



KEGG PATHWAYを中心とした KEGGおよびGenomeNetの使い方

京都大学化学研究所
附属バイオインフォマティクスセンター
化学生命科学領域

時松 敏明

2013/5/28

統合データベース講習会 AJACS筑波3 2013/5/28

内容

- KEGGとゲノムネットの概要
- KEGG, ゲノムネットの配列データ
(GENES, ORTHOLOGY(KO), OC)
- KEGG, ゲノムネットの化合物情報
(LIGAND)
- KEGGのシステム情報、ツール
(MEDICUS, PATHWAY, BRITE, MODULE, KEGG Mapper)
- そのほかのKEGGのツール群

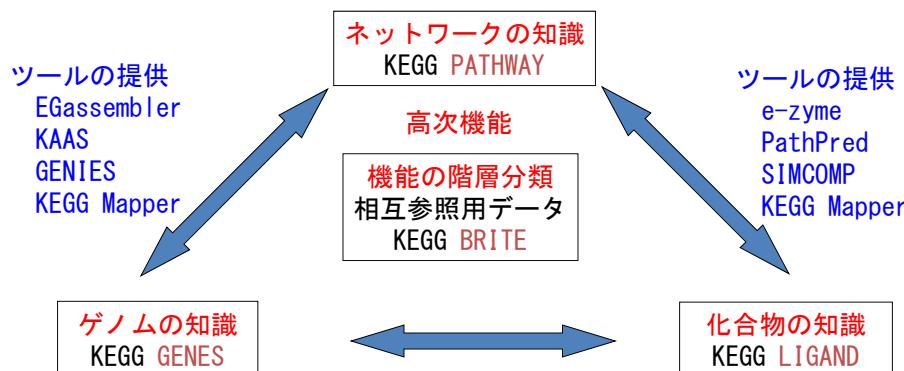
2013/5/28

KEGGとゲノムネットの概要

2013/5/28

KEGG: Kyoto Encyclopedia of Genes and Genomes

様々な種類のデータを「生命現象の総体」として再構築



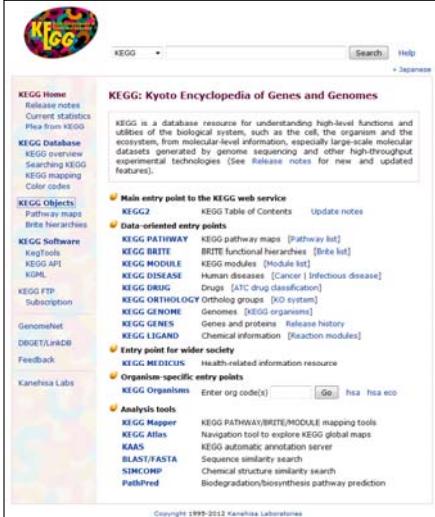
研究者の知識をゲノムレベルのデータと結びつける

2013/5/28

KEGG でカバーしているデータ、いないデータ

<ul style="list-style-type: none"> • いる <p>ゲノム</p> <table style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 50%; vertical-align: top; padding: 5px;"> 遺伝子 アノテーション </td> <td style="width: 50%; vertical-align: top; padding: 5px;"> 制御領域 バリエーション タンパク質立体構造 </td> </tr> </table> <p>パスウェイ</p> <table style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 50%; vertical-align: top; padding: 5px;"> 代謝、制御マップ 病気、薬開発マップ </td> <td style="width: 50%; vertical-align: top; padding: 5px;"> シミュレーション用の パラメータ </td> </tr> </table> <p>化合物</p> <table style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 50%; vertical-align: top; padding: 5px;"> 2次元構造 薬、代謝物、反応 反応パターン </td> <td style="width: 50%; vertical-align: top; padding: 5px;"> 物性、立体構造 速度定数 </td> </tr> </table>	遺伝子 アノテーション	制御領域 バリエーション タンパク質立体構造	代謝、制御マップ 病気、薬開発マップ	シミュレーション用の パラメータ	2次元構造 薬、代謝物、反応 反応パターン	物性、立体構造 速度定数	<ul style="list-style-type: none"> • いない
遺伝子 アノテーション	制御領域 バリエーション タンパク質立体構造						
代謝、制御マップ 病気、薬開発マップ	シミュレーション用の パラメータ						
2次元構造 薬、代謝物、反応 反応パターン	物性、立体構造 速度定数						
データ間のリンク	外部データベースへのリンク						
2013/5/28							

KEGG: Kyoto Encyclopedia of Genes and Genomes



Copyright 1999-2012 Kanehisa Laboratories

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



Copyright 1999-2012 Kanehisa Laboratories

http://www.kegg.jp/kegg_ja.html

KEGG: Kyoto Encyclopedia of Genes and Genomes

Current Statistics

KEGG Database as of 2013/5/23

Database	Description	Count
KEGG PATHWAY	Pathway maps, reference (total)	0 (245,370)
KEGG BRITE	Functional hierarchies, reference (total)	140 (78,544)
KEGG MODULE	KEGG modules, reference (total)	570 (183,985)
KEGG DISEASE	Human diseases	1,301
KEGG DRUG	Drugs	9,694
KEGG ENVIRON	Crude drugs and health-related substances	843
KEGG ORTHOLOGY	KEGG Orthology (KO) groups	16,646
KEGG GENOME	KEGG Organisms	2,684
KEGG GENES	Genes in high-quality genomes (190 eukaryotes, 2324 bacteria, 153 archaea)	10,765,444
KEGG 55DB	Best hit relations within GENES Bi-directional best hit relations within GENES	129,877,428,994 2,953,717,497
KEGG GENES	Genes in draft genomes (17 eukaryotes)	379,706
KEGG EGENES	Genes as EST contigs (99 eukaryotes)	3,792,883
KEGG MGenes	Genes in metagenomes (716 samples)	90,754,416
KEGG COMPOUND	Metabolites and other small molecules	17,011
KEGG GLYCAN	Glycans	10,985
KEGG REACTION	Biochemical reactions	9,317
KEGG RPAIR	Reactant pair chemical transformations	14,043
KEGG RCLASS	Reaction class	2,797
KEGG ENZYME	Enzyme nomenclature	5,973

Copyright 1995-2012 Kanehisa Laboratories.

2013/5/28 <http://www.kegg.jp/kegg/docs/statistics.html>

システムの知識

KEGG PATHWAY
KEGG BRITE
KEGG MODULE
KEGG MEDICUS

ゲノムの知識

KEGG GENES

化合物の知識

KEGG LIGAND

KEGG Objects Identifier

KEGG Objects

KEGG objects are biological entities from molecular to higher levels that are represented as database entries in KEGG, such as genes and proteins, small molecules, reactions, pathways, diseases and drugs.

KEGG Object Identifier

The identifier of each database entry is generally in the form of `db:entry` where "db" is the database name and "entry" is the entry name or the accession number (see [DBGET](#) for the list of database names and abbreviations). However, "db" may be omitted in thirteen of the fifteen databases. In these cases, the identifier is called a **KEGG object identifier** consisting of a database-dependent prefix and a five-digit number, which is unique across the databases. The KEGG object identifiers are often called D numbers. For example, the KEGG pathway identifier shown below, D01441, is thus equivalent to dr:D01441 or drug:D01441.

Database	Object	Prefix	Example
KEGG PATHWAY	Pathway map	map	keg04990
KEGG BRITE	Functional hierarchy	br	ko01003
KEGG MODULE	KEGG module	M	M00010
KEGG DISEASE	Human disease	H	H0002
KEGG DRUG	Drug	D	D01441
KEGG ENVIRON	Crude drug, etc.	E	E00048
KEGG ORTHOLOGY	KO group	K	K04527
KEGG GENOME	KEGG organism	T	T01001 (hsa)
KEGG GENES	Gene / protein	hsa	hsa3643
KEGG COMPOUND	Small molecule	C	C00001
KEGG GLYCAN	Glycan	G	G00109
KEGG REACTION	Reaction	R	R00259
KEGG RPAIR	Reactant pair	RP	RP04458
KEGG RCLASS	Reaction class	RC	RC00046
KEGG ENZYME	Enzyme	EC	EC2.7.10.1

The first two databases are manually curated by KEGG. The remaining two databases, KEGG GENES derived from RefSeq and KEGG ENZYME derived from ExplorEnz, are also given KEGG-original annotations.

2013/5/28

基本は Prefix + 5桁の数字

GENESは、
生物種コード：遺伝子コード

ゲノムネット

GenomeNet

KEGG KEGG2 PATHWAY DRITE MEDICUS DBGET [English | Japanese]

Search 統合データベース for Go Clear

ゲノムネット
ゲノムネットとは
お知らせ
強調

統合アーティベース
統合DBの概要
KEGGの概要
リリース情報

仮想品データベース
KLUG
varUB

研究支援データベース
計画ツール
他のツール
HPI
ノードパック

ゲノムネット統合アーティベース
DBGET search
LinkDB search Homepage

KEGG MeSH: 化学/医療用語辞典リソース
医薬品/医療機器/お薬手帳
KEGG: 生命システム情報統合アーティベース
KEGG2: 目次カタログ
KEGG PATHWAY: システム情報をバイオウェイ
KEGG BRITE: リスクアソシエーション
KEGG Organisms: 生物種ごとの入口
KEGG CENES: ゲノム情報
KLUG LIGAND: ゲノム情報

Reaction Ontology: 反応オントロジー分類
varDB: 抗原変異データベース
研究支援データベース
CYORF: シアノゾンテリア
UUJU: -株会社
EXPRESSION: 遺伝子発現プロファイル

KEGG パスバイオロジイ
BRITE 機能階層(日本語)一覧
KEGG 生物種一覧

データベース人間リンク

反応オントロジーデータベース

その他のプロジェクトデータベース

2013/5/28 [http://www.genome.jp/ \(英語\)](http://www.genome.jp/)
[http://www.genome.jp/ja/ \(日本語\)](http://www.genome.jp/ja/)

LinkDBによるデータベース間のリンク

LinkDB: database of link information

DIRECT LinkDB KEGG

Search all → links from in format

(Note) specify a single entry or multiple entries in the form of semicolon separated URLs.

Links are categorized into the following 4 types.

- original links are extracted from the database entries provided by the GenomeNet DBGET system.
- reverse links are derived from the original links by exchanging a source entry and its target entry.
- synthetic links are special original links to logically equivalent contexts between KEGG entries, COMPOUND, LIGAND, REACTION databases and databases other than KEGG.
- indirect links are derived by combining two or more original links. Currently, links from REACTION to REACTION via KEGG and/or COMPOUND via REACTION are supported.

