

猫も杓子も

Geneious Prime でシーケンス解析



第 13 回 アンプリコンメタゲノミクス(その 5)

メタゲノミクスは、環境サンプルから直接回収された遺伝物質の研究です。前回([第 12 回 アンプリコンメタゲノミクス\(その 4\)](#))から引き続き、発酵プロセスに関連する細菌群をプロファイルするため、自然発酵したザワークラウトから PCR 増幅された 16S rRNA 遺伝子配列を解析する手法をご紹介します。

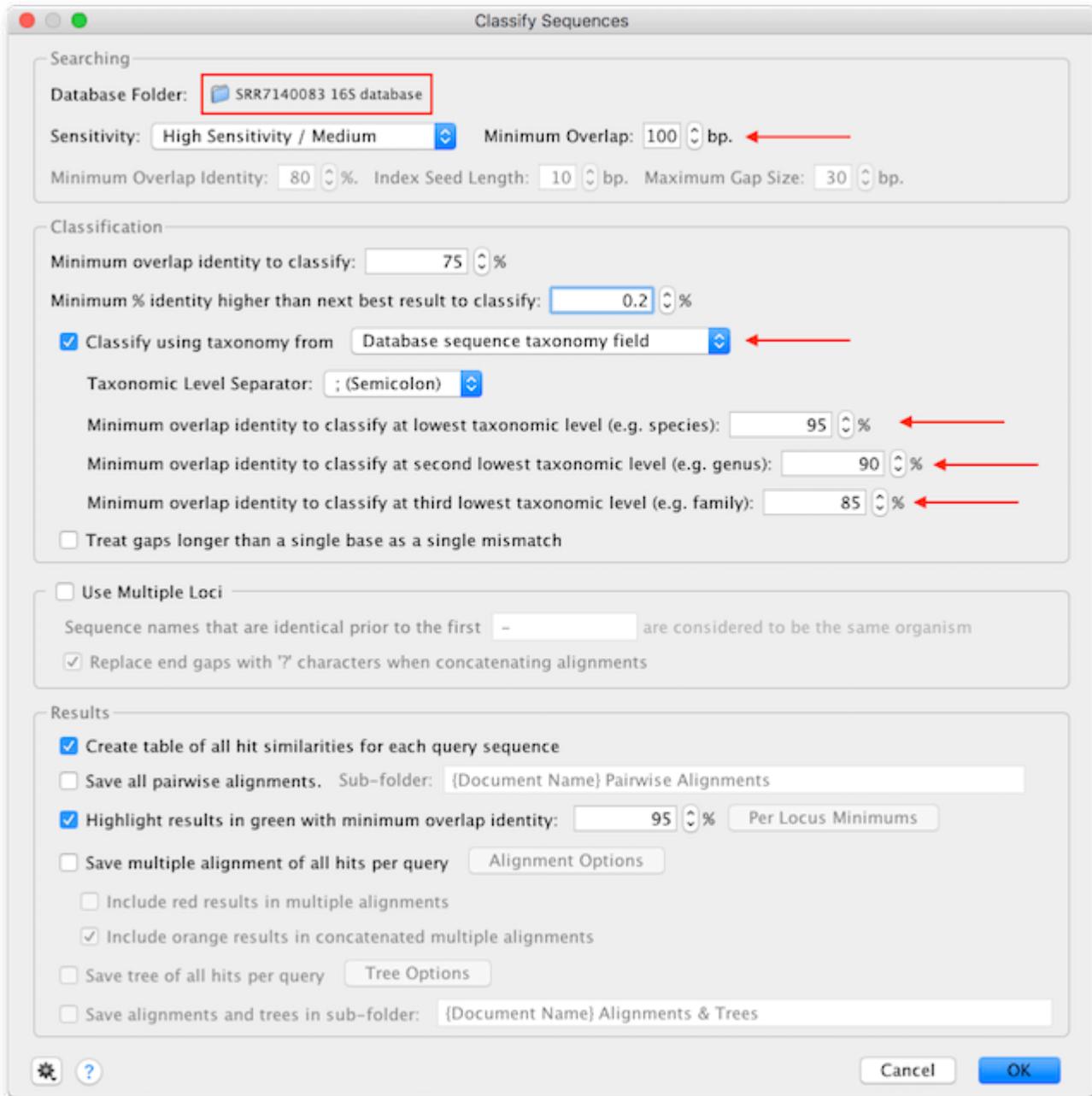
今回は Sequence Classifier プラグインを使用して、前回までに作成した 16S データベースを用いて、全アンプリコンデータセットを解析する手法についてです。

