

猪流感病毒血凝素基因的克隆和序列测定*

石成华 G. W. Both

本文报道了以两种化学合成的寡脱氧核苷酸引物合成了猪流感病毒 NJ/11/76 (X53a) 株血凝素基因的全长拷贝。并将它克隆到大肠杆菌 RRI 细胞,进行了核苷酸全序列测定。同时将血凝素的氨基酸序列与同一亚型的另两个人甲型流感毒株的氨基酸序列进行了比较。

关键词 流感病毒;反转录;克隆;序列分析;抗原变异

流感病毒的基因组由 8 个单链 RNA 片段组成。它的总核苷酸数约 14,000^[1,2],每一基因片段的 3' 末端有 12 个相同的核苷酸序列,5' 末端有 13 个相同的核苷酸序列^[3,4]。由于这一结构上的特点, Winter 等^[5,6,7]以固相磷酸三酯法分别合成了 12 和 13 个核苷酸,以它们作为引物,合成了 A/PR/8/34 流感株编码非结构蛋白 NS₁ 和 NS₂ 基因的全长拷贝。本文应用这两个引物合成了编码血凝素蛋白 HA 基因的全长拷贝,并进行了克隆和序列测定。

材料和方 法

(一) 病毒 RNA

为 1976 年由病人分离的猪流感 NJ/11/76 的重组株 X53a 中提取的基因组 RNA,由 Dr. Kilbourne 赠送^[11]。

(二) 引物

12 寡脱氧核苷酸 d(AGCAAAAGCAGG), 13 寡脱氧核苷酸 d(AGTAGAACAAGG)。

(三) 双链 DNA 的制备

50 μ g 病毒 RNA (约 12 p mol 3' 末端) 加 250 p mol 的 12 寡脱氧核苷酸引物一起煮沸 1 分钟、放冰内冷却后,加入 80 单位 AMV 逆转录酶,于 200 μ l 反应体积中,含 50mM Tris·Cl pH 8.3、70mM KCl、10mM MgCl₂、10 mM DTT、0.5 mM dNTPs 和 14 p mol α -³²P-dATP。42 $^{\circ}$ C 下保温 1 小时。取出煮沸 1 分钟,冰浴冷却使合成的 ssDNA 和 RNA 分离,加入 RNase A 使终浓度为 0.125 μ g/ μ l,

42 $^{\circ}$ C 水浴消化 RNA 10 分钟。再用酚抽提、乙醇沉淀 ssDNA,并溶于适量 TE 缓冲液 (10mM Tris·Cl pH7.6、0.1mM EDTA)。将此溶液通过 Sephadex G-100 柱层析以去除未掺入的核苷酸,测各管的放射性,合并第一峰的各管,乙醇沉淀并溶于适量 TE 缓冲液。

总的 ssDNA 加入 2.5 p mol 的 13 寡脱氧核苷酸引物,一起煮沸 1 分钟,冰内冷却,在 100 μ l 反应体积中含 10 mM Tris·Cl pH 8.0、10 mM MgCl₂、75mM NaCl、1mM DTT、0.5mM dNTPs (如合成的 ssDNA 计数很高,在合成 dsDNA 时,可不加 α -³²P-dATP) 和大肠杆菌 DNA 聚合酶 I 的 Klenow 片段 2 单位。37 $^{\circ}$ C 水浴保温 1 小时,继而进行酚抽提、乙醇沉淀 dsDNA,并溶于适量 TE 缓冲液。将此溶液通过 Sephadex G-100 柱层析,同前测各管放射性,合并高峰管,经乙醇沉淀后,溶于适量 TE 缓冲液。

(四) 双链 DNA 的接尾、克隆及转化子筛选

经 dC 接尾后的 DNA 在 3% 聚丙烯酰胺凝胶上电泳,由胶上切下第 4 条血凝素基因带,与经 Pst I 酶切、dG 接尾的 pBR322 载体进行重组,并转化至大肠杆菌 RRI 受体中。具体方法参阅 Sleight 等 1979 年文献 [8]。

(五) 血凝素基因的序列测定

采用下述三种方法:(1) 将血凝素基因的 *Sma* 3A 酶切的 5 个片段次级克隆于 M13mp7.1 噬菌体中,以 Sanger 链终止法进行序列测定^[6,10];(2)

本文于 1984 年 5 月 17 日收到。

* 本工作在澳大利亚联邦科学院分子生物实验室完成,由 Both 博士指导。

Maxam-Gilbert 化学降解法^[11]; (3)用病毒基因组 RNA 为模板,以克隆的血凝素基因的酶切片段为引物,在逆转录酶和 dNTP 和 ddNTP 的参与下,分析 cDNA 的序列^[12]。