LinksDB information

Released 2013-07-15, 11:11
KEGG UNIVERSAL Information Center
2,341,110,174 entries

Database and link statistics

DBGET entries
Indirect database
HML prefixes

KEGGを核として、国内外の様々なDBとリンク

LinkDBによるデータベース間のリンク

2013/5/28 Download link information from to in format

ゲノムネット計算ツール

varDB

研究支援データベース

計算ツール
「他のツール」

FTP

フィードバック

KEGG BRITE - システム情報、オントロジー
KEGG Organome - 生物種ごとの入口
KEGG RNFDB - デノム情報
KEGG LIGAND - ケミカル情報

Reaction Ontology - 反応のサーンと分類
varDB - 批量実験データベース
研究支援データベース
CYORT - シアノバクテリア
BSORF - 植草菌
EXPRESSION - 遺伝子発現プロファイル

ゲノムネット計算ツール

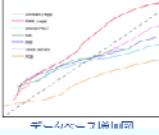
配列解析
BLAST / FASTA - ホモジジー検索
MOTIF - モチーフ検索
CLUSALW / MAJESTI / NN - 配列アライメント

ゲノム構造解析
OC-Viewer - オン・スクリーン表示
KAAS - KEGG自動アノテーションサーバー
LGassemblyer - LUIコノイグリゲート
GRNIF3 - 遺伝子ネットワーキング予測
DINIES - 医薬品-標的的キットワーク予測 *New!*

リミカル基團解析
GIMCOMI / SUUCOMI - 化合物構造検索
KCalM - 热力学活性度
PathCmp - 可逆性反応経路の計算
PathSearch - 類似反応経路検索
PathPred - 分解・合成反応経路予測
Ezyme - 化合物間の酵素反応予測



データベース間のリンク



データベース間のリンク

ホモジニー検索、モチーフ検索
マルチプルアライメント

遺伝子機能アノテーション、予測など

化学構造比較、パス予測など

支那人生物学研究所バイオインフォマティクスセンター

<http://www.genome.jp/> (英語)

<http://www.genome.jp/ja/> (日本語)

2013/5/28

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

支那人生物学研究所バイオインフォマティクスセンター

2013/5/28

KEGG: Kyoto Encyclopedia of Genes and Genomes

The screenshot shows the KEGG homepage with a sidebar containing links to KEGG Home, Database, Software, Objects, and more. The main content area features a search bar and links to KEGG2, PATHWAY, BRITE, MODULE, KO, GENOME, GENES, SSDB, and ORTHOLOGY (KO). A central box displays 'ORTHOLOGY (KO) 機能アノテーション情報' and 'GENES 配列データベース'. Below this is a cluster diagram labeled 'OC オーソログクラスター'. To the right is a 'GenomeNet Database Resources' section with a list of databases and a 'GenomeNet Bioinformatics Tools' section with various bioinformatics tools.

2013/5/28

KEGG GENES データベース

The screenshot shows the KEGG GENES database interface. It includes a navigation bar with links to KEGG2, PATHWAY, BRITE, MODULE, KO, GENOME, GENES, SSDB, and Organisms. Below the navigation bar is a search bar for 'Enter org:gene' with buttons for Entry, Gene cluster, Ortholog, Paralog, Motif, and Clear. A 'Gene Catalogs' section provides a detailed description of the KEGG GENES collection. A table lists gene catalogs categorized by 'Category' (e.g., Complete genomes, EST datasets, Metagenomes, Viruses) and 'Remark' (e.g., High-quality genomes with KOALA and manual annotation, Draft genomes with automatic (KAAS) annotation, EST contigs with automatic (KAAS) annotation, Metagenomes with automatic (KAAS) annotation). Below the table are two search boxes: one for 'Search GENES for' and another for 'Search Organism hsa for'. A green box highlights the 'GENES' category in the table. A blue box highlights the 'Remark' column. A red box highlights the 'Viruses' entry in the table. A blue box also highlights the 'Search GENES for' search box.

KEGG GENES の
サブカテゴリ

登録データの一覧表へのリンク
生物種単位、サンプル単位

2013/5/28

KEGG Organisms - GENESに登録されている生物種 -



KEGG Organisms: Complete Genomes

Eukaryotes: 190 Bacteria: 2323 Archaea: 153

[Genomes | Draft | ESTs | Meta | Pan]

Eukaryotes	分類	生物種コード	生物種名	データソース																																								
<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th>Category</th> <th>Species</th> <th>Source</th> </tr> </thead> <tbody> <tr> <td rowspan="18" style="vertical-align: top; text-align: center;"> Eukaryotes Mammals Vertebrates </td> <td>hsa: Homo sapiens (human)</td> <td>RefSeq</td> </tr> <tr> <td>ptr: Pan troglodytes (chimpanzee)</td> <td>RefSeq</td> </tr> <tr> <td>pan: Pan paniscus (bonobo)</td> <td>RefSeq</td> </tr> <tr> <td>gor: Gorilla gorilla gorilla (western lowland gorilla)</td> <td>RefSeq</td> </tr> <tr> <td>pon: Pongo abelii (Sumatran orangutan)</td> <td>RefSeq</td> </tr> <tr> <td>mac: Macaca mulatta (rhesus monkey)</td> <td>RefSeq</td> </tr> <tr> <td>mmu: Mus musculus (mouse)</td> <td>RefSeq</td> </tr> <tr> <td>rno: Rattus norvegicus (rat)</td> <td>RefSeq</td> </tr> <tr> <td>clu: Canis familiaris (dog)</td> <td>RefSeq</td> </tr> <tr> <td>amn: Ailuropoda melanoleuca (giant panda)</td> <td>RefSeq</td> </tr> <tr> <td>fca: Felis catus (domestic cat)</td> <td>RefSeq</td> </tr> <tr> <td>bta: Bos taurus (cow)</td> <td>RefSeq</td> </tr> <tr> <td>ssc: Sus scrofa (pig)</td> <td>RefSeq</td> </tr> <tr> <td>cab: Equus caballus (horse)</td> <td>RefSeq</td> </tr> <tr> <td>mdu: Monodelphis domestica (possum)</td> <td>RefSeq</td> </tr> <tr> <td>shr: Sarcophilus harrisii (tasmanian devil)</td> <td>RefSeq</td> </tr> <tr> <td>oaa: Ornithorhynchus anatinus (platypus)</td> <td>RefSeq</td> </tr> <tr> <td>qua: Gallus gallus (chicken)</td> <td>RefSeq</td> </tr> </tbody> </table>					Category	Species	Source	Eukaryotes Mammals Vertebrates	hsa: Homo sapiens (human)	RefSeq	ptr: Pan troglodytes (chimpanzee)	RefSeq	pan: Pan paniscus (bonobo)	RefSeq	gor: Gorilla gorilla gorilla (western lowland gorilla)	RefSeq	pon: Pongo abelii (Sumatran orangutan)	RefSeq	mac: Macaca mulatta (rhesus monkey)	RefSeq	mmu: Mus musculus (mouse)	RefSeq	rno: Rattus norvegicus (rat)	RefSeq	clu: Canis familiaris (dog)	RefSeq	amn: Ailuropoda melanoleuca (giant panda)	RefSeq	fca: Felis catus (domestic cat)	RefSeq	bta: Bos taurus (cow)	RefSeq	ssc: Sus scrofa (pig)	RefSeq	cab: Equus caballus (horse)	RefSeq	mdu: Monodelphis domestica (possum)	RefSeq	shr: Sarcophilus harrisii (tasmanian devil)	RefSeq	oaa: Ornithorhynchus anatinus (platypus)	RefSeq	qua: Gallus gallus (chicken)	RefSeq
Category	Species	Source																																										
Eukaryotes Mammals Vertebrates	hsa: Homo sapiens (human)	RefSeq																																										
	ptr: Pan troglodytes (chimpanzee)	RefSeq																																										
	pan: Pan paniscus (bonobo)	RefSeq																																										
	gor: Gorilla gorilla gorilla (western lowland gorilla)	RefSeq																																										
	pon: Pongo abelii (Sumatran orangutan)	RefSeq																																										
	mac: Macaca mulatta (rhesus monkey)	RefSeq																																										
	mmu: Mus musculus (mouse)	RefSeq																																										
	rno: Rattus norvegicus (rat)	RefSeq																																										
	clu: Canis familiaris (dog)	RefSeq																																										
	amn: Ailuropoda melanoleuca (giant panda)	RefSeq																																										
	fca: Felis catus (domestic cat)	RefSeq																																										
	bta: Bos taurus (cow)	RefSeq																																										
	ssc: Sus scrofa (pig)	RefSeq																																										
	cab: Equus caballus (horse)	RefSeq																																										
	mdu: Monodelphis domestica (possum)	RefSeq																																										
	shr: Sarcophilus harrisii (tasmanian devil)	RefSeq																																										
	oaa: Ornithorhynchus anatinus (platypus)	RefSeq																																										
	qua: Gallus gallus (chicken)	RefSeq																																										

2013/5/28

KEGG Metagenomes - MGENESに収録されているサンプル -



KEGG Metagenomes

[Genomes | ESTs | Meta | Pan]

