

IBV 青岛腺胃分离株(SD/97/02)S1 蛋白基因的序列测定和分析*

潘杰彦,陈德胜,戴亚斌,陈溥言**

(南京农业大学农业部动物疫病诊断与免疫重点开放实验室,江苏南京 210095)

Sequencing and Sequence Analysis of S1 Spike Protein Gene from Avian Infectious Bronchitis Virus Proventriculus Strain Isolated in Qingdao(SD/97/02)

PAN Jie-yan, CHEN De-sheng, DAI Ya-bin, CHEN Pu-yan**

(Key Laboratory of Animal Diseases Diagnostic and Immunology, Ministry of Agriculture, Nanjing Agriculture University, Nanjing 210095, China)

Abstract: A pair of primers were designed and synthesized according to the reported S1 spike protein gene sequence of IBV. The viral RNA of SD/97/02 strain was amplified by reverse transcription-polymerase chain reaction(RT-PCR). The amplified product was analyzed by agarose gel electrophoresis, all appeared a fragment of 1657 bp as expected. The RT-PCR product was cloned into pMD18-T vector, and then the recombinant plasmid was sequenced. The result suggested that it was the S1 spike gene of IBV. The recombinant plasmid was designated pMDQXS1. The sequence has been published in the Genbank, and the accession number is AF193423. Sequence analysis showed that the S1 gene has low G + C content at about 37.0%. There existed several RE cleavage sites such as *Hind* III, *Bam* HI, *Bgl* III, *Sac* I and *Sal* I, but have no *Eco* R I site, which is different from that of other strains. Its homology compared with other IBV strains was between 87.02% and 91.21%. In the site between 154 and 429 of this S1 gene is a highly variation region. The S1 protein is composed of 540 amino acid residues after the gene was translated into amino acids, the pI is 8.24. In the inner of the protein there have 18 Cys residues. The sequence of the cleavage site between S1 and S2 subunits is HRRRR, which is different from that other IBV strains(RRF/SRR). Three high conservative regions were located at 169-181, 230-250, 485-506 amino acids residues. Comparative analysis of the antigenicity of SD/97/02 with that of other strains showed epitope disappearance in 320-326aa and 390-401aa, two strong epitopes appeared in 325-345aa and 379-389aa; in 438-444aa, the epitope which existed in other strains of IBV(except ZJ971 strain) disappeared; and the antigenicity between the amino acids residues 53-65 became weak.

Key words: Infectious bronchitis virus(IBV); Proventriculus strain; S1 spike protein gene; Sequence analysis

摘要:参考 Genbank 上发表的 IBV S1 纤突蛋白基因序列,设计了一对引物,对鸡传染性支气管炎病毒青岛腺胃分离株(SD/97/02)RNA 进行 RT-PCR 扩增。将 PCR 产物克隆入 pMD18-T 载体中进行序列测定和分析。序列分析表明,

收稿日期:2001-03-08,修回日期:2001-06-13

* 基金项目:国家自然科学基金重大项目(39893290 2-2)

作者简介:潘杰彦(1973-),男,浙江省宁波籍,博士,研究方向为动物分子病毒学,E-mail:jieyanpan@yahoo.com,现在地址:中国科学院上海细胞生物学研究所,上海,200031

