

相互作用・パスウェイ (KEGG pathway、他)

奥田修二郎

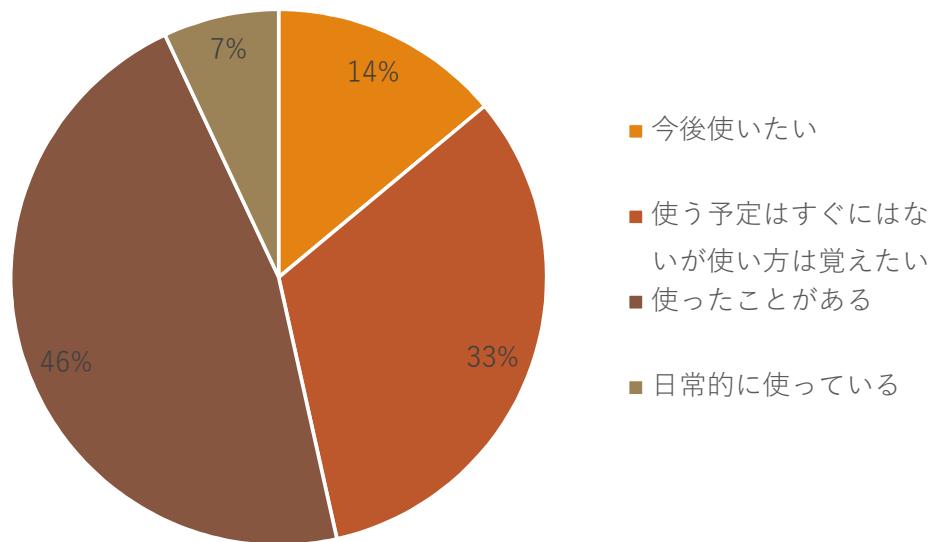
新潟大学大学院医歯学総合研究科

AJACS徳島@大塚製薬株式会社徳島研究所2研ホール

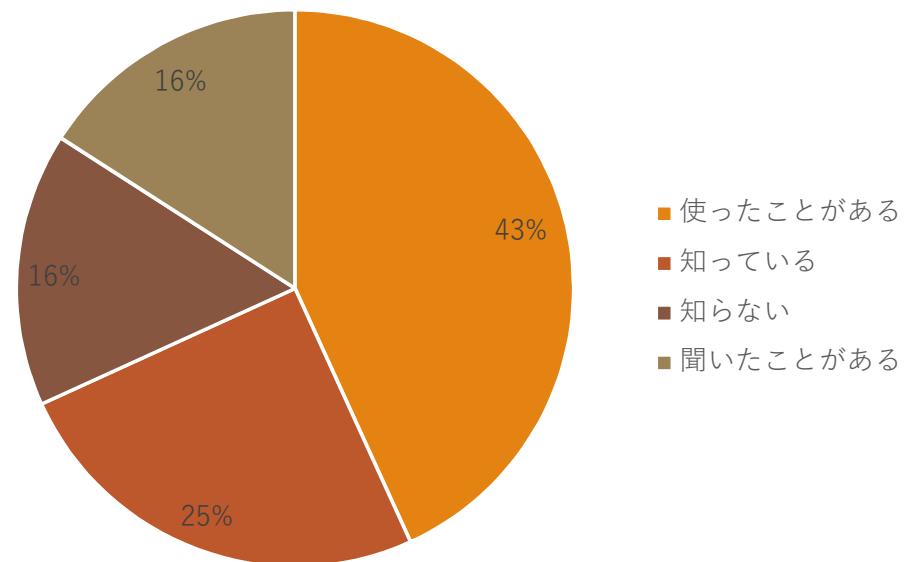
2019年6月5日

事前アンケート

【設問3-1】 パスウェイ関連のデータベースを使ったことがありますか。



【設問3-2】 KEGGを使ったことがありますか。



パスウェイデータベースとは

- パスウェイとは、生体内での遺伝子やタンパク質、その他の化合物等の分子間相互作用を "経路" として表現したものです。相互作用の知識を集積して可視化、電子化したものがパスウェイデータベースと呼ばれています。歴史的には、代謝経路の表現から始まりました
 - ベーリンガーマンハイム社（現ロシュ社）の代謝マップ (1965~)
 - 最初は紙媒体での出版。のちに電子化されています
- 計算機上で表現することで、様々な可視化が行えるようになり、パスウェイ全体を俯瞰したり、一部の相互作用に注目したり、生命現象の理解が行い易くなります。また、データベース化することで網羅的に扱えるようになり、コンピューターでの利用が可能になります
 - ゲノムアノテーションや種間比較、進化解析
 - 遺伝子発現などのエンリッチメント解析
 - モデル化、シミュレーション、予測

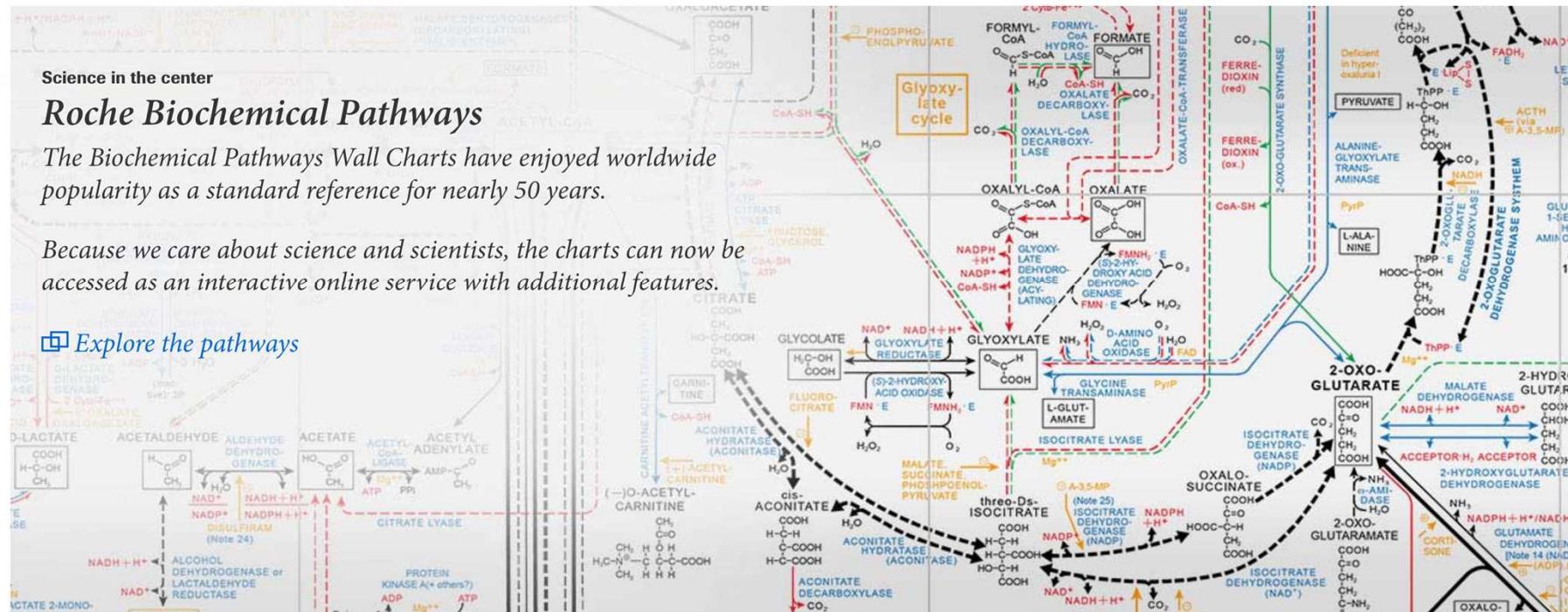
Science in the center

Roche Biochemical Pathways

The Biochemical Pathways Wall Charts have enjoyed worldwide popularity as a standard reference for nearly 50 years.

Because we care about science and scientists, the charts can now be accessed as an interactive online service with additional features.

Explore the pathways



Share Rate

Mapping the paths of life

Biochemical Pathways provide an overview of the chemical reactions of cells in various species and organs. Dr. Michal first compiled the Pathways Chart in 1965 and has been fine-tuning it ever since. Today, and with the collaboration of Roche, the two enormous posters can be found hanging in just about every research institute from Argentina to New Zealand.

“ You have to be someone with tenacity and patience. And love for science

Dr Gerhard Michal
Editor of the Roche Biochemical Pathways

By the numbers

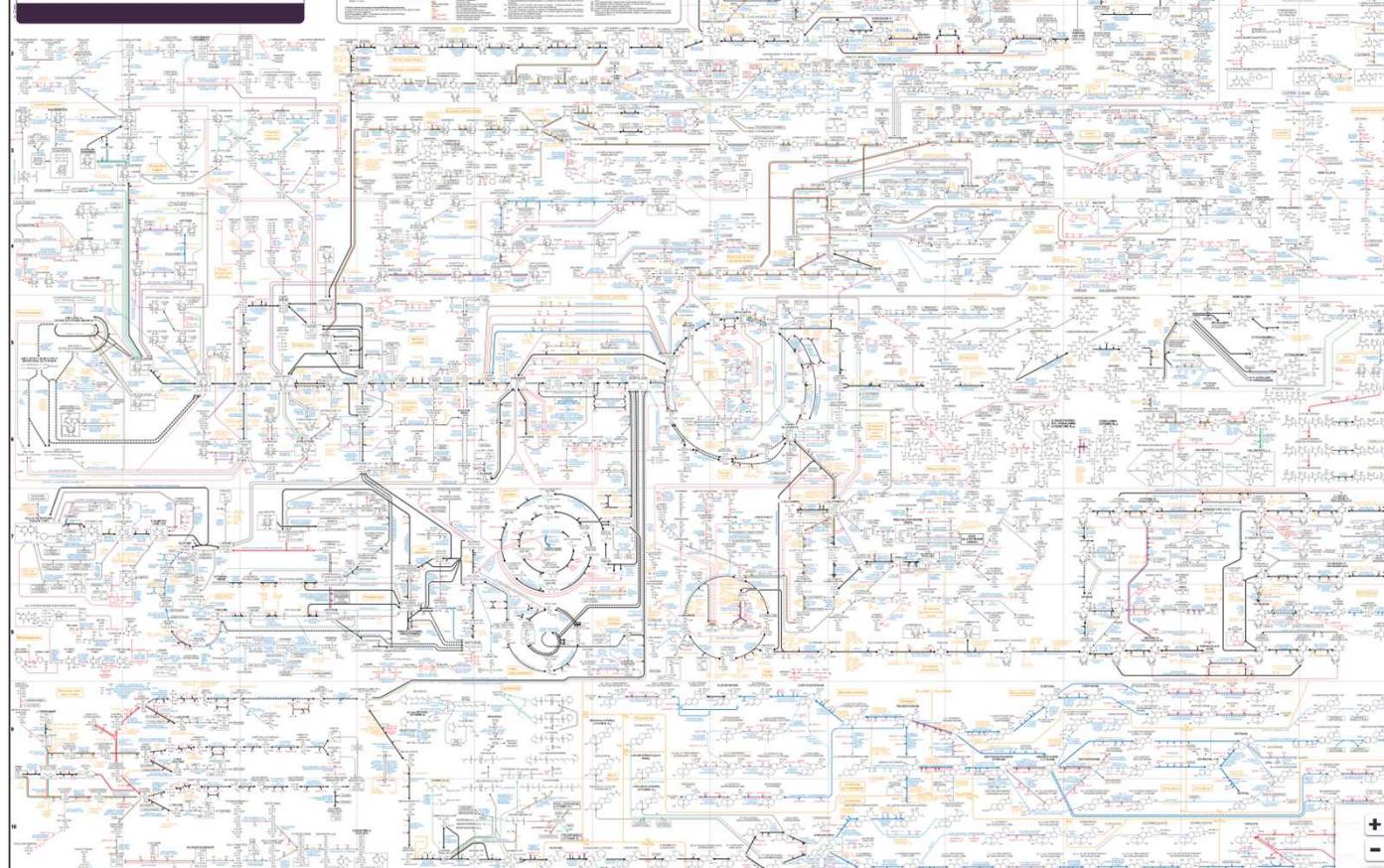
49 years
continuously fine-tuned by the editor himself.

Part 1: Metabolic Pathways Part 2: Cellular and Molecular Processes

Click here to open the menu

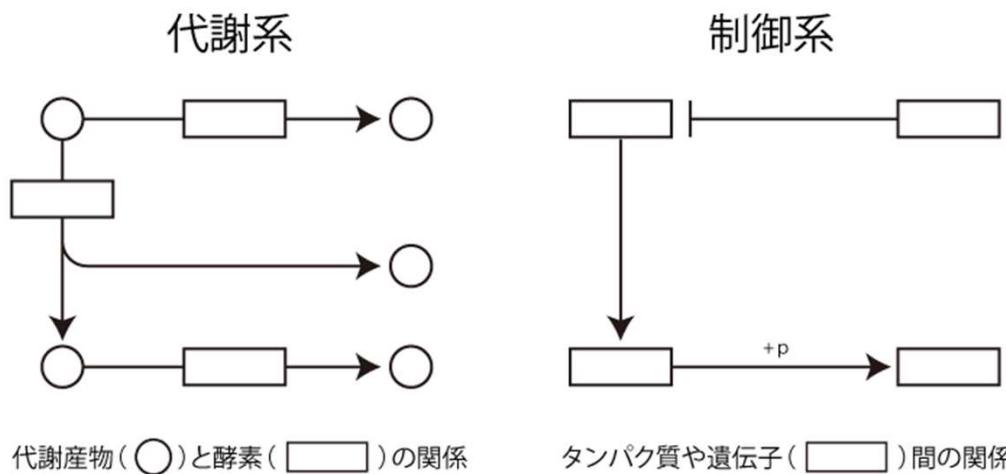
Part 1
Metabolic Pathways

Roche Biochemical Pathways
4th Edition, Part 1 - Editor: Gerhard Michal



パスウェイデータベースでの表現

- 見やすくするために、ダイアグラムで表現されていることが多いっています。
- KEGG Pathway での表現例
 - 代謝パスウェイでは代謝産物をノード、酵素反応をエッジとして表現されています
 - 制御系ではタンパク質や遺伝子、その他の小分子をノード、その関係性（活性化、抑制、リン酸化など）がエッジとして表現されています



