论著

文章编号:1000-5404(2005)01-0036-03

# 幽门螺杆菌 *cagA* 基因全长克隆及序列分析

周建嫦[1],张建中[2],徐采朴[1],何利华[2] ([1]第三军医大学西南医院全军消化内科中心,重庆 400038;[2]中国疾病预防控制中心传染病控制所腹泻研究室,北京 102206)

**提 要:目的** 探索扩增 *cagA* 基因全长的PCR方法,克隆中国2株胃癌幽门螺杆菌(*Helicobacter pylori*, Hp)分离株和1株十二指肠溃疡分离株及国际标准株SS1的 *cagA* 基因,初步探讨 *cagA* 序列和磷酸化位点变异及其临床意义。**方法** 针对已知 *cagA* 基因两侧及cag致病岛右侧反转重复序列设计PCR引物,扩增 *cagA* 基因并克隆测序,分析已知 *cagA* 序列同源性及其酪氨酸磷酸化位点。**结果** 3株中国Hp分离株 *cagA* 氨基酸序列的同源性约为94%(93.9%、94%、94.5%),并都有pY117/118/122和EPIYA-A、B、D 4个酪氨酸磷酸化位点。SS1株 *cagA* 与西方菌株同源性大于85%,而与中国菌株的同源性小于80%,具有pY117/118/122和EPIYA-A、B、C 4个酪氨酸磷酸化位点。**结论** 建立的PCR方法可扩增 *cagA* 基因全长。CagA序列变异和磷酸化位点具有地区聚类特征,但与Hp感染临床结局无特殊关联。

**关键词:** 幽门螺杆菌; *cagA* 基因; CagA蛋白; 克隆; 序列分析

**中图法分类号:** R377; R394-33; R394.3 **文献标识码:** A

# Cloning and sequence analysis of *cagA* gene in full-length of *Helicobacter pylori*

ZHOU Jian-chang[1], ZHANG Jian-zhong[2], XU Cai-pu[1], HE Li-hua[2] ([1]Gastroenterology Research Center, Southwest Hospital, Third Military Medical University, Chongqing 400038; [2]Department of Diarrhea Research, Institute of Infectious Diseases Control, Center for Diseases Control and Prevention, Beijing 102206, China)

**Abstract: Objective** To establish PCR method for amplifying *cagA* gene in full-length, and clone the *cagA* genes of two Chinese *H. pylori* strains isolated from patients with gastric cancer and one from patients with duodenal ulcer as well as that of Sydney strain 1 so as to investigate the heterogeneity of the amino acid sequences and phosphorylation sites of CagA and its clinical significance. **Methods** PCR primers designed according to the nucleotide acid sequence flanking the open reading frame of *cagA* and the lower reverse repeat sequence of cag pathological island were used to amplify *cagA* genes of *H. pylori* strains. PCR products were then cloned and sequenced, and the homology of all published CagA amino acid sequences and their putative phosphorylation sites were analyzed. **Results** The amino acid sequences of the three Chinese *H. pylori* strains revealed about 94% homologous to each other (93.9%, 94%, and 94.5%), and four tyrosine phosphorylation motifs (pY117/118/122, EPIYA-A, B, and D) present in CagA protein of all three Chinese isolates. The SS1 CagA showed more than 85% homologous to Western *H. pylori* strains while less than 80% to Chinese *H. pylori* strains, and four tyrosine phosphorylation motifs (pY117/118/122, EPIYA-A, B, and C) present in it. **Conclusion** A PCR method to amplify *cagA* genes in full-length has been established. Although distinct geographical cluster characteristics exist, the heterogeneity of amino acid sequence and phosphorylation sites of CagA are not specifically associated with the manifestation of *H. pylori* infection.

