

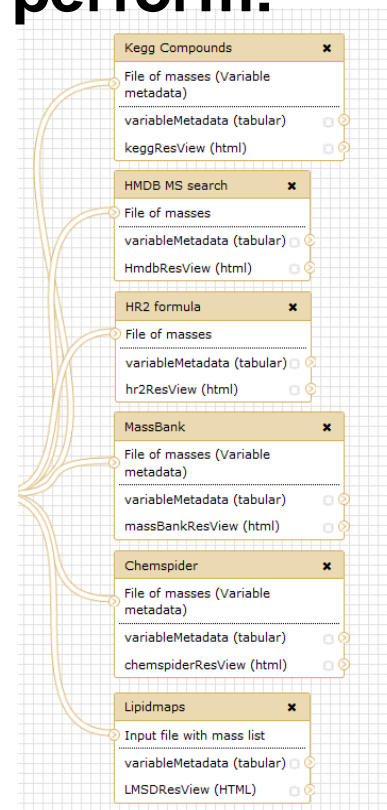
# HOW TO PERFORM LC-MS ANNOTATIONS?

W4M Core Team

# « LC-MS ANNOTATION » MODULES IN W4M

The « LCMS Annotation » modules allow you to perform:

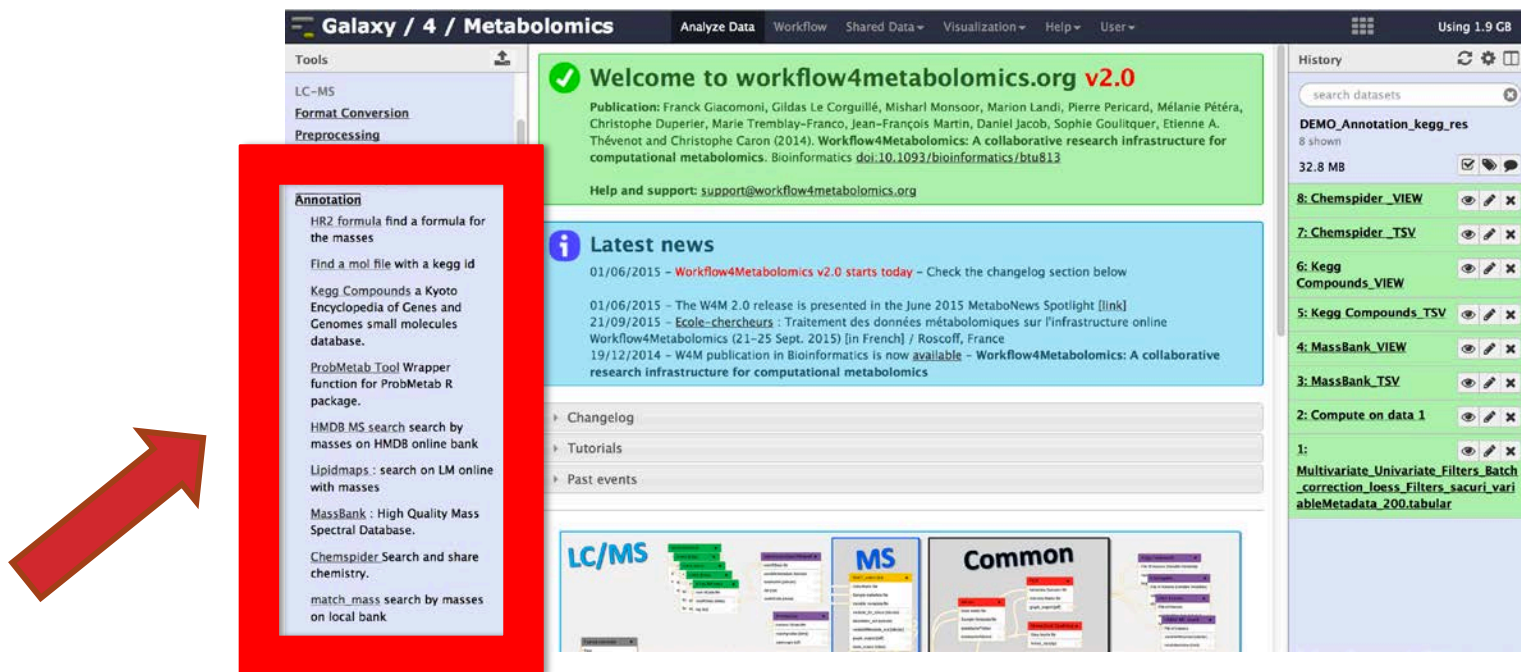
- « Adducts » annotation
- « Chemical » annotation:
  - ✓ Meta-engine (Chemspider)
  - ✓ General banks (Kegg, ChEBI, PubChem...)
  - ✓ Spectral banks (MassBank)
  - ✓ Specialized banks (HMDB, Lipidmaps, ...)
  - ✓ « De novo » tools (HR2 for chemical composition)
- Automatic probabilistic lcms-based metabolome annotation (ProbMetab)



# « LCMS ANNOTATION » MODULES IN W4M

## Selection of the modules:

- LCMS annotation modules are accessible via the left panel:
- > LC-MS
- > 6-Annotation



The screenshot shows the Galaxy / 4 / Metabolomics interface. The left panel lists tools under the 'Annotation' category, which is highlighted with a red box and a red arrow. The main panel displays a welcome message for workflow4metabolomics.org v2.0, latest news, and a workflow diagram. The right panel shows a history of datasets, including 'DEMO\_Annotation\_kegg\_res'.

**Annotation tools:**

- HR2 formula find a formula for the masses
- Find a mol file with a kegg id
- Kegg Compounds a Kyoto Encyclopedia of Genes and Genomes small molecules database.
- ProbMetab Tool Wrapper function for ProbMetab R package.
- HMDB.MS search search by masses on HMDB online bank
- Lipidmaps : search on LM online with masses
- MassBank : High Quality Mass Spectral Database.
- Chemspider Search and share chemistry.
- match\_mass search by masses on local bank

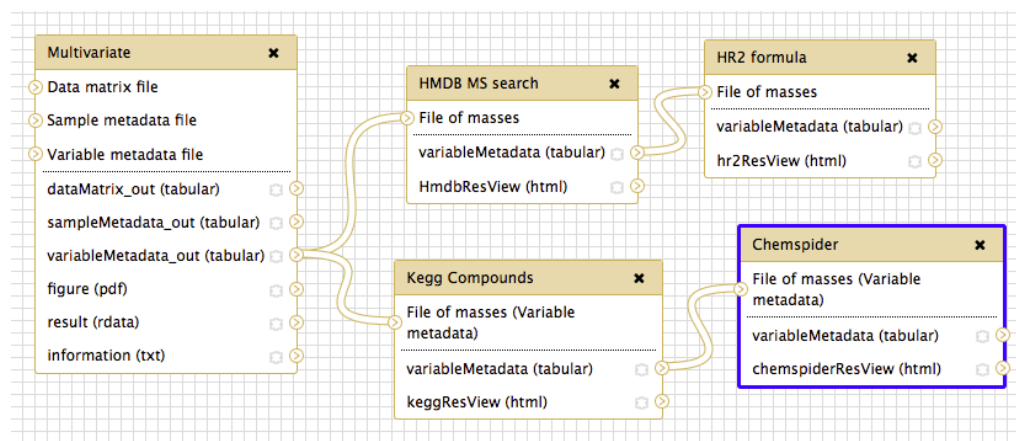
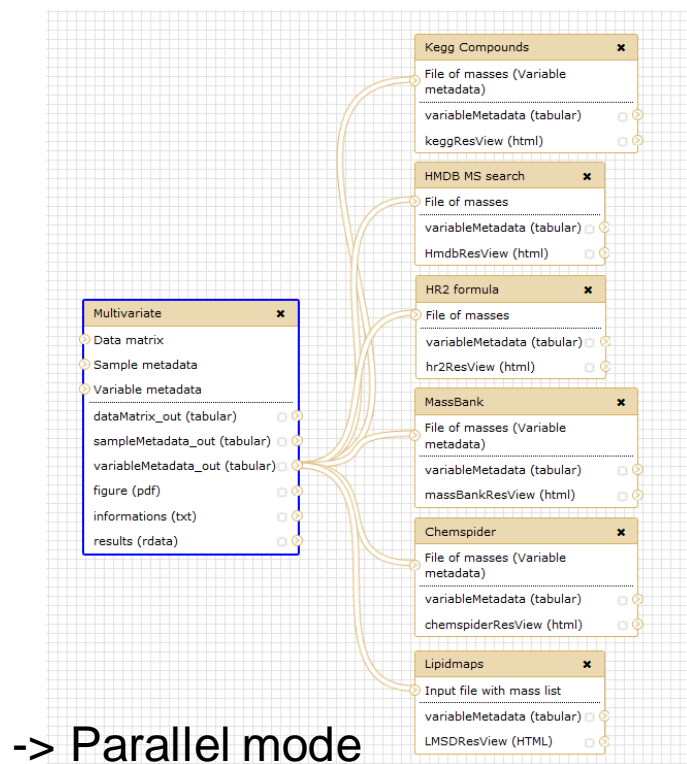
**Workflow diagram:**

