维普资讯 http://www.cqvip.com

文章编号:0253 - 3626(2005)06 - 0808 - 05

SARS 冠状病毒 N 蛋白的真核细胞定位研究

陶 鹏,陈维贤,郭进军,黄爱龙

(1. 重庆医科大学病毒性肝炎研究所;2. 感染性疾病分子生物学教育部重点实验室 重庆,400010)

【摘 要】目的:为进一步探讨 SARS - CoV 核衣壳蛋白(N 蛋白)在病毒复制及细胞通路中的作用,观察了核衣壳蛋白不同区段在细胞内的定位。方法:应用核定位序列分析软件分析 N 蛋白序列,找出其中的核定位信号序列;构建增强型绿色荧光蛋白(EGFP)与 SARS - CoV 不同长度核衣壳蛋白区段融合表达的 pEGFP - C1 重组载体,对 293 细胞进行瞬时转染,在荧光显微镜下观察核衣壳蛋白不同区段在细胞内的定位。结果:不同的核定位信号分析软件分析的结果不同,Prosite 找到的核定位信号位于 N 蛋白第 373 - 389 或 374 - 390aa,而 PredictNLS server 找到的核定位信号位于 N 蛋白第 36 - 44aa;全长核衣壳蛋白及第 161 - 422aa 区段均定位于细胞质,第 1 - 187aa 区段定位于细胞质和细胞核。结论:SARS - CoV 核衣壳蛋白的核定位信号可能位于 N 蛋白第 36 - 44aa。

【关键词】SARS - CoV;核衣壳蛋白;细胞定位

【中图分类号】R511 【文献标识码】A 【收稿日期】2005 - 3 - 23

Cellular localization of SARS - CoV nucleocapsid protein in 293 cells

TAO Peng et al

(*Institute for Viral Hepatitis & Key Laboratory of Molecular Biology on Infectious Diseases of The Ministry of Education, Chongqing Medical University*)

【Abstract】Objective: To study the role of nucleocapsid protein in virus replication and cell signaling passway, and to observe the cellular localization of SARS - CoV nucleocapsid protein. Methods: The sequence of SARS - CoV nucleocapsid protein was analyzed by protein analysis programs to identify nuclear localization signals(NLS); 293 cells were transfected by the pEGFP - C1 recombinant plasmids which expressed fusion protein of enhanced green fluorescence protein(EGFP) and nucleocapsid protein. At 24, 48, 72 hours post transfection (p. t.), the Green Fluorescent Proteins (GFP) were observed by fluorescence microscope. Results: By using PROSITE, we identified a nuclear localization signal sequence from amino acids 373 to 389 or from amino acids 374 to 390 in the nucleocapsid protein sequence. However, the analysis of PredictNLS server revealed that the nuclear localization signal region was located between amino acids 36 and 44. N1 protein(1 - 187aa) was located in both cytoplasm and cell nucleu; N protein(1 - 422aa) and N2 protein(161 - 422aa) were located in cytoplasms. Conclusion: The nuclear localization signal is likely to be located between amino acids 36 and 44 in SARS - CoV nucleocapsid protein.

【Key words】SARS - CoV; Nucleocapsid protein; Localization

严重急性呼吸综合症(SARS)是一种由 SARS 冠状病毒(SARS - CoV)引起的急性传染病。SARS - CoV 的基因组为单股正链 RNA,全长约 30kb,有 S,E,M,N 等结构蛋白,其中 N 蛋白在冠状病毒的生理过程中有两个方面的作用:在病毒包装的过程中和病毒 RNA 结合并与 M 蛋白(small membrane protein)相互作用使病毒 RNA 和 N 蛋白复合体被包裹进病毒核衣壳中;在病毒的转录和复制过程中与 mRNA 相互作用,可能影响病毒 RNA 的转录和复制[1,2]。此外,N 蛋白还有可能参与了细胞信号通路 AP - 1 的传递[3]。通过生物信息学分析,发现 N 蛋白序列中存在有核定位信号(nuclear localization signals, NLS)[4],本研究拟构建不同长度 N 蛋白的重组表达载体,转染细胞后通过荧光显微镜观察 N 蛋白在细胞内的定位。

