GPS-YNO2 Manual

Prediction of Tyrosine Nitration Sites

Version 1.0.1
14/08/2014

Author: Zexian Liu, Jian Ren & Yu Xue

Contact:
Dr. Zexian Liu, lzx.bioinfo@gmail.com
Dr. Jian Ren, renjian.sysu@gmail.com
Dr. Yu Xue, xueyu@mail.hust.edu.cn
The software is only free for academic research.
The latest version of GPS-YNO2 software is available from http://yno2.biocuckoo.org

Index

Statement

1. **Implementation**. The softwares of the CUCKOO Workgroup are implemented in JAVA (J2SE). Usually, both of online service and local stand-alone packages will be provided.
2. **Availability**. Our softwares are freely available for academic researches. For non-profit users, you can copy, distribute and use the softwares for your scientific studies. Our softwares are not free for commercial usage.
3. **GPS**. Previously, we used the GPS to denote our Group-based Phosphorylation Scoring algorithm. Currently, we are developing an integrated computational platform for post-translational modifications (PTMs) of proteins. We re-denote the GPS as Group-based Prediction Systems. This software is an indispensable part of GPS.
4. **Usage**. Our softwares are designed in an easy-to-use manner. Also, we invite you to read the manual before using the softwares.
5. **Updation**. Our softwares will be updated routinely based on users' suggestions and advices. Thus, your feedback is greatly important for our future updation. Please do not hesitate to contact with us if you have any concerns.
6. **Citation**. Usually, the latest published articles will be shown on the software websites. We wish you could cite the article if the software has been helpful for your work.
7. **Acknowledgements**. The work of CUCKOO Workgroup is supported by grants from the National Basic Research Program (973 project) (2006CB933300, 2007CB947401, 2007CB914503, and 2010CB912103), Natural Science Foundation of China (90919001, 30700138, 30900835, 30830036, 30721002, 30871236, and 90913016), Chinese Academy of Sciences (KSCX1-YW-R65, KSCX2-YW-R-139, INFO-115-C01-SDB4-36), and National Science Foundation for Post-doctoral Scientists (20080430100).

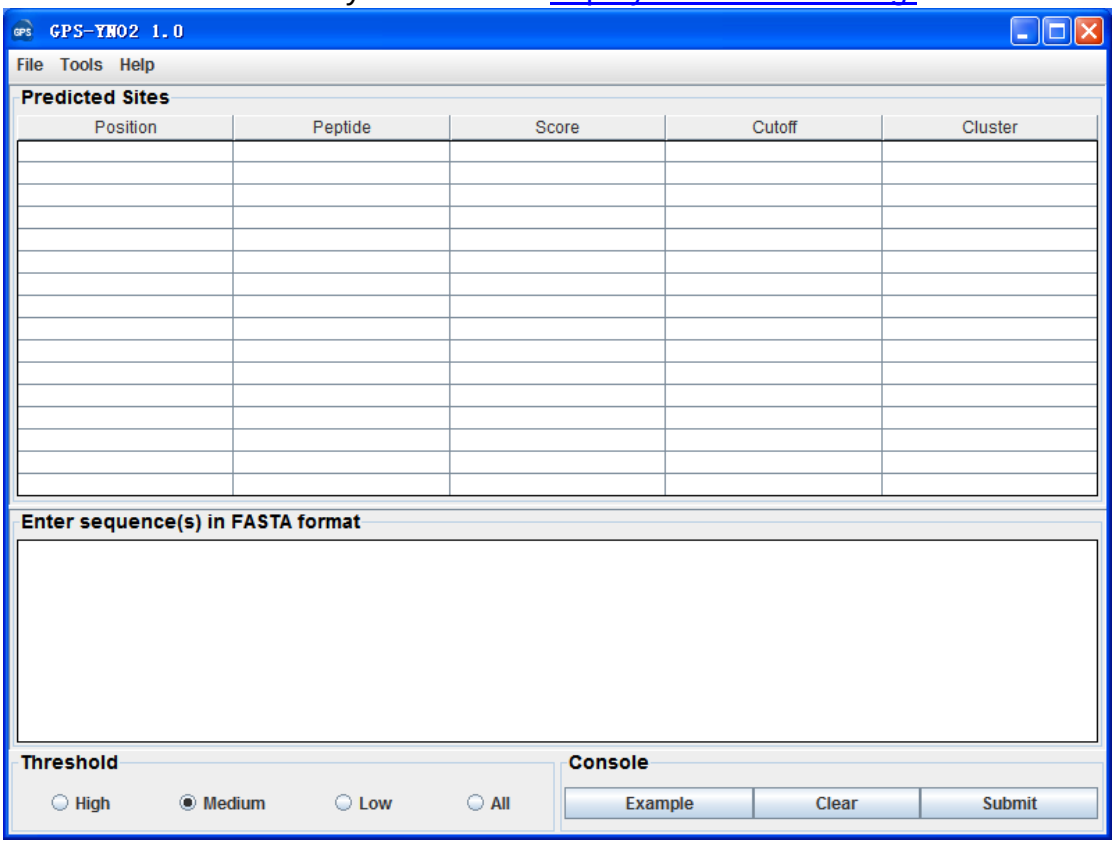
Introduction

The 1998 Nobel Prize in Physiology or Medicine was awarded for discoveries concerning nitric oxide (NO) as a signaling molecule in the cardiovascular system. NO acts as a freely-diffusible signaling molecule and second messenger which can regulates the production of cyclic GMP (cGMP). The following studies showed that an interplay involving excess NO, transition metal centers and oxidants may generate protein cysteine *S*-nitrosylation as well as protein tyrosine nitration (PTN) (1-5). Though lots of work contributed to dissect the mechanisms of PTN previously (6-8), our understanding of PTN is still fragmentary. However, it is noticed that when oxidants, such as superoxide radicals ($O_2^{\bullet-}$) and hydrogen peroxide (H_2O_2), are presented with transition metal centers in NO metabolism, reactive nitrogen species such as peroxynitrite anion ($ONOO^-$) and nitrogen dioxide (NO_2) will format, which resulted in nitration of protein tyrosines (Figure 1). Although this process was questioned by the complexity of the cellular environment, for instance, the transition metal could be replace by heme peroxidase, it was assured that PTN could be triggered *in vivo*. Though originally addressed in early *in vitro* protein chemistry studies, recent studies led to the discovery that cellular PTN has important implications on histone modification (9), protein activity regulation (10), epitope recognition (11) and so on. Furthermore, PTN showed its significance in biochemical processes critical for physiology and pathology including signal transduction, immune response, cell death, aging, neurodegeneration and so on (1-5). However, as a covalent modification linked to NO signaling pathway, our understanding of PTN and its relationship with *S*-nitrosylation is still fragmentary.

Combined with the conventional experimental identification of substrates for PTN (3,10), large-scale detection of cellular nitrated proteins was introduced for a global view with the development of the biotechnology including the antibodies to recognize the nitrotyrosine, the method for selective enrichment of nitrotyrosine-containing peptides and the mass spectrometer (12-14). In light of the advance of biotechnologies, lots of studies touched on systematic analysis of nitrated proteins to provide insights into the biological roles of PTN (13,15-20). Previously, Souza *et al*. and Elfering *et al*. investigated the consensus sequence around the nitration sites, which resulted in opposed conclusions (21-22). However, a recent proteomic work with 335 PTN sites in 267 proteins from human Jurkat cells of which lysate was treated by peroxynitrite resulted in no conclusive sequence preference for PTN (16). Nowadays, the desire to map the PTM sites without the time-consuming and expensive experimental methods has motivated the development of computational approaches, which was shown to be able to rapidly generate helpful information for further experimental verification.

