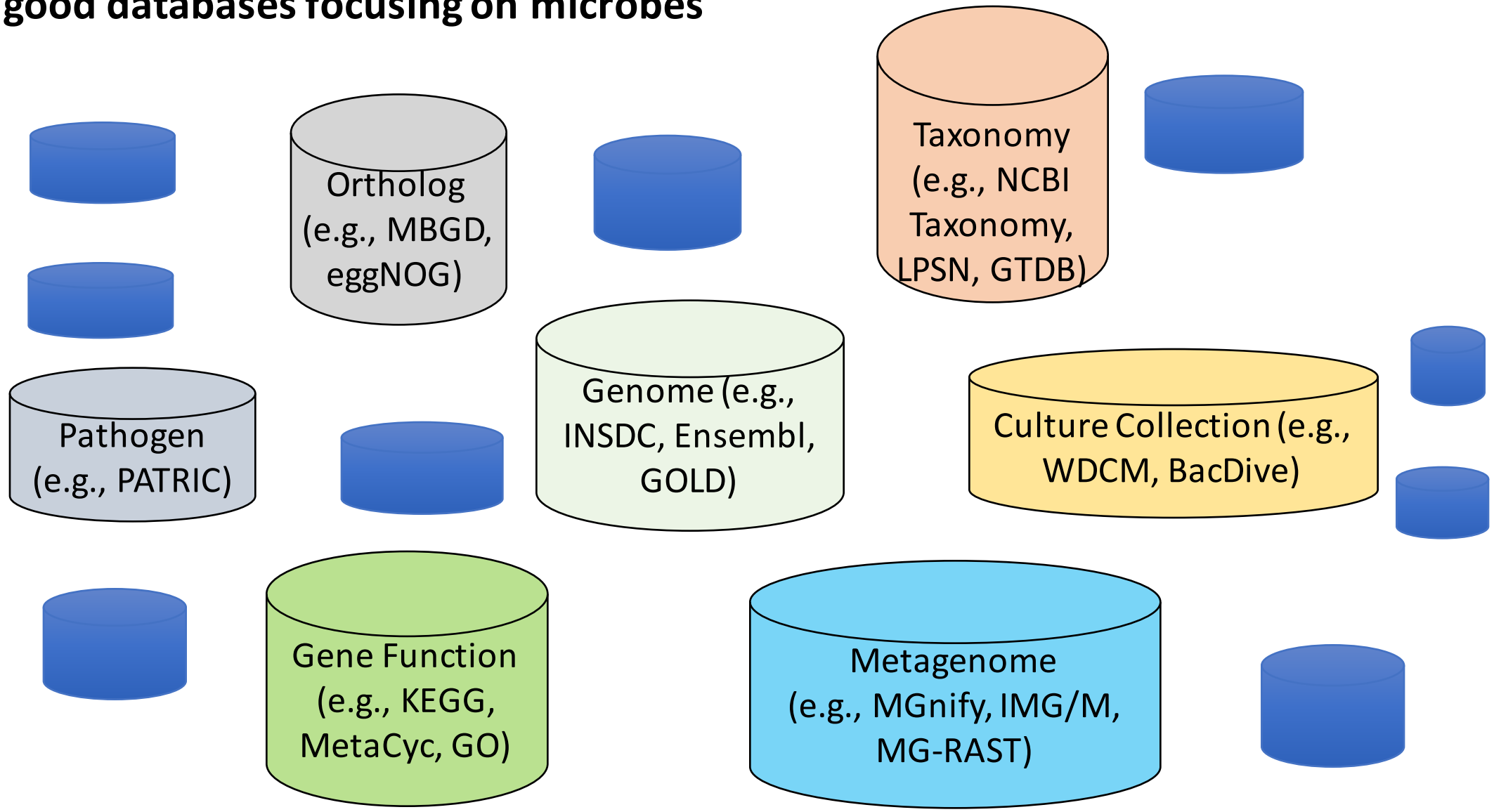
微生物統合データベースMicrobeDB.jp version 3

〇森宙史[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

Licensed under a Creative Commons 表示4.0国際ライセンス
©2020森宙史（国立遺伝学研究所）

Many good databases focusing on microbes



Integration of data is necessary especially for beginners

Licensed under a Creative Commons 表示4.0国際ライセンス
©2020森宙史（国立遺伝学研究所）

Licensed under a Creative Commons 表示4.0国際ライセンス
©2020森宙史（国立遺伝学研究所）

MicrobeDB.jp version 3 data

Data category	Data sources	Ontologies