** 通讯作者:陈溥言(1942-),男,辽宁省籍,博导,教授,研究方向为动物传染病学及病毒基因工程。Correspondence author.

该毒株的 S1 基因的 G+C% 含量较少,为 37.0%,存在 *Hind* III, *Bam* HI, *Bgl* I, *Sac* I 和 *Sal* I 位点,无 *Eco* R I 位点,与其他毒株的同源性在 87.02%~94.21% 之间,在第 154~429nt 处为高度的变异区;将基因序列翻译成氨基酸后,假定的 S1 蛋白由 540 个氨基酸组成,等电点 8.24,在蛋白质内部存在 18 个 Cys,在 S1 与 S2 蛋白之间的剪切位点为 HRRRR,这与大多数 IBV 毒株(RRF/SRR)不一样,有三个区域的氨基酸序列高度保守;169~181aa,230~250aa,485~506aa;与其他毒株进行抗原性比较后发现,在该毒株的 320~326aa 及 390~401aa 处的抗原表位消失,而在 325~345aa,379~389aa 处则出现了很强的抗原表位;第 438~444aa 处,其他 IBV 毒株(除 ZJ971 株外)原来存在的强抗原位点在本毒株中消失;在 53~65 位的氨基酸抗原性与其他毒株相比明显变弱。

关键词:IBV;腺胃株;S1 纤突蛋白基因;序列分析

中图分类号:S852.65⁺9.6 文献标识码:A 文章编号:1003-5125(2001)04-0377-05

鸡传染性支气管炎病毒(IBV)属于套式病毒目(又叫尼多病毒目(Nidovirales))^[1],为冠状病毒科,冠状病毒属的代表成员。该病毒可引起鸡的严重的传染病,此病以气管啰音、咳嗽、打喷嚏为特征^[2]。IBV 由于具有独特的先导-引物转录机制,虽然能使如此大的基因组(约 27.6kb)保持其种属的稳定性,但也由此造就了 IBV 的另一特性-多变异性,使 IBV 具有多种血清型,并且也产生了很多的组织嗜型,如最经典的呼吸型,稍后出现的肾型,以及最近出现的产蛋下降型和肠型。而中国从 1995 年以来,出现了一种以腺胃病变为典型特征的鸡急性传染病^[3],一些人认为它是由变异株的 IBV 引起的一种新的传染病,但也有人有不同的观点,而在国外尚未见报道。

IBV 的结构蛋白主要为三种:纤突糖蛋白(S)、基质蛋白(又叫膜蛋白,M)、核衣壳蛋白(N)。纤突蛋白在病毒成熟后裂解成两个亚单位:S1 和 S2,其中 S1 亚单位位于病毒粒子的最外端,比较容易脱落,它能介导病毒与细胞受体的结合并进入细胞;另外,S1 蛋白还介导细胞融合,诱发产生中和抗体和血凝抑制抗体。正是由于 S1 蛋白具有如此重要的功能,现在对 S1 蛋白及其基因的研究是最多的,而且比较深入。从以往的研究来看,S1 蛋白基因的变异还可能导致组织嗜性的改变本试验中所用的病毒分离自青岛当地腺胃严重病变鸡的腺胃组织,已确认为冠状病毒^[4]。本试验拟从分子生物学的角度出发,直接测定该病毒的 S1 基因序列,试图从分子水平来研究该病毒的变异机制。

1 材料和方法

1.1 引物设计和合成

利用 NCBI 网站 Genbank 上的 BLAST 软件对大量的 IBV S1 基因序列进行了同源性比较,选取较为保守的部位作为引物,其中 5'引物(S11)位于 S1 基因起始密码子以前的约 40nt 处,该处的序列经序列

同源性比较为保守的区域,且包含了前导序列的共有部分:CTGAACAA,这样可以使该引物的适用性更广;3'端引物则位于 S2 亚单位基因内,同样比较保守。引物序列如下:

S11 5'-GCAAACCTGAACAAAAGAC(CA)GACT-TAG-3'

S12 5'-CTT(AG)CCATAACTAACATAAGGGCA-3'

(注:括号内(CA)的意思为此个碱基为 C 或 A,同样(AG)为 A 或 G,在引物中以相同的比例出现,在此引入简并碱基的目的是更好的扩增 RNA,使引物的引发功能更为强大。此对引物中未加酶切点,可通过 T 载体来进行克隆。)

