

 **Microbe DB^{.JP}** データを用いた
メタゲノム解析Webアプリケーションの開発

森 宙史
Hiroshi Mori

東京工業大学
大学院生命理工学研究科
生命情報専攻 黒川・山田研 助教

OUR MICROBIAL PLANET

MICROBES—life forms too tiny to see—play a surprisingly large role in life on Earth. Microbes are everywhere, and they do a lot of good for human health and our planet. In fact, disease-causing microbes make up only a very tiny fraction of the millions of types of microbes. Microbes...

Think microbes are bad guys? Think again.



Keep us healthy. Amazingly, only about 1 out of 10 cells in the human body is actually a human cell; most of the cells in our bodies are microbes! Some of the microbes living in our bodies actually help us fight disease-causing microbes by competing against them for space. This mutually beneficial relationship helps to protect us from getting diseases while giving the "good" microbes a place to live.

Make air breathable. Without microbes, we wouldn't have oxygen to breathe. This is because many microbes are photosynthetic—like plants, they harvest their energy from the sun, releasing oxygen into the air. Billions of years ago, photosynthetic microbes gradually added oxygen to Earth's atmosphere, making it possible for larger forms of life—including humans—to live.

Provide sources of new medicines. Hundreds of medicines available today were derived from chemicals first found in microbes. Microbes naturally produce an amazing variety of chemicals, which scientists can use to create new medicines.

Help us digest food. Many of the foods we eat would be indigestible without the 10-100 billion microbes living within our guts. Microbes also play a major role in creating many of the foods we love, such as cheese, yogurt and bread.



Keep our environment clean. Because of their special adaptations, some microbes can help clean up gasoline leaks, oil spills, sewage, nuclear waste, and many other types of pollution.



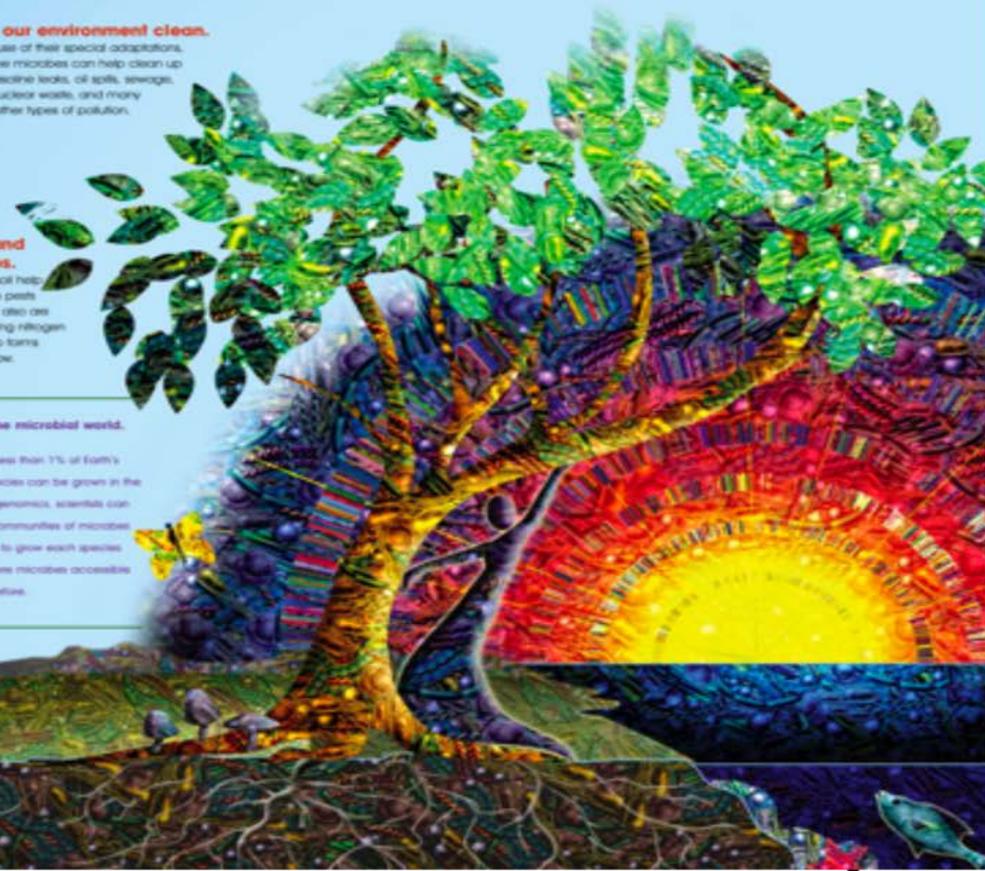
Support and protect crops. Microbes living in soil help protect plants from pests and diseases. They also are essential for converting nitrogen and other nutrients into forms that plants can use to grow.

The science of metagenomics is shedding new light on the microbial world.



Scientists estimate that less than 1% of Earth's millions of microbial species can be grown in the laboratory. Using metagenomics, scientists can now study how whole communities of microbes function without having to grow each species separately—making more microbes accessible to science than ever before.

Living in a microbial world...



