

統合データベース講習会：AJACS十勝
2014年9月12日

KEGG PATHWAYを中心とした パスウェイデータベースの紹介

情報・システム研究機構 (ROIS)
ライフサイエンス統合データベースセンター (DBCLS)/
科学技術振興機構 (JST)
バイオサイエンスデータベースセンター (NBDC)

河野 信

1

©2014 統合データベース講習会 Licensed Under CC 表示 2.1

GenomeNet

- ◆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を中心とするDB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
- ◆DBGET/LinkDB

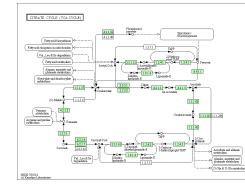
The screenshot shows the GenomeNet homepage with a search bar for "統合データベース" and a dropdown menu for "for". Below the search bar are links to KEGG, DBGET, BioSystems, and other resources. A central panel displays a network diagram of biological pathways and a link to "データベース増加図". On the left, there's a sidebar with links to BLAST, FASTA, and various sequence analysis tools.

©2014 統合データベース講習会 Licensed Under CC 表示 2.1

本日紹介するデータベース (DB)

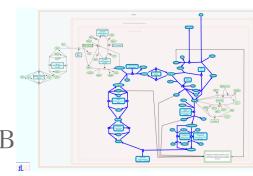
◆ KEGG (Kyoto Encyclopedia of Genes and Genomes)

- KEGGとGenomeNetの概要
- ゲノム情報（配列情報）
- ケミカル情報（化合物と反応データ）
- システム情報（パスウェイと階層分類）



◆ Reactome

- ヒトを中心としたパスウェイDB



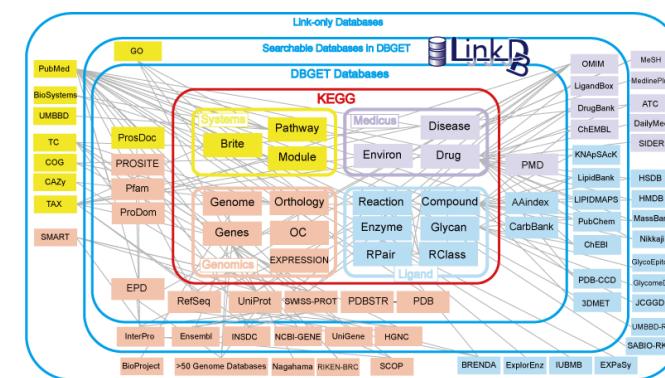
◆ BioCyc

- さまざまな生物のパスウェイを整備したDB

©2014 統合データベース講習会 Licensed Under CC 表示 2.1

LinkDB : <http://www.genome.jp/linkdb/>

- ◆KEGGを核として、国内外のさまざまなDBとリンク



4

©2014 統合データベース講習会 Licensed Under CC 表示 2.1

演習1: IDの対応関係の一覧表を取得する

- ◆KEGG GENESデータベースとUniProtデータベースのヒトのID対応表をダウンロードしてみましょう

- ・"GenomeNet"などで検索してページを開きます
- ・"Genes"の箱をクリックします⇒リンク情報を取得可能なデータベースがハイライトされます
- ・"UniProt"の箱をクリックします⇒図の下に検索条件がセットされます
- ・KEGGでのヒトの生物種コードは"hsa"です⇒生物種を変更するときは"Organism"をクリックして目的の生物種を検索
- ・"Download"でダウンロード⇒メモ帳などで開けます

The screenshot shows the KEGG LinkDB interface. At the top, there's a search bar for 'Searchable Database in DBGET'. Below it is a grid of KEGG databases: Pathway, Brite, Module, Disease, Genes, and others like Prokaryote, PDB, and Drug. A green box highlights the 'Genes' and 'UniProt' boxes. Below the grid, there are three download buttons: 'Click database names to download link information between two databases', 'Download equivalent link information for a selected organism', and 'Download equivalent link information for COMPOUND, DRUG and REACTION'. The URL at the bottom is <http://www.genome.jp/dbget-bin/www-kegg-linkdb?hsa+UniProt>.

©2014 統合データベース講習会 Licensed Under CC 表示 2.1

KEGG

Kyoto Encyclopedia of Genes and Genomes

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

6

©2014 統合データベース講習会 Licensed Under CC 表示 2.1

KEGG

- ◆ゲノムや分子レベルの情報から、細胞・個体・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DBリソース

- ◆遺伝子や化合物などの分子部品のDBと、それらをつなぐネットワークのDBからなっている

The screenshot shows the main KEGG homepage. It features a navigation bar with links like 'KEGG Home', 'KEGG Database', 'KEGG Objects', 'KEGG FTP', 'KEGG Net', 'DBGET/LinkDB', 'Feedback', and 'Kanehisa Labs'. The central content area has sections for 'KEGG: Kyoto Encyclopedia of Genes and Genomes' (with a brief description), 'KEGG PATHWAY' (with a pathway diagram), 'KEGG GENOME' (with a genome map), 'KEGG GENES' (with a gene map), 'KEGG COMPOUND' (with a compound map), 'KEGG REACTION' (with a reaction map), 'KEGG DISEASE' (with a disease map), 'KEGG DRUG' (with a drug map), 'KEGG MEDICUS' (with a medical map), and 'KEGG ORTHOLOGY' (with an orthology map). There are also links for 'BlastKOALA' and 'KEGG Annotations'.

©2014 統合データベース講習会 Licensed Under CC 表示 2.1

KEGGとは？



さまざまな種類のデータを
リンクすることによって
生命現象を再構築

研究者の知識をゲノム・分子レベルのデータと結びつける

8

©2014 統合データベース講習会 Licensed Under CC 表示 2.1 © 2012 時松敏明 (京都大学化学研究所) licensed under CC表示2.1日本

KEGGトップページ <http://www.kegg.jp/>

KEGG Kyoto Encyclopedia of Genes and Genomes

KEGG 分子レベルの遺伝子から細胞、個体、エコシステムといった三次元生命システムへのゲノム学や地図を構築するためのリソースです。ご意見をよろしくお聞かしください。
KEGG の概要
Searching KEGG
KEGG Map
Color codes

KEGG Objects
Pathway maps
Brite hierarchies

KEGG Software
KeggTools
KEGG API
KGML
KEGG FTP
リモート接続

ゲノムストック
DBGET/LinkDB
Feedback

Kanehisa Labs

KEGG Home
Release notes
Current statistics
Plea from KEGG
KEGG Database
KEGG の概要
Searching KEGG
KEGG Map
Color codes

KEGG Objects
Pathway maps
Brite hierarchies

KEGG Software
KeggTools
KEGG API
KGML
KEGG FTP
リモート接続

ゲノムストック
DBGET/LinkDB
Feedback

KEGG: Kyoto Encyclopedia of Genes and Genomes

KEGG の主な機能

KEGG2 PATHWAY BRITE MODULE KO GENOME GENES LIGAND DISEASE DRUG DBGET

Search : KEGG t for [Go] [Clear]

Category Entry Point Search & Compute DBGET Search

System Information
KEGG PATHWAY Search Pathway PATHWAY
KEGG MODULE Search Brite BRITE
KEGG MAPPER Search Taxonomy MODULE
KEGG ATLAS BlastKOALA New!
BlastKOALA サポート マツマノアマゾン サービスを開始しました。 more ...

Genomic information
KEGG ORTHOLOGY KO system, New! ORTHOLOGY
KEGG Annotation Search Pathway PATHWAY
KEGG MODULE Search Brite BRITE
KEGG MAPPER Search Taxonomy MODULE
KEGG ATLAS BlastKOALA New!
BlastKOALA サポート マツマノアマゾン サービスを開始しました。 more ...

Chemical information
KEGG LIGAND SSDB search GENOME
KEGG COMPOUND OC viewer GENES
KEGG REACTION BLAST+ / FASTA+ GENOME+
KEGG METACYC KAAS+ GENOMEST
SIMCOMP

KEGG REACTION KEGG COMP+ / SUBCOMP+ COMPOUND
KEGG DRUG PathSearch+ GLYCAN
KEGG ENVIRON PathComp+ REACTION
KEGG MEDICUS E-zyme+ RPAIR RCCLASS

Health information
KEGG DISEASE MEDICUS search DRUG INTERACTION CHECKER
KEGG DRUG Drug Interaction Checker Human diseases DRUG DRUGPAIR ENVIRON
KEGG ENVIRON Infectious Diseases DRUG GROUP ENVIRON
KEGG MEDICUS ATC drug classification

KEGGはさまざまなDBの集合体

システムの知識
ゲノムの知識
化合物の知識
(健康情報) 9

©2014 統合データベース講習会 Licensed Under CC 表示 2.1

ゲノムの知識データベース・ツール

10

ゲノムの知識データベース・ツール

◆KEGG GENES

- ゲノムが解読された生物種の配列に対して、KEGG
独自の機能アノテーションを付けたDB

◆KEGG ORTHOLOGY (KO)

- 各生物のオーソログ（異なる生物に存在する相同
な機能を持った遺伝子群）を集めたDB

◆KAAS (KEGG Automatic Annotation Sserver)