パスウェイデータベースのデータ形式

- 計算機での取り扱いを目的として XML (Extensible Markup Language) で記述されていることが多いっています。
 - [KGML](#) (KEGG Markup Language) は分子間の関係とダイアグラムのレイアウトを取り扱うための KEGG 独自のフォーマット
 - [SBML](#) (Systems Biology Markup Language)、[CellML](#)、[CSML](#) (Cell System Markup Language) はパスウェイのシミュレーションやモデリングを行うためのフォーマット
 - [PSI-MI](#) (Proteomics Standards Initiative Molecular Interaction XML Format) はタンパク質間相互作用を記述するためのフォーマット
 - [BioPAX](#) (Biological Pathways Exchange) は様々なパスウェイデータを統合したり、データ交換を行うことを目的として策定された標準化を目指したフォーマット
- これらのデータ形式を扱うことのできるネットワーク可視化ソフトウェアには [Cytoscape](#) や [VisANT](#) などがあります。
 - [AJACS58 : Cytoscapeを使ったデータの可視化](#)
- どのパスウェイデータベースを研究に使えば良いかは、対象生物や対象パスウェイ、目的によって異なってきます。
- 今回は BioCyc, Reactome, KEGG PATHWAY のブラウザでの使い方を紹介します。

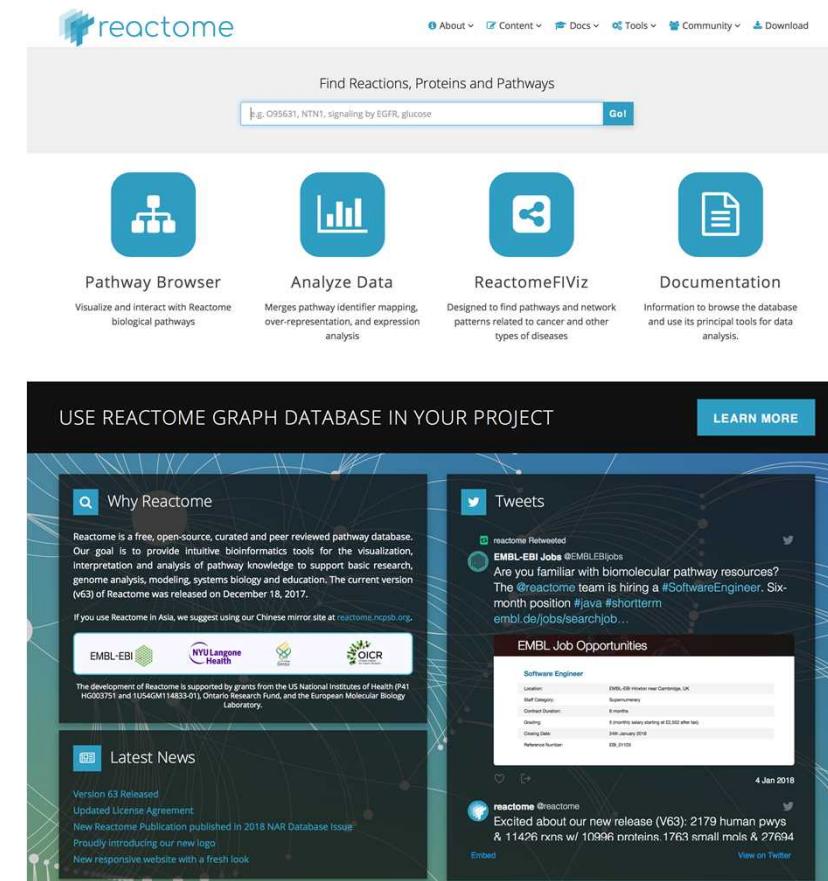
BioCyc

- ウェブサイト：<https://biocyc.org/>
- 開発：SRIインターナショナル
(Stanford Research Institute)
- 対象：大腸菌からヒトまで、異株を含めて 7,600 種以上
 - 専門家が手作業で作成した文献ベースのデータ + 自動ツール
 - 代謝パスウェイ、制御系
- 利用：アカデミックフリー
- データ形式：BioPAX

The screenshot shows the BioCyc Database Collection homepage. At the top, there's a navigation bar with links for LOGIN, Why Log in?, Create New Account, Site Search, Quick Search, Gene Search, and a search bar for "Enter a gene, protein, metabolite or pathway". Below the search bar, it says "Searching Escherichia coli K-12 substr. MG1655 (EcoCyc) change organism database". The main content area is titled "BioCyc Database Collection". It features a diagram illustrating metabolic pathways, specifically showing three routes from a glucose derivative to a phenylpyruvate derivative: "Long Route", "Loss of Atoms", and "Best Route: Atoms Conserved". The "Best Route" is highlighted in blue. To the right of the diagram, there's a section titled "RouteSearch: Search for Paths through the Metabolic Network" with a "Learn More" button and a page navigation bar (1-10). Below the diagram, there's a "Tools" section with a bulleted list of features: Genome browser, Display of individual metabolic pathways, and of full metabolic maps; Multiple analysis methods for user-supplied omics and multi-omics datasets including painting onto metabolic maps, regulatory maps, and genome maps; Store groups of genes and pathways in your account as SmartTables; share, analyze, transform those groups; Comparative analysis tools. There's also a "BioCyc Databases" section with a note about tiered databases and funding information from NIH.

Reactome

- ウェブサイト：<https://www.reactome.org/>
- 開発：EMBLE-EBI 他
- 対象：ヒトを中心に脊椎動物、酵母、植物、19種
 - ヒト：専門家が手作業で作成した文献ベースのデータ
 - その他：計算機での推定
(<https://reactome.org/documentation/inferred-events/>)
 - 代謝パスウェイ、シグナル伝達系、他
- 利用：フリー
- データ形式：BioPAX, SBML



KEGG PATHWAY

- ウェブサイト：<https://www.kegg.jp/>
- 開発：京都大学
- 対象：ゲノムの決まった全生物種（異株を含む）4,200種以上（>300真核生物、>3,700真正細菌、>220古細菌）、真核ドラフトゲノム25種、環境メタゲノム300サンプル、生体メタゲノム700サンプル
 - 専門家が手作業で作成した文献ベースのデータ + 自動ツール
 - リファレンスパスウェイ：専門家が手作業で文献ベースから作成
 - 生物種パスウェイ
 - 自動ツールでリファレンスパスウェイから作成し、手作業でキュレーション
 - 自動ツールで作成（自動ツールにも段階があります）
 - 代謝パスウェイ、シグナル伝達系、他
- 利用：アカデミックフリー
- データ形式：KGML
 - [KCPAVS KEGG-XML converter](#) などで代謝パスウェイ、シグナル伝達などの多くのパスウェイを標準形式に変換可能
 - [KEGGscape] (<http://apps.cytoscape.org/apps/keggscape>) でネットワーク可視化ソフト Cytoscape に読み込み可能

KEGGはデータベースの集合

- KEGG2をクリック
- KEGG PATHWAY を含むシステム情報データベースの他に、遺伝情報、化学情報、健康情報などのデータベースがリンクしています。

The screenshot shows the KEGG homepage with a navigation bar at the top. The main content area features a banner for 'Kyoto Encyclopedia of Genes and Genomes' and a section about the KEGG database. Below this is a 'Data-oriented entry points' section with links to KEGG PATHWAY, BRITE, MODULE, ORTHOLOGY, GENOME, GENES, COMPOUND, GLYCAN, REACTION, ENZYME, NETWORK, DISEASE, DRUG, and MEDICUS. A red box highlights the 'KEGG2' link in the 'Main entry point to the KEGG web service' section.

The screenshot shows a search results page for 'KEGG' with a table titled 'Data-oriented entry points'. The table maps categories to specific KEGG entries and their DBGET search equivalents. The categories include Systems information, Genomic information, Chemical information, and Health information. The table lists various KEGG databases like PATHWAY, BRITE, MODULE, KO, GENES, GENOME, GENES, COMPOUND, GLYCAN, REACTION, ENZYME, NETWORK, DISEASE, DRUG, and MEDICUS, along with their corresponding DBGET search terms.

Category	Entry Point	Content	DBGET Search
Systems information	KEGG PATHWAY KEGG BRITE KEGG MODULE	KEGG pathway maps BRITE hierarchies and tables KEGG modules	PATHWAY BRITE MODULE
	KO (KEGG Orthology)	Functional orthologs	ORTHOLOGY
Genomic information	KEGG GENOME KEGG GENES KEGG SSDB	KEGG organisms (complete genomes) Genes and proteins GENE sequence similarity	GENOME GENES
Chemical information (KEGG LIGAND)	KEGG COMPOUND KEGG GLYCAN KEGG REACTION KEGG ENZYME	Small molecules Glycans Reactions and reaction classes Enzyme nomenclature	COMPOUND GLYCAN REACTION RCLASS ENZYME
Health information	KEGG NETWORK <i>New!</i> KEGG DISEASE KEGG DRUG KEGG ENVIRON KEGG MEDICUS	Disease-related network elements Human diseases Drugs and drug groups Health related substances Japanese drug labels (JAPIC) FDA drug labels (DailyMed)	NETWORK VARIANT DISEASE DRUG DGROUP ENVIRON

対象生物種を見る

- データベースのテーブルの下、[KEGG organisms](#) をクリック

Subject-oriented entry points

KEGG Cancer	Cancer research
KEGG Pathogen	Pathogen genomes and antimicrobial resistance
KEGG Virus	Virus research
KEGG Plant	Plant research
KEGG Annotation	KO annotation of genes and proteins
KEGG SeqData	KEGG sequence data collection
KEGG RModule	Architecture of metabolic networks

Organism-specific entry points

[KEGG organisms](#) the list of currently available organisms

Select Organism (examples) hsa mmu sce eco bsu syn

Organism group	Pangenome		Genome
Eukaryotes	Animals Plants Fungi Protists	Genus level pangenomes	Species level pangenomes
Prokaryotes	Bacteria Archaea		KEGG organisms

See KEGG GENOME for more details.

KEGG Organisms

- KEGG では 3-4 文字の独自の生物種コードを使用している

Eukaryotes

Category	Organisms	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
	nle Nomascus leucogenys (northern white-cheeked gibbon)	RefSeq
	mcc Macaca mulatta (rhesus monkey)	RefSeq
	mcf Macaca fascicularis (crab-eating macaque)	RefSeq
	csab Chlorocebus sabaeus (green monkey)	RefSeq
	rro Rhinopithecus roxellana (golden snub-nosed monkey)	RefSeq
	rbb Rhinopithecus bieti (black snub-nosed monkey)	RefSeq
	cjc Callithrix jacchus (white-tufted-ear marmoset)	RefSeq
	sbq Saimiri boliviensis boliviensis (Bolivian squirrel monkey)	RefSeq
	mmu Mus musculus (mouse)	RefSeq
	rno Rattus norvegicus (rat)	RefSeq
	cge Cricetulus griseus (Chinese hamster)	RefSeq
	ngi Nannospalax galili (Upper Galilee mountains blind mole rat)	RefSeq
	hgl Heterocephalus glaber (naked mole rat)	RefSeq
	ccan Castor canadensis (American beaver)	RefSeq
	ocu Oryctolagus cuniculus (rabbit)	RefSeq
	tup Tupaia chinensis (Chinese tree shrew)	RefSeq

KEGG Organisms

- 生物種コードのリンクをクリックすると、種の情報が表示されます
 - Annotation
 - manual : 手作業によるアノテーション（ヒト(hsa)等）
 - KOALA : SSEARCH ベースの自動ツールによるアノテーション（ゴリラ(ggo)等）
 - BlastKOALA : BLAST ベースの自動ツールによるアノテーション（ドラフトゲノム）
 - GhostKOALA : GhostX ベースの自動ツールによるアノテーション（メタゲノム）

The image displays two side-by-side screenshots of the KEGG Organisms interface, illustrating the genome information for two different species: Homo sapiens (human) and Pan paniscus (bonobo).

Homo sapiens (human) Screenshot:

- Annotation:** manual (highlighted with a red box)

Pan paniscus (bonobo) Screenshot:

- Annotation:** KOALA (highlighted with a red box)

In both screenshots, the "Annotation" field is explicitly highlighted with a red box, demonstrating the difference in annotation methods between the two species.

データベースリスト

- Genomes : 主に NCBI Refseq、GenBank に登録された生物種
- Species : 異株を一つにまとめたデータベース
- Genus : Genus レベルでまとめたデータベース
- Draft : Genomes に入っていない真核生物
- Meta : メタゲノム



KEGG Organisms: Complete Genomes

Eukaryotes: 429 Bacteria: 4521 Archaea: 263

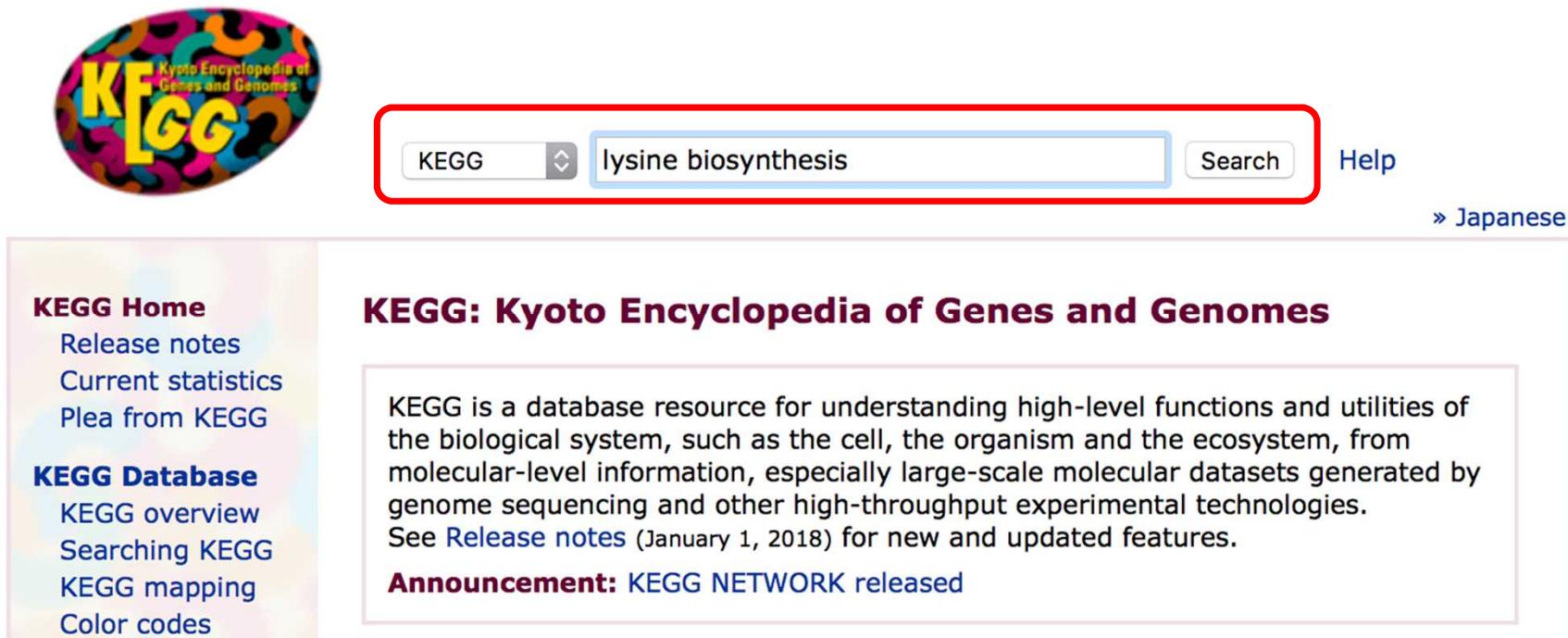
[Genomes | Species | Genus | Viruses | Meta]

Eukaryotes

Category	Organisms		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
	nle	Nomascus leucomoanus (northern white-cheeked gibbon)	RefSeq

KEGGパスウェイマップを見る

- [トップページ](#) 上方の検索ボックスで "lysine biosynthesis" や "glycolysis" やなどの生命現象関連の単語を入力し、Search ボタンをクリック



The screenshot shows the KEGG homepage. On the left, there is a navigation sidebar with links to "KEGG Home", "Release notes", "Current statistics", "Plea from KEGG", "KEGG Database", "KEGG overview", "Searching KEGG", "KEGG mapping", and "Color codes". The main content area features the KEGG logo and a search bar. The search bar has the text "lysine biosynthesis" entered into it, which is highlighted with a red rectangle. To the right of the search bar are buttons for "Search" and "Help", and a link to "» Japanese". Below the search bar, the text "KEGG: Kyoto Encyclopedia of Genes and Genomes" is displayed, followed by a descriptive paragraph about the KEGG database and an announcement about the KEGG NETWORK release.

lysine biosynthesis

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 other high-throughput experimental technologies. See [Release notes](#) (January 1, 2018) for new and updated features.