**Key words:** *Helicobacter pylori*; *cagA*; CagA; cloning; sequence analysis

CagA(cytotoxin associated protein A)在幽门螺杆菌(*Helicobacter pylori*, Hp)致病中的作用已受到广泛的关注,研究显示大部分欧美地区 *cagA* 阳性Hp感染十二指肠溃疡和胃癌发生显著相关[1,2],而 *cagA* 存在于多数亚洲菌株中,不能作为区分Hp致病力差异的单一毒力标志[3,4]。由于 *cagA* 具有序列多态性及磷酸化位点数目及位置的差异,目前认为只有某一类CagA才具有致病性。国外有报道胃癌分离Hp株的 *cagA* 基因3′端可变区较长,而十二指肠溃疡分离株较短[5]。但我们发现我国大部分Hp菌株 *cagA* 基因3′端可变区

**基金项目:** 国家自然科学基金资助项目(39870032)
Supported by National Natural Science Foundation of China (39870032)

**作者简介:** 周建嫦(1972-),女,湖南省道县人,博士,讲师,主要从事幽门螺杆菌分子生物学及致病机制方面的研究。电话:(023)65423260, E-mail: zhoujch@sohu.com

**收稿日期:** 2004-08-05; **修回日期:** 2004-12-11

较短，与临床表现无显著相关[3]。新近研究表明，具有磷酸化位点的 CagA 被 Hp 的Ⅳ型分泌系统转运入宿主细胞，并由 Src 酪氨酸激酶进行磷酸化修饰后，可干扰细胞信号转导，使细胞骨架排列紊乱、转录因子表达改变，导致细胞增生[6]。因此，克隆分析 *cagA* 基因全长，对研究 CagA 功能、明确致病 CagA 的分子特征具有重要意义。我们探索了 PCR 扩增 *cagA* 基因全长的方法，并克隆了 3 株中国 Hp 分离株和国际标准株 SS1 的 *cagA* 基因，对相应的 CagA 氨基酸序列及磷酸化位点进行了分析。现报告如下。

## 1　材料与方法

### 1.1　Hp 菌株及培养条件

CAPM N62、CAPM N111 为胃癌分离株，CAPM N93 为十二指肠溃疡分离株，SS1（Sydney Strain 1）为国际标准株，均由中国 Hp 协作组菌种保藏中心提供。细菌在含 10% 羊血的 Hp 选择性培养基上，37 ℃微需氧环境下培养 3～4 d。大肠杆菌 JM109 购自华美生物工程公司。

### 1.2　主要试剂和试剂盒

*Taq* DNA 聚合酶、λDNA/*Eco*R Ⅰ + *Hind* Ⅲ分子量标准、dNTPs 购自华美生物工程公司；PCR 产物纯化试剂盒、PinPoint™ Xa-1 T 载体连接系统分别为 Mannheim Boehringer 和 Promega 公司产品。

