

パスウェイ情報を中心とした KEGG/GenomeNet Webサービスの紹介

統合データベース講習会：AJACS千里

奈良先端科学技術大学院大学

武藤 愛 muto@bs.naist.jp

Table of Contents

- GenomeNet/KEGGの概要説明
- ブラウザを使ったGenomeNet/KEGGの利用法
 - GenomeNet実習
 - bfind, bget, LinkDB
 - KEGG実習（KEGG PATHWAYを中心に）
 - KEGG Mapper, Blast KOALA

GenomeNet

- ゲノム情報を基盤とした新しい生命科学研究と創薬・医療・環境保全への応用を推進するためのインターネットサービス
- KEGGを主とするデータベース群と、それらを利用した解析を行なうための計算ツール群からなる

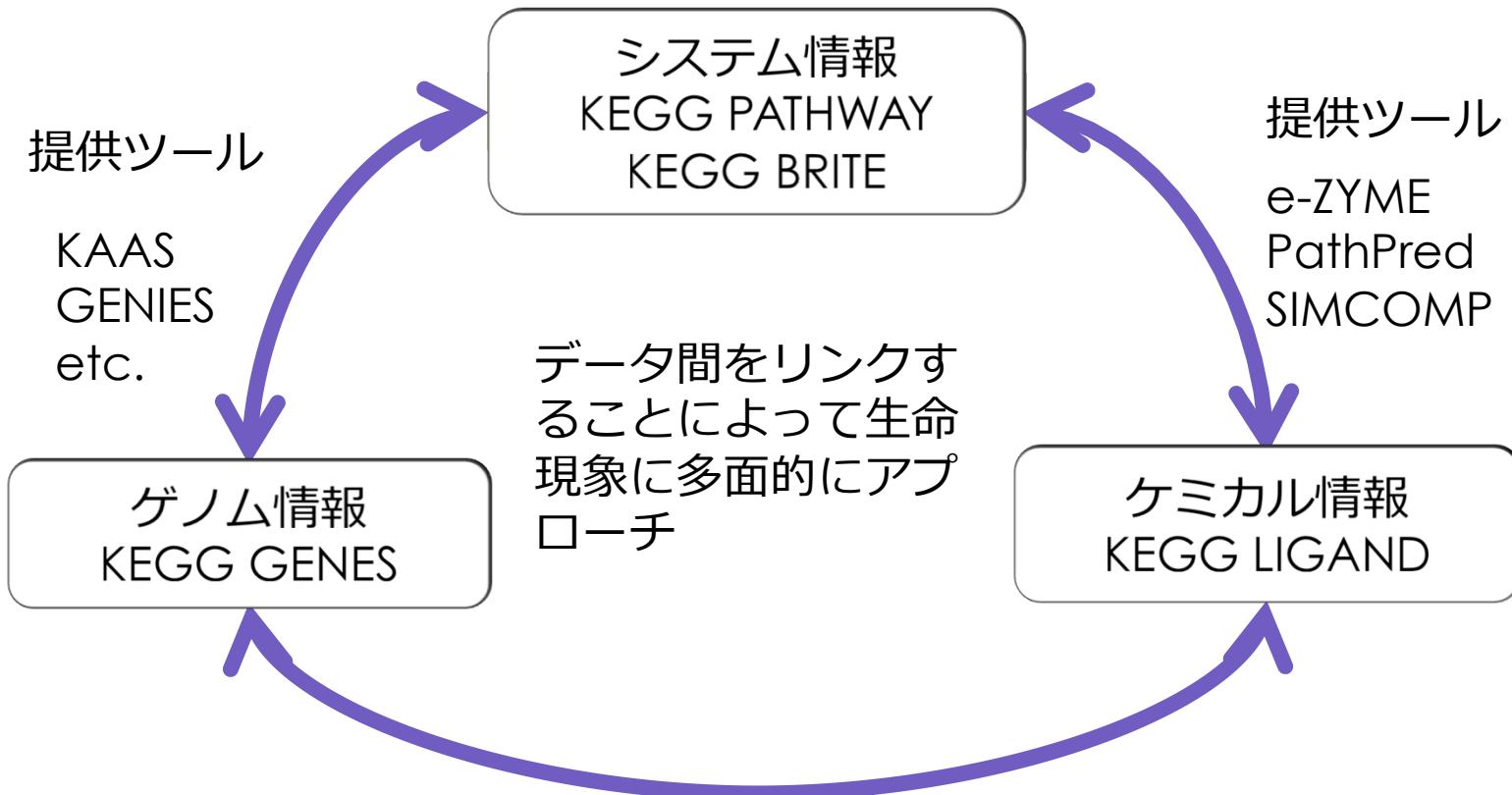
The screenshot shows the homepage of the GenomeNet website. At the top, there is a navigation bar with back and forward arrows, a search icon, and the URL www.genome.jp/ja/. The main title "GenomeNet" is displayed prominently. To the right, the KEGG logo is visible. Below the title, there is a circular logo featuring stylized DNA helixes and the letters T, G, C, A. A search bar is present with the placeholder "Search 統合データベース". A red banner at the bottom of the page reads "タラ海洋探査のメタゲノムデータを MGENES データベースで検索". On the left side of the main content area, there is a sidebar with links to various sections: "ゲノムネット", "統合データベース", "医薬品データベース", "KEGG", "varDB", "研究支援データベース", and "計算ツール". On the right side, there is a larger list of links under the heading "ゲノムネットデータベースリソース": "ゲノムネット統合データベース", "DBGET search", "LinkDB search", "SPARQL エンドポイント", "KEGG: 生命システム情報統合データベース", "KEGG2 - 目次のページ", "KEGG PATHWAY - システム情報: パスウェイ", "KEGG BRITE - システム情報: オントロジ", "KEGG Organisms - 生物種ごとの入口", "KEGG GENES - ゲノム情報", "KEGG LIGAND - ケミカル情報", "KEGG MEDICUS - 疾患・医薬品情報", "KEGG MGENES: メタゲノム情報 New!", "Reaction Ontology: 反応パターンと分類", and "varDB: 抗原変異データベース".

KEGG(Kyoto Encyclopedia of Genes and Genomes)

- 分子レベルの情報から細胞、個体、エコシステムといった高次生命システムレベルの機能や有用性を理解するためのリソース
- 遺伝子や化合物などの生体関連物質のデータベースと、それらをつなぐネットワークのデータベースからなる

The screenshot shows the KEGG homepage with a dark header containing the title. Below the header is a navigation bar with a back arrow, a globe icon, and the URL www.kegg.jp/ja/. The main content area features the KEGG logo (a colorful puzzle piece design with the letters 'KEGG' in yellow) and the text 'Kyoto Encyclopedia of Genes and Genomes'. A search bar with the placeholder 'KEGG' is also visible. The page is divided into several sections: 'KEGG Home' with links to 'Release notes', 'Current statistics', and 'Plea from KEGG'; 'KEGG Database' with links to 'KEGG の概要', 'Searching KEGG', 'KEGG mapping', and 'Color codes'; 'KEGG Objects' with links to 'Pathway maps' and 'Brite hierarchies'; and 'KEGG Software' with links to 'KegTools' and 'KEGG API'. On the right side, there is a large box titled 'KEGG: Kyoto Encyclopedia' which contains a summary of what KEGG is and a link to 'お知らせ' (Announcements). At the bottom, there are two sections: 'KEGG の主要エントリーポイント' (Main entry points for KEGG) listing 'KEGG2' and 'KEGG 全体の目' (General index); and 'データタイプごとのエントリーポイント' (Entry points by data type) listing 'KEGG PATHWAY' and 'KEGG BRITE'.

KEGG



KEGGのシステム情報

- PATHWAY-パスウェイマップ
- BRITE-機能階層・オントロジー
- MODULE-機能ユニット
- Mapper-PATHWAY/BRITE/MODULE
へのマッピングツール
- Atlas-Global mapのviewer



