GPS-SNO Manual

S-nitrosylation sites Prediction

Version 1.0
27/10/2009

Author: Yu Xue & Jian Ren

Contact: Dr. Yu Xue, xueyu@ustc.edu.cn; Dr. Jian Ren, renjian@ustc.edu.cn
The software is only free for academic research.
The latest version of GPS-SNO software is available from http://sno.biocuckoo.org

Index

STATEMENT **2**

INTRODUCTION **3**

DOWNLOAD & INSTALLATION **5**

PREDICTION OF S-NITROSYLATION SITES **7**

1. A SINGLE PROTEIN SEQUENCE IN FASTA FORMAT 7
2. MULTIPLE PROTEIN SEQUENCES IN FASTA FORMAT 11

REFERENCES **15**

RELEASE NOTE **16**

Statement

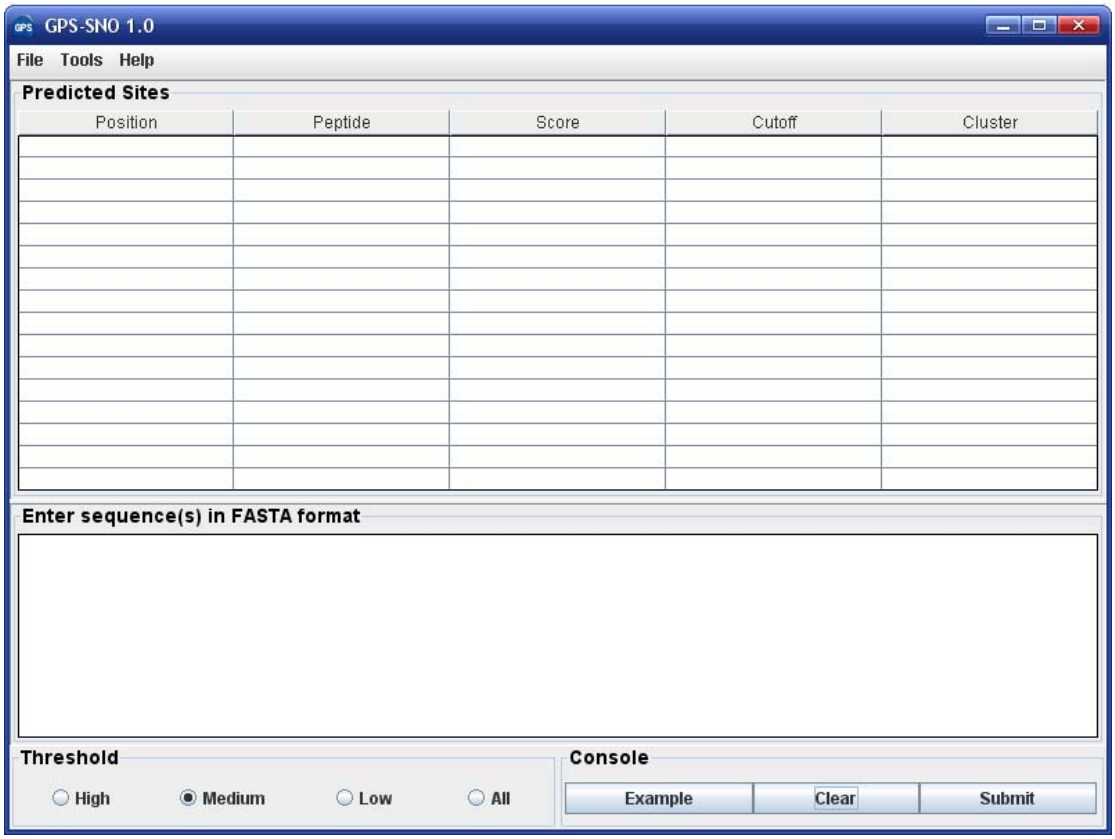
1. **Implementation**. The softwares of the CUCKOO Workgroup are implemented in JAVA (J2SE). Usually, both of online service and local stand-alone packages will be provided.
2. **Availability**. Our softwares are freely available for academic researches. For non-profit users, you can copy, distribute 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), National Natural Science Foundation of China (90919001, 30700138, 30900835, 30830036, 30721002, and 30871236), Chinese Academy of Sciences (INFO-115-C01-SDB4-36, KSCX2-YW-R-139), the China High Technology Research Program (2008ZX1002-020), and National Science Foundation for Post-doctoral Scientists (20080430100).

Introduction

The 1998 Nobel Prize in Physiology or Medicine was awarded for seminal discoveries that nitric oxide (NO) is a freely-diffusible signalling molecule and second messenger, regulates the production of cyclic GMP (cGMP), and plays essential roles in the cardiovascular system. Later, flood of studies challenged this fundamental view by observing that NO could spatially and temporally target specific cysteine thiols and transition metals of proteins, a reversibly post-translational modification (PTM) termed as S-nitrosylation [1-6]. Usually, NO synthases (NOSs) catalyze the reaction of arginines and O2 to produce citrullines and endogenous NO in most cell types (Figure 1). The NO could be further oxidated into NO2 then processed into N2O3 (Figure 1). By direct interactions or through scaffold and adaptor proteins, protein targets closely associated with NOSs might be in situ S-nitrosylated to form S-nitrosothiols (SNOs) (Figure 1) [1-4]. Although enzymatic mechanisms of protein S-nitrosylation were still elusive, several enzymes were proven to facilitate S-nitrosylation or de-nitrosylation reactions. For example, Cu, Zn superoxide dismutase (SOD) and thioredoxin (TRX) could promote S-nitrosylation, while protein disulfide isomerase (PDI) might regulate de-nitrosylation [3,4]. Current progresses proposed that S-nitrosylation could modulate proteins' stabilities [7], activities [8] and trafficking [9,10], and play important roles in a variety of biological processes, including transcriptional regulation [7], cell signalling [11], apoptosis [8], chromatin remodeling [12] and so on. Moreover, aberrant S-nitrosylation has been implicated in numerous diseases and cancers [1,2,8]. In this regard, experimental identification of S-nitrosylated proteins with their sites will be a foundation of understanding the molecular mechanisms and regulatory roles of S-nitrosylation.

