

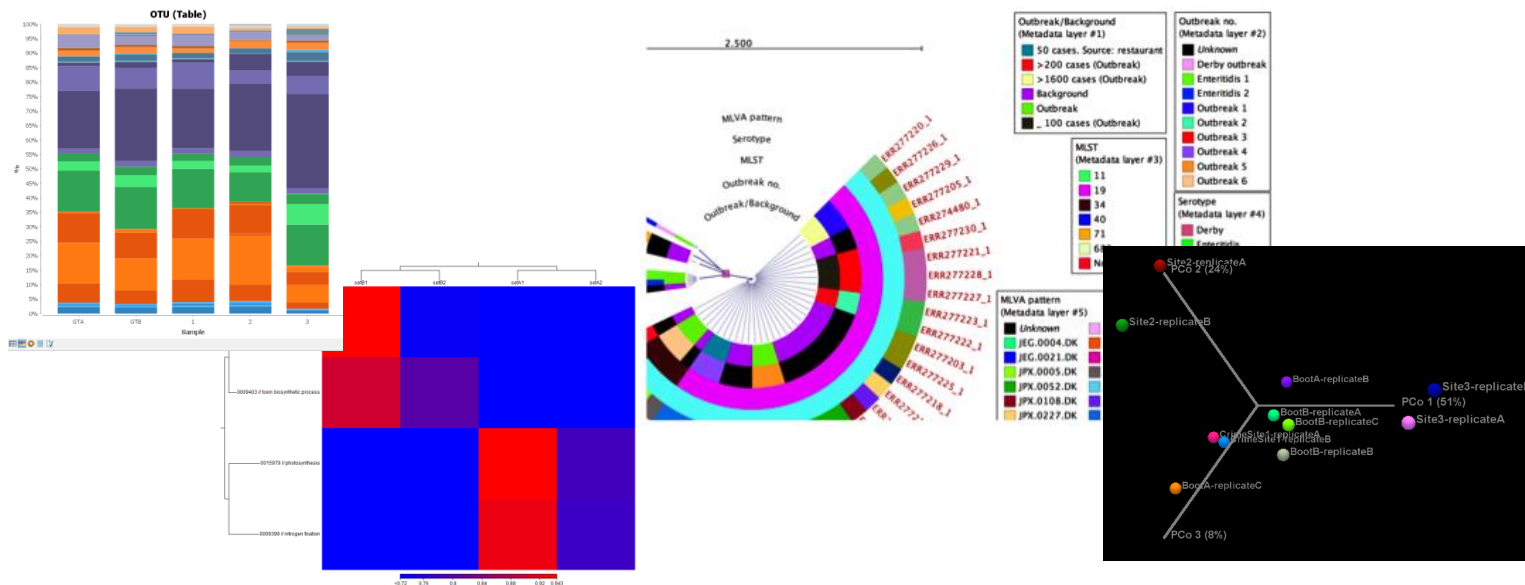
CLC Microbial Genomics Module

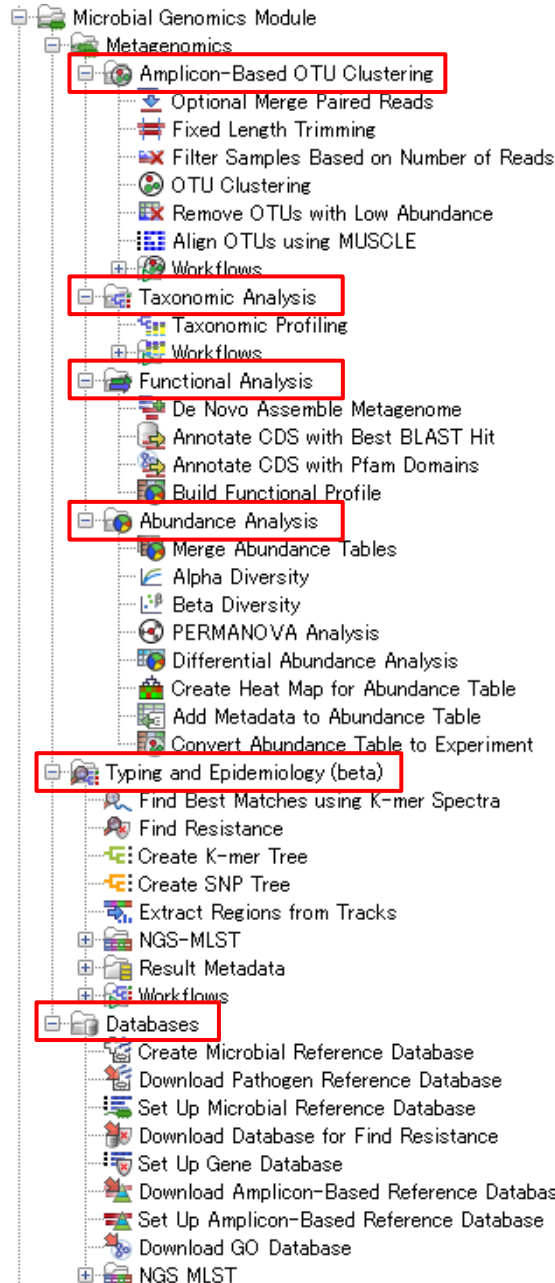
ウェブトレーニングセミナー：微生物プロファイル解析編

フィルジェン株式会社 バイオサイエンス部
(biosupport@filgen.jp)



- CLC bio (QIAGEN社) Workbenchシリーズの微生物ゲノム解析用プラグイン
- 16S rRNAやショットガンメタゲノムデータを用いた菌種組成解析と、病原菌のタイピングや疫学解析ツール、遺伝子機能解析ツールが利用可能になる
- CLC bio Workbench上のデータを、そのまま解析に用いることが可能
- 解析パイプラインのワークフローが最初から組み込まれており、簡単な操作で解析を行うことができる
- バクテリアやウイルスなど、解析に使用するリファレンスゲノムデータなども、専用ツールから簡単にダウンロードが可能





使用可能になるツール群

Amplicon-Based OTU Clustering

- 16S rRNAなどアンプリコンシーケンスデータの各種QCチェックおよびOTUクラスタリングによる菌種組成解析

Taxonomic Analysis

- ショットガンメタゲノムデータを用いた、宿主ゲノム配列データの除去および菌種組成解析

Functional Analysis (別途有償プラグイン「MetaGeneMark」が必要)

- メタゲノムシーケンスデータのDe Novoアセンブル
- BLAST検索、Pfamドメイン検索による遺伝子機能アノテーション付けと組成解析

Abundance Analysis

- 菌種組成データからの α 多様性と β 多様性の計算
- 菌種または遺伝子機能組成データからの、サンプル間比較やヒートマップ作成

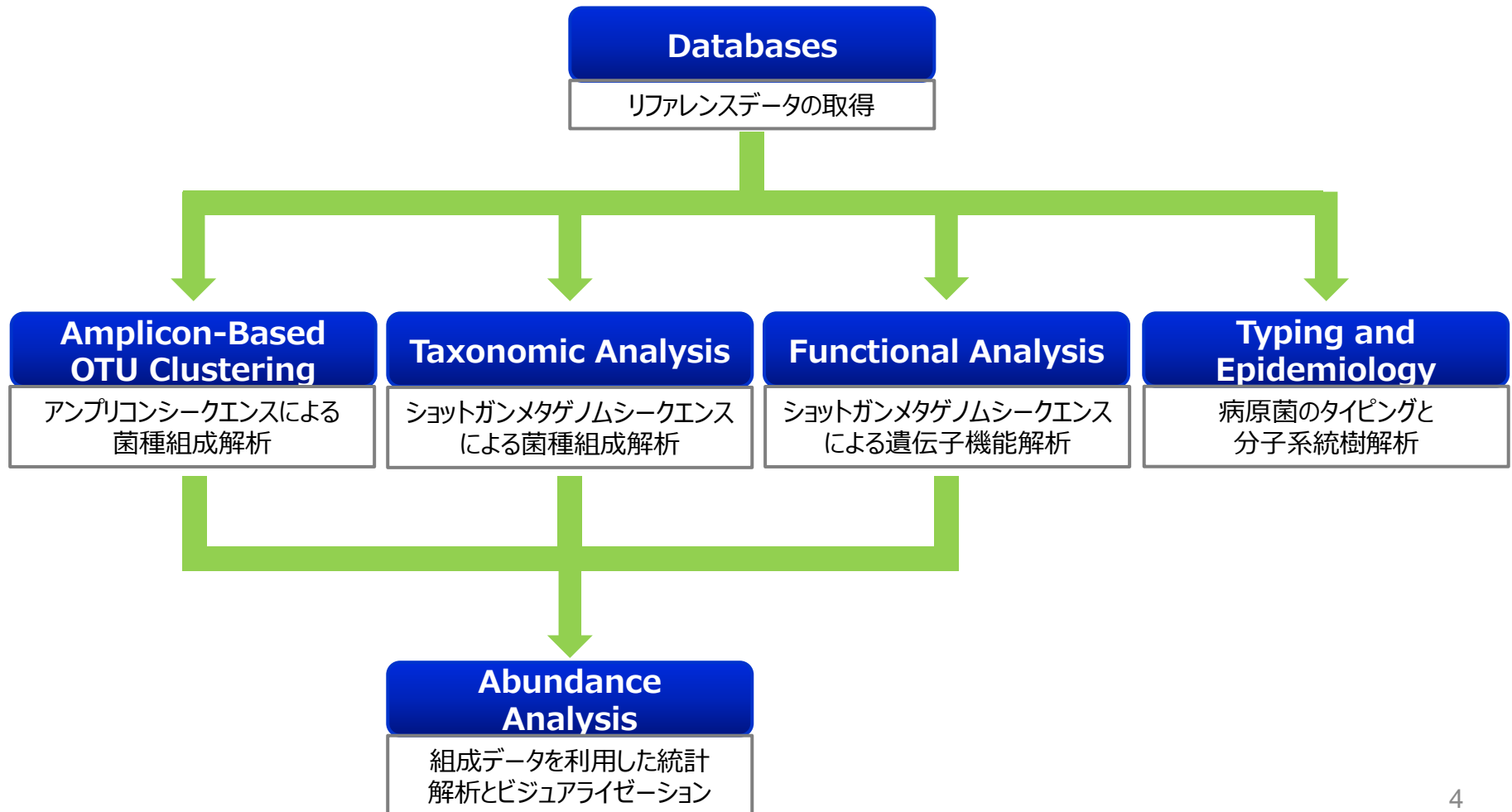
Typing and Epidemiology (ベータ版)

- NGS-MLST (Multi Locus Sequence Typing) 解析による病原菌のタイピングおよび薬剤耐性の確認
- K-mer Treeによる複数菌種のゲノム配列の類似度の比較
- SNP Treeによる分子系統樹の作成

Databases

- NCBIからの、バクテリアやウイルスなどのゲノム配列データの一括ダウンロード
- Greengenes, SILVA, UNITEなどのOTU配列データ、MLSTスキーマや薬剤耐性遺伝子配列データのダウンロード
- カスタムデータからのデータベース作成

- どのアプリケーションを使用する場合も、最初にDatabasesツールでリファレンスデータの取得が必要
- 「Amplicon-Based OTU Clustering」「Taxonomic Analysis」「Functional Analysis」では、解析結果のデータを使用し、「Abundance Analysis」による2次解析が可能



微生物プロフィール解析

Amplicon-Based OTU Clustering

- 16S-, 18S-, ITS rRNA配列のDe Novo / ReferenceベースのOTUクラスタリング
- Greengenes, Silva, UNITEの菌種分類データベース、またはカスタムデータベースをサポート

Taxonomic Analysis

- メタゲノムサンプル内の菌種組成の決定
- 宿主ゲノムDNAのコンタミネーションの除去



1. Amplicon-Based OTU Clustering

手順1 : リファレンスデータの取得

- Greengenesなどのリファレンスデータをダウンロード

手順2 : 1次解析

- リード配列のクオリティチェック（アダプター除去、ペアリードの結合など）
- OTUクラスタリング

手順3 : 2次解析

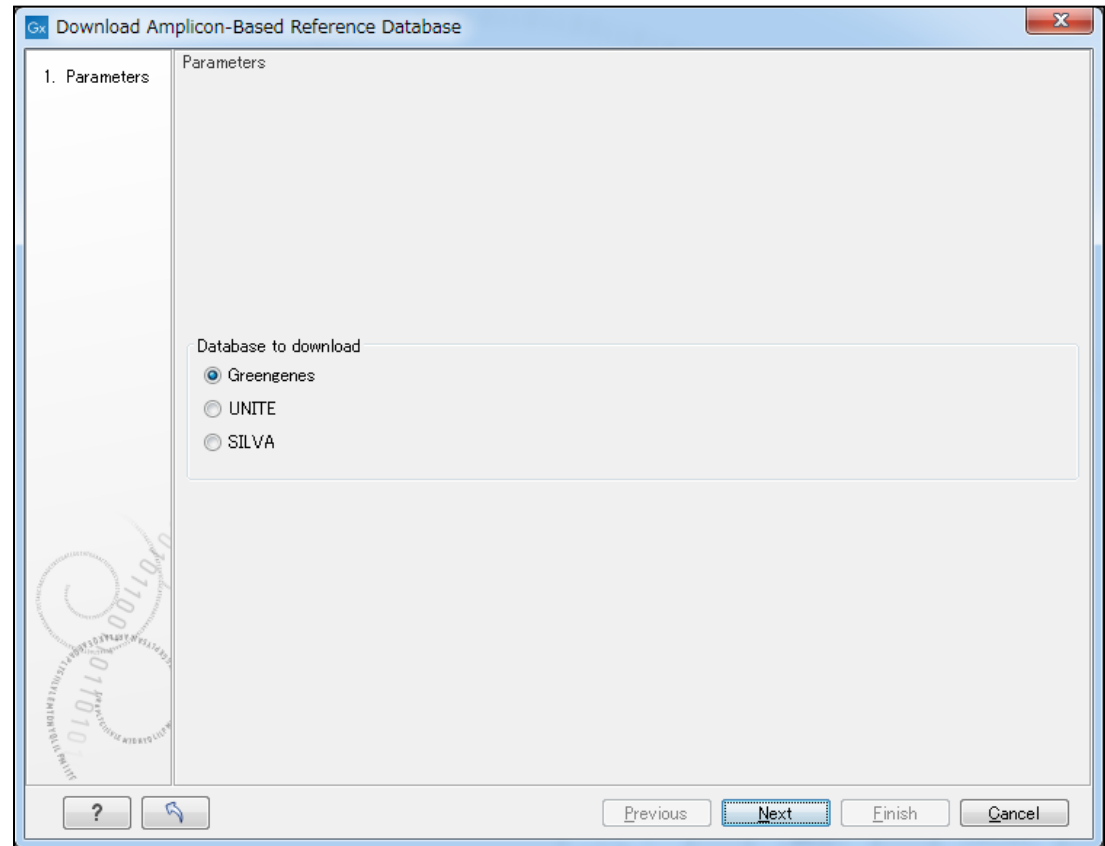
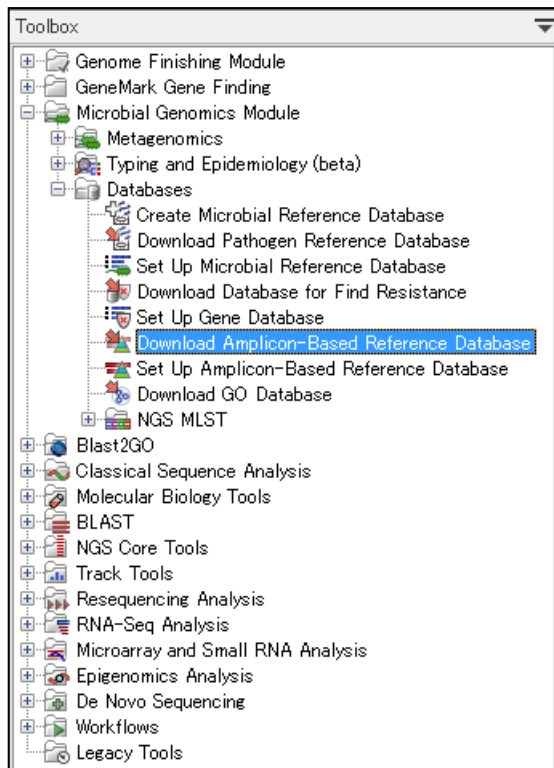
- アルファ多様性、ベータ多様性の計算

