

湖沼に生息する細菌・ウイルスのメタゲノム解析
Metagenomics of prokaryotes and viruses in freshwater systems

京都大学 化学研究所 化学生命科学研究領域 岡崎 友輔

研究成果概要

本研究では、京都大学化学研究所スーパーコンピュータシステムを利用し、湖沼に生息する細菌およびウイルスのメタゲノム解析を行っている。今年度は主に、琵琶湖沖において2水深、12カ月にわたり採集した原核生物画分の24サンプルについて、ショートリードおよびロングリードシーケンサーを併用したハイブリッドメタゲノム解析を実施した。まずロングリードをFlyeまたはRavenでアセンブリし、RaconおよびMedakaを用いたロングリードベースでのエラー補正、次いでpilonおよびRaconを用いたショートリードベースでのエラー補正を行った。得られたコンティグはMetaBATまたはMaxBinを用いてビニングし、metagenome assembled genome (MAG)を得た。それぞれのアセンブラ・ビニングソフトで得られたMAGをfastANIを用いて塩基配列相同性に基づいてクラスタリングし、quickmergeを用いて統合することで冗長性を取り除いた。さらに異なるサンプルで得られたMAG間にも同様の手法を適用し、24サンプル間での冗長性も取り除き、最終的に4648個のMAGが得られた。これらのMAGに元のシーケンスリードをマッピングしたところ、8割以上のリードが張り付いたことから、得られたMAGはサンプル中の生物多様性をおおむね網羅していることが示された。さらに得られたMAGのうち303個が高品質(completeness>90%, contamination<5%)の細菌ゲノムであり、そのうちの106個はコンティグ数10個以下、31個は単一のコンティグにアセンブルされた細菌ゲノムであった。一方で、completenessが低いMAGを精査すると巨大核質DNAウイルス(NCLDV)や、jumbo phage等のウイルスゲノムが多数含まれており、これらの多くは100 kbを超える長さにも関わらず、1〜数本のコンティグにアセンブルされていた。総じて、従来のショートリードメタゲノム解析では得られなかった高品質のゲノム情報が得られ、環境微生物・ウイルスのゲノム解析におけるハイブリッドメタゲノム解析の有効性を示すことができた。今後は、得られた高品質なデータを用いて、従来法ではアプローチできなかった、環境中の未培養細菌・ウイルスを対象としたゲノムの微小多様性(Microdiversity)の解明を進める計画である。

発表論文

なし