Evaluation of a Genome Editing Approach in Lactic Acid Bacteria Based on the CRISPR/Cas9 System

· Inês da Costa Gonçalves[1] ·

[1] Laboratories of Bioengineering Research Group (BERG), Instituto Superior Técnico, Av. Rovisco Pais, 1049-001 Lisbon, Portugal

Chairperson: Prof. Arsénio do Carmo Sales Mendes Fialho

Supervisor: Prof. Gabriel António Amaro Monteiro

Member of the Committee: Prof. Leonilde de Fátima Morais Moreira

Abstract DNA vaccination based on the delivery of plasmid DNA is a promising tool to treat, autoimmune, infectious and genetic-related diseases. The current favourite host for plasmid production is E. coli. However, the majority of pDNA extraction methods co-purify lipopolysaccharides (LPS) present in *E. coli* cell wall, which generate inflammatory responses in mammal hosts. Lactic acid bacteria (LAB) are LPS-free GRAS organisms, used in human health and nutrition from ancient years. The use of LAB as microbial cell factories for pDNA production is proposed. Hurdles associated with LAB for this mean include the lack of plasmids and features that allows them to do so. Therefore, the present work aims to genetically modify *Lactococcus lactis* LMG 19460, using the CRISPR/Cas9 system, to diminish harsh host terms for pDNA production. Applications of this improved strain include industrial production of pDNA and DNA delivery for gene therapy purposes. The previously designed nthCRISPRa plasmid, modified from pKCcas9dO, was unsuccessfully transformed by electroporation in L. lactis LMG 19460 to attempt nth knockout in a single step and a PCR reaction using 250,000 cells was developed for the assessment of both genome edition and nthCRISPRa presence. The reconstruction of nthCRISPRa into nthCRISPRe was attempted by Gibson assembly to change the selection marker from apmR to eryR. The CRISPR/Cas9 based editing plasmid containing the erythromycin resistant gene is yet to be achieved. Theoretical predictions were made for future advantageous gene knockouts. Furthermore, the OptFlux software was used to predict non-obvious knockouts that should increase DNA production.

Key Words Lactic Acid Bacteria, Gene Therapy, plasmid DNA, CRISPR/Cas9, PCR, Gibson Assembly.

Introduction

Lactic acid bacteria (LAB) are a non-taxonomic group of Gram-positive, nonsporing, low GC content, aerotolerant anaerobe and non-motile bacteria, in which are included different species such as *Lactobacillus* and *Lactococcus*.[1–3] Most of these bacteria are characterized by their capability to convert fermentable carbohydrates, mainly glucose, to lactic acid. These strictly fermentative bacteria have limited biosynthetic capabilities, especially amino acid, vitamin and nitrogenous bases production, thereby requiring complex growth media, which limits LAB natural niches to foodstuffs, plant surfaces and animal mucosal cavities. [4–6]
These bacteria have a long history of exploitation by humans, being used for centuries in food production and preservation and as probiotic agents to promote human health. Given their long history of food-grade applications and non-pathogenicity, LAB are classified as Generally Recognized As Safe (GRAS) by the United States Food and Drug Administration (FDA) and fulfil criteria of the Qualified Presumption of Safety (QPS) according to the European Food Safety Authority (EFSA).[7–9] LAB are also used as starter cultures to modify the organoleptic characteristics of foods..[9]
Although being the first-choice microorganism for the expression of heterologous systems, *E. coli* may present pyrogenic endotoxins and lipopolysaccharides (LPS) in its derived products, limiting the final application of its products, particularly those with pharmaceutical means.[10,11] Given these limitations, an increased number of scientists are considering Gram-positive bacteria as safer microorganisms for heterologous expression and shuttle systems.[10]
In general, LAB hold great promise in DNA and protein production as they constitute an endotoxin and LPS free platform and efficient delivery system for humans and animals.[12,13] The use of *L. lactis*, particularly, seems to be attractive for plasmid DNA (pDNA) delivery to epithelial membranes, not only due to already developed and established molecular tools, but also due to is GRAS, food-grade, transient and non-invasive features.[11] DNA vaccination mediated by LAB has proven to induce humoral and cellular immune responses.[6,11] This gene therapy approach consists in the transfection of a therapeutic vector from LAB to target cells, generically aiming the target gene addition, alteration, correction and/or knockdown or the stable expression of a therapeutic protein[6,14,15], aspiring the treatment of genetic-related disorders, the immunization promotion and the impairment of bacterial, viral and fungal infections[6,8,15]. This non-viral vector should circumvent some of the problems associated with customarily studied viral vector and attenuated pathogen delivery, including endogenous viral recombination, viral oncogenic effects, unexpected immune response and virulent state reversion.[7,16] The limited availability of LAB plasmids and theirs low-copy number nature has been one of the big hurdles in using LAB as pDNA and protein production hosts.[17] As such, these bacteria may benefit genome alterations in order to achieve higher plasmid production. Alongside genome alterations that improve pDNA production in *E. coli* are the *endA*, *recA*, *pgi* and *pyk* knockouts, genes which code for endonuclease I, bacterial DNA recombination protein A, phosphoglucose isomerase and pyruvate kinase, respectively.[18,19] It seems reasonable that the deletion of these genes' homologs in *L. lactis* could have the same impact on pDNA production, if orthology is confirmed for this feature.
Early methods for genome engineering were based on DNA double-strand break (DSB)-inducing protein nucleases, to persuade endogenous DNA repair machinery and promote edition in specific target sites.[20] CRISPR/Cas9 technology is currently into the spotlight of the genome-editing field as it only depends on an RNA moiety to target the Cas9 nuclease to the desired sequence, thus not requiring different protein structures for different edition events.[21–23] The recent unveilment of the molecular machinery of this

adaptive immune-derivative system allowed the development of engineering strategies for genetic modifications.[24,25] The type II CRISPR/Cas9-based genome engineering requires a Cas9 protein and a single-guide RNA (sgRNA) to induce the DSB.[26] A repair template with the desired modification is used to induce Homologous Recombination (HR) repair and, consequentially, genomic alteration. Other approaches take advantage of Non-Homologous End Joining (NHEJ) to disrupt genomic regions.[25]

The present study aims the attainment of *Lactococcus lactis* LMG 19460 capable of producing high quantity and high quality DNA, either for large scale DNA production and purification, or for alternative gene delivery purposes *in situ*. The knockout of the endonuclease III coding gene, *nth*, is attempted using a Cas9/sgRNA-based plasmid nthCRISPRa, according to Huang *et al.*[27]. The previously designed plasmid transformation in LAB is struggled by electroporation aspiring the posterior expression of the Cas9 system and induction of *nth* deletion by homology repair. The reconstruction of the nthCRISPRa into nthCRISPRe is attempted by Gibson assembly, aiming to change the plasmid selection marker from apramycin to the erythromycin resistance coding gene.

Materials and Methods

Bacterial Strains and Plasmids

Throughout the development of the present work the LAB *Lactococcus lactis* LMG 19460 and three *Escherichia coli* strains, DH5α, SCS 110 and GM 2163, were used.

Plasmids used include pTRKH3, an *E. coli*/LAB shuttle, with the erythromycin resistant marker; pKCcas9dO, used as vector by Huang *et al.*[27], designed for CRISPR-based genome editing, with the apramycin resistant marker; nthCRISPRa, designed from pKCcas9dO by Sofia Duarte (unpublished work) (Figure 1A); and nthCRISPRe designed from nthCRISPRa in the present work to change the selection marker to erythromycin (Figure 1B).

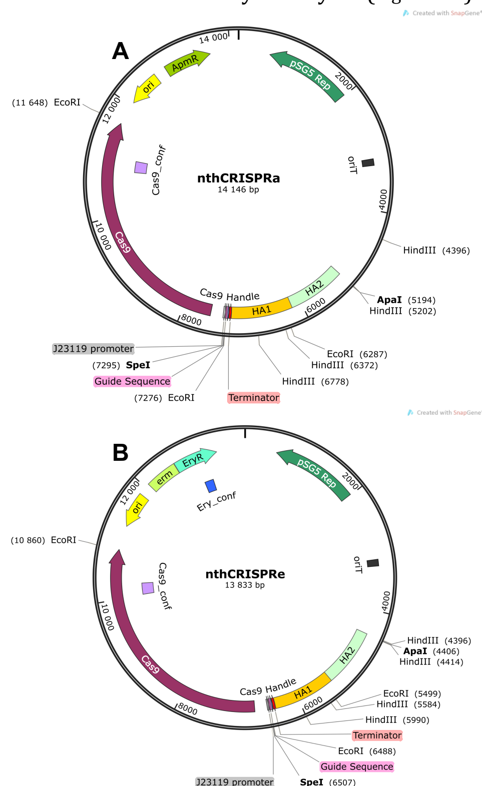

Figure 1 Schematic representation of plasmids nthCRISPRa (A) and nthCRISPRe (B) and its constituents.

Both plasmids (Figure 1) contain a Cas9 coding gene, a 20 bp antisense RNA sgRNA for *nth* gene targeting and two homology repair fragments, corresponding to the flanking regions of the *nth* gene, in *L. lactis* LMG 19460.

Growth conditions and transformation of *E. coli* DH5α

The *E. coli* bacterial strains were cultured in Luria Bertani (LB) medium (25 g/L) from NZYTech, supplemented with erythromycin 500 μg/mL, at 37ºC and 250 rpm agitation, when harbouring the pTRKH3 plasmid or with apramycin 25 μg/mL, at 30ºC and 250 rpm agitation, when harbouring the nthCRISPRa plasmid. Cells were grown in a pre-inoculum overnight, in 15 mL falcon tubes containing 5 mL of LB medium supplemented with antibiotic. After reaching the exponential growth phase cells were transferred to 15 mL falcon tubes containing 5 mL fresh LB medium supplemented with antibiotic, and incubated at the respective conditions until reaching an OD600nm of approximately 1 for cell banks and late exponential for pDNA purification.

