*my*ProMS

User's Guide

(02/03/18)

Table Index

Introduction

Pro eomic Mass Spec rome r (MS) genera es comple da a ha req ire m l iple s eps of comp a ional and man al processing o be ransla ed in o meaningf l biological informa ion. To be s ccessf l, his process req ires he skills of MS specialis s, bioinforma icians and biologis s. To facili a e s ch collabora ion, e ha e de eloped m ProMS (Po lle e al., 2007), a eb-based ool ha ra ionali es his da a processing orkflo hile allo ing m l iple sers o in erac i h he da a according o heir e per ise le el.

T picall , o p files from MS ools s ch as Masco or Pro eome Disco erer are impor ed in o m ProMS da abase i hin a defined e perimen al con e . Spec r m in erpre a ions, pro ein a rib ions and ariable modifica ion (e.g. phosphor la ion) posi ions can be alida ed b MS specialis s ei her a oma icall hro gh dedica ed algori hms or man all b is al inspec ion of each spec r m. Onl alida ed da a become accessible o biologis s for f r her in es iga ion. Differen q an ifica ion and differen ial anal sis me hods are pro ided hro gh in i i e in erfaces. Res l s are displa ed as in erac i e graphics o help sers is ali e and mine he da a. F r her biological in erpre a ion is possible sing in egra ed Gene On olog anal ses and e ensi e linking o e ernal reso rces.

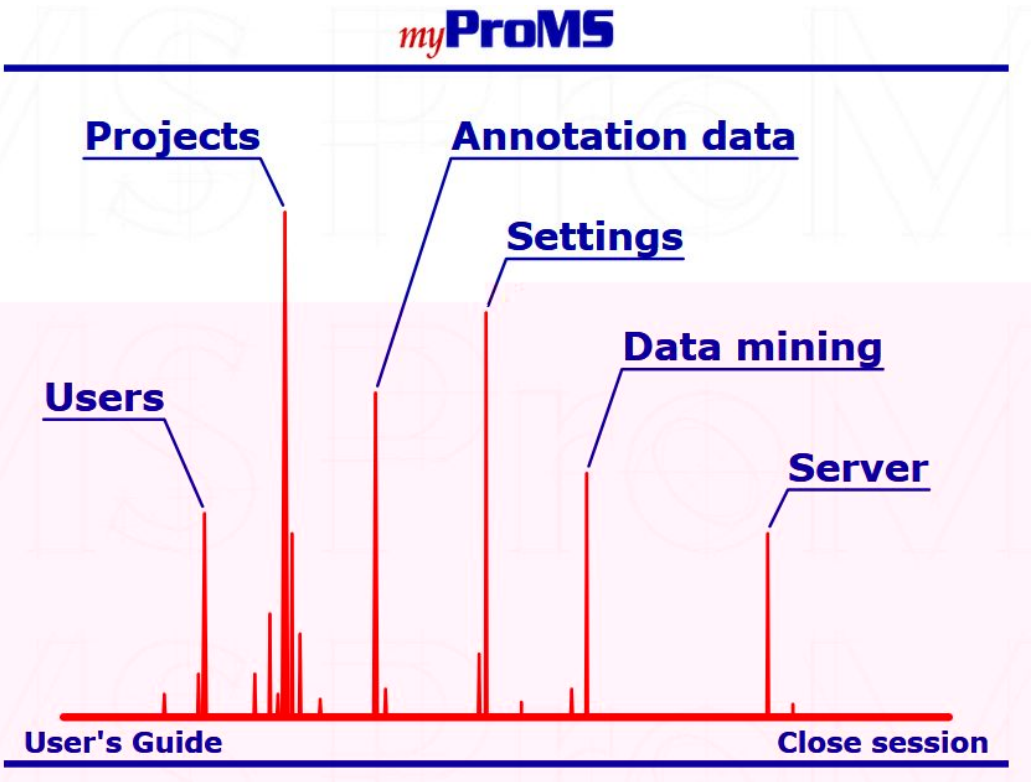
Connection to myProMS server

Login

Access o m ProMS da a req ires a login and pass ord. Con ac o r local m ProMS adminis ra or(s) o req es an acco n . Yo ill hen be able o login o he ser er b clicking on he S ar Session b on from he home page. D ring login o can choose be een ab- or f ll screen- displa modes.

Server main sections

Once logged o m ProMS, o r login, ser class (see he **Users management** chap er belo for more informa ion) and connec ion da e are displa ed a he op of he bro ser indo . Depending of o r ser class, ei her **myProMS main window** (Massis and bioinforma ician) or he **Project selection window** (biologis) ill be displa ed as sho n belo . Users (incl ding biologis s) can access he main indo a an ime b selec ing *Main Window* from he Projec selec ion indo .



The main indo displa s links o he 6 areas of m ProMS:

- Users managemen ,
- Projec s access and managemen ,
- Pro ein anno a ion da a managemen ,
- Se ings managemen ,
- *Da a mining sec ion,
- Ser er managemen ,

**Not yet available.*

Each of hese sec ions is described in a dedica ed chap er (see belo).

Some sec ions migh no be accessible sers depending on he access pri ileges. T picall , mos end sers ill ha e access onl o heir acco n and projec s. See **Users management** chap er belo for more informa ion.

Users management

User classes and access privileges

Da a access in m ProMS is igh l con rolled a he ser-le el o ins re da a pri ac and in egri . M l iple classes of sers are defined based on e per ise req ired o perform he differen da a/ sers managemen and processing asks a ailable in m ProMS. In addi ion, i hin cer ain classes, sers can be gran ed addi ional pri ileges if heir e per ise MS da a processing j s ifies i .

There are 4 classes of sers defined in m ProMS ordered b decreasing access pri ileges: bioinforma icians, massis s, da a managers and biologis s.

Bioinformaticians

This class of sers is in ended for ser er adminis ra ion and anno a ion da a managemen . Al ho gh bioinforma icians ha e f ll access o all f nc ionali ies of m ProMS, he sho ld no be sed o perform ro ine da a processing s ch as MS da a alida ion has he migh no ha e he necessar e per ise. We recommend o keep he n mber of bioinforma ician acco n s as lo as possible d e o heir e ended abili o modif he da a,

Massists

The massis class represen s MS e per s ho are in charge of MS da a impor , alida ion and repor ing. Massis s also manage ser acco n s and projec s crea ion. B defa l massis s ha e access o all m ProMS f nc ionali ies e cep hose normall dedica ed o bioinforma icians.

Data managers and workgroups

Da a managers ha e he same pri ileges as massis s b res ric ed o projec s and sers of heir orkgro p. Workgro p sage is op ional b is par ic larl sef l for m l iple MS-based research labs sharing a common MS facili . In ha case, a single ins ance of m ProMS i h a orkgro p a rib ed o each lab ill ins re da a pri ac hile main aining managemen cen ralisa ion b he MS facili .

Biologists

Biologis s are end sers of m ProMS. The ha e access o he projec s he par icipa e o i h ario s le els of pri ileges depending of heir e per ise and in ol emen in each projec . Projec access pri ileges for biologis s come in m l iple fla ors:

+ Projec in ol emen -based pri ileges:

- **Guest**: A g es ser can onl access he projec da a b canno modif hem.
- **User**: Can access and modif projec da a.
- **Administrator**: Same as a **user**. In addi ion, a projec Adminis ra or can gran o her sers access o projec .

+ E per ise-based pri ileges:

Biologis s (sers and adminis ra ors) can be gran ed par ial access o MS da a alida ion f heir kno ledge of he proced res in ol ed is j dged s fficien : d

- **Power** (User/Adminis ra or): Can er Va alida ion mode o alida a pro ed ide ifica ion da a onl .
- **Super** (User/Adminis ra or.

Projects

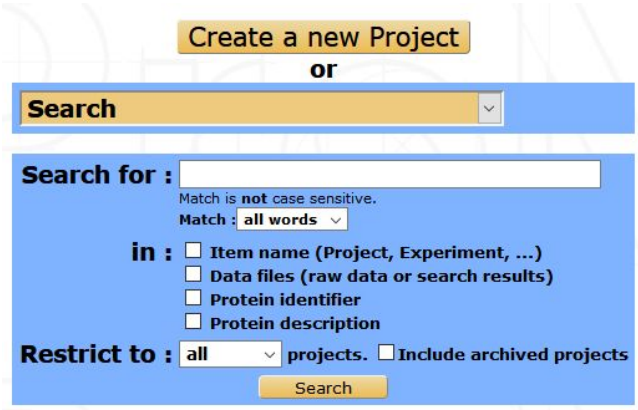
All MS search res l s and da a s bseq en l genera ed are organi ed in projec s. A projec regro ps se s of da a ha belong o he same ser or gro p of sers and genera ed in he con e of a defined scien ific projec . End- ser (biologis s and managers o side heir orkgro p) accessibili o he da a is defined a he projec le el.

Project selection

From m ProMS main indo , follo he **Projects** link o displa he **Project selection** in erface.

<Fig re Projec selec ion>

For biologis s, a s raigh for ard lis of he projec s he ha e access o ill be displa ed.
For o her classes of sers, projec s can be organi ed based on he follo ing opics: *Workgroups* (defa l), *Project owners*, *Active projects* and *Archived projects*. Corresponding projec s are lis ed b ascending name. Their descrip ion, o ner and/or orkgro p oge her i h he access creden ials of c rren ser.
If *On-going analyses* is selec ed, he lis genera ed is composed of Anal ses s ill ndergoing alida ion. Their name, descrip ion, da a file name, crea ion da e and corresponding projec are displa ed. This lis can be sor ed b Impor da e, Name, MS pe, Valida ion s a s and Da a file name.
Al erna i el , a *Search* can be performed sing ario s cri erias:

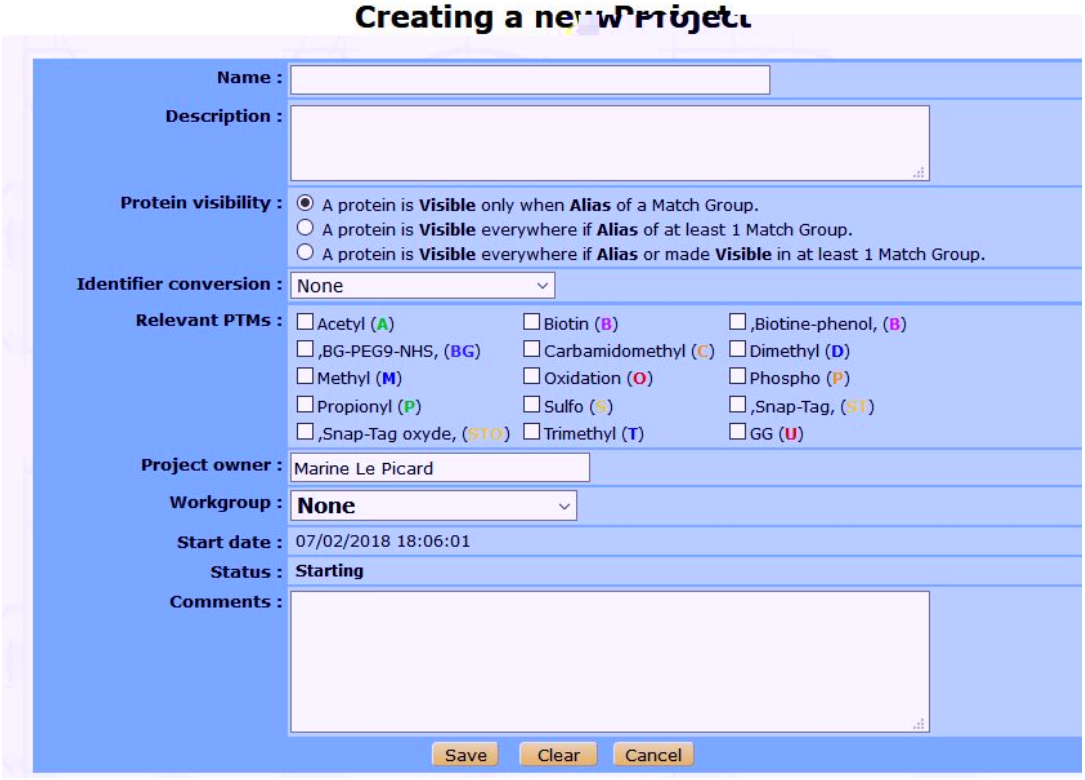


Projec s are hen lis ed oge her i h he i ems ha ere ma ched d ring he search.
Once lis ed, click on he *Open* b on corresponding o he projec an ed o access i .

Project creation and settings

Onl bioinforma icians, massis s and da a managers can crea e projec s.

From he **Project selection** in erface click on he *Create a new Project* b on. The follo ing form is hen displa ed:



- **Name**: Pro ide a manda or name for he projec .

- **Description**: an op ional descrip ion for he projec .

- he projec - ide pro ein isibili r le o be sed. See **Match groups** and **visibility** belo for de ailed informa ion on his concep .

- **Comments**: an op ional commen s for he projec .
Click on he *Save* b on o crea e projec .
Projec s can be edi ed a an ime o modif an of hese se ings.

Project life span

- **On-going**: Once crea ed, a projec is se as **active** and **on-going**. This means ha i can be pop la ed i h ne i ems and da a. On-going projec s are flagged i h a ello icon in he projec selec ion indo .
- **Ended**: If he projec is j dged comple ed, i can be edi ed and **ended** b clicking on he *End* b on a he bo om of he edi ion form. Ending a projec ill a oma icall end all par iall - alida ed anal ses i ho ne repor ing (see **Validations** and **Reporting** sec ions in he **Analysis management** chap e**r** belo for more informa ion). Once ended, a projec is s ill ac i e and accessible b can no longer be edi ed or pop la ed. Ended projec s are flagged i h a green icon in he projec selec ion indo .
- **Archived**: As ime passes, some projec migh no longer be accessed b an sers. These projec s can be archi ed o sa e space on he ser er. All da a files s ored o side he da abase ill be compressed. Archi ed projec s are flagged i h a red icon and are no longer accessible for da a e plora ion. The can ho e er be lis ed in he Projec selec ion indo b selec ing *List of: Archived projects* .
- **Restoration**: Archi ed and Ended projec s can be f ll res ored o an ac i i s a e if necessar b clicking on he appropria e b on in he projec home page.

Accessibility

Bioinforma icians and massis s ha e f ll access o all projec s recorded in m ProMS. Da a managers ha e f ll access o all projec s i hin heir orkgro p. Biologis s and managers o side heir orkgro p m s be e plici l gran ed access o projec s hen needed. The projec access managemen in erface is accessible from he projec s home page b clicking on he *Project Accessibility* b on in he **option frame**.

The in erface s mmari es he lis of sers able o access he projec oge her i h heir creden ials. Ne sers can be added one a a ime. Once added o he access lis , sers are g es s b defa l . Selec he creden ials o ish o pro ided each ser i h. The access righ s a ailable are lis ed belo he ser access form. See also **User classes and access privileges** abo e for more informa ion.
Click on he *Save* b on o alida e an changes.

Project navigation

Navigation frame

Sub-navigation frame

Option frame

Results frame

Project organization

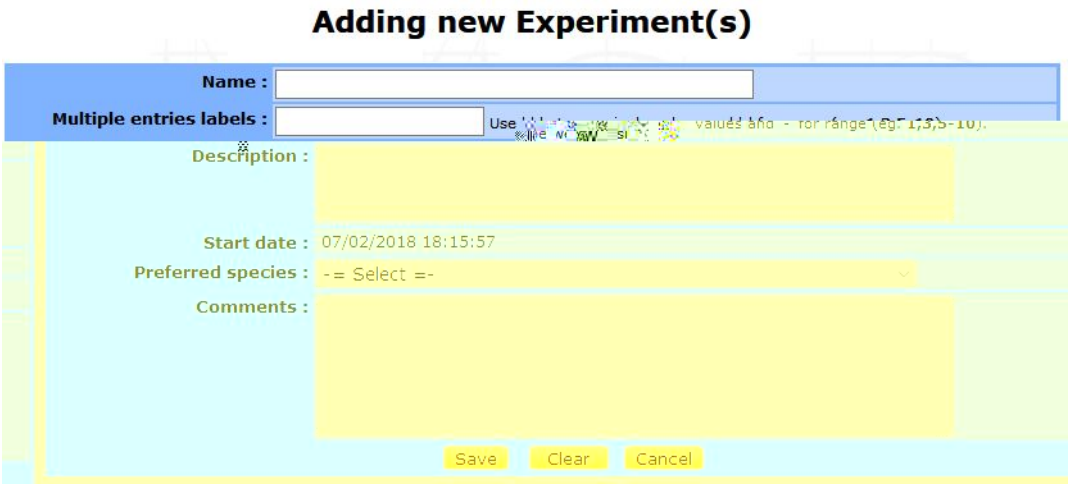
Da a in a projec are hierarchicall organi ed as sho n in he fig re belo :

<Fig re Projec hierarch >

Experiments

An **Experiment** i em represen s an ac al biological e perimen for hich MS da a ill be collec ed.

To crea e a ne e perimen , selec he projec elemen in he op lef na iga ion frame and click on *Add Experiment(s)* in he op ion frame.

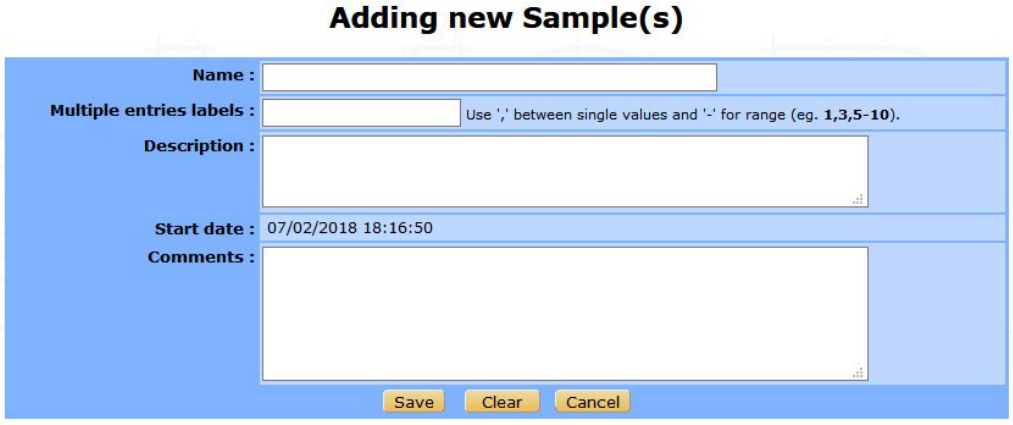


Pro ide a name and op ional descrip ion and/or commen s. M l iple e perimen s can be crea ed a once if he field M l iple en ries labels is filled in. Labels defined in his field ill be seq en iall appended o he name each e perimen crea ed. Labels can be defined indi id all sing a comma-separa ed s ring (eg. A,D,G) or a range s ring sing a - (eg. 1-5);

Samples

A **Sample** i em is a loose en i ha can represen a single or m l iple mi ed (e.g. for labelled q an ifica ion) biological samples. I can be ie ed as a s b-e perimen or Anal sis-con aining i em. I is p he ser o define i s f nc ion depending on he e perimen al con e of he anal ses i con ains.

To crea e a ne sample, selec i s paren e perimen in he na iga ion frame and click on *Add Sample(s)* in he op ion frame.



Pro ide a name and op ional descrip ion and/or commen s. M l iple samples can be crea ed as described for e perimen s (see he **Experiments** paragraph abo e).

Analyses

An **Analysis** corresponds o a da ase impor ed from a single search engine res l file: mos l he MS/MS spec ra (e cep for PMF r ns), he pep ide/pro ein iden ifica ions and associa ed q an ifica ions hen presen in he file. Anal sis da a m s be impor ed, alida ed and repor ed before end sers can access hem and f r her process heir res l s. These proced res are described in he chap er **Analysis data import and validation** belo .

