

# KEGG PATHWAYを中心としたKEGG、 GenomeNetの使い方、および メタボローム関連データベースの紹介

情報・システム研究機構  
ライフサイエンス統合データベースセンター

時松 敏明  
([tokimatu@dbcls.rois.ac.jp](mailto:tokimatu@dbcls.rois.ac.jp))

2014/8/20

統合データベース講習会 AJACS阿波 2014/8/20 14:00-15:30

1

## 内容

- ブラウザを使ったKEGG/GenomeNetの利用法
  - KEGGとゲノムネットの概要
  - KEGG, ゲノムネットの配列データ  
(GENES, ORTHOLOGY(KO), OC)
  - KEGG, ゲノムネットの化合物情報  
(LIGAND)
  - KEGGのシステム情報、ツール  
(PATHWAY, BRITE, MODULE, KEGG mapper)
- メタボローム関連データベースの紹介

2014/8/20

2

# KEGGとゲノムネットの概要

2014/8/20

3

## KEGG: Kyoto Encyclopedia of Genes and Genomes

日本語 : [http://www.kegg.jp/kegg/kegg\\_ja.html](http://www.kegg.jp/kegg/kegg_ja.html)  
英語 : <http://www.kegg.jp/>

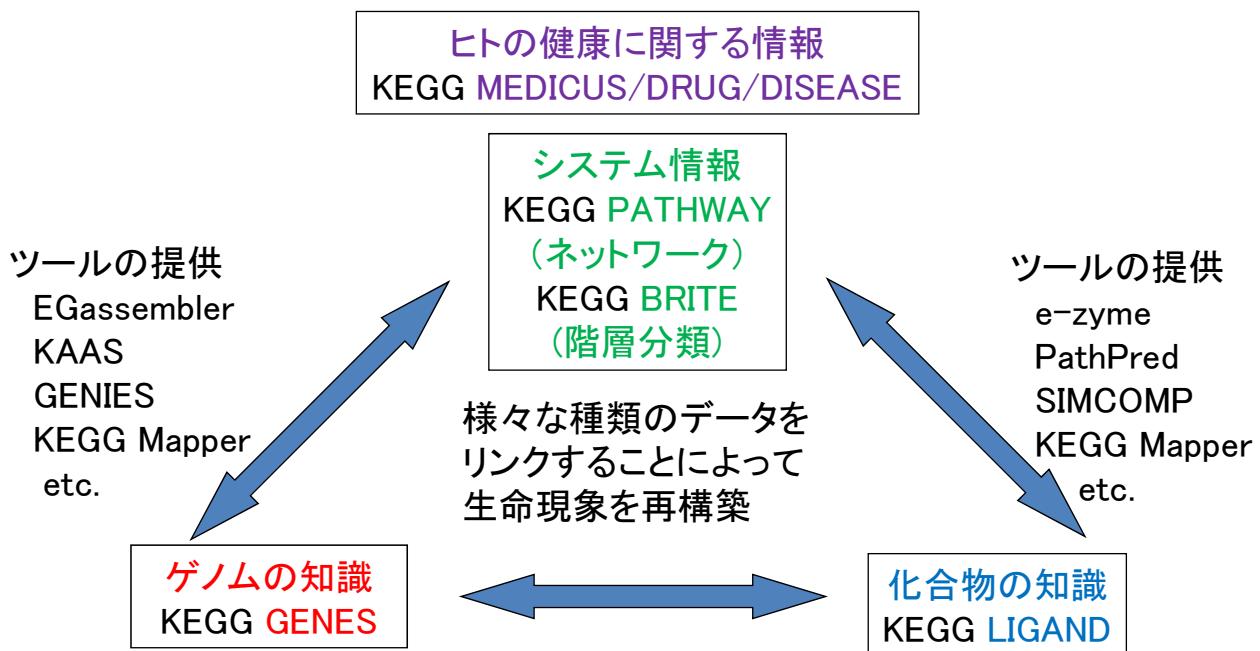
The screenshot shows the KEGG homepage with a navigation bar at the top. The main content area displays the KEGG logo and the title "KEGG: Kyoto Encyclopedia of Genes and Genomes". Below the title is a message about the KEGG database being a molecular-level information resource. The page is divided into several sections: "KEGG の主要エントリーポイント" (including KEGG2, KEGG PATHWAY, KEGG BRITE, KEGG MODULE, KEGG ORTHOLOGY, KEGG GENOME, KEGG GENES, KEGG COMPOUND, KEGG REACTION, KEGG DISEASE, KEGG DRUG, KEGG MEDICUS), "生物種ごとのエントリーポイント" (including KEGG Organisms, KEGG Mapper, KEGG Atlas, KAAS, BLAST/FASTA, SIMCOMP, PathPred), and "解析ツール" (including KEGG Pathway/Module Mapper, KEGG Genome Browser, KEGG Orthology, KEGG BLAST, KEGG SIMCOMP). A search bar is located at the top right, and a copyright notice at the bottom states "Copyright 1995-2014 Kanehisa Laboratories".

- ゲノムや分子レベルの情報から細胞、固体、エコシステムといった生命システムレベルの機能を理解するためのデータベースリソース
- 遺伝子や化合物などの分子部品のデータベースと、それらをつなぐネットワークのデータベースからなっている

2014/8/20

4

# KEGG: Kyoto Encyclopedia of Genes and Genomes



2014/8/20

5

## KEGG でカバーしているデータ、いないデータ

• いる

• いない

### ゲノム

遺伝子  
アノテーション

制御領域  
バリエーション  
タンパク質立体構造

### パスウェイ

代謝、制御マップ  
病気、薬開発マップ

シミュレーション用の  
パラメータ

### 化合物

2次元構造  
薬、代謝物、反応  
反応パターン

物性、立体構造  
速度定数

データ間のリンク

外部データベースへのリンク

2014/8/20

6

# KEGG Objects Identifier

2014/8/20

The screenshot shows the KEGG Objects Identifier page. On the left, there's a sidebar with links like KEGG Home, KEGG Database, KEGG Software, and KEGG FTP. A red box highlights the "KEGG Objects" section under "KEGG Database". Below it is a table of database entries:

Database	Object	Prefix	Example
KEGG PATHWAY	Pathway map	map, ko, ec, rn, (org)	hsa04930
KEGG BRITE	Functional hierarchy	br, jp, ko, (org)	ko01003
KEGG MODULE	KEGG module	M, (org)_M	M00010
KEGG ORTHOLOGY	KO group	K	K04527
KEGG GENOME	KEGG organism	T	T01001 (hsa)
KEGG GENES	Gene / protein		hsa:3643
KEGG COMPOUND	Small molecule	C	C00031
KEGG GLYCAN	Glycan	G	G00109
KEGG REACTION	Reaction	R	R00259
KEGG RPAIR	Reactant pair	RP	RP04458
KEGG RCLASS	Reaction class	RC	RC00046
KEGG ENZYME	Enzyme	ec	ec:2.7.10.1
KEGG DISEASE	Human disease	H	H00004
KEGG DRUG	Drug	D	D01441
KEGG DGROUP	Drug group	DG	DG00710
KEGG ENVIRON	Crude drug, etc.	E	E00048

(org) represents three-, four-, or five-letter organism code

The page also contains text explaining the identifier format: "The identifier of each database entry is generally in the form of db:entry where "db" is the database name and "entry" is the entry name or the accession number (see DBGET for the list of database names and abbreviations). However, "db" may be omitted in fourteen of the sixteen databases shown below, because the entry name, called the **KEGG object identifier** consisting of a database-dependent prefix and a five-digit number, is unique across the databases. The KEGG object identifiers are often called D numbers, K numbers, C numbers, etc. The D number example shown below D01441 is thus equivalent to dr:D01441 or drug:D01441."

基本は Prefix + 5桁の数字

GENESは、  
生物種コード：遺伝子コード

7

# GenomeNet

[http://www.genome.jp/ja/ \(日本語\)](http://www.genome.jp/ja/)

[http://www.genome.jp/ \(英語\)](http://www.genome.jp/)

The screenshot shows the GenomeNet homepage. It features a search bar at the top with options for English and Japanese. Below the search bar, there are several sections:

- 統合データベース**: Includes KEGG, varDB, and research databases.
- ゲノムネット**: Includes a search interface for various databases.
- ゲノムネット計算ツール**: Includes BLAST/FASTA, MOTIF, CLUSTALW/MAFFT/PRRN, and various assembly and annotation tools.
- データベース関連**: Includes a diagram titled "データベース間のリンク" (Links between databases) showing relationships between KEGG, BRITE, and other databases.

- ゲノム情報を基盤とした生命科学研究を促進するためのインターネットサービス

- KEGGを主幹とするデータベース群と、それらデータを解析するための計算ツール群からなっている

2014/8/20

8

# GenomeNet

GenomeNet

KEGG KEGG2 PATHWAY BRITE MEDICUS DBGET LinkDB [ English | Japanese ]

Search 統合データベース for Go Clear

## データベース検索システム: DBGET/LinkDB

### KEGGと国内外の様々なデータベースを統合的に検索するシステム

ゲノムネット  
ゲノムネットとお知らせ  
謝辞

統合データベース  
統合DBの概要  
DBGETの概要  
リリース情報

医薬品データベース  
KEGG  
varDB

研究支援データベース  
計算ツール  
その他のツール  
FTP

フィードバック

ゲノムネット統合データベース  
DBGET search REST サービス

KEGG: 生命システム情報統合データベース  
KEGG2 - 目次のページ  
KEGG PATHWAY - システム情報: パスウェイ  
KEGG BRITE - システム情報: オントロジー  
KEGG Organisms - 生物種ごとの入口  
KEGG GENES - ゲノム情報  
KEGG LIGAND - ケミカル情報  
KEGG MEDICUS - 疾患・医薬品情報

Reaction Ontology: 反応パターンと分類  
varDB: 抗原変異データベース  
研究支援データベース  
CYORF - シアノウクテリア  
BSORF - 枯草菌  
EXPRESSION - 遺伝子発現プロファイル

BRITE 機能階層 (日本語)  
KEGG バスウェイマップ  
KEGG 生物種

生物種ごとのゲノム一覧  
完全 (KEGG GENES)  
ドラフト (KEGG DGENES)  
メタ (KEGG MGENES)  
ウイルス (KEGG VGENES)  
EST (EGENES)

オーソログクラスタ (OC)

### 反応オントロジーデータベース

### その他のプロジェクトデータベース



2014/8/20

9

## LinkDBによるデータベース間のリンク

KEGGを核として、国内外の様々なDBとリンク

LinkDB: database of link information

DBGET LinkDB KEGG2

Search all links from in html format Go Clear

(Note) Specify a single entry or multiple entries in the form of dbentry, such as hsa126 hsa127

Links are categorized into the following 4 types.

original links are extracted from the database entries provided by the GenomeNet DBGET system.

reverse links are derived from the original links by exchanging a source entry and its target entry.

equivalent links are special original links to signify equivalent contents between KEGG GENES, COMPOUND, DRUG, REACTION databases and databases other than KEGG.

indirect links are derived by combining two or more original links. Currently, links from KEGG GENES to REACTION via KO, and to COMPOUND via REACTION are available.

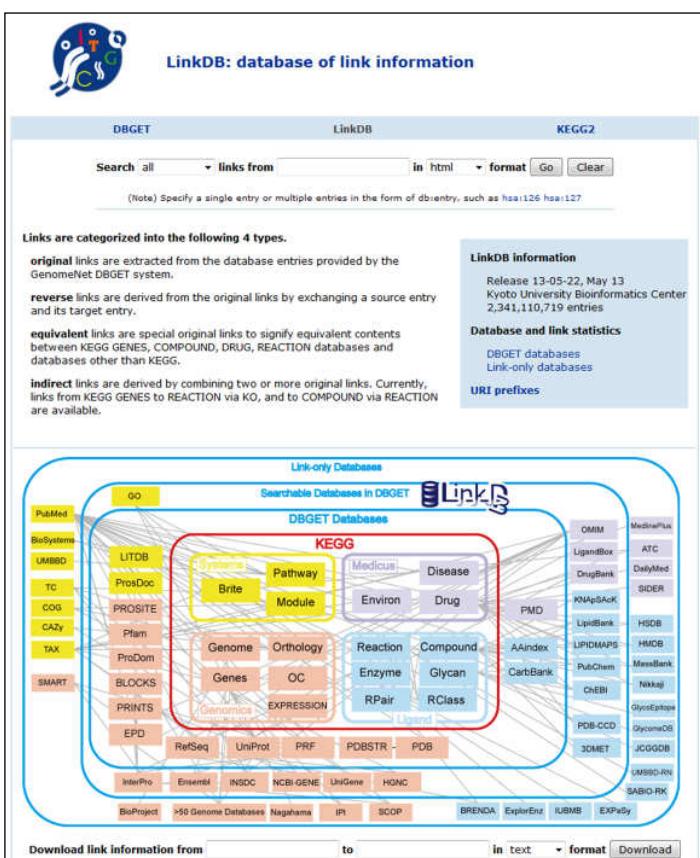
LinkDB information

Release 13-05-22, May 13  
Kyoto University Bioinformatics Center  
2,341,110,719 entries

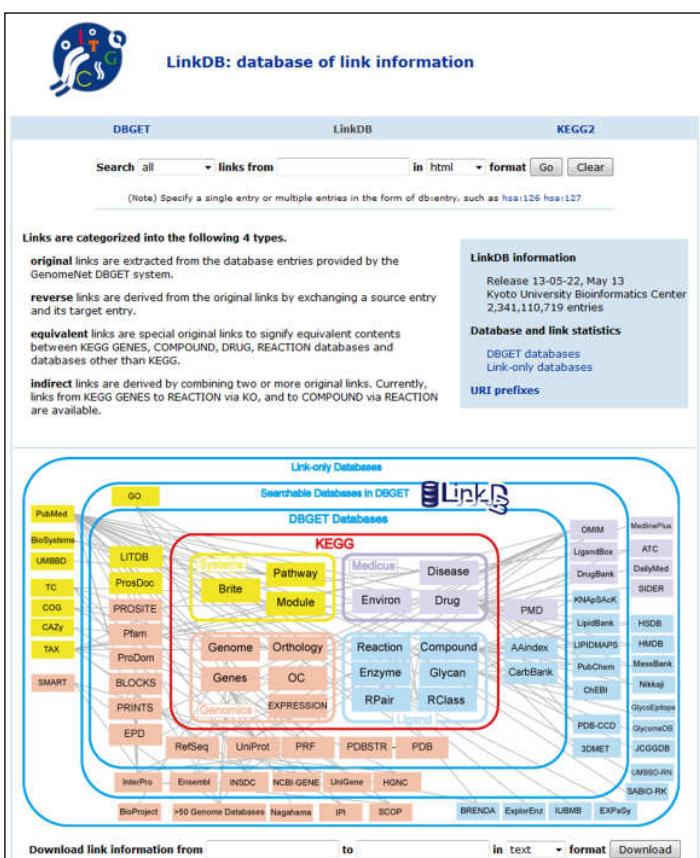
Database and link statistics

DBGET databases  
Link-only databases  
URI prefixes

LinkDB Database Map



Download link information from to in text format Download



2014/8/20

10

# ゲノムネット計算ツール

ゲノムネット計算ツール

**配列解析**

BLAST / FASTA - ホモロジー検索  
MOTIF - モチーフ検索  
CLUSTALW / MAFFT / PRRN - 複数アライメント

**ゲノム情報解析**

OC Viewer - オンコログクラスタ *Updated!*  
REST サービス  
KAAS - KEGG自動アノテーションサーバー  
MAPLE - 代謝・生理機能評価サーバー  
EGassembler - ESTコンティグ作成  
GENIES - 遺伝子ネットワーク予測  
DINIES - 医薬品-標的ネットワーク予測

**ケミカル情報解析**

SIMCOMP / SUBCOMP - 化合物構造検索  
REST サービス  
KCalM - 糖鎖構造検索  
PathComp - 可能な反応経路の計算  
PathSearch - 類似反応経路検索  
PathPred - 分解・合成反応経路予測  
E-zyme - 化合物間の酵素反応予測 *Updated!*



**ホモロジー検索、モチーフ検索  
マルチプルアライメント**



**遺伝子機能アノテーション、予測など**



**データベース増加図**

**化学構造比較、パスウェイ予測など**

京都大学化学研究所バイオインフォマティクスセンター

2014/8/20

11

## KEGG, ゲノムネットの配列データ (GENES, ORTHOLOGY(KO), OC)

2014/8/20

12

# KEGG: Kyoto Encyclopedia of Genes and Genomes

The screenshot shows the KEGG homepage with several databases highlighted:

- Orthology (KO)**: 機能アノテーション情報
- GENES**: 配列データベース
- OC**: オーソログクラスター
- KEGG GENES**: Genes and proteins
- KEGG GENOME**: Genomes [KEGG or NCBI]
- KEGG MODULE**: KEGG modules [NCBI]
- KEGG BRITE**: BRITE functional hierarchy
- KEGG BRW**: Drugs [ATC drug classification]
- KEGG DBGET**: Ortholog groups
- KEGG LIGAND**: Chemical information
- KEGG MEDICUS**: Health-related information resource
- KEGG MAPPER**: Navigation tool to explore KEGG pathways
- KEGG Atlas**: KEGG automatic annotation
- KAAS**: KEGG automatic annotation
- BLAST/Fasta**: Sequence similarity search
- SIMCOMP**: Chemical structure similarity search
- PathPred**: Biodegradation/biosynthesis pathway prediction

The right side of the page shows the GenomeNet interface with various databases and tools.