结果和讨论

(一) 猪流感血凝素基因 ssDNA 和 dsDNA 的合成

以 12 和 13 寡脱氧核苷酸合成了猪流感 NJ/11/76(X53a) 株 8 个基因全长的 ssDNA 和 dsDNA(图 1 和图 2)。在合成反应中,出现一些短的 ssDNA(由于上样量少,图 1 中不明显)和 dsDNA 拷贝,可能是 12 寡聚核苷酸引物在病毒基因组上发生非特异性的结合,从而产生带有相同 3' 末端短的转录物。这些大小不等的转录物在 13 寡聚核苷酸的作用下产生带平末端短的 dsDNA 拷贝。另一可能是 ssDNA 在 3' 末端形成返折环的自身引物,产生了在 DNA 一端带有一个环的短的 dsDNA 拷贝。

以化学合成的寡核苷酸作为引物合成

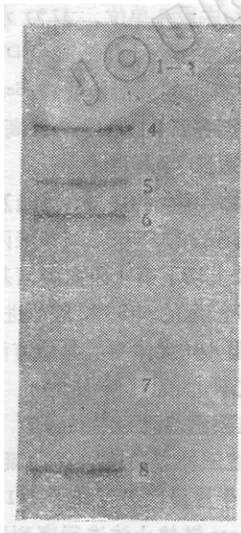


图 1 由病毒 RNA 合成的 ssDNA 聚丙烯酰胺凝胶电泳

Fig. 1 Polyacrylamide gel electrophoresis of ssDNA synthesized from virus RNA

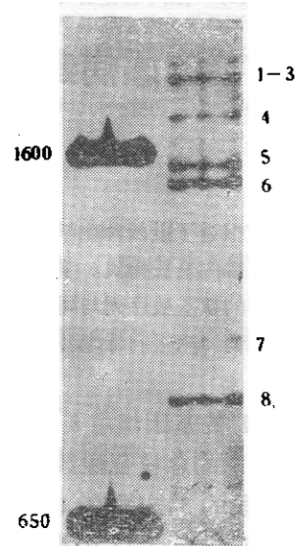


图 2 由病毒 ssDNA 合成的 dsDNA 聚丙烯酰胺凝胶电泳,左侧是碱基大小标准(bp)

Fig. 2 Polyacrylamide gel electrophoresis of dsDNA synthesized from ssDNA. Size standards were written in bp on the left

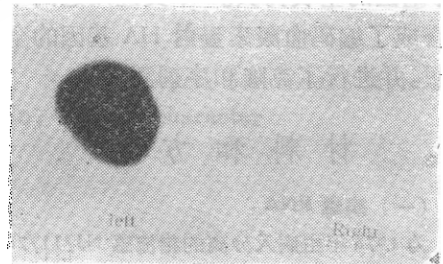


图 3 以 ^{32}P 标记的单链 cDNA (来自病毒 RNA 的拷贝) 探针与重组 DNA 的点杂交(左侧),右侧为质粒 DNA 对照,无杂交点。

Fig. 3 Dot hybridization of recombinant DNA with ^{32}P labelled cDNA probe copied from virus RNA (Left). Plasmid DNA control (Right).

dsDNA 的优点是不需要以 S1 酶切去返折环的步骤,所产生的不完整的 dsDNA 较少。

(二) 猪流感病毒血凝集基因的克隆和酶切图谱

以末端转移酶分别在 dsDNA 的 3' 末端和 Pst I 酶切后的 pBR322 DNA 的 3' 末

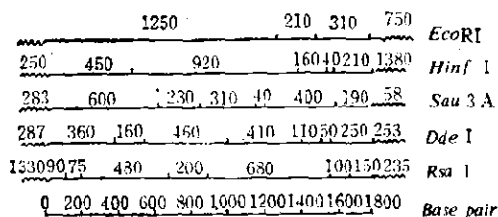


图4 NJ/11/76(X53a)株血凝素基因酶切图谱。波浪线条表示与血凝素基因末端相连的质粒DNA

Fig. 4 Preliminary restriction enzyme cleavage map of HA gene from strain NJ/11/76(X53a). Wavy lines are plasmid DNA attached to the ends of the inserts

