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.

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) to introduce a double-stranded break

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

98 # Materials and Methods

99 ## Bacterial strains, transformations and growth conditions

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

104 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 +pDS05 [*pG*$^+$*host ori*(Ts)-*ermR-cloDF13ori-specR-*P_{Zn}*_wtcas9-*P_3*_gfp-sgRNA*]	This study
VL3656	*SPV_2146-lacZ-chl-aliA* +pDS07 [*pG*$^+$*host ori*(Ts)-*ermR-cloDF13ori-specR-pZn_wtcas9-*P_3*_sgRNA lacZ*]	This study
VL3657	*ΔlacZ* +pDS07 [*pG*$^+$*host ori*(Ts)-*ermR-cloDF13ori-specR-pZn_wtcas9-*P_3*_sgRNA lacZ*]	This study
VL3658	*ΔlacZ*	This study
VL3659	*Δcps* + pDS07 [*pG*$^+$*host ori*(Ts)-*ermR-cloDF13ori-specR-*P_{Zn}*_wtcas9-*P_3*_sgRNA lacZ*]	This study
VL3660	*Δcps*	This study
VL3661	*Δcps* + 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* + 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

128 synthetic CSP^{-1} (100 ng mL^{-1}) and incubated for 20 min at 30°C with the HR template. After
129 incubation, cells were grown in C+Y medium at 30°C for 20 min. Transformants were selected
130 by plating inside Columbia agar supplemented with 3% of defibrinated sheep blood (Thermo
131 Scientific), followed by CRISPR-mediated counter selection, using 1 mM $ZnCl_2/MnSO_4$. Plates
132 were incubated at 30°C. Correct transformation was verified by PCR and sequencing. Working
133 stocks of cells were prepared by growing cells in C+Y (pH 6.8), until an OD_{595} of 0.4. Cells
134 were collected by centrifugation (1595 × *g* for 10 min) and resuspended in fresh C+Y medium
135 with 15% glycerol and stored at −80°C.
136

137 ## Plasmid curing

138 After the plasmid has been transformed into pneumococcus and successful deletion has been
139 performed with the HR template and CRISPR-mediated counter selection, the plasmid can be
140 eliminated from the pneumococcal cells. To achieve that, we first grow the strain with the
141 plasmid at the non-permissive temperature, 40°C in C+Y, with a starting inoculum 1/10.000.
142 Next, we plate the liquid culture in Columbia blood agar in several dilutions, to obtain single
143 colonies after overnight incubation at 40°C. Single colonies were screened and 99% of them
144 had successfully cured the plasmid from the strain.
145

146 ## Recombinant DNA techniques

147 Oligonucleotides were ordered from Sigma and are listed in Table 2. Phanta Max Super-
148 Fidelity DNA Polymerase (Vazyme) was used in PCR amplifications, restriction enzymes
149 (ThermoFisher Scientific) were used for digestions and T4 DNA Ligase (Vazyme) was used
150 for ligations.
151

152 *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

153

154 ## Strain construction

155 pDS05 (*pG$^+$host ori(Ts)-ermR-cloDF13ori-specR-P$_{Zn}$_wtcas9-P$_3$_gfp-sgRNA*). Gram-positive,
156 temperature sensitive origin of replication *pG$^+$host* (Maguin et al. 1992) and gene *ermR*, which
157 confers resistance to erythromycin, were amplified from plasmid pGh9::ISS1 (Maguin et al.
158 1996) using the primers OVL4739_pGh F and OVL4740_pGh R (fragment 1). Gram-negative
159 origin of replication *CloDF13* (CDF) and gene *specR*, which confers resistance to
160 spectinomycin, were amplified from plasmid pCDF-1b (Nijkamp et al. 1986) with primers
161 OVL4741_pCDF-1b F and OVL4742_pCDF-1b R (fragment 2). The gene which encodes
162 wtCas9 under the control of the Zinc-inducible promoter was amplified from plasmid
163 pJWV102-spCas9wt(van Raaphorst, Kjos, and Veening 2017), using the primers
164 OVL4743_wtcas9 F and OVL4744_wtcas9 R (fragment 3). The sgRNA sequence in which the

165 20 base-pairing region of the sgRNA is replaced by the *gfp* gene was amplified from strain
166 VL2172 with primers OVL4745_gRNA F and OVL4746_gRNA R (fragment 4). The four
167 fragments were digested all together with restriction enzyme *AarI* and ligated. The ligation
168 product was transformed into *E. coli* NEB Turbo and transformants were selected on LB agar
169 with spectinomycin. Correct assembly was confirmed by PCR and sequencing.
170
171 pDS07 (*pG$^+$host ori(Ts)-ermR-cloDF13ori-specR-P_{Zn}_wtcas9-P_3_sgRNA lacZ*). pDS05 was
172 amplified with primers OVL2143_lin pGh F and OVL2142_lin pGh R (fragment 1). Spacer
173 sequence of *sgRNA lacZ* was constructed by annealing primers OVL2132_GG-lacZ-F and
174 OVL2133_GG-lacZ-R. Amplified pDS05 was digested with restriction enzyme *BsaI* and ligated
175 with the annealed oligos. The ligation product was transformed into *E. coli* NEB Turbo and
176 transformants were selected on LB agar with spectinomycin. Correct assembly was confirmed
177 by PCR and sequencing.
178
179 pDS12 (*pG$^+$host ori(Ts)-ermR-cloDF13ori-specR- P_{Zn} _wtcas9-P_3_sgRNA ply*). pDS05 was
180 amplified with primers OVL2143_lin pGh F and OVL2142_lin pGh R. Spacer sequence of
181 *sgRNA ply* was constructed by annealing primers OVL2250_GG-ply-F and OVL2251_GG-ply-
182 R. Amplified pDS05 was digested with restriction enzyme *BsaI* and ligated with the annealed
183 oligos. The ligation product was transformed into *E. coli* NEB Turbo and transformants were
184 selected on LB agar with spectinomycin. Correct assembly was confirmed by PCR and
185 sequencing.
186
187 pDS13 (*pG$^+$host ori(Ts)-ermR-cloDF13ori-specR- P_{Zn} _wtcas9-P_3_sgRNA lytA*). pDS05 was
188 amplified with primers OVL2143_lin pGh F and OVL2142_lin pGh R. Spacer sequence of
189 *sgRNA lytA* was constructed by annealing primers OVL2813_GG-sgRNAlytA-F and
190 OVL2814_GG-sgRNAlytA-R. Amplified pDS05 was digested with restriction enzyme *BsaI* and
191 ligated with the annealed oligos. The ligation product was transformed into *E. coli* NEB Turbo
192 and transformants were selected on LB agar with spectinomycin. Correct assembly was
193 confirmed by PCR and sequencing.
194
195 VL321 (*SPV_2146*-P_{32}-*lacZ-chl-aliA*). *rfbD* and *SPV_2146* were amplified from chromosomal
196 DNA with primers rfbD-F and rfbD-R+BglII (fragment 1). P_{32}-*lacZ* was amplified from pRAS2
197 (lab collection) with primers P32-lacZ-BglII and lacZ_R-AscI (fragment 2). A chloramphenicol
198 resistance marker and *aliA* was amplified from strain VL588 with primers cam-F+AscI and
199 del_CSP.dn-R. Fragment 1 was digested with restriction enzyme *Bgl*II, fragment 2 was
200 digested with restriction enzymes *Bgl*II and *Asc*I and fragment 3 was digested with restriction

