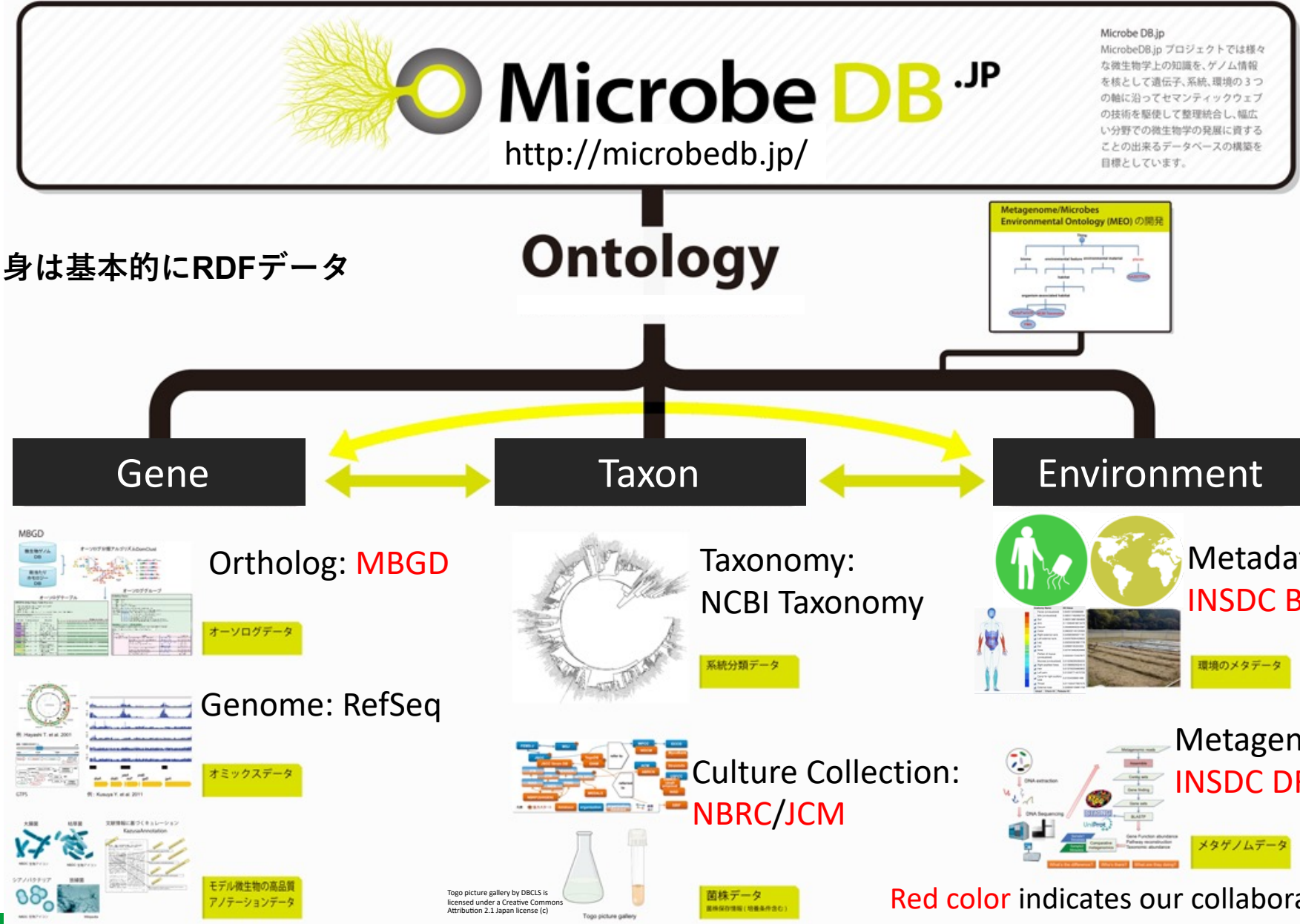


○森宙史¹、藤澤貴智¹、西出浩世²、矢口貴志³、高橋弘喜³、中川善一⁴、山田拓司⁴、内山郁夫²、中村保一¹、黒川顕¹

1. 情報・システム研究機構 国立遺伝学研究所、2. 自然科学研究機構 基礎生物学研究所、3. 千葉大学真菌医学研究センター、
4. 東京工業大学生命理工学院

MicrobeDB.jp integrates microbial data that can be linked to **genomes**. DB has opened **since 2011**