For the preparation of chemically competent *E. coli* cells, an overnight growth in LB broth (25 g/L), at 37ºC, 250 rpm, was performed and cells were transferred to a 100 mL Erlenmeyer flask containing 20 mL LB broth (25 g/L), as to start the culture at an OD_{600nm} of 0.1. Cells were then grown as to reach an OD_{600nm} of 1 and were posteriorly centrifuged at 1,000 × g, for 10 minutes at 4ºC. Supernatants were discarded and the pellets were resuspended in 2 mL sterile (0.22 μm filter) TSS. After the preparation of 100 μL aliquots, cell preparations were kept on ice for 10 minutes, before storage at -80ºC. Chemically competent *E. coli* DH5α cells were transformed by heat shock. Each aliquot of competent cells was incubated with the desired DNA mass in the desired volume on ice for 30 minutes. The mixture was then placed on a 42ºC dry bath for 1 minute, and on ice for 2 minutes, immediately after. The cell mixture was then resuspended with 900 μL LB broth (25 g/L) and left for incubation at 30ºC, 250 rpm. Ultimately, these cells were plated in LB agar medium (40 g/L), from NZYTech, supplemented with erythromycin 500 μg/mL and incubated at 30ºC for 48h-60h. Exceptions regarding the antibiotic concentration are referred throughout the present work. Transformation candidates were grown as explained, until visible growth. Cell banks were made after cell growth reached an OD_{600nm} of ~1, if possible. Plasmid DNA was purified later after cells reached a higher OD_{600nm} value, if possible. Some transformation candidates did not exhibit an $OD_{600nm}>1$ and pDNA extraction was proceeded anyway.

E. coli DH5α cells were grown as explained in the section's first paragraph, until they reached an OD_{600nm} of 1. After a 3-minute centrifugation step at 6,000 × g, 4ºC, cells were washed three times with 1 mL deionized water and finally resuspended with 100 μL of a deionized water solution containing 20% glycerol. Similarly, after the preparation of 50 μL aliquots, cell preparations were kept on ice for 10 minutes, before storage at -80ºC. Electrocompetent *E. coli* DH5α cells were transformed by electroporation. Each cell aliquot together with the desired DNA mixture was transferred to 2 mm electroporation cuvettes from BTX and incubated for 30 minutes on ice. Cuvettes were dried with laboratory paper tissue and an electric field of 12.5 kV/cm was applied during 5 ms, using the pulse controller ECM 399 Electroporation System, BTX. The cell mixture was then resuspended with 950 μL LB broth (25 g/L) and incubated at 30ºC, 250 rpm. Ultimately, these cells were plated in LB agar medium (40 g/L) supplemented with erythromycin 250 or 500 μg/mL and incubated at 30ºC for 48h-60h. Electroporation candidates were grown as explained for chemically transformation candidates.

Growth conditions and transformation of *Lactococcus lactis* LMG 19460

LAB cells were cultured overnight in a pre-inoculum, in 15 mL falcon tubes with 5 mL M17 broth microbiology medium (42 g/L) from SIGMA-ALDRICH supplemented with glucose 0.5% (w/v) and 5 µg/mL erythromycin, at 37ºC, 250 rpm, when cells contained the pTRKH3 plasmid. After reaching the exponential growth phase cells were transferred to 15 mL falcon tubes containing 5 mL fresh LB medium supplemented with glucose and antibiotic, and incubated at the respective conditions until reaching an OD_{600nm} of approximately 1 for cell banks and late exponential for pDNA purification.

Lactococcus lactis electrocompetent cells were obtained using a growth protocol with glycine in osmotically stabilized media. After an overnight growth in 5 mL M17 broth (42 g/L), supplemented with 0.5% (w/v) glucose, at 30ºC, 100 rpm, in a 15 mL falcon tube, cells were transferred to a 100 mL Erlenmeyer flask containing 75 mL M17 broth (42 g/L) supplemented with 0.5% (w/v) glucose (GM17 medium), as to start the culture at an OD_{600nm} of 0.1. After the culture reached an OD_{600nm} between 0.5 and 0.8, a 100-fold dilution was made into a new Erlenmeyer flask containing M17 broth (42 g/L), 0.5% (w/v) glucose, 0.5 M sucrose and glycine 2% (SGGM17 medium). For this 100-fold dilution, 750 µL of cell suspension grown in GM17 was transferred into a 100 mL Erlenmeyer flask containing 75 mL SGGM17 medium. Later one this electrocompetent protocol was also preformed using SGGM containing glycine 1%. Cells were grown in SGGM17 until reaching an OD_{600nm}>1. Afterwards, cells were centrifuged at 5,000 × g, 4ºC, for 10 minutes and then washed twice with a 1 mL water washing solution, composed by sucrose 0.5 M and glycerol 10% (v/v). Cells were finally resuspended in a particular volume of washing solution in order to have 8.89×10^{10} cells/mL in each 100 µL cell bank. After the preparation of 100 µL aliquots, cells were kept on ice until storage at -80ºC. Competent cells aliquots were diluted in a 1:3 ratio using PCR grade water and divided into 40 µL aliquots, which contained 1.185×10^{10} cells and which were mixed with the desired DNA and incubated on ice for 30 minutes in 1 mm electroporation cuvettes from BTX. The transformation process was conducted using a pulse controller BTX ECM 399 electroporator, and the conditions used are shown in Table 1.

Table 1 Electroporation conditions and antibiotic concentration selection of 8.89×10^{10} cells/mL of L. lactis LMG 19460.

pDNA mass (ng)	Electric pulse conditions (# of pulses × kV/cm) during 5 ms	Apramycin concentration (µg/mL)
10	3 × 10	500 and 1 000
100	1 × 10	500 and 1 000
	3 × 10	
	3 × 20	1 000 and 2 000
250	3 ×10	500
500	1 × 10	500 and 1000

After the electroporation procedure, cells were immediately resuspended in 960 uL recovery medium, composed by M17 broth (42 g/L), glucose 0.5% (w/v), Sucrose 0.5 M, $MgCl_2$ 20 mM and $CaCl_2$ 2 mM, and incubated in 30ºC, without agitation, for 3 hours. Afterwards, these cells were pelleted for 3 minutes at 6000 × g, 4ºC and the pellet was transferred to 15 mL falcon tubes containing 5 mL of M17, supplemented with glucose 0.5% (w/v) and 250, 300 or 500 µg/mL apramycin, in which they grew overnight, in a Memmert incubator at 30ºC. After, cells were incubated at 30ºC on microplates containing complex agar medium constituted by tryptone 10 g/L, yeast extract 5 g/L, sucrose 200 g/L, glucose 10 g/L, gelatin 25 g/L, agar-agar 15 g/L, $MgCl_2.6H_20$ 0.5 g/L and $CaCl_2$ 0.3 g/L, supplemented with apramycin 500, 1 000 or 2 000 µg/mL, until visible growth (48-72 hours). Throughout electroporation tests, nthCRISPRa plasmid from three different *E. coli* strains DH5α, *SCS 110 and GM 2163* were used as to see whether the transformation of methylated and non-methylated DNA had an influence in transformation efficiencies, DNA copy number or quality. As transformations were not achieved, DNA origin is not specified.

For the <u>spectinomycin-induced pDNA production protocol</u>, cell bank deriving *L. lactis* cells containing the plasmid pTRKH3 or nthCRISPRa (a transformation candidate) were cultured in 15 mL falcon tubes containing 5 mL of M17 medium (42 g/L) supplemented with glucose 0.5% (w/v) and erythromycin 5 µg/mL or apramycin 250 µg/mL, as previously explained. Cultures were grown until reaching mid-exponential phase, with as OD_{600nm} of approximately 2. The falcon tubes containing cultures were supplemented with spectionomycin 10 mg/mL and left for incubation for 16 hours, after which the plasmid was purified. Cells harbouring the pTRKH3 vector were incubated at 37ºC, 250 rpm, while candidate cells for nthCRISPRa transformation were grown at 30ºC, 250 rpm.

Cell banks

Bacterial cell banks, excluding competent cell banks (explained above), were prepared using 80 µL cell suspension solution and 20 µL glycerol 99.99%, in 1.5 mL microcentrifuge tubes. Promptly, the 100 µL cell preparations were kept on ice until storage at -80ºC.

Relation between optical density and cell number

To ascertain the number of cells in a given growth broth, the relation of OD_{600} = 0.1 or 7×10^7 cells, used by Jones *et al.*[28] for *L. reuteri* cell counting, was also used in the present work. This relation was confirmed for *L. lactis* LMG 19460 in the laboratory by Martins, M.[29]. In the present work, it was used to determine the resuspension volume of washing solution necessary for 100 mL cell banks of *L. lactis* LMG19460 electrocompetent cells to have 8.89×10^{10} cells/mL and to determine the necessary pellet volume and dilutions for cell PCR reactions (explained above).

DNA purification, quantification and quality assessment

1. Purification of plasmid DNA from *E. coli* DH5α:

DNA was purified using the NZYMiniprep kit from NZYTech following the manufacturer's instructions. Some manufacture recommendations were performed: using double the A1, A2 and A3 buffer volumes and preheat the elution solution until 70ºC. For DNA purification procedures, the elution buffer used was PCR grade water (filtered and autoclaved). The use of PCR grade water, instead of buffer AE in the elution step avoids inhibitory effects that may happen in PCR reactions and circumvents electric arcs during electroporation due to the elution buffers' salt composition. Purified pDNA mixtures were stored at 4ºC until use.