2014/8/20

13

## KEGG GENES データベース

KEGG GENESの  
サブカテゴリ

配列の種類  
アノテーションの種類  
(manual, KOALA, KAAS)

The screenshot shows the KEGG GENES search interface with the following sections:

- Gene Catalogs**: A detailed description of KEGG GENES gene catalogs.
- Search**: Search fields for org:gene, Gene cluster, Ortholog, Paralog, Motif, and Clear.
- Search**: Search fields for Organism (hsa) and for, with options bfind mode and bget mode.

Gene catalog	Category	Remark
GENES	Complete genomes	High-quality genomes with KOALA and manual annotations
DGENES	EST datasets	Draft genomes with automatic (KAAS) annotation
EGENES	Metagenomes	EST contigs with automatic (KAAS) annotation
MGENES	Viruses	Metagenomes with automatic (KAAS) annotation
VGENES		No annotation; available only in DBGET

2014/8/20

14

登録データの一覧表へのリンク  
生物種単位、サンプル単位

# KEGG Organisms

## - GENESに登録されている生物種 -



### KEGG Organisms: Complete Genomes

Eukaryotes: 190 Bacteria: 2323 Archaea: 153

[ Genomes | Draft | ESTs | Meta | Pan ]

Eukaryotes	分類	生物種コード	生物種名	データソース
2014/8/20				
Category		Species		Source
		hsa	Homo sapiens (human)	RefSeq
		ptr	Pan troglodytes (chimpanzee)	RefSeq
		pps	Pan paniscus (bonobo)	RefSeq
		ggo	Gorilla gorilla gorilla (western lowland gorilla)	RefSeq
		pon	Pongo abelii (Sumatran orangutan)	RefSeq
		mcc	Macaca mulatta (rhesus monkey)	RefSeq
		mmu	Mus musculus (mouse)	RefSeq
		rno	Rattus norvegicus (rat)	RefSeq
		cfa	Canis familiaris (dog)	RefSeq
		aml	Ailuropoda melanoleuca (giant panda)	RefSeq
		fca	Felis catus (domestic cat)	RefSeq
		bta	Bos taurus (cow)	RefSeq
		ssc	Sus scrofa (pig)	RefSeq
		ecb	Equus caballus (horse)	RefSeq
		mdo	Monodelphis domestica (opossum)	RefSeq
		shr	Sarcophilus harrisii (Tasmanian devil)	RefSeq
		oaa	Ornithorhynchus anatinus (platypus)	RefSeq
		gga	Gallus gallus (chicken)	RefSeq

15

# KEGG Metagenomes

## - MGENESに収録されているサンプル -



### KEGG Metagenomes

[ Genomes | ESTs | Meta | Pan ]

Environmental samples	分類	生物種コード	生物種名	データソース
2014/8/20				
Category		Project		Source
Ocean	T30001		Planktonic microbial communities from North Pacific Subtropical Gyre	NCBI
	T30002		Planktonic microbial communities from Monterey Bay, CA	NCBI
	T30003		Human gut metagenome collected from healthy human sample F1-S (male adult)	Metagenome.jp
	T30004		Human gut metagenome collected from healthy human sample F1-T (female adult)	Metagenome.jp
	T30005		Human gut metagenome collected from healthy human sample F1-U (infant female)	Metagenome.jp
	T30006		Human gut metagenome collected from healthy human sample F2-V (male adult)	Metagenome.jp
	T30007		Human gut metagenome collected from healthy human sample F2-W (female adult)	Metagenome.jp
	T30008		Human gut metagenome collected from healthy human sample F2-X (male child)	Metagenome.jp
	T30009		Human gut metagenome collected from healthy human sample F2-Y (female child)	Metagenome.jp
	T30010		Human gut metagenome collected from healthy human sample In-A (male adult)	Metagenome.jp
	T30011		Human gut metagenome collected from healthy human sample In-B (male infant)	Metagenome.jp
	T30012		Human gut metagenome collected from healthy human sample In-D (male adult)	Metagenome.jp
	T30013		Human gut metagenome collected from healthy human sample In-E (male infant)	Metagenome.jp
	T30014		Human gut metagenome collected from healthy human sample In-M (infant female)	Metagenome.jp
	T30015		Human gut metagenome collected from healthy human sample In-R (female adult)	Metagenome.jp
	T30016		MH0001 MetaHIT sample from healthy Danish female	MetaHIT
	T30017		MH0002 MetaHIT sample from healthy Danish female	MetaHIT
	T30018		MH0003 MetaHIT sample from healthy Danish male	MetaHIT

16

# 生物種のゲノム情報

- 種の系統、データソース、アノテーションレベル、文献などの詳細情報 -



Genome info    Pathway map    Brite hierarchy    Module    Genome map    Blast    Taxonomy

Search genes:  Go Clear 遺伝子のキーワード検索

**Genome information**

**T number** T01001  
**Org code** hsa  
**Aliases** HUMAN, 9606  
**Full name** Homo sapiens (human)  
**Definition** Homo sapiens (human)  
**Annotation** manual  
**Taxonomy** TAX: 9606  
**Lineage** Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo  
**Data source** RefSeq (Project: 168)  
**Original DB** NCBI, OMIM, HGNC, HPRD, Ensembl  
**Statistics** Number of protein genes: 19761  
Number of RNA genes: 1431  
**Reference** PMID: 11237011  
2014/8/20  
**Authors** Lander ES, et al.

17

## bfind/bget – ワード、エントリー検索

- データ検索、データ取得のためのシステム
- bfind: キーワード検索
- bget: エントリの取得

「Histidine」で GENESにbget

「eco:b4139」で bfind

2014/8/20

18

# KEGG GENESのIDを探したいときは？

## - KEGG Objects -

The left screenshot shows the KEGG Objects page where a user has entered KEGG GENES IDs (hsa-3843, G00109, R00259, R004458, RC00046, ec:2.7.10.1) and clicked 'Convert'. The right screenshot shows the resulting BCONV Result page with the converted NCBI GeneIDs.

KEGG Object ID	NCBI GeneID
hsa-3843	3775638
G00109	3737440
R00259	3743551
R004458	3743551
RC00046	3743551
ec:2.7.10.1	3902295

2014/8/20

19

KEGG Objectのページで、  
NCBI GeneID, NCBI gi, UniProt から、  
KEGG GENES entry の IDに変換可能

# KEGG GENESと他DBのIDとの対応関係の一括取得 - GenomeNet LinkDB -

The search results page for KEGG GENES entries (hsa-3843, G00109, R00259, R004458, RC00046, ec:2.7.10.1) displays a grid of equivalent links to various databases like NCBI, UniProt, and ChEMBL.

LinkDBから、  
NCBI-GenID, NCBI-gi, Genbank,  
UniGene, UniProt とKEGG GENESの  
特定の生物種の遺伝子IDとの対  
応リストを取得可能

2014/8/20

1	ath:AT1G01010	ncbi-geneid:339580	equivalent
2	ath:AT1G01020	ncbi-geneid:339568	equivalent
3	ath:AT1G01030	ncbi-geneid:339321	equivalent
4	ath:AT1G01050	ncbi-geneid:339574	equivalent
5	ath:AT1G01050	ncbi-geneid:339579	equivalent
6	ath:AT1G01060	ncbi-geneid:339341	equivalent
7	ath:AT1G01070	ncbi-geneid:339550	equivalent
8	ath:AT1G01070	ncbi-geneid:7922546	equivalent
9	ath:AT1G01080	ncbi-geneid:339463	equivalent
10	ath:AT1G01090	ncbi-geneid:339429	equivalent
11	ath:AT1G01100	ncbi-geneid:339410	equivalent
12	ath:AT1G01110	ncbi-geneid:339394	equivalent
13	ath:AT1G01115	ncbi-geneid:7922461	equivalent
14	ath:AT1G01120	ncbi-geneid:339395	equivalent
15	ath:AT1G01130	ncbi-geneid:339359	equivalent
16	ath:AT1G01140	ncbi-geneid:339349	equivalent
17	ath:AT1G01150	ncbi-geneid:339334	equivalent
18	ath:AT1G01160	ncbi-geneid:339278	equivalent
19	ath:AT1G01170	ncbi-geneid:339263	equivalent
20	ath:AT1G01180	ncbi-geneid:339232	equivalent
21	ath:AT1G01190	ncbi-geneid:339233	equivalent
22	ath:AT1G01200	ncbi-geneid:339499	equivalent
23	ath:AT1G01210	ncbi-geneid:339481	equivalent
24	ath:AT1G01220	ncbi-geneid:339420	equivalent
25	ath:AT1G01225	ncbi-geneid:339358	equivalent
26	ath:AT1G01230	ncbi-geneid:339331	equivalent
27	ath:AT1G01240	ncbi-geneid:339274	equivalent
28	ath:AT1G01250	ncbi-geneid:339322	equivalent
29	ath:AT1G01280	ncbi-geneid:339345	equivalent
30	ath:AT1G01280	ncbi-geneid:339470	equivalent
31	ath:AT1G01290	ncbi-geneid:339445	equivalent
32	ath:AT1G01300	ncbi-geneid:339375	equivalent
33	ath:AT1G01305	ncbi-geneid:5240609	equivalent
34	ath:AT1G01310	ncbi-geneid:339333	equivalent
35	ath:AT1G01320	ncbi-geneid:339293	equivalent
36	ath:AT1G01340	ncbi-geneid:339882	equivalent
37	ath:AT1G01350	ncbi-geneid:339864	equivalent
38	ath:AT1G01355	ncbi-geneid:5241268	equivalent
39	ath:AT1G01360	ncbi-geneid:338452	equivalent
40	ath:AT1G01370	ncbi-geneid:339104	equivalent
41	ath:AT1G01380	ncbi-geneid:338096	equivalent
42	ath:AT1G01390	ncbi-geneid:337790	equivalent
43	ath:AT1G01400	ncbi-geneid:337049	equivalent
44	ath:AT1G01410	ncbi-geneid:337547	equivalent
45	ath:AT1G01420	ncbi-geneid:837003	equivalent
46	ath:AT1G01430	ncbi-geneid:337247	equivalent
47	ath:AT1G01440	ncbi-geneid:337155	equivalent
48	ath:AT1G01450	ncbi-geneid:337060	equivalent

20

# GENESをbfindで検索

- 例: aspartate ammonia-lyase -

The left window shows the KEGG GENES Database search page. It includes a search bar for 'aspartate ammonia-lyase' and two radio button options: 'bfnd mode' (selected) and 'bget mode'. The right window shows the DBGET search results for 'aspartate ammonia-lyase' (Total 801 hits). A red box highlights the entry for 'aspartate ammonia-lyase (EC:4.3.1.1)'.

2014/8/20 21

## KEGG GENES Entry (例: eco:b4139)

The left window displays the KEGG GENES entry for 'aspartate ammonia-lyase' (EC:4.3.1.1). It includes sections for Definition, Orthology, Pathway, Class, SSDB, Motif, Other DBs, and Structure. The right window shows the DBGET integrated database retrieval system, specifically the 'Position' tab for 'aspartate ammonia-lyase'.

Entry: エントリ名、種類、生物種名

Gene name: 遺伝子・タンパク質名、別名

Definition: オリジナルDBの機能アノテーション

Ortholog: KEGGでアサインしたKEGG Orthology (KO)

アノテーション(後述)

Pathway: エントリの遺伝子が機能するPathwayへのリンク

Class: KEGGにおける機能カテゴリBRITE(階層テキスト)  
20ページのリンク

SSDB: SSDB (Similarity Score Database)からの情報抽出  
用のリンク。(Ortholog/Paralog推定、Gene Cluster)

Motif: エントリが持つドメイン・モチーフの情報

Other DBs: 他の配列DBへのリンク

Structure: PDBへのリンク

Position: エントリのゲノム上の位置(ゲノム情報がある場合)

AA seq: アミノ酸配列取得、ホモロジー検索

NT seq: 塩基配列取得。

All links: LinkDBでリンクされている内部および外部DB<sup>22</sup>

# KO (KEGG Orthology)

- KEGGでは同じ機能を有していると考えられるオーソログ遺伝子を集め、同一のID(K番号)をつけることでその機能を表現している

- 新規ゲノムが新しくGenesに登録されるとき、遺伝子をオーソロググループ(KO)に追加することで、遺伝子の機能アノテーションを行っている

2014/8/20

23

## KO Entry

Entry	K01744	KO
Name	aspA	
Definition	aspartate ammonia-lyase [EC:4.3.1.1]	
Pathway	ko00250 Alanine, aspartate and glutamate metabolism ko00910 Nitrogen metabolism	
Class	Metabolism; Energy Metabolism; Nitrogen metabolism [PATH:ko00910] Metabolism; Amino Acid Metabolism; Alanine, aspartate and glutamate metabolism [PATH:ko00250] <a href="#">BRITE hierarchy</a>	
Other DBs	RN: R00490 COG: COG0127 GO: 0008797	
Genes	b4139 (aspA) ECJ: JW4099 (aspA) ECO: ECOH10B_432 (aspA) EBI: BWG_3852 (aspA) ECE: Z5744 (aspA) ECS: ECs5120 (aspA) EOF: EOF74115_5655 (aspA) ETW: ECSP_5239 (aspA) EOJ: EO206_5305 (aspA) EOI: ECO111_5074 (aspA) EOH: ECO103_4934 (aspA) ECG: E234BC_4465 (aspA) EOK: G2583_4986 (aspA) ECC: c5222 (aspA) ECP: ECP_4383 (aspA) ECI: UTI89_C4736 (aspA) ECV: APEC01_2250 (aspA)	

DBGET integrated database retrieval system

2014/8/20

24

# ゲノムネットゲノム情報解析ツール

varDB  
研究支援データベース  
計算ツール  
その他のツール  
FTP  
フィードバック

KEGG BRITE - システム情報: オントロジー  
KEGG Organisms - 生物種ごとの入口  
KEGG GENES - ゲノム情報  
KEGG LIGAND - ケミカル情報  
Reaction Ontology: 反応パターンと分類  
varDB: 抗原変異データベース  
研究支援データベース  
CYORF - シアノバクテリア  
BSORF - 枯草菌  
EXPRESSION - 遺伝子発現プロファイル

ゲノムネット計算ツール

配列解析  
BLAST / FASTA - ホモロジー検索  
MOTIF - モチーフ検索  
CLUSTALW / MAFFT / PRRN - 複数アライメント

**ゲノム情報解析**

- OC Viewer - オーソログクラスタ
- KAAS - KEGG自動アノテーションサーバー
- EGassembler - ESTコンティング作成
- GENIES - 遺伝子ネットワーク予測
- DINIES - 医薬品-標的ネットワーク予測 *New!*

ケミカル情報解析  
SIMCOMP / SUBCOMP - 化合物構造検索  
KCal - 糖鎖構造検索  
PathComp - 可能な反応経路の計算  
PathSearch - 類似反応経路検索  
PathPred - 分解・合成反応経路予測  
E-zyme - 化合物間の酵素反応予測

京都大学化学研究所バイオインフォマティクスセンター

オーソログクラスタ、自動アノテーション、ネットワーク予測など

2014/8/20 [http://www.genome.jp/ \(英語\)](http://www.genome.jp/)  
[http://www.genome.jp/ja/ \(日本語\)](http://www.genome.jp/ja/)

25

## KAAS (KEGG 自動アノテーションサーバー)

- 入力
  - 大規模シークエンスによって得られた生物種のアミノ酸や塩基配列
- 出力
  - 配列名とKOの対応表
  - PATHWAYマップ
  - BRITE階層分類

KAAS - KEGG Automatic Annotation Server  
for ortholog assignment and pathway mapping

**About KAAS**

KAAS (KEGG Automatic Annotation Server) provides functional annotation of genes by BLAST comparisons against the manually curated KEGG GENES database. The result contains KO (KEGG Orthology) assignments and automatically generated KEGG pathways.

