

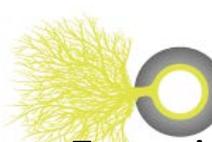
平成26年度ライフサイエンスデータベース統合推進事業  
トーゴーの日シンポジウム2014

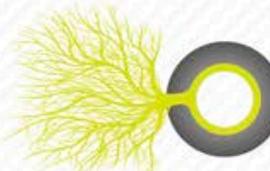
# ゲノム・メタゲノム情報統合による 微生物DBの超高度化推進

東京工業大学地球生命研究所  
黒川 顕

# 微生物研究を取り巻く現状

- 微生物は地球上のいたる所に存在し環境と密接に関与している
- 微生物研究はバイオ分野のみならず、他の多くの分野と連携可能
- 既に多様なDBが多数存在する
- しかし、微生物と環境との関連性を記述しているDBは存在しない
- さらに、専門知識を持っていない**バイオ分野以外の人**は利用困難

 **Microbe DB .JP** integrates lots of data related to microbes.  
Especially, we integrate the microbial data that can be linked to **genomes**.



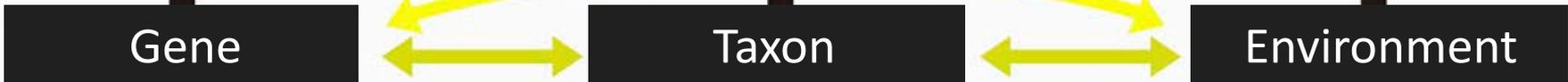
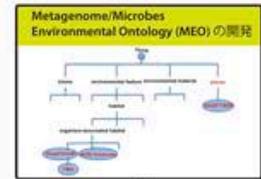
# Microbe DB .JP

<http://microbedb.jp/>

Microbe DB.jp  
MicrobeDB.jp プロジェクトでは様々な微生物学上の知識を、ゲノム情報を核として遺伝子、系統、環境の3つの軸に沿ってセマンティックウェブの技術を使用して整理統合し、幅広い分野での微生物学の見解に資することの出来るデータベースの構築を目標としています。

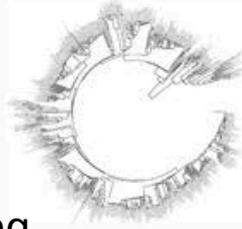
## Ontology

オントロジー: 検索タームの柔軟化&明確化



MBGD  
オースログデータ

Ortholog: **MBGD**



Taxonomy: **NCBI Taxonomy**

環境のメタデータ

Metadata: **INSDC SRA**

オミックスデータ

Genome: **GTPS/RefSeq**

菌株データ  
菌株保存情報 | 培養条件等 |

Culture Collection: **NBRC/JCM**

メタゲノムデータ

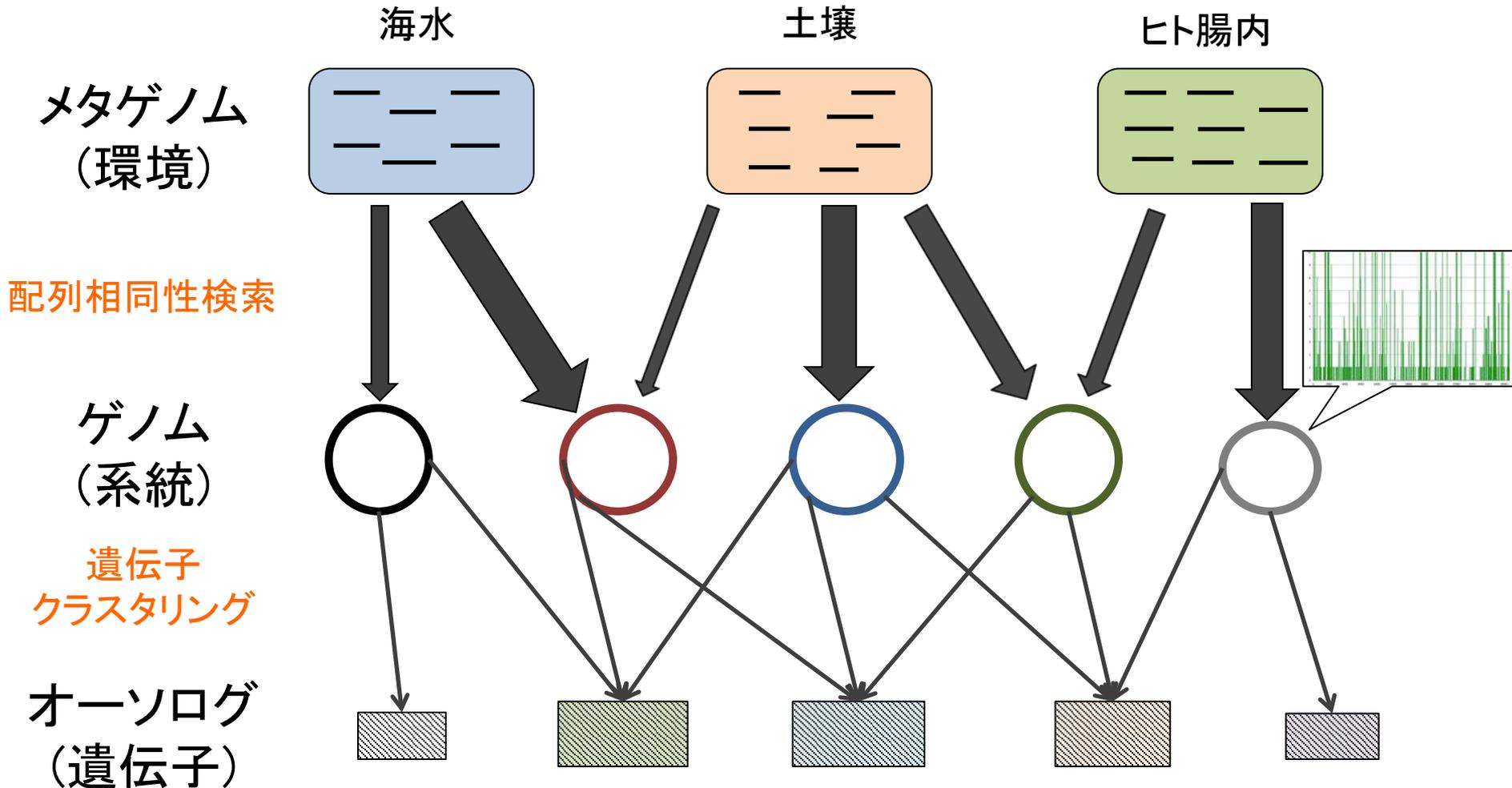
Metagenome: **INSDC SRA**

モデル微生物の高品質アノテーションデータ

Annotation: **TogoAnnotation**

Red color indicates our collaborators.

# 全てのデータをゲノムを核として統合する



# 微生物における各種DBを統合化し、環境情報との連携を徹底的に記述した新規DB構築を実現

- 異分野データの統合化のため、セマンティックウェブの技術を徹底的に導入
- データ間をリンクするためのゲノム情報、オーソログ遺伝子情報、メタゲノム情報の整備
- 全データのRDF化、各データID間のリンク構築
- 各種オントロジーの開発、各データにマッピング
- アノテーション高度化システムの開発
- ユーザ認証システム基盤開発
- 検索結果可視化アプリケーション「Stanza」の開発
- ゲノム自動アノテーションシステムMiGAPとの連携

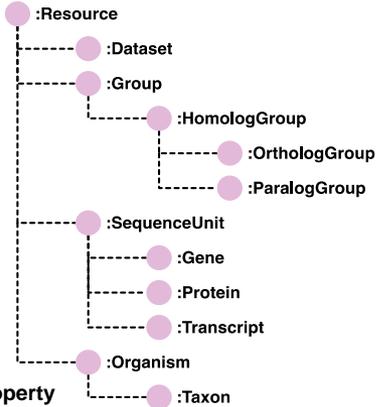
# 各種データのトリプル数

グラフ名	説明	作成元	トリプル数
refseq	RefSeq Prokaryoteゲノムデータ	DBCLS	550,273,744
mbgd	MBGD Orthologデータ	基生研	291,714,037
gtps	GTPSゲノムデータ	遺伝研	197,069,932
taxonomy	SPARQLthonで作成したNCBI Taxonomyオントロジー改良版	DBCLS,遺伝研,東工大	10,183,714
meta16S	各SRSメタ16Sの系統組成データ	東工大	9,831,600
gazetteer	地理オントロジー	外部機関	7,062,536
srs_metadata	SRSメタ16S・メタゲノムの様々なメタデータ	東工大	4,982,739
srs_ortholog	各SRSメタゲノムのMBGD Ortholog組成	東工大,基生研	2,026,746
go	Geneオントロジー	外部機関	1,211,571
brc	JCM/NBRC菌株データ with NCBI Taxonomy ID	遺伝研,東工大,DBCLS	903,319
gold	GOLDの個別ゲノムのMEO等へのオントロジーマッピングデータ	東工大,DBCLS	150,899
srs	SRSメタ16S・メタゲノムのMEO等へのオントロジーマッピングデータ	東工大	53,691
so	Sequenceオントロジー	外部機関	43,060
pdo	感染症オントロジー + 症状オントロジー + ゲノムへのオントロジーマッピングデータ	東工大	8,809
meo	微生物の生息環境オントロジー	東工大	4,975
msv	SRSメタ16S・メタゲノムのメタデータオントロジー	東工大	1,601
mpo	微生物フェノタイプオントロジー	DBCLS	734
mccv	菌株オントロジー	東工大,DBCLS	293
その他中間データ	いくつかのデータ集計系のSPARQLクエリは遅いため、MSSが集計結果のデータを作成		440,773
合計			1,075,964,773

