

# パスウェイデータベースを使って ゲノムを理解する

奥田修二郎

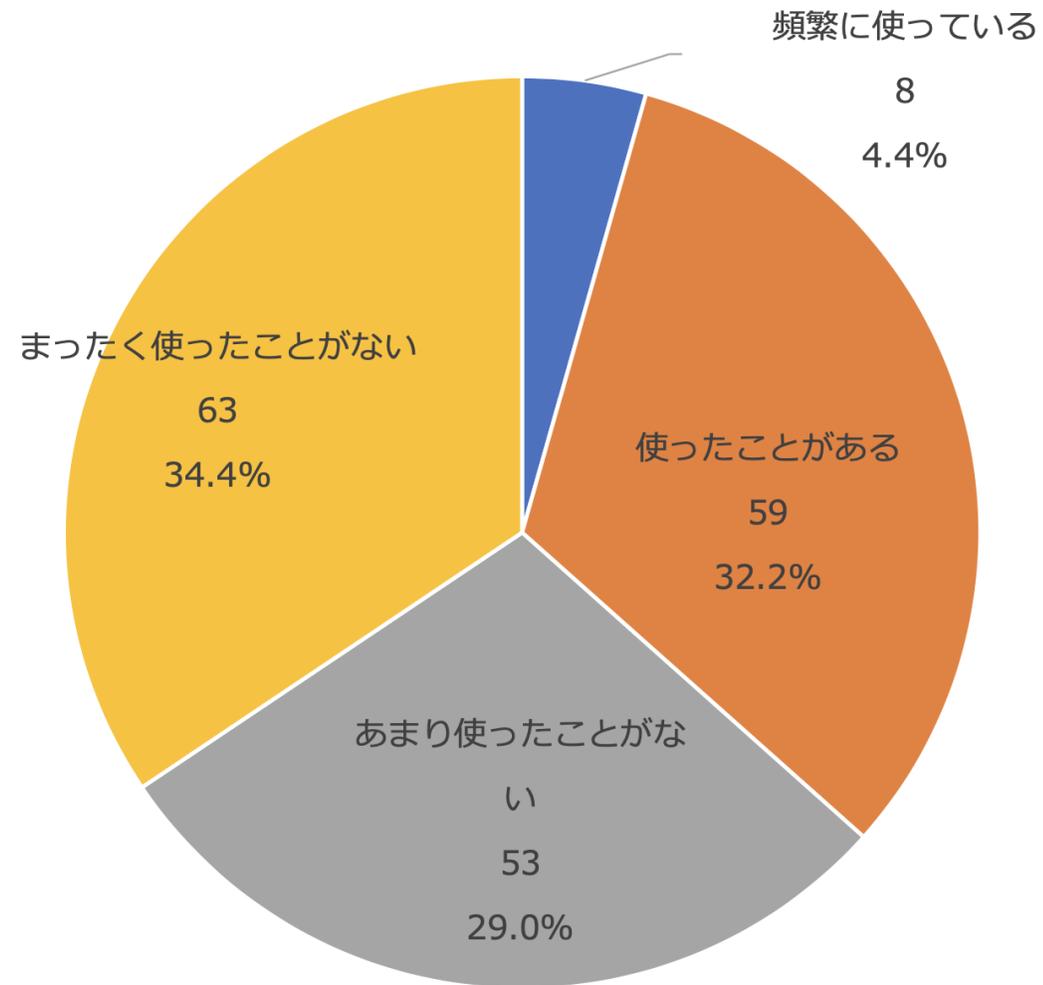
新潟大学医学部メディカルAIセンター

オンライン

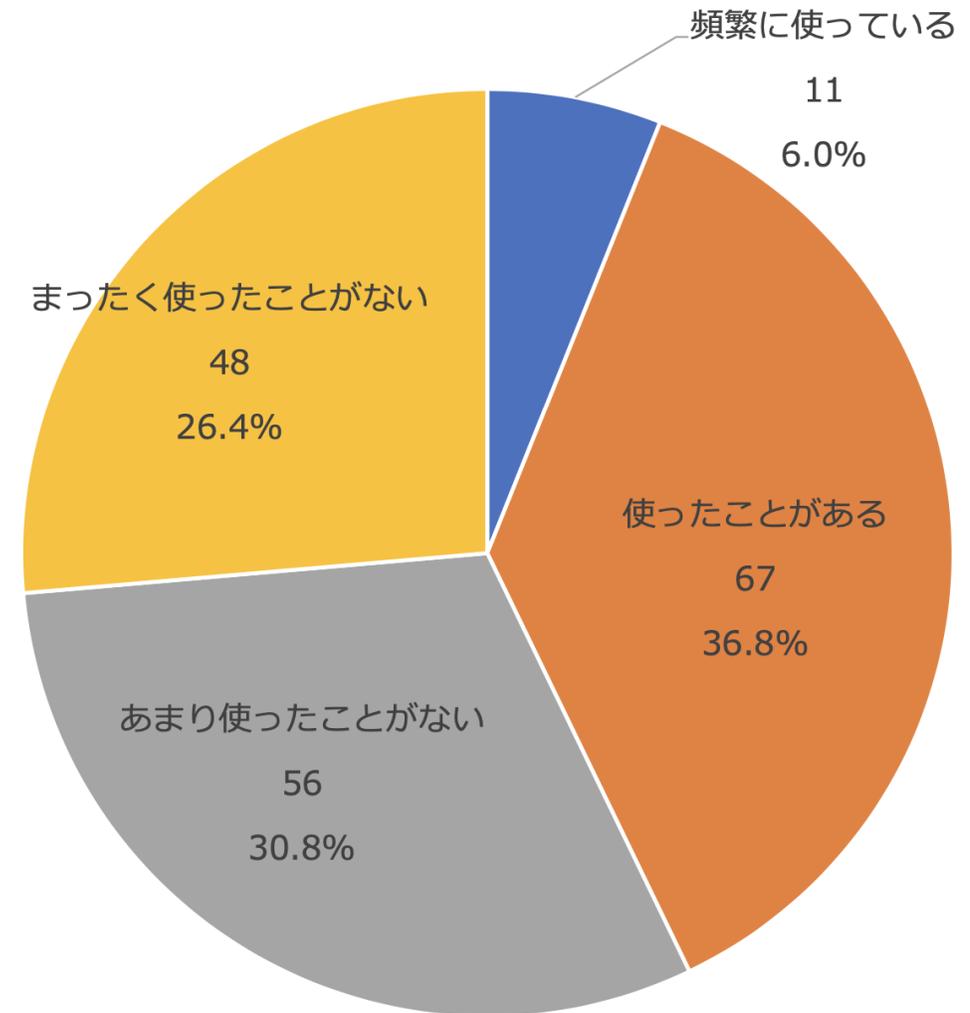
2022年8月25日

# 事前アンケート

【講義1-1】 パスウェイデータベースを使ったことがありますか？



【講義1-2】 KEGGデータベース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

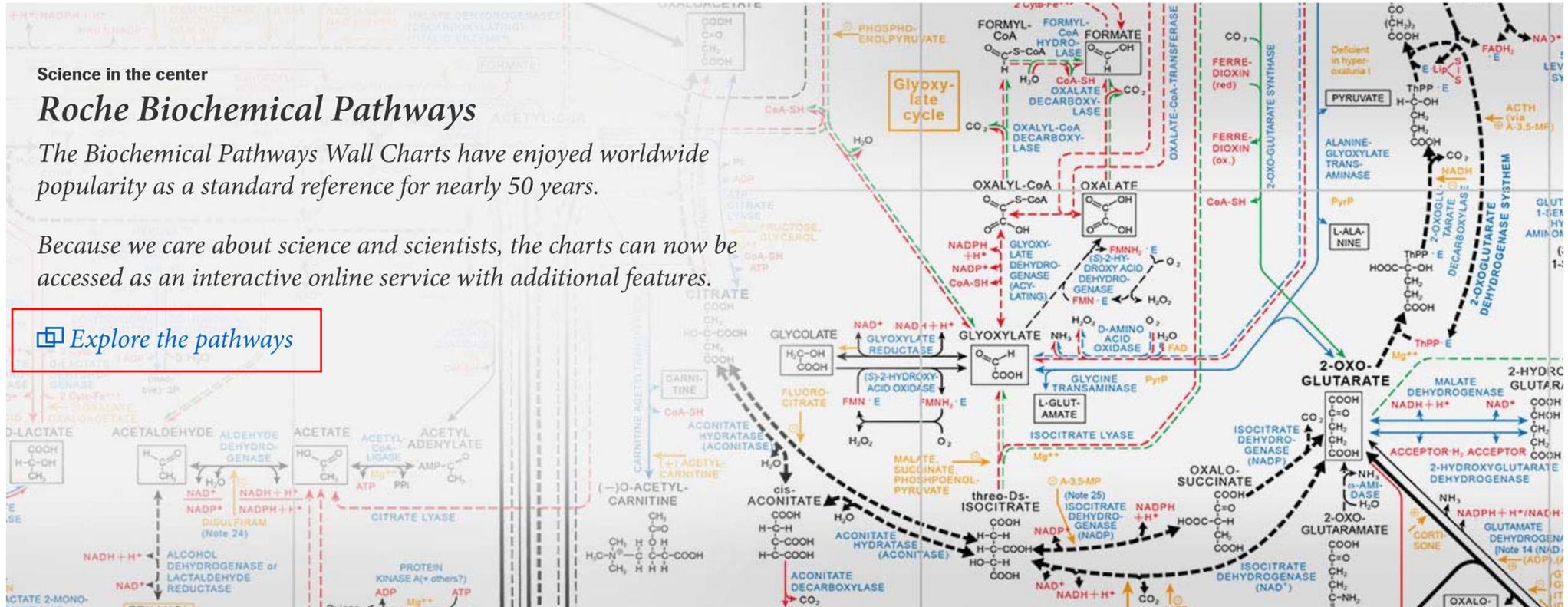
Science in the center

## 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

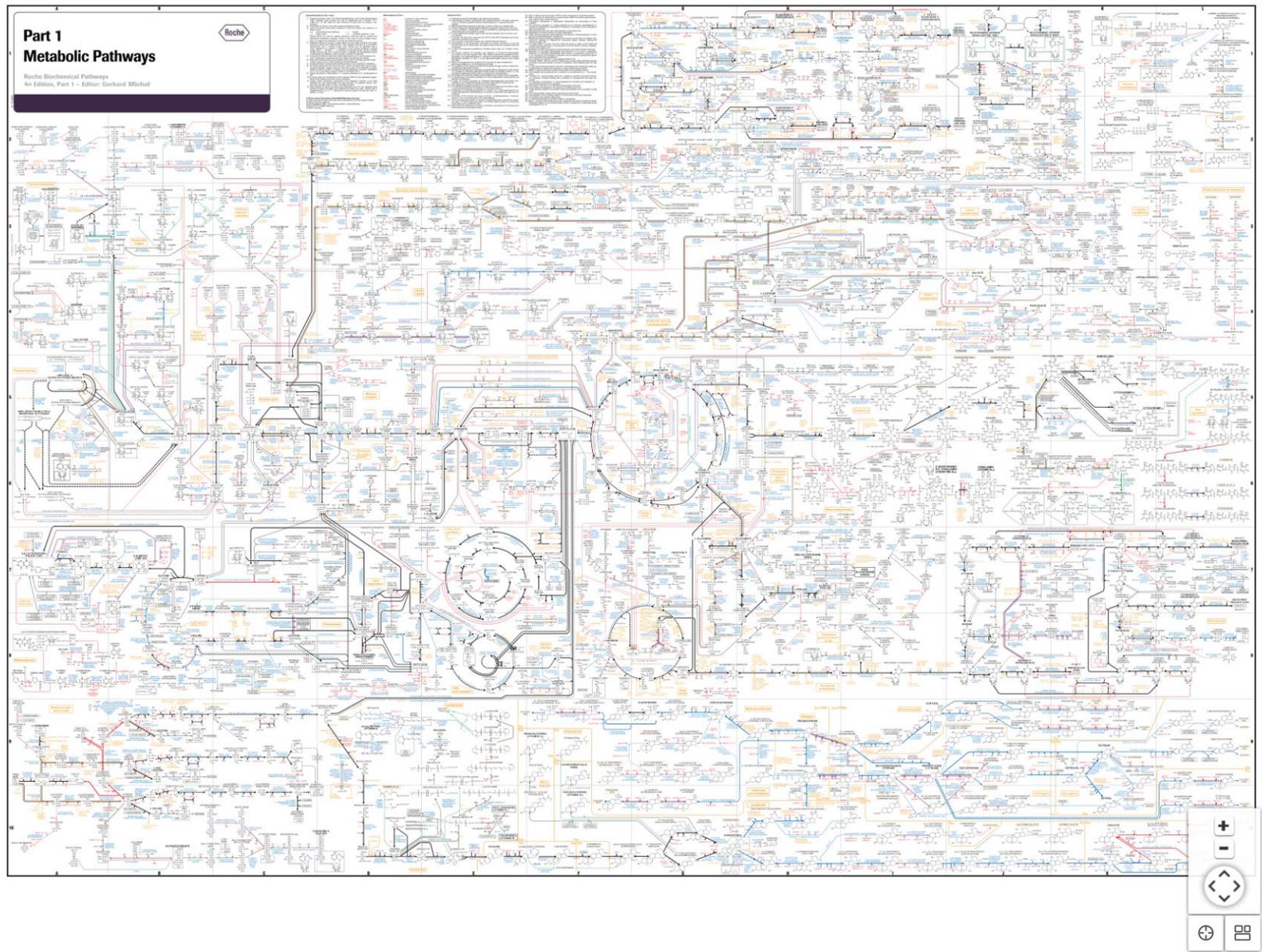
**Dr Gerhard Michal**  
Editor of the Roche Biochemical Pathways