Metagenome



This Figure is obtained from National Research Council (USA)

16S rRNA gene amplicon sequencing analysis (メタ16S解析)



DNA extraction

PCR amplification

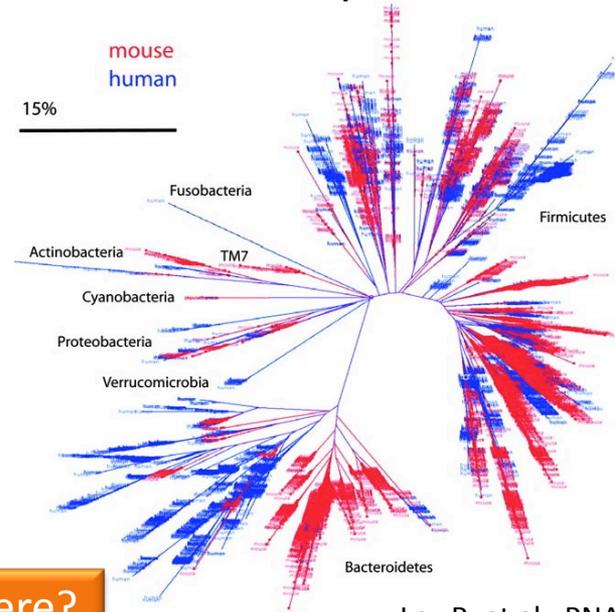
DNA Sequencing

Pre-analysis (Remove Primer, Chimera etc.)

Multiple sequence alignment
or sequence clustering

```
CGTAACAA-GGTAGCCGTACCGGAA-GGTGTGGCTGGATCACCTCCTT  
CGTAACAA-GGTAGCCGTACCGGAA-GGTGTGGCTGGATCACCTCCTT  
CGTAACAA-GGTAGCCGTACCGGAA-GGTGTGGCTGGATCACCTCCTT  
CGTAACAA-GGTAGCCGTACCGGAA-GGTGTGGCTGGATCACCTCCTT  
CGTAACAA-GGTAGCCGTACCGGAA-GGTGTGGCTGGATCACCTCCTT  
CGTAACAA-GGTAGCCGTACCGGAA-GGTGTGGCTGGATCACCTCCTT  
CGTAACAA-GGTAGCCGTACCGGAA-GGTGTGGCTGGATCACCTCCTT  
CGTAACAA-GGTAGCCGTACCGGAA-GGTGTGGCTGGATCACCTCCTT  
CGTAACAA-GGTAGCCGTACCGGAA-GGTGTGGCTGGATCACCTCCTT  
CGTAACAA-GGTAGCCGTACCGGAA-GGTGTGGCTGGATCACCTCCTT
```

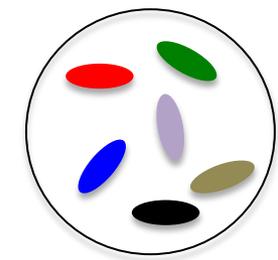
Distance calculation
or similarity search



Who's there?

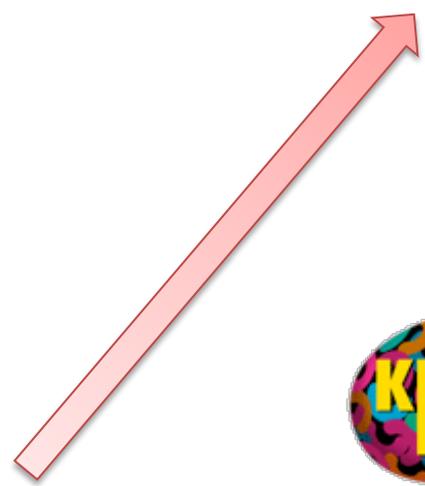
Ley R. et al., PNAS 2007

Metagenomic sequencing analysis (メタゲノム解析)



DNA extraction

DNA Sequencing



Sample1
Metadata

Sample2
Metadata

Comparative
metagenomics

Metagenomic reads

Assemble

SOAPdenovo2, Velvet,
IDBA-UD etc.

Contig sets

Gene finding

MGA, MetaGeneMark

Gene sets

BLASTP

Gene Function abundance
Pathway abundance
Pathway reconstruction
Taxonomic abundance

Who's there?

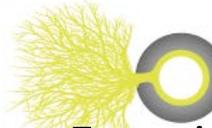
What are they doing?

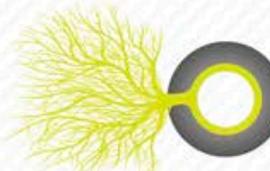
- メタゲノム解析データから、簡便に
系統組成と遺伝子機能組成の情報を抽出したい
- 組成情報を既存のメタゲノム解析サンプルと比較したい



その環境の細菌群集は他のサンプルと比較するとどのような共通性・違いがあるか？

その環境のメタデータと自分のサンプルのメタデータを比べて、何が異なるのか？

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



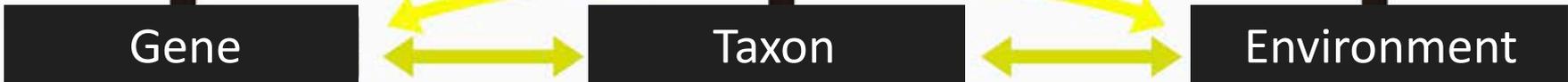
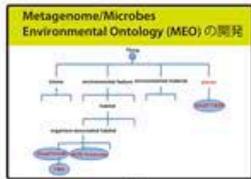
Microbe DB .JP

