
PROGRAMA INTEGRADO (UNESP, USP E UNICAMP) DE PÓS-GRADUAÇÃO
EM BIOENERGIA

RECONSTRUÇÃO EM *ESCHERICHIA COLI* DE ALELOS IDENTIFICADOS POR
EVOLUÇÃO ADAPTATIVA EM SACAROSE USANDO O SISTEMA EASYGUIDE
CRISPR

JONECLEI ALVES BARRETO

Tese apresentada ao Instituto de Pesquisa em Bioenergia de Rio Claro, Universidade Estadual Paulista, como parte dos requisitos para obtenção do título de Doutor em Ciências, área de Bioenergia.

Orientador: Dr. Jeferson Gross

Coorientador: Dr. Jonas Contiero

Rio Claro – SP
2025

**PROGRAMA INTEGRADO (UNESP, USP E UNICAMP) DE PÓS-GRADUAÇÃO
EM BIOENERGIA**

**RECONSTRUÇÃO EM *ESCHERICHIA COLI* DE ALELOS IDENTIFICADOS POR
EVOLUÇÃO ADAPTATIVA EM SACAROSE USANDO O SISTEMA EASYGUIDE
CRISPR**

JONECLEI ALVES BARRETO

Tese apresentada ao Instituto de Pesquisa em Bioenergia, Universidade Estadual Paulista (UNESP), Programa de Pós-Graduação em Bioenergia, USP/UNICAMP/UNESP, como parte dos requisitos para obtenção do título de Doutor em Ciências, área de Bioenergia.

Orientador: Dr. Jeferson Gross

Coorientador: Prof. Jonas Contiero

**Rio Claro – SP
2025**

B273r Barreto, Joneclei Alves
Reconstrução em Escherichia coli de alelos identificados por
evolução adaptativa em sacarose usando o sistema EasyGuide
CRISPR / Joneclei Alves Barreto. -- Rio Claro, 2025
191 p. : il., tabs., fotos

Tese (doutorado) - Universidade Estadual Paulista (UNESP),
Instituto de Pesquisa em Bioenergia, Rio Claro
Orientador: Jeferson Gross
Coorientador: Jonas Contiero

1. CRISPR/Cas9. 2. Evolução adaptativa em laboratório. 3.
Escherichia coli. 4. Sacarose. 5. Engenharia metabólica. I. Título.

CERTIFICADO DE APROVAÇÃO

TÍTULO DA TESE: RECONSTRUÇÃO EM ESCHERICHIA COLI DE ALELOS IDENTIFICADOS POR EVOLUÇÃO ADAPTATIVA EM SACAROSE USANDO O SISTEMA EASYGUIDE CRISPR

AUTOR: JONECLEI ALVES BARRETO

ORIENTADOR: JEFERSON GROSS

COORIENTADOR: JONAS CONTIERO

Aprovado como parte das exigências para obtenção do Título de Doutor em Ciências, área: Bioenergia pela Comissão Examinadora:

Documento assinado digitalmente
 **JEFERSON GROSS**
Data: 31/01/2025 12:54:35-0300
Verifique em <https://validar.iti.gov.br>

Pesquisador JEFERSON GROSS (Participação Virtual)
Instituto de Pesquisa em Bioenergia IPBEN / Instituto de Pesquisa em Bioenergia IPBEN

Documento assinado digitalmente
 **JOSE GREGORIO CABRERA GOMEZ**
Data: 04/02/2025 18:06:27-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. JOSÉ GREGÓRIO CABRERA GOMEZ (Participação Virtual)
Departamento de Microbiologia / Instituto de Ciências Biomédicas da Universidade de São Paulo (USP)

Documento assinado digitalmente
 **DANIELLE BISCARO PEDROLI**
Data: 06/02/2025 12:11:38-0300
Verifique em <https://validar.iti.gov.br>

Profa. Dra. DANIELLE BISCARO PEDROLI (Participação Virtual)
Departamento de Engenharia de Bioprocessos e Biotecnologia / Faculdade de Ciências Farmacêuticas do Campus de Araraquara da Unesp

Documento assinado digitalmente
 **LEANDRO VIEIRA DOS SANTOS**
Data: 31/01/2025 15:12:32-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. LEANDRO VIEIRA DOS SANTOS (Participação Virtual)
Chemical Biology and Biological Chemistry / Instituição: Universidade de Manchester, UK

Rio Claro, 31 de janeiro de 2025

AGRADECIMENTOS

Primeiramente gostaria de agradecer à bolsa recebida processo nº 2020/02246-0, Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) e ao apoio financeiro do projeto temático 2017/50249-6, Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), e à UNESP e ao IPBEN e pela disponibilização do espaço físico para a execução desse projeto.

Ao Dr. Jeferson Gross principalmente pela oportunidade concedida e pela orientação desse trabalho, além de ser o idealizador de estratégias fundamentais para execução do mesmo. Além de aprendizados dentro e fora do laboratório. Sendo que, ao final desse trabalho, tornou-se mais que um orientador, um grande amigo.

À Dra Ana Paula Jacobus por me acolher no laboratório na minha chegada e pelo apoio, principalmente no aprendizado de técnicas que pouco conhecia e hoje tenho domínio, além do apoio moral e emocional nesse período.

Ao meu colega de laboratório e apartamento Lucas de Bem e seus gatos (Mazda, Geraldo, Melanina e Amálgama), que após esse período tornou-se um grande amigo.

Ao professor Dr. Jonas Contiero pela oportunidade de trabalharmos juntos e por todo o apoio no período.

Ao Dr. Michel Brienzo e a Daniele Marin pelo auxílio nas quantificações de sacarose.

Ao Matheus Lacôrte pelas quantificações e produção dos filmes de PHB.

Aos meus colegas de laboratório Yasmine, Isabelle, Fernanda, Sara, Stella e Dandara que mesmo que, com alguns, passamos pouco tempo juntos, mas foram momentos marcantes.

A Dona Maria, Edna, Roberto, Álvaro, Cleber, Luís e aos demais pessoas do corpo técnico e administrativo do IPBEN-Rio Claro pelo grande auxílio nesse período.

Aos meus colegas de laboratório do SENAI Dione, Juliana e Diego e que nesses últimos meses me suportaram e me auxiliaram nos momentos difíceis.

A Fernanda Andrade e ao Prof. Gregório pelo compartilhamento de cepas e plasmídeos usados.

E por último e não menos importantes gostaria de agradecer à minha família pelo apoio moral, mesmo tão distantes e com seus problemas não deixam de me ouvir nos momentos difíceis.

RESUMO

A evolução adaptativa em laboratório é uma poderosa ferramenta para o desenvolvimento de cepas de microrganismos que tenham capacidade de crescimento superior em fontes de carbono não tradicionais, de baixo custo e renováveis. Entretanto, para utilizar as informações obtidas, é necessário selecionar, reconstruir e testar as mutações encontradas durante a evolução na cepa progenitora, um processo conhecido como engenharia reversa. Para a reconstrução desses alelos mutados é imprescindível o uso de técnicas de modificação genômica de fácil execução e baixo custo. Nesse trabalho foram desenvolvidas técnicas de edição genômica usando CRISPR/Cas9 para serem realizadas para a reconstrução e teste de mutações observadas em experimento de evolução realizada em sacarose com populações de *Escherichia coli*. Nesse experimento, três populações de *E. coli* DH5 α contendo o plasmídeo pCscBKA para consumo de sacarose e quatro populações de *E. coli* E2348/69 foram cultivadas através de passagens sucessivas em meio mínimo contendo sacarose como única fonte de carbono, ao longo de 637 a 690 gerações. Ao final do processo, o DNA genômico das populações finais foi sequenciado e mutações encontradas foram selecionadas para reconstrução. Para a construção dessas linhagens mutadas foram desenvolvidos sistemas para clonagem de gRNAs e montagem de *donors* que foram chamados *EasyGuide*. Nesse contexto, foram projetados cinco plasmídeos base para as diferentes estratégias de clonagem de gRNAs. Para a produção de *donors* foi empregado um sistema para a montagem de fragmentos de DNA *in vivo* em *Saccharomyces cerevisiae*. Outra inovação implementada foi a utilização de oligonucleotídeos para montagem de gRNAs e *donors* para deleções e modificações de nucleotídeos únicos (abordagem *EasyOligo*). Com as técnicas desenvolvidas foi possível construir linhagens para testar mutações individuais e introduzir diferentes construções gênicas voltadas para o consumo de sacarose. Entre as mutações identificadas nas populações, a integração do operon *cscBKA* e a deleção do *cluster* formador do flagelo foram responsáveis pelos maiores ganhos de *fitness* (30-35% e 5%, respectivamente), medidos por ensaios de competição. Outra mutação que chamou a atenção foi a deleção de 81 pb no gene *pcnB*, associada a redução de número de cópias do plasmídeo pCscBKA e consequente aumento de *fitness* propagativo. Para demonstrar a capacidade do sistema *EasyGuide* CRISPR foram integrados ao genoma da DH5 α os operons *cscBKA* (consumo de sacarose) e *phaCAB*, para produção de polihidroxibutirato (PHB) a partir da sacarose. Com essas integrações, foi possível produzir ~31% massa celular seca de PHB a partir de 20 g/L sacarose.

Palavras-chave: CRISPR/Cas9; evolução adaptativa em laboratório; *Escherichia coli*; engenharia metabólica; sacarose; operon *cscBKA*.

ABSTRACT

Adaptive evolution is a powerful tool for developing strains of microorganisms that have superior growth capacity on non-traditional, low-cost, and renewable carbon sources. However, to make use of the obtained information, it is necessary to select, reconstruct, and test the mutations found during evolution in the progenitor strain, a process called reverse engineering. To reconstruct these mutated alleles, it is necessary to use easy-to-perform and low-cost genomic modification techniques. Genomic modification techniques using CRISPR/Cas9 were developed in this work for the reconstruction and testing of mutations obtained in an evolution experiment carried out in sucrose with *Escherichia coli*. In this experiment, three populations of *E. coli* DH5 α , carrying the plasmid pCscBKA for sucrose consumption, and four populations of *E. coli* E2348/69 were cultivated through successive passages in minimal medium containing sucrose as the sole carbon source for 637-690 generations. At the end of the process, the genomic DNA of the final populations was sequenced, and mutations found were selected for reconstruction. To obtain these mutated strains, gRNA cloning systems and donor assembly systems called EasyGuide were developed. Five base plasmids were developed for different gRNA cloning strategies. To produce donors, a system was used for the assembly of DNA fragments in vivo in *Saccharomyces cerevisiae*. Another advance was the use of oligonucleotides for the assembly of gRNAs and donors for deletions and modifications of single nucleotides (The EasyOligo approach). With these developed techniques, it was possible to construct strains for testing individual mutations and different genetic constructs for sucrose consumption. Among the mutations found in the populations, the integration of the *cscBKA* operon and the deletion of the flagellum assembly cluster were responsible for the highest fitness obtained (30-35% and 5%, respectively), measured by competition assays. Another mutation that attracted attention was the 81 bp deletion in the *pcnB* gene, which is associated to a reduction in the plasmid pCscBKA copy numbers and an increase in growth fitness. To demonstrate the capacity of the EasyGuide system, the *cscBKA* (sucrose consumption) and *phaCAB* operons were integrated into the genome of DH5 α cells to produce polyhydroxybutyrate (PHB) from sucrose. These integrations made it possible to produce ~31% of dry cell weight of PHB.

Keywords: CRISPR/Cas9; adaptive laboratory evolution; *Escherichia coli*; metabolic engineering; sucrose; operon *cscBKA*.

SUMÁRIO

1 INTRODUÇÃO.....	11
2 REVISÃO.....	13
2.1 <i>Escherichia coli</i>	13
2.2 Evolução adaptativa.....	14
2.2.1 Objetivos dos experimentos de evolução adaptativa com microrganismos	15
2.3 Engenharia reversa.....	16
2.4 CRISPR/Cas9.....	17
2.5 <i>In vivo cloning</i> em <i>E. coli</i>	23
2.6 Competições.....	24
2.7 Polihidroxialcanoatos	26
2.8 Sacarose	28
3 OBJETIVOS.....	30
4 MATERIAIS E MÉTODOS	31
4.1 Linhagens celulares, armazenamento, condições de crescimento e plasmídeos.....	31
4.2 Procedimentos de clonagem.....	31
4.3 Preparação das células bacterianas para transformações	32
4.4 Modificações genômicas por CRISPR/Cas9	33
4.5 Experimento de ALE	34
4.5.1 Meio mínimo suplementado com sacarose	34
4.5.2 Clonagem do <i>operon cscBKA</i>	35
4.5.3 Construção das linhagens parentais (progenitoras) para a ALE.....	36
4.5.4 Condução do experimento de ALE.....	37
4.5.5 Sequenciamento genômico das populações evoluídas.....	38
4.5.5 Ensaios de competição	39
4.6 Quantificação de sacarose.....	40
4.6 Quantificação de PHB.....	40
4.6.1 Produção dos filmes de PHB	42
5 RESULTADOS	43
5.1 Experimento de ALE em meio mínimo com sacarose	43
5.1.4 Avaliação das taxas de crescimento das populações finais do experimento de Evolução	43

5.2.1 Mutações encontradas nas populações evoluídas de DH5α	45
5.2.2 Mutações encontradas nas populações evoluídas de E2348/69	49
5.3 Estabelecimento de técnicas de genética molecular e edição genômica	50
5.3.1 Técnica de CRISPR/Cas9.....	50
5.3.2 Otimização da clonagem de gRNAs na técnica de CRISPR/Cas9	52
5.3.3 Método de clonagem em <i>Saccharomyces cerevisiae</i> de DNAs <i>donors</i> para uso em edições CRISPR/Cas9 em <i>E. coli</i>	57
5.3.4 Abordagem EasyOligo para a geração de <i>Donors</i>	60
5.4 Engenharia reversa das mutações encontradas no experimento de evolução	64
5.4.1 Avaliação de <i>fitness</i>	65
5.4.2 Construção do sistema de competição	65
5.4.4 Ensaios de competição e <i>fitness</i> das populações evoluídas	66
5.4.5 <i>Fitness</i> das linhagens com alelos evoluídos	67
5.4.6 Variações estruturais do operon para consumo da sacarose	70
5.4.7 Efeito do número de cópias de plasmídeos.....	71
5.4.8 Integração do operon <i>phaCAB</i>	73
6 DISCUSSÃO.....	75
7 CONCLUSÕES.....	85
REFERÊNCIAS	86
APÊNDICE A- Linhagens utilizadas no trabalho	98
APÊNDICE B – Plasmídeos usados neste trabalho	101
APÊNDICE C – Figuras detalhadas de plasmídeos	104
pTargetAmp.....	104
pSpec5'	105
pSpec3'	106
pOliSpec5'	107
pOliSpec3'	108
pCscBKA	109
APÊNDICE D – Mutações encontradas no experimento de Evolução	110
APÊNDICE E – Procedimentos de genética molecular	116
APÊNDICE F – Primers usados nas mutações e clonagens.....	133
ANEXO A – Publicações em periódicos especializados.....	146

1 INTRODUÇÃO

A *Escherichia coli* é uma bactéria gram-negativa que é considerada um organismo modelo por ter diversas características biológicas bem estudadas e estabelecidas, além de possuir o seu genoma caracterizado. Essa bactéria também apresenta grande importância para estudos que visam a construção de plataformas para a produção de diversos bioprodutos (MOHAMED; MUNDHADA; LANDBERG; CANN *et al.*, 2019).

A sacarose é um dissacarídeo formado pela ligação de uma molécula de glicose a uma de frutose. É um açúcar que, além de ser usado para a alimentação, pode ser empregado como fonte de carbono para produção industrial, tendo como principal destaque a produção do etanol. Principalmente quando derivada da cana-de-açúcar, a sacarose é tida como uma fonte de energia barata e renovável, também atrativa para a produção de outros bens além do etanol. Em adição a isso, a exploração desse carboidrato com destino a síntese de bioprodutos com maior valor agregado é interessante para nós, em virtude do Brasil ser um dos maiores produtores mundiais de açúcar da cana-de-açúcar (MOHAMED; MUNDHADA; LANDBERG; CANN *et al.*, 2019; OLAVARRIA; FINA; VELASCO; VAN LOOSDRECHT *et al.*, 2019).

Entretanto, somente algumas linhagens de *E. coli* metabolizam a sacarose, sendo que as laboratoriais ou industriais normalmente não utilizam o dissacarídeo (BRUSCHI; BOYES; SUGIARTO; NIELSEN *et al.*, 2012). Com isso, técnicas de engenharia genética são empregadas para transformar linhagens de interesse em cepas capazes de metabolizar esse açúcar. Mas, tanto linhagens de *E. coli* geneticamente modificadas, como cepas que naturalmente metabolizam a sacarose, não tem uma taxa alta de crescimento se comparada ao uso da glicose como fonte de carbono (MOHAMED; MUNDHADA; LANDBERG; CANN *et al.*, 2019). Com isso, são necessários aprimoramentos para que a bactéria possa utilizar a sacarose de forma eficiente.

Nesse sentido, a evolução adaptativa em laboratório (em inglês: *Adaptive Laboratory Evolution*, ALE) é uma poderosa ferramenta que pode ser usada para se entender as possibilidades de adaptação de um microrganismo quando desafiado a crescer em condições que não são as ideais. Desta forma, a ALE pode fornecer possibilidades para o desenvolvimento de cepas industriais capazes de ter a sua

produção otimizada (DRAGOSITS; MATTANOVICH, 2013). Assim, o uso da ALE para construir cepas de *E. coli* que apresentem um desenvolvimento adequado na sacarose é uma empolgante possibilidade para a obtenção de cepas superiores e plataformas industriais para bioprodutos interessantes comercialmente.

Para explorar as mutações observadas ao final de um experimento ALE é importante reconstruir e testar tais alterações nas cepas progenitoras do experimento (OUD; FLORES; GANCEDO; ZHANG *et al.*, 2012). Esse processo é chamado de engenharia reversa e é neste momento em que as mutações-chave são investigadas em relação ao *fitness*. As mutações que contribuem positivamente podem ser assim selecionadas e usadas para construção de cepas superiores (JACOBUS; CAVASSANA; DE; BARRETO *et al.*, 2024).

2 REVISÃO

2.1 *Escherichia coli*

Isolada inicialmente por Theodor Escherich a partir de intestino de bebês, *Escherichia coli* é uma bactéria gram-negativa e um dos organismos mais estudados, com fácil manipulação, rápido crescimento e suas características genéticas bem elucidadas (BLOUNT, 2015). Essa bactéria é conhecida pela capacidade de algumas cepas de serem patogênicas. Entretanto, trata-se de um microrganismo comensal humano, que coloniza o intestino poucas horas após o nascimento (KAPER; NATARO; MOBLEY, 2004; VILA; SAEZ-LOPEZ; JOHNSON; ROMLING *et al.*, 2016). No campo da biotecnologia, há um destaque especial para manipulação de DNA, como clonagem e manipulação de genes (RUIZ; SILHAVY, 2022). Além disso, é utilizada na produção de bioprodutos como biocombustíveis, bioplásticos, aminoácidos e vitaminas. Não menos importante é o uso de *E. coli* na produção de proteínas recombinantes, sendo o exemplo mais famoso a insulina (BAESHEN; BAESHEN; SHEIKH; BORA *et al.*, 2014; FANG; KANG; ZHANG, 2017; LIANG; LIU; FREED; ECKERT, 2020; RODRIGUEZ; MARTINEZ; FLORES; ESCALANTE *et al.*, 2014; ROSANO; CECCARELLI, 2014; WU; ZHOU; PEI; JIANG *et al.*, 2022). Apesar de não ser o organismo ideal para diversos bioprodutos, a compreensão básica para diversas vias biossintéticas passa inicialmente pela bactéria *E. coli*.

A primeira publicação relatando o sequenciamento do genoma completo dessa bactéria foi em 1997 por Blattner e colaboradores (BLATTNER; PLUNKETT; BLOCH; PERNA *et al.*, 1997). A cepa usada foi a K-12, com aproximadamente 4,6 Mb de extensão e com um total de 4.288 regiões codificadoras, das quais, na época, não se conhecia a função de 38% (BLATTNER; PLUNKETT; BLOCH; PERNA *et al.*, 1997). Desde então, com o avanço das tecnologias de sequenciamento, o genoma de diversas linhagens de *E. coli* foi sequenciado (ANTON; RALEIGH, 2016; ARCHER; KIM; JEONG; PARK *et al.*, 2011; DUNNE; CHAUDHURI; ROSSITER; BERIOTTO *et al.*, 2017; IGUCHI; THOMSON; OGURA; SAUNDERS *et al.*, 2009; MALDONADO; LOPEZ-HERNANDEZ; GARCIA-

MONTANER; LOPEZ-CORTES *et al.*, 2024; REID; BLAU; JECHALKE; SMALLA *et al.*, 2019).

Como mencionado, um dos papéis mais importantes na biotecnologia exercido por essa bactéria é na manipulação de material genético em laboratório. Praticamente toda manipulação de DNA em laboratórios envolve, em algum momento, o uso de *E. coli*, seja na clonagem ou na manutenção de plasmídeos (RUIZ; SILHAVY, 2022; ZHANG; MUYRERS; TESTA; STEWART, 2000). Cepas como DH5α, BL21(DE3), JM109, XL1-Blue e Top10 são amplamente utilizadas devido a características genéticas específicas que facilitam esses processos.

Como a bactéria possui condições de crescimento conhecidas, além de diversas linhagens poderem ser cultivadas em condições mais baixas de biossegurança, torna-se uma plataforma interessante para produção industrial de bioprodutos. Uma dessas possíveis bioprodutos são as proteínas e enzimas recombinantes (RIGGS, 2021).

2.2 Evolução adaptativa

A evolução adaptativa laboratorial é uma estratégia na qual o microrganismo é adaptado a condições de crescimento não tradicionais (DRAGOSITS; MATTANOVICH, 2013). Por meio desse processo, é possível selecionar fenótipos de crescimento superior além de isolar modificações genéticas essenciais para a adaptação às condições utilizadas (JACOBUS; CAVASSANA; DE; BARRETO *et al.*, 2024). Essa abordagem permite adaptar microrganismos para o consumo de fontes de carbono não usuais, como xilose para leveduras e sacarose e glicerol para *E. coli* (DOS SANTOS; CARAZZOLLE; NAGAMATSU; SAMPAIO *et al.*, 2016; KIM; HOU; CHOE; KANG *et al.*, 2022). Também é possível melhorar a capacidade de crescimento em meios com elevada concentração de inibidores de crescimento, como no uso de hidrolisados lignocelulósicos para crescimento de *Saccharomyces cerevisiae* (MENEGON; GROSS; JACOBUS, 2022; STOVICEK; DATO; ALMQVIST; SCHOPPING *et al.*, 2022).

No contexto de experimentos evolutivos com *Escherichia coli*, destaca-se o trabalho de Richard Lenski. Chamado de Experimento de Evolução de Longa Duração (LTEE, do inglês *Long-Term Evolution Experiment*), teve início em 1988 com doze

populações de *E. coli*, submetidas diariamente a transferências para um meio de cultura contendo glicose limitada como única fonte de carbono. Atualmente esse experimento já ultrapassou 75.000 gerações, permitindo a observação de adaptações específicas às condições de cultivo (LENSKI, 2017; 2023). Uma das adaptações mais marcantes foi a capacidade de consumo de citrato em condições aeróbicas em uma das populações, algo que a população progenitora e demais populações não são capazes de fazer nessas condições (BLOUNT; BORLAND; LENSKI, 2008).

2.2.1 Objetivos dos experimentos de evolução adaptativa com microrganismos

A realização de experimentos de evolução com foco específico no desenvolvimento de microrganismos superiores pode visar diferentes objetivos. Um deles é a adaptação a consumo de substratos não tradicionais. Nessa situação pode-se ter duas situações. Na primeira, o organismo já possui o arcabouço genético para consumo da fonte carbono, mas não a consome de forma eficiente (KIM; HOU; CHOE; KANG *et al.*, 2022). No caso de organismos que não possuem a capacidade de utilizar a fonte de carbono, é necessária a inserção de estruturas genéticas para catabolismo do novo nutriente (DOS SANTOS; CARAZZOLLE; NAGAMATSU; SAMPAIO *et al.*, 2016). Para ambas as situações visa-se alavancar mecanismos que melhorem a captação e/ou o metabolismo de consumo do substrato não tradicional, o que culmina na melhoria da capacidade de crescimento do microrganismo.

Outro foco está na adaptação de microrganismos a substratos de interesse que possuem certa toxicidade. Alguns meios de cultivo, ou substratos de fermentação alternativos, como os derivados de materiais lignocelulósicos, contém compostos inibidores que reduzem a capacidade de crescimento do microrganismo (LI; LIU; ZHAO, 2022; MORENO; GONZÁLEZ-FERNÁNDEZ; TOMÁS-PEJÓ, 2022). As populações submetidas a esse crescimento restritivo podem evoluir mecanismos de tolerância, ou até mesmo metabolizar os compostos tóxicos (ALMARIO; REYES; KAO, 2013). Situações semelhantes ocorrem quando a acumulação de um bioproduto produzido inibe o crescimento do microrganismo usado para a sua produção, como o biocombustível etanol produzido por leveduras (JACOBUS; CAVASSANA; DE; BARRETO *et al.*, 2024).

O organismo usado também pode ser exposto a concentrações elevadas do bioproduto para o desenvolvimento da tolerância (JANG; JU; JEONG; JUNG *et al.*, 2021).

Em se tratando do processo de realização da evolução das populações de microrganismos, o método de transferências usando Erlenmeyers, no qual a fração de um cultivo é passada para outro frasco com meio de cultivo fresco, é o mais tradicional, sendo usado, por exemplo, no LTEE de Richard Lenski (LENSKI; ROSE; SIMPSON; TADLER, 1991). Atualmente há diversos incrementos na realização desses processos, como o uso de sistemas robotizados, permitindo a evolução de diversas populações ao mesmo tempo, com passagens em momentos específicos do crescimento (MENEGON; GROSS; JACOBUS, 2022; MOHAMED; MUNDHADA; LANDBERG; CANN *et al.*, 2019).

2.3 Engenharia reversa

Durante ou após o processo de evolução adaptativa, o DNA genômico pode ser sequenciado com o objetivo de identificar mutações em comparação ao genoma do microrganismo antes do processo. Para o sequenciamento genômico pode-se realizar o sequenciamento do DNA de toda a população, ou ainda de colônias selecionadas (JACOBUS; CAVASSANA; DE; BARRETO *et al.*, 2024; MOHAMED; MUNDHADA; LANDBERG; CANN *et al.*, 2019).

Após a obtenção de dados de sequenciamento, é realizada a identificação das mutações ocorridas no processo de evolução. No caso do sequenciamento de populações evoluídas o primeiro critério para seleção de mutações é a frequência, o que implica que, se há uma mutação com alta presença na população, há a possibilidade daquele alelo mutado ter influência na evolução (DEATHERAGE; BARRICK, 2021; MOHAMED; MUNDHADA; LANDBERG; CANN *et al.*, 2019). Entretanto, tanto para a análise de populações como de colônias isoladas, é necessária uma revisão bibliográfica extensa para compreender o papel exercido por mutações na adaptação à situação em que o microrganismo foi exposto. Essa revisão pode ser usada para inicialmente montar cenários que expliquem a direção que o processo evolutivo tomou, selecionando assim mutações candidatas para o processo engenharia reversa. Outro critério a ser usado quando há populações independentes é o uso de mutações que possuem paralelismo

com outras. Se uma região mutada, sendo a parte codificante, região promotora ou terminadora de um gene, está mutada em mais de uma população, este tende a ser um alvo de grande interesse para investigação (MENEGON; GROSS; JACOBUS, 2022). Em certas situações, podem ocorrer, não necessariamente no mesmo gene, mas na mesma rota metabólica ou de sinal (JOHNSON; GOPALAKRISHNAN; GOYAL; DILLINGHAM *et al.*, 2021).

A engenharia reversa de mutações visa validar as mutações-chave responsáveis por incrementar a capacidade de crescimento na condição de evolução usada através da reconstrução de mutações individuais ou em associação no microrganismo progenitor (OUD; FLORES; GANCEDO; ZHANG *et al.*, 2012). Um fato interessante que deve ser pontuado é que a maioria das mutações que são encontradas em um experimento de evolução não são responsáveis por incrementar a capacidade de crescimento do microrganismo na condição analisada. Essas mutações são chamadas “caroneiras” (*Hitchhiking*), e a sua frequência é aumentada na população por co-ocorrência no genoma com mutações positivas que incrementam o *fitness* celular (BUSKIRK; PEACE; LANG, 2017).

2.4 CRISPR/Cas9

Para a reconstrução de mutações individuais é necessário o uso de técnicas de biologia molecular eficientes para permitir a reconstrução tanto de mutações pontuais quanto estruturais (WANG; LI; ZHANG; YIN *et al.*, 2023). Apesar de haverem diversas técnicas clássicas de modificações genômicas, a técnica de CRISPR/Cas9 é uma das mais robustas e que mais cumprem às necessidades para reconstrução de mutações (JACOBUS; CAVASSANA; DE; BARRETO *et al.*, 2024).

A técnica de CRISPR/Cas9 (do inglês *Clustered Regularly Interspaced Short Palindromic Repeats*) é uma das mais poderosas técnicas de modificação genômica atuais (LI; WANG; XU; WANG *et al.*, 2023; WEI; LI, 2023). Com ela já foi possível editar o genoma de bactérias, leveduras, plantas, animais e até células humanas (SHABBIR; SHABBIR; WU; MAHMOOD *et al.*, 2019; WEI; LI, 2023).

O sistema CRISPR foi identificado inicialmente como parte do sistema imunológico de bactérias que usam o sistema CRISPR-Cas para combater infecções virais, presente em cerca de 50% das bactérias e 90% das arqueas (WEI; LI, 2023). Para que funcione dessa forma são necessários três passos: adaptação, processamento e interferência. Na fase de adaptação é adquirido o DNA invasor (como DNA viral, por exemplo) (WEI; LI, 2023) (SHABBIR; SHABBIR; WU; MAHMOOD *et al.*, 2019). Esse material genético é incorporado ao arranjo do CRISPR como um novo *spacer* (SHABBIR; SHABBIR; WU; MAHMOOD *et al.*, 2019). No processamento, esse arranjo é expresso como um pre-crRNA (Pre-CRISPR RNA), o qual é maturado para o crRNA (CRISPR RNA). Na fase de interferência, o crRNA maturado liga-se à proteína efetora, que pode ser um complexo de proteínas Cas ou apenas uma unidade, direcionando para o alvo específico, realizando a clivagem do DNA ou RNA invasor (WEI; LI, 2023). Devido a essa especificidade, o sistema CRISPR/Cas foi cooptado como uma ferramenta biotecnológica formidável para modificação genômica de diversos organismos.

Para a realização da técnica, são necessários três componentes: a Cas9, o gRNA e o *donor*. A Cas9 mais usada é a de *Streptococcus pyogenes*. A enzima Cas9 (ou CRISPR-associated protein 9) é uma endonuclease que desempenha um papel central na técnica de CRISPR/Cas9, sendo a efetora na clivagem do genoma (BASSALO; GARST; HALWEG-EDWARDS; GRAU *et al.*, 2016). O local da clivagem é direcionado pelo gRNA (*guide* RNA), uma sequência de 20-nt homóloga com o genoma alvo. No genoma, essa sequência alvo precisa ser flanqueada pela sequência PAM (*protospacer adjacent motif*), a qual para o reconhecimento com a Cas9 de *S. pyogenes* é formada pela sequência NGG, onde N é qualquer nucleotídeo (A, T, C e G). A modificação genômica é especificada pelo *donor*, um fragmento de DNA desenhado pelo experimentalista para reparar o DNA clivado por recombinação homóloga (BASSALO; GARST; HALWEG-EDWARDS; GRAU *et al.*, 2016).

Diferentemente de outras espécies, para a realização da técnica de CRISPR/Cas9 em *E. coli* são necessárias algumas modificações específicas (BASSALO; GARST; HALWEG-EDWARDS; GRAU *et al.*, 2016). A primeira é em relação ao sistema de recombinação da bactéria, o qual não é tão eficiente. Sendo assim, há a necessidade de uso de um sistema de recombinação exógeno, no caso o Lambda Red (JIANG; CHEN;

DUAN; SUN *et al.*, 2015). Derivado do bacteriófago Lambda, esse sistema já é amplamente usado em inativação de genes em *E. coli* através da substituição por marcadores de resistência (DATSENKO; WANNER, 2000). O sistema Lambda Red é formado por três enzimas, Exo, Beta e Gam. A enzima Exo realiza uma atividade exonucleotídica da extremidade 5' para 3', transformando o DNA em fita simples. Essa extremidade com fita única pode ser degradada pela célula, e para que isso não ocorra a subunidade Beta a protege da degradação. A unidade Gam inibe as nucleases celulares RecBCD e SbcCD, que podem degradar o DNA para inserção (MURPHY, 1998). Com essas funcionalidades, o sistema Lambda Red, em associação com a clivagem do genoma realizada pela Cas9 guiada pelo gRNA, juntamente com o DNA *donor* para reparação do genoma clivado, torna possível a realização da técnica CRISPR em *E. coli*.

Outro gargalo encontrado para a realização da técnica em *E. coli* é relativo ao *donor* usado. A maioria dos trabalhos recomenda o uso de regiões de homologia grandes, de 200 pb a 1000 pb, sendo o ideal recomendado de 400 pb (BASSALO; GARST; HALWEG-EDWARDS; GRAU *et al.*, 2016; JIANG; CHEN; DUAN; SUN *et al.*, 2015). Apesar de alguns trabalhos relatarem o uso de regiões homólogas menores (SHUKAL; LIM; ZHANG; CHEN, 2022), o uso desses braços de homologia maiores são para aumentar a eficiência da modificação genômica. Assim, considerando a estrutura ideal, é necessário haver pré-clonagens para amplificação dos fragmentos para serem usados como *donor*, aumentando o valor e o tempo gasto na execução da técnica.

A clonagem do gRNA também pode ser destacada como outro gargalo. Isso porque, tratando-se principalmente de processos de engenharia reversa, nos quais há a necessidade de clonagem rápida de vários gRNAs em pouco tempo. As técnicas tradicionais de clonagem de gRNA usam kits de GoldenGate, Gibson Assembly ou métodos tradicionais com enzimas de restrição, elevando o custo de cada clonagem (LI; SUN; LV; ZUO *et al.*, 2023; PYNE; MOO-YOUNG; CHUNG; CHOU, 2015; ZERBINI; ZANELLA; FRACCASCIA; KONIG *et al.*, 2017).

Na Tabela 1 estão mostradas abordagens de CRISPR/Cas9 para *E. coli*.

Tabela 1: Principais abordagens CRISPR usando endonuclease Cas9 para edição de genoma em *E. coli*

Referência	Plasmídeo da Cas9 e clonagem do gRNA	Produção do <i>donor</i>
(JIANG; BIKARD; COX; ZHANG <i>et al.</i> , 2013)	Sistema de dois plasmídeos (RNA Crispr e Cas9). gRNA clonado via restrição e ligação.	<i>Donor</i> de oligonucleotídeos lineares.
(JIANG; CHEN; DUAN; SUN <i>et al.</i> , 2015)	pTargetF (gRNA) e pCas9 (Cas9). gRNA clonado por PCR inversa ou via restrição e ligação.	Donor de plasmídeo (HA 250/550 pb), clonado por PCR de sobreposição e restrição e ligação. <i>Donors</i> linear.
(LI; SUN; CHEN; ZHANG <i>et al.</i> , 2021)	Sistema de dois plasmídeos. Oligos complementares de gRNA ligados ao plasmídeo preparado com BsaI.	<i>Donor</i> linear
(LI; SUN; LV; ZUO <i>et al.</i> , 2023)	Dois plasmídeos. p15A-Cas9 (alto número de cópias). Oligos de gRNA ligados a um plasmídeo digerido por BsaI.	<i>Donor</i> linear.
(REISCH; PRATHER, 2015)	Sistema de dois plasmídeos (gRNA/ λ -Red e Cas9). gRNA clonado via CPEC.	<i>Donors</i> oligonucleotídeo: 60 – 90 nts ssDNA ou dsDNA oligos.
(PYNE; MOO-YOUNG; CHUNG; CHOU, 2015)	Sistema de três plasmídeos (gRNA, Cas9 e λ -Red). gRNA oligos ligados em plasmídeo preparado com BsaI.	60-nts ssDNA <i>donors</i> . <i>Donors</i> dsDNA a partir do <i>lacZ</i> amplificado por PCR. Primers fornecem 40 pb de homologia.

(BASSALO; GARST; HALWEG-EDWARDS; GRAU <i>et al.</i> , 2016)	Plasmídeo de gRNA montado por CPEC. Quatro plasmídeos para gRNA, Cas9, λ -Red, e <i>donor</i> .	<i>Donor</i> circular clonado com CPEC (50 – 600 bp de homologia). Integração de fragmento linear de 10 kb.
(ZERBINI; ZANELLA; FRACCASCIA; KONIG <i>et al.</i> , 2017)	Sistema de dois plasmídeos. Oligos complementares de gRNA ligados ao plasmídeo preparado com Bsal.	ssDNA e dsDNA oligos (70 nt and 120 nts).
(SHUKAL; LIM; ZHANG; CHEN, 2022)	gRNA preparado por clonagem sem restrições Sistema de dois plasmídeos (gRNA e Cas9/ λ -Red).	Donor linear assimétrico com 50 bp / 500 bp, preparado via PCR.
(ZHAO; YUAN; XIONG; SUN <i>et al.</i> , 2016)	gRNA e <i>donor</i> montados via Golden Gate. Um plasmídeo para Cas9, λ -Red, gRNA e <i>donor</i> .	<i>Donor</i> com braços de homologia de 41 – 300 bp, clonado com o gRNA via Golden Gate assembly.
(LI; LIN; HUANG; ZHANG <i>et al.</i> , 2015)	Sistema de dois plasmídeos. Plasmídeo amplificado por PCR com gRNA codificado por primer, Golden Gate assembly.	<i>Donors</i> com braços de homologia de 300 – 500 bp produzidos via <i>fusion PCR</i> .
(FANG; ZHAO; ZHAO; DONG <i>et al.</i> , 2024)	Plasmídeo com gRNA e Cas9 combinados. gRNA clonado através de MetClo assembly.	Plasmídeo de <i>donor</i> clonado usando MetClo assembly.
(WANG; SUI; DING; FU <i>et al.</i> , 2021)	Sistema de dois plasmídeos. gRNA clonado via PCR. Abordagem para cura de plasmídeo Rock-Paper-Scissors.	<i>Donor</i> clonado no plasmídeo de gRNA usando kits de clonagem contínua ou restrição e ligação.

(LIM; LEE; KIM; LEE, 2023)	Cepa portadora de Cas9. Clonagem de plasmídeo para gRNA usando Gibson <i>assembly</i> . Uso de 5'-truncado gRNAs.	<i>Donors</i> como oligonucleotídeos de 41–70-mer.
(FENG; ZHAO; ZHANG; DING <i>et al.</i> , 2018)	Sistema de três plasmídeos (Cas9/ λ -Red, gRNA e <i>donor</i>). Golden Gate <i>assembly</i> .	<i>Donors</i> clonados e entregues em plasmídeos. Golden Gate <i>assembly</i> .
(MOREB; HOOVER; YASEEN; VALYASEVI <i>et al.</i> , 2017)	Sistema de dois plasmídeos (gRNA/ λ -Red e Cas9/recA56). Clonagem de gRNA com kits comerciais.	<i>Donors</i> como oligos ssDNA e dsDNA.
(SUN; WANG; MAO; FEI <i>et al.</i> , 2019)	Sistema de dois plasmídeos (Cas9/ λ -Red e gRNA). gRNA clonado por <i>inverse PCR</i> .	<i>Donors</i> lineares montados com <i>fusion PCR</i> .
(TANG; LOU; LIU, 2017)	Sistema de três plasmídeos (Cas9, λ -Red e gRNA/SceI). gRNA clonado através PCR.	Donor dsDNA linear fornecido como um produto de PCR montado por meio de PCR de extensão de sobreposição.
(ZHAO; GUO, 2024)	Sistema de dois plasmídeos: Cas9/Beta e gRNA, clonado por restrição e ligação e recombinação homóloga.	Linear dsDNA gerado por meio de PCR de sobreposição de extensão.

2.5 *In vivo cloning* em *E. coli*

In vivo cloning ou *in vivo assembly* em *E. coli* é uma técnica que permite a ligação de fragmentos de DNA diretamente no interior da célula (WATSON; GARCIA-NAFRIA, 2019). Nessa estratégia de clonagem não são realizadas ligações *in vitro* mas sim *in vivo*, usufruindo assim dos sistemas de recombinação das células (KOSTYLEV; OTWELL; RICHARDSON; SUZUKI, 2015). Há trabalhos da década de 1990 que relatam o uso dessa ferramenta de clonagem em *E. coli*, como os de Bubeck e colaboradores (BUBECK; WINKLER; BAUTSCH, 1993) e Oliner (OLINER; KINZLER; VOGELSTEIN, 1993). Desde de então, há muitos trabalhos que relatam essa técnica de clonagem.

Para a execução da técnica é necessária a produção de fragmentos através da amplificação do DNA template por PCR. Quando necessário, recomenda-se o tratamento com DpnI. A recomendação do método de transformação é o uso do choque-térmico (BUBECK; WINKLER; BAUTSCH, 1993). Entretanto, Oliner (1993) sugere que a eletroporação pode ser usada para a transformação de outras cepas de *E. coli* diferentes da DH5α.

Uma das partes mais importantes para a efetiva utilização da técnica de *in vivo cloning* em *E. coli* é o desenho dos primers para a amplificação dos fragmentos por PCR. Esses primers precisam ter caudas 5' homólogas com os fragmentos que serão ligados. Vários trabalhos sugerem diferentes extensões dessas caudas. Bubeck e colaboradores (1993) relatam o uso de um tamanho mínimo de 5 pb, chegando a 48 pb. Oliner (1993) relata um tamanho de homologies entre o vetor e o inserto variando de 48 até 88 pb. Outros trabalhos sugerem o uso de outros tamanhos de homologia de 10-50 pb (JACOBUS; GROSS, 2015; KOSTYLEV; OTWELL; RICHARDSON; SUZUKI, 2015).

Os mecanismos responsáveis por esse processo de montagem de fragmentos no interior da célula ainda não estão bem caracterizados (WATSON; GARCIA-NAFRIA, 2019). Acredita-se que há uma atividade exonucleotídica 3'-5', independente de RecA, que seja responsável por gerar regiões de DNA simples fita (CONLEY; SAUNDERS; SAUNDERS, 1986; NOZAKI; NIKI, 2019). Essas regiões correspondem a extremidades coesivas entre os fragmentos de clonagem e, portanto, podem se anelar e depois serem reparadas por DNA polimerase e ligase (NOZAKI; NIKI, 2019).

Apesar dessa grande possibilidade de aplicações, um dos fatores que podem influenciar a aplicabilidade da técnica é o número de fragmentos que ela é capaz de ligar. O número ideal, para ligar usando a cepa DH5 α , com alta eficiência, são dois fragmentos, com a possibilidade de uso de mais fragmentos em determinadas situações. Para contornar esse problema podem ser usadas cepas com recombinases, como o Lambda Red, ou células ultracompetentes para transformação. Entretanto, há o problema do custo de adquirir essas células e a possibilidade de ocorrência de combinações indesejadas. Outra alternativa é a possibilidade do uso de clonagem usando a capacidade de montagens de fragmentos de DNA em levedura *Saccharomyces cerevisiae* (BASSALO; GARST; HALWEG-EDWARDS; GRAU *et al.*, 2016).

2.6 Competições

No processo de engenharia reversa, que segue o experimento de evolução adaptativa em laboratório, para avaliar a contribuição de mutações-chave encontradas, é necessário haver um sistema com alta precisão para aferição dos efeitos proporcionados pelas alterações genéticas (JACOBUS; CAVASSANA; DE; BARRETO *et al.*, 2024). Em um primeiro momento, a análise de curvas de crescimento pode ser uma ferramenta que cumpra essa necessidade. Entretanto, as variações na capacidade de crescimento causadas por mutações reconstruídas por engenharia reversa podem não causar variações tão proeminentes que possam ser observadas em uma curva de crescimento (GULLBERG; CAO; BERG; ILBACK *et al.*, 2011). Assim, há a necessidade de sistemas mais robustos para a avaliação dos efeitos das reconstruções dos alelos mutados no progenitor do experimento de evolução.

Uma alternativa que cumpre os requisitos é o ensaio de competição. Nesse tipo de experimento é possível fazer uma caracterização com extrema precisão dos efeitos fenotípicos de mutações reconstruídas (JACOBUS; CAVASSANA; DE; BARRETO *et al.*, 2024). O que torna o ensaio de competição mais robusto é o fato de ser o produto de todo o crescimento, passando por todas as suas fases, *lag*, exponencial e estacionária (GULLBERG; CAO; BERG; ILBACK *et al.*, 2011). Já quando é realizada uma análise uma curva de crescimento, foca-se apenas no perfil da curva, ou na fase exponencial.

O ensaio envolve a competição entre a população evoluída (ou um clone amostrado dessa), ou a linhagem submetida a engenharia reversa, contra o parental do experimento de evolução (JACOBUS; CAVASSANA; DE; BARRETO *et al.*, 2024). Os dois são misturados em uma proporção definida e propagados em condições específicas, geralmente refletindo o ambiente do experimento de evolução adaptativa (COOPER; LENSKI, 2000). Devido à mistura de células, é necessário haver uma forma de diferenciar os dois competidores, de maneira que seja possível a contagem de células e a aferição da proporção entre os competidores (COOPER; LENSKI, 2000). Essa aferição pode ser realizada por meio de plaqueamento em ágar diferencial, se houver uma diferença metabólica entre os competidores, ou pela etiquetagem de pelo menos um dos competidores com uma marca de fluorescência que permita quantificação das células pelo uso de citometria de fluxo (GULLBERG; CAO; BERG; ILBACK *et al.*, 2011; JACOBUS; CAVASSANA; DE; BARRETO *et al.*, 2024).

Um exemplo de aplicação de ensaios de competição é no LTEE (LENSKI; ROSE; SIMPSON; TADLER, 1991). Nesse experimento as cepas ancestrais REL606 e REL607 são *ara*⁺ (metabolizam arabinose) ou *ara*⁻ (não metabolizam arabinose), respectivamente. Quando colocadas para crescer em meio sólido com arabinose e tetrazólio (indicador de pH), aquelas que consomem o açúcar acidificam o meio e formam colônias com um vermelho intenso. Aquelas colônias que não consomem arabinose mantêm uma aparência esbranquiçada. Assim, por exemplo, após misturar e crescer na condição analisada as células que consomem a arabinose com as populações que não consomem e consomem a arabinose é possível realizar o plaqueamento e contar as colônias e verificar a proporção de células (LENSKI; MONGOLD; SNIEGOWSKI; TRAVISANO *et al.*, 1998). Apesar de ser uma estratégia relativamente simples, há a necessidade de um grande número de plaqueamentos. Além disso, a avaliação é feita após o crescimento das colônias e assim, se houver algum erro durante o plaqueamento o resultado pode ser mensurado erroneamente.

Uma forma de contornar esse problema é o uso de células marcadas, por meio da expressão de proteínas fluorescentes. Assim, pode-se, com o uso da citometria de fluxo separar e estabelecer a proporção de células marcadas e não marcadas, além de contá-

las (GULLBERG; CAO; BERG; ILBACK *et al.*, 2011). A principal vantagem dessa estratégia é que ela permite uma análise direta da cultura em análise.

Através do ensaio de competição é possível determinar o aumento da capacidade de crescimento, tanto das populações evoluídas, como das mutações individuais reconstruídas. Para expressar em números esse aumento do *fitness* é expresso usando o coeficiente de seleção (S), o qual é expresso pela equação $S = \ln[Cf/GFPf] - \ln[Ci/GFPi]$, onde Cf representa a proporção de células que estão sob análise no fim do crescimento, Ci a proporção dessas células no início, enquanto o GFPi e GFPf representam as células marcadas por GFP no início e fim do experimento (JACOBUS; CAVASSANA; DE; BARRETO *et al.*, 2024). Sabendo o número de gerações é possível calcular o S/doubling, correspondendo ao coeficiente de seleção por cada geração. Se o S/doubling for positivo há um incremento na capacidade de crescimento das células analisadas (populações/clones evoluídos ou linhagens engenheiradas). Quando for negativo há a redução na capacidade de crescimento.

2.7 Polihidroxialcanoatos

Os PHAs são polímeros naturalmente produzidos por bactérias como uma forma de armazenar carbono em situação de falta de certos nutrientes (LI; YANG; LOH, 2016). Esses organismos armazenam esse polímero na forma de grânulos intracelulares, que, quando extraídos, podem ser usados na substituição de plásticos convencionais, fabricação de filmes biodegradáveis e controladores de liberação de compostos, dentre outros (LI; YANG; LOH, 2016; PANDEY; ADAMA; ADJALLE; BLAIS, 2022). Com essas aplicações são possíveis candidatos a substituição de produtos oriundo de petróleo.

Estruturalmente são poliésteres que apresentam alta biodegradabilidade por microrganismos, transformando-se em dióxido de carbono, água e biomassa quando degradados (SUZUKI; TACHIBANA; KASUYA, 2021). As unidades monoméricas podem ser fixas ou variáveis, como poli(3-hidroxibutirato) (PHB) e o poli(3-hidroxivalerato) (PHV). Dependendo do número de carbonos nas unidades monoméricas, pode ser classificado em PHAs de cadeia curta com 3 a 5 carbonos; os de cadeia média possuem de 6 a 14 carbonos na unidade monomérica; e aqueles com a partir de 15 carbonos são chamados

de cadeia longa (LI; YANG; LOH, 2016; SHARMA; SEHGAL; GUPTA, 2021). Em virtude dessas diferenças na sua composição polimérica, esses bioplásticos podem diferir de forma significativa as suas características físicas (BEHERA; PRIYADARSHANEE; VANDANA; DAS, 2022).

No grupo dos PHAs, o mais estudado é o PHB, isso em virtude de suas propriedades físicas, mecânicas e imunológicas (TRAKUNJAE; BOONDAENG; APIWATANAPIWAT; KOSUGI *et al.*, 2021). Esse biopolímero é formado por subunidades de 3-hidroxibutirato. Tendo como principal produtor a bactéria *Cupriavidus necator* ou *Ralstonia eutropha*, esse polímero possui um alto ponto de *melting* (~175 °C to 180 °C) e baixa temperatura de degradação (~220 °C) permitindo o uso de técnicas de moldagem como termoformagem e injeção (ANBUKARASU; SAUVAGEAU; ELIAS, 2015; MCADAM; BRENNAN FOURNET; MCDONALD; MOJICEVIC, 2020). Devido a sua alta biodegradabilidade, o seu emprego pode ser feito tanto em embalagens quanto em itens descartáveis, além de poder ser usado em aplicações médicas, como fios de suturas, curativos, malhas de regeneração (MCADAM; BRENNAN FOURNET; MCDONALD; MOJICEVIC, 2020).

Apesar dessas possíveis aplicações, o custo de produção do PHB, comparado aos polímeros baseados em petróleo, ainda não é competitivo, apresentando baixo rendimento e complexidade tecnológicas (ADNAN; SIDDIQUI; ASHRAF; SNOUSSI *et al.*, 2022; MCADAM; BRENNAN FOURNET; MCDONALD; MOJICEVIC, 2020). Por exemplo, o custo do propileno, que possui propriedades semelhantes ao PHB, pode ser algumas vezes menor que o PHB (AKKOYUNLU; GABARRE; DALY; CASEY *et al.*, 2024). Como a manufatura desse biopolímero precisa ainda de melhorias, há uma necessidade de tecnologias que facilitem o desenvolvimento da produção industrial desse polímero.

Uma das ferramentas para melhoria dessa produção é o uso de organismos geneticamente modificados (MCADAM; BRENNAN FOURNET; MCDONALD; MOJICEVIC, 2020). Nesse contexto, o uso da bactéria *E. coli* é interessante por diversos motivos. Um primeiro ponto é o fato dessa bactéria possuir crescimento rápido e facilidade de cultivo, podendo ser crescer em diversos meios de cultivo, com características de crescimento bem conhecidas e sistemas industriais de produção bem desenvolvidos (JIANG; WANG; SONG; YU *et al.*, 2023). Além disso, as ferramentas

genéticas básicas para manipulação desse microrganismo são bem avançadas, além de possuir o seu genoma bem caracterizado.

No cultivo de *E. coli* engenheirada para a produção de PHB não há a necessidade de estressá-la com redução de concentração de nutrientes durante o crescimento do microrganismo para que comece a produzir o bioplástico (LEE, 1996). Outro ponto positivo da produção do PHB em *E. coli* é a capacidade dessa bactéria de sintetizar esse polímero em quantidades superiores às observadas em microrganismos produtores naturais, perfazendo 80-90% da massa celular seca (LEE, 1996; PENA; CASTILLO; GARCIA; MILLAN *et al.*, 2014). Além disso, usando determinadas condições de cultivo é capaz de produzir PHB de alto peso molar, o qual tem suas características físicas alteradas, podendo ter aplicações específicas e nobres, como o uso médico (BOCANEGRA; DA CRUZ PRADELLA; DA SILVA; TACIRO *et al.*, 2013; PEÑA; LÓPEZ; GARCÍA; ESPÍN *et al.*, 2014).

Em adição às modificações genéticas, o uso de fontes de carbono alternativas, sustentáveis e de baixo custo também podem ser um foco para a melhoria da produção do PHB (WU; ZHOU; PEI; JIANG *et al.*, 2022). Uma dessas alternativas é a sacarose.

2.8 Sacarose

A sacarose é um dissacarídeo formado por glicose e frutose, produzida comercialmente a partir da cana e beterraba (GETZ, 2000; WACLAWOVSKY; SATO; LEMBKE; MOORE *et al.*, 2010). O açúcar produzido pela cana-de-açúcar apresenta uma grande interesse para o Brasil em virtude da grande capacidade de cultivo e os seus níveis de produção (ALMEIDA; SILVA; SILVA; SILVA *et al.*, 2024).

O uso da sacarose é destacada principalmente para alimentação humana. Entretanto, o açúcar oriundo da cana pode ser usado como fonte de carbono em bioprocessos, especialmente para a produção de etanol (LIN; ZHANG; PAKRASI, 2020). Além de álcool, esse açúcar pode ser usado na produção de aminoácidos, polifenóis, bioplásticos, dentre outros (NI; CHEN; TIAN; XU *et al.*, 2022). Como essa é uma fonte de carbono considerada renovável e de baixo custo, pode ser uma opção para uso como fonte de carbono para processos biológicos na produção de bioprodutos como o PHB

(MOHAMED; MUNDHADA; LANDBERG; CANN *et al.*, 2019; WU; ZHOU; PEI; JIANG *et al.*, 2022).

3 OBJETIVOS

- Realizar a evolução adaptativa em sacarose das linhagens de *E. coli* DH5 α e E2348/69;
- Identificar por sequenciamento *Illumina* mutações possivelmente responsáveis pelo aumento do fitness durante evolução;
- Desenvolver ferramentas CRISPR/Cas9 para realização de engenharia reversa de mutações identificadas no experimento de evolução;
- Teste de *fitness* por ensaios de competição das mutações engenheiradas;
- Construção de linhagem de *E. coli* produtora de PHB a partir da sacarose.

4 MATERIAIS E MÉTODOS

4.1 Linhagens celulares, armazenamento, condições de crescimento e plasmídeos

As linhagens celulares e os plasmídeos usados estão descritos no Apêndices A e B. O armazenamento das células foi realizado em freezer -80 °C, em glicerol a 25% (v/v). A pureza das culturas foi avaliada através de crescimento em ágar MacConkey, preparado conforme instruções do fabricante.

Para todas as manipulações, as células bacterianas foram crescidas em meio *Lysogenic Broth* (LB) (10 g/L triptona, 5 g/L extrato de levedura, 10 g/L NaCl) a 37 °C, adicionando-se ágar (15 g/L) quando há a necessidade de meio sólido. Quando necessário, realizou-se a adição de antibiótico(s) Ampicilina (100 µg/mL), Canamicina (50 µg/mL), Espectinomicina (50 µg/mL) e Cloranfenicol (50 µg/mL).

Para obtenção de plasmídeos a serem usados nas clonagens ou como produto destas a partir de *E. coli*, 50 µL de células bacterianas hospedando o plasmídeo de interesse foram postas para crescer em meio em 5 mL de LB sob agitação com adição do respectivo antibiótico, entre 12-16 h. Após o crescimento, foi realizada a extração plasmidial usando o kit QIAprep® Miniprep da Qiagen, seguindo as instruções do fabricante. Após a extração, o material foi quantificado em *Qubit 2.0 Fluorometer* (Invitrogen), seguindo instruções do fabricante. Sendo em seguida mantidos a -20 °C para posterior utilização.

Para purificar os plasmídeos clonados em levedura, as células foram postas para crescer em YPS (10 g/L de extrato de levedura, 20 g/L de peptona e 20 g/L sacarose), com o respectivo antibiótico, e purificados usando o kit QIAprep® Miniprep da Qiagen, com adaptações.

4.2 Procedimentos de clonagem

Todos os fragmentos de DNA usados para as clonagens gerados através de PCRs foram realizadas com oligonucleotídeos (primers) adquiridos da empresa Exxtend. Também foram encomendados dessa empresa primers para as PCRs confirmatórias de

ligações entre os fragmentos. Para os procedimentos de clonagem usando *E. coli* foi usada a técnica de *in vivo cloning* (JACOBUS; GROSS, 2015). Os primers e suas sequências estão descritos no Apêndice F.

Todas as PCRs destinadas a obter fragmentos para clonagens de plasmídeos ou como *donor* para integrações cromossômicas foram realizadas usando Phusion High-Fidelity DNA Polymerase (Thermo Fisher Scientific ou NEB) ou Genie Fusion Ultra High-Fidelity DNA Polymerase (AssayGenie) seguindo as instruções do fabricante. PCRs confirmatórias de integrações foram realizadas com GoTaq™ DNA Polymerase (Promega), seguindo as instruções do fabricante. Os produtos das PCRs foram analisados em eletroforese em agarose 0,8-3%, em buffer TBE (Tris-Borato-EDTA), com brometo de etídio a 0,01% para visualização do DNA. A corrida foi realizada com voltagem constante de 90V. O gel obtido foi fotografado em transluminador UV.