- KAAS Help

**Complete or Draft Genome**

KAAS works best when a complete set of genes in a genome is known. Prepare query amino acid sequences and use the SBH (bi-directional best hit) method to assign orthologs.

- KAAS job request (BBH method)

**Partial Genome**

KAAS can also be used for a limited number of genes. Prepare query amino acid sequences and use the SBH (single-directional best hit) method to assign orthologs.

- KAAS job request (SBH method)  
- KAAS interactive

**ESTs**

When ESTs are comprehensive enough, a set of consensus contigs can be generated by the EGassembler server and used as a gene set for KAAS with the BBH method. Otherwise, use ESTs as they are with the SBH method.

- KAAS job request (BBH method)  
- KAAS job request (SBH method)

**Example of Results**

**KO assignment**

Query gene : KO assignment
query_2001
query_3002
query_3003
query_3004
query_3005
query_3006
query_3007
query_3008
query_3009
query_3010
query_3011
query_3012
query_3013
query_3014
query_3015

**KEGG pathway mapping**

Diagram illustrating the KEGG pathway mapping process, showing the flow from ESTs to pathway assignment.

2014/8/20

26

# OC Viewer

## - KEGG OC (Ortholog Cluster) -

KEGGにある全生物種の遺伝子の、  
Ortholog Clusterのデータベース

The screenshot shows the OC Viewer interface. At the top, there's a logo for 'Ortholog Cluster' and the title 'OC Viewer'. Below that is a 'Home' section with a brief introduction about the KEGG OC database. Under 'OC search', there are two input fields: 'ID or terms query' containing 'pyruvate dehydrogenase org=bsu' and 'ex. hsa:362, K04517, M00170, pyruvate dehydrogenase org=bsu'. Below these is an 'Amino acid sequence query' field with 'ex. hsa:5162'. On the right side of the search results, there's a list of entries starting with 'OC.174882 (947 / 948) - Bacteria - K00163(932)'.

2014/8/20

27

# OC Viewer

## - KEGG OC (Ortholog Cluster) -

This screenshot shows a detailed view of the OC.174882 cluster. At the top, it says 'OC.174882 (948 sequences)'. Below that is a table listing various organisms and their corresponding KEGG IDs. The table includes columns for 'Organism' and 'KO ID'. A green highlight covers the first few rows. At the bottom of the table, there's a note: '(view text of OC.174882)'. Below the table, there's a 'Refinement of cluster' section with a dropdown menu set to 'OC Alpha'. To the right is an 'Organism selection' tree view showing a hierarchy of taxonomic groups. At the bottom left is a 'Distribution of taxonomic categories' pie chart. On the right, there's a 'Neighbor clusters' section showing a dendrogram of related clusters.

2014/8/20

28

# GenomeNetからの遺伝子データベース一括検索機能

- 例: aspartate ammonia-lyase -

The screenshot shows two browser windows side-by-side. The left window is the GenomeNet homepage, featuring a navigation bar with links to KEGG, KEGG2, PATHWAY, BRITE, MEDICUS, and DBGET. Below the navigation is a search bar with the query 'aspartate ammonia-lyase'. To the right of the search bar is a 'DB growth curve' graph. The right window is a detailed search result titled 'DBGET Search Result: All databases...'. It lists various entries related to the enzyme, including chemical reactions, gene identifiers (e.g., K01744, K04835), and protein identifiers (e.g., P00487, P00488). The results are categorized under 'Gene' and 'KEGG ORTHOLOGY'.

2014/8/20

29

## KEGG, ゲノムネットの化合物情報 (LIGAND)

2014/8/20

30

# KEGG: Kyoto Encyclopedia of Genes and Genomes

**KEGG Home**  
Release notes  
Current statistics  
Plea from KEGG

**KEGG Database**  
KEGG overview  
Searching KEGG  
KEGG mapping  
Color codes

**KEGG Objects**  
Pathway maps  
Brite hierarchies

**KEGG Software**  
KegTools  
KEGG API  
KGML

**KEGG FTP**  
Subscription

**GenomeNet**  
DBGET/LinkDB  
Feedback

**Kanehisa Labs**

**KEGG: Kyoto Encyclopedia of Genes and Genomes**

KEGG is a database resource for understanding high-level functions and utilities of the biological system, such as the cell, the organism and the ecosystem, from molecular-level information especially large-scale molecular datasets generated by genome sequencing and experimental technologies (See [Release features](#)).

**LIGAND**  
**化合物情報データベース**

**SIMCOMP/SUBCOMP**  
**化合物類似構造・部分構造検索**

**GenomeNet Database Resources**

DBGET: Integrated Database Retrieval System  
DBGET search  
LinkDB search [Renewal!](#)

KEGG: Kyoto Encyclopedia of Genes and Genomes  
KEGG2 - Table of contents  
KEGG PATHWAY - Systems information pathways  
KEGG BRITE - Systems information ontologies  
KEGG MODULE - Organism-specific entry points  
KEGG ORTHOLOGY - Ortholog groups [KO system]  
KEGG GENOME - Genomes [KEGG organisms]  
KEGG GENES - Genes and proteins - Release history  
Chemical information [Reaction modules]

**Community Use**  
Bioinformatics tools  
Other tools  
FTP  
Feedback

**KEGG LIGAND** [chemical information]  
Kegg Metabolism [Health-related info resource]  
MEDICUS search  
Reaction Ontology: Reaction classifications  
verb0: Antigenic variation database  
Community Databases  
CYORF: Cyanobacteria annotation database  
BSORF: Bacillus subtilis genome database  
EXPRESSION: Gene expression profile database

**GenomeNet Bioinformatics Tools**

Sequence Analysis  
BLAST/Fasta - Sequence similarity search  
MOTIF - Sequence motif search  
CLUSTALW / MAFFT / PRRN - Multiple alignment

Genome Analysis  
OC Viewer - KEGG ortholog clusters  
KAAS - KEGG automatic server  
EGassembler - EST consensus contigs  
GENIES - Gene network prediction  
DINIES - Drug-target network prediction [New!](#)

**SIMCOMP/SUBCOMP** [chemical structure search]  
Kegg organic structure search  
PathComp - Possible reaction path computation  
PathSearch - Similar reaction path search  
PathPred - Reaction pathway prediction  
Ezyme - Enzymatic reaction prediction

Copyright 1995-2012 Kanehisa Laboratories

Kyoto University Bioinformatics Center

2014/8/20

31

## KEGG LIGAND データベース ( <http://www.genome.jp/kegg/ligand.html> )

KEGG LIGANDの  
サブカテゴリ

化合物、糖鎖、反応  
に特化した入口

**KEGG2 PATHWAY BRITE MODULE LIGAND COMPOUND GLYCAN REACTION**

Enter C/G/R/RP/RC numbers (Example) C00389 C05903 C06562 C08650 C09727 C09726

Filter Pathway mapping Brite mapping Get title Get entry Clear

**Chemical Substances and Reactions**

KEGG LIGAND contains our knowledge on the universe of chemical substances and reactions that are relevant to life. It is a composite database consisting of COMPOUND, GLYCAN, REACTION, RPAIR, RCLASS, and ENZYME databases, whose entries are identified by C, G, R, RP, RC, and EC numbers, respectively. ENZYME is derived from the IUBMB/IUPAC Enzyme Nomenclature, but the others are internally developed and maintained.

Database	Identifier	Content	Specialized entry point
COMPOUND	C number	Chemical compound structures	KEGG COMPOUND
GLYCAN	G number	Glycan structures	KEGG GLYCAN
REACTION	R number	Biochemical reactions	KEGG REACTION
RPAIR	RP number	Reactant pair alignments	
RCLASS	RC number	Reaction class	
ENZYME	EC number	Enzyme nomenclature	

Search LIGAND for Go Clear

bfind mode bget mode

ID番号のIdentifierと  
各DBに含まれるデータの説明

KEGGにおける「Chem(o)-」の部分、すなわち生化学的な情報全般を扱う

2014/8/20

32

# KEGG COMPOUND (代謝化合物情報)

- 主に代謝化合物を収録したデータベース
  - 中心代謝、二次代謝物 etc.
- 構造は2次元のグラフとして持つ
  - ファイルフォーマットは、MDL/mol and KCF
- KCFフォーマットの特徴
  - KEGG Atom-Type による原子表記

2014/8/20

33

## KCFフォーマットとKEGG Atom type

- KCF (KEGG Chemical Function)
  - KEGGにおける化合物データのフォーマット
  - MDL/mol類似のフォーマットで、原子情報にKEGG atom typesを使用
- KEGG atom types
  - 原子の周辺環境を反映させて、68種類に原子タイプを細分類
    - 1)結合パターン
    - 2)周辺原子種
    - 3)リング、芳香性
  - SIMCOMPなどの構造比較計算に利用

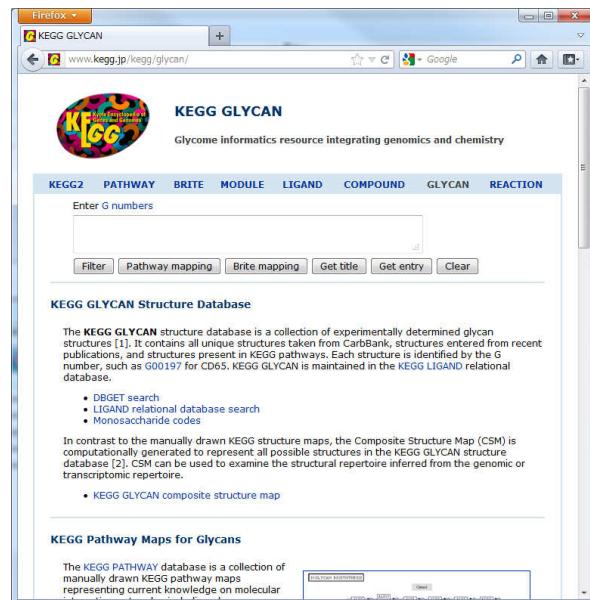
Atom	Functional group	Atom type	Description	Frequency
C	Alkane	C1a	R-CH3	16473
		C1b	R-CH2-R	20193
		C1c	R-CH(R)-R	4964
		C1d	R-C(=O)-R	1595
Cyclic alkane		C1x	ring-CH2-ring	14910
		C1y	ring-CH(R)-ring	27376
		C1z	ring-CH(R)2-ring	4463
		C2a	R=CH2	634
Alkene		C2b	R=CH-R	3965
		C2d	R=C(R)2	1914
		C2x	ring-CH=ring	2964
		C2y	ring-C(=R)=ring or ring-C(=R)-ring	3722
Cyclic alkene		C3a	R=CH	43
		C3b	R=C(R)	282
		C5a	R-C(=O)-O	350
		C5g	R-C(=O)-R	3995
Alkyne		C5x	ring-C(=O)-ring	2257
		C6a	R-C(=O)-OH	3190
		C7a	R-C(=O)-O-R	1691
		C7x	ring-C(=O)-O-ring	669
Aldehyde		C8a	ring-CH=ring	1990
		C8v	ring-C(=R)=ring	20511
		C9		8
		N1a	R-NH2	2440
Ketone	Amine	N1b	R-NH-R	3003
		N1c	R-(N(R))2	374
		N1d	R-(N(R))+	105
		N1v	ring-N(R)-ring	406
Cyclic ketone	Cyclic amine	N1y	ring-(N(R))2-ring	1464
		N2a	R=N-H	230
		N2b	R=N-R	163
		N2x	ring-N=ring	357
Carboxylic acid	Imine	N2y	ring-N(R)+=ring	14
	Cyclic imine	N3a	R=NH	110
		N4x	ring-NH=ring	785
		N4y	ring-N(R)-ring	840
Carboxylic ester	Cyan	N5a	ring-N=ring	2131
	Aromatic ring	N5y	ring-N(R)+=ring	59
		N6		194
		O1a	R-OH	18369
Aromatic ring	Hydroxy	O1b	R-OH	199
		O1c	P-OH	3113
		O1d	S-OH	332
		O2a	R-O-R	4199
Undefined N	Ether	O2b	P-O-R	2481
		O2c	P-O-P	607

2014/8/20

34

## KEGG GLYCAN (糖鎖構造情報)

- 糖鎖分子の構造を収録したデータベース
  - エントリー数は、約1万1000
- 構造は2次元のグラフとして持つ
  - グラフは、実際にはツリーとなる
- ファイルフォーマットはKCF(Glycan)
  - Compound/Drugとの互換性はない

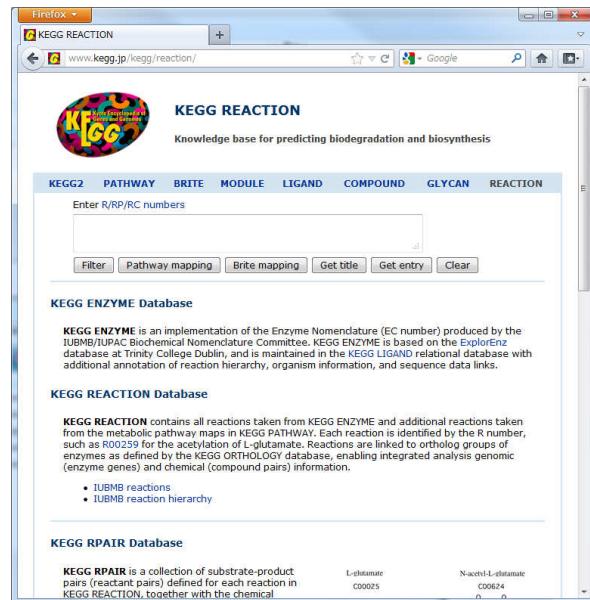


2014/8/20

35

## KEGG REACTION, RPAIR, RCLASS, ENZYME (生化学反応情報)

- REACTAION (生化学反応)
  - 生体内反応のデータベース
    - 代謝系の酵素反応がメイン
- RPAIR(基質-生成物変化)
- RCLASS(反応分類)
  - RPAIRのパターンによる反応分類
- ENZYME(酵素番号)
  - IUBMBの酵素番号



2014/8/20

36

# REACTION, RPAIR, RCLASS

## - R00259 acetyl-CoA:L-glutamate N-acetyltransferase を例に -

Firefox X

KEGG REACTION: R00259

www.kegg.jp/dbget-bin/www\_bget?r=R00259

Google

**REACTION: R00259**

Entry R00259 Reaction

Name acetyl-CoA:L-glutamate N-acetyltransferase

Definition Acetyl-CoA + L-Glutamate  $\leftrightarrow$  CoA + N-Acetyl-L-glutamate

Equation C00024 + C00025  $\leftrightarrow$  C00010 + C00624

RPair R000007\_C00010\_C00024\_main  
R00458\_C00025\_C00624\_main  
R00475\_C00024\_C00624\_trans

Enzyme 2.3.1.1

Pathway rr00330 Arginine and proline metabolism  
rr01100 Metabolic pathways  
rr01110 Biosynthesis of secondary metabolites

Orthology K00618 amino-acid N-acetyltransferase [EC:2.3.1.1]  
K00619 amino-acid N-acetyltransferase [EC:2.3.1.1]  
K00620 glutamate N-acetyltransferase / amino-acid N-acetyltransferase [EC:2.3.1.35 2.3.1.1]  
KI1087 N-acetylglutamate synthase [EC:2.3.1.1]  
KI4681 argininosuccinate lyase / amino-acid N-acetyltransferase [EC:4.3.2.1 2.3.1.1]  
KI4682 amino-acid N-acetyltransferase [EC:2.3.1.1]

