

猪瘟病毒的分子生物学研究进展

王 镇 丁明孝

(北京大学生命科学院细胞与遗传学系 北京 100871)

猪瘟病毒(Classical swine fever virus, CSFV; 或称 Hog cholera virus, HCV), 属黄病毒科(Flaviviridae)瘟病毒属(Pestivirus)^[1]。CSFV 与同属的牛病毒性腹泻病毒(Bovine viral diarrhoea virus, BVDV)及羊边界病病毒(Border disease virus, BDV)在结构上有很高的相似性, 在血清学上有交叉反应^[2]。感染猪可引起高热、皮肤变色、大量内出血等及神经系统症状^[3], 死亡率很高。猪瘟病毒是世界上损害养猪业的重要病原之一^[4]。

猪瘟病毒是有囊膜, 直径 40~60nm 的正链 RNA 病毒^[2,5]。在蔗糖中浮力密度为 1.12g/ml, 沉降系数 $S_{20} = 140 \sim 150$ 。能在多种动物细胞中增殖, 一般不使培养细胞产生病理变化(Cytopathic effect, CPE)^[5]。

近年来, 随着对 CSFV 研究的深入, 许多 CSFV 毒株基因组核酸序列已被全部或部分测定出来, 对 CSFV 基因组结构的研究也有了较大的进展^[6~9]。

1 CSFV 基因组结构

CSFV 的基因组核酸是一个 12.3kb 的有传染性的正链 RNA 分子, 由 5' 非编码区(5'-uncoding region, 5'-UCR)、一个可读框(ORF)和 3' 非编码区(3'-uncoding region, 3'-UCR)构成。

CSFV 的 5'-UCR 由 360~373 个核苷酸组成。5' 端缺乏甲基化的“帽子”结构, 序列保守性很高^[6], 可形成稳定的二级结构。5' UCR 中有多个不能启动翻译的隐性 AUG 密码子^[10], 可通过其内部的核糖体进入位点(internal ribosome entry site, IRES 或 ribosome landing pad, RLP)与核糖体结合, 并以一种不依赖于“帽”结构的机制(cap-independent mechanism)在 ORF 的 AUG 起始翻译^[10,11]。1992 年 Brown 等绘出了 CSFV 5'-UCR 的二级结构模式图^[10]。1996 年 Le 等人提出 IRES 与宿主核糖体的结合是以碱基配对模式(base-pairing model)进行的。核糖体中 18s rRNA 3' 保守序列 GUCGUAACAAGG 与瘟病毒 5'-UCR 3' 端 ORF 起始密码子 AUG 上游的单链区核苷酸互补, 通过碱基配对而结合^[11], 这段单链区是 IRES 功能发挥所必需的。

CSFV 5'-UCR 二级结构在起始及调控病毒蛋白质的合成中是必需的。序列比较发现: CSFV 5'-UCR 5' 端的两段序列 GTATACGAG 及 CTCGTATAC 在属内高度保守, 推测这两段序列在病毒 5'-UCR 稳定二级结构的形成及行使功能中具有一定作用^[6]。

CSFV 的 3'-UCR 由 226~243 个核苷酸组成, 3' 端缺乏 poly(A) 尾结构, 在属内保守性很高, 在病毒复制中起一定作用^[6~9]。Moormann 等人报道 CSFV 强毒 Brescia 株 3'-UCR 中有一个 50 个核苷酸的富含 A、T 的区域, 且 ATTT 在该区域中有 8 次重复。序列比较还发现: 弱毒力的 C 株与强毒 Brescia 及 Alfort 株相比, 其 3'-UCR 中有一段 13 个核苷酸的插入序列(TTTTCITTTTTT)^[6]。Khromykh 等已证实黄病毒 3'-及 5'-UCR 参与了病毒的复制调控, 与病毒复制酶对 RNA 模板链的识别有关^[12]。Moormann 等人测定 CSFV Brescia 株及 C 株感染细胞中病毒 RNA 的量, 发现前者中 CSFV 的 RNA 量明显高于后者^[6]。这些毒株生长能力与毒力的差异是否与 3'-UCR 中 13 个核苷酸的插入序列有关, 有待进一步实验证实。对 CSFV 5'-UCR 在病毒蛋白质合成、3'-UCR 对病毒 RNA 合成的调控中所起的作用等问题的研究对搞清 CSFV 的复制过程及其毒力和致病性是十分必要的, 并且对猪瘟的防治也是很有实用意义的。

