

iProX: an integrated proteome resource

Supplementary File

1. An example dataset at iProX

Please refer to project IPX0000951000 as an example of complete submission dataset at iProX, which is composed of the raw data and analysis results of the published paper (1).

The project URL is <http://www.iprox.org/page/PDV014.html?projectId=IPX0000951000>, and the details can also be found in the example dataset link in iProX (<http://www.iprox.org/page/exampleDataset.html>).

In this study, five analysis strategies were developed and compared for human plasma proteome profiling, and five sub-datasets were generated. The data from each sub-dataset is composed of a subproject in IPX0000951000. Figure 1 shows the overview information of this project. All data files of this project can be downloaded by clicking the "Aspera Download" or "Http download" button on this page. Five subprojects are shown at the bottom of this page (red rectangle in Figure 1). Taking the first subproject IPX0000951001 as an example, click the subproject name and jump to the page of subproject information (<http://www.iprox.org/page/SDV015.html?subprojectId=IPX0000951001>).

Figure 2 shows the overview information of this subproject. All data files of this subproject are shown in the bottom of this page (red rectangle in Figure 2A). Select one or several files and click the "Http" or "Aspera" button for downloading (Figure 2B); the http download pages may be blocked by the browser, please configure your web browser to allow pop-ups from iProX. By clicking the "View identification results" button on this page or clicking the 'view' button beside each standard format result file (Figure 2C), the protein and peptide identifications of the result files with standard format can be viewed in iProX (Figure 2D).

iProX integrated proteome resources

Home My Space Browse Search Submit Help

English Sign in Register

Search

Project Information

Project ID: IPX0000951000

Project Title: Human Plasma Proteome with Three-dimensional Strategies by High-resolution Quadrupole Orbitrap Mass Spectrometry

XML File: Pride_IPX0000951000.xml [Download](#)

Download All Files (101.13G): [Aspera Download](#) [Http Download](#)

Submission Type: Complete

Project Status: submit

Description: Five three-dimensional strategies combining HAP depletion (the 1st dimension) and protein fractionation (the 2nd dimension), followed by LC-MS/MS analysis (the 3rd dimension) were developed and compared for human plasma proteome profiling. Pros and cons of the five strategies are discussed for two issues: HAP depletion and complexity reduction. Strategies A and B used proteome equalization and tandem Seppro IgY14 immunodepletion, respectively, as the first dimension. Proteome equalization strategy A) was biased toward the enrichment of basic and low-molecular weight proteins and had limited ability to enrich low-abundance proteins. By tandem removal of HAPs (strategy B), the efficiency of HAP depletion was significantly increased, whereas more off-target proteins were subtracted simultaneously. In the comparison of complexity reduction, strategy D involved a deglycosylation step before high-pH RPLC separation. However, the increase in sequence coverage did not increase the protein number as expected. Strategy E introduced SDS-PAGE separation of proteins, and the results showed oversampling of HAPs and identification of fewer proteins. Strategy C combined single Seppro IgY14 immunodepletion, high-pH RPLC fractionation and LC-MS/MS analysis. It generated the largest dataset, containing 1544 plasma protein groups and 258 newly identified proteins in a 30-h machine-time analysis, making it the optimum three-dimensional strategy in our study. Further analysis of the integrated data from the five strategies showed identical distribution patterns in terms of sequence features and GO functional analysis with the 1505 plasma-protein dataset, further supporting the reliability of our plasma protein identifications. The characterization of 20 cytokines in the concentration range from sub-nanograms/milliliter to micrograms/milliliter demonstrated the sensitivity of the current strategies.

Project Tag: Plasma, Proteome profiling, IgY14 immunodepletion, Biomarker

Institution: National Center for Protein Sciences (Beijing), Beijing Proteome Research Center

Submitter:

Contact name	Contact email	Contact affiliation
Chang Cheng	1987ccpacer@163.com	National Center for Protein Sciences (Beijing), Beijing Proteome Research Center.

Lab Head:

Contact name	Contact email	Contact affiliation
Prof. Xiaohong Qian	qianxh1@163.com	National Center for Protein Sciences (Beijing), Beijing Proteome Research Center.

Publication:

Pubmed ID	Reference
26724764	Mining the Human Plasma Proteome with Three-dimensional Strategies by High-resolution Quadrupole Orbitrap Mass Spectrometry. Anal Chem Acta. 2016 Jan 21;904:65-75.

