

微生物ゲノムクイックスタート

このミニコースでは、微生物のゲノム配列とアノテーションにどのようにアクセスし、どのようにデータを見ればよいか、また遺伝子やアミノ酸配列情報をどのようにダウンロードすればよいかを紹介すると共に、NCBIから提供されている個別ゲノムや比較ゲノム解析ツールについても紹介します。このコースを通して、「水平伝播したと思われる遺伝子をバクテリアゲノム中から見つける」であるとか「病原菌と近縁種の常在菌との間の違いは何ですか?」といった実践的な問題に対応できるようになることが期待されます。

課題1

■ 課題1-1

NCBIホームページから、“Genomic Biology”⇒“Genome Resource : Microbial”とリンクをたどって微生物ゲノムプロジェクト一覧のページへと移ってください。

ゲノム解読が終了した微生物ゲノムは何種類ありますか？

現在解読中の微生物ゲノムは何種類ありますか？その中でアセンブリ中の状態のものは何種類ありますか？

“Firmicutes”に限ると何種のゲノムが解読終了していますか？

■ 課題1-2

解読が終了した微生物ゲノムプロジェクト一覧のページに戻ってください。大腸菌株の箇所までスクロールしてください。

大腸菌では何株のゲノムが完全解読されていますか？

0157堺株を選択して、ゲノムプロジェクトの概要のページに移ってください。

■ 課題1-3

堺株は、ヒトに対して病原性を持っていますか？

病原性株と非病原性株の違いについてゲノム情動的にどのように違うと書いてありますか？

アクセッション番号NC_002695をクリックしてEntrez Genomeの概要ページへと移動してください。

■ 課題1-4

Entrez Genomeの概要ページからRefSeqのゲノム情報や、遺伝子情報、タンパク質一覧テーブルへアクセスしてみてください。

ゲノム上にいくつの遺伝子がアノテーションされていますか？

ゲノム上の位置や、向き、遺伝子名が書かれているタンパク質一覧の詳細ページへ移動して下さい。

構造RNA一覧の詳細ページへ移動して、FASTAファイルを取得してください。

アルギニンに対応したtRNAはいくつありますか？

Entrez Genomeの概要ページへ戻ってください。

■ 課題1-5

続いて、Entrez Genomeの概要ページから“COG”と書かれたリンクへと移動してください。COGとは、NCBIが提供している遺伝子の機能分類データベースであり、各ゲノムのタンパク質コード遺伝子にCOG番号を割り当てることで、遺伝子が持つ機能的な情報も合わせて提供しています。

表中の“% in genome”と“% in genus”の違いに注目してください。何が重要か分かりますか？

表より“Intracellular trafficking and secretion”に分類される遺伝子リストをダウンロードしてください。

Entrez Genomeの概要ページへ戻ってください。

■ 課題1-6

続いて、Entrez Genomeの概要ページから“TaxPlot”と書かれたリンクへと移動してください。TaxPlotでは現在見ている生物種の全遺伝子（染色体、プラスミドすべて）を他の二生物種全遺伝子と比較し、その結果を相同性に基づいてプロットした情報を提供しています。

病原性大腸菌である0157の特徴を調べるために、大腸菌K12株と他のH157株との比較を行ってみてください。

プルダウンメニューより、カテゴリを選択することで結果に対して機能的な分類を行ってみてください。（現在この機能は提供されていません。）

“Translation and Cell envelope biogenesis”や“Cell motility”に分類された遺伝子の分布を比較してみてください。（現在この機能は提供されていません。）

何が分かりましたか

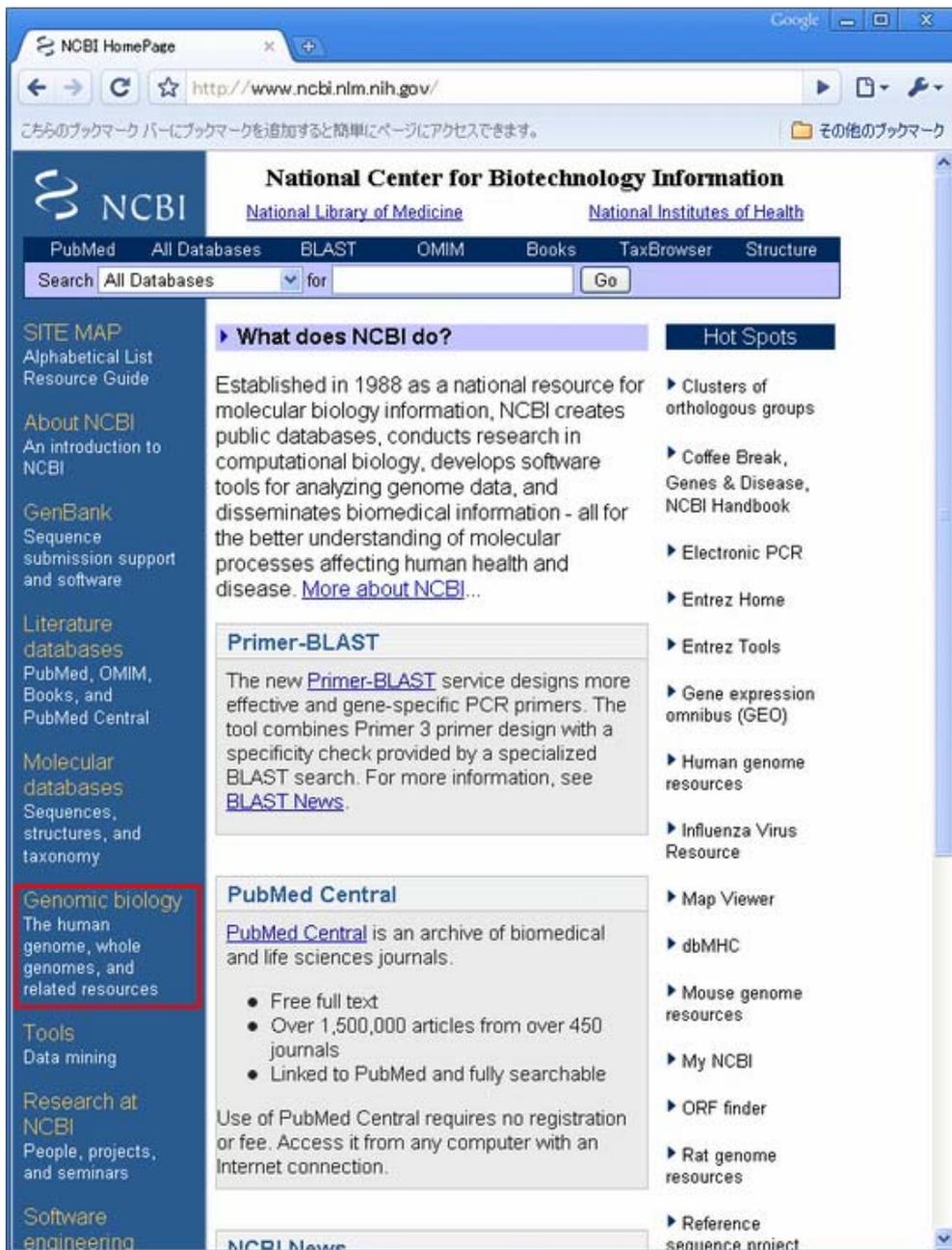
Entrez Genomeの概要ページへ戻ってください。

解答・解説

■ 解答1-1

[微生物ゲノムプロジェクト一覧ページの立ち上げ](#)

NCBIのトップページから画面左にある“Genomic biology”をクリックし



画面右上の“Genome Resources”から“Microbial”をクリックして、微生物ゲノムプロジェクト一覧ページへと移動してください。

Genome biology

http://www.ncbi.nlm.nih.gov/Genomes/

Genomic Biology

PubMed All Databases BLAST OMIM Books TaxBrowser Structure

Search All Databases for Go

NCBI

Site Map
guide to NCBI resources

Cancer
Chromosomes
chromosomal abnormalities

Clusters of
Orthologous Groups
analysis of complete genomes

Gene
gene-related information

Genome
complete genome sequences

GEO
gene expression data

HomoloGene
orthologs between pairs of organisms

Map Viewer
map and genome displays

RefSeq
the reference sequence project

Software
for genetic analysis

Genomic Biology

NCBI provides several genomic biology tools and resources, including organism-specific pages that include links to many web sites and databases relevant to that species. We invite you to explore the links provided on this page.