Genome Metadata	INSDC BioSample, NCBI Assembly report	MEO, HMADO, NCBITAX
Ortholog	MBGD	ORTH
Culture collection	JCM (RDF-Portal)	MCCV, MPO, MEO, PDO, CSSO
Culture collection	NBRC (RDF-Portal)	MCCV, MPO
Metagenome	INSDC DRA	NCBITAX, KEGG Orthology
Metagenome metadata	INSDC BioSample	MEO, HMADO, SIO

Licensed under a Creative Commons 表示4.0国際ライセンス
©2020森宙史（国立遺伝学研究所）

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

Licensed under a Creative Commons 表示4.0国際ライセンス
©2020森宙史（国立遺伝学研究所）

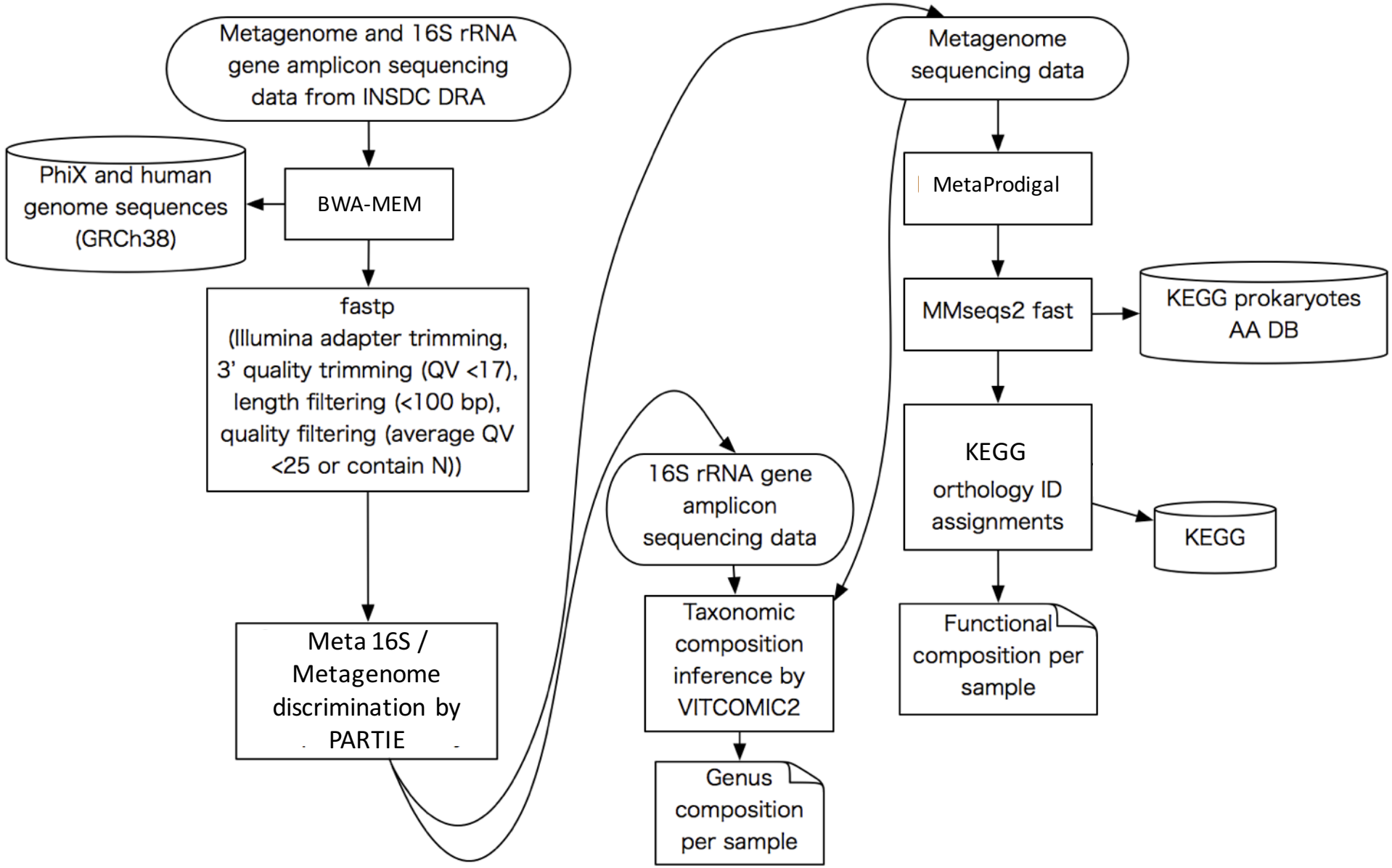
INSDC DRA/ERA/SRAで公開されたマイクロバイオームサンプル数（2020年8月時点）