In this work, we manually collected 467 experimentally verified S-nitrosylation sites in 302 unique proteins from scientific literature (Supplementary Table S1). Previously, we developed an algorithm of GPS 2.0 (Group-based Prediction System) for prediction of kinase-specific phosphorylation sites [9]. Here, we greatly improved the method and released GPS 3.0 algorithm. Then we developed a novel computational software of GPS-SNO 1.0 for prediction of S-nitrosylation sites. The leave-one-out validation and 4-, 6-, 8-, 10-fold cross-validations were calculated to evaluate the prediction performance and system robustness. By comparison, the performance of GPS 3.0 algorithm was much better than several other approaches, with an accuracy of 75.70%, a sensitivity of 55.32% and a specificity of 80.11% under the low threshold. As applications of GPS-SNO 1.0, we also collected 485 potentially S-nitrosylated substrates from PubMed (Supplementary Table S2). These proteins were detected from large-scale or small-scale studies, while the exact S-nitrosylation sites were still not experimentally determined. Successfully, we predicted 371 (~76%) of these targets with at lease one potential S-nitrosylation site. These prediction results might be a useful reservoir for further experimental verification. Finally, the online service and local packages of GPS-SNO 1.0 were implemented in Java SE and freely

available at: http://sno.biocuckoo.org/ .

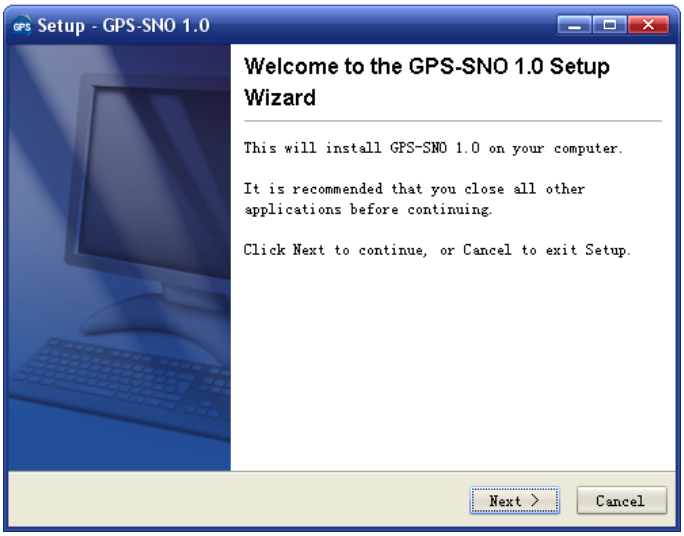


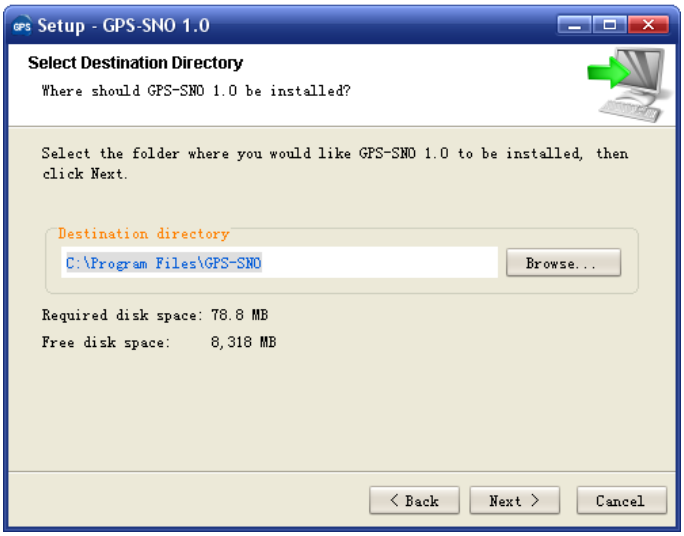
GPS-SNO 1.0 User Interface

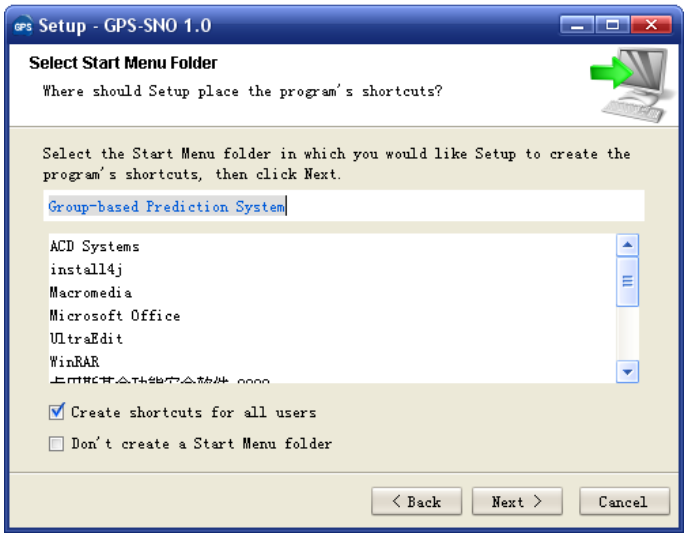
Download & Installation

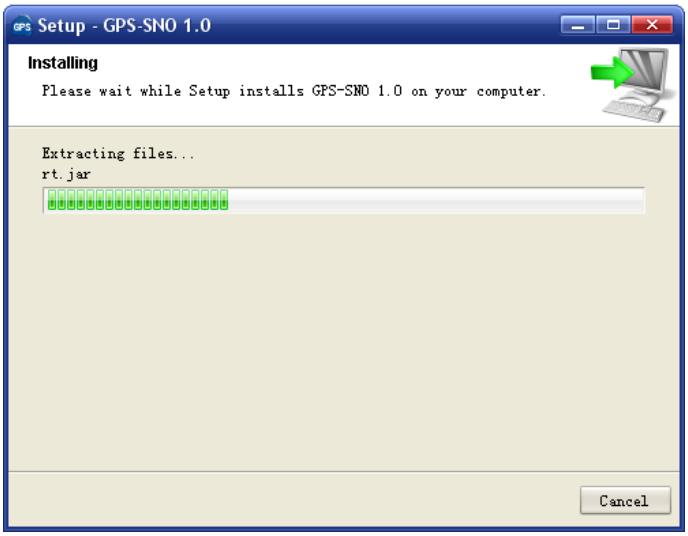
The GPS-SNO 1.0 was implemented in Java SE, and could support three major Operating Systems (OS), including Windows, Linux/Unix or Mac OS X systems. Both of online web service and local stand-alone packages are available from: http://sno.biocuckoo.org/prediction.php. We recommend that users could download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

