禽减蛋综合征病毒纤维蛋白基因的克隆与编码产物结构的分析

金 奇　朱 雷　章金钢④　李茂祥④　曾力宇
殷 震④　侯云德

(　中国预防医学科学院病毒研究所病毒基因工程国家重点实验室, 北京 100052;
④中国人民解放军农牧大学, 长春 130062)

摘要　克隆了禽减蛋综合征病毒纤维蛋白基因, 并对其进行了核苷酸序列测定及所编码产物的结构分析. 结果表明禽减蛋综合征病毒纤维蛋白具有所有哺乳类腺病毒纤维蛋白的特征结构, 但是与除羊腺病毒以外的其他腺病毒纤维蛋白在核苷酸及氨基酸序列上有较大差别.

关键词　**禽减蛋综合征病毒　纤维蛋白　基因克隆**

腺病毒是一群具有广泛宿主范围的双链 DNA 病毒. 禽减蛋综合征病毒(EDSV_76)为禽㈣群腺病毒中的唯一成员, 是导致鸡产蛋量下降的主要病原体[1]. 由于 EDSV_76 在生物学和免疫学特性等方面与禽 iv群、禽㈦群及其他动物腺病毒有较大的差别, 所以有关其基因组结构与功能的研究不仅有利于 EDSV_76 及整个腺病毒基因组结构特点、基因调控机制和进化途径等方面的探索, 而且为阐明该病毒的致病机理奠定了基础. 然而, 目前国际上从基因水平对 EDSV_76 的研究基本处于空白.

我们利用从中国分离的 EDSV_76 毒株 AA2, 先后构建了全基因组限制性内切酶 Hind ㈣和 Pst iv的基因文库, 分析了病毒 E3 区及六邻体、p Ⅺ蛋白和巯基内肽酶等基因及编码产物的特点[2,3], 揭示了 EDSV_76 与其他腺病毒有明显差异. 鉴于纤维蛋白是腺病毒编码的一个涉及病毒组装、病毒与受体结合及病毒免疫原性的多种功能蛋白, 本文首次报道了 EDSV_76 纤维蛋白基因和编码产物的序列, 并就其结构特点和与其他腺病毒纤维蛋白的比较进行了分析.

1　材料与方法

EDSV_76/AA2 由中国人民解放军农牧大学从发病鸡体内分离. 病毒接种鸭胚绒毛尿囊腔, 利用超速离心法(Beckman L8_M 超速离心机, SW_40Ti 转子, 40 000 r/min, 3 h)收集尿液中的病毒颗粒, 以终浓度为 1% SDS 和 1 mg/mL 蛋白酶 K 的液体裂解病毒外壳(37 ℃, 过夜). 病毒核酸经酚/氯仿抽提, 无水乙醇沉淀, 溶于 TE 缓冲液.

限制性内切酶 Hind ㈣(Promega 公司)水解的病毒 DNA 在 0.8% 的琼脂糖凝胶电泳分离后, 回收长度为 5.9 kb 的 Hind ㈣B 片段, 经 SPIN_X™ Centrifuge filter unit (COSTER 公司)纯化, 直接克隆至质粒 pBluescript ㈦SK(+)(Stratagene 公司)的相应位点. 重组体的鉴定采用酶切电泳和 Digoxigenin_dUTP 标记的全病毒 DNA 打点杂交等方法.

利用 Exonuclease ㈣的定向降解制备完整片段的双向嵌套质粒, 经 QIAwell_8 plasmid purification mini kit(QIAGEN 公司)纯化的 DNA 序列测定模板, 以 T7 及 T3 寡聚核苷酸为引物, 按照 Taq DyeDeoxy™ Terminator Cycle Sequencing Kit (Perkin Elmer 公司)提供的条件进

行荧光素标记的 DNA 序列测定反应. DNA 序列测定装置为 ABI 373A 型自动测序仪(Applied Biosystems 公司).

全部核苷酸及氨基酸序列的排列、分析和同源比较等工作均利用美国威斯康星大学的遗传分析软件包 GCG 以及 DNASIS 和 PROSIS 分别在 VAX 和 PC 计算机上完成.

2 结果与讨论

EDSV_76 病毒基因组与其他腺病毒差异较大. 为了了解基因组的特点, 我们对采用鸟枪法随机克隆的一些短的片段进行了序列测定, 并对这些片段可能的编码产物进行了分析. 结果发现, 由其中一个片段推导出的氨基酸序列与羊腺病毒纤维蛋白有部分同源性. 利用这个片段作为探针, 对我们所建立的 AA2 株基因组文库进行杂交, 选择 Hind ㉔B 片段进行全序列测定.