<https://microbedb.jp/>**MicrobeDB.jp**

Integrating and representing genome, metagenome, taxonomy resources and the analysis datasets with Semantic Web Technologies.

[Learn more >>](#)

Features

Data sources of MicrobeDB.jp ver. 3

Metagenome and Microbes Environmental Ontology 2401

Taxonomy 129342

Ortholog Groups 4203173

Microbial Phenotype Ontology 277

Genome and Metagenome Sample 1920339

Culture collections in Japan 38414

Pathogenic Disease Ontology 387

Human Microbiome Associated Disease Ontology 305

KEGG Orthology 22421

Last Modified date: 2020-02-16

🔍 Keyword Search

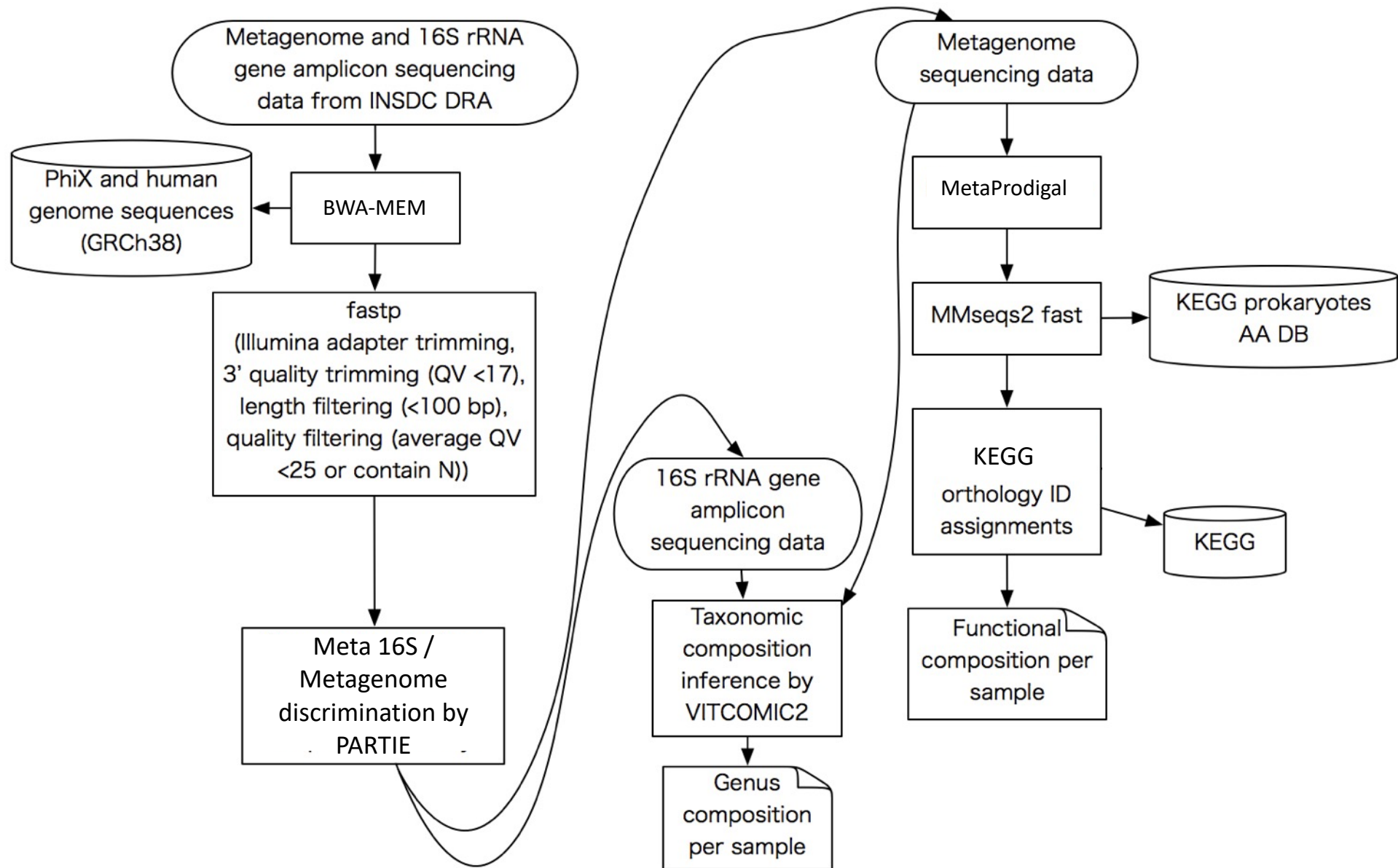
MicrobeDB.jp provides a keyword search function with a simple interface. The keyword search gives the user free-text access to the literal fields of all RDF/OWL resources on MicrobeDB.jp. Click