<http://microbedb.jp/>

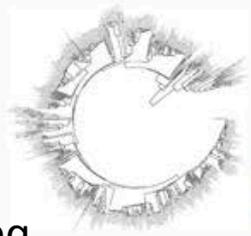
Microbe DB.jp
MicrobeDB.jp プロジェクトでは様々な微生物学上の知識を、ゲノム情報を核として遺伝子、系統、環境の3つの軸に沿ってセマンティックウェブの技術を使用して整理統合し、幅広い分野での微生物学の見解に資することの出来るデータベースの構築を目標としています。

Ontology

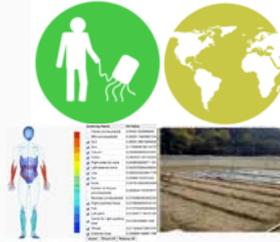
オントロジー: 検索タームの柔軟化&明確化



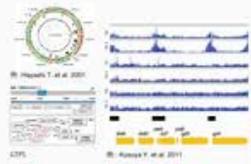
Ortholog: **MBGD**
オソログデータ



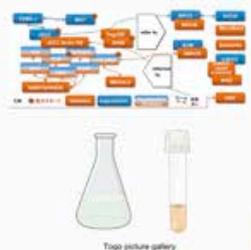
Taxonomy: **NCBI Taxonomy**
系統分類データ



Metadata: **INSDC SRA**
環境のメタデータ



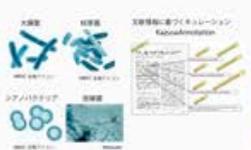
Genome: **GTPS/RefSeq**



Culture Collection: **NBRC/JCM**
菌株データ



Metagenome: **INSDC SRA**
メタゲノムデータ



Annotation: **TogoAnnotation**
モデル微生物の高品質アノテーションデータ

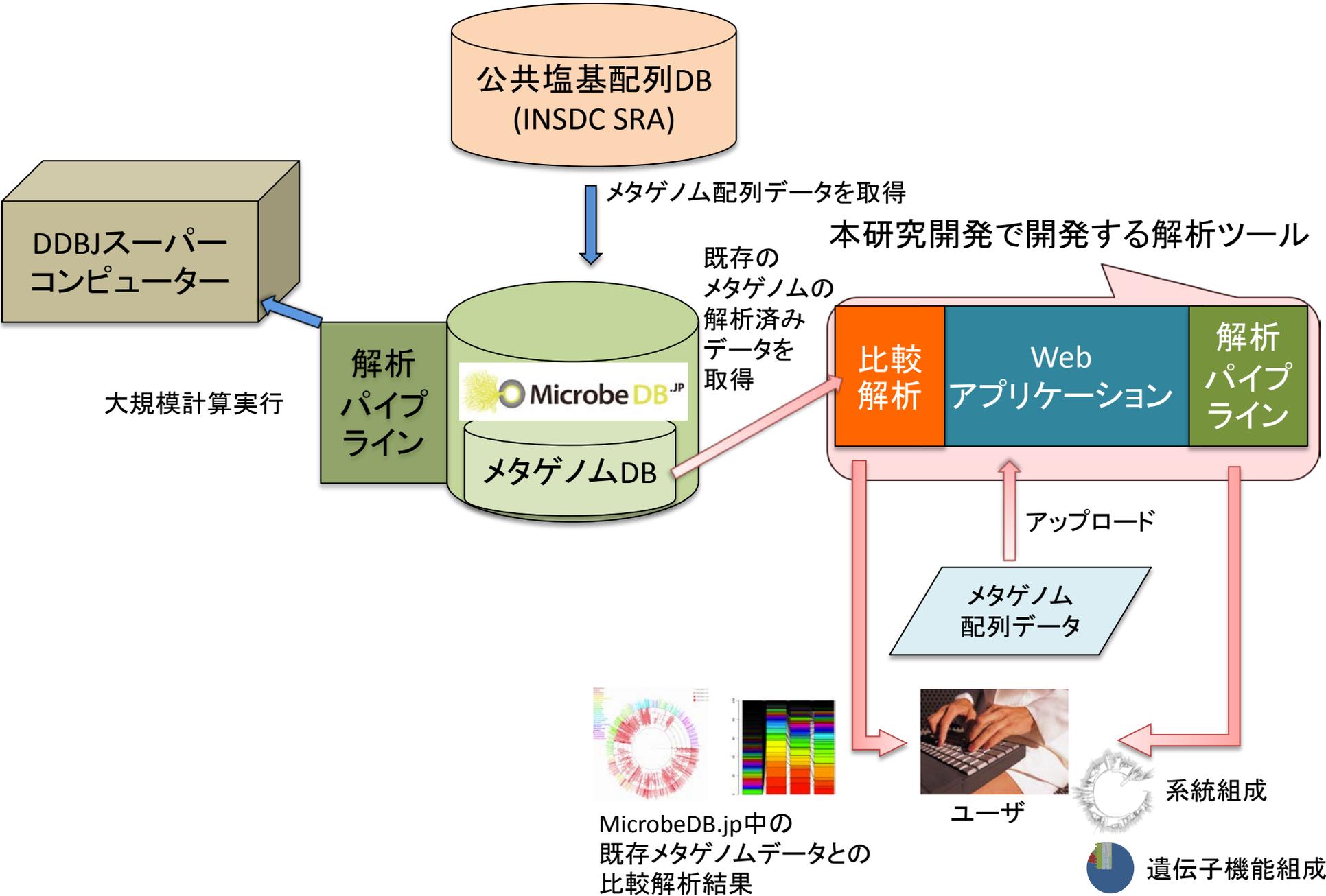
Red color indicates our collaborators.