### By the numbers

**49 years**  
continuously fine-tuned by the editor himself.

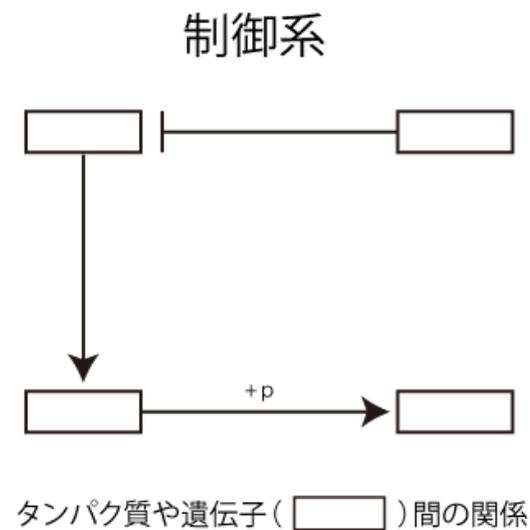
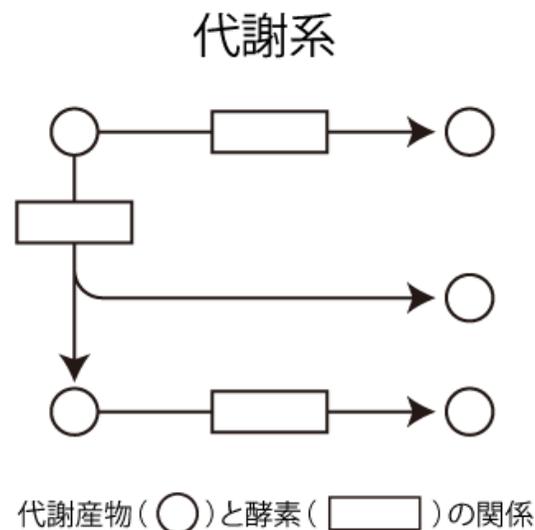
検索：Roche pathway

[Click here to open the menu](#)



# 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くなっています。
- 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くなっています。
  - [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 主にKEGG PATHWAY のブラウザ上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 18,000 種以上
  - 専門家が手作業で作成した文献ベースのデータ + 自動解析ツール
  - 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 displays the BioCyc website interface. At the top, there is a navigation bar with 'BioCyc Database Collection' and a search box. Below the navigation bar, the main content area is divided into several sections:

- BioCyc Database Collection:** A section with a heading and a paragraph explaining that subscriptions are now required to access BioCyc. It provides a link for more information on obtaining a subscription and notes that users do not need to purchase a subscription if their institution's name is displayed in the top right corner.
- Getting Started:** A section titled 'New to BioCyc? Typical usage is:' followed by a list of instructions: 'Select one or more databases (genomes) to search. To do so, click "change organism database" in the box in the top right of every page. By default, BioCyc searches *Escherichia coli* K-12 substr. MG1655.', 'Search for a gene or pathway using the Quick Search, or see the Search menu for more options.', and a 'New User Guide >>' button.
- Tools:** A section titled 'BioCyc provides tools for navigating, visualizing, and analyzing the underlying databases, and for analyzing omics data:' followed by a list of tools: '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', and 'Comparative analysis tools'.
- RouteSearch:** A section with a heading and a diagram illustrating the search for paths through the metabolic network. The diagram shows a 'Long Route' (orange), a 'Loss of Atoms' route (green dashed line), and a 'Best Route: Atoms Conserved' (blue). Below the diagram is a 'RouteSearch: Search for Paths through the Metabolic Network' section with a 'Learn More' button and a pagination bar (1-10).
- BioCyc Databases:** A section titled 'The BioCyc databases are divided into three tiers, based on their quality.' It describes 'Tier 1 databases' (received at least one person-year of literature-based curation) and 'Tier 2 and Tier 3 databases' (contain computationally predicted metabolic pathways).
- Funding Sources:** A section stating 'The development of BioCyc is funded by NIH grant GM080746 from the NIH National Institute of General Medical Sciences.'

## Metabolism -> Cellular Overview



[LOGIN](#) | [Why Login?](#) | [Create New Account](#)

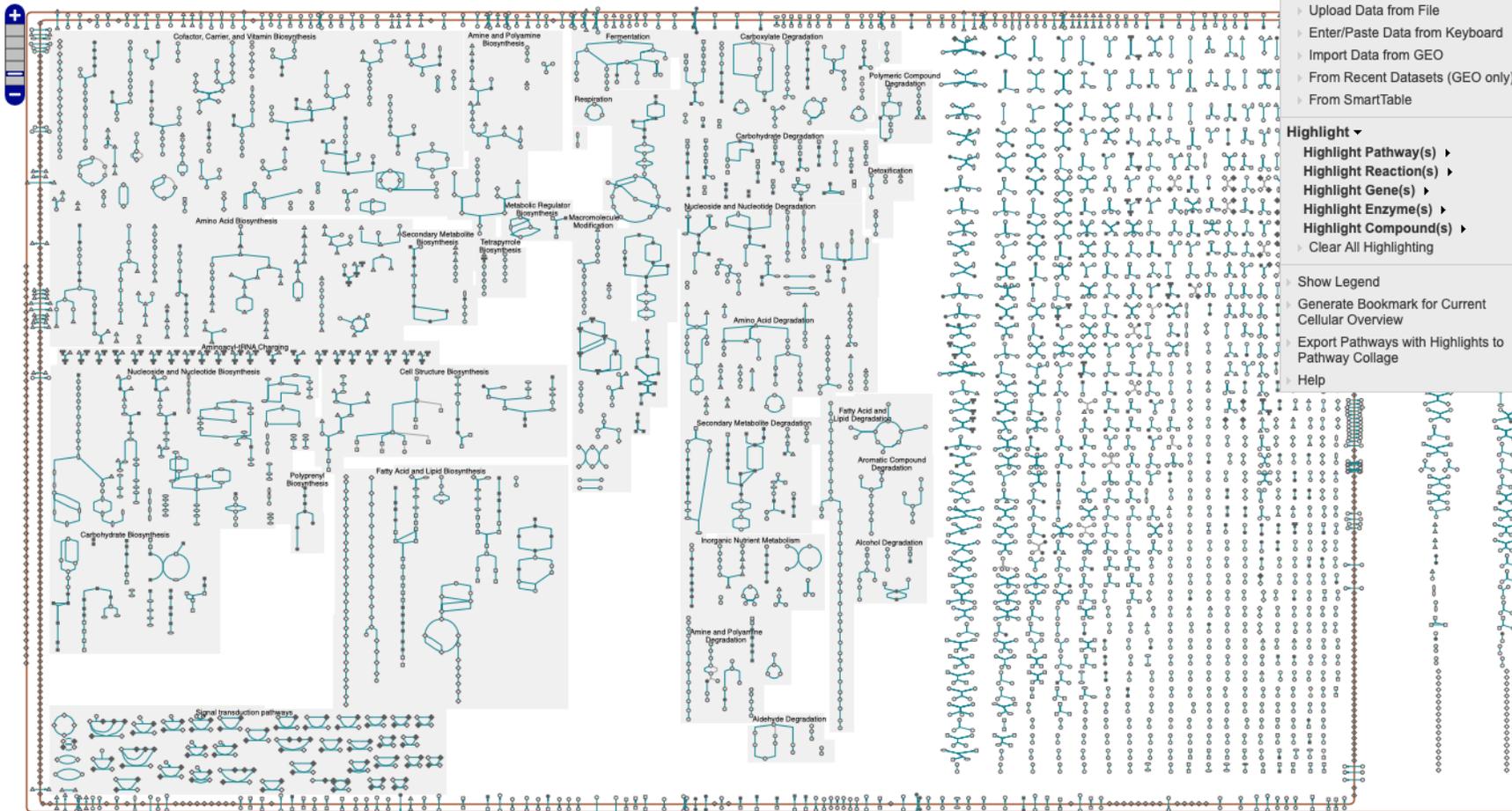
Enter a gene, protein, metabolite or pathway... [Quick Search](#) [Gene Search](#)  
 Searching *Escherichia coli K-12 substr. MG1655 reference genome (EcoCyc)* [change organism database](#)

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

Pan left/right/up/down the entire diagram by holding the left mouse button, click on an object for more info, right-click (ctrl-click for Mac) for menu

*Escherichia coli K-12 substr. MG1655 reference genome (EcoCyc)*

|                      |                      |                            |  |
|----------------------|----------------------|----------------------------|--|
| Opacity              | Edge Thickness       | Highlighted Edge Thickness | Base Layer   |
| <input type="text"/> | <input type="text"/> | <input type="text"/>       | <input checked="" type="radio"/> Cellular Overview |



- OPERATIONS**
- Overlay Experimental Data (Omics Viewer) ▾
    - Upload Data from File
    - Enter/Paste Data from Keyboard
    - Import Data from GEO
    - From Recent Datasets (GEO only)
    - From SmartTable
  - Highlight ▾
    - Highlight Pathway(s) ▸
    - Highlight Reaction(s) ▸
    - Highlight Gene(s) ▸
    - Highlight Enzyme(s) ▸
    - Highlight Compound(s) ▸
    - Clear All Highlighting
  - Show Legend
  - Generate Bookmark for Current Cellular Overview
  - Export Pathways with Highlights to Pathway Collage
  - Help

# Reactome

- ウェブサイト：<https://www.reactome.org/>
- 開発：EMBL-EBI 他
- 対象：ヒトを中心に脊椎動物、酵母、植物、16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

The screenshot shows the Reactome homepage with the following elements:

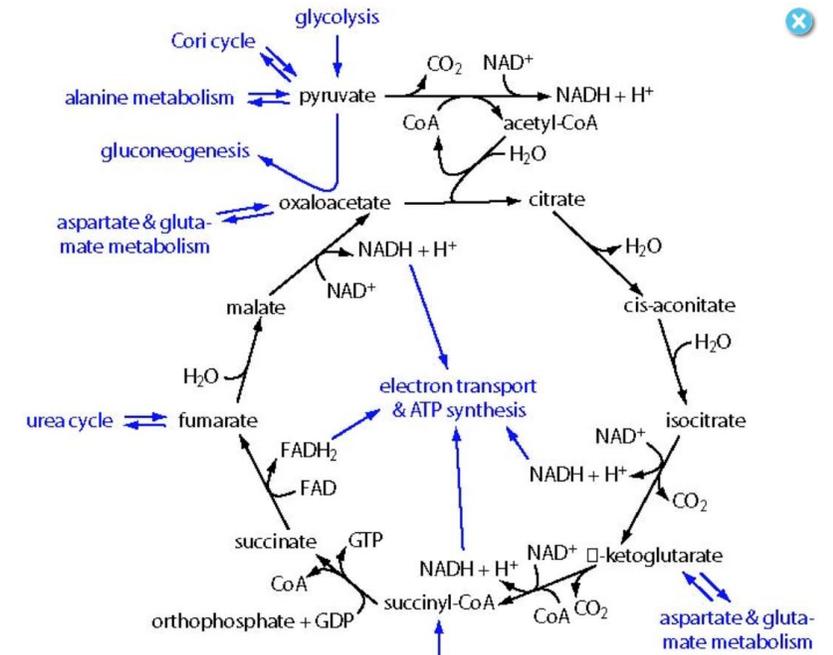
- Logo: reactome
- Navigation: About, Content, Docs, Tools, Community, Download
- Search bar: Find Reactions, Proteins and Pathways. Example input: "e.g. O95631, NTN1, signaling by EGFR, glucose".
- Four main service tiles:
  - Pathway Browser** (highlighted with a red box): Visualize and interact with Reactome biological pathways.
  - Analyze Data**: Merges pathway identifier mapping, over-representation, and expression analysis.
  - ReactomeFIViz**: Designed to find pathways and network patterns related to cancer and other types of diseases.
  - Documentation**: Information to browse the database and use its principal tools for data analysis.

The screenshot shows a project page with the following content:

- Header: USE REACTOME GRAPH DATABASE IN YOUR PROJECT (with a LEARN MORE button).
- Section: Why Reactome. Text: "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, modelling, systems biology and education. The current version (v63) of Reactome was released on December 18, 2017." Logos for EMBL-EBI, NYU Langone Health, and OICR are shown.
- Section: Latest News. Text: "Version 63 Released", "Updated License Agreement", "New Reactome Publication published in 2018 NAR Database Issue", "Proudly introducing our new logo", "New responsive website with a fresh look".
- Section: Tweets. A tweet from @reactome: "Excited about our new release (V63): 2179 human pwys & 11426 rxns w/ 10996 proteins, 1763 small mols & 27694".

## Pathway Browser

The screenshot displays the Reactome Pathway Browser interface. At the top, it shows the Reactome logo, version 3.7, and the selected organism: **Homo sapiens**. A search bar is available with the placeholder text "Search for a term, e.g. pten ...". Below the search bar is a navigation menu with icons for Citation, Analysis, Tour, and Layout. On the left side, there is an "Event Hierarchy" panel listing various biological processes such as Autophagy, Cell Cycle, and Metabolism. The main area shows a complex, interconnected network of pathways represented by blue nodes and lines. At the bottom, there is a navigation bar with tabs for Description, Molecules, Structures, Expression, Analysis, and Downloads. A description box at the bottom left explains that the interface displays details when an item is selected, including input and output molecules, summary, and references.



