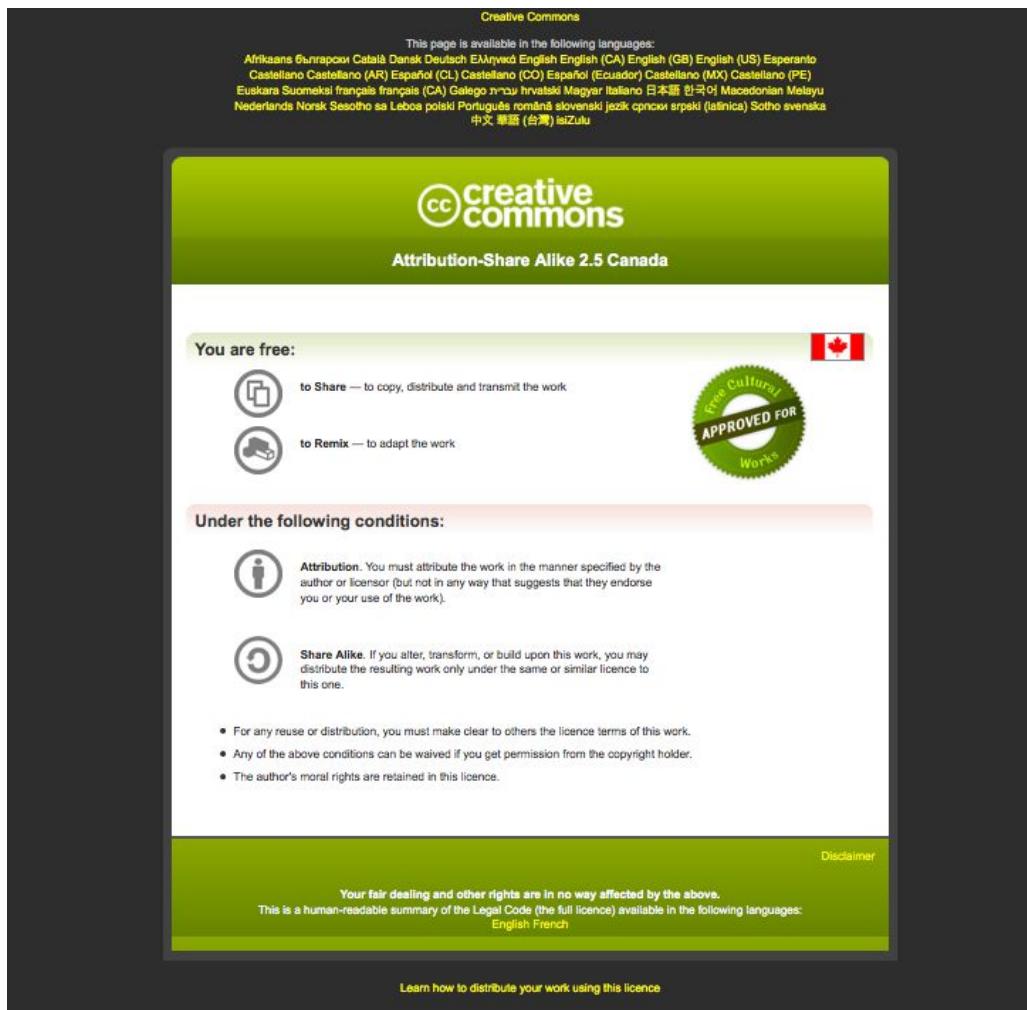




Canadian Bioinformatics Workshops

www.bioinformatics.ca

bioinformaticsdotca.github.io



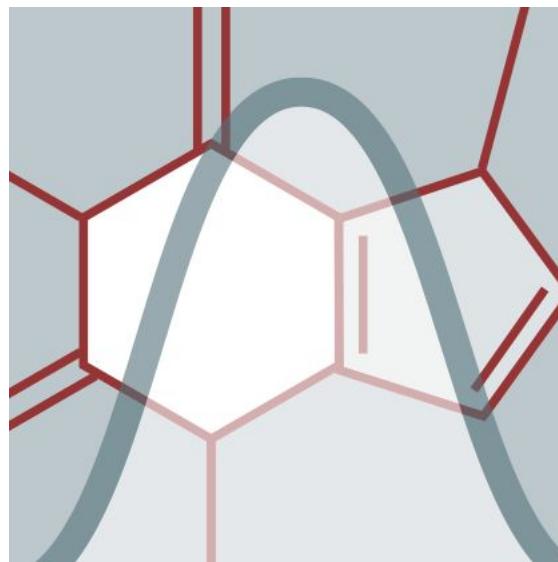
Databases for Biological Interpretation



David Wishart

Informatics and Statistics for Metabolomics

July 6-7, 2023



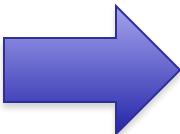
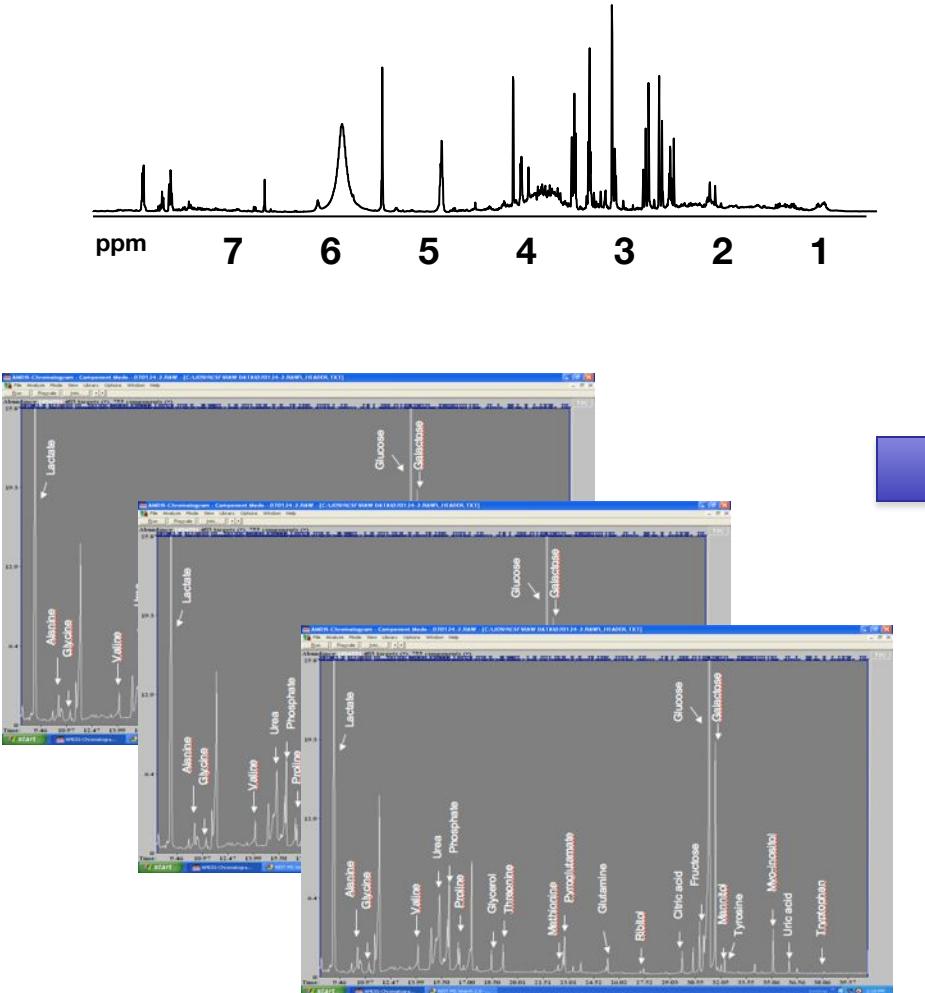
Schedule For July 6, 2023

Time	Module
8:00 (MST)/10:00 (EST)	Arrival & Check-in
8:30 (MST)/10:30 (EST)	Welcome (Nia Hughes)
9:00 (MST)/11:00 (EST)	Module 1: Introduction to Metabolomics (David Wishart)
10:30 (MST)/12:30 (EST)	Break/Lunch (45 min)
11:15 (MST)/13:15 (EST)	Module 2: Targeted, Quantitative Metabolomics (David Wishart)
12:15 (MST)/14:15 (EST)	Lunch/Break (45 min)
13:00 (MST)/15:00 (EST)	Module 3 (Lab): Quantitative Metabolomics (David Wishart)
15:00 (MST)/17:00 (EST)	Break (30 min)
15:30 (MST)/17:30 (EST)	Module 4: Databases for Biological Interpretation (David Wishart)
17:00 (MST)/19:00 (EST)	Finish

Learning Objectives

- To learn about the wide variety of organism-specific and purpose-specific databases now available to help with metabolite data interpretation
- To learn about the potential applications of these specialized databases
- To learn about various pathway databases and how these resources can aid in both metabolomic data and biological interpretation

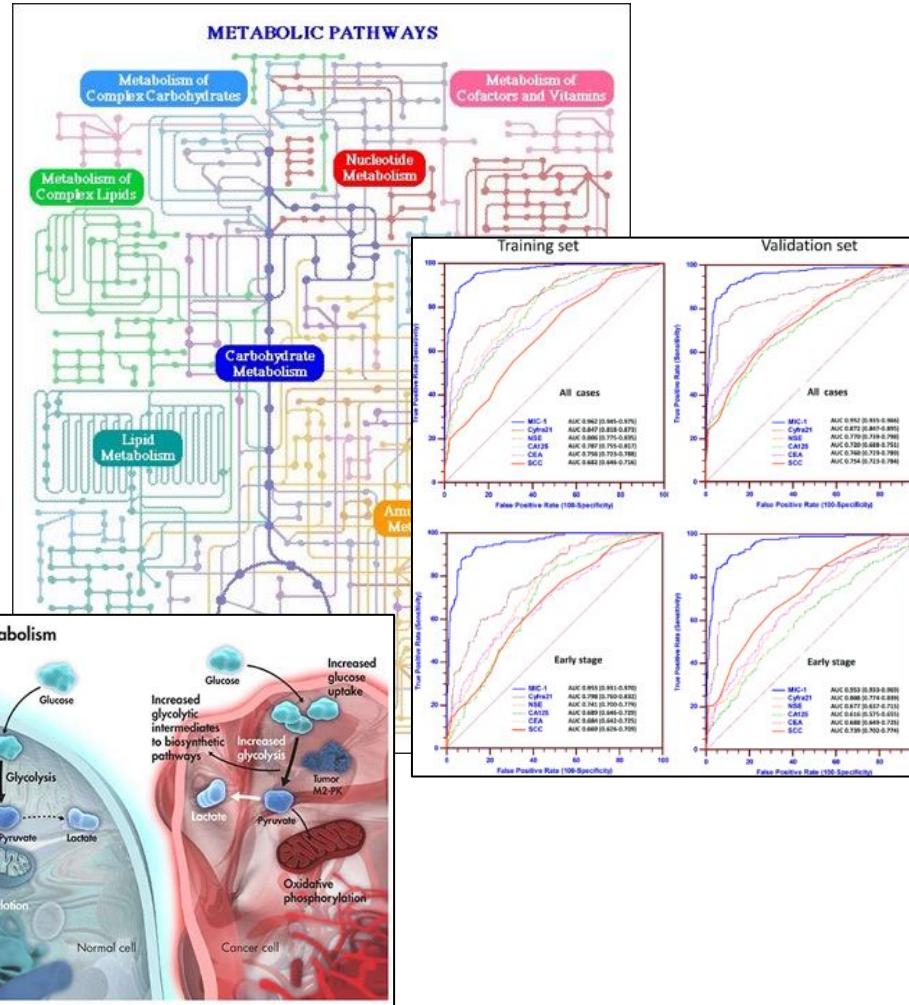
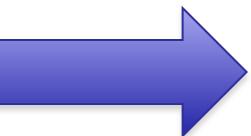
From Spectra to Lists



Compound	Retention Time (min)	Conc. in Urine (μ M)	Compound	Retention Time (min)	Conc. in Urine (μ M)
Dns-o-phospho-L-serine	0.92	<DL*	Dns-Ile	6.35	25
Dns-o-phospho-L-tyrosine	0.95	<DL	Dns-3-aminosalicylic acid	6.44	0.5
Dns-adenosine monophosphate	0.99	<DL	Dns-pipeolic acid	6.50	0.5
Dns-o-phosphoethanolamine	1.06	16	Dns-Leu	6.54	54
Dns-glucosamine	1.06	22	Dns-cystathione	6.64	0.3
Dns-o-phospho-L-threonine	1.09	<DL	Dns-Leu-Pro	6.60	0.4
Dns-6-dimethylamino purine	1.20	<DL	Dns-5-hydroxylsine	6.65	1.6
Dns-3-methyl-histidine	1.22	80	Dns-Cysteine	6.73	160
Dns-taurine	1.25	834	Dns-N-norleucine	6.81	0.1
Dns-carnosine	1.34	28	Dns-5-hydroxydopamine	7.17	<DL
Dns-Arg	1.53	36	Dns-dimethylamine	7.33	293
Dns-Asn	1.55	133	Dns-5-HIAA	7.46	18
Dns-hypotaurine	1.58	10	Dns-umbelliferone	7.47	1.9
Dns-homocarnosine	1.61	3.9	Dns-2,3-diaminopropionic acid	7.63	<DL
Dns-guanidine	1.62	<DL	Dns-L-orntinine	7.70	15
Dns-Gln	1.72	633	Dns-4-acetylamidophenol	7.73	51
Dns-allantoin	1.83	3.8	Dns-procaine	7.73	8.9
Dns-L-citrulline	1.87	2.9	Dns-homocysteine	7.76	3.3
Dns-1-(or 3-)methylhistidine	1.94	1.9	Dns-acetaminophen	7.97	82
Dns-adenosine	2.06	2.6	Dns-Phe-Phe	8.03	0.4
Dns-methylguanidine	2.20	<DL	Dns-5-methoxy salicylic acid	8.04	2.1
Dns-Ser	2.24	511	Dns-Lys	8.16	184
Dns-aspartic acid amide	2.44	26	Dns-aniline	8.17	<DL
Dns-4-hydroxy-proline	2.56	2.3	Dns-leu-Phe	8.22	0.3
Dns-Glu	2.57	21	Dns-His	8.35	1560
Dns-Asp	2.60	90	Dns-4-thiolsine	8.37	<DL
Dns-Thr	3.03	157	Dns-benzylamine	8.38	<DL
Dns-epinephrine	3.05	<DL	Dns-1-ephedrine	8.50	0.6
Dns-ethanolamine	3.11	471	Dns-tryptamine	8.63	0.4
Dns-aminoadipic acid	3.17	70	Dns-pyridoxamine	8.94	<DL
Dns-Gly	3.43	2510	Dns-2-methyl-benzylamine	9.24	<DL
Dns-Ala	3.68	593	Dns-5-hydroxytryptophan	9.25	0.12
Dns-aminolevulinic acid	3.97	30	Dns-13-diaminopropane	9.44	0.23
Dns-r-amino-butyric acid	3.98	4.6	Dns-p-tetrescine	9.60	0.5
Dns-p-amino-hippuric acid	3.98	2.9	Dns-1,2-diaminopropane	9.66	0.1
Dns-5-hydroxymethyluric acid	4.58	1.9	Dns-tyrosinamide	9.79	29
Dns-tryptophanide	4.70	5.5	Dns-dopamine	10.08	140
Dns-isoguanine	4.75	<DL	Dns-cadaverine	10.08	0.08
Dns-5-amino pentanoic acid	4.79	1.6	Dns-histamine	10.19	0.4
Dns-sarcosine	4.81	7.2	Dns-3-methoxy-tyramine	10.19	9.2
Dns-3-amino-isobutyrate	4.81	85	Dns-Tyr	10.28	321
Dns-2-aminobutyric acid	4.91	17	Dns-cysteamine	10.44	<DL

From Lists to Biology

Compound	Retention Time (min)	Conc. in Urine (μM)	Compound	Retention Time (min)	Conc. in Urine (μM)
Dns-o-phospho -L-serine	0.92	<DL*	Dns-Lle	6.35	25
Dns-o-phospho -L-tyrosine	0.95	<DL	Dns-3-aminosalicylic acid	6.44	0.5
Dns-adenosine monophosphate	0.99	<DL	Dns-pipeolic acid	6.50	0.5
Dns-o-phosphoethanolamine	1.06	16	Dns-Leu	6.54	54
Dns-glucosamine	1.06	22	Dns-cystathione	6.54	0.3
Dns-o-phospho -L-threonine	1.09	<DL	Dns-Leu-Pro	6.60	0.4
Dns-6-dimethyllysine	1.20	<DL	Dns-5-hydroxylysine	6.65	1.6
Dns-3-methyl -histidine	1.22	80	Dns-Cysteine	6.73	160
Dns-taurine	1.25	834	Dns-N-norleucine	6.81	0.1
Dns-carnosine	1.34	28	Dns-5-hydroxydopamine	7.17	<DL
Dns-Arg	1.53	36	Dns-dimethylamine	7.33	293
Dns-Asn	1.55	133	Dns-5-HIAA	7.46	18
Dns-hypotaurine	1.58	10	Dns-umbelliferone	7.47	1.9
Dns-homocarnosine	1.61	3.9	Dns-2,3-diaminopropionic acid	7.63	<DL
Dns-guanidine	1.62	<DL	Dns-L-ornithine	7.70	15
Dns-Gln	1.72	633	Dns-4-acetylamidophenol	7.73	51
Dns-allantoin	1.83	3.8	Dns-procaine	7.73	8.9
Dns-L-citrulline	1.87	2.9	Dns-homocystine	7.76	3.3
Dns-1 (or 3-)methylhistamine	1.94	1.9	Dns-acetaminophen	7.97	82
Dns-adenosine	2.06	2.6	Dns-Phe-Phe	8.03	0.4
Dns-methylguanidine	2.20	<DL	Dns-5-methoxy salicylic acid	8.04	21
Dns-Ser	2.24	511	Dns-Lys	8.16	184
Dns-aspartic acid amide	2.44	26	Dns-aniline	8.17	<DL
Dns-4-hydroxy -proline	2.56	2.3	Dns-Ieu-Phe	8.22	0.3
Dns-Glu	2.57	21	Dns-His	8.35	1550
Dns-Asp	2.60	90	Dns-4-thialysine	8.37	<DL
Dns-Thr	3.03	157	Dns-benzylamine	8.38	<DL
Dns-epinephrine	3.05	<DL	Dns-1-ephedrine	8.50	0.6
Dns-ethanolamine	3.11	471	Dns-tryptamine	8.63	0.4
Dns-aminoacidic acid	3.17	70	Dns-pyridoxamine	8.94	<DL
Dns-Gly	3.43	2510	Dns-2-methyl -benzylamine	9.24	<DL
Dns-Ala	3.88	593	Dns-5-hydroxytryptophan	9.25	0.12
Dns-aminolevulinic acid	3.97	30	Dns-1,3-diaminopropane	9.44	0.23
Dns- <i>l</i> -amino-butyric acid	3.98	4.6	Dns-putrescine	9.60	0.5
Dns- <i>d</i> -amino-hippuric acid	3.98	2.9	Dns-1,2-diaminopropane	9.66	0.1
Dns-5-hydroxymethyluric acid	4.58	1.9	Dns-tyrosinamide	9.79	29
Dns-tryptophanamide	4.70	5.5	Dns-dopamine	10.08	140
Dns-isoguanine	4.75	<DL	Dns-cadaverine	10.08	0.08
Dns-5-aminopentanoic acid	4.79	1.6	Dns-histamine	10.19	0.4
Dns-sarcosine	4.81	7.2	Dns-3-methoxy -tyramine	10.19	9.2
Dns-3-amino -isobutyrate	4.81	85	Dns-Tyr	10.28	321
Dns-2-aminobutyric acid	4.91	17	Dns-cysteamine	10.44	<DL



Interpreting Metabolomics Data

- Most ‘omics data is just a long list of names (genes, proteins, metabolites) and either relative or absolute concentrations
- How do we determine if certain metabolite levels are too high or too low?
- How do we determine the origin of the compounds we measure (food? microbes? endogenous?)
- How do we determine what pathways or pathologies these metabolites are involved?
- How do we determine which genes or proteins are responsible for catabolizing or anabolizing these metabolites?