Quando necessário, os fragmentos obtidos por PCR foram purificados das reações usando o QIAquick PCR Purification Kit (Qiagen), seguindo as instruções do fabricante. Quando necessário purificar o fragmento a partir do gel de agarose, foi usado o QIAquick Gel Extraction Kit (Qiagen), seguindo as instruções do fabricante.

Para a geração de oligos gRNA e *donor*, os oligos foram diluídos a 10 pmol/μL. Para PCR, foram usados 2,5 μL de oligos flanqueadores e 1,25 μL do oligo central. A PCR envolveu uma desnaturação inicial a 95 °C, 30 ciclos de 95 °C por 10 segundos, 50 °C por 15 segundos e 72 °C por 10 segundos, com uma extensão final a 72 °C por 1 minuto. Os produtos foram analisados por eletroforese em géis de agarose a 3%.

Os detalhes das modificações clonagens estão descritos nos Apêndices A e B.

4.3 Preparação das células bacterianas para transformações

As transformações foram realizadas com método de choque térmico e eletroporação (MANIATIS, 1982). O preparo de células competentes para transformação usando choque térmico foi realizado seguindo o protocolo de preparação com cloreto de rubídio Promega *Protocols and Applications Guide (3rd edition)* (protocolo adaptado de (HANAHAHAN, 1983)). Para as transformações por choque térmico, o DNA foi incubado por 30 min no gelo com 50 μL de células competentes. Em seguida, por 1 min, as células

foram incubadas em banho a 42 °C; após isso, foi realizada mais uma imersão de gelo por 20 min. Após esse tempo as células foram recuperadas em 1mL de meio LB por 1 hora sob agitação a 37 °C. Por fim, as células foram centrifugadas por 3 min a 8.000 rpm, tendo o seu sobrenadante descartado e as células ressuspensas e plaqueadas com alça de Drigalsky na placa com antibiótico(s) necessário(s) e incubadas a 37 °C por 12 h a 16 h.

Para a eletroporação, a partir de um pré-cultivo overnight de células, foram postas para crescer em LB, na proporção 1:100. Após atingir a OD de 0,5-0,7, o cultivo foi centrifugado para a precipitação das células. Depois da remoção do sobrenadante, o pellet foi ressuspensado em glicerol 10% (v/v), sendo novamente centrifugado. Após três dessas lavagens, o pellet formado foi ressuspensado em 500 µL de glicerol 10% (v/v), sendo em seguida transferidos 50 µL dessas células para microtubos de 1,5 mL, os quais foram usados para transformação por eletroporação ou armazenados a -80 °C para posterior utilização.

Todas as eletroporações foram realizadas a 2,5 kV/5 ms (eletroporador Bio-Rad Micropulser e cubetas Bio-Rad). Após a eletroporação, as células foram recuperadas em 1 mL de meio LB por 1 hora sob agitação a 37 °C. Por fim, as células foram centrifugadas por 3 min a 8.000 rpm, tendo o seu sobrenadante descartado e as células ressuspensas e espalhadas na placa com antibiótico(s) de interesse e postas para crescer em 37 °C.

4.4 Modificações genômicas por CRISPR/Cas9

Células de *E. coli* DH5α foram transformadas através de choque térmico com o plasmídeo CAS9BAC1P-1EA e foram selecionadas em placas de LB com canamicina. As colônias foram selecionadas por PCR. Das colônias positivas, uma foi posta para crescer *overnight* em LB. Após isso, a cultura foi inoculada em 30 mL de LB na proporção 1:100, na presença de canamicina, e posta para crescer a 30°. Quando a OD alcançou a um valor entre 0,2-0,3, foi adicionada L-arabinose à concentração final de 10 mM. Quando o crescimento atingiu a OD de 0,5-0,7 foram preparadas para transformação por eletroporação.

Para a realização da técnica de CRISPR/Cas9 usando eletroporação, foram utilizados 400 ng de *donor* e 40 ng do plasmídeo de gRNA para transformar as células abrigando o plasmídeo CAS9BAC1P-1EA. Quando foi usado choque térmico, foram transformados 40 ng do plasmídeo de gRNA purificado e 10 µL do produto de PCR do *donor*. A seleção foi feita em placas de LB contendo canamicina, espectinomicina e L-arabinose. Para a controle foi realizada a transformação das células com o plasmídeo da Cas9 apenas com 40 ng do plasmídeo do gRNA. Como controle de transformação, foi realizada a transformação com 40 ng do pSpec (plasmídeo controle). As células transformadas foram postas para recuperar por 1 h, em 1 mL de meio, com L-arabinose a 10 mM, a 30 °C em agitação. Após isso, esse cultivo foi plaqueado como descrito anteriormente. Os detalhes das modificações genômicas estão descritos nos Apêndices E e F.

4.5 Experimento de ALE

4.5.1 Meio mínimo suplementado com sacarose

O meio M9 foi preparado com concentração de 5x, contendo os sais $\text{NaH}_2\text{PO}_4 \cdot \text{H}_2\text{O}$ (36,4 g/L), KH_2PO_4 (15 g/L), NaCl (2,5 g/L) e NH_4Cl (5 g/L). Após autoclavagem e antes do uso (no momento de diluir para 1x) foram adicionados MgSO_4 a 1 M, CaCl_2 a 2 M (concentração final 2 mM e 100 µM). Além disso, foram adicionadas soluções de sacarose a 200 g/L, tiamina 1% (m/v), minerais traço 10.000x (Tabela 1). Ao final, a concentração da sacarose foi de 20 g/L, da tiamina 0,01% (m/v).

Tabela 1: Solução de minerais 10.000x

Composto	Concentração
H_3BO_3	5 g/L
MnSO_4	4 g/L
$\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$	4 g/L
$\text{FeCl}_2 \cdot 6\text{H}_2\text{O}$	2 g/L

Ácido molíbdico	0,4 g/L
KI	1 g/L
CuSO ₄ ·5H ₂ O	0,4 g/L
Ácido Cítrico	10 g/L

4.5.2 Clonagem do operon *cscBKA*

Como a maioria das cepas industriais ou laboratoriais de *E. coli* normalmente não metabolizam sacarose, foi necessária a clonagem e inserção dos genes necessários à metabolização desse dissacarídeo. Existem diferentes estratégias que os microrganismos usam para utilização de sacarose (LEE; CHOI; PARK; VICKERS *et al.*, 2010). Neste estudo foi usado o operon *cscBKA* para a metabolização desse açúcar. Ele é formado pelos genes *cscB*, o qual codifica o transportador para a sacarose, *cscK*, especificando uma frutocinase e *cscA*, codificando uma invertase (Figura 1 A). A expressão desse operon é realizada por um promotor bidirecional, o qual é naturalmente regulado pelo repressor *cscR* que é ativo apenas quando a concentração de sacarose no meio fica abaixo de 0,2% (m/v) (MOHAMED; MUNDHADA; LANDBERG; CANN *et al.*, 2019; SABRI; NIELSEN; VICKERS, 2013). Esse repressor não foi usado neste trabalho. Dessa forma, foram clonados apenas os genes responsáveis pela metabolização da sacarose.

A linhagem *E. coli* E2348/69 tem no seu genoma o operon para metabolização da sacarose (IGUCHI; THOMSON; OGURA; SAUNDERS *et al.*, 2009). Dessa maneira foi decidido realizar a clonagem desses genes em um plasmídeo através da técnica de *in vivo cloning* (JACOBUS; GROSS, 2015). O operon *cscBKA* foi amplificado do DNA genômico de *E. coli* E2348/69, sem terminador. O plasmídeo pUC57_*sucP* foi ponto de partida para amplificação por PCR do *backbone* plasmidial para a clonagem, contendo a origem de replicação, a marca de resistência à ampicilina e o terminador B0015 (Figura 1 A). Duas regiões com 400 pb de homologia com a região genômica SS9 (BASSALO; GARST; HALWEG-EDWARDS; GRAU *et al.*, 2016) também foram incluídas visando uma futura integração cromossômica. A funcionalidade desse operon clonado no plasmídeo

foi avaliada através do crescimento em meio mínimo, com sacarose como única fonte de carbono, em comparação com controle, sem a presença do plasmídeo. Além disso, o sequenciamento Sanger da região do operon *cscBKA* desse plasmídeo, com as respectivas regiões de homologia, mostrou que o mesmo estava clonado da forma esperada. Foi também quantificado o consumo de sacarose por células albergando esse plasmídeo, confirmando assim a sua funcionalidade.

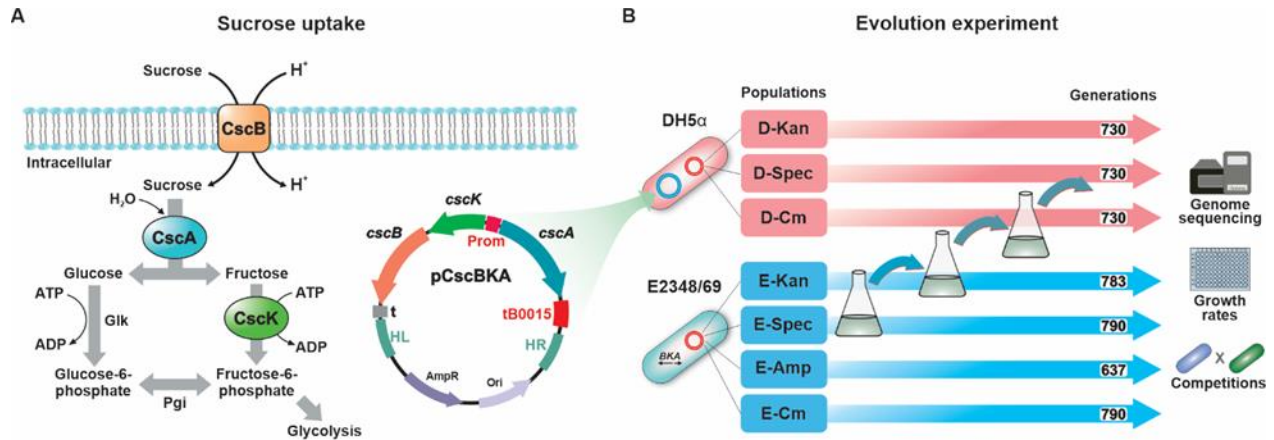
4.5.3 Construção das linhagens parentais (progenitoras) para a ALE

Para dar início ao experimento de evolução foi necessária a montagem de linhagens parentais com resistência a diferentes antibióticos. Isso teve por objetivo evitar contaminações cruzadas no decorrer do processo. Para o estabelecimento dessas populações, foram clonados vetores plasmidiais que contivessem apenas a marca de resistência a um antibiótico e a origem de replicação como unidades funcionais. Assim foram clonados *in vivo* os plasmídeos pKan (resistência a canamicina), pSpec (espectinomicina) e pAmp (ampicilina). Para a resistência ao cloranfenicol, foi usado o plasmídeo pGGA, presente no kit comercial Golden Gate (NEB).

Estes plasmídeos foram transformados em sete populações progenitoras para serem usadas no experimento de evolução (Figura 1 B). Para a *E. coli* E2348/69, foram fundadas quatro populações, sendo chamadas E-Kan, E-Spec, E-Amp e E-Cm. Não foi necessária a transformação do plasmídeo pCscBKA em virtude dessa cepa já metabolizar de forma natural a sacarose.

As outras três populações foram construídas na linhagem *E. coli* DH5α contendo o plasmídeo pCscBKA (resistência à ampicilina), constituindo as populações D-Kan (resistência a canamicina e ampicilina), D-Spec (espectinomicina e ampicilina) e D-Cm (cloranfenicol e ampicilina).

Figura 1: Experimento de evolução adaptativa de laboratório realizado em sacarose a 20 g/L



A Estrutura da via catabólica da sacarose formada pela expressão do operon *cscBKA* do genoma de *E. coli* E234869 clonado no plasmídeo pCscBKA. *cscB* codifica um transportador de sacarose, *cscA* codifica uma invertase e *cscK* uma frutocinase. **B** Experimento de evolução. Três populações de *E. coli* DH5α e quatro de *E. coli* E2348/69 foram fundadas e são diferenciadas por marcadores de resistência codificados por plasmídeo para Kanamicina (Kan), Espectinomicina (Spec), Ampicilina (Amp) ou Cloranfenicol (Cm). As populações DH5α foram adicionalmente transformadas com o plasmídeo pCscBKA. O experimento foi conduzido por meio de transferências diárias em sacarose a 20 g/L por 637-790 gerações, conforme indicado. As populações finais foram submetidas a sequenciamento de DNA genômico, crescimento em microplacas e ensaios de competição.

4.5.4 Condução do experimento de ALE

O processo de evolução foi realizado em meio mínimo contendo como única fonte de carbono 20 g/L de sacarose. Foi realizado em Erlenmeyers de 50 mL, com volume de 20 mL de meio, os quais foram mantidos sob agitação de 80 rpm a 37 °C (Figura 1 B). Diariamente, foram feitas passagens na diluição de 1:100 (v) para meio fresco, além de ser medida a densidade óptica (OD) dos crescimentos. Semanalmente foram feitos estoques de cada população em glicerol 25% (v/v), armazenados em *ultrafreezer*.

Uma bateria de testes foi estabelecida para verificar contaminações das populações em evolução. Semanalmente (ou sempre que havia suspeita de alguma possível contaminação) eram feitos testes em ágar MacConkey, coloração de Gram e

testes de crescimento em diferentes antibióticos. Também, a OD na fase estacionária era monitorada em cada passagem, e quando essa medida aumentava de forma incomum, eram feitos os testes de contaminação. Em caso de contaminação, a cultura era reiniciada a partir do estoque de glicerol conservado em *ultrafreezer*. Contaminações foram detectadas duas vezes com a população E-Amp, e esse protocolo foi aplicado e o experimento retomado.

Ao final do processo, o DNA genômico das populações finais e progenitor foi purificado usando o método fenol:clorofórmio, tendo a sua qualidade analisada por eletroforese em agarose, quantificado e evaporado para posterior sequenciamento em plataforma *Illumina* pela empresa GenOne.

4.5.5 Sequenciamento genômico das populações evoluídas

Após o término do experimento de evolução realizou-se o sequenciamento genômico das populações DH5 α e E2348/69. Primeiramente, o DNA genômico foi extraído das populações evoluídas e dos progenitores. O material foi enviado para a empresa GenOne: Soluções em Biotecnologia, para sequenciamento na plataforma *Illumina*. Para identificar mutações (SNPs e InDels), as *reads* das populações evoluídas foram filtradas por seleção de tamanho (> 50 bases) e pontuações de qualidade geral Q>30. O mapeamento foi feito contra o genoma de referência da DH5 α e E2348/69 (números de acesso no genbank CP026085.1 e FM180568, respectivamente). Os mapeamentos das *reads* foram feitos utilizando o algoritmo *Burrows-Wheeler transform* implementado na plataforma CLC Genomics Workbench 8.01 (QIAGEN). Foram especificados limites de seleção de 0,8 para fração mínima da *read* a ser alinhada e 0,8 para similaridade mínima de alinhamento. A identificação de SNPs/InDels foi realizada diretamente a partir de mapeamentos de *reads* usando algoritmo de detecção de variantes dentro do pacote CLC Genomics Workbench 8.01. Um corte de presença de variantes em pelo menos 20% das leituras foi observado. As mutações encontradas estão apresentadas nas Figuras 4 A e B e no Apêndice D. As mutações encontradas foram avaliadas para analisar os seus possíveis efeitos e a possibilidade de construção de

mutantes com essas variações. As *reads* do sequenciamento foram depositadas no NCBI sob o *BioProject number* PRJNA1200122.

4.5.5 Ensaios de competição

Para leituras em citômetro de fluxo, 10 µL da cultura foram adicionados em 990 µL de PBS e agitado em vórtex. Após filtragem, desse material foi realizada a leitura em no citômetro Attune NxT (Thermo Fisher Scientific), previamente ajustado.

Para os ensaios de competição foram realizadas três passagens dos competidores em meio mínimo contendo sacarose para adaptação, tanto para as células em análise como as células expressando o GFP.

A montagem do experimento consistiu na mistura das células do parental com GFP e das população evoluídas, ou os mutantes obtidos da engenharia reversa. Em seguida foi realizada a leitura em citômetro. Esse passo foi repetido até o mix de células atingir a proporção inicial desejada de células com e sem GFP. Para análise de reengenharia reversa a proporção inicial foi ajustada para 50% com GFP e 50% das células com a mutação analisada. Para algumas situações as proporções entre os competidores precisaram ser adequadas de forma desigual. Por exemplo, a proporção para as células de DH5α com o plasmídeo pCscBKA foi de 95%, enquanto que o parental marcado com GFP foi ajustado para 5%. Para as populações evoluídas de E2348/69, as proporções da referência com GFP e populações evoluídas foram mantidas em 1:1. Essa adaptação foi necessária em virtude da necessidade de realizar leituras nos intervalos adequados, e as células com menor *fitness* não desaparecerem completamente durante o experimento.

Após isso, 200 µL do mix de células contendo a referência GFP e o competidor em análise foram usados para inocular 20 mL nas mesmas condições de condução do experimento de ALE, em triplicata. Após crescimento por 24h, essas triplicatas foram analisadas em citômetro de fluxo e novamente foi realizada uma passagem de 200 µL para inocular 20 mL de meio mínimo frasco, mantendo as triplicatas. Esse processo foi repetido três vezes, ou até um dos competidores não ser mais detectado pelo citômetro.

Ao final as proporções dos competidores foram aferidas e usadas para calcular o coeficiente de seleção (S) (DELUNA; VETSIGIAN; SHORESH; HEGRENESS *et al.*, 2008).

4.6 Quantificação de sacarose

A quantificação de sacarose foi realizada com o auxílio do professor Dr. Michel Brienzo e da Dra. Daniele Marin.

Para a quantificação de sacarose, as células foram pré-adaptadas em meio mínimo contendo sacarose como única fonte de carbono, em Erlenmeyer de 125 mL, a 37 °C e 100 rpm. Foram feitas três passagens nessas condições antes de iniciar o cultivo para quantificação. Esse foi realizado em triplicata e, após 24 horas de cultivo, foi coletado o sobrenadante, sendo usado para quantificação de sacarose. Para quantificação, as amostras foram diluídas com água ultrapura, filtradas em filtro de seringa 0,22 µm.

A quantificação foi realizada usando cromatografia líquida de alta eficiência (HPLC), com protocolo adaptado de Freitas e Brienzo (2023) (DE FREITAS; BRIENZO, 2023). O equipamento usado foi o Waters 2414, com a coluna Aminex HPX-87C BIO-RAD (300 × 7.8 mm) à 80 °C. A fase móvel usada foi água ultrapura, com o fluxo de 0,6 mL/min. 20 µL de amostra foram usados para quantificação.

4.6 Quantificação de PHB

A quantificação de PHB foi realizada com o auxílio do mestrando Matheus Victor Maso Lacôrte e Silva.

Após o término do cultivo das cepas o produto da cultura foi centrifugado a 8.000 x g por 20 min. a 4 °C e o sobrenadante descartado, a biomassa recuperada foi liofilizada. Cada 10 mg do liofilizado de células foi submetido à metanólise. Essa etapa foi realizada da seguinte maneira: em tubos de ensaio de vidro foram adicionados 2 mL de clorofórmio, 2 mL de metanol acidificado com 15% de ácido sulfúrico (85:15 v / v) e 100 µL de ácido benzóico em metanol, com proporção de 0,02 g/mL, sendo agitados em vórtex a 2500

rpm por 1 min após cada adição; em seguida os tubos foram colocados em água homeotérmica a 85 °C por 4 h, agitando os tubos novamente no decorrer de 1 h. Após a reação de metanólise os tubos foram resfriados à temperatura ambiente e foram adicionadas 2 mL de água MilliQ seguidos de 30 segundos de agitação em vórtex, seguido de 4 h de incubação para que ocorra a separação das fases orgânicas e aquosa. Por fim, a fase aquosa foi descartada e a fase orgânica, contendo os metil ésteres, foi estocada a -20 °C para posterior análise em GCMS-QP2010 Ultra Mass Spectrometer (Shimadzu), equipado com uma coluna Rtx®-5MS, utilizando como padrão interno o ácido benzóico.

Para a análise em GCMS, a temperatura do injetor foi mantida a 250 °C e com início da aplicação na coluna a 100° C por 3 min, com aumento da temperatura de 8° C/min até 210° C, que foi mantida por 15 min. A temperatura do detector foi de 280° C, com um fluxo constante de 0,8 mL/min de gás hélio, utilizado como gás de arraste. Usando como padrão interno o ácido benzóico, e padrões externos PHB e PHBV, com código Sigma 36,350-2 e 40,311-3 respectivamente.

Para quantificação da porcentagem de massa seca de PHB produzido por cada cepa, os resultados das áreas obtidas no GC-MS foram submetidos a 3 fórmulas:

Massa de PHB (g) (M):

$$\frac{A \times 5 \times 10^{-8}}{B} \times 40.000$$

Onde A representa o valor da área de PHB e B representa a área do ácido benzóico.

Massa de PHB ajustando com o rendimento dos padrões internos (R):

$$\frac{M \times 100}{21,5}$$

Onde M representa o valor da massa de PHB (g).

PHB produzido (%) em relação massa seca celular (MSC) (P):

$$\frac{R}{0,01} \times 100$$

Onde R representa a massa de PHB ajustado com o rendimento dos padrões internos.

4.6.1 Produção dos filmes de PHB

Para a produção de filmes de PHB, as células liofilizadas foram pesadas e dissolvidas em clorofórmio, na proporção 1g:20 mL, durante 48 horas, a 60 °C e 200 rpm de agitação. Os detritos celulares em suspensão foram removidos por ultrafiltração em filtro PTFE 0,2 µm e a fase líquida obtida foi depositada em uma superfície de vidro. A partir disso, o filme obtido foi dissolvido em uma pequena quantidade de clorofórmio e precipitado com adição de metanol gelado a -20 °C. O polímero precipitado foi recuperado, embrulhado em papel alumínio e submetido a uma prensa hidráulica MA098/A (Marconi) com pressão de 1 tonelada por 1 minuto, com 1 dia de resfriamento a temperatura ambiente para formação dos filmes padronizados.

5 RESULTADOS

5.1 Experimento de ALE em meio mínimo com sacarose

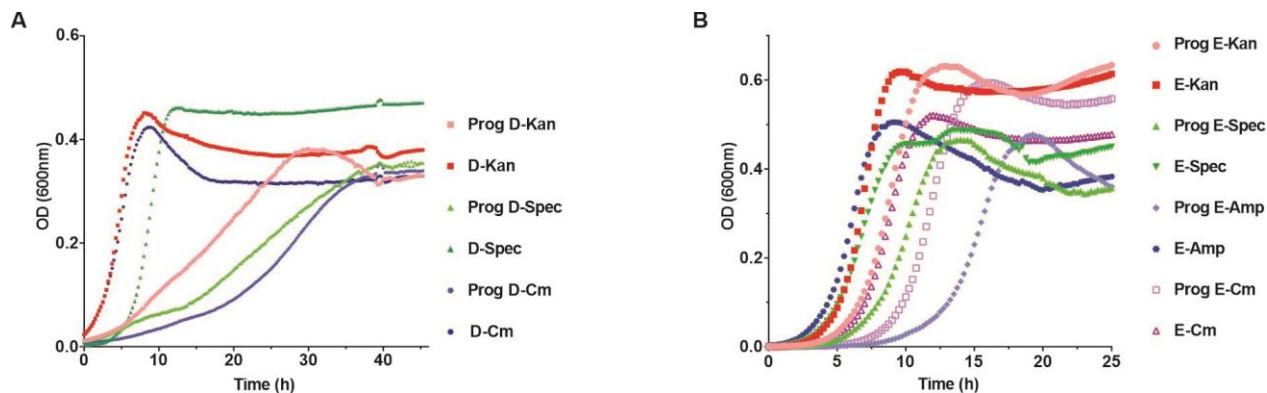
A maioria das cepas laboratoriais e industriais de *E. coli* não metabolizam sacarose de forma natural, sendo necessárias etapas de modificações genéticas para que essa bactéria utilize esse açúcar. Apesar disso, a capacidade de crescimento das linhagens geneticamente modificadas em sacarose não se assemelha ao desempenho em meio rico ou até mesmo em meio mínimo com glicose (MOHAMED; MUNDHADA; LANDBERG; CANN *et al.*, 2019). Considerando isso, a obtenção de cepas com capacidade de crescimento superior em sacarose utilizando ALE é uma possibilidade promissora. Para isso, foram fundadas populações com as cepas de *E. coli* DH5 α e E2348/69 e essas foram usadas no experimento de evolução. Após isso, foram realizadas análises fenotípicas de todas as populações finais, além do sequenciamento do DNA genômico.

5.1.4 Avaliação das taxas de crescimento das populações finais do experimento de Evolução

Para avaliar o efeito do processo evolutivo, testes fenotípicos para a caracterização das populações finais em relação às iniciais, foram realizados. O primeiro teste realizado foi o crescimento em meio mínimo, com a sacarose como única fonte de carbono, em leitor de microplacas (Tecan, Sunrise), sendo em seguida calculada a taxa específica máxima de crescimento ($\mu_{\max}$) nessas condições, tanto da população progenitora quanto das populações finais. As curvas de crescimento para as populações de DH5 α e E2348/69 estão presentes nas Figuras 2 A e B, respectivamente. Observa-se uma grande diferença de tempo para alcançar a fase estacionária dos progenitores em relação aos evoluídos, nas condições de crescimento do aparelho usado. Essa diferença nas curvas de crescimento também é refletida no $\mu_{\max}$ presente no gráfico dos painéis

Figura 3 A, nas quais é possível observar a diferença na velocidade de crescimento em meio mínimo.

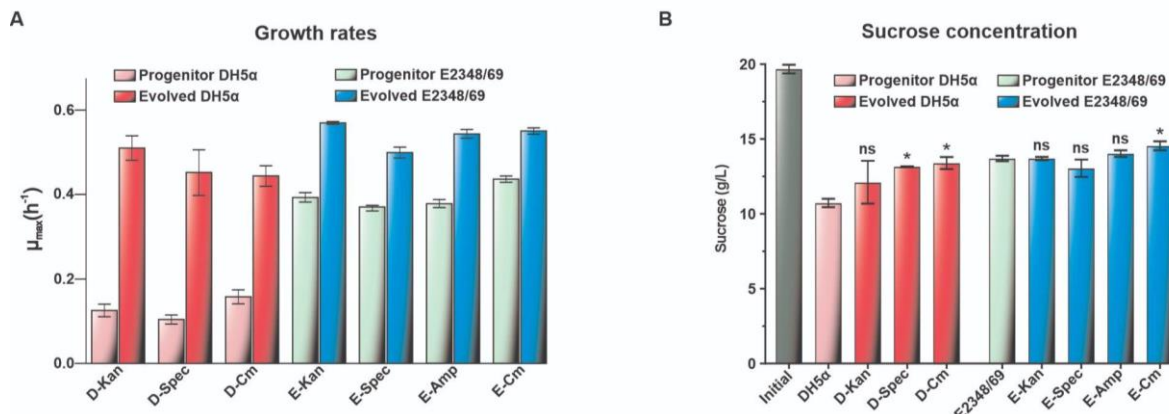
Figura 2: Análise de crescimento em microplaca das populações progenitoras (Prog) e evoluídas de DH5 α e E2348/69.



A Curvas de crescimento para as populações DH5 α . **(B)** curvas de crescimento para as populações E2348/69.

Para verificar se o processo evolutivo teve alguma influência direta no consumo de sacarose (Figura 3 B), foi realizada a quantificação da sacarose ao final de um crescimento de 24 h. Com os valores da quantificação pôde-se observar que não houve diferença significativa no consumo do açúcar do progenitor e das populações evoluídas de E2348/69. Já para as populações evoluídas de DH5 α houve uma redução no consumo do açúcar em comparação com a do plasmídeo pCscBKA. Isso pode sugerir que o direcionamento da evolução nas condições realizadas não foi focado no consumo de açúcar.

Figura 3 - Velocidade máxima de crescimento e a sacarose residual após o crescimento em populações evoluídas e progenitoras



A Velocidade máxima de crescimento (μ_{max}) para as populações DH5α e E2348/69. O teste foi realizado em microplacas de 96 poços, com o progenitor e as respectivas populações evoluídas. **B** Consumo de sacarose em populações evoluídas e progenitoras. A concentração inicial de sacarose foi de 20 g/L. DH5α foi transformado com o plasmídeo pCscBKA. A significância estatística foi avaliada em relação ao respectivo progenitor, ns: diferença não significativa e (*) diferença significativa ($p < 0,05$).

5.2.1 Mutações encontradas nas populações evoluídas de DH5α

O experimento de evolução com a DH5α foi realizado utilizando dois plasmídeos em cada população: o plasmídeo para o operon *cscBKA* e o respectivo plasmídeo para resistência a antibióticos (pKan, pSpec e pGGA). Nas populações finais D-Kan e D-Cm a cobertura de *reads* de sequências *Illumina* para o plasmídeo pCscBKA é igual a encontrada para o DNA genômico, sugerindo que estes plasmídeos se inseriram no genoma. A inspeção das *reads* revelou que o plasmídeo pCscBKA se integrou em regiões de sequências de inserção (IS).

Entretanto, para a população final de D-Spec não houve integração. Nesta população a cobertura de *reads* de sequência *Illumina* nos plasmídeos pCscBKA e pSpec foi menor do que dos vetores mínimos das populações D-Kan e D-Cm (Figura C). Dessa forma, especulamos sobre a possibilidade de que algumas mutações influenciaram o número de cópias desses plasmídeos, o que parece ser o caso da mutação no gene *pcnB* (com deleção de 81 pb em 100% dessa população). Esse gene codifica a poly(A)

polymerase (PAP), responsável pela montagem da cauda *poli-A* em mRNAs de *E. coli*. Foi relatado que a sua completa inativação está relacionada à redução do número de cópias de plasmídeos, além de aumentar a meia-vida de mRNAs (LOPILATO; BORTNER; BECKWITH, 1986; NADRATOWSKA-WESOŁOWSKA; SLOMINSKA-WOJEWODZKA; LYZEN; WĘGRZYN *et al.*, 2010). Dessa forma, pode-se verificar que o sentido de algumas adaptações nas células de DH5 α foi a redução do número de cópias do plasmídeo pCscBKA.

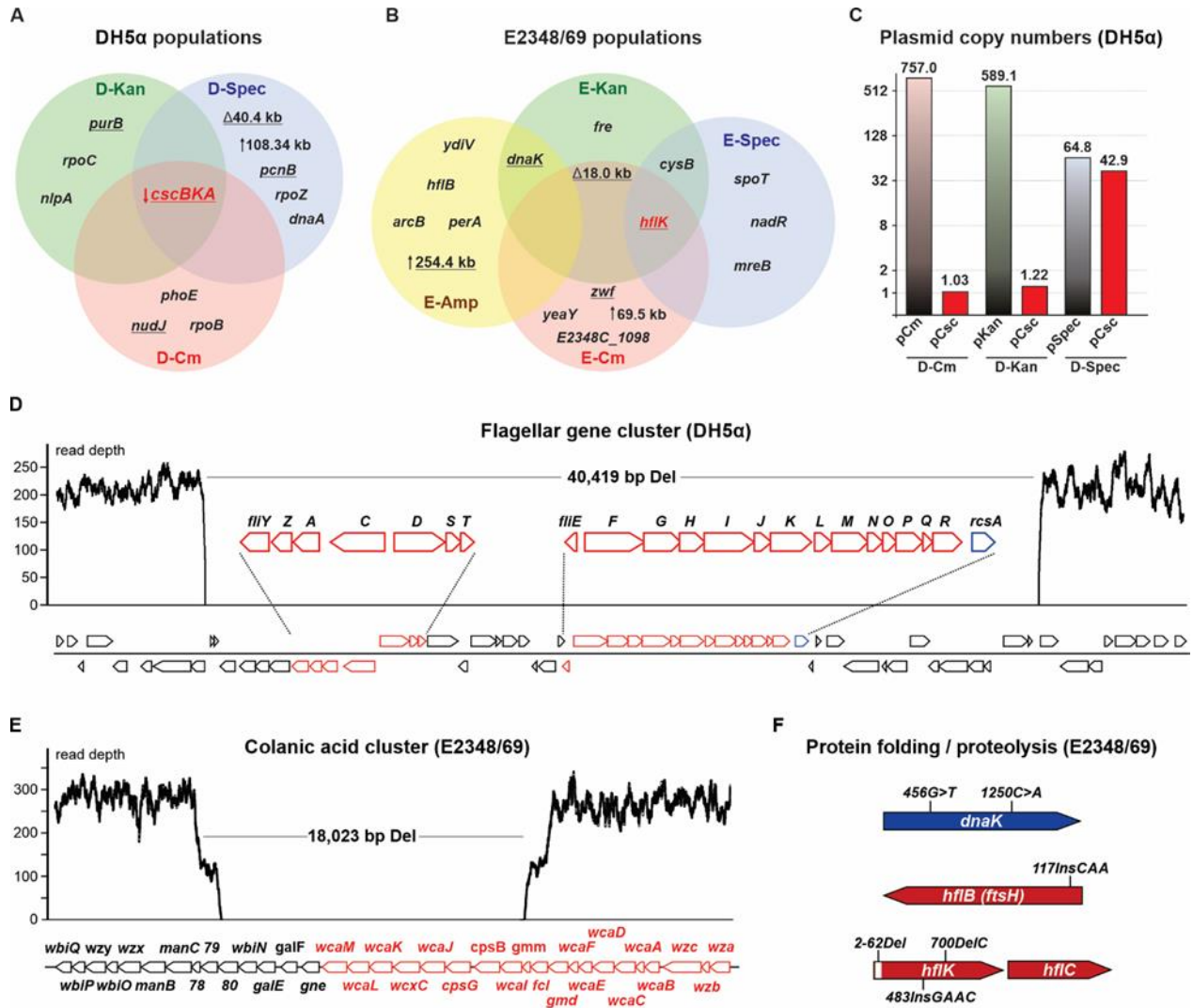
Outras mutações com paralelismo funcional em todas as populações de DH5 α referem-se à transcrição. Observou-se que havia mutações nas três populações, em subunidades diferentes da RNA polimerase: para a população D-Kan subunidade beta', D-Spec subunidade ômega e, para D-Cm beta. Mutações nas subunidades beta e beta' da RNA polimerase de *E. coli* já foram relatadas como adaptação para meios mínimos (CONRAD; FRAZIER; JOYCE; CHO *et al.*, 2010; MOHAMED; MUNDHADA; LANDBERG; CANN *et al.*, 2019). No caso da beta', por exemplo, mutações em uma região próxima ao local da variante encontrada neste estudo pode aumentar a velocidade de transcrição do complexo RNA polimerase (CONRAD; FRAZIER; JOYCE; CHO *et al.*, 2010). No caso da população D-Spec, foi encontrado um polimorfismo na subunidade ômega da RNA polimerase. Polimorfismos nessa subunidade não são comuns e a sua deleção não afeta o crescimento da célula bacteriana (VRENTAS; GAAL; ROSS; EBRIGHT *et al.*, 2005). Entretanto, em estudos mais recentes, quando as células são postas para crescer em meio de cultura o desenvolvimento das células com mutações nesse gene têm alterações (YAMAMOTO; YAMANAKA; SHIMADA; SARKAR *et al.*, 2018). Apesar dessas informações, a função dessa subunidade ainda não parece bem clara, sugerindo um possível papel regulatório no complexo da RNA polimerase.

Além das mutações já citadas, a população D-Kan apresentou mutações nos genes *nlpA* e *purB*. No caso de *purB*, a mesma mutação (Lys115Glu) foi demonstrada por Jung e colaboradores (JUNG; SMITH; LEE; HONG *et al.*, 2010) como restauradora do fenótipo de crescimento em meio mínimo da DH5 α , quando comparado a outras cepas. O gene *nlpA* é responsável por expressar uma lipoproteína que estabiliza a formação de vesículas da membrana externa em determinadas condições, como, por exemplo, na presença de antibióticos e meios de crescimento não tradicionais; a

formação dessas vesículas está relacionada à extrusão de moléculas nocivas ao meio extracelular, além de envolvimento na comunicação entre as células no meio de cultura (SCHWECHHEIMER; KULP; KUEHN, 2014; YAMAGUCHI; INOUE, 1988).

Na população D-Spec houve mais duas mutações a serem consideradas. A primeira dessas foi na região codificadora para *dnaA*, o qual especifica uma proteína iniciadora de replicação cromossômica de *E. coli* (KATAYAMA; KASHO; KAWAKAMI, 2017). Zheng e colaboradores (2001)(ZHENG; LI; SKARSTAD; CROOKE, 2001) relatam que mutações nesta proteína podem estar relacionados a restauração de fenótipo de crescimento em bactérias com deficiência em fosfolípidios ácidos. A *dnaA* também regula a replicação de plasmídeos ColE, como o pCscBKA, agindo no sítio OriC de iniciação. Portanto, a mutação no gene *dnaA* também pode estar causando diminuição no número de cópias do pCscBKA nesta população. Outra mutação encontrada na população D-Spec foi a deleção de segmento cromossômico de 40.418 pb (Figura4 C), especificando grupos de proteínas envolvidas na montagem de flagelo. Para a população D-Cm também houve duas mutações no gene *nudJ*, responsável pela codificação de uma enzima do grupo das NUDIX hidrolases, um grupo de enzimas com atividade pirofosforolítica, envolvida nas vias de diversos nucleosídeos difosfato (BESSMAN, 2019; LEI; QIU; WU; QIAO *et al.*, 2021).

Figura 4- Resumo das mutações encontradas na análise de sequenciamento genômico de populações evoluídas.



(A) e (B) Diagramas de Venn mostrando mutações comuns entre populações. As setas para cima indicam duplicações e as setas para baixo indicam uma redução no número de cópias. (C) Número de cópias de plasmídeos em relação ao genoma das populações finais de DH5 α com base no número de leituras de plasmídeos em comparação com o genoma. (D) Deleção de 40,4 kb na população D-Spec com destaques em vermelho para genes relacionados à montagem do flagelo. Destaque também em azul para o gene *rcsA*, regulador de transcrição do ácido colânico (CA). (E) Deleção de parte da região responsável pela expressão de CA nas populações E-Kan e E-Cm. (F) Mutações em genes relacionados ao dobramento de proteínas e proteólise em *dnaK* (E-Kan e E-Amp), *hflB* (E-Amp) e *hflK* (E-Kan, E-Spec e E-Cm).

5.2.2 Mutações encontradas nas populações evoluídas de E2348/69

Dentre as mutações encontradas nas populações de E2348/69 algumas merecem destaque. Para o *dnaK*, foram encontrados poliformismos em duas populações (E-Kan e E-Amp). Esse gene codifica a chaperona mais representativa do grupo das Hsp70 em *E. coli*. A expressão desse gene está relacionada a situações de estresse celular, relacionando-se com aproximadamente 700 proteínas na bactéria (CALLONI; CHEN; SCHERMANN; CHANG *et al.*, 2012).

Outro gene que aparece mutado ao mesmo tempo em diferentes populações é o *hflK* (E-Kan, E-Spec e E-Cm). Esse gene é relacionado ao processo lisogênico do fago lambda (KIHARA; AKIYAMA; ITO, 1996). Outro gene que pode ser relacionado à atividade lisogênica é o *hflB* (E-Amp). HflB é uma metaloprotease inibida pelo complexo formado por HflK e HflC (HflKC). Essas proteínas influenciam se a célula infectada por fagos vai seguir o ciclo lítico ou lisogênico, sendo que mutações defetivas em *hflK* e *hflB* estão relacionadas a promoção da lisogenia, ou seja, a manutenção do profago inativo no genoma (BANDYOPADHYAY; PARUA; DATTA; PARRACK, 2011; HERMAN; THEVENET; D'ARI; BOULOC, 1997).

Também foram encontradas mutações relacionadas ao gene *cysB* em duas populações (E-Kan e E-Spec). Essas duas mutações, diferentemente das anteriores citadas, situavam-se na região promotora deste gene. O produto de *cysB* participa da regulação da via de síntese de cisteína (LILIC; JOVANOVIĆ; JOVANOVIĆ; SAVIC, 2003).

Nas populações E-Spec e E-Amp também houve duas mutações interessantes, nos genes *relA* e *ydiV* respectivamente. *relA*, juntamente com *spoT*, é responsável pela síntese de guanosina tetra e penta fosfato ((p)ppGpp), a qual atua como uma alarmona, influenciando na regulação da resposta estrigente em situações de stress nutricional (TRAXLER; SUMMERS; NGUYEN; ZACHARIA *et al.*, 2008). Um fato interessante é que *relA* localiza-se no mesmo operon que o *rpoZ* (SPIRA; OSPINO, 2020). Esse gene está mutado na população final de DH5alpha D-Spec. *ydiV* tem relação a expressão das proteínas formadoras do flagelo em baixas concentrações de nutrientes (WADA; HATAMOTO; KUTSUKAKE, 2012), o que é interessante em virtude de na população

DH5alpha D-Spec o conjunto de genes responsáveis pela montagem do flagelo terem sido deletados.

Os genes *zwf* e *nadR* foram encontrados mutados em apenas uma população cada (E-Cm e E-Spec, respectivamente). *zwf* codifica a glucose-6-fosfato-desidrogenase, a qual está na via das pentoses fosfato (ZHAO; BABA; MORI; SHIMIZU, 2004). *nadR* expressa um regulador negativo que também age como controlador do transporte externo de nicotinamida mononucleotídeo, imediato precursor de NAD (RAFFAELLI; LORENZI; MARIANI; EMANUELLI *et al.*, 1999).

5.3 Estabelecimento de técnicas de genética molecular e edição genômica

Quando é proposto o desenvolvimento de uma pesquisa cujo objetivo seja a produção de um microrganismo como plataforma para síntese de um ou mais bioprodutos ou para o teste de mutações encontradas em um experimento de evolução, o desenvolvimento de técnicas de genética molecular de fácil e rápida utilização é de importância crucial para se chegar a uma cepa superior para a produção industrial. Da mesma forma, estudos de evolução adaptativa requerem ferramentas de modificação genética que facilitem a engenharia reversa de alelos evoluídos para quantificação da adaptação conferida por essas mutações. Assim, um dos focos iniciais deste doutorado foi o desenvolvimento de técnicas de edição genômica e montagem múltipla de fragmentos DNA na estrutura desejada.

5.3.1 Técnica de CRISPR/Cas9

A técnica de CRISPR/Cas9 é uma das mais versáteis e promissoras estratégias para modificações genéticas em praticamente qualquer local do genoma, tanto de eucariotos como de procariotos. Esse recurso de edição genômica pode ser usado para inserção de sequências heterólogas, para a deleção de regiões do genoma do organismo (genes ou sequências reguladoras), além de permitir a mudança individual de pares de bases. Todas essas modificações são realizadas de forma precisa e específica, tornando CRISPR/Cas9 uma técnica robusta para ser usada para qualquer aplicação na qual

deseja-se fazer modificações genômicas (JIANG; CHEN; DUAN; SUN *et al.*, 2015; RONDA; PEDERSEN; SOMMER; NIELSEN, 2016).

A Cas9 é uma nuclease que foi isolada para fins biotecnológicos da bactéria *Streptococcus pyogenes*. Essa enzima promove uma quebra na dupla fita de DNA em um local específico. Para a definição desse ponto de clivagem, a Cas9 utiliza uma pequena molécula de RNA para guiá-la até o local da clivagem. Essa molécula de RNA é chamada de gRNA (RNA-guia) e é composta de 20 pares de bases complementares na região do genoma em que se deseja realizar a quebra de dupla fita do DNA. Para que uma sequência do genoma seja usada como gRNA é necessário que a mesma esteja adjacente a uma sequência PAM. No caso da Cas9 de *S. pyogenes*, as sequências PAM têm o formato NGG (sendo N qualquer nucleotídeo). Quando a sequência PAM é alterada por edição genômica, por exemplo, o gRNA não funcionará para a clivagem do genoma naquele ponto (BASSALO; GARST; HALWEG-EDWARDS; GRAU *et al.*, 2016; RONDA; PEDERSEN; SOMMER; NIELSEN, 2016).

Após a clivagem do DNA genômico pela Cas9, a célula pode seguir dois caminhos. Se a quebra na dupla fita não for reparada a célula morre. Ou então, o dano ao DNA pode ser reparado por mecanismos como “non-homologous end joining”, ou por recombinação homóloga. É pela via da recombinação homóloga em que é possível inserir o fragmento *donor* destinado a fazer a modificação genômica. Desta forma, o reparo da quebra de dupla fita do DNA é feito e pelo fragmento *donor* e a célula com a edição do DNA sobreviverá (BASSALO; GARST; HALWEG-EDWARDS; GRAU *et al.*, 2016).

Em tratando-se da execução da técnica de CRISPR/Cas9 em *E. coli* especificamente, além da Cas9, do gRNA e do fragmento *donor* de DNA, é necessário o uso do sistema Lambda Red para realizar a recombinação do fragmento de DNA a ser inserido (JIANG; CHEN; DUAN; SUN *et al.*, 2015). Isso ocorre porque, diferentemente de outros organismos, o mecanismo intrínseco de recombinação homóloga dessa bactéria tem uma baixa eficiência. O sistema Lambda Red é composto por três enzimas: Exo, Beta e Gama que, atuando em conjunto, aumentam a efetividade da recombinação da molécula de DNA para a alteração genômica (BASSALO; GARST; HALWEG-EDWARDS; GRAU *et al.*, 2016; JIANG; CHEN; DUAN; SUN *et al.*, 2015). Para a execução desse

projeto, foi adquirido o plasmídeo CAS9BAC1P-1EA (Sigma Aldrich), que expressa a SpCas9 e o sistema Lambda Red.

5.3.2 Otimização da clonagem de gRNAs na técnica de CRISPR/Cas9

O gRNA direciona o local no genoma onde a Cas9 promove o corte no DNA. Assim, em virtude da necessidade de serem realizadas diversas modificações genômicas em *E. coli*, um sistema simples para a clonagem de um plasmídeo para expressar essa molécula apresenta grande importância.

Inicialmente, foi desenvolvida uma técnica para clonagem de plasmídeos para expressar o gRNA. Esta é baseada na recombinação *in vivo* em *E. coli*, conforme previamente descrita em nosso grupo (JACOBUS; BARRETO; DE BEM; MENEGON *et al.*, 2022; JACOBUS; GROSS, 2015). A primeira abordagem foi a construção do plasmídeo pTargetAmp (Figura 5 A). A técnica se baseia na amplificação do plasmídeo pTargetAmp, contendo o gRNA, em um único PCR usando um primer *forward* (que se liga ao gRNA) e outro reverso (que liga no promotor pJ23119) (Figura 5 A). Ambos os primers *forward* e reverso têm caudas 5' especificando a sequência alvo de 20 pb de homologia (i.e., o *spacer*). O fragmento de PCR de 1889 pb (Figura 5 D, PCR1), gerado com estes primers, após tratamento com DpnI para eliminar o plasmídeo *template*, é transformado por choque térmico em células de DH5 α , para o mesmo se fechar via recombinação *in vivo* entre as sequências homólogas *spacer* de 20 pb. Uma vez confirmada a clonagem, o plasmídeo do gRNA é purificado e posteriormente usado na técnica de CRISPR/Cas9.

Embora esta abordagem inicial tenha trazido vários resultados promissores, para alguns gRNAs essa estratégia clonagem não foi efetiva; após a transformação, as colônias testadas por PCR não eram positivas para a clonagem do gRNA. Diversas alternativas foram tentadas para solucionar este problema, como aumentar a concentração de DpnI, reduzir a quantidade de plasmídeo *template*, dentre outras, sem o sucesso desejado. Com isso, foi demonstrado que esse sistema não era otimizado para os fins pretendidos.

Para otimização do método, uma nova estratégia foi idealizada para diminuir (ou eliminar) falsos positivos na clonagem do gRNA. Na nova abordagem, o crescimento de colônias só ocorre quando o plasmídeo do gRNA é fechado da forma correta a partir de dois fragmentos. A ideia encontrada foi a divisão do gene especificando resistência a espectinomicina em duas partes, sendo a região 5' no plasmídeo pSpec5' e a região 3' em outro plasmídeo chamado pSpec3' (Figura 5 B). Portanto, para a construção do plasmídeo para expressão do gRNA dois PCRs são necessários. No primeiro PCR (primers P1 e P4), a partir do template pSpec3' é amplificado um fragmento de 1.114 pb com o promotor do gRNA (pJ23119), origem de replicação do plasmídeo e a região 3' do gene de resistência a espectinomicina (Figura 5 B e D, PCR3). Em um segundo PCR (primers P3 e P2), a partir do plasmídeo pSpec5', a outra parte 5' do gene de resistência à espectinomicina e o gRNA-*scaffold* são amplificados em um fragmento de 882 pb (Figura 5 B e D, PCR2). Os primers que são usados e que anelam no gRNA-*scaffold* (P2) e no promotor J23119 (P1), possuem cada um uma cauda de 20 pb homólogas entre si que servem para especificar a sequência *spacer* e nuclear uma recombinação entre os dois produtos de PCR. A recombinação também ocorre na outra extremidade, em uma região de 92 pb de homologia entre os fragmentos 5' e 3', os quais são fusionados gerando uma fase de leitura completa especificando a resistência à espectinomicina. Após confirmação da amplificação por eletroforese, os dois fragmentos de PCR amplificados são misturados e cotransformados em células *E. coli* DH5 α , as quais são plaqueadas em meio sólido contendo espectinomicina. O crescimento de colônias é indicativo da formação de um plasmídeo funcional a partir dos dois produtos de PCR por recombinação *in vivo* em *E. coli* (JACOBUS; GROSS, 2015).

As vantagens dessa nova estratégia residem principalmente no fato de que colônias são formadas apenas se a montagem do plasmídeo de gRNA e da marca de resistência à espectinomicina estiverem corretas. Isso ocorre porque tanto o pSpec3' e o pSpec5' não possuem resistência à espectinomicina, apenas parte da região codificadora em cada um. Outras vantagens são a não necessidade de tratamento por DpnI (ou qualquer outra enzima de restrição) e a purificação de fragmentos, sendo transformados diretamente os produtos de reação de PCR. Esse sistema demonstrou uma eficiência de

100% na clonagem de plasmídeos para expressão de gRNA em todos os testes realizados.

Para otimizar a técnica de clonagem do plasmídeo de gRNA foi elaborada uma terceira estratégia com base no uso de oligonucleotídeos, oligo PCR cloning, ou abordagem EasyOligo (Figura 5 C). Nesse caso, foram clonados dois plasmídeos, pOliSec5' e pOliSpec3'. No primeiro plasmídeo foi clonada em outro vetor com resistência a Cloranfenicol a região SpecR 5' juntamente com o gRNA-Scaffold. O pOliSpec3' apresenta uma estrutura semelhante ao pSpec3', entretanto entre o promotor e a marca de resistência a ampicilina e o final da marca de ampicilina foram adicionados sítios de restrição para *Bam*HI. Para o uso, o plasmídeo pOliSpec5' é digerido com *Pvu*II e *Pst*I e o plasmídeo pOliSpec3', digerido com *Bam*HI. Após purificação em coluna, esses fragmentos são amplificados com os primers P3/P5 (pOliSec5') e P4/P6 (pOliSec3'). Nesse caso, nenhum dos primers que amplificaram as partes do *backbone* possuem a caudas 5' (Figura 5 D, PCRs 4 e 5).

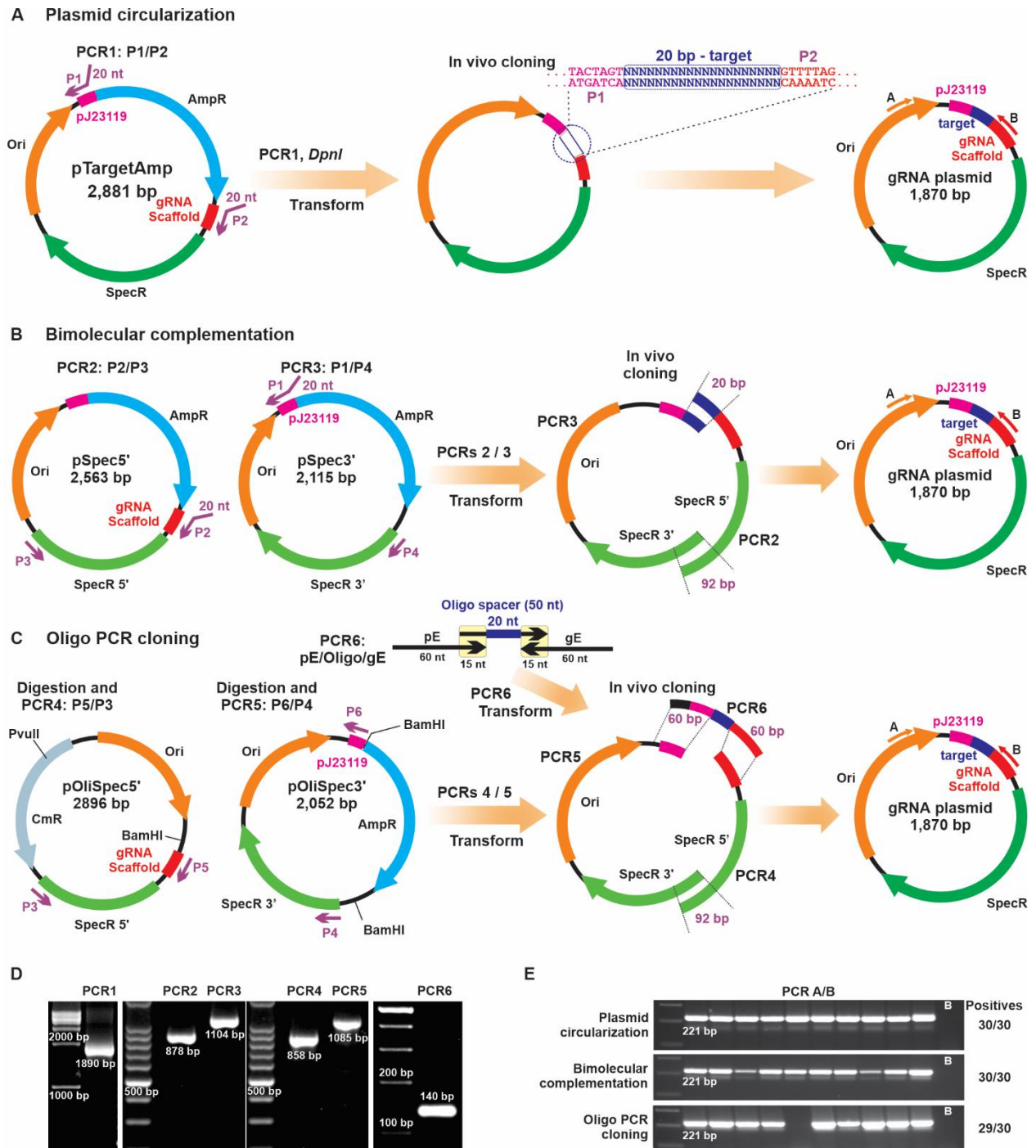
A sequência do *spacer* é desenhada em um oligonucleotídeo central (oligo-*Spacer*), o qual é amplificado por PCR pelos primers pE e gE. Após a amplificação, o produto (PCR6, Figura 5 C e D) é transformado juntamente com os fragmentos do *backbone* plasmidial (PCRs 4 e 5). A recombinação *in vivo* envolvendo as regiões homólogas monta o *spacer* corretamente entre o promotor e o gRNA *scaffold*. A seleção para espectinomomicina é indicativa que a montagem uniu as duas partes do *backbone* e fechou um plasmídeo *in vivo*.

Em um primeiro momento, a utilização do sistema EasyOligo para clonagem de gRNAs parece ser desvantajosa em comparação ao sistema de complementação bimolecular. Entretanto, a ideia é que os fragmentos do *backbone* sejam produzidos em grandes quantidades, purificados e armazenados na geladeira para diversos usos. Dessa forma, para a utilização desse sistema, é necessário apenas a realização da PCR do oligonucleotídeo *spacer* (PCR 6). Esse produto é misturado com os fragmentos previamente produzidos, sendo o mix transformado.

Um teste comparativo entre os três sistemas de clonagem de gRNAs foi realizado com clonagens do gRNA *galK*, como demonstrado na Figura 5 E. Todos os sistemas foram efetivos na clonagem com alta eficiência dos plasmídeos de gRNA. Todos os cinco

plasmídeos do sistema EasyGuide para *E. coli* que foram gerados neste trabalho estão disponíveis na Addgene (<https://www.addgene.org/>), com os números de catálogo 228548 a 228552. Figuras detalhadas dos mapas desses plasmídeos estão disponíveis no Apêndice C.

Figura 5- Diferentes estratégias para produção de plasmídeos para expressão de gRNA.