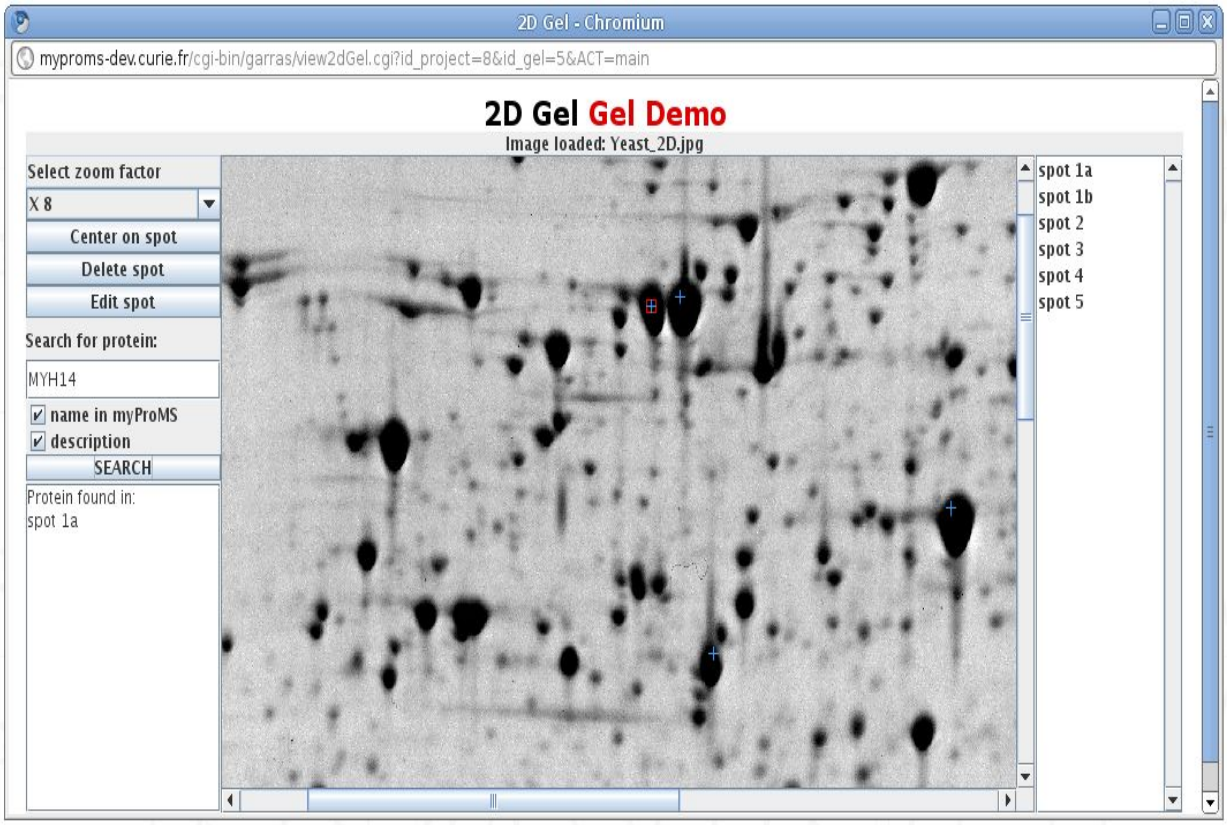
2D gels and spots

2D-Gel (T o-dimensional gel elec rophoresis) can be recorded and is ali ed in m ProMS. A pic re of he gel (JPEG forma onl) m s be ploaded in o m ProMS so as o keep a is al record.

To crea e a 2D-gel, selec i s paren e perimen in he pper na iga ion frame and click on *Add 2D-Gel* in he op ion frame. The form sho n belo ill be displa ed.

Pro ide a name, op ional descrip ion/commen s and selec he JPEG image file of he gel o be ploaded. Click on *Save* o s ore he ne 2D-Gel elemen in m ProMS.

To is ali e a 2D-gel, selec he gel elemen in he pper lef na iga ion frame and click on *Display 2D Gel* in he op ion frame. A pop- p indo ill displa he gel image as sho n belo (JAVA m s be enabled in o r bro ser and o m s accep he sec ri iss es).

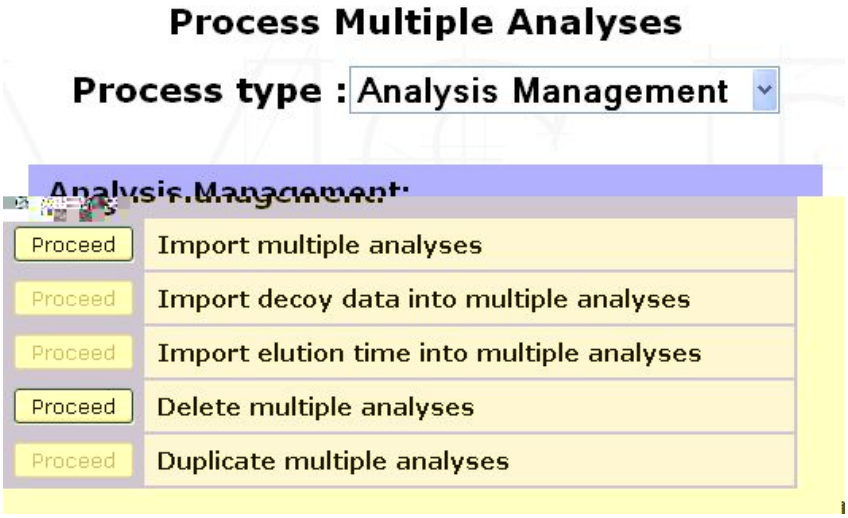


To record a ne spo , do ble click on he gel image here he spo is loca ed. se he pop- p bo o pro ide a manda or name and o her op ional informa ion (isoelec ric poin , molec lar eigh , e ernal iden ifier if o se ano her image processing sof are like ImageMas er,...). A bl e-cross is hen displa ed o represen he spor recorded. Yo can selec / nselec a spo b clicking on he corresponding cross. When selec ed, a spo can be edi ed or dele ed. Pro ein iden ifica ion da a can be appended o a spo b linking his spo o an e is ing **Sample**. This sample ill no longer be lis ed in he na iga ion indo and iden ifica ion da a ill become accessible hro gh he spo onl . When mo sing-o er a spo i h linked iden ifica ion da a, he op-pro ein (bes iden ifica ion in he associa ed anal sis) is displa ed is he spo pop- p bo .
Yo also can search a pro ein of in eres (in he abo e e ample, MYH14) and spo (s) con aining his pro ein ill be highligh ed i h a red sq are.
In addi ion o he graphical is ali a ion of he gel, all spo s are lis ed in he lo er na iga ion frame once a 2D-gel is selec ed. Each spo can be edi ed/dele ed from he s al in erface.

Search results data import: MS Analysis

The collec ion of spec ra/pep ides/pro eins(/q an ifica ion) da a con ained in a search res l file are impor ed in o m ProMS as an **Analysis**. Onl bioinforma icians, massis s and managers can impor Anal ses.

Selec he **Experiment** or **Sample** or **2D Gel** in o hich he Anal ses m s be impor ed and click *Process Analyses* in he op ion frame. From he selec ion men displa ed, selec *Analysis Management* o displa he lis of a ailable op ions.



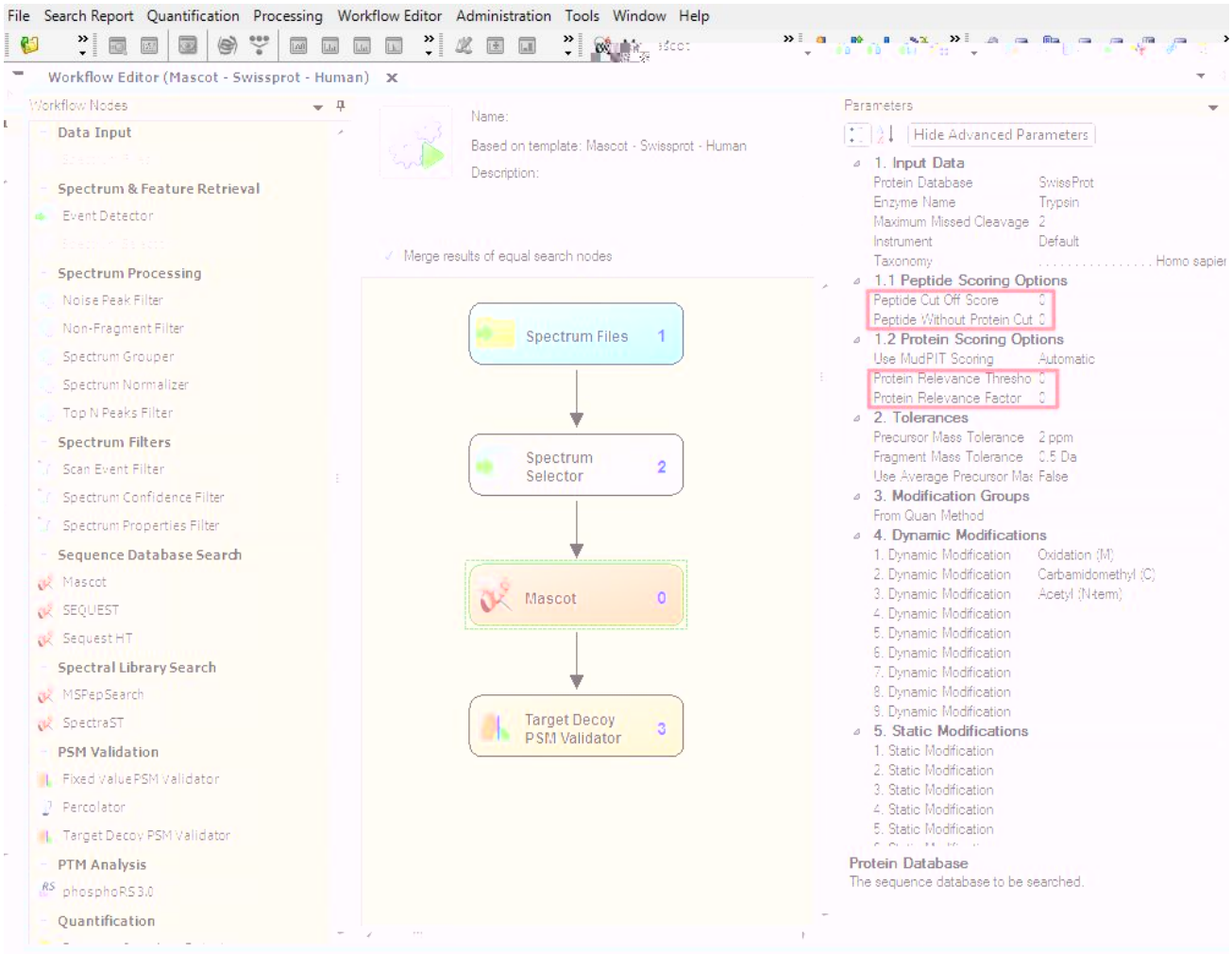
Supported search engines

m ProMS allo s o impor from ario s search engines :

- Masco (DAT file or MSF file from Pro eome Disco erer Sof are b Thermo Scien ific).
- Paragon (XML genera ed from Pro einPilo ™ Sof are b AB SCIEX, gro p2 ml.e e).
- Seq es (MSF file from Pro eome Disco erer Sof are b Thermo Scien ific).
- Phen (XML file genera ed hro gh Phen pla form b GeneBio). DEPRECATED!
- Andromeda/Ma Q an (mqpar. ml and 3-4 files are req ired).
- X!Tandem (XML file from X! Tandem pipeline (PAPPSO) or from Trans-Pro eomic Pipeline (TPP)).
- PeakVie (E por ed E cel XLXS file for SWATH da a).
- OpenS a h (TSV file from OpenSWATH orkflo).
- Spec rona (TSV file genera ed from Spec ra no ™ b Biognos s).

Important note: if o perform Masco searches i h Pro eome Disco erer (PD) Sof are, make s re o di no se Pro ein and Pep ide fil ers. In he **Workflow Editor**, click on Masco node and hen, se he fo r fil ers **Peptide Cutt Off Score**, **Peptide Without Protein Cut**, **Protein Relevance Threshold** and **Protein Relevance Factor** o 0. If hese fil ers are no rned-off m ProMS impor op ions s ch as predefined False Disco er Ra e (FDR) ill no be

acc ra e.



Collecting search files

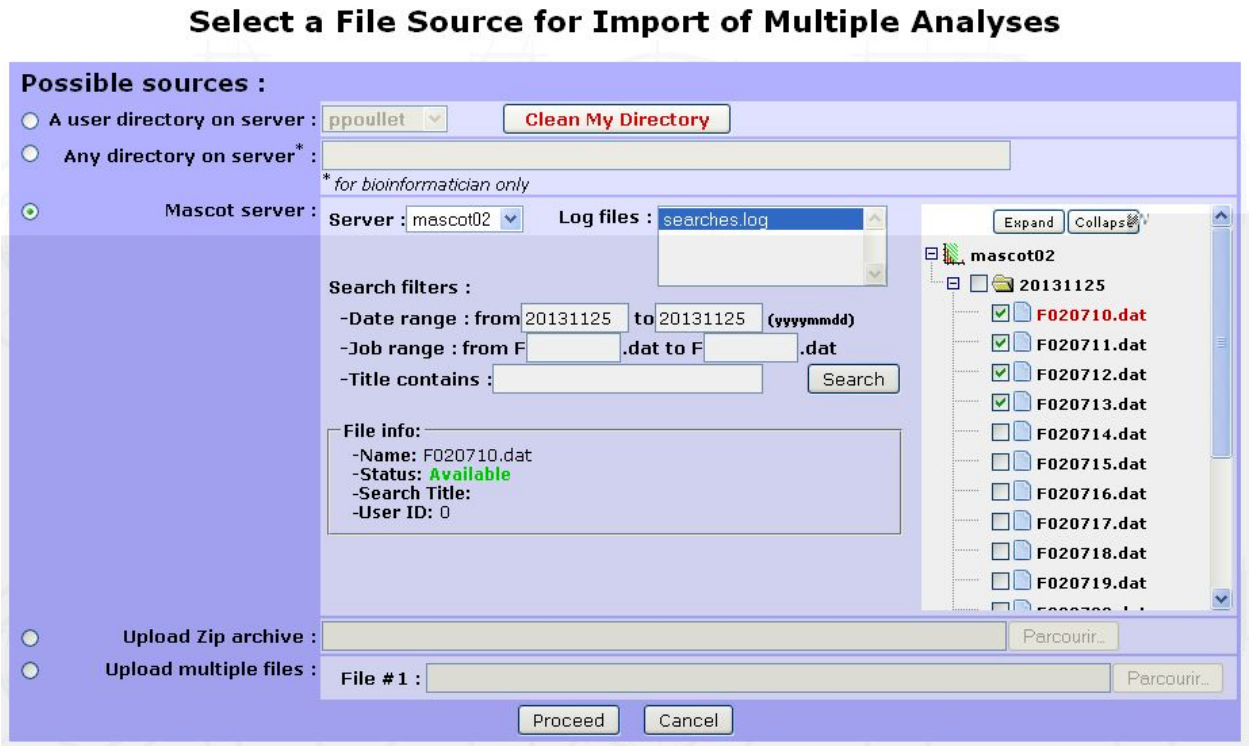
M l iple **Analyses** can be impor ed a once as long as he corresponding searches ere performed i h he same search engine and pro ein da abank(s).

Selecting data files (Mascot, Proteome Discoverer and X!Tandem)

Click on *Proceed* ne o he **Import multiple analyses** process as sho n belo . The follo ing form ill be displa ed o selec he so rce of he search files o be impor ed.

M l iple impor so rces are a ailable:

- **A user directory on server**: Follo ing pload, files are s ored in a ser-dedica ed direc or on ser er. These files ill s a a ailable o he ser n il he decides o dele e hem; ei her j s af er impor or la er. In he la er case he ser can s ill access his direc or for file managemen p rpose b clicking on he ***Clean My Directory*** b on.
- **Any directory on server**: This op ion is a ailable o bioinforma icians onl . The ser can pro ide an pa h on he ser er here m ProMS sho ld look for search res l s files.
- **Mascot server**: If a Masco ser er is declared in m ProMS config ra ion file, i can be accessed, searched for specific search res l s and files direc l ploaded in o m ProMS ser direc or .



User can search res l s files b **date** or **job number** range or ke ords in he files **search title**. The lis ma ching files is hen displa ed and gro ped b da of crea ion. Specific informa ion on a file (name, a ailabili , search i le and ser ID) can be displa ed b clicking on he file name. If access o Masco is res ric ed (ser acco n s se p), he Masco serID m s be defined in

m ProMS as ell (see Acco n managemen sec ion abo e). In his case, onl Masco files accessible o he ser ill be displa ed.

- **Upload Zip archive**: If a large n mber of files m s be impor ed, he can be ploaded a once as he ip archi e. The archi e ill be n ipped on he ser er.
- **Upload multiple files**: Al erna i el , p o 10 files can be ploaded as s para e files.

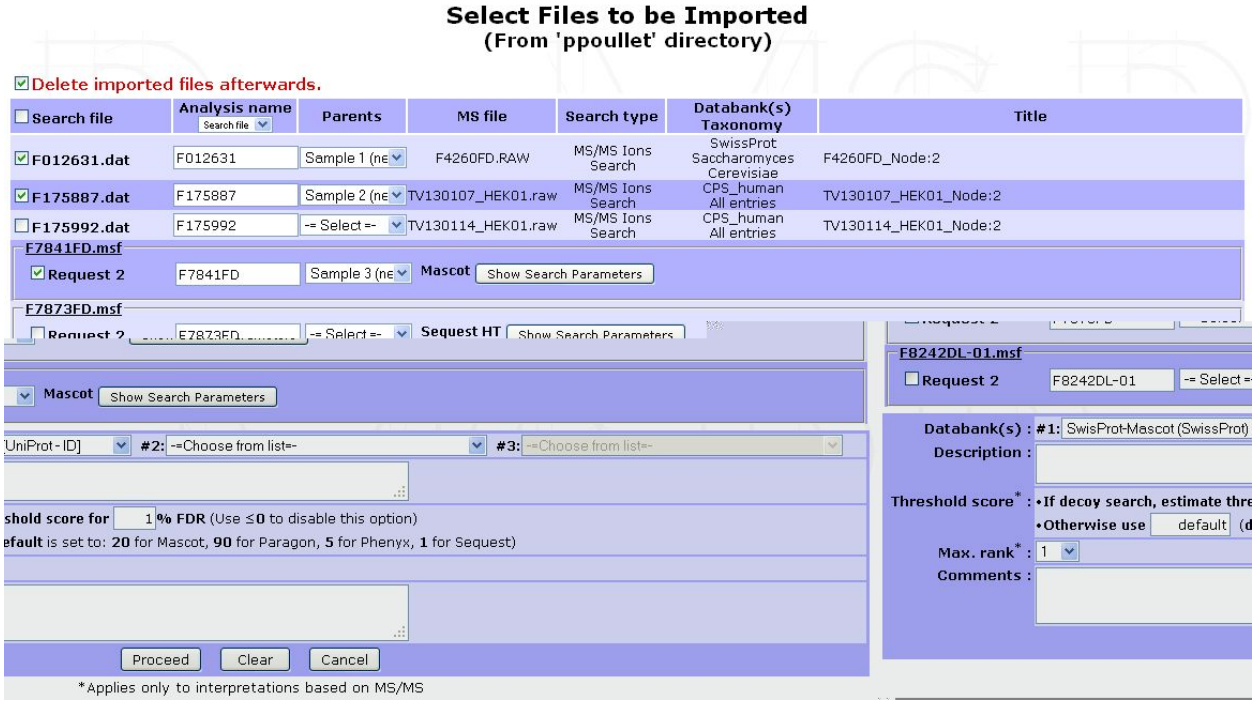
Once o r files ha e been selec ed, click on *Proceed* o ini ia e file re rie al from he selec ed so rce. This proced re ma ake a fe min es depending of he n mber and si e of he files. Once he ransfer is comple e, a file impor in erface ill be displa ed.