端加上约 30 个 dC 和 dG 尾。从聚丙烯酰胺凝胶上切下血凝素基因带, 即第 4 条带与 pBR322 DNA 进行体外重组, 转化至大肠杆菌 RRI 受体。从四环素平皿上选出的重组子经琼脂糖电泳鉴定含有大小近似血凝素的插入物。以 ^{32}P 标记的单链 cDNA (来自 NJ/11/76 X53a 株 RNA 的 cDNA 拷贝) 作探针能与重组子 DNA 发生杂交 (图 3)。以各种核酸内切酶对重组 DNA 进行单酶和双酶消化, 初步确定重组 DNA 内含有血凝素基因的全长拷贝, 约 1770bp,

	1	10	20	30	40	50
0	UCGUUUUCGU	CCCCUUUUUAU	UUUCGGUGGC	UUUACUUCGG	UUUUGAUAAU	
50	CAGAACAUAU	CAUGUAAACG	UCGGUGUUUA	CGUCUGUGUG	AUACAUAUCC	
100	AAUAGUACGU	UUAAUUAAGUU	GACUGUGACA	ACUAUGUCAU	GAUCUUUUCU	
150	UACAUUGUCA	UUGUGUGAGA	CAAUUAGAAG	AACUUCUGUC	UGUAUUGCCC	
200	UUUGAUACAU	UUGACCCCCC	CUAUCGGGGU	AACGUGAAC	CAUUUACAUU	
250	GUAACGGCCA	UCCGAAAACC	CUUUGGGUCU	GAUUAUAAU	GAUGAGUGUC	
300	AUUCGAGUAC	CAGAAUGUAA	CACCUUUUGA	GCUUUAGUCU	GUUACCCUGU	
350	ACAAUGGGUC	CUCUAAAGUA	GUUAAUACUU	CUCGACUCUC	UCGUCAACUC	
400	GAGUCACAGU	AGUAAACUUU	CUAAGCUCUA	UAAGGGGUUC	UGUUCACGCA	
450	CCGGGUUAGU	ACUUGUUUG	UCUCCACACU	GCCGUCGUAC	GGGAUACGA	
500	CCUCGUUUGU	CGAAGAUUC	UUUAAAUAU	ACCGACCACU	UUUUUCUUUU	
550	AAGUAVGGGU	UUCGAGUCGU	UUAGGAVACA	AUUGUUAUUU	CCCUUCCUUC	
600	AGGAACACGA	UAGCCCGUAA	GUAGUAGGUG	GAUGGUCAUG	ACUAGUUGU	
650	UCAGAGAGUG	UUUACGUCU	ACGGAUACAA	AAACAUCCCA	GUAGUUUUUA	
700	AUUGUCUUUC	AAGUUCGGUC	UUUAUCGUCG	UUCUGGGUUC	CACUCUCCAG	
750	UUCGUCCCUC	CUACUCGAUA	AUGACCUUGG	AUUAUCUUGG	ACCUCUGUGU	
800	UAUUGUAAAC	UUCGUUGACC	UUUAGAUCAC	CAUGGUUCUA	UACGAAGCG	
850	UUACUUUAUCU	CCAAAGACUA	GACCAUUUAU	UUAAACCCUA	CGGGUACAAG	
900	UACUAAACAU	GUGCUUCACA	GUUUGUGGGU	UCCCACGGUA	UUUGUGGUCG	
950	CAGGGUAAAG	UCUUAUAUGU	AGGUCGGUGU	UAACCUUUAU	CAGGUUUUAU	
1000	ACAGUUUUCG	UGUUUUUAACU	CCUACCGAUG	UCCUGAUUCC	UUACAGGGCA	
1050	GAUAAAGUAG	AUCUCCAGAC	AAACCCUGGU	AACGACCAAA	AUAACUCCCC	
1100	CCUACCUUGUC	CUUACUAUCU	ACCUACCAUA	CCAUAUGUGG	UAGUCUUACU	
1150	UGUCCCUAGU	CCUAUACGUC	GCCUGGUUUC	CUCGUGUGUU	UUACGGUJAC	
1200	UGCCCUAGUG	AUUGUUUCAU	UUUAGACAAU	AACUUUUUCA	CUUGUGCGUU	
1250	AAGUGCCGUC	ACCCAUUUUCU	UAAGUUGGUG	AACCUUUUUU	CUUAUCUCUU	
1300	AAAUUUUGUU	UUCCAACUAC	UACCAAAAGA	CCUAUAAACC	UGUAUUGUAC	
1350	GUCUUAGCAA	CCAAGAUAAAC	CUUUUACUUU	CUUGAGACCU	AAAGGUGCGU	
1400	AGCUUACAUU	UCUUGGACAU	ACUCUUUCAU	UCUUCGGUUC	AUCCUUGUUG	
1450	ACGGUUUCUU	UAACCUUUAC	CGACGAACAU	UAAGAUAGUG	UUUACACUAC	
1500	UAUGUACGUA	CCUCUCGCAG	UUUUUACCCU	GAAUACUAAU	GGGUUUCAUG	
1550	AGUCUUCUUA	GUUUCGAUUU	GUCUCUCCUC	UAUCUACCUC	AUUUCAACCU	
1600	CAGUUGUUCU	UAAAUUGUCU	AAAACCGCUA	GAUAAGUUGA	CACCGGUCAA	
1650	GUAACCAUGA	CAACCGAGAG	GACCCCGGUU	AGUCAAAAAC	CUACACGAGA	
1700	UUACCCAGAA	AUGUCACGUG	UUUAACAUAU	AUUUUGAUCC	UAAAGUCUCU	
1750	GUACUCUUUU	UUGUGGGAAC	AAAGAU			