Assembly and Annotation Information

- [The Genome Reference Consortium \(GRC\)](#) NEW
- [AGP Resources](#)
- [Annotation Information](#)
- [Assembly Information](#)
- [Genome Glossary](#)
- [NCBI Handbook Chapter 14: Genome Assembly and Annotation Process](#)

Announcements

Map Viewer - genome annotation updates:

Species	Build	Map Viewer Release
Physcomitrella patens (moss)	1.1	January 8, 2009
Caenorhabditis elegans (nematode)	WS190	October 10, 2008
Anopheles gambiae (mosquito)	AgamP3.3	October 10, 2008
Arabidopsis thaliana (mouse-ear cress)	8.1	September 29, 2008
Ciona intestinalis	1.1	September 16, 2008
Acyrthosiphon pisum (pea aphid)	1.1	August 6, 2008
Bos taurus (cow)	Btau 4.0	August 5, 2008

Genome Resources

- ▶ [Entrez Genome](#)
- ▶ [Fungal Genomes Central](#)
- ▶ [Genome Projects Database](#)
 - [Eukaryotic](#)
 - [Fungi](#)
 - [Insects](#)
 - [Mammals](#)
 - [Microbial](#)
 - [Plants](#)
- ▶ [Map Viewer](#)
- ▶ [Organelles](#)
- ▶ [Plant Genomes Central](#)
- ▶ [Viral Resources](#)
 - [Influenza Virus Resource](#)
 - [Retroviruses](#)
 - [Viral Genomes](#)

Organism-Specific

- ▶ [Aphid](#) NEW
- ▶ [Arabidopsis](#)
- ▶ [Aspergillus](#)
- ▶ [Bee](#)
- ▶ [Beetle](#)
- ▶ [Buffalo](#)
- ▶ [Cat](#)
- ▶ [Chicken](#)

Legend: G Genome Resources, B BLAST, M Map Viewer, P Genome Project DB

Complete Microbial Genomes

http://www.ncbi.nlm.nih.gov/genomes/entrez.cgi

NCBI ENTREZ Genome Project

Search: Genome Project

Organism info Complete genomes Genomes in progress

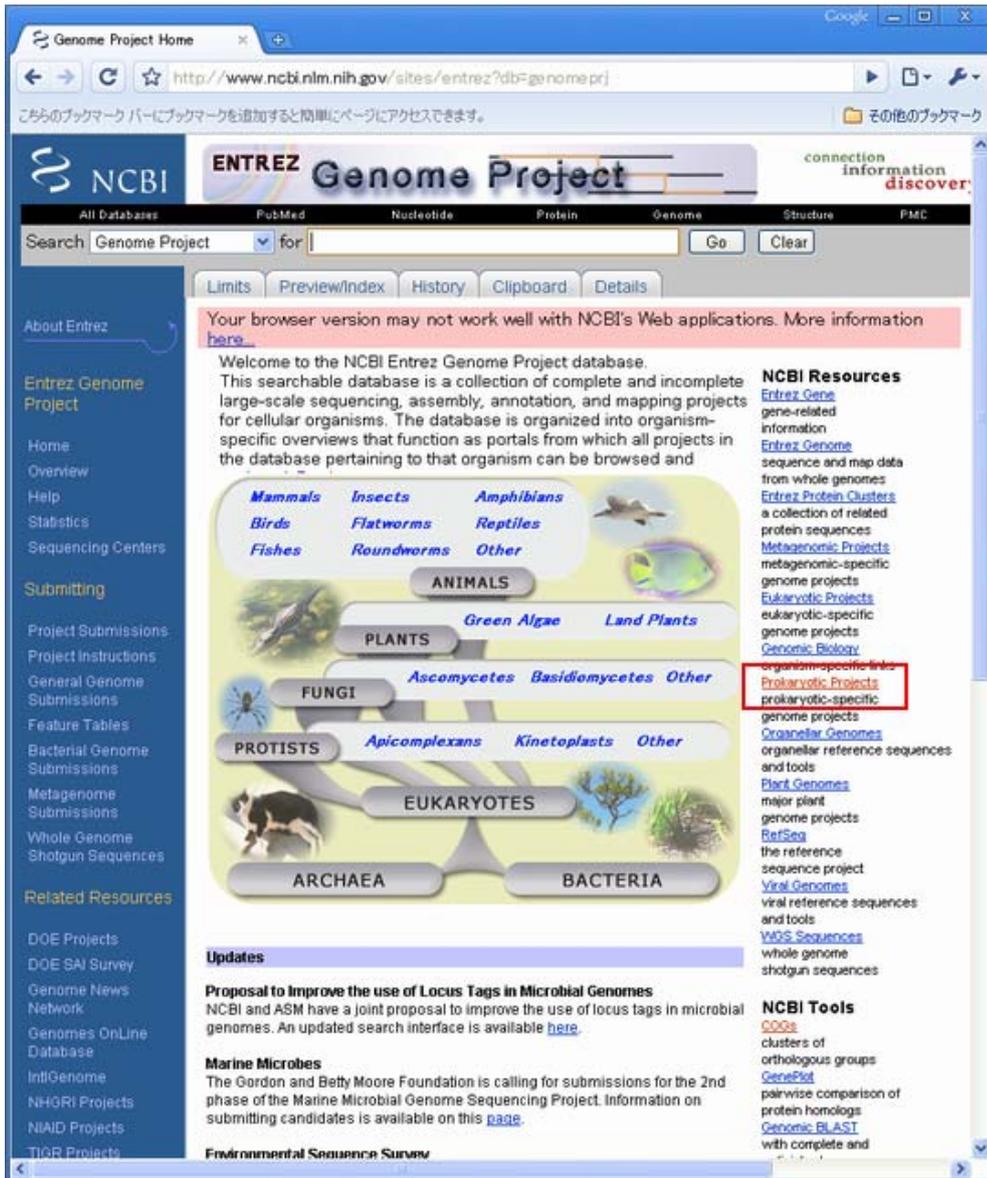
organism group: -- All --

Tools legend: T - TaxMap; P - ProtTable; G - COG Table; D - 3-D neighbors; L - BLAST; S - COG search; G - GenePlot; X - TaxPlot; M - aMap; F - FTP; R - Publications.
* size is estimated, otherwise genome size is calculated based on existing sequences

821 Complete Microbial Genomes selected: [A] - 55, [B] - 766

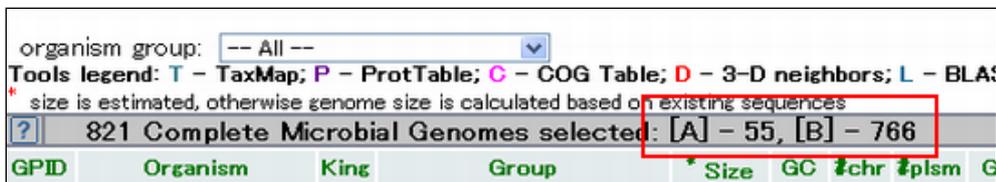
GenBank	Organism	King	Group	Size	GC	#chr	#plasm	GenBank	RefSeq	Released	Modified	Center	Tools
12001	<i>Anaerobaculum mobile</i> (DSMZ 11011)	B	Cyanobacteria	8.36	47.0	1	9	CP000828	NC_009025.1	10/16/07	07/25/08	Washington University (WashU) (more)	I E D L S G X E R
10060	<i>Acholeplasma laidlawii</i> (ATCC 35064)	B	Firmicutes	1.5	31.9	1		CP000896.1	NC_010160.1	12/07/07	07/30/08	Research Institute for Innovation in Chemical Medicine, Federal Agency of Public Health and Social Development of the Russian Federation (more)	I E D L S G X E
15720	<i>Acidithiobacillus ferrooxidans</i> ATCC 29621	B	Alphaproteobacteria	3.97	67.1	1	8	CP000927.1	NC_009484.1	05/11/07	07/26/08	DOE Joint Genome Institute	I E D L S G X E
53	<i>Acidithiobacillus ferrooxidans</i> ATCC 29621	B	Gammaproteobacteria	3	58.8	1		CP001219.1	NC_011761.1	12/19/08	12/22/08	TIGR	E L S G X M E R
16695	<i>Acidithiobacillus ferrooxidans</i> ATCC 29621	B	Other Bacteria	2.9	58.9	1		CP001192.1	NC_011206.1	09/05/08	09/08/08	DOE Joint Genome Institute	I E D L S G X M E
15771	<i>Acidithiobacillus bacterium</i> F6rQ45	B	Acidobacteria	5.7	58.4	1		CP000660.1	NC_008009.1	05/04/06	07/20/08	DOE Joint Genome Institute	I E D L S G X E
16007	<i>Acidithermus caldophilus</i> 119	B	Actinobacteria	2.4	66.9	1		CP000481.1	NC_008570.1	11/09/06	07/29/08	DOE Joint Genome Institute	I E D L S G X E
15700	<i>Acidithiobacillus ferrooxidans</i> ATCC 29621	B	Betaproteobacteria	5.4	60.5	1		CP000512.1	NC_008752.1	01/04/07	07/29/08	DOE Joint Genome Institute	I E D L S G X M E
15685	<i>Acidithiobacillus</i> sp. J542	B	Betaproteobacteria	4.54	66.1	1	2	CP000539.1	NC_008782.1	01/04/07	07/21/08	DOE Joint Genome Institute (more)	I E D L S G X M E
21111	<i>Acidithiobacillus</i> sp. AD0067	B	Gammaproteobacteria	4.11	39.2	1	1	CP001182.1	NC_011586.1	11/17/08	11/17/08	Cornell University (more)	E L S G X M E R
30069	<i>Acidithiobacillus</i> sp. AD007-0004	B	Gammaproteobacteria	3.8	39.0	1		CP001172.1	NC_011585.1	11/24/08	12/11/08	SRM (Salk Institute) (more)	E L S G X M E R
17827	<i>Acidithiobacillus</i> sp. AD007-0004	B	Gammaproteobacteria	3.99	39.9	1	2	CP000863.1	NC_010611.1	04/15/08	11/12/08	Institute of Molecular Biology and Genetics, VIB (more)	I E D L S G X M E R
17472	<i>Acidithiobacillus</i> sp. ATCC 17928	B	Gammaproteobacteria	4.02	38.9	1	2	CP000507.1	NC_009005.1	03/01/07	07/29/08	Yale Univ.	I E D L S G X M E R
28921	<i>Acidithiobacillus</i> sp. ATCC 17928	B	Gammaproteobacteria	3.99	39.3	1	4	CU89141.1	NC_010410.1	02/29/08	07/30/08	Genoscope	I E D L S G X M E R