Subproject Count: 5

Access level: Public

Public Start Date: 2017-06-20

Groups: Chinese liver proteome project Group

Grants:

Agency	Agency Abbr	Grant ID	Grant Title
Ministry of Science and Technology of the People's Republic of China	863 programme	2012AA020201	Research and development of key technology for quantitative analysis of protein function network

Show All Information

Overall Rating: 4.5 Stars [view comments](#)

Show SubProject

Show 5 entries

Subproject Name	Subproject ID	PXID	Species	Instrument	File Count (size)	Release Date
Dataset generat...	IPX0000951001				70 (194.7G)	2017-06-20 09:44:54
Dataset generat...	IPX0000951002		Homo sapiens	Q Exactive	70 (194.7G)	2017-06-20 09:44:54
Dataset generat...	IPX0000951003		Homo sapiens	Q Exactive	70 (20.86G)	2017-06-20 09:44:54
Dataset generat...	IPX0000951004		Homo sapiens	Q Exactive	70 (21.16G)	2017-06-20 09:44:54
Dataset generat...	IPX0000951005		Homo sapiens	Q Exactive	70 (21.16G)	2017-06-20 09:44:54

Showing 1 to 5 of 5 entries

Evaluation (Please sign in for score and comments.)

Score: Not Rated

Comments:

[Submit](#)

Show Log

RequestTime	Operator	Reason	AgreeTime	SubmitTime	OperateType	Status
2017-06-20 09:44:54	ccpacer		2017-06-18 14:11:29	2017-06-20 09:45:51	Initialization	
2017-06-20 09:44:54	ccpacer		2017-06-20 09:45:51	2017-06-20 09:45:51	Submit project	Submit

Support by Beijing Proteome Research Center (BPRC)
京ICP备11036954号-3

Figure 1. Overview of the example project IPX0000951000 in iProX.

Subproject Information

Subproject ID: IPX000951001

Project ID: IPX000951000

Subproject Title: Dataset generated by Strategy A

Subproject Status: submit

Subproject Tag: Strategy A

Species: Homo sapiens

Tissue: blood plasma

Quantitation: Label free

Digestion: Trypsin

Experiment Type: Shotgun proteomics

Subproject Summary: Dataset generated by Strategy A: LAP Enrichment (Proteominer kit), In-solution Digestion and High-pH RPLC Separation (23 Fractions)

Access Level: Public

Public Start Date: 2017-06-20

Groups: Chinese liver proteome project Group

MS Instrument: Q Exactive

Experiment protocol: Human plasma samples were supplied by the China Human Liver Proteome Project (CNHLPP) sample bank. Seven healthy plasma samples were pooled and prepared for the subsequent experiment.

Informatics Protocol: All the raw files from Q-Exactive were converted into mgf files using the msconvert module in Trans-Proteomic Pipeline (TPP v4.5.2) and then searched against the UniProt human database (release201304_09601 entries) using the Mascot v2.3.2 local server. Precursor and product ion spectra were searched with an initial mass tolerance of 20 ppm and 0.05 Da, respectively. Tryptic cleavage was selected, and up to two missed cleavages were allowed. Carbamidomethylation on cysteine (+57.0203) was set as a fixed modification, and oxidation (+15.99Da) on methionine was set as a variable modification. The target-decoy-based strategy was applied to control both peptide and protein-level false discovery rates (FDRs) at lower than 1%. Based on the highly confident identifications mentioned above, the area under the extracted ion chromatograms (XICs) of each peptide was calculated as the peptide abundance using the label-free quantification function in SILVER. The protein abundance was defined as the sum of all unique peptide abundances. For normalizing the different protein abundances in different experiments, the protein IBAQ index was calculated and then median-normalized. All the further analyses were based on the normalized results.

