Unbiased homeologous recombination during pneumococcal transformation allows for multiple chromosomal integration events

Jun Kurushima[1], Nathalie Campo[2], Renske van Raaphorst[1], Guillaume Cerckel[1], Patrice Polard[2], Jan-Willem Veening[1*]

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

[2]Laboratoire de Microbiologie et Génétique Moléculaires (LMGM), Centre de Biologie Intégrative (CBI), Toulouse, France

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

Abstract

The rapid spread of antimicrobial resistance and vaccine escape in the opportunistic human pathogen *Streptococcus pneumoniae* can be largely attributed to competence-induced transformation. To better understand the dynamics of competence-induced transformation, we studied this process at the single-cell level. We show that within isogenic populations, all cells become naturally competent and bind exogenous DNA. In addition, we find that transformation is highly efficient and that the chromosomal location of the integration site or whether the transformed gene is encoded on the leading or lagging strand has limited influence on recombination efficiency. Indeed, we have observed multiple recombination events in single recipients in real-time. However, because of saturation of the DNA uptake and integration machinery and because a single stranded donor DNA replaces the original allele, we find that transformation efficiency has an upper threshold of approximately 50% of the population. Counterintuitively, in the presence of multiple transforming DNAs, the fraction of untransformed cells increases to more than 50%. The fixed mechanism of transformation results in a fail-safe strategy for the population as half of the population generally keeps an intact copy of the original genome. Together, this work advances our understanding of pneumococcal genome plasticity.

Keywords: *Streptococcus pneumoniae*, natural transformation, competence development, homologous recombination, homeologous recombination, single cell analysis, antibiotic resistance, horizontal gene transfer

Introduction

The opportunistic human pathogen *Streptococcus pneumoniae* (the pneumococcus) kills over a million individuals each year, despite the introduction of several vaccines targeting its capsule (Croucher et al., 2018; O'Brien et al., 2009; Prina et al., 2015). Because of its ability to take up DNA from its environment by competence activation, genes associated with capsule biosynthesis are rapidly transferred from one strain to the other thereby contributing to vaccine escape (Salvadori et al., 2019). In addition, antibiotic resistance remains a cause of concern and competence-dependent recombination plays an important role in the spread of drug resistance (Lo et al., 2019). For example, one of the main genetic sources for penicillin resistance in *S. pneumoniae* is DNA acquired from non-pathogenic *Streptococci* from the viridans group such as *S. mitis* that also lives in the human nasal and oral cavities (Bryskier, 2002; Janoir et al., 1999). Consistently, antibiotic resistant pneumococci and vaccine-escape variants remain an important cause of invasive infections in spite of the introduction of the conjugate vaccines (Fenoll et al., 2018; Levy et al., 2019; Ouldali et al., 2018).

Although pneumococcal competence is one of the best studied bacterial regulatory system (Gómez-Mejia et al., 2018; Johnston et al., 2014; Lin and Kussell, 2017; Salvadori et al., 2019; Shanker and Federle, 2017; Straume et al., 2015; Veening and Blokesch, 2017), and pneumococcal transformation was already discovered in the early twentieth century (Avery et al., 1944; Griffith, 1928), we have a poor understanding on how competence-dependent transformation drives pneumococcal population dynamics, serotype displacement and the spread of antibiotic resistance. Importantly, horizontal gene transfer (HGT) via natural transformation is not only conserved in Streptococci but is present in many human pathogens where it promotes the spread of virulence determinants and antibiotic resistance (Brockhurst et al., 2019; Dubnau and Blokesch, 2019; Johnston et al., 2014). For this reason,

it is crucial to understand what the main bottlenecks are during the take-up and recombination of exogenous DNA that leads to transformed new genotypes.

In contrast to many other competent pathogens such as *Acinetobacter* spp. and *Neisseria meningitidis* in which competence is constitutively expressed, competence development in *S. pneumoniae* is only activated under specific conditions (Blokesch, 2016; Claverys et al., 2006). Pneumococcal competence is under control of a two-component quorum sensing system (**Figure 1**). ComC is cleaved and exported by the peptidase-containing ATP-binding cassette transporter ComAB (Chandler and Morrison, 1988; Håvarstein et al., 1995; Hui et al., 1995). Cleaved ComC autoinducer is commonly referred to as CSP, for Competence Stimulating Peptide (Alloing et al., 1996; Håvarstein et al., 1996, 1995). CSP is recognized by the membrane-bound histidine kinase ComD(Håvarstein et al., 1996). Once a certain threshold level of CSP has been reached, as the culture reaches higher densities, or when other environmental factors increase local CSP concentrations (Domenech et al., 2018; Moreno-Gámez et al., 2017), ComD will autophosphorylate and transfer the phosphoryl group to the response regulator ComE (Martin et al., 2013). Phosphorylated ComE then dimerizes (Boudes et al., 2014; Sanchez et al., 2015) and binds to a specific DNA sequence (Martin et al., 2013; Pestova et al., 1996; Slager et al., 2019; Ween et al., 1999). The *comCDE* and *comAB* operons are under direct control of ComE, setting up a positive feedback loop. The genes under control of ComE are called the early *com* genes (**Figure 1**). Importantly, phosphorylated ComE also activates expression of the gene encoding the alternative sigma factor ComX. ComX activates transcription of the so-called late *com* genes, which includes the genes required for DNA uptake and integration (Campbell et al., 1998; Dagkessamanskaia et al., 2004; Luo et al., 2003; Pestova and Morrison, 1998; Slager et al., 2019) (**Figure 1**). While regulation of competence is highly diverse between naturally

transformable bacteria, the actual DNA uptake and integration machinery is largely conserved (Chen and Dubnau, 2004; Johnston et al., 2014).

During pneumococcal competence, exogenous double stranded DNA (dsDNA) is bound by a type IV-like pilus (Laurenceau et al., 2013) and subsequently sequestered to the DNA uptake machinery (**Figure 1**). Note that in contrast to some other competent bacteria, pneumococcus binds and takes up DNA of any sequence, including non-kin DNA(Mell and Redfield, 2014). Next, the dsDNA is processed into single stranded DNA (ssDNA) by the EndA nuclease and internalized through a membrane pore consisting of ComEC. Once inside, the ssDNA is bound by a competence-specific ssDNA binding protein, SsbB, and stabilized by DprA and RecA (Attaiech et al., 2011; Berge et al., 2003) (**Figure 1**). This complex undergoes homology scanning and forms a temporal hetero-duplex during strand invasion which can lead to homologous recombination (Mortier-Barrière et al., 2007). The exact details on the kinetics of this process, as well as how the heteroduplex is resolved in most cells remains elusive. The competent transformation state in *S. pneumoniae* is transient as DprA interacts with phosphorylated ComE to inhibit its activity (Mirouze et al., 2013). In addition, several key Com proteins are rapidly turned over after their synthesis, leading to a window of DNA uptake of approximately 15 min (Liu et al., 2019; Tomasz, 1966; Weng et al., 2013).

As most work on pneumococcal competence and transformation has been performed using bulk assays, it is unclear what the actual bottlenecks are during competence development and why one cell will be transformed whereas another one will not. Here, we have set up single cell transformation assays that allow us to quantify successful recombination events in real-time. This study provides direct evidence for several decades-old models underpinning bacterial transformation, and offers new insights that help explain

why competence-induced transformation is so effective in changing global pneumococcal population structures.

Results

All pneumococci become competent and bind exogenous DNA

To quantify pneumococcal transformation efficiency and determine at which step potential bottlenecks arise, we systematically analyzed every stage during the process: 1) competence development, 2) production of the DNA uptake machinery, 3) binding of exogenous DNA, and 4) recombination and expression of the newly acquired genetic information (**Figure 1**). While competence development in *B. subtilis* is limited to approximately 10% of the population (Maamar and Dubnau, 2005; Smits et al., 2005), up to 100% of cells within pneumococcal populations have been reported to become competent when induced with exogenously added synthetic CSP or when grown on semi-solid surfaces (Bergé et al., 2017; Domenech et al., 2018; Litt et al., 1958; Martin et al., 2010; Moreno-Gámez et al., 2017; Slager et al., 2014).

To quantify competence development in clonal pneumococcal populations in a systematic fashion, we constructed a set of reporters. First, we assessed the timing of both naturally induced and artificially induced competence (by the addition of synthetic CSP) at the population level utilizing a firefly luciferase reporter under the control of the late competence *ssbB* promoter (strain DLA3). Cells were grown in C+Y medium at 37°C (see Methods) and growth and luciferase activity were measured every 10 min. As expected, under these experimental conditions, the population rapidly activates *ssbB* in the presence of added CSP, while in the absence of externally added CSP, the *ssbB* promoter peaks after approximately 100 min (**Figure 2A**). To determine which fraction of the cells switch on the competence pathway, we fused the *ssbB* promoter to a fast folding yellow fluorescent protein (msfYFP) and integrated this construct at the native *ssbB* locus (strain VL2219). As shown in **Figures 2B-C**, ~97% of the population was positive for *ssbB* expression 20 minutes after addition of synthetic CSP as determined by fluorescence microscopy followed by automated

image analysis (see Methods for details). Importantly, spontaneous competence without the addition of synthetic CSP was reached in 92% of the population showing that almost all pneumococci, regardless of their cell length and cell cycle status become naturally competent (**Figure 2–figure supplement 1**).

To test whether competent cells actually produce the machinery required for DNA uptake, we constructed translational msfYFP fusions to three essential components of the transformation machinery: ComGA (ATPase driving the DNA uptake pilus), ComEA (DNA receptor) and ComFA (ATPase driving DNA import) as the only copy integrated at their native locus. After 20 min of incubation with synthetic CSP, cells were collected for fluorescence microscopy. In line with the fraction of cells that become competent, msfYFP-ComEA, ComFA-msfYFP and ComGA-msfYFP were also expressed in the majority of the cells (~92%, ~99% and ~99%, respectively) (**Figure 2D** and **Figure 2–figure supplement 2**). A double-labeled strain (strain OVL2536: P*ssbB*-mScarlet-I, ComGA-msfYFP) demonstrated that all competent cells indeed produce the DNA uptake machinery (**Figure 2E**).

Finally, to assess whether the proteins required for recombination and chromosomal integration of exogenous DNA also were expressed in the majority of the population, we constructed translational fusions to RecA and the recombination mediator protein DprA. Similar to the DNA-uptake proteins, RecA and DprA were induced in most competent cells (**Figure 2D** and **Figure 2–figure supplement 2**).

During pneumococcal competence, the capture of extracellular DNA by the ComGC pilus is an essential step for transformation (Berge et al., 2002). To examine which proportion of cells is capable of binding DNA during competence, we labeled extracellular DNA (285 bp *S. pneumoniae* DNA fragment, see Methods) fluorescently with the Cy3 dye. After induction of competence with synthetic CSP of cells mutated for EndA (to prevent

degradation of the exogenous DNA), ~90% of the population bound extracellular DNA as visualized by fluorescence microscopy (**Figure 2F**). Even without additional CSP, spontaneous competence also led to most cells (89.6%) binding exogenous DNA (**Figure 2G**). As observed before in an unencapsulated R6 strain (Bergé et al., 2013), we note that also in the encapsulated serotype 2 D39V strain, DNA mainly bound to the mid-cell positions of the cell, corresponding to the localization of the DNA uptake machinery particularly the ComEA receptor (**Figure 2F** and **Figure 2–figure supplement 1**) (Bergé et al., 2013). Collectively, these data validate by direct single cell observations that pneumococcal competence development, the subsequent production of the DNA uptake and integration machinery, as well as DNA binding is highly efficient and occurs in nearly every cell of the population regardless of their cell cycle state.

Real-time single-cell analysis of homeologous recombination during transformation

Having established that there are no significant bottlenecks during the process of both induced and natural competence development and DNA uptake, we next set out to develop a system that allows for the direct visualization of successful recombination. Traditionally, transformation efficiencies are evaluated using antibiotic selection methods. However, these selection methods have limitations because they depend on the counting of colony forming units, which can lead to the overestimation of transformation efficiencies, due to inefficient separation of transformed from non-transformed daughter cells and nongenetic inheritance of antibiotic resistance (Dalia and Dalia, 2019; Domenech et al., 2018; Ephrussi-Taylor, 1962, 1958) (**Figure 3–figure supplement 1**). In order to overcome these concerns and analyze successful recombination events during transformation at the single cell level, we developed a fluorescence-based reporter system inspired by a system previously used to observe natural transformation in *S. pneumoniae* (Bergé et al., 2013) and other bacterial species (Boonstra et

al., 2018; Corbinais et al., 2016; Godeux et al., 2018). To do so, we utilized a fluorescent donor strain in which the gene encoding the abundant histone-like protein HlpA (aka HU) was fused in frame with the gene encoding the red fluorescent protein mScarlet-I integrated at the native *hlpA* locus at 169° on the circular chromosome (Keller et al., 2019) (strain VL1780) (**Figure 3A**). A recipient, non-fluorescent strain was constructed (strain VL1784) in which *hlpA* was separated from *mScarlet-I* by a stop codon mutation (G>T) (**Figure 3 A–C** and **Figure 3–figure supplement 2A**, *hlpA-stop-mScarlet-I*). Upon uptake, integration and expression of exogenous transforming DNA (tDNA) containing the donor construct (intact *hlpA-mScarlet-I*), successfully transformed recipient cells will produce functional HlpA-mScarlet-I that can be quantified by fluorescence microscopy or flow cytometry (**Figure 3B, 3C** and **Figure 3–figure supplement 2B**). As this is a recombination event between highly similar but not identical DNA (except for the SNP causing a stop codon), this is called a homeologous recombination event (Humbert et al., 1995; Petit et al., 1991). Note that this reporter system does not affect growth regardless of the presence of the stop codon and that flow cytometry analysis slightly overestimates the real transformation efficiencies due to cell chaining (**Figure 5** and **Figure 3–figure supplement 4**, see below).