2. Purification of plasmid DNA from L. lactis LMG19460:

DNA was purified based on the Plasmid DNA Purification user manual of the Nucleo Spin Plasmid, DNA, RNA and protein purification kit, from Macherey Nagel Bioanalysis. The first two steps of purification were based on "Isolation of plasmids from Gram-positive bacteria", which includes a lysis step with lysozyme 10 mg/mL to aid cell wall lysis. For the remaining purification steps the protocol "Isolation of low-copy plasmids" was considered.

For the <u>alternative lysis step</u> (boiling) procedure, the harvesting of cells from growth cultures was made by centrifugation at 4 000 × g and 4ºC for 15 minutes. The pelleted cells were resuspended in 1

mL of "ice cold" STE solution, containing 0.1 M NaCl, 10 mM Tris.Cl (pH 8) and 1 mM EDTA (pH 8). The resuspended pellet was then re-centrifuged using the same conditions. The lysis by boiling was then performed as an adapted protocol from Holmes and Quigley.[30] The pellet was resuspended in 100 μL STET containing 0.1 M NaCl, 10 mM Tris.Cl (pH 8) and 1 mM EDTA (pH 8) and 5% (v/v) Triton X-100. The resuspended pellet was transferred to an Erlenmeyer and 10 μL of lysozyme 10 mg/mL, prepared with 10 mM Tris.Cl (pH 8), was added. The samples were boiled under a Bunsen burner under continuous manual stirring. After the samples started to boil, erlenmeyers were immerged in a 99ºC wet bath for 40 seconds. The samples were then cooled by placing the Erlenmeyers in "ice-cold water" for 5 minutes. The viscous content was transferred to centrifuge flasks and centrifuged at 13,000 rpm, 4ºC, for 30 minutes. The remaining purification steps were made according to the previous mentioned protocol, excluding the lysis step.

3. Purification of genomic DNA from *L. lactis* LMG 19460:

The protocol "Isolation of Genomic DNA from Gram Positive and Gram Negative Bacteria" from Wizard Genomic DNA Purification Kit, Promega, using the Gram-positive bacteria specifications, was used. However, in the last purification step, the rehydration of the DNA pellet was done using PCR grade water, instead of Rehydration Solution.

4. PCR clean up and purification of DNA from TAE agarose gels:

the NZYGelpure kit from NZYTech was used based on the "Protocol for PCR clean-up or DNA purification from enzymatic reactions". DNA originated from PCR reactions that could not be purified using the PCR clean up protocol, due to unspecified amplifications, had to be purified using the same kit, but using the "Protocol for plasmid DNA purification from Agarose Gels". Once more, the elution step of both protocols was made using PCR grade water rather than Elution Buffer.

5. Quantification of purified DNA products:

Following DNA purification, DNA concentration was accessed using the NanoVue Plus from GE Lifesciences. The purification degree was determined considering the absorbance values (A) ratio between DNA and protein and DNA and salts and solvents, measured in the wavelengths for which these components have maximum absorption (A(260/280) and A(260/230), respectively). To confirm the quality of the extracted DNA and to make sure that the DNA was present in the purified mixture, 500 ng to 1 μg of DNA was run in agarose gels for visualization.

6. Digestion of nthCRISPRa:

To access DNA quality and integrity of the purified plasmid nthCRISPRa, a digestion using the Promega restriction enzyme EcoRI, with restriction site G^AATTC, was conducted. This enzymes's action leads to the digestion of the plasmid into three different fragments with 8,785 bp, 4,372 bp and 989 bp. The digestion was performed using 500 ng of nthCRISPRa, 0.3 μL (12 U/μL) EcoRI, 1.5 μL 10x buffer H and PCR grade water, making a total volume of 15 μL. The mixture was incubated at 37ºC for 1 hour and 30 minutes. Ultimately, the digested samples were run in 1% agarose gels to confirm if the fragments size were the ones expected.

PCR Amplifications

1. *E. coli* DH5α colony PCR

Colony PCR was conducted to confirm the presence of nthCRISPRe plasmid after Gibson Assembly (explained above) and cell transformation. The primers Ery_conf (forward:5'-CCATGCGTCTGACATCTATCT-3' and reverse:5'-CTGTGGTATGGCGGGTAAGT-3') and Cas9_conf (forward: 5'-CCCAGGTCAACATCGTCAAG - 3' and reverse: 5'-TCCATGATGGTGATGCCCAG -3') were either used to amplify a region inside the erythromycin resistance gene (190 bp) or the Scocas9 gene (241 bp), respectively, both only present in the plasmid. The transformation cfu candidates were picked to 70 μL PCR grade water. After careful resuspension, 50 μL of the cell mixture was incubated in 99ºC for 5 minutes, to promote cell rupture. The remaining 20 μL were kept under sterile conditions in 4ºC to be either grown in liquid medium or solid medium supplemented with erythromycin, if PCR amplifications were positive and had the expected base pair length. fter the heat treatment, cells were centrifuged at 12,000 × g, 4ºC for 1 minute. 10 μL of supernatant were used for the PCR reaction, which was carried out with the NovaTaq™ Hot Start Master Mix kit, from Merck Millipore. Each 25 μL reaction contained 10 μL of supernatant (containing the pDNA, if present), 12.5 μL master mix (final concentration of 1×), 1 μL of each forward and reverse primers (0.4 μM final concentration) and 0.5 μL PCR grade water. PCR conditions of Ery_conf and Cas9_conf were, respectively: 2 min at 95ºC; 35 cycles of: 30s at 95ºC, 30s at 55ºC and 30s at 72ºC; 10min at 72ºC; and 2 min at 95ºC; 40 cycles of: 30s at 95ºC, 30s at 56ºC and 3min at 70ºC.

2. PCR amplification for confirmation purposes using purified DNA as template:

To ascertain the *nth* genomic knockout in *L. lactis* LMG 19460, the primers Nth_conf (forward: 5'- CAAGGGCAAGTCAAATATAC - 3' and reverse: 5'- GTATGGCTGGTATAGACAGCA -3') that prime the upstream and downstream flanking regions of the homology arms sequences of the *nth* gene in the LAB genome were used. If the knockout had happened, the amplicon would be 657 bp (*nth* gene size) smaller than the amplicon containing the intact *nth* gene (2,790 bp). The Nth_conf amplicon is only amplified using gDNA as template. This PCR procedure was done using purified genomic DNA, before testing the cell PCR protocol. PCR conditions were: 2 min at 95ºC; 35 cycles of: 30s at 95ºC, 30s at 53ºC and 3min at 72ºC. Early in the present work, the presence of the nthCRISPRa plasmid in transformed and grown *L. lactis* LMG 19460 cells was assessed using the HA1 primers (forward: 5'- GAACCGTTGTAGAGTAAGTC - 3' and reverse: 5'- CGTGCACGTCCATTATCTCT -3') which amplify the homology arm 1 (1,028 bp), downstream of the *nth* sgRNA in the nthCRISPRa plasmid (Figure 1). As the HA1 sequence was also present in the *L. lactis* LMG 19460 genome, upstream of the *nth* gene, a purified mixture of pDNA was used as template. The PCR conditions used were: 2 min at 95ºC; 35 cycles of: 30s at 95ºC, 30s at 65ºC and 1.5min at 70ºC. Both reactions were performed using KOD Hot Start DNA Polymerase kit, from Merck Millipore and each 25 μL reaction mixture contained 2.5 μL buffer 10× (final concentration 1×), 2.5 μL dNTP mixture (final concentration 0.2 mM each), 2.3 μL $MgSO_4$ (final concentration 2.25 mM), 0.8 μL of each forward and reverse primers (final concentration 0.32 mM each), 0.5 μL KOD Hot Start Polymerase (final concentration 0.02 U/μL), DNA and water. For genomic (Nth_conf primers) and plasmid (HA1 primers) amplifications, 100-200 ng and 10-100 ng DNA mass were used, respectively.

3. *L. lactis* LMG 19460 cells PCR

A protocol using determined number of cells from pellet was used after optimization. The same PCR protocol - 2 min at 95ºC; 40 cycles of: 30s at 95ºC, 30s at 56ºC and 3min at 70ºC -amplified both Scocas9 internal sequence and the sequence amplified by Nth_conf primers, in the control sequences. The Scocas9 fragment was

amplified using the Cas9_conf primers. Control samples of purified pDNA (extracted from *E. coli*) and LAB gDNA used were 1 ng and 200 ng, respectively. The number of cells to be chosen for further tests was 250,000. The PCR was carried out using the NovaTaq™ Hot Start Master Mix kit, from Merck Millipore. Each 25 μL reaction contained 12.5 μL master mix (final concentration 1×), 1 μL of each forward and reverse primers (0.4 μM final concentration), the desired volume of diluted pelleted cells and the remaining volume of PCR grade water.