201 enzyme *Asc*I. All three fragments were ligated together, and the ligation product was
202 transformed into D39V. Correct assembly was confirmed by PCR and sequencing.
203
204 VL3655 (D39V + pDS05 [*pG*$^{+}$*host ori(Ts)-ermR-cloDF13ori-specR-P*$_{Zn}$*_wtcas9-P*$_{3}$*_gfp-*
205 *sgRNA*]). Plasmid pDS05 was transformed into D39V and transformants were selected on
206 Columbia blood agar with erythromycin to produce the strain VL3655. The presence of the
207 plasmid was confirmed by PCR and plasmid extraction.
208
209 VL3656 (*SPV_2146-lacZ-chl-aliA* + pDS07 [*pG*$^{+}$*host ori(Ts)-ermR-cloDF13ori-specR-*
210 *p*$_{Zn}$*_wtcas9-P*$_{3}$*_sgRNA lacZ*]). Plasmid pDS07, was transformed into VL321(*SPV_2146-lacZ-*
211 *chl-aliA*) and transformants were selected on Columbia blood agar with erythromycin to
212 produce the strain VL3656. The presence of the plasmid was confirmed by PCR and plasmid
213 extraction.
214
215 VL3657 (Δ*lacZ* + pDS07 [*pG*$^{+}$*host ori(Ts)-ermR-cloDF13ori-specR-P*$_{Zn}$*_wtcas9-P*$_{3}$*_sgRNA*
216 *lacZ*]). HR template Δ*lacZ* was transformed into VL3656 and transformants were selected on
217 Columbia blood agar with $ZnCl_2$/$MnSO_4$ to produce the strain VL3657. Correct integration was
218 confirmed by PCR.
219
220 VL3658 (*ΔlacZ*). Strain VL3657 was cured from the plasmid, as described, resulting in strain
221 VL3658.
222
223 VL3659 (Δ*cps* + pDS07 [*pG*$^{+}$*host ori(Ts)-ermR-cloDF13ori-specR-P*$_{Zn}$*_wtcas9-P*$_{3}$*_sgRNA*
224 *lacZ*]). HR template Δ*cps* was transformed into VL3656 and transformants were selected on
225 Columbia blood agar with $ZnCl_2$/$MnSO_4$ to produce the strain VL3659. Correct integration was
226 confirmed by PCR.
227
228 VL3660 (*Δcps*). Strain VL3659 was cured from the plasmid, as described, resulting in strain
229 VL3660.
230
231 VL3661 (*Δcps* + pDS12 [*pG*$^{+}$*host ori(Ts)-ermR-cloDF13ori-specR-P*$_{Zn}$*_wtcas9- P*$_{3}$*-sgRNA*
232 *ply*]). Plasmid pDS12 was transformed into VL3660 and transformants were selected on
233 Columbia blood agar with erythromycin to produce the strain VL3661. The presence of the
234 plasmid was confirmed by PCR and plasmid extraction.
235
236 VL3662 (*Δcps, Δply* + pDS12 [*pG*$^{+}$*host ori(Ts)-ermR-cloDF13ori-specR-P*$_{Zn}$*_wtcas9-*
237 *P*$_{3}$*_sgRNA ply*]). HR template Δ*ply* was transformed into VL3661 and transformants were

238 selected on Columbia blood agar with $ZnCl_2/MnSO_4$ to produce the strain VL3662. Correct
239 integration was confirmed by PCR.
240
241 VL3663 (*Δcps, Δply*). Strain VL3662 was cured from the plasmid, as described, resulting in
242 strain VL3663.
243
244 VL3664 (*Δcps, Δply* + pDS13 [*pG*$^+$*host ori(Ts)-ermR-cloDF13ori-specR-*P_{Zn}*_wtcas9-*
245 P_3*_sgRNA lytA*]). Plasmid pDS13 was transformed into VL3663 and transformants were
246 selected on Columbia blood agar with erythromycin to produce the strain VL3664. The
247 presence of the plasmid was confirmed by PCR and plasmid extraction.
248
249 VL3665 (*Δcps, Δply, ΔlytA* + pDS13 [*pG*$^+$*host ori(Ts)-ermR-cloDF13ori-specR-* P_{Zn}*_wtcas9-*
250 P_3*_sgRNA lytA*]). HR template Δ*lytA* was transformed into VL3664 and transformants were
251 selected on Columbia blood agar with $ZnCl_2/MnSO_4$ to produce the strain VL3665. Correct
252 integration was confirmed by PCR.
253
254 ## Microscopy
255 *S. pneumoniae* cells were stored as exponential phase frozen cultures. Frozen stock was
256 inoculated 1:100 in C+Y medium and pre-grown to OD600 ~ 0.1. Cells were diluted once
257 again 1:100 in fresh C+Y (with antibiotic, if applicable) and grown to exponential phase to
258 achieve balanced growth.
259 Cells were grown as described above to achieve balanced growth and subsequently
260 concentrated and brought onto a multi test slide carrying a thin layer of 1.2% agarose in C+Y.
261 Imaging was performed on Fluorescence microscopy was performed on a Leica DMi8 through
262 a 100x phase contrast objective (NA 1.40) with a SOLA Light Engine (lumencor) light source.
263 Light was filtered through external excitation filters 470/40 nm (Chroma ET470/40x) for
264 visualization of GFP. Light passed through a cube (Leica 11536022) with a GFP/RFP
265 polychroic mirror (498/564 nm). External emission filters used were from Chroma ET520/40m.
266 Images were captured using LasX software (Leica) and exported to ImageJ(Schneider,
267 Rasband, and Eliceiri 2012) for final preparation.
268 Cell outlines were detected using MicrobeJ (Ducret et al. 2016). For all microscopy
269 experiments, random image frames were used for analysis. The cell outline, object detection,
270 and fluorescent intensity data were further processed using the R-package BactMAP
271 (Raaphorst, Kjos, and Veening 2020).
272

