



Universidade de Brasília
Instituto de Ciências Biológicas
Departamento de Biologia Celular
Programa de Pós-graduação em Biologia Molecular

**Desenvolvimento de uma plataforma de metilação de
DNA *in vivo* para transformação de *Clostridium
acetobutylicum***

Catarina Vargas Cunha

Orientador: Prof. Fernando Araripe Gonçalves Torres

**Brasília
Julho de 2019**

Universidade de Brasília
Instituto de Ciências Biológicas
Departamento de Biologia Celular
Programa de Pós-graduação em Biologia Molecular

**Desenvolvimento de uma plataforma de metilação de
DNA *in vivo* para transformação de *Clostridium
acetobutylicum***

Catarina Vargas Cunha

Orientador: Prof. Fernando Araripe Gonçalves Torres

*Dissertação de mestrado apresentada ao Programa
de Pós-Graduação em Biologia Molecular da
Universidade de Brasília como requisito parcial para
obtenção do título de mestre em Biologia Molecular.*

**Brasília
Julho de 2019**

Trabalho realizado no Laboratório de Biologia Molecular, Departamento de Biologia Celular, Instituto de Ciências Biológicas, Universidade de Brasília, sob orientação do professor Fernando Araripe Gonçalves Torres e com apoio financeiro do CNPq e CAPES.

Banca Examinadora

Professor Dr. Fernando Araripe Gonçalves Torres
Universidade de Brasília
Presidente da Banca

Professora Dra. Cintia Marques Coelho
Universidade de Brasília
Membro titular

Professora Dra. Eliane Ferreira Noronha
Universidade de Brasília
Membro titular

Professora Dra. Lídia Maria Pepe de Moraes
Universidade de Brasília
Membro suplente

*“Great things never came
from comfort zone.”*

AGRADECIMENTOS

À Deus por ter me dado força para chegar até aqui.

Aos meus pais por todo o incentivo que tive dentro de casa, por nunca me deixarem desistir e, principalmente, por serem um enorme exemplo para mim.

À minha irmã Isadora, por todo o apoio e companheirismo durante as noites em claro.

Ao meu namorado Rafael por sempre me tranquilizar nos momentos de ansiedade.

Ao professor Fernando Araripe pelo prestimoso auxílio e orientação. Por estar sempre presente e por ter me dado a chance de ingressar no programa de pós-graduação e retornar à academia sem me conhecer.

À Myrna por ter me ajudado em todos os momentos difíceis. Por ter sido sempre tão prestativa e cuidadosa. Por ter me ensinado tanto e me conduzido em cada etapa até aqui, e pela amizade sincera.

À professora Lídia por todas as recomendações e pelo conhecimento transmitido.

À professora Janice pela preocupação e carinho.

À todos os amigos do Lab2 por toda a ajuda, pelo convívio e pelo acolhimento:
Vivis, Chris, Vanessa, Roberta, Igor, Francisco, Thiago, Pedro e Júlia.

À Lu pelo auxílio e preocupação. Obrigada por ter me ajudado na correção e por ter me orientado várias vezes até aqui.

À Ana Laura pela disposição em ajudar e por ter me dado sempre as melhores dicas.

À Yasmin por ter sido tão companheira. Por toda a animação do dia a dia. Por alegrar o laboratório e me fazer sentir em casa.

Aos membros da banca examinadora pela disponibilidade e paciência.

Ao CNPq pelo apoio financeiro.

À todos que de alguma forma me ajudaram na realização deste trabalho.

ÍNDICE

ÍNDICE DE FIGURAS.....	i
ÍNDICE DE TABELAS.....	ii
LISTA DE ABREVIATURAS.....	iii
RESUMO.....	v
ABSTRACT.....	vi
1. INTRODUÇÃO.....	1
1.1. Clostrídios.....	1
1.1.1. <i>Clostridium acetobutylicum</i>	1
1.1.2. Busca por novas fontes de energia renovável: Fermentação ABE.....	1
1.1.3. Melhoramentos genéticos em <i>C. acetobutylicum</i>	2
1.1.4. Sistema de restrição de <i>C. acetobutylicum</i>	3
1.1.5. Suplantando o sistema de restrição de <i>C. acetobutylicum</i>	4
1.2. <i>Escherichia coli</i>	5
1.2.1. Expressão heteróloga em <i>E. coli</i>	5
1.2.2. Linhagens de <i>E. coli</i>	9
1.2.3. Técnicas moleculares para integração no genoma de <i>E. coli</i>	10
1.2.3.1. Reparo por HDR (<i>Homology Direct Repair</i>)	10
1.2.3.2. Utilização do sistema CRISPR/Cas9.....	12
1.2.4. Construção de uma plataforma de metilação em <i>E. coli</i>	14
2. OBJETIVO	16
2.1. Objetivo geral.....	16
2.2. Objetivos específicos.....	16
3. ESTRATÉGIA.....	17
4. MATERIAIS E MÉTODOS.....	18
4.1. Materiais.....	18
4.1.1. Bactérias.....	18
4.1.2. Meios de cultura.....	18
4.1.3. Soluções.....	20
4.1.4. Enzimas.....	23
4.1.5. Marcadores de massa molecular para DNA.....	25
4.1.6. Kits.....	25
4.1.7. Oligonucleotídeos (<i>primers</i>)	26

4.1.8.	Vetores.....	27
4.1.9.	Ferramentas de bioinformática.....	27
4.2.	Métodos.....	28
4.2.1.	Cultivo de micro-organismos.....	28
4.2.2.	Amplificação por PCR.....	29
4.2.3.	Preparação de células bacterianas competentes.....	29
4.2.4.	Manipulação de DNA.....	31
4.2.5.	Preparação de DNA plasmidial.....	32
4.2.6.	Extração de DNA genômico de bactérias.....	34
4.2.7.	Transformação de bactérias.....	34
4.2.8.	Ativação da endonuclease Cas9 quando utilizado o cassete de integração ligado ao vetor pBluescript SK (pBlue- $\phi 3Tim$).....	37
4.2.9.	PCR de colônia.....	37
4.2.10.	TAIL-PCR e sequenciamento dos fragmentos obtidos.....	38
4.2.11.	Otimização do gene $\phi 3Ti$	39
5.	RESULTADOS E DISCUSSÃO.....	40
5.1.	Construção de fragmentos de DNA	40
5.2.	Construção de plasmídeos.....	46
5.3.	Transformação de <i>E. coli</i> DH10B.....	49
5.4.	Confirmação da integração cromossomal do cassete de integração.....	54
5.5.	Confirmação do local de integração do fragmento <i>spec-$\phi 3Tim$</i>	55
5.5.1.	PCR da região alvo do sistema CRISPR/Cas9	55
5.5.2.	Utilização da técnica TAIL PCR e sequenciamento.....	58
5.6.	Remoção do gene que confere resistência a espectinomicina.....	65
5.7.	Análise da funcionalidade de $\phi 3Ti$ -MTP integrada em <i>E. coli</i>	67
5.7.1.	Análise <i>in vitro</i>	67
5.7.2.	Análise <i>in vivo</i>	79
6.	CONCLUSÕES E PERSPECTIVAS.....	72
7.	REFERÊNCIAS.....	73

ÍNDICE DE FIGURAS

Figura 1. Sistema RecA.....	11
Figura 2. Mecanismo de atuação do sistema CRISPR/Cas9. de <i>S. pyogenes</i>	13
Figura 3. Estratégia para a construção de um sistema de metilação em <i>E. coli</i>	17
Figura 4. Esquema do fragmento <i>spec-φ3Tim</i>	41
Figura 5. Estrutura da molécula de RNA guia.....	42
Figura 6. Sequência completa do fragmento que codifica para o sgRNA.....	44
Figura 7. Representação esquemática da construção do sgRNA.....	45
Figura 8. Amplificação do fragmento que expressa o sgRNA.....	45
Figura 9. Confirmação da integração do sgRNA em AraC_pBAD_Cas9.....	47
Figura 10. Mapa físico do vetor AraC_pBAD_CRISPR.....	47
Figura 11. Mapa físico do vetor pBlue- ϕ 3Tim.....	48
Figura 12. Análise de transformantes de <i>E. coli</i> com o cassete de integração linear.....	51
Figura 13. Análise de transformantes de <i>E. coli</i> com o vetor pBlue- Φ 3Tim.....	53
Figura 14. Análise da integração por PCR com DNA genômico de transformantes.....	55
Figura 15. PCR de colônia de clones de <i>E. coli</i>	56
Figura 16. Estratégia para determinar o local de integração (1ª etapa)	59
Figura 17. Amplicons obtidos após a realização da primeira etapa de TAIL-PCR.....	60
Figura 18. Estratégia para determinar o local de integração (2ª etapa)	60
Figura 19. Amplicons obtidos após a realização da segunda etapa de TAIL-PCR.....	61
Figura 20. Estratégia para determinar o local de integração (3ª etapa)	62
Figura 21. Amplicons obtidos após a realização da terceira etapa do processo.....	63
Figura 22. Resultado das regiões no DNA de <i>E. coli</i> reveladas por sequenciamento.....	63
Figura 23. Cepas <i>E. coli</i> recombinadas com o vetor 705-Cre.....	66
Figura 24. Amplicons obtidos PCR de cepas recombinadas com 705-Cre.....	66
Figura 25. Análise da digestão de PACYC184 com SatI.....	67
Figura 26. Mapa físico do vetor pACYC184.....	68
Figura 27. Mapa físico do vetor pMTL-mazF-PTA.....	69
Figura 28. Resultado da transformação de <i>C. acetobutylicum</i>	70

ÍNDICE DE TABELAS

Tabela 1. Enzimas de restrição utilizadas.....	23
Tabela 2. <i>Primers</i> utilizados.....	26
Tabela 3. Vetores utilizados.....	27
Tabela 4. Ferramentas de bioinformática utilizadas.....	28
Tabela 5. Sequências usadas para a montagem do sgRNA.....	44

LISTA DE ABREVIATURAS

CRISPR	Repetições palindrômicas curtas agrupadas e regularmente interespaçadas
crRNA	CRISPR RNA
DNA	ácido desoxirribonucleico
EDTA	ácido etileno-diamino-tetra-acético
g	grama
<i>g</i>	força gravitacional
HDR	<i>homology direct repair</i>
kb	quilobase
L	litro
LB	Luria-Bertani medium
M	molar
mg	miligrama
mL	mililitro
mM	milimolar
ng	nanograma
OD ₆₀₀	densidade celular óptica medida a 600nm
PAM	motivo adjacente proto-espaçador
pb	par de base
PCR	reação em cadeia de polimerase
pH	potencial hidrogeniônico
p/v	peso por volume
TAIL-PCR	PCR entrelaçada térmica assimétrica

REase	ribonuclease A
RM	restrição e modificação
RNA	ácido ribonucleico
sgRNA	<i>single guide</i> RNA
TAE	tampão tris-acetato-EDTA
tracrRNA	<i>trans-activating</i> CRISPR RNA
s	segundo
V	volts
v/v	volume por volume
YM	<i>yeast extract minimal medium</i>
°C	graus Celsius
μF	microfaraday
μg	micrograma
μ L	microlitro
Ω	Ohm

RESUMO

Bactérias pertencentes à espécie *Clostridium acetobutylicum* têm sido amplamente utilizadas para a produção de solventes por realizarem um processo fermentativo conhecido como fermentação ABE (*Acetone/Butanol /Ethanol*). Através desta via, este micro-organismo é capaz de produzir acetona, butanol e etanol a partir de carboidratos simples. Dessa forma, *C. acetobutylicum* é considerado um micro-organismo promissor na produção de bioálcoois, sobretudo, de butanol em escala industrial. Diferentes cepas de *C. acetobutylicum* já foram estudadas e exibiram capacidade de produzir solventes, no entanto, bactérias deste gênero apresentam resistência à manipulação genética decorrente de um sistema endógeno de restrição capaz de clivar sequências de DNA não metilado, impedindo o melhoramento genético desses organismos. O presente estudo teve o objetivo de desenvolver uma cepa de *Escherichia coli* expressando a metiltransferase $\phi 3Ti$ -MTP, isolada do fago $\phi 3T$ de *Bacillus subtilis*, que se mostrou eficiente para a proteção de sequências de DNA contra a ação do sistema de restrição de *C. acetobutylicum*. Para tanto, foram utilizados como estratégia de edição genômica o sistema CRISPR/Cas9 e o sistema de reparo por HDR (*Homology Direct Repair*) de *E. coli* com o objetivo de integrar o gene $\phi 3Ti$ expressando a metiltransferase no genoma de *E. coli*. As cepas mutantes foram confirmadas por PCR do fragmento integrado ao DNA genômico de *E. coli* e a funcionalidade de $\phi 3Ti$ *in vitro* foi confirmada pela digestão de vetores previamente introduzidos em células mutadas de *E. coli* com a enzima SatI, isoesquizômero da enzima de restrição Cac824I, presente no sistema de restrição de *C. acetobutylicum*. Em seguida a capacidade de metilação e proteção de material plasmidial desses clones foi avaliada *in vivo* pela transformação de *C. acetobutylicum* com vetores previamente submetidos aos clones de *E. coli*. Finalmente, uma técnica de PCR com princípio semelhante ao da TAIL-PCR e o subsequente sequenciamento das regiões amplificadas foram utilizados como estratégia para determinar o local aproximado de integração do gene $\phi 3Tim$. Os resultados obtidos mostraram que o gene $\phi 3Ti$ foi integrado e expresso com sucesso em *E. coli*, gerando uma plataforma universal de metilação com a finalidade de permitir a manipulação genética de *C. acetobutylicum*.

ABSTRACT

Bacteria belonging to the *Clostridium acetobutylicum* specie have been widely used for the production of solvents due to their capacity of performing an uncommon fermentative process known as ABE fermentation (Acetone/Butanol/Ethanol). Through this pathway this microorganism is capable of fermenting simple carbohydrates obtaining as final products the acetone, butanol and ethanol. Thus, *C. acetobutylicum* is considered promisor for the production of bioalcohols, specially, of butanol in an industrial scale. Different strains of *C. acetobutylicum* had already been studied and showed capacity of producing solvents; however, bacteria belonging to this genus present resistance to genetic manipulation because of their endogenous restriction system capable of breaking down DNA sequences there are not methylated, hampering the genetic improvement of these organisms. The present study had as objective the development of an *Escherichia coli* expressing the methyltransferase $\phi 3Ti$ -MTP, isolated from the $\phi 3T$ phage of *Bacillus subtilis*, which showed to be efficient for the protection of DNA sequences against the cleavage action of the *C. acetobutylicum* restriction system. Therefore, the CRISPR/Cas9 system as well as the HDR (Homology Direct Repair) system present in *E. coli* cells were used as genomic edition strategies for the integration of the gene $\phi 3Ti$ expressing the methyltransferase in *E. coli* DH10B genomic DNA. The mutant strains were confirmed by a PCR of the integrated fragment on the *E. coli* genomic DNA and the functionality *in vitro* of the $\phi 3Ti$ gene was confirmed by the digestion of vectors previously introduced in mutant *E. coli* cells with SatI, an isoschizomer of the Cac824I restriction enzyme, present on the *C. acetobutylicum* restriction system. Then, the capacity of methylating and protecting plasmidial material of the tested clones was evaluated *in vivo* through the transformation of *C. acetobutylicum* with vectors previously subjected to the *E. coli* clones. Finally, a PCR technique with a principal similar to the one of TAIL-PCR and the subsequent sequencing of the amplified regions were used as strategy to identify the approximate site of integration of the $\phi 3Ti$ gene. The results obtained showed that the $\phi 3Ti$ gene was integrated and successfully expressed in *E. coli*, generating an universal methylation system with the purpose of permitting the genetic manipulation of *C. acetobutylicum*.

1. INTRODUÇÃO

1.1. Clostrídios

Bactérias do gênero *Clostridium* são gram-positivas, anaeróbias estritas, formadoras de esporos e algumas de suas espécies são conhecidas por serem produtoras naturais de etanol e butanol. No entanto, apenas algumas espécies são capazes de produzir quantidades significativas desses solventes durante a fermentação sob condições específicas (Kharkwal *et al.*, 2009).

1.1.1. *Clostridium acetobutylicum*

Dentre as espécies de clostrídios mais estudadas, *Clostridium acetobutylicum* tem chamado a atenção de pesquisadores por ter o genoma completamente sequenciado, fornecendo uma riqueza de informações da sua maquinaria de síntese proteica, assim como a compreensão da via metabólica de produção de solventes durante a fermentação, permitindo que vários métodos de manipulação genética fossem determinados para essa espécie (Croux *et al.*, 2016).

1.1.2. Busca por novas fontes de energia renovável: Fermentação ABE

Ao longo da história, em períodos de escassez de petróleo, micro-organismos capazes de naturalmente produzir álcoois de interesse foram intensamente estudados e utilizados. É o caso de algumas espécies de clostrídios, como *C. acetobutylicum*, capaz de realizar um processo metabólico conhecido como fermentação ABE (*Acetone/ Butanol/Ethanol*), obtendo como produtos finais a acetona, o butanol e o etanol, com proporções típicas de 3 : 6 :1, respectivamente (Awang, Jones e Ingledew, 1988; Weber, Farwick e Benisch, 2010; Gheshlaghi *et al.*, 2009).

Com o aumento da preocupação voltada para os impactos da poluição ambiental, o aquecimento global e as reservas limitadas de combustíveis

fósseis, muita atenção tem sido dada para a busca de fontes de energia renováveis. Várias opções para a produção de biocombustíveis sustentáveis têm sido exploradas e entre elas destaca-se o uso de bioálcoois (Weber, Farwick e Benisch, 2010). Bioálcoois têm grande potencial para reduzir a dependência de combustíveis fósseis e diminuir a emissão de gases do efeito estufa por serem fontes de energia renováveis e produtos químicos que podem ser estocados (Kharkwal *et al.*, 2009).

O butanol é um álcool que tem ganhado destaque por apresentar vantagens em relação ao etanol tradicionalmente utilizado na indústria de combustíveis, principalmente em termos de densidade energética, por ser um composto estruturalmente mais parecido com a gasolina (Lee *et al.*, 2009). As propriedades químicas do butanol lhe conferem menor higroscopicidade (menor susceptibilidade a absorver a água do meio) e volatilidade, o que facilita o armazenamento e a distribuição desse produto. Além disso, o butanol pode ser utilizado como um composto químico intermediário em outras indústrias, como na produção de plásticos, detergentes e conservantes (Li *et al.*, 2009; Weber, Farwick e Benisch, 2010). Contudo, a maior parte da produção de butanol na atualidade é derivada de uma rota petroquímica baseada na síntese de óxido de propileno, tornando os custos da produção ligados ao mercado de propileno e, portanto sensíveis ao preço do petróleo bruto (Green, 2008).

Com isso, a ideia de aprofundar as aplicações biotecnológicas de *C. acetobutylicum* a fim de produzir álcoois de interesse em massa, em especial o biobutanol, tem se tornado alvo de muitos pesquisadores (Li *et al.*, 2009).

1.1.3. Melhoramentos genéticos em *C. acetobutylicum*

Apesar dos resultados promissores para a produção de biobutanol e bioetanol por *C. acetobutylicum*, existem ainda algumas desvantagens que dificultam a sua aplicação a níveis comerciais, fazendo com que o desenvolvimento de melhoramentos genéticos se torne necessário a fim de otimizar essa produção. Existem várias questões a serem mais bem estudadas: modificações que permitam um maior aproveitamento de substratos, incluindo

fontes alternativas derivadas de bagaços e outros resíduos lignocelulósicos; a produção de linhagens capazes de produzir maior concentração de solventes e com desempenho e produtividade contínua em sistemas de células imóveis; o desenvolvimento de cepas que tenham melhor tolerância aos produtos finais da fermentação e ao oxigênio; e o controle do processo de esporulação (Woods, 1995).

Para que a fermentação ABE seja viável em escala industrial sendo capaz de competir com outros processos químicos, o rendimento de solventes obtidos deve ser elevado. Contudo, estudos anteriores mostraram que esses micro-organismos apresentam grande sensibilidade a altas concentrações de butanol (maiores que 2%) limitando a utilização de fontes de carbono pelas células quando na presença deste álcool (Knoshaug e Zhang, 2009). Além disso, a identificação e o estudo de diferentes genes em *C. acetobutylicum* confirmaram que a proteína do estágio zero de esporulação A (Spo0A) controla diretamente a formação de solventes durante a fase “lag” de crescimento. Assim, é de particular interesse entender e regular genes envolvidos no processo de esporulação a fim de prolongar o metabolismo celular e a fase de solventogênese (Alsaker, Spitzer e Papoutsakis, 2004); (Yang *et al.*, 2018). Ainda, processos fermentativos realizados sob condições anaeróbicas resultam na baixa formação de biomassa devido aos baixos ganhos de ATP, por isso, para que haja uma produção de solventes em larga escala, o melhoramento genético de *C. acetobutylicum* a fim de aumentar a tolerância ao oxigênio poderia ampliar as aplicações biotecnológicas desses micro-organismos (Kharkwal *et al.*, 2009).

1.1.4. Sistema de restrição de *C. acetobutylicum*

Apesar da necessidade de redirecionar o metabolismo celular de *C. acetobutylicum*, existe uma enorme dificuldade em modificar esses micro-organismos devido à existência de um sistema de restrição e modificação (RM) altamente eficiente, impedindo que esses indivíduos aceitem DNA não-metilado eficientemente (Dong *et al.*, 2010). A ocorrência desse tipo de sistema pode ser consistente com um sistema imune capaz de atacar qualquer DNA exógeno

que não apresente características específicas, permitindo que a célula diferencie aquilo que é intrínseco do que não é (Berndt, Meier e Wackernagel, 2003). Esse mecanismo de reconhecimento acaba por conferir resistência à inserção de DNA plasmidial e integração no genoma, impactando a modificação genética desses micro-organismos (L. Mermelstein, N. Welker, G. Bennett, 1992).

Sistemas RM compreendem metiltransferases (MTases), que modificam o DNA endógeno pela adição de grupos metil em sítios específicos, e endonucleases de restrição (REases), que clivam o DNA exógeno não modificado (Nikolajewa *et al.*, 2005).

O sistema RM de *C. acetobutylicum* ATCC 824 é composto pela REase do tipo II Cac824I que tem como sítio de restrição o fragmento 5'-GCNGC-3' quando a citosina interna dessas sequências não está corretamente metilada. (L. Mermelstein, N. Welker, G. Bennett, 1992). Já foi observado que bactérias contendo sistemas RM apresentam uma frequência diminuída do sítio de restrição em seu genoma para aquela sequência de reconhecimento do seu próprio sistema, um fenômeno denominado “evitação” do local de restrição (Vasu e Nagaraja, 2013). Este é o caso do genoma de *C. acetobutylicum*, no qual o sítio de reconhecimento de Cac824I é pouco frequente em razão da grande quantidade de A + T (72%), mas estão frequentemente presentes em plasmídeos, fazendo com que este tipo de DNA exógeno seja clivado com uma alta frequência (L. Mermelstein, N. Welker, G. Bennett, 1992).

