

# CMAUP Database User Manual

(Collective Molecular Activities of Useful Plants)

June 2018

## Table of Content

### FUNCTIONS OF CMAUP INTERFACE

1. BROWSE THE DATABASE.....	2
1.1. Browse CMAUP by plant family.....	2
1.2. Browse CMAUP by plant usage class.....	3
1.3. Browse CMAUP by target genes.....	3
1.4. Browse CMAUP by target gene ontology.....	4
1.5. Browse CMAUP by target KEGG pathway.....	4
1.6. Browse CMAUP by related human disease.....	5
2. SEARCH BY KEYWORDS.....	5
3. ACCESS PLANTS BY WORLD MAP.....	6
4. DOWNLOAD DATA.....	6

### DESCRIPTION OF WEBPAGES

5. WEBPAGE OF A SPECIFIC PLANT.....	7
6. COLLECTIVE MOLECULAR ACTIVITIES AT 10uM ACTIVITY CUTOFF.....	10
7. WEBPAGE OF PLANT'S CHEMICAL CONSTITUENTS.....	10
8. WEBPAGE OF PLANT'S BIOLOGICAL ACTIVITY LANDSCAPE.....	12

# 1. BROWSE THE DATABASE

The homepage of CMAUP database can be accessed at <http://bidd2.nus.edu.sg/CMAUP/>. By clicking the image under the “Browse CMAUP Database”, users will be redirected to a new page where one can browse the database by multiple ways.

**Keyword Search**

Search Plant By: plant name, target, disease, GO terms, or KEGG pathways .....

Try Examples: [Hypericum perforatum](#), [Foeniculum vulgare](#), [GSK3B](#), [ALK](#), [Alcoholic hepatitis](#), [Arteriosclerosis](#), [GO:0006286](#), [hsa04620](#)

**Browse CMAUP Database**

- Medicinal Plant
- Agricultural Plant
- Human Edible Plant
- Garden Plant
- By Plant Usage Class
- By Target Genes
- By Target Genes Affiliated Gene Ontology
- By Target Genes Affiliated KEGG Pathway
- By Target Genes Related Disease

**Access Plants by Geographical Location**

Access Plants from Specific Country/Region by Clicking the World Map

## 1.1 Browse CMAUP by plant family

The current version of CMAUP includes 5754 plant species which from 345 taxonomic families. Users can access plants of a specific family by click the family name in the first column of the table.

**Browse Plants**

By Plant Family | By Plant Usage Class | By Target Genes | By Target Gene Ontology | By Target KEGG Pathway | By Related Disease

**Access Plants in Each Plant Family**

Show 10 entries

Search:

Plant Family	# Plant Genus	# Plants	# Medicinal Plants	# Human Edible Plants	# Agricultural Plants	# Garden Plants
<a href="#">Acanthaceae</a>	16	30	18	3	0	0
<a href="#">Aceraceae</a>	2	2	0	0	0	0
<a href="#">Achariaceae</a>	3	5	1	0	0	0
<a href="#">Acoraceae</a>	1	3	3	2	0	0
<a href="#">Actinidiaceae</a>	1	6	5	3	0	0
<a href="#">Adelanthaceae</a>	1	1	0	0	0	0
<a href="#">Adoxaceae</a>	2	15	7	4	0	0
<a href="#">Alzooaceae</a>	5	7	6	0	1	0
<a href="#">Alismataceae</a>	3	4	3	2	0	0
<a href="#">Alstroemeriaceae</a>	1	1	1	0	0	0

Showing 1 to 10 of 345 entries

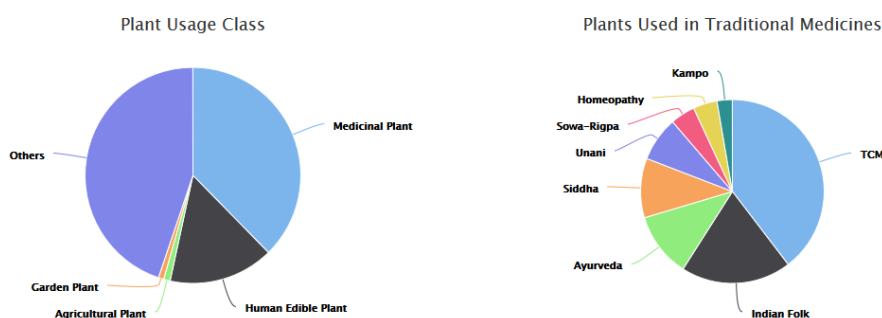
Previous 1 2 3 4 5 ... 35 Next

## 1.2 Browse CMAUP by plant usage class

Generally, plants are categorized into several usage classes: Medicinal Used Plant; Human Edible Plant; Agricultural Plant; Garden Plant; and Others. Moreover, the medicinal used plants are further labelled with traditional medicine systems that uses the plant as medicinal herb. These traditional medicine systems include: TCM (Traditional Chinese Medicine); Indian Folk; Ayurveda; Siddha; Unani; Sowa-Rigpa; Homeopathy; and Kampo. Users can access plants of each category by clicking the pie charts.



Access Plants in Each Plant Usage Class (Click the Pie Charts)



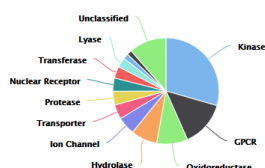
## 1.3 Browse CMAUP by target genes

Targets of plants are classified into few categories: Kinase; Oxidoreductase; Protease; GPCR; Ion Channel; Transporter; Nuclear Receptor; Unclassified and so on. Users can re-order the targets by each table columns such as “target Category”, then access plants by a specific target (click the target gene symbol).



