

幽门螺杆菌基因组研究

梁 钧* 袁志明 梁布锋

(中国科学院武汉病毒研究所, 武汉 430071)

摘要 幽门螺杆菌 *Helicobacter pylori* (Hp) 是引起人慢性活动性胃炎和消化性溃疡的病原菌, 同时与胃癌的发生密切相关. 近年来 Hp 感染过程和其致病性的分子机理研究受到广泛的关注. 根据其全基因组测序及基因功能解析结果阐述了 Hp 基因组中重要的结构特征, 基因的表达调控方式, 以及毒力相关基因在 Hp 感染中的作用等.

关键词 幽门螺杆菌, 基因组, 基因功能分析

学科分类号 R573

幽门螺杆菌 (*Helicobacter pylori*, Hp) 是一种微需氧, 弯曲呈 S 状或螺旋状, 专性寄生于人胃粘膜上的革兰氏阴性细菌. 1983 年 Warren 等首次发现 Hp 以来, 经胃粘膜组织样品培养及针对性治疗确证, Hp 是引起人慢性活动性胃炎的病原菌和消化性溃疡的重要致病因子, 同时与胃癌的发生密切相关. 另外, 流行病学调查结果显示, 西方国家大约 30%~50% 的人感染 Hp, 在中国, 其感染率更高达 50% 以上. Hp 的高感染率及严重的致病性, 使之成为医学微生物领域研究的热点之一. 目前, Hp 的全基因组测序工作已完成, 通过对全基因组序列信息的分析比较, 有望阐明其适应胃微环境生长的特殊生理机制、生化代谢途径、基因调控机制、与宿主相互作用过程等. 同时也有助于研究者分离到抗 Hp 的新药物靶位, 筛选到 Hp 疫苗的最佳抗原蛋白组分.

1 Hp 全基因组序列的测定

Hp 26695 株和 Hp J99 株的全基因组序列测定工作分别由 Tomb 等和 Gemone Therapeutic 公司完成. 其全基因组序列数据发表在因特网. 26695 株: <http://www.tigr.org>; J99 株: <http://www.astraboston.com/Hppylori>. 从测序的结果来看^[1], Hp 只有一条环状染色体, 大小分别是 26695 株为 1 667 867 bp, J99 株为 1 643 831 bp, 大约是 *E. coli* 的 1/3. 经预测 J99 有 1 495 个开放阅读框架 (ORF), 其中 89 个 ORF 在 26695 株中不存在, 并且 J99 株比 26695 株少 95 个 ORF. 从基因功能分析来看, Hp 全基因组中约 2/3 的基因可以在基因数据库查到功能基因, 并推测其可能的生物学功

能, 1/5 的基因功能未知, 1/4 的基因为 Hp 特有基因.

26695 株基因组平均 C+G 含量为 39%, 但存在 5 个明显的 C+G 含量异常区^[2]. 其中区域 1, 3 内有一个或多拷贝 IS605, 5S rRNA 基因, 另外区域 1 还含有 VirB4 基因, 与 *Agrobacterium tumefaciens* 中编码与 T-DNA 转移有关的蛋白质同源; 区域 2 内有产生 CagA 抗原等毒素蛋白集中排列的 Hp 基因群致病岛 Cag-PAI (Cag pathogenicity island); 区域 4 中存在编码 RNA 聚合酶 β' , β 亚基的 rpoB, rpoC 基因和编码翻译延伸因子 EF-G 的 fusA 基因; 区域 5 内编码两套独特的修饰-限制酶系统.

另外长度为 0.47~3.8 kb 的 8 个重复序列家族在 Hp 染色体上被发现. 其中重复序列 7 存在于基因间隔区域, 重复序列 1, 2, 3, 6 分布于编码外膜蛋白基因中, 两种转座因子 IS605, IS606 以完整拷贝或不完全形式分别出现 13 次和 4 次.

2 Hp 基因组结构特征

根据全基因组测序的结果显示, Hp 基因组具有一些显著的结构特征: a. 特有基因及毒力相关基因等通常在基因组上呈簇连续排列, 如 Cag-PAI; b. 有些基因虽然在染色体上非连续分布, 但编码的蛋白质具有相似的结构域特征并组成重要的功能群, 如 Omp 家族; c. 双核苷酸重复在启动子区域的出现, 显示 Hp 以一种特殊的机制调控基因的表达.

* 通讯联系人.

