中国科学院计算技术研究所
INSTITUTE OF COMPUTING TECHNOLOGY, CHINESE ACADEMY OF SCIENCES
pFind Studio

pGlyco2 User Guide

pGlyco Team

pFind Lab

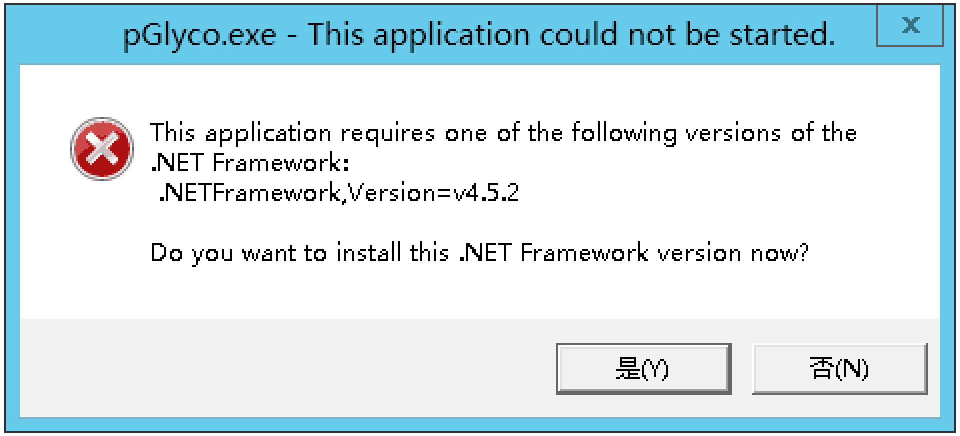
2018/02/01

Requirements

- Windows 7 or above
  - 64 bit version
  - .Net Framework 4.5.2

.Net Framework 4.5.2

- If .Net Framework 4.5.2 was not installed, message below will be shown when running pGlyco.



pGlyco.exe - This application could not be started.

This application requires one of the following versions of the .NET Framework:
.NETFramework,Version=v4.5.2

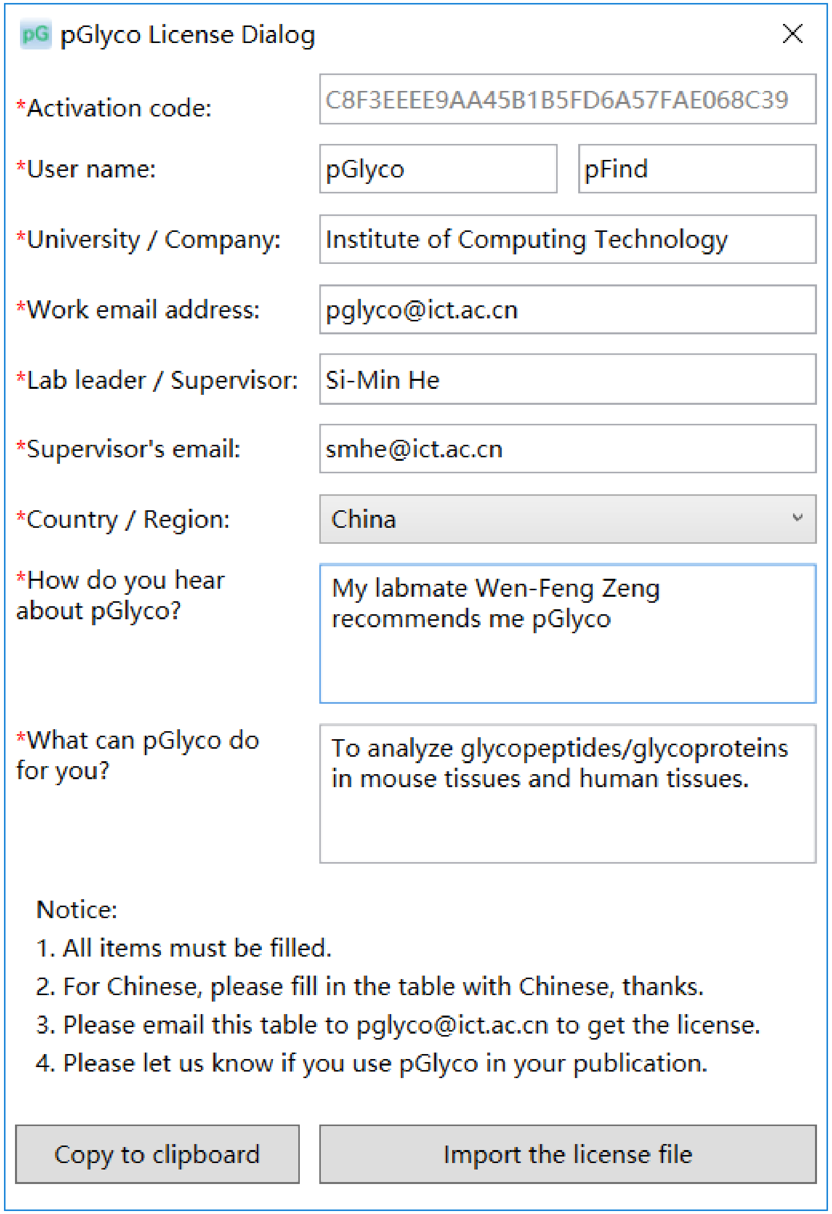
Do you want to install this .NET Framework version now?

- To download and install .Net Framework 4.5.2, please visit https://www.microsoft.com/en-us/download/details.aspx?id=42642

Catalog

- Activation
- Search
- Results
- Use gLabel to annotate GPSM

Activation

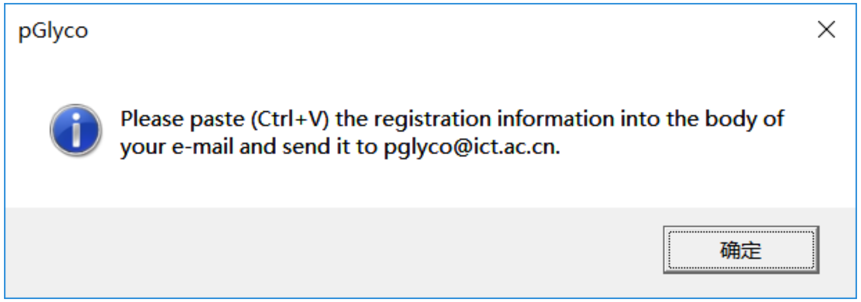


*Activation code:	C8F3EEEE9AA45B1B5FD6A57FAE068C39
*User name:	pGlyco pFind
*University / Company:	Institute of Computing Technology
*Work email address:	pglyco@ict.ac.cn
*Lab leader / Supervisor:	Si-Min He
*Supervisor's email:	smhe@ict.ac.cn
*Country / Region:	China
*How do you hear about pGlyco?	My labmate Wen-Feng Zeng recommends me pGlyco
*What can pGlyco do for you?	To analyze glycopeptides/glycoproteins in mouse tissues and human tissues.

Notice:

1. All items must be filled.
2. For Chinese, please fill in the table with Chinese, thanks.
3. Please email this table to pglyco@ict.ac.cn to get the license.
4. Please let us know if you use pGlyco in your publication.

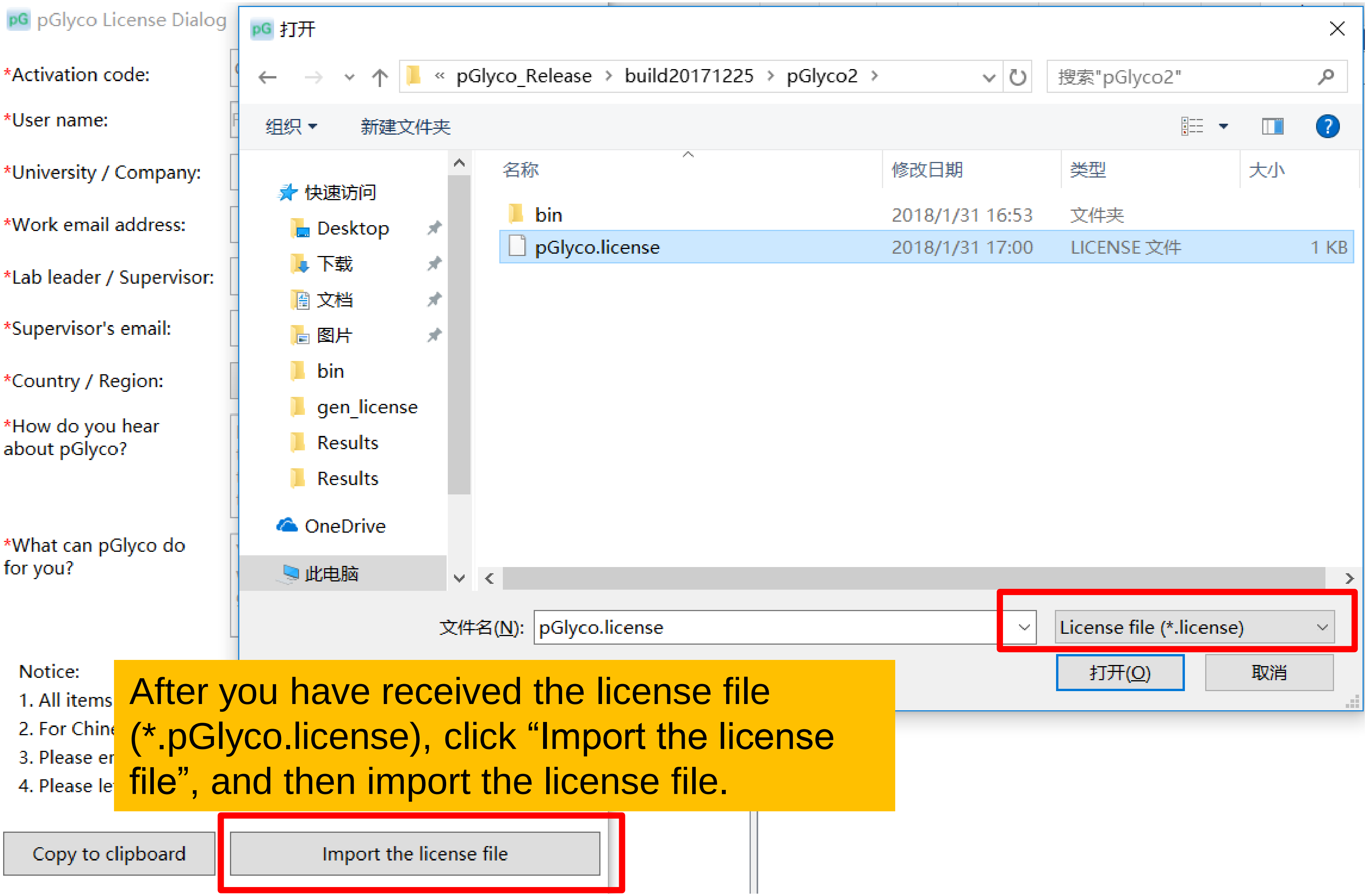
- If it is the first time you run pGlyco, activation dialog will show up. After filling all the blanks, click “Copy to clipboard”



Please paste (Ctrl+V) the registration information into the body of your e-mail and send it to pglyco@ict.ac.cn.

- After we receive the activation information, we will email the license file (a *.pGlyco.license file) to you via pglyco@ict.ac.cn

Activation



名称	修改日期	类型	大小
bin	2018/1/31 16:53	文件夹	
pGlyco.license	2018/1/31 17:00	LICENSE 文件	1 KB

文件名(N): pGlyco.license

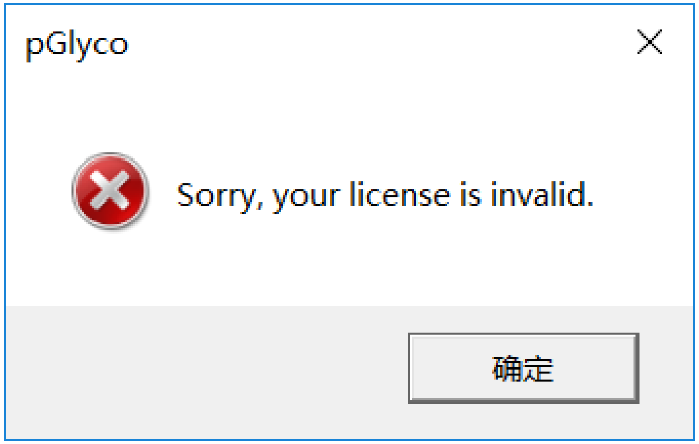
License file (*.license)

After you have received the license file (*.pGlyco.license), click “Import the license file”, and then import the license file.

Activation

* The activation code is **machine related**, if you did the above activation process in a machine, the activation code you received can only activate that machine.