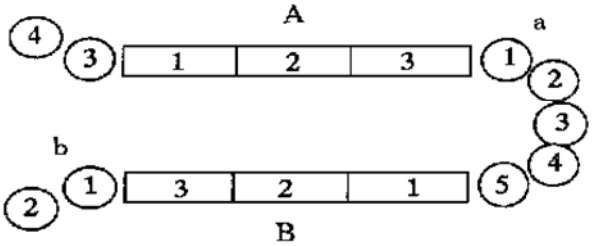
EDSV_76/AA2 株的纤维蛋白基因位于基因组 68.2~ 74.2 图谱单位(m.u.)之间, 全长

```
          10         20         30         40         50         60         70
ATGAAGCGAC TACGGTTGGA CCCTGATCCT GTTTATCCCT TCGAGACGAG CGACACGATC CCAATGCCTC 70
CGTTCATCGA AGCTGGGTCA GGTCTAGCAG TAAATGGACT GCAGCTTTAT ATAACAGCTC AAGCTCCGGT 140
GGGCTTCACC AACAAAGCTG TAACATTAAA ATATGGAGAT GGATTGGAAG TAAATGAAAA TGGAGAACTC 210
ATAGCTACGG CTTCTTCGGC AGTAAAGCCA CCACTCCATT TTGATAGGGG TTATATAGTG TTAAATCTTC 280
AGGATCCATT GGGTGTTATT GATGGGAAGC TTGGGGTCAA GTTAAGCCCT GGGGTTCACA TCAATGGTGA 350
          360        370        380        390        400        410        420
AGGGGCTGTG GCGGTAGAAT CCCCTGTGGA CCCCATTACA CTTGATACGG CTGGTAGAAT TACTTTAAAT 420
TATGGCACAG GTTTAAATGT GAGTGATGGA AAATTACGAC TAGTAAGTCC TGAAAGTCCG CTCACACTTC 490
TTGGAAATGG CAAGGTTGCT CTTAATTTTG GTAACTCAAT GGAGCTTGTG CAAGGGACCT TGCAATTGAA 560
AGCTCCGCTA AATCCTTTGT TCATGACCCC CGCGGGTGCG ATCGGCTTAA GGGTGGATGA CATGTTTAAC 630
ATTTCTGAAG GTTTACTCTC CTTCAAGATG CCATCCGATC CAATTTCGTT TAATGCCGAT GGTATGTTGT 700
          710        720        730        740        750        760        770
CTTTGAACAC AAATGACACA TTGCAAACAA CTGGTGGGCT GTTAGGGTTG ACCGAACCTG CCAAGCCGTT 770
AAAATTGGCC GATGGCAAGT TAGGTGTAAA TGTGGGCCTT GGGTTATCGG TTTCTAATGG GTCATTGACT 840
GTAAATGCAG GGCAGGGGTT GACTATTCGA AATAATGCGG TGGCAGTTAA TGGGGGCAAC ACGCTTGCTT 910
TTAATAATTA TGGAGAGGTG GAAATTAAAA ACCCTAGAAA CCCCATAAGC CTGACCCAAG ATGGTGAATT 980
GGTTTTGATA ATCGGTGATG GCCTAACAAC CCTTGATGGA CGGCTCACTC TACTTACCGC TTCGACCTCT 1050
          1060       1070       1080       1090       1100       1110       1120
CCGATAGCTG TAGGGCCAGC CGGTGTTACA TTTAATGTTA CACCGAGTGA TTTTTACTTT TTATCTCCTA 1120
TATTAGCTCT CAATGTTCAG ACCCGTGGCG GCCTAGAAAA AAGTGACACT GGTTTAAAAA TTAAACGTGC 1190
GGCCCCTCTC AGTATCACAT CTGATGGTGA GTTGACTTTG GCTTATGATT CCACGGATTT TCAGGTGACA 1260
GAAAACGGCC TAGCCCTAAA GGTATCTCCG ACGCAGACCC CTCTCACCAG AATAATTTCT ATGGGAAATA 1330
ACTTGTTTGA TTCTGGTTAT GAGATTTTTG CTTCATGTCC GCAGAACAAA GCACCAAAGG TTGCAGGGTA 1400
          1410       1420       1430       1440       1450       1460       1470
TGTGTATTTA ACATCCGTTG GTGGGCTTGT ACATGGGACC ATTCACATTA AAGCTACTGC GGGGTATTGG 1470
TTTACGGGGG GAAACAGCGT GCACGAAAGT ATCACGTTTG GATTGGTGTT GTGTCCTTTT AATGCTCGCG 1540
ACCCCACTGC TAACCTGTCA AGCTGGCCAA CGCCAGTAGT GTGGAATGGT GATAACAATA CTCCCCTATA 1610
TTTTGCGGCC AATGCCATTA ATTATACCAA TAACCGTGTA AATCTTGCAT TACCGGTACT TTTACAAGGA 1680
GGAAACCGAA TTGCCGGGTT ACACTCGTCA TTCTTTCTGC CCTACCGGGA CCACCGGAAT GAATTTTACA 1750
          1760       1770       1780       1790       1800       1810       1820
GGGGGTAA 1758
```