- 新規ゲノムに対して、自動的にKOをアノテーション

11

アノテーションとは？

◆注釈付け

- 構造アノテーション
 - コーディング領域を特定
- 機能アノテーション
 - 翻訳産物の機能を推定

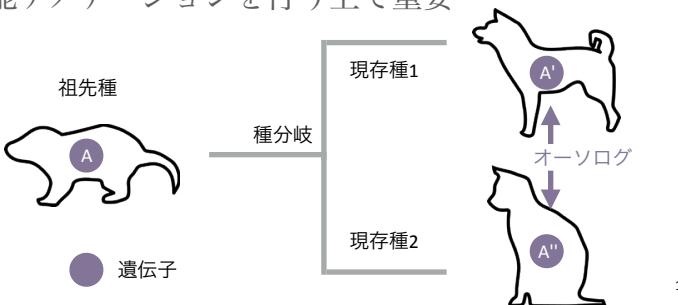
ゲノムから翻訳領域を特定し、
その機能を予測すること

ORIGIN
1 aatgttttgtt gtgtgttgtc gtctgttgtt atccgttgtt gctctgtctt gcaaggaaat
61 gttccgggggg gtgttttttttcc catgttgacc ccgtgttcgg gtgcggggct ttgtataaaat
121 tttttttttt ccccccccgc
181 tgatgtttttt gatccggatcc
241 atgttttgtt gatccggccgcatccatgttgc ctgttttgttgtt gtatgttgtc ttgggtgttg
301 ctttttttgtt gatccgttttttccatgttgc ctgttttgttgc atgttgtttgtt ccataaaaaaa
361 ttctccggccatcc acaatgtgggg tccgcacgcg ttcccttttttccatgtttttttgg ccgcgttttt
421 ggccggcgccatcc gtgtttttttccatgttgc gtt
481 gcaatgtttttttccatgttgc gtt
541 gggccatgttccatgttgc
601 ggccggggatccatgttgc
661 ttt
721 cccccatgttgc
781 ccgttttttttccatgttgc
841 cttagatgttgtccatgttgc
901 caatgttgtccatgttgc
961 atggccatgttccatgttgc
1021 ttccggccatgttgc
1081 accataacttccatgttgc
1261 accatgttgtccatgttgc
1321 tgatgttgtccatgttgc
1381 cccatgttgtccatgttgc
1441 gtttgttttgtccatgttgc
1501 ttcttttgttccatgttgc
機能アノテーション(Functional annotation)
1501 ttcttttgttccatgttgc
12
homology to endoglucanases

©2014 統合データベース講習会 Licensed Under CC 表示 2.1

オーソログとは？

- ◆ 種分岐で生じた生物間で対応する遺伝子
- 同一の生物学的機能を持つと考えられる
- 配列類似性から推定する
- 機能アノテーションを行う上で重要



©2014 統合データベース講習会 Licensed Under CC 表示 2.1

KEGG GENES



KEGG GENES Database

Molecular building blocks of life in the genomic space

KEGG2	PATHWAY	BRITE	MODULE	KO	GENOME	GENES	SSDB	Organisms
<input type="text" value="Enter your gene: (Example: sym:spn3451)"/>	<input type="button" value="Entry"/>	<input type="button" value="Gene cluster"/>	<input type="button" value="Ortholog"/>	<input type="button" value="Paralog"/>	<input type="button" value="Motif"/>	<input type="button" value="Clear"/>		

Gene Catalogs

KEGG GENES is a collection of gene catalogs for all complete genomes (see [reference history](#)) generated from publicly available resources and manually curated by KEGG. They are annotated with KO assignments, generated by KOALA tool. KEGG DBS is a supplementary collection of gene catalogs for eukaryotic draft genomes, which are given automatic KO assignment by KAAS with GENES used as a reference data set. KEGG MGENES contains gene catalogs for metagenomes of environmental samples (see also [KEGG GENOME](#)) with automatic annotation. The viral gene catalog, VGENES, is not yet fully integrated in the KEGG system.

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

Search **[GENES]** for bfind mode bget mode
GENES: コンプリートゲノム
DGENES: ドラフトゲノム
MGENES: メタゲノム
VGENES: ウイルスゲノム

©2014 統合データベース講習会 Licensed Under CC 表示 2.1

生物種リストの表示

Category	Organisms	Source
Homo sapiens	hsa	RefSeq
Pan troglodytes	ptr	RefSeq
Pan paniscus	ppa	RefSeq
Gorilla gorilla gorilla	ggo	RefSeq
Gorilla gorilla gorilla	gor	RefSeq
Macaca mulatta	mmu	RefSeq
Rattus norvegicus	rno	RefSeq
Canis familiaris	cfa	RefSeq
Alouatta macaco	amn	RefSeq
Macaca fasciata	mac	RefSeq
Bos taurus	bos	RefSeq
Ovis aries	oar	RefSeq
Equus caballus	equ	RefSeq
Urotriches macroura	uca	RefSeq
Urotriches macroura	loc	RefSeq
Equus caballus	eqc	RefSeq

- 各生物には3文字もしくは4文字の生物種コードが付けられている
 - 各生物の遺伝子は↓で指定する
生物種コード：遺伝子ID
(例：hsa:1005)
- 14

演習2：KEGG GENESの検索

- ◆ KEGG GENESでキーワード検索をしてみましょう

- 2種類の検索モード
 - bfind: キーワード検索
 - bget: KEGG GENES IDがわかっているとき

Search [GENES] for amylose

bfind mode bget mode

DBのサブカテゴリから検索

Search [Organism] [SSC] for

bfind mode bget mode

生物種を指定して検索

Find three- or four-letter KEGG organism code

pi psp pspf ppi pspif pspifc
 Pichia pastoris [psp]
 Pirellulus stali [psp]
 Pseudonihonium intermedium [psp]
 plg Sus scrofa [pi] psc pscf
 Puccinia graminis var. tritici [pspif]
 Phytophthora infestans [pspifc]

15

KEGG GENES 検索結果

Search [GENES] for amylose

Database: GENES - Search term: amylose (1 - 1000) / Next

hsa:6397
TTL2, C6orf104, NYD-TSPG, D3EN623;3, tubulin tyrosine ligase-like family, member 2; K16600 tubulin polymerase-like TTL2 [EC:6...-]
hsa:20016 AMY2B, AMY2, AMY3, HXA; amylase, alpha 2B (pancreatic) [EC:3.2.1.1]; K01176 alpha-amylase [EC:3.2.1.1]
hsa:20017 AMY2B, AMY2, AMY2B, PA; amylase, alpha 2A (pancreatic) [EC:3.2.1.1]; K01176 alpha-amylase [EC:3.2.1.1]
hsa:104395

KEGG GENES Database for Homo sapiens (human): 279
 Entry: 279 CDS T01001
 Gene name: AMY2A, AMY2, AMY2B, PA
 Definition: amylase, alpha 2A (pancreatic) [EC:3.2.1.1]
 Orthology: K01176 alpha-amylase [EC:3.2.1.1]
 Organism: Homo sapiens (human)
 Pathway: carbohydrate metabolism; carbohydrate digestion and absorption; digestive system; metabolism

Position: 511 aa / 1000 b.p. (91%)
 AA seq: MKFFLLTTGFQWADYSNTQQLTSWFLFEWWVVVQWSRPLNLRKG
 NEVAWWNPFRPWRFTGPTYPSVYLNPSQNGEPRFNMAYTTRGIVPRVY
 AVAAKTKPFYHNRYHPTTNRVLLVYDNLTVLQGKNAKQGKRNQDANPPLD
 QDLMKPNKRYVHKNVHPRQAFGRDLSHHSQADLNLKNNNSWIPFAG
 SKPFRYQVFLDGPKFSQSDFQFNGRTYFEEFKAKLTVKINGENSKYLNWPG
 FVPSRKAUFLVUNHQNKGHGAGGCTTWHDALYKMAWV
 IFOQFOPHRRPQWYRSGQYHGAQHGRDPRRQFADQHQSNDKQVHDFH
 FTVWYONGSNQVAFGRQRDRQFVNNDOWFSLTQJGJL...
 KYVSODKAHISGSAEDEPNAHAEK

他のDBへのリンク

立体構造
 枠内blast検索
 パスウェイ
 階層分類
 アミノ酸配列
 塩基配列
 モチーフ

©2014 統合データベース講習会 Licensed Under CC 表示 2.1

16

KEGG ORTHOLOGY (KO) のエントリ

- 各生物のオーソログ（異なる生物に存在する相同的な機能を持った遺伝子群）を集めたDB
- 新しいゲノムがKEGG GENESに登録されるときKOの対応関係を計算して、KEGG独自のアノテーションをつけている
- K番号
IDはKから始まる
- オーソロググループの名前、機能、各DBへのリンク、グループに分類されている遺伝子のリスト、機能の根拠となる文献などが記載

ORTHOLOGY: K01176

Entry K01176 KO
Name E3.2.1.., amylA, malS
Definition alpha-amylase [EC:3.2.1.1] KEGGでつけられたアノテーション
Pathway ko00500 Starch and sucrose metabolism
ko04973 Carbohydrate digestion and absorption
KEGG Orthology (KO) [BR:ko00001]
Molecular functions
Carbohydrate metabolism
00500 Starch and sucrose metabolism
K01176 E3.2.1.., amylA, malS; alpha-amylase
Organisms
Digestive system
04973 Carbohydrate digestion and absorption
K01176 E3.2.1.., amylA, malS; alpha-amylase
Enzymes [BR:ko01000]
B: Carbohydrates
3.2. Glycosides
3.2.1 Glycosidases, i.e. enzymes that hydrolyse O- and S-glycosyl compounds
3.2.1.1 alpha-amylase
K01176 E3.2.1.., amylA, malS; alpha-amylase
BRITHE hierarchy