MicrobeDB.jp version 3 data

Data category	Number of entry
Prokaryote genome metadata (from RefSeq)	290,208 genomes
Culture collection strain metadata from JCM/NBRC (from RDF-Portal)	38,414 strains
Microbiome metadata (from INSDC BioSample)	1,631,611 samples
Microbiome taxonomic composition data	96,766 samples
Microbiome functional composition data	4,784 samples

Version2と比べて、ゲノムは約20倍、メタゲノムは約10倍

MeGAP3 (MetaGenome Analysis Pipeline for MicrobeDB.jp ver. 3)



Search and compare microbiome samples from soil

MicrobeDB.jp Home Document Analysis e.g. hot spring, Enterococcus faecalis, psbA Search Sign Up Sign In

facet_metagenome_public:8531

hasMetagenomeAnalysis
☒ taxonomy 8531
☐ function 119

Search id ...

attribute name
Search attribute name ...

attribute value
Search attribute value ...

hasMEO (Text)
soil x

hasMEO: Component
Component for environment 8102

hasMEO: Env
Environment for microbes 7117

taxonomy (Text)
Search taxonomy terms ...

taxonomy
root 8531

hasHostTaxonomy (Text)
Search HostTaxonomy...

hasHostTaxonomy
root 1410

pH
0 14

Temperature
-100 150

HMADO (Text)
Search HMADO terms ...

HMADO
Human microbiome associated 87 disease

HostEthnicity
Search HostEthnicity ...