KEGG - Table of Contents

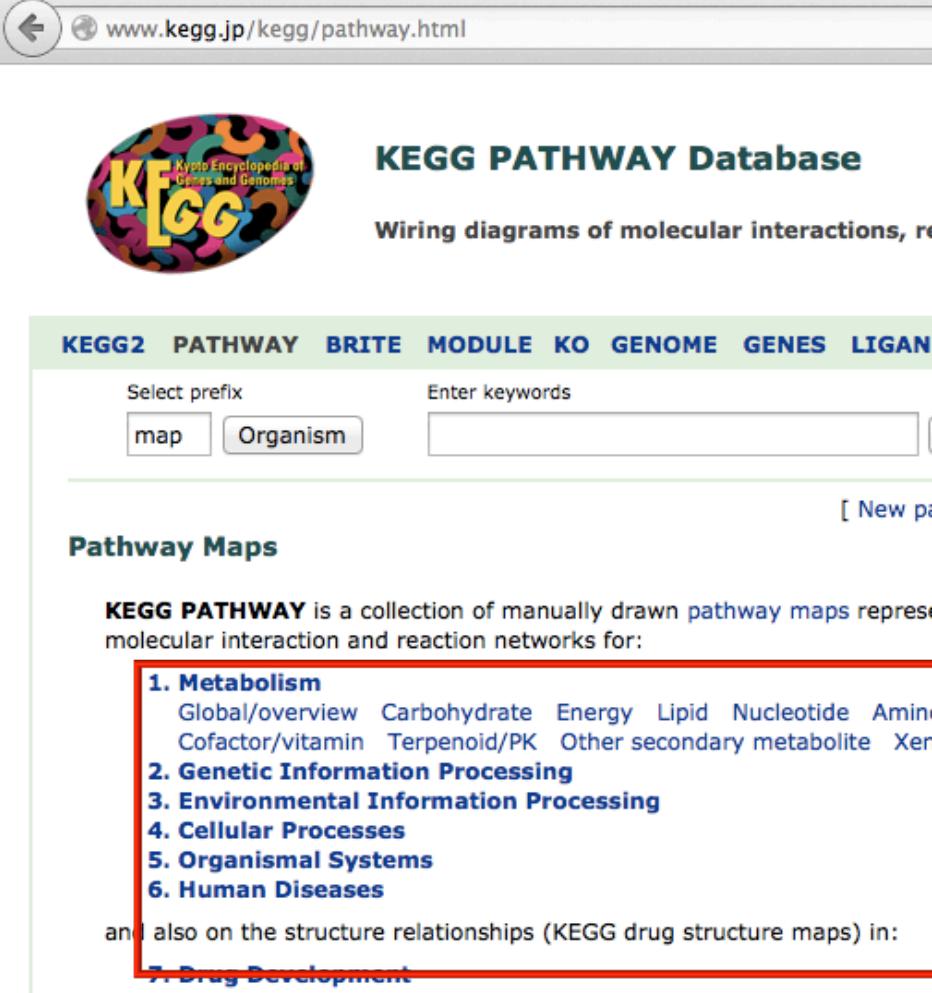
Category	Entry Point	Search & Compare
Systems information	KEGG PATHWAY KEGG BRITE KEGG MODULE KEGG Mapper KEGG Atlas	Search Pathway Search Brite Reconstruct Map Taxonomy
Genomic information	KEGG ORTHOLOGY KEGG Annotation	BlastKOALA GhostKOALA Annotate Sequence KO system
	KEGG GENOME KEGG GENES	SSDB search OC viewer

KEGG PATHWAY

生命システムに関する分子間ネットワーク図のデータベース

- 1. 代謝系
- 2. 遺伝子制御系
- 3. 環境シグナル
- 4. 細胞プロセス
- 5. 生体システム
- 6. 疾患
- 7. 薬剤開発

www.kegg.jp/kegg/pathway.html



The screenshot shows the KEGG PATHWAY Database homepage. At the top right is the URL "www.kegg.jp/kegg/pathway.html". Below it is the KEGG logo, which is a circular emblem with the letters "KEGG" in yellow and red, surrounded by a colorful, abstract pattern. To the right of the logo is the title "KEGG PATHWAY Database" and a subtitle "Wiring diagrams of molecular interactions, reaction networks and biological pathways." A navigation bar at the top includes links for KEGG2, PATHWAY, BRITE, MODULE, KO, GENOME, GENES, and LIGAND. Below the navigation bar are search fields for "Select prefix" (with options "map" and "Organism") and "Enter keywords". A link "[New pathway]" is also present. The main content area is titled "Pathway Maps" and contains a brief description: "KEGG PATHWAY is a collection of manually drawn pathway maps representing molecular interaction and reaction networks for:". A list of pathway categories follows, with the first item, "1. Metabolism", highlighted with a red border. The categories are: 1. Metabolism, 2. Genetic Information Processing, 3. Environmental Information Processing, 4. Cellular Processes, 5. Organismal Systems, 6. Human Diseases. At the bottom, a note states "and also on the structure relationships (KEGG drug structure maps) in: 7. Drug Development".

KEGG PATHWAY



Glycolysis / Gluconeogenesis - Reference pathway

[Pathway menu | Organism menu | Pathway entry | Hide description | User data mapping]

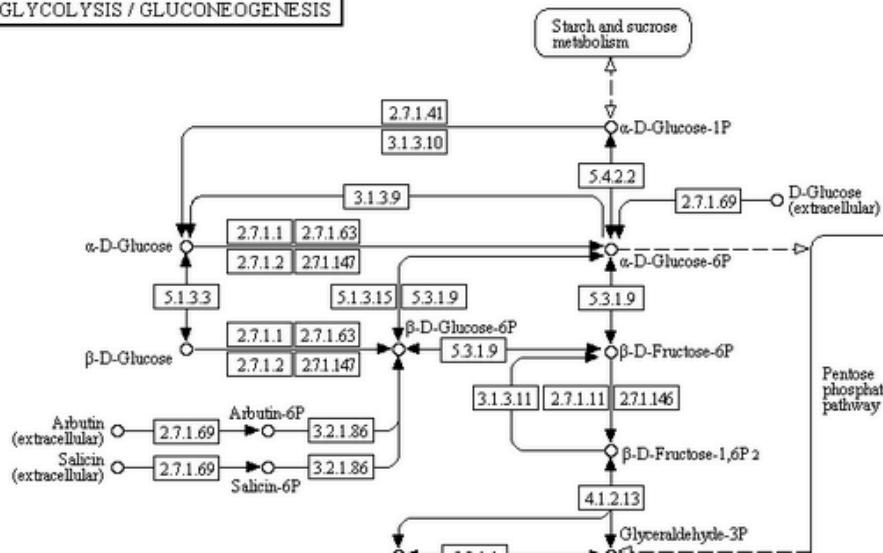
Glycolysis is the process of converting glucose into pyruvate and generating small amounts of ATP (energy) and NADH (reducing power). It is a central pathway that produces important precursor metabolites: six-carbon compounds of glucose-6P and fructose-6P and three-carbon compounds of glyceraldehyde-3P, glycerate-3P, phosphoenolpyruvate, and pyruvate [MD:M00001]. Acetyl-CoA, another important precursor metabolite, is produced by oxidative decarboxylation of pyruvate [MD:M00307]. When the enzyme genes of this pathway are examined in completely sequenced genomes, the reaction steps of three-carbon compounds from glyceraldehyde-3P to pyruvate form a conserved core module [MD:M00002], which is found in almost all organisms and which sometimes contains operon structures in bacterial genomes. Gluconeogenesis is a synthesis pathway of glucose from noncarbohydrate precursors. It is essentially a reversal of glycolysis with minor variations of alternative paths [MD:M00003].