1.2 SD/97/02 株 IBV S1 基因的获取

具体方法见文献^[5],现将实验步骤简要描述如下:病毒→接种 10 日龄鸡胚→40h 后收毒→冻融,离心,纯化→提取 RNA→反转录→PCR→电泳检查;其中反转录的条件为 42℃ 1h,37℃ 50min(用 M-MLV RT),PCR 的条件为,94℃ 预变性 5min,然后 94℃ 1min,47℃ 2min,72℃ 4min,进行 40 个循环,最后 72℃ 延伸 10min。

1.3 目的基因的克隆

同样按文献^[5],将 PCR 产物克隆入 pMD18-T 载体(大连宝生物(TaKaRa)工程公司产品)

1.4 测序

在大连宝生物(TaKaRa)工程公司的协助下,利用 ABI PRISM™377 型自动测序仪,通过通用引物 M13-47 和 RV-M 进行双向测序,然后在序列的中间再合成一个引物将此序列测通,通过自动序列拼接软件进行拼接后即所需的 S1 基因序列。

1.5 序列分析

首先将所得的序列和 NCBI 网站上 Genbank 中的序列进行同源性比较,然后用 DNASTAR 等分子生物学软件进行分析。

2 结果

2.1 引物的设计和目的基因的获取

经过 Goldkey 软件的辅助分析,引物 S11 和 S12 内部均没有连续 4 个以上碱基构成反向互补序列,二者的 3' 末端也无同源性,不会形成引物二聚体,且错误引发的能力也较弱,唯一的缺点就是这对引物的 G+C% 均较低,仅在 35%~40% 之间,这是不符合常规的引物设计原则的,不过整个 S1 基因的全