既存のメタゲノム・メタ16Sデータの整理・統合

微生物の情報とその生息環境の情報を結びつける

1. 公共DB(SRA)からメタ16S・メタゲノム配列データとメタデータを取得
2. 配列データから系統組成・遺伝子機能組成情報を計算
3. メタ16S・メタゲノム論文から微生物の生息環境についてのメタデータを抽出
4. 生息環境を記述するために同義語、語彙間の関連性を記述したオントロジーMEOを構築
5. MEOとメタ16S・メタゲノムのメタデータの対応付け

サンプルの種類	サンプル数
環境メタ16S	2904
ヒトメタ16S	11289
環境メタゲノム	約1100
ヒトメタゲノム	約4900



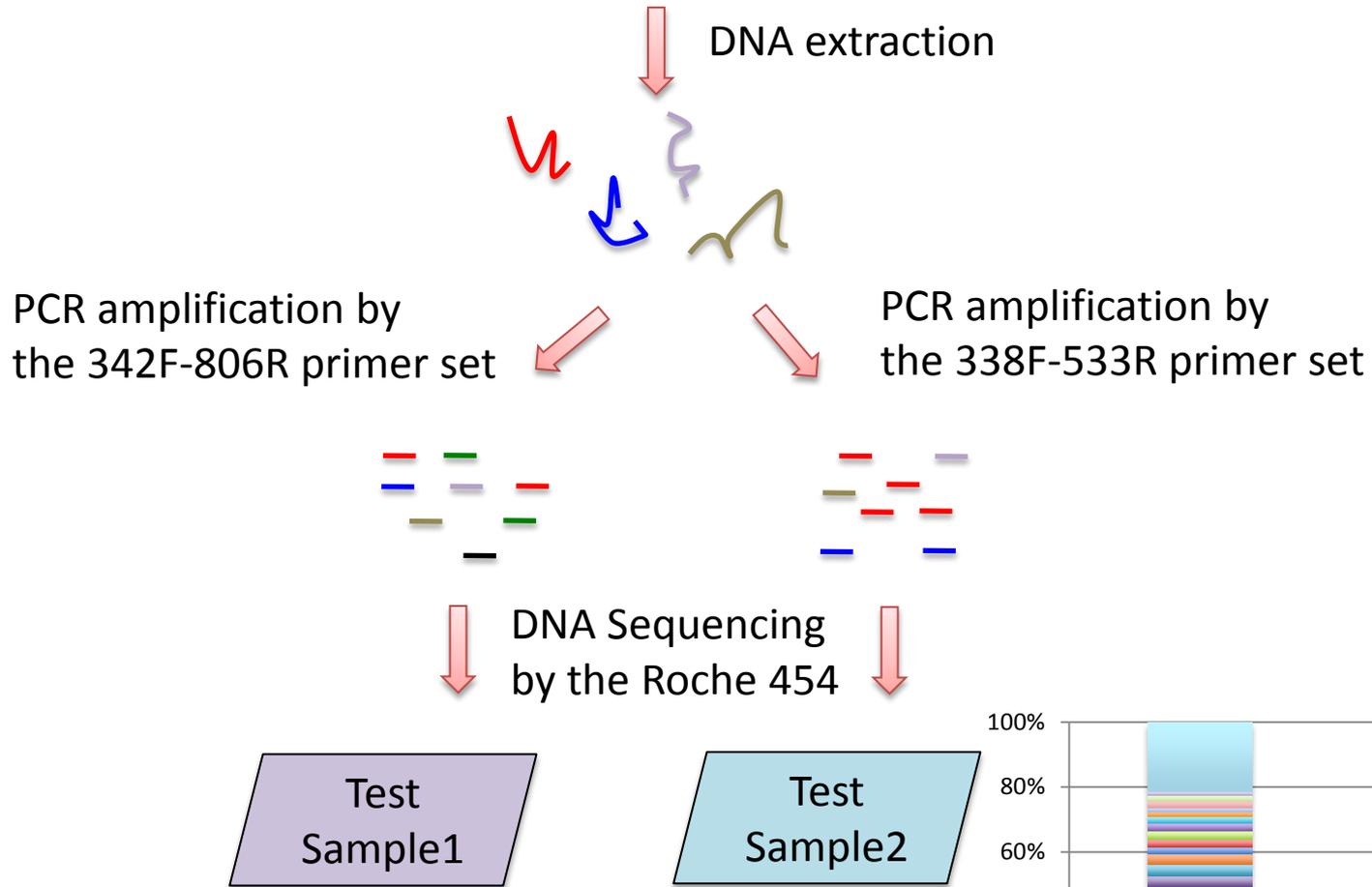
ユーザーのメタゲノム解析データを受け取り、系統組成と遺伝子機能組成を解析し、既存メタゲノムデータと比較解析した結果を返す