As mScarlet-I is a fast folding red fluorescent protein (Bindels et al., 2017), this reporter system should allow for the real-time detection of homeologous recombination during transformation. To test this, we provided competent recipient cells that besides the *hlpA-stop-mScarlet-I* allele also constitutively expressed sfGFP (strain VL1832) with intact *hlpA-mScarlet-I* as donor tDNA in the presence of CSP and then performed time-lapse microscopy (see Methods for details). As shown in **Figure 3D** and **Videos 1 and 2**, recipient cells do not display any red fluorescence in the beginning and then gradually start to express red fluorescence. When quantifying the fluorescence signals and superimposing this on a cell lineage tree constructed using a set of new scripts written in BactMAP (Raaphorst et al.,

2020) (see Methods), it becomes apparent that the initial recipient cell already expresses HlpA-mScarlet-I right after the addition of tDNA before the first cell division as red fluorescent signals above background levels can be detected (**Figure 3E**). Notably, only half of the recipients' descendants appear to strongly express HlpA-mScarlet-I (**Figure 3E**, right lineage). Contrary, after three more divisions the non-transformed lineage no longer expresses red fluorescence (**Figure 3E**, left lineage). These results are in line with a recent study in *Vibrio cholerae* that showed a period of non-genetic inheritance in daughter cells during transformation (Dalia and Dalia, 2019). Similar observations were made when using a different transformation reporter system (**Figure 3–figure supplement 3** and **Video 3**, see below). In line with current models of transformation (Davidoff-Abelson and Dubnau, 1971; Ephrussi-Taylor and Gray, 1966; Fox and ALLEN, 1964; Gabor and Hotchkiss, 1966; LACKS, 1962; Méjean and Claverys, 1984; Piechowska and Fox, 1971), these observations are consistent with a model in which recombination occurs by direct integration of the ssDNA donor and forms a hetero-duplex. Therefore, at least one round of DNA replication and division is required to generate two different homo-duplex chromosomes in progeny cells (**Figure 3F**). The fact that we initially also observe fluorescence in the un-transformed lineage suggests that phenotypic expression derived from the acquired allele might occur prior to forming a homo-duplex. In this case, the transformed ssDNA likely replaced the antisense, noncoding strand so functional *hlpA-mScarlet-I* could be immediately transcribed after integration via RecA-directed homeologous recombination (mismatched pairing between exchanged DNA strands that are tolerated during the process of homologous recombination). Alternatively, phenotypic expression in these cells can occur if the transformed locus gets replicated, resulting in two homo-duplexes (transformed and original allele), and then transcribed before division of the cell (Dalia and Dalia, 2019).

Single cell quantification of homeologous recombination highlights transformation bottlenecks

The constructed system now allows us to quantify successful homeologous recombination events at the single cell level, without the bias introduced by traditional plating assays. Previously, it was shown that the concentration of donor DNA as well as the length of the homology regions strongly influences transformation efficiency (Keller et al., 2019; Lee et al., 1998). To examine recombination bottlenecks in our single cell setup, we treated our reporter recipient strain VL1784 with CSP and used intact *hlpA-mScarlet-I* donor tDNA with various lengths of homology surrounding the stop codon (fragments of 2.7 kb, 5 kb or 7 kb) at a range of different concentrations (0.0032 nM, 0.032 nM, 0.32 nM or 3.2 nM). Then, after 4 h incubation in liquid medium to allow for complete homo-duplex allele formation and dilution of non-genetically inherited HlpA-mScarlet-I, cells were separated from chains by vigorously shaking on a bead beater devise (see **Figure 3–figure supplement 4**). Finally, transformation efficiencies were quantified by flow-cytometry (**Figure 4A**). In line with studies using classical plating methods to assess transformation efficiencies (Keller et al., 2019; Lee et al., 1998), higher transformation frequencies were observed at higher donor DNA concentrations and with longer homology regions (**Figure 4B**). Interestingly, the frequency of transformation plateaued at ~50% regardless of the concentration of donor DNA and sequence homology length (**Figure 4B**). This is in contrast to reported transformation frequencies using traditional plating assays where transformation frequencies of higher than 75% (Ephrussi-Taylor, 1958) and up to 100%(Marie et al., 2017) have been reported. This discrepancy can be explained by the lack of separation of transformed from non-transformed cells within the counted colony (**Figure 3–figure supplement 1**). To exclude the possibility that the observed limitation in transformation frequency is due to an unique feature of the *hlpA-stop-mScarlet-I* reporter, we constructed an alternative reporter cassette in which we

translationally fused the superfolder green fluorescent protein (sfGFP) and SPV_1159, a nonessential small membrane protein under control of the strong constitutive P3 promoter (Keller et al., 2019; Sorg et al., 2015) cloned into the transcriptionally silent CEP locus at 295° on the circular chromosome (**Figure 4C**, strain VL1786). Based on this construct, a recipient strain was constructed containing a stop codon mutation in the linker between *spv_1159* and *sfGFP* (strain VL1788). Indeed, this *spv_1159-sfGFP*-based transformation reporter demonstrated similar transformation characteristics as the *hlpA-mScarlet-I* reporter in time-lapse microscopy and flow-cytometry analysis (**Figure 4–figure supplement 1** and **Video 3**). The transformation frequency of the *spv_1159-sfGFP* reporter was also dependent on donor DNA concentration and never exceeded ~50% (**Figure 4B**).

These data show that there is a limit on the maximum efficiency of transformation, despite the fact that most cells become competent and bind extracellular DNA (**Figure 2**) and support a model in which in general only one of the recipient allele strands is replaced by the donor DNA (Ephrussi-Taylor, 1966) (**Figure 3F**). Importantly, these experiments indicate that during competence-dependent transformation, given the donor DNA is of sufficient (homology) length and concentration (see Discussion), in principle all targeted loci can be replaced at least on one strand.

Recombination with tDNA is RecA-dependent and independent of mismatch repair

Previous work showed that pneumococcal genetic transformation involves the DNA mismatch repair (MMR) system, which is mediated by HexA (Claverys and Lacks, 1986), and it was suggested that certain alleles upon transformation might be particularly prone to repair (Ephrussi-Taylor, 1966). To test whether *hexA* plays a role in our reporter system, we quantified transformation efficiencies in a *hexA* mutant background. This showed no significant recombination differences compared to the wild-type background (**Figure 4–**

figure supplement 2). To test if our transformation reporter system depends on the competence-induced homologous recombinase, RecA, we depleted RecA expression level using CRISPR interference (Liu et al., 2017) (**Figure 4D**). In control strain VL3485 (P_{lac}_*dcas9*, without sgRNA), induction of dCas9 by IPTG did not affect the transformation efficiency with *hlpA-mScarlet-I* tDNA. However, when RecA expression was depleted by induction with IPTG in strain VL3486 (P_{lac}_*dcas9*+sgRNA-recA), the transformation frequency was decreased in an IPTG-dose dependent manner. Note that although RecA is known to be critical for optimal growth in *S. pneumoniae* (Mortier-Barrière et al., 1998), the CRISPRi depletion levels during competence did not affect bacterial growth (**Figure 4–figure supplement 3**). Collectively, this data show that our fluorescence-based transformation assay faithfully reflects RecA-dependent homeologous recombination events.

Cell-cycle independent homeologous recombination

It was previously suggested that the genomic location and the cell cycle state might influence transformation efficiency as each heteroduplex needs to be resolved to a homoduplex by DNA replication and loci located close to *oriC* will have multiple copies (Bergé et al., 2013; Dalia and Dalia, 2019; Ephrussi-Taylor and Gray, 1966; Porter and Guild, 1969). To test whether the genomic location of the recombination site and the read orientation of the newly acquired functional allele influences transformation efficiency, we introduced the *spv_1159-stop-sfGFP* reporter at 4 different positions on the chromosome: on the right arm of the chromosome at 101° of the circular chromosome (*bgaA* locus), near the terminus at 169° (*hlpA* locus), on the left arm of the chromosome at 295° (*cep* locus) and near *oriC* at 359° (*comCDE* locus) (**Figures 5A and B**). In addition, *spv_1159-stop-sfGFP* was introduced on both the positive and negative strand on the left arm of the chromosome (*cep* locus at 295°) and on the right arm of the chromosome (*bgaA* locus at 101°) (**Figure 5C**). As shown in

Figure 5, and **Figure 5–figure supplement 1**, transformation efficiencies were of a similar order across all tested loci and genetic orientations, with a maximal recombination efficiency of approximately 50%. We do note that certain loci consistently demonstrate higher transformation efficiencies than others (cf. CEP locus vs *bgaA* locus, Fig. 5-supplement 1), but no significant differences were observed regarding the orientation of the construct (see Discussion).

By performing time-lapse microscopy and tracking cell fates across several generations, we can, in principle, tell whether there was a preference for integration at either the leading or lagging strand (**Figure 5–figure supplement 2A**). By placing the direction of transcription of the reporter on the leading strand, RNAP will thus use the noncoding strand as template. In this situation, only if the noncoding strand is replaced by the donor DNA, fluorescence will be apparent during the first cell cycle upon transformation. If the donor DNA is integrated at the coding strand, it will take one more replication cycle before the heteroduplex is resolved and the noncoding strand contains the intact reporter and fluorescence will be observed later than in the first case (**Figure 5–figure supplement 2B**). Indeed, we can observe all different scenarios with transformants rapidly expressing HlpA-mScarlet-I (possible non-coding strand or double stranded recombinants) and cells that only express HlpA-mScarlet-I after the first cell division (possible coding strand recombinants) (**Figure 5–figure supplement 2C**). As we did not simultaneously track DNA replication in these cells, we cannot exclude the possibility that after transformation, a round of replication occurs before phenotypic expression. Nevertheless, together with the 'bulk' (FACS) single-cell transformation experiments described above, the time-lapse data strongly suggest that there is no preference for integration at either the leading or lagging strand and that this is an unbiased event. These findings correspond with work done in the 1960s and 1970s that showed that either strand of the incoming dsDNA is degraded randomly by EndA and either

strand has a similar chance of being integrated (Puyet et al., 1990). Recent work in *V. cholerae* demonstrated that 7% of transformation events occurred at both strands of the integration site, and it was speculated that this was because of integration of multiple donor ssDNA's replacing both the leading and lagging strand of the recipient. By recording 76 single cell transformation events using time-lapse microscopy, we found 6 cases in which both daughter cells (7.8%) expressed fluorescence, suggestive of double transformation events on both strands. These findings also indirectly indicate that heteroduplex DNA can be transcribed by RNAP and do not necessarily require a round of DNA replication to form homoduplex DNA (see below and (Uptain and Chamberlin, 1997)). Together, this data show that heteroduplexes with exogenous DNA are made across all available loci regardless of reading strand or distance to *oriC*.

Direct observations of multiple recombination events in single recipients

The previous experiments demonstrated that, under ideal conditions with long flanking homology regions and high DNA concentrations, all available recombination sites are transformed on at least one of the strands. Previous studies demonstrated that pneumococcal natural transformation is capable to deal with multiple donor DNAs for genetic recombination (Dalia et al., 2014; Lam et al., 2020). Also, it has been reported that the DNA-uptake and recombination process in *S. pneumoniae* is complete within 15 min (Berge et al., 2003), which is a shorter time window than the doubling time (Ephrussi-Taylor, 1966). In order to investigate the possibility of visualizing multiple recombination events, we constructed a dual reporter strain (strain VL1803), which harbors both *hlpA-stop-mScarlet-I* and *spv_1159-stop-sfGFP* at distinct chromosomal locations (**Figure 6A**). Transformation efficiencies of this reporter strain with each single donor DNA at the saturated concentration typically reached 50% for both *hlpA-mScarlet-I* and *spv_1159-sfGFP* as quantified by

361 microscopy (**Figures 6C-D**). When both donor DNA's were provided, double transformants
362 were observed (15.6 ± 4.4%) as well as single *hlpA-mScarlet-I* transformants (20.2 ± 8.9%)
363 and single *spv_1159-sfGFP* transformants (15.2 ± 4.9%). Time-lapse imaging of competent
364 recipient VL1803 cells with both donor DNAs clearly demonstrated that single recipients
365 could successfully recombine both fragments (**Figures 6B, 6C, Videos 4 and 5**). We note
366 that, on average, the fraction of non-transformed cells is close to 50% (48.9 ± 9.5%),
367 implying that each recombination event is not independent from the next or that there is an
368 upper limit to the number of successful recombinations, otherwise we would expect the
369 fraction of non-transformed cells to decrease with multiple donor DNAs (**Figure 6–figure**
370 **supplement 1**). An alternative model could be that each recombination event is independent
371 from the next but due to recombination events outside the stop codon SNP, which cannot be
372 quantified in our setup, a reduced transformation efficiency is recorded (see Discussion).
373 To further explore whether transformation efficiency indeed has a plateau, we
374 constructed a triple reporter strain (VL3127) that harbors *ftsZ-stop-mTurquoise2*, *spv_1159-*
375 *stop-msfYFP* and *hlpA-stop-mScarlet-I* at three different genomic locations (**Figure 7A**).
376 Beside the fact that the fluorescent proteins used are spectrally distinct, every fluorescent
377 reporter also has a specific cellular localization, facilitating automated image analyses of
378 successful recombination. The triple reporter strain was transformed with donor tDNA
379 fragments *ftsZ-mTurquoise2*, *spv_1159-msfYFP* and *hlpA-mScarlet-I*. After 4 h of incubation
380 for fluorescent protein maturation and chromosomal segregation, cells were assessed by
381 fluorescence microscopy. As shown in **Figure 7B** and **Video 5**, multiple transformed cells
382 with double or triple acquired fluorescence signals were readily observed. Next, we
383 performed single cell transformation assays with strain VL3127 providing one tDNA or all
384 three tDNAs and automatically quantified recombination efficiencies using Oufti and
385 BactMAP-based image analysis (Paintdakhi et al., 2016; Raaphorst et al., 2020) (**Figure 7C**).

In line with our previous observations, each single transformation with a saturated concentration of donor tDNA resulted in a recombination efficiency not higher than 50% (**Figure 7C**). Interestingly, every possible recombination event happened within the population: cells were observed in which just a single recombination event took place (the most occurring type of transformation), two recombination events (2.2 ± 0.9%, 4.1 ± 2.7% or 2.1 ± 1.8% for each possible combination) or even three recombination events (1.5 ± 1.1% of all cells). Nevertheless, more than half of the population (58.7 ± 13.4%) did not show any fluorescence when simultaneously transformed with three tDNAs. These observations support a model in which each transformation is in principle independent from the next (**Figure 6–figure supplement 1**).

Non-homologous DNA competes with homologous DNA to reduce transformation efficiency

The data so far show that in principle every locus in *S. pneumoniae* can be efficiently transformed to a maximum of 50% of the cells when providing tDNA at high DNA concentrations and with long homology arms. However, when providing multiple tDNAs, the untransformed fraction even increases and becomes greater than 50%. Recently, it was shown using whole genome sequencing of transformation events occurring during contact-dependent DNA uptake, that a single recipient could have at least 29 different recombination events (Cowley et al., 2018). Together, this suggests that many recombination events are likely going unnoticed in our single cell transformation assay and that these recombination events become limiting, as we can only detect successful recombination when the stop codon in our fluorescent reporter is replaced for a functional allele. If this is true, adding non-specific DNA would compete with donor tDNA resulting in reduced transformation efficiencies. To test this, we utilized homology-unrelated *E. coli*-derived DNA fragments of 5 kb with a

similar GC content to *S. pneumoniae* as competing donor DNA. Indeed, as shown in **Figure 7E**, co-transformation of strain VL1803 (*hlpA-stop-mScarlet-I*, *spv_1159-stop-sfGFP*) with *E. coli* DNA significantly reduced the transformation efficiency. When 0.32 μM of *hlpA-mScarlet-I* tDNA alone was used as 7 kb donor DNA, approximately 43 % of cells were transformed. However, when 0.32 μM of *hlpA-mScarlet-I* tDNA was given in the presence of saturating amounts of *E. coli* DNA (3.2 μM), only 3% of transformants were observed. Together, this data suggests that the fact that we never reach transformation efficiencies higher than 50% of the population even in the presence of multiple tDNAs, is because of saturation of the DNA uptake and integration machinery. The saturation can be caused by non-successful recombination events with donor DNA or successful recombination events with the donor DNA but outside of the stop SNP that cannot be detected in the single cell assay (**Figure 6–figure supplement 1**).