The workflow diagram shows three main sections: LC/MS, MS, and Common. The LC/MS section includes tools like 'LC-MS', 'Format Conversion', and 'Preprocessing'. The MS section includes tools like 'MassBank\_VIEW', 'MassBank\_TSV', and 'Compute on data 1'. The Common section includes tools like 'Multivariate\_Univariate\_Filters\_Batch\_correction\_loess\_Filters\_sacuri\_variableMetadata\_200.tabular'.

# CHAINING THE « ANNOTATION » MODULES

The « LCMS Annotation » modules can be chained with « **statistical** » modules and the **Filters** module. ProbMetab can be directly connected with XCMS outputs modules.

Query modules can be run sequentially or in parallel.



# YOUR INPUT DATA

« Chemical Annotation » modules take as input either:

- a list of masses entered manually
- or
- an input file

HMDB MS search (version 2014-05-07)

Would you use a file :

If 'NO' is selected then one or more mass(es) must be entered manually



mass to submit :

You should use dot (.) like decimal separator

HMDB MS search (version 2014-05-07)

Would you use a file :

If 'NO' is selected then one or more mass(es) must be entered manually

File of masses :  

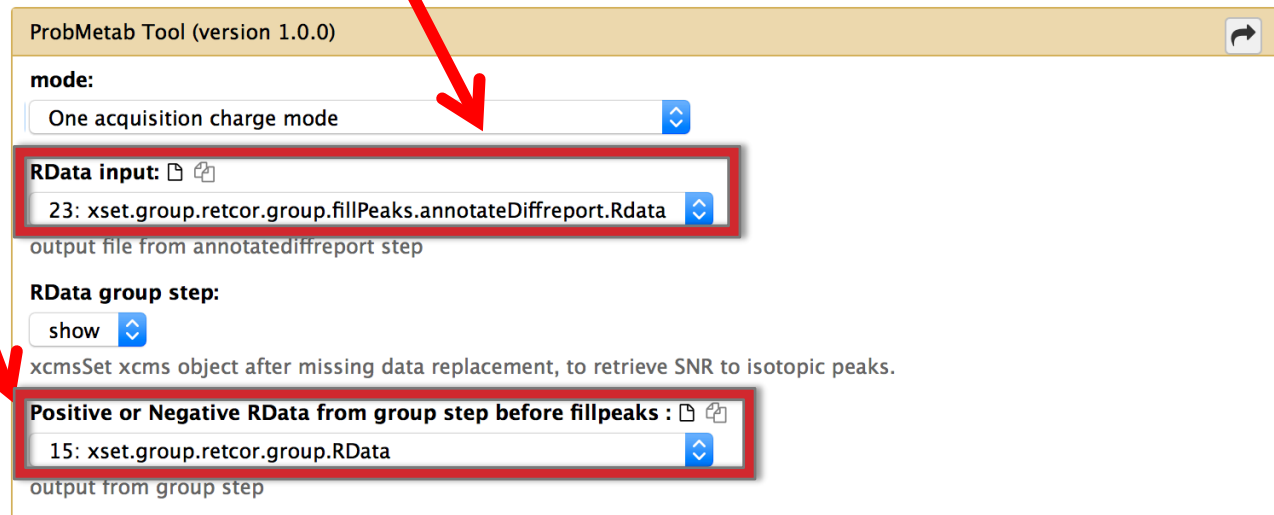
1: sacuri\_variableMetadata\_annotated.tsv

Please note : HMDB allows maximum 150 query masses per request

# YOUR INPUT DATA


« ProbMetab » module take as input either:

- output file from annotatediffreport
- and
- xcmsSet xcms object after missing data replacement, to retrieve SNR to isotopic peaks.



ProbMetab Tool (version 1.0.0)


mode:  
One acquisition charge mode

**RData input:**   
23: xset.group.retcor.group.fillPeaks.annotateDiffreport.Rdata

output file from annotatediffreport step

**RData group step:**  
show

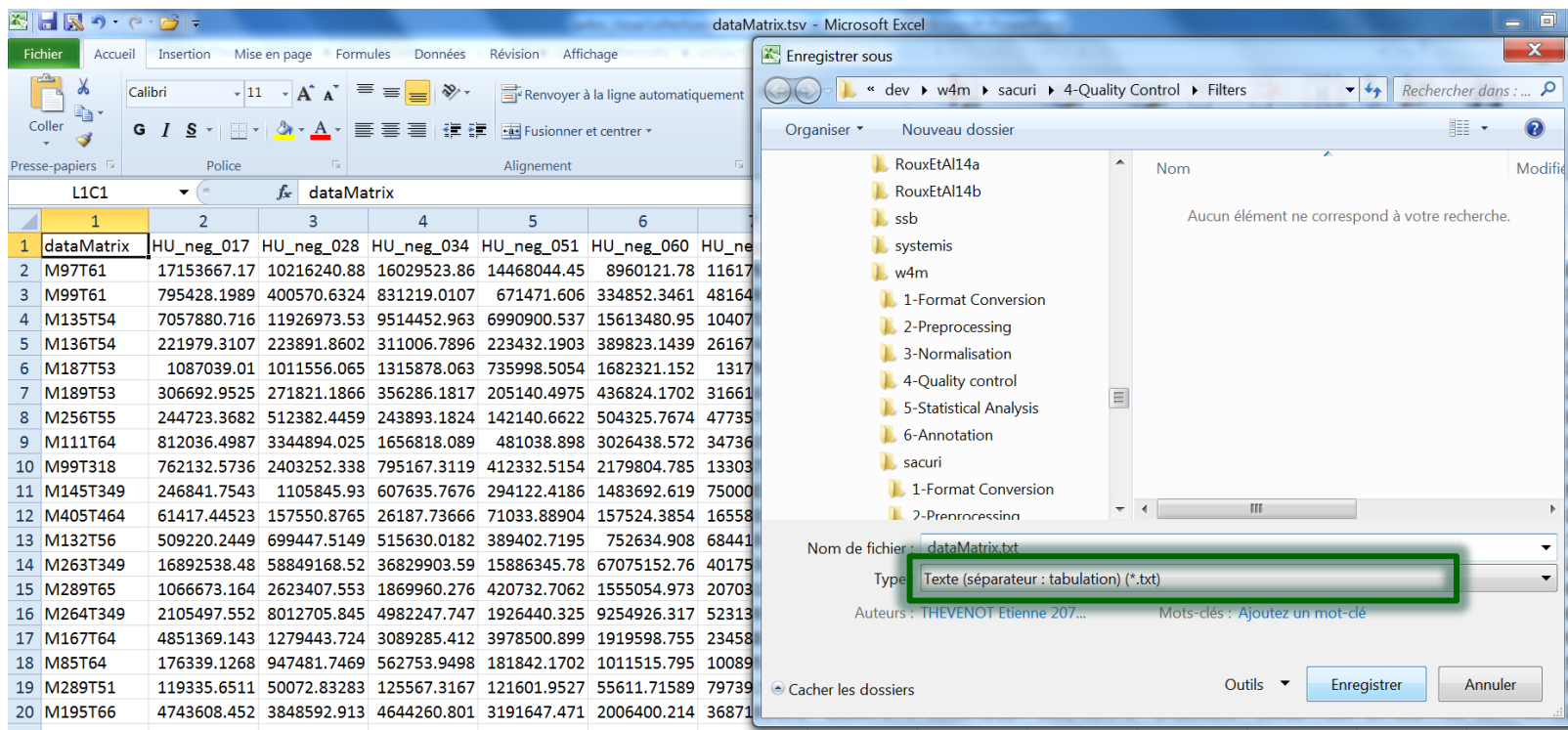
xcmsSet xcms object after missing data replacement, to retrieve SNR to isotopic peaks.