2014/8/20

L-glutamate C00025 → N-acetyl-L-glutamate C00624

Reaction center N1a → N1b  
Difference atom (H) → C5a  
Matched atom C1c → C1c

- Reactionの基質と生成物の間で、原子の由来関係を持つ化合物のペアを Reactant Pair と定義
- RPAIRは、Reactant Pairを原子アラインメントして、反応前後の変化を RDMパターンとしてデータ化
- Reaction は RPAIRの組み合わせで表現できる
- Reactant Pairは、生化学的意味に基づきラベル付け可能
- RCLASSは、mainとラベル付けされた Reactant pairの変化の類似性に基づいた反応の分類

37

# LIGANDをキーワード検索 (1)

## - 例: phenylalanine -

Firefox X

KEGG LIGAND Database

www.kegg.jp/kegg\_ligand.htm

Search LIGAND for Phenylalanine

Database: LIGAND - Search term: Phenylalanine

KEGG COMPOUND

C00779 L-Phenylalanine; (S)-alpha-Amino-beta-phenylpropionic acid; C00780 D-Phenylalanine; 3-Hydroxy-L-lysine; L-beta-(3,4-Dihydroxyphenyl)alanine; Levodopa; D,L-phenylalanine; Phenylalanine; DL-Phenylalanine; alpha-Amino-beta-phenylpropionic acid; C00781 Phenylalanine-DL-Phenylalanine; alpha-Amino-beta-phenylpropionic acid

...+ display all

R00031 L-Phenylalanine dehydrogenase; Oxygen + L-Tryptone  $\rightarrow$  2,3-Dihydroxy-L-phenylalanine + 2 CO<sub>2</sub>; R00032 3-O-hydroxy-L-phenylalanine oxygen endotranslocase; Oxygen + 2,3-Dihydroxy-L-phenylalanine  $\rightarrow$  2,3-Dihydroxy-L-phenylalanine + 2 H<sub>2</sub>O; R00033 ...+ display all

R00034 L-Phenylalanine lactamase (ATP hydrolase); ATP + L-Phenylalanine + H<sub>2</sub>O  $\rightarrow$  AMP + Diphosphate + L-Leucine + NH<sub>4</sub><sup>+</sup>; R00035 L-Phenylalanine aminopeptidase/transaminase; ATP + L-Phenylalanine  $\leftrightarrow$  Diphosphate + N-Acetyl-L-phenylalanine; R00036 L-Phenylalanine H<sub>2</sub>O + nitroreductase (deaminating); L-Phenylalanine + H<sub>2</sub>O + NADH  $\rightarrow$  Phenylpyruvate + H<sub>2</sub>O + NAD<sup>+</sup>; R00037 ...+ display all

KEGG ENZYME

1.4.1.20 L-phenylalanine dehydrogenase; L-phenylalanine dehydrogenase; PHD; Phenylalanine 2-monooxygenase; L-phenylalanine oxidase (deaminating and decarboxylating); Phenylalanine-decarboxylase; Phenylalanine-dehydrogenase

Database: COMPOUND - Search term: Phenylalanine (Total 20 hits)

C00079 KegDraw Jmol L-Phenylalanine; (S)-alpha-Amino-beta-phenylpropionic acid

...+ display all

C00282 KegDraw Jmol 3-O-Dihydroxy-L-phenylalanine; L-Dopa; 3-Hydroxy-L-tyrosine; L-beta-(3,4-Dihydroxyphenyl)alanine; Levodopa; D,L-phenylalanine; Phenylalanine; DL-Phenylalanine; alpha-Amino-beta-phenylpropionic acid

C003519 KegDraw Jmol Phenylalanine; DL-Phenylalanine; alpha-Amino-beta-phenylpropionic acid

C003708 KegDraw Jmol Phenylalanine-DL-Phenylalanine; alpha-Amino-beta-phenylpropionic acid

C00620 KegDraw Jmol N-Acetyl-L-phenylalanine

C011217 KegDraw Jmol 2,3-Dihydroxy-L-phenylalanine

C11341 KegDraw Jmol N-Acetylphenylalanine-beta-naphthyl ester; N-Acetyl-DL-phenylalanine beta-naphthyl ester

C11615 KegDraw Jmol FMPL; N-(2-Methyl-4-methoxy-L-leucyl)-L-phenylalanine

C11616 KegDraw Jmol Pristinamycin IA; Vernamycin B alpha; Mikanycin B; 4-(4-(Dimethylamino)-N-methyl-L-phenylmethyl)amino; Streptogramin B

C11617 KegDraw Jmol Pristinamycin IB; Vernamycin B beta; Epefrisin; 4-(N-Methyl-4-(methylamino)-L-phenylalanine)arginamycin S1

C11618 KegDraw Jmol Pristinamycin IC; Vernamycin B gamma; 2-D-Manane-4-(4-(dimethylamino)-N-methyl-L-phenylalanine)arginamycin S1

C12031 KegDraw Jmol 3-Aminophenylalanine; p-Aminophenylalanine

C17235 KegDraw Jmol L-Homophenylalanine

C18819 KegDraw Jmol 4-Omethylaminohomophenylalanine

2014/8/20

38

# KEGG COMPOUND Entry

STRUCTURE: C00079  
Name: L-Phenylalanine; (S)-alpha-Amino-beta-phenylproprionic acid  
Formula: C9H11NO2  
Exact mass: 165.079  
Mol weight: 165.1891  
Structure: C[C@H](C(=O)Nc1ccccc1)C

**Pathway:**  
ko00360 Phenylalanine metabolism  
ko00400 Phenylalanine, tyrosine and tryptophan biosynthesis  
ko00840 Phenylpropanoid biosynthesis  
ko00930 Phenylpropanoid and tyrosine alkaloid biosynthesis  
ko00988 Glucosinolate biosynthesis  
ko00970 Aminoacyl-tRNA biosynthesis  
ko01080 Biosynthesis of plant secondary metabolites  
ko01061 Biosynthesis of phenylpropanoids  
ko01093 Biosynthesis of alkaloids derived from shikimate pathway  
ko01084 Biosynthesis of alkaloids derived from ornithine, lysine and nicotinic acid  
ko01070 Biosynthesis of plant hormones  
ko01100 Metabolic pathways  
ko01110 Biosynthesis of secondary metabolites  
ko02000 ABC transporters  
ko04614 Lipid digestion and absorption  
ko04970 Mineral absorption

**Enzyme:**  
1.4.1.20 1.4.3.2 1.11.1.7 1.11.1.21  
1.13.12.9 1.14.13.124 1.14.16.1 2.3.1.53  
2.6.1.1 2.6.1.5 2.6.1.8 2.6.1.28  
2.6.1.57 2.6.1.58 2.6.1.64 2.6.1.70  
2.7.7.54 3.4.17.23 4.1.1.28 4.1.1.53  
4.2.1.51 4.2.1.61 4.3.1.24 4.3.1.25  
5.3.1.1 5.3.1.1 5.3.1.1 5.3.1.1

**Other DBs:**  
QAS: 89-91-2  
PubChem: 3379  
ChEBI: 17295  
KNAcSack: C00001386  
PDB-CDD: PHE  
SMILES: B01151  
NIKKAJI: J8\_175H

**KCF data:** Show

構造表示、  
構造情報(mol, KCF)、  
構造検索、構造表示(&編集)、

各種データベースへのリンク  
REACTION, PATHWAY, ENZYME

外部データベースへのリンク

PubChem, ChEBI, etc.

LinkDB情報

GenomeNet内部でリンクされているDBや  
対応関係のとれる外部DBを表示

39

## LIGANDをbfindでキーワード検索 (2) - 例: phenylalanine -

Search LIGAND - for Phenylalanine

Database: LIGAND - Search term: Phenylalanine

REACT: REACTION - Search term: Phenylalanine (Total 52 hits)

R00031 1,2-Benzodiol oxygen oxidoreductase; Oxygen + 2-L-Tyrosine <=> 2,3-Dihydroxy-L-phenylalanine  
R00045 2,3-Dihydroxy-L-phenylalanine oxygen oxidoreductase; Oxygen + 2,3-Dihydroxy-L-phenylalanine <=> 2-Dopamine + 2 H2O  
R00046 L-Phenylalanine racemase (ATP-hydrolyzing); ATP + L-Phenylalanine <=> AMP + Diphosphate + D-Phenylalanine  
R00047 ATP:L-phenylalanine adenyltransferase; ATP + L-Phenylalanine <=> Diphosphate + N-Adenylyl-L-phenylalanine  
R00048 L-phenylalanine NAD+ oxidoreductase (deaminating); L-Phenylalanine + H2O + NAD+ <=> Phenylpyruvate + NADH + H+  
R00049 L-phenylalanine oxygen oxidoreductase (deaminating); L-Phenylalanine + H2O + Oxygen <=> Phenylpyruvate + H2O + Hydrogen peroxide  
R00050 L-phenylalanine oxygen oxidoreductase (decarboxylating); L-Phenylalanine + CO2 + H2O  
R00051 L-arginate hydro-lyase (decarboxylating); L-phenylalanine-forming; L-Arogenate <=> L-Phenylalanine + H2O + CO2  
R00052 L-phenylalanine pyruvate aminotransferase; L-Phenylalanine + Pyruvate <=> Phenylpyruvate + L-Alanine  
R00053 Acetyl-CoA:L-phenylalanine N-acetyltransferase; Acetyl-CoA + L-Phenylalanine <=> CoA + N-Acetyl-L-phenylalanine  
R00054 L-Phenylalanine 2-oxoglutarate aminotransferase; L-Phenylalanine + 2-Oxoglutarate <=> Phenylpyruvate + L-Glutamate  
R00055 L-aspartate phenylpyruvate aminotransferase; L-Aspartate + Phenylpyruvate <=> Oxaloacetate + Phenylalanine  
R00056 L-phenylalanine ammonia-lyase (trans-cinnamate-forming); L-Phenylalanine <=> trans-Cinnamate + NH3

2014/8/20

40

# KEGG REACTION Entry

Firefox - KEGG REACTION: R00697

www.kegg.jp/dbget-bin/www\_bget?rn:R00697

**REACTION: R00697**

<b>Entry</b>	R00697	<b>Reaction</b>
<b>Name</b>	L-phenylalanine ammonia-lyase (trans-cinnamate-forming)	
<b>Definition</b>	L-Phenylalanine <=> trans-Cinnamate + NH <sub>3</sub>	
<b>Equation</b>	C00079 <=> C00423 + C00014	
<b>RPair</b>	RP01058 C00079_C00423 main RP05821 C00014_C00079 leave	
<b>Enzyme</b>	4.3.1.24 4.3.1.25	
<b>Pathway</b>	rn00360 Phenylalanine metabolism rn00910 Nitrogen metabolism rn00840 Phenylpropanoid biosynthesis rn01100 Metabolic pathways rn01110 Biosynthesis of secondary metabolites	
<b>Orthology</b>	KI0775 phenylalanine ammonia-lyase [EC:4.3.1.24] KI3064 phenylalanine/tirosine ammonia-lyase [EC:4.3.1.25]	

DBGET integrated database retrieval system

反応式(物質名、C番号、図)、

各種データベースへのリンク  
RPAIR, PATHWAY, ENZYME, KO

LinkDB情報

GenomeNet内部でリンクされているDBや  
対応関係のとれる外部DBを表示

2014/8/20

41

# KEGG RPAIR Entry

Firefox - KEGG RPAIR: RP01058

www.kegg.jp/dbget-bin/www\_bget?rp:RP01058

**RPAIR: RP01058**

<b>Entry</b>	RP01058	<b>RPair</b>
<b>Name</b>	C00079_C00423	
<b>Compound</b>	C00079 L-Phenylalanine C00423 trans-Cinnamate	
<b>Type</b>	main	
<b>RDM</b>	2 1 C1b-C2b:-*-*:C1c+C8y-C2b+C8y 2 C1c-C2b:N1a:-*-*:C1b+C8a-C2b+C8a	
<b>RClass</b>	R00361	
<b>Related pair</b>	RP01074 RP01342	
<b>Reaction</b>	R00697	
<b>Enzyme</b>	4.3.1.24 4.3.1.25	
<b>KCF data</b>	Show	

アライメントの結果(図)

ペアのフラグ情報(生化学的意味)

RDMの文字列表記

各種データベースへのリンク

(COMPOUND, RCLASS, Related RPAIR,  
REACTION, ENZYME)

2014/8/20

42

# KEGG RCLASS Entry

Firefox > KEGG RCLASS: RC00361

www.kegg.jp/dbget-bin/www\_bget?c:RC00361

**RCLASS: RC00361**

Entry	RC00361	RClass
Definition	C1b-C2b:-x-C1c-C8y-C2b+C8y C1c-C2b:N1a-*:C1b+C6a-C2b+C6a	
RPAIR	RPP01058 RPD01074 RPD01342 (DB search)	
Related class	(DB search)	
Reaction	RD0097 RD00737 RD1168	
Enzyme	4.3.1.3    4.3.1.23    4.3.1.24    4.3.1.25	
Pathway	rn00340 Histidine metabolism rn00380 Phenylalanine metabolism rn00910 Nitrogen metabolism rn00940 Phenylpropanoid biosynthesis rn01100 Metabolic pathways rn01110 Biosynthesis of secondary metabolites	
Orthology	K01745 histidine ammonia-lyase [EC:4.3.1.3] K10744 tyrosine ammonia-lyase [EC:4.3.1.23] K10775 phenylalanine ammonia-lyase [EC:4.3.1.24] K13064 phenylalanine/tyrosine ammonia-lyase [EC:4.3.1.25]	

RDMのパターンで、RPAIRを分類

RDM変換パターン(文字列、図)

RCLASSに含まれる RPAIR  
各種データベースへのリンク  
(RPAIR, Related RCLASS,  
REACTION, ENZYME, PATHWAY, KO)

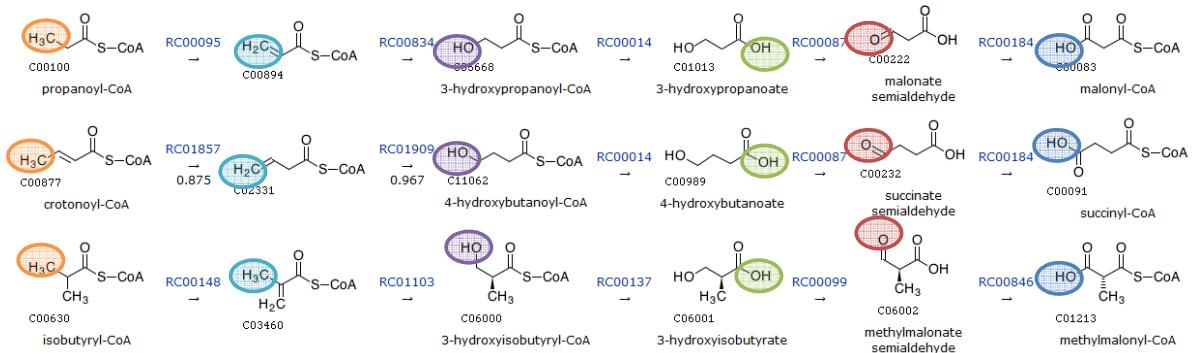
2014/8/20

43

## KEGG reaction modules



構造変化パターンの分類から、類似した連続反応によって代謝が進む代謝経路上の連続反応を集めたデータベース



2014/8/20

44

# ゲノムネットケミカル情報解析ツール

KEGG  
varDB  
研究支援データベース  
計算ツール  
その他のツール  
FTP  
フィードバック

KEGG BRITE - システム情報: オントロジー  
KEGG Organisms - 生物種ごとの入口  
KEGG GENES - ゲノム情報  
KEGG LIGAND - ケミカル情報  
Reaction Ontology: 反応パターンと分類  
varDB: 抗原変異データベース  
研究支援データベース  
CYORF - シアノバクテリア  
BSORF - 枯草菌  
EXPRESSION - 遺伝子発現プロファイル

ゲノムネット計算ツール

配列解析  
BLAST / FASTA - ホモロジー検索  
MOTIF - モチーフ検索  
CLUSTALW / MAFFT / PRRN - 複数アライメント

ゲノム情報解析  
OC Viewer - オーソログクラスタ  
KAAS - KEGG自動アノテーションサーバー  
EGassembler - ESTコンティング作成  
GENIES - 遺伝子ネットワーク予測  
DINIES - 医薬品-標的ネットワーク予測 *New!*

ケミカル情報解析  
SIMCOMP / SUBCOMP - 化合物構造検索  
KCal - 糖鎖構造検索  
PathComp - 可能な反応経路の計算  
PathSearch - 類似反応経路検索  
PathPred - 分解・合成反応経路予測  
E-zyme - 化合物間の酵素反応予測

京都大学化学研究所バイオインフォマティクスセンター

化学構造比較、パスウェイ予測など

2014/8/20 [http://www.genome.jp/ \(英語\)](http://www.genome.jp/)  
[http://www.genome.jp/ja/ \(日本語\)](http://www.genome.jp/ja/)

45

## SIMCOMP/SUBCOMP

- 化合物の類似性検索・部分構造検索ツール
- SIMCOMP
  - データベースへの化合物の類似性検索ツール
- SIMCOMP2
  - 問合せ化合物間の類似性検索
- SUBCOMP
  - 化合物の部分構造検索

SIMCOMP Search

SIMCOMP SUBCOMP SIMCOMP2 API KEGG2

Enter query compound: (in one of the four forms)

Compound ID (Example) C00022 View structure  
MOL File Name 参照... ファイルが選択されていません。

MOL File Text

SMILES

Select target database:

COMPOUND DRUG KNAPSAck REACTION

Search options:

Global search Local search Customized search  
Option details

SIMCOMP help  
SIMCOMP REST API

Download KegDraw  
KegDraw is a Java application for drawing compound structure.