もし Sequence Classifier プラグインをインストールしていない場合は、Tools → Plugins で表示されるプラグインの一覧から、Sequence Classifier を選択してインストールします。

第 9 回アンプリコンメタゲノミクス(その 1)で作成したリードファイルを選択します。デモデータでは **SRR7140083_50000 (trimmed) (merged) - length 150 to 260** というファイル名で、トリム、マージ、length フィルタリング済みのリードセットです。OTU リストではなく、全リードのリストに戻って使用することにご注意ください。

Tools → Classify Sequences に進み、Database Folder で、前回作成した **SRR7140083 16S database** を選択します。デモデータセットに含まれる 16S アンプリコン配列は、16S の V4 領域のみから得られたものであるため、種レベルに分類するために十分な解像度を有していません。そこで、Database sequences taxonomy field を利用して、データベース配列との相同性から、アンプリコンリードを属レベルに分類することにします。

そのためには、以下のスクリーンショットに示すような設定を行います。まず、**Sensitivity** を **High Sensitivity/Medium** に設定し、**Minimum Overlap** を 100 bp に設定します。Classification の設定では、分類元として **Database sequence taxonomy field** を選択し、minimum overlap identity for the lowest taxonomic level を **95%**に、それ以降は **90%** および **85%**に設定します。上述のようにデモデータでは、最も低い分類レベルが種ではなく属であることに注意してください。



OK をクリックすると、Sequence Classifier が実行され、完了すると結果がレポートドキュメントに保存されます。レポートは Summary、Classifications、Results の 3 つで構成されており、**Summary** にはデータベースを使用して基準に従って分類された配列の数が表示されます。また、未分類の配列の数もリストで表示されます。Classifications には、解析に用いられたすべての配列が一覧表示され、マッチした分類の情報が表示されます。Classifications で任意の配列を選択すると、その配列に対する個々の詳細が Results に表示されます。

Classify Sequences Results Text View Lineage Info

Export table

Classified 12,458 out of 12,465 sequences

Summary			Results for SRR7140083.21 21/1			
Classification	Frequency	% Frequency	Database Sequence Name	Overlap Identity	Query Identity	Taxonomy
Bacteria; Firmicutes; Bacilli; Lactobacillales; Leuconostocaceae; Leuconostoc	7,097	56.94%	NR_133769 - BLAST Hit	100%	100%	Bacteria; Firmicutes; B...
Bacteria; Firmicutes; Bacilli; Lactobacillales; Lactobacillaceae; Lactobacillus	5,019	40.26%	NR_075014 - BLAST Hit	99.21%	99.21%	Bacteria; Firmicutes; B...
Bacteria; Firmicutes; Bacilli; Lactobacillales; Leuconostocaceae	135	1.08%	NR_074997 - BLAST Hit	99.21%	99.21%	Bacteria; Firmicutes; B...
Bacteria; Firmicutes; Bacilli; Lactobacillales; Lactobacillaceae	95	0.76%	NR_157602 - BLAST Hit	98.02%	98.02%	Bacteria; Firmicutes; B...
Bacteria; Firmicutes; Bacilli; Lactobacillales	25	0.20%	NR_113255 - BLAST Hit	96.84%	96.84%	Bacteria; Firmicutes; B...
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Yersin...	22	0.18%	NR_136428 - BLAST Hit	95.26%	95.26%	Bacteria; Firmicutes; B...
Bacteria; Firmicutes; Bacilli; Lactobacillales; Streptococcaceae; Lactococcus	14	0.11%	NR_114312 - BLAST Hit	93.28%	93.28%	Bacteria; Firmicutes; B...
Bacteria; Cyanobacteria; Nostocales	14	0.11%	NR_136437 - BLAST Hit	92.49%	92.49%	Bacteria; Firmicutes; B...
Bacteria; Firmicutes; Bacilli; Lactobacillales; Leuconostocaceae; Weissella	9	0.07%	NR_029041 - BLAST Hit	92.09%	92.09%	Bacteria; Firmicutes; B...
Unclassified (no match)	7	0.06%	NR_042194 - BLAST Hit	89.72%	89.72%	Bacteria; Firmicutes; B...
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales; Pseu...	2	0.02%	NR_113289 - BLAST Hit	88.54%	88.54%	Bacteria; Firmicutes; B...
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enter...	2	0.02%	NR_145899 - BLAST Hit	88.14%	88.14%	Bacteria; Firmicutes; B...
Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales; Oxalobact...	2	0.02%	NR_125575 - BLAST Hit	88.14%	88.14%	Bacteria; Firmicutes; B...
Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales; Methylobacte...	2	0.02%	NR_125562 - BLAST Hit	88.14%	88.14%	Bacteria; Firmicutes; B...
Bacteria; Cyanobacteria	2	0.02%	NR_041509 - BLAST Hit	88.14%	88.14%	Bacteria; Firmicutes; B...