Efficient horizontal gene transfer in sessile co-cultures

So far, we demonstrated that transformation is highly efficient under ideal and saturated experimental settings in which pure PCR products were used as donor DNA. Previous studies showed that natural environments also promote efficient HGT (Cowley et al., 2018; Domenech et al., 2020). To assess transformation potential under more realistic conditions, we tested transformation efficiency in a co-culture system in which two pneumococcal strains are grown together on a semi-solid surface without adding synthetic CSP (Fig. 8A; see Materials and Methods section for detail procedure). HGT in such systems is based on genomic DNA released by dead cells followed by DNA uptake and transformation of competent recipient cells (Domenech et al., 2020). Here we used strain R895 (*cmR*) as recipient and strain R4692 (Δ*comCDE*, *smR*, *nov*^{R}) as donor. Both are genetically identical unencapsulated R800 derivatives (Lefevre et al., 1979) except for a single SNP conferring

streptomycin (SNP in *rpsL*) or novobiocin (SNP in *gyrB*) resistance and a chloramphenicol resistance cassette present in the recipient R895 strain (Fig. 8A). Strain R4692 is also unable to activate competence due to a *comCDE* deletion so transformation can only occur in one direction from donor (R4692) to recipient (R895). R895 and R4692 were pre-cultured separately until early exponential phase and then mixed in an approximate ratio of 3:7 (see Methods). The mixture was immediately spotted on agar plates followed by incubation at 37°C for 4 h to allow spontaneous competence development and transformation between strains. Cells were collected by scraping them from the plates and separated by sonication. Serial dilutions of the resulting cell suspension were plated with 4.5 µg/mL of chloramphenicol (for the recovery of the total number of viable recipient cells) and with combinations of chloramphenicol plus streptomycin (10 µg/mL) and/or novobiocin (4 µg/mL) (for the recovery of the single or double transformed recipient cells). As shown in Figure 8B, also in this more realistic model, very high transformation efficiencies are obtained with a single transformation efficiency with *smR* or nov^R of 5.70×10^{-2} (SD, 5.70×10^{-2}) or 1.75×10^{-2} (SD, 1.68×10^{-2}), respectively. Double transformation efficiency with both *smR* and nov^R was 8.01×10^{-5} (SD, 9.92×10^{-5}), which is close to the product of the single transformation efficiencies ($5.70 \times 10^{-2} \times 1.75 \times 10^{-2} = 9.9 \times 10^{-4}$). As a control, we also performed experiments using strain R4574 as donor (same genotype as R4692, but not harboring *smR* or *novR* allele), which never generated streptomycin nor novobiocin resistant R895, demonstrating that *de novo* mutations conferring resistance do not occur in this experimental setup. Together, these experiments support our single cell observations that multiple transformation events occur efficiently and independently even in more realistic settings within sessile co-cultures.

Discussion

The species of *Streptococcus pneumoniae* is vastly diverse with a core genome of approximately 500-1100 orthologues genes and a pan-genome of 5000-7000 orthologs (Hiller and Sá-Leão, 2018). In addition, many genes are mosaic such as several genes encoding for penicillin-binding proteins in penicillin-resistant clinical strains (Hakenbeck et al., 2012). One of the main reasons for the high level of genome plasticity and rapidly changing population dynamics is because of the highly conserved competence-based transformation system present in nearly all pneumococcal genomes (Croucher et al., 2016). Indeed, rapid spread of antibiotic resistance alleles and capsule loci have been observed among human populations under selective pressure (Chewapreecha et al., 2014). Here, we investigated the molecular basis for competence-dependent transformation at the single-cell level and show that the uptake, integration and expression of tDNA is highly efficient and is largely independent from the recipient's cell cycle stage or of the chromosomal position of the target locus. This was made possible by the setup of a sensitive real-time detection system to quantify successful homeologous recombination events. A major benefit of the here established single cell approach over traditional plate-based assays is that it allows for the detection of more subtle effects and offers better resolution to study the kinetics of the processes involved. Indeed, using the system developed here, we could visualize and quantify the recombination of three different tDNAs in single recipient cells demonstrating the efficiency of the pneumococcal transformation process.

Genome sequencing has indicated that up to 29 recombination events may have taken place in a single round of transformation in the same cell when selecting for the transfer of an antibiotic resistance allele in *S. pneumoniae* (Cowley et al., 2018), while 40 recombination events have been reported in *B. subtilis* (Carrasco et al., 2016). Our work now provides direct evidence that this is not an anomaly and that multiple recombination events are possible

during a single transformation episode, even in the absence of selection. Besides shedding light on the efficiency by which transformation can happen in *S. pneumoniae*, by imaging transformation at the single cell level, we provide direct evidence that typically only one recipient strand is replaced during competence-dependent transformation, and that there is no bias towards replacement of the leading or lagging strand. As observed in *V. cholerae*, in approximately 7% of transformants, both strands can be replaced, which is likely caused by DNA repair leading to removal of the recipient strand on the heteroduplex or by integration of multiple tDNAs (Dalia and Dalia, 2019). This is in line with predictions made using unlinked antibiotic resistance alleles (Porter and Guild, 1969). In addition, our single cell observations suggest that the replaced noncoding strand by recombination within the heteroduplex is immediately transcribed by RNAP and can lead to lineages of cells with non-genetic inherited phenotypes, or that the transformed allele is replicated and transcribed well before cell division occurs (**Figure 3**).

We show that any site regardless of its chromosomal position or orientation with regards to DNA replication can be efficiently transformed, although not with the exact same efficiencies (**Figure 5–figure supplement 1**). Possible explanations for local difference in recombination efficiency could be the levels of DNA compaction or transcription activity. As RecA-mediated DNA strand exchange is a reversible reaction *in vitro* (Dutreix et al., 1991; Konforti and Davis, 1990), under steady state conditions DNA strand exchange rarely reaches 50% efficiency. However, *in vivo*, when providing a single tDNA to competent cells, we readily reach 50% DNA strand exchange, again highlighting that this process is highly efficient under our experimental conditions.

Interestingly, we find that the percentage of untransformed cells is lower when three tDNAs are provided instead of two tDNAs (~58% vs ~49% of untransformed cells, respectively: **Figures 6 and 7**). Together with the observation that the presence of non-

510 homologous DNA reduced our observed transformation efficiency (**Figure 7E**), suggests
511 that, in principle, every recombination event is independent of the next, but that many
512 unsuccessful recombination events and successful recombination events outside the stop
513 codon of our reporter are taking place and that this limits the efficiency of site-specific
514 recombination (**Figure 6–figure supplement 1** and **Figure 7F**).

515 The overall biological implication of the limitation on competence-dependent
516 transformation is that this mechanism ensures that in most cases one copy of the original
517 recipient DNA remains unaltered. This might represent a fail-safe scenario so that in case a
518 deleterious tDNA is incorporated, at least one daughter cell will survive. While this might be
519 considered as a "spandrel' effect: a characteristic that flows inevitably from a selected
520 phenotype but has not been selected for directly (Gould and Lewontin, 1979), being able to
521 safely sample from a large pan-genome might contribute to the vast genome plasticity and
522 genome diversity as observed in natural pneumococcal populations. Interestingly, we also
523 find highly efficient HGT and independent transfer of genetic markers between a donor and
524 recipient pneumococcal strain growing together on agar plates (Fig. 8), indicating that our
525 single cell observations under laboratory conditions also reflect settings that depend on lysis
526 of the donor cell and uptake of chromosomal DNA. It will be interesting to see how efficient
527 competence-dependent transformation and horizontal gene transfer is under more realistic
528 conditions such as within polymicrobial community within a host. Future single-cell work
529 will allow the investigation of the localization of the enzymes involved in transformation,
530 how strand exchange during transformation occurs and what the dynamics of the molecular
531 machines are during DNA uptake, integration and expression of tDNA.
532

Materials and Methods

Key Resources Table

Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
strain, strain background (*Streptococcus pneumoniae*)	Various	This paper	NCBI Taxon: 1313	See Supplementary file 1
sequence-based reagent	Various oligonucleotides	This paper (Sigma-Aldrich)	Primers for cloning	See Supplementary file 2
chemical compound, drug	D-Luciferine	Synchem	bc219; CAS: 115144-35-9	
software, algorithm	FIJI	doi: 10.1038/nmeth.2019.	RRID:SCR_002285	
software, algorithm	Oufti	doi: 10.1111/mmi.13264	RRID:SCR_016244	
software, algorithm	BactMAP	doi.org/10.1111/mmi.14417	https://github.com/veeninglab/BactMAP	
software, algorithm	SuperSegger	doi: 10.1111/mmi.13486	https://github.com/wiggins-lab/SuperSegger	

Bacterial strains and growth condition

All pneumococcal strains used in this study are derivatives of serotype 2 *S. pneumoniae* D39V (Avery et al., 1944; Slager et al., 2018) unless specified otherwise. See Table S1 for a list of the strains used and the Supplemental information for details on the construction of the strains. *S. pneumoniae* was grown in C+Y (pH 6.8) medium at 37°C. C+Y was adapted from Adams and Roe (Adams and Roe 1945) and contained the following compounds: adenosine (68.2 mM), uridine (74.6 mM), L-asparagine (302 mM), L-cysteine (84.6 mM), L-glutamine (137 mM), L-tryptophan (26.8 mM), casein hydrolysate (4.56 g L^{-1}), BSA (729 mg L^{-1}), biotin (2.24 mM), nicotinic acid (4.44 mM), pyridoxine (3.10 mM), calcium pantothenate (4.59 mM), thiamin (1.73 mM), riboflavin (0.678 mM), choline (43.7 mM), $CaCl_2$ (103 mM), K2HPO4 (44.5 mM), $MgCl_2$ (2.24 mM), $FeSO_4$ (1.64 mM), $CuSO_4$ (1.82 mM), $ZnSO_4$ (1.58 mM), $MnCl_2$ (1.29 mM), glucose (10.1 mM), sodium pyruvate (2.48 mM), saccharose (861 mM), sodium acetate (22.2 mM) and yeast extract (2.28 g L^{-1}).

Strain construction

Construction of *ssbB*::*ssbB_msfYFP* (VL2219)

To construct YFP reporter for *ssbB* transcription, monomeric *yfp* (*myfp*) was introduced immediately downstream of *ssbB* at the native *ssbB* locus together with an RBS. *myfp* gene was amplified with OVL1414 and OVL1417 from genomic DNA of MK308 (*parB::parB-yfp*) (Raaphorst et al., 2017). Upstream and downstream fragments were amplified with primer pairs of OVL166/OVL1196 and OVL1199/OVL167 using genomic DNA of VL599 (ssbB::ssbB_luc_kanR) (Slager et al., 2014) as template, respectively. The three resulting fragments were digested with BsmBI, ligated and transformed into *S. pneumoniae* D39V to obtain strain VL2219.

Construction of *comGA*::*comGA-msfYFP* (VL2536)

562 To construct translational fusion of *comGA* and *msfYFP*, *linker-mYFP* was amplified
563 by PCR with OVL351/OVL690 from genomic DNA of MK308 (parB::parB-yfp) (Raaphorst
564 et al., 2017) . 'Upper' and 'downer' fragments containing *comGA* were amplified with primer
565 pairs OVL391/OVL392 and OVL691/OVL394 using genomic DNA of D39V as template,
566 respectively. The three resulting fragments were fused by overlap PCR and transformed into
567 *S. pneumoniae* D39V. Transformed clones were screened by PCR and sequenced. Resulting
568 strains were additionally transformed by *ssbB::ssbB_mScarlet-I,kan* fragments, obtaining
569 strain VL2536.
570
571 **Construction of *ssbB:ssbB_mScarlet-I, kanR* (VL2536)**
572 To construct transcriptional fusion of *ssbB* and *mScarlet-I*, *mScarlet-I* was amplified
573 by PCR with OVL1415/OVL1418 from genomic DNA of VL1787 (*cep:spv_1159-mScarlet-*
574 *I, spcR*) (Keller et al., 2019). Upper and downer fragments were amplified with
575 OVL166/OVL1168 and OVL1199/OVL167 using genomic DNA of VL599 (*ssbB::ssbB_luc,*
576 *kanR*) as template, respectively. The three resulting fragments were fused by Golden Gate
577 assembly using BsmBI and transformed into *S. pneumoniae* strain, *comGA::comGA-msfYFP*.
578 Transformed clones were selected with kanamycin and sequenced, obtaining strain VL2536.
579
580 **Construction of *comEA::msfYFP-comEA* (VL2537)**
581 To construct translational fusion of *comEA* and *msfYFP*, *mYFP-linker* was amplified
582 with OVL2029/OVL2028 from genomic DNA of VL1818 (*comEC::msfYFP-*
583 *comEC*)(Veening lab collection). Upper and downer fragments containing *comEA* was
584 amplified with OVL354/OVL1664 and OVL2026/OVL357 using genomic DNA of VL870
585 (*comEA::mNeonGreen-comEA*)(Veening lab collection) as template, respectively. The three
586 resulting fragments were fused by Golden Gate assembly using BsmBI and transformed into

587 *S. pneumoniae* D39V. Transformed clones were screened by PCR and sequenced, obtaining
588 strain VL2537.
589
590 **Construction of *comFA*::*comFA-msfYFP* (VL2538)**
591 To construct a translational fusion of *comGA* and *msfYFP*, *linker-mYFP* gene was
592 amplified with OVL351/OVL690 from genomic DNA of MK308 (parB::parB-yfp)
593 (Raaphorst et al., 2017). Upper and downer fragments containing *comFA* were amplified with
594 OVL358/OVL521 and OVL1129/OVL361 using genomic DNA of D39V as template,
595 respectively. The three resulting fragments were fused by overlap PCR and transformed into
596 *S. pneumoniae* D39V. Transformed clones were screened by colony PCR and sequenced,
597 obtaining VL2538 strain.
598
599 **Construction of *dprA*::*dprA-msfYFP, eryR* (VL3355)**
600 To construct a translational fusion of *dprA* and *msfYFP*, *linker-mYFP* gene was
601 amplified with OVL3481/OVL3482 from genomic DNA of *cep::spcR, P3_spv_1159-msfYFP*
602 (codon-optimized) strain (Rueff AS and Veening JW, unpublished). Upper and downer
603 fragments containing *dprA* were amplified with OVL3487/OVL3488 and
604 OVL3489/OVL3490 using genomic DNA of D39V as template, respectively. Erythromycin
605 resistance marker (*eryR*) was amplified OVL2549/OVL2771 using genomic DNA of
606 *hexA::eryR* strain (Veening lab collection). The four fragments were fused by Golden Gate
607 assembly using BsmBI and transformed into *S. pneumoniae* D39V. Transformed clones were
608 selected by erythromycin and sequenced, obtaining strain VL3355.
609
610 **Construction of *hlpA::hlpA_hlpA-mScarlet-I* (VL1780)**

To construct *hlpA-mScarlet-I*, the *hlpA-mScarlet-I* gene was introduced downstream of the original *hlpA* gene at its own locus as a second copy of *hlpA*. Upper and downer fragments were amplified by PCR with OVL43/OVL44 and OVL45/OVL46 using genomic DNA of MK119 (*hlpA::hlpA_hlpA-mKate2_cmR*) (Beilharz et al., 2015), respectively. *mScarlet-I* gene was amplified by PCR with OVL55 and OVL56 using codon-optimized synthetic *mScarlet-I* gene as template (Keller et al., 2019) . The three resulting fragments were fused by overlap PCR and transformed into *S. pneumoniae*. Transformed clone was selected by chloramphenicol, obtaining VL1780.