1.1.5. Suplantando o sistema de restrição de *C. acetobutylicum*

Uma alternativa para escapar deste sistema seria a deleção do gene que expressa a endonuclease de restrição Cac824I (CA_C1502), resultando em uma cepa incapaz de clivar o DNA exógeno não metilado. No entanto, já foi reportado por Vasu & Nagaraja, 2013, que genes envolvidos na expressão de elementos de sistemas RM são complexos e residem geralmente em regiões cromossômicas de difícil acesso. Além disso, foi observado que sistemas RM estão frequentemente ligados a elementos móveis ou que são adquiridos por

transferência horizontal de genes. Tudo isso abre questões acerca das vantagens adaptativas conferidas por esses sistemas para a célula em adição às funções de defesa celular, como um possível papel na estabilização cromossomal. (Vasu & Nagaraja, 2013)

Além disso, os sistemas RM são importantes para a manutenção da identidade da espécie. Em organismos mais complexos esse controle é feito por isolamento reprodutivo e/ou geográfico, mas no caso de alguns procariotos esta limitação é dada por meio de sistemas de isolamento gênico como o caso de sistemas RM, que impedem trocas frequentes de material genético e aquisição de novas propriedades biológicas. Já foi reportada, a ocorrência de infecções por fagos em produções industriais com fermentação por *C. acetobutylicum*, sugerindo que a supressão de sistemas de restrição pode resultar na produção de cepas mais susceptíveis (Vasu e Nagaraja, 2013).

Assim, outra forma de possibilitar a transformação destes micro-organismos seria a alteração no sítio de reconhecimento da Cac824I em vetores a serem introduzidos em *C. acetobutylicum*, protegendo-os da atuação deste sistema. Estudos com diferentes genes para metiltransferases (MTases) capazes de modificar o sítio de reconhecimento de Cac824I já foram realizados (M. Noyer-Weidner, S. Jentsch, J. Kupsch, M. Bergbauer, 1985). Foi demonstrado que a metiltransferase $\phi 3Ti$ clonada do fago $\phi 3T$ de *Bacillus subtilis* é capaz de proteger suficientemente bem fragmentos de DNA da clivagem por Cac824I, possibilitando que vetores de expressão previamente metilados por esta metiltransferase possam ser introduzidos de forma eficiente em *C. acetobutylicum* ATCC 824 por eletroporação (Papoutsakis e Mermelstein, 1993).

1.2. *Escherichia coli*

1.2.1. Expressão heteróloga em *E. coli*

O sistema procariótico de expressão de proteínas heterólogas é um dos mais bem conhecidos, tanto em escala laboratorial quanto industrial (Terpe, 2006). Particularmente, a bactéria *Escherichia coli* apresenta um dos sistemas

mais bem estudados e descritos tanto genética quanto fisiologicamente dentre os micro-organismos comumente utilizados para expressão heteróloga. Suas propriedades fisiológicas observadas em laboratório podem servir como base para o cultivo em alta escala e produção comercial (Olins e Lee, 1993). Vários sistemas de expressão estão disponíveis atualmente, no entanto, diferentes linhagens de *E. coli* continuam sendo extensivamente usadas: aproximadamente 80% das proteínas com estruturas bem conhecidas e submetidas ao banco de dados de proteínas (PDB) em 2003 foram expressas em *E. coli* (Kaur, Kumar e Kaur, 2018).

O sistema de expressão de *E. coli* confere vantagens relevantes como o rápido crescimento e a capacidade de alcançar alta densidade celular em meios de cultura simples e sob condições ambientais, tornando a seleção de mutantes conveniente (Baneyx, 1999). Ainda, os métodos de transformação com DNA exógeno para esses indivíduos são bem descritos (Rosano e Ceccarelli, 2014).

Por fim, outra característica primordial de células de *E. coli* é a sua fácil manipulação genética. Seu genoma pode ser modificado de forma rápida e precisa. Além disso, há uma grande disponibilidade de promotores induzíveis envolvidos na regulação da expressão gênica e de vetores de expressão compatíveis para estes organismos. Também, o número de cópias de plasmídeos pode ser facilmente alterado (Demain e Vaishnav, 2011).

Em geral, proteínas recombinantes superexpressas em *E. coli* se acumulam no citoplasma, tornando este o local preferencialmente utilizado para a produção de proteínas heterólogas com um bom rendimento (Terpe, 2006). A notável compreensão das regulações e interações celulares assim como a grande disponibilidade de ferramentas genéticas para esses micro-organismos podem garantir um grande rendimento de proteínas recombinantes biologicamente ativas em um curto espaço de tempo (Baneyx, 1999; Olins e Lee, 1993).

Não existe um sistema de expressão universal para a produção de proteínas recombinantes. O micro-organismo utilizado para a expressão heteróloga vai definir toda a tecnologia necessária para o projeto, devido à

variedade de ferramentas moleculares presentes nos diferentes tipos de sistema de expressão, assim, o objetivo de determinado estudo definirá qual o melhor sistema de expressão a ser utilizado (Rosano e Ceccarelli, 2014).

Os vetores de expressão são construídos de tal maneira que o gene presente possa ser transcrito e/ou traduzido. Para células de *E. coli*, esse sistema deve compreender todos os componentes para a expressão de genes clonados com elementos específicos, permitindo a sua aplicação para a expressão adequada de proteínas em *E. coli*, como promotores, marcas de seleção e sítios de múltipla clonagem (MCS). Existe uma grande disponibilidade de vetores comerciais para *E. coli*, assim, é importante que cada componente seja avaliado cuidadosamente de acordo com as necessidades de determinado projeto de pesquisa (Kaur, Kumar e Kaur, 2018; Rosano e Ceccarelli, 2014).

Os promotores são componentes essenciais de um sistema de expressão por serem responsáveis por controlar a iniciação da transcrição de genes associados (Kaur, Kumar e Kaur, 2018). A regulação da expressão gênica deve requerer sistemas induzíveis e controláveis, e, por essa razão, um promotor eficiente precisa ser forte, porém deve ser rigidamente reprimido. Promotores constitutivos, por exemplo, não são adequados para a produção de proteínas tóxicas ou deletérias para a célula quando continuamente expressas (Schumann e Ferreira, 2004). Ainda, é desejável que sistemas de indução de promotores sejam simples e acessíveis, como por exemplo, a indução através de nutrientes independentes da composição de dado meio de cultura que podem ser adicionados em determinado momento do cultivo (Terpe, 2006).

Muitos promotores são descritos para *E. coli*, alguns dos mais conhecidos foram obtidos de diferentes origens, sejam elas origens bacterianas, como o caso dos promotores lac, Trp e araBAD, ou de bacteriófagos, como os promotores R7 e SP6 por exemplo (Kaur, Kumar e Kaur, 2018). O promotor lac e seus derivados estão entre os mais comuns para os vetores de expressão em *E. coli* por ser um sistema eficiente e sujeito a fácil regulação por lactose ou IPTG (isopropil- β -D-1-tiol-galactopiranosídeo) (Gileadi, 2017).

Dentre os promotores induzíveis através da presença de nutrientes no meio, o promotor pBAD ou araBAD se destaca. Esse sistema faz uso da L-arabinose, um açúcar abundante na natureza, como indutor (Baneyx, 1999). Quando um gene é clonado na região *downstream* do promotor pBAD, a sua expressão é controlada pelo ativador araC, induzido por arabinose. Já na ausência de arabinose, ou presença de glicose, a mudança conformacional de araC faz com que este atue como um repressor, fazendo com que genes controlados pelo promotor pBAD não sejam expressos (Terpe, 2006). Ademais, estudos avançados com o emprego de biologia sintética têm trazido grandes contribuições na construção de componentes da maquinaria celular, como o uso de promotores sintéticos (Davy, Kildegaard e Andersen, 2017). Este é o caso do promotor constitutivo sintético PJ23119 comumente utilizado para a expressão de moléculas de sgRNA do sistema CRISPR/Cas, em razão do seu sítio de início de transcrição bem definido (Larson, M. H. *et al.*, 2013).

Para a seleção de células transformadas com um gene de interesse é essencial a utilização de um marcador de resistência. Para tanto, é necessário entender o genoma do micro-organismo e a linhagem com a qual se irá trabalhar. Em *E. coli* são usualmente utilizados genes que conferem resistência a determinado antibiótico, possibilitando que ocorra apenas o crescimento de clones mutantes (Rosano e Ceccarelli, 2014)

Cloranfenicol é um antibiótico comumente usado para a seleção de clones de *E. coli* transformados. Ele atua inibindo a síntese proteica bacteriana, e geralmente o gene que lhe confere resistência, Cm^R , atua através da inativação deste antibiótico pela enzima cloranfenicol acetiltransferase (CAT). Este marcador é eficiente e bem caracterizado para sinais de expressão para *E. coli*, o que o torna uma marca versátil e facilmente subclonada na construção de vetores (Shaw, 1975; Hadfield, Cashmore e Meacock, 1987).

Espectinomicina é um antibiótico menos frequente, porém também utilizado como marca de seleção para *E. coli*. Essa substância atua inibindo a biossíntese proteica de células bacterianas se ligando à subunidade 30S ribossomal (Kehrenberg *et al.*, 2005). Geralmente, a resistência a espectinomicina é conferida por enzimas capazes de inativar a droga por

adenilação. A adenililtransferase mais utilizada em vetores de expressão é a aminoglicosídeo adenililtransferase expressa pelo gene *aadA1*, capaz de inativar este antibiótico (Stern *et al.*, 2018).

Existem também marcas auxotróficas utilizadas para a seleção de *E. coli*, como é o caso da contrasseleção baseada na interrupção do gene *pyrF*. Este gene, responsável pela expressão da enzima orotidina-5-fosfato descarboxilase, é essencial na via de biossíntese de pirimidinas e também é responsável pela metabolização do ácido 5-fluoroótico (5FOA) ao seu produto final tóxico, o 5-fluorouracil (Suzuki, Murakami e Yoshida, 2012). A interrupção do gene através da integração de um fragmento de DNA na sua região codante pode ser usada como uma marca de contrasseleção, uma vez que as cepas resultantes da modificação se tornam auxotróficas para uracila e resistentes ao 5-FOA. Assim, quando cultivadas em um meio suplementado com uracila, apenas linhagens com o gene deletado serão capazes de sobreviver na presença do 5-FOA (Eberhart, Knutson e Barney, 2016).

Por fim, origens de replicação, formadas por um sítio específico onde é iniciada a replicação do DNA, são elementos essenciais para a propagação de um vetor, permitindo que ele faça outras cópias na célula hospedeira. Plasmídeos com origens de replicação iguais são incapazes de coexistir na mesma célula por serem considerados incompatíveis, o que significa que eles competem pela mesma maquinaria de replicação, gerando interferência um no outro e desestabilizando a frequência de replicação (Novick, 1987).

1.2.2. Linhagens de *E. coli*

E. coli é um micro-organismo modelo para o estudo de processos envolvidos no genoma procariótico. Por ser a espécie bacteriana mais bem estudada, existe hoje uma enorme disponibilidade de dados de sequenciamento genômico para múltiplas linhagens, muitos publicados e comercialmente disponíveis (Lukjancenko, Wassenaar e Ussery, 2010).

Devido à sua posição como modelo preferido para aplicações biotecnológicas em geral, a linhagem *E. coli* K-12 foi o organismo mais antigo a

ser sugerido como candidato ao sequenciamento completo (Blattner *et al.*, 1997). Trata-se de uma cepa não patogênica devido à ausência de estruturas superficiais que conferem fator de virulência e de genes que expressam toxinas responsáveis pela maioria das infecções. Seu genoma é composto por um cromossomo circular de 4,6 megabases, contendo 4401 genes que codificam para 116 RNAs e 4285 proteínas. O conhecimento relevante acerca dessa linhagem confere muitas vantagens para o uso experimental deste micro organismo (Cronan, 2006).

A linhagem de *E. coli* DH10B, um derivado de K-12, foi desenvolvida para a propagação de grandes vetores. Devido a sua alta eficiência de transformação de DNA e manutenção de grandes plasmídeos, essa cepa tem sido extensivamente usada para a clonagem. Seu genoma é largamente colinear com o da estirpe selvagem K-12, embora seja substancialmente mais complexo, permitindo que seja mais extensivamente utilizado para fins biotecnológicos. O genoma circular de DH10B possui 4.686.137 pares de bases, com 4305 genes codificadores de proteínas. Esta linhagem é incapaz de crescer em meio mínimo sintético sem a adição de leucina devido à deleção do operon *leuLABCD* que expressa a enzima responsável pela biossíntese deste aminoácido e pode ser utilizada para a seleção azul e branco, baseada na complementação *lacZ α* (Durfee *et al.*, 2008).

1.2.3. Técnicas moleculares para integração no genoma de *E. coli*

1.2.3.1. Reparo por HDR (*Homology Direct Repair*)

Proteínas pertencentes à classe de “proteínas de troca de fita de DNA” são moléculas capazes de eficientemente buscar e identificar sequências homólogas ao longo de uma grande sequência de DNA, possibilitando que a célula realize reparos de danos cromossomais por homologia, mantendo assim, a integridade genômica. Em *E. coli* encontra-se a proteína recombinatória de troca de fita de DNA denominada RecA. Para a atuação de RecA, inicialmente o DNA clivado é processado por um complexo de helicases e nucleases (RecBCD) que se liga nas duas regiões adjacente ao local de rompimento,

separando a dupla fita nessas regiões e degradando, para cada um dos lados, uma das fitas do dsDNA no sentido 5'→3,' produzindo assim dois braços de ssDNA nas regiões adjacente. Este complexo RecBCD catalisa a quebra da dupla fita até encontrar uma sequência de reconhecimento bacteriano chamado sítio Chi, que é espaçado ao longo do genoma de *E. coli* a aproximadamente cada 4-5 kb. As proteínas RecA atuam então nas regiões de ssDNA geradas. Nesta etapa, é montado um filamento formado pela união de várias proteínas RecA ao redor de cada região de fita simples de DNA no cromossomo quebrado, que por sua vez usa a sequência de ssDNA para procurar uma região homóloga ao longo do genoma, podendo utilizar, inclusive, a cromátide-irmã como modelo de reparo durante a replicação (Bell e Kowalczykowski, 2016).

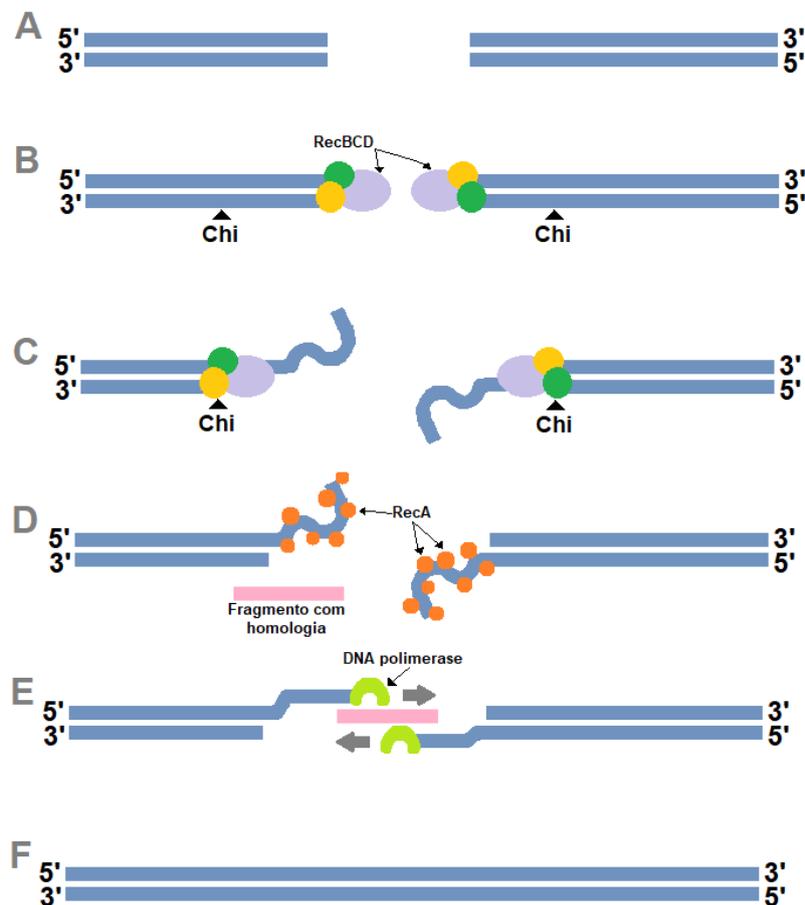


Figura 1. Sistema RecA. Representação esquemática de reparo da quebra de dupla fita através de recombinação homóloga mediada pelo sistema RecA. (A) região no DNA com quebra na dupla fita, (B) chegada do complexo RecBCD à região de reparo, (C) processamento do DNA pelo complexo RecBCD, (D) formação do filamento de proteínas RecA nas regiões de ssDNA e aproximação de um fragmento homólogo à região danificada, (E) polimerização da

fita complementar de DNA utilizando o fragmento homólogo como DNA molde, (F) dsDNA reparada.

O sistema celular de reparo por homologia induzido por estresse e ruptura da dupla fita de DNA (DSB, *Double Strand Break*) é bem descrito em *E. coli*. Este processo pode, no entanto, causar mutagênese e ser propenso a erros, o que levou ao desenvolvimento de estudos sobre mecanismos de mutação induzidos para essas células como forma de edição genômica (Shee *et al.*, 2012). Assim, estratégias de edição visando à inserção de fragmentos em sítios alvo no DNA e à introdução de mutações nocaute em *locus* de interesse tornaram-se objetos de estudo para esses micro-organismos (Cui e Bikard, 2016).

1.2.3.2. Utilização do sistema CRISPR/Cas9

A integração de um gene ao genoma pode ser obtida através da utilização da tecnologia CRISPR (“Clustered Regularly *Interspaced Short Palindromic Repeats*”) associada à endonuclease Cas9. A ferramenta CRISPR/Cas9, derivado do sistema imune adaptativo de *Streptococcus pyogenes*, consiste em um conjunto de duas estruturas: Cas9, capaz de clivar a dupla fita de DNA quando direcionada pela outra estrutura do sistema, o RNA guia (sgRNA, do inglês *single guide RNA*), formada pela união de duas moléculas de RNA, o crRNA (CRISPR RNA), que possui uma sequência complementar à da região alvo no DNA (sequência proto-espaçadora), e pelo tracrRNA (*trans-activating CRISPR RNA*) capaz de se ligar a Cas9. A endonuclease deste sistema, quando direcionada a um sítio alvo no DNA, procura por uma sequência PAM (Motivo adjacente proto-espaçador) nessa região, um seguimento de três nucleotídeos (NGG) localizado em uma região adjacente ao sítio de complementaridade com o RNA guia (Y. Li, Z. Lin, C. Huang, Y. Zhang, Z. Wang, Y. T. T. Chen, 2015; Marraffini, 2015).

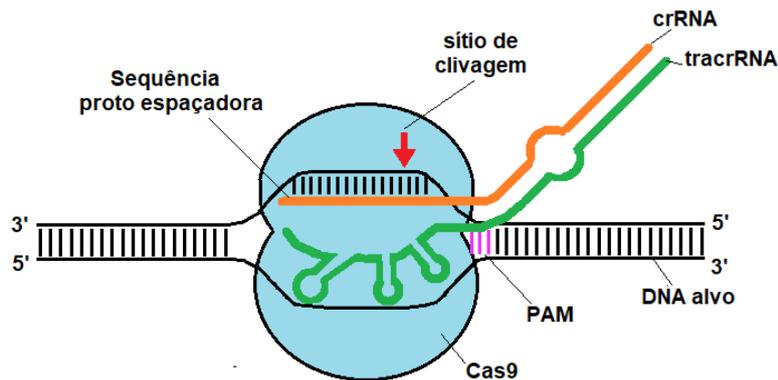


Figura 2. Mecanismo de atuação do sistema CRISPR/Cas9 de *S. pyogenes*. São mostradas as estruturas que compõem o sistema e a atuação deste na quebra da dupla fita de DNA

Recentemente, tem sido descrito o uso do sistema CRISPR/Cas9 como ferramenta de edição precisa no genoma de *E. coli*, levando em consideração a capacidade destes micro-organismos de realizar reparo celular por recombinação homóloga, assim como a fácil manipulação genética (Jiang *et al.*, 2013). Já foi descrita uma eficiência de 65% de modificação na edição por CRISPR/Cas9 em *E. coli* e até 100% quando o DNA doador foi inserido na célula sob a forma de dupla fita de DNA (Jiang *et al.*, 2013)(Pines *et al.*, 2015). Outras vantagens da utilização desse sistema são a possibilidade de se inserir fragmentos maiores, com até 2kb de tamanho, e a velocidade do processo (Y. Li, Z. Lin, C. Huang, Y. Zhang, Z. Wang, Y. T. T. Chen, 2015).

Esse complexo se tornou uma ferramenta útil para a edição genômica, fazendo com que a região de complementaridade do complexo de RNA possa ser devidamente programada para o direcionamento preciso da endonuclease Cas9 a uma região de interesse no DNA do hospedeiro (Y. Li, Z. Lin, C. Huang, Y. Zhang, Z. Wang, Y. T. T. Chen, 2015). Como a quebra da dupla fita é letal, células com modificações específicas geradas por recombinação homóloga com um plasmídeo, produto de PCR ou oligonucleotídeo sintético contendo regiões de homologia, podem ser facilmente obtidas, resultando em linhagens com mutações cromossômicas específicas (Pyne *et al.*, 2015).

Outra vantagem desse método de mutação induzida é a estabilidade das modificações pontuais em determinadas regiões do genoma. Modificações genéticas que visam a introdução e expressão de genes heterólogos se mostraram frequentemente mais estáveis e causaram menos sobrecarga

metabólica para a célula quando o gene de interesse é integrado ao genoma (Y. Li, Z. Lin, C. Huang, Y. Zhang, Z. Wang, Y. T. T. Chen, 2015).

Embora a especificidade de 22 pb da sequência protoespaçadora juntamente com a sequência PAM possam ser suficientes para garantir a mutação direcionada, alguns estudos levantaram preocupações de que este sistema pode causar danos colaterais no genoma. O sistema CRISPR/Cas9 tolera desequilíbrios especialmente na região 5' da sequência protoespaçadora, mas não na região entre 6-11 nt, mais próxima da sequência PAM. Recentemente, alguns estudos mostraram que, em células de humanos, este sistema pode induzir mutações fora do alvo, mesmo em locais que diferem por até 5 nt da região alvo (Cho *et al.*, 2014). Ainda que o direcionamento para *off-targets* seja menos provável em bactérias, por terem um genoma menor, já foi mostrado em *E. coli* que sequências protoespaçadoras bem como sequências PAM podem promover interferências, causando o direcionamento da endonuclease do sistema para um local incorreto (Xue *et al.*, 2015; Luo, Leenay e Beisel, 2016).

Além da existência de *off-targets*, alguns outros fatores podem influenciar na eficácia deste sistema, como a seleção do local do DNA alvo, o design do sgRNA, a incidência e eficiência do HDR e atividade da Cas9. Diferentes estudos mostraram que a seleção do local de destino e o design do sgRNA não são tão simples. Como já mencionado, a incompatibilidade de algumas bases pode ser tolerada, mas o desequilíbrio pode ser aumentado quando se aproxima da sequência PAM. O design do sgRNA tem sido objeto de vários trabalhos, resultando em muitos critérios para a sua construção (Lino *et al.*, 2018). O direcionamento eficiente da Cas9 permitindo alcançar diferentes *locus* gênicos de forma específica fornece um gargalo crítico para a edição genômica baseada na utilização do sistema CRISPR.

