巨大ウイルスのゲノムに保存された制御配列の系統的解明
Systematic investigation of conserved regulatory sequences in the genomes of giant viruses

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　緒方研　　TANG WEI

研究成果概要

Nucleo-cytoplasmic large DNA viruses (NCLDVs) are characterized by their large genome sizes, double-stranded DNA (dsDNA), and high coding capacities. A crucial element in their transcriptional process is the promoter, a region located upstream of transcription start sites (TSS) on the DNA. Promoters facilitate the assembly of the transcription machinery through the recruitment of transcription factors and RNA polymerase, enabling the transcription process. Unlike cellular genomes, which are typically composed of dsDNA and exhibit conserved transcriptional mechanisms, viral genomes display significant diversity in composition, structure, replication, and transcription strategies. These differences have profound implications for virus biology, particularly in the context of virus-host interactions. Most giant viruses encode their own transcriptional machinery, rendering them relatively independent of host systems. Studies have shown that transcription in large DNA viruses follows a temporal pattern in the host cytoplasm, where early, middle, and late-stage genes are regulated by distinct promoter sequences. While promoter sequences differ among these temporal gene classes, conservation patterns within each class suggest evolutionary selection to ensure temporal gene expression during replication in host cells. In Mimivirus, a type of NCLDV, a stem-loop structure from gene downstream has been identified as a key site for polyadenylation, an essential post-transcriptional mRNA modification. This highlights the diverse regulatory mechanisms in NCLDVs. To investigate conserved regulatory sequences in NCLDVs, we analyzed 1,706 metagenome-assembled genomes (MAGs) and reference genomes from the Global Ocean Eukaryotic Viral (GOEV) database. Using the MEME Suite, we extracted promoter motifs from gene upstream, and Rnamotif was employed to identify stem-loop structures from gene downstream. We also quantified the proportion of genes with promoter motifs and stem-loop structures across different NCLDVs. Our findings reveal that, in addition to Mimivirus, some of other NCLDVs also exhibit a high prevalence of stem-loop structures. Notably, certain giant viruses, such as Mirusvirus, show a high expression of promoter motifs. Further investigation is ongoing to explore the correlation between the phylogenetic relationships of conserved regulatory sequences and core genes in NCLDVs, aiming to deepen our understanding of their transcriptional regulation and evolutionary adaptation.