Although there have been ~170 databases and computational tools developed for PTM analyses especially phosphorylation (http://www.biocuckoo.org/link.php) (23), *in silico* prediction of tyrosine nitration sites in proteins is still a great challenge.

In this work, 1,066 experimentally verified PTN sites within 554 unique proteins were collected from the scientific literature and public databases. In our previous study, the GPS 2.0 (Group-based Prediction System) algorithm was developed for the prediction of kinase-specific phosphorylation sites (24), and we just improved the method substantially and released the GPS 3.0 algorithm for prediction of *S*-nitrosylation sites (25). Here we developed a novel computational software of GPS-YNO2 1.0 for prediction of tyrosine nitration sites. The leave-one-out validation and 4-, 6-, 8- and 10-fold cross-validations were adopted to evaluate the prediction performance and system robustness. The performance of the GPS-YNO2 1.0 was satisfying, with an accuracy of 76.51%, a sensitivity of 50.09% and a specificity of 80.18% under the low threshold condition. As applications of GPS-YNO2 1.0, we predicted 325 (~88%) of these the collected substrates without *bona fide* sites with at least one potential nitration site. Obviously, these prediction results might be of great help for further experimental verification. Finally, the online service and local packages of GPS-YNO2 1.0 were implemented in JAVA 1.4.2 and are freely available at: http://yno2.biocuckoo.org/.

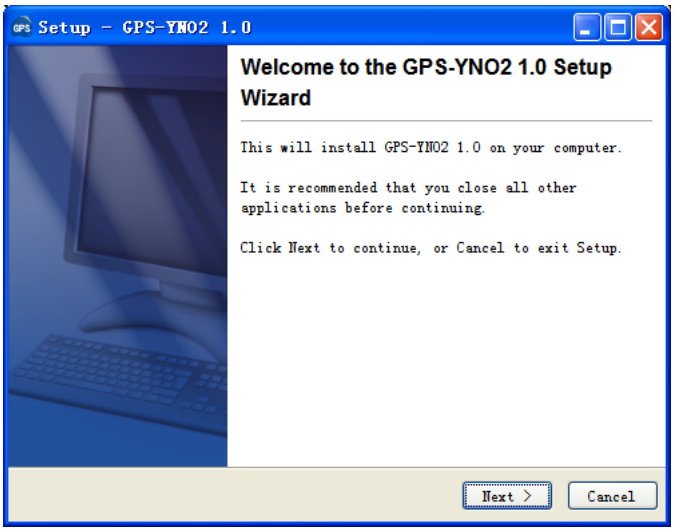


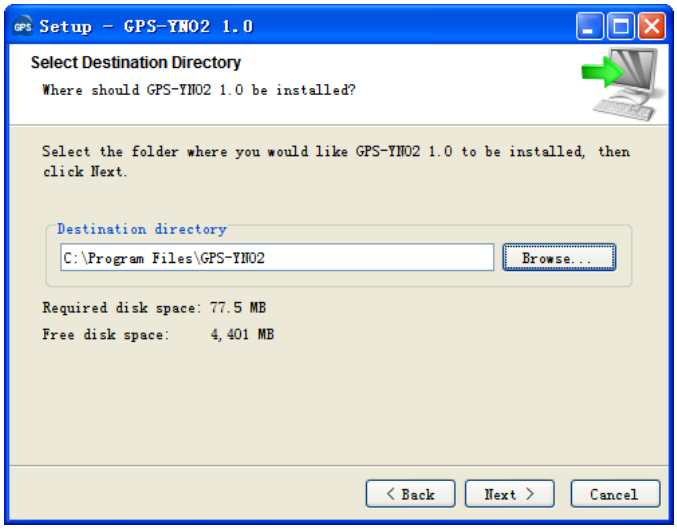
GPS-YNO2 1.0 User Interface

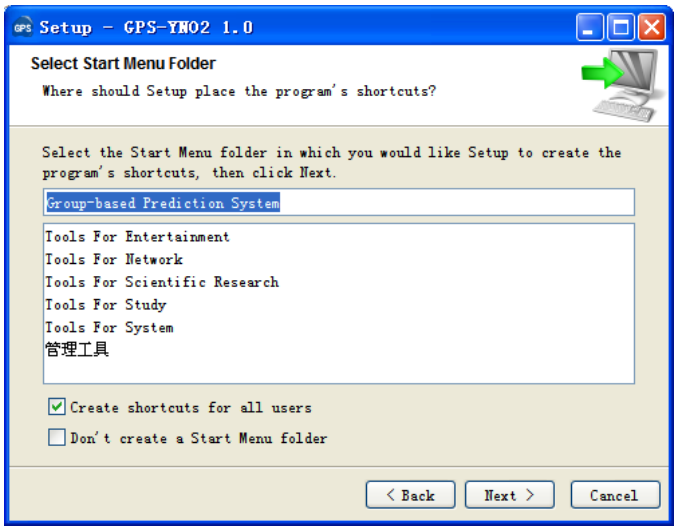
Download & Installation

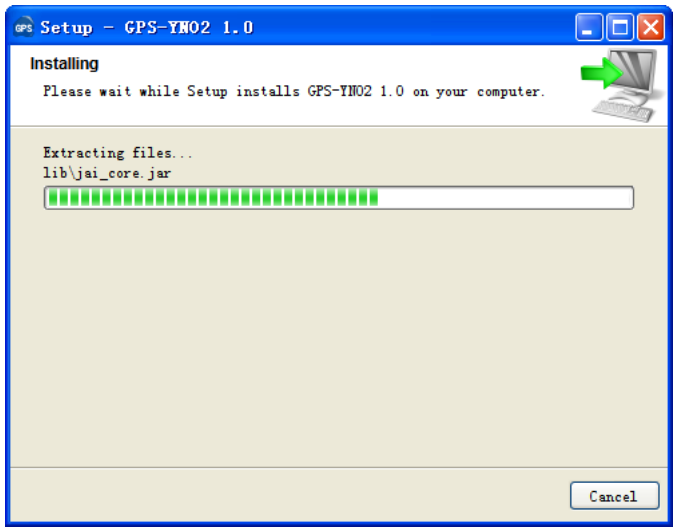
The GPS-YNO2 1.0 was implemented in JAVA (J2SE), and could support three major Operating Systems (OS), including Windows, Linux/Unix or Mac OS X systems. Both of online web service and local stand-alone packages are available from: http://yno2.biocuckoo.org/ . We recommend that users could download the latest release.

