

GBS 724
2-10-16

Pathway software and databases

Stephen Barnes

Synopsis

- **Public and proprietary metabolite databases**
 - Instrument manufacturers provide software and search tools for their proprietary databases
 - Usually a single copy of the software is provided
 - Sciex and Waters provide additional copies at >\$35,000 for each
 - Metabolomics service companies have proprietary databases and do not allow investigators submitting samples to have access to them
 - This discourages deeper research to be performed on metabolomics data

Public databases

- North America
 - Metlin (<https://metlin.scripps.edu/index.php>)
 - Human metabolomics database <http://www.hmdb.ca/>
 - PubChem <http://pubchem.ncbi.nlm.nih.gov/>
- Europe
 - ChemSpider (<http://www.chemspider.com/>)
- Asia
 - Kyoto Encyclopedia of Genes and Genomes (KEGG) (<http://www.genome.jp/kegg/>)

Choosing the ions to search

- Go to the Download file from Metaboanalyst and open the file "peak_normalized_rt_mz.csv"
- Use Excel to calculate the p-values for all the metabolites (they're ordered by m/z in the file)
- Now sort the entire file by the p-values
- Choose the ions with p-values <0.001 (there should be 80) and copy them into a Word file and save the file
 - We'll search these in multiple databases

Metlin

METLIN: Metabolite Search

Batch

[Simple \(Saved Searches\)](#) |
 [Advanced](#) |
 [Batch](#) |
 [Fragment](#) |
 [Neutral Loss](#) |
 [MS/MS Spectrum Match](#) |
 [Unknowns](#)

Masses:

Charge:

Accuracy (ppm):

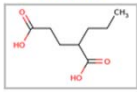
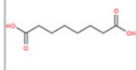
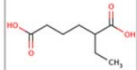
Display Structure:

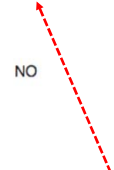
Remove peptides from search:

*To select multiple Adducts:
 - Hit Ctrl + Adducts
 - Hit Command + Adducts
 Select: **all** | none

Metlin output-1

173.08163 m/z
 (174.0872 - 174.0906 daltons): 7 Metabolites [M-H]⁻

MetlinID	Mass	Δppm	Name	Formula	CAS	MS/MS	Structure
2998	174.0892	1	2-Propylglutaric acid	C8H14O4	32806-62-5	NO	
4243	174.0892	1	Suberic acid	C8H14O4	505-48-6	<input type="button" value="View"/>	
45928	174.0892	1	Ethyladipic acid	C8H14O4		NO	

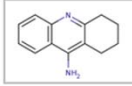
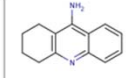
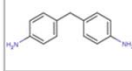
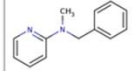


Has MSMS data in this record

Metlin output-2

197.11032 m/z

(198.1156 - 198.1196 daltons): 4 Metabolites [M-H]⁻

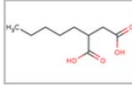
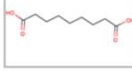
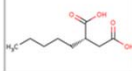
MetlinID	Mass	Δppm	Name	Formula	CAS	MS/MS	Structure
44521	198.1157	9	9-AMINO-1,2,3,4-TETRAHYDROACRIDINE	C13H14N2	NA	View	
2616	198.1157	9	Tacrine	C13H14N2	321-64-2	View	
69942	198.1157	9	4,4'-Methylenedianiline	C13H14N2	101-77-9	NO	
70631	198.1157	9	N-(Phenylmethyl)-N-methyl-2-pyridinamine	C13H14N2		NO	

Unlikely fits – high Δ ppm

Metlin output-3

187.09745 m/z

(188.1029 - 188.1066 daltons): 8 Metabolites [M-H]⁻

MetlinID	Mass	Δppm	Name	Formula	CAS	MS/MS	Structure
44183	188.1049	0	NONIC ACID	C9H16O4		View	
5750	188.1049	0	Nonanedioic acid	C9H16O4		View	
62450	188.1049	0	Nonate	C9H16O4		NO	

Spectra Search Mass Spectrum

MS Search **MS/MS Search** GC/MS Search 1D NMR Search 2D NMR Search **HMDB search**