图 1 EDSV_76/AA2 纤维蛋白基因核苷酸序列

1 758 bp, 共编码 585 个氨基酸残基(aa) (图 1) . 所推导蛋白的分子量大约为 62 ku, 含有 6 个 N 型糖基化位点. 基因的转录方向从左向右. 在其翻译起始信号的上游和终止信号的下游没有腺病毒基因组中常见的 TATA 盒(TATA box) 及多聚腺苷酸加尾信号 AATAAA.

腺病毒纤维蛋白具有一个共同的结构特点, 即由蛋白的 N 氨基端至羧基端依次排列着与五邻体基座非共价结合的纤维尾部、柄及顶端球型结 3 个结构域(结构域 iv, ㈦和 ㈣)[4]. 氨基酸序列同源排列和含有特征序列的结构基本单位分析表明: EDSV_76/ AA2 纤维蛋白的 3 个结构域分别由 35, 401 及 149 个 aa 组成, 其中结构域 ㈦中含有 25 个拟重复基序(motif) (图 2) . 每个拟重复基序中 A_1, A_3, B_1 和 B_3 位为疏水性氨基酸残基. 由 a_1~ a_5 和 b_1~ b_4 位氨基酸残基构成的五肽和四肽 β 转角分别将由 A_1~ A_3 和 B_1~ B_3 位氨基酸残基组成的 2 个三肽 β 链以反向平行的方式连接, 形成腺病毒纤维同源三聚体中单肽链的一个螺旋单位. 在 25 个拟重复基序中有 18 个基序的 b_4 位为甘氨酸残基, 以及大部分基序的长度均超过 15 aa 等构成了 EDSV_76 纤维蛋白拟重复基序的特征. 不同动物腺病毒纤维蛋白的拟重复基序的数量不等, 导致纤维突起的长度也有所不同. 目前已知包括人腺病毒在内的大多数哺乳类动物腺病毒纤



	b_4	b_3	A_1	A_2	A_3	a_1	a_2	a_3	a_4	a_5	B_1	B_2	B_3	b_1	b_2
1			G	L	Q	L	Y	I	T	A	Q	A	P	V	G	F
2	T	N	K	A	V	T	L	K	Y	G	D	G	L	E	V	N	E
3		N	G	E	L	I	A	T	A	S	S	A	V	K	P	P	L
4	H F	D	R	G	Y	I	V	L	N	L	Q	D	P	L	G	V	I	D
5			G	K	L	G	V	K	L	S	P	G	V	H	I	N	G
6		E	G	A	V	A	V	E	S	P	V	D	P	I	T	L	D
7	T	A	G	R	I	T	L	N	Y	G	T	G	L	N	V	S	D
8			G	K	L	R	I	V	S	P	E	S	P	L	T	L	L
9	G	N	G	K	V	A	L	N	F	G	N	S	M	E	L	V	Q
10			G	T	L	Q	L	K	A	P	L	N	P	L	F	M	T
11	P	A	G	A	I	G	L	R	V	D	D	M	F	N	I	S	E
12			G	L	L	S	F	K	M	P	S	D	P	I	S	F	N
13	A	D	G	M	L	S	L	N	T	N	D	T	L	Q	T	T	G
14			G	L	L	G	L	T	E	P	A	K	P	L	K	L	A
15		D	G	K	L	G	V	N	V	G	L	G	L	S	V	S	N
16			G	S	L	T	V	N	A	G	Q	G	L	T	I	R	N
17			N	A	V	A	V	N	G	G	N	T	L	A	F	N	N
18		Y	G	E	V	E	I	K	N	P	R	N	P	I	S	L	T
19	Q	D	G	E	L	V	L	I	I	G	D	G	L	T	T	L	D
20			G	R	L	T	L	L	T	A	S	T	S	P	I	A	V
21	G	P	A	G	V	T	F	N	V	T	P	S	D	F	Y	F	L
22		S	P	I	L	A	L	N	V	Q	T	R	G	G	L	E	K
23	S	D	T	G	L	K	I	K	R	A	A	P	L	S	I	T	S
24		D	G	E	L	T	L	A	Y	D	S	T	D	F	Q	V	T
25		E	N	G	L	A	L	K	V	S	P	T	Q	T	P	L	T