作者介绍:陶 鹏(1973 -),男,主治医师,博士,主要研究方向:感染性疾病分子生物学。

1 材料与方法

维普资讯 http://www.cqvip.com

1.1 材料

1.1.1 质粒和细胞:增强型绿色荧光蛋白表达载体 pEGFP-C1 由美国 University of Chicago Medical Center 的 Tong-Chuan He 博士惠赠。含 SARS-CoV N 基因的质粒 pCDNA-N 为加拿大国立微生物实验室的 Run-Tao He 博士惠赠。大肠杆菌 DH5α 和人胚肾细胞系 293 细胞由本研究所保存。

1.1.2 主要试剂:T4 DNA 连接酶购自 Promega 公司、各种限制性内切酶和 PCR 试剂盒为 Takara 公司产品,脂质体转染试剂 Lipofectamine2000 购于 Invitrogen 公司。

1.2 方法

由于用两种核定位序列分析软件分析的结果不一致,核定位信号分别位于 N 蛋白的 N 端和 C 端,因此本研究不仅观察了 N 蛋白全长的细胞定位,还观察了 N1 蛋白(N 蛋白的第 1-187aa,由 N 基因 1-561bp 翻译)及 N2 蛋白(N 蛋白的第 161-422aa,由 N 基因 483-1269bp 翻译)的细胞定位。

含 SARS-CoV N 基因全长的质粒 pcDNA-N 为国外惠赠,而分别含有 N 基因全长、N1(1-561bp)片段和 N2(483-1269bp)片段的重组质粒 pEGFP-C1-N、pEGFP-C1-N1 和 pEGFP-C1-N2 则由自己构建。

用 PCR 方法扩增目的片段,利用 PCR 产物两端携带的 Sal I 和 BamH I 酶切位点和 pEGFP-C1 载体连接,得到能表达完整 N 蛋白的重组质粒 pEGFP-C1-N。

化学合成 SARS-CoV N1、N2 片段(位于载体 pet32a(+)),利用目的片段两端的 BamH I 和 Hind III 酶切位点将其从 pet32a(+)切下,利用 Bgl II 与 BamH I 有相同的粘端,绿色荧光蛋白表达载体 pEGFP-C1 用 Bgl II 和 Hind III 消化,目的片段和载体连接,得到能表达 N 蛋白片段的重组质粒 pEGFP-C1-N1 和 pEGFP-C1-N2,同时目的基因的读码框不变。

1.2.1 SARS-CoV N 蛋白核定位序列分析

从 NCBI 获得 SARS 冠状病毒核衣壳蛋白的氨基酸序列,将其提交到 prosite 及 PredictNLS server (prediction and analysis of nuclear localization signals, Columbia University Bioinformatics Center)进行核定位信号分析。

1.2.2 融合表达载体 pEGFP-C1-N 的构建

通过 PCR 扩增得到 SARS-CoV N 全基因。以含 SARS-CoV N 全基因的质粒 pCDNA-N 为模板,上游引物 5′-CGCGTCGACTCTGATAATGGACCCCAATC-3′,下游引物 5′-GTACGGATCCTTATGCCTGAGTTGAATCAGCAGA-3′,划线部分为酶切位点(Sal I, BamH I), PCR 产物经 Sal I 和 BamH I 双酶切后克隆于增强型绿色荧光蛋白表达载体 pEGFP-C1。Sal I 和 BamH I 双酶切鉴定,鉴定正确的克隆命名为 pEGFP-C1-N。将筛选的重组质粒送上海申友生物技术有限公司进行序列分析,以证实目的 DNA 片段克隆入 pEGFP-C1 载体。

1.2.3 融合表达载体 pEGFP-C1-N1 的构建

根据公布的 SARS-CoV 基因序列(GenBank 序号为 AY274119),人工合成部分 N 基因(第 1-561bp),其 5 端和 3 端分别带有 BamH I 和 Hind III 酶切位点,由上海生工生物工程技术公司完成。将目的片段 N1 用 BamH I 和 Hind III 从合成时的表达载体 PET32a(+)切下;因 BamH I 与 Bgl II 有相同的粘端,绿色荧光蛋白表达载体 pEGFP-C1 用 Bgl II 和 Hind III 双酶切,二者连接后目的片段 N1 位于 GFP 下游,与 GFP 构成融合基因。Hind III 单酶切,Bgl II 单酶切,Hind III + Xho I 双酶切鉴定,鉴定正确的克隆命名为 pEGFP-C1-N1。将筛选的重组质粒送上海申友生物技术有限公司进行序列分析,以证实目的 DNA 片段克隆入 pEGFP-C1 载体。