Query Masses (Da)

165.0555
425.04502
658.45859
201.11197
501.1806
329.11214
177.05641
269.04555
426.03226
226.03738
199.09784
227.09225
193.05045
490.13636
608.13368

Enter one mass per line (maximum 150 query masses per request)

Ionization

Ion Mode
Negative

Adduct Type
M-H
M+Na-2H
M+Cl
M+K-2H
M+FA-H
M+Hac-H
M+Br

Hold Ctrl (⌘) or Command (⌘) to select multiple adducts

Molecular Weight Tolerance ±
10 ppm

Search **Load Example**

HMDB search output-1

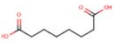
MS search for 173.08163 m/z Delta = abs(query mass - adduct mass)

Show 10 entries Search

Compound	Name	Adduct	Adduct MW (Da)	Compound MW (Da)	Delta
HMDB00893	Suberic acid	M-H	173.081933	174.089208936	0.000303
HMDB02023	Ethyladipic acid	M-H	173.081933	174.089208936	0.000303
HMDB33838	Diethyl succinate	M-H	173.081933	174.089208936	0.000303
HMDB59727	2,4-Dimethyladipic acid	M-H	173.081933	174.089208936	0.000303
HMDB59757	3-Methylpimelic acid	M-H	173.081933	174.089208936	0.000303
HMDB60684	2-Propylglutaric acid	M-H	173.081933	174.089208936	0.000303

Showing 1 to 6 of 6 entries Previous 1 Next

Click on HMDB record

Record Information	
Version	3.6
Creation Date	2005-11-16 15:48:42 UTC
Update Date	2015-12-02 05:10:07 UTC
HMDB ID	HMDB00893
Secondary Accession Numbers	<ul style="list-style-type: none"> • HMDB00837
Metabolite Identification	
Common Name	Suberic acid
Description	Suberic acid, also octanedioic acid, is a dicarboxylic acid, with formula $C_8H_{12}(COOH)_2$. It is present in the urine of patients with Fatty Acid Oxidation Disorders (PMID 10404733 ↗). A metabolic breakdown product derived from oleic acid. Elevated levels of this unsaturated dicarboxylic acid are found in individuals with dicarboxylic acid and medium-chain acyl-CoA dehydrogenase deficiency (MCAD).
Structure	 <div style="border: 1px solid black; padding: 2px; display: flex; gap: 5px;"> 🔍 MOL SDF PDB SMILES InChI </div>
Synonyms	<ol style="list-style-type: none"> 1,6-Dicarboxyhexane 1,6-Hexanedicarboxylate 1,6-Hexanedicarboxylic acid 1,8-Octanedioate 1,8-Octanedioic acid Cork acid

ChemSpider
Search ChemSpider

Search and share chemistry

Simple Structure Advanced History

Advanced search

Structure

Identifier

Elements

Intrinsic Properties ←

Calculated Properties

Data Source

Lasso Similarity

Supplementary Info

Tags

FILTER ▾
Search Hits Limit: 100 ▾
CLEAR FORM
SEARCH

Searching in ChemSpider

Monoisotopic Mass: ±
 min/max +/-

Calculated Properties

Data Source

Lasso Similarity

Supplementary Info

Tags

FILTER

Search Hits Limit: 100

CLEAR FORM

SEARCH

ChemSpider output

Found 1374 results

Search term: MM >= 174.083906 AND MM <= 174.093906 AND abs(Monoisotopic_Mass - 174.088906451991) as mass_defect



1 2 3 4 5

ID	Structure	Molecular Formula	Molecular Weight	# of Data Sources	# of References	# of PubMed	# of RSC	Mass defect
10025 		C ₈ H ₁₄ O ₄	174.1944	109	264	73	275	0.0003
11364 		C ₈ H ₁₄ O ₄	174.1944	76	121	0	173	0.0003
11539		C ₈ H ₁₄ O ₄	174.1944	17	18	0	4	0.0003

MetaboSearch

- Find MetaboSearch using your browser
 - it's at <http://omics.georgetown.edu/metabosearch.html>

The screenshot shows the MetaboSearch website. At the top is the Resson Omics LAB logo. Below the logo is a navigation menu with links for Home, Projects, Publications, Members, Resources, and Software Tools. The main content area is titled METABOSEARCH and features an Introduction section. The Introduction text describes mass-based search for metabolite identification. A sidebar on the right contains a CONTENT menu with links to Introduction, Downloads, User Guide, Input File Format, MetaboSearch Interface, Steps To Run MetaboSearch, Output File Format, and Questions & Answers.