Answer: Databases

Databases

- Many kinds of chemical databases exist
- Most of us are aware of ChEBI, PubChem or ChemSpider but these are general chemical structure/nomenclature databases and have very little biology or organism specificity
- What is needed for metabolomics are databases that link biology to chemistry and that link chemistry to specific applications or disciplines (i.e., purpose-specific DBs)

Need For Organism/Purpose-Specific Databases



[Metabolites](#). 2016 Mar; 6(1): 8.

PMCID: PMC4812337

Published online 2016 Feb 15. doi: [10.3390/metabo6010008](https://doi.org/10.3390/metabo6010008)

PMID: [26891337](#)

The Time Is Right to Focus on Model Organism Metabolomes

Arthur S. Edison,¹ Robert D. Hall,² Christophe Junot,³ Peter D. Karp,⁴ Irwin J. Kurland,⁵ Robert Mistrik,⁶ Laura K. Reed,⁷ Kazuki Saito,⁸ Reza M. Salek,⁹ Christoph Steinbeck,⁹ Lloyd W. Sumner,¹⁰ and Mark R. Viant^{11,*}

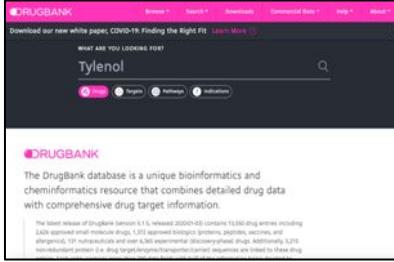
Peter Meikle, Academic Editor

► Author information ► Article notes ► Copyright and License information [Disclaimer](#)

Organism/Purpose-Specific Metabolome Databases



www.hmdb.ca



www.drugbank.ca



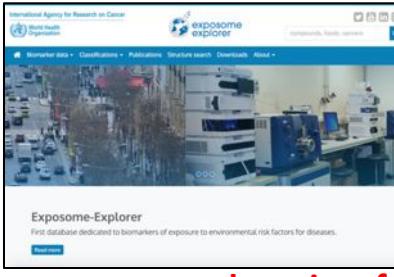
www.ymdb.ca



www.smpdb.ca



www.ecmdb.ca



exposome-explorer.iarc.fr



www.phenol-explorer.eu



www.t3db.org



www.foodb.ca



www.phytobank.ca



www.pathbank.org



www.contaminantdb.ca

The Human Metabolome Database (HMDB)



- A comprehensive metabolomics database containing detailed information about human metabolites, their structures, pathways, origins, concentrations, functions and reference spectra
- HMDB has 248,855 metabolites, 132,335 pathways, 3.1 million MS and NMR spectra, metabolite biomarker data on >600 diseases
- A resource established to provide reference metabolite values for human disease, human exposures & population health
- Captures both **targeted** and **untargeted** metabolomics (and exposomics) data

<http://www.hmdb.ca>

Inside the HMDB...

The screenshot displays the HMDB interface. At the top, there's a navigation bar with links for 'HMDB', 'Browse', 'Search', 'Downloads', 'About', and 'Contact Us'. Below the navigation, a search bar is present with placeholder text 'Search for a compound...'. The main content area is divided into several sections:

- Physical Properties**: Shows the state as 'Solid'.
- Experimental Molecular Properties**: Includes Melting Point (249 °C), Boiling Point (Not Available), Water Solubility (200 g/kg), and LogP (Not Available). Reference links point to 'Properties of Amino Acids.pdf'.
- Experimental Collision Cross Sections**: Lists adduct types [M-H]-, [M-H]2+, [M+H]+, and [M+H]2+ with their CCS values (139.639, 129.0, 133.103) and references (30932474, 30932474, 30932474).
- Predicted Molecular Properties**: Shows Water Solubility (0.93 g/L), logP (-3), and logS (-3.1).
- GC-MS Spectra**: A detailed view of the mass spectrum for 1-Methylhistidine. It includes a plot from 1 to -1 ppm, an assignment table with peak multiplets at 7.67, 7.00, 3.96, 3.88, 3.18, 3.06 ppm, and a cluster midpoint table.
- Spectra View**: A list of deposited spectra with columns for Deposition Date, Source, and View.

- Detailed descriptions of every metabolite
- Reference (measured or predicted) NMR, MS/MS and EI-MS spectra of individual chemicals along with measured (or predicted) retention indices, retention times, collision cross sectional area, etc.
- Search tools for matching structures (by Tanimoto substructure search), matching formulas, matching masses, matching NMR spectra or matching MS spectra
- Designed to help identify or confirm the identity of compounds through a whole range of metabolomics experiments or metabolomics measurements

<http://www.hmdb.ca>

MS Spectral Searching

Human Metabolome Database: Spectra Search Mass Spectrum
www.hmdb.ca/spectra/ms/search

Google Apple iCloud Facebook Twitter Wikipedia Yahoo News Popular

HMDB Browse Search Downloads About Contact Us Search metabolites

Query Masses (Da)
175.01
238.19
420.16
780.32
956.25
1100.45

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

Ionization Ion Mode Adduct Type
Positive Unknown M+H M+2H M+3H 2M+H M+K 2M+K

Molecular Weight Tolerance ± 0.05 Da

Search

MS search for 175.01 m/z

Delta = abs(query mass - adduct mass)

Compound	Name	Adduct	Adduct MW (Da)	Compound MW (Da)	Delta
HMDB60293	Hydroxidiodoxidosulfidosulfate	M+IsoProp+H	175.009875	113.944535	0.000125
HMDB31436	Silicic acid	M+DMSO+H	175.009105	95.987885	0.000895
HMDB33657	De-O-methylsterigmatocystin	M+H+K	175.009086	310.047738	0.000914
HMDB35230	Aurantrichlide B	M+H+K	175.009086	310.047738	0.000914
HMDB34155	Thiourea	2M+Na	175.008256	76.009519	0.001744
HMDB01570	Thymidine 3',5'-cyclic monophosphate	M+2Na	175.012237	304.046037	0.002237
HMDB01270	Glyceric acid 1,3-biphosphate	M+2ACN+2H	175.013458	265.95927	0.003458
HMDB01294	2,3-Diphosphoglyceric acid	M+2ACN+2H	175.013458	265.95927	0.003458
HMDB60394	5-Fluorodeoxyuridine monophosphate	M+H+Na	175.014012	326.03153	0.004012
HMDB60015	Phenol sulphate	M+H	175.005955	173.998679	0.004045

Showing 1 to 10 of 302 entries

Previous 1 2 3 4 5 ... 31 Next

MS/MS Spectral Searching

Human Metabolome Database: Spectra Search Tandem Mass Spectrum
www.hmdb.ca /spectra/ms_ms/search

HMDB Browse Search Downloads About Contact Us Search metabolites Q Search

MS Search MS/MS Search GC/MS Search 1D NMR Search 2D NMR Search

Parent Ion Mass (Da) 146.0

Parent Ion Mass Tolerance ± 0.1 Da

Ionization Positive

CID Energy Low

MS/MS Peak list (M/Z RT)

40.948 0.174
56.022 0.424
84.37 53.488
101.50 8.285
102.401 0.775
129.670 100.000
146.966 20.070

Mass/Charge (m/z) Tolerance ± 0.5

Include predicted spectra?

HMDB Browse Search Downloads About Contact Us Search metabolites Q Search

Name CAS Number	Formula Weight	Structure	Fit(%)	RFit(%)	Purity(%)
EXPERIMENTAL 2-Methylglutaric acid (HMDB00422) View Spectra 617-62-9 C ₆ H ₁₀ O ₄	146.1412		0.00	0.54	0.50
EXPERIMENTAL L-Glutamine (HMDB00641) View Spectra 56-85-9 C ₆ H ₁₀ N ₂ O ₃	146.1445		1.00	1.00	1.00
EXPERIMENTAL Methylglutaric acid (HMDB00752) View Spectra 626-51-7 C ₆ H ₁₀ O ₄	146.1412		0.56	0.63	0.66
EXPERIMENTAL Coumarin (HMDB01218) View Spectra 91-64-5 C ₉ H ₆ O ₂	146.1427		0.79	0.36	0.50

NMR Spectral Searching

Human Metabolome Database: Spectra Search NMR Spectrum
www.hmdb.ca/spectra/nmr_one_d/search

HMDB Browse Search Downloads About Contact Us Search metabolites Search

Spectra Search NMR Spectrum

MS Search MS/MS Search GC/MS Search 1D NMR Search 2D NMR Search

Spectra Library 1H NMR

Peak List

3.81	
3.82	
3.83	
3.85	
3.89	
3.90	
3.91	
4.25	
4.26	
4.27	
4.41	
8.19	
8.31	

Peak Tolerance \pm 0.02

Search

Name CAS Number Formula Weight Structure Library Matches

Glycerol 3-phosphate (HMDB00126)	172.0737	20/34
57-03-4	C ₃ H ₆ O ₃ P	
Inosine (HMDB00195)	268.2261	12/21
58-63-9	C ₁₀ H ₁₂ N ₄ O ₆	
Xanthosine (HMDB00299)	284.2255	12/22
146-80-5	C ₁₀ H ₁₂ N ₄ O ₆	
Ribonolactone (HMDB01900)	148.1114	8/10
5336-08-3	C ₃ H ₆ O ₅	

Chemical Structures:

- Glycerol 3-phosphate (HMDB00126): A three-carbon chain with a phosphate group at the third position.
- Inosine (HMDB00195): A purine nucleoside with a ribose sugar.
- Xanthosine (HMDB00299): A purine nucleoside with a xanthosine ring system.
- Ribonolactone (HMDB01900): A cyclic lactone with hydroxyl groups.
- 5336-08-3: A cyclic lactone with hydroxyl groups.

HMDB Structure Searching

Human Metabolome Database: ChemQuery Search by structure
www.hmdb.ca/structures/search/metabolites/structure

Google Apple iCloud Facebook Twitter Wikipedia Yahoo News Popular

HMDB Browse Search Downloads About Contact Us

Structure Search Molecular Weight

Absolute

Search Options

Similarity Substrate

Similarity threshold: 0.7

Molecular Weight Filter: e.g. 100

Maximum Results: 100

Status (default all):

- Detected and quantified
- Detected but not quantified
- Expected but not quantified

Search

← Previous

1-Methylhistidine
332-80-9
Detected and Quantified

L-Histidine
71-00-1
Detected and Quantified

3-Methylhistidine
368-16-1
Detected and Quantified

(S,S)-Nt-Histidinylalanine
65428-77-5
Expected but not Quantified

The HMDB Biofluids Database

- Reference metabolite concentrations for >650 different diseases & conditions
- Abnormal and normal metabolite concentrations for >15 biofluids and >5000 different metabolites
- Designed for clinical chemists & physicians
- Largest & most complete resource of its kind

The screenshot shows the HMDB Biofluids Database interface. At the top, there's a navigation bar with links for Browse, Search, Downloads, About, and Contact Us. The search bar contains the text "metabolites". Below the navigation is a logo for TMIC (The Metabolomics Innovation Centre) and a tagline "Specializing in ready to use metabolomics kits". The main content area is titled "Browsing Biofluids". It features three filter sections: "Filter by metabolite status" (with options for Detected and quantified, Detected but not quantified, Expected but not quantified), "Filter by biofluid" (with options for Other Fluids, Saliva, Cerebrospinal Fluid, Urine, Blood, Feces, and Sweat), and "Filter by identifiers" (with options for Metabolite and Disease). Below the filters, a message says "Displaying metabolites 1 - 25 of 74461 in total". A navigation bar below shows pages 1 through 5, with "Next" and "Last" buttons. The results table for "1-Methylhistidine" (HMDB00001) has columns for Biofluid, Concentration, Patient Status, Conditions, Age, Sex, Reference, and Details. Four rows of data are shown for Blood samples:

Biofluid	Concentration	Patient Status	Conditions	Age	Sex	Reference	Details
Blood	7.7 +/- 1.9 μM	Normal	Normal	Adult (>18 years old)	Both	7061274	
Blood	14.4 +/- 2.3 μM	Normal	Normal	Adult (>18 years old)	Both	7061274	
Blood	19.6 +/- 2.6 μM	Normal	Normal	Adult (>18 years old)	Both	7061274	
Blood	12.7 +/- 2.9 μM	Normal	Normal	Adult (>18 years old)	Both	7061274	

HMDB Applications

- Widely used to learn more about metabolites identified in metabolomics studies (biology, chemistry, roles)
- Commonly used to determine normal and abnormal concentrations in different biofluids or tissues
- Used to determine provenance or origin of many metabolites
- Used to determine pathways, enzymes or receptors used by metabolites

MarkerDB: A Biomarker Database

The screenshot displays the MarkerDB homepage and several search results pages for different biomarker types:

- MarkerDB Home:** Shows a heatmap of biomarkers across various conditions and a navigation bar with categories: Condition-specific Biomarkers, Chemical Biomarkers, Genetic Biomarkers, Protein Biomarkers, Karyotype Biomarkers, Diagnostic Biomarkers, Prognostic Biomarkers, Predictive Biomarkers, and Exposure Biomarkers.
- Chemical Biomarkers:** A search result page for "Chemical Biomarkers" showing three entries: (R)-3-Hydroxybutyric acid, (R)-3-hydroxyisobutyric acid, and (R)-4-Methylsalicinal. Each entry includes a chemical structure, marker type (single), and condition name (e.g., Multiple conditions, Diabetes Mellitus Type 2, Parkinson's Disease).
- Karyotype Biomarkers:** A search result page for "Karyotype Biomarkers" showing four entries: IGH, BCL2, BCL2L1, and FOXP1. Each entry includes a karyogram, marker type (single), and condition name (e.g., Malt Lymphoma, Acute Lymphoblastic Leukemia, Non-Hodgkin Lymphoma, Somatic, or Malt Lymphoma).
- Protein Biomarkers:** A search result page for "Protein Biomarkers" showing two entries: Glycated hemoglobin and Angiotensin-converting enzyme. Each entry includes a protein structure, marker type (single), and condition name (e.g., Diabetes Mellitus Type 2 Clinical, Diabetes Mellitus Type 2 Clinical, or Serotonin (Clinical)).
- Specific Condition:** A detailed view for "Diabetes Mellitus Type 2" showing general conditions, biofluids, disease concentrations, and an ROC curve plot of Sensitivity vs. 1 - Specificity.

- An integrated biomarker database of validated molecular diagnostic, prognostic, predictive and exposure biomarkers
- Includes metabolite, protein, and gene biomarkers
- Covers 1089 metabolite-disease/diet biomarkers
- Covers 142 protein-disease biomarkers
- Covers 26,121 genetic-disease biomarkers
- Can search by sequence, structures and text

<https://markerdb.ca>

MarkerDB Applications

- Useful for finding known disease associations for many metabolites
- Used to identify (or rediscover) previously known biomarkers
- Used to identify food consumption and exposure biomarkers and their origins in human biofluids and tissues
- Allows assessment of newly discovered biomarkers against known biomarkers for same/similar conditions

The Drug Database (DrugBank v. 5.0)

The DrugBank database is a cheminformatics resource with comprehensive data on small molecule drugs, biologics, nutraceuticals, and experimental drugs. It includes detailed ADMET, MOA, and pharmacokinetic data, along with metabolizing enzyme data, drug metabolites, MS+NMR spectra, unique drug targets, and data fields per drug. The database supports sequence, spectral, structure, and text searches, as well as compound browsing.

Drug Record for Acetaminophen (DB00316):

- Name: Acetaminophen
- Accession Number: DB00316 (APRD00252)
- Type: Small Molecule
- Groups: Approved
- Description: Acetaminophen, also known as paracetamol, is commonly used for its analgesic and antipyretic effects. Its therapeutic effects are similar to salicylates, but it lacks anti-inflammatory, antiplatelet, and gastric ulcerative effects.
- Structure: O=C(Oc1ccc(C(=O)O)cc1)C
- Synonyms: 4-(Acetylamino)phenol, 4-acetamidophenol, 4-hydroxyacetanilide

- **2645 small molecule drugs, 1403 biologics, 130 nutraceuticals**
- **>6300 experimental drugs**
- **Detailed ADMET, MOA and pharmacokinetic data**
- **>3850 drugs with metabolizing enzyme data**
- **>1360 drug metabolites**
- **>6000 MS+NMR spectra**
- **>4560 unique drug targets**
- **215 data fields/drug**
- **Supports sequence, spectral, structure and text searches as well as compound browsing**

<https://go.drugbank.com/>

DrugBank Applications

- Useful for finding origins of drugs, drug-like molecules or drug metabolites found in human biosamples
- Used to identify possible drug-like roles of metabolites or natural products
- Used to identify pathways and mechanisms for many small molecules
- Used frequently for drug discovery and drug repurposing

The Food Constituent Database (FooDB)

The screenshot displays three main pages of the FooDB website:

- Homepage:** Shows a banner with various fruits and vegetables, the FooDB logo, and links for "Browse Foods >>" and "Browse Compounds >>".
- Listing foods:** A table showing a list of foods, including Angelica and Savoy cabbage, with columns for Name, Scientific name, Food group, and Food subgroup, accompanied by small food images.
- Showing Compound Mulberrofuran P (FDB000001):** A detailed record page for the compound Mulberrofuran P, showing its record information (version 1.0, creation date 2010-04-09 22:04:23 UTC), chemical information (FooDB Name: Mulberrofuran P, Description: Constituent of Morus alba (white mulberry) [CC05], CAS Number: 101365-02-0), and its chemical structure (a triterpenoid saponin).

