# Harnessing CRISPR-Cas9 for genome editing in *Streptococcus pneumoniae*

Dimitra Synefiaridou and Jan-Willem Veening*

Department of Fundamental Microbiology, Faculty of Biology and Medicine, University of Lausanne, Biophore Building, CH-1015 Lausanne, Switzerland

*Correspondence to Jan-Willem Veening: Jan-Willem.Veening@unil.ch, tel: +41 (0)21 6925625, Twitter: @JWVeening

# Abstract

CRISPR systems provide bacteria and archaea with adaptive immunity against viruses and plasmids by detection and cleavage of invading foreign DNA. Modified versions of this system can be exploited as a biotechnological tool for precise genome editing at a targeted locus. Here, we developed a novel, replicative plasmid that carries the CRISPR-Cas9 system for RNA-programmable, genome editing by counterselection in the opportunistic human pathogen *Streptococcus pneumoniae.* Specifically, we demonstrate an approach for making targeted, marker-less gene knockouts and large genome deletions. After a precise double-stranded break (DSB) is introduced, the cells' DNA repair mechanism of homology-directed repair (HDR) pathway is being exploited to select successful transformants. This is achieved through the transformation of a template DNA fragment that will recombine in the genome and eliminate recognition of the target of the Cas9 endonuclease. Next, the newly engineered strain, can be easily cured from the plasmid that is temperature-sensitive for replication, by growing it at the non-permissive temperature. This allows for consecutive rounds of genome editing. Using this system, we engineered a strain with three major virulence factors deleted. The here developed approaches should be readily transportable to other Gram-positive bacteria.


# Importance

*Streptococcus pneumoniae* (the pneumococcus) is an important opportunistic human pathogen killing over a million people each year. Having the availability of a system capable of easy genome editing would significantly facilitate drug discovery and vaccine candidate efforts. Here, we introduced an easy to use system to perform multiple rounds of genome editing in the pneumococcus by putting the CRISPR-Cas9 system on a temperature-sensitive replicative plasmid. The here used approaches will advance genome editing projects in this important human pathogen.

# Introduction

*Streptococcus pneumoniae* (the pneumococcus) is a Gram-positive, human commensal that colonizes asymptomatically the mucosal surfaces of the upper respiratory tract (UTR) (Kadioglu et al. 2008). However, in susceptible groups like children, the elderly and the immunocompromised, it can occasionally become pathogenic causing diseases that range from a mild upper respiratory tract infection, acute otitis media and sinusitis, to a severe and potentially life-threatening condition such as pneumonia, bacteremia and meningitis (Simell et al. 2012). It is responsible for more than one million deaths annually (O'Brien et al. 2009) and in 2017, the World Health Organization (WHO) classified *S. pneumoniae* as one of twelve priority pathogens for which new antibiotics are urgently needed.

Historically, *S. pneumoniae* research played a central role in advancing molecular biology. While trying to develop a vaccine against the pneumococcus, Griffith discovered natural transformation (Griffith 1928). This was followed by research of Avery, MacLeod and McCarty to establish that DNA is the genetic material (Avery et al. 1944). Over the last decade, the pneumococcus has become a valuable model to study the cell biology of ovoid-shaped bacteria and several cell biological tools such as integration vectors, fluorescent reporters, inducible promoters and CRISPR interference have been established for this organism (Massidda et al. 2013; Keller et al. 2019; Liu et al. 2017). In addition, many selection and counterselection methods are available making it relatively easy to generate gene deletions, gene complementation mutants or point mutations in the pneumococcal genome (Sung et al. 2001; Halfmann et al. 2007; Y. Li et al. 2014; Sorg et al. 2019). However, all current gene deletion methods established for *S. pneumoniae* are poorly scalable and often require a specific genetic background to function (e.g. the *rpsL*+ background in the janus system (Sung et al. 2001)).

In the case of gene replacement by selection markers, while powerful, this also has drawbacks, preventing further modifications of the genome when there are no further selectable markers available for additional strain development. Also, many important categories of gene mutation, such as missense substitutions and in-frame deletions, usually present no selectable phenotype (Sung et al. 2001). To circumvent these issues, we here established CRISPR genome editing for use as counterselection in the pneumococcus.

Clustered regularly interspaced short palindromic repeats (CRISPR) are present in many bacteria and most archaea (Jansen et al. 2002). Naturally, the system provides resistance against foreign genetic elements (e.g. phages or plasmids) via small noncoding RNAs that are derived from CRISPR loci. In class 2 type II CRISPR systems, the mature crRNA that is base-paired to a trans-activating crRNA (tracrRNA) forms a two-RNA structure that directs the CRISPR-associated proteins (e.g. Cas9 from *Streptococcus pyogenes*) to

75 introduce a double-stranded break (DSB) into the target DNA locus. Site-specific cleavage
76 occurs at locations determined by both base-pairing complementarity between the crRNA and
77 the target protospacer DNA and a short protospacer adjacent motif (PAM) (Jinek et al. 2012).
78 It has been demonstrated that the endonuclease can be programmed by engineering the
79 mature dual-tracrRNA: crRNA as a single RNA chimera (sgRNA for single guide RNA), to
80 cleave specific DNA sites. Thereby, modified versions of the system can be exploited as a
81 biotechnological tool for precise, RNA-programmable genome targeting and editing (Jinek et
82 al. 2012).
83 After the DSB has been introduced, the cell can utilize two major pathways in order to
84 repair the break and survive: homologous recombination (HR) or non-homologous end-joining
85 (NHEJ). In HR, a second intact copy of the broken chromosome segment, homologous to the
86 DSB site, serves as a template for DNA synthesis across the break. In this mechanism, the
87 crucial process of locating and recombining the homologous sequence is performed by RecA
88 (Shuman and Glickman 2007). NHEJ does not rely on a homologous DNA template, as the
89 two DNA ends are rejoined directly together. Most bacteria such as *S. pneumoniae* cannot
90 perform NHEJ, while it is capable to perform HR (Prudhomme et al. 2002; 2014). DSB repair
91 can be used as a way to generate mutants or desired changes to the genome by providing a
92 HR template, and forms the basis of CRISPR engineering (Adli 2018). Indeed, early work,
93 using integrative vectors and tracrRNAs, showed that Cas9 can be used to make markerless
94 gene deletions in *S. pneumoniae* (Jiang et al. 2013).
95 In this study, we set out to establish a CRISPR engineering framework for *S.*
96 *pneumoniae*. Specifically, we constructed a novel replicative plasmid containing a
97 temperature-sensitive origin of replication (facilitating curing of the plasmid) carrying a genetic
98 system for making targeted, marker-less gene knockouts and large genome deletions, which
99 works with high efficiency in *S. pneumoniae*. The here developed plasmid system should be
100 readily transportable to other Gram-positive bacteria as the used origin of replication was
101 shown to be functional in *L. lactis* and *B. subtilis* (Bijlsma et al. 2007). While similar approaches
102 have recently been undertaken to perform genome engineering in certain Gram-positive
103 organisms such as *Enterococcus faecium* (de Maat et al. 2019), Clostridium (IC Cañadas et
104 al. 2019) and *L. lactis* (Guo et al. 2019), a CRISPR-Cas9 gene editing system was not yet
105 available for *S. pneumoniae* and the here described vector has the advantage of being readily
106 curable due to its temperature sensitive origin of replication.
107

108 # Materials and Methods

109 ## Bacterial strains, transformations and growth conditions

110 All pneumococcal strains used in this study are derivatives of the serotype 2 *S. pneumoniae*
111 strain D39V (Avery et al. 1944, Slager et al. 2018). All plasmids where cloned in *NEB® Turbo*
112 *Competent E. coli* (catalog number C2984; New England BioLabs). All the strains are shown
113 in *Table 1*.

114 Table 1: Strain and plasmid list

| *S. pneumoniae* Strains | Relevant Genotype | Reference |
|---|---|---|
| D39V | Serotype 2 strain, wild type | Slager et al.,2018 |
| VL321 | *SPV_2146*-$P_{32}$-*lacZ-chl-aliA* | This study |
| VL588 | *ssbB-luc-kan, cps::chl* | Lab collection |
| VL2172 | pPEPZ-read1-$P_3$-*BsaI-gfp-BsaI*-read2-N701-p7 | Lab collection |
| VL3655 | D39V<br>+pDS05 [*pG*$^+$*host ori*(Ts)-*ermR-cloDF13ori-specR-*$P_{Zn}$*_wtcas9-*$P_3$*_gfp-sgRNA*] | This study |
| VL3656 | *SPV_2146-lacZ-chl-aliA*<br>+pDS07 [*pG*$^+$*host ori*(Ts)-*ermR-cloDF13ori-specR-pZn_wtcas9-*$P_3$*_sgRNA lacZ*] | This study |
| VL3657 | *ΔlacZ*<br>+pDS07 [*pG*$^+$*host ori*(Ts)-*ermR-cloDF13ori-specR-pZn_wtcas9-*$P_3$*_sgRNA lacZ*] | This study |
| VL3658 | *ΔlacZ* | This study |
| VL3659 | *Δcps*<br>+ pDS07 [*pG*$^+$*host ori*(Ts)-*ermR-cloDF13ori-specR-*$P_{Zn}$*_wtcas9-*$P_3$*_sgRNA lacZ*] | This study |
| VL3660 | *Δcps* | This study |
| VL3661 | *Δcps*<br>+ pDS12 [*pG*$^+$*host ori*(Ts)-*ermR-cloDF13ori-specR-*$P_{Zn}$*_wtcas9- *$P_3$*-sgRNA ply*] | This study |
| VL3662 | Δcps, Δply + pDS12 [*pG*$^+$*host ori*(Ts)-*ermR-cloDF13ori-specR-*$P_{Zn}$*_wtcas9- *$P_3$*-sgRNA ply*] | This study |
| VL3663 | *Δcps, Δply* | This study |
| VL3664 | *Δcps, Δply*<br>+ pDS13 [*pG*$^+$*host ori*(Ts)-*ermR-cloDF13ori-specR-*$P_{Zn}$*_wtcas9-*$P_3$*_sgRNA lytA*] | This study |

| | | |
|---|---|---|
| VL3665 | *Δcps, Δply, ΔlytA* + pDS13 [*pG⁺host ori*(Ts)-*ermR-cloDF13ori-specR-* $P_{Zn}$*_wtcas9-*$P_3$*_sgRNA lytA*] | This study |
| **Plasmids** | | |
| pDS05 | *pG⁺host ori*(Ts)-*ermR-cloDF13ori-specR-* $P_{Zn}$*_wtcas9-*$P_3$*_gfp-sgRNA* | This study |
| pDS07 | *pG⁺host ori*(Ts)-*ermR-cloDF13ori-specR-* $P_{Zn}$*_wtcas9-*$P_3$*_sgRNA lacZ* | This study |
| pDS12 | *pG⁺host ori*(Ts)-*ermR-cloDF13ori-specR-* $P_{Zn}$ *_wtcas9-*$P_3$*_sgRNA ply* | This study |
| pDS13 | *pG⁺host ori*(Ts)-*ermR-cloDF13ori-specR-* $P_{Zn}$*_wtcas9-*$P_3$*_sgRNA lytA* | This study |
| pRAS2 | pJWV01-$P_{32}$-*lacZ* | Lab collection |