	サンプル数
マイクロバイオーム	約 2,300,000
ヒト	約 650,000
マウス	約 120,000
土壌	約 280,000
水環境	約 240,000
人工環境	約 160,000
植物共生	約 150,000

230万サンプルはヒトや土壌等マイクロバイオームサンプルの総数

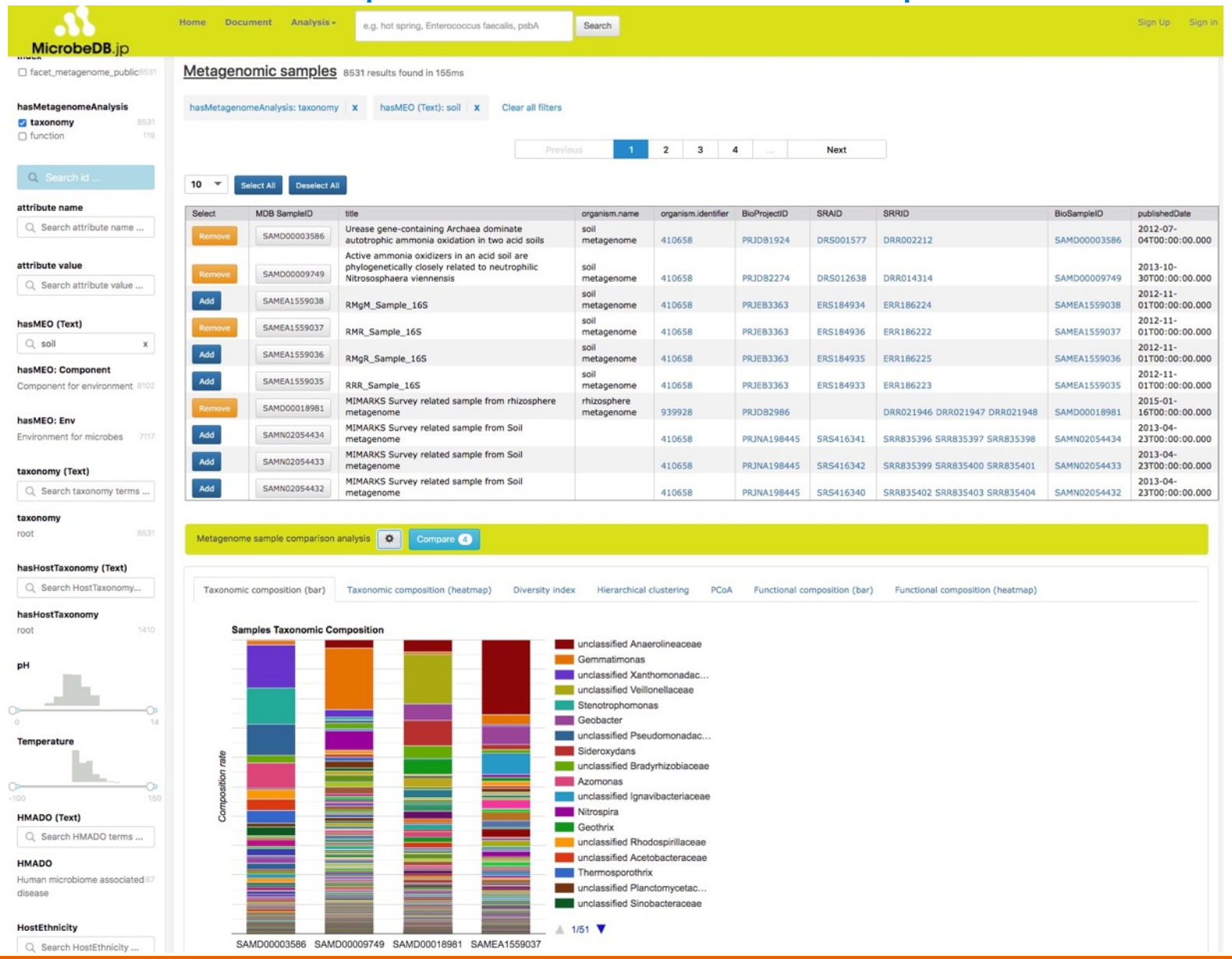
Licensed under a Creative Commons 表示4.0国際ライセンス
©2020森宙史（国立遺伝学研究所）