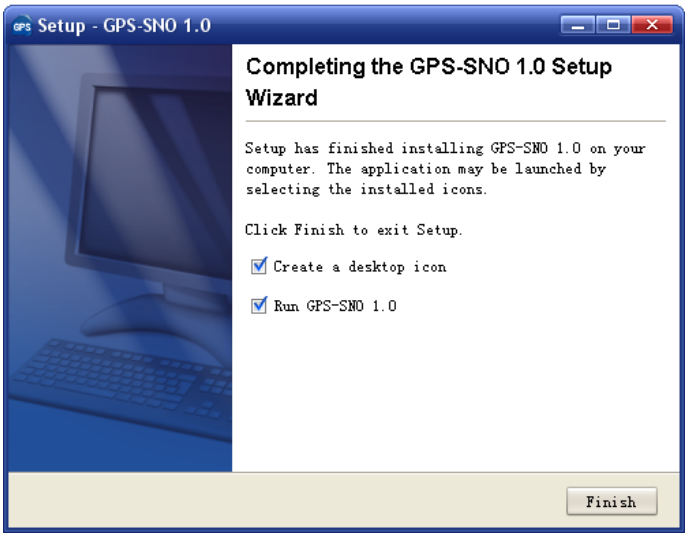
Please choose the proper package to download. After downloading, please double-click on the software package to begin installation, following the user prompts through the installation. And snapshots of the setup program for windows are shown below:











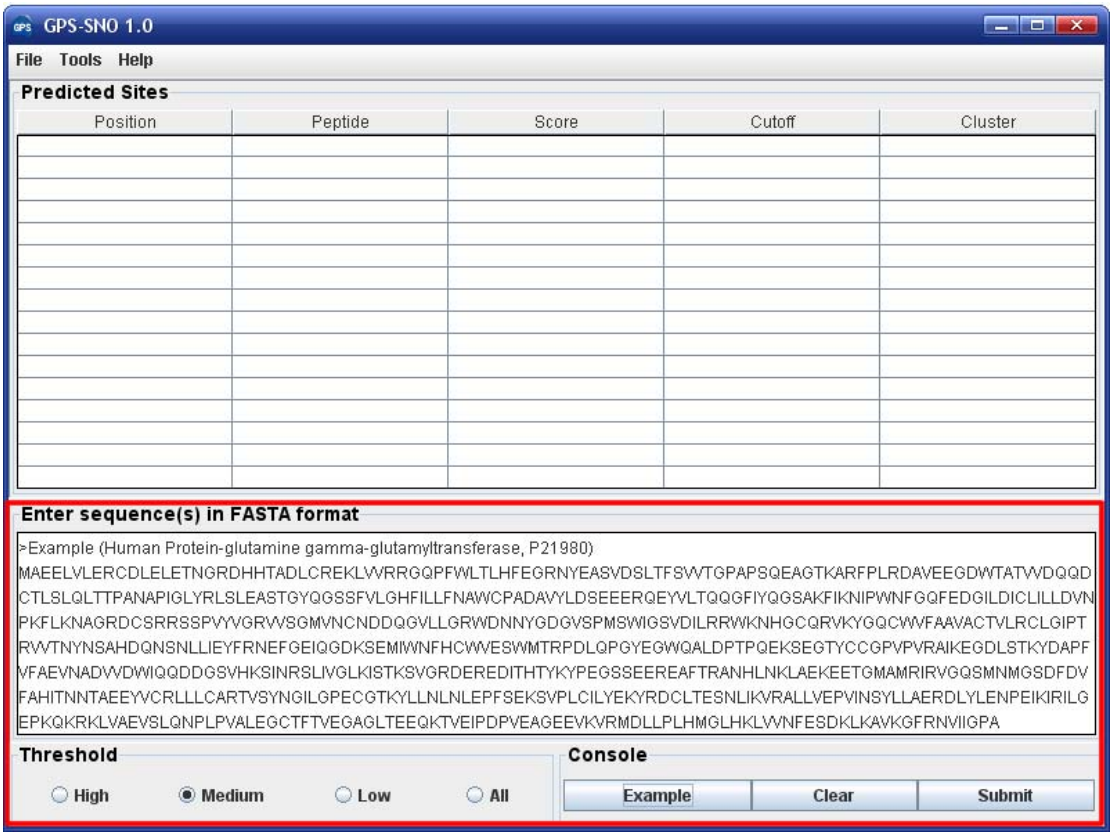
Finally, please click on the **Finish** button to complete the setup program.