CSFV ORF 可编码 3898 个氨基酸的多聚蛋白^[6]。多聚蛋白在翻译的同时及翻译后, 经宿主细胞及病毒编码酶的加工, 形成多个病毒蛋白。

2 CSFV 的结构蛋白

编码 CSFV 结构蛋白的基因主要定位于 CSFV 基因组 5' 端三分之一部分(图 1)。CSFV ORF 编码的多聚蛋白经加工, 可形成四种 CSFV 的结构蛋白, 即组成核衣壳的 p14 及囊膜上的糖蛋白 gp44/48、gp33 和 gp55。p14, 也称 C, 是病毒的衣壳蛋白, 由病毒 ORF 中的 Ser₁₆₉~Ala₂₆₇ 组成。其 N 端由 CSFV 非结构蛋白 p23 的水解作用

产生, C端可能是由宿主细胞的信号肽酶作用所致^[13]。

CSFV的囊膜糖蛋白是由约130kDa的前体(E^m12), 在宿主细胞信号肽酶、定位于高尔基复合体上的蛋白水解酶及糖基化酶的作用下加工成熟的。成熟的蛋白被运送到宿主细胞膜上, 并随着病毒的释放装配到病毒上^[13]。

gp44/48, 也称E^m, 旧称E0, 是病毒的一种囊膜糖蛋白^[14]。有9个可能的糖基化位点, 去糖基的蛋白骨架分子量约26kDa^[13,14], 由病毒ORF中Glu₂₆₃~Ala₄₉₄组成^[7,13]。在感染细胞中gp44/48在内质网内积累, 并可存在于细胞膜表面或被分泌到胞外^[13,15]。在病毒粒子中gp44/48以97kDa的同源二聚体形式存在于囊膜表面。由于其分子内缺乏疏水的膜锚定区, 与病毒囊膜结合力弱, 易从囊膜上游离下来^[16]。gp44/48可诱导产生猪瘟的中和抗体, 免疫猪可诱导产生对致死量CSFV的保护性免疫^[15]。由于编码gp44/48的核酸序列在属内比编码gp55的序列保守程度高, 因此E^m可作为防治猪瘟的一种靶蛋白^[17]。

E^m具有RNase活性, 作用于单链的RNA, 且对富含U的序列活性最高^[14]。4ng的E^m在体外即可将CSFV基因组RNA完全降解。推测E^m没有将其自身RNA降解掉的原因可能是由于空间上的隔离。在宿主细胞中, 新合成的E^m在内质网腔内, 而其基因组RNA则在胞浆中; 在病毒粒子中, 定位于囊膜的E^m与其基因组RNA间隔着衣壳蛋白, 因此E^m是不会降解其自身RNA的。编码E^m的基因是瘟病毒属病毒所特有的, 黄病毒科中其它病毒基因组内无此基因, 推测该基因是病毒与宿主细胞的RNase基因重组得到的^[16]。

gp33, CSFV的囊膜糖蛋白, 也称E1^[13], 去糖基的蛋白骨架分子量21.8kDa, 由195个氨基酸残基(ORF编码的Leu₄₉₅至Gly₆₈₉)组成, 内有3个可能的N-连接的糖基化位点。其蛋白骨架中有2段疏水序列(Leu₅₄₈~Pro₅₇₉, Thr₆₅₉~Gly₆₈₉)。在多聚蛋白转位过程中, 前一段是E1向内质网腔转位的终止信号, 后一段则充当E2转位的信号肽; 除此之外, 这二段疏水区还充当E1的膜锚定位点^[15]。E1埋在病毒囊膜内^[21], 不能诱导猪产生抗体。

gp55, CSFV的囊膜糖蛋白, 称为E2, 是病毒主要的抗原蛋白^[15]。也是三个病毒糖蛋白中保守性最低, 最易变异的分子^[17]。gp55常以100kDa的同源二聚体及与gp33形成75kDa的异源二聚体形式存在于病毒粒子及CSFV感染的细胞表面。在体外gp55可诱导产生病毒