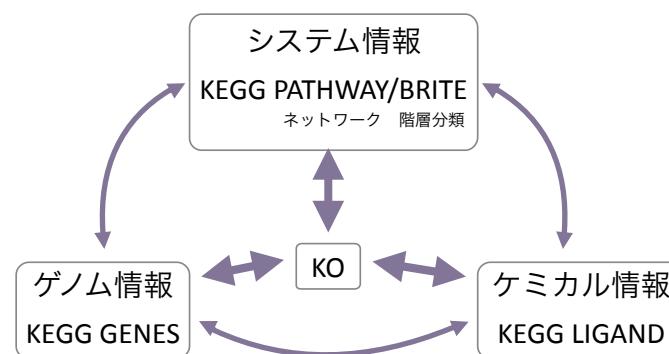
Other DBs R02108
GO: 0004556
Genes HSA: 276(AMY1A) 277(AMY1B) 278(AMY1C) 279(AMY2A) 280(AMY2B)
PTC: 457057(AMY2B) 469397(AMY2A) 736558(AMY2A)
PPS: 100980682 100981583 100983088(AMY2B) 100983778(AMY2A)
100993245(AMY1A)
GGD: 101133335(AMY2A) 101134742(AMY2B) 101135108(AMY1B)
100980682
PON: 100446394(AMY1A) 100447008
MCC: 712893(AMY2B) 713027(AMY2A)
MCF: 101866053 102123076(AMY2A)
MMU: 100436255 100436256 100436257(Amy2a3) 100436888(Amy2a2)
100980682(Amy2a3) 11723(Amy1)
RNO: 100911494 24203(Amy1a) 497039(Amy2a3)
CGE: 100766318 100767001
show all
Taxonomy | KOALA | UniProt

Reference PMID:941418
Authors Janecek S
Title alpha-Amylase family: molecular biology and evolution.
Journal Prog Biophys Mol Biol 67:67-97 (1997)

17 文献

(c) KEGG ©2014 統合データベース講習会 Licensed Under CC 表示 2.1

KEGGシステムはKOを中心に構成されている



3つのタイプのDBの多くは、KOを介してリンクされているため、KEGGを利用する上で非常に重要
KOグループに割り当てられていない、KEGG外部の配列データ、ユーザが自分でシーケンスした配列データからはKEGGのシステムが利用できない？

18

(c) KEGG ©2014 統合データベース講習会 Licensed Under CC 表示 2.1

KAAS

Gene Annotation
The annotation of KEGG GENES involves assignment of KO identifiers (K numbers). Internally, this is done using the KOALA and GFTT annotation tools based on the SSDb database (see: Genome Annotation in KEGG). The annotation of KEGG DGENES and MGENES is done automatically using the KAAS program shown below.

Annotate genomes using KEGG GENES
KAAS: automatic annotation (KO assignment) and pathway reconstruction [reference]

Search similar sequences in KEGG GENES

BLAST: sequence similarity search by BLAST
FASTA: sequence similarity search by FASTA

完全ゲノム/ドラフトゲノム用

部分ゲノム/メタゲノム用

EST用

・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
パスウェイマップ
BRITE階層分類

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

Request

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 Ortholog) assignments and automatically generated KEGG pathways.

KAAS also provides a set of tools for ortholog assignment and pathway reconstruction.

KAAS job request (Blast method)

KAAS job request (SBB method)

KAAS interactive

Example of Results

KO assignment

KEGG pathway mapping

EST

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

KAAS job request (Blast method)

KAAS job request (SBB method)

(c) KEGG ©2014 統合データベース講習会 Licensed Under CC 表示 2.1

KAAS入力画面

KAAS Job Request

Home Help

Query sequences (in multi-FASTA)

Text data (in Nucleotide)
+>P04327;1 hypothetical protein GAT_0005
(Glycoside hydrolase family 23 member 2)
+>P04327;2 hypothetical protein GAT_0006
(Glycoside hydrolase family 23 member 3)
+>P04327;3 hypothetical protein GAT_0007
(Glycoside hydrolase family 23 member 4)
File upload (in Nucleotide)
ファイルを選択 読み込まれていません

Query name query

E-mail address

GENES data set (presubmit button or text box)

Representative set
for GENES for eukaryotes for Prokaryotes

Manual selection
Organism list (drop-down window)
Selected organisms (up to 40 organisms)
has, dme, cel, ssc, ore, cho, eco, nre, hga, rpr, bsa, lla, lac, mru, cfr, bba, syn, dha, dca, rpa, rpe

(Example) eco, bsa, syn

Assignment method
B: Blast (Bi-directional best Hit)
S: SBB (Single-direction best Hit)

Compute Clear

Multi FASTA配列（ゲノム配列そのものは×）

コピー&ペースト or ファイルアップロード

メールアドレス

計算に時間がかかるため、終わったらメールで知らせてくれる

blast検索する近縁の生物種

なるべく近縁の生物種を選ぶ

双方向 or 片方向ベストヒット

20

(c) KEGG ©2014 統合データベース講習会 Licensed Under CC 表示 2.1

KAAS入力と結果例

酢酸菌Gluconacetobacter hansenii ATCC23769完全ゲノムをKAASで自動アノテーション

```
>gi|296114119|ref|ZP_06832775.1| hypothetical protein GXY_00005 [Gluconacetobacter hansenii ATCC 23769]
MRWGLVMIQACMIVLKGIEHLGSRDAAAPAGVV
>gi|296114121|ref|ZP_06832776.1| chaperone clpB [Gluconacetobacter hansenii ATCC 23769]
MNIEKFTERSRGFLQAAQTIAIMREYNQLTPEHLLKALDDQGAAASLURAAGGQPAAIAAVDTALAK
LPKVQGGGAGQPSATPLVRLDAEQAQAKGDVEVAQDRLLAIAASETPGGALRAGGTQALDKA
IATIRKGRTVTSNEAEASFDALKYARDVTEIALQGKLDPVGRDEERRAIQLVLRASNK
>gi|296114122|ref|ZP_06832777.1| hypothetical protein GXY_00019 [Gluconacetobacter hansenii ATCC 23769]
MGAALIMITTLIAGYALCEWPGEETPALMQIK
```

マルチFASTAファイルを入力

TCAサイクル

KO Assignment Results

Home [KO list] [BRITE hierarchies] [Pathway map] [Threshold change] [Download]

Query gene : KO assignment

query

```
gi|296114119|ref|ZP_06832775.1|
gi|296114121|ref|ZP_06832776.1| K03695
gi|296114122|ref|ZP_06832777.1|
gi|296114123|ref|ZP_06832778.1|
gi|296114124|ref|ZP_06832779.1|
gi|296114125|ref|ZP_06832780.1| K00053
gi|296114126|ref|ZP_06832782.1| K01653
gi|296114127|ref|ZP_06832783.1| K01652
gi|296114128|ref|ZP_06832784.1| K00791
```

©2014 統合データベース講習会 Licensed Under CC 表示 2.1

21

化合物の知識データベース・ツール

22

©2014 統合データベース講習会 Licensed Under CC 表示 2.1

KEGG LIGAND

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

Search (LIGAND for [LIGAND]) Go Clear

bfind mode bget mode

LIGAND Relational Database

The primary database of KEGG LIGAND is a relational database with the KegDraw interface, which is used to generate the secondary (flat file) database for DBGET. A read-only copy of the LIGAND relational database is made publicly accessible.

Search COMPOUND

Compound ID : example) C00010 Go Clear

Search GLYCAN

Glycan ID : example) G10596 Go Clear

Search REACTION

Reaction ID : example) R00259 Go Clear

Search RPAIR

Rpair ID : example) RP00010 Go Clear

©2014 統合データベース講習会 Licensed Under CC 表示 2.1

サブデータベース

COMPOUND 中間代謝物、二次代謝産物などの化合物DB

GLYCAN

糖鎖構造DB

REACTION/RPAIR/RCLASS/ENZYME 生化学反応DB

キーワード/ID検索

化合物の特徴による検索
分子式、分子量など

23

化合物の知識データベース・ツール

◆KEGG LIGAND

●化合物関係のDBを総称してLIGANDと呼んでいる

◆KEGG COMPOUND

●主に代謝化合物を収録したDB

◆KEGG REACTION

●生体内反応のDB、これらをつなぎ合わせると
パスウェイになる

24

©2014 統合データベース講習会 Licensed Under CC 表示 2.1

演習3

◆ KEGG LIGANDでキーワード検索してみましょう



KEGG LIGAND Database

Molecular building blocks of life in the chemical space

KEGG2 PATHWAY BRITE MODULE LIGAND COMPOUND GLYCAN REACTION

Enter C/G/R/RP/RC numbers (Example) C00389 C05903 C06562 C08650 C09727 C09762

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 TUBMB/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 phenylalanin Go Clear

bind mode bget mode

25

(c) KEGG ©2014 統合データベース講習会 Licensed Under CC 表示 2.1

KEGG LIGAND検索結果

KEGG Search LIGAND for phenylalanin Go Clear

Database: LIGAND - Search term: phenylalanin

KEGG COMPOUND

C00079	KeplDraw	1mol
L-Phenylalanine; (S)-alpha-Amino-beta-phenylpropionic acid		
C00355	KeplDraw	1mol
3,4-Dihydroxy-L-phenylalanine; L-Dopa; 3-Hydroxy-L-tyrosine; L-beta-(3,4-Dihydroxyphenyl)alanine; Levodopa; Dihydroxy-L-phenylalanine		
C02057	KeplDraw	1mol
Phenylalanine; alpha-Amino-beta-phenylpropionic acid		
C02061	KeplDraw	1mol
D-Phenylalanine; D-alpha-Amino-beta-phenylpropionic acid		
C03519	KeplDraw	1mol
N-Acetyl-phenylalanine		