Important Note: Mos bro sers do no s ppor pload of files i h (o al) si e > **2 Gb**. If files larger han 2 Gb m s be ploaded, e recommend o se **Google Chrome**. This limi a ion does no appl hen re rie ing files direc l from a Masco ser er.

Importing analyses

Import parameters (Mascot, Proteome Discoverer and X!Tandem)

Files re rie ed are lis ed in alphabe ical order oge her i h per inen informa ion abo he search performed: he MS file, search pe (MS2, MS1 or mi of bo h), da abank(s) and a onom and search i le sed. See fig re belo .



Note: The same Pro eome Disco erer msf file can con ain m l iple searches res l s (e.g. a search performed i h Masco and ano her i h Seq es). In his case, separa e en ries ill be lis ed for each search performed oge her i h more de ails on he parame ers sed. Each search res l can be impor ed separa el and dis inc Anal sis i ems ill be crea ed.

Proceed as follo o con in e da a impor :

1. Selec he files o be impor ed from he lis b checking he bo es on he lef -hand side of he files name.

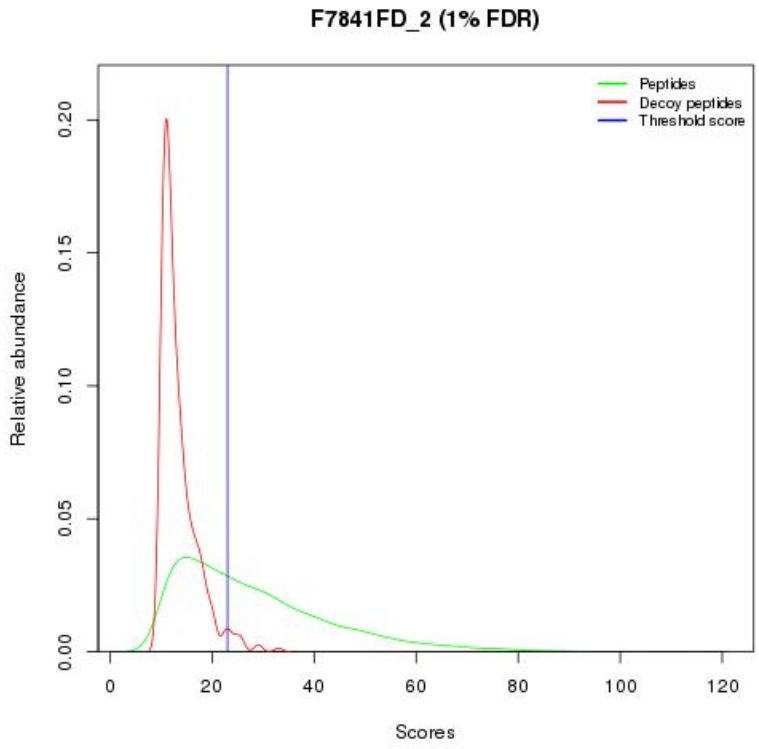
2. Pro ide a name for he Anal sis o be crea ed (**Analysis name** col mn). This name can be ped b he ser or se o ma ch ei her he name of he search file or ha of he original MS file sed for he search.

3. If he paren i em as an E perimen (or a 2D Gel), each ne Anal sis m s be associa ed i h a specific Sample (or Spo). Pre-e is ing Samples (or Spo s) can be selec ed from a dropdo n men or he can be crea ed on he fl (Samples onl): To crea e a ne Sample, selec Ne from he dropdo n men of **Parents** col mn. A pop p indo ill ask o o pro ide a name for he ne Sample.

4. Selec he da abanks o be sed o e rac pro ein anno a ions (**Databank(s)** field). If m l iple da abanks ere sed d ring he search (possible i h Masco for ins ance), he corresponding n mber sho ld also be selec ed here. All search files o be impor ed in he same ba ch sho ld ha e been performed sing he same or eq i alen (se of) da abank(s).

5. Define a fil ering r le for he da a o be impor ed (**Threshold score**):

If a deco search as performed, da a can be fil ered o based on a ser-defined **False Discovery Rate (FDR)** on pep ide iden ifica ion (defa l is 1%). A hreshold score for pep ide iden ifica ion ill be de ermined so ha he da a impor ed ill (en a i el) ma ch he defined FDR al e as ill s ra ed in he fig re belo . Threshold score calc la ion can ei her ses he **qvality** algori hm (K ll e *al.* Bioinforma ics 2008,25(7)), he **Mayu** algori hm or he **DT count** algori hm. In DT co n mode, deco (D) and arge (T) pep ides are sim l aneo sl co n ed in descending score order n il he propor ion of he 2 pop la ions ma ches he selec ed FDR al e.



If no deco search as performed or he FDR al e as se eq al or less han 0, he fil ering ill be performed according o a minim m (**threshold**) score for pep ide iden ifica ion. A defa l (search engine-specific) hreshold score ill be applied nless a differen one is pro ided b he ser.

6. Selec he ma im m n mber of in erpre a ions allo ed of he same fragmen a ion spec r m (**Max. rank**). The defa l is 1, b p o 10 can be chosen.

Note: When performing FDR fil ering, i is recommended o se his al e o 1 since he FDR calc la ion is based on 1 in erpre a ion per spec r m.

7. Pro ide op ional **description** and **comments**.

8. Decide he her he files sho ld be dele ed af er impor or no (**Delete imported files afterwards**). Unless selec ed for dele ion, file ill remain on he ser er for ne impor n il he ser decides o dele e hem.

9. Click on *Proceed* o ini ia e he da a impor in o m ProMS da abase.

Importing MaxQuant data

3 o 5 files are req ired o impor a Ma Q an search/q an ifica ion in o m ProMS:

1. mqpar. ml (s all loca ed in he roo direc or of he Ma Q an search),
2. e idence. (file from he Combined/ direc or),
3. Pep ides. (idem),
4. pro einGro ps. (idem. Op ional, onl o impor pro ein q an ifica ion da a),
5. msms. (idem. Op ional, onl o displa pep ide fragmen a ion spec ra).

Files 2 o 5 m s be compressed in a common archi e before impor .
Selec an E perimen in hich o impor he da a. From he *Process Analyses* indo , selec ei her he **Import multiple Analyses** or I**mport quantification** processes and click on *Proceed* ne o **Import MaxQuant quantification** o displa he form belo .

<Image>

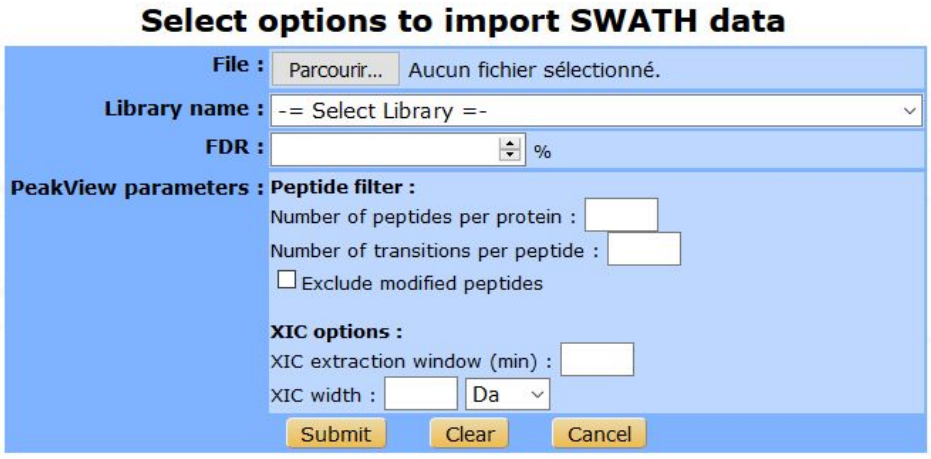
Pro ide he files men ioned abo e and he pro ein seq ence da abank(s) sed for he search. If con aminan s ere searched, pro ide a ma ching con aminan da abank. Finall , selec he r le o ish o se for pro ein aggrega ion in o ma ch gro ps: m ProMS or Ma Q an . Specifi also if o ish o impor pro ein q an ifica ion da a (he pro einGro ps. file m s be pro ided in he archi e in ha case). S bmi he form o s ar da a impor . Da a impor ill ake a fe min es. Samples, Anal ses and an e perimen al Design, a pep ide q an ifica ion and 1 o m l iple pro ein q an ifica ions ill be added o he selec ed E perimen according o he informa ion e rac ed from he files ploaded. Pep ides and pro eins ill be a oma icall alida ed since he ere sed in pro ein q an ifica ion.

Importing DIA quantification data

User can impor DIA q an ifica ion da a from hree differen sof are : PeakVie , OpenSWATH and Spec rona . In o m ProMS selec an E perimen and from he Process Anal ses indo , selec he **Analysis quantification** process and click on Proceed ne o **Import PeakView/OpenSWATH/Spectronaut data** o displa he associa ed form.

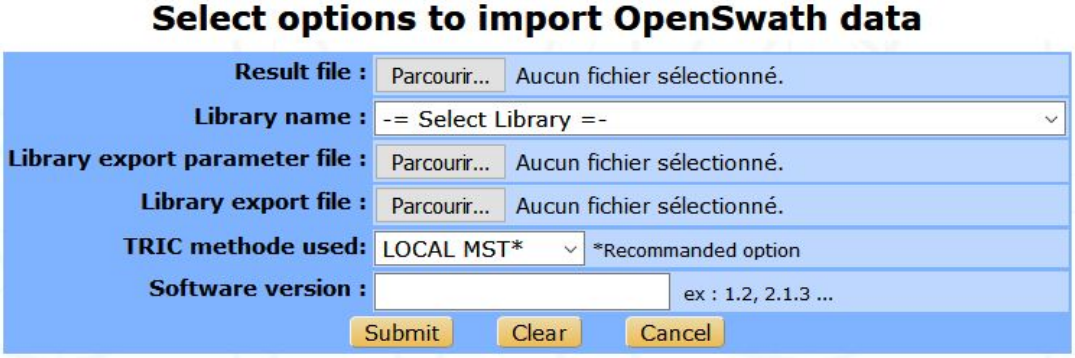
- *From PeakView*

T o files are req ired : he E cel orkshee file genera ed b PeakVie and he spec ral librar . In o PeakVie , once he e perimen al SWATH da a anal sis is o er, o can e por he res l in o an E cel file b clicking on he **Quantitation** ab on he oolbar and selec ing **SWATH Processing/Export/All**.
The PeakVie search parame ers can be filled in he follo ing form o be sa ed in m ProMS da abase o ens re raceabili . Then s bmi he form o la nch da a impor .



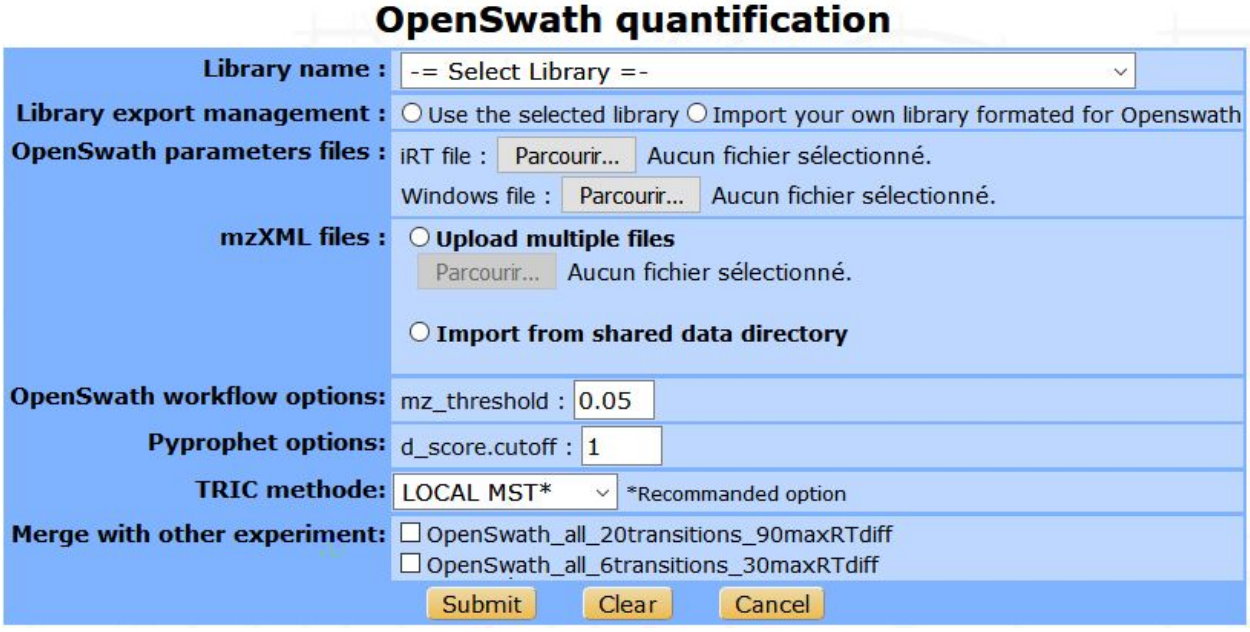
From OpenSWATH

The file from OpenSWATH is req ired, his is a TSV file genera ed b he las s ep of he orkflo (TRIC). The librar ha ill be sed o anal se he e perimen al da a is also needed as he associa ed e por parame er file. Then, he ser m s pro ide he TRIC me hod sed and he n mber ersion of OpenSWATH.



Running OpenSWATH quantification

The OpenSWATH orkflo can be la nched direc l from m ProMS. The process ill anal se he e perimen al files and impor he res l in a same s ep. Some parame ers are req ired as he librar name, he m XML res l s files, he iRT file (in TraML forma) and he DIA indo s file. The ser can impor his o n librar con er ed for OpenSWATH, and choose o merge his res l i h o her e is ing anal sis.



From Spectronaut

The impor Spec rona da a form is similar o OpenS a h s form.

Select options to import Spectronaut data

Result file :	Parcourir... Aucun fichier sélectionné.
Library name :	-= Select Library =-
Library export parameter file :	Parcourir... Aucun fichier sélectionné.
Library export file :	Parcourir... Aucun fichier sélectionné.
Software version :	ex : 1.2, 2.1.3 ...

Submit Clear Cancel

Once he process is o er, samples and anal ses ill be added o he selec ed E perimen . Pep ides and pro eins ill be a oma icall alida ed. Transi ion q an ifica ion da a ill be impor ed b no he pep ide or he pro ein q an ifica ion res l s as a dedica ed pipeline is a ailable in m ProMS o perform his ask (see he **Protein quantification** chap er belo for more informa ion).

Analysis summary

Files ill be impor ed seq en iall and he Anal ses (and possibl Samples) ne l genera ed ill appear in he op lef na iga ion indo . The Anal sis s mmar is sho n belo :

Analysis F7228FD

Name :	F7228FD
Description :	
Position :	1
Start date :	12/12/2012 14:29:05
MS type :	MS/MS Ions Search Export Elution Data
Search engine :	Mascot from Proteome Discoverer
Search file :	F7228FD_2.pdm Show Search Parameters
MS data file :	F7228FD.RAW
Databank(s) :	•SwisProt-Mascot **Type:** UniProt - ID
Taxonomy :	Homo sapiens (human)
Labeling :	None
Max. rank :	1
Min. score :	20.52 [FDR-based] Show Score Distribution
False discovery :	**Targeted FDR:** 1% - **Observed FDR:** [Data not reported: No peptides validated.] **Decoy method:** Automated decoy search.
Status :	**Not validated.** Show validation history
Comments :	

Di erse informa ion is a ailable o he ser s ch as he MS pe, search engine, da abank(s) (selec ed in m ProMS), pro ein iden ifier pe, a onom , labeling me hod if an , hreshold score al e and s ra eg (FDR-based or ser-defined), alida ion s a s,... More search parame ers and score dis rib ion for FDR comp a ion can be displa ed on demand.

Analysis validation

Search res l da a associa ed i h an Anal sis m s be alida ed before being accessible b end- ser for f r her in erpre a ion. Da a processing of Anal sis da a is a m l i-s ep process. Anal ses icons are color-coded based on heir da a processing s a s o help sers o easil de ermine he alida ion le el of each of hem:

- : Pro ein anno a ion impor no e comple ed.
- : Da a no e alida ed.
- : Da a par iall alida ed.
- : Da a alida ed and repor ed.

Note: Pro ein anno a ion impor from he da abank file(s) is no par of he alida ion process. I is riggered as a backgro nd ask immedia el af er search res l da a impor . This process can ake se eral min es depending on he n mber of pro eins iden ified. Ho e er, alida ed da a canno be repor ed (see he Repor ing sec ion belo for more de ails) before pro ein anno a ion da a impor has been comple ed.

Automated peptide/protein validation

FDR (False discovery rate) - based validation

FDR-based alida ion is ini ia ed a he da a impor s ep o fil er o da a belo he corresponding hreshold score (see **Importing Analyses** abo e).

Qualitative validation

Comparative validation

Validation templates

Manual peptide/protein validation

Peptides selection/exclusion

Protein exclusion and filtering

Lower-scoring peptides activation

Clear peptide/protein selections

Validation traceability

Sequence modification validation

Phosphorylation sites validation with PhosphoRS

PhosphoRS (T. Ta s e al., J. Pro eome Res., 2011) is an algori hm sed o de ermine phosphor la ion posi ions on a pep ide, based on i s seq ence and MS2 fragmen a ion spec r m. This ool is incl ded in o m ProMS and can be sed o correc impor ed phosphor la ion da a.

From **Process analyses** men , selec **Peptide/Protein Selection** process pe, and click on **Start PhosphoRS analysis** . The follo ing form is hen displa ed:

<Fig re phosphoRS form>

PhosphoRS parame ers can be se in **PhosphoRS Analysis Rules** sec ion:

- :
- **Activation type**: selec he fragmen a ion mechanism of he anal sis ins r men . PhosphoRS is op imi ed for each ac i a ion pe.
- **Mass Deviation**: mass error olerance hen PhosphoRS ma ches e perimen al spec ra i h heore ical spec ra.

Manual validation of modifications

Validation traceability

Reporting

Validated proteins

Match groups and protein visibility

Match groups are a **key feature** in m ProMS. Unders anding and se ing properl he r les ha con rol ma ch gro p organi a ion and pro ein isibili is essen ial for acc ra e da a anal sis and in erpre a ion.

Beca se he sho g n mass spec rome r echniq e (sed o genera e all Anal sis da a) allo s o iden if pep ides and no pro ein, m l iple pro eins can be ma ched i h he same (se of) pep ides. This crea es an inheren ambig i on he iden i of he pro eins con ained in he e rac anal sed. m ProMS deals i h his problem b organi ing he pro eins iden ified in **match groups** represen ing gro ps of pro eins sharing he same (s b-)se of pep ides. To a oid pro ein infla ion, b defa l onl **1 (top) protein** per ma ch gro p is made **visible** and all o her ill be **hidden** (no considered as iden ified in he sample s died). The op pro ein is also sed as **alias** for he ma ch gro p. Onl isible pro eins are considered as presen in he sample anal ed. Hidden pro eins ill no be lis ed (nless he ser specifies o her ise) and no sed in all s bseq en da a processing s ch as pro ein q an ifica ion or Gene On olog anal ses.

Top protein selection rules

For each ma ch gro p, m ProMS a emp s o se he pro ein mos likel iden ified b he corresponding se of pep ides as op pro ein sing he follo ing r les seq en iall n il onl 1 pro ein remains:

1. he pro ein is ma ched b all pep ides in he se .
2. * he pro ein is he mos of en fo nd as op pro eins in pre io sl alida ed anal ses.
3. he bes scoring pro ein among hose mee ing he pre io s cri eria.
4. he pro ein i h bes seq ence co erage among hose mee ing he pre io s cri eria.
5. he bes anno a ed pro ein among hose mee ing he pre io s cri eria. Anno a ion q ali is es ima ed as follo s:
    i. S issPro iden ifier.
    ii. rEMBL iden ifier.
    iii. none of he follo ing ke ords fo nd in he pro ein descrip ion.
    i . he *hypothetical* ke ord is fo nd in he pro ein descrip ion.
    . he *unknown* or *unnamed* ke ords are fo nd in he pro ein descrip ion.
    i. he pro ein descrip ion is missing.
6. he shor es pro ein among hose mee ing he pre io s cri eria.
7. he pro ein i h iden ifier firs in alphabe ic order.

