


Hands on with



 H. Paul Benton
Scripps Research

What we're going to do

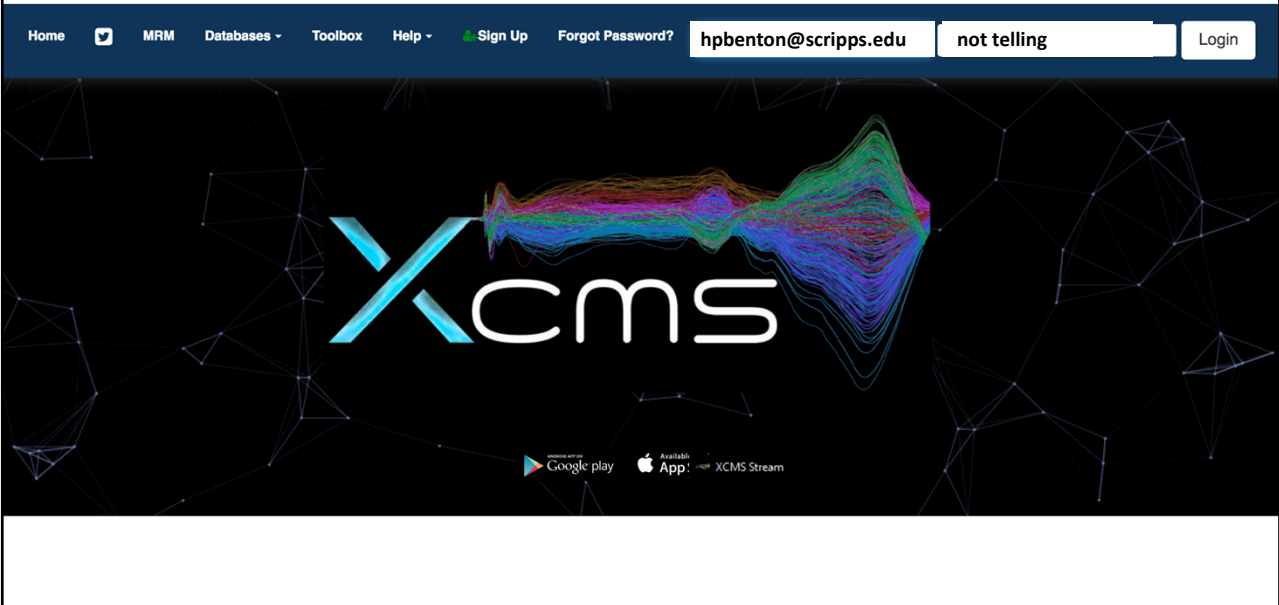
- It's your choice
 - A – Upload data and run a pairwise/multigroup jobs
 - B – Use a premade cake and find some interesting metabolites
 - C – Run over to xcms-mrm and analyse data there

First lets check your User



The screenshot shows the XCMS website's home page. At the top, there is a navigation menu with links for Home, MRM, Databases, Toolbox, Help, Sign Up, and Forgot Password?. Below the menu are two input fields: "Enter email address" and "Enter password", followed by a "Login" button. The main content area features a large, stylized "XCMS" logo in white and cyan, set against a dark background with a colorful, abstract wave pattern. At the bottom of the main area, there are icons for Google Play and the App Store, with the text "Available on the App Store" and "XCMS Stream".

First lets check your User



This screenshot is identical to the one above, but the login fields are now filled. The "Enter email address" field contains the text "hpbenton@scripps.edu" and the "Enter password" field contains "not telling". The "Login" button remains visible to the right of the password field.