... display all

KEGG REACTION

R00331	KeplDraw	1mol
1,2-benzenediol:ox		
R00441	KeplDraw	1mol
3,4-dihydroxy-L-phenylalanine:Dopaglutamate + 2 H		
R00686	KeplDraw	1mol
L-Phenylalanine racemase (ATP-hydrolysing); ATP + L-Phenylalanine + H2O <=> AMP + Diphosphate + D-Phenylalanine		
R00688	KeplDraw	1mol
L-phenylalanine:NAD+ oxidoreductase (deaminating); L-Phenylalanine + H2O + NAD+ <=> Phenylpyruvate + Ammonia + NADH + H+		
R00689	KeplDraw	1mol
L-phenylalanine:oxygen oxidoreductase (deaminating); L-Phenylalanine + H2O + Oxygen <=> Phenylpyruvate + Ammonia + Hydrogen peroxide		
... display all		
KEGG ENZYME		
1.4.1.20	KeplDraw	1mol
phenylalanine dehydrogenase; L-phenylalanine dehydrogenase; PHD		
1.1.3.2.9	KeplDraw	1mol
phenylalanine 2-monooxygenase; L-phenylalanine oxidase (deaminating and decarboxylating); phenylalanine (deaminating, decarboxylating)oxidase		
1.1.3.2.15	KeplDraw	1mol
3,4-dihydroxyphenylalanine oxidative deaminase; 3,4-dihydroxy-L-phenylalanine: oxidative deaminase; ortho-deaminase; DOPA oxidative deaminase; DOPADA		
1.1.3.12.24	KeplDraw	1mol
phenylalanine N-monoxygenase; phenylalanine N-hydroxylase; CYP79A2		

(c) KEGG ©2014 統合データベース講習会 Licensed Under CC 表示 2.1

26

KEGG COMPOUNDのエントリ

COMPOND: C00079

Entry C00079 Compound

Name L-Phenylalanine; (S)-alpha-Amino-beta-phenylpropionic acid

Formula C9H11NO2

Exact mass 165.079

Mol weight 165.1891

Structure

Remark Same as: D00021
Reaction R00656 R00658 R00659 R00660 R00661 R006592 R00693 R00694
R00659 R00697 R00698 R00699 R01372 R01376 R01
R00744 R07211 R08435 R08463 R08652 R08660 R09
R09579 R10286 R10287 R10495 R10499 R10729

Pathway map0360: Phenylalanine metabolism
map00400: Phenylalanine, tyrosine and tryptophan biosynthesis
map00460: Cysteine and methionine metabolism
map00940: Phenylpropanoid biosynthesis
map00960: Tropane, piperidine and pyridine alkaloid biosynthesis
map00966: Glucosinolate biosynthesis
map00970: Aminoacyl-tRNA biosynthesis
map01060: Biosynthesis of plant secondary metabolites
map01061: Biosynthesis of phenylpropanoids
map01063: Biosynthesis of alkaloids derived from shikimate pathway
map01064: Biosynthesis of alkaloids derived from ornithine, lysine and nicotinic acid
map01070: Biosynthesis of plant hormones
map01100: Metabolic pathways
map01110: Biosynthesis of secondary metabolites
map01210: 2-Oxocarboxylic acid metabolism
map01230: Biosynthesis of amino acids

KEGG2 PATHWAY BRITE MODULE LIGAND COMPOUND GLYCAN REACTION

Enter C/G/R/RP/RC numbers (Example) C00389 C05903 C06562 C08650 C09727 C09762

Filter Pathway mapping Brite mapping Get title Get entry Clear

モジュール：生物種間でよく保存されている一連の反応

Module M00024: Phenylalanine biosynthesis, chorismate => phenylalanine
M00039: Monolignol biosynthesis, phenylalanine/tyrosine => monolignol
M00137: Flavanon biosynthesis, phenylalanine => naringenin
M00350: Capsaicin biosynthesis, L-Phenylalanine => Capsaicin

Enzyme C00079

酵素番号DBへのリンク

Compounds with biological roles [Brbr08001]
Peptides
Amino acids
Common amino acids
C00079: Phenylalanine (Phe)
Risk category of Japanese OTC drugs [Brbr08312]
Third-class OTC drugs
map00360: Phenylalanine metabolism
map00400: Phenylalanine, tyrosine and tryptophan biosynthesis
map00460: Cysteine and methionine metabolism
map00940: Phenylpropanoid biosynthesis
map00960: Tropane, piperidine and pyridine alkaloid biosynthesis
map00966: Glucosinolate biosynthesis
map00970: Aminoacyl-tRNA biosynthesis
map01060: Biosynthesis of plant secondary metabolites
map01061: Biosynthesis of phenylpropanoids
map01063: Biosynthesis of alkaloids derived from shikimate pathway
map01064: Biosynthesis of alkaloids derived from ornithine, lysine and nicotinic acid
map01070: Biosynthesis of plant hormones
map01100: Metabolic pathways
map01110: Biosynthesis of secondary metabolites
map01210: 2-Oxocarboxylic acid metabolism
map01230: Biosynthesis of amino acids

階層分類

パスウェイ

他DBへのリンク

KCF data Show

(c) KEGG ©2014 統合データベース講習会 Licensed Under CC 表示 2.1

KEGG REACTIONのエントリ

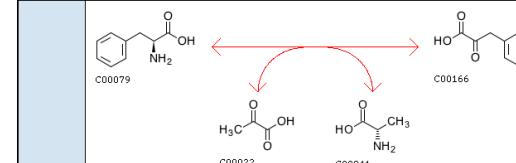
KEGG REACTION: R00692

Entry R00692 Reaction

Name L-phenylalanine:pyruvate aminotransferase

Definition L-Phenylalanine + Pyruvate <=> Phenylpyruvate + L-Alanine

Equation C00079 + C00022 <=> C00166 + C00041



RPair

R00021 C00022_C00041 main [RC:R00008]

R000057 C00079_C00166 main [RC:R00006]

R006279 C00022_C00166 trans

R006280 C00041_C00079 trans

Enzyme 2.6.1.58

Pathway rn00360: Phenylalanine metabolism

rn00400: Phenylalanine, tyrosine and tryptophan biosynthesis

rn00960: Tropane, piperidine and pyridine alkaloid biosynthesis

rn01100: Metabolic pathways

rn01110: Biosynthesis of secondary metabolites

Orthology K05821: aromatic amino acid aminotransferase II [EC:2.6.1.58 2.6.1.28]

R番号

COMPOUND IDを使った反応式

反応式

RC (Reaction Class):反応中心の分類

RPAIR: 基質・生成物間での原子のアラインメント
酵素番号DBへのリンク

パスウェイ

触媒する酵素のKO番号

28

(c) KEGG ©2014 統合データベース講習会 Licensed Under CC 表示 2.1

KEGG RPAIR: リアクションペア

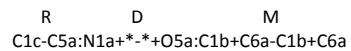
RP番号

基質・生成物間での原子のアラインメント



KEGGでは原子をまわりの環境（隣の原子とかー重結合二重結合などの結合様式）で原子を細かく分類している

反応中心近辺の反応をRDMで定義



同じRDMパターンの反応を集めてDB化

⇒KEGG RCLASS

©2014 統合データベース講習会 Licensed Under CC 表示

Kegg RPAIR: RP00057

Entry	RP00057	RPair	Help
Name	C00079_C00166		
Compound	C00079 L-Phenylalanine C00166 Phenylpyruvate		
Type	main		
RDM	1 1 C1c-C5a:N1a+*-+O5a:C1b+C6a-C1b+C6a		
RClass	RC00006		
Related pair	RP00014 RP00059 RP00084 RP00101 RP00126 RP00133 RP00204 RP00271 RP00289 RP00298 RP00309 RP00356 RP00424 RP00550 RP00558 RP00621 RP00622 RP00683 RP00813 RP00842 RP00904 RP01339 RP01721 RP01859 RP01955 RP02210 RP02418 RP02426 RP02474 RP02552 RP02676 RP02724 RP02734 RP02841 RP02865 RP02876 RP02897 RP03571 RP03653 RP03678 RP03804 RP03806 RP03972 RP04106 RP04114 RP04518 RP04680 RP11549 RP12924 RP13568 RP13577 RP13586 RP13595 RP13604 RP13613 RP13688		
<input type="button" value="DB search"/>			