*S. pneumoniae* was grown either at 30°C, 37°C or 40°C (indicated) without shaking in liquid C+Y medium adapted from Adams and Roe (Martin et al. 1995) and contained the following compounds: adenosine (68.2 µM), uridine (74.6 µM), L-asparagine (302 µM), L-cysteine (84.6 µM), L-glutamine (137 µM), L-tryptophan (26.8 µM), casein hydrolysate (4.56 g L-1 ), BSA (729 mg L-1 ), biotin (2.24 µM), nicotinic acid (4.44 µM), pyridoxine (3.10 µM), calcium pantothenate (4.59 µM), thiamin (1.73 µM), riboflavin (0.678 µM), choline (43.7 µM), CaCl2 (103 µM), K2HPO4 (44.5 mM), MgCl2 (2.24 mM), FeSO4 (1.64 µM), CuSO4 (1.82 µM), ZnSO4 (1.58 µM), MnCl2 (1.29 µM), glucose (10.1 mM), sodium pyruvate (2.48 mM), saccharose (861 µM), sodium acetate (22.2 mM) and yeast extract (2.28 g L-1).
*E. coli* strains were cultivated in LB at 37°C with shaking. When appropriate, 100 µg/ml spectinomycin (spec) was added.

## Transformation

To transform the different plasmid variants into *S. pneumoniae*, cells were grown in C+Y medium (pH 6.8) at 37 °C to an $OD_{595}$ of 0.1. Then, cells were treated for 12 min at 37°C with synthetic CSP-1 (100 ng $mL^{-1}$) and incubated for 20 min at 30°C with the plasmid. After incubation, cells were grown in C+Y medium at the permissive temperature of 30°C for 120 min. *S. pneumoniae* transformants were selected by plating inside Columbia agar supplemented with 3% of defibrinated sheep blood (Thermo Scientific), followed by antibiotic selection, using erythromycin at concentration 0.25 µg/ml. Plates were incubated at 30°C.

To transform the HR template, cells were grown in C+Y medium (pH 6.8) at 30°C with 0.1 µg/ml erythromycin to an $OD_{595}$ of 0.1. Then, cells were treated for 12 min at 37°C with

138 synthetic $CSP^{-1}$ (100 ng $mL^{-1}$) and incubated for 20 min at 30°C with the HR template. After
139 incubation, cells were grown in C+Y medium at 30°C for 20 min. Transformants were selected
140 by plating inside Columbia agar supplemented with 3% of defibrinated sheep blood (Thermo
141 Scientific), followed by CRISPR-mediated counter selection, using 1 mM $ZnCl_2/MnSO_4$. Plates
142 were incubated at 30°C. Correct transformation was verified by PCR and sequencing. Working
143 stocks of cells were prepared by growing cells in C+Y (pH 6.8), until an $OD_{595}$ of 0.4. Cells
144 were collected by centrifugation (1595 × *g* for 10 min) and resuspended in fresh C+Y medium
145 with 15% glycerol and stored at −80°C.
146

147 ## Plasmid curing

148 After the plasmid has been transformed into pneumococcus and successful deletion has been
149 performed with the HR template and CRISPR-mediated counter selection, the plasmid can be
150 eliminated from the pneumococcal cells. To achieve that, we first grow the strain with the
151 plasmid at the non-permissive temperature, 40°C in C+Y, with a starting inoculum 1/10.000.
152 Next, we plate the liquid culture in Columbia blood agar in several dilutions, to obtain single
153 colonies after overnight incubation at 40°C. Single colonies were screened and 99% of them
154 had successfully cured the plasmid from the strain.
155

156 ## Recombinant DNA techniques

157 Oligonucleotides were ordered from Sigma and are listed in Table 2. Phanta Max Super-
158 Fidelity DNA Polymerase (Vazyme) was used in PCR amplifications, restriction enzymes
159 (ThermoFisher Scientific) were used for digestions and T4 DNA Ligase (Vazyme) was used
160 for ligations.
161

162 *Table 2: Oligonucleotides used in this study*

| Name | Sequence (5′–3′); restriction site (underlined) |
|---|---|
| OVL4739_pGh F | CTCTCA<u>CACCTGC</u>CTGTCAATCGCAACATCAAACCAAAATA AAAAC |
| OVL4740_pGh R | CTCTCA<u>CACCTGC</u>CTGTTTCAAAAGCGACTCATAGAATTAT TTC |
| OVL4741_pCDF-1b F | CTCTCA<u>CACCTGC</u>CGTATGAATCTAGAGCGGTTCAGTAGA AAAG |
| OVL4742_pCDF-1b R | CTCTCA<u>CACCTGCC</u>GTATACTTGAACGAATTGTTAGACATT ATTTG |

| | |
|---|---|
| OVL4743_wtcas9 F | AGATGGCACCTGCCAGAAGTACAAGCACTTTGGGACGTTCTCCCTTAG |
| OVL4744_wtcas9 R | AGATGGCACCTGCCAGAACGCTAAATACGCTTCACAGTTTCTTCTTC |
| OVL4745_gRNA F | CTCTCACACCTGCTCACGCGTATAAGAGACAGCCATTCTACAG |
| OVL4746_gRNA R | CTCTCACACCTGCTCACATTGAGACAGAAAAAAAGCACCGACTC |
| OVL2132_GG-lacZ-F | TATAGGATGAAGACCAGCCCTTCC |
| OVL2133_GG-lacZ-R | AAACGGAAGGGCTGGTCTTCATCC |
| OVL2142_lin pGh R | CCTAGGTCTCATATAGTTATTATACCAGGG |
| OVL2143_lin pGh F | GTAAGGTCTCGGTTTAAGAGCTATG |
| OVL2250_GG-ply-F | TATACCGAGTTGTAACAGGCAAGG |
| OVL2251_GG-ply-R | AAACCCTTGCCTGTTACAACTCGG |
| OVL2813_GG-sgRNAlytA-F | TATAAACCAAAGAAGAGTTCATGA |
| OVL2814_GG-sgRNAlytA-R | AAACTCATGAACTCTTCTTTGGTT |
| rfbD-F | TCATGACCTACCTAGCTGAAAATCG |
| rfbD-R+BglII | GGCCAGATCTAAGCGCCCAATAACGAAGTATATTG |
| P32-lacZ-BglII | ATGCAGATCTAGGCCGGCCGATATGATAAG |
| lacZ_R-AscI | ATCACGGGCGCGCCTTATTTTTGACACCAGACCAACTG |
| cam-F+AscI | ACGTGGCGCGCCAGGAGGCATATCAAATGAAC |
| del_CSP.dn-R | GATAGAGACGAGCTGCTGTAAGGC |

163

164 ## Strain construction

165 pDS05 (*pG$^+$host ori(Ts)-ermR-cloDF13ori-specR-P$_{Zn}$_wtcas9-P$_3$_gfp-sgRNA*). Gram-positive,
166 temperature sensitive origin of replication *pG$^+$host* (Maguin et al. 1992) and gene *ermR*, which
167 confers resistance to erythromycin, were amplified from plasmid pGh9::ISS1 (Maguin et al.
168 1996) using the primers OVL4739_pGh F and OVL4740_pGh R (fragment 1). Gram-negative
169 origin of replication *CloDF13* (CDF) and gene *specR*, which confers resistance to
170 spectinomycin, were amplified from plasmid pCDF-1b (Nijkamp et al. 1986) with primers
171 OVL4741_pCDF-1b F and OVL4742_pCDF-1b R (fragment 2). The gene which encodes
172 wtCas9 under the control of the Zinc-inducible promoter was amplified from plasmid
173 pJWV102-spCas9wt(van Raaphorst, Kjos, and Veening 2017), using the primers
174 OVL4743_wtcas9 F and OVL4744_wtcas9 R (fragment 3). The sgRNA sequence in which the