If wrong license



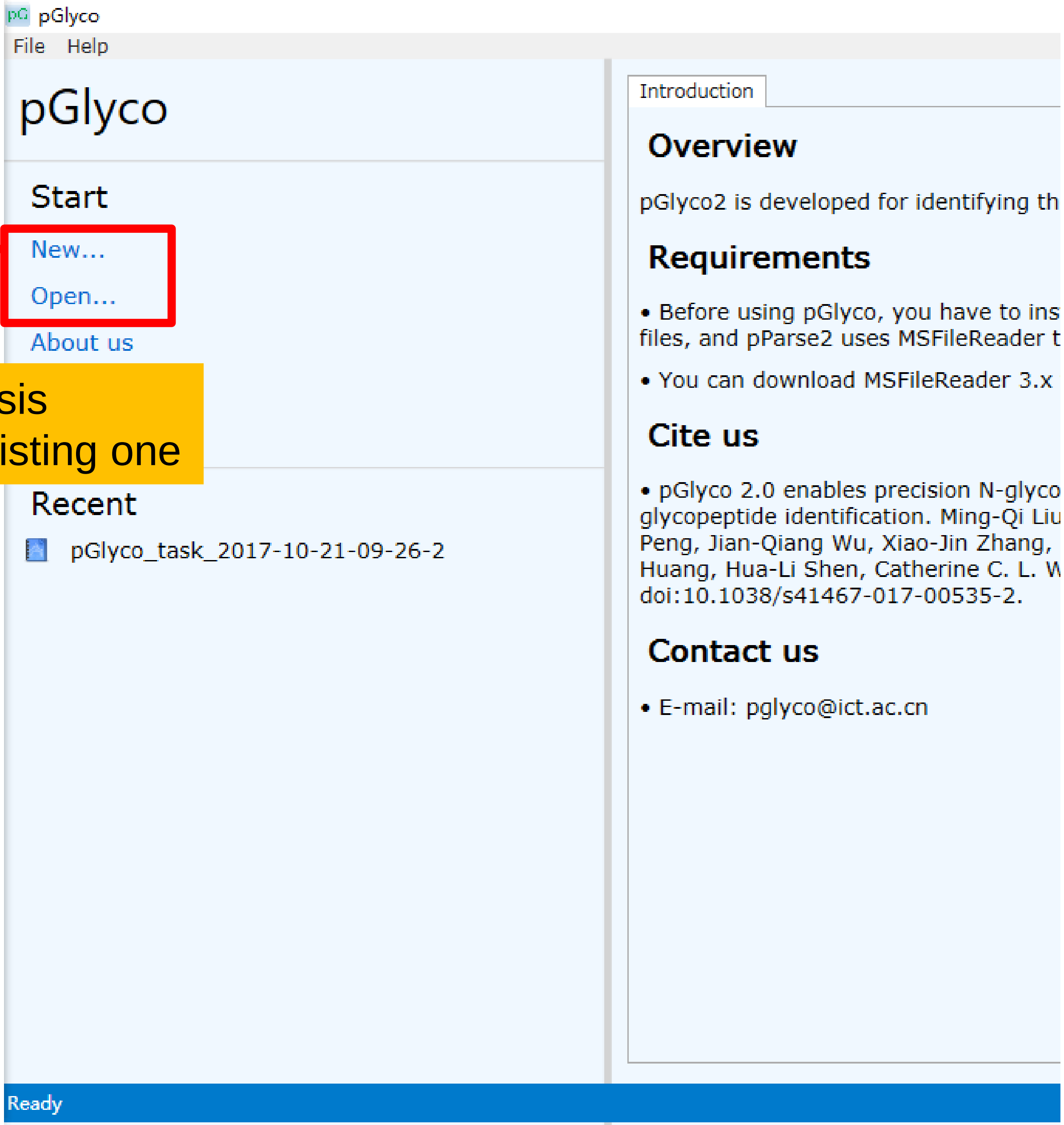
pGlyco

Sorry, your license is invalid.

Catalog

- **Activation**
- **Search**
- **Results**
- **Use gLabel to annotate GPSM**

Search



Overview

pGlyco2 is developed for identifying th

Requirements

- Before using pGlyco, you have to ins files, and pParse2 uses MSFileReader t
- You can download MSFileReader 3.x

Cite us

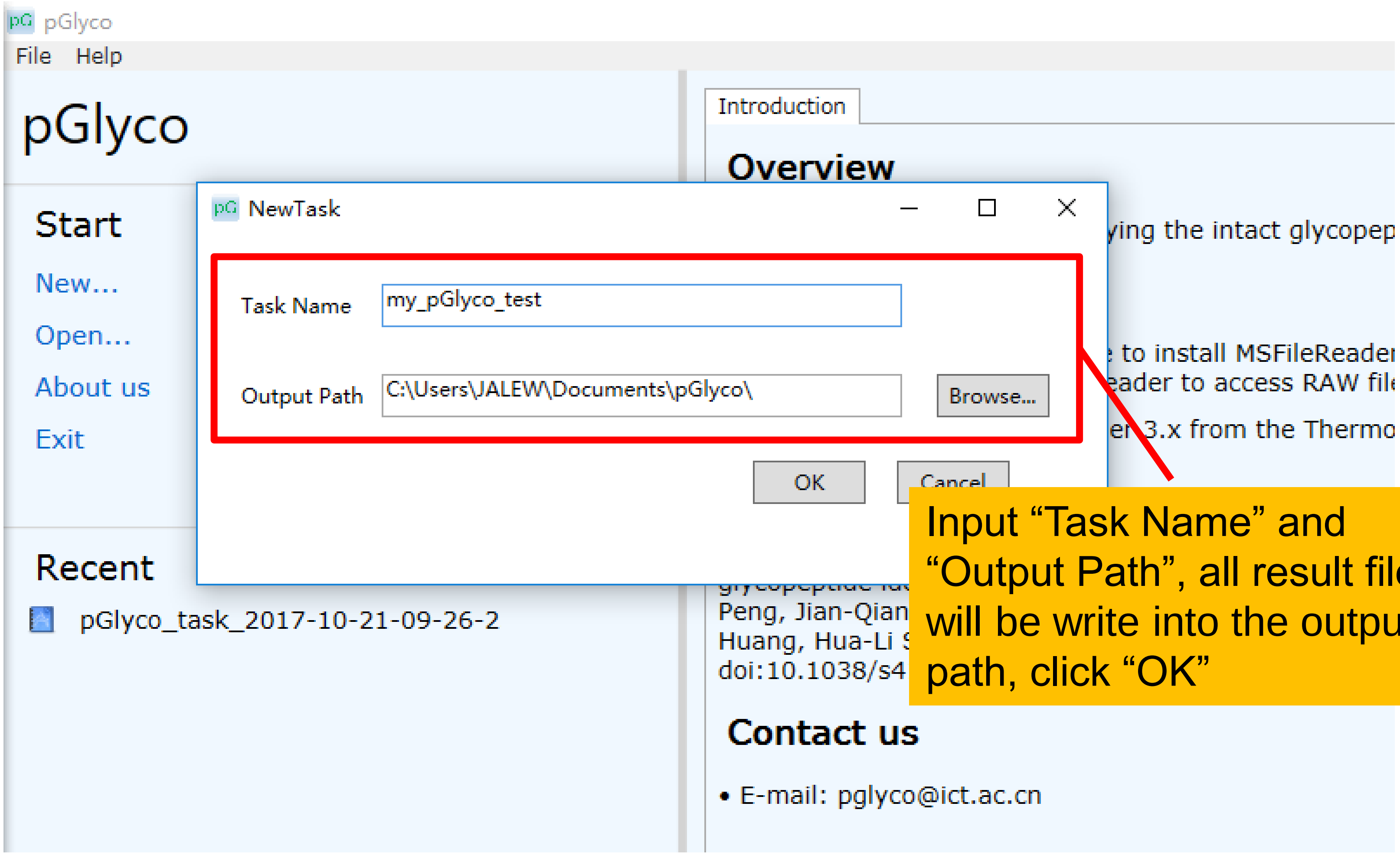
- pGlyco 2.0 enables precision N-glyco glycopeptide identification. Ming-Qi Liu Peng, Jian-Qiang Wu, Xiao-Jin Zhang, Huang, Hua-Li Shen, Catherine C. L. W doi:10.1038/s41467-017-00535-2.

Contact us

- E-mail: pglyco@ict.ac.cn

New an analysis
or open an existing one

Search

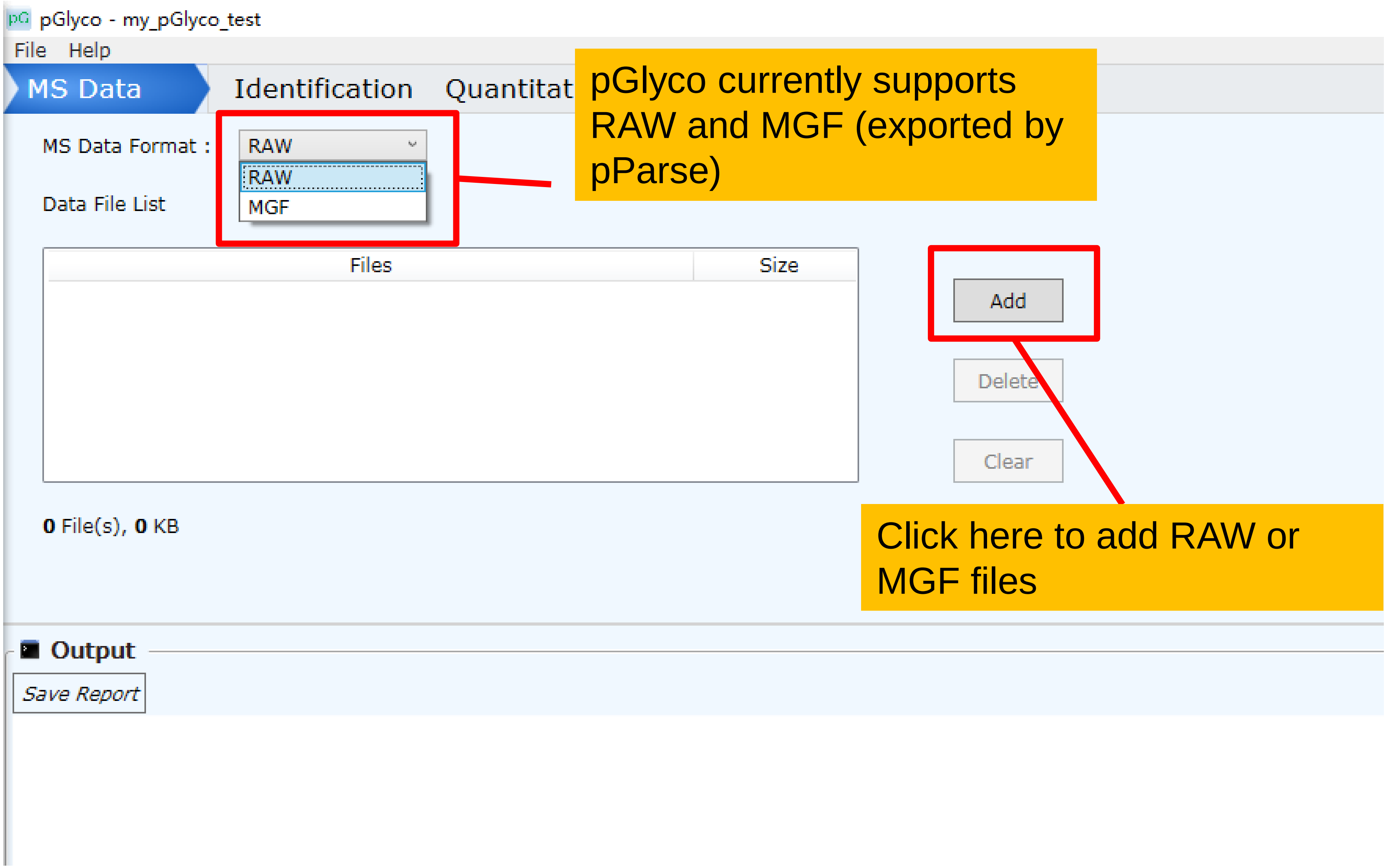


Task Name: my_pGlyco_test

Output Path: C:\Users\JALEW\Documents\pGlyco\

Input “Task Name” and “Output Path”, all result files will be write into the output path, click “OK”

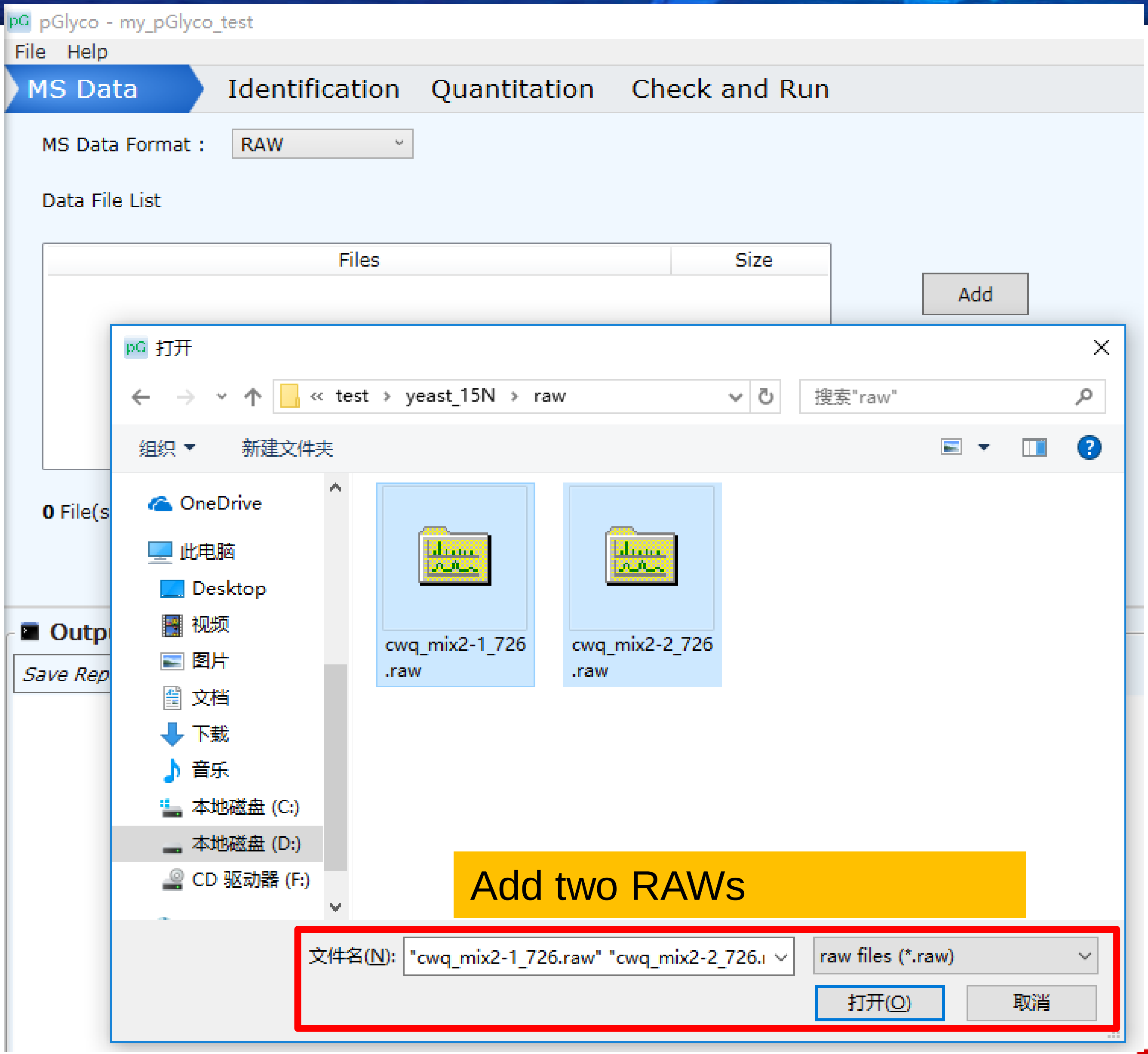
Search



pGlyco currently supports RAW and MGF (exported by pParse)

