RES-210

A genomic data mining study of nine major non-polio enterovirus serotypes

Qian Yang, Xianfeng Zhou

Abstract

Objectives: This study aimed to elucidate the evolutionary characteristics of nine major non-polio enterovirus serotypes (CVA2, CVA4, CVA6, CVA10, CVA16, CVB3, CVB5, EV-A71, and EV-D68) through genomic data mining, focusing on their spatiotemporal distribution and evolutionary dynamics.

Methods: We employed a data mining framework integrating programming, phylogenetic analysis, Bayesian evolutionary modeling, and selection pressure assessment. Over 40,000 genomic sequences from GenBank were analyzed to reconstruct temporal phylogenies, estimate evolutionary rates, and characterize amino acid variability in the capsid protein. Seasonal decomposition and spatial-temporal trend modeling were applied to evaluate epidemic patterns across six WHO regions.

Results: This study yielded several key findings that include: (1) Distinct biennial/triennial epidemic cycles for EV-D68 and seasonal peaks for HFMD-associated serotypes; (2) The novel "60% to transcend" theory: when the cumulative VP1 nucleotide mutations reach 60%, the cumulative non-synonymous amino acid mutations begin to exceed 60%, particularly in antigenic loops; (3) Episodic positive selection at critical VP1 codons, suggesting immune-driven evolution; (4) The relative genetic diversity trends, with EV-A71, CVA16 and CVA6 have shown sustained expansion, while CVB5 and EV-D68 diversity has plummeted.

Conclusions: This study provides a data-driven framework for global pathogen spectrum monitoring and informs public health strategies against enterovirus threats.

Keywords: Enteroviruses; evolutionary dynamics; selection pressure; capsid proteins; molecular surveillance