1.2.4. Construção de uma plataforma de metilação em *E. coli*

Estudos anteriores para a análise e caracterização da metiltransferases de sistemas de restrição, como ϕ 3Ti-MTP, já foram realizados com a

transformação de células de *E. coli* com vetores de alta expressão para a produção dessas MTases (Reiners e Keiners, 1986). Já foi demonstrado também que o gene $\phi 3Ti$ é bem expresso em *E. coli* com seu promotor natural e suficientemente bem expresso para metilar completamente o DNA plasmidial *in vivo*, para a posterior inserção em *C. acetobutylicum*, demonstrando que este pode ser um eficiente sistema para a proteção de vetores da enzima de restrição Cac824I (Papoutsakis e Mermelstein, 1993; Gomes, 2018).

Técnicas de engenharia genética que utilizam o sistema de expressão de *E. coli* para a obtenção de melhoramentos genéticos em *C. acetobutylicum* já foram bem estudadas e descritas no laboratório de genética e biotecnologia de leveduras. O trabalho de Gomes, 2018 demonstrou que a aplicação de linhagens de *E. coli* contendo o plasmídeo pACYC184 contendo o gene da metiltransferase $\phi 3Ti$ -MTP otimizado ($\phi 3Tim$) resultou em um eficiente sistema de metilação de vetores para *C. acetobutylicum*, tornando-os resistentes à ação do sistema de restrição deste micro-organismo. A funcionalidade e ampla aplicabilidade destes sistemas de expressão foram confirmadas, no entanto, a manutenção de células contendo o gene $\phi 3Ti$ ligado a um plasmídeo depende da utilização de marcas de seleção para aquele vetor. Também, estes sistemas restringem a sua aplicação para a metilação de vetores que não contenham a mesma marca de seleção ou origem de replicação do plasmídeo contendo o gene $\phi 3T$, evitando a ocorrência de incompatibilidade entre eles.

Com o intuito de desenvolver uma plataforma universal de metilação de vetores *in vivo* para a transformação de clostrídios, que seja estável, não dependendo do uso de antibióticos para a preservação do sistema, e que seja aplicável a diferentes vetores, com diversas origens de replicação e marcas de seleção, a expressão da metiltransferase $\phi 3Ti$ -MTP integrada ao DNA cromossomal de *E. coli* poderá ser uma alternativa eficiente, resultando na formação de um sistema estável e que possa ser amplamente utilizado, capaz de metilar múltiplos vetores sem causar incompatibilidade com suas origens de replicação ou interferência na seleção com determinado antibiótico. Vetores previamente submetidos a células de *E. coli* modificadas poderão ser metilados e, portanto, protegidos contra a atuação do sistema de restrição de *C. acetobutylicum*.

2. OBJETIVO

2.1. Objetivo geral

Desenvolver um sistema de metilação de vetores em *E. coli* para a transformação genética de *Clostridium acetobutylicum*

2.2. Objetivos específicos

- Integrar o gene $\phi 3Ti$ no genoma de *E. coli*.
- Analisar a funcionalidade de $\phi 3Ti$ através da digestão de plasmídeos isolados de *E. coli* modificadas.
- Remover a marca de seleção utilizada com o propósito de universalizar a aplicabilidade do sistema de metilação.
- Avaliar *in vivo* a proteção de plasmídeos metilados contra o sistema de restrição de *C. acetobutylicum*.

3. ESTRATÉGIA

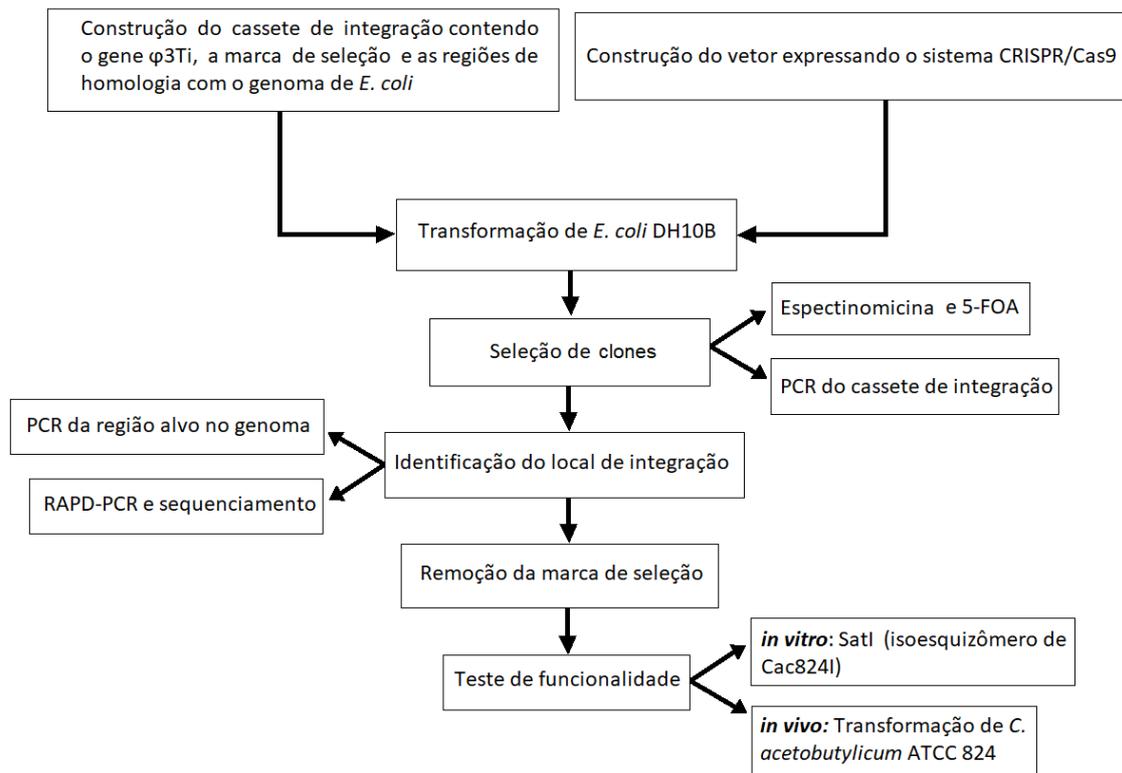


Figura 3. Estratégia para a construção e validação de um sistema de metilação em *E. coli*. Estratégia de trabalho para a criação de um sistema de metilação de vetores em *E. coli* para a transformação de *C. acetobutylicum*, com a utilização do gene $\phi 3Ti$ otimizado expressando a metiltransferase $\phi 3Ti$ -MTP.

4. MATERIAIS E MÉTODOS

4.1. Materiais

4.1.1. Bactérias

- *Escherichia coli* DH10B (Invitrogen, EUA): F^- *mcrA* $\Delta(mrr-hsdRMS-mcrBC)$ $\Phi80lacZ\Delta M15$ $\Delta lacX74$ *recA1* *endA1* *araD139* $\Delta(ara\ leu)$ 7697 *galU* *galK* *rpsL* *nupG* λ^-
- *Escherichia coli* DH10B $\Phi3Tim$ (derivada da linhagem DH10B, construída no nosso laboratório)
- *Clostridium acetobutylicum* DSM792 (DSMZ, Alemanha), doada pela Professora Dr. Cláudia Lareo da Universidad de la República Uruguay.

4.1.2. Meios de cultura

Meio Luria-Bertani – LB

- Extrato de levedura 0,5% (p/v)
- Peptona de caseína 1% (p/v)
- NaCl 1% (p/v)

O pH foi ajustado para 7,2 e o meio foi autoclavado. Para o meio sólido, foi adicionado ágar bacteriológico 1,4% (p/v)

Meio SOB

- Triptona 20 g/L
- Extrato de levedura 5 g/L
- NaCl 0,6 g/L
- KCl 0,5 g/L
- MgCl₂ 10 mM
- MgSO₄ 10 mM

O meio contendo a triptona, extrato de levedura e os sais foi autoclavado

e as soluções filtradas de $MgCl_2$ e $MgSO_4$ foram adicionadas posteriormente.

Meio SOC

- Glicose 2M 1mL
- SOB 99mL

Meio SB

- Peptona de caseína 3,0% (p/v)
- Extrato de levedura 2,0% (p/v)
- MOPS 1,0% (p/v)

O pH foi ajustado para 7,0 e o meio foi autoclavado

Meio mínimo M9

- Na_2HPO_4 12 g/L
- KH_2PO_4 6 g/L
- NaCl 1 g/L
- NH_4Cl 2 g/L

Meio YM (*Yeast extract minimal medium*)

- Meio M9 98% (v/v)
- Uracila 0,2% 1,12% (v/v)
- Extrato de levedura 2% 0,5% (v/v)
- $CaCl_2$ 50mM 0,2% (v/v)
- $MgSO_4$ 1M 0,1% (v/v)
- $ZnSO_4$ 100mM 0,08% (v/v)

Meio YM-ágar (seleção em 5-FOA)

- Ágar 2% 82,75% (v/v)
- YM 9,5% (v/v)
- Glicose 20% 5% (v/v)
- Tiamina 2% 1,6% (v/v)
- Leucina 1% 1% (v/v)

- FeSO_4 100mM 0,15% (v/v)

Foi adicionado 5-FOA à mistura final nas concentrações de 0,2mM, 0,16mM, 0,14mM e 0,12mM, separadamente para diferentes cultivos.

Meio YTG

- Triptona 16 g/L
- Extrato de levedura 10 g/L
- Glicose 5 g/L
- NaCl 4 g/L

Meio CGM (*Clostridium* Growth Medium)

- KH_2PO_4 5,5 mM
- K_2HPO_4 4 mM
- $\text{MgSO}_4 \cdot \text{H}_2\text{O}$ 3 mM
- $\text{MnSO}_4 \cdot \text{H}_2\text{O}$ 0,6 M
- $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$ 0,6 M
- NaCl 0,2 M
- Asparagina 0,15 M
- Extrato de levedura 0,5% (p/v)
- $(\text{NH}_4)_2\text{SO}_4$ 0,15 M
- Glicose 0,5 % (p/v)

4.1.3. Soluções

Soluções estoque para meios de cultura

Solução de glicose

- Glicose 2 M

Solução Mg 2M

- $\text{MgCl}_2 \cdot 6\text{H}_2\text{O}$ 1 M
- $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ 1 M

Extrato de levedura

- Extrato de levedura 2% (p/v)

Solução de arabinose

- L (-) arabinose 40% (p/v)

Antibióticos

Solução de Cloranfenicol (1000X)

- Cloranfenicol 50 mg/mL

Dissolvido em etanol 100% e esterilizado por filtração

Solução de Espectinomicina (1000X)

- Espectinomicina 50 mg/mL

Dissolvida em água milli Q e esterilizada por filtração

Solução de Ampicilina (1000X)

- Ampicilina 100 mg/mL

Dissolvida em água milli Q e esterilizada por filtração

Soluções para células bacterianas competentes

Tampão de transformação I

- RbCl 12 g/L
- $MnCl_2 \cdot 4H_2O$ 9,9 g/L
- Acetato de potássio 0,03 M
- $CaCl_2 \cdot 2H_2O$ 1,5 g/L
- Glicerol 150 g/L

Foi utilizada uma solução estoque de acetato de potássio a 1 M, pH 7,5

(ajustado com ácido acético) e esterilizada por filtração. O pH do tampão de transformação I foi ajustado para 5,8 com ácido acético 0,2 M e a solução foi esterilizada por filtração.

Tampão de transformação II

- MOPS 0,02 M
- RbCl 1,2 g/L
- CaCl₂ 11 g/L
- Glicerol 150 g/L

Foi utilizada uma solução estoque de MOPS a 1 M, com o pH ajustado para 6,8 com NaOH e esterilizada por filtração. O pH do tampão de transformação foi ajustado para 6,8 com NaOH e a solução foi esterilizada por filtração.

Solução de glicerol

- Glicerol 10% (v/v)

Soluções para extração de DNA plasmidial

Solução I (TE)

- Tris-HCl pH 8,0 10 mM
- EDTA pH 8,0 1 mM

Diluídos em água destilada.

Solução II

- NaOH 0,2 M
- SDS 1% (p/v)

Dissolvidos em água destilada. Solução preparada no momento de uso.

Solução III

- Acetato de potássio 5 M
- Acido acético glacial 2 M

O pH foi ajustado para 4,8

Soluções para eletroforese em gel de agarose

Brometo de Etídio (EtBr)

- EtBr 10 mg/mL

Tampão Tris-acetato EDTA (TAE) 50X

- Tris 2 M
- Ácido acético 1 M
- EDTA pH 8,0 0,05 M

Tampão de amostra

- TAE 20X 50% (v/v)
- Glicerol 30% (v/v)
- Azul de bromofenol 0,25% (p/v)

4.1.4. Enzimas

Enzimas de restrição

A tabela 1 apresenta as enzimas de restrição usadas neste trabalho, indicando seus sítios de clivagem, temperatura de incubação e o tampão utilizado para cada enzima.

Tabela 1. Enzimas de restrição utilizadas.

Enzima	Sítio de Clivagem	Tampão	Temperatura de incubação	Fornecedor
EcoRI	G ^A AATTC	NEB CutSmart	37°C	Jena Bioscience
XbaI	T ^A CTAGA	NEB CutSmart	37°C	Roche
SacI	C ^A TCGAG	NEB CutSmart	37°C	New England Biolabs

BamHI	G [^] GATCC	NEB	37°C	New England Biolabs
		Buffer 3.1		
EcoRV	GAT [^] ATC	NEB	37°C	New England Biolabs
		Buffer 3.1		
SatI (Fnu4HI)	GC [^] NGC	Tampão G	37°C	Thermo Scientific

Enzima de ligação

- Enzima T4 ligase (New England Biolabs): Utilizada na concentração de 1U/μL

Polimerases

- *Taq* DNA Polymerase, recombinant (Invitrogen)

Utilizada na concentração de 1 U/50μL de reação

- Platinum® *Taq* DNA Polymerase High Fidelity (Invitrogen)

Utilizada na concentração de 1 U/50μL de reação

- Phusion® High-Fidelity DNA Polymerase (New England Biolabs)

Utilizada na concentração de 1 U/50μL de reação

- DNA Polymerase I, Large (Klenow) Fragment (New England Biolabs)

Utilizada na concentração de 1U/μg de DNA

RNase A

- RNase A (DNase-free) 10 mg/mL

4.1.5. Marcadores de massa molecular para DNA

- 100bp DNA ladder, 100 - 1.500 bp (Sinapse inc.)
- 2-log DNA ladder, 100 - 10.000 bp (New England Biolabs)
- 1 kb Plus DNA Ladder 75 - 20.000 bp (Thermo Scientific)

4.1.6. Kits

- Wizard® Genomic DNA Purification Kit (Promega): Isolamento de DNA genômico de bactérias.
- NucleoSpin® Gel and PCR Clean-up (Macherey-Nagel): purificação de fragmentos de DNA a partir de gel de agarose e purificação de produtos de PCR.
- Wizard SV Gel and PCR Clean-Up System (Promega): purificação de fragmentos de DNA a partir de gel de agarose e purificação de produtos de PCR.
- Wizard® Plus Midipreps DNA Purification System (Promega): Isolamento de DNA plasmidial
- Wizard® Plus Minipreps DNA Purification System (Promega): Isolamento de DNA plasmidial
- In-Fusion® HD EcoDry™ Cloning Plus (Takara Bio): clonagem de fragmentos por recombinação homóloga *in vitro*.
- QIAGEN Plasmid Midi Kit (Qiagen): Purificação de DNA plasmidial em larga escala.

4.1.7. Oligonucleotídeos (*primers*)

A tabela 2 apresenta os *primers* utilizados neste trabalho com os respectivos sítios para enzimas de restrição quando presentes na sequência.

Tabela 2. *Primers* utilizados.

Primer	Sequência 5'→3'	Sítios de Restrição
sgRNAF	CAGT GAATTC TTGACAGCTAGCTCAGTCCTAGGTATA ATCAG CTCGAG TCATCAACCATTAACCGATTTCGGTT TTAGAGCTAGAAATAGCAAGTTAAAATAAGGC	EcoRI XhoI
sgRNAR	CTAGT CTAG AAAAAAGCACCGACTCGGTGCCACTTT TTCAAGTTGATAACGGACTAGCCTTATTTAACTTGCT ATTCT	XbaI
infusionF	ATGATTTCTGGAATT GAATTC TTGACAGCTAGCTCAGT CC	EcoRI
infusionR	AAGCGGCCGCGAATTT CTAG AAAAAAGCACCGACT CG	XbaI
sphi_F2	TTACCGCATAACCCCTCTACTGGCATTGTCCGTCTTG TCGGGCCTGGTCAATGCGCTCACGCAACTGGTCCAG	//
sphi_R3	GGATTTTTTTTTAAAAATTACAGGCCATCAAGACCGCGA ATCGGTTTAATGAGAGCGTTCACCGACAAACAAC	//
specF	GGCGGTTTTTCATGGCTTGTT	//
spec200R	GATACTTCGGCGATCACCGC	//
phiF	TCTGGTTTCCCGTGTCCAA	//
phiR	GTCGTGGCCCTTCAGATCAA	//
pyr3.2	AGTAAGCCTCGTTATCGTTGC	//
pyr4	CGCTTATCGCCCTGGATTTCAA	//
specFR	AACAAGCCATGAAAACCGCC	//

As sequências dos sítios de restrição estão em negrito e sublinhadas enquanto as regiões de anelamento dos primers estão em itálico.

4.1.8 Vetores

A tabela 3 exibe todos os vetores usados neste trabalho e apresenta uma breve descrição de cada um.

Tabela 3. Vetores utilizados

Vetor	Descrição
AraC_pBAD_Cas9	Vetor cedido pelo professor Spartaco Astolfi Filho da Universidade Federal do Amazonas, contendo o gene para a expressão da proteína Cas9 do sistema CRISPR/Cas9 e o gene de resistência a cloranfenicol como marca de seleção.
AraC_pBAD_CRISPR	Vetor para a expressão da endonuclease de restrição Cas9 juntamente com o complexo sgRNA (RNA guia) do sistema CRISPR/Cas9 (construído neste trabalho, a partir do vetor AraC_pBAD_Cas9).
pBluescript II SK (+)	Vetor de clonagem comercial para expressão em bactéria utilizado para se ligar ao cassete de integração <i>spec-ϕ3Tim</i> .
pBlue-ϕ3Tim	Vetor pBluescript II SK(+) ligado ao cassete de integração. Este sistema de ligação foi utilizado para a transformação de <i>E. coli</i> por eletroporação. (construído neste trabalho).
705-Cre	Vetor de expressão para a Cre-recombinase, utilizado para retirar o gene de resistência a espectinomicina de cepas de <i>E. coli</i> transformadas. Contém um promotor termossensível possibilitando que a recombinação seja controlada pela temperatura de incubação. Fonte: Gene Bridges
pACYC_184	Vetor de clonagem comercial para <i>E. coli</i> com baixo número de cópias, utilizado para testar a eficiência da atividade da metiltransferase Φ 3Ti-MTP integrada ao genoma.

4.1.9. Ferramentas de bioinformática

Neste trabalho foram utilizadas algumas ferramentas de bioinformática para a busca e o alinhamento de sequências, bem como para a análise de possíveis regiões a serem utilizadas como sítio alvo do sgRNA. Estas estão listadas na tabela 4.

Tabela 4. Ferramentas de bioinformática utilizadas

Ferramenta	Endereço eletrônico	Aplicação
GenBank (NCBI)	http://www.ncbi.nlm.nih.gov	Busca de sequências
BLAST (NCBI)	http://blast.ncbi.nlm.nih.gov/Blast.cgi1	Alinhamento de sequências
Crispr RGEN Tools	http://www.rgenome.net/	Montagem do fragmento que codifica para a sequência proto espaçadora do crRNA do sistema CRISPR/Cas9
In-Fusion Cloning Primer Design Tool	http://www.takabio.com/learning-centers/cloning/in-fusion-cloning-tools	Montagem dos primers utilizados com o kit In-Fusion HD EcoDry Cloning Plus (Takara)

4.2. Métodos

4.2.1. Cultivo de micro-organismos

Culturas de *E. coli* foram cultivadas a 37°C com exceção do cultivo de cepas transformadas com o plasmídeo 705-Cre, que foram cultivadas a 30°C para que o plasmídeo se replicasse antes da recombinação. Para o crescimento em meio líquido, as culturas foram mantidas a uma velocidade de agitação de 200 rpm.

A cepa do *C. acetobutylicum* DSM792 foi doada pela Prof.^a Claudia Lareo da Universidad de la República Uruguay e mantida em meio líquido em frasco de cultivo de 50mL adicionado de nitrogênio gasoso (N₂) para

manutenção da condição anaeróbia, fechado com tampa butírica e selados com lacre de alumínio. Os inóculos foram realizados a partir da injeção de 1% do estoque mantido a 4°C em meio líquido apropriado e incubados em shaker a 37°C por 48 h.

4.2.2. Amplificação por PCR

Os *primers* utilizados neste trabalho estão listados na tabela 2. As amplificações foram realizadas utilizando-se as enzimas *Taq* DNA Polymerase, recombinant (Invitrogen), Platinum® *Taq* DNA Polymerase High Fidelity (Invitrogen) e Phusion® High-Fidelity DNA Polymerase (New England Biolabs) de acordo com as instruções dos fabricantes.

Durante os ciclos de PCR foram mantidas as temperaturas de desnaturação e extensão especificadas pelos fabricantes das enzimas, bem como os tempos sugeridos para cada etapa do processo. Foram realizados 30 ciclos para todas as reações de PCR. As temperaturas de anelamento foram mantidas entre 55°C e 57°C com exceção das etapas de TAIL-PCR que tiveram temperaturas de anelamento de 40°C para a primeira etapa e 63°C para etapas subsequentes.

O fragmento que codifica para o RNA guia do sistema CRISPR/Cas9 foi gerado também por polimerização de DNA, com a utilização da enzima DNA Polymerase I, Large (Klenow) Fragment (New England Biolabs), uma polimerase desprovida de atividade exonucleásica 5'→3', e de dois oligonucleotídeos específicos com regiões que se complementam.

4.2.3 Preparação de células bacterianas competentes

Células de *E. coli* competentes para choque térmico

Células de *E. coli* DH10B foram semeadas em meio LB ágar a partir de um estoque armazenado a -80°C e cultivadas a 37°C por 18 horas. Uma única

colônia isolada foi inoculada em 10 mL de meio SOB em um Erlenmeyer de 125 mL a 37°C, sob agitação de 200rpm por 16h. Dois mililitros deste pré-inóculo foram inoculados em 100 mL de meio SOB em frascos Erlenmeyer de 500 mL e então incubados a 37°C sob agitação de 200 rpm até que a cultura atingisse OD₆₀₀ de 0,3 (~ 1 hora). A cultura foi então dividida em dois tubos tipo Falcons de 50 mL e rapidamente resfriada em banho de água/gelo por aproximadamente 15 minutos. A cultura foi coletada por centrifugação a 3.000 g por 5 minutos a 4°C e o sobrenadante foi descartado. O pellet foi ressuscitado em 16 mL de tampão de transformação I (para cada 50 mL da cultura original) e novamente incubado em gelo por 15 minutos. Em seguida, as células foram coletadas por centrifugação a 3.000 g por 5 minutos a 4°C e ressuscitadas em 2 mL de tampão de transformação II para cada 50 mL da cultura original. Neste ponto, as células consideradas termo competentes foram armazenadas em tubos Eppendorf de 1,5 mL em alíquotas de 100 µL e congeladas imediatamente em gelo seco. Por fim, os tubos foram guardados a -80°C.

