

巨大ウイルスのプロモーターの配列解析
Analysis of the promoter sequence of giant virus

京都大学 化学研究所 バイオインフォマティクスセンター 唐 威

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、プロモーターの配列解析を通して、NCLDV あるいは GV のプロモーターの共通祖先を支持する根拠が得られるかどうかを精査する。NCLDV は単系統であると提唱されており、その根拠として、多くのNCLDV に共通する遺伝子(コア遺伝子)の存在が上げられている。一方、こうした単系統性に着目し、プロモーター配列も共通祖先から進化してきたという仮説がある。しかし、プロモーター配列の共通祖先を示唆する根拠は未だ提示されていない。

問題点としては、プロモーター配列は一般に短いため、配列類似性に基づくホモロジーの検出で、どれほど進化の深いところまで、プロモーターの相同性を言えるのか、疑問である。同時に、プロモーターの共通祖先を統計的に検出できなかった場合でも、そのことが共通祖先の否定にはつながらない。

今がやっている作業は、7 個のコア遺伝子に関して、GOEV データベースの全てのゲノムで、promoter や stem-loop が保存されているかどうかを $7 \times N$ のマトリックスで表示する(N はゲノムの数)。次に、上記各遺伝子の上流配列のセットを 7 セット用意し、それぞれのセットに対して motif 抽出と検出を行う。1つの遺伝子に複数の motif が検出されたら motif-1, motif-2 などと分類する。この分類を用いて、 $7 \times N$ のマトリックスを色分けで表示し、上記のマトリックスと比較し、同じような感じか調べる。これにより、コア遺伝子における promoter や stem-loop の保存性及び多様性が少し分かる。