Metagenomic samples

8531 results found in 155ms

hasMetagenomeAnalysis: taxonomy x hasMEO (Text): soil x Clear all filters

Previous 1 2 3 4 ... Next

10 Select All Deselect All

Select	MDB SampleID	title	organism.name	organism.identifier	BioProjectID	SRAID	SRRID	BioSampleID	publishedDate
Remove	SAMD00003586	Urease gene-containing Archaea dominate autotrophic ammonia oxidation in two acid soils	soil metagenome	410658	PRJD81924	DRS001577	DRR002212	SAMD00003586	2012-07-04T00:00:00.000
Remove	SAMD00009749	Active ammonia oxidizers in an acid soil are phylogenetically closely related to neutrophilic Nitrososphaera viennensis	soil metagenome	410658	PRJD82274	DRS012638	DRR014314	SAMD00009749	2013-10-30T00:00:00.000
Add	SAMEA1559038	RMgM_Sample_16S	soil metagenome	410658	PRJB3363	ERS184934	ERR186224	SAMEA1559038	2012-11-01T00:00:00.000
Remove	SAMEA1559037	RMR_Sample_16S	soil metagenome	410658	PRJB3363	ERS184936	ERR186222	SAMEA1559037	2012-11-01T00:00:00.000
Add	SAMEA1559036	RMgR_Sample_16S	soil metagenome	410658	PRJB3363	ERS184935	ERR186225	SAMEA1559036	2012-11-01T00:00:00.000
Add	SAMEA1559035	RRR_Sample_16S	soil metagenome	410658	PRJB3363	ERS184933	ERR186223	SAMEA1559035	2012-11-01T00:00:00.000
Remove	SAMD00018981	MIMARKS Survey related sample from rhizosphere metagenome	rhizosphere metagenome	939928	PRJD82986		DRR021946 DRR021947 DRR021948	SAMD00018981	2015-01-16T00:00:00.000
Add	SAMN02054434	MIMARKS Survey related sample from Soil metagenome		410658	PRJNA198445	SRS416341	SRR835396 SRR835397 SRR835398	SAMN02054434	2013-04-23T00:00:00.000
Add	SAMN02054433	MIMARKS Survey related sample from Soil metagenome		410658	PRJNA198445	SRS416342	SRR835399 SRR835400 SRR835401	SAMN02054433	2013-04-23T00:00:00.000
Add	SAMN02054432	MIMARKS Survey related sample from Soil metagenome		410658	PRJNA198445	SRS416340	SRR835402 SRR835403 SRR835404	SAMN02054432	2013-04-23T00:00:00.000

Metagenome sample comparison analysis 4

Taxonomic composition (bar) Taxonomic composition (heatmap) Diversity index Hierarchical clustering PCoA Functional composition (bar) Functional composition (heatmap)

Samples Taxonomic Composition

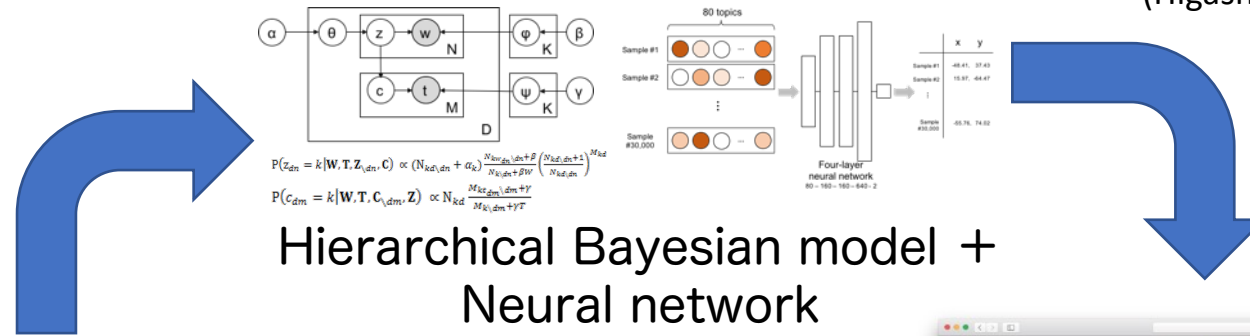
Composition rate

SAMD00003586 SAMD00009749 SAMD00018981 SAMEA1559037

1/51

Latent Environment Allocation (LEA)

(Higashi K et al. 2018, PLoS Comp Biol)



