



L1.LtrB 内含子编码蛋白反转录结构域关键催化位点分析及功能验证

陈相好^{1,2}, 张峥嵘^{1,3}, 刘芳¹, 陈峥宏^{1,3}, 洪伟⁴, 綦廷娜^{1,3}, 谷俊莹^{2,5*}, 崔古贞^{1,3*}

¹ 贵州医科大学基础医学院, 贵州 贵阳 550025

² 贵州医科大学医学检验学院, 贵州 贵阳 550004

³ 贵州省普通高等学校病原生物学特色重点实验室, 贵州 贵阳 550025

⁴ 贵州医科大学分子生物学重点实验室, 贵州 贵阳 550004

⁵ 贵州医科大学附属医院, 贵州 贵阳 550004

摘要: 【目的】筛选影响 L1.LtrB 内含子编码蛋白(Intron encoded protein, IEP)反转录功能的关键催化位点, 并获得无反转录活性的 IEP 突变体。【方法】首先, 利用 NCBI 数据库, 通过序列比对及同源建模方法筛选影响 IEP 反转录功能的关键氨基酸催化位点; 然后, 对筛选获得的关键催化位点进行定点突变, 同时以 Targetron 载体为模板, 构建无反转录功能的突变型 Targetron 打靶系统; 最后, 以大肠杆菌 *lacZ* 基因为例, 体内验证 IEP 突变体的功能及其对 II 型内含子“归巢”效率的影响。【结果】筛选到 C164 和 G214 两个位点是影响内含子编码蛋白反转录功能的关键氨基酸残基, 并获得 C164K 和 G214W 两个突变体。体内功能分析表明, 此两个位点突变完全失活了 II 型内含子的“归巢”功能。【结论】筛选并获得了失活反转录功能的 L1.LtrB 内含子编码蛋白突变体, 为深入研究 II 型内含子的结构和“归巢”机理奠定了基础。

关键词: L1.LtrB, II 型内含子, 内含子编码蛋白, 反转录, Targetron

II 型内含子(group II intron)是一类反转录转座子, 由具有催化功能的内含子 RNA 和具有反转录酶活性的内含子编码蛋白(intron encoded

protein, IEP)两部分组成^[1-4]。其中, 内含子 RNA 具有核酶活性, 能够通过碱基互补配对原则识别双链 DNA, 并切割 DNA 靶位点; IEP 蛋白具有

基金项目: 国家自然科学基金(31760318, 31500078, 31560318, 31601012); 贵州省科技计划项目(黔科合基础[2018]1132, [2019]1441); 贵州省教育厅自然科学研究项目(黔教合 KY 字[2014]216); 贵州省研究生科研基金立项项目(11348)

*通信作者。Tel: +86-851-88174015; E-mail: 谷俊莹, 921469044@qq.com, 崔古贞, cuiguzhen@hotmail.com

收稿日期: 2019-01-28; 修回日期: 2019-04-04; 网络出版日期: 2019-04-22

反转录酶功能，与内含子 RNA 结合成为核糖核蛋白复合体，辅助 II 型内含子的识别和切割；而且，IEP 蛋白能以内含子 RNA 为模板合成互补 cDNA，从而实现内含子 RNA 在染色体上的“归巢”(retrohoming)^[5-7]。

II 型内含子主要存在于细菌、真核生物细胞器(线粒体、叶绿体)及部分古细菌中^[8]。目前，研究较多的是来源于乳酸乳球菌(*Lactococcus lactis*)的 Ll.LtrB 的 II 型内含子，其“归巢”过程如图 1 所示。首先，内含子 RNA 与内含子编码蛋白组装为核糖核蛋白复合体，通过碱基互补配对

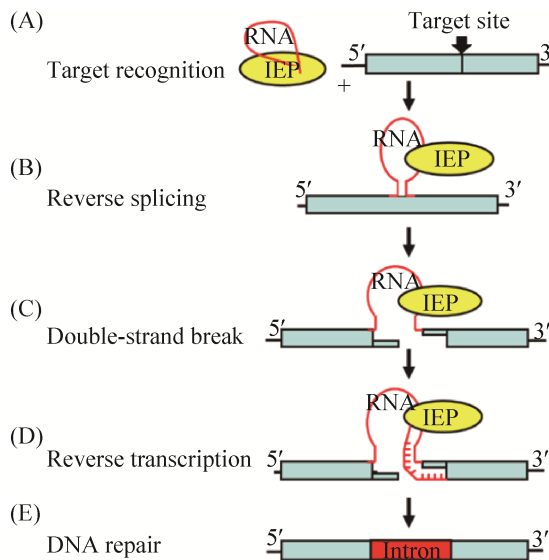


图 1. Ll.LtrB II 型内含子“归巢”原理示意图

Figure 1. Retrohoming of group II intron. A: target recognition, the complex composed of intron RNA and IEP recognizes target sites through base complementary pairing; B: reverse splicing, with the assistance of IEP, intron RNA is inserted into a single DNA strand; C: double-strand break, the endonuclease of IEP cleaves the other single strand of DNA; D: reverse transcription, group II introns synthesize cDNA using intron RNA as template with the reverse transcriptase activity of IEP; E: DNA repair, cells synthesize complementary DNA using cDNA as template through DNA repair mechanism.