Construction of *hlpA::hlpA_hlpA-stop-mScarlet-I* (VL1784, VL1832)

To disrupt translation between *hlpA* and *mScarlet-I,* on *hlpA::hlpA_hlpA-mScarlet-I, cmR* construct, single nucleotide mutation was introduced in domain breaking linker between *hlpA* and *mScarlet-I*. Upper or downer fragments was amplified by PCR with OVL43/OVL724 or OVL873/OVL46 using genomic DNA of VL1780 as template. The resulting fragments were fused by overlap PCR and transformed into *S. pneumoniae* D39V. Transformed clone was selected by chloramphenicol, obtaining VL1784.

To obtain VL1832 (*hlpA::hlpA_hlpA-mScarlet-I, cmR; CEP::sfGFP, spcR*), *CEP::P3_sfGFP, spcR* fragment was amplified by PCR with OVL37/OVL40 using genomic DNA of D-PEP33 (*CEP::spcR,P3_sfGFP*) (Sorg et al., 2015). The fragment was transformed into VL1784, and transformed clone was selected by spectinomycin, obtaining VL1832.

Construction of *CEP::spcR, P3_spv_1159-sfGFP* (VL1785, VL1800)

To construct membrane localizing sfGFP, hypothetical protein with transmembrane domain, spv_1159, was translationally fused to sfGFP under the control of synthetic constitutive promoter P3 at the CEP locus of the *S. pneumoniae* chromosome. Upper and downer

636 fragments were amplified by PCR with OVL37/OVL631 and OVL634/OVL40 using
637 genomic DNA of VL430 (*CEP::spcR,P3_spv_1159-sfGFP*) (Keller et al., 2019), respectively.
638 *spv_1159* was amplified by PCR with OVL632 and OVL633 using genomic DNA of D39V.
639 The three resulting fragments were fused by Golden Gate assembly with BsmBI and
640 transformed into *S. pneumoniae* D39V. Transformants were selected on Colombia agar plates
641 containing spectinomycin, obtaining strain VL1785.

642 To obtain a dual labeled strain, the *CEP::spcR,P3_spv_1159-sfGFP* fragment was
643 amplified by PCR with OVL37/OVL40 using genomic DNA of VL1785, and transformed
644 into VL1780 to obtain VL1800 (*hlpA::hlpA_hlpA-mScarlet-I,cmR*; *CEP::spcR,*
645 *P3_spv_1159-sfGFP*).

646

647 **Construction of *CEP::spcR, P3_spv_1159-stop-sfGFP* (VL1788, VL1803, VL1930)**
648 To disrupt translation between *spv_1159* and *sfGFP,* on the *CEP::spcR,P3_spv_1159-sfGFP*
649 construct, a single nucleotide mutation was introduced in domain breaking linker between
650 *spv_1159* and *sfGFP*. Upper and downer fragments were amplified by PCR with
651 OVL37/OVL724 or OVL723/OVL40 using genomic DNA of VL1785 as template. The
652 resulting fragments were fused by overlap PCR and transformed into *S. pneumoniae* D39V.
653 Transformants were selected by chloramphenicol, obtaining strain VL1786.

654 To obtain dual reporter strain VL1800, the *CEP::spcR,P3_spv_1159-stop-sfGFP*
655 fragment was amplified by PCR with OVL37/40 using genomic DNA of VL1786, and
656 transformed into VL1784 to obtain VL1800 (*hlpA::hlpA_hlpA-stop-mScarlet-I,cmR*;
657 *CEP::spcR, P3_spv_1159-stop-sfGFP*)).

658

659 **Construction of *hlpA::spcR, P3_spv_1159-sfGFP* (VL3096)**

660 To insert the *spcR, P3_spv_1159-sfGFP* construct downstream of the *hlpA* locus, *spcR,*
661 *P3_spv_1159-sfGFP* was amplified by PCR with OVL2855/OVL2856 using genomic DNA
662 of VL1785 (*CEP::spcR,P3_spv_1159-sfGFP*) as template. Upper and downer fragments
663 were amplified by PCR with OVL2868/OVL2859 and OVL2860/OVL2869 using genomic
664 DNA of D39V, respectively. The three resulting fragments were fused by Golden Gate
665 assembly with BsmBI and transformed into *S. pneumoniae* D39V. Transformants were
666 selected by spectinomycin, obtaining strain VL3096.
667
668 **Construction of *hlpA::spcR, P3_spv_1159-stop-sfGFP* (VL3097)**
669 To disrupt translation between *spv_1159* and *sfGFP,* on *hlpA::spcR, P3_spv_1159-*
670 *sfGFP* construct, single nucleotide mutation was introduced in domain breaking linker
671 between *spv_1159* and *sfGFP*. Upper or downer fragments were amplified by PCR with
672 OVL2868/OVL724 or OVL723/OVL2869 using genomic DNA of VL3096 as a template.
673 The resulting fragments were fused by overlap PCR and transformed into *S. pneumoniae*
674 D39V. Transformants were selected on Colombia agar plates containing chloramphenicol,
675 obtaining VL3097.
676
677 **Construction of *comCDE::spcR, P3_spv_1159-sfGFP* (VL3098)**
678 To insert the *spcR, P3_spv_1159-sfGFP* construct right downstream of the *comCDE* locus,
679 *spcR, P3_spv_1159-sfGFP* was amplified by PCR with OVL2855/OVL2856 using genomic
680 DNA of VL1785 (*CEP::spcR, P3_spv_1159-sfGFP*) as template. Upper and downer
681 fragments were amplified by PCR with OVL371/OVL2861 and OVL2862/OVL2870 using
682 genomic DNA of D39V, respectively. The three resulting fragments were fused by Golden
683 Gate assembly with BsmBI and transformed into *S. pneumoniae* D39V. Transformed clone
684 was selected by spectinomycin, obtaining VL3098.

685

686 **Construction of *comCDE::spcR, P3_spv_1159-stop-sfGFP* (VL3099)**

687 To disrupt translation between *spv_1159* and *sfGFP,* on
688 *comCDE::spcR,P3_spv_1159-sfGFP* construct, single nucleotide mutation was introduced in
689 domain breaking linker between *spv_1159* and *sfGFP*. Upper or downer fragments were
690 amplified by PCR with primers OVL371/OVL724 or OVL723/OVL2870 using genomic
691 DNA of VL3096 as template. The resulting fragments were fused by overlap PCR and
692 transformed into *S. pneumoniae* D39V. Transformants were selected on Colombia agar plates
693 containing, obtaining VL3099.

694

695 **Construction of *bgaA::spcR, P3_spv_1159-sfGFP* (VL3100, VL3348)**

696 To insert the *spcR, P3_spv_1159-sfGFP* construct right at the *bgaA* locus, *spcR,*
697 *P3_spv_1159-sfGFP* was amplified by PCR with OVL2855/OVL2856 using genomic DNA
698 of VL1785 (*CEP::spcR, P3_spv_1159-sfGFP*) as template. Upper and downer fragments
699 were amplified by PCR with OVL1312/OVL2863 and OVL2864/OVL2871 using genomic
700 DNA of D39V, respectively. The three resulting fragments were fused by Golden Gate
701 assembly with BsmBI and transformed into *S. pneumoniae* D39V. Transformed clone was
702 selected by spectinomycin, obtaining VL3100.

703 To obtain strain VL3348, the *bgaA::spcR, P3_spv_1159-sfGFP* fragment was
704 amplified by PCR with OVL1312/2871 using genomic DNA of VL3100, and transformed
705 into VL1780 to obtain VL3348 (*hlpA::hlpA_hlpA--mScarlet-I,cmR*; *bgaA::spcR,*
706 *P3_spv_1159-sfGFP*)).

707

708 **Construction of *bgaA::spcR, P3_spv_1159-stop-sfGFP* (VL3101, VL3349)**

To disrupt translation between *spv_1159* and *sfGFP,* on *bgaA::spcR, P3_spv_1159-sfGFP* construct, single nucleotide mutation was introduced in domain breaking linker between *spv_1159* and *sfGFP*. Upper or downer fragments was amplified by PCR with OVL1312/OVL724 or OVL723/OVL2871 using genomic DNA of VL3096 as template. The resulting fragments were fused each other by overlap PCR and transformed into *S. pneumoniae* D39V. Transformed clone was selected by chloramphenicol, obtaining VL3101.

To obtain dual reporter strain, the *bgaA::spcR, P3_spv_1159-stop-sfGFP* fragment was amplified by PCR with OVL1312/2871 using genomic DNA of VL1786, and transformed into VL1784 to obtain VL3349 (*hlpA::hlpA_hlpA-stop-mScarlet-I, cmR*; *bgaA::spcR, P3_spv_1159-stop-sfGFP*).

Construction of *CEP::spcR,P3_spv_1159-sfGFP(inverted)* (VL3346)

To re-introduce *spcR, P3_spv_1159-sfGFP* in inverted direction at *CEP* locus, *spcR, P3_spv_1159-sfGFP* was amplified by PCR with OVL3358/OVL3359 using genomic DNA of VL1785 (*CEP::spcR, P3_spv_1159-sfGFP*) as template. Upper and downer fragments were amplified by PCR with OVL37/OVL3390 and OVL3391/OVL40 using genomic DNA of D39V, respectively. The three resulting fragments were fused by Golden Gate assembly with BsmBI and transformed into *S. pneumoniae* VL1784. Transformed clone was selected by spectinomycin, obtaining VL3346 (*hlpA::hlpA_hlpA-mScarlet-I,cmR*; *CEP::spcR, P3_spv_1159-sfGFP(inverted)*).

Construction of *CEP::spcR, P3_spv_1159-stop-sfGFP(inverted)* (VL3347)

To disrupt translation between *spv_1159* and *sfGFP,* on *CEP::spcR, P3_spv_1159-sfGFP(inverted)* construct, single nucleotide mutation was introduced in domain breaking linker between *spv_1159* and *sfGFP*. Upper or downer fragments was amplified by PCR with

OVL37/OVL723 or OVL724/OVL40 using genomic DNA of VL3346 as template. The resulting fragments were fused each other by overlap PCR and transformed into *S. pneumoniae* VL1784. Transformed clone was selected by chloramphenicol, obtaining VL3347 (*hlpA::hlpA_hlpA-stop-mScarlet-I,cmR*; *CEP::spcR,P3_spv_1159-stop-sfGFP(inverted)*).

Construction of *bgaA::spcR, P3_spv_1159-sfGFP(inverted)* (VL3350)

To re-introduce *spcR, P3_spv_1159-sfGFP* in inverted direction at *CEP* locus, *spcR, P3_spv_1159-sfGFP* was amplified by PCR with OVL3358/OVL3359 using genomic DNA of VL1785 (*CEP::spcR, P3_spv_1159-sfGFP*) as template. Upper and downer fragments were amplified by PCR with OVL1312/OVL2863 and OVL2864/OVL2871 using genomic DNA of D39V, respectively. The three resulting fragments were fused by Golden Gate assembly with BsmBI and transformed into *S. pneumoniae* VL1784. Transformed clone was selected by spectinomycin, obtaining VL3346 (*hlpA::hlpA_hlpA-mScarlet-I,cmR* ;*CEP::spcR, P3_spv_1159-sfGFP(inverted)*).

Construction of *bgaA::spcR, P3_spv_1159-stop-sfGFP(inverted)* (VL3351)

To disrupt translation between *spv_1159* and *sfGFP,* on *bgaA::spcR, P3_spv_1159-sfGFP(inverted)* construct, single nucleotide mutation was introduced in domain breaking linker between *spv_1159* and *sfGFP*. Upper or downer fragments was amplified by PCR with OVL1312/OVL723 or OVL724/OVL2871 using genomic DNA of VL3346 as template. The resulting fragments were fused each other by overlap PCR and transformed into *S. pneumoniae* VL1784. Transformed clone was selected by chloramphenicol, obtaining VL3347 (*hlpA::hlpA_hlpA-stop-mScarlet-I, cmR*; *CEP::spcR, P3_spv_1159-stop-sfGFP(inverted)*).

759

760 **Construction of *ftsZ-mTurquoise2* (VL3126)**

761 To construct a triple labeled strain, upper or downer fragments was amplified with primer
762 pair OVL452/OVL1921 or OVL1922/OVL1441 using genomic DNA of *ftsZ::ftsZ-*
763 *mTurquoise2, spcR* strain (Gallay C and Veening JW, unpublished) as template. These
764 fragments were fused by overlap PCR to remove the spcR gene. The fused fragment was
765 used for transformation in D39V and spectinomycin-susceptible clone was selected.

766 To construct triple labeled strain, *hlpA::hlpA_hlpA-mScarlet-I,cmR* was amplified
767 with OVL43/OVL46 using genomic DNA of VL1780 and *cep::spcR,P3_spv_1159-msfYFP*
768 was amplified with OVL37/OVL40 using genomic DNA of *cep::spcR,P3_spv_1159-msfYFP*
769 *(codon-optimized)* strain (Rueff AS and Veening JW, unpublished). These two DNA
770 fragments were transformed into the strain (*ftsZ::ftsZ-mTurquoise2*) and transformed clone
771 was selected by chloramphenicol and spectinomycin, obtaining triple labeled strain VL3126
772 (*ftsZ::ftsZ-mTurquoise2; hlpA::hlpA_hlpA-mScarlet-I,cmR; cep::spcR, P3_spv_1159-*
773 *msfYFP*).

774

775 **Construction of *ftsZ-stop-mTurquoise2* (VL3127)**

776 To construct triple reporter strain, upper or downer fragments was amplified with primer pair
777 of OVL452/OVL724 or OVL723/OVL1441 using genomic DNA of VL3126 as template.
778 These fragments were fused each other by overlap PCR to introduce stop codon between *ftsZ*
779 and *mTurquoise2*. The fused fragment *ftsZ::ftsZ-stop-mTurquoise2* was transformed in
780 *ftsZ::ftsZ-mTurquoise2* strain and clone that lost mTurquoise2 fluorescence was screened by
781 fluorescence microscopy.

782 To disrupt translation between *spv_1159* and *msfYFP,* on *CEP::spcR, P3_spv_1159-*
783 *msfYFP* construct, single nucleotide mutation was introduced in domain breaking linker

between *spv_1159* and *msfYFP*. Upper or downer fragments was amplified by PCR with OVL37/OVL724 or OVL723/OVL40 using genomic DNA of VL3126 as template. While, *hlpA-::hlpA-stop-mScarlet-I, cmR* was amplified by PCR with OVL43/OVL46 using VL1784. The resulting two fragments were transformed into *ftsZ::ftsZ-stop-mTurquoise2* and transformed clone was selected by chloramphenicol and spectinomycin, obtaining VL3128.