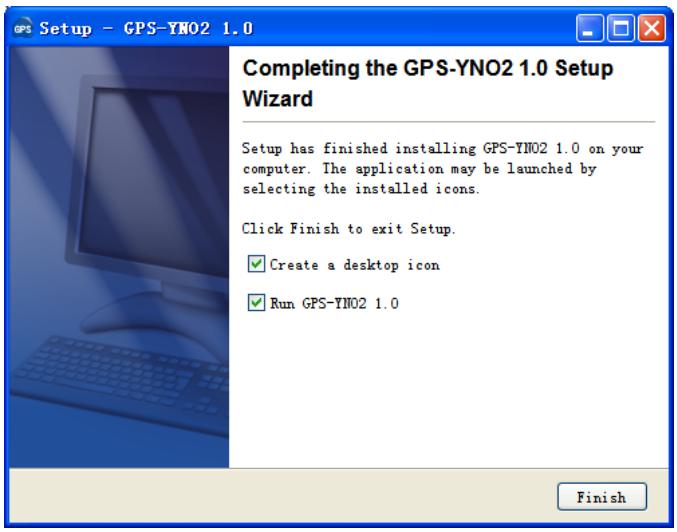
Please choose the proper package to download. After downloading, please double-click on the software package to begin installation, following the user prompts through the installation. And snapshots of the setup program for windows are shown below:











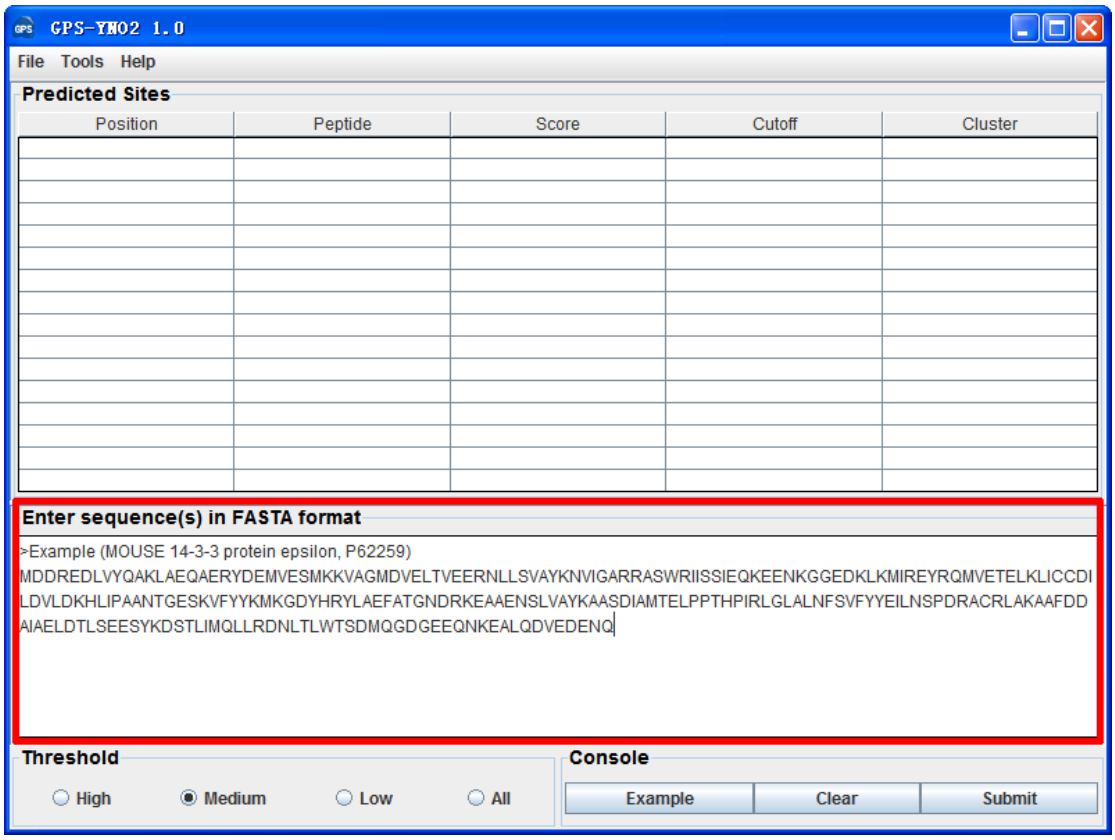
Finally, please click on the **Finish** button to complete the setup program.

Prediction of Tyrosine Nitration Sites

1. A single protein sequence in FASTA format

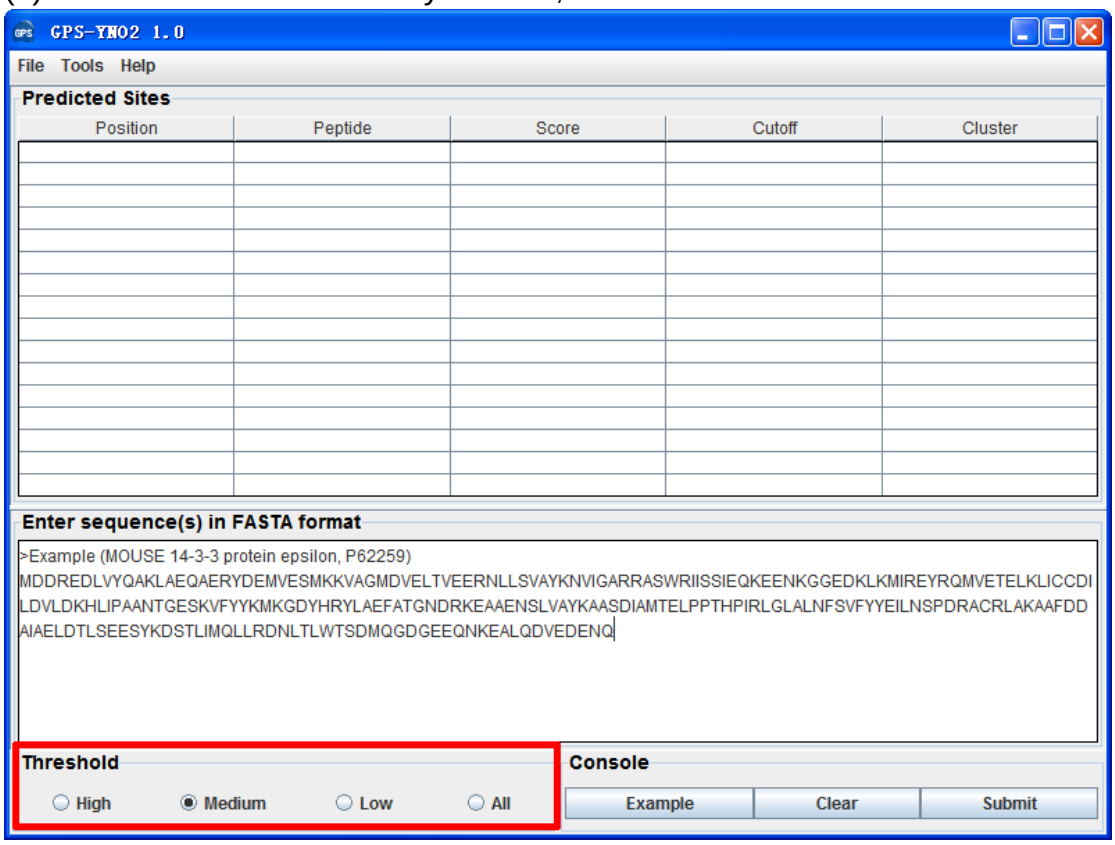
The following steps show you how to use the GPS-YNO2 1.0 to predict tyrosine nitration sites for a single protein sequence in FASTA format.