1.2.4 融合表达载体 pEGFP-C1-N2 的构建

根据公布的 SARS-CoV 基因序列(GenBank 序号为 AY274119),人工合成部分 N 基因(第 483-1269bp),其 5′端和 3 端分别带有 BamH I 和 Hind III 酶切位点,由上海生工生物工程技术公司完成。将目的片段 N2 用 BamH I 和 Hind III 从合成时的表达载体 PET32a(+)切下;因 BamH I 与 Bgl II 有相同的粘端,绿色荧光蛋白表达载体 pEGFP-C1 用 Bgl II 和 Hind III 双酶切,二者连接后目的片段 N2 位于 GFP 下游,与 GFP 构成融合基因。Bgl II 单酶切,Bgl II + Hind III 双酶切鉴定,鉴定正确的克隆命名为 pEGFP-C1-N2。将筛选的重组质粒送上海申友生物技术有限公司进行序列分析,以证实目的 DNA 片段克隆入 pEGFP-C1 载体。

1.2.5 转染 293 细胞

将 pEGFP-C1-N ,pEGFP-C1-N1 和 pEGFP-C1-N2 分别转染 293 细胞,同时设立以下对照:

(1)阳性对照 pEGFP-HCV-C1 组,即针对丙型肝炎病毒核心蛋白构建的重组质粒。

(2)空载体 pEGFP-C1 组

置 37℃培养,5h 后用含 10% 胎牛血清的 RPMI-1640 培养液换液,观察细胞生长情况及荧光表达。

1.2.6 N 蛋白细胞内定位的观察

由于 GFP 与 N 蛋白表达为融合蛋白,根据 GFP 在 293 细胞中表达的位置可以间接反映 N 蛋白的细胞内定位。于转染 24~72h,在荧光显微镜下观察绿色荧光蛋白的表达情况。

2 结 果

2.1 N 蛋白核定位序列分析

用 **prosite** 分析出的核定位信号位于 N 蛋白 C 端(373~389 或 374~390aa),长度 18aa,富含赖氨酸 K,呈碱性特征。

MSDNGPQSNQRSAPRITFGGPTDSTDNNQNGGRNGA
RPKQRRPQGLPNNTASWFTALTQHGKEELRFPRGQ
GVPINTNSGPDDQIGYYRRATRRVRGGDGKMKELSP
RWYFYYLGTGPEASLPYGANKEGIVWVATEGALNT

维普资讯 http://www.cqvip.com

PKDHIGTRNPNNNAATVLQLPQGTTLPKGFYAEGSR
GGSQASSRSSSSRSRGNSRNSTPGSSRGNSPARMASGG
GETALALLLLDRLNQLESKVSGKGQQQQGQTVTKKS
AAEASKKPRQKRTATKQYNVTQAFGRRGPEQTQGN
FGDQDLIRQGTDYKHWPQIAQFAPSASAFFGMSRIG
MEVTPSGTWLTYHGAIKLDDKDPQFKDNVILLNKHI
DAYKTFPPTEPKKD***KKKKTDEAQPLPQRQKKQ***
PTVTLLPAADMDDFSRQLQNSMSGASADSTQA