Access Plants That Show Activities on Each Target

Number of Targets in Each Target Class



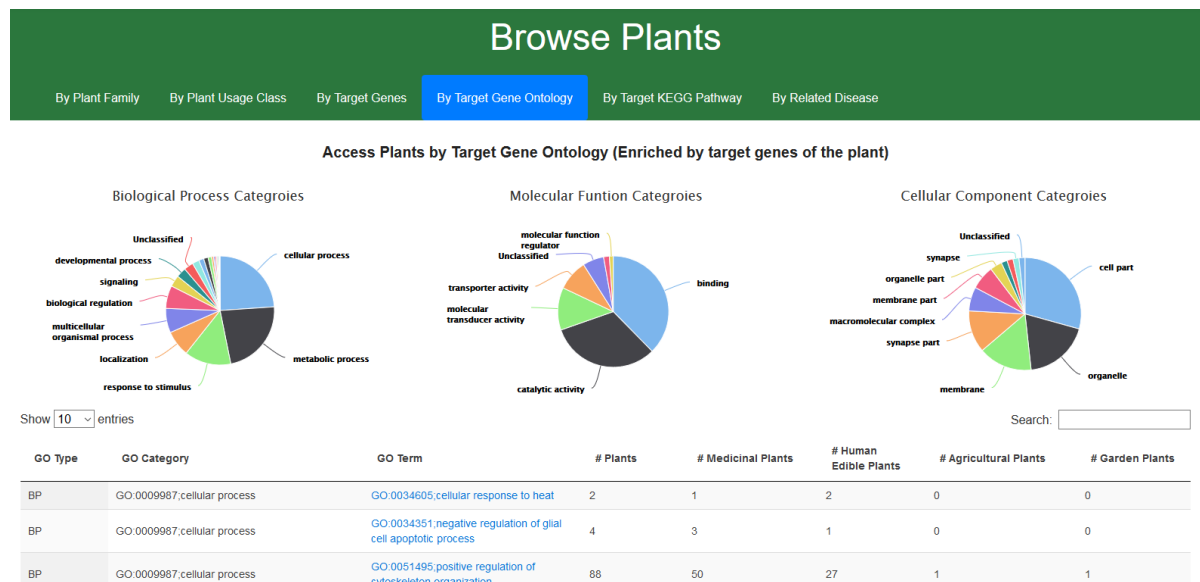
Show  entries

Search:

Target Category	Target Gene Symbol	Target Name	# Plants	# Medicinal Plants	# Human Edible Plants	# Agricultural Plants	# Garden Plants
G Protein-Coupled Receptor	<a href="#">SMO</a>	Smoothened homolog	9	2	1	0	0
G Protein-Coupled Receptor	<a href="#">CCR6</a>	C-C chemokine receptor type 6	1	1	0	0	0
G Protein-Coupled Receptor	<a href="#">GRM1</a>	Metabotropic glutamate receptor 1	30	16	12	3	0
G Protein-Coupled Receptor	<a href="#">CNR2</a>	Cannabinoid CB2 receptor	183	72	31	2	1
G Protein-Coupled Receptor	<a href="#">NPSR1</a>	Neuropeptide S receptor	1418	642	285	19	13
G Protein-Coupled Receptor	<a href="#">OPRM1</a>	Mu opioid receptor	21	8	4	0	1
G Protein-Coupled Receptor	<a href="#">GPR35</a>	G-protein coupled receptor 35	329	145	68	3	2

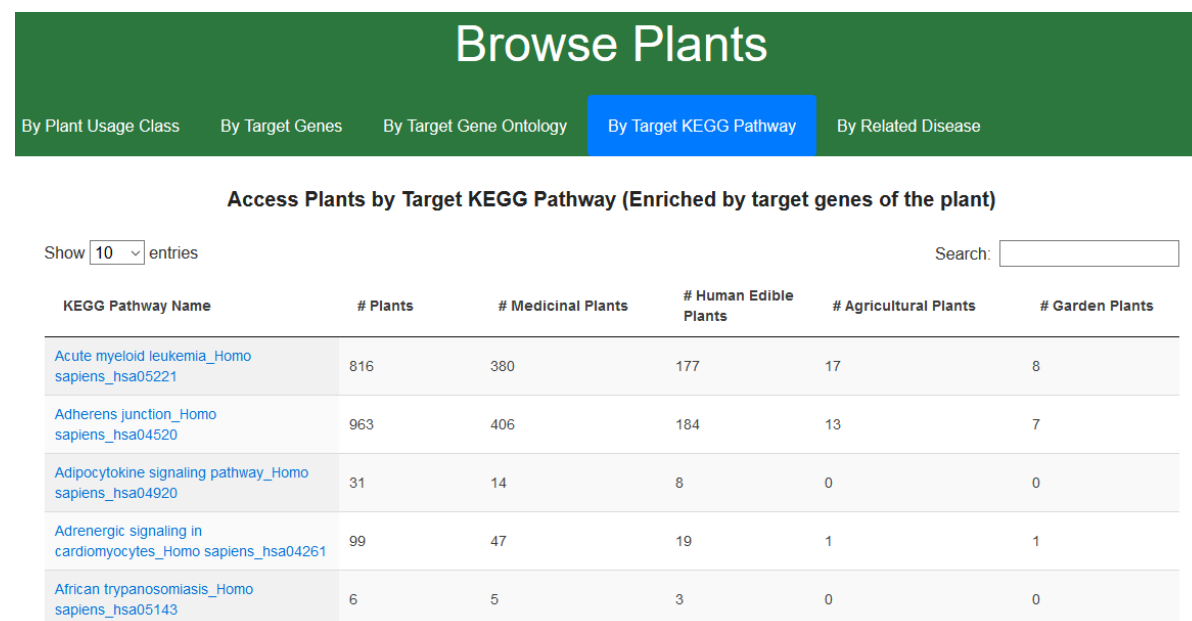
## 1.4 Browse CMAUP by target gene ontology