テストデータ

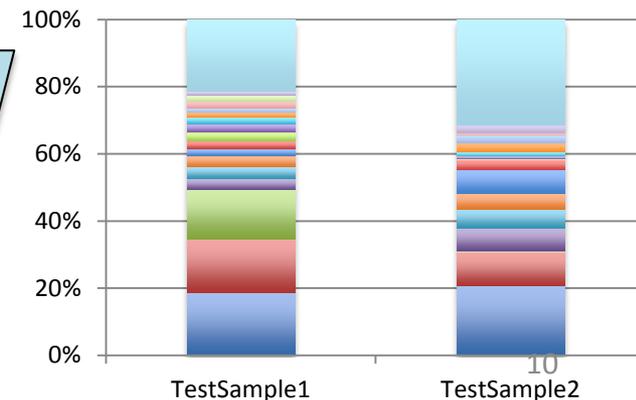
Mori H. et al., 2013, DNA Res.
の2種類のメタ16Sデータを使用



畑の土壌

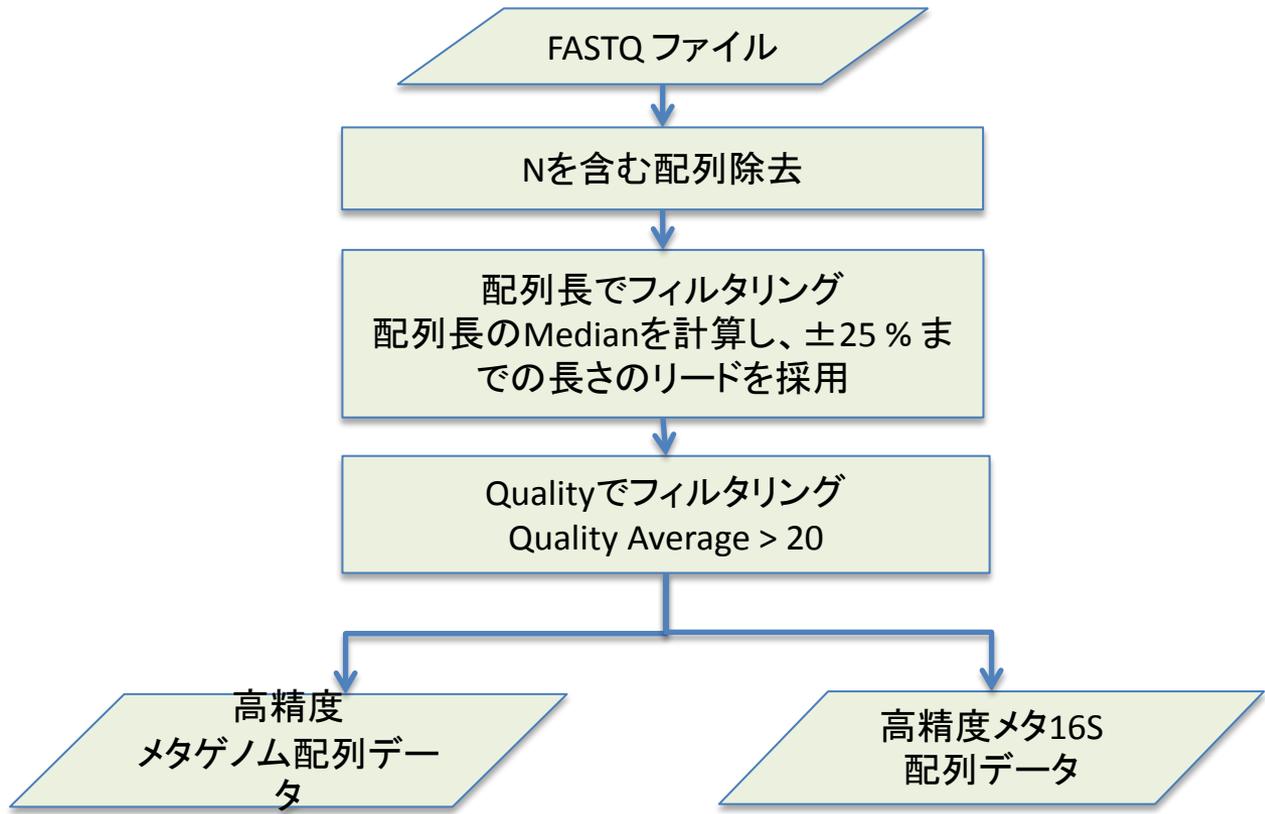