175 20 base-pairing region of the sgRNA is replaced by the *gfp* gene was amplified from strain
176 VL2172 with primers OVL4745_gRNA F and OVL4746_gRNA R (fragment 4). The four
177 fragments were digested all together with restriction enzyme *AarI* and ligated. The ligation
178 product was transformed into *E. coli* NEB Turbo and transformants were selected on LB agar
179 with spectinomycin. Correct assembly was confirmed by PCR and sequencing.
180
181 pDS07 (*pG$^+$host ori(Ts)-ermR-cloDF13ori-specR-$P_{Zn}$_wtcas9-$P_3$_sgRNA lacZ*). pDS05 was
182 amplified with primers OVL2143_lin pGh F and OVL2142_lin pGh R (fragment 1). Spacer
183 sequence of *sgRNA lacZ* was constructed by annealing primers OVL2132_GG-lacZ-F and
184 OVL2133_GG-lacZ-R. Amplified pDS05 was digested with restriction enzyme *BsaI* and ligated
185 with the annealed oligos. The ligation product was transformed into *E. coli* NEB Turbo and
186 transformants were selected on LB agar with spectinomycin. Correct assembly was confirmed
187 by PCR and sequencing.
188
189 pDS12 (*pG$^+$host ori(Ts)-ermR-cloDF13ori-specR- $P_{Zn}$ _wtcas9-$P_3$_sgRNA ply*). pDS05 was
190 amplified with primers OVL2143_lin pGh F and OVL2142_lin pGh R. Spacer sequence of
191 *sgRNA ply* was constructed by annealing primers OVL2250_GG-ply-F and OVL2251_GG-ply-
192 R. Amplified pDS05 was digested with restriction enzyme *BsaI* and ligated with the annealed
193 oligos. The ligation product was transformed into *E. coli* NEB Turbo and transformants were
194 selected on LB agar with spectinomycin. Correct assembly was confirmed by PCR and
195 sequencing.
196
197 pDS13 (*pG$^+$host ori(Ts)-ermR-cloDF13ori-specR- $P_{Zn}$ _wtcas9-$P_3$_sgRNA lytA*). pDS05 was
198 amplified with primers OVL2143_lin pGh F and OVL2142_lin pGh R. Spacer sequence of
199 *sgRNA lytA* was constructed by annealing primers OVL2813_GG-sgRNAlytA-F and
200 OVL2814_GG-sgRNAlytA-R. Amplified pDS05 was digested with restriction enzyme *BsaI* and
201 ligated with the annealed oligos. The ligation product was transformed into *E. coli* NEB Turbo
202 and transformants were selected on LB agar with spectinomycin. Correct assembly was
203 confirmed by PCR and sequencing.
204
205 VL321 (*SPV_2146*-$P_{32}$-*lacZ-chl-aliA*). *rfbD* and *SPV_2146* were amplified from chromosomal
206 DNA with primers rfbD-F and rfbD-R+BglII (fragment 1). $P_{32}$-*lacZ* was amplified from pRAS2
207 (lab collection) with primers P32-lacZ-BglII and lacZ_R-AscI (fragment 2). A chloramphenicol
208 resistance marker and *aliA* was amplified from strain VL588 with primers cam-F+AscI and
209 del_CSP.dn-R. Fragment 1 was digested with restriction enzyme *Bgl*II, fragment 2 was
210 digested with restriction enzymes *Bgl*II and *Asc*I and fragment 3 was digested with restriction

211 enzyme *Asc*I. All three fragments were ligated together, and the ligation product was
212 transformed into D39V. Correct assembly was confirmed by PCR and sequencing.
213
214 VL3655 (D39V + pDS05 [*pG*$^{+}$*host ori(Ts)-ermR-cloDF13ori-specR-P*$_{Zn}$*_wtcas9-P*$_{3}$*_gfp-*
215 *sgRNA*]). Plasmid pDS05 was transformed into D39V and transformants were selected on
216 Columbia blood agar with erythromycin to produce the strain VL3655. The presence of the
217 plasmid was confirmed by PCR and plasmid extraction.
218
219 VL3656 (*SPV_2146-lacZ-chl-aliA* + pDS07 [*pG*$^{+}$*host ori(Ts)-ermR-cloDF13ori-specR-*
220 *p*$_{Zn}$*_wtcas9-P*$_{3}$*_sgRNA lacZ*]). Plasmid pDS07, was transformed into VL321(*SPV_2146-lacZ-*
221 *chl-aliA*) and transformants were selected on Columbia blood agar with erythromycin to
222 produce the strain VL3656. The presence of the plasmid was confirmed by PCR and plasmid
223 extraction.
224
225 VL3657 (Δ*lacZ* + pDS07 [*pG*$^{+}$*host ori(Ts)-ermR-cloDF13ori-specR-P*$_{Zn}$*_wtcas9-P*$_{3}$*_sgRNA*
226 *lacZ*]). HR template Δ*lacZ* was transformed into VL3656 and transformants were selected on
227 Columbia blood agar with $ZnCl_2$/$MnSO_4$ to produce the strain VL3657. Correct integration was
228 confirmed by PCR.
229
230 VL3658 (*ΔlacZ*). Strain VL3657 was cured from the plasmid, as described, resulting in strain
231 VL3658.
232
233 VL3659 (Δ*cps* + pDS07 [*pG*$^{+}$*host ori(Ts)-ermR-cloDF13ori-specR-P*$_{Zn}$*_wtcas9-P*$_{3}$*_sgRNA*
234 *lacZ*]). HR template Δ*cps* was transformed into VL3656 and transformants were selected on
235 Columbia blood agar with $ZnCl_2$/$MnSO_4$ to produce the strain VL3659. Correct integration was
236 confirmed by PCR.
237
238 VL3660 (*Δcps*). Strain VL3659 was cured from the plasmid, as described, resulting in strain
239 VL3660.
240
241 VL3661 (*Δcps* + pDS12 [*pG*$^{+}$*host ori(Ts)-ermR-cloDF13ori-specR-P*$_{Zn}$*_wtcas9- P*$_{3}$*-sgRNA*
242 *ply*]). Plasmid pDS12 was transformed into VL3660 and transformants were selected on
243 Columbia blood agar with erythromycin to produce the strain VL3661. The presence of the
244 plasmid was confirmed by PCR and plasmid extraction.
245
246 VL3662 (*Δcps, Δply* + pDS12 [*pG*$^{+}$*host ori(Ts)-ermR-cloDF13ori-specR-P*$_{Zn}$*_wtcas9-*
247 *P*$_{3}$*_sgRNA ply*]). HR template Δ*ply* was transformed into VL3661 and transformants were

248 selected on Columbia blood agar with $ZnCl_2/MnSO_4$ to produce the strain VL3662. Correct
249 integration was confirmed by PCR.
250
251 VL3663 (*Δcps, Δply*). Strain VL3662 was cured from the plasmid, as described, resulting in
252 strain VL3663.
253
254 VL3664 (*Δcps, Δply* + pDS13 [*pG*$^+$*host ori(Ts)-ermR-cloDF13ori-specR-*$P_{Zn}$*_wtcas9-*
255 $P_3$*_sgRNA lytA*]). Plasmid pDS13 was transformed into VL3663 and transformants were
256 selected on Columbia blood agar with erythromycin to produce the strain VL3664. The
257 presence of the plasmid was confirmed by PCR and plasmid extraction.
258
259 VL3665 (*Δcps, Δply, ΔlytA* + pDS13 [*pG*$^+$*host ori(Ts)-ermR-cloDF13ori-specR-* $P_{Zn}$*_wtcas9-*
260 $P_3$*_sgRNA lytA*]). HR template Δ*lytA* was transformed into VL3664 and transformants were
261 selected on Columbia blood agar with $ZnCl_2/MnSO_4$ to produce the strain VL3665. Correct
262 integration was confirmed by PCR.
263

264 ## Microscopy

265 *S. pneumoniae* cells were stored as exponential phase frozen cultures. Frozen stock was
266 inoculated 1:100 in C+Y medium and pre-grown to OD600 ~ 0.1. Cells were diluted once
267 again 1:100 in fresh C+Y (with antibiotic, if applicable) and grown to exponential phase to
268 achieve balanced growth.
269 Cells were grown as described above to achieve balanced growth and subsequently
270 concentrated and brought onto a multi test slide carrying a thin layer of 1.2% agarose in C+Y.
271 Imaging was performed on Fluorescence microscopy was performed on a Leica DMi8 through
272 a 100x phase contrast objective (NA 1.40) with a SOLA Light Engine (lumencor) light source.
273 Light was filtered through external excitation filters 470/40 nm (Chroma ET470/40x) for
274 visualization of GFP. Light passed through a cube (Leica 11536022) with a GFP/RFP
275 polychroic mirror (498/564 nm). External emission filters used were from Chroma ET520/40m.
276 Images were captured using LasX software (Leica) and exported to ImageJ(Schneider,
277 Rasband, and Eliceiri 2012) for final preparation.
278 Cell outlines were detected using MicrobeJ (Ducret et al. 2016). For all microscopy
279 experiments, random image frames were used for analysis. The cell outline, object detection,
280 and fluorescent intensity data were further processed using the R-package BactMAP
281 (Raaphorst, Kjos, and Veening 2020).
282

283 ## Transformation efficiency assays

284 To calculate the transformation efficiency, 1 µg/ml of PCR product of DNA fragment containing
285 the upstream and the downstream region of the deletion target was added. Serial dilutions
286 were plated either with or without 1 mM $ZnCl_2$/$MnSO_4$, and the transformation efficiency was
287 calculated by dividing the number of transformants by the total viable count. All transformation
288 efficiency values are averages of three biologically independent experiments.
289

290 ## Whole genome sequencing and variant analysis

291 Genomic DNA was isolated using the FastPure Bacteria DNA Isolation Mini Kit (Vazyme)
292 according to the manufacturers' protocol and sent for whole genome sequencing. Illumina
293 library prep and sequencing were carried out by Novogene (sequencing in PE150 mode).
294 Reads were trimmed using Trimmomatic (Bolger, Lohse, and Usadel 2014), then assembled
295 using SPAdes (Nurk et al. 2013) and contigs were reordered in Mauve (Darling et al. 2004)
296 using D39V as a reference (Slager et al NAR 2018). Reads were mapped onto the scaffold
297 using bwa (H. Li and Durbin 2009), read depth was determined in samtools (H. Li et al. 2009),
298 and plotted in R (Team 2014). In order to detect small variants, raw reads were mapped onto
299 the reference using bwa, and variants were detected using Freebayes (Garrison and Marth
300 2012). Potential variants with PHRED scores greater than 30 were filtered out on DP >5 using
301 vcflib (Garrison n.d.), then intersected with the D39V annotation using Bedtools (Quinlan and
302 Hall 2010)
303