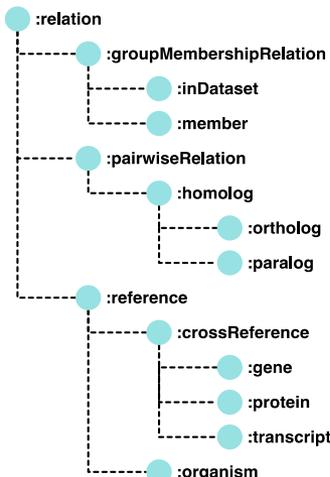
# オーソログ情報のRDF化

## 汎用 Ortholog Ontology (OrthO)の開発

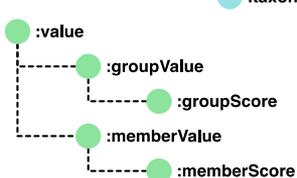
### owl:Class



### owl:ObjectProperty



### owl:DatatypeProperty



## RDF版MBGD に対する SPARQL query のポータルサイト

**MBGD SPARQL Search**
MBGD | Ontology | RDF Archive | Documents

**SPARQL Query**

[A simple example](#)

```

PREFIX orth: <http://purl.jp/bio/11/orth#>

SELECT *
WHERE {
  ?group a orth:OrthologGroup .
}
LIMIT 10
                
```

**Examples (Click to Execute)**

- List of ortholog dataset
 

```

PREFIX orth: <http://purl.jp/bio/11/orth#>
PREFIX dct: <http://purl.org/dc/terms/>
PREFIX pav: <http://purl.org/pav/>

SELECT ?dataset ?description ?derived_from
WHERE {
  ?dataset a orth:Dataset ;
    dct:description ?description ;
    pav:derivedFrom ?derived_from .
}
ORDER BY ?dataset
                    
```
- Search from group descriptions ("thiamin\*")
 

```

PREFIX orth: <http://purl.jp/bio/11/orth#>
PREFIX dct: <http://purl.org/dc/terms/>

SELECT ?dataset ?cluster_id
(COUNT(?member) AS ?size) ?description
WHERE {
  ?group a orth:OrthologGroup ;
    orth:inDataset ?dataset ;
    dct:identifier ?cluster_id ;
    dct:description ?description ;
    orth:member ?member ;
    ?description bif:contains "thiamin*" .
}
ORDER BY ?dataset DESC(?size)
                    
```
- Application1: GO annotation using orthologs (uniprot:K9Z723)
 

```

define input:inference "mbgd"

PREFIX dct: <http://purl.org/dc/terms/>
PREFIX orth: <http://purl.jp/bio/11/orth#>
PREFIX mbgdr: <http://mbgd.genome.ad.jp/rdf/resource/>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX mbgd: <http://purl.jp/bio/11/mbgd#>
PREFIX uniprot: <http://purl.uniprot.org/uniprot/>
PREFIX goa: <http://purl.jp/bio/11/goa#>
PREFIX skos: <http://www.w3.org/2004/02/skos/core#>

SELECT DISTINCT ?cluster_id ?gene_id ?organism
?evidence ?go
WHERE {
  ?group a orth:OrthologGroup ;
    orth:inDataset mbgdr:2014-01_default ;
    dct:identifier ?cluster_id ;
    orth:member/orth:crossReference+ uniprot:K9Z723 ;
    orth:member/orth:gene ?gene .
  ?gene mbgd:uniprot ?uniprot ;
    dct:identifier ?gene_id ;
    orth:organism/rdfs:label ?organism .
  ?uniprot ?goa ?go_id .
  ?goa rdfs:subPropertyOf+ goa:goaExperimental ;
    rdfs:label ?evidence .
  ?go_id skos:prefLabel ?go .
}
                    
```

**MBGD RDF Model (Click to Magnify)**

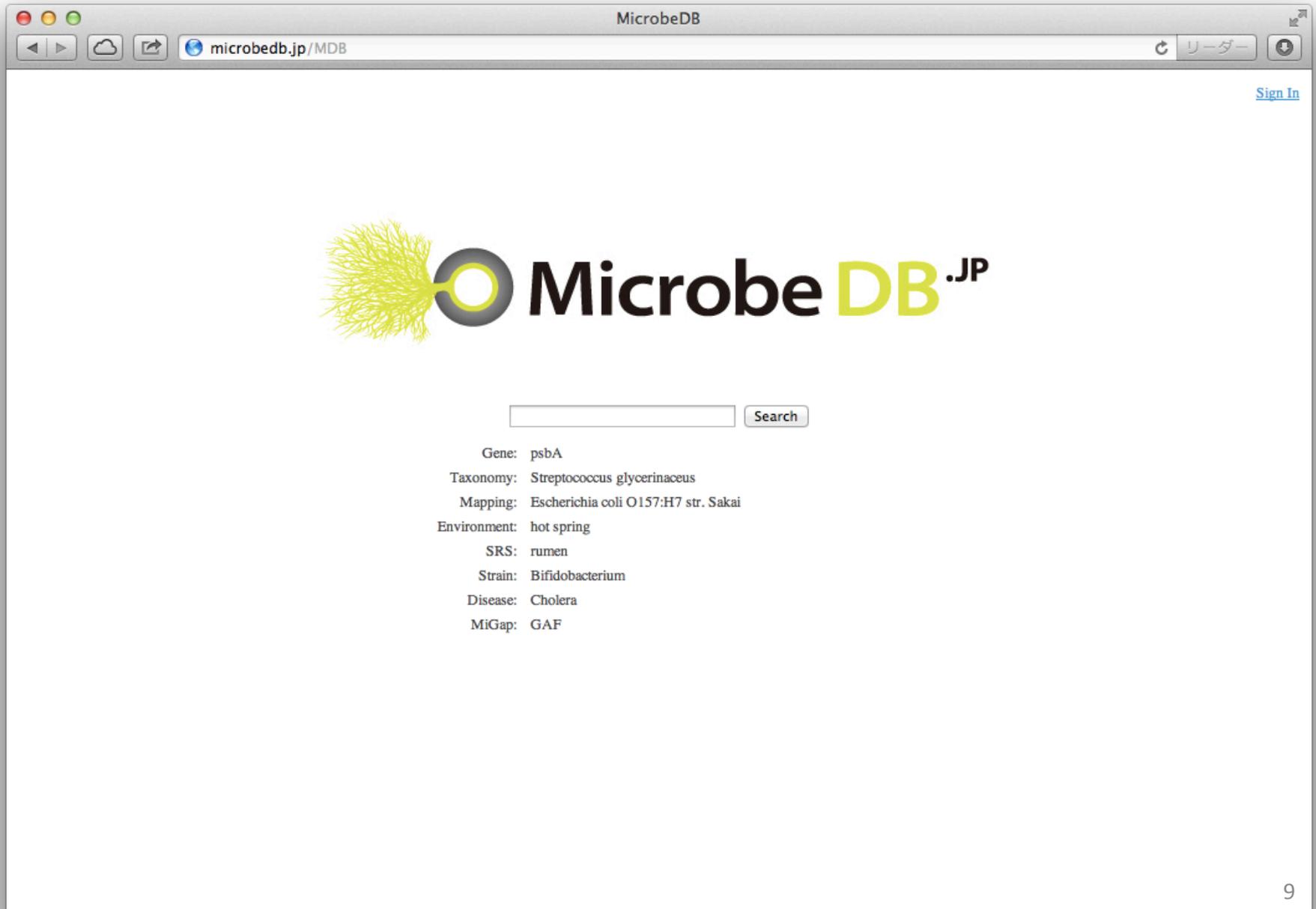
Back-end of this search page is powered by [Virtuoso Open-Source Edition 7.1.0](#) .

The content is available under [Creative Commons Attribution Share Alike](#) unless otherwise noted.