的中和抗体^[17], 体内可诱导抗CSFV的保护性免疫, 保护猪抵抗致死量CSFV的攻击^[15,17,18]。Hulst用20μg昆虫细胞表达的E2双相油包水乳剂免疫猪, 可使之获得对100LD₅₀(Lethal dose) CSFV Brescia毒株强毒的抵抗力^[18]。gp55的蛋白骨架由370个氨基酸(ORF编码的690~1060氨基酸残基)组成, 并以其C端的40个疏水氨基酸锚定在膜上^[13~19]。由于糖基化程度不同, E2的分子量可为51~58kDa。E2分子内的15个Cys残基在属内均保守, 其中N端6个Cys残基参与抗原结构域的形成, C端9个Cys残基则参与同源、异源二聚体的形成^[19]。

gp55的抗原决定区在其N端部分(ORF编码的第690~866氨基酸), 可分为2个独立的抗原结构单元^[19], 四个抗原结构域(domain)(A、B、C及D)。其中结构域A又可分为A₁、A₂、A₃三个亚域(Subdomain)^[19,20]。结构域B、C属于一个结构单元, 均为中和抗体的结合位点, 由ORF编码的690~800氨基酸残基参与构成。Cys₆₉₃与Cys₇₃₇形成的二硫键对结构域B、C的空间结构至关重要^[20]。另一个结构单元含A、D抗原结构域。A₁、A₂亚区定位于第795~851氨基酸残基, 在种内保守, A₁是一个抗体的中和位点^[19,20]。亚区A₃定位于766~813氨基酸残基, 而氨基酸残基766~800则参与了结构域D的构成^[19]。Cys₇₉₂~Cys₈₅₆及Cys₈₁₈~Cys₈₂₈间的二硫键对这个结构单元空间结构的形成是必需的。连接这二个结构单元的是一段属内保守的疏水序列, 推测这段序列参与了病毒侵入细胞过程中的膜融合过程^[20]。

3 CSFV的非结构蛋白

CSFV的ORF可能编码6~7种非结构蛋白(图1)。除N^{ro}的编码序列, 编码其它非结构蛋白的序列定位于CSFV基因组3'端三分之二部分。

N^{ro}, 分子量23kDa病毒非结构蛋白。是ORF编码的多聚蛋白N端的第一个蛋白质, 由168个氨基酸构成, 具有蛋白水解酶活性, 它以自催化的方式从正在翻译的多聚蛋白上断裂下来, 成为成熟的病毒蛋白。断裂位点在Cys₁₆₈与其下游的病毒核衣壳p14N端第一个氨基酸Ser₁₆₉间。序列比较发现, 编码p23的基因在瘟病毒属中是较为保守的。Cys₆₉和His₁₃₀位于p23的活性中心, 可水解断裂Cys与Ala, 及Gly与Ser间的肽键, 断裂位点位于活性中心His残基下游38个氨基酸残基处^[21]。N^{ro}不参与CSFV的结构及非结构蛋白的加工过程, 其蛋白酶活性在CSFV增殖过程中的作用还不清楚。

p120位于CSFV结构蛋白下游, 分子量为120kDa,

与 BVDV 非结构蛋白 p125 / p80 在血清学上有交叉反应^[18,22]。编码该蛋白 C 端的序列在是瘟病毒属中保守性最高的一段区域^[2]。该段序列在 BVDV 中编码的 p80 是多功能蛋白,具有丝氨酸型的蛋白水解酶、RNA 刺激的核苷三磷酸酶及 RNA 解旋酶等活性,在病毒非结构蛋白的加工、成熟基因组复制与转录过程中,以及宿主细胞 CPE 的产生中起重要作用^[23~25]。在 CSFV 中该段序列编码的蛋白在病毒增殖过程及与宿主细胞相互关系中的作用目前还不太清楚。Meyers 等发现可以使培养细胞产生 CPE 的 cp 型 CSFV Alfort / M 株、ATCC 株及 Steriomer 株是由含 8kb 左右,缺失了 CSFV 基因组 5' 三分之一序列的 DI 颗粒 (defective interfering particles) 及含 12.3kb 全基因组 RNA 的辅助病毒 (helper virus) 构成。在病毒的增殖过程中,8kb 核酸表达的 NS3 导致了培养细胞 CPE 的产生^[26]。他们用 CSFV cDNA 体外缺失转录出同样长度的 RNA 转染细胞,也可促使 CSFV 感染的细胞产生 CPE^[27],表明这几株 cp 型 CSFV 中的 DI 颗粒是造成 CSFV 感染细胞产生 CPE 的因子。在 BVDV 中 p125 C 端序列转译出的 p80 是 cp 型 BVDV 使培养细胞产生 CPE 的因素。推测同 BVDV 类似,CSFV 使培养细胞产生 CPE 需要表达一定量的 NS3。CSFV 的 NS3 是否同 BVDV 的 p80 一样参与病毒非结构蛋白的加工及基因组复制过程,还有待更多更直接的实验证据。