このページには、Entrez Genome Projectのトップページから右側の“Prokaryotic Projects”をクリックすることでも移動することが可能です。

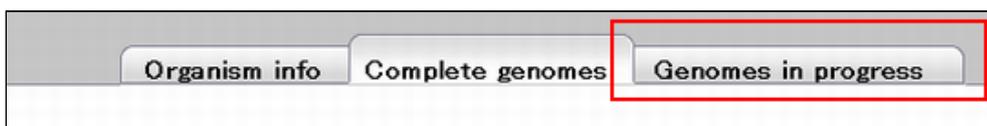


ゲノムプロジェクトの一覧ページでは、1行につき1微生物ゲノムプロジェクトが記載されており、ゲノムサイズやGC含量アクセッション番号とそのリンクや、遺伝子リストへのリンクなどが表示されています。

表の一番上の記載からArchaea55ゲノム、Bacteria766ゲノムの合計821ゲノムについて完成配列が出ていることがわかります。



また、表右上から“Genomes in Progress”タブをクリックすると



現在解説中のゲノムプロジェクトの一覧ページへと移動します。

Microbial Genome Sequencing

http://www.ncbi.nlm.nih.gov/genomes/grpka.cgi

NCBI ENTREZ Genome Project

Search: Genome Project

Organism info Complete genomes Genomes in progress

organism group: All

Legend: wgs assembly; no sequence available.

sequencing availability filter: all; wgs assembly; no sequence available

1401 Microbial Genomes In-Progress selected: [A] - 41, [B] - 1360

GenID	Organism	King	Group	Conts	CDS	Size	GC	Accession	Blast	Released	Center
30802	Nitococcus anitae DSM 10331	II	Cyanobacteria	--	--	--	--	--	--	--	DOE Joint Genome Institute
16702	Acetobacterium aceti DSM 5410	II	Cyanobacteria	--	--	--	--	--	--	--	J. Craig Venter Institute
92799	Acetobacterium arabaticum DSM 5501	II	Firmicutes	--	--	--	--	--	--	--	DOE Joint Genome Institute
92215	Acetivibrio sp. ATCC 49095	II	Firmicutes	--	--	--	--	--	--	--	University of Valencia
192	Acidithiobacillus thiooxidans DSM 10331	A	Crenarchaeota	--	--	1.8	--	--	--	--	University of Copenhagen
28525	Acidithiobacillus ferrooxidans DSM 10331	II	Actinobacteria	--	--	67	69	--	--	--	DOE Joint Genome Institute
1	Acidobacterium capsulatum ATCC 51196	II	Acidobacteria	--	--	4.15	--	--	--	--	ITIG
28085	Acidobacterium capsulatum ATCC 51196	II	Acidobacteria	--	--	--	--	--	--	--	J. Craig Venter Institute
32000	Acidobacterium delfieldii DSM	II	Betaproteobacteria	--	--	65	66	--	--	--	DOE Joint Genome Institute
19272	Aciduliferobacterium boonei T-169	A	Euryarchaeota	63	--	39.1	--	NC_085000000000	BLAST	07/15/08	J. Craig Venter Institute
13000	Acinetobacter baumannii	II	Gammaproteobacteria	--	--	3.9	--	--	--	--	Genoscreen
30995	Acinetobacter baumannii AB900	II	Gammaproteobacteria	60	--	38.9	--	NC_080000000000	BLAST	12/05/08	Infectious Disease and Genomics Laboratory, Center of Excellence in Biotechnology and Life Sciences, SUNY-Duffalo
20333	Acinetobacter baumannii MCR-2_090	II	Gammaproteobacteria	--	--	--	--	--	--	--	Zhejiang University, China
29639	Acinetobacter baumannii HSA0098	II	Gammaproteobacteria	--	--	38.9	--	--	--	--	Naval Medical Research Center
29635	Acinetobacter baumannii HSL7215	II	Gammaproteobacteria	--	--	38.9	--	--	--	--	Naval Medical Research Center
29633	Acinetobacter baumannii HSL7292	II	Gammaproteobacteria	--	--	38.9	--	--	--	--	Naval Medical Research Center
30949	Acinetobacter sp. ATCC 27284	II	Gammaproteobacteria	255	--	39.4	--	NC_080000000000	BLAST	10/23/08	Baylor College of Medicine
6117	Actinobacillus pleuropneumoniae serovar 1 str. 6074	II	Gammaproteobacteria	--	--	2.19	37	--	--	--	University of Oklahoma
196	Actinobacillus pleuropneumoniae serovar 1 str. 6074	II	Gammaproteobacteria	140	2132	2.29	41.4	NC_080000000000	BLAST	04/21/03	University of Oklahoma
202	Actinobacillus pleuropneumoniae serovar 5	II	Gammaproteobacteria	--	--	2.18	37	--	--	--	University of Oklahoma
332	Actinobacillus pleuropneumoniae serovar 7	II	Gammaproteobacteria	--	--	2.08	37	--	--	--	University of Oklahoma
13510	Actinomyces israelii 1401	II	Actinobacteria	--	--	3.06	--	--	--	--	ITIG
18171	Actinomyces odontolyticus ATCC 17982	II	Actinobacteria	4	2159	65.4	--	NC_080000000000	BLAST	03/16/07	Washington University (WashU)
31620	Actinomyces viscosus DSM 15424	II	Actinobacteria	--	--	--	--	--	--	--	BCM
21072	Actinomyces robitiae DSM 43927	II	Actinobacteria	--	--	70.8	--	--	--	--	DOE Joint Genome Institute
19795	Actinomyces minutus DSM 43927	II	Actinobacteria	--	--	--	--	--	--	--	DOE Joint Genome Institute
32178	Actinomyces actinomycetensiformis H1163	II	Gammaproteobacteria	--	--	--	--	--	--	--	University of Oklahoma

このページの上部を見ると、現在解読中のゲノムはArchaea41ゲノム、Bacteria1360ゲノムの合計1401ゲノムであることがわかります。

Genome size is estimated, otherwise genome size is calculated based on existing sequences

1401 Microbial Genomes In-Progress selected: [A] - 41, [B] - 1360

Organism King Group Conts CDS

また、画面右上のラジオボタンから“wgs assembly”フィルタを選択することで、

Legend: wgs assembly; no sequence available.

sequencing availability filter: all; wgs assembly; no sequence available.

save

現在ドラフト配列が入手できるゲノムのみ絞り込まれます。その数は、Archaea8ゲノム、Bacteria608ゲノムの合計616ゲノムになります。

Genome size is estimated, otherwise genome size is calculated based on existing sequences

616 Microbial Genomes In-Progress selected: [A] - 8, [B] - 608

Organism King Group Conts CDS

プロジェクト一覧の絞り込み

右上のタブから“Complete Genomes”をクリックして、解読完了ゲノムの一覧に戻ってください。

Organism info Complete genomes Genomes in progress

今度は左上の“organism group:”横のプルダウンメニューから“Firmicutes”を選択して、表示されるゲノムをFirmicutesのものに限ってください。

organism group: -- All --

Tools legend: T - TaxMap; P - ProtTable; C - COG Table; D - 3-D neighbors; L - BLAST

size is estimated, otherwise genome size is calculated based on existing sequences

767 Complete Microbial Genomes selected: [A] - 52, [B] - 715

GPID	Organism	King	Group	* Size	GC	#chr	#plsm	GenBa
12997	Acaryochloris MBIC11017		bacteria	8.36	47.0	1	9	CP0008
19259	Acholeplasma PG-8A		utes	1.5	31.9	1		CP0008
15753	Acidiphilium or 5		obacteria	3.97	67.1	1	8	CP0008
16689	Acidithiobacillus ferrooxidans A		acteria	2.9	58.9	1		CP0011
	Acidobacteria bacterium							

164ゲノムがFirmicutesで解読が終了していることが分かります。

organism group: Firmicutes

Tools legend: T - TaxMap; P - ProtTable; C - COG Table; D - 3-D neighbors; L - BLAST

size is estimated, otherwise genome size is calculated based on existing sequences

164 Complete Microbial Genomes selected: [A] - 0, [B] - 164

GPID	Organism	King	Group	* Size	GC	#chr	#plsm	GenBa
------	----------	------	-------	--------	----	------	-------	-------

■ 解答1-2

[ゲノムプロジェクトの概要ページの立ち上げ](#)

organism group:横のプルダウンメニューからAllを選択して解読が終了している全ゲノムを表示してください。

organism group: -- All --

Tools legend: T - TaxMap; P - ProtTable; C - COG Table; D - 3-D neighbors; L - BLAST

size is estimated, otherwise genome size is calculated based on existing sequences