用 **PredictNLS server** 分析出的核定位信号位于 N 蛋白 N 端(36 ~ 44aa),长度 9aa,富含精氨酸 R,呈碱性特征。

MSDNGPQSNQRSAPRITFGGPTDSTDNNQNGGRNG
ARPKQRRPQGLPNNTASWFTALTQHGKEELRFPRG
QGVPINTNSGPDDQIGYYRRATRRVRGGDGKMKEL
SPRWYFYYLGTGPEASLPYGANKEGIVWVATEGAL
NTPKDHIGTRNPNNNAATVLQLPQGTTLPKGFYAEG
SRGGSQASSRSSSSRSRGNSRNSTPGSSRGNSPARMAS
GGGETALALLLLDRLNQLESKVSGKGQQQQGQTVT
KKSAAEASKKPRQKRTATKQYNVTQAFGRRGPEQT
QGNFGDQDLIRQGTDYKHWPQIAQFAPSASAFFGM
SRIGMEVTPSGTWLTYHGAIKLDDKDPQFKDNVILLN
KHIDAYKTFPPTEPKKDKKKKTDEAQPLPQRQKKQP
TVTLLPAADMDDFSRQLQNSMSGASADSTQA

2.2 pEGFP - C1 - N 的鉴定

PCR 产物电泳结果见图 1,酶切电泳鉴定结果见图 2。从图 1 和图 2 可见,PCR 产物长约 1300bp,pEGFP - C1 - N 经 Bam H I 和 Sal I 双酶切得到的小片段和 PCR 产物长度相当。测序证实插入序列正确。

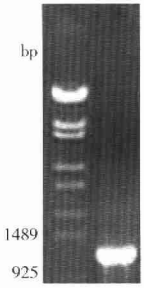
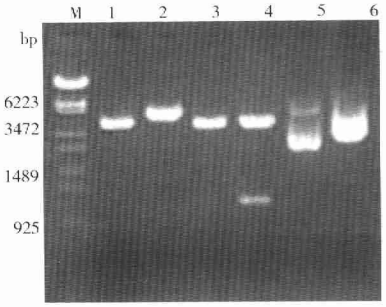


图 1 N 基因 PCR 产物电泳图

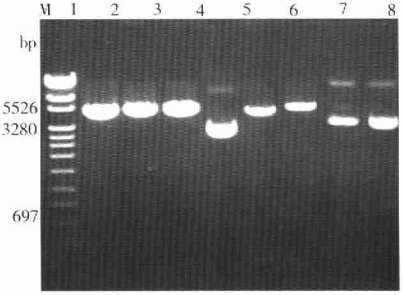
2.3 pEGFP - C1 - N1 的鉴定



M:marker; 1:pEGFP-C1/BamH I
2:pEGFP-C1-N/BamH I;3:pEGFP-C1-N/BamH I + Sal I;
4:pEGFP-C1-N/BamH I + Sal I; 5:pEGFP-C1;

图 2 重组 pEGFP - C1 - N 的酶切鉴定

酶切电泳鉴定结果见图 3。重组质粒 pEGFP - C1 - N1 Bgl II 酶切位点丢失,但可以被 Hind III 酶切。而空载体可分别被 Bgl II 和 Hind III 酶切。用 Xho I 和 Hind III 双酶切后,由于在 N1 内部 203bp 处(N1 长度 561bp)有一 Xho I 酶切位点,重组质粒产生 4691bp 和 359bp 两个条带,而空载体 pEGFP - C1 产生 4691bp 和 9bp 两个片段,但电泳只有 4691bp 带可见。测序证实插入序列正确。



M:Marker; 1:pEGFP-C1/XHo I + Hind Ⅲ;
2:pEGFP-C1/Hind 3:pEGFP-C1/Bgl Ⅱ;
4:pEGFP-C1-N1/Hind Ⅲ;5:pEGFP-C1-N1/XHo I + Hind Ⅲ;
8:pEGFP-C1-N1

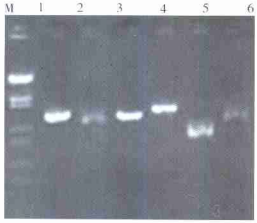
图 3 重组 pEGFP-C1-N1 的酶切鉴定

2.4 pEGFP - C1 - N2 的鉴定

酶切电泳鉴定结果见图 4。重组质粒 pEGFP - C1 - N2 Bgl II 酶切位点丢失,但可以被 Hind III 酶切。而空载体可分别被 Bgl II 和 Hind III 酶切。用

维普资讯 http://www.cqvip.com

Bgl Ⅱ 和 Hind Ⅲ 双酶切后，重组质粒产生一个条带，而空载体 pEGFP－C1 产生 4691bp 和 9bp 两个片段，但电泳只有 4691bp 带可见。测序证实插入序列正确。