手順1. リファレンスデータの取得

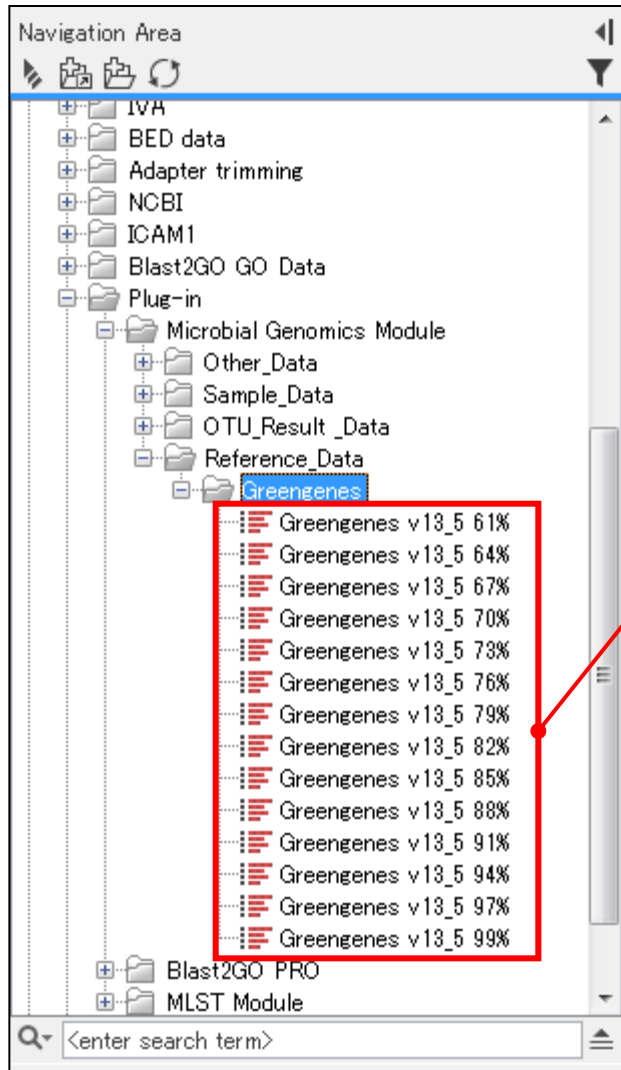
- リファレンスデータとして、データベースのOTU配列データが必要になり、専用のダウンロードツールを使ってダウンロードできる。
- ダウンロードを行う際は、コンピュータがインターネットに接続されている必要がある。

取得できるデータベースの種類

- **Greengenes** (原核生物の16S rRNA配列データ)
- **SILVA** (原核生物と真核生物の、16S/18S rRNA配列データ)
- **UNITE** (ITS spacer配列データ)

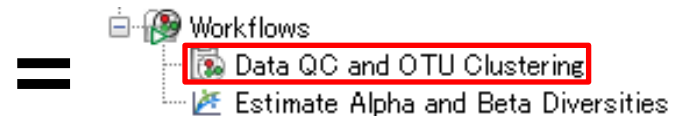
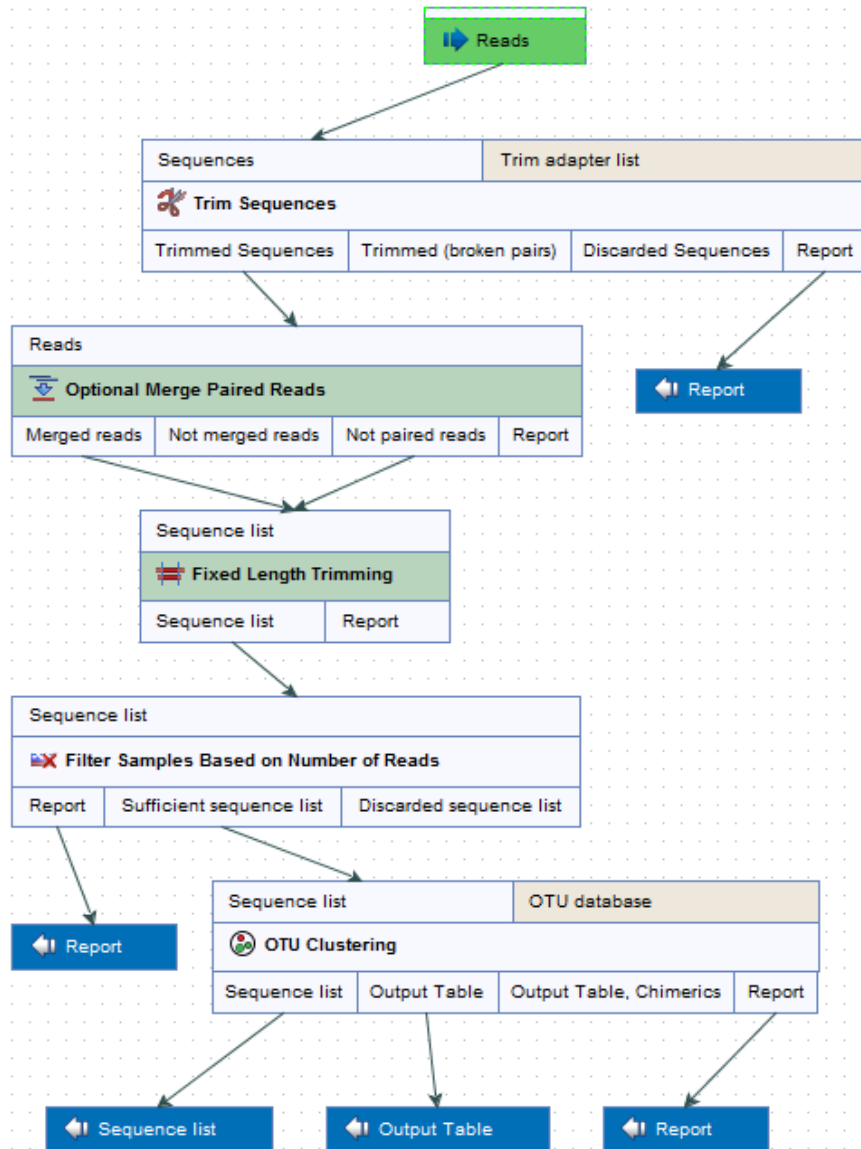


1. Download Amplicon-Based Reference Databaseを選択してダブルクリック。
2. ダウンロードするデータベースを選択。



ダウンロードが終了すると、保存先に指定したフォルダ内に、リファレンスデータが作成される。

手順2. **1次解析**



Trim Sequences

- アダプターの除去、およびリード配列のトリミングを行う

Optional Merge Paired Reads

- オーバーラップしているペアリード配列の結合を行う

Fixed Length Trimming

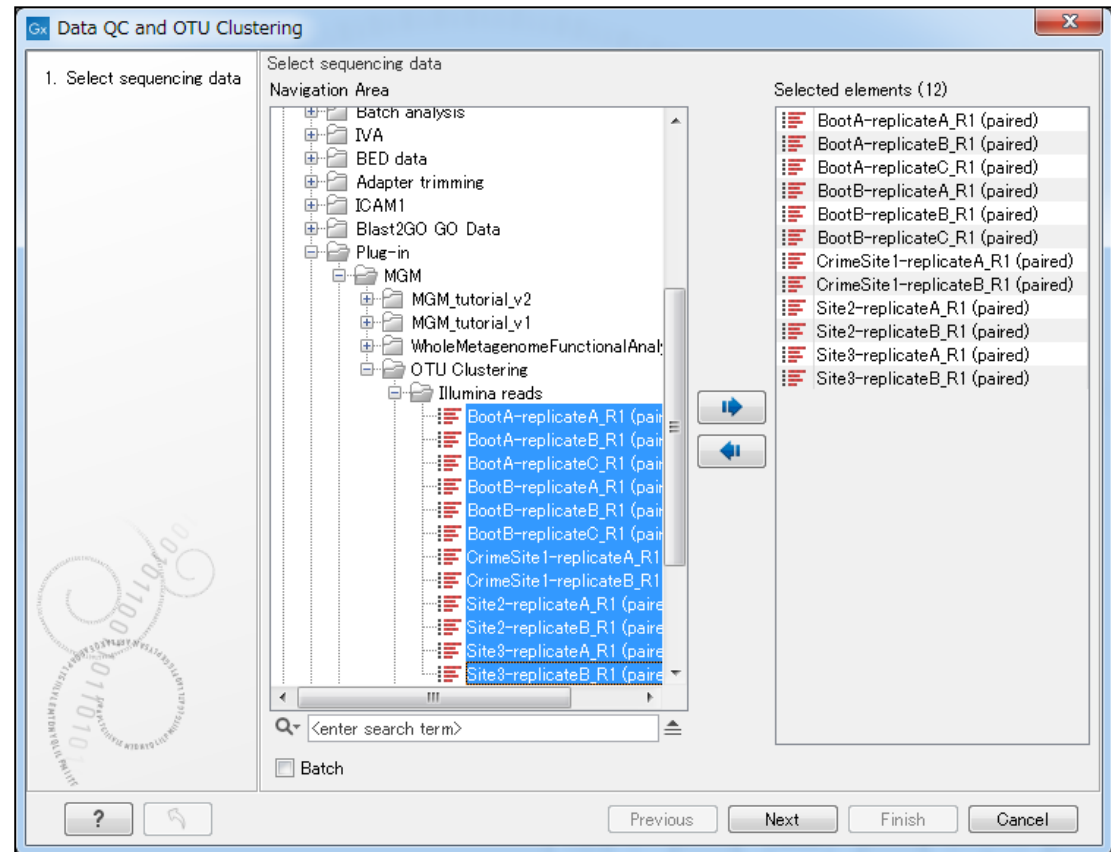
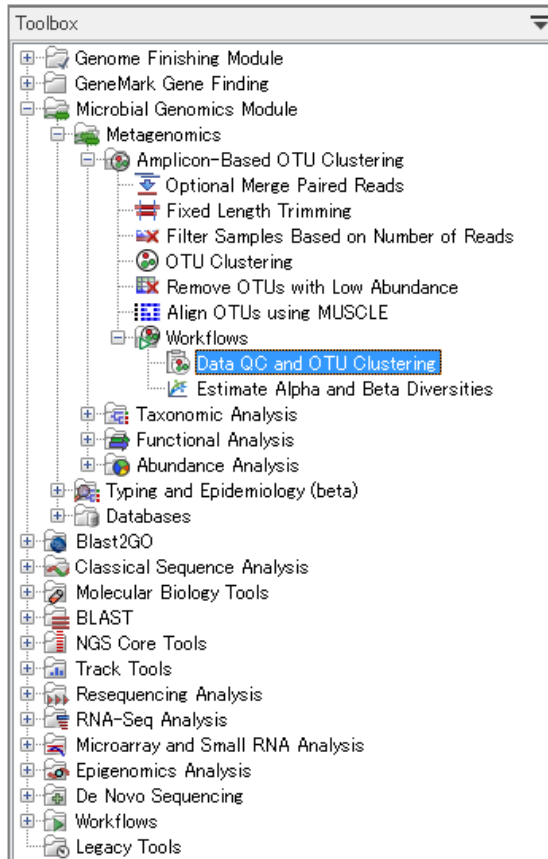
- リード配列の長さを揃える

Filter Samples Based on Number of Reads

- リード配列数の少ないサンプルを除去する

OTU Clustering

- OTUクラスタリングを行う



1. Data QC and OTU Clusteringを選択してダブルクリック。
2. リード配列データを選択。

1. Select sequencing data

2. Trim Sequences

Trim Sequences

Configurable Parameters

Trim adapter list: 16S_primers_round_robin

Quality limit: 0.05

Also search on reversed sequence:

Locked Settings

Previous Next Finish Cancel

- Trim adapter list: アダプターリストデータを指定する。
- Quality limit: Quality scoreの信頼性の閾値を指定する。
- Also search on reversed sequence: リード配列のアンチセンス鎖からも、アダプターのトリミングを行うかを選択する。

1. Select sequencing data

2. Trim Sequences

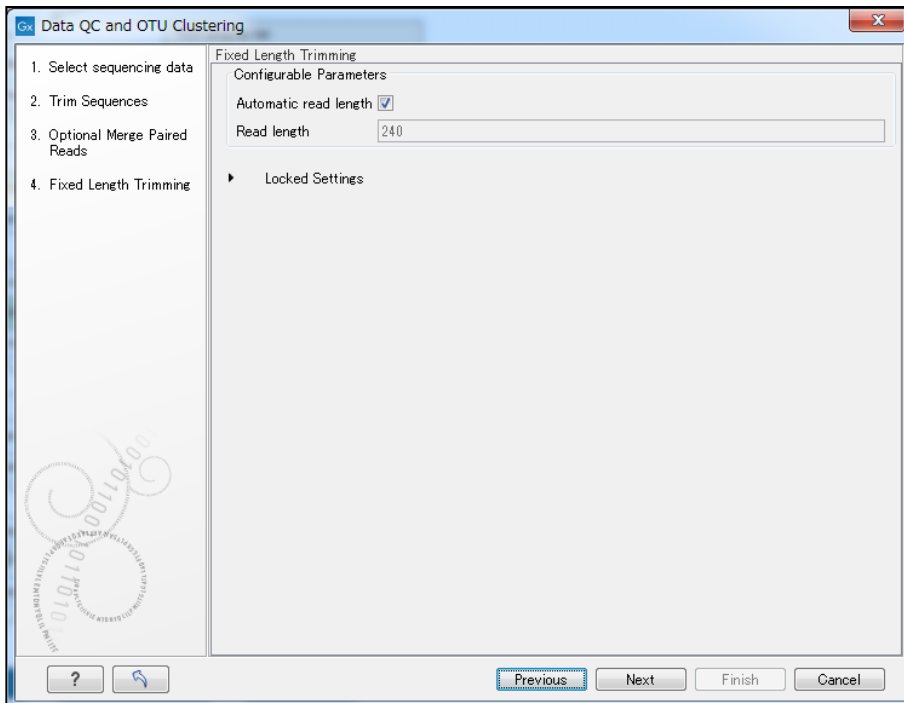
3. Optional Merge Paired Reads

Optional Merge Paired Reads

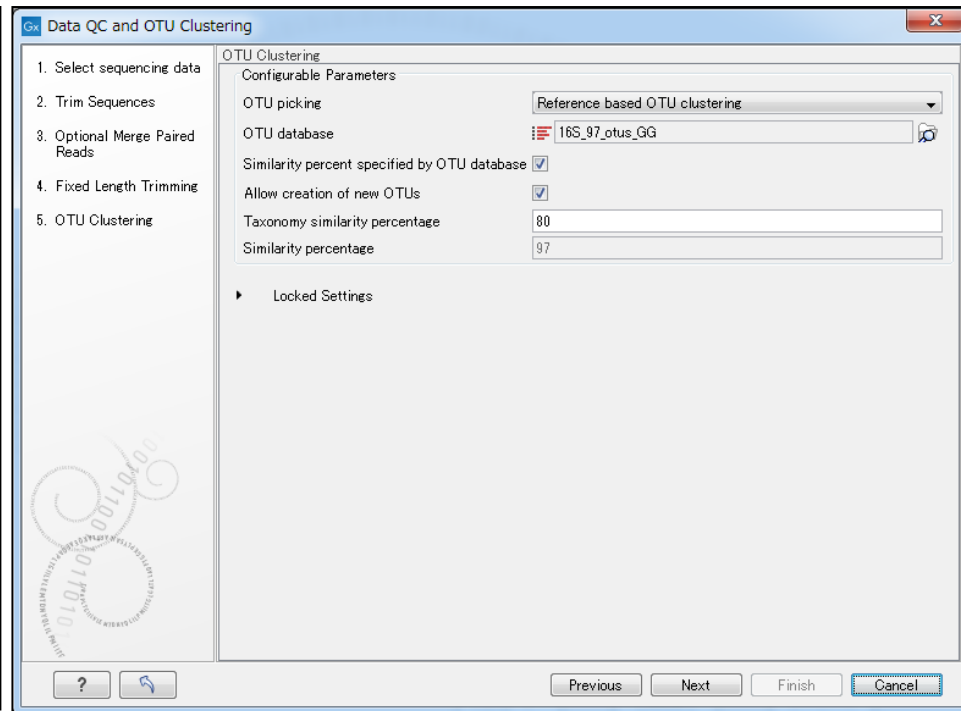
Mismatch cost	1
Minimum score	40
Gap cost	4
Maximum unaligned end mismatches	5