(1) Firstly, please use “Ctrl+C & Ctrl+V” (Windows & Linux/Unix) or “Command+C & Command+V” (Mac) to copy and paste your sequence into the text form of GPS-YNO2 1.0

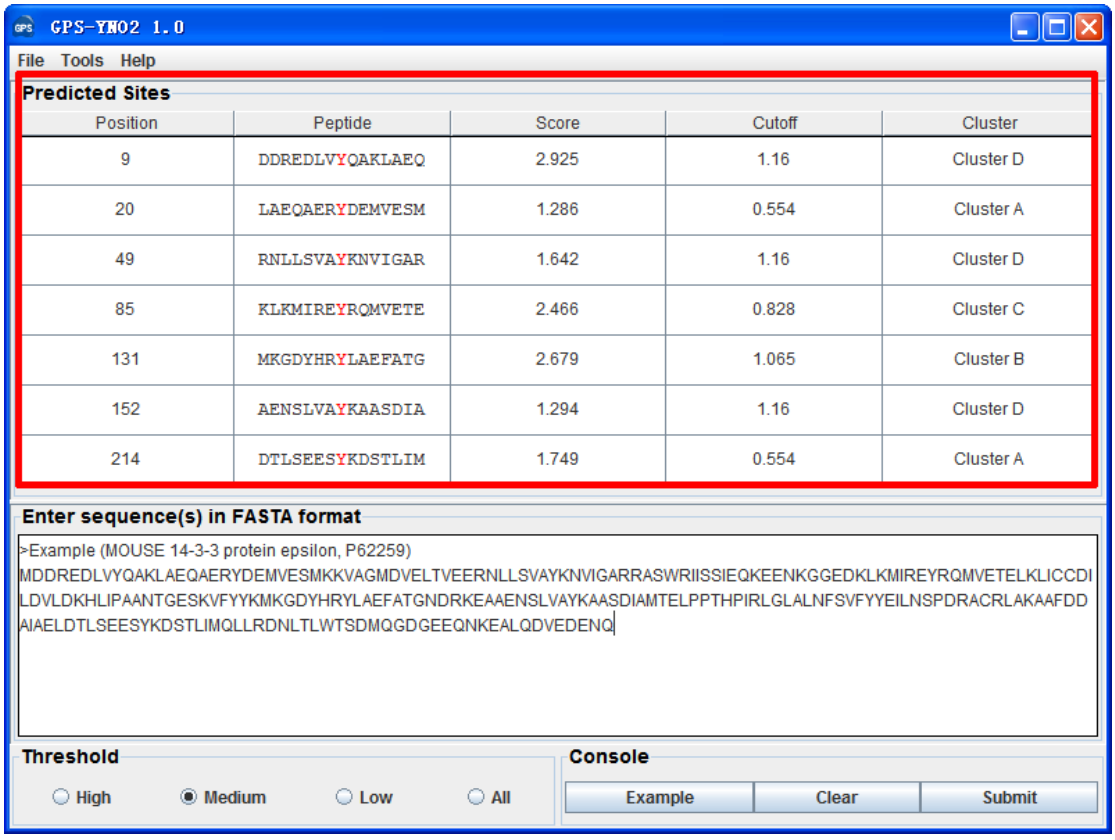


Note: for a single protein, the sequence without a name in raw format is also OK. However, for multiple sequences, the name of each protein should be presented.

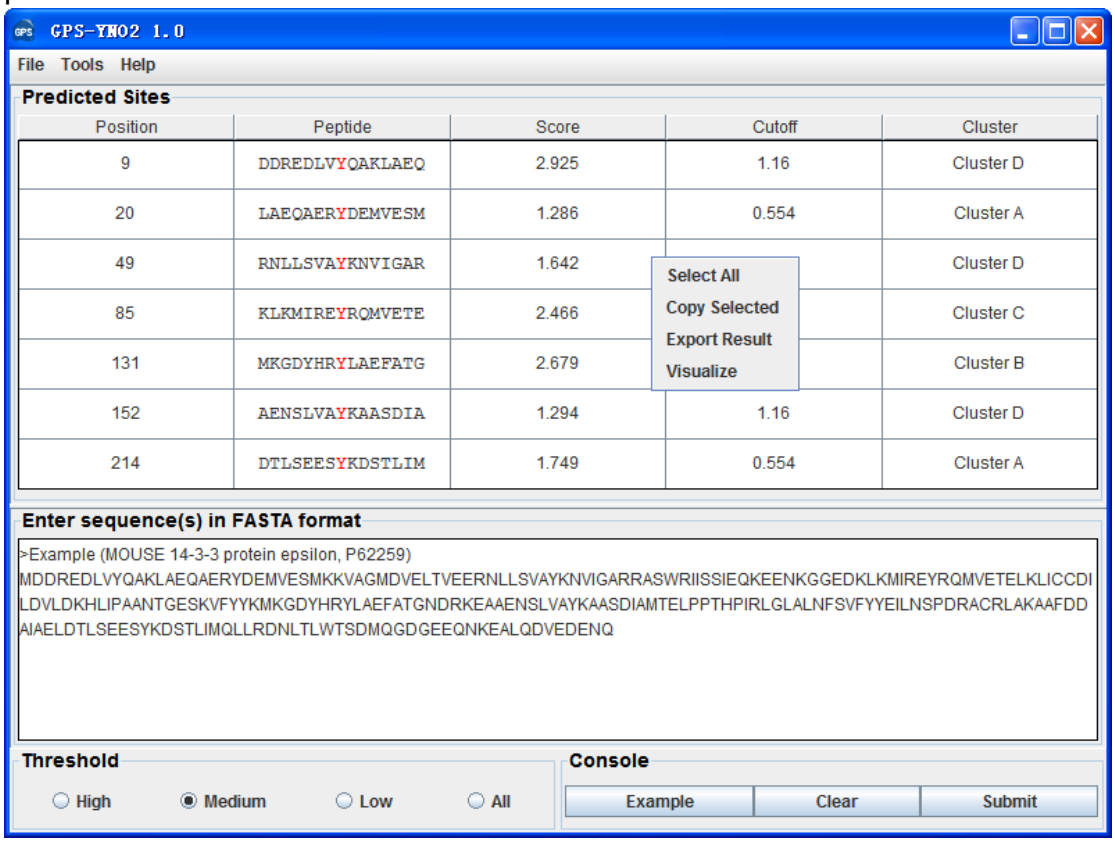
(2) Choose a **Threshold** that you need, the default cut-off is **Medium**.



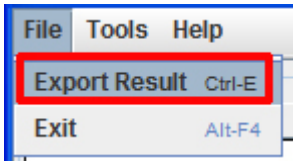
(3) Click on the **Submit** button, then the predicted tyrosine nitration sites will be shown.