Click here to add RAW or MGF files

Search



pGlyco - my_pGlyco_test

MS Data Format : RAW

Data File List

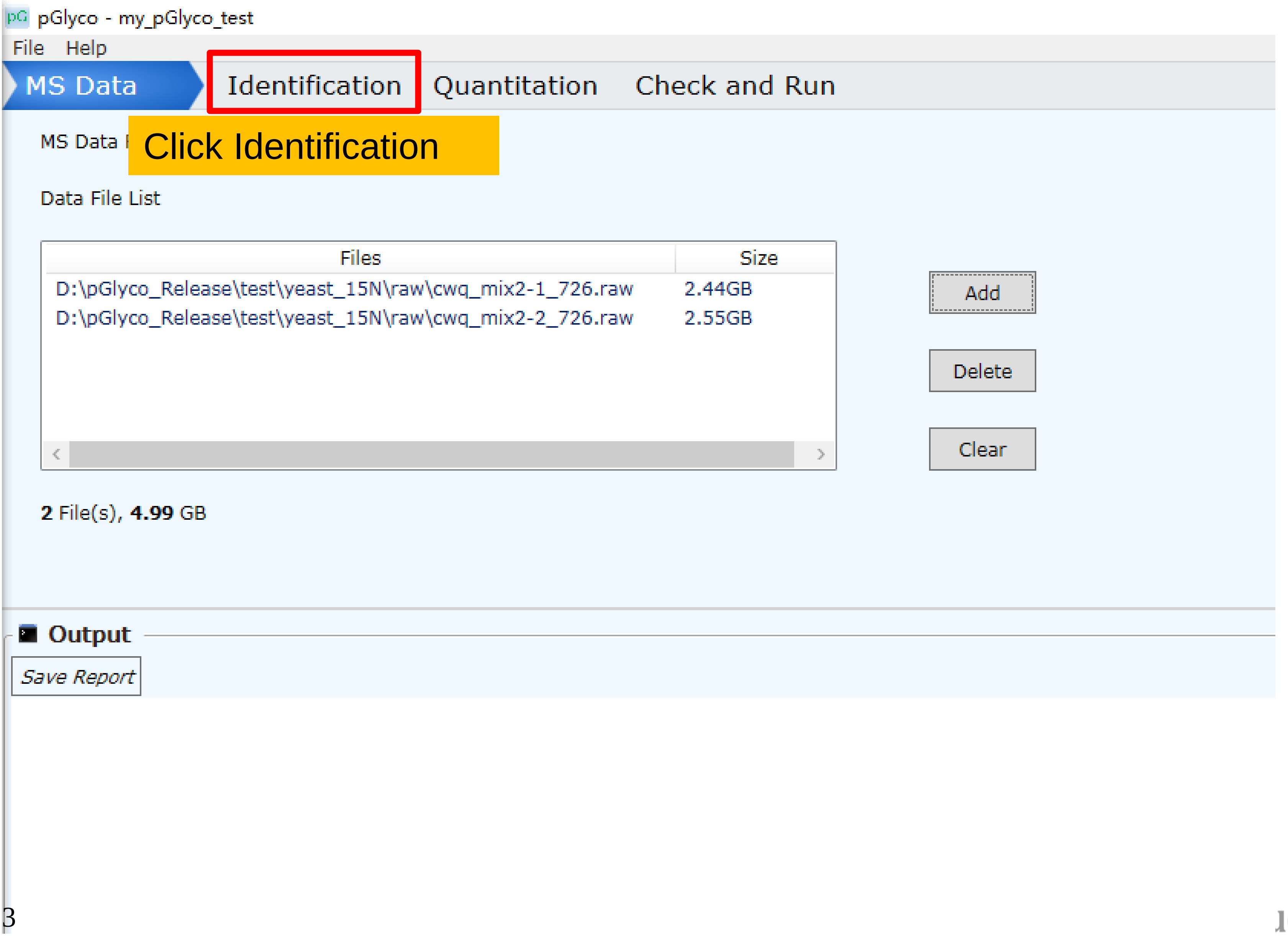
0 File(s

文件名(N): "cwq_mix2-1_726.raw" "cwq_mix2-2_726.r

raw files (*.raw)

Add two RAWs

Search

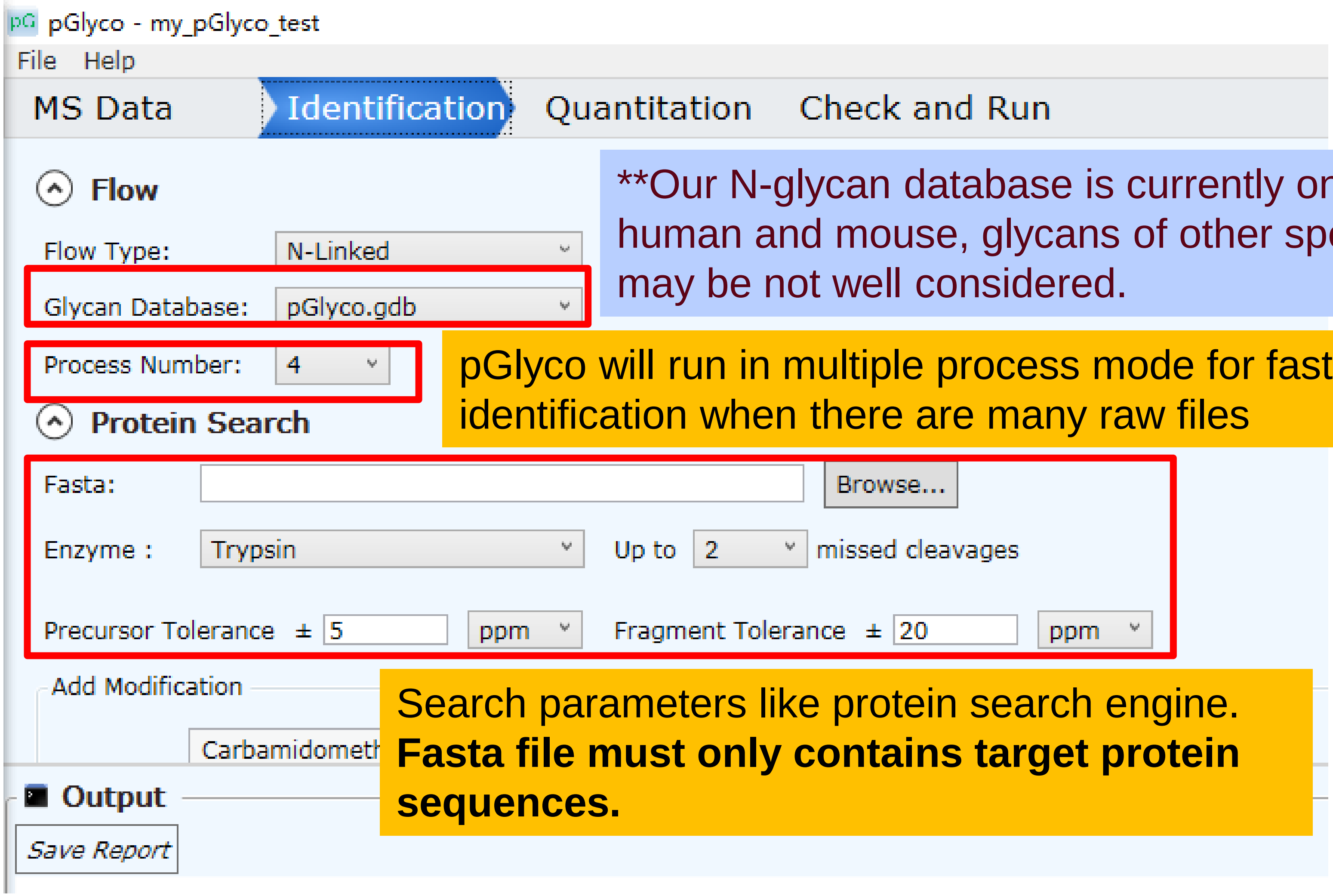


Click Identification

Files	Size
D:\pGlyco_Release\test\yeast_15N\raw\cwq_mix2-1_726.raw	2.44GB
D:\pGlyco_Release\test\yeast_15N\raw\cwq_mix2-2_726.raw	2.55GB

2 File(s), **4.99** GB

Search



Flow Type: N-Linked

Glycan Database: pGlyco.gdb

Process Number: 4

Fasta:

Enzyme : Trypsin Up to 2 missed cleavages

Precursor Tolerance ± 5 ppm Fragment Tolerance ± 20 ppm

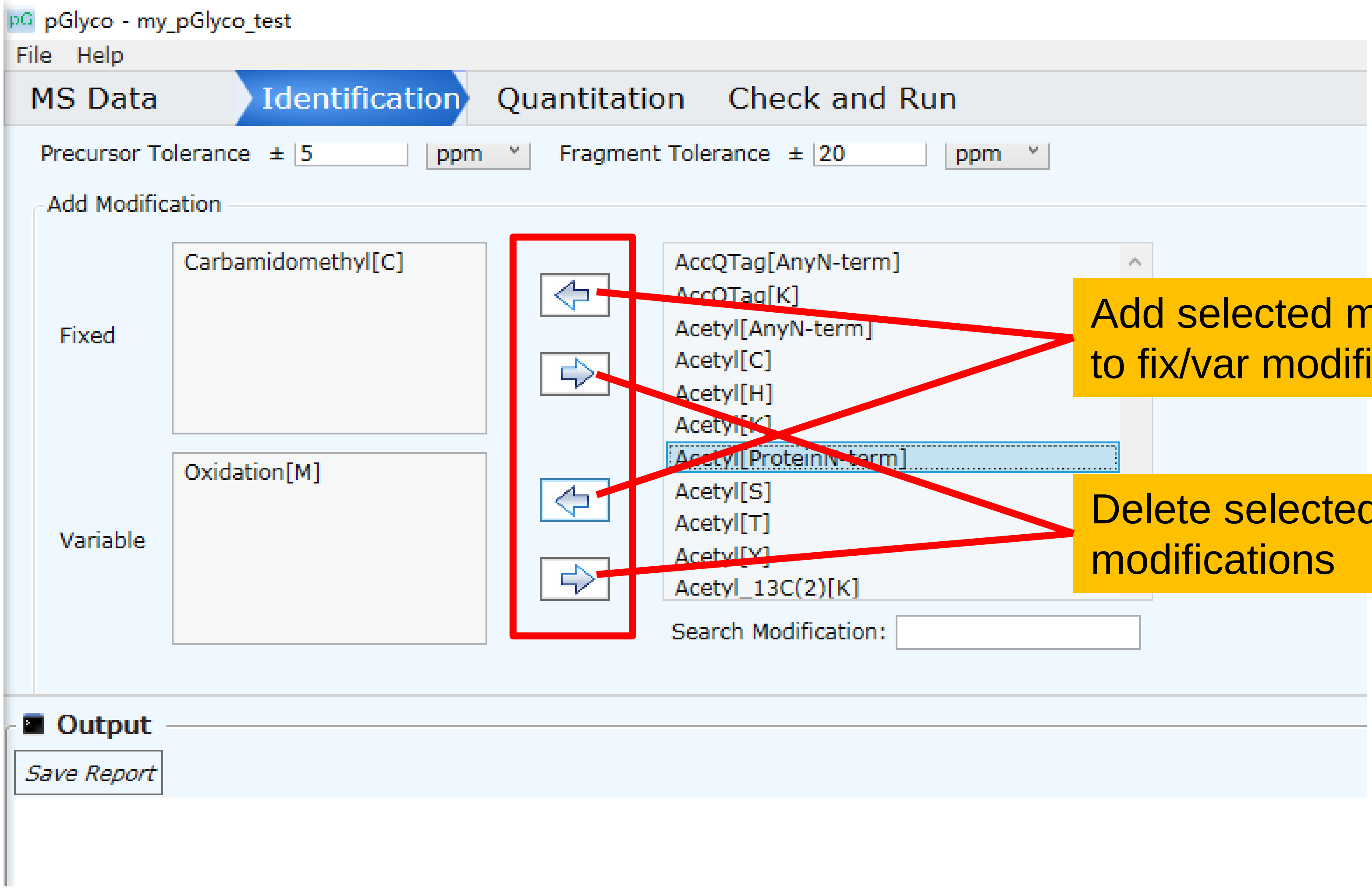
Carbamidometh

**Our N-glycan database is currently only for human and mouse, glycans of other species may be not well considered.

pGlyco will run in multiple process mode for fast identification when there are many raw files

Search parameters like protein search engine. **Fasta file must only contains target protein sequences.**

Search



Precursor Tolerance ± 5 ppm

Fragment Tolerance ± 20 ppm

Add Modification

Fixed: Carbamidomethyl[C]

Variable: Oxidation[M]