# 開発したオントロジー

- FALDO (Feature Annotation Location Description Ontology)
  - ゲノム中の各featureの位置情報を記述するためのオントロジー (w/BioHackathon)
- INSDC Ontology
  - INSDCエントリのfeatureとqualifierのターム記述のためのオントロジー (w/DBCLS)
- MCCV (Microbial Culture Collection Vocabulary)
  - 菌株データを記述するためのオントロジー
- MEO (Metagenome/Microbe Environmental Ontology)
  - 細菌の生息環境を記述するためのオントロジー
- PDO/CSSO (Pathogenic Disease Ontology with Symptom)
  - 細菌が引き起こす感染症の情報および感染症の症状を連結したオントロジー
- GMO (Growth Media Ontology)
  - 細菌の培地情報を記述するためのオントロジー (w/DBCLS)

<http://microbedb.jp/>



The screenshot shows a web browser window titled "MicrobeDB" with the address bar containing "microbedb.jp/MDB". The page features the MicrobeDB logo, which consists of a yellow, branching, tree-like structure on the left and the text "Microbe DB .JP" in black and yellow. Below the logo is a search bar with a "Search" button. The search results are displayed as follows:

- Gene: psbA
- Taxonomy: Streptococcus glycerinaceus
- Mapping: Escherichia coli O157:H7 str. Sakai
- Environment: hot spring
- SRS: rumen
- Strain: Bifidobacterium
- Disease: Cholera
- MiGap: GAF

In the top right corner of the browser window, there is a "Sign In" link. The browser's address bar also contains navigation icons and the text "リーダー".

# MicrobeDB.jpの開発で実現したこと

1. 既存のゲノム中の各遺伝子の情報（オーソログ、系統プロファイル、環境プロファイル）
2. 菌株保存機関に存在する菌株の情報（生育培地、表現型情報、遺伝子機能組成）
3. 様々な環境中の細菌群集の情報（系統組成、遺伝子機能組成）
4. 上記の情報をシームレスに統合



問合せ例：

高温環境に多く存在する遺伝子はどのような遺伝子か？その遺伝子は、どの系統が主に持っているのか？

# 本研究開発の目標・ねらい

## MicrobeDB.jpを

- より広い微生物種を対象として拡張
- データ収集や更新自動化による持続可能なシステム
- 最先端解析プロトコルを実装した解析結果の可視化

研究者コミュニティだけでなく不特定多数のイノベータを対象とした利用性の向上を徹底する

単なる統計量の羅列ではなく、大規模データから新規知識発見を容易に行う事が可能な、今までのDBを超えた**DBシステム**を構築する事を目標とする。

# 主たる共同研究者

## 東京工業大学

黒川 顕: 微生物DBにおける研究統括

山田拓司: ゲノム・メタゲノムDBの構築

森 宙史: 真菌類データの整備、DB自動更新システムの開発

山本 希: 解析Stanza & オントロジー開発

鈴木真也: 高度解析Stanzaの開発

## 国立遺伝学研究所

中村保一: 藻類データの整備

菅原秀明: MiGAPとの連携強化

神沼英里: MeGAPとの連携強化

藤澤貴智: アクセスレベルの制限システムの開発

## 基礎生物学研究所

内山郁夫: 真核生物に対するオーソログ解析手法の開発

千葉啓和: ドラフトゲノムのオーソログ解析

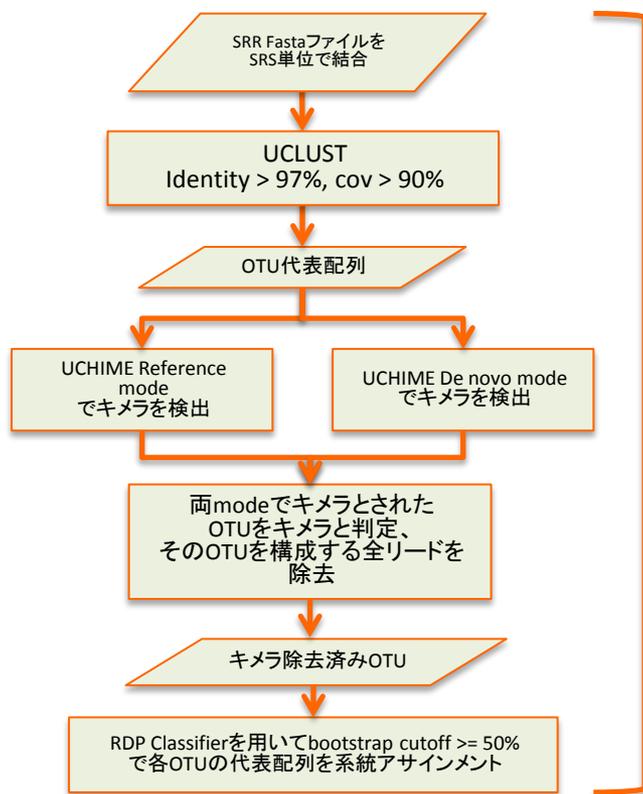
西出浩世: オーソログを基軸とした各種データ統合の推進

# 具体的な研究開発項目（3年間）

1. 各種オミックスデータへの対応
2. 真菌類および藻類を対象とした拡張
3. 各種オントロジー、ボキャブラリの開発
4. 解析プロトコルを実装した各種Stanzaの開発
5. データの収集およびクオリティコントロール、更新の自動化など持続可能なシステムの構築
6. データ共有、公開におけるアクセスレベルの制御システムの構築
7. 構築したシステムを幅広い分野の研究者に活用してもらうためのユーザビリティの向上

# 4. 解析プロトコルを実装した 各種Stanzaの開発

多種多様な情報が混在しているゲノムやメタゲノム等の複雑なデータから知識発見をするために、比較ゲノム解析や比較メタゲノム解析など様々な解析Stanzaを開発する



解析Stanzaによる  
結果の可視化

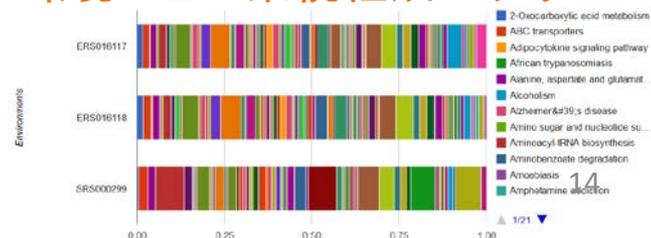


## 各系統と温度との相関係数リスト

メタデータ: temperature  
表示種別: pathway  
全メタデータ幅平均: 10.8063  
ダウンロード  
件数: 224

function ID	機能名	相関係数	サンプル数	メタデータ幅平均
http://www.genome.jp/dbget-bin/www_tget?i=000565	Ether lipid metabolism	0.9655333020905	3	8.5
http://www.genome.jp/dbget-bin/www_tget?i=000072	Synthesis and degradation of ketone bodies	0.6179014416164	25	11.3
http://www.genome.jp/dbget-bin/www_tget?i=002020	Two-component system	0.44564123374872	33	12.6
http://www.genome.jp/dbget-bin/www_tget?i=000910	Nitrogen metabolism	0.4390991977256	34	13.1
http://www.genome.jp/dbget-bin/www_tget?i=000643	Styrene degradation	0.3512843632747	17	10.2
http://www.genome.jp/dbget-bin/www_tget?i=000780	Biotin metabolism	0.34291921434483	26	11.3
http://www.genome.jp/dbget-bin/www_tget?i=003450	Non-homologous end-joining	0.27328119544247	7	12.5
http://www.genome.jp/dbget-bin/www_tget?i=000253	Tetracycline biosynthesis	0.26271210448501	26	11.3
http://www.genome.jp/dbget-bin/www_tget?i=000206	MicroRNAs in cancer	0.19776999332279	12	11.1
http://www.genome.jp/dbget-bin/www_tget?i=000550	Butanoate metabolism	0.18878105249684	26	11.3
http://www.genome.jp/dbget-bin/www_tget?i=004940	Type 1 diabetes mellitus	0.17906842297637	27	11.2
http://www.genome.jp/dbget-bin/www_tget?i=005349	Primary immunodeficiency	0.15863302443048	6	12.3
http://www.genome.jp/dbget-bin/www_tget?i=001552	Tuberculosis	0.15430685893007	29	11.4
http://www.genome.jp/dbget-bin/www_tget?i=001134	Legionellosis	0.1418962728288	28	11.2
http://www.genome.jp/dbget-bin/www_tget?i=003018	RNA degradation	0.08409246481617	29	11.4
http://www.genome.jp/dbget-bin/www_tget?i=000051	Fructose and mannose metabolism	0.07385683214052	27	11.1
http://www.genome.jp/dbget-bin/www_tget?i=000280	Valine, leucine and isoleucine degradation	0.05254845491534	26	11.3
http://www.genome.jp/dbget-bin/www_tget?i=000030	Pentose phosphate pathway	0.059475153688065	27	10.6
http://www.genome.jp/dbget-bin/www_tget?i=000190	Oxidative phosphorylation	0.03063403416373	28	11.4
http://www.genome.jp/dbget-bin/www_tget?i=000110	Carbon fixation in photosynthetic organisms	0.018619571474291	28	10.7
http://www.genome.jp/dbget-bin/www_tget?i=000051	Fatty acid biosynthesis	0.0091971570069616	26	11.3
http://www.genome.jp/dbget-bin/www_tget?i=000945	Stilbenoid, diarylheptanoid and gingerol biosynthesis	2	0.1	
http://www.genome.jp/dbget-bin/www_tget?i=000504	Glycosphingolipid biosynthesis - ganglio series	2	10.4	
http://www.genome.jp/dbget-bin/www_tget?i=001150	Staphylococcus aureus infection	1	10.7	