A Plasmid circularization. Utilizando o plasmídeo pTargetAmp como modelo, a amplificação por PCR é realizada com os iniciadores P1 e P2 (PCR 1). As extremidades 5' desses primers são complementares e especificam o *spacer* de gRNA de 20 nts. Após o tratamento com DpnI, o produto de PCR é diretamente transformado e ligado *in vivo* em *E. coli*. **B** Bimolecular complementation. Os fragmentos de PCR são

amplificados a partir do plasmídeo pSpec5' com primers P2/P3 (PCR2) e do plasmídeo pSpec3' com primers P1/P4 (PCR3). Estes fragmentos de PCR são misturados e transformados diretamente em *E. coli* e o plasmídeo gRNA é montado *in vivo* por complementação das duas partes. **C** Oligo PCR cloning. Antes da sua utilização, os plasmídeos pOliSpec5' e pOliSpec3' são digeridos com *Bam*HI-*Pvu*II e *Bam*HI, respectivamente. Após a purificação, as estruturas digeridas pOliSpec5' e pOliSpec3' são amplificadas por PCR com combinações de primers P5/P3 e P4/P6, respectivamente. Para gerar a região gRNA, o oligo-*spacer* é amplificado com os primers pE e gE. O produto de PCR de 140 pb resultante é misturado com os dois fragmentos da estrutura do plasmídeo e transformado em *E. coli* para montagem *in vivo* do plasmídeo gRNA. **D** Eletroforese em géis de agarose a 1% (PCR 1), 1,5% (PCRs 2-5) e 3% (PCR 6) mostrando os produtos de PCR necessários para clonagem dos plasmídeos de gRNA de acordo com as três estratégias apresentadas. Os comprimentos dos produtos de PCR são indicados e comparados com marcadores de peso molecular (marcador de peso molecular de 1kb para PCR1 e escada de 100 pb para PCRs 2-6). **E** Eletroforese em gel de agarose 2% mostrando produtos de PCR de 221 pb com primers A e B confirmando a clonagem de gRNA *galK* usando as três estratégias apresentadas. Triplicatas de 10 colônias cada foram testadas para cada estratégia de clonagem. Taxas de confirmação positivas e estão indicadas no lado direito. *Ladder* de 100 pb foi utilizada como marcador de peso molecular. Os detalhes das clonagens genômicas estão descritos nos Apêndices A e B.

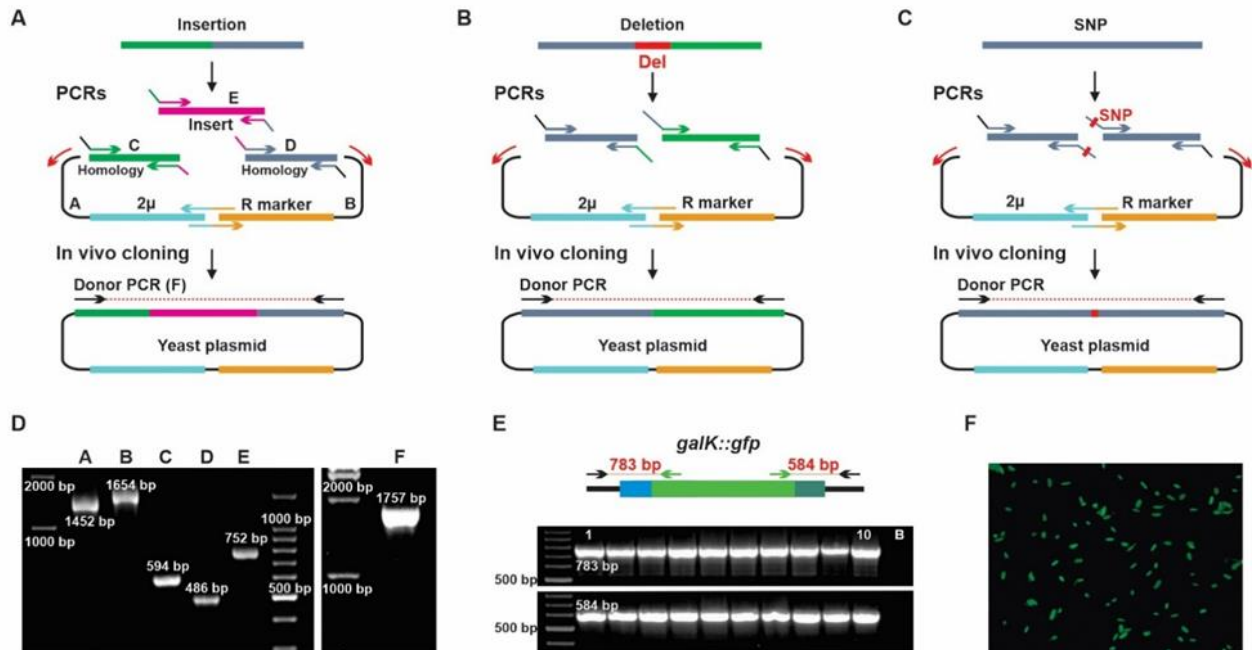
5.3.3 Método de clonagem em *Saccharomyces cerevisiae* de DNAs *donors* para uso em edições CRISPR/Cas9 em *E. coli*

A capacidade de montagem orientada de fragmentos de DNA em plasmídeos, de forma rápida e eficiente, é de grande importância para diversas aplicações em biologia sintética. Além disso, ao se realizar pesquisas com experimentos de evolução é necessário dispor de técnicas eficientes e de baixo custo para realização do processo de engenharia reversa, de suma importância para a testagem de possíveis mutações que possam contribuir para a construção de cepas superiores. No caso de edições genômicas com CRISPR/Cas9 em *E. coli*, os fragmentos para fazer modificações genéticas (i.e., sequências *donors*) precisam ter regiões de homologia maiores do que aquelas usadas em leveduras para aumentar a eficiência das modificações genômicas. O tamanho dos braços de homologia com o genoma recomendados são de 400 pb em cada lado para maximização de efetividade (BASSALO; GARST; HALWEG-EDWARDS; GRAU *et al.*, 2016; JIANG; CHEN; DUAN; SUN *et al.*, 2015).

Dessa forma, para a realização de inserções genômicas em *E. coli* pela técnica de CRISPR/Cas9, são necessárias pré-clonagens para montagem adequada dos fragmentos *donors*, que em seguida são amplificados e usados na edição CRISPR. Para a finalidade de ligar vários fragmentos de DNA a abordagem com melhores resultados por nós testada foi a transformação em *S. cerevisiae*, devido à alta eficiência de recombinação homóloga pelo sistema Rad51 da levedura (SHINOHARA; OGAWA; OGAWA, 1992). Assim, tem-se a possibilidade de dispensar uso de kits e enzimas de modificação de DNA porque levedura tem a capacidade de ligar vários fragmentos de DNA de forma natural (BASSALO; GARST; HALWEG-EDWARDS; GRAU *et al.*, 2016).

A estratégia proposta aqui explora o sistema de plasmídeos *EasyGuide* proposto por Jacobus e colaboradores (JACOBUS; BARRETO; DE BEM; MENEGON *et al.*, 2022). Essa estratégia consiste, inicialmente, na amplificação, por PCR, da origem de replicação para levedura (2 micron), marca de resistência a antibiótico (ambos a partir de plasmídeos que não se replicam em *S. cerevisiae*) e dos fragmentos que especificam a mutação a ser realizada em *E. coli* (Figura 6). Esses produtos de PCR são transformados (sem purificar) em *S. cerevisiae*. Após seleção no antibiótico, colônias são testadas por PCR para verificar se a clonagem dos fragmentos foi realizada de forma correta, sendo postas para crescer. A realização da PCR para obtenção do fragmento a ser usado como *donor* em *E. coli* pode ser feita a partir da própria colônia da transformação. Para otimização do PCR, frequentemente optamos por uma extração de DNA total de alta qualidade de culturas celulares, ou ainda a purificação do plasmídeo adaptando protocolos de kit de purificação plasmidial de bactérias (QIAprep Spin Miniprep Kit, QIAGEN). Com essa estratégia, foram montados e amplificados por PCR *donors* para diversas modificações genômicas. Como exemplo da utilização desse sistema de para clonagem de *donors*, pode-se citar a inserção da sequência codificadora de GFP; na Figura 6 é ilustrado todo o processo de PCRs e clonagens e verificação da expressão de *gfp* integrado ao genoma.

Figura 6- Montagem de sequências *donors* por recombinação homóloga *in vivo* de levedura.



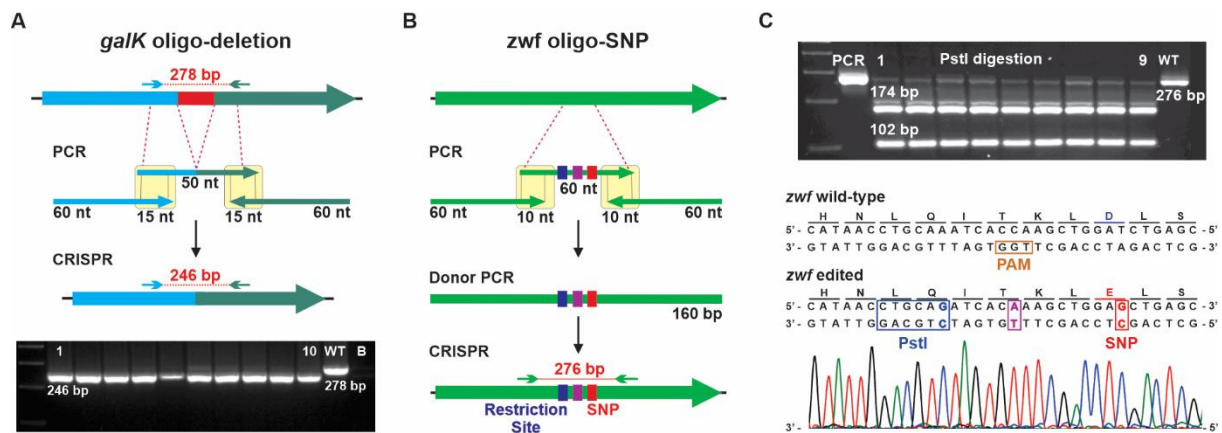
A origem de replicação do plasmídeo de levedura (2μ) e o marcador de resistência a antibióticos (marcador R) são amplificados através de dois PCRs (A e B, respectivamente) da série de plasmídeos pEasyG3. Estas partes são combinadas com PCRs para gerar diferentes doadores contendo plasmídeos (painéis A-C) através de recombinação homóloga em levedura. **(A)** Pré-clonagem de *donor* para inserir um gene no genoma de *E. coli*. As regiões esquerda e direita que flanqueiam o sítio de integração e o gene para inserção são amplificadas através de PCRs C, D e E, respectivamente. Os oligonucleotídeos usados para PCRs especificam regiões de sobreposição entre os fragmentos de PCR adjacentes e com a estrutura do plasmídeo, permitindo a montagem correta das partes em um único plasmídeo na levedura. **(B)** Montagem de *donor* para especificar deleções cromossômicas. As regiões flanqueadoras da deleção são amplificadas por PCR e a deleção é especificada por primers centrais que possuem caudas homólogas. **(C)** Montagem de *donors* para engenharia reversa de mutações de nucleotídeo único. As regiões flanqueadoras de recombinação homóloga são amplificadas por PCR e o SNP é especificado nas partes 5' dos oligonucleotídeos que conectam as regiões flanqueadoras. Os *donors* montados nos painéis A-C são amplificados por PCR e transformados em *E. coli* para integração do genoma mediada por CRISPR/Cas9. **(D)** Eletroforese em gel de agarose a 1,5% mostrando amplificações por PCR de partes (A-E) necessárias para a montagem de *donor* para inserir o gene que codifica GFP no locus *galK* de *E. coli*. Após a clonagem, a amplificação por PCR do *donor* gera um produto de 1.767 pb (PCR F). Ladder de 1 kb e 100 pb foram usadas como marcadores de peso molecular. **E** Confirmação por PCR da integração *galK::gfp* em *E. coli* DH5α. **F** Expressão de GFP em células de *E. coli* *galK::gfp*.

5.3.4 Abordagem EasyOligo para a geração de *Donors*

Apesar da clonagem em plasmídeos em levedura ser uma grande alternativa para a montagem de fragmentos destinados à *donors*, ela ainda demanda um tempo de 2/3 dias para a obtenção de transformantes, mais um dia para extração de DNA e realização do PCR. Para reduzir o tempo, foi estruturada uma estratégia de produção rápida de *donors* com base na amplificação de oligonucleotídeos, semelhante ao gRNA (abordagem EasyOligo).

A primeira concepção foi a deleção de 32 pb do gene *galK* (Figura 7 A). Para a construção do *donor*, a região da deleção é especificada por um oligo central (50 nt), o qual é amplificado por PCR com oligos laterais (60 nt) com 15 nt de anelamento. O produto dessa PCR é diretamente transformado, juntamente com o plasmídeo do gRNA, em células carregando o plasmídeo da Cas9.

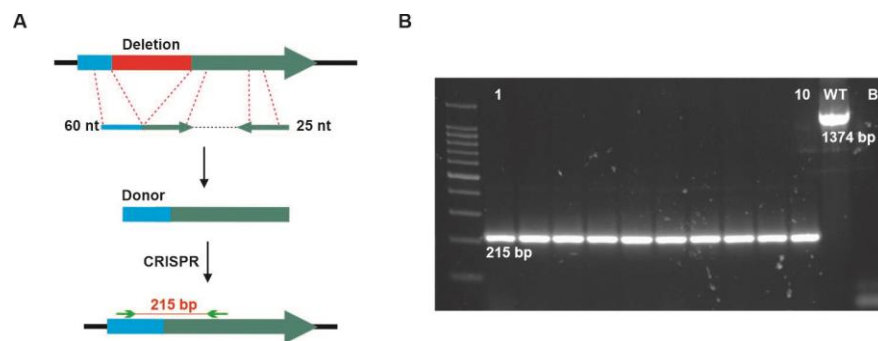
Para a realização de mutações de nucleotídeos únicos esses, estes são especificadas no oligo central o qual é usado para amplificação com os oligos laterais. O exemplo apresentado foi para o SNP no gene *zwf*, cuja mutação foi especificada no Oligo central (60 nt), sendo este amplificado pelos oligos laterais para composição do *donor* para ser usado no experimento de CRISPR-Cas9 (Figura 7 B). Para rastrear as colônias em que ocorreram modificações, foi inserido no *donor* um sítio de restrição para *PstI*, sem mudança do aminoácido produzido (Figura 7 C). Para a interrupção da PAM foi realizada também uma alteração, sem mudar o aminoácido na proteína produzida. Após confirmar a modificação genômica através de PCRs de colônias, foi realizado o sequenciamento Sanger dessa região, verificando que todas as mutações planejadas foram realizadas, apenas com uma troca de aminoácido, conforme desejado.

Figura 7- Produção de *donors* utilizando oligonucleotídeos.

A Geração de um *donor* para produzir uma deleção de 32 pb no gene *galk*. A deleção é especificada por um oligo central de 50 nt, que é usado como modelo para gerar um produto de PCR de 120 pb com oligos direito e esquerdo apresentando sítios de ligação de 15 nts. O *donor* é diretamente transformado em *E. coli* para o experimento CRISPR. A eletroforese em gel de agarose a 2% mostra um fragmento de PCR de 246 pb confirmando a deleção de *galk*, enquanto um produto de PCR de 278 pb surge do locus WT. **B** A edição de SNPs no gene *zwf*. A mutação *zwf* T/G (caixa vermelha), a obliteração da sequência PAM (roxo) e a geração de um sítio de restrição *PstI* (azul) são especificadas no oligo central de 60 nts, que é usado como modelo para a amplificação por PCR de um *donor* de 160 pb, transformado diretamente em *E. coli* para edição do genoma CRISPR/Cas9. **C** Confirmação da edição CRISPR através da digestão *PstI* e sequenciamento Sanger. A edição é positiva quando fragmentos de 174 pb e 102 pb são produzidos através da digestão *PstI* de um produto de PCR de 302 pb. Para o locus WT é produzido um fragmento de 276 pb após digestão. O sequenciamento Sanger confirmou a edição correta dos três SNPs. Tanto a introdução do sítio de restrição *PstI* como a modificação do PAM não alteraram o aminoácido codificado.

Outra forma de estruturação de *donor* que foi usar é usar os braços de homologia assimétricos. Para essa estratégia o *donor* foi amplificado por PCR a partir do genoma com um dos primers apresentando uma cauda na qual é especificada a mutação. Foi feita uma deleção teste do gene *fliA* (Figura 8) e outra do gene *hflK*.

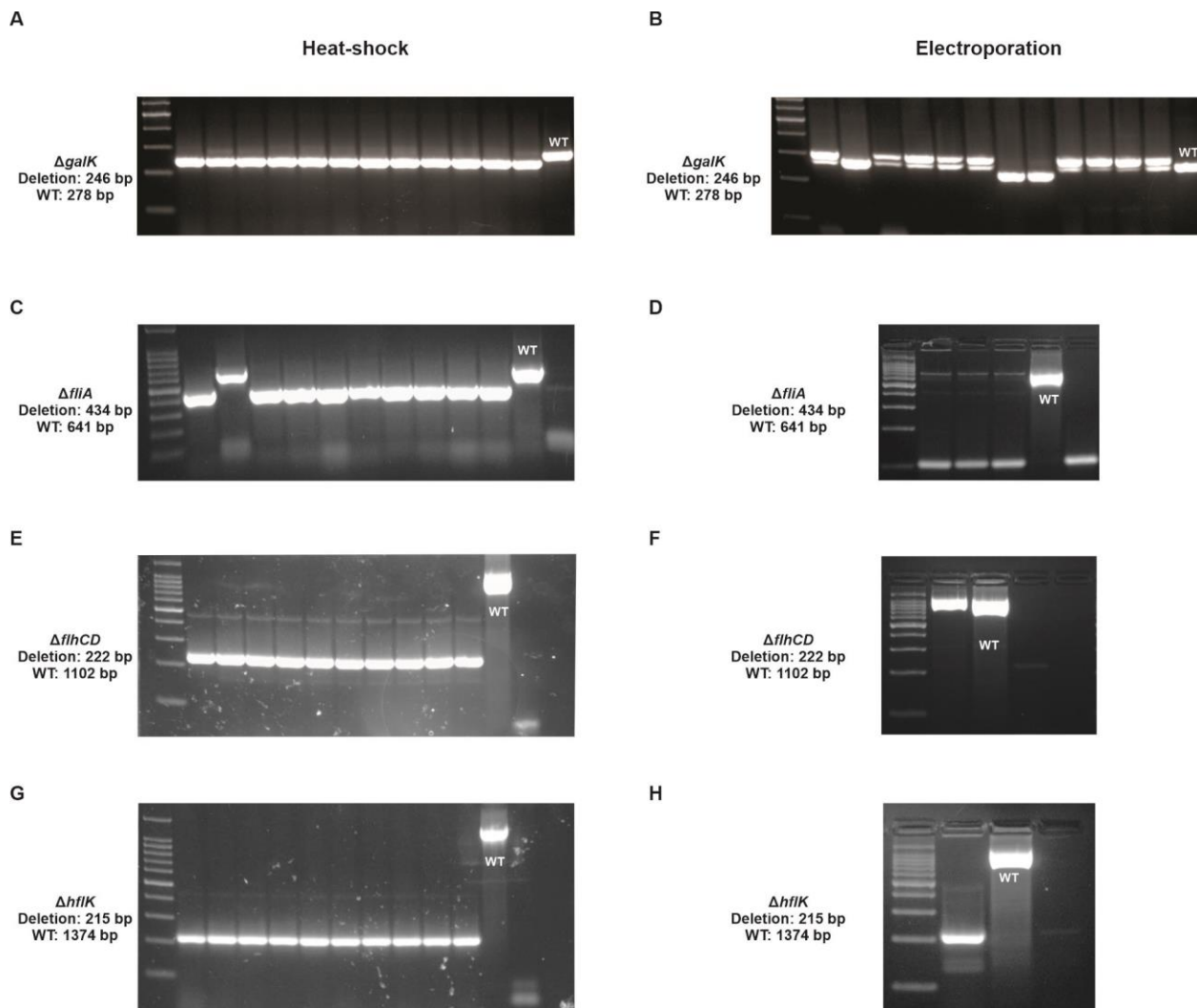
Figura 8- Deleção usando um *donor* com braços de homologia assimétrico.



A Estrutura da deleção de gene com um *donor* com braços de homologia assimétricos. Para amplificar o doador por PCR, um primer especifica a deleção, enquanto um primer de menor alcance é usado. Os braços de homologia, embora desiguais, permitem a recombinação para deletar a região especificada pelo primer mais longo. **B** Eletroforese em gel (agarose 2%) mostrando resultados de PCR da deleção de *hflK* em 10 colônias. Deleções positivas mostram uma banda de 215 pb, enquanto o tipo selvagem (WT) permanece intacto.

A utilização dos oligos como *donors* para realização de deleções e modificações de nucleotídeo único são realizadas com fragmentos de DNA pequenos, contrariando o que foi citado anteriormente. De fato, os primeiros testes realizados principalmente com a deleção de 32 pb do *galK* não foram efetivos para deleção. Entretanto, como para a técnica de *in vivo* cloning usa-se choque térmico para a transformação, resolveu-se testar essa estratégia para a execução da técnica. Assim foi observada uma alta eficiência na modificação genômica usando essa estrutura de *donor* com choque térmico transformando diretamente o produto de PCR. Ao comparar com a eletroporação pôde-se observar uma maior eficiência usando essa estratégia (Figura 9).

Figura 9- Comparação de métodos de transformação para deleções usando choque térmico (A, C, E, G) e eletroporação (B, D, F, H) para os genes *galK*, *fliA*, *flhCD* e *hflK*



Um *donor* oriundo da amplificação de oligos foi usado para deleções de *galK* e *flhCD*, enquanto um *donor* com braços assimétricos foi usado para *fliA* e *hflK*. Nas transformações de eletroporação para *fliA*, *flhCD* e *hflK*, todas as colônias resultantes foram testadas.

O resumo de todas as modificações genômicas com as técnicas usadas nesse estudo está representado na Tabela 2. Também estão presentes as eficiências de modificações genômicas nas transformações. As fotos das respectivas eletroforeses em gel de agarose comprovando essas modificações são apresentados no Apêndice E.

Tabela 2: Modificações genéticas mediadas por CRISPR/Cas9 através da abordagem EasyGuide

Genetic modification	gRNA cloning	Donor assembly	Efficiency
<i>SS9::cscBKA</i>	Plasmid circularization	In vivo cloning in <i>E. coli</i>	10/10
<i>galK::cscBKA</i>	Plasmid circularization	In vivo cloning <i>S. cerevisiae</i>	10/10
<i>cscA::sucP</i>	Bimolecular complementation	In vivo cloning <i>S. cerevisiae</i>	4/4
<i>cscA::pJ23119-sucP</i>	Bimolecular complementation	In vivo cloning <i>S. cerevisiae</i>	4/4
<i>lacZ::sucP</i>	Bimolecular complementation	In vivo cloning <i>S. cerevisiae</i>	3/14
<i>galK::gfp</i>	Plasmid circularization	In vivo cloning <i>S. cerevisiae</i>	10/10
<i>lacZ::gfp</i>	Bimolecular complementation	In vivo cloning <i>S. cerevisiae</i>	1/10
<i>lacZ::phaCAB</i>	Bimolecular complementation	In vivo cloning <i>S. cerevisiae</i>	6/14
<i>galK::ampR</i>	Plasmid circularization	In vivo cloning <i>S. cerevisiae</i>	5/5
<i>lacZ::ampR</i>	Bimolecular complementation	In vivo cloning <i>S. cerevisiae</i>	3/5
$\Delta 40.4$ kb	Bimolecular complementation	PCR Amplification from D-Spec DNA	4/4
$\Delta fliA$ (205 bp)	Bimolecular complementation	In vivo cloning <i>S. cerevisiae</i>	10/10
$\Delta fliA$ (205 bp)	Bimolecular complementation	Oligo PCR amplification	9/10
$\Delta fliHDC$	Oligo PCR cloning	Oligo PCR amplification	10/10
$\Delta pcnB$ (81 bp)	Bimolecular complementation	PCR amplification from the D-Spec DNA	1/10
$\Delta hflK$	Oligo PCR cloning	PCR of a donor with asymmetric arms	10/10
ΔCA (12 kb)	Oligo PCR cloning	Oligo PCR amplification	9/10
$\Delta galK$ (32 bp)	Oligo PCR cloning	Oligo PCR amplification	10/10
<i>nudJ</i> (G151V)	Bimolecular complementation	In vivo cloning <i>S. cerevisiae</i>	2/2
<i>purB</i> (L115G)	Plasmid circularization	In vivo cloning <i>S. cerevisiae</i>	5/5
<i>dnaK</i> (T417N)	Oligo PCR cloning	Oligo PCR amplification	3/10
<i>zwf</i> (D405E)	Oligo PCR cloning	Oligo PCR amplification	9/9

5.4 Engenharia reversa das mutações encontradas no experimento de evolução

Para avaliar os efeitos de mutações individuais na aptidão celular, uma linhagem base foi construída para que modificações genômicas pudessem ser analisadas sob as

mesmas condições de crescimento, sem o uso de plasmídeos. Assim, o operon *cscBKA* foi inserido no genoma de *E. coli* DH5 α na região intergênica SS9 (BASSALO; GARST; HALWEG-EDWARDS; GRAU *et al.*, 2016). A construção do operon *cscBKA* com duas regiões de homologia de 400 pb foi amplificada por PCR a partir do plasmídeo pCscBKA (Figura 1A).

E2348/69 não foi usado para reconstrução de mutação porque é refratária à modificação genética usando CRISPR/Cas9. Várias transformações foram tentadas com o plasmídeo Cas9, mas esta linhagem não foi capaz de abrigar tal plasmídeo.

Nesta linhagem base, mutações foram selecionadas para serem avaliadas separadamente de acordo com o paralelismo entre populações e com base em estudos ou função gênica que sugerem atividades desses genes com aptidão celular.

5.4.1 Avaliação de *fitness*

5.4.2 Construção do sistema de competição

E. coli DH5 α foi escolhida como a cepa para testar as mutações encontradas durante o experimento de evolução. Essa cepa foi escolhida por ser o progenitor do experimento de evolução. Além disso, essa linhagem tem a capacidade de manutenção estável de plasmídeos necessários às modificações genômicas.

Para a construção da cepa base para o teste das mutações encontradas foi realizada a inserção do operon *cscBKA* no seu genoma, na região intergênica SS9, relatada com sendo neutra para integrações (BASSALO; GARST; HALWEG-EDWARDS; GRAU *et al.*, 2016) Nesta linhagem com o operon *cscBKA* integrado, foram inseridas, via CRISPR/Cas9, cópias do gene expressando a GFP em dois locais distintos do genoma da DH5 α (*galk::gfp* e *lacZ::gfp*). A expressão dessas inserções, mediada por promotor constitutivo, reforça o sinal de fluorescência detectada por citometria de fluxo. Essa cepa com as inserções da GFP e com o operon *cscBKA* foi considerada a linhagem tester-GFP, diante da qual as populações evoluídas e com mutações estão sendo testadas.

O primeiro passo a ser realizado foi a competição entre as linhagem SS9::*cscBKA* *versus* a tester-GFP, realizada em meio mínimo com sacarose. Essa competição indica

que houve uma redução no *fitness* das células com as duas cópias de GFP integradas de menos de 1% (Figura 10 C). Esse valor está dentro do que é encontrado em outros trabalhos (GULLBERG; CAO; BERG; ILBACK *et al.*, 2011). Além disso, esse valor baixo na diferença de *fitness* pode reduzir possíveis ruídos durante os ensaios de competição. O coeficiente de seleção (S) é calculado pela fórmula $S = \ln(E_f/T_f) - \ln(E_i/T_i)$, como citado anteriormente. O valor obtido pode ser dividido pelo número de gerações no experimento, obtendo assim o S por geração.

5.4.4 Ensaios de competição e *fitness* das populações evoluídas

Para o cálculo do S das populações do experimento de evolução, foram realizados ensaios de competição com os progenitores e, após isso, com os evoluídos, ambos diante das células com 2 cópias de GFP no genoma (tester-GFP). A diferença $S(\text{evoluídos}) - S(\text{progenitores})$ estabelece uma normalização para o coeficiente de seleção para as populações finais do experimento de evolução em relação aos progenitores (Figura 10 B). Para as populações evoluídas, os progenitores são aquelas células antes de iniciar o processo de evolução. Em relação reconstrução das mutações individuais em DH5α, o parental de referência foi considerado a linhagem com integração do operon *cscBKA* no genoma, visto que no experimento de evolução a maior força adaptativa foi a integração ao genoma do plasmídeo contendo o operon da metabolização da sacarose (populações D-Kan e D-Cm), ou a redução no número de cópias do plasmídeo (população D-Spec).

Para as competições das populações evoluídas o coeficiente de seleção foi calculado em relação ao progenitor. As populações de DH5α apresentaram uma variação maior de *fitness* em comparação a população progenitora em virtude dessas populações terem sido comparadas a essa cepa albergando o plasmídeo com o operon *cscBKA* (progenitor). Células que albergam plasmídeos normalmente têm a sua capacidade de crescimento reduzida (ainda mais se considerar o operon *cscBKA* em um plasmídeo de alto número de cópias). Isso pode ser observado nas curvas de crescimento (Figura 2) e na análise da velocidade máxima de crescimento das populações evoluídas e progenitoras (Figura 3 A). No caso das populações E2348/69, a diferença no coeficiente

de seleção foi menor provavelmente em virtude desta cepa já ter o operon naturalmente integrado no genoma.

5.4.5 *Fitness* das linhagens com alelos evoluídos

As competições das mutações testadas estão representadas na Figura 10C. A primeira mutação testada foi a integração do operon *cscBKA* no genoma (Figura 10 C). Foi observada possível integração do plasmídeo com este operon nas populações D-Kan e D-Cm. A comparação desta mutação com o tester-GFP foi considerada como o ponto zero e todas as outras mutações foram comparadas a esta.

A mutação encontrada no experimento de evolução em que houve maior variação positiva com diferença significativa no *fitness* foi $\Delta 40,4$ kb, da população D-Spec. Nesta região, há vários genes relacionados a diferentes funções, destacando-se o conjunto de operons envolvidos na formação do flagelo. Diante disto, como a síntese do flagelo envolve um grande consumo de energia pela célula (SUN; WANG; LIU; WU *et al.*, 2022), decidiu-se investigar mais precisamente se este é o ponto principal deste aumento do *fitness* nesta população.

Inicialmente, foi deletado o *fliA*, que é um regulador de transcrição que controla a expressão tardia do flagelo e também é encontrado na região de 40,4 kb (SUN; WANG; LIU; WU *et al.*, 2022). No entanto, após a competição, a deleção do regulador de expressão não afetou a aptidão celular (Figura 10 C). Investigando mais profundamente, o operon *flhDC* foi deletado, que codifica os reguladores primários da síntese do flagelo (MONDAY; MINNICH; FENG, 2004; SUN; WANG; LIU; WU *et al.*, 2022). Novamente, não houve diferença significativa em relação à integração do operon *cscBKA* sozinho, mas houve diferença para a deleção de 40,4 kb. Portanto, pode-se sugerir que há influência da deleção de outros genes nesta região.

Uma possibilidade a ser interpretada é o regulador de transcrição *rcaA*, que é encontrado nessa região deletada, e que é responsável pela transcrição de operons relacionados à síntese de ácido colânico (CA). Curiosamente, a região do genoma E2348-69 responsável pela síntese de CA foi deletada na população E-Cm. Portanto, decidimos testar a exclusão dessa região codificadora para CA em DH5 α . No entanto,

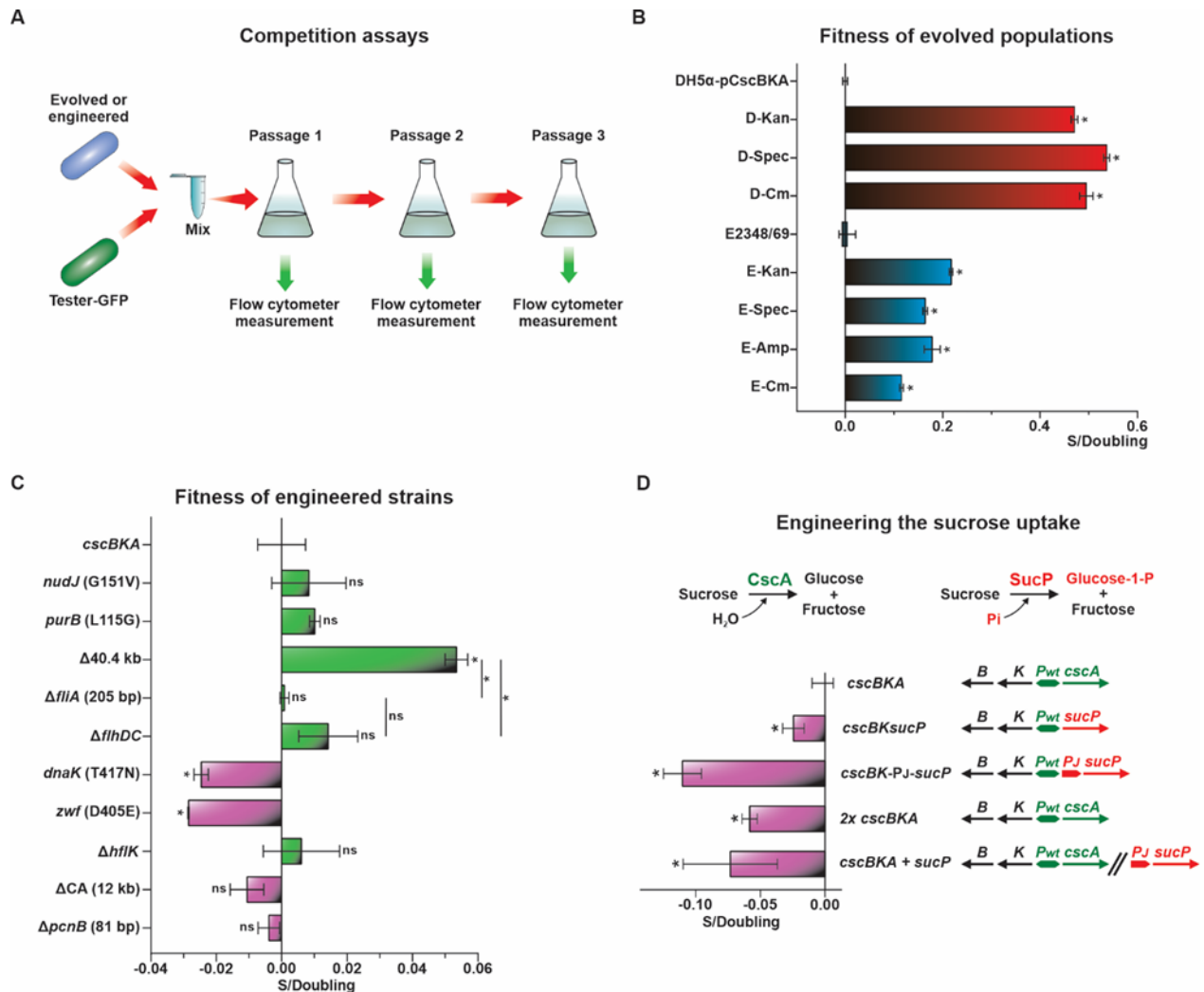
essa exclusão mostrou aptidão negativa quando em competição. Assim, pode-se supor que há uma sinergia na região de exclusão de 40,4 kb em D-Spec.

Outra mutação na população D-Spec testada foi a deleção de 81 pb do gene *pcnB*. A competição direta por sacarose dessa exclusão contra o testador GFP não representou um aumento de *fitness* (Figura 10 C), reforçando a ideia de que essa deleção pode estar relacionada a manutenção do plasmídeo pCscBKA, o que é discutido em mais detalhes abaixo.

Os SNPs nos genes *purB* e *nudJ* das populações D-Kan e D-Cm, respectivamente, também foram testados. Ambos os genes não mostraram diferenças significativas de *fitness* (Figura 10 C).

Outras mutações encontradas nas populações E2348-69 testadas foram a exclusão de *hflK* e os SNPs em *zwf* e *dnaK*. Para os SNPs, houve redução no *fitness*, enquanto para a deleção do *hflK* não houve diferença significativa em relação às células com inserção do operon *cscBKA* no genoma nas condições testadas.

Figura 10- Ensaios de competição para populações evoluídas e mutações selecionadas.



A Um esquema dos experimentos de competição. As cepas modificadas por engenharia reversa a serem analisadas e o tester-GFP são misturados em uma proporção de 1:1. Essa mistura é inoculada em meio fresco em triplicatas para iniciar a competição. Durante três passagens, as medições do citômetro de fluxo estabelecem a proporção de células com e sem GFP e a estimativa de gerações por passagem. Isso permite o cálculo de um S por duplicação (S/d). **B** Estimativas de ganho de *fitness* por meio de competições realizadas com populações evoluídas. Cada competição foi normalizada em relação ao seu respectivo progenitor do experimento de evolução (*E. coli* DH5α-pCscBKA ou E2348/69). **C** Estimativas de aptidão por meio de competições realizadas com cepas modificadas por engenharia reversa. **D** Estimativas de aptidão de cepas transformadas com diferentes estruturas de operon *cscBKA* e com o gene *sucP*. S em (C) e (D) foram normalizados com o S obtido para uma competição do tester-GFP vs. *E. coli* DH5α

SS9:cscBKA. A significância estatística foi avaliada em comparação com *E. coli* DH5 α SS9:cscBKA, ns: diferença não significativa; (*) diferença significativa ($p < 0,05$).

5.4.6 Variações estruturais do operon para consumo da sacarose

O operon *cscBKA* é formado por três genes, que codificam um transportador (*cscB*), uma frutocinase (*cscK*) e uma invertase (*cscA*). A invertase realiza a quebra (hidrólise) da sacarose em frutose e glicose, essa sendo fosforilada em seguida por uma hexocinase usando uma molécula de ATP. Entretanto, essa quebra da sacarose pode ser feita pelo produto do gene *sucP* (sacarose fosforilase), produzindo a glicose já fosforilada, usando fosfato inorgânico e economizando assim uma molécula de ATP (OLAVARRIA; FINA; VELASCO; VAN LOOSDRECHT *et al.*, 2019).

Com o intuito de melhorar o catabolismo da sacarose, economizando ATP, foi realizada a troca do *cscA* no operon *cscBKA* pelo gene *sucP* (*cscBKsucP*). Foram elaboradas duas configurações estruturais de operon com o gene *sucP* substituindo *cscA* diretamente no operon *cscBKA* previamente integrado no cromossomo de *E. coli* DH5 α . Ambas as estruturas foram previamente clonadas em levedura e os plasmídeos formados usados para amplificar os *donors*. Na primeira estrutura, o promotor bidirecional nativo do operon foi mantido. No segundo arranjo, também foi usado o promotor original, mas, além disso, foi inserido o promotor sintético J23119 para impulsionar a expressão do *sucP*. Para realizar a integração dessas estruturas, foi desenhado e clonado em plasmídeo o gRNA especificando como alvo o gene *cscA*. A montagem dos constructos que foram usados como *donor* foram realizadas em *S. cerevisiae*. Após amplificação por PCR, foi realizada a inserção no genoma, substituindo o gene *cscA* pelo *sucP*, com as duas estruturas. Além dessas duas formulações, uma terceira formulação também foi testada. Nessa alternativa foi mantido o operon *cscBKA* intacto, e no *lacZ* foi inserida uma cópia do gene *sucP*, sob regulação do promotor J23119.

Com essas construções, foi realizada a competição contra tester-GFP em meio mínimo com sacarose. Todas as variações das integrações do *sucP* apresentaram redução na capacidade de crescimento, demonstrando que com esse arranjo a integração desse gene não apresentou um incremento na capacidade de crescimento

(Figura 10 D). A integração com a associação *J23119-sucP* apresentou uma redução maior no *fitness* do que a integração apenas com o *sucP*. Essa redução mais intensa é provavelmente em virtude da intensa expressão do *sucP* com associação dos dois promotores.

Outro teste realizado foi com a inserção de duas cópias do operon *cscBKA*. A segunda cópia foi inserida no locus *galK* (2C-cscBKA). Após a cura dos plasmídeos usados para essa modificação genômica, foi realizada a competição contra a linhagem tester-GFP. O coeficiente de seleção para essa configuração também foi negativo (Figura 10 D).

5.4.7 Efeito do número de cópias de plasmídeos

Como apresentado anteriormente, a deleção de 81 pb do gene *pcnB* não apresentou um coeficiente de seleção positivo. Como esse gene já foi relacionado ao número de cópias de plasmídeos (LOPILATO, BORTNER, BECKWITH, 1986), resolveu-se realizar novos testes para verificar tais efeitos. Para essa avaliação, foi clonado o plasmídeo pGFP, em que a expressão de GFP é regulada pelo promotor do operon *lac*, o qual é induzido pela presença de IPTG. Dessa forma, ao analisar em citômetro de fluxo, pode-se verificar a intensidade de expressão de GFP, de acordo à quantidade de cópias do plasmídeo expressando a proteína fluorescente.

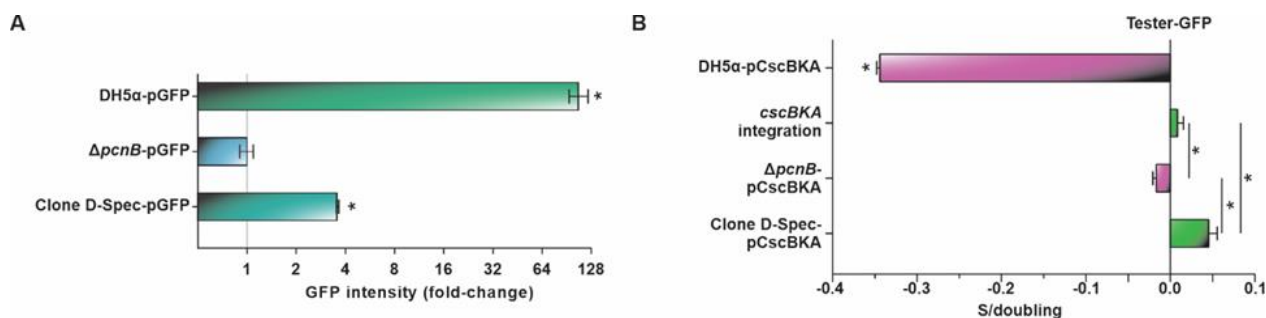
Para essa análise, inicialmente foi observada a expressão em células de DH5 α (DH5 α -pGFP, Figura 11 A), e foi observado o nível de expressão de GFP. Quando foi feita a análise para células de DH5 α com o plasmídeo pGFP e com a deleção do *pcnB* ($\Delta pcnB$ -pGFP), o nível de expressão da proteína fluorescente estava reduzido em relação às células de DH5 α (Figura 11 A). Outra avaliação realizada foi com um clone (colônia) isolado da população D-Spec, o qual primeiramente teve os plasmídeos curados. Depois o clone foi transformado com o pGFP (Evolved Clone D-Spec-pGFP). Nesse caso, apesar de haver uma intensificação da expressão de GFP em relação a deleção do *pcnB*, em comparação com as células de DH5 α houve redução na intensidade da fluorescência (Figura 11 A).

Para avaliar os efeitos dessas observações no crescimento, foram realizadas competições com essas linhagens com o plasmídeo pCscBKA, sendo todas comparados com o tester-GFP. As células de DH5 α com o plasmídeo (pCscBKA, Figura 11 B) apresentaram o menor *fitness*. Entretanto, quando se analisa a integração do operon no genoma, há um aumento da capacidade de crescimento.

Para as células com $\Delta pcnB$ -pCscBKA, houve um aumento na capacidade de crescimento em relação às células de DH5 α com o plasmídeo pCscBKA (Figura 11 B). Quando foi feita a análise de um clone isolado da população D-Spec (Evolved Clone D-Spec pCscBKA) foi observado que também aumentou o *fitness*, acima do observado para a integração cromossômica. Isso é devido, provavelmente, porque esse clone isolado possui outras mutações além da deleção do *pcnB*, como a deleção de 40,4 kb, que foi observado que aumenta o *fitness* celular.

Dessa forma, com os dados observados para o número de *reads* dos plasmídeos, a integração nas populações D-Kan e D-Cm e com os dados observados nas competições pode-se supor que a direção da evolução das populações de DH5 α foi para a redução do número de cópias do operon *cscBKA*, sendo uma cópia responsável por melhorar o crescimento.

Figura 11- Efeito de mutações nos números de cópias do plasmídeo.



A A expressão de GFP baseada em plasmídeo pode ser relacionada aos números de cópias do plasmídeo. O plasmídeo pGFP, de alto número de cópias, foi transformado em *E. coli* DH5 α (referência), DH5 α $\Delta pcnB$ e no clone evoluído D-Spec. A expressão de GFP foi medida por citometria de fluxo após a indução de IPTG. Os experimentos foram realizados em triplicatas. **B** Impacto no *fitness* dos números de cópias do plasmídeo pCscBKA. As seguintes cepas de *E. coli* foram submetidas a competições com o tester-GFP para estimar o *fitness* de crescimento em sacarose: pCscBKA (*E. coli* DH5 α com plasmídeo pCscBKA), SS9:*cscBKA* (DH5 α com o operon *cscBKA* integrado ao locus SS9), $\Delta pcnB$ -pCscBKA (DH5 α $\Delta pcnB$

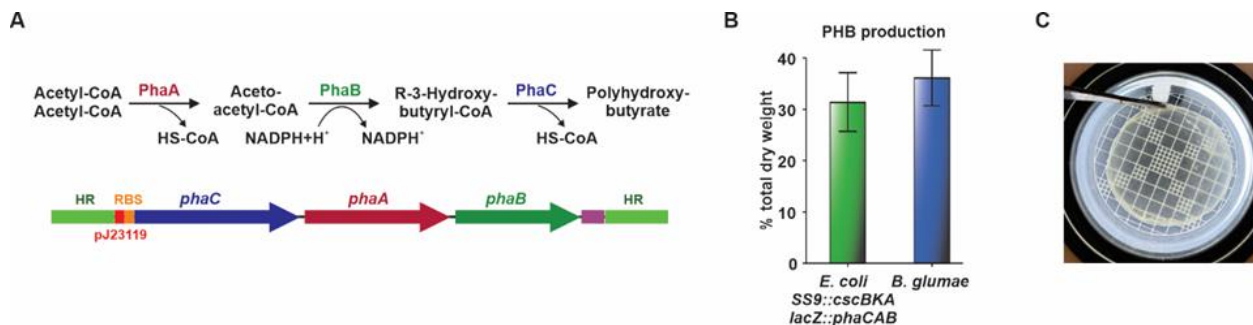
transformado com o pCscBKA), D-Spec-pCscBKA (clone evoluído D-Spec com o pCscBKA). (*) indicam diferenças significativas ($p < 0,05$) entre as linhagens.

5.4.8 Integração do operon *phaCAB*

A via metabólica mais estudada para produção de PHB é através do operon *phaCAB* de *Cupriavidus necator* (ou *Ralstonia eutropha*). Em nosso laboratório, temos esse operon clonado em plasmídeo (pBBRMCS2-*phaCAB*), o qual foi gentilmente cedido pelo Professor Dr. José Gregório Cabrera Gomez. Esse plasmídeo foi sequenciado e testes preliminares demonstram a sua capacidade de produção do PHB quando transformado em DH5 α . O operon *phaCAB* no pBBRMCS2-*phaCAB* encontra-se sob regulação do promotor e terminador naturais de *C. necator*.

Como mostrado nos experimentos de competição e no experimento de evolução, a manutenção de plasmídeo pode ser extremamente custosa para a célula, além de ser necessário o uso de antibióticos. Dessa forma, a integração do operon *phaCAB* ao genoma é provavelmente a melhor alternativa para uma produção eficiente do bioplástico. Considerando isso, e como teste do sistema de montagem de fragmentos de DNA, foi clonado um constructo desse operon em *S. cerevisiae* para inserção do mesmo no genoma de *E. coli* (Figura 12 A). Esse operon foi inserido no *lacZ*, sob regulação do promotor J23119 em células de DH5 α com o operon *cscBKA* no genoma. A produção de PHB foi detectada através da coloração de colônias por Sudan Black e por GC-MS, confirmando assim a capacidade de produção do composto pelo esse construto gênico inserido (CHOTCHINDAKUN; PATHOM-AREE; DUMRI; RUANGSURIYA *et al.*, 2021). A quantificação de PHB produzido pela linhagem modificada usando sacarose foi de aproximadamente de 31% do peso seco. Esse valor foi semelhante a quantidade de cerca de 36% do peso seco produzido por *Burkholderia glumae* (Figura 12 B). A Figura 12 C mostra o filme de PHB produzido pela *E. coli* transgênica.

Figura 12- Produção de poli-hidroxibutirato (PHB) em *E. coli* DH5 α com integração de operons *cscBKA* e *phaCAB* da sacarose.



A Via metabólica para produção de PHB e estrutura da construção usada para inserção. Esta estrutura foi montada em levedura. Após a clonagem, o *donor* foi amplificado e usado para inserção genômica no locus *lacZ*. A expressão do operon *phaCAB* é regulada por um promotor forte constitutivo (*pJ23119*). **B** Produção de PHB medida como a porcentagem de massa celular seca que representa o bioplástico *E. coli* DH5 α com o operon *cscBKA* e o operon *phaCAB* integrados ao genoma e células de *Burkholderia glumae* como comparação. **C** Filme de PHB produzido a partir de células *E. coli* DH5 α com o operon *cscBKA* e o operon *phaCAB* integrados ao genoma.

6 DISCUSSÃO

O desenvolvimento de uma indústria verde passa pelo entendimento da capacidade e uso de microrganismos para síntese de biomateriais. Nesse sentido, para que esses microrganismos cresçam e produzam de forma financeiramente sustentável, é necessário o uso de fontes de energia de baixo custo, como a sacarose (LI; WILKINS, 2020; MOHAMED; MUNDHADA; LANDBERG; CANN *et al.*, 2019). Esse dissacarídeo apresenta um interesse especial ao Brasil devido à grande capacidade de produção em território nacional.

A bactéria *E. coli* é considerada um organismo modelo e possui diversas das suas características, como genoma e condições de crescimento, bem elucidadas. Dessa forma, esse microrganismo é um interessante alvo para estudos que visem a utilização como chassi para produção de bioprodutos. Apesar de possuir tais características, a capacidade de metabolizar a sacarose não é igualmente representada entre as diversas cepas, sendo mais frequentemente encontrada nas cepas enteropatogênicas, ao contrário das laboratoriais ou industriais, sendo representadas principalmente pela *E. coli* W (MOHAMED; MUNDHADA; LANDBERG; CANN *et al.*, 2019; SABRI; NIELSEN; VICKERS, 2013).

O consumo de sacarose em cepas de *E. coli* é mediado por dois sistemas. O primeiro é formado pelo regulador *scr*, que expressa um sistema de fosfotransferase para sacarose. A segunda forma de utilização é através do operon *csc*. Esse operon é formado pelos genes *cscB*, *cscK* e *cscA*, além do repressor *cscR* (OLAVARRIA; FINA; VELASCO; VAN LOOSDRECHT *et al.*, 2019; SABRI; NIELSEN; VICKERS, 2013). O repressor desse operon é ativo apenas quando a concentração da sacarose é inferior a 2%.

Interessante também ressaltar que nem todos os genes que compõem esse cluster são necessários para que a células que o alberguem cresçam em sacarose, mas são necessários para um crescimento eficiente. Isso foi mostrado por Sabri e colaboradores (SABRI; NIELSEN; VICKERS, 2013), que relatam que a deleção da fructocinase (*cscK*) não afeta o crescimento em sacarose. Entretanto, no caso da inativação desse gene, a frutose que seria fosforilada acumula-se no meio, indicando assim a necessidade dessa enzima para a utilização completa da sacarose. Além disso,

nesse trabalho também é demonstrada a não necessidade da utilização do repressor (*cscR*). Com base nessas informações, para que a capacidade de uso da sacarose como fonte de carbono possa ser estendida para DH5 α , os genes *cscBKA* da cepa E2348/69 foram clonados em um plasmídeo. Após testes de crescimento em meio contendo esse açúcar, sequenciamento e quantificação do consumo da sacarose, a integridade desse operon foi confirmada.

Apesar das células *E. coli* albergando o operon *cscBKA* apresentarem crescimento em sacarose, essa capacidade de crescimento é reduzida quando comparado a glicose (MOHAMED; MUNDHADA; LANDBERG; CANN *et al.*, 2019). Diante disso, para tentar otimizar essa característica, foi realizado o experimento de ALE, no qual a sacarose foi a única fonte de carbono. Para esse experimento foram usadas populações das cepas DH5 α , albergando o plasmídeo com o operon *cscBKA*, e E2348/69, a qual metaboliza sacarose naturalmente. Esse experimento foi realizado por cerca de 600 a 800 gerações, com transferências diárias. Ao final, as populações (três de DH5 α e quatro de E2348/69) foram caracterizadas fenotipicamente para crescimento em meio mínimo sacarose em comparação aos progenitores. O DNA genômico de todas as populações evoluídas foi sequenciado.

Uma observação realizada que chama a atenção é em relação ao consumo de sacarose dessas populações finais e progenitores. Mesmo indicando um aumento na capacidade de crescimento nas condições testadas, o consumo de sacarose não alterou (no caso das populações de E2348/69) ou reduziu-se (nas populações evoluídas de DH5 α). Com essa observação fica claro que a direção da evolução que as células das populações do experimento encontraram não foi relacionada a um maior consumo de sacarose nas condições analisadas.

Não foram realizados experimentos para tentar entender o que pode estar causando esse consumo não eficiente do açúcar. Entretanto, ao analisar a composição do meio de cultivo observa-se que determinados componentes estão em quantidades até vezes inferiores ao que é observado em trabalhos que usam o meio mínimo para cultivo de *E. coli* (ARIFIN; ARCHER; LIM; QUEK *et al.*, 2014; MOHAMED; MUNDHADA; LANDBERG; CANN *et al.*, 2019). Dentre esses componentes, o ferro é o que chama mais atenção. Os íons ferro estão presentes em diversas proteínas servindo como cofatores

(BAEZ; SHARMA; BRYUKHANOV; ANDERSON *et al.*, 2022). Em condições de falta de ferro, as células bacterianas podem ter vários distúrbios, desde alterações metabólicas até afetar a citocinese durante a divisão celular (BAEZ; SHARMA; BRYUKHANOV; ANDERSON *et al.*, 2022; SANTOS; LAMMERS; ZHOU; SPARKS *et al.*, 2018). Os efeitos práticos dessa falta íon no meio de crescimento é o baixo acúmulo de biomassa e determinados subprodutos, como acetato, em condições mais brandas de falta de ferro e lactato, em condições mais severas (BAEZ; SHARMA; BRYUKHANOV; ANDERSON *et al.*, 2022; FOLSOM; PARKER; CARLSON, 2014). O acúmulo de acetato no meio de cultivo reduz a eficiência da utilização da fonte de carbono e pode inibir o crescimento (GECSE; LABUNSKAITE; PEDERSEN; KILSTRUP *et al.*, 2024).

A engenharia reversa de mutações em um experimento de ALE compreende a reconstrução das mutações individuais ou em conjunto no progenitor do experimento de evolução. Dessa forma é possível entender a direção das adaptações que as populações evoluídas seguiram para adquirir um fenótipo de crescimento superior, além de selecionar possíveis mutações que possam contribuir de forma positiva para esse fenótipo (GASSLER; BAUMSCHABL; SALLABERGER; EGERMEIER *et al.*, 2022; LEE; PARK; PARK; KIM *et al.*, 2020).

Para que a engenharia reversa seja realizada é necessário haver um conjunto de técnicas de modificações genômicas bem estabelecidas e de fácil execução, as quais permitam a reconstrução de várias mutações. Por isso, uma parte importante desse projeto foi dedicada ao desenvolvimento de técnicas de modificação genômica em *E. coli*.

A técnica base incorporada a esse projeto para edição de genomas é o CRISPR associada a Cas9 (CRISPR/Cas9). Essa é uma das mais promissoras técnicas para modificações genômicas sendo possível com a mesma fazer deleções, inserções e até modificações de nucleotídeos únicos. Para a execução da técnica de CRISPR/Cas9 são necessários três componentes: a Cas9, o gRNA e o fragmento de DNA para fazer a modificação genômica (*donor*). Além disso, como *E. coli* não possui sistema de recombinação homóloga eficiente, é necessário o uso do sistema de recombinases Lambda Red (BASSALO; GARST; HALWEG-EDWARDS; GRAU *et al.*, 2016). Para suprir a necessidade da Cas9 e do sistema Lambda Red, foi adquirido comercialmente um plasmídeo com esses componentes.

O gRNA na técnica de CRISPR/Cas9 é uma das partes mais importantes, tendo em vista que é esse fragmento de RNA que estabelece o local no genoma onde a Cas9 irá realizar a clivagem (JACOBUS; BARRETO; DE BEM; MENEGON *et al.*, 2022). Dessa forma, foi necessário o desenvolvimento de um sistema para clonagem de plasmídeos para a expressão dessa molécula. A clonagem rápida e de forma simples de plasmídeos para expressão do gRNA é uma necessidade inerente a qualquer projeto que use a técnica de CRISPR/Cas9. Normalmente a clonagem de plasmídeo para expressão do guia usa-se técnicas que empreguem enzimas de restrição/ligase (JIANG; CHEN; DUAN; SUN *et al.*, 2015), ou Gibson Assembly (TONG; JORGENSEN; WHITFORD; WEBER *et al.*, 2021). Como esses métodos citados usam de kits comerciais para clonagem, ou ainda etapas de purificação e tratamento com DpnI, para se reduzir os custos de execução da técnica, foi montado um sistema de dois plasmídeos. Com essa ideia, são amplificados por PCR em dois fragmentos, sendo os mesmos transformados em células competentes de *E. coli*, sendo formado assim o plasmídeo para expressão do gRNA por *in vivo* cloning (JACOBUS; GROSS, 2015). Essa estratégia possui como vantagens a não necessidade do uso de enzimas de restrição ou purificação dos fragmentos, sendo transformados diretamente os fragmentos de PCR em células competentes, reduzindo assim o custo para a clonagem do plasmídeo para expressão de gRNA.

A confecção de fragmentos de DNA para serem usados como *donors* são de grande importância para a execução da técnica de CRISPR/Cas9. Como citado anteriormente, *E. coli* não possui um sistema de recombinação homóloga com alta eficiência, sendo necessário para a execução de técnica de CRISPR/Cas9 a associação do sistema Lambda Red. Para otimizar a eficiência da técnica uma estratégia usada é aumentar o tamanho dos braços de homologia (ZHAO; YUAN; XIONG; SUN *et al.*, 2016). Bassalo e colaboradores (BASSALO; GARST; HALWEG-EDWARDS; GRAU *et al.*, 2016) usam tamanhos de braços de homologia que vão desde 1 kb a 300 pb de homologia com alta eficiência de modificação; esses autores também conseguiram uma eficiência próxima a 50%, quando usaram 100 pb de homologia para integrar ao genoma de *E. coli* a via de produção do butanol, com 10 kb de tamanho. Já Jiang e colaboradores (JIANG; CHEN; DUAN; SUN *et al.*, 2015) testaram diferentes tamanhos de braços de homologia no *donor*, variando desde 40 pb a mais de 500 pb de cada lado, em diferentes loci, o que

demonstrou maior eficiência quando são usados braços de homologia maiores. Padrão semelhante também é observado no trabalho de Zhao e colaboradores (ZHAO; YUAN; XIONG; SUN *et al.*, 2016), no qual, quando usados fragmentos com 51 pb de homologia de ambos os lados, a eficiência decaía para 0%, sendo aumentada para valores próximos a 100% quando as regiões de homologia com o genoma foram aumentadas. Dessa forma, fica evidente a necessidade de etapas de pré-clonagem para a confecção dos *donors*. Diante desses resultados também foi decidido usar como tamanho-base de braços de homologia com o genoma 400 pb de cada lado.