273 ## Transformation efficiency assays

274 To calculate the transformation efficiency, 1 µg/ml of PCR product of DNA fragment containing
275 the upstream and the downstream region of the deletion target was added. Serial dilutions
276 were plated either with or without 1 mM $ZnCl_2$/$MnSO_4$, and the transformation efficiency was
277 calculated by dividing the number of transformants by the total viable count. All transformation
278 efficiency values are averages of three biologically independent experiments.
279

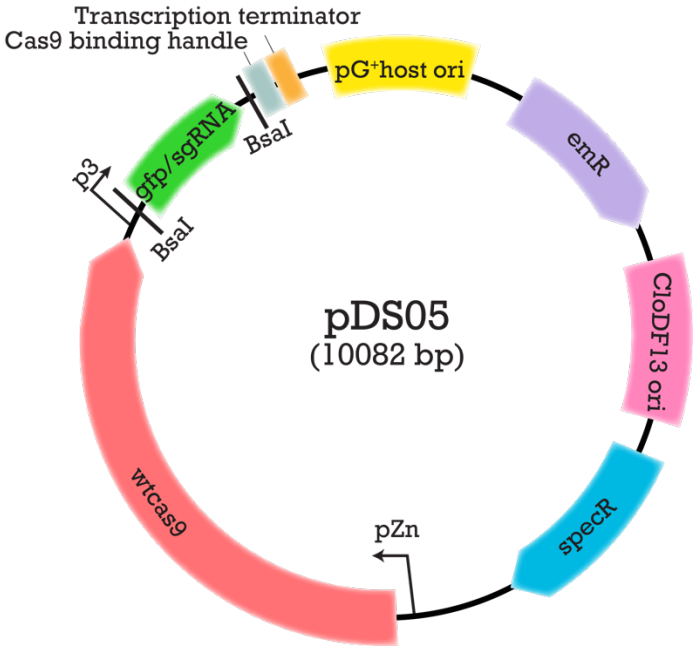
280 ## Whole genome sequencing and variant analysis

281 Genomic DNA was isolated using the FastPure Bacteria DNA Isolation Mini Kit (Vazyme)
282 according to the manufacturers' protocol and sent for whole genome sequencing. Illumina
283 library prep and sequencing were carried out by Novogene (sequencing in PE150 mode).
284 Reads were trimmed using Trimmomatic (Bolger, Lohse, and Usadel 2014), then assembled
285 using SPAdes (Nurk et al. 2013) and contigs were reordered in Mauve (Darling et al. 2004)
286 using D39V as a reference (Slager et al NAR 2018). Reads were mapped onto the scaffold
287 using bwa (H. Li and Durbin 2009), read depth was determined in samtools (H. Li et al. 2009),
288 and plotted in R (Team 2014). In order to detect small variants, raw reads were mapped onto
289 the reference using bwa, and variants were detected using Freebayes (Garrison and Marth
290 2012). Potential variants with PHRED scores greater than 30 were filtered out on DP >5 using
291 vcflib (Garrison n.d.), then intersected with the D39V annotation using Bedtools (Quinlan and
292 Hall 2010)
293

294 # Results

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

296 First, a replicative plasmid was designed and constructed (Figure 1), by combining PCR
297 amplified genomic and plasmid parts. The main idea behind the choice for individual vector
298 components relied in creating a platform with the CRISPR-Cas9 system in *S. pneumoniae*
299 while at the same time allowing for plasmid propagation in both Gram-positive and Gram-
300 negative hosts. The modular vector consists of six individual components. Two origins of
301 replication; the high-copy pG^{+}host replicon, which is a replication thermosensitive derivative
302 of pWV01 (Otto et al. 1982) that in *L. lactis*, (and other Gram-positive bacteria) replicates at
303 28$^{\circ C}$ but is lost above 37$^{\circ C}$, and the low-copy *CloDF13* (CDF) replicon for propagation in *E.*
304 *coli.* By combining these two origins of replication, it ensures low copy numbers at 37$^{\circ C}$ in *E.*
305 *coli* thereby preventing toxicity of the CRISPR-Cas9 system while cloning. Additionally, it has
306 the gene which encodes wtCas9 under the control of the Zinc-inducible promoter Pzn
307 (Eberhardt et al. 2009) (plasmid pDS05) and genes conferring spectinomycin (*E. coli*) and
308 erythromycin (*S. pneumoniae*) resistance. Finally, it has the strong synthetic constitutive P3
309 promoter (Sorg et al. 2015) driving the sgRNA sequence in which the 20 base-pairing region
310 of the sgRNA is replaced by the *gfp* gene flanked by two *Bsa*I restriction sites. This allows for
311 easy replacement of *gfp* by the spacer sequence of sgRNA with golden gate cloning. This
312 way, successful cloning of the sgRNA allows for easy selection by absence of GFP
313 fluorescence, giving us a versatile vector for different sgRNAs (see below).
314