リード数はそれぞれ約29,000本と約15,000本

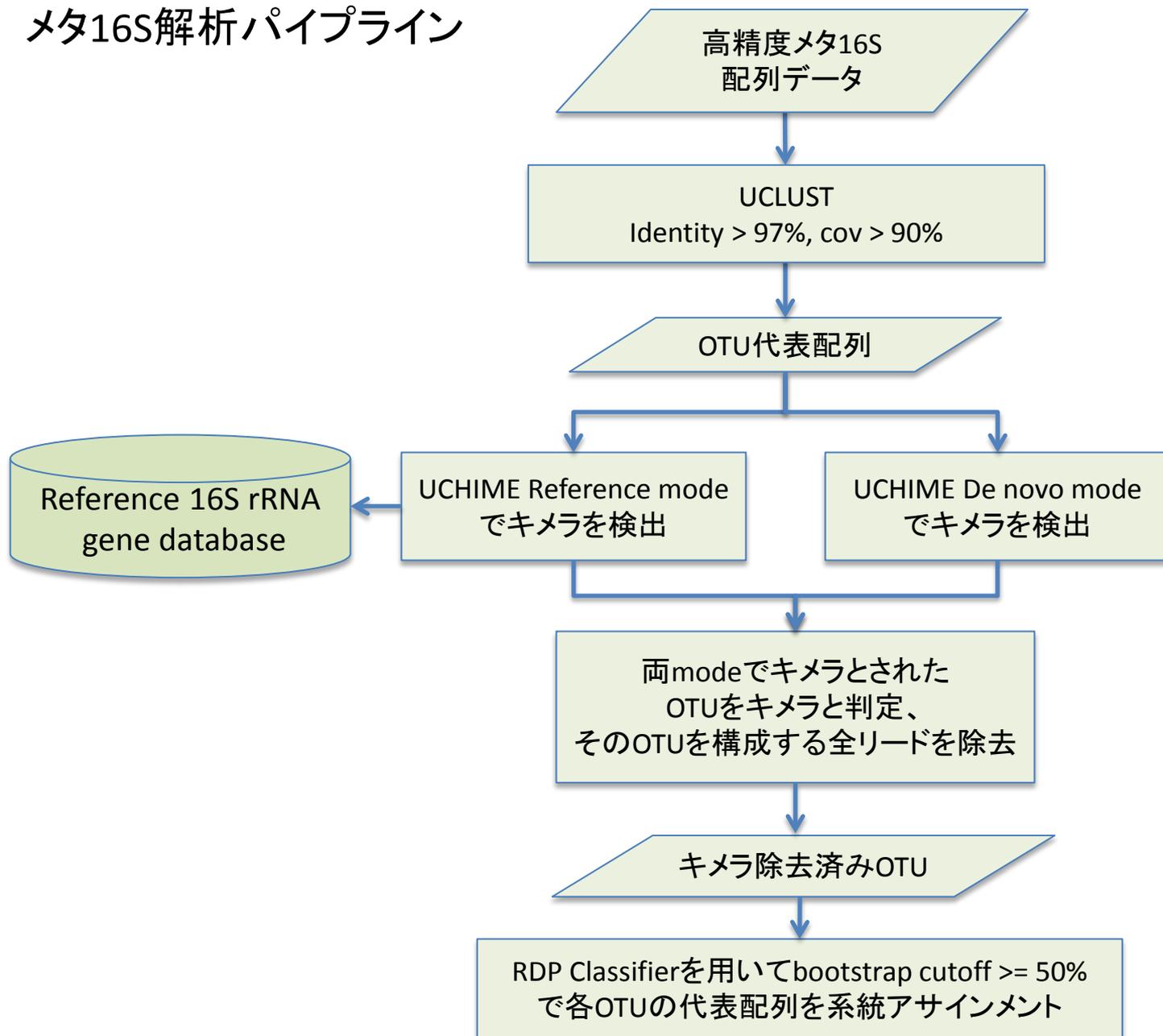


配列の前処理 (目的: メタゲノムとメタ16Sデータの区別 & 高精度配列データの抽出)