Classifications			
Query Sequence	Overlap Identity	Classification	Closest Sequence(s)
SRR7140083.2...	100%	Bacteria; Firmicutes; Bacilli; La...	NR_133769 - BLAST Hit
SRR7140083.3...	95.67%	Bacteria; Firmicutes; Bacilli; La...	NR_042194 - BLAST Hit
SRR7140083.4...	100%	Bacteria; Firmicutes; Bacilli; La...	NR_133769 - BLAST Hit
SRR7140083.4...	100%	Bacteria; Firmicutes; Bacilli; La...	NR_074997 - BLAST Hit
SRR7140083.5...	97.06%	Bacteria; Firmicutes; Bacilli; La...	NR_133769 - BLAST Hit
SRR7140083.5...	99.13%	Bacteria; Firmicutes; Bacilli; La...	NR_157602 - BLAST Hit
SRR7140083.5...	98.06%	Bacteria; Firmicutes; Bacilli; La...	2 best matches. Names: [NR_...
SRR7140083.5...	98.41%	Bacteria; Firmicutes; Bacilli; La...	NR_114844 - BLAST Hit
SRR7140083.6...	95.32%	Bacteria; Firmicutes; Bacilli; La...	NR_114844 - BLAST Hit
SRR7140083.6...	98.02%	Bacteria; Firmicutes; Bacilli; La...	NR_042456 - BLAST Hit
SRR7140083.7...	99.60%	Bacteria; Firmicutes; Bacilli; La...	NR_113821 - BLAST Hit
SRR7140083.7...	95.65%	Bacteria; Firmicutes; Bacilli; La...	NR_133769 - BLAST Hit
SRR7140083.8...	99.21%	Bacteria; Firmicutes; Bacilli; La...	2 best matches. Names: [NR_...
SRR7140083.8...	93.96%	Bacteria; Firmicutes; Bacilli; La...	2 best matches. Names: [NR_...
SRR7140083.9...	99.07%	Bacteria; Firmicutes; Bacilli; La...	2 best matches. Names: [NR_...

デモデータセットは主に *Leuconostoc* と *Lactobacillus* という、ザワークラウト発酵で優勢であることが知られている属で構成されていることがわかります。

データをエクスポートするには、**Summary**、**Classifications**、**Results** の各テーブルを選択し、左上の Export Table をクリックすることで.csv フォーマットでエクスポートすることができます。

Classify Sequences ツールの詳細なマニュアルについては[こちら](#)からダウンロードできます。Sequence Classifier のチュートリアルは[こちら](#)から入手できます。

次回はアンプリコンメタゲノミクスのおまけとして、16S Biodiversity ツールを用いた配列の分類についてご紹介する予定です。

Geneious 製品概要・トライアルリクエストについては[こちら](#)

『Geneious Prime でシーケンス解析』の過去の記事は[こちらでチェック!](#)