Resson Omics LAB

Home Projects Publications Members Resources **Software Tools**

METABOSEARCH

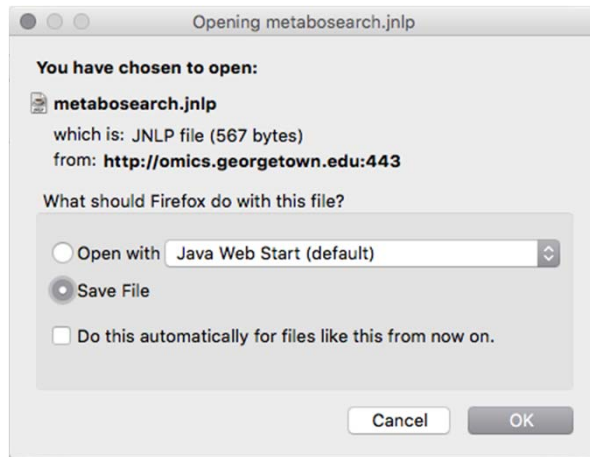
Introduction

- Mass-based search is an important step for metabolite identification in mass-spectrometry-based metabolomic analysis. The mass-to-charge ratio (m/z) value of a molecular ion of interest is searched against metabolite database(s). The metabolites having molecular weights within a specified tolerance to the query m/z value are retrieved from the databases as putative identifications. These putative identifications serve as a foundation for further metabolite verification. In addition to searching with m/z values only, the ion annotation information can be used to aid the mass-based search. Ion annotation groups the ions originating from the same metabolite together and annotates them as adducts/isotopes/in-source fragments. R package **CAMERA** (Collection of Algorithms for MEtabolite pRofile Annotation) was previously developed for ion annotation by Kuhl etc (Carsten Kuhl etc. CAMERA: Collection of annotation related methods for mass spectrometry data. R package version 1.10.0.). Using the ion annotation information, the appropriate mass values of ions can be calculated. Then the calculated mass values are searched against databases. This approach is expected to improve the accuracy for metabolite identification.

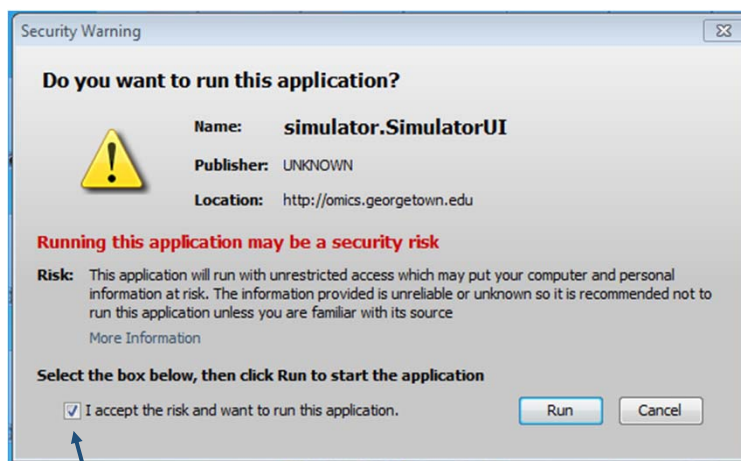
CONTENT

- Introduction
- Downloads
- User Guide
 - Input File Format
 - MetaboSearch Interface
 - Steps To Run MetaboSearch
 - Output File Format
- Questions & Answers

Download at Resson lab



Opening MetaboSearch gives you this



MetaboSearch set up page

MetaboSearch Tool V 1.2

CUSTOMIZED DATABASES

URL	Description	Check
http://www.hmdb.ca/search/spectra?type=ms_search	HMDB database	<input checked="" type="checkbox"/>
http://metlin.scripps.edu/metabo_batch_list.php	Metlin database	<input checked="" type="checkbox"/>
http://mmcd.nmr.fam.wisc.edu/	MMCD database	<input checked="" type="checkbox"/>
http://www.lipidmaps.org/downloads/index.html	Lipidmaps database	<input checked="" type="checkbox"/>

Plug In Manager

INPUT

Positive/Negative Mode
 Positive Negative

MW Tolerance in ppm

Input Mass Data
 Please insert m/z value here .