原则识别双链 DNA 靶位点；然后，内含子 RNA 的核酶活性反向切割并整合于 DNA 靶位点；随后，IEP 蛋白的核酸内切酶活性切开双链 DNA 的反义链，并以切口处下游的 3'末端为引物，以插入的内含子 RNA 为模板合成 cDNA；最后，受损伤的 DNA 借助细胞自身修复机制，实现内含子 RNA 在靶位点的“归巢”^[6,9-10]。由于 II 型内含子的特异性主要是由内含子 RNA 的碱基配对决定的，因此，通过设计内含子 RNA 的识别序列，即可实现内含子 RNA 在染色体不同位点“归巢”^[9]。基于此原理，目前已开发出高效基因打靶技术——Targetron 技术，该技术已在革兰阴性菌(如大肠杆菌、鼠伤寒沙门氏菌、弗氏志贺菌)^[11-13]和革兰阳性菌(如艰难梭菌、肉毒梭菌、产气荚膜梭菌、金黄色葡萄糖菌、乳酸乳球菌)^[14-18]以及人类细胞^[19]等多个物种中获得广泛应用。

然而，上述 II 型内含子的“归巢”过程是根据现有实验数据进行的推测，仍缺乏详细的实验证据。近年来，由于冷冻电镜技术的发展，II 型内含子复合体的结构得以解析，为其机理研究提供了更深入的依据^[20]。由于 II 型内含子“归巢”过程中靶位点的识别、切割和反转录等过程是一个连续发生的过程，在时间上无法分离，在空间上彼此接近，因此，很难深入了解其每一步的详细过程。如果能够将 II 型内含子识别、切割、反转录等过程加以分离，就有可能深入了解其“归巢”机制。

II 型内含子“归巢”过程中最关键的一步是反转录，如果能够失活 IEP 蛋白的反转录功能，而又不影响 IEP 蛋白的其他功能，就有可能将 II 型内含子的“归巢”过程加以分离。因此，本研究首

先利用序列分析方法, 分析并筛选了影响 IEP 蛋白反转录活性的关键氨基酸残基, 然后通过定点突变获得失活反转录功能的 IEP 蛋白突变体, 最后, 通过构建突变型 Targetron 系统, 在大肠杆菌中验证突变型 IEP 蛋白的功能, 为深入了解 II 型内含子的“归巢”机制奠定了基础。

1 材料和方法

1.1 材料

1.1.1 主要试剂与仪器: TRYPTONE、YEAST EXTRACT、Agar Powder 购自 OXOID 公司; IPTG、X-gal、抗生素购自北京索莱宝公司; Trans Start FastPfu DNA Polymerase、DNA Marker 和 T4 DNA 连接酶购自北京全式金生物技术有限公司; 限制性内切酶购自 Thermo Fisher Scientific; PCR 产物纯化试剂盒、琼脂糖凝胶 DNA 回收试剂盒、定点突变试剂盒、质粒提取试剂盒购自北京天根公司。

PCR 仪和高速冷冻离心机, Thermo Fisher Scientific 公司; 电泳仪和凝胶成像仪, 北京六一公司; 恒温仪, 杭州米欧公司。

1.1.2 培养基及培养方法: LB 液体培养基(10 g/L TRYPTONE、5 g/L YEAST EXTRACT、10 g/L NaCl); LB 固体培养基(10 g/L TRYPTONE、5 g/L YEAST EXTRACT、10 g/L NaCl, 15 g/L Agar Powder)。液体培养基置于 37 °C、200 r/min 摇床中培养, 固体培养基置于 37 °C 恒温箱培养。需要时加入氨苄青霉素(100 µg/mL)进行筛选, IPTG (100 µmol/L)和 X-gal (40 µg/mL)加入 LB 固体培养基中用于蓝白斑筛选和菌落计数。

1.1.3 引物: 本研究用到的引物由生工生物工程(上海)股份有限公司合成, 引物及序列详见表 1。

1.1.4 菌株及质粒: 本研究中用的菌株及质粒详见表 2。

1.2 内含子编码蛋白生物信息学分析

以来源于乳酸乳球菌 L1.LtrB 的 II 型内含子编码蛋白序列为模板, 利用 NCBI 数据库进行同源序列比对和分析, 筛选可能影响 II 型内含子反转录功能的关键氨基酸残基, 同时在 <https://swissmodel.expasy.org/>检索获得 IEP 蛋白晶体结构, 利用 PyMOL 和 Swiss-PdbViewer 软件分析 IEP 蛋白三维结构图。

表 1. 本研究所用引物

Table 1. The primers used in this study

Primers name	Sequence(5'→3')	Description
lacZ-F	GGCCCGCACCGATCGCCCTTC	Detection primers of <i>lacZ</i>
lacZ-R	GCCATTTTTTGATGGACCATTC	
IEP-F	ATCGAGGCTAGCGCTATATGCGTTGATG	Amplification primers of IEP, the underline indicates enzyme loci of <i>Nhe I</i>
IEP-R	CGTTCAGATCTCCTTACTCGTA	Amplification primers of IEP, the underline indicates enzyme loci of <i>Bgl II</i>
C164K-F	GGAGATATAAAAGGCAAATTCGATAATATAGAC	Mutation primers of C164K
C164K-R	TTTGCCTTTTATATCTCCCTCCACAAACCATCT	
G214W-F	AGCGGAACACCTCAATGGGGAATTCTATCTCCT	Mutation primers of G214W
G214W-R	CCATTGAGGTGTTCCGCTGTAAGTTTTGTGATA	