767 Complete Microbial Genomes selected: [A] - 52, [B] - 715

GPID	Organism	King	Group	* Size	GC	#chr	#plsm	GenBa
12997	Acaryochloris MBIC11017		bacteria	8.36	47.0	1	9	CP0008
19259	Acholeplasma PG-8A		utes	1.5	31.9	1		CP0008
15753	Acidiphilium or 5		obacteria	3.97	67.1	1	8	CP0008
16689	Acidithiobacillus ferrooxidans A		acteria	2.9	58.9	1		CP0011
	Acidobacteria bacterium							

画面をスクロールして、大腸菌 (Escherichia coli) を見てみましょう。

13480	Erythrobacter litoralis HTCC2594	B	Alphaproteobacteria	3.05	63.1	1	
16235	Escherichia coli 536	B	Gammaproteobacteria	4.9	50.5	1	
33413	Escherichia coli 55989	B	Gammaproteobacteria	5.2	50.7	1	
16718	Escherichia coli APEC O1	B	Gammaproteobacteria	5.51	50.3	1	2
18083	Escherichia coli ATCC 8739	B	Gammaproteobacteria	4.7	50.9	1	
313	Escherichia coli OFT073	B	Gammaproteobacteria	5.2	50.5	1	
13960	Escherichia coli E24377A	B	Gammaproteobacteria	5.27	50.6	1	6
33409	Escherichia coli ED1a	B	Gammaproteobacteria	5.2	50.7	1	
13959	Escherichia coli HS	B	Gammaproteobacteria	4.6	50.8	1	
33373	Escherichia coli IAI1	B	Gammaproteobacteria	4.7	50.8	1	
33411	Escherichia coli IAI39	B	Gammaproteobacteria	5.1	50.6	1	
32571	Escherichia coli O127:H6 str. E2348/69	B	Gammaproteobacteria	5.1	50.5	1	2
259	Escherichia coli O157:H7 EDL933	B	Gammaproteobacteria	5.59	50.3	1	1
27739	Escherichia coli O157:H7 str. EC4115	B	Gammaproteobacteria	5.73	50.4	1	2
226	Escherichia coli O157:H7 str. Sakai	B	Gammaproteobacteria	5.6	50.5	1	2
33375	Escherichia coli S88	B	Gammaproteobacteria	5	50.7	1	
18057	Escherichia coli SE11	B	Gammaproteobacteria	5.17	50.7	1	6
19469	Escherichia coli SMS-3-5	B	Gammaproteobacteria	5.25	50.5	1	4
33415	Escherichia coli UMN026	B	Gammaproteobacteria	5.2	50.7	1	
16259	Escherichia coli UTI89	B	Gammaproteobacteria	5.21	50.6	1	1
20079	Escherichia coli str. K-12 substr. DH10B	B	Gammaproteobacteria	4.7	50.8	1	
225	Escherichia coli str. K-12 substr. MG1655	B	Gammaproteobacteria	4.6	50.8	1	
16351	Escherichia coli str. K-12 substr. W3110	B	Gammaproteobacteria	* 4.65	50.8	1	
33369	Escherichia fergusonii ATCC 35469	B	Gammaproteobacteria	4.6	49.9	1	
10649	Exiguobacterium sibiricum 255-15	B	Firmicutes	3.01	47.7	1	2

図中赤四角で囲んだ23ゲノムが解読されていることがわかります。

この中から青四角で囲んだO157 堺株のリンクをクリックして、ゲノムプロジェクトの概要のページに移ってみましょう。

Genome Project > *Escherichia coli* > *Escherichia coli* O157:H7 str. Sakai O157:H7 project at GIRC

Resource Links: Enterohemorrhagic *Escherichia coli*. Project data

NCBI Resources: BLAST, genome, FTP, TaxPlot

Organism data in GenBank: Genomic, mRNA, Protein, WGS

Sequencing Centers: Osaka Univ, Sanger, U. Wash

Related Resources: ADAP, BDB, DRS, COSC, Colibri, CyberCell, E. coli Ontology, E. coli TF, E. coli WMI, ECCE, ECDC, ECORDEC, EPD, ERIC, Extrablast

Lineage: Bacteria, Proteobacteria, Gammaproteobacteria, Enterobacteriales, Enterobacteriaceae, *Escherichia*, *Escherichia coli* O157:H7 str. Sakai

Photo: CDC

Genome Projects: Plasmid genome: *Escherichia coli* O157 at Molecular Biology, Institute of Biochemistry and Physiology of Microorganisms, IBPM RAS, RUSSIA; Institute of Biochemistry and Physiology of Microorganisms, Molecular Biology, Russia, Pushchino; *Escherichia coli* at University of Bristol, Department of Pathology and Microbiology, UK

Genome sequencing: *Escherichia coli* O82 at Sanger Institute [In progress]; *Escherichia coli* O157 at JGI [Assembly]; *Escherichia coli* S88 at University of Goettingen [Complete]; *Escherichia coli* O157 at JGI [Assembly]; *Escherichia coli* APEC O1 at Iowa State University [Complete]; *Escherichia coli* ATCC 35067 at US DOE Joint Genome Institute (JGI-PGF) [Complete]

Expand to show the list of Centers: *Escherichia coli* B str. O157:H7 at International E. coli B Consortium [In progress]; *Escherichia coli* O157 at JGI [Assembly]; *Escherichia coli* O157 at JGI [Assembly]; *Escherichia coli* O157:H7 at DOE Joint Genome Institute [In progress]; *Escherichia coli* O157:H7 at Korea Research Institute of Bioscience and Biotechnology [In progress]

Genome information:

Name	RefSeq	GenBank	Publications	Length (kb)	GC content	Proteins	RNAs	TestMap	CDD	COG
Chromosomal	NC_002955	BA000007	2	5.5	50.5%	3253	141	✓	✓	✓
Plasmid pO157	NC_002128	AB011549	1	0.092721	47.6%	85	--			
Plasmid pO54K1	NC_002127	AB011548	1	0.003306	43.4%	3	--			

Publications: [Thangtham T et al. "Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genomic comparison with a laboratory strain K-12." *DNA Res* 2001 Feb 28;8\(1\):11-22](#)

***Escherichia coli* O157:H7 str. Sakai O157:H7**

Escherichia coli This organism was named for its discoverer, Theodore Escherich, and is one of the premier model organisms used in the study of bacterial genetics, physiology, and biochemistry. This enteric organism is typically present in the lower intestine of humans, where it is the dominant facultative anaerobe present, but it is only one minor constituent of the complete intestinal microflora. *E. coli* is capable of causing various diseases in its host, especially when they acquire virulence traits. Strains of *E. coli* can cause urinary tract infections, neonatal meningitis, and many different intestinal diseases, usually by attaching to the host cell and introducing toxins that disrupt normal cellular processes. Virulence proteins may be encoded on extrachromosomal plasmids or within bacteriophages and distinct DNA segments termed pathogenicity islands (PAIs). PAIs are likely to have been transferred horizontally and may even have integrated into the chromosome through bacteriophage or plasmid integration or transposition.

Escherichia coli O157:H7 This strain is associated with Hamburger disease, which is caused by the contamination of meat products by enterohemorrhagic *E. coli* (EHEC). The identifier O157:H7 refers to the serotype of EHEC, and reflects the specific antigenic markers found on the surface of the cell. EHEC attaches and effaces to cells in the large intestine.

There are numerous differences that distinguish O157:H7 from K-12, hundreds of them associated with genomic islands in either strain, including at least 9 large PAIs in O157:H7 that encode virulence factors. A type III secretion system, the locus of enterocyte effacement, numerous toxins and adhesins, as well as fimbrial gene clusters and iron uptake systems are found in these PAIs. Pathogenicity genes are also found on the plasmid pO157.

Escherichia coli O157:H7 strain Sakai This strain of O157:H7 was isolated in a 1997 outbreak in Sakai, Japan.

Cellular features				Environment			Temperature		
Gram stain	Shape	Arrangement	Endospores	Motility	Salinity	Oxygen Req.	Habitat	Opt. temp.	Range
-	Rod	Singles, Pairs		Yes		Facultative	Host-associated	37C	Mesophilic

Pathogenic in: Human Disease: Hemorrhagic colitis

上図のような概要と、様々な情報へのリンクが記されたページが立ち上がります。

■ 解答1-3

[ゲノムプロジェクト概要ページ](#)

このO157:H7 str. Sakaiのゲノムに関する概要ページを読んでいきましょう。一番下の赤四角で囲んだ部分に“Pathogenic in:Human”と書いてあることから、このゲノム解読株はヒトに対して病原性を有していることがわかります。

***Escherichia coli* O157:H7 str. Sakai O157:H7**

Escherichia coli This organism was named for its discoverer, Theodore Escherich, and is one of the premier model organisms used in the study of bacterial genetics, physiology, and biochemistry. This enteric organism is typically present in the lower intestine of humans, where it is the dominant facultative anaerobe present, but it is only one minor constituent of the complete intestinal microflora. *E. coli* is capable of causing various diseases in its host, especially when they acquire virulence traits. Strains of *E. coli* can cause urinary tract infections, neonatal meningitis, and many different intestinal diseases, usually by attaching to the host cell and introducing toxins that disrupt normal cellular processes. Virulence proteins may be encoded on extrachromosomal plasmids or within bacteriophages and distinct DNA segments termed pathogenicity islands (PAIs). PAIs are likely to have been transferred horizontally and may even have integrated into the chromosome through bacteriophage or plasmid integration or transposition.