目前对 CSFV 大部分非结构蛋白,尤其是 CSFV 基因组 3' 部分编码的蛋白质还缺乏了解。经核酸序列分析,CSFV ORF 编码的多聚蛋白 C 端存在 RNA 聚合酶活性中心的典型序列 Gly-Asp-Asp,可能是病毒的复制酶^[8,13],但要确定 CSFV 基因组 3' 端 2 / 3 的序列编码的病毒非结构蛋白的结构功能还需更多的实验证据。

在分子水平上研究 CSFV,了解其基因组结构及蛋白表达策略,促进了人们对 CSF 增殖过程及其与宿主细胞间关系等细节的了解,对揭示 CSFV 增殖过程及致病机制起了很大的促进作用,必将为猪瘟病毒的防治工作提供理论指导。

参 考 文 献

- [1] Wengler G. Family Flaviviridae. In Franchi R I B *et al* (eds), Classification and nomenclature of viruses, Fifth report of the International Committee on Taxonomy of Viruses, Berlin, Springer Verlag, 1991, 223~233.
- [2] Moenig V. Vet Microbiol, 1990, 23: 35~54.
- [3] Susa M, Konig M, Saalmueller A, *et al*. J Virol, 1992, 66: 1171~1175.
- [4] Pan I C, Huang T S, Pan C H, *et al*. Arch Virol, 1993, 131: 475~481.
- [5] 殷震, 刘景华. 动物病毒学. 北京: 科学出版社, 1983, 517~529.
- [6] Moormann R J M, van Gennip H G P, Miedema G K W, *et al*. J Virol, 1996, 70: 763~770.
- [7] Meyers G, Rumenapf T, Thiel H-J, *et al*. Virol, 1989, 171: 555~567.
- [8] Moormann R J M, Warnerham P A M, van der Meer B, *et al*. Virol, 1990, 177: 184~198.
- [9] Ishikawa K, Nagai H, Katayama K, *et al*. Arch Virol, 1995, 140: 1385~1391.
- [10] Brown E A, Zhang H, Ping L H, *et al*. Nucleic Acids Res, 1992, 20: 5031~5045.
- [11] Le S Y, Sonenberg N and Jr Maize J V. Gene, 1995, 154: 137~143.
- [12] Khromykh A A, Westaway E G. J Virol, 1994, 68: 4580~4588.
- [13] Rumenapf T, Unger G, Strauss J H, *et al*. J Virol, 1993, 67: 3288~3294.
- [14] Windisch J M, Schneider R, Stark R, *et al*. J Virol, 1996, 70: 352~358.
- [15] Konig M, Lengsfeld T, Pauly T, *et al*. J Virol, 1995, 69: 6479~6486.
- [16] Hulst M M, Hines G, Newbigin E, *et al*. Virol, 1994, 200: 558~565.
- [17] Weiland E, Stark R, Haas B, *et al*. J Virol, 1990, 64: 3563~3569.
- [18] Hulst M M, Westra D F, Wensvoort G, *et al*. J Virol, 1993, 67: 5435~5442.
- [19] van Rijn P A, de Meijer E J, van Gennip H G P, *et al*. J Gen Virol, 1993, 74: 2053~2060.
- [20] van Rijn P A, Miedema G K W, Wensvoort G, *et al*. J Virol, 1994, 68: 3934~3942.
- [21] Stark R, Meyers G, Rumenapf T, *et al*. J Virol, 1993, 67: 7088~7095.
- [22] Greiser-Wilke I, Dittmar K E, Liess B, *et al*. J Gen Virol, 1992, 73: 47~52.
- [23] Wiskerchen M, Collett M S. Virol, 1991, 184: 341~350.
- [24] Tamura J K, Warrenner P, Collett M S, Virol, 1993, 193: 1~10.
- [25] Warrenner P, Collett M S. J Virol, 1995, 69: 1720~1726.
- [26] Meyers G and Thiel H-J. J Virol, 1995, 69: 3683~3689.
- [27] Meyers G, Thiel H-J., Rumenapf T. J Virol, 1996, 70: 1588~1595.