Reference pathway

Go

100%

GLYCOLYSIS / GLUCONEOGENESIS



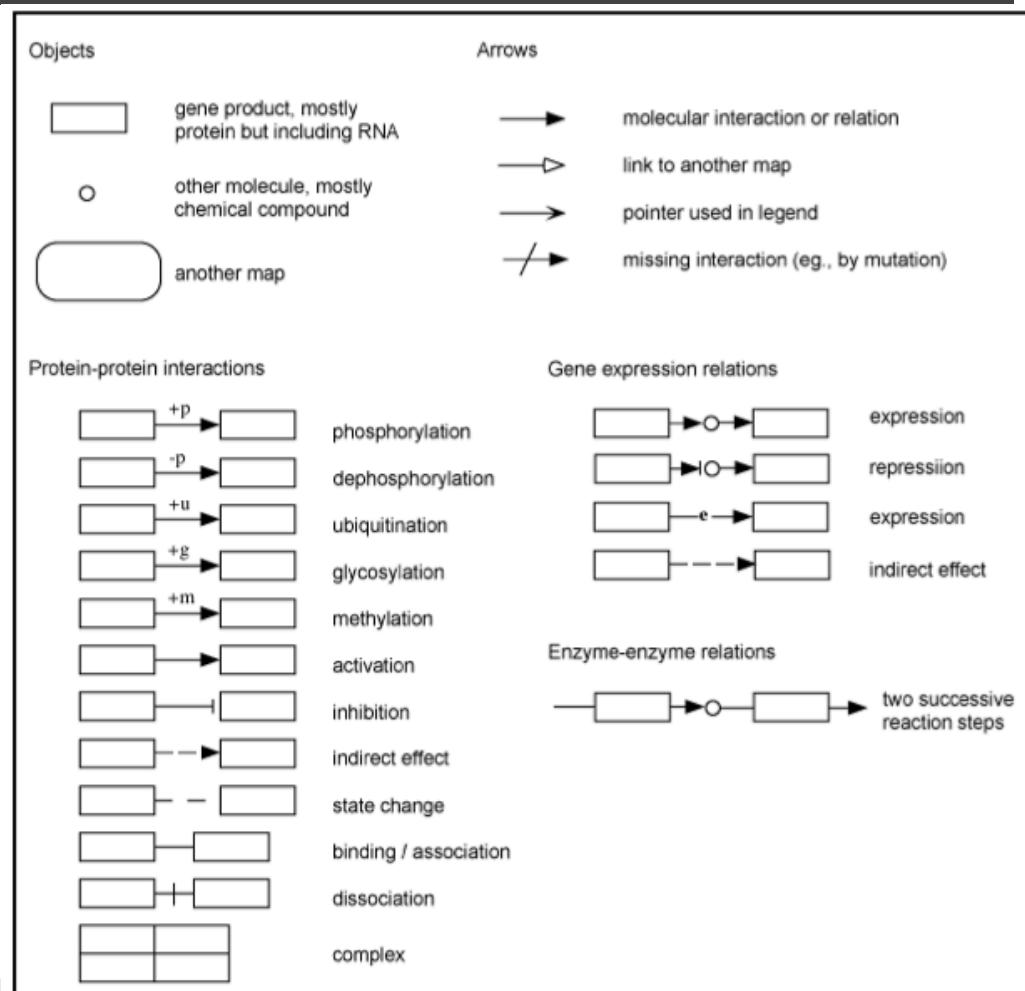
KEGG PATHWAY

□ 箱：遺伝子 丸：化合物

□ 矢印：

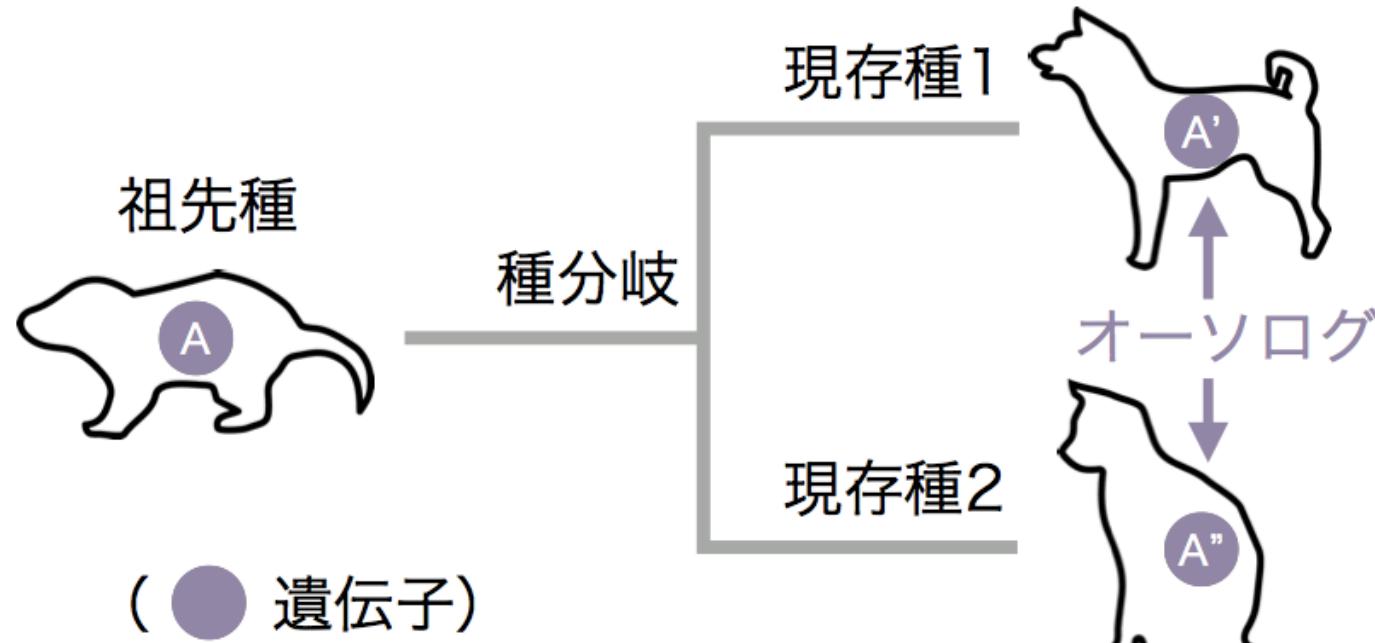
代謝系では代謝反応

それ以外の系ではリン酸化、活性化など様々なタイプがある



KEGG ORTHOLOGY (KO)

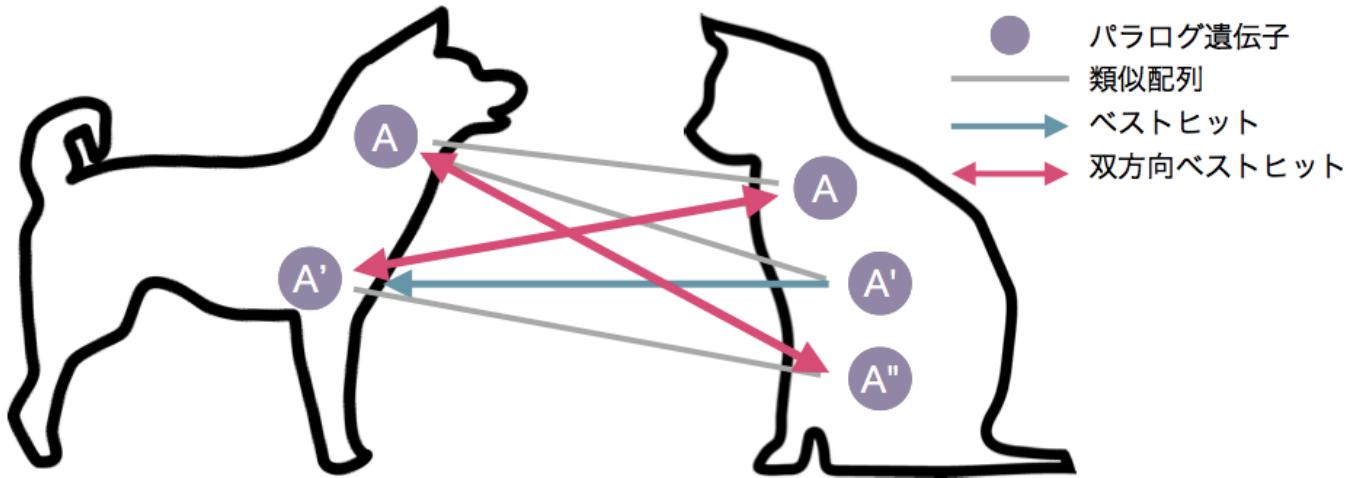
□ オーソログ遺伝子



種分岐によって生じた生物間で対応する遺伝子で、配列相同性を持ち、同一の生物学的機能を持っていると考えられる遺伝子。オーソログ遺伝子を見つけることは機能アノテーションを行う上で重要

KEGG ORTHOLOGY (KO)

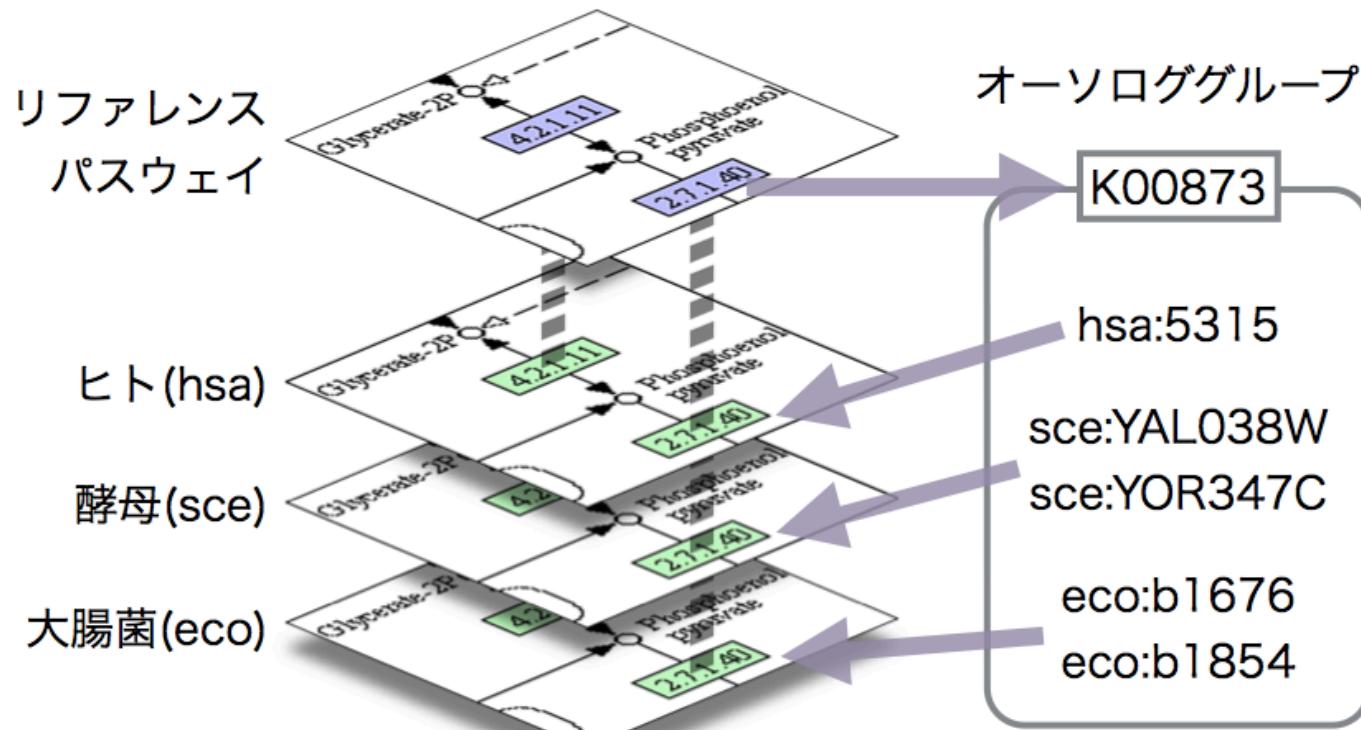
□ KOを介してシステム情報へマッピング



ただ似ているだけでなく、双向性ベストヒットである遺伝子がオーソログの関係にある可能性が高い

KEGG PATHWAY

□ リファレンスパスウェイ



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

まずはゲノムネットにアクセスしましょう

□ <http://www.genome.jp/ja/>

実習 1 :

bget/bfindで検索してみましょう

- bget: GenomeNetのデータベースエントリーID検索
- bfind: GenomeNetのデータベースに対するキーワード検索

ページ上部のテキスト検索ボックスに、好きなキーワードを入れて検索してみる。DB名:エントリーID という形式で入力するとbgetモード（例 eco:b0002） 、それ以外では、 bfindモードで検索が行われる(例 Histidine kinase, Cyanamide等)。