Escherichia coli O157:H7 This strain is associated with Hamburger disease, which is caused by the contamination of meat products by enterohemorrhagic *E. coli* (EHEC). The identifier O157:H7 refers to the serotype of EHEC, and reflects the specific antigenic markers found on the surface of the cell. EHEC attaches and effaces to cells in the large intestine.

There are numerous differences that distinguish O157:H7 from K-12, hundreds of them associated with genomic islands in either strain, including at least 9 large PAIs in O157:H7 that encode virulence factors. A type III secretion system, the locus of enterocyte effacement, numerous toxins and adhesins, as well as fimbrial gene clusters and iron uptake systems are found in these PAIs. Pathogenicity genes are also found on the plasmid pO157.

Escherichia coli O157:H7 strain Sakai This strain of O157:H7 was isolated in a 1997 outbreak in Sakai, Japan.

Cellular features				Environment			Temperature		
Gram stain	Shape	Arrangement	Endospores	Motility	Salinity	Oxygen Req.	Habitat	Opt. temp.	Range
-	Rod	Singles, Pairs		Yes		Facultative	Host-associated	37C	Mesophilic

Pathogenic in: Human Disease: Hemorrhagic colitis

また、その上の青く囲んだ領域に0157など病原性株と非病原性である大腸菌K12のゲノム情報的な違いが述べられています。病原性に関連する遺伝子群は、外来性のプラスミド、バクテリオファージなどによりゲノム中に取り込まれ、周辺配列と明らかに異なる特徴を持った“pathogenicity islands (PAIs) (病原性遺伝子クラスター)”を形成しています。このPAIs内には、III型分泌系遺伝子群、LEEクラスター (Locus of Enterocyte Effacement)、多くの毒素関連、接着関連因子、鞭毛関連遺伝子、鉄分取り込み遺伝子などが含まれていることが書かれています。

Genome information:						
Name	RefSeq	GenBank	Publications	Length (Mbp)	GC content	Proteins
Chromosome	NC_002695	BA000007	4	5.5	50.5%	5253
Plasmid pO157	NC_002128	AB011549	1	0.092721	47.6%	85
Plasmid pOSAK1	NC_002127	AB011548	1	0.003306	43.4%	3

Publications:

- [Havashi T et al.](#), "Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and laboratory strain K-12.", *DNA Res*, 2001 Feb 28;8(1):11-22

“Genome information”に書かれたリンクから染色体ゲノムのRefSeqである、NC_002695をクリックして “Entrez Genome”の概要ページへ移動して下さい。

Genome > Bacteria > Escherichia coli O157:H7 str. Sakai, complete genome

Lineage: [Bacteria](#); [Proteobacteria](#); [Gammaproteobacteria](#); [Enterobacteriales](#); [Enterobacteriaceae](#); [Escherichia](#); [Escherichia coli](#); [Escherichia coli O157:H7](#); [Escherichia coli O157:H7 str. Sakai](#)

Chromosomes: [genome](#)
Plasmids: [pOSAK1](#), [pO157](#)

Genome Info:	Features:	BLAST homologs:	Links:	Review Info:
Refseq: NC_002695	Genes: 5372	COG	Genome Project	Publications: [4]
GenBank: BA000007	Protein coding: 5230	TaxMap	Refseq FTP	Refseq Status: Provisional
Length: 5,498,450 nt	Structural RNAs: 141	TaxPlot	GenBank FTP	Seq Status: Completed
GC Content: 50%	Pseudo genes: None	GenePlot	BLAST	Sequencing center: GIRC
% Coding: 85%	Others: 170	gMap	TraceAssembly	Completed: 2001/10/02
Topology: circular	Contigs: 1		COG	Organism Group
Molecule: DNA			Other genomes for species: 115	

Gene Classification based on [COG functional categories](#) Search gene, GeneID or locus_tag:

Zoom: 1 nt to 9,910 nt

Click [here](#) for Sequence Viewer presentation (base sequence and aligned amino acids) of selected region

■ 解答1-4

[ゲノム概要のページ](#)

表示された“Entrez Genome”概要のページでは、このゲノムに関連した様々な情報が表示されているとともに遺伝子情報などへとリンクが張られています。

Escherichia coli O157:H7 str. Sakai, complete genome

Position: from to Length: from to

Length histogram.
Click on a bar to select length range.

1 2905

141 RNA(s) shown
Legends: ◆ DNA region in flatfile format ◆ DNA region in FASTA format

Product Name	Start	End	Strand	Length	GeneID	Locus	Locus_tag	Links
16S ribosomal RNA	227102	228643	+	1542	913977	rrsH	ECs5366	◆ ◆
Ile tRNA	228712	228788	+	77	913984	ileV	ECs5367	◆ ◆
Ala tRNA	228831	228906	+	76	913988	alaV	ECs5368	◆ ◆
23S ribosomal RNA	229090	231992	+	2903	913990	rrlH	ECs5369	◆ ◆
5S ribosomal RNA	232085	232204	+	120	914016	rrfH	ECs5370	◆ ◆
Asp tRNA	232257	232333	+	77	914020	aspU	ECs5371	◆ ◆
Asp tRNA	240481	240557	+	77	914060	aspV	ECs5372	◆ ◆
Thr tRNA	299982	300057	+	76	914369	thrW	ECs5373	◆ ◆
misc_RNA	542513	542626	+	114	914612	ffs	ECs5378	◆ ◆
Arg tRNA	657182	657258	+	77	916955	argU	ECs5379	◆ ◆
Gln tRNA	776591	776665	-	75	917064	glnX	ECs5384	◆ ◆
Gln tRNA	776703	776777	-	75	917065	glnV	ECs5385	◆ ◆
Met tRNA	776826	776902	-	77	917066	metU	ECs5386	◆ ◆

一番右のLinksと書かれた箇所のうち黄色のダイヤモンド印（青四角で囲まれた箇所）をクリックして見てください。FASTA形式で配列が表示されます。

1: [NC_002695](#). Reports Escherichia coli ...[gi:15829254]

```
>gi|15829254:227102-228643 Escherichia coli O157:H7 str. Sakai, complete genome
AAATTGAAGAGTTTGATCATGGCTCAGATTGAACGCTGGCGGACGGCCTAACACATGCAAGTCGAACGGT
AACAGGAAAGAAAGCTTGCTCTTTGCTGACGAGTGGGGACGGGTGAGTAATGTCTGGGAAACTGCCTGAT
GGAGAGGGATAAATACTGGAAACGGTACGCTAATACCGCATAACGTCGCAAGACCAAGAGGGGGACCTTC
GGCCCTCTGCCATCGGATGTGCCAGATGGGATTAGCTAGTAGGTGGGTAAACGGCTCACCTAGGGGAC
GATCCCTAGCTGGTCTGAGAGGATGACCAAGCCACACTGGAACTGAGACACGGTCCAGACTOCTACGGGAG
GCAGCAGTGGGGAATATTGCACAATGGCGCAAGCCTGATGCAGCCATGCCGGTGTATGAAAGAGGCCT
TCGGGTTGTAAGTACTTTACGGGGGAGGAAAGGAGTAAAGTTAATACCTTTGCTCATTGACGTACCC
GCAGAAGAAGCACCGGCTAACTCGTGCACGACGCGCGGTAAACGGAGGGTGAAGCGTTAATCGGAA
TTACTGGCGTAAAGCGCAAGCAGGCGGTTTGTAAAGTCAGATGTGAAATCCCGGGCTCAACCTGGGAA
TGCATCTGATACTGGCAAGCTTGAGTCTCGTAGAGGGGGTGAATTCAGGTGTAGCGGTGAAATGCG
TAGAGATCTGGAGGAATACCGGTGGCGAAAGCGGCCCTTGAAGAAGACTGACGCTCAGGTGCGAAAGC
GTGGGAGCAAAACAGGATTAGTACCCCTGGTGTGACAGCGGTAAAAGATGTGACTTGGAGGTTGTGCC
CTTGAGGCGTGGCTCCCGAGCTAACCGGTTAAGTCGACCGCCTGGGGAGTACGGCCGCAAGGTTAAAAC
TCAAAATGAATTGACGGGGGCCGCAAGCGGTGGAGCATGTGGTTTAAATCGATGCAACCGCAAGAACCC
TTACCTGGTCTTGACATCCACAGAACTTCCAGAGATGGATTGGTGCCTTCGGGAACGTGAGACAGGTG
CTGCATGGCTGTCTGACGCTCGTGTGTGAAAATGTGGGTAAAGTCCCGCAACGAGCGCAACCCCTATCC
TTTGTGCCAGCGGTCCGGCCGGAACTCAAAAGGAGACTGCCAGTGAATAACTGGAGGAAGGTGGGATG
ACGTCAAGTCATCATGGCCCTTACGACCAAGGCTACACAGTGTCAATGGCGCATACAAAGAGAAAGCG
ACCTCGGAGAGCAAGCGGACCTCATAAAGTGGTGTAGTCCGATTGGAGTCTGCAACTCGACTCCAT
GAAGTCGGAATCGCTAGTAACTGGTCAAGATGCCAGGTGAATACGTTCCCGGGCCTTGTACACACC
GCCCGTCAACCATGGGAGTGGGTTCGAAAAGAAAGTGGTACGTTAACTTCCGTTCCGGAGGGCGCTTACCACT
TTGTGATTCATGACTGGGTAAGTCGTAACAAGGTAACCGTAGGGGAACTGCGGTTGGATCACCTCCT
TA
```