# Results

## A novel *S. pneumoniae* replicative plasmid that carries the CRISPR-Cas9 system

First, a replicative plasmid was designed and constructed (Figure 1), by combining PCR amplified genomic and plasmid parts. The main idea behind the choice for individual vector components relied in creating a platform with the CRISPR-Cas9 system in *S. pneumoniae* while at the same time allowing for plasmid propagation in both Gram-positive and Gram-negative hosts. The modular vector consists of six individual components. Two origins of replication; the high-copy pG$^{+}$host replicon, which is a replication thermosensitive derivative of pWV01 (Otto et al. 1982) that in *L. lactis*, (and other Gram-positive bacteria) replicates at 28$^{\circ C}$ but is lost above 37$^{\circ C}$, and the low-copy *CloDF13* (CDF) replicon for propagation in *E. coli.* By combining these two origins of replication, it ensures low copy numbers at 37$^{\circ C}$ in *E. coli* thereby preventing toxicity of the CRISPR-Cas9 system while cloning. Additionally, it has the gene which encodes wtCas9 under the control of the Zinc-inducible promoter Pzn (Eberhardt et al. 2009) (plasmid pDS05) and genes conferring spectinomycin (*E. coli*) and erythromycin (*S. pneumoniae*) resistance. Finally, it has the strong synthetic constitutive P3 promoter (Sorg et al. 2015) driving the sgRNA sequence in which the 20 base-pairing region of the sgRNA is replaced by the *gfp* gene flanked by two *Bsa*I restriction sites. This allows for easy replacement of *gfp* by the spacer sequence of sgRNA with golden gate cloning. This way, successful cloning of the sgRNA allows for easy selection by absence of GFP fluorescence, giving us a versatile vector for different sgRNAs (see below).

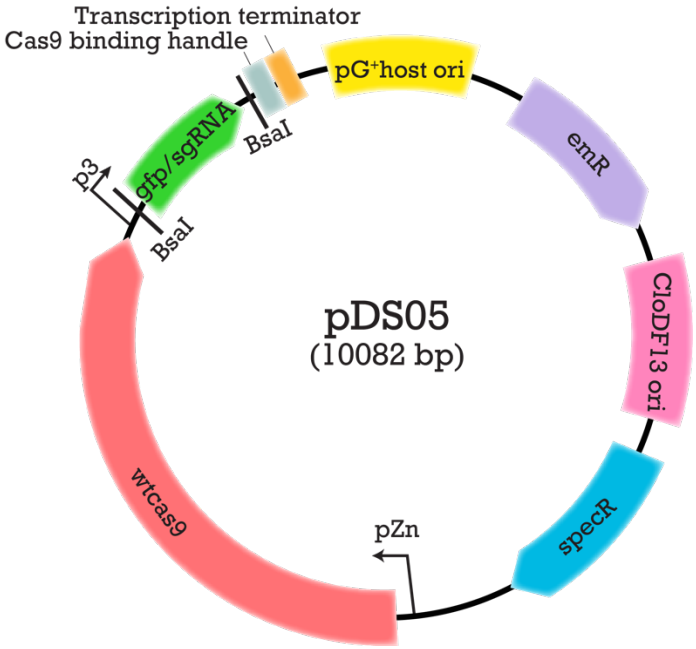
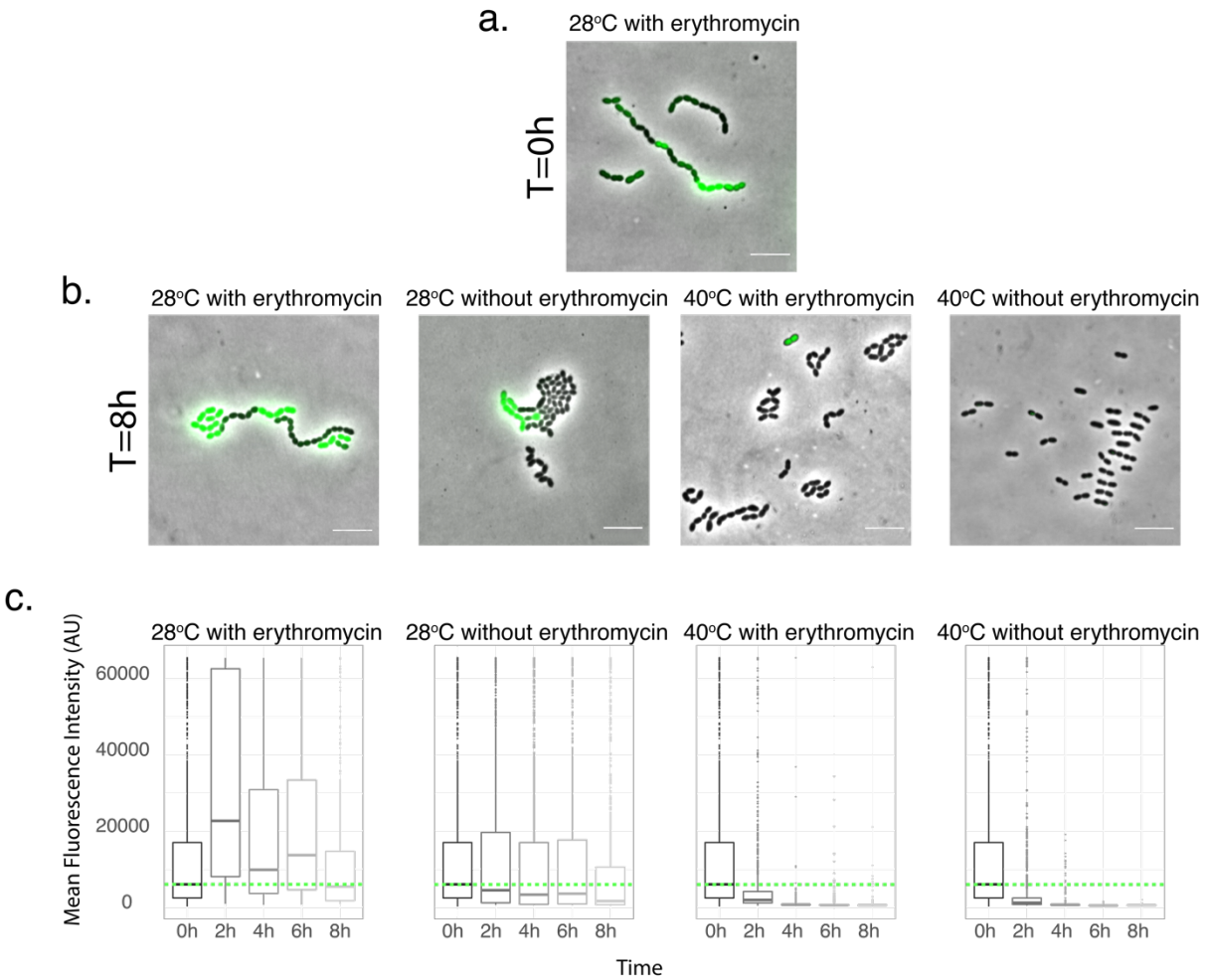



**Figure 1: Schematic representation of plasmid pDS05**

327 ### Successful plasmid propagation and plasmid curing in *S. pneumoniae*

328 To test that the newly constructed plasmid was being replicated and genes were expressed
329 in *S. pneumoniae*, we grew strain VL3655 (carrying plasmid pDS05, see Fig. 1), which
330 encodes GFP, in C+Y medium at 28°C in the presence of erythromycin. Fluorescence
331 microscopy demonstrated that all cells produced GFP, although significant cell-to-cell
332 variability was observed (Figure 2a). GFP intensity levels were determined in exponentially
333 growing cells (balanced growth). Additionally, cells pre-grown in 28°C in presence of
334 erythromycin (T=0) were split and grown under four different conditions. The permissive 28°C
335 with and without erythromycin in the growth medium and the non-permissive 40°C with and
336 without erythromycin. Note that growth was balanced by re-diluting exponentially growing cells
337 several times. Cells were being collected every two hours for 8 hours and GFP intensity levels
338 were determined using fluorescence microscopy and images were analyzed using MicrobeJ
339 and BactMAP (Ducret et al. 2016; Raaphorst et al. 2020) (Figure 2b and c). The results show
340 that GFP levels and, by extension plasmid copy number, stay stable at 28°C with
341 erythromycin, and slowly decreases in the absence of antibiotic pressure. Furthermore, GFP
342 levels decrease significantly in cells grown at 40°C, confirming that this is a non-permissive
343 temperature for propagation of the plasmid. Absence of antibiotic pressure seems also to
344 facilitate the decrease of the intensity levels of GFP, suggesting that the plasmid gets
345 eliminated successfully under these conditions.




346

347 **Figure 2: Microscopy analysis of strain VL3655 (D39V, pDS05).** Overlay of GFP signals
348 with phase contrast image shows GFP expression. a. Preculture grown at 28°C with
349 erythromycin (T=0h), b. Images are shown of cells grown for 8h as exponentially growing cells
350 (balanced growth) in four different conditions: 28°C with erythromycin, 28°C without
351 erythromycin, 40°C with erythromycin, and 40°C without erythromycin. Scale bar in all images
352 = 6 µm. c. Quantification of mean fluorescence intensity of GFP of cells grown under four
353 different conditions: 28°C with erythromycin, 28°C without erythromycin, 40°C with
354 erythromycin, and 40°C without erythromycin) in time points 0, 2, 4, 6 and 8 hours after dilution
355 from the 28°C with erythromycin condition. Fluorescence microscopy of ±1000 cells per
356 condition per time point were quantified and analyzed using MicrobeJ and BactMap and
357 plotted as box plots, (box size and line represent the average intensity per cell) (see Materials
358 and Methods). The green dotted horizontal line indicates the mean fluorescence of cells from
359 the preculture harboring pDS05.

360

## CRISPR/Cas9-Mediated Counterselection

Once a deletion target has been selected, the plasmid with the specific sgRNA needs to be constructed. The targeting of Cas9 to a locus of interest is achieved by cloning two annealed 24-bp DNA oligonucleotides (containing the 20 bp protospacer element) into the sgRNA backbone that matches the specified locus. First, *gfp* is removed from pDS05 by digesting with *Bsa*I. Complementary oligos that carry the spacer sequence are annealed together. They are designed in a way that after annealing, they have overhangs complementary with those left after digestion of the backbone, as previously described (Liu et al. 2020) (Figure 3). The desired plasmid is obtained after ligation and transformed to *E. coli*. False positive transformants are easily identified, since they still carry the *gfp* and produce detectable fluorescent green colonies.

