

研究経過報告書

1. 研究課題名 ロングリードメタゲノム解析の臨床応用を目指した検査基盤の創出

2. 研究代表者氏名 太田 悠介

3. 研究発表

1. Yusuke Ota, Kei Yasunaga, Samiratu Mahazu, Isaac Prah, Satoshi Nagai, Yoshiaki Gu, Ryoichi Saito. Comparative evaluation of analytical pipelines for Illumina short- and Nanopore long-read 16S rRNA gene amplicon sequencing with mock microbial communities. In submission.
2. 太田 悠介, 齋藤 良一. 16S rRNA アンプリコン解析におけるショート及びロングリード解析パイプラインの比較評価. 第 70 回日本臨床検査医学会学会学術集会. 一般演題採択済み.

4. 研究実績

1. 背景

近年、有病率や死亡率の高い糖尿病や大腸ガンなどの疾患において、体内微生物叢の乱れと疾患との関係性が明らかとなり、腸内フローラ移植療法などの微生物叢をターゲットとした治療法が確立されつつある。微生物群集の解析には、次世代シーケンサーにより細菌の 16S rRNA 遺伝子を網羅的に調べる 16S メタゲノム解析が活用され、主にシーケンス長は短い正確性の高いショートリードシーケンサーとその解析ソフトウェアである QIIME2 が研究レベルで用いられるが、導入コストや装置サイズなどの問題から一般的な検査施設での実用化は困難である。一方、正確性に課題のあったロングリードシーケンサーは、近年の技術改良によりその解析精度が向上した。Oxford 社のナノポアロングリードシーケンサーは、従来機器と比べ安価且つ小型なデバイスであることから、一般検査施設での 16S メタゲノム解析実装の可能性を秘めているが、現在一般的に活用されている EPI2ME ソフトウェアによるロングリードデータの解析では、低品質のリードが生成され細菌の誤分類を引き起こす。近年、ロングリードデータの新規解析パイプラインである Emu が登場し、期待値最大化アルゴリズムによりシーケンスエラーを低減し、正確な分類が可能であることが報告されている。しかし、このワークフローが従来の解析手法と比較し有用かどうか検討されていない。本研究では、16S rRNA 遺伝子の構成が既知の模擬サンプルを用いて、Emu パイプラインの分類結果を現行のショートリード (QIIME2) およびロングリード解析パイプライン (EPI2ME) の分類結果と比較することにより、Emu パイプラインの有用性を評価した。

2. 対象と方法

細菌 8 菌種の正確な 16S rRNA 遺伝子比率が既知の市販試料と、各細菌の 16S rRNA 遺伝子をプラスミドへクローン化し、それぞれを等量混合することで 16S rRNA 遺伝子のコピー数を統一した自家試料を使用した。MiSeq を用いた 16S rRNA V3-V4 領域のショートリードシーケンス後、QIIME2 により細菌叢解析を行った。MinION による 16S rRNA 全長のロングリードシーケンス後、EPI2ME と Emu により細菌叢解析を行った。種レベルにおいて適切に同定された分類群数、誤分類・未分類リードの割合と、精度の指標である F 値 (1 に近いほど高精度) を求め、方法間で比較した。

3. 結果及び考察

各パイプラインの解析結果を下図に示す。市販試料を対象とした各パイプラインでの分類群数において、Emu では全 8 菌種を検出できた一方で、QIIME2 では 1 菌種、EPI2ME では 7 菌種のみ同定された。未分類・誤分類リードの割合は、QIIME2、Emu、EPI2ME で各 84.1%、0.1%、46.1%であった。F 値は、QIIME2、Emu、EPI2ME で各 0.274、0.854、0.465 であった。自家試料を対象とした各パイプラインでの分類群数において、Emu と EPI2ME では全 8 菌種を検出できた一方で、QIIME2 では 2 菌種のみ同定された。誤分類・未分類リードの割合は、QIIME2、Emu、EPI2ME で各 63.3%、0.1%、10.3%であった。F 値は、QIIME2、Emu、EPI2ME で各 0.366、0.831、0.608 であった。以上より、Emu による解析が最も高い分類精度を示すと考えられた。

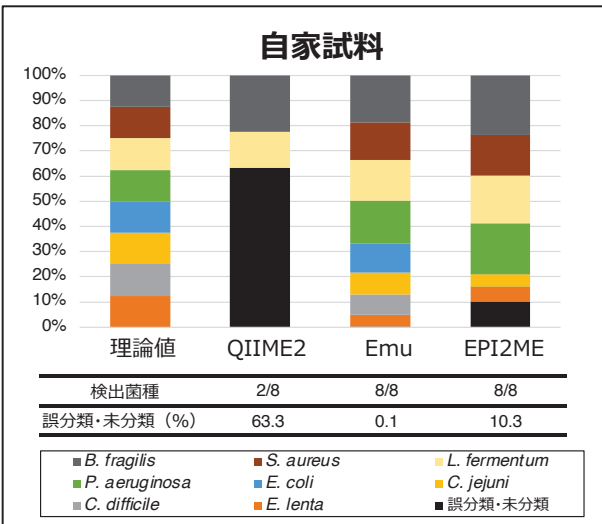
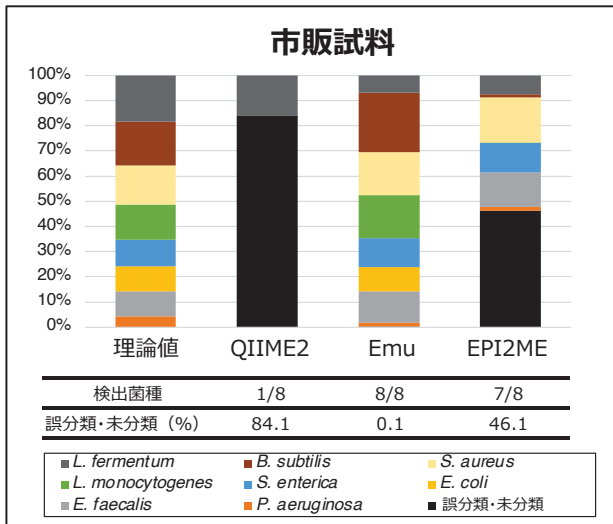


図. 各パイプラインによる解析結果の比較

4. 今後の展望

16S メタゲノム解析におけるロングリードシーケンサーと解析パイプラインである Emu の有用性を確認できたことから、今後は方法の更なる迅速化・最適化のためヒトゲノム解析センターのスーパーコンピュータシステム SHIROKANE の活用も検討する。更には、16S rRNA 遺伝子領域のみならず、微生物の全ゲノム配列を対象としたショットガンメタゲノム解析を用いた検査法の確立を試みる。ショットガンメタゲノム解析では、菌の病原性や薬剤耐性に関わる遺伝子を含むゲノム全体を網羅的に取得し解析するため、微生物叢全体が持つ機能についての知見を得ることが予想できる。本研究成果である 16S メタゲノムデータを基に、微生物叢解析におけるショットガンメタゲノムデータの有用性を評価し、微生物叢の組成のみに留まらず微生物叢が担う役割を解明するための検査基盤を構築する。