令和3年度　京都大学化学研究所 スーパーコンピュータシステム 利用報告書

バイオインフォマティクスを用いた生命生態・進化の解明
Using bioinformatics to understand the ecology and evolution of life

京都大学 薬学部 薬科学科 統合ゲノミクス　井出詩菜

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，琵琶湖から採取された核細胞質性大型DNAウイルスと推定される巨大ウイルスの系統的新規性および遺伝子機能の解明を試みた．なお，巨大ウイルスは粒子径が大きいため，原核生物のサイズ画分を対象としたサンプルから副次的に得られた．

次世代シーケンシング(NGS)技術の発達により，培養を介さずに微生物群集サンプル中に存在するゲノムを包括的に解析する「メタゲノム解析」が近年可能になってきた．その結果，様々な系統の巨大ウイルスが世界中の水圏および土壌に存在していることが明らかにされた．また，巨大ウイルスが宿主の集団動態の決定を介して，物質循環や赤潮動態にも深く関係している可能性も示されており，巨大ウイルスの生物学的機能の解明は微生物生態学の重要な課題である．

本研究において対象としたコンティグ(sbin546_03_JunE)は，先行研究[i]のロングリードメタゲノム解析で得られたウイルスゲノム配列である.815kb のコンティグ長，一部の巨大ウイルスに特有の高いAT含量，多数のリピートといった特徴から，ショートリードを基盤とした従来法ではアセンブルが困難であった新規巨大ウイルス系統である可能性が高いと考え，未知のウイルスコンティグについて，アノテーションや系統解析を用いてそのホストや遺伝子を探索した．コンティグにコードされる遺伝子は *Mimiviridae* に高い類似性を示すものが多く，PolB の系統解析からも，このコンティグが由来するウイルスは *Imitervirales* に属するものと判断された．ホストの予測には至らなかった．リピート配列の最大長がショートリードによりアセンブリされたウイルスよりも顕著に長いと予測したが，リピートに関しては，sbin546_03_JunE の平均配列長および最大配列長は予想よりも短かった．しかし，同じリピート配列が多く含まれているという点は興味深い．

本研究成果（の一部）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て得られたものです。

---

[i] 岡崎友輔，湖沼に生息する細菌・ウイルスのメタゲノム解析，京都大学化学研究所スーパーコンピュータシステム研究成果報告書, vol.2020, pp.34