GenomeNet

KEGG KEGG2 PATHWAY BRITE MEDICUS DBGET LinkDB

[English | Japanese]



Search 統合データベース for **cytochrome** Go Clear

タラ海洋探査のメタゲノムデータを MGENES データベースに収録しました。

bfindで"cytochrome"を検索した結果

- キーワードがヒットする全てのエントリが表示される

GenomeNet

Search 統合データベース for cytochrome Go Clear

Database: 統合データベース - Search term: cytochrome

オントロジー

データベース名 KEGG BRITE

エントリー

- ko00001 KO; KEGG Orthology (KO)
- ko00002 Module; KEGG modules
- ko00194 Photosynthesis; Photosynthesis proteins
- ko00199 Cytochrome P450
- ko01000 Enzyme; Enzymes
- ... » display all

GO

0000277 [GO]

The screenshot shows the GenomeNet search interface. In the search bar, 'Search' is set to '統合データベース' and 'for' is set to 'cytochrome'. Below the search bar, the database is specified as '統合データベース' and the search term is 'cytochrome'. A blue header bar says 'オントロジー'. On the left, there's a sidebar with 'データベース名' and 'KEGG BRITE' under it. A yellow bracket labeled 'エントリー' groups several entries: 'ko00001 KO; KEGG Orthology (KO)', 'ko00002 Module; KEGG modules', 'ko00194 Photosynthesis; Photosynthesis proteins', 'ko00199 Cytochrome P450', 'ko01000 Enzyme; Enzymes', and '... » display all'. At the bottom, there's a 'GO' section with a single entry '0000277 [GO]'. The overall background is light blue.

実習2： LinkDBでID対応表を取得しましょう

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

The screenshot shows the LinkDB interface at <http://www.genome.jp/linkdb/>. The page displays a complex network of interconnected databases, primarily centered around KEGG. A red box highlights the KEGG database cluster, which includes Systems (Brite, Pathway, Module), Medicus, Disease, Environ, Drug, Reaction, Compound, Enzyme, Glycan, RPair, and RCClass. The network also connects to various other databases listed on the left and right sides, such as GO, PubMed, BioSystems, UMBBD, TC, COG, CAZy, TAX, SMART, EPD, RefSeq, UniProt, SWISS-PROT, PDBSTR, PDB, InterPro, Ensembl, INSDC, NCBI-GENE, UniGene, HGNC, BioProject, >50 Genome Databases, Nagahama, RIKEN-BRC, SCOP, BRENDA, ExplorEnz, IUBMB, and EXPaSy. On the right, there is a sidebar listing many more databases like OMIM, MeSH, LigandBox, DrugBank, ChEMBL, KNApSack, LipidBank, AAindex, CarbBank, LIPIDMAPS, PubChem, ChEBI, PDB-CCD, 3DDMET, GlycomeDB, JCGGDB, UMBBD-RN, and SABIO-RK.

LinkDB

Link-only Databases

Searchable Databases In DBGET

DBGET Databases

KEGG

Systems

Brite

Pathway

Module

Medicus

Disease

Environ

Drug

Reaction

Compound

Enzyme

Glycan

RPair

RClass

Ligand

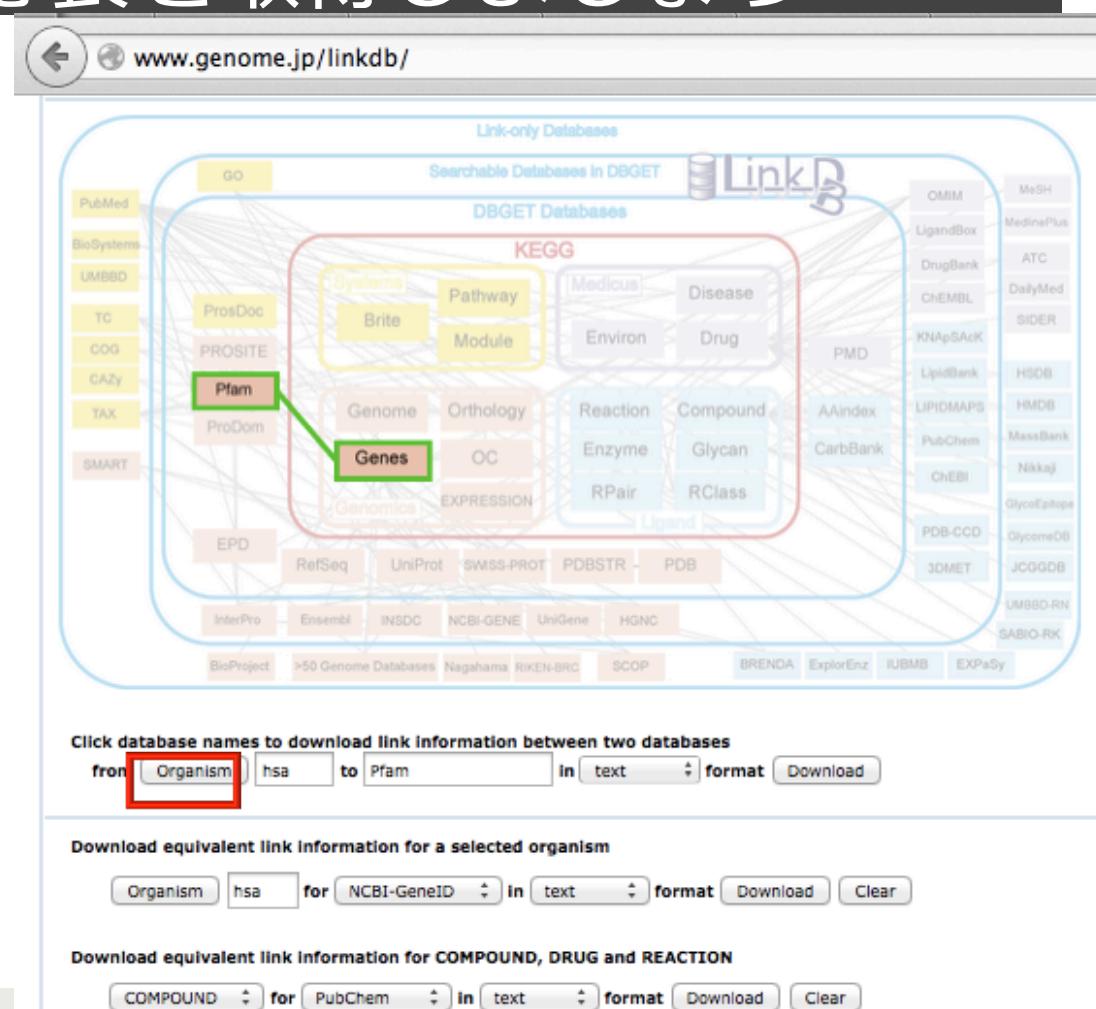
Click database names to download link information between two databases
from to in format

Download equivalent link information for a selected organism
Organism: hsa for NCBI-GeneID in format

Download equivalent link information for COMPOUND, DRUG and REACTION
COMPOUND for PubChem in format

実習2： LinkDBでID対応表を取得しましょう

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



実習2： LinkDBでID対応表を取得しましょう

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

The screenshot shows the LinkDB interface at <http://www.genome.jp/linkdb/>. A search dialog box is open, prompting for a three- or four-letter KEGG organism code. The input field contains "es". A red box highlights the "Select" button. Below the input field, a list of organisms is shown, with "Escherichia coli K-12 MG1655 [eco]" highlighted. Other options include "Escherichia coli K-12 W3110 [ecj]", "Escherichia coli K-12 DH10B [ecd]", "Escherichia coli K-12 MC4100(MuLac) BW2952 [ebw]", "Escherichia coli K-12 MDS42 [ecok]", and "Escherichia coli O157:H7 EDL933 (EHEC) [ece]". The background shows a network diagram of biological databases and a sidebar listing various databases like OMIM, DrugBank, and PDB.