**Positive or Negative RData from group step before fillpeaks :**   
15: xset.group.retcor.group.RData

output from group step

# PREPARING YOUR INPUT FILE (1/3)

The input file format is TSV (tabulation separated values) or TABULAR. It is simple text file that can be prepared by using Excel and saved using the tabulated type format.

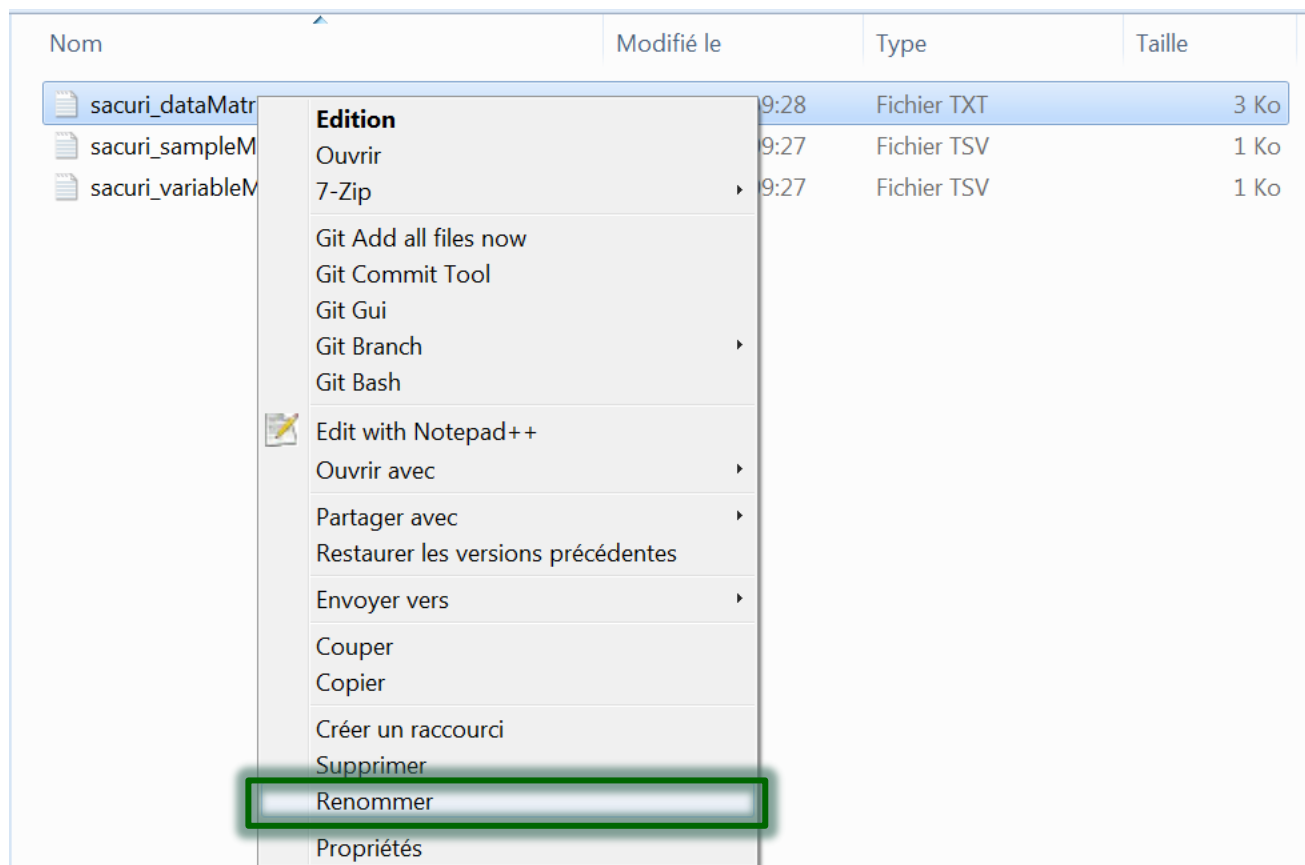


The screenshot shows a Microsoft Excel window titled 'dataMatrix.tsv - Microsoft Excel'. The spreadsheet contains a data matrix with columns labeled 'L1C1', '1', '2', '3', '4', '5', '6', and '7'. The first row is 'dataMatrix', and subsequent rows contain numerical data. A 'Fichier' (File) menu is open, showing options like 'Enregistrer' (Save), 'Enregistrer sous' (Save As), 'Ouvrir' (Open), etc. The 'Enregistrer sous' dialog box is displayed, showing a file explorer view of the 'dev' directory. The file type is set to 'Texte (séparateur : tabulation) (\*.txt)'.

L1C1	1	2	3	4	5	6	7
1	dataMatrix	HU_neg_017	HU_neg_028	HU_neg_034	HU_neg_051	HU_neg_060	HU_neg_071
2	M97T61	17153667.17	10216240.88	16029523.86	14468044.45	8960121.78	11617
3	M99T61	795428.1989	400570.6324	831219.0107	671471.606	334852.3461	48164
4	M135T54	7057880.716	11926973.53	9514452.963	6990900.537	15613480.95	10407
5	M136T54	221979.3107	223891.8602	311006.7896	223432.1903	389823.1439	26167
6	M187T53	1087039.01	1011556.065	1315878.063	735998.5054	1682321.152	1317
7	M189T53	306692.9525	271821.1866	356286.1817	205140.4975	436824.1702	31661
8	M256T55	244723.3682	512382.4459	243893.1824	142140.6622	504325.7674	47735
9	M111T64	812036.4987	3344894.025	1656818.089	481038.898	3026438.572	34736
10	M99T318	762132.5736	2403252.338	795167.3119	412332.5154	2179804.785	13303
11	M145T349	246841.7543	1105845.93	607635.7676	294122.4186	1483692.619	75000
12	M405T464	61417.44523	157550.8765	26187.73666	71033.88904	157524.3854	16558
13	M132T56	509220.2449	699447.5149	515630.0182	389402.7195	752634.908	68441
14	M263T349	16892538.48	58849168.52	36829903.59	15886345.78	67075152.76	40175
15	M289T65	1066673.164	2623407.553	1869960.276	420732.7062	1555054.973	20703
16	M264T349	2105497.552	8012705.845	4982247.747	1926440.325	9254926.317	52313
17	M167T64	4851369.143	1279443.724	3089285.412	3978500.899	1919598.755	23458
18	M85T64	176339.1268	947481.7469	562753.9498	181842.1702	1011515.795	10089
19	M289T51	119335.6511	50072.83283	125567.3167	121601.9527	55611.71589	79739
20	M195T66	4743608.452	3848592.913	4644260.801	3191647.471	2006400.214	36871

# PREPARING YOUR FILES (2/3)

You can then rename your file with the .tsv extension (instead of .txt) by right-clicking on the file (and ignoring the warning):



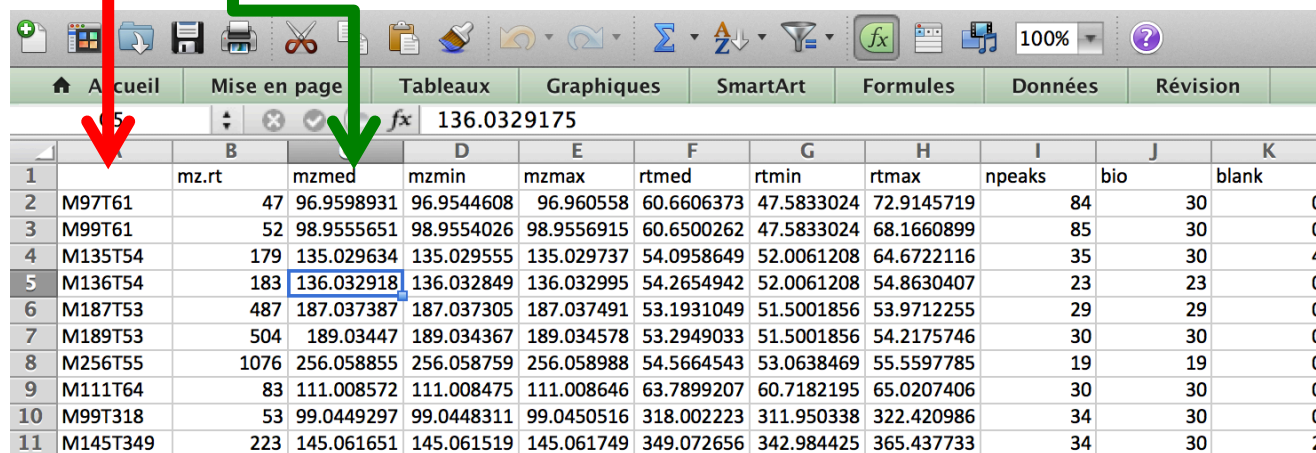


# PREPARING YOUR INPUT FILE (3/3)

Such « .tsv » files (i.e. tabular separated; e.g. **variableMetadata.tsv**) can be handled correctly both by Excel and Galaxy.