315
316 **Figure 1: Schematic representation of plasmid pDS05**

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

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

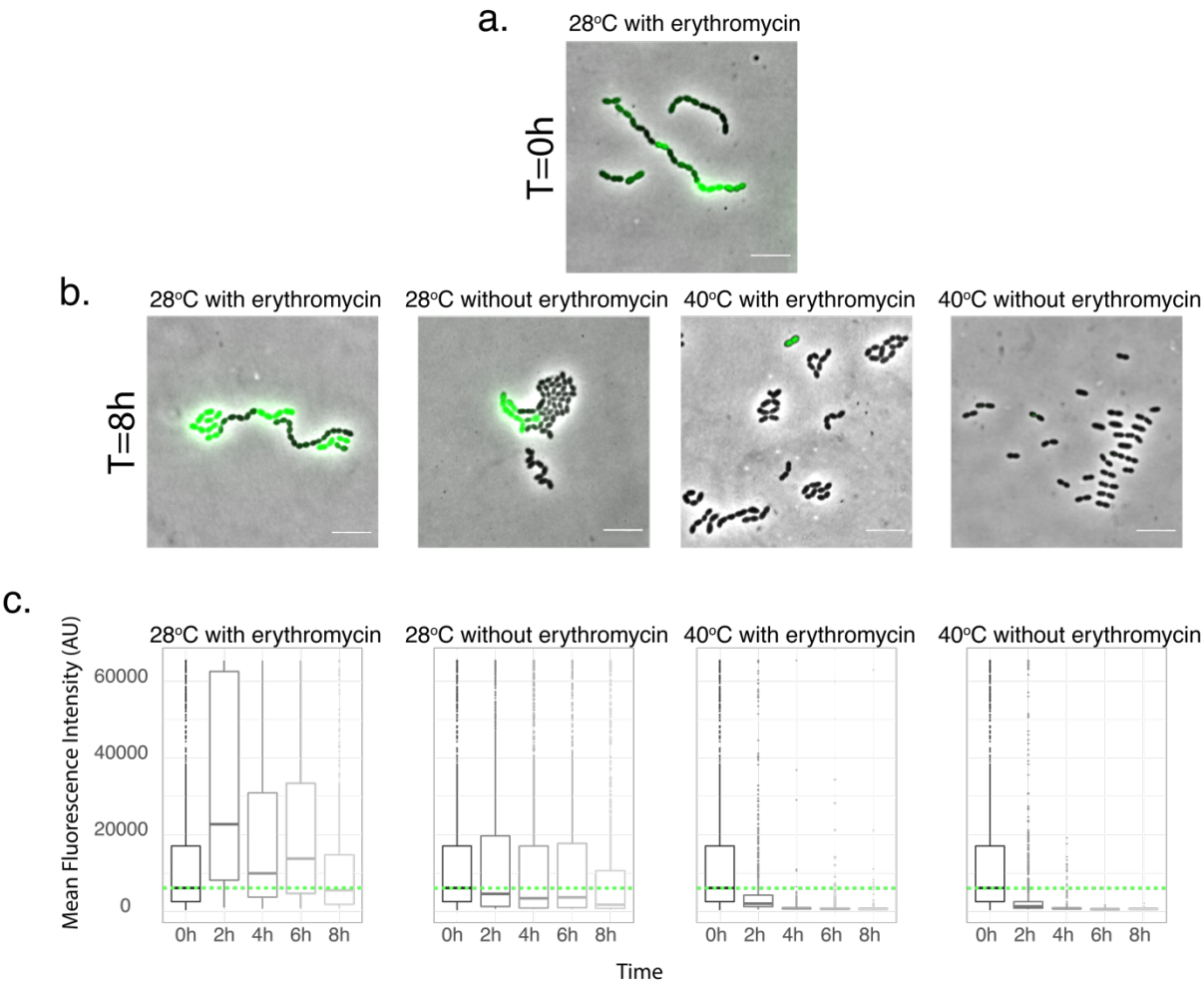


Figure 2: Microscopy analysis of strain VL3655 (D39V, pDS05). Overlay of GFP signals with phase contrast image shows GFP expression. a. Preculture grown at 28°C with erythromycin (T=0h), b. Images are shown of cells grown for 8h as exponentially growing cells (balanced growth) in four different conditions: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

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).

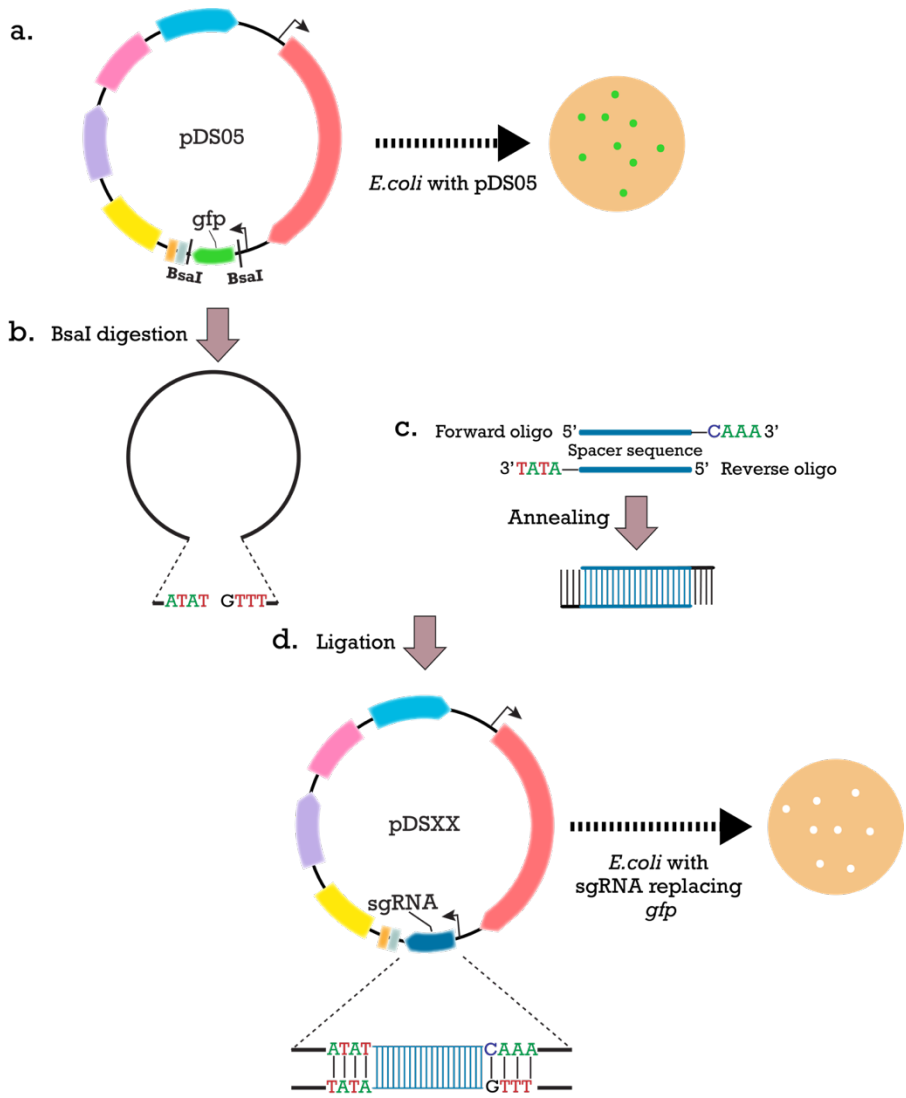


Figure 3: Workflow of sgRNA replacement. a. pDS05 was designed to facilitate easy replacement of *gfp* by the spacer sequence of the desired sgRNA with golden gate cloning, allowing also for detection of false positive transformants. *gfp,* which encodes a green fluorescent protein, is in place of the 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.