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



Licensed under a Creative Commons 表示4.0国際ライセンス
©2020森宙史（国立遺伝学研究所）

Search and compare microbiome samples from **soil**



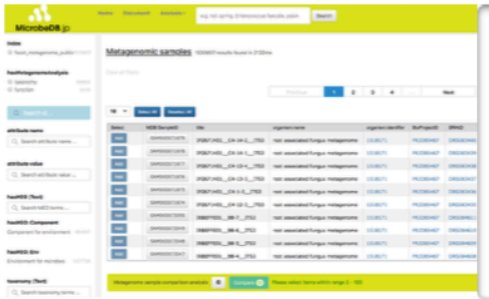
Licensed under a Creative Commons 表示4.0国際ライセンス
©2020森宙史（国立遺伝学研究所）

MicrobeDB.jp ver. 3の使い方

マイクロバイオームサンプルの絞り込み検索

1. Webブラウザで下記にアクセス
https://microbedb.jp

2. 最上段のAnalysis —> Metagenomic samples を選択



3. サンプル（BioSample）の絞り込み検索画面が表示

4. 左側のファセット検索用の選択窓でサンプルを絞り込む（詳細は裏ページ）

5. 絞り込んだサンプルのうち、左上のhasMetagenomeAnalysisで taxonomy（系統組成）か function（遺伝子機能組成）を選択し、それらのデータが計算済みのサンプルのみ表示する