Para a clonagem de *donors*, foram testados alguns kits de clonagem em *E. coli*, os quais para as necessidades aqui presentes, não se mostraram práticos e efetivos, além de possuírem custo elevado. Dessa forma foram avaliadas outras alternativas, sendo a mais efetiva o desenvolvimento de uma estratégia usando *in vivo* cloning em *Saccharomyces cerevisiae*. Essa levedura possui a capacidade de ligar fragmentos de DNA com alta especificidade, através da especificação de regiões de homologia especificada por primers. Com essa técnica, na execução desse projeto, foram clonados *donors* para inserções (*gfp*, operon *csc*, operon *phaCAB*, *sucP*), deleção ($\Delta fliA$) e SNPs (*nudJ* (G151V) e *purB* (L115G)). Essa estratégia apresenta vantagens em comparação às técnicas tradicionais em virtude de não usar enzimas de restrição e não ser necessária a purificação dos fragmentos gerados por PCR, sendo a reação transformada diretamente nas células competentes.

As mutações selecionadas após análise do sequenciamento do DNA genômico estão sendo reconstruídas em DH5 α . Isso porque a cepa E2348/69, além de ser uma ETEC, não mantém alguns plasmídeos de forma estável, como o próprio plasmídeo da Cas9 (HOBSON; PRICE; WARD; RAIVIO, 2008). Também, em virtude disso, as mutações em DH5 α foram as primeiras a serem testadas.

Uma das primeiras observações do sequenciamento das populações evoluídas de DH5 α foi que em duas dessas populações (D-Kan e D-Cm) o plasmídeo albergando o operon *cscBKA* foi inserido no cromossomo bacteriano. Para que as células mantenham plasmídeos há normalmente um custo metabólico associado que pode ser aumentado em virtude das proteínas que estão sendo expressas por esse plasmídeo, o que pode ser observado no *fitness* geral no crescimento (AHMAD; PRENSKY; BALESTRIERI;

ELNAGGAR *et al.*, 2023; LI; WANG; GENG; LI *et al.*, 2012). Se considerarmos que o plasmídeo pCscBKA possui os genes responsáveis pelo metabolismo da sacarose com expressão constitutiva (e os ensaios de competição realizados demonstrando o grande custo metabólico para manutenção desse plasmídeo) fica evidente que as células de *E. coli* nessas populações resolveram esse problema integrando esse plasmídeo ao genoma (também observado através de ensaios de competição). Assim, a integração do plasmídeo no genoma representa uma vantagem competitiva para essas populações e pode indicar que a direção da evolução foi no sentido de reduzir o número de cópias do operon.

A ideia da redução do número de cópias do *operon* cscBKA pode ser suportada também pelo fato de a população D-Spec, apesar de não ter tido a integração no genoma desse plasmídeo, houve uma redução na cobertura das *reads* de ambos plasmídeos que essa população alberga, indicando assim uma redução do número de cópias desses plasmídeos nessas células, as quais podem ser associadas a deleção na região codificadora do gene *pcnB* e, potencialmente, a mutação pontual no gene *dnaA*. Mutações no primeiro são relacionadas a manutenção do número de cópias de plasmídeos dentro da célula e o segundo é relacionado a replicação de plasmídeos (MA; CAMPBELL, 1988; MASTERS; COLLOMS; OLIVER; HE *et al.*, 1993). Isso corrobora o fato de que para a célula manter um plasmídeo em seu interior há alterações na capacidade de crescimento (LI; WANG; GENG; LI *et al.*, 2012).

Assim, diante dessas observações, foi realizada a integração do operon cscBKA no sítio intergênico SS9 no genoma de DH5 α , sendo esta constituída a cepa base para testes das mutações. Essa região foi escolhida porque no trabalho de Bassalo e colaboradores (BASSALO; GARST; HALWEG-EDWARDS; GRAU *et al.*, 2016) é relatado que essa região foi efetiva para a inserção de uma sequência codificadora para GFP e a expressão nesse *locus* desse gene não foi reduzida em comparação ao demais locais testados por esses autores.

Após terem sido desenvolvidas tecnologias para reconstrução das mutações selecionadas, o uso de um sistema para avaliar se essas mudanças genômicas eram efetivas para melhorar o *fitness* tornou-se evidente. Para isso foi escolhido o ensaio de

competição com células marcadas com sequências de *gfp* integradas ao genoma, em virtude da robustez dessa técnica (GULLBERG; CAO; BERG; ILBACK *et al.*, 2011).

As primeiras competições realizadas foram com as populações evoluídas. Pode-se observar diferenças entre as populações evoluídas e as progenitoras, indicando um *fitness* positivo. Esse incremento na capacidade de crescimento celular das populações evoluídas confirma que o processo de ALE foi efetivo para adaptar as populações para crescimento em sacarose nas condições do experimento.

Os SNPs nos genes *nudJ* (G151V) e *purB* (L115G) foram testados por ensaios de competição. Para essas mutações não foram observadas vantagens no *fitness* celular. Buskirk, Peace e Lang (BUSKIRK; PEACE; LANG, 2017) citam que cerca de 80% das mutações encontradas em experimentos de evolução não causam mudanças significativas no crescimento celular. Essas mutações possuem sua frequência alta na população, mas isso provavelmente é pelo fato de as células com essas mutações terem sido carregadas por outra(s) positiva(s) para o *fitness* celular.

Uma mutação interessante identificada na população D-Spec foi a deleção completa da região do *cluster* responsável pela montagem do flagelo (c. 40,4 kb). Após o ensaio de competição, essa mutação apresentou *fitness* positivo para o crescimento celular. Já foi relatado que a inativação de regulador da expressão de *cluster* causa aumento na produção de proteínas recombinantes (HAN; JUNG; OH, 2021). O flagelo controla a motilidade celular de *E. coli* e é necessária a expressão de cerca de 60 proteínas para a sua síntese (BAREMBRUCH; HENGGE, 2007; CHILCOTT; HUGHES, 2000). Esse emprego de recursos celulares para montagem do flagelo requer 5,0% de energia, juntamente com aproximadamente mais 5,2% de energia para o seu funcionamento (BAREMBRUCH; HENGGE, 2007; CHILCOTT; HUGHES, 2000). A inibição da expressão do flagelo pode ser realizada pela repressão catabólica de carbono (*Carbon catabolite repression* - CCR) a qual é ativada na presença de carboidratos, como glicose (GORKE; STULKE, 2008; SOUTOURINA; BERTIN, 2003).

A estrutura gênica para expressão do flagelo que se encontra deletada na população D-Spec tem c. 40,4 kb. Essa região é grande, contendo vários outros genes além dos relacionados à montagem do flagelo. Assim, para investigar somente o efeito na montagem do flagelo, foi realizada a deleção do regulador *fliA*. A proteína FliA é um

fator sigma relacionado à regulação da expressão da terceira classe das proteínas flagelares, estas sendo responsáveis pelo restante da estrutura do flagelo e o sistema de quimiotaxia (BAREMBRUCH; HENGGE, 2007; SUN; WANG; LIU; WU *et al.*, 2022). A deleção do fator *fliA* não demonstrou, durante os ensaios de competição, a mesma melhoria no *fitness* verificado para a deleção do cluster de c. 41 Kb. Em vista desse resultado, também realizamos a deleção dos fatores *flhDC*, relacionados ao controle da expressão das proteínas flagelares da classe II (SUN; WANG; LIU; WU *et al.*, 2022). Da mesma forma, o ganho de *fitness* não foi da mesma magnitude do que a deleção 40,4 kb.

Outras mutações encontradas nas populações DH5 α dizem respeito às subunidades da RNA polimerase. Em *E. coli*, a RNA polimerase é formada por cinco subunidades: α (duas subunidades), β , β' (beta prime) e ω , codificadas pelos genes *rpoA*, *rpoB*, *rpoC* e *rpoZ*, respectivamente (SUTHERLAND; MURAKAMI, 2018). Experimentos de evolução em meio mínimo relatam mutações nas subunidades dessa polimerase (HERRING; RAGHUNATHAN; HONISCH; PATEL *et al.*, 2006; MOHAMED; MUNDHADA; LANDBERG; CANN *et al.*, 2019). No estudo aqui em questão, foram encontradas mutações no gene *rpoC*, *rpoZ* e *rpoB* (populações D-Kan, D-Spec e D-Cm, respectivamente). No trabalho de Herring e colaboradores (HERRING; RAGHUNATHAN; HONISCH; PATEL *et al.*, 2006) é relatado um experimento de evolução em meio mínimo com glicerol e que as mutações nos genes *rpoC* e *rpoB* aumentaram a taxa de crescimento em 48-65%. Apesar desses achados indicarem um aumento na capacidade de crescimento de *E. coli* em meio mínimo, quando é comparado com meio rico, há um decréscimo do *fitness* celular (CONRAD; FRAZIER; JOYCE; CHO *et al.*, 2010). O papel da subunidade ω (*rpoZ*) ainda não é claro, exceto pela interação com a alarmona ppGpp (PATEL; GAUTAM; CHATTERJI, 2019).

Apesar de mutações isoladas nas subunidades da RNA polimerase proporcionarem um incremento no crescimento de *E. coli*, o crescimento em meio rico é prejudicado. Dessa forma, como, para o desenvolvimento de uma cepa chassi é interessante o crescimento em diferentes meios de cultivo, as mutações nas subunidades da polimerase não foram testadas nesse estudo.

Um fato interessante nas mutações das populações E2348/69 ocorreu nos *loci hfl*. Mutações nos genes *hflK* foram encontradas nas populações E-Kan, E-Spec e E-Cm. Na população E-Amp, foi encontrada uma mutação no *hflB*. Esses genes podem ser relacionados ao ciclo lítico-lisogênico do bacteriófago lambda. A decisão para a célula bacteriana com bacteriófagos entrar em ciclo lítico ou lisogênico depende da ação de proteínas do próprio fago (CI, CII e CIII) e do hospedeiro (proteínas Hfl) (BANDYOPADHYAY; PARUA; DATTA; PARRACK, 2011; ECHOLS; GREEN, 1971). Quando o bacteriófago se encontra em estado de latência na célula bacteriana, a proteína CI encontra-se ativa, inibindo todos os promotores de atividades virais. A CII é uma proteína instável, que, na concentração adequada, é o principal sinalizador para o fago entrar no ciclo lisogênico. Já a proteína CIII controla a estabilidade da proteína CII, inibindo a ação protease expressa pelo gene *hflB* (BANDYOPADHYAY; PARUA; DATTA; PARRACK, 2011; HERMAN; THEVENET; D'ARI; BOULOC, 1997). HflB é uma metaloprotease que cliva proteínas com problemas de dobra quando expressas muito rapidamente quando nutrientes estão em grande disponibilidade (BANDYOPADHYAY; PARUA; DATTA; PARRACK, 2011). A CII funciona como um substrato para essa enzima, que, quando clivada, libera a atividade lítica do bacteriófago. A proteína HflK atua juntamente com HflC, formando complexo HflKC, o qual modula a da atividade de HflB, através da formação de complexos com o mesmo.

Mutações em genes *hfl* estão relacionadas ao aumentar a probabilidade do fago entrar no ciclo lisogênico (BANDYOPADHYAY; PARUA; DATTA; PARRACK, 2011). Como a cepa E2348/69 possui 13 profagos em seu genoma (IGUCHI; THOMSON; OGURA; SAUNDERS *et al.*, 2009), essas mutações podem contribuir para o estabelecimento de células com mais resistência à entrada no ciclo lítico dos bacteriófagos.

A integração do GFP no genoma bacteriano, além de fornecer a linhagem base para os ensaios de competição, também serviu como prova de conceito para o desenho da estrutura gênica de integração do operon *phaCAB* no genoma de *E. coli*, pois indicou a escolha do promotor forte, da RBS e do terminador para a inserção desse operon. Isso porque Yin e colaboradores (YIN; WANG; FU; GAO *et al.*, 2015) relatam que só foi detectada produção de PHB após a integração de quatro cópias do operon em múltiplos

locais do genoma. Em contraste, foi conseguida a produção de PHB com inserção de apenas uma cópia do operon *phaCAB*. Li e colaboradores (LI; WANG; GENG; LI *et al.*, 2012) também relatam algo semelhante: a integração cromossômica em células de DH5 α com apenas uma cópia do operon produziu apenas 3,6% da massa celular seca de PHB, conseguindo um aumento para 23,7% usando associação de promotores em *tandem*. A maior produção de PHB conseguida na nossa linhagem provavelmente deve-se ao fato de que a expressão do operon *phaCAB* foi dirigida sob a ação de um promotor constitutivo (como na linhagem expressando GFP), e não foi usado o promotor natural da *R. eutropha* (YIN; WANG; FU; GAO *et al.*, 2015). Portanto, os nossos resultados iniciais, quanto à engenharia da síntese de PHB, são bastante promissores.

7 CONCLUSÕES

Nesse trabalho, foram propostas melhorias no sistema de CRISPR/Cas9 para serem aplicadas na edição genômica de *E. coli*. A primeira melhoria citada foi a implementação de três sistemas de clonagem de plasmídeos para expressão de gRNA, sendo o sistema mais avançado o uso do sistema com oligos, que aumenta a velocidade de clonagem da molécula de RNA guia. Também foram propostos sistemas de clonagem de *donor* usando amplificações de oligos e montagem de fragmentos de DNA em levedura. Com essas ferramentas foi possível a reconstrução de mutações identificados em experimento de evolução de populações de *E. coli* em sacarose. A avaliação das mutações reconstruídas demonstrou que a redução do número de cópias do plasmídeo *cscBKA* e a inativação do flagelo aumentaram a capacidade de crescimento. Com os sistemas para otimização de CRISPR/Cas9 apresentados foi possível integrar a via metabólica para produção de PHB e foi detectada a produção de ~31% da massa celular seca, partir da sacarose.

REFERÊNCIAS

- ADNAN, M.; SIDDIQUI, A. J.; ASHRAF, S. A.; SNOUSSI, M. *et al.* Polyhydroxybutyrate (PHB)-Based Biodegradable Polymer from *Agromyces indicus*: Enhanced Production, Characterization, and Optimization. **Polymers (Basel)**, 14, n. 19, Sep 23 2022.
- AHMAD, M.; PRENSKY, H.; BALESTRIERI, J.; ELNAGGAR, S. *et al.* Tradeoff between lag time and growth rate drives the plasmid acquisition cost. **Nat Commun**, 14, n. 1, p. 2343, Apr 24 2023.
- AKKOYUNLU, B.; GABARRE, C.; DALY, S.; CASEY, E. *et al.* Process modelling for industrial scale polyhydroxybutyrate production using fructose, formic acid and CO₂: Assessing carbon sources and economic viability. **Bioresour Technol**, 393, p. 130139, Feb 2024.
- ALMARIO, M. P.; REYES, L. H.; KAO, K. C. Evolutionary engineering of *Saccharomyces cerevisiae* for enhanced tolerance to hydrolysates of lignocellulosic biomass. **Biotechnol Bioeng**, 110, n. 10, p. 2616-2623, Oct 2013.
- ALMEIDA, L. J. M.; SILVA, A. V.; SILVA, J. S. L.; SILVA, J. F. *et al.* Sugarcane productivity and economic viability in response to planting density. **Brazilian Journal of Biology**, 84, 2024.
- ANBUKARASU, P.; SAUVAGEAU, D.; ELIAS, A. Tuning the properties of polyhydroxybutyrate films using acetic acid via solvent casting. **Sci Rep**, 5, p. 17884, Dec 7 2015.
- ANTON, B. P.; RALEIGH, E. A. Complete Genome Sequence of NEB 5-alpha, a Derivative of *Escherichia coli* K-12 DH5alpha. **Genome Announc**, 4, n. 6, Nov 10 2016.
- ARCHER, C. T.; KIM, J. F.; JEONG, H.; PARK, J. H. *et al.* The genome sequence of *E. coli* W (ATCC 9637): comparative genome analysis and an improved genome-scale reconstruction of *E. coli*. **BMC Genomics**, 12, p. 9, Jan 6 2011.
- ARIFIN, Y.; ARCHER, C.; LIM, S.; QUEK, L. E. *et al.* *Escherichia coli* W shows fast, highly oxidative sucrose metabolism and low acetate formation. **Appl Microbiol Biotechnol**, 98, n. 21, p. 9033-9044, Nov 2014.
- BAESHEN, N. A.; BAESHEN, M. N.; SHEIKH, A.; BORA, R. S. *et al.* Cell factories for insulin production. **Microb Cell Fact**, 13, p. 141, Oct 2 2014.
- BAEZ, A.; SHARMA, A. K.; BRYUKHANOV, A.; ANDERSON, E. D. *et al.* Iron availability enhances the cellular energetics of aerobic *Escherichia coli* cultures while upregulating anaerobic respiratory chains. **N Biotechnol**, 71, p. 11-20, Nov 25 2022.
- BANDYOPADHYAY, K.; PARUA, P. K.; DATTA, A. B.; PARRACK, P. Studies on *Escherichia coli* HflKC suggest the presence of an unidentified lambda factor that influences the lysis-lysogeny switch. **BMC Microbiol**, 11, p. 34, Feb 17 2011.
- BAREMBRUCH, C.; HENGGE, R. Cellular levels and activity of the flagellar sigma factor FliA of *Escherichia coli* are controlled by FlgM-modulated proteolysis. **Mol Microbiol**, 65, n. 1, p. 76-89, Jul 2007.

BASSALO, M. C.; GARST, A. D.; HALWEG-EDWARDS, A. L.; GRAU, W. C. *et al.* Rapid and Efficient One-Step Metabolic Pathway Integration in *E. coli*. **ACS Synth Biol**, 5, n. 7, p. 561-568, Jul 15 2016.

BEHERA, S.; PRIYADARSHANEE, M.; VANDANA; DAS, S. Polyhydroxyalkanoates, the bioplastics of microbial origin: Properties, biochemical synthesis, and their applications. **Chemosphere**, 294, p. 133723, May 2022.

BESSMAN, M. J. A cryptic activity in the Nudix hydrolase superfamily. **Protein Sci**, 28, n. 8, p. 1494-1500, Aug 2019.

BLATTNER, F. R.; PLUNKETT, G., 3rd; BLOCH, C. A.; PERNA, N. T. *et al.* The complete genome sequence of *Escherichia coli* K-12. **Science**, 277, n. 5331, p. 1453-1462, Sep 5 1997.

BLOUNT, Z. D. The unexhausted potential of *E. coli*. **Elife**, 4, Mar 25 2015.

BLOUNT, Z. D.; BORLAND, C. Z.; LENSKI, R. E. Historical contingency and the evolution of a key innovation in an experimental population of *Escherichia coli*. **Proc Natl Acad Sci U S A**, 105, n. 23, p. 7899-7906, Jun 10 2008.

BOCANEGRA, J. K.; DA CRUZ PRADELLA, J. G.; DA SILVA, L. F.; TACIRO, M. K. *et al.* Influence of pH on the molecular weight of poly-3-hydroxybutyric acid (P3HB) produced by recombinant *Escherichia coli*. **Appl Biochem Biotechnol**, 170, n. 6, p. 1336-1347, Jul 2013.

BRUSCHI, M.; BOYES, S. J.; SUGIARTO, H.; NIELSEN, L. K. *et al.* A transferable sucrose utilization approach for non-sucrose-utilizing *Escherichia coli* strains. **Biotechnol Adv**, 30, n. 5, p. 1001-1010, Sep-Oct 2012.

BUBECK, P.; WINKLER, M.; BAUTSCH, W. Rapid cloning by homologous recombination in vivo. **Nucleic Acids Res**, 21, n. 15, p. 3601-3602, Jul 25 1993.

BUSKIRK, S. W.; PEACE, R. E.; LANG, G. I. Hitchhiking and epistasis give rise to cohort dynamics in adapting populations. **Proc Natl Acad Sci U S A**, 114, n. 31, p. 8330-8335, Aug 1 2017.

CALLONI, G.; CHEN, T.; SCHERMANN, S. M.; CHANG, H. C. *et al.* DnaK functions as a central hub in the *E. coli* chaperone network. **Cell Rep**, 1, n. 3, p. 251-264, Mar 29 2012.

CHILCOTT, G. S.; HUGHES, K. T. Coupling of flagellar gene expression to flagellar assembly in *Salmonella enterica* serovar typhimurium and *Escherichia coli*. **Microbiol Mol Biol Rev**, 64, n. 4, p. 694-708, Dec 2000.

CHOTCHINDAKUN, K.; PATHOM-AREE, W.; DUMRI, K.; RUANGSURIYA, J. *et al.* Low Crystallinity of Poly(3-Hydroxybutyrate-co-3-Hydroxyvalerate) Bioproduction by Hot Spring Cyanobacterium *Cyanosarcina* sp. AARL T020. **Plants (Basel)**, 10, n. 3, Mar 8 2021.

CONLEY, E. C.; SAUNDERS, V. A.; SAUNDERS, J. R. Deletion and rearrangement of plasmid DNA during transformation of *Escherichia coli* with linear plasmid molecules. **Nucleic Acids Res**, 14, n. 22, p. 8905-8917, Nov 25 1986.

CONRAD, T. M.; FRAZIER, M.; JOYCE, A. R.; CHO, B. K. *et al.* RNA polymerase mutants found through adaptive evolution reprogram *Escherichia coli* for optimal growth in minimal media. **Proc Natl Acad Sci U S A**, 107, n. 47, p. 20500-20505, Nov 23 2010.

COOPER, V. S.; LENSKI, R. E. The population genetics of ecological specialization in evolving *Escherichia coli* populations. **Nature**, 407, n. 6805, p. 736-739, Oct 12 2000.

DATSENKO, K. A.; WANNER, B. L. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. **Proc Natl Acad Sci U S A**, 97, n. 12, p. 6640-6645, Jun 6 2000.

DE FREITAS, C.; BRIENZO, M. Enzymatic Hydrolysis Applied to Banana Pseudostem Biomass Compared to Solubilized Xylan for Xylooligosaccharides Production with High Substrate Concentration. **BioEnergy Research**, 16, n. 2, p. 1040-1050, 2023/06/01 2023.

DEATHERAGE, D. E.; BARRICK, J. E. High-throughput characterization of mutations in genes that drive clonal evolution using multiplex adaptome capture sequencing. **Cell Syst**, 12, n. 12, p. 1187-1200 e1184, Dec 15 2021.

DELUNA, A.; VETSIGIAN, K.; SHORESH, N.; HEGRENESS, M. *et al.* Exposing the fitness contribution of duplicated genes. **Nat Genet**, 40, n. 5, p. 676-681, May 2008.

DOS SANTOS, L. V.; CARAZZOLLE, M. F.; NAGAMATSU, S. T.; SAMPAIO, N. M. *et al.* Unraveling the genetic basis of xylose consumption in engineered *Saccharomyces cerevisiae* strains. **Sci Rep**, 6, p. 38676, Dec 21 2016.

DRAGOSITS, M.; MATTANOVICH, D. Adaptive laboratory evolution -- principles and applications for biotechnology. **Microb Cell Fact**, 12, p. 64, Jul 1 2013.

DUNNE, K. A.; CHAUDHURI, R. R.; ROSSITER, A. E.; BERIOTTO, I. *et al.* Sequencing a piece of history: complete genome sequence of the original *Escherichia coli* strain. **Microb Genom**, 3, n. 3, p. mgen000106, Mar 2017.

ECHOLS, H.; GREEN, L. Establishment and maintenance of repression by bacteriophage lambda: the role of the *cl*, *cII*, and *c3* proteins. **Proc Natl Acad Sci U S A**, 68, n. 9, p. 2190-2194, Sep 1971.

FANG, H.; KANG, J.; ZHANG, D. Microbial production of vitamin B12: a review and future perspectives. **Microbial Cell Factories**, 16, n. 1, p. 15, 2017/01/30 2017.

FANG, H.; ZHAO, J.; ZHAO, X.; DONG, N. *et al.* Standardized Iterative Genome Editing Method for *Escherichia coli* Based on CRISPR-Cas9. **ACS Synth Biol**, 13, n. 2, p. 613-623, Feb 16 2024.

FENG, X.; ZHAO, D.; ZHANG, X.; DING, X. *et al.* CRISPR/Cas9 Assisted Multiplex Genome Editing Technique in *Escherichia coli*. **Biotechnol J**, 13, n. 9, p. e1700604, Sep 2018.

FOLSOM, J. P.; PARKER, A. E.; CARLSON, R. P. Physiological and proteomic analysis of *Escherichia coli* iron-limited chemostat growth. **J Bacteriol**, 196, n. 15, p. 2748-2761, Aug 2014.

GASSLER, T.; BAUMSCHABL, M.; SALLABERGER, J.; EGERMEIER, M. *et al.* Adaptive laboratory evolution and reverse engineering enhances autotrophic growth in *Pichia pastoris*. **Metab Eng**, 69, p. 112-121, Jan 2022.

GECSE, G.; LABUNSKAITE, R.; PEDERSEN, M.; KILSTRUP, M. *et al.* Minimizing acetate formation from overflow metabolism in *Escherichia coli*: comparison of genetic engineering strategies to improve robustness toward sugar gradients in large-scale fermentation processes. **Front Bioeng Biotechnol**, 12, p. 1339054, 2024.

GETZ, H. P. Sucrose accumulation and synthesis in sugar beet. *In*: GUPTA, A. K. e KAUR, N. (Ed.). **Developments in Crop Science**: Elsevier, 2000. v. 26, p. 55-77.

GORKE, B.; STULKE, J. Carbon catabolite repression in bacteria: many ways to make the most out of nutrients. **Nat Rev Microbiol**, 6, n. 8, p. 613-624, Aug 2008.

GULLBERG, E.; CAO, S.; BERG, O. G.; ILBACK, C. *et al.* Selection of resistant bacteria at very low antibiotic concentrations. **PLoS Pathog**, 7, n. 7, p. e1002158, Jul 2011.

HAN, J. H.; JUNG, S. T.; OH, M. K. Improved Yield of Recombinant Protein via Flagella Regulator Deletion in *Escherichia coli*. **Front Microbiol**, 12, p. 655072, 2021.

HANAHAN, D. J. J. o. m. b. Studies on transformation of *Escherichia coli* with plasmids. 166, n. 4, p. 557-580, 1983.

HERMAN, C.; THEVENET, D.; D'ARI, R.; BOULOC, P. The HflB protease of *Escherichia coli* degrades its inhibitor lambda cIII. **J Bacteriol**, 179, n. 2, p. 358-363, Jan 1997.

HERRING, C. D.; RAGHUNATHAN, A.; HONISCH, C.; PATEL, T. *et al.* Comparative genome sequencing of *Escherichia coli* allows observation of bacterial evolution on a laboratory timescale. **Nat Genet**, 38, n. 12, p. 1406-1412, Dec 2006.

HOBSON, N.; PRICE, N. L.; WARD, J. D.; RAIVIO, T. L. Generation of a restriction minus enteropathogenic *Escherichia coli* E2348/69 strain that is efficiently transformed with large, low copy plasmids. **BMC Microbiol**, 8, p. 134, Aug 5 2008.

IGUCHI, A.; THOMSON, N. R.; OGURA, Y.; SAUNDERS, D. *et al.* Complete genome sequence and comparative genome analysis of enteropathogenic *Escherichia coli* O127:H6 strain E2348/69. **J Bacteriol**, 191, n. 1, p. 347-354, Jan 2009.

JACOBUS, A. P.; BARRETO, J. A.; DE BEM, L. S.; MENEGON, Y. A. *et al.* EasyGuide Plasmids Support in Vivo Assembly of gRNAs for CRISPR/Cas9 Applications in *Saccharomyces cerevisiae*. **ACS Synth Biol**, 11, n. 11, p. 3886-3891, Nov 18 2022.

JACOBUS, A. P.; CAVASSANA, S. D.; DE, O., II; BARRETO, J. A. *et al.* Optimal trade-off between boosted tolerance and growth fitness during adaptive evolution of yeast to ethanol shocks. **Biotechnol Biofuels Bioprod**, 17, n. 1, p. 63, May 10 2024.

JACOBUS, A. P.; GROSS, J. Optimal cloning of PCR fragments by homologous recombination in *Escherichia coli*. **PLoS One**, 10, n. 3, p. e0119221, 2015.

JANG, B. K.; JU, Y.; JEONG, D.; JUNG, S. K. *et al.* L-Lactic Acid Production Using Engineered *Saccharomyces cerevisiae* with Improved Organic Acid Tolerance. **J Fungi (Basel)**, 7, n. 11, Oct 31 2021.

JIANG, N.; WANG, M.; SONG, L.; YU, D. *et al.* Polyhydroxybutyrate production by recombinant *Escherichia coli* based on genes related to synthesis pathway of PHB from *Massilia* sp. UMI-21. **Microbial Cell Factories**, 22, n. 1, p. 129, 2023/07/14 2023.

JIANG, W.; BIKARD, D.; COX, D.; ZHANG, F. *et al.* RNA-guided editing of bacterial genomes using CRISPR-Cas systems. **Nat Biotechnol**, 31, n. 3, p. 233-239, Mar 2013.

JIANG, Y.; CHEN, B.; DUAN, C.; SUN, B. *et al.* Multigene editing in the *Escherichia coli* genome via the CRISPR-Cas9 system. **Appl Environ Microbiol**, 81, n. 7, p. 2506-2514, Apr 2015.

JOHNSON, M. S.; GOPALAKRISHNAN, S.; GOYAL, J.; DILLINGHAM, M. E. *et al.* Phenotypic and molecular evolution across 10,000 generations in laboratory budding yeast populations. **eLife**, 10, p. e63910, 2021/01/19 2021.

JUNG, S. C.; SMITH, C. L.; LEE, K. S.; HONG, M. E. *et al.* Restoration of growth phenotypes of *Escherichia coli* DH5alpha in minimal media through reversal of a point mutation in *purB*. **Appl Environ Microbiol**, 76, n. 18, p. 6307-6309, Sep 2010.

KAPER, J. B.; NATARO, J. P.; MOBLEY, H. L. Pathogenic *Escherichia coli*. **Nat Rev Microbiol**, 2, n. 2, p. 123-140, Feb 2004.

KATAYAMA, T.; KASHO, K.; KAWAKAMI, H. The DnaA Cycle in *Escherichia coli*: Activation, Function and Inactivation of the Initiator Protein. **Front Microbiol**, 8, p. 2496, 2017.

KIHARA, A.; AKIYAMA, Y.; ITO, K. A protease complex in the *Escherichia coli* plasma membrane: HflKC (HflA) forms a complex with FtsH (HflB), regulating its proteolytic activity against SecY. **EMBO J**, 15, n. 22, p. 6122-6131, Nov 15 1996.

KIM, K.; HOU, C. Y.; CHOE, D.; KANG, M. *et al.* Adaptive laboratory evolution of *Escherichia coli* W enhances gamma-aminobutyric acid production using glycerol as the carbon source. **Metab Eng**, 69, p. 59-72, Jan 2022.

KOSTYLEV, M.; OTWELL, A. E.; RICHARDSON, R. E.; SUZUKI, Y. Cloning Should Be Simple: *Escherichia coli* DH5alpha-Mediated Assembly of Multiple DNA Fragments with Short End Homologies. **PLoS One**, 10, n. 9, p. e0137466, 2015.

LEE, J. W.; CHOI, S.; PARK, J. H.; VICKERS, C. E. *et al.* Development of sucrose-utilizing *Escherichia coli* K-12 strain by cloning beta-fructofuranosidases and its application for L-threonine production. **Appl Microbiol Biotechnol**, 88, n. 4, p. 905-913, Oct 2010.

LEE, M. J.; PARK, J.; PARK, K.; KIM, J. F. *et al.* Reverse Engineering Targets for Recombinant Protein Production in *Corynebacterium glutamicum* Inspired by a Fast-Growing Evolved Descendant. **Front Bioeng Biotechnol**, 8, p. 588070, 2020.

LEE, S. Y. Bacterial polyhydroxyalkanoates. **Biotechnol Bioeng**, 49, n. 1, p. 1-14, Jan 5 1996.

LEI, D.; QIU, Z.; WU, J.; QIAO, B. *et al.* Combining Metabolic and Monoterpene Synthase Engineering for de Novo Production of Monoterpene Alcohols in *Escherichia coli*. **ACS Synth Biol**, 10, n. 6, p. 1531-1544, Jun 18 2021.

LENSKI, R. E. Experimental evolution and the dynamics of adaptation and genome evolution in microbial populations. **ISME J**, 11, n. 10, p. 2181-2194, Oct 2017.

LENSKI, R. E. Revisiting the Design of the Long-Term Evolution Experiment with *Escherichia coli*. **J Mol Evol**, 91, n. 3, p. 241-253, Jun 2023.

LENSKI, R. E.; MONGOLD, J. A.; SNIEGOWSKI, P. D.; TRAVISANO, M. *et al.* Evolution of competitive fitness in experimental populations of *E. coli*: what makes one genotype a better competitor than another? **Antonie Van Leeuwenhoek**, 73, n. 1, p. 35-47, Jan 1998.

LENSKI, R. E.; ROSE, M. R.; SIMPSON, S. C.; TADLER, S. C. Long-Term Experimental Evolution in *Escherichia coli*. I. Adaptation and Divergence During 2,000 Generations. **The American Naturalist**, 138, n. 6, p. 1315-1341, 1991.

LI, B.; LIU, N.; ZHAO, X. Response mechanisms of *Saccharomyces cerevisiae* to the stress factors present in lignocellulose hydrolysate and strategies for constructing robust strains. **Biotechnol Biofuels Bioprod**, 15, n. 1, p. 28, Mar 15 2022.

LI, M.; WANG, J.; GENG, Y.; LI, Y. *et al.* A strategy of gene overexpression based on tandem repetitive promoters in *Escherichia coli*. **Microb Cell Fact**, 11, p. 19, Feb 6 2012.

LI, M.; WILKINS, M. R. Recent advances in polyhydroxyalkanoate production: Feedstocks, strains and process developments. **Int J Biol Macromol**, 156, p. 691-703, Aug 1 2020.

LI, Q.; SUN, B.; CHEN, J.; ZHANG, Y. *et al.* A modified pCas/pTargetF system for CRISPR-Cas9-assisted genome editing in *Escherichia coli*. **Acta Biochim Biophys Sin (Shanghai)**, 53, n. 5, p. 620-627, Apr 15 2021.

LI, Q.; SUN, M.; LV, L.; ZUO, Y. *et al.* Improving the Editing Efficiency of CRISPR-Cas9 by Reducing the Generation of Escapers Based on the Surviving Mechanism. **ACS Synth Biol**, 12, n. 3, p. 672-680, Mar 17 2023.

LI, Y.; LIN, Z.; HUANG, C.; ZHANG, Y. *et al.* Metabolic engineering of *Escherichia coli* using CRISPR-Cas9 mediated genome editing. **Metab Eng**, 31, p. 13-21, Sep 2015.

LI, Z.; YANG, J.; LOH, X. J. Polyhydroxyalkanoates: opening doors for a sustainable future. **NPG Asia Materials**, 8, n. 4, p. e265-e265, 2016/04/01 2016.

LI, Z. H.; WANG, J.; XU, J. P.; WANG, J. *et al.* Recent advances in CRISPR-based genome editing technology and its applications in cardiovascular research. **Mil Med Res**, 10, n. 1, p. 12, Mar 10 2023.

LIANG, L.; LIU, R.; FREED, E. F.; ECKERT, C. A. Synthetic Biology and Metabolic Engineering Employing *Escherichia coli* for C2-C6 Bioalcohol Production. **Front Bioeng Biotechnol**, 8, p. 710, 2020.

LILIC, M.; JOVANOVIĆ, M.; JOVANOVIĆ, G.; SAVIC, D. J. Identification of the CysB-regulated gene, hslJ, related to the Escherichia coli novobiocin resistance phenotype. **FEMS Microbiol Lett**, 224, n. 2, p. 239-246, Jul 29 2003.

LIM, S. R.; LEE, H. J.; KIM, H. J.; LEE, S. J. Multiplex Single-Nucleotide Microbial Genome Editing Achieved by CRISPR-Cas9 Using 5'-End-Truncated sgRNAs. **ACS Synth Biol**, 12, n. 7, p. 2203-2207, Jul 21 2023.

LIN, P.-C.; ZHANG, F.; PAKRASI, H. B. Enhanced production of sucrose in the fast-growing cyanobacterium Synechococcus elongatus UTEX 2973. **Scientific Reports**, 10, n. 1, p. 390, 2020/01/15 2020.

LOPILATO, J.; BORTNER, S.; BECKWITH, J. Mutations in a new chromosomal gene of Escherichia coli K-12, pcnB, reduce plasmid copy number of pBR322 and its derivatives. **Mol Gen Genet**, 205, n. 2, p. 285-290, Nov 1986.

MA, D.; CAMPBELL, J. L. The effect of dnaA protein and n' sites on the replication of plasmid ColE1. **J Biol Chem**, 263, n. 29, p. 15008-15015, Oct 15 1988.

MALDONADO, N.; LOPEZ-HERNANDEZ, I.; GARCIA-MONTANER, A.; LOPEZ-CORTES, L. E. *et al.* Whole-genome characterisation of Escherichia coli isolates from patients with bacteraemia presenting with sepsis or septic shock in Spain: a multicentre cross-sectional study. **Lancet Microbe**, 5, n. 4, p. e390-e399, Apr 2024.

MANIATIS, T. J. A. I. m. Molecular cloning. 1982.

MASTERS, M.; COLLOMS, M. D.; OLIVER, I. R.; HE, L. *et al.* The pcnB gene of Escherichia coli, which is required for ColE1 copy number maintenance, is dispensable. **J Bacteriol**, 175, n. 14, p. 4405-4413, Jul 1993.

MCADAM, B.; BRENNAN FOURNET, M.; MCDONALD, P.; MOJICEVIC, M. Production of Polyhydroxybutyrate (PHB) and Factors Impacting Its Chemical and Mechanical Characteristics. 12, n. 12, p. 2908, 2020.

MENEGON, Y. A.; GROSS, J.; JACOBUS, A. P. How adaptive laboratory evolution can boost yeast tolerance to lignocellulosic hydrolyses. **Curr Genet**, 68, n. 3-4, p. 319-342, Aug 2022.

MOHAMED, E. T.; MUNDHADA, H.; LANDBERG, J.; CANN, I. *et al.* Generation of an E. coli platform strain for improved sucrose utilization using adaptive laboratory evolution. **Microb Cell Fact**, 18, n. 1, p. 116, Jun 29 2019.

MONDAY, S. R.; MINNICH, S. A.; FENG, P. C. A 12-base-pair deletion in the flagellar master control gene flhC causes nonmotility of the pathogenic German sorbitol-fermenting Escherichia coli O157:H- strains. **J Bacteriol**, 186, n. 8, p. 2319-2327, Apr 2004.

MOREB, E. A.; HOOVER, B.; YASEEN, A.; VALYASEVI, N. *et al.* Managing the SOS Response for Enhanced CRISPR-Cas-Based Recombineering in E. coli through Transient Inhibition of Host RecA Activity. **ACS Synth Biol**, 6, n. 12, p. 2209-2218, Dec 15 2017.

MORENO, A. D.; GONZÁLEZ-FERNÁNDEZ, C.; TOMÁS-PEJÓ, E. Insights into cell robustness against lignocellulosic inhibitors and insoluble solids in bioethanol production processes. **Scientific Reports**, 12, n. 1, p. 557, 2022/01/11 2022.

MURPHY, K. C. Use of bacteriophage lambda recombination functions to promote gene replacement in *Escherichia coli*. **J Bacteriol**, 180, n. 8, p. 2063-2071, Apr 1998.

NADRATOWSKA-WESOŁOWSKA, B.; SŁOMINSKA-WOJEWODZKA, M.; LYZEN, R.; WĘGRZYN, A. *et al.* Transcription regulation of the *Escherichia coli* *pcnB* gene coding for poly(A) polymerase I: roles of ppGpp, DksA and sigma factors. **Mol Genet Genomics**, 284, n. 4, p. 289-305, Oct 2010.

NI, D.; CHEN, Z.; TIAN, Y.; XU, W. *et al.* Comprehensive utilization of sucrose resources via chemical and biotechnological processes: A review. **Biotechnology Advances**, 60, p. 107990, 2022/11/01/ 2022.

NOZAKI, S.; NIKI, H. Exonuclease III (XthA) Enforces In Vivo DNA Cloning of *Escherichia coli* To Create Cohesive Ends. **J Bacteriol**, 201, n. 5, Mar 1 2019.

OLAVARRIA, K.; FINA, A.; VELASCO, M. I.; VAN LOOSDRECHT, M. C. M. *et al.* Metabolism of sucrose in a non-fermentative *Escherichia coli* under oxygen limitation. **Appl Microbiol Biotechnol**, 103, n. 15, p. 6245-6256, Aug 2019.

OLINER, J. D.; KINZLER, K. W.; VOGELSTEIN, B. In vivo cloning of PCR products in *E. coli*. **Nucleic Acids Res**, 21, n. 22, p. 5192-5197, Nov 11 1993.

ODD, B.; FLORES, C. L.; GANCEDO, C.; ZHANG, X. *et al.* An internal deletion in MTH1 enables growth on glucose of pyruvate-decarboxylase negative, non-fermentative *Saccharomyces cerevisiae*. **Microb Cell Fact**, 11, p. 131, Sep 15 2012.

PANDEY, A.; ADAMA, N.; ADJALLE, K.; BLAIS, J. F. Sustainable applications of polyhydroxyalkanoates in various fields: A critical review. **Int J Biol Macromol**, 221, p. 1184-1201, Nov 30 2022.

PATEL, U. R.; GAUTAM, S.; CHATTERJI, D. Unraveling the Role of Silent Mutation in the omega-Subunit of *Escherichia coli* RNA Polymerase: Structure Transition Inhibits Transcription. **ACS Omega**, 4, n. 18, p. 17714-17725, Oct 29 2019.

PENA, C.; CASTILLO, T.; GARCIA, A.; MILLAN, M. *et al.* Biotechnological strategies to improve production of microbial poly-(3-hydroxybutyrate): a review of recent research work. **Microb Biotechnol**, 7, n. 4, p. 278-293, Jul 2014.

PEÑA, C.; LÓPEZ, S.; GARCÍA, A.; ESPÍN, G. *et al.* Biosynthesis of poly- β -hydroxybutyrate (PHB) with a high molecular mass by a mutant strain of *Azotobacter vinelandii* (OPN). **Annals of Microbiology**, 64, n. 1, p. 39-47, 2014/03/01 2014.

PYNE, M. E.; MOO-YOUNG, M.; CHUNG, D. A.; CHOU, C. P. Coupling the CRISPR/Cas9 System with Lambda Red Recombineering Enables Simplified Chromosomal Gene Replacement in *Escherichia coli*. **Appl Environ Microbiol**, 81, n. 15, p. 5103-5114, Aug 2015.

- RAFFAELLI, N.; LORENZI, T.; MARIANI, P. L.; EMANUELLI, M. *et al.* The Escherichia coli NadR regulator is endowed with nicotinamide mononucleotide adenyltransferase activity. **J Bacteriol**, 181, n. 17, p. 5509-5511, Sep 1999.
- REID, C. J.; BLAU, K.; JECHALKE, S.; SMALLA, K. *et al.* Whole Genome Sequencing of Escherichia coli From Store-Bought Produce. **Front Microbiol**, 10, p. 3050, 2019.
- REISCH, C. R.; PRATHER, K. L. The no-SCAR (Scarless Cas9 Assisted Recombineering) system for genome editing in Escherichia coli. **Sci Rep**, 5, p. 15096, Oct 14 2015.
- RIGGS, A. D. Making, Cloning, and the Expression of Human Insulin Genes in Bacteria: The Path to Humulin. **Endocr Rev**, 42, n. 3, p. 374-380, May 25 2021.
- RODRIGUEZ, A.; MARTINEZ, J. A.; FLORES, N.; ESCALANTE, A. *et al.* Engineering Escherichia coli to overproduce aromatic amino acids and derived compounds. **Microb Cell Fact**, 13, n. 1, p. 126, Sep 9 2014.
- RONDA, C.; PEDERSEN, L. E.; SOMMER, M. O.; NIELSEN, A. T. CRMAGE: CRISPR Optimized MAGE Recombineering. **Sci Rep**, 6, p. 19452, Jan 22 2016.
- ROSANO, G. L.; CECCARELLI, E. A. Recombinant protein expression in Escherichia coli: advances and challenges. **Front Microbiol**, 5, p. 172, 2014.
- RUIZ, N.; SILHAVY, T. J. How Escherichia coli Became the Flagship Bacterium of Molecular Biology. 204, n. 9, p. e00230-00222, 2022.
- SABRI, S.; NIELSEN, L. K.; VICKERS, C. E. Molecular control of sucrose utilization in Escherichia coli W, an efficient sucrose-utilizing strain. **Appl Environ Microbiol**, 79, n. 2, p. 478-487, Jan 2013.
- SANTOS, T. M. A.; LAMMERS, M. G.; ZHOU, M.; SPARKS, I. L. *et al.* Small Molecule Chelators Reveal That Iron Starvation Inhibits Late Stages of Bacterial Cytokinesis. **ACS Chem Biol**, 13, n. 1, p. 235-246, Jan 19 2018.
- SCHWECHHEIMER, C.; KULP, A.; KUEHN, M. J. Modulation of bacterial outer membrane vesicle production by envelope structure and content. **BMC Microbiol**, 14, p. 324, Dec 21 2014.
- SHABBIR, M. A. B.; SHABBIR, M. Z.; WU, Q.; MAHMOOD, S. *et al.* CRISPR-cas system: biological function in microbes and its use to treat antimicrobial resistant pathogens. **Ann Clin Microbiol Antimicrob**, 18, n. 1, p. 21, Jul 5 2019.
- SHARMA, V.; SEHGAL, R.; GUPTA, R. Polyhydroxyalkanoate (PHA): Properties and Modifications. **Polymer**, 212, p. 123161, 2021/01/06/ 2021.
- SHINOHARA, A.; OGAWA, H.; OGAWA, T. Rad51 protein involved in repair and recombination in S. cerevisiae is a RecA-like protein. **Cell**, 69, n. 3, p. 457-470, May 1 1992.
- SHUKAL, S.; LIM, X. H.; ZHANG, C.; CHEN, X. Metabolic engineering of Escherichia coli BL21 strain using simplified CRISPR-Cas9 and asymmetric homology arms recombineering. **Microb Cell Fact**, 21, n. 1, p. 19, Feb 5 2022.

SOUTOURINA, O. A.; BERTIN, P. N. Regulation cascade of flagellar expression in Gram-negative bacteria. **FEMS Microbiol Rev**, 27, n. 4, p. 505-523, Oct 2003.

SPIRA, B.; OSPINO, K. Diversity in *E. coli* (p)ppGpp Levels and Its Consequences. 11, 2020-August-12 2020. Review.

STOVICEK, V.; DATO, L.; ALMQVIST, H.; SCHOPPING, M. *et al.* Rational and evolutionary engineering of *Saccharomyces cerevisiae* for production of dicarboxylic acids from lignocellulosic biomass and exploring genetic mechanisms of the yeast tolerance to the biomass hydrolysate. **Biotechnol Biofuels Bioprod**, 15, n. 1, p. 22, Feb 27 2022.

SUN, D.; WANG, L.; MAO, X.; FEI, M. *et al.* Chemical transformation mediated CRISPR/Cas9 genome editing in *Escherichia coli*. **Biotechnol Lett**, 41, n. 2, p. 293-303, Feb 2019.

SUN, H.; WANG, M.; LIU, Y.; WU, P. *et al.* Regulation of flagellar motility and biosynthesis in enterohemorrhagic *Escherichia coli* O157:H7. **Gut Microbes**, 14, n. 1, p. 2110822, Jan-Dec 2022.

SUTHERLAND, C.; MURAKAMI, K. S. An Introduction to the Structure and Function of the Catalytic Core Enzyme of *Escherichia coli* RNA Polymerase. **EcoSal Plus**, 8, n. 1, Aug 2018.

SUZUKI, M.; TACHIBANA, Y.; KASUYA, K.-i. Biodegradability of poly(3-hydroxyalkanoate) and poly(ϵ -caprolactone) via biological carbon cycles in marine environments. **Polymer Journal**, 53, n. 1, p. 47-66, 2021/01/01 2021.

TANG, Q.; LOU, C.; LIU, S. J. Construction of an easy-to-use CRISPR-Cas9 system by patching a newly designed EXIT circuit. **J Biol Eng**, 11, p. 32, 2017.

TONG, Y.; JORGENSEN, T. S.; WHITFORD, C. M.; WEBER, T. *et al.* A versatile genetic engineering toolkit for *E. coli* based on CRISPR-prime editing. **Nat Commun**, 12, n. 1, p. 5206, Sep 1 2021.

TRAKUNJAE, C.; BOONDAENG, A.; APIWATANAPIWAT, W.; KOSUGI, A. *et al.* Enhanced polyhydroxybutyrate (PHB) production by newly isolated rare actinomycetes *Rhodococcus* sp. strain BSRT1-1 using response surface methodology. **Scientific Reports**, 11, n. 1, p. 1896, 2021/01/21 2021.

TRAXLER, M. F.; SUMMERS, S. M.; NGUYEN, H. T.; ZACHARIA, V. M. *et al.* The global, ppGpp-mediated stringent response to amino acid starvation in *Escherichia coli*. **Mol Microbiol**, 68, n. 5, p. 1128-1148, Jun 2008.

VILA, J.; SAEZ-LOPEZ, E.; JOHNSON, J. R.; ROMLING, U. *et al.* *Escherichia coli*: an old friend with new tidings. **FEMS Microbiol Rev**, 40, n. 4, p. 437-463, Jul 1 2016.

VRENTAS, C. E.; GAAL, T.; ROSS, W.; EBRIGHT, R. H. *et al.* Response of RNA polymerase to ppGpp: requirement for the omega subunit and relief of this requirement by DksA. **Genes Dev**, 19, n. 19, p. 2378-2387, Oct 1 2005.

WACLAWOVSKY, A. J.; SATO, P. M.; LEMBKE, C. G.; MOORE, P. H. *et al.* Sugarcane for bioenergy production: an assessment of yield and regulation of sucrose content. **Plant Biotechnol J**, 8, n. 3, p. 263-276, Apr 2010.

WADA, T.; HATAMOTO, Y.; KUTSUKAKE, K. Functional and expressional analyses of the anti-FlhD4C2 factor gene *ydiV* in *Escherichia coli*. **Microbiology (Reading)**, 158, n. Pt 6, p. 1533-1542, Jun 2012.

WANG, G.; LI, Q.; ZHANG, Z.; YIN, X. *et al.* Recent progress in adaptive laboratory evolution of industrial microorganisms. **J Ind Microbiol Biotechnol**, 50, n. 1, Feb 17 2023.

WANG, J.; SUI, X.; DING, Y.; FU, Y. *et al.* A fast and robust iterative genome-editing method based on a Rock-Paper-Scissors strategy. **Nucleic Acids Res**, 49, n. 2, p. e12, Jan 25 2021.

WATSON, J. F.; GARCIA-NAFRIA, J. In vivo DNA assembly using common laboratory bacteria: A re-emerging tool to simplify molecular cloning. **J Biol Chem**, 294, n. 42, p. 15271-15281, Oct 18 2019.

WEI, J.; LI, Y. CRISPR-based gene editing technology and its application in microbial engineering. **Eng Microbiol**, 3, n. 4, p. 100101, Dec 2023.

WU, F.; ZHOU, Y.; PEI, W.; JIANG, Y. *et al.* Biosynthesis of Poly-(3-hydroxybutyrate) under the Control of an Anaerobically Induced Promoter by Recombinant *Escherichia coli* from Sucrose. **Molecules**, 27, n. 1, Jan 4 2022.

YAMAGUCHI, K.; INOUE, M. Lipoprotein 28, an inner membrane protein of *Escherichia coli* encoded by *nlpA*, is not essential for growth. **J Bacteriol**, 170, n. 8, p. 3747-3749, Aug 1988.

YAMAMOTO, K.; YAMANAKA, Y.; SHIMADA, T.; SARKAR, P. *et al.* Altered Distribution of RNA Polymerase Lacking the Omega Subunit within the Prophages along the *Escherichia coli* K-12 Genome. **mSystems**, 3, n. 1, Jan-Feb 2018.

YIN, J.; WANG, H.; FU, X. Z.; GAO, X. *et al.* Effects of chromosomal gene copy number and locations on polyhydroxyalkanoate synthesis by *Escherichia coli* and *Halomonas* sp. **Appl Microbiol Biotechnol**, 99, n. 13, p. 5523-5534, Jul 2015.

ZERBINI, F.; ZANELLA, I.; FRACCASCIA, D.; KONIG, E. *et al.* Large scale validation of an efficient CRISPR/Cas-based multi gene editing protocol in *Escherichia coli*. **Microb Cell Fact**, 16, n. 1, p. 68, Apr 24 2017.

ZHANG, Y.; MUYRERS, J. P.; TESTA, G.; STEWART, A. F. DNA cloning by homologous recombination in *Escherichia coli*. **Nat Biotechnol**, 18, n. 12, p. 1314-1317, Dec 2000.

ZHAO, D.; YUAN, S.; XIONG, B.; SUN, H. *et al.* Development of a fast and easy method for *Escherichia coli* genome editing with CRISPR/Cas9. **Microb Cell Fact**, 15, n. 1, p. 205, Dec 1 2016.

ZHAO, J.; BABA, T.; MORI, H.; SHIMIZU, K. Effect of *zwf* gene knockout on the metabolism of *Escherichia coli* grown on glucose or acetate. **Metab Eng**, 6, n. 2, p. 164-174, Apr 2004.

ZHAO, W.; GUO, Y. Increasing the efficiency of gene editing with CRISPR-Cas9 via concurrent expression of the Beta protein. **Int J Biol Macromol**, 270, n. Pt 2, p. 132431, Jun 2024.

ZHENG, W.; LI, Z.; SKARSTAD, K.; CROOKE, E. Mutations in DnaA protein suppress the growth arrest of acidic phospholipid-deficient Escherichia coli cells. **EMBO J**, 20, n. 5, p. 1164-1172, Mar 1 2001.