图5 猪流感病毒 A/NJ/11/76 (X53a) HA 基因的核苷酸序列(病毒 RNA 序列)

Fig. 5 Nucleotide sequence of the HA gene of swine influenza virus strain A/NJ/11/76 (X53a) (vRNA sequence)

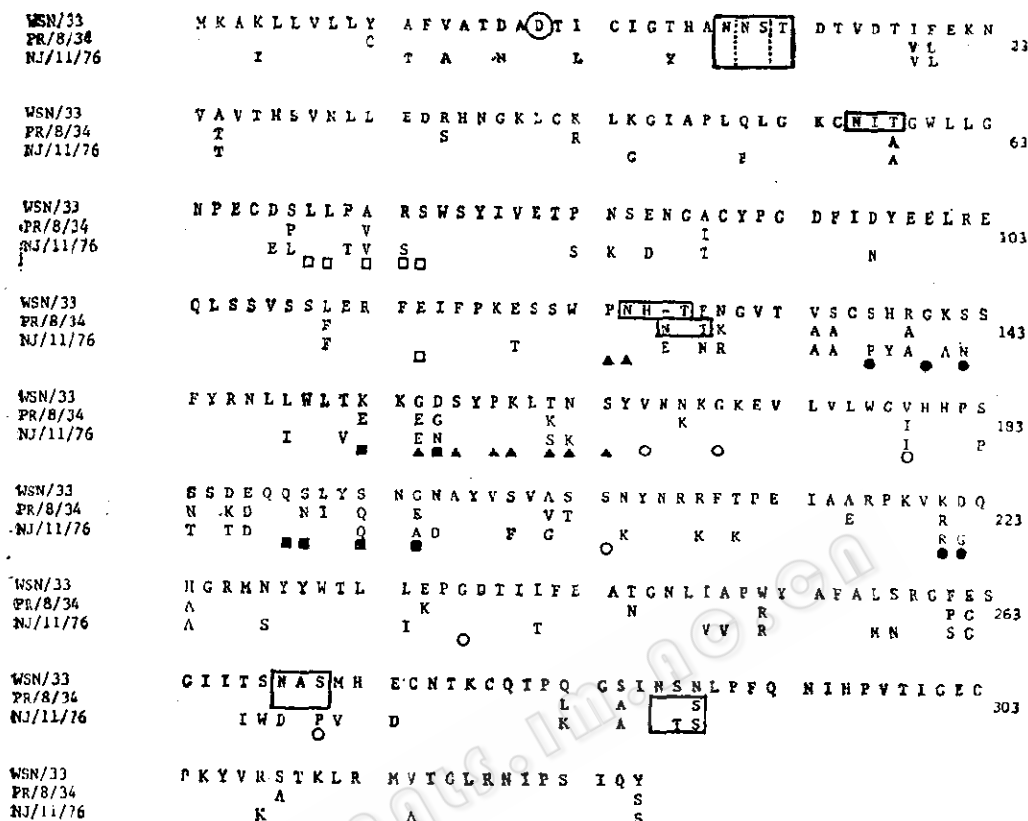


图6 猪流感和两株人流感血凝素 HA1 的氨基酸序列的比较

大□表示碳水化合物附着位点 大○成熟血凝素N末端氨基酸

以A/PR/8/34 (Mt. Sinai)定名的抗原区: ▲Sa ■Sb □Cb ●Ca2 ○Ca1

Fig. 6 The comparison of amino acid sequences of HA1 between two human strains and a swine strain. □ indicate potential sites for carbohydrate attachment.

