Exploiting Endogenous CRISPR-Cas9 System for Functional Engineering of Probiotic *Lacticaseibacillus rhamnosus* GG (LGG)

Authors: Zifan Xie[1,2], Yong-Su Jin[1,2]*, and Michael J. Miller[1,2]*

Affiliation:

[1]Department of Food Science and Human Nutrition, University of Illinois Urbana-Champaign, Urbana, IL, USA

[2] Carl R. Woese Institute for Genomic Biology, University of Illinois Urbana-Champaign, Urbana, IL, USA

E-mail: Zifan Xie (zifanxie@illinois.edu), Yong-Su Jin (ysjin@illinois.edu), Michael J. Miller (mille216@illinois.edu)

***Corresponding author:**

Prof. Yong-Su Jin (ysjin@illinois.edu)

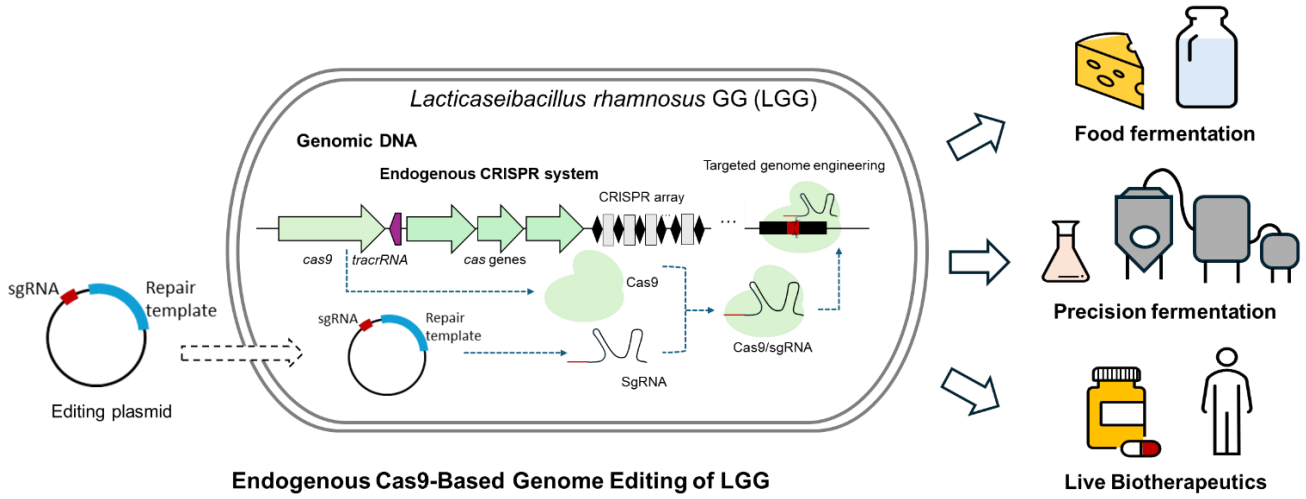
Prof. Michael J. Miller (mille216@illinois.edu)

ABSTRACT

Lacticaseibacillus rhamnosus GG (LGG) is one of the most extensively studied probiotic strains, widely used in food and health applications. However, the absence of efficient, precise genome editing methods has limited its broader potential and functional versatility. Here, we present an endogenous type II-A CRISPR-Cas9 genome editing workflow for LGG designed for functional strain construction. Using a plasmid interference assay together with single-nucleotide substitutions, we confirm the precise PAM requirement as 5′-NGAAA-3′. We pair a synthetic sgRNA cassette with homology-directed repair donors to enable precise deletions and insertions across multiple loci. Using this optimized and precise genome engineering method, we generated a β-glucuronidase (GUS)-expressing LGG strain for robust strain tracking within complex gut microbial communities. This work removes barriers to LGG engineering, expands the probiotic CRISPR toolkit, and provides broadly applicable strategies for designing next-generation probiotics with applications in food biotechnology and microbial therapeutics.

Keywords: *CRISPR-Cas9*, *probiotic, synthetic biology*, *lactic acid bacteria*

30 **GRAPHICAL ABSTRACT**



31 **Endogenous Cas9-Based Genome Editing of LGG**

32

INTRODUCTION

Lacticaseibacillus rhamnosus GG (LGG) is one of the most extensively studied probiotic strains due to its well-documented therapeutic and industrial benefits. Originally isolated from the human gastrointestinal tract, LGG demonstrates robust tolerance to gastric acid and bile salts, strong adherence to intestinal epithelial cells, and exceptional stability throughout food processing (Segers & Lebeer, 2014; Capurso, 2019). Since its discovery, numerous clinical and preclinical studies have demonstrated LGG's ability to enhance gut barrier function, modulate host immune system, and exert antimicrobial activity (Hudault et al., 1997; Yan & Polk, 2002; Sindhu et al., 2014; Gao et al., 2019). Given its Generally Recognized as Safe (GRAS) status from the US Food and Drug Administration (GRAS Notice No. 845), LGG has been widely used in dietary supplements, fermented food products, and clinical therapies aimed at modulating gut health, immune modulation, and microbial homeostasis (Jia et al., 2016; Chan et al., 2021). More recently, LGG's well-established safety profile and versatile functionality have attracted significant interest in its development as a microbial chassis for next-generation live biotherapeutics and precision fermentation applications (Günaydın et al., 2014; Wan et al., 2016; Petrova et al., 2018; Li et al., 2025).

Despite over 300 human-focused publications underscoring LGG's extensive health benefits, its genetic manipulation and strain optimization have been significantly limited by a lack of efficient genome editing tools. Because of limited replicable plasmids, chromosomal modification in LGG has typically relied on non-replicative suicide vectors (Lebeer et al., 2012; S. Zhang et al., 2018). These approaches depend on low-frequency double-crossover events and usually require multiple rounds of selection or counterselection, making them labor-intensive, time-consuming, and poorly suited to complex or high-throughput engineering. The absence of precise genome-editing technologies has thus

55 hindered mechanistic studies of LGG's probiotic actions and restricted the development of advanced
56 strains for therapeutic and industrial applications.

57 Recently, many innovative genome modification approaches have been developed and applied in
58 lactic acid bacteria to determine relevant phenotypes and to enhance beneficial properties (Xie et al.,
59 2024). Notably, the CRISPR-Cas (clustered regularly interspaced short palindromic repeats and
60 CRISPR-associated proteins) system, originally discovered as a bacterial adaptive immune mechanism,
61 has emerged as a powerful tool for precise and efficient genome modification in various organisms (Hsu
62 et al., 2014; Adli, 2018). Heterologous CRISPR-Cas systems have been successfully employed in many
63 lactobacilli, such as *Limosilactobacillus reuteri*, *Lactiplantibacillus plantarum*, *Lacticaseibacillus casei*,
64 *Lactobacillus acidophilus*, *Lactobacillus gasseri*, and *Lacticaseibacillus paracasei* (Song et al., 2017;
65 Leenay et al., 2018; Huang et al., 2019; Goh & Barrangou, 2021). Despite these successes, heterologous
66 CRISPR systems present several challenges. These systems generally require at least one plasmid for the
67 expression of large Cas9 protein, most commonly derived from *Streptococcus pyogenes* (Cong et al.,
68 2013). Such large plasmid constructs often exhibit reduced plasmid stability, lower transformation
69 efficiency, and impose metabolic burdens on host cells. Additionally, heterologous Cas proteins may
70 exhibit suboptimal activity or cytotoxicity in certain strains, further limiting editing efficiency and
71 viability (J. Zhang et al., 2018; Vento et al., 2019). To overcome these limitations, recent research has
72 increasingly focused on endogenous CRISPR-Cas systems encoded within the host genomes. These
73 native systems eliminate the need for exogenous proteins and exhibit enhanced compatibility with host
74 machinery, resulting in improved editing precision and reduced cellular toxicity. Functional CRISPR-
75 Cas systems have been identified in many lactobacilli strains (Crawley et al., 2018), and some of them
76 have been successfully repurposed for precise genome editing. For example, the endogenous subtype I-E
77 CRIPSR-Cas system has been utilized for efficient genetic manipulation of *Lactobacillus crispatus*

78 (Hidalgo-Cantabrana, Goh, Pan, et al., 2019), and the subtype II-A CRISPR-Cas9 system has been used
79 to achieve genome engineering in *Lcb. paracasei* (Gu et al., 2022). Recently, type II-A CRISPR-Cas9
80 system was identified in LGG (Crawley et al., 2018; Panahi et al., 2023), and the Cas9 protein from
81 LGG has been heterologously expressed for plant genome engineering (Zhong et al., 2023). However,
82 those studies did not establish a practical genome-editing workflow directly inside LGG.

83 Here, we establish an endogenous type II-A CRISPR-Cas9 genome-editing system for LGG that
84 enables precise gene deletions and insertions. Using a plasmid-interference assay with single-nucleotide
85 protospacer adjacent motif (PAM) variants, we confirmed the specificity of the native Cas9 PAM as 5′-
86 NGAAA-3′. A replicative editing plasmid was constructed to carry synthetic single-guide RNA
87 (sgRNA) cassettes and homologous repair templates, enabling efficient genome modifications across
88 multiple loci. To demonstrate the practical applicability of this system, we engineered a β-glucuronidase
89 (gusA) expressing strain that allows precise tracking in complex microbial communities and host
90 models. Animal experiments further confirmed that the gusA-positive LGG strain serves as a suitable
91 model for studying LGG-associated health benefits. Together, these results establish an efficient LGG
92 genome-editing platform that facilitates mechanistic investigations and accelerates the development of
93 functional probiotic strains for food biotechnology and microbial therapeutics.

94

95 **RESULTS**

96 **Specificity of the Endogenous CRISPR-Cas9 System in LGG**

97 LGG harbors a complete endogenous type II-A CRISPR-Cas9 locus containing four Cas protein-
98 encoding genes (*cas1*, *cas2*, *csn2*, and *cas9*) adjacent to a CRISPR array (Crawley et al., 2018; Panahi et
99 al., 2023). The array contains 24 distinct 30-base pair (bp) spacers interspaced by 36-bp direct repeats

100 (Figure 1A). Consistent with previous reports, alignment of protospacer-flanking sequences in LGG
101 indicates a preferred PAM of 5′-NGAAA-3′ (Figure 1B) (Crawley et al., 2018; Zhong et al., 2023)

102 To validate the specificity of the PAM sequence, we performed a plasmid interference assay
103 incorporating single-nucleotide PAM substitutions (Figure 1C). Interference plasmids were constructed
104 based on vector pTRK870 (Supplementary Table S2), a derivative of pTRK830 with a multiple cloning
105 site (Xie et al., 2025). Each plasmid carried a protospacer matching the most recently acquired spacer 24
106 (5'-AAGAAACGATCCCGAGTTTCTGGATTGCTT-3') in the CRISPR array, either accompanied by
107 or lacking the predicted PAM sequence.