表 2. 本研究所用到的菌株及质粒

Table 2. The strains and plasmids used in this study

Strains and plasmids	Relevant features	Source or reference
Strains		
<i>E. coli</i> DH5 α	Clone strain, F ⁻ <i>lacZ</i> Δ M15 Δ (<i>lacZYA</i> -argF) relA1	TakaRa
<i>E. coli</i> BL21(DE3)	F ⁻ , ompT, hsdS (rBB-mB), gal, dcm (DE3)	TakaRa
BL21-pSY7-lacZ-635s	Derived from BL21(DE3), carrying plasmid pSY7-lacZ-635s	This lab ^[21]
BL21-pSY7-lacZ-1063a	Derived from BL21(DE3), carrying plasmid pSY7-lacZ-1063a	This lab ^[21]
BL21-pSY7-M1-lacZ-635s	Derived from BL21(DE3), carrying plasmid pSY7-M1-lacZ-635s	This study
BL21-pSY7-M2-lacZ-635s	Derived from BL21(DE3), carrying plasmid pSY7-M2-lacZ-635s	This study
BL21-pSY7-M1-lacZ-1063a	Derived from BL21(DE3), carrying plasmid pSY7-M1-lacZ-1063a	This study
BL21-pSY7-M2-lacZ-1063a	Derived from BL21(DE3), carrying plasmid pSY7-M2-lacZ-1063a	This study
Plasmids		
pMD19T	TA clone vector, Amp ^r	TakaRa
pMD19T-IEP	pMD19T ligated with part IEP sequence	This study
pMD19T-M1-IEP	Derived from pMD19T-IEP, IEP C164K mutation	This study
pMD19T-M2-IEP	Derived from pMD19T-IEP, IEP G214W mutation	This study
pSY7	Derived from pSY6, <i>lacI</i> , T7 promoter, Amp ^r	This lab ^[21]
pSY7-lacZ-635s	Derived from pSY7, targeting the sense strand 635 site of <i>lacZ</i> in BL21(DE3)	This lab ^[21]
pSY7-lacZ-1063a	Derived from pSY7, targeting the antisense strand 1063 site of <i>lacZ</i> in BL21(DE3)	This lab ^[21]
pSY7-M1-lacZ-635s	Derived from pSY7-lacZ-635s, IEP C164K site mutation	This study
pSY7-M2-lacZ-635s	Derived from pSY7-lacZ-635s, IEP G214W site mutation	This study
pSY7-M1-lacZ-1063a	Derived from pSY7-lacZ-1063a, IEP C164K site mutation	This study
pSY7-M2-lacZ-1063a	Derived from pSY7-lacZ-1063a, IEP G214W site mutation	This study

1.3 定点突变及突变型 Targetron 载体构建

以质粒 pSY7 为模板, IEP-F/IEP-R 为引物, 利用 TaqDNA 聚合酶扩增获得 1.2 kb IEP 蛋白的基因序列(包含拟突变核苷酸位点), 将所获 PCR 产物与 pMD19T 载体连接获得 pMD19T-IEP 载体。

以 pMD19T-IEP 为模板, 用定点突变引物扩增整个质粒。扩增条件为: 95 °C 2 min; 95 °C 30s, 55 °C 30 s, 72 °C 4 min, 循环 15 次; 72 °C 5 min。反应结束后, PCR 产物用 *Dpn* I 消化除去甲基化的质粒模板; 然后, 转化 *E. coli* DH5 α , 测序获得突变型 pMD19T-M-IEP 质粒。

用 *Nhe* I 和 *Bgl* II 双酶切 pMD19T-M-IEP 质粒, 回收目的片段, 与经同样双酶切的 pSY7-lacZ-

635s、pSY7-lacZ-1063a 载体连接, 获得内含子编码蛋白反转录催化位点突变的 Targetron 基因打靶载体 pSY7-M-lacZ-635s 和 pSY7-M-lacZ-1063a, 详细构建流程如图 2 所示。

1.4 遗传筛选及打靶效率计算

将质粒 pSY7-lacZ-635s、pSY7-lacZ-1063a 及 IEP 蛋白突变的质粒 pSY7-M-lacZ-635s、pSY7-M-lacZ-1063a 分别转化 *E. coli* BL21(DE3), 在含 100 μ g/mL 氨苄青霉素的 LB 液体培养基中 37 °C、200 r/min 过夜培养。40 μ L 的过夜培养物转接到 4 mL 含氨苄青霉素的 LB 培养基中 37 °C、200 r/min 培养 1 h, 加入 IPTG 使其终浓度为 0.5 mmol/L, 于 37 °C、200 r/min 诱导培养 45 min,