<http://leamicrobe.jp/>

Taxonomic composition
calculated from
DNA sequence

MicrobeDB.jp

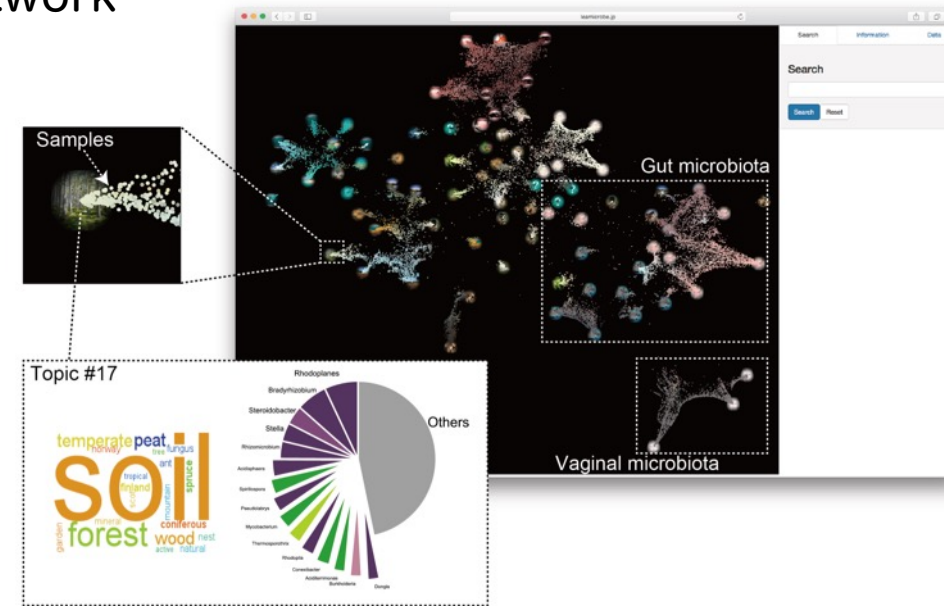
Natural language data
related to
environments

Study Detail

Title	Metatranscriptomic analysis for eukaryotic enriched forest soil sample. A total approach. This method is widely as
Study Type	Metatranscriptomics
Abstract	To explore the eukaryotic functions for eukaryotic enriched forest soil sample. A total approach. This method is widely as
Description	The aim of this study is the exploration of eukaryotic functions for eukaryotic enriched forest soil sample. A total approach. This method is widely as
Center Name	AIST

Center Name AIST

Center Name AIST



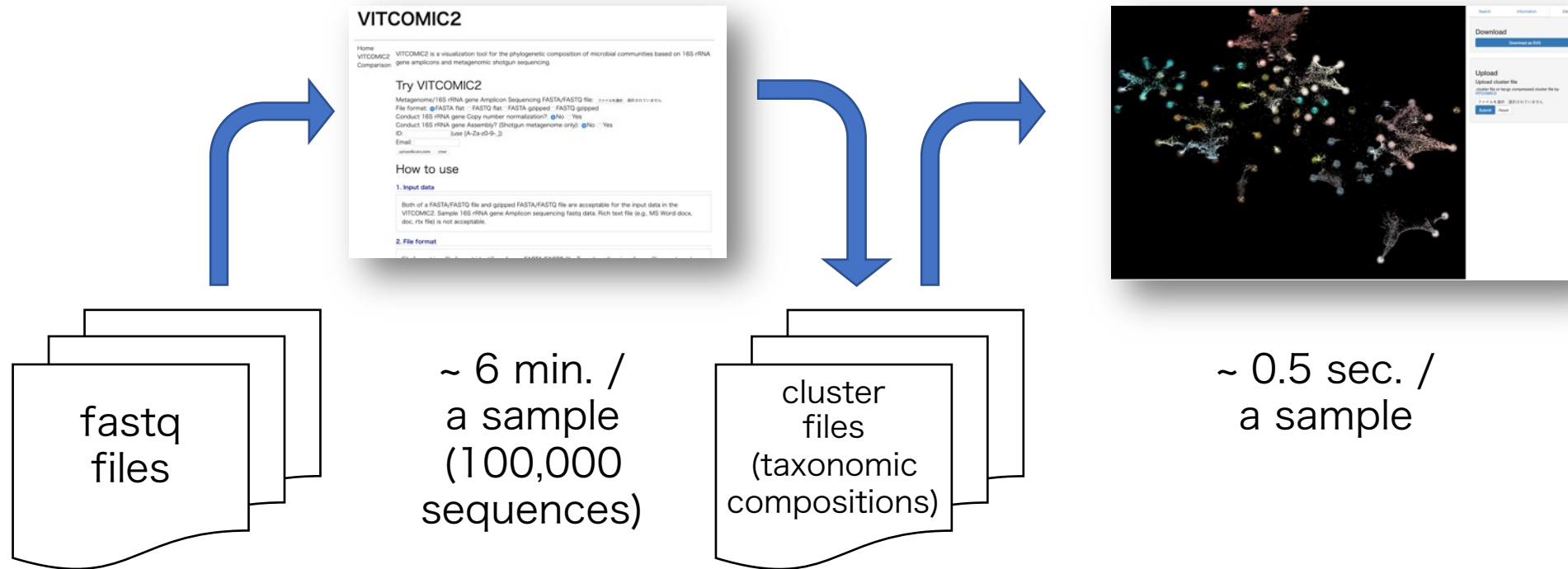
Visualization of microbial community compositions with environment topics

