

· 实验室研究 ·

一株未知黏附基因型的尿道致病性
大肠埃希菌 UPEC4030 株的基因分析

吴庆刚 张敬平 陈锦英 郑明寰 尤凤兴 姜永

【摘要】 目的 比较尿道致病性大肠埃希菌 UPEC4030 株 *papG* 和 *papA* 基因与相关序列的差异,以明确 UPEC4030 *papA* 基因变异情况及与 *papG* 的组合形式,确定是否为未知基因型。方法 采用基因克隆测序方法,分别对 UPEC4030 株 *papG* 和 *papA* 基因进行序列分析,并与相关序列进行比较。结果 UPEC4030 *papA* 由 722 个碱基对组成,编码 192 个氨基酸的多肽。与 10 个基因型 UPEC *papA* 核苷酸序列和推导的氨基酸序列同源性比较,分别为 36.11%~77.95% 和 22.20%~78.34%;与 *papA* 基因分型多重 PCR 引物序列比较,下游分型引物在 UPEC4030 *papA* 相应位置上的序列同源性只有 10.00%~66.67%,表明该菌株 *papA* 确实为未知的基因型。UPEC4030 菌株 *papG* 全基因片段由 1100 个碱基组成,编码 337 个氨基酸的多肽,其核苷酸及推导的氨基酸序列与 II 型 *papG* 代表株 UPEC IA2 相比较,同源性分别为 99.00% 和 99.11%,表明 UPEC4030 *papG* 属于 II 型。结论 UPEC4030 *papA* 为未知的基因型,其 *papG* 属于 II 型。

【关键词】 尿道致病性大肠埃希菌; 基因测序; 未知基因型

Genomic analysis of an unknown *pap* genotype from uropathogenic *Escherichia coli* 4030 strain WU Qing-gang*, ZHANG Jing-ping, CHEN Jin-ying, ZHENG Ming-huan, YOU Feng-xing, JIANG Yong. *Wuxi Center for Disease Control and Prevention, Wuxi 214023, China

【Abstract】 Objective To investigate the differences between the sequences of *papA* and *papG* of UPEC4030 strain and the related genes, to better understand the genetic variation of UPEC4030 *papA* and its combination forms with *papG* so as to identify if it was a new genotype. Methods Cloning and sequencing methods were used to analyze the sequences of *papG* and *papA* of UPEC4030 strain and to compare their related sequences. Results Through sequence analysis of *papA*, it was revealed that there was a 722 bp gene, encoding 192 amino acid polypeptide. The overall homology of the *papA* genes between UPEC4030 and the standard strains of ten F types were 36.11%-77.95% and 22.20%-78.34% at nucleotide and deduced amino acid levels. Homology between the sequences of reverse primers and the corresponding sequence of UPEC4030 *papA* was 10.00%-66.67%. The results confirmed that UPEC4030 strain contained a novel *papA* variant. Through sequence analysis of UPEC4030 *papG*, we revealed a 1100 bp gene, encoding 337 amino acid polypeptide. The homology of the *papG* genes between UPEC4030 and UPEC IA2, the standard strain, was 99.00% at nucleotide level and 99.11% at deduced amino acid and UPEC4030 strain carried class II genotype of *papG*. Conclusion UPEC4030 strain contained an unknown *papA* variant or the new genotype and carried class II genotype of *papG*. The pathogenic mechanism and epidemiology call for further study.

【Key words】 Uropathogenic *Escherichia coli*; Gene sequencing; Unkown genotype

大肠埃希菌为尿路感染的主要致病菌,从尿路感染者尿液中分离的大肠埃希菌 80% 以上具有 P 菌毛,称为尿道致病性大肠埃希菌(Uropathogenic *Escherichia coli*, UPEC)^[1]。P 菌毛与尿道上皮细胞

的特异性黏附是 UPEC 致病的基础,也是肾盂肾炎发病的关键。近年来,主要菌毛蛋白亚单位 PapA 与次要菌毛蛋白亚单位 PapG 因在 P 菌毛组成和黏附中的主导作用而成为研究的焦点,尤其 *papA* 与 *papG* 基因分型的研究在 UPEC 流行病学、致病机制、种系关系和变异及免疫防治研究中的作用引起人们极大的兴趣^[2-4]。本研究对国内 45 株 UPEC 的 *papA* 和 *papG* 基因分型,采用 PCR 克隆技术对 2 种基因进行克隆和序列分析。