Células de *E. coli* competentes para eletroporação

Células de *E. coli* DH10B foram semeadas em meio LB ágar a partir de um estoque armazenado a -80°C e cultivadas a 37°C por 18 horas. Uma única colônia isolada foi inoculada em 10 mL de meio SB em um frasco Erlenmeyer de 125 mL a 37°C, sob agitação de 200rpm durante a noite. Quinhentos microlitros dessa cultura foram inoculados em 500 mL de meio SB contendo 2,5 mL de solução estoque de glicose 2M e 2,5 mL de solução estoque de Mg 2M em um frasco de 1 litro. A cultura foi incubada sob agitação de 200rpm a 37°C até atingir uma OD₆₀₀ de 0,7 a 0,9 (meio da fase exponencial) – aproximadamente 3 horas de incubação, e então, resfriada em gelo picado seguindo-se centrifugação a 3.000 g por 20 minutos a 4°C. O sobrenadante foi descartado e o pellet foi ressuscitado cuidadosamente em 25 mL de glicerol 10% (v/v) gelado com o auxílio de pipetas pré-resfriadas. Em seguida foram adicionados mais 75 mL de glicerol 10% (v/v) seguindo-se centrifugação a 3.000 g por 20 minutos a 4°C. O sobrenadante foi descartado e este

procedimento de lavagem com glicerol foi repetido novamente. As células foram ressuspensas uma terceira vez em 25 mL de glicerol 10% (v/v) gelado com o auxílio de pipetas pré-resfriadas, transferidas para tubos Falcons de 50 mL e coletadas por centrifugação a 3.000 g por 20 minutos a 4°C. Finalmente, as células foram ressuspensas no volume residual de glicerol no tubo para uma OD₆₀₀ de aproximadamente 200-250 (em torno de 1 – 2 mL de solução). Neste ponto, as células consideradas eletro competentes foram armazenadas em tubos Eppendorf de 1,5 mL em alíquotas de 100 µL, congeladas imediatamente em gelo seco e por fim mantidas a -80°C.

Células de *C. acetobutylicum* competentes para eletroporação

Dois mL de uma cultura estoque mantida a 4°C foram adicionados ao meio CGM. O sistema foi incubado a 37°C até atingir uma OD₆₀₀ de 0,6. A preparação do inóculo foi feita com a adição de 10% do pré-inóculo a 50 mL de meio CGM e incubada a 37 °C até atingir a DO₆₀₀ 1,2. O cultivo foi transferido para tubos Falcon de 50 mL e centrifugados a 6000 g por 15 min. As células coletadas foram lavadas duas vezes e ressuspensas em 1,8 mL de uma solução de NaH₂PO₄ 5 mM e glicose 270 mM. Neste ponto, as células consideradas eletro competentes foram, em seguida, utilizadas para eletroporação.

4.2.4 Manipulação de DNA

Digestão de DNA com enzimas de restrição

As digestões de DNA foram realizadas de acordo com as informações listadas para cada uma das enzimas na tabela 1, como tampões e temperaturas utilizadas, com base no grau de eficiência de cada reação de digestão. A quantidade de enzima para cada digestão foi ajustada à quantidade de DNA a ser digerido considerando que o volume da enzima não deveria

ultrapassar 10% do volume total da reação. Todas as reações foram incubadas por, pelo menos, 2 horas.

Análise e purificação de fragmentos de DNA

Após as etapas de digestão ou amplificação de DNA a análise dos fragmentos gerados foi feita utilizando-se a técnica de eletroforese em gel de agarose adaptada de Sambrook et al. (1989). Todos os géis foram preparados na concentração de 0,8% (p/v) de agarose para a análise de fragmentos > 1kb e 2% (p/v) de agarose para a análise de fragmentos pequenos (< 1kb). Para o preparo dos géis, a agarose foi diluída em tampão de corrida TAE 1X e adicionado brometo de etídeo na concentração final de 0,5 µg/mL. As amostras foram aplicadas no gel e submetidas à corrente elétrica (70 V) para possibilitar a migração e separação dos fragmentos de DNA. A visualização das bandas de DNA foi feita com a exposição do gel à luz ultravioleta.

As purificações dos fragmentos de DNA extraídos de gel de agarose e dos produtos de PCR foram realizadas com os kits NucleoSpin® Gel and PCR Clean-up (Macherey-Nagel) ou Wizard SV Gel and PCR Clean-Up System (Promega) seguindo o manual dos respectivos fabricantes.

4.2.5. Preparação de DNA plasmidial

Preparação em pequena escala (miniprep) (adaptado de Sambrook et al., 1989)

Uma colônia isolada de *E. coli* foi inoculada em 3 mL de meio LB com o antibiótico adequado e o inóculo foi cultivado a 37°C por 18 h. As células foram concentradas por centrifugação a 10.000 g por 2 minutos e o sobrenadante foi descartado. O pellet foi ressuspendido em 200µL de TE (solução I) e, em seguida, foram adicionados 360µL de solução II. O material ressuspendido foi mantido à temperatura ambiente por 5 minutos e, então, foram adicionados 300

μL de solução III gelada. A mistura foi então incubada em gelo por 5 minutos e, posteriormente, coletada por centrifugação a 10.000 g por 5 minutos. O sobrenadante foi transferido para um tubo Eppendorf contendo 750 μL de isopropanol, e o sistema foi homogeneizado por inversão e submetido a centrifugação a 10.000 g por 5 minutos. O sobrenadante foi descartado e o pellet restante foi ressuspensionado em 200 μL de TE. Seguidamente, foram adicionados 110 μL de acetato de amônio 7,5 M, o sistema foi misturado vigorosamente e, então, submetido a centrifugação a 10.000 g por 10 minutos. O sobrenadante foi transferido para outro tubo contendo 750 μL de etanol 100%, homogeneizado por inversão, e submetido a centrifugação a 10.000 g por 5 minutos. O sobrenadante resultante desta segunda centrifugação foi descartado, o pellet lavado com 1 mL de etanol 70% (sem ressuspensão) e esta mistura foi submetida a centrifugação a 10.000 g por 2 minutos. Finalmente, o sobrenadante foi descartado, e o excesso de etanol foi evaporado a temperatura ambiente. O precipitado foi ressuspensionado em água milli Q contendo RNase A (0,1 mg/mL) e armazenado a -20°C para uso posterior.

Preparação em pequena escala (miniprep) com o kit Wizard® Plus Minipreps DNA Purification System (Promega)

O DNA plasmidial foi extraído de acordo com as orientações do fabricante.

Preparação em média escala (midiprep) (adaptado de Sambrook *et al.*, 1989)

O kit QIAGEN Plasmid Midi Kit (Qiagen) foi utilizado até o momento da adição da solução III e centrifugação do material (etapa 5). Até este ponto foram seguidas as orientações do fabricante para cada etapa do processo já realizada. Em seguida, o sobrenadante formado foi separado do pellet e misturado a um volume igual de fenol: clorofórmio (1:1). A mistura foi

homogeneizada vigorosamente e submetida a centrifugação a 8.000 *g* por 1 hora. A fase superior da solução bifásica formada foi transferida para um novo tubo e misturada a um volume igual de clorofil (álcool isoamílico: clorofórmio, 1:24). A amostra foi submetida a centrifugação a 8000 *g* por 45 minutos e novamente a parte superior da solução bifásica formada foi transferida para um novo tubo e misturada a um volume de isopropanol igual a 70% do volume da amostra. Essa mistura foi submetida a centrifugação a 15.000 *g* por 1 hora, o sobrenadante foi descartado e o pellet foi lavado com 2 mL de etanol 70%. Por fim, a mistura foi submetida a centrifugação a 15.000 *g* por 15 minutos, o sobrenadante foi mais uma vez descartado, o etanol residual foi removido por evaporação e o DNA foi ressuspenso em água Milli Q.

Preparação em média escala (midiprep) com o kit Wizard® Plus Midipreps DNA Purification System (Promega)

O DNA plasmidial foi extraído de acordo com as orientações do fabricante.

4.2.6. Extração de DNA genômico de bactérias

A extração do DNA genômico foi feita com o kit Wizard® Genomic DNA Purification Kit (Promega) de acordo com as orientações do fabricante.

4.2.7. Transformação de bactérias

Transformação de *E. coli* por choque térmico

Uma alíquota contendo 100 μ L de células previamente preparadas foi utilizada para cada inserção de DNA plasmidial. As células termo competentes foram retiradas do freezer -80°C e incubadas no gelo por 5 minutos para que descongelassem. Foram adicionados aproximadamente 200-300 ng de solução

plasmidial a essa mistura e, então, incubou-se no gelo por 30 minutos. Em seguida, as células foram submetidas ao choque térmico a 42°C por 90 segundos e foram adicionados 900 µL de meio LB ao sistema seguindo-se incubação a 37°C por 1 hora. Por fim, as células foram semeadas em placas contendo meio LB e uma adição de antibiótico adequado. As placas foram incubadas por 18 h a 37°C.

Transformação de *E. coli* por eletroporação

Uma alíquota contendo 100µL de células previamente preparadas foi utilizada para cada incorporação de DNA exógeno. As células eletrocompetentes foram retiradas do freezer -80°C e incubadas no gelo por 5 minutos para que descongelassem. As amostras de DNA exógeno foram adicionadas aos tubos Eppendorfs contendo as células eletrocompetentes na concentração de aproximadamente 200-300 ng para DNA plasmidial e 800-900 ng para fragmentos de DNA linear. Em seguida, os sistemas de transformação foram transferidos para cubetas de 0,2cm pré-resfriadas, e a eletroporação foi realizada com os seguintes parâmetros: 2,5kV; 25 µF e 200 Ω. O tempo de pulso foi de 4 ou 5 segundos. Imediatamente após este procedimento foi adicionado 1 mL de meio SOC à cada cubeta e o sistema foi então transferido para um tubo Falcon estéril de 15 mL contendo 2 mL de meio SOC. As células foram incubadas a 37°C sob agitação de 200 rpm por uma hora e meia para recuperação. Finalmente, as células foram semeadas em placas contendo meio LB ágar e o antibiótico adequado para a seleção de transformantes. As placas foram incubadas a 37°C por 18 h, com exceção das placas contendo células transformadas com o plasmídeo 705-Cre que foram incubadas a 30°C.

Transformação de *E. coli* e concomitante ativação da proteína Cas9 quando utilizado o cassete de integração na forma de fragmento linear (método não convencional de preparação de células competentes para eletroporação) - adaptado de Wang *et al.*, (2006).

Antes da transformação com o fragmento de DNA linear, uma transformação por eletroporação convencional foi realizada para transformar *E. coli* DH10B com o vetor AraC_pBAD_CRISPR. Para a transformação usada para a inserção do fragmento, uma colônia isolada da cepa previamente transformada com AraC_pBAD_CRISPR foi inoculada em 5mL de meio LB com cloranfenicol (para assegurar que a célula irá manter o plasmídeo, que confere resistência ao antibiótico) e incubada a 37°C, sob agitação de 200 rpm durante a noite. Um volume de 750 µL dessa cultura foi inoculado em 50 mL de meio LB contendo cloranfenicol em um frasco Erlenmeyer de 250 mL e seguindo-se incubação a 37°C, sob agitação de 200 rpm, até que atingisse uma OD₆₀₀ de aproximadamente 0,3 (~ 4 h). Em seguida, adicionou-se arabinose para uma concentração final de 1% e a cultura foi incubada novamente até que chegasse a uma OD₆₀₀ de 0,45. Ao final desse processo, a cultura foi resfriada em gelo e depois transferida para um tubo Falcon estéril de 50 mL. A partir desse momento, as células foram sempre mantidas a 0°C. A cultura foi então coletada por centrifugação a 6.000 g por 20 minutos. Logo após, o pellet foi ressuscitado em 20 mL de glicerol 10% (v/v) gelado sendo em seguida submetida a centrifugação a 6.000 g por 15 minutos. O sobrenadante foi descartado e esse processo de lavagem com glicerol foi repetido. Por fim, o pellet foi ressuscitado no glicerol residual do tubo e as células concentradas foram adicionadas do produto de PCR contendo o cassete de integração. A mistura foi transferida para uma cubeta resfriada de 0,2 cm e o choque foi aplicado nas mesmas condições já especificadas na transformação convencional por eletroporação. O processo de recuperação celular foi feito em 3 mL de meio SOB com adição de arabinose para uma concentração final de 1% e sob as mesmas condições descritas para a recuperação de células eletroporadas pelo método convencional. Ao final deste processo, as células foram semeadas em placas contendo meio LB ágar com adição de espectinomicina ou meio YM-ágar.

Transformação de *C. acetobutylicum* por eletroporação

Células de *C. acetobutylicum* eletro competentes foram adicionadas a cubetas de 2 mm (aproximadamente 900 μ L por cubeta) e incubadas no gelo por 5 min. Após os 2 primeiros minutos, 600 ng de plasmídeo foram adicionados. As condições de eletroporação foram: 2kV, 25 μ F e 600 Ω . O tempo de pulso foi de 6 segundos. Imediatamente após este procedimento as amostras eletroporadas foram injetadas em 15 mL de meio YTG incubado por 24 h a 37 °C para recuperação. Dois mL dessa cultura foram transferidos para 15mL de meio YTG contendo tianfenicol (15 μ g/mL) e o sistema foi incubado a 37 °C, sob agitação de 200rpm por 48 h.

4.2.8. Ativação da endonuclease Cas9 quando utilizado o cassete de integração ligado ao vetor pBluescript SK (pBlue- Φ 3Tim)

Para a ativação do sistema CRISPR/Cas9 em *E. coli* transformadas conjuntamente com os vetores AraC_pBAD_CRISPR e pBlue- Φ 3Tim, o crescimento em placas após a transformação baseou-se em um método adaptado de Zhao *et al.*, (2016). Neste procedimento, uma colônia isolada crescida em placa contendo os antibióticos cloranfenicol e ampicilina (marcas de seleção para cada um dos plasmídeos introduzidos), foi inoculada em 5 mL de meio LB acrescido dos mesmos antibióticos. A cultura foi incubada a 37°C por 2 horas sob agitação de 200 rpm e depois adicionou-se 2 g/L de arabinose. O cultivo foi mantido sob as mesmas condições por mais 6 horas e, por fim, as células foram semeadas em meio LB acrescido de cloranfenicol, ampicilina e arabinose 2 g/L. As placas foram incubadas a 37°C durante a noite e, ao fim do cultivo, as colônias crescidas transferidas para uma placa em grade contendo LB acrescido de espectinomicina, como uma segunda etapa de seleção das células transformadas contendo o fragmento *spec- Φ 3Tim*.

4.2.9. PCR de colônia

A PCR de colônia foi realizada com um conjunto de até 10 colônias por reação. A preparação do DNA molde a ser utilizado foi feita da seguinte maneira: cada uma das 10 colônias a ser analisadas foi coletada minimamente

com a ponta de um palito estéril e inoculada em 50 μ L de água Milli-Q. O sistema foi fervido a 100°C por 5 minutos e coletado por centrifugação a 6.000 g por 2 minutos. Foram utilizados 2 μ L do sobrenadante deste sistema como DNA molde para a etapa de PCR, realizada com a enzima *Taq* DNA Polymerase, recombinant (Invitrogen) de acordo com as instruções do fabricante.

4.2.10. TAIL-PCR e sequenciamento dos fragmentos obtidos

Neste trabalho, uma técnica com princípio baseado em TAIL-PCR foi utilizada como estratégia para identificar o local aproximado de integração do cassete contendo o fragmento *spec- Φ 3Tim* no genoma de *E. coli*, uma vez que não foi possível até aqui determinar onde este fragmento estava integrado. Para tal, foi utilizado inicialmente um único primer de sequência complementar (reverso) a uma sequência presente no início do cassete de integração.

Para a primeira etapa, a PCR foi conduzida com uma baixa temperatura de anelamento permitindo que o primer se ligasse com baixa especificidade ao DNA molde. Foi utilizada a enzima *Taq* DNA Polymerase, recombinant (Invitrogen) de acordo com as instruções do fabricante. As bandas de DNA geradas pela PCR foram analisadas em gel de agarose 0,8% e extraídas do gel com o auxílio de uma micropipeta. Esse material extraído foi utilizado como DNA molde separadamente para a próxima etapa de PCR, permitindo analisar isoladamente cada uma das bandas extraídas.

Na segunda etapa, mais uma vez a técnica TAIL-PCR foi empregada, utilizando-se o mesmo primer usado anteriormente, conduzida com uma alta temperatura de anelamento para assegurar que o primer se anelasse com maior especificidade aos DNA molde.

Após uma terceira etapa do processo, foi realizada uma PCR comum. Para tal, as amostras que amplificaram uma banda única na etapa anterior foram empregadas como DNA molde. Neste procedimento, foi usado um par de primers internos ao cassete de integração, separados por uma região de aproximadamente 200 pb.

Finalmente, depois de revelada qual banda continha o cassete de integração, a segunda etapa da técnica de TAIL-PCR foi realizada novamente, com uma maior quantidade de material, para purificação e sequenciamento. A purificação foi realizada com o kit Wizard SV Gel and PCR Clean-Up System (Promega) e o sequenciamento do fragmento foi realizado pela empresa Myleus Facility.

4.2.11. Otimização do gene $\phi 3Ti$

O gene $\phi 3Ti$, que codifica para a metiltransferase $\phi 3Ti$ -MTP produzida pelo fago $\Phi 3T$ de *B. subtilis* foi otimizado por Gomes, (2018) pela substituição do códon de iniciação atípico, TTG, pelo códon de iniciação canônico, ATG, afim de ser devidamente expresso em *E. coli*. Já foi demonstrado que a utilização de um sistema com o gene $\phi 3Ti$ formado pelo códon de iniciação atípico TTG não foi corretamente expresso em *E. coli* DH5 α , resultando na digestão parcial de vetores quando submetidos à ação da enzima Fnu4HI, um isosquizômero de Cac824I, após a metilação em cepas transformadas com plasmídeo contendo este gene (GOMES, 2018).

Assim, o gene otimizado foi quimicamente sintetizado pela empresa IDT DNA (Iowa, USA) na forma de um fragmento de 1,535 kb. O fragmento de DNA sintético resultante foi formado por uma sequência correspondente à região codante do gene $\phi 3Ti$ com otimização do códon de iniciação e sob o controle do promotor do gene *lacI* e da sequência terminadora de transcrição T1 do gene *rrnB*, ambos de *E. coli*. Este gene sintético foi denominado $\phi 3Tim$.

5. RESULTADOS E DISCUSSÃO

5.1. Construção de fragmentos de DNA

Construção do cassete de integração

O gene sintético $\phi 3Tim$ foi ligado ao gene que confere resistência a espectinomicina, denominado *spec*. Ao final deste processo, foi gerado um fragmento de 2,657 kb contendo, nesta ordem, os genes *spec* e $\phi 3Tim$. Esse fragmento foi denominado *spec- $\phi 3Tim$* .

A utilização do gene que confere resistência a espectinomicina foi prevista para a seleção de clones transformados com o fragmento. É esperado que, para a construção de um sistema universal e eficiente de metilação a marca de seleção seja removida. Assim, o gene *spec* foi construído entre sítios LoxP para a remoção do gene via recombinação Cre-Lox. No entanto, como a espectinomicina é um antibiótico pouco utilizado, o uso do cassete completo *spec- $\phi 3Tim$* também pode ser considerado.

A construção do cassete de integração foi baseada no processo de reparo celular por HDR, que utiliza sequências de homologia para mediar eventos de reparação da quebra da dupla fita por recombinação homóloga. Assim, o cassete de integração foi constituído de dois braços de homologia, com 50 pb cada, com complementariedade específica à região adjacente ao sítio de clivagem da proteína Cas9 em *E. coli*, no gene *pyrF*, de modo a promover o reparo da quebra da dupla fita por recombinação com essas regiões. Todo este processo de edição genômica tem o intuito de empregar ambas as ferramentas moleculares descritas anteriormente, em duas etapas: o sistema CRISPR/Cas9 e a maquinaria de reparo por HDR presente em células de *E. coli*.

O cassete foi obtido por PCR com a enzima *Taq* DNA Polymerase, recombinant (Invitrogen) e com a utilização dos *primers* *sphi_F2* e *sphi_R3*, que continham uma cauda de 50 pb referentes às sequências de homologia

com o DNA alvo (um braço de homologia no primer 5' e outro no primer 3') e 20 pb que se anelavam à sequência do fragmento *spec-φ3Tim*, utilizado como DNA molde. Por fim, o cassete de DNA a ser integrado no genoma de *E. coli* foi formado pelo fragmento *spec-φ3Tim* flanqueado pelas regiões de homologia, com um tamanho de 2,757 kb.

A figura 4 ilustra como o fragmento *spec-φ3Tim* foi formado e mostra os *primers* utilizados para a construção do cassete de integração contendo as regiões de homologia.

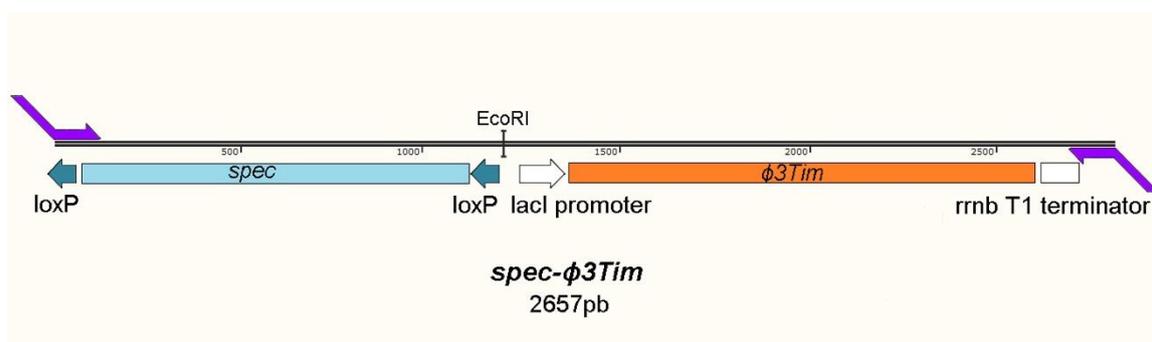


Figura 4. Esquema do fragmento *spec-φ3Tim*. Gene *spec* em azul após a ligação com o gene *φ3Ti* otimizado, em laranja, sob controle do promotor *LacI* e da região terminadora de transcrição *rrnB T1* (branco). Os *primers* utilizados na PCR para a construção do cassete de integração estão indicados em roxo.

