

MS-Format 使用说明 (FuLab)

1、收到平台蛋白搜库结果后，下载 XML 结果

User : Zhen Li
E-mail :
MS data file : G:\UserData\QE1\2021\202110\20211013\fangxiaomin-gfp0a_20211013141321.raw
Databases : 1: contaminants (517 sequences; 241,338 residues)
2: Flybase_fangxiaomin 20210121 (30,724 sequences; 20,308,477 residues) ← 2 关注这里是搜库时所使用的数据库
Timestamp : 14 Oct 2021 at 03:32:22 GMT

Re-search ☒ All ☐ Non-significant ☐ Unassigned [help] Export As XML

Not what you expected? Try [the select summary](#).

► Search parameters
► Score distribution
► Modification statistics for all protein families
► Legend

Protein Family Summary

Format	Significance threshold p<	0.05
	Display non-sig. matches	<input type="checkbox"/>
	Show Percolator scores	<input type="checkbox"/>
	Preferred taxonomy	All entries

► Sensitivity and FDR (reversed protein sequence)

Proteins (211) | [Report Builder](#) | [Unassigned \(11,227\)](#)

发件人: 中国农业大学生物质谱实验室
发送时间: 2021年10月14日 11:37
收件人: [redacted]
主题: 蛋白鉴定结果

方晓敏您好, 蛋白鉴定结果见链接, 请查收:
http://10.2.149.230/mascot/cgi/master_results_2.pl?file=.%2Fdata%2F20211014%2FF008158.dat

实验数据和搜索结果仅在生物质谱实验室服务器保存一年, 请及时备份和保存实验结果。
使用生物质谱实验室数据发表论文时请在文章中致谢生物质谱实验室。

中国农业大学生物质谱实验室生命科学研究中心1062室
电话: 62731128

Export search results

Help

Export format XML

Significance threshold p< 0.05 at ☐ identity ☒ homology

Target FDR (overrides significance threshold if set) <not set>

FDR type Distinct PSMs

Display non-significant matches ☐ ← 5 这个不勾选, 其余使用默认设置, 勾选全部信息

Max. number of hits AUTO

Min. number of sig. unique sequences 1

Protein scoring ☐ Standard ☒ MudPIT

Include same-set protein hits (additional proteins that span the same set of peptides) ☐

Include sub-set protein hits (additional proteins that span a sub-set of peptides) 1

**MATRIX
SCIENCE**

Search this site

[Home](#) [Access Mascot Server](#) [Database search help](#) [Contact](#)