Or Input File From Local:

OUTPUT

Output Results

STATUS

Data loaded – note negative ions

MetaboSearch Tool V 1.2

CUSTOMIZED DATABASES

URL	Description	Check
http://www.hmdb.ca/search/spectra?type=ms_search	HMDB database	<input checked="" type="checkbox"/>
http://metlin.scripps.edu/metabo_batch_list.php	Metlin database	<input checked="" type="checkbox"/>
http://mmcd.nmr.fam.wisc.edu/	MMCD database	<input checked="" type="checkbox"/>
http://www.lipidmaps.org/downloads/index.html	Lipidmaps database	<input checked="" type="checkbox"/>

Plug In Manager

INPUT

Positive/Negative Mode
 Positive Negative

MW Tolerance in ppm

Input Mass Data
 329.11214
 177.05641
 269.04555
 426.03226
 226.03738
 199.09784
 227.09225
 193.05045
 490.13636
 608.13368

Or Input File From Local:

OUTPUT

Output Results

STATUS

MetaboSearch completed

CUSTOMIZED DATABASES

URL	Description	Check
http://www.hmdb.ca/search/spectra?type=ms_search	HMDB database	<input checked="" type="checkbox"/>
http://metlin.scripps.edu/metabo_batch_list.php	Metlin database	<input checked="" type="checkbox"/>
http://mmcd.nmfam.wisc.edu/	MMCD database	<input checked="" type="checkbox"/>
http://www.lipidmaps.org/downloads/index.html	LipidMaps database	<input checked="" type="checkbox"/>

INPUT

Positive/Negative Mode: Positive Negative

MW Tolerance in ppm: 10

Input Mass Data

329.11214
177.05641
269.04555
426.03226
226.03738
199.09784
227.09225
193.05045
490.13636
608.13368

Or Input File From Local:

OUTPUT

The directory of the result is: C:\Users\Stephen Barnes\Downloads\tempforsearch//
Searching databases, please wait...
Grasped Metlin is coming...
Grasped MMCD is coming...
Please save the result by clicking 'Export' button!

QueryID	Name	Score	Compound Name
2	cq_04971	0.000302805800033	Ethyladipic acid C8H
2	cq_11996	0.000302805800033	Dimethyl adipate, Di
2	cq_13690	0.00180496030001	Diphenylcarbazine
2	cq_17537	0.00100951620004	Cefteram pivoxil, T 2
2	cq_16240	0.00258718569995	1,2-HYDRO-1-OXY-3
3	cq_07883	0.00258718569995	Athamantin C24
5	cq_09397	0.00189767219999	4,4'-Methylenedianil
7	cq_13056	0.00189767219999	Tacrine C13
7	cq_05799	0.00189767219999	N-(Phenylmethyl)-N
8	cq_10129		
8	cq_00955		
8	cq_16789		

Opening Metabosearch file

CUSTOMIZED DATABASES

URL	Description	Check
http://www.hmdb.ca/search/spectra?type=ms_search	HMDB database	<input checked="" type="checkbox"/>
http://metlin.scripps.edu/metabo_batch_list.php	Metlin database	<input checked="" type="checkbox"/>
http://mmcd.nmfam.wisc.edu/	MMCD database	<input checked="" type="checkbox"/>
http://www.lipidmaps.org/downloads/index.html	LipidMaps database	<input checked="" type="checkbox"/>

INPUT

Positive/Negative Mode: Positive Negative

MW Tolerance in ppm: 10

Input Mass Data

329.11214
177.05641
269.04555
426.03226
226.03738
199.09784
227.09225
193.05045
490.13636
608.13368

Or Input File From Local:

OUTPUT

The directory of the result is: C:\Users\Stephen Barnes\Downloads\tempforsearch//
Searching databases, please wait...
Grasped Metlin is coming...
Grasped MMCD is coming...
Would you like to open the output file?