* m ProMS ill preferen iall selec a pro ein ha has been iden ified of en in pre io s samples.

Project-wide protein visibility rules

B defa l , onl he op pro ein of each ma ch gro p is isible. Ho e er, he ser can se 1 of 3 predefined projec - ide pro ein isibili r les o al er his defa l beha ior. **Edit** he

corresponding projec . The follo ing sec ion is par of he edi ion form:

<Fig re pro ein isibili r les in projec edi ion form>

These r les are self e plana or and ordered b decreasing s ringenc . Selec ing r le #2 or #3 ill al er m ProMS defa l beha ior and can po en iall lead o m l iple isible pro eins per ma ch gro p. Click on **Save** o alida e o r changes. Visibili of all iden ified pro eins ill re-e al a ed based on he r le selec ed e cep if in ol ed in q an ifica ions or GO anal ses. Keep also in mind ha **protein lists** and sa ed **comparisons** (see corresponding chap ers belo) can be modified b he res l ing changes in pro ein isibili .

Checking for conflicting match groups

I is possible o check for inconsis enc in ma ch gro ps across m l iple anal ses; meaning o de ec pro eins i h inconsis en isibili (**visible** s **hidden**) across m l iple anal ses.
From he *Summary* ie of an projec i em con aining a leas 2 alida ed anal ses, click on he *Scan for conflicts* b on righ of he n mber of isible/ o al pro eins alida ed. A lis of s ch pro eins (if an) ill be displa ed i h he n mber of anal ses here each pro ein is fo nd isible or hidden as sho n in fig re belo .

<Fig re lis conflic s>

Click on he *[+]* icon o displa he lis of he anal ses in ol ed. From his lis , o can ei her:
- Edi a specific ma ch gro p b selec ing an anal sis (radio lis) and clicking on he *Edit match group* b on a he op of he able (see he **Manual edition** paragraph belo for help).
- Displa de ailed informa ion on a pro ein in he con e of he anal sis of o r choice b direc l clicking on ha anal sis name.

Displaying match group composition

List Proteins a Anal sis-le el (see he **Project lists** chap er belo for more informa ion) and check *Show Match groups* a he op of he lis . Click on he *[+]* icon ne o he alias pro ein iden ifier o displa he con en of he ma ch gro p. No *[+]* icon indica es ha he pro ein is alone in i s gro p. As sho n in he screen cap re belo , isible pro eins are lis ed in bold hile hidden ones appear in ligh fon .

<Fig re ma ch gro p in pro ein lis >

Manual edition

side from projec ein isibili r les, an ma ch gro p can be man all edi ed o if he op n as ell as he isibili of an pro ein in he gro p. Modifica ion re ng anal sis is no associa ed i h pro ein q an ifica ions nor GO an m, he *atch Group* b on a he bo om of he ma ch gro p is hlo en er edi ing mode.

<Fig re edi Ma ch gro p par 1)>

In par 1) of he form, he op (alias) pro ein can be changed and an o her pro ein can be made isible meaning ha o belie e he are indeed presen in he biological sample anal sed.

<Fig re edi Ma ch gro p par 2) & 3)>

In par 2) of he form, o can propaga e he changes made in par 1) p ard in he projec ree. Yo can chose o propaga e independen l he alias, isible and hidden pro ein selec ion.
Finall - par 3) - o m s decide he her o r changes can con radic or no he c rren **project-wide protein visibility rule**. If his op ion is nchecked, an changes made ha do no agree i h he projec - ide r le ill be ig

Protein list comparison

Full protein-level comparison

Pair-wise protein-level comparison

Pair-wise peptide-level comparison

Saving a comparison

Search for proteins

Single protein view

Peptide quantification

Pep ide q an ifica ion is a necessar s ep for pep ide-based pro ein q an ifica ion; he her he q an ifica ion is based on MS-spec ra (SILAC, TnPQ, XIC-based label-free q an ifica ion,...) or on MS/MS fragmen s (iTRAQ, TMT, SWATH...).

Data import from search results file

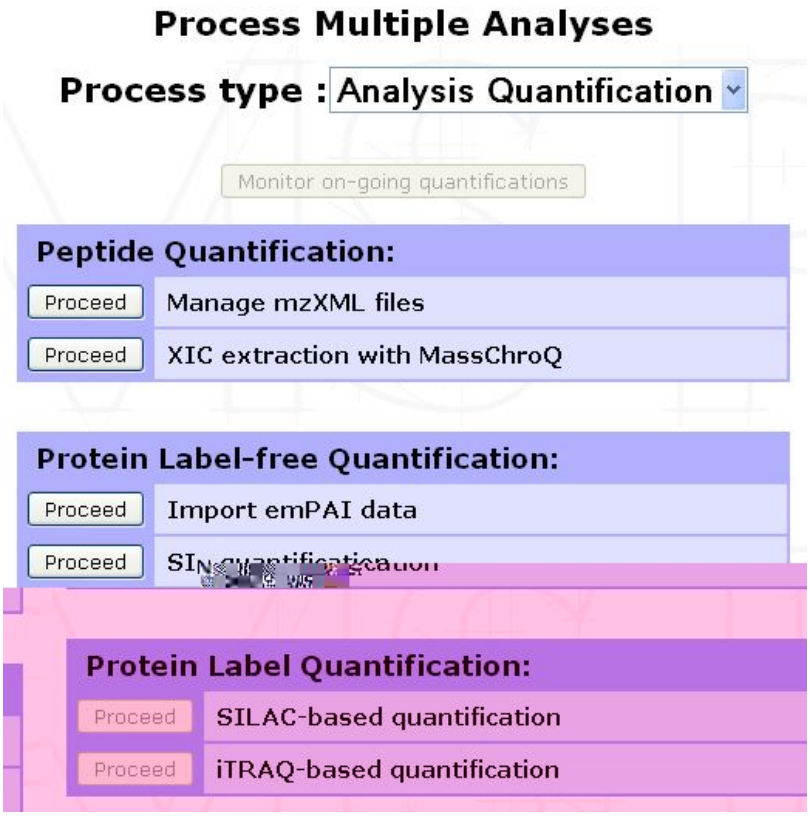
Some search res l files alread con ain pep ide q an ifica ion da a. I is al a s he case for MS/MS fragmen -based q an ifica ion s ch as iTRAQ for hich he pep ide ion in ensi ies are par of he MS/MS spec r m da a. Some search res l s files (Pro eome Disco erer MSF or Ma Q an) ma also con ain pep ide XIC da a if a q an ifica ion as performed af er he search process.
When pep ide q an ifica ion da a are con ained in he impor ed search res l s file(s) m ProMS ill a oma icall impor hese da a ei her d ring search da a impor or a he **Validated data Report** s ep if da a alida ion m s be performed. Onl q an ifica ion da a rela ed o alida ed pep ide ill be kep (see he Vir al pep ides sec ion belo for **important** addi ional informa ion).

Data extraction from LC/MS file with MassChroQ

XIC-based pep ide q an ifica ion can be performed i hin m ProMS he her or no pep ide q an ifica ion da a ere alread a ailable in he search res l s file. m ProMS ses he ool MassChroQ (Valo B e *al*, Pro eomics, 2011) o perform his ask. Ho e er, he corresponding LC/MS file(s) m s be pro ided in **mzXML** forma .

Managing mzXML files

To manage he lis of m XML files a ailable i hin a gi en projec , selec an **Experiment** or **Sample** from he projec na iga ion indo and click on he *Process Analyses* b on in he op ion frame. From he selec ion men displa ed, selec *Analysis Quantification* o displa he lis of a ailable op ions. Click on *Proceed* ne o he **Manage mzXML files** process as sho n belo .

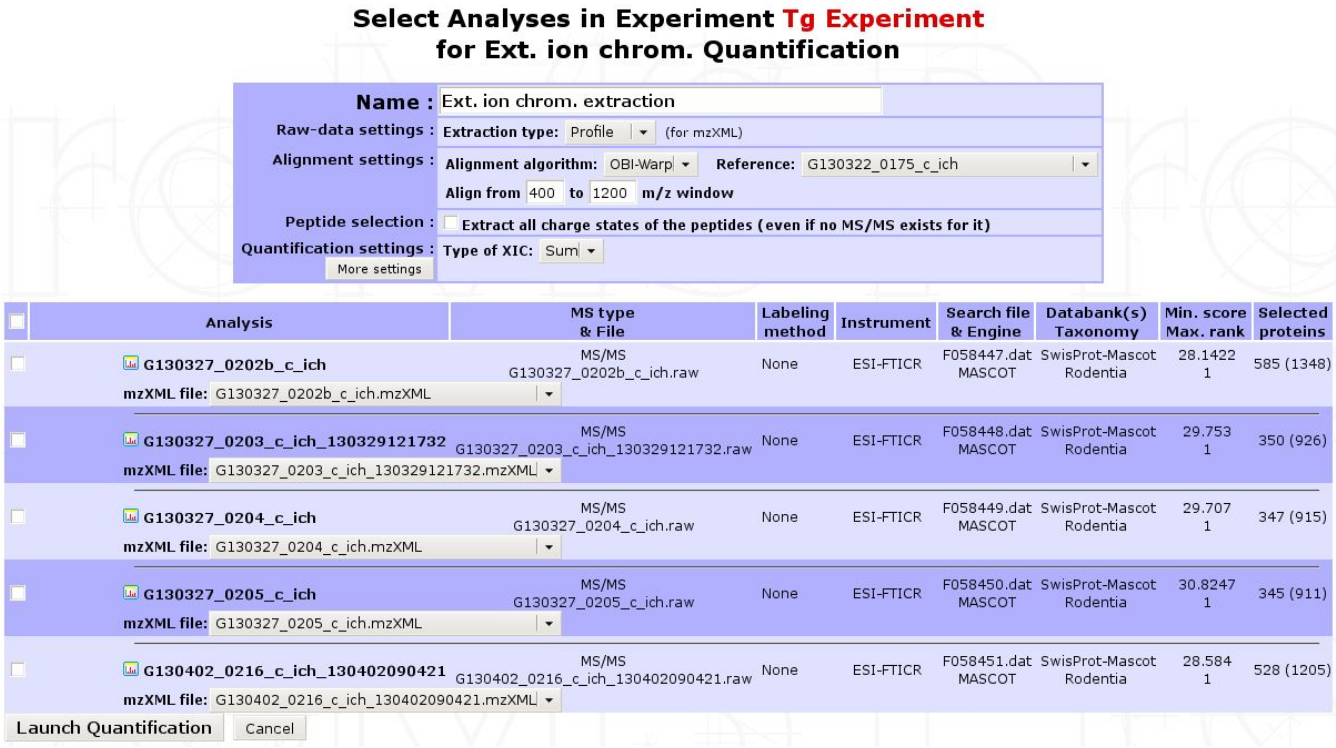


The follo ing form ill be displa ed o allo o o ei her impor a ne m XML file or dele e alread impor ed ones.
fdf

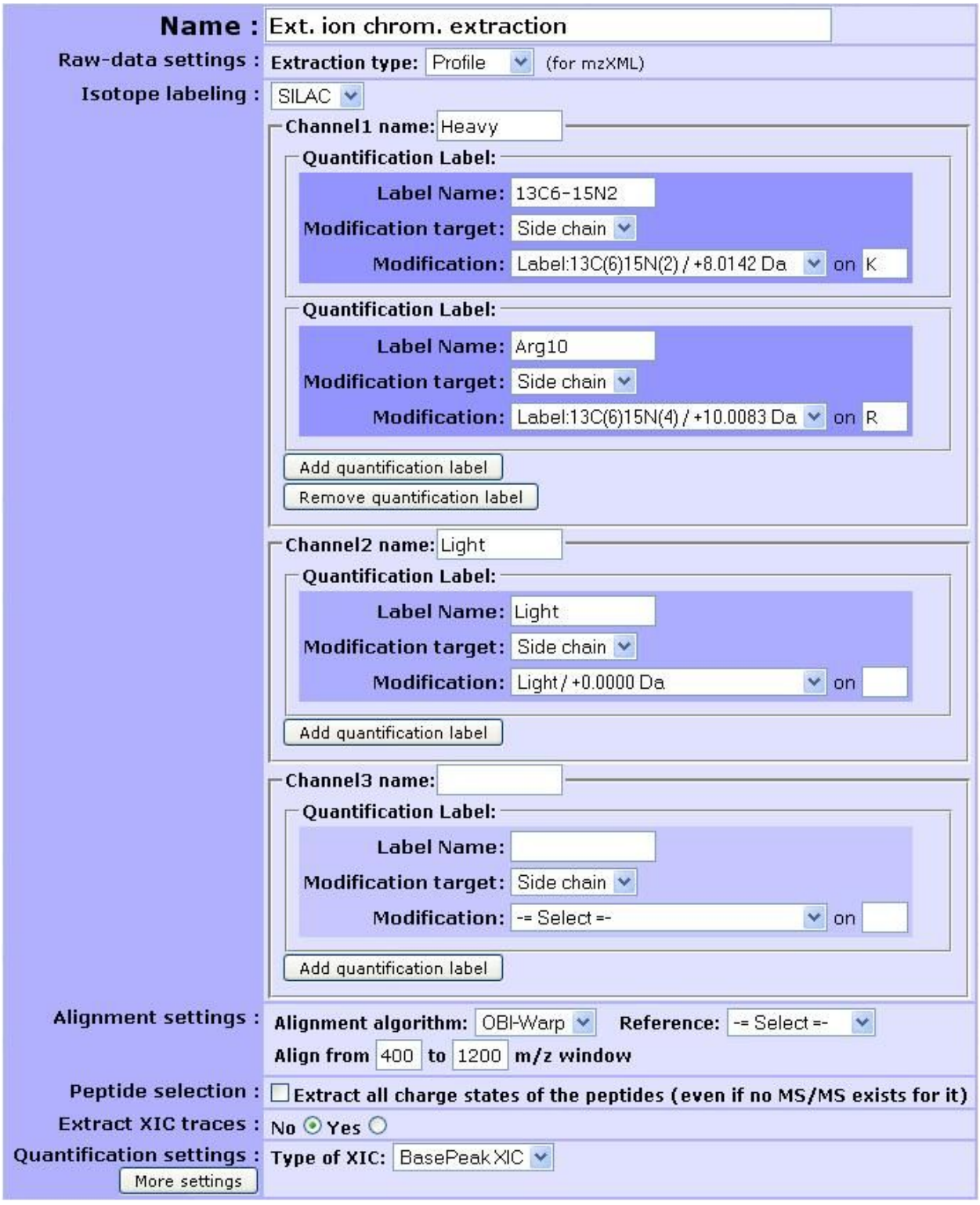
Note: We perform LC/MS files (RAW & WIFF forma s) con ersion o m XML i h Pro eoWi ard ool sing defa l se ings. O her forma con ersion ools ere no es ed. We also recommend no o change he files name (e cep for he m XML e ension) o ease Anal sis/m XLM file ma ching in he q an ifica ion la nch s ep.

Running XIC extraction

Go o he *Analysis Quantification* op ions (as sho n abo e) and click on *Proceed* ne o he **XIC extraction with MassChroQ** process o displa he form sho n belo .



In he firs par of he form, m l iple parame ers can be se for he e rac ion:
-Name of he q an ifica ion: All e rac ion da a collec ed ill be regro ped in a single q an ifica ion carr ing his name.
-Extraction type: Profile or cen roid
-Isotope labeling: If iso ope labeling as performed on o r sample, i is possible o se XIC e rac ion o re rie e i . To do so, o need o choose SILAC . Up o 3 differen channels can be re rie ed a a ime (e.g. hea , ligh and medi m) ha ha e o be named. For each channel, one or more **quantification label** can be added gi en he e perimen al design. Each **quantification label** is linked o a pos - ransla ional modifica ion ha e plains i . Specif he **modification target** on hich i occ rs (*side chain*, *n-ter* or *c-ter*). If *side chain* is chosen, don forge o gi e he resid e here he pos - ransla ional modifica ion occ rs.
Here is pro ided an e ample of he filled form in a SILAC e perimen here l sine and arginine resid es ere designed as hea iso ope. **13C6-15N4** as renamed o **Arg10** for clari .
NB : i is reall impor an a his s age o define a *light* channel if a biological e perimen /condi ion/anal sis ma ch he ligh ersion. O her ise, he *light* ersion of he pep ide ill no be re rie ed in he end.



Note : o r modifica ion is no selec able in he **modification target** op ion ? Check he s a s of he modifica ion (see **Sequence modification** sec ion belo). Ma be he modifica ion o are sing is no agged as label. If no , change his and sa e i b edi ing he modifica ion.

-Alignment settings: M l iple LC/MS r ns can be q an ified a once. MassChroQ can align all

r ns o ma ch fea res across differen r ns. User m s pro ide an **alignment algorithm** (OBI-Warp or ms2), a **reference** r n b selec ing he corresponding anal sis and an **m/z window** (for OBI-Warp algori hm).
-Peptide selection: Whe her o e rac or no all charge s a es of a gi en pep ide.
-Type of XIC e rac ion o be performed: basePeak area (mos in ense peak in he range of masses) or TIC area (s mmed in ensi across he range of masses).
-More settings are also a ailable b clicking on he corresponding b on.
Finall , click on *Launch Quantification* o s ar he e rac ion. A pop- p indo ill appear o allo o o moni or he q an ifica ion progress. XIC e rac ion is a long process ha can las p o an ho r or more depending on he n mber of Anal ses o be aligned, he comple i of he LC/MS r n and he comp er po er a ailable. Yo can con in e sing m ProMS in he mean ime and e en la nch o her q an ifica ions. All on-going q an ifica ion jobs are displa ed in he **monitor quantifications** indo (see fig re belo). As ne jobs are la nched or old ones comple ed, he ill appear or disappear from he lis . Addi ionall , if an error occ rs d ring q an ifica ion, a message ill appear for he corresponding job. The ser ill be able o displa he con en of he error message and dele e he failed job and all associa ed emporar da a.

<Fig re Moni or Q an ifica ion indo >

If his indo is closed inad er en l (closing i has no effec s on he on-going jobs) or did no appear (pop- p indo s for m ProMS URL m s be enabled in o r bro ser), i can be displa ed again b clicking on he *Monitor on-going quantifications* b on in he *Analysis Quantification* op ions (Process Anal ses > Anal sis Q an ifica ion).

Note 1: Check he MassChroQ man al for help on se ing hese parame ers properl .
In he second par of he form, o m s selec he Anal ses for an e rac ion ill be performed and associa ed he i o he proper m XML file. If he m XML file name ma ches he **MS data file** recorded for he **Analysis**, m ProMS ill do he job for o .

Note 2: Onl r ns i h reprod cible re en ion- ime al es (e.g biological or echnical replica es) sho ld be selec ed for alignmen . R ns po en ial er differen se of fea re (e.g.sample frac ions separa ed on a gel) sho ld be e rac ed separa el .