After isolating the plasmid from *E. coli*, it needs to enter the pneumococcal cells. To achieve this, competence is induced by adding synthetic CSP and the transformation machinery is utilized. Note that competence-dependent transformation with a replicative plasmid is less efficient than transforming with linear homologous DNA (Johnston et al. 2014), so transformation efficiencies with pDS05 are typically low. However, in this case, only 1 successful transformant is required. Next, an HR template is constructed that consists of the upstream and downstream region of the deletion target. Following again induction of competence by CSP, the pneumococcal cells harboring the pDS05-derivative are transformed with this template. Transformants are selected by plating with $ZnCl_2/MnSO_4$, the inducer of Cas9, offering CRISPR-mediated counter-selection. Only cells that have taken up and integrated the HR template, thereby eliminated the recognition target of the sgRNA/Cas9 complex, would be able to survive, while untransformed cells will undergo DNA cleavage and die (Figure 4).

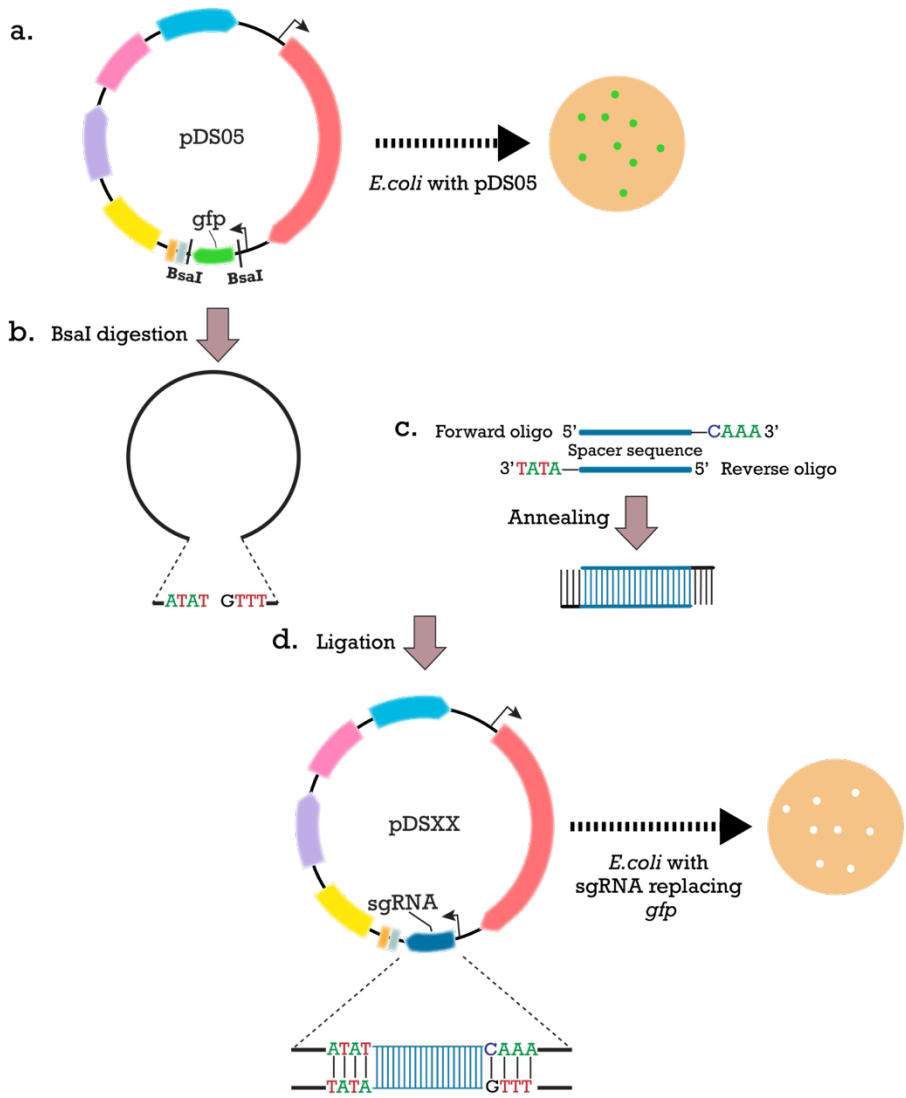



**Figure 3: Workflow of sgRNA replacement.** a. pDS05 was designed to facilitate easy replacement of *gfp* by the spacer sequence of the desired sgRNA with golden gate cloning, allowing also for detection of false positive transformants. *gfp,* which encodes a green fluorescent protein, is in place of the spacer sequence of sgRNA and flanked by *Bsa*I sites. *E. coli* with pDS05 produces green fluorescent colonies b. *Bsa*I digestion of the vector exposes 4 nt overhangs c. For each sgRNA, forward and reverse oligos were designed, as a reverse complement of each other, which after being annealed together, were containing the 20 bp spacer sequence and 4 nt overhangs, that can be specifically annealed with the digested vector. d. Ligation of the digested vector with the sgRNA annealed product was transformed into *E. coli*, producing white colonies.

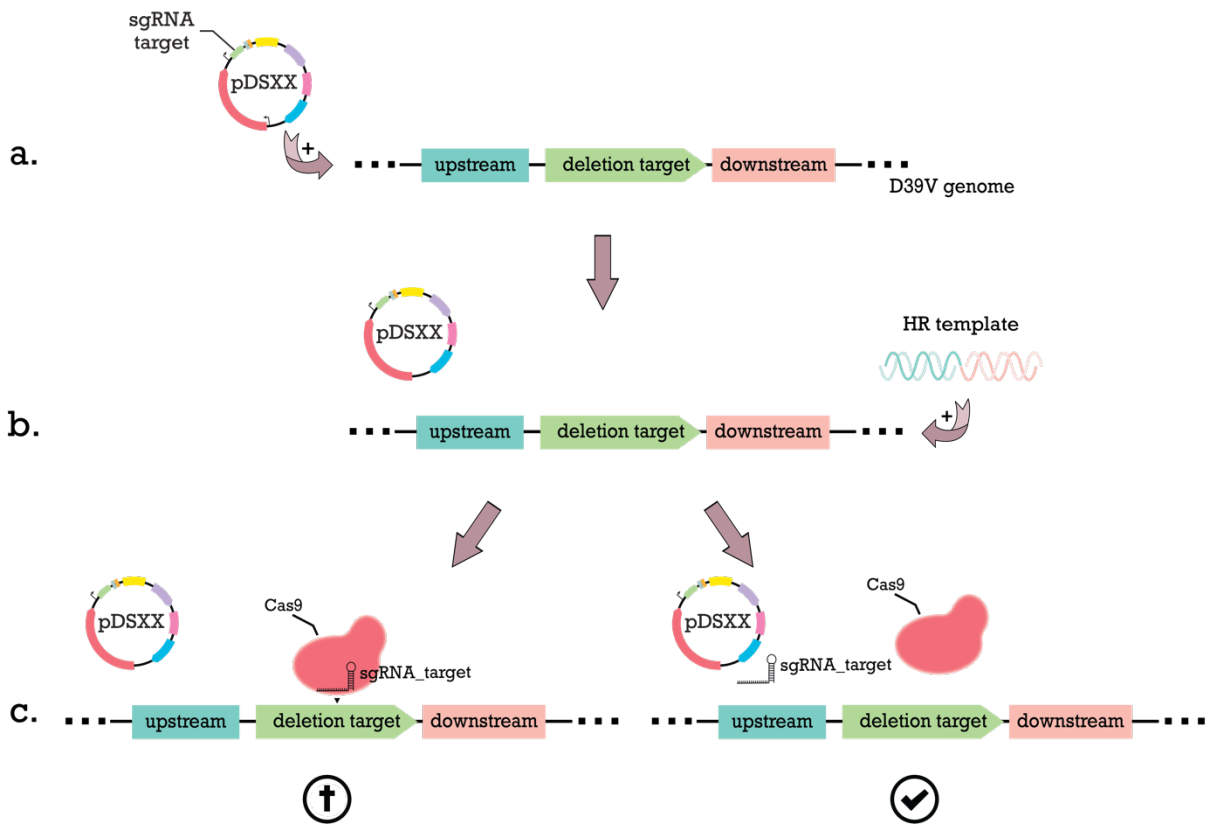



**Figure 4: Workflow for markerless deletions.** a. Uptake of the plasmid by the strain with the sgRNA sequence for the desired deletion. b. Transformation of a homology recombination (HR) template consisting of the ligation of the upstream and downstream region of the deletion target. c. Plating transformants with $Zn^{2+}$ to induce expression of Cas9. Only the cells that have taken up the HR template eliminating the recognition target are able to survive.

405 Deleting genes and large chromosomal regions from the *S. pneumoniae* genome

406 To assess the efficiency of the system, we first constructed a strain (strain VL3656; Figure 5a)
407 in which we placed the *E. coli lacZ* gene under a constitutive promoter behind the *S.*
408 *pneumoniae* D39V *cps* locus (encoding the capsule). *lacZ* encodes for a β-galactosidase that
409 hydrolyzes X-gal to produce a blue product, allowing for blue/white screening on plates.
410 Colonies with blue color would still carry the *lacZ* gene, while colonies with the standard
411 white/green (on blood agar) color would indicate that the gene has been deleted from the
412 chromosome.

413 Strain VL3656 also carries the pDS07 plasmid, which contains a sgRNA targeting *lacZ.*
414 Next, we constructed an HR template that consisted of the 1000 bp upstream and 1000 bp
415 downstream region of *lacZ* (excluding *lacZ*) (**Figure 5**a) and we transformed VL3656 with this
416 template. Transformants were selected by plating on agar containing $ZnCl_2$ to induce
417 expression of Cas9.