Previous Next Finish Cancel

- Mismatch cost: アライメントにマッチしない配列があった場合のコスト。
- Minimum score: リードを結合する場合の最低スコア。
- Gap cost: アライメントにギャップがある場合のコスト。
- Maximum unaligned end mismatches: リード末端において、許容するアライメントされない塩基数の設定。

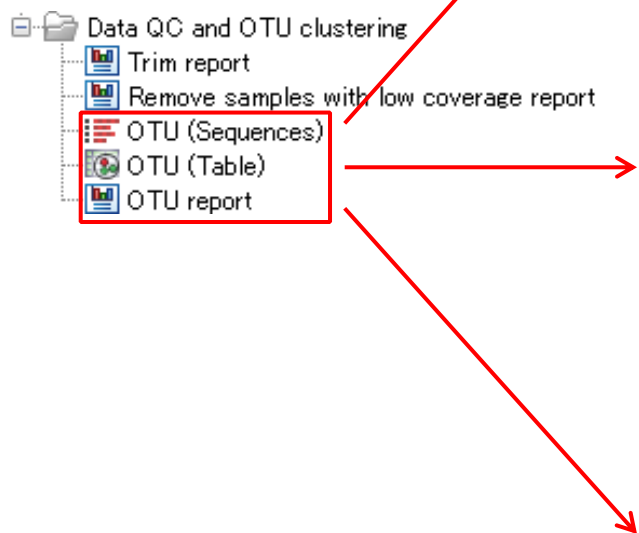


- **Automatic read length:** インputデータから自動で決定されたリード配列長に揃える。
* 別々にインputしたリード配列データでは、違う長さになることがあります。
- **Read length:** 任意のリード配列長を指定する。



- **OTU picking:** OTUクラスタリングの手法を選択する。
- **OTU database:** リファレンス配列データを指定する。
- **Similarity percent specified by OTU database:** リファレンス配列データで設定されている相同性の値を使い、OTUクラスタリングを行う。
- **Allow creation of new OTUs:** OTU databaseにマッチしなかったリード配列から、条件を再設定してOTUクラスタリングを行う。
- **Taxonomy similarity percentage:** マッチしなかったリード配列を使ってのOTUクラスタリングの相同性を再設定する。
- **Similarity percentage:** リファレンス配列データで設定されている相同性以外の値を使って、OTUクラスタリングを行う。

- 計算結果として、以下のデータが出力される。



OTU配列データ

OTU (Sequences) x

	20	40	60	80
1109867	TCGAGAATTTTCTCAATGGGGAAACCTGAGGAGCGACGCCGCTGGGGGATGAATGGCTTCGGCCCGTAAACCCCT			
1109766	TGGGGAATCTTGACAATGGGCGAAAGCCTGATGCAGCGACGCCGCTGGGGGATGAAGCATTTCGGTGTGTAAACCCCT			
1108954	TGGGGAATATTGCGCAATGGGCGAAAGCCTGACGCAGCGACGCCGCTGGGGGATGAAGGCCCTTCGGTGTGTAAACCCCT			
1108431	TCGAGAATTTTCTCAATGGGCGAAAGCCTGATGGAGCGACGCCGCTGGGGGATGAATGGCTTCGGCCCGTAAACCCCT			
1108199	TGGGGAATTTGCGCAATGGGGAAACCTGACGCAGCAACGCCGCTGGAGGATGAAGTCCTTGGGACGTAAACTCCT			

OTU組成データ

OTU (Table) x

Rows: 1,256

Name	Taxonomy	Combined A...	GT-A-B
1109867	k_Bacteria, p_Verrucomicrobia, c_[Pedosphaerae], o_[Pedosphaerales], f_auto67_4W, g_s_	2	0
1109766	k_Bacteria, p_Acidobacteria, c_DA052, o_Ellin6513, f_g_s_	12	0
1108954	k_Bacteria, p_Actinobacteria, c_Actinobacteria, o_Actinomycetales, f_g_s_	116	16
1108431	k_Bacteria, p_Verrucomicrobia, c_[Pedosphaerae], o_[Pedosphaerales], f_Ellin515, g_s_	34	0
1108199	k_Bacteria, p_Acidobacteria, c_Acidobacteriia, o_Acidobacteriales, f_Koribacteraceae, g_s_	12	2
1107608	k_Bacteria, p-Proteobacteria, c_Deltaproteobacteria, o_Myxococcales, f_g_s_	15	1
1107128	k_Bacteria, p_Verrucomicrobia, c_[Spartobacteria], o_[Chthoniobacteriales], f_[Chthoniobacteraceae], g_DA101, s_	2	0
1107044	k_Bacteria, p_Verrucomicrobia, c_[Pedosphaerae], o_[Pedosphaerales], f_auto67_4W, g_s_	4	0
1107029	k_Bacteria, p_Acidobacteria, c_DA052, o_Ellin6513, f_g_s_	19	0
1105309	k_Bacteria, p-Proteobacteria, c_Alphaproteobacteria, o_Rhodospirillales, f_Rhodospirillaceae, g_s_	5	0

レポートデータ

OTU report x

1 Overall summary

Input database size	Filtered database size	OTUs based on database	De novo OTUs	Total predicted OTUs
99,322	1,237	974	474	1,448

2 Read summary

Number of reads	Filtered reads	Unique reads after filtering	Chimeric reads	Unique chimeric reads	Reads in OTUs
82,371	47,870	6,646	514	172	33,987

Rows: 1,785

Name	Taxonomy	Combined A...	GT-A-B	GT-A-C	GT-B-A	GT-B-B
1109867	k_Bacteria, p_Verrucomicrobia, c_[Pedosphaerae], o_[Pedosphaerales], f_auto67_4W, g_s_	2	0	0	0	0
1109766	k_Bacteria, p_Acidobacteria, c_DA052, o_Ellin6513, f_g_s_	12	0	0	0	0
1108954	k_Bacteria, p_Actinobacteria, c_Actinobacteria, o_Actinomycetales, f_g_s_	116	16	15	9	0
1108431	k_Bacteria, p_Verrucomicrobia, c_[Pedosphaerae], o_[Pedosphaerales], f_Ellin515, g_s_	34	0	9	2	0
1108199	k_Bacteria, p_Acidobacteria, c_Acidobacteria, o_Acidobacteriales, f_Koribacteriaceae, g_s_	12	2	1	1	0
1107608	k_Bacteria, p_Proteobacteria, c_Deltaproteobacteria, o_Myxococcales, f_g_s_	15	1	0	0	0
1107128	k_Bacteria, p_Verrucomicrobia, c_[Spartobacteria], o_[Chthoniobacteriales], f_[Chthoniobacteraceae], g_DA101, s_	2	0	0	0	0
1107044	k_Bacteria, p_Verrucomicrobia, c_[Pedosphaerae], o_[Pedosphaerales], f_auto67_4W, g_s_	4	0	1	0	0
1107029	k_Bacteria, p_Acidobacteria, c_DA052, o_Ellin6513, f_g_s_	19	0	0	0	0
1105389	k_Bacteria, p_Proteobacteria, c_Alphaproteobacteria, o_Rhodospirillales, f_Rhodospirillaceae, g_s_	5	0	0	0	0
1105075	k_Bacteria, p_Proteobacteria, c_Alphaproteobacteria, o_Rhodospirillales, f_Rhodospirillaceae, g_s_	2	0	0	0	0
1104688	k_Bacteria, p_Acidobacteria, c_DA052, o_Ellin6513, f_g_s_	96	4	18	8	0
1102974	k_Bacteria, p_Armatimonadetes, c_Chthonomonadetes, o_Chthonomonadales, f_Chthonomonadaceae, g_s_	10	0	0	0	0
1100012	k_Bacteria, p_Actinobacteria, c_Actinobacteria, o_Actinomycetales, f_Micromonosporaceae, g_Actinoplanes, s_	4	0	0	0	0
1094340	k_Bacteria, p_Proteobacteria, c_Gammaproteobacteria, o_Xanthomonadales, f_Sinobacteraceae, g_s_	111	5	9	7	0
1088939	k_Bacteria, p_Bacteroidetes, c_Sphingobacteriia, o_Sphingobacteriales, f_g_s_	5	0	0	1	0
1088618	k_Bacteria, p_Proteobacteria, c_Alphaproteobacteria, o_Caulobacteriales, f_Caulobacteraceae, g_Phenylobacterium, s_	18	1	3	4	0
1087325	k_Bacteria, p_Actinobacteria, c_Actinobacteria, o_Actinomycetales, f_Streptomycetaceae, g_s_	2	0	0	1	0
1086554	k_Bacteria, p_Bacteroidetes, c_Sphingobacteriia, o_Sphingobacteriales, f_g_s_	2	0	0	0	0
1085127	k_Bacteria, p_Proteobacteria, c_Betaproteobacteria, o_SC-I-84, f_g_s_	125	0	1	2	0
1083723	k_Bacteria, p_Plantomycetes, c_Plantomycetia, o_Gemmatales, f_Gemmataceae, g_s_	31	2	5	3	0
1081231	k_Bacteria, p_Bacteroidetes, c_[Saprospirae], o_[Saprospirales], f_Chitinophagaceae, g_s_	19	2	2	1	0
1075732	k_Bacteria, p_Actinobacteria, c_Actinobacteria, o_Actinomycetales, f_g_s_	449	33	94	27	0
1068847	k_Bacteria, p_Bacteroidetes, c_[Saprospirae], o_[Saprospirales], f_Chitinophagaceae, g_s_	2	0	0	0	0
1062748	k_Bacteria, p_Actinobacteria, c_Actinobacteria, o_Actinomycetales, f_Mycobacteriaceae, g_Mycobacterium, s_	8	0	1	1	0
1060353	k_Bacteria, p_Actinobacteria, c_Thermoleophilia, o_Solirubrobacteriales, f_Conexibacteraceae, g_s_	9	0	0	0	0
1058769	k_Bacteria, p_Plantomycetes, c_Plantomycetia, o_Gemmatales, f_Gemmataceae, g_s_	12	0	1	1	0
1052559	k_Bacteria, p_Proteobacteria, c_Alphaproteobacteria, o_Sphingomonadales, f_Sphingomonadaceae, g_s_	5	0	0	1	0
1052235	k_Bacteria, p_Proteobacteria, c_Deltaproteobacteria, o_Myxococcales, f_Haliangiaceae, g_s_	2	0	0	0	0
1051459	k_Bacteria, p_Proteobacteria, c_Betaproteobacteria, o_Burkholderiales, f_Comamonadaceae, g_s_	4	0	0	0	0
1043812	k_Bacteria, p_Actinobacteria, c_Actinobacteria, o_Actinomycetales, f_g_s_	143	17	21	11	0
1034969	k_Bacteria, p_Proteobacteria, c_Alphaproteobacteria, o_Ellin329, f_g_s_	7	0	0	0	0
1033426	k_Bacteria, p_Actinobacteria, c_Actinobacteria, o_Actinomycetales, f_Nocardioideae, g_Nocardioides, s_	28	0	0	2	0
1029922	k_Bacteria, p_Proteobacteria, c_Alphaproteobacteria, o_Ellin329, f_g_s_	3	0	0	0	0
1026261	k_Bacteria, p_Proteobacteria, c_Deltaproteobacteria, o_Myxococcales, f_g_s_	3	0	0	0	0
1024056	k_Bacteria, p_Proteobacteria, c_Alphaproteobacteria, o_Rhodospirillales, f_Rhodospirillaceae, g_s_	2	0	0	0	0
1021984	k_Bacteria, p_Proteobacteria, c_Deltaproteobacteria, o_Myxococcales, f_g_s_	17	1	0	4	0
1019914	k_Bacteria, p_Bacteroidetes, c_Cytophagia, o_Cytophagales, f_Cytophagaceae, g_s_	14	0	0	4	0
1018122	k_Bacteria, p_Actinobacteria, c_Actinobacteria, o_Actinomycetales, f_g_s_	15	0	0	1	0
1017063	k_Bacteria, p_Proteobacteria, c_Deltaproteobacteria, o_Myxococcales, f_Polyangiaceae, g_s_	16	0	1	3	0
1012668	k_Bacteria, p_Actinobacteria, c_Acidimicrobiia, o_Acidimicrobiales, f_EB1017, g_s_	5	0	0	2	0

Table Settings

Column width: Manual

Show column:

- Name
- Taxonomy
- Combined Abundance
- Min
- Max
- Mean
- Median
- Std
- GT-A-B
- GT-A-C
- GT-B-A
- GT-B-B
- GT-B-C
- Site-1-A
- Site-1-B
- Site-2-A
- Site-2-B
- Site-3-A
- Site-3-B
- Sequence

Select All / Deselect All

Data

Show OTU abundance values as:

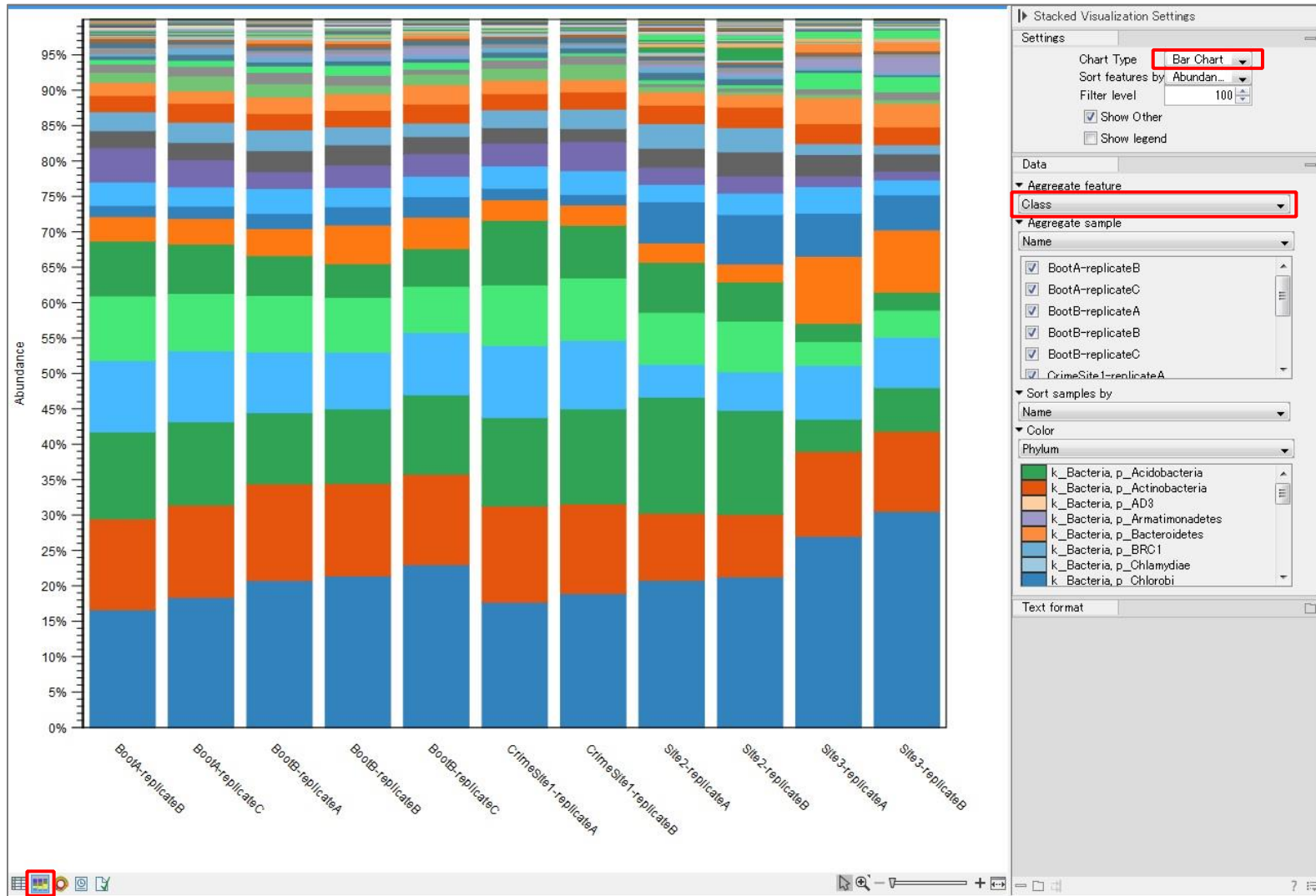
- Counts
- Relative abundance

Aggregate taxonomy:

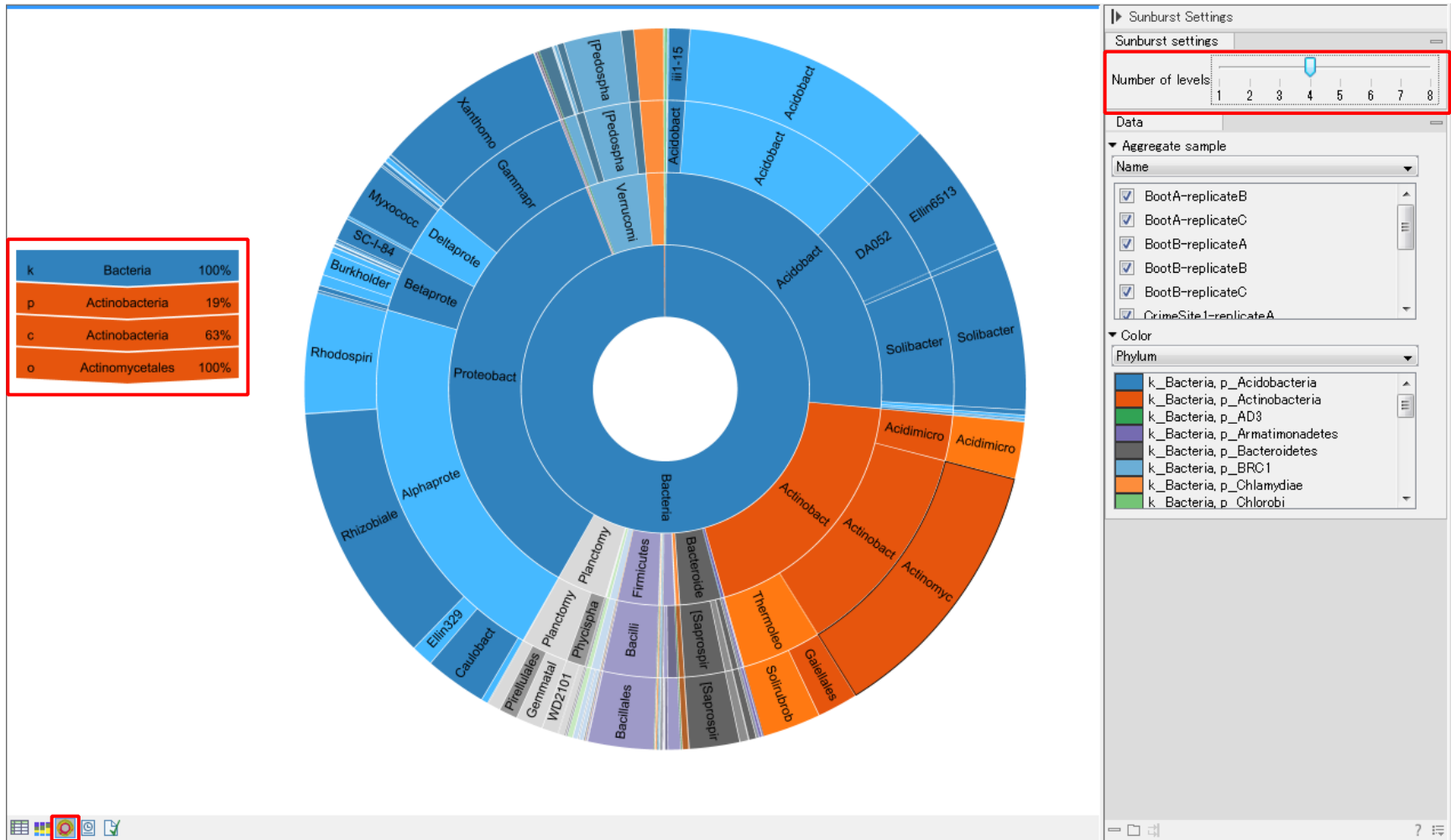
- OTU name
- Kingdom
- Phylum
- Class
- Order
- Family
- Genus
- Species

Create OTU abundance table from Selection

- 組成データのShow Tableアイコンから、サンプルごとの組成データをリスト形式で確認できる。
- Aggregate taxonomy項目から、界～種などのカテゴリー分類を切り替えることが可能。

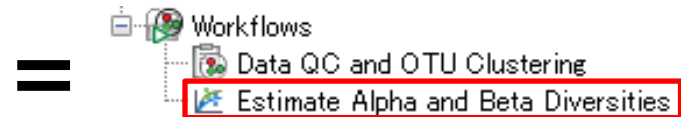
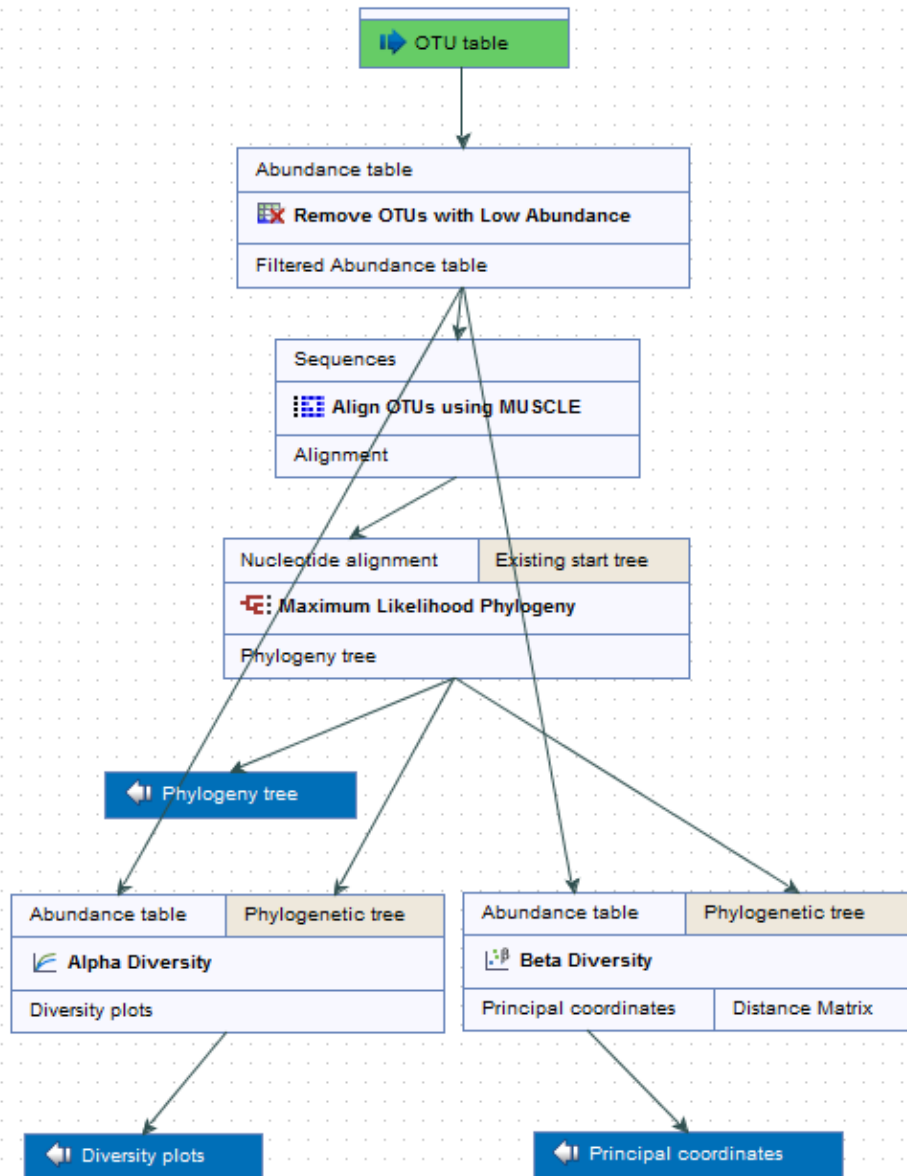


- 組成データのShow Stacked Visualizationアイコンから、サンプルごとの組成データをグラフ形式で確認できる。
- グラフの種類、またカテゴリー分類の切り替えが可能。



- 組成データのShow Sunburst editorアイコンから、サンプルごとの組成データをサンバースト図で確認できる。
- 表示するカテゴリーレベルの変更が可能。
- 任意の箇所にマウスカーソルを合わせると、組成比データが表示される。

手順3. 2次解析



Remove OTUs with Low Abundance

・組成が低いOTUデータを除去する

Align OTUs using MUSCLE

・OTU配列データのアライメントを行う。

Maximum Likelihood Phylogeny

・アライメント結果から系統樹を作成する。

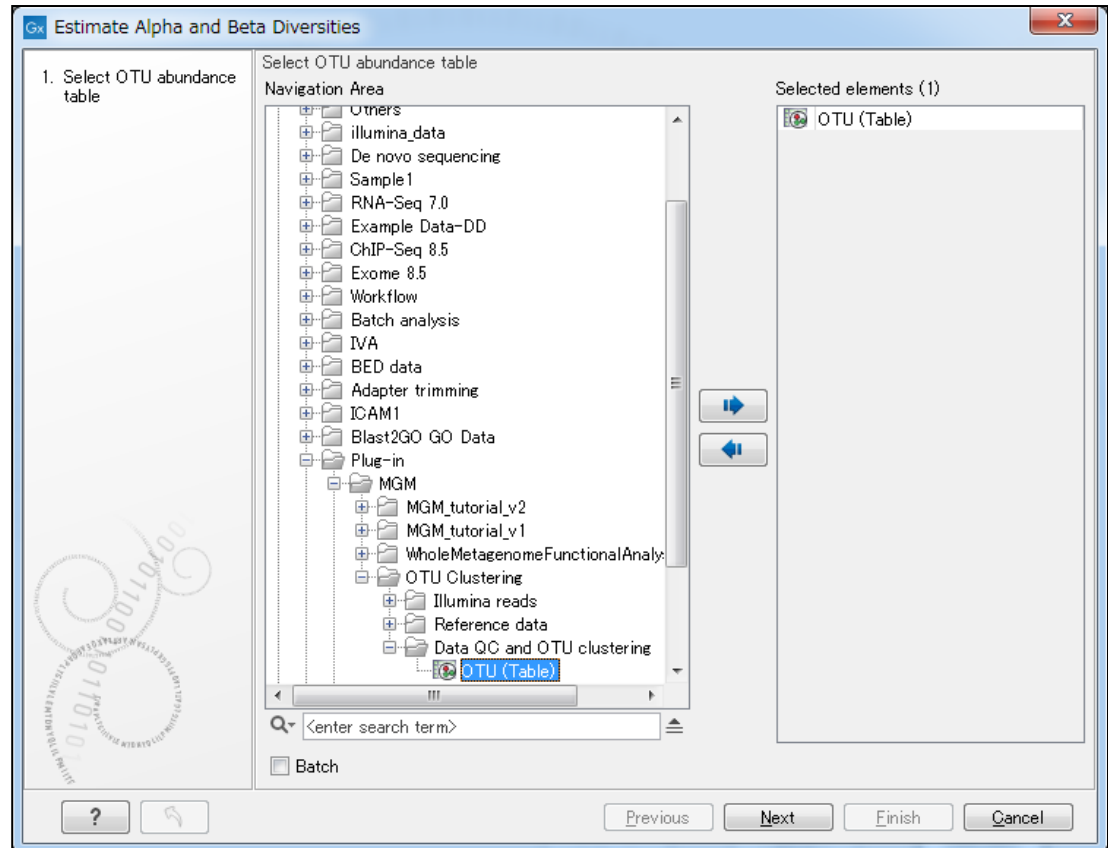
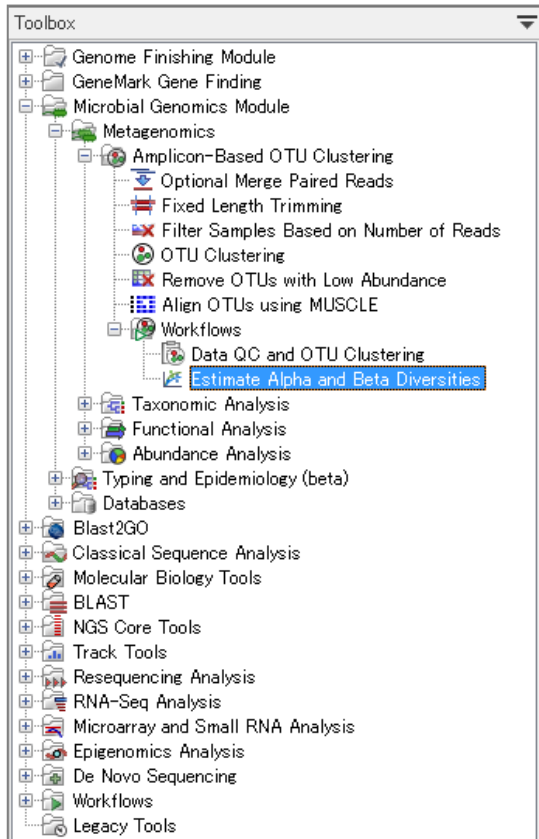
Alpha Diversity

・アルファ多様性を計算する。

Beta Diversity

・ベータ多様性を計算する。

Estimate Alpha and Beta Diversities



1. Estimate Alpha and Beta Diversitiesを選択してダブルクリック。
2. 組成データを選択。

Estimate Alpha and Beta Diversities

Estimate Alpha and Beta Diversities

1. Select OTU abundance table

2. Alpha Diversity

Alpha Diversity

Configurable Parameters

Number of OTUs

Chao 1 bias-corrected

Chao 1

Simpson's index

Shannon entropy

Phylogenetic diversity

Locked Settings

Previous Next Finish Cancel

- Number of OTUs ~ Shannon entropy: アルファ多様性の計算アルゴリズムを選択する。
- Phylogenetic diversity: 系統的多様性の計算を行う。

Estimate Alpha and Beta Diversities

1. Select OTU abundance table

2. Alpha Diversity

3. Beta Diversity

Beta Diversity

Configurable Parameters

Bray-Curtis

Jaccard

Euclidean

Unweighted UniFrac

Weighted UniFrac

Weighted UniFrac not normalized

D_0 UniFrac

D_0.5 UniFrac

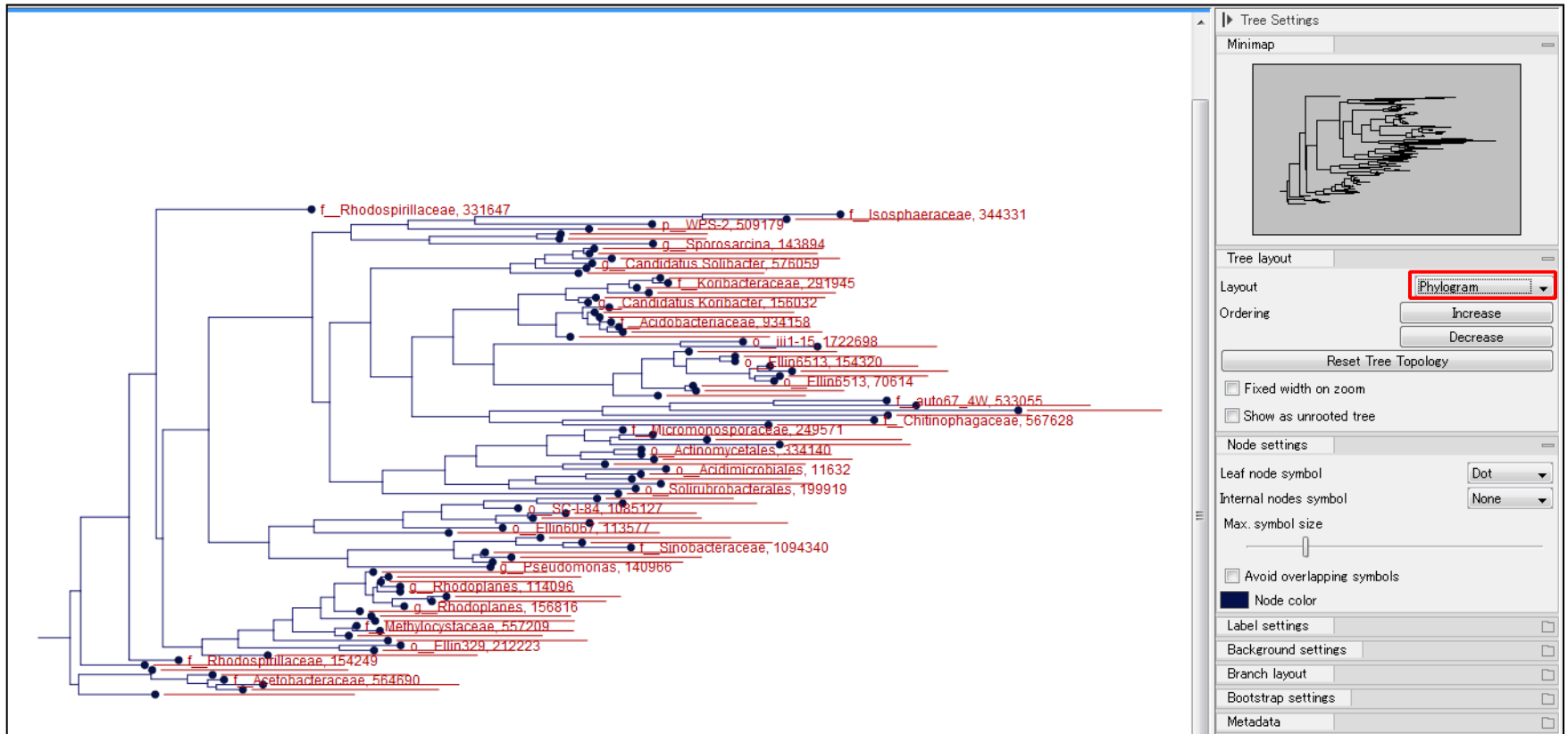
Locked Settings

Previous Next Finish Cancel

- Bray-Curtis ~ Euclidean: ベータ多様性の計算アルゴリズムを選択する。
- Unweighted UniFrac ~ D_0.5 UniFrac: 距離の計算アルゴリズムを選択する。