Construction of *recA::recA-mCheryy, eryR* (VL361)

mCherry-eryR was amplified with RR93/RR94 using VL371 (RR27) (Raaphorst et al., 2017). Upper or downer fragments was amplified by PCR with RR91/RR92 or RR93/RR94 using genomic DNA of D39V as template. The three fragments were assembled using Gibson one-step ISO assembly (Gibson, 2011) and transformed into D39V. Transformed clone was selected by erythromycin.

Construction of R4692 (Δ*comCDE::trmpR; strR; rifR; novR*)

R304 (*strR; rifR; novR*) (Chastanet et al., 2001) strain was transformed with genomic DNA from R4574 (Δ*comCDE::trmpR*) (Johnston et al., 2020) and was selected by trimethoprim, obtaining R4692.

Luminescence assays of competence development

To monitor competence development, strains containing a transcriptional fusion of the firefly *luc* gene with the late competence gene *ssbB* were used. Cells were pre-cultured in C+Y (pH 6.8) at 37ºC to an OD595 nm of 0.2. Right before inoculation, cells were collected by centrifugation (6,000 xg for 3 minutes) and resuspended in fresh C+Y at pH 7.9, which is permissive for natural competence. Luciferase assays were performed in 96-wells plates with a Tecan Infinite 200 PRO illuminometer (TECAN) at 37ºC as described before (Slager et al.,

2014). Luciferin was added at a concentration of 0.45 mg/mL to monitor competence by means of luciferase activity. Optical density (OD595nm) and luminescence (relative luminescence units [RLU]) were measured every 10 minutes.

Phase contrast and fluorescence microscopy

Microscopy acquisition was performed using a Leica DMi8 microscope with a sCMOS DFC9000 (Leica) camera and a SOLA light engine (Lumencor) and a 100x/1.40 oil-immersion objective. Images were primarily processed using LAS X (Leica). For snap shot imaging, cells were concentrated 10x by centrifugation (6,000 xg, 3 min) and 0.5 µl of cells were spotted on 1% agarose/PBS. For time-lapse microscopy, a semi-solid growth surface was prepared with C+Y (pH 7.9) containing 1% agarose in Gene Frame (Thermo Fischer) (Jong et al., 2011). As C+Y medium has some background fluorescence, the C+Y agar pad was pre-exposed on a UV illuminator for 1 min to bleach the background fluorescence.

Phase contrast images were acquired using transmission light with 100 ms exposure for snap shot and 50 ms exposure for time-lapse. Fluorescence was usually acquired with 700 ms exposure for snap shot, and 200–500 ms exposure (17–30% of power from Sola light engine) for time-lapse using filter settings described below. Time-lapses images were recorded by taking images every 5 or 10 minutes.

Leica DMi8 filters set used are as followed: mTurquoise2 (Ex: 430/24 nm Chroma, BS: LP 455 Leica 11536022, Em: 470/24 nm Chroma ET470/24 nm or Ex: 430/29 nm Chroma, BS: 455 (450–490) Chroma 69008, Em: 470/26), sfGFP (Ex: 470/40 nm Chroma ET470/40x, BS: LP 498 Leica 11536022, Em: 520/40 nm Chroma ET520/40m), msfYFP (Ex: 500/20 nm Chroma ET500/20x, BS: LP 520 Leica 11536022, Em: 535/30 nm Chroma ET535/30m or Ex:495/25 nm Chroma ET495/25x, BS520 (510–560) Chroma 69008, Em: 533/30 nm) and mScarlet-I (Chroma 49017, Ex: 560/40 nm, BS: LP 590 nm, Em: LP 590 nm

or Ex: 575/35 nm, BS: 595 (590–670) nm Chroma 69008, Em: 635/70 nm). Microscopy image are available at the BioImages Archive (accession S-BIAD26).

Quantitative image analysis

For quantitative image analysis of single cells, obtained microscopic images were processed by FIJI software (Schindelin et al., 2012). Single cell segmentation and fluorescence signal intensity measurement were performed by Oufti (Paintdakhi et al., 2016). The generated celllist files were analyzed in R (https://www.r-project.org/), using BactMAP (Raaphorst et al., 2020) for statistical analysis and visualization. After celllist file were imported into R, cells were filtered between 0.7–1.2 µm in width length to exclude false events derived from noise or miss-segmentation. Threshold of fluorescence of signal intensity was defined based on negative or positive control for each experiment setting. >500 cells were analyzed at least for each replicate. To exclude the possibility of overlap in detection of fluorescence (particularly mTurquoise2/msfYFP and msfYFP/mScarlet-I) in multi-fragments transformation, we ensured that single transformation experiments did not show any signal in the other channels and this was confirmed by looking at the protein localization patterns.

For generating cell lineage trees from time-lapse imaging, the stacked time-lapse images were processed by FIJI and stabilization between time frames was performed by Huygens (Scientific volume imaging). Single cell segmentation and fluorescence intensity acquisition were performed by SuperSegger (Stylianidou et al., 2016). The resultant data set was analyzed using BactMAP (Raaphorst et al., 2020).

DNA binding assays

Analysis of DNA binding was performed in an *endA* mutant background (strain D39V *ssbB::luc (cam) endA::kan*), to favor accumulation of transforming DNA at the surface of

competent cells. In wild type, *endA*$^+$ cells, surface-bound DNA is immediately internalized into the cytosol or degraded otherwise, which makes surface-bound DNA accumulation hard to be visualized as previously shown (Bergé et al., 2013). After gently thawing stock cultures, aliquots were inoculated at an OD550 of 0.006 in C+Y medium, supplemented with 20 mM HCl to prevent spontaneous competence development, and grown at 37°C to an OD550 of 0.3. These precultures were inoculated (1/50) in C+Y medium (pH 7.8) and incubated at 37°C. In these conditions, competence developed spontaneously and reached its maximal level in the population after 55-60 minutes. At 35 minutes, 1 ml samples were collected and induced, or not, with synthetic CSP (50 ng/ml). At 50 minutes, these samples were incubated for 5 minutes with 10 ng of a 285 bp DNA fragment labelled with a Cy3 fluorophore at its 5' extremities (Bergé et al., 2013). Cells were pelleted (3,000 xg, 3 min), washed twice in 500µl C+Y, and resuspended in 20 to 50 µL C+Y medium before microscopy. Two µl of this suspension was spotted on a microscope slide containing a slab of 1.2% C+Y agarose as described previously (Jong et al., 2011).

Phase contrast and fluorescence microscopy were performed with an automated inverted epifluorescence microscope Nikon Ti-E/B, a phase contrast objective (CFI Plan Apo Lambda DM 100X, NA1.45), a Semrock filter set for Cy3 (Ex: 531BP40; DM: 562; Em: 593BP40), a LED light source (Spectra X Light Engine, Lumencor), and a sCMOS camera (Neo sCMOS, Andor). Images were captured and processed using the Nis-Elements AR software (Nikon). Cy3 fluorescence images were false colored red and overlaid on phase contrast images. Overlaid images were further analyzed to quantify the number of cells bound with Cy3-labelled DNA. Single cells were first detected using the threshold command from Nis-Elements and cells bound or not to DNA were manually classified using the taxonomy tool.

884 **Evaluation of transformation frequency using the fluorescence reporter**

885 To quantify the efficiency of transformation with tDNA fragments, reporter cells were pre-
886 cultured in C+Y (pH 6.8) at 37ºC to an OD595 nm of 0.2. Right before inoculation, cells
887 were collected by centrifugation (6,000 xg for 3 minutes) and resuspended in fresh C+Y at
888 pH 7.9, adjusted to OD=0.1. Competence was induced by incubation in the presence of CSP
889 (100 ng/µl) at 37°C for 10 min and then donor tDNA was provided the indicated
890 concentration. After an additional 4 h of incubation at 37°C for complete cell division to form
891 homo-duplex and maturation of fluorescence proteins, cells were placed on ice to stop cell
892 growth and were directly analyzed by fluorescence microscopy or flow-cytometry.

893 Donor tDNA was designed in such a way that the single nucleotide mutation is
894 positioned in the middle of the entire fragment so that the left and right homology arms are of
895 equal length. Preparation of the tDNA was performed by PCR using primer pairs indicated in
896 Supplementary table S2, using the corresponding parent strain as template. For competition
897 experiments using unrelated tDNA as shown in **Figure 7**, a DNA fragment that has no
898 homology to the pneumococcal genome but is of equal size (5 kb) and GC content (~40% of
899 GC) to the *hlpA-mScarlet-I*/*spv_1159-sfGFP* fragments, was amplified by PCR from
900 genomic DNA of *E. coli* DH5alpha (Hanahan et al., 1991), using primers OVL3527 and
901 OVL3528.

902

903 **Flow-cytometry analysis**

904 Cells were collected by centrifugation (6,000 xg for 3 minutes) and resuspended in filtered
905 (0.22 µm) PBS adjusted to a cell density of approximately $1.0x10^5$~$1.0x10^6$ cells/mL. As
906 encapsulated *S. pneumoniae* D39V cells tend to form chains particularly during competence,
907 we separated cells by bead beating (BioSpec) without any glass beads. At least >$1.0x10^4$
908 events were analyzed on a Novocyte flow cytometer (ACEA bioscience) harboring 488 nm

and 561 nm lasers. Fluorescence filters used were: FITC (Ex: 488nm, Em: 530/45nm) for sfGFP and PE.Texas.Red (Ex: 561 nm, Em: 615/20 nm) for mScarlet-I. Obtained raw data were imported and analyzed in R. Non-bacterial particles were excluded by gating the FSC and SSC values. A threshold was determined so that positive events counted in a negative control strain were <1% and validated with both negative (no DNA control) and positive (parent strain without the point mutation) control for each experimental setting.

Transformation assays between *S. pneumoniae* strains

We used strain R895 as a recipient strain and R4692 and R4574 as donor strains. R895 is a naturally competent derivative of unencapsulated strain R6 and carries a chloramphenicol resistance marker (Chastanet et al., 2001). R4692 is unable to develop competence and carries point mutations conferring resistance to streptomycin, novobiocin and rifampicin. R4574 was used as a negative control donor strain. All strains were pre-cultured in C+Y (pH 6.8) at 37°C to an OD550 nm of 0.1. Cells were then collected by centrifugation (6,000 xg for 3 minutes) and concentrated 3-fold in fresh C+Y at pH 7.9. Mixed inoculates containing 30 µL of donor strain R895 and 70 µL of recipient strain (R4574 or R4692) were subsequently spotted onto Petri dishes containing 3% horse blood CAT-agar (1% agar) supplemented with catalase (300 U/ mL). Different ratios of donor and recipient were tested (1:1, 1:10 and 3:7) and a 3:7 ratio showed the least HGT variation across experiments. After 4 hours incubation at 37°C, cells were scraped off the plates and separated by sonication in an ultrasonic cleaner (90 sec, 50 Hz). Serial dilutions were plated with 4.5 µg/mL of chloramphenicol and with combinations of chloramphenicol plus streptomycin (10 µg/mL) and/or novobiocin (4 µg/mL). Transformation efficiency was calculated by dividing the number of transformants by the total number of viable recipient count. Four independent biological replicates were performed.

Author Contributions and Notes

J.K. and J.W.V. designed research, J.K., N.C., R.v.R. and G.C. performed research, J.K. and N.C. analyzed data; and J.K. P.P. and J.W.V. wrote the paper.

The authors declare no conflict of interest.

This article contains supporting information online.

Acknowledgments

We thank Jelle Slager and Arnau Domenech for critically reading this manuscript, two anonymous referees, Juan Carlos Alonso and all members of the Veening lab for stimulating discussions. Work in the Veening lab is supported by the Swiss National Science Foundation (SNSF) (project grant 31003A_172861), a JPIAMR grant (40AR40_185533) from SNSF and ERC consolidator grant 771534-PneumoCaTChER. Work in the Polard lab is supported by the Centre National de la Recherche Scientifique, Université Paul Sabatier, and the Agence Nationale de la Recherche (grants ANR-10-BLAN-1331 and ANR-17-CE13-0031). Jun Kurushima was supported by The Naito Foundation.

952 **Figure legends**

953 **Figure 1. Regulation of pneumococcal competence and transformation.**

954 Schematic overview showing representative competence-related genes involved in
955 pneumococcal transformation. Competence development is initiated by activation of the early
956 *com* genes (shown in blue area). ComAB exports ComC and processes it into the competence
957 stimulating peptide (CSP). The two-component system ComDE recognizes CSP and
958 positively regulates the early *com* genes. Subsequently, the alternative sigma factor ComX,
959 activates late *com* gene expression (shown in orange area). ComGA, GB and GC are
960 assembled to form the DNA-binding pilus. EndA is the endonuclease that cleaves dsDNA
961 into ssDNA. ComEA, ComEC, ComFA and ComFC form the ssDNA uptake channel.
962 Internalized foreign ssDNA is protected by SsbB and DprA. DprA ensures the loading of
963 RecA on single strand tDNA to form a presynaptic filament and the resulting DNA scanning
964 complex is capable of homologous (or homeologous) recombination with the recipient
965 chromosome.

966

967 **Figure 2. Single cell analysis of competence activation and DNA binding**

968 **A.** Kinetics of bacterial growth and competence development. Growth curves (top) and OD-
969 normalized bioluminescence activity (bottom) of strain DLA3 (P*ssbB-luc*) in the presence
970 (orange) or absence (grey) of CSP. Arrow indicates the moment after addition of CSP (0
971 min). Lines and confidence bands represent means of three replicates and SD, respectively.
972 **B.** Single cell imaging of fluorescence competence reporter cells. VL2219 (P*ssbB- msfYFP*)
973 was treated with (top, grey frame) or without (bottom, orange frame) CSP for 20 min and
974 analyzed by fluorescence microscopy. Images are overlays of phase contrast and YFP signal.
975 Scale bar: 20 μm. **C.** Quantification of the imaging. Scatter plots of single cell YFP signal
976 intensity (y axis) against cell length (x axis), based on microscopy images. Red line indicates

the threshold used to score YFP positive cells. Proportion of positive cells (%) is shown. **D.** Quantification of cells expressing the transformation machinery. Fluorescence signal intensity for indicated strain harboring *comGA-msfYFP* (VL2536), *msfYFP-comEA* (VL2537), *comFA-msfYFP* (VL2538) or *dprA-msfYFP* (VL3355) treated with (orange) or without (grey) CSP for 20 min. Red line indicates threshold for YFP positive cells. Proportion of positive cells (%) is shown. **E.** Correlation between competence activation and ComGA production. VL2536 (*comGA-msfYFP*, P*ssbB_mScarlet-I*) was incubated with or without CSP. Scatter plot of single cell YFP signal intensity (translational fusion of ComGA) against mScarlet-I signal (transcriptional fusion to *ssbB*). Red line indicates threshold used. Proportion of positive cells (%) is shown on each plot. **F** and **G.** DNA binding analysis using Cy3-labeled DNA added to induced- or spontaneous competent cells. **f.** Representative image of Cy3-labeled DNA-bound TD290 (*ssbB_luc*, Δ*endA*) cells. Scale bar: 2 μm. **G.** Quantification of microscopy images of Cy3-labeled DNA bound D39V cells. Bacteria were treated with CSP for 20 min to induce competence in C+Y (pH 7.9). Efficiency of spontaneous competent cells (no CSP) is also represented. Total of 6027 cells (without added CSP) and 3082 cells (with CSP) were collected over three independent experiments.