Tel: 027-87641963, E-mail: jliang@public.wh.hb.cn

收稿日期: 2001-06-26, 接受日期: 2001-07-27

Cag-PAI: 编码毒素蛋白 CagA 基因的上游序列约 40 kb 含 29 个 ORF 的区域被定义为 Hp 的基因群致病岛。Cag-PAI 基因产物是 Hp 感染特异宿主细胞反应所必需, 包括引起宿主蛋白的 Tyr 位点磷酸化, 诱导胃上皮细胞产生白介素 8 (IL-8), 影响宿主细胞表面肌动蛋白叶枕形成过程的细胞骨架重排等^[3]。虽然 Cag-PAI 基因产物引起宿主致病反应过程未知, 但其中 5 个 ORF 和 *Agrobacterium tumefaciens* 中以 type IV 分泌系统机制转移 DNA 的 VirB 蛋白家族同源。根据推测 Cag-PAI 编码的基因可能具有与 type III 分泌系统类似的联系调控分泌形式, 将调控 DNA, 蛋白质影响因子通过特定的通道传递到宿主细胞靶位引起上述反应^[4]。

Omp 家族: 1% 的 Hp 基因组编码 32 种外膜蛋白 (Omp)。Omp 家族的大多数成员在基因数据库中找不到同源基因, 并且各成员相互间具有相似的结构特征^[5]。Omp 组成 Hp 生存所必需且特有的重要功能群。如构成 Hp 细胞膜成分, 具有转运蛋白质和结合蛋白质的功能, 以及作为粘附因子调控与 Lewis^b 血型组抗原的粘附过程。另外这个蛋白质家族中被称为 porins 的蛋白质与 Hp 耐受抗菌素的敏感性密切相关。

双核苷酸重复序列: Hp 中编码细胞表面相关蛋白, 脂多糖生物合成有关酶系, 以及限制与修饰系统相关酶类基因的启动子序列中 CT 或 AG 双核苷酸重复序列大量出现^[2]。它们可能与以滑脱链修补机制 (slipped-strand repair) 开关 Hp 中不同基因表达, 调控表达时相有关。目前虽然还不清楚这种特征序列是否在 Hp 所有基因中发挥类似的功能。但在 Le 抗原中, Le^x 和 Le^y 表达的时相性, 可用 $\alpha 1, 2 \rightarrow \alpha 1, 3$ 墨角藻糖转运酶基因表达活性开关的链滑脱配对机制来解释^[6]。

3 Hp 基因表达调控

为了对外界环境刺激作出反应, 如氮源缺乏、细胞密度、pH 值、温度等。微生物必须具有调控基因表达的能力。Hp 基因组中已证实具有调控作用的蛋白质比 *E. coli* 基因组中的少。说明 Hp 对胃微环境已具有高度的适应性, 并且缺乏来自其他微生物种群的生存竞争压力。除了发现具有与 *E. coli* 相同各种类型的调控基因的形式外, Hp 基因组中还编码几种特有的调控因子。

三个具有螺旋-拐角-螺旋典型结构域的调控蛋白在 Hp 基因组中被发现。两个分别由 Hp1124ORF

和 Hp1349ORF 编码, 另一个由 HspR 基因编码, 在 *Streptomyces coelicolor* 中专性识别 Dnak 基因启动子。另外, Hp 中缺乏典型的热休克蛋白基因转录调控因子 $\sigma 37$ 。但是不意味 Hp 缺少热休克蛋白调控因子, 事实上热休克蛋白 HspR、GroEL、GrpE 被 $\sigma 70$ 调控。同时发现了两种严谨反应调控因子 SpoT 和 CstA 分别调控氨基酸饥饿反应和碳饥饿反应^[2]。

微生物对环境反应需要环境传感器的变化和将信息向细胞内部的基因调控网络传递。两种调控因子家族成员: DNA 结合应答调控因子 (DNA binding response regulator) 和膜组氨酸激酶感受体 (membrane histidine kinase sensor) 在 Hp 基因组中被发现^[2], 完成上述信号传导机制。如 FleR 基因在 Hp 中是鞭毛合成转录激活因子, 同其他调控蛋白一起结合在 $\sigma 54$ 识别的启动子上游作为增强组件激活鞭毛基因的转录。Hp 中的组氨酸激酶感受体蛋白由 AtoS 基因编码, AtoS 以磷酸化方式激活 $\sigma 54$ RNA 聚合酶的转录调控因子 AtoC, 调控短链脂肪酸的代谢。

4 Hp 基因组中的致病因子

Hp 的鞭毛、粘附因子、毒素蛋白三者的协同作用引起胃粘膜细胞的严重致病效应。从基因组测序结果分析至少有 40 个基因与鞭毛的结构体形成、鞭毛蛋白的分泌、组装和调控有关。Hp 感染过程中鞭毛丝复合体嵌入宿主细胞表面, 形成了向宿主内释放毒素蛋白、DNA 影响因子的特定通道。另外 Hp 鞭毛的游动性对形成胃粘膜菌落克隆是必需的, 使 Hp 能够延胃粘膜上皮细胞侵染并建立长期持续感染状态。MotA、MotB 基因编码的整合蛋白是鞭毛的运动所必需, 而 FliG、FliN、FliM、FliY 基因产物形成的鞭毛运动开关复合物, 则介导运动能量传导途径以及与趋化系统信息间的应答。Hp 识别和对最适宜生长环境条件作出反应的感觉适应机制系统即趋化作用 (chemotaxis)^[7], 由 3 种趋化途径作用酶 CheW、CheA、CheY 和 4 种识别特异性配体甲基受体趋化蛋白 Mcps 组成, 它对 Hp 的菌落克隆化形成至关重要。