4. Reconstruction of the nthCRISPRa plasmid into nthCRISPRe

Aiming to change the selection marker sequence, the reconstruction of nthCRISPRa into nthCRISPRe was attempted. Two fragments (Apa_Ery and Ery_Apa) were amplified by PCR from nthCRISPRa (with 7,686 bp and 4,908, respectively) and the fragment containing the erythromycin resistance gene and promotor sequence (erm cassette) (with 1,299 bp) was amplified by PCR from pTRKH3. A long optimization procedure was performed. Primers Ery (forward: 5'-TGATCGACTGAGTCTAGAATCGATACGATTTTG-3' and reverse: 5'-GCTCATGAGCTTATTTCCTCCCGTTAAATAATAG-3'), Apa_Ery (forward: 5'-TGCCAAGCTTGGGCCCATATATAAGCTTC-3' and reverse: 5'-ATTCTAGACTCAGTCGATCATAGCACGATC-3'), Ery_Apa (forward: 5'-GAGGAAATAAGCTCATGAGCGGAGAACG-3' and reverse: 5'-ATATGGGCCCAAGCTTGGCACTGGCCGT-3') and Ery_Apa2 (forward: 5'-GAGGAAATAAGCTCATGAGCGGAGAACGagatgacgtt-3' and reverse: 5'-ATATGGGCCCAAGCTTGGCACTGGCCGTcgttttacaa-3') were designed using the available online software NEB Builder Assembly tool v1.12.16.[31] Highlighted nucleotides correspond to overlapping sequences for posterior Gibson ligation protocol. Ery_Apa_2 primers were designed to have ten additional nucleotides in the seeding region (non-overlapping region), which correspond to the lowercase nucleotide representative letters. The Ery and Apa_Ery fragments were amplified using KOD Hot Start Kit, the Ery_Apa using the Nova Taq Hot Start Master Mix and the Ery_Apa2 using the KOD Hot Start Master Mix. The first three PCR reaction mixtures were prepared as described before for each polymerase kit. The KOD Hot Start Master Mix for Ery_Apa2 PCR amplification 25 μL mixture contained 10 μL master mix, 0.6 μL of each forward and reverse primers (0.4 μM final concentration), the desired volume of DNA mixture and the remaining volume of PCR grade water. The PCR conditions for the Ery, Apa_Ery, Ery_Apa and Ery_Apa2 fragments were, respectively: 2 min at 95ºC; 40 cycles of: 1min at 95ºC, 1min at 59ºC and 2min at 70ºC**;** 2 min at 95ºC; 40 cycles of: 1min at 95ºC, 1min at 62ºC and 8min at 70ºC**;** 2 min at 94ºC; 40 cycles of: 30sec at 94ºC, 30sec at 62ºC and 5min at 72ºC**;** and 2 min at 95ºC; 40 cycles of: 30sec at 95ºC, 30sec at 54ºC and 2.25min at 70ºC.

Gibson Assembly

Prior to DNA ligation protocol, the DpnI enzyme from Promega was used to digest methylated template DNA, leaving PCR amplified products intact. This digestion aimed the degradation of template DNA, either pTRKH3 or nthCRISPRa, in order to not interfere in subsequent transformations. Digestion was done by adding 1 μL (10 U/μL) DpnI to 10 μL of each PCR reaction product mixture and incubating the mixture at 37ºC for 1 hour and 30 minutes. DpnI inactivation was accomplished by submitting the samples to 80ºC for 20 minutes. the ligation of the amplified and purified PCR products Ery, Apa_Ery and Ery_Apa or Ery_Apa_2, the protocol of NEBuilder HiFi DNA Assembly Master Mix/NEBuilder HiFi DNA Assembly Cloning Kit Instruction Manual was used. Ligations using different molarities and masses were performed (Table 2). In each 20 μL reaction volume, DNA fragments volume accounted for a maximum of 10 μL and the other 10 μL for NEBuilder HiFi DNA Assembly Master Mix. PCR grade water was added to the fragments mixture, until a total volume of 10 μL was reached.

Table 2 DNA masses (ng and pmol) used for different attempts of Gibson assembly-based ligation of the cited fragments.

Fragments	Gibson #1		Gibson #2-#5		Gibson #6	
	Mass (ng)	Mass (pmol)	Mass (ng)	Mass (pmol)	Mass (ng)	Mass (pmol)
Ery	120.00	0.142	120.00	0.142	84.44	0.100
Apa_Ery	60.00	0.012	80.00	0.016	250.00	0.050
Ery_Apa	60.00	0.019	40.00	0.013	159.10	0.050
Total Mass		0.173		0.171		0.200

Gibson #1 was performed under the recommendations of the protocol, using the DNA mass ratio of vector:insert=1:2, assuming the larger fragments (Apa_Ery and Ery_Apa) as a vector fragments and the shorter fragment (Ery) as the insert. Gibson #2-#5 were made under the assumption of Apa_Ery and Ery_Apa as a unique vector fragment, hence the approximate equimolarity (pmol per reaction volume). The double of the larger fragments mass mean was used to determine the Ery fragment mass. Gibson #6 was made using the DNA molarity ratio of vector:insert=1:2. Every Gibson protocol took into account the maximum molarity of 0.2 pmol for each ligation protocol. Samples were incubated at 50ºC and taken and stored after 15 minutes, 1 hour, 2 hours and 4 hours of incubation, aimed at further *E. coli* DH5α transformation. Either 2 μL or 1 μL were used for chemical transformation or electroporation, respectively.

Gel Electrophoresis

Gel electrophoresis was performed to visualize DNA after purification or from PCR origins, by separating the products according to their size. Agarose gels were prepared using 1 × concentrated TAE buffer and Seakem LE agarose (Lonza) to obtain 1% concentrated gels. The molecular weight marker NZYDNA ladder III was used to infer samples' base pair number. Electrophoresis was performed at 100 V for 1 hour, when using small gels and at 120 V for 1 hour and 30 minutes when using bigger gels. The DNA was stained with ethidium bromide for 15 to 40 minutes, depending on the agents' stability. Visualization was made on the Eagle Eye II image acquisition system (Stratagene).

OptFlux Simulation

A briefly metabolic modelling approach was conducted using the OptFlux 3.3.3 software, aiming the discovery of new knockout targets for pDNA optimized production. To this mean, the iAP358 metabolic model from *Lactococcus lactis*, subspecies lactis Il1403 was imported to the modelling software. This model was the one available with more similarity to what should be the metabolic map of *L. lactis* LMG19460. The critical genes/reactions were determined and the "Optimization-Evolutionary" tool was used to predict new possible knockouts for DNA production optimization. For the simulations, the maximum number of solutions evaluations was limited to 5,000 and the maximum number modifications were chosen to be 6. Critical and drain reactions were excluded from the simulation. The used method was "SPEA2 reaction knockout" and the objective function setup was determined as "BPCY: Biomass-Product Coupled Yield". The chosen product whose production would be optimized was nrdD_1, involved in the pyrimidine and purine synthesis pathways. A series of possible genes to be knocked out in a near future in *L. lactis* was obtained.

Results and Discussion

OptFlux *Lactococcus lactis* Metabolic Optimization

The 2005 metabolic model iAP358 of the strain species *L. lactis* IL1403[33] was used to predict potential knockouts considering a given objective function in the OptFlux software. Although not being the same strain, *L. lactis* IL1403 is the most phylogenetically closed organism to *L. lactis* LMG 19460, among other *L. lactis* species.[34]

The simulation of critical reactions is extremely helpful in the sense that it allows the user to exclude gene/reaction deletions that would be lethal to the cell, decreasing the vast amount of generated data that the human eye has to critically process. The "Optimization-Evolutionary" tool allows the optimization of a determined selected feature by different methods, which include reaction knockout and reaction under/over expression. This software allows the user to choose the number of possible combinations of alterations for the given objective function. Using all the data in metabolic models would be an almost impossible thing to solve, mathematically speaking, even for high processing computers. In order to solve this problem OptFlux uses an algorithm that makes all possible combinations of gene/reaction alterations within a group of limited number. The number of solutions in that group is defined by the user before the simulation. Therefore, different simulations using the same parameters can follow different outputs and solutions, as the system "randomly" changes the group composition. It seems that different simulations using the same parameters can be advantageous in the sense that solutions that are more likely to promote a given function could appear more times and in different combinations. In addition to critical reactions, the drain and transport reactions were also excluded from the simulation, because transport reactions usually play important roles in the cell stress response, nutrient uptake, signalling pathways, etc., and, although these "reactions" may be virtually possible, in practice, could have nefarious consequences for cell growth and manipulation. The chosen objective function setup was the BPCY and the chosen "reaction" to be optimized (product) was the *nrdD* gene product. The *nrdD* gene codes for an anaerobic ribonucleoside-triphosphate reductase (EC number 1.17.4.2)[35] and it is involved in the pyrimidine and purine metabolic pathways, responsible for the dCTP, dGTP, dATP and dGTP production, which leads to DNA synthesis. This was the closest product to the DNA synthesis pathway that one could find in the software and have results, which does not mean that other genes/reactions could not offer better results. After a couple simulations using the same parameters, the solutions that were more common between knockout predictions were chosen and inferred about their function (Table 3). The biological significance standpoint of each proposed deletion was analysed and their function information was taken from NCBI data base[34] and Bolotin *et al*[36].

Table 3 List of chosen genes/reactions outputted for knockout using Biomass-nrdD Coupled Yield in OptFlux Software. Gene functions were assessed using the NCBI ID number.

