



尿路致病性大肠杆菌中 I 型限制-修饰系统 C5423-5425 的鉴定

李火明, 夏颖, 李干武*, 蔡文通*

中国农业科学院哈尔滨兽医研究所, 兽医生物技术国家重点实验室, 黑龙江 哈尔滨 150069

摘要:【目的】多方面鉴定尿路致病性大肠杆菌中的 I 型限制-修饰(restriction-modification, RM)系统 C5423-5425, 并揭示其功能与生理意义。【方法】利用 Lambda Red 重组系统构建 CFT073 DNA 甲基化转移酶基因缺失株 $\Delta c5424$; 利用单分子实时测序分析以获得 C5424 甲基化修饰的位点与基序; 利用转化效率试验说明 RM 系统抵御外源 DNA 的转化; 利用转录组测序和荧光定量 RT-PCR 鉴定 C5424 的调控基因; 利用软琼脂平板运动试验等方法研究细菌生理功能的变化。【结果】用生物信息学方法鉴定了一个 I 型 RM 系统; 获得了 $c5424$ 的突变株; 鉴定了 C5424 修饰的靶序列 $G^mAGNNNNNNNGTCA/TG^mACNNNNNNNCTC$, 并揭示了靶序列在基因组中的分布; C5424 系统可以抵御外源 DNA 的进入; $c5424$ 的缺失显著影响 17 个基因的表达, 包括运动相关基因 *motB* 和抗活性氯基因 *rclR* 等; $c5424$ 的缺失显著影响 CFT073 的运动能力和抗次氯酸的能力。【结论】本文从多方面鉴定了尿路致病性大肠杆菌 CFT073 中的 I 型 RM 系统 C5423-5425, 这个系统对细菌具有重要的生理意义。本文对研究细菌的表观遗传学具有参考价值。

关键词: 尿路致病性大肠杆菌, DNA 甲基化转移酶, 单分子实时测序, 限制-修饰系统

在细菌中, 限制-修饰(restriction-modification, RM)系统是普遍存在的, 常被认为是细菌最原始的免疫系统, 用来抵御外来的 DNA, 尤其是噬菌体^[1-2]。它们通常由两部分组成: 核酸内切酶, 负责位点特异性的 DNA 切割; 甲基化转移酶, 可以甲基化 DNA, 其修饰时的识别位点通常与内切酶切割的识别位点一致。根据亚基组成、识别靶序列特征、切割位点特征和辅因子需求等, RM 系统

被划分为 I (如 *EcoK* I)、II (如 *EcoR* I)、III (如 *EcoP15* I)和IV (如 *Mcr* 和 *Mrr*) 4 个不同的类型^[3]。I 型 RM 系统通常由 HsdS (S 亚基)、HsdM (M 亚基)、HsdR (R 亚基) 3 个亚基组成, 形成催化限制性内切和修饰的五聚体蛋白 R_2M_2S ^[4]。其中, S 亚基特异识别 DNA 序列, R 亚基发挥限制性酶切的作用, M 亚基催化甲基化转移^[5]。三聚体 M_2S 作为 DNA 甲基化转移酶可以独立存在。*hsdR* 和

基金项目: 国家自然科学基金(31902242); 国家重点研发计划(2018YFD0500500)

*通信作者。Tel/Fax: +86-451-51051728; E-mail: 李干武, liganwu@caas.cn, 蔡文通, caiwentong@caas.cn

收稿日期: 2020-09-20; 修回日期: 2021-01-14; 网络出版日期: 2021-03-12

hsdM 基因高度保守, 而 *hsdS* 则高度可变, 导致特异性不同。I 型 RM 的识别序列包括两部分, 通常被 N_{5-8} 隔开, 比如 *EcoK* I 的识别序列为 $A^mACNNNNNNGTGC/GC^mACNNNNNNGTT$ (互为反向重复序列), 两部分中的腺嘌呤都会被甲基化。这种识别方式是因为 S 亚基含有 2 个靶位点识别结构域(TRD)。很多 I 型系统含有重复的 *hsdS* 基因, 2 个 *hsdS* 可以通过反向重复序列发生翻转, 导致甲基化特异性的变化, 产生“相变”(phase variation) 现象^[6]。另外, 在多种病原菌有报道发现 I 型 RM 系统可以调控毒力基因的表达, 进而影响毒力^[7-8]。

尿路致病性大肠杆菌(uropathogenic *Escherichia coli*, UPEC)是导致尿路感染的主要病原细菌, 感染后会导致肾盂肾炎、膀胱炎和尿道炎等疾病, 严重影响人们的健康^[9]。UPEC 编码多种毒力因子, 包括负责黏附和定殖的菌毛、负责侵袭和运动的鞭毛、负责调控宿主免疫反应的溶血素、负责摄取铁离子的摄铁蛋白等^[9-10]。毒力调控使得病原菌能够适应宿主环境并适当地表达毒力因子以建立感染^[11]。鉴于 I 型 RM 系统的多样性和毒力调控作用, 本文鉴定了 UPEC CFT073 菌株中的一个 I 型 RM 系统 *c5423-c5424-c5425*, 为细菌的表观遗传学和 UPEC 病原学研究提供了重要参考。

1 材料和方法

1.1 材料

1.1.1 菌株和质粒: UPEC CFT073, Lambda Red 同源重组系统质粒 pKD46、pKD3、pKD4、pCP20 和 pCJ112 均由本实验室保存。

1.1.2 主要试剂和仪器: LB 固体和液体培养基均购自北京博润莱特科技有限公司; 所有的 PCR 试剂购自宝生物工程(大连)有限公司(TaKaRa);

限制性内切酶购自赛默飞世尔科技(中国)有限公司; 总 RNA 提取试剂盒(Promega EastepTM Super Total RNA Extraction Kit)购自北京宝林科生物科技有限公司; 细菌全基因组提取试剂盒购自哈尔滨牧乐生物试剂有限公司; 氨苄青霉素(Amp)和卡那霉素(Kan)购自武汉盖云天生物技术有限公司。

1.2 细菌培养条件

1.2.1 静置培养: 将菌种接种到 LB 液体培养基中, 置于恒温培养箱中 37 °C 静置培养 12 h, 然后, 1:100 转接到新的 LB 液体培养基中, 37 °C 静置培养 6 h 至菌液 $OD_{600} \approx 1$ 。