Environmental samples	分類	生物種コード	生物種名	データソース																																								
<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th>Category</th> <th>Project</th> <th>Source</th> </tr> </thead> <tbody> <tr> <td rowspan="18" style="vertical-align: top; text-align: center;"> Ocean </td> <td>I30001: Planktonic microbial communities from North Pacific Subtropical Gyre</td> <td>NCBI</td> </tr> <tr> <td>I30002: Planktonic microbial communities from Monterey Bay, CA</td> <td>NCBI</td> </tr> <tr> <td>I30003: Human gut metagenome collected from healthy human sample F1-S (male adult)</td> <td>Metagenome.jp</td> </tr> <tr> <td>I30004: Human gut metagenome collected from healthy human sample F1-T (female adult)</td> <td>Metagenome.jp</td> </tr> <tr> <td>I30005: Human gut metagenome collected from healthy human sample F1-U (infant female)</td> <td>Metagenome.jp</td> </tr> <tr> <td>I30006: Human gut metagenome collected from healthy human sample F2-V (male adult)</td> <td>Metagenome.jp</td> </tr> <tr> <td>I30007: Human gut metagenome collected from healthy human sample F2-W (female adult)</td> <td>Metagenome.jp</td> </tr> <tr> <td>I30008: Human gut metagenome collected from healthy human sample F2-X (male child)</td> <td>Metagenome.jp</td> </tr> <tr> <td>I30009: Human gut metagenome collected from healthy human sample F2-Y (female child)</td> <td>Metagenome.jp</td> </tr> <tr> <td>I30010: Human gut metagenome collected from healthy human sample In-A (male adult)</td> <td>Metagenome.jp</td> </tr> <tr> <td>I30011: Human gut metagenome collected from healthy human sample In-B (male infant)</td> <td>Metagenome.jp</td> </tr> <tr> <td>I30012: Human gut metagenome collected from healthy human sample In-D (male adult)</td> <td>Metagenome.jp</td> </tr> <tr> <td>I30013: Human gut metagenome collected from healthy human sample In-E (male infant)</td> <td>Metagenome.jp</td> </tr> <tr> <td>I30014: Human gut metagenome collected from healthy human sample In-M (infant female)</td> <td>Metagenome.jp</td> </tr> <tr> <td>I30015: Human gut metagenome collected from healthy human sample In-R (female adult)</td> <td>Metagenome.jp</td> </tr> <tr> <td>I30016: VH0001 MetaHIT sample from healthy Danish female</td> <td>MetaHIT</td> </tr> <tr> <td>I30017: VH0002 MetaHIT sample from healthy Danish female</td> <td>MetaHIT</td> </tr> <tr> <td>I30018: VH0003 MetaHIT sample from healthy Danish male</td> <td>MetaHIT</td> </tr> </tbody> </table>					Category	Project	Source	Ocean	I30001: Planktonic microbial communities from North Pacific Subtropical Gyre	NCBI	I30002: Planktonic microbial communities from Monterey Bay, CA	NCBI	I30003: Human gut metagenome collected from healthy human sample F1-S (male adult)	Metagenome.jp	I30004: Human gut metagenome collected from healthy human sample F1-T (female adult)	Metagenome.jp	I30005: Human gut metagenome collected from healthy human sample F1-U (infant female)	Metagenome.jp	I30006: Human gut metagenome collected from healthy human sample F2-V (male adult)	Metagenome.jp	I30007: Human gut metagenome collected from healthy human sample F2-W (female adult)	Metagenome.jp	I30008: Human gut metagenome collected from healthy human sample F2-X (male child)	Metagenome.jp	I30009: Human gut metagenome collected from healthy human sample F2-Y (female child)	Metagenome.jp	I30010: Human gut metagenome collected from healthy human sample In-A (male adult)	Metagenome.jp	I30011: Human gut metagenome collected from healthy human sample In-B (male infant)	Metagenome.jp	I30012: Human gut metagenome collected from healthy human sample In-D (male adult)	Metagenome.jp	I30013: Human gut metagenome collected from healthy human sample In-E (male infant)	Metagenome.jp	I30014: Human gut metagenome collected from healthy human sample In-M (infant female)	Metagenome.jp	I30015: Human gut metagenome collected from healthy human sample In-R (female adult)	Metagenome.jp	I30016: VH0001 MetaHIT sample from healthy Danish female	MetaHIT	I30017: VH0002 MetaHIT sample from healthy Danish female	MetaHIT	I30018: VH0003 MetaHIT sample from healthy Danish male	MetaHIT
Category	Project	Source																																										
Ocean	I30001: Planktonic microbial communities from North Pacific Subtropical Gyre	NCBI																																										
	I30002: Planktonic microbial communities from Monterey Bay, CA	NCBI																																										
	I30003: Human gut metagenome collected from healthy human sample F1-S (male adult)	Metagenome.jp																																										
	I30004: Human gut metagenome collected from healthy human sample F1-T (female adult)	Metagenome.jp																																										
	I30005: Human gut metagenome collected from healthy human sample F1-U (infant female)	Metagenome.jp																																										
	I30006: Human gut metagenome collected from healthy human sample F2-V (male adult)	Metagenome.jp																																										
	I30007: Human gut metagenome collected from healthy human sample F2-W (female adult)	Metagenome.jp																																										
	I30008: Human gut metagenome collected from healthy human sample F2-X (male child)	Metagenome.jp																																										
	I30009: Human gut metagenome collected from healthy human sample F2-Y (female child)	Metagenome.jp																																										
	I30010: Human gut metagenome collected from healthy human sample In-A (male adult)	Metagenome.jp																																										
	I30011: Human gut metagenome collected from healthy human sample In-B (male infant)	Metagenome.jp																																										
	I30012: Human gut metagenome collected from healthy human sample In-D (male adult)	Metagenome.jp																																										
	I30013: Human gut metagenome collected from healthy human sample In-E (male infant)	Metagenome.jp																																										
	I30014: Human gut metagenome collected from healthy human sample In-M (infant female)	Metagenome.jp																																										
	I30015: Human gut metagenome collected from healthy human sample In-R (female adult)	Metagenome.jp																																										
	I30016: VH0001 MetaHIT sample from healthy Danish female	MetaHIT																																										
	I30017: VH0002 MetaHIT sample from healthy Danish female	MetaHIT																																										
	I30018: VH0003 MetaHIT sample from healthy Danish male	MetaHIT																																										

2013/5/28

生物種のゲノム情報

- 種の系統、データソース、アノテーションレベル、文献などの詳細情報 -

KEGG Homo sapiens (human)

Genome Info Pathway map Brite hierarchy Module Genome map Blast Taxonomy

Search genes: Go Clear 遺伝子のキーワード検索

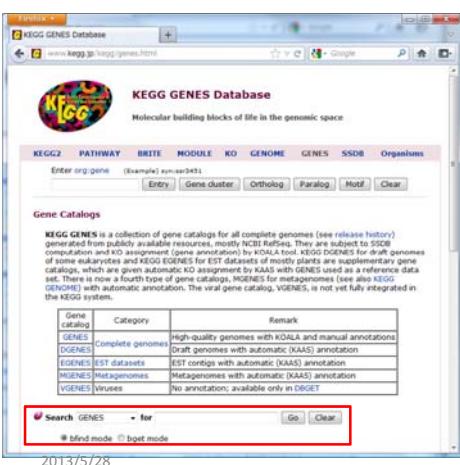
Genome information

T number: T01001
 Org code: hsa
 Aliases: HUMAN, 9606
 Full name: Homo sapiens (human)
 Definition: Homo sapiens (human)
 Annotation: manual
 Taxonomy: TAX: 9606
 Lineage: Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo
 Data source: RefSeq (Project: 168)
 Original DB: NCBI, OMIM, HGNC, HPRD, Ensembl
 Statistics: Number of protein genes: 19761
 Number of RNA genes: 1431
 Reference: PMID: 11237011
 Authors: Lander ES, et al.
 2013/5/28

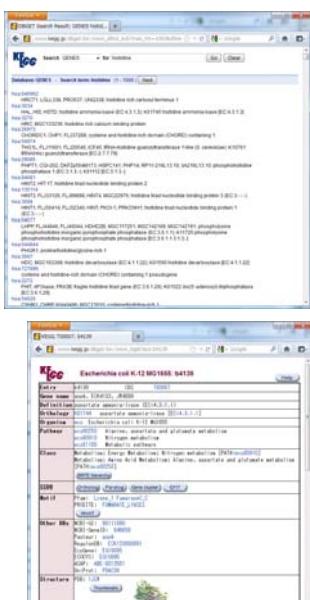
bfind/bget – ワード、エントリー検索

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

「Histidine」で GENESにbget



「eco:b4139」で bfind



KEGG GENESのIDを探したいときは？ - KEGG Objects -

Links to Outside Databases

Enter valid GO accession numbers to convert to KEGG GENES entries:

NCBI GenelD NCBI gi UniProt

(Example) 3775638 3737440 3743551 3897645 3902295

Convert Entry list Close

Last updated: January 4, 2012 Copyright 1999-2012 Kavli Institute of Nanoscience

2013/5/28

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

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

LinkDB from

Download link information from to In text - format Download

Download equivalent link information for a selected component

Download equivalent link information for LUMIUNO, UNIPROT, and REACTIONS

LUMIUNO for hubchrom In text - format Download Close

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

2013/5/28

GENESをbfindで検索

- 例: aspartate ammonia-lyase -

The left window shows the KEGG GENES Database homepage with a search bar containing 'aspartate ammonia-lyase'. The right window shows the search results page with a list of KEGG IDs and their corresponding gene names and EC numbers.

2013/5/28

KEGG GENES Entry (例: eco:b4139)

Entry: エントリ名、種類、生物種名
 Gene name: 遺伝子・タンパク質名、別名
 Definition: オリジナルDBの機能アノテーション
 Ortholog: KEGGでアサインしたKEGG Orthology (KO) アノテーション(後述)
 Pathway: エントリの遺伝子が機能するPathwayへのリンク
 Class: KEGGにおける機能カテゴリBRITE(階層テキスト) 203へのリンク

SSDB: SSDB (Similarity Score Database)からの情報抽出用のリンク。(Ortholog/Paralog推定、Gene Cluster)
 Motif: エントリが持つドメイン・モチーフの情報
 Other DBs: 他の配列DBへのリンク
 Structure: PDBへのリンク
 Position: エントリのゲノム上の位置(ゲノム情報がある場合)
 AA seq: アミノ酸配列取得、ホモロジー検索
 NT seq: 塩基配列取得
 All links: LinkDBでリンクされている内部および外部DB

KO (KEGG Orthology)

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

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

The screenshot shows a window titled "KEGG Orthology (KO)". The main content is a list of genes grouped under the heading "Carbohydrate Metabolism". Each entry includes a gene ID like "ko00001", a name like "D-Glutamate kinase [EC:2.7.1.1]", and a detailed description. For example, the first entry is "ko00001 D-Glutamate kinase [EC:2.7.1.1]". The list continues with other enzymes and their respective details.

2013/5/28

KO Entry

The screenshot shows the KEGG Orthology entry for KO1744. The top section displays basic information: Entry (KO1744), Name (aspA), Definition (aspartate ammonia-lyase [EC:4.3.1.1]), Pathway (ko0250: Alanine, aspartate and glutamate metabolism; ko00910: Nitrogen metabolism), Class (Metabolism: Energy Metabolism; Nitrogen metabolism [PATH:ko00910]), and Other DBs (RN: R00490, COG: COG1027, GO: GO000797). Below this, the "Genes" section lists numerous EC numbers and protein IDs, such as EC01: b4139(aspa), EC01: JN4099(aspa), EC01: 251744(aspa), EC01: E045120(aspa), EC01: E0474115_5055(aspa), EC01: EC08P_5238(aspa), EC01: E02490_4465(aspa), EC01: E00111_5074(aspa), EC01: E00105_4934(aspa), EC01: E22490_4465(aspa), EC01: G2583_4968(aspa), EC01: s5222(aspa), EC01: E0P_4933(aspa), EC11: HT168_24798(asca), and EC01: MP1001_2250(aspa). A "Taxonomy" section at the bottom lists MMB: Mash_0724, MEL: Metabo_2111, NMIC: Nair_1260, and TSV: TSV_0042. At the bottom, there are links for Taxonomy, KOALA, and UniProt.

Definition: KEGGで定義された機能アノテーション
(GENESエントリにおけるOrthologyフィールド)

Genes: 同一の機能アノテーションを付与されている遺伝子リスト

The screenshot shows the KEGG Organisms interface. The main content is a list of organisms grouped under the heading "Eukaryotes (313)". Each entry includes a name like "Klebsiella pneumoniae [KO1744]" and a detailed description. For example, the first entry is "Klebsiella pneumoniae [KO1744]". The list continues with many other eukaryotic species.