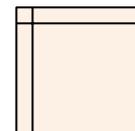
ユーザがメタ16S or メタゲノムの情報を入力



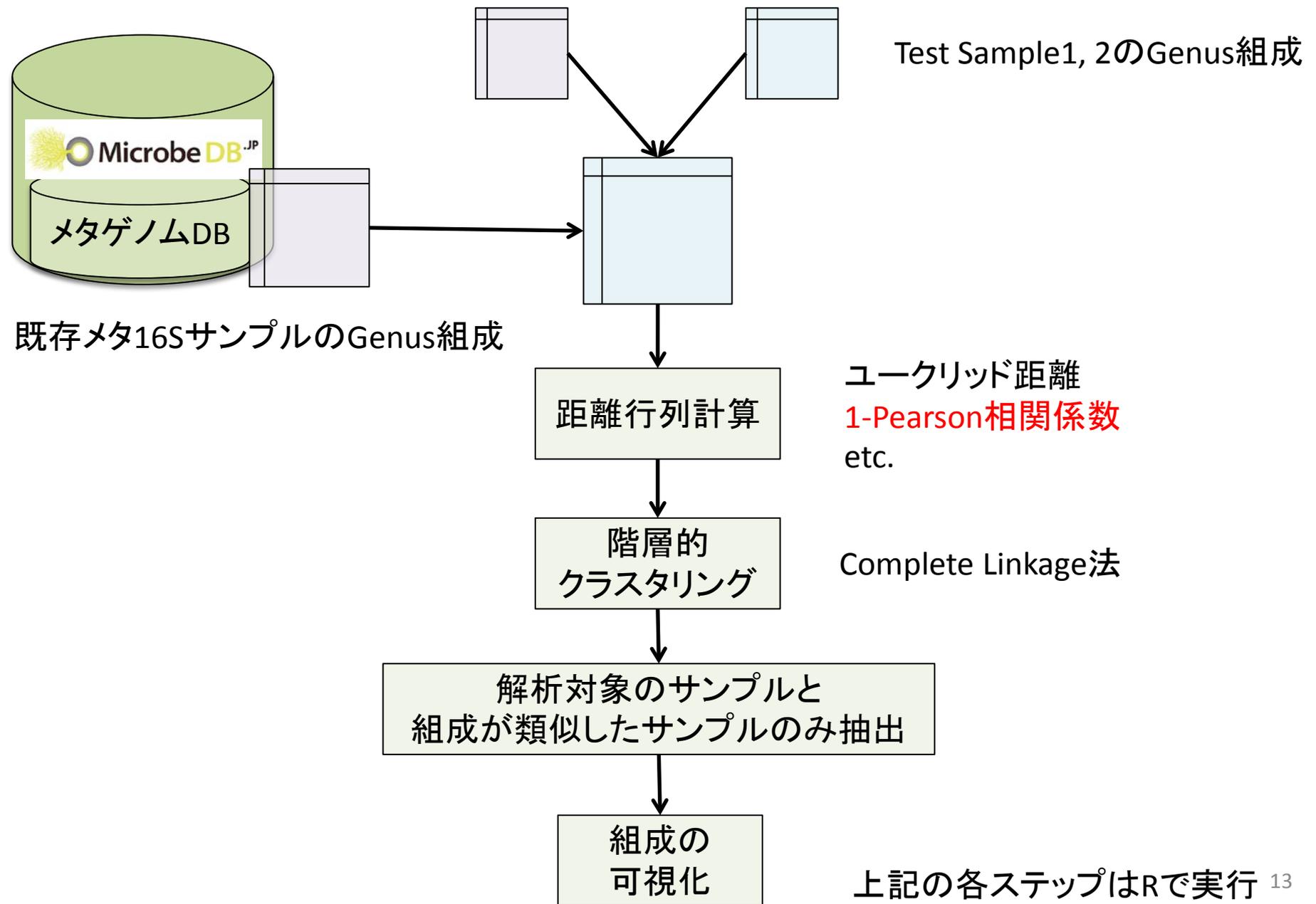
メタ16S解析パイプライン



Genus-Phylumレベルの系統組成

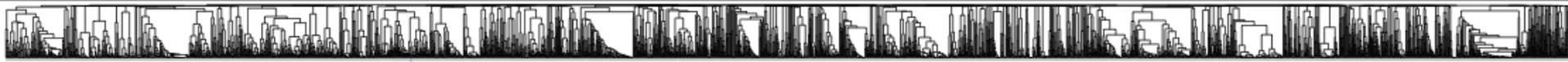


MicrobeDB.jp中の既存サンプルとの比較解析

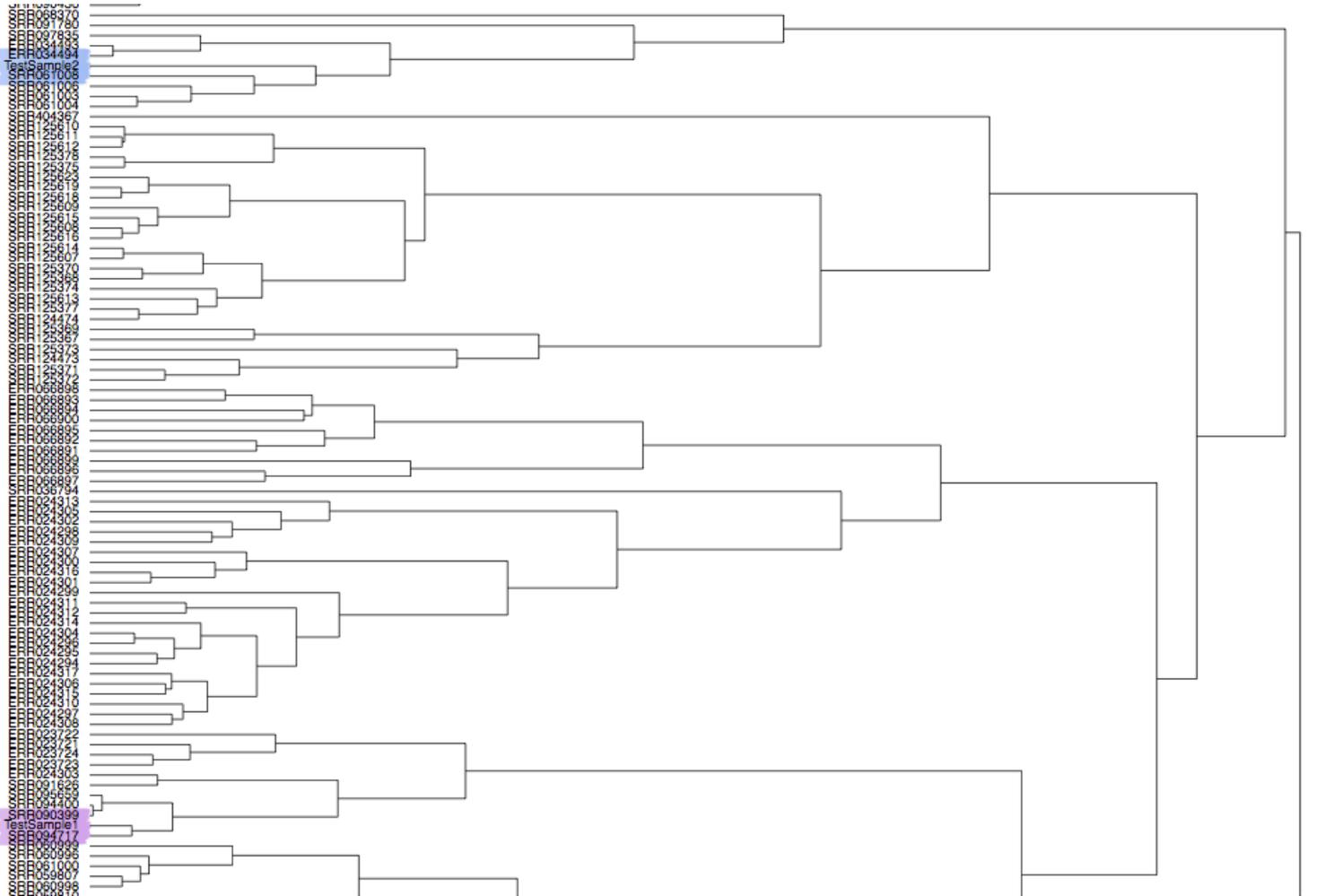


Genus組成によるサンプル間階層的クラスタリングの結果