基金项目:卫生部科研基金资助项目(WKJ2006-2-10)

作者单位:214023 无锡市疾病预防控制中心(吴庆刚、张敬平、尤凤兴、姜永);天津医科大学基础医学院(陈锦英);中国疾病预防控制中心传染病预防控制所(郑明寰)

材料与方法

1. 菌株、质粒和培养基: UPEC4030 菌株由天津医科大学总医院临床尿路感染患者尿液标本分离, 经常规细菌学鉴定为大肠埃希菌, MRHA 血凝试验和 *papC*、*papG* PCR 法检测为阳性, 确定为 P 菌株, *papA* 多重 PCR 基因分型方法不能分型。UPEC132 为天津地区临床尿路感染患者尿液标本分离所得, 经血凝、血清学、免疫电泳、*papA* 和 *papG* 基因序列克隆测序和分析, P 菌毛为 F13 型 PapG 为 II 型。IA2 为 II 型 *papG* 代表株, UPECJ96 为 F13 型 P 菌毛代表株及 *E. coli* K-12 p678-54 为无菌毛代表株。质粒载体 pMD18-T Vector 购自 TaKaRa 公司。实验中使用 LB、SOB、SOC 液体和固体培养基液体及固体培养基。

2. 试剂和仪器: DNA 提取应用 genomic DNA isolation kit (博大泰克公司), PCR 引物由上海 Sangon 公司合成, dNTP、Taq DNA 聚合酶、限制性内切酶和 DNA 标准分子质量参照购自 TaKaRa 公司。DNA 序列分别由上海 Sangon 和上海博亚生物技术公司完成。PCR 扩增仪为 PE2400。基因分析计算机软件为 Primer 5.0。

3. UPEC4030 菌株 *papA* 基因的克隆和序列分析: 根据已发表 *papA* 基因序列设计引物^[3]。为了定向克隆的需要在二引物的 5' 端加上限制性内切酶 EcoR I 和 BamH I 的识别序列, 引物序列为 5'-CGG GAA TTC GAT AAA TAA CCT GCC CTG A 和 5'-CCG GAT CCG GAT TAT CAG TAT TAC TGA。基因克隆、质粒提取和感受态细胞制备与重组质粒的酶切鉴定按常规方法进行^[5]。由上海 Sangon 公司完成重组质粒上插入片段 UPEC4030 *papA* 序列测序。

4. UPEC4030 菌株 *papG* 基因的克隆和序列分析: 根据发表的 II 型 *papG* 序列设计引物^[6], 上下游引物序列分别为 5'-ATG AGC ATG ATT TAT AAC TGA 和 5'-ACA ACA CTC TAT CCG GCT CC。PCR 产物末端加 A 后与 T 载体连接, 转化感受态细胞, 经酶切鉴定正确后, 由上海博亚生物技术公司完成 UPEC4030 *papG* 序列测序。