### 1.3　方法

1.3.1　Hp DNA 提取　　从培养基上刮取生长良好的 Hp 菌落，1×TE 洗涤后采用酚/氯仿混合提取法抽提。

1.3.2　PCR 扩增 *cagA* 全长　　根据 GenBank 中 *cagA* 基因 5′和 3′端保守序列及 cag 致病岛 3′端反转重复序列设计引物。引物序列如下：

上游：*cagA* F 5′-AAC AAT GAC TAA CGA AAC CA-3′

下游：*cagA* R1 5′-TTT TGG TAT TCC TTA ATC CT-3′

*cagA* R2 5′-TAA AGA ATG GCT CAA ATT GT-3′

用 *cagA* F-R1 扩增 CAPM N62 和 SS1 的 *cagA*，*cagA* F-R2 扩增另两株菌的 *cagA*。

扩增条件为：94 ℃ 1 min，54 ℃ 1 min，72 ℃ 5 min，20 轮循环，最后于 72 ℃延伸 10 min。

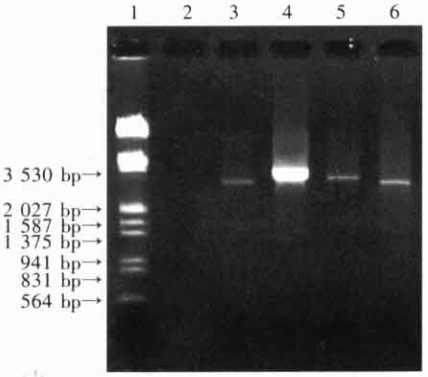
1.3.3　*cagA* 基因克隆、序列测定及分析　　PCR 产物用纯化试剂盒纯化后，按外源片段与 PinPoint™ Xa-1 T 载体分子比 3∶1 的比例加至 10 μl 连接反应体系中，16 ℃连接 12 h，按常规方法转化感受态 JM109 细胞。PCR 鉴定出含外源片段的阳性重组克隆，送上海博亚生物技术有限公司（BioAsia）完成 *cagA* 基因测序。上 GenBank 查询出已登录的另 40 个 *cagA* 基因序列，采用 Winstar 软件包中 Megalign 进行序列同源性分析并构建系统发生树。

1.3.4　CagA 磷酸化位点分析　　据 CagA 磷酸化模体[RK]-X(2,3)-[DE]-X(2,3)-Y[Arg/Lys-X(2,3)-Asp/Glu-X(2,3)-Y]和 EPIYA(Glu-Pro-Ile-Tyr-Ala)，确定 CagA 的酪氨酸磷酸化位点，后者依据其所处位置及周围氨基酸序列分为 A、B、C 和 D 4 类[7]。

## 2　结果

### 2.1　PCR 扩增 *cagA* 基因全长

CAPM N62 和 SS1 的 *cagA* 基因扩增产物长约 3.5 kb，CAPM N93、CAPM N111 株的 *cagA* 扩增产物长度均为 3.8 kb 左右（图 1）。




1：λDNA/*Eco*R Ⅰ + *Hind* Ⅲ standard molecular weight marker；2：Blank control；3：PCR product of *cagA* gene of strain SS1；4：PCR product of *cagA* gene of strain CAPM N111；5：PCR product of *cagA* gene of strain CAPM N93；6：PCR product of *cag*A gene of strain CAPM N62

**图 1　*cagA* 基因的 PCR 产物（1% 琼脂糖凝胶电泳）**

**Fig 1　PCR products of *cagA* genes of different *H. pylori* strains on 1% agarose gel**

### 2.2　*cagA* 基因测序结果及推定氨基酸序列同源性聚类分析

CAPM N62、CAPM N93、CAPM N111 和 SS1 株 *cagA* 基因序列已登录 GenBank（序列号分别为 AF249275、AF367250、AF367251、AF247651），其 ORF 长度分别为3 522 bp、3 519 bp、3 516 bp、3 540 bp，相应编码含1 173、1 172、1 171及1 179个氨基酸残基的 CagA 蛋白质。CagA 氨基酸序列同源性比较结果显示，CAPM N62、CAPM N93、CAPMN111 之间的同源性为 94.0%、93.9%、94.5%，与 31 株日本菌株中 28 株的同源性均高于 90.0%，并分别与其中的胃癌分离株 J-566(95.6%)、慢性萎缩性胃炎分离株 J-198 和慢性胃炎分离株 OK129(96.2%)、消化性溃疡分离株 J207 和慢性胃炎分离株 OK129(94.4%)的同源性最高，而与欧美菌株 ATCC53726(76.0%)、ATCC43526(75.9%)、ATCC43526(76.8%)最低；而 SS1 株则与同为澳大利亚分离株 NCTC11638 同源性最高(94.2%)，与日本菌株 CPY3401 同源性最低(76.1%)。SS1 与 3 株中国菌株 CagA 的同源性约为 80.0%。在系统发生树上，3 株中国菌株与 28 株日本菌株的 CagA 聚成一群，10 个西方菌株及另 3 株日本菌株聚成另一群，呈较明显地区聚类关系，而即使是同一群内，也未显示出明显的疾病聚类关系（图 2）。