DB search software name: Mascot DB search software version: 2.3.2

Searched database Type: SwissProt Searched database Version: release201304

Decoy database Type: Reversed(separated)

Quality estimation method: Other Peptide and protein-level false discovery rates (FD)

Type	CV	Accession	Name	Value
Fixed modifications	PRIDE	PRIDE:0000282	Fixed modification	Carbamidomethylation on cysteine (+57.02Da)
Fragment mass tolerance setting	PRIDE	PRIDE:0000161	Fragment mass tolerance setting	0.05 Da
Max. missed cleavages	PRIDE	PRIDE:0000074	Maximum Missed Cleavages Setting	2
Peptide mass tolerance setting	PRIDE	PRIDE:0000078	Peptide mass tolerance setting	20 ppm
Variable modifications	PRIDE	PRIDE:0000283	Variable modification	oxidation (+15.99Da) on methionine

File Size: 70 (19.41G)

[View identification results](#)

Show All Information

Click and jump to the identifications view page

Data File

5 records

File Name	File Type	Related File	File Size	Status	View
201300926_ZY_Human_pool7_LP_RP_10.dat	SEARCH	201300926_ZY_Human_pool7_LP_RP_10.raw	84.65M	✓	
201300926_ZY_Human_pool7_LP_RP_10.dat-pride-filtered.xml.gz	PRIDEXML	201300926_ZY_Human_pool7_LP_RP_10.raw	3.09M	✓	View
201300926_ZY_Human_pool7_LP_RP_10.raw	RAW	201300926_ZY_Human_pool7_LP_RP_10.dat;201300926_ZY_Human_pool7_LP_RP_10.dat-pride-filtered.xml.gz	770.56M	✓	
201300926_ZY_Human_pool7_LP_RP_11.dat	SEARCH	201300926_ZY_Human_pool7_LP_RP_11.raw	87.36M	✓	
201300926_ZY_Human_pool7_LP_RP_11.dat-pride-filtered.xml.gz	PRIDEXML	201300926_ZY_Human_pool7_LP_RP_11.raw	3.66M	✓	View

Showing 1 to 5 of 70 records

Show Log

Request Time	Operator	Reason	Agree Time	Submit Time	Operation Type	Status
2017-06-20 09:44:55	ccpacer			2017-06-18 14:12:48	Initialization	
2017-06-20 09:44:55	ccpacer		2017-06-20 09:46:36	2017-06-20 09:46:36	Submit subproject	Submit

Data File Http Aspera

5 records Search:

<input type="checkbox"/>	File Name	File Type	Relate File	File Size	Status
<input checked="" type="checkbox"/>	201300926_ZY_Human_pool7_LP_RP_10.dat	SEARCH	201300926_ZY_Human_pool7_LP_RP_10.raw	84.65M	✓
<input checked="" type="checkbox"/>	201300926_ZY_Human_pool7_LP_RP_10.dat-pride-filtered.xml.gz	PRIDEXML	201300926_ZY_Human_pool7_LP_RP_10.raw	3.09M	✓
<input checked="" type="checkbox"/>	201300926_ZY_Human_pool7_LP_RP_10.raw	RAW	201300926_ZY_Human_pool7_LP_RP_10.dat,201300926_ZY_Human_pool7_LP_RP_10.dat-pride-filtered.xml.gz	770.56M	✓
<input type="checkbox"/>	201300926_ZY_Human_pool7_LP_RP_11.dat	SEARCH	201300926_ZY_Human_pool7_LP_RP_11.raw	87.36M	✓
<input type="checkbox"/>	201300926_ZY_Human_pool7_LP_RP_11.dat-pride-filtered.xml.gz	PRIDEXML	201300926_ZY_Human_pool7_LP_RP_11.raw	3.66M	✓

Showing 1 to 5 of 70 records < 1 2 3 4 5 ... 14 >

(B)