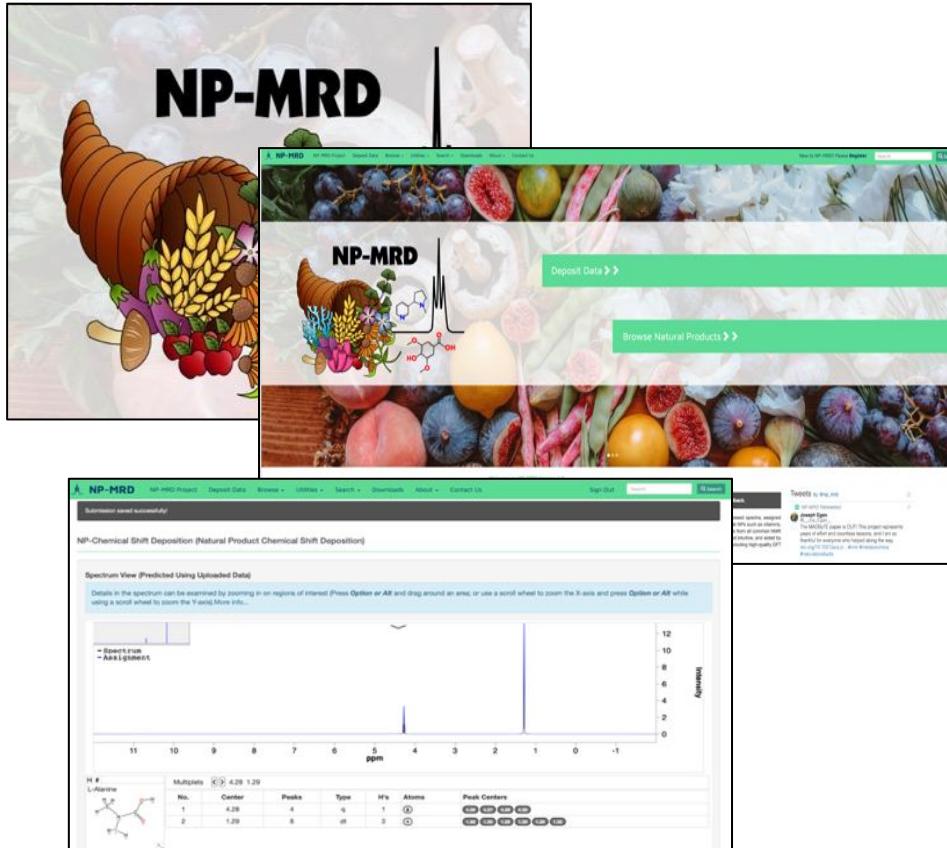
- **Database of 70,000+ compounds found in 727 foods and their effects on flavour, aroma, colour and human health**
- **Comprehensive concentration information to ID foods that are rich in particular micronutrients**
- **Links chemistry to food types (biological species) to flavour, aroma, colour and human health**
- **Supports sequence, spectral, structure and text searches**

<http://www.foodb.ca>

FooDB Applications

- **Useful for finding origins of food molecules or food metabolites or other “unknowns” found in human biosamples**
- **Used to identify possible drug-like roles of food chemicals or natural products**
- **Used to identify organoleptic and health effects properties of many food chemicals**

NP-MRD The Natural Product Database



- A natural products database containing edible plant, medicinal plant and microbial natural products and their corresponding NMR data (NP=Natural Products, MRD=Magnetic Resonance Database)
- Supports user NMR spectral deposition (like BMRB or PDB)
- Includes NMR assignments, structures, spectra, descriptions, medicinal uses, chemical properties, species of origin
- 297,290 compounds with 1,210,074 NMR spectra (1975 expt. measured, 408,293 simulated, 799,806 predicted)
- Supports spectral, structure & text searches
- Uses a variety of AI and ML methods to curate, annotate and maintain the DB

<http://www.np-mrd.org/>

NP-MRD Applications

- Useful for finding origins of natural product medicines or natural product molecules or other “unknowns” found in human biosamples
- Used to identify possible drug-like roles of natural products
- Used to help characterize the metabolomes of plants and microbes
- Useful for chemotaxonomy (finding the biological or species origins of natural chemicals)

The Toxic Exposome Database (T3DB)

The screenshot shows two views of the T3DB website. The top view is the homepage, featuring a banner with two researchers in protective suits, a search bar, and a sidebar with tweets from the Alberta Research Twitter account. The bottom view is a detailed record for the compound Digoxin (T3D2670), displaying its predicted LogP (2.3667), route of exposure (Injection or dermal contact), mechanism of action (binds to the alpha-subunit of the Na+/K+ ATPase pump), metabolism (hepatic), toxicity values, carcinogenicity, uses/sources, minimum risk level, health effects, symptoms, and treatment. It also includes a note about its use in foxglove plants.

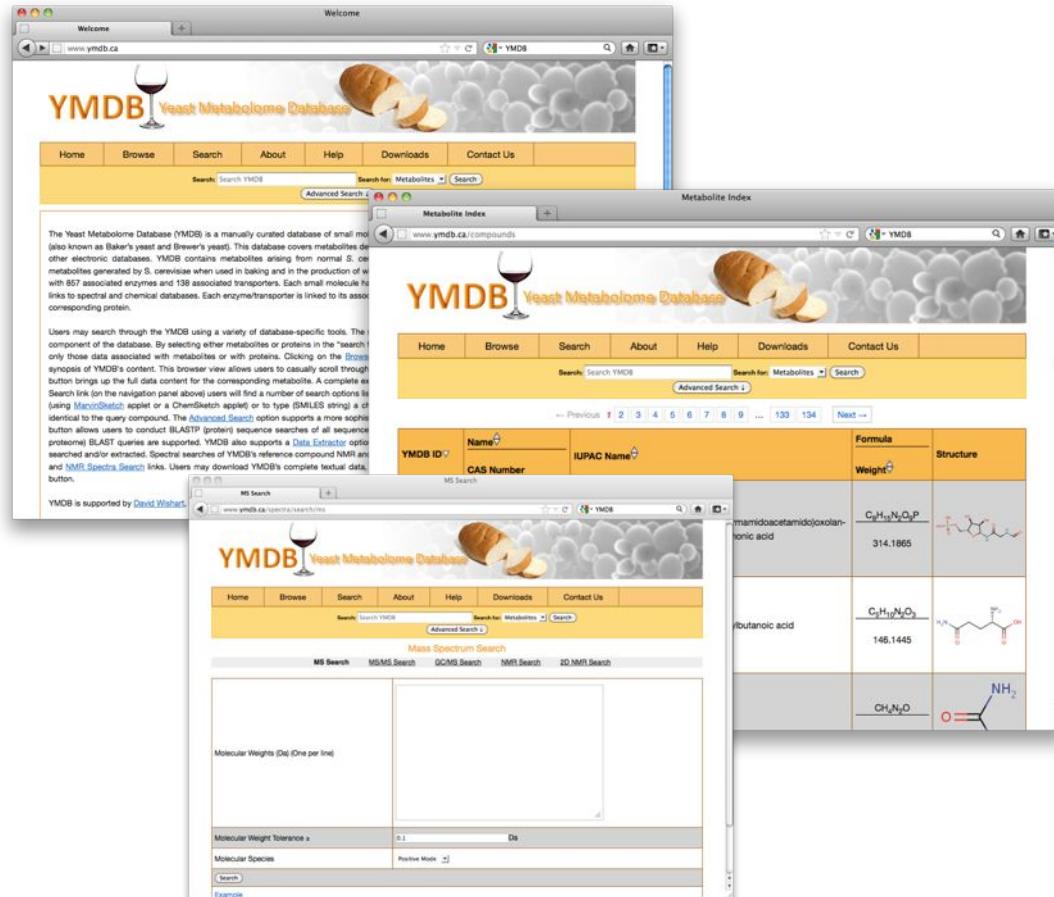
- Comprehensive data on toxic compounds (pesticides, herbicides, endocrine disruptors, solvents, carcinogens, etc.)
- Detailed mechanisms, binding constants, target info, lots of ToxCast data
- >3600 toxic compounds
- >1900 reference spectra
- ~2100 toxic targets
- Supports sequence, spectral, structure, text searches as well as compound browsing
- Full data downloads

<http://www.t3db.ca>

T3DB Applications

- Useful for finding origins of exposure compounds (household chemicals, LPVs, herbicides, pesticides, pollutants, drugs) or other “unknowns” found in human biosamples
- Useful for understanding mechanism of toxicity for many man-made or synthetic compounds (household chemicals, LPVs, herbicides, pesticides, pollutants, drugs)

The Yeast Metabolome Database (YMDB)



- >16,042 yeast metabolites from 57 different growth substrates
- 1058 protein/enzyme metabolite associations
- 52,550 NMR or GC-MS or LC-MS reference spectra
- 31,624 reactions, 9547 pathways
- Supports sequence, spectral, structure and text searches as well as compound browsing
- 78 data fields per compound

<http://www.ymdb.ca>

The E. coli Metabolome Database (ECMDB)

ECMDB ID	Name	Formula	Structure
ECMDB00005	2-Ketobutyric acid MetaboCard 600-18-0	C ₄ H ₆ O ₃ 102.0886	<chem>CC(=O)C(=O)C(O)C</chem>
ECMDB00012	Deoxyuridine MetaboCard 951-78-0	C ₁₀ H ₁₂ N ₂ O ₅ 228.202	<chem>C1=CNC2=C1C(=O)C(=O)N2C</chem>
ECMDB00014	Deoxycytidine MetaboCard	C ₁₀ H ₁₃ N ₃ O ₄	<chem>C1=CNC2=C1C(=O)C(=O)N2C</chem>

- **3755 E. coli metabolites**
- **1789 genes (1402 enzymes, 387 transporters)**
- **3145 chemical reactions**
- **4300 references**
- **1542 pathways**
- **Supports sequence, structure & text searches as well as compound browsing**
- **80 data fields per compound**
- **19,294 NMR and MS spectra**
- **Supports sequence, spectral, structure and text searches**

<http://www.ecmdb.ca>

The Human Microbial Metabolome Database: MiMeDB

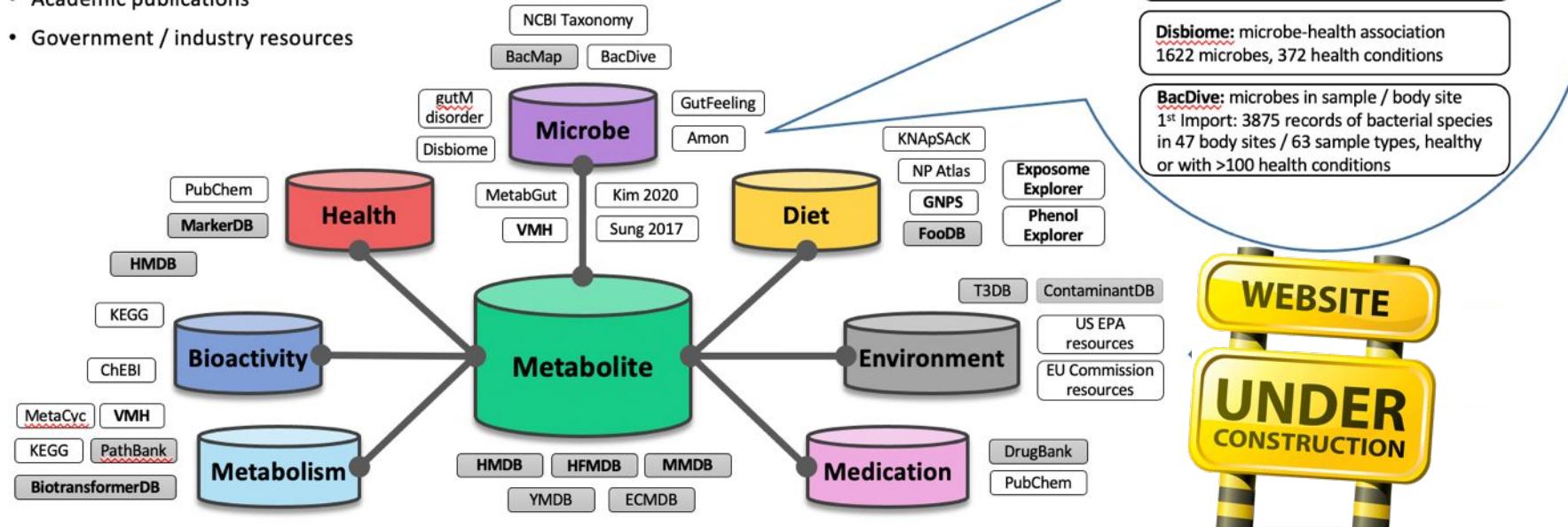
The figure shows two screenshots of the MiMeDB website. The top screenshot is the homepage, featuring a blue header with navigation links (Browse, Search, Visualize, Downloads, About, Contact Us) and a search bar. Below the header is a banner with a blue background and white text. The bottom screenshot shows a detailed view of the database, with a sidebar for 'Browsing Metabolites' and a main panel displaying a metabolic pathway diagram. The diagram illustrates the biosynthesis of indole compounds, starting from tryptophol and tryptamine, leading through indoleacrylic acid, indolepropionic acid, and indoleactic acid to various metabolites like indoleacetic acid and 3-methylindole. Other nodes include indoxyl sulfate, indoxyl, 5,6-diol-indole, indole, indolylglycine, and indoleacryloylglycine.

<https://mimedb.org/>

- A database of microbially-derived metabolites, with information about their microbial origins, their food origins, their reactions and their effects on human health
- 24,254 unique metabolites with detailed descriptions, physicochemical & spectral data
- 1904 unique microbial species with annotated genomic data and genome maps
- 626 human diseases & mechanisms & associations

MiMeDB – A Comprehensive Microbial Metabolite DB

- Connecting resources in one portal, with wide overview and flexible query tool
- Assist in defining and prioritising potential biomarkers
- Investigate results within their biological context (both host and microbes)
- A comprehensive data repository (>30 sources; all linked to source).
 - Extensive databases (in grey - Wishart Lab databases)
 - Text mining (tools developed in the Wishart Lab)
 - Published datasets
 - Academic publications
 - Government / industry resources



MiMeDB – A Look Inside

Host source view

Host and Biospecimen	Status	Concentration	Age	Sex	Health condition	Reference
Human urine	Detected but not Quantified		adult	Both	type 2 diabetes, high CVD risk	28692847
Human urine	Detected but not Quantified		adult		unknown	30087103
Human blood	Detected but not Quantified		adult	Both	CKD	30087103
Human urine	Detected and Quantified	36.932 micromol/mmol	1-13yo	Both	Eosinophilic esophagitis	HMDB
Human blood	Detected and Quantified	4.42 micromolar	adult	Male	healthy	HMDB
Human blood	Detected and Quantified	2.49 +/- 1.36 micromolar	1-13yo	Both	healthy	HMDB
Human urine	Detected and Quantified	129 (45-191) micromol/mmol creatinine	adult	Both	healthy	HMDB
Human urine	Detected and Quantified	19.74 +/- 5.26 micromol/mmol creatinine	adult	Male	healthy	HMDB
Human urine	Detected and Quantified	17.76 +/- 6.58 micromol/mmol creatinine	adult	Female	healthy	HMDB

Health effects view

Health outcome	Metabolite Response	Related Health Condition	Evidence Type	Measured in Matrix	Reference
Mild cognitive impairment	Increased	CKD	Association	serum	PMID:26797588

Bioactivity	Metabolite effect	Related Health Condition	Evidence Type	Measured in Matrix	Reference
Nuclear factor-kb	agonist	Neuroinflammation CKD	Cell culture	astrocyte cells	PMID:26659803
Aryl hydrocarbon receptor	agonist	Neuroinflammation	Cell culture	Mouse CaCo-2 intestinal cells	PMID:34137156
Aryl hydrocarbon receptor	agonist	Neuroinflammation	Cell culture	astrocyte cells	PMID:30336612
Neurotoxin		Neuroinflammation CKD	Animal_model	Mouse brain tissue	PMID:26659803
Pro-inflammatory		Neuroinflammation	Cell culture	astrocyte cells	PMID:30336612
Pro-inflammatory		Neuroinflammation CKD	Cell culture	astrocyte cells	PMID:26659803
ROS production	induces	Neuroinflammation	Cell culture	astrocyte cells	PMID:26659803
ROS production	induces	CKD	Cell culture	renal tubular cells	PMID:29474405

Reactions view

Reaction ID	Precursor	Product	Enzyme	Reaction type	Reference
MMDBr0000005	Tryptophan	Indole	tRNA	Carbon-carbon lyase (elimination)	PathBank KEGG
MMDBr0000008	Indole	Indoxyl sulfate	CYP2E1	Oxidation; Phase-I biotransformation	PMID: 11808865 MetaCyc
MMDBr0000009			SULT1A1	Conjugation; Phase-II biotransformation	PathBank

Pathway view

Indoxyl sulfate is an indole compound that is formed through gut microbial metabolism from dietary tryptophan and a sulfation reaction in liver hepatic cells. After being transported into gut microbes, tryptophan undergoes a reaction with the enzyme tryptophanase to form indole. Indole that is produced from the gut microbes then enters systemic circulation. Ultimately this compound undergoes a sulfation reaction in a liver hepatic cell by a sulfotransferase enzyme to form Indoxyl sulfate, which this compound returns back into systemic circulation. It is shown to be a major uremic toxin through high levels of retention. Indoxyl sulfate, like indoxyl glucuronide, is shown to cause a reduction in Erythropoietin production which ultimately results in renal anemia. It is also shown to cause vascular calcification and disrupt the electron transport chain and oxidative phosphorylation causing muscle atrophy.