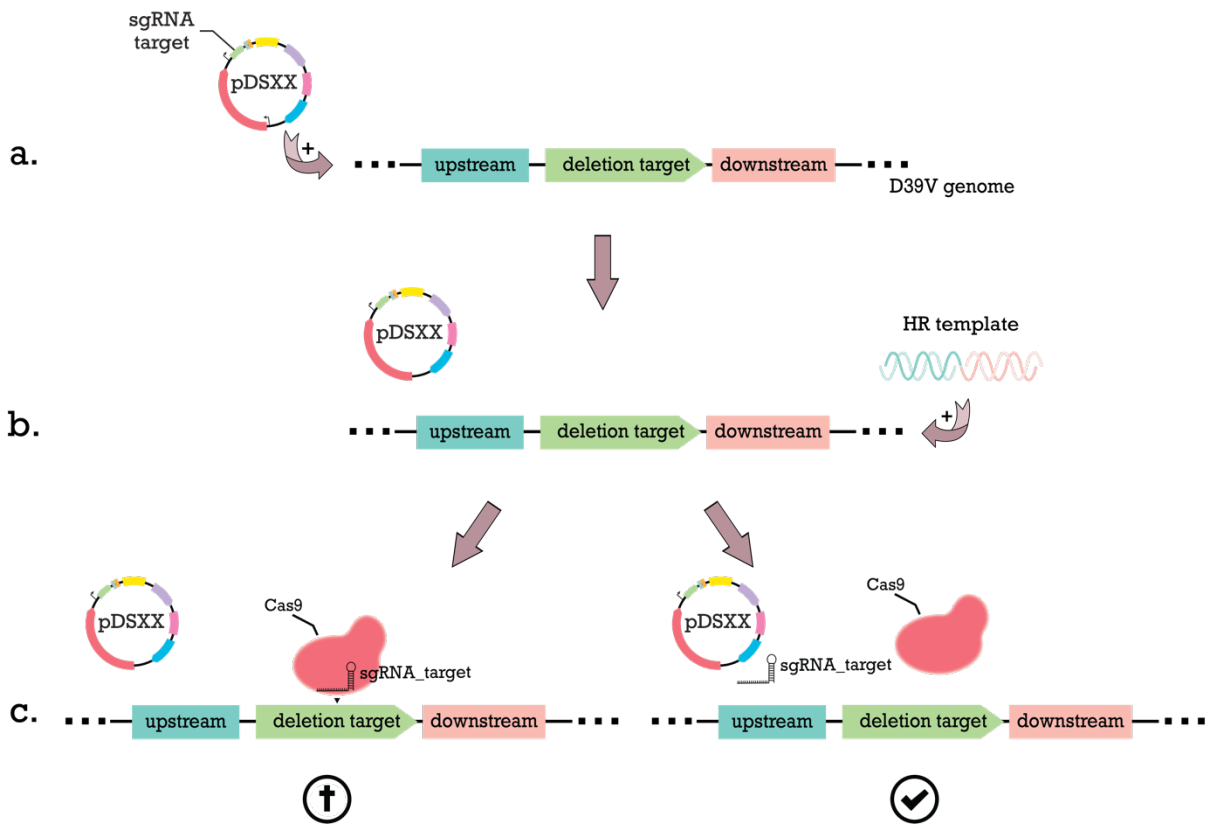


Figure 4: Workflow for markerless deletions. a. Uptake of the plasmid by the strain with the sgRNA sequence for the desired deletion. b. Transformation of a homology recombination (HR) template consisting of the ligation of the upstream and downstream region of the deletion target. c. Plating transformants 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.

395 Deleting genes and large chromosomal regions from the *S. pneumoniae* genome
396 To assess the efficiency of the system, we first constructed a strain (strain VL3656; Figure 5a)
397 in which we placed the *E. coli lacZ* gene under a constitutive promoter behind the *S.*
398 *pneumoniae* D39V *cps* locus (encoding the capsule). *lacZ* encodes for a β-galactosidase that
399 hydrolyzes X-gal to produce a blue product, allowing for blue/white screening on plates.
400 Colonies with blue color would still carry the *lacZ* gene, while colonies with the standard
401 white/green (on blood agar) color would indicate that the gene has been deleted from the
402 chromosome.
403 Strain VL3656 also carries the pDS07 plasmid, which contains a sgRNA targeting *lacZ.*
404 Next, we constructed an HR template that consisted of the 1000 bp upstream and 1000 bp
405 downstream region of *lacZ* (excluding *lacZ*) (**Figure 5**a) and we transformed VL3656 with this
406 template. Transformants were selected by plating on agar containing $ZnCl_2$ to induce
407 expression of Cas9.
408 After transformation of strain VL3656 with the HR template, transformation efficiency
409 was calculated (Figure 5b). The CRISPR-mediated counter-selection, offered by the system,
410 worked successfully. The selection efficiency was high and almost all colonies in the
411 transformation where the HR template was given and Cas9 was induced had their original
412 color, indicating the *lacZ* gene has been successfully deleted.
413 Transformants were tested for correct deletion of *lacZ* by colony PCR. The primers
414 used were binding 1000bp upstream and downstream of *lacZ*, setting the correct PCR product
415 of the successful deletion at the 2000bp (5000bp if *lacZ* was not deleted) (Figure 5c). All the
416 tested colonies had the expected product demonstrating successful deletion of *lacZ*, resulting
417 in strain VL3657 (Figure 5d).
418 Additionally, we also used the system to delete an even larger chromosomal fragment.
419 For this, we targeted the operon that encodes the capsule and the *lacZ* gene that had been
420 inserted downstream of it, which is around 24Kbp long, allowing for blue/white screening.
421 Once again, selection efficiency was very high and almost all colonies had their original color
422 (Figure 5b). Colony PCR verified correct deletion of the *cps-lacZ* chromosomal region (see
423 below).
424 Using the same HR template to delete *lacZ,* we also performed transformation assays
425 without the counterselection offered, by inducing our CRISPR system (Figure 5f). Thousands
426 of colonies needed to be screened to find successful transformants with the original colony
427 color, among the blue colonies. In contrast, by using the system, almost with absolute success
428 rate, all the colonies on our plates are the correct transformants, demonstrating how efficient
429 our system is to easily select edited cells.
430