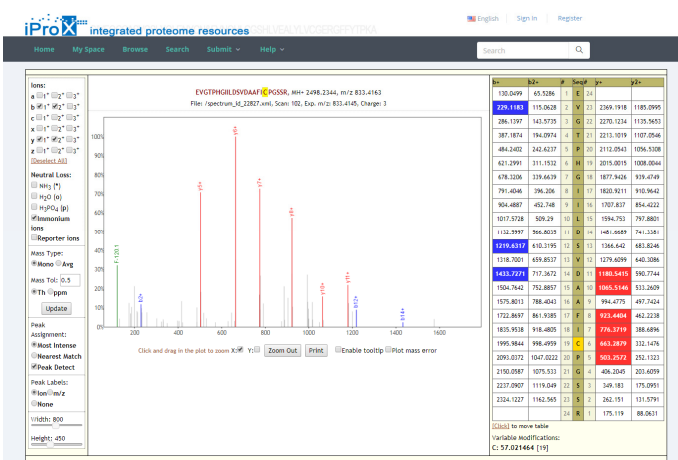
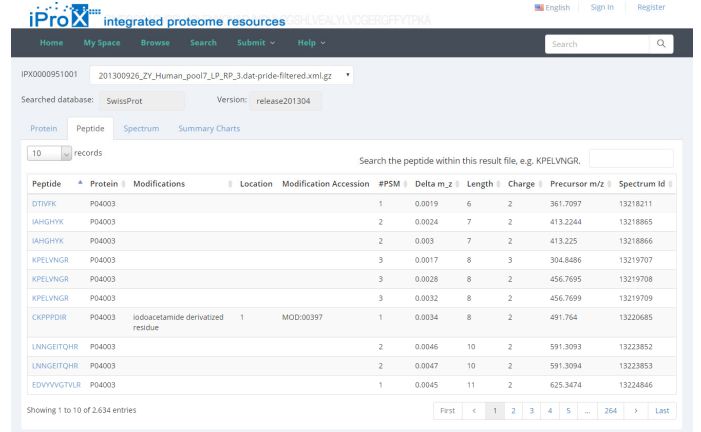
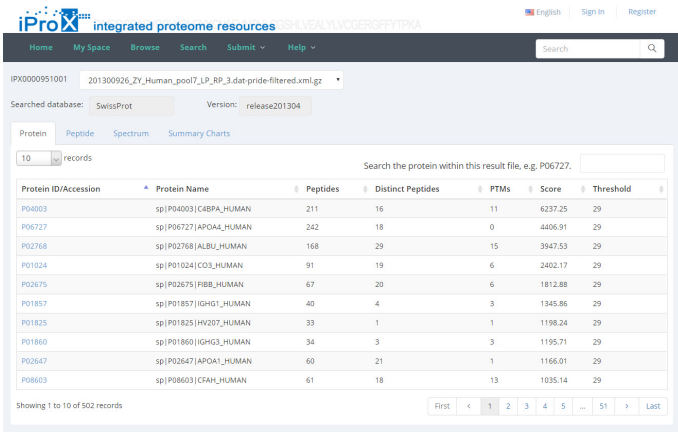
Data File Http Aspera

5 records Search:

<input type="checkbox"/>	File Name	File Type	Related File	File Size	Status	View
<input type="checkbox"/>	201300926_ZY_Human_pool7_LP_RP_10.dat	SEARCH	201300926_ZY_Human_pool7_LP_RP_10.raw	84.65M	✓	
<input type="checkbox"/>	201300926_ZY_Human_pool7_LP_RP_10.dat-pride-filtered.xml.gz	PRIDEXML	201300926_ZY_Human_pool7_LP_RP_10.raw	3.09M	✓	view
<input type="checkbox"/>	201300926_ZY_Human_pool7_LP_RP_10.raw	RAW	201300926_ZY_Human_pool7_LP_RP_10.dat,201300926_ZY_Human_pool7_LP_RP_10.dat-pride-filtered.xml.gz	770.56M	✓	
<input type="checkbox"/>	201300926_ZY_Human_pool7_LP_RP_11.dat	SEARCH	201300926_ZY_Human_pool7_LP_RP_11.raw	87.36M	✓	
<input type="checkbox"/>	201300926_ZY_Human_pool7_LP_RP_11.dat-pride-filtered.xml.gz	PRIDEXML	201300926_ZY_Human_pool7_LP_RP_11.raw	3.66M	✓	view

Showing 1 to 5 of 70 records < 1 2 3 4 5 ... 14 >

(C)



(D)

Figure 2. Overview of the example subproject IPX0000951001 in iProX.

2. Access of the shared dataset in iProX

In iProX, submitters can create a sharing URL with a password for peer-review for their private datasets, the validity is limited to 30 days, 90 days, 180 days or 360 days. As shown in Figure 3A, choose the private project in "My Space" ("Project Tree" of the user) and click the "Share" button. Then, a share URL with its password is generated in the pop-up box, and user can also select the valid period of the share link in this box (Figure 3B). Copy and paste the link into the browser (Figure 4A), enter the password, and you will see the shared project (Figure 4B). All data files can be downloaded on this page (red box in Figure 4B). Then click the subproject title and the subproject information is shown (Figure 4C).