APÊNDICE A- Linhagens utilizadas no trabalho

Strain	Genotype	Description
<i>E. coli</i> DH5α	F– endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG purB20 φ80dlacZΔM15 Δ(lacZYA-argF) U169, hsdR17(rK–mK+), λ–	Cloning and maintenance of plasmids; testing of genetic modifications; conduction of the evolution experiment
<i>E. coli</i> E2348/69	Wild-type EPEC, Serotype O127:H6	Cloning of the <i>cscBKA</i> operon; conducting evolution experiments
<i>S. cerevisiae</i> CEN.PK113-7D	MATa; prototrophic laboratory yeast strain	Used for plasmid assembly for donor amplification for CRISPR/Cas9
DH5α pCscBKA	<i>E. coli</i> DH5α with pCscBKA plasmid	Cells with plasmid pCscBKA for consumption of sucrose
D-Kan	Evolved population from <i>E. coli</i> DH5α with pCscBKA and MV_Kan plasmids	Final populations of evolution experiment
D-Spec	Evolved population from <i>E. coli</i> DH5α with pCscBKA and MV_Spec plasmids	Final populations of evolution experiment
D-Cm	Evolved population from <i>E. coli</i> DH5α with pCscBKA and pGGA plasmids	Final populations of evolution experiment
E-Kan	Evolved population from <i>E. coli</i> E2348/69 with MV_Kan plasmid	Final populations of evolution experiment
E-Spec	Evolved population from <i>E. coli</i> E2348/69 with MV_Spec plasmid	Final populations of evolution experiment
E-Amp	Evolved population from <i>E. coli</i> E2348/69 with MV_Amp plasmid	Final populations of evolution experiment
E-Cm	Evolved population from <i>E. coli</i> E2348/69 with pGGA plasmid	Final populations of evolution experiment
<i>cscBKA</i>	<i>E. coli</i> DH5α SS9::cscBKA	Base Strain for construction of all genomic mutations
Tester-GFP	<i>E. coli</i> DH5α SS9::cscBKA <i>galK::gfp lacZ::gfp</i>	Comparative Strain for competition experiment

<i>nudJ</i> (G151V)	<i>E. coli</i> DH5 α SS9:: <i>cscBKA nudJ</i> (G151V)	Strain with <i>cscBKA</i> operon inserted in SS9 intergenic region with pontual mutation in <i>nudJ</i> gene
<i>purB</i> (L115G)	<i>E. coli</i> DH5 α SS9:: <i>cscBKA purB</i> (L115G)	Strain with <i>cscBKA</i> operon inserted in SS9 intergenic region with pontual mutation in <i>purB</i> gene
Δ 40.4 kb	<i>E. coli</i> DH5 α SS9:: <i>cscBKA</i> Δ 40 kb	Strain with <i>cscBKA</i> operon inserted in SS9 intergenic region with deletion of 40.4 kb of partial region for flagellum synthesis
Δ <i>fliA</i> (205 bp)	<i>E. coli</i> DH5 α SS9:: <i>cscBKA</i> Δ <i>fliA</i> (205 bp)	Strain with <i>cscBKA</i> operon inserted in SS9 intergenic region with deletion of 205 bp of <i>fliA</i> gene
Δ <i>flhDC</i>	<i>E. coli</i> DH5 α SS9:: <i>cscBKA</i> Δ <i>flhDC</i>	Strain with <i>cscBKA</i> operon inserted in SS9 intergenic region with complete deletion of <i>flhDC</i> genes
<i>dnaK</i> (T417N)	<i>E. coli</i> DH5 α SS9:: <i>cscBKA dnaK</i> (T417N)	Strain with <i>cscBKA</i> operon inserted in SS9 intergenic region with pontual mutation in <i>dnaK</i> gene
<i>zwf</i> (D405E)	<i>E. coli</i> DH5 α SS9:: <i>cscBKA zwf</i> (D405E)	Strain with <i>cscBKA</i> operon inserted in SS9 intergenic region with pontual mutation in <i>zwf</i> gene
Δ <i>hflK</i>	<i>E. coli</i> DH5 α SS9:: <i>cscBKA</i> Δ <i>hflK</i>	Strain with <i>cscBKA</i> operon inserted in SS9 intergenic region with complete deletion of Δ <i>hflK</i> gene
Δ CA	<i>E. coli</i> DH5 α SS9:: <i>cscBKA</i> Δ 12 kb	Strain with <i>cscBKA</i> operon inserted in SS9 intergenic region with deletion of 12kb for colanic acid synthesis
Δ <i>pcnB</i> (81 bp)	<i>E. coli</i> DH5 α SS9:: <i>cscBKA</i> Δ <i>pcnB</i> (81 bp)	Strain with <i>cscBKA</i> operon inserted in SS9 intergenic region with deletion of 81 bp o codificant region of <i>pcnB</i> gene
<i>cscBKsucP</i>	<i>E. coli</i> DH5 α SS9:: <i>cscBKsucP</i>	Strain with <i>cscBKsucP</i> operon inserted in SS9 intergenic region
<i>cscBK-PJ-sucP</i>	<i>E. coli</i> DH5 α SS9:: <i>cscBK-PJ-sucP</i>	Strain with <i>cscBK-PJ-sucP</i> operon inserted in SS9 intergenic region. The expression <i>sucP</i> is controled by

		action of natural promoter of <i>csc</i> operon and with J23119 promoter
2x <i>cscBKA</i>	<i>E. coli</i> DH5α SS9:: <i>cscBKA galK::cscBKA</i>	Strain with two insertion of <i>cscBKA</i> operon (SS9 intergenic region and <i>galK</i> gene)
<i>cscBKA</i> + <i>sucP</i>	<i>E. coli</i> DH5α SS9:: <i>cscBKA lacZ::sucP</i>	Strain with <i>cscBKA</i> operon inserted in SS9 intergenic region and <i>sucP</i> on <i>lacZ</i>
DH5α pGFP	<i>E. coli</i> DH5α pGFP	Strain with pGFP plasmid, without genomic mutations
$\Delta pcnB$ pGFP	<i>E. coli</i> DH5α $\Delta pcnB$ (81 bp) pGFP	Strain with pGFP plasmid and 81 bp deletion of <i>pcnB</i> gene
Evolved clone D-Spec pGFP	Clone from Evolved population from <i>E. coli</i> DH5α with pCscBKA and MV_Spec plasmids pGFP	Colony isolated after curing process of both plasmids from D-Spec population expressing pGFP plasmid
$\Delta pcnB$ pCscBKA	<i>E. coli</i> DH5α $\Delta pcnB$ (81 bp) pCscBKA	Strain with pGFP plasmid and 81 bp deletion of <i>pcnB</i> gene
Evolved clone D-Spec pCscBKA	Clone from Evolved population from <i>E. coli</i> DH5α with pCscBKA and MV_Spec plasmids pCscBKA	Colony isolated after curing process of MV_Spec plasmid from D-Spec population expressing pCscBKA plasmid

APÊNCIDE B – Plasmídeos usados neste trabalho

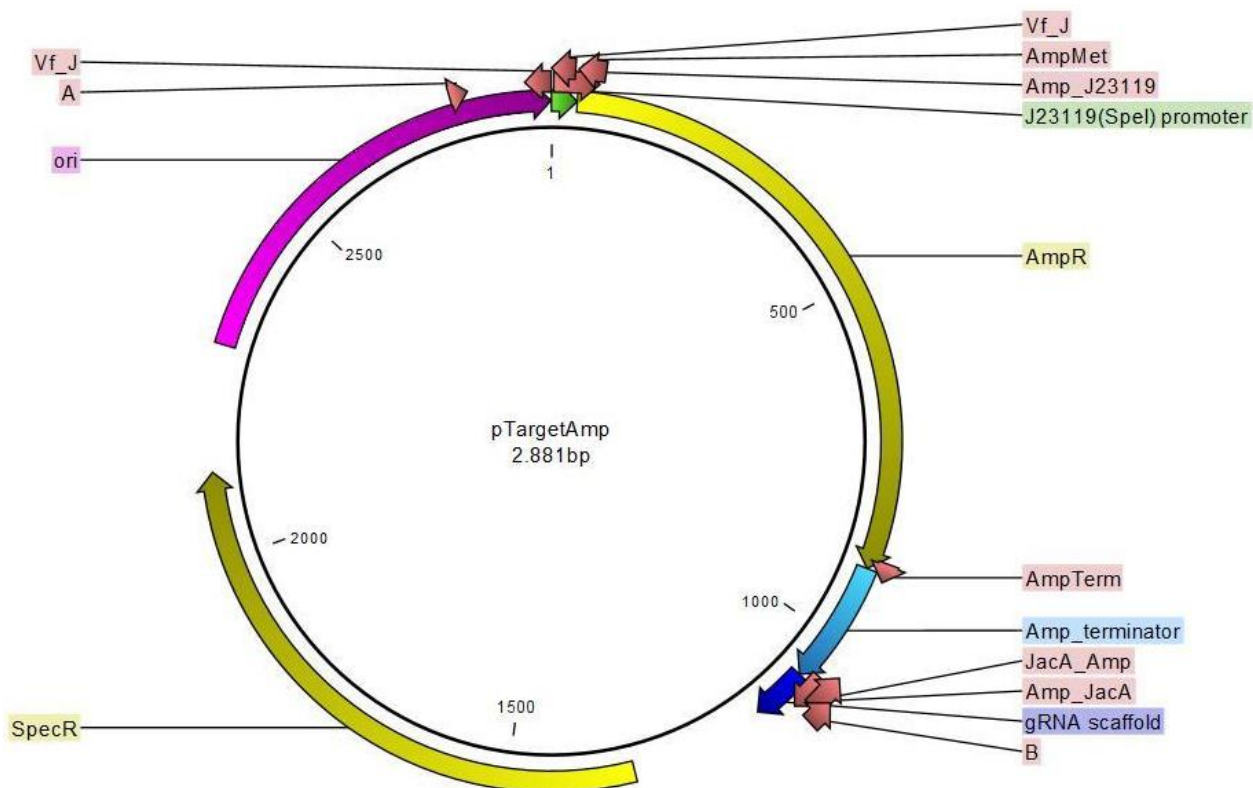
Plasmid	Used for / characteristics	Source
pTargeF	Plasmid used for cloning pTargeg_Amp and MV-Spec. Spectinomycin resistance marker.	Jiang et al. (2015)
pTarget_Amp	Plasmid used for generation of plasmids for gRNA expression; Marker for resistance to Spectinomycin and Ampicillin.	This study
pUC19	Used for cloning pUC19_B0015 and MV-Amp. Ampicillin resistance marker.	
pUC19_B0015	It has an IPTG-inducible promoter and the B0015 terminator. Used to host the mEGFP sequence. Marker for Ampicillin resistance.	This study
pUC57_sucP	Commercially synthesized plasmid with the coding sequence for the <i>Bifidobacterium adolescentis</i> <i>sucP</i> gene (codon optimized for <i>E. coli</i>). Under the action of the J23119 promoter and B0015 terminator, this construct was flanked by two 400 bp regions of the SS9 genomic region. Ampicillin resistance marker.	This study
pUC57_cscBKA	Cloned from pUC57_sucP, with <i>cscBKA</i> operon in place of the <i>sucP</i> gene. Ampicillin resistance marker.	This study
MV-Kan	Cloned from pUC57, possessing only the origin of replication and the resistance kanamycin.	This study
MV-Spec	Cloned from pTargeF, possessing only the origin of replication and the Spectinomycin resistance marker.	This study
MV-Amp	Cloned from pUC19, possessing only the origin of replication and the Ampicillin resistance marker.	This study
pGGA	Plasmid present in the Golden Gate Assembly (NEB) cloning kit. Chloramphenicol resistance marker.	New England Biolabs Inc.
CAS9BAC1P-1EA	Commercial plasmid containing coding sequence for Cas9 and Lambda Red recombinases. It is inducible by L-arabinose, with replication at 30 °C. Marker of resistance to kanamycin.	Sigma Aldrich

mEGFP_pBAD	Plasmid containing the sequence of the GFP version mEGFP, under the action of an arabinose-inducible promoter. Resistance to Ampicillin.	Michael Davidson/Addgene (não publicado)
pUC19_GFP	Plasmid used to host the mEGFP sequence, under the action of the lac operon promoter, inducible by IPTG. Resistance to Ampicillin.	This study
pEasyG3-mic	Plasmid used for amplification of the origin of replication to be used in cloning of the plasmids used for donor assembly in yeast. Chloramphenicol resistance.	Jacobus et al. (2022)
pEasyG3-Nat	Plasmid used for amplification of the nourseothricin resistance marker to be used in cloning the plasmids used for donor assembly in yeast. Chloramphenicol resistance.	Jacobus et al. (2022)
pEasyG3-Hph	Plasmid used for amplification of the hygromycin resistance marker to be used in cloning the plasmids used for donor assembly in yeast. Chloramphenicol resistance.	Jacobus et al. (2022)
pMIC_cscKsucP	Plasmid cloned in yeast for the assembly of the fragment to be used as a donor for insertion of the <i>sucP</i> gene into the <i>E. coli</i> genome. Resistance to nourseothricin (yeast).	This study
pMIC_cscKsucP(J23119)	Plasmid cloned in yeast for the assembly of the fragment to be used as a donor for insertion of the <i>sucP</i> gene together with the J23119 promoter into the <i>E. coli</i> genome. Resistance to nourseothricin (yeast).	This study
pMIC_SNP_nudJ	Plasmid cloned in yeast for the assembly of the fragment to be used as a donor for base change in the <i>E. coli</i> genome. Resistance to nourseothricin (yeast).	This study
pMIC_SNP_purB	Plasmid cloned in yeast for the assembly of the fragment to be used as a donor for base change in the <i>E. coli</i> genome. Resistance to hygromycin (yeast).	This study
pSpec5'	Precursor plasmid for cloning plasmids for gRNA expression. Resistance to Ampicillin.	This study
pSpec3'	Precursor plasmid for cloning plasmids for gRNA expression. Resistance to Ampicillin.	This study

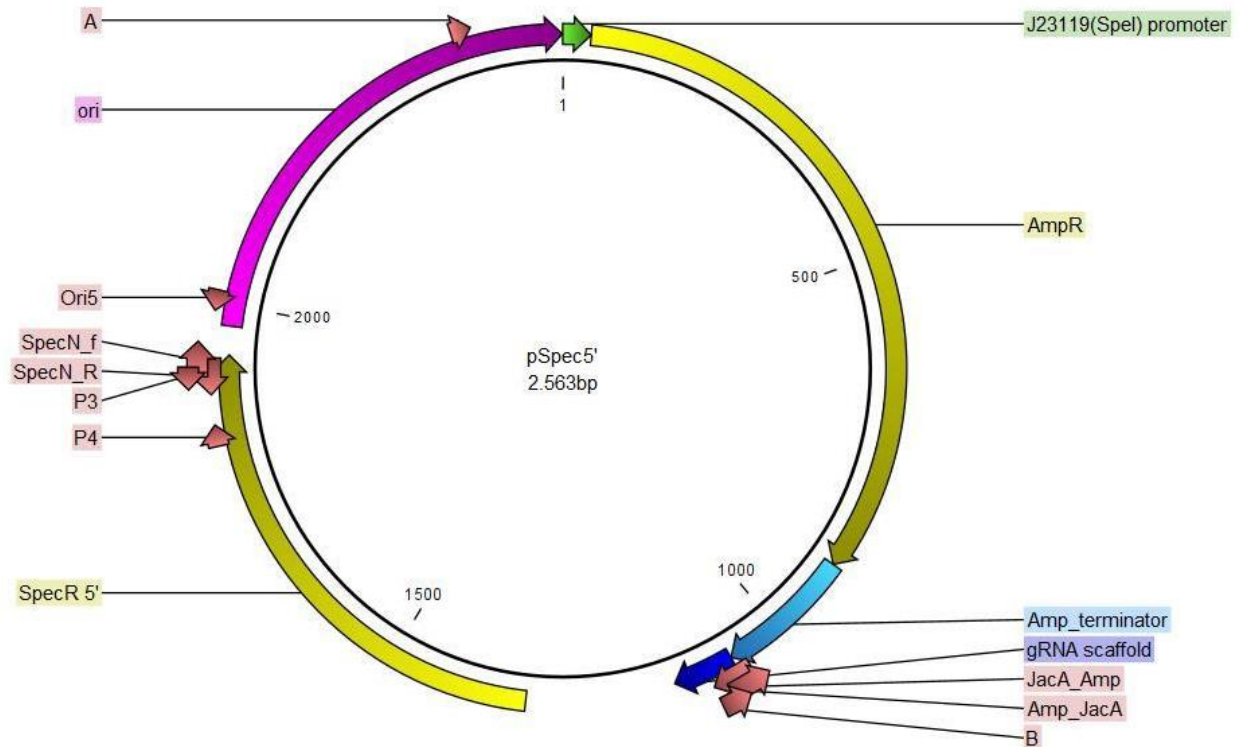
pOliSpec5'	Precursor plasmid for cloning plasmids for gRNA expression. Resistance to Chloramphenicol .	This study
pOliSpec3'	Precursor plasmid for cloning plasmids for gRNA expression. Resistance to Ampicillin.	This study
pMIC_GFP_ΔgalK	Plasmid cloned in yeast for the assembly of the fragment to be used as a donor for insertion of the GFP into <i>galK</i> gene the <i>E. coli</i> genome. Resistance to nourseothricin (yeast).	This study
pMIC_GFP_lacZ	Plasmid cloned in yeast for the assembly of the fragment to be used as a donor for insertion of the GFP in <i>lacZ</i> region in the the <i>E. coli</i> genome. Resistance to nourseothricin (yeast).	This study
pMIC_cscBKA_ΔgalK	Plasmid cloned in yeast for the assembly of the fragment to be used as a donor for insertion of the second copy of cscBKA operon into <i>galK</i> of the <i>E. coli</i> genome. Resistance to nourseothricin (yeast).	This study
pMIC_sucP_lacZ	Plasmid cloned in yeast for the assembly of the fragment to be used as a donor for insertion of the <i>sucP</i> in <i>lacZ</i> region gene into the <i>E. coli</i> genome. Resistance to nourseothricin (yeast).	This study
pMIC_phaCAB_lacZ	Plasmid cloned in yeast for the assembly of the fragment to be used as a donor for insertion of the <i>phaCAB</i> operon into the <i>E. coli</i> genome. Resistance to nourseothricin (yeast).	This study
pMIC_ΔfliA	Plasmid cloned in yeast for the assembly of the fragment to be used as a donor for deletion of <i>fliA</i> gene into the <i>E. coli</i> . Resistance to nourseothricin (yeast).	This study

APÊNCIDE C – Figuras detalhadas de plasmídeos

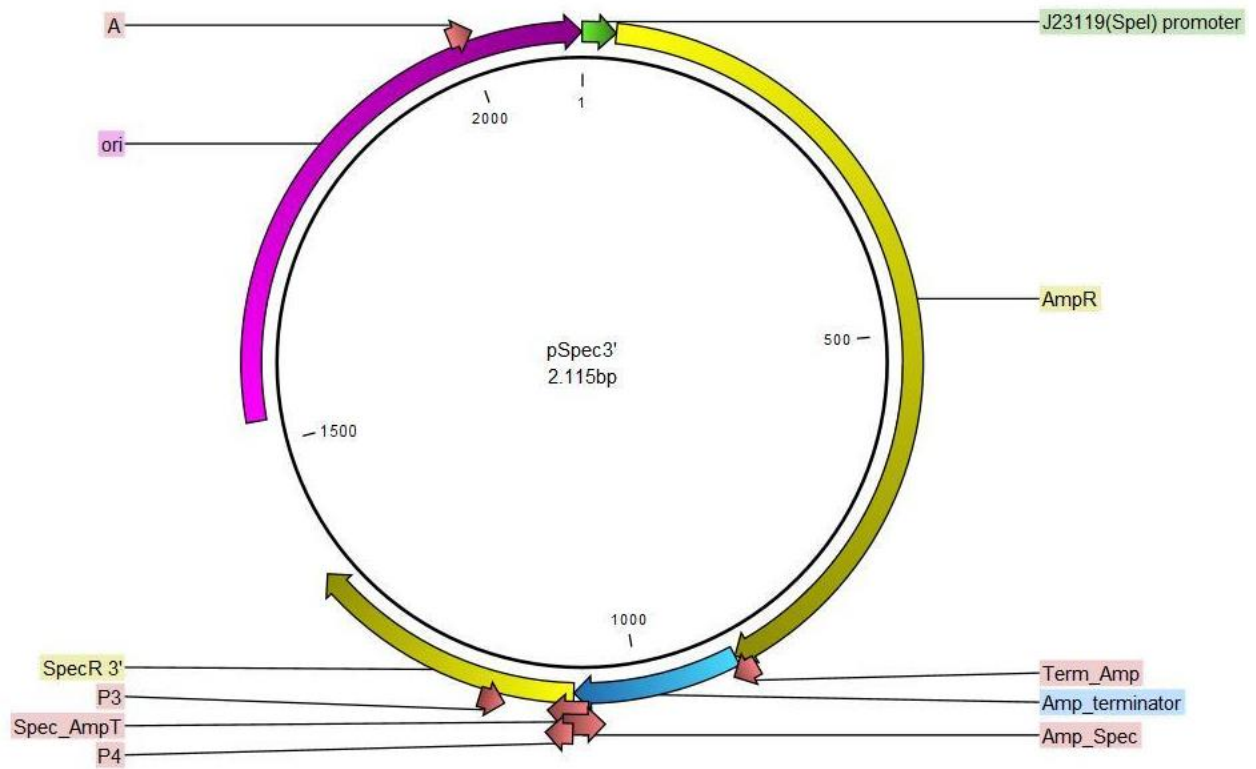
pTargetAmp



pTargetAmp- plasmídeo para a estratégia inicial de clonagem de gRNA. O promotor prJ23119 é mostrado em verde, e o gRNA *scaffold* em azul escuro. Para clonagem de plasmídeo de gRNA, ori, SpecR, andaime de gRNA e prJ23119 são amplificados por PCR usando primers homólogos de extremidade 5'.

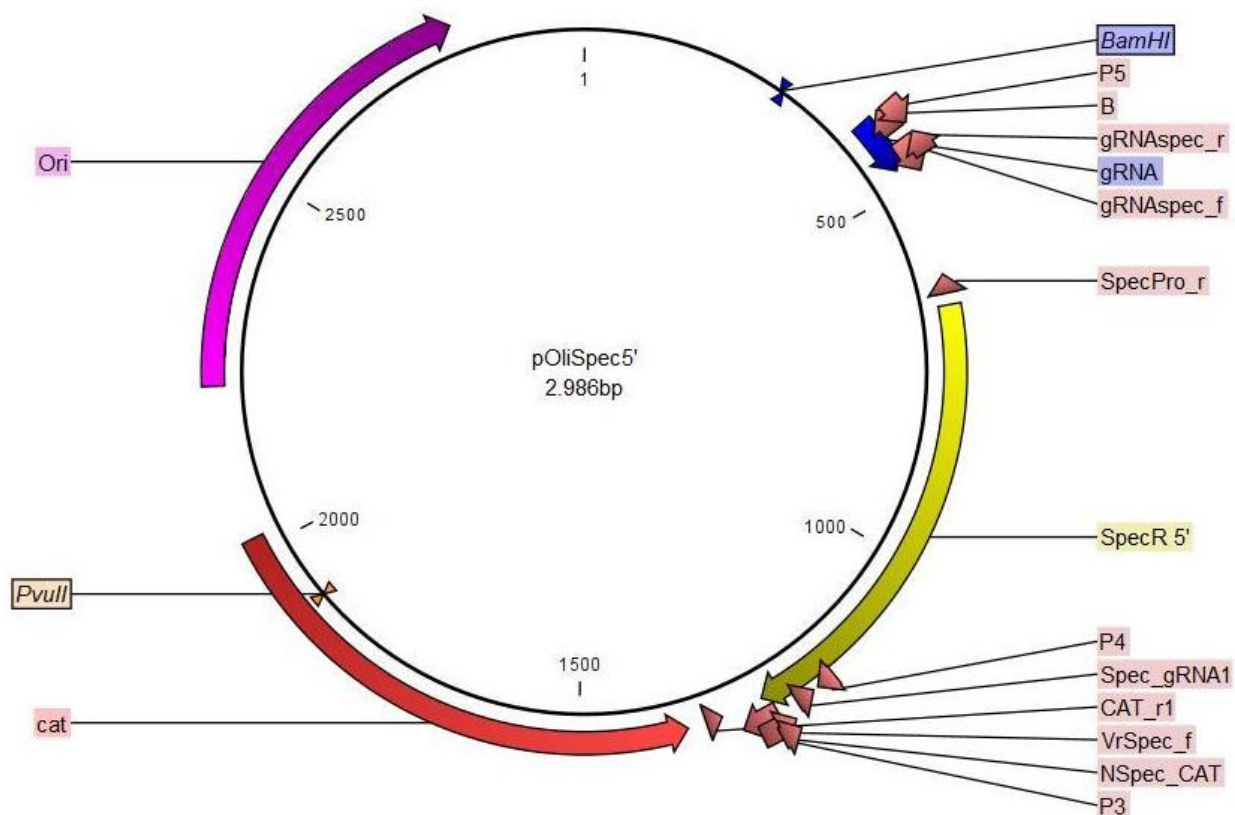
pSpec5'

pSpec5'. Primeiro plasmídeo para a estratégia de complementação bimolecular na clonagem de gRNA. Para clonagem de plasmídeo de gRNA, a extremidade 5' do SpecR e o gRNA scaffold são amplificados por PCR com primers 5' gRNA e P3.

pSpec3'

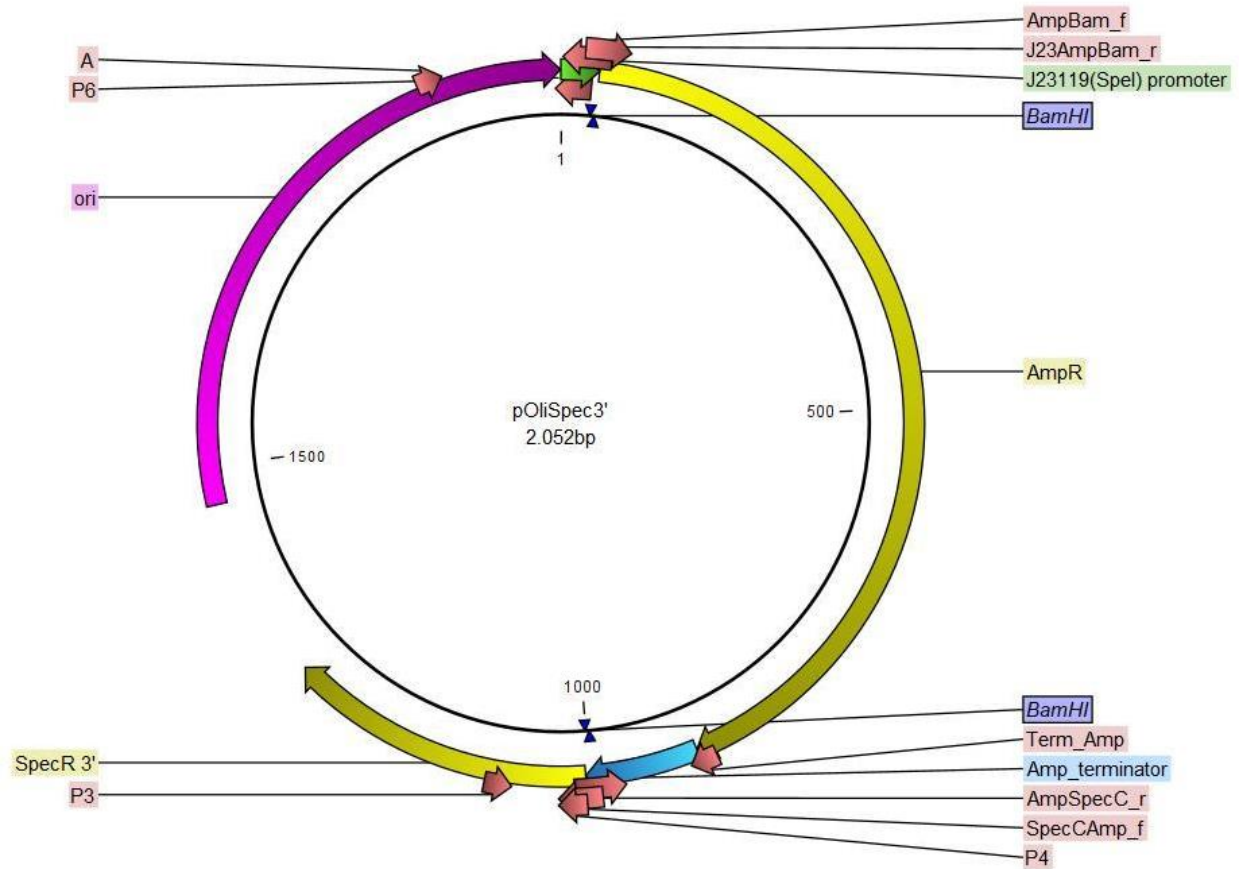
pSpec3'. Segundo plasmídeo para a estratégia de bimolecular complementation na clonagem de gRNA. Para clonagem de plasmídeo de gRNA, a extremidade 3' do SpecR, ori e prJ23119 são amplificadas por PCR com primers P4 e 5'end gRNA.

pOliSpec5'



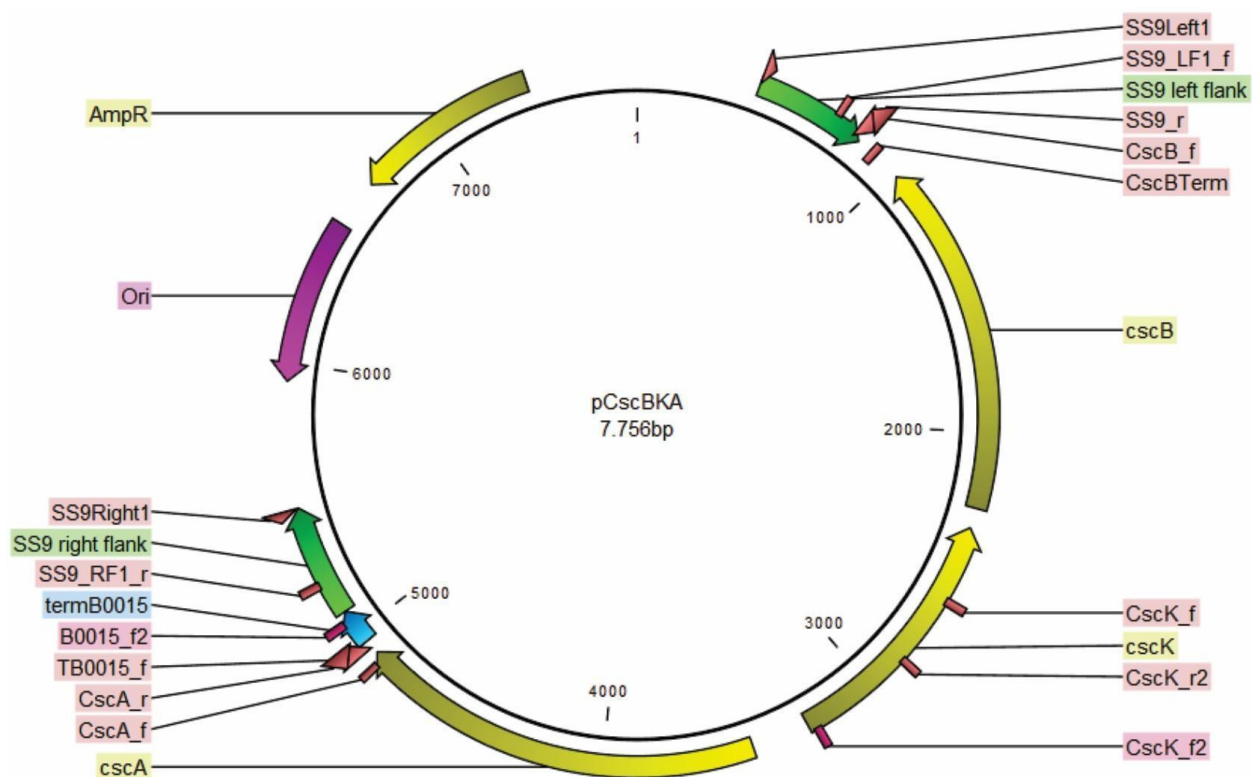
pOliSpec5'. Primeiro plasmídeo para clonagem de oligo gRNA. Para clonagem de plasmídeo gRNA, a extremidade 5' do SpecR e o *scaffold* do gRNA são amplificados por PCR usando primers P5 e P3, após restrição com enzimas *Bam*HI e *Pvu*II.

pOliSpec3'



pOliSpec3'. Segundo plasmídeo para clonagem de oligo gRNA. Para clonagem de plasmídeo gRNA, a extremidade 3' do SpecR, ori e prJ23119 são amplificadas por PCR usando primers P4 e P6, após restrição com enzima *Bam*HI.

pCscBKA



pCscBKA. Plasmídeo para expressão do operon cscBKA. As regiões de flaqueamento SS9 são mostradas em verde. Para uso como doador, as regiões de flaqueamento e o operon cscBKA são amplificados com primers SS9Left1/SS9Right1.

APÊNDICE D – Mutações encontradas no experimento de Evolução

Tabela Suplementar 1: Mutações nas populações evoluídas

DH5α-Kan							
Overlapping annotations	Reference Position	Reference	Allele	Frequency	Coding region change	Amino acid change	Annotation
NEB5A_00100, NEB5A_00105	19796..20563		Amplification				IS1 family transposase
NEB5A_00770	168800..170008		Amplification				IS4-like element ISVsa5 family transposase
NEB5A_01240	259482..260690		Amplification				IS4-like element ISVsa5 family transposase
NEB5A_01255	262664..262776		Amplification				IS4 family transposase
NEB5A_01425, NEB5A_01430	287258..288589		Amplification				IS2 repressor TnpA, transposase
CDS: <i>purB</i> , NEB5A_05385, Gene: NEB5A_05385	1088135	T	C	100	AOO69410.1:c.343A>G	AOO69410.1:p.Lys115Glu	adenylosuccinate lyase
NEB5A_05995	1196613..1197380		Amplification				IS1 family transposase
NEB5A_17590	3523998..3524765		Amplification				IS1 transposase; pfam03400
CDS: <i>nlpA</i> , Gene: NEB5A_18775	3780165		Insertion	30.8	AOO71868.1:c.399Ins	disruption	lipoprotein NlpA; cytoplasmic membrane lipoprotein-28. Insertion of a IS4-like element ISVsa5 family transposase, probably encoded by NEB5A_01240.
CDS: <i>rpoC</i> , Gene: NEB5A_20470	4129541	G	T	72.9	AOO72163.1:c.3622G>T	AOO72163.1:p.Asp1208T	DNA-directed RNA polymerase subunit beta'
NEB5A_21675	4375064..4375839		Amplification				IS1 family transposase
NEB5A_22220	4483285..4484052		Amplification				IS1 family transposase
<i>pCscBKA</i> plasmid			<i>pCscBKA</i> plasmid chromosomal integration				genome integration of the cscBKA-encoding plasmid mediated by IS1 family transposons; possible IS1 insertion site at plasmid position 549..557.

DH5α-Cm

Overlapping annotations	Reference Position	Reference	Allele	Frequency	Coding region change	Amino acid change	Annotation
NEB5A_00100, NEB5A_00105	19796..20563		Amplification				IS1 family transposase
102 nts upstream NEB5A_04295 (phosphoprotein <i>phoE</i>)	882195		Insertion	74.4			IS1 family transposase - phosphoprotein PhoE
NEB5A_00770	168779..170034		Amplification				IS4-like element ISVsa5 family transposase
NEB5A_00770	169800..170034		Amplification				IS4-like element ISVsa5 family transposase
Gene: NEB5A_01240	259469..260716		Amplification				IS4-like element ISVsa5 family transposase
Gene: NEB5A_01240	260481..260716		Amplification				IS4-like element ISVsa5 family transposase
phosphoprotein <i>phoE</i> , Gene: NEB5A_01245	260801..261651		deletion NEB5A_01245 and NEB5A_01250				incomplete; partial on complete genome; missing start
sigma factor-binding protein Crl ; Gene: NEB5A_01250							involved in the expression of csgBA which is involved in curli formation; interacts with sigmaS; incomplete; partial on complete genome; missing start
NEB5A_01255	262664..262898		Amplification				IS4-like element ISVsa5 family transposase
CDS: <i>nudJ</i> , Gene: NEB5A_05400	1090328	C	A	100	AOO69412.1:c.452G>T	AOO69412.1:p.Gly151Val	NUDIX hydrolase; phosphatase NudJ
NEB5A_05995	1196613..1197380		Amplification				IS1 family transposase
NEB5A_17590	3523998..3524765		Amplification				IS1 transposase; pfam03400
CDS: <i>rpoB</i> , Gene: NEB5A_20465	4121843	G	A	82	AOO72162.1:c.29G>A	AOO72162.1:p.Arg10His	DNA-directed RNA polymerase subunit beta
NEB5A_21675	4375064..4375839		Amplification				IS1 family transposase
NEB5A_22220	4483285..4484052		Amplification				IS1 family transposase

pCscBKA plasmid			pCscBKA plasmid chromosomal integration				genome integration of the cscBKA-encoding plasmid mediated by IS4 family transposons; possible IS4 integration sites at plasmid positions 5748 and 5942.
-----------------	--	--	---	--	--	--	--

DH5α-Spec							
Overlapping annotations	Reference Position	Reference	Allele	Frequency	Coding region change	Amino acid change	Annotation
NEB5A_00100, NEB5A_00105	19796..20563		Amplification				IS1 family transposase
CDS: <i>pcnB</i> , Gene: NEB5A_00725	158456..158536		81 bp deletion	100	558..638 deletion	27 amino acids deletion	polynucleotide adenylyltransferase; poly(A) polymerase
122 predicted genes	986288..1094632		108,345 bp duplication				flanked by transposon NEB5A_04815 and NEB5A_05420 (e14 prophage; predicted inner membrane protein). Consistent with a tandem duplication.
NEB5A_06770	1340821..1341801		Amplification				IS5 family transposase
NEB5A_06790	1344371..1345351		Amplification				IS5 family transposase
NEB5A_09385	1855905..1856885		Amplification				IS5 family transposase
50 predicted genes	1941726..1982144		40,419 bp deletion	100			50 predicted genes, including flagella assembly cluster
NEB5A_10410	2048226..2049206		Amplification				IS5 family transposase
NEB5A_10700	2116862..2117728		Amplification				IS3-like element IS3 family transposase
NEB5A_10720	2120759..2121739		Amplification				IS5 family transposase
CDS: <i>rpoZ</i> , Gene: NEB5A_18710	3762833	G	A	100	AOO71856.1:c.157G>A	AOO71856.1:p.Glu53Lys	DNA-directed RNA polymerase subunit omega
CDS: <i>dnaA</i> , Gene: NEB5A_19010	3823754	G	A	94.92	AOO71913.1:c.547C>T	AOO71913.1:p.His183Tyr	chromosomal replication initiation protein DnaA
NEB5A_21675	4375111..4375614		Amplification				IS1 family transposase
NEB5A_22220	4483534..4484037		Amplification				IS1 family transposase
			pCscBKA plasmid lower copy number				lower copy number of the plasmid cscBKA

E2348/69-
progenitor

Overlapping annotations	Reference Position	Reference	Allele	Frequency	Coding region change	Amino acid change	Annotation
182 predicted genes	1270565..1429913		159,349 bp duplication				region in between prophages E2348_PP5 and E2348_PP6
			Absence				

E2348/69-Cm

Overlapping annotations	Reference Position	Reference	Allele	Frequency	Coding region change	Amino acid change	Annotation
CDS: E2348C_1098 , Gene: E2348C_1098	1161079		18 bp deletion	100	22..39 deletion	6 amino acids deletion	CDS inside the integrative element E2348_IE2
CDS: yeaY , Gene: E2348C_1929	1989108		insertion prophage E2348_PP9, Mu-like	100	disruption	disruption	outer membrane lipoprotein, Slp family; TIGR00752
CDS: zwf , Gene: E2348C_1977	2034067	A	C	99.57	CAS09525.1:c.1215T>G	CAS09525.1:p.Asp405Glu	glucose-6-phosphate dehydrogenase
15 CDS: E2348C_2179; E2348C_2180; wbiN; galE; galF; gne; wcaM; wcaL; wcaK; wzxC; wcaJ; cpsG; cpsB; wcaI; gmm	2227996..2246018		18,023 bp deletion	100	Deletion of 15 genes	deletion	15 genes involved in colanic acid biosynthesis
prophage E2348_PP9, Mu-like	2715938..2751675		Amplification				amplification mediated by the prophage E2348_PP9, Mu-like
63 CDS	3086086..3155640		69,555 bp amplification				
CDS: hflK , Gene: E2348C_4497	4750118	C	-	94.25	CAS12045.1:c.700delC	CAS12045.1:p.Leu234fs	modulator for HflB protease specific for phage lambda cII repressor

E2348/69-Kan

Overlapping annotations	Reference Position	Reference	Allele	Frequency	Coding region change	Amino acid change	Annotation
CDS: <i>dnaK</i> , Gene: E2348C_0014	14067	G	T	90.1	CAS07562.1:c.456G>T	CAS07562.1:p.Gln152His	chaperone Hsp70, co-chaperone with DnaJ
96 bp upstream CDS: <i>cysB</i> , Gene: E2348C_1465	1500338	T	C	77.6			DNA-binding transcriptional dual regulator, O-acetyl-L-serine-binding
15 CDS: E2348C_2179; E2348C_2180; <i>wbiN</i> ; <i>galE</i> ; <i>galF</i> ; <i>gne</i> ; <i>wcaM</i> ; <i>wcaL</i> ; <i>wcaK</i> ; <i>wzxC</i> ; <i>wcaJ</i> ; <i>cpsG</i> ; <i>cpsB</i> ; <i>wcaI</i> ; <i>gmm</i>	2227996..2246018		18,023 bp deletion	100	Deletion of 15 genes	deletion	15 genes involved in colanic acid biosynthesis
CDS: <i>fre</i> , Gene: E2348C_4156	4343318	G	A	48.2	CAS11704.1:c.416G>A	CAS11704.1:p.Gly139Asp	flavin reductase
CDS: <i>hfIK</i> , Gene: E2348C_4497	4749901	-	GAAC	90.9	CAS12045.1:c.482_483insGAAC	CAS12045.1:p.Val163fs	modulator for HflB protease specific for phage lambda cII repressor

E2348/69-Amp							
Overlapping annotations	Reference Position	Reference	Allele	Frequency	Coding region change	Amino acid change	Annotation
CDS: <i>dnaK</i> , Gene: E2348C_0014	14861	C	A	91.48	CAS07562.1:c.1250C>A	CAS07562.1:p.Thr417Asn	chaperone Hsp70, co-chaperone with DnaJ
275 predicted genes	1292769..1547205		254,437 bp amplification				the 254437 bp amplification Includes the same region amplified in the ancestor chromosome (1270565..1429913), however it extends beyond and has different boundaries. Flanked by insertion sequences E2348C_IS_31 and E2348C_IS_42.
CDS: <i>ydiV</i> , Gene: E2348C_1792	1863125	T	-	90	CAS09340.1:c.318delA	CAS09340.1:p.Lys106fs	predicted protein
CDS: <i>hflB</i> , Gene: E2348C_3457	3605098	-	TTG	95.4	CAS11005.1:c.117_118insCAA	CAS11005.1:p.Gln39_Glu40insGln	protease, ATP-dependent zinc-metallo

CDS: <i>arcB</i> , Gene: E2348C_3489	3630417	A	T	46.04	CAS11037.1:c.748T>A	CAS11037.1:p.Tyr250Asn	hybrid sensory histidine kinase in two-component regulatory system with ArcA
CDS: <i>perA</i> , Gene: E2348_P1_020	17952	-	A	100	CAS07456.1:c.776_777insA	CAS07456.1:p.Tyr262fs	transcriptional activator of the bfp operon

E2348/69-Spec

Overlapping annotations	Reference Position	Reference	Allele	Frequency	Coding region change	Amino acid change	Annotation
79 nt upstream CDS: <i>cysB</i> , Gene: E2348C_1465	1500355	A	C	99.58			DNA-binding transcriptional dual regulator, O-acetyl-L-serine-binding
CDS: <i>mreB</i> , Gene: E2348C_3522	3670290	C	G	28.83	CAS11070.1:c.202G>C	CAS11070.1:p.Gly68Arg	cell wall structural complex MreBCD, actin-like component MreB
CDS: <i>spoT</i> , Gene: E2348C_3914	4082579	C	T	100	CAS11462.1:c.1217C>T	CAS11462.1:p.Pro406Leu	bifunctional (p)ppGpp synthetase II/guanosine-3',5'-bis pyrophosphate 3'-pyrophosphohydrolase
CDS: <i>hfik</i> , Gene: E2348C_4497	4749420..4749480		61 bp deletion	100	CAS12045.1:c.2TDel 61 bp	CAS12045.1:p.Trp3fs	modulator for HflB protease specific for phage lambda cII repressor
CDS: <i>nadR</i> , Gene: E2348C_4688	4951063	C	-	94.72	CAS12236.1:c.373delC	CAS12236.1:p.Arg125fs	bifunctional DNA-binding transcriptional repressor/NMN adenyllyltransferase

APÊNDICE E – Procedimentos de genética molecular

gRNA Plasmids cloning

Cloning pTargetAmp

The ampicillin resistance marker and terminator were amplified via PCR using primers Amp_J23119/Amp_JacA from plasmid pUC57_sucP, generating a 1,082 bp fragment. The J23119 promoter, origin of replication, spectinomycin resistance marker, and gRNA scaffold were amplified by PCR from plasmid pTargetF using primers JacA_Amp/Vf_J, producing a 1,844 bp fragment. After DpnI treatment, 200 ng of each fragment were mixed and used to transform *E. coli* DH5 α cells, using in vivo cloning approach. Colonies were screened with colony PCR using primers A/AmpMet (200 bp).

For use in gRNA plasmid cloning, amplification is performed with primers P2_XXXX/P1_XXXX (1870 bp). where the gRNA sequence is specified by P2_XXXX, and the complementary sequence by P1_XXXX. After treatment with DpnI, the fragment produced was transformed into competent *E. coli* DH5 α cells. Colony selection on spectinomycin was followed by confirmation via colony PCR with primers A/B (221 bp).

pSpec5' and pSpec3' cloning and use for plasmid gRNA

For pSpec5' cloning, the ampicillin resistance marker and the origin of replication were amplified via PCR with primers SpecN_f/B (1750 bp) from pTargetAmp. The initial part of the spectinomycin resistance marker and gRNA scaffold were amplified with primers SpecN_R/JacA_Amp (884 bp) from pTargetAmp. After DpnI treatment, the PCR products were mixed and transformed into competent *E. coli* DH5 α cells. Positive clones, selected on ampicillin plates, were confirmed by colony PCR with Ori5/P4 primers (182 bp).

For pSpec3' cloning, the ampicillin resistance mark was amplified via PCR with primers Amp_J23119/Amp_Spec (1076 bp) from pTargetAmp, while the terminal part of the spectinomycin resistance marker, replication origin, and J23119 promoter were amplified using Spec_AmpT/AmpMet primers (1,132 bp) from pTargetAmp. After DpnI treatment, with DpnI, the PCR products were mixed and transformed into competent *E.*

coli DH5 α cells. Positive clones, selected on ampicillin plates, were confirmed via colony PCR with Term_Amp/Spec_gRNA1 primers (247 bp).

For gRNA plasmid cloning, PCR amplification is performed from pSpec5' with primers P3/P2_XXXX (888 bp); PCR amplification is performed from plasmid pSpec3' with primers P1_XXXX/P4 (1114 bp). The gRNA is specified in primer P2_XXXX and the complementary sequence in primer P1_XXXX. These fragments are mixed and transformed into competent *E. coli* DH5 α cells. After colony selection in spectinomycin, cloning confirmation is through colony PCR with the primers A/B (221 bp).

pOliSpec5' and pOliSpec3' cloning and use for plasmid gRNA

For pOliSpec5' cloning, the gRNA scaffold and the initial part of the spectinomycin resistance marker were amplified via PCR using the gRNAspec_f/NSpec_CAT_r primers r(832 bp) from pSpec5'. The chloramphenicol resistance marker and origin of replication were amplified from pEasyG3_mic by PCR, followed by DpnI treatment. These PCR products were mixed and transformed into competent *E. coli* DH5 α cells. After colony selection on chloramphenicol plates, positive clones were selected by colony PCR with P4/CAT_r1 (169 bp) P5/SpecPro_r (246 bp) primers.

For pOliSpec3' cloning, the terminal part of spectinomycin, replication origin, and J23119 promoter were amplified by PCR using SpecCAmp_f/J23AmpBam_r (1109 bp) primers from pTargetAmp. The ampicillin resistance and gRNA scaffold were amplified by PCR from pTargetAmp with AmpBam_f/AmpSpecC_r (922 bp) primers. The PCR products were DpnI-treated, mixed, and transformed into competent *E. coli* DH5 α cells. After colony selection on ampicillin plates, positive clones were selected by colony PCR with P3/Term_Amp (213 bp) primers.

Prior to amplification for use in gRNA cloning, plasmid pOliSpec5' is digested with *Pst*I and *Bam*HI and pOliSpec3' is digested with *Bam*HI. PCR amplification from digested pOliSpec5' used P5/P3 primers (858 bp) and from pOliSpec3', P4/P6 primers (1,084 bp). These fragments are purified and stored. The gRNA-oligo is amplified by PCR from the central oligo with the primers pE/gE. These three fragments are mixed and transformed into competent *E. coli* DH5 α cells. After colony selection in spectinomycin, cloning confirmation is through colony PCR with the A/B (221 bp) primers.

Cloning plasmids for tagging ALE populations.

For pKan, two fragments were amplified by PCR from pUC57_phaB using the primers combinations Vf_kan/AmpT_f (641 bp) and pGEX_kan/Or2_r (1,280 bp).

For pSpec, two fragments were amplified by PCR from pTargetAmp, with the primer combinations Vf_spec/AmpT_f (641 bp) and spec_pro/Or2_r (1,253 bp).

For pAmp, a single fragment was amplified from pUC19 using Vf_min/Vr_min primers (1,751 bp).

The PCR products were DpnI-treated and purified. For pKan and pSpec, 200 ng of each fragment were mixed and this mixture was used to transform competent *E. coli* DH5 α cells, which were plated on LB plates with the respective antibiotic. For pAmp, 200 ng of the single fragment were used to transform *E. coli* DH5 α cells, plated on LB with ampicillin. Colonies were verified by PCR: pKan: prokan_f/Ori_Term_PhaB_R, 164 bp; pSpec: Ori_Term_PhaB_R/prospec_R, 210 bp; pAmp: AmpMet/Ori_Term_PhaB_R, 205 bp.

Cloning pUC19_B0015 and pGFP

The Gibson (NEB) method was used to clone pUC19_B0015. The B0015 terminator was amplified by PCR from plasmid pUC57_sucP using B0015for/B0015rev (173 bp) primers, while the vector fragment was amplified from pUC19 using Vf_B0015/pUC19_B0015 primers (2,649 bp). Both products were treated with DpnI. The Gibson followed the manufacturer's instructions, and the product was used to transform competent *E. coli* DH5 α cells by heat shock, followed by plating on LB with ampicillin. Colonies were screened by PCR with B0015for/AmpPro (221 bp) primers.

For GFP expression plasmid cloning, the mEGFP sequence was amplified from mEGFP_pBAD plasmid using GFP19for/GFP_B15rev (752 bp) primers. The backbone, containing the lac promoter and B0015 terminator, was amplified from the pUC19_B0015 B15-GFPfor/19GFPprev primers. These fragments were transformed into DH5 α cells for cloning and selected for ampicillin. Positive cloning was evaluated using colony PCR using GFP_f/B0015rev (271 bp) and pUC19-2for/GFP_r2 (354 bp) primers. The colonies were cultured, and GFP expression was analyzed by flow cytometry.

Donor assembly and CRISPR/Cas9-mediated genomic modifications

pCscBKA cloning and SS9::cscBKA insertion

The *cscBKA* operon was amplified by PCR from *E. coli* E2348/69 genomic DNA of using the CscB_f/CscA_r (4,178 bp) primers. The cloning vector, pUC57_SucP, was amplified with SS9_r/TB0015_f (3,695 bp) primers, followed by DpnI treatment. The PCR products were purified, and 200 ng of each were mixed to transform competent *E. coli* DH5 α cells, which were plated on LB with ampicillin. Colonies were screened by PCR using SS9_LF1_f/cscB_Term (202 bp) and CscA_f/SS9_RF1_r (362 bp) primers.

From this cloned plasmid, the donor was amplified by PCR for genome insertion of the *cscBKA* operon using the primers SS9Left1/SS9Right1 (5046 bp) purified and 400 ng was used for the CRISPR/Cas9 technique. The insertion of the operon was confirmed by colony PCR with FE1/CscBTerm (582 bp) and B0015_f2/FD1 (524 bp) primers were used. The gel for confirmation of this insertion is represented in Fig. S1.

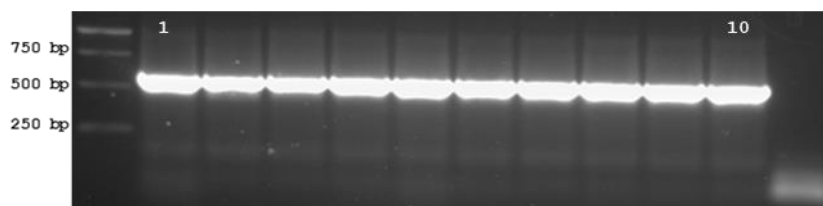


Fig. S1. Colony PCR for confirmation of insertion of *cscBKA* operon in the SS9 site. Primers: FE1/CscBTerm (582 bp). The numbers (1-10) indicate the colonies tested.

Δ pcnB (81 bp)

The donor was amplified by PCR with primers donor_pcnB_F/donor_pcnB_R (739 bp) from genomic DNA of D-Spec population. and used in CRISPR/Cas9 with the gRNA plasmid in *E. coli* DH5 α cells containing the *cscBKA* operon. Deletion was confirmed by colony PCR with pcnB_F/pcnB_R (228 bp for deletion; wild type: 309 bp) primers. Confirmation of this deletion is shown in Fig. S12.

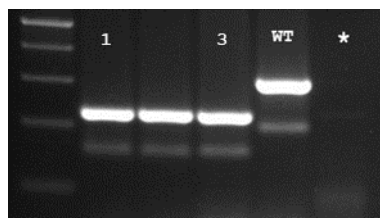


Fig. S2. PCR from genomic DNA for confirmation of deletion of 81 bp of *pcnB* gene. Primers: *pcnB_F/pcnB_R* (228 bp for deletion; wild type: 309 bp). The numbers (1-3) indicate the colonies tested. (*) represents blank; WT: wild type. 100 bp ladder.

Δ40.4 kb

Primers for PCR amplification of the donor for a 40.4 kb deletion (*Flagel_left_F/Flagel_right_R*) were designed 1,000 bp upstream and downstream of the target region in the D-Spec genome. The amplified fragment was used for the CRISPR/Cas9 with gRNA plasmid in *E. coli* DH5α cells with *cscBKA* operon inserted. The deletion was confirmed by colony PCR with primers *del_flagel_F/Del_flagel_R*. The gel for confirmation of this deletion is represented in Fig. S3.

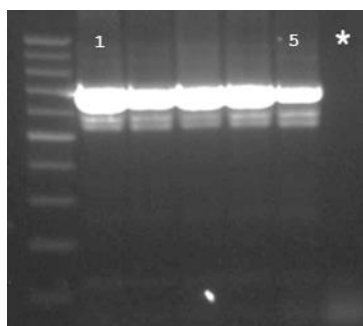


Fig. S3. Colony PCR for confirmation of deletion of 40.4 kb. Primers: *flagel_seq_F/flagel_seq_R*. (*) represents blank. The numbers (1-5) indicate the colonies tested. 100 bp ladder.

Cloning in *Saccharomyces cerevisiae*

For yeast plasmid cloning for donor amplification, pEasyG3-mic and pEasyG3-nat or pEasyG4-hph were used as backbone for cloning. The 2-micron origin of replication was amplified by PCR using 2Mic_f/2micMX primers from pEasyG3-mic (1,452 bp), while the resistance marker was amplified with tMX_2m/MXProUni primers from pEasyG3-nat or hph (1,449 bp or 1,645 bp).

galK::cscBKA

The left and right of *galK* flanking regions were amplified by PCR with mx_galK/AmpTerm_R (594 bp) and cscB_galK_F/galK_2mic (506 bp) primers from the genomic DNA of cells containing the two GFP copies of integrated. The *cscBKA* operon was amplified by PCR from the lineage with this operon integrated with primers term_cscA_F/galK_cscB_R (3,880 bp). These reactions were transformed together with the origin of replication and the resistance marker into *S. cerevisiae* cells. Colony PCR confirmed plasmid assembly using cscBterm/right_R (318 bp) and galK_amp/cscA_f (285 bp) primers.

The donor was amplified by PCR with galK_left_f/Right_rev (5,087 bp) primers and used in the CRISPR/Cas9 technique. Colony PCR confirmed insertion using cscBterm/galT_rev (540 bp) and galM_f/cscA_f (762 bp) primers. Gel confirmation is shown in Fig. S4.

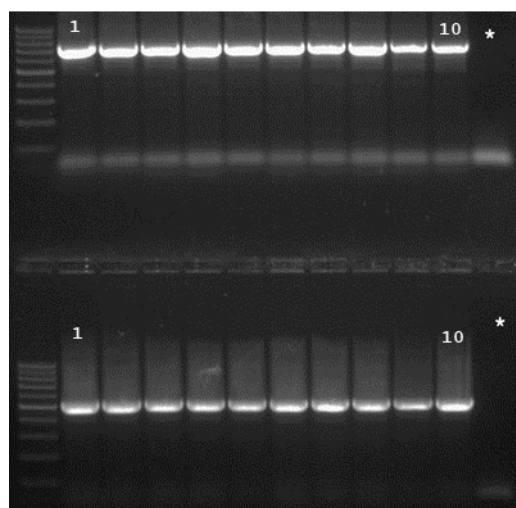


Fig. S4. Colony PCR for confirmation of insertion of *cscBKA* operon in the *galK* gene. Primers: galM_f/cscA_f (762 bp, up) and cscBterm/galT_rev (540 bp, down). The numbers (1-10) indicate the colonies tested. (*) represents blank. 100 bp ladder.

cscA::sucP* and *cscA::pJ23119-sucP

For cloning of pMIC_cscBKsucP, the *sucP* gene with the flanking SS9 region was amplified by PCR from plasmid pUC57_sucP with SS9R_2M/SucP_fw (2,056 bp) primers.

The fragment containing *cscBK* genes and the other flanking region was amplified by PCR from pCscBKA, using SS9L_MX/sucP_rev (2,999 bp) primers. For pMIC_cscBKsucP(J23119), primers J23_f/SS9R_2M (2,116 bp) from pUC57_sucP and primers SS9L_MX/CscK_r (3,048 bp) were used. These fragments were transformed into yeast. Colony PCR using CscA_f2/SucP_r (312 bp for *cscA::sucP* and 347 bp for *cscA::pJ23119-sucP*) primers, confirmed plasmid assembly. The donor was amplified by PCR with primers csc_F6/SS9Right1 (2,413 bp for *cscA::sucP*; 2,438 bp for *cscA::pJ23119-sucP*) and used for CRISPR/Cas9 technique in *E. coli* DH5 α cells with *cscBKA* operon inserted. Insertion was confirmed by colony PCR with primers csc_F6/SucP_r (423 bp for pMIC_cscBKsucP(J23119) and 388 bp for pMIC_cscBKsucP) and B0015_f2/FD1 (524 bp). The gel for confirmation of these insertions shown in Fig. S5.

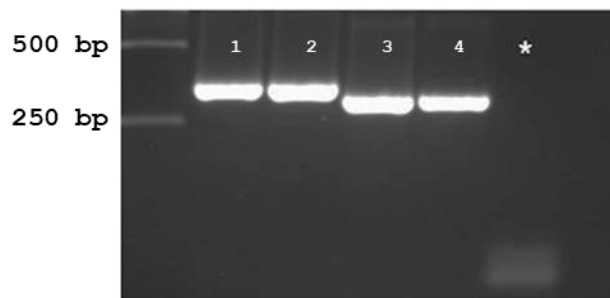


Fig. S5. PCR from genomic DNA for confirmation of insertion of *sucP* operon in the *cscA* gene, with prJ23119 (1 and 2) and only with natural promoter (3-4). Primers: CscA_f2/SucP_r (312 bp for *cscA::sucP* and 347 bp for *cscA::pJ23119-sucP*) . (*) represents blank.

lacZ::sucP

For this cloning, the left and right flanking regions were amplified from genomic DNA of cells with two integrated GFP copies using mx_lacZ/Ampterm_R (594 bp) and J23_lacZ_F/lacZ_2mic (442 bp) primers. The *sucP* gene was amplified from pUC57_sucP with galK_sucP_F/galK_J23_R (1,602 bp) primers. These reactions were transformed together with the origin of replication and the resistance marker into *S. cerevisiae* cells. Plasmid assembly was confirmed by PCR with lacZ_amp_f/SucP_R2 (259 bp) and SucP_r/donor_lacZ_r (501 bp) primers. The donor was amplified by PCR with lacZ_left_f/donor_lacZ_r (2,530 bp) primers and used in the CRISPR/Cas9 technique.