Announcement: KEGG NETWORK released

KEGGパスウェイマップ

- KEGG データベース全体でヒットしたエントリーが全てリストアップされ、KEGG PATHWAY にヒットがあれば、一番上に表示されます。

The screenshot shows the KEGG search interface with the query "lysine biosynthesis". The results page displays two main sections: KEGG PATHWAY and KEGG MODULE.

KEGG PATHWAY

- map00300 Lysine biosynthesis (highlighted with a red box)
- map01064 Biosynthesis of alkaloids derived from ornithine, lysine and nicotinic acid

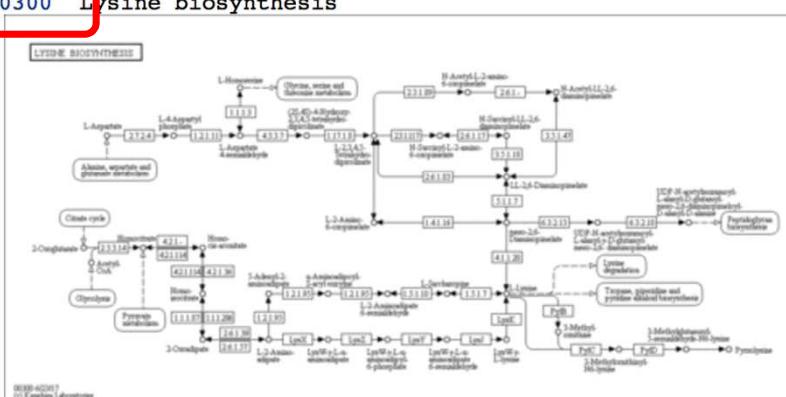
KEGG MODULE

- M00016 Lysine biosynthesis, succinyl-DAP pathway, aspartate => lysine
- M00030 Lysine biosynthesis, AAA pathway, 2-oxoglutarate => 2-aminoadipate => lysine
- M00031 Lysine biosynthesis, mediated by LysW, 2-aminoadipate => lysine
- M00433 Lysine biosynthesis, 2-oxoglutarate => 2-oxoadipate
- M00525 Lysine biosynthesis, acetyl-DAP pathway, aspartate => lysine

KEGGパスウェイマップ

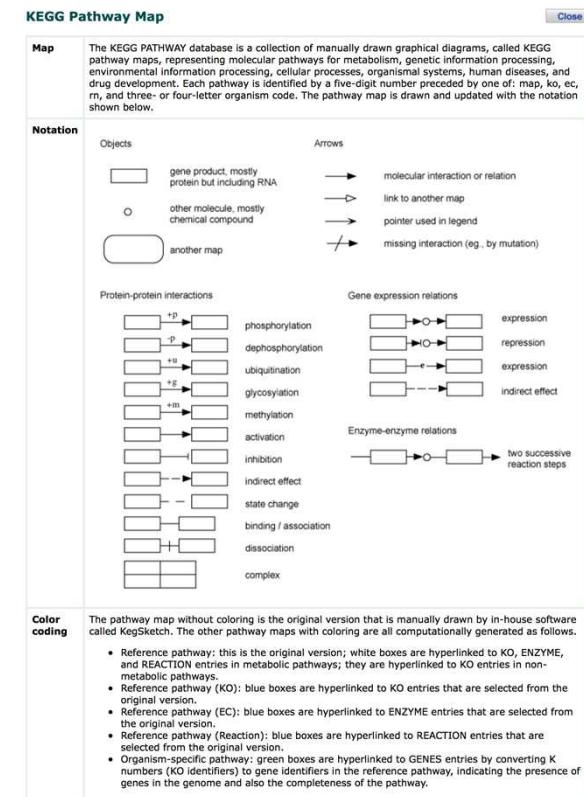
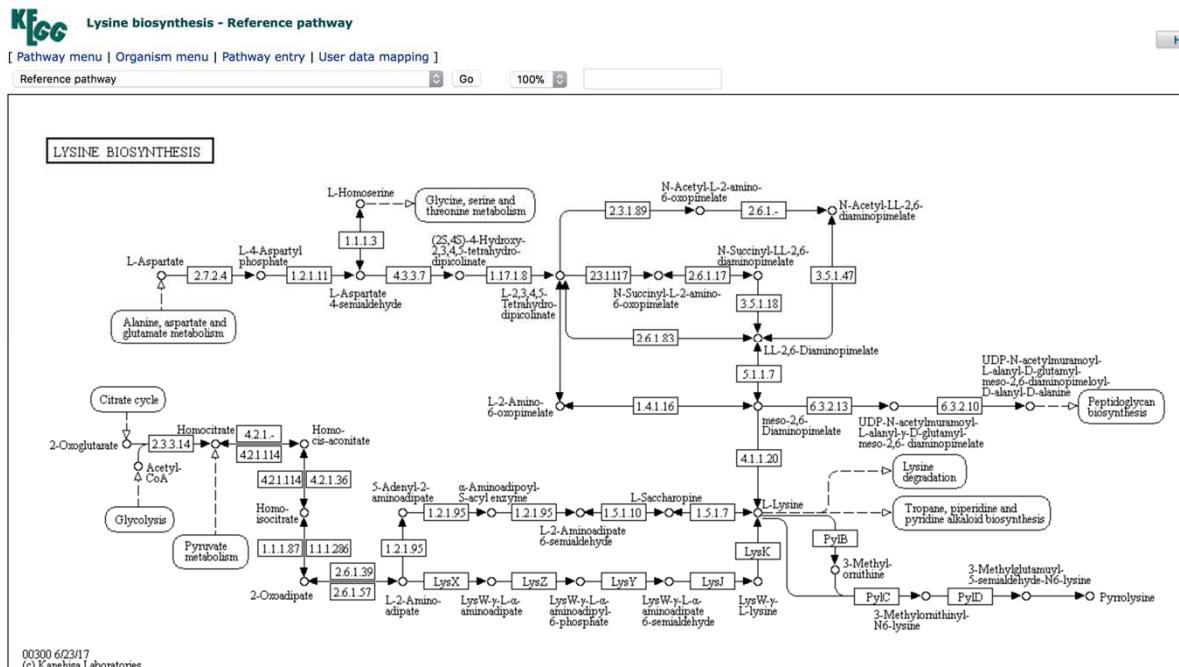
- 各パスウェイの情報が表示されます。
 - KEGGにおけるパスウェイの最小単位で、ダイアグラム画像を“マップ”と呼んでいます。

KEGG PATHWAY: map00300 Help

Entry	map00300 Pathway
Name	Lysine biosynthesis
Class	Metabolism; Amino acid metabolism
DRIVE Hierarchy	
Pathway map map00300 Lysine biosynthesis	
	
All organisms Ortholog table	
Module	M00016 Lysine biosynthesis, succinyl-DAP pathway, aspartate => lysine [PATH:map00300] M00030 Lysine biosynthesis, AAA pathway, 2-oxoglutarate => 2-aminoadipate => lysine [PATH:map00300]

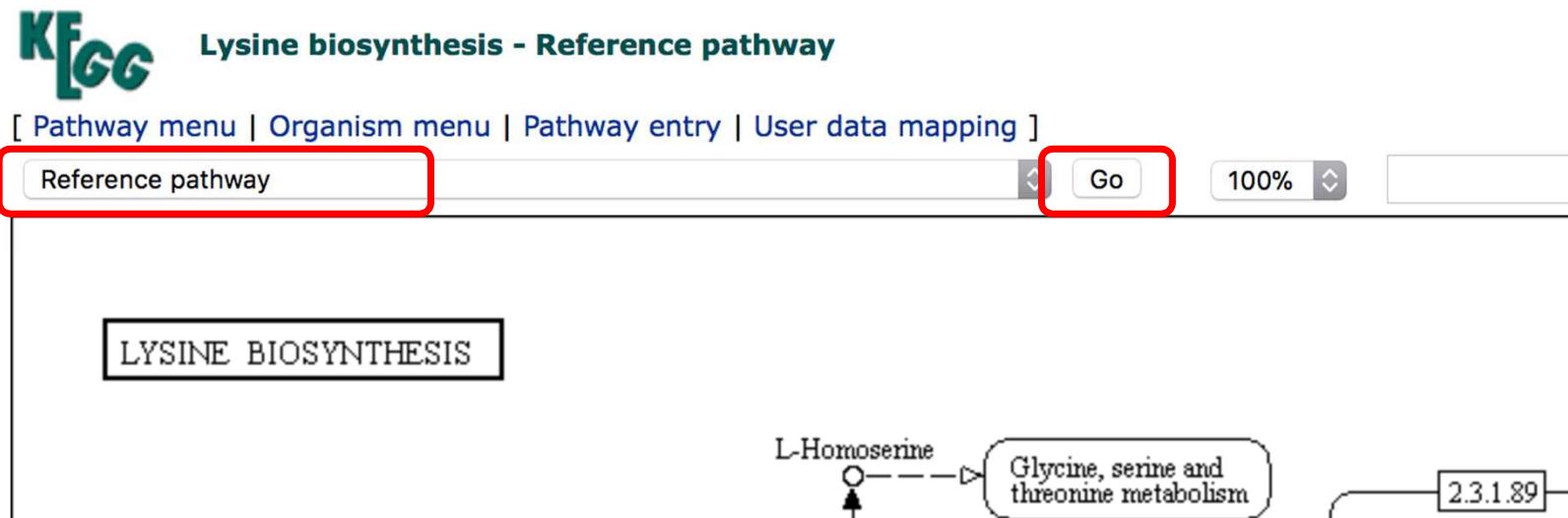
KEGGパスウェイマップ

- この色のついていない白いダイアグラムが、専門家が手作業で文献ベースから作成したリファレンスパスウェイになります
 - ボックスが遺伝子やタンパク質などの配列情報、丸が代謝産物、環境物質などの化合物
 - 各図形の説明は右上の Help から見られます



好きな生物のパスウェイを見る

- プルダウンメニューから好きな生物を選択して Go をクリック
 - リストが多すぎて選びにくいので
 - < Sort below by alphabet > を選択して Go をクリックでリストをソート
 - < Set personalized menu > を選択して Go をクリックでポップアップウィンドウからリストの絞り込み



別のアプローチ

- 種、属でまとめたパスウェイ、ドラフトゲノム、メタゲノムのパスウェイはここからは選べないので、生物種リストのページから、種のページ、パスウェイリストへ移動する必要があります。

Prokaryotes

Category	Organism
	Escherichia coli K-12 MG1655
	Escherichia coli K-12 W3110
	Escherichia coli K-12 DH10B
	Escherichia coli BW2952
	Escherichia coli K-12 MDS42
	Escherichia coli O157:H7 EDL933 (EHEC)
	Escherichia coli O157:H7 Sakai (EHEC)
	Escherichia coli O157:H7 EC4115 (EHEC)
	Escherichia coli O157:H7 TW14359 (EHEC)
	Escherichia coli O157:H7 Xuzhou21 (EHEC)
	Escherichia coli O26:H11 11368 (EHEC)
	Escherichia coli O111:H- 11128 (EHEC)

KEGG Homo sapiens (human)