MiMeDB – A Look Inside

MiMeDB Browse ▾ Search ▾ Downloads About ▾ Contact Us Metabolites ▾ Search Search

Showing metabocard for Indoxyl sulfate (MMDBc0000661)

Record Information	
Version	1.0
Created at	2021-07-26 19:24:25 UTC
Updated at	2021-07-26 19:24:25 UTC
Mime	MMDBc0000661

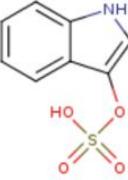
Metabolite Identification	
Common Name	Indoxyl sulfate
Description	Indoxyl sulfate, also known as 3-indoxyl sulfate, is a member of the class of organic compounds known as arylsulfates. These are organic compounds containing a sulfate group that carries an aryl group through an ether group. Indoxyl sulfate is a metabolite of the common amino acid tryptophan and is derived through the consumption, digestion and microbial processing of protein-rich foods. Indoxyl sulfate is technically a bacterial co-metabolite, meaning that it is derived from both bacterial and host metabolism. Specifically, it is generated from dietary L-tryptophan which is converted to indole in the large intestine via tryptophanase-expressing gastrointestinal bacteria (PMID: 27102537). The resulting indole is converted to indoxyl in the liver via enzyme-mediated hydroxylation by the CYP450 enzyme CYP2E1 (PMID 11808865). Subsequently, indoxyl is converted into indoxyl sulfate by the SULT1A1 sulfotransferase enzyme in the liver (PMID: 12064372). Indoxyl sulfate has been identified as a uremic toxin according to the European Uremic Toxin Working Group (PMID: 22626821) and is classified as a protein-bound uremic solute. Indoxyl sulfate is known to bind to serum albumin (PMID: 22626821), 22626821  , to be transported by the OAT1 transporter (PMID: 34678967 ) and to be an agonist for the arylhydrocarbon receptor (AhR) (PMID: 32527975 ). High concentrations of indoxyl sulfate in whole blood or blood plasma are known to be associated with the development and progression of chronic kidney disease (CKD) as well as the development of cardiovascular disease (CVD) in humans and other mammals (PMID: 28754616 ). As a uremic toxin, indoxyl sulfate is known to stimulate glomerular sclerosis (PMID: 8035108  , interstitial fibrosis (PMID: 33138205 ). Indoxyl sulfate upregulates signal transducers and activators of transcription 3 phosphorylation leading to increases in TGF- β 1, monocyte chemotactic protein-1 and alpha-smooth muscle actin production, all of which participate in interstitial inflammation, renal fibrosis and, consequently, CKD progression (PMID: 33138205 ). Indoxyl sulfate is also a known cardiotoxin (PMID: 30200452 ). In plasma, indoxyl

[Read more...](#)

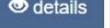
MiMeDB – Metabolite Page

MiMeDB Browse ▾ Search ▾ Downloads About ▾ Contact Us Metabolites ▾ Search Search

Showing metabocard for Indoxyl sulfate (MMDBc0000661)

Structure	 Read more...														
	 MOL 3D MOL SDF 3D SDF PDB 3D PDB SMILES InChI														
Synonyms	<table><thead><tr><th>Value</th><th>Source</th></tr></thead><tbody><tr><td>3-Indolyl hydrogen sulfate</td><td>ChEBI</td></tr><tr><td>3-Indolyl sulfate</td><td>ChEBI</td></tr><tr><td>3-Indoxyl sulfate</td><td>ChEBI</td></tr><tr><td>3-Indoxylsulfuric acid</td><td>ChEBI</td></tr><tr><td>Indican</td><td>ChEBI</td></tr><tr><td>indol-3-yl Hydrogen sulfate</td><td>ChEBI</td></tr></tbody></table> Show more...	Value	Source	3-Indolyl hydrogen sulfate	ChEBI	3-Indolyl sulfate	ChEBI	3-Indoxyl sulfate	ChEBI	3-Indoxylsulfuric acid	ChEBI	Indican	ChEBI	indol-3-yl Hydrogen sulfate	ChEBI
Value	Source														
3-Indolyl hydrogen sulfate	ChEBI														
3-Indolyl sulfate	ChEBI														
3-Indoxyl sulfate	ChEBI														
3-Indoxylsulfuric acid	ChEBI														
Indican	ChEBI														
indol-3-yl Hydrogen sulfate	ChEBI														
Chemical Formula	C ₈ H ₇ NO ₄ S														
Average Molecular Weight	213.21														
Monoisotopic Molecular Weight	213.009578407														
IUPAC Name	(1H-indol-3-yl)oxidanesulfonic acid														
Traditional Name	3-sulfooxy-1H-indole														

MiMeDB – Host Source

MiMeDB	Browse ▾	Search ▾	Downloads	About ▾	Contact Us	Metabolites ▾	Search	Search
Detected in Host								▲
	Host and Biospecimen	Status	Concentration	Age	Sex	Health condition	Reference	
Human urine	Detected but not Quantified			adult	Both	type 2 diabetes; high CVD risk	28692847	 details
Human urine	Detected but not Quantified			adult	Both	unknown	30087103	 details
Human blood	Detected but not Quantified			adult	Both	CKD	30087103	 details
Human urine	Detected and Quantified	36.932 +/- 21.826 micromol/mmol creatinine	1-13yo		Both	Eosinophilic esophagitis	HMDB	 details
Human blood	Detected and Quantified	14 +/- 4.2 micromolar	adult		Male	healthy	HMDB	 details
Human blood	Detected and Quantified	2.49 +/- 1.36 micromolar	1-13yo		Both	healthy	HMDB	 details
Human urine	Detected and Quantified	129 (45-191) micromol/mmol creatinine	adult		Both	healthy	HMDB	 details
Human urine	Detected and Quantified	19.74 +/- 5.26 micromol/mmol creatinine	adult		Male	healthy	HMDB	 details
Human urine	Detected and Quantified	17.76 +/- 6.58 micromol/mmol creatinine	adult		Female	healthy	HMDB	 details

MiMeDB – Health Effects

MiMeDB Browse ▾ Search ▾ Downloads About ▾ Contact Us Metabolites ▾ Search Search

Health and Bioactivity

	Health outcome	Metabolite Response	Related Health Condition	Evidence Type	Measured in Matrix	Reference	
	Mild cognitive impairment	Increased	CKD	Association	Human serum	PMID:26797588	<button>details</button>
	Bioactivity	Metabolite effect	Related Health Condition	Evidence Type	Measured in Matrix	Reference	
	Nuclear factor-kb	agonist	Neuroinflammation CKD	Cell culture	astrocyte cells	PMID:28659803	<button>details</button>
	Aryl hydrocarbon receptor	agonist		Cell culture	Mouse CaCo-2 intestinal cells	PMID:34137156	<button>details</button>
	Aryl hydrocarbon receptor	agonist	Neuroinflammation	Cell culture	astrocyte cells	PMID:30336612	<button>details</button>
	Neurotoxin		Neuroinflammation CKD	Animal_model	Mouse brain tissue	PMID:28659803	<button>details</button>
	Pro-inflammatory		CKD	Cell culture	renal tubular cells	PMID: 22610984	<button>details</button>
	Pro-inflammatory		Neuroinflammation	Cell culture	astrocyte cells	PMID:30336612	<button>details</button>
	Pro-inflammatory		Neuroinflammation CKD	Animal_model	Mouse brain tissue	PMID:28659803	<button>details</button>
	ROS production	induces	Neuroinflammation CKD	Cell culture	astrocyte cells	PMID:28659803	<button>details</button>
	ROS production	induces	CKD	Cell culture	renal tubular cells	PMID: 29474405	<button>details</button>

MiMeDB – Reactions

MiMeDB Browse ▾ Search ▾ Downloads About ▾ Contact Us Metabolites ▾ Search Search

Showing metabocard for Indoxyl sulfate (MMDBc0000661)

Record Information

Metabolite Identification

Related Microbes

	Microbe ID	Organism	Kingdom	Phylum	Host and Body site	Microbial Link	Reference
--	------------	----------	---------	--------	--------------------	----------------	-----------

Metabolic Reactions

	Reaction ID	Precursor	Product	Enzyme	Reaction type	Reference
	MMDBr0000005	Tryptophan	Indole	tnaA	Carbon-carbon lyase (elimination)	PathBank KEGG
	MMDBr0000008	Indole	Indoxyl	CYP2E1	Oxidation; Phase-I biotransformation	PMID:11808865 MetaCyc
	MMDBr0000009	Indoxyl	Indoxyl sulfate	SULT1A1	Conjugation; Phase-II biotransformation	PathBank

Exposure Sources

	Exposure Type	Exposure Source	Metabolite Response	Measured in Matrix	Health Condition	Evidence Type	Reference	
	Diet	high dietary protein:fiber ratio	increased	Human serum	CKD	correlation	PMID:26026209	details
	Diet	Mediterranean diet	decreased	Human urine	type 2 diabetes; high CVD risk	correlation	PMID:28692847	details

MiMeDB – Food Sources

MiMeDB Browse ▾ Search ▾ Downloads About ▾ Contact Us Metabolites ▾ Search Search

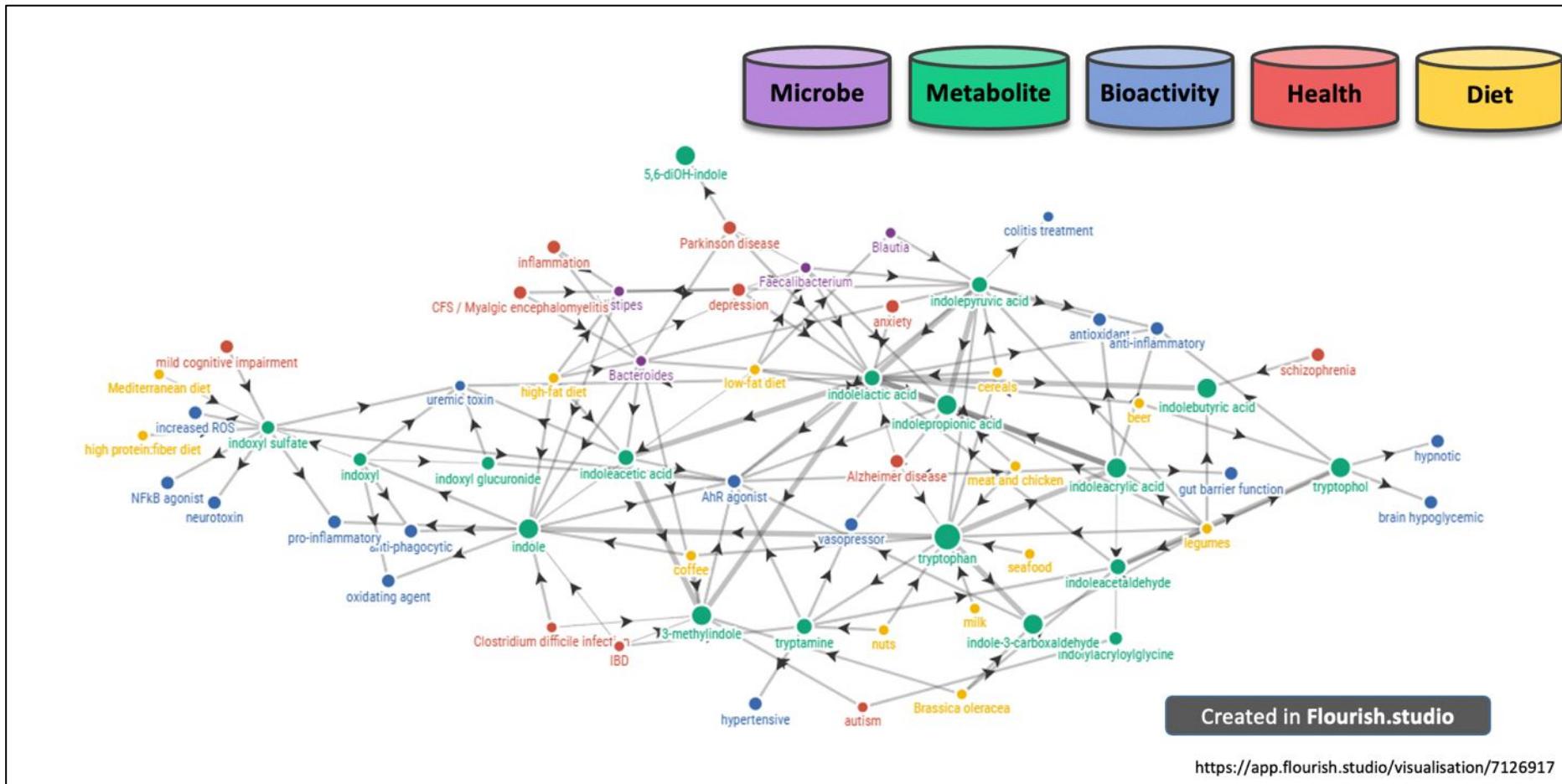
Health and Bioactivity

Health Type	Evidence Type	Related Health Condition	Metabolite Response	Measured in Matrix	Reference	
Health outcome	Association	Alzheimer disease	decreased	Human plasma	PMID:33269100	

Exposure Sources

Exposure Type	Exposure Source	Source Sub-type	Species	Taxonomy	Reference	
Ingestion	Animal	Fish	Salmon	NCBI:504568	FooDB	
Ingestion	Animal	Fish	Smelt	NCBI:8012	FooDB	
Ingestion	Animal	Crustaceans	Snow crab	NCBI:41210	FooDB	
Ingestion	Animal		Cow	NCBI:9913	FooDB	
Ingestion	Food	Coffee	Coffea arabica	NCBI:13443	FooDB	
Ingestion	Food	Tofu	Soy	NCBI:3847	FooDB	
Ingestion	Plant		Cashew nuts	NCBI:171929	FooDB	
Ingestion	Plant	Bean	Cocoa	NCBI:3641	FooDB	
Ingestion	Plant		Fennel	NCBI:48038	FooDB	
Ingestion	Plant		Oat	NCBI:50455	FooDB	

MiMeDB - Network View



MiMeDB Applications

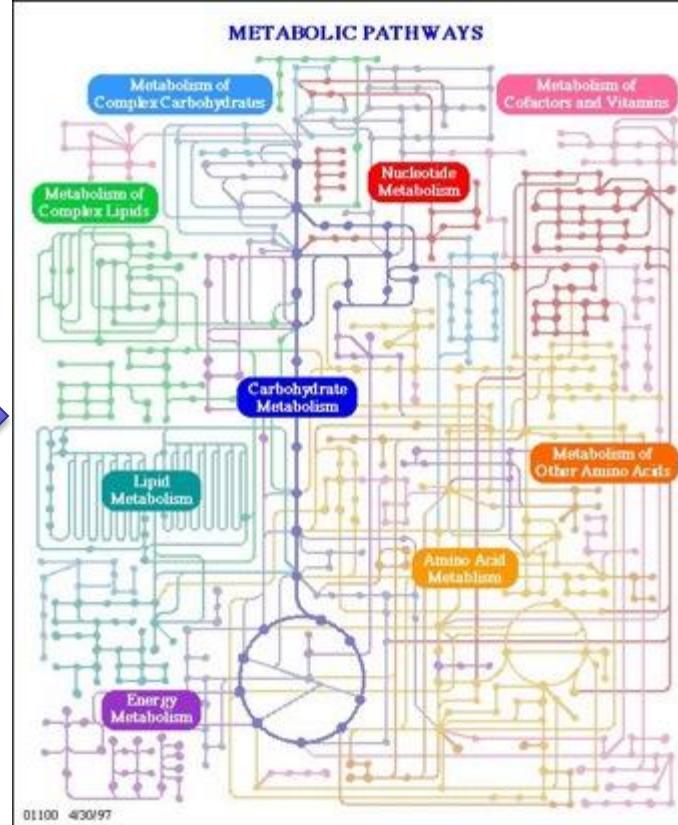
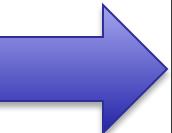
- Used to help characterize microbially derived metabolites
- Used to help characterize microbially derived food metabolites
- Used to characterize microbial genomes and to determine their metabolic functions
- Used to connect microbes and microbial metabolites to health and disease

Organism & Purpose-Specific Databases

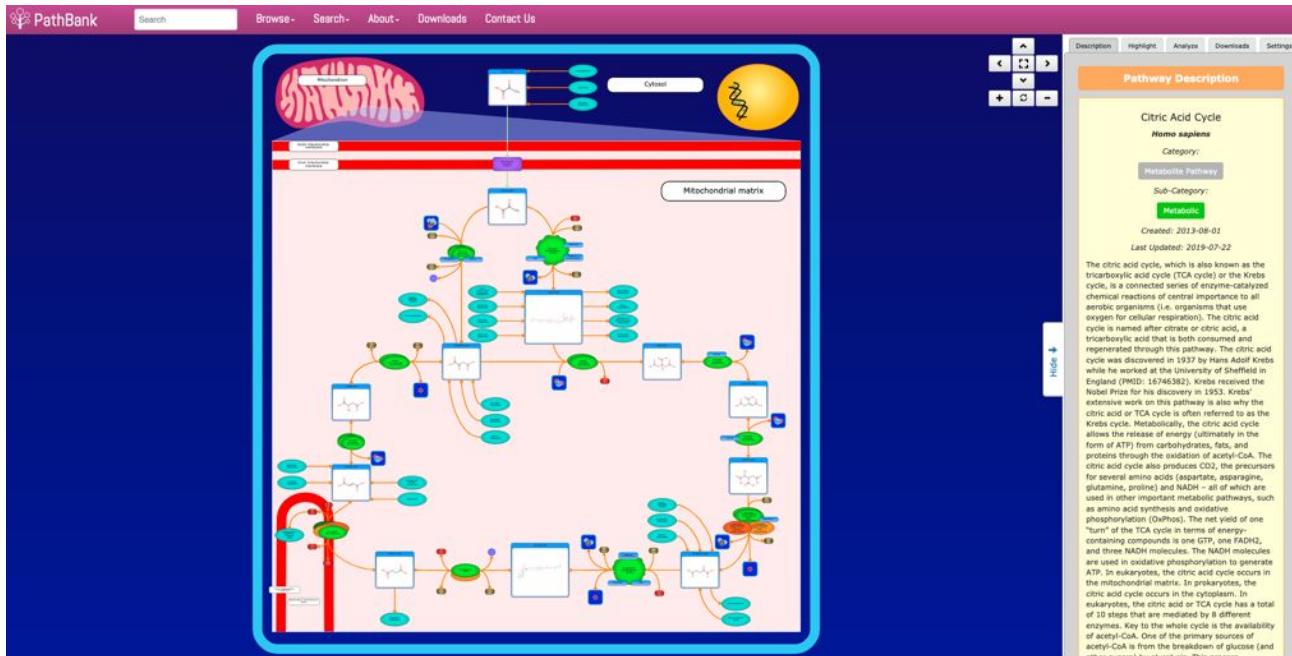
- These databases provide key information on chemical provenance and origin for many metabolites
- Specific to the biological system being studied
- Avoid the trap of mis-identifying biological impossible or chemically infeasible compounds
- Also allows more direct interpretation of biology or biological context