108 Transformation of LGG with plasmids (pPS1AGAAA, pPS1TGAAA, and pPS1GGAAA)
109 containing the predicted PAM sequences resulted in no detectable transformants, demonstrating robust
110 interference mediated by the endogenous CRISPR-Cas9 system (Figure 1D). To further evaluate PAM
111 specificity, we introduced single-nucleotide substitutions at conserved positions (positions 2-5) of the
112 predicted PAM sequence, generating seven PAM variants AAAAA, ACAAA, TCAAA, AGGAA,
113 AGTAA, AGATA, and AGAAG. As shown in Figure 1D, plasmids with altered PAM sequences,
114 pPS1AAAAA, pPS1ACAAA, and pPS1TCAAA, failed to produce a detectable interference response,
115 indicating that guanine at position 2 is essential for effective PAM recognition. Additionally, plasmid
116 pPS1AGATA, carrying a thymine substitution at position 4, showed significantly reduced interference,
117 indicating the importance of adenine conservation at this position. Notably, plasmids pPS1AGGAA and
118 pPS1AGAAG, each harboring substitutions at positions 3 and 5, exhibited partial interference,
119 suggesting a moderate tolerance of these positions to guanine substitutions. However, significantly
120 fewer colonies were recovered using the optimal PAM (pPS1AGAAA), confirming a strong preference
121 for adenines at positions 3–5.

122 Collectively, these results confirm the PAM recognized by the endogenous Cas9 in LGG as 5′-
123 NGAAA-3′ and demonstrate its high sequence specificity, with enhanced interference activity observed
124 under our assay conditions.

125 **Construction of Editing Plasmids for Reprogramming the Endogenous CRIPSR-Cas9 System**

126 After confirming the specificity of PAM sequence, we then explored its potential for genome
127 editing in LGG. Initially, we designed an editing plasmid derived from pTRK870, which contains a
128 homologous repair template and a CRISPR RNA (crRNA) expression cassette under control of the
129 native CRISPR array leader sequence (Supplementary Figure S1). However, we found that due to the
130 inherent instability of rolling circle plasmids (Xie et al., 2025), the space region between the repeats was
131 frequently looped out from the plasmid during plasmid replication, resulting in spacer-deficient plasmids
132 and a high frequency of wild-type (WT) survivors (Supplementary Figure S1).

133 To overcome this limitation, we transitioned to a sgRNA-based approach for directing precise
134 genome targeting by the endogenous CRISPR-Cas9. Based on the predicted structure of the crRNA/
135 trans-activating CRISPR RNA (tracrRNA) duplex and prior transcriptional sequencing data (Crawley et
136 al., 2018), we engineered an artificial sgRNA by linking the crRNA and tracrRNA with an AAUC loop
137 (Figure 1E). We then constructed a targeting plasmid encoding the sgRNA under the control of the
138 strong constitutive promoter *Ppgm* from *L. acidophilus* NCFM (Figure 1F) (Xie et al., 2025).

139 To validate our artificial sgRNA design, we targeted a highly conserved region within the 16S
140 ribosomal RNA (rRNA) gene. Introduction of this 16S-targeting plasmid into LGG significantly reduced
141 viable transformants compared to the pTRK870 control plasmid (Supplementary Figure S2),
142 demonstrating effective Cas9-mediated chromosomal cleavage and providing direct evidence of
143 functional sgRNA expression. To evaluate the broader applicability of this sgRNA system, we tested the

144 same 16S-targeting plasmid in several additional *L. rhamnosus* strains (Supplementary Table S1). Unlike
145 LGG, these other strains yielded viable transformants without evidence of Cas9-mediated lethality
146 (Supplementary Figure S2), suggesting limited portability of our platform across closely related strains.
147 Possible reasons include differences in plasmid transformation efficiency, strain-specific PAM sequence
148 preferences, or variations in CRISPR architecture among strains (Supplementary Figure S2).

149 Building upon this validated targeting plasmid, we next designed a comprehensive editing
150 plasmid incorporating both the sgRNA expression cassette and homologous repair templates (HR1 and
151 HR2) flanking the intended genomic modification site (Figure 1F).

152 **Reprogramming the Endogenous CRISPR-Cas9 System for Gene Deletion**

153 As a test, we first attempted to target the *fucI* gene (LGG_02685), which encodes L-fucose
154 isomerase (Figure 2A). Deletion of *fucI* is expected to significantly impair LGG growth when fucose is
155 the sole carbon source (Becerra et al., 2015). A 30-bp DNA sequence complementary to *fucI* was
156 selected as the spacer and cloned into the pTRK870 backbone, together with two 1-kb homologous
157 arms, to generate the editing plasmid pfucI-sgRNA-RT (Supplementary Table S2). As controls, we also
158 constructed plasmids containing only the repair template (pfucI-RT) or only the sgRNA (pfucI-sgRNA)
159 (Supplementary Table S2). Transformation of LGG with the self-targeting plasmid (pfucI-sgRNA)
160 yielded no colonies (Figure 2B), confirming effective genome cleavage triggered by the endogenous
161 CRISPR-Cas9 system guided by the sgRNA. In contrast, transformation with the repair template-only
162 plasmid (pfucI-RT) generated numerous colonies (Figure 2B), but none of screened colonies carried the
163 intended deletion genotype (Supplementary Figure S3), indicating that the repair template alone was
164 insufficient for generating targeted deletions efficiently. Importantly, transformation with the complete
165 editing plasmid (pfucI-sgRNA-RT) resulted in significantly fewer colonies (four colonies). PCR
166 screening of these transformants identified one colony (Colony No. 1) displaying the expected 2.1-kb

167 PCR product, indicative of the deletion genotype, whereas the other three exhibited wild-type PCR
168 products (2.7 kb; Figure 2C). Sanger sequencing of the 2.1-kb amplicon confirmed the precise deletion,
169 and the successfully engineered strain was designated MJM568 (Figure 2D). Phenotypic analysis
170 further revealed that the MJM568 strain could not utilize fucose as the sole sugar source (Figure 2E).

171 To further evaluate the versatility of this system, we successfully deleted a larger 1.5-kb genomic
172 region encoding coenzyme A (CoA) synthase (*acs1*) using the same strategy (Supplementary Figure S4).
173 This demonstrated the platform's capability for targeting and deleting genomic regions of varying
174 lengths. Collectively, these results confirm that the LGG endogenous CRISPR-Cas9 system can be
175 effectively harnessed for targeted gene deletions, with both the sgRNA-guided Cas9 cleavage and
176 homologous repair templates essential for achieving precise genomic modifications.

177 **Reprogramming the Endogenous CRISPR-Cas9 System for Gene Insertion**

178 After demonstrating the feasibility of gene deletions using the native CRISPR-Cas9 system in
179 LGG, we next aimed to evaluate the system's capacity for gene insertion. The ability to stably integrate
180 heterologous genes is a critical feature for engineering LGG as a microbial cell factory or live
181 biotherapeutic. However, the previously used plasmid backbone pTRK870, which replicates via a
182 rolling-circle mechanism, is inherently unstable and has limited insert capacity (Xie et al., 2025). To
183 address these limitations, we constructed a new editing backbone, pMJM100 (Supplementary Figure
184 S5), which is a smaller and more stable plasmid with improved capacity for accommodating larger
185 genetic inserts. This redesigned backbone provides a robust platform for gene insertion mediated by the
186 native CRISPR-Cas9 machinery in LGG.

187 To test this platform, we repurposed the *fucI*-targeting sgRNA expression cassette previously
188 used for gene deletion and combined it with a homologous repair template designed to partially delete

189 *fucI* and simultaneously insert a chloramphenicol resistance gene (*cam*) expression cassette (Figure 3A).
190 The *cam* cassette was flanked by 0.5 kb homology arms derived from the original *fucI* deletion template.
191 The use of shorter arms reduced the plasmid size and improved its stability during cloning and
192 transformation. The resulting construct, pfucI-sgRNA-Cm-RT2, was assembled using the pMJM100
193 backbone.

194 Following transformation of pfucI-sgRNA-Cm-RT2 into LGG, six erythromycin-resistant
195 transformants were obtained (Figure 3B). PCR screening revealed that one transformant (Colony No. 1)
196 displayed the expected amplicon size for the *cam* gene insertion (Figure 3C), which was further
197 confirmed by Sanger sequencing (Figure 3D). To eliminate the editing plasmid, the verified mutant was
198 passaged twice under non-selective conditions, yielding an erythromycin-sensitive derivative
199 (Supplementary Figure S6), designated MJM655. As excepted, MJM655 showed robust growth on MRS
200 plates supplemented with 5 μg/mL chloramphenicol, while the WT LGG strain remained
201 chloramphenicol-sensitive (Figure 3E).

202 **Construction of GUS-Positive LGG**

203 To demonstrate the application of our endogenous CRISPR-Cas9 system for functional gene
204 insertion, we engineered a GUS-positive LGG strain by inserting the *gusA* gene from *Lactobacillus*
205 *gasseri* ADH, a GRAS strain. The *gusA* gene encodes β-glucuronidase and is widely used as a reporter
206 in lactic acid bacteria. We aimed to create a chromosomally integrated reporter strain suitable for
207 tracking in complex microbial communities.

208 To achieve this, we constructed the editing plasmid prha-sgRNA-gusA-RT2, designed to
209 introduce *gusA* gene expression cassette into the intergenic region between the fucose and rhamnose
210 operons (Figure 4A). Following transformation and selection, colony PCR of one transformant (Colony

211 No. 1) showed both wild-type and mutant bands (Figure 4B), indicating a mixed population. Re-
212 isolation of colony No.1 yielded colonies with a single PCR product of the expected insertion size
213 (Supplementary Figure S7), and Sanger sequencing confirmed precise genomic integration of the *gusA*
214 cassette at the target locus (Figure 4C). After plasmid curing through non-selective subculturing, the
215 resulting genome-edited strain was designated MJM670. To verify *gusA* expression, we plated MJM670
216 and WT LGG on MRS agar supplemented with X-gluc. As expected, MJM670 formed distinct blue
217 colonies, while the WT LGG colonies remained white (Figure 4D), confirming successful chromosomal
218 expression of *gusA*. Moreover, MJM670 maintained consistent GUS expression over serial passages, in
219 contrast to a plasmid-based *gusA* construct, which showed progressive loss of expression (Figure 4E),
220 underscoring the advantage of chromosomal integration for long-term genetic stability.