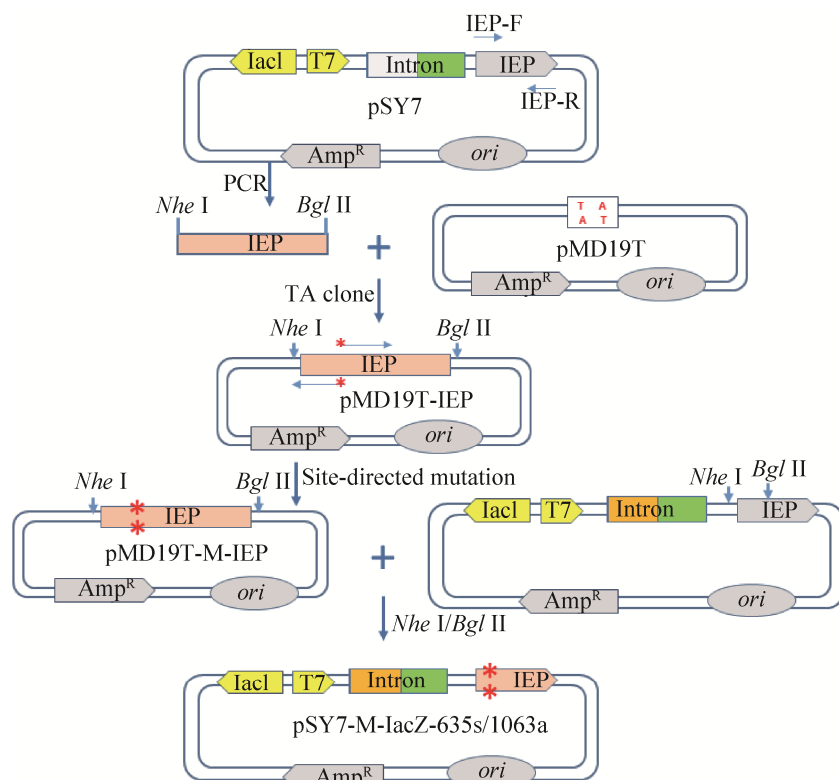


图 2. 突变型 Targetron 载体构建流程

Figure 2. Flow chart of mutant Targetron vectors construction. The red asterisk indicates the mutation sites.

离心收集菌体，以不同的稀释浓度涂布 LB 固体培养基(含氨苄青霉素 100 $\mu\text{g}/\text{mL}$ 、X-gal 40 $\mu\text{g}/\text{mL}$ 、IPTG 0.1 mmol/L)，置于 37 $^{\circ}\text{C}$ 过夜培养，通过蓝白斑计数计算其“归巢”效率，“归巢”效率=白斑/(白斑+蓝斑) $\times 100\%$ 。

1.5 菌落 PCR 鉴定

分别挑取蓝色菌落和白色菌落，以 lacZ-F/lacZ-R 为引物进行菌落 PCR 扩增，检测内含子 RNA 是否特异性插入到 DNA 靶位点，进一步验证 lacZ 基因突变情况。

2 结果和分析

2.1 内含子编码蛋白同源序列分析及反转录核心催化位点筛选

以乳酸乳球菌 L1.LtrB 的 II 型内含子编码蛋

白氨基酸序列为模板，利用同源序列比对分析，结合其蛋白结构解析，筛选到 13 个可能影响内含子编码蛋白反转录功能的核心催化位点，分别是：D160、I161、K162、G163、C164、F165、Q213、G214、Y306、D308、D309、L357、G358(图 3 亮黄色背景显示序列)，其中，预测 D160、I161、K162、G163、C164、F165、Q213、D308 可能与底物 NTP 结合，G214 可能与染色体上的核酸结合。

2.2 C164 和 G214 在 IEP 蛋白中的相互作用分析

IEP 蛋白的空间结构已被解析，将 IEP 蛋白氨基酸序列提交至 <https://swissmodel.expasy.org/> 检索获得 IEP 蛋白晶体结构，利用 PyMOL 软件分析得到 IEP 三维结构图(图 4)。本研究分别选择与底物 NTP 结合的 C164 残基和与 DNA 核酸结合的 G214 残基为靶点，分别将其定点突变为