1.2.2 振荡培养: 直接挑取菌种划线到 LB 固体培养基上, 恒温培养箱中 37 °C 倒置培养 12 h, 挑取单菌落接种到 LB 液体培养基中, 37 °C、180 r/min 培养 3 h 至菌液 $OD_{600} \approx 1$ 。

1.3 引物设计与合成

根据 GenBank 上公布的 UPEC CFT073 的全基因序列, 利用软件 Clonemanager 设计所有本研究中用到的特异性引物(表 1)。克隆带有目的基因(*c5424*)同源臂的替换引物 MT54-F 和 MT54-R 由两部分组成, 引物的 5'端前 55 bp 是和缺失目的基因(*c5424*)两侧同源的, 后边 3'端部分是以 pKD4 为模板克隆卡那霉素的基因特异性引物。荧光定量 RT-PCR 引物均使用网站(<https://sg.idtdna.com/Primerquest/>)设计。以上所有的引物均由吉林库美生物有限公司合成。

1.4 生物信息学分析

根据 GenBank 上发表的 *CFT073* 基因组序列, 利用 Rebase 数据库(<http://rebase.neb.com/rebase/rebase.html>)分析基因组中 RM 系统的分布信息, 通过与大肠杆菌 K12、UTI89、EC958 和 O157 等菌株 RM 系统的分布相互比较分析, 鉴定并阐述 I 型 RM 系统在 UPEC 上的分布特征。

表 1. 本研究中用到的引物
Table 1. Primers used in this study

Primers	Sequences (5'→3')	Target genes
MT54-F	TGTGCCTTCTGCAACCATCATGTTGCAGAAGGCCTTGTTTTTCATGGAATACAA CGTGTAGGCTGGAGCTGCTTCGA	
MT54-R	TGCAACGCCGAGGTCCAGGTTTCCATATGATCAACGATCAGCTTTTCCACACTC ACATATGAATATCCTCCTTAG	
Check-F	TATGGTCAGCCGCCGTTAAC	
Check-R	ATTTCCGCAATTCTAGTTAGC	
Kan-R	CCTCGTCCTGCAGTTCATTC	
Kan-F	GCTTGCCGAATATCATGGTG	
4cut-F	CGCGCCCGGTACCAGCATGCTGACAATTGCGCTCGAATTCTGGTACACCTGCTG TGTC	
4cut-R	TAGCTGTTTCTGTCAGTCATGACCTAGTTACTCTCTAGAAACCTGCCTTATCAT CAAAG	
qRT-PCR		
c0418-F	CCATCAGCCCACCCAATAAA	c0418 (forward)
c0418-R	AGCATGGCAAACGGCTAATA	c0418 (reverse)
c0421-F	CCACAAGGAACGATCGATAAGA	c0421 (forward)
c0421-R	AACGCATGCCAACGAATAAC	c0421 (reverse)
c2304-F	GAGGCTAATTCGGTTGGGAATA	c2304 (forward)
c2304-R	CAGAAATCGCCCGATGTTTAGA	c2304 (reverse)
c5423-F	CTGCACCTGCCTTTGTTTAC	c5423 (forward)
c5423-R	CAAAGTGATGGAAGGCTGTATTG	c5423 (reverse)
c4327-F	AGGATACTGCAACTCGCTTAAT	c4327 (forward)
c4327-R	CGCGAAGAAGGTACGTCATATT	c4327 (reverse)
rpoB-F	GCATCATCCCTTACCGTGGTTC	rpoB (forward)
rpoB-R	GGATCTGCTCTGTGGTGTAGTTCA	rpoB (reverse)

1.5 Illumina 平台与单分子实时(SMRT)测序

采用静置培养的方式,待细菌生长至平台期,使用 DNA 提取试剂盒 (TaKaRa) 提取野生株 CFT073 和突变株 $\Delta c5424$ 的全基因组,将样本保存在无 RNA 酶的 ddH₂O 中。每个样本使用声裂法,将约 100 ng 的全基因组处理成小于 500 bp 的片段。使用末端修复酶对片段进行末端修复,选择修复好的片段回收。然后,每个样本使用 P5 和 P7 引物进行 PCR 扩增,PCR 产物用 Agilent 2100 生物分析仪进行分析,并且使用 Qubit 3.0 荧光计进行定量。随后,将不同的 DNA 文库加载到

Illumina HiSeq 仪器上,测序采用 2×150 双端(PE)测序,使用 HiSeq 控制软件(HCS)+OLB+gapipine-1.6 (Illumina)对 HiSeq 进行图像分析和碱基调用。

对于 SMRT 测序,剪切全基因组 DNA 约 10 μg,然后选择 10 kb 的双链 DNA 片段,用通用发夹接头对 DNA 片段进行末端修复。根据要求准备文库,文库在 PacBio Sequel 仪器上进行测序,后续测序分析由苏州金唯智生物技术有限公司完成,数据已上传至国家微生物科学数据中心(NMDC10017678)。

1.6 基因缺失株的构建

采用 Lambda Red 同源重组^[12]进行基因突变

株的构建, 首先利用 MT54-F 和 MT54-R 引物, 以 pKD4 为模板克隆带有目的基因同源臂的 Kan 基因(Kan^R)片段, PCR 产物经过 1.5%琼脂糖凝胶电泳检测后回收, 片段保存在-20 °C 中。

将带有 pKD46 的 CFT073 菌种接种带有 Amp 抗性的 LB 液体培养基中, 30 °C、180 r/min 振荡培养 OD_{600} 达到 0.2, 加入 1%的 L-阿拉伯糖溶液, 继续培养直到菌液 OD_{600} 达到 0.6–0.8。将菌液制备成电转化感受态细胞, 电转化之前回收的带同源臂的片段, 将转化好的菌液均匀涂布到带有 Kan 基因抗性的 LB 固体培养基上, 37 °C 倒置培养 12 h, 直到长出肉眼可见的单菌落, 使用 Check-F、Kan-R 和 Check-R、Kan-F 鉴定目的基因缺失是否成功。筛选出带有抗性的 $\Delta c5424::Kan^R$ 阳性重组子, 将其制成电转化感受态细胞, 将 pCP20 质粒电转化到 $\Delta c5424::Kan^R$ 感受态细胞中, 37 °C 培养去掉 Kan^R 和 pCP20 质粒, 获得 UPEC 目的基因突变株 $\Delta c5424$ 。