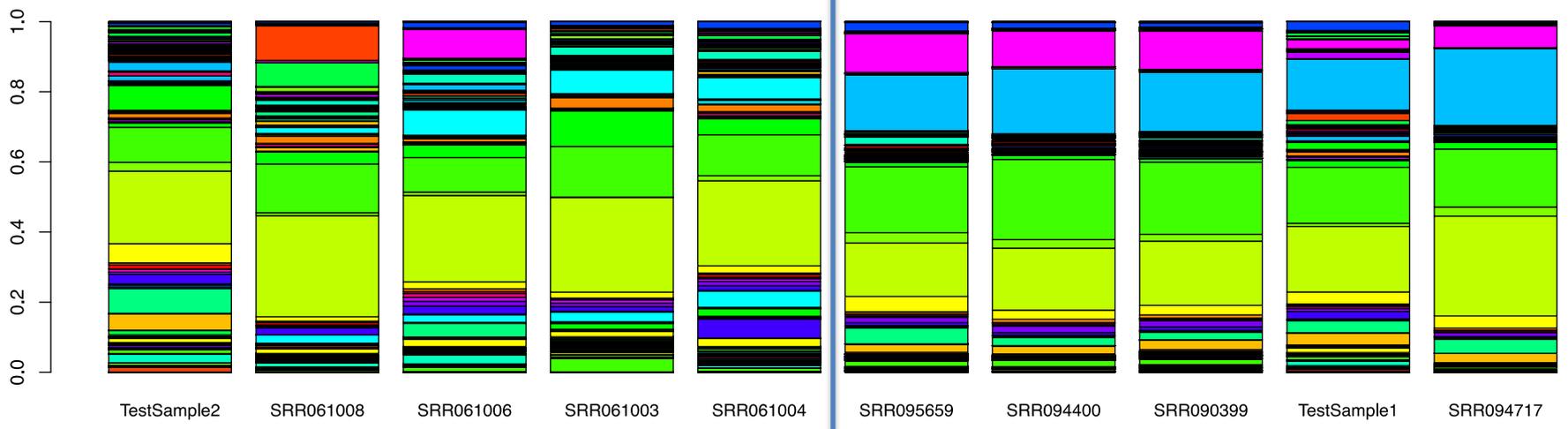
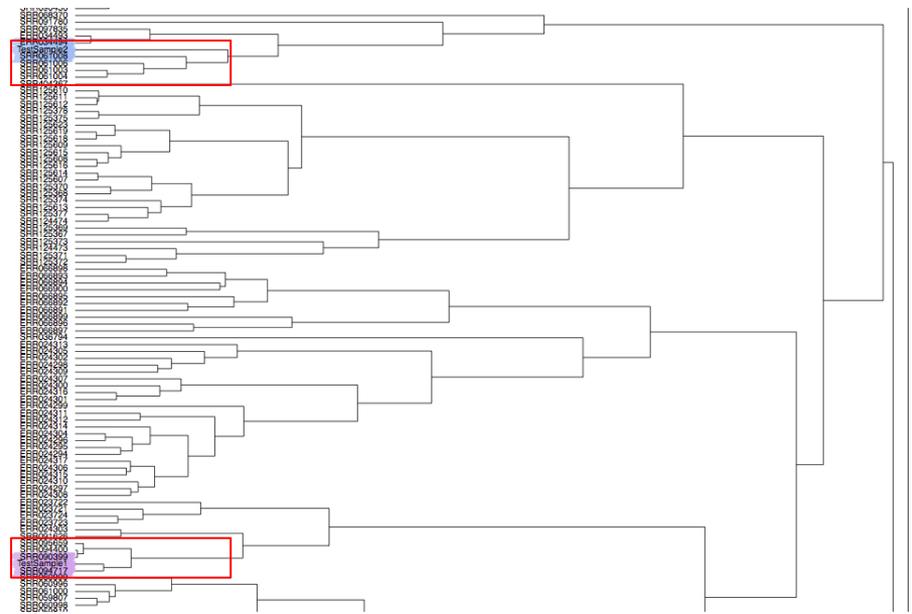
全体像 約3000ノード



Test Sample 1, 2の付近のみ拡大



Genus組成の可視化



MicrobeDB.jp中の既存サンプルの環境メタデータやそのGenusの情報をどう整理して提供する？

1. ユーザが入力した環境メタデータを利用した解析
2. 比較解析機能の充実
3. サーバ上でWebアプリケーションを構築
4. ユーザが持つ巨大な配列データをどう扱うか