221 To evaluate whether the insertion affected adjacent gene function, we compared the growth
222 profiles of MJM670 and WT LGG in media supplemented with glucose, fucose or rhamnose. As shown
223 in Supplementary Figure S8, MJM670 exhibited growth patterns nearly identical to the wild-type strain
224 under all tested sugar conditions, indicating that *gusA* integration did not disrupt the function of the
225 flanking fucose and rhamnose operons. Together, these results demonstrate that the intergenic region
226 between the fucose and rhamnose operons is a permissive site for gene insertion, and that prha-sgRNA-
227 RT2 is an effective vector for functional gene integration in LGG. This strategy offers a valuable tool for
228 constructing recombinant probiotic strains for metabolic enhancement and therapeutic applications.

229 **Quantification of GUS-Positive LGG in Mouse Fecal Samples**

230 Analyzing LGG abundance in fecal samples offers a non-invasive and comprehensive view of
231 bacterial persistence through the gastrointestinal (GI) tract. Fecal enumeration is particularly valuable
232 for assessing long-term colonization and microbial dynamics, providing insight into the overall stability
233 and retention of probiotic strains in vivo. To evaluate the persistence and viability of engineered GUS-

234 positive LGG, MJM670, we conducted both plating- and qPCR-based analysis of fecal samples
235 collected from C57BL/6J mice following a single oral gavage (Figure 5A).

236 As previously shown in Figure 4D, MJM670 expresses β-glucuronidase, enabling colony
237 visualization on MRS agar supplemented with X-Gluc. GUS activity resulted in distinct blue colonies,
238 allowing clear differentiation from native gut microbiota (Supplementary Figure S9). Based on plating
239 analysis (Figure 5B), approximately 10^7 CFU/g MJM670 were detected in feces 12 hours after one-time
240 gavage, which decreased to ~10^5 CFU/g at 24 hours. By 48 hours, MJM670 was undetectable in all
241 groups, indicating complete clearance. In parallel, we performed qPCR using LGG-specific primers to
242 assess total LGG abundance regardless of cell viability (Figure 5C). The qPCR results revealed that
243 MJM670 exhibited similar survival dynamics in the GI tract as the WT LGG strain (Figure 5C), with
244 both showing a rapid decline over time. These results further support the observation that both WT LGG
245 and MJM670 exhibit transient colonization following oral administration.

246 When comparing the two methods (Figure 5B and 5C), the plating method consistently yielded
247 higher LGG counts than qPCR. This is likely due to incomplete DNA recovery from fecal samples using
248 commercial stool DNA kits, which may underestimate total bacterial abundance. Moreover, while qPCR
249 can detect both viable and nonviable cells, plating selectively recovers only viable LGG cells capable of
250 colony formation, providing a direct measure of functional probiotic survival. Despite its higher
251 detection threshold, the plating method provides unique advantages for engineered strains like MJM670.
252 Its visual specificity enables clear identification of the target strain, and it allows isolation of individual
253 colonies for downstream phenotypic or genomic characterization, which is an essential step for verifying
254 genetic stability or strain behavior after going through the GI tract.

255 Overall, the GUS-positive LGG strain, MJM670, demonstrated reliable detection via plating and
256 enabled viable cell recovery, supporting its application in colonization tracking and functional studies
257 within the gastrointestinal microbiome.

258

259 **DISCUSSION**

260 LGG is one of the most widely used probiotics in fermented foods and dietary supplements, yet
261 the development of enhanced LGG strains has been constrained by the lack of efficient genome editing
262 tools. Existing methods based on non-replicable vectors are inefficient, labor-intensive, and require
263 antibiotic selection. While heterologous CRISPR-Cas9 systems have been applied in other *Lactobacillus*
264 species, their use in LGG has been restricted due to the lack of stable plasmids capable of
265 accommodating the large *cas9* gene. In this situation, repurposing the endogenous CRISPR–Cas
266 machinery offers a promising approach for editing in LGG. A type II-A CRISPR–Cas9 system has
267 previously been identified in LGG, and its Cas9 protein has been heterologously expressed for plant
268 genome engineering (Crawley et al., 2018; Zhong et al., 2023). However, these studies did not establish
269 a practical genome-editing workflow in LGG. Critical parameters, such as PAM sequence specificity in
270 the LGG context, sgRNA expression system, homologous repair templates design, plasmid curing, and
271 edit stability, have remained undefined. Here, we addressed these gaps and developed an endogenous
272 type II-A CRISPR-Cas9 platform for LGG that enables efficient genome editing to construct functional
273 strains for fermentation and therapeutic applications.

274 We first confirmed 5'-NGAAA-3' as the preferred PAM, exhibiting strict sequence requirements
275 for target recognition. To assess the reprogrammability of the native system, we designed an artificial
276 sgRNA targeting the essential 16S rRNA gene. Introduction of this construct resulted in a marked loss of
277 viable transformants in LGG, confirming the high activity and specificity of the native system and

278 demonstrating that artificial sgRNAs can be correctly recognized and processed in LGG. In contrast, the
279 same construct failed to induce lethality in other *Lcb. rhamnosus* strains, despite some encoding type II-
280 A CRISPR loci. These strains showed no detectable targeting activity, likely due to low transformation
281 efficiency or variations in PAM specificity or CRISPR locus architecture. These findings highlight the
282 strain-dependent nature of endogenous CRISPR systems and underscore the importance of strain-
283 specific validation for successful repurposing.

284 Using the validated system, we developed replicable plasmids for sgRNA expression and repair
285 template delivery, achieving precise deletions and insertions without scars or antibiotic-resistant genes,
286 making them suitable for food and therapeutic applications. Although overall editing efficiency in LGG
287 was lower than in some other lactobacilli (Gu et al., 2022; Hidalgo-Cantabrana, Goh, Pan, et al., 2019),
288 likely due to the inherent instability of rolling-circle plasmids and relatively lower homologous
289 recombination efficiency, this workflow offers several practical advantages. The system enables scarless
290 gene deletions and insertions using a single plasmid, without the need for heterologous Cas9 expression
291 or laborious selection. In practice, desired mutants can be routinely obtained in one transformation and a
292 brief colony screen, typically completed within three days, without multiple rounds of subculturing.
293 Moreover, the editing plasmids are easily curable, leaving no selection or counterselection markers in
294 the final strains, which support iterative marker-free editing to stack multiple modifications in the same
295 genetic background. Together, these attributes make the platform well-suited for functional genetics,
296 strain optimization, and medium-throughput genome engineering in LGG, where speed, precision, and
297 regulatory compatibility are at a premium.

298 To demonstrate the application of this genome editing platform, we engineered a GUS-
299 expressing LGG strain (MJM670) to enable specific and robust tracking within mixed microbial
300 populations and animal models. In the gastrointestinal tract, endogenous β-glucuronidase activity is

301 present in only a limited number of bacterial species beyond *Escherichia coli*, making *gusA* a reliable
302 and specific reporter gene. Genomic integration into LGG enables clear visual differentiation from
303 native gut microbes and provides a direct alternative to qPCR-based detection. While qPCR detects
304 DNA from both viable and nonviable cells, plating recovers only viable LGG, providing a direct
305 measure of functional probiotic survival. The *gusA* gene, sourced from the GRAS strain *L. gasseri* ADH,
306 supports biosafety and avoids antibiotic resistance genes, reducing the risk of horizontal gene transfer
307 and the spread of antibiotic resistance within the gut environment. This GUS-positive LGG enables both
308 qualitative and quantitative monitoring in complex communities, like co-culture systems, fecal samples,
309 and host tissues. Therefore, these features make GUS-positive LGG a valuable tool for mechanistic
310 probiotic research and a promising chassis for therapeutic protein or antigen delivery in live
311 biotherapeutics.

312 In conclusion, we present a functional genome editing platform for LGG based on its
313 endogenous CRISPR-Cas9 system. This system enables precise, scarless gene deletions and insertions
314 using a single plasmid. The platform is rapid, convenient, and broadly applicable, making it a valuable
315 tool for LGG functional studies, probiotic engineering, and industrial strain development. The successful
316 engineering of GUS-expressing LGG strains highlights the potential of this system for live
317 biotherapeutic applications. In the future, this platform has the potential to broaden LGG's role as a
318 programmable chassis for therapeutic delivery and precision fermentation. By enabling efficient and
319 scarless genome editing, our platform paves the way for future research and strain development in LGG
320 and related *Lactobacillus* species, offering a powerful tool to enhance probiotic functionality and
321 industrial utility.

322

323 **MATERIAL AND METHODS**

324 **Strains and Culture Conditions**

325 All *Lcb. rhamnosus* strains (listed in Supplementary Table S1) were cultivated in de-Man-
326 Rogosa-Sharpe (MRS; Hardy Diagnostics, Santa Maria, CA, USA) broth or on MRS agar (1.5%) at
327 37 °C under anaerobic conditions (90% N_2, 5% CO_2, and 5% H_2). *Escherichia coli* MC1061 was
328 cultured in Luria-Bertani (LB) broth with agitation at 250 rpm or on LB agar (1.5%) and incubated at
329 37 °C. For GUS activity detection, MRS agar plates were supplemented with 200 µg/mL 5-Bromo-4-
330 chloro-1*H*-indol-3-yl β-D-glucopyranosiduronic acid (X-gluc; Gold Biotechnology, USA). When
331 appropriate, antibiotics were added at the following concentrations: erythromycin (Em; Gold
332 Biotechnology, St. Louis, MO, USA) at 5 µg/ml for LGG and 150 µg/ml for *E. coli*, and
333 chloramphenicol (Cm; Sigma, St. Louis, MO, USA) at 5 µg/ml for LGG.