1.7 生长曲线测定

分别将野生株 CFT073 和突变株 $\Delta c5424$ 接种到 LB 液体培养基中, 37 °C、180 r/min 培养过夜, 然后使用新的 LB 培养基将菌液 OD_{600} 稀释到 0.05, 之后将稀释好的菌液同时置于相同的条件下培养, 每 0.5 h 测量一次菌液 OD_{600} 值。汇总数据, 绘制生长曲线, 比较野生株 CFT073 和突变株 $\Delta c5424$ 生长速度的差异。

1.8 转化效率试验

转化效率试验参考文献[8]: 用带有识别基序的引物 4cut-F/R 扩增 *pdeH* 基因, 用 C112 试剂盒将 DNA 片段克隆至 pCJ112 质粒, 形成 p4cut。制备各菌株的电转化感受态细胞, 每管约含有 5×10^9 个细胞。通过电转化(2.5 kV)将 0.3 μ g 质粒转入野

生株和突变株感受态细胞中, 将转化后的菌液平均分为两部分, 均匀涂布在 Kan 抗性和无抗性的 LB 平板上, 37 °C 培养, 统计长出的菌落数, 转化效率% = Kan^R 的转化子 CFU/无抗性 LB 的 CFU $\times 100\%$ 。

1.9 转录组测序(RNA-seq)

采用静置培养的方式, 将野生株 CFT073 和突变株 $\Delta c5424$ 培养至 $OD_{600} \approx 1$, 使用总 RNA 提取试剂盒提取总 RNA 进行转录组测序。转录组测序由金唯智生物技术有限公司完成, 数据已上传至国家微生物科学数据中心(NMDC10017678)。

1.10 荧光定量 RT-PCR(qRT-PCR)

静止培养野生株 CFT073 和突变株 $\Delta c5424$, 提取总 RNA, 通过 qRT-PCR 验证转录组测序结果。以基因 *rpoB* 为内参基因, 引物见表 1。

1.11 细菌运动性试验

细菌运动性试验^[13], 采用静置培养的方式, 将野生株 CFT073、突变株 $\Delta c5424$ 培养至菌液 $OD_{600} \approx 1$, 用无菌牙签蘸取菌液垂直插进 LB 半固体培养基中, 但是不要触及底部, 37 °C 静置培养 12 h, 测量细菌菌落生长浑浊区域的直径大小, 所有试验重复 3 次测量, 计算浑浊区域的直径平均值并进行统计。

1.12 细菌抗氧化试验

细菌抗氧化试验^[14-16], 采用静置培养的方式, 细菌过夜培养, 1:100 转接到新的 LB 培养基中, 同样的条件下, 培养细菌至 $OD_{600} \approx 0.4-0.6$ 。取 1 mL 菌液, 4 °C、5000 r/min 离心 10 min, 收集菌体, 使用预冷的 PBS 缓冲液清洗菌体 2 次, 使用等体积的 PBS 重悬菌体, 然后 1:5 稀释菌体至 $OD_{600} = 0.1$ 。于 96 孔板中, 100 μ L 菌液和 100 μ L

2 倍预期浓度(200 $\mu\text{mol/L}$)次氯酸溶液混合(迅速取样后终止反应,涂板计数,此数据为 0 min), 37 $^{\circ}\text{C}$ 避光培养 20 min, 加入终浓度 9 mmol/L 的甲硫氨酸溶液终止反应, 最后使用 PBS 溶液稀释涂板, 计数。存活率计算公式为: 存活率=(CFU_{20 min}/CFU_{0 min}) \times 100%。

2 结果和分析

2.1 生物信息学分析鉴定 CFT073 中的 I 型 RM 系统

基于 RM 系统中 DNA 甲基化转移酶保守的酶活位点(DPPY 和 FxGxG)和基因构成预测到 UPEC CFT073 包含有一个新的 I 型 RM 系统(*c5423-c5424-c5425*), 分别编码特异性识别蛋白(HsdS)、甲基化转移酶(HsdM)和核酸内切酶(HsdR)(图 1-A)。对 *c5424* 进行 BLASTp 分析发现:

在非致病性大肠杆菌 MG1655 菌株中, 不存在 *c5424* 的直系同源蛋白; 在某些 UPEC 和 EHEC 菌株中存在 *c5424* 的直系同源蛋白。值得一提的是, 虽然在 MG1655 和 UPEC 膀胱炎分离株 UTI89 中不存在 *c5424* 的直系同源蛋白, 但是在染色体的相同位置, 存在另外一套 I 型 RM 系统, 由 3 个基因构成, 分别为 UTI89_C5050、UTI89_C5051 和 UTI89_C5053 (图 1-B)。

2.2 *c5424* 基因缺失株构建及生长特性测定

利用 Lambda Red 重组系统进行缺失株构建, 获得带有 Kan 抗性的 *c5424* 的缺失株 $\Delta c5424::kan^R$ 。进一步利用带有翻转重组酶的 pCP20 质粒, 去掉 Kan 抗性基因, 最终获得 $\Delta c5424$ 菌株。

通过测定野生株 CFT073 和突变株 $\Delta c5424$ 在 LB 液体培养基中的生长情况, 绘制生长曲线, 由图 2 可知, *c5424* 基因对菌株生长无明显影响。