From Lists to Pathways

Compound	Retention Time (min)	Conc. in Urine (μM)	Compound	Retention Time (min)	Conc. in Urine (μM)
Dns-o-phospho -L-serine	0.92	<DL *	Dns-le	6.35	25
Dns-o-phospho -L-tyrosine	0.95	<DL	Dns-3-aminosalicylic acid	6.44	0.5
Dns-adenosine monophosphate	0.99	<DL	Dns-pipeolic acid	6.50	0.5
Dns-o-phosphoethanolamine	1.06	16	Dns-Leu	6.54	54
Dns-glucosamine	1.06	22	Dns-cystathione	6.54	0.3
Dns-o-phospho -L-threonine	1.09	<DL	Dns-Leu-Pro	6.60	0.4
Dns-6-dimethyllysine purine	1.20	<DL	Dns-5-hydroxylysine	6.65	1.6
Dns-3-methyl -histidine	1.22	80	Dns-Cysteine	6.73	160
Dns-taurine	1.25	834	Dns-N-norleucine	6.81	0.1
Dns-carnosine	1.34	28	Dns-5-hydroxydopamine	7.17	<DL
Dns-Arg	1.53	36	Dns-dimethylamine	7.33	293
Dns-Asn	1.55	133	Dns-5-HIAA	7.46	18
Dns-hypotaurine	1.58	10	Dns-umbelliferone	7.47	1.9
Dns-homocarnosine	1.61	3.9	Dns-2,3-diaminopropionic acid	7.63	<DL
Dns-guanidine	1.62	<DL	Dns-L-ornithine	7.70	15
Dns-Gln	1.72	633	Dns-4-acetylphenol	7.73	51
Dns-allantoin	1.83	3.8	Dns-procaine	7.73	8.9
Dns-L-citrulline	1.87	2.9	Dns-homocysteine	7.76	3.3
Dns-1-(or 3-)methylhistamine	1.94	1.9	Dns-acetaminophen	7.97	82
Dns-adenosine	2.06	2.6	Dns-Phe-Phe	8.03	0.4
Dns-methylguanidine	2.20	<DL	Dns-5-methoxy salicylic acid	8.04	2.1
Dns-Ser	2.24	511	Dns-Lys	8.16	184
Dns-aspartic acid amide	2.44	26	Dns-aniline	8.17	<DL
Dns-4-hydroxy -proline	2.56	2.3	Dns-leu-Phe	8.22	0.3
Dns-Glu	2.57	21	Dns-His	8.35	1560
Dns-Asp	2.60	90	Dns-4-hialsine	8.37	<DL
Dns-Thr	3.03	157	Dns-benzylamine	8.38	<DL
Dns-epinephrine	3.05	<DL	Dns-1-ephedrine	8.50	0.6
Dns-ethanolamine	3.11	471	Dns-tryptamine	8.63	0.4
Dns-aminoadipic acid	3.17	70	Dns-pyridoxamine	8.94	<DL
Dns-Gly	3.43	2510	Dns-2-methyl -benzylamine	9.24	<DL
Dns-Ala	3.68	593	Dns-5-hydroxytryptophan	9.25	0.12
Dns-aminolevulinic acid	3.97	30	Dns-1,3-diaminopropane	9.44	0.23
Dns-r-amino -butyric acid	3.98	4.6	Dns-purescine	9.60	0.5
Dns-p-amino -hippuric acid	3.98	2.9	Dns-1,2-diaminopropane	9.66	0.1
Dns-5-hydroxy methyluricil	4.58	1.9	Dns-tyrosinamide	9.79	29
Dns-tryptophanamide	4.70	5.5	Dns-dopamine	10.08	140
Dns-isoguanine	4.75	<DL	Dns-cadaverine	10.08	0.08
Dns-5-amino pentanoic acid	4.79	1.6	Dns-histamine	10.19	0.4
Dns-sarcosine	4.81	7.2	Dns-3-methoxy -tyramine	10.19	9.2
Dns-3-amino -isobutyrate	4.81	85	Dns-Tyr	10.28	321
Dns-2-aminobutyric acid	4.91	17	Dns-cysteamine	10.44	<DL



Pathway Databases



Pathway DBs

- Different from “organism-specific” or purpose-specific databases which are primarily text based
- Pathway DBs provide a rich source of visual data that relates metabolites to genes, proteins, diseases, signaling events and processes
- Provide various tools to permit visualization and gene/metabolite mapping
- Often cover multiple species

KEGG – Kyoto Encyclopedia of Genes and Genomes

The screenshot displays the KEGG homepage and a detailed view of a metabolic pathway.

KEGG Home: Includes links to Release notes, Current statistics, Plea from KEGG, KEGG Database, KEGG overview, Searching KEGG, KEGG mapping, Color codes, KEGG Objects, Pathway maps, Brite hierarchies, KEGG DB links, KEGG Software, KEGG API, KGML, KEGG FTP, Subscription, GenomeNet, DBGET/LinkDB, Feedback, and Copyright request.

KEGG: Kyoto Encyclopedia of Genes and Genomes: A brief introduction stating KEGG is a database resource for understanding the biological system, molecular-level information, genome sequencing, and other high-throughput data. It also links to Release notes (May 1, 2019).

Main entry point to the KEGG: KEGG2, KEGG Table.

Data-oriented entry points:

- KEGG PATHWAY:** KEGG pathway maps
- KEGG BRITE:** BRITE hierarchies and tables
- KEGG MODULE:** KEGG modules
- KEGG ORTHOLOGY:** KO functional orthologs [Annotation]
- KEGG GENOME:** Genomes [Pathogen | Virus | Plant]
- KEGG GENES:** Genes and proteins [SeqData]
- KEGG COMPOUND:** Small molecules
- KEGG GLYCAN:** Glycans
- KEGG REACTION:** Biochemical reactions [RModule]
- KEGG ENZYME:** Enzyme nomenclature
- KEGG NETWORK:** Disease-related network elements
- KEGG DISEASE:** Human diseases [Cancer]
- KEGG DRUG:** Drugs [New drug approvals]

Classification: Pathway, Brite, Brite table, Module, KO (Function), Organism, Compound, Network, Disease (ICD), Drug (ATC), Drug (Target).

Pathway Map: A complex diagram showing the biosynthesis of 2-Acetamidoethyl-phosphonate. Key nodes include Phosphono-alanine, Rhizococcin, FomC, PhpC, Fom3, 2-Hydroxyethyl-phosphonate, 2-Aminoethyl-phosphonate, Glycine, Lipophosphoglycan, Ciliatocolate, Ceramide ciliatine, Diacylglyceryl-2-aminoethyl-phosphonate, and various methylphosphonates. Enzymes involved are 3.11.1.1, 3.11.1.2, 1.2.1., 2.6.1., 2.3.3.19, 4.1.1.82, 2.6.1.37, 1.14.11.45, 1.13.11.78, 23.1.280, 1.13.11.72, 1.13.11.73, 2.7.8.37, 3.6.1.63, 2.7.7.14, 2.7.8.1, and 2.7.8.3.

Compound Details: A detailed view for compound C00198, D-Glucuronic acid-1,5-lactone. It includes the entry name, formula (C6H10O6), mass (178.0479), and structure (a five-membered lactone ring with hydroxyl groups at positions 1 and 5). It also lists related pathways, enzymes, and diseases.

<http://www.genome.jp/kegg/>

KEGG

- The “Go-to” Metabolic Pathway Database
- Has 535 “canonical” pathway diagrams or maps covering 5994 organisms for a total of 604,808 pathways
- ~170 metabolic pathways covering 18,553 compounds, includes many disease pathways (80), protein signaling (70) pathways, and biological process pathways (70)
- Metabolic pathways are highly schematized and mostly limited to catabolic and anabolic processes

Reactome

The screenshot shows the Reactome Pathway Browser interface for the Pentose phosphate pathway (R-HSA-71336) in Homo sapiens. The main panel displays a complex network of biological reactions and their interactions. Key components include D-Glucose 6-phosphate, D-Glucono-1,5-lactone 6-phosphate, Fructose 6-phosphate, and Ribulose 5-phosphate. Reactions are catalyzed by enzymes such as G6PDH, PGL, RPE, TKT, and FBPase. NADPH and NADH are shown as electron donors and acceptors. A sidebar on the left provides an event hierarchy and a detailed description of the pathway, mentioning its role in biosynthetic reactions and nucleotide synthesis.

PB | Pentose phosphate pathway

Pathways for: Homo sapiens

Analysis: Tour: Layout:

Event Hierarchy:

- Autophagy
- Cell Cycle
- Cell-Cell communication
- Cellular responses to external stimuli
- Chromatin organization
- Circadian Clock
- Developmental Biology
- Digestion and absorption
- Disease
- DNA Repair
- DNA Replication
- Extracellular matrix organization
- Gene expression (Transcription)
- Hemostasis
- Immune System
- Metabolism
 - Metabolism of carbohydrates
 - Glycogen metabolism
 - Glucose metabolism
 - Fructose metabolism
 - Lactose synthesis
 - Galactose catabolism
 - Pentose phosphate pathway

Search for a term, e.g. pten ...

Description Molecules 46 Structures Expression Analysis Downloads

Pentose phosphate pathway Id: R-HSA-71336.6 Species: Homo sapiens

DOI: 10.3180/REACT_1859.1 Summation

The pentose phosphate pathway is responsible for the generation of a substantial fraction of the cytoplasmic NADPH required for biosynthetic reactions, and for the generation of ribose 5-phosphate for nucleotide synthesis. Although the pentose phosphate pathway and glycolysis are distinct, they involve three common intermediates, glucose 6-phosphate, glyceraldehyde 3-phosphate, and fructose 6-phosphate, so the two pathways are interconnected. The pentose phosphate pathway consists of eight reactions: 1. Conversion glucose 6-phosphate to D-glucono-1,5-lactone 6-phosphate, with the formation of NADPH; 2. Conversion of D-glucono-1,5-lactone 6-phosphate to 6-phospho-D-gluconate; 3. Conversion of 6-phospho-D-gluconate to ribulose 5-phosphate, with the formation of NADPH; 4. Conversion of ribulose 5-phosphate to xylulose 5-phosphate; 5. Conversion of xylulose 5-phosphate to 2-keto-3-deoxy-6-phosphogluconate; 6. Conversion of 2-keto-3-deoxy-6-phosphogluconate to ribose 5-phosphate, with the formation of NADPH; 7. Conversion of ribose 5-phosphate to D-glyceraldehyde 3-phosphate, with the formation of NADPH; 8. Conversion of D-glyceraldehyde 3-phosphate to glyceraldehyde 3-phosphate, with the formation of NADPH.

<https://reactome.org/>

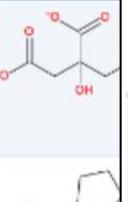
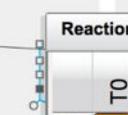
Reactome

- Pathway database maintained by the EBI and the Ontario Institute for Cancer Research (OICR)
- Covers 15 model organisms with an average of 1500 pathways each
- Includes disease, metabolism, signaling, apoptosis, transcription pathways
- Mostly focused on protein pathways or protein-protein interactions

BioCyc

ADD TRANSFORM COLUMN ADD PROPERTY COLUMN ENRICHMENTS

Prev 1 2 3 4 5 6 Next Show all

	column 1	column 2	column 3	Monoisotopic-Molecular-Weight	Structure of comp
citrate	6.03	-6.76	192.02701		
nicotine-1'-N-oxide					
glutarate semialdehyde					

Enrichment parameters

Analysis type: Enrichment Depletion Enrichment and Depletion

Include results whose p-value less than: 0.1

Algorithm details:

Statistic: Fisher Exact Fisher Exact Parent-Child Union Fisher Exact Parent-Child Intersection None Bonferroni Correction Benjamini-Hochberg Correction Benjamini-Yekutieli Correction

Correction

Fermentation **Carbohydrates Degradation** **Amino Acids Degradation**

Reaction Summary

pykA T0 T0.5 T1 T2

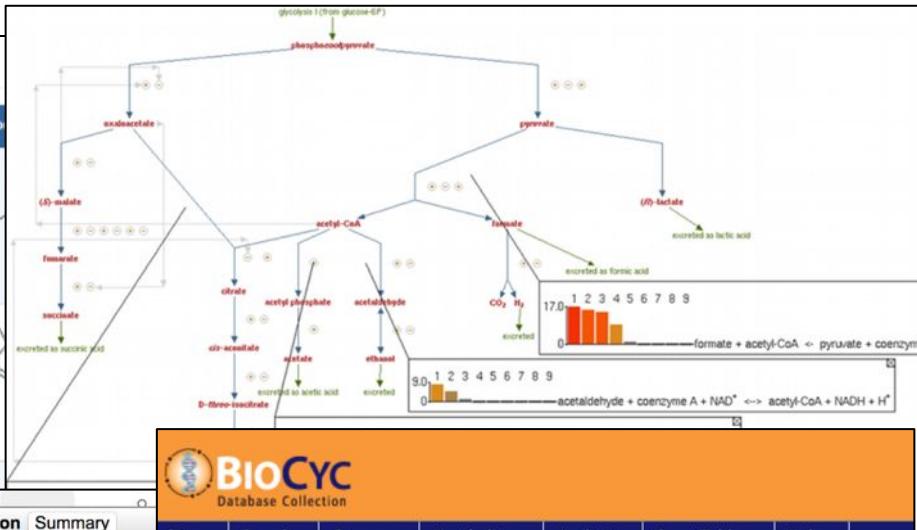
pykF T0 T0.5 T1 T2

Fatty Acids and Lipids Degradation **Reaction Summary**

pflB T0 T0.5 T1 T2

aceB T0 T0.5 T1 T2

Figure showing reaction summaries for pykA, pykF, and pflB across time points T0, T0.5, T1, and T2.



glycolysis I (from glucose-6P)

phosphoenolpyruvate

pyruvate

(R) lactate

(S) lactate

excreted as lactic acid

formate

succinate

acetyl-CoA

acetate

CO₂ + H₂

excreted as formic acid

D-threo-isocitrate

citrate

acetyl phosphate

acetate

ethanol

excreted

excreted as acetic acid

17.0 1 2 3 4 5 6 7 8 9

formate + acetyl-CoA <- pyruvate + coenzyme A

9.0 1 2 3 4 5 6 7 8 9

acetaldehyde + coenzyme A + NAD⁺ <- acetyl-CoA + NADH + H⁺

BIOCYC
Database Collection

Sites ▾ | Search ▾ | Genome ▾ | Metabolism ▾ | Analysis ▾ | SmartTables ▾ | Help ▾

BioCyc Database Collection

BioCyc is a collection of 14728 Pathway/Genome Databases (PGDBs), plus software tools for exploring them [Karp17]. Key aspects of BioCyc data:

- Quality data curated from tens of thousands of publications, including curated databases for *E. coli*, *B. subtilis*, *H. sapiens*, and *S. cerevisiae*.
- Computationally predicted metabolic pathways and operons.
- Data integrated from other databases including gene essentiality, regulatory networks, protein features, and GO annotations.

Current version: 23.0, released on April 29, 2019.

[Update History](#)

BioCyc Collection

- Contains 14,728 pathway databases including 7 manually annotated databases, 71 semi-manual and 14,650 automatically generated DBs
- Largely focused on metabolic pathways, does include some operon pathways
- Most pathways are very small/short compared to KEGG pathways
- Includes 250 pathways for humans, EcoCyc has 431 pathways for 2846 compounds

Limitations with Today's Pathway Databases

- Most pathway databases are limited to showing metabolic (catabolic or anabolic) events for metabolites
- Metabolites are more than just “bricks and mortar”, they play key roles in signaling, immune function, inflammation, homeostasis, epigenetic events, disease processes, drug action, tissue repair, etc.
- Is any of this information captured anywhere???

The Small Molecule Pathway Database (SMPDB)

The screenshot displays two views of the SMPDB website. The top view shows the homepage with a green 'Welcome to the Small Molecule Pathway Database' banner, a central molecular network diagram, and a pink 'Version 2.0' badge. The bottom view shows a detailed pathway page for '11-beta-hydroxylase deficiency' (SMP00575) and '17-Beta Hydroxysteroid Dehydrogenase III Deficiency' (SMP00356). Each page includes a small diagram, a list of associated metabolites and diseases, and links to BioPax and SVG+BioPax representations.

SMPDB

Welcome to the Small Molecule Pathway Database Version 2.0

Brought to you by the creators of the Human Metabolome Database (HMDB) and DrugBank

SMPDB (The Small Molecule Pathway Database) is an interactive, visual database containing pathways. More than 70% of these pathways (>433) are not found in any other pathway database. SMPDB pathways include information on protein cofactors, protein locations, metabolite locations, chemical structures and protein hyperlinked to detailed descriptions contained in the HMDB or DrugBank and each protein SMPDB pathways are accompanied with detailed descriptions and references, providing a wealth of information. The database is easily browsed and supports full text, sequence and metabolite queries. SMPDB with lists of metabolite names, drug names, genes/protein names, Swiss-Prot IDs, UniProt IDs, and microarray IDs. These queries will produce lists of matching pathways and highlight the matching pathways. Gene, metabolite and protein concentration data can also be visualized through SMPDB's pathway maps, descriptions and tables are downloadable.

Get started now: ★ Browse Pathway

SMP00575 11-beta-hydroxylase deficiency disease 11-Dehydrocorticosterone 3 beta-hydroxysteroid dehydrogenase...
11b,17a,21-Trihydroxyprog-neno... 3-oxo-5-beta-steroid 4-dehydro...
11b,21-Dihydroxy-3,20-oxo-5b-p... Aldo-keto reductase family 1 member...
11b,21-Dihydroxy-5b-pregnane-3... Cholesterol side-chain cleavage...
11b-Hydroxyprogesterone Corticosteroid 11-beta-dehydro...
(show all) (show all)

SMP00356 17-Beta Hydroxysteroid Dehydrogenase III Deficiency disease 17-beta-Estradiol-3-glucuronide 3 beta-hydroxysteroid dehydrogenase...
17-Hydroxyprogesterone 3-oxo-5-alpha-steroid 4-dehydro...
17a-Hydroxy pregnenolone 3-oxo-5-beta-steroid 4-dehydro...
19-Hydroxyandrost-4-ene-3,17-d... Cytochrome P450 19A1
(show all) Estradiol 17-beta-dehydrogenase...
(show all)

<http://www.smpdb.ca>

SMPDB

- **Nearly 48,900 hand-drawn small molecule pathways**
 - 404 drug action pathways
 - 20,251 metabolic disease pathways
 - 27,876 metabolic pathways
 - 160+ signaling and other pathways
- **Depicts organs, cell compartments, organelles, protein locations, and protein quaternary structures**
- **Maps gene chip & metabolomic data**
- **Converts gene, protein or chemical lists to pathways or disease diagnoses**

Exploring Pathways with SMPDB

The screenshot shows the SMPDB pathway viewer interface for the "Alanine Metabolism" pathway (SMP00055). The main view displays a complex metabolic network diagram with various metabolites (e.g., Alanine, Pyruvate, Alpha-ketoglutarate) and enzymes (e.g., Alanine transaminase, Alanine-glyoxylate transaminase) represented by icons. A large, semi-transparent red rectangle highlights a central portion of the pathway diagram, specifically the conversion of pyruvate to alanine. To the left of this highlighted area, a sidebar provides a detailed description of the pathway:

Pathway Description

Alanine Metabolism

Metabolic Pathway

Alanine is most commonly produced by the reductive amination of pyruvate via alanine transaminase. This reversible reaction involves the interconversion of alanine and pyruvate, coupled to the interconversion of alpha-ketoglutarate (2-oxoglutarate) and glutamate. Because transamination reactions are readily reversible and pyruvate is widespread, alanine can be easily formed in most tissues. Another route to the production of alanine is through the enzyme called alanine-glyoxylate transaminase. This reaction involves the interconversion of alanine and pyruvate, coupled to the interconversion of glyoxylate and glycine. Once synthesized, alanine can be coupled to alanyl tRNA via alanyl-tRNA synthetase and used by the body in protein synthesis. Alanine constitutes about 8% of human proteins. Under fasting conditions, alanine, derived from protein breakdown, can be converted to pyruvate and used to synthesize glucose via

On the left side of the main viewer, there is a sidebar titled "Chemical Structure" which lists various organic acids and their properties. At the top of the page, the URL "www.smpdb.ca/view/SMP00055" is visible in the browser's address bar.