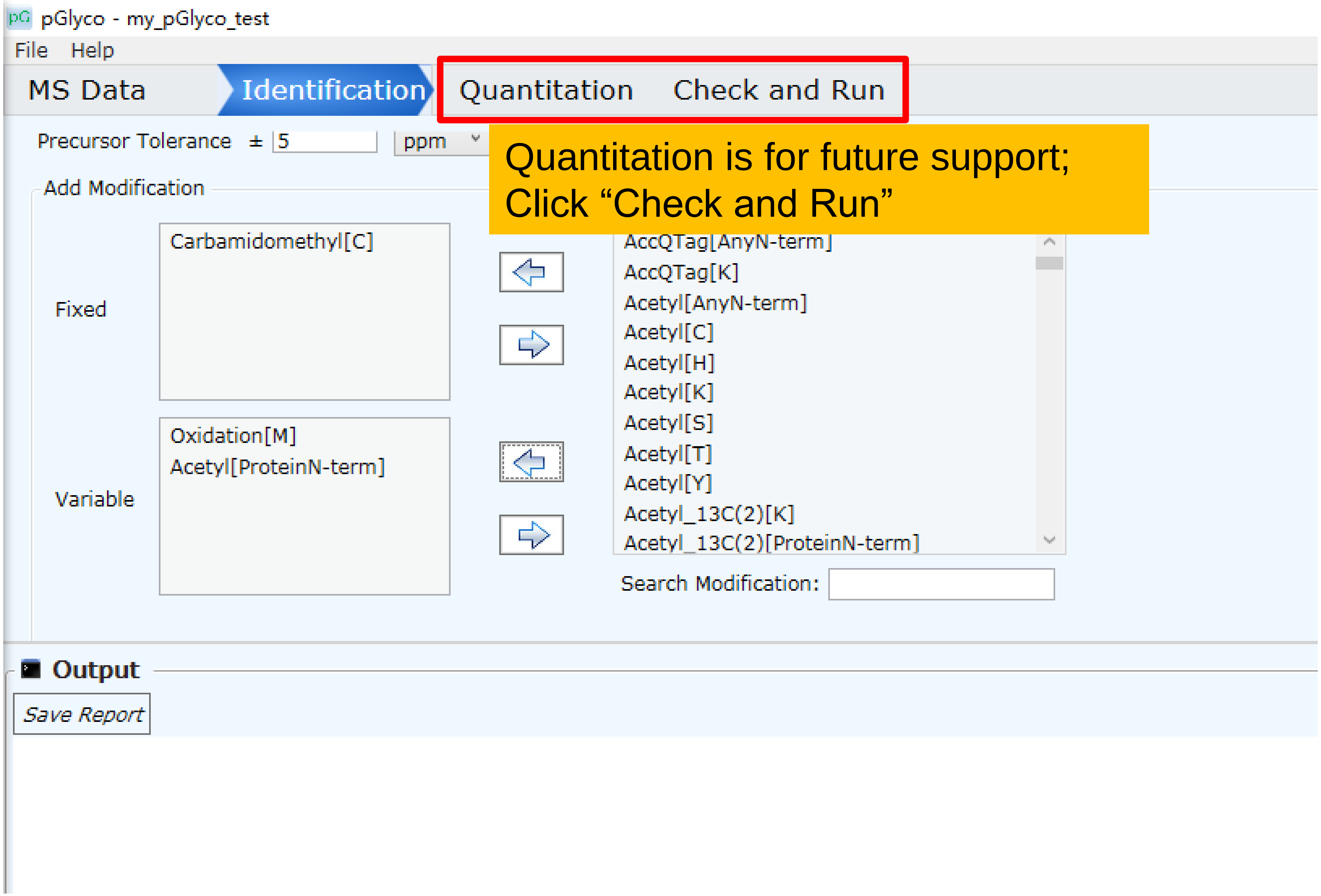
AccQTag[AnyN-term]
AccQTag[K]
Acetyl[AnyN-term]
Acetyl[C]
Acetyl[H]
Acetyl[K]
Acetyl[ProteinN-term]
Acetyl[S]
Acetyl[T]
Acetyl[Y]
Acetyl_13C(2)[K]

Search Modification:

Add selected modification to fix/var modifications

Delete selected fix/var modifications

Search



Precursor Tolerance ± 5 ppm

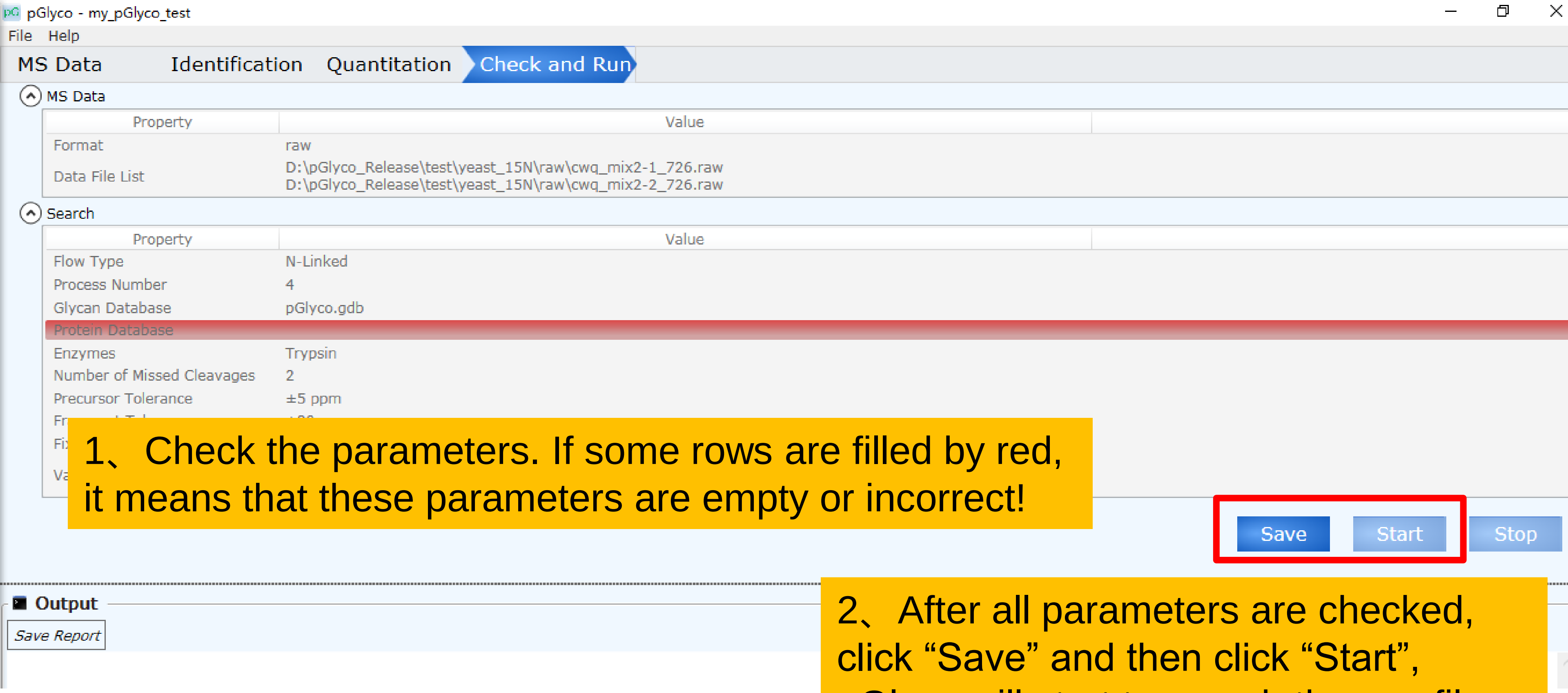
Add Modification

Fixed: Carbamidomethyl[C]

Variable: Oxidation[M], Acetyl[ProteinN-term]

Quantitation is for future support;
Click “Check and Run”

Search



MS Data

Property	Value
Format	raw
Data File List	D:\pGlyco_Release\test\yeast_15N\raw\cwq_mix2-1_726.raw D:\pGlyco_Release\test\yeast_15N\raw\cwq_mix2-2_726.raw

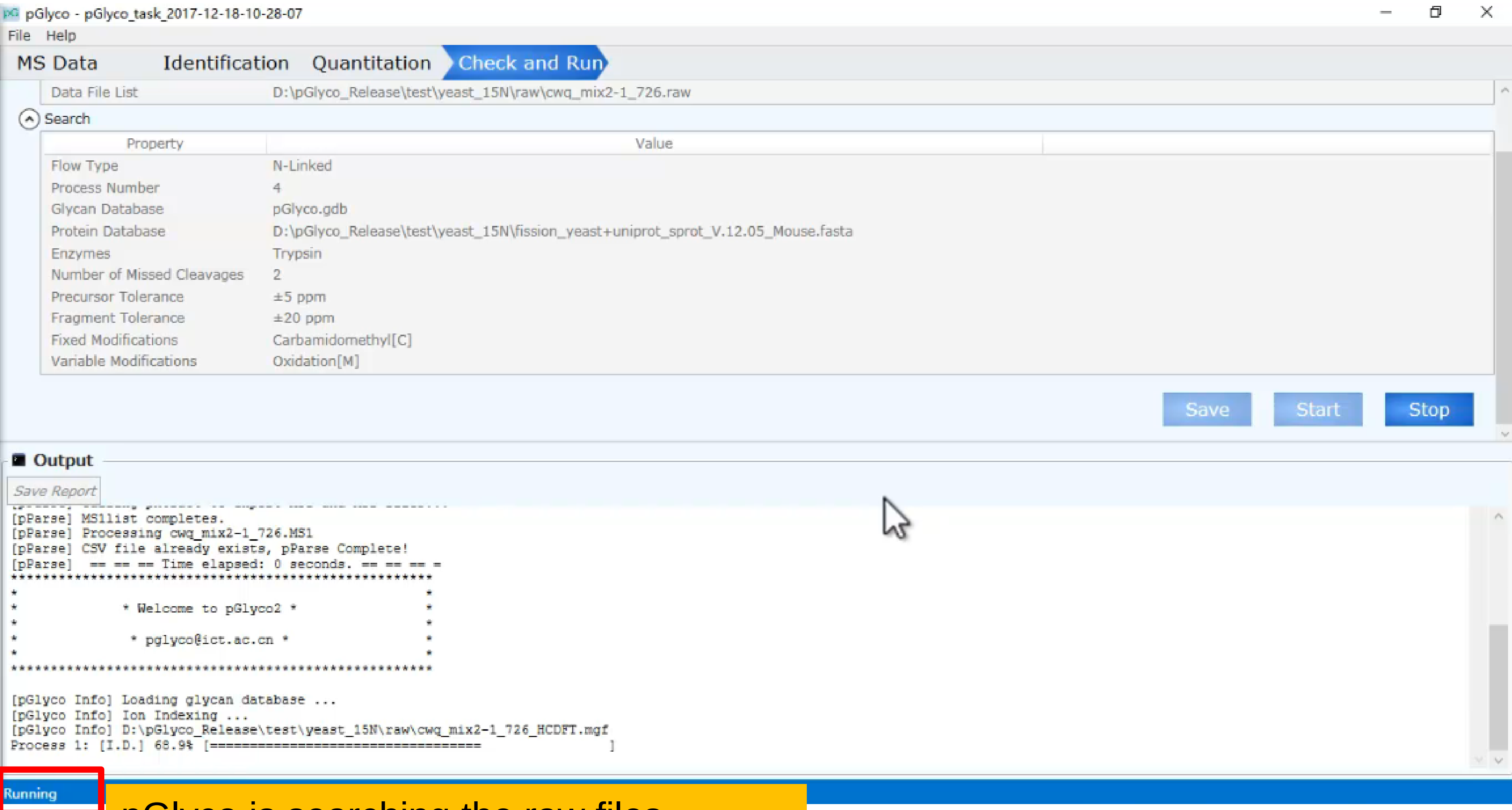
Search

Property	Value
Flow Type	N-Linked
Process Number	4
Glycan Database	pGlyco.gdb
Protein Database	
Enzymes	Trypsin
Number of Missed Cleavages	2
Precursor Tolerance	±5 ppm

1、 Check the parameters. If some rows are filled by red, it means that these parameters are empty or incorrect!

2、 After all parameters are checked, click “Save” and then click “Start”, pGlyco will start to search the raw files