KEGG RCLASS

RC番号

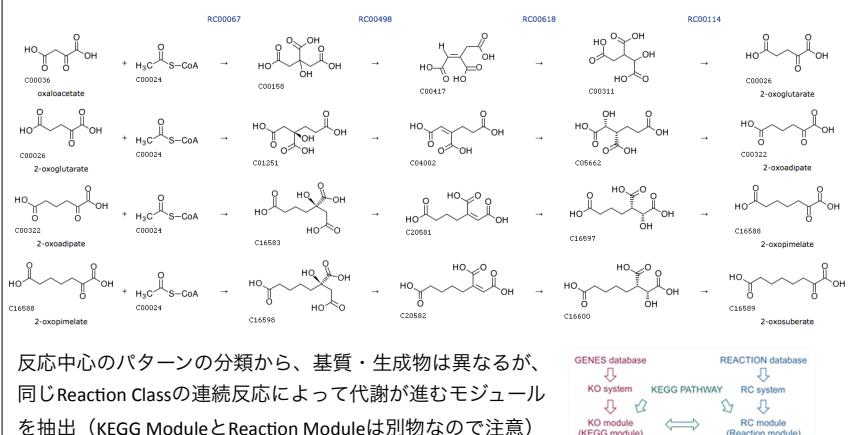
RDMのパターンでRPAIRを分類

Kegg RCLASS: RC00006

Entry	RC00006	RClass	Help
Definition	C1c-C5a:N1a+*-+O5a:C1b+C6a-C1b+C6a		
RPair	RP00014 RP00057 RP00059 RP00084 RP00101 RP00126 RP00133 RP00204 RP00271 RP00289 RP00298 RP00309 RP00356 RP00424 RP00550 RP00558 RP00621 RP00622 RP00683 RP00813 RP00842 RP00904 RP01339 RP01721 RP01859 RP01955 RP02210 RP02418 RP02426 RP02474 RP02552 RP02676 RP02724 RP02734 RP02841 RP02865 RP02876 RP02897 RP03571 RP03653 RP03678 RP03804 RP03806 RP03972 RP04106 RP04114 RP04518 RP04680 RP11549 RP12924 RP13568 RP13577 RP13586 RP13595 RP13604 RP13613 RP13688 RP14045 RP14666 RP16440		
Related class	<input type="button" value="DB search"/>		
Reaction	RP00021 RO0093 RO0114 RO0243 RO0248 RO0250 RO0258 RO0279 RO0355 RO0357 RO0359 RO0372 RO0373 RO0400 RO0447 RO0457 RO0493 RO0576 RO0581 RO0585 RO0588 RO0648 RO0652 RO0667 RO0668 RO0675 RO0676 RO0677 RO0684 RO0688 RO0689 RO0692 RO0694 RO0695 RO0729 RO0734 RO0895 RO0896 RO0908 RO1088 RO1090 RO1148 RO1155 RO1161 RO1214 RO1339 RO1346 RO1374 RO1375 RO1376 RO1380 RO1582 RO1648 RO1684 RO1713 RO1731 RO1939 RO1956 RO1957 RO2077 RO2199 RO2274 RO2283 RO2433 RO2457 RO2459 RO2618 RO2619 RO2700 RO2713 RO2755 RO2772 RO2773 RO2781 RO2851 RO2923 RO2924 RO3001 RO3053 RO3063 » show all		
Enzyme	1.4.1.2 1.4.1.3 1.4.1.4 1.4.1.5 1.4.1.6 1.4.1.9 1.4.1.20 1.4.1.21 1.4.1.1		

©2014 統合データベース講習会 Licensed Under CC 表示

Reaction Module



31

SIMCOMP

類似した化合物を検索する
 BLASTの化合物版

SIMCOMP: DB中から類似構造を検索

SIMCOMP2: 2個の化合物の間の類似度を計算

SUBCOMP: 部分構造検索

KegDraw (化合物お絵かきソフト) で

化合物を描いてSIMCOMPで検索することも可能

SIMCOMP Search

SUBCOMP	SIMCOMP	SIMCOMP2	API	KEGG2
Enter query compound: (in one of the four forms) Compound ID (Example) C00222 View structure MOL File Name ファイルを選択 検索されません MOL File Text SMILES Select target database: <input checked="" type="radio"/> COMPOUND <input type="radio"/> DRUG <input type="radio"/> KNAPSAck <input type="radio"/> REACTION Search options: <input checked="" type="radio"/> Global search <input type="radio"/> Local search <input type="radio"/> Customized search SIMCOMP Help SIMCOMP REST API Download KegDraw KegDraw is a Java application for drawing compound structure. Feedback KEGG GenomeNet Kyoto University Bioinformatics Center				

©2014 統合データベース講習会 Licensed Under CC 表示 2.1

システムの知識データベース・ツール

33

©2014 統合データベース講習会 Licensed Under CC 表示 2.1

システムの知識データベース・ツール

◆KEGG PATHWAY

- 代謝系、制御系、疾患、薬の開発などのネットワークを収集したDB

◆Global Map

- 代謝系ネットワークを俯瞰する

◆KEGG MODULE

- 生物種間で保存されている連続的な反応を集めたD

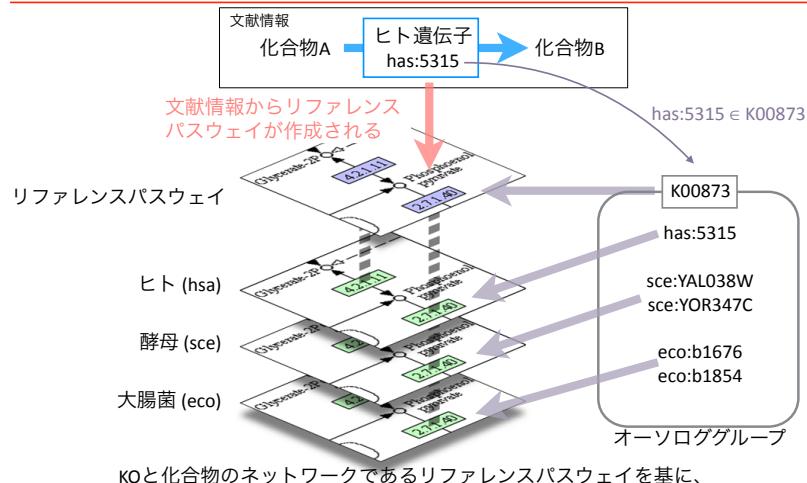
◆KEGG MAPPER

- それぞれの酵素や化合物の色を指定して表示

34

©2014 統合データベース講習会 Licensed Under CC 表示 2.1

KEGG PATHWAY



35

©2014 統合データベース講習会 Licensed Under CC 表示 2.1

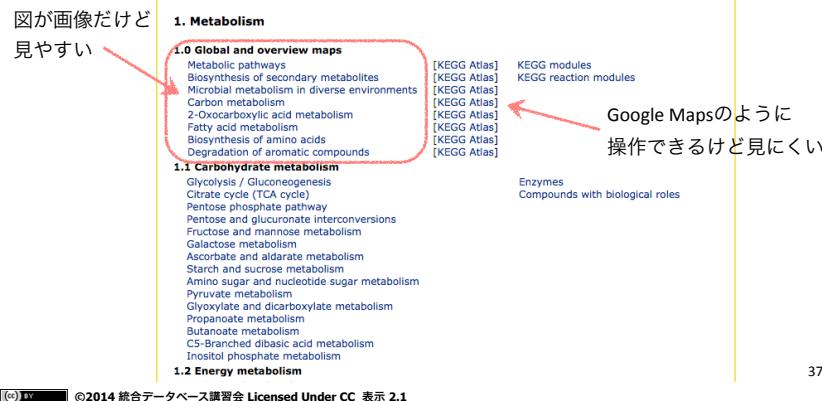
KEGG PATHWAY



©2014 統合データベース講習会 Licensed Under CC 表示 2.1

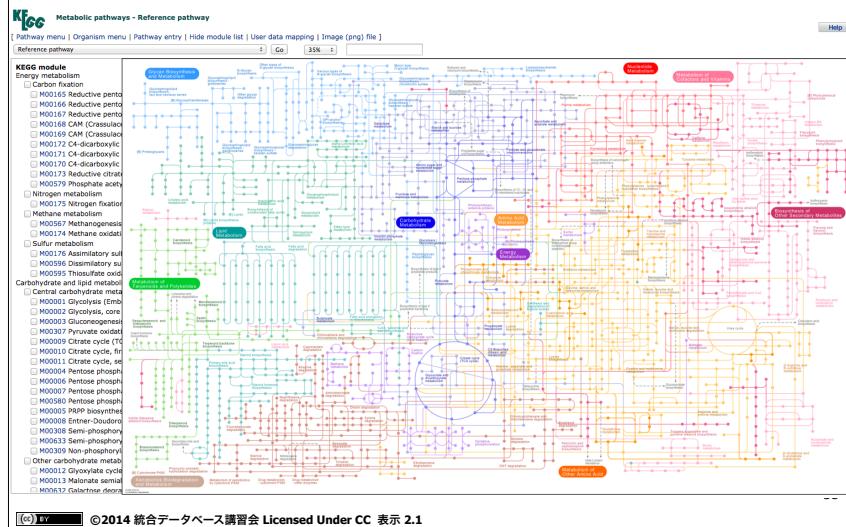
演習4

◆KEGG PATHWAYで様々な生物種のGlobal Mapを表示して、使われているパスウェイの違いを俯瞰してみましょう



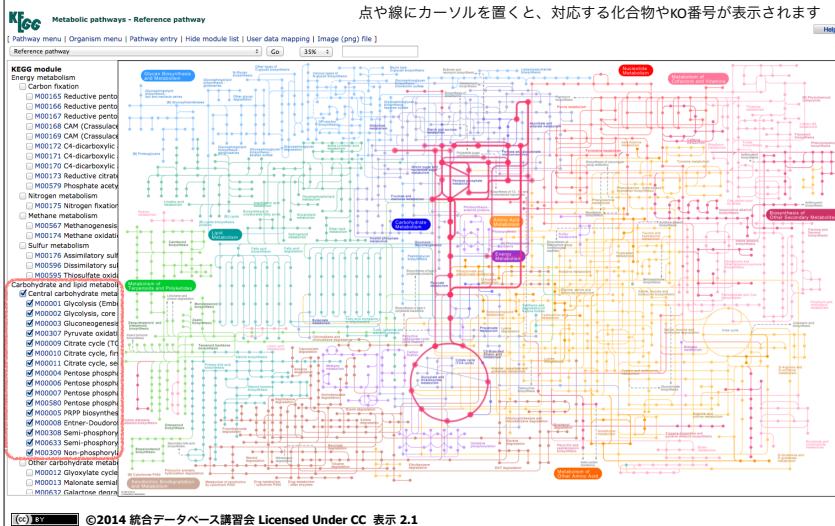
37

Global Map (Reference pathway)