2013/5/28

OC Viewer - KEGG OC (Ortholog Cluster) -



2013/5/28

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



OC Viewer

Home

terms : pyruvate dehydrogenase

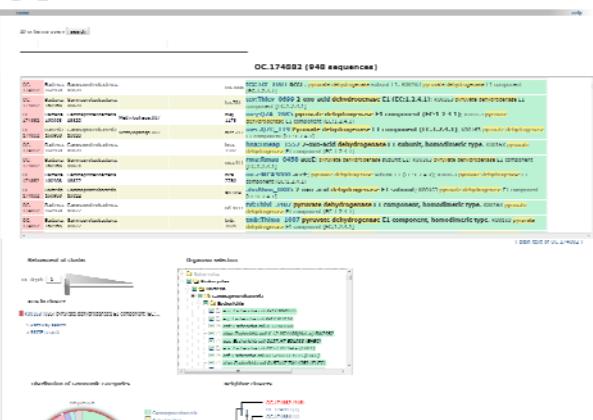
OC_174582 (947 / 948) - Dehydrogenase [EC:1.2.4.1]
OC_174583 (948 / 948) - Pyruvate dehydrogenase [EC:1.2.4.1]
OC_174584 (949 / 948) - Pyruvate dehydrogenase [EC:1.2.4.1]
OC_174585 (950 / 948) - Pyruvate dehydrogenase [EC:1.2.4.1]
OC_174586 (951 / 948) - Pyruvate dehydrogenase [EC:1.2.4.1]
OC_174587 (952 / 948) - Pyruvate dehydrogenase [EC:1.2.4.1]
OC_174588 (953 / 948) - Pyruvate dehydrogenase [EC:1.2.4.1]
OC_174589 (954 / 948) - Pyruvate dehydrogenase [EC:1.2.4.1]
OC_174590 (955 / 948) - Pyruvate dehydrogenase [EC:1.2.4.1]
OC_174591 (956 / 948) - Pyruvate dehydrogenase [EC:1.2.4.1]
OC_174592 (957 / 948) - Pyruvate dehydrogenase [EC:1.2.4.1]
OC_174593 (958 / 948) - Pyruvate dehydrogenase [EC:1.2.4.1]
OC_174594 (959 / 948) - Pyruvate dehydrogenase [EC:1.2.4.1]
OC_174595 (960 / 948) - Pyruvate dehydrogenase [EC:1.2.4.1]
OC_174596 (961 / 948) - Pyruvate dehydrogenase [EC:1.2.4.1]
OC_174597 (962 / 948) - Pyruvate dehydrogenase [EC:1.2.4.1]
OC_174598 (963 / 948) - Pyruvate dehydrogenase [EC:1.2.4.1]
OC_174599 (964 / 948) - Pyruvate dehydrogenase [EC:1.2.4.1]
OC_174600 (965 / 948) - Pyruvate dehydrogenase [EC:1.2.4.1]
OC_174601 (966 / 948) - Pyruvate dehydrogenase [EC:1.2.4.1]
OC_174602 (967 / 948) - Pyruvate dehydrogenase [EC:1.2.4.1]
OC_174603 (968 / 948) - Pyruvate dehydrogenase [EC:1.2.4.1]
OC_174604 (969 / 948) - Pyruvate dehydrogenase [EC:1.2.4.1]
OC_174605 (970 / 948) - Pyruvate dehydrogenase [EC:1.2.4.1]
OC_174606 (971 / 948) - Pyruvate dehydrogenase [EC:1.2.4.1]
OC_174607 (972 / 948) - Pyruvate dehydrogenase [EC:1.2.4.1]
OC_174608 (973 / 948) - Pyruvate dehydrogenase [EC:1.2.4.1]
OC_174609 (974 / 948) - Pyruvate dehydrogenase [EC:1.2.4.1]
OC_174610 (975 / 948) - Pyruvate dehydrogenase [EC:1.2.4.1]
OC_174611 (976 / 948) - Pyruvate dehydrogenase [EC:1.2.4.1]
OC_174612 (977 / 948) - Pyruvate dehydrogenase [EC:1.2.4.1]
OC_174613 (978 / 948) - Pyruvate dehydrogenase [EC:1.2.4.1]
OC_174614 (979 / 948) - Pyruvate dehydrogenase [EC:1.2.4.1]
OC_174615 (980 / 948) - Pyruvate dehydrogenase [EC:1.2.4.1]
OC_174616 (981 / 948) - Pyruvate dehydrogenase [EC:1.2.4.1]
OC_174617 (982 / 948) - Pyruvate dehydrogenase [EC:1.2.4.1]
OC_174618 (983 / 948) - Pyruvate dehydrogenase [EC:1.2.4.1]
OC_174619 (984 / 948) - Pyruvate dehydrogenase [EC:1.2.4.1]
OC_174620 (985 / 948) - Pyruvate dehydrogenase [EC:1.2.4.1]
OC_174621 (986 / 948) - Pyruvate dehydrogenase [EC:1.2.4.1]
OC_174622 (987 / 948) - Pyruvate dehydrogenase [EC:1.2.4.1]
OC_174623 (988 / 948) - Pyruvate dehydrogenase [EC:1.2.4.1]
OC_174624 (989 / 948) - Pyruvate dehydrogenase [EC:1.2.4.1]
OC_174625 (990 / 948) - Pyruvate dehydrogenase [EC:1.2.4.1]
OC_174626 (991 / 948) - Pyruvate dehydrogenase [EC:1.2.4.1]
OC_174627 (992 / 948) - Pyruvate dehydrogenase [EC:1.2.4.1]
OC_174628 (993 / 948) - Pyruvate dehydrogenase [EC:1.2.4.1]
OC_174629 (994 / 948) - Pyruvate dehydrogenase [EC:1.2.4.1]
OC_174630 (995 / 948) - Pyruvate dehydrogenase [EC:1.2.4.1]
OC_174631 (996 / 948) - Pyruvate dehydrogenase [EC:1.2.4.1]
OC_174632 (997 / 948) - Pyruvate dehydrogenase [EC:1.2.4.1]
OC_174633 (998 / 948) - Pyruvate dehydrogenase [EC:1.2.4.1]
OC_174634 (999 / 948) - Pyruvate dehydrogenase [EC:1.2.4.1]
OC_174635 (1000 / 948) - Pyruvate dehydrogenase [EC:1.2.4.1]

Show all (947)

Search by ID

Feedback KEGG Genome Wiki

OC Viewer - KEGG OC (Ortholog Cluster) -



2013/5/28

GenomeNetからの遺伝子データベース一括検索機能 - 例: aspartate ammonia-lyase -

The screenshot shows two windows side-by-side. The left window is the GenomeNet homepage with a search bar at the top containing 'aspartate ammonia-lyase'. The right window is a search results page from the DEBNET database, also showing 'aspartate ammonia-lyase' in the search bar. Both windows display a list of search results, which include various KEGG identifiers and descriptions related to the enzyme.

2013/5/28

KEGG, ゲノムネットの化合物情報 (LIGAND)

2013/5/28

KEGG: Kyoto Encyclopedia of Genes and Genomes

LIGAND
化合物情報データベース

SIMCOMP/SUBCOMP
化合物類似構造・部分構造検索

Copyright 1999-2012 Kanehisa Laboratories

2013/5/28

KEGG LIGAND データベース

KEGG2 PATHWAY BRITE MODULE LIGAND COMPOUND GLYCAN REACTION

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

Filter Pathway mapping Brite mapping Get title Get entry Clear

Chemical Substances and Reactions

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

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

Search LIGAND for Go Clear

bfind mode bget mode

ID番号のIdentifierと各DBに含まれるデータの説明

化合物、糖鎖、反応に特化した入口

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

2013/5/28

KEGG COMPOUND (代謝化合物情報)

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

The screenshot shows the KEGG Compound database interface. At the top, there's a search bar with 'Enter C numbers' and several buttons: Filter, Pathway mapping, Links mapping, Get title, Get entry, and Clear. Below the search bar is a navigation menu with tabs: KEGG2, PATHWAY, BRITE, MODULE, LIGAND, COMPOUND, GLYCAN, and REACTION. The main content area displays a list of compounds with their names and IDs. A specific compound entry is shown in detail, including its structure, chemical name, and various properties. The bottom of the page contains sections like 'Biosynthetic Codes' and 'KEGG RELATIONAL DATABASE'.

2013/5/28

KCFフォーマットとKEGG Atom type

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

The screenshot shows the KEGG atom types database interface. It lists various atom functional groups with their descriptions and frequencies. The table has columns for 'Atom functional group', 'Description', and 'Frequency'. A red box highlights the first few rows of the table, showing examples like 'C-S', 'C=O', and 'C-C'. The frequency column shows values such as 10473, 9039, 4964, etc.

Atom functional group	Description	Frequency
C-S	W=CH-S	10473
	W=CH-O-S	9039
C=O	W=CH=O	4964
	W=CH-C(=O)-	498
Cx	Ring-CH2-ring	14930
Cy	Ring-CH2-ring	27776
Cz	Ring-CH2-C(=O)-ring	4483
Amine	W=C-NH2	934
	W=C-NH-C(=O)-	2969
	W=C-NH-C(=O)-NH2	1914
Cx2	Ring-CH2-Hydroxyl	2964
Cy2	Ring-CH2-Hydroxyl	7732
Cz2	Ring-CH2-C(=O)-ring or ring-C(=O)-ring	43
Alkane	W=C	29
	W=C-C(=O)-	350
Cx3	R=C(=O)-R	3595
Cx4	Ring-CH2-C(=O)-ring	2237
Cx5	Ring-CH2-C(=O)-OH	2196
Cx6	Ring-CH2-C(=O)-OH	1991
Cx7	Ring-CH2-C(=O)-O-ring	890
Cx8	Ring-CH2-C(=O)-O-ring	1990
Cx9	Ring-CH2-C(=O)-ring	2551
Cx10	Ring-CH2-C(=O)-ring	1976
N		
Nx	W=C-NH2	2440
	W=C-NH-C(=O)-	2005
	W=C-NH-C(=O)-NH2	374
Ny	W=C-NH-C(=O)-	215
	W=C-NH-C(=O)-NH2	806
Nz	Ring-NH2-ring	18434
Nx2	W=C-NH-C(=O)-	230
Ny2	W=C-NH-C(=O)-	182
Nz2	Ring-NH2-ring	357
Nx3	W=C-NH-C(=O)-ring	34
Ny3	W=C-NH-C(=O)-NH2	795
Nz3	Ring-NH2-ring	840
Nx4	W=C-NH-C(=O)-NH2	2131
Ny4	W=C-NH-C(=O)-ring	36
Nz4	Ring-NH2-ring	194
O		
Ox	W=C-OH	15599
	W=C-O-C(=O)-	198
Oy	W=C-O-C(=O)-	3111
	W=C-O-C(=O)-NH2	332
Oz	W=C-O-C(=O)-NH2	4199
Ox2	Ring-OH	2411
Oy2	Ring-O-C(=O)-	505
Oz2	Ring-O-C(=O)-NH2	101
Ether		