Find three- or four-letter KEGG organism code

es Cancel

Escherichia coli K-12 MG1655 [eco]
Escherichia coli K-12 W3110 [ecj]
Escherichia coli K-12 DH10B [ecd]
Escherichia coli K-12 MC4100(MuLac) BW2952 [ebw]
Escherichia coli K-12 MDS42 [ecok]
Escherichia coli O157:H7 EDL933 (EHEC) [ece]

LinkDB

Disease
Drug
Compound
AAIndex
Glycan
CarbBank
PMD
LipidBank
LIPIDMAPS
PubChem
ChEBI
Nikkaj
GlycoEpitope
PDB-CCD
JCGGDB
UMBBDRN
SABIO-RK

Find organism

www.genome.jp/kegg-bin/find_org_www?mode=

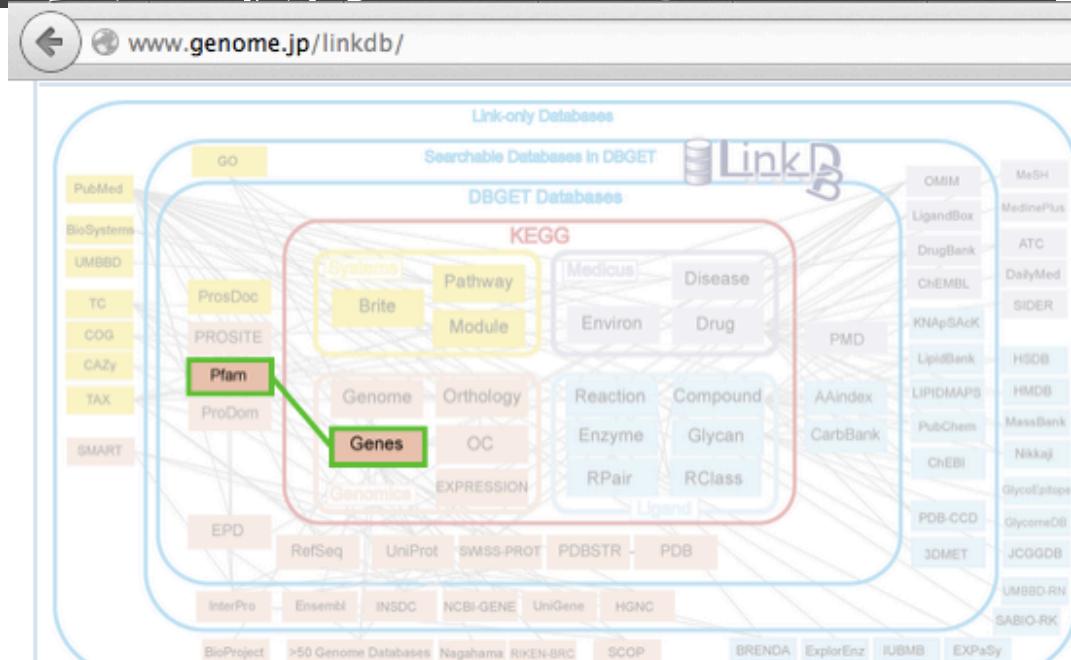
Click database names to download link information between two databases
from to in format

Download equivalent link information for a selected organism
 hsa for in format

Download equivalent link information for COMPOUND, DRUG and REACTION
 for in format

実習2： LinkDBでID対応表を取得しましょう

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



Click database names to download link information between two databases

from eco to in format

Download equivalent link information for a selected organism

eco for in format

Download equivalent link information for COMPOUND, DRUG and REACTION

for in format

実習2： LinkDBでID対応表を取得しましょう

- 遺伝子IDとPfam IDの対応表が取れる

eco:b0001	pf:PF08254	original
eco:b0002	pf:PF00696	original
eco:b0002	pf:PF00742	original
eco:b0002	pf:PF01842	original
eco:b0002	pf:PF03447	original
eco:b0002	pf:PF13840	original
eco:b0003	pf:PF00288	original
eco:b0003	pf:PF08544	original
eco:b0004	pf:PF00291	original
eco:b0004	pf:PF14821	original
eco:b0005	pf:PF10697	original
eco:b0006	pf:PF03883	original
eco:b0007	pf:PF01235	original
eco:b0008	pf:PF00923	original
eco:b0009	pf:PF00994	original
eco:b0010	pf:PF01184	original
eco:b0011	pf:PF03937	original
eco:b0011	pf:PF03981	original
eco:b0011	pf:PF13099	original