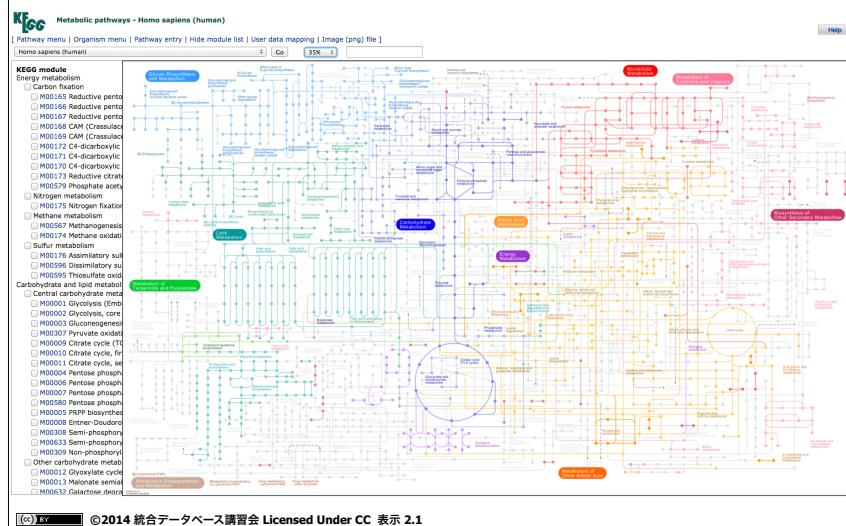
©2014 統合データベース講習会 Licensed Under CC 表示 2.1

糖代謝を強調表示



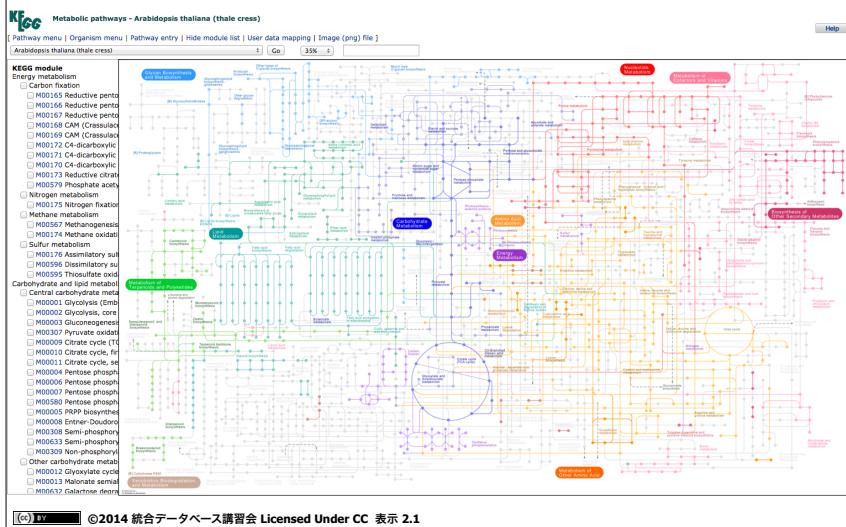
37

Global Map (Homo sapiens)



©2014 統合データベース講習会 Licensed Under CC 表示 2.1

Global Map (Arabidopsis thaliana)

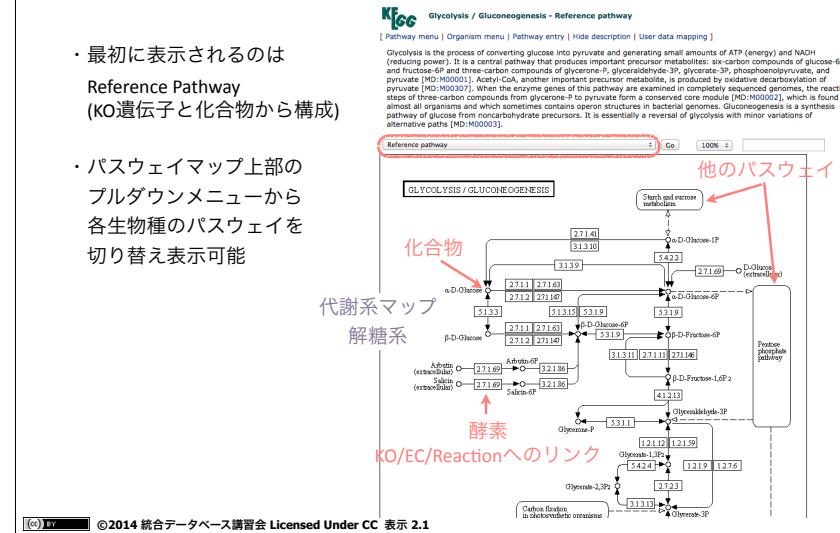


KEGG PATHWAY - Reference pathway

最初に表示されるのは

Reference Pathway
(KO遺伝子と化合物から構成)

・パスウェイマップ上部の
ブルダウンメニューから
各生物種のパスウェイを
切り替え表示可能

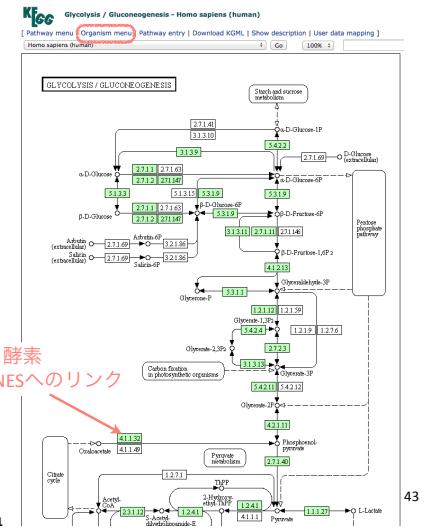


実習5: いろいろな生物種のパスウェイを表示してみましょう

・ヒトの解糖系を表示した場合

- ・パスウェイマップ上部の
ブルダウンメニューから
各生物種のパスウェイを
切り替え表示可能
- ・目的の生物種を探しにくい時には、
ブルダウンメニューの上にある
"Organism menu" のリストから
選ぶこともできます

酵素
KEGG GENESへのリンク



KEGG MODULE

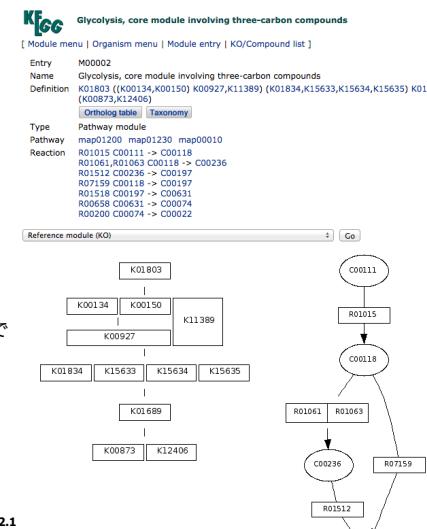
・パスウェイモジュール

・M番号

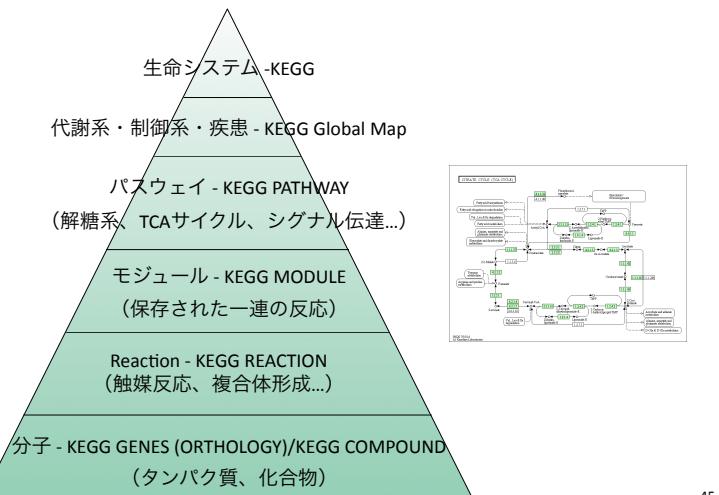
・PATHWAYマップより小さい単位

・生物種間での保存度合い、複合体形成、
オペロンなどを考慮した機能単位

・オーソロググループ (KO) の組み合わせで
表現



システムの階層性



©2014 統合データベース講習会 Licensed Under CC 表示 2.1

KEGG Mapper

The KEGG Mapper interface allows users to map their data onto KEGG Pathway maps.

Search Bar: Select prefix (map or organism), Enter keywords, Go, Help, New maps, Update history.