Genome info Pathway map Brite hierarchy Module Genome map Blast Taxonomy

Search genes: Go Clear

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

2009	GenBank
2012	GenBank
2009	GenBank
2009	GenBank

Lysine biosynthesis マップ

- 一部のボックスが緑色で塗られる、その生物（またはサンプル）の持つ遺伝子を示しています。

00240 Pyrimidine metabolism

Amino acid metabolism

00250 Alanine, aspartate and glutamate metabolism

00260 Glycine, serine and threonine metabolism

00270 Cysteine and methionine metabolism

00280 Valine, leucine and isoleucine degradation

00290 Valine, leucine and isoleucine biosynthesis

00300 Lysine biosynthesis

00310 Lysine degradation

00220 Arginine biosynthesis

00330 Arginine and proline metabolism

00340 Histidine metabolism

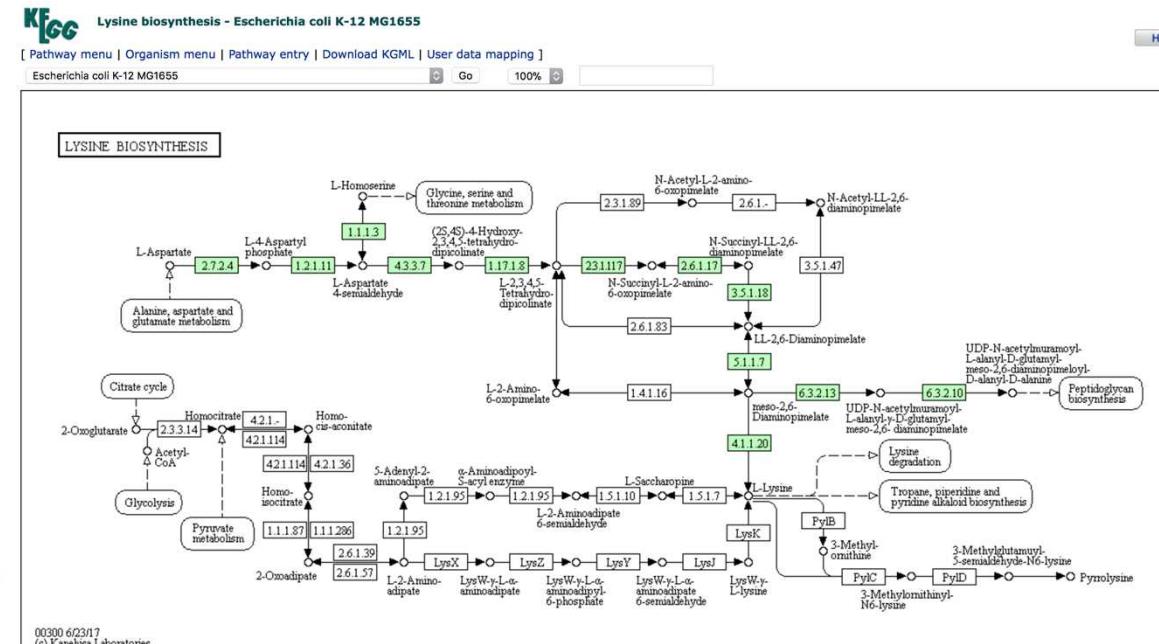
00350 Tyrosine metabolism

00360 Phenylalanine metabolism

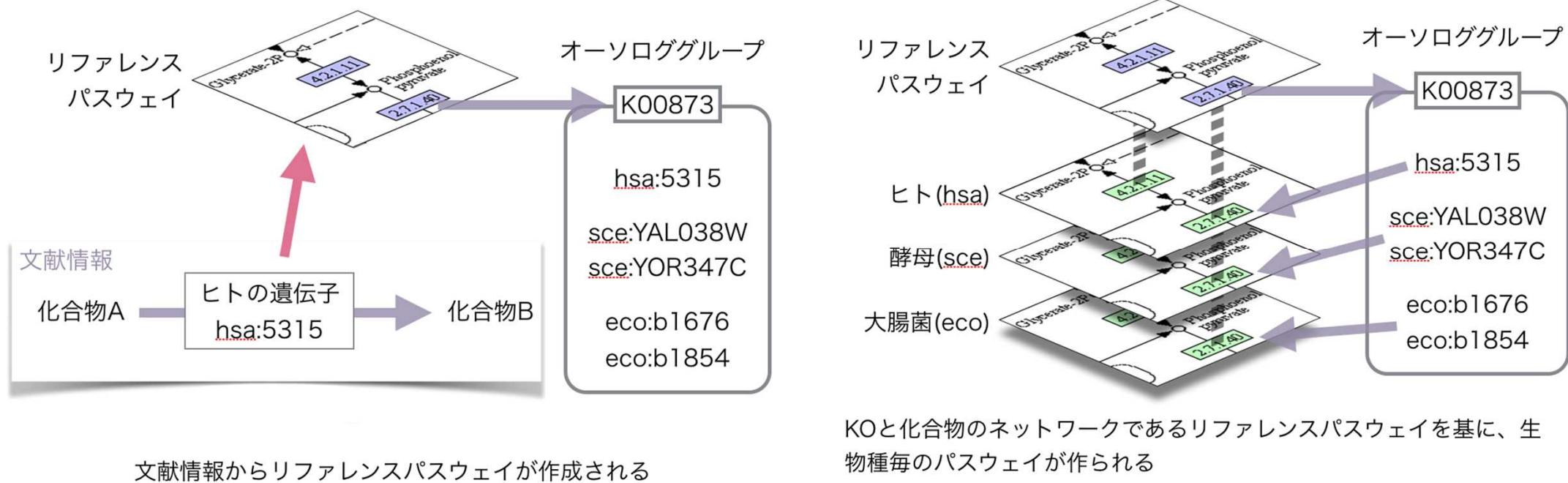
00380 Tryptophan metabolism

00400 Phenylalanine, tyrosine and tryptophan biosynthesis

Metabolism of other amino acids



リファレンスパスウェイと種毎のパスウェイの関係



Overviewマップを見る

- <https://www.kegg.jp/kegg/pathway.html>
- 1.0 Global and overview maps の [Metabolic pathways](#) をクリック
 - 右の [KEGG Atlas] は Java で動くビューワーで、自由度が少し高い分、動作が重たい
- 左にモジュールのリスト（KEGG におけるパスウェイの小さい機能単位）、右にマップが表示
- 機能単位毎にパスウェイを強調表示できる

1. Metabolism

1.0 Global and overview maps

01100	Metabolic pathways	[KEGG Atlas]
01110	Biosynthesis of secondary metabolites	[KEGG Atlas]
01120	Microbial metabolism in diverse environments	[KEGG Atlas]
01130	Biosynthesis of antibiotics	[KEGG Atlas]
01200	Carbon metabolism	[KEGG Atlas]
01210	2-Oxocarboxylic acid metabolism	[KEGG Atlas]
01212	Fatty acid metabolism	[KEGG Atlas]
01230	Biosynthesis of amino acids	[KEGG Atlas]
01220	Degradation of aromatic compounds	[KEGG Atlas]

1.1 Carbohydrate metabolism

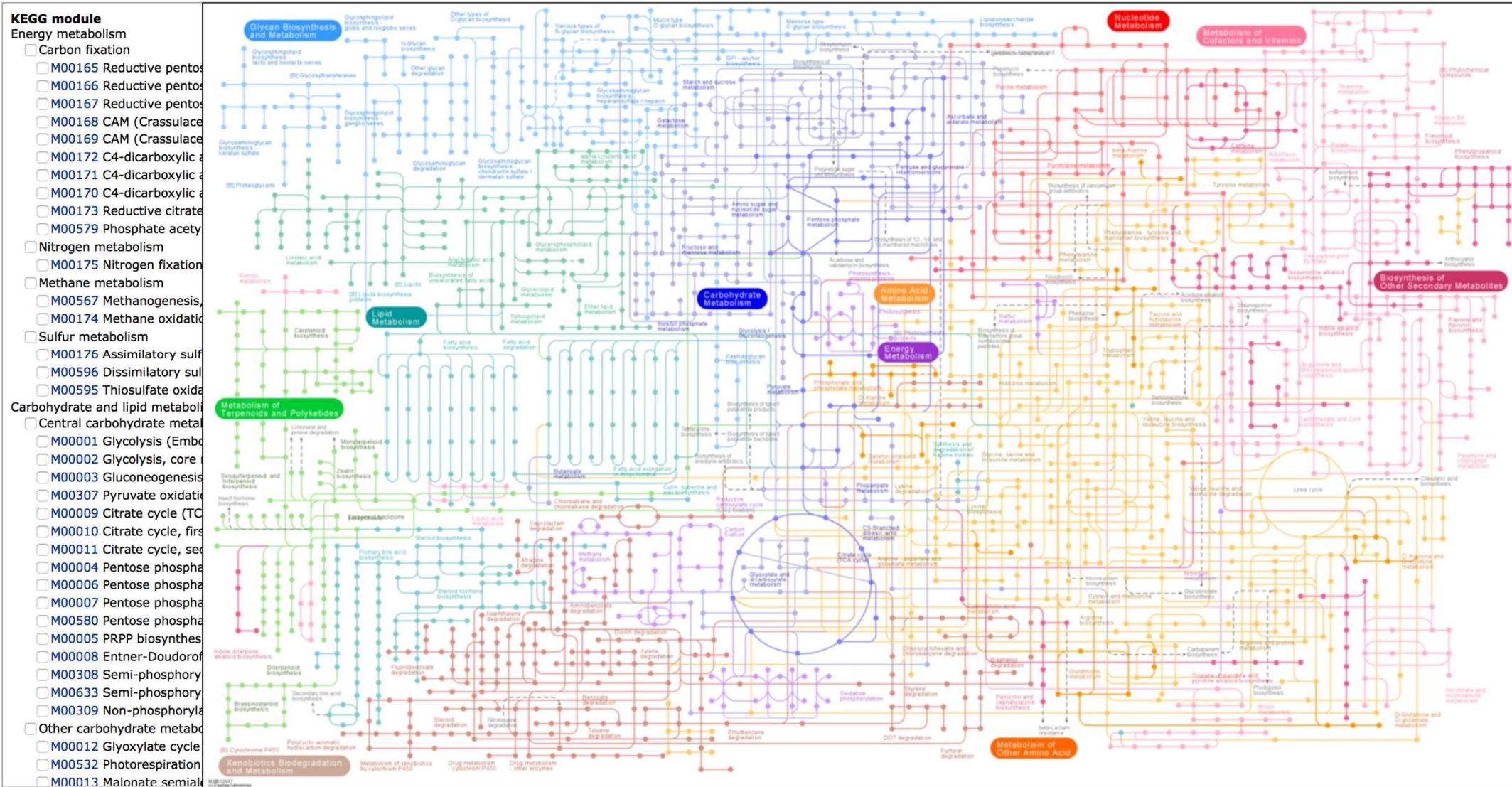
00010	Glycolysis / Gluconeogenesis
00020	Citrate cycle (TCA cycle)

[Pathway menu | Organism menu | Pathway entry | Hide module list | User data mapping | Image (png) file]

Reference pathway

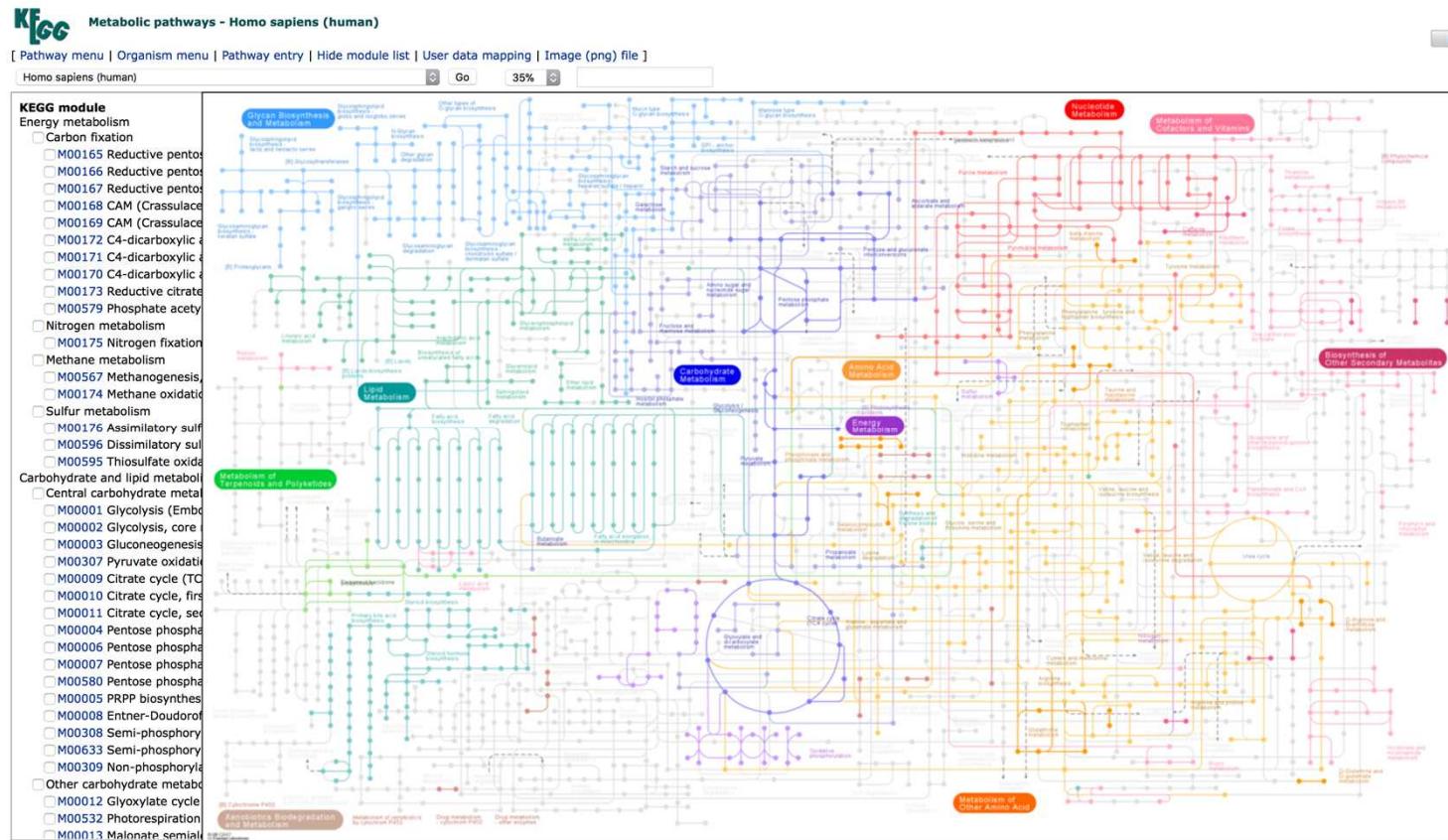
Go

35%



生物種毎のOverviewマップを見る

- ・プルダウンメニューから生物を選択し、Go をクリック
- ・生物の持っていない経路は灰色になる



ヒトの疾患パスウェイを見る

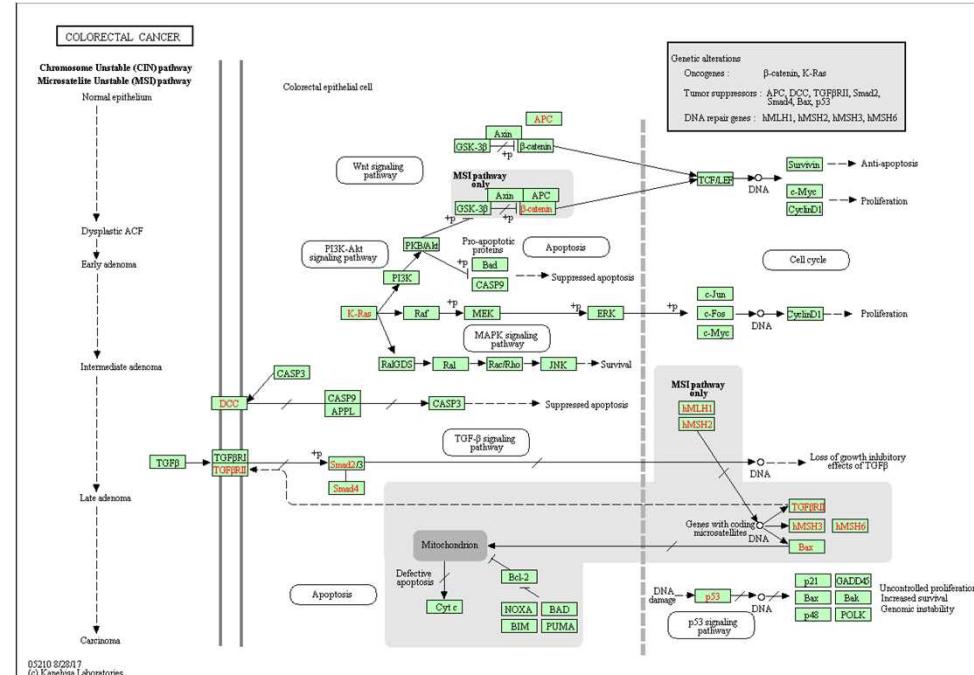
- <https://www.genome.jp/kegg/pathway.html#disease>
- がん、免疫系疾患、神経変性疾患など多因子性の疾患
- 好きな疾患パスウェイをクリック（例：[大腸がん](#)）
 - 赤字の遺伝子が疾患の病因遺伝子を示しています。