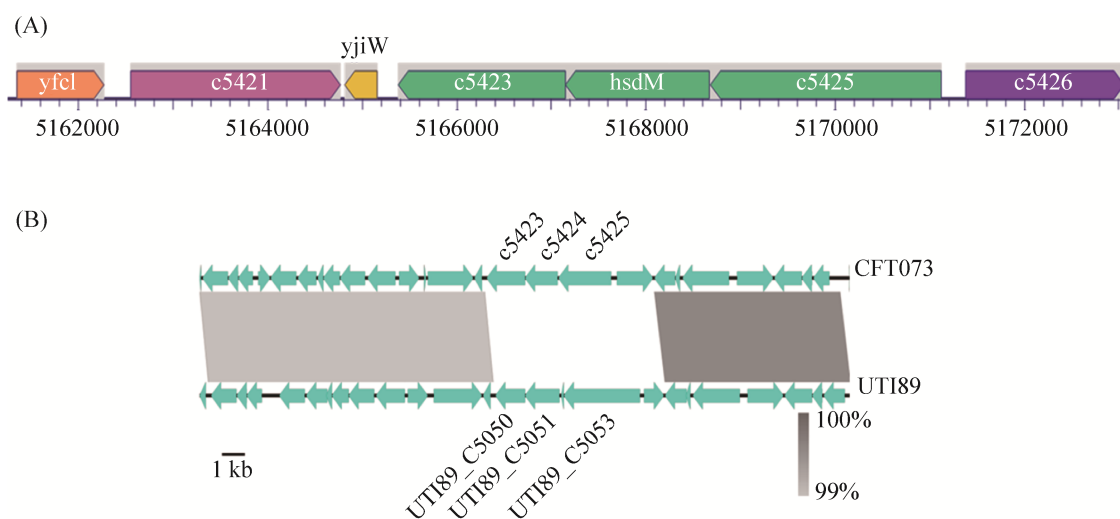


图 1. *c5423-c5425* 基因位置信息与同源比较

Figure 1. Genetic organization and sequence comparison of *c5423-c5425* genes. A: genetic organization of the type I RM system genes in UPEC CFT073 genome; B: sequence comparison of the RM systems and their adjacent genes between the UPEC CFT073 and UTI89 strains. The grey scale indicates the homology between two DNA molecules.

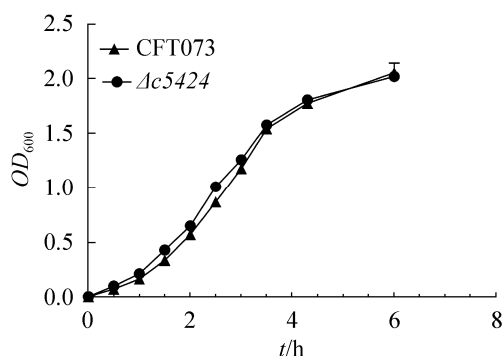


图 2. $c5424$ 基因缺失株的生长曲线

Figure 2. Growth curve of the $\Delta c5424$ mutant.

2.3 DNA 甲基化转移酶 $c5424$ 修饰基序的鉴定

为了鉴定 $C5424$ 甲基化修饰的基序, 使用 SMRT 单分子实时测序对野生株 CFT073 及 $c5424$ 缺失株进行全基因组测序, 同时利用 Illumina HiSeq 进行序列校准。根据碱基的 IPD 值, 通过 SMRTlink 软件分析, 确定碱基的修饰类型。分析结果表明野生株中共测到 3 种 m6A 的甲基化基序, 而 $c5424$ 突变株中仅有 2 种基序(表 2), 表明

$C5424$ 负责 $G^mAGNNNNNNGTCA$ 基序的甲基化。更进一步, 我们对 427 个基序作了定位, 按照正义链/反义链、基因编码区/基因间区等参数进行排布, 以展示基序的位置规律(图 3)。比如, 在 3700–3800 kb 的位置, 基序倾向于分布在正义链上; 在 300–400 kb 的位置, 基序倾向于分布在基因间区, 这种成簇的在基因间区的出现, 可能参与直接调控。

2.4 RM 系统对转化效率的影响

为了研究 RM 系统($C5423$ - 5425)是否参与抵御外来 DNA, 我们首先构建了带有识别基序的质粒 p4cut; 将 p4cut 质粒通过电穿孔分别转化至野生株 CFT073 和突变株 $\Delta c5424$ 中, 发现突变株能够得到更多的转化子(图 4-A); 与野生株比, 转化效率提高约 30 倍(图 4-B), 说明 RM 系统($C5423$ - 5425)参与抵御外来的 DNA。作为对照, 我们在野生株和 $lacZ$ 基因突变株中同时转化 p4cut 质粒, 转化效率无显著差别(结果未展示)。

表 2. 单分子实时测序分析 CFT073 和 $\Delta c5424$ 的 m6A 甲基化结果

Table 2. The single molecule real time sequencing analysis of m6A methylation in CFT073 and $\Delta c5424$

MS	CP	MT	fraction	nD	nG	GT	MIR
Motif detected in WT							
GATC	2	m6A	0.998	41465	41538	GATC	5.71
CACAG	4	m6A	1	8010	8010	CACAG	5.97
GAGNNNNNNGTCA	2	m6A	0.81	355	427	GAGNNNNNNGTCA/TGA CNNNNNNNCTC	4.49
TGACNNNNNNNCTC	3	m6A	0.791	341	427	GAGNNNNNNGTCA/TGA CNNNNNNNCTC	3.73
Motif detected in mutant							
GATC	2	m6A	0.998	41465	41538	GATC	5.71
CACAG	4	m6A	1	8010	8010	CACAG	5.97

MS (motif string): motif sequence; CP (center position): the location of modification base in motif; MT: modification type; fraction: percentage of detected modified motif relative to the amount of all motif sequences in the genome; nD (ndetected): number of modified motif sequences detected; nG (ngenome): number of motif sequences in the genome; GT (group tag): the presence of bipartite motif; MIR (mean Ipd ratio): the average value of IPD.