構造RNAの一覧表をもう一度見てみましょう。その中でArg tRNAと書かれているRNA遺伝子はいくつあるでしょうか。上の一覧表で赤く囲んだ遺伝子など合計18個あることがわかります。

■ 解答1-5

[COGテーブルの見方](#)

Backボタンを使って“Entrez Genome”概要のページへ戻ってください。

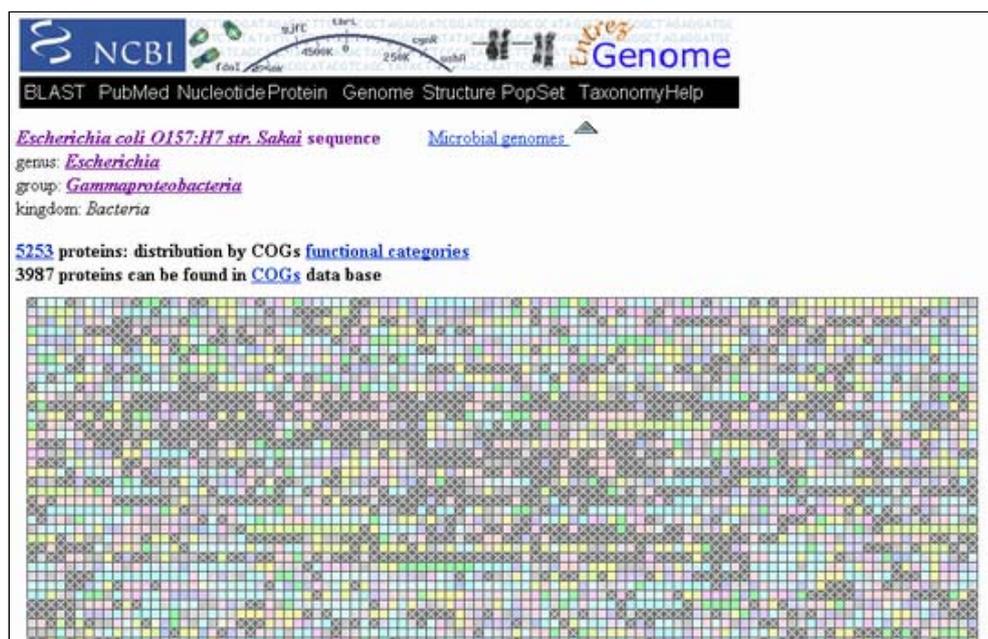
その表の中からCOGと書かれたリンクをクリックしてください。

Genome Info:	Features:	BLAST homologs:	Links:	Review Info:
Refseq: NC_002695	Genes: 5372	COG	Genome Project	Publications: [4]
GenBank: BA000007	Protein coding: 5230	TaxMap	Refseq FTP	Refseq Status: Provisional
Length: 5,498,450 nt	Structural RNAs: 141	TaxPlot	GenBank FTP	Seq Status: Completed
GC Content: 50%	Pseudo genes: None	GenePlot	BLAST	Sequencing center: GIRC
% Coding: 85%	Others: 170	gMap	TraceAssembly	Completed: 2001/10/02
Topology: circular	Contigs: 1		CDD	Organism Group
Molecule: DNA			Other genomes for species: 133	

以下のような0157 Sakai株の各遺伝子にCOG番号をアサインしたデータが表示されたページへと移動します。

COGとは、NCBIが提供している遺伝子の機能分類データベースであり、各ゲノムのタンパク質コード遺伝子にCOG番号を割り当てることで、遺伝子が持つ機能的な情報を提供しているものです。同じCOG番号を持つ遺伝子は同じような機能を持つことが期待されます。

COGによる機能分類は約6000種類ほどに分かれており、さらにその上位階層として25種類に分類されています。以下の図では、上部に0157 Sakaiの各遺伝子がどの25種類に分類されるのかを色分けした情報と、下部にその集計情報とが表示されています。



Code	COGs	Description	(% in sequence)	(% in genome)	(% in genus)	(% in Gammaproteobacteria)	(% in Bacteria)	
<input type="checkbox"/>	J	194	Translation	3.2247	3.1751	3.3715	4.2982	4.3421
<input type="checkbox"/>	A	2	RNA processing and modification	0.0332	0.0327	0.0357	0.0311	0.0144
<input type="checkbox"/>	K	352	Transcription	5.8511	5.8101	6.1126	5.9460	5.9565
<input type="checkbox"/>	L	311	Replication, recombination and repair	5.1695	5.3682	5.1322	5.6096	4.9420
<input type="checkbox"/>	B	0	Chromatin structure and dynamics	0.0000	0.0000	0.0000	0.0162	0.0279
<input type="checkbox"/>	D	36	Cell cycle control, mitosis and meiosis	0.5984	0.6056	0.6661	0.7877	0.7632
<input type="checkbox"/>	Y	0	Nuclear structure	0.0000	0.0000	0.0000	0.0000	0.0000
<input type="checkbox"/>	V	53	Defense mechanisms	0.8810	0.9002	0.9607	1.1040	1.2533
<input type="checkbox"/>	T	203	Signal transduction mechanisms	3.3743	3.3224	3.4376	4.3089	3.9496
<input type="checkbox"/>	M	259	Cell wall/membrane biogenesis	4.3052	4.3208	4.5019	4.6689	4.3540
<input type="checkbox"/>	N	116	Cell motility	1.9282	1.9967	1.9590	2.1616	1.4129
<input type="checkbox"/>	Z	0	Cytoskeleton	0.0000	0.0000	0.0000	0.0107	0.0120
<input type="checkbox"/>	W	4	Extracellular structures	0.0665	0.0655	0.0411	0.0241	0.0145
<input type="checkbox"/>	U	150	Intracellular trafficking and secretion	2.4934	2.6514	2.5143	2.4173	1.7598

COG分類の概要表を読み取る

COG分類の概要表をよく見てみましょう。

Code	COGs	Description	(% in sequence)	(% in genome)	(% in genus)	(% in Gammaproteobacteria)	(% in Bacteria)	
<input type="checkbox"/>	J	194	Translation	3.2247	3.1751	3.3715	4.2982	4.3421
<input type="checkbox"/>	A	2	RNA processing and modification	0.0332	0.0327	0.0357	0.0311	0.0144
<input type="checkbox"/>	K	352	Transcription	5.8511	5.8101	6.1126	5.9460	5.9565
<input type="checkbox"/>	L	311	Replication, recombination and repair	5.1695	5.3682	5.1322	5.6096	4.9420
<input type="checkbox"/>	B	0	Chromatin structure and dynamics	0.0000	0.0000	0.0000	0.0162	0.0279
<input type="checkbox"/>	D	36	Cell cycle control, mitosis and meiosis	0.5984	0.6056	0.6661	0.7877	0.7632
<input type="checkbox"/>	Y	0	Nuclear structure	0.0000	0.0000	0.0000	0.0000	0.0000
<input type="checkbox"/>	V	53	Defense mechanisms	0.8810	0.9002	0.9607	1.1040	1.2533
<input type="checkbox"/>	T	203	Signal transduction mechanisms	3.3743	3.3224	3.4376	4.3089	3.9496
<input type="checkbox"/>	M	259	Cell wall/membrane biogenesis	4.3052	4.3208	4.5019	4.6689	4.3540
<input type="checkbox"/>	N	116	Cell motility	1.9282	1.9967	1.9590	2.1616	1.4129
<input type="checkbox"/>	Z	0	Cytoskeleton	0.0000	0.0000	0.0000	0.0107	0.0120
<input type="checkbox"/>	W	4	Extracellular structures	0.0665	0.0655	0.0411	0.0241	0.0145
<input type="checkbox"/>	U	150	Intracellular trafficking and secretion	2.4934	2.6514	2.5143	2.4173	1.7598
<input type="checkbox"/>	O	163	Posttranslational modification, protein turnover, chaperones	2.7094	2.7005	2.7697	3.3036	2.9523
<input type="checkbox"/>	C	300	Energy production and conversion	4.9867	4.9100	5.3947	4.9092	4.7983
<input type="checkbox"/>	G	389	Carbohydrate transport and metabolism	6.4661	6.3830	7.4358	5.1274	4.8607
<input type="checkbox"/>	E	436	Amino acid transport and metabolism	7.2473	7.1358	7.7859	7.5364	7.2550
<input type="checkbox"/>	F	91	Nucleotide transport and metabolism	1.5126	1.4894	1.7072	1.7399	1.7834
<input type="checkbox"/>	H	160	Coenzyme transport and metabolism	2.6596	2.6187	2.8161	3.1682	2.9938
<input type="checkbox"/>	I	113	Lipid transport and metabolism	1.8783	1.8494	1.9643	2.5132	2.9319
<input type="checkbox"/>	P	314	Inorganic ion transport and metabolism	5.2194	5.1391	5.2715	4.9139	4.5646
<input type="checkbox"/>	Q	87	Secondary metabolites biosynthesis, transport and catabolism	1.4461	1.4403	1.5929	1.9177	2.3300
<input type="checkbox"/>	R	595	General function prediction only	9.8903	9.7872	9.8645	9.9953	10.4301
<input type="checkbox"/>	S	422	Function unknown	7.0146	6.9067	6.4519	6.5640	6.1355
<input checked="" type="checkbox"/>	-	1266	Not in COGs	21.0439	21.3912	18.2128	16.9267	20.1622