6.1 Cancers: Overview

- 05200 Pathways in cancer
- 05230 Central carbon metabolism in cancer
- 05231 Choline metabolism in cancer
- 05202 Transcriptional misregulation in cancer
- 05206 MicroRNAs in cancer
- 05205 Proteoglycans in cancer
- 05204 Chemical carcinogenesis
- 05203 Viral carcinogenesis

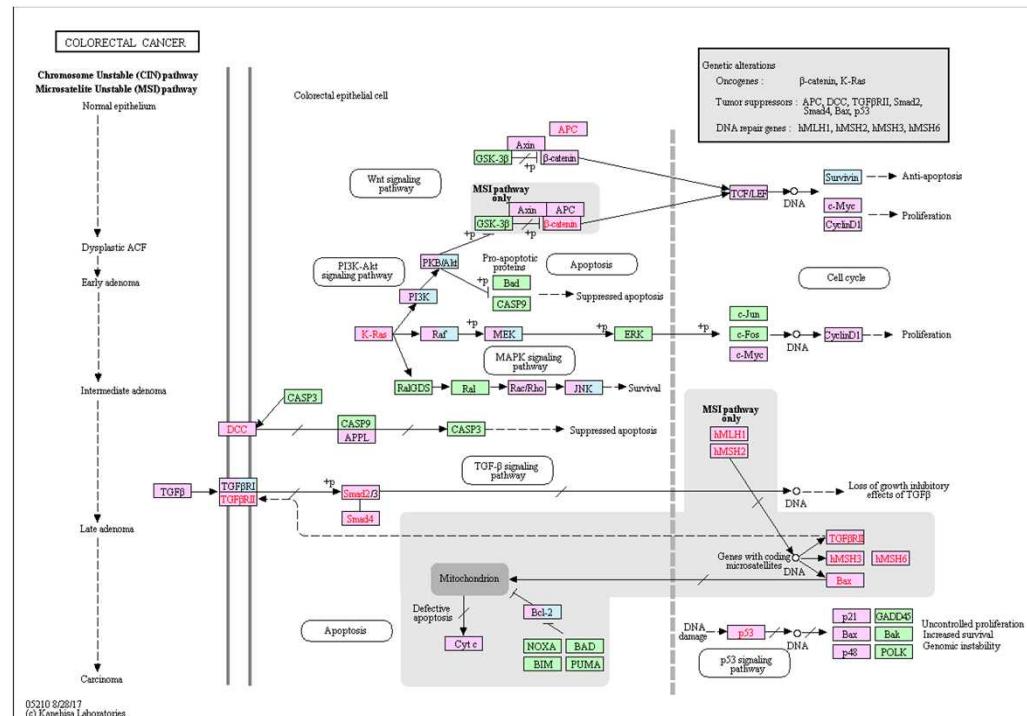
6.2 Cancers: Specific types

- 05210 Colorectal cancer
- 05212 Pancreatic cancer
- 05225 Hepatocellular carcinoma *New!*
- 05226 Gastric cancer *New!*



ヒトの疾患パスウェイ

- プルダウンメニューから Homo sapiens (human) + Disease/drug を選択
 - ピンクのボックスは何らかの疾患で病因遺伝子となっている遺伝子を示しています
 - ライトブルーのボックスは何らかの疾患で医薬品のターゲットとなっている遺伝子を示しています
- このように、KEGG では正常な状態のパスウェイの他に、病原因子や医薬品、そのターゲットなどの情報もパスウェイとして表現されています



種間比較をする

- <https://www.genome.jp/kegg/kegg2.html>
- KEGG GENOMEのリンクをクリック
- KEGG Mapping for Genome Comparison and Combination のテキストボックスに "eco ecs" と入力して Go をクリック
 - eco : 非病原性大腸菌 k-12
 - ecs : 病原性大腸菌 O157

Data-oriented entry points

Category	Entry Point	Content	DBGET Search
Systems information	KEGG PATHWAY KEGG BRITE KEGG MODULE	KEGG pathway maps BRITE hierarchies and tables KEGG modules	PATHWAY BRITE MODULE
Genomic information	KO (KEGG Orthology) KEGG GENOME KEGG GENES KEGG SSDB	Functional orthologs KEGG organisms (complete genomes) Genes and proteins GENES sequence similarity	ORTHOLOGY GENOME GENES
Chemical information (KEGG LIGAND)	KEGG COMPOUND KEGG GLYCAN KEGG REACTION KEGG ENZYME	Small molecules Glycans Reactions and reaction classes Enzyme nomenclature	COMPOUND GLYCAN REACTION RCLASS ENZYME
Health information	KEGG NETWORK <i>New!</i> KEGG DISEASE KEGG DRUG KEGG ENVIRON KEGG MEDICUS	Disease-related network elements Human diseases Drugs and drug groups Health related substances Japanese drug labels (JAPIC) FDA drug labels (DailyMed)	NETWORK VARIANT DISEASE DRUG DGROUP ENVIRON

KEGG is developed by Kanehisa Laboratories. See [Kanehisa et al. \(2017\)](#) for updates of KEGG.

KEGG Mapping for Genome Comparison and Combination

An organism group may be defined to compare or combine KEGG organisms, other organisms and environmental samples, enabling the analysis of combined pathway maps for the group.

Define organism group (enter organism codes or T numbers):

(Genome comparison)

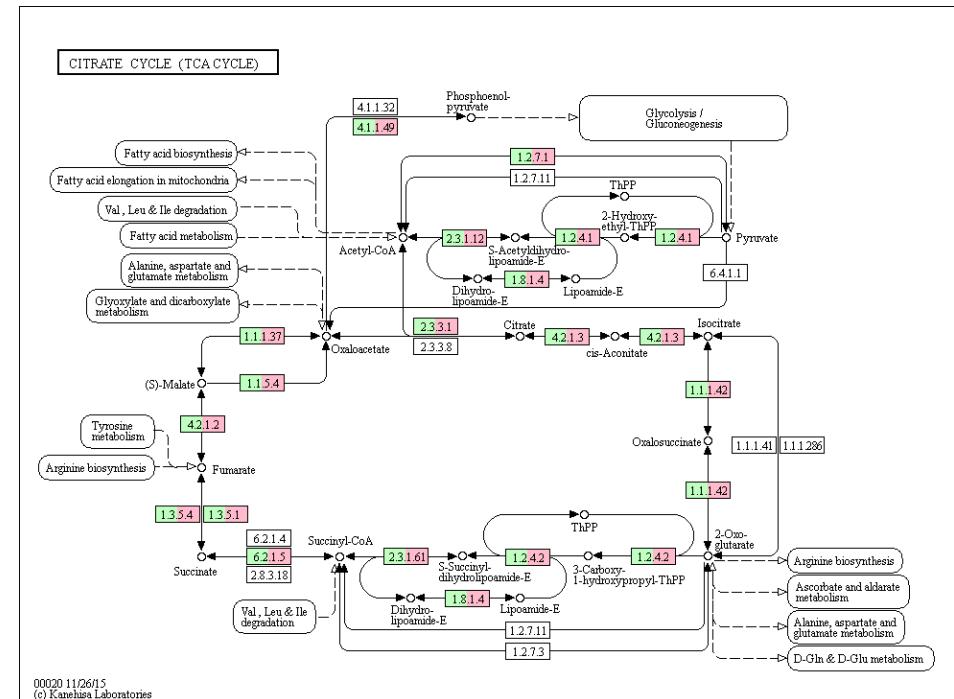
hsa eco Homo sapiens (human) vs. Escherichia coli K-12 strain
such as [Terpenoid backbone biosynthesis](#)
where green for human, red for E.coli, and split for both

種間比較

- 上のメニューの Pathway map をクリック
- 好きなパスウェイをクリック
 - 前者が持っている遺伝子が緑、後者が持っている遺伝子がピンクで表示
 - Overview パスウェイの場合、両者が持っている遺伝子はライトブルーで表示

Kegg Organism group: eco ecs

Category info	Pathway map	Brite hierarchy	Taxonomy				
Search genes:	<input type="text"/>	<input type="button" value="Go"/>	<input type="button" value="Clear"/>				
eco ecs							
T00007 eco Escherichia coli K-12 MG1655							
T00048 ecs Escherichia coli O157:H7 Sakai (EHEC)							
KEGG2	PATHWAY	BRITE	GENES	GENOME	LIGAND	DISEASE	DRUG



Bacterial secretion system マップ

- 03070 : Bacterial secretion system パスウェイでは病原性に関わる III型、VI型分泌装置が O157 側だけが持っているのがわかる

Environmental Information Processing

Membrane transport

02010 ABC transporters

02060 Phosphotransferase system (PTS)

03070 Bacterial secretion system

Signal transduction

02020 Two-component system

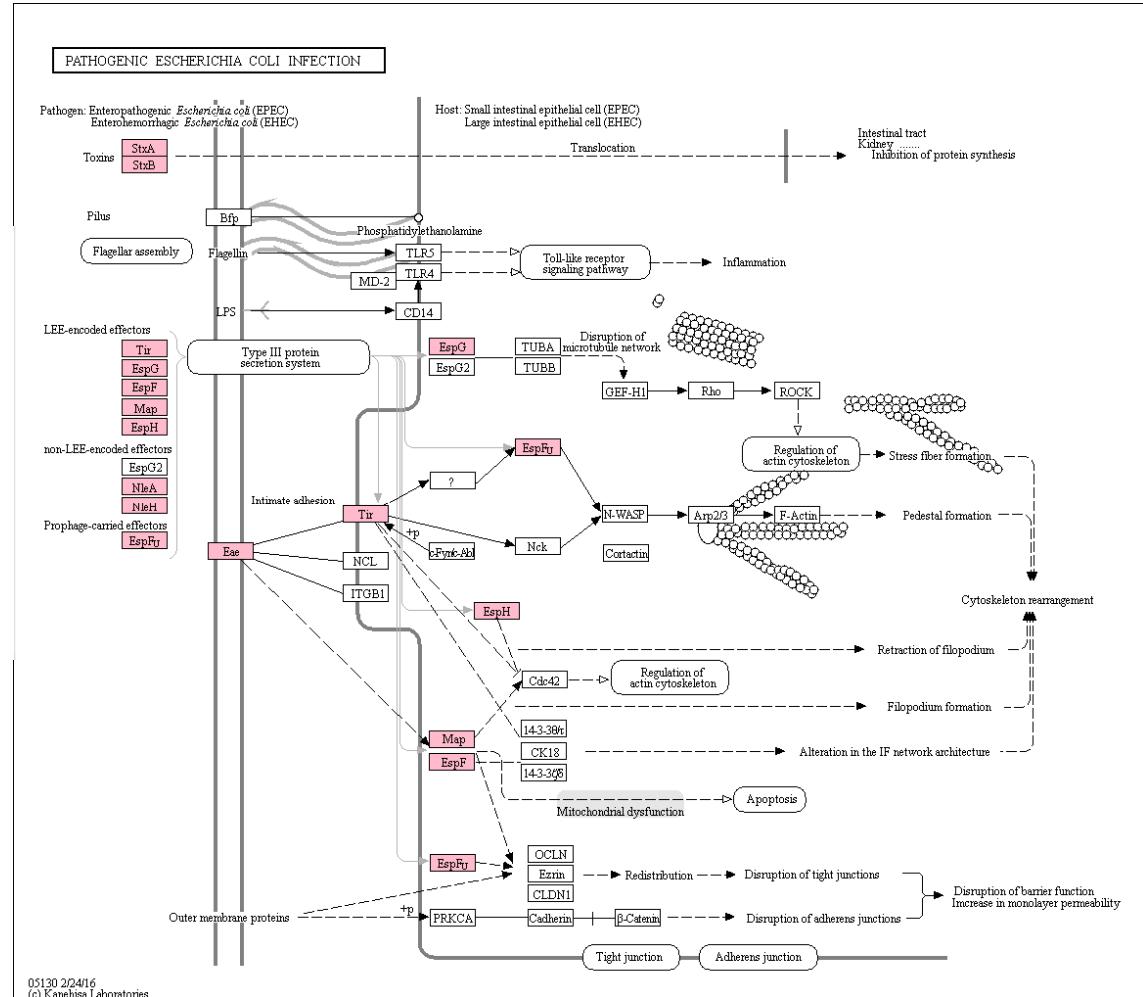
02011 ~~Two-component system~~

Type I	Type III	Type II
Outer membrane protein (OMP) TolC	Needle YscF	Secretin GspD
Membrane fusion protein HlyD	YscO YscP YscX	GspS
ABC transporter HlyB	Secretin YscC	GspC GspF GspG
	OMP YscW	GspH GspI GspJ
	Inner membrane protein (IMP) YscJ YscR YscS	GspK GspL GspM
	ATPase YscT YscU YscV	GspE
	ATPase-associated protein YscN	Leaderpeptidase GspO
	YscQ YscL	
Sec-SRP		
IMP SecD/F	SecE	SecG
SecY	YajC	YidC
ATPase SecA		
Secretion monitor SecM		
SRP receptor FtsY		
Targeting protein SecB ffh		