Gene ontology (GO) terms were enriched for genes of protein targets of each plant. GO terms are categorized by the categories of GO hierarchical system. Users can re-order the GO terms by their categories and access plants of a specific GO term by clicking the “GO Term”.



## 1.5 Browse CMAUP by target KEGG pathway

Genes of protein targets of plants were also enriched for KEGG pathways. Users can access plants of a specific pathway by clicking the pathway name.



## 1.6 Browse CMAUP by related human disease

Plants are linked to human diseases via the targets. Diseases are categorized by the ICD-10 system. Users can access plants of a related human disease by clicking the disease name.

### Browse Plants

Plant Usage Class   By Target Genes   By Target Gene Ontology   By Target KEGG Pathway   **By Related Disease**

Access Plants by Related Human Diseases (Linked by target genes of the plant)

Show  entries Search:

Disease Category	Disease	# Plants	# Medicinal Plants	# Human Edible Plants	# Agricultural Plants	# Garden Plants
A00-B99: Certain infectious and parasitic diseases	<a href="#">Cryptosporidium infection</a>	402	201	94	4	3
A00-B99: Certain infectious and parasitic diseases	<a href="#">Dengue fever</a>	13	3	3	0	0
A00-B99: Certain infectious and parasitic diseases	<a href="#">Diarrhea</a>	28	8	4	0	1
A00-B99: Certain infectious and parasitic diseases	<a href="#">Dutch elm disease</a>	33	19	8	0	0
A00-B99: Certain infectious and parasitic diseases	<a href="#">Adult varicella zoster virus infection</a>	83	39	18	2	0

## 2 SEARCH BY KEYWORDS

Users can search plants by multiple type of keywords: plant name, target gene symbol/protein name, disease, GO term ID/name, and KEGG pathway.

### Keyword Search

Search Plant By: plant name, target, disease, GO terms, or KEGG pathways ..... 

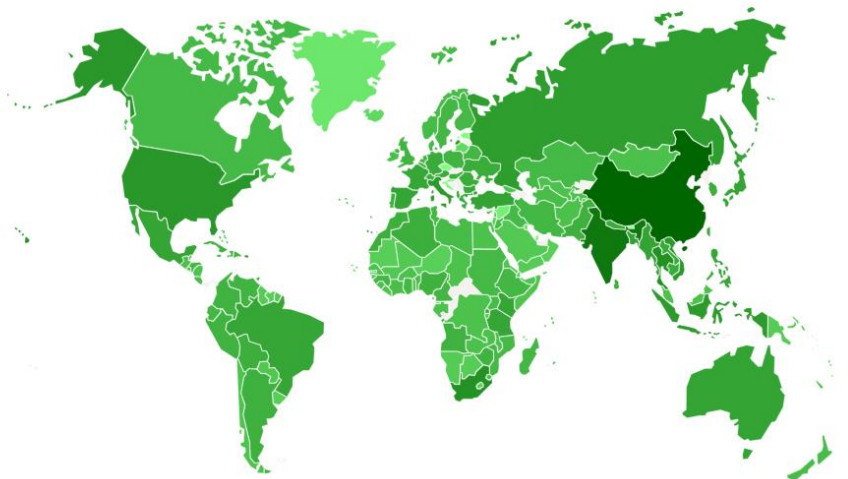
Try Examples: [Hypericum perforatum](#), [Foeniculum vulgare](#), [GSK3B](#), [ALK](#), [Alcoholic hepatitis](#), [Arteriosclerosis](#), [GO:0006286](#), [hsa04620](#)

### 3 ACCESS PLANTS BY WORLD MAP

Users can click country/region on the world map to access related plants.






#### Access Plants by Geographical Location

Access Plants from Specific Country/Region by Clicking the World Map 







### 4 DOWNLOAD DATA

All data can be freely downloaded at each individual webpage.


 Human Target Genes    Target Gene Ontology    Target KEGG Pathway    Human Disease    Download Data

Data Section	Select to Download
1. General Info & Plant Usage & Geographical Distribution	<input type="checkbox"/>
2. Human Target Genes	<input type="checkbox"/>
3. Target Gene Ontology	<input type="checkbox"/>
4. Target KEGG Pathway	<input type="checkbox"/>
5. Related Human Diseases	<input type="checkbox"/>



 Compound Cards    Physico-Chemical Properties    Download Data

Data Section	Select to Download
1. General information & structure data (InChI, InChIKey, SMILES)	<input type="checkbox"/>
2. Physico-chemical Properties	<input type="checkbox"/>



Download All Activity Values of Compounds of the Plant



## 5 WEBPAGE OF A SPECIFIC PLANT

A typical example of plant page is shown below:

**Plant Name: *Hypericum perforatum***

**Taxonomic Information:**  
 Herb ID: NPO4391  
 Herb Latin Name: *Hypericum perforatum*  
 NCBI Taxonomy ID: 65561  
 Taxonomy Genus: Hypericum  
 Taxonomy Family: Hypericaceae


**Used in Medicines:**  
**Country/Region:**  
 Turkey, Australia, Italy, Turkmenistan, South Africa; India, Lithuania; United States; Albania; Germany; Chile; Greece; Lebanon; Argentina; China; New Zealand; Spain

**Traditional Medicine System:**  
 TCM; Homeopathy; Siddha

**Medicinal Functions:**  
 Analgesic; Antidepressant; Antiseptic; Antispasmodic; Aromatic; Astringent; Cholagogue; Digestive; Diuretic; Expectorant; Homeopathy; Nervine; Resolvent; Sedative; Stimulant; Vermifuge; Vulnerary


**Geographical Distribution:**  
 Turkey, Australia, Italy, Turkmenistan, South Africa; Lebanon; India; United States; Albania; Germany; Spain, Chile; Greece; China; Argentina; New Zealand; Lithuania

**Example Image**



**Geographical Distribution**

■ Medicinal Used ■ Other Distributed Region



Part 1. Plant General Information

## Part 2. Plant Collective Molecular Activities

Human Target Genes
Target Gene Ontology
Target KEGG Pathway
Human Disease
Download Data

**List of Genes Collectively Targeted by the Plant**

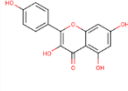
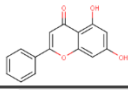
Target Activity Range: no more than 1uM  
Click Gene ID to Access Activity Profile of Individual Target

Target Class	Target Genes
G Protein-Coupled Receptor	TSHR, OPRD1, CRHR1, DRD3, NPSR1;
Hydrolase	RECQL, ALPL, ACHE;
Kinase	MET, AXL, CDK6, FLT3, CDK1, PIM1, KDR, IGF1R, AURKB;
Lyase	CA2, CA12, CA14, CA7, CA4;
Nuclear Receptor	NR1H4, NR1I2;
Oxidoreductase	TYR, CYP2C19, CYP19A1, ALOX15, CYP1A2, SBR1, CYP3A4, HSD17B2, HSD17B10, MAOA, CYP1B1, NOX4, CYP1A1;
Protease	MMP9, MMP1, MMP2;
Transcription Factor	AHR, TP53, NFKB1;
Transporter	SLCO1B1;
Unclassified	LMNA, XDH, APEX1, POLB, SMN1, SMN2;

**Detailed Information of Human Protein Targets**

Plant: *Hypericum perforatum*  
 Target: Cytochrome P450 1A2 [Gene ID: CYP1A2]

Show 10 entries

Compound ID	Compound Structure	No. of Activity Values	Activity Value (nM)	Activity Type	Uniprot ID
IP0116775		1	716	IC50	P35462
IP078640		2	Median: 69 Range: 54 to 84	IC50	Q6W5P4

Kinase KDR Vascular endothelial growth factor receptor 2 P35968  
 Kinase AXL Tyrosine-protein kinase receptor UFO P30530

>> Browse chemical constituents, all targets and activity landscape

Chemical  
Constituents

All Targets &  
Activity Landscape

>> Browse collective molecular activities at 10uM activity cutoff

Collective Molecular Activities at 10µM

## Part 3. Other Information

- Chemical Constituents
- All targets & Activity Landscape
- Collective Molecular Activities at 10uM cutoff

### Part 1. Plant General Information Section:

In this section, the taxonomy information (species, genus, and family names) was provided, a link to NCBI Taxonomy Database was provided if available. Moreover, countries/regions or traditional medicine system where use the plant as medicinal herbs are listed if available. Geographic distribution of the plant is provided along with a highlighted world map. An example image of the plant will be displayed when available.

## Part 2. Collective Molecular Activities Section:

This section includes “Human Target Genes”, “Target Gene Ontology”, “Target KEGG Pathway”, “Human Disease”, and “Download Data” subsections.

### A. Human Target Genes

**Collective Molecular Activities of the Plant**

Human Target Genes
Target Gene Ontology
Target KEGG Pathway
Human Disease
Download Data

**List of Genes Collectively Targeted by the Plant**  
Target Activity Range: no more than 1uM  
Click Gene ID to Access Activity Profile of Individual Target

Target Class	Target Genes
G Protein-Coupled Receptor	TSHR, OPRD1, CRHR1, DRD3, NPSR1,
Hydrolase	RECQL, ALPL, ACHE,
Kinase	MET, AXL, CDK6, FLT3, CDK1, PIM1, KDR, IGF1R, AURKB,
Lyase	CA2, CA12, CA14, CA7, CA4,
Nuclear Receptor	NR1H4, NR1I2,
Oxidoreductase	TYR, CYP2C19, CYP19A1, ALOX12, CYP1A2, CBR1, CYP3A4, HSD17B2, HSD17B10, MAOA, CYP1B1, NOX4, CYP1A1,
Protease	MMP9, MMP1, MMP2,
Transcription Factor	AHR, TP53, NFKB1,
Transporter	SLCO1B1,
Unclassified	LMNA, XDH, APEX1, POLB, SMN1, SMN2,

**Detailed Information of Human Protein Targets**

Show 10 entries Search:

Protein Class	Gene ID	Protein Name	Uniprot ID
G Protein-Coupled Receptor	DRD3	Dopamine D3 receptor	P35482
G Protein-Coupled Receptor	NPSR1	Neuropeptide S receptor	Q6W5P4
G Protein-Coupled Receptor	CRHR1	Corticotropin releasing factor receptor 1	P34998
G Protein-Coupled Receptor	OPRD1	Delta opioid receptor	P41143
G Protein-Coupled Receptor	TSHR	Thyroid stimulating hormone receptor	P16473
Hydrolase	RECQL	ATP-dependent DNA helicase Q1	P46963
Hydrolase	ALPL	Alkaline phosphatase, tissue-nonspecific isozyme	P05186
Hydrolase	ACHE	Acetylcholinesterase	P22303
Kinase	KDR	Vascular endothelial growth factor receptor 2	P35968
Kinase	AXL	Tyrosine-protein kinase receptor UFO	P30530

Showing 1 to 10 of 49 entries Previous 1 2 3 4 5 Next

In this subsection, all targets are displayed in a grid graph with the highlights of target category. On the right side, the detailed information of targets is presented in an interactive table.

### B. Target Gene Ontology

Human Target Genes
Target Gene Ontology
Target KEGG Pathway
Human Disease
Download Data

**Enrichment of Gene Ontology of Human Protein Targets (Activity<=1uM)**

**Biological Process Categories**

- GO:0010896 response to stimulus
- GO:0008152 metabolic process
- GO:0050987 cellular process
- GO:0032502 developmental process
- GO:0051179 localization
- GO:0062866 base-excision repair: base-free sugar-phosphate removal
- GO:0062854 positive regulation of cell proliferation

**Molecular Function Categories**

- GO:0006089 molecular transducer activity
- GO:0003284 catalytic activity
- GO:0005448 binding

**Cellular Component Categories**

**Detailed Information of Enriched Gene Ontology**

Show 10 entries Search:

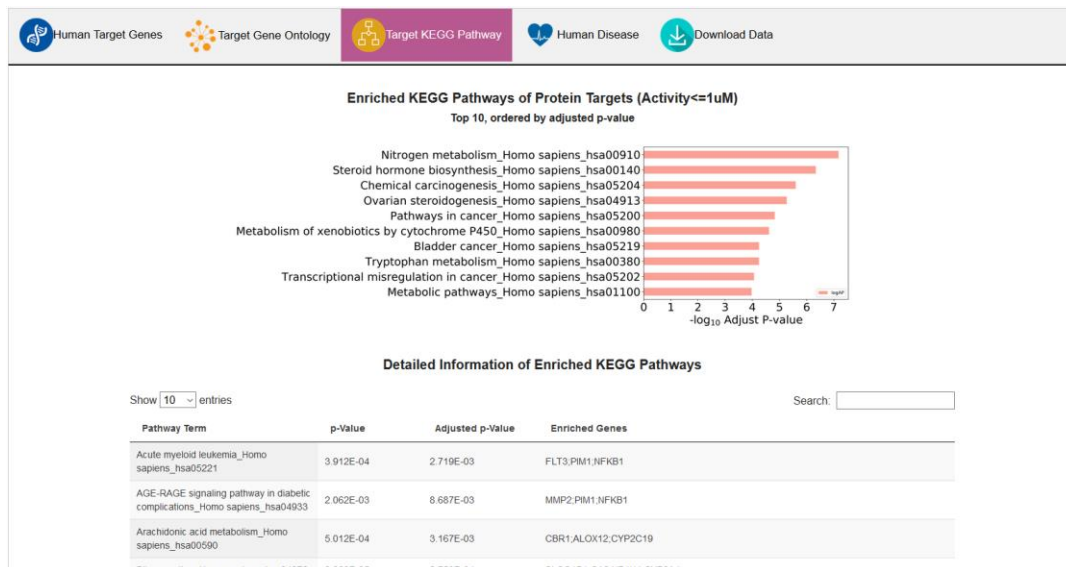
GO Type	GO Category	Enriched GO Terms	p-Value	Adjusted p-Value	Enriched Genes
BP	GO:0008152:metabolic process	GO:0097267:omega-hydroxylase P450 pathway	8.246E-09	7.481E-06	CYP1A1, CYP1A2, CYP1B1, CYP2C19
BP	GO:0050986:response to stimulus	GO:0042738:exogenous drug catabolic process	1.936E-08	1.448E-05	CYP1A2, CYP2C19, CYP3A4, NR1I2
BP	GO:0051179:localization	GO:0015701:bicarbonate transport	9.524E-08	5.408E-05	CA12, CA14, CA2, CA4, CA7
BP	GO:0008152:metabolic process	GO:0019373:epoxygenase P450 pathway	2.301E-07	1.044E-04	CYP1A1, CYP1A2, CYP1B1, CYP2C19
BP	GO:0008152:metabolic process	GO:0016098:monoterpenoid metabolic process	3.246E-07	1.334E-04	CYP1A2, CYP2C19, CYP3A4
BP	GO:0009987:cellular process	GO:0043066:negative regulation of apoptotic process	1.377E-06	4.003E-04	ALOX12, AURKB, AXL, CDK1, IGF1R, KDR, LMNA, MMP9, NFKB1, NR1H4, PIM1, TP53
BP	GO:0032502:developmental process	GO:0007566:aging	2.060E-06	5.608E-04	APEX1, AURKB, CDK1, CYP1A1, NOX4, POLB, TP53
BP	GO:0008152:metabolic process	GO:0006730:one-carbon metabolic process	2.201E-06	5.779E-04	CA12, CA2, CA4, CA7
BP	GO:0009987:cellular process	GO:0006286:base-excision repair: base-free sugar-phosphate removal	6.567E-06	1.571E-03	APEX1, POLB
BP	GO:0006285:cell proliferation	GO:0006284:positive regulation of cell proliferation	1.297E-05	2.742E-03	ALOX12, CDK1, CDK6, DRD3, FLT3, IGF1R, KDR, MMP2, MMP9, PIM1, TSHR