M：marker；　1：pEGFP-C1/Bgl Ⅱ；
2：pEGFP-C1-N2/Bgl Ⅱ；　3：pEGFP-C1/Bgl Ⅱ＋Hind Ⅲ；
4：pEGFP-C1-N2/Bgl Ⅱ＋Hind Ⅲ；　5：pEGFP-C1；
6：pEGFP-C1-N2

图 4　重组 pEGFP-C1-N2 的酶切鉴定

2.5　N 蛋白的细胞定位

转染细胞 24、48、72h 观察绿色荧光，以 48h 荧光最强。从图 5 可以看出，全长 N 蛋白及其第 161～422aa 区段（N2 蛋白）均位于细胞浆内，第 1～187aa 区段（N1 蛋白）定位于细胞浆和细胞核，阳性对照丙型肝炎病毒的核心蛋白 C1 段位于细胞核内，而空载体 pEGFP－C1 转染后细胞核和细胞浆都有荧光，呈均匀分布。



图5-1　pEGFP-C1-N

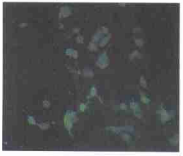


图5-2　pEGFP-C1-N1

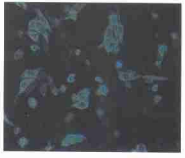


图5-3　pEGFP-C1-N2

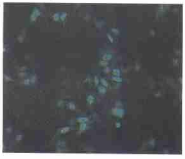


图5-4　pEGFP-HCV-C1

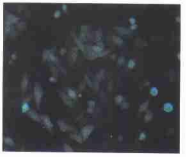


图5-5　pEGFP-C1

图5 荧光显微镜下N蛋白的细胞内定位

3　讨　论

SARS－CoV 核衣壳蛋白（N 蛋白）全长 1269bp，编码 422 个氨基酸，富含 lysine，在病毒复制过程中大量表达，有很高的抗原性，目前已经制备出高效价的抗体（包括多抗和单抗），作为检测 SARS 感染的试剂[5～8]。SARS－CoV 核衣壳蛋白是其 4 种结构蛋白中唯一和基因组 RNA 结合在一起的（其余三种为包膜蛋白），位于病毒的核心颗粒，它不仅参与了病毒的包装，还对病毒基因组的复制和转录有调控作用。此外，He R 的研究发现 SARS－CoV 核衣壳蛋白在细胞中的表达可以激活 c－Fos，FosB 和 ATF－2 等转录因子，从而参与细胞 AP－1 信号通路的转导，可能在发病过程中起作用[3]。

维普资讯 http://www.cqvip.com

核定位信号(nuclear localization signal,NLS)是存在于亲核蛋白内的一些短的氨基酸序列片段,富含碱性氨基酸残基,如 Lys、Arg,此外还常含有 Pro。它与信号肽不同的是,在不同的蛋白质中分布的位置不同,并且蛋白入核后不被切除。从现有的核信号序列可找到一个统一的 bipartite motif,Dingwall C[9]提出了确定 bipartite motif 的三条规则:(1)两个相邻的碱性氨基酸(Arg or Lys);(2)一个含十个氨基酸残基的 spacer;(3)在 spacer 之后的 5 个 10 位置至少有三个碱性氨基酸残基(Arg or Lys)。我们通过两种蛋白分析软件的分析,发现结果不同。Prosite 分析出的核定位信号位于 N 蛋白 373 ~ 389 aa 或 374 ~ 390aa(N2 段),而 PredictNLS server 分析出的核定位信号位于 36 ~ 44aa(N1 段)。理论上具有核定位信号的蛋白质应该定位在细胞核内,而本实验却发现 N1 段位于细胞核和细胞质(与 pEGFP - C1 空载体转染后荧光分布一致),N 蛋白全长和 N2 片段都位于细胞质内。从实验结果看,似乎 PredictNLS server 的分析更可信。实验结果与理论不符的可能原因如下:1 .核定位信号预测软件设计缺陷,导致分析结果的不确定性;2 细胞中由于生物大分子之间的相互作用,屏蔽了 N 蛋白的核定位区域,从而影响了 N 蛋白细胞定位。已有研究发现,在哺乳动物细胞内,N 蛋白不仅可以因自身相互作用形成二聚体[10,11],而且可以和细胞内的 HAb18G/CD147 以及异质核蛋白 A1(Heterogeneous nuclear ribonucleoprotein A1 ,hnRNP A1)结合[12,13],虽然与它们的结合位点不在核定位区域,我们不排除有其他蛋白通过核定位区域与 N 蛋白结合的可能;3 融合蛋白与天然蛋白之间结构的差异也可能影响 N 蛋白的细胞定位;4 GFP 本身在细胞内的定位影响了 N 蛋白的细胞定位。