Feature 1		####	
gi_44004428	101	KRVFIPK. [5]. RPLGI. [2]. IEDRIIQMMKQVLEPVLEAQF. [3]. SFGFRPERTTYHALE. [11]. HWIVEGDIRQ 180	
query	85	RRMYIAK. [6]. RPLGI. [2]. FTDKLIQEAIVRILESIYEPVF. [3]. SHGFRPQRSCHTALK. [9]. RWFVEGDIKG 163	
gi_22536745	96	KRVYIPK. [5]. RPLGI. [2]. VRDRVVQTAVKIVIEPIFEADF. [3]. SYGFRPKRSANQAIR. [10]. EWVIDADLKG 174	
gi_81096456	84	RRTYIPK. [5]. RPLGI. [2]. FDDKLLQVQVIMILESIYEQF. [3]. SHGFRPNKSCHSALT. [9]. KWFIEGDIKS 161	
gi_89892796	74	RRTYIPK. [5]. RPLGI. [2]. FDDKLVQECVRLLEAVYEGSF. [3]. SHGFRPNHSCHTALS. [9]. KWFVEGDIKG 151	
gi_2495490	82	RRVYIPK. [5]. RPLGI. [2]. LRDRIVQRAMLMEPIWESDF. [3]. SYGFRPERSVHHAIR. [14]. RWVIEGDLS 164	
gi_7545209	243	RRVYIPK. [5]. RPLGI. [2]. PRDKIVQEVFRAILEQVLEPRF. [3]. SHGFRPGRGCHSALA. [8]. KWFIEGDIKS 319	
gi_6746623	272	RRVEIPK. [5]. RPLGV. [2]. PRDKIVQKALHAILEAIFEPLF. [3]. SHGFRPNRSTHSALL. [9]. NWVIQGDITK 349	
gi_113939021	101	RRVYIPK. [5]. RPLGI. [2]. IKDRVVQEMVRLILDPIYESTF. [3]. SYGFRPYRATHHAVV. [11]. QMALEGDIRA 180	
gi_68231786	165	RRVHIPK. [5]. RPLGL. [2]. WSDKLVGEVVRLLLEAYEPTF. [3]. SHGFRPRRGCHTALR. [9]. AWFIEGDITD 242	
Feature 1	##	##	##
gi_44004428	181	FFDKVNHRIKIKL. [31]. GTPQGG. [1]. LSPLLANVYLHSD. [35]. LIRYADDVWLVTNN. [5]. RWKTVIKNFL 310	
query	164	CFDNIDHVTILIGLI. [32]. GTPQGG. [1]. LSPLLANIYLHELD. [72]. YVRYADDFIISVKG. [6]. WIKEQLKFLI 332	
gi_22536745	175	YFDTIPHDKLLLV. [32]. GTPQGG. [1]. ISPLLANIYLNALD. [16]. LIRYADDFVILCSN. [3]. KYYQYAKQRT 284	
gi_81096456	162	FFDNINHDVIIQIL. [32]. GTPQGG. [1]. ISPILANIYLDKFD. [76]. YTRYADDFICGVIG. [6]. RIKADIKDYL 334	
gi_89892796	152	FFDNINHEVMIGIL. [32]. GTPQGG. [1]. ISPILANIYLDKLD. [76]. YVRYADDFLIGVIG. [6]. AIKEQVKAFV 324	
gi_2495490	165	YFDTVHHRLLMKAV. [32]. GVPQGG. [1]. ISPLLSNIMLNEFD. [42]. YCRYADDFVLIVKG. [6]. AIREECRGVL 303	
gi_7545209	320	FFDNIDHHILEKLL. [32]. GVPQGG. [1]. ASPILSNLVNELD. [81]. YVRYADDFWLVGAG. [6]. AIKERIAAYL 497	
gi_6746623	350	CFDSIPHSIILKRI. [34]. GTPQGG. [1]. LSPILSNIVLHEFD. [62]. YVRYADDFVVFISG. [6]. FIRNNDKDYL 510	
gi_113939021	181	CFDRIHHTLIRIL. [32]. GTPQGG. [1]. VSPLLANIYLNELD. [28]. ITRYADDFVLLHG. [6]. TLKTALATFL 305	
gi_68231786	243	CFGSLNHDLMIGIL. [32]. GAPQGG. [1]. ASPILSSIYLHKLD. [66]. YCRYADDFHLLGFTG. [6]. EIRQLAEFL 405	
Feature 1		##	
gi_44004428	311.	[2]. ELKLELSEBKTRI. [7]. IEPFLGFKY 340	
query	333.	[2]. KLKMELESEKTLI. [6]. ARFLGYDI 361	
gi_22536745	285.	[1]. KLGLTLNEEKTRI. [6]. FDFLGYTL 312	
gi_81096456	335.	[2]. VLKLELSEKTLI. [6]. ARFLGYHL 363	
gi_89892796	325.	[2]. MLILELSEKTLI. [6]. ARFLGYDI 353	
gi_2495490	304.	[2]. SLKLRNMDKTKI. [6]. FIFLGHRL 332	
gi_7545209	498.	[2]. ILKLELSEKTLI. [7]. AYFLGTBI 527	
gi_6746623	511.	[2]. NCGLELNVDKTAL. [7]. WKFLGABL 540	
gi_113939021	306.	[2]. HLHLELSEKTLI. [6]. FDFLGFHI 334	
gi_68231786	406.	[2]. DLKLELSEKTLI. [7]. ARFLGYBI 435	

图 3. 内含子编码蛋白同源序列分析

Figure 3. Analysed of homology sequence with intron encoded protein. The yellow background are key amino acids which related to the activity of reverse transcriptase, the grey number indicates the number of amino acids omitted.

C164K 和 G214W, 研究此两个位点突变对 II 型内含子反转录功能的影响。

然后, 我们分别将 C164K 和 G214W 两个位点突变的 IEP 蛋白氨基酸序列提交至 <https://swissmodel.expasy.org/>, 以野生型 IEP 蛋白为模板进行同源建模, 分别获得 C164K 和 G214W 突变的 IEP 蛋白三维结构, 利用 Swiss-PdbViewer 软件分析 C164 和 G214 位点突变前后空间结构的变化(图 4)。从空间结构上看, C164 与 Y208 通过氢键相互作用, 当 C164 突变后, C164K 仍与 Y208 保持氢键的连接, 但其氢键的距离在空间结构上较突变前缩短 0.11 Å; G214 与 T211、L217、S218 通过氢键相互

作用, 当 G214 突变后失去与 L217 和 S218 的氢键相互作用。因此从结构上推测 C164 和 G214 两个位点可能影响 IEP 蛋白反转录结构域的空间结构。

2.3 IEP 蛋白反转录结构域关键氨基酸位点突变对内含子“归巢”效率的影响

以筛选到的 IEP 蛋白 C164 和 G214 为靶位点, 利用定点突变方法构建 IEP 蛋白突变体, 构建流程如图 2 所示。将野生型 Targetron 载体及突变型 Targetron 载体分别转化 *E. coli* BL21(DE3), 通过菌落 PCR 检测(图 5)及蓝白斑筛选(图 6), 分析内含子编码蛋白反转录结构域关键氨基酸位点突变对 II 型内含子“归巢”的影响。菌落 PCR

结果显示, 经过诱导表达后, 野生型 II 型内含子可通过“归巢”插入到 DNA 靶位点, 突变型 II 型内含子不能插入 DNA 靶位点(图 5)。前期的统计结果表明, 野生型 II 型内含子在 lacZ-635s 及 lacZ-1063a 位点的归巢效率分别为 $90.845\% \pm 6.792\%$ 、 $92.582\% \pm 2.898\%$ ^[21]; 然而, 当 IEP 蛋白反转录结

构域中两个关键催化位点(C164、G214)突变后, 其归巢效率均为 0%, 归巢功能完全丧失(图 6)。上述结果表明 C164 和 G214 两个位点是 IEP 蛋白反转录结构域关键催化位点, 该位点突变完全失活了 IEP 蛋白的反转录功能, II 型内含子不能通过反转录合成 cDNA 而插入到靶位点。