## 環境ごとの系統組成のグラフ



# Stanza Example

- Gene Annotation

### Feature

[[http://purl.obolibrary.org/obo/SO\\_0000704](http://purl.obolibrary.org/obo/SO_0000704)]

dbxref	<a href="http://www.ncbi.nlm.nih.gov/gene/897644">http://www.ncbi.nlm.nih.gov/gene/897644</a>
feature_gene	polC
feature_locus_tag	TM0576
location	605923..610026
isPartOf	<a href="http://genome.db/uuid/b4d48cd7-00ef-4e03-9adb-fda7de39e078">http://genome.db/uuid/b4d48cd7-00ef-4e03-9adb-fda7de39e078</a>
type	<a href="http://purl.obolibrary.org/obo/SO_0000704">http://purl.obolibrary.org/obo/SO_0000704</a>
label	TM0576

[[http://purl.obolibrary.org/obo/SO\\_0000316](http://purl.obolibrary.org/obo/SO_0000316)]

dbxref	<a href="http://www.ncbi.nlm.nih.gov/gene/897644">http://www.ncbi.nlm.nih.gov/gene/897644</a>
dbxref	<a href="http://www.ncbi.nlm.nih.gov/nucleotide/15643342">http://www.ncbi.nlm.nih.gov/nucleotide/15643342</a>
exons	nodeID://b71582

- Ortholog list

### Ortholog

ID	Genome	Description	Protein	UniProt	GTFS	RefSeq
<a href="#">aac:AACT_1427</a>	<a href="#">aac</a>	DNA polymerase III subunit alpha	<a href="#">YP_003184842.1</a>	<a href="#">C8WWI2</a>	<a href="#">AACT_ACIDOCALDARIUSD46:ST2344</a>	NC_013205.1
<a href="#">aac:ACEAR_1599</a>	<a href="#">aac</a>	DNA polymerase III catalytic subunit, PolC type	<a href="#">YP_003828170.1</a>		<a href="#">AARA_DSM5501:ST105</a>	
<a href="#">acl:ACL_0247</a>	<a href="#">acl</a>	DNA polymerase III subunit alpha	<a href="#">YP_001620249.1</a>	<a href="#">A9NEU3</a>	<a href="#">ALAI_PG8A:ST588</a>	NC_010163.1
<a href="#">af:AFLY_1700</a>	<a href="#">af</a>	DNA polymerase III PolC	<a href="#">YP_002316046.1</a>	<a href="#">B7GG80</a>	<a href="#">AFLA_WK1:ST2505</a>	NC_011567.1
<a href="#">afn:ACFER_1370</a>	<a href="#">afn</a>	DNA polymerase III subunit alpha	<a href="#">YP_003399045.1</a>		<a href="#">AFER_DSM20731:ST1519</a>	NC_013740.1
<a href="#">amt:AMET_2678</a>	<a href="#">amt</a>	DNA polymerase III subunit alpha	<a href="#">YP_001320489.1</a>	<a href="#">A6TRL2</a>	<a href="#">AMET_QYME:ST2214</a>	NC_009633.1

- Genome Information

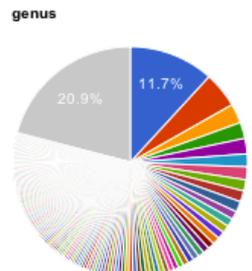
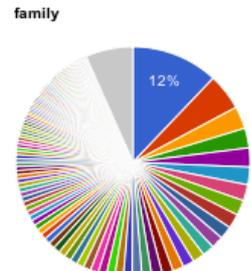
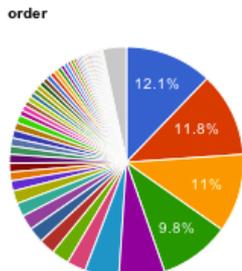
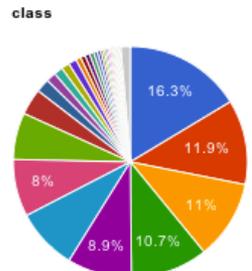
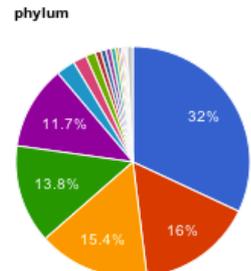
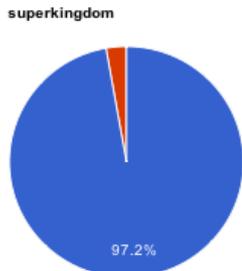
### Genome

length	1860725
location	1..1860725
molecularType	genomic DNA
organism	Thermotoga maritima MSB8
sequence	<a href="http://genome.db/uuid/b4d48cd7-00ef-4e03-9adb-fda7de39e078.fasta">http://genome.db/uuid/b4d48cd7-00ef-4e03-9adb-fda7de39e078.fasta</a>
start	1
stop	1860725
strain	MSB8
version	NC_000853.1
modified	2012-02-13
type	<a href="http://purl.obolibrary.org/obo/SO_0000340">http://purl.obolibrary.org/obo/SO_0000340</a>
type	<a href="http://purl.obolibrary.org/obo/SO_0000988">http://purl.obolibrary.org/obo/SO_0000988</a>
comment	Thermotoga maritima MSB8 chromosome, complete genome.

# Stanza Example

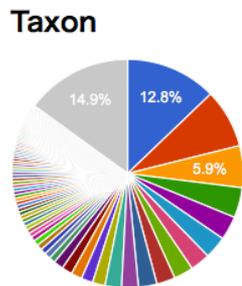
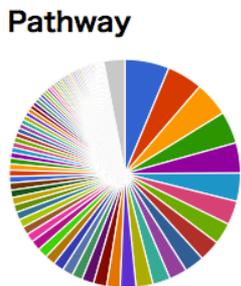
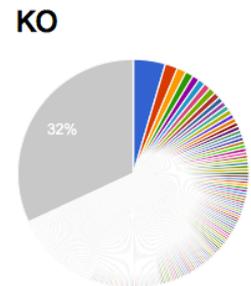
## Taxonomy Composition

- Taxonomic composition of 16S rRNA gene amplicon sequencing analysis



- Functional and taxonomic composition of a metagenome sample

## Sample Function



# 検索ワード: lake

MEO OWL

菌株-MEOマッピングRDF

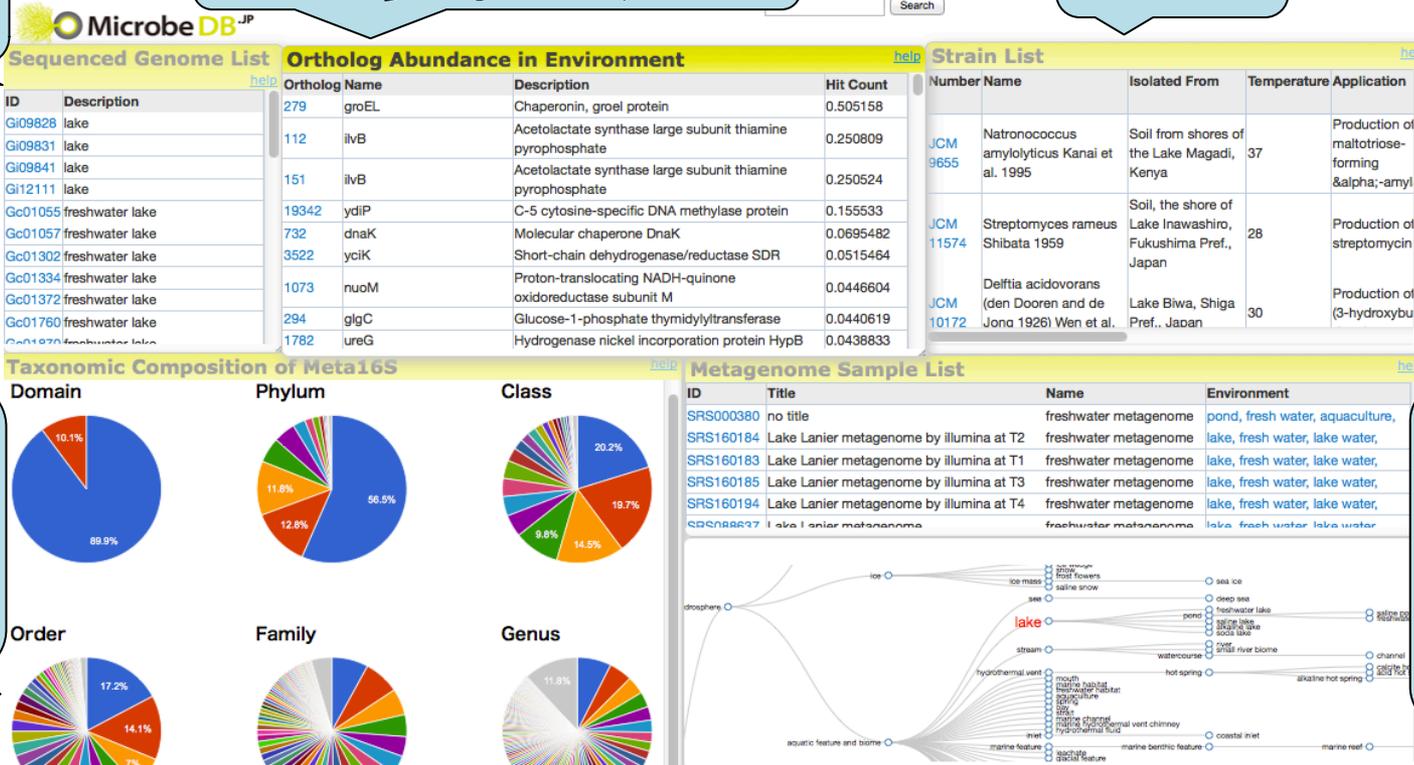
lake → meo:pond is\_a meo:lake → Strain\_A mccv:isolation\_source meo:pond → Strain\_A