Construção do fragmento que codifica para o sgRNA

O fragmento de DNA contendo o RNA guia (sgRNA) do sistema CRISPR/Cas9 utilizado na ligação com o vetor *AraC_pBAD_Cas9* e coexpresso com a proteína *Cas9* consistiu de 4 segmentos, a saber:

- i. o promotor constitutivo sintético *pj23119*;
- ii. o crRNA constituído por 20 nucleotídeos complementares à região alvo no genoma de *E. coli* – sequência proto espaçadora (parte variável do crRNA) e 12 nucleotídeos pertencentes à porção não variável do crRNA (forma um grampo com parte do tracrRNA);
- iii. 30 nucleotídeos referentes ao tracrRNA dos quais 4 nucleotídeos formam a parte truncada deste segmento, responsável por produzir a

- extremidade do grampo da estrutura do sgRNA;
- iv. 40 nucleotídeos derivados de um terminador de *S. pyogenes*, que demonstrou ser funcional para células bacterianas e importante na estabilização da estrutura do RNA guia (Larson, M. *et al.*, 2013).

A molécula de RNA guia formada pelo fragmento construído neste trabalho trata-se de um RNA guia quimérico gerado pela fusão da extremidade 3' do crRNA com a extremidade 5' de uma versão truncada de tracrRNA de *S. pyogenes*. Já foi descrito por Jinek *et al.* (2012) que a proteína Cas9 pode ser eficientemente programada para a clivagem correta no DNA alvo utilizando-se uma única molécula de RNA guia modificada. Diferentes construções de RNA guia fusionadas já foram projetadas e testadas, cada uma com diferentes comprimentos de sequências para o tracrRNA em sua extremidade 3' e a capacidade de edição com fragmentos únicos de sgRNA foi confirmada (Chu *et al.*, 2014).

A figura 5 demonstra a estrutura original da molécula de sgRNA do sistema CRISPR/Cas9, proveniente de *S. pyogenes*, assim como a molécula quimérica deste sgRNA desenvolvida por outros estudos para a edição genômica com este sistema.

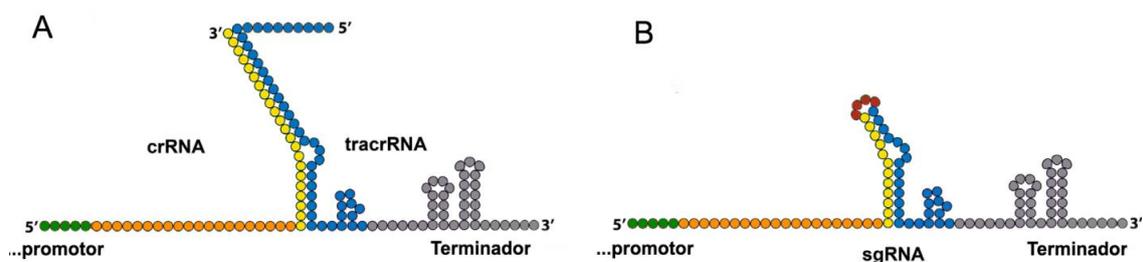


Figura 5. Estrutura da molécula de RNA guia. (a) Nucleotídeos para o RNA guia formado por moléculas separadamente expressas de tracrRNA e crRNA, originalmente presente em *S. pyogenes*, e (b) RNA guia formado por um único RNA sintético com todos os componentes expressos por somente um fragmento de DNA, como o utilizado neste trabalho. Figura adaptada de (Nowak *et al.*, 2016) com modificações.

A sequência de sgRNA foi montada com sítios para as enzimas de restrição EcoRI e XbaI para que, após digerido, o fragmento pudesse se ligar ao vetor AraC_pBAD_Cas9 linearizado, em um desses sítios. Também um sítio de restrição para XhoI foi adicionado entre o promotor pj23119 e o início da região de complementaridade com o DNA alvo, permitindo que futuramente o sgRNA pudesse ser utilizado com outro promotor, caso necessário.

O fragmento de 20 pb referente à região proto-espaçadora que compõe o crRNA foi desenhado com a ajuda do sítio eletrônico Crispr RGEN Tools. Essa ferramenta de bioinformática é capaz de buscar possíveis sequências alvo para o RNA guia levando em consideração o micro-organismo a ser modificado, a região alvo que se quer editar, a sequência PAM necessária para a ação da endonuclease Cas9 (NGG) e a probabilidade de ocorrência de *off-targets* (clivagem em sítios incorretos) para aquela sequência.

O tracrRNA é responsável por formar a maior parte da estrutura tridimensional do RNA guia, conhecida como o arcabouço do arranjo estrutural da molécula. Este trabalho utilizou uma sequência padrão que codifica para um tracrRNA truncado, retirada da literatura, assim como o promotor sintético pj23119 comumente usado para a expressão do RNA guia do sistema CRISPR/Cas9. A escolha deste promotor deveu-se a utilização frequente deste elemento para a expressão de tracrRNA, que por ser capaz de iniciar a transcrição em um sítio bem definido, se mostra ideal para a expressão de moléculas pequenas como o sgRNA (Jiang *et al.*, 2015).

A tabela 5 mostra as sequências utilizadas para cada parte do fragmento que expressa o sgRNA e de onde estas sequências foram obtidas, e a figura 6 apresenta a construção completa deste fragmento.

Tabela 5. Sequências usadas para a montagem do sgRNA

Fragmento	Sequência de DNA	Fonte
Promotor (pj23119)	TTGACAGCTAGCTC AGTCCTAGGTATAAT	(Jiang <i>et al.</i> , 2015)
crRNA (Região proto-espaçadora) (Região conservada)	<u>TCAACCATTAAACC</u> <u>GATTCGGTTTTAGA</u> GCTA	Proto-espaçadora: este trabalho Região conservada: (Jinek <i>et al.</i> , 2012)
tracrRNA (Região truncada) (Região conservada)	<u>GAAATAGCAAGTTAA</u> AATAAGGCTAGTCCG	(Jinek <i>et al.</i> , 2012)
Terminador derivado de <i>S. pyogenes</i>	TTATCAACTTGAAAA AGTGGCACCGAGTC GGTGCTTTTTT	(Larson, M. <i>et al.</i> , 2013)
Sítio alvo escolhido (Gene <i>pyrF</i> de <i>E. coli</i>) (Região complementar ao crRNA) (^ = sítio de clivagem da Cas9) (Sequência PAM)	<u>TCAACCATTAAACC</u> <u>GAT^TCGCGG</u>	Este trabalho

**CAGTGAATTCTTGACAGCTAGCTCAGTCCTAGGTATAATCAGCTCGAGTCATCAA
CCATTAAACCGAT^TCGGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAG
TCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTCTAGACTAG**

**EcoRI + promotor + XhoI + N20 + crRNA_{conservado} + TracrRNA_{região de ligação} +
TracrRNA_{conservado} + terminador + XbaI**

Figura 6. Sequência completa do fragmento que codifica para o sgRNA. A região representada por (^) indica a região complementar ao sítio de clivagem da endonuclease Cas9 no DNA alvo.

O fragmento completo de sgRNA contendo as sequências que codificam para o crRNA e tracrRNA, juntamente com o promotor e o terminador, foi obtido pelo uso de dois *primers* específicos com uma região complementar entre eles que, com a utilização da enzima DNA Polymerase I, Large (Klenow) Fragment (New England Biolabs), foram anelados e as regiões que não se

complementavam foram preenchidas para a formação de uma dupla fita, produzindo o fragmento completo de dupla fita de DNA. O protocolo de montagem da reação e os ciclos utilizados estiveram de acordo com o descrito pelo fabricante da enzima Klenow.

A figura 7 ilustra como foi a estratégia utilizada para a construção do fragmento de DNA contendo o sgRNA.

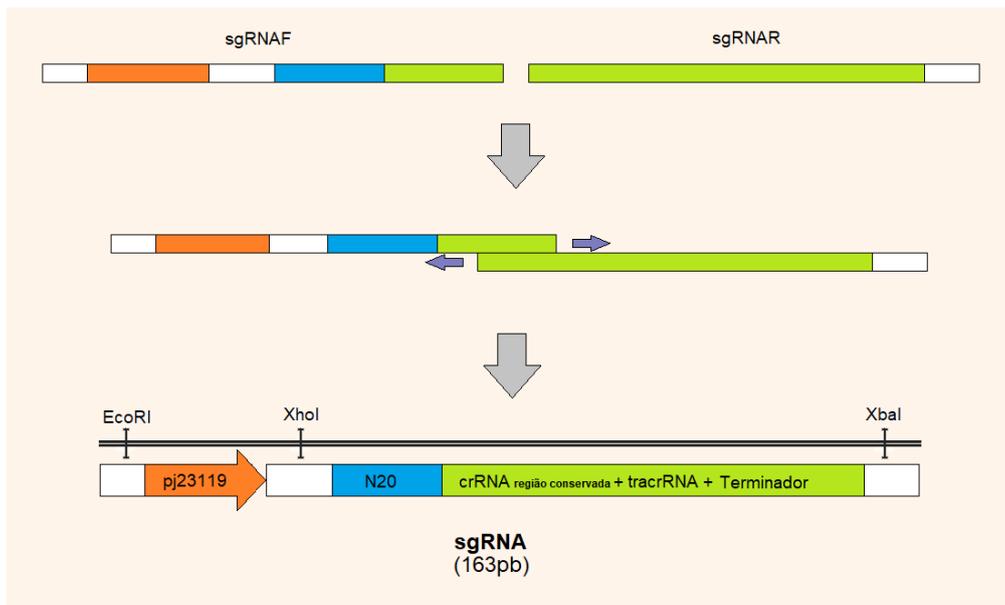


Figura 7. Representação esquemática da construção do sgRNA. Esquema da estratégia utilizada para a síntese do fragmento de dupla fita de DNA expressando o sgRNA a ser utilizado no sistema CRISPR/Cas9 deste trabalho.

A figura 8 mostra a visualização em gel de agarose deste fragmento formado.

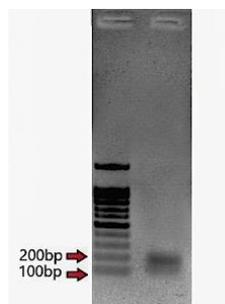


Figura 8. Amplificação do fragmento que expressa o sgRNA. Análise em gel de agarose do produto da polimerização realizada com a enzima DNA Polymerase (Klenow) Fragment, utilizando-se os *primers* sgRNAF e sgRNAR. O tamanho do fragmento gerado foi de acordo

com o esperado para o produto final da reação (163 bp). Marcador: 100 bp DNA ladder (Sinapse inc.)

5.2. Construção de plasmídeos

Plasmídeo AraC_pBAD_CRISPR

A montagem do fragmento sgRNA com o vetor AraC_pBAD_Cas9 digerido com a enzima EcoRI foi realizada utilizando o kit In-Fusion HD EcoDry Cloning Plus (Takara). Foram utilizados os *primers* infusionF e infusionR para a amplificação do fragmento de DNA contendo o sgRNA e as regiões de homologia com o vetor AraC_pBAD_Cas9, e as condições da reação estiveram de acordo com as instruções do fabricante.

O sistema foi utilizado para transformar *E. coli* DH10B competentes pelo método de choque térmico. O DNA plasmidial foi extraído e a montagem do fragmento foi confirmada por PCR e digestão do produto de midiprep com EcoRI e XbaI: com o fragmento clonado, era esperado que a dupla digestão com EcoRI e XbaI liberasse um fragmento de 149 pb. A digestão com somente uma dessas enzimas deveria linearizar o vetor. A PCR realizada com os mesmos *primers* utilizados com o kit In-Fusion deveria amplificar um fragmento de 193bp, referente ao tamanho do fragmento de sgRNA de 163 bp e os 30 pb para as regiões de homologia com o vetor, presentes nos *primers*.

A figura 9 mostra o gel de confirmação da montagem do plasmídeo AraC_pBAD_CRISPR contendo o sgRNA.

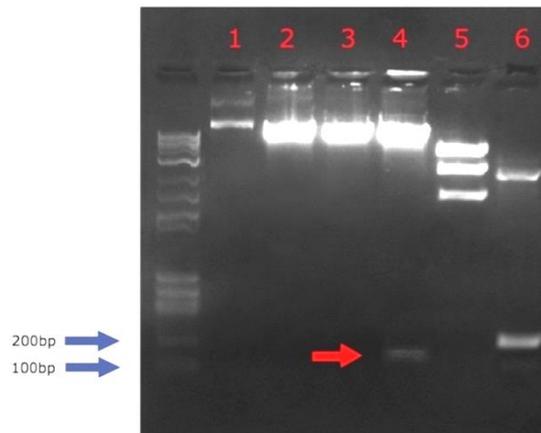


Figura 9. Confirmação da integração do sgRNA em AraC_pBAD_Cas9. Eletroforese em gel de agarose 2% corado com brometo de etídeo apresentando o vetor (1) intacto, (2) digerido com EcoRI, (3) digerido com XbaI, (4) digerido duplamente com EcoRI e XbaI e (5) digerido com SacI para confirmar o padrão de digestão para o plasmídeo AraC_pBAD_Cas9 (3 sítios para SacI) e (6) a PCR do fragmento de sgRNA com os primers utilizados para a clonagem por in-Fusion. M: 100 bp DNA ladder (Sinapse inc.).

Como esperado, foi liberado o fragmento de DNA para o sgRNA na digestão dupla, assim como houve a amplificação deste fragmento pela PCR. O vetor resultante da ligação com o fragmento sgRNA foi denominado AraC_pBAD_CRISPR.

A figura 10 ilustra o mapa físico do vetor AraC_pBAD_CRISPR.

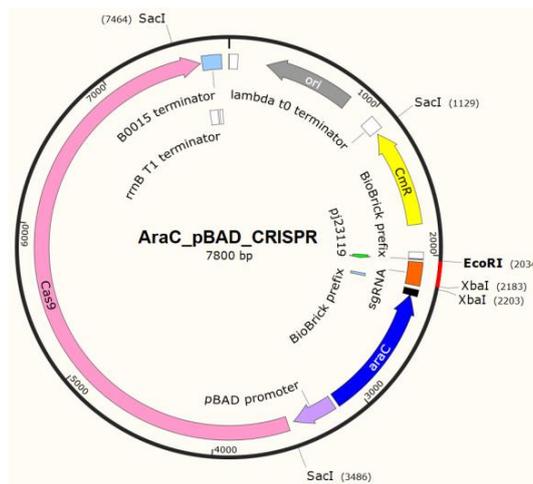


Figura 10. Mapa físico do vetor AraC_pBAD_CRISPR. Mapa físico do vetor AraC_pBAD_CRISPR expressando a endonuclease Cas9 sob controle do promotor induzível pBAD e o sgRNA, sob controle do promotor constitutivo pj23119. São mostradas as enzimas de restrição e os sítios utilizados para as digestões.

Plasmídeo pBlueScript II SK (+) contendo o cassete de integração

A construção do plasmídeo pBlueScript II SK (+) contendo o cassete de integração foi realizada por ligação do vetor linearizado com EcoRV com este fragmento de DNA. A reação de ligação foi incubada a 4°C por 16 h. A clonagem e replicação do vetor foi realizada em *E. coli* DH10B pelo método de choque térmico, a seleção de células transformadas foi realizada em antibiótico adequado (ampicilina e espectinomicina como marca de seleção para o plasmídeo pBlueScript II SK (+) e para o fragmento, respectivamente) e o DNA plasmidial foi posteriormente extraído com o kit Wizard® Plus Minipreps DNA Purification System (Promega). A ligação do fragmento foi confirmada pelo padrão de digestão do produto de miniprep com a enzima BamHI e o vetor resultante foi denominado pBlue- ϕ 3Tim.

O mapa físico do vetor pBlueScript II SK (+) contendo o cassete de integração está representado na figura 11.

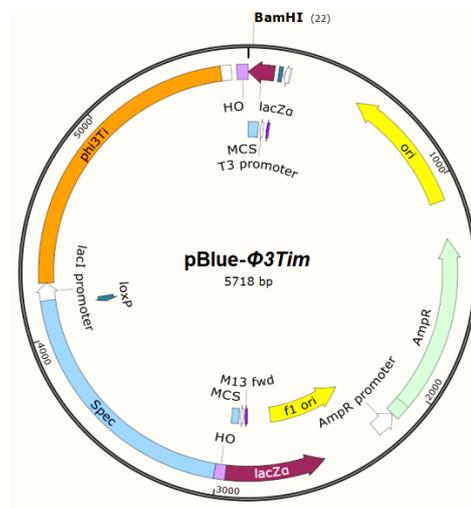


Figura 11. Mapa físico do vetor pBlue- ϕ 3Tim. Vetor pBluescript II SK(+) ligado ao cassete de integração, utilizado na transformação de *E. coli*. É mostrado o sítio utilizado para a digestão.

5.3. Transformação de *E. coli* DH10B

Transformação com o cassete de integração na forma de fragmento linear

Com o objetivo de diminuir a atividade exonuclease da enzima RecBCD de *E. coli* e, conseqüentemente, evitar a degradação do cassete de integração a ser transformado na forma de dupla fita de DNA linear, foi utilizado o método de transformação por eletroporação para essas células. Foi mostrado também que além da reduzida degradação de fragmentos lineares, a recombinação cromossomal em células de *E. coli* é eficiente quando utilizado este método de transformação (Karoui, El *et al.*, 1999).

Para demonstrar a atividade do sistema CRISPR/Cas9 como ferramenta determinante no processo de integração do fragmento no local esperado, foi feito um controle com células não transformadas com o vetor AraC_pBAD_CRISPR e células sem a indução da transcrição por arabinose da endonuclease Cas9 deste sistema.

Para a transformação de *E. coli* com o cassete de integração na forma de fragmento linear de DNA, as células foram previamente transformadas com o vetor AraC_pBAD_CRISPR e selecionadas em meio contendo cloranfenicol. Quando então transformadas com o cassete de integração, as células foram semeadas em meio YM-ágar contendo 5-FOA. Foi observado o aparecimento de colônias apenas quando a concentração de 5-FOA foi reduzida a pouco mais da metade descrita por Meng *et al.*, (2006), no entanto, foi observado um crescimento similar na placa controle, com células não transformadas com o cassete de integração, sugerindo que, embora várias concentrações de 5-FOA tenham sido testadas, as cepas selvagens também eram capazes de sobreviver.

Posteriormente ao reanalisar a construção do sgRNA utilizado, foi observado que a região proto-espaçadora estava direcionada para um sítio adjacente à região codante do gene *pyrF* no genoma de *E. coli*, indicando que este gene não havia sido interrompido com a atuação do sistema CRISPR/Cas9. Esses dados sugeriram que o crescimento obtido em placa

poderia ter sido decorrente das baixas concentrações de 5-FOA, permitindo o crescimento de colônias resistentes à baixa toxicidade deste composto para aquelas concentrações.

O gene *pyrF*, responsável para codificar a orotidina-5'-fosfato descarboxilase, uma enzima envolvida na síntese de pirimidinas, quando inativado é conhecido por conferir resistência a 5-FOA e auxotrofia a uracila, funcionando então como uma marca de seleção positiva e negativa. A utilização desta ferramenta genética seria particularmente interessante neste estudo por permitir a construção de mutantes selecionáveis com ausência de marcadores (Bocioaga *et al.*, 2014), no entanto, apesar deste sistema de seleção ser comumente utilizado, sobretudo em leveduras, alguns estudos demonstraram que a pressão seletiva em 5-FOA para a seleção de crossover duplo é um processo demorado e com baixa eficiência de nocaute para este gene, assim, a velocidade e facilidade do processo poderiam ser significativamente aumentados usando um sistema de seleção mais forte (Mougiakos *et al.*, 2017).

Em virtude de tudo isso, a seleção passou a ser realizada apenas em meio adicionado de espectinomicina. Na placa controle, onde foi semeada a bactéria submetida à eletroporação sem o cassete de integração, não houve crescimento.

A figura 12 mostra as placas resultantes da transformação de *E. coli* realizada com o cassete de integração na forma de fragmento de DNA linear.

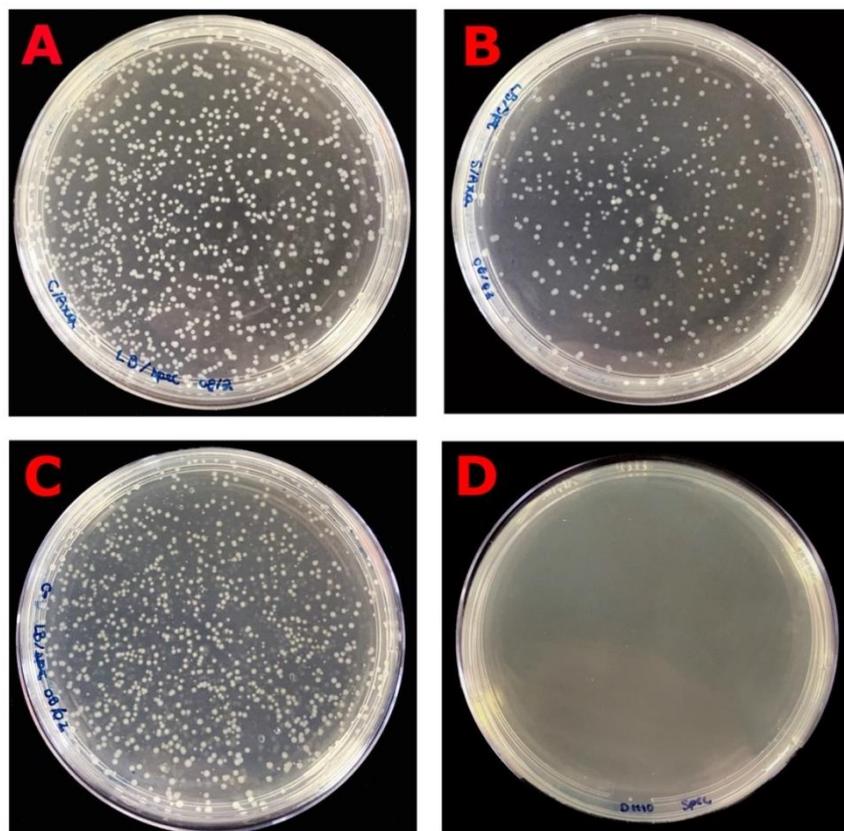


Figura 12. Análise de transformantes de *E. coli* com o cassete de integração linear. Resultado da transformação de *E. coli* DH10B com o fragmento linear contendo o cassete de integração. (A) Células transformadas com o vetor AraC_pBAD_CRISPR induzido por arabinose, (B) células transformadas com o vetor AraC_pBAD_CRISPR não induzido (C) células não transformadas com o vetor AraC_pBAD_CRISPR e (D) controle: células eletroporadas sem adição do cassete de integração. Colônias selecionadas em placas com meio LB adicionado de espectinomicina após 16 horas de incubação a 37°C.

Contudo, o crescimento foi similar para as transformações com e sem indução da proteína Cas9, assim como na transformação sem o vetor AraC_pBAD_CRISPR, sugerindo que o sistema CRISPR/Cas9 poderia não ter tido influência na integração do cassete contendo o fragmento *spec-Φ3Tim*. Estes resultados sugeriram que o sistema CRISPR/Cas9 pode ser desnecessário para a modificação genômica em *E. coli* quando o sistema de reparo por HDR também está sendo utilizado como ferramenta de edição.