Prediction of S-nitrosylation Sites

1. A single protein sequence in FASTA format

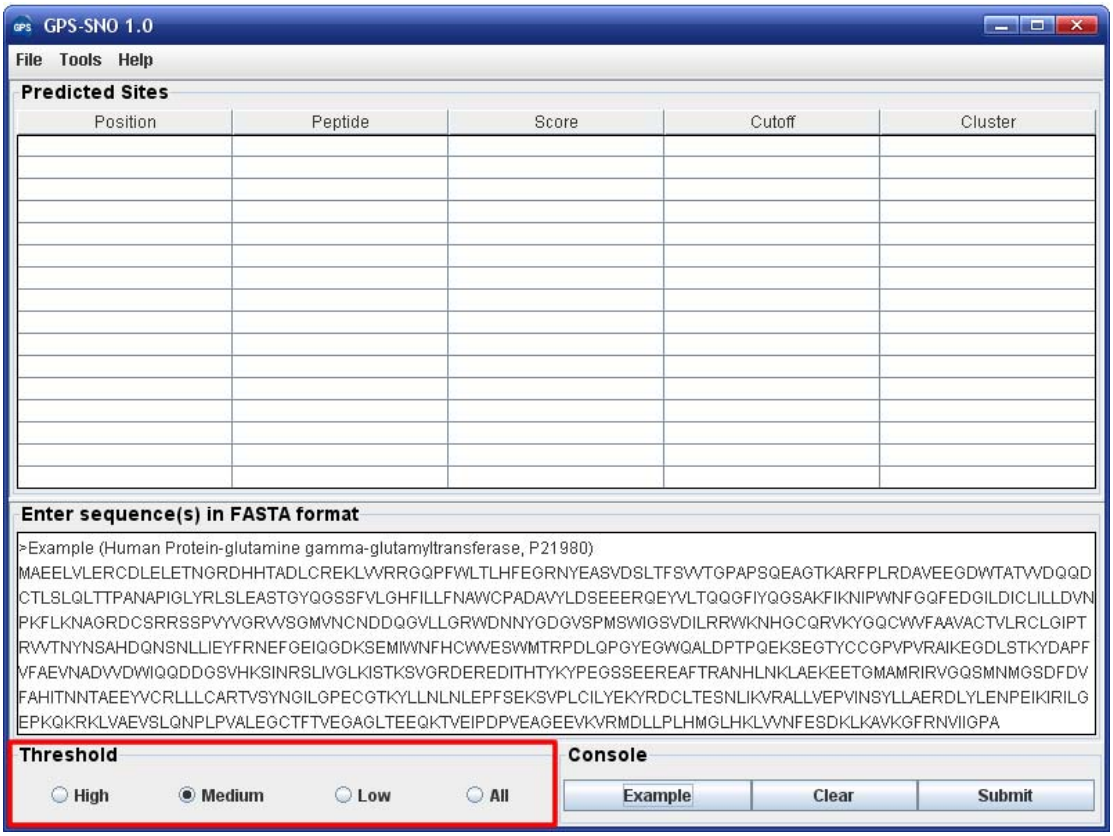
The following steps show you how to use the GPS-SNO 1.0 to predict S-nitrosylation sites for a single protein sequence in FASTA format.

(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-SNO 1.0

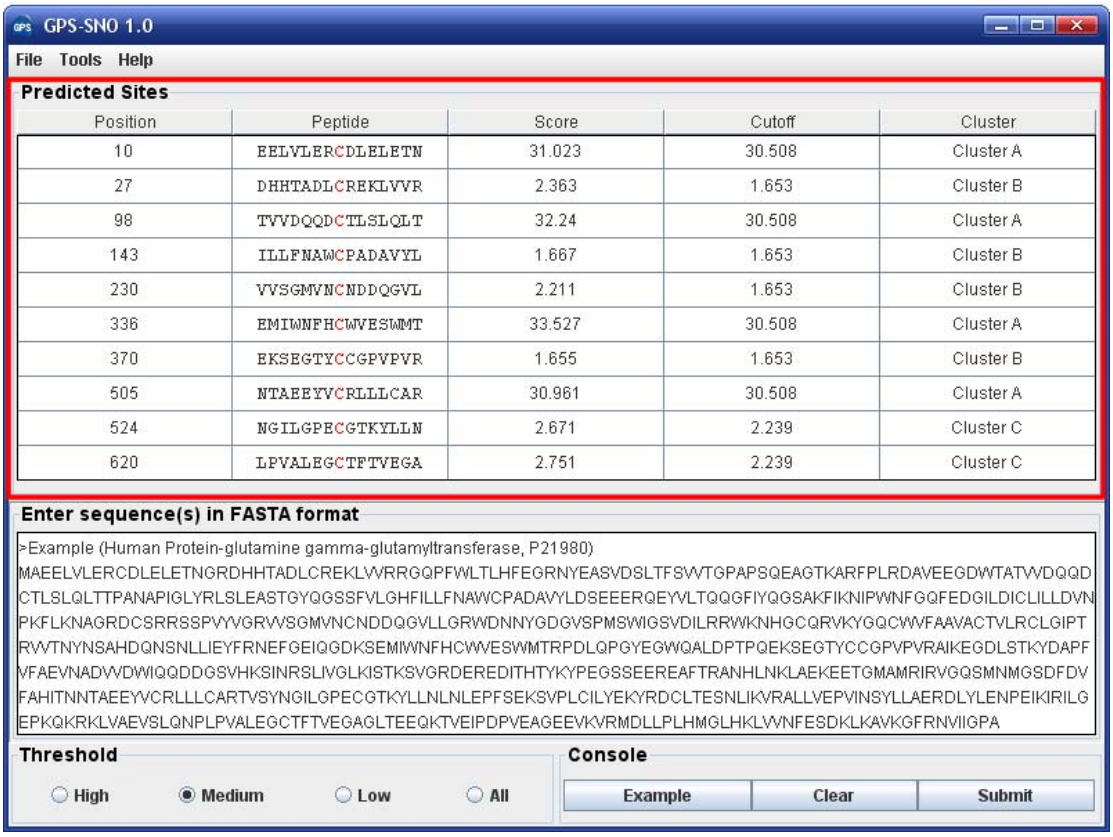


Note: for a single protein, the sequence without a name in raw format is also OK. However, for multiple sequences, the name of each protein should be presented.