Estimate Alpha and Beta Diversities

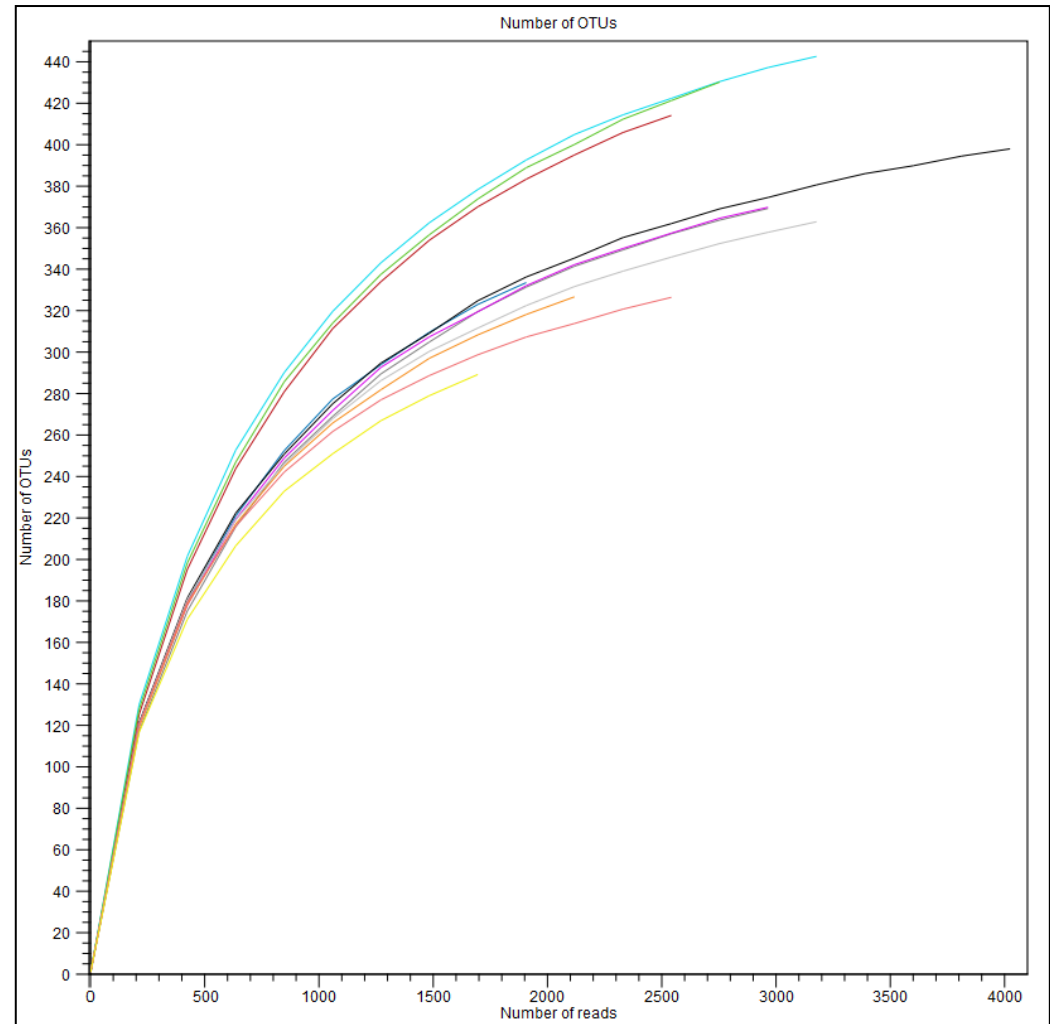
- Estimate Alpha and Beta Diversities
 - OTU (Table) (Filtered) alignment tree**
 - OTU (Table) (Filtered) (PCoA - Bray-Curtis)
 - OTU (Table) (Filtered) (PCoA - Jaccard)
 - OTU (Table) (Filtered) (PCoA - Unweighted UniFrac)
 - OTU (Table) (Filtered) (PCoA - Weighted UniFrac)
 - OTU (Table) (Filtered) (Alpha Diversity - Number of OTUs)
 - OTU (Table) (Filtered) (Alpha Diversity - Chao 1 bias-corrected)
 - OTU (Table) (Filtered) (Alpha Diversity - Phylogenetic diversity)



- 系統樹データでは、サークル表示などのレイアウトの変更が可能。

Estimate Alpha and Beta Diversities

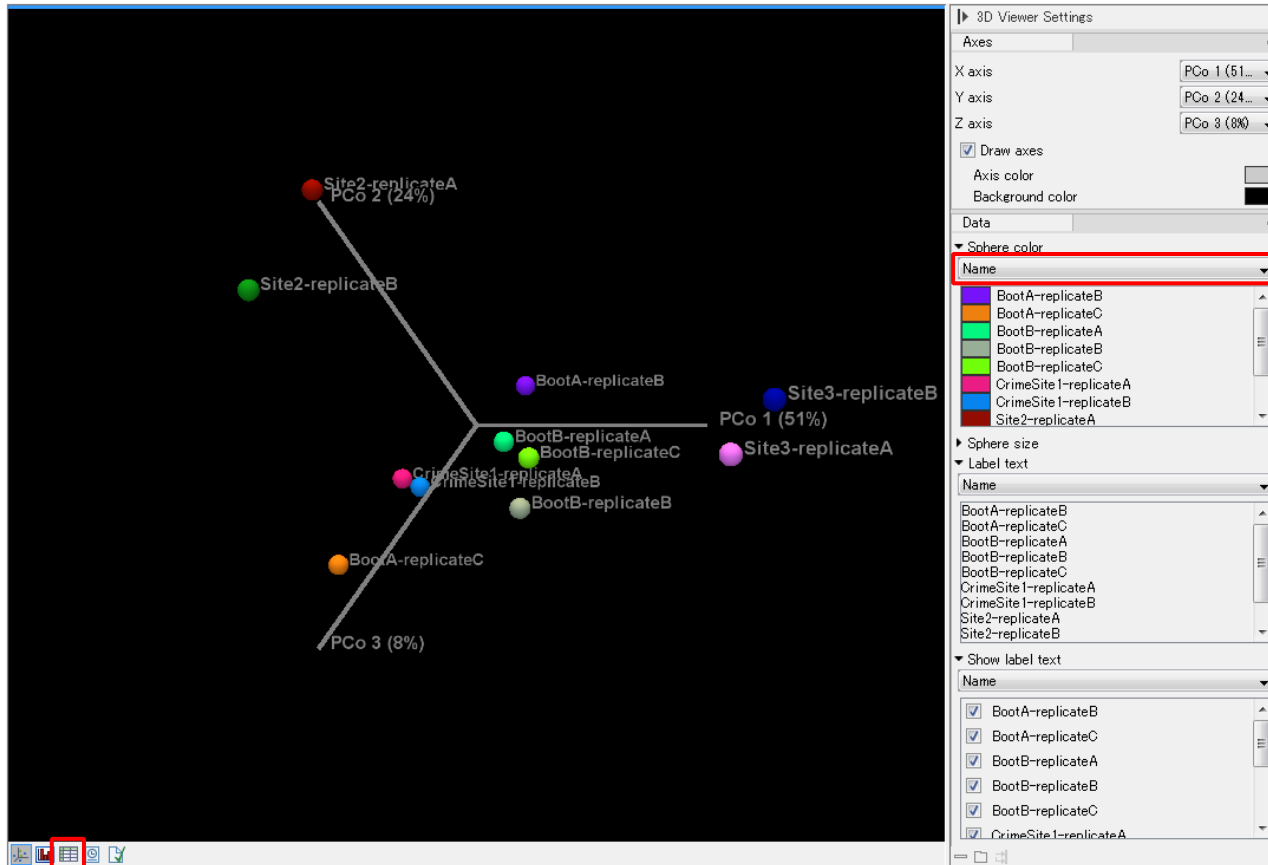
- Estimate Alpha and Beta Diversities
 - OTU (Table) (Filtered) alignment_tree
 - OTU (Table) (Filtered) (PCoA - Bray-Curtis)
 - OTU (Table) (Filtered) (PCoA - Jaccard)
 - OTU (Table) (Filtered) (PCoA - Unweighted UniFrac)
 - OTU (Table) (Filtered) (PCoA - Weighted UniFrac)
 - OTU (Table) (Filtered) (Alpha Diversity - Number of OTUs)
 - OTU (Table) (Filtered) (Alpha Diversity - Chao 1 bias-corrected)
 - OTU (Table) (Filtered) (Alpha Diversity - Phylogenetic diversity)



- アルファ多様性データをまとめたレポートが、選択したアルゴリズム別に出力される。

Estimate Alpha and Beta Diversities

- Estimate Alpha and Beta Diversities
 - OTU (Table) (Filtered) alignment_tree
 - OTU (Table) (Filtered) (PCoA - Bray-Curtis)
 - OTU (Table) (Filtered) (PCoA - Jaccard)
 - OTU (Table) (Filtered) (PCoA - Unweighted UniFrac)
 - OTU (Table) (Filtered) (PCoA - Weighted UniFrac)
 - OTU (Table) (Filtered) (Alpha Diversity - Number of OTUs)
 - OTU (Table) (Filtered) (Alpha Diversity - Chao 1 bias-corrected)
 - OTU (Table) (Filtered) (Alpha Diversity - Phylogenetic diversity)



- 選択した計算アルゴリズムごとに、ベータ多様性データが出力される。
- メタデータに基づいたラベルの表示分類や、テーブル表示への切り替えで、主成分の寄与度などを確認できる。

2. Taxonomic Analysis

手順1： リファレンスデータの取得

- バクテリア、ウイルスなどのリファレンスゲノム配列データをダウンロード

手順2： 1次解析

- リード配列のクオリティチェック（アダプター除去、低クオリティ配列のカットなど）
- 宿主ゲノムデータの除去とTaxonomicプロファイリング

手順3： 2次解析

- サンプルごとの菌種組成データの統合
- アルファ多様性、ベータ多様性の計算

手順1. リファレンスデータの取得

- リファレンスデータとして、バクテリアやウイルスなどのゲノム配列データが必要となり、専用のダウンロードツールを使ってダウンロードできる。
- ダウンロードを行う際は、コンピュータがインターネットに接続されている必要がある。
- データ量が大きいため、コンピュータのハードディスク空き容量に注意すること。

取得できるデータの種類

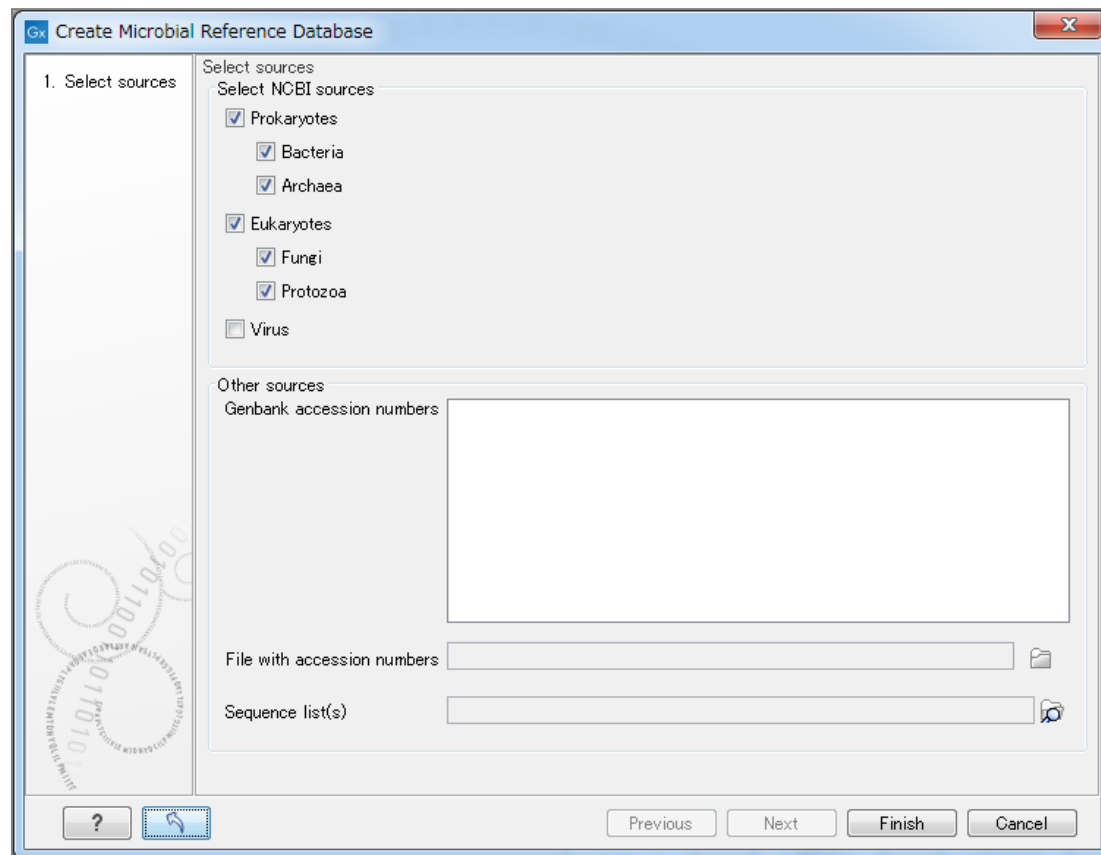
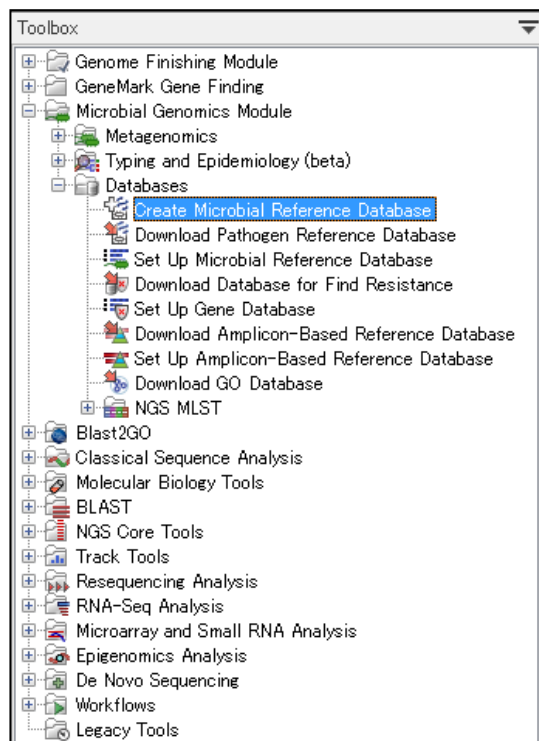
原核生物