Search



Data File List D:\pGlyco_Release\test\yeast_15N\raw\cwq_mix2-1_726.raw

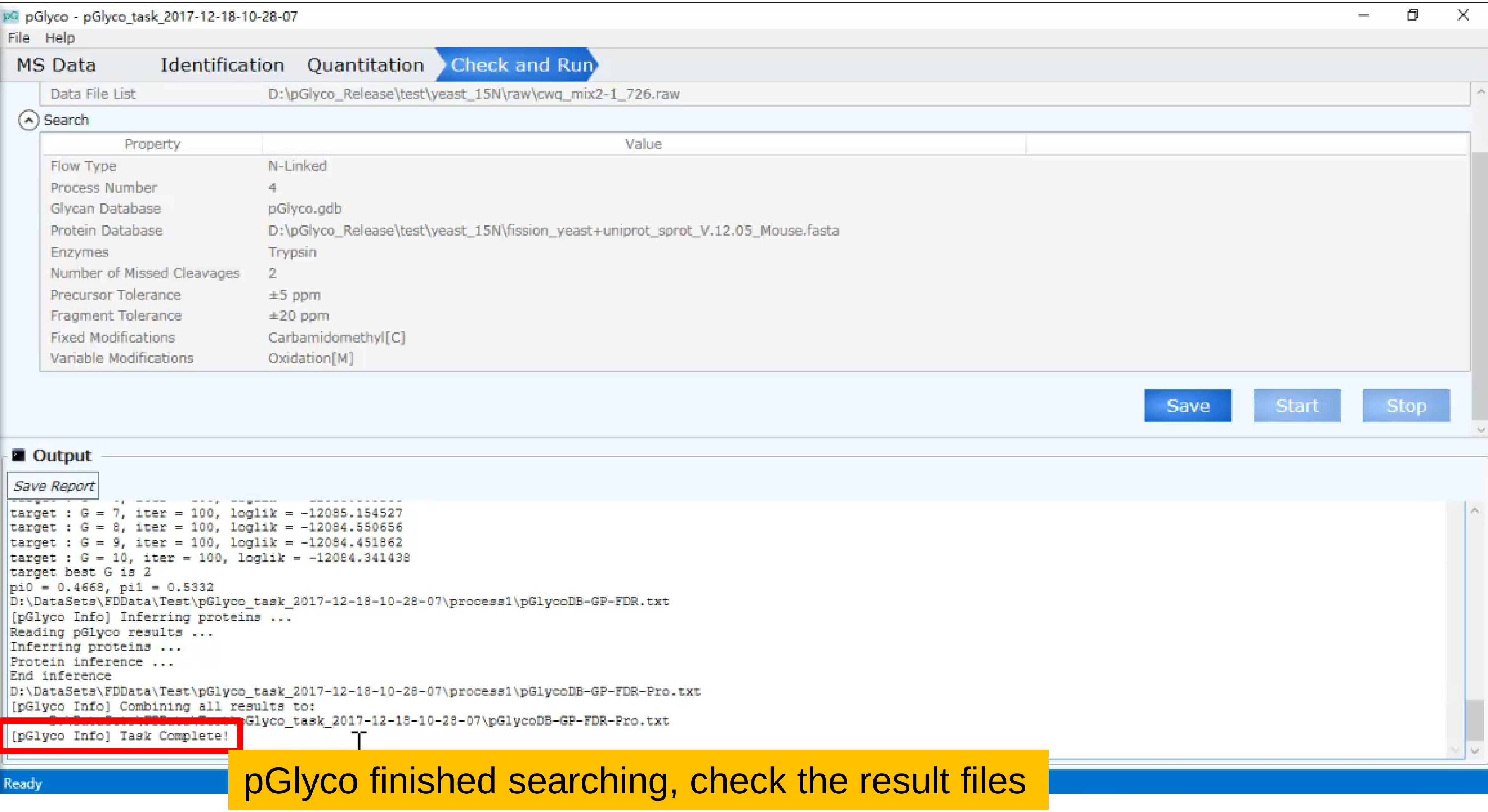
Property	Value
Flow Type	N-Linked
Process Number	4
Glycan Database	pGlyco.gdb
Protein Database	D:\pGlyco_Release\test\yeast_15N\fission_yeast+uniprot_sprot_V.12.05_Mouse.fasta
Enzymes	Trypsin
Number of Missed Cleavages	2
Precursor Tolerance	±5 ppm
Fragment Tolerance	±20 ppm
Fixed Modifications	Carbamidomethyl[C]
Variable Modifications	Oxidation[M]

```
[pParse] MS1list completes.
[pParse] Processing cwq_mix2-1_726.MS1
[pParse] CSV file already exists, pParse Complete!
[pParse]  == == == Time elapsed: 0 seconds. == == == =
*****************************************************
*                                                   *
*             * Welcome to pGlyco2 *                *
*                                                   *
*              * pglyco@ict.ac.cn *                 *
*                                                   *
*****************************************************

[pGlyco Info] Loading glycan database ...
[pGlyco Info] Ion Indexing ...
[pGlyco Info] D:\pGlyco_Release\test\yeast_15N\raw\cwq_mix2-1_726_HCDFT.mgf
Process 1: [I.D.] 68.9% [==================================                ]
```

pGlyco is searching the raw files

Search



Data File List: D:\pGlyco_Release\test\yeast_15N\raw\cwq_mix2-1_726.raw

Property	Value
Flow Type	N-Linked
Process Number	4
Glycan Database	pGlyco.gdb
Protein Database	D:\pGlyco_Release\test\yeast_15N\fission_yeast+uniprot_sprot_V.12.05_Mouse.fasta
Enzymes	Trypsin
Number of Missed Cleavages	2
Precursor Tolerance	±5 ppm
Fragment Tolerance	±20 ppm
Fixed Modifications	Carbamidomethyl[C]
Variable Modifications	Oxidation[M]

```
target : G = 7, iter = 100, loglik = -12085.154527
target : G = 8, iter = 100, loglik = -12084.550656
target : G = 9, iter = 100, loglik = -12084.451862
target : G = 10, iter = 100, loglik = -12084.341438
target best G is 2
pi0 = 0.4668, pi1 = 0.5332
D:\DataSets\FDData\Test\pGlyco_task_2017-12-18-10-28-07\process1\pGlycoDB-GP-FDR.txt
[pGlyco Info] Inferring proteins ...
Reading pGlyco results ...
Inferring proteins ...
Protein inference ...
End inference
D:\DataSets\FDData\Test\pGlyco_task_2017-12-18-10-28-07\process1\pGlycoDB-GP-FDR-Pro.txt
[pGlyco Info] Combining all results to:
    D:\DataSets\FDData\Test\pGlyco_task_2017-12-18-10-28-07\pGlycoDB-GP-FDR-Pro.txt
[pGlyco Info] Task Complete!
```

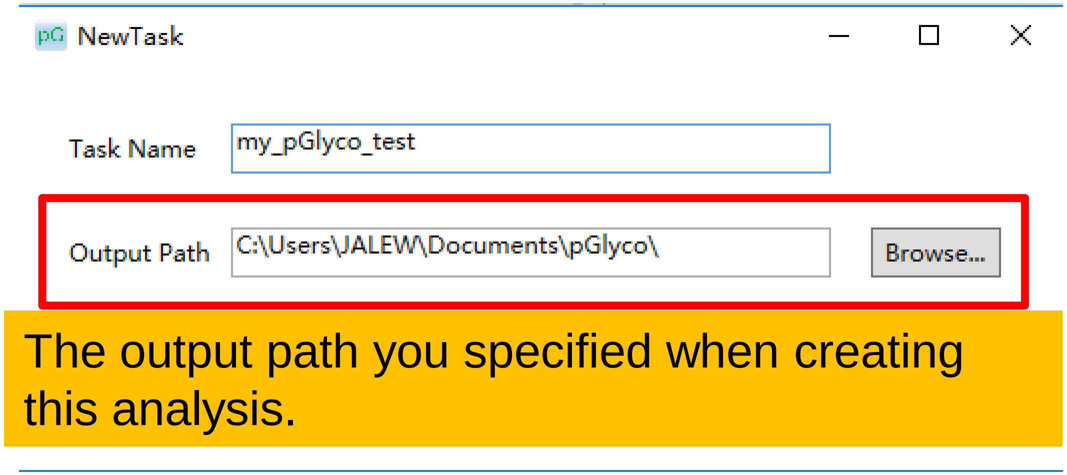
pGlyco finished searching, check the result files

Catalog

- Activation
- Search
- Results
- Use gLabel to annotate GPSM

Results

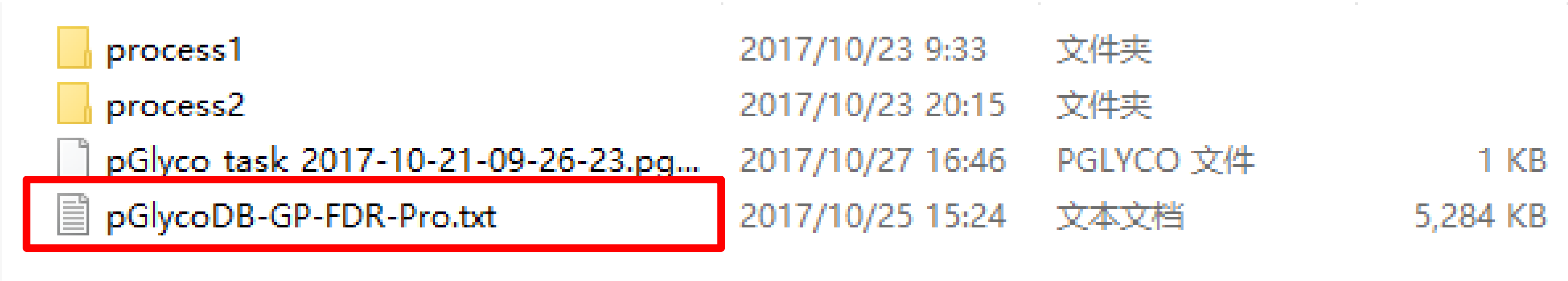
- Final result file is located in “Output Path”, the file name is pGlycoDB-GP-FDR-Pro.txt



Task Name: my_pGlyco_test

Output Path: C:\Users\JALEW\Documents\pGlyco\

The output path you specified when creating this analysis.

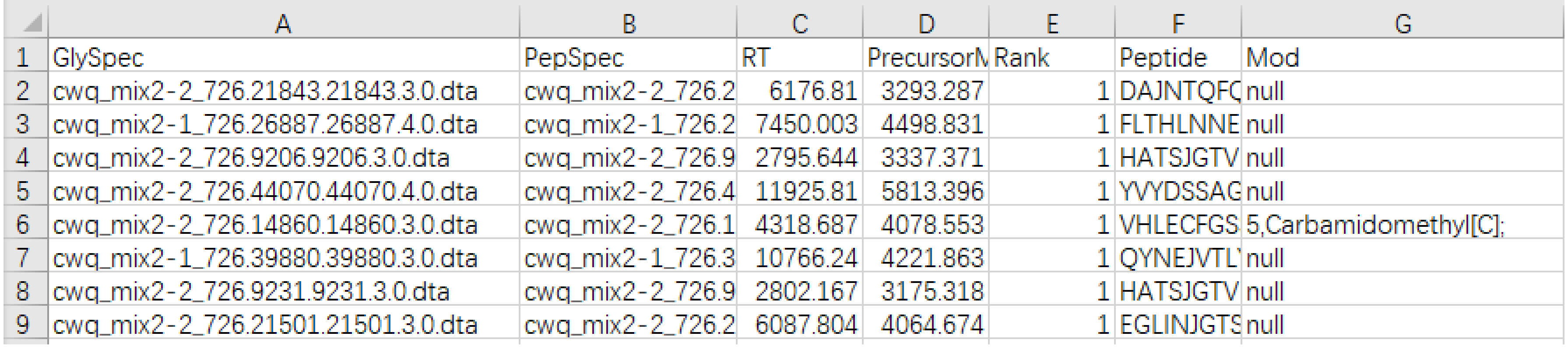


process1	2017/10/23 9:33	文件夹	
process2	2017/10/23 20:15	文件夹	
pGlyco task 2017-10-21-09-26-23.pg...	2017/10/27 16:46	PGLYCO 文件	1 KB
pGlycoDB-GP-FDR-Pro.txt	2017/10/25 15:24	文本文档	5,284 KB

Results

- Open pGlycoDB-GP-FDR-Pro.txt in Microsoft Excel

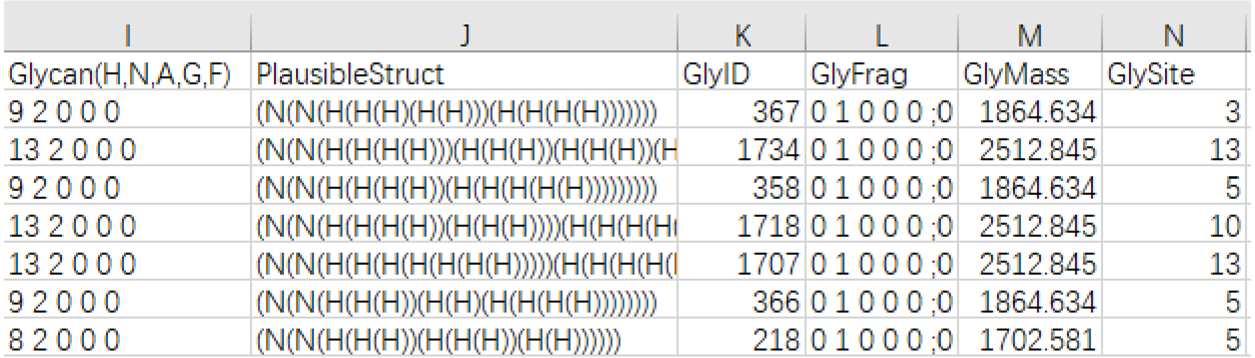
- Column information
  - GlySpec: spectrum name (raw_name.scan.scan.charge.mixid.dta)
  - PepSpec: same as GlySpec
  - RT: retention time in seconds
  - Peptide / Mod: peptide and modification information
    - pGlyco will automatically convert N in N-X-S/T/C to J



	A	B	C	D	E	F	G
1	GlySpec	PepSpec	RT	PrecursorM	Rank	Peptide	Mod
2	cwq_mix2-2_726.21843.21843.3.0.dta	cwq_mix2-2_726.2	6176.81	3293.287	1	DAJNTQFC	null
3	cwq_mix2-1_726.26887.26887.4.0.dta	cwq_mix2-1_726.2	7450.003	4498.831	1	FLTHLNNE	null
4	cwq_mix2-2_726.9206.9206.3.0.dta	cwq_mix2-2_726.9	2795.644	3337.371	1	HATSJGTV	null
5	cwq_mix2-2_726.44070.44070.4.0.dta	cwq_mix2-2_726.4	11925.81	5813.396	1	YVYDSSAG	null
6	cwq_mix2-2_726.14860.14860.3.0.dta	cwq_mix2-2_726.1	4318.687	4078.553	1	VHLECFGS	5,Carbamidomethyl[C];
7	cwq_mix2-1_726.39880.39880.3.0.dta	cwq_mix2-1_726.3	10766.24	4221.863	1	QYNEJVTL	null
8	cwq_mix2-2_726.9231.9231.3.0.dta	cwq_mix2-2_726.9	2802.167	3175.318	1	HATSJGTV	null
9	cwq_mix2-2_726.21501.21501.3.0.dta	cwq_mix2-2_726.2	6087.804	4064.674	1	EGLINJGTS	null