Mapping Metabolites with SMPDB

The image displays two screenshots of the SMPDB (Systems Pathway Markup Language Database) platform. The left screenshot shows the SMP-MAP search interface, which includes a search bar, a list of search examples (e.g., TCA cycle), and a large central area showing a complex metabolic pathway map with various metabolites and enzymes. The right screenshot shows the SMPDB viewer interface, featuring a navigation menu (Search, About, Downloads, Contact Us), a toolbar (Description, Highlight, Analyze, Downloads, Settings), and a sidebar for highlighting specific compounds like Coenzyme Q10 and Hydrogen carbonate.

SMP-MAP New Search

You can enter your search terms, one on each line as shown in example.
1. You can also add concentration data which will show up in SMP-Analyzer by entering two columns, as shown in example.
2. The first column is the query term, the second column is the associated concentration value. These two columns must be divided by a tab character or a colon followed by a space (";").

Select a "Search by" type from the form below then click on an example to populate the form with an example search.

Example 1 (TCA cycle) Example 2 (TCA cycle with concentrations)

Search by Compound Names

Citric acid
Fumaric acid
L-Malic acid
Isocitric acid
Oxoglutaric acid
Oxaloacetic acid
Pyruvic acid
Succinic acid

SMPDB

Search About Downloads Contact Us

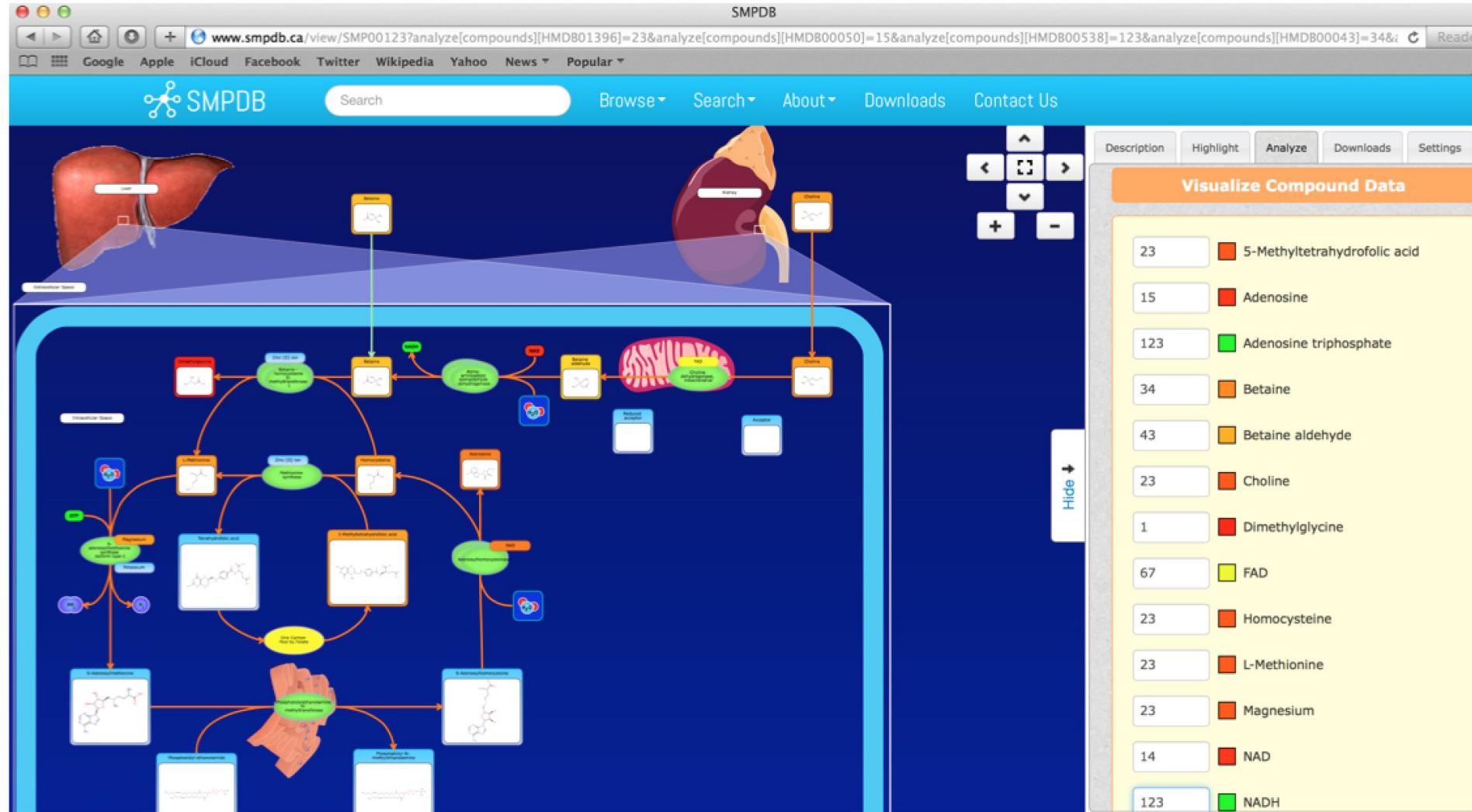
Description Highlight Analyze Downloads Settings

Highlighted elements will appear in red.

Highlight Compounds

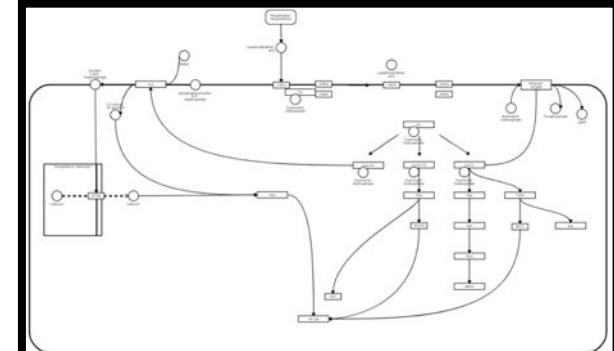
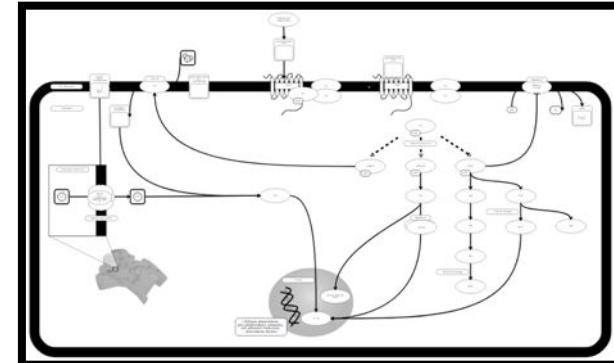
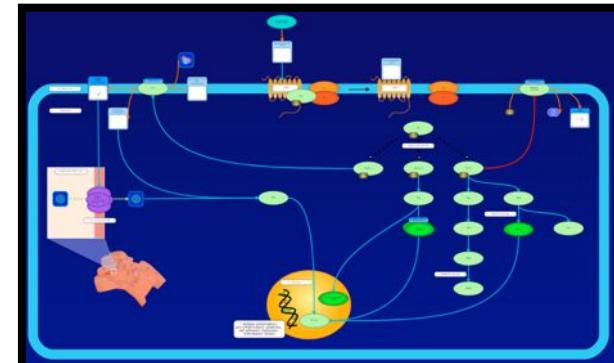
- 4Fe-4S
- Acetyl-CoA
- Adenosine triphosphate
- ADP
- Biotin
- Carbon dioxide
- cis-Aconitic acid
- Citric acid
- Coenzyme A
- Coenzyme Q10
- FAD
- FADH
- Fumaric acid
- Guanosine diphosphate
- Guanosine triphosphate
- Hydrogen carbonate

Mapping Metabolite/Gene Concentrations with SMPDB



SMPDB/HMDB Pathways

- Machine readable (SBML, SBGN, BioPAX, PWML)
- Multiple image formats (SVG, PNG)
- Multiple rendering styles (colour, B+W, KEGG-like, printer-friendly)
- All pathways have text descriptions & references and are linked to HMDB or UniProt
- All pathways are downloadable and community editable via PathWhiz



Citric Acid Cycle

PathBank

Search

Browse- Search- About- Downloads Contact Us

Description Highlight Analyze Downloads Settings

Mitochondrion

Cytosol

Mitochondrial matrix

Mitochondrial inner membrane

Mitochondrial outer membrane

Pathway Description

Citric Acid Cycle
Homo sapiens

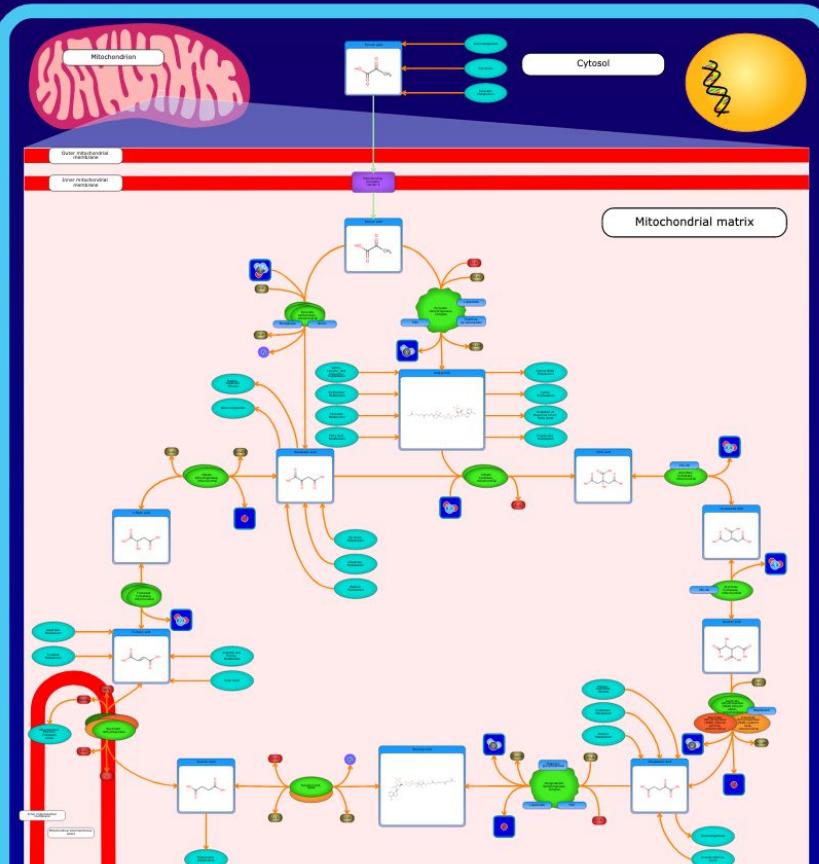
Category: Metabolite Pathway

Sub-Category: Metabolic

Created: 2013-08-01

Last Updated: 2019-07-22

The citric acid cycle, which is also known as the tricarboxylic acid cycle (TCA cycle) or the Krebs cycle, is a connected series of enzyme-catalyzed chemical reactions of central importance to all aerobic organisms (i.e. organisms that use oxygen for cellular respiration). The citric acid cycle is named after citrate or citric acid, a tricarboxylic acid that is both consumed and regenerated through this pathway. The citric acid cycle was discovered in 1937 by Hans Adolf Krebs while he worked at the University of Sheffield in England (PMID: 16746382). Krebs received the Nobel Prize for his discovery in 1953. Krebs' extensive work on this pathway is also why the citric acid or TCA cycle is often referred to as the Krebs cycle. Metabolically, the citric acid cycle allows the release of energy (ultimately in the form of ATP) from carbohydrates, fats, and proteins through the oxidation of acetyl-CoA. The citric acid cycle also produces CO₂, the precursors for several amino acids (aspartate, asparagine, glutamine, proline) and NADH – all of which are used in other important metabolic pathways, such as amino acid synthesis and oxidative phosphorylation (OxPhos). The net yield of one "turn" of the TCA cycle in terms of energy-containing compounds is one GTP, one FADH₂, and three NADH molecules. The NADH molecules are used in oxidative phosphorylation to generate ATP. In eukaryotes, the citric acid cycle occurs in the mitochondrial matrix. In prokaryotes, the citric acid cycle occurs in the cytoplasm. In eukaryotes, the citric acid or TCA cycle has a total of 10 steps that are mediated by 8 different enzymes. Key to the whole cycle is the availability of acetyl-CoA. One of the primary sources of acetyl-CoA is from the breakdown of glucose (and other sugars) by glycolysis. This process



Neuron Function

PathBank

Search Browse Search About Downloads Contact Us

Description Highlight Analyze Downloads Settings

Pathway Description

Neuron Function
Homo sapiens
Category: Metabolite Pathway
Sub-Category: Physiological
Created: 2018-09-27
Last Updated: 2019-06-06

Hide

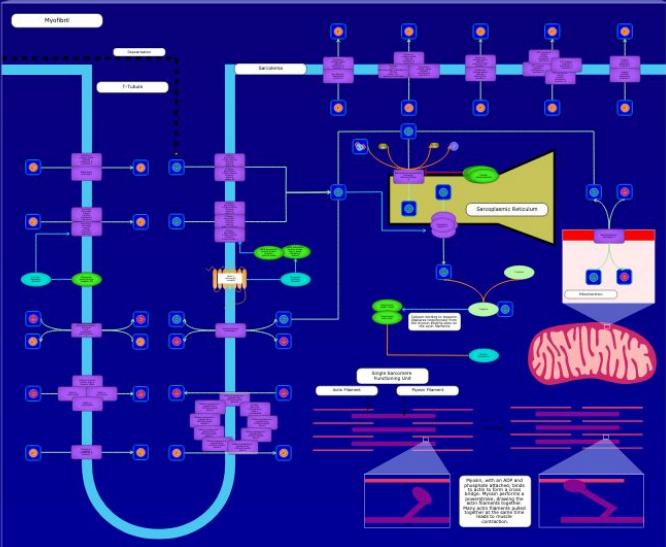
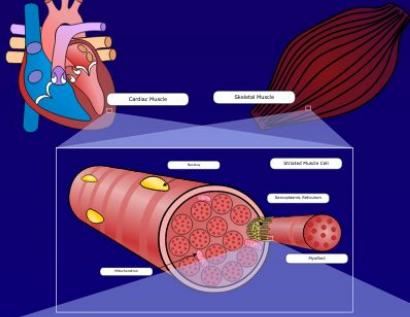
Neurons are electrically excitable cells that process and transmit information through electrical and chemical signals. A neuron consists of a cell body, branched dendrites to receive sensory information, and a long singular axon to transmit motor information. Signals travel from the axon of one neuron to the dendrite of another via a synapse. Neurons maintain a voltage gradient across their membrane using metabolically driven ion pumps and ion channels for charge-carrying ions, including sodium (Na^+), potassium (K^+), chloride (Cl^-), and calcium (Ca^{2+}). The resting membrane potential (charge) of a neuron is about -70 mV because there is an accumulation of more sodium ions outside the neuron compared to the number of potassium ions inside. If the membrane potential changes by a large enough amount, an electrochemical pulse called an action potential is generated. Stimuli such as pressure, stretch, and chemical transmitters can activate a neuron by causing specific ion-channels to open, changing the membrane potential. During this period, called depolarization, the sodium channels open to allow sodium to rush into the cell which results in the membrane potential to increase. Once the interior of the neuron becomes more positively charged, the sodium channels close and the potassium channels open to allow potassium to move out of the cell to try and restore the resting membrane potential (this stage is called repolarization).

Gastric Acid Function

PathBank

Search Browse Search About Downloads Contact Us

Description Highlight Analyze Downloads Settings



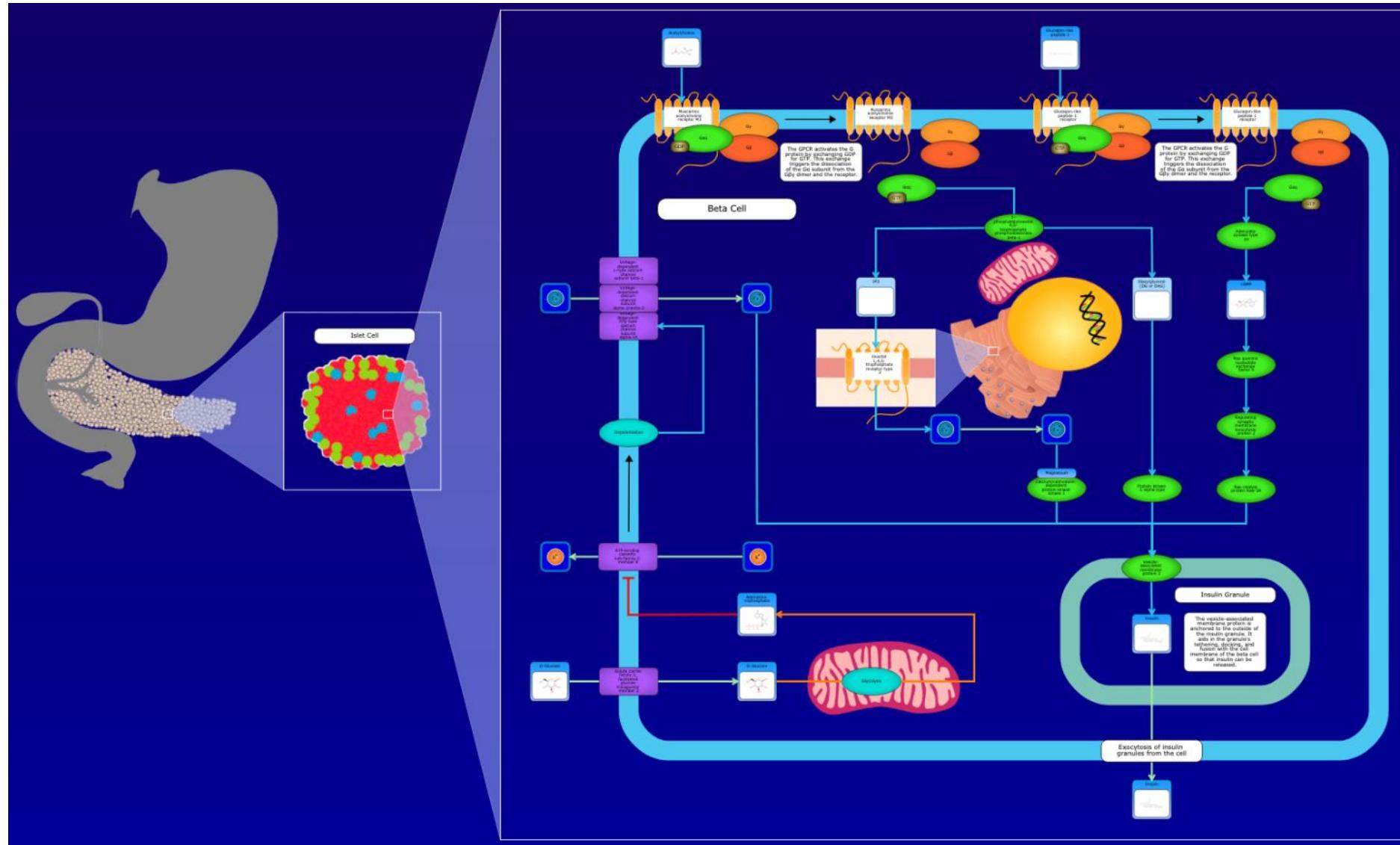
Pathway Description

Striated Muscle Contraction
Homo sapiens
Category: Metabolite Pathway
Sub-Category: Physiological
Created: 2013-09-04
Last Updated: 2019-06-06