结 果

1. UPEC4030 菌株 *papA* 基因的克隆和序列分析: 用 PCR 方法扩增 UPEC4030 株 *papA* 基因,

经酶切、连接转化、筛选, 共得到 3 个白色阳性菌落, 经 PCR 鉴定, 其中 3 个阳性菌落均可得到 700~750 bp 的插入片段。摇菌提取质粒经 EcoR I 和 BamH I 双酶切, 可得到 2.69 kb 的载体 DNA 带和约 700~750 bp 大小的插入片段。阳性克隆子测序结果表明, UPEC4030 株 *papA* 由 722 个碱基对组成, 具有典型的原核基因结构特征(图 1)。起始密码子 ATG 位于碱基的 130~132 处, 终止密码子 TAA 位于碱基 706~708 处, TAA 稍后又紧跟着 2 个终止密码子读码框 TGA 和 TAA, 从起始密码子到终止密码子的开放阅读框(ORF)全长为 579 bp, 共翻译 192 个氨基酸组成的蛋白质。在 ATG 上游 16 碱基起始由 -AAA GAG GT-组成公认的核糖体结合位点(即 S-D 序列)。推导信号肽序列由成熟肽的 NH₂-端丙氨酸上游 21 个密码子组成, 特征性正电荷赖氨酸残基位于 -19 位。*papA* 信号肽序列的 -1, -3 位切割三联体为 ANV。UPEC4030 株的 *papA* 序列分析符合已有报道的规律, 由核苷酸推导氨基酸的组成为四个部分, 即 NH₂-端(R₁-R₂₁), Cys-Cys 环区(R₂₂-R₆₁), 可变区(R₆₂-R₁₅₉)和 COOH-端(R₁₆₀-R₁₇₁)。其中 NH₂ 端、COOH 端和信号肽序列为遗传高度保守区域。与 F13 型血清型代表株 UPECJ96 和 UPEC132 相比较, UPEC4030 *papA* 核苷酸序列在 Cys-Cys 环区和可变区差别明显。UPEC4030、UPECJ96、UPEC132 *papA* 基因编码氨基酸残基数分别为 192、185、184 个。三者起始密码子、终止密码子和 S-D 序列均相同, 而信号肽序列 UPECJ96 株由 22 个密码子组成, UPEC132 和 UPEC4030 株由 21 个密码子组成, 是因为在 -3 位缺失编码 Asn 的密码子 AAT, 导致 *papA* 信号肽序列为 -1, -3 位切割三联体, UPECJ96 株为 ANN, UPEC132 株和 UPEC4030 株为 ANV。此外 UPEC4030 株 Cys-Cys 环区和 UPECJ96、UPEC132 编码氨基酸数相同, 而在可变区比后两者多编码 6 个氨基酸, 位于 UPEC4030 肽链的 +148~+153 处, -COOH 端也多编码 3 个氨基酸(+163~+165), 见图 1。

2. UPEC4030 株 *papA* 序列与 *papA* 基因分型多重 PCR 引物序列的比较^[3]: 应用 Primer 5.0 软件将 11 个基因型 *papA* 基因的上游通用引物 Ff 和下游分型引物进行比对。在 UPEC4030 *papA* 基因序列的 165~184 bp 处可找到与 Ff 完全相同的序列, 同源率为 100%; F7-1r、F7-2r、F8r、F9r、F10r、F11r、

F12r、F13r、F14r、F15r、F12/15r、F16r 12 个基因型下游分型引物在 UPEC4030 *papA* 相应位置上的序列同源性分别为 30.00%、66.67%、40.00%、31.58%、25.00%、25.00%、30.00%、21.74%、40.91%、10.00%、21.74% 和 15.00%，即 10.00%~66.67% 之间。由此可以解释前述 UPEC4030 株 *papA* 多重 PCR 分型结果阴性的原因所在。

个基因型 UPEC 代表株的 *papA* 核苷酸序列比较发现，其间有大量的插入或缺失碱基序列，因此推导的氨基酸序列也有差异。

表1 UPEC4030 *papA* 与其他基因型 UPEC *papA* 核苷酸序列和推导氨基酸序列同源性(%)比较

UPEC <i>papA</i>	核苷酸序列同源性	推导氨基酸序列同源性
F7-2	77.95	78.34
F10	55.56	46.12
F14	50.40	49.44
F13	49.62	46.11
F8	49.26	46.01
F7-1	49.07	46.11
F11	47.38	45.56
F15	40.19	23.89
F9	39.17	21.11
F16	36.11	22.20