Gene/ Reaction	Function	NCBI ID number
araT	Aromatic amino acid aminotransferase.[37]	1113660
nucA	Nucleotidase: nucleotide and nucleoside interconversions.	1114734
yjhF	Phosphoglycerate mutase (glycolysis pathway).	1114580
gltB	Glutamate synthase large subunit: brings together nitrogen and carbon metabolism.	1114935
nrdF	Ribonucleoside-diphosphate reductase beta chain (pyrimidine pathway).	1114605
adhE	Bifunctional acetaldehyde-CoA/alcohol dehydrogenase (energy metabolism; fermentation).	1115832
argH	Arginine biosynthesis - argininosuccinate lyase.	1113731
argG	Arginine biosynthesis - argininosuccinate synthase.	1113730
purH	*De novo* purine biosynthesis.	1115175

Most of the outputted genes/reactions are involved amino acid modifications and energy metabolism. The knockout proposals of carbon metabolism may have some sense as it can be explained by the need to channel carbon flux to DNA synthesis, instead of fermentative pathways. The knockout proposal of *purH* gene/reaction does not seem realistic as this gene/reaction is directly involved DNA production. *purH* is involved in the *de novo* purine biosynthesis by directing carbon coming from de pentose phosphate pathway to other pathways that lead to DNA production, including pathways in which nrdD is involved.[38] In principle, it does not seem a good approach to knockout the purH gene in a strain aiming the increase of DNA production. This latter result evidences the importance of human eye analysis and judgment when working with simulation programs. It is also important to denote that most of the genes/reactions here proposed are involved in other metabolic pathways and that their knockout could result in pleiotropic effects. Also, these kinds of metabolic modulations oversee genes as reactions and does not know non-translated elements; thus, does mot consider genic regulation issues, which can be extremely important in real life application. The obtained proposed gene deletions could be potentially used in further strain optimization aiming to obtain a highly DNA productive bacteria, suitable for gene therapy applications. It is to consider that the metabolic model here used is from 2005 and may not be fully updated. These tools are as limited as their metabolic acquaintance and the detail that it has been put when designing and programing them. Therefore, some simulation outputs may not correspond be realistic. Nonetheless, OptFlux constitutes a very promising tool for knockout prediction for a given purpose and the development of new metabolic models considering other issues than reactions are much awaited in the scientific community.

Potential gene knockouts for increased plasmid production and quality

Theoretically, the target genes for knockouts, aiming the increase of pDNA yield production and quality, should include genes that are associated with DNA degradation and unspecific recombination, such as endonuclease and recombinase genes. Genes whose inactivation redirect the carbon flux to the pentose phosphate pathway, leading to an increased production of nucleotides also seem to be pertinent. In *E. coli* the genes *endA*, *recA*, *pgi*, *pykA* and *pyk*F, have been related to pDNA production, whose knockout increase DNA production and/or quality.[19,18] Indeed, the first plasmid construction for CRISPR/Cas9-based knockout in *L. lactis* LMG 19460 here presented is targeted at the endonuclease III gene, *nth*, which may have a similar function to the *E. coli*'s *endA* gene. Given the remaining knockout possibilities in *L. lactis* LMG 19460, the primers for new sgRNAs and homologous arms for recombination, as well as primers for KO confirmation, were designed (data not shown), according to the same strategy here presented. The fragment to be inserted in the vector was planned to

be cloned using the SpeI and ApaI restriction sites. As the SpeI and ApaI restriction sites are not found in the fragment constructions of *recA*, *pgi* and *pyk*, these restriction sites could be again used for fragment design, and these enzymes can be used for restriction digestion of the fragments and vector for posterior ligation. One and two nucleotides were added in the 5' end of the ApaI and SpeI restriction sequences, respectively, in order to increase higher efficiency of docking and digestion, as proposed by NEB.[32] The PCR amplification of HA1 and HA2 require *L. lactis* LMG 19460 genomic DNA as template. A SOEing PCR of the HA1 and HA2 fragments is made before ligating the whole fragment together (sgRNA and homology arms). If SpeI and ApaI restriction sites were present in the designed fragments (recognition sequence or homology arms), the fragment could be inserted in the vector by Gibson Assembly, after the vector SpeI and ApaI digestion.

Lactococcus lactis LMG 19460 transformation and _nth_ deletion assessments

Transformation of *Lactococcus lactis* LMG 19460 competent cells with the nthCRISPRa plasmid was attempted using different pDNA masses (10-500ng), electric pulse conditions (1 or 3 pulses and 1000 and 2000 V), apramicin recovery concentrations (250-500 ug/mL) and apramicin solid medium growth selection (500, 1000 or 2000 ug/mL), in different combinations. The first transformation attempts were being tested using primers HA1 in a PCR reaction using purified pDNA, as the template fragment was present not only in the plasmid DNA, but also in the genomic DNA of *L. lactis*. Genomic alterations were being tested using purified gDNA and Nth_conf primers. Typical results consisted of amplification of a 2500-3000 bp amplicon using the Nth_conf primers, indicating that no edition event had occurred; amplification of a 1,028 bp amplicon using the HA1 primers, supposedly confirming the presence of the plasmid DNA; and electrophoresis analysis of the samples in which pDNA was not visible. It was hypothesised that the absence of pDNA could be related to its low copy number nature. Aiming to increase the plasmid DNA copy number, the spectinomycin-induced pDNA production protocol[39,40] was used. The alternative lysis step was also performed. Both growth and purification parameters were used independently and in combination. pTRKH3-containing cells were used as controls, as the presence of the plasmid had already been confirmed. Neither of these protocols increased the nthCRISPRa plasmid copy number, as no pDNA structure was observed after growth protocols, purification and gel electrophoresis analysis. In fact, the pDNA apparent plasmid copy number did not increase in *L. lactis* harbouring the pTRKH3; the pDNA was visible and with good quality when using standard growth and pDNA extraction protocols. Using the spectinomycin protocol and standard purification decreased pDNA yield. The other two protocols did not render pDNA visualization after purification. This protocol did not serve its purposed and may require optimizations. Transformation attempts proceeded.

The use of two independently purified nthCRISPRa samples has led the conclusion that purified pDNA samples were contaminated with genomic DNA: The ~2,800 bp fragment was amplified when using primers Nth_conf and purified pDNA as template (data not shown). The Nth_conf primers prime to regions upstream and downstream of the homologous arms in the *L. lactis* genome, which means that this would not be possible using pure pDNA template. Therefore, the conclusions made until now would not be applicable. One cannot confirm whether the plasmid has entered cells using these set of primers. Therefore, a new pair of primers that amplify a small region (241 bp) of the internal sequence of Scocas9, only present in the plasmid, was designed. An optimization protocol allowed for the determination of a PCR programme suitable for the amplification of both Cas9_conf and Nth_conf, using a suspension containing a known number of cells as template (described bellow). This PCR protocol would allow a less time-consuming protocol to test a larger set of transformation candidates. Several transformation tests using the new assessment test were made. However, tests rendered the same results: approximately 30-500 colonies grew after respective apramycin selection and tested candidates did not underwent genome edition and did not contain the nthCRISPRa plasmid. These results suggested that the transformation events were unsuccessful. Given the observed growth in high apramycin concentrations, and the absence of pDNA, it was hypothesized that apramycin was not a suitable selection marker for *L. lactis*.

Optimizing the cell number for _Nth_conf_ and _Cas9_conf_ PCR amplification

Aiming the development of a fast approach to test *L. lactis* candidates for CRISPR modification and plasmid presence, a PCR protocol using cells as templates was developed. The elaboration of such protocol would enable to test more candidates at a time, without the need to resort time and resources at purifying plasmid and genomic DNA. In this assay, the volume of cell pellet, number of cells and needed dilutions was calculated.

The amplification using the Nth_conf primers was performed using 50,000 and 10,000 cells as PCR templates. Results rendered tenuous amplification of the 2,790 bp fragment with higher intense profile in PCR reactions using 10,000 cells, which may indicate that, although containing more genetic material, higher number of cells may cause inhibition or efficiency loss of PCR components at least when using KOD Hot Start DNA Polymerase, probably due to the presence of high cellular debris content. A PCR test was performed using KOD Hot Start Polymerase and Nova Taq Hot Start Master Mix. PCR reactions, using Nova Taq Hot Start Master Mix not only improved the Nth_conf amplification intensity in agarose gel, but also decreased the smear that was characteristic of KOD amplifications (data not shown). A NovaTaq PCR using 250,000 cells and a 40-cycle programme was tested using three different annealing temperatures: 50ºC, 53ºC and 56ºC. The number of cells was increased using the new DNA polymerase, as to increase the acuteness of the genomic fragment amplification and to increase the chance of amplification of the fragment of the nthCRISPRa low copy number plasmid, if present. The procedure provided good amplifications, specially using 56ºC of annealing temperature (data not shown). The use of 250,000 cells did not have a negative impact in the amplification of the fragment using NovaTaq polymerase and agreed with a more intense amplification of the genomic fragment. Although functioning with less cells, 250,000 cells were chosen as being ideal for, at least the Nth_conf amplified fragment. Conclusion remarks about the amplification of the internal fragment of Scocas9 were not possible, as no tested candidates encompassed the nthCRISPRa plasmid. As soon as the internal fragment of the nthCRISPR is detected, the number of cells can be restudied, if necessary, to assure a clear amplification of both Nth_conf and Cas9_conf resulting fragments. To assure that plasmid sequences would be amplified using cells as templates, a PCR test using *L. lactis* LMG 19460 containing the high-copy number pTRKH3 plasmid was performed using 10,000 and 1,000 cells as templates. For the present PCR reaction, the polymerase NovaTaq Master Mix and Ery_conf primers were used. All three samples tested rendered positive for the 190 bp fragment amplification and 10,000 cells showed to be better for confirmation purposes, as a clearer amplification was obtained (data not shown). Although PCR amplifications may depend on DNA template sequence, size,

structure and quality, it is possible to conclude that PCR protocols to amplify pDNA sequences using cells are possible and that, when present in *L. lactis* cells, the nthCRISPRa should be able to be amplified and detected.