The input file structure must contain:


- column with ID of the ion - example : M(mz)T(rt)
- column with masses of ions
- Decimal separator must be "."

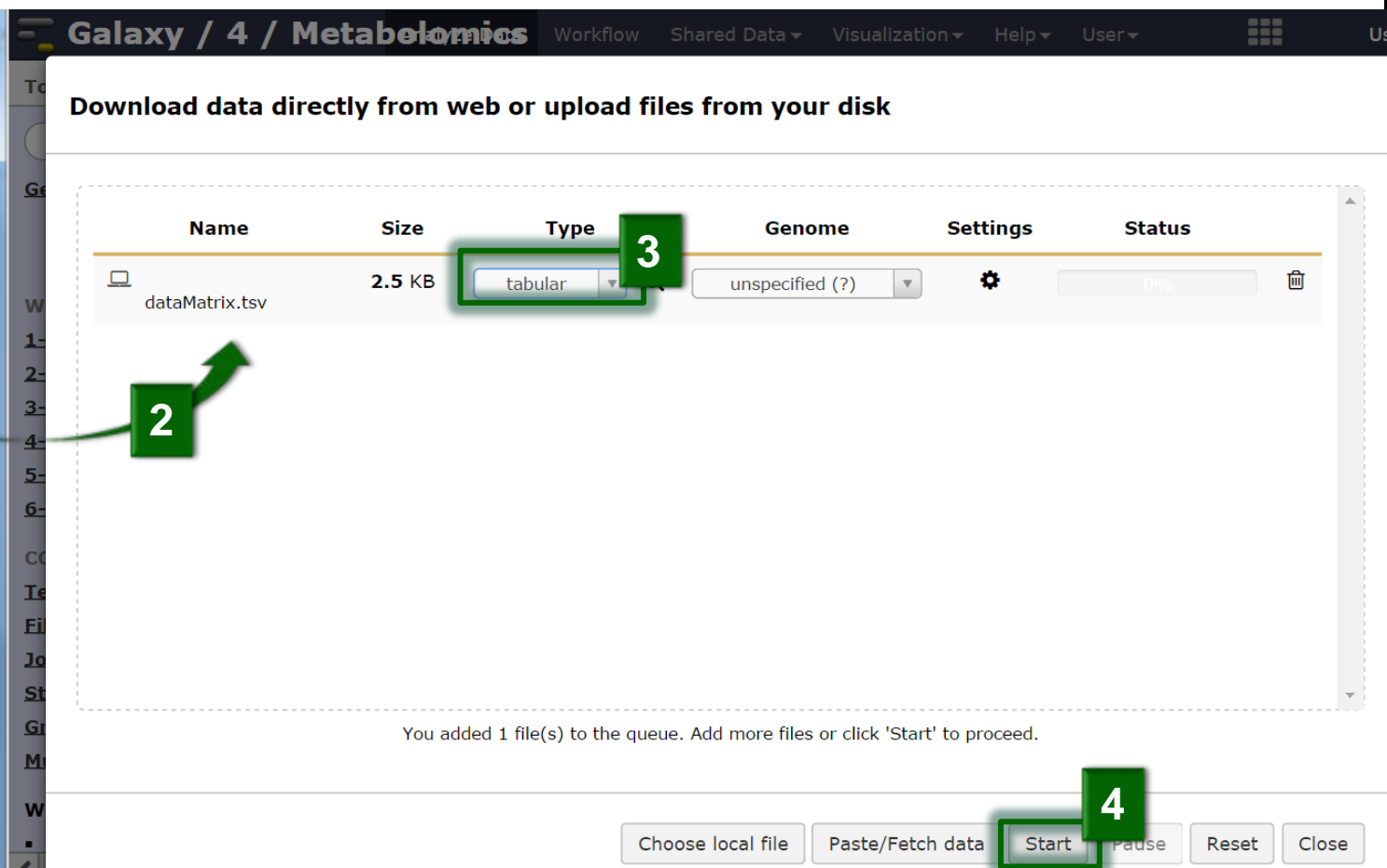
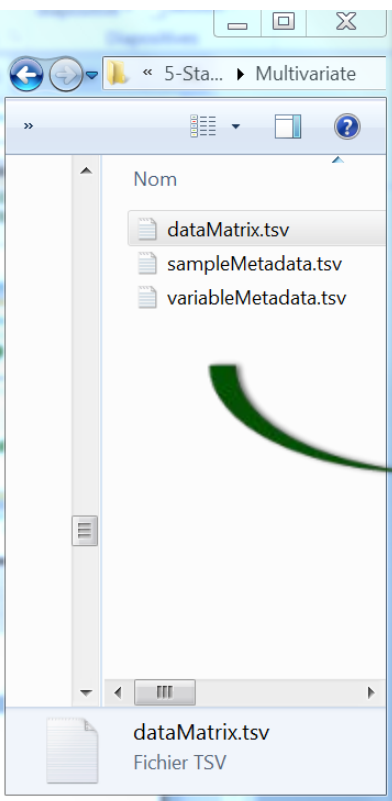


	B	C	D	E	F	G	H	I	J	K
	mz.rt	mzmed	mzmin	mzmax	rtmed	rtmin	rtmax	npeaks	bio	blank
1										
2	M97T61	47	96.9598931	96.9544608	96.960558	60.6606373	47.5833024	72.9145719	84	30
3	M99T61	52	98.9555651	98.9554026	98.9556915	60.6500262	47.5833024	68.1660899	85	30
4	M135T54	179	135.029634	135.029555	135.029737	54.0958649	52.0061208	64.6722116	35	30
5	M136T54	183	136.032918	136.032849	136.032995	54.2654942	52.0061208	54.8630407	23	23
6	M187T53	487	187.037387	187.037305	187.037491	53.1931049	51.5001856	53.9712255	29	29
7	M189T53	504	189.03447	189.034367	189.034578	53.2949033	51.5001856	54.2175746	30	30
8	M256T55	1076	256.058855	256.058759	256.058988	54.5664543	53.0638469	55.5597785	19	19
9	M111T64	83	111.008572	111.008475	111.008646	63.7899207	60.7182195	65.0207406	30	30
10	M99T318	53	99.0449297	99.0448311	99.0450516	318.002223	311.950338	322.420986	34	30
11	M145T349	223	145.061651	145.061519	145.061749	349.072656	342.984425	365.437733	34	30

# LOADING YOUR FILES INTO GALAXY (1/2)

Upload your file (variableMetadata.tsv)

- either by using the icon  and « drag & drop » the file:



- or with the Get Data / Upload File

**Galaxy / 4 / Metabolomics** Workflow Shared Data Visualization Help User

**Tools**

search tools

**Get Data**

Upload File from your computer

**WORKFLOW 4 METABOLOMICS**

**1-Format Conversion**

**2-Preprocessing**

**3-Normalisation**

**4-Quality Control**

**5-Statistical Analysis**

**6-Annotation**

**COMMON TOOLS**

**Text Manipulation**

**Filter and Sort**

**Upload File (version 1.1.4)**

**File Format:** tabular

Which format? See help below

**File:** Choisissez un fichier

TIP: Due to browser limitations, a file larger than 2GB is guaranteed to fail. For very large files, use the URL method if enabled by the site administrator.