注：F12 为 UPEC Prs 菌株基因型，严格讲不属于 P 菌毛范围，未在表中显示。

paI
1. TCGATAAATAACCTGCCCTG

1. AAAATACGAGAATATTATTTGTATGTATCTGGTATTAAAGGTAATCGGGTCATT
2. M I K S V I A G A V A M A V V S

1. TTAATTTGCCAGATATCTCTGGTGTGTTTCAGTAATGAAAGAGGTGTGTTATTT
2. F G V N A A P T I P Q G Q G K

1. ATG ATT AAG TCG GTT ATC GCC GGT GCG GTA GCT ATG GCA GTG GTG TCT
2. M I K S V I A G A V A M A V V S

1. TTT GGT GTA AAT GCT GCT CCA ACT ATT CCA CAG GGG CAG GGT AAG
2. F G V N A A P T I P Q G Q G K

1. GTA ACT TTT AGC GGA ACT GTG GTT GAT GCA CCA TGT GGT ATT GAT GCG
2. V T F S G T V V D A P C G I S A

1. CAG TCA GCT GAT GAC TCT GTC GAT TTT GGC CAG ATA TCA AAA GTC TTT
2. Q S A D Q S V D F G Q I S K V F

1. CTG GAC AAT GAT GGG CAA ACC ACA CCT AAG GCT TTT GAT ATT AAG CTC
2. L D N D G Q T T P K A F D I K L

1. GTG AAT TGC GAT ATT ACT AAC TAT AAG ATA GCA GTC GCT GGT GGT GGC
2. V N C D I T N Y K I A V A G G G

1. GGC GCA CCT AAA GCA GGA ACA GTA TCC ATG ACA TTC TCC GGT GTT ACA
2. G A P K A G T V S M T F S G V T

1. GCA GGT AAT ACC CAT GAC ATG TTA CAG ACG AAT GGC GAG ACC AAT ACA
2. A G N T H D M L Q T N G E T N T

1. GCT ATC GTT GTT ATC CAG GCG CAT GGC AAA CGT GTG AAG TTT GAT GGT
2. A I V I D R H G K R V K F D G

1. CAG ACC TCT ACC GAG GTT TCA AAT CTG ATT AAT GGT GAT AAC ACG ATT
2. Q T S T E V S N L I N G D N T I

1. CAT TTC ACT GCT GCA GTT ATG AAA GAT GGT AAT GGC GAC CCT GTA AAA
2. H F T A A V M K D G N G D P V K

1. GAG GGC GCC TTC TCA GCT GTC GCG AAT TTC AAC CTT ACT TAT CAG TAT
2. E G A F S A V A N F N L T Y Q Y

1. CAG TAATACTGATAATCCGG 722
2. Q

注：1: UPEC4030 株的 *papA* 基因序列；2: UPEC4030 株的 *PapA* 蛋白的氨基酸序列；paI 与 par 为引物序列

图1 UPEC4030 *papA* 基因序列和推导氨基酸序列

3. UPEC4030 *papA* 与其他基因型 UPEC *papA* 核苷酸序列和推导的氨基酸序列同源性比较：在 Primer 5.0 基因分析软件辅助下，将 UPEC4030 *papA* 与 GenBank 登录的 10 个基因型 UPEC *papA* 核苷酸序列比较，同源性为 36.11%~77.95%，从高到低依次为 F7-2、F10、F14、F13、F8、F7-1、F11、F15、F9 和 F16；根据核苷酸序列推导的氨基酸序列比较，同源性为 22.20%~78.34%，从高到低依次为 F7-2、F14、F10、F13 与 F7-1、F8、F11、F15、F16 和 F9(表 1)。另外，从起始密码 ATG(130~132)下游 31~33 处的密码子 ATG 为起点 (*papA* 多重 PCR 基因分型的上游通用引物 Ff 起点处)，UPEC4030 *papA* 与 10

4. UPEC4030 菌株 *papG* 基因的克隆和序列分析：用 PCR 方法扩增 UPEC4030 株 *papG* 基因，经纯化回收、连接转化、筛选鉴定、阳性克隆子测序，UPEC4030 菌株 *papG* 全基因片段由 1100 个碱基组成，含有 337 个密码子构成的 ORF。UPEC4030 *papG* 核苷酸及推导的氨基酸序列与 II 型 *papG* 代表株 UPEC IA2 相比较，其同源性分别为 99.00% 和 99.11%，表明 UPEC4030 *papG* 属于 II 型。起始密码子 ATG 位于序列的 65~67 处，终止密码子 TAA 位于 1076~1078 处。在 ATG 上游 14~18 bp 由 -GGAAG- 组成特征性核糖体结合位点 (S-D 序列)。推导信号肽序列由成熟肽的 NH₂-端色氨酸上游 20 个密码子组成，信号肽切割位点位于 -1 位丙氨酸和 +1 位色氨酸之间。该序列推导的氨基酸组成具有 *PapG* 共有的特征：① 4 个半胱氨酸 Cys45、Cys119、Cys198 和 Cys230 构成 2 个 Cys-Cys 环区；② 142 位为保守的天冬氨酸残基 (Asp142)，包括 Asp142 在内的近 NH₂-端 2/3 区域为特征性受体结合区，其中心部位为 Cys45-Cys119 构成的环区；③ 近 COOH-端 Cys198 和 Cys230 之间为典型的 31 个氨基酸的残基，这些部位构成 *PapD* 蛋白的作用区。与 II 型 *papG* 代表株 UPEC IA2 的 *papG* 相比，UPEC4030 *papG* 在 R₅₂、R₆₃、R₁₂₁、R₁₃₂、R₁₈₁ 和 R₂₈₉ 存在 6 处同义突变，在特征性受体结合区和 *PapD* 蛋白的作用区分别存在 1 处和 2 处错义突变，即 +57 位 TAG→TCC (Thy→Ser) 和 +250 位 CAT→CTT (His→Leu)、+380 位 GGT→AGT (Gly→Ser)，见图 2。