PCR optimization for the amplification of the nthCRISPRe constituent fragments

The transformation of *L. lactis* LMG 19460 cells with the plasmid nthCRISPRa was not achieved, even using higher apramycin concentrations and different electroporation conditions. The lack of detection of nthCRSIPRa plasmid in *L. lactis* cells led to the debate whether apramycin was a good selection marker for this system, although previous minimum inhibitory concentration studies had been made for solid and liquid medium growth (unpublished). It was decided to change the selection marker apramycin to a selection marker already established for lactic acid bacteria, namely *L. lactis*, erythromycin, which is the pTRKH3 selective marker. Erythromycin has been used in the laboratory for a long time and it embodies in the list of ten antimicrobials that the EFSA has set as a basic requirement for the LAB group.[41] As the plasmid sequence did not contain restriction sites that would allow the removal of the *apmR* cassette and the insertion of erythromycin resistance gene and its promoter (erm), the assembly of desired fragments had to be done using a Gibson Assembly protocol. Due to the large size of nthCRISPRa, the amplification of the plasmid was decided to be made by the amplification of two different fragments. One fragment narrowed from the only restriction site of the ApaI restriction enzyme until the beginning of the *apmR* cassette, where the sequence of the erm cassette would start (Apa_Ery with 7,666 bp); and the other from the end of the *apmR* cassette, where the sequence of the *eryR* would finish until the ApaI restriction site (Ery_Apa with 4,888 bp). The 1,279 bp erm and *eryR* sequence would be placed in the same location as the previous resistance marker. To shorten the resulting plasmid sequence, the deletion of the region between the ApaI and the HindIII restriction sites (Figure 1) in the Huang and colleagues[27] pKCcas9dO vector and nthCRISPRa, was removed by designing primers to do as such: the Apa_Ery primers have an addition of 10 nucleotides that overlap the HindIII restriction region and the Ery_Apa primers have an addition of nucleotides that overlap the ApaI restriction region (Figure 1).

The Ery (erm and *eryR*) amplification from pTRKH3 was promptly amplified using a single gradient test for three independent annealing temperatures of 59ºC, 60ºC, and 61ºC. The annealing temperature of 59ºC managed a clearer and more intense amplification (data not shown) and no further optimizations were necessary.

The amplification of the Apa_Ery fragment was attempted using annealing temperatures ranging from 58ºC to 63ºC. The amplification of the desired 7,686 bp product was achieved using 60.5ºC annealing temperature. However, the PCR reaction also yielded other unspecific amplifications, some of which having higher intensity in the agarose gel comparatively to the desired product. More problematic was the amplification of fragments a few base pairs larger and a few base pairs smaller, rendering it difficult to posteriorly purify the DNA band from the TAE agarose gel (data not shown). Several PCR reaction tests followed, and some alterations aiming the increase of the PCR stringency included, in addition to annealing temperature variation: magnesium salt concentration variation, increment of DNA template mass and use of the previously amplified fragment mixture as template. Increasing the DNA template mass and the annealing temperature revealed an apparent decrease of the amplification of the products upstream and downstream of the desired product, which was advantageous for future purification of the DNA band from the TAE agarose gel, as it was decided that no further optimizations would be attempted and that the desired Apa_Ery product would be amplified using an annealing temperature of 62ºC and purified from the TAE agarose gel. Later in the work, this fragment was purified using a PCR clean up protocol, although having one unspecific product.

The Ery_Apa fragment amplification was very laborious as the variation of annealing temperatures (from 59.5ºC to 70.5ºC), different DNA template masses (0.05, 0.5 and 5 ng) and MgSO4 concentrations (0.75, 1, 1.25, 1.75, 2.25, 2.75 and 3.25 mM), using the KOD Hot Start Polymerase did not follow positive amplicons. Although not ideal due to its non-proofreading and lower fidelity nature, a PCR protocol using NovaTaq Hot Start Master Mix was attempted and 0.5 ng nthCRISPRa template. Of the 50ºC, 54ªC, 58ºC, 62ºC, 66ºC and 70ºC annealing temperatures, the 4,908 bp product amplification was possible using 58ºC and 62ºC annealing temperatures, although characterized by an intense generalized smear (data not shown). Aiming to increase the binding specificity of the primers to the DNA template, the Ery_Apa2 pair of primers was designed as to have 10 additional seeding nucleotides in its sequence. The DNA template increment to 20 ng proved to be advantageous for the amplification profile and 62ºC was chosen as annealing temperature. Later, another set of tests was made using different DNA polymerases (KOD Hot Start Master Mix, BIOTAQ, Platinum Supermix HiFi DNA Polymerase, Pfu Turbo, Nzylong and Supreme NzyLong) employing the manufacturer's recommended conditions for each enzyme, using 5 ng of template and a gradient programme with independent annealing temperatures of 58ºC, 60ºC and 62ºC. KOD Hot Start Master Mix yielded the desired product using annealing temperatures of 58ºC and 60ºC, together with seven unspecific products varying from approximately 2,500 bp to 400 bp (data not shown). As a proofreading high fidelity enzyme, a KOD Hot Star Master Mix amplification would be preferable over an amplification using the error prone Nova Taq Polymerase, when projecting a cloning protocol. A test using 50ºC and 60ºC annealing temperatures proved that the pair of primers Ery_Apa2 was preferable over Ery_Apa primers. Finally, an annealing temperature gradient test was performed using 50ºC, 54ºC and 58ºC. As the PCR reaction yielded one other unspecific products, the product would be purified from the TAE agarose gel, after a 54ºC annealing temperature PCR protocol.

Gibson Assembly

Six different Gibson Assembly reactions were independently performed and transformed in *E. coli* DH5α cells. Gibson reactions #1 to #4 were performed using the Apa_Ery fragment mixture, purified from TAE agarose gel and the Ery_Apa fragment amplified using the low-fidelity NovaTaq Polymerase, and purified from TAE agarose gel. Gibson reactions #5 and #6 was performed using the Apa_Ery fragment mixture obtained by PCR cleanup, and the Ery_Apa fragment amplified using the high-fidelity KOD Hot Start Master Mix, and purified from TAE agarose gel.

The reaction mixtures of Gibson Assembly #1 were used in three different transformation events, two heat shock chemical transformations and an electroporation protocol, which, after selection in erythromycin 500 μg/mL, yielded 25 cfu, in total. Of analysed colonies, the ones that exhibited growth did not contain pDNA products. Three candidates were assessed for Cas9_conf fragment present, and the result was negative. Gibson Assembly #2 was also used for two heat shock chemical transformations and an electroporation protocol. The electroporation protocol yielded 3 colonies after selection in erythromycin 500 μg/mL, 2 of which did

not contain the Cas9_conf fragment. One heat shock event was performed using erythromycin 250 μg/mL. The reasoning behind the antibiotic concentration decrease was the possibility of the 13,833 bp resulting nthCRISPRe plasmid to be too large to produce sufficient number of copies to produce enough erythromycin resistance protein to cope with the high concentration of antibiotic in the medium. This transformation protocol yielded a total 742 candidates (from three different transformations: 142, 313 and 287 cfu), 12 of which were submitted to colony PCR to assess the presence of the cas9 containing fragment. After electrophoresis of PCR products, the presence of such fragment was confirmed in 11 of the total 12 transformation candidates. Of the 6 grown candidates, pDNA was not observed and the samples appeared degraded, with a generalized smear starting from a tenuous band above 10,000 bp. Other 3 candidates were analysed, and the results were similar; however, the Cas9_conf and Ery_conf primers were used to confirm the fragments presence by PCR. The test turned out to be positive, suggesting that cells could have taken up the linear DNA fragments, instead of the pDNA to be constructed, which allowed the erythromycin resistance feature. The other heat shock event using Gibson #2 reaction mixture yielded 42 cfu after selection in erythromycin 350 μg/mL, of which 8 were picked to liquid medium, 1 exhibited growth, but did not contain the pDNA product. Gibson #3 products were used to transform *E. coli* either by chemical means (heat shock) or by electroporation. The heat shock protocol lasted 90 seconds, instead of one minute. After transformation and recovery, cells were divided and grown in solid medium either containing erythromycin 250 μg/mL or 500 μg/mL. Chemical transformation yielded 1 cfu after selection in ery 500 μg/mL, and electroporation yield 60 cfu after selection in ery 250 μg/mL and 1 cfu after selection in ery 500 μg/mL. Although the 6 analysed cfu rendered positive results for the presence of the Scocas9 241 bp fragment (data not shown), none exhibited growth in liquid medium. After electroporation of *E. coli* with Gibson #4 reaction products, and selection in ery 250 μg/mL containing medium, 253 cfu were perceived, 9 of which gave positive results concerning the presence of the internal region of the erythromycin resistant gene, and 1 of which exhibited growth in liquid medium, but rendered negative for the presence of the Cas_9 fragment. Once again, it seems that the fragments are present in cells that undergo transformation, but the plasmid is not forming itself during the Gibson Assembly Protocol. Other electroporation event using the Gibson #4 reaction mixture yielded no colonies, after selection in ery 500 μg/mL. Chemical transformants of Gibson #5 and #6 accounted for 23 cfu, after selection in ery 500 μg/mL. The only candidate that exhibited growth in liquid medium (from Gibson #5), rendered negative results regarding pDNA presence.

With all the above-mentioned results one can conclude that something appears to allow cell growth without the presence of the plasmid, and that could be the presence of non-ligated fragments. It is difficult to counteract this issue using the present system, as one cannot make sure that the ligation is happening before the transformation process. The concentration of DNA used in the present protocol of Gibson Assembly is so low that the ligation product cannot be visualized in an agarose gel electrophoresis. One must trust ligation probability events until the cloning of the desired product is obtained. Moreover, the ligation itself might be working, but transformation efficiency may not be high enough to see it. To increase Gibson Assembly efficiencies one can increase the number of overlapping nucleotides in each fragment. Although not making use of one of the major advantages of Gibson Assembly (ligation of multiple fragments in a single ligation reaction), the assembly protocol may be done using two fragments at a time, instead of the ligation of three fragments at once.

nthCRISPRa and nthCRISPRe considerations

It is important to highlight that the pKCcas9dO plasmid was used to induce genome editing in *Streptomyces coelicolor* M145, and although being a Gram-positive bacterium as well, the system is yet to be proved to work in *L. lactis*.