Results

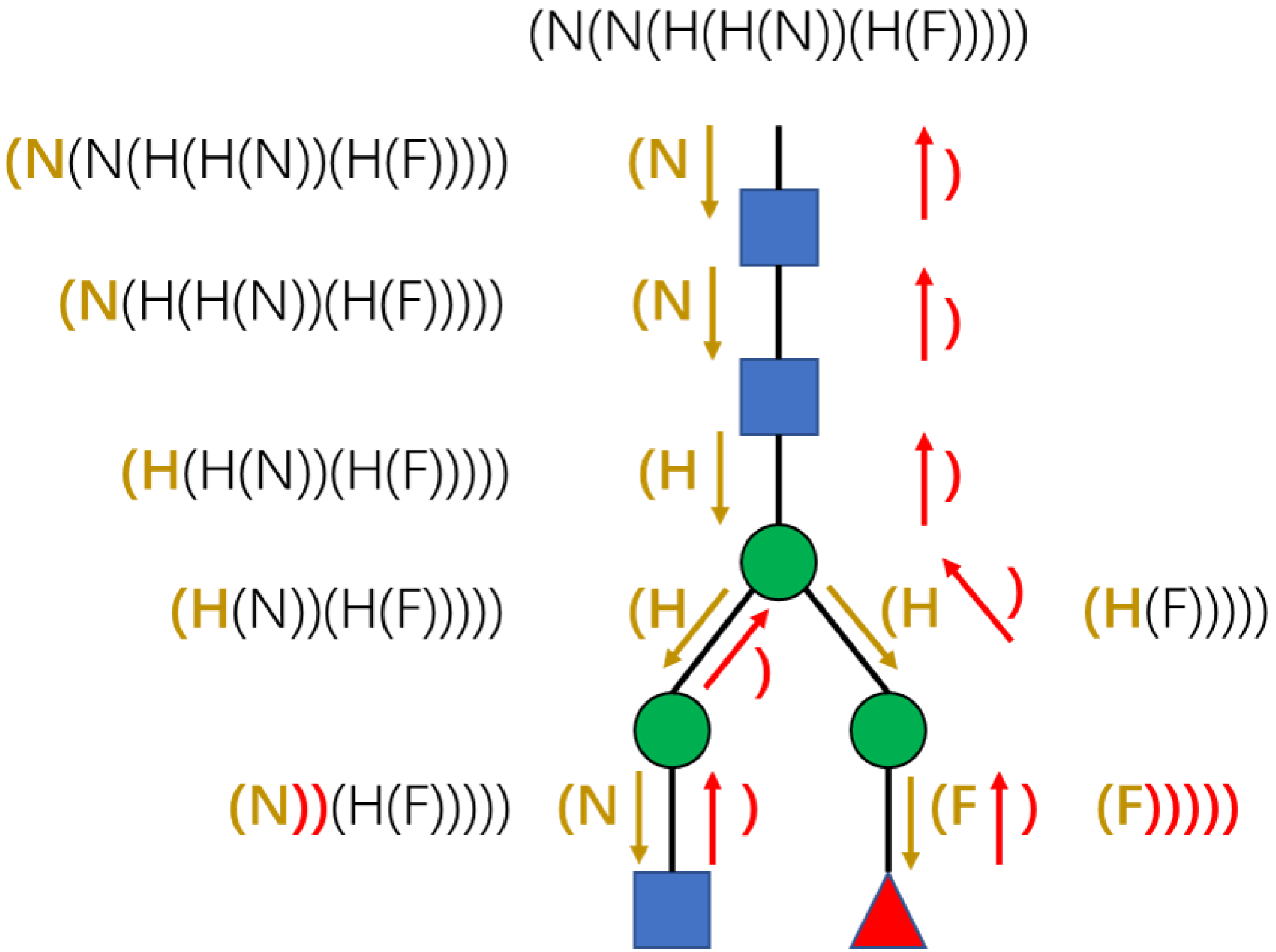
- Column information:
  - Glycan(H,N,A,G,F): glycan composition,
    - H = #Hex, N = #HexNAc, A = #NeuAc, G = #NeuGc, F = #Fuc
  - PlausibleStruct: plausible glycan structure in canonical form, we can decode the canonical string into the structure (see next slide).
  - GlySite: glycosylated site of peptide sequence



I	J	K	L	M	N
Glycan(H,N,A,G,F)	PlausibleStruct	GlyID	GlyFrag	GlyMass	GlySite
9 2 0 0 0	(N(N(H(H(H)(H(H)))(H(H(H(H)))))))	367	0 1 0 0 0 ;0	1864.634	3
13 2 0 0 0	(N(N(H(H(H(H)))(H(H(H))(H(H(H))(H	1734	0 1 0 0 0 ;0	2512.845	13
9 2 0 0 0	(N(N(H(H(H(H))(H(H(H(H(H)))))))))	358	0 1 0 0 0 ;0	1864.634	5
13 2 0 0 0	(N(N(H(H(H(H))(H(H(H))))(H(H(H(H	1718	0 1 0 0 0 ;0	2512.845	10
13 2 0 0 0	(N(N(H(H(H(H(H(H(H)))))(H(H(H(H(	1707	0 1 0 0 0 ;0	2512.845	13
9 2 0 0 0	(N(N(H(H(H))(H(H)(H(H(H(H)))))))	366	0 1 0 0 0 ;0	1864.634	5
8 2 0 0 0	(N(N(H(H(H))(H(H(H))(H(H))))))	218	0 1 0 0 0 ;0	1702.581	5

Results

- Decoding the canonical structure



(N(N(H(H(N))(H(F)))))

(N(N(H(H(N))(H(F)))))

(N(H(H(N))(H(F)))))

(H(H(N))(H(F)))))

(H(N))(H(F)))))

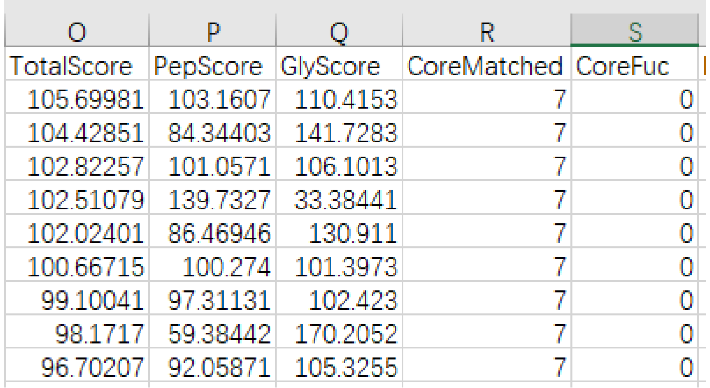
(N))(H(F)))))

(H(F)))))

(F)))))

Results

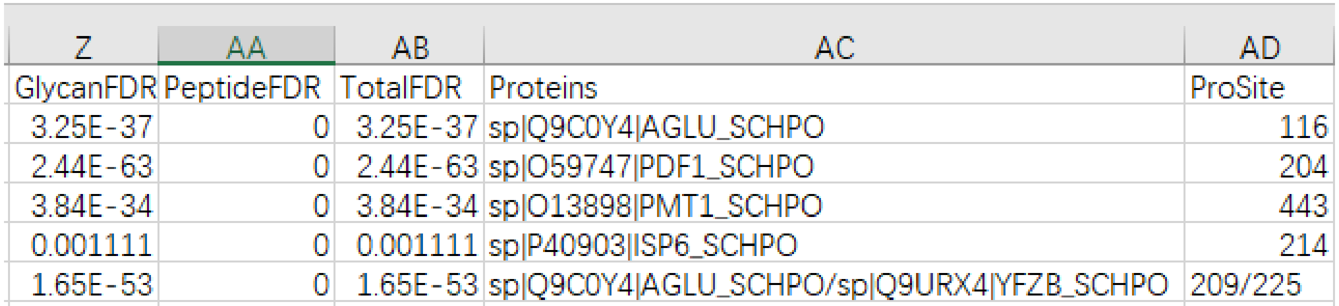
- Column information
  - TotalScore: glycopeptide score
  - PepScore: peptide score
  - GlyScore: glycan score
  - CoreMatched: How many trimannosyl-core ions matched
  - CoreFuc: core-fucosylated ions, N1F1 and N2F1 in pGlyco
    - 0: not core-fuc
    - 10: core-fuc, but no core-fuc ions matched
    - 11: core-fuc, matched one core-fuc ion
    - 12: core-fuc, matched two core-fuc ions



O	P	Q	R	S
TotalScore	PepScore	GlyScore	CoreMatched	CoreFuc
105.69981	103.1607	110.4153	7	0
104.42851	84.34403	141.7283	7	0
102.82257	101.0571	106.1013	7	0
102.51079	139.7327	33.38441	7	0
102.02401	86.46946	130.911	7	0
100.66715	100.274	101.3973	7	0
99.10041	97.31131	102.423	7	0
98.1717	59.38442	170.2052	7	0
96.70207	92.05871	105.3255	7	0

Results

- Column information
  - GlycanFDR
  - PeptideFDR
  - TotalFDR: glycopeptide FDR at spectrum level
  - Proteins: inferred protein AC (separated by “/” if multiple proteins inferred)
  - ProSite: glycosylated sites of inferred proteins, separated by “/”



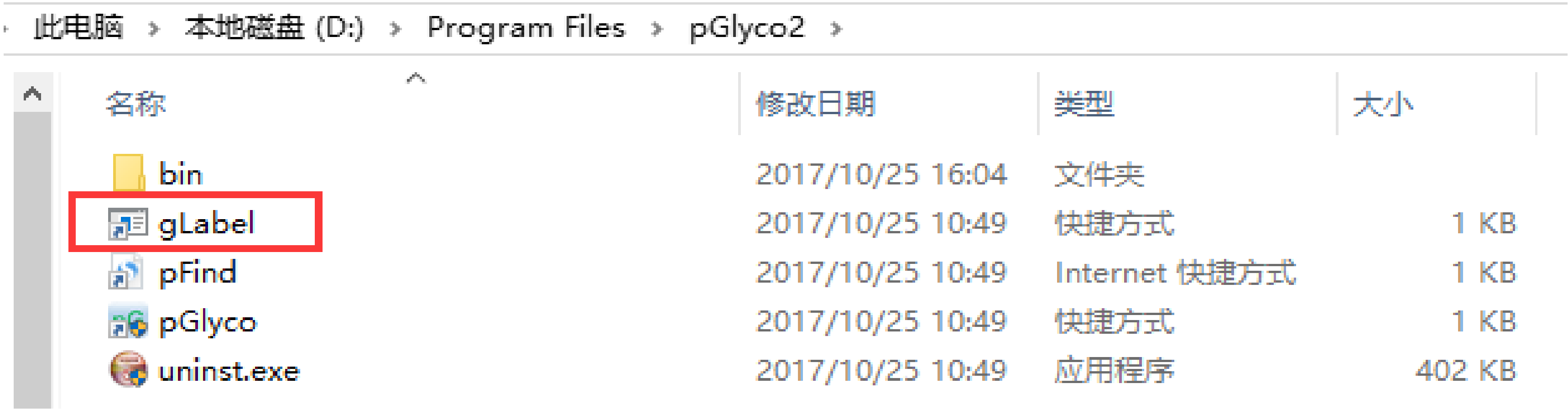
Z	AA	AB	AC	AD
GlycanFDR	PeptideFDR	TotalFDR	Proteins	ProSite
3.25E-37	0	3.25E-37	sp|Q9C0Y4|AGLU_SCHPO	116
2.44E-63	0	2.44E-63	sp|O59747|PDF1_SCHPO	204
3.84E-34	0	3.84E-34	sp|O13898|PMT1_SCHPO	443
0.001111	0	0.001111	sp|P40903|ISP6_SCHPO	214
1.65E-53	0	1.65E-53	sp|Q9C0Y4|AGLU_SCHPO/sp|Q9URX4|YFZB_SCHPO	209/225

Catalog

- **Activation**
- **Search**
- **Results**
- **Use gLabel to annotate GPSM**

gLabel

- Use gLabel to annotate the spectrum

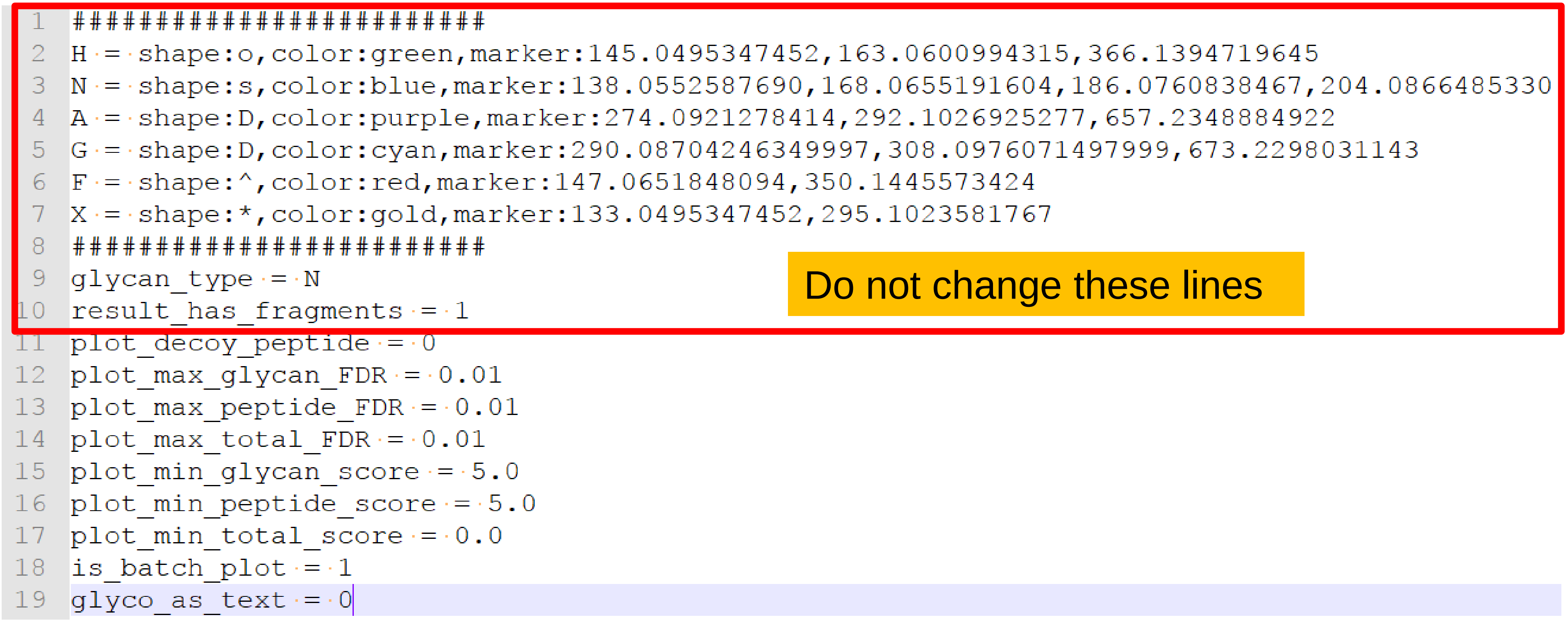


此电脑 › 本地磁盘 (D:) › Program Files › pGlyco2 ›

名称	修改日期	类型	大小
bin	2017/10/25 16:04	文件夹	
gLabel	2017/10/25 10:49	快捷方式	1 KB
pFind	2017/10/25 10:49	Internet 快捷方式	1 KB
pGlyco	2017/10/25 10:49	快捷方式	1 KB
uninst.exe	2017/10/25 10:49	应用程序	402 KB