Antigenic regions defined for A/PR/8/34 (Mt. Sinai) are indicated:

▲Sa ■Sb □Cb ●Ca2 ○Ca1

并建立了酶切图谱(图4)。

(三) 猪流感病毒血凝素基因的全序列测定

图5为血凝素基因的全部核苷酸序列,共由1776个核苷酸组成。3'末端的12个核苷酸和其他7个基因片段的3'末端序列是一致的,即UCGUUUUCGUCC。整个序列编码17个氨基酸的信号肽,326个氨基酸的血凝素HA1和222个氨基酸的血凝素HA2。HA1和HA2之间由一个精氨酸相连。

(四) 猪流感血凝素与两株人流感血凝素氨基酸序列的比较

由于血凝素HA2结构的保守性和流感病毒抗原位点集中于HA1区,故仅将HA1区氨基酸序列进行比较(图6)。可见猪流感A/NJ/11/76和两株人流感A/PR/8/34、A/WSN/33的氨基酸序列具有高度的同源性,HA1约为80%,HA2约为94%,它们同属于H1N1亚型。同时可见猪流感血凝素氨基酸的改变比较分散,但基本上位于或靠近Caton 1982年⁽¹⁰⁾提出的H1N1

亚型的已知抗原位点区,特别在 Ca2、Cb 和 Sb 位点区的氨基酸改变较为明显。与 A/WSN/33 相比,猪流感病毒血凝素失去三个碳水化合物附着位点,分别在第 58、127 和 269 氨基酸处,得到一个碳水化合物附着位点,在第 287 氨基酸处。一般认为碳水化合物附着位点(其氨基酸序列为 Asn-X-Ser 或 Asn-X-Thr)能掩盖抗原位点,从而妨碍抗体识别抗原,起调节抗原性的作用^[13,14]。猪流感血凝素缺乏能掩盖抗原位点的碳水化合物附着位点。从比较三株的 127—142 和 267—272 区的氨基酸(即碳水化合物附着位点区),可发现猪流感血凝素氨基酸序列的漂移。267—272 区在三向立体结构上相当于流感病毒 H3N2 亚型的 269—274 氨基酸区,即 C 抗原位点。

参 考 文 献

[1] Desselberger, U. et al.: *Virology*, 88: 394—

399, 1978.

- [2] Sleight, M. J. et al.: *Nucleic. Acid. Res.*, 6: 1309—1321, 1979.
- [3] Robertson, J. S.: *Nucleic. Acid. Res.*, 6: 3745—3757, 1979.
- [4] Skehel, J. J. et al.: *Nucleic. Acid. Res.*, 5: 1207—1218, 1978.
- [5] Winter, G. et al.: *ibid.*, 9: 237—245, 1981.
- [6] Gait, M. J. et al.: *ibid.*, 8: 1081—1096, 1980.
- [7] Gait, M. J. et al.: *Nucleic. Acid. Res. Symposium Series.*, No. 7: 243—258, 1980.
- [8] Sleight, M. J. et al.: *Nucleic. Acid. Res.*, 7: 879—893, 1979.
- [9] Messing, J. et al.: *ibid.*, 9: 309—321, 1981.
- [10] Sanger, F. et al.: *PNAS.*, 74: 5463—5467, 1977.
- [11] Maxam, A. and W. Gilbert: *Ibid.*, 74: 560—564, 1977.
- [12] Both, G. W. et al.: *J. Virol.*, 39: 663—672, 1981.
- [13] Wiley, D. C. et al.: *Nature*, 289: 373—378, 1981.
- [14] Caton, A. J. et al.: *Cell*, 31: 417—427, 1982.
- [15] Kilbourne, E. D.: *PNAS.*, 75: 6258—6262, 1978.

CLONING AND SEQUENCING OF THE HEMAGGLUTININ GENE OF SWINE INFLUENZA VIRUS*

Shi Chenghua G. W. Both

Complete double-stranded DNA copies of the RNAs of swine influenza virus NJ/11/76(X53a) have been synthesized by using two synthetic oligodeoxynucleotide primers. The gene encoding the hemagglutinin, prepared with these primers, has been cloned and sequenced. A comparison is made on the predicted amino acid sequence of swine influenza hemagglutinin with those of two other human strains and

the results are discussed.

Key words

Influenza virus RNA; Reverse transcription; Cloning; Nucleotide sequence; Antigenic variation

* This article was completed in CSIRO Molecular Biology Laboratory, Australia.