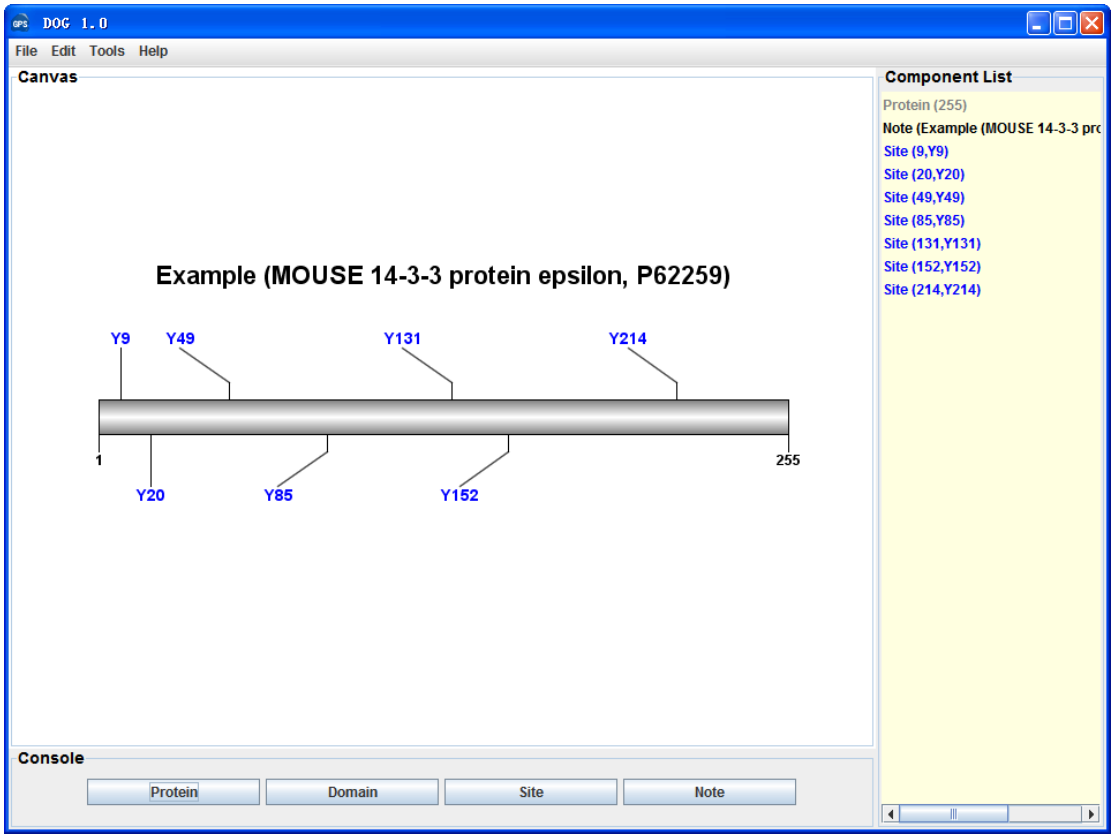
(4) Then please click on the **RIGHT** button in the prediction form. You can use the "**Select All**" and "**Copy Selected**" to copy the selected results into Clipboard. Then please copy the results into a file, eg., an EXCEL file for further consideration. Also, you can choose "**Export Prediction**" to export the prediction results into a tab-delimited text file.



Again, you can also click the "**Export Prediction**" in **File** menu to export the results.



If you choose the Visualize function, the given protein and its predicted sites will be visualized with DOG (Domain Graph, Version 1.0), an illustrator of protein domain structures.

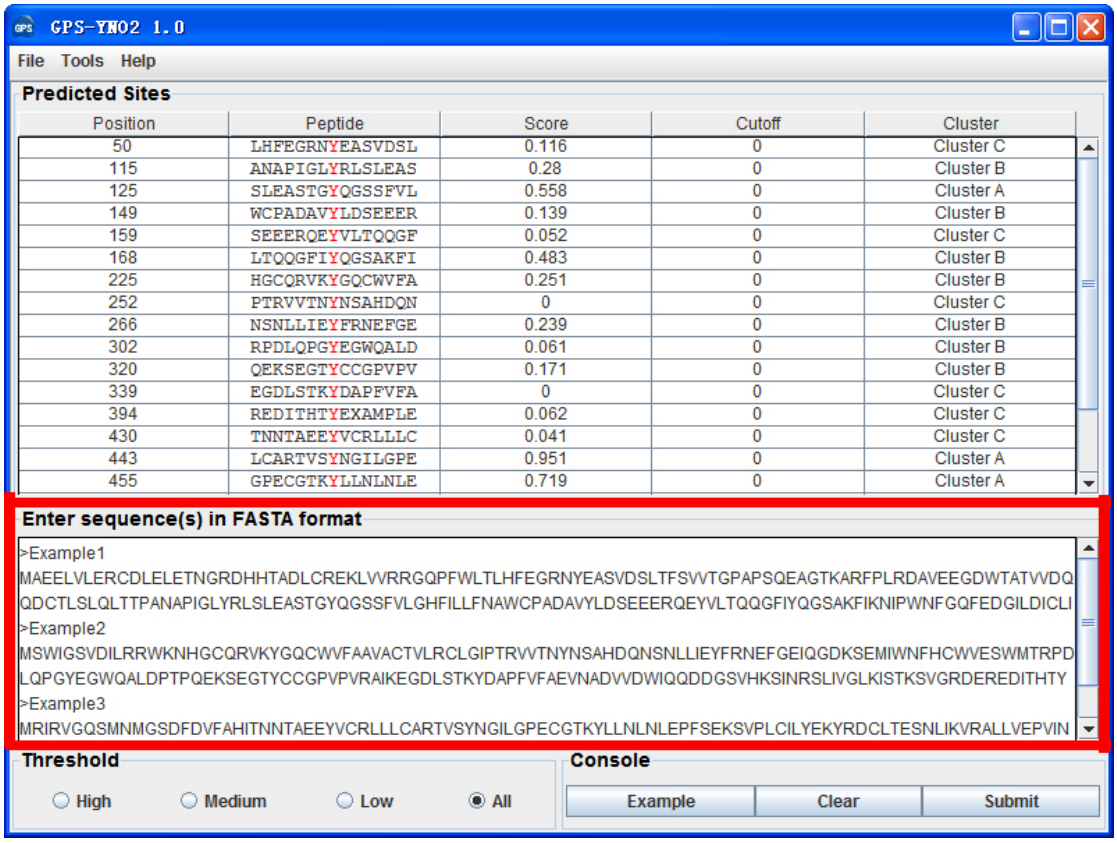


2. Multiple protein sequences in FASTA format

For multiple protein sequences, there are two ways to use the GPS-YNO2 1.0.

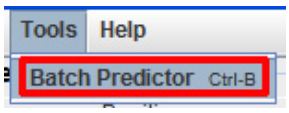
A. Input the sequences into text form directly. (Num. of Seq ≤ 2,000)

If the number of total protein sequences is not greater than 2,000, you can just use “Ctrl+C & Ctrl+V” (Windows & Linux/Unix) or “Command+C & Command+V” (Mac) to copy and paste your sequences into the text form of GPS-YNO2 1.0 for prediction.



B. Use Batch Predictor tool.

If the number of protein sequences is very large, eg., yeast or human proteome, please use the **Batch Predictor**. Please click on the “**Batch Predictor**” button in the **Tools** menu.