gLabel

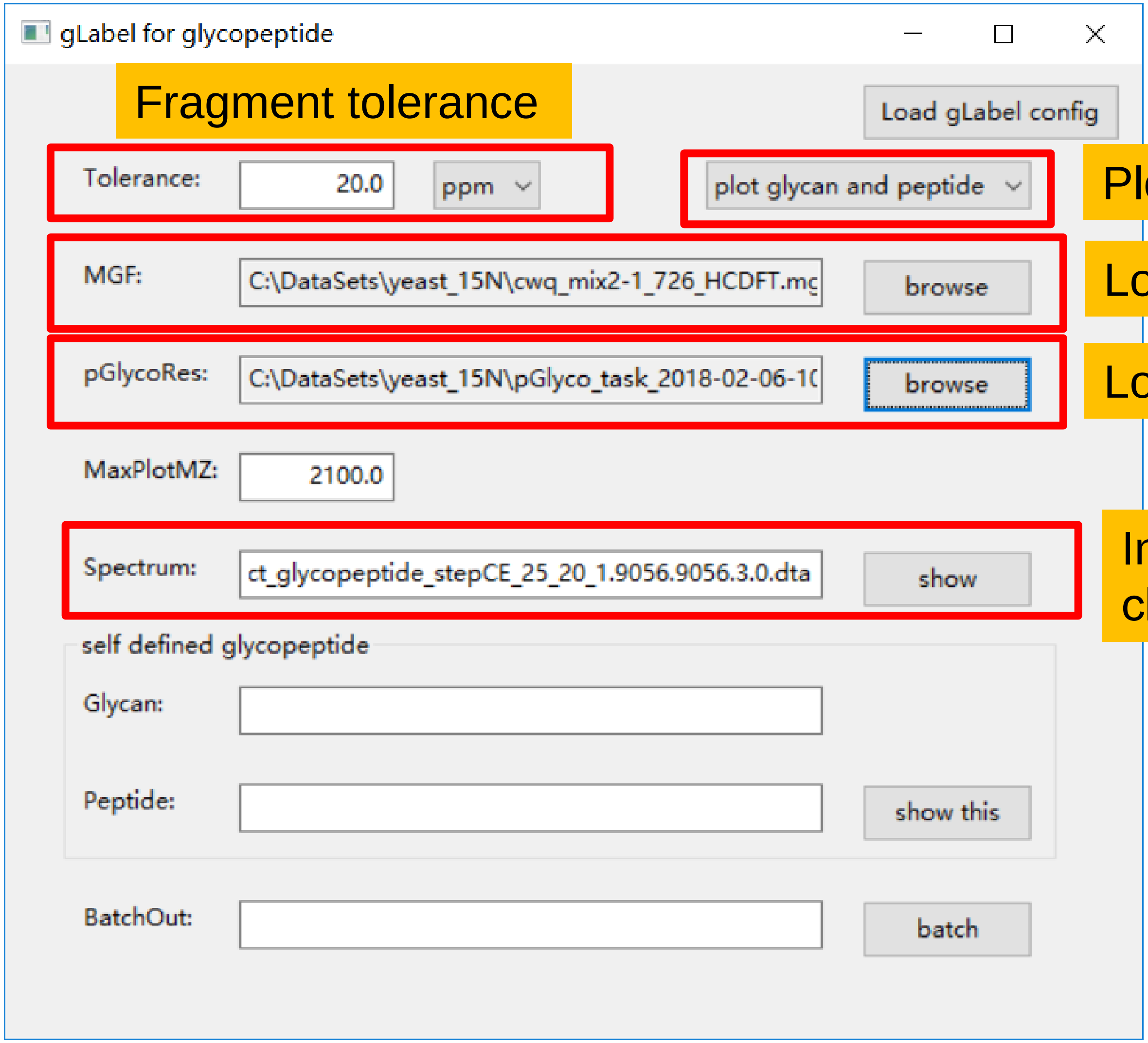
- We used glabel.gconf file in the bin folder to control the the plot, you can modify this file and then run gLabel if necessary.



```
#########################
H = shape:o,color:green,marker:145.0495347452,163.0600994315,366.1394719645
N = shape:s,color:blue,marker:138.0552587690,168.0655191604,186.0760838467,204.0866485330
A = shape:D,color:purple,marker:274.0921278414,292.1026925277,657.2348884922
G = shape:D,color:cyan,marker:290.08704246349997,308.0976071497999,673.2298031143
F = shape:^,color:red,marker:147.0651848094,350.1445573424
X = shape:*,color:gold,marker:133.0495347452,295.1023581767
#########################
glycan_type = N
result_has_fragments = 1
plot_decoy_peptide = 0
plot_max_glycan_FDR = 0.01
plot_max_peptide_FDR = 0.01
plot_max_total_FDR = 0.01
plot_min_glycan_score = 5.0
plot_min_peptide_score = 5.0
plot_min_total_score = 0.0
is_batch_plot = 1
glyco_as_text = 0
```

Do not change these lines

gLabel – show a GPSM



gLabel for glycopeptide

Fragment tolerance

Tolerance: 20.0 ppm

plot glycan and peptide

Plotting option

MGF: C:\DataSets\yeast_15N\cwq_mix2-1_726_HCDFT.mc

Load MGF

pGlycoRes: C:\DataSets\yeast_15N\pGlyco_task_2018-02-06-1(

Load result file

MaxPlotMZ: 2100.0

Spectrum: ct_glycopeptide_stepCE_25_20_1.9056.9056.3.0.dta

Input spectrum name and click “show”

self defined glycopeptide

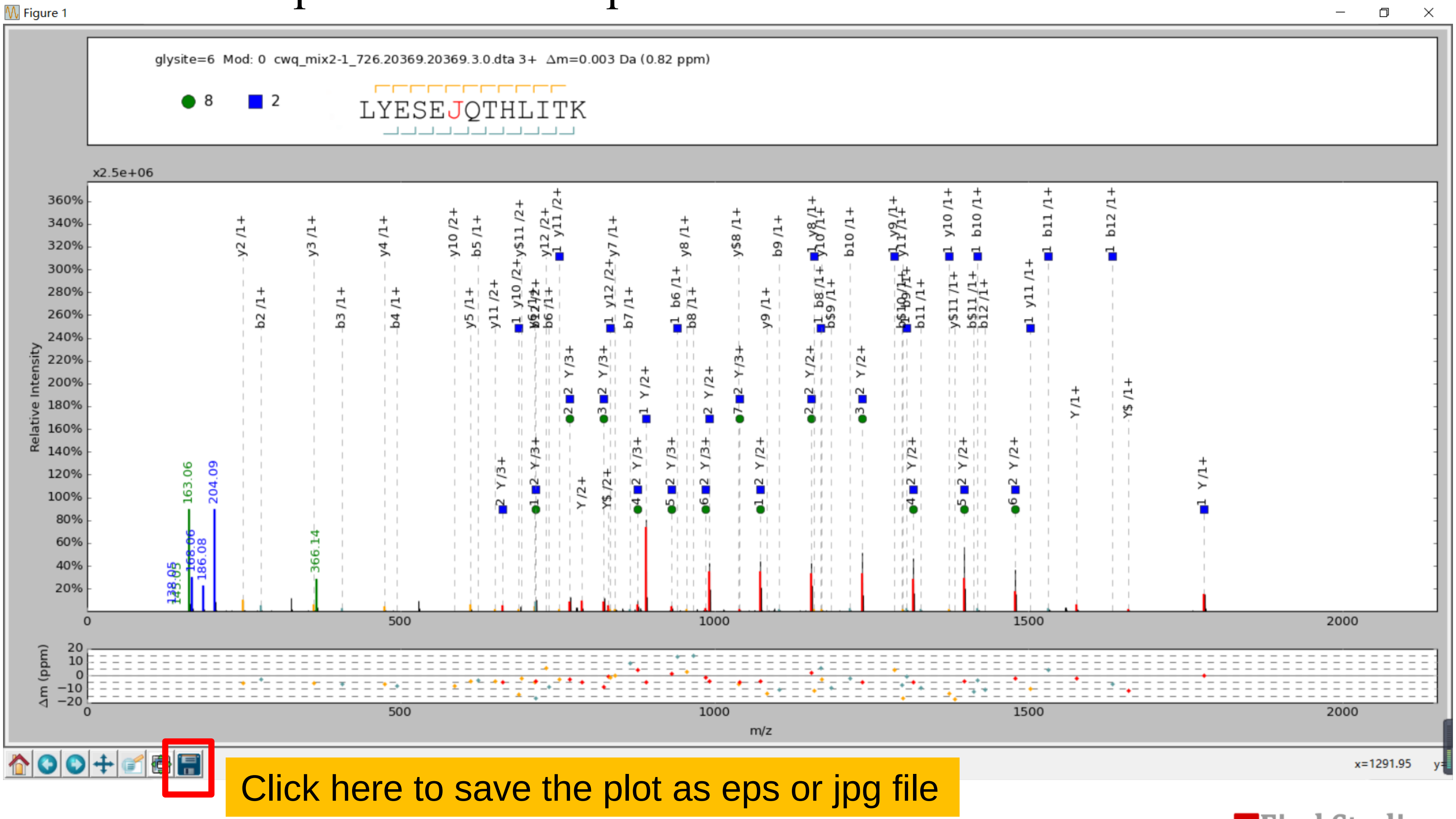
Glycan:

Peptide:

BatchOut:

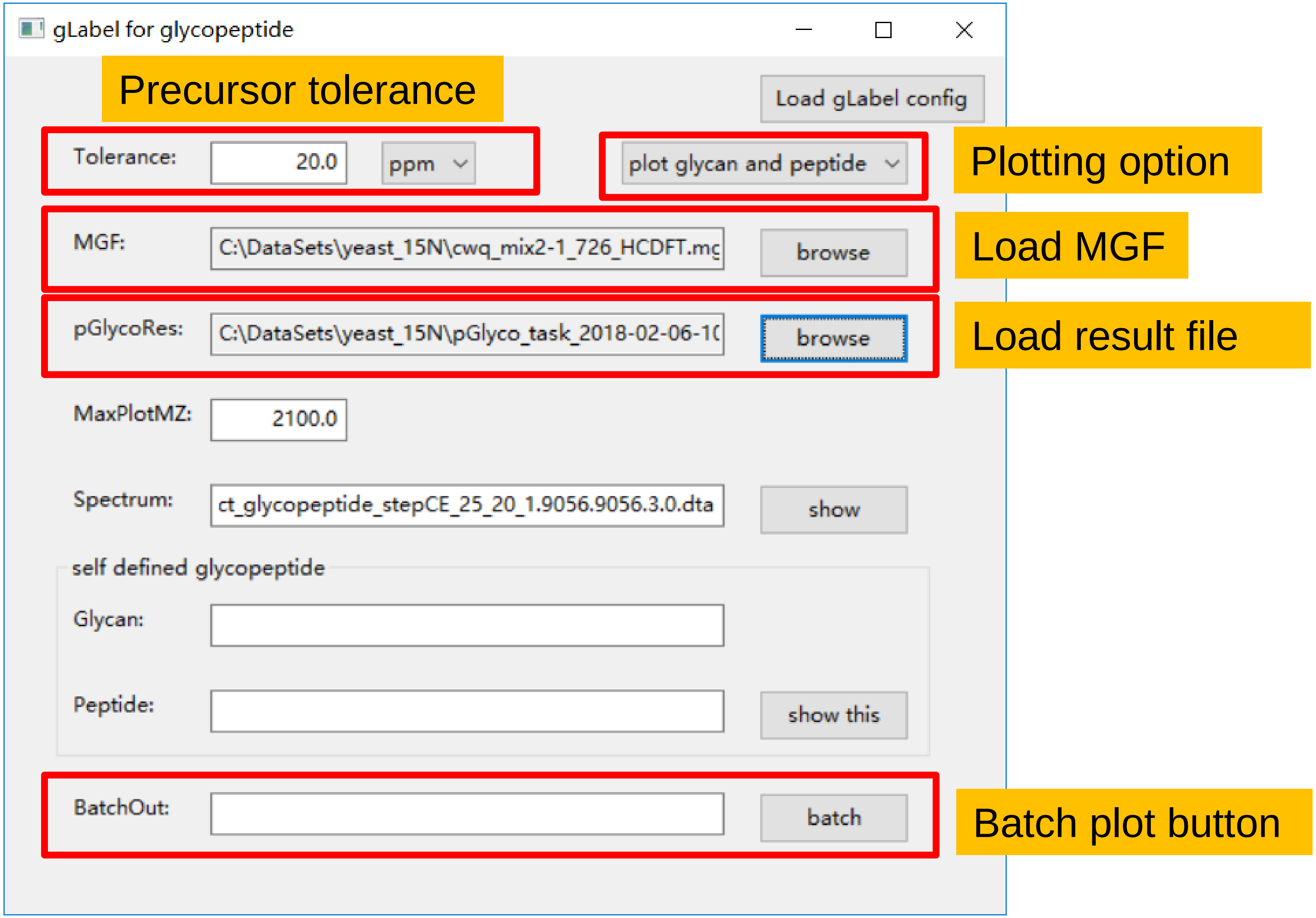
gLabel – show a GPSM

- After input the correct spectrum name and click “show”