**URL/Text:**

Here you may specify a list of URLs or paste the contents of a file.

**History**

**Unnamed history**

0 bytes

**Ouvrir**

« w4m » 5-Statistical Analysis » multivariate » input

Rechercher dans : ...

Organiser Nouveau dossier

Nom

sacuri\_dataMatrix.tsv

sacuri\_sampleMetadata.tsv

sacuri\_variableMetadata.tsv

Nom du fichier : sacuri\_dataMatrix.tsv

Ouvrir Annuler

# CHECK THAT YOUR DATA HAVE BEEN UPLOADED CORRECTLY

**Galaxy / 4 / Metabolomics**
Analyze Data
Workflow
Shared Data
Visualization
Help
User
Using 1.7 GB

**Tools**

search tools

**Get Data**

Upload File from your computer

**WORKFLOW 4 METABOLOMICS**

1-Format Conversion
2-Preprocessing
3-Normalisation
4-Quality Control
5-Statistical Analysis
6-Annotation

**COMMON TOOLS**

Text Manipulation
Filter and Sort
Join, Subtract and Group
Statistics
Graph/Display Data
Multiple regression

""	"mz.rt"	"mzmed"	"mzmin"	"mzmax"	"rtmed"	"rtmin"
"M97T61"	47	96.95989309	96.9544608	96.96055802	60.66063731	47.583302
"M99T61"	52	98.9555651	98.9554026	98.95569154	60.65002621	47.583302
"M135T54"	179	135.0296344	135.0295548	135.0297374	54.09586485	52.00612
"M136T54"	183	136.0329175	136.0328493	136.0329949	54.26549421	52.00612
"M187T53"	487	187.0373874	187.0373051	187.037491	53.19310488	51.50018
"M189T53"	504	189.0344696	189.0343673	189.0345776	53.29490329	51.50018
"M256T55"	1076	256.0588553	256.0587585	256.0589882	54.5664543	53.063846
"M111T64"	83	111.0085715	111.0084745	111.0086463	63.7899207	60.718219
"M99T318"	53	99.04492968	99.04483112	99.04505157	318.0022226	311.95033
"M145T349"	223	145.0616508	145.0615194	145.061749	349.0726556	342.98442
"M405T464"	2027	405.1910066	405.1902768	405.1915084	463.9998039	462.4033
"M132T56"	168	132.0300721	132.0299974	132.0301813	55.93629944	53.977446
"M263T349"	1138	263.1031444	263.1030089	263.1032901	348.9179671	348.01836
"M289T65"	1335	288.9865308	288.9863773	288.9868076	64.78969478	61.541278
"M264T349"	1143	264.106299	264.1061244	264.1064523	348.9415214	347.45344
"M167T64"	358	167.0207709	167.020477	167.0208833	63.67500027	60.718219
"M85T64"	29	85.02943231	85.02936548	85.02948246	63.79149641	60.718219
"M289T51"	1332	288.8159429	288.8157194	288.8161154	51.11847144	49.704977
"M195T66"	549	194.9269118	194.921922	194.9271839	66.16606849	48.1399
"M285T57"	1312	285.0822653	285.0818636	285.0826499	57.38433388	55.653269

**History**

263.6 KB

2: Galaxy1-

[sacuri\_variableMetadata\_annotated.tsv].tabular

2,346 lines
format: tabular, database: ?

uploaded tabular file

1 2 3 4

"" "mz.rt" "mzmed" "mzmin"

"M97T61" 47 96.95989309 96.95446

"M99T61" 52 98.9555651 98.95540

"M135T54" 179 135.0296344 135.0295

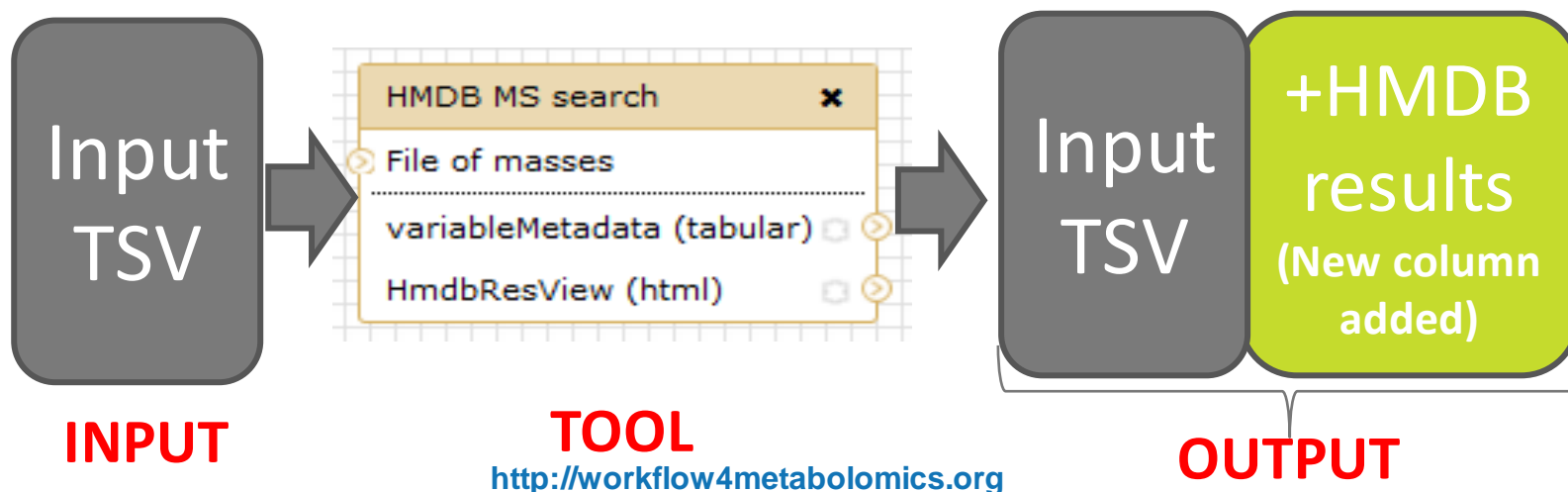
"M136T54" 183 136.0329175 136.0328

"M187T53" 487 187.0373874 187.0373

# USE « LCMS ANNOTATION » MODULES

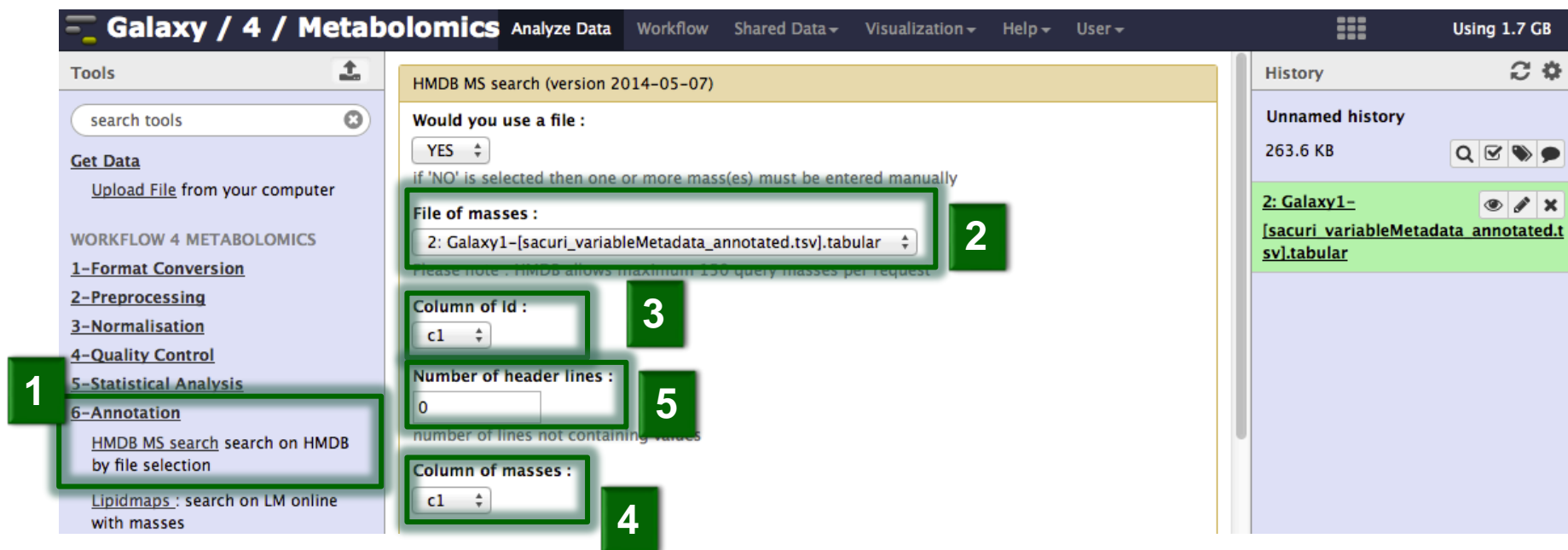
Each module proposes the same input and output formats, and the same type of parameters :