What we're going to do

- Can everyone login ?
 - If you cannot please let me know!
- Otherwise –
 - Have a look at our help section. Please have a scan over this -
 - We have an FAQ page that we constantly update with questions that have been posted
 - We have a help page that we have need to get back to
 - Contact us page - E-mail us here- Not instantaneous!
 - Lets check out the account tab

The screenshot displays the 'Account' page for user 'hpbenton'. The main section is 'Platform Utilization', which includes a 'User Profile View' on the left and a 'Platform Utilization' dashboard in the center. The dashboard shows 67 created jobs, 34 completed jobs, 65 datasets (total files: 618), and 85.51 GB of aggregate storage. A pie chart shows job status: Active Datasets (70%), Active Results (30%), Pending Removal (0%), Filesystem Reserved (0%), and Other log/temp (0%). A table on the right lists datasets by file count, with 'non-preg-pos' having the highest count at 140 files.

Dataset	Files
non-preg-pos (339336)	140
Preg_pos (339335)	104
T24 (232088)	16
T0 (232087)	16
T96 (232089)	16
T196 (232090)	16
18months (317182)	12
cotex-24months (317183)	12
24months-Hippocampus (317185)	11
servier-ted-t1 (338577)	10

What we're going to do



- Datasets to use –
 1. Chronic Pain – SHAM rats vs (TNT) tibial nerve transected rats
 2. *Halobacterium salinarum* KO models – VNG111
 3. Coke Vs Pepsi – A dataset to find the secret ingredients
 4. MRM Plasma with 10 Standards

A bit about the datasets



- Chronic Pain – SHAM rats vs (TNT) tibial nerve transected rats
 - **Alterations in Spinal Cord Metabolism during Treatment of Neuropathic Pain**
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4548716>
 - **Metabolomics implicates altered sphingolipids in chronic pain of neuropathic origin**
<https://www.nature.com/articles/nchembio.767>
 - Main metabolites were identified
 - *N,N*-dimethylsphingosine (DMS) 328.321
 - Due to the Sphingomyeline-ceramide pathway having a non-reversible pathway from sphingosine to DMS
- https://www.dropbox.com/sh/wk4qggnzugphqkt/AACSln5rmiUVMFIQ_oxhn4fEa?dl=0

A bit about the datasets



- *Halobacterium salinarum* KO models – VNG111
 - **Meta-analysis of untargeted metabolomic data from multiple profiling experiments**
<https://www.nature.com/articles/nprot.2011.454>
 - Knock out models of *Halobacterium Salinarum* at VNG1816G, VNG1179c & VNG2094G vs the wild type URA3
 - Glutamic acid was seen to be important in some of the mutants but not all mutants. Copper trafficking was shown to be responsible for this change. (overlap analysis was able to find these results)
- <https://www.dropbox.com/sh/vkhja1ss21pear5/AADA1ggF2HN PcZ4yJIGWEIKYa?dl=0>

A bit about the datasets



- Coke Vs Pepsi – A dataset to find the secret ingredients
 - A highly complex study hidden in trade and lab secrets
 - Coke and pepsi was extracted using sea water type extraction
 - 10ml was injected for each
- <https://www.dropbox.com/sh/xviewq92n9mpsck/AADqflpNK 9bGvs6NTUZqTTeKa?dl=0>

The lab was never legally pursued by either company as long as the ingredients were never declared.

A bit about the datasets

SELECT DATASET 1
(See File Formats for more information)

Load New Dataset Select Dataset

ID	Dataset Name	File Count
Please upload or select dataset(s)		

Next

- Here you can upload the data and you'll be greeted by a upload box.
- Within the upload window you can drag and drop the files directly onto the uploader or select and find the files
- Wait till all the data is uploaded. This means a green check mark next to each dataset.
- Name the dataset something useful that you'll remember for later.
- Hit 'Next'

A bit about the datasets

SELECT DATASET 2
(See File Formats for more information)

Load New Dataset Select Dataset

ID	Dataset Name	File Count
343335	VNG1179	3

Previous Next

- Do the same again for the 2nd dataset you want to compare to.
- Once uploaded you will see the dataset with the saved name and the number of files that were uploaded.
- Again lets hit next to proceed.