图 2　腺病毒拟重复基序示意图及 EDSV 纤维蛋白结构域 ㈦中 25 个拟重复基序的排列

维蛋白的拟重复基序一般在6~ 22个之间. 尽管EDSV_76纤维突起的长度仅为25 nm, 但是EDSV_76同羊腺病毒纤维蛋白的重复基序一样为25个. EDSV_76纤维蛋白结构域㈣所含aa的数量明显少于人腺病毒(平均为180 aa)、犬腺病毒(CAV, 178 aa)、牛腺病毒(BAV, 175 aa)和猪腺病毒(PAV, 584 aa), 但是与羊腺病毒(OAV, 120 aa)相近[5~ 10].

氨基酸序列同源性比较的结果显示: EDSV_76纤维蛋白与其他哺乳类动物腺病毒纤维蛋白的同源性均低于20%. 其中EDSV_76纤维蛋白结构域iv和㈦的N端序列与人腺病毒12型及CAV纤维蛋白的相应区域同源性高于35%, 而其余部分的同源性低于5%. 与此相反, 除去完全不相同的结构域iv部分, EDSV_76纤维蛋白结构域㈦和㈣与鼠及禽iv群腺病毒纤维蛋白氨基酸序列的同源性超过30%.

本文首次克隆并分析了EDSV_76纤维蛋白基因和编码产物的结构特点. 结果表明该病毒具有所有腺病毒纤维蛋白的特征结构, 但是与除OAV以外的其他哺乳类动物腺病毒纤维蛋白在核苷酸及氨基酸序列上差别较大, 显示该蛋白在进化途径、所呈现的抗原表型及与受体结合的特异性等方面明显不同. EDSV_76纤维蛋白的序列同源性、拟重复基序的结构和数量等与OAV纤维蛋白比较相近, 而且在与五邻体基座结合的结构域iv中同样具有PVYPFE样保守序列(在其他哺乳类腺病毒纤维蛋白中多为PVYPYD), 提示EDSV_76与OAV的亲缘关系较近, 有可能来源于共同的祖先. 有关EDSV_76与OAV之间的演变过程还有待于进一步探讨.

参 考 文 献

1 Van Eck J H H, Davclaur F G, Van den Hcuvel Plesman T A M, et al. Dropped egg production, soft shelled and shell_less eggs associated with appearance of precipitins to adenovirus in flocks of laying fowl. Avian Pathol, 1976, 5: 261~ 272
2 曾力宇, 金 奇, 章金钢, 等. 鸡减蛋综合征病毒六邻体蛋白基因的序列分析. 中国兽医学报, 1997, 17(3): 256~ 263
3 曾力宇, 金 奇, 朱 雷, 等. 鸡减蛋综合征病毒疏基内肽酶基因的序列分析. 病毒学报, 1997, 13(2): 134~ 137
4 Stouten P F W, Sander C. New triple_helical model for the shaft of the adenovirus fiber. J Mol Biol, 1992, 226: 1 073~ 1 084
5 Chroboczek J, Jacrot B. The sequence of adenovirus fiber: similarities and differences between serotypes 2 and 5. Virology, 1987, 161: 549~ 554
6 Kleibocker S B. Sequence analysis of the fiber genomic region of a porcine adenovirus predicts a novel protein. Virus Res, 1995, 39: 299~ 309
7 Dragulev B P, Sira S, Abouhaidar M G, et al. Sequence analysis of putative E3 and fiber genomic regions of two strains of canine adenovirus type 1. Virology, 1991, 183: 298~ 305
8 Raviprakash K S, Grunhaus A, EL Kholy M A, et al. The mouse adenovirus type 1 contains an unusual E3 region. J Virol, 1989, 63: 5 455~ 5 458
9 Vrti A, Boyle D, Kocherhans R, et al. Sequence of ovine adenovirus homologs for 100K hexon assernbly, 33 K, p㈡and fiber genes: early region E3 is not in the expected location. Virology, 1995, 209: 400~ 408
10 Chiocca S, Kurzbauer R, Schafliner G, et al. The complete DNA sequence and genomic organization of the avian adenovirus CELO. J Virol, 1996, 70: 2 939~ 2 949

(1997-09-01收稿, 1997-12-09收修改稿)