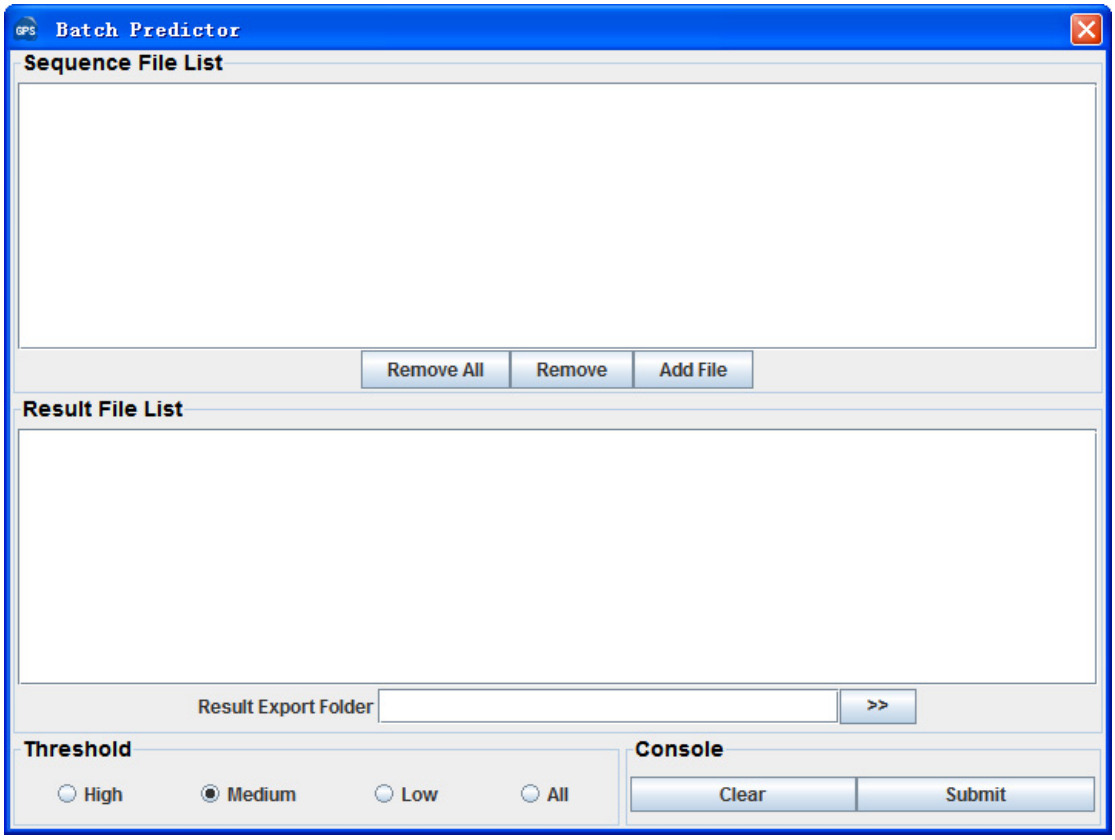
The following steps show you how to use it:

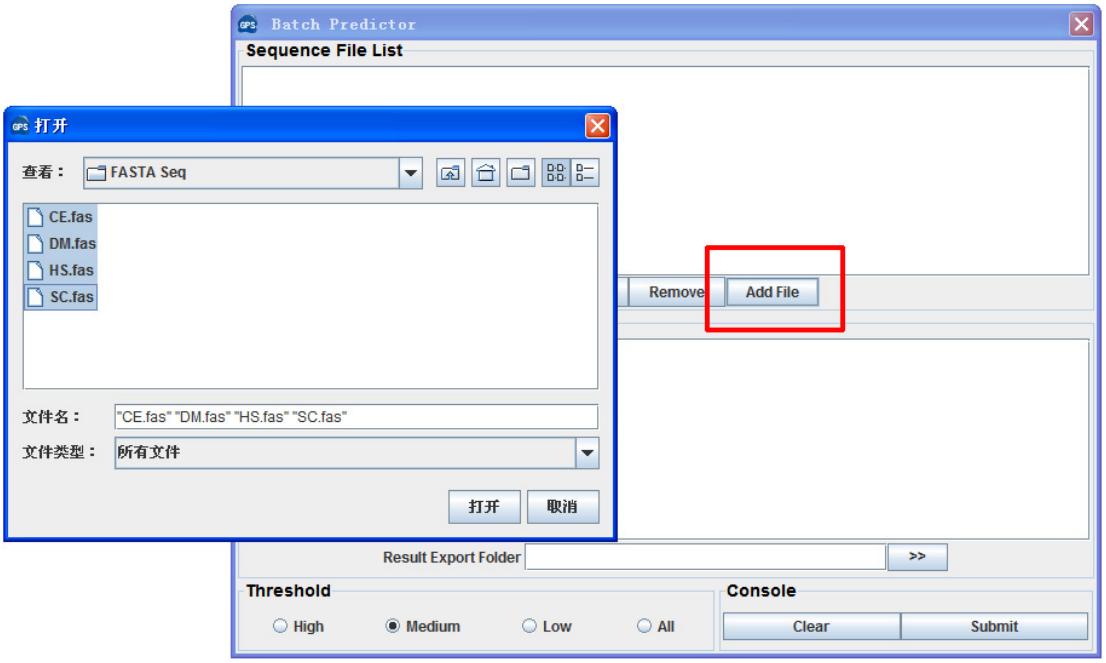
(1) Put protein sequences into one or several files (eg., SC.fas, CE.fas, and etc)

with FATSA format as below:

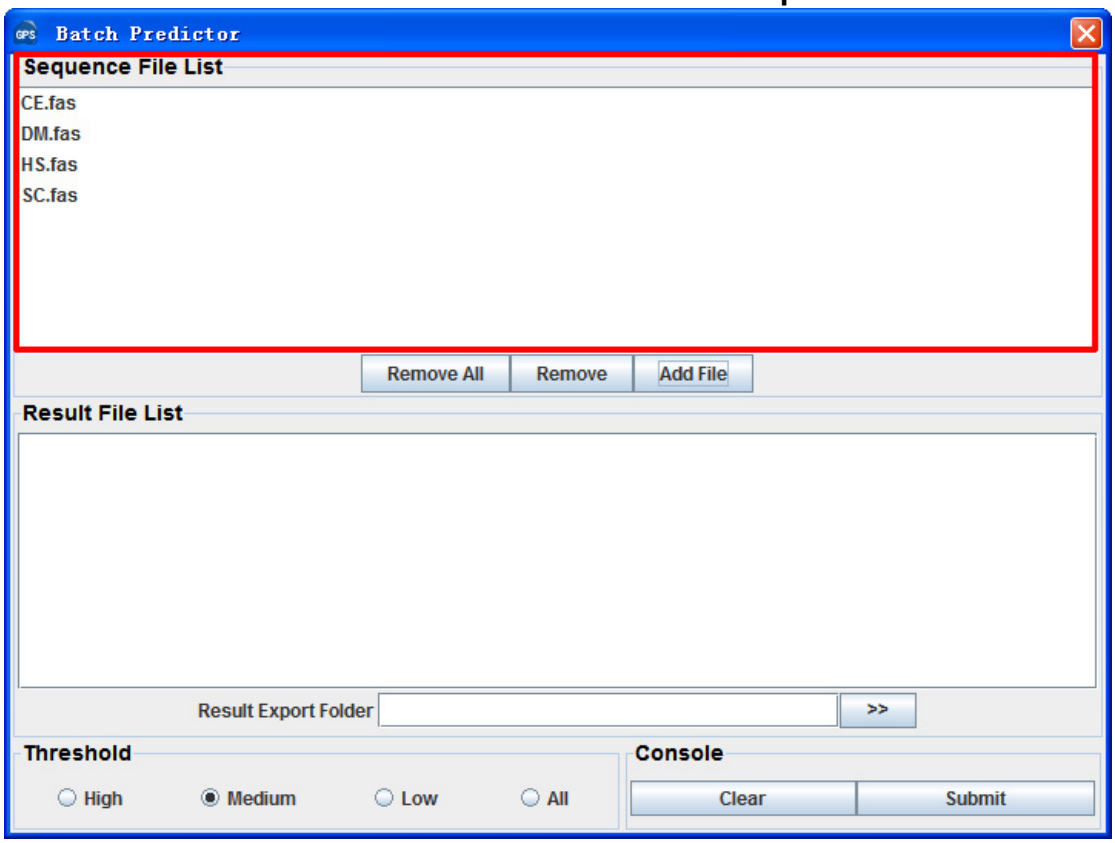
>protein1
XXXXXXXXXXXX
XXXXXXXX
>protein2
XXXXXXXXXXXXXXX…
>protein3
XXXXXXXXXXX
...
Most importantly, the name of each protein should be presented.