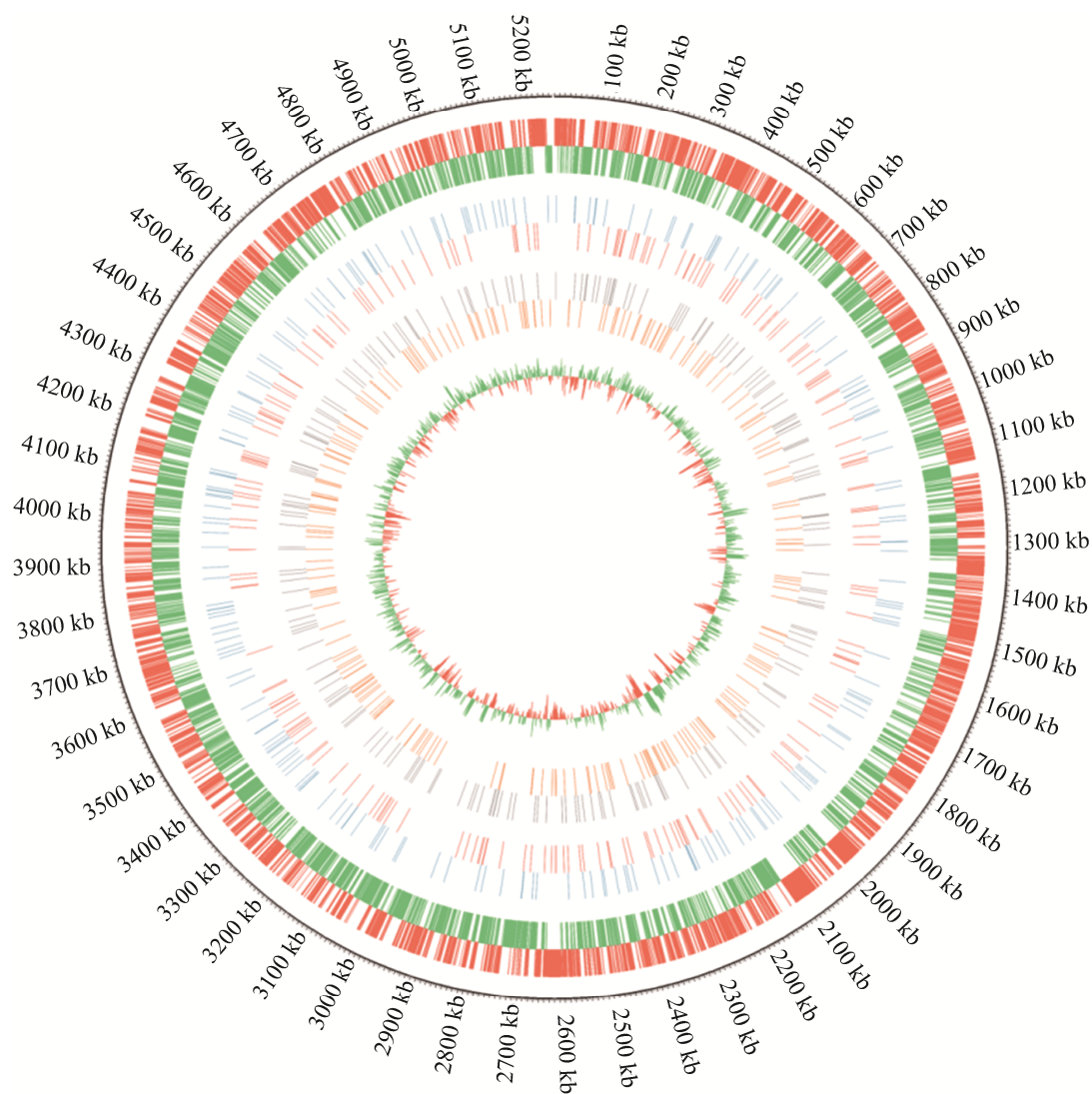


图 3. *c5424* 修饰的基序在 CFT073 基因组中的分布

Figure 3. Distribution of C5424 modified motifs in the CFT073 genome. The five concentric circles represent (from outside to inside) genomic positions; ORF on the positive and negative strands; C5424 motifs on the positive and negative strands; C5424 motifs on the CDS regions and intergenic regions; GC content, peaks facing outside indicate higher GC content than the average genomic GC content, peaks facing inside indicate lower GC content than the average genomic GC content, the higher the peaks, the bigger the difference.

2.5 转录组测序分析

为了确定基因 *c5424* 调控的靶基因, 我们比较了突变株 $\Delta c5424$ 和野生株 CFT073 的基因表达谱, 转录组测序数据表明 *c5424* 影响 17 个基因的表达 (表 3), 随后, 我们挑选了其中 5 个基因做了 qRT-PCR 验证, 结果与转录组测序结果基本相符 (图 5)。

2.6 *c5424* 影响细菌运动能力

C5424 影响 *motB* 的表达大于 2 倍, 而 *motB* 与细菌运动能力有关。为了研究 *c5424* 的缺失是否影响细菌运动能力, 测量并比较 LB 半固体培养基上野生株 CFT073 与突变株 $\Delta c5424$ 菌落浑浊区域大小。如图 6, 野生株 CFT073 浑浊区域大

小均数(1.5 ± 0.3) cm, 突变菌株 $\Delta c5424$ 浑浊区域大小均值(0.9 ± 0.2) cm, 两者呈现显著性差异 ($P < 0.05$)。因此, $c5424$ 对于维持细菌的运动能力很重要。

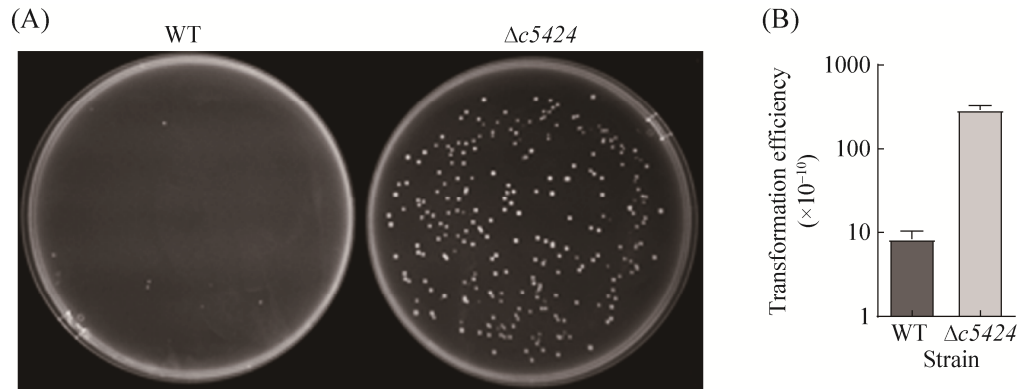


图 4. C5424 对转化效率的影响

Figure 4. The effect of C5424 on plasmid transformation efficiency. A: a plate view of transformants after transformation of p4cut into the wild-type strain CFT073 and mutant strain $\Delta c5424$; B: transformation efficiency of plasmid p4cut in the wild-type strain CFT073 and mutant strain $\Delta c5424$.