Pathogenic Escherichia coli infection

Infectious diseases: Bacterial

- 05110 Vibrio cholerae infection
- 05120 Epithelial cell signaling in Helicobacter pylori infection
- 05130 Pathogenic Escherichia coli infection
- 05132 Salmonella infection
- 05131 Shigellosis
- 05133 Pertussis
- 05134 Legionellosis
- 05150 Staphylococcus aureus infection
- 05152 Tuberculosis
- 05100 Bacterial invasion of epithelial cells



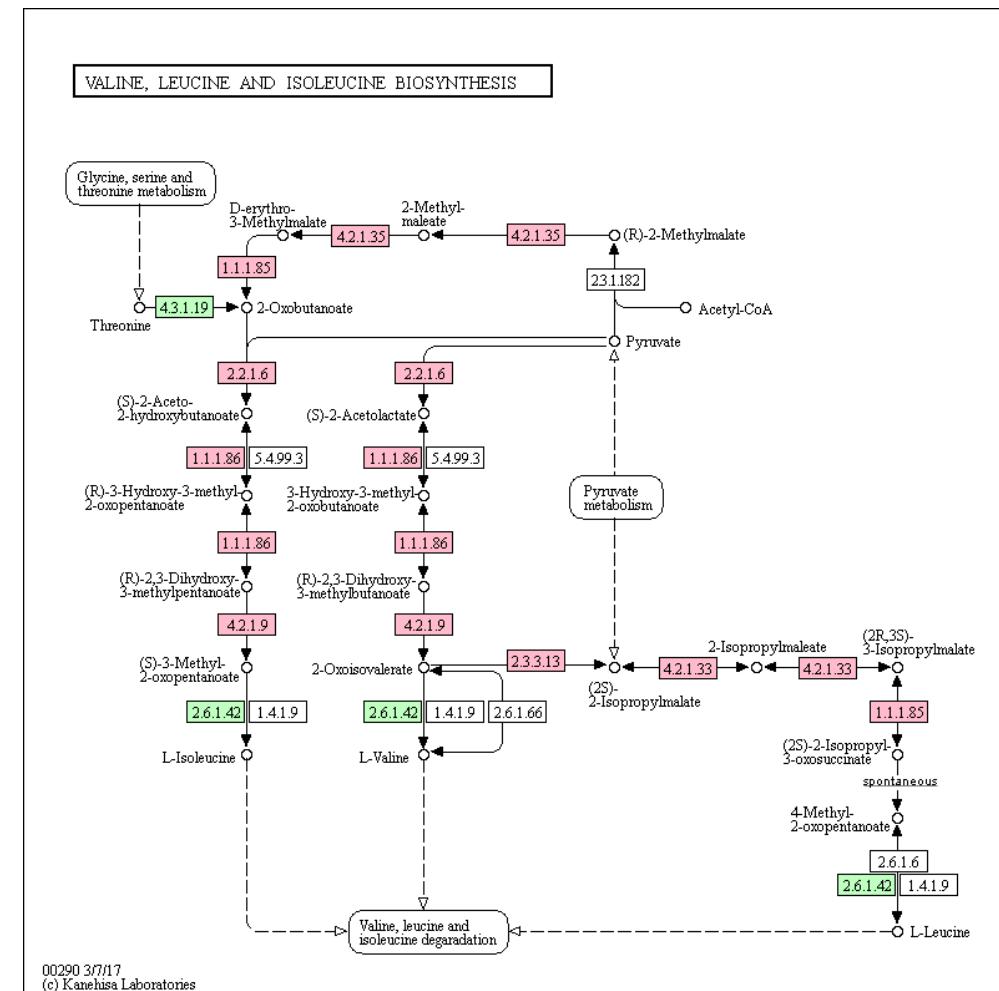
アブラムシとブフネラで種間比較

- 00290 : Valine, leucine and isoleucine biosynthesis マップ
 - アブラムシ : api
 - ブフネラ : buc
- 共生生物間のパスウェイ補完によってアミノ酸合成が可能になっていることがわかる

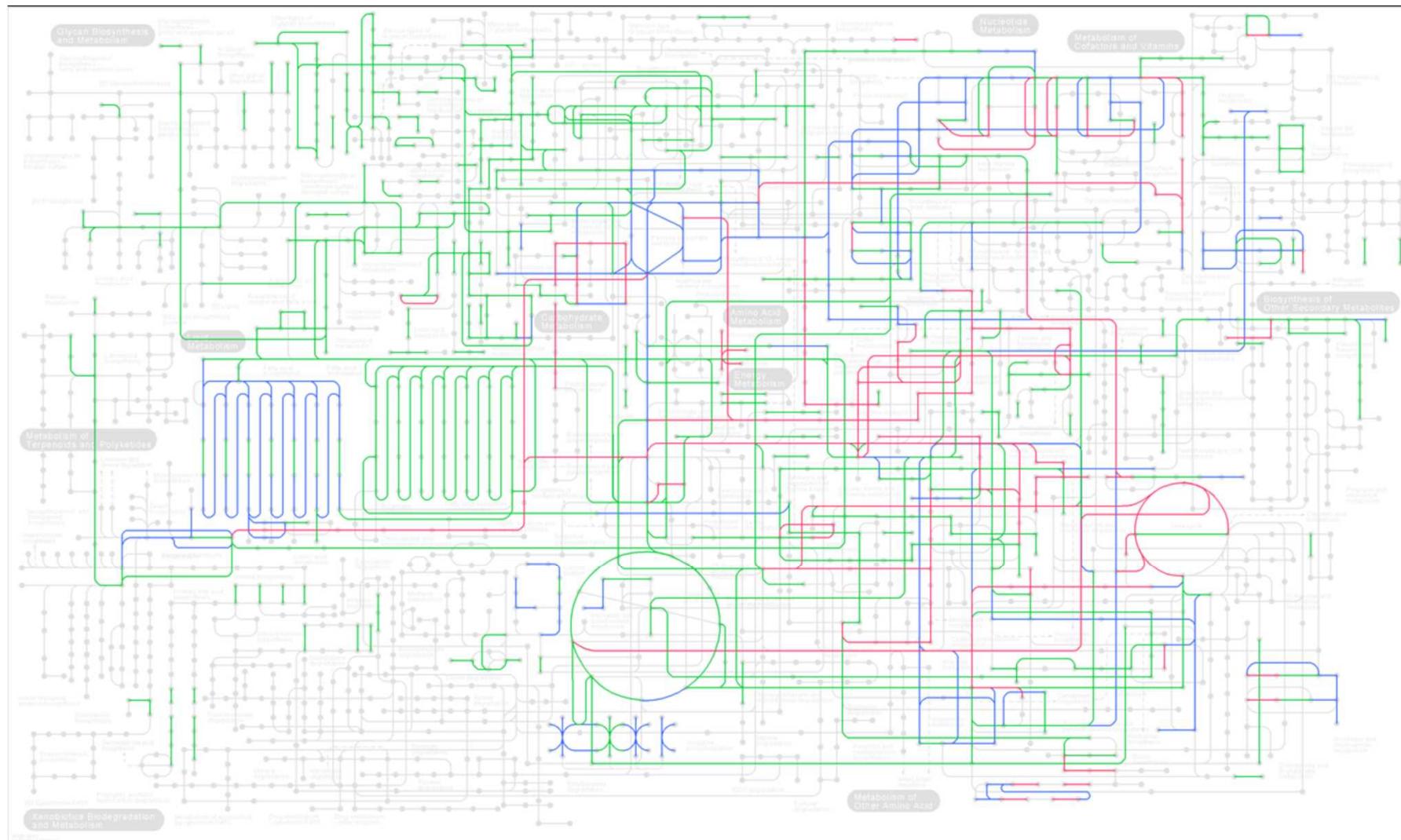
00240 Pyrimidine metabolism

Amino acid metabolism

- 00250 Alanine, aspartate and glutamate metabolism
- 00260 Glycine, serine and threonine metabolism
- 00270 Cysteine and methionine metabolism
- 00280 Valine, leucine and isoleucine degradation
- 00290 Valine, leucine and isoleucine biosynthesis**
- 00300 Lysine biosynthesis
- 00310 Lysine degradation
- 00220 Arginine biosynthesis
- 00330 Arginine and proline metabolism
- 00340 Histidine metabolism
- 00350 Tyrosine metabolism
- 00360 Phenylalanine metabolism
- 00380 Tryptophan metabolism



Overviewマップの場合



アブラムシ
ブフネラ
共通

サンプルデータのマッピング

- KEGG には遺伝子リストからパスウェイをマッピングするツールが組み込まれています。

- [KEGG Mapper](https://www.kegg.jp/kegg/mapper.html)
(<https://www.kegg.jp/kegg/mapper.html>)
- Pathway mapping tool の 2 番目の [Search&Color Pathway](#) をクリック
 - Search against : データベースコード
 - Primary ID : ID 種類 (KEGG ID, NCBI-GenID, NCBI-ProteinID, UniProt)
 - テキストエリア : 要素のリスト (遺伝子、タンパク質、化合物)
 - [配列 ID or 代謝産物 ID] 塗りつぶし色[,線の色]
 - 配列 ID は KEGG gene ID, NCBI-GenID, NCBI-ProteinID, UniProt ID
 - 代謝産物 ID は KEGG Compound ID (C番号) のみ
 - 線の色はオプション
 - 色は16進数表記か基本的なカラー名で記述



KEGG Mapper

A suite of KEGG mapping tools



About 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).

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 - 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)
- 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 color gradation or 3D bar graph

KEGG Mapper

- テキストエリア右の Example を選択して Exec ボタンをクリックすると、ヒットしたパスウェイのリストが表示されます（カッコの中はヒットした要素の数）

The screenshot shows the KEGG Mapper interface for searching and coloring pathways. On the left is a sidebar with links like About KEGG Mapper, Search Pathway, Search Brite, Search Module, Search Disease, Reconstruct Pathway, Convert ID, etc. The main area has a search bar with 'Search against: hsa' and an 'Examples:' dropdown containing 'Homo sapience pathway' (which is highlighted with a red box). Below the search bar is a text input field with '7167 red,blue' and 'C00118 pink' entered. There's also a section for file uploads and a checkbox for including aliases.

KEGG Mapper – Search&Color Pathway

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

Primary ID: KEGG identifiers (Outside IDs for organism-specific pathways only)

Enter objects one per line followed by bgcolor, fgcolor:

Examples: Homo sapience pathway

7167 red,blue
C00118 pink

Alternatively, enter the file name containing the data:

ファイルを選択 ファイル未選択

If necessary, change default bgcolor: pink

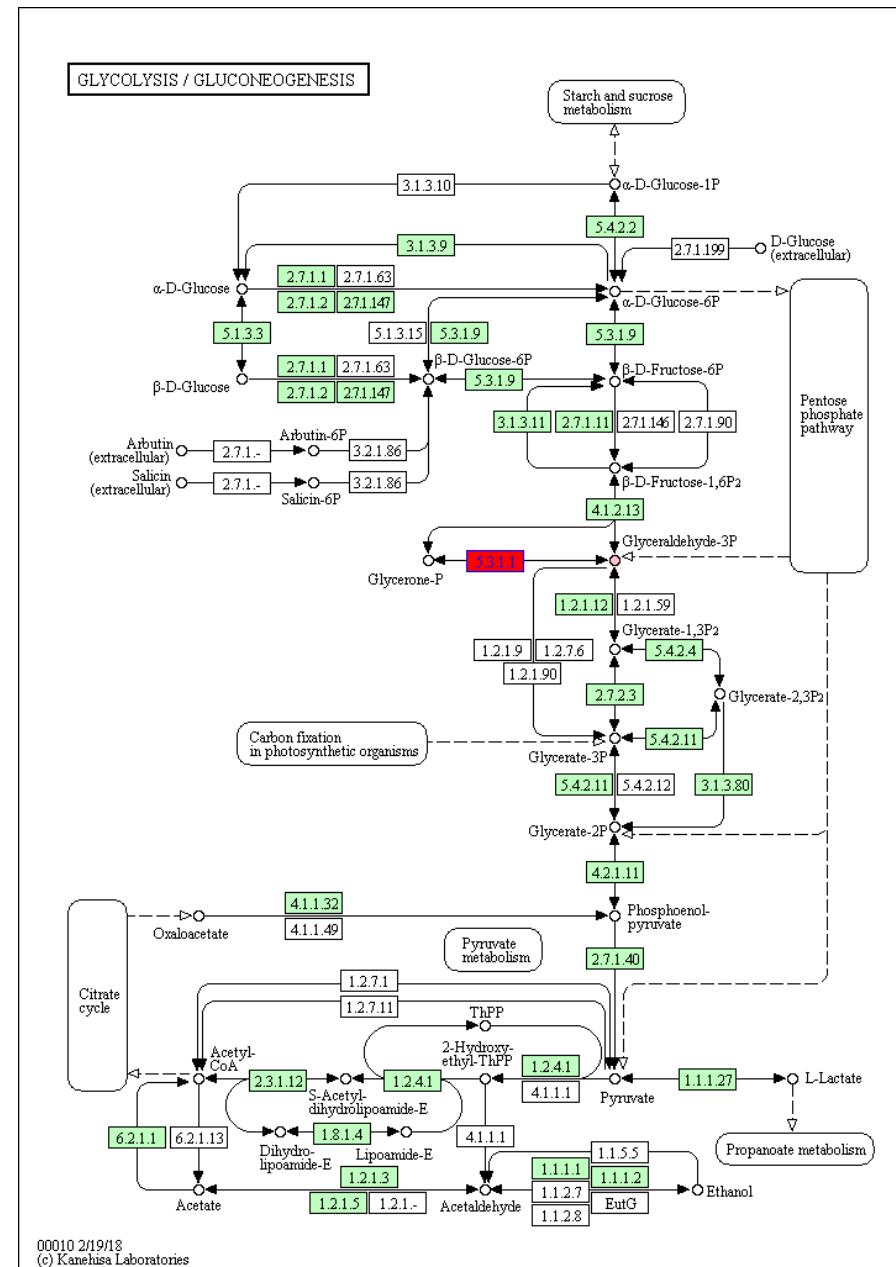
Include aliases

Pathway Search Result

[Sort by the pathway list](#)

[Show all objects](#)

- hsa01230 Biosynthesis of amino acids - Homo sapiens (human) (2)
- hsa01100 Metabolic pathways - Homo sapiens (human) (2)
- **hsa00010 Glycolysis / Gluconeogenesis - Homo sapiens (human) (2)**
- hsa00562 Inositol phosphate metabolism - Homo sapiens (human) (2)
- hsa00051 Fructose and mannose metabolism - Homo sapiens (human) (2)
- hsa01200 Carbon metabolism - Homo sapiens (human) (2)
- hsa00900 Terpenoid backbone biosynthesis - Homo sapiens (human) (1)
- hsa00750 Vitamin B6 metabolism - Homo sapiens (human) (1)
- hsa00052 Galactose metabolism - Homo sapiens (human) (1)
- hsa00030 Pentose phosphate pathway - Homo sapiens (human) (1)
- hsa00730 Thiamine metabolism - Homo sapiens (human) (1)