A bit about the datasets

SELECT DATASET 2
(See File Formats for more information)

Load New Dataset Select Dataset

ID	Dataset Name	File Count
343335	VNC1179	3

Previous Next

- Do the same again for the 2nd dataset you want to compare to.
- Once uploaded you will see the dataset with the saved name and the number of files that were uploaded.
- Again lets hit next to proceed.

A bit about the datasets

SELECT PARAMETERS

Parameters -

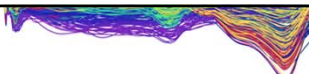
- HPLC / Q-TOF (1)
- HPLC / UHD Q-TOF (2)
- UPLC / UHD Q-TOF (155)
- HPLC / UHD Q-TOF (HILIC, neg, mode) (6674)
- HPLC / Bruker Q-TOF neg (5787)
- UPLC / Bruker Q-TOF pos (6675)
- UPLC / TripleTOF pos (769)
- HPLC / Single Quad (261)
- UPLC - High Res POS (Waters) (11025)
- HPLC - UHD Qtof pairs (7288)

Note: You may need to...

View/edit

- So now we have to look at parameters !!!
- We can either choose a general selection of parameters that will most likely work pretty well but for fine tuning select a close set of parameters and then click the 'View/edit'
- Before proceeding please note that statistics and model type are saved with the job and not with the parameter set.
- If you save the parameters and then go back to view it within the job the parameters will reset to default settings for the statistics and model type

A bit about the datasets



View/Edit Parameters for Job

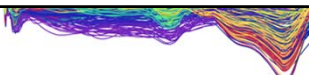
Polarity is defined on the General tab and will affect values on the Annotation and Identification (adducts) tabs. Job results will be misleading if this value is not correctly defined.
 The current parameter set is read-only. Use **Create New** button below to modify parameters to suit your job.

General **Feature Detection** Retention Time Correction Alignment Statistics Annotation Identification Visualization Miscellaneous

Method: centWave

Option	Value	Note:
ppm	30	maximal tolerated m/z deviation in consecutive scans, in ppm (parts per million)
minimum peak width	10	minimum chromatographic peak width in seconds note: must be less than max peak width. See also here .
maximum peak width	60	maximum chromatographic peak width in seconds note: must be greater than min peak width. See also here .
View Advanced Options		
Signal/Noise threshold	6	Signal/Noise threshold
mzdiff	0.01	minimum difference in m/z for peaks with overlapping retention times, can be negative to allow overlap
Integration method	2	Integration method. If =1 peak limits are found through descent on the mexican hat filtered data, if =2 the descent is done on the real data. Method 2 is very accurate but prone to noise, while method 1 is more robust to noise but less exact.
prefilter peaks	3	Prefilter step for the first phase. Mass traces are only retained if they contain at least [prefilter peaks] peaks with intensity >= [prefilter intensity]
prefilter intensity	500	Prefilter step for the first phase. Mass traces are only retained if they contain at least [prefilter peaks] peaks with intensity >= [prefilter intensity]
Noise Filter	0	optional argument which is useful for data that was centroided without any intensity threshold, centroids with intensity < noise are omitted from ROI detection

A bit about the datasets



View/Edit Parameters for Job

Polarity is defined on the General tab and will affect values on the Annotation and Identification (adducts) tabs. Job results will be misleading if this value is not correctly defined.
 The current parameter set is read-only. Use **Create New** button below to modify parameters to suit your job.