- バクテリア
- 古細菌

真核生物

- カビ
- 原生動物

ウイルス



1. Create Microbial Reference Databaseを選択してダブルクリック。
2. ダウンロードするゲノムデータの生物の種類などを設定。

Create Microbial Reference Database

Included	Name	Taxonomy	Fraction in...	Assembly ID	Scaffolds	Size (Mbp)	In RefSeq	Status
<input type="checkbox"/>	Acanthamoeba polyphaga-duplicate-0001	0 / 1	GCA_001567625.1	18098	49.35	Yes	Contig
<input type="checkbox"/>	Acanthamoeba polyphaga-duplicate-0002	0 / 1	GCA_000826345.1	224482	120.42	Yes	Scaffold
<input type="checkbox"/>	Acanthamoeba quina	0 / 1	GCA_000826445.1	60490	83.59	Yes	Scaffold
<input type="checkbox"/>	Acanthamoeba rhyodes	0 / 1	GCA_000826385.1	62836	75.82	Yes	Scaffold
<input type="checkbox"/>	Acanthamoeba royreba	0 / 1	GCA_000826365.1	24098	79.54	Yes	Scaffold
<input type="checkbox"/>	Acaricomes phytoseiuli DSM 14247	0 / 1	GCA_000376245.1	64	2.42	Yes	Contig (refr)
<input type="checkbox"/>	Acaryochloris marina MBIC11017	0 / 1	GCA_000018105.1	10	8.36	Yes	Complete Genome (refr)
<input type="checkbox"/>	Acaryochloris sp. OCMEE 5410	0 / 1	GCA_000238775.2	511	7.88	Yes	Contig
<input type="checkbox"/>	Acetanaerobacterium elongatum	0 / 1	GCA_900103835.1	73	2.92	No	Scaffold
<input type="checkbox"/>	Acetitumaculum ruminis DSM 5522	0 / 1	GCA_900112085.1	54	3.08	No	Scaffold
<input type="checkbox"/>	Acetivibrio cellulolyticus CD2	0 / 1	GCA_000179595.2	10	6.16	Yes	Scaffold (refr)
<input type="checkbox"/>	Acetivibrio ethanoligenens	0 / 1	GCA_001461035.1	229	3.66	Yes	Scaffold (refr)
<input type="checkbox"/>	Acetoanaerobium noterae	0 / 1	GCA_900168025.1	19	2.81	Yes	Scaffold
<input type="checkbox"/>	Acetoanaerobium sticklandii	0 / 1	GCA_000196455.1	1	2.72	Yes	Complete Genome (refr)
<input type="checkbox"/>	Acetobacter aceti	0 / 1	GCA_002005445.1	1	3.73	Yes	Complete Genome
<input type="checkbox"/>	Acetobacter aceti 1023	0 / 1	GCA_000691125.1	33	3.01	Yes	Scaffold
<input type="checkbox"/>	Acetobacter aceti ATCC 23746	0 / 1	GCA_000379545.1	3	3.69	Yes	Scaffold
<input type="checkbox"/>	Acetobacter aceti NBRC 14818-duplicate-0001	0 / 1	GCA_000963905.1	209	3.55	Yes	Contig (refr)

Quick Selection Include All Exclude All

Selected references: 0
Size of selection: 0

Download selection

Single scaffold complete genomes in RefSeq
Complete genomes in RefSeq
All complete genomes

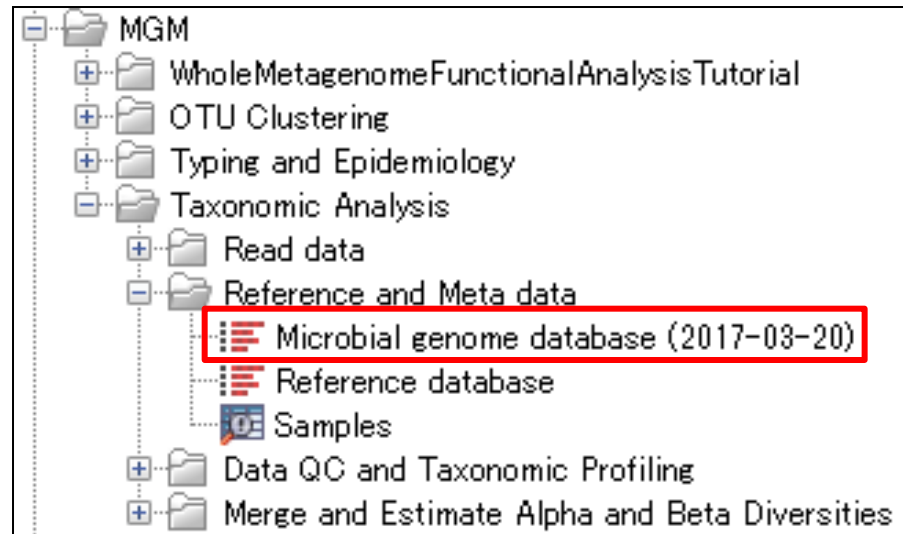
Included	Name	Taxonomy	Fraction in...	Assembly ID	Scaffolds	Size (Mbp)	In RefSeq	Status
<input type="checkbox"/>	Acanthamoeba polyphaga-duplicate-0001	0 / 1	GCA_001567625.1	18098	49.35	Yes	Contig
<input type="checkbox"/>	Acanthamoeba polyphaga-duplicate-0002	0 / 1	GCA_000826345.1	224482	120.42	Yes	Scaffold
<input type="checkbox"/>	Acanthamoeba quina	0 / 1	GCA_000826445.1	60490	83.59	Yes	Scaffold
<input type="checkbox"/>	Acanthamoeba rhyodes	0 / 1	GCA_000826385.1	62836	75.82	Yes	Scaffold
<input type="checkbox"/>	Acanthamoeba royreba	0 / 1	GCA_000826365.1	24098	79.54	Yes	Scaffold
<input type="checkbox"/>	Acaricomes phytoseiuli DSM 14247	0 / 1	GCA_000376245.1	64	2.42	Yes	Contig (refr)
<input checked="" type="checkbox"/>	Acaryochloris marina MBIC11017	1 / 1	GCA_000018105.1	10	8.36	Yes	Complete Genome (refr)
<input type="checkbox"/>	Acaryochloris sp. OCMEE 5410	0 / 1	GCA_000238775.2	511	7.88	Yes	Contig
<input type="checkbox"/>	Acetanaerobacterium elongatum	0 / 1	GCA_900103835.1	73	2.92	No	Scaffold
<input type="checkbox"/>	Acetitumaculum ruminis DSM 5522	0 / 1	GCA_900112085.1	54	3.08	No	Scaffold
<input type="checkbox"/>	Acetivibrio cellulolyticus CD2	0 / 1	GCA_000179595.2	10	6.16	Yes	Scaffold (refr)
<input type="checkbox"/>	Acetivibrio ethanoligenens	0 / 1	GCA_001461035.1	229	3.66	Yes	Scaffold (refr)
<input type="checkbox"/>	Acetoanaerobium noterae	0 / 1	GCA_900168025.1	19	2.81	Yes	Scaffold
<input checked="" type="checkbox"/>	Acetoanaerobium sticklandii	1 / 1	GCA_000196455.1	1	2.72	Yes	Complete Genome (refr)
<input checked="" type="checkbox"/>	Acetobacter aceti	1 / 1	GCA_002005445.1	1	3.73	Yes	Complete Genome
<input type="checkbox"/>	Acetobacter aceti 1023	0 / 1	GCA_000691125.1	33	3.01	Yes	Scaffold
<input type="checkbox"/>	Acetobacter aceti ATCC 23746	0 / 1	GCA_000379545.1	3	3.69	Yes	Scaffold
<input type="checkbox"/>	Acetobacter aceti NBRC 14818-duplicate-0001	0 / 1	GCA_000963905.1	209	3.55	Yes	Contig (refr)

Quick Selection Include All Exclude All

Selected references: 8,465
Size of selection: 39 (Gbp)

Download selection

3. データの一覧が表示されたら、Quick Selectionより、ダウンロードするデータのカテゴリを選択する。
4. チェックが表示されたデータと合計データ容量を確認し、Download selectionをクリックする。

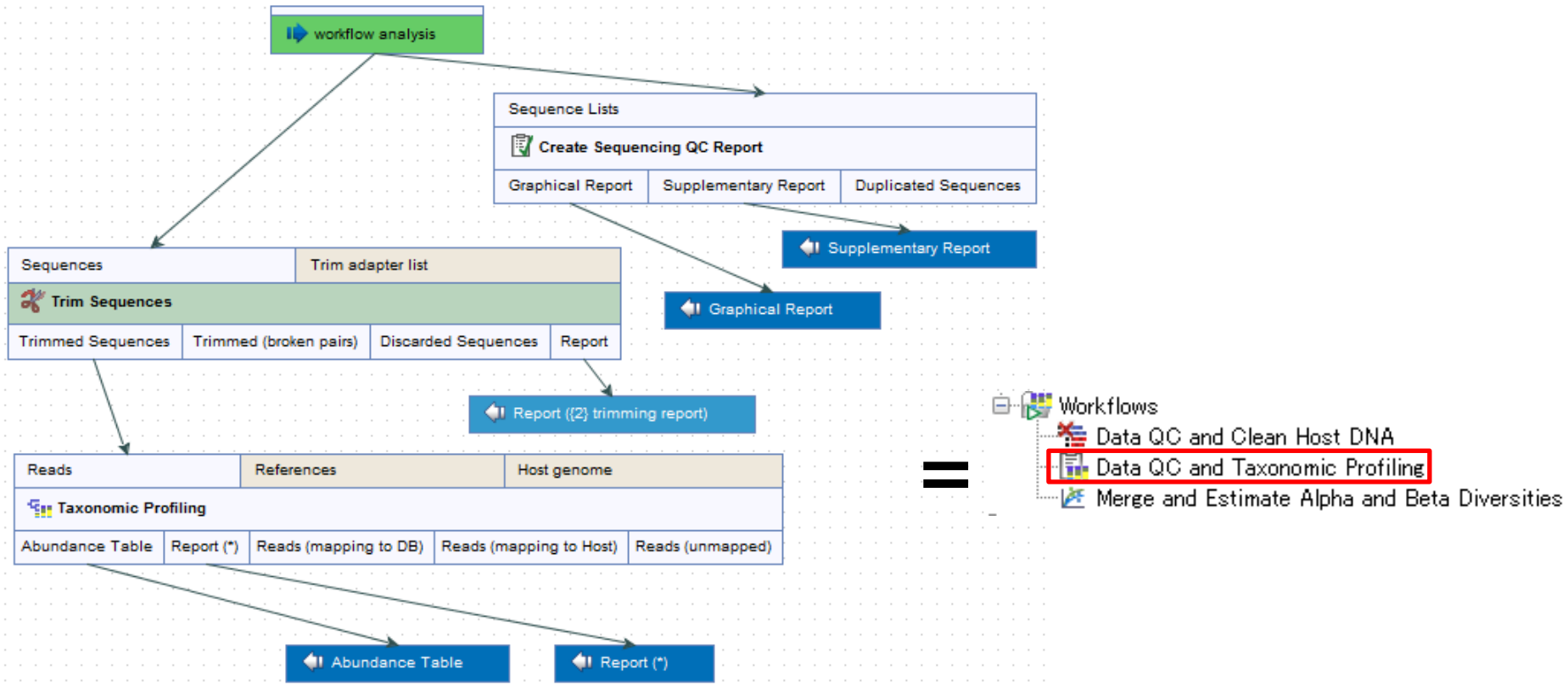


5. ダウンロードされたゲノムデータが作成される。

* ダウンロードには数時間かかることもあり、データベースの更新作業によって、ダウンロードが途中で中断されることもある。その場合は、再度実行し直す必要がある。

手順2. **1次解析**

1次解析の解析パイプライン



Create Sequencing QC Report

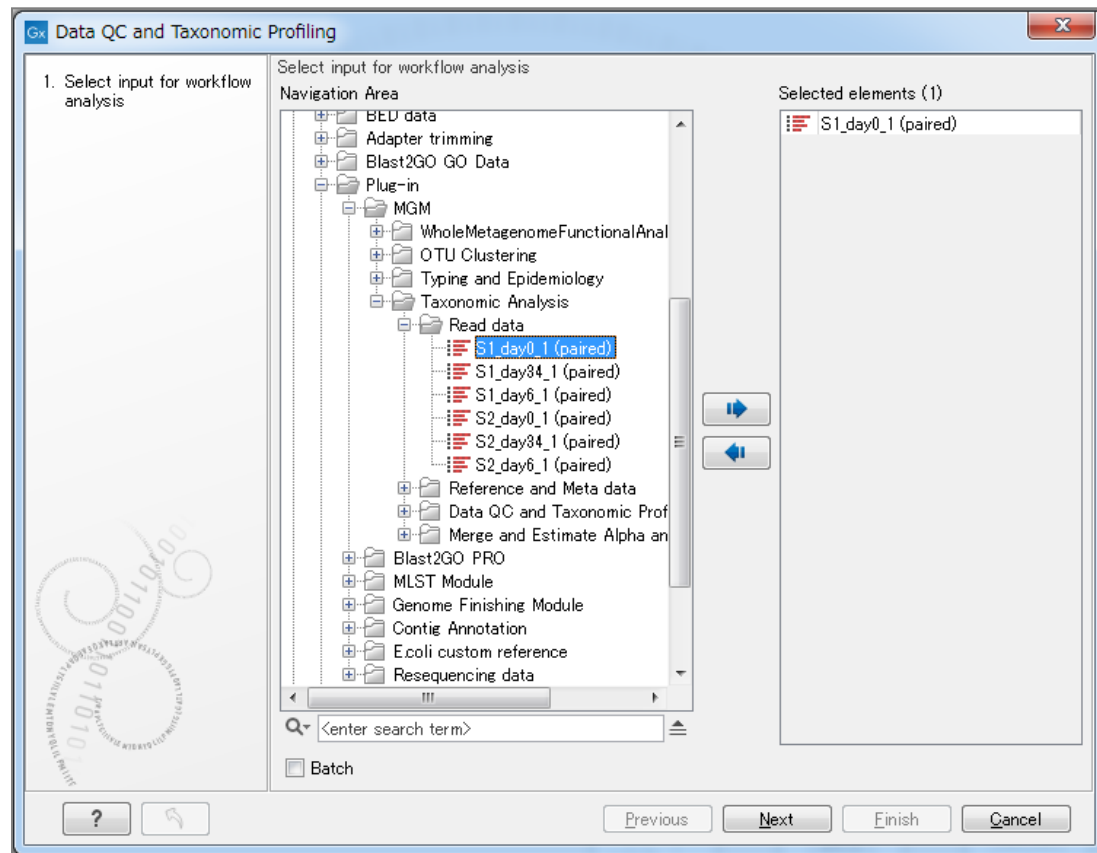
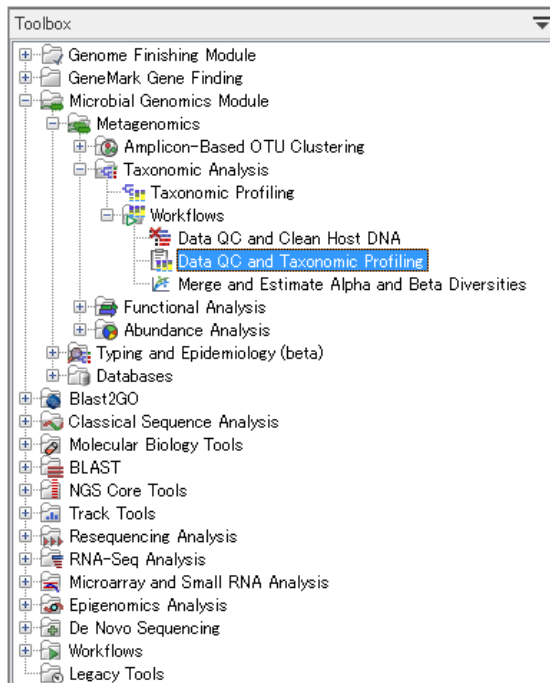
・リード配列データのQCチェック結果のレポートを作成する