334 **CRISPR-Cas System Characterization *in silico***

335 The complete genome sequence of LGG (GenBank: CP031290.1) is available at the National
336 Center for Biotechnology Information (NCBI) database. CRISPR array and associated Cas proteins were
337 identified using the CRISPRCasFinder platform (Couvin et al., 2018). PAM sequences were predicted
338 by aligning LGG spacer to target sequences with CRISPRTarget (Biswas et al., 2013), and consensus
339 PAM sequences were presented as sequence logo generated by WebLogo (Crooks et al., 2004). For type
340 II CRISPR systems, tracrRNA contains a region complementary to the repeat sequence of the crRNA;
341 thus, the repeat sequence was queried against the genome using the Basic Local Alignment Search Tool
342 (BLAST) to locate the tracrRNA (Hidalgo-Cantabrana, Goh, & Barrangou, 2019). The crRNA–
343 tracrRNA duplex was further modeled using the NCPACK web server and manually illustrated (Zadeh
344 et al., 2011).

345 **DNA Manipulations and Molecular Cloning**

346 Genomic DNA was isolated using the Wizard genomic DNA purification kit (Promega
347 Corporation, Madison, WI, USA). Plasmid DNA from *E. coli* was prepared with the QIAGEN plasmid
348 mini-prep kit (QIAGEN, Germantown, MD, USA). Restriction enzymes and T4 DNA ligase (New
349 England Biolabs, Ipswich, MA, USA) were used according to the supplier's protocols. Polymerase chain
350 reactions (PCR) were conducted following standard procedures with either DreamTaq DNA polymerase
351 (Thermo Fisher Scientific, Waltham, MA, USA) or Phusion DNA polymerase (NEB, Ipswich, MA,
352 USA). Oligonucleotide primers (Supplementary Table S3) were ordered from Integrated DNA
353 Technologies (IDT, Coralville, IA, USA) and synthetic gene fragments (Supplementary Table S4) were
354 obtained from Twist Bioscience (Twist Bioscience, San Francisco, CA, USA). PCR amplicons were
355 electrophoresed on 1.0% agarose gels and purified using the QIAquick PCR purification kit (QIAGEN,
356 Germantown, MD, USA). Sanger sequencing was performed at the UIUC Core Sequencing Facility
357 (Urbana, IL, USA), and whole plasmid sequencing was conducted by Plasmidsaurus, an Oxford
358 Nanopore certified service.

359 **Construction of Interference Plasmids**

360 pTRK830 is a replicating shuttle vector for *E. coli* and lactobacilli strains (Xie et al., 2025). The
361 multiple cloning site (MCS) from pBluescript KS (+) was ligated into the *StuI* site of pTRK830, which
362 generated pTRK870 (Supplementary Table S2). Plasmid pTRK870 was used for all following plasmid
363 constructions. The interference plasmids were constructed as described before (Crawley et al., 2018).
364 The double-stranded DNA protospacers, with or without the PAM, were generated by PCR using and
365 then ligated into XbaI-HindIII-digested pTRK870 to construct interference plasmids. The resulting
366 interference plasmids were PCR-screened in *E. coli* transformants using PS1 and PS2 primers
367 (Supplementary Table S3) for the presence of the insert and sequenced to confirm sequence content.

368 **Construction of Editing Plasmids**

369 Each editing plasmid consisted of two key components: a sgRNA expression cassette and a
370 repair template. The sgRNA expression cassettes included a promoter *Ppgm* of the phosphoglycerate
371 mutase from *L. acidophilus* NCFM, a designed sgRNA containing a 30-bp sequence targeting a specific
372 genomic locus of LGG strain, and a Rho-independent terminator.

373 ***fucI* gene deletion**: The *fucI* targeting sgRNA expression cassette (Supplementary Table S4) was
374 synthesized and amplified using primers sgRNA-XbaI and sgRNA-XmaI (Supplementary Table S3),
375 then cloned into XbaI-XmaI-digested pTRK870 to generate targeting plasmid pfucI-sgRNA
376 (Supplementary Table S2). The repair template with two 1-kb homologous regions for *fucI* gene editing
377 was amplified from the genomic DNA of LGG strain. For this, two 1-kb homologous regions were fused
378 by overlap extension PCR to obtain the 2-kb repair template to the *fucI* gene. This repair template was
379 digested with SalI and ApaI, then cloned into SalI-ApaI-digested plasmids pfucI-sgRNA and pTRK870
380 to generate plasmids pfucI-sgRNA-RT and pfucI-RT (Supplementary Table S2), respectively. All these
381 final plasmid constructs were PCR-screened using PS1 and PS2 primers (pTRK870 based) and
382 sequenced to confirm sequence content.

383 ***acs1* gene deletion**: The *acs1* targeting sgRNA expression cassettes (Supplementary Table S4)
384 were synthesized and amplified using primers sgRNA-XbaI and sgRNA-XmaI (Supplementary Table
385 S3), then cloned into XbaI-XmaI-digested pTRK870 to generate targeting plasmids pacs1-sgRNA
386 (Supplementary Table S2). The repair template (Supplementary Table S4) with two 1-kb homologous
387 regions was synthesized and cloned into the PstI-ApaI-digested pacs1-sgRNA generating pacs1-sgRNA-
388 RT (Supplementary Table S2). All these final plasmid constructs were PCR-screened using PS1 and PS2
389 primers (pTRK870 based) and sequenced to confirm sequence content.

390 ***cam* expression cassette insertion:** To construct a compact and stable cloning vector for gene
391 insertion, a multiple cloning site sequence was inserted into StuI-MfeI-digested pTRK829

392 (Supplementary Table S2), removing the entire *cam* gene and other non-essential regions, to generate
393 pMJM100 (Supplementary Figure S2). The *fucI* targeting sgRNA expression cassette (Supplementary
394 Table S4) was cloned into XbaI-XmaI-digested pMJM100 to construct a new *fucI* targeting plasmid
395 pMJM100-fucI-sgRNA (Supplementary Table S2). The repair template, containing two 0.5-kb
396 homologous regions, was amplified from pfucI-RT using primers fucI-RT2-SalI and fucI-RT2-ApaI
397 (Table S3), and cloned into a TOPO vector using the Zero Blunt® TOPO® PCR Cloning Kit (Thermo
398 Fisher Scientific, Waltham, MA, USA), resulting in pTOPO-fucI-RT2 (Supplementary Table S2). The
399 *cam* gene expression cassette was amplified from pTRK829 using primers Cm-F and Cm-R and inserted
400 into pTOPO-fucI-RT2 at the *HindIII* site, generating pTOPO-fucI-Cm-RT2 (Supplementary Table S2).
401 Finally, the complete repair template, containing the *cam* gene expression cassette flanked by two 0.5-kb
402 homologous regions, was cloned into SalI-ApaI-digested pMJM100-fucI-sgRNA, resulting in the final
403 editing plasmid pfucI-sgRNA-Cm-RT2 (Supplementary Table S2). pfucI-sgRNA-Cm-RT2 was PCR-
404 screened using PS3 and PS4 primers (pMJM100 based) and sequenced to confirm sequence content.

405 ***gusA* expression cassette insertion**: The sgRNA expression cassette targeting to intergenic
406 region between fucose and rhamnose operons (Supplementary Table S4) was synthesized and amplified
407 using primers sgRNA-XbaI and sgRNA-XmaI (Supplementary Table S3), then cloned into XbaI-XmaI-
408 digested pMJM100 to generate targeting plasmid pMJM100-rha-sgRNA (Supplementary Table S2). The
409 repair template (Supplementary Table S4) with two 0.5-kb homologous regions was synthesized and
410 cloned into the SalI-ApaI-digested pMJM100-rha-sgRNA generating plasmid prha-sgRNA-RT2
411 (Supplementary Table S2). The *gusA* gene expression cassette was amplified from *L. gasseri* ADH using
412 primers gusA-SacII-F and gusA-SacII-R and inserted into prha-sgRNA-RT2 at the *SacII* site, generating
413 prha-sgRNA-gusA-RT2 (Supplementary Table S2). All these final plasmid constructs were PCR-
414 screened using PS3 and PS4 primers (pMJM100 based) and sequenced to confirm sequence content.

415 **Transformation of *Lacticaseibacillus* strains**

416 Electrotransformation of LGG and other *Lcb. rhamnosus* strains was performed following our
417 previously described protocol (Xie et al., 2025), with minor modifications. Briefly, competent cells were
418 prepared as previously outlined and electroporated under identical conditions (2.5 kV, 25 μF, 200 Ω in 2-
419 mm cuvettes). The main adjustments included using 1 μg of plasmid DNA per 100 μL of cells and
420 decreasing the recovery volume to 900 μL of pre-warmed MRS broth. All other steps, including cell
421 preparation, washing, and post-electroporation plating, were carried out as reported previously.

422 Transformants were screened by PCR and confirmed by Sanger sequencing using gene-specific
423 primers (Supplementary Table S3). For the fucose isomerase gene *fucI*, the primers fucI-check-F and
424 fucI-check-R were used for the chromosomal PCR amplification and sequencing (2.7-kb in the WT and
425 2.1-kb in deletion mutant). For the CoA synthase gene *acs1*, the primers acs1-check-F and acs1-check-R
426 were used for the chromosomal PCR amplification and sequencing (3.6-kb in the WT and 2.1-kb in
427 deletion mutant). For the insertion of *cam* expression cassette, the primers fucI-check-F and fucI-check-
428 R were used for the chromosomal PCR amplification and sequencing (2.7-kb in the WT and 3.2-kb in
429 insertion mutant). For the insertion of *gusA* expression cassette, the primers fuc-rha-check-F and fuc-
430 rha-check-R were used for the chromosomal PCR amplification and sequencing (2.1-kb in the WT and
431 4.2-kb in insertion mutant).

432 **Growth Curve Measurement**

433 The growth of the LGG strain and its mutant *L. rhamnosus* MJM568 and MJM670 was analyzed
434 in sugar-free MRS media (US Biological, Swampscott, MA, USA) containing 2% glucose, fucose, or
435 rhamnose (Sigma, St. Louis, MO, USA). Strains were picked and cultured in MRS broth overnight.
436 Then the cultures were harvested and washed three times with phosphate-buffer solution (PBS, pH 7.4)

437 and resuspended in same volume of PBS buffer. Cells were inoculated 2% into MRS containing
438 different sugar sources and transferred into a sterile 96-well plate. Each condition was tested in
439 triplicate. The plates were incubated at 37 °C without agitation for 24 h in FilterMax F5 Multi-Mode
440 Microplate Reader (Molecular Devices, LLC, Sunnyvale, CA, USA). The absorbance at 595nm
441 (Abs_{595nm}) in each well was recorded every 20 or 30 min. The maximum specific growth rate (μ_{max}) was
442 defined as the slope of the tangent line at the inflection point of the logistic regression curve and was
443 reported in $hours^{-1}$.