表 3. 转录组测序分析得到突变株 $\Delta c5424$ 中差异表达基因

Table 3. RNA-seq analysis revealed differentially expressed genes in the mutant $\Delta c5424$ compared to the wild-type strain

Genes	Gene products ^a	Log ₂ (fold change) ^b ($\Delta c5424$ VS CFT073)
gene-c1234	Conserved hypothetical protein	-2.57
gene-c0418	Hypothetical protein ykgB	-2.35
gene-c3692	Hypothetical protein	-2.02
gene-c2304	Chemotaxis motB protein	-1.82
gene-c4327	Transcriptional regulator gadX	-1.69
gene-c1233	Hypothetical protein	-1.57
gene-c0421	Hypothetical transcriptional regulator ykgD	-1.56
gene-c4323	Hypothetical protein yhiE	-1.53
gene-c5426	Conserved hypothetical protein	-1.44
gene-c5423	Putative restriction modification enzyme S subunit	-1.44
gene-c3145	Hypothetical protein ydfK	-1.38
gene-c4326	Hypothetical transcriptional regulator yhiW	-1.24
gene-c5395	Chaperone protein fimC precursor	-1.23
gene-c1123	Cold shock-like protein cspG	-1.21
gene-c3184	Cold shock-like protein cspB	-1.18
gene-c2301	Methyl-accepting chemotaxis protein II	-1.15
gene-c3816	30S ribosomal protein S21	1.24

^a: according to GenBank annotation; ^b: values in this column represent the changes in the $\Delta c5424$ in comparison to the parental strain, CFT073. Transcriptome analysis was carried out with three biological replicates.

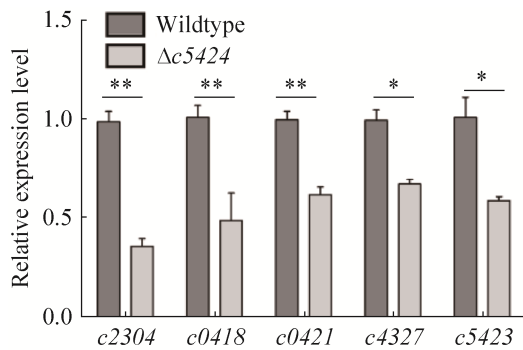


图 5. qRT-PCR 验证部分转录组测序得到的差异表达基因

Figure 5. Validation of differentially expressed genes by qRT-PCR. Transcriptional levels were normalized to CFT073 *rpoB*; data are mean \pm SD from three biological replicates ($n=3$). *: $P<0.05$; **: $P<0.01$ (Student's *t* test).

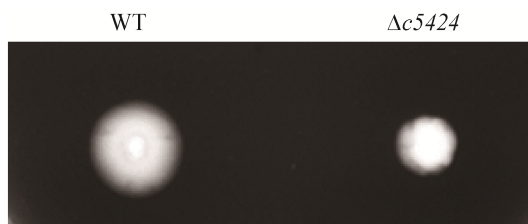


图 6. 野生株 CFT073 和 $\Delta c5424$ 突变株在 LB 半固体培养基上的运动能力检测

Figure 6. Bacterial motility of *E. coli* CFT073 and $\Delta c5424$ in semi-solid agar.

2.7 *c5424* 促进细菌抵抗次氯酸杀伤

C5424 调控基因 *c0421* 的表达; 其编码的基因属于 AraC 家族转录调控因子^[15], 与细菌抗次氯

酸杀伤有关, 为了说明 *c5424* 参与细菌抵抗次氯酸, 比较野生株 CFT073 与突变株 $\Delta c5424$ 被次氯酸杀伤后的存活能力。如图 7, 未经过次氯酸处理的野生株和突变株, 细菌数无显著差别; 次氯酸处理后, 野生株 CFT073 的存活率为 0.0042%, 突变株 $\Delta c5424$ 存活率约为 0.0006%, 呈现明显下降趋势。表明 *c5424* 促进细菌抵抗次氯酸杀伤。

3 讨论

依据致病部位、症状和毒力因子等, 致病性大肠杆菌被分为多个致病型, 如肠出血型、致脑膜炎型和尿路致病型。这种多样性一般被认为是由于大肠杆菌基因组的可塑性导致的, 比如质粒、转座子和噬菌体等可移动元件极大地改变了大肠杆菌基因组的构成^[17]。本研究关注的 *C5424* 系统, 在非致病性大肠杆菌中不存在, 而只存在于部分 UPEC 和 EHEC 的菌株中, 暗示其可能是通过基因水平转移获得的, 其对某些菌株有特殊的作用。的确, 我们发现 *C5424* 影响 *motB* 的表达以及运动性, *motB* 是鞭毛运动蛋白复合体的主要组成成分; 鞭毛是 UPEC 和 EHEC 中重要的粘附因子和侵入因子, 尤其是, 它被认为在 UPEC 从膀胱上行至肾脏感染的过程中发挥重要作用^[18]。编码 RclR 的基因 *c0421* 在大肠杆菌抵抗活性氯的过程中发挥重要的作用。RclR 是 AraC 族转录激活因子, 其

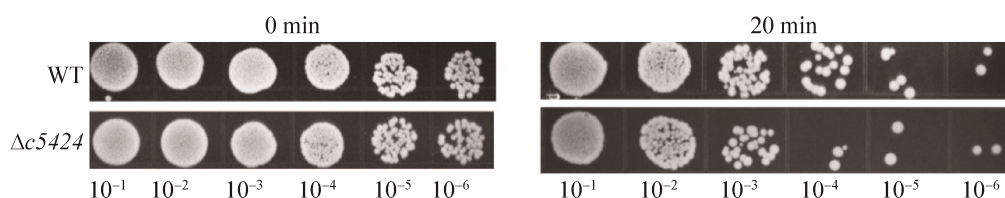


图 7. 野生株 CFT073 和 $\Delta c5424$ 突变株抗次氯酸能力测定

Figure 7. Hypochlorous acid resistance assay of *E. coli* CFT073 and $\Delta c5424$.