2013/5/28

KEGG GLYCAN (糖鎖構造情報)

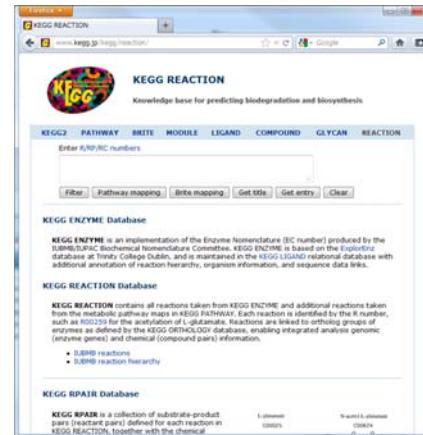
- 糖鎖分子の構造を収録したデータベース
 - エントリー数は、約1万1000
- 構造は2次元のグラフとして持つ
 - グラフは、実際にはツリーとなる
- ファイルフォーマットはKCF (Glycan)
 - Compound/Drugとの互換性はない



2013/5/28

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

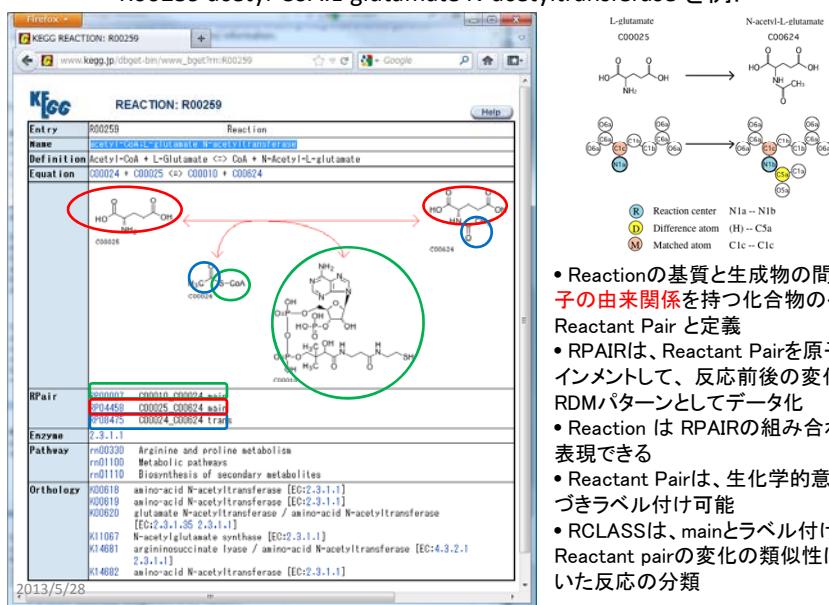
- REACTAION (生化学反応)
 - 生体内反応のデータベース
 - 代謝系の酵素反応がメイン
- RPAIR(基質-生成物変化)
- RCLASS(反応分類)
 - RPAIRのパターンによる反応分類
- ENZYME(酵素番号)
 - IUBMBの酵素番号



2013/5/28

REACTION, RPAIR, RCLASS

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



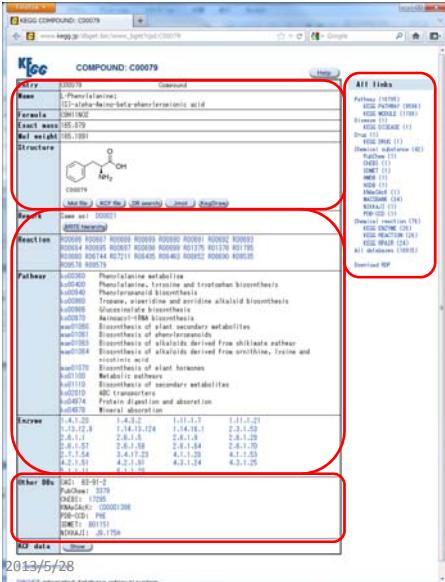
LIGANDをbfindでキーワード検索 (1)

- 例: phenylalanine -

The screenshots show the KEGG LIGAND Database interface. The first window shows the search bar with "phenylalanine". The second window shows the search results, listing various phenylalanine-related compounds. The third window shows a detailed view of one of the results, highlighting "Phenylalanine" in red.

2013/5/28

KEGG COMPOUND Entry



構造表示、構造情報(mol, KCF)、構造検索、構造表示(&編集)、

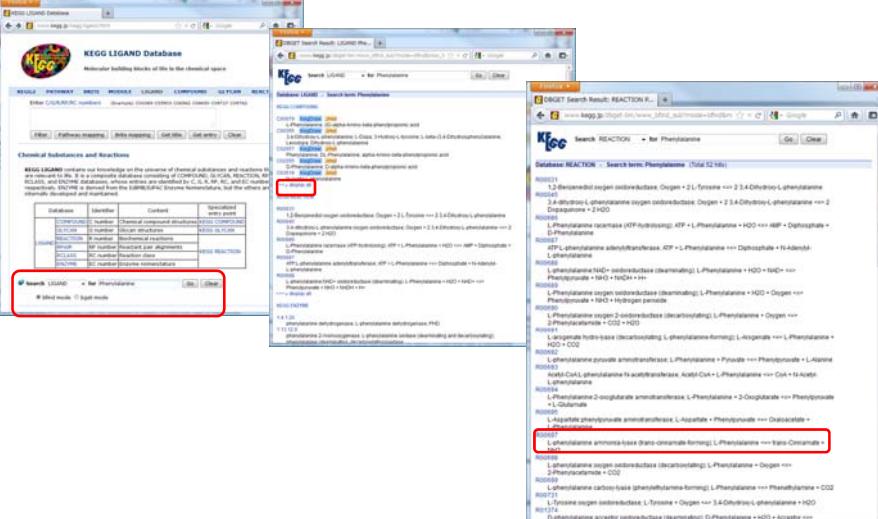
各種データベースへのリンク
REACTION, PATHWAY, ENZYME

外部データベースへのリンク
PubChem, ChEBI, etc.

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

LIGANDをbfindでキーワード検索 (2)

- 例: phenylalanine -



2013/5/28

KEGG REACTION Entry

反応式(物質名、C番号、図)、各種データベースへのリンク RPAIR, PATHWAY, ENZYME, KO

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

2013/5/28

KEGG RPAIR Entry

アライメントの結果(図)
ペアのフラグ情報(生化学的意味)
RDMの文字列表記
各種データベースへのリンク (COMPOUND, RCLASS, Related RPAIR, REACTION, ENZYME)

2013/5/28

KEGG RCLASS Entry

RPAIR

- RC00361 RPAIR RPAIR
- (DB search)

Related class

Reaction

Enzyme

Pathway

Orthology

RDM変換パターン(文字列、図)
RCLASSに含まれる RPAIR
各種データベースへのリンク
(RPAIR, Related RCLASS, REACTION, ENZYME, PATHWAY, KO)

2013/5/28

GenomeNetからの化合物データベース一括検索機能

カタゴリー 1 ~ 3 のデータベース

統合データベースを選択しキーワードを
入力するとメニューに表示されている
データベース全てに対する検索となる。

英語: <http://www.genome.jp/>
日本語: <http://www.genome.jp/ja/>

2013/5/28

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

2013/5/28

KEGG: Kyoto Encyclopedia of Genes and Genomes

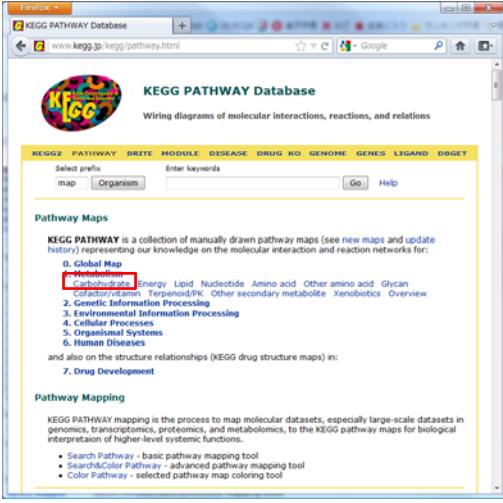
The screenshot displays the KEGG homepage and several overlapping windows illustrating its components:

- KEGG Home:** Shows links to Recent notes, Current statistics, Plea from KEGG, KEGG Database, KEGG overview, Searching KEGG, KEGG mapping, Color codes, KEGG Objects, Pathway maps, Brite hierarchies, KEGG Software, KegTools, KEGG API, KDM, KEGG FTP, Subscription, GenomeNet, OBO/LinkDB, Feedback, and Kanehisa Labs.
- KEGG Encyclopedia of Genes and Genomes:** Describes KEGG as a database resource for understanding the biological system, mentioning its focus on molecular-level information and experimental technologies.
- PATHWAY パスウェイマップ:** A window showing a pathway map with nodes and connections.
- BRITE 機能階層・オントロジー:** A window showing a hierarchical tree of functional categories.
- MODULE モジュール (機能ユニット):** A window showing a modular organization of pathways.
- Mapper PATHWAY/BRITE/MODULEへのマッピングツール群:** A window showing a collection of tools for mapping between these KEGG databases.
- GenomeNet Database Resources:** Lists various KEGG databases including KEGG PATHWAY, KEGG BRITE, KEGG GENOME, KEGG GENES, KEGG LIGAND, KEGG MEDICUS, KEGG Organisms, KEGG Atlas, KAAS, BLAST/Fasta, SINCERITY, and PathPred.
- GenomeNet Bioinformatics Tools:** Lists various bioinformatics tools such as Sequence analysis, BLAST search, Sequence similarity search, BLASTP/BLASTN, Multiple alignment, Genome Analysis, CO viewer, KEGG ortholog clusters, and KEGG ortholog search.