# 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）約8,000種（>780真核生物、>7,100真正細菌、>390古細菌）、ウイルス 359ゲノム、環境メタゲノム 477サンプル、生体メタゲノム 712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。



KEGG  Search Help [» Japanese](#)

## 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 DB links](#)

## KEGG Software

[KegTools](#)  
[KEGG API](#)  
[KGML](#)

## KEGG FTP

[Subscription](#)

[GenomeNet](#)

## 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

### Main entry point to the KEGG web service

**KEGG2** [KEGG Table of Contents](#) [[Update notes](#) | [Release history](#)]

### Data-oriented entry points

**KEGG PATHWAY** KEGG pathway maps  
**KEGG BRITE** BRITE hierarchies and tables  
**KEGG MODULE** KEGG modules  
**KEGG ORTHOLOGY** KO functional orthologs [[Annotation](#)]  
**KEGG GENOME** Genomes [[Pathogen](#) | [Virus](#) | [Plant](#)]  
**KEGG GENES** Genes and proteins [[SeqData](#)]  
**KEGG COMPOUND** Small molecules  
**KEGG GLYCAN** Glycans

### Classification

[Pathway](#)  
[Brite](#)  
[Brite table](#)  
[Module](#)  
[KO \(Function\)](#)  
[Organism](#)  
[Compound](#)  
[Network](#)  
[Disease \(ICD\)](#)

Menu [PATHWAY](#) [BRITE](#) [MODULE](#) [KO](#) [GENES](#) [LIGAND](#) [NETWORK](#) [DISEASE](#) [DRUG](#) [DBGET](#)

Search  KEGG for

### Data-oriented entry points

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

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

# 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

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

## Organism-specific entry points

**KEGG organisms** the list of currently available organisms

Select     (examples) [hsa](#) [mmu](#) [sce](#) [eco](#) [bsu](#) [syn](#)

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

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 sabaues (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

 **Homo sapiens (human)**

Genome info   Pathway map   Brite hierarchy   Module   Genome map   Blast   Taxonomy

Search genes:

---

**Genome information**

**T number** [T01001](#)

**Org code** hsa

**Aliases** HUMAN, 9606

**Full name** Homo sapiens (human)

**Definition** Homo sapiens (human)

**Category** Reference genome

**Annotation** yes

**Taxonomy** TAX: [9606](#)

**Lineage** Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo

**Data source** [RefSeq](#) (Assembly: [GCF\\_000001405.39](#))  
[BioProject: 168](#)

**Original DB** [NCBI](#), [OMIM](#), [HGNC](#), [HPRD](#), [Ensembl](#), [Vega](#), [Pharos](#)

**Statistics** Number of protein genes: 19768  
Number of RNA genes: 2641

**Created** 2000

**Reference** PMID: [11237011](#)

**Authors** Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, et al.

**Title** Initial sequencing and analysis of the human genome.

**Journal** Nature 409:860-921 (2001)  
DOI: [10.1038/35057062](#)

# 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
- Virus : ウイルスゲノム
- Meta : メタゲノム



## KEGG Organisms: Complete Genomes

Eukaryotes: 782    Bacteria: 7141    Archaea: 395

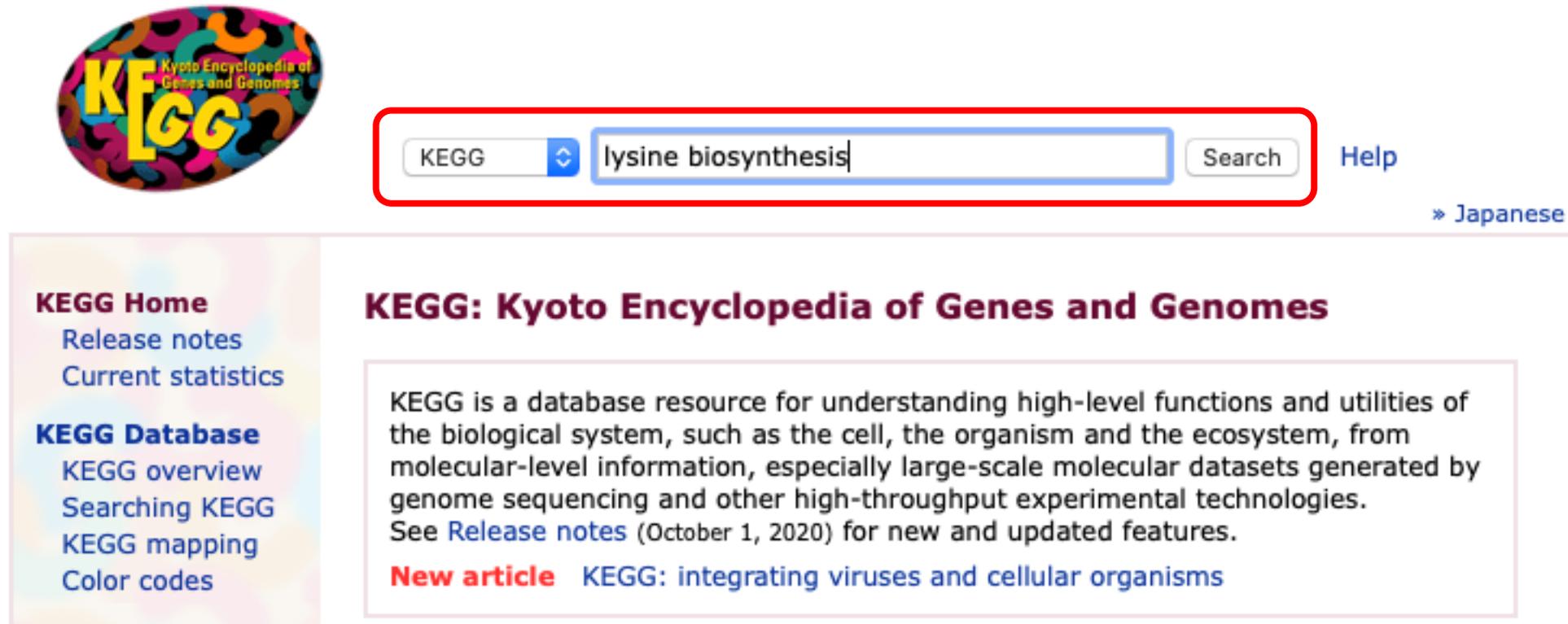
[ [Genomes](#) | [Species](#) | [Genus](#) | [Viruses](#) | [Meta](#) ]

### Eukaryotes

| Category |  | Organisms |     |   | Source |
|----------|--|-----------|-----|---|--------|
|          |  | hsa       | KGB | <a href="#">Homo sapiens (human)</a>                              | RefSeq |
|          |  | ptr       | KGB | <a href="#">Pan troglodytes (chimpanzee)</a>                      | RefSeq |
|          |  | pps       | KGB | <a href="#">Pan paniscus (bonobo)</a>                             | RefSeq |
|          |  | gga       | KGB | <a href="#">Gorilla gorilla gorilla (western lowland gorilla)</a> | RefSeq |

# KEGGパスウェイマップを見る

- [トップページ](https://www.kegg.jp) (<https://www.kegg.jp>) 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 website interface. On the left is the KEGG logo, which is a colorful, abstract shape with the text "KEGG Kyoto Encyclopedia of Genes and Genomes". To the right of the logo is a search bar with a dropdown menu set to "KEGG", a search input field containing "lysine biosynthesis", and a "Search" button. A red rectangular box highlights the search bar area. To the right of the search bar are links for "Help" and "» Japanese". Below the search bar is a navigation menu with the following items:

- KEGG Home**
  - [Release notes](#)
  - [Current statistics](#)
- KEGG Database**
  - [KEGG overview](#)
  - [Searching KEGG](#)
  - [KEGG mapping](#)
  - [Color codes](#)

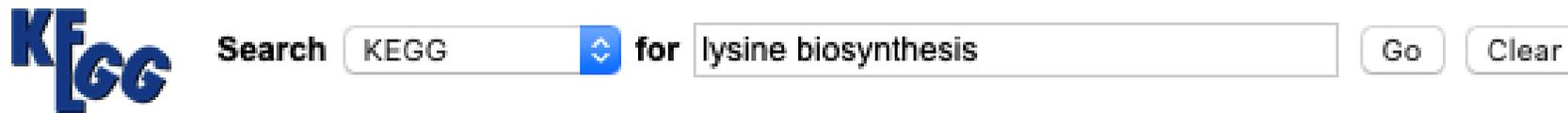
The main content area features the title "KEGG: Kyoto Encyclopedia of Genes and Genomes" in a large, bold, dark red font. Below the title is a text box containing the following information:

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](#) (October 1, 2020) for new and updated features.

**New article** [KEGG: integrating viruses and cellular organisms](#)

# 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。



Database: KEGG - Search term: lysine biosynthesis

## KEGG PATHWAY

[map00300](#)

Lysine biosynthesis

[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



# 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

KEGG Lysine biosynthesis - Reference pathway

[ Pathway menu | Pathway entry | Image (png) file | **Help** ]

Change pathway type

▼ Option

Scale: 100%

▼ Search

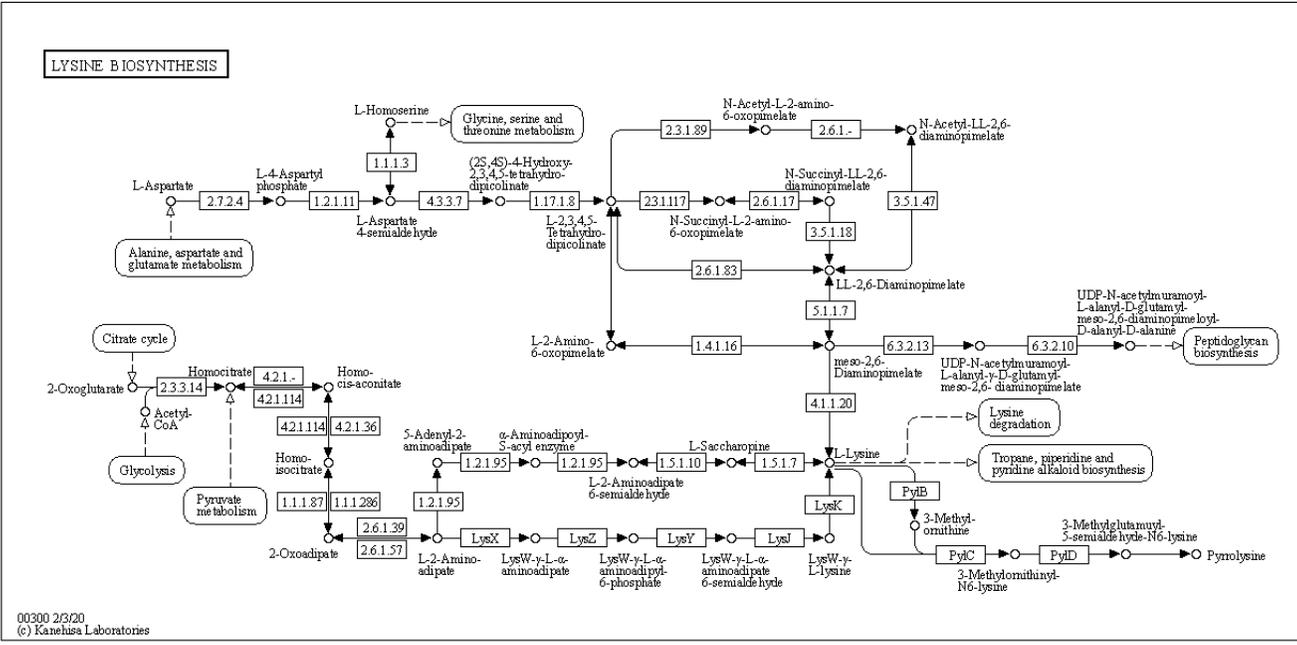
Go Clear

▼ User data

▼ Module

Pathway modules

- Energy metabolism
- Methane metabolism
- M00608 2-Oxocarboxylic acid metabolism
- Amino acid metabolism
- Lysine metabolism
- M00016 Lysine biosynthesis
- M00525 Lysine biosynthesis
- M00526 Lysine biosynthesis
- M00527 Lysine biosynthesis
- M00030 Lysine biosynthesis
- M00433 Lysine biosynthesis
- M00031 Lysine biosynthesis



KEGG Pathway Map

The KEGG PATHWAY database is a collection of manually drawn graphical diagrams, called KEGG pathway maps, representing molecular pathways for metabolism, genetic information processing, environmental information processing, cellular processes, organismal systems, human diseases, and drug development. Each pathway is identified by a five-digit number preceded by one of: map, ko, ec, rn, and three- or four-letter organism code. The pathway map is drawn and updated with the notation shown below.

| Objects |  | Arrows |                                       |
|---------|--|--------|---------------------------------------|
|         | gene product, mostly protein but including RNA |        | molecular interaction or relation     |
|         | other molecule, mostly chemical compound       |        | link to another map                   |
|         | another map                                    |        | pointer used in legend                |
|         |  |        | missing interaction (eg, by mutation) |

| Protein-protein interactions |                       | Gene expression relations |                               |
|------------------------------|-----------------------|---------------------------|-------------------------------|
|                              | phosphorylation       |                           | expression                    |
|                              | dephosphorylation     |                           | repression                    |
|                              | ubiquitination        |                           | expression                    |
|                              | glycosylation         |                           | indirect effect               |
|                              | methylation           |                           |                               |
|                              | activation            |                           |                               |
|                              | inhibition            |                           | two successive reaction steps |
|                              | indirect effect       |                           |                               |
|                              | state change          |                           |                               |
|                              | binding / association |                           |                               |
|                              | dissociation          |                           |                               |
|                              | complex               |                           |                               |

**Color coding**

The pathway map without coloring is the original version that is manually drawn by in-house software called KegSketch. The other pathway maps with coloring are all computationally generated as follows.

- Reference pathway: this is the original version; white boxes are hyperlinked to KO, ENZYME, and REACTION entries in metabolic pathways; they are hyperlinked to KO entries in non-metabolic pathways.
- Reference pathway (KO): blue boxes are hyperlinked to KO entries that are selected from the original version.
- Reference pathway (EC): blue boxes are hyperlinked to ENZYME entries that are selected from the original version.
- Reference pathway (Reaction): blue boxes are hyperlinked to REACTION entries that are selected from the original version.
- Organism-specific pathway: green boxes are hyperlinked to GENES entries by converting K numbers (KO identifiers) to gene identifiers in the reference pathway, indicating the presence of genes in the genome and also the completeness of the pathway.