Home
My Space
Browse
Search
Submit
Project
Group
Super Project
Help
Search

Project Tree

- ccpacer
 - Project
 - The MS data for accu...(2 subprojects)
 - The DDA data for acc...
 - The DIA data for acc...
 - Quantitative and In...(3 subprojects)
 - UPS2 only dataset
 - UPS2 yeast datasets
 - UPS2 mouse datasets
 - Mouse TAM proteomic...(1 subprojects) ↗
 - MS data for TAM and ...
 - Mining the Human Pla...(5 subprojects)
 - 20130614_ZY_Human_po...
 - 20130614_ZY_Human_po...
 - 20130804_ZY_Human_po...
 - 201300926_ZY_Human_p...
 - 201300929_ZY_Human_p...
 - Human Plasma Proteom...(5 subprojects)
 - Dataset generated by...
 - Dataset generated by...
 - Dataset generated by...
 - Dataset generated by...
 - Dataset generated by...
 - The MS data for prec...(2 subprojects)
 - MS data of Q Activ...
 - MS data of Orbitrap ...
 - Rat-liver(1 subprojects)
 - RNCD
 - Groups(2 groups)
 - Public Project

Project Information

Share
New SubProject
Edit
Delete

Project ID: IPX000782000

ProteomeXchange ID: none

Project Title: Mouse TAM proteomic analysis

XML File: PX_IPX000782000.xml Download

Download All Files (22.36G): Aspera Download *recommended
Http Download

Project Status: submit

Description: Comprehensive Proteomics Analysis Reveals Metabolic Reprogramming of Tumor-Associated Macrophages Stimulated by the Tumor Microenvironment

Project Tag: tumor-associated macrophages, proteomics, tumor microenvironment

Institution: National Center for Protein Sciences (Beijing)

Dataset ID	Dataset URL

Contact name	Contact email	Contact affiliation
Cheng Chang	1987ccpacer@163.com	National Center for Protein Sciences (Beijing)

Contact name	Contact email	Contact affiliation
Yunping Zhu	zhuyunping@gmail.com	National Center for Protein Sciences (Beijing)
Li Tang	tangli08@aliyun.com	National Center for Protein Sciences (Beijing)
Fuchu He	hefc@nic.bmi.ac.cn	National Center for Protein Sciences (Beijing)

Pubmed ID	Reference
	Comprehensive Proteomic Analysis Reveals Metabolic Reprogramming of Tumor-Associated Macrophages Stimulated by the Tumor Microenvironment.

Subproject Count: 1

Access level: Private

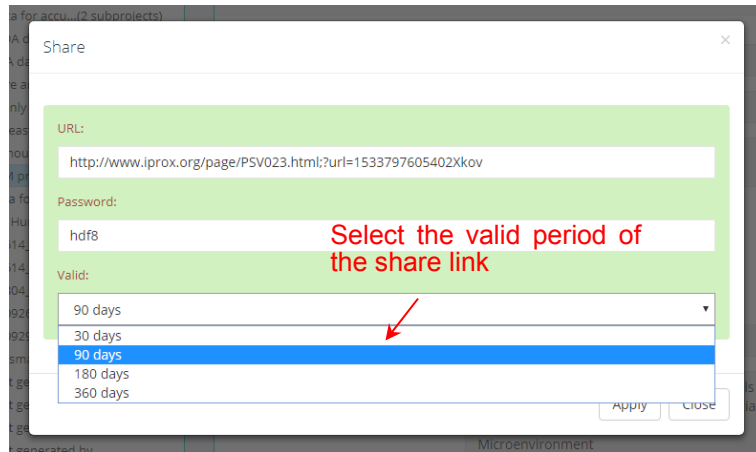
Public Start Date: N/A

Super Project: Super-project for NSFC project 21475150

Agency	Agency Abbr	Grant ID	Grant Title
National Natural Science Foundation of China	NSFC General Project	21475150	Study on Quantitative and Reliability Evaluation Strategy and Absolute Quantification Algorithm for Proteomics Based High Resolution Mass Spectrometry