(2) Click on the **Batch Predictor** button and then click on the **Add File** button and add one or more protein sequence files in your hard disk.

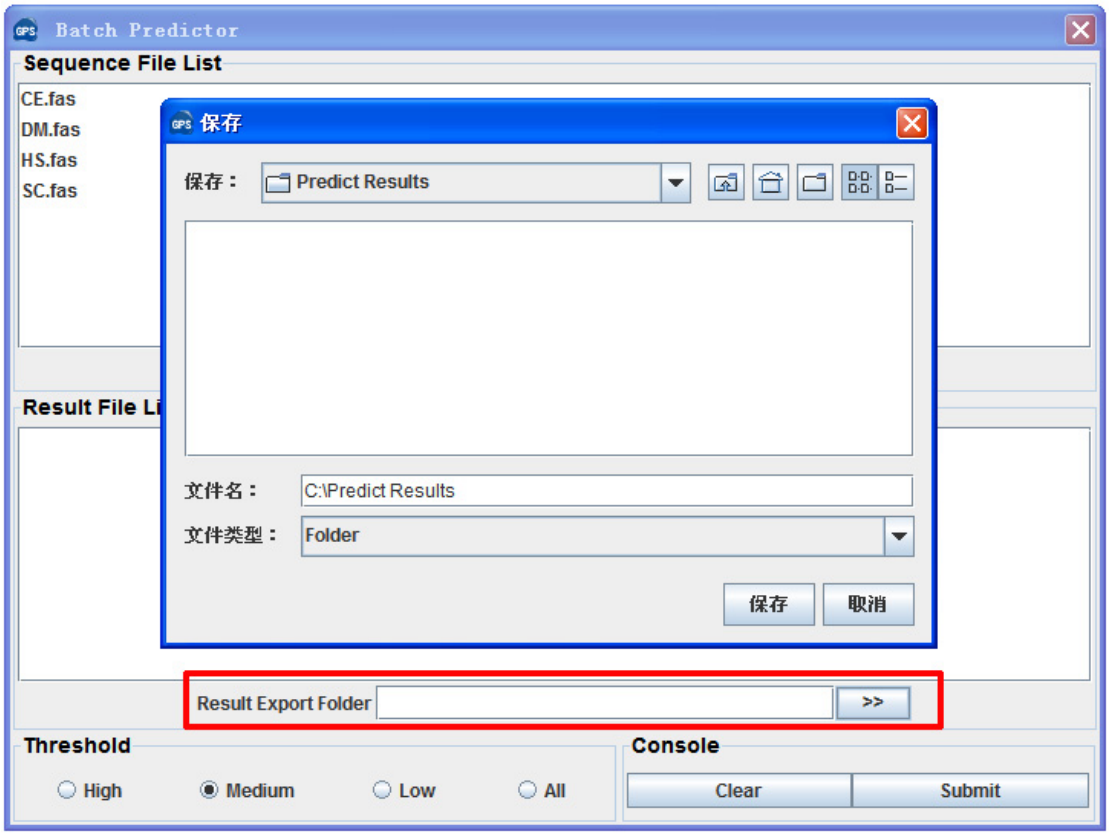




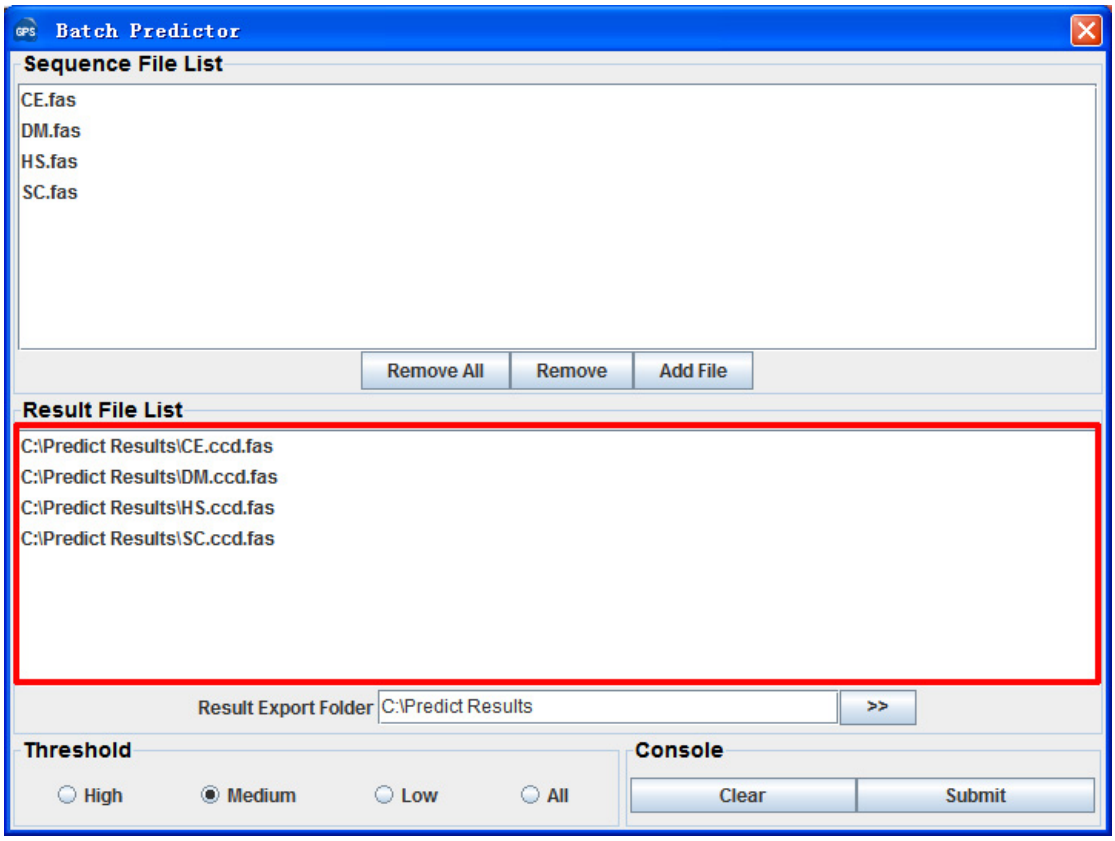
Then the names of added files will be shown in the **Sequence File List**.



(3) The output directory of prediction results should also be defined. Please click on the **>>** button to specify the export fold.



(4) Please choose a proper threshold before prediction. Then please click on the **Submit** button, then the **Batch Predictor** begin to process all of the sequence files that have been added to the list. The result of prediction will be export to the **Prediction Export Fold,** and the name of result files will be shown in the **Prediction File List**.