体 G+C% 也仅在 36% 左右,后来也表明 PCR 过程中非特异扩增较多,这也是 S1 基因较难扩增的原因之一。

2.2 S1 全基因序列的测定

SD/97/02 株 S1 基因的全序列和相应推导出的氨基酸序列见图 1(该序列已提交 Genbank 注册,注册号为 AF193423)。

2.3 序列分析

```

ATG TTG GGG AAG TCA CTG TTT TTA GTG ACC ATT TTG TGT GCA CTA TGT AGT GCA AAT TTG TTC GAT TCT GCT AAT AAT TAT GTG TAC TAC CAA AGT GGC TTT AAG CC 110
M L G K S L F L V T I L C A L C S A N L F D S A N N Y V Y Y Y Q S A F R P
TCCA AAT GGA TOG CAT TTG CAA GGG GGT GCT TAT GCA GTA GTG AAT TCC ACT AAT TAT AGT AAT AAT GCA GGT TCT GCA CCT CAG TCC ACT GTT GGT GTT ATT AAG GAC G 220
P N G W H L Q G G A Y A V V N S T N Y S N N A G S A P Q C T V G V I K D
TC TAT AAT CAA AGT GCG GCT TCT ATA GCT ATG ACA GCA CCT CTT CAG GGT ATG GCT TGG TCT AAG TCA CAA TTT TGT AGT GCA CAC TGT AAC TTT TCT GAA ATT ACA GTT 330
Y Y N Q S A S I A M T A P L Q G M A W S K S Q F C S A H C N F S E I T V
TTT GTC ACA CAT TGT TAT AGT AGT GGT AGCGGG TCT TGT CCT ATA ACA GGC ATG ATT CCA CGT GAT CAT ATT CGT ATT TCT GCA ATG AAA AAT GGT TCT TTA TTT TAT AA 440
F V T H D Y S S G S G S C P L T G M I P R D H I R I S A M K N G S L F Y
T TTA ACA GTT AGC GTA TCT AAA TAC CCT AAT TTT AAA TCT TTT CAA TGT GTT AAC AAC TTC ACA TCT GTT TAT TTA AAT GGT GAT CTT GTT TTT ACT TCCAACAAA ACT A 550
N L T V S V S K Y P N F K S F Q C Y N N F T S V Y L N G D L V P T S N K T
CT GAT GTT ACG TCA GCA GGT GTG TAT TTT AAA GCA GGT GGA CCT GTA AAT TAT AAT ATT ATG AAA GAA TTT AAG GTT CTT GCT TAC TTT GTT AAT GGT ACA GCA C AA GAT 660
T D V T S A G V Y F K A G G P V N Y N I M K E F K V L A Y F V N G T A Q D
GTA ATT TTG TGC GAT AAT TCC CCC AAG GGT TTG CTA GCC TGT CAA TAT AAC ACT GGC AAT TTT TCA GAT GGC TTT TAT CCT TTT ACT AAT AGT ACT TTA GTT AAG GAA AA 770
V I L C D N S P K G L L A C Q Y N T G N F S D G F Y P F T N S T L V R E
G TTC ATT GTC TAT GGC GAA AGT AGT GTT AAT ACT ACT CTGGGG TTA ACT AAT TTC ACT TTT ACT AAT GTA AGT AAT GCA CAG CCT AAT AGT GGT GGT GTT AAT ACT TTT C 880
K F I V Y R E S S V N T T L A L T N F T F T N V S N A Q P N S G G V N T F
AT TTA TAT CAA ACA CAA ACA GCT CAG AGT GGT TAT TAT AAT TTT AAT TTG TCA TTT CTG AGT CAG TTT GTG TAT AAG GCA AGT GAT TTT ATG TAT GGG TCT TAC CAC PCT 990
H L Y Q T Q T A Q S G Y Y N F N L S P L S Q F V Y K A S D F M Y G S Y H P
AGT TLT TCT TTT AGA CCA GAA ACC ATT AAT AGT GGT TTG TGG TTT AAT TCC TTG TCA GTT TCT CTT ACT TAT GGA CCC CTA CAG GGA GGG TGT AAG CAA TCT GTT TTT AG 1100
S C S F R P E T I V S G L W F N S L S V S L T Y G P L Q G G C K Q S V F
T GGT AAG GCA ACG TGT TGT TAT GCC TAC TCT TAT AAT GGC CCA AGC GCA TGT AAA GGT GTT TAT TCA GGT GAA TTA AOC ATG AAT TTT GAA TGT GGA TTG CTG GTT TAT G 1210
S G K A T C C Y A Y S Y N G P R A C K G V Y S G E L S M N F E C G L L V Y
TT ACT AAG AGT CAT GGC TCT CGT ATA CAG ACT AGA ACG GAG CCC TTA GTA TTA ACG CAA CAC AAT TAT AAT AAT ATT ACT TTA GAT AAG TGT GTT GCT TAT AAT ATA TAT 1320
V T K S H G S R I Q T R T E P L V L T Q H N Y N N I T L D K C V A Y N I Y
GOC AGA GTA GOC CAA GGT TTT ATT ACT AAT GTG ACT GAT TCT GCT OCT AAT TTT AGT TAT TTA GCA GAT GGT GGG TTA GCT ATT TTA GAT ACG TCG GGT GOC ATA GAT GT 1430
G R Y G Q G F I T N Y T D S A A N F S Y L A D G G L A I L D T S G A I D
T TTT GTT TGA AAG GGC AGC TAT GGT CTT AAT TAT TAC AAG GTT AAT OCT TGT GAA GAT GTT AAC CAA CAG TTT GTA GTG TCT GGT GGC AAT ATA GTT GGC AT TCT ACT T 1540
V F V V K G S Y G L N Y Y K V N P C C D V N Q Q F V V S G G N I V G I L T
CT AGA AAT GAA ACA GGT TCT GAA CAG GTT GAG AAC CAG TTT TAT GTT AAG TTA ACC AAT AOC TCA CAT CGT GOC AGG CGT 1620
S R N E T G S E Q V E N Q F Y V K L T N S S H R R R R

```

图 1 SD/97/02 株 IBV S1 基因序列及其所推导的氨基酸序列

Fig. 1 Sequence of IBV SD/97/02 strain S1 gene and its deduced amino acids sequence