由于核定位信号序列的规则并不能适用于所有的核蛋白,不同的核定位信号分析软件分析结果不一样,而且有小部分非核蛋白会有阳性结果,对 SARS 冠状病毒 N 蛋白的序列还需要进行分析,它在细胞中和其他蛋白以及自身的相互作用如何,其功能和在 SARS 发病中的意义等问题还需要更进一步的研究。

参 考 文 献

[1] Marra MA, Jones SJ, Astell CR, et al. The genome sequence of the SARS - associated coronavirus[J]. Science,2003;300(5624):1399 - 1404.

[2] Paul A, Rota, M. Steven Oberste Stephan S, et al. Characterization of a Novel Coronavirus Associated with Severe Acute Respiratory Syndrome[J]. Science,2003; 300(5624): 1394 - 1399.

[3] He R, Leeson A, Andonov A,et al. Activation of AP - 1 signal transduction pathway by SARS coronavirus nucleocapsid protein[J]. Biochem Biophys Res Commun, 2003 ;311(4): 870 - 876.

[4] 石 磊,张其鹏,芮 伟. SARS 冠状病毒 N 蛋白序列中的核定位信号网站. 中国生物医学信息,SARS 病毒基因组序列 CMBI - PUCHG 生物信息学分析

[5] Zakhartchouk AN, Viswanathan S, Mahony JB, et al. Severe acute respiratory syndrome coronavirus nucleocapsid protein expressed by an adenovirus vector is phosphorylated and immunogenic in mice[J]. J Gen Virol, 2005 ;86(Pt 1):211 - 215.

[6] Ohnishi K, Sakaguchi M, Kaji T, et al. Immunological detection of severe acute respiratory syndrome coronavirus by monoclonal antibodies[J]. Jpn J Infect Dis, 2005 ;58(2):88 - 94.

[7] 王彦斌,常昭瑞,王健伟,等. SARS 冠状病毒 N 蛋白单克隆抗体的制备及鉴定[J]. 中华实验和临床病毒学杂志, 2004 ;18(4):316 - 320.

[8] Di B, Hao W, Gao Y, et al. Monoclonal antibody - based antigen capture enzyme - linked immunosorbent assay reveals high sensitivity of the nucleocapsid protein in acute - phase sera of severe acute respiratory syndrome patients[J]. Clin Diagn Lab Immunol, 2005 ;12(1):135 - 140.

[9] Dingwall C, Laskey RA. Nuclear targeting sequences—a consensus[J]? Trends Biochem Sci, 1991 ;16(12):478 - 481.

[10] He R, Dobie F, Ballantine M, et al. Analysis of multimerization of the SARS coronavirus nucleocapsid protein[J]. Biochem Biophys Res Commun, 2004 ;316(2):476 - 483.

[11] Surjit M, Liu B, Kumar P, et al. The nucleocapsid protein of the SARS coronavirus is capable of self - association through a C - terminal 209 amino acid interaction domain[J]. Biochem Biophys Res Commun, 2004 ;317(4): 1030 - 1036.

[12] Chen Z, Mi L, Xu J, et al. Function of HAb18G/CD147 in invasion of host cells by severe acute respiratory syndrome coronavirus[J]. J Infect Dis, 2005 ;191(5):755 - 760.

[13] Luo H, Chen Q, Chen J, et al. The nucleocapsid protein of SARS coronavirus has a high binding affinity to the human cellular heterogeneous nuclear ribonucleoprotein A1[J]. FEBS Lett, 2005 ;579(12):2623 - 2628.