How to map your metagenome data on LEA

(Mori H et al. 2018, BMC Syst Biol)

VITCOMIC2
(<http://vitcomic.org/>)

LEA
(<http://leamicrobe.jp/>)



Ultrafast amino acid sequence similarity search web server against metagenome sequences (PZLAST)

(Mori et al. 2021, Bioinformatics)

<https://pzlast.riken.jp/meta>

30 MEO IDs associated with 59 samples in top 5447 hits.

Click MEO labels to jump to MicrobeDB.jp!

PZLAST

Query Result History Help



Discover sequences from microbial life on this planet.

- What is PZLAST?
- How to use PZLAST
- PZLAST API
- Example output!

Server status at
2020/12/01 21:21 (UTC):
Waiting jobs = 0
Running jobs = 0

Input (Multi-)FASTA protein sequences

```
>LC068768.1 Bifidobacterium breve FL-SBP gene for fucosyl-  
tose ABC transporter substrate-binding protein  
MKKKSRLVAAIAALAMTAGISACGSDTNGSQAKSEVTAKESALTDTSKMVELTVWAY  
SAKQMEPTVKAFFEEKYPHIKFVNTGAADHFTKFNQVQANKIDPVVQMSANKFQF  
AVSGALLNFANDSIDKMSKLYTKTAWQVHYAGGLYGVQDATFLANYVRDILDEHNL  
QVPESWEDIYNQGVKLEKEDSNKYMGILGSDSIFTNLYRSVGARLWKVNSVNDVELTMN  
SGKAKEFTFELQKCLDKGVLEGTVTDEFNRSVNDGRYATFISENWMGMKYDKQNPSLK  
GKMVVTTPSWNNHFQSSSVGSMMSVAAACPKKQAAALAFINWLDSPDAISQWQDNT  
SGNYFMAASVYQDDKTVREKKETDGYFANTDVNSVYFDSMDNVNTDWEYLPFMSQVEVVF  
NDVIVFEMNENGDLVGAKAKQKQKAYAEADNGFKVTTDAD  
>WP_162105629.1 Rhodovulum sulfidophilum chlorophyllide a re-  
ductase subunit Y  
MTEEVKYDEAHTAEVIEGTPREAGNPFEGADPRLDGMGCHAGAAEMEKARAAGASEIL  
DQYAKDYPQPHDRPQSMCPAFGSLRVGLRMRRVATVLSGSAACCVYGLSFVSHFYGARRS  
ICVYFESSETLVTKGLFERIDAVHELADPDYDAVVVNLCVPTASGVFLRLPDEING  
VRIVGIDVPQGVPHAEAKDVLGAMLYARKEIEAGPVAAPGEGVSDRPTVTLGEMF  
PADPIVIGQMLDPLGLACGPPVFAREWRELYAALDCGAVAAIHFFYTAAWREFFDAAGRM  
VGSAPVGHDTAAWLANIGDAFLGSLADKAVAKNMFLPTIREALAGAPVKGITLISGYEG  
SELLVARLLIESGANIPYVGTACPRTKAEADRWLESRGVVFVFRASLEDKAAVDATIR  
PDLAVGTTTPVQHAKEGLIPSLYFTNLISARFLMSGAGAGSLAQVVAIAANKGRMASMK  
EFFEGVGEAGTAGVWGDPNLHFEFRALQKQKLDKKRAEKAQEMT
```