Transformação com o vetor PBlue- Φ 3Tim

Já foi observado que durante eventos de recombinação homóloga em *E. coli* a utilização de plasmídeos linearizados foi mais eficaz do que em sua forma circular. Isso pode estar associado com a quebra aleatória de plasmídeos circulares em locais não desejados, como nos braços de homologia (Song e Stieger, 2017).

Apesar de a maioria dos estudos com o sistema de reparo por homologia terem utilizado DNA doador na forma de fragmento linear, existe uma taxa considerável de degradação destes fragmentos pelo complexo recBCD citado anteriormente. Tendo em vista esta instabilidade de fragmentos lineares, foi testada também a transformação de *E. coli* com o cassete de integração ligado a um vetor circular (Lovett *et al.*, 2002).

Outra vantagem da transformação com o fragmento clonado em um vetor é a possibilidade de realizar duas etapas de seleção, com a marca do vetor e a do fragmento *spec-φ3Tim*. Além disso, foi possível submeter as células transformadas à indução por arabinose por mais tempo, com menos chances de as células perderem o fragmento linear neste processo.

Para a transformação com o vetor PBlue- $\phi 3Tim$, apenas uma etapa de eletroporação foi realizada para a introdução de ambos os vetores AraC_pBAD_CRISPR e PBlue- $\phi 3Tim$, também como já descrito. Para esta transformação, as células foram semeadas em meio LB adicionado de ampicilina (marca de seleção para o vetor pBlueScript II SK (+) e, posteriormente, as colônias crescidas foram transferidas para uma placa adicionada de espectinomicina como uma segunda etapa de seleção de transformantes. Para ambos os cultivos, as colônias começaram a surgir após 16 h de incubação. Também foi realizado um controle com células não transformadas com o vetor AraC_pBAD_CRISPR para analisar a atividade do sistema CRISPR/Cas9 no processo de integração do fragmento. Para as células submetidas a eletroporação sem PBlue- $\phi 3Tim$, não houve crescimento.

A figura 13 mostra a segunda etapa de seleção (em espectinomicina) de transformantes de *E. coli* com o vetor PBlue- $\phi 3Tim$.

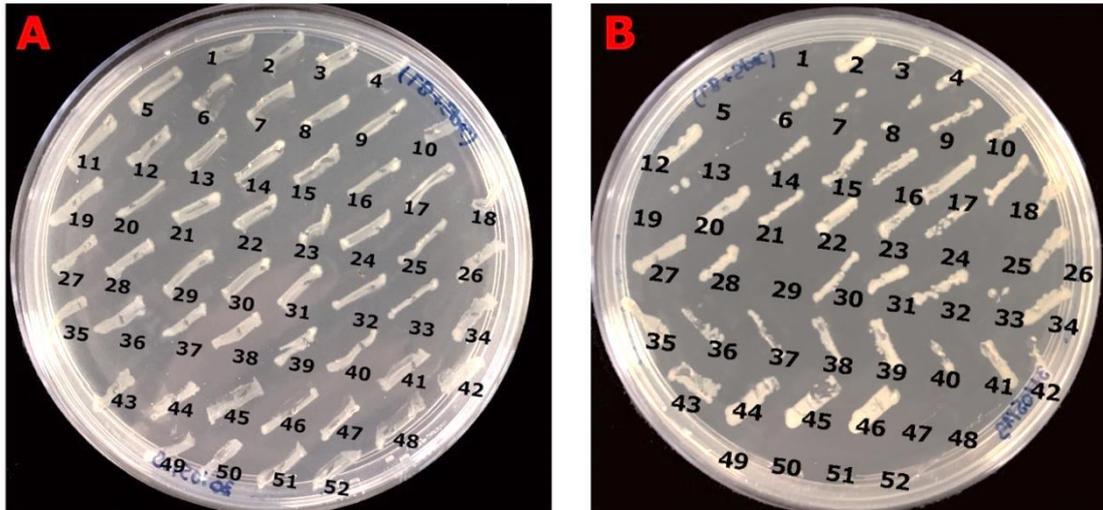


Figura 13. Análise de transformantes de *E. coli* com o vetor pBlue- ϕ 3Tim. (A) Resultado da transformação de *E. coli* DH10B com o vetor AraC_pBAD_CRISPR e indução da transcrição da endonuclease Cas9, (B) 1–43: transformação com o vetor AraC_pBAD_CRISPR e indução da Cas9, 44–46: células não transformadas com o vetor AraC_pBAD_CRISPR, 47–52: controle negativo, transformação sem o vetor pBlue- ϕ 3Tim. Colônias selecionadas em placas com meio LB adicionado de espectinomicina após 16 h de incubação a 37°C.

Igualmente ao crescimento observado para a transformação com o cassete de integração na forma de fragmento linear, o crescimento também ocorreu para as transformações sem o vetor AraC_pBAD_CRISPR, sugerindo novamente que o sistema CRISPR/Cas9 não influenciou na integração do cassete.

Trabalhos anteriores demonstraram que, em comparação com outras enzimas, a atividade de Cas9 é baixa, com pouca rotatividade. Essa análise feita em células bacterianas demonstrou que, uma vez ligada à sequência de DNA alvo, o deslocamento de Cas9 na fita de DNA pode ainda ser um desafio. Assim, a Cas9 acaba atuando como uma molécula de disparo único, o que pode ser indesejável para aplicações que requerem a sua atividade catalítica eficiente (Jinek *et al.*, 2012; Lino *et al.*, 2018).

Além da atividade da proteína Cas9, outros fatores podem influenciar a atividade do sistema CRISPR/Cas9 em *E. coli*. Já foi demonstrado em células de humanos que o aumento da concentração relativa de sgRNA para a proteína Cas9, por exemplo, revelou estar ligado ao aumento da atividade de

corde no alvo, presumivelmente assegurando que todas as proteínas Cas9 estavam formando o complexo ribonucleoproteico ativo, no entanto, este mesmo estudo demonstrou que o sgRNA excessivo também pode aumentar os efeitos fora do alvo (Fu et al., 2013).

Já foi demonstrado que diferentes RNAs guia possuem diferentes eficiências de direcionamento, podendo variar grandemente quanto a sua capacidade para guiar a Cas9 (Xu, Duan e Chen, 2017). Estudos recentes analisaram diferentes RNAs guia selecionados aleatoriamente e identificaram diversas características sequência específicas e da posição alvo no genoma que são preditivas para a construção de RNAs guia funcionais, como por exemplo a acessibilidade à cromatina, a formação da estrutura de ligação entre o DNA alvo e do sgRNA e os parâmetros de energia e entropia essenciais para a ligação da molécula de sgRNA (Xu, Duan e Chen, 2017; Wong, Liu e Wang, 2015;). Ainda é um desafio crítico nos estudos moleculares baseados em CRISPR separar RNAs guia ativos de inativos para garantir e quantificar a edição genômica (Schwartz *et al.*, 2018). A possibilidade de se ter utilizado um RNA guia pouco eficaz ou não funcional é uma opção a ser considerada, uma vez que o uso de sgRNAs potentes é um gargalo essencial para a edição eficiente. Tipicamente, múltiplos RNAs guia são projetados e testados para direcionar a Cas9 para um mesmo gene de interesse, levando em consideração a possibilidade de que alguns destes podem não ser funcionais. Assim uma alternativa futura para garantir a funcionalidade do sistema CRISPR/Cas9 é a construção de mais de uma molécula de RNA guia.

5.4. Confirmação da integração cromossomal do cassete

Para confirmar se o cassete estava sendo integrado no DNA cromossomal de *E. coli* foram feitas extrações de DNA genômico de 4 clones resultantes da transformação com o fragmento linear de DNA e 4 clones resultantes da transformação com pBlue- ϕ 3Tim para realização de PCRs com *primers* que se anelavam à região interna do cassete de integração. A correta

integração revelaria fragmentos amplificados de 2.1 kb para o par de *primers* specF e phiR, conforme apresentado na figura 14.

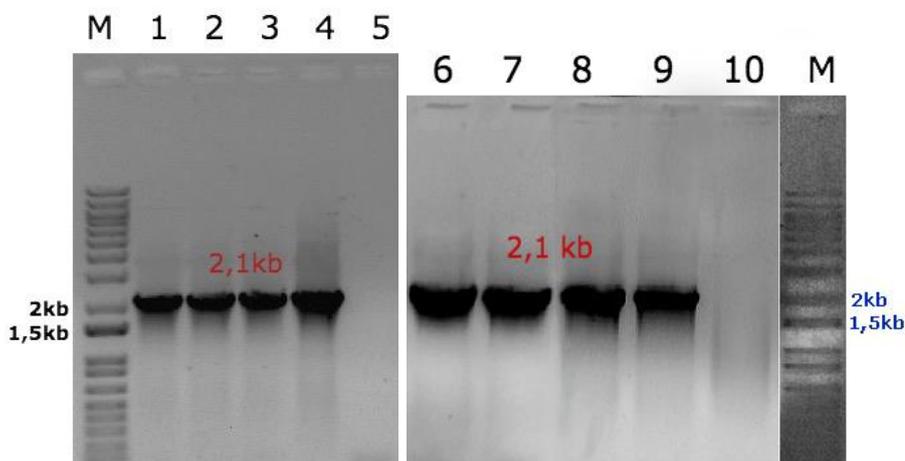


Figura 14. Análise da integração por PCR com DNA genômico de transformantes. Eletroforese em gel de agarose 0,8% corado com brometo de etídeo. M:1 kb Plus DNA Ladder (Thermo Scientific). PCR com DNA genômico dos clones transformados com o cassete de integração. 5 e 10: PCR com o DNA genômico de *E. coli* DH10 não transformadas, 1-4: PCR de colônias transformadas com pBlue- ϕ 3Tim utilizando os primers specF/phiR, 6-9: PCR de colônias transformadas com o cassete de integração na forma de fragmento linear utilizando os pares de *primers* specF/phiR.

Como observado na figura 13, todos os clones apresentaram as bandas correspondentes ao tamanho esperado para a amplificação do fragmento *spec- ϕ 3Tim* com os *primers* em questão. Assim, os clones recombinantes foram selecionados e armazenados para aplicação posterior.

5.5. Confirmação do local de integração do fragmento *spec- ϕ 3Tim*

5.5.1. PCR da região alvo do sistema CRISPR/Cas9

Para verificar se a integração cromossomal ocorreu no local desejado foram feitas PCRs com primers que se anelavam em sítios adjacentes à região de clivagem da proteína Cas9, no gene *pyrF* do genoma de *E. coli*. Era esperado que esta endonuclease fosse direcionada pelo sgRNA modificado

com a sequência guia complementar ao sítio alvo na sequência do gene *pyrF*, sem que houvesse *off-targets* para outras regiões do DNA, como determinado pelo site RGEN Tools no momento de escolha da sequência protoespaçadora da molécula de RNA guia.

Com o propósito de testar o máximo de colônias transformadas com o cassete de integração foram realizadas PCRs de colônia com o material genético de até 10 colônias por reação de PCR, permitindo a análise da integração de fragmentos em até 10 colônias por vez. Foram testadas 100 colônias obtidas da transformação com o cassete de integração na forma de fragmento linear de DNA, e 93 colônias para a transformação com o vetor pBlue- ϕ 3Tim.

Foram utilizados os *primers* pyr3.2 e pyr4 para a amplificação da região alvo no genoma de *E. coli*. Era esperado que fosse gerado um fragmento de aproximadamente 4kb caso a integração por HDR tivesse ocorrido no local desejado. Caso não tivesse ocorrido a integração para aquela região, deveria ser amplificado somente um fragmento de 1,3kb.

A figura 15 mostra o resultado da PCR de diferentes transformantes com os *primers* pyr3.2 e pyr4.

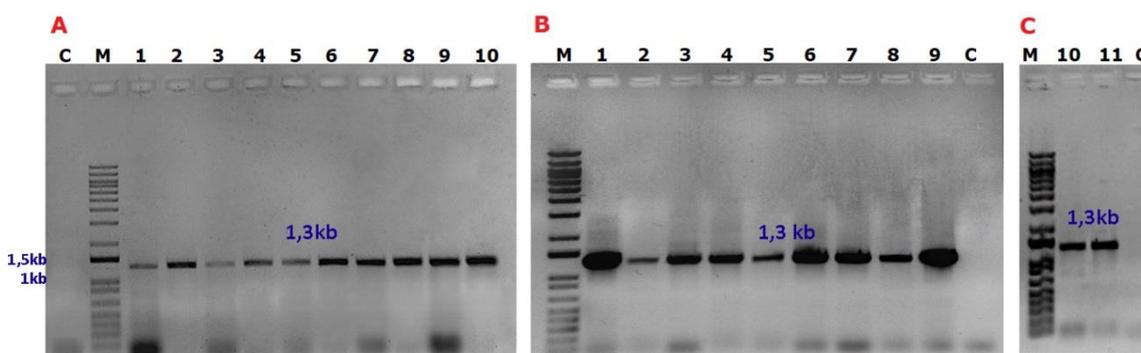


Figura 15. PCR de colônia de clones de *E. coli*. Eletroforese em gel de agarose 0,8% corado com brometo de etídeo. **M:** 1 kb Plus DNA Ladder (Thermo Scientific). PCR de colônia de vários clones de *E. coli* transformados com o cassete de integração. **A:** Clones obtidos através das transformações com o fragmento linear de DNA (1-9: Cas9 induzida, 10: Transformações sem o vetor AraC_pBAD_CRISPR ou sem a indução da transcrição da Cas9). **B e C:** Transformações com o vetor pBlue- ϕ 3Tim (1-10: Indução da transcrição da Cas9, 11: Controles sem a transformação com o vetor AraC_pBAD_CRISPR)

Como observado na figura 15, não foi verificada a banda de 4kb para nenhum dos clones avaliados, mostrando que não houve integração do fragmento *spec-φ3Tim* no sítio desejado. Esses resultados indicaram que o sistema CRISPR/Cas9 não atuou na clivagem do DNA alvo, ou que o reparo por HDR não ocorreu com as regiões de homologia do cassete de integração.

A utilização de regiões curtas de homologia para o DNA doador já foi reportada em diferentes artigos que aplicaram um sistema eficiente de recombinação em *E. coli* (Pyne *et al.*, 2015; Zhao *et al.*, 2016). Foi demonstrado que a utilização de braços homólogos de curta sequência é capaz de formar um método rápido de edição, no entanto, acaba por limitar alguma eficiência deste processo. Um aumento no tamanho das regiões de homologia poderia representar uma melhora na eficiência de recombinação, permitindo então a detecção de eventos recombinatórios de baixa eficiência. Já foi descrito por Jiang *et al.* (2015) um aumento significativo no percentual de recombinação para eventos de edição genômica utilizando-se o sistema CRISPR/Cas9 em conjunto com o DNA doador formado por braços de homologia com mais de 200pb. Assim, explorar a estratégia de edição utilizada como o aumento das regiões de homologia do cassete de integração poderá ser um método de melhorar a capacidade de recombinação.

Outro estudo sobre clonagens por recombinação em *E. coli* baseada no reparo de pequenos intervalos de DNA de fita simples mostrou que vetores e insertos digeridos com endonucleases de restrição podem ser facilmente recombinados em células de *E. coli*. Embora esta situação seja diferente da integração cromossomal de fragmentos, este estudo explorou o envolvimento da maquinaria celular no processamento de dsDNA e formação de ssDNA *in vivo*, facilitando assim a recombinação de fragmentos por homologia. Este estudo utilizou cepas de *E. coli* DH5α *recA* negativas, sugerindo que um mecanismo de recombinação homóloga não convencional poderia estar envolvido (Jacobus e Gross, 2015). De fato, uma via de recombinação homóloga independente de *recA*, caracterizada por funcionar em substratos com extensões limitadas de homologia, foi descoberta em *E. coli*. Em ensaios específicos de recombinação, esta via é a única responsável por recombinar plasmídeos com pequenas regiões de homologia. A recombinação

independente de RecA, via conversão gênica ou crossing-over, é quase tão eficiente quanto a catalisada por RecA. Este tipo de recombinação emprega regiões de ssDNA. Assim como a recombinação dependente de RecA, sua eficiência também aumenta com o aumento da região de homologia, porém com menor dependência. A recombinação dependente de RecA pode precisar de uma região de homologia de até 150 pb, enquanto que recombinação independente de RecA, por outro lado, tem um limiar inferior a 50 pb para apresentar uma boa eficiência. (Dutra, Sutura e Lovett, 2007). Este procedimento pode ser útil para gerar mutações específicas, como deleções e inserções em um gene, sendo uma maneira rápida e eficaz para produzir proteínas recombinantes (Jacobus e Gross, 2015). Seria de interesse elucidar se esta via de recombinação homóloga enigmática em *E. coli* poderia ser responsável pela clonagem do fragmento *spec-φ3Tim* no genoma de *E. coli*, explicando a integração via recombinação em outro local do genoma com pequenas regiões de homologia.

A partir desses resultados, 6 clones foram selecionados para determinar qual havia sido o local de integração.

5.5.2. Utilização da técnica TAIL-PCR e sequenciamento dos fragmentos obtidos

Tendo como resultados a ausência de bandas confirmatórias da integração no *locus* desejado, porém, a presença do fragmento *spec-φ3Tim* no genoma, foi realizada uma técnica que permitiria encontrar o local de integração do cassete.

Assim, 6 clones cujo DNA genômico já havia sido extraído e a integração do fragmento *spec-φ3Tim* confirmada, foram selecionados para a aplicação de uma técnica cujo princípio se assemelha ao de TAIL-PCR e sequenciamento dos fragmentos obtidos.

A técnica de amplificação aleatória do DNA genômico aqui utilizada foi empregada com o objetivo de localizar, ao longo de uma grande sequência de DNA, a região de integração de um fragmento específico. Essa técnica não

requer grande conhecimento prévio do DNA molde a ser analisado e utiliza-se um único primer que, ao longo da PCR, é anelado com baixa especificidade a este DNA, produzindo diferentes bandas de DNA formadas por fragmentos menores visualizáveis por eletroforese que agora poderão ser extraídas do gel de agarose e analisadas separadamente.

A técnica utilizada se baseia na premissa de que um *primer* irá hibridizar em locais aleatórios com os quais ele tenha alguma complementariedade e iniciar a polimerização de DNA em mais de uma região ao longo do genoma.

Nas duas primeiras etapas deste procedimento, foi utilizado unicamente o *primer* specR200, que se anelava no início da sequência do fragmento *spec-φ3Tim*. A primeira etapa gerou várias bandas inespecíficas, como esperado, devido à baixa especificidade da reação de anelamento entre o *primer* e o DNA molde. Foram extraídas 4 a 6 bandas de cada amostra para realização da segunda etapa na qual cada banda selecionada foi amplificada por PCR com o mesmo primer.

Para todas as amostras testadas, as bandas mais evidentes geradas com a primeira etapa do processo tinham tamanhos de aproximadamente 2,1 kb, 1,6 kb, 1,2 kb, 800 pb e 575 pb.

As figuras 16 e 17 demonstram a estratégia e os resultados para a primeira etapa de TAIL-PCR.

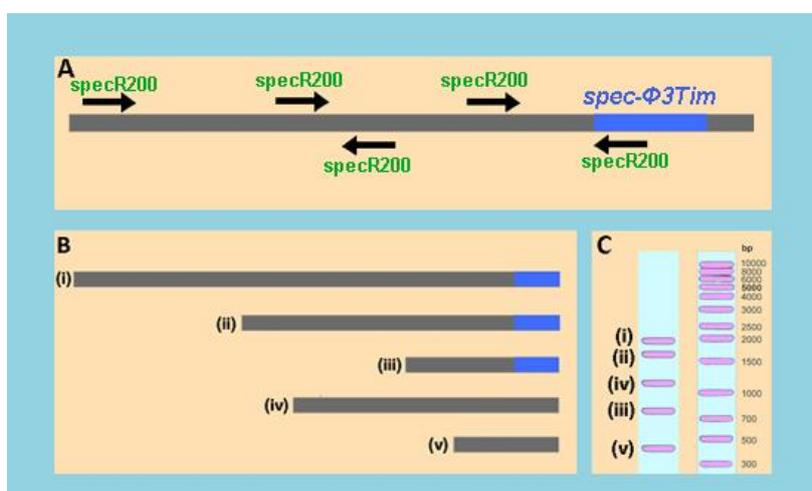


Figura 16. Estratégia para determinar o local de integração (1ª etapa). Estratégia utilizada para determinar o local de integração do fragmento *spec-φ3Tim* baseada na técnica TAIL-

PCR (1ª etapa). **A:** amplificação aleatória de fragmentos do DNA genômico (em azul a região do inserto), **B:** fragmentos gerados, **C:** visualização das bandas em gel de agarose

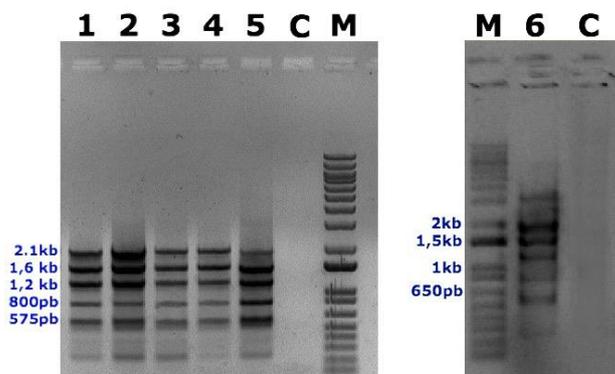


Figura 17. Amplicons obtidos após a realização da primeira etapa de TAIL-PCR. Eletroforese em gel de agarose 0,8% corado com brometo de etídeo. C: Controle negativo, M: 1 kb Plus DNA Ladder (Thermo Scientific). Primeira etapa de TAIL-PCR com 6 clones de *E. coli* DH10B contendo o fragmento *spec-φ3Tim* integrado ao genoma.

As figuras 18 e 19 mostram a estratégia utilizada e os resultados para cada amostra da segunda etapa de TAIL-PCR.

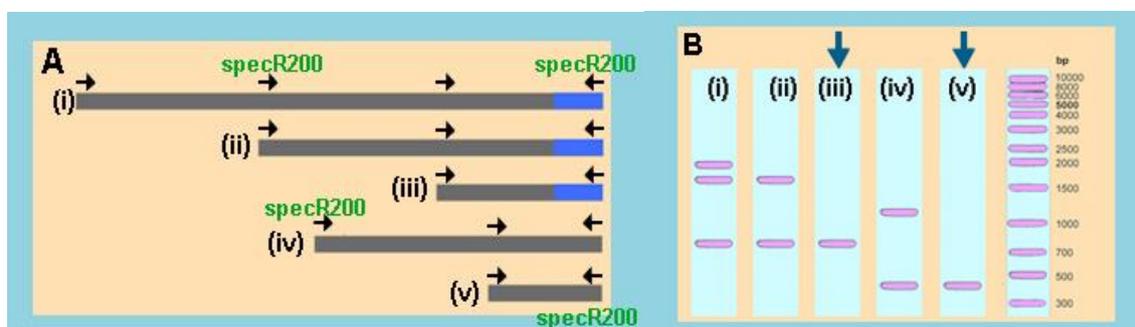


Figura 18. Estratégia para determinar o local de integração (2ª etapa). Estratégia utilizada para determinar o local de integração do fragmento *spec-φ3Tim* baseada na técnica TAIL-PCR (2ª etapa) **A:** amplificação de fragmentos para cada banda gerada anteriormente, utilizada como DNA molde, **B:** visualização dos fragmentos amplificados gel de agarose.