# KEGGパスウェイマップ

Change pathway type

▼ Option

Scale: 100%

▼ Search

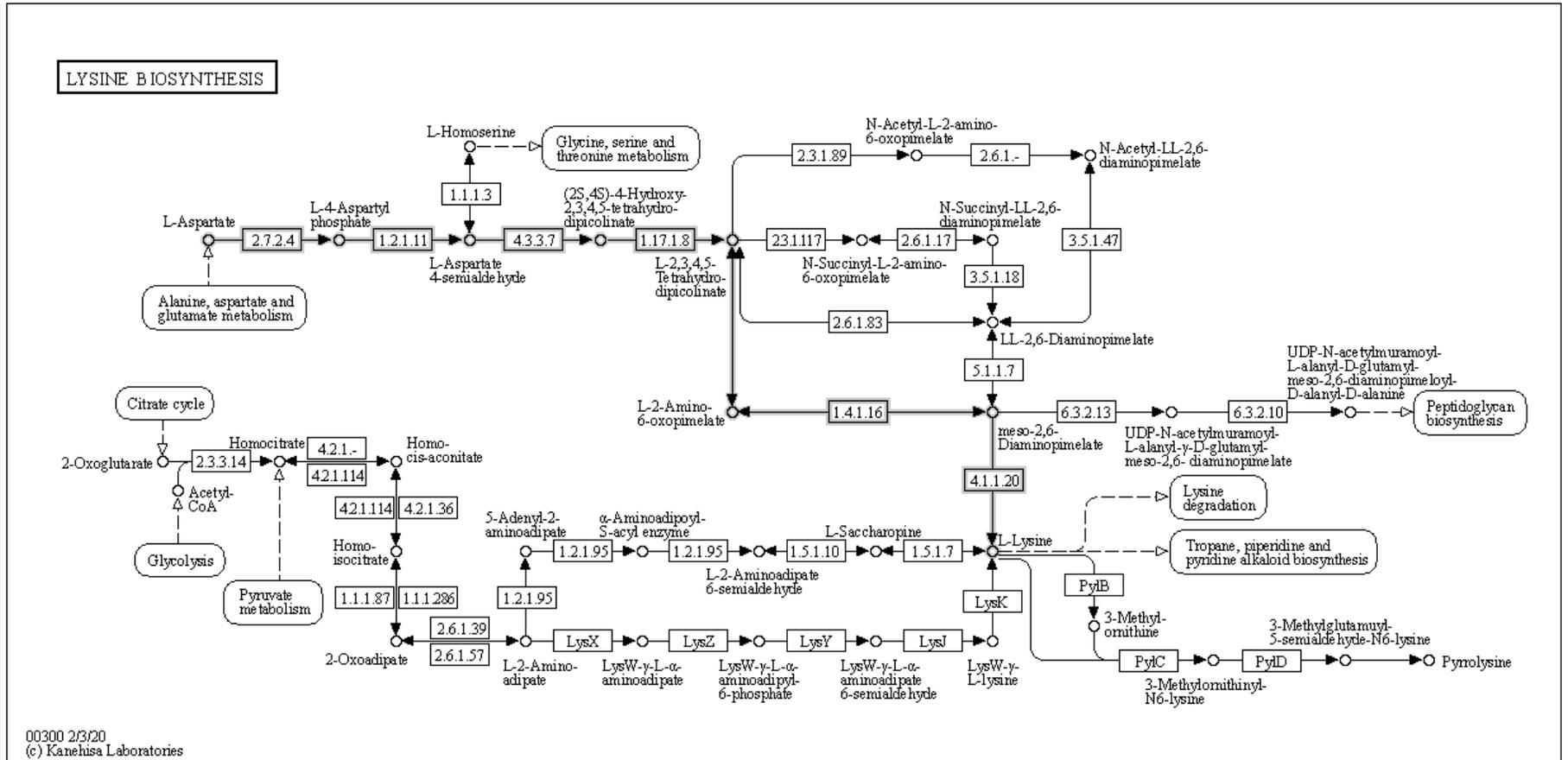
Go Clear

▼ User data +

▼ Module

- Pathway modules
- Energy metabolism
- Methane metabolism
  - M00608 2-Oxocarboxylic acid metabolism
- Amino acid metabolism
  - Lysine metabolism
    - M00016 Lysine biosynthesis
    - M00525 Lysine biosynthesis
    - M00526 Lysine biosynthesis
    - M00527 Lysine biosynthesis
  - M00030 Lysine biosynthesis
  - M00031 Lysine biosynthesis

オンマウスオーバーで、KEGG MODULEの表示





# 好きな生物のパスウェイを見る

- Change pathway type をクリック

**KEGG Lysine biosynthesis - Reference pathway**

[ Pathway menu | Pathway entry | Image (png) file | Help ]

Change pathway type

Scale: 100%

Search

User data

Module

- Pathway modules
  - Energy metabolism
  - Methane metabolism
    - M00608 2-Oxocarboxylic acid metabolism
  - Amino acid metabolism
    - Lysine metabolism
      - M00016 Lysine biosynthesis**
      - M00525 Lysine biosynthesis
      - M00526 Lysine biosynthesis
      - M00527 Lysine biosynthesis
      - M00030 Lysine biosynthesis
      - M00433 Lysine biosynthesis
      - M00031 Lysine biosynthesis

LYSINE BIOSYNTHESIS

00300 2/3/20  
(c) Kanehisa Laboratories

# 好きな生物のパスウェイを見る

## Select pathway type for 00300

▼▼▼▼

### ▼ Reference

- [map](#) Reference pathway
- [ko](#) Reference pathway (KO only)
- [ec](#) Reference pathway (EC only)
- [rn](#) Reference pathway (Reaction only)

### ▼ Organism specific

- ▼ Animals
  - ▼ Mammals
    - [chx](#) *Capra hircus* (goat) 2/39
    - [pcad](#) *Physeter catodon* (sperm whale) 2/39
  - ▼ Lancelets
    - [bfo](#) *Branchiostoma floridae* (Florida lancelet) 2/39
  - ▼ Insects
    - [mpa](#) *Monomorium pharaonis* (pharaoh ant) 2/39
    - [fcd](#) *Folsomia candida* 2/39
  - ▼ Crustaceans
    - [dpx](#) *Daphnia pulex* (common water flea) 2/39
    - [pvm](#) *Penaeus vannamei* (Pacific white shrimp) 3/39
  - ▼ Mollusks
    - [lgi](#) *Lottia gigantea* (owl limpet) 2/39
    - [pcan](#) *Pomacea canaliculata* (golden apple snail) 3/39
    - [crg](#) *Crassostrea gigas* (Pacific oyster) 2/39
    - [myi](#) *Mizuhopecten yessoensis* (Yesso scallop) 2/39
  - ▼ Brachiopods
    - [lak](#) *Lingula anatina* 2/39
  - ▼ Cnidarians
    - [nve](#) *Nematostella vectensis* (starlet sea anemon) 2/39
    - [spis](#) *Stylophora pistillata* 2/39
- ▼ Plants
  - ▼ Eudicots
    - [ath](#) *Arabidopsis thaliana* (thale cress) 8/39
    - [aly](#) *Arabidopsis lyrata* (lyrate rockcress) 8/39
    - [crb](#) *Capsella rubella* 8/39
    - [csat](#) *Camelina sativa* (false flax) 8/39
    - [eus](#) *Eutrema salsugineum* 8/39
    - [brp](#) *Brassica rapa* (field mustard) 8/39
    - [bna](#) *Brassica napus* (rape) 8/39
    - [boe](#) *Brassica oleracea* (wild cabbage) 8/39
    - [rsz](#) *Raphanus sativus* (radish) 8/39
    - [thj](#) *Tarenaya hassleriana* (spider flower) 8/39
    - [cpap](#) *Carica papaya* (papaya) 8/39
    - [cit](#) *Citrus sinensis* (Valencia orange) 8/39
    - [cic](#) *Citrus clementina* (mandarin orange) 8/39
    - [tcs](#) *Thalictrum flavum* (poppy) 8/39

Select pathway typeのパネルがポップアップ

見たい生物種の名前で検索 (例えば、Escherichia coli)

ecoをクリック

- [ngr](#) *Naegleria gruberi* 3/39
- ▼ Bacteria
  - ▼ Gammaproteobacteria - Enterobacteria
    - eco** *Escherichia coli* K-12 MG1655 12/39
    - [ecj](#) *Escherichia coli* K-12 W3110 12/39
    - [ecd](#) *Escherichia coli* K-12 DH10B 12/39
    - [ebw](#) *Escherichia coli* K-12 BW2952 12/39
    - [ecok](#) *Escherichia coli* K-12 MDS42 12/39
    - [ece](#) *Escherichia coli* O157:H7 EDL933 (EHEC) 12/39
    - [ecs](#) *Escherichia coli* O157:H7 Sakai (EHEC) 12/39
    - [ecf](#) *Escherichia coli* O157:H7 EC4115 (EHEC) 12/39
    - [etw](#) *Escherichia coli* O157:H7 TW14359 (EHEC) 12/39
    - [elx](#) *Escherichia coli* O157:H7 Xuzhou21 (EHEC) 12/39
    - [eoi](#) *Escherichia coli* O111:H- 11128 (EHEC) 12/39
    - [eoj](#) *Escherichia coli* O26:H11 11368 (EHEC) 12/39
    - [eoh](#) *Escherichia coli* O103:H2 12009 (EHEC) 12/39
    - [ecoo](#) *Escherichia coli* O145:H28 RM13514 (EHEC) 12/39
    - [ecoh](#) *Escherichia coli* O145:H28 RM13516 (EHEC) 12/39
    - [esl](#) *Escherichia coli* O104:H4 2011C-3493 (EAEC) 12/39
    - [eso](#) *Escherichia coli* O104:H4 2009EL-2071 (EAEC) 12/39
    - [esm](#) *Escherichia coli* O104:H4 2009EL-2050 (EAEC) 12/39
    - [eck](#) *Escherichia coli* 55989 (EAEC) 12/39
    - [ecg](#) *Escherichia coli* O127:H6 E2348/69 (EPEC) 12/39
    - [eok](#) *Escherichia coli* O55:H7 CB9615 (EPEC) 12/39
    - [elr](#) *Escherichia coli* O55:H7 RM12579 (EPEC) 12/39
    - [elh](#) *Escherichia coli* O78:H11:K80 H10407 (ETEC) 12/39

# Escherichia coli K-12 MG1655ゲノムでマッピング

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。



Lysine biosynthesis - Escherichia coli K-12 MG1655

[ [Pathway menu](#) | [Pathway entry](#) | [Download KGML](#) | [Image \(png\) file](#) | [Help](#) ]

Change pathway type

▼ Option

Scale: 100%

▼ Search

Go Clear

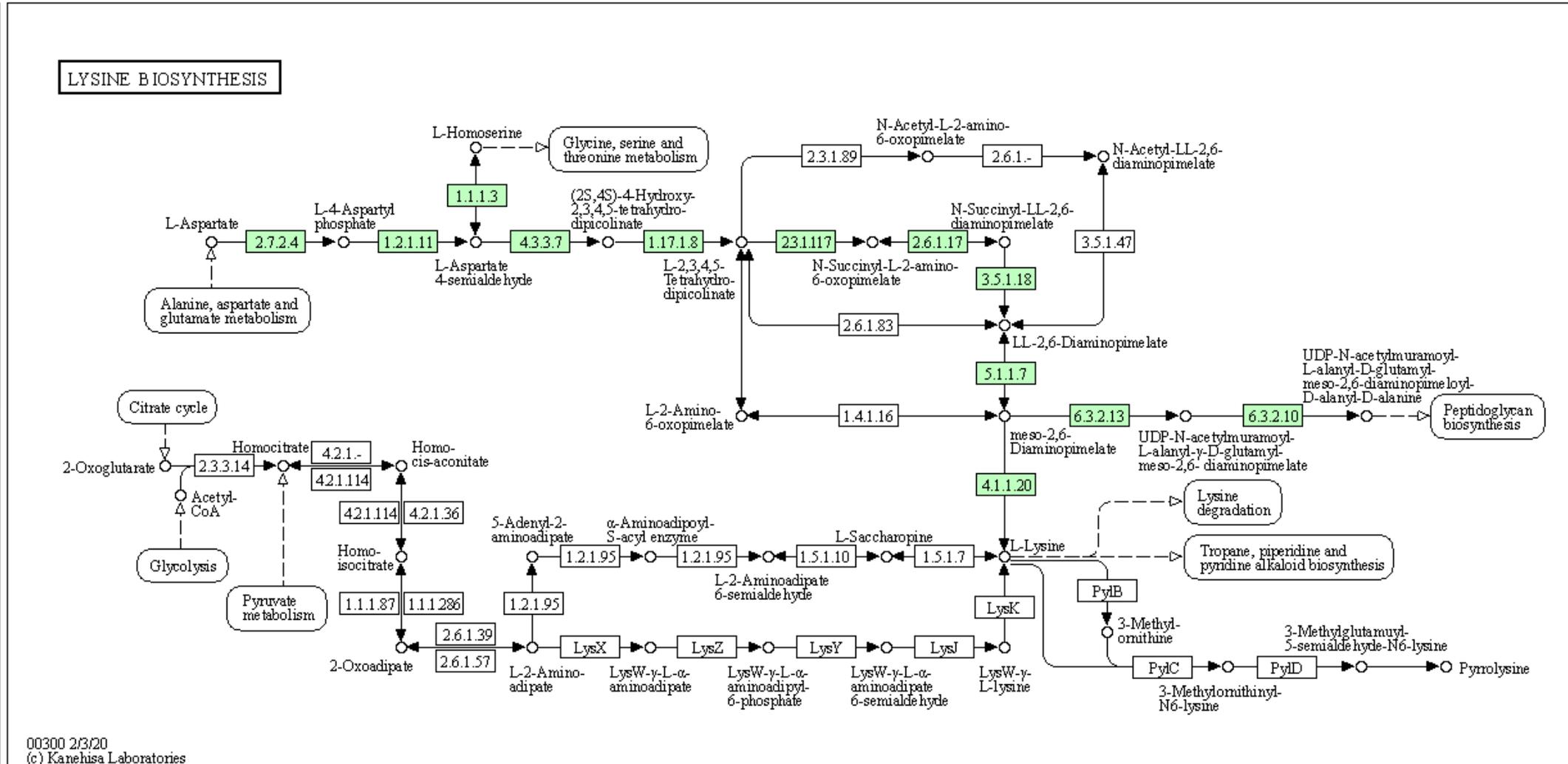
▼ User data +

▼ Module

Complete only

Pathway modules

- Energy metabolism
- Methane metabolism
- Amino acid metabolism
- Lysine metabolism
- M00016 Lysine biosynthesis



# 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   |
|----------------------|--|
| <a href="#">eco</a>  | <a href="#">Escherichia coli K-12 MG1655</a>             |
| <a href="#">ecj</a>  | <a href="#">Escherichia coli K-12 W3110</a>              |
| <a href="#">ecd</a>  | <a href="#">Escherichia coli K-12 DH10B</a>              |
| <a href="#">ebw</a>  | <a href="#">Escherichia coli BW2952</a>                  |
| <a href="#">ecok</a> | <a href="#">Escherichia coli K-12 MDS42</a>              |
| <a href="#">ece</a>  | <a href="#">Escherichia coli O157:H7 EDL933 (EHEC)</a>   |
| <a href="#">ecs</a>  | <a href="#">Escherichia coli O157:H7 Sakai (EHEC)</a>    |
| <a href="#">ecf</a>  | <a href="#">Escherichia coli O157:H7 EC4115 (EHEC)</a>   |
| <a href="#">etw</a>  | <a href="#">Escherichia coli O157:H7 TW14359 (EHEC)</a>  |
| <a href="#">elx</a>  | <a href="#">Escherichia coli O157:H7 Xuzhou21 (EHEC)</a> |
| <a href="#">eoj</a>  | <a href="#">Escherichia coli O26:H11 11368 (EHEC)</a>    |
| <a href="#">eoi</a>  | <a href="#">Escherichia coli O111:H- 11128 (EHEC)</a>    |