Showing 1 to 10 of 34 entries Previous 1 2 3 4 Next

In this subsection, enriched GO terms (biological process, molecular function, and cellular component) are grouped by GO categories. Detailed data (GO term, p-value, adjusted p-value, enriched genes) was presented in an interactive table.

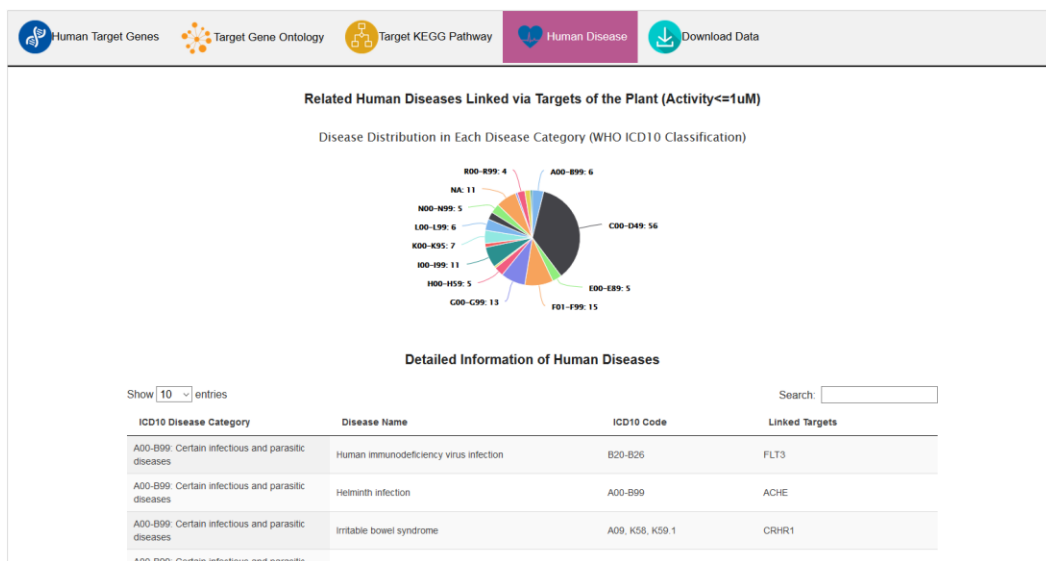


### C. Target KEGG Pathway



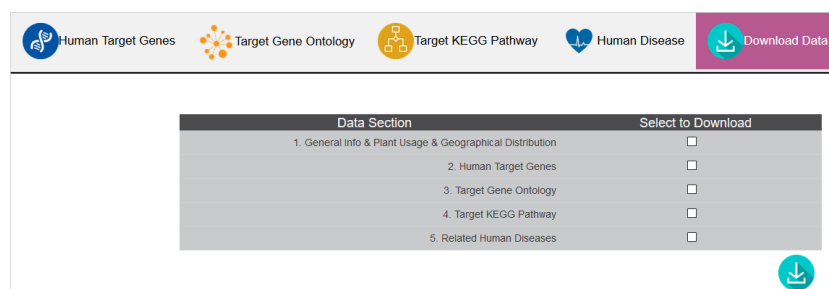
In this subsection, top 10 enriched KEGG pathways were presented in a bar chart, detailed information was provided in the below table.

### D. Human Disease



In this subsection, plant related human diseases were categorized based on ICD-10 code system of the World Health Organization.

### E. Download Data



## 6 COLLECTIVE MOLECULAR ACTIVITIES AT 10uM ACTIVITY CUTOFF

Default activity cut-off is 1uM. Besides, collective molecular activities of plant at activity cut-off of 10uM are also provided in another webpage.

Plant Name: *Hypericum perforatum*

### Collective Molecular Activities at 10uM activity Cutoff

Human Target Genes Target Gene Ontology Target KEGG Pathway Human Disease Download Data

#### List of Genes Collectively Targeted by the Plant

Target Activity Range: no more than 10uM  
Click Gene ID to Access Activity Profile of Individual Target