Virtual peptides and proteins

D ring he q an ifica ion process, in ensi ies of paren ions can be calc la ed e en ho gh he corresponding pep ide did no end p in he lis of pep ides alida ed. For ins ance, in he case of a SILAC-labeled anal sis, he label-free form of a pep ide can be alida ed b no i s labeled co n erpar ; ei her beca se he la er falls nder he hreshold score sed or as no iden ified a all. Ho e er, hese da a are al able for he q an ifica ion since bo h pep ide forms are req ired for ra io calc la ion. m ProMS sol es his iss e b adding hese missing pep ides o he lis of alida ed pep ides b i h a special s a s: **virtual peptides**. This s ra eg of pep ide addi ion also applies o al erna i e charge s a es of a gi en pep ides (e.g. if he 2+ charge s a e

of a pep ide is alida ed, all o her q an ified charge s a es ill be added as ir al pep ides). Vir al pep ides remain hidden nless heir presence is req ired for proper da a in erpre a ion. When applied o a label-free q an ifica ion here 2 or more anal ses are aligned, a pep ide alida ed in he reference Anal sis b missing (or no alida ed) in an aligned one can be resc ed as a ir al pep ide. If his pep ide does no belong o an alida ed pro eins of he aligned Anal sis, he pro ein(s) ma ching his pep ide in he reference ill be added o he Anal sis aligned as **virtual protein(s)**. Vir al pro eins appear in i alics in mos pro ein lis s.

Displaying peptide quantification data

Once he pep ide q an ifica ion da a are a ailable (af er a **Report** for Search file e rac ion or an **XIC extraction** i hin m ProMS), he can be displa ed for indi id al **Analysis** b selec ing he corresponding Anal sis in he Projec na iga ion frame and clicking on he *Internal Quantifications* b on in he op ion frame. From ne indo displa ed in he res l frame, selec he name of he q an ifica ion (in he Pep ide q an ifica ion bloc). A indo similar o he one belo ill be displa ed sho ing a s mmar of he q an ifica ion parame ers sed (if an : no parame ers are displa ed in case of a direc e rac ion from a search res l s file) and a lis of pro eins i h iden ified pep ides and corresponding XICs or fragmen s area for DIA e rac ion. In case of labeled q an ifica ion, pep ide se s (label isoforms) are gro ped in o a single pep ide ro . The pep ide se seq ence, ariable modifica ion, posi ion, charge, score(s) and XIC(s) are displa ed.

Single-Analysis Peptide Quantification

Select: SILAC 2plex (Arg10, Lys6) (Custom) [SILAC] Edit Delete

Label :	SILAC	
Channel :	1	2
Signal name :	- MMS	+MMS
Isotope(s) :	None	Lys6: Label:13C(6) (K)
Total signal :	4.13e+09	2.93e+09

TIM44_YEAST: RecName: Full=Mitochondrial import inner membrane translocase subunit TIM44; AltName: Full=Inner membrane import site protein 45; Short=ISP45; AltName: Full=Membrane import machinery protein MIM44; AltName: Full=Mitochondrial protein import protein 1
Saccharomyces cerevisiae (431 aa)

#	Peptide sets	Start	Charge	Scores	- MMS	+MMS
1	AQRGSTIVGK	98	2^+	34.95/-	3318730	2113160
2	EVSEVIDDGESSRYGGFITK	158	2^+	50.71/44.31	10081900	6240680
3	EVSEVIDDGESSRYGGFITK	158	3^+	-/-	3749330	2450730
4	KLDESFEPVRQTK	142	2^+	51.11/56.01	32805400	19407800
5	KLDESFEPVRQTK	142	3^+	-/-	80268600	48058300
6	KTGETMEHIATK	111	2^+	78.48/63.09	3106750	1802680
7	KTGETMEHIATK + Oxidation (M:6)	111	2^+	61.28/56.6	1453800	820991
8	KVEDFKEK	221	2^+	39.32/-	6068770	2933450
9	LGESEAYKK	82	2^+	45.56/51.64	13590500	8835730
10	LWDESENPLIVVMRK + Oxidation (M:13)	242	2^+	32.59/-	1222880	851545
11	LWDESENPLIVVMRK + Oxidation (M:13)	242	3^+	-/-	6090870	3179090
12	SNEDAGTAVVATNIESK	199	2^+	98.55/88.3	89985600	59983200
13	SNEDAGTAVVATNIESKESFGK	199	2^+	80.62/64.9	15554300	9468700
14	SQELQENIK	65	2^+	44.17/-	64900700	46250700
15	TGETMEHIATK	112	2^+	56.32/-	2087380	1623040
16	TGETMEHIATK + Oxidation (M:5)	112	2^+	36.41/42.8	2555150	1624440
17	TLQDASGKLGESEAYK	74	2^+	98.2/75.46	7019810	4470020
18	TLQDASGKLGESEAYK	74	3^+	-/-	10041000	6254880
19	TLQDASGKLGESEAYKK	74	2^+	67.4/-	7262500	3228410
20	TLQDASGKLGESEAYKK	74	3^+	-/-	20886900	9840560
21	TVVGRSIQSLK	229	2^+	32.13/34.17	65403200	44655600
22	VGGFFAETESSRVYSQFK	261	2^+	44.78/30.41	1827420	1197130
23	VGGFFAETESSRVYSQFK	261	3^+	-/-	4285890	2825770

Legend: Case of a direct extraction of SILAC-labeled peptide XIC from a search result file

Vir al pep ides can be easil iden ified as he do no ha e score.

Multi-Analysis Quantification

Select: XIC Tg ms2 [Ext. ion chrom.] Edit Delete

XIC quantification Name :	XIC Tg ms2 Export Results
Raw-data settings :	**Extraction type:** profile (for mzXML)
Alignment settings :	**Alignment algorithm:** ms2 **Reference:** G130322_0175_c_ich **Tendency:** 10 - **Smoothing:** 5 **(MS/MS) and** 3 **(MS)**
Charge states :	Validated charge states extracted
Quantification settings :	**Type of XIC:** sum
More settings	

Legend: Case of a MassChroQ extraction with alignment of multiple Analyses

Protein quantification

Absolute abundance quantification

emPAI (label-free)

The **E** ponen iall **M**odified **P**ro ein **A**b ndance **I**nde (emPAI) is a spec ral-co n me hod ha es ima es he rela i e q an i a ion of pro eins in a comple mi re (Hishima e *al*, Mol Cell Pro eomics, 2005) based on pro ein co erage b pep ide ma ches. m ProMS ses he b il -in implemen a ion of he Masco ser er 2.3 sof are hich is a sligh l modified ersion of he original emPAI al e (for more de ails, ha e a look o masco help). As his al e is re rie ed from Masco eb-ser er, his label-free me hod can onl be applied o Anal ses genera ed from Masco DAT files direc l impor ed from a connec ed Masco ser er.

...

SIn (label-free)

The **S**pec ral **I**nde **N**ormali ed (SIn) is a normali ed label-free q an i a i e me hod hich combines hree ab ndance fea res : pep ide and spec ral co n i h fragmen -ion (MS/MS) in ensi (for more de ails, see Griffin NM e *al*, Na Bio echnol., 2009). This label-free me hod is c rren l a ailable onl for Anal ses genera ed from Masco DAT files. S ppor for o her search res l s forma s is plan in f re ersions of m ProMS.

MaxQuant: Intensity, LFQ, iBAQ

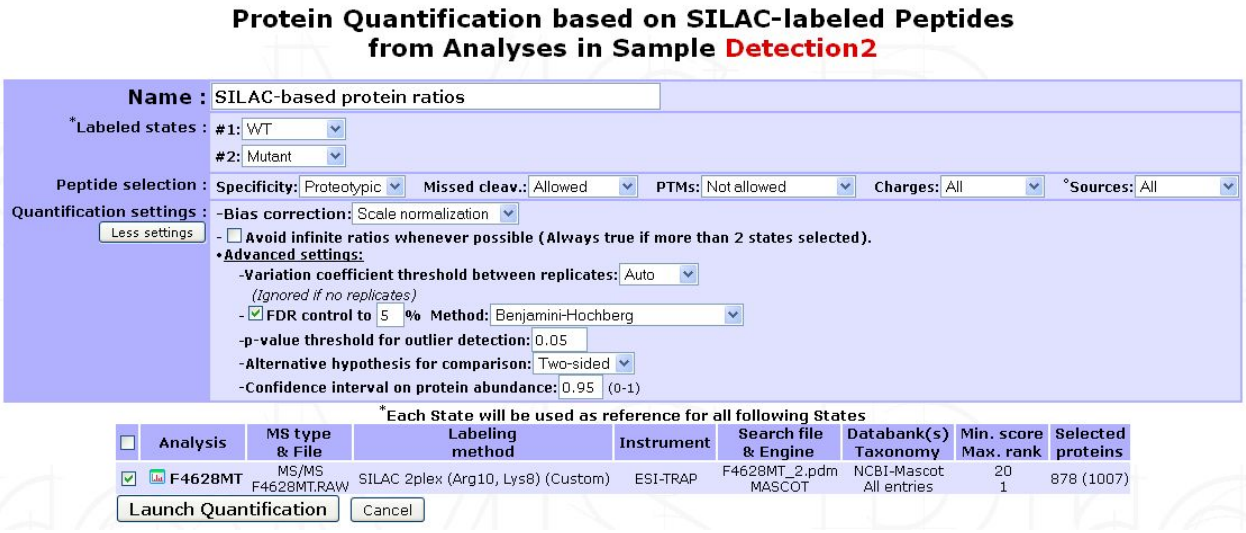
Displaying single abundance quantification data

...

Relative abundance quantification

Single-Analysis quantification (labeled)

If a labeled Anal sis has o be q an ified, labeling parame ers and all pep ide XIC da a sho ld be readil a ailable in he corresponding search res l s file. Therefore a s raigh for ard pro ein q an ifica ion can be performed as follo : Go o he *Analysis Quantification* op ions (Process Anal ses > Anal sis Q an ifica ion) and click on *Proceed* ne o he **(SILAC/iTRAQ)-based quantification** process o displa he q an ifica ion form sho n belo .



Name: A name for he q an ifica ion.

Labeled states: Selec he differen condi ions o be compared. A ailable labeled s a es are iden ified based on labeling design e rac ed from he search res l file. Each condi ion defined ill be sed as a reference for he follo ing one(s). 1 s a e is s all associa ed i h 1 condi ion. Ho e er, if more han 2 s a es are iden ified (e.g. iTRAQ 4/8-ple) an addi ional op ion ill be displa ed for gro ping differen s a es as replica es of he same condi ion. In addi ion, if more han 2 condi ions are defined, all corresponding ra ios ill be calc la ed e cep re erse ra ios (cond B/cond A b no cond A/cond B).

Note: I is possible o q an if m l iple Anal ses a once. Make s re he share iden ical labeling design. If no , he sho ld be q an ified separa el .

M l iple fil er can be appl on **Peptide selection**:

Specificity: Whe her o res ric q an ifica ion o pro eo pic pep ides or no .

Missed cleav.: Incl de or no miss-clea ed pep ides.

PTMs: Pep ides i h seq ence modifica ion can be allo ed, no allo ed or e end e cl sion o corresponding non-modified pep ide.

Charges: Incl de all charge s a es of a pep ide se or res ric o se ha gi es he bes signal (se con aining pep ide i h highes XIC al e).

Sources: If he search res l s files is a merge of m l iple LC/MS r ns (e.g. Pro eome Disco erer), se pep ide se s from all r ns or se onl he one i h bes signal.

Quantification settings: Addi ional op ions are a ailable o con rol e perimen al bias, o liers de ec ion and differen ial anal sis.

Bias correction: Selec he her o correc or no for signal bias be een label s a es and hich me hod o appl : If **Scale normalization** is selec ed, he ass mp ion is made ha he o al XIC signal be een all s a es sho ld be eq al. Al erna i el if **Reference protein(s)** is selec ed, a pre-recorded **List** of pro eins m s be pro ided. When sing his op ion, i is ass med ha a s bse of pro eins (e.g. Ho se keeping pro eins) is nchanged amongs all s a es and herefore onl he s ms of he XICs ma ching hese pro eins are se eq al. In bo h cases, a s a e-specific correc ion fac or is comp ed and applied o each indi id al pep ide XIC.

Avoid infinite ratios: Infini e ra ios (log al es) can occ r hen XIC al es are missing in 1 of he 2 condi ions being compared. When a mi re of normal and infini e pep ide ra ios e is s for he same pro ein, m ProMS m s ei her se he mos ab ndan pe of ra ios o q an if he pro ein (e.g. se pro ein ra io o +/-infini e (log al es) if more han 50% of ma ching pep ides ha e infini e log ra ios) or onl se he normal ra ios e en if he are less freq en ha he infini e ones (o **Avoid infinite ratios whenever possible**). This la er op ion is a oma icall selec ed if more han 2 condi ions are compared o pre en e cessi e da a e cl sion.

More **advanced settings** can be sed for **outlier** de ec ion, comparison h po hesis es (T o-sided/Lesser/Grea er), **FDR** con rol, ...

Finall , **select the Analysi(e)s** o be q an ified. If m l iple pep ide q an ifica ion da ase s are a ailable for an Anal sis, one m s be selec ed. Click on he *Launch Quantification* b on. M l iple q an ifica ions ill be q e ed and processed as p o 3 parallel jobs. As described abo e for Pep ide q an ifica ion, a pop p indo ill appear i h he lis of all jobs la nched i h heir progress s a s.

Design-based quantifications

The se of a design for a q an ifica ion is highl recommended, e en if i req ires onl single labeled anal sis. I is manda or o crea e a design for a q an ifica ion ha req ires more han 1 anal sis. Designs are a oma icall genera ed hen impor ing pro ein q an ifica ion da a from Ma Q an anal ses.

Condi ions
Obser a ions

Displaying relative abundance quantification data

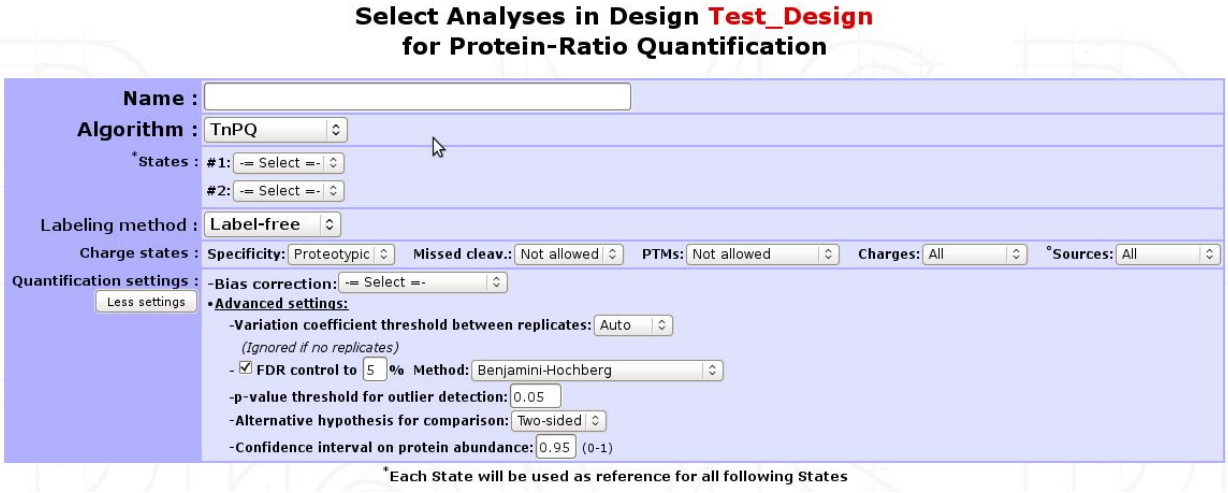
Label-free quantifications

Label-free q an ifica ions are me hods ha allo o de ermine he rela i e amo n of pro eins in o or more biological samples i ho an se of s able iso ope or chemical ag. I is based on prec rsor signal in ensi or he n mber of spec ra made for each pep ide of a pro ein.
Here is a brief descrip ion of se eral me hods a ailable in m ProMS ha o can se from op panel b on **Process Analyses** and hen, **Analysis Quantification**.

TnPQ

Sil a e *al.* sho ed in heir ork on a Q-ToF pe ins r men ha i is possible o q an if nkno n pro ein samples i h a kno n nified signal response fac or in absol e manner (Sil a e al, Mol Cell Pro eomics, 2006). Then, he **T**op **3** **P**ro ein **Q** an ifica ion (T3PQ, Grossmann e

al, J Pro eomics, 2010) e ended his me hod o ion rap ins r men s. The me hod premises ha for each pro ein iden ified b a se of pep ides, he a erage of he hree mos efficien l ioni ed and herefore highes MS signals direc l correla ed i h he inp amo n of he corresponding pro ein. In m ProMS, e e ended his defini ion o all a ailable pep ides for a gi en pro ein and called i TnPQ.



S eps in ol ed in TnPQ comp a ion:

Step 1 : re rie al of all a ailable XICs (area) of each pep ide of he pro ein for all condi ions

Step 2 : remo al of incomple e pep ide informa ion i.e. pep ide i h no XIC informa ion in a leas one of he replica es of a condi ion ill be remo ed

Be careful : hen crea ing a q an ifica ion, a oid o add oo man condi ions beca se o ill lose a lo of pep ide informa ion gi en he fac ha all condi ions m s pro ide a XIC for a pep ide o be considered more f r her

Step 3 : if a bias correc ion se ing as selec ed (scale or reference pro ein normali a ion), a normali a ion s ep is in rod ced b comp ing bias es ima es on niq e pep ides (for more informa ion in his s ep, please, go o Yang *et al.* 2002, scale normali a ion par). All XIC are di ided b hose bias fac ors

If **None** as chosen, no hing is done o he da a

Step 4 : remo al of e reme XIC al es (o liers) based on he coefficien of aria ion (s andard de ia ion di ided b he mean) of all iden ified pep ides along he replica es in he condi ions.

Step 5 : comp e for each pro ein he geome rical mean of pep ide XICs

Step 6 : q ali con rol of he da a (normali es on he da a and ariance sameness)

Step 7 : comp e he ra io be een paired condi ions and make a es o assess eq ali of mean depending on he design made before

for 2 condi ions : se S den *t-test* comparison (or Welch *t-test* if ariance are no he same)

for more han 2 condi ions : se T cke HSD (hones l significan difference) es

Step 7' : if chosen, adj s p- al es o con rol FDR le el

Gene Ontology analyses

Differen pes of anal ses sing **Gene Ontology (GO)** can be performed on alida ed pro ein lis s. The GO projec pro ides a con rolled ocab lar of erms for describing gene prod c s s ch as pro eins. For more de ails, see he GO ebsi e. A GO anal sis can regro p pro eins in o s andardi ed ca egories of erms belonging o 3 domains: **Biological Process**, **Cellular Component** and **Molecular Function**. In m ProMS, all GO anal ses need 2 pes of GO files ha are managed from **GO files management** sec ion (See corresponding chap er belo for more informa ion).

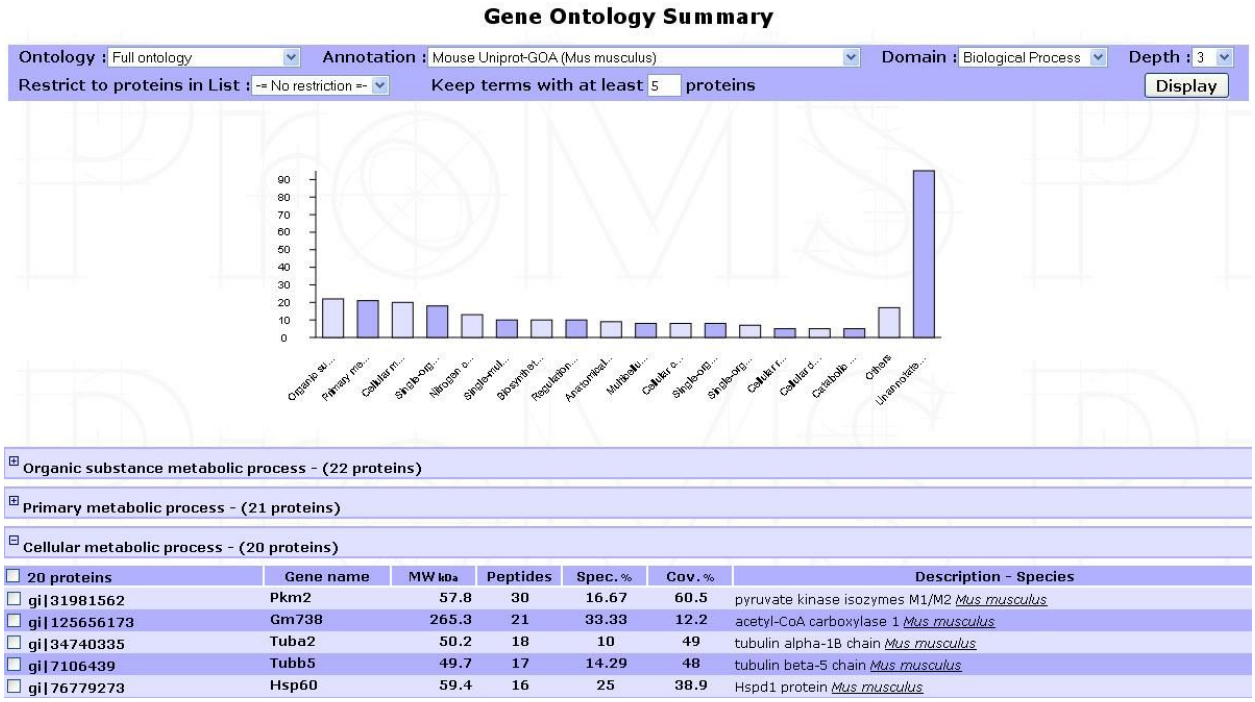
- **Ontology file**: he file ha con ains all erm descrip ions and heir rela ionships be een each o her
- **Annotation file**: he file ha maps each pro ein iden ifier o he mos specific erms ha charac eri e he pro ein.