Feedback KEGG GenomeNet Kyoto University Bioinformatics Center

2014/8/20

46

# GenomeNetからの化合物データベース一括検索機能

The screenshot shows two browser windows side-by-side. The left window is the main GenomeNet homepage, and the right window is a search result for the query "Phenylalanine".

**Left Window (GenomeNet Home):**

- Header: GenomeNet, www.genome.jp
- Navigation: KEGG, KEGG2, PATHWAY, BRITE, MEDICUS, DBGET
- Search bar: Search All databases for Phenylalanine
- Resources: KEGG COMPOUND, KEGG GLYCAN, PubChem, ChEBI, PDB-CCD, 3D-MET, LIPIDMAPS, KEGG BRITE, KNapsack, LigandBox, Chemical reaction, KEGG ENZYME, KEGG REACTION, KEGG RPAN, Genome, KEGG GENOME, KEGG MGENOME, varDB, Community Databases, Bioinformatics tools, Feedback.
- Information: GenomeNet Overview, Release notes, Acknowledgments, DBGET Overview, DB release info, KEGG varDB, Community DBs, Bioinformatics tools, Other tools, Feedback.
- Date: 2014/8/20
- Footer: Kyoto University

**Right Window (Search Result):**

- Title: DBGET Search Result: All databases... [Chemical substance]
- Search term: Phenylalanine
- Result count: 1003055 [KegDraw] [Jmol]
- Chemical structure: C<sub>9</sub>H<sub>11</sub>NO<sub>2</sub> (Phenylalanine)
- Text: L-phenylalanine, C<sub>7</sub>-[alpha;beta;-amino-beta;phenyl]propanoic acid
- Chemical formula: C<sub>9</sub>H<sub>11</sub>NO<sub>2</sub>
- SMILES: [3R]-Dihydroxy-L-phenylalanine, L-Dopa, 3-Hydroxyl-tyrosine, L-beta-(3,4-Dihydroxyphenyl)alanine;
- DB growth curve: A line graph showing the growth of the database over time.
- List of results:
  - 103024702 [PubChem] D-Phenylalanine, N-[(phenylmethoxy)carbonyl]-, 4-chloro-2-[[4-chlorophenyl]amino]carbonyl]phenyl ester
  - 103025974 [PubChem] 4-[2-((2S)-2-amino-3-phenylpropyl)ester, L-Phenylalanine, N-[(1,1-dimethyllethoxycarbonyl)phenyl]-, (1R,2S)-2-[(2S)-4-amino-1,4-dioxo-2-[(2-quinolonecarbonylamino)butyl]amino]-1-[(3S,4aS,8aS)-3-[(1,1-dimethyllethyl)amino]carbonyl]octahydro-2(1H,8aH,9aH)-3-phenyl-2-trimethylammonio]propanoate, L-Phenylalanine betaine
  - 103032560 [PubChem] 2-Methylpropyl N-[(2R,5R)-5-6-amino-9H-purin-9-yl)-4-fluoro-2,5-dihydrofuran-2-yl]methyl] (phenoxy)phosphoryl-L-phenylalaninate; Mono(isobutyl-phenylalanine) amidate
  - 103044993 [PubChem] polylysine-co-phenylalanine
- More options: ... > display all
- ChEBI link: CHEBI
- Other links: 41654 [CHEBI], 53490 [KEGG-Link], 2197 [CHEBI], 61957 [CHEBI], 29997 [CHEBI]
- Bottom text: 英語: http://www.genome.jp/ 日本語: http://www.genome.jp/ja/
- Page number: 47

## KEGGのシステム情報、ツール (PATHWAY, BRITE, MODULE, KEGG MAPPER)

# KEGG: Kyoto Encyclopedia of Genes and Genomes

The screenshot shows the KEGG homepage with several highlighted sections:

- PATHWAY**: パスウェイマップ (Pathway Map)
- BRITE**: 機能階層・オントロジー (Functional Hierarchy)
- MODULE**: モジュール (機能ユニット) (Module)
- Mapper**: PATHWAY/BRITE/MODULEへのマッピングツール群 (Mapping Tools)

2014/8/20

49

## KEGG PATHWAY

生体内(外)の分子間ネットワーク図

The screenshot shows the KEGG PATHWAY Database interface with the following sections:

- Pathway Maps**: A collection of manually drawn pathway maps.
- Mapper**: Tools for mapping molecular datasets onto pathway maps.

<http://www.kegg.jp/kegg/pathway.html>

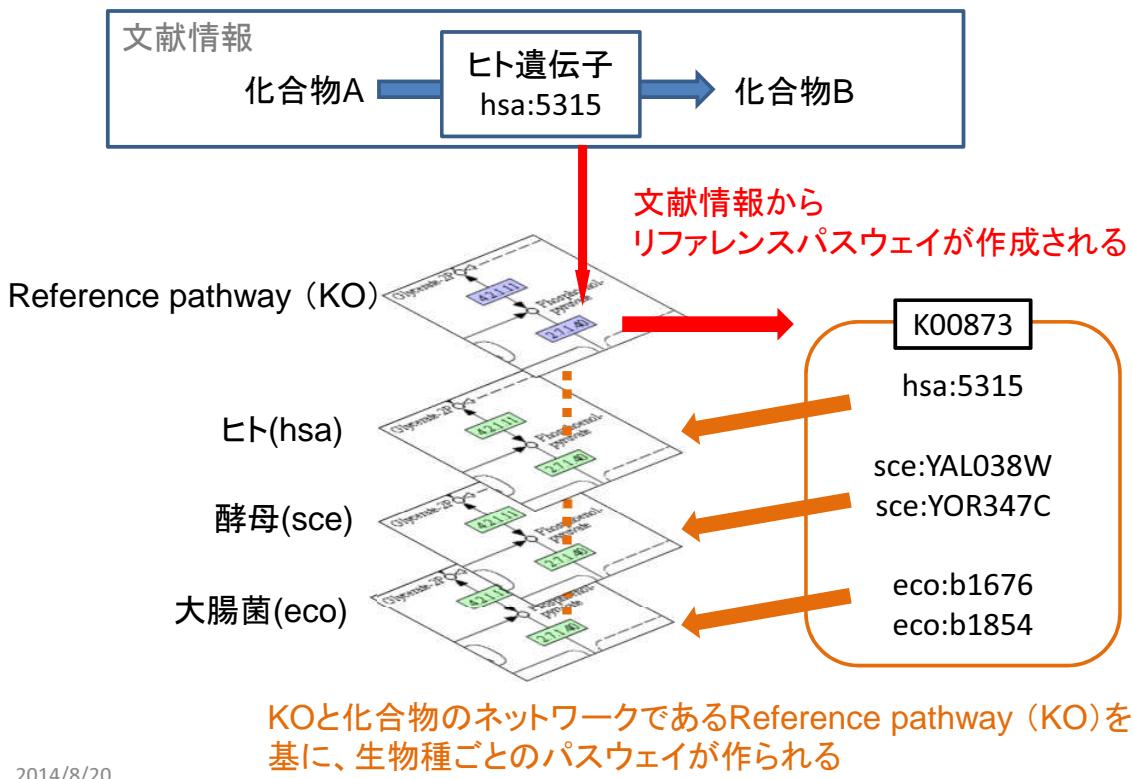
Carbohydrate → Glycolysis / Gluconeogenesis

2014/8/20

- 代謝系(1)
  - 12カテゴリ
  - 中間代謝、二次代謝、薬の代謝、全体像
- 制御系(2-5)
  - 20カテゴリ
  - 遺伝制御、環境シグナル、細胞プロセス、生体システム他
- 疾患(6)
  - がん、免疫・神経変性・循環器・代謝疾患、感染症
- 薬の開発(7)
  - 開発の歴史、標的ベース、構造ベース

50

# KEGG PATHWAY



2014/8/20

51

# KEGG PATHWAY

生体内(外)の分子間ネットワーク図

The screenshot shows the KEGG PATHWAY Database interface. At the top, there's a logo and a navigation bar with links like "PATHWAY", "BRITE", "MODULE", "DISEASE", "DRUG", "KO", "GENOME", "GENES", "LIGAND", and "DBGET". Below the navigation bar, there's a search bar with "Select prefix" dropdown set to "map" and "Enter keywords" input field. To the right of the search bar are "Organism" dropdown and "Go" button. Underneath the search bar, there's a section titled "Pathway Maps" with a brief description of what KEGG PATHWAY is. It lists categories such as "Global Map", "Metabolism", "Genetic Information Processing", "Environmental Information Processing", "Cellular Processes", "Organismal Systems", "Human Diseases", and "Drug Development". There's also a "Pathway Mapping" section with a description of how it maps molecular datasets to KEGG pathway maps.

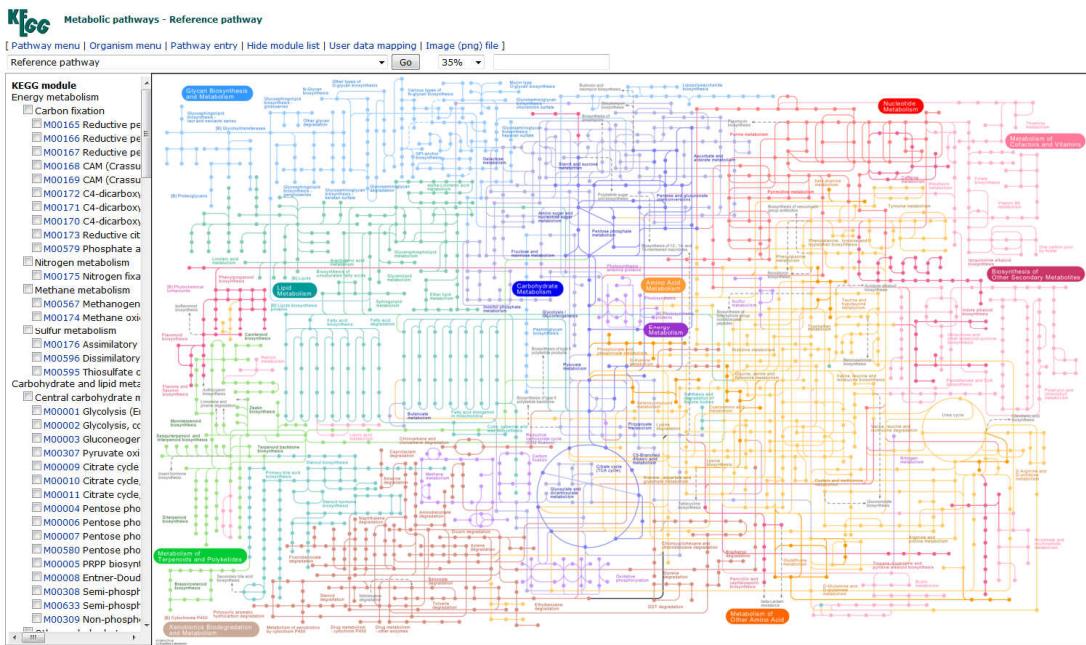
<http://www.kegg.jp/kegg/pathway.html>

2014/8/20

- 代謝系(1)
  - 12カテゴリ
  - 中間代謝、二次代謝、薬の代謝、全体像
- 制御系(2-5)
  - 20カテゴリ
  - 遺伝制御、環境シグナル、細胞プロセス、生体システム他
- 疾患(6)
  - がん、免疫・神経変性・循環器・代謝疾患、感染症
- 薬の開発(7)
  - 開発の歴史、標的ベース、構造ベース