Batch Predictor
Sequence File List
CE.fas
DM.fas
HS.fas
SC.fas
Remove All
Remove
Add File
Result File List
C:\Predict Results\CE.ccd.fas
C:\Predict Results\DM.ccd.fas
C:\Predict Results\HS.ccd.fas
C:\Predict Results\SC.ccd.fas
Result Export Folder
C:\Predict Results
>>
Threshold
High
Medium
Low
All
Console
Clear
Submit

References

1. Radi, R. (2004) Nitric oxide, oxidants, and protein tyrosine nitration. *Proc Natl Acad Sci U S A*, **101**, 4003-4008.
2. Nagaraj, S., Gupta, K., Pisarev, V., Kinarsky, L., Sherman, S., Kang, L., Herber, D.L., Schneck, J. and Gabrilovich, D.I. (2007) Altered recognition of antigen is a mechanism of CD8+ T cell tolerance in cancer. *Nat Med*, **13**, 828-835.
3. Kers, J.A., Wach, M.J., Krasnoff, S.B., Widom, J., Cameron, K.D., Bukhalid, R.A., Gibson, D.M., Crane, B.R. and Loria, R. (2004) Nitration of a peptide phytotoxin by bacterial nitric oxide synthase. *Nature*, **429**, 79-82.
4. Beckman, J.S. and Koppenol, W.H. (1996) Nitric oxide, superoxide, and peroxynitrite: the good, the bad, and ugly. *Am J Physiol*, **271**, C1424-1437.
5. Schopfer, F.J., Baker, P.R. and Freeman, B.A. (2003) NO-dependent protein nitration: a cell signaling event or an oxidative inflammatory response? *Trends Biochem Sci*, **28**, 646-654.
6. Thomas, D.D., Espey, M.G., Vitek, M.P., Miranda, K.M. and Wink, D.A. (2002) Protein nitration is mediated by heme and free metals through Fenton-type chemistry: an alternative to the NO/O2- reaction. *Proc Natl Acad Sci U S A*, **99**, 12691-12696.
7. Bian, K., Gao, Z., Weisbrodt, N. and Murad, F. (2003) The nature of heme/iron-induced protein tyrosine nitration. *Proc Natl Acad Sci U S A*, **100**, 5712-5717.
8. Goldstein, S., Czapski, G., Lind, J. and Merenyi, G. (2000) Tyrosine nitration by simultaneous generation of (.)NO and O-(2) under physiological conditions. How the radicals do the job. *J Biol Chem*, **275**, 3031-3036.
9. Dixit, K., Khan, M.A., Sharma, Y.D., Uddin, M. and Alam, K. (2009) Physicochemical studies on peroxynitrite-modified H3 histone. *Int J Biol Macromol*.
10. Zaragoza, R., Torres, L., Garcia, C., Eroles, P., Corrales, F., Bosch, A., Lluch, A., Garcia-Trevijano, E.R. and Vina, J.R. (2009) Nitration of cathepsin D enhances its proteolytic activity during mammary gland remodelling after lactation. *Biochem J*, **419**, 279-288.
11. Hardy, L.L., Wick, D.A. and Webb, J.R. (2008) Conversion of tyrosine to the inflammation-associated analog 3'-nitrotyrosine at either TCR- or MHC-contact positions can profoundly affect recognition of the MHC class I-restricted epitope of lymphocytic choriomeningitis virus glycoprotein 33 by CD8 T cells. *J Immunol*, **180**, 5956-5962.
12. Aulak, K.S., Miyagi, M., Yan, L., West, K.A., Massillon, D., Crabb, J.W. and Stuehr, D.J. (2001) Proteomic method identifies proteins nitrated in vivo during inflammatory challenge. *Proc Natl Acad Sci U S A*, **98**, 12056-12061.
13. Tyther, R., Ahmeda, A., Johns, E. and Sheehan, D. (2007) Proteomic identification of tyrosine nitration targets in kidney of spontaneously hypertensive rats. *Proteomics*, **7**, 4555-4564.
14. Zhang, Q., Qian, W.J., Knyushko, T.V., Clauss, T.R., Purvine, S.O., Moore, R.J., Sacksteder, C.A., Chin, M.H., Smith, D.J., Camp, D.G., 2nd *et al.* (2007) A method for selective enrichment and analysis of nitrotyrosine-containing peptides in complex proteome samples. *J Proteome Res*, **6**, 2257-2268.
15. Casoni, F., Basso, M., Massignan, T., Gianazza, E., Cheroni, C., Salmona, M., Bendotti, C. and Bonetto, V. (2005) Protein nitration in a mouse model of familial amyotrophic lateral

sclerosis: possible multifunctional role in the pathogenesis. *J Biol Chem*, **280**, 16295-16304.

16. Ghesquiere, B., Colaert, N., Helsens, K., Dejager, L., Vanhaute, C., Verleysen, K., Kas, K., Timmerman, E., Goethals, M., Libert, C. *et al.* (2009) In vitro and in vivo protein-bound tyrosine nitration characterized by diagonal chromatography. *Mol Cell Proteomics*, **8**, 2642-2652.
17. Kanski, J., Hong, S.J. and Schoneich, C. (2005) Proteomic analysis of protein nitration in aging skeletal muscle and identification of nitrotyrosine-containing sequences in vivo by nanoelectrospray ionization tandem mass spectrometry. *J Biol Chem*, **280**, 24261-24266.
18. Miyagi, M., Sakaguchi, H., Darrow, R.M., Yan, L., West, K.A., Aulak, K.S., Stuehr, D.J., Hollyfield, J.G., Organisciak, D.T. and Crabb, J.W. (2002) Evidence that light modulates protein nitration in rat retina. *Mol Cell Proteomics*, **1**, 293-303.
19. Turko, I.V., Li, L., Aulak, K.S., Stuehr, D.J., Chang, J.Y. and Murad, F. (2003) Protein tyrosine nitration in the mitochondria from diabetic mouse heart. Implications to dysfunctional mitochondria in diabetes. *J Biol Chem*, **278**, 33972-33977.
20. Zhan, X., Du, Y., Crabb, J.S., Gu, X., Kern, T.S. and Crabb, J.W. (2008) Targets of tyrosine nitration in diabetic rat retina. *Mol Cell Proteomics*, **7**, 864-874.
21. Souza, J.M., Daikhin, E., Yudkoff, M., Raman, C.S. and Ischiropoulos, H. (1999) Factors determining the selectivity of protein tyrosine nitration. *Arch Biochem Biophys*, **371**, 169-178.
22. Elfering, S.L., Haynes, V.L., Traaseth, N.J., Ettl, A. and Giulivi, C. (2004) Aspects, mechanism, and biological relevance of mitochondrial protein nitration sustained by mitochondrial nitric oxide synthase. *Am J Physiol Heart Circ Physiol*, **286**, H22-29.
23. Xue, Y., Gao, X., Cao, J., Liu, Z., Jin, C., Wen, L., Yao, X. and Ren, J. (2010) A Summary of Computational Resources for Protein Phosphorylation. *Curr Protein Pept Sci*.
24. Xue, Y., Ren, J., Gao, X., Jin, C., Wen, L. and Yao, X. (2008) GPS 2.0, a tool to predict kinase-specific phosphorylation sites in hierarchy. *Mol Cell Proteomics*, **7**, 1598-1608.
25. Xue, Y., Liu, Z., Gao, X., Jin, C., Wen, L., Yao, X. and Ren, J. (2010) GPS-SNO: Computational Prediction of Protein S-Nitrosylation Sites with a Modified GPS Algorithm. *PLoS ONE*, **5**, e11290.

Release Note

1. Dec. 20th, 2009, the online service and the local stand-alone packages of GPS-YNO2 1.0 were released.
2. Aug. 14th, 2010, GPS-YNO2 was updated for the new sites and retraining.
3. Aug. 14th, 2014, GPS-YNO2 was updated for the curation of nitration sites.