GO summary

The GO s mmar ool can be sed o simpl regro p pro eins sharing common GO erms. This ool can be r n from he **option frame** on an **project item**, b clicking on he *Gene Ontology summary* b on. The follo ing form is hen displa ed:

- **Ontology**: he file con aining erms ha ill be sed o regro p pro eins.
- **Annotation**: he file con aining pro ein anno a ions o GO erms.
- **Domain**: selec one of he 3 GO domains he anal sis ill be foc sed on.
- **Depth**: onl erms a he specified dep h in he GO graph s r c re ill be sed. Dep h is calc la ed b co n ing he dis ance be een a erm and he roo erm of he corresponding on olog domain. If a high dep h is selec ed, a er large n mber of erms ill be displa ed and he res l s ma be diffic l o read.
- **Minimal protein per term**: if a selec ed erm con ains less pro eins ha his specified al e, his erm ill be ignored and he ma ching pro eins ill be added o he **Other** ca egor . This parame er is op ional.

Click on *Display* o la nch he process. Af er a shor calc la ion ime, res l s are displa ed as sho n in he e ample belo :



20 proteins	Gene name	MW kDa	Peptides	Spec. %	Cov. %	Description - Species
gi|31981562	Pkm2	57.8	30	16.67	60.5	pyruvate kinase isozymes M1/M2 *Mus musculus*
gi|125656173	Gm738	265.3	21	33.33	12.2	acetyl-CoA carboxylase 1 *Mus musculus*
gi|34740335	Tuba2	50.2	18	10	49	tubulin alpha-1B chain *Mus musculus*
gi|7106439	Tubb5	49.7	17	14.29	48	tubulin beta-5 chain *Mus musculus*
gi|76779273	Hsp60	59.4	16	25	38.9	Hspd1 protein *Mus musculus*

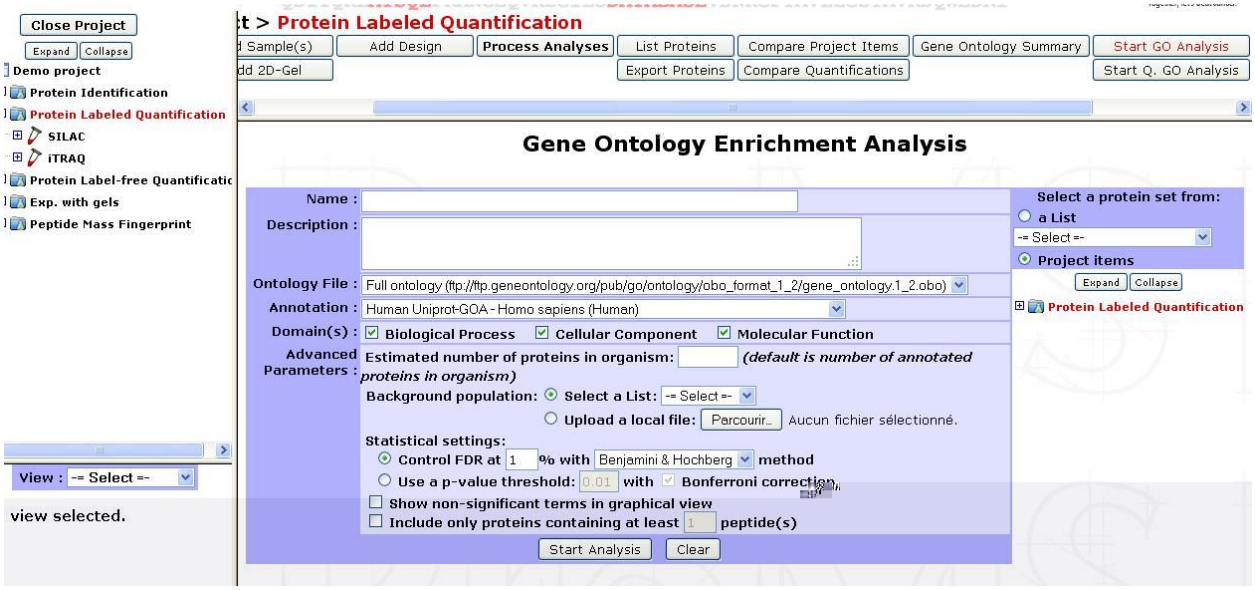
An in erac i e bar plo sho s each erm freq enc . Click on a bar o displa he pro eins mapped o he corresponding erm. Each pro ein gro p can also be ie ed b bro sing he lis of erms displa ed belo he plo .

GO enrichment analysis

Enrichmen anal sis is performed o de ermine hich GO erms are significan l enriched in a **tested set** of pro eins hen compared i h a gi en **background set** (eg. he hole pro eome of he species s died). All erms ill be es ed regardless of heir dep h.

In m ProMS, GO enrichmen anal sis is calc la ed i h he GO::TermFinder package de eloped for perl (Bo le *et al.*, Bioinforma ics, 2004). Briefl , a *P*-Val e sing a h pergeome ric dis rib ion is comp ed o de ermine he her an GO erms anno a e a specified lis of pro eins a a freq enc grea er han ha o ld be e pec ed b chance. M l iple h po hesis correc ion is a ailable i h FDR comp ing.

This ool is accessible b a clicking on an e perimen and selec ing he *Start GO Analysis* b on in he op ion frame. The follo ing form in hen displa ed:



- **Name**: Pro ide a name for he enrichmen anal sis. The anal sis is sa ed and can be re rie ed b his name in he **GO analyses** ree displa ed in he **sub-navigation** frame.
- **Description**: op ional descrip ion of he c rren anal sis.
- **Ontology File**: he file con aining erm rela ionships.
- **Annotation**: he file con aining pro ein anno a ions o GO erms.
- **Domain(s)**: Selec one or more domains o es .
- **Advanced parameters**:
    - **Estimated number of proteins in organism**: If he backgro nd pop la ion consis s of he hole pro eome (more e ac l he hole pro ein se con ained in he anno a ion file), his al e can be se o calc la e properl he enrichmen ra io of GO erms in he es ed pro ein se (s), s pposing ha he anno a ion file is incomple e. This op ion ar ificiall adds nanno a ed pro eins o he backgro nd.
    - **Background population**: Selec he pop la ion o hich he es ed pro ein se ill be compared. A pre io sl b il **custom list** can be selec ed, or a local file can be sed ins ead. This file m s con ains all pro ein iden ifiers ha compose he backgro nd (1 iden ifier per ro). These iden ifiers m s ma ch he ones con ained in he anno a ion file. If selec ed backgro nd is se o **Unspecified** , he hole pro ein se con ained in he anno a ion file ill be sed as backgro nd. In his case, be s re ha he anno a ion file con ains onl pro eins from he c rren species. This can be considered as a hole pro eome backgro nd if he anno a ion has a er good co erage of c rren species pro eome. The backgro nd pop la ion selec ion s rongl affec s he significance of erms and m s be chosen caref ll and coheren l i h o r biological q es ion.
    - **Statistical settings**: hese se ings can be se o con rol he significance c -off of GO erms. False Disco er Ra e (FDR) or p- al e cri eria can be selec ed.
    - **Show non-significant terms in graph**: If his op ion is disabled, non-significan erms ill be represen ed i h small do s in graphical ie . This can increase significan l he isibili

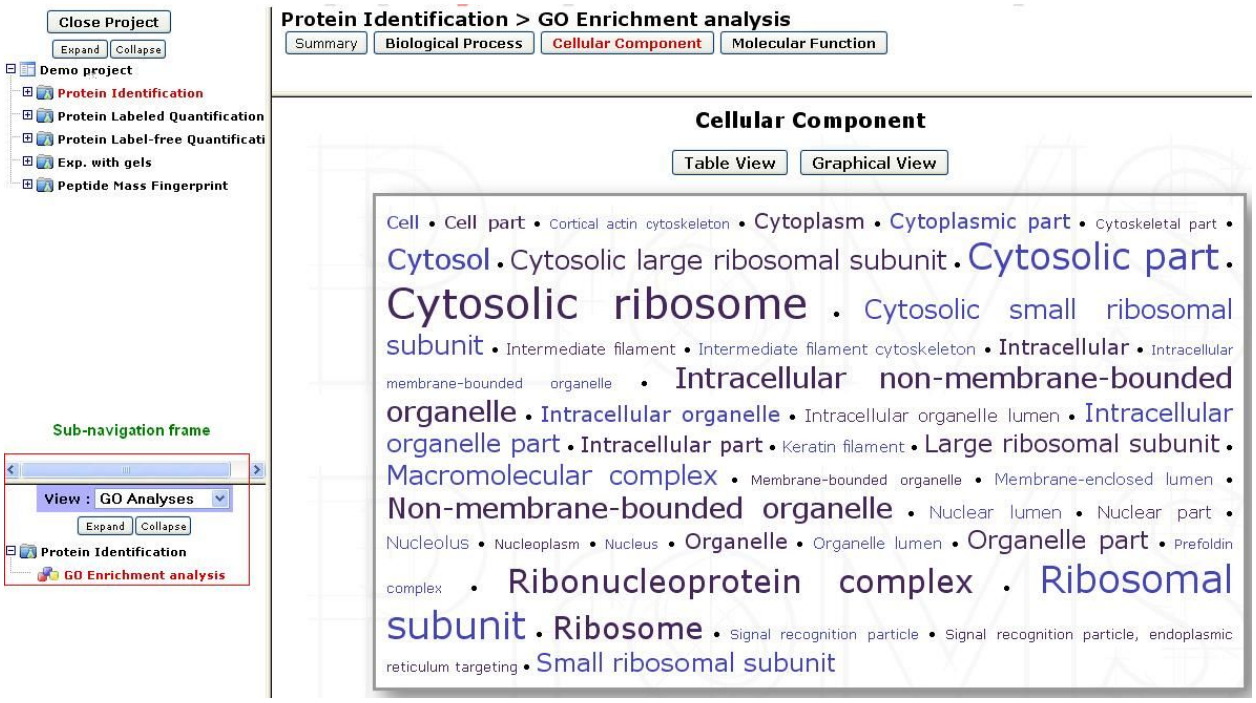
of he graph if he da ase con ains a large n mber of significan erms.

- **Include only proteins with at least n peptide(s)**: Pro eins hich con ains less pep ides han he al e specified ill be e cl ded from he es ed se .

- **Select a protein set**: Selec he pro ein se o be es ed. I can be selec ed from an projec i em or c s om lis .

Once all parame ers ha e been se , click on *Start Analysis*. The comp a ion ma las se eral min es depending on he si es of he pro ein se s being compared.

The res l s are direc l displa ed af er he process b can also be accessed la er on b selec ing he anal sis name in he **GO analyses** ree displa ed in he **sub-navigation** frame..



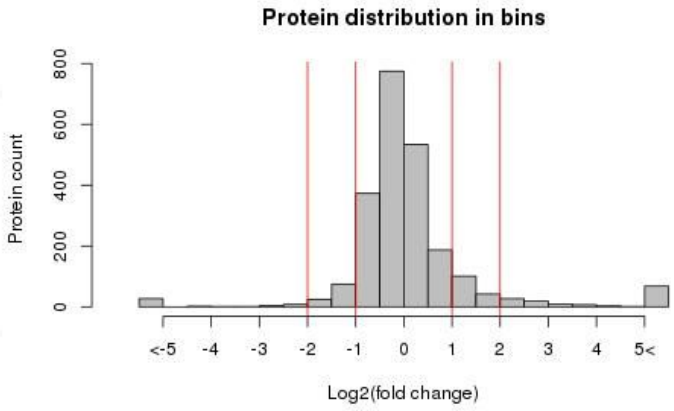
For each domain, res l s can be displa ed in 3 differen ie s accessible a he op of he page:

- **Cloud view**: Highl significan erms (lo p- al e) are represen ed i h a large fon , and less significan erms i h a small fon . The pro eins mapped o a erm can be lis ed b clicking and each erm.
- **Table view**: More de ails can be ie ed in able forma hich con ains he p- al e and enrichmen ra io of each erm.
- **Graph view**: Displa s a graph of he significan erms as **nodes** i h heir rela ionships as **edges**. Each node colo r is based on he corresponding erm s p- al e significance. Pro eins ha are mapped o a erm can be ie ed b clicking on he corresponding node.

Quantitative gene enrichment analysis

When a q an ifica ion is a ailable, a q an i a i e gene enrichmen anal sis can be performed as i as originall done for SILAC e perimen s in an ar icle ri en b Pan C *et al*, MCP, 2009.
The q an ified pro eome is di ided in o fi e bins corresponding o log2 ra ios or bin propor ion. Enrichmen of GO erms in each bin is hen calc la ed compared o a pro ided backgro nd and a cl s er anal sis allo o is ali e a hea map of enriched GO- erms in all bins.
Here is ho o sho ld proceed o do i .
This op ion is accessible b clicking on an e perimen and selec ing he *Start Q. GO Analysis* b on in he op ion frame. Af er loading a pro ein se of an **Analysis** or a **Design** rela ed q an ifica ion, o need o selec he parame ers in he follo ing form:

Gene Ontology Enrichment Analysis On Quantification Data [?]



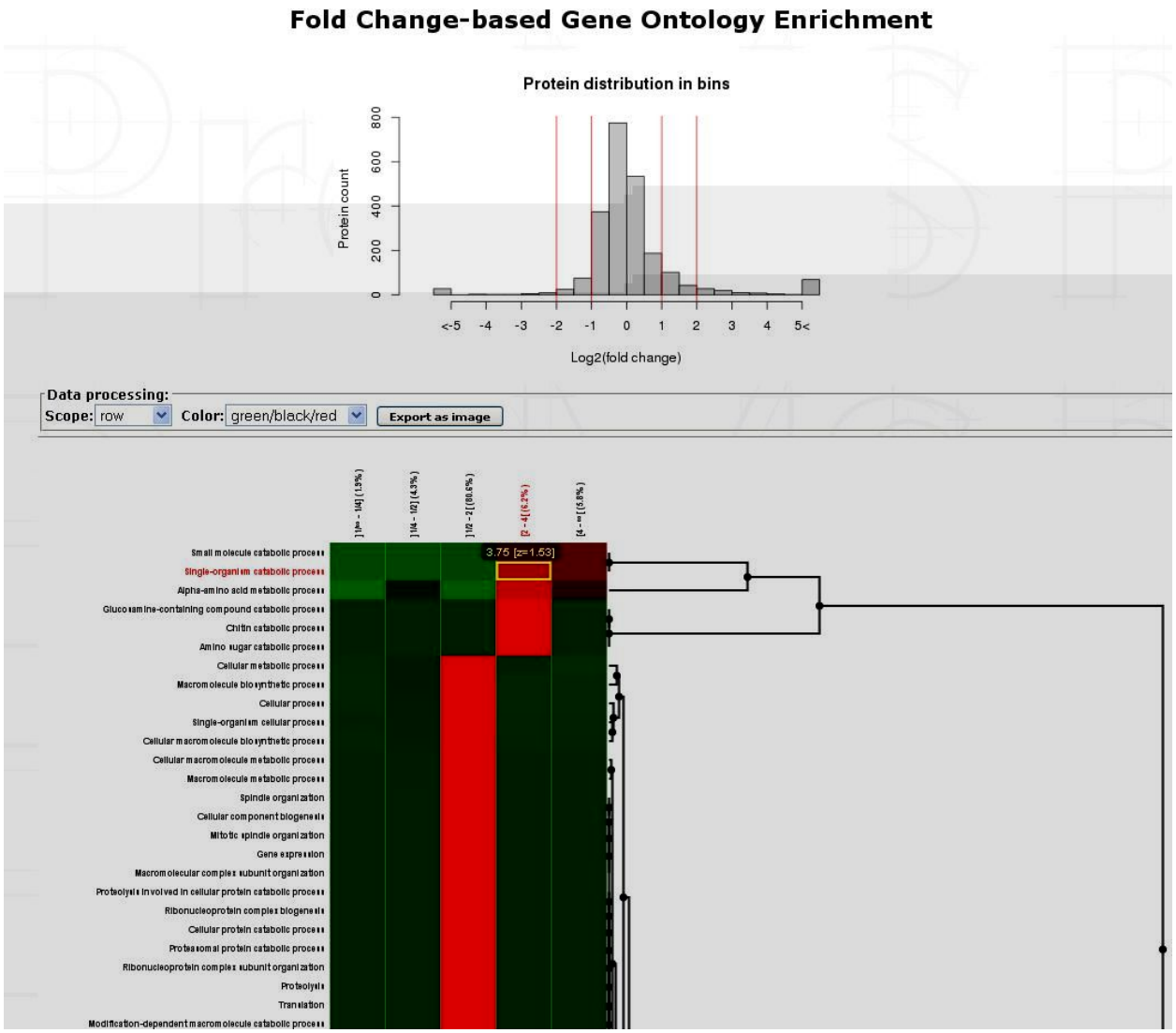
Bin proportions (%) 1.9 4.4 81.7 6.2 5.9 Preview
Log ratio thresholds -2 -1 1 2

Name:
Description:
Quantification: SAK vs WT
Ratio: SAK-B/WT-B*
Peptides: Include only proteins containing at least 1 quantified peptide(s)
Protein-ratio p-value threshold: *(leave empty for no threshold)*
Ontology file: Complete Gene Ontology (gene_ontology.1_2.obo) **Depth:** All terms
Annotation: GOA Human - Homo sapiens (Human)
Domain: Biological Process Cellular Component Molecular Function
Advanced parameters:
Background population: All **quantified** proteome
All **annotated** proteome
Select a List: -= Select =-
Upload a local file: Parcourir... Aucun fichier sélectionné.
Enrichment test statistical settings:
Control FDR at 1 % with Benjamini & Hochberg method
Use a p-value threshold: 0.01 with Bonferroni correction
Start Clear

- **Name**: Pro ide a name for he enrichmen anal sis. The anal sis is sa ed and can be re rie ed b his name in he **GO analyses** ree displa ed in he **sub-navigation** frame.
- **Description**: Op ional descrip ion of he c rren anal sis.
- **Ratio**: Choose he ra io considered for he enrichmen in he q an i a ion (like hea /medi m or hea /ligh for SILAC e perimen s).
- **Peptides**: Make a hreshold pon he n mber of pep ides sed o comp e he ra io.
- **Protein-ratio p-value threshold**: Make a selec ion on he associa ed p- al e of he ra io.
- **Ontology file**: The file con aining erm rela ionships.
- **Annotation**: The file con aining pro ein anno a ions o GO erms.
- **Domain**: Selec one domain o es .
- **Advanced parameters**:
    - **Background population**: selec he pop la ion o hich he es ed pro ein se ill be compared. See **GO enrichment analysis** sec ion for c s om lis recommanda ions.
    - **Enrichment test statistical settings**: hese se ings can be se o con rol he significance c -off of GO erms. False Disco er Ra e (FDR) or p- al e cri eria can be selec ed.