52

# KEGG PATHWAY (Global map)

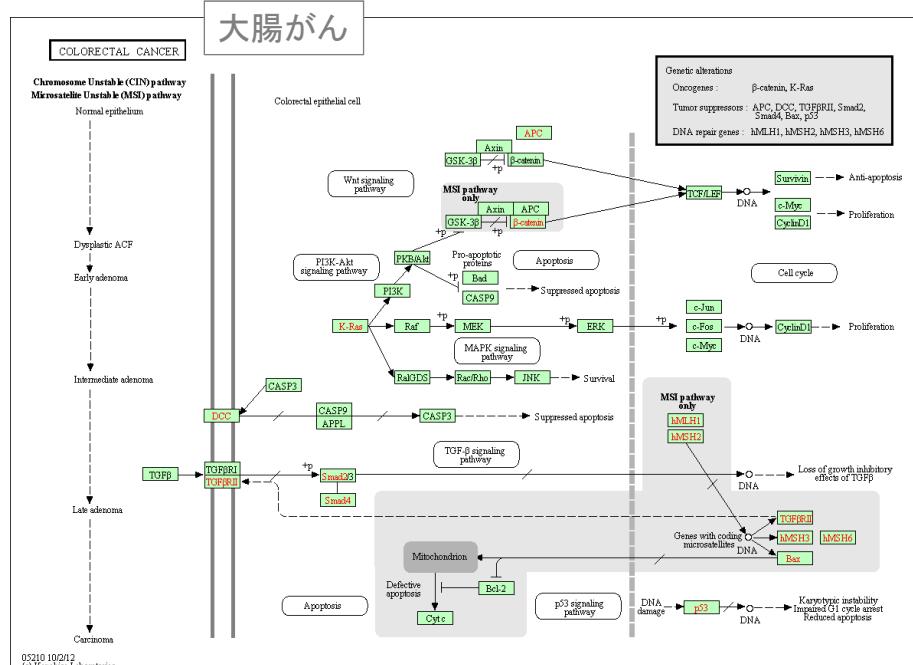


2014/8/20

53

# KEGG PATHWAY (疾患マップ)

病気の進行

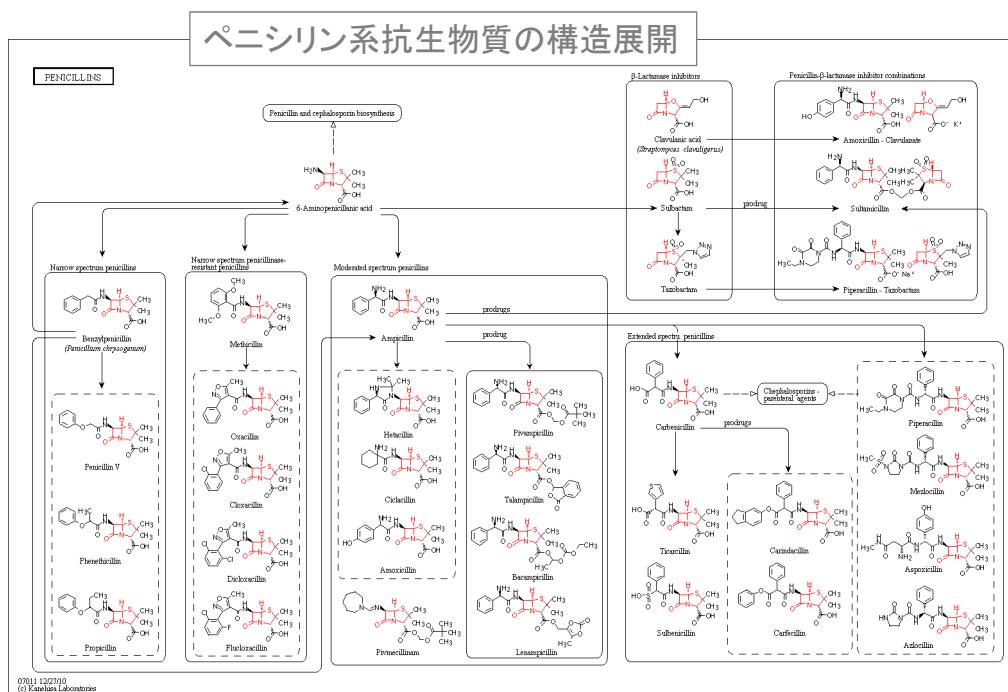


2014/8/20

シグナル伝達

54

# KEGG PATHWAY (薬の開発)

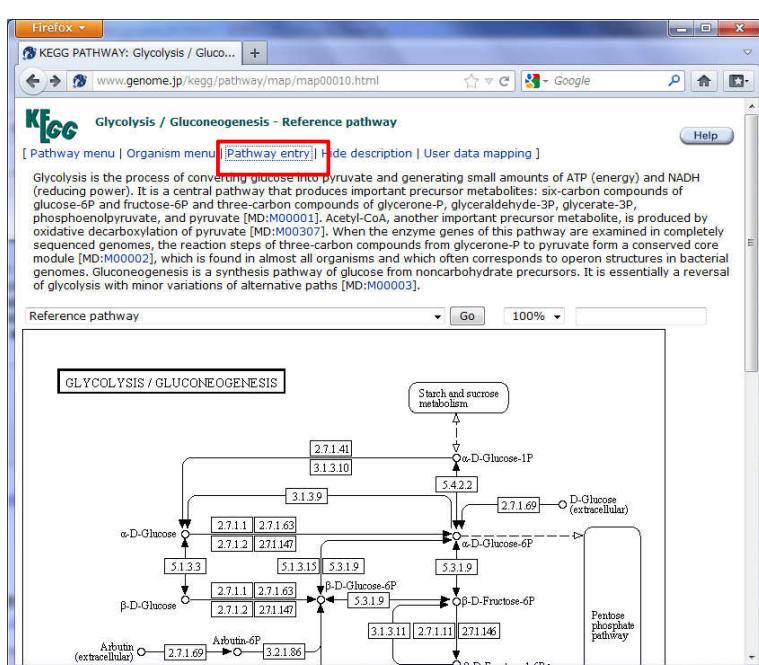


2014/8/20

55

# KEGG PATHWAY

## マップの例: 解糖系



2014/8/20  
<http://www.genome.jp/kegg/pathway/map/map00010.html>

- 酶素／遺伝子と化合物のネットワーク
- Pathway menu
  - BRITE 形式の階層分類
- Organism menu
  - 生物種の階層分類
- Pathway entry
  - パスウェイデータベースのテキストバージョン
- Hide description
  - マップの説明を隠す
- User data mapping
  - マップ中のオブジェクトへの色付け
  - マップの拡大縮小

56

# KEGG PATHWAY

テキストエントリー

- Pathway entry

- Entry
  - マップ番号(map|ko|ec|rn|生物種コード+番号)
- Name, Description
  - パスウェイの説明とモジュールとの関係
- Class
  - 階層分類情報
- Pathway map
  - マップとオーソログテーブルへのリンク
- Disease
- Reference

- モジュール

- 生物種間での保存、複合体、オペロンを考慮した機能単位

[http://www.genome.jp/dbget-bin/www\\_bget?pathway+map00010](http://www.genome.jp/dbget-bin/www_bget?pathway+map00010)

57

# KEGG PATHWAY

マップの例: 解糖系

- Reference pathway

- KO, EC, Reactionへのリンク
- (KO) オーソログエントリーへのリンク
- (EC) 酵素エントリーへのリンク
- (Reaction) 反応エントリーへのリンク

- 生物種名

- 各生物種の遺伝子エントリーへのリンク

- Set personalized menu

- 生物種の選択

- Sort below by

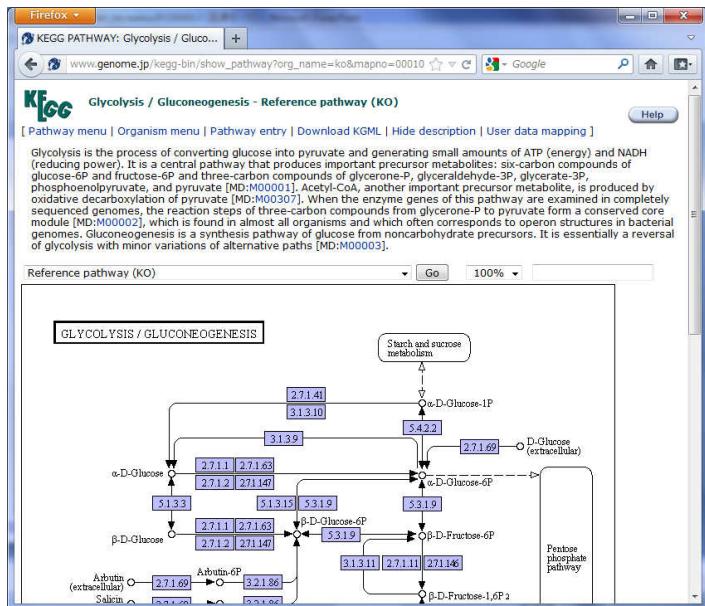
- 生物種名のソート

2014/8/20  
<http://www.genome.jp/kegg/pathway/map/map00010.html>

58

# KEGG PATHWAY

マップの例：解糖系 (EC)

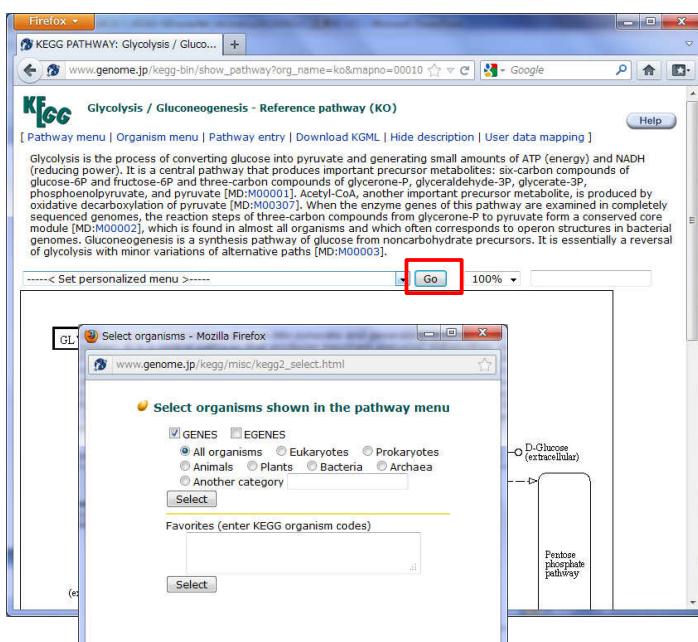


2014/8/20 <http://www.genome.jp/kegg/pathway/ko/ko00010.html>

59

# KEGG PATHWAY

生物種の選択



2014/8/20

[http://www.genome.jp/kegg/catalog/org\\_list.html](http://www.genome.jp/kegg/catalog/org_list.html)

- リストの生物種を限定する

- カテゴリーは KEGG 生物種一覧のものが指定可能
  - Mammals, Protists, Actinobacteria など
- 生物種コードも KEGG 生物種一覧のものが指定可能
  - hsa(ヒト), mmu(マウス), eco(大腸菌)など

60

# PATHWAY の検索とマッピング

## キーワード検索

- Entry, Name, Description フィールドとマップ中のオブジェクト(遺伝子、オーソログ、反応、化合物)や注釈を対象とした検索
- 複数キーワードは AND 検索

## パスウェイマッピング

- マップ中のオブジェクトを指定してパスウェイにマッピング
- 複数オブジェクトを指定するとマッチしたものすべてをマッピング
- Search Pathway
  - 指定したオブジェクトを赤く色づけ
- Search & Color Pathway
  - 指定したオブジェクトを自由に色づけ
- Color Pathway
  - 指定したパスウェイにオブジェクトの色データを与えて色づけ、数値を与えることも可能

2014/8/20

61

## PATHWAY のキーワード検索

2014/8/20

62

# PATHWAY のオブジェクトに好きな色を付ける

The screenshot shows two windows side-by-side. On the left is the 'KEGG PATHWAY Database' window, which includes a search bar for 'Organism' (set to 'hsa') and 'Keywords' (set to 'cancer'), and a 'Pathway Maps' section with a list of categories like Global Map, Metabolism, etc. Below these are links for 'Search Pathway' and 'Color Pathway'. A red box highlights the 'Color Pathway' link. On the right is the 'KEGG Mapper: Color Objects in KEGG Pathways' window, featuring a search bar for 'Reference pathway (KO)' and a text input field where 'KO1803 red,blue' and 'C00118 pink' are entered. A red box highlights this input field. Below the input field are examples of color specifications and a button labeled 'Exec'.

- オブジェクトの色を指定する(htmlで使える色指定ならOK)
- Example をコピペ

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

63

# PATHWAY のオブジェクトに好きな色を付ける (KEGG mapper)

The screenshot shows two windows. On the left is the 'Pathway Search Result' window, listing various metabolic pathways. A red box highlights the entry for 'ko00010 Glycolysis / Gluconeogenesis (2)'. On the right is a detailed 'KEGG PATHWAY: Glycolysis / Gluconeogenesis - Reference pathway (KO)' diagram. This diagram shows the complex network of glycolytic and gluconeogenic reactions, with various enzymes and metabolites labeled with their KEGG IDs. A red box highlights a specific node in the pathway diagram.

2014/8/20

64

# KEGG GENOME

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

The screenshot shows the KEGG GENOME Database homepage. At the top is a logo and the title "KEGG GENOME Database". Below it is a sub-header "Organisms and ecosystems with genome sequence information". A navigation bar includes links for KEGG2, PATHWAY, BRITE, MODULE, KO, GENOME, GENES, SSDB, and Organisms. A search bar at the top allows entering a KEGG organism code or selecting an organism from a dropdown menu. Below the search bar is a table titled "KEGG Organisms" with columns for Category, Genome, Identifiers, Gene catalog, and Annotation. The table lists categories like Complete genomes, Draft genomes, Metagenomes, Pangenomes, and Viruses, each with their respective genome types (GENOME, MGENOME, VGENOME) and identifier details. A search bar below the table allows searching for "Saccharomyces" with options for bfind mode and bgrep mode.

2014/8/20

- KEGG 生物種ごとの入り口

- GENES/DGENES: ゲノムが決定された生物種
- EGENES: ESTで作成された遺伝子セット
- MGENES: メタゲノムデータ
- Pangenomes: 近縁生物種をまとめたもの
- Viruses
- 生物種の組み合わせ

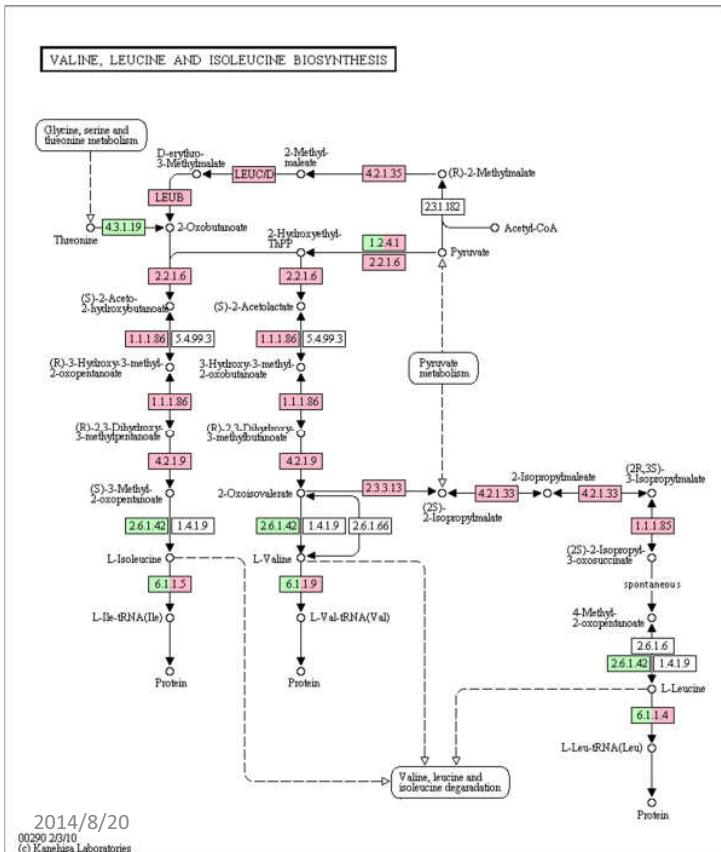
- KEGG Mapper の  
Reconstrcut pathway で実現

This screenshot shows the "KEGG Organism Groups" page. It explains that an organism group is defined as a combination of KEGG organisms for analysis. A search bar at the top allows defining an organism group by entering three-letter organism codes. Below the search bar are examples of different groupings:

- (Metagenome example) hsa+T30003 Homo sapiens (human) + gut microbiome [New!](#)
- (Symbiosis examples)
  - dja+mlo Lotus japonicus (lotus) + Mesorhizobium loti
  - apl+bac Acyrthosiphon pisum (pea aphid) + Buchnera aphidicola
  - bmy+wbm Brugia malayi (flaria) + Wolbachia
- (Pathogenesis examples)
  - hsa+pfa Homo sapiens (human) + Plasmodium falciparum
  - aga+pfa Anopheles gambiae (mosquito) + Plasmodium falciparum