General Feature Detection Retention Time Correction Alignment **Statistics** Annotation Identification Visualization Miscellaneous

Option	Value	Note:
Statistical test	Paired parametric t-test	Statistical test method: Welch t-test (unequal variances) or Wilcoxon Rank Sum test
Perform paired test	VIEW PAIRS1	The selected statistical test is performed as a paired test. The sample pairs need to be specified.
Perform post-hoc analysis	True	Perform post-hoc analysis [multigroup only]
p-value threshold (highly significant features)	0.01	Features with a p-value less than this threshold are considered highly significant. Some statistical figures (e.g. Mirror plot) are generated using only the dysregulated features according to this threshold.
fold change threshold (highly significant features)	1.5	Features with a fold change greater than this threshold are considered highly significant. Some statistical figures (e.g. Mirror plot) are generated using only the dysregulated features according to this threshold.
p-value threshold (significant features)	0.01	Features with a p-value less than this threshold are not considered significant and are omitted from some calculations to save time and space. EIC's, annotations and database ID's are not generated for features with p-values above this threshold.
View Advanced Options		
value	into	intensity values to be used for the diffreport. If value="into", integrated peak intensities are used. If value="maxo", maximum peak intensities are used.
Normalization	None	Normalize the intensity values by either probabilistic quotient or cyclic loess normalization.

A bit about the datasets

General Feature Detection Retention Time Correction Alignment Statistics Annotation Identification Visualization Miscellaneous

Option **Value** **Note:**

ppm 10 tolerance for database search

adducts [M+H]⁺
[M+NH4]⁺
[M+Na]⁺
[M+H+2O]⁺
[M+H+2H2O]⁺
[M+K]⁺
[M+ACN+H]⁺
[M+ACN+Na]⁺
[M+2Na-H]⁺
[M+2H]²⁺
adducts to be considered for database search

sample biosource SELECT BIOSOURCE set default
SELECTED: HSAL478009 Select your species/cell line, etc. that correspond to your samples. Default human

pathway ppm deviation 5 metabolite pathway lookup

input intensity threshold minimum intensity cut-off for pathway analysis

significant list p-value cutoff AUTO significant list p-value cut-off

[Create New](#) [Cancel](#)

Show entries Search:

Select	ID	Biosource	Strain
<input type="button" value="SELECT"/>	GCF_000515255	R. salinarum DSM 9154	
<input type="button" value="SELECT"/>	GCF_000753715	I. salinarum	
<input type="button" value="SELECT"/>	HSAL478009	H. salinarum R1	

Showing 1 to 3 of 3 entries (filtered from 7,627 total entries) Previous Next

A bit about the datasets

1 SELECT DATASET 1 2 SELECT DATASET 2 3 SELECT PARAMETERS 4 REVIEW & SUBMIT

REVIEW & SUBMIT

User ID
11405

Job ID
1245668

Job Name
[P_2018-07-18_22.51](#)

Dataset1
343336


Dataset2
343335

Parameter ID
HPLC/Q-TOF (1)

[Previous](#) [Submit](#)

So now what

- Normally at this point you would have to wait until your job was finished.
- For today I have given everyone multi-submit access and setup our own processing server so things should be fast for us.
- If you were following the above you should see something like the below image -



The screenshot shows a job management interface with a table of jobs. The table has columns for JobID, Status, Progress, JobName, Datasets/Sources, Created, Parameters (ID#), Group, Share, and Delete. A single job is visible with the following details:

JobID	Status	Progress	JobName	Datasets / Sources	Created	Parameters (ID#)	Group	Share	Delete
1245668	PROCESSING	SUBMITTED 0%	P_2018-07-18_22:51	VNG1816 (#343336) VNG1179 (#343335)	2018-07-18 22:51:27	1			

Running the other job types ...