418 After transformation of strain VL3656 with the HR template, transformation efficiency
419 was calculated (Figure 5b). The CRISPR-mediated counter-selection, offered by the system,
420 worked successfully. The selection efficiency was high and almost all colonies in the
421 transformation where the HR template was given and Cas9 was induced had their original
422 color, indicating the *lacZ* gene has been successfully deleted.

423 Transformants were tested for correct deletion of *lacZ* by colony PCR. The primers
424 used were binding 1000bp upstream and downstream of *lacZ*, setting the correct PCR product
425 of the successful deletion at the 2000bp (5000bp if *lacZ* was not deleted) (Figure 5c). All the
426 tested colonies had the expected product demonstrating successful deletion of *lacZ*, resulting
427 in strain VL3657 (Figure 5d).

428 Additionally, we also used the system to delete an even larger chromosomal fragment.
429 For this, we targeted the operon that encodes the capsule and the *lacZ* gene that had been
430 inserted downstream of it, which is around 24Kbp long, allowing for blue/white screening.
431 Once again, selection efficiency was very high and almost all colonies had their original color
432 (Figure 5b). Colony PCR verified correct deletion of the *cps-lacZ* chromosomal region (see
433 below).

434 Using the same HR template to delete *lacZ,* we also performed transformation assays
435 without the counterselection offered, by inducing our CRISPR system (Figure 5f). Thousands
436 of colonies needed to be screened to find successful transformants with the original colony
437 color, among the blue colonies. In contrast, by using the system, almost with absolute success
438 rate, all the colonies on our plates are the correct transformants, demonstrating how efficient
439 our system is to easily select edited cells.
440

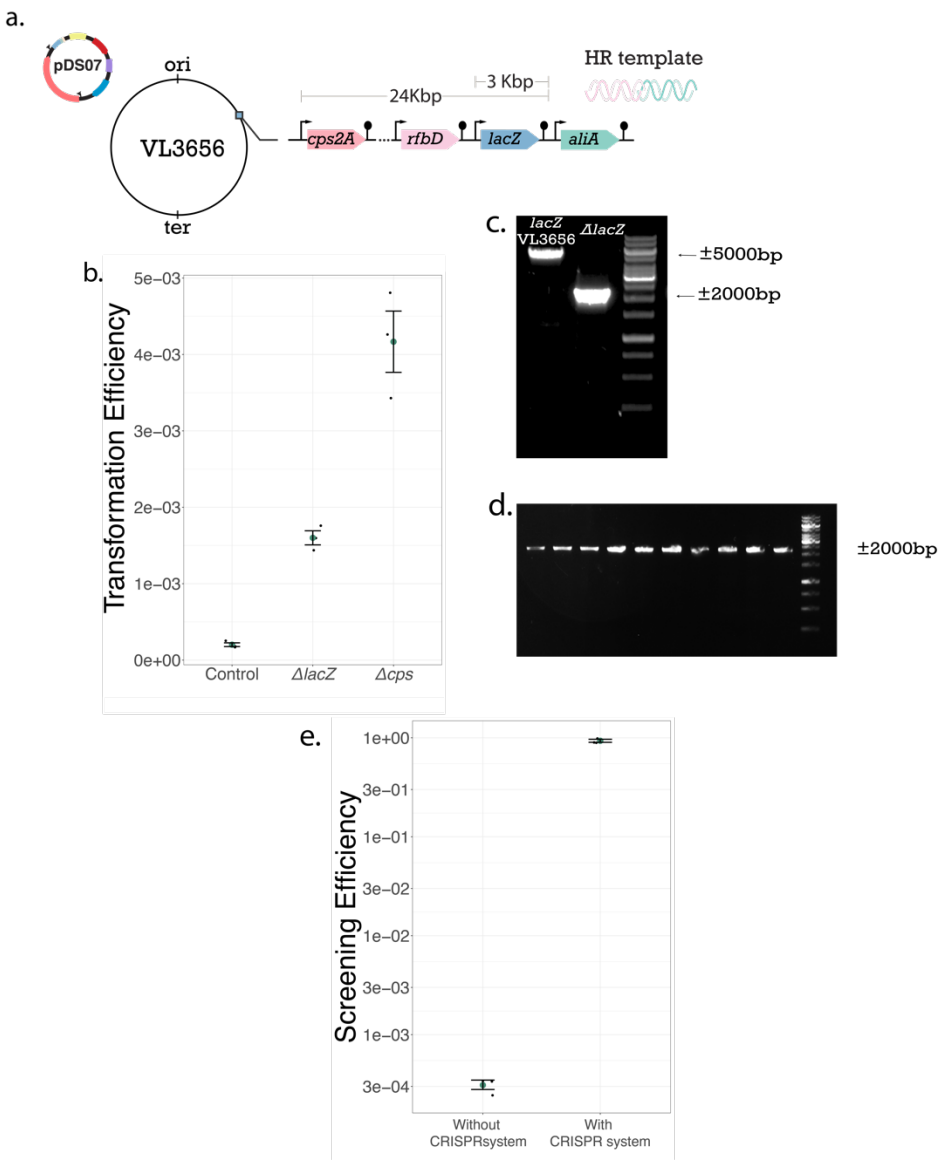



**Figure 5: Genome editing in *S. pneumoniae* using CRISPR/Cas9.** a. Schematic representation of strain VL3656. The *lacZ* gene has been inserted downstream the capsule operon and a version of the plasmid with a sgRNA targeting *lacZ* has been transformed in the strain. Control is the transformation assay of strain VL3656 in the absence of HR template DNA. b. Transformation efficiency of *lacZ* and capsule operon deletion. The transformation efficiency was calculated by dividing the total number of cells as counted on plates without Cas9 inducer (1 mM $Zn^{2+}$) by the number of colonies in the presence of inducer. c. Colony PCR analysis of expected sizes. d. Eight randomly selected transformants of *lacZ* deletion. e. Efficiency of successful transformants screened for integration of the *lacZ* deletion when using no selection and when using the CRISPR system. Data represent the average of three independent experiments (± SE).

453 ## Consecutive deletions of virulence factors of *S. pneumoniae*

454 Once the capsule operon and *lacZ* were removed from the chromosome, it was confirmed by
455 colony PCR. All tested colonies demonstrated the expected PCR product. One such colony
456 was picked resulting in strain VL3659. Next, we grew the new strain at the non-permissive
457 temperature (40°C), eliminating the plasmid, resulting in strain VL3660 (*Δcps*).

458 To examine whether the system could be used in multiple rounds of genome editing,
459 we attempted to delete the virulence factor pneumolysin. To delete the *ply* gene, we designed
460 a sgRNA targeting pneumolysin and constructed plasmid pDS12, which we transformed into
461 VL3660. Following the same procedure as used to delete the *cps* operon, we deleted *ply*.
462 Again, to confirm the successful deletion, the same principle for the primers set was used. All
463 the colonies from the transformation plate had the expected PCR product demonstrating
464 extremely high selection efficiency using the CRISPR-Cas9 system. Finally, following the
465 same strategy, we also deleted another important virulence factor, *lytA* resulting in strain
466 VL3665 (**Figure 6**d; *Δcps, Δply, ΔlytA* + pDS13).

467

468 ## Cas9-dependent genome editing is specific without evidence for off-target cutting in *S.*
469 *pneumoniae*

470 After three consecutive deletions, using our novel plasmid with the CRISPR-Cas9 system, the
471 final result was strain VL3665. It has been previously shown that Cas9 tolerates mismatches
472 between guide RNA and target DNA at different positions in a sequence-dependent manner,
473 resulting in off-target DSB (Hsu et al. 2013). To examine the fidelity of our CRISPR system
474 and whether there were detectable genome-wide off-target effects, we performed whole-
475 genome sequencing (WGS). The analysis detected one single SNP in the genome, in the
476 gene *psaA* (*SPV_1463*), a manganese ABC transporter. The mutation results in a D137E
477 amino acid change. Using Sanger sequencing we confirmed that this SNP occurred only in
478 the last strain of the consecutive deletions and it has not been present in the intermediate
479 steps. There is no evidence to believe that this mutation is associated with an off-target effect
480 as the sequence surrounding the SNP is completely different from the used sgRNA present in
481 plasmid pDS13 and most probably happened randomly during growth, without affecting the
482 fitness (**Error! Reference source not found.**).