QueryID	Name	Score	Compound Name
2	cq_04971	0.000302805800033	Ethyladipic acid C8H
2	cq_11996	0.000302805800033	Dimethyl adipate, Di
2	cq_13690	0.00180496030001	Diphenylcarbazine
2	cq_17537	0.00100951620004	Cefteram pivoxil, T 2
2	cq_16240	0.00258718569995	1,2-HYDRO-1-OXY-3
3	cq_07883	0.00258718569995	Athamantin C24
5	cq_09397	0.00189767219999	4,4'-Methylenedianil
7	cq_13056	0.00189767219999	Tacrine C13
7	cq_05799	0.00189767219999	N-(Phenylmethyl)-N
8	cq_10129		
8	cq_00955		
8	cq_16789		

MetaboSearch output

A	B	C	D	E	F	G	H	I	J	K	L	M
Query_ID	Query_m/	Input_RT	Name	Formula	Exact_Ma	KEGG ID	PubChem	PubChem	HMDB ID	Databases	dppm	Delta
1	417.1029	-	-	-	-	-	-	-	-	-	-	-
2	173.0816	-	[O4]-ACETOXY-2,3-DIDEOXYFI	C8H14O4	174.0892	-	-	-	-	MMCD	1.665336	2.90E-04
2	173.0816	-	2-(ALPHA-HYDROXYISOVALER	C8H14O4	174.0892	-	-	-	-	MMCD	1.665336	2.90E-04
2	173.0816	-	Ethyladipic acid	C8H14O4	174.0892	-	-	152459	HMDB020	MMCD	1.665336	2.90E-04
2	173.0816	-	Dimethyl adipate;Dimethyl h	C8H14O4	174.0892	C14570	-	-	-	MMCD	1.665336	2.90E-04
2	173.0816	-	Suberic acid	C8H14O4	174.0892	C08278	10457;-	10457	HMDB008	LIPIDMaps	1.665336	2.90E-04
2	173.0816	-	Ethyladipic acid	C8H14O4	174.0892	-	152459;-	152459	HMDB020	LIPIDMaps	1.665336	2.90E-04
2	173.0816	-	Diethyl succinate	C8H14O4	174.0892	-	-	31249	HMDB338	HMDB	1.665336	2.90E-04
2	173.0816	-	Suberic acid;Cork acid;1,8-Oct	C8H14O4	174.0892	C08278	153742	10457	HMDB008	MMCD	1.665336	2.90E-04
3	241.1077	-	Diphenylcarbazine	C13H14N4	242.1168	C11232	151941	-	-	MMCD	7.436648	0.001801
4	499.7199	-	-	-	-	-	-	-	-	-	-	-
5	592.1392	-	Cefteram pivoxil;T 2588	C22H27N9	593.1475	C13147	192387	54885	-	MMCD	1.646408	9.77E-04
6	155.0716	-	-	C8H12O3	156.0786	-	5312943	-	-	LIPIDMaps	1.857506	2.90E-04
6	155.0716	-	8-Hydroxy-5,6-octadienoic ac	C8H12O3	156.0786	-	-	-	HMDB3111	HMDB	1.564215	2.44E-04
6	155.0716	-	2,5-Dimethyl-4-ethoxy-3(2H)	C8H12O3	156.0786	-	-	3017596	HMDB322	HMDB	1.564215	2.44E-04
7	429.1893	-	Mammea E/BA	C24H30O7	430.1992	-	-	11517592	HMDB308	HMDB	6.029752	0.002594
7	429.1893	-	Armillarilin	C24H30O7	430.1992	-	-	21126389	HMDB316	HMDB	6.029752	0.002594
7	429.1893	-	Athamantin	C24H30O7	430.1992	C09123	11315	442051	-	MMCD	6.029752	0.002594
7	429.1893	-	1,2-HYDRO-1-OXY-3,4-HYDRO	C24H30O7	430.1992	-	-	-	-	MMCD	6.029752	0.002594