Trim Sequences

・アダプターの除去、およびリード配列のトリミングを行う

Taxonomic Profiling

・宿主ゲノム配列の除去と、菌種組成解析を行う



1. Data QC and Taxonomic Profilingを選択してダブルクリック。
2. リード配列データを選択。

1. Select input for workflow analysis

2. Trim Sequences

Trim Sequences

Configurable Parameters

Trim adapter list

Quality limit

Also search on reversed sequence

Locked Settings

Previous Next Finish Cancel

1. Select input for workflow analysis

2. Trim Sequences

3. Taxonomic Profiling

Taxonomic Profiling

References

Filter host reads

Host genome

Previous Next Finish Cancel

- Trim adapter list: アダプターリストデータを指定する。
- Quality limit: Quality scoreの信頼性の閾値を指定する。
- Also search on reversed sequence: リード配列のアンチセンス鎖からも、アダプターのトリミングを行うかを選択する。

- References: リファレンスゲノム配列データを指定する。
- Filter host reads: 宿主ゲノム配列データの除去を行うかどうかを選択する。
- Host genome: 宿主ゲノム配列データを指定する。

Name	Taxonomy	Assembly ID	Combined ...	S1_day0	Confidence...	Total reads ...
Alistipes putredinis (Unknown)	Bacteria; Bacteroidetes; Bacteroidia; Bacteroidales; Rikenellaceae; Alistipes; Alistipes putredinis		11.84	11.84	0.50	99283.0
Bacteroides coprocola DSM 17136	Bacteria; Bacteroidetes; Bacteroidia; Bacteroidales; Bacteroidaceae; Bacteroides; Bacteroides coprocola	GCA_000154845.1	8.06	8.06	0.45	135629.0
Faecalibacterium prausnitzii (Unknown)	Bacteria; Firmicutes; Clostridia; Clostridiales; Ruminococcaceae; Faecalibacterium; Faecalibacterium prausnitzii		6.44	6.44	0.37	59175.0
Bacteroides stercoris (Unknown)	Bacteria; Bacteroidetes; Bacteroidia; Bacteroidales; Bacteroidaceae; Bacteroides; Bacteroides stercoris		5.44	5.44	0.32	27317.0
Bacteroides plebeius (Unknown)	Bacteria; Bacteroidetes; Bacteroidia; Bacteroidales; Bacteroidaceae; Bacteroides; Bacteroides plebeius		4.94	4.94	0.31	16978.0
Parabacteroides johnsonii (Unknown)	Bacteria; Bacteroidetes; Bacteroidia; Bacteroidales; Porphyromonadaceae; Parabacteroides; Parabacteroides johnsonii		4.75	4.75	0.07	53711.0
Prevotella sp. 109 (Unknown)	Bacteria; Bacteroidetes; Bacteroidia; Bacteroidales; Prevotellaceae; Prevotella; Prevotella sp. 109		4.67	4.67	0.36	8492.0
Barnesiella intestinihominis YIT 11860	Bacteria; Bacteroidetes; Bacteroidia; Bacteroidales; Porphyromonadaceae; Barnesiella; Barnesiella intestinihominis	GCA_000296465.1	3.63	3.63	4.13E-9	124731.0
Bacteroides vulgatus (Unknown)	Bacteria; Bacteroidetes; Bacteroidia; Bacteroidales; Bacteroidaceae; Bacteroides; Bacteroides vulgatus		3.43	3.43	6.54E-16	177314.0
[Eubacterium] siraeum (Unknown)	Bacteria; Firmicutes; Clostridia; Clostridiales; Ruminococcaceae; Ruminiclostridium; [Eubacterium] siraeum		3.29	3.29	1.78E-7	66423.0
Roseburia faecis (Unknown)	Bacteria; Firmicutes; Clostridia; Clostridiales; Lachnospiraceae; Roseburia; Roseburia faecis		3.02	3.02	0.04	7539.0
Sutterella wadsworthensis (Unknown)	Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales; Sutterellaceae; Sutterella; Sutterella wadsworthensis		2.97	2.97	3.33E-3	15049.0
Alistipes putredinis DSM 17216	Bacteria; Bacteroidetes; Bacteroidia; Bacteroidales; Rikenellaceae; Alistipes; Alistipes putredinis	GCA_000154465.1	2.94	2.94	3.41E-13	66253.0
Alistipes shahii (Unknown)	Bacteria; Bacteroidetes; Bacteroidia; Bacteroidales; Rikenellaceae; Alistipes; Alistipes shahii		2.79	2.79	3.90E-21	105018.0
Bacteroides coprophilus (Unknown)	Bacteria; Bacteroidetes; Bacteroidia; Bacteroidales; Bacteroidaceae; Bacteroides; Bacteroides coprophilus		2.79	2.79	0.02	7411.0
Clostridium sp. L2-50 (Unknown)	Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae; Clostridium; Clostridium sp. L2-50		2.26	2.26	1.49E-3	6318.0
Alistipes timonensis (Unknown)	Bacteria; Bacteroidetes; Bacteroidia; Bacteroidales; Rikenellaceae; Alistipes; Alistipes timonensis		1.95	1.95	2.44E-5	6171.0
Roseburia intestinalis (Unknown)	Bacteria; Firmicutes; Clostridia; Clostridiales; Lachnospiraceae; Roseburia; Roseburia intestinalis		1.82	1.82	6.08E-18	19795.0
[Eubacterium] siraeum DSM 15702	Bacteria; Firmicutes; Clostridia; Clostridiales; Ruminococcaceae; Ruminiclostridium; [Eubacterium] siraeum	GCA_000382085.1	1.75	1.75	2.54E-32	16028.0
Bacteroides fluxus (Unknown)	Bacteria; Bacteroidetes; Bacteroidia; Bacteroidales; Bacteroidaceae; Bacteroides; Bacteroides fluxus		1.46	1.46	8.03E-5	2621.0
Subdoligranulum sp. 4_3_54A2FAA (Unk...)	Bacteria; Firmicutes; Clostridia; Clostridiales; Ruminococcaceae; Subdoligranulum; Subdoligranulum sp. 4_3_54A2FAA		1.37	1.37	1.32E-13	7051.0
Oscillibacter sp. ER4	Bacteria; Firmicutes; Clostridia; Clostridiales; Oscillospiraceae; Oscillibacter; Oscillibacter sp. ER4	GCA_000765235.1	1.31	1.31	6.15E-7	2877.0
Bacteroides oleiciplenus (Unknown)	Bacteria; Bacteroidetes; Bacteroidia; Bacteroidales; Bacteroidaceae; Bacteroides; Bacteroides oleiciplenus		1.27	1.27	2.78E-25	11113.0
Alistipes obesi (Unknown)	Bacteria; Bacteroidetes; Bacteroidia; Bacteroidales; Rikenellaceae; Alistipes; Alistipes obesi		1.27	1.27	5.16E-45	19989.0
Pseudoflavonifractor capillosus (Unkno...)	Bacteria; Firmicutes; Clostridia; Clostridiales; ; Pseudoflavonifractor; Pseudoflavonifractor capillosus		1.23	1.23	3.50E-7	2438.0
Coprobacter secundus (Unknown)	Bacteria; Bacteroidetes; Bacteroidia; Bacteroidales; Porphyromonadaceae; Coprobacter; Coprobacter secundus		1.18	1.18	5.27E-8	2629.0
[Eubacterium] rectale (Unknown)	Bacteria; Firmicutes; Clostridia; Clostridiales; Lachnospiraceae; ; [Eubacterium] rectale		1.16	1.16	1.19E-111	40449.0
Bacteroides ovatus (Unknown)	Bacteria; Bacteroidetes; Bacteroidia; Bacteroidales; Bacteroidaceae; Bacteroides; Bacteroides ovatus		1.14	1.14	0.00	73783.0
Riemerella columbina (Unknown)	Bacteria; Bacteroidetes; Flavobacteriia; Flavobacteriales; Flavobacteriaceae; Riemerella; Riemerella columbina		1.12	1.12	4.18E-5	1357.0
Dorea formicigenerans (Unknown)	Bacteria; Firmicutes; Clostridia; Clostridiales; Lachnospiraceae; Dorea; Dorea formicigenerans		1.04	1.04	2.90E-13	3214.0
Bacteroides uniformis (Unknown)	Bacteria; Bacteroidetes; Bacteroidia; Bacteroidales; Bacteroidaceae; Bacteroides; Bacteroides uniformis		1.01	1.01	0.00	49818.0
Faecalibacterium prausnitzii A2-165	Bacteria; Firmicutes; Clostridia; Clostridiales; Ruminococcaceae; Faecalibacterium; Faecalibacterium prausnitzii	GCA_000162015.1	0.72	0.72	0.00	15656.0
Butyriconas virosa DSM 23226	Bacteria; Bacteroidetes; Bacteroidia; Bacteroidales; Odoribacteraceae; Butyriconas; Butyriconas virosa	GCA_000519105.1	0.67	0.67	0.00	16225.0

Table Settings

Column width: Manual

Show column:

- Name
- Taxonomy
- Assembly ID
- Combined Abundance
- Min
- Max
- Mean
- Median
- Std
- S1_day0
- Confidence score (S1_day0)
- Total reads (S1_day0)

Select All / Deselect All

Data

Show abundance values as:

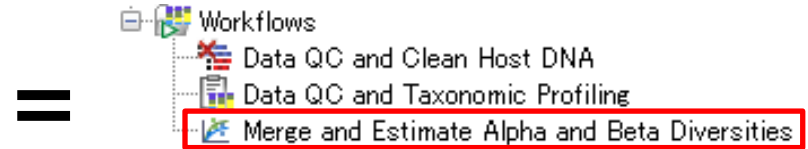
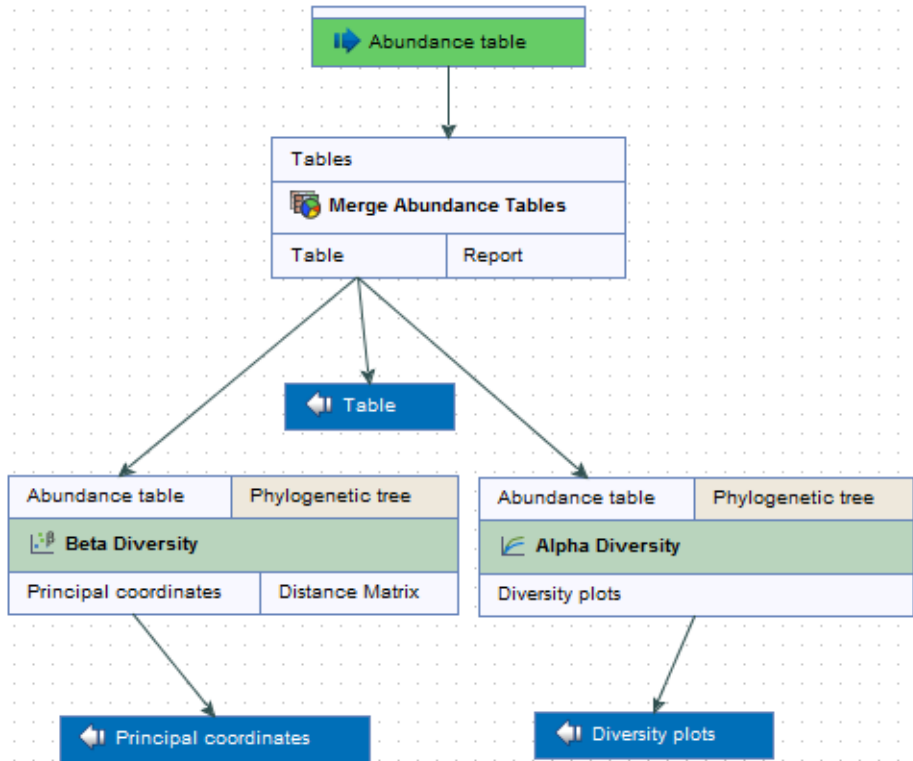
- Raw
- Relative

Aggregate feature: Name

Aggregate sample: Name

- Amplicon-Based OTU Clusteringと違い、各サンプルごとにデータが作成される。
- データの表示については、Amplicon-Based OTU Clusteringの場合と同じで、Aggregate taxonomy 項目から、界～種などのカテゴリー分類を切り替えたり、バーチャートグラフやサンバースト図での表示が可能。

手順3. 2次解析



Merge Abundance Table

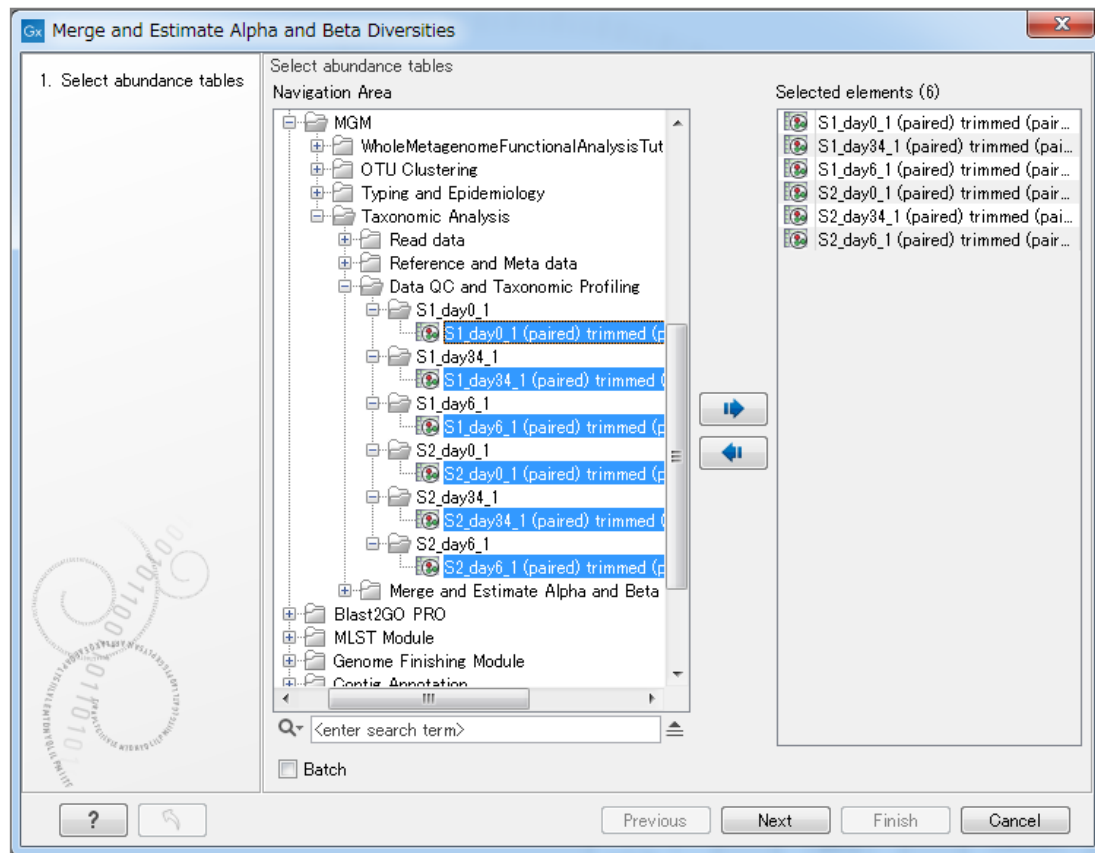
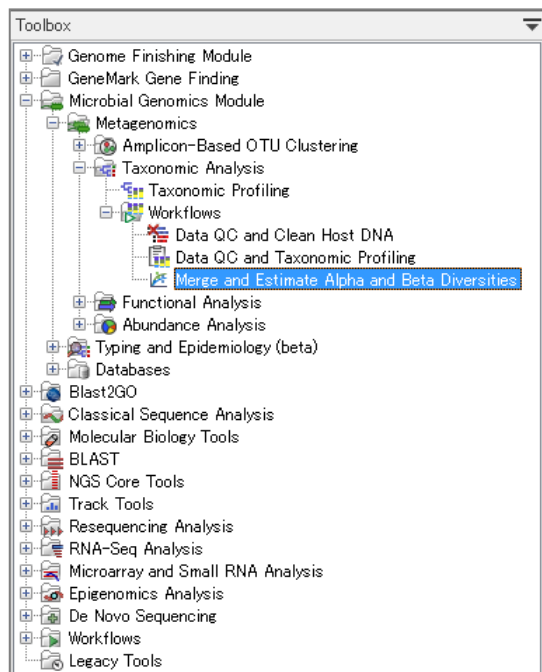
・各サンプルの組成データを統合する。