この表では、左から1文字コード、現在見ているゲノム配列に含まれる遺伝子数、分類の説明、現在見ているゲノム配列中に占める割合、現在見ている生物種ゲノム全体に占める割合、現在見ている生物種が属する“属”全体に占める割合、γプロテオバクテリア中で占める割合、バクテリア中で占める割合を示しています。

例えば、現在見ている生物種ゲノム全体に占める割合と現在見ている生物種が属する“属”全体に占める割合とを比べることで、この生物種が持つ機能的な特徴を浮き上がらせることができます。

0157 Sakaiでは、細胞外構造物に分類される(W)遺伝子の割合が高いことや、糖代謝、輸送に関する遺伝子の割合が少ないこと(G)などがわかります。

Intracellular trafficking and secretion(U:細胞内輸送、分泌)に分類される遺伝子の一覧を、Uと書かれた右横の150をクリックすることで取得してみましょう。

クリックすると下図の様な遺伝子一覧が得られます。

Escherichia coli O157:H7 str. Sakai

Intracellular trafficking and secretion genes:

Location	strand	Length	PID	Locus	LocusTag	COG	Product
19298.21748	-	816	15829276	=	ECs0022	COG3188NU	putative outer membrane usher protein precursor
21761.22444	-	227	15829277	=	ECs0023	COG3121NU	putative fimbrial chaperone
22494.23027	-	177	15829278	=	ECs0024	COG3539NU	putative fimbrial protein
29618.30112	+	164	15829284	<i>lspA</i>	ECs0030	COG0597MU	lipoprotein signal peptidase
112884.115589	+	901	15829356	<i>secA</i>	ECs0102	COG0653U	preprotein translocase subunit SecA
119011.120213	-	400	15829364	=	ECs0110	COG1459NU	type IV pilin biogenesis protein
120203.121588	-	461	15829365	=	ECs0111	COG2804NU	hypothetical protein ECs0111
121598.122038	-	146	15829366	=	ECs0112	COG4969NU	putative major pilin subunit
157241.159841	-	886	15829397	=	ECs0143	COG3188NU	putative outer membrane usher protein
159883.160608	-	241	15829398	=	ECs0144	COG3121NU	putative chaperone protein EcpD
286244.287983	-	579	15829510	=	ECs0257	COG1298NU	FliA
368426.369520	-	384	15829604	=	ECs0350	COG3468MU	putative adhesin
379734.383717	+	1327	15829616	=	ECs0362	COG3468MU	AidA-I adhesin-like protein
379734.383717	+	1327	15829616	=	ECs0362	COG5295UW	AidA-I adhesin-like protein
451255.454197	+	980	15829678	=	ECs0424	COG3468MU	putative flagellin structural protein
490916.491248	+	110	15829712	<i>yajC</i>	ECs0458	COG1862U	preprotein translocase subunit YajC
491276.493123	+	615	15829713	<i>secD</i>	ECs0459	COG0342U	preprotein translocase subunit SecD
493134.494105	+	323	15829714	<i>secE</i>	ECs0460	COG0341U	preprotein translocase subunit SecE
522729.523352	+	207	15829745	<i>clpP</i>	ECs0491	COG07400U	ATP-dependent Clp protease proteolytic subunit
579895.581250	+	451	15829794	=	ECs0540	COG1538MU	putative outer membrane transport protein
606274.607290	-	338	15829802	=	ECs0548	COG5295UW	adhesin/invasin-like protein
608198.608653	-	151	15829805	=	ECs0551	COG15850U	hypothetical protein ECs0551
651426.652118	+	230	15829847	=	ECs0593	COG3121NU	putative chaperone
652149.654758	+	869	15829848	=	ECs0594	COG3188NU	putative outer membrane protein
655789.656304	+	171	15829850	=	ECs0596	COG3539NU	putative fimbrial protein
662098.664335	-	745	15829855	<i>nfrB</i>	ECs0601	COG2804NU	bacteriophage N4 adsorption protein B
678260.679642	+	460	15829864	=	ECs0610	COG1538MU	copper/silver efflux system outer membrane protein CusC
740303.740506	+	87	15829919	<i>tatE</i>	ECs0665	COG1826U	twin arginine translocase protein E
825622.826353	-	243	15829996	=	ECs0742	COG3121NU	putative chaperone
826656.828827	-	723	15829997	=	ECs0743	COG3188NU	truncated outer membrane protein
859659.860351	+	230	15830026	=	ECs0772	COG0611U	colicin uptake protein TolQ

■ 解答1-6

[TaxPlot情報の見方](#)

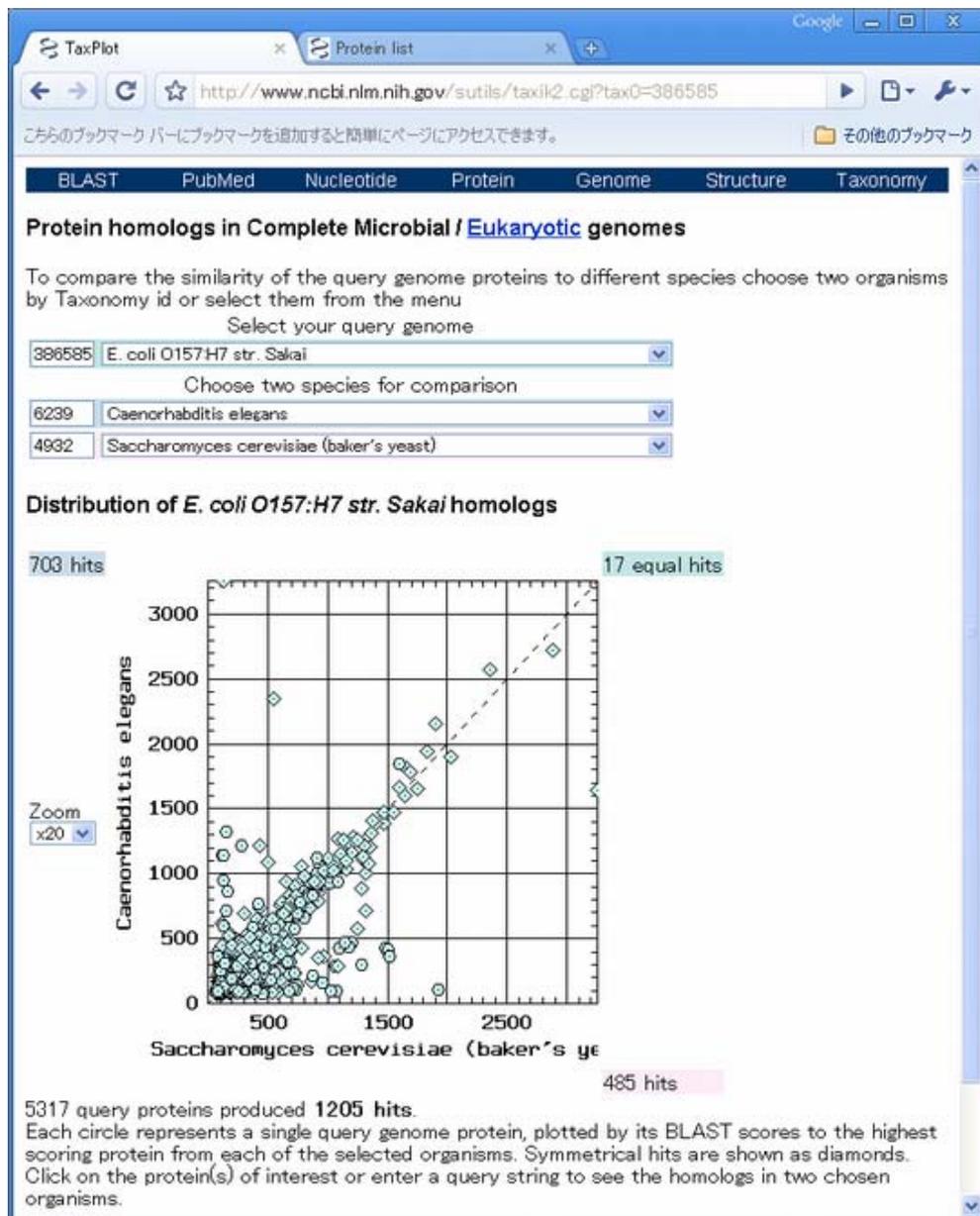
Backボタンを使って“Entrez Genome”概要のページへ戻ってください。

その表の中からTaxPlotと書かれたリンクをクリックしてください。

Genome Info:	Features:	BLAST homologs:	Links:	Review Info:
Refseq: NC_002695	Genes: 5372	COG	Genome Project	Publications: [4]
GenBank: BA000007	Protein coding: 5230	TaxMap	Refseq FTP	Refseq Status: Provisional
Length: 5,498,450 nt	Structural RNAs: 141	TaxPlot	GenBank FTP	Seq Status: Completed
GC Content: 50%	Pseudo genes: None	GenePlot	TraceAssembly	Sequencing center: GIRC
% Coding: 85%	Others: 170	gMap		Completed: 2001/10/02
Topology: circular	Contigs: 1		CDD	Organism Group:
Molecule: DNA			Other genomes for species: 133	