## KEGG Escherichia coli K-12 MG1655

Genome info **Pathway map** Brite hierarchy Module Genome map Blast Taxonomy

Search genes:

### Genome information

|                   |                              |
|-------------------|------------------------------|
| <b>T number</b>   | T00007                       |
| <b>Org code</b>   | eco                          |
| <b>Aliases</b>    | ECOLI, 511145                |
| <b>Full name</b>  | Escherichia coli K-12 MG1655 |
| <b>Definition</b> | Escherichia coli K-12 MG1655 |
| <b>Category</b>   | Reference genome             |
| <b>Annotation</b> | yes                          |
| <b>Taxonomy</b>   | TAX: 511145                  |

|  |      |         |
|--|------|---------|
|  | 2012 | GenBank |
|  | 2009 | GenBank |
|  | 2009 | GenBank |

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

# Lysine biosynthesis マップ

- 00230 Purine metabolism
- 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
  - 00410 beta-Alanine metabolism
  - 00430 Taurine and hypotaurine metabolism

KEGG Lysine biosynthesis - Escherichia coli K-12 MG1655

[ Pathway menu | Pathway entry | Download KGML | Image (png) file | Help ]

Change pathway type

▼ Option

Scale:  100%

▼ Search

▼ User data

▼ Module

Complete only

Pathway modules

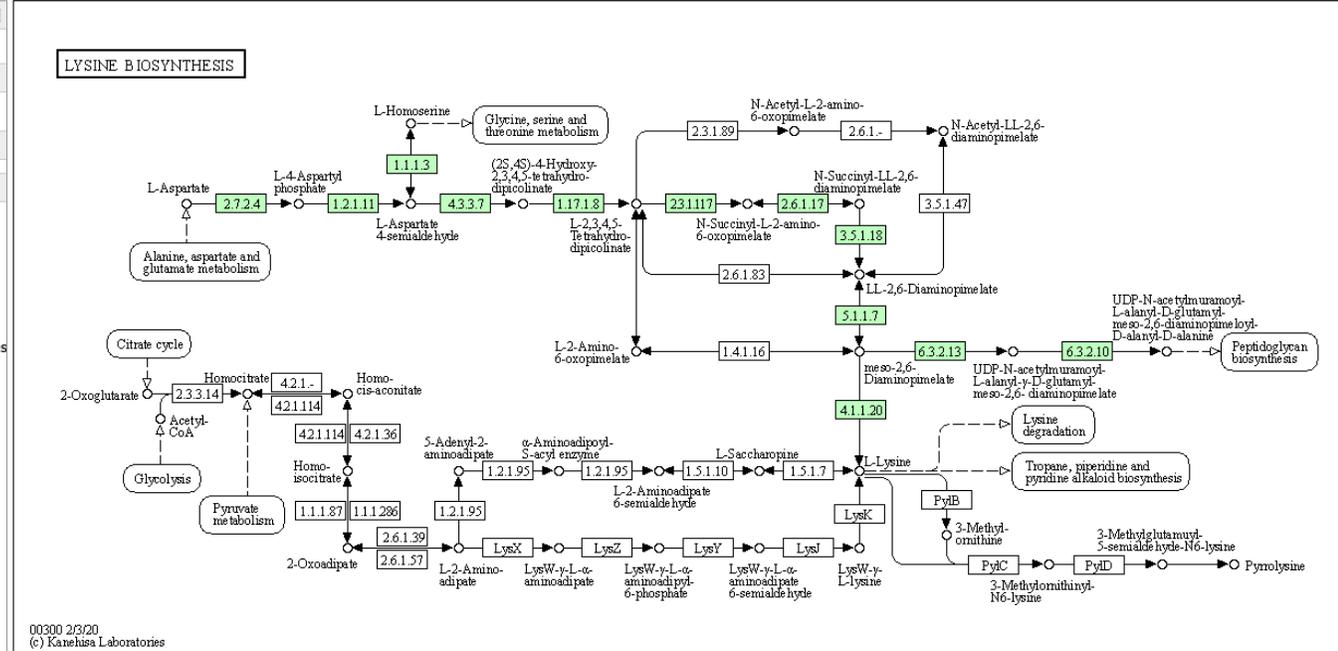
Energy metabolism

Methane metabolism

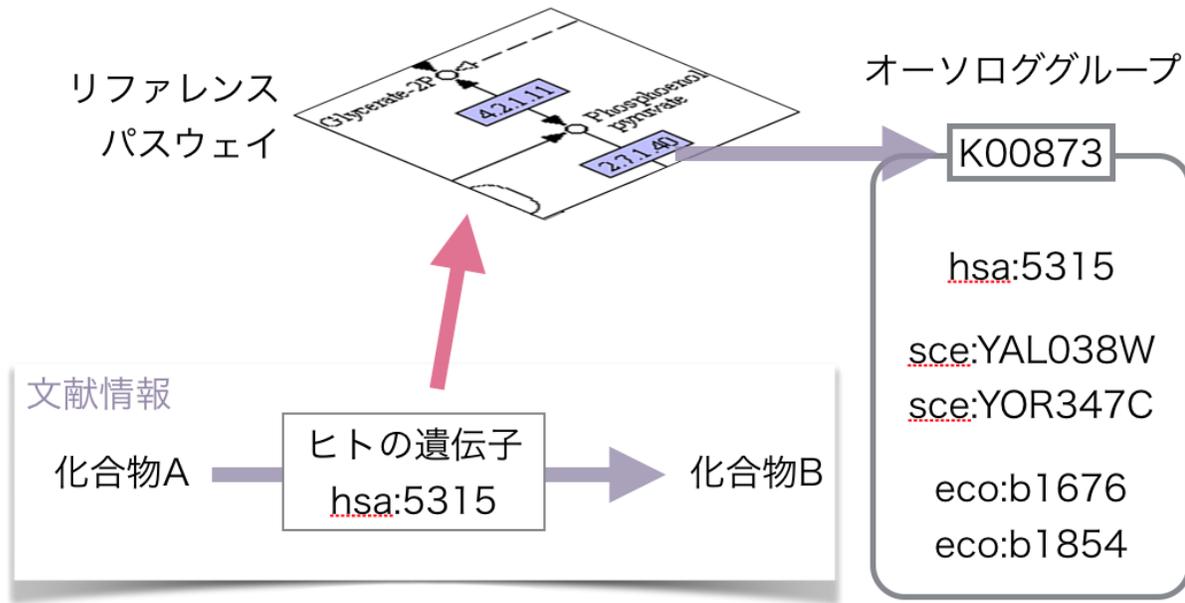
Amino acid metabolism

Lysine metabolism

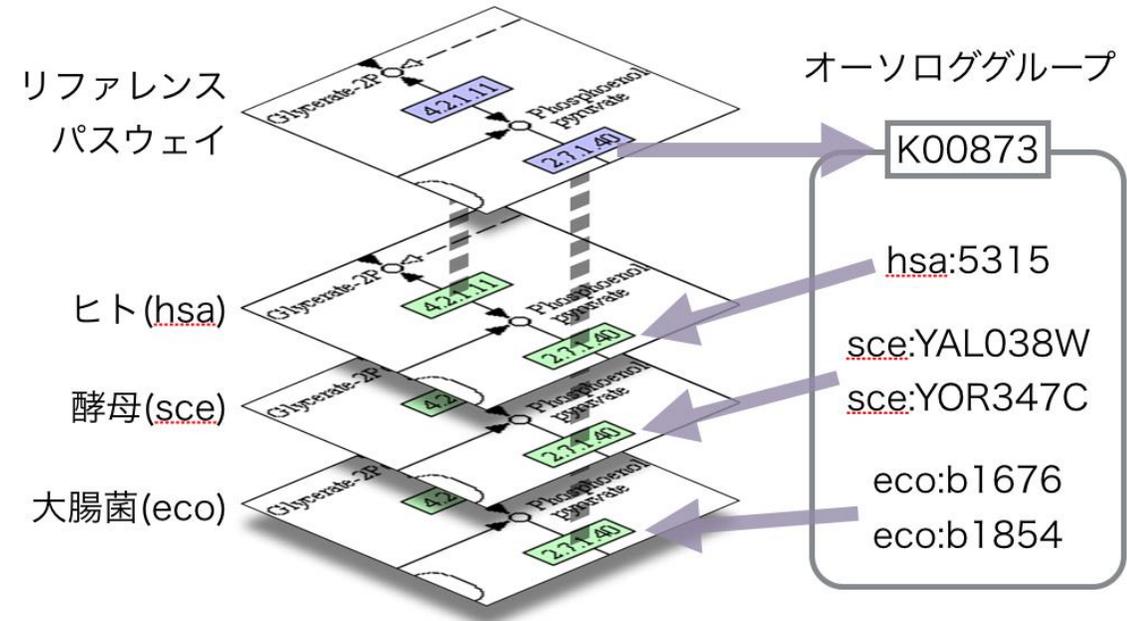
M00016 Lysine biosynthesis



# リファレンスパスウェイと種毎のパスウェイの関係



文献情報からリファレンスパスウェイが作成される



KOと化合物のネットワークであるリファレンスパスウェイを基に、生物種毎のパスウェイが作られる

# 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
- 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 M [Metabolic pathways](#)
- 01110 M [Biosynthesis of secondary metabolites](#)
- 01120 M [Microbial metabolism in diverse environments](#)
- 01200 M R [Carbon metabolism](#)
- 01210 M R [2-Oxocarboxylic acid metabolism](#)
- 01212 M R [Fatty acid metabolism](#)
- 01230 M R [Biosynthesis of amino acids](#)
- 01220 M R [Degradation of aromatic compounds](#)

### 1.1 Carbohydrate metabolism

- 00010 M N [Glycolysis / Gluconeogenesis](#)

Change pathway type

**▼ Option**

Scale:  30%

Link:

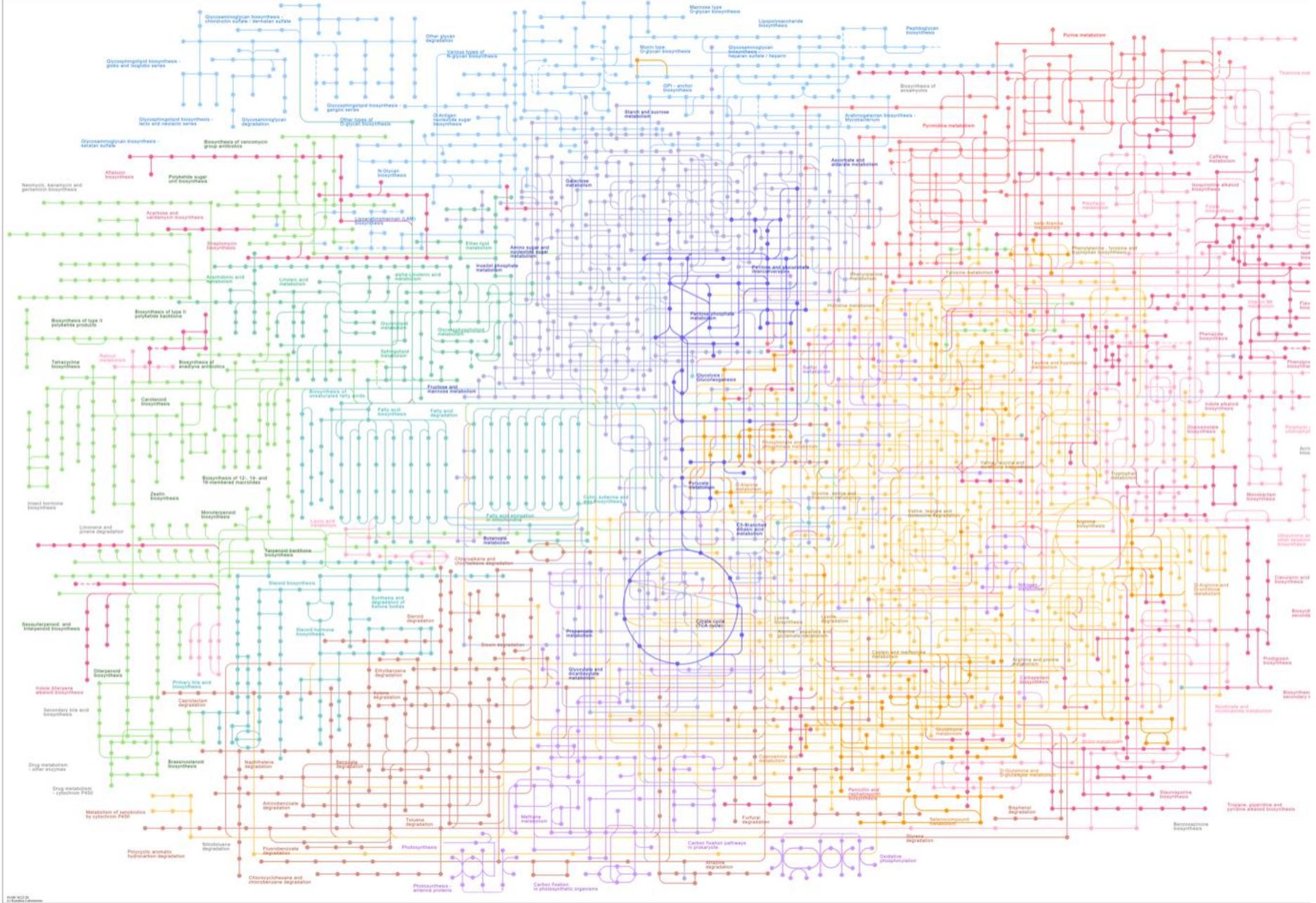
**▼ Search**

**▼ User data** +

**▼ Module**

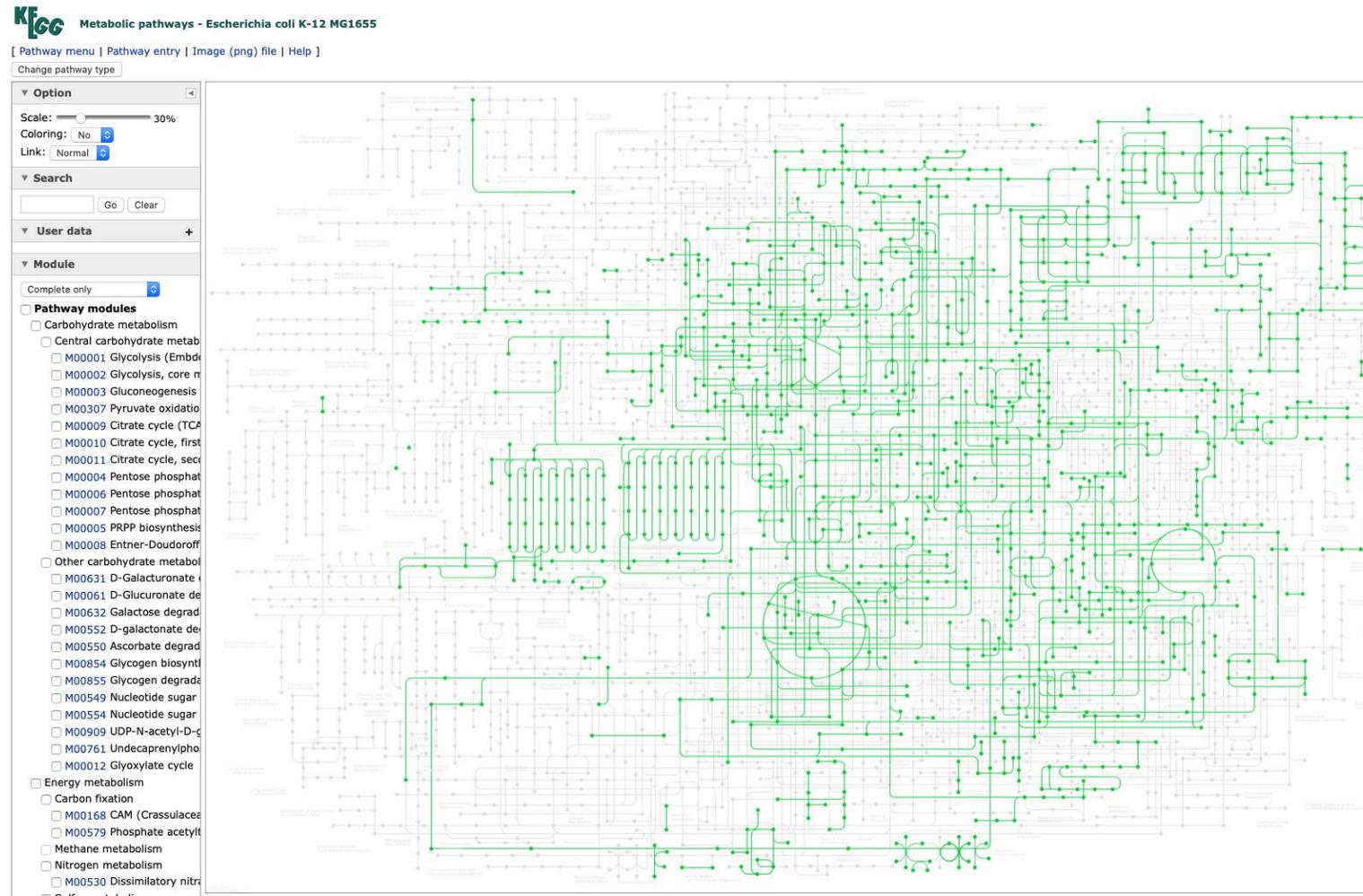
**Pathway modules**

- Carbohydrate metabolism
  - Central carbohydrate metab
    - M00001 Glycolysis (Embde
    - M00002 Glycolysis, core n
    - M00003 Gluconeogenesis
    - M00307 Pyruvate oxidatio
    - M00009 Citrate cycle (TCA
    - M00010 Citrate cycle, first
    - M00011 Citrate cycle, seco
    - M00004 Pentose phosphat
    - M00006 Pentose phosphat
    - M00007 Pentose phosphat
    - M00580 Pentose phosphat
    - M00005 PRPP biosynthesis
    - M00008 Entner-Doudoroff
    - M00308 Semi-phosphoryla
    - M00633 Semi-phosphoryla
    - M00309 Non-phosphorylat
  - Other carbohydrate metabol
    - M00014 Glucuronate path
    - M00630 D-Galacturonate
    - M00631 D-Galacturonate
    - M00061 D-Glucuronate de
    - M00081 Pectin degradatio
    - M00632 Galactose degrad
    - M00552 D-galactonate de
    - M00129 Ascorbate biosynt
    - M00114 Ascorbate biosynt
    - M00550 Ascorbate degrad
    - M00854 Glycogen biosynt
    - M00855 Glycogen degrad
    - M00565 Trehalose biosynt
    - M00549 Nucleotide sugar
    - M00554 Nucleotide sugar
    - M00892 UDP-N-acetyl-D-c
    - M00909 UDP-N-acetyl-D-c
    - M00364 UDP-glucose-6-ph



# 生物種毎のOverviewマップを見る

- Change pathway type から生物コードをクリック
- 生物の持っていない経路は灰色になる (ecoの例)



# ヒトの疾患パスウェイを見る

- <https://www.kegg.jp/kegg/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。

## ▼ Human Diseases

### ▼ Cancer: overview

- 05200 Pathways in cancer
- 05202 Transcriptional misregulation
- 05206 MicroRNAs in cancer
- 05205 Proteoglycans in cancer
- 05204 Chemical carcinogenesis - DNA
- 05207 Chemical carcinogenesis - rece
- 05208 Chemical carcinogenesis - reac
- 05203 Viral carcinogenesis
- 05230 Central carbon metabolism in c
- 05231 Choline metabolism in cancer
- 05235 PD-L1 expression and PD-1 chec

### ▼ Cancer: specific types

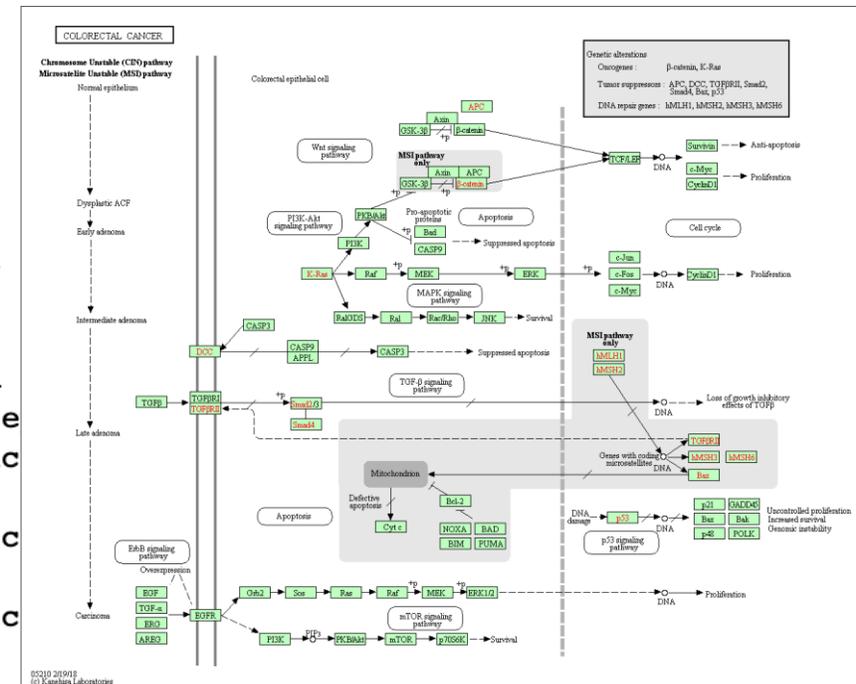
- 05210 Colorectal cancer
- 05212 Pancreatic cancer
- 05225 Hepatocellular carcinoma

## Disease Pathway Maps

The Human Diseases category of the KEGG PATHM

### KEGG Pathway Maps: Human Diseases

It contains multifactorial diseases such as cancers, diseases, and metabolic diseases where known dis interacting molecular networks of both pathogens



# ヒトの疾患パスウェイ

- [Disease genes and drug targets in KEGG pathways](#)から「Colorectal cancer」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

## Disease Mapping

The [Search Disease](#) tool in [KEGG Mapper](#) is a mapping tool against disease genes accumulated in KEGG DISEASE entries, together with related KEGG pathway maps. The user may upload a set of human genes or KOs to check if any diseases may be associated with the dataset.

[KEGG Mapper - Search Disease](#)

## Pathway/Brite mapping of disease genes and drug targets

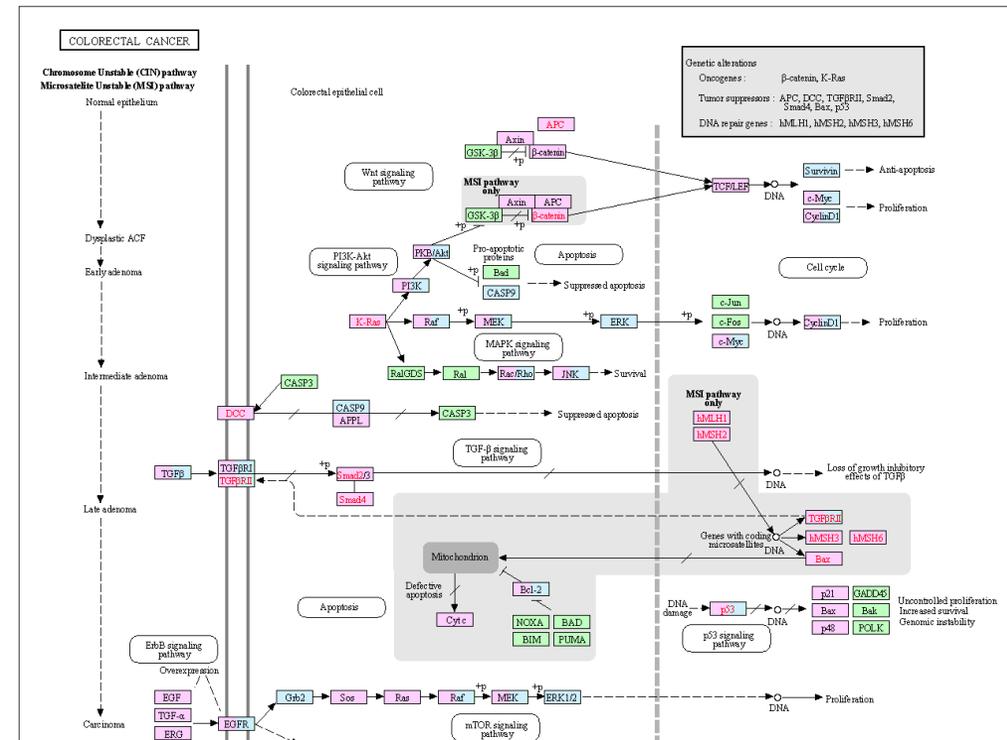
Disease genes accumulated in the KEGG DISEASE database and drug targets stored in the KEGG DRUG database are often represented in the KEGG PATHWAY and BRITE databases as well. The pathway maps and BRITE hierarchy files with mapping of disease genes and drug targets are identified by the five-letter organism code "hsadd" and the extension code "\_dd", respectively.

For example, [hsadd04620](#) represents disease/drug mapped toll-like receptor signaling pathway with the coloring convention as follows:

- When the gene is associated with a disease, it is marked in pink.
- When the gene (product) is a drug target, it is marked in light blue.
- When the gene is both a disease gene and a drug target, its coloring is split into pink and light blue.

The disease/drug mapped version is precomputed in the daily KEGG database update procedure, and each map or BRITE hierarchy can be selected from the organisms selection menu. For the disease/drug pathway maps the following list may also be used.

[Disease genes and drug targets in KEGG pathways](#)



# 種間比較をする

- <https://www.kegg.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                                     | <a href="#">KEGG PATHWAY</a><br><a href="#">KEGG BRITE</a><br><a href="#">KEGG MODULE</a>   | KEGG pathway maps<br>BRITE hierarchies and tables<br>KEGG modules  | <a href="#">PATHWAY</a><br><a href="#">BRITE</a><br><a href="#">MODULE</a>   |
| Genomic information                                     | <a href="#">KO (KEGG Orthology)</a>   | Functional orthologs   | <a href="#">ORTHOLOGY</a>  |
|   | <a href="#">KEGG GENOME</a><br><a href="#">KEGG GENES</a><br><a href="#">KEGG SSDB</a>  | KEGG organisms (complete genomes)<br>Genes and proteins<br>GENES sequence similarity   | <a href="#">GENOME</a><br><a href="#">GENES</a>  |
| Chemical information<br>( <a href="#">KEGG LIGAND</a> ) | <a href="#">KEGG COMPOUND</a><br><a href="#">KEGG GLYCAN</a><br><a href="#">KEGG REACTION</a><br><a href="#">KEGG ENZYME</a>                              | Small molecules<br>Glycans<br>Reactions and reaction classes<br>Enzyme nomenclature  | <a href="#">COMPOUND</a><br><a href="#">GLYCAN</a><br><a href="#">REACTION</a> <a href="#">RCLASS</a><br><a href="#">ENZYME</a>                      |
| Health information                                      | <a href="#">KEGG NETWORK</a><br><a href="#">KEGG DISEASE</a><br><a href="#">KEGG DRUG</a><br><a href="#">KEGG ENVIRON</a><br><a href="#">KEGG MEDICUS</a> | Disease-related network elements<br>Human diseases<br>Drugs and drug groups<br>Health related substances<br>Japanese drug labels ( <a href="#">JAPIC</a> )<br>FDA drug labels ( <a href="#">DailyMed</a> ) | <a href="#">NETWORK</a> <a href="#">VARIANT</a><br><a href="#">DISEASE</a><br><a href="#">DRUG</a> <a href="#">DGROUP</a><br><a href="#">ENVIRON</a> |

KEGG is developed by Kanehisa Laboratories. See [Kanehisa et al. \(2021\)](#) 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

(Host-microbiome relationship)

[hsa+T30003](#) Homo sapiens (human) + gut microbiome such as [Global metabolism map](#) where green for human, red for gut microbiome, and blue 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
  - 前者の遺伝子が緑、後者の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