2.3.1 SD/97/02 株 IBVS1 基因的特征及其与其他 IBV 毒株的同源性比较 在 S1 基因中,各种核苷酸的比例为:A-27.6% (457/1657), G-20.3% (337/1657), C-15.7% (260/1657), T-36.4% (603-1657),这一比率和其他毒株的 IBV 几乎一致。此基因的特征就是它的 G+C% 含量较少。

对该毒株的 S1 基因的酶切位点进行分析表明存在 *Hind* III, *Bam* HI, *Bgl* I, *Sac* I 和 *Sal* I 位点,但无 *Eco*RI 位点;该毒株的 S1 基因与 Genbank 上的其他 IBV 毒株 S1 基因的同源性在 87.02%~94.21% 之间,在第 154~429 为高度的变异区,该毒株与大多

数毒株的差异即在此。

2.3.2 SD/97/02 株 IBV S1 蛋白一些特征 用 DNASTAR 中的强大功能,分析了以下特征,见图 2。在亲水性上,我们发现前 18 个氨基酸具有强烈的疏水性,这儿应该是信号肽区域,这与其他 IBV 毒株完全一致。

S1 蛋白由 540 个氨基酸组成,等电点为 8.24,在蛋白质内部存在 18 个 Cys,在 S1 与 S2 蛋白之间的剪切位点为 HRRRR(剪切位点位于最后一个 R 后);进行多序列氨基酸同源性比较后发现,有三个区域的氨基酸序列高度保守:169~181aa, 230~

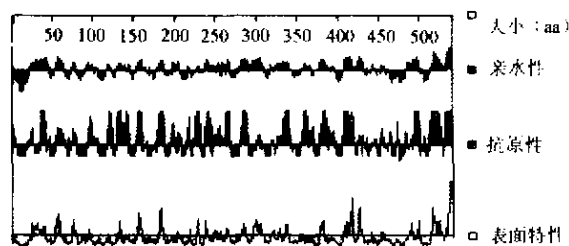


图2 SD/97/02株IBV S1蛋白的一些特性

Fig.2 Some characteristics of IBV(SD/97/02 strain)S1 protein.

250aa, 485 ~ 506aa。

2.3.3 SD/97/02株IBV与其他IBV的进化树分析

(参见图3,可见SD/97/02株与其他毒株相比较为独立,在基因分型上可将其独立地分为一个型。此图列举了几种在组织嗜型上各不相同的毒株,其中IBV D41为我国分离的肾型毒株,IBV H120为来自荷兰的标准呼吸型毒株,M41株为美国麻省标准呼吸型毒株,IBV Beau为美国另一标准呼吸型毒株,

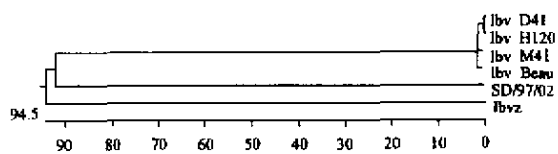


图3 IBV的S1基因进化树分析

Fig.3 The phylogenetic tree of IBV S1 gene

IBVZ为分离自浙江杭州的腺胃表现型毒株^[12]。

2.3.4 SD/97/02株S1蛋白与其他毒株进行抗原性分析 参见图4,对SD/97/02株的S1蛋白的比较分析表明,在320~326aa及390~401aa处的抗原表位消失,而在325~345aa,379~389aa处则出现了很强的抗原表位;第438~444位,其他毒株(除IBV ZJ971毒株,为另一个腺胃表现型的IBV毒株)都有一个强的抗原表位,而在此毒株中不存在;在53~65位的氨基酸抗原性与其他毒株相比明显变弱。