高度保守的半胱氨酸残基对活性氯极度敏感,当细菌受到活性氯的氧化胁迫后会激活 RclR 以增加细菌抗氧化的能力^[14]。本文中我们发现 C5424 调控 *rclR* 并参与 CFT073 抵抗次氯酸的杀伤;巨噬细胞和嗜中性粒细胞中都有髓过氧化物酶 (myeloperoxidase), 它可将过氧化氢和氯离子转化为次氯酸, 以起到杀灭微生物的作用^[19]。因此, C5424 系统可能使得 UPEC 在被巨噬细胞和嗜中性粒细胞吞噬后能够更好地存活。此外, C5424 还调控抗酸基因 *c4327*^[20]、菌毛蛋白基因 *c5395*^[21] 等 UPEC 体内适应性相关的基因。

虽然 DNA 甲基化转移酶在病原菌中有广泛的调控作用, 但是 RM 系统本身的主要功能是抵御外来 DNA 的入侵, 比如质粒或噬菌体。在链球菌中, I 型 RM 系统介导抵御外源 DNA 通过转化进入细胞^[8,22]。我们利用 SMRT 测序方法比较野生株 CFT073 和缺失株 $\Delta c5424$ 的甲基化组, 确定了 C5424 甲基化修饰的基序, 这个基序成对出现, 符合 I 型系统的特征^[23]。接下来我们构建了包含识别基序的质粒。很明显, C5424 的缺失极大地增加了质粒 DNA 的转化, 表明 C5424 系统可以用来抵御外来 DNA, 很有可能是 C5425 介导了质粒的切割。UPEC 通常定殖于肠道中, 一旦污染了尿道口, 就可能引起尿路感染^[24]。人体肠道内存在大量的噬菌体 ($>10^{10}/g$ 粪便)^[25], 因此, C5424 系统介导的外源 DNA 切割可能有助于大肠杆菌抵御肠道内噬菌体从而更好地定殖。另外, 值得一提的是, 生物信息学预测与 SMRT 测序结果暗示, 在 CFT073 中存在其他的 RM 系统, 也发挥抵御外来 DNA 的功能。这些不同的系统很可能有不同的修饰基序, 用来抵御不同来源的 DNA。

但是, C5424 系统调控靶基因的分子机制仍

是未解之谜。DNA 甲基化转移酶调控的常见方式有: (1) 通过甲基化靶基因的启动子区域, 影响 RNA 聚合酶或调控蛋白的结合; (2) DNA 甲基化转移酶内部的重复序列的重组导致相变的发生, 进而引起甲基化识别位点变化以及靶基因变化; (3) DNA 的甲基化导致染色体构象发生变化, 进而影响转录^[26]。经过搜索, 我们没有发现靶基因的启动子区域含有识别基序, 也没有发现 *c5424* 基因内部有正向或反向重复序列, 因此我们推测 C5424 影响靶基因的转录很有可能是通过第三种机制, 即间接调控方式。

最后, 本研究鉴定了大肠杆菌 CFT073 菌株中的一套 RM 系统 C5423-5425: 鉴定出 DNA 甲基化转移酶 C5424 的修饰基序; 确定了 RM 系统可以抵御外源 DNA; 从全局层面定义了 C5424 的调控元, 这些靶基因很有可能促进了 UPEC 的体内适应性或毒力。

参考文献

- [1] Vasu K, Nagaraja V. Diverse functions of restriction-modification systems in addition to cellular defense. *Microbiology and Molecular Biology Reviews*, 2013, 77(1): 53–72.
- [2] Tong T, Wang LR. Epigenetic regulation role of DNA methylation in bacteria. *Acta Microbiologica Sinica*, 2017, 57(11): 1688–1697. (in Chinese)
童童, 王连荣. 甲基化修饰在细菌表观调控中的功能. *微生物学报*, 2017, 57(11): 1688–1697.
- [3] Roberts RJ, Belfort M, Bestor T, Bhagwat AS, Bickle TA, Bitinaite J, Blumenthal RM, Degtyarev SK, Dryden DTF, Dybvig K, Firman K, Gromova ES, Gumport RI, Halford SE, Hattman S, Heitman J, Hornby DP, Janulaitis A, Jeltsch A, Josephsen J, Kiss A, Klaenhammer TR, Kobayashi I, Kong H, Krüger DH, Lacks S, Marinus MG, Miyahara M, Morgan RD, Murray NE, Nagaraja V, Piekarowics A, Pingoud A, Raleigh E, Rao DN, Reich N, Repin VE, Selker EU, Shaw PC, Stein DC, Stoddard BL, Szybalski W, Trautner TA,