444 **Assessment of *gusA* Gene Expression Cassette Stability**

445 To evaluate the genetic stability of the *gusA* gene, both plasmid-based (LGG with pMJM100-
446 gusA) and genome-integrated (MJM670) constructs were examined over multiple generations. The *gusA*
447 gene expression cassette was amplified from *L. gasseri* ADH using primers gusA-SpeI-F and gusA-
448 SpeI-R and inserted into pMJM100 at the *SpeI* site, generating pMJM100-gusA (Supplementary Table
449 S2). LGG harboring pMJM100-gusA was cultured 12 hours in MRS broth supplemented with Em (0
450 generation), while MJM670 was grown 12 hours in MRS broth without antibiotics (0 generation). Then,
451 0.1% of each overnight culture was then inoculated into 10 mL of fresh MRS broth without antibiotics.
452 This process was repeated every 12 hours for a total of 2 days (~40 generations). At defined intervals
453 (e.g., 0, 10, 20, 30, and 40 generations), cultures were serially diluted in PBS and plated on MRS agar
454 supplemented with 200 µg/mL X-gluc. The proportion of blue (GUS-expressing) and white (non-
455 expressing) colonies was quantified to assess the retention of *gusA* expression.

456 **In vivo Mice Study**

457 All animal procedures were approved by the University of Illinois Institutional Animal Care and
458 Use Committee (protocol #23108). Male C57BL/6J mice (8 weeks old, n = 18; Jackson Laboratories,

459 Bar Harbor, ME, USA) were acclimated for 4 days and then randomly assigned to three groups (n = 6
460 each). Mice were housed individually on wire grids with free access to water and an AIN-93M diet
461 (Research Diets, Inc., New Brunswick, NJ, USA) under a 12-h light/dark cycle. The three experimental
462 groups of mice were administered a single 0.2 mL oral gavage containing either 0.85% NaCl (control),
463 WT LGG (GUS-, 4.6×10^9 CFU/mL), or MJM670 (GUS+, 3.2×10^9 CFU/mL). Fecal pellets were
464 collected at designated time points and placed into sterile microcentrifuge tubes. For plating analysis,
465 fecal pellets were vortexed vigorously in sterile PBS, followed by serial dilution. Aliquots from each
466 dilution were plated on MRS agar supplemented with X-Gluc and aerobically incubated at 37 °C for
467 colony counting. Plates with colony counts between 25 and 250 were considered within the acceptable
468 range for enumeration and counts outside of this range were excluded from analysis. CFU per gram of
469 feces was calculated based on the dilution factor. Due to the lowest countable dilution being 1:100, the
470 detection limit for this assay was estimated at 2.5×10^4 CFU/g feces.

471 **Quantitative PCR Analysis**

472 Total DNA was extracted using the PureLink™ Microbiome DNA Purification Kit (Invitrogen,
473 Thermo Fisher Scientific, Waltham, MA, USA) according to the manufacturer's instructions. DNA
474 concentration was measured with a Nanodrop spectrophotometer (Thermo Scientific, Waltham, MA,
475 USA). LGG was detected using strain-specific primers Lrhamn1 and Lrhamn2 (Brandt & Alatossava,
476 2003; L. Zhang et al., 2020). qPCR was carried out using our previous reported protocol (Lei et al.,
477 2024) on a Q instrument (Quantabio, Beverly, MA, USA). Reactions (10 μL) contained PowerUp SYBR
478 Green Master Mix (Thermo Fisher Scientific, Waltham, MA, USA), LGG-specific primers, and
479 extracted DNA in the same proportions as described earlier. Cycling conditions included an initial
480 incubation (50 °C, 2 min), denaturation (95 °C, 2 min), and 40 amplification cycles (95 °C for 15 s,
481 54.5 °C for 15 s, 72 °C for 1 min). Each sample was analyzed in triplicate. Standard curves were

482 generated using serial dilutions of known concentrations of LGG genomic DNA, and results were
483 expressed as colony-forming units per gram of feces (CFU/g). For samples with *Ct* values greater than
484 35, a fixed *Ct* value of 35—corresponding to the detection limit of 1.1×10^4 CFU/g feces—was assigned
485 to facilitate consistent data analysis.

486 **Statistical Analysis**

487 Results are presented as mean values ± standard deviation from independent experiments.
488 Differences between groups were evaluated using unpaired *t*-tests and one-way ANOVA, with statistical
489 significance defined as $p < 0.05$. In figures, distinct letters were used to indicate significant different
490 groups. All analyses and graph preparation were performed using GraphPad Prism (GraphPad Software,
491 Boston, MA, USA).

492

493 **ACKNOWLEDGEMENTS**

494 Zifan Xie thanks support from the Investment for Growth Program from the University of
495 Illinois, Synthetic Biology for Food and Nutrition Innovation (SynFONI). We would like to express our
496 deep appreciation to Molly Black for her invaluable assistance and unwavering support during the mice
497 experiment, which significantly contributed to the success of our research.

498

499 **DATA AVAILABILITY STATEMENT**

500 The authors confirm that the data supporting the findings of this study are available within the
501 article and its supplementary materials.

502

FIGURES

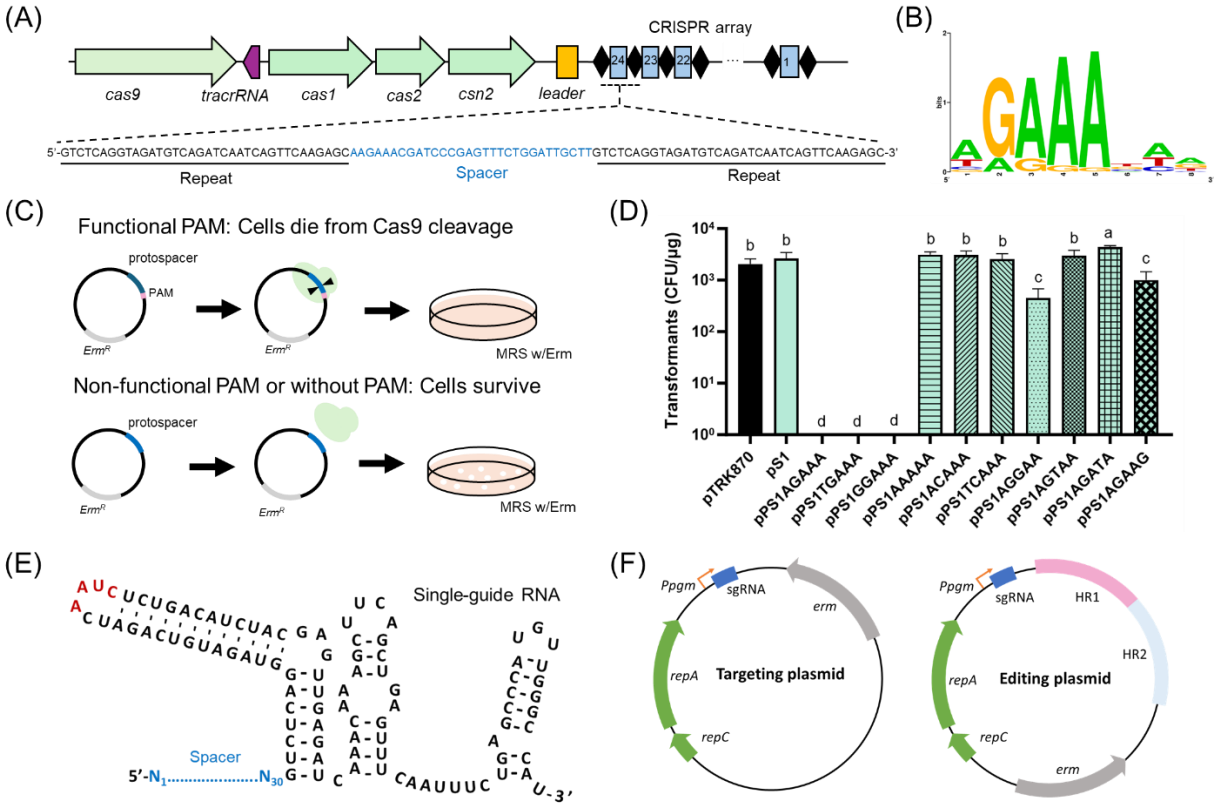


Figure 1. Genome-editing platform based on the endogenous type II-A CRISPR–Cas9 system in LGG. (A) Architecture of the CRISPR-Cas locus in LGG. The type II-A CRISPR-Cas9 system contains four Cas protein-encoding genes (*cas1, cas2, csn2*, and the critical *cas9*). The CRISPR array includes 24 spacers (30 bp, blue rectangles) interspaced with repeats (black diamonds). (B) PAM prediction visualized using a frequency plot generated using WebLogo. (C) Schematic of the plasmid interference assay. Plasmids carrying target sequences with functional or mutated PAMs were introduced into LGG to assess CRISPR-Cas9-mediated interference. A reduction in colony recovery indicates effective targeting by the native CRISPR-Cas9 system. (D) Recovery of transformants following introduction of interference plasmids compared to the control plasmid pTRK870. CFU: colony forming unit. Distinct letters were employed to denote significant differences ($p < 0.05$). (E) Structure of the SgRNA, designed by linking the crRNA and tracrRNA with an artificial AAUC loop (red). (F) Schematic of targeting and editing plasmids. *repA* and *repC*, plasmid replication gene; *erm*, erythromycin resistance gene; HR1 and HR2, the homologous repair templates; Ppgm, the promoter of phosphoglycerate mutase (*Ppgm*) from *L. acidophilus* NCFM.

