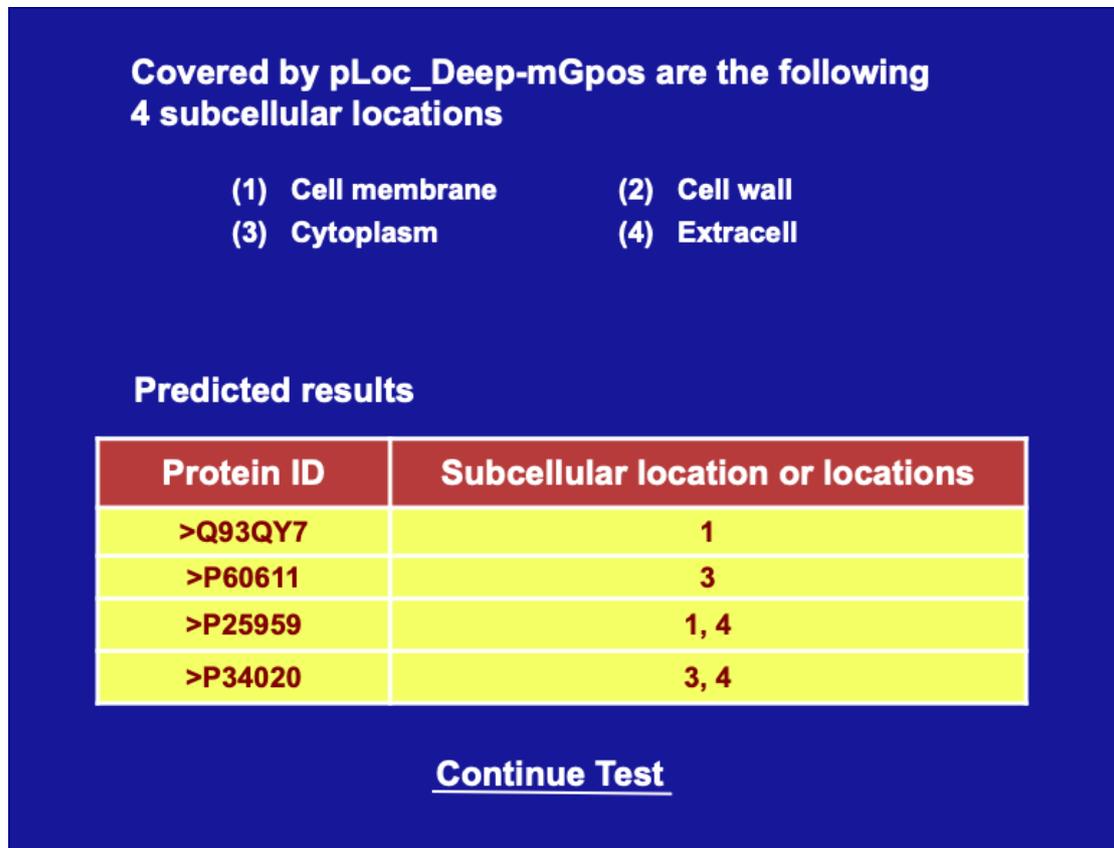


Figure 2

Step 4: As shown on the lower panel of **Fig.2**, you may also choose the batch prediction by entering your e-mail address and your desired batch input file (in FASTA format of course) via the [Browse](#) button. To see the sample of batch input file, click on the button [Batch-example](#). After clicking the button [Batch-submit](#), you will see “Your batch job is under computation; once the results are available, you will be notified by e-mail.”

Step 5: Click on the [Citation](#) button to find the papers that have played the key role in developing the current predictor of pLoc_Deep-mGpos.

Step 6: Click the Supporting Information button to download the Supporting Informations mentioned in this paper.

REFERENCES RÉFÉRENCES REFERENCIAS

1. Z. Lu, K.C. Chou, Showcase to illustrate how the web-server pLoc_Deep-mGpos is working, Journal of Biomedical science and Engineering (JBiSE) 13 (2020) 55-65.
2. K.C. Chou, Advance in predicting subcellular localization of multi-label proteins and its implication for developing multi-target drugs, Current Medicinal Chemistry 26 (2019) 4918-4943.
3. A. Maxwell, R. Li, B. Yang, H. Weng, A. Ou, H. Hong, Z. Zhou, P. Gong, C. Zhang, Deep learning architectures for multi-label classification of intelligent health risk prediction. , Bmc Bioinformatics, 18 (2017) 523.
4. S. Khan, M. Khan, N. Iqbal, T. Hussain, S.A. Khan, K.C. Chou, A Two-Level Computation Model Based on Deep Learning Algorithm for Identification of piRNA and Their Functions via Chou's 5-Steps Rule (2019) <https://doi.org/10.1007/s10989-019-09887-3>, Human Genetics (2019).

5. I. Nazari, M. Tahir, H. Tayari, K.T. Chong, iN6-Methyl (5-step): Identifying RNA N6-methyladenosine sites using deep learning mode via Chou's 5-step rules and Chou's general PseKNC, Chemometrics and Intelligent Laboratory Systems (CHEMOLAB), <https://doi.org/10.1016/j.chemolab.2019.103811> (2019).
6. Z.U. Khan, F. Ali, I.A. Khan, Y. Hussain, D. Pi, iRSpot-SPI: Deep learning-based recombination spots prediction by incorporating secondary sequence information coupled with physio-chemical properties via Chou's 5-step rule and pseudo components, Chemometrics and Intelligent Laboratory Systems (CHEMOLAB), 189 (2019) 169-180.