Insertion was confirmed by PCR with SucP_r/LacZ_R (553 bp) and SucP_R2/LacZ_F (698 bp) primers. Confirmation of the insertion is shown in Fig. S6.

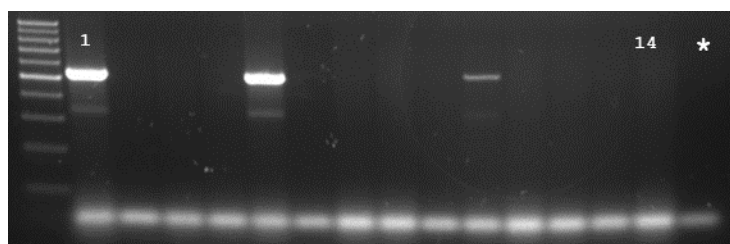


Fig. S6. Colony PCR for confirmation of insertion of *sucP* operon in the *lacZ* locus. Primers: SucP_r/LacZ_R (553 bp). The numbers (1-14) indicate the colonies tested. (*) represents blank. 100 bp ladder.

galK::gfp

The mEGFP coding sequence was amplified by PCR from the mEGFP_pBAD plasmid with the primers termAmp_GFP_f/RBS_GFP_F (772 bp). The right and left flanking regions, with the J23119 promoter and AmpR terminator, respectively, were amplified by pMIC_AmpR_ΔgalK, with mx_galK/Ampterm_R (594 bp) and RBS_J23_R/galK_2mic (486 bp) primers. These fragments were transformed into *S. cerevisiae*, and positive cloning were selected by colony PCR with galK_amp/GFP_f (307 bp) and GFP_r/donor_lacZ_r (548 bp) primers. After growth and purification, the donor was amplified by PCR using lacZ_left_f/Right_rev (1,757 bp) primers, which was co-transformed with gRNA plasmid into *E. coli* DH5α with the *cscBKA* operon integrated into the genome and Cas9 plasmid. Insertions confirmation was performed by colony PCR using primers galM_f/GFP_f (783 bp) and GFP_r/galT_rev (608 bp) primers. Confirmation of the insertion is shown in Fig. S7.

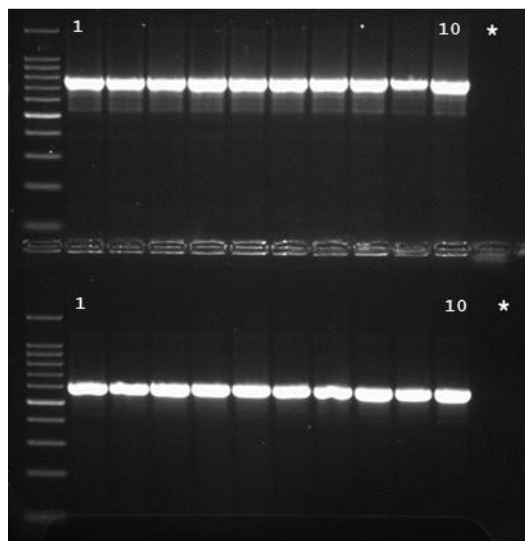


Fig. S7. Colony PCR for confirmation of insertion of *gfp* in the *galK* locus. Primers: *galM_f/GFP_f* (783 bp, up) and *GFP_r/galT_rev* (608 bp, down). The numbers (1-10) indicate the colonies tested. (*) represents blank. 100 bp ladder.

lacZ::gfp

The mEGFP coding sequence was amplified by PCR from the mEGFP_pBAD plasmid using *termAmp_GFP_f/RBS_GFP_F* (772 bp) primers. The right and left flanking regions, containing the J23119 promoter and AmpR terminator, respectively, were amplified by from pMIC_AmpR_lacZ, with *mx_lacZ/Amp_term_R* (594 bp) and *RBS_J23_R/lacZ_2mic* (486 bp) primers. These fragments were transformed into *S. cerevisiae*, and positive cloning was selected by colony PCR with *lacZ_amp_f/GFP_f* (307 bp) and *GFP_r/donor_lacZ_r* (548 bp) primers. After growth and purification, the donor was amplified by PCR using *lacZ_left_f/donor_lacZ_r* (1,757 bp) primers and co-transformed with gRNA plasmid into *E. coli* DH5 α with integrated *cscBKA* operon and Cas9 plasmid. Insertions confirmation was performed made by colony PCR using *LacZ_F/GFP_f* (746 bp) and *GFP_r/galT_rev* (600 bp) primers. Confirmation of the insertion is shown in Fig. S8.

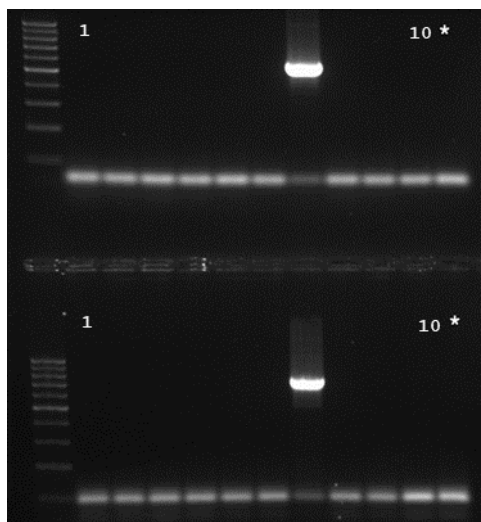


Fig. S8. Colony PCR for confirmation of insertion of *gfp* in the *lacZ* locus. Primers: GFP_r/LacZ_R (600 bp, up) LacZ_F/GFP_f (746 bp, down). The numbers (1-10) indicate the colonies tested. (*) represents blank. 100 bp ladder.

lacZ::phaCAB

For this cloning, the left and right flanking regions were amplified by PCR using mx_lacZ/Amp_term_R (594 bp) and phaC_RBS_F/lacZ_2mic (506 bp) primers from genomic DNA of cells with two copies of GFP integrated into the genome. The *phaCAB* operon was amplified from the pBBRMCS-2::*phaCAB* plasmid with term_phaB_F/RBS_phaC_R (3,880 bp) primers. These reactions were transformed together with the origin of replication and the resistance marker into *S. cerevisiae* cells. Plasmid assembly was confirmed by colony PCR using lacZ_amp_f/phaB_rev (254 bp) and phaC_fwd/donor_lacZ_r (529 bp) primers. The donor was amplified by PCR with the lacZ_left_f/donor_lacZ_r (4,888 bp) primers and used in the CRISPR/Cas9 technique. The insertion was confirmed with primer pairs phaC_fwd/LacZ_R (581 bp) and phaB_rev/LacZ_F (698 bp). Confirmation of this insertion is shown in Fig. S9.

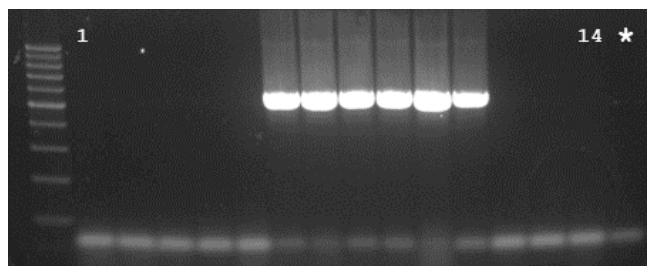


Fig. S9. Colony PCR for confirmation of insertion of *phaCAB* operon in the *lacZ* locus. Primers: phaC_fwd/LacZ_R (581 bp). The numbers (1-14) indicate the colonies tested. (*) represents blank. 100 bp ladder.

galk::ampR

The ampicillin resistance marker was amplified by PCR from the plasmid pTarget_Amp with galk_amp/galk_J23_R (1,103 bp) primers, with the promoter (J23119), the coding region and the terminator. The left and right flanking regions were amplified from the genomic DNA of *E. coli* DH5α of *galk* region, with primer pairs mx_galk/Amp_galk_R (439 bp) and J23_galk_F/galk_2mic (442 bp), respectively. The obtained fragments were mixed with the backbone fragments and used to transform competent *S. cerevisiae* cells. Colony PCR confirmed correct cloning with Left_f/AmpT_f (364 bp) and AmpMet/right_R (306 bp) primers. The donor was amplified by PCR with galk_left_f/Right_rev (1,866 bp) primers, purified and used to transform with gRNA plasmid in competent cells of *E. coli* DH5α with *cscBKA* operon, expressing Cas9. Colony PCR was used to demonstrate positive insertion with galM_f/AmpT_f (530 bp) and AmpMet/galT_rev(602 bp) primers. Confirmation of this insertion is shown in Fig. S8.

lacZ::ampR

J23119 promoter, coding region, and terminator, was amplified by PCR from the pTarget_Amp plasmid using lacZ_amp_f/lacZ_J23_R (1,103 bp) primers. The left and right flanking regions of the *lacZ* locus were amplified from *E. coli* DH5α genomic DNA using mx_lacZ/Amp_lacZ_R (439 bp) and J23_lacZ_F/lacZ_2mic (442 bp) primers. The fragments were mixed with the backbone fragments and used to transform competent *S. cerevisiae* cells. Colony PCR was used to confirm correct cloning with lacZ_left_f/AmpT_f (437 bp) and AmpMet/donor_lacZ_r (468 bp) primers. The donor was amplified by PCR

with lacZ_left_f/donor_lacZ_r (1,866 bp) primers, purified, and co-transformed with gRNA plasmid into *E. coli* DH5 α cells with *cscCBKA* integrated containing the Cas9 plasmid. Colony PCR confirmed positive insertion using LacZ_F/AmpT_f (494 bp) and AmpMet/LacZ_R (520 bp) primers. Confirmation of this insertion is shown in Fig. S10.

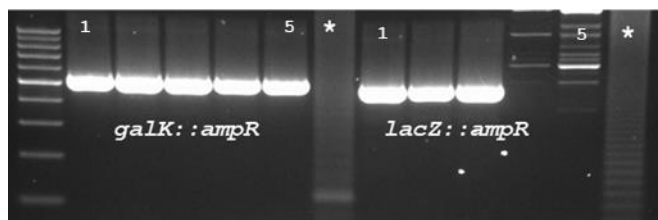


Fig. S10. Colony PCR for confirmation of insertion of *ampR* in the *lacZ* and *galk*. Primers for *galk*: galM_f/AmpT_f (530 bp); for *lacZ* LacZ_F/AmpT_f (494 bp). The numbers (1-5) indicate the colonies tested. (*) represents blank. 100 bp ladder.

Δ fliA (205 bp)

The left and right flanking regions were amplified by PCR from *E. coli* DH5 α genomic DNA using mx_fliA/FliA_R (380 bp) and FliA_F/mic_fliA (348 bp) primers, respectively. These fragments were mixed with the origin of replication and the resistance marker and transformed into *S. cerevisiae* cells. Colony PCR using Del_fliA_F/del_fliA_R (279 bp) primers was used to confirm correct cloning. The donor fragment was amplified using Donor_FliA_F/Donor_FliA_R (634 bp) primers and co-transformed with gRNA plasmid into *E. coli* DH5 α with the *cscBKA* operon integrated and Cas9 plasmid. Genomic modification was verified by PCR using primers Del_fliA_F/Check_FliA_R (474 bp for deletion, 679 bp for WT). Confirmation of this deletion is shown in Fig. S11.

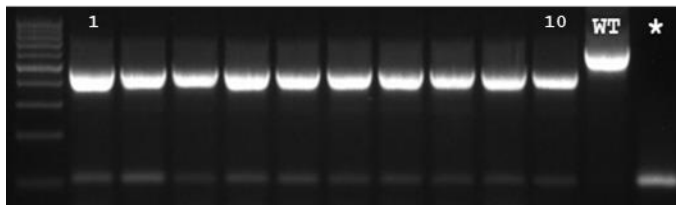


Fig. S11. Colony PCR for confirmation of deletion of 205 bp on *fliA*. Primers: Del_fliA_F/Check_FliA_R (deletion 474 bp; WT: 679 bp). The numbers (1-10) indicate the colonies tested. (*) represents blank; WT: wild type. 100 bp ladder.

***nudJ* (G115V)**

To clone the SNP region of the *nudJ* gene, the flanking regions were amplified by PCR using the primers pairs mx_*nudJ*/*nudJ*_R (443 bp) and *nudJ*_F/mic_*nudJ*_R (452 bp). Mutation, restriction site, and PAM decharacterization was introduced by homologous primers *nudJ*_F and *nudJ*_R. These products were transformed with the origin of replication and the resistance marker into *S. cerevisiae* cells. PCR confirmation of the plasmid assembly was done using *nudJ*_seq_F/*nudJ*_seq_R (309 bp) primers. The donor was amplified by PCR with Donor_*nudJ*_F/donor_*nudJ*_R (814 bp) primers and co-transformed with gRNA plasmid into *E. coli* DH5 α with the *cscBKA* operon integrated and Cas9-expressing plasmid. Selection of bacterial colonies with genomic modification was performed by colony PCR using primers XbaI_SNP_*nudJ*/*nudJ*_XbaI (269 bp). After confirming amplification, fragments were digested with XbaI. The positive modification was confirmed by forming a 210 bp fragment. The gel confirming this single nucleotide change is shown in Fig. S12.

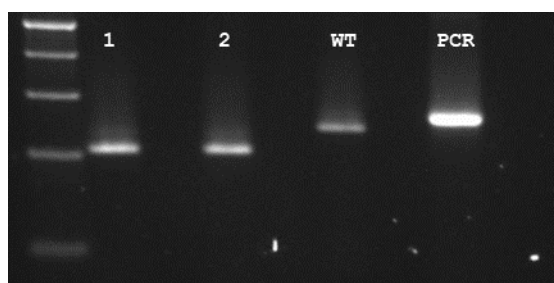


Fig. S12. Digestion of fragment amplified by PCR from genomic DNA for confirmation of nucleotide change in *nudJ* gene. The positive modification was confirmed forming a 210 bp fragment. The numbers (1-2) indicate the clones tested. Wd: without digestion; WT: wild type digested. 100 bp ladder.

***purB* (L115G)**

To clone the region with the SNP of the *purB* gene, the flanking regions were amplified by PCR with primers pairs mx_*purB*/*purB*_R (375 bp) and 2*purB*_F2/mic_*purB*_R (409 bp). These reactions were co-transformed with the origin of replication and the resistance mark into *S. cerevisiae* cells. Mutation and PAM decharacterization was specified by homologous primers 2*purB*_F2 and *purB*_R. Mutation

in the *purB* gene naturally generated a restriction site for *EcoRI*. Colony PCR with primers *purB_seq_F/purB_seq_R* (263 bp) confirmed cloning. The donor was amplified by PCR with *donor_purB_F/Donor_purB_R* (695 bp) primers, and was co-transformed with gRNA plasmid into *E. coli* DH5 α with the *cscBKA* operon integrated and Cas9-expressing plasmid. Selection of bacterial colonies with genomic modification was performed with colony PCR with *purB_seq_F/EcoRI_purb* (325 bp) primers. Corrected amplified fragments were digested with *EcoRI*. The positive modification was confirmed by forming a 179 bp fragment. The gel confirming this single nucleotide change is shown in Fig. S13.

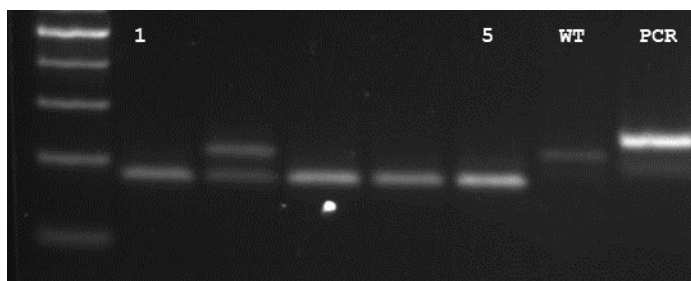


Fig. S13. Digestion of fragment amplified by colony PCR for confirmation of nucleotide change in *purB* gene. The positive modification was confirmed, forming a 179 bp fragment. The numbers (1-5) indicate the colonies tested. Wd: without digestion; WT: wild type digested. 100 bp ladder.

Modifications using the EasyOligo approach

$\Delta flhDC$

The oligo-donor was amplified by PCR with primers L-FlhC/mid-FlhCD/R-FlhD (160 bp). This PCR product was used in CRISPR/Cas9 with corresponding gRNA plasmid in *E. coli* DH5 α cells with the *cscBKA* operon integrated. The confirmation of deletion was made using primers *Check_FlhCD_F/Check_FlhCD_R* (deletion: 222 bp; wild type: 1,103 bp) in colony PCR. The gel for confirmation of this deletion is represented in Fig. S14.

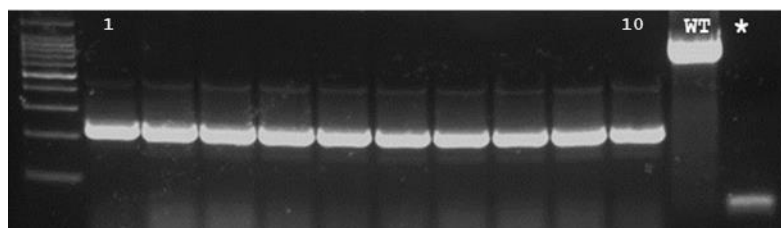


Fig. S14. Colony PCR for confirmation of deletion of *flhCD*. Primers: Check_FlhCD_F/Check_FlhCD_R (deletion: 222 bp; wild type: 1,103 bp). The numbers (1-10) indicate the colonies tested. (*) represents blank; WT: wild type. 100 bp ladder.

ΔhflK

The oligo-donor was amplified by PCR with primers L-hflK/Mid-hflK/R-hflK (160 bp) and employed in CRISPR/Cas9 with the corresponding gRNA plasmid in *E. coli* DH5α cells containing the *cscBKA* operon. Deletion was confirmed by colony PCR using primers Check_hflK_F/Check_hflK_R (deletion: 215 bp; wild type: 1,374 bp). Confirmation of this deletion is shown in Fig. S15.

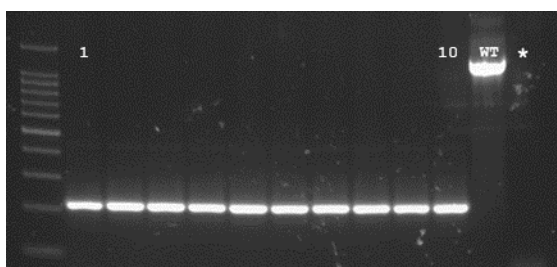


Fig. S15. Colony PCR for confirmation of deletion of *hflK*. Primers: Check_hflK_F/Check_hflK_R (deletion: 215 bp; wild type: 1,374 bp). The numbers (1-10) indicate the colonies tested. (*) represents blank; WT: wild type. 100 bp ladder.

***ΔCA* (12 kb)**

The oligo-donor was amplified by PCR using primers L-galF/Mid-colonic/R-wcal (160 bp). and used in CRISPR/Cas9 with the corresponding gRNA plasmid in *E. coli* DH5α cells containing the *cscBKA* operon. Deletion was confirmed. Confirmation of this deletion is shown in Fig. S14 colony PCR using primers Check_GalK_F/Check_GalK_R (deletion: 239 bp). The gel for confirmation of this deletion is represented in Fig. S16.

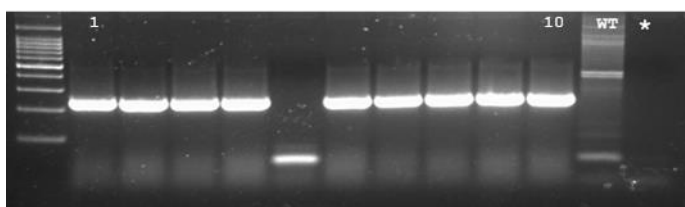


Fig. S16. Colony PCR for confirmation of deletion of CA. Primers: Check_GalK_F/Check_GalK_R (deletion: 239 bp; wild type: ~12 kb). The numbers (1-10) indicate the colonies tested. (*) represents blank; WT: wild type. 100 bp ladder.

ΔgalK

The oligo-donor was amplified by PCR with primers galK_L/Oligo central galK/galK_R (140 bp) and used in CRISPR/Cas9 with corresponding gRNA plasmid in *E. coli* DH5α cells with the *cscBKA* operon integrated. Deletion was confirmed by colony PCR using primers Check_GalK_F/Check_GalK_R (deletion: 246 bp; wild type: 278 bp). The gel confirming this deletion is shown in Fig. 5.

dnaK (T417N)

The oligo-donor was amplified by PCR with primers L2-dnaK/Mid2-dnaK/R2-dnaK (157 bp). and applied in CRISPR/Cas9 with the corresponding gRNA plasmid in *E. coli* DH5α cells with the *cscBKA* integrated. Nucleotide change was made by digestion with PvuII of fragment amplified by colony PCR using primers DnaK2_PvuII_F/DnaK2_seq_R (264 bp). Positive mutants were identified by the presence of a 159 bp fragment after digestion. Sanger sequencing of positive transformants was performed using PCR amplification with primers DnaK2_seq_F/DnaK2_seq_R (268 bp) in both directions. The gel confirming this single nucleotide change is shown in Fig. S17.

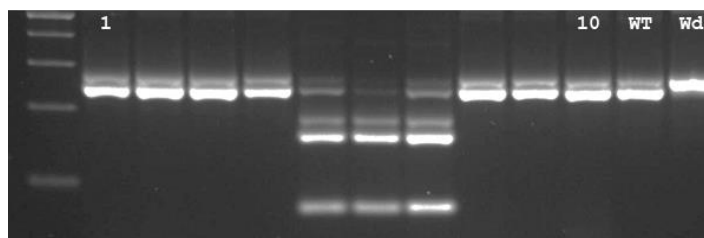


Fig. S17. Digestion of fragment amplified by colony PCR for confirmation of nucleotide change in *dnaK* gene. The positive modification was confirmed forming a 159 bp fragment. The numbers (1-10) indicate the colonies tested. Wd: without digestion; WT: wild type digested. 100 bp ladder.

zwf (D405E)

The oligo-donor was amplified by PCR with primers L-zwf/Mid-zwf/R-zwf (155 bp). This PCR product was used in CRISPR/Cas9 with respective gRNA plasmid in cells with *cscBKA* operon integrated. Nucleotide change was confirmed by digestion with PstI of fragment amplified by colony PCR using primers zwf_PstI_F/zwf_seq_R (302 bp). Positive mutants were selected by 174 bp and 102 fragments presence. Sanger sequencing of positive transformants was made using amplification by PCR with zwf_seq_F/zwf_seq_R (218 bp) primers, in both directions.

APÊNDICE F – Primers usados nas mutações e clonagens

Tabela Suplementar 2- Oligonucleotídeos usados para construção de plasmídeo	
pTargetAmp cloning	
Primer	DNA sequence (5'->3')
Amp_J23119	ACAGCTAGCTCAGTCCTAGGTATAATACTAGTATGAGTATTCAACATTTCCGTGTCGCC
Amp_JacA	CTATTTCTAGCTCTAAAACGAAGATCCTTTGATCTTTTCTACGGGGTCT
Vf_J	TATTATACCTAGGACTGAGCTAGCTGTCAATTTCCATAGGCTCCGCCCCCTGACGAGCATC AC
JacA_Amp	GATCAAAGGATCTTCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGG
A	GGAACAGGAGAGCGCACGAGGGAG
AmpMet	AATAAGGGCGACACGGAAATGTTGAATACTCATA
B	CCTTATTTTAACTTGCTATTTCTAGCTCTAAAAC
pSpec5'	
Primer	DNA sequence (5'->3')
SpecN_f	GCGCAGTAACCTTTTCGTTCCACTGAGCGTCAGACCCCGT
SpecN_R	TGGAACGAAAAGGTTACTGCGCTGTACCAAATGCGGGACAACG
A_jac	CCTTATTTTAACTTGCTATTTCTAGCTCTAAAAC
JacA_Amp	GATCAAAGGATCTTCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGG
P4	CCTTAACGCTATGGAACTCGCCGCC
P3	GGTTACTGCGCTGTACCAAATGCGG
Ori5	TTGCAAGCAGCAGATTACGCGCAG
pSpec3'	
Primer	DNA sequence (5'->3')
Amp_Spec	AGCGTTAAGGGAAGATCCTTTGATCTTTTCTACGGGGTCT
Spec_AmpT	GATCAAAGGATCTTCCCTTAACGCTATGGAACTCGCCGCC
AmpMet	AATAAGGGCGACACGGAAATGTTGAATACTCATA
Amp_J23119	ACAGCTAGCTCAGTCCTAGGTATAATACTAGTATGAGTATTCAACATTTCCGTGTCGCC
Term_Amp	ATTAAGCATTGGTAACTGTCAGACC
Spec_gRNA1	CACTACATTTGCTCATCGCCAGCC
pOliSpec5'	
Primer	DNA sequence (5'->3')
gRNAspec_f	CAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTTTGcATTCTC
NSpec_CAT_r	TAAACAAATAGGGGTTCCGCGGTTACTGCGCTGTACCAAATGCGG

VrSpec_f	CGCAGTAACCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTC
gRNAspec_r	GCACCGACTCGGTGCCACTTTTTCAAGTTG
P4	CCTTAACGCTATGGAACTCGCCGCC
CAT_r1	CCTTAAACGCCTGGTGCTACTC
P5	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGG
SpecPro_r	CGCTGATGACATCAGTCGATCATAG
P4	CCTTAACGCTATGGAACTCGCCGCC
P5	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGG
pOliSpec3'	
Primer	DNA sequence (5'->3')
AmpBam_f	GTATAATAggatccATGAGTATTCAACATTTCCGTGTCGC
AmpSpecC_r	TAGCGTTAAGGGAGGaTcCATGAGATTATCAAAAAGGATCTTCACCT
SpecCAmp_f	TCTCATGgAtCCTCCCTTAACGCTATGGAACTCGCCGCC
J23AmpBam_r	GAATACTCATggatccTATTATACCTAGGACTGAGCTAGCTGTC
P3	GGTTACTGCGCTGTACCAAATGCGG
Term_Amp	ATTAAGCATTGGTAACTGTCAGACC
P6	TATTATACCTAGGACTGAGCTAGCTGTCAATTTCC
P3	GGTTACTGCGCTGTACCAAATGCGG
P6	TATTATACCTAGGACTGAGCTAGCTGTCAATTTCC
pCscBKA cloning	
Primer	DNA sequence (5'->3')
CscB_f	ATCAATAATCAACGCGATATAATAAACCTGGAATGACGACAATGTCCTGGAAATCAGCG
CscA_r	GACTGAGCCTTTTCGTTTTATTTGATGCCTGGGTGACGCGGATCCCTTGCCCGCTGTTGA
SS9_r	GCTGATTTCCAGGACATTGTCGTCATTCCAGGTTTATTATATCGCGTTGATTATTGAT
TB0015_f	TCAACAGCGGGCAAGGGATCCGCGTCACCCAGGCATCAAATAAAACGAAAGGCTCAGTC
SS9_LF1_f	CCTGAGCTTGATCCTACACATACTT
CscBTerm	CCTCATTCATGTTCTGGTACAGTCGCTC
CscA_f	GGAGTGGCTGTGCTGCAACATGGAG
SS9_RF1_r	GATGACGTAATTAAGCATTGATAATTGAG
Cloning minimal vectors	
Primer	DNA sequence (5'->3')
kana_f	CAGTCTAGCTATCGCCATGTAAG
Vf_min	AATAGGGGTTCCGCGTTTTCCATAGGCTCCGCCCCCTGACGAGCATCAC
Vr_min	GCCTATGGAAACGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTC
spec_pro	GGGGGCGGAGCCTATGGAAATCGACCTGCAGAAGCTTAGATCTATTACC

Vf_spec	GCTTCTGCAGGTCGATTTCCATAGGCTCCGCCCCCCTGACGAGCATCAC
prospec_R	GACATCAGTCGATCATAGCACGATCAACGG
Ori_Term_PhaB_R	GTTTCGCCACCTCTGACTTGAGC
AmpT_f	GAGCGTCAGACCCCGTAGAAAAGATCAAAGGATC
Or2_r	AAAAAGAGTTGGTAGCTCTTGATCC
Vf_kan	GCTCCCGGAGACGGTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCAC
pGEX_kan	GGGGCGGAGCCTATGGAAAACCGTCTCCGGGAGCTGCATGTGTCAGAGG
prokan_f	TGTAACATCAGAGATTTTGAGACACGGGCC
pUC19_B0015 Cloning	
Primer	DNA sequence (5'->3')
B0015for	GGGCCTCGTGATACGCCTATTCCAGGCATCAAATAAAACGAAAGG
B0015rev	AATAAACAAATAGGGGTTCCGCGTATAAACGCAGAAAGGCCACCCGAA
Vf_B0015	CTGCGTTTATACGCGGAACCCCTATTTGTTTATTTTCTAAATACATTC
pUC19_B0015	CGTTTTATTTGATGCCTGGAATAGGCGTATCACGAGGCCCTTTCGTCTC
AmpPro	CAGGGTTATTGTCTCATGAGCGGATA
pGFP Cloning	
Primer	DNA sequence (5'->3')
GFP19for	TTTCACACAGGAAACAGCTATGGTGAGCAAGGGCGAGGAGCTG
GFP_B15rev	ATTTGATGCCTGGTTACTTGTACAGCTCGTCCATGCCGAGAGT
B15-GFPfor	GGACGAGCTGTACAAGTAACCAGGCATCAAATAAAACGAAAGG
19GFPprev	GCCCTTGCTCACCATAGCTGTTTCCTGTGTGAAATTGTTATCC
GFP_f	ACCTGAGCACCCAGTCCAAGCTGAG
GFP_r	CTTGTGGCCGTTTACGTCGCCGTCC
pUC19-2for	CCGCTCGCCGCAGCCGAACGACCGAGCGC

Tabela Suplementar 3- Oligonucleotídeos usados para modificações cromossômicas. Os nucleotídeos sublinhados indicam sequência espaçadora de gRNA. Os nucleotídeos sublinhados, em negrito e em itálico indicam sequência espaçadora de gRNA complementar.

SS9::cscBKA		
Primer	DNA sequence (5'->3')	Finality
P2_SS9	TCTGGCGCAGTTGATATGTAGTTTTAGAGCTAGAAATAGCAAGTTAAATAAGG	gRNA cloning with plasmid circularization
P1_SS9	TACATATCAACTGCGCCAGAACTAGTATTATACCTAGGACTGAGCTAGCTGTCAA	
CscB_f	ATCAATAATCAACGCGATATAATAAACCTGGAATGACGACAATGTCTTGGAAATCAGCG	Donor assembly in vivo in <i>E. coli</i>
CscA_r	GACTGAGCCTTTTCGTTTTATTTGATGCCTGGGTGACGCGGATCCCTTGCCCGCTGTTGA	
SS9_r	GCTGATTTCCAGGACATTGTCGTCATTCCAGGTTTATTATATCGCGTTGATTATTGAT	
TB0015_f	TCAACAGCGGGCAAGGGATCCGCGTCACCCAGGCATCAAATAAAACGAAAGGCTCAGTC	
SS9_LF1_f	CCTGAGCTTGATCCTACACATACTT	
CscBTerm	CCTCATTCATGTTCTGGTACAGTCGCTC	
CscA_f	GGAGTGGCTGTGCTGCAACATGGAG	
SS9_RF1_r	GATGACGTAATTAAGCATTGATAATTGAG	
SS9Left1	GGTCTGTATGACAACAAGTGGGTGCTTCACTCAACG	Donor amplification
SS9Right1	AGTTCTGCCAGCTTCTGCTGGGTTTCTTTATCCTGC	
FE1	ACCCGCCACAGTAGTTCCTTCTCACC	Confirmation of insertion
B0015_f2	CTCTACTAGAGTCACACTGGCTCACCTTC	
FD1	CCGACGTCCATCCAGCCACATCGTCC	
CscBTerm	CCTCATTCATGTTCTGGTACAGTCGCTC	
galK::ampR		
Primer	DNA sequence (5'->3')	Finality
P2_galK2	CAAGGCGACCTGAAACGTATGTTTTAGAGCTAGAAATAGCAAGTT	gRNA cloning with plasmid circularization
P1_galK2	ATACGTTTCAGGTCGCCTTGACTAGTATTATACCTAGGACTGAGCTAG	
mx_galK	GAGGGTATTCTGGGCCCTCCATGTCGCTGGTTAACGCCCTGACTTGGCG	Donor assembly in vivo in <i>S. cerevisiae</i>
Amp_galK_R	GATCAAAGGATCTTCTGCCGTACAGCAAGCTGTCGCAGAACAAATATGA	
galK_amp	ACAGCTTGCTGTACGGCAGAAGATCCTTTGATCTTTTCTACGGGGTC	
galK_J23_R	TGCAACTGCGTAACAACAGTTGACAGCTAGCTCAGTCCTAGG	
J23_galK_F	GACTGAGCTAGCTGTCAACTGTTGTTACGCAGTTGCAGATGTTTC	
galK_2mic	TATAAAAATAGGCGTATCACGAGGTGACCGCAGAACAGGCAGCAGAGCG	
2Mic_f	CCTCGTGATACGCCATTTTTATAGGTTAATGTC	
2micMX	CATCCAGTGTCGAAAACGAGCTCTCCACATTTCCCGGAAAAGTGCCACC	
tMX_2m	GCACTTTTCGGGGAAATGTGGAGAGCTCGTTTTTCGACACTGGATGGCGG	
MXProUni	GACATGGAGGCCCAAGAATACCCTCCTTGACAGTCTT	
galK_left_f	GCTGGTTAACGCCCTGACTTGGCGAAAGCG	Donor amplification
Right rev	TGACCGCAGAACAGGCAGCAGAGCGTTTGC	

galM_f	CCATCATCTGAACTCAGGGCAAACAGCACC	Confirmation of insertion
galT_rev	CCGCCACCGTACGTAAATTTATGGTTGG	
AmpMet	AATAAGGGCGACACGAAATGTTGAATACTCATA	
AmpT_f	GAGCGTCAGACCCCGTAGAAAAGATCAAAGGATC	
lacZ::ampR		
Primer	DNA sequence (5'->3')	Finality
P2_lacZ	ATCAGCGGGCGCGTCTCTCCGTTTTAGAGCTAGAAATAGCAAGTT	gRNA cloning with bimolecular complementation
P1_lacZ	GGAGAGACGCGCCCGCTGATACTAGTATTATACCTAGGACTGAGCTAGC	
P3	GGTTACTGCGCTGTACCAAATGCGG	
P4	CCTTAACGCTATGGAACTCGCCGCC	
mx_lacZ	GAGGGTATTCTGGGCCTCCATGTCGCTCTCCGGCTGCGGTAGTTCAGG	Donor assembly in vivo in S. cerevisiae
Amp_lacZ_R	GATCAAAGGATCTTCTCAGTATCCCCGTTTACAGGGCGGCTTCGTCTG	
lacZ_amp_f	TGTAAACGGGGATACTGAGAAGATCCTTTGATCTTTTCTACGGGGTC	
lacZ_J23_R	GTAATCACCCGAGTGTGATTTGACAGCTAGCTCAGTCCTAGG	
J23_lacZ_F	GACTGAGCTAGCTGTCAAATCACACTCGGGTGATTACGATCGCGC	
lacZ_2mic	TATAAAAATAGGCGTATCACGAGGCGCTGATTGAAGCAGAAGCCTGCG	Donor amplification
Donor assembly in vivo in S. cerevisiae	GCTCTCCGGCTGCGGTAGTTCAGGCAGTTC	
donor_lacZ_r	CGCTGATTGAAGCAGAAGCCTGCGATGTCTG	
AmpMet	AATAAGGGCGACACGAAATGTTGAATACTCATA	Confirmation of insertion
AmpT_f	GAGCGTCAGACCCCGTAGAAAAGATCAAAGGATC	
LacZ_F	TGACCATGCGGTCGCGTTTCGGTTGC	
LacZ_R	GAAATCCCGAATCTCTATCGTGCGG	
galk::gfp		
Primer	DNA sequence (5'->3')	Finality
P2_galk2	CAAGGCGACCTGAAACGTATGTTTTAGAGCTAGAAATAGCAAGTT	gRNA cloning with plasmid circularization
P1_galk2	ATACGTTTCAGGTCGCCTTGACTAGTATTATACCTAGGACTGAGCTAG	
termAmp_GFP_f	ATATGAGTAAACTTGGTCTGACAGTTACTTGTACAGCTCGTCCATGCCGAGAGT	Donor assembly in vivo in S. cerevisiae
RBS_J23_R	TACCTCCTTAATTTGATTTCTCTAGTAGCTAGCATTATACCTAGGACTGAGC	
RBS_GFP_F	GAGAAATCAAATTAAGGAGGTAAGATAATGGTGAGCAAGGGCGAGGAGCTG	
AmpTerm_R	CTGTCAGACCAAGTTTACTCATATATACTTTAGATTG	
mx_galk	GAGGGTATTCTGGGCCTCCATGTCGCTGGTTAACGCCCTGACTTGGCG	
galk_2mic	TATAAAAATAGGCGTATCACGAGGTGACCGCAGAACAGGCAGCAGAGCG	
2Mic_f	CCTCGTGATACGCCTATTTTTATAGGTTAATGTC	
2micMX	CATCCAGTGTCGAAAACGAGCTCTCCACATTTCCCGAAAAGTGCCACC	
tMX_2m	GCACTTTTCGGGGAAATGTGGAGAGCTCGTTTTTCGACACTGGATGGCGG	

MXProUni	GACATGGAGGCCCAGAATACCCTCCTTGACAGTCTT	
lacZ_left_f	GCTCTCCGGCTGCGGTAGTTCAGGCAGTTC	Donor amplification
donor_lacZ_r	CGCTGATTGAAGCAGAAGCCTGCGATGTCG	
GFP_f	ACCTGAGCACCCAGTCCAAGCTGAG	Confirmation of insertion
GFP_r	CTTGTGGCCGTTTACGTCGCCGTCC	
LacZ_F	TGACCATGCGGTCGCGTTTCGGTTGC	
LacZ_R	GAAATCCCGAATCTCTATCGTGCGG	
<i>lacZ::gfp</i>		
Primer	DNA sequence (5'→3')	Finality
P2_lacZ	ATCAGCGGGCGCTCTCTCCGTTTTAGAGCTAGAAATAGCAAGTT	gRNA cloning with bimolecular complementation
P1_lacZ	GGAGAGACGCGCCCGCTGATACTAGTATTATACCTAGGACTGAGCTAGC	
P3	GGTTACTGCGCTGTACCAAATGCGG	
P4	CCTTAACGCTATGGAACGCGCCGCC	
termAmp_GFP_f	ATATGAGTAACTTGGTCTGACAGTTACTTGTACAGCTCGTCCATGCCGAGAGT	Donor assembly in vivo in <i>S. cerevisiae</i>
RBS_J23_R	TACCTCCTTAATTTGATTTCTCTAGTAGCTAGCATTATACCTAGGACTGAGC	
RBS_GFP_F	GAGAAATCAAATTAAGGAGGTAAGATAATGGTGAGCAAGGGCGAGGAGCTG	
AmpTerm_R	CTGTCAGACCAAGTTTACTCATATATACTTTAGATTG	
mx_lacZ	GAGGGTATTCTGGGCCTCCATGTCGCTCTCCGGCTGCGGTAGTTCAGG	
lacZ_2mic	TATAAAAATAGGCGTATCACGAGGCGCTGATTGAAGCAGAAGCCTGCG	
lacZ_amp_f	TGTAAACGGGGATACTGAGAAGATCCTTTGATCTTTTCTACGGGGTC	
2Mic_f	CCTCGTGATACGCCTATTTTTATAGGTTAATGTC	
2micMX	CATCCAGTGTGAAAACGAGCTCTCCACATTTCCCGAAAAGTGCCACC	
tMX_2m	GCACTTTTCGGGGAAATGTGGAGAGCTCGTTTTTCGACACTGGATGGCGG	
MXProUni	GACATGGAGGCCCAGAATACCCTCCTTGACAGTCTT	
lacZ_left_f	GCTCTCCGGCTGCGGTAGTTCAGGCAGTTC	
donor_lacZ_r	CGCTGATTGAAGCAGAAGCCTGCGATGTCG	Donor amplification
GFP_f	ACCTGAGCACCCAGTCCAAGCTGAG	Confirmation of insertion
GFP_r	CTTGTGGCCGTTTACGTCGCCGTCC	
LacZ_F	TGACCATGCGGTCGCGTTTCGGTTGC	
LacZ_R	GAAATCCCGAATCTCTATCGTGCGG	
<i>nudJ (G151V)</i>		
Primer	DNA sequence (5'→3')	Finality
P2_nudJ	TTTTAACTGGCCTTTTACAAGTTTTAGAGCTAGAAATAGCAAGTT	gRNA cloning with bimolecular complementation
P1_nudJ	TTGTAAAAGGCCAGTTAAAACTAGTATTATACCTAGGACTGAGCTAGC	
P3	GGTTACTGCGCTGTACCAAATGCGG	
P4	CCTTAACGCTATGGAACGCGCCGCC	

mx_nudJ	GAGGGTATTCTGGGCCTCCATGTCATCTTCGGCGGCAAATTCGAGGAAG	Donor assembly in vivo in <i>S. cerevisiae</i>
nudJ_R	TGGCCTTTTACAAAAGTTGTCATCTAGAAGGCTGCGTGATAGAATACGC	
nudJ_F	TATTCTATCACGCAGCCTTCTAGATGACAACTTTGTAAAAGGCCAG	
mic_nudJ_R	TATAAAAATAGGCGTATCACGAGGTAGTCGTTGAAGAGACGATTAATGG	
2Mic_f	CCTCGTGATACGCCTATTTTTATAGGTTAATGTC	
2micMX	CATCCAGTGTCGAAAACGAGCTCTCCACATTTCCCGAAAAGTGCCACC	
tMX_2m	GCACTTTTCGGGGAAATGTGGAGAGCTCGTTTTCGACACTGGATGGCGG	
MXProUni	GACATGGAGGCCCAGAATACCCCTCCTTGACAGTCTT	Donor amplification
Donor_nudJ_F	CTTCGGCGGCAAATTCGAGGAAGGC	
donor_nudJ_R	TGAAGAGACGATTAATGGTAAAGCG	Confirmation of mutation
XbaI_SNP_nudJ	AACTTCAAGGCGGCGTATTCTATCACG	
nudJ_XbaI	AACGCCGTTTTTTCGCTTTCTCTAGACCATTGAGCTTGAGCAAATATGC	Sequencing
nudJ_seq_F	CCCTGTTGTTGCAACAGCCAGG	
nudJ_seq_R	ACAGCGATATCGACTGCTGCCG	
<i>purB (L115G)</i>		
Primer	DNA sequence (5'→3')	Finality
P2_purB	ATTGATATCTTCCGAAGTACGTTTTCAGAGCTAGAAATAGCAAGTT	gRNA cloning with plasmid circularization
P1_purB	GTACTTCGGAAGATATCAATACTAGTATTATACCTAGGACTGAGCTAGC	
purb_R	TGTACTTCGGAAGATATCAATAACCTCTCC	Donor assembly in vivo in <i>S. cerevisiae</i>
purB_F	GGTTATTGATATCTTCCGAAGTACAGACAAAGTGG	
purB_2mic	CTATAAAAATAGGCGTATCACGAGGTGGAATTATCCTCACTGACCGCCG	
donor_purB_F	ACGAACTCTTCGCTGAACTGATGCC	
Donor_purb_R	TTATCCTCACTGACCGCCGTTTCCC	
2purB_F2	GGTTATTGATATCTTCCGAAGTACATGCAAAGTGG	
2Mic_f	CCTCGTGATACGCCTATTTTTATAGGTTAATGTC	
2micMX	CATCCAGTGTCGAAAACGAGCTCTCCACATTTCCCGAAAAGTGCCACC	
tMX_2m	GCACTTTTCGGGGAAATGTGGAGAGCTCGTTTTTCGACACTGGATGGCGG	
MXProUni	GACATGGAGGCCCAGAATACCCCTCCTTGACAGTCTT	
EcoRI_purb	TGGCCGCGCACGCAGCGATCAAGGGAATTCCTGCTTTTGCTGCCGACGC	Confirmation of cloning and sequencing
purB_seq_F	ATGCCATCAATCAGTTGACGCC	
purB_seq_R	GCAATCGTCGCCAGTTTCAGCG	
<i>cscA::sucP</i>		
Primer	DNA sequence (5'→3')	Finality
P2_cscA1	GGCCCCAGTGTTTCGCTCATGTTTTCAGAGCTAGAAATAGCAAGTTAAAATAAGG	gRNA cloning with bimolecular complementation
P1_cscA1	ATGAGCGAACACTGGGGCCACTAGTATTATACCTAGGACTGAGCTAGCTGTCAA	
P3	GGTTACTGCGCTGTACCAAATGCGG	

P4	CCTTAACGCTATGGAACTCGCCGCC	
SS9L_MX	GGAGGGTATTCTGGGCCTCCATGTCTGGTCTGTATGACAACAAGTGGGTG	Donor assembly in vivo in <i>S. cerevisiae</i>
SS9R_2M	CCTATAAAAATAGGCGTATCACGAGGAGTTCTGCCAGCTTCTGCTGGGT	
sucP_rev	CCTGGTTAAGCAAATTTGGTGAATTGTTATTTTTCATTTTCTCCTCTTT	
SucP_fw	AAAGAGGAGAAAAATGAAAAATAAGGATTCACCAAATTTGCTTAACCAGG	
2Mic_f	CCTCGTGATACGCCTATTTTTATAGGTTAATGTC	
2micMX	CATCCAGTGTGCGAAAACGAGCTCTCCACATTTCCCGAAAAGTGCCACC	
tMX_2m	GCACTTTTCGGGGAAATGTGGAGAGCTCGTTTTTCGACACTGGATGGCGG	
MXProUni	GACATGGAGGCCCAGAATACCCTCCTTGACAGTCTT	
SucP_r	CGTGCCATCACCTAAACGATCCGC	
CscK_f2	CTACGACCGCATCCCCTAAAACCCATAC	
csc_F6	CCCGACCTATAAACCACCTTGTT	Donor amplification
SS9Right1	AGTTCTGCCAGCTTCTGCTGGGTTTCTTTATCCTGC	Confirmation of insertion
SucP_r	CGTGCCATCACCTAAACGATCCGC	
csc_F6	CCCGACCTATAAACCACCTTGTT	
cscA::PJsucP		
Primer	DNA sequence (5'→3')	Finality
P2_cscA1	GGCCCCAGTGTTTCGCTCATGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGG	gRNA cloning with bimolecular complementation
P1_cscA1	ATGAGCGAACACTGGGGGCCACTAGTATTATACCTAGGACTGAGCTAGCTGTCAA	
P3	GGTTACTGCGCTGTACCAAATGCGG	
P4	CCTTAACGCTATGGAACTCGCCGCC	
J23_f	TTCACCAAATTTGCTTAACCAGGTTGACAGCTAGCTCAGTCCTAGGTATAATGCTAG	Donor assembly in vivo in <i>S. cerevisiae</i>
SS9R_2M	CCTATAAAAATAGGCGTATCACGAGGAGTTCTGCCAGCTTCTGCTGGGT	
CscK_r	ATACCTAGGACTGAGCTAGCTGTCAACCTGGTTAAGCAAATTTGGTGAATTGTTAAC	
SS9L_MX	GGAGGGTATTCTGGGCCTCCATGTCTGGTCTGTATGACAACAAGTGGGTG	
2Mic_f	CCTCGTGATACGCCTATTTTTATAGGTTAATGTC	
2micMX	CATCCAGTGTGCGAAAACGAGCTCTCCACATTTCCCGAAAAGTGCCACC	
tMX_2m	GCACTTTTCGGGGAAATGTGGAGAGCTCGTTTTTCGACACTGGATGGCGG	
MXProUni	GACATGGAGGCCCAGAATACCCTCCTTGACAGTCTT	
SucP_r	CGTGCCATCACCTAAACGATCCGC	
CscK_f2	CTACGACCGCATCCCCTAAAACCCATAC	
csc_F6	CCCGACCTATAAACCACCTTGTT	Donor amplification
SS9Right1	AGTTCTGCCAGCTTCTGCTGGGTTTCTTTATCCTGC	Confirmation of insertion
SucP_r	CGTGCCATCACCTAAACGATCCGC	
csc_F6	CCCGACCTATAAACCACCTTGTT	
Δ40.4 kb		

Primer	DNA sequence (5'→3')	Finality
P2_FliA	TCACGAAGCATTGCGCCTGCGTTTTAGAGCTAGAAATAGCAAGTT	gRNA cloning with bimolecular complementation
P1_FliA	GCAGGCGCAATGCTTCGTGAAGTAGTATTATACCTAGGACTGAGCTAGC	
P3	GGTTACTGCGCTGTACCAAATGCGG	
P4	CCTTAACGCTATGGAACCTGCCGCC	
Flage_left_F	AAAGGATATGATTTATCATCGCGTAGC	Donor amplification
Flage_right_R	CGTTACCGCAAAGAAATTAACATAGGGC	
flagel_seq_F	ATCTCCAGAAATAGGGATAACGC	Confirmation of deletion and sequencing
flagel_seq_R	TTGAACAGTCATAGTGATTCC	
<i>ΔfliA</i>		
Primer	DNA sequence (5'→3')	Finality
P2_FliA	TCACGAAGCATTGCGCCTGCGTTTTAGAGCTAGAAATAGCAAGTT	gRNA cloning with bimolecular complementation
P1_FliA	GCAGGCGCAATGCTTCGTGAAGTAGTATTATACCTAGGACTGAGCTAGC	
P3	GGTTACTGCGCTGTACCAAATGCGG	
P4	CCTTAACGCTATGGAACCTGCCGCC	
mx_fliA	GGAGGGTATTCTGGGCCTCCATGTCTCCGGCAACGTTTCGATGGCTTC	Donor assembly in vivo in <i>S. cerevisiae</i>
FliA_R	TATACCGCTGAAGGTGTAAGCTGGATGAACTTCGCAGCCGTGACTG	
FliA_F	CACGGCTGCGAAGTTCATCCAGCTTACACCTTCAGCGGTATAGAGTG	
mic_fliA	CTATAAAAATAGGCGTATCACGAGGCGCTGCGCTTATCAGGCCTACAAG	
2Mic_f	CCTCGTGATACGCCTATTTTTATAGGTTAATGTC	
2micMX	CATCCAGTGTCGAAAACGAGCTCTCCACATTTCCCGAAAAGTGCCACC	
tMX_2m	GCACTTTTCGGGGAAATGTGGAGAGCTCGTTTTTCGACACTGGATGGCGG	
MXProUni	GACATGGAGGCCCAAGATACCTCCTTGACAGTCTT	
Donor_FliA_F	TCCGGCAACGTTTCGATGGCTTCCATCACC	Donor amplification
Donor_FliA_R	CTGCGCTTATCAGGCCTACAAGTTGAATTGC	
Del_fliA_F	TAAACGTTCCGCTACCTCAGTTTCC	Confirmation of deletion
Check_FliA_R	ATTAAGTGAAGTACGCGCAACGCC	
<i>galK::cscBKA</i>		
Primer	DNA sequence (5'→3')	Finality
P2_galK2	CAAGGCGACCTGAAACGTATGTTTTAGAGCTAGAAATAGCAAGTT	gRNA cloning with plasmid circularization
P1_galK2	ATACGTTTCAGGTCGCCTTGACTAGTATTATACCTAGGACTGAGCTAG	
mx_galK	GAGGGTATTCTGGGCCTCCATGTCGCTGGTTAACGCCCTGACTTGGCG	Donor assembly in vivo in <i>S. cerevisiae</i>
AmpTerm_R	CTGTCAGACCAAGTTTACTCATATATACTTTAGATTG	
cscB_galK_F	TTCCAGGACATTGTCGTCATTCTGTTGTTACGCAGTTGCAGATGTTTC	
galK_2mic	TATAAAAATAGGCGTATCACGAGGTGACCGCAGAACAGGCAGCAGAGCG	
term_cscA_F	TATATGAGTAACTTGGTCTGACAGGTGACGCGGATCCCTTGCCCGCTGTTG	

galK_cscB_R	TGCGTAACAACAGGAATGACGACAATGTCTCGGAAATCAGCG	
2Mic_f	CCTCGTGATACGCCTATTTTTATAGGTTAATGTC	
2micMX	CATCCAGTGTGCGAAAACGAGCTCTCCACATTTCCCCGAAAAGTGCCACC	
tMX_2m	GCACTTTTCGGGGAAATGTGGAGAGCTCGTTTTTCGACACTGGATGGCGG	
MXProUni	GACATGGAGGCCGAGAATACCCTCCTTGACAGTCTT	
CscBTerm	CCTCATTCATGTTCTGGTACAGTCGCTC	
right_R	ATTTGATTGGTGAACACACCGAC	
galK_amp	ACAGCTTGCTGTACGGCAGAAGATCCTTTGATCTTTTCTACGGGGTC	
CscA_f	GGAGTGGCTGTGCTGCAACATGGAG	
galK_left_f	GCTGGTTAACGCCCTGACTTGGCGAAAGCG	Donor amplification
Right_rev	TGACCGCAGAACAGGCAGCAGAGCGTTTGC	Confirmation of insertion
CscBTerm	CCTCATTCATGTTCTGGTACAGTCGCTC	
CscA_f	GGAGTGGCTGTGCTGCAACATGGAG	
galM_f	CCATCATCTGAACTCAGGGCAAACAGCACC	
galT_rev	CCGCCACCGTACGTAAATTTATGGTTGG	
<i>lacZ::sucP</i>		
Primer	DNA sequence (5'→3')	Finality
P2_lacZ	ATCAGCGGGCGCGTCTCTCCGTTTTAGAGCTAGAAATAGCAAGTT	gRNA cloning with bimolecular complementation
P1_lacZ	GGAGAGACGCGCCCGCTGATACTAGTATTATACCTAGGACTGAGCTAGC	
P3	GGTTACTGCGCTGTACCAAATGCGG	
P4	CCTTAACGCTATGGAACTCGCCGCC	
mx_lacZ	GAGGGTATTCTGGGCCTCCATGTCGCTCTCCGGCTGCGGTAGTTCAGG	Donor assembly in vivo in <i>S. cerevisiae</i>
AmpTerm_R	CTGTCAGACCAAGTTTACTCATATATACTTTAGATTG	
J23_lacZ_F	GACTGAGCTAGCTGTCAAATCACACTCGGGTGATTACGATCGCGC	
lacZ_2mic	TATAAAAATAGGCGTATCACGAGGCGCTGATTGAAGCAGAAGCCTGCG	
galK_sucP_F	TATGAGTAAACTTGGTCTGACAGTTACGCTACCACAGGCGGGTTGGCG	
galK_J23_R	TGCAACTGCGTAACAACAGTTGACAGCTAGCTCAGTCCTAGG	
2Mic_f	CCTCGTGATACGCCTATTTTTATAGGTTAATGTC	
2micMX	CATCCAGTGTGCGAAAACGAGCTCTCCACATTTCCCCGAAAAGTGCCACC	
tMX_2m	GCACTTTTCGGGGAAATGTGGAGAGCTCGTTTTTCGACACTGGATGGCGG	
MXProUni	GACATGGAGGCCGAGAATACCCTCCTTGACAGTCTT	
lacZ_amp_f	TGTAAACGGGGTACTGAGAAGATCCTTTGATCTTTTCTACGGGGTC	
sucP_R2	TGGAATGGGAGGATAGCGCAGGCG	
SucP_r	CGTGCCATCACCTAAACGATCCGC	
donor_lacZ_r	CGCTGATTGAAGCAGAAGCCTGCGATGTCTG	
lacZ_left_f	GCTCTCCGGCTGCGGTAGTTTCAGGCAGTTT	Donor amplification
donor_lacZ_r	CGCTGATTGAAGCAGAAGCCTGCGATGTCTG	

sucP_R2	TGGAATGGGAGGATAGCGCAGGCG	Confirmation of insertion
SucP_r	CGTGCCATCACCTAAACGATCCGC	
LacZ_F	TGACCATGCGGTCGCGTTCGGTTGC	
LacZ_R	GAAATCCCGAATCTCTATCGTGCGG	
<i>lacZ::phaCAB</i>		
Primer	DNA sequence (5'→3')	Finality
P2_lacZ	ATCAGCGGGCGCGTCTCTCCGTTTATAGAGCTAGAAATAGCAAGTT	gRNA cloning with bimolecular complementation
P1_lacZ	GGAGAGACGCGCCGCTGATACTAGTATTATACCTAGGACTGAGCTAGC	
P3	GGTTACTGCGCTGTACCAAATGCGG	
P4	CCTTAACGCTATGGAACGCGCCGCC	
mx_lacZ	GAGGGTATTCTGGGCTCCATGTCGCTCTCCGGCTGCGGTAGTTCAGG	Donor assembly in vivo in <i>S. cerevisiae</i>
AmpTerm_R	CTGTCAGACCAAGTTTACTCATATATACTTTAGATTG	
lacZ_2mic	TATAAAAATAGGCGTATCACGAGGCGCTGATTGAAGCAGAAGCCTGCG	
phaC_RBS_F	TTTGCCGGTCGCCATTATCTTACCTCCTTAATTTGATTTCTCTAGTAGC	
term phaB_F	AACCTGGTCTGACAGTCAGCCCATATGCAGGCCGCCGTTGAGCG	
RBS phaC_R	AAGGAGGTAAGATAATGGCGACCGGCAAAGGCGCGGCAGCTTCCACG	
2Mic_f	CCTCGTGATACGCCTATTTTTATAGGTTAATGTC	
2micMX	CATCCAGTGTCGAAAACGAGCTCTCCACATTTCCCGAAAAGTGCCACC	
tMX_2m	GCACTTTTCGGGAAATGTGGAGAGCTCGTTTTTCGACACTGGATGGCGG	
MXProUni	GACATGGAGGCCCAGAATACCCTCCTTGACAGTCTT	
lacZ_amp_f	TGTAAACGGGGATACTGAGAAGATCCTTTGATCTTTTCTACGGGGTC	
phaB_rev	TCGGAGGAGTCCGGTTTCTCGACCGG	
phaC_fwd	ACCTTGAATGGTTGGGACTTGCCTTCC	
donor_lacZ_r	CGCTGATTGAAGCAGAAGCCTGCGATGTCTG	
lacZ_left_f	GCTCTCCGGCTGCGGTAGTTCAGGCAGTTC	Donor amplification
donor_lacZ_r	CGCTGATTGAAGCAGAAGCCTGCGATGTCTG	
phaB_rev	TCGGAGGAGTCCGGTTTCTCGACCGG	Confirmation of insertion
phaC_fwd	ACCTTGAATGGTTGGGACTTGCCTTCC	
LacZ_F	TGACCATGCGGTCGCGTTCGGTTGC	
LacZ_R	GAAATCCCGAATCTCTATCGTGCGG	
<i>ΔpcnB (81)</i>		
Primer	DNA sequence (5'→3')	Finality
P2_pcnB	GCTGATGCGCATACCCAATTGTTTTATAGAGCTAGAAATAGCAAGTT	gRNA cloning with bimolecular complementation
P1_pcnB	AATTGGGTATGCGCATCAGCACTAGTATTATACCTAGGACTGAGCTAGC	
P3	GGTTACTGCGCTGTACCAAATGCGG	
P4	CCTTAACGCTATGGAACGCGCCGCC	