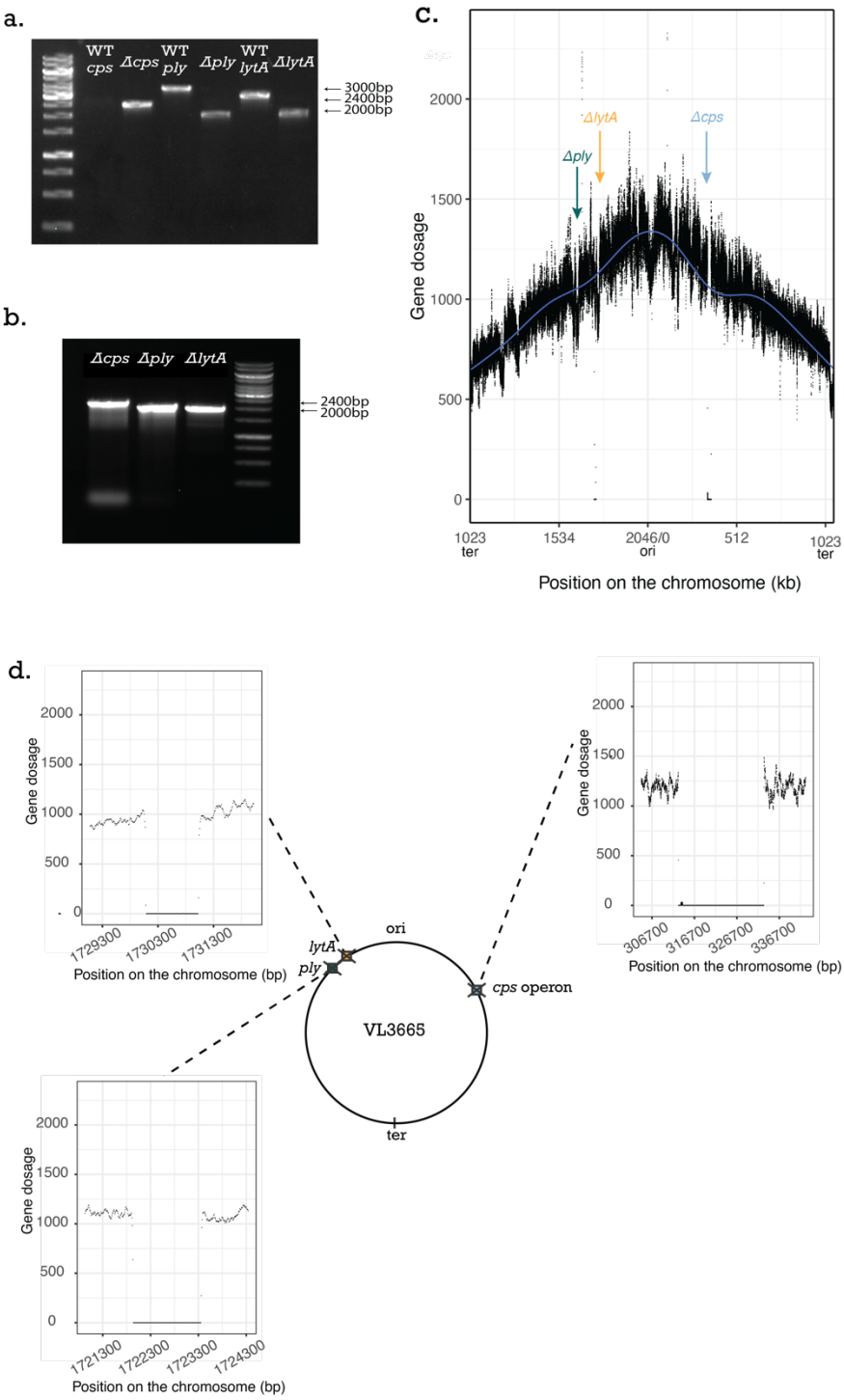
483

484

485

486

487 On strain VL3665, with the three major virulence factors deleted, we performed the final
488 confirmations. Colony PCR showed that the chromosomal fragments have successfully been
489 deleted from the chromosome (Figure 6a and b). Additionally, reads from whole-genome
490 sequencing were competitively mapped onto the reference genome, our wild type lab strain
491 D39V (Figure 6c). Direct comparison between the genomes reveals the three chromosomal
492 positions that the deletions have taken place, since in these positions, the chromosomal
493 dosage drops. Therefore, we confirmed that we had successfully performed markerless
494 deletions of these three genes (Figure 6d).




495

496 **Figure 6: Genome analysis of the Δ*cps*, Δ*ply*, Δ*lytA* triple mutant generated using**
497 **CRISPR-Cas9 editing.** a. Colony PCR analysis of expected sizes for deletion of three
498 virulence genes. WT vs VL3665. b. Colony PCR analysis of three virulence gene deletions in
499 the final strain VL3665. c. Whole genome marker frequency analysis of strain VL3665. d.
500 Schematic representation of strain VL3665 with three virulence gene deletions and zoom in
501 10kb upstream and downstream of the regions we deleted. The number of mapped reads
502 (gene dosage) is plotted as a function of the position on the circular chromosome.

503 # Discussion

504 Genetic manipulation of microorganisms has been pivotal for the development of
505 biotechnological tools and the study of microorganisms themselves. In this study, we have
506 developed a novel, replicative plasmid with a temperature-sensitive origin of replication
507 carrying a CRISPR-Cas9 based system for advanced and markerless genome engineering in
508 the bacterium *S. pneumoniae.* In particular, we demonstrate that we have successfully deleted
509 genes and large chromosomal regions in a precise and sequential way.

510 The here designed plasmid has the temperature sensitive origin of replication *pG*$^{+}$*host*,
511 which is a derivative of pWV01 of *L. lactis* and can be successfully propagated in
512 pneumococcus at 30°C, while it is not stable at 40°C. Indeed, we show that our *pG*$^{+}$*host*
513 derivative, pDS05, is rapidly lost at 40°C (Fig. 2). We used this feature to eliminate the plasmid
514 from the strains, upon the desired deletion. The fact that the copies of the plasmid vary per
515 cell does not affect our system, since even one copy of *cas9* seems to be sufficient to perform
516 the DSB (van Raaphorst et al. 2017).

517 Specifically, our approach is to harness this CRISPR and the homologous
518 recombination system, to perform CRISPR/Cas9-Mediated Counterselection. Following the same
519 principle of transformation with antibiotic selection, successful transformants survive the
520 CRISPR/Cas9 induced DSB, like they survive growth in antibiotics, if they uptake the rescue HR
521 template. Applying this, our CRISPR system manages to select for transformants in which
522 single genes or even large chromosomal regions were deleted with very high efficiency.
523 Comparing this to just performing natural transformation without any counterselection, which
524 would be an alternative for clean deletions, we show the advantages of our system (Fig. 4f).
525 Without it, we would need to screen many colonies to find correct transformants, depending
526 on the target. This will have to be done by colony PCR, since in most of the cases, the desired
527 deletion will not give any phenotypic difference in the colonies of the successful transformant,
528 which is a costly and time demanding process. On the other hand, with the CRISPR/Cas9-
529 Mediated counterselection, nearly all the colonies that we obtained were the desired
530 transformant, since very few false positives have been observed.

531 Since we are ultimately interested to remove multiple genes and chromosomal regions
532 from the genome, we also needed to demonstrate that our system is capable of consecutive
533 deletions. The key for this was to easily eliminate the plasmid from the newly constructed
534 strain. By growing the strain still carrying the plasmid at the high, non-permissive temperature,
535 we manage to easily cure it. Next, we can transform a new plasmid and proceed further with
536 our deletions. Specifically, after we deleted the capsule, we next deleted virulence factors *ply*
537 and *lytA,* proving that our CRISPR-Cas9 system has flexibility in genetic manipulation of the
538 bacterial genome. Together, the here described plasmid and approach will be useful for the

539 pneumococcal research community and may be applicable to other Gram-positive bacteria as
540 well. Plasmid pDS05 is available from Addgene (accession number pending).
541

## 542 Acknowledgements

543 We are grateful to Paddy Gibson for help with WGS, Renske van Raaphorst for help with
544 image analysis and all the members of the Veening laboratory for stimulating discussions.
545 Work in the Veening lab is supported by the Swiss National Science Foundation (SNSF)
546 (project grant 31003A_172861), a JPIAMR grant (40AR40_185533) from SNSF and ERC
547 consolidator grant 771534-PneumoCaTChER.

548 # References


549 Adli, Mazhar. 2018. “The CRISPR Tool Kit for Genome Editing and Beyond.” *Nature*
550 *Communications* 9(1): 1–13.
551 Avery, O T, C M Macleod, and Maclyn Mccarty. 1944. “Studies on the chemical
552 nature of the substance inducing transformation of pneumococcal types :
553 induction of transformation by a desoxyribonucleic acid fraction isolated from
554 pneumococcus type III.” *The Journal of experimental medicine* 79(2): 137–58.
555 http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2135445&tool=pmcen
556 trez&rendertype=abstract (October 16, 2014).
557 Bijlsma, Jetta J E et al. 2007. “Development of Genomic Array Footprinting for
558 Identification of Conditionally Essential Genes in Streptococcus Pneumoniae.”
559 *Applied and environmental microbiology* 73(5): 1514–24.
560 http://www.ncbi.nlm.nih.gov/pubmed/17261526 (April 1, 2018).
561 Bolger, Anthony M, Marc Lohse, and Bjoern Usadel. 2014. “Trimmomatic: A Flexible
562 Trimmer for Illumina Sequence Data.” *Bioinformatics (Oxford, England)* 30(15):
563 2114–20. https://pubmed.ncbi.nlm.nih.gov/24695404.
564 Darling, Aaron C E, Bob Mau, Frederick R Blattner, and Nicole T Perna. 2004.
565 “Mauve: Multiple Alignment of Conserved Genomic Sequence with
566 Rearrangements.” *Genome research* 14(7): 1394–1403.
567 de Maat, Vincent et al. 2019. “CRISPR-Cas9-Mediated Genome Editing in
568 Vancomycin-Resistant Enterococcus Faecium.” *FEMS Microbiology Letters*
569 366(22).
570 https://academic.oup.com/femsle/article/doi/10.1093/femsle/fnz256/5697197
571 (May 28, 2020).
572 Ducret, Adrien, Ellen M. Quardokus, and Yves V. Brun. 2016. “MicrobeJ, a Tool for
573 High Throughput Bacterial Cell Detection and Quantitative Analysis.” *Nature*
574 *microbiology* 1(7): 16077. http://www.ncbi.nlm.nih.gov/pubmed/27572972 (April
575 17, 2020).
576 Eberhardt, Alice et al. 2009. “Cellular Localization of Choline-Utilization Proteins in
577 *Streptococcus Pneumoniae* Using Novel Fluorescent Reporter Systems.”
578 *Molecular Microbiology* 74(2): 395–408. http://doi.wiley.com/10.1111/j.1365-
579 2958.2009.06872.x (April 27, 2020).
580 Garrison, Erik. “GitHub - Vcflib/Vcflib: C++ Library and Cmdline Tools for Parsing