- **The common input and output formats** allow you to run multiple annotation tools sequentially or in parallel.
- **Commons outputs :**
  - HTML view: includes hyperlinks; multiple annotation results for the same variable are split into consecutive rows; next releases will provide more interactivity with the results.
  - .tsv format :



# USE « LCMS ANNOTATION » MODULES

Open the module of your choice and select your input file.  
Select position of your id and mass columns.



The screenshot shows the Galaxy 4 Metabolomics interface. The left sidebar contains a 'Tools' section with a search bar and a list of modules. The '6-Annotation' module is highlighted with a green box and a green circle with the number 1. The main panel shows the 'HMDB MS search (version 2014-05-07)' tool. The 'Would you use a file' dropdown is set to 'YES'. The 'File of masses' dropdown is set to '2: Galaxy1-[sacuri\_variableMetadata\_annotated.tsv].tabular' (highlighted with a green box and a green circle with the number 2). The 'Column of Id' dropdown is set to 'c1' (highlighted with a green box and a green circle with the number 3). The 'Number of header lines' input field is set to '0' (highlighted with a green box and a green circle with the number 5). The 'Column of masses' dropdown is set to 'c1' (highlighted with a green box and a green circle with the number 4). The right sidebar shows the 'History' section with an 'Unnamed history' entry of 263.6 KB, which is highlighted with a green box and a green circle with the number 2.

Indicate if your file contains header lines.

**You are now ready to configure your annotation tool!**

## Example of HMDB input and configuration :

HMDB MS search (version 2014-05-07)

**Would you use a file :**  
   
 if 'NO' is selected then one or more mass(es) must be entered manually

**File of masses :**  
   
 Please note : HMDB allows maximum 150 query masses per request

**Column of Id :**

**Number of header lines :**  
   
 number of lines not containing values

**Column of masses :**

**Molecular Weight Tolerance +/- :**  
   
 Default value is 0.05 for HMDB (Da)

**Molecular Species :**  
☐ Positif Mode  
☐ Negatif Mode  
☒ Neutral Mass

Data input: a variableMetadata file with mandatory id and mz columns

Search parameters: tolerance and molecular species

**Authors** Marion Landi [marion.land@clermont.inra.fr](mailto:marion.land@clermont.inra.fr) and Franck Giacomoni [franck.giacomoni@clermont.inra.fr](mailto:franck.giacomoni@clermont.inra.fr)

**Please cite If you use this tool, please cite**

Wishart DS, Jewison T, Guo AC, Wilson M, Knox C, et al., HMDB 3.0—The Human Metabolome Database in 2013. Nucleic Acids Res. 2013.



## Annotation tool results – HTML view

Results of HMDB queries - Search params : Molecular specie = negative / delta = 0.001

Metadata  
of your query

Prev 1 2 3 4 Next

ID from input	Mass (m/z)	Compound_ID	Formula	Compound MW (Da)	Adduct	Adduct_Type	Adduct MW (Da)	Delta
M99T318	99.04492968							
		<a href="#">HMDB41792</a>	C4H6	54.04695	M+FA-H	-	99.045151	0.00022132
		<a href="#">HMDB01470</a>	C5H8O2	100.05243	M-H	-	99.045154	0.00022432
		<a href="#">HMDB32351</a>	C5H8O2	100.05243	M-H	-	99.045154	0.00022432
		<a href="#">HMDB01862</a>	C5H8O2	100.05243	M-H	-	99.045154	0.00022432
		<a href="#">HMDB32459</a>	C5H8O2	100.05243	M-H	-	99.045154	0.00022432
		<a href="#">HMDB30989</a>	C10H16O4	200.104859	M-2H	-	99.045154	0.00022432
M145T349	145.0616508							
		<a href="#">HMDB33952</a>	C4H8N2O	100.063663	M+FA-H	-	145.061864	0.0002132
		<a href="#">HMDB31642</a>	C4H8N2O	100.063663	M+FA-H	-	145.061864	0.0002132
		<a href="#">HMDB02031</a>	C5H10N2O3	146.069142	M-H	-	145.061866	0.0002152
		<a href="#">HMDB00641</a>	C5H10N2O3	146.069142	M-H	-	145.061866	0.0002152
		<a href="#">HMDB06899</a>	C5H10N2O3	146.069142	M-H	-	145.061866	0.0002152
		<a href="#">HMDB28687</a>	C5H10N2O3	146.069142	M-H	-	145.061866	0.0002152
		<a href="#">HMDB03423</a>	C5H10N2O3	146.069142	M-H	-	145.061866	0.0002152
		<a href="#">HMDB40613</a>	C26H30O6	438.204239	M-3H	-	145.060804	0.0008468
		<a href="#">HMDB36597</a>	C26H30O6	438.204239	M-3H	-	145.060804	0.0008468
M405T464	405.1910066							
		<a href="#">HMDB39013</a>	C22H32O8	424.209718	M-H2O-H	-	405.191328	0.0003214
		<a href="#">HMDB34492</a>	C22H32O8	424.209718	M-H2O-H	-	405.191328	0.0003214
		<a href="#">HMDB40795</a>	C20H26O5	346.178024	M+Hac-H	-	405.191875	0.0008684
		<a href="#">HMDB12613</a>	C21H28O5	360.193674	M+FA-H	-	405.191875	0.0008684
		<a href="#">HMDB41795</a>	C20H26O5	346.178024	M+Hac-H	-	405.191875	0.0008684
		<a href="#">HMDB36901</a>	C20H26O5	346.178024	M+Hac-H	-	405.191875	0.0008684



## Annotation tool results – HTML links

HMDB Browse Search Downloads About Contact Us Search Metabolites Q Search

Showing metabocard for Tiglic acid (HMDB01470)

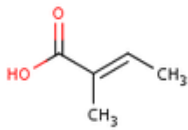
Identification Taxonomy Ontology Physical properties Spectra Biological properties Concentrations Links References XML

Show Metabolites with Similar Structures

**Record Information**

Version	3.6
Creation Date	2005-11-16 15:48:42 UTC
Update Date	2013-02-09 00:10:36 UTC
HMDB ID	HMDB01470
Secondary Accession Numbers	None

**Metabolite Identification**

Common Name	Tiglic acid
Description	Tiglic acid is a thick, syrupy poisonous liquid, C <sub>5</sub> H <sub>8</sub> O <sub>2</sub> , derived from croton oil, having a spicy odor and used in making perfumes and flavoring agents.
Structure	 <p>MOL SDF PDB SMILES InChI View Structure</p>

Product W (Da)	Delta
104.045151	0.00022132
104.045154	0.00022432
104.045154	0.00022432
104.045154	0.00022432
104.045154	0.00022432
104.045154	0.00022432
106.061864	0.0002132
106.061864	0.0002132
106.061866	0.0002152
106.061866	0.0002152
106.061866	0.0002152
106.061866	0.0002152
106.061866	0.0002152
106.061866	0.0002152
106.060804	0.0008468
106.060804	0.0008468
118.06191328	0.0003214
118.06191328	0.0003214
118.06191875	0.0008684
118.06191875	0.0008684
118.06191875	0.0008684
118.06191875	0.0008684