## 00020 [Citrate cycle \(TCA cycle\)](#)

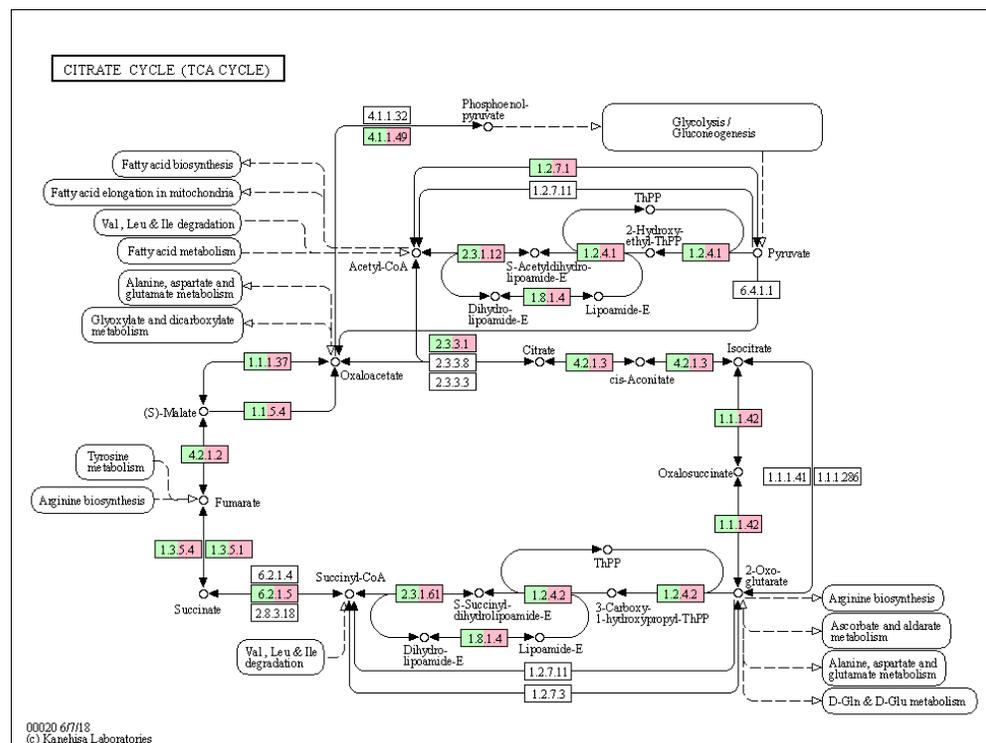
KEGG Organism group: eco ecs

Category info **Pathway map** Brite hierarchy Taxonomy

Search genes:  Go Clear

**eco ecs**

|        |     |                                       |
|--------|-----|---------------------------------------|
| T00007 | eco | Escherichia coli K-12 MG1655          |
| T00048 | ecs | Escherichia coli O157:H7 Sakai (EHEC) |

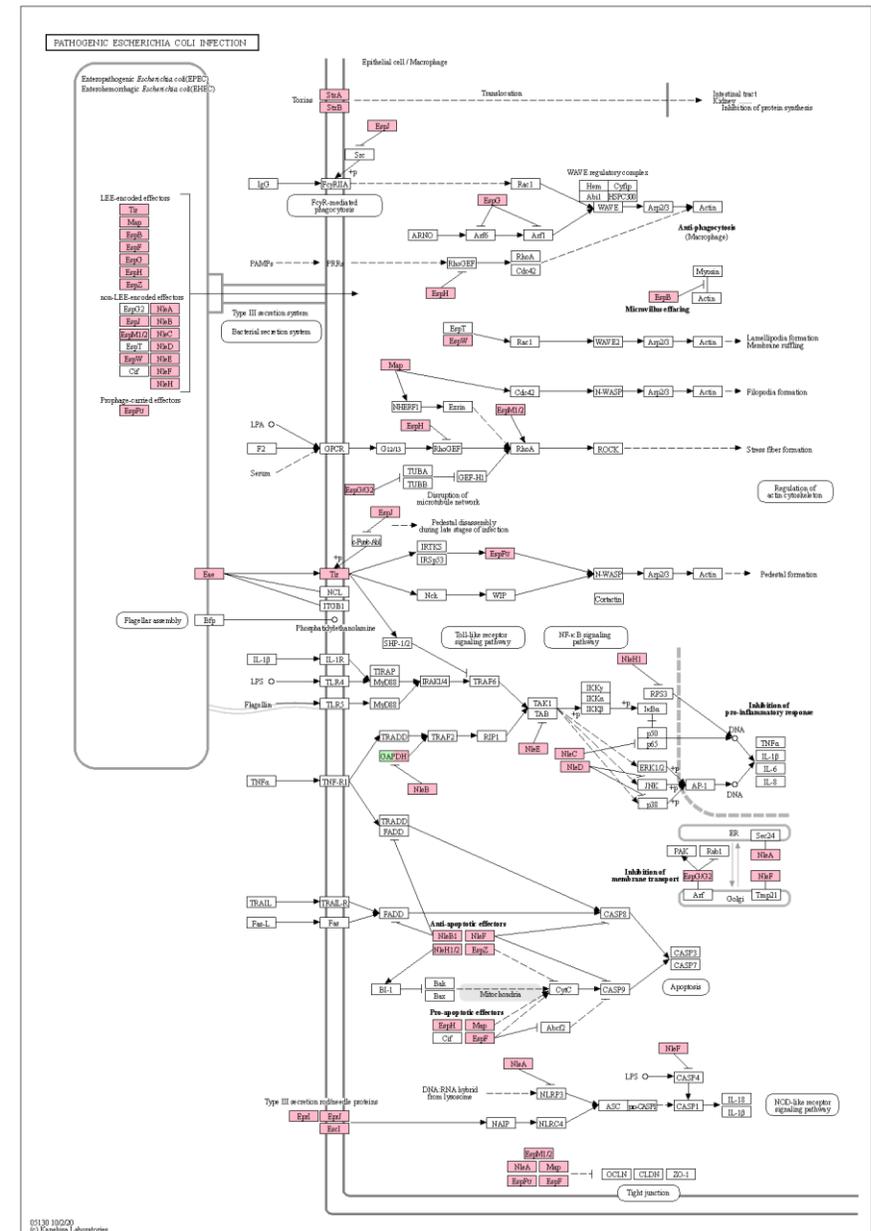




# Pathogenic Escherichia coli infection

Infectious disease: bacterial

- 05110 Vibrio cholerae infection
- 05120 Epithelial cell signaling in Helicobacter pylori infection
- 05130 Pathogenic Escherichia coli infection
- 05132 Salmonella infection
- 05131 Shigellosis
- 05135 Yersinia infection
- 05133 Pertussis
- 05134 Legionellosis
- 05150 Staphylococcus aureus infection
- 05152 Tuberculosis
- 05100 Bacterial invasion of epithelial cells

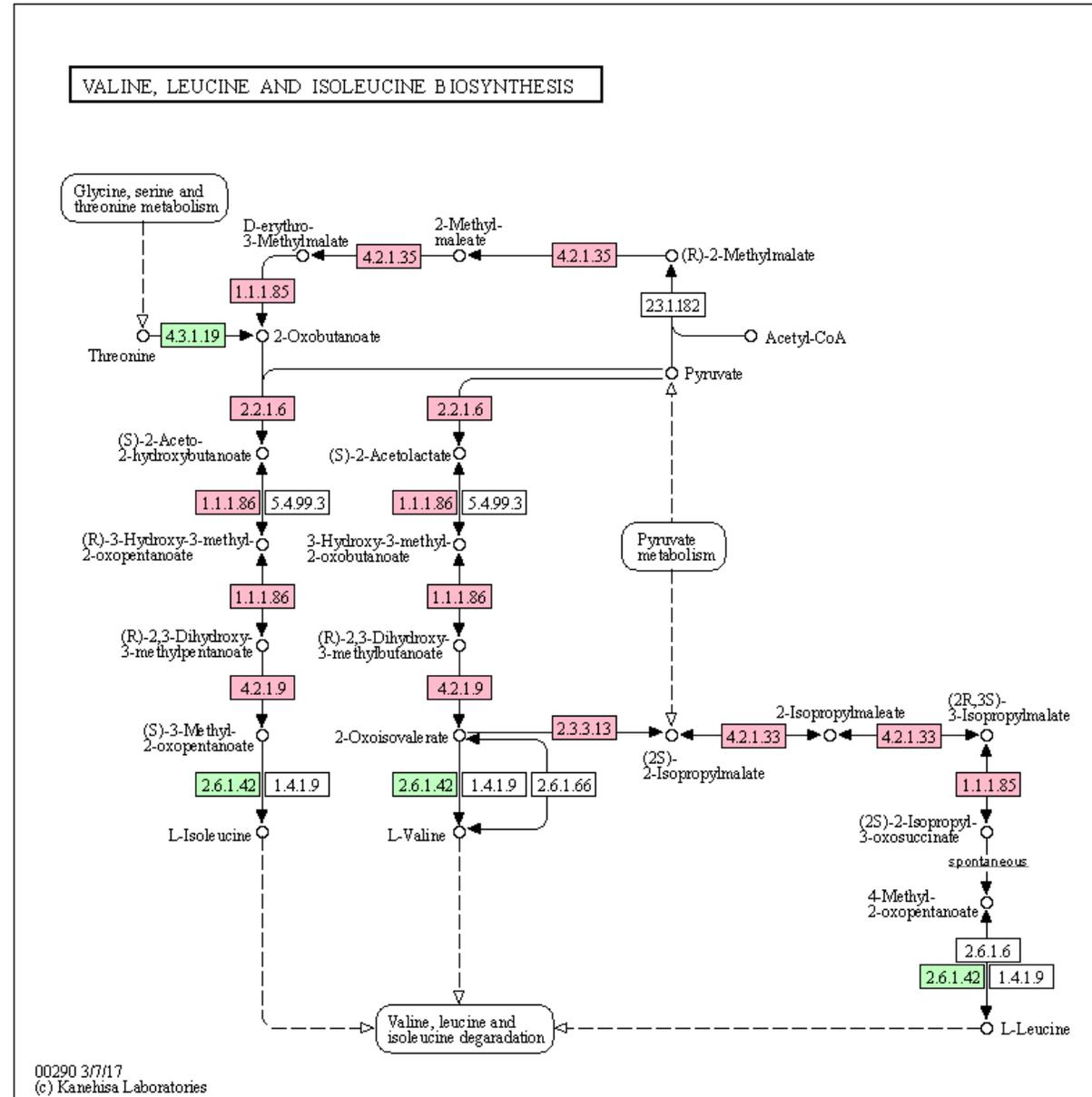


# アブラムシとブフネラで種間比較

- 00290 : Valine, leucine and isoleucine biosynthesisマップ
  - アブラムシ : api (緑)
  - ブフネラ : buc (ピンク)
- 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

## 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



# 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3番目の [Color](#) をクリック
    - Search mode : データベースコード
    - 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-GeneID, 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ネームで記述



## KEGG Mapper

A suite of KEGG mapping tools

[KEGG2](#) [About](#) [Reconstruct](#) [Search](#) [Color](#) [Join](#) [Convert ID](#) [Assign KO](#) [Taxonomy](#)

### About KEGG Mapper

**KEGG Mapper** is a collection of tools for [KEGG mapping](#) including popular KEGG pathway mapping, JOIN BRITE operations and MODULE completeness checks. Historically, "Search Pathway" and "Search&Color Pathway" tools were introduced at the beginning of the KEGG project. As the KEGG database contents expanded, so did the mapping tools. In the last version released in July 2019 [1], fourteen existing tools were reorganized into just five tools, allowing multiple mapping operations to be done at the same time and the result page shown in multiple tabs. This new version 5 of KEGG Mapper [2] has a further simplified architecture together with the enhancement of the join tool.

See new article: [KEGG mapping tools for uncovering hidden features in biological data](#)

### KEGG mapping tools

There are four KEGG mapping tools as summarized below.

**Reconstruct** (used to be called Reconstruct Pathway) is the basic mapping tool used for linking KO annotation (K number assignment) data to KEGG pathway maps, BRITE hierarchies and tables, and KEGG modules.

**Search** (used to be called Search Pathway) is the traditional tool for searching mapped objects in the user's dataset and mark them in red.

**Color** (used to be called Search&Color Pathway) is another traditional tool for searching mapped objects in the user's dataset and mark them in any combination of background and foreground colors. This tool now applies only to KEGG pathway maps. Use the Join tool for coloring of Brite hierarchies.

**Join** (used to be called Join Brite) is a tool to combine a Brite hierarchy file and a binary relation file, effectively adding a new column to the hierarchy file. This tool has been significantly expanded in this release.

# 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）



## KEGG Mapper – Color

[KEGG2](#) [About](#) [Reconstruct](#) [Search](#) [Color](#) [Join](#) [Convert ID](#) [Assign KO](#) [Taxonomy](#)

### Color tool

The Color tool searches various KEGG objects, including genes, KOs, EC numbers, metabolites and drugs, against KEGG pathway maps. Found objects may be marked in any combination of background and foreground colors.