581 and Manipulating VCF Files.”
582 Garrison, Erik, and Gabor Marth. 2012. “Haplotype-Based Variant Detection from
583 Short-Read Sequencing.”
584 Griffith, F. 1928. “The Significance of Pneumococcal Types.” *The Journal of hygiene*
585 27(2): 333–39. http://www.ncbi.nlm.nih.gov/pubmed/20474956 (February 7,
586 2019).
587 Guo, Tingting et al. 2019. “A Rapid and Versatile Tool for Genomic Engineering in
588 Lactococcus Lactis.” *Microbial Cell Factories* 18(1): 22.
589 https://microbialcellfactories.biomedcentral.com/articles/10.1186/s12934-019-
590 1075-3 (May 28, 2020).
591 Halfmann, Alexander, Regine Hakenbeck, and Reinhold Brückner. 2007. “A New
592 Integrative Reporter Plasmid for Streptococcus Pneumoniae.” *FEMS*
593 *Microbiology Letters* 268(2): 217–24. https://academic.oup.com/femsle/article-
594 lookup/doi/10.1111/j.1574-6968.2006.00584.x (April 27, 2020).
595 Hsu, Patrick D et al. 2013. “DNA Targeting Specificity of RNA-Guided Cas9
596 Nucleases.” *Nature Biotechnology* 31(9): 827–32.
597 http://www.nature.com/articles/nbt.2647 (November 30, 2018).
598 IC, Cañadas et al. 2019. “RiboCas: A Universal CRISPR-Based Editing Tool for
599 Clostridium.” *ACS synthetic biology* 8(6).
600 https://pubmed.ncbi.nlm.nih.gov/31181894/ (May 28, 2020).
601 Jansen, Ruud, Jan D A van Embden, Wim Gaastra, and Leo M Schouls. 2002.
602 “Identification of Genes That Are Associated with DNA Repeats in Prokaryotes.”
603 *Molecular microbiology* 43(6): 1565–75.
604 http://www.ncbi.nlm.nih.gov/pubmed/11952905 (May 13, 2018).
605 Jiang, Wenyang et al. 2013. "RNA-guided editing of bacterial genomes using
606 CRISPR-Cas systems." *Nature Biotechnology*. 31: 233–239.
607 https://doi.org/10.1038/nbt.2508
608 Jinek, Martin et al. 2012. “A Programmable Dual-RNA-Guided DNA Endonuclease in
609 Adaptive Bacterial Immunity.” *Science (New York, N.Y.)* 337(6096): 816–21.
610 http://www.ncbi.nlm.nih.gov/pubmed/22745249 (July 9, 2014).
611 Johnston, Calum et al. 2014. “Bacterial Transformation: Distribution, Shared
612 Mechanisms and Divergent Control.” *Nature Reviews Microbiology* 12(3): 181–
613 96. http://www.nature.com/doifinder/10.1038/nrmicro3199.
614 Kadioglu, Aras, Jeffrey N Weiser, James C Paton, and Peter W Andrew. 2008. “The

615 Role of Streptococcus Pneumoniae Virulence Factors in Host Respiratory
616 Colonization and Disease." *Nature reviews. Microbiology* 6(4): 288–301.
617 http://www.ncbi.nlm.nih.gov/pubmed/18340341 (February 19, 2014).
618 Keller, Lance E., Anne Stéphanie Rueff, Jun Kurushima, and Jan Willem Veening.
619 2019. "Three New Integration Vectors and Fluorescent Proteins for Use in the
620 Opportunistic Human Pathogen Streptococcus Pneumoniae." *Genes* 10(5).
621 Li, H. et al. 2009. "The Sequence Alignment/Map Format and SAMtools."
622 *Bioinformatics* 25(16): 2078–79.
623 Li, H., and R. Durbin. 2009. "Fast and Accurate Short Read Alignment with Burrows-
624 Wheeler Transform." *Bioinformatics* 25(14): 1754–60.
625 Li, Yuan, Claudette M. Thompson, and Marc Lipsitch. 2014. "A Modified Janus
626 Cassette (Sweet Janus) to Improve Allelic Replacement Efficiency by High-
627 Stringency Negative Selection in Streptococcus Pneumoniae." *PLoS ONE* 9(6):
628 2–7.
629 Liu, Xue et al. 2017. " High-throughput CRISPRi Phenotyping Identifies New
630 Essential Genes in Streptococcus Pneumoniae ." *Molecular Systems Biology*
631 13(5): 931.
632 ———. 2020. "Exploration of Bacterial Bottlenecks and Streptococcus Pneumoniae
633 Pathogenesis by CRISPRi-Seq." *bioRxiv*: 2020.04.22.055319.
634 Maguin, E et al. 1992. "New Thermosensitive Plasmid for Gram-Positive Bacteria."
635 *Journal of bacteriology* 174(17): 5633–38.
636 http://www.ncbi.nlm.nih.gov/pubmed/1324906 (October 21, 2019).
637 Maguin, E, H Prévost, S D Ehrlich, and A Gruss. 1996. "Efficient Insertional
638 Mutagenesis in Lactococci and Other Gram-Positive Bacteria." *Journal of*
639 *Bacteriology* 178(3): 931 LP – 935. http://jb.asm.org/content/178/3/931.abstract.
640 Martin, B, P García, M P Castanié, and J P Claverys. 1995. "The RecA Gene of
641 Streptococcus Pneumoniae Is Part of a Competence-Induced Operon and
642 Controls Lysogenic Induction." *Molecular microbiology* 15(2): 367–79.
643 http://www.ncbi.nlm.nih.gov/pubmed/7538190 (October 27, 2014).
644 Massidda, Orietta, Linda Nováková, and Waldemar Vollmer. 2013. "From Models to
645 Pathogens: How Much Have We Learned about *Streptococcus Pneumoniae*
646 Cell Division?" *Environmental Microbiology* 15(12): 3133–57.
647 http://doi.wiley.com/10.1111/1462-2920.12189 (April 27, 2020).
648 Nijkamp, H. J.J. et al. 1986. "The Complete Nucleotide Sequence of the

649 Bacteriocinogenic Plasmid CloDF13." *Plasmid* 16(2): 135–60.
650 Nurk, Sergey et al. 2013. "Assembling Genomes and Mini-Metagenomes from Highly
651 Chimeric Reads." In Springer, Berlin, Heidelberg, 158–70.
652 O'Brien, Katherine L et al. 2009. "Burden of Disease Caused by Streptococcus
653 Pneumoniae in Children Younger than 5 Years: Global Estimates." *Lancet*
654 374(9693): 893–902. http://www.ncbi.nlm.nih.gov/pubmed/19748398 (October
655 22, 2014).
656 Otto, R, W M de Vos, and J Gavrieli. 1982. "Plasmid DNA in Streptococcus Cremoris
657 Wg2: Influence of PH on Selection in Chemostats of a Variant Lacking a
658 Protease Plasmid." *Applied and environmental microbiology* 43(6): 1272–77.
659 http://www.ncbi.nlm.nih.gov/pubmed/16346027 (October 21, 2019).
660 Prudhomme, Marc, Virginie Libante, and Jean Pierre Claverys. 2002. "Homologous
661 Recombination at the Border: Insertion-Deletions and the Trapping of Foreign
662 DNA in Streptococcus Pneumoniae." *Proceedings of the National Academy of*
663 *Sciences of the United States of America* 99(4): 2100–2105.
664 Quinlan, Aaron R., and Ira M. Hall. 2010. "BEDTools: A Flexible Suite of Utilities for
665 Comparing Genomic Features." *Bioinformatics* 26(6): 841–42.
666 Raaphorst, Renske, Morten Kjos, and Jan-Willem Veening. 2020. "BactMAP: An R
667 Package for Integrating, Analyzing and Visualizing Bacterial Microscopy Data."
668 *Molecular Microbiology* 113(1): 297–308.
669 https://onlinelibrary.wiley.com/doi/abs/10.1111/mmi.14417 (April 17, 2020).
670 van Raaphorst, Renske, Morten Kjos, and Jan-Willem Veening. 2017. "Chromosome
671 Segregation Drives Division Site Selection in *Streptococcus Pneumoniae*."
672 *Proceedings of the National Academy of Sciences* 114(29): E5959–68.
673 http://www.pnas.org/lookup/doi/10.1073/pnas.1620608114.
674 Schneider, Caroline A, Wayne S Rasband, and Kevin W Eliceiri. 2012. "NIH Image
675 to ImageJ: 25 Years of Image Analysis." *Nature methods* 9(7): 671–75.
676 http://www.ncbi.nlm.nih.gov/pubmed/22930834 (July 17, 2014).
677 Shuman, Stewart, and Michael S. Glickman. 2007. "Bacterial DNA Repair by Non-
678 Homologous End Joining." *Nature Reviews Microbiology* 5(11): 852–61.
679 http://www.nature.com/doifinder/10.1038/nrmicro1768.
680 Simell, Birgit et al. 2012. "The Fundamental Link between Pneumococcal Carriage
681 and Disease." *Expert review of vaccines* 11(7): 841–55.
682 http://www.ncbi.nlm.nih.gov/pubmed/22913260.

683 Slager, Jelle, Rieza Aprianto, and Jan Willem Veening. 2018. “Deep Genome
684 Annotation of the Opportunistic Human Pathogen Streptococcus Pneumoniae
685 D39.” *Nucleic Acids Research* 46(19): 9971–89.
686 Sorg, Robin A., Clement Gallay, and Jan-Willem Veening. 2019. “Synthetic Gene
687 Regulatory Networks in the Opportunistic Human Pathogen Streptococcus
688 Pneumoniae.” *bioRxiv*: 834689.
689 https://www.biorxiv.org/content/10.1101/834689v1 (April 27, 2020).
690 Sorg, Robin A., Oscar P. Kuipers, and Jan-Willem Veening. 2015. “Gene Expression
691 Platform for Synthetic Biology in the Human Pathogen *Streptococcus*
692 *Pneumoniae*.” *ACS Synthetic Biology* 4(3): 228–39.
693 http://pubs.acs.org/doi/abs/10.1021/sb500229s.
694 Sung, C K, H Li, J P Claverys, and D A Morrison. 2001. “An RpsL Cassette , Janus ,
695 for Gene Replacement through Negative Selection in Streptococcus
696 Pneumoniae An RpsL Cassette , Janus , for Gene Replacement through
697 Negative Selection in Streptococcus Pneumoniae.” *Applied and environmental*
698 *microbiology* 67(11): 5190–96.
699 Team, R core. 2014. “R: A Language and Environment for Statistical ComputingNo
700 Title.” *R Foundation for Statistical Computing, Vienna, Austria*.
701