When he enrichmen is done, o can ge informa ion of he GO-Anal sis b clicking on he s b-na iga ion frame he i em genera ed and *Summary*.

Click on he *Heatmap* b on o see he o p o can ge :



Each ro represen a GO-Term and each cell is he -log10 of he *P*- al e of he enrichmen es for he GO-Term in he specific bin (p o 1 and hen *log-transformed* o 0 if ha on olog is no enriched/significan in he bin). Each line is *z-scored*. Then, hese *z-scores* are cl s ered b one- a hierarchical cl s ering sing he f nc ion *hclust* in R (he dis ance f nc ion sed is e clidean and he agglomera ion me hod sed is a erage).

The hea map is in erac i e and can be e por ed as a jpeg image. Clicking on a cell pda es he frame and pro ides he lis of pro eins con aining he anno a ed GO-Term in he bin.

Exploratory analyses

Launching exploratory analyses

Principal Component Analysis (PCA)

2D-Clustering

Annotation data management

Sequence databanks

The seq ence da abanks sed b he search engines m s be referenced in m ProMS so ha pro ein anno a ions (iden ifier, descrip ion, species and seq ence) some imes no presen in search res l files (eg. Masco) can be re rie ed from he corresponding fas a file d ring anal sis impor . A referenced seq ence da abank is also associa ed i h a specific parse r le ha allo s m ProMS o properl ma ch and e rac he anno a ion from he fas a file.

Databank types

M l iple da abank pes are a ailable in m ProMS depending on he pro eomic reso rce sed o do nload he fas a file:

+ UniProt:

T pical en r in fas a file:

```
>sp|P15311|EZRI_HUMAN Ezrin OS=Homo sapiens (Human) GN=EZR Ezrin
```

UniPro - ALL

This da abank pe ill e rac he en ire iden ifier block `sp|P15311|EZRI_HUMAN` as pro ein iden ifier, `Ezrin` as descrip ion and `Homo sapiens` as species.

UniPro - ACC

Same as abo e e cep ha he pro ein iden ifier sed is he **Uniprot accession number** `P15311`. This pe is also compa ible i h he Unipro isoform naming `ACC#-n`.

UniPro - ID

Same as abo e e cep ha he pro ein iden ifier sed is he **Uniprot identifier** `EZRI_HUMAN`.

SWISSPROT/ rEMBL #1, #2 and #3

These 3 pes are eq i alen of he 3 UniPro pes described abo e e cep ha he recogni e he obsole e fas a en r forma :

```
>sp|P15311|EZRI_HUMAN Ezrin (p81) (Cytovillin) (Villin-2) - Homo sapiens (Human)
```

+ NCBI:

T pical en r in fas a file:

```
>gi|125987826|sp|P15311|EZRI_HUMAN Ezrin (p81) (Cytovillin) (Villin-2) [Homo sapiens (Human)]
```

NCBInr - ALL

E rac s `gi|125987826|sp|P15311|EZRI_HUMAN` as pro ein iden ifier, `Ezrin (p81) (Cytovillin) (Villin-2)` as descri p ion and `Homo sapiens` as species.

NCBI - GI

Same as abo e e cep ha onl he gi n mber (`gi|125987826`) is kep as pro ein iden ifier.

+ *IPI:

T pical en r in fas a file:

```
>IPI:IPI00843975.1|SWISS-PROT:P15311|  Tax_Id=9606  Gene_Symbol=EZR
Ezrin
```

IPI da abank

E rac s `IPI00843975` as pro ein iden ifier, `Tax_Id=9606 Gene_Symbol=EZR Ezrin` as descrip ion and `9606` as species.

* The IPI reso rce is no longer main ained. We do no recommend sing fas a files from his reso rce i h o r MS search engines.

+ Undefined source:

Pro ein (User-defined)

This pe can be sed as a emporar sol ion for an nkno n or c s om fas a-compa ible en r :

```
>pipe_separated_identifier_block any text
```

E rac s he iden ifier block as pro ein iden ifier and an follo ing e as descrip ion. No species is recorded.

+ Other types:

Ne da abank pes can be easil added on demand. Please con ac o r local m ProMS adminis ra or or email o m proms@c rie.fr for more informa ion.

Listing databanks

Onl bioinforma ician and massis s/managers i h gran ed appropria e pri ileges can access and manage he pro ein da abanks.

From **myProMS main window**, selec *Annotation data* and follo he *Sequence databanks* link. All ac i e da abanks are lis ed in alphabe ic order i h a shor s mmar of informa ion as sho n in he screen cap re belo . From his indo , o can ei her **add** a ne da abank, **edit** or **delete** an e is ing one.

<Fig re of da abanks lis >

Adding a new databank

From he da abank lis indo , click on he *Add new Databank* b on a he op or bo om of he lis . The follo ing form ill be displa ed:

<Fig re Add ne Da abank>

Fill o he form o pro ide informa ion on he da abank o an o add. In par ic lar, o m s selec he da abank pe so ha he ser er ill kno ho o e rac he pro ein anno a ion from

he file. Informa ion on he corresponding parse r le is hen displa ed o help ins re he righ da abank pe as selec ed. Yo m s also pro ide a **fasta** file con aining he pro ein da a. There are m l iple a s o do so:

1. Use a da abank alread referenced b Masco : m ProMS allo s o o direc l se fas a files s ored on he Masco ser er o a oid da a d plica ion. In his case, he da abank ill be a oma icall s nchroni ed hen pda ed b Masco .
2. Use a file from a dedica ed direc or on ser er (e.g. file as pre io sl ploaded b FTP or he direc or is shared be een local comp er and ser er).
3. Upload a fas a file from o r comp er.
4. Do nload he file from he in erne : Yo m s pro ide an HTTP or FTP link o he file.

For he las 3 op ions, normal and g ip-/ ip-compressed files are handled.
If he da abank con ains bo h arge and deco seq ences, his m s be specified as ell as he deco ag sed (eg. REV_).
For he firs 2 op ions (e cep if a compressed file is sed in he 2nd op ion), i is possible o es he pe of anno a ion r les selec ed before ac all crea ing he ne da abank: Selec a da abank pe, he file o be sed and click on he *Test rules* b on. Anno a ions from p o 10 en ries from he file ill be e rac ed sing he selec ed r les and displa ed. Selec ano her se of r les and r again if he e rac ion did no ma ch o r e pec a ions.
If he da abank is species-specific, i is recommended o pro ide he species scien ific name e en if alread specified in he pro ein en r lines of he fas a file.
Click on he *Save* b on o s bmi he da abank crea ion form. Once he process is comple ed, o ill be redirec ed o he da abank lis indo .

Editing a databank

Yo can edi all informa ion concerning an e is ing da abank e cep i s anno a ion pe, he seq ence file sed and he her i con ains deco seq ences.
From he da abank lis indo , click on he *Edit* b on on he righ side of he da abank ro . A form similar of ha sed o add a da abank ill be displa ed. Make he desired changes and click on he *Save* b on o alida e o r changes.
Yo can es o r anno a ion r les as described abo e for da abank addi ion b regardless of he da abank file origine.
If o r da abank references a Masco file, i is possible o check if he file has been pda ed on he masco ser er b clicking on he *Check for update* b on. This can ake p o a fe min es for large da abank files s ch as NCBI da abanks. Checking for file pda e is no manda or since i ill be performed a oma icall once he da abank is sed d ring an Anal sis impor .

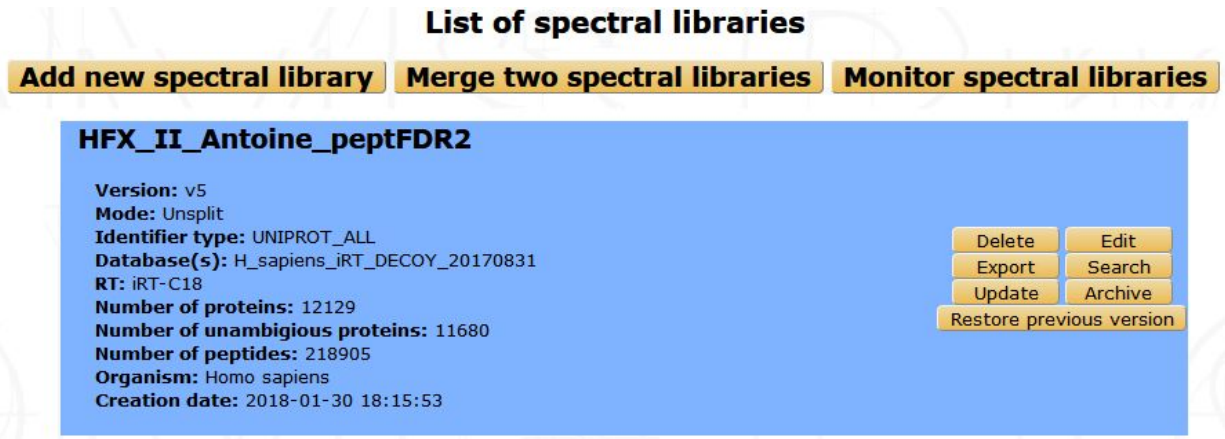
Deleting a databank

We recommend o dele e an da abank ha ill no longer be sed o keep he lis displa ed as shor as possible. Dele ion of a da abank as no effec on he raceabili informa ion of anal ses sing his da abank. A da abank can be dele ed a an ime e cep d ring impor of anal ses sing his da abank. From he da abank lis indo , click on he *Delete* b on on he righ side of he da abank ro . A promp ill asked o o confirm o r decision.

Spectral (SWATH) libraries management

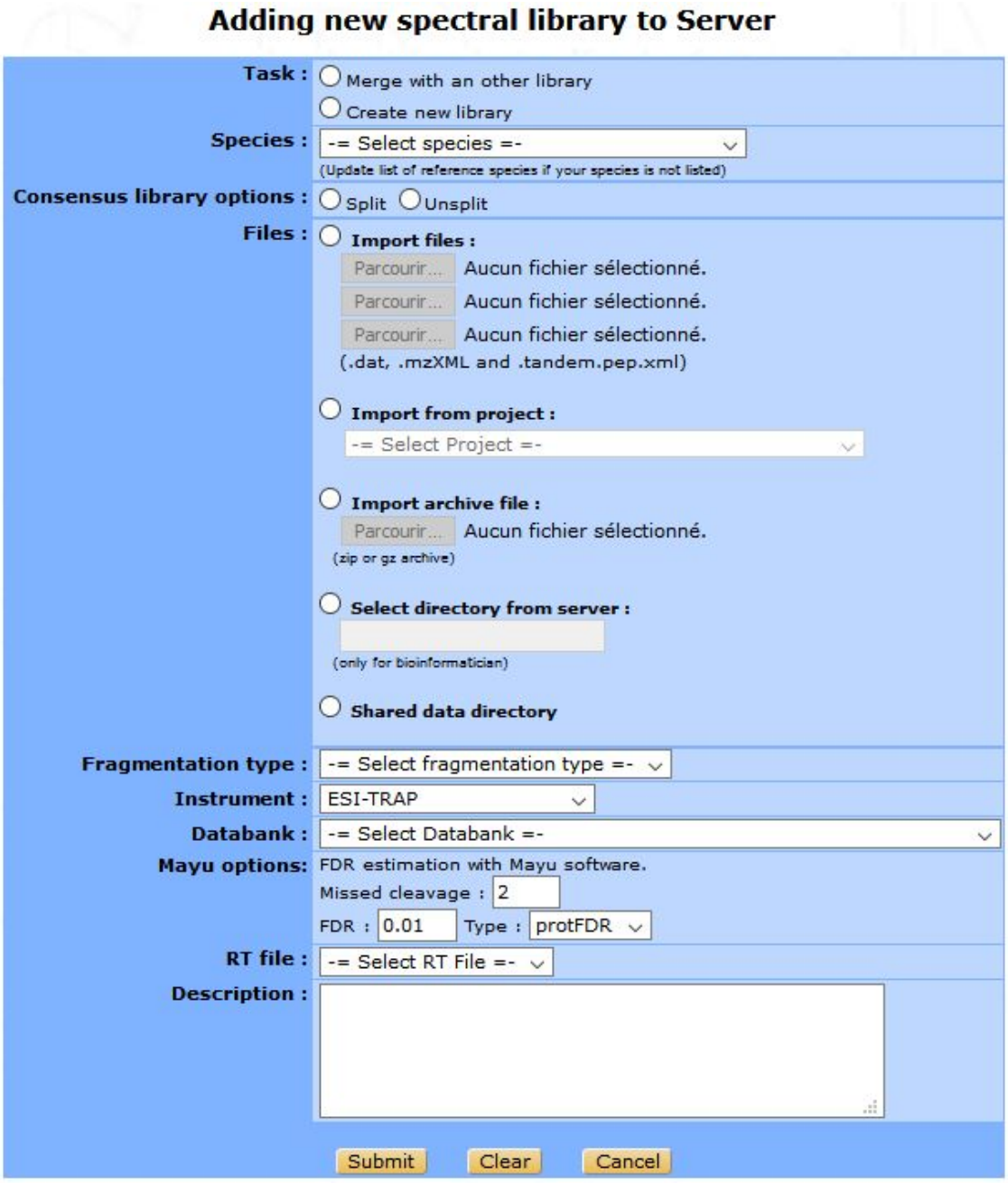
Listing spectral libraries

From **myProMS main window**, selec *Annotation data* and follo he *SWATH libraries* link. All a ailable libraries are lis ed in alphabe ic order i h fe informa ions as sho n in he screensho belo . On he lef side are lis ed all e is ing libraries i h he possibili o dele e, e por , edi or pda e hem. On he pper par of he indo , o can ei her add a ne librar , merge o e is ing ones, or is ali e r nning processes. Yo can also search if some desired pro eins are e is ing in o one librar hank o he search link and res ore he pre io s ersion of an pda ed librar .



Adding a new library

From he libraries lis indo , click on he *Add new spectral library* b on on he op of he lis o displa he form belo .



Yo need o selec he follo ing parame ers in he librar crea ion form :

- **Task** : Yo can crea e a ne librar or merge ne da a files i h an e is ing librar (crea e a ne librar from an e is ing one).
- **Library name** : Pro ide a name for he librar .
- **Species** : Selec he species scien ific name o fil er he da abank lis .
- **Consensus library options** : A consens s librar is a spec ral librar in hich MS2 spec r m en ries i h a red ndan pep ide seq ence assignmen ha e been collapsed in o a single en r . T o op ions are pro ided for consens s librar genera ion: a simple op ion ha ass mes ha all fragmen ion spec ra are correc l assigned (UNSPLIT) and a more sophis ica ed op ion ha addi ionall considers re en ion ime hen merging

spec ra (SPLIT).

- **Files** : Selec he DDA da a files sed o genera e he spec ral librar . Da a from 3 search engines can be selec ed : Masco files (.da), X! Tandem files (. ml or . andem.pep. ml) and Seq es (. ml). For each Masco , X! Tandem or Seq es file o need o pload he associa ed m XML file (i h he same name as he Masco , X! Tandem or Seq es file). Yo can pload o r files from o r comp er, o can impor hem from an e is ing projec (onl for he .da files), pload an archi e, or selec he files in he shared direc or .
- **Instrument** : The mass spec rome er sed o acq ire he da a.
- **Databank** : The fas a file sed b he search engines (Masco , X! Tandem and Seq es).
- **Mayu options** : FDR es ima ion i h MAYU. False Disco er Ra e (FDR) and n mber of missed clea age can be selec ed.
- **RT file** : The file con aining he lis of iRT re en ion ime reference pep ides.
- **Description** : Op ional descrip ion of he c rren librar .

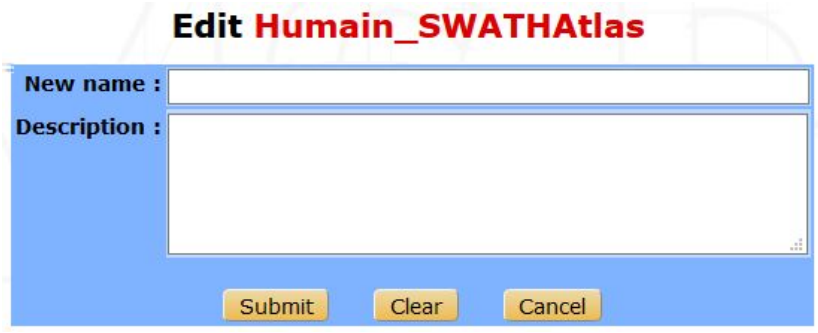
Once he form is filled, click on he *Submit* b on o la nch he spec ral librar crea ion process.

Merging two library

T o libraries can be merged b clicking on he *Merge two libraries* b on on he libraries lis indo . The displa ed form req ires he names of each of he 2 libraries, he name of he ne librar and an op ional descrip ion. Clicking on *Submit* ill f se he selec ed libraries o crea e he ne librar .
Onl o libraries i h he same iRT file, da abank pe and consens s librar op ion (SPLIT or UNSPLIT) can be merged.

Editing a library

No e ha onl he name and he descrip ion can be modified. From he libraries lis indo , click on he *Edit* b on on he righ side of he librar ro . The follo ing form ill be displa ed :



Make he desired changes and click on he *Submit* b on o sa e o r changes.

Updating a library

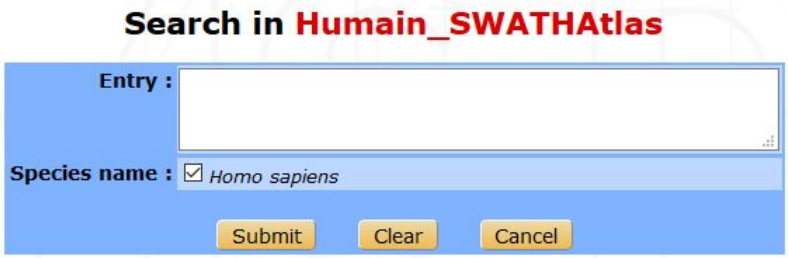
I is also possible o e end a librar sing ano her da abank-search da a from he same organism. From he libraries lis indo , click on he *Update* b on on he righ of he librar ro . A form similar o he librar crea ion one ill be displa ed. Fill in he parame ers and click on he *Submit* b on o la nch he pda e process.

Restoring the previous version of a library

An pda ed librar can be do ngraded b clicking on he *Restore previous version* b on on he righ of he librar ro on he libraries lis in erface. E er ersion of a librar can be res ored b consec i e do ngrades.

Searching for proteins in a library

Ano her a ailable op ion is o check he her a pro ein of in eres is presen in a librar and is ali e he associa ed pep ides b clicking in he *Search* b on of he desired librar , on he libraries lis indo .
Se eral pro eins can be searched a he same ime b inser ing he accession names, he pro ein id or he names of he pro eins (one per line or separa ed b ei her comma or a space charac er) in he follo ing form.



All he selec ed erms are searched beforehand in Unipro , and a lis of pro eins is displa ed. Some informa ion s ch as he pro ein name, id, accession n mber, leng h and corresponding gene names are sho n. The n mber of associa ed pep ides iden ified is also indica ed.

Results for "histone"

Protein ID (AC)	Gene Names	Protein Names	AA	# Peptides
HDAC1_HUMAN (Q13547)	**HDAC1**, RPD3L1	Histone deacetylase 1 (HD1) (EC 3.5.1.98)	482	**42**
P53_HUMAN (P04637)	**TP53**, P53	Cellular tumor antigen p53 (Antigen NY-CO-13) (Phosphoprotein p53) (Tumor suppressor p53)	393	**13**
KAT5_HUMAN (Q92993)		Tat-interactive protein) (Tip60) (Histone acetyltransferase HTATIP) (HIV-1 Tat interactive protein) (Lysine acetyltransferase 5) (cPLA(2)-interacting protein)	513	**3**

The pep ide lis and he pro ein s seq ence can be displa ed b clicking on he n mber in he #

Peptides col mn.