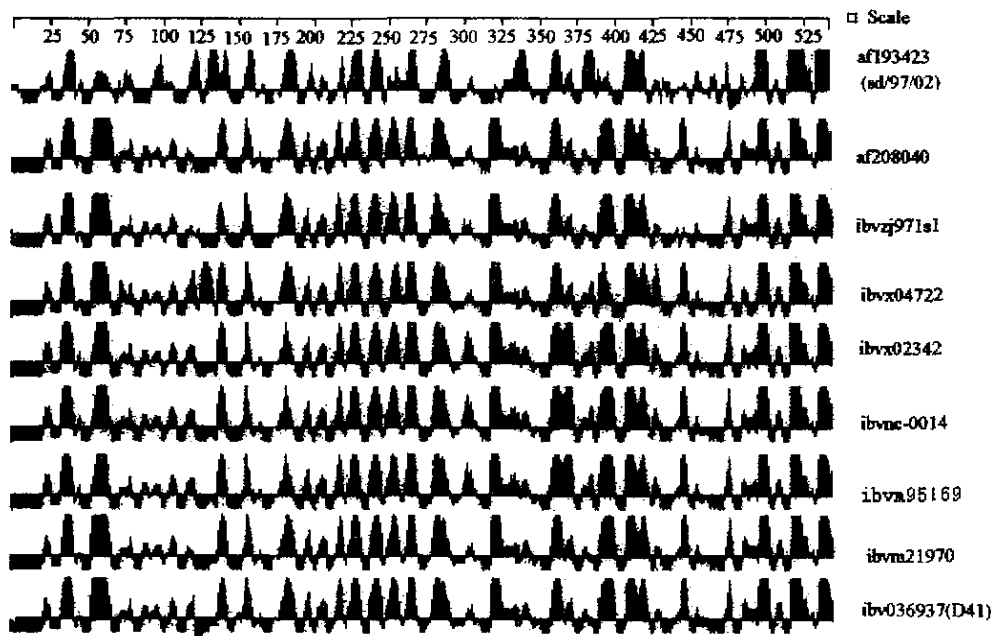


图4 IBV SD/97/02株S1蛋白的抗原性比较分析

Fig.4 Comparative Analysis of the Antigenicity of IBV SD/97/02 strain

3 讨论

IBV的S1蛋白一直被认为能激发中和抗体和血凝抑制抗体^[6,7],并与病毒的组织嗜性有关,S蛋白与宿主细胞膜上糖蛋白受体的结合可能是IBV及吸附细胞的前提条件^[8],S1基因的变异,将直接影

响到病毒与细胞膜受体的结合。挑选有代表性的毒株(指不同组织嗜性的病毒株),克隆其S1基因,并进行测序及序列分析,或许会从中找出一些规律,从而为阐明IBV的致病机理奠定坚实的基础。

我们分析了SD/97/02毒株S1基因的限制酶切位点,发现存在Hind III, BamH I, Bgl II, Sac I和Sal I

位点,但无 *EcoR* I 位点,这与已报道的不致^[9]。将基因序列翻译成氨基酸,发现 S1 蛋白由 540 个氨基酸组成,等电点为 8.24,在蛋白质内部存在 18 个 Cys (在 IBVS1 蛋白中,Cys 在 16~19 个之间),而 Cys 能形成二硫键,因此在 Cys 上的变异会导致蛋白功能较大的变化。在 S1 与 S2 蛋白之间的剪切位点 HRRRR,这与大多数 IBV 毒株(RRF/SRR)不一样。在禽流感病毒中的血凝素切割位点为 R-X-R/K-R,这与 IBV 的 S 蛋白切割位点相类似。在 NDV 的 F 蛋白切割位点的变异可导致毒力的变化^[10],或许这也可以推广到 IBV 的 S 蛋白,但这还需要实验的证实。

我们将 SD/97/02 株的氨基酸序列与 Genbank 中的序列进行了比较发现,有三个区域的氨基酸序列高度保守:169~181aa,230~250aa,485~506aa。Wang 等^[11]报道在 S1 蛋白的第 240~255 位为高度保守区。我们的结果与此比较接近。在第 24 位的氨基酸处,SD/97/02 株与其他毒株相比多了一个氨基酸(Ala),这是否对 S1 蛋白的功能会产生影响还有待进一步研究。