Figure 3. Development of a fluorescence-based real-time reporter for genetic transformation in *S. pneumoniae*

A. Schematic representation of the reporter system. *hlpA-mScarlet-I* was inserted downstream of the native *hlpA* locus as a second copy (strain VL1780), resulting in red-fluorescently marked nucleoids as shown by fluorescence microscopy (Left: phase contrast, middle: red fluorescence, right: overlay, scale bar: 4 μm). A single nucleotide mutation generating a stop codon was introduced in the linker sequence between *hlpA* and *mScarlet-I*,

resulting in non-fluorescent strain VL1784 (*hlpA-stop-mScarlet-I*). **B.** Flow-cytometry measurement of *hlpA-mScarlet-I* signal of VL1780 (*hlpA-mScarlet-I*, white) and VL1784 (*hlpA-stop-mScarlet-I*, grey). **C.** CSP-treated (right) or untreated (left) VL1784 was provided with tDNA (*hlpA-mScarlet-I*) and analyzed by fluorescence microscopy after 4 h of incubation. Scale bar, 20 μm. **D.** Time-lapse visualization of transformation with *hlpA-mScarlet-I* in VL1832 (VL1784+constitutively expressing cytoplasmic sfGFP). VL1832 was treated with CSP for 10 min, tDNA added (*hlpA-mScarlet-I*) for 10 min, and then spotted on C+Y agarose pad to start time-lapse imaging with a 5 min interval. Signal of constitutively expressed cytoplasmic sfGFP (top, green in the overlay) was used for cell segmentation in image analysis. Successfully transformed cells were detected by expression of HlpA-mScarlet-I (middle, red in the overlay). Scale bar, 10 μm. Also see **Video 1**. **E.** Cell lineage tree with superimposed fluorescence intensity was built based on the time-lapse image shown in E. The quantified mean mScarlet-I signal intensity of each cell during its cell cycle was plotted as a color-coded dot onto the lineage tree with each dot corresponding to the moment of 'birth'. Note that the tree represents the lineage from only one of the two progenitor cells (indicated by white arrow in panel d). **F.** Working model for DNA integration and chromosomal segregation of the transformed allele. tDNA is internalized as ssDNA, and recombines to replace one strand on the host chromosome forming a hetero-duplex after recombination. Following chromosomal replication and segregation, the two daughter cells have distinct homo-duplexes (either the original allele or the tDNA allele). Note that initial hetero-duplex formation might permit for phenotypic expression from the newly acquired allele if the noncoding strand is replaced by tDNA (see main text).

Figure 4. Single cell quantification of recombination reveals an upper level of transformation efficiency

A. Quantification of transformation frequency by flow-cytometry. CSP-treated VL1784 was transformed with *hlpA-mScarlet-I* tDNA of various lengths (2.7 kb, 5 kb or 7 kb) at differing tDNA concentrations (0.0032 nM, 0.032 nM, 0.32 nM or 3.2 nM). The single nucleotide variant to repair the point mutation is located in the middle of each fragment (**Figure 3–figure supplement 2**). After 4h incubation post tDNA addition, cell chains were separated (see **Figure 3–figure supplement 4**) and analyzed by flow-cytometry. Negative control without any donor DNA (filled grey histogram) and positive control (VL1780, open histogram) is shown in all panels. Red vertical line indicates the threshold used to score mScarlet-I positive cells. **B.** Correlation between transformation frequency and donor DNA concentration. Transformation frequency was plotted against final concentration of donor DNA. Frequency was calculated by dividing the number of cells with a FL intensity above the threshold by the total number of cells based on flow-cytometry data, as shown in panel A. **C.** Alternative transformation reporter present on a different chromosomal position. The *spv_1159-sfGFP* reporter fusion was cloned into the CEP locus (295°; VL1786, VL1788). A point mutation resulting in a stop codon was introduced in the linker sequence separating *spv_1159* and *sfGFP* in VL1788. **D.** Reduction of transformation efficiency by CRISPRi-based *recA* depletion. CRISPRi-based depletion strains VL3485 (P_{lac}_*dcas9*, no sgRNA control) and VL3486 (P_{lac}_*dcas9*, sgRNA targeting *recA*) were introduced in the *hlpA-stop-mScarlet-I* reporter strain. Strains were pre-grown with or without IPTG (0, 0.01, 0.1 or 1 mM) in acidic C+Y (pH 6.8), and then incubated with CSP (100 ng/µl) in fresh C+Y (pH 7.8) provided with donor tDNA (5 kb length, 0.32 nM). After 4 h of phenotypic expression, transformed cells were analyzed by flow-cytometry.

Figure 5. Effect of chromosomal position and strand on recombination potential

A. The spv_1159-sfGFP reporter was cloned into various loci; CEP (295°; VL1786, VL1788), *hlpA* (169°; VL3096, VL3097), *comCDE* (359°; VL3098, VL3099) or *bgaA* (101°; VL3100, VL3101). A point mutation that generates a stop codon was introduced in the linker sequence between *spv_1159* and *sfGFP* for each strain (VL1788, VL3097, VL3099 or VL3101). **B.** Flow cytometry analysis on transformations with intact *spv_1159-sfGFP* tDNA. Strain VL1788, VL3097, VL3099 or VL3101 was transformed with intact *spv_1159-sfGFP* tDNA with 5 kb homology arm at the final concentration of 3.2 nM. 4 h post transformation, cells were separated by beat beating and analyzed by flow-cytometry. The red vertical line indicates the threshold of positive cells in spv_1159-sfGFP signal expression. **C.** Genetic orientation effect on transformation efficiencies. All dual reporter strains VL1803, VL3347, VL3349 or VL3351 harbor both *hlpA-stop-mScarlet-I* and *spv_1159-stop-sfGFP* reporters, but *spv_1159-sfGFP* was cloned at distinct chromosomal positions and different reading directions. *spv_1159-stop-sfGFP* was cloned at the CEP (295°) or *bgaA* (101°) locus resulting in strains VL1803/VL3347 or VL3349/3351, respectively. The coding strand of spv_1159-sfGFP was cloned either in the same direction as the DNA replication fork (read direction on the leading strand) (green triangle, strains VL3347/3351) or in the opposite direction (read direction on the lagging strand) (green square, strains VL1803/3349). The strains were treated with CSP and transformed with corresponding *spv_1159-sfGFP* tDNA (5 kb, 3.2 nM) alone. Transformants were analyzed by flow-cytometry. Vertical blue lines represent the threshold for green fluorescence intensity. Experiments were performed at least three times and FACS analysis of a typical experiment are shown.

Figure 6. Dual transformation at distinct chromosomal positions

A. Graphical representation of dual reporter strains. Dual transformation reporter strain VL1804 harbors the two transformation reporter constructs *hlpA-stop-mScarlet-I* and *spv_1159-stop-sfGFP*, at the *hlpA* and *CEP* loci, respectively. Donor tDNA (5 kb) *hlpA-mScarlet-I* and *spv_1159-sfGFP* were amplified from strain VL1800 and used for transformation. **B.** Time-lapse visualization of double transformation. Dual reporter strain VL1803 (*hlpA-stop-mScarlet-I, spv_1159-stop-sfGFP*) was treated with CSP for 10 min, provided with 3.2 nM of both *hlpA-mScarlet-I* and *spv_1159-sfGFP* tDNAs (5 kb) for 10 min, and then spotted on a C+Y agarose pad to start time-lapse imaging at 5 min intervals. Successfully transformed cells were detected by expression of HlpA-mScarlet-I (middle, red in merge) and spv_1159-sfGFP (bottom panels, green in merge). Red and Green arrows indicate single transformed cells with *hlpA-mScarlet-I* and *spv_1159-sfGFP* tDNA, respectively. Yellow arrows indicate doubly transformed cells. Scale bar: 4 μm. See **Video 4**. **C.** Cell lineage tree with superimposed fluorescence intensity was built based on the time-lapse image shown in B. Means of mScarlet-I (pink) and sfGFP (green) signal intensity of each cell was calculated and displayed with a color bar. Yellow, green and red arrows indicate double transformed, single *spv_1159-sfGFP*-transformed and single *hlpA-mScarlet-I*-transformed lineages, respectively. **D.** Snap shots and quantitative image analysis of transformed populations. Strain VL1803 was transformed with single (*hlpA-mScarlet-I* or *spv_1159-sfGFP*) or double (*hlpA-mScarlet-I*/*spv_1159-sfGFP*) tDNA(s) (5 kb) at final concentration of 3.2 nM. After 4 h of incubation, still images were obtained and the fluorescence intensities were quantified and plotted. Scale bar: 10 μm. Experiments were performed at least three times and microscopy analysis of a typical experiment are shown. **E.** Proportion of transformed phenotypes. Stacked bars represent the fraction of single transformed (red or green), double transformed (yellow) and non-transformed (grey) cells. Population of each transformed phenotype was quantified from microscopy images. Bars

represent mean ± SD of three independent replicates. Analyzed data of the positive control strain VL1800 (*hlpA-mScarlet-I, spv_1159-sfGFP*) and negative control non-transformed strain VL1803 (*hlpA-stop-mScarlet-I, spv_1159-stop-sfGFP*) are also shown, demonstrating the accuracy of the threshold used to score positive transformants.

Figure 7. Direct observation of recombination of three separate tDNAs during a single transformation event

A. Schematic representation of the triple labeled strain VL3126 harboring three reporter cassettes: *hlpA-mScarlet-I*, *spv_1159-msfYFP* and *ftsZ-mTurquoise2* at the *hlpA*, *CEP* and *ftsZ* loci, respectively. Strain VL3127 contains stop codon mutations in the linker between each of the fluorescent fusion proteins. Gray arrows indicate the direction of the DNA replication fork relative to the reporter cassette. **B.** Microscope image of strain VL3127 treated with CSP and transformed with the tDNAs of *hlpA-mScarlet-I*, *spv_1159-sfGFP* and *ftsZ-mTurquioise2* (3.2 nM each) amplified from VL3126. Merge image of phase contrast, cyan (FtsZ-mTurquoise2), yellow (spv_1159-msYFP) and red (HlpA-mScarlet-I) fluorescence is shown. Scale bar: 10 μm. **C** and **D.** Proportion of transformed phenotypes. Population of each transformed phenotype was quantified from microscope images. Representative images for each phenotype are shown. Scale bar: 2 μm. **D.** Stacked bars represent proportion of single transformed (cyan, yellow or red), double transformed [green (cyan+yellow), blue (cyan+red), orange (yellow+red)], triple transformed (white) and non-transformed (grey) cells. Bars represent mean ± SD of three independent replicates. Analyzed data of the positive control strain VL3126 (*ftsZ-mTurquoise2, spv_1159-msfYFP, hlpA-mScarlet-I*) and negative control non-transformed strain VL3127 (*ftsZ-stop-mTurquoise2,*

spv_1159-stop-msfYFP, hlpA-stop-mScarlet-I) are also shown, demonstrating the accuracy of the threshold used to score positive transformants.

E. Competition effect of unrelated DNA on transformation frequency. CSP-treated VL1803 was transformed with 7 kb *hlpA-mScarlet-I* tDNA at the final concentration of 0.32 nM in the absence or the presence of an unrelated DNA fragment (0.32 nM or 3.2 nM) amplified from *E. coli*. After incubation of 4 h post transformation, cells were separated and analyzed by flow-cytometry. Red vertical line indicates the threshold of positive cells in mScarlet-I signal expression. The proportion of positive cells (%) is depicted in the plots. **F**. Fragmented tDNA recombination model. The fluorescence-based reporters used in this study rely on replacement of the stop codon SNP (grey star) by intact (amino acid coding) SNP (yellow star) that is located in the middle of the tDNA fragment (orange line). All prepared tDNA molecules have obviously intact SNP, but, integration into host chromosome may take place outside the SNP, which is never distinguished from true untransformed cells by the fluorescence-based system and effectively acting as competing DNA for tDNA's that transform the SNP.

Figure 8. Horizontal gene transfer between *S. pneumoniae* strains

A. Schematic representation of the transformation assay between *S. pneumoniae* strains. Pre-cultured recipient strain R895 (*cmR*) and donor strain R4692 *(ΔcomCDE, smR*, nov^R) were mixed in approximately 3:7 ratio, and spotted on 3% horse blood CAT-agar (see Methods). After 4h incubation at 37°C to allow strain-to-strain transformation, cells were scraped and separated by sonication. Then, serial dilutions of cell suspensions were plated with 4.5 µg/mL of chloramphenicol (Cm, for the recovery of the total number of viable recipient cells) and with combinations of chloramphenicol plus streptomycin (Sm, 10 µg/mL) and/or novobiocin (Nov, 4 µg/mL). **B**. Transformation efficiency was calculated by dividing the number of

transformants by the total number of viable recipient count. Four independent biological replicates were performed and box plots demonstrate the average efficiencies.

Figure supplements

Figure 2–figure supplement 1. Single cell quantification of spontaneously activated competence.

A. Competence reporter strain VL2219 (cytoplasmic msfYFP reporter of transcriptional fusion to *ssbB*) was incubated in C+Y in the absence of CSP. After 100 min (peak of spontaneous competence, see **Figure 2A**), the cells were analyzed by fluorescence microscopy. Images are overlays of phase contrast and YFP signal. Scale bar, 20 µm. **B.** Quantification of competence positive cells. Scatter plot of single cell YFP signal intensity (y axis) against cell length (x axis), based on microscope images shown in panel a. Red line indicates threshold used to score YFP positive cells. Proportion of positive cells (%) is shown.

Figure 2–figure supplement 2.

Fluorescence microscopy images of VL2536 (*msfYFP-comEA*), VL2537 (*comFA-msfYFP*), VL2538 (*comGA-msfYFP*), VL3355 (*dprA::dprA-msfYFP*) and VL361 (*recA::recA-mCherry*) treated with (orange frame) or without (grey frame) CSP for 20 min. Images are overlays of phase contrast and YFP signal. Scale bar, 20 µm. Magnified images of induced condition are also shown in the inset. Scale bar, 2 µm.

Figure 3–figure supplement 1. Classical methodology for transformation frequency estimation by antibiotics selection.

A. Workflow representation of the experiment. D39V wild type (Sm susceptible) or its *rpsL* mutant derivative (Sm resistant) were treated with synthetic CSP and transformed with Sm resistance (*rpsL**) or Sm susceptible ($rpsL_{WT}$) of *rpsL* allele, respectively (5kb, SNP in the middle, 3.2 nM of DNA). With or without additional incubation for 2 h in liquid C+Y medium, then serially diluted bacterial suspensions were plated onto two Columbia agar plates for each. For phenotypic expression, agar plates were incubated 2 h in 37°C and then overlaid with additional Columbia agar containing streptomycin or no antibiotics, followed by over-night incubation. Colony forming units (cfu) were counted for each plate and the ratio of Sm resistant cfu (on +Sm plate) divided by total cfu (on plate without antibiotic) was calculated to evaluate the transformation frequency. **B.** Histogram represents mean and SD (from three replicates) of Sm resistant cfu against total cfu for each transformation setting. Note that when transforming the *smR* host, *smR* cfu reflects the un-transformed population.

