

海洋メタゲノム解析
Marine metagenome analysis

京都大学 化学研究所 化学生命科学研究領域 隈部 彰彦

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、ウイルスと原核生物遺伝子の多様性を比較した。本研究は、以下の 3 ステップで構成されている。

- ステップ1) データベースに登録しているウイルス遺伝子と原核生物遺伝子の多様性比較
- ステップ2) ウイルス遺伝子の多様性について、データベース中と環境サンプル中での比較
- ステップ3) 環境サンプルにおけるウイルス遺伝子と原核生物遺伝子の多様性比較

ステップ1では、データベース中に登録している全ウイルス遺伝子と全原核生物遺伝子をクラスタリングし、その結果からウイルス遺伝子と原核生物遺伝子の推定値を得た。クラスタリングとは、遺伝子を配列類似度によって分類し、グループ化することで、出来たグループをクラスターと呼ぶ。推定値とは、得られたクラスターの数や出来たクラスターの特徴から、サンプルの母集団の数を推定した値である。ステップ1では、ウイルス遺伝子のクラスターの数を実測値で 111,981、推定値で 403,476 であった。一方、原核生物遺伝子のクラスターの数、実測値で 2,951,860、推定値で 9,536,656 であった。ウイルス遺伝子のクラスターの数、原核生物遺伝子のクラスターの数の実測値で 3.79%、推定値で 4.23%であった。

ステップ1で、データベース中の遺伝子多様性を比較したが、データベースに登録されている遺伝子には偏りがある。なぜならデータベースに登録されている生物やウイルスは基本的に実験室で培養可能なものを対象としているからだ。そこでステップ2で、ウイルス遺伝子について、データベース中と環境サンプルで多様性を比較した。

ステップ3では、環境サンプル中で、ウイルス遺伝子と原核生物遺伝子の多様性を比較した。ある環境サンプルにおいて、ウイルス遺伝子と原核生物遺伝子を比較すると、ウイルス遺伝子のクラスターの数、実測値で 121,401、推定値で 486,500 であった。一方、原核生物遺伝子は実測値で 132,083、推定値で 586,438 であった。ウイルス遺伝子のクラスターの数の方は原核生物遺伝子のクラスターの数の実測値で 91.9%、推定値で 82.9%であった。依然、ウイルス遺伝子のクラスターの数の方が原核生物遺伝子よりも少なかったが、データベース中の差よりは小さいことがわかった。

発表論文:なし