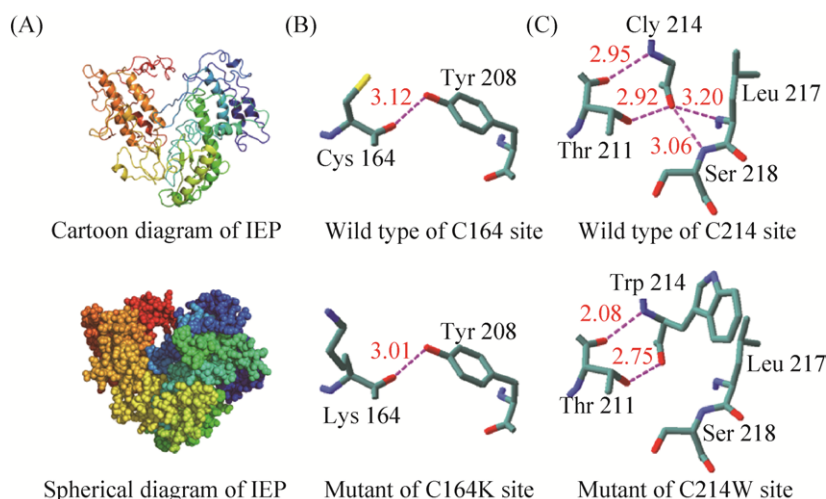


图 4. IEP 候选位点野生型与突变型的空间结构图

Figure 4. Spatial structure of wild and mutant IEP candidate sites. A: Cartoon and spherical diagram of the three-dimensional structure of IEP; B: The three-dimensional structure of C164 wild type and mutant type, the dotted line in purple represents hydrogen bonds, the red number indicates the hydrogen bond distance; C: The three-dimensional structure of G214 wild type and mutant type, the dotted line in purple represents hydrogen bonds, the red number indicates the hydrogen bond distance.

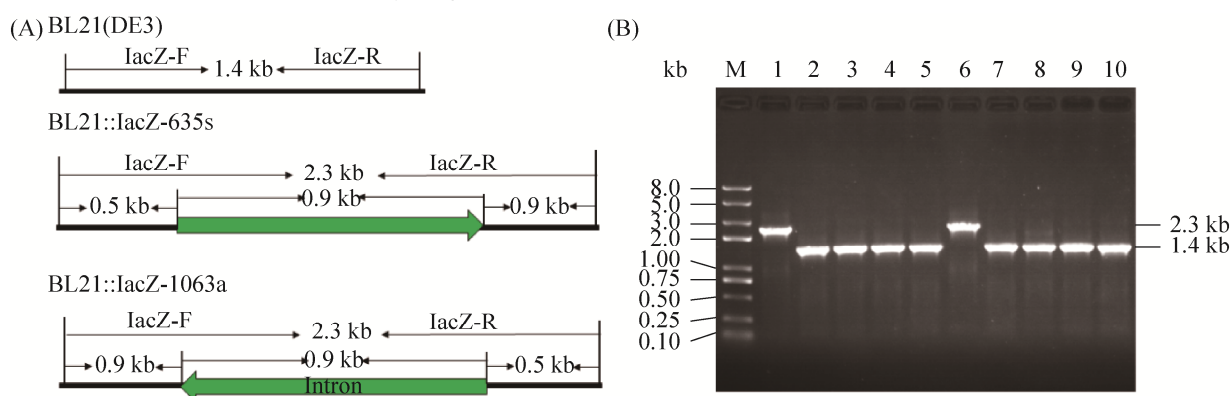


图 5. 突变型 II 型内含子功能验证电泳图

Figure 5. Functional verification of mutant group II introns. A: Diagram of primer binding sites. B: Functional verification electrophoresis of mutant group II intron; M: 8000 bp DNA marker; 1: White colony of BL21-pSY7-lacZ-635s; 2: Blue colony of BL21-pSY7-lacZ-635s; 3: Blue colony of BL21-pSY7-M1-lacZ-635s; 4: Blue colony of BL21-pSY7-M2-lacZ-635s; 5: BL21(DE3) ; 6: White colony of BL21-pSY7-lacZ-1063a; 7: Blue colony of BL21-pSY7-lacZ-1063a; 8: Blue colony of BL21-pSY7-M1-lacZ-1063a; 9: Blue colony of BL21-pSY7-M2-lacZ-1063a; 10: BL21(DE3).