Pathway Maps:

- KEGG PATHWAY Database:** A collection of manually drawn pathway maps representing knowledge on the molecular interactions and reaction networks for:
 - Metabolism
 - Genetic Information Processing
 - Environmental Information Processing
 - Cellular Processes
 - Organismal Systems
 - Human Diseases
- Pathway Mapping:** KEGG PATHWAY mapping is the process to map molecular datasets, especially large-scale datasets in genomics, transcriptomics, proteomics, and metabolomics, to the KEGG pathway maps for biological interpretation of higher-level systemic functions.
 - Search Pathway - basic pathway mapping tool
 - Search&Color Pathway - advanced pathway mapping tool
 - Color Pathway - selected pathway map coloring tool
- 1. Metabolism:** 1.0 Global and overview maps

©2014 統合データベース講習会 Licensed Under CC 表示 2.1

Search Pathway

The KEGG Mapper - Search Pathway interface allows users to search for pathways based on various criteria.

Search Fields: Search against: hsa, Enter: map, ko, ec, rn, hasdd, or [org].

Enter Objects: Examples: Homo sapiens pathway. Input field: hsa:7167 hsa:GPI cpd:CD0118 ALDOA 1.2.1.12 CD0236.

Search Options:

- Alternatively, enter the file name containing the data: ファイルを選択: 対象を選択していません.
- Filter1: Filter2 (to extract K/C/G/O/R/P/B numbers)
- Include aliases
- Display objects not found in the search
- Search pathways containing all the objects (AND search)

Buttons: Exec, Clear.

47

©2014 統合データベース講習会 Licensed Under CC 表示 2.1

Search&Color Pathway

The KEGG Mapper - Search&Color Pathway interface allows users to color KEGG pathway maps based on specific identifiers.

Search Fields: Search against: hsa, Enter: map, ko, ec, rn, hasdd, or [org]. Primary ID: KEGG identifiers. Examples: Homo sapiens pathway. Input field: 7167:red,blue C00118:pink.

Search Options:

- Enter objects one line followed by bcolor, fcicolor:
- If necessary, change default bcolor: pink.
- Include aliases
- Use uncoupled diagrams
- Display objects not found in the search
- Search pathways containing all the objects (AND search)

Buttons: Exec, Clear.

48

©2014 統合データベース講習会 Licensed Under CC 表示 2.1

Color Pathway

定量データをパスウェイ上で可視化する

KEGG Mapper - Color Pathway

2. 表示するパスウェイを指定
Select KEGG pathway map: hsa05200

Enter file name containing the data:
ファイルを選択 hsa_CML-COSMIC.txt

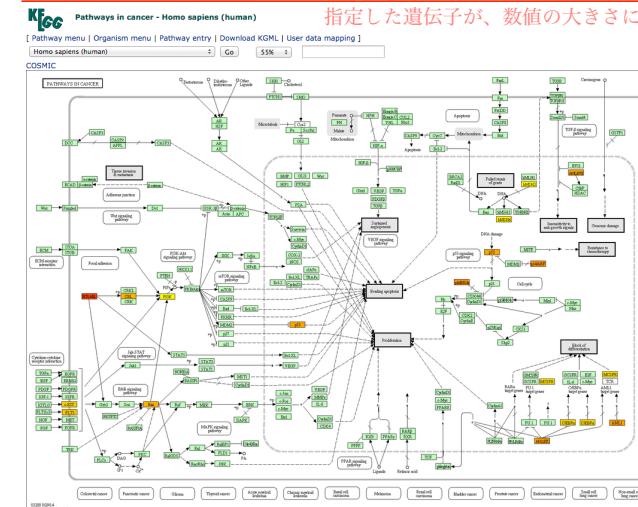
3. 保存したファイルを選択
File type: Color specification
Numerical value (Convert to log scale) minimum-maximum negative-zero-positive
Minimum: #ffff00 Maximum: #ff0000 Check

4. File typeをNumerical value, log scaleに変更
1. 右クリックで保存

色の指定、チェックができる 49

KEGG Mapperの結果

指定した遺伝子が、数値の大きさによって色付けされる



50

KEGG Mapperの結果

KEGG Pathways in cancer - Homo sapiens (human)

CLP/CMP - Blast_phase - All

Color pathwayでは複数の色付けマップを切り替えて比較可能

#hsa CLP/CMP_Blast_phase_All
613 #ddccff,red #ddccff,red
25 #ddccff,red #ddccff,red
1398 #ddccff #ddccff
1399 #ddccff #ddccff
23624 #ddccff #ddccff
867 #ddccff #ddccff
868 #ddccff #ddccff
5290 #ddccff #ddccff
5291 #ddccff #ddccff
5292 #ddccff #ddccff
5293 #ddccff #ddccff
23533 #ddccff #ddccff
5295 #ddccff #ddccff
5296 #ddccff #ddccff
8503 #ddccff #ddccff
201 #ddccff #ddccff
208 #ddccff #ddccff
10000 #ddccff #ddccff
572 #ddccff #ddccff
598 #ddccff #ddccff
1147 #ddccff #ddccff
353 #ddccff #ddccff
8517 #ddccff #ddccff
4792 #ddccff #ddccff
4790 #ddccff #ddccff
4791 #ddccff #ddccff
5970 #ddccff #ddccff
4120 #ddccff,red #ddccff,red
7157 #ddccff,red #ddccff,red
1029 #ddccff #ddccff
1026 #ddccff #ddccff
595 #ddccff #ddccff
1019 #ddccff #ddccff
1021 #ddccff #ddccff
5925 #ddccff,red #ddccff,red
1869 #ddccff #ddccff

51

複数生物種のパスウェイを比較

KEGGトップページのAnalysis toolsから

Reconstruct Pathwayを選択

Analysis tools
KEGG Mapper
KEGG Atlas
BlastKOALA New
BLAST/Fasta
SIMCOMP

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.
Multiple search objects may be specified in multiple lines.

Search&Color Pathway - pathway mapping tool with coloring options
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).
Background colors are specified by color names.
Numerical values 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.,

numerical background and foreground colors in multiple columns.

Results are shown in multiple snapshots of the map.

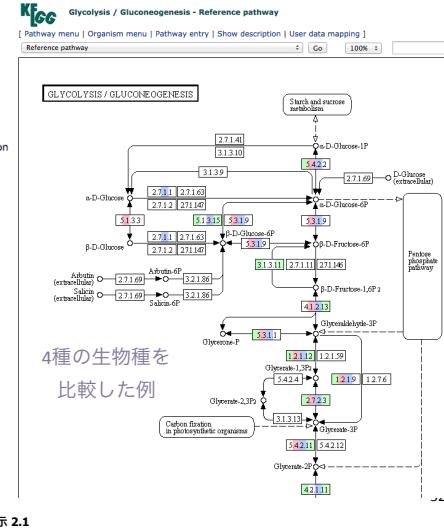
When numerical values are given as attributes, they are displayed by orientation or by using WebGL graph.

Reconstruct Pathway - pathway mapping tool to assist genome annotation

Search against all pathway maps in the ko category.

Mapped objects are marked in green for a single genome or in multiple colors (see: KEGG color codes) for multiple genomes

©2014 統合データベース講習会 Licensed Under CC 表示 2.1



4種の生物種を
比較した例

©2014 統合データベース講習会 Licensed Under CC 表示 2.1

Reactome

ヒトを中心としたパスウェイデータベース

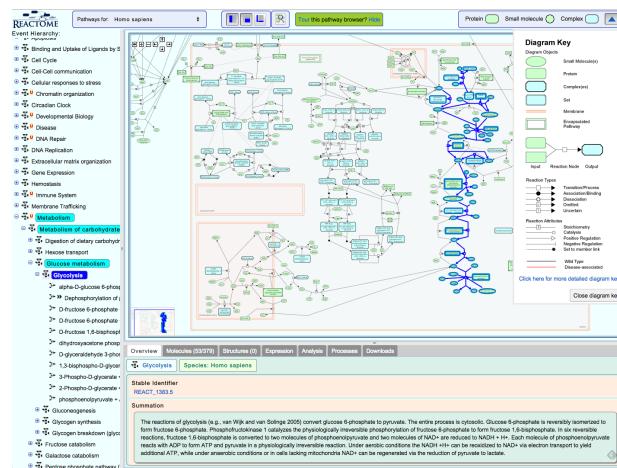
<http://www.reactome.org/>

53

©2014 統合データベース講習会 Licensed Under CC 表示 2.1

Browse Pathway

階層をたどってパスウェイを表示 (≒KEGG Global Atlas)



55

©2014 統合データベース講習会 Licensed Under CC 表示 2.1

Reactome



About Content Documentation Tools Community Download Contact e.g. O95631, NTN1, signa Search

Browse Pathways Analyze Data Reactome FI Network

User Guide Data Download Contact Us

About Reactome

Reactome is a free, open-source, curated and peer reviewed pathway database. Our goal is to provide intuitive bioinformatics tools for the visualization, interpretation and analysis of pathway knowledge to support basic research, genome analysis, modeling, systems biology and education. The current version (v49) of Reactome was released on July 15, 2014.

The development of Reactome is supported by a grant from the US National Institutes of Health (P41 HG003751), Ontario Research Fund, and the European Molecular Biology Laboratory.

reactome @reactome 3 Sep Pop by to see @EMBLComplex Portal poster with links to @reactome, #FEBSEMBIO2014, #EMBLEBI, #INTACT_PROJECT

reactome @reactome 2 Sep Join @robinshaw tomorrow @10:15 in Rm242AB, @Reactome talk, modeling biological process session, #FEBSEMBIO2014 #OICR_news #EMBLEBI #NYULMC Expand

reactome @reactome 25 Jul New Full-text Search Tool! Reactome is pleased to announce the release of a new search based upon the

[Tweet to @reactome](#)

54

©2014 統合データベース講習会 Licensed Under CC 表示 2.1

Analysis Tools

遺伝子、タンパク質、化合物の名前、IDを指定してパスウェイにマッピング (≒KEGG Mapper)
1行に1IDを記載したファイルをアップロード or "Click here ~"から入力 (サンプルデータあり)

Analysis Tools

This tool merges pathway identifier mapping, overrepresentation and expression analysis into a single tabbed data analysis portal, with integrated visualization and summary features.