のようなオントロジーを介した推論検索を実行

lake由来の  
ゲノム解読  
済み株

lake由来のメタゲノムサンプル  
で多いオーソログ

lake由来  
の菌株

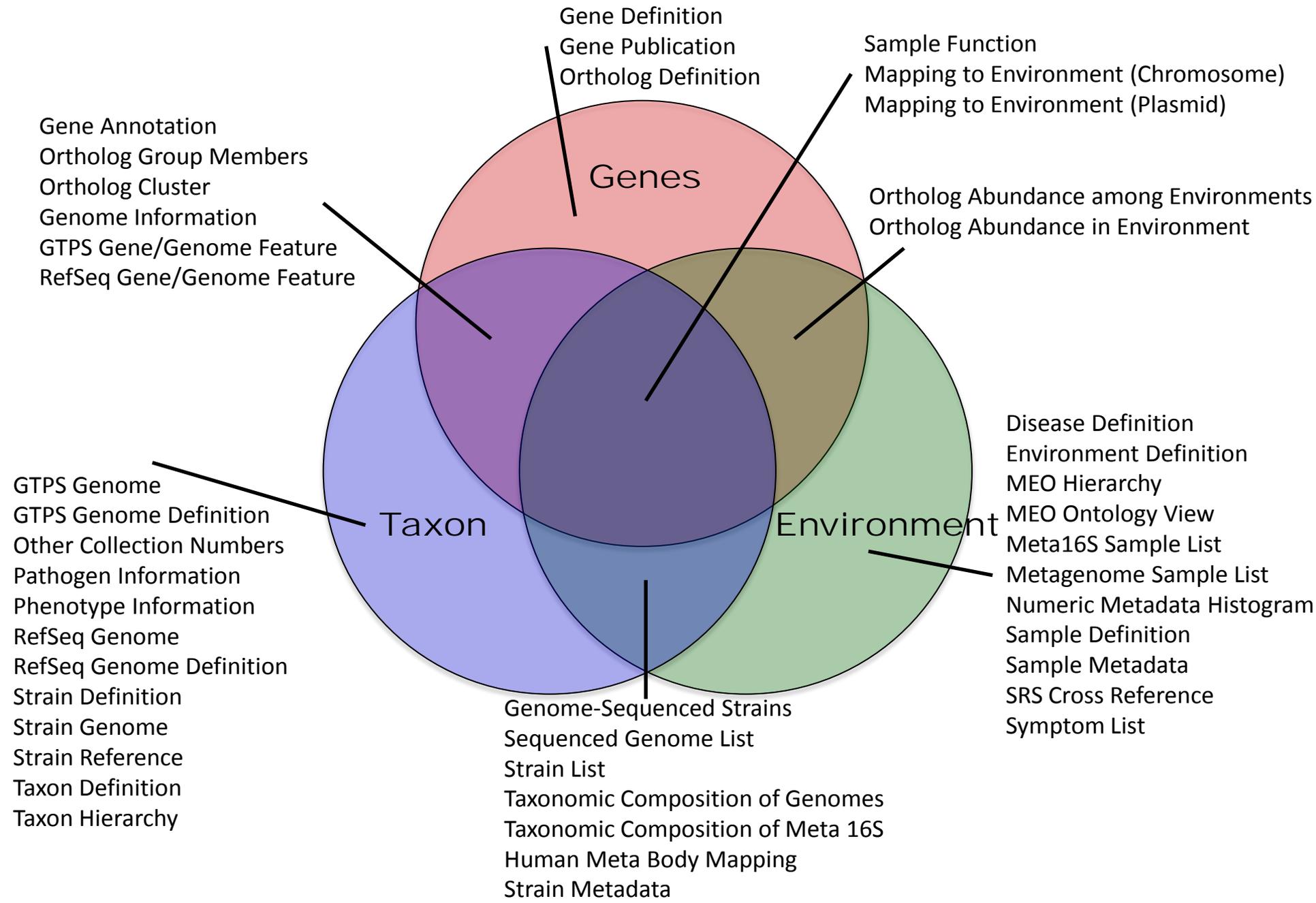


lake由来  
のメタゲ  
ノムサン  
プル

lake由来の  
メタ16Sサン  
プルの平均  
的な系統組  
成

lakeと他  
の語彙の  
MEO中で  
の位置関  
係

# Stanza categories in MicrobeDB.jp



# 【研究テーマ】工業排水のメタゲノム解析

1. 工業廃水における細菌群集構造は？

→細菌群集構造を明らかにする

2. サンプル間でどのような変化がある？

→サンプル(時系列等)ごとに細菌群集を比較する

3. どのような細菌種、遺伝子が環境因子と強い相関を持つ？

→環境因子と強い相関を持つ遺伝子・種の探索

4. 同様の環境ではどのような細菌群集が存在する？

→同様のメタデータを持つサンプルとの比較

5. 同様の細菌群集はどのような環境に存在する？

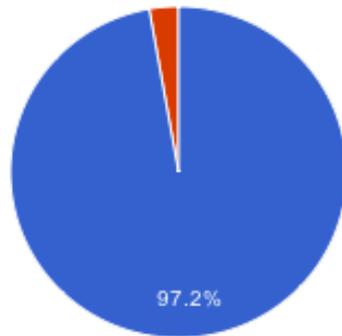
→同様の細菌群集構造を示すサンプルとの比較

# 細菌群集構造スタンザ

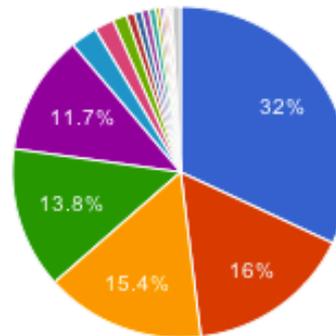