2013/5/28

KEGG PATHWAY

生体内(外)の分子間ネットワーク図



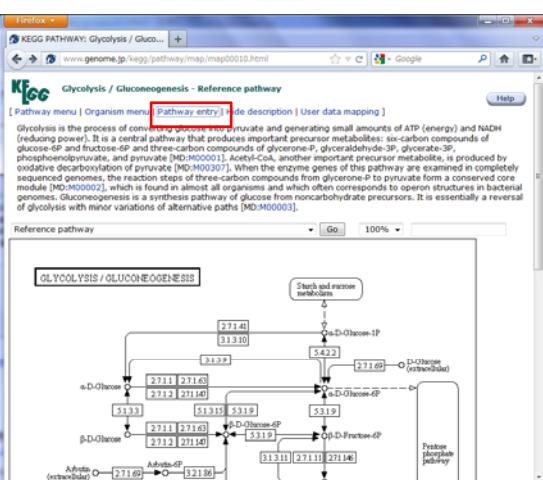
<http://www.kegg.jp/kegg/pathway.html>

Carbohydrate
Glycolysis / Gluconeogenesis

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

KEGG PATHWAY

マップの例：解糖系



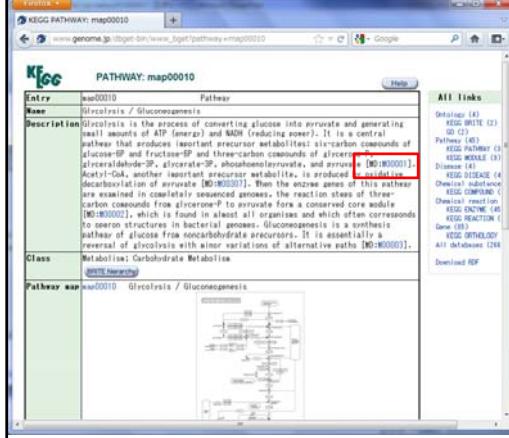
- 酵素／遺伝子と化合物のネットワーク
- Pathway menu
 - BRITE 形式の階層分類
- Organism menu
 - 生物種の階層分類
- Pathway entry
 - パスウェイデータベースのテキストバージョン
- Hide description
 - マップの説明を隠す
- User data mapping
 - マップ中のオブジェクトへの色付け
 - マップの拡大縮小

2013/5/28
<http://www.kegg.jp/kegg/pathway/map/map00010.html>

KEGG PATHWAY

テキストエントリー

- Pathway entry
 - Entry
 - マップ番号 (map|ko|ec|rn|生物種コード+番号)
 - Name, Description
 - パスウェイの説明とモジュールとの関係
 - Class
 - 階層分類情報
 - Pathway map
 - マップとオーソログテーブルへのリンク
 - Disease
 - Reference
- モジュール
 - 生物種間での保存、複合体、オペロンを考慮した機能単位

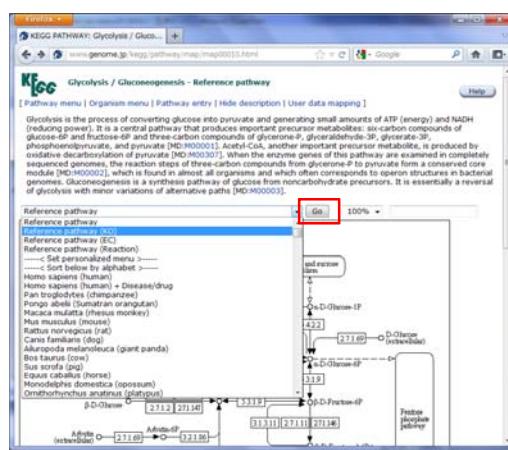


http://www.genome.jp/dbget-bin/www_bget?pathway+map00010

KEGG PATHWAY

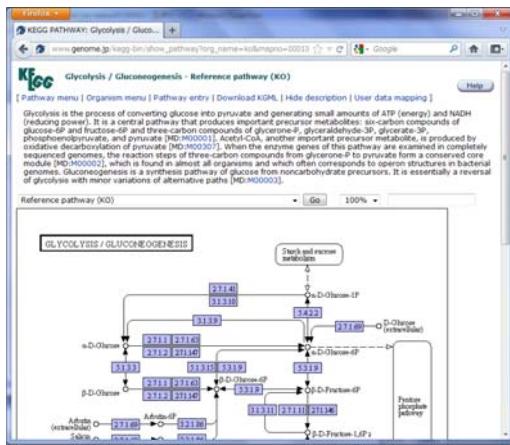
マップの例：解糖系

- Reference pathway
 - K0, EC, Reactionへのリンク
 - (KO) オーソログエントリーへのリンク
 - (EC) 酵素エントリーへのリンク
 - (Reaction) 反応エントリーへのリンク
- 生物種名
 - 各生物種の遺伝子エントリーへのリンク
- Set personalized menu
 - 生物種の選択
- Sort below by
 - 生物種名のソート



2013/5/28
<http://www.genome.jp/kegg/pathway/map/map00010.html>

KEGG PATHWAY



マップの例：解糖系（EC）

- Reference pathway (EC)

- 従来のリファレンスに対応
- 対応する酵素エントリーのある箱に色づけ

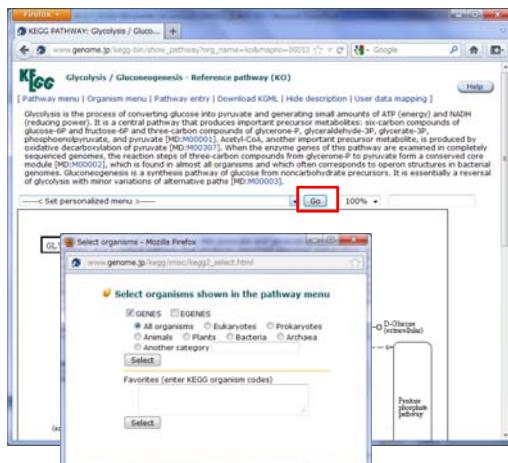
- Reference pathway (KO)

- 対応するオーソログエントリーのある箱に色づけ
- KEGG GENES に登録されている生物種が持つ遺伝子に関して配列の類似度を元に定義されているオーソログ情報
- 酵素によってはオーソログが定義できないものもある

<http://www.kegg.jp/kegg/pathway/ko/ko00010.html>

KEGG PATHWAY

生物種の選択



- リストの生物種を限定する

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

2013/5/28

http://www.kegg.jp/kegg/catalog/org_list.html

PATHWAY の検索とマッピング



キーワード検索

- Entry, Name, Description フィールドとマップ中のオブジェクト（遺伝子、オーソログ、反応、化合物）や注釈を対象とした検索
- 複数キーワードは AND 検索
- **パスウェイマッピング**
- マップ中のオブジェクトを指定してパスウェイにマッピング
- 複数オブジェクトを指定するとマッチしたもののすべてをマッピング
- Search Pathway
 - 指定したオブジェクトを赤く色づけ
- Search & Color Pathway
 - 指定したオブジェクトを自由に色づけ
- Color Pathway
 - 指定したパスウェイにオブジェクトの色データを与えて色づけ、数値を与えることも可能

2013/5/28

PATHWAY のキーワード検索

Entry	Thumbnail Image	Name	Description	Object
hsa05215		Prostate cancer - Homo sapiens (human)	Prostate cancer constitutes a major health problem in Western countries. It is the second most frequently diagnosed cancer in men.	...010: MAPK signaling pathway, receptor interacting protein kinase 1; cancer hsa04110: Cell cycle
hsa05223		Non-small cell lung cancer - Homo sapiens (human)	Lung cancer is a leading cause of cancer death among men and women in industrialized countries. Non-s... ...RAF 3843 (XKRAS hsa05223: Non-smoking related lung cancer) is the most common genetic alteration in non-smoking lung cancer. ...p33 signaling p... ...RAF 3843 (XKRAS hsa05223: Non-smoking related lung cancer) is the most common genetic alteration in non-smoking lung cancer. ...p33 signaling p...	...RAF 3843 (XKRAS hsa05223: Non-smoking related lung cancer) is the most common genetic alteration in non-smoking lung cancer. ...p33 signaling p...
hsa0513		Endometrial cancer - Homo sapiens (human)	Endometrial cancer (EC) is the most common gynecological malignancy and the fourth... ...MAPK signaling pathway hsa04110: Endometrial cancer hsa04110: Wnt signaling pathway hsa04110: p53 signaling p...	...MAPK signaling pathway hsa04110: Endometrial cancer hsa04110: Wnt signaling pathway hsa04110: p53 signaling p...

2013/5/28

PATHWAY のオブジェクトに好きな色を付ける

KEGG2 PATHWAY BRITE MODULE DISEASE DRUG KO GENOME GENES LIGAND DISEASE

Select profile: hsa [Organism] Enter keywords: cancer Go Help

Pathway Maps

KEGG PATHWAY is a collection of manually drawn pathway maps (see new maps and update history) representing our knowledge on the molecular interaction and reaction networks for:

- Global Map
- Metabolism: Carbohydrate, Energy, Lipid, Nucleotide, Amino acid, Other amino acid, Glycan, Cell wall, Terpenoid/PK, Other secondary metabolite, Xenobiotics, Overview
- Genetic Information Processing
- Environmental Information Processing
- Cellular Processes
- Organismal Systems
- Human Diseases

and also on the structure relationships (KEGG drug structure maps) in:

- Drug Development

Pathway Mapping

KEGG PATHWAY mapping is the process to map molecular datasets, especially large-scale datasets in genomics, transcriptomics, proteomics, and metabolomics, to the KEGG pathway maps for biological interpretation of higher-level systemic functions.

- Search Pathway - basic pathway mapping tool
- SearchColor Pathway - advanced pathway mapping tool
- Color Pathway - selected pathway map coloring tool

KEGG Mapper: Color Objects in KEGG Pathways

KO01B03 red,blue
C00118 pink

Alternatively, enter the file name containing the data:
ファイルを選択 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

Execute Clear

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

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

2013/5/28

PATHWAY のオブジェクトに好きな色を付ける

Pathway Search Result

Show all objects

- ko000362 Inositol phosphate metabolism (2)
- ko00010 Glycolysis / Gluconeogenesis (2)
- ko00103 TPI, tpIa, triosephosphate isomerase (TIM) [EC:5.3.1.1]
- ko00051 Fructose and mannose metabolism (2)
- ko01100 Metabolic pathways (2)
- ko00710 Carbon fixation in photosynthetic organisms (2)
- ko01110 Biosynthesis of secondary metabolites (2)
- ko01120 Microbial metabolism in diverse environments (2)
- ko00900 Terpenoid backbone biosynthesis (1)
- ko00730 Thiamine metabolism (1)
- ko00052 Galactose metabolism (1)
- ko00030 Pentose phosphate pathway (1)
- ko00040 Pentose and glucuronate interconversions (1)
- ko00750 Vitamin B6 metabolism (1)
- ko00331 Clavulanic acid biosynthesis (1)
- ko00680 Methane metabolism (1)

KEGG PATHWAY: Glycolysis / Gluconeogenesis - Reference pathway (KO)

GLYCOLYSIS / GLUCONEOGENESIS

Pathway menu | Pathway entry | Show description | User data mapping | Reference pathway (KO) Go

Detailed pathway diagram showing the flow from glucose to glyceraldehyde-3P and fructose-1,6P2, involving enzymes like 2.7.1.41, 3.1.3.10, 5.4.22, etc.

2013/5/28

KEGG GENOME

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

KEGG GENOME Database
Organisms and ecosystems with genome sequence information

KEGG2 PATHWAY BRITE KO GENES SSDB GENOME Organisms

Enter KEGG organism code or use Organism button to select
Organism: Go Clear (Examples) has mmu dmz cel ath zce evo bau smj

KEGG Organisms and Ecosystems

KEGG GENOME has been a collection of KEGG organisms (see release history), which are the organisms with known complete genome sequences supplemented by those with massive EST datasets (see also KEGG GENES). KEGG GENOME is being extended in three ways. First, it now contains metagenomes representing collections of genomes from a variety of environments. Second, KEGG organisms with high-quality KAALA annotation can now be combined computationally to examine, for example, reconstructed pathways for pangenomes and organism groups. Third, virus genomes will be integrated and virus genes will be annotated.