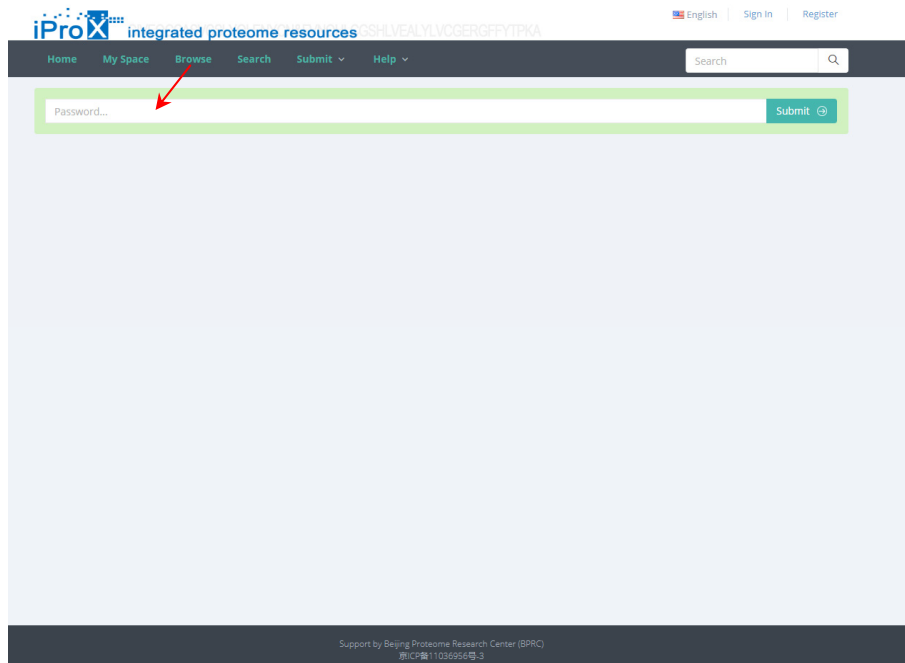
Overall Rating: ★ ★ ★ ★ ★ 4.6 Stars view comments

(A)



(B)

Figure 3. Creating a share URL for peer-review for the private dataset in iProX.



(A)

Project Tree

- Project
 - Mouse TAM proteomic...(1 subprojects)
 - MS data for TAM and ...

Project Information

Project ID: IPX0000782000

Project Title: Mouse TAM proteomic analysis

XML File: Pride_IPX0000782000.xml [Download](#)

Download All Files (22.36G)

- [Aspera Download](#) *Recommended
- [Http Download](#)

Project Status: submit

Description: Comprehensive Proteomics Analysis Reveals Metabolic Reprogramming of Tumor-Associated Macrophages Stimulated by the Tumor Microenvironment

Project Tag: tumor-associated macrophages, proteomics, tumor microenvironment

Institution: National Center for Protein Sciences (Beijing)

Submitter

Contact name	Contact email	Contact affiliation
Cheng Chang	1987ccpacer@163.com	National Center for Protein Sciences (Beijing)

Lab Head

Contact name	Contact email	Contact affiliation
Yunping Zhu	zhuyunping@gmail.com	National Center for Protein Sciences (Beijing)
Li Tang	tangli08@aliyun.com	National Center for Protein Sciences (Beijing)
Fuchu He	hefc@nic.bmi.ac.cn	National Center for Protein Sciences (Beijing)

Publication

Pubmed ID	Reference
	Comprehensive Proteomics Analysis Reveals Metabolic Reprogramming of Tumor-Associated Macrophages Stimulated by the Tumor Microenvironment.

Subproject Count: 1

Access level: Private

Public Start Date: N/A

Grants

Agency	Agency Abbr	Grant ID	Grant Title
National Natural Science Foundation of China	NSFC General Project	21475150	Study on Quantitative and Reliability Evaluation Strategy and Absolute Quantification Algorithm for Proteomics Based High Resolution Mass Spectrometry

Show All Information

Overall Rating: 4.6 Stars

(B)

iProX integrated proteome resources

English | Sign In | Register

Home | My Space | Browse | Search | Submit | Help

Search

Project Tree

- Project
 - Mouse TAM proteomic ... (1 subprojects)
 - MS data for TAM and ...

Click the subproject title to view the subproject information

Subproject Information

Upload | Download | Edit | Delete

Subproject ID: IPX000782001

Subproject Title: MS data for TAM and BMDM cells

Subproject Status: submit

Subproject Tag: Tumor-associated macrophages

Species: Mus musculus

Tissue: bone marrow-derived macrophage

Quantitation: Label free

Digestion: Trypsin

Experiment type: Shotgun proteomics

Subproject Summary: Here we used the proteome responses of the mouse BMDMs treated with tumor extract to demonstrate the relative differential effects of TAMs stimulated by the tumor microenvironment.

Access level: Private

Public Start Date: N/A

superProject: Super-project for NSFC project 21475150

MSInstrument: Orbitrap Fusion

Informatics Protocol: all Thermo raw data were directly processed using MaxQuant software (version 1.5.4.1).