Alpha Diversity

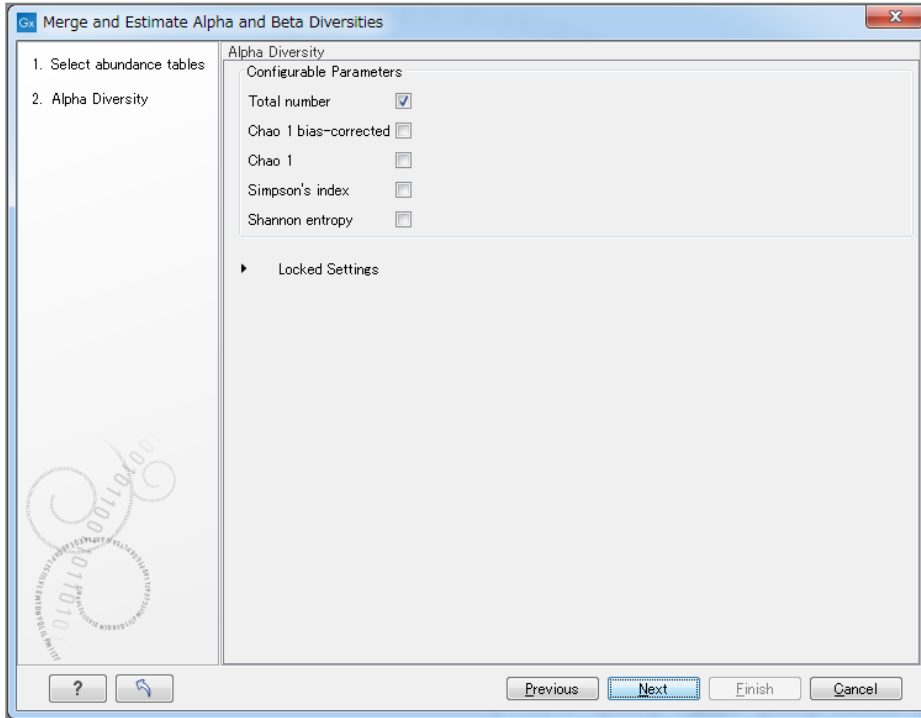
・アルファ多様性を計算する。

Beta Diversity

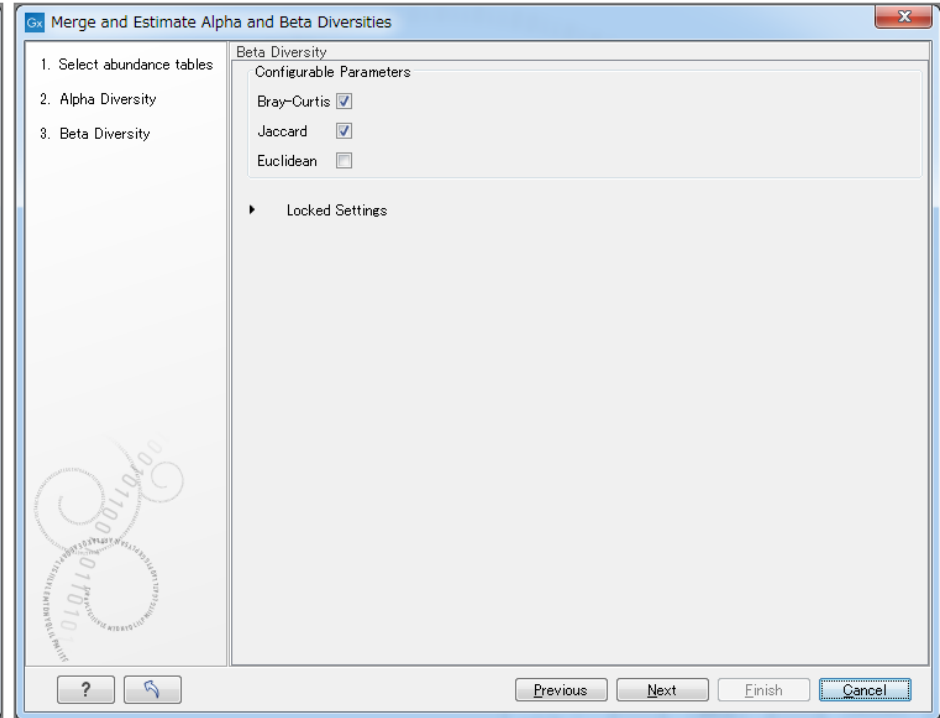
・ベータ多様性を計算する。



1. Merge and Estimate Alpha and Beta Diversitiesを選択してダブルクリック。
2. 組成データを選択。



- Total number ~ Shannon entropy: アルファ多様性の計算アルゴリズムを選択する。



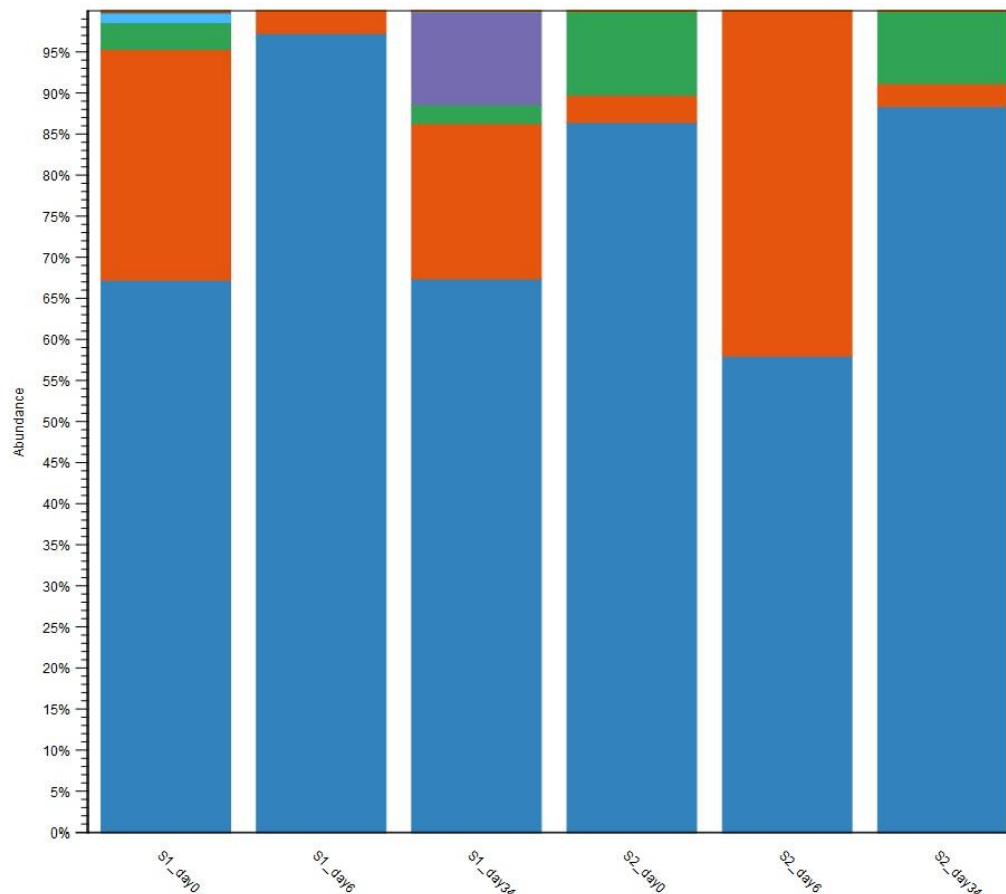
- Bray-Curtis ~ Euclidean: ベータ多様性の計算アルゴリズムを選択する。

Merge and Estimate Alpha and Beta Diversities

merged

merged (Alpha Diversity - Total number)

merged (PCoA - Bray-Curtis)



Stacked Visualization Settings

Settings

Chart Type: Bar Chart

Sort features by: Abund...

Filter level: 100

Show Other

Show legend

Data

Aggregate feature: Class

Aggregate sample

Name

S1_day0

S1_day6

S1_day34

S2_day0

S2_day6

S2_day34

Sort samples by

Name

Color

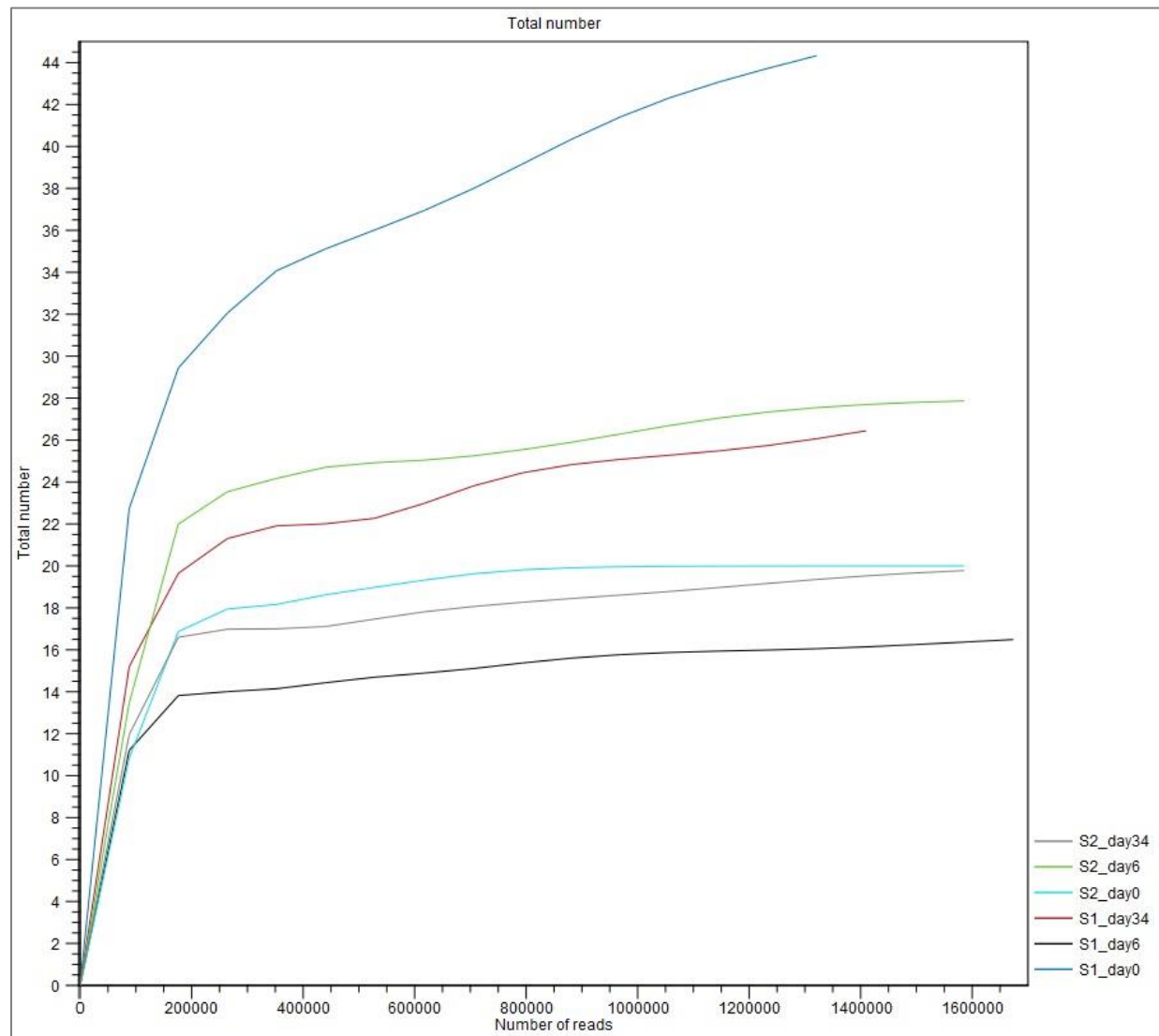
Phylum

- Bacteria;
- Bacteria; Actinobacteria
- Bacteria; Bacteroidetes
- Bacteria; Firmicutes
- Bacteria; Proteobacteria
- Bacteria; Verrucomicrobia

Text format

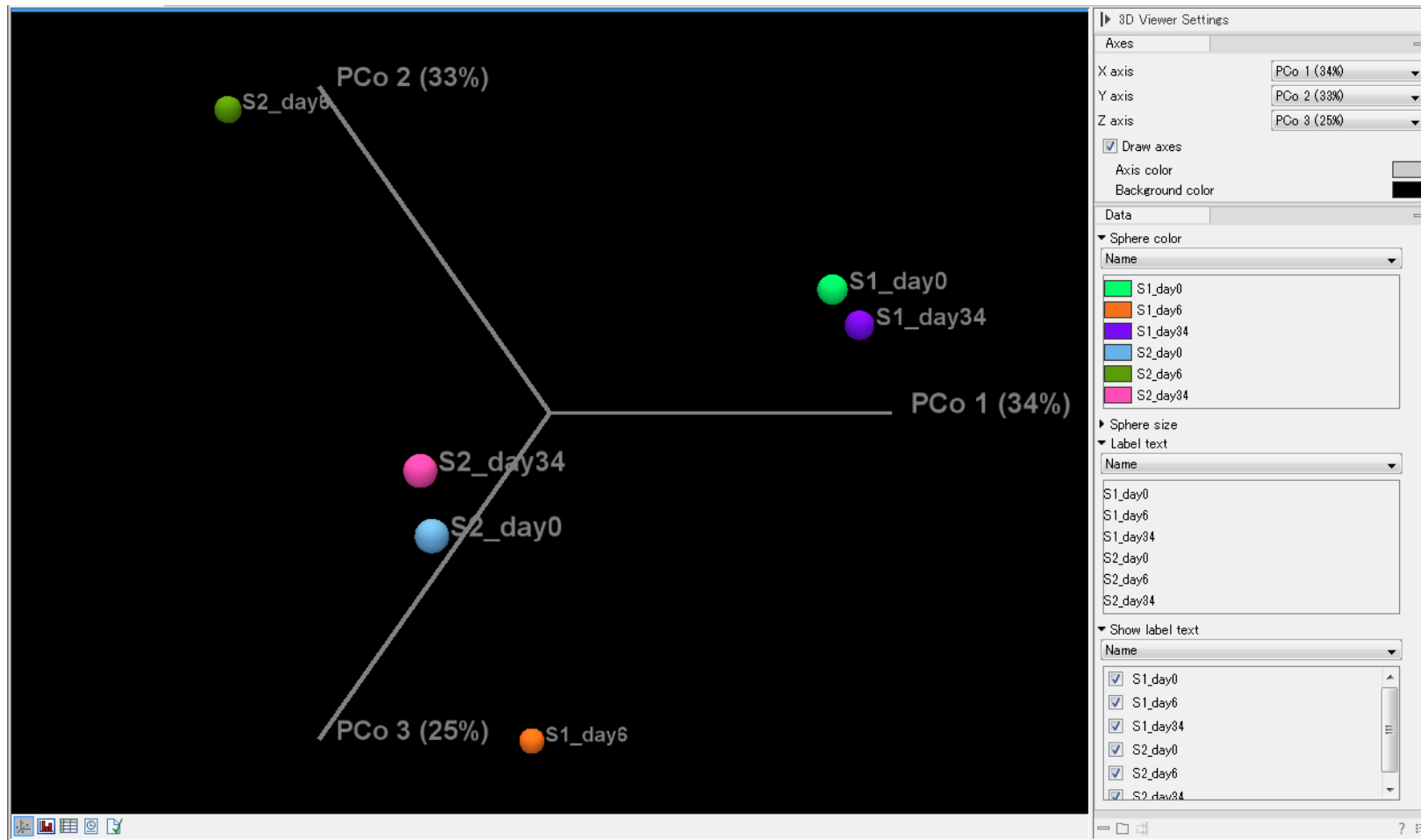
- 統合された組成データでは、界～種などのカテゴリー分類の切り替えや、バーチャートグラフやサンバースト図での表示が可能。

- Merge and Estimate Alpha and Beta Diversities
 - merged
 - merged (Alpha Diversity - Total number)
 - merged (PCoA - Bray-Curtis)



- アルファ多様性データをまとめたレポートが、選択したアルゴリズム別に出力される。

- Merge and Estimate Alpha and Beta Diversities
 - merged
 - merged (Alpha Diversity - Total number)
 - merged (PCoA - Bray-Curtis)



- 選択した計算アルゴリズムごとに、ベータ多様性データが出力される。

お問い合わせ先：フィルジエン株式会社

TEL 052-624-4388 (9:00～17:00)

FAX 052-624-4389

E-mail: biosupport@filgen.jp