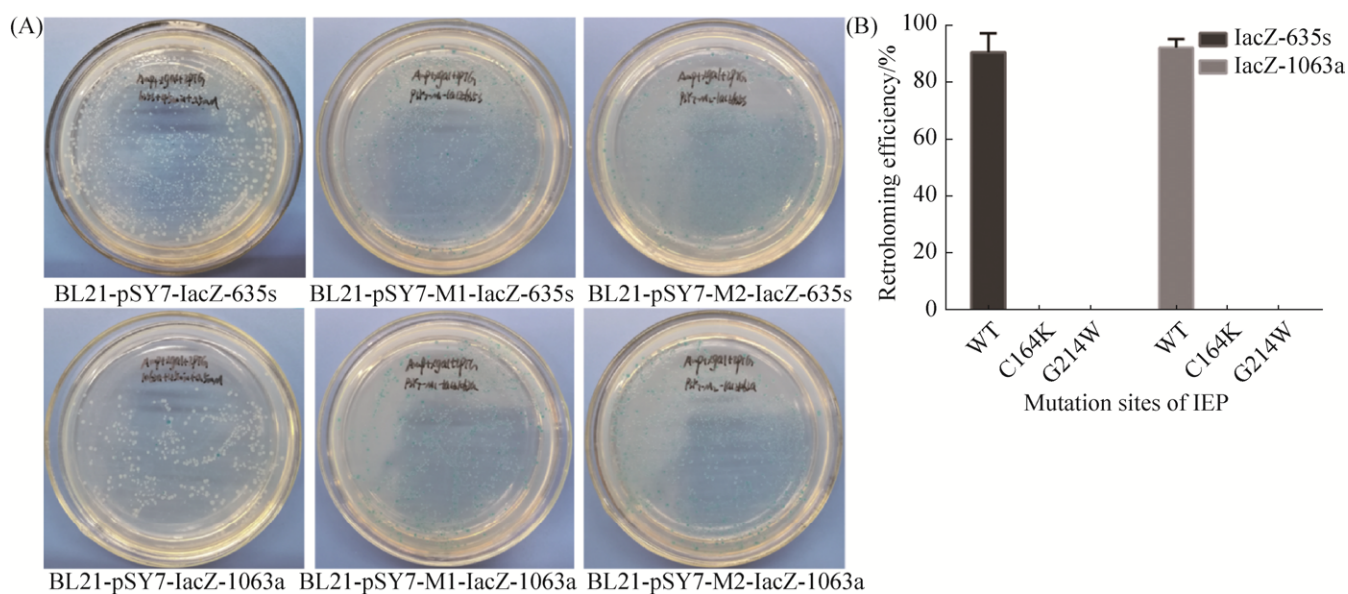


图 6. 突变型 II 型内含子对归巢效率的影响

Figure 6. Effect of mutant group II introns on retrohoming efficiency. A: Blue and white spot screening; B: Comparison of retrohoming efficiency between wild and mutant group II introns.

3 讨论

本研究中, 我们筛选到影响 II 型内含子“归巢”的两个关键氨基酸位点(C164 和 G214), 并通过定点突变和体内功能验证, 确定此两个位点的突变完全失活了 II 型内含子“归巢”的功能。此外, 由于 IEP 蛋白是由反转录结构域、成熟酶结构域、DNA 结合域、核酸内切酶结构域四个不同的部分组成的多功能蛋白^[1,6,22], 且不同结构域在空间上和功能上相对独立^[1], 因此, IEP 蛋白反转录结构域关键催化位点的定点突变, 理论上不会影响其他结构域的功能, 即: 反转录结构域突变, 只可能影响 IEP 蛋白的反转录功能, 不会影响其与 DNA 结合及核酸内切酶的功能。因此, 本研究获得的突变体, 理论上能够将 II 型内含子识别、切割和反转录的连续过程加以分离。

在后续研究中, 如果能够获得上述 IEP 蛋白

突变体与内含子 RNA、底物 NTP 和 DNA 复合体的结晶体, 我们将有可能从其空间结构上进一步解析 II 型内含子不同过程的作用机理, 为深入研究 II 型内含子的“归巢”机制和广泛应用提供理论模型。

参考文献

- [1] Lambowitz AM, Zimmerly S. Mobile group II introns. *Annual Review of Genetics*, 2004, 38: 1–35.
- [2] Pyle AM. Group II intron self-splicing. *Annual Review of Biophysics*, 2016, 45: 183–205.
- [3] Zimmerly S, Semper C. Evolution of group II introns. *Mobile DNA*, 2015, 6: 7.
- [4] Zhao C, Pyle AM. The group II intron maturase: a reverse transcriptase and splicing factor go hand in hand. *Current Opinion in Structural Biology*, 2017, 47: 30–39.
- [5] Lambowitz AM, Zimmerly S. Group II introns: mobile ribozymes that invade DNA. *Cold Spring Harbor Perspectives in Biology*, 2011, 3(8): a003616.
- [6] McNeil BA, Semper C, Zimmerly S. Group II introns:

- versatile ribozymes and retroelements. *Wiley Interdisciplinary Reviews: RNA*, 2016, 7(3): 341–355.
- [7] Toro N, Molina-Sánchez MD, Nisa-Martínez R, Martínez-Abarca F, García-Rodríguez FM. Bacterial group II introns: identification and mobility assay. *Methods in Molecular Biology*, 2016, 1400: 21–32.
- [8] Lambowitz AM, Belfort M. Mobile bacterial group II introns at the crux of eukaryotic evolution. *Microbiology Spectrum*, 2015, 3(1): MDNA3-0050-2014.
- [9] Enyeart PJ, Mohr G, Ellington AD, Lambowitz AM. Biotechnological applications of mobile group II introns and their reverse transcriptases: gene targeting, RNA-seq, and non-coding RNA analysis. *Mobile DNA*, 2014, 5: 2.
- [10] Liu YJ, Zhang J, Cui GZ, Cui Q. Current progress of targetron technology: development, improvement and application in metabolic engineering. *Biotechnology Journal*, 2015, 10(6): 855–865.
- [11] Yao J, Lambowitz AM. Gene targeting in gram-negative bacteria by use of a mobile group II intron (“Targetron”) expressed from a broad-host-range vector. *Applied and Environmental Microbiology*, 2007, 73(8): 2735–2743.
- [12] Yao J, Zhong J, Lambowitz AM. Gene targeting using randomly inserted group II introns (targetrons) recovered from an *Escherichia coli* gene disruption library. *Nucleic Acids Research*, 2005, 33(10): 3351–3362.
- [13] Karberg M, Guo HT, Zhong J, Coon R, Perutka J, Lambowitz AM. Group II introns as controllable gene targeting vectors for genetic manipulation of bacteria. *Nature Biotechnology*, 2001, 19(12): 1162–1167.
- [14] Francis MB, Allen CA, Shrestha R, Sorg JA. Bile acid recognition by the *Clostridium difficile* germinant receptor, CspC, is important for establishing infection. *PLoS Pathogens*, 2013, 9(5): e1003356.
- [15] Heap JT, Pennington OJ, Cartman ST, Carter GP, Minton NP. The ClosTron: a universal gene knock-out system for the genus *Clostridium*. *Journal of Microbiological Methods*, 2007, 70(3): 452–464.
- [16] Chen Y, McClane BA, Fisher DJ, Rood JI, Gupta P. Construction of an alpha toxin gene knockout mutant of *Clostridium perfringens* type A by use of a mobile group II intron. *Applied and Environmental Microbiology*, 2005, 71(11): 7542–7547.
- [17] Yao J, Zhong J, Fang Y, Geisinger E, Novick RP, Lambowitz AM. Use of targetrons to disrupt essential and nonessential genes in *Staphylococcus aureus* reveals temperature sensitivity of LtrB group II intron splicing. *RNA*, 2006, 12(7): 1271–1281.
- [18] Frazier CL, San Filippo J, Lambowitz AM, Mills DA. Genetic manipulation of *Lactococcus lactis* by using targeted group II introns: generation of stable insertions without selection. *Applied and Environmental Microbiology*, 2003, 69(2): 1121–1128.
- [19] Truong DM, Hewitt FC, Hanson JH, Cui XX, Lambowitz AM. Retrohoming of a mobile group II intron in human cells suggests how eukaryotes limit group II intron proliferation. *PLoS Genetics*, 2015, 11(8): e1005422.
- [20] Qu GS, Kaushal PS, Wang J, Shigematsu H, Piazza CL, Agrawal RK, Belfort M, Wang HW. Structure of a group II intron in complex with its reverse transcriptase. *Nature Structural & Molecular Biology*, 2016, 23(6): 549–557.
- [21] Chen XH, Liu F, Wang CX, Chen ZH, Hong W, Cai MD, Zhang ZR, Qi TN, Liao YH, Gu JY, Cui GZ. Construction of highly efficient and rigorous targetron system in *Escherichia coli*. *Biotechnology Bulletin*, 2019, 35(6): 213–220. (in Chinese)
- 陈相好, 刘芳, 王彩霞, 陈峥宏, 洪伟, 蔡梦迪, 张峥嵘, 綦廷娜, 廖永慧, 谷俊莹, 崔古贞. 高效严谨型大肠杆菌 Targetron 基因打靶系统的构建. *生物技术通报*, 2019, 35(6): 213–220.
- [22] Zimmerly S, Guo HT, Perlman PS, Lambowitz AM. Group II intron mobility occurs by target DNA-primed reverse transcription. *Cell*, 1995, 82(4): 545–554.

Key catalytic sites in the reverse transcription domain of Ll.LtrB intron encoded protein

Xianghao Chen^{1,2}, Zhengrong Zhang^{1,3}, Fang Liu¹, Zhenghong Chen^{1,3}, Wei Hong⁴, Tingna Qi^{1,3}, Junying Gu^{2,5*}, Guzhen Cui^{1,3*}

¹ School of Basic Medical Science, Guizhou Medical University, Guiyang 550025, Guizhou Province, China

² School of Clinical Laboratory Science, Guizhou Medical University, Guiyang 550004, Guizhou Province, China

³ Key Laboratory of Medical Microbiology and Parasitology of Education Department of Guizhou, Guiyang 550025, Guizhou Province, China

⁴ Key Laboratory of Molecular Biology, Guizhou Medical University, Guiyang 550004, Guizhou Province, China

⁵ Affiliated Hospital of Guizhou Medical University, Guiyang 550004, Guizhou Province, China

Abstract: [Objective] To screen the key catalytic sites that affect the reverse transcription function of intron-encoded protein (IEP) from Ll.LtrB, and to obtain the IEP mutant without reverse transcription activity.

[Methods] The key catalytic sites affecting the reverse transcription of IEP were screened by sequence alignment and homology modeling methods. Then, the screened key catalytic sites were subjected to site-directed mutagenesis. The mutated IEP was combined with the Targetron vector to construct the mutated Targetron targeting system without reverse transcription function. Finally, the function of mutant IEP was verified by using the *lacZ* gene in *Escherichia coli*. [Results] The sites C164 and G214 of IEP were screened and mutated, completely inactivated the “retrohoming” function *in vivo*. [Conclusion] The IEP mutants of Ll.LtrB obtained in our study laid a solid foundation for further research on the structure and mechanism of group II intron.

Keywords: Ll.LtrB, group II intron, intron encode protein, reverse transcription, Targetron

(本文责编: 张晓丽)

Supported by the National Natural Science Foundation of China (31760318, 31500078, 31560318, 31601012), by the Science and Technology Plan Program of Guizhou Province ([2018]1132, [2019]1441), by the Natural Science Research Project of Guizhou Education Department ([2014]216) and by the Project of Guizhou Graduate Research Fund (11348)

*Corresponding author. Tel: +86-851-88174015, E-mail: Junying Gu, 921469044@qq.com, Guzhen Cui, cuiguzhen@hotmail.com

Received: 28 January, 2019; Revised: 4 April 2019; Published online: 22 April 2019