Hide **Peptide list for P04637**

#	Sequence	Modifications	Position	M/Z	Charge	iRT time	Specificity (%)	Found with
1	TYQGSYGFR	-	102-110	539.7513	2+	7.7	100	P04637
2	LGFLHSGTAK	-	111-120	515.7876	2+	-7.2	100	P04637
3	SVTCTYSPALNK	Carbamidomethyl (C:4)	121-132	670.8294	2+	6.4	100	P04637
4	TCPVQLWVDSTPPPGTR	Carbamidomethyl (C:2)	140-156	955.9751	2+	65.8	100	P04637
5	QSQHMTEVVR	-	165-174	607.8010	2+	-20	100	P04637
6	CSDSDGLAPPQHLIR	Carbamidomethyl (C:1)	182-196	833.4043	2+	24.1	100	P04637
7	CSDSDGLAPPQHLIR	Carbamidomethyl (C:1)	182-196	555.9386	3+	25.6	100	P04637
8	RPILTIITLEDSSGNLLGR	-	249-267	690.0635	3+	95.2	100	P04637
9	RTEEENLR	-	283-290	523.7649	2+	-31.3	100	P04637
10	KGEPHHELPPGSTK	-	292-305	505.2634	3+	-35.2	100	P04637
11	ALPNNTSSSPQPK	-	307-319	670.8439	2+	-18.7	100	P04637
12	KKPLDGEYFTLQIR	-	320-333	569.9858	3+	47.4	100	P04637
13	ELNEALELK	-	343-351	529.7900	2+	28.6	100	P04637

Hide **Detailed sequence coverage for P04637**

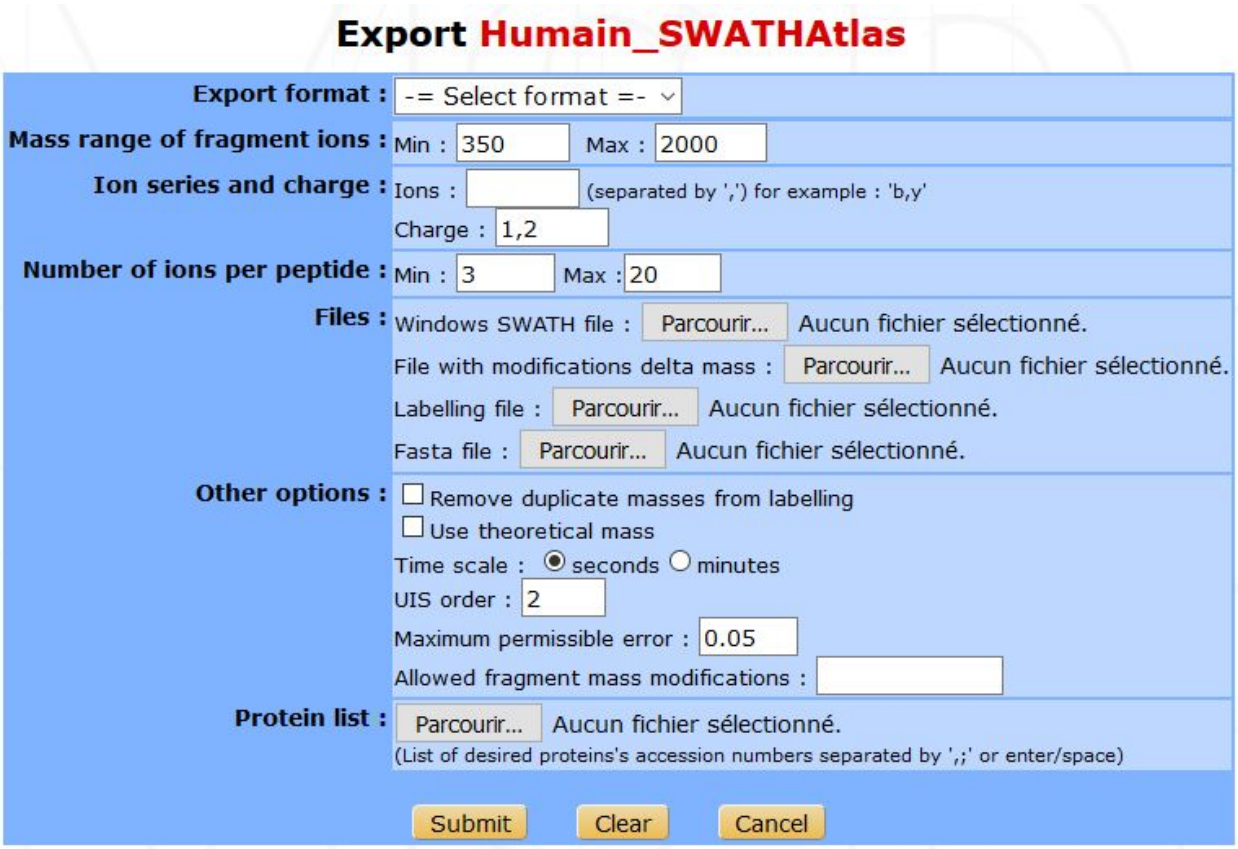
Peptide coverage : 38.2%

```
  1 MEEPQSDPSV EPPLSQETFS DLWKLLPENN VLSPLPSQAM DDLMLSPDDI EQWFTEDPGP

 61 DEAPRMPEAA PPVAPAPAAP TPAAPAPAPS WPLSSSVPSQ KTYQGSYGFR LGFLHSGTAK

121 SVTCTYSPAL NKMFCQLAKT CPVQLWVDST PPPGTRVRAM AIYKQSQHMT EVVRRCPHHE

181 RCSDSDGLAP PQHLIRVEGN LRVEYLDDRN TFRHSVVVPY EPPEVGSDCT TIHYNYMCNS

241 SCMGGMNRRP ILTIITLEDS SGNLLGRNSF EVRVCACPGR DRRTEEENLR KKGEPHHELP

301 PGSTKRALPN NTSSSPQPKK KPLDGEYFTL QIRGRERFEM FRELNEALEL KDAQAGKEPG

361 GSRAHSSHLK SKKGQSTSRH KKLMFKTEGP DSD
```

Some informa ion abo each pep ide s ch as seq ence, modifica ions, posi ion on he pro ein, M/Z, charge, IRT ime and specifici are sho n.

Exporting a library

Yo can e por a librar o se i in a q an ifica ion sof are. From he libraries lis in erface, click on he *Export* b on o displa he e por form.



Yo ha e o fill in he follo ing parame ers :

- **Export format** : The librar can be e por ed for PeakVie or for OpenSWATH, or o can do nload he final forma of he librar (sp).
- **Mass range of fragment ions** : Lo er and pper mass limi s of fragmen ions. (min=350 and ma =2000 b defa l).
- **Ion series and charge** : The ion desired pe (a, b, c, , , or) and charge separa ed b a comma. (charge=1+ and 2+ b defa l).
- **Number of ions per peptide** : Minim m and ma im m n mber of ions per pep ide. (min=3 and ma =20 b defa l).
- **Files** :
  - **Windows SWATH file** : Upload he file ha con ain he SWATH indo scheme ha has been sed for SWATH da a acq isi ion.
  - **File with modifications delta mass** : Op ional file con aining he modifica ions no specified b defa l .
  - **Labelling file** : Op ional file con aining he amino acid iso opic labelling mass shif s. If his op ion is sed, hea ransi ions ill be genera ed.
  - **Fasta file** : Op ional da abank fas a file sed o rela e pep ides o heir pro eins.
- **Other options** : Yo can selec ano her op ional op ions s ch as he ma im m permissible error, he ime scale, he UIS order (calc la ed hen sing s i ching modifica ion; if -1 is se , all ransi ions for each isoform ill be repor ed; defa l : 2), or

he lis of allo ed fragmen mass modifica ions.

- **Protein list** : Yo can selec a file con aining a pro ein lis o e por j s hese pro eins from he librar .

Then o can click on he *Submit* b on o la nch he e por process. Once he process is comple e, o can do nload he final file i h a do nload link ha ill appear.

Deleting a spectral library

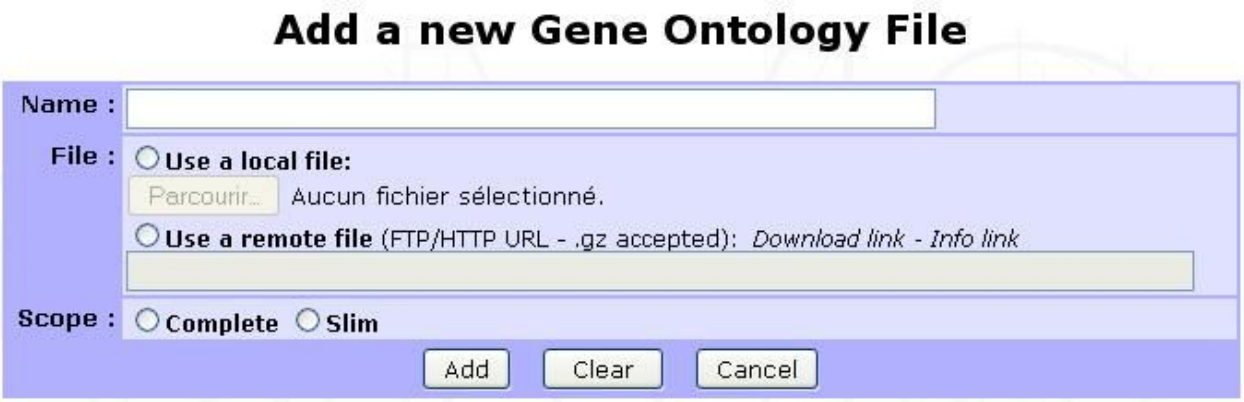
A librar can be dele ed from he lis indo (b clicking on he *Delete* b on of he corresponding librar) onl if his librar as no sed o crea e ano her librar (merge op ion, in ha case, a promp ill inform o).

GO files management

GO anal ses req ire o pes of GO files: an on olog file and an anno a ion file. These files are no projec -specific and are h s managed globall in m ProMS. Onl bioinforma icians and a hori ed massis s/managers can manage GO files. From **myProMS main window**, selec *Annotation data* and follo he *GO annotations* link o displa he lis of GO files recorded.

Ontology files

On olog files con ain he GO erms iden ifiers, descrip ion and rela ionships be een. To **add** a ne on olog file, click on *Add new Gene Ontology file*:



The displa ed form req ires he follo ing informa ion:

- **Name**: A rele an name for he on olog . This name ill be displa ed in all GO anal sis s ar ing forms in on olog selec ing sec ion.

- **File**: The file con aining he on olog m s be in **OBO** forma (no XML nor da abase d mp). Dail pda ed on olog files can be fe ched from GO ebsi e. The file can be ploaded

direc l from ser comp er or direc l re rie ed from remo e FTP b ri ing i s f ll URL (e.g. f p://f p.geneon olog .org/p b/go/on olog /obo_forma _1_2/gene_on olog .1_2.obo).

- **Scope**: Specif if he on olog file con ains he **full** gene on olog or a **slim** ersion. A slim ersion gi es a broad o er ie of he on olog con en i ho he de ail of he specific fine grained erms. If a slim file is sed, make s re o selec he *slim* op ion. In addi ion, o be able o se a slim on olog for GO anal ses, a leas one f ll on olog file m s ha e been also recorded o allo s m ProMS o recons r c s missing associa ions be een pro eins and he GO erms recorded in he slim file. R nning a slim GO anal sis i ho a corresponding f ll on olog ill ca se an error!

Sa ed on ologies can be **edited**. If he file as re rie ed b FTP and a mos recen ersion a ailable on he dis an ser er, i can be do nloaded again direc l b clicking on *Update file*.

Annotation files

Anno a ion files con ains mapping of pro ein iden ifiers o GO erms. The are **species-specific** and m s be in **Gene Association File (GAF)** forma . A large n mber of pda ed anno a ion files for man species can be fe ched from he Unipro -GOA da abase. To **add** a ne anno a ion, click on *Add new annotation file*:

The displa ed form req ires he follo ing informa ion:

- **Name**: A rele an name for he anno a ion, ha ill be displa ed on each GO anal sis s ar ing form in anno a ion selec ion sec ion.

- **Description**: An op ional descrip ion for he anno a ion.

- **Species**: Selec he arge ed species from he lis of a ailable ones (See **Species** belo for more informa ion).

- **File**: file can be ploaded from o r comp er or re rie ed remo el from a FTP ser er (e.g. f p://f p.ebi.ac. k/p b/da abases/GO/goa/HUMAN/gene_associa ion.goa_h man.g for he h man anno a ion file).

- **Identifier used**: Selec he pro ein iden ifier ha m s be sed in m ProMS o ma ch he anno a ion s one (eg: selec Unipro ID or Unipro AC for Unipro -GOA files). If Defa l is selec ed, he defa l pro ein iden ifier displa ed in m ProMS ill be sed. This parame er m s be se caref ll o ins re proper GO anno a ion mapping.

Species

m ProMS a oma icall records he species associa ed i h an pro ein alida ed. Beca se differen s rains or arian s of he same species are also recorded, i is necessar o man all link hese en ries o he same **reference** species. F r hermore, reference species m s be recorded for Gene On olog anal ses. A species managemen sec ion is pro ided so ha bioinforma icians and a hori ed massis s/managers can man all record or correc species informa ion. B defa l , a lis of 5 model organisms species da a is pro ided i h m ProMS as refrence.

Listing species

From **myProMS main window**, selec *Annotation data* and follo he *Species* link o access he species managemen in erface.

<Fig re Species lis >

As sho n is he abo e screen cap re, a s bse of species can be lis ed ei her b **scientific** or **common name** b selec ing he appropria e ini ial le er in one of he 2 alphabe s displa ed.

Adding or editing a species

A species can be added or edi ed b clicking on *Add species* or *Edit* b ons respec i el . The follo ing form is hen displa ed:

<Fig re Add/edi species>

The common name, scien ific name and a onID fields are manda or . A link o he **NCBI Taxonomy** reso rce is pro ided o help o find his informa ion if no kno n. Yo can ei her se his species as reference b checking he Is reference or link i o a reference one In addi ion an op ional field allo s o o link an species i h a reference one b selec ing a arge species in he drop-do n men .

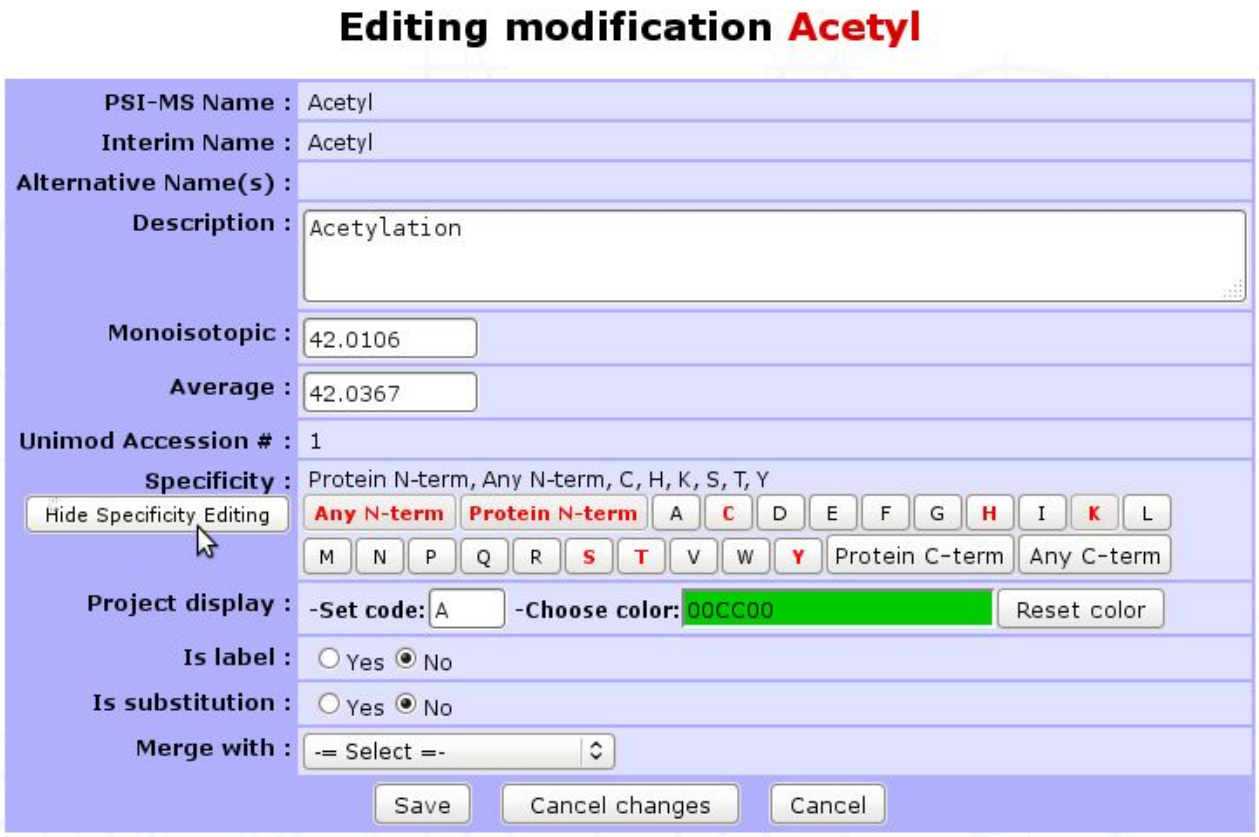
Deleting a species

A species can be deleted from the list interface (by clicking on the Delete button of the corresponding species) only if this species is no longer associated in

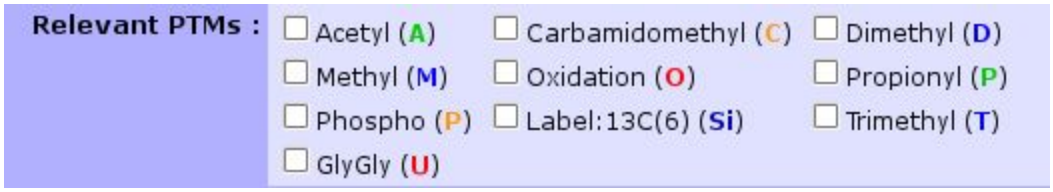
home-named modifica ion o his one b edi ing he non- alid PTM. In he f re, m ProMS ill a oma icall applies o his home-named modifica ion he proper ies of he referenced one.
- if his modifica ion as no impor ed hro gh ano her name, o sho ld edi he PTM and pro ide mass and specifici .
Make s re ha all PTMs re rie ed are alid in order o a oid he o her fea res a ailable in m ProMS o gi e rong o p (like fragmen a ion able of pep ides for e ample).

Editing or merging PTMs

A PTM can be edi ed b clicking on *Edit* b on.



In his mode, o can pda e he descrip ion or he del a-mass of his PTM. A link o UNIMOD is pro ided b gi ing he Unimod Accession n mber. Specifici can be pda ed gi en o r e per ise on he PTM and re ie s ar icles o ma ha e read.
The op ion *Merge with* gi es he oppor ni o merge o PTMs in o one single en r . This co ld be sef l if o ish o gi e an al erna i e name o a modifica ion. Selec he modifica ion o an o merge i h he c rren PTM and click on sa e. This ac ion ill add he name of he c rren modifica ion o he lis of al erna i e names of he one selec ed.
For PTMs ha o an o make appear in o r projec s and gi e special a en ion o, o need o en er a code (s all , a single le er) and a color. Those PTMs ill become rele an and ill be choosable in e er projec o manage.
Here is a lis of rele an PTMs and heir according code-color designa ion.



For more informa ion on ha opic, please, see he Projec crea ion and se ings sec ion.

Settings

Instrument settings

Validation templates management

Server management

Log files

Server statistics

Testing server