The naturally temperature sensitive replicon pSG5, from *Streptomyces ghanaensis* DSM2932, is stably inherited under growth conditions below 34ºC. Above this temperature, the plasmid does not replicate and is lost after some replication cycles. Therefore, the use of this plasmid as a mutational platform seems extremely useful, as these vectors can be eliminated from the host cells very efficiently by increasing the incubation temperature, after edition.[42] In the present work, *L. lactis* cells that would render positive results for the *nth* deletion would be then cultivated at 37ºC to lost the nthCRISPRa/nthCRISPRe plasmid, to proceed with the other transformations for the study of plasmid copy number or to proceed with other knockout experiments. The pSG5 replicon is of wide host range within *Sptreptomyces* species; however, it has not been shown to work in lactic acid bacteria. If such origin is not applicable to LAB, the present system would resemble a non-replicative editing system, which is much less effective. The increased efficiency of the CRISPR/Cas editing system relies in proper induction of DSB; therefore, using a non-replicative plasmid, the probability of enough expression of Scocas9 and sgRNA for gene targeting and cleavage is decreased and, consequentially, editing events are decreased. Alternatives to this replication origins are replicons known to function in LAB, such as pAMβ1 and pWV01 replicons.[43,44] The pvW01-derivative pVE6002 thermosensitive replicon, non-functional from 35ºC, is also a possible alternative.[45] In fact, thermosensitive replicons may be advantageous in this kind of systems as the plasmid curing efficacy is increased.[46]

The insertion of the fragment containing the sgRNA and homology arms was made by the SpeI restriction site, present in the J23119 promoter. Hence, the distance between the end of the promoter and the beginning of the guide RNA sequence is 2 bp, whereas the distance between the promoter and the sgRNA of Huang *et al*. design is 19 bp[27]. This may not be problematic as the objective is to obtain a non-translatable guide transcript. In fact, the base pair reduction between these components may reduce some instability effects that larger intermediated fragments may cause.

The tipA promoter should be responsible, in the nthCRISPRa plasmid, for the expression of Scocas9 *in vivo*. The transcription of this promoter is known to be highly induced by thiostrepton and other thiostrepton-like antibiotics, even when using extremely low concentrations.[47–49] This system was already tested in thiostrepton non-producing *Streptomyces* organisms, such as *Streptomyces lividans*.[50] in this organism, tipA promoter reveals expression with and without thiostrepton, but is highly inducible by the latter.[50] In the present work, the thiostrepton antibiotic was not used. As no *L. lactis* transformation candidates were obtained, one cannot conjecture about the activity of the tipA promoter in this organism. If colonies did contain nthCRISPRa and were not edited, it could be due to low expression of the Scocas9 and lack of DSB events. In this case, the addition of thiostrepton to the medium could be tested to see whether edited cells would grow. The best-case scenario would be that tipA promoter would have a basal expression, enough to produce the sufficient amounts of Scocas9 to induce the desired DSB. An alternative would be to change this promoter with native promoters, such as the inducible nisin A promoter or the constitutive promoters P45 or P32.[51] The use of conditional or weak

promoter seems advantageous when considering that the overexpression of Cas9 proteins can render toxic effects.[52] other problematic is the codon usage of the Scocas9 coding gene itself. In nthCRISPRa, the sequence that codes for the Cas9 protein is codon optimized for *S. coelicor*. Although being both Gram-positive bacteria, *S. coelicor* and *L. lactis* have very dissimilar GC contents of approximately 70% and 35%, respectively[53,54], and given this disparity, the preferred codons and regulatory elements used by one organism in one end of the GC content spectrum may not be recognized by organisms on the other end.[55] If the editing system fails to work with nthCRISPRe, the Scocas9 coding sequence may need to be reconstructed using *L. lactis* codon usage bias. The commonly used SpyCas9 could also be used, as the GC content of *Streptococcus pyogenes* is of approximately 36%[56], resembling more the one observed in *L. lactis*.
The present expression and editing system also contains an oriT region, which conduces to the consideration that this plasmid(s) may be transformed in *L. lactis* LMG 19460 by conjugation, in the future, if transformation fails to work.

Conclusions and Future Remarks

Transforming the nthCRISPRa plasmid did not prove to be an easy task and transformants were not obtained. The desired edition and the nthCRISPRa plasmid were not detected in the supposedly transformed cells that manifested growth, using previously defined minimal inhibitory concentrations and higher concentrations of apramycin. This led to the conclusion that apramycin is not the indicated antibiotic to be used in lactic acid bacteria selection, as medium containing this compound is either selecting non-sensible cells or spontaneous resistant organisms. In what concerns the assessment of edition and plasmid presence after transformation, a PCR protocol using 250,000 candidate cells and NovaTaq Hot Start Master Mix was successfully established. The PCR protocol achieved the simultaneous desired amplification of Nth_conf and Cas9_conf fragments using purified gDNA and pDNA, respectively. However, the amplification of the Cas9_conf fragment using cells as template is yet to be achieved and, as soon as the internal fragment of the nthCRISPR is detected, the number of cells may need to be restudied.
PCR amplifications for Gibson assembly proven that PCR amplifications may be laborious when they are not optimized and that they depend on a variety of factors, such as template DNA quality and concentration, DNA polymerase, annealing temperature, co-factors presence, etc. Although using similar conditions, generation of amplicons were not always totally reproducible, indicating that each PCR reaction preparation may harbour small differences in sample preparations that cause changes in amplification profiles. Gibson assembly is yet to be achieved, and we may conclude that the ligation has not taken place properly as Scocas9 or erythromycin resistance internal fragments are detectable by colony PCR, but cells do not contain plasmid DNA. The successful transformation of the linear fragments inside cells would explain this fact and the cell growth in erythromycin containing media. To ascertain fragments ligation before growing candidates in liquid medium, primers that prime forwardly one of the fragments and reversely other fragment can be designed.
Once the *nth* knockout of *L. lactis* LMG 19460 and the editing plasmid curing is accomplished, a qRT-PCR relative method should be performed, to quantify plasmid DNA inside cells. To analyse pDNA copy number, pTRKH3 could be used. A comparison between *L. lactis* LMG 10460 wild type cells and *L. lactis* LMG 10460 *Δnth* would be made to infer whether the knockout had improved pTRKH3 plasmid copy number. This qRT-PCR method would be made based on the ratio between the *erm* gene, present in pTRKH3, and the amplification of the single-copy genome reference gene, *feoA*. Furthermore, pDNA supercoiled fractions would be analysed by image analysis of agarose gel electrophoresis.
The main concern of the present work is the system feasibility. As it was designed to suit *Streptomyces* genome edition, replicative, transcriptional, expression and regulatory elements of *L. lactis* LMG 19460 may not recognize some plasmid elements in the nthCRISPRa/nthCRISPRe, which is corroborated with the fact that *Streptomyces* and *Lactococcus* have very dissimilar GC contents. However, such conclusions can only be made after the transformation of the plasmid is achieved. Alternatively, the system could be redesigned to incorporate native or previously tested regulatory and expression elements from *L. lactis* or from more closely related organisms. Other editing systems such as CRISPR/Cas9 selection combined with ssDNA recombineering can also be used. The latter has indeed been successfully used for mutagenesis and deletions in the LAB *Lactobacillus reuteri*.
CRISPR-based genome editing techniques were quickly adopted for eukaryotic applications. However, their use in prokaryotic applications have lagged behind, mainly due to technical shortcomings, as transformation efficiencies and the still lacking characterization of many prokaryotic regulatory mechanisms and optimal growth conditions. Basic microbiology and molecular biology are very important to develop new genome editing techniques, and will allow the extended application of CRISPR/Cas applications. There is no definitive answer for which is the best genome editing technique, but CRISPR/Cas9 seems to be of great promising applications as soon as the system is established. One great possibility of this tool is the multiplexing, which will allow for the edition of different target sites in a single editing event.

References

1. Zhu, Y., Zhang, Y. & Li, Y. Understanding the industrial application potential of lactic acid bacteria through genomics. *Appl. Microbiol. Biotechnol.* **83,** 597–610 (2009).
2. Makarova, K. *et al.* Comparative genomics of the lactic acid bacteria. *Proc. Natl. Acad. Sci. U. S. A.* **103,** 15611–15616 (2006).
3. *Biology of Microorganisms on Grapes, in Must and in Wine*. (Springer-Verlag Berlin Heidelberg, 2009).
4. Hayek, S. A. & Ibrahim, S. A. Current limitations and challenges with lactic acid bacteria: a review. *Food Nutr. Sci.* **4,** 73 (2013).
5. Pessione, E. Lactic acid bacteria contribution to gut microbiota complexity : lights and shadows. *Front. Cell. Infect. Microbiol.* **2,** (2012).
6. Wells, J. M. & Mercenier, A. Mucosal delivery of therapeutic and prophylactic molecules using lactic acid bacteria. *Nat. Rev. Microbiol.* **6,** 349–362 (2008).
7. Cano-Garrido, O., Seras-Franzoso, J. & Garcia-Fruitós, E. Lactic acid bacteria : reviewing the potential of a promising delivery live vector for biomedical purposes. *Microb. Cell Fact.* **14,** 137 (2015).
8. Steidler, L. & Rottiers, P. Therapeutic drug delivery by genetically modified *Lactococcus lactis*. *Ann. N. Y. Acad. Sci.* **1072,** 176–186 (2006).
9. Panagiota, F.-P., Christaki, E. & Bonos, E. Lactic acid bacteria as source of functional ingredients. in *Lactic Acid Bacteria-R & D for Food, Health and Livestock Purposes* 489–614 (2013).
10. García-fruitós, E. Lactic acid bacteria: a promising alternative for recombinant protein production. *Microb. Cell Fact.* **11,** 157 (2012).
11. Bermúdez-Humarán, L. G. *et al.* Engineering lactococci and lactobacilli for human health. *Curr. Opin. Microbiol.* **16,** 278–283 (2013).
12. Faurez, F., Dory, D., Moigne, V. Le, Gravier, R. & Jestin, A. Biosafety

