Nucleotide sequence of the region coding for 100K and 33K proteins of human enteric adenovirus type 41 (Tak)

Susan B. Siemenda, Norman J. Pieniazek, Jorge Velarde, Jr., Danuta Pieniazek and Ronald B. Luftig

Department of Microbiology, Immunology and Parasitology, Louisiana State University Medical Center, 1901 Perdido Street, New Orleans, LA 70112-1393, 1Parasitic Diseases Branch and 2AIDS Program, Centers for Disease Control, Atlanta, GA 30333, USA

Submitted March 23, 1990

EMBL accession no X52532

Previous sequencing (1, 2) of human adenovirus type 41 strain Tak (3) DNA only partially identified the regions coding for L4 100K and 33K proteins. These two proteins are involved in virion morphogenesis (4, 5). We have now completed this identification by cloning an Ad41 3.1 Kb MstI-EcoRI fragment (map position 64.9% to 74%) in pBluescriptII and sequencing it directly using custom oligonucleotide primers. By sequence analysis and GenBank homology searches we identified an open reading frame of 777 amino acids (base 125 to 2458) coding for the 100K protein. For the 33 K protein gene (bases 2145-2976), consensus donor and acceptor splice sequences were located, which predict the removal of a 178 nucleotides long intron from 33K protein gene transcripts (bases 2503 to 2680) as observed for Ad2 (5).

In addition, five more ORFs were identified (the last three on the complementary strand): (i) 16.5K (bases 480-911) (ii) 7.3K (1618-1815), (iii) 8.0K (1138-914), (iv) 6.9K (1473-1291), and (v) 12.4K (2474-2108). Finally, at base 3017 is the start codon for the pVIT protein (2) and at base 100 (on the complementary strand) is the start codon for the DNA binding protein (1).

REFERENCES