6. サンプルリストから、比較解析したいサンプルを複数選択し、Compareを押す

7. 系統組成等のグラフが表示される

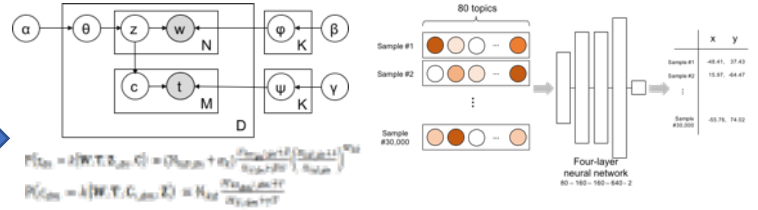
ファセット検索用の検索窓の説明

1. **Search id: BioSample ID**でサンプルを検索
2. **attribute name: RDF**の述語名で検索（玄人向け）
3. **attribute value: RDF**の目的語で検索（玄人向け）
4. **hasMEO (Text):** 微生物の生息環境オントロジーである**MEO**に対するテキスト検索 (例: **soil**)
5. **hasMEO: Component: MEO**の階層構造から**MEO Component**クラスを選択して検索
6. **hasMEO: Env: MEO**の階層構造から**MEO Environment**クラスを選択して検索
7. **taxonomy (Text):** メタゲノムの**NCBI Taxonomy**に対するテキスト検索 (例: **soil metagenome**)
8. **taxonomy: NCBI Taxonomy**のクラスを選択して検索
9. **hasHostTaxonomy (Text):** 宿主の系統について**NCBI Taxonomy**をテキスト検索 (**Homo sapiens**)
10. **hasHostTaxonomy:** 宿主の系統について**NCBI Taxonomy**のクラスを選択して検索
11. **pH:** サンプルが由来した環境の**pH**を範囲指定
12. **Temperature:** サンプルが由来した環境の温度を範囲指定
13. **HMADO (Text):** 微生物が関係するヒトの病気の病名オントロジーである**HMADO**に対するテキスト検索 (例: **cancer**)
14. **HMADO: HMADO**のクラスを選択して検索
15. **HostEthnicity:** サンプルが由来したヒトの**ethnicity**に対するテキスト検索 (例: **Japanese**)
16. **HostAge:** サンプルが由来した宿主の**age**を範囲指定
17. **HostBMI:** サンプルが由来した宿主の**BMI**を範囲指定
18. **AntibioticRegimen:** サンプルが由来したヒトの抗生物質投与の有無を選択

Licensed under a Creative Commons 表示4.0国際ライセンス
©2020森宙史（国立遺伝学研究所）

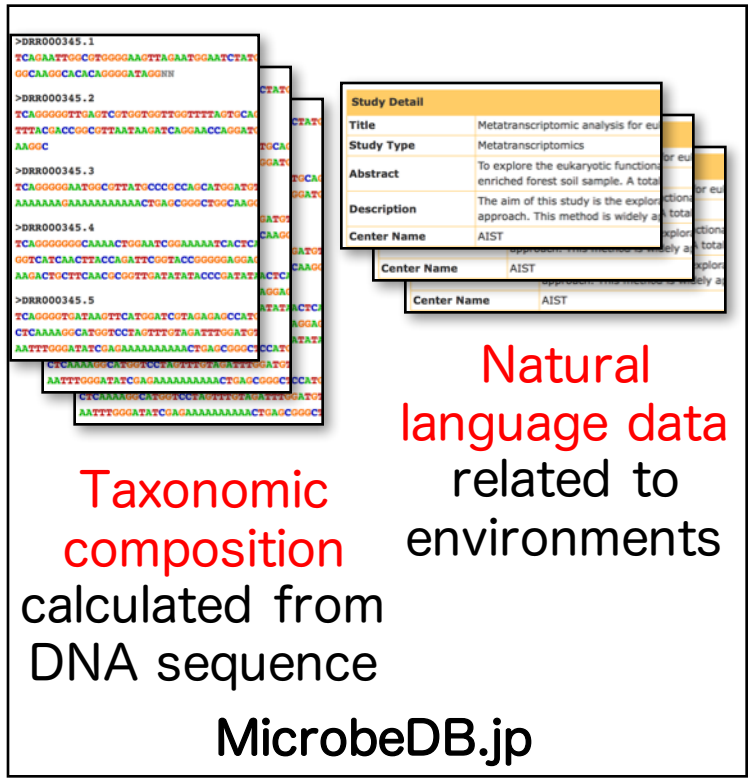
Latent Environment Allocation (LEA)