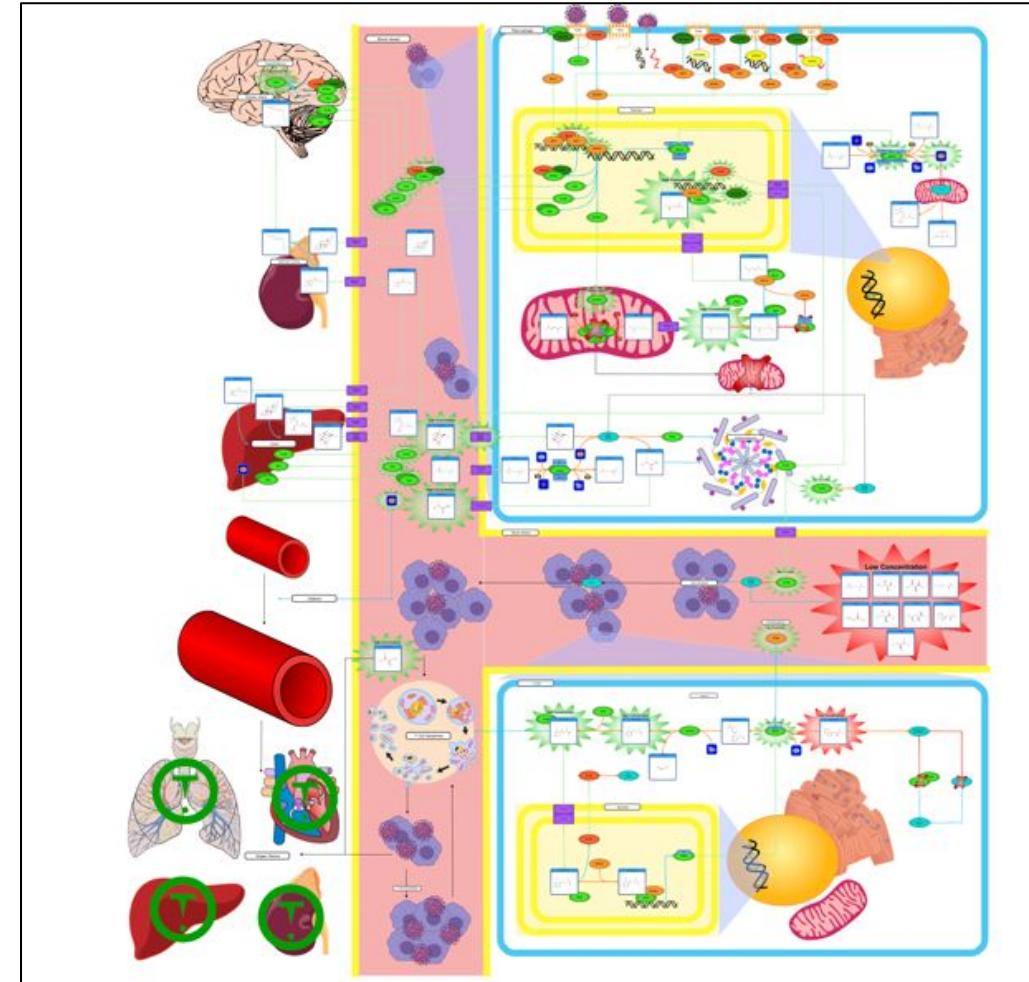
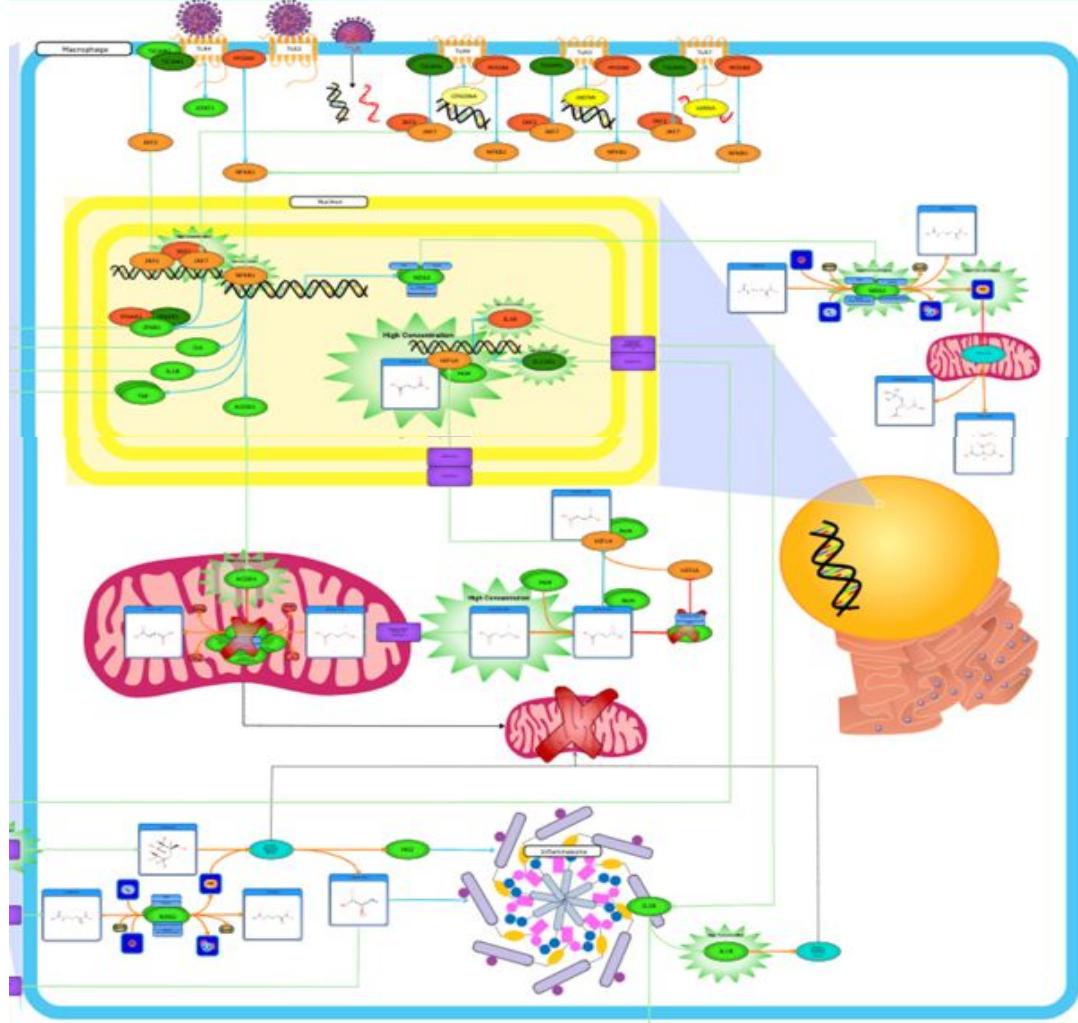
Hide ↴

Tubular striated muscle cells (i.e. skeletal and cardiac myocytes) are composed of bundles of rod-like myofibrils. Each individual myofibril consists of many repeating units called sarcomeres. These functional units, in turn, are composed of many alternating actin and myosin protein filaments that produce muscle contraction. The muscle contraction process is initiated when the muscle cell is depolarized enough for an action potential to occur. When acetylcholine is released from the motor neuron axon terminals that are adjacent to the muscle cells, it binds to receptors on the sarcolemma (muscle cell membrane), causing nicotinic acetylcholine receptors to be activated and the sodium/potassium channels to be opened. The fast influx of sodium and slow efflux of potassium through the channel causes depolarization. The resulting action potential that is generated travels along the sarcolemma and down the T-tubule, activating the L-type voltage-dependent calcium channels on the sarcolemma and ryanodine receptors on the sarcoplasmic reticulum. When these are activated, it triggers the release of calcium ions from the sarcoplasmic reticulum into the cytosol. From there, the calcium ions bind to the protein tropomodulin which displaces the tropomyosin filaments from the binding sites on the actin filaments. This allows for myosin filaments to be able to bind to the actin. According to the Sliding Filament Theory, the myosin heads that have an ADP and phosphate attached binds to the actin forming a cross-bridge. Once attached,

Pancreas Function – Beta Cell



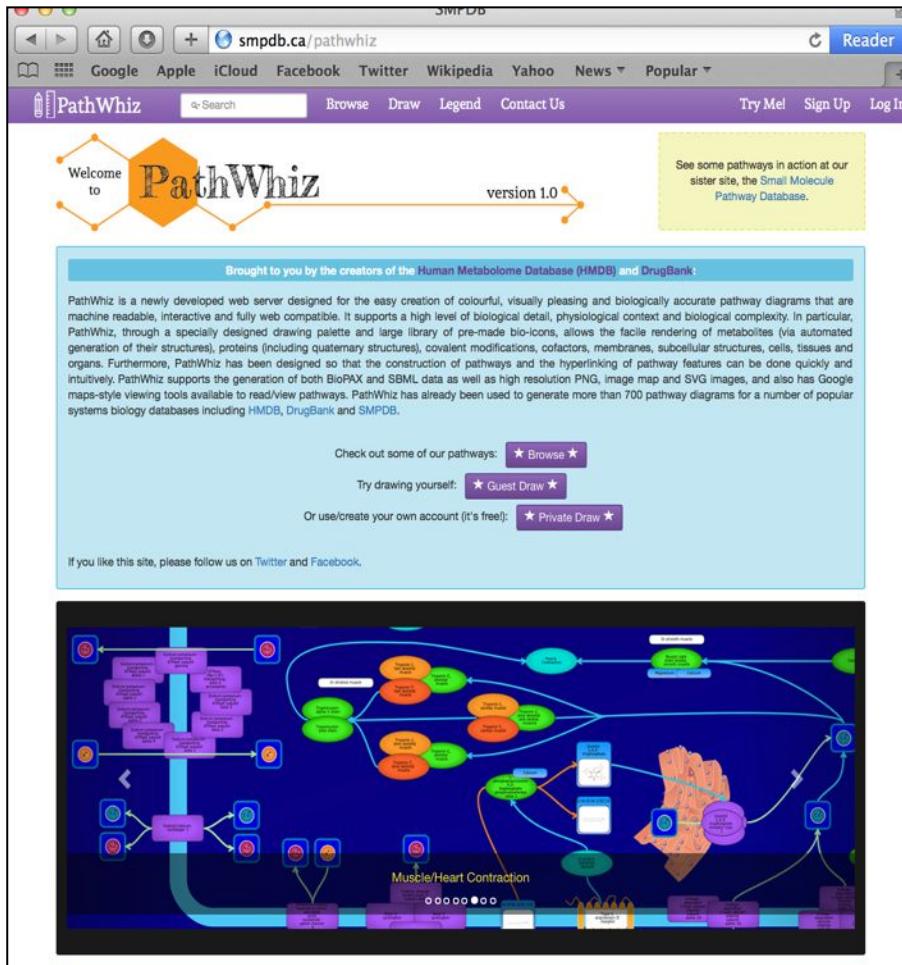
COVID-19 Induced Sepsis



How To Make Your Own Machine-Readable Pathway

**And Contribute to Common Pathway
Resources**

PathWhiz



- **Webswerver designed to permit creation of colourful, biologically accurate pathway diagrams that are machine readable and interactive**
- **Supports pathway replication and propagation along with many other palette-based tools for rapid, facile rendering & annotation**
- **Supports BioPAX, SBML and SBGN conversion as well as SVG and PNG image generation**

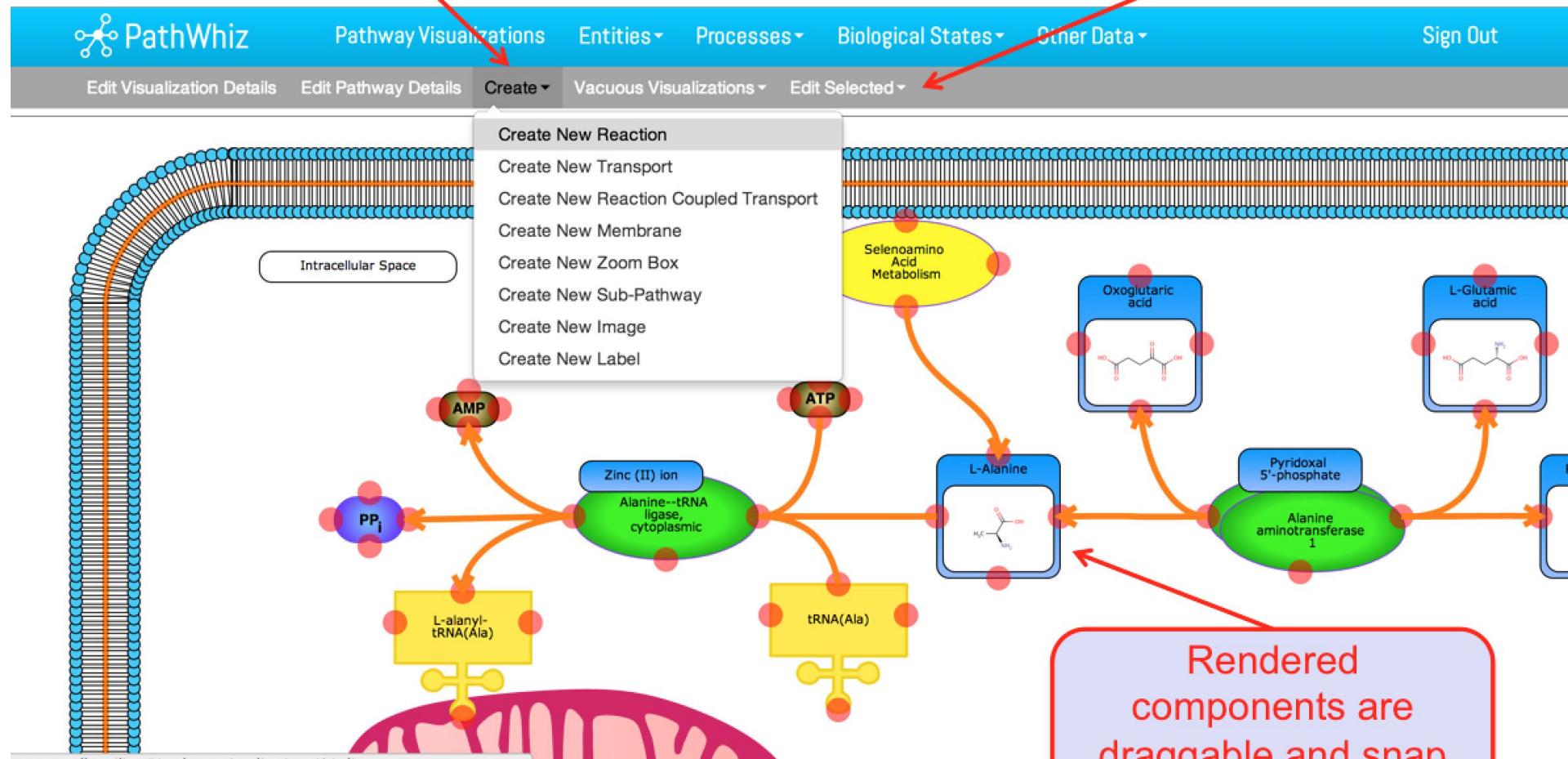
<http://smpdb.ca/pathwhiz>

Building Pathways with PathWhiz

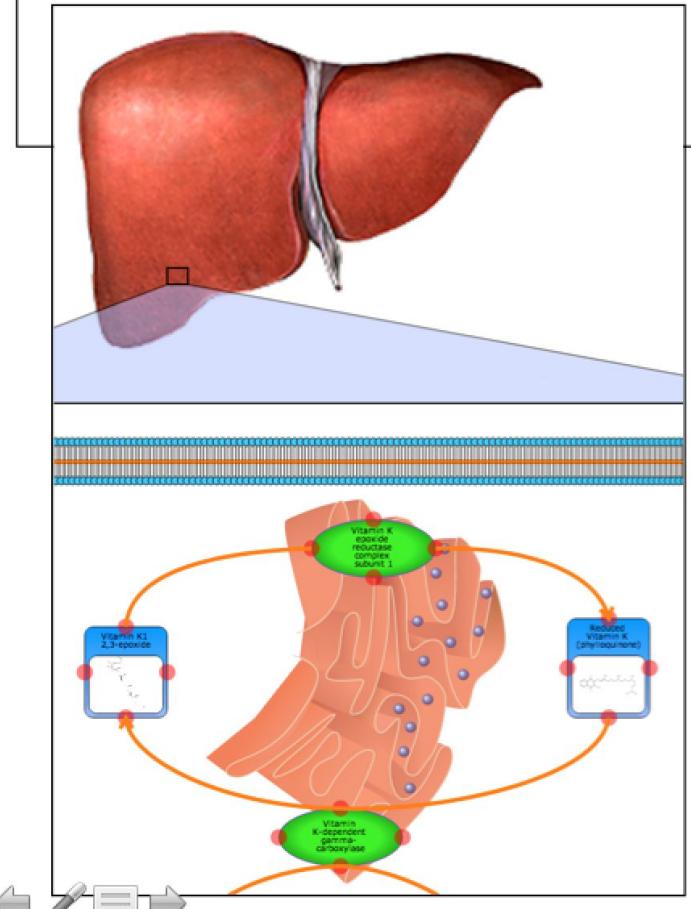
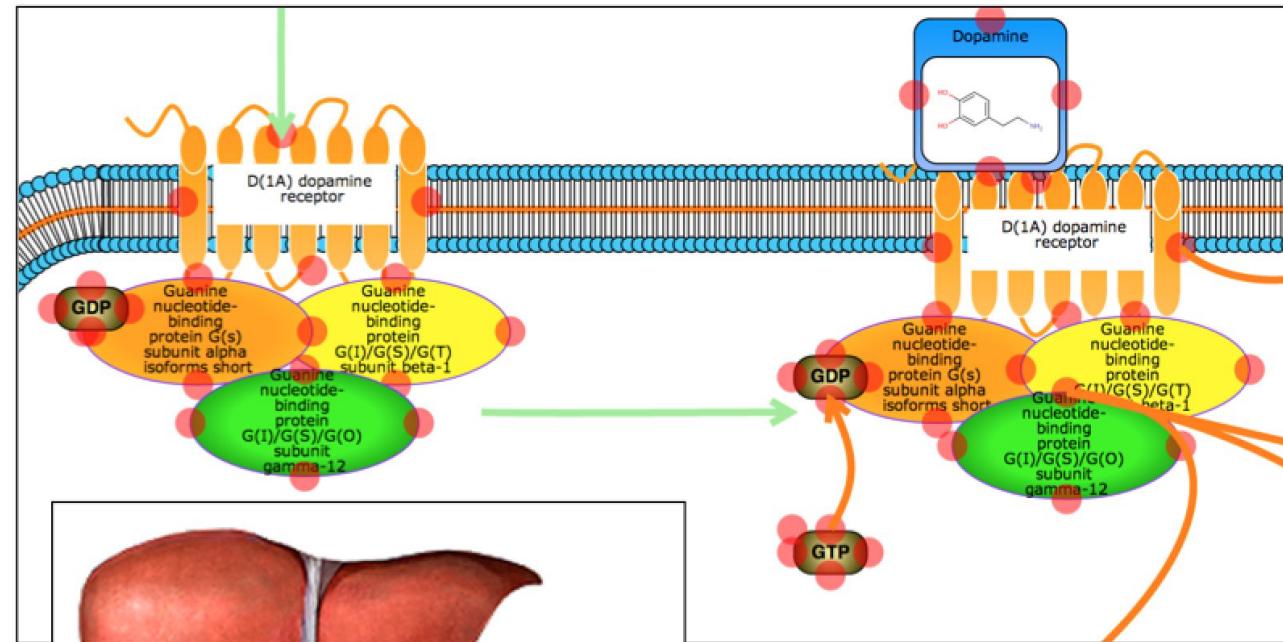
The screenshot illustrates the PathWhiz interface for building pathways. At the top, a navigation bar includes links for Pathway Visualizations, Entities, Processes, Biological States, Other Data, and Sign Out. Below the navigation bar are three main category panels: Entities, Processes, and Biological States. The Entities panel lists Compounds, Element Collections, Nucleic Acids (which is highlighted in blue), Proteins, Polypeptides, and Bound Elements. The Processes panel lists Reactions (which is highlighted in blue), Transports, and Reaction Coupled Transports. The Biological States panel lists Biological States, Organisms, Tissues, Cell Types (which is highlighted in blue), and Subcellular Locations. A large central window titled "Edit reaction" shows a reaction template with "Left Elements" containing L-Alanine and Glyoxylic acid, and "Right Elements" containing Glycine and Pyruvic acid. The "Direction" section shows a double-headed arrow between the left and right sides. Below the reaction template, there are sections for "Enzymes" (listing Serine--pyruvate aminotransferase) and "Update Reaction" (with "Update Reaction" and "Discard Changes" buttons). Two red callout boxes point to the "Add reaction elements" button (near the right side of the reaction template) and the "Add reaction enzymes" button (near the bottom left of the reaction template).