サンプルにおける細菌群集構造を分類群ごとに円グラフで可視化

## Taxonomy Composition

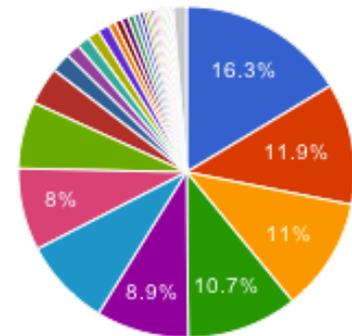
superkingdom



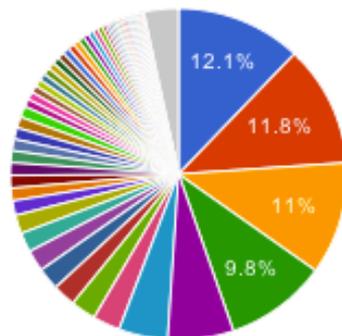
phylum



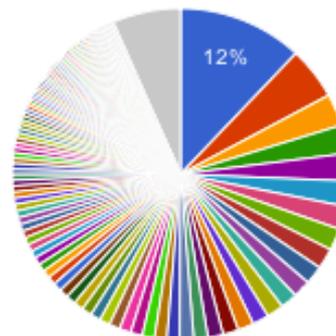
class



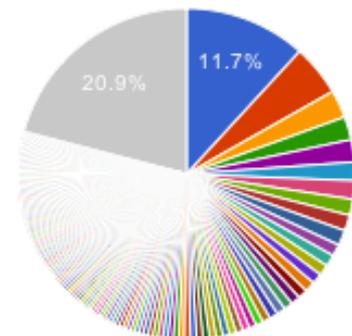
order



family



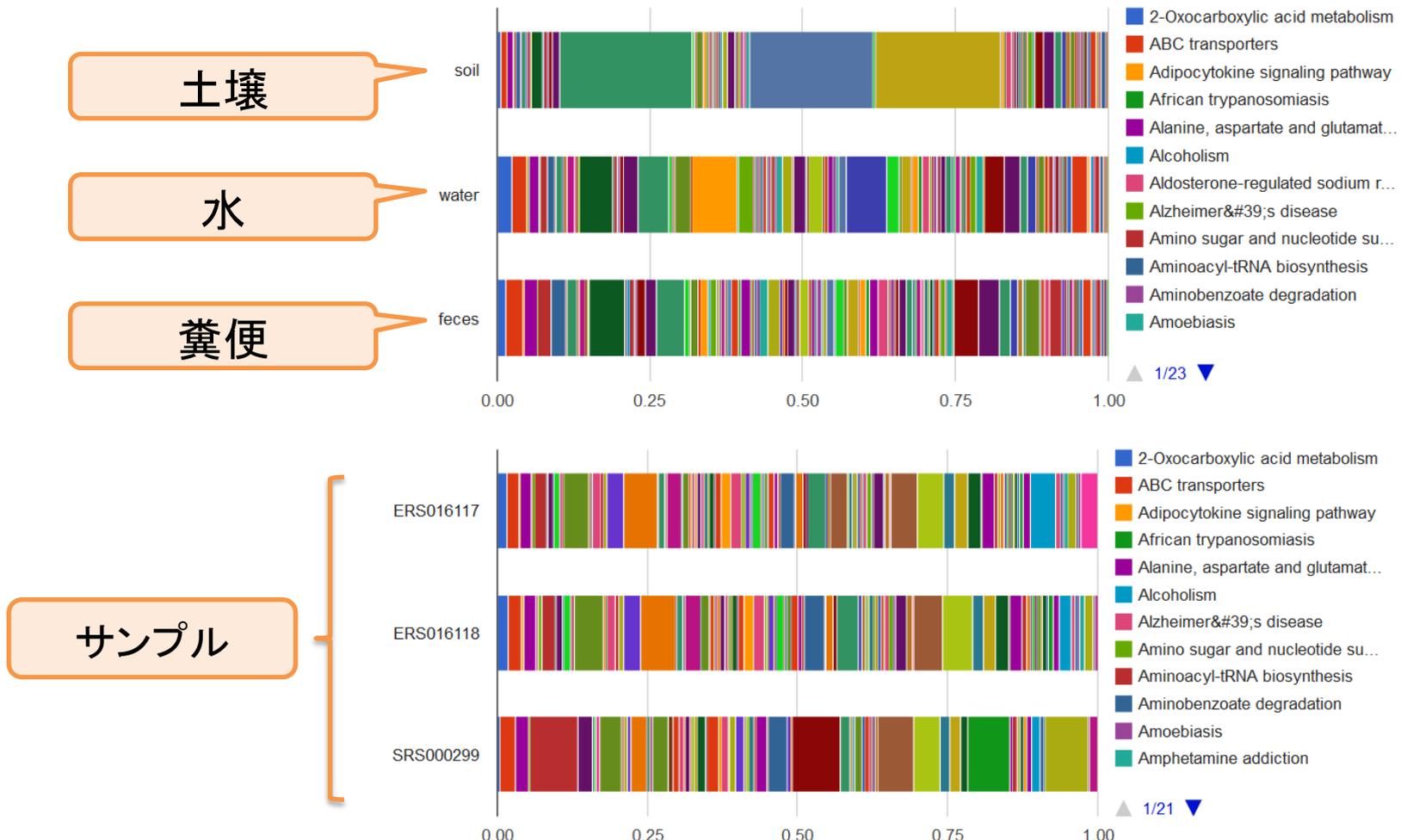
genus



# 系統・遺伝子機能組成の比較解析スタンザ

複数の環境やサンプルの系統組成・遺伝子機能組成を比較できるスタンザ

→ 環境、サンプルの系統的・機能的な特徴を概観できる



# 散布図スタンザ

数値メタデータと系統や遺伝子の存在量の散布図を表示するスタンザ

(1) *Thermocrinis* (genus) は温度と高い相関を示す

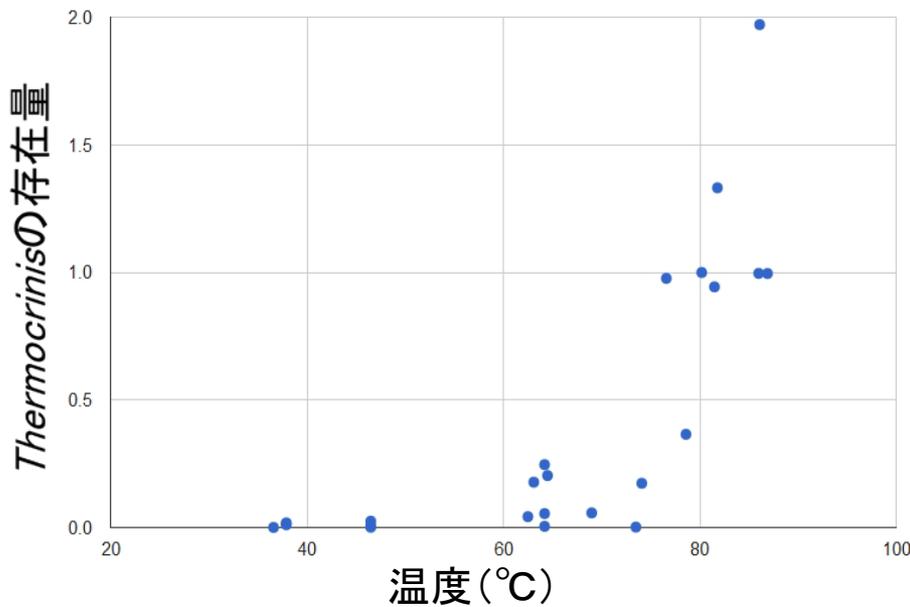


好熱性細菌

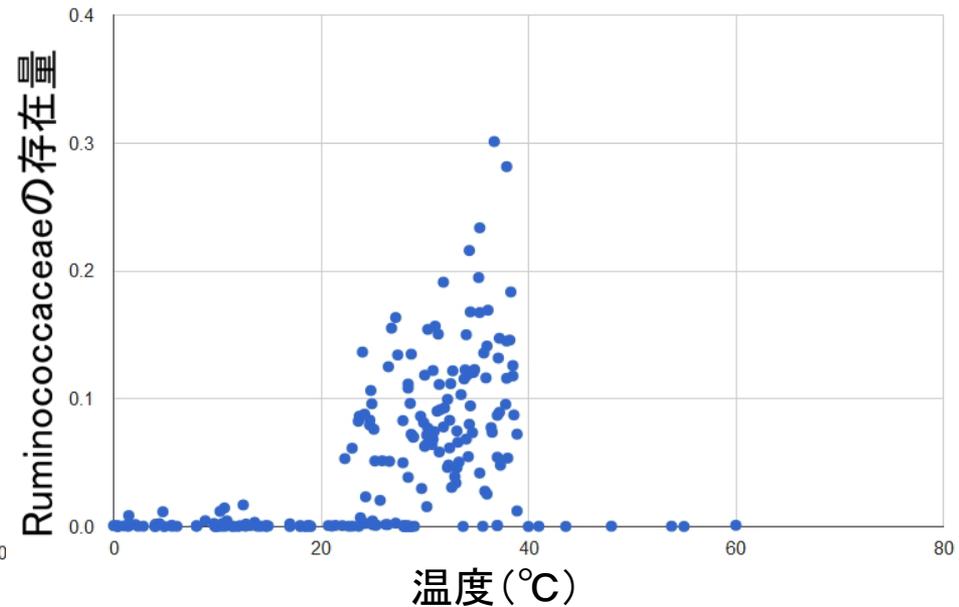
(2) Ruminococcaceae (family) は20~40°Cにサンプルが集中