**.tsv output format: multiple annotations for the same variable are concatenated within the same cell using "::" as separator (Metlin format like)**

	mz.rt	mzmed	rtmed	npeaks	bio	blank	isotopes	adduct	pcgroup	hmdb
M97T61	47	96.95989309	60.66063731	84	30	0		NA	57	"No_result_found_on HMDB"
M99T61	52	98.9555651	60.65002621	85	30	0		NA	57	"No_result_found_on HMDB"
M135T54	179	135.0296344	54.09586485	35	30	4	[13][M]+	NA	81	0.0002606::(C3H6O3)::HMDB01882 0.0002606::(C3H6O3)
M136T54	183	136.0329175	54.26549421	23	23	0	[13][M+1]+	NA	81	0.0001435::(C10H14N2O7)::HMDB28818
M187T53	487	187.0373874	53.19310488	29	29	0		NA	138	0.0002726::(C9H10O3)::HMDB32030 0.0002726::(C9H10O3)
M189T53	504	189.0344696	53.29490329	30	30	0		NA	413	"No_result_found_on HMDB"
M256T55	1076	256.0588553	54.5664543	19	19	0		NA	81	0.0002297::(C10H9N)::HMDB33115 0.0002297::(C10H9N)
M111T64	83	111.0085715	63.7899207	30	30	0	[6][M]+	NA	4	0.0001965::(C5H4O3)::HMDB00617 0.0001965::(C5H4O3)
M99T318	53	99.04492968	318.0022226	34	30	0	[3][M]+	NA	164	0.00022132::(C4H6)::HMDB41792 0.00022432::(C5H8O2)
M145T349	223	145.0616508	349.0726556	34	30	2		NA	5	0.0002132::(C4H8N2O)::HMDB33952 0.0002132::(C4H8N2O)
M405T464	2027	405.1910066	463.9998039	29	29	0	[219][M]+	NA	175	0.0003214::(C22H32O8)::HMDB39013 0.0003214::(C22H32O8)
M132T56	168	132.0300721	55.93629944	30	30	0		NA	109	0.0001569::(C3H5NO2)::HMDB39426 0.0001569::(C3H5NO2)
M263T349	1138	263.1031444	348.9179671	35	30	5	[121][M]+	NA	5	0.0000376::(C13H18N2O5)::HMDB61146 0.0000376::(C13H18N2O5)
M289T65	1335	288.9865308	64.78969478	30	30	0		NA	167	"No_result_found_on HMDB"
M264T349	1143	264.106299	348.9415214	34	30	4	[121][M+1]+	NA	5	0.000072::(C13H17NS)::HMDB40032

"No\_result\_found\_on HMDB"

0.0002606::(C3H6O3)::HMDB01882|0.0002606::(C3H6O3)

**« METLIN FORMAT LIKE »**

0.003604744502::(Clitocine)::HMDB33718

0.003605255498::(Citrusinine II)::HMDB30373

**DELTA**

**COMPOUND\_NAME**

**HMDB\_ID**

# « COMPARE » ANNOTATION

You can use the multi-view tool of W4M and visualize your annotation files simultaneously.

Using 860.8 MB

History

RFMF\_TP03\_kegg\_res  
32.8 MB

8: Chempidder VIEW

7: Chempidder VIEW

6: Kegg Compounds VIEW

5: Kegg Compounds TSV

4: MassBank VIEW

3: MassBank TSV

2: Compute on data

Galaxy / Metabolomics

workflow4metabolomics.org/root?workflow\_id=706ef150c22aa1d1

Galaxy / Metabolomics

Analyze Data Workflow Shared Data Visualization Help User

Data Viewer: Kegg Compounds\_TSV

massbank

Data Viewer: MassBank\_TSV

Data Viewer: MassBank\_VIEW

input	96.95989309							
M97T61	96.95989309							
		<a href="#">KO000546</a>	Cimetidine	C10H16N6S	252.11572	-155.1558	1.0000	LC-ESI-QQ; MS2; CE:30 V; [M-H]-
		<a href="#">KO001997</a>	Valproic acid	C8H16O2	144.11503	-47.1551	1.0000	LC-ESI-QQ; MS2; CE:20 V; [M-H]-
		<a href="#">KNA00765</a>	Glycerone phosphate	C3H7O6P	169.99802	-73.0381	1.0000	LC-ESI-ITFT; MS2; m/z:168.99; NEG
		<a href="#">EA034757</a>	Terbutylazine-2-hydroxy	C9H17N5O	211.1433	-114.1834	1.0000	LC-ESI-ITFT; MS2; CE:90%; R=7500; [M-H]-
		<a href="#">KO000547</a>	Cimetidine	C10H16N6S	252.11572	-155.1558	0.9786	LC-ESI-QQ; MS2; CE:40 V; [M-H]-
		<a href="#">JP005200</a>	FURFURAL	C5H4O2	96.02113	0.9388	0.9693	CI-B; MS
		<a href="#">KO000756</a>	Fructose 1-phosphate	C6H13O9P	260.02972	-163.0698	0.9601	LC-ESI-QQ; MS2; CE:30 V; [M-H]-

# FORMAT YOUR MZ (IF NECESSARY)

## positive / negative ionization modes:

- Some tools are not mass spectrometry oriented (ex: KEGG)
- It is therefore necessary to format your mz values with COMMON TOOLS / Compute prior to using them
  - positive mode:  $C3 - 1.007825$  (Proton) +  $0.0005486$  (electron)
  - negative mode:  $C3 + 1.007825$  (Proton) -  $0.0005486$  (electron)

Compute (version 1.1.0)

**Add expression:**

**as a new column to:**  

  
Dataset missing? See TIP below

**Round result?:**

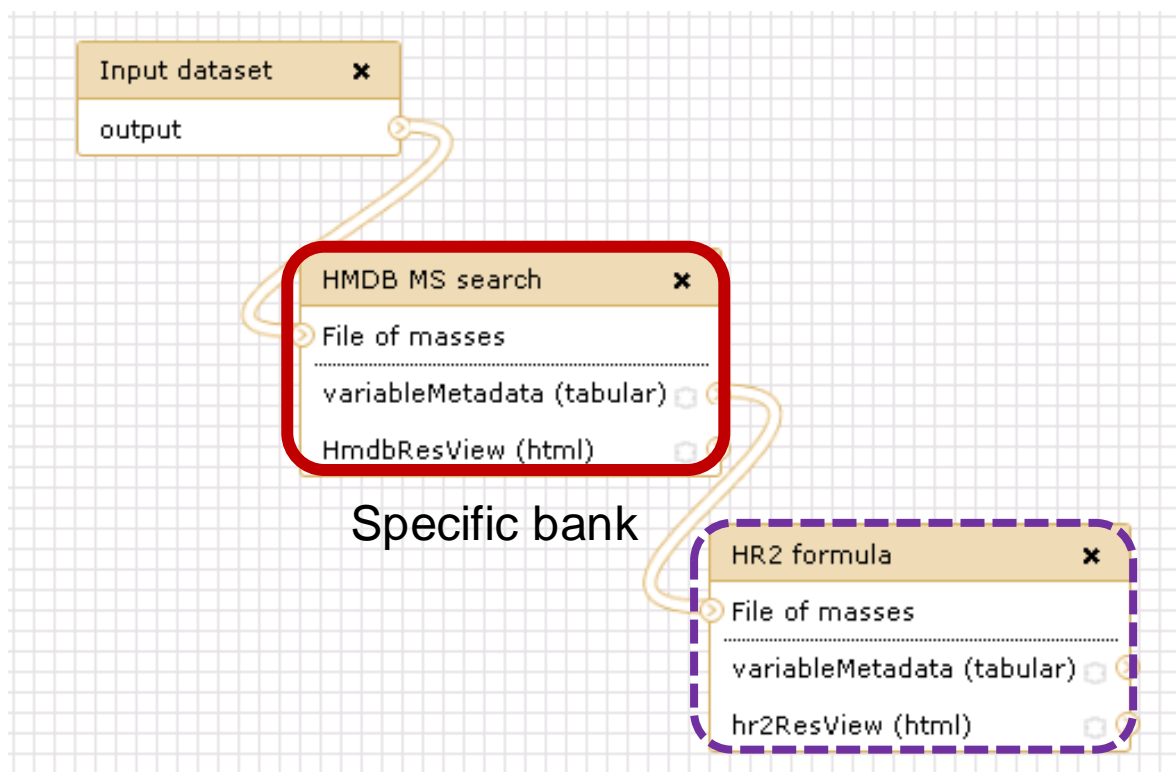
Execute



Mzmed [M+H]	M
132.0300721	131.0227957
263.1031444	262.095868
288.9865308	287.9792544
264.106299	263.0990226

## Example of a sequential « Chemical » annotation workflow:

Chain a specific bank query and a formula elucidator



### 6-Annotation

- [HMDB MS search](#) search on HMDB by file selection
- [Lipidmaps](#) : search on LM online with masses
- [HR2 formula](#) find a formula for the masses
- [Find a mol file](#) with a kegg id
- [Kegg Compounds](#) a Kyoto Encyclopedia of Genes and Genomes small molecules database.
- [Chemspider](#) Search and share chemistry.
- [MassBank](#) : High Quality Mass Spectral Database.
- [CAMERA.annotateDiffreport](#) Wrapper function for the xcms diffreport and the annotate function. Returns a diffreport within the annotation results.