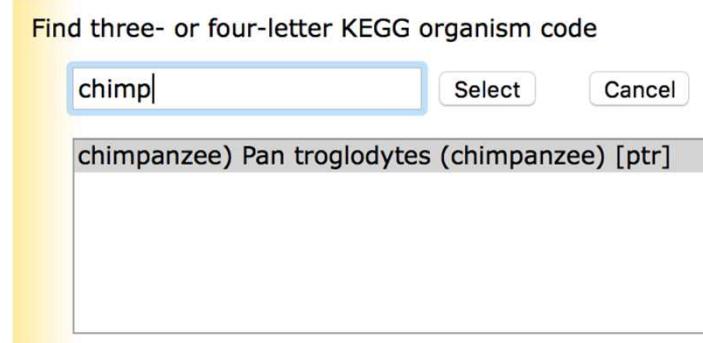


チンパンジーの遺伝子をマッピング

- テキストボックスに「例」のIDを記入
- Search against: にチンパンジーの生物種コードを入れる
 - コードがわからないので、org ボタンをクリック
 - ポップアップウィンドウでに 種名を入力すると、下のボックスに候補が出るので、選択したあと Select をクリック
 - チンパンジーのコード "ptr" が入力されていることを確認

例

453039 red
104003784 coral
453645 gray,red
453565 blue,yellow
450453 #fbfb88
463861 #88ffbb



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

Primary ID: KEGG identifiers (Outside IDs for organism-specific p)

Enter objects one per line followed by bgcolor, fgcolor:

453039 red 104003784 coral 453645 gray,red 453565 blue,yellow 450453 #fbfb88 463861 #88ffbb	Examples: Select
--	------------------

Alternatively, enter the file name containing the data:

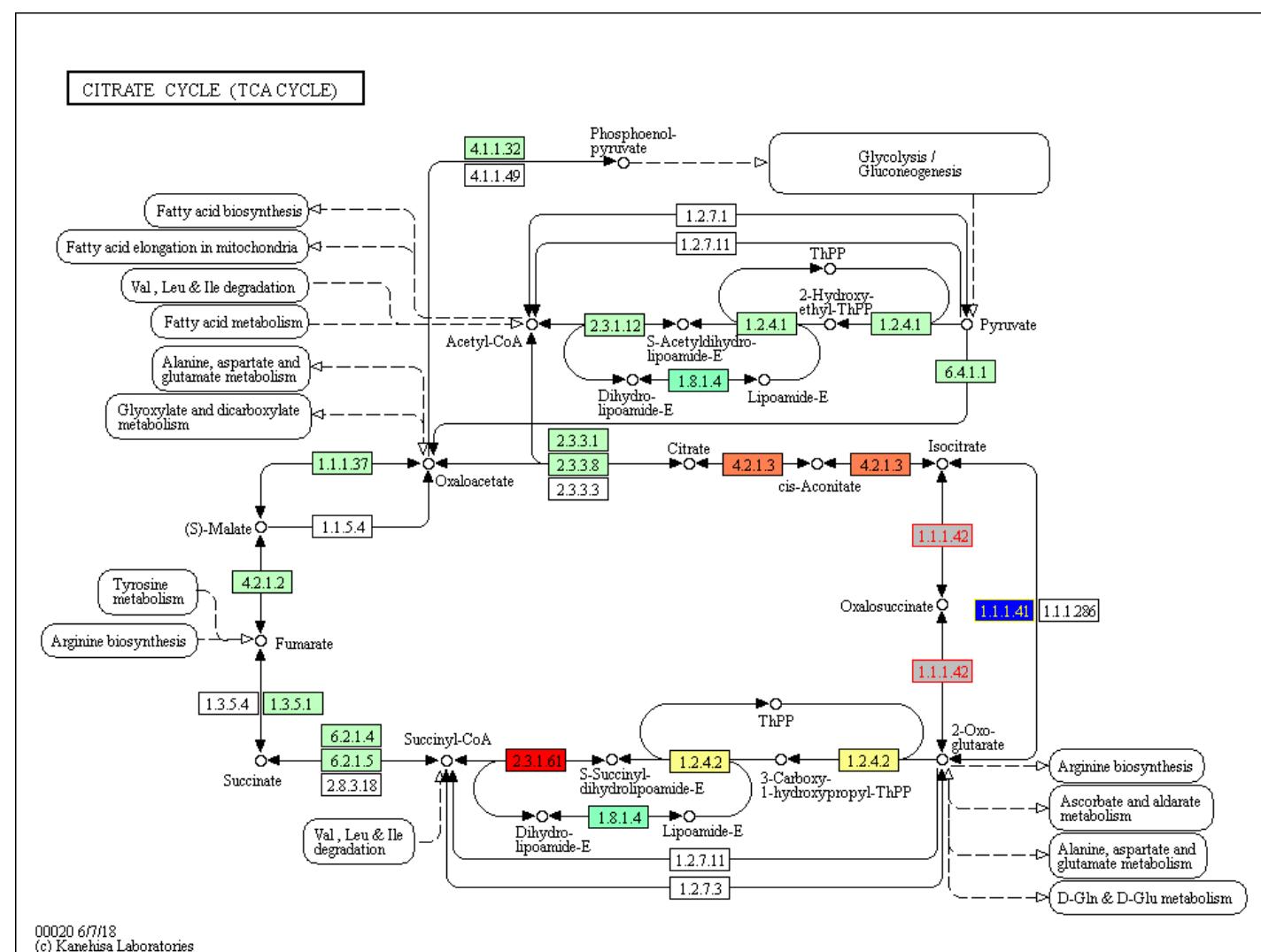
ファイルを選択 ファイル未選択

Pathway Search Result

Sort by the pathway list

Show all objects

- ptr00020 Citrate cycle (TCA cycle) - Pan troglodytes (chimpanzee) (6)**
- ptr01200 Carbon metabolism - Pan troglodytes (chimpanzee) (6)
- ptr01100 Metabolic pathways - Pan troglodytes (chimpanzee) (6)
- ptr01210 2-Oxocarboxylic acid metabolism - Pan troglodytes (chimpanzee) (3)
- ptr01230 Biosynthesis of amino acids - Pan troglodytes (chimpanzee) (3)
- ptr00630 Glyoxylate and dicarboxylate metabolism - Pan troglodytes (chimpanzee) (2)
- ptr00310 Lysine degradation - Pan troglodytes (chimpanzee) (2)
- ptr00380 Tryptophan metabolism - Pan troglodytes (chimpanzee) (1)
- ptr00280 Valine, leucine and isoleucine degradation - Pan troglodytes (chimpanzee) (1)
- ptr00640 Propanoate metabolism - Pan troglodytes (chimpanzee) (1)
- ptr00480 Glutathione metabolism - Pan troglodytes (chimpanzee) (1)
- ptr00260 Glycine, serine and threonine metabolism - Pan troglodytes (chimpanzee) (1)
- ptr00620 Pyruvate metabolism - Pan troglodytes (chimpanzee) (1)
- ptr00010 Glycolysis / Gluconeogenesis - Pan troglodytes (chimpanzee) (1)
- ptr04146 Peroxisome - Pan troglodytes (chimpanzee) (1)



数値データをマッピング

- [Color Pathway](https://www.kegg.jp/kegg/tool/map_pathway3.html) (https://www.kegg.jp/kegg/tool/map_pathway3.html) をクリック
- 右のサンプル Numerical values to colorを選択
 - 中身は配列 ID と数値の対応リスト
- Select KEGG pathway map: でパスウェイを指定 (hsa05200)
- Option: で Numerical values converted to color gradationを選択
- Exec ボタンをクリック

Select KEGG pathway map: hsa05200

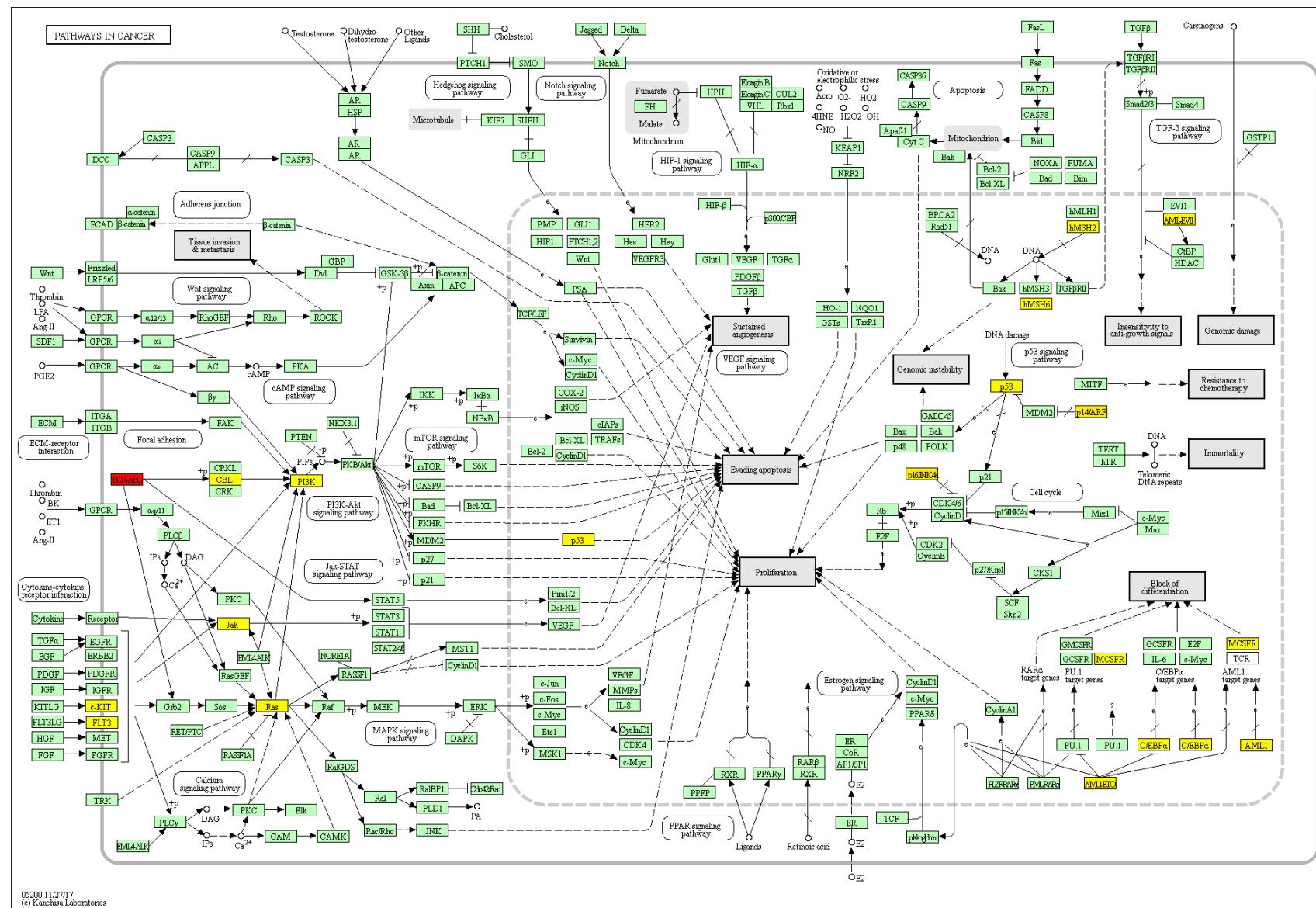
Enter data:

```
#hsa COSMIC
hsa:25    678
hsa:861   26
hsa:867   13
hsa:1029  44
hsa:1050  4
hsa:1436  3
hsa:2146  9
hsa:2322  7
hsa:2623  2
```

Examples:
Numerical values to color

Alternatively, enter the file name containing the data:

ファイルを選択 ファイル未選択



3Dグラフマッピング

- 右のサンプルから「Numerical values to color」
- 下のオプションから「Numerical values converted to 3D graph」を選択する

Select KEGG pathway map: hsa05200

Enter data:

```
#hsa COSMIC
hsa:25 678
hsa:861 26
hsa:867 13
hsa:1029 44
hsa:1050 4
hsa:1436 3
hsa:2146 9
hsa:2322 7
hsa:2623 2
```

Example: Numerical values to graph

Alternatively, enter the file name containing the data:

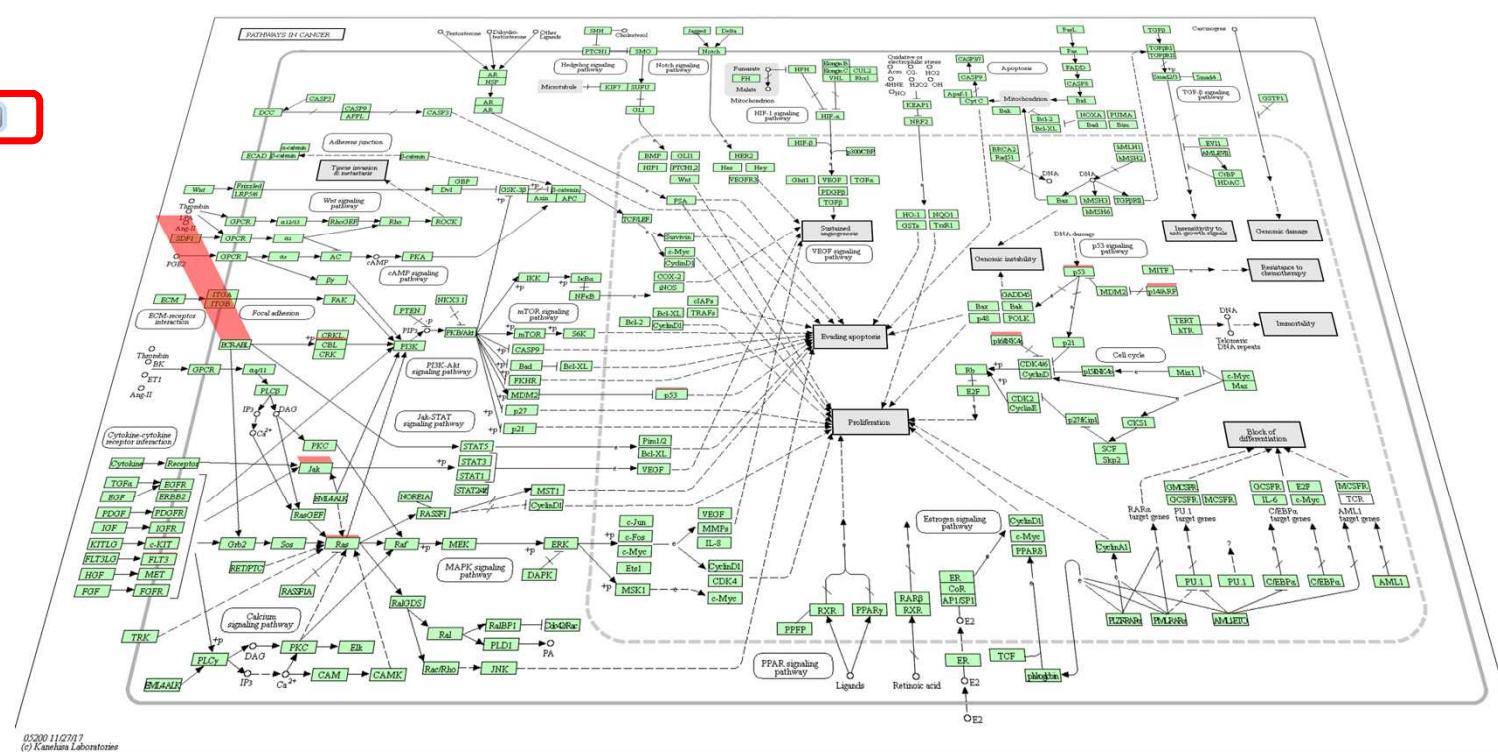
ファイルを選択 ファイル未選択

Option: Color specification
 Numerical values converted to color gradation
Range: minimum-maximum negative-zero-positive
minimum #ffff00 maixnum #ff0000 Check

