

Metagenomics

Taxonomic Classification

各シーケンスの分類群を識別します。OmicsBoxは、16SやWGSに対応し、優れたベンチマークスコアを示します。

豊富な視覚化

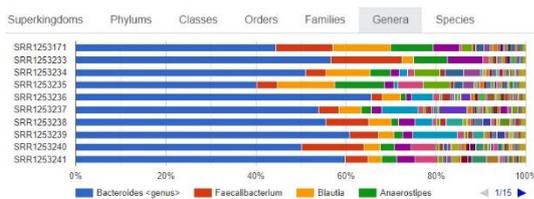
カラフルでインタラクティブなチャートは、データを直感的に探索するのに役立ちます。メタゲノム分類の複雑な階層内で、信頼性スコアを使用して解析は行われます。

サンプル間比較

棒グラフは、たとえば属または種のレベルでのサンプル間の関係を理解することに役立ちます。異なるサンプル間の分類学的な豊富な組成の概要を簡単に示します。

OTU Differential Abundance Testing

edgeRを使用して、異なるグループ内の異なるレベルのOTUを相互にテストし、豊富な差異を判断します。結果に加えて、PDFレポートとヒートマップおよびバブルチャートなどが作成され、調査結果の解釈に役立ちます。またアノテーションが付与されたサンプル間の豊富さをテストし、バブルチャートなどで視覚化をできます。



Taxonomic Classification

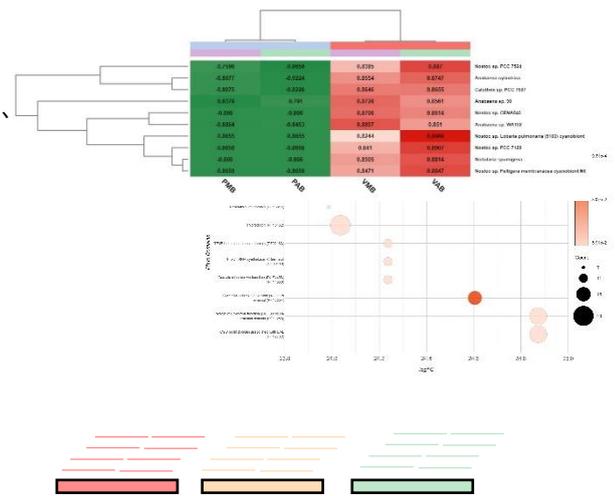
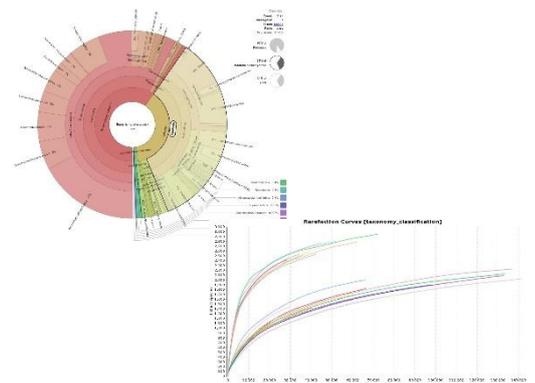
種(バクテリア、古細菌、ウイルス)をKraken2で識別し、マルチレベル円グラフ(Krona)とサンプル間比較棒グラフで結果を視覚化します。他にもChao1の多様性評価、Rare fraction curves、OTU Differential Abundance Testing(edgeR)、PCoAプロットにて評価できます。



Gene Prediction

遺伝子およびタンパク質を予測するためのアプリケーションです。同定し抽出するために、断片化されたリードにはFragGeneScanを使用し、組み立てられたデータにはProdigalを使用します。

Rank	TaxID	Scientific Name	PAB1	PAB2	PAB3
1	2	Bacteria <phylum>	11404	8832	10183
2	231174	Actinobacteria <phylum>	1421	998	1180
3	80864	Comamonadaceae			1604
4	1760	Actinobacteria <class>			1012
5	356	Rhizobiales			850
6	119060	Burkholderiaceae			553
7	976	Bacteroidetes <phylum>			377
8	1239	Firmicutes			458
9	72274	Pseudomonadales			444
10	135621	Pseudomonadaceae			430
11	204457	Sphingomonadales			285
12	286	Pseudomonas			408
13	28221	Deltaaproteobacteria	405	272	283



Metagenomic Assembly

通常のリードでは長さが短すぎて機能の特性評価が困難です。したがって、前のステップとして、メタゲノムアセンブリツールを実行します。MetaSPAdesとMEGAHITのどちらかを選択すると、大規模なデータセットを簡単かつ迅速にクラウドで構築できます。

GO ID	Name	Definition
GO:0000155	phosphotransferase activity	Catalysis of the phosphorylation of a substrate molecule in response to selection of an extracellular signal such as a chemical signal or change in environment, to initiate a change in cell state or activity. The non-covalent source of a phosphate kinase that autophosphorylates a histidine residue in its active site. The phosphate is then transferred to an separate residue in a downstream response molecule, to trigger a response.
GO:0016021	integral component of membrane	The component of a membrane consisting of the gene products and proteins complexes having at least some part of their molecule covalently embedded in the hydrophobic region of the membrane.
GO:0000160	phosphotransferase activity	A conserved system of molecular signals known as phosphorelay and P_{2U} involves autophosphorylation of a histidine kinase and the transfer of the phosphate group to an separate that then acts as a phosphate donor to response regulator proteins.