Nesta etapa era esperado que ao menos uma das reações de PCR gerasse um fragmento único de DNA, sugerindo que a banda utilizada como DNA molde para aquela reação incluía apenas dois sítios adjacente de anelamento do *primer* *specR200* ao genoma.

Para cada uma das amostras analisadas apenas uma reação de PCR gerou banda única. Essas bandas únicas foram novamente extraídas do gel e utilizadas para a realização da terceira etapa do processo: uma PCR com dois *primers* internos ao fragmento *spec-φ3Tim*, denominados specF e specR200.

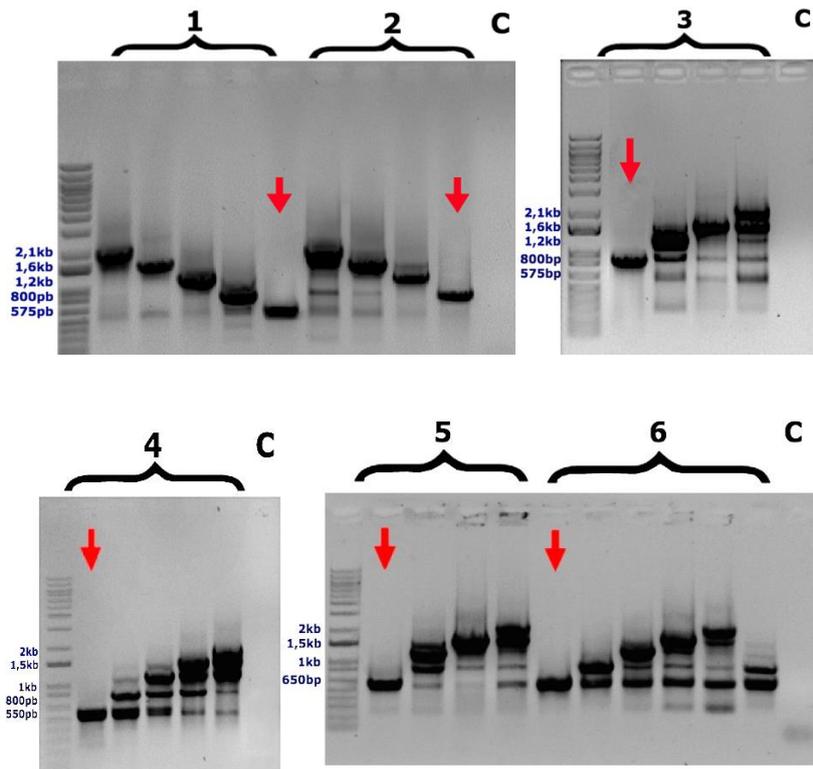


Figura 19. Amplicons obtidos após a realização da segunda etapa de TAIL-PCR. Eletroforese em gel de agarose 0,8% corado com brometo de etídeo. C: Controle negativo, M: 1 kb Plus DNA Ladder (Thermo Scientific). Segunda etapa de TAIL-PCR para as bandas extraídas de cada um dos 6 clones utilizados na etapa anterior. Estão indicadas em seta vermelha as reações que geraram uma única banda: estas foram extraídas e utilizadas como DNA molde para a PCR com os *primers* internos ao fragmento *spec-φ3Tim*.

As figuras 20 e 21 mostram a estratégia e os resultados para a terceira etapa do procedimento utilizado para determinar a localização aproximada do fragmento *spec-φ3Tim* no genoma de *E. coli* modificadas geneticamente.

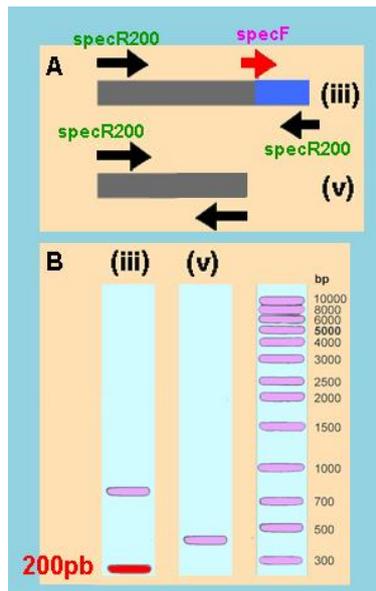


Figura 20. Estratégia para determinar o local de integração (3ª etapa). Estratégia utilizada para determinar o local de integração do fragmento *spec-φ3Tim* baseada na técnica TAIL-PCR (3ª etapa) **A:** PCR com dois *primers* internos ao cassete de integração para as amostras que geraram uma única banda na etapa anterior, **B:** visualização dos fragmentos amplificados em gel de agarose e confirmação da banda que contém o fragmento de interesse com base no tamanho esperado para esta PCR.

Para esta PCR, era esperado que fosse amplificado um fragmento de 200 pb caso parte do fragmento *spec-φ3Tim* estivesse presente na banda empregada como DNA molde.

Uma banda de 200 pb foi gerada para todas as amostras analisadas, de acordo com o esperado, confirmando então a presença de parte do fragmento *spec-φ3Tim* para as bandas utilizadas como DNA molde.

A figura 21 mostra o resultado para a PCR realizada com a utilização dos *primers* internos ao inserto.

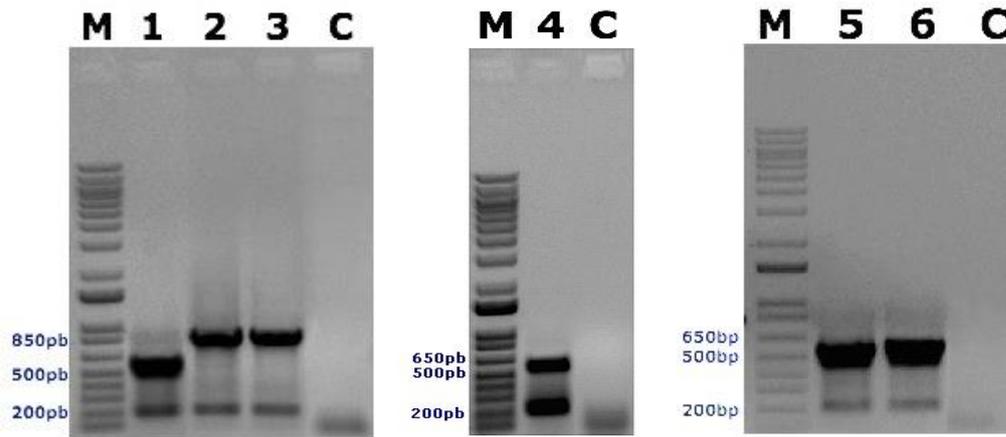


Figura 21. Amplicons obtidos após a realização da terceira etapa deste procedimento. Eletroforese em gel de agarose 2% corado com brometo de etídeo. M: 1 kb Plus DNA Ladder (Thermo Scientific), C: Controle negativo. Terceira etapa da estratégia utilizada para localizar o fragmento de *spec-φ3Tim*, com os *primers* internos ao fragmento

As bandas contendo este fragmento, com tamanhos aproximados de 850bp ou 575bp para as diferentes amostras, foram então purificadas e enviadas para o sequenciamento.

As regiões de integração que foram indicadas pelo sequenciamento estão representadas na figura 22.

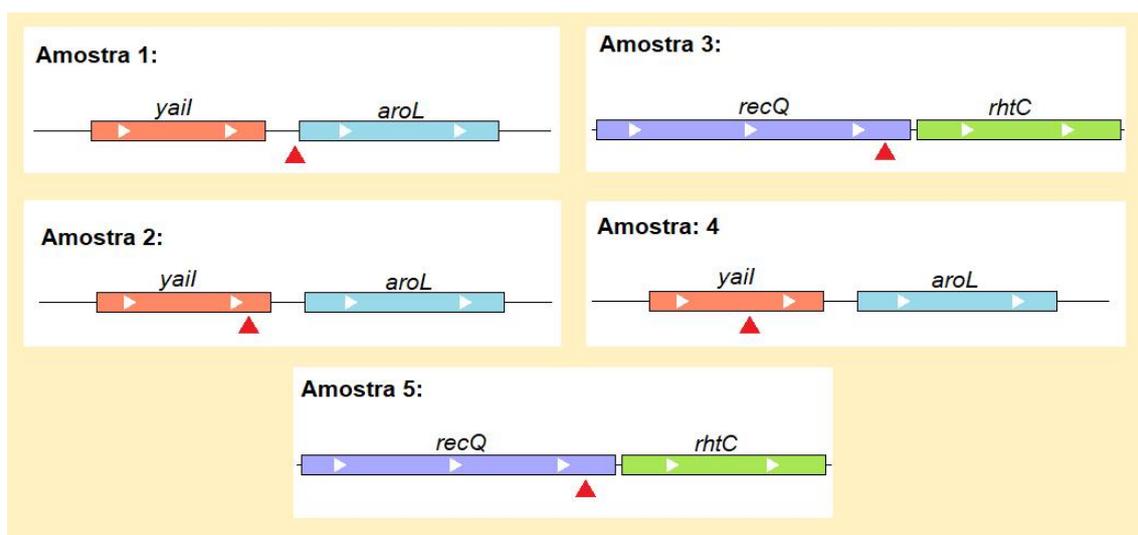


Figura 22. Resultado das regiões no DNA de *E. coli* DH10B reveladas pelo sequenciamento. As amostras 1, 2 e 4 indicaram seqüências de DNA complementares à região entre os genes *Yail* e *aroL* e as amostras 3 e 5 indicaram seqüências complementares à região final do gene *recQ*.

Os sítios de integração indicados pelo método de sequenciamento estão localizados em regiões no genoma de *E. coli* distantes da região alvo de integração desejada para este estudo, que havia sido determinada pelo RNA guia do sistema CRISPR e pelas regiões de homologia do cassete de integração: genes *Yail* e *aroL* localizados entre 344.319 – 344.777 nt, gene *recQ* localizado entre 4.102.807 – 4.104.636 nt e gene *pyrF* localizado entre 1.429.341 – 1.430.078 nt. Os resultados apresentados indicaram, ainda, não haver aleatoriedade no sítio de integração uma vez que esta integração ocorreu mais de uma vez no mesmo local.

No entanto, não foi encontrada nenhuma relação de homologia entre as sequências do cassete de integração e as regiões onde possivelmente esta integração ocorreu, bem como com a sequência do sgRNA utilizado. Uma possível explicação para estes resultados seria a ocorrência de outro evento recombinatório distinto de HDR bem como a acessibilidade e afinidade com o cassete de integração da estrutura do DNA de *E. coli* nas regiões mencionadas.

O estudo de Chayot *et al.* (2010) demonstrou que células de *E. coli* possuem uma atividade distinta de end-joining, também capaz de reparar DSBs e gerar rearranjos genômicos. Este mecanismo, denominado alternativo end-join (A-EJ), diferentemente do NHEJ clássico, é caracterizado pela remoção de regiões do genoma devido ao RecBCD e pelo uso de micro homologia. Embora geralmente mutagênico, o A-EJ é capaz de gerar aquisição de genes, refutando a ideia de que bactérias ganham sequências exógenas apenas por mecanismos dependentes de recombinação. A capacidade de *E. coli* de realizar reparo DSBs via A-EJ foi analisada pela introdução de uma variedade de fragmentos lineares de DNA nestas células para analisar se a reparação estava ocorrendo. Foi demonstrado que a eficiência do reparo foi comparável para extremidades cromossômicas com e sem complementariedade, confirmando que A-EJ é um mecanismo de reparo independente de recombinação e distinto do NHEJ, assim, células de *E. coli* podem integrar sequências exógenas não homogêneas, não relacionadas, por meio de junção de extremidade (Chayot *et al.*, 2010).

5.6. Remoção do gene que confere resistência a espectinomicina

Com o objetivo de melhor abranger a utilização das cepas mutantes para futuramente possibilitar a utilização desta plataforma de metilação com diferentes vetores para *Clostridium*, o gene *spec* utilizado como marca de seleção para estas linhagens foi removido por recombinação mediada pelo sistema Cre-Lox. O nocaute do gene *spec* flanqueado por sítos LoxP foi realizado pela recombinação com o vetor 705-Cre (Gene Bridges) de acordo com as instruções do fabricante. As cepas recombinantes foram crescidas em meio com antibiótico para analisar a sensibilidade das células. Também, a remoção do gene foi confirmada por PCR, utilizando-se dois *primers* internos ao gene da espectinomicina, *specF* e *specR200*.

Caso o gene *spec* tivesse sido nocauteado, não deveriam apresentar bandas para essas amostras, se o gene *spec* ainda estivesse presente, era esperado que um fragmento de 200 bp fosse amplificado. Para confirmar também que o gene $\phi 3Tim$ foi mantido após a recombinação com o vetor 705-Cre, foi realizada uma PCR com primers que se anelavam à região interna do gene da metiltransferase. Era esperado um fragmento de 684 pb para esta PCR caso o gene $\phi 3Tim$ tivesse sido mantido.

A figura 23 mostra a sensibilidade de cepas recombinadas pelo sistema Cre-Lox à espectinomicina. A figura 24 mostra a PCR realizada para estas cepas recombinadas, utilizando *primers* internos ao gene *spec* e *primers* internos ao gene $\phi 3Tim$.

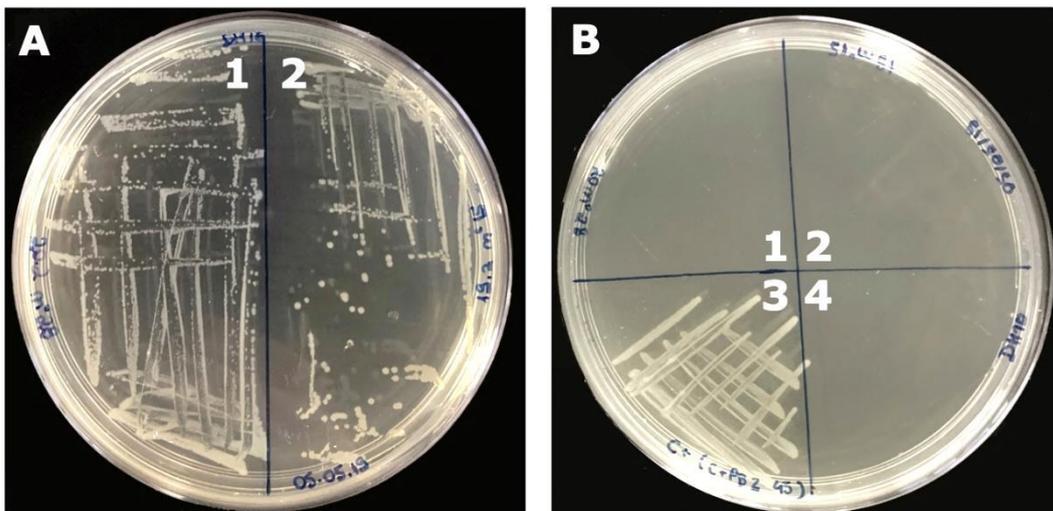


Figura 23. Cepas *E. coli* recombinadas com o vetor 705-Cre. (A) Placa controle contendo apenas meio LB, (B) placa contendo meio LB adicionado de espectinomicina. 1 e 2: Cepas recombinadas para remoção de *spec*, 3: Cepa mutante antes da recombinação, ainda com o gene *spec* e 4: *E. coli* DH10B selvagem. Colônias crescidas a 37°C por 18 horas.

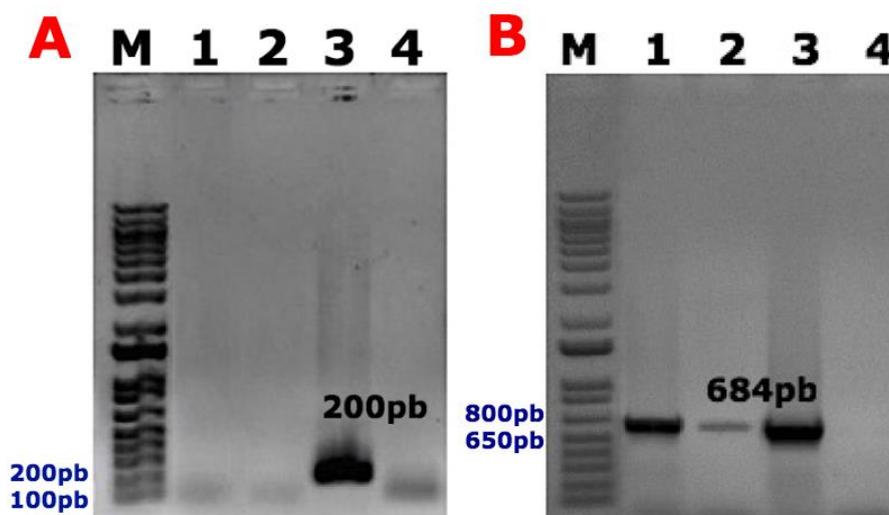


Figura 24. Amplicons obtidos por PCR de cepas recombinadas com 705-Cre. Eletroforese em gel de agarose 2% corado com brometo de etídeo. M:1 kb Plus DNA Ladder (*Thermo Scientific*). PCR para a confirmação da remoção de *spec* e preservação de $\phi 3Tim$. (A) PCR de fragmento interno ao gene *spec* e (B) PCR de fragmento interno ao gene $\phi 3Tim$. 1 e 2: Cepas recombinadas com o vetor 705-Cre. 3: Controle positivo – cepa contendo o gene *spec*, 4: Controle da reação.

Os resultados obtidos mostraram que a marca de seleção para espectinomicina foi removida com sucesso, assim como o gene $\phi 3Tim$ foi

mantido. Este processo foi importante para garantir a aplicação ainda mais ampla deste sistema, possibilitando, por exemplo, a sua utilização para a metilação de vetores de expressão contendo a mesma marca de seleção para espectinomicina.

5.7. Análise da funcionalidade de $\phi 3\text{Tim}$ -MTP integrada em *E. coli*

5.7.1. Análise *in vitro*

Para determinar se a metiltransferase $\phi 3\text{Ti}$ -MTP integrada no genoma de *E. coli* estava funcional foi realizada uma primeira investigação *in vitro* com a utilização da enzima *Sat*I (Fnu4HI), um isosquizômero da endonuclease de restrição do sistema RM de *C. acetobutylicum*, Cac824I. Três clones de *E. coli* contendo o gene $\phi 3\text{Tim}$, foram então transformadas com o vetor pACYC184 juntamente com uma cepa de *E. coli* DH10 selvagem, utilizada como controle. Os transformantes foram selecionados em cloranfenicol e o DNA plasmidial foi extraído. O plasmídeo obtido em cada uma das 4 linhagens em questão foi submetido à digestão com a enzima *Sat*I. Era esperado que o plasmídeo pACYC184 fosse completamente digerido pela enzima *Sat*I caso este não estivesse devidamente metilado impedindo a ação da enzima.

A figura 25 mostra a digestão do plasmídeo pACYC184 extraído de linhagens de *E. coli* mutadas e selvagem quando digerido com a enzima *Sat*I.

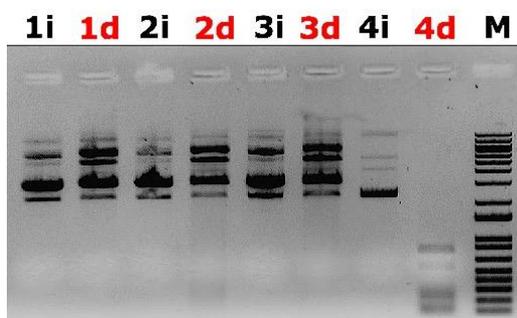


Figura 25. Análise da digestão de PACYC184 com *Sat*I. Eletroforese em gel de agarose 0,8% corado com brometo de etídeo. M:1 kb Plus DNA Ladder (*Thermo Scientific*). 1, 2 e 3: foi utilizado o DNA plasmidial extraído de três cepas de *E. coli* contendo o gene $\phi 3\text{Tim}$ integrado

no genoma, 4: foi usado o DNA plasmidial extraído de *E. coli* DH10 não mutada. Os produtos de miniprep intactos estão indicados em preto e os produtos digeridos em vermelho.

A figura 26 mostra os sítios de restrição para a enzima *SatI* presentes no vetor pACYC184. Pode-se observar que este plasmídeo seria digerido em pequenos fragmentos caso não estivesse protegido contra a ação de *SatI*, como foi observado para a digestão da amostra número 4 da figura 24.

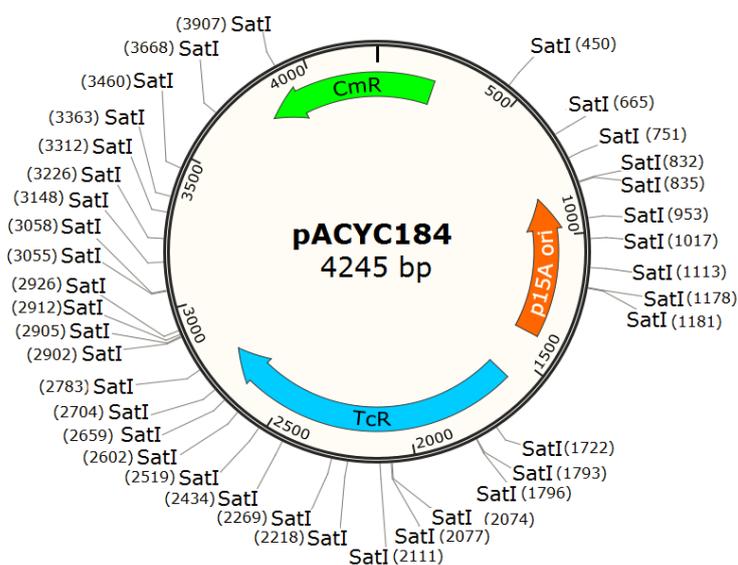


Figura 26. Mapa físico do vetor pACYC184. São mostrados os 37 sítios de restrição para a enzima *SatI*.

Os resultados deste experimento mostraram que o gene $\phi 3Tim$ presente no genoma de *E. coli* consegue proteger o material plasmidial introduzido nessas células contra a ação da enzima *SatI*, uma vez que o vetor pACYC184 não sofreu digestão quando transformado em linhagens de *E. coli* modificadas geneticamente, diferentemente do mesmo vetor transformado em *E. coli* DH10B selvagem, que foi completamente digerido.

A partir desses resultados, resolveu-se fazer um teste de funcionalidade *in vivo* utilizando células de *C. acetobutylicum* para a transformação com vetores previamente introduzidos nas cepas mutantes.

5.7.2. Análise *in vivo*

Uma vez que os testes *in vitro* indicaram que houve proteção plasmidial contra a ação de *SatI*, foi realizada a análise *in vivo* para demonstrar a eficácia deste sistema contra a ação do sistema de *C. acetobutylicum*.

Para a análise *in vivo*, cepas de *E. coli* DH10B modificadas geneticamente foram transformadas com o vetor de expressão em *C. acetobutylicum*, pMTL-mazF-PTA. Como controle, uma cepa de *E. coli* DH10B selvagem também foi transformada com o mesmo plasmídeo. O DNA plasmidial de transformantes foi extraído e utilizado para transformar células de *C. acetobutylicum*.