粘附因子在介导 Hp 与目标细胞的粘附过程中起着关键的作用。目前已证实的三种粘附因子属于 Omp 家族: AlpA、AlpB 基因编码的蛋白质能识别不同的目标细胞表面受体, BabA 基因产物调控参与 Lewis^b 抗原的粘附过程^[8]。最近的研究还发现

一些与 Omp 家族成员结构类似的蛋白质也可以发挥粘附因子的功能。如分离到一种 16 ku 的 Hp 表面蛋白能粘附于寡糖酶配体上, 还有从胃粘膜细胞表面可以分离到一种结合在 PE 型甘油酯上的 63 ku 胞外酶 S 层分子。上述的研究结果显示 Hp 可能运用几种不同的粘附机制成功地粘附于胃上皮细胞上, 并完成其侵染过程。

尽管 Hp 的携带率高, 但 Hp 引起的胃部疾病除了人群中的个体差异外, 主要与 Hp 的重要的毒力因子是否存在有关。这些毒力因子包括: 尿素酶、空泡毒素 (VacA 编码)、细胞毒素相关蛋白 A (CagA 编码)、热休克蛋白、超氧化物歧化酶和粘附蛋白等, 其中尿素酶、空泡毒素和细胞毒素相关蛋白 A 是引起胃严重病患的主要因素^[9]。VacA 蛋白的致病效应主要是通过导致靶细胞的离子转运功能失常, 从而引起靶细胞出现空泡变性等病理改变, 直至发生不可逆的细胞损伤。同时, Hp 大量产生的尿素酶, 通过分解内源性尿素产生的氨可以单独诱导或促进空泡毒素的活性, 从而扩大空泡的变性范围和数量。细胞毒素相关蛋白 A 是 VacA 的一个协同表达因子, 同时也是 Hp 菌株具有高毒力的标志。CagA 蛋白和 Cag-PAI 中的 PicB 蛋白联系作用, 诱导胃上皮细胞产生 IL-1 β 、IL-6、TNF- α 、IL-8 等细胞因子。其中 IL-8 诱导中性粒细胞和淋

巴细胞的趋化性, IL-6 及 TNF- α 则诱导单核细胞和多核巨细胞浸润。上述炎症反应细胞释放的各种蛋白酶及胶原酶是导致胃组织损伤的重要因素。

参 考 文 献

- 1 Alm R A, Lee L S, Moir B L, *et al.* Genomic sequence comparison of two unrelated isolates of the human gastric pathogen *Helicobacter pylori*. *Nature*, 1999, **397** (6715): 176~180
- 2 Tomb J F, White O, Kerlavage A R, *et al.* The complete genome sequence of the gastric pathogen *Helicobacter pylori*. *Nature*, 1997, **388** (6642): 539~547
- 3 Segal E D, Lange C A. Induction of host signal transduction pathways by *Helicobacter pylori*. *Proc Natl Acad Sci USA*, 1997, **94** (14): 7595~7599
- 4 Figura N, Valassina M. *Helicobacter pylori* determinants of pathogenicity. *J Chemother*, 1999, **11** (6): 591~600
- 5 Alm R A, Bina J, Andrews B M, *et al.* Comparative genomics of *Helicobacter pylori*: analysis of the outer membrane protein families. *Infect Immun*, 2000, **68** (7): 4155~4168
- 6 Martin S M, Edbrooke T, Hodgman D, *et al.* Lewis X biosynthesis in *Helicobacter pylori*. Molecular cloning of an α (1, 3)-fucosyltransferase gene. *J Biol Chem*, 1997, **272** (34): 21349~21356
- 7 Yoshiyama H, Nakamura H, Kimoto M, *et al.* Chemotaxis and motility of *Helicobacter pylori* in a viscous environment. *J Gastroenterol*, 1999, **34** Suppl (11): 18~23
- 8 Ilver D A, Arnqvist A, Ogren J, *et al.* *Helicobacter pylori* adhesin binding fucosylated histo-blood group antigens revealed by retagging. *Science*, 1998, **279** (5349): 373~377
- 9 Dundon W G, Bernard M, Montecucco C. Virulence factors of *Helicobacter pylori*. *J Med Microbiol*, 2001, **290** (8): 647~658

Progress on *Helicobacter pylori* Genome

LIANG Jun*, YUAN Zhi-Ming, LIANG Bu-Feng

(Wuhan Institute of Virology, The Chinese Academy of Sciences, Wuhan 430071, China)

Abstract *Helicobacter pylori* (Hp) causes gastritis and peptic ulcer disease of a human being, and is associated with certain types of gastric cancer. In recent years, much attention has been focused on its molecular mechanisms of *H. pylori* infection. As an important human pathogen, Hp genome has been sequenced. A detail summary of the literature, significant genome structure feature, gene expression regulation, and identification of virulence factors is provided.

Key words *Helicobacter pylori*, genome, gene function analysis

* Corresponding author. Tel: 86-27-87641963, E-mail: jliang@public.wh.hb.cn

Received: June 26, 2001 Accepted: July 27, 2001