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



Hierarchical Bayesian model + Neural network

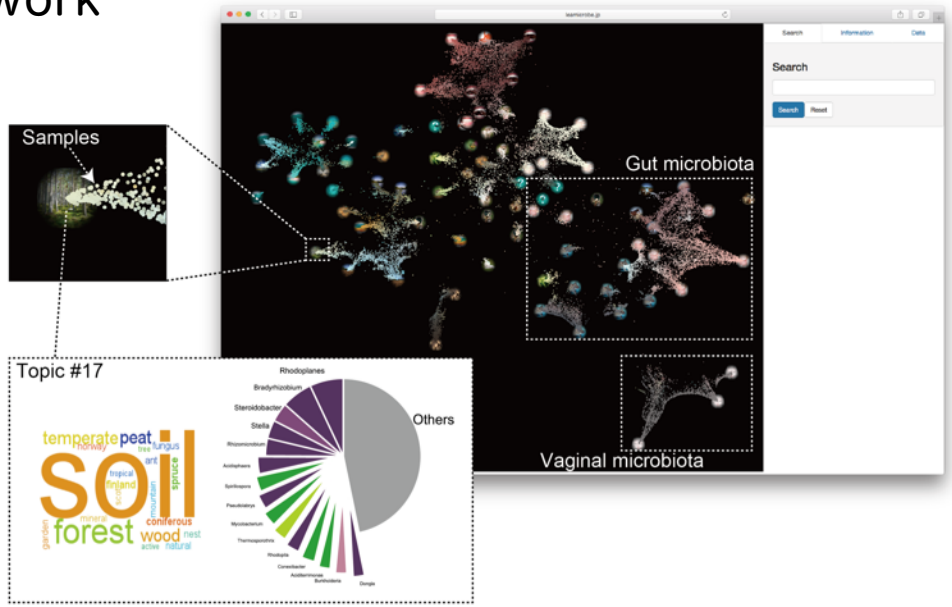
http://leamicrobe.jp/



Natural language data related to environments

Taxonomic composition calculated from DNA sequence

MicrobeDB.jp



Visualization of microbial community compositions with environment topics

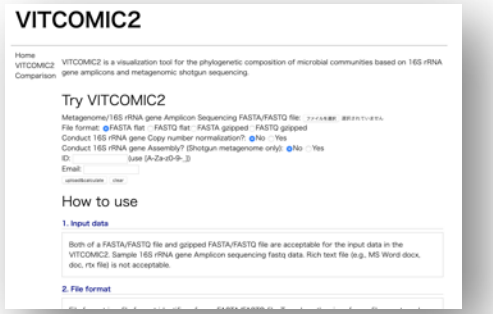
Licensed under a Creative Commons 表示4.0国際ライセンス
©2020森宙史（国立遺伝学研究所）

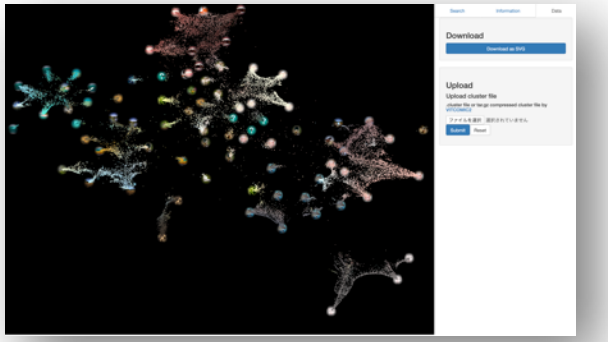
How to map your metagenome data on LEA

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

VITCOMIC2
(http://vitcomic.org/)

LEA
(http://leamicrobe.jp/)





fastq files

~ 6 min. / a sample (100,000 sequences)

cluster files (taxonomic compositions)

~ 0.5 sec. / a sample

Licensed under a Creative Commons 表示4.0国際ライセンス
©2020森宙史（国立遺伝学研究所）