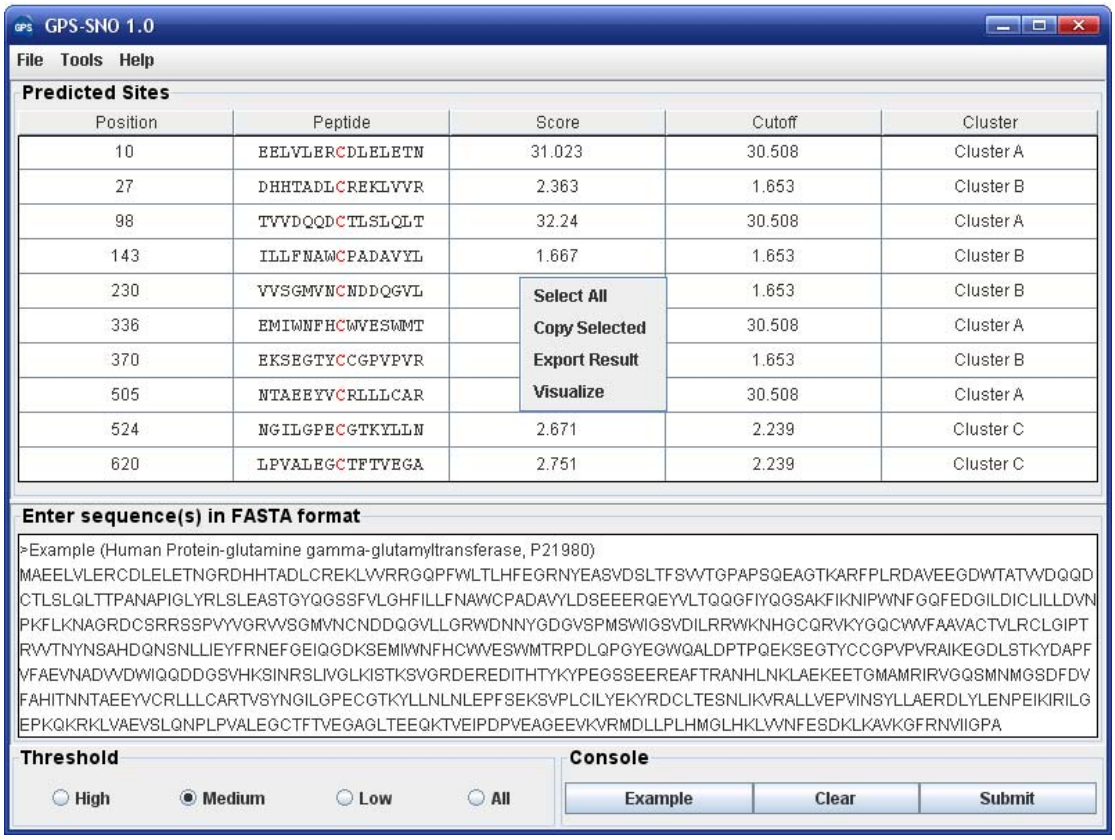
(2) Choose a **Threshold** that you need, the default cut-off is **Medium**.



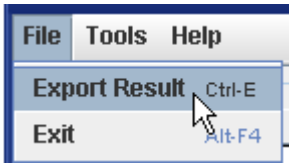
(3) Click on the **Submit** button, then the predicted S-nitrosylation sites will be shown.



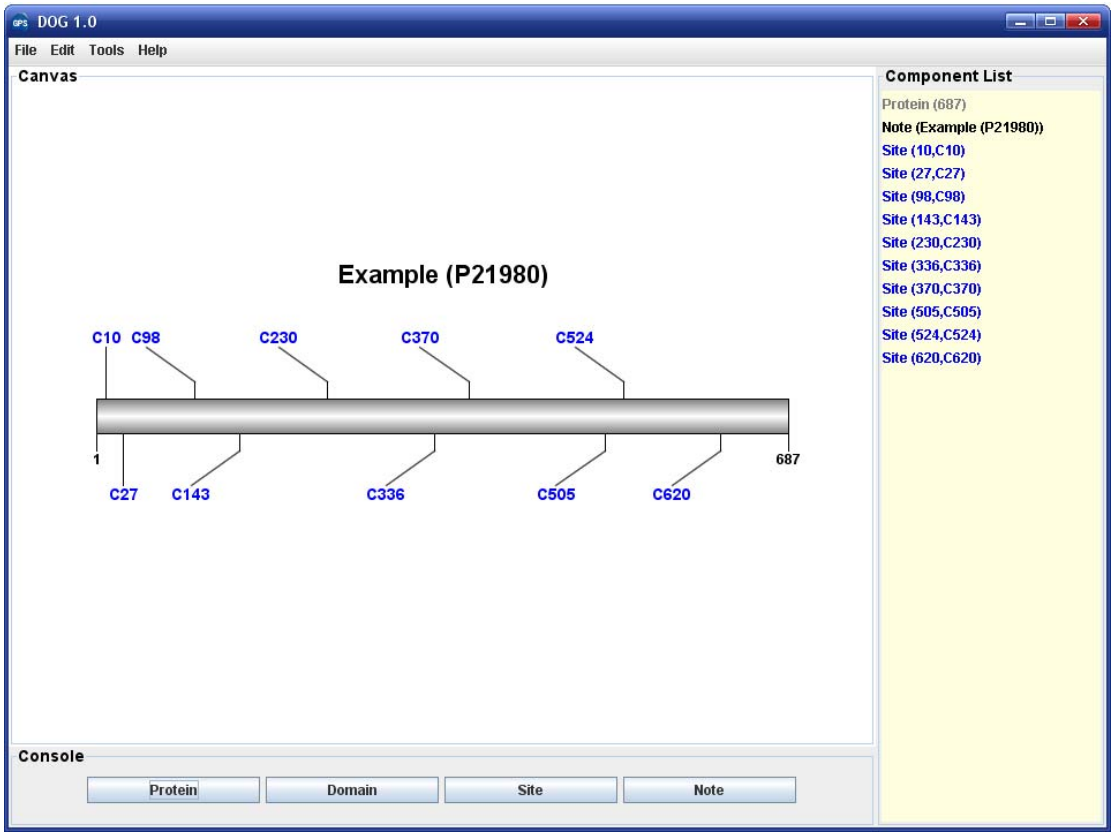
(4) Then please click on the **RIGHT** button in the prediction form. You can use the "**Select All**" and "**Copy Selected**" to copy the selected results into Clipboard. Then please copy the results into a file, e.g., an EXCEL file for further consideration. Also, you can choose "**Export Result**"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



Again, you can also click the "**Export Result**" 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.