		引物序列	S-D 序列
IA2		ATGACATGATTATAACTEAGTCATACCTAAATGAATAACTGTAATACGGAGCTGATTTCTG	
4030	
IA2	65	ATG AAA AAA TGG TTC CCA GCT TTG TTA TTT TCC TTG TGT GTG TCT GGT GAG TCC TCT GCA TGG AAT CAC AAT ATT GTC TTT TAC TCC CTT	-20
4030	
AA		M K K W F P A L L F S L C V S G E S S A W N H N J V F Y S L	-10
IA2	155	GGA AAC GTT AAC TCT TAT CAG GGA GGG AAT GTG GTG ATT ACT CAA AGG CCA CAA TTT ATA ACT TCG TGG CGC COG GGC ATT GCT ACG GTA	-10
4030	
AA		G N V N S Y Q G G N V V I T Q R P Q F I T S W R P G I A T V	-10
IA2	245	ACC TGG AAT CAG TGT AAT GGT CCT GAG TTC GCT GAT GGC TCC TGG GCT TAC TAC AGG GAG TAT ATT GCG TGG GTA GTA TTC CCC AAA AAG	+20
4030	
AA		T W N Q C N G P E F A D G S W A Y Y(S) R E Y I A W V V F P K K	+10
IA2	335	GTT ATG ACC CAA AAT GGA TAT CCC TTA TTT ATT GAG GTT CAT AAT AAA GGT AGC TGG ACT GAG GAG AAT ACT GGT GAC AAT GAC AGC TAT	-50
4030	
AA		V M T Q N G Y P L F I E V H N K G S W S E E N T G D N D S Y	-70
IA2	425	TTT TTT CTC AAG GGG TAT AAG TGG GAT GAG CGG GCC TTT GAT GCA GGT AAT TTC TGT CAG AAA CCA GGA GAA ACA ACT CGT CTG ACT CAG	+80
4030	
AA		F F L K G Y K W D E R A F D A G N L C Q K P Q E T T R L T E	+100
IA2	515	AAA TTT GAC GAT ATT ATT TTT AAA GTC GCC TTA CCT GCA GAT CTT CCT TTA GGG GAT TAT TCT GTT ACA ATT CCA TAC ACT TCC GGC ATA	+110
4030	
AA		K F D D I I F K V A L P A D L P L G D Y S V T I P Y T S G I	+120
IA2	605	CAG CGT CAT TTC GCG AGT TAC TTG GGG GCC CGT TTT AAA ATC CCA TAC AAT GTG GCC AAA ACC CTC CCA AGA GAG AAT GAA ATG TTA TTC	+140
4030	
AA		Q R H F A S Y L G A R F K I P Y N V A K T L P R E N E M L F	+150
IA2	695	TTA TTT AAG AAT ATC GGC GGA TCC CGT CCT TCT GCA CAG TCT CTG GAA ATA AAG CAT GGT GAT CTG TCT ATT AAT AGC GCT AAT AAT CAT	+170
4030	
AA		L F K N I G G C R P S A Q S L E I K H G D L S I N S A N N H	+180
IA2	785	TAT GCG GCT CAG ACT CTT TCT GTG TCT TGC GAT GTG CCT GCA AAT ATT CGT TTT ATG CTG TTA AGA AAT ACA ACT CCG ACA TAC AGC CAT	+190
4030	
AA		Y A A Q T L S V S C D V P A N I R F M L L R N T T P T Y S H(L)	+200
IA2	875	GGT AAG AAA TTT TCG GTT GGT CTG GGG CAT GGC TGG GAC TCC ATT GTT TCA GTT AAC GGG GTG GAC GCA GGA GAG ACA ACG ATG AGA TGG	+220
4030	
AA		G K K F S V G L G H G W D S I V S V N G V D T G E T T M R W	+230
IA2	965	TAC AAA GCA GGT ACA CAA AAC CTG ACC ATC GGC AGT CGC CTC TAT GGT GAA TCT TCA AAG ATA CAA CCA GGA GTA CTA TCT GGT TCA GCA	+250
4030	
AA		Y K A G T Q N L T I G S R L Y G E S S K I Q P G V L S G(S) S A	+260
IA2	1055	ACG CTG CTC ATG ATA TTG CCA TAA ATG GTT	+280
4030	
AA		T L L M I L P ***	+290
IA2	1085	TATCCGGAGCCGATAGACTCT(1100)	+300
		+310