Select a file from your computer and click on the "Analyze" button to perform the analysis.

Select data file for analysis ファイルを選択 選択されていません Project to human

Click here to paste your data or try example data sets...

Paste the data to analyse

#CBM Uniprot
P01823
Q99758
Q15439
Q43184
Q13444
P82087
P04683
Q7ZSR6
P27548
Q13315
P36543
Q13535
Q96604
Q13145
Q32342

Some examples:

Uniprot accession list
Gene name list
Gene NCBI / Entrez list
Small molecules (ChEBI)
Small molecules (KEGG)
Microarray data
Metabolomics data

Project to human

©2014 統合データベース講習会 Licensed Under CC 表示 2.1

Analysis Toolsマッピング結果

入力した分子に色がつく

Pathways for: Homo sapiens

Event Hierarchy:

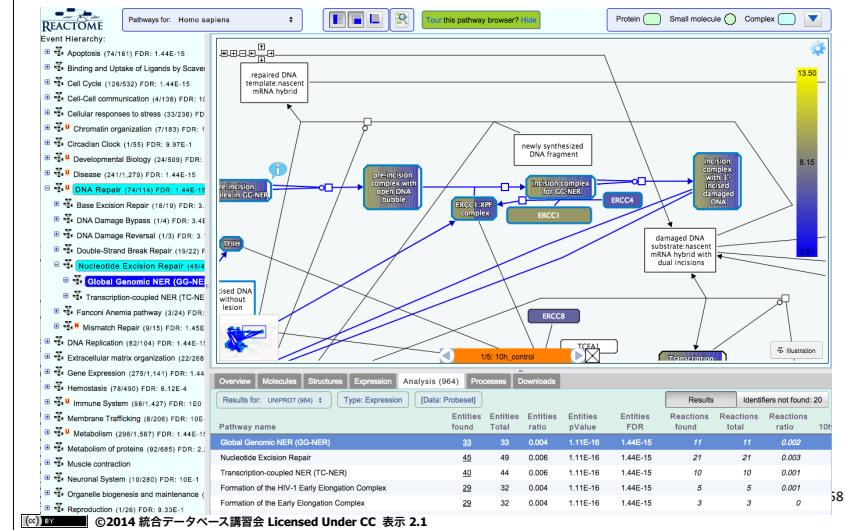
- Apoptosis (6/161) FDR: 3.4E-1
- Binding and Uptake of Ligands by Sceves (3/161) FDR: 9.8E-2
- Cell Cycle (20/532) FDR: 9.8E-2
- Cell-Cell communication (6/181) FDR: 5
- Cellular responses to stress (17/236) FDR: 5
- Chromatin organization (5/93) FDR: 5
- Circadian Clock (25/17) FDR: 1.1E-1
- Developmental Biology (29/599) FDR: 3.6E-1
- DNA Replication (22/29) FDR: 6.8E-1
- Influenza infection (1/12) FDR: 6
- Latent infection of Homo sapiens w/ viruses (1/12) FDR: 6
- Uptake and actions of bacterial toxins (1/12) FDR: 6
- Signaling by EGFR in Cancer (5/15) FDR: 7.2E-1
- Signaling by FGR in disease (36/1) FDR: 8.3E-1
- Diseases associated with visual impairment (3/12) FDR: 8.3E-1
- PISKART Signaling in Cancer (2/4) FDR: 8.3E-1
- Signaling by NOTCH1 in Cancer (5/15) FDR: 8.3E-1
- Abnormal conversion of 2'-oxoglutarate (1/26) FDR: 8.3E-1
- Glycogen storage diseases (3/26) FDR: 8.3E-1
- Defects in vitamin and cofactor metabolism (1/26) FDR: 8.3E-1
- Signaling by TGF-beta Receptor (9/21) FDR: 8.3E-1
- Signaling by WNT in Cancer (6/21) FDR: 8.3E-1
- Processing-defective IgH variants (1/26) FDR: 8.3E-1
- DNA Repair (11/64) FDR: 6.0E-2
- DNA Replication (1/104) FDR: 9.3E-1
- Extracellular matrix organization (15/28) FDR: 9.3E-1

Results for: UNIPROT (39)

Pathway name	Entries found	Total	FDR	pValue	Entities found	Reactions found	Reactions total	Reactions ratio
Signaling by EGFR in Cancer	31	192	2.07E-14	1.11E-16	89	137	1.02	Hor
Signaling by SCF-KT	34	147	0.019	1.11E-16	46	80	0.011	Hor
Signaling by FGFR in disease	35	184	0.024	1.11E-16	101	161	0.026	Hor
Signaling by PGF	34	191	0.025	1.11E-16	39	87	0.013	Hor

©2014 総合データベース講習会 Licensed Under CC 表示 2.1

Analysis Toolsマッピング（定量データ）結果



Species Comparison

Analysis Tools

Select a file for analysis ファイルを選択、選択されていません Project to human Analyse

Click here to paste your data or try example data sets...

Species Comparison

This tool allows you to compare human pathways with those in any of the other species inferred from Reactome by orthology.

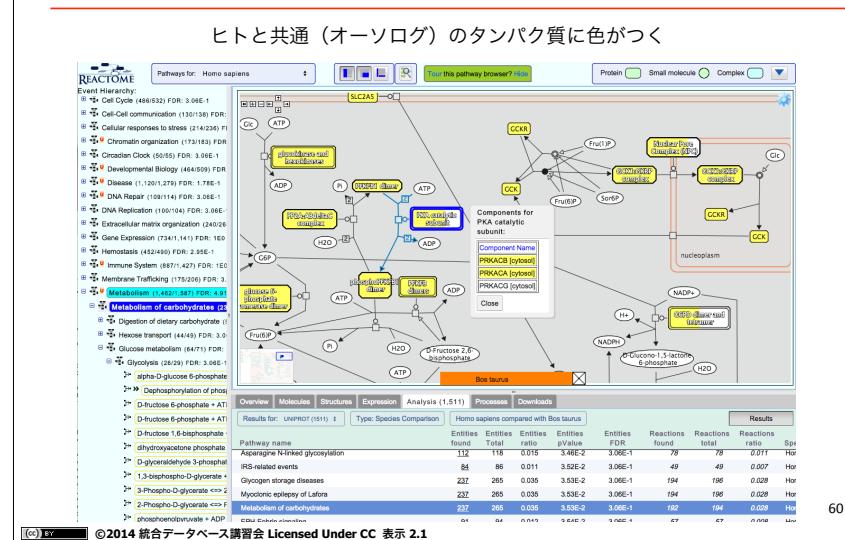
Use the species selector to choose a species... Select a species... Compare

Compare Homo sapiens with... Arabidopsis thaliana

Overview Molecules Structures Expression Analysis (668) Processes Downloads

©2014 総合データベース講習会 Licensed Under CC 表示 2.1

Species Comparisonの結果



BioCyc

<http://biocyc.org/>

61

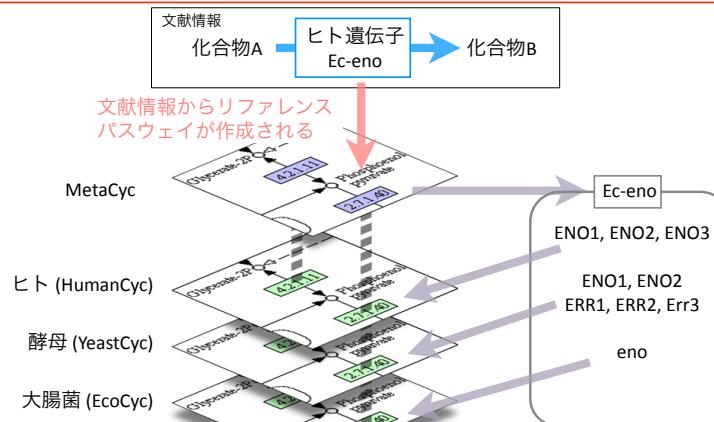
(c) BY ©2014 統合データベース講習会 Licensed Under CC 表示 2.1

BioCyc

The screenshot shows the BioCyc Database Collection homepage. At the top, there's a search bar with the query "Searching Escherichia coli K-12 substr. MG1655 change organism database". Below the search bar, there's a "BioCyc Database Collection" section with a brief description and a "Getting Started" section. To the right, there's a large metabolic pathway diagram for E. coli K-12 substr. MG1655, showing various metabolites and enzymes. Below the diagram, there's a "Build quantitative metabolic models for BioCyc organisms using the MetaFlux metabolic modeling component of Pathway Tools" section. At the bottom, there are navigation links for pages 1 through 10.

62

*Cyc



(c) BY ©2014 統合データベース講習会 Licensed Under CC 表示 2.1

63

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:
EcoCyc, HumanCyc...)

- PATHWAY

↔
• Superpathway

- MODULE

↔
• Pathway

64

(c) BY ©2014 統合データベース講習会 Licensed Under CC 表示 2.1