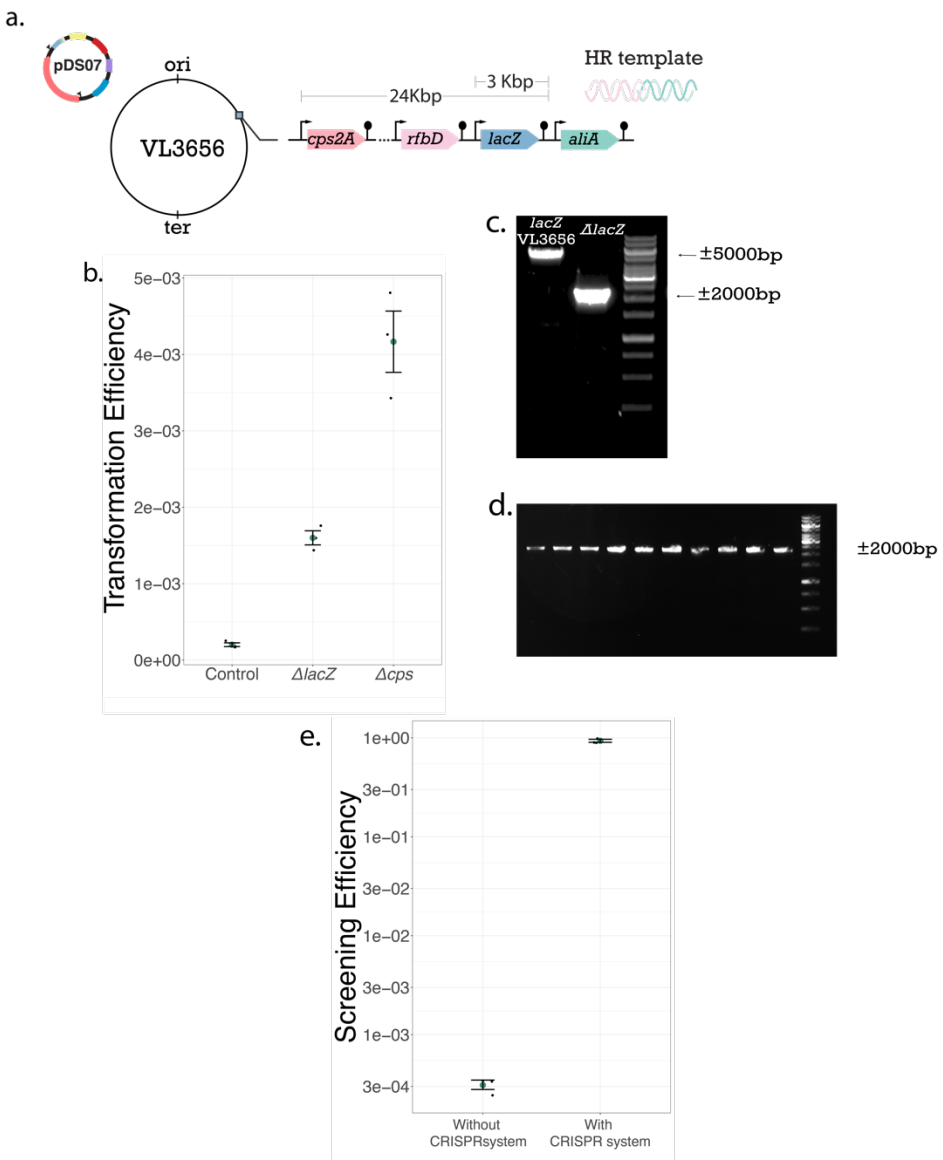


Figure 5: Genome editing in *S. pneumoniae* using CRISPR/Cas9. a. Schematic representation of strain VL3656. The *lacZ* gene has been inserted downstream the capsule operon and a version of the plasmid with a sgRNA targeting *lacZ* has been transformed in the strain. Control is the transformation assay 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).