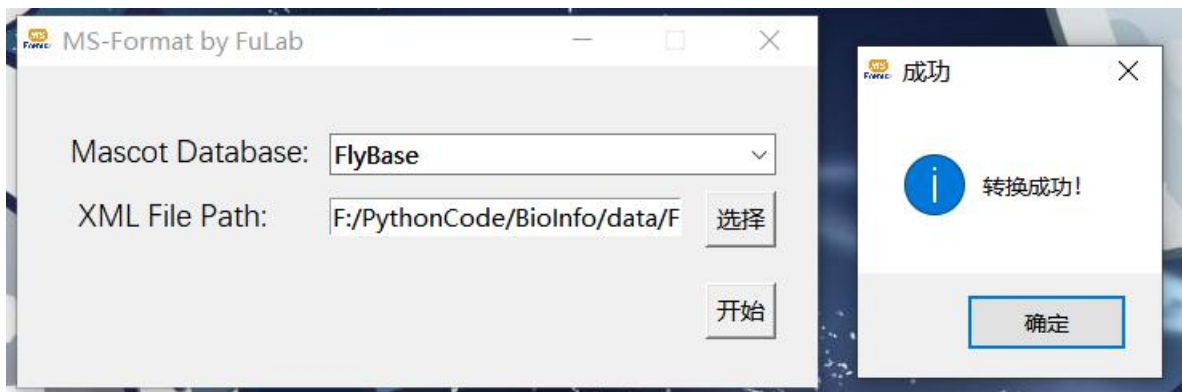
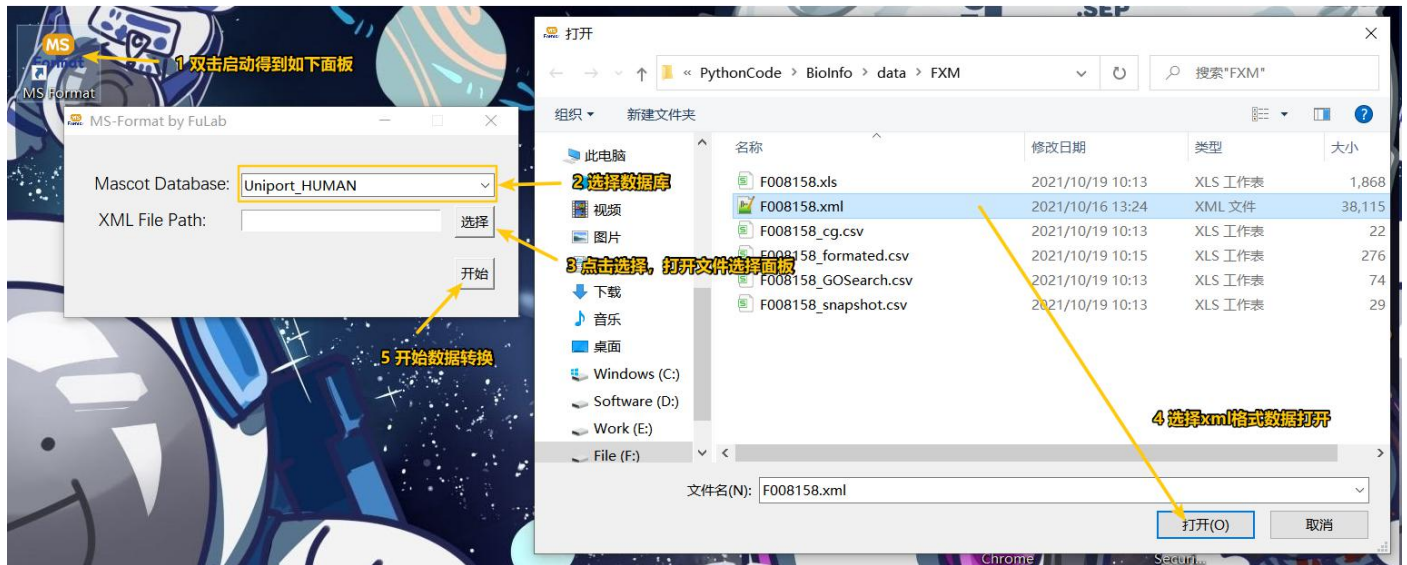
Mascot database search > Access Mascot Server > Export search results

Exporting is ready to be started.

Download

6 下载导出结果

3、使用 MS-Format 整理质谱搜库结果



4、此时同目录中会多出一些文件:

其中与输入文件同名的 xls 文件为 XML 中信息的不等价转化, 将 Mascot 搜库得到的蛋白、肽段信息整理为两个表格; 其中后缀为 "_formatted.csv" 文件为格式转化、信息整合后的文件, 建议大家调整格式后将其保存为 xls 格式文件。

FlyBase 结果表头:

CG	Protein name	Score	Mass	#of spectra	#of peptides	coverage	biological	cellular	molecular	snapshot
CG6699	β' COP-PA	339	103618	29	19	29.4	intracellular	COPI vesicle	structural	Contributions welcome
CG7961	α COP-PB	313	140952	34	23	21.4	intracellular	COPI vesicle	structural	Contributions welcome
CG7961	α COP-PA	313	140952	34	23	21.4	intracellular	COPI vesicle	structural	Contributions welcome

Uniprot 结果表头:

hit#	Accession	score	mass	#of peptides	Gene names	Protein name	Function	Gene ontology (cellular component)
1	Q8IW35	6328	97776	301	CEP97 LRR1	Centrosome	FUNCTION: ciliary base	moduliferous centriolar satellite [GO:0034451];
1	E9PG22	5859	91052	278	CEP97	Centrosomal protein of 97 kDa		centriolar satellite [GO:0034451];
1	H7C4Y5	867	19727	55	CEP97	Centrosomal protein of 97 kDa (Fragment)		