動物共生細菌が多い



*Thermocrinis*の存在量の散布図



温度とRuminococcaceaeの存在量の散布図

# 相関解析スタンザ

数値メタデータと系統・遺伝子機能の存在量の相関を概観できる

系統ID/  
遺伝子機能ID

メタデータ: temperature  
表示種別: pathway  
全メタデータ値平均: 10.8063  
[ダウンロード](#)  
件数: 224

系統名/  
遺伝子機能名

Pearson  
相関係数

サンプル数

メタデータ値  
平均

function ID	機能名	相関係数	サンプル数	メタデータ値平均
<a href="http://www.genome.jp/dbget-bin/www_bget?ko00565">http://www.genome.jp/dbget-bin/www_bget?ko00565</a>	Ether lipid metabolism	0.95655330320905	3	8.5
<a href="http://www.genome.jp/dbget-bin/www_bget?ko00072">http://www.genome.jp/dbget-bin/www_bget?ko00072</a>	Synthesis and degradation of ketone bodies	0.61790144161164	25	11.3
<a href="http://www.genome.jp/dbget-bin/www_bget?ko02020">http://www.genome.jp/dbget-bin/www_bget?ko02020</a>	Two-component system	0.44564123374872	33	12.6
<a href="http://www.genome.jp/dbget-bin/www_bget?ko00910">http://www.genome.jp/dbget-bin/www_bget?ko00910</a>	Nitrogen metabolism	0.4092991977256	34	13.1
<a href="http://www.genome.jp/dbget-bin/www_bget?ko00643">http://www.genome.jp/dbget-bin/www_bget?ko00643</a>	Styrene degradation	0.3512943632747	17	10.2
<a href="http://www.genome.jp/dbget-bin/www_bget?ko00780">http://www.genome.jp/dbget-bin/www_bget?ko00780</a>	Biotin metabolism	0.34291921434483	26	11.3
<a href="http://www.genome.jp/dbget-bin/www_bget?ko03450">http://www.genome.jp/dbget-bin/www_bget?ko03450</a>	Non-homologous end-joining	0.27328119544247	7	12.5
<a href="http://www.genome.jp/dbget-bin/www_bget?ko00253">http://www.genome.jp/dbget-bin/www_bget?ko00253</a>	Tetracycline biosynthesis	0.26271210448501	26	11.3
<a href="http://www.genome.jp/dbget-bin/www_bget?ko05206">http://www.genome.jp/dbget-bin/www_bget?ko05206</a>	MicroRNAs in cancer	0.19776999332279	12	11.1
<a href="http://www.genome.jp/dbget-bin/www_bget?ko00650">http://www.genome.jp/dbget-bin/www_bget?ko00650</a>	Butanoate metabolism	0.18878105249684	26	11.3
<a href="http://www.genome.jp/dbget-bin/www_bget?ko04940">http://www.genome.jp/dbget-bin/www_bget?ko04940</a>	Type I diabetes mellitus	0.17906842287637	27	11.2
<a href="http://www.genome.jp/dbget-bin/www_bget?ko05340">http://www.genome.jp/dbget-bin/www_bget?ko05340</a>	Primary immunodeficiency	0.15863303443048	8	12.3
<a href="http://www.genome.jp/dbget-bin/www_bget?ko05152">http://www.genome.jp/dbget-bin/www_bget?ko05152</a>	Tuberculosis	0.15430685080307	29	11.4
<a href="http://www.genome.jp/dbget-bin/www_bget?ko05134">http://www.genome.jp/dbget-bin/www_bget?ko05134</a>	Legionellosis	0.14189627282888	28	11.2
<a href="http://www.genome.jp/dbget-bin/www_bget?ko03018">http://www.genome.jp/dbget-bin/www_bget?ko03018</a>	RNA degradation	0.088409246481617	29	11.4
<a href="http://www.genome.jp/dbget-bin/www_bget?ko00051">http://www.genome.jp/dbget-bin/www_bget?ko00051</a>	Fructose and mannose metabolism	0.087385683214052	27	11.1
<a href="http://www.genome.jp/dbget-bin/www_bget?ko00280">http://www.genome.jp/dbget-bin/www_bget?ko00280</a>	Valine, leucine and isoleucine degradation	0.06284845491534	26	11.3
<a href="http://www.genome.jp/dbget-bin/www_bget?ko00030">http://www.genome.jp/dbget-bin/www_bget?ko00030</a>	Pentose phosphate pathway	0.059475153688065	27	10.6
<a href="http://www.genome.jp/dbget-bin/www_bget?ko00190">http://www.genome.jp/dbget-bin/www_bget?ko00190</a>	Oxidative phosphorylation	0.03063403416373	28	11.4
<a href="http://www.genome.jp/dbget-bin/www_bget?ko00710">http://www.genome.jp/dbget-bin/www_bget?ko00710</a>	Carbon fixation in photosynthetic organisms	0.018619571474591	28	10.7
<a href="http://www.genome.jp/dbget-bin/www_bget?ko00061">http://www.genome.jp/dbget-bin/www_bget?ko00061</a>	Fatty acid biosynthesis	0.0091971570069616	26	11.3
<a href="http://www.genome.jp/dbget-bin/www_bget?ko00945">http://www.genome.jp/dbget-bin/www_bget?ko00945</a>	Stilbenoid, diarylheptanoid and gingerol biosynthesis		2	9.1
<a href="http://www.genome.jp/dbget-bin/www_bget?ko00604">http://www.genome.jp/dbget-bin/www_bget?ko00604</a>	Glycosphingolipid biosynthesis - ganglio series		2	10.4
<a href="http://www.genome.jp/dbget-bin/www_bget?ko05150">http://www.genome.jp/dbget-bin/www_bget?ko05150</a>	Staphylococcus aureus infection		1	10.7

# Numeric Metadata Stanza

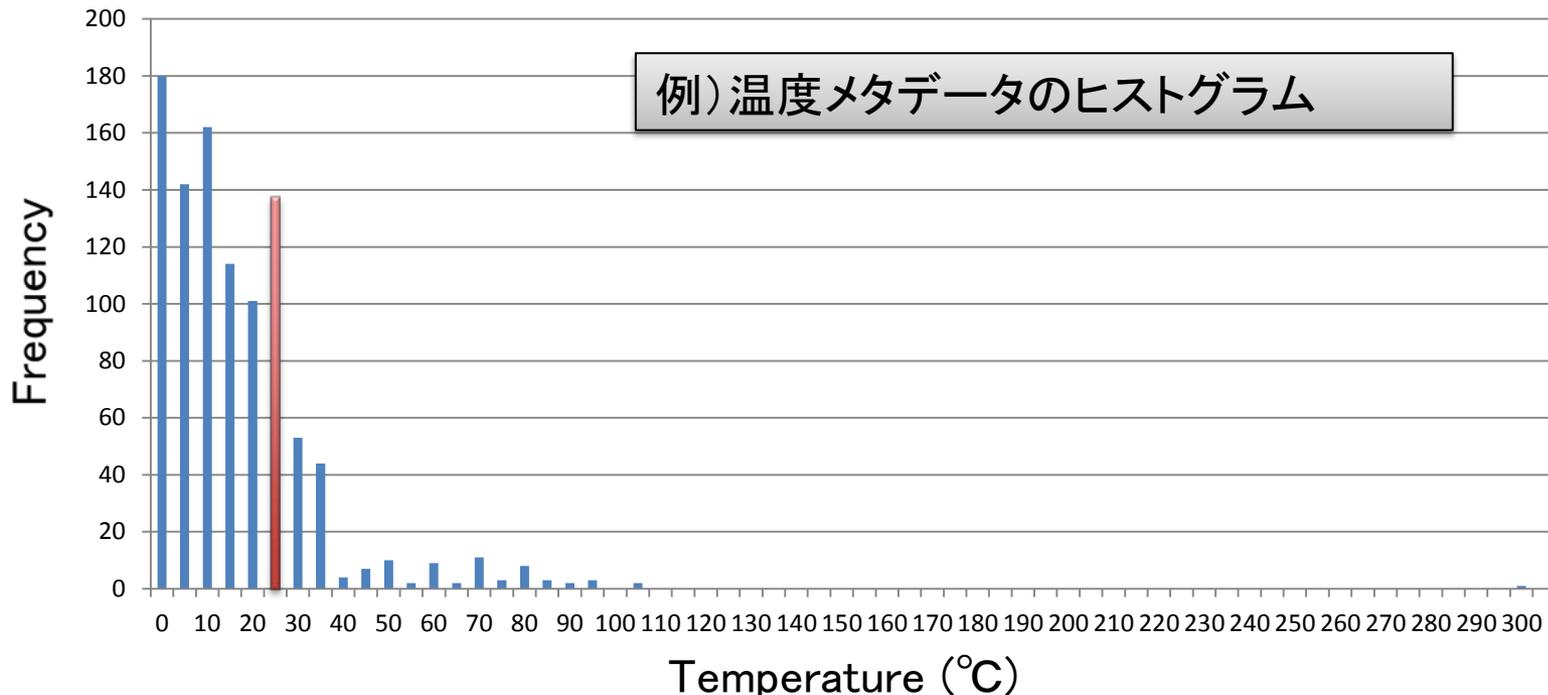
数値メタデータ※の分布をヒストグラムで表示するスタンザ

※温度、年齢、BMI etc.