See new article: [KEGG mapping tools for uncovering hidden features in biological data](#)

**Search mode:**  Reference  hsa  other org

### Enter KEGG identifiers followed by color specification

```
C00079 red
C00423 red
C00811 red
C00223 red
C06561 red
C00509 red
K00487 yellow
K00660 pink
K01859 cyan
```

Examples:

Reference

Or upload file:

Default bgcolor:

Use uncolored diagrams

Include aliases (for hsa and other org modes)

# KEGG Mapper Search Result

## Pathway (34)

Sort by the pathway list

Show matched objects

map01110 Biosynthesis of secondary metabolites (9)

map01100 Metabolic pathways (9)

**map00941 Flavonoid biosynthesis (6)**

map00940 Phenylpropanoid biosynthesis (5)

map01061 Biosynthesis of phenylpropanoids (5)

map00130 Ubiquinone and other terpenoid-quinone biosynthesis (5)

map01220 Degradation of aromatic compounds (4)

map00996 Biosynthesis of various alkaloids (3)

map00999 Biosynthesis of various plant secondary metabolites (3)

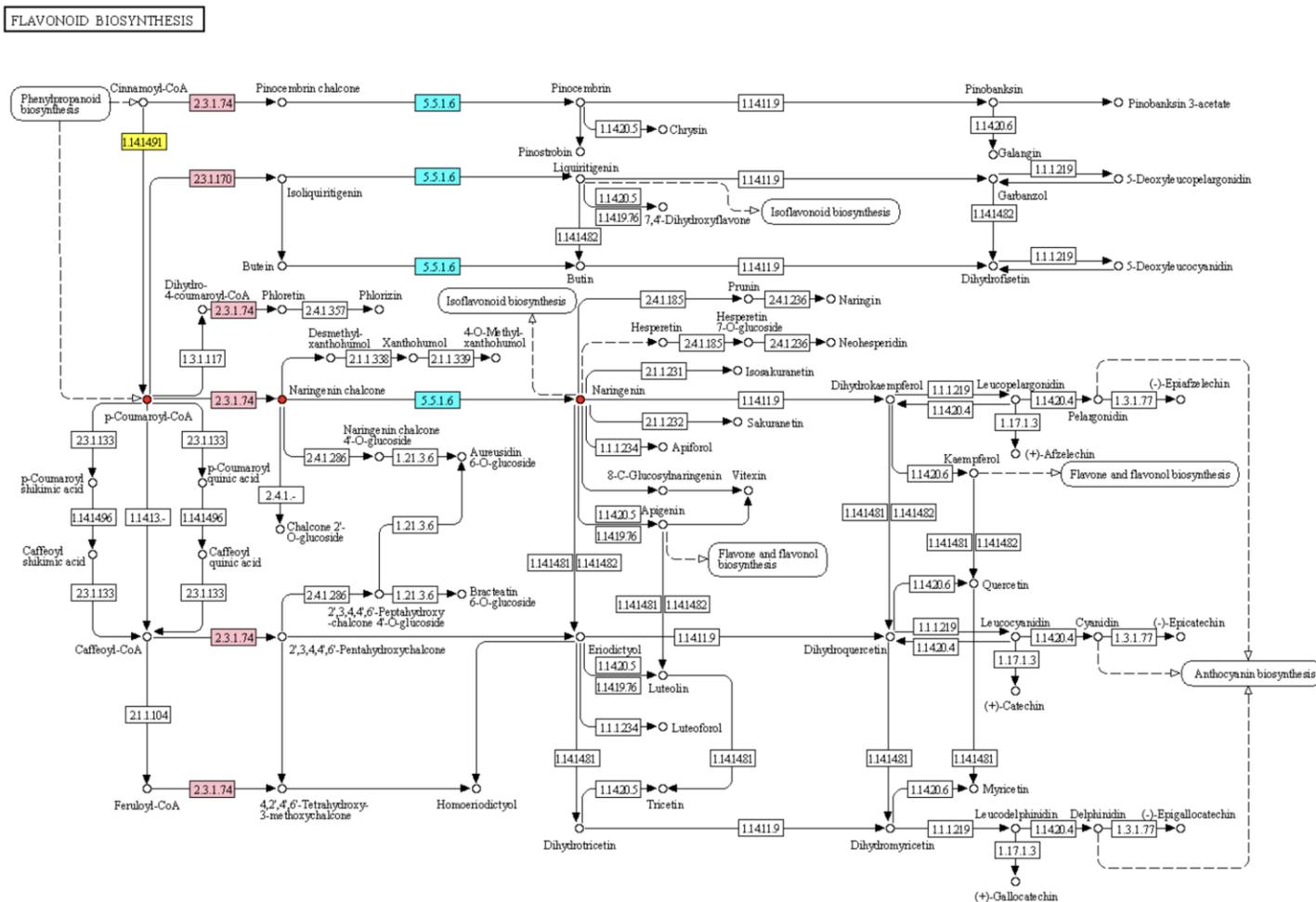
map01060 Biosynthesis of plant secondary metabolites (2)

map00360 Phenylalanine metabolism (2)

map00960 Tropane, piperidine and pyridine alkaloid biosynthesis (2)

map01070 Biosynthesis of plant hormones (2)

map00945 Stilbenoid, diarylheptanoid and gingerol biosynthesis (2)



# 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

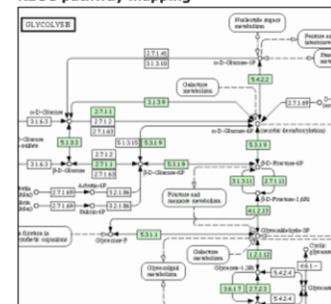
### KO assignment

KAAS KO Assignment Results

Query gene : KO assignment

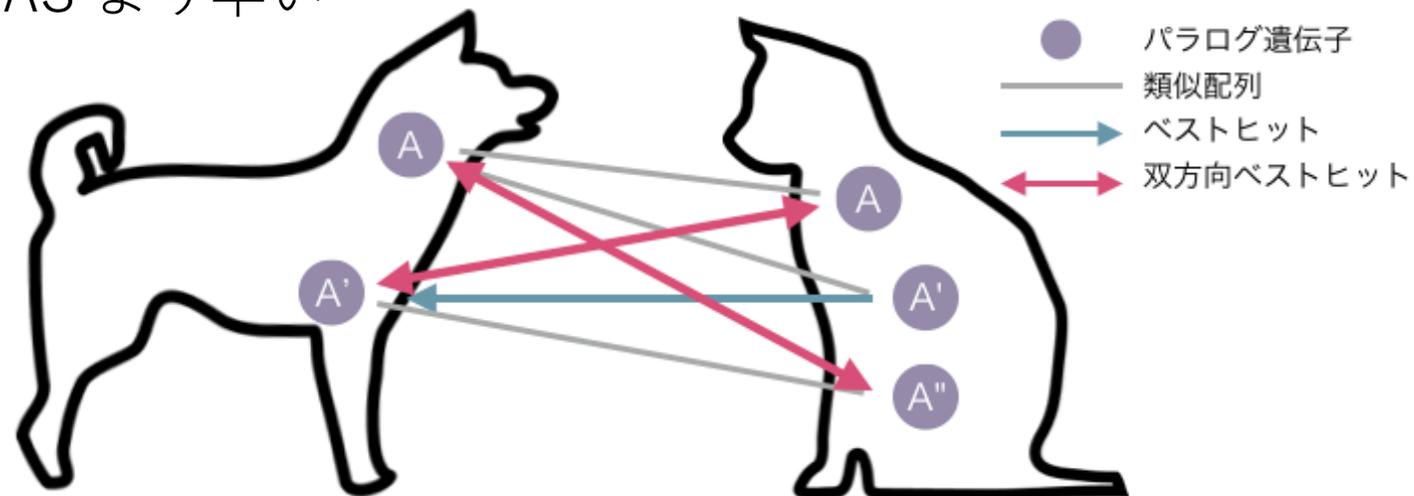
|            |        |
|------------|--------|
| query_0001 | K00003 |
| query_0002 | K00003 |
| query_0003 | K00003 |
| query_0004 | K00003 |
| query_0005 | K00003 |
| query_0006 | K00003 |
| query_0007 | K00003 |
| query_0008 | K00003 |
| query_0009 | K00003 |
| query_0010 | K00003 |
| query_0011 | K00003 |
| query_0012 | K00003 |
| query_0013 | K00003 |
| query_0014 | K00003 |
| query_0015 | K00003 |
| query_0016 | K00003 |

### KEGG pathway mapping



# 配列相同性検索

- [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

- **KEGG Mapper – Assign KO**

([https://www.kegg.jp/kegg/mapper/assign\\_ko.html](https://www.kegg.jp/kegg/mapper/assign_ko.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 – Assign KO

[KEGG2](#) [About](#) [Reconstruct](#) [Search](#) [Color](#) [Join](#) [Convert ID](#) **Assign KO** [Taxonomy](#)

### Assign KO tool

The Assign KO tool is an interface to the [BlastKOALA](#) server in the interactive mode. BlastKOALA assigns KO identifiers (numbers) to a given set of sequence data for subsequent analysis with the [Reconstruct](#) tool. This tool may also be used converting users' sequence IDs to K numbers for use in the Search and Color tools.

### Upload query amino acid sequences in FASTA format

Enter FASTA sequences

```
YLINFFYKTLKIKGTPIQIQFDNENPYVKKNK
>|cl|NZ_ACFK01000001.1_prot_WP_009874555.1_575 [gene=C5S_RS0103170]
[protein=/inference=EXISTENCE: similar to AA sequence:SwissProt:P57663.1]
[protein_id=WP_009874555.1] [location=complement(641210..641791)]
MLNISKKNIIFFILFLLIISLILFNWKYFSLVKNKENLESLKYEKIIKKIKKKSKNLYEVENFIVQ
NTSI
YGTLLTALSIAKKYVECNLNDKALLQLNNSLKYTKEENLKNLLKINIAKIOIKNENNKAMNLL
```

Example:

Query data: [sequence.txt](#)  
Select Buchnera (32199)

|                |       |                            |
|----------------|-------|----------------------------|
| 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-9h          |

- Exec ボタンをクリック
- 数分待つ



**BlastKOALA**

**Job Request**

**Request accepted**

Your job ID is 196cfba9f578d1b20d2aae154e896c678f5d9866

Started at Tue Nov 3 12:54:38 JST 2020

This page reloads automatically every 5 seconds

Your result will be displayed in the following page

[https://www.kegg.jp/kegg-bin/blastkoala\\_result?](https://www.kegg.jp/kegg-bin/blastkoala_result?id=196cfba9f578d1b20d2aae154e896c678f5d9866&passwd=cMFA3A&mode=mapper)

[id=196cfba9f578d1b20d2aae154e896c678f5d9866&passwd=cMFA3A&mode=mapper](https://www.kegg.jp/kegg-bin/blastkoala_result?id=196cfba9f578d1b20d2aae154e896c678f5d9866&passwd=cMFA3A&mode=mapper)

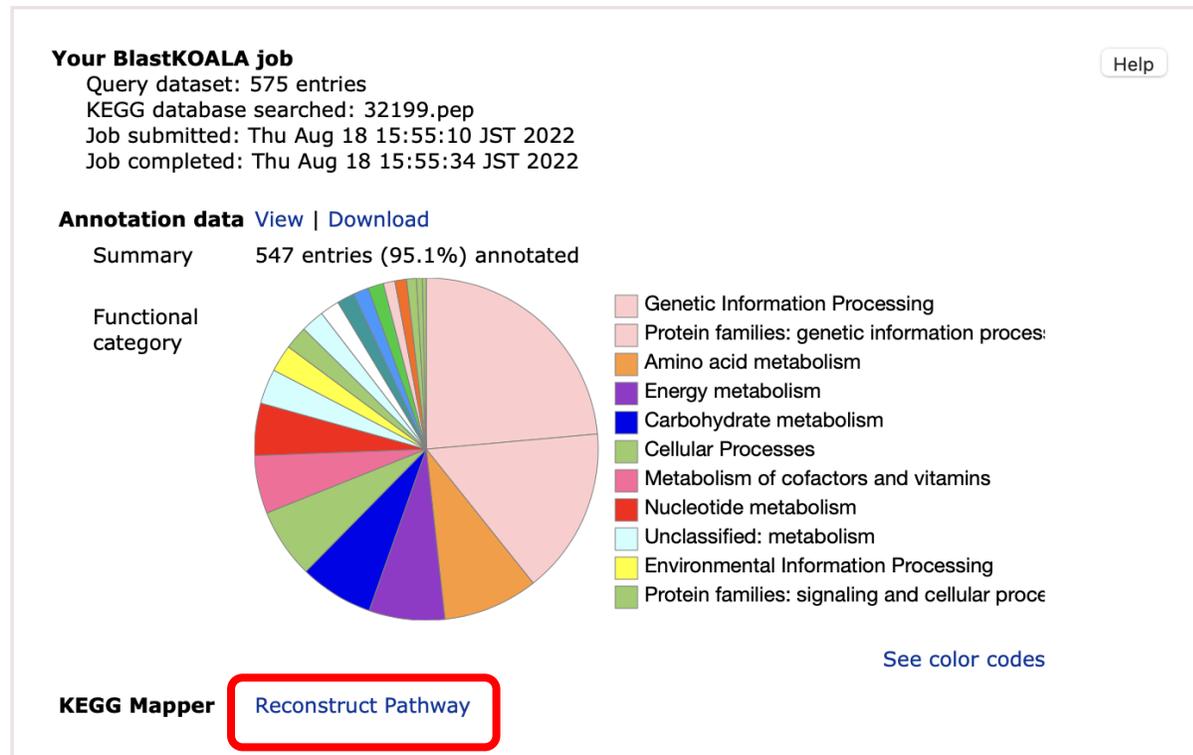
# 計算結果

- 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



**BlastKOALA**

**Result**



# Overviewマップの再構築結果

## KEGG Mapper Reconstruction Result

Pathway (144)   Brite (32)   Brite Table (5)

- Show matched objects
- ### Metabolism
- Global and overview maps
  - 01100 Metabolic pathways (207)**
  - 01110 Biosynthesis of secondary metabolites (105)
  - 01120 Microbial metabolism in diverse environments (52)
  - 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 (14)
  - 00020 Citrate cycle (TCA cycle) (5)
  - 00030 Pentose phosphate pathway (12)
  - 00040 Pentose and glucuronate interconversions (1)
  - 00051 Fructose and mannose metabolism (5)
  - 00052 Galactose metabolism (1)
  - 00500 Starch and sucrose metabolism (2)
  - 00520 Amino sugar and nucleotide sugar metabolism (8)
  - 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 (1)
- ### Nucleotide metabolism
- 00230 Purine metabolism (18)
  - 00240 Pyrimidine metabolism (18)
- ### Amino acid metabolism

KEGG Metabolic pathways - Reference pathway  
[ Pathway menu | Pathway entry | Image (png) file | Help ]

Change pathway type

Option

Scale: 30%  
Link: Normal

Search

Module

Complete only

- Pathway modules
  - Carbohydrate metabolism
    - Central carbohydrate metab
      - M00002 Glycolysis, core m
      - M00307 Pyruvate oxidatio
      - M00004 Pentose phosphat
      - M00006 Pentose phosphat
      - M00007 Pentose phosphat
      - M00005 PRPP biosynthesi
    - Other carbohydrate metabol
  - Energy metabolism
    - Carbon fixation
      - M00579 Phosphate acetyl
    - Methane metabolism
    - Nitrogen metabolism
    - Sulfur metabolism
      - M00176 Assimilatory sulfa
  - Lipid metabolism
    - Fatty acid metabolism
      - M00083 Fatty acid biosynt
    - Sterol biosynthesis
    - Lipid metabolism
  - Nucleotide metabolism
    - Purine metabolism
      - M00049 Adenine ribonucle
    - Pyrimidine metabolism
      - M00051 Uridine monophos
  - Amino acid metabolism
    - Serine and threonine metab
      - M00018 Threonine biosynt
    - Cysteine and methionine me
      - M00021 Cysteine biosynth
    - Branched-chain amino acid i
    - Lysine metabolism
      - M00016 Lysine biosynthes
    - Arginine and proline metabo
      - M00028 Ornithine biosynt
      - M00844 Arginine biosynth
    - Polyamine biosynthesis
    - Histidine metabolism



ご清聴ありがとうございました。