

Sequence of human enteric adenovirus type 41 Tak fiber protein gene

Norman J.Pieniazek¹, Susan B.Slemenda¹, Danuta Pieniazek², Jorge Velarde Jr 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 October 17, 1989

EMBL accession no. X16583

The 1.9 kb SmaI-EcoRI DNA fragment (map position 86.4% to 92%) of the human enteric adenovirus type 41 strain Tak(1) was cloned in pBluescriptII and sequenced directly using custom oligonucleotide primers. The gene coding for the fiber protein (protein IV) was identified using the sequence of Ad5 fiber protein gene as reference(2). Ad41 fiber shows high degree of homology with Ad40 fiber(3) except for the shaft region. As found for Ad2(4) and Ad5(2), Ad41 fiber protein gene shaft contains 22 typical amino acid repeats, whereas Ad40 has only 21 such repeats(3); there is 97.7% homology between the amino acid sequences of Ad41 and Ad40 in the knob region.

CCGGCCAACTGCTCATCAAAATCGCCCTAACATCACTTCACTGCTGTACACAGATATAAACAGTGGGTATGCTTTTACTTTAAATGGTCAGCCGAAACCCGTTTCA 120
Fiber Protein: M K R A R L E D D F N P V Y
CCGACCTACCGCTGTATTTTTCYACATAACTGAACATAAATAATCATTCAGCAGCACAATCTTCGCATTCTTTTTCCGAGATGAAACGAGCCAGACTTGAGATGACTCAACCCCGCTC 240
sp100s knobp100
P Y E N H Y F L D I D P F I T T P P F A S S H G L G E K F P G V L S L K Y T D P L Y
ACCTTAGCAGACTACATACCTCCCTTGACATGCCATTTTATACACCCCGCTTTGCTCTCCACAGCGCTTCGACAGAAACACCCGCGGCTCCAGCTTGAAATACACTGATCCACTTA 360
T R N G A L T F L K L G T G L N I D E N G D L S S D A S V E V S A P I T K T F N K I
CAACAAAACAGCGGCTTTAACCTTAAACTG99CACGGGACTAAACATTTGATGAAAATGGAGATCTTTCTCAGATGCTAGCGTGGAASTTAGCCGCCCTATTACTAAACCAACAAA 480
V G L N Y T F P L A L R S N A L T L T S Y N A P L N V V H N H L A L N I S Q P V Y
TCGTAGGTTTAAATACACTAAACCTTCGCCCTCGAAGTAAACGGCTCACTCTTCTTCAACAGCCACCTTAAACGTAGTAAATAAACAAATTTAGCTTTAAATATCTCCAAACCTGTCA 600
V N A N H E L S L L I D A P L N A D T G Y L R L Q S A A P L G L V D K T L K V L
CTGTAAATGCAAAACAGCAACTTCTCTCTCTTAAATAGACGCCCCACTTAAATGCTACACAGGGCACCTCTTCGCCCTCAAAGTCTGCACCTCTCTGGACTAGTGGACAAAACACTAAAGTTT 720
F S S P L Y L D N H F L T L A I E R P L A L S S S R A V Y L K Y S P P L K I E N
TGTTTTCTAGCCCGCTCTATCTAGATAAATAACTTCTTACACTAGCCATTGAACGCCCGCTAGCTCTATCCAGTAGCAGAGCAGTACCCCTTAAGTATTCACCCACTTTAAAAATAGAAA 840
E N L T L S T G G P F T V S G G H L N L T T S A P L S V Q N H S L S L V I T S P
ACGAAACTTAACCTTAAGCACAGCGCGGCTTTACTGTAAAGCGGGGAAATCTAAACTTAACAACAATCGGCACCTCTCCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 960
L K V I N S M L A V G V N P F F T I T D S G L A M D L G D G L A L G G S K L I I
CTTAAAGATTAATATCTATGATGACCTTGGGGTTAACCCGCTTTTACCATCACTGACTCTGGATTAAGCTATGGACTAGGAGCGCTCTGCACCTAGGTGGCTTAAGTTAATAA 1080
N L G P G L Q M S N G A I T L A L D A A L P L O Y R D N O L Q L R I G S T S G L
TCAACTCTGGCTCAGGTTTACAAATGCTAAATGGAGCTTACTTTAGCAGTAAATGCGGCTGCTCTTCCAAATATAGAGACAACTTCAACTCAGAAATGGCTCAACACTGGCT 1200
I N S G V T Q T L N V N A N T G K G L A V E N H S L V T V K L G N G L R F D S W G
TAATATAGCGGATTAACAGCAAACTTAACAGTCAATACCGCAAGGCTCTGCTGTGAAAACAACTCAGTATGTTAGCTGGAAACGGCTCTGCTCTTGTAGCTGG 1320
S I E V S P T T T T P T T L L W T T A D D S P N A T F Y E S L D A K V N L V L V K
GAGCATATAGCTCTGCTCACTACCTACTCCCTACCCACTTACGACTACCCGACAGCCACTACCTAAGCCCACTTTTATGAATCACTAGACCCCAAGTGGCTAGTGTATTAGTAA 1440
C N G N V M H T S I S I E A Q E G I L L R P T A S F I S F V N Y F Y S D G T W R K
AATGCAACCGCATGTTAACGGGACATATCCATTAAGCTCAGAAAGGCAATTTACTTAGCCCTACAGCTAGTTTTATTCTCTTGTGATCTATCTACAGCATGGACATAGAAAG 1560
Y Y S V F D D E G I L A N S A T V N G Y R Q G C S A N T H V S N A V E F M P S S K
AAAACATATCCGCTTTTGCACAGAGGAGTACTAGCAAAAGTGGCCAGCTGGGGTATCTGCAACAGGACGCTTCGCCCACTAACCTTTTCTAATGCTGTAGAAATTTATGCTAGCTCTA 1680
R Y P N H G K G S E V Q N H A L T Y T F L Q G D P N H A I S F Q S I Y H E A L E G
AAGATATCCCAATCAAAAAGGTTGAAATGCTAGAGACTATGCTCTTACCTACCTCTTTTTCAGAGGTATCTTACACTATGCTCTTACAGTATTTATAATCATATGCAATAGAAAG 1800
Y S L K F T W R V R N H E R P D I F C C S F S Y V T E Q
GCTACTCATTAATAATTTACCTGGCGGTTTGGAAATATATGAACTTTTACACTCCCGCTGCTCACTTTTCTATGTATACAGAACTAATAAATAATGTTGTTTTTTTATAACTTTAT 1920
p101A
TGACTTTTACAGAAATTC 1939
EcoRI

Ad41 fiber protein gene. The putative fiber mRNA splice and polyadenylation sites are underlined.

*To whom correspondence should be addressed

References

- 1. van Loon,A.E., Rozijn,T.H., de Jong,J.C., and Sussenbach,J.S. (1985) Virology 140, 197-200.
2. Chroboczek,J., and Jacrot,B. (1987) Virology 161, 549-554.
3. Kidd,A.H., and Erasmus,M.J. (1989) Virology 172, 134-144.
4. Herisse,J., Rigolet,M., Dupont de Dinechin,S., and Galibert,F. (1981) Nucl. Acids Res. 9, 4023-4042.