- I recommend to run the VNG-URA datasets as both a pairwise setup
 - VNG1816G VS URA (WT)
 - VNG1179c VS URA (WT)
 - VNG2094G VS URA (WT)
- Then again as a multigroup and finally
- The above pairwise as a meta-XCMS analysis to see the differences between the multigroup and meta analysis

System Biology module

Systems Biology Matching Parameters

JOB ID:

JOB NAME:

FILES UPLOADED: [↑ UPLOAD LIST](#)

FileID	Filename	Upload Date	List Type	Accession ID	Metabolic Genes	Remove
664274	VNG_genell	2018-07-18 23:37:35	Genes	Gene symbol	View	×

[Run matching subjobs](#)

METABOLIC PATHWAY MATCHING

PROGRESS (100%)

SUBJOB ID

[View Log](#)

POPULATE METABOLIC PATHWAY TABLE

PROGRESS (100%)

SUBJOB ID

[View Log](#)

System Biology module

Home MRM Databases - Create Job - View Results XCMS Public XCMS Institute Stored Datasets Account Toolbox Help - Logout [hpbenton]

Metabolic Pathway Results

[Pathway Cloud Plot](#)[Show 25 rows](#)[VALUES <> PERCENT](#)[TSV](#)[PDF](#)[Print](#)Search: [Predictive Metabolites Results](#)

Pathway	Overlapping genes	All genes	Overlapping proteins	All proteins	Overlapping putative metabolites ¹	All metabolites ^{2*}	p-values
pyrimidine deoxyribonucleotides <i>de novo</i> biosynthesis	3	35	0	15	3	3	4.6e-3
gluconeogenesis	3	24	0	24	2	2	2.5e-2
pyrimidine deoxyribonucleotides biosynthesis from CTP	3	24	0	17	1	1	1.0e+0
guanosine deoxyribonucleotides <i>de novo</i> biosynthesis	2	12	0	9	0	0	1.0e+0
adenosine deoxyribonucleotides <i>de novo</i> biosynthesis	2	12	0	9	0	0	1.0e+0
glycolysis	1	24	0	24	2	2	2.5e-2
L-dopa degradation	1	1	0	2	2	4	1.2e-1

XCMS-mrm & METLIN-MRM

- First you want to get your transitions



XCMS-mrm & METLIN-MRM

- Search for a metabolite that you want the transition for

METLIN-MRM Search

Show 10 entries

Search:

Search Tip: You do not have to fill out every field. Fields left blank will be ignored during the search.

MID:

Smiles:

Smiles Exact Match

Mass: Min Max

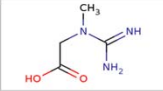
Name:

Name Exact Match

Formula:

CAS:

KEGG:

METLIN ID	Mass	Name	Formula	CAS	KEGG	MRM	MS/MS	Structure
7	131.069476547	Creatine	C4H9N3O2	57-00-1	C00300	View	View	

Showing 1 to 1 of 1 entries

Previous [1](#) Next

XCMS-mrm & METLIN-MRM

- Decided which is the best transition for you...

Statistically Optimized Experimental Transitions

View Selected Fragment(s)

Name: Creatine, MID: 7

Show 10 entries

Precursor	Adduct	Mode	Col. E.	MZ	Rating
<input type="checkbox"/> 130.1	M-H	-	10	88	<input type="checkbox"/> (0) <input type="checkbox"/> (0)
<input type="checkbox"/> 130.1	M-H	-	20	112	<input type="checkbox"/> (0) <input type="checkbox"/> (0)
<input type="checkbox"/> 132.1	M+H	+	10	90.1	<input type="checkbox"/> (0) <input type="checkbox"/> (0)
<input type="checkbox"/> 132.1	M+H	+	10	87.1	<input type="checkbox"/> (0) <input type="checkbox"/> (0)
Precursor	Adduct	Mode	Col. E.	MZ	Rating

Showing 1 to 4 of 4 entries

Previous 1 Next

Agilent

View Selected Fragment(s)

Name: Creatine, MID: 7

Show 10 entries

Precursor	Adduct	Mode	Col. E.	MZ	Rating
<input type="checkbox"/> 130.06168	M-H	-	10	88	<input type="checkbox"/> (0) <input type="checkbox"/> (0)
<input type="checkbox"/> 130.06168	M-H	-	20	112.06	<input type="checkbox"/> (0) <input type="checkbox"/> (0)
<input type="checkbox"/> 132.07728	M+H	+	24	44.1	<input type="checkbox"/> (0) <input type="checkbox"/> (0)
<input type="checkbox"/> 132.07728	M+H	+	12	90	<input type="checkbox"/> (0) <input type="checkbox"/> (0)
Precursor	Adduct	Mode	Col. E.	MZ	Rating