注: IA2:UPEC IA2 株的 *papG* 基因序列; 4030:UPEC4030 株的 *papG* 基因序列(.....表示与 IA2 株碱基相同); AA:UPEC IA2 株的 *PapG* 蛋白的氨基酸序列,()表示 UPEC4030 株的 *PapG* 蛋白的氨基酸序列与 UPEC IA2 株的差异; * * *:终止密码子

图2 UPEC4030 *papG* 基因序列与 II 型 *papG* 的比较

讨 论

UPEC 是尿道感染的主要致病菌,其中 P 菌毛的黏附是致病的关键^[1,2,6]。本课题组在对国内分离的 45 株 UPEC 菌株进行 *papG* 和 *papA* 的多重 PCR 基因分型研究中发现,所有 45 株 UPEC *papG* 分型均为 II 型,44 株 UPEC *papA* 分型有结果,UPEC4030 *papA* 基因用多重 PCR 方法不能分型。基于以上研究基础,为了探讨 UPEC4030 株 *papA* 基因是否存在变异及变异程度,参考已发表的 *papA* 序列,设计合成可扩增 *papA* 全基因的引物,用于 UPEC4030 *papA* 序列的扩增。应用 PCR 定向克隆方法成功克隆了 UPEC4030 的 *papA* 基因,并进行序列分析。结果表明其 *papA* 由 722 个碱基

对组成,基因结构具有典型的原核基因结构特征,共翻译 192 个氨基酸组成的蛋白质。将 UPEC4030 *papA* 与 *papA* 基因分型多重 PCR 引物序列比较,在其 165~184 bp 处可找到与上游通用引物完全相同的序列,下游分型引物在 UPEC4030 *papA* 相应位置上的序列同源性只有 10.00%~66.67%。这一方面证实 UPEC4030 *papA* 应用多重 PCR 分型方法确实不能分型,另一方面也证实了分型方法的可靠性。比较 UPEC4030 *papA* 与其他基因型 UPEC *papA* 核苷酸序列和推导的氨基酸序列同源性分别为 36.11%~77.95% 和 22.20%~78.34%。编码免疫原性表位的 Cys-Cys 环区和编码抗原性表位的可变区碱基差异明显,尤其在可变区的末端由于核苷酸的插入导致 UPEC4030 *papA* 肽链长度有所改

变,但未发现移码突变,这种改变难以用 PCR 扩增中 Taq DNA 聚合酶可能出现的碱基错掺来解释^[7]。根据以上结果,可初步确定 UPEC4030 *papA* 属于未知基因型,即新的基因型。

业已证明,C 区疏水性的稳定是信号肽酶准确切割的依据^[8]。本研究组在研究 UPEC132 与同基因型代表株 UPECJ96 的 PapA 相对分子质量(M_r)时发现,2 菌株 PapA 分子质量相差 3000,认为其差异可能与 UPEC132 PapA 的信号肽编码基因变异造成不准确切割有关。UPEC4030 *papA* 信号肽识别窗口与 UPEC132 相同,-3 位也缺失了三联密码 AAT,即-3 位的切割位点由天冬酰胺残基变为缬氨酸残基。由于天冬酰胺为极性中性氨基酸,缬氨酸为非极性疏水性氨基酸,C 区的疏水性改变,这样 UPEC4030 *papA* 信号肽酶识别及剪切是否出现类似 UPEC132 的现象,仍需进一步研究。诚然,相关研究对明确保守的信号肽酶识别窗口序列缺失或突变时信号肽切割的共同规律无疑具有重要意义。