入力: MSV ID (メタデータID), SRS ID (メタゲノムサンプルID))

出力: ヒストグラム

・SRS IDも入力すると、そのサンプルが位置する部分を強調

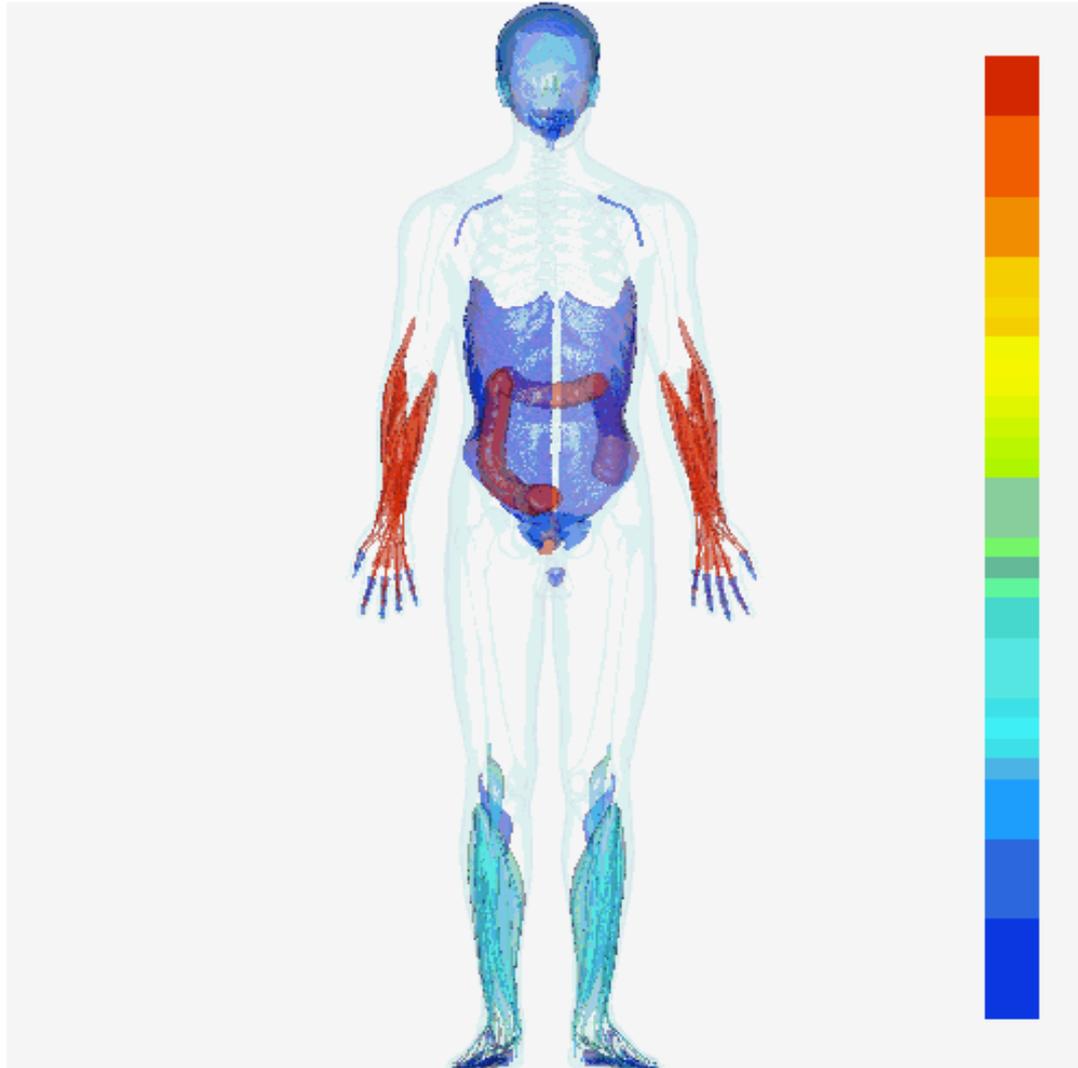




# Human Meta Body Map Stanza

自分の興味がある系統・遺伝子はヒトの体のどこに多いのか？

## Body Mapping



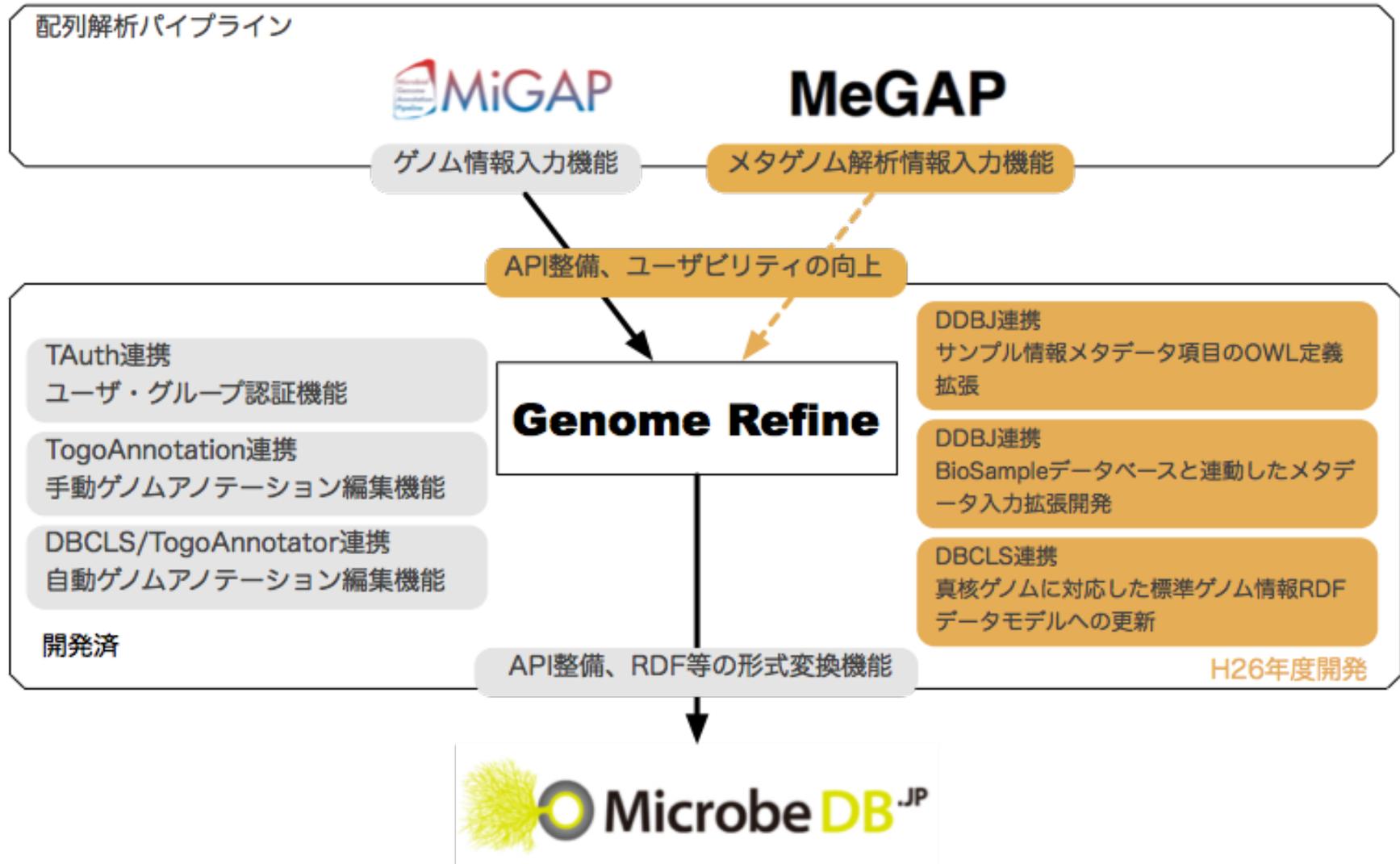
Anatomy Name	Hit Value
<input type="checkbox"/> Feces (unvisualized)	0.84051325986689
<input type="checkbox"/> Milk (unvisualized)	0.585317462682724
<input checked="" type="checkbox"/> Gut	0.362513891694809
<input checked="" type="checkbox"/> Arm	0.110064819613473
<input checked="" type="checkbox"/> Cecum	0.093869656324387
<input checked="" type="checkbox"/> Colon	0.090222145120303
<input checked="" type="checkbox"/> Right external naris	0.045853955671191
<input checked="" type="checkbox"/> Left external naris	0.043376094428822
<input checked="" type="checkbox"/> Leg	0.032520323991776
<input checked="" type="checkbox"/> Ear	0.02869155324353
<input checked="" type="checkbox"/> Nose	0.027815662800968
<input type="checkbox"/> Portion of mucus (unvisualized)	0.022045172407877
<input type="checkbox"/> Mucosa (unvisualized)	0.015296295285225
<input checked="" type="checkbox"/> Right popliteal fossa	0.013866629324515
<input checked="" type="checkbox"/> Hair	0.013750255660852
<input checked="" type="checkbox"/> Left palm	0.012587714675729
<input checked="" type="checkbox"/> Canal for right auditory tube	0.01234339691499
<input checked="" type="checkbox"/> Throat	0.011534477907475
<input checked="" type="checkbox"/> External nose	0.009956159861758

reload

Check All

Release All

# MicrobeDB.jpへのゲートウェイGenome Refine拡張開発



メタゲノム解析パイプラインMeGAPデータ入力について対応するとともに、ユーザビリティの向上を目的としたウェブリソースを整備し、データの公開・再利用の促進のための強化を図る

# 【研究テーマ】工業排水のメタゲノム解析

1. 工業廃水における細菌群集構造は？

→ 細菌群集構造を明らかにする

2. サンプル

→ サンプル

3. どのよ

つ？

→ 環境

4. 同様の

→ 同様の

5. 同様の細菌群集はこのよ様な環境に存在する？

→ 同様の細菌群集構造を示すサンプルとの比較

水処理関連企業A:  
分析ソリューション事業

- 分析サービスの差別化
- 膨大な分析結果を視覚化

H27年度から事業化

相関を持

る？