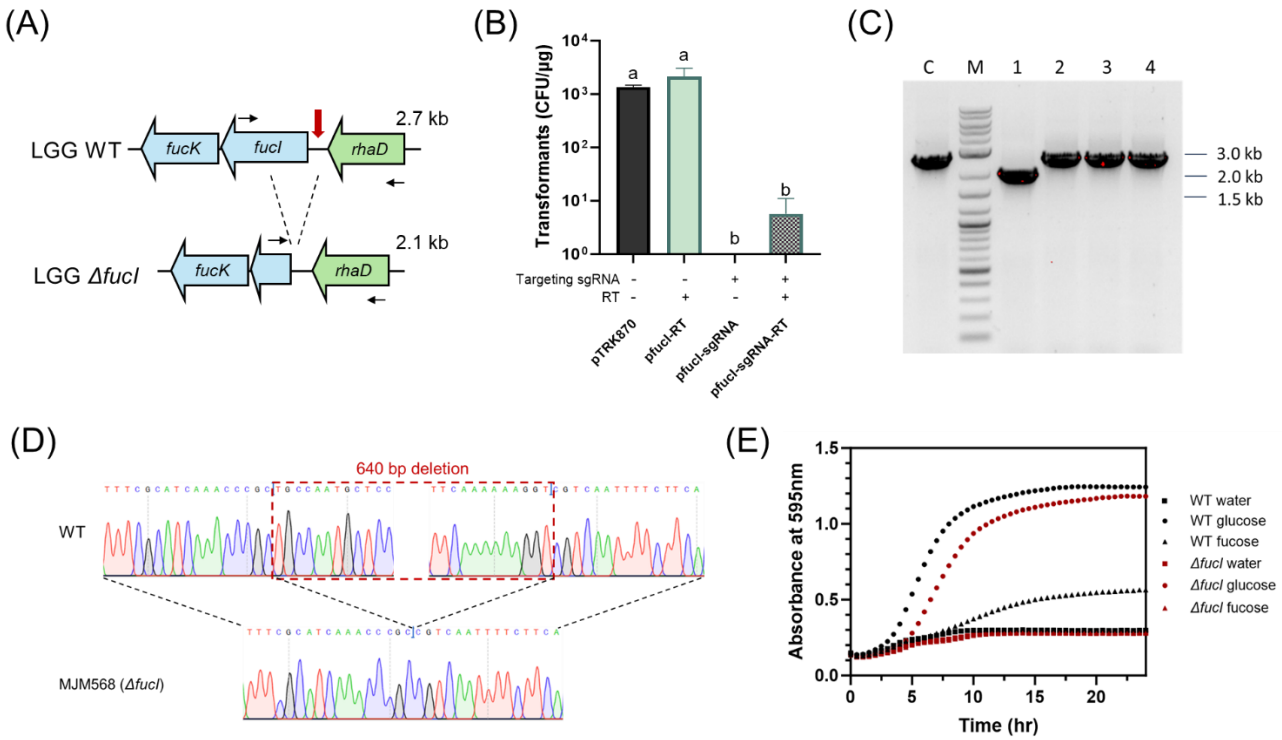


Figure 2. Deletion of L-fucose isomerase (*fucI*) by reprogramming the endogenous CRISPR-Cas9 system in LGG. (A) Schematic representation of the genotypes of the wild type (WT) strain and the *ΔfucI* mutant. (B) Transformation efficiencies of control plasmid pTRK870, homologous repair template containing plasmid pfucI-RT, self-targeting plasmid pfucI-sgRNA, and editing plasmid pfucI-sgRNA-RT. Distinct letters were employed to denote significant differences ($p < 0.05$). (C) Colony PCR screening of four surviving transformants carrying the editing plasmid pfucI-sgRNA-RT to confirm the deletion genotype. Expected PCR product sizes for WT and *ΔfucI* are 2.7 kb and 2.1 kb, respectively. C: WT control; M: DNA marker; 1-4: surviving transformants. (D) Sanger sequencing results confirming the 640-bp deletion in the Δ*fucI* strain (MJM568) compared to WT. (E) Growth curve of WT (black) and MJM568 (red) strains in MRS broth containing water (square), 2% glucose (circle) or 2% fucose (triangle), measured by absorbance at 595nm.

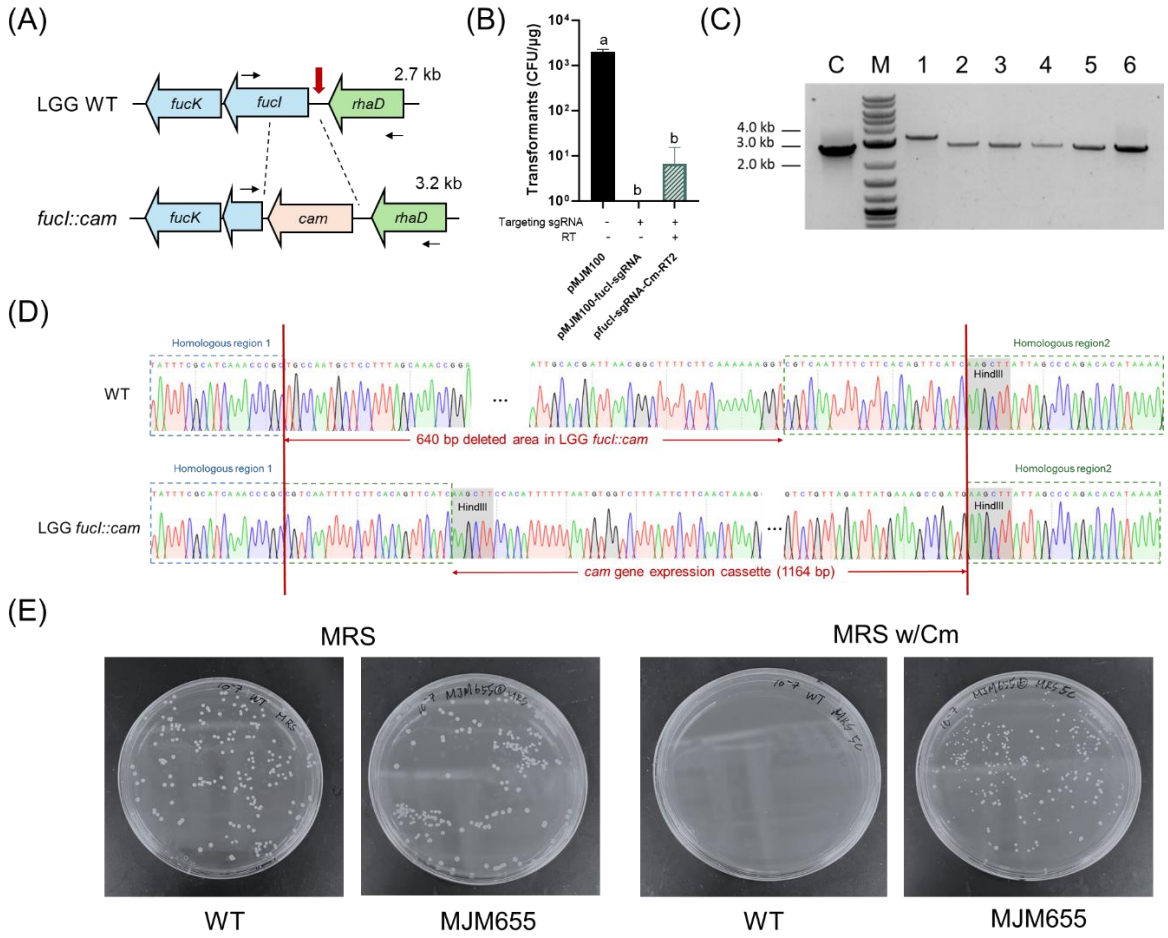


Figure 3. Insertion of the chloramphenicol resistance (*cam*) gene by reprogramming the endogenous CRISPR-Cas9 system in LGG. (A) Schematic representation of the genotypes of the wild type (WT) strain and the LGG *fucI::cam* mutant. (B) Transformation efficiencies of control plasmid pMJM100, self-targeting plasmid pMJM100-fucI-sgRNA, and editing plasmid pfucI-sgRNA-Cm-RT2. Distinct letters were employed to denote significant differences ($p < 0.05$). (C) Colony PCR screening of six surviving transformants carrying the editing plasmid pfucI-sgRNA-Cm-RT2 to confirm the insertion genotype. Expected PCR product sizes for WT and *fucI::cam* are 2.7 kb and 3.2 kb, respectively. C: WT control; M: DNA marker; 1-6: surviving transformants. (D) Sanger sequencing results of the WT and LGG *fucI::cam* strains. In *fucI::cam*, a 1164-bp *cam* gene expression cassette was inserted into the genome, replacing a 640-bp segment of the *fucI* gene. Homologous region 1 is enclosed in a dashed blue box, and homologous region 2 is enclosed in a dashed green box. The *HindIII* restriction site, located at the junction, is shaded in grey. (E) Growth of WT LGG and MJM655 strains on MRS plates or MRS plates supplemented with 5 μg/mL chloramphenicol.