Figure 3–figure supplement 2. Fluorescence-based detection of successful genetic transformation.

A. Stop codon generating SNP introduced in reporter strain. Nucleotide sequence of linker between *hlpA* (grey) and *mScarlet-I* (red) in strains VL1780 (up) and VL1784 (bottom) are shown. Bold letter under the nucleotide sequence indicates translated amino acids. In parent strain VL1780 (*hlpA-mScarlet-I*), translation of *hlpA* and *mScarlet-I* are bridged by a linker sequence (39 bp, 13 amino acids). In reporter strain VL1784 (*hlpA-stop-mScarlet-I*), a single nucleotide mutation (G>T) was induced, resulting in a stop codon in the middle of the linker so that *mScarlet-I* is not translated unless it is transformed with the intact sequence (i.e. *hlpA-mScarlet-I*). Note that other reporter genes used in this study have the identical linker sequence and SNP. **B.** Schematic representation of the fluorescence-based transformation reporter. Once genetic recombination of donor DNA (intact PCR fragment of *hlpA-mScarlet-*

I from VL1780) into strain VL1784 succeeds, the stop codon between *hlpA* and *mScarlet-I* is repaired, which allows for production of the fluorescent protein.

Figure 3–figure supplement 3. Growth curves of the reporter strains.

Pre-cultured *S. pneumoniae* strains D39V (wild type), VL1800 (*hlpA-mScarlet-I, spv_1159-sfGFP*). VL1803 (*hlpA-stop-mScarlet-I*, *spv_1159-stop-sfGFP*), VL3126 (*ftsZ-mTurquoise2, spv_1159-msfYFP, hlpA-mScarlet-I*) and VL3127 (*ftsZ-stop-mTurquoise2, spv_1159-stop-msfYFP, hlpA-stop-mScarlet-I*), were inoculated into fresh C+Y media at initial concentration of OD = 0.1 with or without CSP, which is the exact same cultivation condition to transformation experiment implemented in this study. Lines and confidence bands represent means of three replicates and SD, respectively.

Figure 3–figure supplement 4. Effect of multicellular filament disruption by beadbeater on estimation of transformation frequency in flow-cytometry analysis.

A. Effect of cell separation treatment by beadbeater on FSC-SSC profile. Dual reporter strain VL1803 treated with CSP was transformed with mock (no DNA) or single (*hlpA-mScarlet-I* fragment or *spv_1159-sfGFP*) or double tDNA fragments. After 4 h incubation for phenotypic expression and chromosome segregation, cells were analyzed by flow-cytometry. Without beadbeater treatment (bottom panels), FSC-SSC profile showed increased value and heterogeneity compared to beadbeater cells (top panels). It is noted that one event in flow-cytometry does not indicate one single cell but one particle, which might consist of a multicellular filament. **B.** Microscopic images of cells prepared in panel a. Bead beater-untreated cells showed chained phenotype consisting of >4 cells, meanwhile bead beater-treated cells showed shorter chain consisting of 2–4 cells, in line with FSC-SSC profile in flow-cytometry (panel A). Scale bar, 20 μm. **C.** Effect of cell separation treatment by

beadbeater on apparent transformation frequency in flow-cytometry. Note that bead beater-untreated cells tend to show generally more population of fluorescence positive cells compared to beadbeater-treated cells. This may be explained by masking of the negative cells by positive cells within the same chain as height values (highest peak) are measured for fluorescence. Vertical or horizontal red line represent threshold for green or red fluorescence intensity, respectively. **D.** Correlation between particle size and apparent transformation frequency. Distributions of FSC value (particle size in flow-cytometry) grouped by transformed- (including both of each single and double transformed population, orange bars) or un-transformed (grey bars) population are shown. In separated cells, there is no significant difference in FSC profile between un-transformed and transformed particles. Meanwhile, in beadbeater-untreated cells, transformed group tends to show higher FCS value than untransformed group. events with high FSC (longer filament) tend to have more chance to be estimated as transformed.

Figure 4–figure supplement 1. Time-lapse imaging of transformation with *spv_1159-sfGFP* fragment in VL1788.

A. Time-lapse visualization of transformation with *spv_1159-sfGFP* fragment in VL1788 (CEP::P3_*spv_1159-stop-sfGFP*). VL1788 was treated with CSP for 10 min, added with tDNA (*spv_1159-sfGFP*) for 10 min, and then spotted on C+Y agarose pad to start time-lapse imaging with 5 min intervals. Phase contrast images (left, panels) were obtained for cell segmentation in image analysis. Successfully transformed cells were detected by expression of SPV_1159-sfGFP (middle panels). Also see **Video 2** for original movie. **B.** Cell lineage tree with fluorescence intensity was built based on the time-lapse image in panel A. Means of sfGFP signal intensity of each cells were calculated and displayed with ball color. Note that

fluorescence signals detected in the initial few time points were mostly background from C+Y medium rather than that derived from SPV_1159-sfGFP.

Figure 4–figure supplement 2. Effect of *hexA* knockout on transformation frequency in the fluorescence-based reporter. Dual reporter strain VL1803 (*hlpA-stop-mScarlet-I*, *spv_1159-stop-sfGFP*) and its derivative VL1930 (VL1803+Δ*hexA* were treated with CSP and transformed with *hlpA-mScarlet-I* or/and *spv_1159-sfGFP* tDNA (3.2 μM each). The transformed cells were analyzed by flow-cytometry. Vertical or horizontal red line represent threshold for green or red fluorescence intensity, respectively.

Figure 4–figure supplement 3. Effect of *recA* knock-down on bacterial growth. Pre-cultured *S. pneumoniae* strains, DL3485 (*prs1::lacI; bgaA::Plac_dcas9*, *hlpA::hlpA_hlpA-stop-mScarlet-I*) and DL3485 (*prs1::lacI; bgaA::Plac_dcas9*, *hlpA::hlpA_hlpA-stop-mScarlet-I*) were inoculated into fresh C+Y media at initial concentration of OD = 0.01 with or without CSP and IPTG. Lines and confidence bands represent means of three replicates and SD, respectively.

Figure 5–figure supplement 1. Chromosomal position effect on correlation between transformation frequency and tDNA concentration. The spv_1159-sfGFP reporter was cloned into various loci; CEP (295°; VL1788), *hlpA* (169°; VL3097), *comCDE* (359°; VL3099) or *bgaA* (101°; VL3101). A point mutation that generates a stop codon was introduced in the linker sequence between *spv_1159* and *sfGFP* for each strain (**Figure 5A**). Flow cytometry analysis on transformations with intact *spv_1159-sfGFP* tDNA. Strains VL1788, VL3097, VL3099 or VL3101 were transformed with intact *spv_1159-sfGFP* tDNA

with 5 kb homology arm at the final concentration of 0, 0.032, 0.32 or 3.2 nM. 4 h post transformation, cells were separated by bead beating and analyzed by flow-cytometry. Transformation frequency was plotted against final concentration of donor DNA. Frequency was calculated by dividing the number of cells with a FL intensity above the threshold by the total number of cells based on flow-cytometry data.

Figure 5–figure supplement 2. Effect of chromosomal strand to be replaced by tDNA on progression of phenotypic expression during cell division.

A. Two possible recombination outcomes of single stranded tDNA into the host chromosome are shown. As tDNA (orange line) is processed to a single stranded DNA molecule during internalization, recombination can only occur on one strand of the host chromosome. Therefore, if the tDNA encodes new phenotypic information, which of (i) coding or (ii) noncoding strand to be replaced will affect phenotypic expression dynamics derived from the newly acquired allele. **B.** Hypothetical model of phenotypic expression dynamics depending on which strand is replaced by tDNA. In case of coding strand replacement (i), RNAP uses the original allele as template to synthesize transcripts (grey dashed line), and transcription of tDNA-derived sequence occurs only once the homo-duplex forms after one replication cycle. Also, it may take more cell divisions to replace the original product (proteins, etc.) (grey balls) by tDNA-derived products (orange balls) to express visible phenotype. Meanwhile, in case of non-coding strand replacement (ii), transcription from tDNA is permitted as RNAP recognizes the tDNA sequence and synthesizes transcripts from it (orange dashed line). Consequently, tDNA-derived products (orange balls) come up right after recombination before cell division without forming the homo-duplex. After cell division, tDNA-derived products can be non-genetically distributed even in progenitor with homo-duplex of original allele not only in one with homo-duplex of tDNA-derived allele. Therefore, in this case non-

genetic inheritance is observed. This phenotypically intermediate state should be solved according to a couple of cell divisions as non-genetic inheritance will be diluted out. **C.** Representative example of two types of phenotypic expression. The model described in panel B is also suggested by actual observations. Shown is a time-lapse montage of VL1784 (*hlpA-stop-mScarlet-I*) transformed with *hlpA-mScarlet-I* tDNA where two transformed events (indicated by white and yellow arrows, respectively) are visualized side by side. Similarly, as already shown in **Figure 3**, the cell indicated by the white arrow expresses fluorescence relatively early during the time -lapse, and also gradual reduction of fluorescence is observed in progenitors in one side of the lineage. On the other hand, the cells indicated by yellow started expressing fluorescence a little bit later than cells indicated by white arrow, suggesting that this lineage might need cell division to form homo-duplex to transcribe tDNA allele. Importantly, this cell lineage did not show evidence of non-genetic phenotypic expression. Collectively, according to the model described in panel B, it seems that the cell lineage indicated by the white arrow was transformed at the noncoding strand and the cell indicated by the yellow arrow was transformed at the coding strand. Scale bar, 4 μm.

Figure 6–figure supplement 1. Expected outcome of genetic proportion after saturated transformation.

A. Expected genetic proportion after saturated single transformation. In single fragment (allele) transformation the proportion of progenitors with a tDNA-replaced allele or the original allele becomes 50:50. This can be predicted and explained by the principle mechanism of natural genetic transformation in which tDNA is processed into single stranded DNA and is integrated into one strand of the host chromosome. **B, C.** Expected genetic proportion after saturated multiple transformations in case each recombination event is independent from the other. Note that for simplification, the single circular chromosome is

depicted as separate linear fragments. As shown in this study, multiple transformation events can occur at the distinct loci at the same time in single cell. If multiple transformation events do not interfere with each other and take place independently, most cells in the population would have a higher chance to acquire a tDNA, and the proportion of untransformed cells that did not acquire any tDNA fragments should be reduced. In the case of 2 tDNAs (panel b), the expected outcome would be that 25% of the transformants obtained both new alleles, 25% just one of the two tDNAs and 25% would have the wild type genome. In the case of 3 tDNAs (panel c), the expected outcome, when every tDNA would be converted on one of the recipient strands independent from the next tDNA, would be that 12.5% of the transformants would have the wild type genome. Note that the theoretical outcomes showcased in panels a-c do not match our experimental observations (**Figure 6E** and **7C–D**) in which in general more than 50% of the cells did not acquire the SNP. However, the relative transformation efficiencies do seem to be independent. For instance, from the data of **Figure 6** (double tDNA), the observed transformation frequencies were for single *hlpA-mScarlet-I* red transformants (R) 20.2 %. The observed transformation frequency to become green (G) by the take up and integration of *spv_1159-sfGFP* was 15.2 % (P(G)=0.152. The observed transformation frequencies for both *hlpA-mScarlet-I* (R) and *spv_1159-sfGFP* (G) was 15.6 %. So, the observed number of R cells were 20.2 + 15.6 = 35.8% and the observed number of G cells were 15.2 + 15.6= 30.8%. These values are largely in line if transformation of both *mScarlet-I* (R) and *spv_1159-sfGFP* (G) are independent as the chance (P) to become R = 0.358*(1-0.308) = 24.7% (observed 20.2%). For G: 0.308*(1-0.358) = 19.7% (observed 15.2%). For both R and G: 0.358*0.308=11.0% (observed 15.6%). Finally, the proportion of untransformed cells: (1-0.308)*(1-0.358) = 44.4% (observed 49%). A similar calculation can be done for the triple tDNA transformation from **Figure 7**: *hlpA-mScarlet-I* (Red, R), *spv_1159-msfYFP* (Yellow, Y) and *ftsZ-mTurquoise2* (Cyan, C). Observed values were: P(C)

= 12.7+2.2+4.1+1.5 = 20.5 --> 0.205. P(Y) = 7.7+2.2+2.1+1.5 = 13.5 --> 0.135. P(R) = 11.0+4.1+2.1+1.5=18.7 --> 0.187. If these tDNAs are taken up independent of each other the expected percentages would be: C+Y: 0.205*0.135*(1-0.187) = 2.2% (observed: 2.2%). C+R: 0.205*0.187*(1-0.135) = 3.3% (observed: 4.1%). Y+R: 0.135*0.187*(1-0.205) = 2.0% (observed: 2.1%). C+Y+R: 0.205*0.135*0.187 = 0.5% (observed: 1.5%). Untransformed: (1-0.205)(1-0.135)(1-0.187) = 55.9% (observed: 58.7%). Together, this data suggests that in principle each transformation event is independent from each other but, since non-homologous DNA can compete for successful recombination events (**Figure 7E**), saturation of the uptake and integration system become limiting as well as off target recombination events that cannot be measured in our SNP assay. This last point is illustrated in **Figure 7E** of the main text.

Videos

Video 1. Visualization of transformation using *hlpA-mScarlet-I* reporter.

Shown is a movie of transformation with *hlpA-mScarlet-I* tDNA fragment in VL1832 (*hlpA-stop-mScarlet-I*) depicted in **Figure 3E**. Images are sfGFP (left), mScarlet-I (middle) and merged image (right). Frame interval, 5 min. Scale bar, 4 µm.

Vide 2. Visualization of population dynamics after transformation using *hlpA-mScarlet-I* reporter.

Shown is a movie of transformation with *hlpA-mScarlet-I* tDNA fragment in VL1832 (*hlpA-stop-mScarlet-I*) in lower magnification. Images are merged of phase contrast and mScarlet-I fluorescence signal. Frame interval, 5 min. Scale bar, 20 µm.

Video 3. Visualization of transformation using *spv_1159-sfGFP* reporter.

Shown is a movie of transformation with *spv_1159-sfGFP* tDNA fragment in VL1788 (*spv_1159-stop-sfGFP*) depicted in **Figure 4–figure supplement 1**. Images are phase contrast (left), sfGFP (middle), merged image (right) Frame interval, 5 min. Scale bar, 4 μm.

Video 4. Visualization of double transformation in the population.

Shown is a movie of double transformation with *hlpA-mScarlet-I* and *spv_1159-sfGFP* tDNA fragments in dual reporter strain VL1803 (*hlpA-stop-mScarlet-I*, *spv_1159-stop-sfGFP*). White arrow indicates a double transformed cell lineage. Frame interval, 5 min. Scale bar, 20 μm.