### 2.3　CagA 磷酸化位点

CAPM N62、CAPM N93 及 CAPM N111 的 CagA 除有 1 个氨基端磷酸化位点（pY117/118/122）外，还有 EPIYA-A（pY883/884）、EPIYA-B（pY902/903）和 EPIYA-D（pY956/957）3 个羧基端磷酸化位点，与日本菌株中的消化性溃疡、胃癌、慢性萎缩性胃炎分离株各 3 株及 7 株慢性胃炎分离株一致。SS1 的 CagA 则与多数西方菌株类似，有 pY117、EPIYA-A（pY893）、EPIYA-B（pY912）及 1 个 EPIYA-C（pY965）位点。

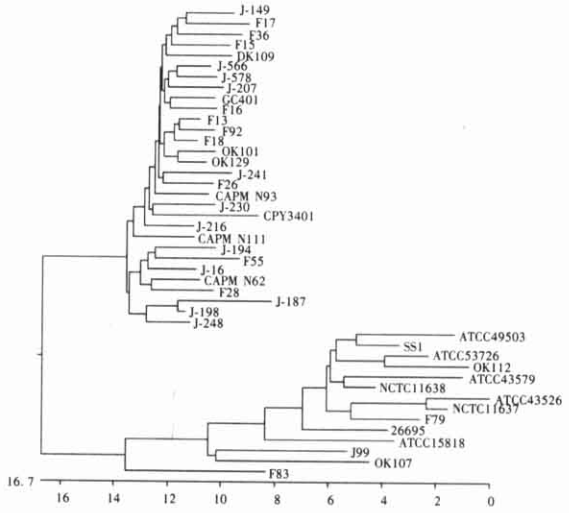



**图2 已知 CagA 氨基酸序列的系统发生树分析**

**Fig 2 Phylogenetic tree of all known CagA amino acid sequences created with Megalign**

## 3 讨论

由于 CagA 在 Hp 致病中具有重要作用，Hoshino[8]、彭颖等[9]进行了 *cagA* 基因克隆和序列分析研究。其中 Hoshino 等[8]采用超长 PCR 方法，针对其上下游的结构基因序列设计引物来扩增 *cagA* 基因全长，这无疑不经济；彭颖等[9]结合巢式 PCR 和 TD-PCR 方法，虽可扩出部分临床分离株的 *cagA* 基因，但操作十分繁琐。我们比较了已知 *cagA* 基因序列，发现 ORF 5′端相对保守，3′末端形成稳定的茎环结构，不能用于设计引物，3′端外侧的核苷酸序列变异较大。我们设计的引物 *cagA* R1 虽能扩增 SS1 株和 CAMP N62 株的 *cagA* 基因，但不能扩增出 CAPM N93 和 CAPM N111 的 *cagA* 基因。因而，我们利用 *cagA* 基因位于 cag 致病岛的最右侧端，其 3′端与 cag 致病岛的 31 bp 下游反转重复序列间仅间隔约 400 bp 的特点，设计了互补于 *cag* 致病岛的下游反转重复序列的引物 *cagA* R2，结果显示可有效扩增后两株 Hp 的 *cagA* 基因。由于该反转重复序列是 cag 致病岛的右侧边界，在所有 *cagA* 阳性菌株中保守，因而可用于扩增所有 Hp 菌株的 *cagA* 基因。这种用普通 PCR 扩增 *cagA* 基因的方法，简便而有效，为 *cagA* 检测、以及研究不同疾病相关 Hp 菌株的 *cagA*/CagA 分子特征提供了有力工具。