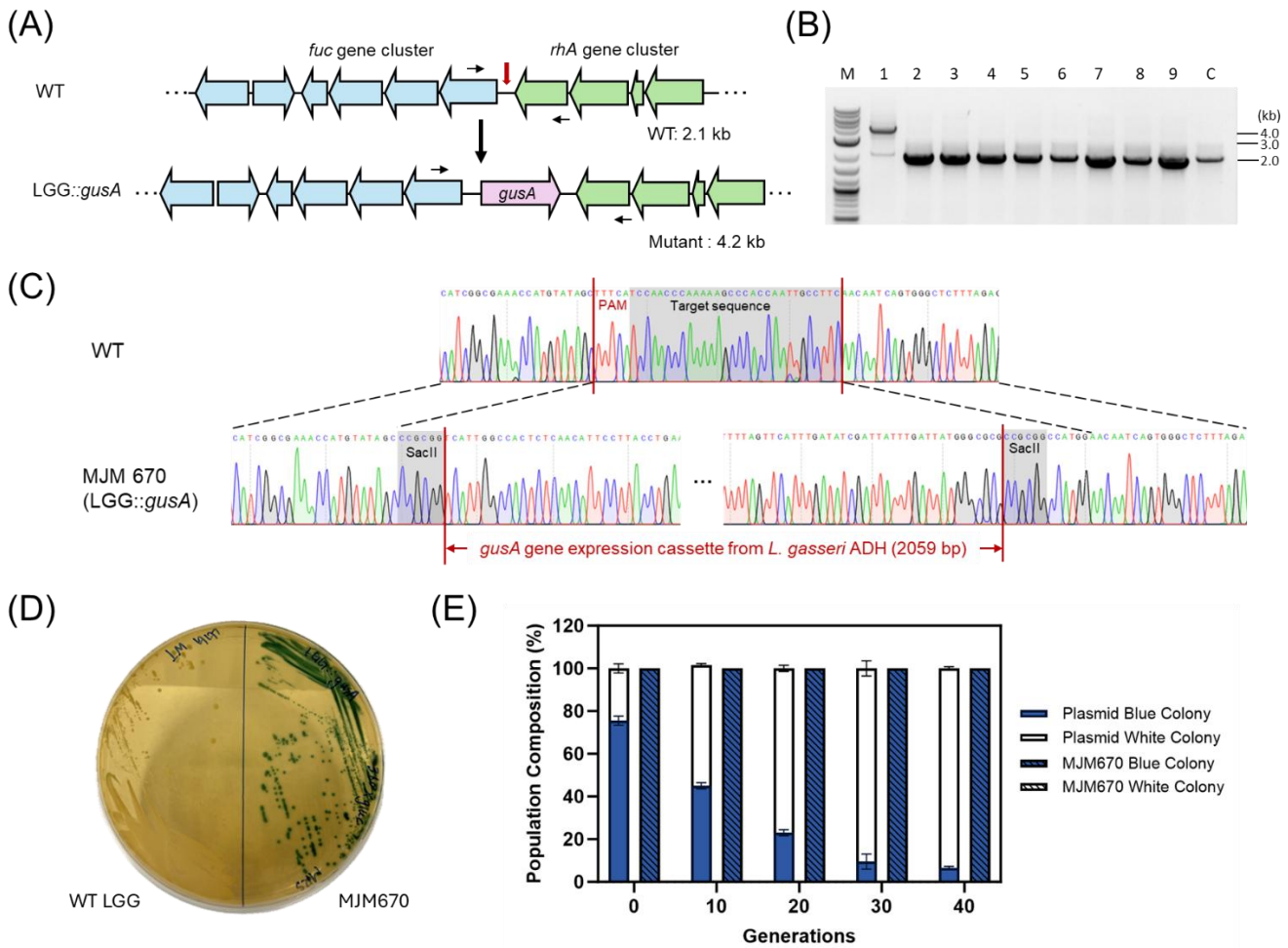


Figure 4. Insertion of the β-glucuronidase encoding gene *gusA* by reprogramming the endogenous CRISPR-Cas9 system in LGG. (A) Schematic representation of the genomic locus in the wild type (WT) strain and the LGG::*gusA* mutant, showing insertion of the *gusA* expression cassette from *Lactobaillus gasseri* ADH between the fucose and rhamnose operons. (B) Colony PCR screening of nine surviving transformants carrying the editing plasmid prha-sgRNA-gusA-RT2 to confirm the insertion genotype. Expected PCR product sizes for WT and LGG::*gusA* are 2.1 kb and 4.2 kb, respectively. C: WT control; M: DNA marker; 1-9: surviving transformants. (C) Sanger sequencing results confirming precise *gusA* insertion in the mutant strain MJM670 (LGG::*gusA*) compared to WT. (D) Growth phenotype of WT LGG and MJM670 on MRS agar supplemented with X-gluc. MJM670 exhibits a blue colony phenotype indicative of β-glucuronidase activity, while WT colonies remain white. (E) Stability of *gusA* gene expression across generations for plasmid-based and genome-integrated constructs.

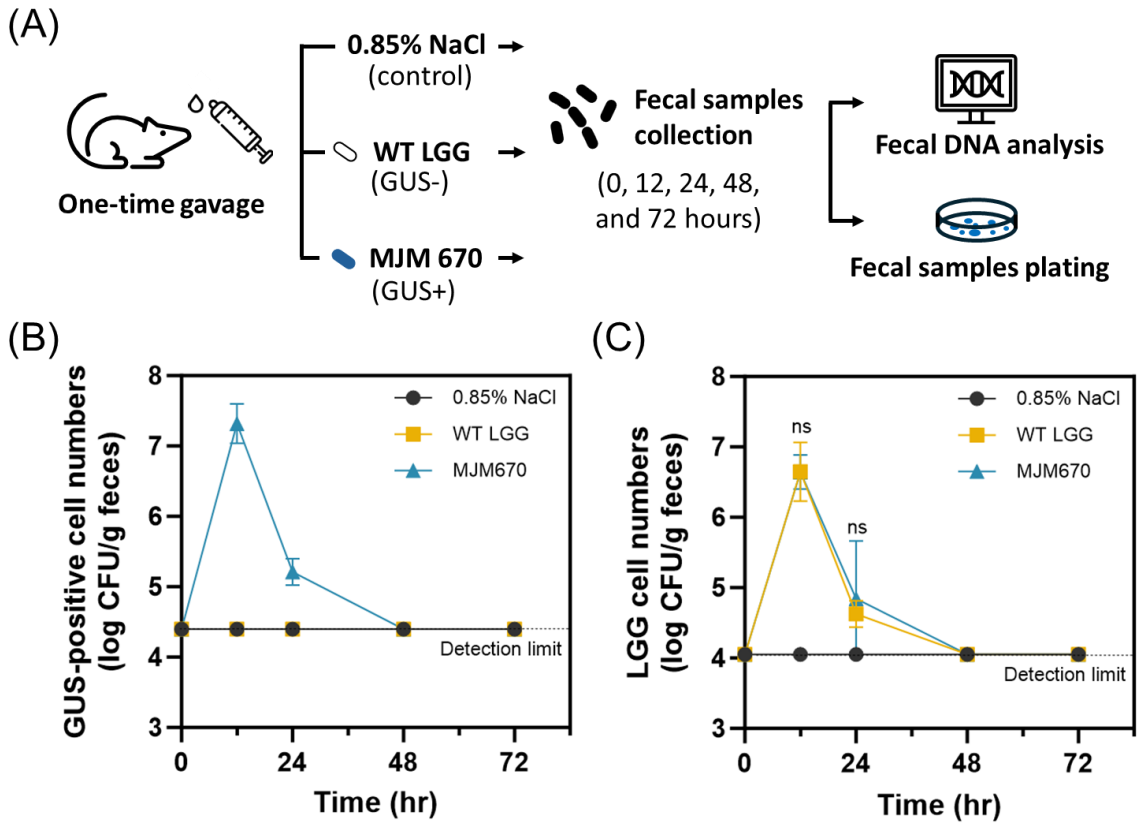


Figure 5. ***In vivo*** **tracking and quantification of GUS-positive LGG in mice.** (A) Mice were orally gavaged with saline, WT LGG (GUS- LGG), or MJM670 (GUS+ LGG). Fecal samples were collected at 0, 12-, 24-, 48-, and 72-hours post-gavage. (B) MJM670 was quantified by plating on MRS + X-gluc agar to enumerate blue colonies (viable GUS+ LGG). (C) In parallel, total LGG abundance was assessed in all groups by LGG-specific qPCR. Data represents mean ± SD from biological replicates (n=6).

563 **REFERENCES**

564 Adli, M. (2018). The CRISPR tool kit for genome editing and beyond. *Nature Communications*, *9*(1),
565 1911. https://doi.org/10.1038/s41467-018-04252-2

566 Becerra, J. E., Yebra, M. J., & Monedero, V. (2015). An l-Fucose Operon in the Probiotic Lactobacillus
567 rhamnosus GG Is Involved in Adaptation to Gastrointestinal Conditions. *Applied and*
568 *Environmental Microbiology*, *81*(11), 3880–3888. https://doi.org/10.1128/AEM.00260-15

569 Biswas, A., Gagnon, J. N., Brouns, S. J. J., Fineran, P. C., & Brown, C. M. (2013). CRISPRTarget. *RNA*
570 *Biology*, *10*(5), 817–827. https://doi.org/10.4161/rna.24046

571 Brandt, K., & Alatossava, T. (2003). Specific identification of certain probiotic *Lactobacillus rhamnosus*
572 strains with PCR primers based on phage-related sequences. *International Journal of Food*
573 *Microbiology*, *84*(2), 189–196. https://doi.org/10.1016/S0168-1605(02)00419-1

574 Capurso, L. (2019). Thirty Years of Lactobacillus rhamnosus GG: A Review. *Journal of Clinical*
575 *Gastroenterology*, *53*, S1. https://doi.org/10.1097/MCG.0000000000001170

576 Chan, M. Z. A., Toh, M., & Liu, S.-Q. (2021). Growth, survival, and metabolic activities of probiotics
577 *Lactobacillus rhamnosus* GG and *Saccharomyces cerevisiae* var. *Boulardii* CNCM-I745 in
578 fermented coffee brews. *International Journal of Food Microbiology*, *350*, 109229.
579 https://doi.org/10.1016/j.ijfoodmicro.2021.109229

580 Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P. D., Wu, X., Jiang, W., Marraffini,
581 L. A., & Zhang, F. (2013). Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science*,
582 *339*(6121), 819–823. https://doi.org/10.1126/science.1231143

583 Couvin, D., Bernheim, A., Toffano-Nioche, C., Touchon, M., Michalik, J., Néron, B., Rocha, E. P. C.,
584 Vergnaud, G., Gautheret, D., & Pourcel, C. (2018). CRISPRCasFinder, an update of

585 CRISRFinder, includes a portable version, enhanced performance and integrates search for Cas
586 proteins. *Nucleic Acids Research*, *46*(W1), W246–W251. https://doi.org/10.1093/nar/gky425
587 Crawley, A. B., Henriksen, E. D., Stout, E., Brandt, K., & Barrangou, R. (2018). Characterizing the
588 activity of abundant, diverse and active CRISPR-Cas systems in lactobacilli. *Scientific Reports*,
589 *8*(1), Article 1. https://doi.org/10.1038/s41598-018-29746-3
590 Crooks, G. E., Hon, G., Chandonia, J.-M., & Brenner, S. E. (2004). WebLogo: A Sequence Logo
591 Generator. *Genome Research*, *14*(6), 1188–1190. https://doi.org/10.1101/gr.849004
592 Gao, J., Li, Y., Wan, Y., Hu, T., Liu, L., Yang, S., Gong, Z., Zeng, Q., Wei, Y., Yang, W., Zeng, Z., He,
593 X., Huang, S.-H., & Cao, H. (2019). A Novel Postbiotic From Lactobacillus rhamnosus GG With
594 a Beneficial Effect on Intestinal Barrier Function. *Frontiers in Microbiology*, *10*.
595 https://doi.org/10.3389/fmicb.2019.00477
596 Goh, Y. J., & Barrangou, R. (2021). Portable CRISPR-Cas9N System for Flexible Genome Engineering
597 in Lactobacillus acidophilus, Lactobacillus gasseri, and Lactobacillus paracasei. *Applied and*
598 *Environmental Microbiology*, *87*(6). https://doi.org/10.1128/AEM.02669-20
599 Gu, S., Zhang, J., Li, L., & Zhong, J. (2022). Repurposing the Endogenous CRISPR-Cas9 System for
600 High-Efficiency Genome Editing in Lacticaseibacillus paracasei. *ACS Synthetic Biology*, *11*(12),
601 4031–4042. https://doi.org/10.1021/acssynbio.2c00374
602 Günaydın, G., Zhang, R., Hammarström, L., & Marcotte, H. (2014). Engineered *Lactobacillus*
603 *rhamnosus* GG expressing IgG-binding domains of protein G: Capture of hyperimmune bovine
604 colostrum antibodies and protection against diarrhea in a mouse pup rotavirus infection model.
605 *Vaccine*, *32*(4), 470–477. https://doi.org/10.1016/j.vaccine.2013.11.057

