

1、数据准备

打开 IP2 的结果，在 export each protein individual 后面选择 excel 格式下载文件

打开文件后保存成 csv 文件。（如果直接选择 csv 选项下载数据，表头会丢失）

The screenshot displays the IP2 web interface with a table of protein data. The table has columns for protein ID, name, BioGPS score, and other identifiers. Below the table, there are export options for the entire group and for each protein individually. The 'Excel' option is highlighted in the 'Export each protein as individual' section. A red box labeled '打开文件' (Open File) points to the 'protein_group_list (5)...' file in the Windows file explorer.

view	Q61702	BioGPS	2.3%	1	1	0.001320248	0.05005484	Zinc-alpha-2-glycoprotein OS=Mus musculus GN=A2Gp1 PE=1 SV=2	0	1
view	Q61702	BioGPS	2.4%	2	2	8.937512E-4	0.05681753	Inter-alpha-trypsin inhibitor heavy chain H1 OS=Mus musculus GN=Ith1 PE=1 SV=2	0	2
	Q91WU7	BioGPS	2.4%			8.9178473E-4	0.05681753	Inter-alpha trypsin inhibitor, heavy chain 1 (Fragment) OS=Mus musculus GN=Ith1 PE=2 SV=1		
	Q8K159	BioGPS	2.4%			8.898269E-4	0.05681753	Ith1 protein OS=Mus musculus GN=Ith1 PE=2 SV=1		
	F8WJ05	BioGPS	2.4%			8.898269E-4	0.05681753	Inter-alpha-trypsin inhibitor heavy chain H1 OS=Mus musculus GN=Ith1 PE=1 SV=1		
view	Q3UV17	BioGPS	1.5%	2	2	0.0013647008	0.035142183	Keratin, type II cytoskeletal 2 oral OS=Mus musculus GN=Krt6 PE=1 SV=1	0	2
view	Q8BM75	BioGPS	1.2%	2	2	6.823504E-4	0.028016329	AT-rich interactive domain-containing protein 5B OS=Mus musculus GN=And5b PE=1 SV=3	0	2
view	Q61329	BioGPS	0.2%	1	1	1.087805E-4	0.0046157837	Zinc finger homeobox protein 3 OS=Mus musculus GN=Zfx3 PE=1 SV=1	0	1
	E9QMD3	BioGPS	0.2%			1.0886816E-4	0.0046157837	Zinc finger homeobox protein 3 OS=Mus musculus GN=Zfx3 PE=1 SV=1		

Export proteins in group: CSV | Excel | XML | PDF

Export each protein as individual: CSV | **Excel** | XML | PDF

Version: 4.0 | Logged in as: calzlw

© 2008 Integrated Proteomics Applications, Inc.

10.202.30.50/so2/searchResult.html?experimentId=1145863787067274-1&searchName=id-TMT-6&searchId=26788&id=141&d-2711775-e=2&sampleName=TMT-6%2B8BprojectName=chen+huahao&createdDate=2016-09-18+14%3AM%3A04.9

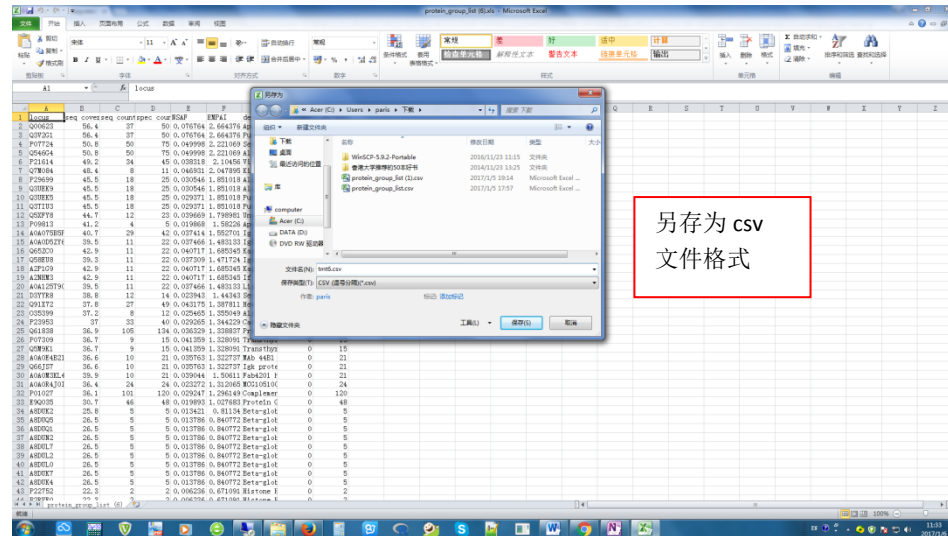
protein_group_list (5)...