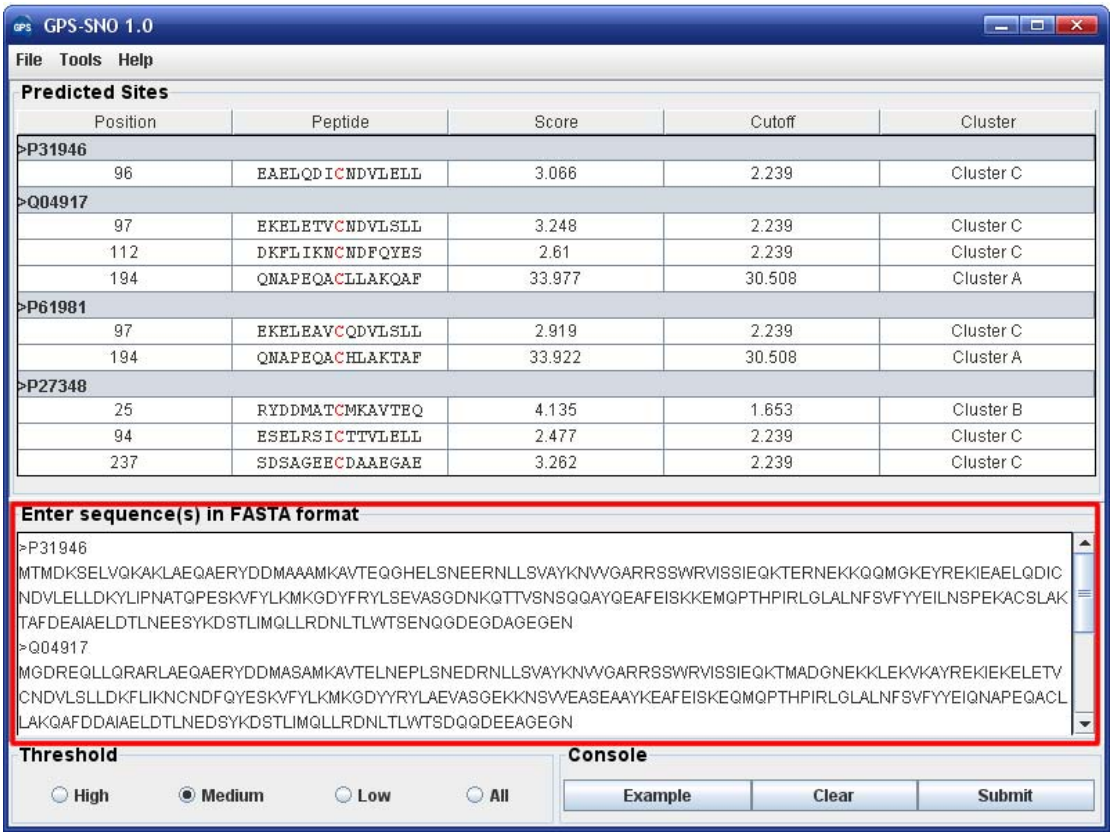


2. Multiple protein sequences in FASTA format

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

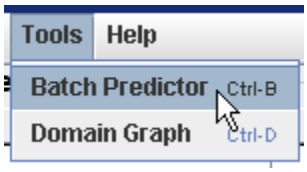
A. Input the sequences into text form directly. (Num. of Seq ⩽1,000)

If the number of total protein sequences is not greater than 1,000, 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-SNO 1.0 for prediction.



B. Use Batch Predictor tool.

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



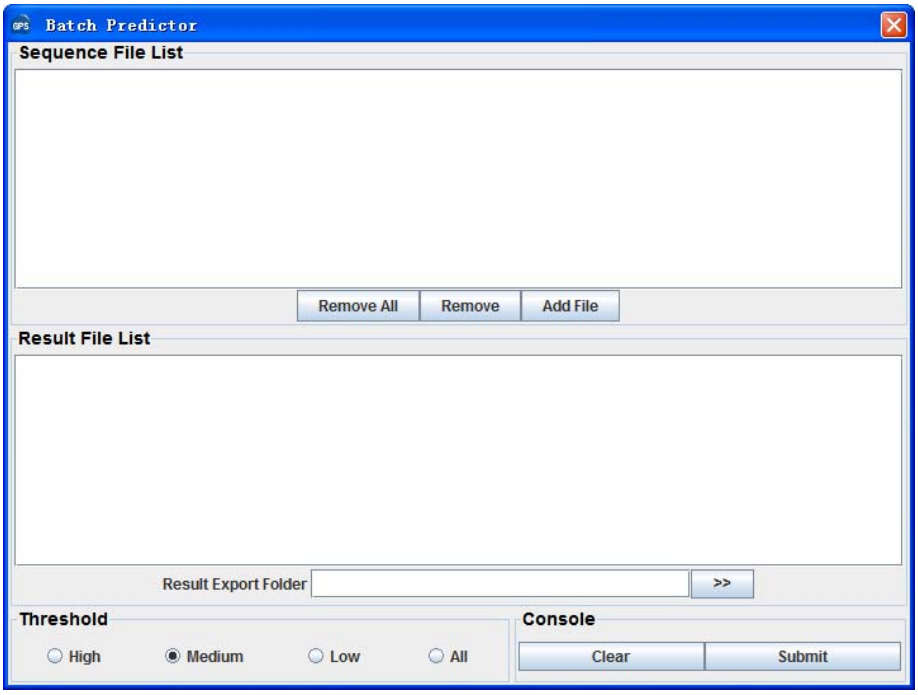
The following steps show you how to use it:

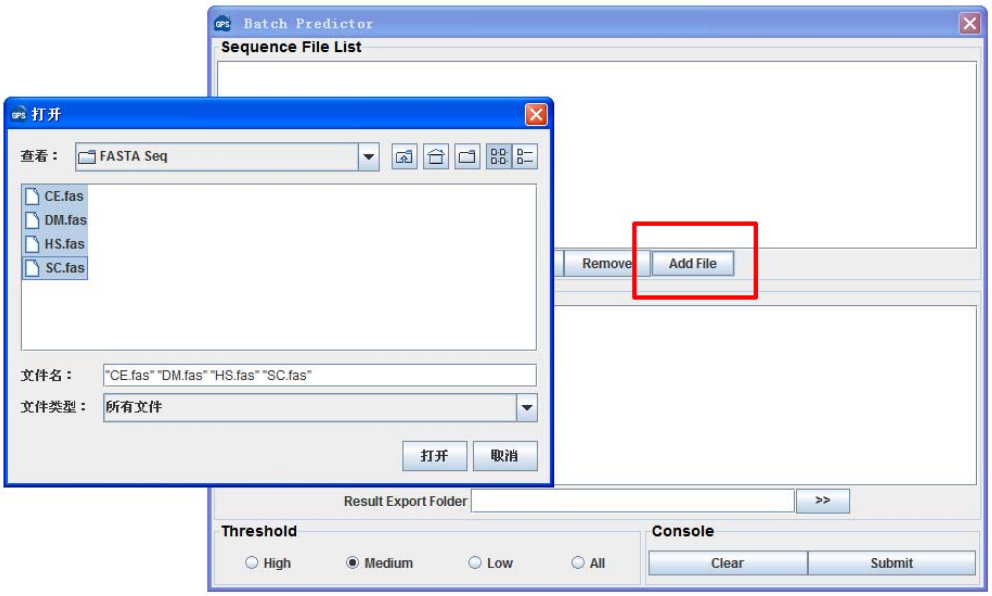
(1) Put protein sequences into one or several files (e.g., SC.fas, CE.fas, and etc) with FATSA format as below:

>protein1
XXXXXXXXXXXXX
XXXXXXXX
>protein2
XXXXXXXXXXXXXXXX…
>protein3
XXXXXXXXXXXX
...

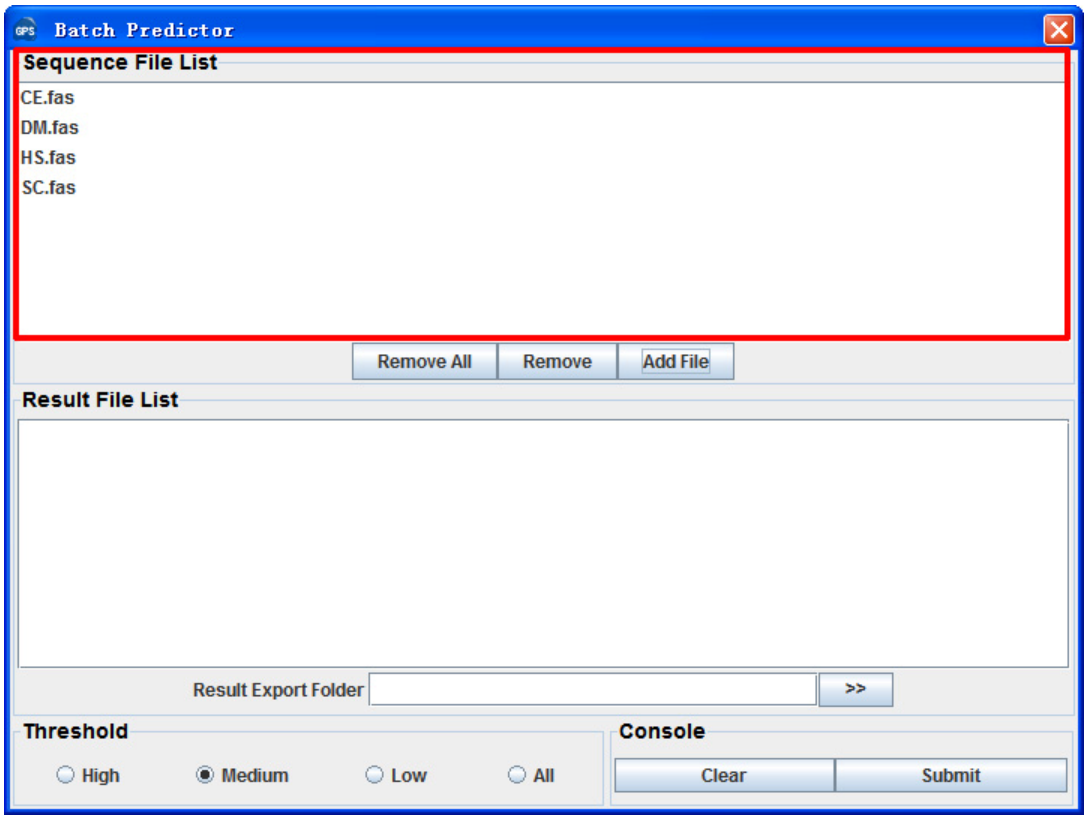
Most importantly, the name of each protein should be presented.

(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.