- Etten JL, Vitor JMB, Wilson GG, Xu SY. A nomenclature for restriction enzymes, DNA methyltransferases, homing endonucleases, and their genes restriction endonucleases. *Nucleic Acids Research*, 2003, 31(7): 1805–1812.
- [4] Dryden DTF, Murray NE, Rao DN. Nucleoside triphosphate-dependent restriction enzymes. *Nucleic Acids Research*, 2001, 29(18): 3728–3741.
- [5] Murray NE. Type I restriction systems: sophisticated molecular machines (a legacy of Bertani and Weigle). *Microbiology and Molecular Biology Reviews*, 2000, 64(2): 412–434.
- [6] de Ste Croix M, Vacca I, Kwun MJ, Ralph JD, Bentley SD, Haigh R, Croucher NJ, Oggioni MR. Phase-variable methylation and epigenetic regulation by type I restriction-modification systems. *FEMS Microbiology Reviews*, 2017, 41(Supp_1): S3–S15.
- [7] Doberenz S, Eckweiler D, Reichert O, Jensen V, Bunk B, Spröer C, Kordes A, Frangipani E, Luong K, Korfach J, Heeb S, Overmann J, Kaefer V, Häussler S. Identification of a *Pseudomonas aeruginosa* PAO1 DNA methyltransferase, its targets, and physiological roles. *mBio*, 2017, 8(1): e02312–16.
- [8] Nye TM, Jacob KM, Holley EK, Nevarez JM, Dawid S, Simmons LA, Watson ME. DNA methylation from a Type I restriction modification system influences gene expression and virulence in *Streptococcus pyogenes*. *PLoS Pathogens*, 2019, 15(6): e1007841.
- [9] Shea AE, Marzoa J, Himpls SD, Smith SN, Zhao LL, Tran L, Mobley HLT. *Escherichia coli* CFT073 fitness factors during urinary tract infection: identification using an ordered transposon library. *Applied and Environmental Microbiology*, 2020, 86(13): e00691–20.
- [10] Zalewska-Piątek BM, Piątek RJ. Alternative treatment approaches of urinary tract infections caused by uropathogenic *Escherichia coli* strains. *Acta Biochimica Polonica*, 2019, 66(2): 129–138.
- [11] Cotter PA, DiRita VJ. Bacterial virulence gene regulation: an evolutionary perspective. *Annual Review of Microbiology*, 2000, 54: 519–565.
- [12] Datsenko KA, Wanner BL. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proceedings of the National Academy of Sciences of the United States of America*, 2000, 97(12): 6640–6645.
- [13] Spurbeck RR, Tarrion RJ, Mobley HLT. Enzymatically active and inactive phosphodiesterases and diguanylate cyclases are involved in regulation of motility or sessility in *Escherichia coli* CFT073. *mBio*, 2012, 3(5): e00307–12.
- [14] Winter J, Ilbert M, Graf PCF, Özcelik D, Jakob U. Bleach activates a redox-regulated chaperone by oxidative protein unfolding. *Cell*, 2008, 135(4): 691–701.
- [15] Parker BW, Schwessinger EA, Jakob U, Gray MJ. The RclR protein is a reactive chlorine-specific transcription factor in *Escherichia coli*. *Journal of Biological Chemistry*, 2013, 288(45): 32574–32584.
- [16] Gebendorfer KM, Drazic A, Le Y, Gundlach J, Bepperling A, Kastenmüller A, Ganzinger KA, Braun N, Franzmann TM, Winter J. Identification of a hypochlorite-specific transcription factor from *Escherichia coli*. *The Journal of Biological Chemistry*, 2012, 287(9): 6892–6903.
- [17] Kaper JB, Nataro JP, Mobley HL. Pathogenic *Escherichia coli*. *Nature Reviews Microbiology*, 2004, 2(2): 123–140.
- [18] Lane MC, Alteri CJ, Smith SN, Mobley HLT. Expression of flagella is coincident with uropathogenic *Escherichia coli* ascension to the upper urinary tract. *Proceedings of the National Academy of Sciences of the United States of America*, 2007, 104(42): 16669–16674.
- [19] Albrecht AM, Ashby LV, Dickerhof N, Kettle AJ, Winterbourn CC. Heterogeneity of hypochlorous acid production in individual neutrophil phagosomes revealed by a rhodamine-based probe. *The Journal of Biological Chemistry*, 2018, 293(40): 15715–15724.
- [20] Hu JL, Lei P, Mohsin A, Liu XY, Huang MZ, Li L, Hu JH, Hang HF, Zhuang YP, Guo MJ. Mixomics analysis of *Bacillus subtilis*: effect of oxygen availability on riboflavin production. *Microbial Cell Factories*, 2017, 16: 150. <https://doi.org/10.1186/s12934-017-0764-z>
- [21] Fekrirad Z, Kashef N, Arefian E. Photodynamic inactivation diminishes quorum sensing-mediated virulence factor production and biofilm formation of *Serratia marcescens*. *World Journal of Microbiology and Biotechnology*, 2019, 35: 191. <https://doi.org/10.1007/s11274-019-2768-9>
- [22] Oliver MB, Swords WE. Comparative analysis of *Streptococcus pneumoniae* type I restriction-modification loci: variation in hsdS gene target recognition domains. *Pathogens*, 2020, 9(9): 712.
- [23] Loenen WAM, Dryden DTF, Raleigh EA, Wilson GG. Type I restriction enzymes and their relatives. *Nucleic Acids Research*, 2014, 42(1): 20–44.
- [24] Alteri CJ, Mobley HLT. Metabolism and fitness of urinary

tract pathogens.//Conway T, Cohen PS. Metabolism and Bacterial Pathogenesis. American Society of Microbiology, 2015: 215–230.

[25] Sutton TDS, Hill C. Gut bacteriophage: current understanding

and challenges. *Frontiers in Endocrinology*, 2019, 10: 784.

[26] Beaulaurier J, Schadt EE, Fang G. Deciphering bacterial epigenomes using modern sequencing technologies. *Nature Reviews Genetics*, 2019, 20(3): 157–172.

Characterization of a type I restriction-modification system C5423-5425 in uropathogenic *Escherichia coli*

Huoming Li, Ying Xia, Ganwu Li*, Wentong Cai*

State Key Laboratory of Veterinary Biotechnology, Harbin Veterinary Research Institute, Chinese Academy of Agricultural Sciences, Harbin 150069, Heilongjiang Province, China

Abstract: [Objective] To characterize a type I restriction-modification (RM) system C5423-5425 in uropathogenic *Escherichia coli* CFT073 and to determine its function and physiological significance. [Methods] Lambda Red recombination system was used to construct methyltransferase gene deletion mutant $\Delta c5424$ in CFT073. The modification sites and recognition motif of C5424 methyltransferase were obtained by single molecule real-time sequencing analysis. A transformation efficiency assay was used to test that the RM system C5423-5425 can block transformation of exogenous DNA. Genes regulated by C5424 were identified by transcriptome sequencing and real-time quantitative PCR. The physiological function of *c5424* was studied by soft agar plate motility assay. [Results] A type I RM system was identified by bioinformatics methods. A *c5424* deletion mutant was constructed. We found that the C5424 recognition motif was $G^mAGNNNNNNGTCA/TG^mACNNNNNNNCTC$, and its distribution in the genome was presented. C5424 contributed to reducing transformation of foreign DNA. Deletion of *c5424* significantly affected the expression of 17 genes, including *motB* and *rcIR*. Lacking *c5424* significantly affected bacterial motility and the resistance to hypochlorous acid in CFT073. [Conclusion] In this paper, the type I RM system C5423-5425 was characterized in detail, and it is physiologically important for uropathogenic *E. coli* CFT073. Thus, our data provided valuable information for bacterial epigenetics studies.

Keywords: uropathogenic *Escherichia coli*, DNA methyltransferase, single molecule real-time sequencing, restriction-modification system

(本文责编: 李磊)

Supported by the National Natural Science Foundation of China (31902242) and by the National Key Research and Development Project (2018YFD0500500)

*Corresponding authors. Tel/Fax: +86-451-51051728; E-mail: Ganwu Li, liganwu@caas.cn, Wentong Cai, caiwentong@caas.cn
Received: 20 September 2020; Revised: 14 January 2021; Published online: 12 March 2021