Showing 1 to 4 of 4 entries

Previous 1 Next

Sciex

View Selected Fragment(s)

Name: Creatine, MID: 7

Show 10 entries

Precursor	Adduct	Mode	Col. E.	MZ	Rating
<input type="checkbox"/> 130.1	M-H	-	15	62.1	<input type="checkbox"/> (0) <input type="checkbox"/> (0)
<input type="checkbox"/> 130.1	M-H	-	15	83.2	<input type="checkbox"/> (0) <input type="checkbox"/> (0)
<input type="checkbox"/> 132.07728	M+H	+	24	44.1	<input type="checkbox"/> (0) <input type="checkbox"/> (0)
<input type="checkbox"/> 132.07728	M+H	+	15	90	<input type="checkbox"/> (0) <input type="checkbox"/> (0)
Precursor	Adduct	Mode	Col. E.	MZ	Rating

Showing 1 to 4 of 4 entries

Previous 1 Next

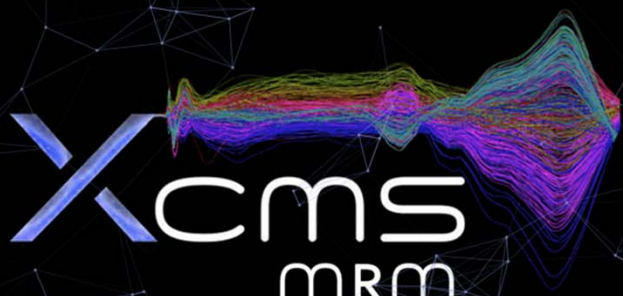
XCMS-mrm & METLIN-MRM

- Decided which is the best transition for you...

https://xcmsonline-mrm.scripps.edu/landing_page.php?pgcontent=mainPage

Getting Started Programming Hacking Science Computer stuff Metabolomics DB Other Stuff Jobs Media Open in Papers Hikes Education

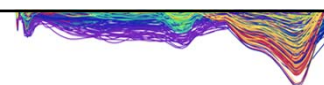
Home Sign Up Forgot Password? Enter email address Enter password Login



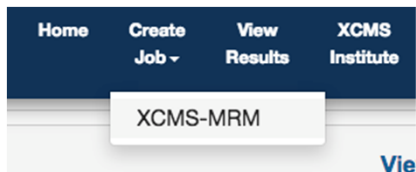
XCMS MRM

XCMS MRM is a cloud-based platform for processing targeted mass spectrometry data

XCMS-mrm & METLIN-MRM



- Run a job the same way we did before



- After uploading your files you'll need a selection of transitions that you have used –

Name	Precursor	Product	RT.min	RT.max	Prec.Labeled	Prod.Labeled
Leucine	132.1	43.096	0.709	1.7		
Leucine	132.1	44.096	0.707	1.7		
Isoleucine	132.1	44.096	0.708	1.705		
Isoleucine	132.1	69.066	0.706	1.703		
Leucine	132.1	86.086	0.704	1.7		
Isoleucine	132.1	86.086	0.705	1.701		
Phenylalanine	166.08	103.096	2.012	3.004		
Phenylalanine	166.08	120.076	2.01	3.003		
Phenylalanine	166.08	130.996	2.007	3.003		
Tyrosine	182.08	136.096	0.607	1.599		
Tyrosine	182.08	146.996	0.604	1.598		
Tyrosine	182.08	165.096	0.601	1.597		
Caffeine	195.08	83.056	2.863	3.867		
Caffeine	195.08	110.066	2.862	3.864		
Caffeine	195.08	138.056	2.861	3.862		