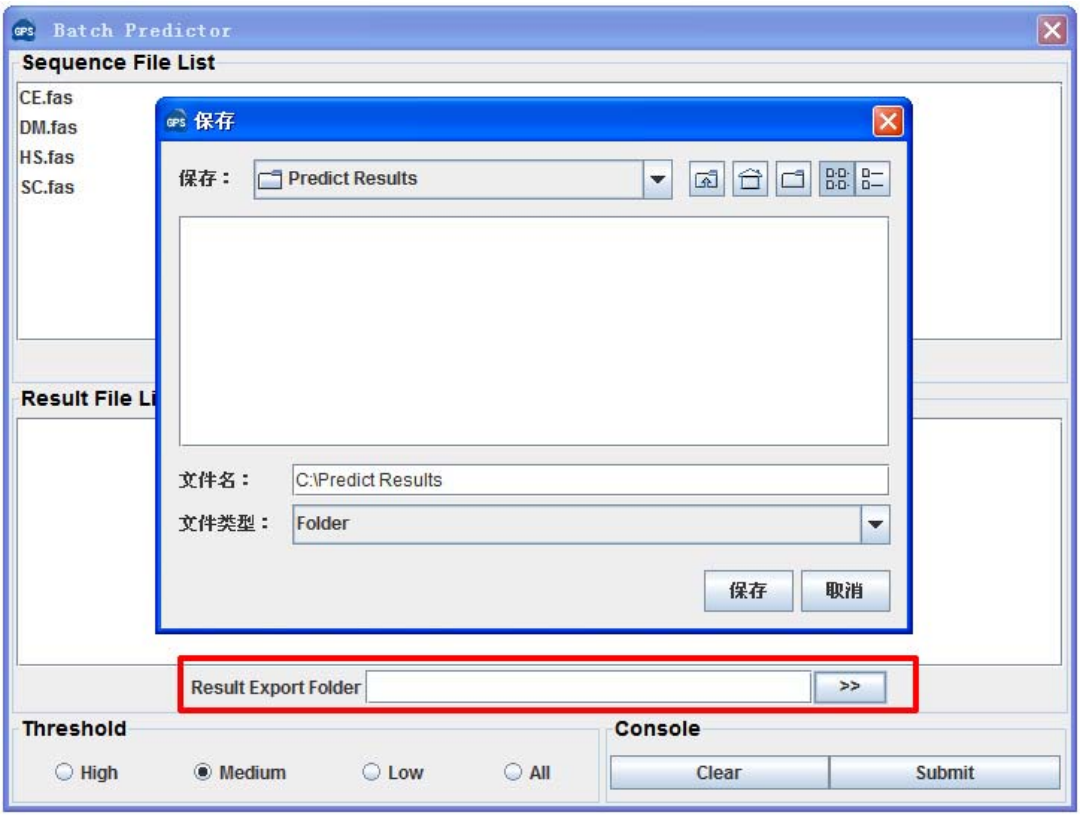




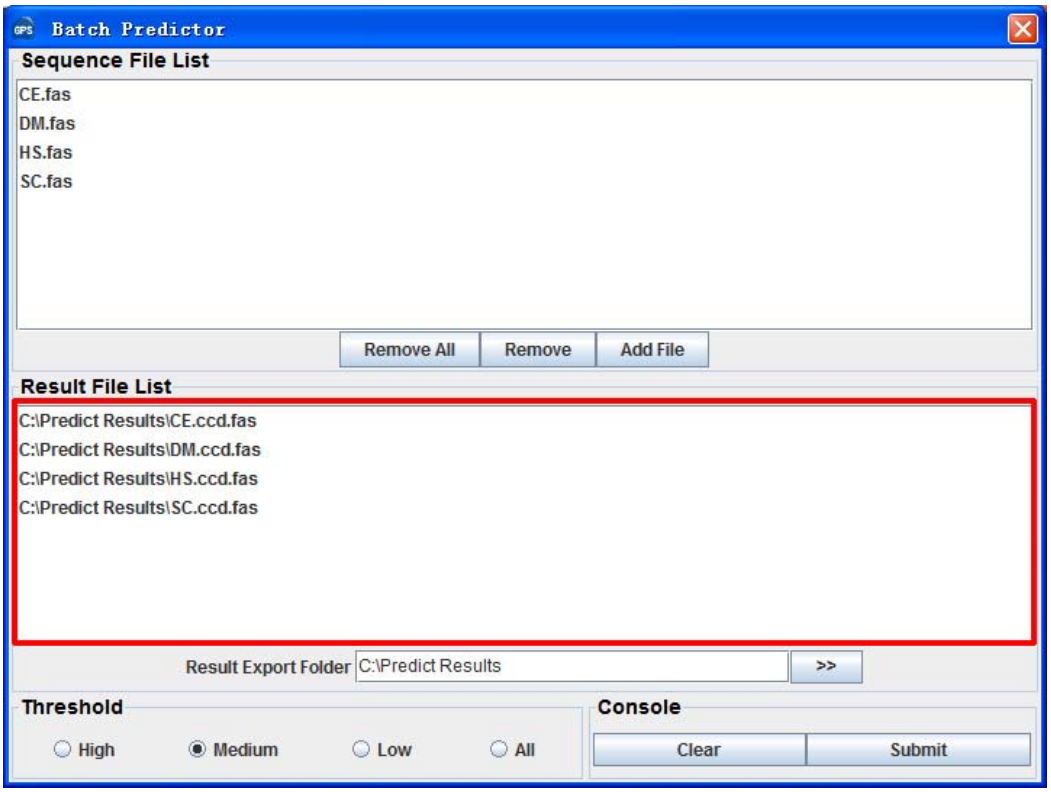
Then the names of added files will be shown in the **Sequence File List**.



(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 Folder,** and the name of result files will be shown in the **Prediction File List**.



References

[1] M.W. Foster, D.T. Hess, and J.S. Stamler, Protein S-nitrosylation in health and disease: a current perspective. Trends Mol Med 15 (2009) 391-404.

[2] M.W. Foster, T.J. McMahon, and J.S. Stamler, S-nitrosylation in health and disease. Trends Mol Med 9 (2003) 160-168.

[3] D.T. Hess, A. Matsumoto, S.O. Kim, H.E. Marshall, and J.S. Stamler, Protein S-nitrosylation: purview and parameters. Nat Rev Mol Cell Biol 6 (2005) 150-166.

[4] S.R. Tannenbaum, and F.M. White, Regulation and specificity of S-nitrosylation and denitrosylation. ACS Chem Biol 1 (2006) 615-618.

[5] D.T. Hess, A. Matsumoto, R. Nudelman, and J.S. Stamler, S-nitrosylation: spectrum and specificity. Nat Cell Biol 3 (2001) E46-49.

[6] J.S. Stamler, S. Lamas, and F.C. Fang, Nitrosylation. the prototypic redox-based signaling mechanism. Cell 106 (2001) 675-683.

[7] F. Li, P. Sonveaux, Z.N. Rabbani, S. Liu, B. Yan, Q. Huang, Z. Vujaskovic, M.W. Dewhirst, and C.Y. Li, Regulation of HIF-1alpha stability through S-nitrosylation. Mol Cell 26 (2007) 63-74.

[8] A.H. Tsang, Y.I. Lee, H.S. Ko, J.M. Savitt, O. Pletnikova, J.C. Troncoso, V.L. Dawson, T.M. Dawson, and K.K. Chung, S-nitrosylation of XIAP compromises neuronal survival in Parkinson's disease. Proc Natl Acad Sci U S A 106 (2009) 4900-4905.

[9] Y. Xue, J. Ren, X. Gao, C. Jin, L. Wen, and X. Yao, GPS 2.0, a tool to predict kinase-specific phosphorylation sites in hierarchy. Mol Cell Proteomics 7 (2008) 1598-1608.

Release Note

1. Oct. 27th, 2009, the online service and the local stand-alone packages of GPS-SNO 1.0 were released.