Numerical values converted to 3D bar graph

Use uncolored diagram

Exec Clear



KEGGデータベースにはない遺伝子のマッピング

- KEGGに登録されている配列データと類似性を計算し、自動で遺伝子機能を推定、パスウェイへのマッピングを行う。
- [KAAS \(<https://www.genome.jp/tools/kaas/>\)](https://www.genome.jp/tools/kaas/)
- [BlastKOALA](#), [GhostKOALA](#)



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 or GHOST 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 BBH (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

Metagenomes

When the query consists of large numbers of sequences and / or sequences from mixture of species such as those from metagenome sequencing project, we recommend the GHOSTX search and SBH method.

- KAAS job request (SBH method for amino acid sequence query)

Example of Results

KO assignment

KAAS KO Assignment Results

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

Query gene : KO assignment

KAAS

70411

KAAS_00001
KAAS_00002 400003
KAAS_0003 400072
KAAS_0004 01733
KAAS_0005
KAAS_0006
KAAS_0007 403310
KAAS_0008 403311
KAAS_0009 403311
KAAS_0010 407024
KAAS_0011
KAAS_0012
KAAS_0013
KAAS_0014 404043
KAAS_0015 403086
KAAS_0016

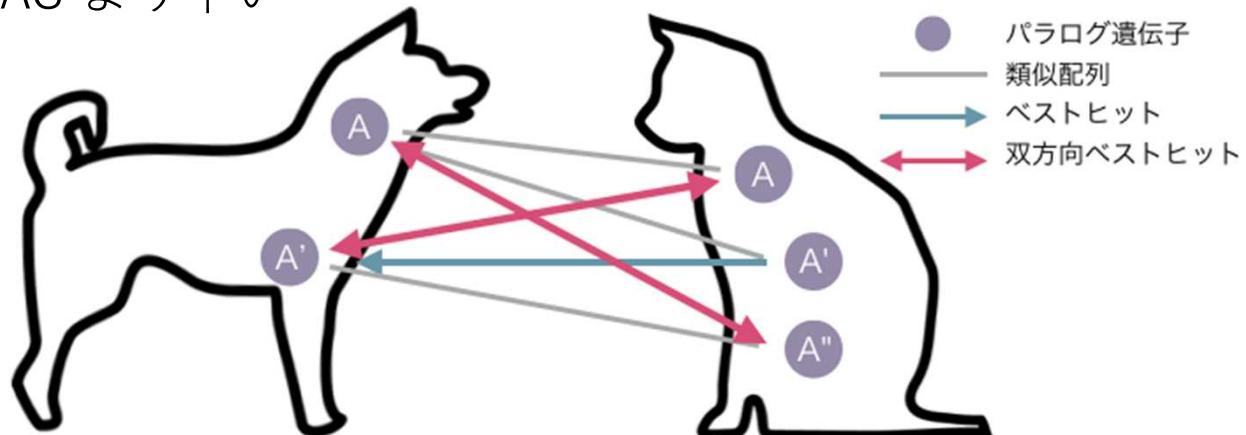
KEGG pathway mapping

GLYCOLYE

Starch and sucrose metabolism
Glycolysis/Gluconeogenesis
Pentose and glucuronate interconversions
Starch and sucrose metabolism
Glycolysis/Gluconeogenesis
Pentose and glucuronate interconversions
Fatty acid metabolism
Amino acid metabolism
Carbohydrate metabolism
Metabolism of other amino acids
Metabolism of aromatic compounds
Metabolism of terpenoids and prenyl compounds
Metabolism of nucleic acids
Metabolism of other organic compounds
Metabolism of inorganic substances
Metabolism of sulfur compounds
Metabolism of other inorganic substances
Metabolism of inorganic ions
Metabolism of inorganic elements

配列相同意検索

- [KAAS](https://www.genome.jp/tools/kaas/) (<https://www.genome.jp/tools/kaas/>)
 - 配列類似性の計算は [BLAST](#), [GhostX](#), [GhostZ](#) ベースの 3つ
 - GhostX は BLAST より精度は劣るが 100 倍早い
 - GhostZ は GhostX より精度は劣るが 2 倍早い
 - 種間で両方向ベストヒットを利用して遺伝子機能を推定（片方向も可能）
- [BlastKOALA](#), [GhostKOALA](#)
 - 配列類似性の計算は BLAST, GhostX ベース
 - クエリーからデータベースへの片方向の計算なのと、データベースを圧縮しているぶん KAAS より早い



BlastKOALAを使ったパスウェイマッピング

- [Annotate Sequence by BlastKOALA](https://www.kegg.jp/kegg/tool/annotate_sequence.html)
(https://www.kegg.jp/kegg/tool/annotate_sequence.html)
- Example: の sequence.txt をコピー & ペースト、もしくはダウンロードしてファイルを選択
- Family/Genus ボタンをクリック
 - サンプルが Buchnera の仲間なので、KEGG の Buchnera データを使う
 - 新たに開いたウィンドウで、Buchnera を探し、Taxonomy番号をクリック

KEGG Mapper – Annotate Sequence by BlastKOALA

About KEGG Mapper Upload query amino acid sequences in FASTA format

Search Pathway Enter FASTA sequences

Search&Color Pathway

Color Pathway

Search Brite

Search&Color Brite

Join Brite

Join Brite Table

Search Module

Search&Color Module

Search Disease

Reconstruct Pathway

Reconstruct Brite

Reconstruct Module

Map Taxonomy

YLINFFYKTLKIKGTPIQIFQFDNENPYVKNNK
>IcINZ_ACFK0100001.1_prot_WP_009874555.1_575 [gene=C55_RS0103170]
[protein=/inference=EXISTENCE: similar to AA sequence:SwissProt:P57663.1]
[protein_id=WP_009874555.1] [location=complement(641210..641791)]
MLNISKKNIFFLIISLILFNWKYFSLVNKENLESKYEKIKKINKKKSKNLYEVENFIVQ
NTSI
YGTLTALSKKYVECNNLDKALLQLNNSLKYTKNEENLKNLLKINIAKIQIQKNENNKKAMNIL
ETIQNHN
WKNIIEHMKGDIFININNKKEAIKSWKSLFIEDSNASKEIIINMKLNELKEQN

Or upload file: ファイル未選択

Select GENES family/genus dataset to be searched

Select Example:
Query data: sequence.txt
Select Buchnera (32199)

Exec Clear

Erwinia	551	Erwinia amylovora
		Erwinia billingiae
		Erwinia sp. Ejp617
		Erwinia gerundensis
Buchnera	32199	Buchnera aphidicola
Wigglesworthia	51228	Wigglesworthia glossinidia
		Pantoea ananatis
		Pantoea vagans
		Pantoea sp. At-Qh

計算中

- Exec ボタンをクリック
- 数分待つ



BlastKOALA
Job Request

Request accepted

Your job ID is 2fed9b298f4efd7e9a9eebed200f16a1fa687edb

Started at Mon Jan 8 15:09:48 JST 2018

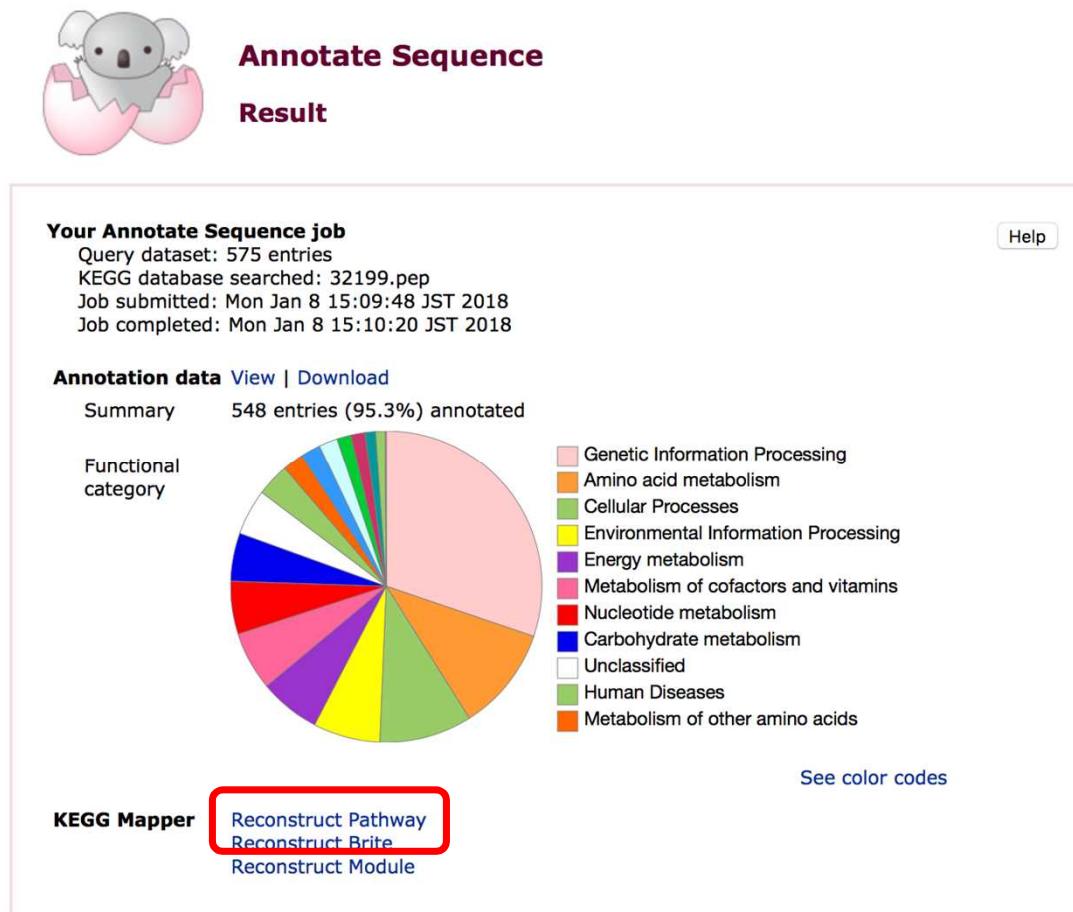
This page reloads automatically every 5 seconds

Your result will be displayed in the following page

[http://www.kegg.jp/kegg-bin/blastkoala_result?
id=2fed9b298f4efd7e9a9eebed200f16a1fa687edb&passwd=mTYL46&mode=mapper](http://www.kegg.jp/kegg-bin/blastkoala_result?id=2fed9b298f4efd7e9a9eebed200f16a1fa687edb&passwd=mTYL46&mode=mapper)

計算結果

- Reconstruct Pathway から遺伝子がマッピングされたパスウェイを見ることができる



Overviewマップの再構築結果

Pathway Reconstruction Result

Show all objects

Metabolism

Glob and overview maps

- [0_100 Metabolic pathways \(205\)](#)
- [0_110 Biosynthesis of secondary metabolites \(97\)](#)
- [01120 Microbial metabolism in diverse environments \(52\)](#)
- [01130 Biosynthesis of antibiotics \(76\)](#)
- [01200 Carbon metabolism \(30\)](#)
- [01210 2-Oxocarboxylic acid metabolism \(10\)](#)
- [01212 Fatty acid metabolism \(5\)](#)
- [01230 Biosynthesis of amino acids \(62\)](#)

Carbohydrate metabolism

- [00010 Glycolysis / Gluconeogenesis \(15\)](#)
- [00020 Citrate cycle \(TCA cycle\) \(5\)](#)
- [00030 Pentose phosphate pathway \(12\)](#)
- [00040 Pentose and glucuronate interconversions \(1\)](#)
- [00051 Fructose and mannose metabolism \(7\)](#)
- [00052 Galactose metabolism \(1\)](#)
- [00500 Starch and sucrose metabolism \(2\)](#)
- [00520 Amino sugar and nucleotide sugar metabolism \(9\)](#)
- [00620 Pyruvate metabolism \(6\)](#)
- [00630 Glyoxylate and dicarboxylate metabolism \(2\)](#)
- [00640 Propanoate metabolism \(3\)](#)
- [00650 Butanoate metabolism \(2\)](#)
- [00660 C5-Branched dibasic acid metabolism \(2\)](#)
- [00562 Inositol phosphate metabolism \(2\)](#)

Energy metabolism

- [00190 Oxidative phosphorylation \(27\)](#)
- [00195 Photosynthesis \(8\)](#)
- [00710 Carbon fixation in photosynthetic organisms \(7\)](#)
- [00720 Carbon fixation pathways in prokaryotes \(4\)](#)
- [00680 Methane metabolism \(8\)](#)
- [00920 Sulfur metabolism \(9\)](#)

Lipid metabolism

- [00061 Fatty acid biosynthesis \(5\)](#)
- [00564 Glycerophospholipid metabolism \(1\)](#)
- [01040 Biosynthesis of unsaturated fatty acids \(2\)](#)

Nucleotide metabolism