已知 CagA 的功能包括两个方面：①通过激活 Elk1、Ras、MEK 途径致血清反应元件(SER)和血清反应因子介导的基因转录活性增加，这是与 Grb2 相互作用的结果，依赖于 CagA 的羧基端序列(CagA 873-1002)，但与其磷酸化修饰无关[10]。②进入宿主细胞并被 Src 酪氨酸激酶磷酸化的 CagA 可锚固 SHP-2(SRC homology 2 domain-containing protein tyrosine phosphatase)并激活后者，导致胞内肌动纤维分布改变，细胞形态改变和增殖，表现出生长因子样效应[6]。有研究表明，分子量越大、磷酸化位点越多的 CagA 与胃癌的关系更为密切[5,7]。但本研究结果显示，3 株中国 Hp 的 CagA 不仅相互间氨基酸序列同源性很高、磷酸化位点一致，与大部分(28/31，90.3%)日本菌株 CagA 的氨基酸序列同源性也高于 90.0%，且磷酸化位点也与大部分(16/31，51.6%)日本菌株一样为 EPIYA-A-B-D 型，而无论氨基酸序列同源性还是磷酸化位点分析，均未显示出明显的疾病聚类关系，提示这类 CagA 在东亚 Hp 中可能占主导地位，分布于邻近地区的菌株主要源于同一祖先。44 个 CagA 全长序列中，60%(6/10)西方菌株有 pY117/118/122、EPIYA-A、EPIYA-B 磷酸化位点缺失，显著高于东亚菌株(5/34)；而 80%(4/5)缺失 pY117/118/122 的 CagA，往往有 3 个 EPIYA-C 位点，或在其他位置出现新的磷酸化位点，其意义尚待进一步研究。在获得 *cagA* 基因全长的基础上，分段表达有代表性的 CagA 片段，或对磷酸化位点进行突变，对它们的功能进行研究，将有助于明确 CagA 序列及其磷酸化位点变异在其功能中的意义。

## 参考文献：

[1] Oliveira A G, Santos A, Guerra S A, *et al*. babA2- and *cagA*-positive *Helicobacter pylori* strains are associated with duodenal ulcer and gastric carcinoma in Brazil[J]. J Clin Microbiol, 2003, 41(8): 3964 - 3966.

[2] Huang J Q, Zheng G F, Sumanac K, *et al*. Meta-analysis of the relationship between CagA seropositivity and gastric cancer[J]. Gastroenterology, 2003, 125(6): 1636 - 1644.

[3] Zhou J, Zhang J, Xu C, *et al*. *cagA* genotype and variants in Chinese *Helicobacter pylori* strains and relationship to gastroduodenal diseases[J]. J Med Microbiol, 2004, 53(Pt 3): 231 - 235.

[4] Maeda S, Ogura K, Yoshida H, *et al*. Major virulence factors, VacA and CagA, are commonly positive in *Helicobacter pylori* isolates in Japan[J]. Gut, 1998, 42(3): 338 - 343.

[5] Yamaoka Y, Kodama T, Kashima K, *et al*. Variants of the 3′ region of the *cagA* gene in *Helicobacter pylori* isolates from patients with different *H. pylori*-associated diseases[J]. J Clin Microbiol, 1998, 36(8): 2258 - 2263.

[6] Yamazaki S, Yamakawa A, Ito Y, *et al*. The CagA protein of *Helicobacter pylori* is translocated into epithelial cells and binds to SHP-2 in human gastric mucosa[J]. J Infect Dis, 2003, 187(2): 334 - 337.

[7] Higashi H, Tsutsumi R, Fujita A, *et al*. Biological activity of the *Helicobacter pylori* virulence factor CagA is determined by variation in the tyrosine phosphorylation sites[J]. Proc Natl Acad Sci U S A, 2002, 99(22): 14428 - 14433.

[8] Hoshino F B, Katayama K, Watanabe K, *et al*. Heterogeneity found in the *cagA* gene of *Helicobacter pylori* from Japanese and non-Japanese isolates [J]. J Gastroenterol, 2000, 35(12): 890 - 897.

[9] 彭 颖，熊汉华，黄庆华，等. 幽门螺杆菌全长基因 *cagA* 扩增及限制性酶切片段多态性研究[J]. 中国微生态学杂志，2001，13(3)：127 - 128，134.

[10] Hirata Y, Maeda S, Mitsuno Y, *et al*. *Helicobacter pylori* CagA protein activates serum response element-driven transcription independently of tyrosine phosphorylation[J]. Gastroenterology, 2002, 123(6): 1962 - 1971.

(编辑 薛国文)