Target Class	Target Genes
G Protein-Coupled Receptor	OPRM1, OPRD1, CRHR1, TSHR, ADRA2C, ADRA2A, NPSR1, HTR2C, CXCR1, NMUR2, DRD3, DRD4;
Hydrolase	ALPL, RECQL; PLA2G1B, TDP1, BLM, PTPN1; ALPI, ACHE, CTDSP1, GAA;
Isomerase	P4HB;
Kinase	AKT1, MET, CAMK2B, PKM, CSNK2A1, DAPK1, NEK2, FLT3, PKN1, SRC, IGF1R, PTK2, AURKB, NEK5, CDK2, GSK3B, KDR, CDK1, MAPK1, PIM1, NUA1; ALK, AXIL, CDK5, PK3R1;
Lyase	CA12, CA14, CA3, CASB, CA1, CA7, CA4, CA6, CA2; GLO1; CA9, CA5A;
Nuclear Receptor	AR, THR, NR1H4, NR1H2, ESRRA; PPARA, PPAR, ESR1, ESR2;
Oxidoreductase	HPGD; MPO, HSD17B1; KDM4E, CBR1, AKR1B1; HSD17B2; HSD17B10, MAOA, ALOX15, AKR1B10, CYP1A2, ALOX5; CYP2C19, CYP2C9, CYP19A1, ALDH1A1, TYR, CYP2C8, ALOX12, CYP3A4, CYP1B1; NQO1, CYP1A1;
Protease	USP2, CASP7, MMP13, CASP1, MMP3, MMP2, MMP9; BACE1, MMP1;
Transcription Factor	AHR, STAT6, HIF1A; NFKB1, TP53;
Transferase	PYGL;
Transporter	VCP, ABCC1, ABCB1, ABCG2, SLC01B3, SLC01B1;
Unclassified	LMNA; APEX1, MCL1, MAPT, POLB, FASN, SMN1, SMN2, XBP1, XDH, THPO;

#### Detailed Information of Human Protein Targets

Show 10 entries Search:

Protein Class	Gene ID	Protein Name	Uniprot ID
G Protein-Coupled Receptor	DRD4	Dopamine D4 receptor	P21917
G Protein-Coupled Receptor	NPSR1	Neuropeptide S receptor	Q6W5P4
G Protein-Coupled Receptor	ADRA2C	Alpha-2c adrenergic receptor	P18825
G Protein-Coupled Receptor	HTR2C	Serotonin 2c (5-HT2c) receptor	P28335
G Protein-Coupled Receptor	DRD3	Dopamine D3 receptor	P35462
G Protein-Coupled Receptor	TSHR	Thyroid stimulating hormone receptor	P16473
G Protein-Coupled Receptor	OPRD1	Delta opioid receptor	P41143
G Protein-Coupled Receptor	CRHR1	Corticotropin releasing factor receptor 1	P34998
G Protein-Coupled Receptor	NMUR2	Neuromedin-U receptor 2	Q9GZQ4
G Protein-Coupled Receptor	ADRA2A	Alpha-2a adrenergic receptor	P08913

Showing 1 to 10 of 124 entries Previous 1 2 3 4 5 ... 13 Next

## 7 WEBPAGE OF PLANT'S CHEMICAL CONSTITUENTS


Chemical constituents of the plant were provided. Only those compounds which have IC50 or EC50 or Ki activity records were labelled as "Activity Available" compound part. Few cards were displayed to show the profile of several important drug-like properties.

Besides, the chemical structure images were displayed in "Compound Cards", physico-chemical properties (calculated by using PaDEL software) were provided in another subsection.

All these data can be downloaded in the "Download Data" subsection.

Compound Cards Physico-Chemical Properties Download Data

Data Section	Select to Download
1. General Information & structure data (InChi, InChiKey, SMILES)	<input type="checkbox"/>
2. Physico-chemical Properties	<input type="checkbox"/>



# Chemical Constituents of Plant: *Hypericum perforatum*

## 382 Total Compounds

Unique compounds have been isolated from this plant

## 83 Compounds Have Activities Available

Unique compounds have activity data available (only IC50/EC50/Ki activity types are included) on human protein targets

Fig 1. Molecular Weight

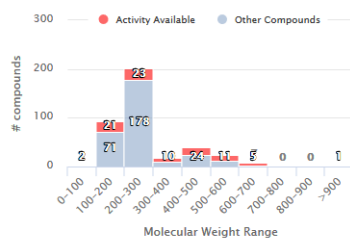


Fig 3. TPSA vs XlogP

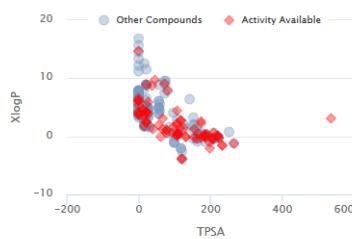


Fig 2. #HBA vs #HBD

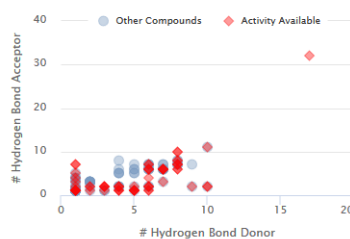
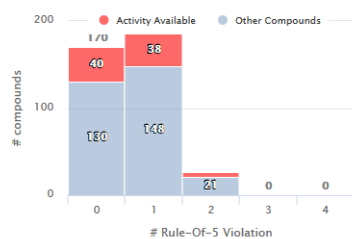
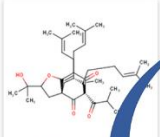


Fig 4. RO5 Violation



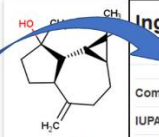
Ingredient Cards    Physico-Chemical Properties    Download Data

### Ingredient Structural Cards



Ingredient ID: NPC9965  
Formula: C35H52O5

[SEE DETAILS](#)



Ingredient ID: NPC99480  
Formula: C15H24O

[SEE DETAILS](#)

**Ingredient ID: NPC99657**

**Detailed Information & References of Ingredients**

**Common Name** Furohyperforin

**IUPAC Name** n.a.

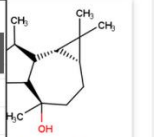
**Canonical SMILES** CC(=CCC[C@]1(C)[C@@H](CC=C(C)C)C[C@]23C(=O)[C@]1(C(=O)C(C)C)C2