Donor_pcnB_F	AAGCCTGGCTGGTTGGCGGCGCGTGGC	Donor amplification
donor_pcnB_R	ACGTCGTTTCATCGCCAGCGCGAAAGCG	
pcnB_F	CCGCGATTTCACTATCAACAGCC	Confirmation of deletion
pcnB_R	TAGGTTTCGTAACCGTAGCCCGC	
<i>ΔgalK</i> (32 bp)		
Primer	DNA sequence (5'→3')	Finality
oligo_gRNA_galK	AGGTATAATACTAGTCAAGGCGACCTGAAACGTATGTTTTAGAGCTAGAA	gRNA cloning with oligo
pE	GTCAGGGGGGCGGAGCCTATGGAAATTGACAGCTAGCTCAGTCCTAGGTATAATACTAGT	
gE	ACTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAC	
Galk_L	AGTGTCAATTTGCGGCACGGTGATTTGAAATCATCGCGCATAGAGGCATGAGACTCCGC	Donor amplification
Galk_R	GGACCCAATCGTGGCGAAACGCGTGCGTCATATCCTGACTGAAAACGCCCCGACCGTTGA	
Oligo central galK	GGCATGAGACTCCGCCATCAACTCGGCTGGCAGCTTCAACGGTGCGGGCG	
Check_K_Galk_F	CATGCGTACGCCACCTTTGTCCG	Confirmation of deletion
Check_K_Galk_R	AGCCAGCTCTGCGCGATGTCACC	
<i>ΔCA</i> (12 kb)		
Primer	DNA sequence (5'→3')	Finality
C-gRNA1-wcaM	AGGTATAATACTAGTATCAACTGGGCGATCGGCATGTTTTAGAGCTAGAA	gRNA cloning with oligo
pE	GTCAGGGGGGCGGAGCCTATGGAAATTGACAGCTAGCTCAGTCCTAGGTATAATACTAGT	
gE	ACTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAC	
L-galF	CATCCGGTCAGATTTTCATTATTCGCTTAACAGCTTCTCAATACCTTTACGGAACCTCGC	Donor amplification
Mid-colonic	GGAACCTTCGCCCCCTCTTTTCAGGTTGCGTAAACTCCGGCGAGTAGTTAATGCCGTAGACC	
R-wcaI	CTATTTTCTCGCTGAGAAGCGTACCGGAGTACCGGATTATGAAAATACTGGTCTACGGC	
Check_galF_F	TACTTTTGCTGCGAATTTTCACTGCCG	Confirmation of deletion and sequencing
Check_wcaI_R	GTGATAATGTTTCATGCTAACAGCCGCG	
<i>dnaK</i>		
Primer	DNA sequence (5'→3')	Finality
C-gRNA2-dnaK	AGGTATAATACTAGTGCTGTGCTTGGTCGGGATAGGTTTTAGAGCTAGAA	gRNA cloning with oligo
pE	GTCAGGGGGGCGGAGCCTATGGAAATTGACAGCTAGCTCAGTCCTAGGTATAATACTAGT	
gE	ACTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAC	
L2-dnaK	CTGCTGGACGTTACCCCGCTGTCTCTGGGTATCGAAACCATGGGCGGTGTGATGACGACG	Donor amplification
Mid2-dnaK	GTGATGACGACGCTGATCGCGAAAAACACAAATATCCCGACCAAGCACAGCCAGGTGTTT	
R2-dnaK	CCTGCAGCACATGGATGGTTACCGCAGACTGGTTGTCTTCAGCTGTAGAGAACACCTGGC	
DnaK2_PvuII_F	ACGAAGCTGTAGCAATCGGTGCAGCTGTTTCAGGGTGGTGTCTGACTGG	Confirmation of mutation
DnaK2_seq_R	GACCCAGAGATTTGTTATCAGCCGC	
DnaK2_seq_F	CCGGACGAAGCTGTAGCAATCGG	Sequencing

DnaK2_seq_R	GACCCAGAGATTTGTTATCAGCCGC	
<i>ΔflhDC</i>		
Primer	DNA sequence (5'→3')	Finality
C-gRNA1-FlhC	AGGTATAATACTAGTCTGTTCAAGGTATAAACGGTGTTTTAGAGCTAGAA	gRNA cloning with oligo
pE	GTCAGGGGGGCGGAGCCTATGGAAATTGACAGCTAGCTCAGTCCTAGGTATAATACTAGT	
gE	ACTTTTTCAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAC	
L-FlhC	ATTGGGGCAAAAAAAGCAGCGGTACGTCGTTACCGCTGCTGGAATGTTGCGCCTCACCG	Donor amplification
mid-FlhCD	CGCCTCACCGTATCAGTTAAACAGCCTGTATGTCATAAATGTGTTTCAGCAACTCGGAGG	
R-FlhD	GAAACCGCATAAAAAATAAGTTGGTTATTCTGGGTGGGAATAATGCATACCTCCGAGTTG	
Check_FlhCD_F	GACAGGATGTTTCAGTCGTCAGGCG	Confirmation of deletion
Check_FlhCD_R	GGCTTCCC GGCGACATCACGGGG	
<i>ΔhflK</i>		
Primer	DNA sequence (5'→3')	Finality
C-gRNA1-hflK	AGGTATAATACTAGTGGCGGTGGCAGTTCATCGCAGTTTTAGAGCTAGAA	gRNA cloning with oligo
pE	GTCAGGGGGGCGGAGCCTATGGAAATTGACAGCTAGCTCAGTCCTAGGTATAATACTAGT	
gE	ACTTTTTCAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAC	
L-hflK	AAAGACAACAGGGATCACCGCATAACAAATATGGAGCACAAACATGGCGTGGAATCAGCC	Donor amplification
Mid-hflK	GGAATCAGCCCGTAATAACGGACAAGACCCGATATTATGGACCAACGCCGCGCCAACGC	
R-hflK	ACTGACTTACGCATCGTTATTCCCCCTGACGCTGGTAGTCGTTACGCTGCGCGTTGGCGC	
Check_hflK_F	GCGTGGAGTCATATCCTCTGGCG	Confirmation of deletion
Check_hflK_R	CTACCAGCACGATGATGATAATCGCG	
<i>zwf</i>		
Primer	DNA sequence (5'→3')	Finality
C-gRNA1-zwf	AGGTATAATACTAGTTGAATAGCTCAGATCCAGCTGTTTTAGAGCTAGAA	gRNA cloning with oligo
pE	GTCAGGGGGGCGGAGCCTATGGAAATTGACAGCTAGCTCAGTCCTAGGTATAATACTAGT	
gE	ACTTTTTCAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAC	
L-zwf	CTATCCGTCTGCAACCTGATGAAGGCGTGGATATCCAGGTACTGAATAAAGTTCCTGGCC	Donor amplification
Mid-zwf	GTTCTTGGCCTTGACCACAAACATAACCTGCAGATCACAAGCTGGAGCTGAGCTATTC	
R-zwf	CAAACGTTTCATAGGCATCCGCCAGATGCGTCTGATTAAAGGTTTCTGAATAGCTCAGCTC	
zwf_PstI_F	TACCTGCGTACTGGTAAACGTCTGCAGACCAAAATGTTCTGAAGTCGTGG	Confirmation of mutation
zwf_seq_R	GCCTGAATACCACGCATGGTTTCC	
Zwf_seq_F	GAATCGTGGCAGGATCTGCCGC	Sequencing
zwf_seq_R	GCCTGAATACCACGCATGGTTTCC	

ANEXO A – Publicações em periódicos especializados

BARRETO, J. A.; E SILVA, M. V. M. L.; MARIN, D. C.; BRIENZO, M. et al. Engineering Adaptive Alleles for *Escherichia coli* Growth on Sucrose Using the EasyGuide CRISPR System. p. 2024.2012.2020.629655, 2024.

doi: <https://doi.org/10.1101/2024.12.20.629655>

Artigo em revisão na Journal of biotechnology. Publicado como preprint na plataforma bioRxiv. Esse manuscrito está apresentado a partir da próxima página.

JACOBUS, A. P.; BARRETO, J. A.; DE BEM, L. S.; MENEGON, Y. A. et al. EasyGuide Plasmids Support in Vivo Assembly of gRNAs for CRISPR/Cas9 Applications in *Saccharomyces cerevisiae*. ACS Synth Biol, 11, n. 11, p. 3886-3891, Nov 18 2022. doi: <https://pubs.acs.org/doi/10.1021/acssynbio.2c00348>

JACOBUS, A. P.; CAVASSANA, S. D.; DE, O., II; BARRETO, J. A. et al. Optimal tradeoff between boosted tolerance and growth fitness during adaptive evolution of yeast to ethanol shocks. Biotechnol Biofuels Bioprod, 17, n. 1, p. 63, May 10 2024. doi: <https://doi.org/10.1186/s13068-024-02503-7>

DE BEM, L. S.; BARRETO, J. A.; SOUZA, D. T.; DOS SANTOS, L. et al. QTL mapping, breeding, and debugging *Saccharomyces cerevisiae* strains through Reiterated Mass Selection and Backcrossing (ReMaSSing). Em preparação para submissão para Biotechnology for Biofuels and Bioproducts..

Engineering Adaptive Alleles for *Escherichia coli* Growth on Sucrose Using the EasyGuide CRISPR System

Joneclei Alves Barreto^{a,b}, Matheus Victor Maso Lacôrte e Silva^{a,c}, Danieli Canaver Marin^{a,b}, Michel Brienzo^{a,b}, Ana Paula Jacobus^{a,b}, Jonas Contiero^{a,b,c}, Jeferson Gross^{a,b*}

^a*São Paulo State University (Unesp), Institute for Research in Bioenergy, Rio Claro, 13500-230, SP, Brazil*

^b*PhD Program in Bioenergy, São Paulo State University (Unesp), Rio Claro, 13500-230, Brazil*

^c*São Paulo State University (Unesp), Institute of Biosciences, Rio Claro, 13506-900, SP, Brazil*

***Corresponding author: jeferson.gross@unesp.br**

Abstract

Adaptive Laboratory Evolution (ALE) is a powerful approach for mining genetic data to engineer industrial microorganisms. This evolution-informed design requires robust genetic tools to incorporate the discovered alleles into target strains. Here, we introduce the EasyGuide CRISPR, a five-plasmid platform that exploits *E. coli*'s natural recombination system to assemble gRNA plasmids from overlapping PCR fragments. The production of gRNAs and donor DNA is further facilitated by using recombination cassettes generated through PCR with 40 to 60-mer oligos. With the new CRISPR toolkit, we constructed 22 gene edits in *E. coli* DH5α, most of which corresponded to alleles

mapped in *E. coli* DH5 α and E2348/69 ALE populations selected for sucrose propagation. For DH5 α ALE, sucrose consumption was supported by the *cscBKA* operon expression from a high-copy plasmid. During ALE, plasmid integration into the chromosome, or its copy number reduction due to the *pcnB* deletion, conferred a 30–35% fitness gain, as demonstrated by CRISPR-engineered strains. A ~5% advantage was also associated with a ~40.4 kb deletion involving *fli* operons for flagella assembly. In E2348/69 ALE, inactivation of the *hfl* system suggested selection pressures for maintaining λ -prophage dormancy (lysogeny). We further enhanced our CRISPR toolkit using yeast for in vivo assembly of donors and expression cassettes, enabling the establishment of polyhydroxybutyrate synthesis from sucrose. Overall, our study highlights the importance of combining ALE with streamlined CRISPR-mediated allele editing to advance microbial production using cost-effective carbon sources.

Keywords: *Escherichia coli*, CRISPR/Cas9, Adaptive laboratory evolution, *cscBKA*, sucrose, metabolic engineering

1. Introduction

Adaptive Laboratory Evolution (ALE) is a dynamic area of biological research where organisms undergo evolutionary changes in controlled laboratory environments (Barrick and Lenski, 2013). By selecting for adaptive mutations and enhancing phenotypes, ALE serves as a powerful tool for developing industrial microbial strains with increased production capacity and improved stress tolerance (Dragosits and Mattanovich, 2013; Menegon et al., 2022). Moreover, the application of next-generation sequencing to ALE populations or evolved clones facilitates the precise identification of adaptive alleles that

are enriched under selection. This knowledge allows for reverse-engineering evolved mutations back into parent strains, providing insights into allele-specific fitness and guiding the rational design of industrial strains (Barrick and Lenski, 2013; Menegon et al., 2022).

Escherichia coli is a preferred microorganism for biotechnological applications due to its ease of cultivation and robust tools for genetic modification (Blount, 2015; Jiang et al., 2015). However, most *E. coli* strains are unable to utilize sucrose, a cost-effective carbon source derived from abundant feedstocks like sugarcane and beet crops (Bruschi et al., 2012). To enable sucrose consumption, industrial and laboratory strains can be engineered with the *cscB*, *cscK*, and *cscA* genes, which are naturally found in the operons of *E. coli* W and enteropathogenic strains (Bruschi et al., 2012; Olavarria et al., 2019). However, implementing a heterologous pathway often requires ALE to ensure seamless integration into the host's metabolic network and to identify genetic targets that can be reverse-engineered to enhance sucrose utilization (Mohamed et al., 2019).

Reconstructing ALE mutations necessitates the use of robust genetic tools (Menegon et al., 2022). Over the past decade, CRISPR/Cas9 has emerged as the premier genome editing technology, enabling gene deletions, the integration of heterologous pathways, and precise single-nucleotide modifications (Dong et al., 2021). Numerous studies have highlighted strategies for genome editing in *E. coli* using CRISPR/Cas9 (Table 1). These typically involve plasmid-based co-expression of *Streptococcus pyogenes* Cas9 and the λ -Red Exo/Bet/Gam recombinase system (Dong et al., 2021). This phage-derived recombinase enhances the chromosomal integration of donor DNAs in *E. coli*, serving as templates for genome edits (Bassalo et al., 2016; Jiang et al., 2015).

72 However, recent research indicates that RecA-proficient *E. coli* strains can achieve high-
73 efficiency gene editing with attenuated Cas9/Cas12 targeting (Collias et al., 2023). A
74 crucial element of the CRISPR protocol is the engineering of guide RNAs (gRNAs), which
75 harbor a 20-nt spacer that directs Cas9 to its genomic targets (Jacobus et al., 2022). Each
76 planned genome modification requires the cloning of a new gRNA into a separate plasmid,
77 which is then introduced into a host strain expressing the Cas9/ λ -Red recombinase
78 system (Table 1).

79 The construction of donor DNAs is also a major bottleneck in CRISPR/Cas9
80 experiments, primarily due to the need for assembling large homology arms (HA) of 100-
81 600 bp flanking the editing sequence (Bassalo et al., 2016; Jiang et al., 2015). This often
82 requires precloning, increasing both cost and time of the genome edit. Despite numerous
83 studies outlining various strategies for cloning gRNAs and donors (Table 1), there is still
84 a demand for simple, cost-effective techniques (Shukal et al., 2022). In vivo cloning in *E.*
85 *coli* provides a practical and economical alternative to traditional restriction-ligation
86 methods, potentially eliminating the need for expensive cloning kits based on Gibson and
87 Golden Gate assemblies (Watson and Garcia-Nafria, 2019). This method leverages *E.*
88 *coli*'s natural ability to recombine DNA fragments with short overlapping homologies (10-
89 50 bp) (Jacobus and Gross, 2015; Kostylev et al., 2015). Although the precise
90 mechanisms of this gap-repair cloning are not fully understood, it likely involves a RecA-
91 independent pathway with 3' to 5' exonucleolytic activity (ExoIII) generating single-
92 stranded DNA (ssDNA) termini (Conley et al., 1986; Nozaki and Niki, 2019). These ssDNA
93 ends can anneal and be repaired, possibly by DNA polymerase I (*polA*) and DNA ligase
94 (Nozaki and Niki, 2019).

In this study, we conducted an ALE experiment to improve *E. coli* growth using sucrose as the sole carbon source. To streamline the CRISPR/Cas9 workflow for reverse-engineering ALE mutations, we adapted the EasyGuide approach—originally developed for yeast—for *E. coli* (Jacobus et al., 2022). We introduce a five-plasmid platform to facilitate PCR generation of DNA components for assembling gRNA-encoding plasmids via in vivo cloning in *E. coli*. Additionally, we present the EasyOligo approach, enabling straightforward PCR amplification of gRNAs and donors. Our strategies for in vivo donor assembly in yeast also facilitates the construction of heterologous expression cassettes, advancing the metabolic engineering of polyhydroxybutyrate (PHB) production from sucrose in *E. coli*.

Table 1

Key CRISPR approaches using Cas9 endonuclease for genome edition in *E. coli*

Reference	Cas9 plasmid and gRNA cloning	Donor production
(Jiang et al., 2013)	Two-plasmid system (Crispr RNA and Cas9). gRNA cloned via restriction-ligation.	Linear oligonucleotide donor.
(Jiang et al., 2015)	pTargetF (gRNA) and pCas9 (Cas9). gRNA cloned by inverse PCR, or restriction-ligation.	Plasmid donor (250/550 bp HA), cloned by overlap PCR and restriction-ligation. Linear PCR Donors.
(Li et al., 2021)	Two-plasmid system. gRNA complementary oligos ligated into BsaI-prepared plasmid.	Linear donor.
(Li et al., 2023)	Two-plasmids. p15A-Cas9 (high-copy number). gRNA oligos ligated into a BsaI-cut plasmid.	Linear donor.
(Reisch and Prather, 2015)	Two-plasmid system (gRNA/ λ -Red and Cas9). gRNA cloned via CPEC.	Donors as oligonucleotides: 60 – 90 nts ssDNA or dsDNA oligos.
(Pyne et al., 2015)	Three-plasmid system (gRNA, Cas9, and λ -Red). gRNA oligos ligated into BsaI-prepared plasmid.	60-nts ssDNA donors. dsDNA donors from <i>lacZ</i> amplification by PCR. The primers provide 40 bp HA.

(Bassalo et al., 2016)	gRNA plasmid assembled through CPEC. Four plasmids for gRNA, Cas9, λ -Red, and donor.	Circular donors cloned with CPEC (50 – 600 bp HA). Linear 10 kb integration cassette cloned in yeast.
(Zerbini et al., 2017)	Two-plasmid system. gRNA oligonucleotides ligated into a Bsal-prepared plasmid.	ssDNA and dsDNA oligos (70 nt and 120 nts).
(Shukal et al., 2022)	gRNA prepared through restriction-free cloning. Two-plasmid system (gRNA and Cas9/ λ -Red).	Linear donors with asymmetric HA (50 bp / 500 bp) prepared by PCR.
(Zhao et al., 2016)	gRNA plus donor assembled via Golden Gate. One plasmid for Cas9, λ -Red, gRNA and donor.	Donor with 41 – 300 bp HA cloned with the gRNA via Golden Gate assembly.
(Li et al., 2015)	Two-plasmid system. PCR-amplified plasmid with primer-encoded gRNA, Golden Gate assembly.	Donors with 300 – 500 bp HA produced via fusion PCR.
(Fang et al., 2024)	Combined plasmid for gRNA and Cas9. gRNA cloning through MetClo assembly.	Plasmid donor cloned through MetClo assembly.
(Wang et al., 2021)	Two-plasmid system. gRNA cloned via PCR. Rock-Paper-Scissors approach for plasmid curing.	Donor cloned into the gRNA plasmid by using seamless clone kits or restriction-ligation.
(Lim et al., 2023)	Cas9-harboring strain. gRNA plasmid cloning via Gibson assembly. Use of 5'-truncated gRNAs.	Donors as 41–70-mer oligonucleotides.
(Feng et al., 2018)	Three-plasmid system (Cas9/ λ -Red, gRNA, and donor). Golden Gate assembly.	Donors cloned and delivered within plasmids. Golden Gate assembly.
(Moreb et al., 2017)	Two-plasmid system (gRNA/ λ -Red and Cas9/recA56). gRNA cloning with commercial kits.	Donors as ssDNA and dsDNA oligos.
(Sun et al., 2019)	Two-plasmid system (Cas9/ λ -Red and gRNA). gRNA cloning through inverse PCR.	Linear donors assembled through fusion PCR.
(Tang et al., 2017)	Three-plasmid system (Cas9, λ -Red and gRNA/SceI). gRNA cloned through PCR.	Linear dsDNA donor supplied as a PCR product assembled through overlap extension PCR.
(Zhao and Guo, 2024)	Two plasmids: Cas9/Beta and gRNA, cloned by restriction-ligation and homologous recombination.	Linear dsDNA generated through extension overlap PCR.

106

107 2. Materials and methods

108 2.1 Strains and growth conditions

E. coli DH5 α , E2348/69, and yeast strains are listed in Supplementary Table 1. The purity of *E. coli* cultures was confirmed using MacConkey agar (Merck KGaA, Darmstadt, Germany) and Gram staining (Laborclin, Pinhais, PR, Brazil). *E. coli* was cultured in liquid LB medium (1% tryptone, 0.5% yeast extract, 1% NaCl) at 37 °C with 200 rpm shaking or on solid medium with 1.5% agar. Antibiotics used included ampicillin (100 μ g/mL), kanamycin (50 μ g/mL), spectinomycin (50 μ g/mL), and chloramphenicol (50 μ g/mL). Cells were stored in 25% glycerol at -80 °C. Yeast cells were cultured in YPS medium (1% yeast extract, 2% peptone, 2% sucrose) with hygromycin (300 μ g/mL) or nourseothricin (100 μ g/mL) as needed.

2.2 PCR and molecular cloning procedures

PCR oligonucleotides were ordered from Exxtend Biotechnology (Campinas, Brazil) and are listed in Supplementary Tables 2 and 3. DNA fragments for cloning and transformations were produced by PCR with Phusion High-Fidelity DNA Polymerase (Thermo Fisher Scientific, Waltham, MA, USA) or Genie Fusion Ultra High-Fidelity DNA Polymerase (Assay Genie, Dublin, Ireland). GoTaq™ DNA Polymerase (Promega, Madison, WI, USA) and Platus Taq DNA Polymerase (Sinapse Inc, São Paulo, Brazil) were used for diagnostic PCRs. PCR products were analyzed on 1–3% agarose gels in TBE buffer with 0.01% ethidium bromide. For colony PCR, cells from a colony were heated in water, and the lysate was used for diagnostic PCRs. PCR fragments were purified with the QIAquick PCR Purification Kit or QIAquick Gel Extraction Kit (QIAGEN, Hilden, Germany) as necessary.

131 In the EasyOligo method for generating gRNAs and donor DNAs, oligos were
132 prepared at 10 pmol/μL. For PCR, 2.5 μL of flanking oligos and 1.25 μL of the central oligo
133 were used. PCR involved an initial denaturation at 95 °C, 30 cycles of 95 °C for 10
134 seconds, 50 °C for 15 seconds, and 72 °C for 10 seconds, with a final extension at 72 °C
135 for 1 minute. Products were analyzed via electrophoresis on 3% agarose gels.

136 Cloning in *E. coli* DH5α employed the in vivo technique (Jacobus and Gross, 2015).
137 For each fragment, 5 μL of the PCR reaction was used for transformation. If needed, 10
138 U of DpnI (New England Biolabs, Ipswich, MA, USA) was added to the PCR tube.

139 Competent cells were prepared using the rubidium chloride method (Hanahan,
140 1983). For heat shock transformations (Maniatis et al., 1982), DNA-mixed cells were
141 incubated at 42 °C for 1 minute, chilled on ice for 20 minutes, then recovered in 1 mL LB
142 medium at 37 °C with shaking at 200 rpm for 1 hour. After recovery, they were centrifuged,
143 plated on LB agar with antibiotic, and incubated at 37 °C for 12 to 16 hours. For *E. coli*
144 electroporation (Lessard, 2013), 500 μL of overnight culture was grown in 50 mL LB until
145 OD600 reached 0.5-0.7. Cells were centrifuged, washed twice with 10% glycerol,
146 resuspended in 500 μL of 10% glycerol, and aliquoted into 50 μL portions for
147 electroporation or freezing. Electroporation was performed at 2.5 kV for 5 ms in a
148 MicroPulser electroporator (Bio-Rad, Hercules, California, USA) using 0.2 cm cuvettes.
149 Post-electroporation, cells were recovered in 1 mL LB at 37 °C with shaking for 1 hour,
150 and 50 μL was plated on antibiotic media for selection.

151 Plasmids are listed in Supplementary Table 4. For DNA plasmid preparation, *E. coli*
152 DH5α cultures were grown in 10 mL of LB medium with the appropriate antibiotic at 200
153 rpm for 12–16 hours. Plasmids were extracted using the QIAprep® Miniprep kit (QIAGEN,

Hilden, Germany) and quantified with a Qubit 2.0 Fluorometer (Life Technologies, Carlsbad, CA, USA). Plasmid DNA was stored at -20 °C.

For cloning donor DNA plasmids in yeast, the 2 μ replication region and resistance markers were amplified from plasmids pEasyG3-mic, pEasyG3-hph, and pEasyG3-nat (Jacobus et al., 2022) using the primer pairs 2Mic_f/2micMX and tMX_2m/MXProUni, respectively. Transformations used 5 μ L of PCR products introduced into *Saccharomyces cerevisiae* CEN.PK113-7D via the lithium acetate method (Gietz and Woods, 2002); transformed cells were selected on YPS plates with hygromycin or nourseothricin. Plasmids were enriched from *S. cerevisiae* using a QIAprep® Miniprep kit (QIAGEN, Hilden, Germany) after yeast lysis with zirconia beads.

2.3 Genomic modifications through the EasyGuide CRISPR system

Detailed procedures for constructing plasmids pTargetAmp, pSpec5', pSpec3', pOliSpec5', and pOliSpec3' are provided in Supplementary File 1. These plasmids are available at Addgene (<https://www.addgene.org/>) under the catalog numbers 228548–228552, and their maps are displayed in Supplementary Figs. 1–5.

To assemble gRNA plasmids using pTargetAmp as a template, PCR was performed with primers P1 and P2, both containing 20-nt spacers. Plasmid pSpec5' was amplified with P2/P3, and pSpec3' with P1/P4. PCR products for pTargetAmp, pSpec5', and pSpec3' were transformed into *E. coli* DH5 α via heat shock (Maniatis et al., 1982). For pOliSpec5' and pOliSpec3', PCR used P2/P3 and P1/P4, respectively, with P1 and P2 serving as priming regions only. These products were cotransformed with oligo-amplified gRNAs, and successful cloning was confirmed by diagnostic PCR with primers A and B.

The gRNA oligonucleotides are listed in Supplementary Table 3, with gRNA assembly details in Supplementary File 1.

The temperature-sensitive plasmid CAS9BAC1P-1EA (Merck KGaA, Darmstadt, Germany), encoding the *Streptococcus pyogenes* Cas9 and λ -Red recombinase, was transformed into *E. coli* DH5 α and used according to the manufacturer's instructions. An overnight DH5 α culture with CAS9BAC1P-1EA was diluted 1:100 in 50 mL LB containing kanamycin and grown at 30 °C and 200 rpm. Upon reaching an OD600 of 0.2–0.3, L-arabinose was added to 10 mM. Cells at OD600 0.5–0.7 were centrifuged and resuspended in 2 mL. Aliquots of 50 μ L were used for the heat shock transformation (Hanahan, 1983), or occasional use of electroporation (Lessard, 2013).

For CRISPR/Cas9 experiments, 50 μ L cells with the CAS9BAC1P-1EA plasmid were transformed via heat shock with 10 μ L of donor PCR product and 40 ng of gRNA plasmid. Controls included only the gRNA plasmid and the pSpec plasmid. After recovery in LB with 10 mM L-arabinose at 30 °C and 200 rpm, cells were plated on LB agar containing kanamycin, spectinomycin, and L-arabinose. Colony PCRs confirmed genetic modifications. The gRNA plasmid was removed through successive LB passages without antibiotics, and the CAS9BAC1P-1EA was eliminated by propagation at 37 °C. All genetic modifications are described in Supplementary File 1.

2.4 Adaptive laboratory evolution

E. coli E2348/69 (Iguchi et al., 2009) was transformed with pKan, pSpec, pAmp, and pGGA plasmids (Supplementary File 1) to create four distinct ALE populations: E-Kan, E-Spec, E-Amp, and E-Cm. *E. coli* DH5 α was first transformed with the pCscBKA plasmid

(Supplementary Fig. 6), followed by the pKan, pSpec, and pGGA plasmids to produce the D-Kan, D-Spec, and D-Cm populations.

ALE was conducted by serially propagating each population in minimal medium with 2% sucrose (Supplementary Methods). The process included daily 1:100 dilutions in 20 mL of medium within 50 mL Erlenmeyer flasks at 37 °C with shaking at 80 rpm. OD600 was measured daily. Each population was maintained on its respective antibiotic to prevent contamination, and ampicillin was used for DH5 α populations to maintain pCscBKA. Purity was periodically checked using MacConkey agar and Gram staining. Weekly glycerol stocks were made, and propagation lasted about three months.

2.5 Genome DNA sequencing and variant calling

Genomic DNA from the final and progenitor populations was purified using the phenol:chloroform (1:1) extraction (Maniatis et al., 1982). DNA quality was checked with 1% agarose gel electrophoresis, and quantification was done with a Qubit 2.0 Fluorometer (Life Technologies, Carlsbad, CA, USA). Libraries were prepared with the NEBNext® Ultra™ II DNA Library Prep Kit (New England BioLabs, Ipswich, MA, USA), and 2 x 150 bp paired-end sequencing (Q30 > 85%) was performed on the Illumina NovaSeq 6000 platform (GenOne, Solutions in Biotechnology, Rio de Janeiro, Brazil). Sequencing reads were deposited in NCBI under BioProject number PRJNA1200122.

Reads were filtered to retain those with a Phred quality score of at least 30 and a length of 75 bases. Mapping was conducted using the Burrows-Wheeler aligner in CLC Genomics Workbench 8.01 (QIAGEN, Aarhus, Denmark) with a cutoff of 0.8 for both read length and identity (Jacobus et al., 2024). DH5 α reads were mapped to the DH5 α

reference genome (GenBank CP017100), and E2348/69 reads were mapped to their reference chromosomes (GenBank FM180568, FM180570, FM180569), excluding duplicate reads. Variants with a frequency of $\geq 30\%$ specific to evolved populations were identified. Read mappings against plasmids were also performed. Plasmid copy numbers for each ALE population were calculated by normalizing read coverage over antibiotic resistance markers to their length and dividing by genomic read coverage (Jacobus et al., 2021).

2.6 Phenotypic analyses

Growth curves of parent and evolved populations were measured using a Tecan Sunrise microplate reader (Tecan, Männedorf, Switzerland). After two preadaptation passages in minimal medium with 2% sucrose, cells were counted with an Attune NxT flow cytometer (Thermo Fisher Scientific, Waltham, MA, USA). A 5 μL dilution ($\sim 10^4$ cells) was used to inoculate 200 μL of the same medium in a 96-well microplate, with four replicates per population. Growth at 37 °C was monitored by recording OD600 every 15 minutes. Maximum specific growth rates (μ_{max}) were determined during the exponential phase by plotting the natural logarithm of OD600 against time and fitting a linear regression. The μ_{max} for each replicate was determined from the slope and averaged (Rodrigues et al., 2021).

To assess the fitness of evolved DH5 α populations, D-Kan, D-Cm, and D-Spec strains (expressing pCscBKA with AmpR) were competed against a tester-GFP strain (SS9:*cscBKA*; *lacZ::gfp*; *galK::gfp*) with respective marker plasmids plus pAmp, which conferred ampicillin resistance. Initial tests showed that the DH5 α ALE progenitor

(carrying the pCscBKA) was quickly outcompeted by the tester-GFP. To address this, after preadaptation passages, we mixed the tester-GFP and the ALE progenitor at a 95%:5% ratio, while evolved populations were mixed in a 1:1 ratio with the tester-GFP. From these mixtures, we inoculated 200 μ L into 20 mL of 2% sucrose minimal medium in triplicate using 125 mL Erlenmeyer flasks. During two growth passages (1:100 dilutions) at 37 °C and 100 rpm, GFP-expressing and non-expressing cells were measured at the stationary phase using an Attune NxT flow cytometer.

The cell proportions allowed us to calculate the fitness coefficient (S) using the formula: $S = \ln[ALE_f/GFP_f] - \ln[ALE_i/GFP_i]$ (Jacobus et al., 2024). Here, ALE_f/GFP_f are the final ALE populations and tester-GFP cell proportions, while ALE_i/GFP_i are the initial proportions of the competitors after the first passage. Cell counts at each point allowed us to normalize the fitness per cell doubling (S/doubling). The S/doubling obtained for the progenitor DH5 α -pCscBKA competing against tester-GFP was used to calculate the fitness burden from pCscBKA expression and to normalize the S/doubling of evolved ALE populations. The cost of GFP expression was also calculated in a competition between the DH5 α SS9::*cscBKA* strain and tester-GFP.

Evolved E2348/69 populations were also competed against the DH5 α tester-GFP. However, as evolved E2348/69 populations exhibited much faster growth rates, the competitions were set up with a 10:1 proportion of the tester-GFP to evolved E2348/69, while the E2348/69 progenitor was competed starting from a 1:1 mixture. Competitions and S/Doubling calculations were performed under the same conditions as described above.

Fitness evaluation for genetically modified strains started from a 1:1 strain/tester-GFP mixture and was recorded during four passages, with the first one serving as a baseline for S/doubling calculations. Normalization was obtained with the S/doubling calculated from a competition of SS9::cscBKA versus the tester-GFP. Statistical analyses were performed using GraphPad Prism 8.01 (GraphPad Software, La Jolla, California, United States).

We measured the fluorescence of the pGFP plasmid transformed into DH5 α , $\Delta pcnB$, and D-Spec backgrounds by recording 10,000 events with the Attune NxT flow cytometer after triplicate overnight cultures and IPTG induction. Fluorescence intensity values were derived from the median of flow cytometry histograms and averaged (Jacobus et al., 2024).

Procedures for sucrose and PHB quantification are detailed in the Supplementary Methods.

3. Results

3.1 ALE experiment

Many *E. coli* strains that naturally utilize sucrose carry the *cscB*, *cscK*, and *cscA* genes within an operon (Fig. 1A). The *cscB* gene encodes a sucrose/H⁺ symporter for the transport of the disaccharide into the cell, where CscA invertase hydrolyzes it into fructose and glucose. CscK kinase phosphorylates fructose to fructose-6-phosphate, a glycolytic intermediate (Bruschi et al., 2012; Olavarria et al., 2019). The *E. coli* strain E2348/69 (Iguchi et al., 2009) contains the *cscBKA* operon. We used *E. coli* in vivo recombination (Jacobus and Gross, 2015) to clone this operon into the pUC backbone, creating the

pCscBKA plasmid (Fig. 1A and Supplementary Fig. 6). The functionality of pCscBKA in DH5 α was confirmed by its capacity to support growth in minimal medium supplemented with sucrose (Supplementary Fig. 7).

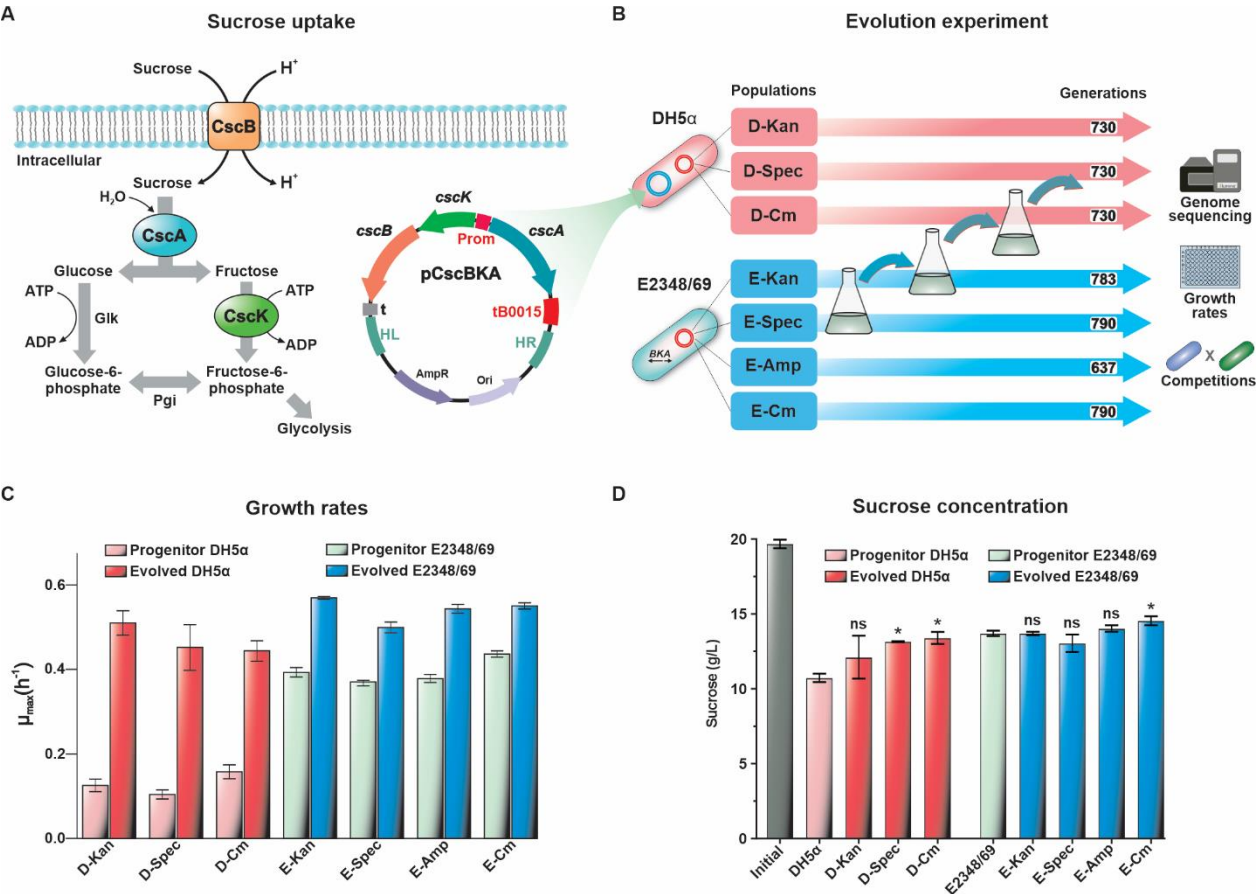


Fig. 1. ALE experiment in 2% sucrose. (A) Sucrose catabolic pathway in DH5 α with the pCscBKA plasmid: *cscB*, *cscA*, and *cscK* encode a sucrose transporter, invertase, and fructokinase, respectively. (B) ALE setup: DH5 α and E2348/69 populations, tagged with plasmid resistance to Kan, Spec, Amp, or Cm, evolved via daily 1:100 transfers on 2% sucrose for 637–790 generations. Final populations underwent DNA sequencing and growth/competition assays. (C) μ_{max} comparisons for progenitor versus evolved populations in 96-well microplates. (D) Sucrose consumption analysis with an initial concentration of 20 g/L (grey bar) and the remaining

sucrose concentration in the medium after 24 h of culture. The DH5 α progenitor was transformed with the pCscBKA plasmid. Statistical significance: ns (non-significant); (*) $p < 0.05$.

To enhance the growth of DH5 α containing pCscBKA on sucrose, we developed an ALE protocol propagating *E. coli* populations on this sugar for hundreds of generations over approximately 120 days (Fig. 2A). Three distinct DH5 α ALE populations were tagged with unique antibiotic resistance plasmids: D-Kan, D-Spec, and D-Cm. To explore alternative evolutionary pathways for improved sucrose utilization, we also established four E2348/69 ALE populations, each distinguished by a plasmid-encoded antibiotic resistance marker: E-Kan, E-Spec, E-Cm, and E-Amp. These antibiotic resistance genes and their corresponding antibiotics also provided protection against contamination during the ALE process.

Upon completion of the experiment, genomic DNA from all final populations was sequenced using the Illumina platform, and phenotypes were analyzed (Fig. 1C and D; Supplementary Fig. 7). The maximum specific growth rate (μ_{max}) of the evolved populations increased compared to their progenitors, with a more pronounced increase in the DH5 α populations (Fig. 1C). However, the DH5 α populations exhibited decreased total sucrose consumption relative to the parent, while the E2348/69 populations showed no significant change in final sucrose concentration compared to the progenitor. These findings suggest that enhanced sucrose consumption was not the primary driver of evolution.

3.2 Genome Sequencing of Evolved Populations

DNA sequencing of the evolved populations and their progenitors was conducted on the Illumina NovaSeq 6000 platform, identifying mutations such as SNPs, segmental amplifications/deletions, and small InDels (summarized in Fig. 2A and B, and Supplementary Table 5). Notably, Illumina read depth for the pCscBKA sequence in DH5 α D-Kan and D-Cm populations was similar to that of the *E. coli* chromosome (Fig. 2C), unlike typical high-copy pUC plasmids that replicate hundreds of copies per cell (Lee et al., 2006). This suggests potential integration of pCscBKA into the chromosome during ALE. Examination of Illumina read mappings revealed broken read pairs over the pCscBKA sequence in D-Kan and D-Cm, with unaligned ends linking to insertion sequences IS1 (D-Kan) and IS4 (D-Cm), indicating possible genomic integrations mediated by these mobile elements (Supplementary Fig. 8).

A different pattern was observed in the evolved D-Spec population, where the coverage depth over the pCscBKA and pSpec plasmids decreased by ~10-fold compared to pKan and pCm (Fig. 2C). Additionally, an 81 bp deletion in *pcnB* was detected, a mutation known to reduce the copy number of ColE1 plasmids (Lopilato et al., 1986), such as the pCscBKA.

Further functionally overlapping mutations in DH5 α populations involved RNA polymerase subunits RpoC, RpoB, and RpoZ, which are linked to a trade-off of increased fitness in minimal media but reduced growth in rich media (Conrad et al., 2010; Mohamed et al., 2019). The D-Spec population had a ~40.4 kb deletion affecting flagellar assembly gene clusters (Fig. 2D). In D-Kan, a SNP in *purB* caused an L115G change, potentially repairing a mutation known to lower DH5 α growth rates in minimal medium (Jung et al.,

2010). Additional mutations in DH5α populations were found in the genes *dnaA*, *phoE*, and *nudJ*.

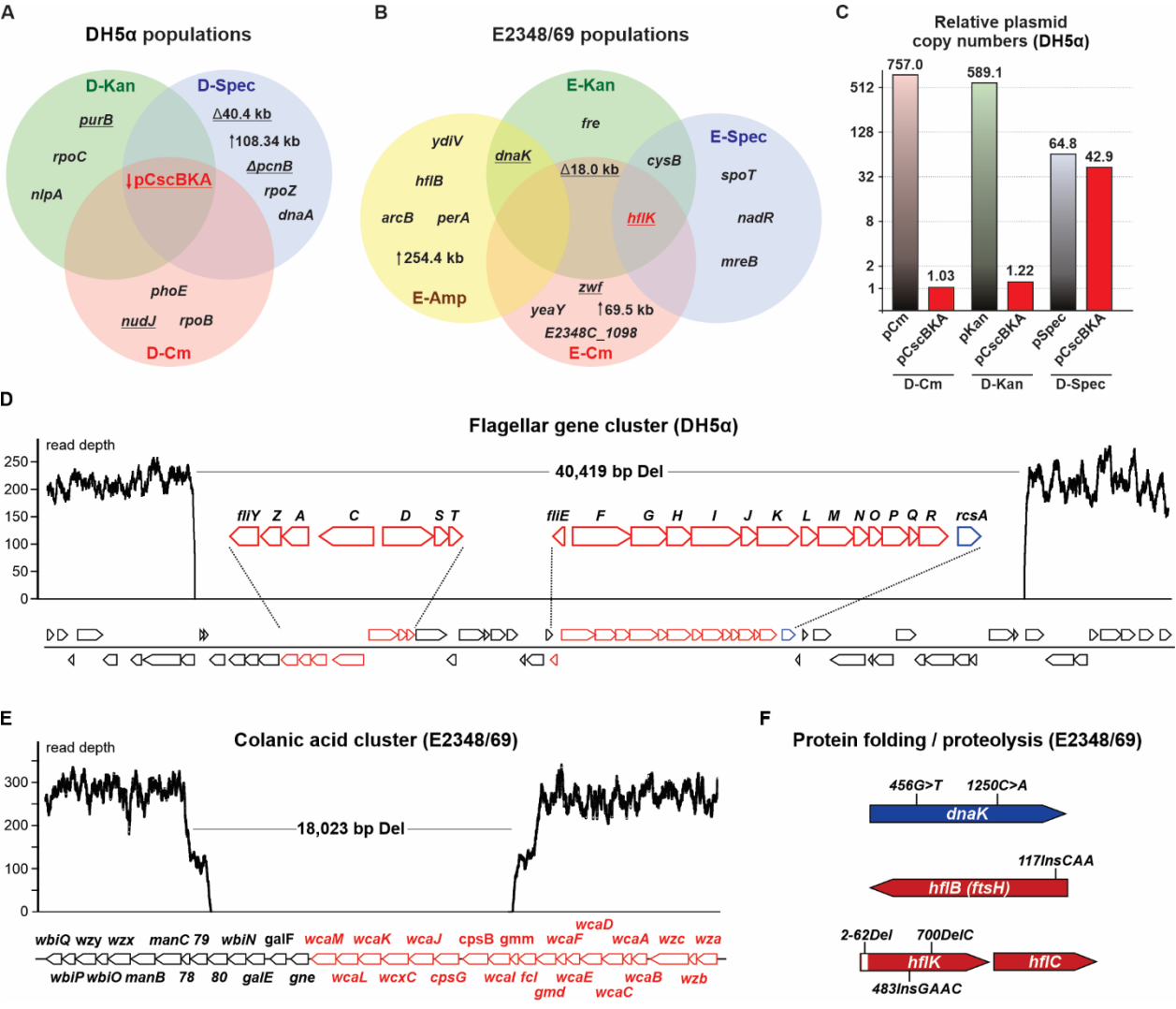


Fig. 2. Summary of mutations in ALE populations. (A, B) Venn diagrams illustrating common and strain-specific mutations; upward arrows indicate segmental amplifications, while downward arrows signify reductions in copy number. Underlined alleles have been selected for reverse engineering. The most frequent targets of the evolution are highlighted in red. (C) Illumina read depth analysis estimating plasmid copy numbers relative to the chromosome in final DH5α ALE populations. (D) Deletion of a ~40.4 kb segment in the DH5α D-Spec population, encompassing

flagellum assembly clusters (*fliA-R* and *fliYZ* genes, red). The *rcsA* gene (blue) is a transcription regulator for CA synthesis. (E) An 18,023 bp deletion in E-Kan and E-Cm, affecting part of the CA biosynthetic cluster (red). (F) Mutations in genes related to protein folding and proteolysis: *dnaK* (E-Kan, E-Amp), *hflB* (E-Amp), and *hflK* (E-Kan, E-Spec, E-Cm).

The evolved alleles in E2348/69 populations are summarized in Fig. 2B. An identical ~18 kb deletion in the colanic acid (CA) biosynthetic cluster was found in both E-Cm and E-Kan populations (Fig. 2E). Similar parallelism appeared in the *hfl* system, with different *hflK* alleles in E-Kan, E-Spec, and E-Cm, and a CAA insertion at position 117 in *hflB* (Fig. 2F). The *hflKC* operon encodes membrane proteins that regulate HflB (FtsH), a protease involved in the turnover of substrates such as the lambda phage CII transcriptional factor, which facilitates phage lysogeny, and the σ^{32} protein, which manages the heat shock response (Bandyopadhyay et al., 2011; Kihara et al., 1996; Tomoyasu et al., 1995).

Parallelism was observed for *dnaK* mutations (E-Kan and E-Amp, Fig. 2F), encoding a chaperone that serves as a protein stability quality control hub (Calloni et al., 2012). SNPs were identified in the *cysB* promoter region in E-Kan and E-Spec populations, which regulate cysteine synthesis (Lilic et al., 2003). Mutations in *zwf* and *nadR* were detected in E-Cm and E-Spec populations, respectively. The *zwf* gene encodes glucose-6-phosphate dehydrogenase, involved in the oxidative branch of the pentose phosphate pathway to generate NADPH (Zhao et al., 2004). NadR functions as a regulator of NAD⁺ biosynthesis and in the uptake of nicotinamide mononucleotide (NMN), the precursor of NAD⁺ (Raffaelli et al., 1999).

3.3 Development of a CRISPR/Cas9 strategy based on the EasyGuide principle

To better understand the adaptive role of key ALE mutations, we established a CRISPR/Cas9 platform for the rapid introduction of evolved alleles into parental strains. Genomic modifications via CRISPR/Cas9 were performed using a two-plasmid system with linear DNA donors. The temperature-sensitive plasmid CAS9BAC1P (Sigma-Aldrich), with optimal replication at 30 °C, expressed *S. pyogenes* Cas9 and the λ -Red recombinase. Specific strategies were developed for the production of gRNAs and donor DNAs.

We previously detailed the EasyGuide approach for the in vivo assembly of functional gRNAs into a vector separate from the Cas9 plasmid (Jacobus et al., 2022). This method uses homologous recombination of PCR products in *S. cerevisiae*, eliminating the need for in vitro ligation. We have now adapted the EasyGuide principle for *E. coli*, enabling in vivo gRNA assembly through three different cloning strategies with varying plasmid sets (Figs. 3A-C).

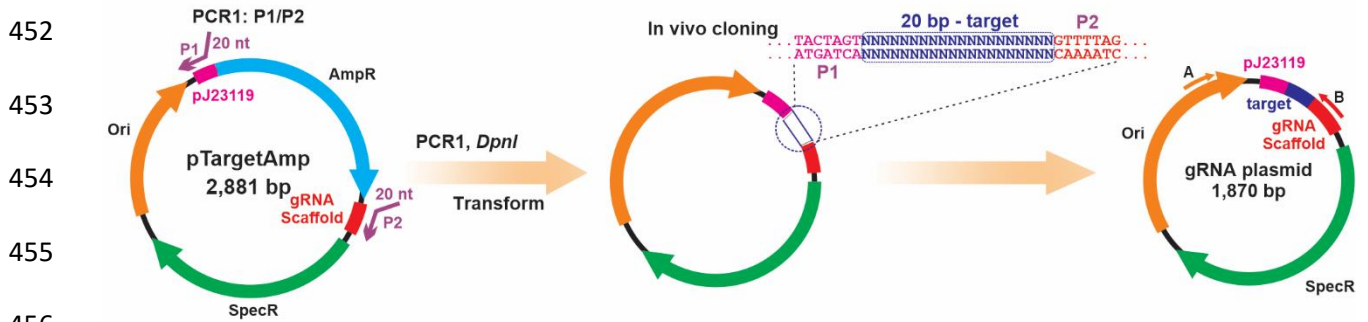
The first gRNA assembly strategy involves plasmid circularization (Fig. 3A). The pTargetAmp plasmid, derived from pTargetF (Jiang et al., 2015), maintains the ColE1 origin and spectinomycin resistance marker, with an ampicillin resistance marker added between the prJ23119 promoter and the gRNA scaffold for spacer sequence insertion. PCR amplifies the pTargetAmp backbone using primers P1 and P2, which have 20-nt 5' overhangs specifying the spacer. After producing an 1,890 bp product (Fig. 3D), DpnI is added to eliminate the template. The linear plasmid is then transformed by heat shock into *E. coli*, where in vivo recombination circularizes it, ligating the spacer. The resulting plasmid expresses a functional gRNA and confers spectinomycin resistance.

Colonies exhibiting ampicillin resistance represent false positives and were excluded from validation, which employs PCR with primers A and B to produce a 221 bp product in positive clones (Fig. 3E). Validated gRNA plasmids are purified and cotransformed with the linear donor into DH5 α cells containing the Cas9/ λ -Red recombinase plasmid. This system successfully cloned gRNAs for *galK* (Fig. 3E), *purB*, and the SS9 region (Table 2). However, for some gRNAs, cloning attempts failed, and in other cases, a high rate of false positives was observed.

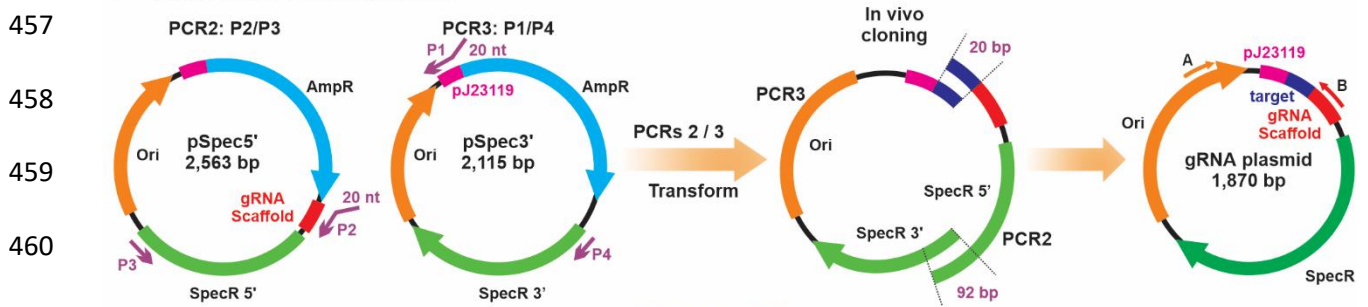
To overcome these problems, a second system was designed using a bimolecular complementation strategy (Fig. 2B). In this system, the gRNA plasmid is functional only through the recombination of two PCR products. The coding sequence for *aadA*, which confers spectinomycin resistance (SpecR) (Murphy, 1985), was divided between two plasmids (Fig. 3B). PCR of pSpec5' with primers P2/P3 (PCR2: 878 bp, Fig. 3D) amplifies the gRNA and *aadA* 5' region (SpecR 5'), while PCR of pSpec3' with primers P1/P4 (PCR3: 1,104 bp, Fig. 3D) amplifies the prJ23119 module and *aadA* 3' end (SpecR 3'). Cotransformation of these PCR products results in in vivo recombination, producing a functional gRNA plasmid that expresses full-length *aadA* (SpecR). This system avoids plasmid carryover, eliminates the need for DpnI treatment, and minimizes false positives. Using this approach, we have achieved 100% efficiency in gRNA cloning attempts.

To streamline the gRNA plasmid cloning procedure, a third strategy based on oligo PCR cloning (the EasyOligo approach, Fig. 3C and D) was developed. This enhances the bimolecular complementation method by enabling the mass preparation and storage of the PCR products (SpecR 5'/gRNA and SpecR 3'/prJ23119) for use as plasmid backbones in multiple gRNA assemblies utilizing PCR products generated from specific oligos.

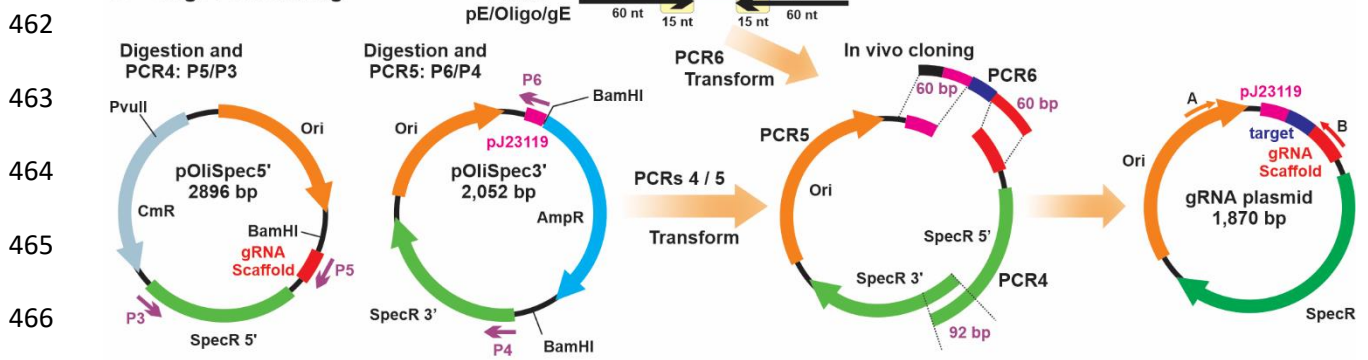
A Plasmid circularization



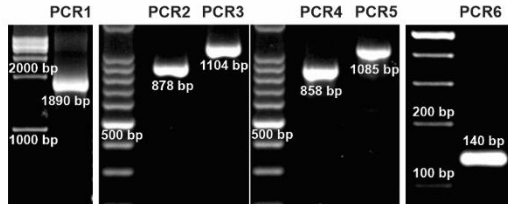
B Bimolecular complementation



C Oligo PCR cloning



D



E

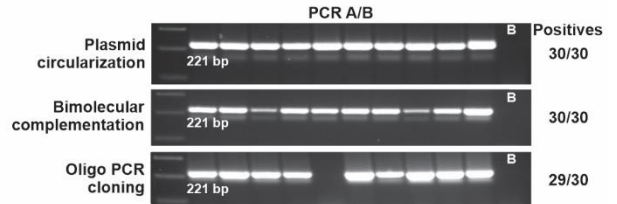


Fig. 3. Different strategies for producing gRNA plasmids. (A) Plasmid circularization: using pTargetAmp as a template, PCR is performed with primers P1 and P2 (PCR1, 1,890 bp), specifying the 20-nt spacer. After DpnI treatment, the product is transformed and ligated in vivo. (B) Bimolecular complementation: PCR fragments from pSpec5' (primers P2/P3) and pSpec3' (primers P1/P4) are cotransformed into *E. coli*, assembling the gRNA plasmid in vivo. (C) Oligo

PCR cloning: plasmids pOliSpec5' and pOliSpec3' are digested and purified, then amplified by PCR using primers P5/P3 and P4/P6. The gRNA region is generated by amplifying the oligo spacer with primers pE and gE. The 140 bp product is mixed with backbone fragments and transformed into *E. coli* for in vivo assembly. (D) Electrophoresis in 1%, 1.5%, and 3% agarose gels showing PCR products for gRNA plasmid cloning. Sizes are indicated (1 kb marker for PCR1, 100 bp ladder for PCRs 2–6). (E) 2% agarose gel electrophoresis showing 221 bp PCR products confirming *galK* gRNA cloning using the three strategies. Triplicate tests were conducted; only one replicate is shown. Positive cloning rates are indicated; a 100 bp ladder was used as a marker.