Different types of components to render

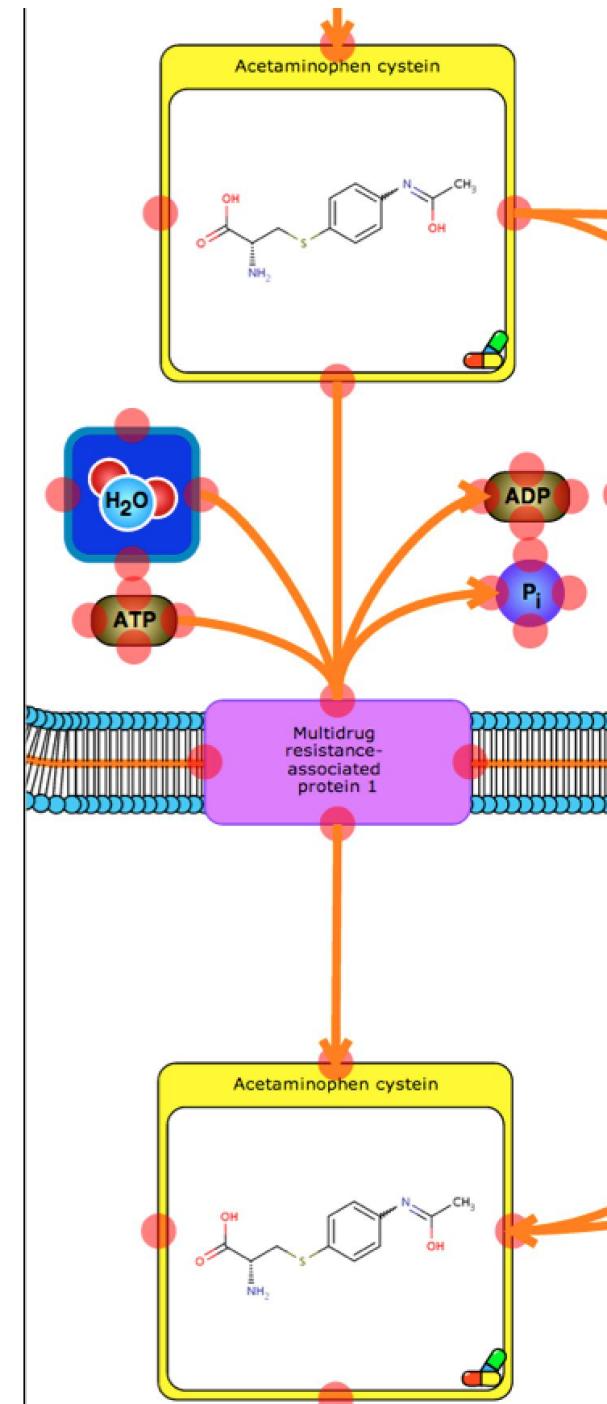
Options to customize components



In the pathway drawing view, components can be rendered and joined together to illustrate different biological processes.



Component visualizations include generic and specialized compounds, drugs, nucleic acids, enzymes, receptors, transporters, organs, organelles, membranes, and interaction arrow types.



PathWhiz

- There are Youtube videos and a JOVE article about making pathways with PathWhiz
- PathWhiz supports pathway propagation (mapping many pathways to other organisms) and replication (making multiple copies of highly similar pathways)
- You can make your own custom organism pathways (hundreds at a time) by “propagating” an existing organism’s pathway (Arabidopsis) to a genome/proteome of interest (Grape) – contact the Wishart Lab for help

PathBank

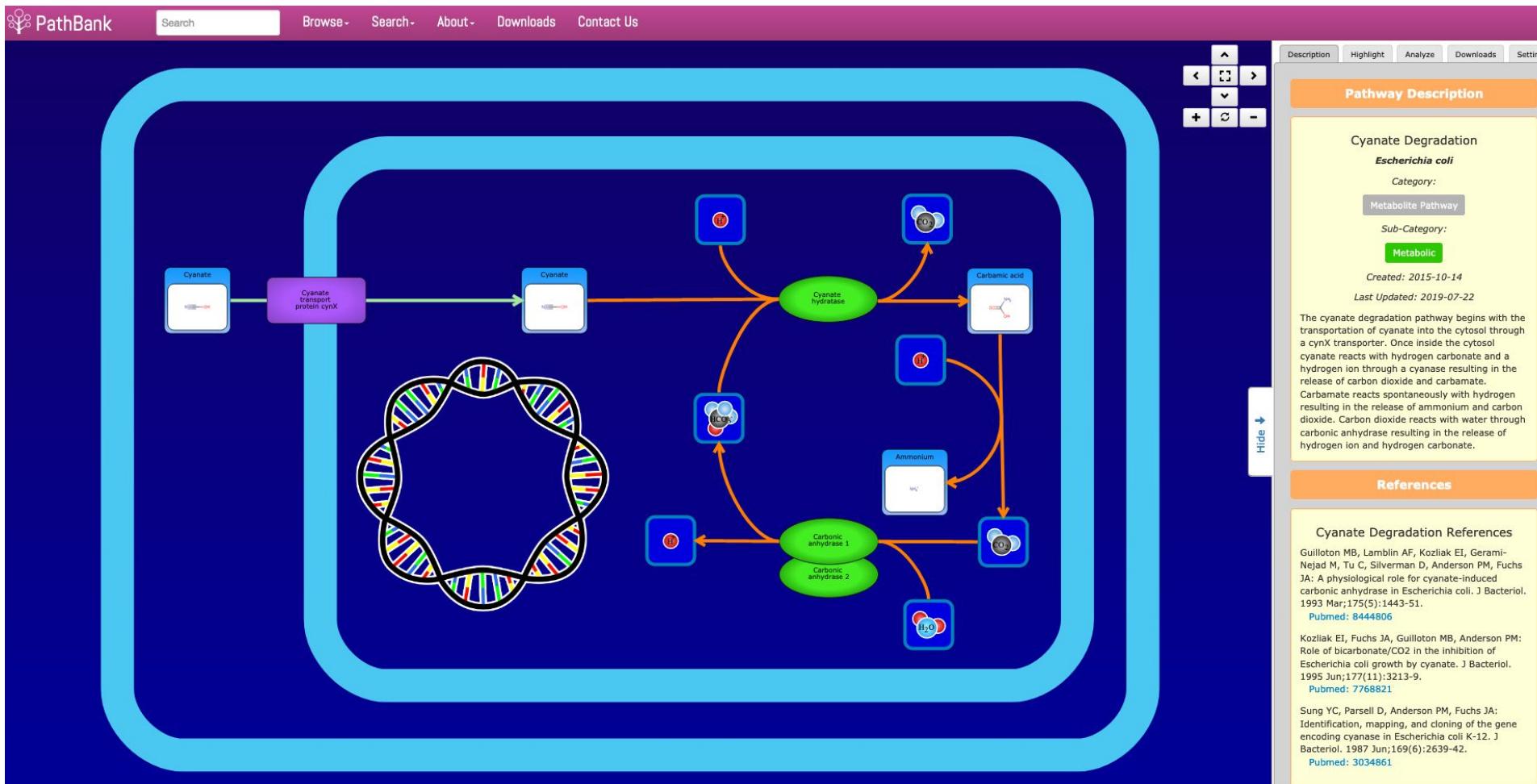
The screenshot shows the homepage of PathBank at pathbank.org. The header features a purple navigation bar with links for Search, Browse, About, Downloads, and Contact Us. A logo for TMIC (The Metabolomics Innovation Centre) is present, along with a banner stating "Specializing in ready to use metabolomics kits." The main content area has a yellow background. It includes a "Welcome to Path Bank" message with three stylized tree icons. A circular badge on the right indicates "Version 1.0". Below this, a descriptive paragraph explains the database's purpose and features. A call-to-action button says "Get started now: ★ Browse Pathways ★". At the bottom, there is a large, detailed diagram of a metabolic pathway involving Pyridoxine-5'-phosphate oxidase, Flavin Mononucleotide, Zinc, and Pyridoxal 5'-phosphate.

<http://pathbank.org/>

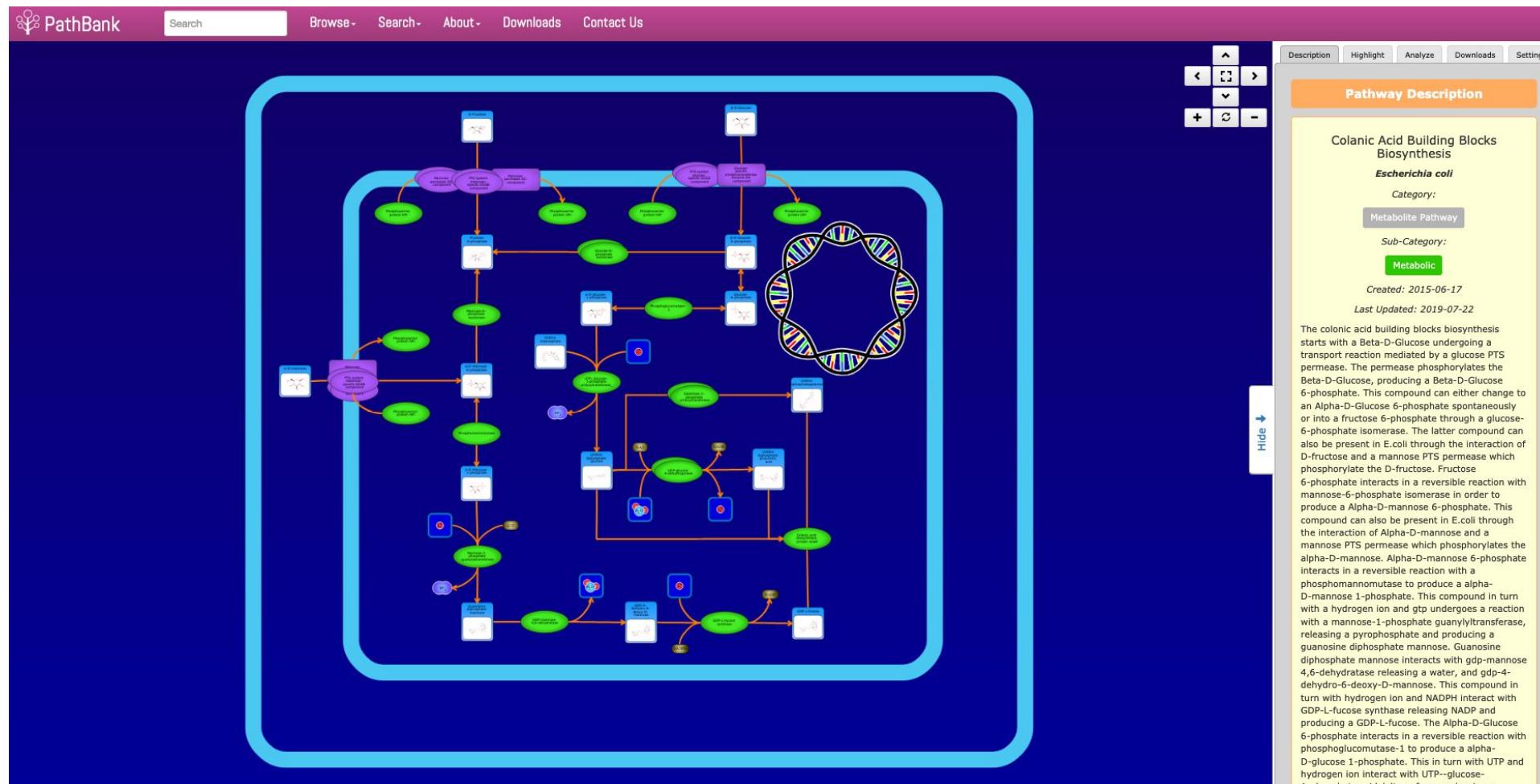
PathBank

- Extension of SMPDB (Human pathways) to 8 other model organisms including yeast, rat, mouse, cow, E. coli, Drosophila, C. elegans, and Arabidopsis
- Now has 109,000 pathways
- Covers disease, drug, metabolite, metabolite signaling pathways and protein signaling pathways

Microbial Cyanate Degradation



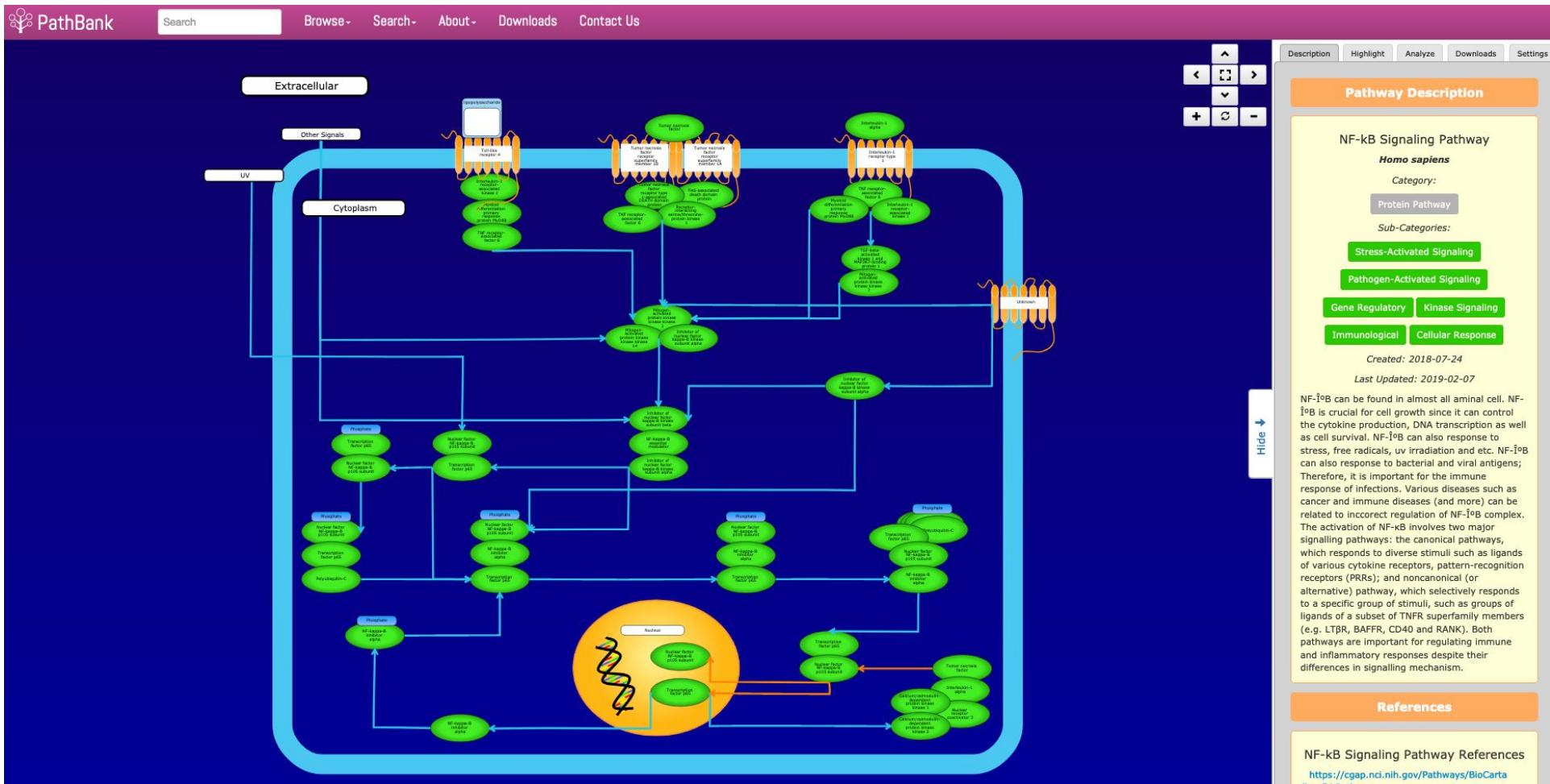
Microbial Colanic Acid Biosynthesis



Microbial Starch and Sucrose Metabolism



Human NF-κB Signaling Pathway (Protein Signaling Pathway)



Conclusions

- Databases – including organism-specific, purpose-specific and pathway databases are key to enabling the biological interpretation of metabolomic data
- Many well-known pathway databases were not developed for metabolomics applications
- The metabolomics community needs its own databases that are designed to better meet its specific and extensive demands
- Some of these "specialized" metabolomics databases have been presented here
- They are ideal for performing biological interpretation