of DNA vaccines: New generation of DNA vectors and current knowledge on the fate of plasmids after injection. *Vaccine* **28,** 3888–3895 (2010).
13. Pontes, D. S. *et al. Lactococcus lactis* as a live vector: Heterologous protein production and DNA delivery systems. *Protein Expr. Purif.* **79,** 165–175 (2011).
14. Kay, M. A. State-of-the-art gene-based therapies: the road ahead. *Nat. Rev. Genet.* **12,** 316–328 (2011).
15. Gardlík, R. *et al.* Vectors and delivery systems in gene therapy. *Med. Sci. Monit.* **11,** RA110-RA121 (2005).
16. Niidome, T. & Huang, L. Gene therapy progress and prospects: nonviral vectors. *Gene Ther.* **9,** 1647 (2002).
17. Son, Y. J., Ryu, A. J., Li, L., Han, N. S. & Jeong, K. J. Development of a high-copy plasmid for enhanced production of recombinant proteins in *Leuconostoc citreum*. *Microb. Cell Fact.* **15,** 12 (2016).
18. Gonçalves, G. A. L., Prather, K. L. J., Monteiro, G. A., Carnes, A. E. & Prazeres, D. M. F. Plasmid DNA production with *Escherichia coli* GALG20, a pgi-gene knockout strain: Fermentation strategies and impact on downstream processing. *J. Biotechnol.* **186,** 119–127 (2014).
19. Gonçalves, G. A. L., Bower, D. M., Prazeres, D. M. F., Monteiro, G. A. & Prather, K. L. J. Rational engineering of Escherichia coli strains for plasmid biopharmaceutical manufacturing. 251–261 (2012). doi:10.1002/biot.201100062
20. Sander, J. D. & Joung, J. K. CRISPR-Cas systems for editing, regulating and targeting genomes. *Nat. Biotechnol.* **32,** 347–355 (2014).
21. Mali, P. *et al.* RNA-guided human genome engineering via Cas9. *Science (80-.).* **339,** 823–826 (2013).
22. Cong, L. *et al.* Multiplex genome engineering using CRISPR/Cas systems. *Science (80-.).* **339,** 819–823 (2013).
23. Cho, S. W., Kim, S., Kim, J. M. & Kim, J.-S. Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nat. Biotechnol.* **31,** 230–232 (2013).
24. Maeder, M. L. & Gersbach, C. A. Genome-editing technologies for gene and cell therapy. *Mol. Ther.* **24,** 430–446 (2016).
25. Doudna, J. A. & Charpentier, E. The new frontier of genome engineering with CRISPR-Cas9. *Science (80-.).* **346,** 1258096 (2014).
26. Jinek, M. *et al.* A programmable dual-RNA–guided DNA endonuclease in adaptive bacterial immunity. *Science (80-.).* **337,** 816–821 (2012).
27. Huang, H., Zheng, G., Jiang, W., Hu, H. & Lu, Y. One-step high-efficiency CRISPR/Cas9-mediated genome editing in Streptomyces. *Acta Biochim. Biophys. Sin. (Shanghai).* **47,** 231–243 (2015).
28. Jones, S. E. & Versalovic, J. Probiotic Lactobacillus reuteri biofilms produce antimicrobial and anti-inflammatory factors. *BMC Microbiol.* **9,** 35 (2009).
29. Martins, M. S. C. Redesign of Lactic Acid Bacteria for plasmid DNA production A strategy for increase the plasmid copy number Examination Committee. (Instituto Superior Técnico, Universidade de Lisboa, 2015).
30. Holmes, D. S. & Quigley, M. A rapid boiling method for the preparation of bacterial plasmids. *Anal. Biochem.* **114,** 193–197 (1981).
31. NEBuilder Assembly Tool. Available at: http://nebuilder.neb.com/. (Accessed: 3rd April 2017)
32. New England BioLabs. Cleavage Close to the End of DNA Fragments. Available at: https://www.neb.com/. (Accessed: 20th July 2017)
33. Oliveira, A. P., Nielsen, J. & Förster, J. Modeling Lactococcus lactis using a genome-scale flux model. **15,** 1–15 (2005).
34. National Center for Biotechnological Information. Available at: https://www.ncbi.nlm.nih.gov/. (Accessed: 20th January 2017)
35. Guillot, A., Gitton, C., Anglade, P. & Mistou, M.-Y. Proteomic analysis of *Lactococcus lactis*, a lactic acid bacterium. *Proteomics* **3,** 337–354 (2003).
36. Bolotin, A. *et al.* The complete genome sequence of the lactic acid bacterium Lactococcus lactis ssp. lactis IL1403. *Genome Res.* **11,** 731–753 (2001).
37. Rijnen, L., Bonneau, S. & Yvon, M. Genetic characterization of the major lactococcal aromatic aminotransferase and its involvement in conversion of amino acids to aroma compounds. *Appl. Environ. Microbiol.* **65,** 4873–4880 (1999).
38. KEGG: Kyoto Encyclopedia of Genes and Genomes. Available at: http://www.kegg.jp/kegg/pathway.html#metabolism. (Accessed: 18th January 2017)
39. Spectinomycine. Available at: http://www.sigmaaldrich.com/catalog/product/sigma/s0692?lang=pt®ion=PT. (Accessed: 2nd February 2017)
40. Sambrook, J., Fritsch, E. F. & Maniatis, T. *Molecular cloning: a laboratory manual. Molecular cloning: a laboratory manual.* (Cold Spring Harbor Laboratory Press, 1989).
41. Karapetkov, N., Georgieva, R., Rumyan, R. & Karaivanova, E. Antibiotic susceptibility of different lactic acid bacteria strains. *Benef. Microbes* **2,** 335–339 (2011).
42. Muth, G., Nußbaumer, B., Wohlleben, W. & Pühler, A. A vector system with temperature-sensitive replication for gene disruption and mutational cloning in streptomycetes. *Mol. Gen. Genet. MGG* **219,** 341–348 (1989).
43. Foley, S., Bron, S., Venema, G., Daly, C. & Fitzgerald, G. F. Molecular Analysis of the Replication Origin of the *Lactococcus lactis* Plasmid pCI305. *Plasmid* **36,** 125–141 (1996).
44. Leenhouts, K. J. *et al.* Nucleotide sequence and characterization of the broad-host-range lactococcal plasmid pWVO1. *Plasmid* **26,** 55–66 (1991).
45. Maguin, E., Duwat, P., Hege, T., Ehrlich, D. & Gruss, A. New thermosensitive plasmid for gram-positive bacteria. *J. Bacteriol.* **174,** 5633–5638 (1992).
46. Oh, J.-H. & van Pijkeren, J.-P. CRISPR–Cas9-assisted recombineering in *Lactobacillus reuteri*. *Nucleic Acids Res.* **42,** e131–e131 (2014).
47. Holmes, D. J., Caso, J. L. & Thompson, C. J. Autogenous transcriptional activation of a thiostrepton-induced gene in *Streptomyces lividans*. *EMBO J.* **12,** 3183 (1993).
48. Murakami, T., Holt, T. G. & Thompson, C. J. Thiostrepton-induced gene expression in Streptomyces lividans. *J. Bacteriol.* **171,** 1459–1466 (1989).
49. Yun, B.-S., Hidaka, T., Furihata, K. & Seto, H. Promothiocins A and B, novel thiopeptides with a tip A promoter inducing activity produced by Streptomyces sp. SF2741. *J. Antibiot. (Tokyo).* **47,** 510–514 (1994).
50. Schmitt-John, T. & Engels, J. W. Promoter constructions for efficient secretion expression in *Streptomyces lividans*. *Appl. Microbiol. Biotechnol.* **36,** 493–498 (1992).
51. Zhu, D. *et al.* Isolation of strong constitutive promoters from *Lactococcus lactis* subsp. lactis N8. *FEMS Microbiol. Lett.* **362,** (2015).
52. Jiang, W., Brueggeman, A. J., Horken, K. M., Plucinak, T. M. & Weeks, D. P. Successful transient expression of Cas9 and single guide RNA genes in Chlamydomonas reinhardtii. *Eukaryot. Cell* **13,** 1465–1469 (2014).
53. Wright, F. & Bibb, M. J. Codon usage in the G+ C-rich Streptomyces genome. *Gene* **113,** 55–65 (1992).
54. Silva, I., Duarte, S., Moreira, L. M. & Monteiro, G. A. Draft Genome Sequence of the Plasmid-Free *Lactococcus lactis* subsp. lactis Strain LMG 19460. *Genome Announc.* **5,** e00210-17 (2017).
55. Primrose, S. B. & Twyman, R. Cloning in Gram-positive bacteria. in *Principles of Gene Manipulation and Genomics* 189–190 (John Wiley & Sons, 2013).
56. Maruyama, F., Watanabe, T. & Nakagawa, I. Streptococcus pyogenes Genomics. in *Streptococcus pyogenes: Basic Biology to Clinical Manifestations* 1–80 (2016).