For this purpose, two new plasmids, pOliSpec5' and pOliSpec3', were constructed with restriction sites to remove the plasmid template backbone prior to PCR amplification (Fig. 3C). This digestion prevents improper recombination events during in vivo cloning (Barreto, personal observations). The pOliSpec5' plasmid is digested with PvuII and BamHI, while the pOliSpec3' plasmid is digested with BamHI. After purification, these templates are PCR-amplified using primers P3/P5 (pOliSpec5', PCR4) and P4/P6 (pOliSpec3', PCR5) (Fig. 3D). Primers P5 and P6 do not include 5'-overhangs; instead, the 20-nt spacer is embedded within a central oligonucleotide amplified with primers pE (for recombining with prJ23119) and gE (overlapping with the gRNA scaffold). The 140 bp dsDNA produced (Fig. 3C and D, PCR6) is transformed with PCR products 4 and 5. Proper recombination of all three fragments assembles a SpecR gRNA plasmid.

3.4 Donor DNA assembly in yeast

Previous CRISPR/Cas9 protocols for *E. coli* have recommended using DNA donors with flanks of 100–600 bp for efficient homologous recombination (Bassalo et al., 2016;

Jiang et al., 2013). We developed a strategy to preassemble donors with large homologous regions using yeast in vivo cloning (Fig. 4). The yeast origin of replication (2 μ) and an antibiotic resistance marker were amplified by PCR from the pEasyG3 plasmid series (Fig. 4A-C) (Jacobus et al., 2022). Homologous flanking regions are PCR-amplified from the chromosomal target site (Fig. 4A), along with the gene or DNA sequence intended for integration. Oligonucleotides used for PCR create 20–40 bp overlaps between products and connections to the yeast plasmid backbone (Fig. 4A). Deletion borders or SNPs are specified between the homologous arms (Figs. 4B and C). PCR products and plasmid fragments are cotransformed into *S. cerevisiae* without DNA cleanup. After selection, positive colonies are verified by PCR. Donor plasmid enrichment occurs through yeast plasmid preparation. PCR amplification from this preparation yields a donor DNA fragment ready for the CRISPR/Cas9 experiment in *E. coli* via heat shock transformation. An example of this approach is the cloning of a GFP cassette into the *galK* locus of *E. coli* DH5 α (Fig. 4D-E).

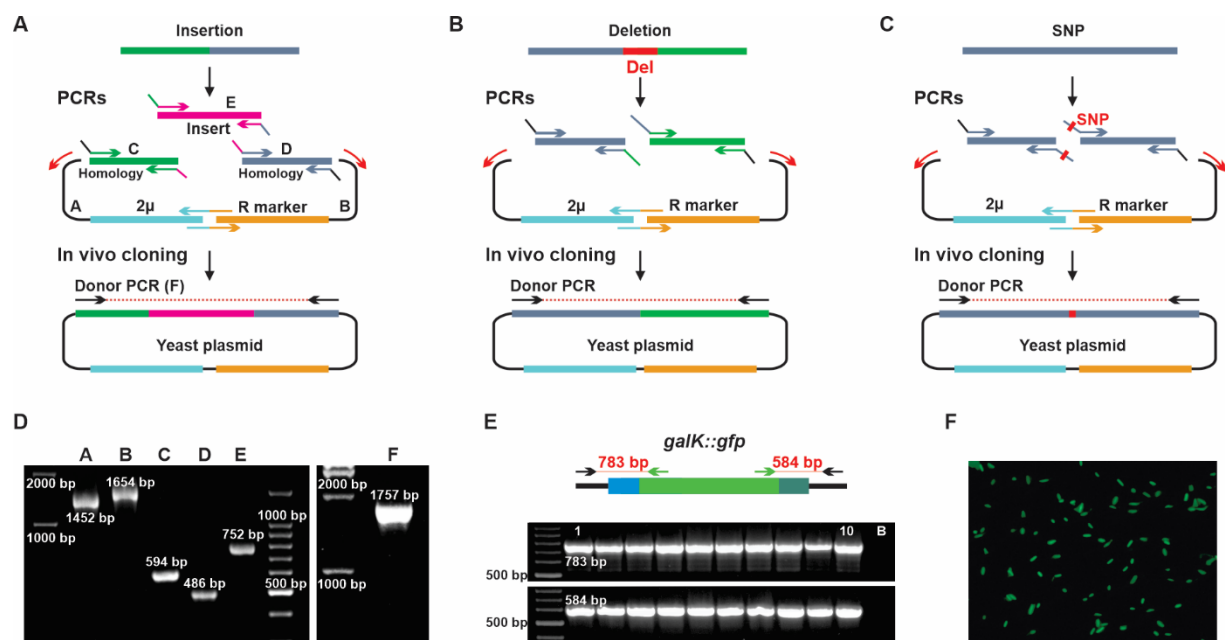


Fig. 4. Assembly of donor sequences by yeast homologous recombination. The yeast plasmid origin of replication (2 μ) and antibiotic resistance marker (R marker) are amplified from pEasyG3 plasmids via two PCRs (A and B). (A-C) Parts are recombined in vivo with PCR products to create plasmids with donors. (A) Donor DNA precloning for the *E. coli* genome insertion. (B) Donor DNA assembly for chromosomal deletions, positioning the deletion border between homologous arms. (C) Donors for reverse engineering SNPs, with SNPs specified at the 5' ends of oligonucleotides used for PCR. Assembled donors are PCR-amplified and transformed into *E. coli* for CRISPR/Cas9 integration. (D) A 1.5% agarose gel showing PCRs of parts (A-E) for inserting GFP into the *galK* locus; donor PCR yields 1,767 bp product (PCR F). 1 kb and 100 bp ladders were used. (E) PCR confirmation of *galK::gfp* integration into *E. coli* DH5 α . (F) Fluorescence microscopy showing enhanced GFP fluorescence observed after *galK::gfp* integration to generate a tester-GFP strain for competition assays.

3.5 The EasyOligo strategy for producing donor DNAs

To eliminate precloning steps, a PCR-based strategy using oligonucleotides, akin to the EasyOligo approach for gRNAs, was developed to directly produce donor DNAs. This method is demonstrated by generating a 32 bp deletion in *galK* (Fig. 5A). The deletion is specified within a central 50-nt oligo, amplified by PCR using 60-nt forward and reverse oligos providing chromosomal recombination flanks. The resulting 140 bp dsDNA and the gRNA plasmid are directly transformed into *E. coli* via heat shock for CRISPR/Cas9-mediated deletion (Fig. 5A).

In another example, a 60-mer central oligo was used to reverse-engineer the T\G transversion in the *zwf* gene from the evolved E-Cm population into DH5 α (Fig. 5B). A PstI restriction site was incorporated into the oligo to identify correctly edited colonies, and

the PAM consensus was altered to prevent Cas9 recognition. These edits did not change the encoded amino acids (Fig. 5C). After transformation, positive colonies were screened using PstI restriction analysis, and DNA modifications were confirmed by Sanger sequencing (Fig. 5C).

Interestingly, attempts to transform oligo-derived PCR products as donors via electroporation failed to produce colonies or resulted in aberrant genomic modifications. Similarly, in vivo cloning of gRNAs was consistently problematic. A comparison of electroporation and heat shock transformations for the deletions of *galK*, *flhCD*, and *hflK* confirmed the drawbacks of electroporation with oligo-derived donors (Fig. S3A-H). These results align with previous reports indicating that in vivo cloning in *E. coli* was more successful with heat shock transformation than with electroporation (Bubeck et al., 1993; Kostylev et al., 2015).

For the *hflK* deletion, a third donor DNA production method was employed involving PCR generation of a repair template with asymmetric arms (Jacobus et al., 2022; Jacobus et al., 2024; Shukal et al., 2022). A minimized 25 bp left flank was utilized. Donor DNA was also directly PCR-amplified from the *E. coli* genome (Table 2).

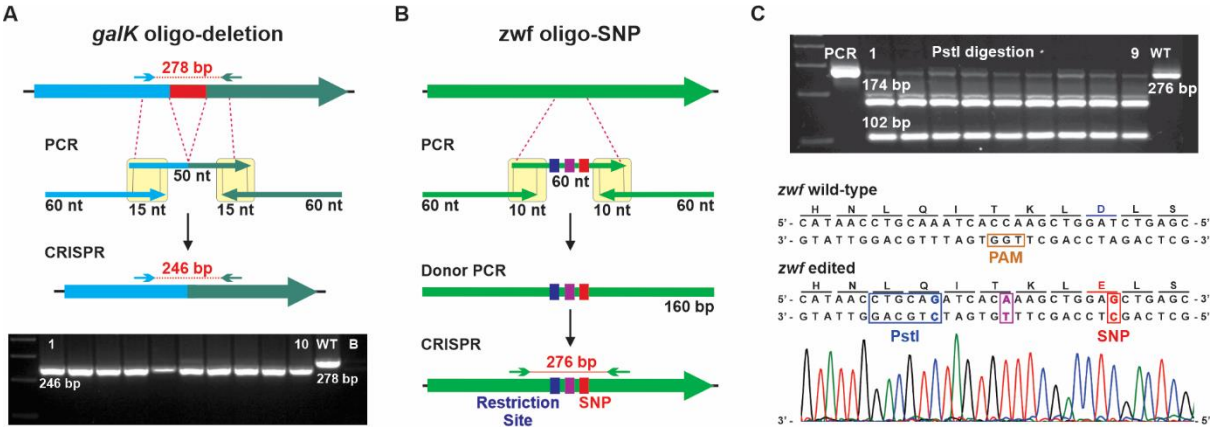


Fig. 5. Production of donors using oligonucleotides. (A) Generation of a donor to produce a 32 bp deletion in *galK*. Deletion borders were specified within a 50-nt central oligo, which was used for amplifying a 140 bp PCR product with forward and reverse oligos containing 15-nt priming sites. 2% agarose gel electrophoresis shows a 246 bp fragment for the deletion and a 278 bp band for WT. (B) The editing of SNPs in the *zwf* gene. The *zwf* T/G mutation (red box), the obliteration of the PAM sequence (purple), and the creation of a PstI restriction site (blue) are specified within the 60-nt central oligo, which serves as a template for the PCR amplification of a 160 bp donor, directly transformed into *E. coli* for CRISPR/Cas9 genome editing. (C) Editing confirmed by PstI digestion and Sanger sequencing; cleaved fragments of 174 bp and 102 bp indicate successful editing, while a 276 bp fragment corresponds to WT.

3.6 Reverse engineering of ALE mutations

We utilized our CRISPR EasyGuide approach to reverse engineer a set of alleles from sucrose-evolved populations into *E. coli* DH5 α (Table 2). To assess the impact of individual mutations on sucrose growth fitness through competition assays (described below), a testing strain was constructed by inserting the *cscBKA* operon into the DH5 α SS9 intergenic region (Bassalo et al., 2016). Despite our efforts, we could not transform the Cas9/ λ -Red recombinase plasmid into *E. coli* E2348/69. Therefore, *E. coli* DH5 α SS9::*cscBKA* was also used to integrate selected mutations from E2348/69 populations. Genes mutated in different ALE populations, indicative of parallel evolution, are reliable markers of adaptive processes (Menegon et al., 2022). Thus, our reverse engineering campaign prioritized alleles from parallel mutated targets (Figure 2A and B) and mutations potentially affecting growth fitness or sucrose consumption under ALE conditions (Table 2).

Following the CRISPR/Cas9 procedure, we screened the transformants to assess positive editing rates through colony PCR. Overall, the efficiency rates achieved were high, with numerous targeted genomic loci being edited with 100% efficiency (Table 2). This high level of precision underscores the effectiveness of our CRISPR/Cas9 approach. However, it is important to note that some loci exhibited low editing efficiency. This discrepancy in efficiency levels may be attributed to the variable performance of specific gRNAs (Bassalo et al., 2016; Shukal et al., 2022). For instance, the editing of the *lacZ* locus demonstrated lower success rates. These observations highlight the importance of gRNA selection in obtaining optimal editing results. Detailed outcomes of each individual CRISPR/Cas9 experiment are documented in the Supplementary File 1.

Table 2

CRISPR/Cas9-mediated genetic modifications through the EasyGuide approach.

Genetic modification	gRNA cloning	Donor assembly	Efficiency
<i>SS9::cscBKA</i>	Plasmid circularization	In vivo cloning in <i>E. coli</i>	10/10
<i>galK::cscBKA</i>	Plasmid circularization	In vivo cloning <i>S. cerevisiae</i>	10/10
<i>cscA::sucP</i>	Bimolecular complementation	In vivo cloning <i>S. cerevisiae</i>	4/4
<i>cscA::pJ23119-sucP</i>	Bimolecular complementation	In vivo cloning <i>S. cerevisiae</i>	4/4
<i>lacZ::sucP</i>	Bimolecular complementation	In vivo cloning <i>S. cerevisiae</i>	3/14
<i>galK::gfp</i>	Plasmid circularization	In vivo cloning <i>S. cerevisiae</i>	10/10
<i>lacZ::gfp</i>	Bimolecular complementation	In vivo cloning <i>S. cerevisiae</i>	1/10
<i>lacZ::phaCAB</i>	Bimolecular complementation	In vivo cloning <i>S. cerevisiae</i>	6/14
<i>galK::ampR</i>	Plasmid circularization	In vivo cloning <i>S. cerevisiae</i>	5/5
<i>lacZ::ampR</i>	Bimolecular complementation	In vivo cloning <i>S. cerevisiae</i>	3/5
$\Delta 40.4$ kb	Bimolecular complementation	PCR Amplification from D-Spec DNA	4/4
$\Delta fliA$ (205 bp)	Bimolecular complementation	In vivo cloning <i>S. cerevisiae</i>	10/10
$\Delta fliA$ (205 bp)	Bimolecular complementation	Oligo PCR amplification	9/10

<i>ΔflhDC</i>	Oligo PCR cloning	Oligo PCR amplification	10/10
<i>ΔpcnB</i> (81 bp)	Bimolecular complementation	PCR amplification from the D-Spec DNA	1/10
<i>ΔhflK</i>	Oligo PCR cloning	PCR of a donor with asymmetric arms	10/10
<i>ΔCA</i> (12 kb)	Oligo PCR cloning	Oligo PCR amplification	9/10
<i>ΔgalK</i> (32 bp)	Oligo PCR cloning	Oligo PCR amplification	10/10
<i>nudJ</i> (G151V)	Bimolecular complementation	In vivo cloning <i>S. cerevisiae</i>	2/2
<i>purB</i> (L115G)	Plasmid circularization	In vivo cloning <i>S. cerevisiae</i>	5/5
<i>dnaK</i> (T417N)	Oligo PCR cloning	Oligo PCR amplification	3/10
<i>zwf</i> (D405E)	Oligo PCR cloning	Oligo PCR amplification	9/9

3.7 Fitness evaluation of evolved populations and engineered strains

We assessed the fitness of the ALE populations and each reverse-engineered strain using a growth competition assay against a GFP-expressing reference strain (Fig. 6A). Flow cytometry was employed to measure the ratio of fluorescent to non-fluorescent cells at the start and end points of propagations, enabling the calculation of a selection coefficient normalized per cell doubling (S/doubling). A positive S/doubling demonstrates that the evolved population or engineered mutant has an enhanced growth capacity (i.e., fitness) compared to the parental strain, while a negative S/doubling indicates that the engineered allele negatively affects growth fitness.

Competitions with the evolved ALE populations revealed increased fitness compared to their evolutionary ancestors (Fig. 6B). The DH5α populations demonstrated a greater fitness improvement than the E2348/69 populations, which aligns with the higher μ_{max} observed in microplate growth assays (Fig. 1B). This fitness boost in DH5α is likely due to the chromosomal integration of the pCscBKA plasmid in the D-Kan and D-Cm populations, along with decreased pCscBKA copy number in D-Spec (Fig. 2C).

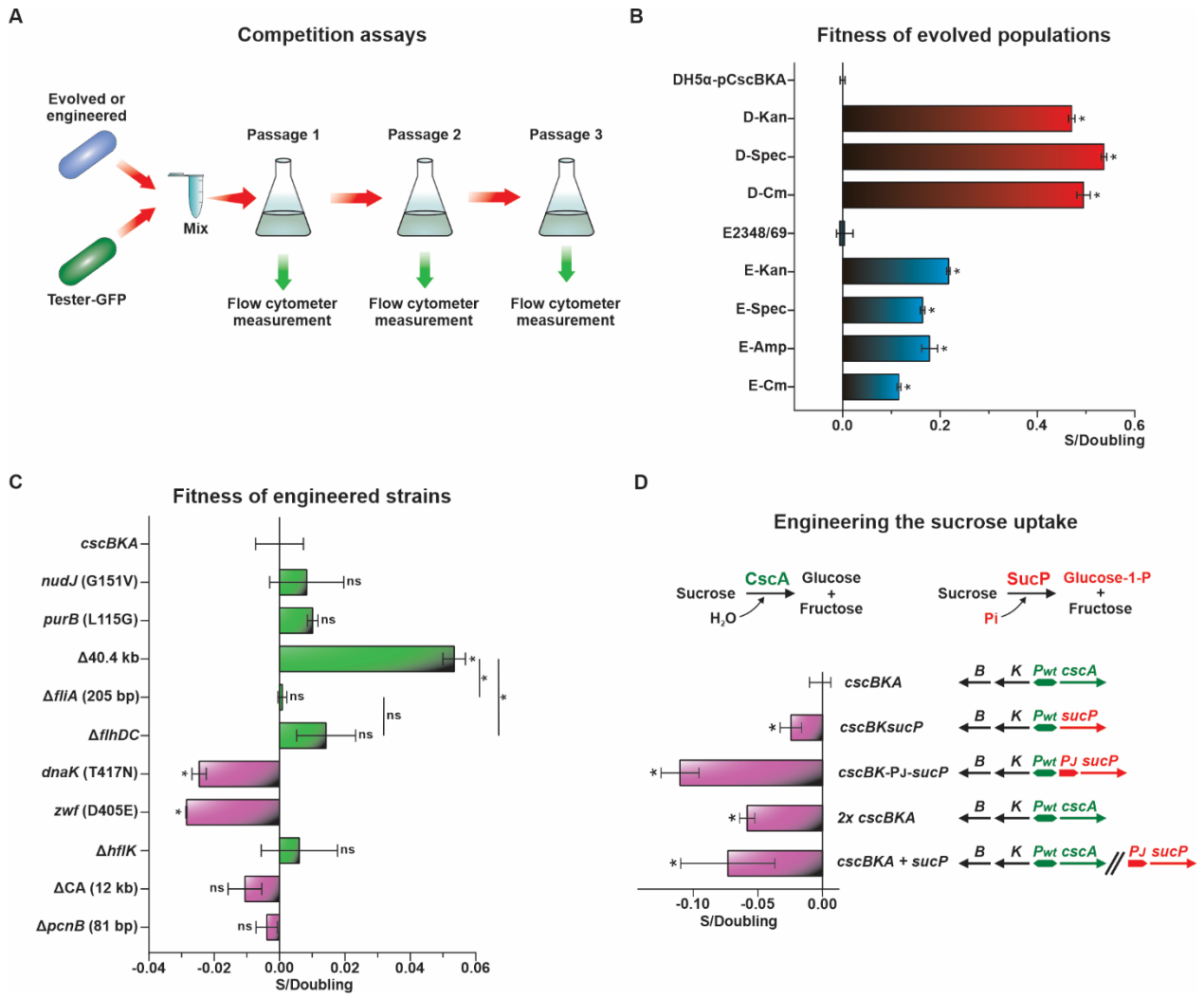


Fig. 6. Fitness estimations for evolved populations and engineered strains. (A) Competition experiment setup: strains and tester-GFP were mixed in specific proportions in minimal medium containing 2% sucrose and propagated in triplicate. Flow cytometry measurements of GFP versus non-GFP cells at the initial point and over one to three passages enabled the calculation of selection coefficients per doubling (S/doubling). (B) Fitness for evolved populations, normalized to the progenitors (*E. coli* DH5α-pCscBKA or E2348/69), was calculated following one passage. (C) Fitness for engineered strains was assessed over three passages. (D) Fitness of strains combining the *cscBKA* and *sucP* genes. *Pwt* (green) represents the E2348/69 *cscBKA* operon promoter; *Pj* (red) is the synthetic prJ23119 promoter. S/doubling values were normalized to the

value obtained for the competition of the tester-GFP versus *E. coli* DH5α SS9::cscBKA. Statistical significance was compared to *E. coli* DH5α SS9::cscBKA and among the *fliA*, *flhDC*, and Δ40.4 kb strains. ns: non-significant; (*) significant ($p < 0.05$).

Among the reverse-engineered ALE mutations, only the 40,419 bp deletion from the D-Spec population demonstrated a significant positive fitness effect (Fig. 6C). This deleted region encompasses the *fli* operons related to flagellum assembly (Juhas et al., 2014). To explore the fitness increase associated with impaired flagellum assembly, we deleted the *fliA* gene, a late-stage flagellum regulator, and the *flhDC* operon, a primary flagellum regulator (Barembuch and Hengge, 2007; Chilcott and Hughes, 2000). Neither deletion resulted in a fitness increase comparable to that of the 40,419 bp deletion. This gene cluster also includes the *rcaA* ORF (Fig. 2D), a positive regulator of CA biosynthesis (Ebel and Trempey, 1999). A partial deletion of the CA biosynthetic cluster was observed in E2348/69 E-Kan and E-Cm populations, suggesting adaptive parallelism. Testing the CA cluster deletion in DH5α revealed a slightly negative fitness effect (Fig. 6C).

Competing reverse-engineered alleles related to *purB*, *nudJ*, *hflK*, *zwf*, and *dnaK* with the tester-GFP exhibited null or negative fitness effects (Fig. 6C). Similarly, an 81 bp deletion in *pcnB* in the D-Spec population slightly hindered growth. However, the Δ*pcnB* mutation positively impacted fitness in DH5α strains expressing the pCscBKA plasmid (see Fig. 7).

3.8 Engineering operons for sucrose catabolism

We leveraged the CRISPR EasyGuide system to engineer variations in the *cscBKA* operon (Fig. 6D), including replacing *cscA* (invertase) with the *Bifidobacterium adolescentis* *sucP* ORF, which encodes sucrose phosphorylase (van den Broek et al., 2004). This phosphorolysis pathway spares one ATP compared to sucrose hydrolysis, where glucokinase (Glk) consumes ATP to phosphorylate glucose (Fig. 1A) (Olavarria et al., 2019). We tested four engineered strains to explore the adaptive effects of encoding *sucP* with the *cscBKA* operon. Competition assays in minimal medium with 2% sucrose (Fig. 7D) showed that expressing *sucP* combined with *cscBK* within the operon or in parallel with *cscBKA* reduced fitness compared to *cscBKA* alone. Additionally, expressing two copies of *cscBKA* also decreased growth fitness on sucrose (Fig. 6D). The results show that under our tested conditions a single copy of the *cscBKA* operon is more effective, yielding greater fitness than its duplication or combinations with *sucP*.

3.9 Reduced fitness associated with *cscBKA* expression from plasmids

We observed a significant reduction in pCscBKA copy numbers in the evolved DH5 α populations. Potential chromosomal integration of pCscBKA was noted in populations D-Kan and D-Cm, while both pCscBKA and pSpec plasmids in D-Spec population had reduced copy numbers (Fig. 2C). The D-Spec population also contains an 81 bp deletion in *pcnB*, which encodes poly(A) polymerase I (PAP I) and was identified in a screening for mutants that lower ColE1 plasmid copy numbers, hence its name (plasmid copy number B) (Lopilato et al., 1986). To investigate the role of *pcnB* in maintaining ColE1 plasmids in DH5 α , we used a plasmid-based GFP reporter (pGFP, a ColE1-type plasmid) for estimating relative differences in plasmid copy numbers. Fluorescence intensity from

GFP-expressing plasmids, measured by flow cytometry, correlates with plasmid copy numbers (Joshi et al., 2022). We isolated a colony from the D-Spec population and propagated it to eliminate the pCscBKA and pSpec plasmids. The curated D-Spec clone, the reverse-engineered strain $\Delta pcnB$, and DH5 α were transformed with pGFP, and fluorescence intensity was measured in triplicate at the stationary growth phase. The D-Spec clone and $\Delta pcnB$ mutant showed a marked decrease in GFP expression compared to DH5 α , consistent with lower pGFP copy numbers (Fig. 7A).

Through competition assays with the tester-GFP, we estimated the fitness associated with plasmid-based (pCscBKA) versus chromosomal expression of *cscBKA* (Fig. 7B). A significant decrease in fitness was observed in DH5 α carrying pCscBKA compared to the tester-GFP with genomic integration of *cscBKA*. Our data suggest a fitness drop of approximately 35% per doubling for *cscBKA* expressed from a high-copy plasmid compared to its chromosomal counterpart. This burden was alleviated by the $\Delta pcnB$ mutation alone or within the D-Spec clone background. Additional mutations, such as the $\Delta 40.4$ kb deletion in the D-Spec background, may contribute to the positive fitness observed in the evolved clone. These findings confirm that the integration or reduction in copy number of the *cscBKA* operon was a primary driver of fitness gains in the evolved DH5 α populations.

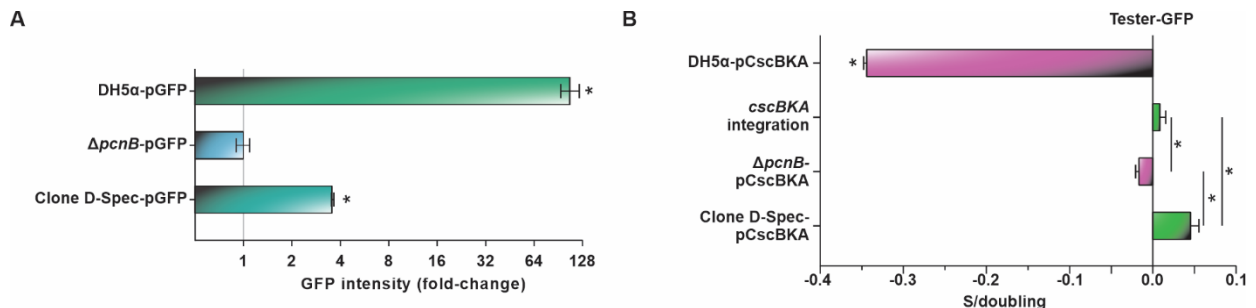


Fig. 7. Impact of mutations on plasmid copy numbers and pCscBKA fitness. (A) GFP expression as a proxy for estimating relative differences in plasmid copy numbers. The high-copy pGFP plasmid was transformed into *E. coli* backgrounds: DH5 α (reference), DH5 α Δ *pcnB*, and the evolved clone D-Spec. GFP levels were measured by flow cytometry, with experiments conducted in triplicate. (B) The fitness impact of pCscBKA copy numbers was assessed through competitions with the tester-GFP. The pCscBKA plasmid was transformed into the *E. coli* backgrounds: DH5 α , DH5 α Δ *pcnB*, and the evolved clone D-Spec. The *cscBKA* integration was into the DH5 α SS9 *locus*. (*) Indicates significant differences ($p < 0.05$) between strains.

3.10 Integration of the *phaCAB* operon for PHB production

To demonstrate the utility of the EasyGuide approach for the metabolic engineering of *E. coli*, we assembled the *phaCAB* cassette and performed its chromosomal integration into the *E. coli* SS9::*cscBKA* strain. The *phaCAB* operon from *Cupriavidus necator* supports the synthesis of PHB, a polyester-type polymer providing a biobased alternative to oil-derived plastics (Hiroe et al., 2012; Sohn et al., 2020). Producing PHB in *E. coli* strains engineered for sucrose consumption enables the use of cost-effective and renewable feedstocks for industrial applications (Sohn et al., 2020). We utilized yeast in vivo recombination to assemble the *C. necator phaCAB* cassette, transforming a PCR product containing the structural genes and flanking regions into *E. coli* DH5 α , targeting the *lacZ* locus for integration. A strong prJ23119 promoter was added to drive *phaCAB* expression (Fig. 8A). PHB production from sucrose in the transformed strain was confirmed through GC-MS analysis, yielding approximately 31% of dry cell mass, close to the 36% produced by *Burkholderia glumae*, a natural PHB producer (Fig. 8B and C).

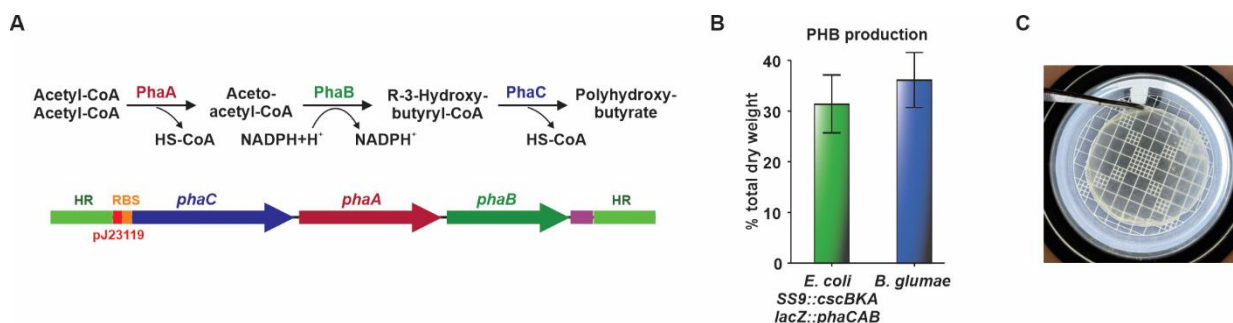


Fig. 8. PHB production from sucrose in *E. coli* DH5α with integrated *cscBKA* and *phaCAB* operons. (A) Metabolic pathway for PHB production and the *phaCAB* construct for insertion, assembled in yeast. After cloning, the donor was amplified for insertion into the *lacZ* locus, with expression driven by a strong constitutive promoter (prJ23119). (B) PHB production as a percentage of dry cell mass in engineered DH5α and *B. glumae*. (C) PHB film produced by engineered *E. coli* DH5α.

4. Discussion

We conducted an ALE experiment to enhance the growth of *E. coli* on sucrose, resulting in the identification of various mutations, including deletions, insertions, and SNPs across seven populations. To rigorously validate these ALE alleles, we used CRISPR/Cas9-mediated reverse engineering alongside innovative strategies for donor DNA production and the adaptation of the EasyGuide system for efficient gRNA recombination cloning in *E. coli* (Jacobus et al., 2022). Our in vivo methods provide a cost-effective and practical alternative to traditional in vitro gRNA/donor preparation techniques, such as restriction-ligation, CPEC, or Golden Gate assembly (Table 1).

The in vivo recombination pathway in *E. coli* was described in the early 1990s for cloning PCR products (Bubeck et al., 1993; Jones and Howard, 1991; Oliner et al., 1993). Despite numerous applications documented since then (Jacobus and Gross, 2015;

Nozaki and Niki, 2019; Watson and Garcia-Nafria, 2019), its potential as a simplified cloning tool has been largely overlooked. In our study, we capitalized on this pathway using three distinct gRNA assembly strategies: (i) Plasmid circularization, which involves PCR amplification of pTargetAmp with primers featuring 5' overlapping spacers for in vivo plasmid closure. While practical, this method is susceptible to false positives due to template plasmid carryover during transformation (Jacobus and Gross, 2015). (ii) Bimolecular complementation effectively reduces false positives. Cotransforming two PCR products from pSpec5' and pSpec3' enables the formation of spectinomycin-resistant colonies if precise recombination assembles both the gRNA and the complete *aadA* gene. This minimizes false positives, as shown in the in vivo assembly of split yeast episomes (Jacobus et al., 2022; Kuijpers et al., 2013). (iii) The EasyOligo approach simplifies the process by using PCR amplification of a single 50-mer oligonucleotide containing the gRNA spacer; the resulting 140 bp dsDNA product is easily cotransformed with pre-prepared pOliSpec5' and pOliSpec3' fragments to generate SpecR colonies.

The production of editing templates for CRISPR applications in *E. coli* has traditionally been limited by the need to preassemble large HA (100–600 bp) for high editing efficiencies (Bassalo et al., 2016; Jiang et al., 2015). To overcome this challenge, we repurposed our pEasyG3 plasmids (Jacobus et al., 2022) to enable in vivo yeast fusion of PCR fragments flanking deletion or editing sites and to assemble insertion cassettes. The time invested in yeast transformation, colony screening, and DNA preparation is compensated by exceptionally high assembly rates, often reaching up to 100%.

As an alternative to preassembling donors, synthetic oligonucleotides (40–120-mer) serve as effective donor DNAs, demonstrating efficient recombineering rates even for

large genomic deletions, with double-stranded oligos outperforming single-stranded ones (Reisch and Prather, 2015; Zerbini et al., 2017). We also used these oligos as templates for PCR production of dsDNA donors with up to 70 bp HA, embedding the editing sequence or deletion borders within the Mid-oligo. This practical approach is streamlined by a thermal-shock transformation protocol, which eliminates the need for PCR product desalting/purification, a required step for electroporation in typical *E. coli* CRISPR applications. This simplified protocol allows direct transfer of DNA from PCR tubes into chemically competent cells.

By employing the EasyGuide approach with yeast-assembled and oligo-derived donors, along with repair templates that were PCR-amplified from ALE populations, we consistently achieved high editing rates, often up to 100% positive outcomes. However, as noted in various studies (Bassalo et al., 2016; Shukal et al., 2022), editing efficiencies varied among different gRNAs, with targets such as *lacZ*, *pcnB*, and *dnaK* showing positive rates as low as 10%. Overall, EasyGuide proved to be a robust method for reverse engineering multiple mutations identified in ALE populations propagated on 2% sucrose. This approach allowed us to assess numerous alleles through competition assays, revealing positive fitness effects for SS9::*cscBKA*, $\Delta pcnB$, and the ~40.4 kb deletion that includes the flagellar assembly cluster. A significant fitness decrease of approximately 35% per cell doubling was observed for plasmid-based expression of *cscBKA* compared to SS9::*cscBKA*, indicating that the evolution of D-Kan and D-Cm populations was primarily driven by the chromosomal integration of pCscBKA. Similarly, the fitness improvement in the D-Spec population likely stemmed from reduced pCscBKA copy numbers due to $\Delta pcnB$. In *E. coli* K-12, decreased growth rates linked to *cscBKA* plasmids

were also alleviated by the operon integration into the genome (Bruschi et al., 2012; Carruthers et al., 2020).

A reduced dosage of CscBKA may alleviate its metabolic burden, while the deletion of a 40,419 bp segment in D-Spec, which includes flagellar assembly regions 3a and 3b (Juhas et al., 2014), likely impacts the cell's energy dynamics. Flagellar assembly in *E. coli* requires the expression of over 60 genes (Barembuch and Hengge, 2007; Chilcott and Hughes, 2000), consuming approximately 5.0% of the cell's energy, with an additional 5.2% dedicated to flagellar rotor operation (Schavemaker and Lynch, 2022). Thus, the observed ~5% fitness gain per doubling from deleting flagellar clusters is expected; however, this improvement exceeds that achieved by individually deleting the key flagella regulators *flhDC* and *fliA* (Barembuch and Hengge, 2007; Chilcott and Hughes, 2000). Other genes within the 40,419 bp segment may also contribute to adaptive effects, such as *rcaA*, a transcription factor involved in CA biosynthesis (Ebel and Trempy, 1999). However, partial deletion of the 12.0 kb CA cluster and additional mutations in *E. coli* DH5 α and E2348/69 have shown negative or neutral impacts on fitness in DH5 α under 2% sucrose conditions. The lack of notable adaptive effects from reverse-engineered ALE alleles may stem from their specific genetic backgrounds or epistatic interactions (Barrick and Lenski, 2013; Menegon et al., 2022). For instance, recreating the loss-of-function mutation in *hflK* from E2348/69 in DH5 α yielded only a negligible fitness gain. Mutations in *hflK* and *hflB*, as observed in E2348/69, have been associated with promoting λ prophage dormancy (Bandyopadhyay et al., 2011), likely helping to mitigate genome instability in the prophage-carrying strain E2348/69. Since the DH5 α genome lacks λ prophages, the *hflK* deletion may not confer any adaptive benefits for this strain.

ALE with DH5 α populations identified mutations in *rpoB*, *rpoC*, and *rpoZ*, which have been linked to enhanced growth rates in minimal media utilizing glucose or sucrose (Conrad et al., 2010; Mohamed et al., 2019). These findings, along with other mutations, suggest that our ALE strategy favored systemic changes that improved overall growth fitness rather than specifically optimizing sucrose utilization pathways. Nonetheless, our study provides a basis for advancing sucrose-utilizing platforms. The competitive fitness observed from sucrose hydrolysis (via CscA) compared to phosphorolysis (via SucP) highlights the need to evaluate various sucrose uptake mechanisms and explore optimized combinations. Furthermore, robust stress tolerance is essential for industrial applications and can be enhanced through ALE (Dragosits and Mattanovich, 2013; Menegon et al., 2022). An ideal *E. coli* sucrose-utilizing platform would integrate growth fitness, effective sugar uptake mechanisms, and increased molasses tolerance, facilitating the efficient production of biobased chemicals and fuels (Kim et al., 2024; Mohamed et al., 2019; Olavarria et al., 2019). In this context, we achieved promising results by enabling PHB production from sucrose in *E. coli* using EasyGuide CRISPR, with yields comparable to that of a naturally producing strain. Ultimately, our study demonstrates that combining CRISPR-mediated genetic engineering with ALE is crucial for unlocking the full catabolic potential of microbial chassis.

Authors contributions

Joneclei A Barreto: Writing – original draft, Visualization, Investigation, Formal analysis, Data curation, Methodology. **Matheus V M L e Silva:** Methodology, Investigation. **Danieli C Marin:** Methodology, Investigation. **Michel Brienzo:**

Methodology, Investigation, Writing – review & editing. **Ana Paula Jacobus:** Methodology, Conceptualization, Writing – review & editing, Funding acquisition. **Jonas Contiero:** Methodology, Investigation, Conceptualization, Writing – review & editing, Funding acquisition. **Jeferson Gross:** Conceptualization, Visualization, Investigation, Formal analysis, Data curation, Methodology, Writing – original draft, Funding acquisition.

Declaration of interest

The authors declare no competing interests.

Data availability

Illumina sequence reads are available at the NCBI under the BioProject number PRJNA1200122.

Acknowledgements

E2348/69 strain was kindly donated by Dr. Fernanda Batista de Andrade (Instituto Butantan, São Paulo, Brazil). We also thank Prof. José Gregório Cabrera Gomez (University of São Paulo, Brazil) for donating the pBBR1MCS-2 plasmid containing the *Cupriavidus necator phaCAB* operon. This work was supported by grants from the São Paulo Research Foundation numbers 17/50249-6, 20/02246-0, 17/22401-8, 17/24453-5, 23/04162-7, and the Brazilian National Council for Scientific and Technological Development (CNPq) grant 404603/2023-8.

Appendix A. Supplementary data

876 Supplementary data to this article can be found online at

877

878 References

- 879 Bandyopadhyay, K., Parua, P. K., Datta, A. B., Parrack, P., 2011. Studies on *Escherichia coli*
880 HflKC suggest the presence of an unidentified lambda factor that influences the lysis-
881 lysogeny switch. *BMC Microbiol.* 11, 34.
- 882 Barembruch, C., Hengge, R., 2007. Cellular levels and activity of the flagellar sigma factor FliA of
883 *Escherichia coli* are controlled by FlgM-modulated proteolysis. *Mol Microbiol.* 65, 76-89.
- 884 Barrick, J. E., Lenski, R. E., 2013. Genome dynamics during experimental evolution. *Nat Rev*
885 *Genet.* 14, 827-39.
- 886 Bassalo, M. C., Garst, A. D., Halweg-Edwards, A. L., Grau, W. C., Domaille, D. W., Mutalik, V. K.,
887 Arkin, A. P., Gill, R. T., 2016. Rapid and Efficient One-Step Metabolic Pathway Integration
888 in *E. coli*. *ACS Synth Biol.* 5, 561-8.
- 889 Blount, Z. D., 2015. The unexhausted potential of *E. coli*. *Elife.* 4.
- 890 Bruschi, M., Boyes, S. J., Sugiarto, H., Nielsen, L. K., Vickers, C. E., 2012. A transferable sucrose
891 utilization approach for non-sucrose-utilizing *Escherichia coli* strains. *Biotechnol Adv.* 30,
892 1001-10.
- 893 Bubeck, P., Winkler, M., Bautsch, W., 1993. Rapid cloning by homologous recombination in vivo.
894 *Nucleic Acids Res.* 21, 3601-2.
- 895 Calloni, G., Chen, T., Schermann, S. M., Chang, H. C., Genevoux, P., Agostini, F., Tartaglia, G.
896 G., Hayer-Hartl, M., Hartl, F. U., 2012. DnaK functions as a central hub in the *E. coli*
897 chaperone network. *Cell Rep.* 1, 251-64.
- 898 Carruthers, D. N., Saleski, T. E., Scholz, S. A., Lin, X. N., 2020. Random Chromosomal Integration
899 and Screening Yields *E. coli* K-12 Derivatives Capable of Efficient Sucrose Utilization. *ACS*
900 *Synth Biol.* 9, 3311-3321.
- 901 Chilcott, G. S., Hughes, K. T., 2000. Coupling of flagellar gene expression to flagellar assembly in
902 *Salmonella enterica* serovar typhimurium and *Escherichia coli*. *Microbiol Mol Biol Rev.* 64,
903 694-708.
- 904 Collias, D., Vialeto, E., Yu, J., Co, K., Almasi, E. D. H., Ruttiger, A. S., Achmedov, T., Strowig, T.,
905 Beisel, C. L., 2023. Systematically attenuating DNA targeting enables CRISPR-driven
906 editing in bacteria. *Nat Commun.* 14, 680.
- 907 Conley, E. C., Saunders, V. A., Saunders, J. R., 1986. Deletion and rearrangement of plasmid
908 DNA during transformation of *Escherichia coli* with linear plasmid molecules. *Nucleic Acids*
909 *Res.* 14, 8905-17.
- 910 Conrad, T. M., Frazier, M., Joyce, A. R., Cho, B. K., Knight, E. M., Lewis, N. E., Landick, R.,
911 Palsson, B. O., 2010. RNA polymerase mutants found through adaptive evolution
912 reprogram *Escherichia coli* for optimal growth in minimal media. *Proc Natl Acad Sci U S A.*
913 107, 20500-5.
- 914 Dong, H., Cui, Y., Zhang, D., 2021. CRISPR/Cas Technologies and Their Applications in
915 *Escherichia coli*. *Front Bioeng Biotechnol.* 9, 762676.
- 916 Dragosits, M., Mattanovich, D., 2013. Adaptive laboratory evolution -- principles and applications
917 for biotechnology. *Microb Cell Fact.* 12, 64.
- 918 Ebel, W., Trempy, J. E., 1999. *Escherichia coli* RcsA, a positive activator of colanic acid capsular
919 polysaccharide synthesis, functions To activate its own expression. *J Bacteriol.* 181, 577-
920 84.

- Fang, H., Zhao, J., Zhao, X., Dong, N., Zhao, Y., Zhang, D., 2024. Standardized Iterative Genome Editing Method for *Escherichia coli* Based on CRISPR-Cas9. *ACS Synth Biol.* 13, 613-623.
- Feng, X., Zhao, D., Zhang, X., Ding, X., Bi, C., 2018. CRISPR/Cas9 Assisted Multiplex Genome Editing Technique in *Escherichia coli*. *Biotechnol J.* 13, e1700604.
- Gietz, R. D., Woods, R. A., 2002. Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. *Methods Enzymol.* 350, 87-96.
- Hanahan, D., 1983. Studies on transformation of *Escherichia coli* with plasmids. *J Mol Biol.* 166, 557-80.
- Hiroe, A., Tsuge, K., Nomura, C. T., Itaya, M., Tsuge, T., 2012. Rearrangement of gene order in the *phaCAB* operon leads to effective production of ultrahigh-molecular-weight poly[(R)-3-hydroxybutyrate] in genetically engineered *Escherichia coli*. *Appl Environ Microbiol.* 78, 3177-84.
- Iguchi, A., Thomson, N. R., Ogura, Y., Saunders, D., Ooka, T., Henderson, I. R., Harris, D., Asadulghani, M., Kurokawa, K., Dean, P., Kenny, B., Quail, M. A., Thurston, S., Dougan, G., Hayashi, T., Parkhill, J., Frankel, G., 2009. Complete genome sequence and comparative genome analysis of enteropathogenic *Escherichia coli* O127:H6 strain E2348/69. *J Bacteriol.* 191, 347-54.
- Jacobus, A. P., Barreto, J. A., de Bem, L. S., Menegon, Y. A., Fier, I., Bueno, J. G. R., Dos Santos, L. V., Gross, J., 2022. EasyGuide Plasmids Support in Vivo Assembly of gRNAs for CRISPR/Cas9 Applications in *Saccharomyces cerevisiae*. *ACS Synth Biol.* 11, 3886-3891.
- Jacobus, A. P., Cavassana, S. D., de, O., Il, Barreto, J. A., Rohwedder, E., Frazzon, J., Basso, T. P., Basso, L. C., Gross, J., 2024. Optimal trade-off between boosted tolerance and growth fitness during adaptive evolution of yeast to ethanol shocks. *Biotechnol Biofuels Bioprod.* 17, 63.
- Jacobus, A. P., Gross, J., 2015. Optimal cloning of PCR fragments by homologous recombination in *Escherichia coli*. *PLoS One.* 10, e0119221.
- Jacobus, A. P., Stephens, T. G., Youssef, P., Gonzalez-Pech, R., Ciccotosto-Camp, M. M., Dougan, K. E., Chen, Y., Basso, L. C., Frazzon, J., Chan, C. X., Gross, J., 2021. Comparative Genomics Supports That Brazilian Bioethanol *Saccharomyces cerevisiae* Comprise a Unified Group of Domesticated Strains Related to Cachaca Spirit Yeasts. *Front Microbiol.* 12, 644089.
- Jiang, W., Bikard, D., Cox, D., Zhang, F., Marraffini, L. A., 2013. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat Biotechnol.* 31, 233-9.
- Jiang, Y., Chen, B., Duan, C., Sun, B., Yang, J., Yang, S., 2015. Multigene editing in the *Escherichia coli* genome via the CRISPR-Cas9 system. *Appl Environ Microbiol.* 81, 2506-14.
- Jones, D. H., Howard, B. H., 1991. A rapid method for recombination and site-specific mutagenesis by placing homologous ends on DNA using polymerase chain reaction. *Biotechniques.* 10, 62-6.
- Joshi, S. H., Yong, C., Gyorgy, A., 2022. Inducible plasmid copy number control for synthetic biology in commonly used *E. coli* strains. *Nat Commun.* 13, 6691.
- Juhas, M., Evans, L. D., Frost, J., Davenport, P. W., Yarkoni, O., Fraser, G. M., Ajioka, J. W., 2014. *Escherichia coli* flagellar genes as target sites for integration and expression of genetic circuits. *PLoS One.* 9, e111451.
- Jung, S. C., Smith, C. L., Lee, K. S., Hong, M. E., Kweon, D. H., Stephanopoulos, G., Jin, Y. S., 2010. Restoration of growth phenotypes of *Escherichia coli* DH5alpha in minimal media through reversal of a point mutation in *purB*. *Appl Environ Microbiol.* 76, 6307-9.

- Kihara, A., Akiyama, Y., Ito, K., 1996. A protease complex in the *Escherichia coli* plasma membrane: HflKC (HflA) forms a complex with FtsH (HflB), regulating its proteolytic activity against SecY. *EMBO J.* 15, 6122-31.
- Kim, G. Y., Yang, J., Han, Y. H., Seo, S. W., 2024. Synthetic redesign of *Escherichia coli* W for faster metabolism of sugarcane molasses. *Microb Cell Fact.* 23, 242.
- Kostylev, M., Otwell, A. E., Richardson, R. E., Suzuki, Y., 2015. Cloning Should Be Simple: *Escherichia coli* DH5alpha-Mediated Assembly of Multiple DNA Fragments with Short End Homologies. *PLoS One.* 10, e0137466.
- Kuijpers, N. G., Solis-Escalante, D., Bosman, L., van den Broek, M., Pronk, J. T., Daran, J. M., Daran-Lapujade, P., 2013. A versatile, efficient strategy for assembly of multi-fragment expression vectors in *Saccharomyces cerevisiae* using 60 bp synthetic recombination sequences. *Microb Cell Fact.* 12, 47.
- Lee, C. L., Ow, D. S., Oh, S. K., 2006. Quantitative real-time polymerase chain reaction for determination of plasmid copy number in bacteria. *J Microbiol Methods.* 65, 258-67.
- Lessard, J. C., 2013. Transformation of *E. coli* via electroporation. *Methods Enzymol.* 529, 321-7.
- Li, Q., Sun, B., Chen, J., Zhang, Y., Jiang, Y., Yang, S., 2021. A modified pCas/pTargetF system for CRISPR-Cas9-assisted genome editing in *Escherichia coli*. *Acta Biochim Biophys Sin (Shanghai).* 53, 620-627.
- Li, Q., Sun, M., Lv, L., Zuo, Y., Zhang, S., Zhang, Y., Yang, S., 2023. Improving the Editing Efficiency of CRISPR-Cas9 by Reducing the Generation of Escapers Based on the Surviving Mechanism. *ACS Synth Biol.* 12, 672-680.
- Li, Y., Lin, Z., Huang, C., Zhang, Y., Wang, Z., Tang, Y. J., Chen, T., Zhao, X., 2015. Metabolic engineering of *Escherichia coli* using CRISPR-Cas9 mediated genome editing. *Metab Eng.* 31, 13-21.
- Lilic, M., Jovanovic, M., Jovanovic, G., Savic, D. J., 2003. Identification of the CysB-regulated gene, hslJ, related to the *Escherichia coli* novobiocin resistance phenotype. *FEMS Microbiol Lett.* 224, 239-46.
- Lim, S. R., Lee, H. J., Kim, H. J., Lee, S. J., 2023. Multiplex Single-Nucleotide Microbial Genome Editing Achieved by CRISPR-Cas9 Using 5'-End-Truncated sgRNAs. *ACS Synth Biol.* 12, 2203-2207.
- Lopilato, J., Bortner, S., Beckwith, J., 1986. Mutations in a new chromosomal gene of *Escherichia coli* K-12, pcnB, reduce plasmid copy number of pBR322 and its derivatives. *Mol Gen Genet.* 205, 285-90.
- Maniatis, T., Fritsch, E. F., Sambrook, J., 1982. Molecular cloning: a laboratory manual. . Cold Spring Harbor: Cold Spring Harbor Laboratory.
- Menegon, Y. A., Gross, J., Jacobus, A. P., 2022. How adaptive laboratory evolution can boost yeast tolerance to lignocellulosic hydrolyses. *Curr Genet.* 68, 319-342.
- Mohamed, E. T., Mundhada, H., Landberg, J., Cann, I., Mackie, R. I., Nielsen, A. T., Herrgard, M. J., Feist, A. M., 2019. Generation of an *E. coli* platform strain for improved sucrose utilization using adaptive laboratory evolution. *Microb Cell Fact.* 18, 116.
- Moreb, E. A., Hoover, B., Yaseen, A., Valyasevi, N., Roecker, Z., Menacho-Melgar, R., Lynch, M. D., 2017. Managing the SOS Response for Enhanced CRISPR-Cas-Based Recombineering in *E. coli* through Transient Inhibition of Host RecA Activity. *ACS Synth Biol.* 6, 2209-2218.
- Murphy, E., 1985. Nucleotide sequence of a spectinomycin adenylyltransferase AAD(9) determinant from *Staphylococcus aureus* and its relationship to AAD(3'') (9). *Mol Gen Genet.* 200, 33-9.
- Nozaki, S., Niki, H., 2019. Exonuclease III (XthA) Enforces In Vivo DNA Cloning of *Escherichia coli* To Create Cohesive Ends. *J Bacteriol.* 201.

- Olavarria, K., Fina, A., Velasco, M. I., van Loosdrecht, M. C. M., Wahl, S. A., 2019. Metabolism of sucrose in a non-fermentative *Escherichia coli* under oxygen limitation. *Appl Microbiol Biotechnol.* 103, 6245-6256.
- Oliner, J. D., Kinzler, K. W., Vogelstein, B., 1993. In vivo cloning of PCR products in *E. coli*. *Nucleic Acids Res.* 21, 5192-7.
- Pyne, M. E., Moo-Young, M., Chung, D. A., Chou, C. P., 2015. Coupling the CRISPR/Cas9 System with Lambda Red Recombineering Enables Simplified Chromosomal Gene Replacement in *Escherichia coli*. *Appl Environ Microbiol.* 81, 5103-14.
- Raffaelli, N., Lorenzi, T., Mariani, P. L., Emanuelli, M., Amici, A., Ruggieri, S., Magni, G., 1999. The *Escherichia coli* NadR regulator is endowed with nicotinamide mononucleotide adenylyltransferase activity. *J Bacteriol.* 181, 5509-11.
- Reisch, C. R., Prather, K. L., 2015. The no-SCAR (Scarless Cas9 Assisted Recombineering) system for genome editing in *Escherichia coli*. *Sci Rep.* 5, 15096.
- Rodrigues, C. I. S., Della-Bianca, B. E., Gombert, A. K., 2021. μ MAX of *Saccharomyces Cerevisiae*: So Often Used, So Seldom Put into Perspective. *Research Square*.
- Schavemaker, P. E., Lynch, M., 2022. Flagellar energy costs across the tree of life. *Elife.* 11.
- Shukal, S., Lim, X. H., Zhang, C., Chen, X., 2022. Metabolic engineering of *Escherichia coli* BL21 strain using simplified CRISPR-Cas9 and asymmetric homology arms recombineering. *Microb Cell Fact.* 21, 19.
- Sohn, Y. J., Kim, H. T., Baritugo, K. A., Song, H. M., Ryu, M. H., Kang, K. H., Jo, S. Y., Kim, H., Kim, Y. J., Choi, J. I., Park, S. K., Joo, J. C., Park, S. J., 2020. Biosynthesis of polyhydroxyalkanoates from sucrose by metabolically engineered *Escherichia coli* strains. *Int J Biol Macromol.* 149, 593-599.
- Sun, D., Wang, L., Mao, X., Fei, M., Chen, Y., Shen, M., Qiu, J., 2019. Chemical transformation mediated CRISPR/Cas9 genome editing in *Escherichia coli*. *Biotechnol Lett.* 41, 293-303.
- Tang, Q., Lou, C., Liu, S. J., 2017. Construction of an easy-to-use CRISPR-Cas9 system by patching a newly designed EXIT circuit. *J Biol Eng.* 11, 32.
- Tomoyasu, T., Gamer, J., Bukau, B., Kanemori, M., Mori, H., Rutman, A. J., Oppenheim, A. B., Yura, T., Yamanaka, K., Niki, H., et al., 1995. *Escherichia coli* FtsH is a membrane-bound, ATP-dependent protease which degrades the heat-shock transcription factor sigma 32. *EMBO J.* 14, 2551-60.
- van den Broek, L. A., van Bostel, E. L., Kievit, R. P., Verhoef, R., Beldman, G., Voragen, A. G., 2004. Physico-chemical and transglucosylation properties of recombinant sucrose phosphorylase from *Bifidobacterium adolescentis* DSM20083. *Appl Microbiol Biotechnol.* 65, 219-27.
- Wang, J., Sui, X., Ding, Y., Fu, Y., Feng, X., Liu, M., Zhang, Y., Xian, M., Zhao, G., 2021. A fast and robust iterative genome-editing method based on a Rock-Paper-Scissors strategy. *Nucleic Acids Res.* 49, e12.
- Watson, J. F., Garcia-Nafria, J., 2019. In vivo DNA assembly using common laboratory bacteria: A re-emerging tool to simplify molecular cloning. *J Biol Chem.* 294, 15271-15281.
- Zerbini, F., Zanella, I., Fraccascia, D., Konig, E., Irene, C., Frattini, L. F., Tomasi, M., Fantappie, L., Ganfini, L., Caproni, E., Parri, M., Grandi, A., Grandi, G., 2017. Large scale validation of an efficient CRISPR/Cas-based multi gene editing protocol in *Escherichia coli*. *Microb Cell Fact.* 16, 68.
- Zhao, D., Yuan, S., Xiong, B., Sun, H., Ye, L., Li, J., Zhang, X., Bi, C., 2016. Development of a fast and easy method for *Escherichia coli* genome editing with CRISPR/Cas9. *Microb Cell Fact.* 15, 205.
- Zhao, J., Baba, T., Mori, H., Shimizu, K., 2004. Effect of zwf gene knockout on the metabolism of *Escherichia coli* grown on glucose or acetate. *Metab Eng.* 6, 164-74.

1067 Zhao, W., Guo, Y., 2024. Increasing the efficiency of gene editing with CRISPR-Cas9 via
1068 concurrent expression of the Beta protein. *Int J Biol Macromol.* 270, 132431.
1069