360浏览器

下载完毕，文件安全

protein_group_list...

打开文件

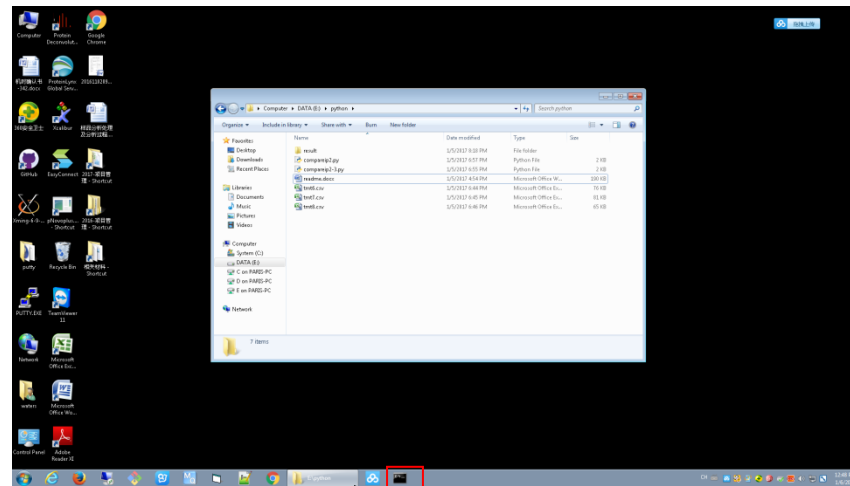


2、运行程序

远程桌面连接：10.200.6.248

目标文件夹：E: \python

✧ 将 csv 格式的文件拷贝到上述文件夹中



✧ 打开桌面任务栏的 cmd 对话框

✧ 比较 2 个数据

输入 `python compareip2.py 文件 1.csv 文件 2.csv`

输好后按回车键，结果在 `E:\python\result` 文件夹中，命名为 `compare2.csv`，用完后请删除该文件。

The screenshot shows a Windows Command Prompt window titled "Administrator: Command Prompt". The text inside the window is as follows:

```
Microsoft Windows [Version 6.1.7601]
Copyright (c) 2009 Microsoft Corporation. All rights reserved.

E:\python>python compareip2.py SHH116.csv SHH117.csv
Traceback (most recent call last):
  File "compareip2.py", line 9, in <module>
    df2=pd.read_csv(SHH117.csv)
NameError: name 'SHH117' is not defined

E:\python>python compareip2.py SHH116.csv SHH117.csv
E:\python>
```

Overlaid on the screenshot is a table with three columns: "Program name", "File1", and "File2". The first row contains the text "compareip2.py", "SHH116.csv", and "SHH117.csv" respectively. The second row is empty.

Program name	File1	File2
compareip2.py	SHH116.csv	SHH117.csv

✧ 比较 3 个数据

输入 `python compareip2-3.py 文件 1.csv 文件 2.csv 文件 3.csv`

输好后按回车键，结果在 `E: \python\result` 文件夹中，命名为 `compareip3.csv`，用完后请删除该文件。

```
Administrator: Command Prompt
Microsoft Windows [Version 6.1.7601]
Copyright (c) 2009 Microsoft Corporation. All rights reserved.

E:\python>python compareip2-3.py tmt6.csv tmt7.csv tmt8.csv
E:\python>
```

✧ 注意事项：手动输入

每个字段后面用一个空格符隔开

如果出现错误，可能是 **result** 里面有同名文件，删除后，重新输入就好

By- Pariswu

wingslovon@gmail.com