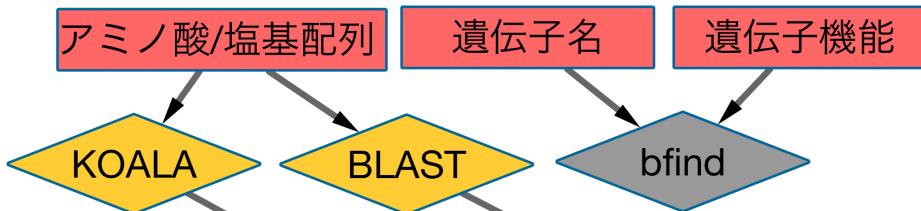
次はKEGGにアクセスしましよう

□<http://www.kegg.jp/ja/>

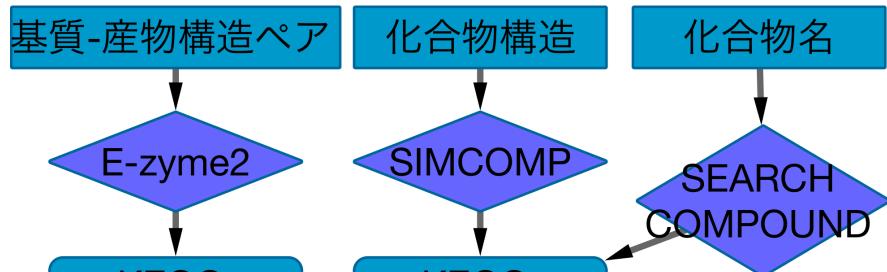
解析の流れ

- あなたの持っているデータはどのデータですか？

遺伝子情報



化合物・反応情報



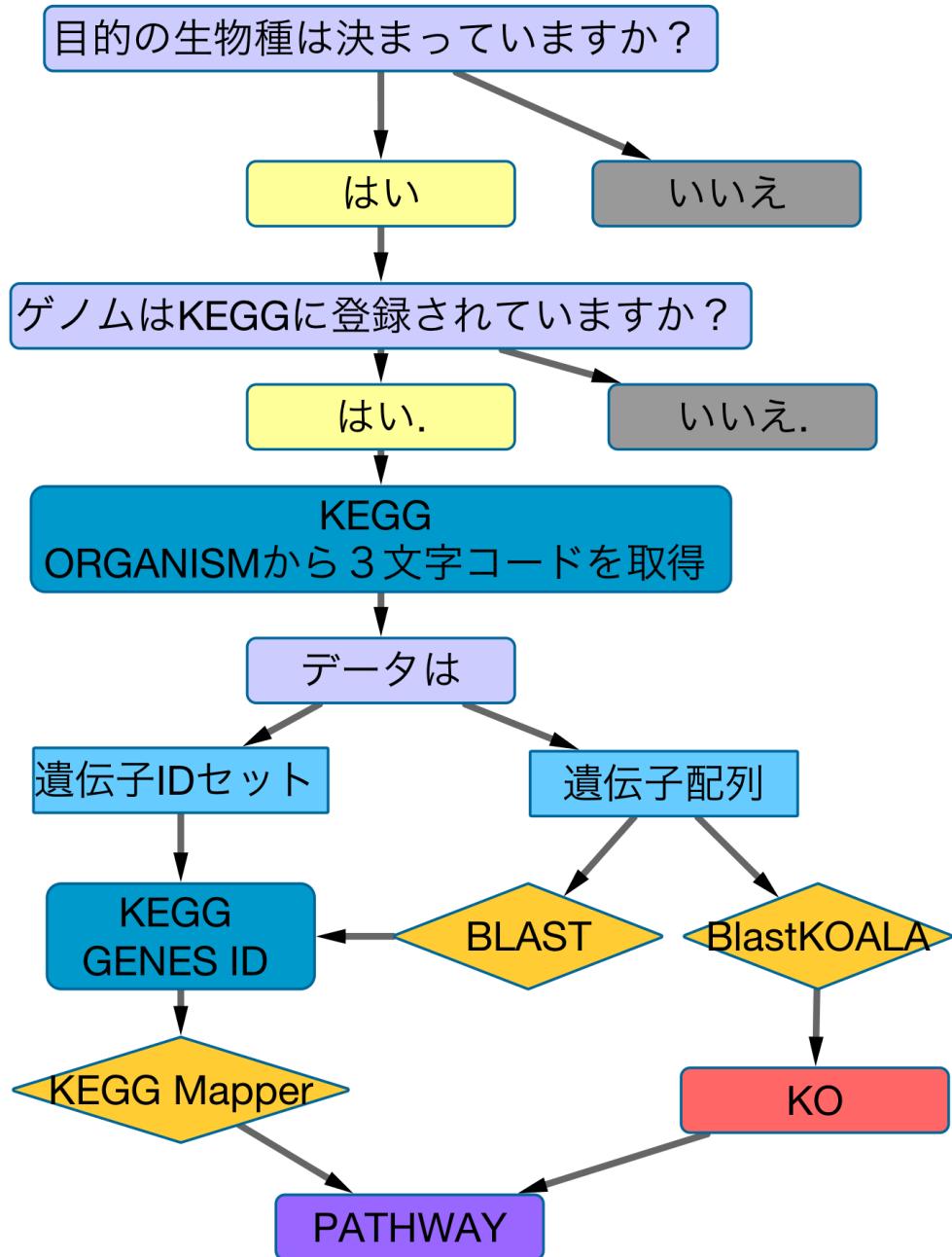
KEGG GENES

KO

PATHWAY

解析の流れ

- ゲノム配列が解読されている生物種の場合



KEGG ORGANISMS

- KEGGにゲノムが登録されている生物種には3文字の生物種コードが与えられている

www.genome.jp/kegg/catalog/org_list.html

検索



KEGG Organisms: Complete Genomes

Eukaryotes: 304 Bacteria: 3393 Archaea: 207

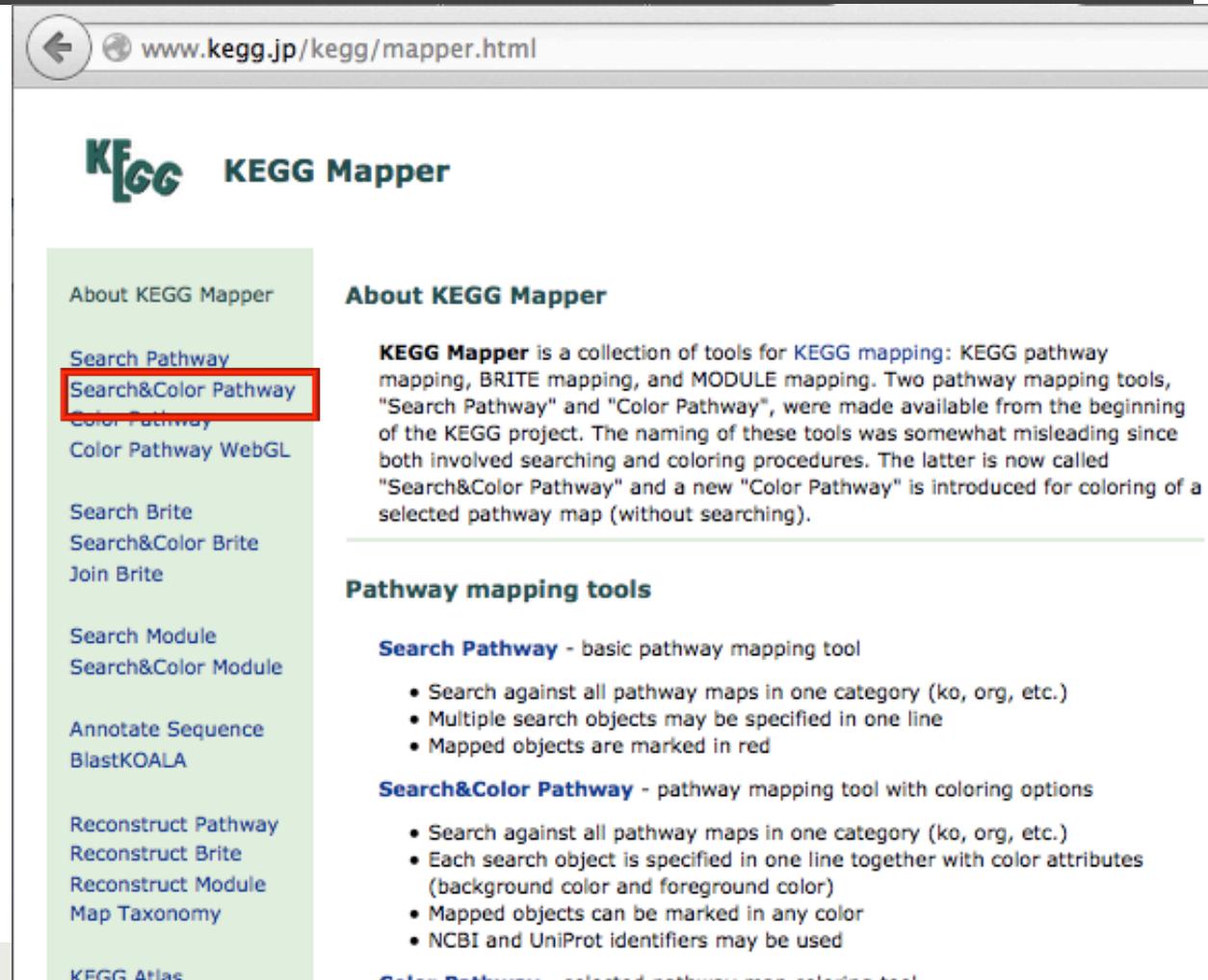
[Genomes | Species | Genus | Draft | Meta]

Eukaryotes

ヒトの3文字コード : hsa

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

実習3：KEGG Mapperを用いた パスウェイ再構築



The screenshot shows the KEGG Mapper homepage. At the top, there is a navigation bar with a back arrow, a globe icon, and the URL www.kegg.jp/kegg/mapper.html. Below the navigation bar, the KEGG logo and the title "KEGG Mapper" are displayed. On the left side, there is a sidebar with various links: "About KEGG Mapper", "Search Pathway", "Search&Color Pathway" (which is highlighted with a red box), "Color Pathway", "Color Pathway WebGL", "Search Brite", "Search&Color Brite", "Join Brite", "Search Module", "Search&Color Module", "Annotate Sequence", "BlastKOALA", "Reconstruct Pathway", "Reconstruct Brite", "Reconstruct Module", "Map Taxonomy", and "KEGG Atlas". The main content area is titled "About KEGG Mapper" and contains a detailed description of the KEGG Mapper tools. It mentions that KEGG Mapper is a collection of tools for KEGG mapping, including pathway, BRITE, and MODULE mapping. It notes that the original "Search Pathway" and "Color Pathway" tools were somewhat misleading as they involved both searching and coloring. The "Search&Color Pathway" tool is described as a new tool introduced for coloring selected pathway maps without searching. Below this, there are two sections: "Pathway mapping tools" and "Search&Color Pathway - pathway mapping tool with coloring options". Each section lists the features of the respective tool.

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

実習
パス[About KEGG Mapper](#)[Search Pathway](#)[Search&Color Pathway](#)[Color Pathway](#)[Color Pathway WebGL](#)[Search Brite](#)[Search&Color Brite](#)[Join Brite](#)[Search Module](#)[Search&Color Module](#)[Annotate Sequence](#)[BlastKOALA](#)[Reconstruct Pathway](#)[Reconstruct Brite](#)[Reconstruct Module](#)[Map Taxonomy](#)[KEGG Atlas](#)[KEGG](#)

生物種 3 文字コード

Search against:

hsa

Enter: map, ko, ec, rn, hsadd, or

org

Primary ID:

KEGG identifiers

IDの種類を選択

(specific pathways only)

Enter objects one per line followed by bgcolor, fgcolor:

7167 red,blue

C00118 pink

Examples:

Homo sapience pathway

遺伝子や化合物のID 塗りつぶし色, 線の色

Alternatively, enter the file name containing the data:

選択...

ファイルが選択されていません。

If necessary, change default bgcolor:

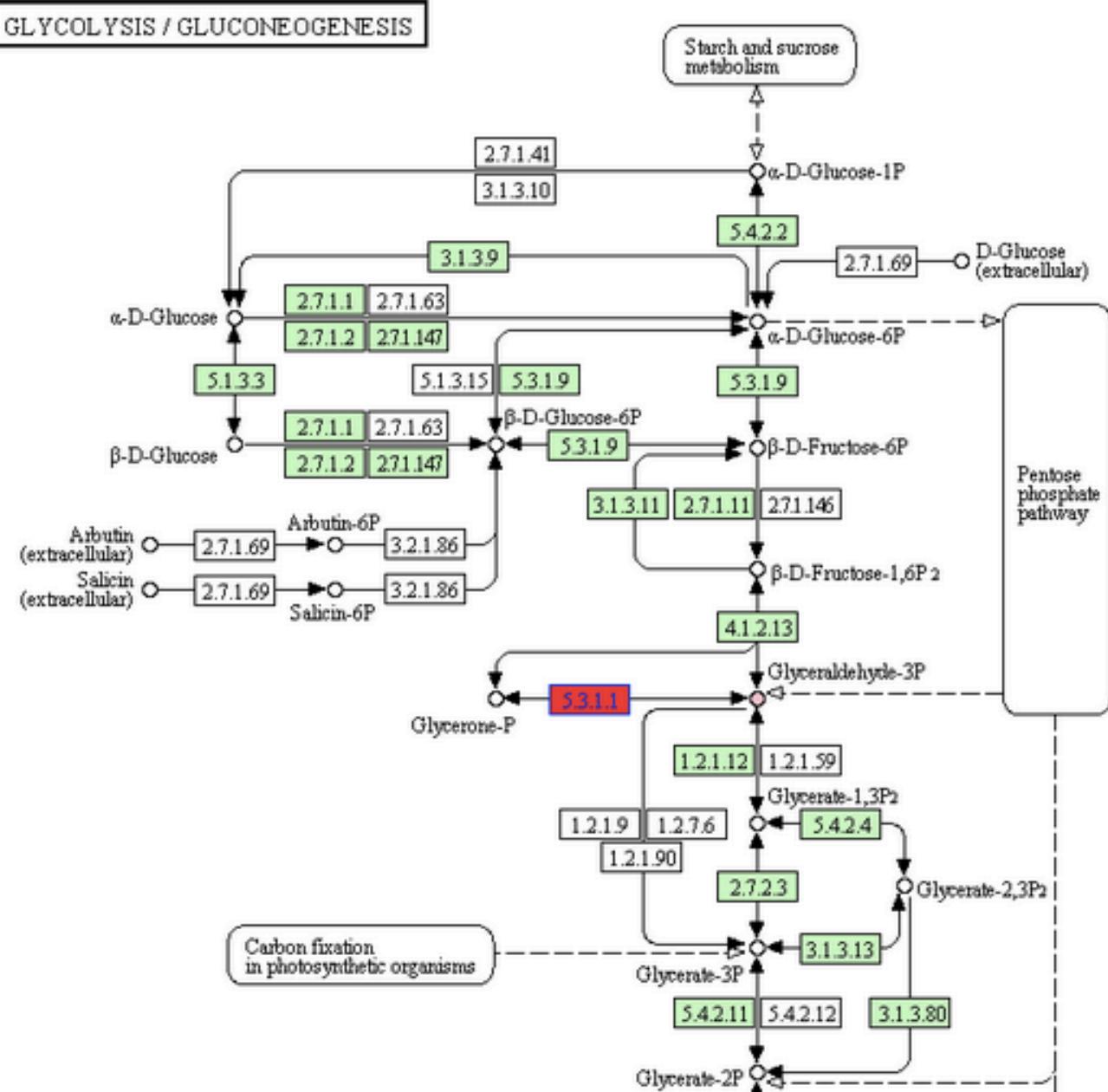
pink

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

実習 パスワード

■ 結果

指定した色でオブジェクトが塗られている



実習 3: KEGG Mapperを用いた パスウェイ再構築

□ 例題

ある研究チームは、大腸菌K-12株の遺伝子「dapA」高発現下で細胞のL-リジン生産量が上がることを発見した。

この結果を説明する代謝経路をKEGG mapperを用いて検索せよ。

自動アノテーションシステム

□ KAAS



input: 大規模シーケンスによって得られた生物種のアミノ酸配列
や塩基配列

output: 配列名とKOの対応表、PATHWAYマップ

□ KOALA



KAASと同様にアミノ酸配列や塩基配列にKOをアサイン

SSEARCHを利用。KEGG内部でKOのアサインにはKOALAを利用

実習4: BlastKOALAを用いた自動アノテーション及pathway再構築

The screenshot shows the KEGG Mapper homepage. At the top, there is a navigation bar with a back arrow, a globe icon, and the URL www.kegg.jp/kegg/mapper.html. Below the navigation bar, the KEGG logo and the text "KEGG Mapper" are displayed. On the left side, there is a sidebar with a green background containing various links: "About KEGG Mapper", "Search Pathway", "Search&Color Pathway", "Color Pathway", "Color Pathway WebGL", "Search Brite", "Search&Color Brite", "Join Brite", "Search Module", "Search&Color Module", "Annotate Sequence" (which is highlighted with a red box), and "BlastKOALA". The main content area has a white background and contains two sections: "About KEGG Mapper" and "Pathway mapping tools". The "About KEGG Mapper" section describes the tool as a collection of KEGG mapping tools. The "Pathway mapping tools" section is divided into three parts: "Search Pathway" (basic pathway mapping tool), "Search&Color Pathway" (pathway mapping tool with coloring options), and "Color Pathway" (selected pathway map coloring tool). Each part lists its features.

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



KEGG Mapper – Annotate Sequence by BlastKOALA

About KEGG Mapper

Search Pathway

Search&Color Pathway

Color Pathway

Color Pathway WebGL

Search Brite

Search&Color Brite

Join Brite

Search Module

Search&Color Module

Annotate Sequence

BlastKOALA

Reconstruct Pathway

Reconstruct Brite

Reconstruct Module

Map Taxonomy

KEGG Atlas

KEGG

Upload query amino acid sequences in FASTA format

Enter FASTA sequences

Or upload file: ファイルが選択されていません。

Select GENES family/genus dataset to be searched

Select

Family/Genus

Example:

Query data: **sequence.txt**

Select Buchnera (32199)

リンク先を全選択してコピペ

Annotate Sequence is an interface to the BlastKOALA server for assigning KO identifiers (K numbers) to a given sequence data set for subsequent analyses with Reconstruct Pathway, Reconstruct Brite, and Reconstruct Module tools. The BlastKOALA computation is performed in an interactive mode using an appropriate subset of KEGG GENES (corresponding to family/genus of your organism). **This tool is useful for converting your sequence IDs to K numbers.**

□ 今回(は)



KEGG Mapper – Annotate Sequence by BlastKOALA

[About KEGG Mapper](#)[Search Pathway](#)[Search&Color Pathway](#)[Color Pathway](#)[Color Pathway WebGL](#)[Search Brite](#)[Search&Color Brite](#)[Join Brite](#)[Search Module](#)[Search&Color Module](#)[Annotate Sequence](#)[BlastKOALA](#)[Reconstruct Pathway](#)[Reconstruct Brite](#)[Reconstruct Module](#)[Map Taxonomy](#)[KEGG Atlas](#)[KEGG](#)

Upload query amino acid sequences in FASTA format

Enter FASTA sequences

```
>|cl|NZ_ACFK01000001.1_prot_WP_009873963.1_1 [gene=C5S_RS0100010]
[protein=/inference=EXISTENCE: similar to AA sequence:SwissProt:P57117.1]
[protein_id=WP_009873963.1] [location=202..2088]
MFNLRNFDVIVVGAGHAGTEAMASSRMGCKTLLTQKISDLGALSCNPAIGGGIGKSHLVKE
IDALGGMM
AKAIDYSGIQFRILNSSKGPAVRSTRAQADKILYHETVKKILKKONNLLILEAEVKDLIFKNYS
VVGVLT
QNEINFYSRSVWLAAGTFLGGKIHIGLKSYSAGRIGDKSAIDLSVRLRELSLRVNRLKTGTPP
RIDINTV
NFNNLIIQNSDTPVPVFSFMGNVSHHPKQIPCYLTHTNEKTHEIIRKNLDKSPIYTGFLKGLG
```

Or upload file: ファイルが選択されていません。

Select GENES family/genus dataset to be searched

Select

Family/Genus

Example:

Query data: sequence.txt

Select Buchnera (32199)

Annotate Sequence is an interface to the BlastKOALA server for assigning KO identifiers (K numbers) to a given sequence data set for subsequent analyses with Reconstruct Pathway, Reconstruct Brite, and Reconstruct Module tools. The BlastKOALA computation is performed in an interactive mode using an appropriate subset of KEGG GENES (corresponding to family/genus of your organism). This tool is useful for converting your sequence IDs to K numbers.

				<i>Yersinia rohdei</i>	29485
				<i>Yersinia ruckeri</i>	29486
	<i>Shigella</i>	620		<i>Shigella flexneri</i>	623
				<i>Shigella sonnei</i>	624
				<i>Shigella boydii</i>	621
				<i>Shigella dysenteriae</i>	622
	<i>Pectobacterium</i>	122277		<i>Pectobacterium atrosepticum</i>	29471
				<i>Pectobacterium carotovorum</i>	554
				<i>Pectobacterium wasabiae</i>	55208
				<i>Pectobacterium</i> sp. SCC3193	1166016
	<i>Erwinia</i>	551		<i>Erwinia tasmaniensis</i>	338565
				<i>Erwinia pyrifoliae</i>	79967
				<i>Erwinia amylovora</i>	552
				<i>Erwinia billingiae</i>	182337
				<i>Erwinia</i> sp. Ejp617	215689
	<i>Photorhabdus</i>	29487		<i>Photorhabdus luminescens</i>	29488
				<i>Photorhabdus asymbiotica</i>	291112
	<i>Buchnera</i>	32199		<i>Buchnera aphidicola</i>	9
	<i>Wigglesworthia</i>	51228		<i>Wigglesworthia glossinidia</i>	51229
	<i>Sodalis</i>	84565		<i>Sodalis glossinidius</i>	63612
				<i>Sodalis</i> sp. HS1	1239307
				<i>Candidatus Sodalis pierantonius</i>	1486991
				<i>Enterobacter</i> sp. 638	399742
	<i>Enterobacter</i>	547		<i>Enterobacter cloacae</i>	550
				<i>Enterobacter lionolyticus</i>	1334193
				<i>Enterobacter</i>	
				<i>Enterobacter</i>	
				<i>Enterobacter</i>	
	<i>Cronobacter</i>	413496		<i>Cronobacter sakazakii</i>	28141
	<i>Siccibacter</i>	1649298		<i>Siccibacter turicensis</i>	413502
	<i>Klebsiella</i>	570		<i>Klebsiella pneumoniae</i>	573
				<i>Klebsiella varicola</i>	244366
				<i>Klebsiella oxytoca</i>	571
				<i>Klebsiella michiganensis</i>	1134687
				<i>Citrobacter koseri</i>	545