**Standard InChI** InChI=1S/C35H52O5/c1-21(2)13-12-18-33(11)25(16-14-22(3)4)19-34-20-27/25-27,33+34-35-m/s1

**Standard InChIKey** SUOQGZCCNMGMYHT-VJRKKGXSA-N

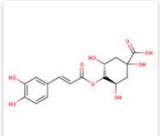
**Ingredient External Links**  
PubChem CID: 44427231  
ChEMBL ID: CHEMBL397752

**Plant-Ingredient Association Reference**  
Database: HerDing  
Database: TM-MC  
Database: TCMID

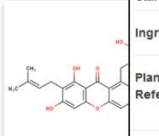


Ingredient ID: NPC9424  
Formula: C15H26O

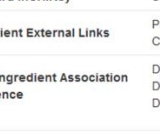
[SEE DETAILS](#)




Ingredient ID: NPC92774  
Formula: C16H18O9



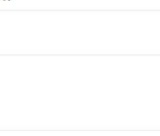
Ingredient ID: NPC92032  
Formula: C23H26O7



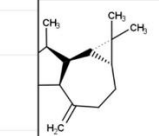
Ingredient ID: NPC91574  
Formula: C15H24



Ingredient ID: NPC90904  
Formula: C16H32O3



Ingredient ID: NPC90889  
Formula: C15H22O



Ingredient ID: NPC90803  
Formula: C15H24

Compound Cards    Physico-Chemical Properties    Download Data

### Properties

Compound ID	Formulae	MW	AlogP	MlogP	XlogP	# HBA	# HBD	PSA	# Rotatable Bond	# Ring	# Heacy Atom	Lipinski RO5 Violation
NPC126029	C15H14O6	290.08	-2.494	2.45	0.406	1	5	110.38	6	3	21	0
NPC138883	C30H44O4	468.3239599	4.5362	4.32	7.382	4	1	71.44	9	2	34	1
NPC157410	C15H14O6	290.0790382	-2.494	2.45	0.406	1	5	110.38	1	3	21	0
NPC171816	C24H26O6	410.1729386	0.1641	3.44	4.04	1	3	96.22	3	4	30	0
NPC173638	C21H20O11	447.0927364	-4.7259	2.56	0.349	6	6	189.2	3	4	32	1
NPC20790	C15H10O7	301.0348276	-3.135	2.34	1.381	2	4	130.28	1	3	22	0
NPC218109	C27H30O16	610.1533849	-6.139	2.67	-1.188	11	10	265.52	6	5	43	2
NPC227839	C30H16O8	502.0688674	-2.673	3.88	1.305	2	4	161.18	0	8	38	0
NPC228383	C19H18O11	422.08	-4.4948	2.34	-1.995	6	8	197.37	10	4	30	1
NPC235260	C21H18O13	478.07	-4.7899	2.34	-0.197	8	8	223.67	12	4	34	1
NPC275463	C21H20O12	463.087651	-5.2259	2.45	-0.354	7	7	209.43	4	4	33	1
NPC290194	C30H16O9	520.08	-3.0961	3.77	-0.248	3	7	175.75	9	8	39	1
NPC41845	C48H32O32	1120.09	-5.994	3.22	3.13	32	17	537.24	23	10	80	2
NPC99657	C35H52O5	552.38	4.6779	4.76	7.914	5	1	80.67	22	3	40	1

## 8 WEBPAGE OF PLANT'S BIOLOGICAL ACTIVITY LANDSCAPE

**Preprocessing of activity values:** when there are multiple activity records for one compound-target pair, multiple activities are processed as following:

**Step 1:** remove outliers

$$\text{Criterion: Median} - 3 * \text{MAD} \leq x \leq \text{Median} + 3 * \text{MAD}$$

'x' indicates each activity record, 'Median' is median of all activity records, 'MAD' is Median Absolute Deviation. (Reference: Leys, C., et al., *Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median, Journal of Experimental Social Psychology (2013), <http://dx.doi.org/10.1016/j.jesp.2013.03.013>*)

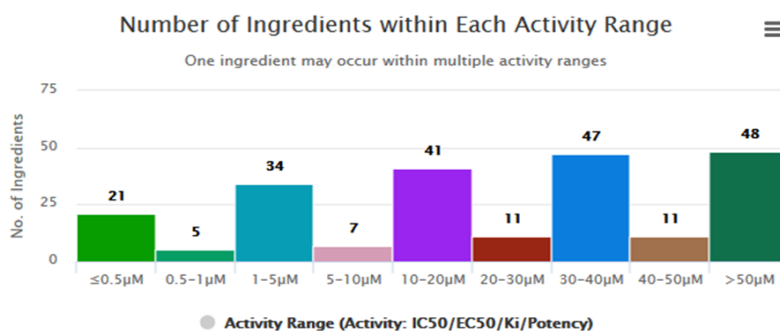
**Step 2:** calculate the median of qualified values

To show the activity landscape of all ingredients of a plant, lower activity value means stronger activity and coloured with the darker red in the heatmap. Clicking the heatmap cells can view detailed activity of certain ingredient against a specific target protein.

Moreover, users can download all the activity values in the "Download Data" subsection.

171

Target Proteins in  
Total



Landscape of Ingredients, Target Proteins and Activity Values

[Click individual heatmap cells to access detailed activity information](#)