File Size: 27 (22.36G)

Show All Information

Data File

5 records

File Name	File Type	Relate File	File Size	Status
6336_TangLi_MUS_liver_macrophage_sRP_10per_F1_R1.raw	RAW	null	990.77M	✓
6336_TangLi_MUS_liver_macrophage_sRP_10per_F2_R1.raw	RAW	null	971.31M	✓
6336_TangLi_MUS_liver_macrophage_sRP_10per_F3_R1.raw	RAW	null	919.41M	✓
6336_TangLi_MUS_liver_macrophage_sRP_10per_F4_R1.raw	RAW	null	934.57M	✓
6336_TangLi_MUS_liver_macrophage_sRP_10per_F5_R1.raw	RAW	null	993.98M	✓

Showing 1 to 5 of 27 records

Show Log

RequestTime	Operator	Reason	AgreeTime	SubmitTime	OperateType	Status
2017-04-22 14:29:59	ccpacer		2017-04-22 14:30:25	2017-04-22 14:30:25	Modify subproject information	Submit
2017-04-22 14:49:40	ccpacer		2017-04-22 14:50:01	2017-04-22 14:50:01	Submit subproject	Submit

(C)

Figure 4. Access to the shared dataset in iProX.

References

1. Zhao, Y., Chang, C., Qin, P., Cao, Q., Tian, F., Jiang, J., Li, X., Yu, W., Zhu, Y., He, F. *et al.* (2016) Mining the human plasma proteome with three-dimensional strategies by high-resolution Quadrupole Orbitrap Mass Spectrometry. *Anal Chim Acta*, **904**, 65-75.

Supplementary Figures

Supplementary Figure S1. The iProX controlled vocabulary viewer. This tool enables data submitters to browse or search for relevant ontology terms directly on the website. Several common values are provided as defaults. Also, users can search and preset their specific CV terms, which can be reused in different data submissions with similar experimental settings.

STEP 2 OF 4

- 1 Add New Project
- 2 Add New Subproject
- 3 Upload Data File
- 4 Submit

Add New SubProject

Subproject ID: IPX0001037001
Subproject ID will be generated automatically.

ProteomeXchange ID: [Empty]

Subproject Title: LM3 proteome dataset
Provide title of the subproject (required).

Subproject Tag: LM3
Provide key words of the subproject (recommended).

Species: Homo sapiens [Choose]

Tissue: LM3 cell [Choose]

Cell Type: [Choose]

Disease: [Choose]

Quantitation: Label free [Choose]

Digestion: Trypsin [Choose]

PTMs: Iodoacetamide derivatized residue [Choose]

Other information: [Empty]

Experiment type: Shotgun proteomics

Summary: The LM3 dataset was generated from a shotgun analysis of the metastatic human hepatocellular carcinoma cell line (HCCLM3) using Q-Exactive.

MS instrument: Q Exactive [more]

Experiment Protocol: [Enter text] [more]

Informatics Protocol: [Enter text] [more]

Buttons: Continue Save Skip Submit

Species Search (NCBI organismal classification)

Commonly used values CV terms Search

Commonly used values:

- Hepatitis C virus subtype 1a
- Homo sapiens
- Hordeum vulgare
- Ictalurus punctatus
- Macaca mulatta
- Merluccius bilinearis
- Mesocricetus auratus
- Mus musculus

Species Search (NCBI organismal classification)

Commonly used values CV terms Search

Homo sapiens [Search]

20 records

id	name	synonym_value	Choose	Add to list
NCBITaxon:9606	Homo sapiens	human,humans,man	[Choose]	[Add to list]
NCBITaxon:83221	Homo sapiens neanderthalensis	Homo neanderthalensis,Neandertal,Neandertal man,Neandertal,Neandertal man	[Choose]	[Add to list]
NCBITaxon:741158	Homo sapiens ssp. Denisova	Denisova hominin,Denisovan,Denisovans,Homo sapiens ssp. 'Denisova','Homo sp. Altai	[Choose]	[Add to list]
NCBITaxon:1131344	Homo sapiens x Mus musculus hybrid cell line		[Choose]	[Add to list]
NCBITaxon:1383439	Homo sapiens/Mus musculus xenograft		[Choose]	[Add to list]
NCBITaxon:1573476	Homo sapiens/Rattus norvegicus xenograft		[Choose]	[Add to list]

Showing 1 to 7 of 7 entries

Add the specific term to user's preset CV list