Category	Genome	Identifiers	Gene catalog	Annotation
Organisms	Complete genomes	T0 numbers or three-letter organism codes	GENES	KAALA
	EST datasets	T1 numbers or four-letter organism codes starting with "e"	DGENES	KAAS
	Metagenomes	T2 numbers or four-letter organism codes starting with "m"	EGENES	KAAS
	Pangenomes	T3 numbers	MGENES	KAAS
Viruses	VGENOME	RefSeq identifiers	VGENES	None

KEGG Organism Groups

An organism group is defined as a combination of KEGG organisms, enabling the analysis of combined pathway maps for the group, for example, in symbiosis or pathogenesis.

Define organism group (enter three-letter organism codes): Go

(Metagenome examples)
has+T30003 Homo sapiens (human) + gut microbiome New!

(Symbiosis examples)
lal+lot Lathyrus japonicus (lotus) + Mesorhizobium loti

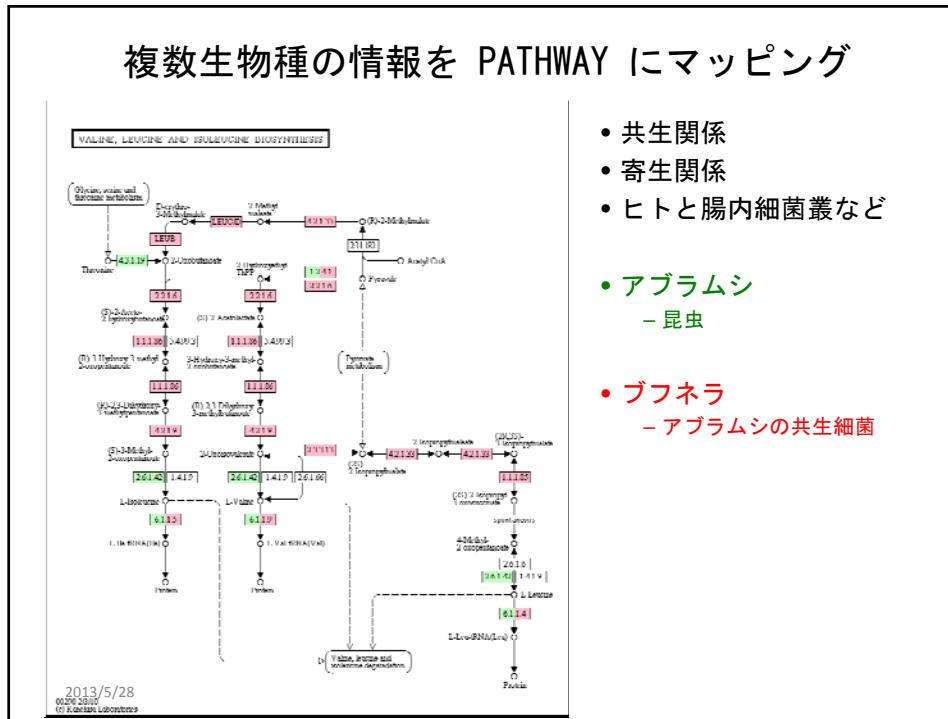
api+buc Anopheles gambiae (mosquito) + Buchnera aphidicola

smi+wol Brugia malayi (filaria) + Wolbachia

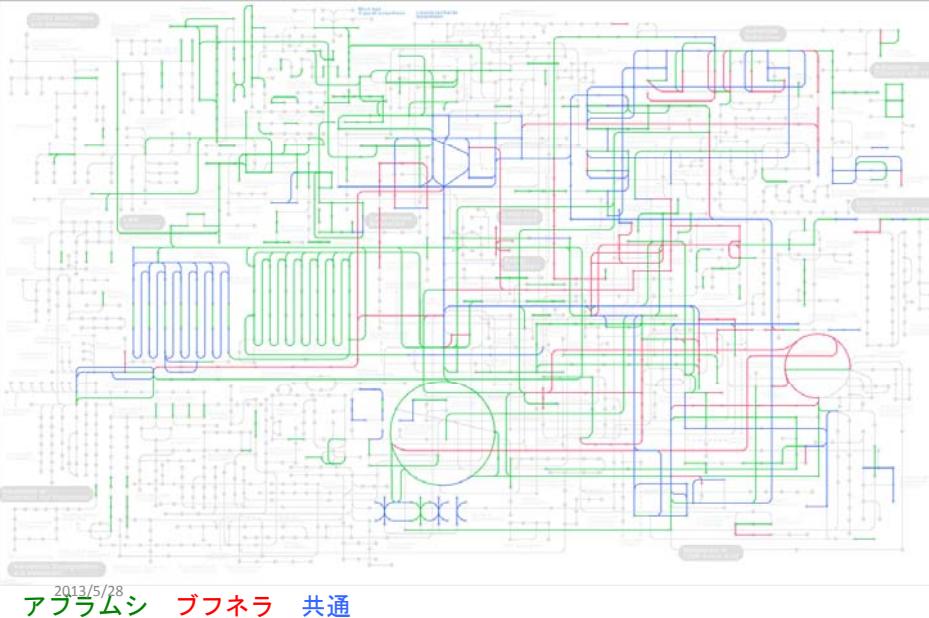
(Pathogenesis examples)
hsa+pfa Homo sapiens (human) + Plasmodium falciparum

age+pfa Anopheles gambiae (mosquito) + Plasmodium falciparum

2013/5/28

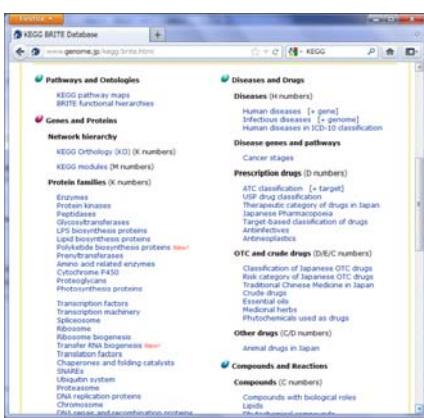


複数生物種の情報を PATHWAY にマッピング



KEGG BRITE

- 機能の階層分類情報 -



- Pathway and ontology
 - KEGG PATHWAY, KEGG BRITE のエントリ一覧を階層で表現
- Genes and Proteins
 - 遺伝子や蛋白質の機能やネットワークの分類。分類対象はKO
- Compound and Reactions
 - 化合物の構造や機能による分類、酵素反応の種類による分類
- Drug and Diseases
 - 薬物や疾患のような医薬系の情報の分類
- Cell and Organisms
 - KEGGの登録生物の系統分類

2013/5/28

KEGG BRITE

- 機能の階層分類情報 -

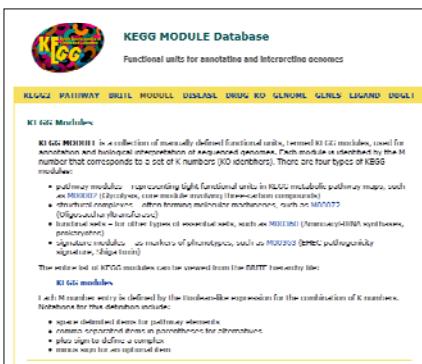


- キーワード検索
 - BRITE階層情報のテキスト検索
- オブジェクトマッピング KEGG Mapper
 - マップ中のオブジェクトを指定してBRITEにマッピング
 - 複数オブジェクトを指定するとマッチしたもののすべてをマッピング
 - Search BRITE
 - 指定したオブジェクトを赤く色づけ
 - Search & Color BRITE
 - 指定したオブジェクトを自由に色づけ
 - Join BRITE
 - BRITEに別の関係情報をマッピング
(例: GPCRのBRITE(ko04030)に、GPCRとリガンドの関係情報をマッピング)

2013/5/28

KEGG MODULE

- マニュアルで定義された生物学的な機能ユニットのコレクション -

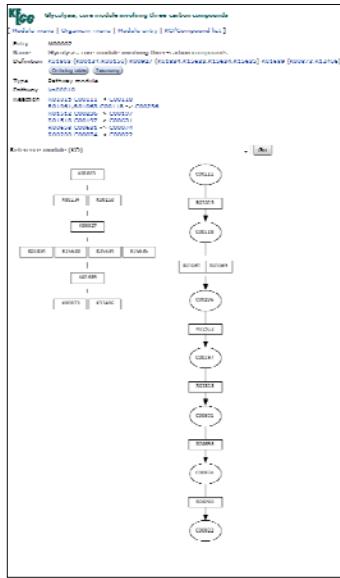


- Pathway modules
 - KEGG PATHWAYの代謝パスウェイで、緊密な機能単位としてなる連続経路
- Structural complexes
 - 分子
- Functional sets
 - その他の必須な機能単位
- Signature module
 - 表現型のマーカーなど

2013/5/28

KEGG MODULE

パスウェイモジュールの例



2013/5/28

- Entry

 - ID

- Name

 - 名前

- Definition

 - モジュールの構造

- Type

 - モジュールのタイプ

- Pathway

 - モジュールが存在するパスウェイ

- Reaction

 - モジュールを構成する反応

KEGG Mapper

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

- KEGG Object (+ 色、あるいは数値) を指定してパスウェイにマッピング

- 目的に合わせて下記の11種類のツールがある

Concepts behind KEGG Mapper

KEGG PATHWAY and KEGG MODULE are the references to knowledge bases for biological interpretation of molecular datasets, especially large-scale datasets generated by high-throughput experimental techniques. KEGG Mapper will try to understand pathway-level evidence functions, based on the relationship between the query data and the reference knowledge as shown below.

Query data	Reference knowledge	Tool
Objects	KEGG PATHWAY database	Search Pathway
Object-set(Brite modules)	KO 1.2.1 PATHWAY database	Search Brite
Object-set(Brite modules)	Single KEGG pathway map	Color Pathway / Color Pathway 3D
Gene KO list	KO 1.2.1 PATHWAY database (KO)	Reconstruct Pathway
Objects	KEGG BRITE database	Search Brite
Object-set(Brite modules)	KO 1.2.1 BRITE database	Search Brite
Object-set(Brite modules)	Single Brite functional hierarchy	Color Brite
Gene KO list	KO 1.2.1 BRITE database (KO)	Reconstruct Brite
Objects	KEGG MODULE database	Search Module
Object-set(Brite modules)	KO 1.2.1 MODULE database	Search Module
Gene KO list	KEGG MODULE database (KO)	Reconstruct Module

The query data may be a collection of molecular objects (genes, proteins, small molecules, etc.) or a more refined set of object-set(Brite modules, search modules).