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

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

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

457

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

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

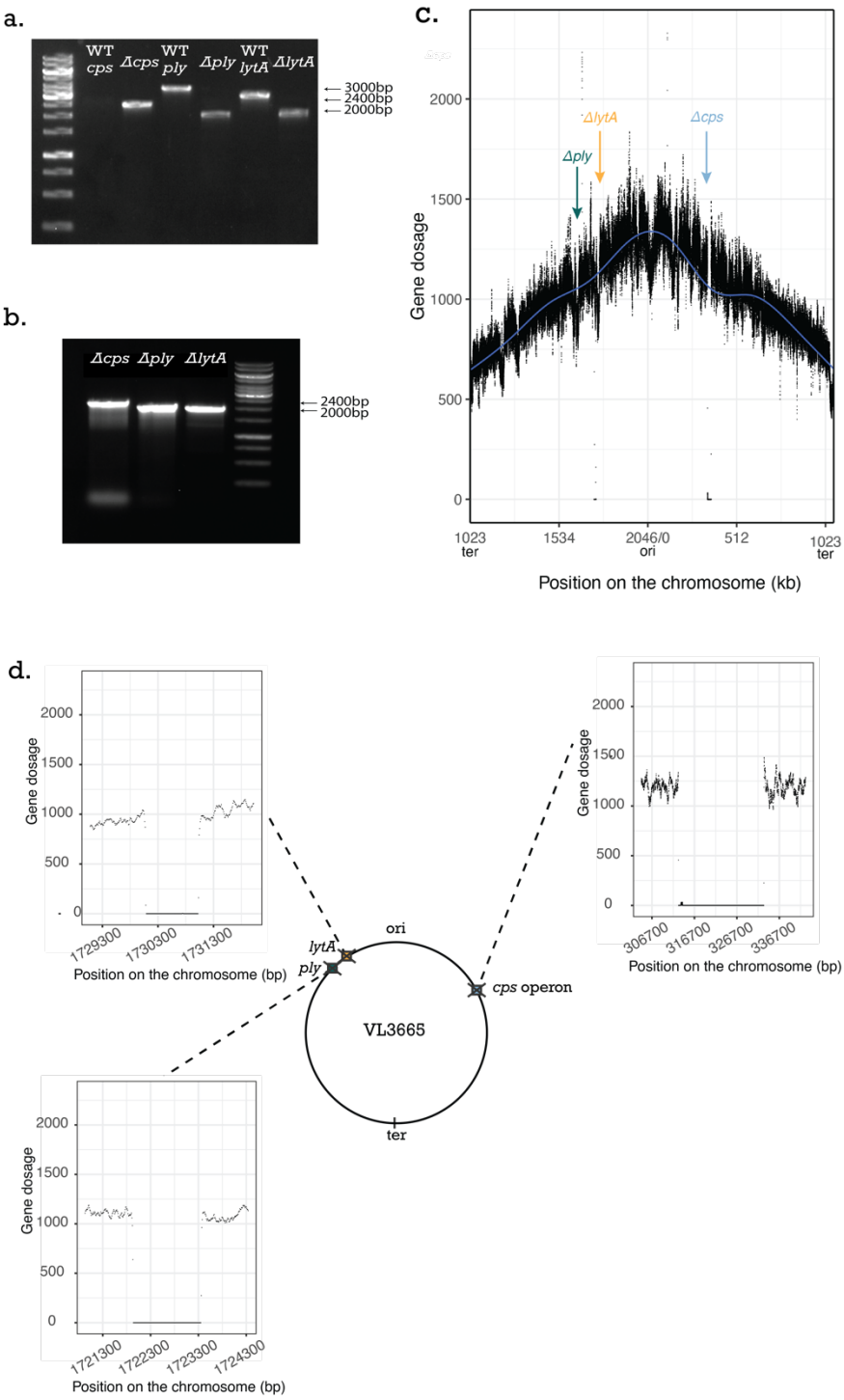
473

474

475

476

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



485

486 **Figure 6: Genome analysis of the Δ*cps*, Δ*ply*, Δ*lytA* triple mutant generated using**
487 **CRISPR-Cas9 editing.** a. Colony PCR analysis of expected sizes for deletion of three
488 virulence genes. WT vs VL3665. b. Colony PCR analysis of three virulence gene deletions in
489 the final strain VL3665. c. Whole genome marker frequency analysis of strain VL3665. d.
490 Schematic representation of strain VL3665 with three virulence gene deletions and zoom in
491 10kb upstream and downstream of the regions we deleted. The number of mapped reads
492 (gene dosage) is plotted as a function of the position on the circular chromosome.

493 # Discussion

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

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

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

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

529 pneumococcal research community and may be applicable to other Gram-positive bacteria as
530 well. Plasmid pDS05 is available from Addgene (accession number pending).
531

532 Acknowledgements

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

538 # References

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

571 and Manipulating VCF Files."
572 Garrison, Erik, and Gabor Marth. 2012. "Haplotype-Based Variant Detection from
573 Short-Read Sequencing."
574 Griffith, F. 1928. "The Significance of Pneumococcal Types." *The Journal of hygiene*
575 27(2): 333–39. http://www.ncbi.nlm.nih.gov/pubmed/20474956 (February 7,
576 2019).
577 Guo, Tingting et al. 2019. "A Rapid and Versatile Tool for Genomic Engineering in
578 Lactococcus Lactis." *Microbial Cell Factories* 18(1): 22.
579 https://microbialcellfactories.biomedcentral.com/articles/10.1186/s12934-019-
580 1075-3 (May 28, 2020).
581 Halfmann, Alexander, Regine Hakenbeck, and Reinhold Brückner. 2007. "A New
582 Integrative Reporter Plasmid for Streptococcus Pneumoniae." *FEMS*
583 *Microbiology Letters* 268(2): 217–24. https://academic.oup.com/femsle/article-
584 lookup/doi/10.1111/j.1574-6968.2006.00584.x (April 27, 2020).
585 Hsu, Patrick D et al. 2013. "DNA Targeting Specificity of RNA-Guided Cas9
586 Nucleases." *Nature Biotechnology* 31(9): 827–32.
587 http://www.nature.com/articles/nbt.2647 (November 30, 2018).
588 IC, Cañadas et al. 2019. "RiboCas: A Universal CRISPR-Based Editing Tool for
589 Clostridium." *ACS synthetic biology* 8(6).
590 https://pubmed.ncbi.nlm.nih.gov/31181894/ (May 28, 2020).
591 Jansen, Ruud, Jan D A van Embden, Wim Gaastra, and Leo M Schouls. 2002.
592 "Identification of Genes That Are Associated with DNA Repeats in Prokaryotes."
593 *Molecular microbiology* 43(6): 1565–75.
594 http://www.ncbi.nlm.nih.gov/pubmed/11952905 (May 13, 2018).
595 Jinek, Martin et al. 2012. "A Programmable Dual-RNA-Guided DNA Endonuclease in
596 Adaptive Bacterial Immunity." *Science (New York, N.Y.)* 337(6096): 816–21.
597 http://www.ncbi.nlm.nih.gov/pubmed/22745249 (July 9, 2014).
598 Johnston, Calum et al. 2014. "Bacterial Transformation: Distribution, Shared
599 Mechanisms and Divergent Control." *Nature Reviews Microbiology* 12(3): 181–
600 96. http://www.nature.com/doifinder/10.1038/nrmicro3199.
601 Kadioglu, Aras, Jeffrey N Weiser, James C Paton, and Peter W Andrew. 2008. "The
602 Role of Streptococcus Pneumoniae Virulence Factors in Host Respiratory
603 Colonization and Disease." *Nature reviews. Microbiology* 6(4): 288–301.
604 http://www.ncbi.nlm.nih.gov/pubmed/18340341 (February 19, 2014).