606 Hidalgo-Cantabrana, C., Goh, Y. J., & Barrangou, R. (2019). Characterization and Repurposing of Type
607 I and Type II CRISPR–Cas Systems in Bacteria. *Journal of Molecular Biology*, *431*(1), 21–33.
608 https://doi.org/10.1016/j.jmb.2018.09.013
609 Hidalgo-Cantabrana, C., Goh, Y. J., Pan, M., Sanozky-Dawes, R., & Barrangou, R. (2019). Genome
610 editing using the endogenous type I CRISPR-Cas system in *Lactobacillus crispatus*. *Proceedings*
611 *of the National Academy of Sciences*, *116*(32), 15774–15783.
612 https://doi.org/10.1073/pnas.1905421116
613 Hsu, P. D., Lander, E. S., & Zhang, F. (2014). Development and Applications of CRISPR-Cas9 for
614 Genome Engineering. *Cell*, *157*(6), 1262–1278. https://doi.org/10.1016/j.cell.2014.05.010
615 Huang, H., Song, X., & Yang, S. (2019). Development of a RecE/T-Assisted CRISPR–Cas9 Toolbox for
616 Lactobacillus. *Biotechnology Journal*, *14*(7), 1800690. https://doi.org/10.1002/biot.201800690
617 Hudault, S., Liévin, V., Bernet-Camard, M. F., & Servin, A. L. (1997). Antagonistic activity exerted in
618 vitro and in vivo by Lactobacillus casei (strain GG) against Salmonella typhimurium C5
619 infection. *Applied and Environmental Microbiology*, *63*(2), 513–518.
620 https://doi.org/10.1128/aem.63.2.513-518.1997
621 Jia, R., Chen, H., Chen, H., & Ding, W. (2016). Effects of fermentation with Lactobacillus rhamnosus
622 GG on product quality and fatty acids of goat milk yogurt. *Journal of Dairy Science*, *99*(1), 221–
623 227. https://doi.org/10.3168/jds.2015-10114
624 Lebeer, S., Claes, I., Tytgat, H. L. P., Verhoeven, T. L. A., Marien, E., von Ossowski, I., Reunanen, J.,
625 Palva, A., de Vos, W. M., De Keersmaecker, S. C. J., & Vanderleyden, J. (2012). Functional
626 Analysis of Lactobacillus rhamnosus GG Pili in Relation to Adhesion and Immunomodulatory
627 Interactions with Intestinal Epithelial Cells. *Applied and Environmental Microbiology*, *78*(1),
628 185–193. https://doi.org/10.1128/AEM.06192-11

629 Leenay, R. T., Vento, J. M., Shah, M., Martino, M. E., Leulier, F., & Beisel, C. L. (2018). Genome
630 Editing with CRISPR-Cas9 in Lactobacillus plantarum Revealed That Editing Outcomes Can
631 Vary Across Strains and Between Methods. *Biotechnology Journal*, *14*(3), 1700583.
632 https://doi.org/10.1002/biot.201700583
633 Lei, Y., Xie, Z., Zhao, A., Colarelli, J. P., Miller, M. J., & Lee, Y. (2024). Layer-by-layer coating of
634 *Lacticaseibacillus rhamnosus* GG (LGG) using chitosan and zein/tween-80/fucoidan
635 nanoparticles to enhance LGG's survival under adverse conditions. *Food Hydrocolloids*, *154*,
636 110039. https://doi.org/10.1016/j.foodhyd.2024.110039
637 Li, K., Li, B., Li, J., Wu, X., Zhao, Y., Yu, J., Guo, J., & Huang, C. (2025). Chairside live biotherapeutic
638 hydrogel for comprehensive periodontitis therapy. *Trends in Biotechnology*, *43*(2), 408–432.
639 https://doi.org/10.1016/j.tibtech.2024.10.001
640 Panahi, B., Dehganzad, B., & Nami, Y. (2023). CRISPR-Cas systems feature and targeting phages
641 diversity in Lacticaseibacillus rhamnosus strains. *Frontiers in Microbiology*, *14*.
642 https://doi.org/10.3389/fmicb.2023.1281307
643 Petrova, M. I., van den Broek, M. F. L., Spacova, I., Verhoeven, T. L. A., Balzarini, J., Vanderleyden, J.,
644 Schols, D., & Lebeer, S. (2018). Engineering Lactobacillus rhamnosus GG and GR-1 to express
645 HIV-inhibiting griffithsin. *International Journal of Antimicrobial Agents*, *52*(5), 599–607.
646 https://doi.org/10.1016/j.ijantimicag.2018.07.013
647 Segers, M. E., & Lebeer, S. (2014). Towards a better understanding of Lactobacillus rhamnosus GG -
648 host interactions. *Microbial Cell Factories*, *13*(Suppl 1), S7. https://doi.org/10.1186/1475-2859-
649 13-S1-S7
650 Sindhu, K. N. C., Sowmyanarayanan, T. V., Paul, A., Babji, S., Ajjampur, S. S. R., Priyadarshini, S.,
651 Sarkar, R., Balasubramanian, K. A., Wanke, C. A., Ward, H. D., & Kang, G. (2014). Immune

Response and Intestinal Permeability in Children With Acute Gastroenteritis Treated With Lactobacillus rhamnosus GG: A Randomized, Double-Blind, Placebo-Controlled Trial. *Clinical Infectious Diseases*, *58*(8), 1107–1115. https://doi.org/10.1093/cid/ciu065

Song, X., Huang, H., Xiong, Z., Ai, L., & Yang, S. (2017). CRISPR-Cas9D10A Nickase-Assisted Genome Editing in Lactobacillus casei. *Applied and Environmental Microbiology*, *83*(22). https://doi.org/10.1128/AEM.01259-17

Vento, J. M., Crook, N., & Beisel, C. L. (2019). Barriers to genome editing with CRISPR in bacteria. *Journal of Industrial Microbiology and Biotechnology*, *46*(9–10), 1327–1341. https://doi.org/10.1007/s10295-019-02195-1

Wan, M. L. Y., Turner, P. C., Allen, K. J., & El-Nezami, H. (2016). *Lactobacillus rhamnosus* GG modulates intestinal mucosal barrier and inflammation in mice following combined dietary exposure to deoxynivalenol and zearalenone. *Journal of Functional Foods*, *22*, 34–43. https://doi.org/10.1016/j.jff.2016.01.014

Xie, Z., Jin, Y.-S., Klaenhammer, T. R., & Miller, M. J. (2025). The insertion of the inverted repeat of an insertion sequence (IS) element from Lacticaseibacillus rhamnosus changes the host range and stability of pGK12, a shuttle vector for lactic acid bacteria. *Applied and Environmental Microbiology*, *91*(4), e01908-24. https://doi.org/10.1128/aem.01908-24

Xie, Z., McAuliffe, O., Jin, Y.-S., & Miller, M. J. (2024). Genomic Modifications of Lactic Acid Bacteria and Their Applications in Dairy Fermentation. *Journal of Dairy Science*. https://doi.org/10.3168/jds.2024-24989

Yan, F., & Polk, D. B. (2002). Probiotic Bacterium Prevents Cytokine-induced Apoptosis in Intestinal Epithelial Cells*. *Journal of Biological Chemistry*, *277*(52), 50959–50965. https://doi.org/10.1074/jbc.M207050200

Zadeh, J. N., Steenberg, C. D., Bois, J. S., Wolfe, B. R., Pierce, M. B., Khan, A. R., Dirks, R. M., & Pierce, N. A. (2011). NUPACK: Analysis and design of nucleic acid systems. *Journal of Computational Chemistry*, *32*(1), 170–173. https://doi.org/10.1002/jcc.21596

Zhang, J., Zong, W., Hong, W., Zhang, Z.-T., & Wang, Y. (2018). Exploiting endogenous CRISPR-Cas system for multiplex genome editing in Clostridium tyrobutyricum and engineer the strain for high-level butanol production. *Metabolic Engineering*, *47*, 49–59. https://doi.org/10.1016/j.ymben.2018.03.007

Zhang, L., Chichlowski, M., Gross, G., Holle, M. J., lbarra-Sánchez, L. A., Wang, S., & Miller, M. J. (2020). Milk Fat Globule Membrane Protects Lactobacillus rhamnosus GG from Bile Stress by Regulating Exopolysaccharide Production and Biofilm Formation. *Journal of Agricultural and Food Chemistry*, *68*(24), 6646–6655. https://doi.org/10.1021/acs.jafc.0c02267

Zhang, S., Oh, J.-H., Alexander, L. M., Özçam, M., & van Pijkeren, J.-P. (2018). D-Alanyl-d-Alanine Ligase as a Broad-Host-Range Counterselection Marker in Vancomycin-Resistant Lactic Acid Bacteria. *Journal of Bacteriology*, *200*(13), 10.1128/jb.00607-17. https://doi.org/10.1128/jb.00607-17

Zhong, Z., Liu, G., Tang, Z., Xiang, S., Yang, L., Huang, L., He, Y., Fan, T., Liu, S., Zheng, X., Zhang, T., Qi, Y., Huang, J., & Zhang, Y. (2023). Efficient plant genome engineering using a probiotic sourced CRISPR-Cas9 system. *Nature Communications*, *14*(1), 6102. https://doi.org/10.1038/s41467-023-41802-9