or upload FASTA file

ファイルを選択

選択されていません

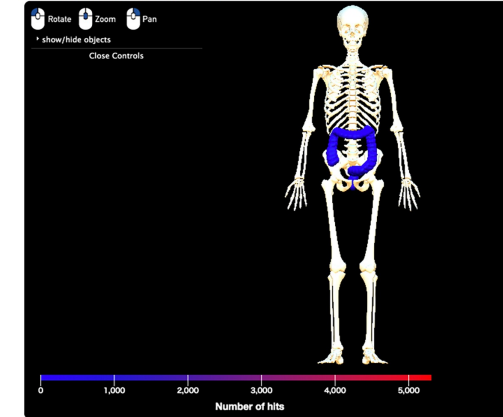
Output top N hits

10000

Submit

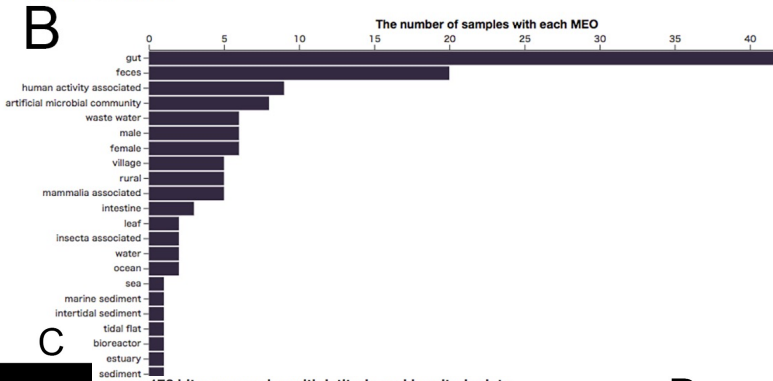
2 FMA IDs found in top 5447 hits.

FMA22905: 5313 hits, FMA14540: 152 hits.



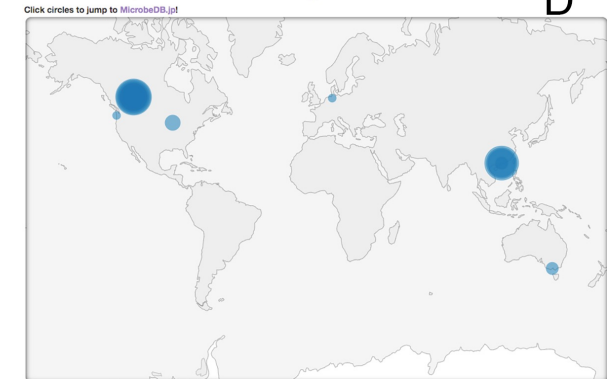
Powered by Three.js and d3.js.
3D model data from
BodyParts3D, © The Database Center for Life Science licensed under CC Attribution-Share Alike 2.1 Japan

Counts Proportions



C

176 hits on samples with latitude and longitude data.



D

Amino acid vs amino acid sequence similarity search against protein sequences of public metagenome data (2.5TB) in MicrobeDB.jp

Current standard of the Fungal microbiome analysis is ...

- **Sequencing strategy:** Amplicon sequencing
- **Taxonomic marker:** ITS
- **Sequencer:** Illumina short reads sequencers
- **Bioinformatics analysis tools:** QIIME2 or DADA2
- **Reference sequence database:** UNITE
- **Taxonomic assignment method:** k-mer frequency

We want to analyze the fungal taxonomic composition in the shotgun metagenomic sequencing data

Mycobiome Taxonomic Assignment Pipeline (MycoTAP)

Genus compositions
of mock community
ITS sequencing data

MycoTAP can also be
applicable to the
shotgun metagenomic
sequencing data