A figura 27 mostra os diversos sítios de restrição para a enzima *Cac824I* do sistema de restrição de *C. acetobutylicum* presentes no plasmídeo pMTL-mazF-PTA. A figura 28 mostra o resultado do crescimento em meio líquido adicionado de tianfenicol, marca de seleção para o plasmídeo pMTL-mazF-PTA.

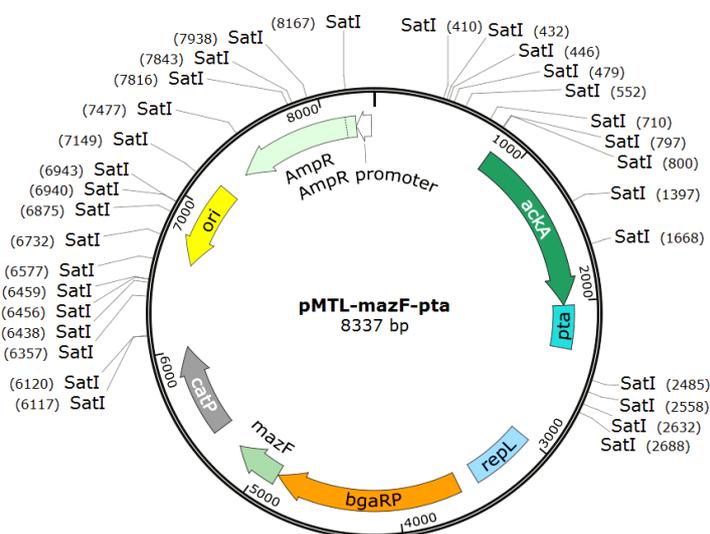


Figura 27. Mapa físico do vetor pMTL-mazF-PTA. São mostrados os 31 sítios de restrição para a enzima *Cac824I*.

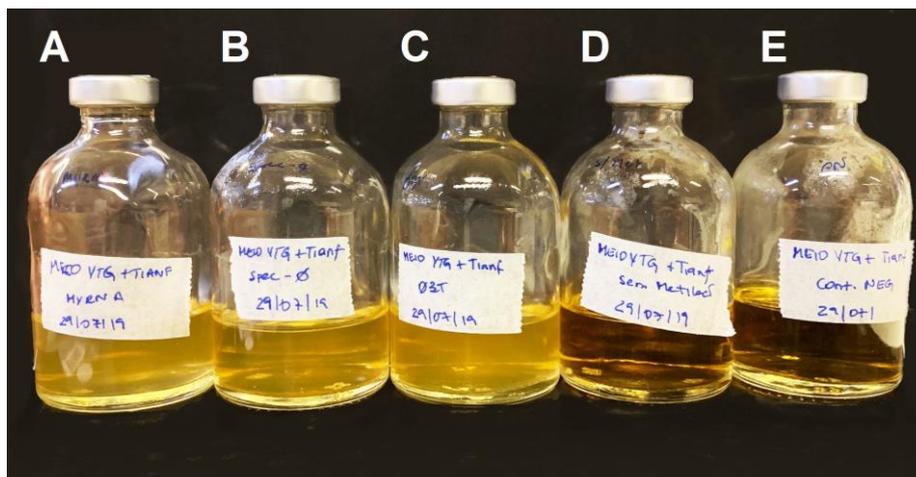


Figura 28. Resultado da transformação de *C. acetobutylicum*. (A) transformação com o vetor pMTL-mazF-PTA extraído de cepas de *E. coli* expressando o gene $\phi 3Tim$ ligado ao plasmídeo pACYC (Controle positivo) – produzido por Gomes, 2018, (B) transformação com o vetor pMTL-mazF-PTA extraído de cepa de *E. coli* modificadas com o fragmento spec- $\phi 3Tim$, (C) transformação com o vetor pMTL-mazF-PTA extraído de cepa de *E. coli* modificadas com o gene $\phi 3Tim$ e recombinadas com 705-Cre para remoção do gene spec, (D) transformação com o vetor pMTL-mazF-PTA extraído de cepa de *E. coli* DH10 selvagem (controle negativo) e (E) cepa de *C. acetobutylicum* eletroporada sem o vetor pMTL-mazF-PTA (controle do experimento).

De acordo com a figura 28, foi indicado que cepas de *C. acetobutylicum* foram eficientemente transformadas com pMTL-mazF-PTA quando este vetor foi previamente submetido a cepas de *E. coli* contendo o gene $\phi 3Tim$, culturas representadas na figura 28 por A, B e C. A resistência ao tianfenicol, determinada pela presença do plasmídeo pMTL-mazF-PTA, foi observada somente nas transformações de *C. acetobutylicum* com este vetor extraído de cepas de *E. coli* modificadas, sugerindo que o plasmídeo foi devidamente metilado e protegido pela metiltransferase $\phi 3Ti$ -MTP. Já os inóculos com células de *C. acetobutylicum* transformadas com pMTL-mazF-PTA extraído de cepas selvagens de *E. coli* (representado em D) ou células de *C. acetobutylicum* não transformadas com pMTL-mazF-PTA (representado em E) não obtiveram crescimento considerável, demonstrando que *C. acetobutylicum* é sensível às concentrações de tianfenicol utilizadas e que o plasmídeo pMTL-mazF-PTA foi perdido pelas células de *C. acetobutylicum* quando este não havia sido submetido à cepas de *E. coli* expressando $\phi 3Ti$ -MTP.

Estes resultados *in vivo* demonstraram que o sistema de metilação aqui desenvolvido é capaz de proteger vetores de expressão para a transformação de *C. acetobutylicum*. Assim, o objetivo mais importante deste trabalho foi realizado com sucesso, que foi a construção de uma linhagem de *E. coli* capaz de metilar vetores *in vivo*, permitindo então a transformação de clostrídios com estes vetores previamente metilados.

O estabelecimento de uma plataforma universal de metilação para *C. acetobutylicum* abrirá portas para o estudo de outros sistemas de restrição. Várias espécies solventogênicas vêm sendo utilizadas para a produção industrial de produtos químicos de interesse, como acetona, butanol e etanol. A falta de um sistema de manipulação genética para estes micro-organismos, como por exemplo, *C. saccharobutylicum*, ainda limita a engenharia da via metabólica para melhorar o rendimento da produção de solventes assim como dificulta os estudos genômicos e fisiológicos dessas bactérias. Criar uma plataforma para a deleção, inserção ou modificação rápida de genes será essencial para a utilização comercial desses micro-organismos. (Huang, Liebl e Ehrenreich, 2018)

6. CONCLUSÕES E PERSPECTIVAS

Neste trabalho foi construída uma linhagem de *E. coli* contendo o gene $\phi 3Tim$, que codifica para a metiltransferase $\phi 3Ti$ -MTP, integrado ao seu genoma.

Foi mostrado que, neste sistema de expressão heteróloga, $\phi 3Ti$ -MTP é capaz de ser expressa para metilar DNA plasmidial em células de *E. coli*, demonstrando que estes clones podem ser um eficiente sistema de proteção de vetores contra o sistema de restrição de *C. acetobutylicum*.

O método de aceitação de DNA exógeno em *C. acetobutylicum* estabelecido neste trabalho tem o intuito de possibilitar o estudo mais aprofundado das ferramentas moleculares desses indivíduos, visando traçar estratégias de melhoramentos genéticos para aperfeiçoar a fermentação.

A funcionalidade *in vivo* foi confirmada, certificando a aplicabilidade deste sistema como uma plataforma para a transformação de *C. acetobutylicum*. Com isso, este estudo compreendeu mais uma ferramenta inovadora de edição de *C. acetobutylicum*

O sistema de metilação aqui construído poderá obter uma patente, estabelecida pela cepa resultante de todo o processo de recombinação, linhagem esta denominada *Escherichia coli* DH10B_ $\phi 3Tim$.

O trabalho desenvolvido para a manipulação de *C. acetobutylicum* poderá ser futuramente estendido para outros micro-organismos, incluindo outros clostrídios fermentadores de interesse econômico que apresentam sistemas de restrição próprios ainda pouco estudados, como *C. saccharobutylicum*.

7. REFERÊNCIAS

ALSAKER, K. V.; SPITZER, T. R.; PAPOUTSAKIS, E. T. Transcriptional Analysis of *spo0A* Overexpression in *Clostridium acetobutylicum* and Its Effect on the Cell's Response to Butanol Stress. **Journal of Bacteriology**. v. 186, n. 7, p. 1959–1971, 2004.

AWANG, G. M.; JONES, G. A.; INGLEDEW, W. M. The acetone-butanol-ethanol fermentation. **Critical Reviews in Microbiology**. v. 15, 1988.

BANEYX, F. Recombinant protein expression in *Escherichia coli*. **Expression vectors and delivery systems**, 1999.

BELL, J. C.; KOWALCZYKOWSKI, S. C. RecA: Regulation and Mechanism of a Molecular Search Engine. **Trends in Biochemical Sciences**, v. 41, n. 6, p. 491–507, 2016.

BERNDT, C.; MEIER, P.; WACKERNAGEL, W. DNA restriction is a barrier to natural transformation in *Pseudomonas stutzeri* JM300 Printed in Great Britain. **Microbiology**. p. 895–901, 2003.

BLATTNER, F. R. *et al.* The Complete Genome Sequence of *Escherichia coli* K-12. **Science**, v. 277, n. 5331, p. 1453–1462, 1997.

BOCIOAGA, D. *et al.* Development of a genetic system for a model manganese-oxidizing proteobacterium, *Leptothrix discophora* SS1. **Microbiology (United Kingdom)**, v. 160, n. 2014, p. 2396–2405, 2014.

CHAYOT, R. *et al.* An end-joining repair mechanism in *Escherichia coli*. **Proceedings of the National Academy of Sciences of the United States of America**, v. 107, n. 5, p. 2141–2146, 2010.

CHO, S. W. *et al.* Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. **Cold Spring Harbor Laboratory Press Method**, p. 132–141, 2014.

CHU, V. T. *et al.* CRISPR-Cas systems for genome editing, regulation and targeting. **Nature**, v. 33, n. 4, p. 347–355, 2014.

CRONAN, J. E. *Escherichia coli* as an Experimental Organism . **Encyclopedia of Life Sciences**, 2006.

CROUX, C. *et al.* Biotechnology for Biofuels Construction of a restriction - less , marker - less mutant useful for functional genomic and metabolic engineering of the biofuel producer *Clostridium acetobutylicum*. **Biotechnology for Biofuels**, p. 1–13, 2016.

CUI, L.; BIKARD, D. Consequences of Cas9 cleavage in the chromosome of *Escherichia coli*. **Nucleic Acids Research**. v. 44, n. 9, p. 4243–4251, 2016.

DAVY, A. M.; KILDEGAARD, H. F.; ANDERSEN, M. R. Cell Factory Engineering. **Cell Systems**, v. 4, n. 3, p. 262–275, 2017.

DEMAIN, A. L.; VAISHNAV, P. Production of Recombinant Proteins by Microbes and Higher Organisms. **Comprehensive Biotechnology, Second Edition**, v. 3, n. 3, p. 333–345, 2011.

DONG, H. *et al.* Engineering *Clostridium* Strain to Accept Unmethylated. **PLoS One**, v. 5, n. 2, p. 1–8, 2010.

DURFEE, T. *et al.* The complete genome sequence of *Escherichia coli* DH10B: Insights into the biology of a laboratory workhorse. **Journal of Bacteriology**, v. 190, n. 7, p. 2597–2606, 2008.

DUTRA, B. E.; SUTERA, V. A.; LOVETT, S. T. RecA-independent recombination is efficient but limited by exonucleases. **Proceedings of the National Academy of Sciences**, v. 104, n. 1, p. 216–221, 2007.

EBERHART, L. J.; KNUTSON, C. M.; BARNEY, B. M. A methodology for markerless genetic modifications in *Azotobacter vinelandii*. **Applied Microbiology**, p. 1595–1604, 2016.

FU, Y. *et al.* High frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. **Nature Biotechnology**, v. 31, n. 9, p. 822–826, 2013.

GHESHLAGHI, R. *et al.* Metabolic pathways of clostridia for producing butanol. **Biotechnology Advances**, v. 27, n. 6, p. 764–781, 2009.

- GILEADI, O. Recombinant Protein Expression in *E. coli*: A Historical Perspective. **Heterologous Gene Expression in *E.coli***, v. 705, p. 3–10, 2017.
- GOMES, M. B. Modificações genéticas em *Clostridium acetobutylicum* visando o aumento da produção de álcoois. 2018.
- GREEN, E. M. Fermentative production of butanol — the industrial perspective. **Current opinion in Biotechnology**, v. 22, p. 337 - 343, 2011.
- HADFIELD, C.; CASHMORE, A. M.; MEACOCK, P. A. Sequence and expression characteristics of a shuttle chloramphenicol-resistance marker for *Saccharomyces cerevisiae* and *Escherichia coli*. **Gene**, v. 52, n. 1, p. 59–70, 1987.
- HUANG, C. N.; LIEBL, W.; EHRENREICH, A. Restriction-deficient mutants and marker-less genomic modification for metabolic engineering of the solvent producer *Clostridium saccharobutylicum*. **Biotechnology for Biofuels**, v. 11, n. 1, p. 1–13, 2018.
- JACOBUS, A. P.; GROSS, J. Optimal cloning of PCR fragments by homologous recombination in *Escherichia coli*. **PLoS ONE**, v. 10, n. 3, p. 1–17, 2015.
- JIANG, W. *et al.* RNA-guided editing of bacterial genomes using CRISPR-Cas systems. **Nature Biotechnology**, v. 31, n. 3, p. 233–239, 2013.
- JIANG, Y. *et al.* Multigene Editing in the *Escherichia coli* Genome via the CRISPR-Cas9 System. **Applied and Environmental Microbiology**, v. 81, n. 7, p. 2506–2514, 2015.
- JINEK, M. *et al.* A Programmable Dual-RNA – Guided. **Science**, v. 337, p. 816–822, 2012.
- KAROUI, M. EL *et al.* Gene replacement with linear DNA in electroporated wild-type *Escherichia coli*. **Nucleic acids research**, v. 27, n. 5, p. 1296, 1999.
- KAUR, JASHANDEEP; KUMAR, A.; KAUR, JAGDEEP. Strategies for optimization of heterologous protein expression in *E. coli*: Roadblocks and reinforcements. **International Journal of Biological Macromolecules**, v. 106, p. 803–822, 2018.

KEHRENBERG, C. *et al.* Novel spectinomycin/streptomycin resistance gene, *aadA14*, from *Pasteurella multocida*. **Antimicrobial Agents and Chemotherapy**, v. 49, n. 7, p. 3046–3049, 2005.

KHARKWAL, S. *et al.* Strain Improvement and Process Development for Biobutanol Production. **Recent Patents on Biotechnology**, p. 202–210, 2009.

KNOSHAUG, E. P.; ZHANG, M. Butanol Tolerance in a Selection of Microorganisms. **Applied Biochemistry and Biotechnology**, v 153, p. 13–20, 2009.

L. MERMELSTEIN, N. WELKER, G. BENNETT, E. P. . _Expression of cloned homologous fermentative gene in *C. acetobutylicum* ATCC 824. **Nature Biotechnology**, v. 10, p. 190 – 195, 1992.

LARSON, M. *et al.* {CRISPR} interference {(CRISPRi)} for sequence-specific control of gene expression. **Nature Protocols**, v.11, p. 2180-2196, 2013.

LARSON, M. H. *et al.* Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression. **Cell**, v.152, p. 1173 – 1183, 2013.

LEE, J. Y. *et al.* Metabolic engineering of *Clostridium acetobutylicum* M5 for highly selective butanol production, **Biotechnology Journal**, p. 1432–1440, 2009.

LI, Y. Z. L. *et al.* Problems with the microbial production of butanol. **Journal of Industrial Microbiology and Biotechnology**, p. 1127–1138, 2009.

LINO, C. A. *et al.* Delivering crispr: A review of the challenges and approaches. **Drug Delivery**, v. 25, n. 1, p. 1234–1257, 2018.

LOVETT, S. T. *et al.* Crossing Over Between Regions of Limited Homology in *Escherichia coli*: RecA-Dependent and RecA-Independent Pathways. **Genetics**, v. 160, p. 851–859, 2002.

LUKJANCENKO, O.; WASSENAAR, T. M.; USSERY, D. W. Comparison of 61 Sequenced *Escherichia coli* Genomes. **Microbial Ecology**, v. 60, n. 4, p. 708–720, 2010.

LUO, M. L.; LEENAY, R. T.; BEISEL, C. L. Current and future prospects for CRISPR-based tools in bacteria. **Biotechnology and Bioengineering**, v. 113, n. 5, p. 930–943, 2016

M. NOYER-WEIDNER, S. JENTSCH, J. KUPSCH, M. BERGBAUER, T. A. T. DNA methyltransferase genes of *Bacillus subtilis* phages : structural relatedness and gene expression. **Gene**, v. 35, p. 143 – 150, 1985.

MARRAFFINI, L. A. CRISPR-Cas immunity in prokaryotes. **Nature**, v. 526, n. 7571, p. 55–61, 2015.

MENG, X. *et al.* Counter-selectable marker for bacterial-based interaction trap systems. **BioTechniques**, v. 40, n.2, p. 179 – 184, 2006.

MOUGIAKOS, I. *et al.* Efficient Genome Editing of a Facultative Thermophile Using Mesophilic spCas9. **ACS Synthetic Biology**, v. 6, n. 5, p. 849–861, 2017.

NIKOLAJEWA, S. *et al.* Common patterns in type II restriction enzyme binding sites. **Nucleic Acids Research**, v. 33, n. 8, p. 2726–2733, 2005.

NOVICK, R. P. Plasmid incompatibility. **Microbiological reviews**, v. 51, n. 4, p. 381–95, 1987.

NOWAK, C. M. *et al.* Guide RNA engineering for versatile Cas9 functionality. **Nucleic Acids Research**, v. 44, n. 20, p. 9555–9564, 2016.

OLINS, P. O.; LEE, S. C. Recent advances in heterologous gene expression in *Escherichia coli*. **Current Opinion in Biotechnology**, p. 520–525, 1993.

PAPOUTSAKIS, E. T.; MERMELSTEIN, L. E. E. D. *In Vivo* Methylation in *Escherichia coli* by the *Bacillus subtilis* Phage + 3T I Methyltransferase to protect plasmids from restriction upon transformation of *Clostridium acetobutylicum* ATCC 824. **Applied and Environmental Microbiology**, v. 59, n. 4, p. 1077–1081, 1993.

PINES, G. *et al.* Codon compression algorithms for saturation mutagenesis. **ACS Synthetic Biology**, v. 4, n. 5, p. 604–614, 2015.

PYNE, M. E. *et al.* Coupling the CRISPR / Cas9 System with Lambda Red Recombineering Enables Simplified Chromosomal Gene Replacement in *Escherichia coli*. **Applied and Environmental Microbiology**, v. 81, n. 15, p. 5103–5114, 2015.

REINERS, L.; KEINERS, L. Multispecific DNA methyltransferases from *Bacillus subtilis* phages properties of wild-type and various mutant enzymes with altered DNA affinity. **European Journal of Biochemistry**, v. 492, p. 485–492, 1986.

ROSANO, G. L.; CECCARELLI, E. A. Recombinant protein expression in *Escherichia coli*: Advances and challenges. **Frontiers in Microbiology**, v. 5, p. 1–17, 2014.

SAMBROOK, J.; FRITSCH, E. F.; MANIATIS, T. Molecular cloning: a laboratory manual, 2^a ed. New York, 1989.

SCHUMANN, W.; FERREIRA, L. C. S. Production of recombinant proteins in *Escherichia coli*. **Genetics and Molecular Biology**, v. 27, n. 3, p. 442–453, 2004.

SCHWARTZ, C. *et al.* Validating genome-wide CRISPR-Cas9 function in the non-conventional yeast *Yarrowia lipolytica*. **bioRxiv**, p. 358630, 2018.

SHAW, W. V. Chloramphenicol Acetyltransferase from Chloramphenicol-Resistant Bacteria. **Methods in Enzymology**, v. 43, n. C, p. 737–755, 1975.

SHEE, C. *et al.* What Limits the Efficiency of Double-Strand in *Escherichia coli*? **Journal of Molecular Microbiology and Biotechnology**. v. 3411, p. 8–19, 2012.

SONG, F.; STIEGER, K. Optimizing the DNA Donor Template for Homology-Directed Repair of Double-Strand Breaks. **Molecular Therapy - Nucleic Acids**, v. 7, p. 53–60, 2017.

STERN, A. L. *et al.* Structural mechanism of AadA, a dual-specificity aminoglycoside adenylyltransferase from *Salmonella enterica*. **Journal of Biological Chemistry**, v. 293, n. 29, p. 11481–11490, 2018.

SUZUKI, H.; MURAKAMI, A.; YOSHIDA, K. Counterselection System for

Geobacillus kaustophilus HTA426 through Disruption of pyrF and pyrR. **Applied and Environmental Microbiology**, v. 78, n. 20, p. 7376–7383, 2012.

TERPE, K. Overview of bacterial expression systems for heterologous protein production: From molecular and biochemical fundamentals to commercial systems. **Applied Microbiology and Biotechnology**, v. 72, n. 2, p. 211–222, 2006.

VASU, K.; NAGARAJA, V. Diverse Functions of Restriction-Modification Systems in Addition to Cellular Defense. **Microbiology and Molecular Biology Reviews**, v. 77, n. 1, p. 53–72, 2013.

WANG, J. *et al.* An Improved Recombineering Approach by Adding RecA to λ Red Recombination. **Molecular Biotechnology**, v. 32, p. 43 – 53, 2006.

WEBER, C.; FARWICK, A.; BENISCH, F. Trends and challenges in the microbial production of lignocellulosic bioalcohol fuels. **Applied Microbiology and Biotechnology**, p. 1303–1315, 2010.

WONG, N.; LIU, W.; WANG, X. WU-CRISPR: Characteristics of functional guide RNAs for the CRISPR/Cas9 system. **Genome Biology**, v. 16, n. 1, p. 1–8, 2015.

WOODS, D. R. The genetic engineering of microbial solvent production. **Trends in Biotechnology**, v. 13, n. 7, p. 259–264, 1995.

XU, X.; DUAN, D.; CHEN, S. J. CRISPR-Cas9 cleavage efficiency correlates strongly with targetsgRNA folding stability: From physical mechanism to off-target assessment. **Scientific Reports**, v. 7, n. 1, p. 1–9, 2017.

XUE, C. *et al.* CRISPR interference and priming varies with individual spacer sequences. **Nucleic Acids Research**, v. 43, n. 22, p. 10831–10847, 2015.

Y. LI, Z. LIN, C. HUANG, Y. ZHANG, Z. WANG, Y. T. T. CHEN, X. Z. Metabolic engineering of *Escherichia coli* using CRISPR-Cas9 mediated genome editing. **Metabolic Engineering**, v. 31, p. 13 – 21, 2015.

YANG, Y. *et al.* Metabolic regulation in solventogenic clostridia: regulators, mechanisms and engineering. **Biotechnology Advances**, v. 36, n. 4, p. 905 –

914, 2018.

ZHAO, D. *et al.* Development of a fast and easy method for *Escherichia coli* genome editing with CRISPR. **Microbial Cell Factories**, p. 1–9, 2016.