Functional Interpretation

EggNOG-Mapper(オルソロググループ) PfamScan(ドメインおよびファミリー)を使用して機能アノテーション情報を付与します。結果はGOグラフやチャートで視覚的に表現し比較することができます。



Metagenomicsモジュールでなにができる？

【使用したデータセット】

2つのソーダ湖(高レベルの炭酸ナトリウムと塩分とpHの上昇)Preta湖・Verde湖より午前と午後の2つのタイムポイントで3サンプルずつサンプルを使用。OmicsBoxに搭載されたFastQCで品質をチェックしTrimmomaticでアダプタトリミングを行いました。

【Taxonomic Classification】

クリーンリードを使用してOmicsBoxではKrakenに基づいたTaxonomic Classificationを行うことができます。またこの結果から様々な図の作成が可能です。
まず、PCoAプロットを作成しました。このプロットによりサンプル間の類似パターンを距離で表すことができます。

次にTaxonomic Classificationの結果を門から種までの構成グラフを作成することができます。本データの場合、門の構成グラフは、湖の間には大きな違いがありますが、サンプリング時間(午前と午後)の間に明確な違いは観察されないことや、Verde湖の支配的な門はシアバクテリア(80%)、Preta湖では、主な門はプロテオバクテリア(75%)であることが視覚的に理解することができます。

OmicsBoxではRarefaction Curvesの作成も可能です。希薄化は、生態学で広く使用されている手法であり、OTU分析に適用されます。この図によりシーケンスのカバレッジが十分に深いかどうかを判断することができます。さらにDiversity Curveの作成もでき、追加のサンプルをデータセットに含めることの微生物の多様性における利点を評価することが可能です。

サンプル間の分類学的な違いをよりよく理解するために、OmicsBoxでDifferential Abundance Testingを実行できます。このツールは2つの微生物群の間で著しく異なる分類単位(OUT)を識別するツールです。階層クラスタリングとして図で表すことも可能です。

【アセンブリと遺伝子発見・機能アノテーション】(一部Functional Analysis Moduleを使います。)

OmicsBoxではMEGAHITを使用して各サンプルを組み立て、FragGeneScanを使用して各アセンブリの遺伝子と遺伝子断片の予測が可能です。本データの場合、これらの解析を行うことでBlastでトップヒットした種の比較や機能情報を得ることができます。

ここまでのデータを使用しサンプル比較チャートの作成が可能です。機能アノテーション情報をまとめサンプル間で比較することでより理解を深めることができます。

Metagenomicsモジュールでは微生物群集の同定から比較までグラフィカルな結果を出力できます。

250000

予測遺伝子

※リードデータ量：6.7M

90%

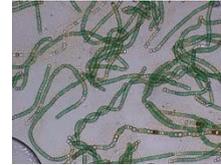
高品質のトランスクリプト

6時間

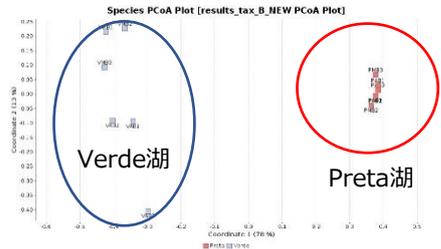
解析時間

比較チャートの結果、Verde湖は、光合成や窒素化合物代謝に関連する機能が特に豊富でありこれらはシアバクテリアの特徴と一致した。対しPreta湖ではこれらの機能は重要ではないことを示した。

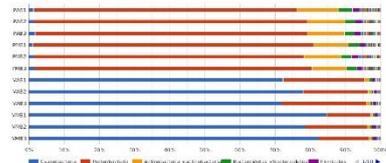
本データはWGSデータですが16SrRNAデータを使用してTaxonomic Classificationを行うことも可能です。



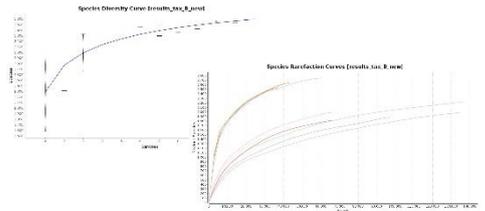
Verde湖でシアバクテリアが多く確認されているのに対しPreta湖ではVerde湖ほど確認されません。したがって、湖間には異なる微生物群集と機能的組成が存在することが予想された。



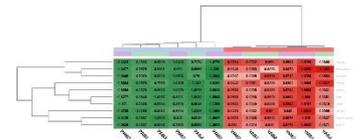
PCoAプロット：2つの湖の分類学的組成に大きな違いが存在することを示唆した。



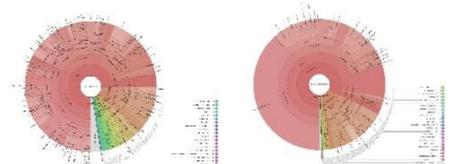
構成グラフ：シアバクテリアの割合はVerde湖で見つかった割合と比較してPreta湖で非常に低かったことがわかる。



データセットにサンプルを追加しても効果が向上しないことを示唆した。特に微生物の多様性のレベルがはるかに高いPreta湖のサンプルの場合、これらの環境に存在するすべての種をカバーするにはシーケンスの深さが十分ではなかったことを示唆している。



Differential Abundance Testingの結果、Preta湖と比較した場合Verde湖で最も豊富な属がシアバクテリアに対応することを示した。



OmicsBoxではKronaチャートの作成もできます。このグラフを使用すると、サンプルを簡単に比較できます。