Click here to save the plot as eps or jpg file

gLabel – batch plot



gLabel for glycopeptide

Tolerance: 20.0 ppm

plot glycan and peptide

MGF: C:\DataSets\yeast_15N\cwq_mix2-1_726_HCDFT.mg

pGlycoRes: C:\DataSets\yeast_15N\pGlyco_task_2018-02-06-1(

MaxPlotMZ: 2100.0

Spectrum: ct_glycopeptide_stepCE_25_20_1.9056.9056.3.0.dta

self defined glycopeptide

Glycan:

Peptide:

BatchOut:

Precursor tolerance

Plotting option

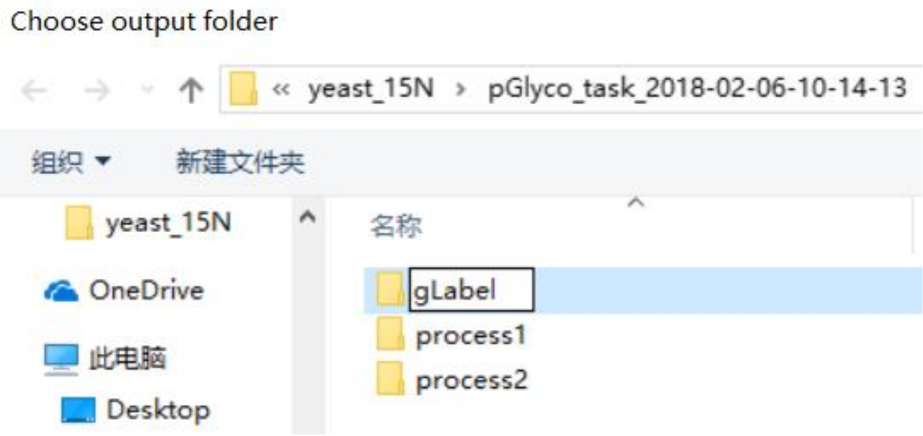
Load MGF

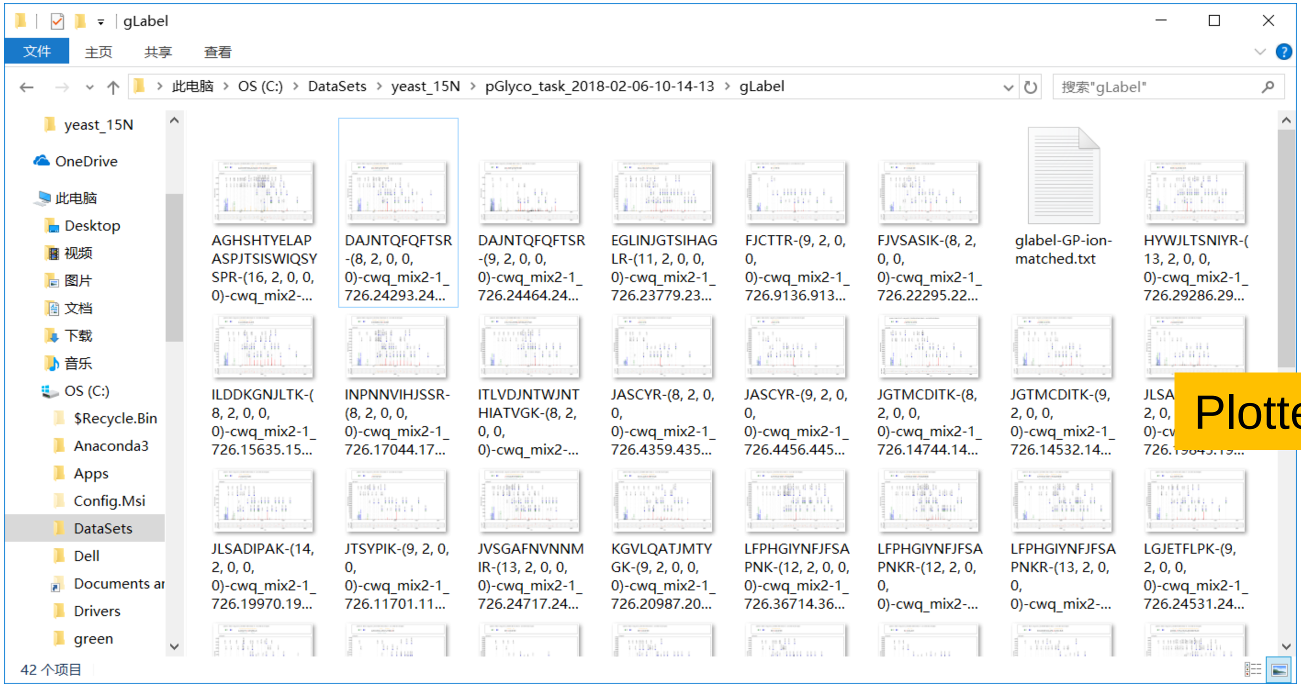
Load result file

Batch plot button

gLabel – batch plot

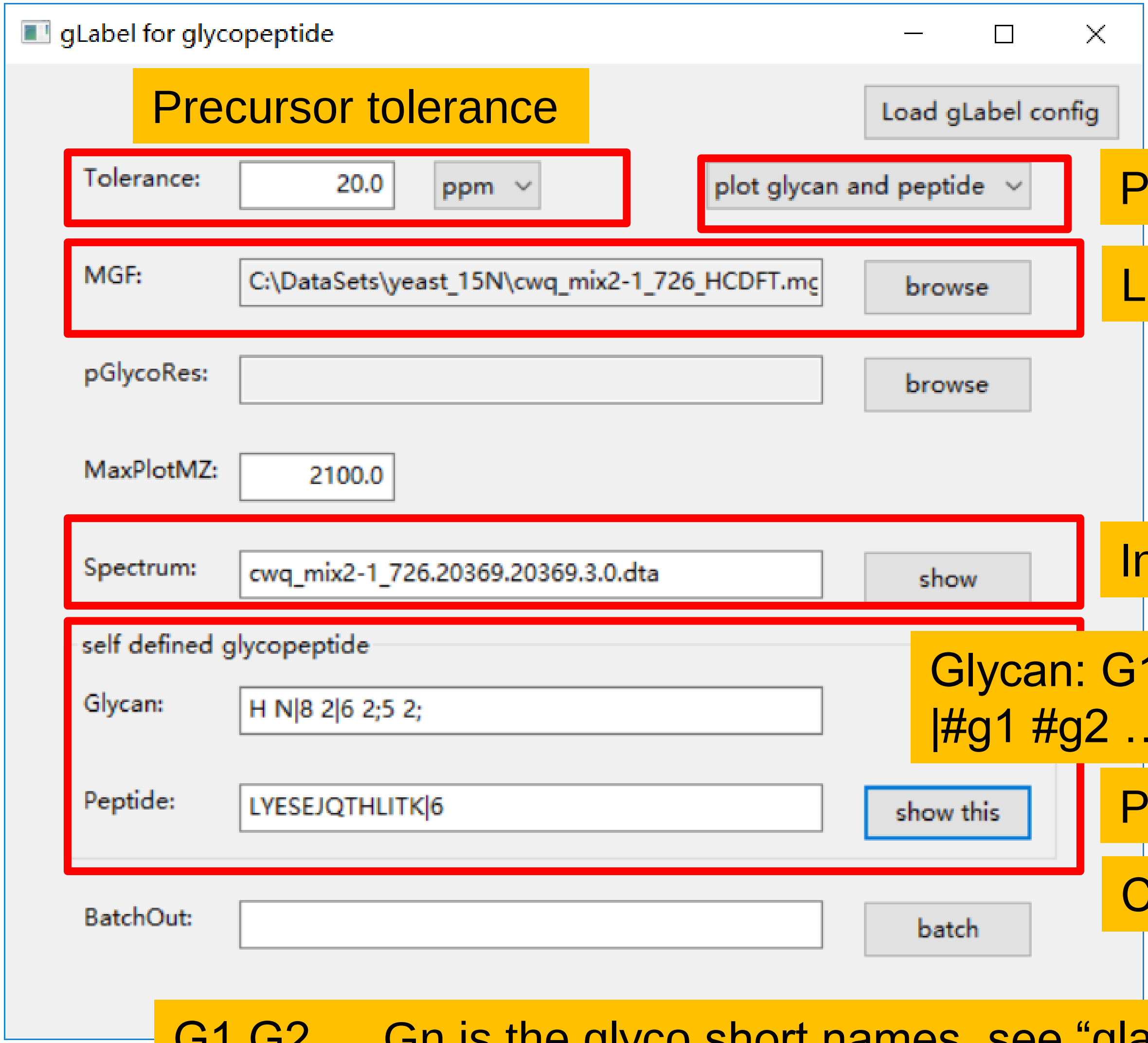
- Click “batch” button, choose an (empty) folder to store the plotted gLabel figures.





Plotted figures in the folder

gLabel – User defined glycopeptides



gLabel for glycopeptide

Tolerance: 20.0 ppm

plot glycan and peptide

MGF: C:\DataSets\yeast_15N\cwq_mix2-1_726_HCDFT.mg

pGlycoRes:

MaxPlotMZ: 2100.0

Spectrum: cwq_mix2-1_726.20369.20369.3.0.dta

self defined glycopeptide

Glycan: H N|8 2|6 2;5 2;

Peptide: LYESEJQTHLITK|6

BatchOut:

Precursor tolerance

Plotting option

Load MGF

Input spectrum name

Glycan: G1 G2 … Gn|#G1 #G2 … #Gn |#g1 #g2 … #gn;#g1 #g2 …

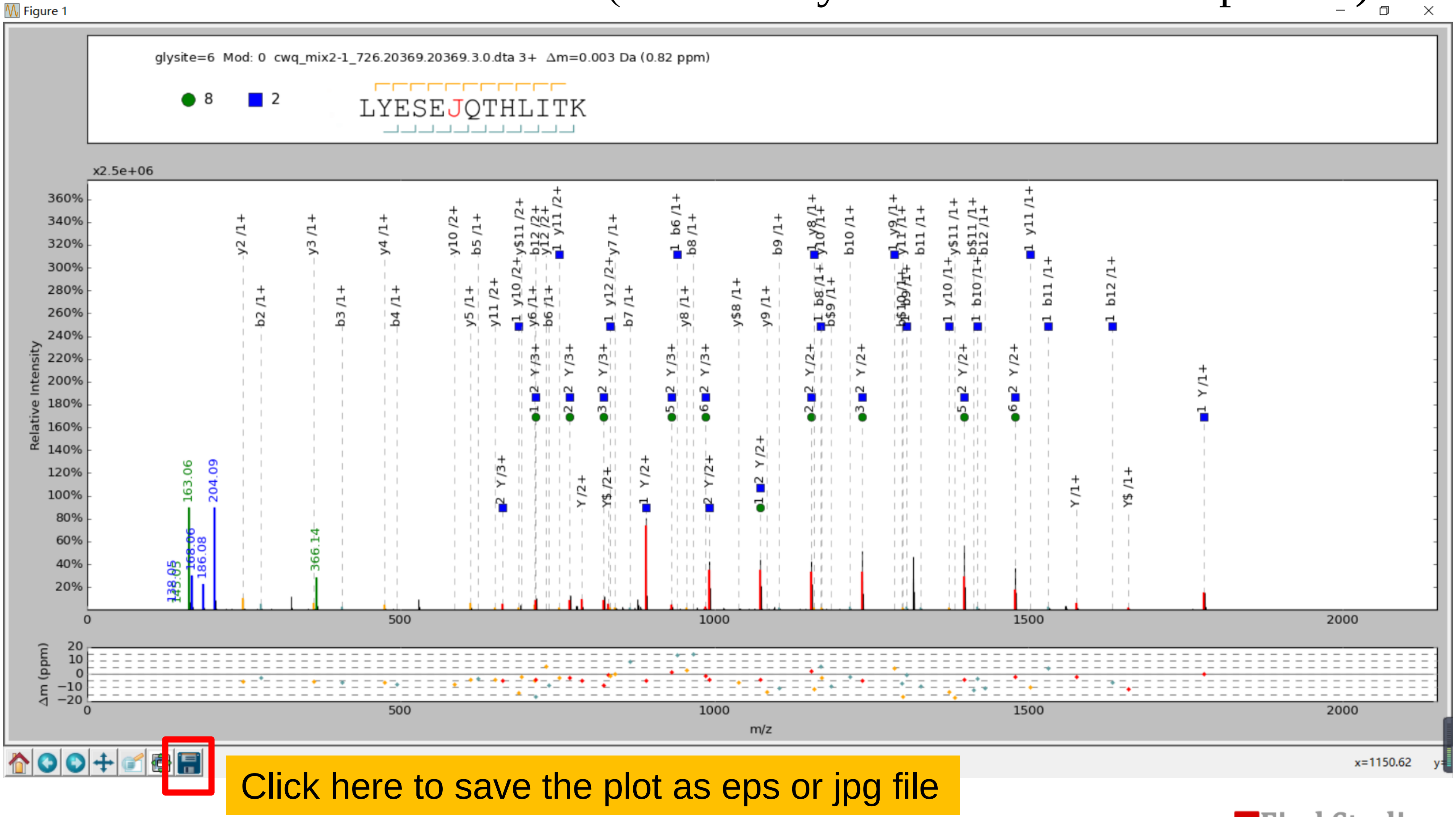
Peptide: sequence|glysite

Click “show this”

G1 G2 … Gn is the glyco short names, see “glabel.gconf” in the bin folder for details

gLabel – User defined glycopeptides

- After click “show this” (trimannosyl core ions also are plotted)



glysite=6 Mod: 0 cwq_mix2-1_726.20369.20369.3.0.dta 3+ Δm=0.003 Da (0.82 ppm)

Click here to save the plot as eps or jpg file

ENDS

- **Thank you for using pGlyco!**

- **If you have any questions, please contact pglyco@ict.ac.cn.**

- **You can also post issues at github for discussion:**
  - **https://github.com/pFindStudio/pGlyco2**
  - **how to post issues at github?**
    - **see http://pfind.ict.ac.cn/file/github.pdf**