Buchneraで検索して数字を
クリック

実習4: 自動アノ

- 少し待っている
と画面が自動で
切り替わる

www.kegg.jp/kegg-bin/blastkoala_result?id=63ad1534d877b373e512a1556



Annotate Sequence Result

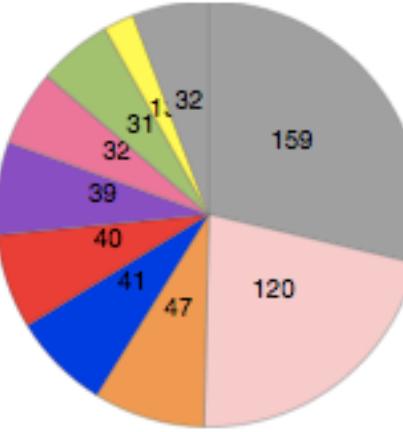
Your Annotate Sequence job

Query dataset: 575 entries
KEGG database searched: 32199.pep
Job submitted: Tue Jun 16 08:00:44 JST 2015
Job completed: Tue Jun 16 08:01:26 JST 2015

Annotation data [View](#) | [Download](#)

Summary 551 entries (95.8%) annotated

Functional category



Functional Category	Count
Unclassified	159
Genetic Information Processing	120
Amino acid metabolism	47
Carbohydrate metabolism	41
Nucleotide metabolism	40
Energy metabolism	39
Metabolism of cofactors and vitamins	32
Cellular Processes	32
Environmental Information Processing	1

KEGG Mapper

[Reconstruct Pathway](#)
[Reconstruct Brite](#)
[Reconstruct Module](#)

[[Annotate Sequence](#) | [KEGG](#) | [Kanehisa Labs](#)]

Annotation query IDs

Assigned KOs

Close

Icl NZ_ACFK01000001.1_prot_WP_009873963.1_1 (628)	K03495	gidA; tRNA uridine 5-carboxymethylaminomethyl modification enzyme
Icl NZ_ACFK01000001.1_prot_WP_009873964.1_2 (274)	K02108	ATPFOA; F-type H+-transporting ATPase subunit a
Icl NZ_ACFK01000001.1_prot_WP_009873965.1_3 (79)	K02110	ATPFOC; F-type H+-transporting ATPase subunit c
Icl NZ_ACFK01000001.1_prot_WP_009873966.1_4 (161)	K02109	ATPFOB; F-type H+-transporting ATPase subunit b
Icl NZ_ACFK01000001.1_prot_WP_009873967.1_5 (177)	K02113	ATPF1D; F-type H+-transporting ATPase subunit delta
Icl NZ_ACFK01000001.1_prot_WP_009873968.1_6 (512)	K02111	ATPF1A; F-type H+-transporting ATPase subunit alpha [EC:3.6.3.14]
Icl NZ_ACFK01000001.1_prot_WP_009873969.1_7 (290)	K02115	ATPF1G; F-type H+-transporting ATPase subunit gamma
Icl NZ_ACFK01000001.1_prot_WP_009873970.1_8 (465)	K02112	ATPF1B; F-type H+-transporting ATPase subunit beta [EC:3.6.3.14]
Icl NZ_ACFK01000001.1_prot_WP_029585514.1_9 (140)	K02114	ATPF1E; F-type H+-transporting ATPase subunit epsilon
Icl NZ_ACFK01000001.1_prot_WP_009873972.1_10 (803)	K02470	gyrB; DNA gyrase subunit B [EC:5.99.1.3]
Icl NZ_ACFK01000001.1_prot_WP_009873973.1_11 (366)	K02338	DPO3B; DNA polymerase III subunit beta [EC:2.7.7.7]
Icl NZ_ACFK01000001.1_prot_WP_009873974.1_12 (454)	K02313	dnaA; chromosomal replication initiator protein
Icl NZ_ACFK01000001.1_prot_WP_009873975.1_13 (47)	K02914	RP-L34; large subunit ribosomal protein L34
Icl NZ_ACFK01000001.1_prot_WP_009873977.1_14 (532)	K03217	yidC; YidC/Oxa1 family membrane protein insertase
Icl NZ_ACFK01000001.1_prot_WP_009873978.1_15 (453)	K03650	mnmE; tRNA modification GTPase [EC:3.6.-.-]
Icl NZ_ACFK01000001.1_prot_WP_009873979.1_16 (96)	K04078	groES; chaperonin GroES
Icl NZ_ACFK01000001.1_prot_WP_009873980.1_17 (548)	K04077	groEL; chaperonin GroEL
Icl NZ_ACFK01000001.1_prot_WP_014498809.1_18 (189)	K02356	efp; elongation factor P
Icl NZ_ACFK01000001.1_prot_WP_009873982.1_19 (246)	K02315	dnaC; DNA replication protein DnaC
Icl NZ_ACFK01000001.1_prot_WP_009873983.1_20 (164)	K02317	dnaT; DNA replication protein DnaT
Icl NZ_ACFK01000001.1_prot_WP_009873984.1_21 (191)	K08316	rsmD; 16S rRNA (guanine966-N2)-methyltransferase [EC:2.1.1.171]
Icl NZ_ACFK01000001.1_prot_WP_009873985.1_22 (378)	K03110	ftsY; fused signal recognition particle receptor
Icl NZ_ACFK01000001.1_prot_WP_009873986.1_23 (284)	K03089	SIG3.3.1; RNA polymerase sigma-32 factor
Icl NZ_ACFK01000001.1_prot_WP_009873987.1_24 (609)	K00820	glmS; glucosamine--fructose-6-phosphate aminotransferase (isomerizing) [EC:2.6.1.1]
Icl NZ_ACFK01000001.1_prot_WP_009873988.1_25 (459)	K04042	glmU; bifunctional UDP-N-acetylglucosamine pyrophosphorylase / Glucosamine-1-pho
Icl NZ_ACFK01000001.1_prot_WP_009873989.1_26 (272)	K07024	
Icl NZ_ACFK01000001.1_prot_WP_009873990.1_27 (758)	K00549	metE; 5-methyltetrahydropteroylglutamate--homocysteine methyltransferase [EC:2.1.1.1]
Icl NZ_ACFK01000001.1_prot_WP_009873991.1_28 (525)	K00602	purH; phosphoribosylaminoimidazolecarboxamide formyltransferase / IMP cyclohydrol
Icl NZ_ACFK01000001.1_prot_WP_009873992.1_29 (92)	K05787	hupA; DNA-binding protein HU-alpha
Icl NZ_ACFK01000001.1_prot_WP_009873993.1_30 (81)		
Icl NZ_ACFK01000001.1_prot_WP_009873994.1_31 (1407)	K03046	rpoC; DNA-directed RNA polymerase subunit beta' [EC:2.7.7.6]
Icl NZ_ACFK01000001.1_prot_WP_009873995.1_32 (1342)	K03043	rpoB; DNA-directed RNA polymerase subunit beta [EC:2.7.7.6]
Icl NZ_ACFK01000001.1_prot_WP_009873996.1_33 (122)	K02935	RP-L7; large subunit ribosomal protein L7/L12
Icl NZ_ACFK01000001.1_prot_WP_009873997.1_34 (165)	K02864	RP-L10; large subunit ribosomal protein L10
Icl NZ_ACFK01000001.1_prot_WP_009873998.1_35 (231)	K02863	RP-L1; large subunit ribosomal protein L1

実習4: 自動アノ

- Reconstruct Pathway をクリック

www.kegg.jp/kegg-bin/blastkoala_result?id=63ad1534d877b373e512a1556

Annotate Sequence Result



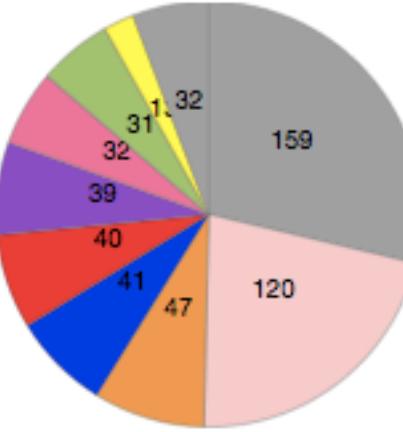
Your Annotate Sequence job

Query dataset: 575 entries
KEGG database searched: 32199.pep
Job submitted: Tue Jun 16 08:00:44 JST 2015
Job completed: Tue Jun 16 08:01:26 JST 2015

Annotation data [View](#) | [Download](#)

Summary 551 entries (95.8%) annotated

Functional category



Functional Category	Count
Unclassified	159
Genetic Information Processing	120
Amino acid metabolism	47
Carbohydrate metabolism	41
Nucleotide metabolism	40
Energy metabolism	39
Metabolism of cofactors and vitamins	32
Cellular Processes	32
Environmental Information Processing	1

KEGG Mapper [Reconstruct Pathway](#) Reconstruct Brite [Reconstruct Module](#)

[\[Annotate Sequence \]](#) [KEGG](#) [Kanehisa Labs](#)

実習4: Blast 自動アノテー

- AssignされたKOが載っているpathway一覧

www.kegg.jp/kegg-bin/find_pathway_object?query_pathwayfile=63ad1534d877

Pathway Reconstruction Result

Show all objects

Metabolism

- Global and overview maps
 - 01100 Metabolic pathways (206)
 - 01110 Biosynthesis of secondary metabolites (98)
 - 01120 Microbial metabolism in diverse environments (48)
 - 01130 Biosynthesis of antibiotics (69)
 - 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 (2)
 - 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
 - 00230 Purine metabolism (29)
 - 00240 Pyrimidine metabolism (30)

実習4: BlastKOALA 自動アノテーション

- AssignされたKOに相当する遺伝子産物のノードが色付けされている