鸡的传染性腺胃炎在中国已存在几年,现大部分研究者认为它是由一种冠状病毒即传染性支气管炎病毒引起的,但这是一种变异的 IBV,这意味着 IBV 的组织嗜性又发生了变化。而其 S1 蛋白的变异直接决定了 IBV 的组织嗜性,我们通过 RT-PCR 技术,取得了 SD/97/02 株的 S1 基因,并通过测序获得了其基因序列,对 SD/97/02 株 S1 基因序列测定结果的初步分析表明,该毒株与其它 IBV 的差异较大,尤其在高变区(5'-HVR)有极大的差异。

本文中所比较的序列中,其中一个(ZJ971 株)^[12]为腺胃表现型的 IBV,对 S1 蛋白的序列比较表明,SD/97/02 株 IBV 与这个毒株的 S1 蛋白的差异也很大,这就说明在病毒的致病型和 S1 等国的基因型间还看不出直接相关性。决定病毒的组织嗜性

的可能是其中的单一或几个抗原表位,这两个都是腺胃型的毒株虽然在整体 S1 基因同源性上相差较大,但在第 438~444aa 处均比其他毒株少了一个表位,该表位的缺失是否会造成 S1 蛋白整体抗原性的改变,有待进行深入的研究。

参考文献

- [1] Cavanagh D. Nidovirales: a new order comprising Coronaviridae and Arteriviridae[J]. Arch Virol, 1997, 142(3): 629-633.
- [2] Calnek B W, John H, Charles W, et al. Diseases of Poultry[M], 10th ed, Ames: Iowa State University Press, 1997, 511-526.
- [3] 王永坤,朱国强,田慧芳,等.鸡传染性腺胃炎的研究[J].江苏农学院学报,1996,17(1):52-53.
- [4] 王玉东,张子春,王永玲,等.鸡腺胃型传染性支气管炎研究初报[J].中国动物检疫,1997,14(3):6-8.
- [5] 潘杰彦,陈德胜,戴亚斌,等. IBV 青岛腺胃分离株 S1 纤突蛋白基因的克隆和鉴定[J].中国兽医科技,2000,30(11):7-9.
- [6] Cavanagh D, Davis P J, Darbyshire J H, et al. Coronavirus IBV: virus retaining spike glycopolyptide S2 but not S1 is unable to induce virus-neutralizing or haemagglutination-inhibiting antibody, or induce chicken tracheal protection[J]. J Gen Virol, 1986, 67(7): 1435-1442.
- [7] Cavanagh D and Davis P J. Coronavirus IBV: removal of spike glycopolyptide S1 by urea abolishes infectivity and haemagglutination but not attachment to cells[J]. J Gen Virol, 1986, 67: 1443-1448.
- [8] Bernard N, David M K. Fields Virology[M]. 2nd ed, New York: Raven Press, 1990, 841-856.
- [9] Sata S, Sato S, Okabe T, et al. Cloning and Sequencing of genes encoding structural proteins of avian infectious bronchitis virus[J]. Virology, 1988, 165: 589-595.
- [10] Glickman R L, Syddall R L, Iorio R M, et al. Quantitative basic residue requirements in the cleavage-activation site of the fusion glycoprotein as a determinant of virulence for Newcastle Disease Virus [J]. J Virol, 1988, 62: 354-356.
- [11] Wang L, Parr R L, King D J, et al. A highly conserved epitope on the spike protein of infectious bronchitis virus[J]. Arch Virol, 1995, 140: 2201-2213.
- [12] 周继勇,沈行燕,丁红梅,等.新的禽传染性支气管炎病毒变异毒株 S 基因克隆[J].中国农业科学,待发表