papG 作为真正的黏附素,介导与尿道上皮细胞受体 Gal-Gal 的特异性结合,本研究应用 *papG* PCR 分型方法已证明 UPEC4030 株 *papG* 为 II 型。UPEC4030 *papG* 核苷酸及推导的氨基酸序列与 II 型 *papG* 代表株 UPEC IA2 相比较,其同源性分别为 99.00% 和 99.11%。参考其他 44 株 UPEC *papG* 的基因分型结果,说明 UPEC 菌株在进化过程中 *papG* 高度保守。这与国外研究结果相一致^[6,9]。

致病岛(pathogenicity island, PAI)编码的基因产物常是分泌性蛋白和细胞表面蛋白,如菌毛、溶血素或干扰和调节宿主细胞信号传导系统的蛋白。PAI 最早是用来描述 UPEC 染色体上 2 个分子质量很大、编码许多毒力相关基因的不稳定的外源 DNA 片段,已报道 3 株 UPEC(536、J96 和 CFT073)的 12 个 PAIs^[10,11]。根据 UPEC4030 具有黏附基因群,暗示其必有 PAI 存在。UPEC4030 株的研究尚处于初级阶段,对其 PAI 和与宿主细胞相互作用的研究将对 UPEC 致病机制有更深入的了解,也会对抗感染

治疗的策略和抗菌药物的设计提供新的思路。随着细菌毒力基因鉴定方法的进展,寻求细菌新的致病基因变得简便快捷。这无疑给下一步探讨 UPEC 的 PAI,揭示 UPEC 与宿主细胞相互作用的分子机制,为细菌 P 菌毛黏附上皮细胞后续致病过程与细胞内信号传导途径的关系指明了方向。

参 考 文 献

- [1] Struve C, Krogfeldt KA. In vivo detection of *Escherichia coli* type 1 fimbrial expression and phase variation during experimental urinary tract infection. *Microbiology*, 1999, 145 (Pt 10): 2683-2690.
- [2] Girón JA, Torres AG, Freer E, et al. The flagella of enteropathogenic *Escherichia coli* mediate adherence to epithelial cells. *Mol Microbiol*, 2002, 44(3): 361-379.
- [3] Johnson JR, Stell AL, Scheutz F, et al. Analysis of the F antigen-specific *papA* alleles of extraintestinal pathogenic *Escherichia coli* using a novel multiplex PCR-based assay. *Infect Immun*, 2000, 68 (3): 1587-1599.
- [4] Lane MC, Simms AN, Mobley HL, et al. Complex interplay between type 1 fimbrial expression and flagellum-mediated motility of uropathogenic *Escherichia coli*. *J Bacteriol*, 2007, 189 (15): 5523-5533.
- [5] Joseph S, David WR, 黄培堂, 等. 分子克隆实验指南. 3 版. 北京: 科学出版社, 2002: 96-99.
- [6] Hagan EC, Mobley HL. Uropathogenic *Escherichia coli*. outer membrane antigens expressed during urinary tract infection. *Infect Immun*, 2007, 75(8): 3941-3949.
- [7] 金冬雁, 金奇, 侯云德. 核酸和蛋白质的化学合成与序列分析. 北京: 科学出版社, 1996: 134-136.
- [8] Hegde RS, Lingappa VR. Membrane protein biogenesis: regulated complexity at the endoplasmic reticulum. *Cell*, 1997, 91(5): 575-582.
- [9] Johnson JR, Brown JJ. A novel multiply primed polymerase chain reaction assay for identification of variant *papG* genes encoding the gal(α 1-4) gal-binding *papG* adhesions of *Escherichia coli*. *J Infect Dis*, 1996, 173(4): 920-926.
- [10] 徐建国. 分子细菌学. 北京: 科学出版社, 2000: 197-200.
- [11] Groisman EA, Ochman H. Pathogenicity islands: bacterial evolution in quantum leaps. *Cell*, 1996, 87(11): 791-794.

(收稿日期: 2008-06-17)

(本文编辑: 尹廉)