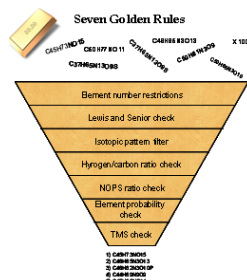
# « CHEMICAL » ANNOTATION WORKFLOW FEATURES

## Human Metabolome DataBase:



- Bank containing small molecule metabolites found in the human body.
- V3.6 with 41,815 entries + 5,688 proteins.

## HR2:



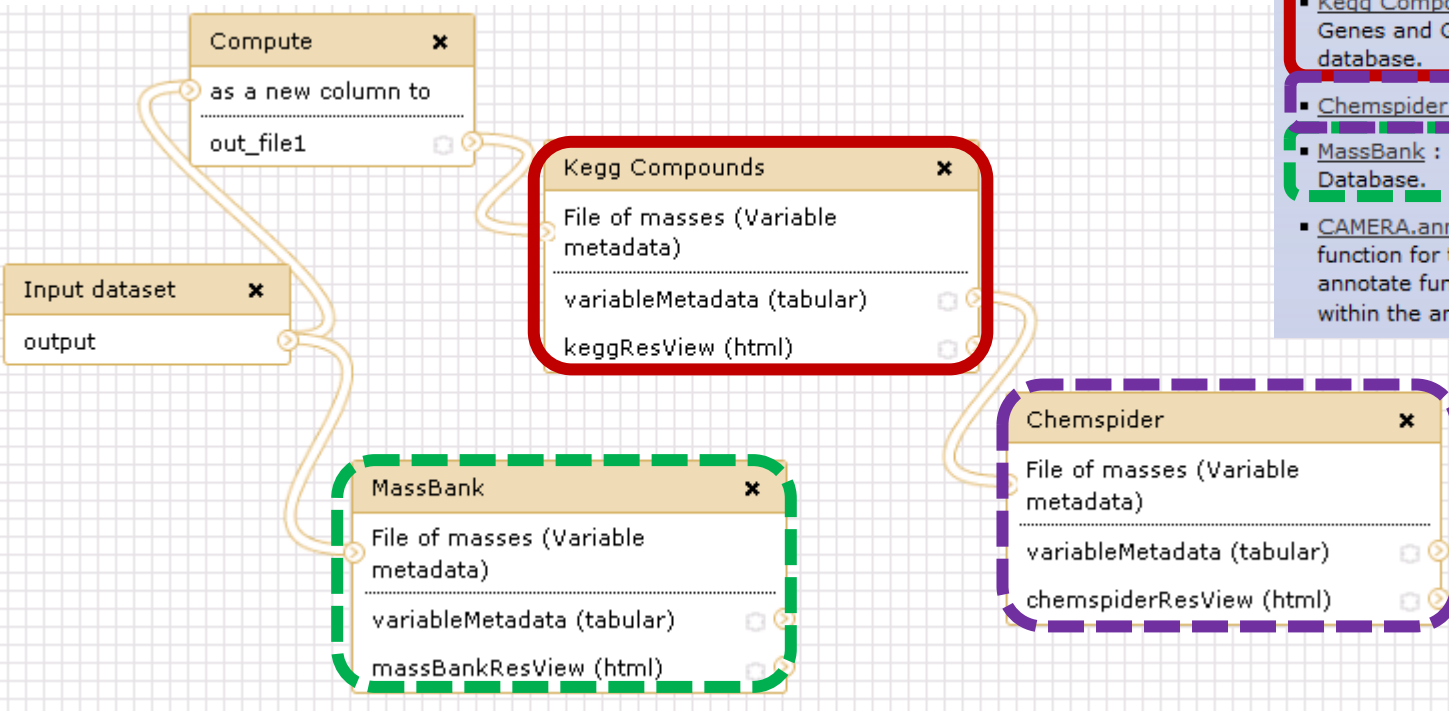
- Molecular formula generator: HiRes ("High Resolution") MS by Joerg Hau.
- Generate random chemical formula and filter them with "the seven golden rules"

## Example of a parallel annotation workflow:

Combine chemical compound libraries  
and a spectral repository.

### 6-Annotation

- [HMDB MS search](#) search on HMDB by file selection
- [Lipidmaps](#) : search on LM online with masses
- [HR2 formula](#) find a formula for the masses
- [Find a mol file](#) with a kegg id
- [Kegg Compounds](#) a Kyoto Encyclopedia of Genes and Genomes small molecules database.
- [Chemspider](#) Search and share chemistry.
- [MassBank](#) : High Quality Mass Spectral Database.
- [CAMERA.annotateDiffreport](#) Wrapper function for the xcms diffreport and the annotate function. Returns a diffreport within the annotation results.



## KEGG Compounds:



- Sub-bank of KEGG containing small molecules, biopolymers, and other chemical substances, relevant to biological systems. V2014-05 with 17,254 entries.

## MassBank:



- Public repository of mass spectra of small chemical compounds for life sciences.
- V2014-03 with 40,889 spectra for 15,775 compounds

## ChemSpider



- Meta engine providing text and structure search to over 30 millions structures from ~500 data sources.



- Wishart DS, Jewison T, Guo AC, Wilson M, Knox C, et al., HMDB 3.0—The Human Metabolome Database in 2013. *Nucleic Acids Res.* 2013.
- Silva, R. R. and Jourdan, F. and Salvanha, D. M. and Letisse, F. and Jamin, E. L. and Guidetti-Gonzalez, S. and Labate, C. A. and Vencio, R. Z. N. (2014). ProbMetab: an R package for Bayesian probabilistic annotation of LC-MS-based metabolomics. In *Bioinformatics*, 30 (9), pp. 1336–1337.
- Sud M., Fahy E., Cotter D., Brown A., Dennis E., Glass C., Murphy R., Raetz C., Russell D., and Subramaniam S. LMSD: LIPID MAPS structure database. *Nucleic Acids Research* 35, D527-32 (2006).
- Tobias Kind and Oliver Fiehn. (2007). "Seven Golden Rules for heuristic filtering of molecular formulas obtained by accurate mass spectrometry." *BMC Bioinformatics* p8:105
- Kanehisa, M., S. Goto, et al. (2012). "KEGG for integration and interpretation of large-scale molecular data sets." *Nucleic Acids Research* 40(D1): D109-D114
- Pence, H.E. and A.J. Williams, (2010). "ChemSpider: An Online Chemical Information Resource." *J.Chem. Educ.*, 87(11): p. 1123-1124.
- H. Horai, M. Arita, S. Kanaya, Y. Nihei, T. Ikeda, K. Suwa. Y. Ojima, K. Tanaka, S. Tanaka, K. Aoshima, Y. Oda, Y. Kakazu, M. Kusano, T. Tohge, F. Matsuda, Y. Sawada, M. Yokota Hirai, H. Nakanishi, K. Ikeda, N. Akimoto, T. Maoka, H. Takahashi, T. Ara, N. Sakurai, H. Suzuki, D. Shibata, S. Neumann, T. Iida, K. Tanaka, K. Funatsu, F. Matsuura, T. Soga, R. Taguchi, K. Saito and T. Nishioka, (2010). "MassBank: A public repository for sharing mass spectral data for life sciences." *J. Mass Spectrom.*, 45, 703-714.

END !