65

## 複数生物種の情報を PATHWAY にマッピング

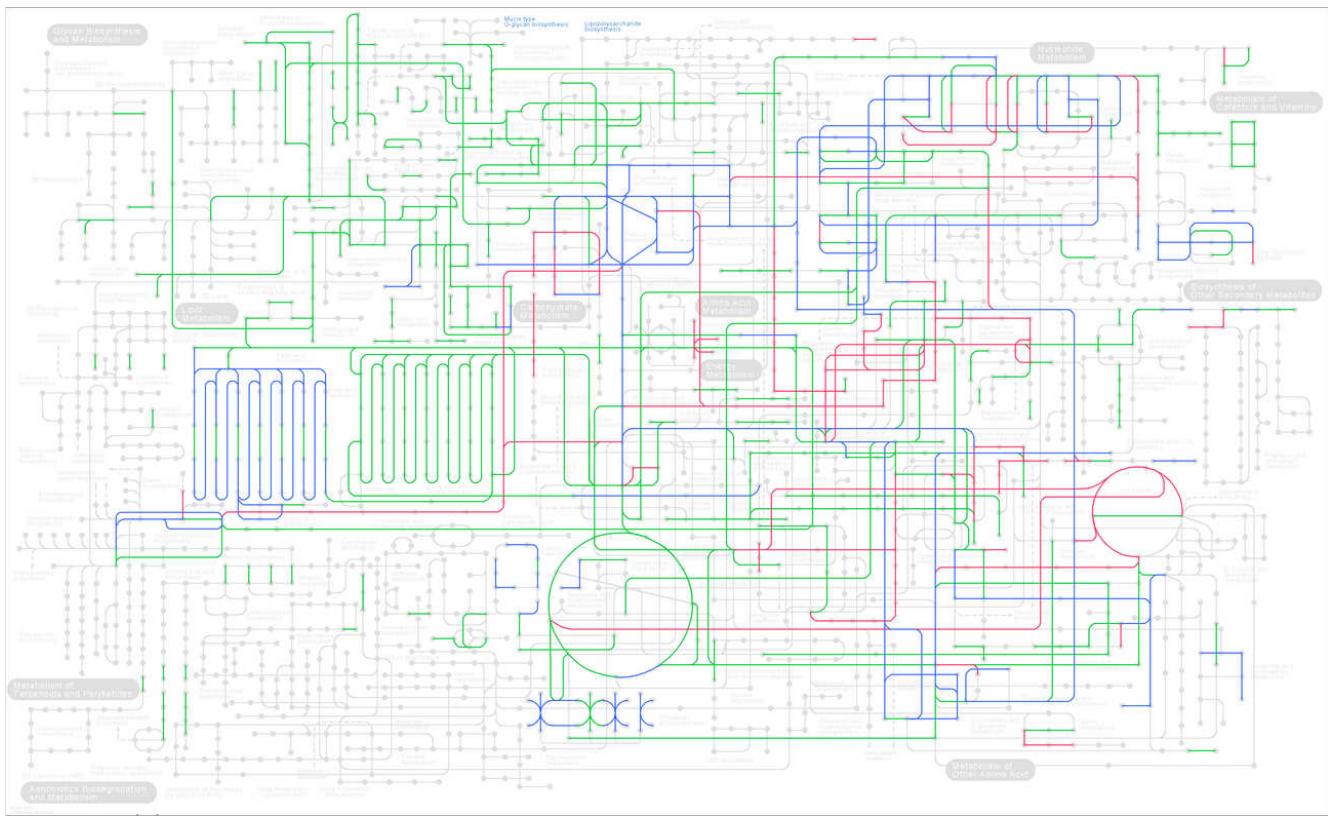


- 共生関係
- 寄生関係
- ヒトと腸内細菌叢など

- アブラムシ  
– 昆虫

- ブフェラ  
– アブラムシの共生細菌

# 複数生物種の情報を PATHWAY にマッピング



67

## KEGG MODULE

- マニュアルで定義された生物学的な機能ユニットのコレクション -

KEGG MODULE Database  
Functional units for annotating and interpreting genomes

KEGG2 PATHWAY BRITE MODULE DISEASE DRUG KO GENOME GENES LIGAND DBGET

**KEGG Modules**

KEGG MODULE is a collection of manually defined functional units, termed KEGG modules, used for annotation and biological interpretation of sequenced genomes. Each module is identified by the M number that corresponds to a set of K numbers (KO identifiers). There are four types of KEGG modules:

- pathway modules – representing tight functional units in KEGG metabolic pathway maps, such as M0002 (citrate), often a module involving three-carbon compounds
- structural complexes – often forming molecular machineries, such as M00072 (Oligosaccharyltransferase)
- functional sets – for other types of essential sets, such as M00360 (Aminocarbonyl-tRNA synthases, prokaryotes)
- signature modules – as markers of phenotypes, such as M00363 (EHEC pathogenicity signature, Shiga toxin)

The entire list of KEGG modules can be viewed from the BRITE hierarchy file:  
KEGG modules

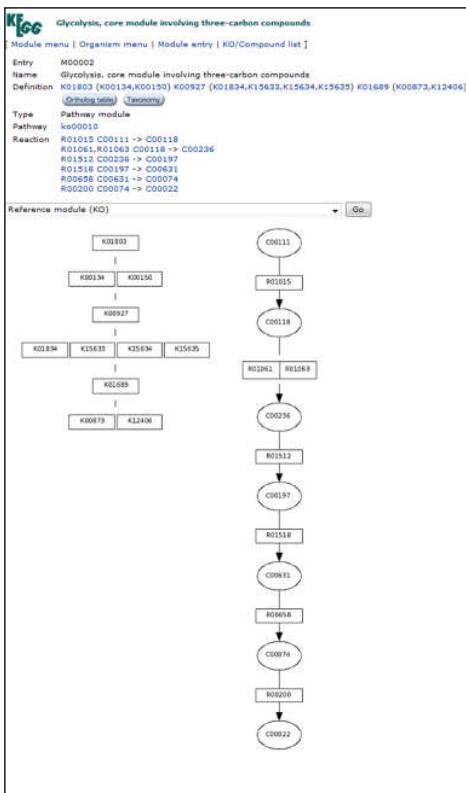
Each M number entry is defined by the Boolean-like expression for the combination of K numbers. Notations for this definition include:

- space delimited items for pathway elements
- comma separated items in parentheses for alternatives
- plus sign to define a complex
- minus sign for an optional item

- Pathway modules
  - KEGG PATHWAYの代謝パスウェイで、緊密な機能単位としてなる連続経路
- Structural complexes
  - 分子
- Functional sets
  - その他の必須な機能単位
- Signature module
  - 表現型のマーカーなど

# KEGG MODULE

## パスウェイモジュールの例



2014/8/20

69

- Entry

-ID

- Name

-名前

- Definition

-モジュールの構造

- Type

-モジュールのタイプ

- Pathway

-モジュールが存在するパスウェイ

- Reaction

-モジュールを構成する反応

## KEGG Mapper

**KEGG Mapper**

KEGG Mapper is a collection of tools for KEGG mapping: KEGG pathway mapping, BRITE mapping, and MODULE mapping. Two pathway mapping tools, "Search Pathway" and "Color Pathway", were made available from the beginning of the KEGG project. The naming of these tools was somewhat misleading since both involved searching and coloring procedures. The latter is now called "Search&Color Pathway", and a new "Color Pathway" is introduced for coloring of a selected pathway map (without searching).

**About KEGG Mapper**

**Pathway mapping tools**

- Search Pathway** - basic pathway mapping tool
  - Search against all pathway maps in one category (ko, org, etc.)
  - Multiple search objects may be specified in one line
  - Mapped objects are marked in red
- Search&Color Pathway** - advanced pathway mapping tool
  - Search against all pathway maps in one category (ko, org, etc.)
  - Each search object is specified in one line together with color attributes (background color and foreground color)
  - Mapped objects can be marked in any color
  - NCBI and UniProt identifiers may be used
- Color Pathway** - selected pathway map coloring tool
  - One selected pathway map is colored in multiple ways
  - Each object is specified in one line together with color attributes, i.e., multiple background and foreground colors in multiple columns
  - Results shown as multiple snapshots of image maps
  - When numerical values are given as attributes, they are displayed by gradation and optionally by 3D representation as well
- Reconstruct Pathway** - pathway mapping tool to assist genome annotation
  - Search against all pathway maps in the ko category
  - Search objects are user-defined genes annotated with K numbers
  - Mapped objects are marked in green for a single genome or in multiple colors (see: [KEGG color codes](#)) for multiple genomes

• PATHWAY/BRITE/MODULEのオブジェクトを探索、色づけ等するツール群

- KEGG Object (+ 色、あるいは数値)を指定してパスウェイにマッピング
- 目的に合わせて下記の11種類のツールがある

### Concepts behind KEGG Mapper

KEGG PATHWAY and KEGG BRITE are the reference knowledge bases for biological interpretation of molecular datasets, especially large-scale datasets generated by high-throughput experimental technologies. KEGG Mapper will help to understand higher-level systemic functions, based on the set operation between the query data and the reference knowledge as shown below.

Query data	Reference knowledge	Tool
Objects	KEGG PATHWAY database	Search Pathway
Object-attributes relations <sup>1</sup>	KEGG PATHWAY database	Search&Color Pathway
Object-attributes relations <sup>2</sup>	Single KEGG pathway map	Color Pathway / Color Pathway 3D
Gene-KO list	KEGG PATHWAY database (KO)	Reconstruct Pathway
Objects	KEGG BRITE database	Search Brite
Object-attributes relations <sup>1</sup>	KEGG BRITE database	Search&Color Brite
Object-attributes relations <sup>3</sup>	Single Brite functional hierarchy	Join Brite
Gene-KO list	KEGG BRITE database (KO)	Reconstruct Brite
Objects	KEGG MODULE database	Search Module
Object-attributes relations <sup>1</sup>	KEGG MODULE database	Search&Color Module
Gene-KO list	KEGG MODULE database (KO)	Reconstruct Module

Attributes are: 1 colors, 2 colors or numerical values, and 3 any attributes.

- The query data may be a collection of molecular objects (genes, proteins, small molecules, etc.) or a more ordered set of object-attributes relations, such as:
- orthologs and their presence in the genome (green boxes in KEGG pathways)
  - genes associated with up/down expression levels (green/yellow/red coloring)
  - genes associated with somatic mutation frequency (3D representation)

2014/8/20

70

# KEGG Mapper

- Search Pathway/Brite/Module
  - 基本的なマッピングツール
  - KEGG Objectを入力して、Pathwayなどにマッピングする
- Search&Color Pathway/Brite/Module
  - より高度なマッピングツール
  - KEGG Objectとbgcolor, fgcolorの色を入力として、Pathwayなどの任意のObjectに指定色をマッピングする
- Color Pathway
  - Search&Color Pathway の拡張
  - KEGG Objectと数値データを入力として、Pathwayマップにカラースケールの色を塗り分けることができる
- Reconstruct Pathway/Brite/Module
  - 一つあるいは複数の遺伝子とKOの対応リストを入力に、Pathwayなどにマッピングしてパスウェイの再構築などを行うことができる
- Join Brite
  - たとえば、薬とそのターゲットのような二項関係のリストを、BRITEの階層ファイルにマッピングして結合するツール

2014/8/20

71

## Search Pathway

**KEGG Mapper – Search Pathway**

Search against: hsa Enter: map, ko, ec, rn, hsadd, or org.

Enter objects:

hsa:7167 hsa:GPI cpd:C00118

Examples:  
(Reference pathway (KO))  
ko:K01803 cpd:C00111 cpd:C00118  
K00134 C00236  
(Homo sapiens pathway)  
hsa:7167 hsa:GPI cpd:C00118  
ALDOA 1.2.1.12 C00236

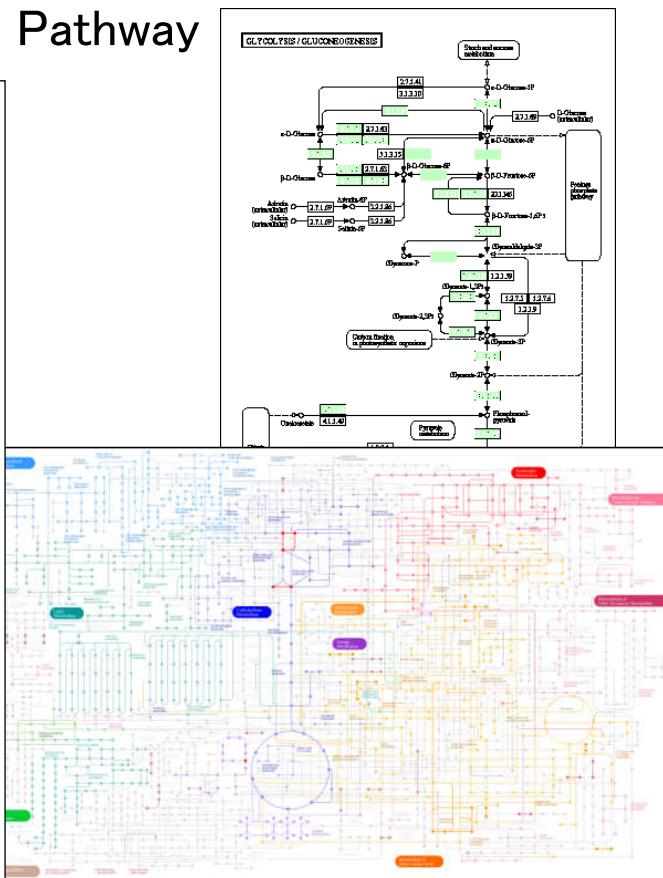
Alternatively, enter the file name containing the data:

Filter1 Filter2 (to extract K/C/G/D/R/RP/RC numbers)

Include aliases  
 Display objects not found in the search  
 Search pathways containing all the objects (AND search)

**Search Pathway** is the basic pathway mapping tool, where given objects (genes, proteins, compounds, glycans, reactions, drugs, etc.) are searched against KEGG pathway maps and found objects are marked in red. The objects in different types of pathway maps are specified by the following KEGG identifiers and aliases.

Prefix Type	KEGG identifier	Alias
map	Reference pathway - metabolic	K/R/EC numbers, C/G/D numbers KO alias
map	Reference pathway - non-metabolic	K number, C/G/D numbers
ko	Reference pathway (KO)	K number, C/G/D numbers
ec	Reference pathway (EC)	EC number, C/G/D numbers
rn	Reference pathway (Reaction)	R number, C/G/D numbers
org	Organism-specific pathway	gene identifier, C/G/D numbers



2014/8/20

72

# Reconstruct Pathway

**KEGG Mapper – Reconstruct Pathway**

Enter gene list with KO annotation:

```
gene16315 K12761
gene16346 K07375
gene16354 K05692
gene16358 K03283
gene16434 K01739
gene16495 K11864
gene16514 K04336
gene16660 K00025
gene16664 K03061
gene16732 K01251
```

Example files:  
genelist.txt (single organism)  
genelist2.txt (multiple organisms)

Alternatively, enter the file name containing the data:

Include global map

**Reconstruct Pathway** is a KEGG PATHWAY mapping tool that assists genome and metagenome annotations. The input data is a single gene list (for a single organism) or multiple gene lists (for multiple organisms) annotated with KEGG Orthology (KO) identifiers or K numbers. Each line of the gene list contains the user-defined gene identifier followed by, if any, the assigned K number. The mapping is performed through the K numbers against the KEGG reference pathways.

When multiple organisms are given in the input file, each organism must be separated by the # line, which may contain user-defined color specification. Default coloring for two organisms is as follows:

org 1	2	1+2
regular map	green	red
global map	green	blue

See KEGG color codes for more details.

Last updated: June 1, 2012

2014/8/20

73

# Search&Color Pathway

**KEGG Mapper – Search&Color Pathway**

Search against:  Enter: map, ko, ec, r, msadd, or

Primary ID: KEGG identifiers  (NCBI and UniProt IDs may be used for organism-specific pathways)

Enter objects one per line followed by bgcolor, fgcolor:

```
7167 red,blue
C00118 pink
```

Examples:  
(Reference pathway (KO))  
K016315 red,blue  
C00118 pink  
  
(Homo sapiens pathway)  
7167 red,blue  
C00118 pink

Alternatively, enter the file name containing the data:

If necessary, change default bgcolor:

Include aliases  
 Use uncolored diagrams  
 Display objects not found in the search  
 Search pathways containing all the objects (AND search)

**Search&Color Pathway** is an advanced version of the KEGG pathway mapping tool, where give objects (genes, proteins, compounds, glycans, reactions, drugs, etc.) are searched against KEG pathway maps and found objects are marked in any background and foreground colors (bgcolor fgcolor). The objects in different types of pathway maps are specified by the following KEGG identifiers and aliases.

Prefix Type	KEGG identifier	Alias
map	Reference pathway - metabolic	K/REC numbers, C/G/D numbers
map	Reference pathway - non-metabolic	K number, C/G/D numbers
ko	Reference pathway (KO)	K number, C/G/D numbers
ec	Reference pathway (EC)	EC number, C/G/D numbers
rm	Reference pathway (Reaction)	R number, C/G/D numbers
		RP/RC numbers

2014/8/20

74

# Color Pathway

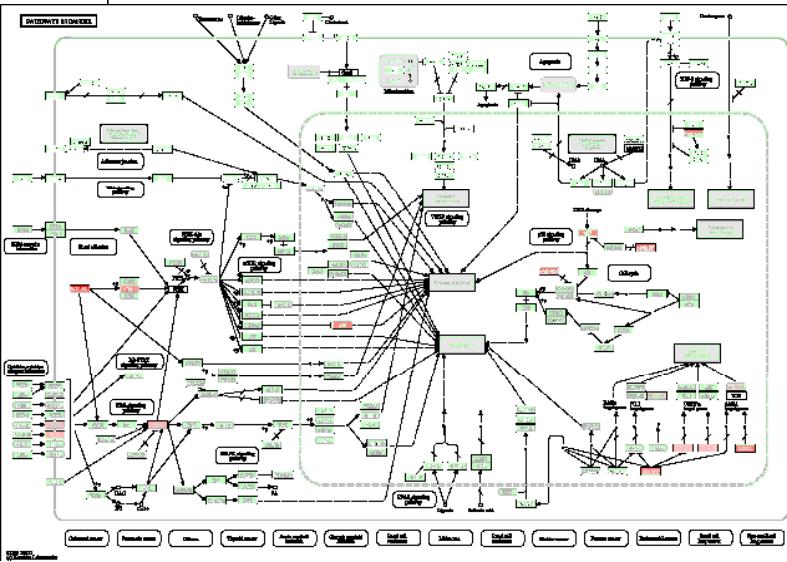
**KEGG Mapper – Color Pathway**

Select KEGG pathway map: hsa05200 Enter pathway map identifier, such as hsa05200  
 Enter file name containing the data: D:\Users\Takimatsu\Downloads\White\_CM [参照]  
 File type:  Color specification  
 Numerical value ( Convert to log scale)  
 Gradation: from #ffffff to #ff0000 Check  
 Use uncolored diagram  
 Exec Clear

**Color Pathway** is an extension of the Search&Color Pathway tool, allowing multiple colored pathway map. Multiple coloring of bcolor (or bcolor[fcolor]), which is specific tab-delimited column, may represent time-dependent or site-dependent changes of gene expressions, pathways active in different cancer stages, etc., and is shown as multiple image maps. Alternatively, numerical values may be given in multiple columns and they by gradation and optionally 3D representation as well. The objects in different types of maps are specified by the following KEGG identifiers and aliases.