- Pathways and their presence in the genome (protein families in KEGG Pathways)
- Genes associated with pathways (proteins in KEGG Pathways)
- Genes associated with pathways (KO representation)

2013/5/28

KEGG Mapper

- Search Pathway/Brite/Module
 - 基本的なマッピングツール
 - KEGG Object を入力して、Pathwayなどにマッピングする
- Search&Color Pathway/Brite/Module
 - より高度なマッピングツール
 - KEGG Object と bgcolor, fgcolor の色を入力として、Pathwayなどの任意のObjectに指定色をマッピングする
- Color Pathway, Color Pathway 3D
 - Search&Color Pathway の拡張
 - KEGG Object と 数値データを入力として、Pathwayマップにカラースケールの色を塗り分けることができる
- Reconstruct Pathway/Brite/Module
 - 一つあるいは複数の遺伝子とKOの対応リストを入力に、Pathwayなどにマッピングしてパスウェイの再構築などを行うことができる
- Join Brite
 - たとえば、薬とそのターゲットのような二項関係のリストを、BRITEの階層ファイルにマッピングして結合するツール

2013/5/28

Search Pathway

KEGG Mapper **Search Pathway**

KEGG Mapper Search Pathway Search&Color Pathway Color Pathway Reconstruct Pathway
 KEGG Object Search Brite Search Brite Color Pathway 3D Reconstruct Brite Reconstruct Module
 BRITE Search Module Search Color Module Color Pathway 3D Reconstruct Brite Reconstruct Module

Search against: KO Enter: map, ko, gene, briteid, or [file]

Enter objects:

```
KO:7107 hsa:SP1 tpd:CO0118
```

Example:
 (Reference pathway (KO))
 ko00100|ko00111|tpd:CO0118
 KU0124|KU0220
 (Home sapiens pathway)
 hsa0107|ko00100|tpd:CO0118
 ALU01.2.1.12|KU0220

Alternatively, enter the file name containing the data:

File1 File2 (or untar K0/K0/K0/K0/K0)

Include aliases

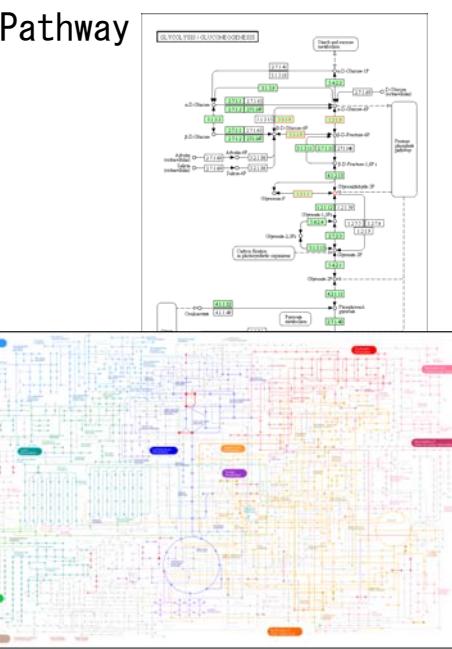
Display objects not found in the search

Search pathways containing all the objects (NWS search)

[Search] [Clear]

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

Pathway	KO ID identifier	Alias
map Reference pathway	K0/K0 numbers, CO/K0 numbers, K0 alias	
map Reference pathway (non-metabolic)	K0/K0 numbers, CO/K0 numbers	K0 alias
map Reference pathway (metabolic)	K0/K0 numbers, CO/K0 numbers	K0 alias
ko Reference pathway (KO)	K0 numbers, CO/K0 numbers	K0 alias, EC numbers
cc Reference pathway (EC)	EC number, CO/K0 numbers	
cc Reference pathway (reaction)	Reaction numbers, CO/K0 numbers	Reaction numbers
org Organism-specific pathway	Species identifier, CO/K0 numbers	K0/K0 numbers



2013/5/28

32

Reconstruct Pathway

KEGG Mapper – Reconstruct Pathway

KEGG Mapper Search Pathway Search&Color Pathway Color Pathway Reconstructed Pathway
KEGG Search Module Search&Color Module Color Pathway 3D Join Brite Reconstructed Module

Enter gene list with KEGG annotations:

```
kw02020 K12393
kw01040 K17373
kw01051 K25924
kw01052 K25925
kw01053 K25926
kw01054 K25927
kw01055 K25928
kw01056 K25929
kw01057 K25930
kw01058 K25931
kw01059 K25932
kw01059 K25933
kw01059 K25934
```

Example file:
kw02020.kml (single organism)
kw02020.kml (multiple organisms)

Alternatively, enter the file name containing the data:

If necessary, change default file type:

Include (global) map

Reconstructed Pathway is a KEGG PATHWAY mapping tool that assists genomic and metagenomic annotation. The input data is a single gene list (for a single organism) or multiple gene lists (for multiple organisms) annotated with KEGG Orthology (KO) identifiers or KEGG Brite. Each line of the gene list must be preceded by the organism identifier, if any, then the KO identifier number. The mapping is performed through the KEGG Brite modules against the KEGG reference pathways.

When multiple organisms are given in the input file, each organism must be separated by the # sign, which may come after defined line separator. Default settings for this step are as follows:

Org 1 2 3 4
regular map global map
See KEGG color codes for more details.

Last updated: June 1, 2012

2013/5/28

Search&Color Pathway

KEGG Mapper – Search&Color Pathway

KEGG Mapper Search Pathway Search&Color Pathway Color Pathway Reconstructed Pathway
KEGG Search Module Search&Color Module Color Pathway 3D Join Brite Reconstructed Module

Search against: hco Enter map, ko, ec, m, hood, or org

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

Enter objects one per line followed by <tgcolor>, <tgcolor>:

```
120 /rodBac
C00118 pink
```

Alternatively, enter the file name containing the data:

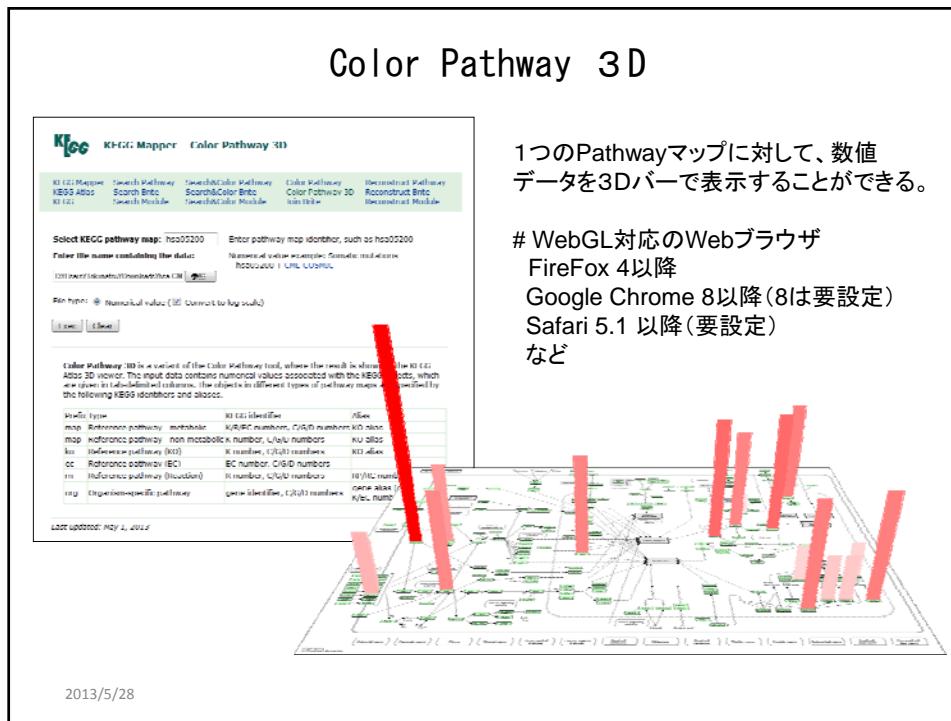
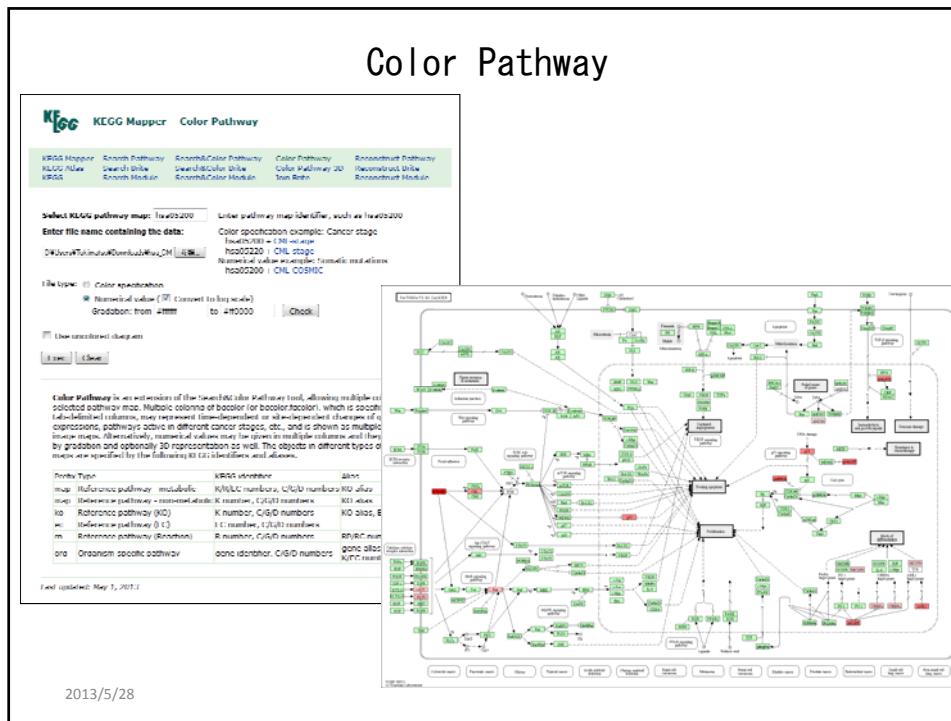
If necessary, change default file type:

Include (global) map Use the standard algorithm Display objects not found in the search Show pathways containing all the objects (AND search)

Search&Color Pathway is an advanced version of the KEGG pathway mapping tool, where genes, objects (genes, proteins, metabolites, enzymes, etc.), maps, and pathways can be mapped onto a pathway. Inputted search objects are mapped in two background and foreground colors (Digital hoodie). The objects in different types of pathways may be specified by the following ID/Path identifier and others.

Path Type	KO ID identifier	Allele
map	Reference pathway, metabolic	KRIBC numbers, C/DB numbers KO alias
map	Reference pathway - received from KEGG, C/DB read/write	KO alias
ko	Reference pathway (KO)	KO number, C/DB numbers
ec	Reference pathway (EC)	EC number, C/DB numbers
me	Reference pathway (reaction)	IPNAME numbers

Last updated: May 20, 2013



そのほかのKEGGのツール群

2013/5/28