605 Keller, Lance E., Anne Stéphanie Rueff, Jun Kurushima, and Jan Willem Veening.
606 2019. “Three New Integration Vectors and Fluorescent Proteins for Use in the
607 Opportunistic Human Pathogen Streptococcus Pneumoniae.” *Genes* 10(5).
608 Li, H. et al. 2009. “The Sequence Alignment/Map Format and SAMtools.”
609 *Bioinformatics* 25(16): 2078–79.
610 Li, H., and R. Durbin. 2009. “Fast and Accurate Short Read Alignment with Burrows-
611 Wheeler Transform.” *Bioinformatics* 25(14): 1754–60.
612 Li, Yuan, Claudette M. Thompson, and Marc Lipsitch. 2014. “A Modified Janus
613 Cassette (Sweet Janus) to Improve Allelic Replacement Efficiency by High-
614 Stringency Negative Selection in Streptococcus Pneumoniae.” *PLoS ONE* 9(6):
615 2–7.
616 Liu, Xue et al. 2017. “ High-throughput CRISPRi Phenotyping Identifies New
617 Essential Genes in Streptococcus Pneumoniae .” *Molecular Systems Biology*
618 13(5): 931.
619 ———. 2020. “Exploration of Bacterial Bottlenecks and Streptococcus Pneumoniae
620 Pathogenesis by CRISPRi-Seq.” *bioRxiv*: 2020.04.22.055319.
621 Maguin, E et al. 1992. “New Thermosensitive Plasmid for Gram-Positive Bacteria.”
622 *Journal of bacteriology* 174(17): 5633–38.
623 http://www.ncbi.nlm.nih.gov/pubmed/1324906 (October 21, 2019).
624 Maguin, E, H Prévost, S D Ehrlich, and A Gruss. 1996. “Efficient Insertional
625 Mutagenesis in Lactococci and Other Gram-Positive Bacteria.” *Journal of*
626 *Bacteriology* 178(3): 931 LP – 935. http://jb.asm.org/content/178/3/931.abstract.
627 Martin, B, P García, M P Castanié, and J P Claverys. 1995. “The RecA Gene of
628 Streptococcus Pneumoniae Is Part of a Competence-Induced Operon and
629 Controls Lysogenic Induction.” *Molecular microbiology* 15(2): 367–79.
630 http://www.ncbi.nlm.nih.gov/pubmed/7538190 (October 27, 2014).
631 Massidda, Orietta, Linda Nováková, and Waldemar Vollmer. 2013. “From Models to
632 Pathogens: How Much Have We Learned about *Streptococcus Pneumoniae*
633 Cell Division?” *Environmental Microbiology* 15(12): 3133–57.
634 http://doi.wiley.com/10.1111/1462-2920.12189 (April 27, 2020).
635 Nijkamp, H. J.J. et al. 1986. “The Complete Nucleotide Sequence of the
636 Bacteriocinogenic Plasmid CloDF13.” *Plasmid* 16(2): 135–60.
637 Nurk, Sergey et al. 2013. “Assembling Genomes and Mini-Metagenomes from Highly
638 Chimeric Reads.” In Springer, Berlin, Heidelberg, 158–70.

639 O'Brien, Katherine L et al. 2009. "Burden of Disease Caused by Streptococcus
640 Pneumoniae in Children Younger than 5 Years: Global Estimates." *Lancet*
641 374(9693): 893–902. http://www.ncbi.nlm.nih.gov/pubmed/19748398 (October
642 22, 2014).
643 Otto, R, W M de Vos, and J Gavrieli. 1982. "Plasmid DNA in Streptococcus Cremoris
644 Wg2: Influence of PH on Selection in Chemostats of a Variant Lacking a
645 Protease Plasmid." *Applied and environmental microbiology* 43(6): 1272–77.
646 http://www.ncbi.nlm.nih.gov/pubmed/16346027 (October 21, 2019).
647 Prudhomme, Marc, Virginie Libante, and Jean Pierre Claverys. 2002. "Homologous
648 Recombination at the Border: Insertion-Deletions and the Trapping of Foreign
649 DNA in Streptococcus Pneumoniae." *Proceedings of the National Academy of*
650 *Sciences of the United States of America* 99(4): 2100–2105.
651 Quinlan, Aaron R., and Ira M. Hall. 2010. "BEDTools: A Flexible Suite of Utilities for
652 Comparing Genomic Features." *Bioinformatics* 26(6): 841–42.
653 Raaphorst, Renske, Morten Kjos, and Jan-Willem Veening. 2020. "BactMAP: An R
654 Package for Integrating, Analyzing and Visualizing Bacterial Microscopy Data."
655 *Molecular Microbiology* 113(1): 297–308.
656 https://onlinelibrary.wiley.com/doi/abs/10.1111/mmi.14417 (April 17, 2020).
657 van Raaphorst, Renske, Morten Kjos, and Jan-Willem Veening. 2017. "Chromosome
658 Segregation Drives Division Site Selection in *Streptococcus Pneumoniae*."
659 *Proceedings of the National Academy of Sciences* 114(29): E5959–68.
660 http://www.pnas.org/lookup/doi/10.1073/pnas.1620608114.
661 Schneider, Caroline A, Wayne S Rasband, and Kevin W Eliceiri. 2012. "NIH Image
662 to ImageJ: 25 Years of Image Analysis." *Nature methods* 9(7): 671–75.
663 http://www.ncbi.nlm.nih.gov/pubmed/22930834 (July 17, 2014).
664 Shuman, Stewart, and Michael S. Glickman. 2007. "Bacterial DNA Repair by Non-
665 Homologous End Joining." *Nature Reviews Microbiology* 5(11): 852–61.
666 http://www.nature.com/doifinder/10.1038/nrmicro1768.
667 Simell, Birgit et al. 2012. "The Fundamental Link between Pneumococcal Carriage
668 and Disease." *Expert review of vaccines* 11(7): 841–55.
669 http://www.ncbi.nlm.nih.gov/pubmed/22913260.
670 Slager, Jelle, Rieza Aprianto, and Jan Willem Veening. 2018. "Deep Genome
671 Annotation of the Opportunistic Human Pathogen Streptococcus Pneumoniae
672 D39." *Nucleic Acids Research* 46(19): 9971–89.

673 Sorg, Robin A., Clement Gallay, and Jan-Willem Veening. 2019. “Synthetic Gene
674 Regulatory Networks in the Opportunistic Human Pathogen Streptococcus
675 Pneumoniae.” *bioRxiv*: 834689.
676 https://www.biorxiv.org/content/10.1101/834689v1 (April 27, 2020).
677 Sorg, Robin A., Oscar P. Kuipers, and Jan-Willem Veening. 2015. “Gene Expression
678 Platform for Synthetic Biology in the Human Pathogen *Streptococcus*
679 *Pneumoniae*.” *ACS Synthetic Biology* 4(3): 228–39.
680 http://pubs.acs.org/doi/abs/10.1021/sb500229s.
681 Sung, C K, H Li, J P Claverys, and D A Morrison. 2001. “An RpsL Cassette , Janus ,
682 for Gene Replacement through Negative Selection in Streptococcus
683 Pneumoniae An RpsL Cassette , Janus , for Gene Replacement through
684 Negative Selection in Streptococcus Pneumoniae.” *Applied and environmental*
685 *microbiology* 67(11): 5190–96.
686 Team, R core. 2014. “R: A Language and Environment for Statistical ComputingNo
687 Title.” *R Foundation for Statistical Computing, Vienna, Austria*.
688