Prefix	Type	KEGG identifier	Alias
map	Reference pathway - metabolic	K/R/EC numbers, C/G/D numbers	KO alias
map	Reference pathway - non-metabolic	K number, C/G/D numbers	KO alias, E
ko	Reference pathway (KO)	K number, C/G/D numbers	KO alias, E
ec	Reference pathway (EC)	EC number, C/G/D numbers	
rn	Reference pathway (Reaction)	R number, C/G/D numbers	RP/RC num
org	Organism-specific pathway	gene identifier, C/G/D numbers	gene alias, K/EC num

Last updated: May 1, 2013



2014/8/20 75

# FeedBack

**GenomeNet**



**GenomeNet Feedback Form**

Please use the form below to send your question or comment to KEGG and GenomeNet. Please enter information in all fields. (Japanese characters accepted)

Name:  Email address:   
 (a copy will be sent to you as well)

Affiliation:

Category:

<input checked="" type="radio"/> KEGG PATHWAY	<input type="radio"/> KEGG MEDICUS	<input type="radio"/> Copyright permission
<input type="radio"/> KEGG BRITE	<input type="radio"/> KEGG API	<input type="radio"/> KEGG license
<input type="radio"/> KEGG MODULE	<input type="radio"/> KAAS	<input type="radio"/> Other
<input type="radio"/> KEGG GENES	<input type="radio"/> GenomeNettools	
<input type="radio"/> KEGG LIGAND	<input type="radio"/> DBGET/LinkDB	

Comments:

Kyoto University Bioinformatics Center

KEGGデータベースや  
 GenomeNetに関する質問や  
 コメントなどを受け付けて  
 います

日本語でOKです

## メタボロミクス関連データベースの紹介

2014/8/20

77

## メタボロミクスで利用される データベースの種類

- 代謝物質リファレンススペクトルデータベース
  - 代謝産物を同定するためのリファレンスデータベース
- 代謝プロファイル・代謝産物データベース
  - 特定生物種についての代謝プロファイル情報
  - 多くの生物種の様々な状態での代謝プロファイル
  - 各生物種について既知の代謝物質をリスト化したもの
- 代謝パスウェイデータベース
  - 特定生物種に関する代謝パスウェイデータベース
  - リファレンス代謝パスウェイデータベース
- などなど

2014/8/20

78

# 主な代謝物質リファレンススペクトル データベース

- NIST 14 (<http://www.nist.gov/srd/nist1a.cfm>)
  - 化合物のEI/MS、MS/MSスペクトル、GC の分析パラメータ・Retention Index ライブライ
- GMD [Golem Metabolome Database] (, <http://gmd.mpimp-golm.mpg.de/>)
  - 代謝産物のGC-MSのデータベース
- Metlin (<http://metlin.scripps.edu/index.php>)
  - 代謝産物と MS/MSスペクトルのデータベース
- MassBank (Public, <http://www.massbank.jp/>)
  - 分散型の高精度マススペクトルデータベース
  - 日本質量分析学会公式データベース

2014/8/20

79

# 主な代謝プロファイル/代謝産物 データベース(1)

- HMDB [Human Metabolome Database] (Public, <http://www.hmdb.ca/>)
  - ヒトの身体に関する代謝産物のデータベース
  - スペクトルデータや代謝経路情報に関する情報を含む
- DrugBank (Public, <http://www.drugbank.ca/>)
  - HMDBの関連データベース
  - 医薬品と医薬品のターゲット情報に関するデータベース
- LipidMaps (Public, <http://www.lipidmaps.org/>)
  - 米国脂質学会による、脂質化合物に関する化合物情報、パスウェイ情報などを扱う Lipidomicsのデータベース
- Lipidbank (Public, <http://lipidbank.jp/>)
  - 日本脂質学会による、脂質化合物のデータベース
- PubChem (Public, <http://pubchem.ncbi.nlm.nih.gov/>)
  - NCBIによる低分子化合物の構造、生理活性情報など

2014/8/20

80

# 主な代謝プロファイル/代謝産物 データベース(2)

- KEGG LIGAND (FTP有償, <http://www.genome.jp/kegg/ligand.html>)
  - 代謝産物を中心とした化合物、代謝反応のデータベース
- Dictionary of Natural Products (商用)
  - CRC Pressによる、25万種以上の天然物情報の辞書
  - 化学情報協会 [JAICI, <http://www.jaici.or.jp/>] 取扱
- KNAPSAcK Family (Public, [http://kanaya.naist.jp/KNAPSAcK\\_Family/](http://kanaya.naist.jp/KNAPSAcK_Family/))
  - 奈良先端大金谷研究室による、天然物と由来生物種、食品、生理活性などを関連付けたデータベース
- MarinLit (商用, <http://pubs.rsc.org/marinlit/>)
  - University of Canterbury, New Zealandで作られ英國化学会(RSC)が提供する、海洋天然化合物のデータベース

# 主なパスウェイデータベース・ツール (1)

- KEGG PATHWAY (FTP有償, <http://www.genome.jp/kegg/pathway.html>)
  - 代謝、制御系、ヒト疾患、医薬品開発に関する知識を集めたパスウェイマップ
- Reactome (Public, <http://www.reactome.org/>)
  - ヒトを中心に代謝系、シグナル伝達系に関する知識を集めたパスウェイマップ
- BioCyc (Public, <http://biocyc.org/>)
  - 3,563生物種のパスウェイ／ゲノムデータベース
- PlantCyc (Public, <http://www.plantcyc.org/>)
  - BioCyc のシステムを用いて作られた、17種の植物のパスウェイ／ゲノムデータベース

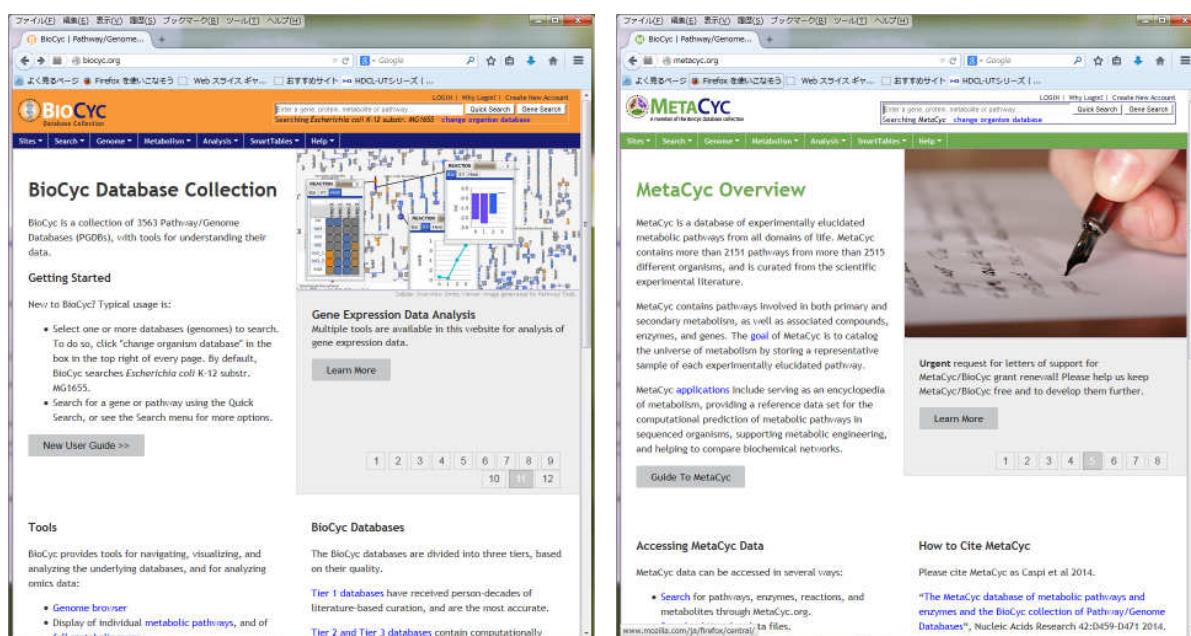
# 主なパスウェイデータベース・ツール (2)

- Pathway Interaction Database (PID, <http://pid.nci.nih.gov/>)
  - ヒトのシグナル伝達系パスウェイのデータベース
- Ingenuity Pathway Analysis (商用, <http://www.ingenuity.com/>)
  - Qiagen系列の世界的に著名なマイクロアレイ、メタボロミクス、プロテオミクス、RNA-Seqなどの実験データをもとにしたパスウェイ解析、などを行えるツール
  - トミーデジタルバイオロジー取扱
- ERGO (商用, <https://ergo.integratedgenomics.com/>)
  - 商用の*in silico* systems biology platform

2014/8/20

83

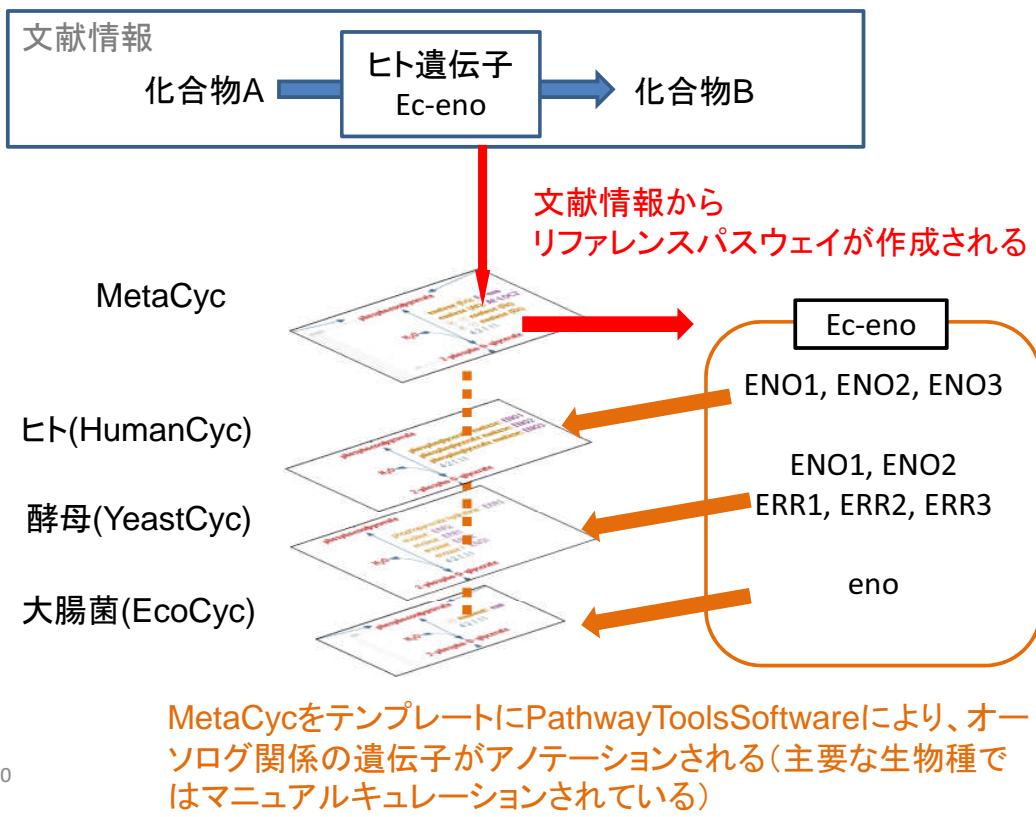
## BioCyc (<http://biocyc.org/>) MetaCyc (<http://metacyc.org/>)



2014/8/20

84

## BioCyc (<http://biocyc.org/>) MetaCyc (<http://metacyc.org/>)



2014/8/20

85

# KEGG PATHWAY/MODULE と BioCyc/MetaCyc

## KEGG PATHWAY

- Reference Pathway



## BioCyc

- MetaCyc

- 個々の生物種マップ



- 個々のPGDB (\*\*Cyc)

- PATHWAY



- Superpathway

- MODULE



- Pathway

2014/8/20

86

# NBDC 研究開発成果データベースリンクにある メタボローム関連データベース (1)

<http://biosciencedbc.jp/db-link/d02-dblink>

メタボローム関連データベース	
DB名	概要
<a href="#">Bio-MassBank</a>	本DBは、植物の組織や微生物試料をLC-, GC-, CE-MSで分析して得られたマススペクトルを代謝物を同定できた、できないにかかわらず収集します。未同定代謝物のマススペクトルをその化学構造を表現する化学的descriptorとして利用することによって、異なる試料間で同じあるいは類似したマススペクトルがあれば同じ未同定代謝物が含まれている、と考えることができます。このように代謝物を同定することができなくても、その存在を知ることができます。現在、シロイヌナズナの葉、ミヤコグサの花をLC-MS, MS2分析したデータそれぞれ664件、636件を公開しています。
<a href="#">KNApSAcK Core DB</a>	本DBは、文献情報をもとに生物種代謝物の関係を収集し、公開しています。現在までに、10万種の生物種-代謝物関係が蓄積されており、メタボロミクス研究の標準データベースとなっています。
<a href="#">KomicMarket</a>	本DBは、化合物ピークのアノテーションを行った質量分析データ（主にLC-MS）を蓄積・公開するためのDBです。
<a href="#">MassBank</a>	本DBは、代謝物質あるいはそれらの関連物質をEI-, FAB-, ESI-, MALDI-MS, MS2などを用いて分析したマススペクトルを収集、公開しています。データの公開状況（2012年3月末）は次のとおりです。21研究グループ（日本15、米国3、ドイツ2、中国1）が13,534化合物について分析した合計29,644マススペクトルデータを9つのデータサーバから公開しています。それらのうち、ESI-MS、MS2データは2,304化合物について測定した16,440件です。

# NBDC 研究開発成果データベースリンクにある メタボローム関連データベース (2)

<http://biosciencedbc.jp/db-link/d02-dblink>

<a href="#">MassBase</a>	本DBは、主に質量分析の未加工データ（生データ）、生データをテキスト型式に変換したデータ、同定・推定（アノテーション）を行わないピークデータを、蓄積・公開するためのDBです。
<a href="#">Metabolome Activity DB</a>	本DBは、文献情報をもとに代謝物の活性情報を収集し、公開しています。現在までに、5千種の代謝物-活性の関係が蓄積されており、メタボロミクス研究の標準データベースとなっています。
<a href="#">metabolomics.jp</a>	本DBは、メタボロミクスを中心とした有田研究室の活動全般を対象としたポータルサイトです。フラボノイド、基礎代謝物、生薬、植物系統分類データベースのほか、ファイトレメディエーションを含めた放射線情報、講義資料も掲載しています。
<a href="#">MFSearcher</a>	本DBは、精密質量分析で得られた精密質量値から、組成式を迅速に推定するためのウェブサービスです。また、KNApSAcK, KEGG, PubChemへ精密質量値からの検索も高速に行うことができます。ハイスクレーブットなピークアノテーションに利用されています。
<a href="#">MS-MS FragmentViewer</a>	本DBは、フラボノイド標品のMS/MSフラグメントを化学構造に帰属させ、分子開裂モデルを提示しているデータベースです。化合物ピークのアノテーションに利用されています。