TaxPlotへのリンク

以下のような画面へと移動します。

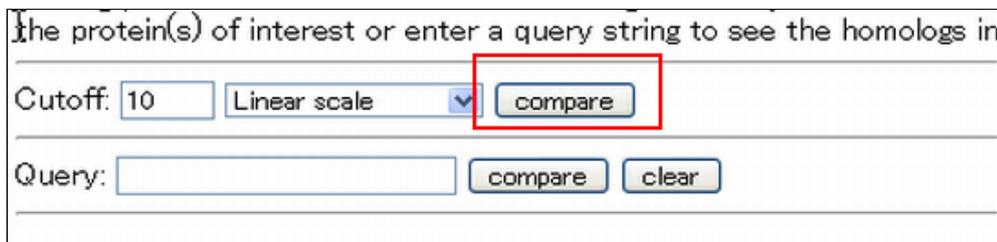
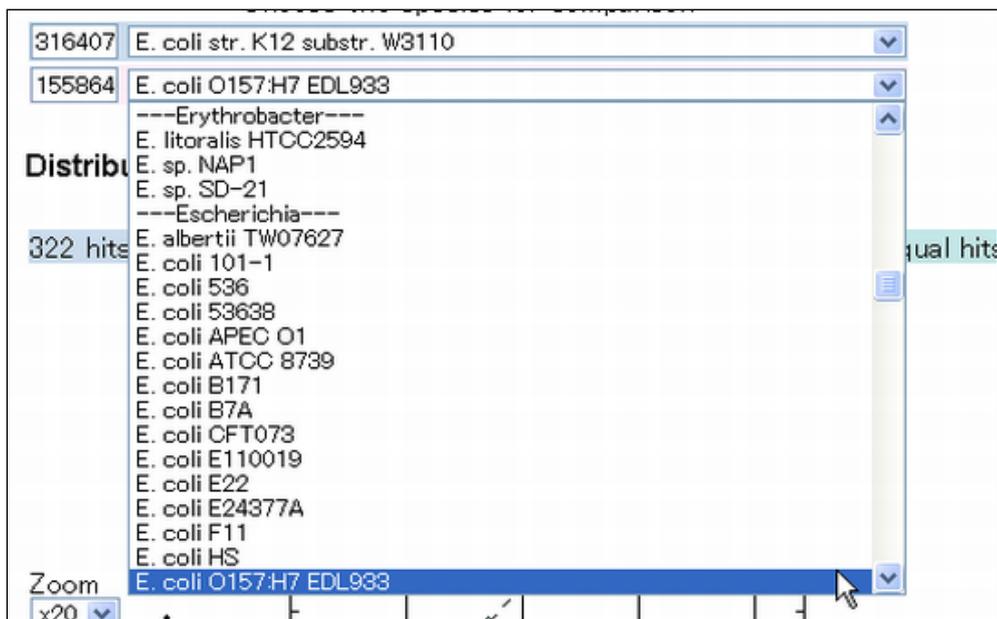
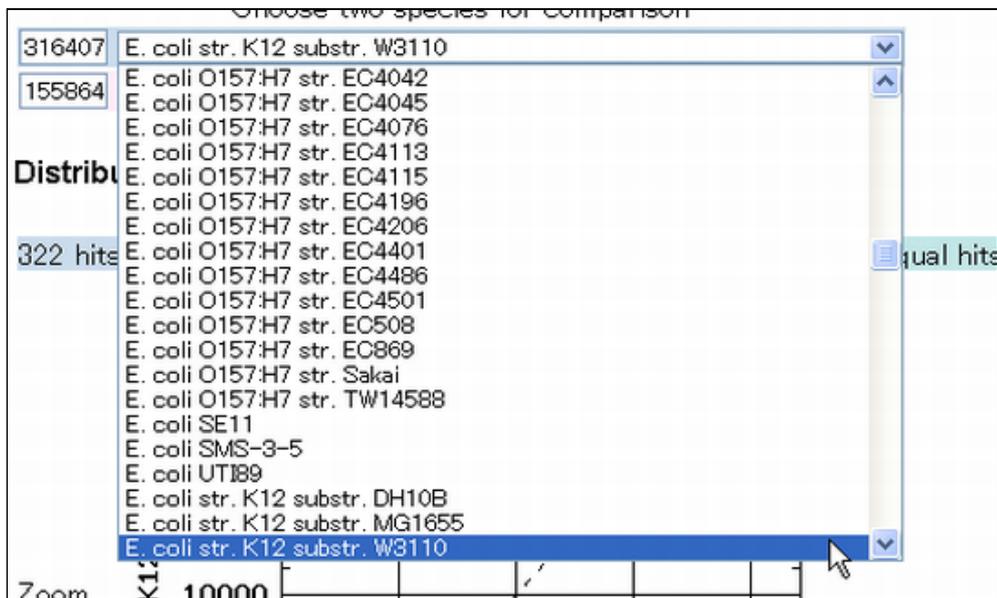


TaxPlotでは現在見ている生物種の全遺伝子（染色体、プラスミドすべて）を他の二生物種の全遺伝子と比較し、その結果を相同性に基づいてプロットした情報を提供しています。

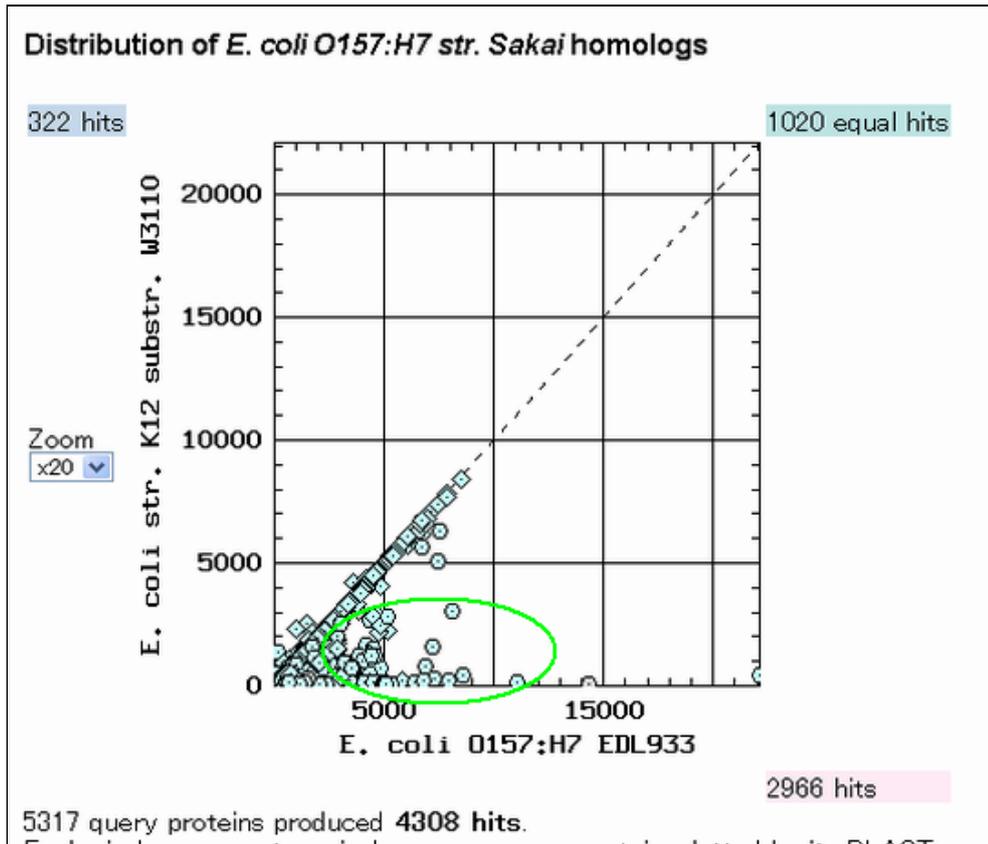
今見ているページでは、0157 Sakai株の全5,317遺伝子を出芽酵母（横軸）と線虫（縦軸）と比較し、その結果を相同性に基づいてプロットした結果が表示されています。各点が遺伝子に対応しており、斜め45度の線より右下に

あるものは酵母により似ている事を、左上にあるものは線虫とより似ている事を示しています。

病原性大腸菌である0157 sakai株の特徴を調べるために、大腸菌K12株と他のH157株との比較を行いましょ。画面上部のChoose two species for comparisonのプルダウンメニューからE. coli K12 W3110とE. coli 0157 EDLとをそれぞれ選択し、画面下部のCompareボタンをクリックして下さい。



すると、下図のように横軸にE. coli 0157 EDL, 縦軸にE. coli K12との相同性を示す結果が得られます。



0157 EDLにのみ類似している遺伝子群が存在することがわかります。(図中緑で囲んだ領域)

原文更新日：2005年2月15日 日本語版更新日：2008年8月23日

All Rights Reserved, Copyright(C) 1997-2008 Japan Science and Technology Agency(JST)