Video 5. Visualization of double transformed cells.

Shown is an enlarged movie of *hlpA-mScarlet-I* and *spv_1159-sfGFP* tDNA fragments in dual reporter strain VL1803 (*hlpA-stop-mScarlet-I*, *spv_1159-stop-sfGFP*) depicted in **Figure 6B**. Images are mScarlet-I (left), sfGFP (middle), merged image (right) Frame interval, 5 min. Scale bar, 4 μm.

Video 6. Visualization of triple transformation in population.

Shown is a movie of triple transformations with *ftsZ-mTurquoise2*, *spv_1159-msfYFP* and *hlpA-mScarlet-I* tDNA fragments in triple reporter strain VL1803 (*ftsZ-stop-mTurquoise2*, *spv_1159-stop-msfYFP*, *hlpA-stop-mScarlet-I*). Frame interval, 10 min. Scale bar, 20 μm.

Video 7. Visualization of differential phenotypic expression timing on sense strand and anti-sense strand transformation.

Shown is a movie of transformation with *hlpA-mScarlet-I* tDNA fragment in VL1832 (*hlpA-stop-mScarlet-I*) depicted in **Figure 5–figure supplement 2**. Images are an overlay between phase contrast and mScarlet-I signal. White arrow indicates a likely anti-sense strand transformed cell and the yellow arrow indicates a likely sense strand transformed cell (See text for details). Frame interval, 5 min. Scale bar, 4 μm.

References

Alloing G, Granadel C, Morrison DA, Claverys JP. 1996. Competence pheromone, oligopeptide permease, and induction of competence in Streptococcus pneumoniae. *Molecular microbiology* **21**:471–478. doi:10.1111/j.1365-2958.1996.tb02556.x

Attaiech L, Olivier A, Mortier-Barrière I, Soulet A-L, Granadel C, Martin B, Polard P, Claverys J-P. 2011. Role of the single-stranded DNA-binding protein SsbB in pneumococcal transformation: maintenance of a reservoir for genetic plasticity. *PLOS Genetics* **7**:e1002156. doi:10.1371/journal.pgen.1002156

Avery OT, Macleod CM, McCarty M. 1944. Studies On The Chemical Nature Of The Substance Inducing Transformation Of Pneumococcal Types : Induction Of Transformation By A Desoxyribonucleic Acid Fraction Isolated From Pneumococcus Type III. *The Journal of experimental medicine* **79**:137–158. doi:10.1084/jem.79.2.137

Beilharz K, Raaphorst R, Kjos M, Veening J. 2015. Red Fluorescent Proteins for Gene Expression and Protein Localization Studies in Streptococcus pneumoniae and Efficient Transformation with DNA Assembled via the Gibson Assembly Method. *Appl Environ Microb* **81**:7244–7252. doi:10.1128/aem.02033-15

Bergé M, Kamgoué A, Martin B, Polard P, Campo N, Claverys J. 2013. Midcell Recruitment of the DNA Uptake and Virulence Nuclease, EndA, for Pneumococcal Transformation. *Plos Pathog* **9**:e1003596. doi:10.1371/journal.ppat.1003596

Berge M, Mortier-Barrière I, Martin B, Claverys J-P. 2003. Transformation of Streptococcus pneumoniae relies on DprA- and RecA-dependent protection of incoming DNA single strands. *Molecular microbiology* **50**:527–536. doi:10.1046/j.1365-2958.2003.03702.x

Berge M, Moscoso M, Prudhomme M, Martin B, Claverys J-P. 2002. Uptake of transforming DNA in Gram-positive bacteria: a view from Streptococcus pneumoniae. *Molecular microbiology* **45**:411–421. doi:10.1046/j.1365-2958.2002.03013.x

Bergé MJ, Mercy C, Mortier-Barrière I, VanNieuwenhze MS, Brun YV, Grangeasse C, Polard P, Campo N. 2017. A programmed cell division delay preserves genome integrity during natural genetic transformation in Streptococcus pneumoniae. *Nature communications* **8**:1621–13. doi:10.1038/s41467-017-01716-9

Bindels DS, Haarbosch L, Weeren L van, Postma M, Wiese KE, Mastop M, Aumonier S, Gotthard G, Royant A, Hink MA, Gadella TWJ. 2017. mScarlet: a bright monomeric red fluorescent protein for cellular imaging. *Nature methods* **14**:53–56. doi:10.1038/nmeth.4074

Blokesch M. 2016. Natural competence for transformation. *Current biology : CB* **26**:R1126–R1130. doi:10.1016/j.cub.2016.08.058

1434 Boonstra M, Vesel N, Kuipers OP. 2018. Fluorescently Labeled DNA Interacts with
1435 Competence and Recombination Proteins and Is Integrated and Expressed Following
1436 Natural Transformation of Bacillus subtilis. *Mbio* **9**:e01161-18. doi:10.1128/mbio.01161-
1437 18

1438 Boudes M, Sanchez D, Graille M, Tilbeurgh H van, Durand D, Quevillon-Cheruel S. 2014.
1439 Structural insights into the dimerization of the response regulator ComE from
1440 Streptococcus pneumoniae. *Nucleic acids research* **42**:5302–5313.
1441 doi:10.1093/nar/gku110

1442 Brockhurst MA, Harrison E, Hall JPJ, Richards T, McNally A, Maclean C. 2019. The
1443 Ecology and Evolution of Pangenomes. *Current biology : CB* **29**:R1094–R1103.
1444 doi:10.1016/j.cub.2019.08.012

1445 Bryskier A. 2002. Viridans group streptococci: a reservoir of resistant bacteria in oral cavities.
1446 *Clinical microbiology and infection : the official publication of the European Society of*
1447 *Clinical Microbiology and Infectious Diseases* **8**:65–69. doi:10.1046/j.1198-
1448 743x.2001.00398.x

1449 Campbell EA, Choi SY, Masure HR. 1998. A competence regulon in Streptococcus
1450 pneumoniae revealed by genomic analysis. *Molecular microbiology* **27**:929–939.
1451 doi:10.1046/j.1365-2958.1998.00737.x

1452 Carrasco B, Serrano E, Sánchez H, Wyman C, Alonso JC. 2016. Chromosomal
1453 transformation in Bacillus subtilis is a non-polar recombination reaction. *Nucleic Acids*
1454 *Res* **44**:2754–68. doi:10.1093/nar/gkv1546

1455 Chandler MS, Morrison DA. 1988. Identification of two proteins encoded by com, a
1456 competence control locus of Streptococcus pneumoniae. *Journal of bacteriology*
1457 **170**:3136–3141. doi:10.1128/jb.170.7.3136-3141.1988

1458 Chastanet A, Prudhomme M, Claverys J-P, Msadek T. 2001. Regulation of Streptococcus
1459 pneumoniae clp Genes and Their Role in Competence Development and Stress Survival. *J*
1460 *Bacteriol* **183**:7295–7307. doi:10.1128/jb.183.24.7295-7307.2001

1461 Chen I, Dubnau D. 2004. DNA uptake during bacterial transformation. *Nature Reviews*
1462 *Microbiology* **2**:241–249. doi:10.1038/nrmicro844

1463 Chewapreecha C, Harris SR, Croucher RMCWPTSDBCFNJ, Turner C, Marttinen P, Cheng
1464 L, Pessia A, Aanensen DM, Mather AE, Page AJ, Salter SJ, Harris D, Nosten F, Goldblatt
1465 D, Corander J, Parkhill J, Turner P, Bentley SD. 2014. Dense genomic sampling identifies
1466 highways of pneumococcal recombination. *Nature genetics* **46**:305–309.
1467 doi:10.1038/ng.2895

1468 Claverys JP, Lacks SA. 1986. Heteroduplex deoxyribonucleic acid base mismatch repair in
1469 bacteria. *Microbiological reviews Baltimore* **50**:133–165.

1470 Claverys J-P, Prudhomme M, Martin B. 2006. Induction of competence regulons as a general
1471 response to stress in gram-positive bacteria. *Annual review of microbiology* **60**:451–475.
1472 doi:10.1146/annurev.micro.60.080805.142139

1473 Corbinais C, Mathieu A, Kortulewski T, Radicella JP, Marsin S. 2016. Following
1474 transforming DNA in Helicobacter pylori from uptake to expression: Visualising H. pylori
1475 transformation. *Mol Microbiol* **101**:1039–1053. doi:10.1111/mmi.13440

1476 Cowley LA, Petersen FC, Junges R, Jimenez MJD, Morrison DA, Hanage WP. 2018.
1477 Evolution via recombination: Cell-to-cell contact facilitates larger recombination events in
1478 Streptococcus pneumoniae. *PLOS Genetics* **14**:e1007410.
1479 doi:10.1371/journal.pgen.1007410

1480 Croucher N, Løchen A, Bentley S. 2018. Pneumococcal Vaccines: Host Interactions,
1481 Population Dynamics, and Design Principles. *Annu Rev Microbiol* **72**:521–549.
1482 doi:10.1146/annurev-micro-090817-062338

1483 Croucher RMCWPTSDBCFNJ, Mostowy R, Wymant C, Turner P, Bentley SD, Fraser C.
1484 2016. Horizontal DNA Transfer Mechanisms of Bacteria as Weapons of Intragenomic
1485 Conflict. *PLOS Biology* **14**:e1002394-42. doi:10.1371/journal.pbio.1002394

1486 Dagkessamanskaia A, Moscoso M, Hénard V, Guiral S, Overweg K, Reuter M, Martin B,
1487 Wells J, Claverys J-P. 2004. Interconnection of competence, stress and CiaR regulons in
1488 Streptococcus pneumoniae: competence triggers stationary phase autolysis of ciaR mutant
1489 cells. *Molecular microbiology* **51**:1071–1086. doi:10.1111/j.1365-2958.2003.03892.x

1490 Dalia AB, Dalia TN. 2019. Spatiotemporal Analysis of DNA Integration during Natural
1491 Transformation Reveals a Mode of Nongenetic Inheritance in Bacteria. *Cell* **179**:1499-
1492 1511.e10. doi:10.1016/j.cell.2019.11.021

1493 Dalia AB, McDonough E, Camilli A. 2014. Multiplex genome editing by natural
1494 transformation. *P Natl Acad Sci Usa* **111**:8937–42. doi:10.1073/pnas.1406478111

1495 Davidoff-Abelson R, Dubnau D. 1971. Fate of transforming DNA after uptake by competent
1496 Bacillus subtilis: failure of donor DNA to replicate in a recombination-deficient recipient.
1497 *Proceedings of the National Academy of Sciences of the United States of America*
1498 **68**:1070–1074. doi:10.1073/pnas.68.5.1070

1499 Domenech A, Brochado AR, Sender V, Hentrich K, Henriques-Normark B, Typas A,
1500 Veening J-W. 2020. Proton Motive Force Disruptors Block Bacterial Competence and
1501 Horizontal Gene Transfer. *Cell Host Microbe* **27**:544-555.e3.
1502 doi:10.1016/j.chom.2020.02.002

1503 Domenech A, Slager J, Veening J. 2018. Antibiotic-Induced Cell Chaining Triggers
1504 Pneumococcal Competence by Reshaping Quorum Sensing to Autocrine-Like Signaling.
1505 *Cell Reports* **25**:2390-2400.e3. doi:10.1016/j.celrep.2018.11.007

1506 Dubnau D, Blokesch M. 2019. Mechanisms of DNA Uptake by Naturally Competent
1507 Bacteria. *Annual review of genetics* **53**:217–237. doi:10.1146/annurev-genet-112618-
1508 043641

1509 Dutreix M, Rao BJ, Radding CM. 1991. The effects on strand exchange of 5′ versus 3′ ends
1510 of single-stranded DNA in RecA nucleoprotein filaments. *J Mol Biol* **219**:645–654.
1511 doi:10.1016/0022-2836(91)90661-o

Ephrussi-Taylor H. 1966. Genetic recombination in DNA-induced transformation of Pneumococcus. IV. The pattern of transmission and phenotypic expression of high and low-efficiency donor sites in the amiA locus. *Genetics* **54**:211–222.

Ephrussi-Taylor H. 1962. Appearance of Streptomycin Resistance Following the Uptake of Transforming Deoxyribonucleic Acid in Pneumococcus. *Nature* **196**:748-752 (1962).

Ephrussi-Taylor H. 1958. The mechanism of deoxyribonucleic acid-induced transformations. *Recent Progress in Microbiology* 5148.

Ephrussi-Taylor H, Gray TC. 1966. Genetic studies of recombining DNA in pneumococcal transformation. *The Journal of general physiology* **49**:211–231. doi:10.1085/jgp.49.6.211

Fenoll A, Ardanuy C, Liñares J, Cercenado E, Marco F, Fleites A, Rodríguez-Mayo M, López-Hontangas J-L, Palop B, Aller A-I, Buendía B, Méndez C, Cifuentes I, Group OS. 2018. Serotypes and genotypes of S. pneumoniae isolates from adult invasive disease in Spain: A 5-year prospective surveillance after pediatric PCV13 licensure. The ODIN study. *Vaccine* **36**:7993–8000. doi:10.1016/j.vaccine.2018.10.098

Fox MS, Allen MK. 1964. On The Mechanism Of Deoxyribonucleate Integration In Pneumococcal Transformation. *Proceedings of the National Academy of Sciences of the United States of America* **52**:412–419. doi:10.1073/pnas.52.2.412

Gabor M, Hotchkiss RD. 1966. Manifestation of linear organization in molecules of pneumococcal transforming DNA. *Proceedings of the National Academy of Sciences of the United States of America* **56**:1441–1448. doi:10.1073/pnas.56.5.1441

Gibson DG. 2011. Enzymatic assembly of overlapping DNA fragments. *Methods Enzymol* **498**:349–61. doi:10.1016/b978-0-12-385120-8.00015-2

Godeux A-S, Lupo A, Haenni M, Guette-Marquet S, Wilharm G, Laaberki M-H, Charpentier X. 2018. Fluorescence-Based Detection of Natural Transformation in Drug-Resistant Acinetobacter baumannii. *Journal of bacteriology* **200**:939. doi:10.1128/jb.00181-18

Gómez-Mejia A, Gámez G, Hammerschmidt S. 2018. Streptococcus pneumoniae two-component regulatory systems: The interplay of the pneumococcus with its environment. *International journal of medical microbiology : IJMM* **308**:722–737. doi:10.1016/j.ijmm.2017.11.012

Gould SJ, Lewontin RC. 1979. The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. *Proc Royal Soc Lond Ser B Biological Sci* **205**:581–598. doi:10.1098/rspb.1979.0086

Griffith F. 1928. The Significance of Pneumococcal Types. *The Journal of hygiene* **27**:113–159. doi:10.1017/s0022172400031879

Hakenbeck R, Brückner R, Denapaite D, Maurer P. 2012. Molecular mechanisms of β-lactam resistance in Streptococcus pneumoniae. *Future microbiology* **7**:395–410. doi:10.2217/fmb.12.2

1549 Hanahan D, Jessee J, Bloom FR. 1991. Plasmid transformation of Escherichia coli and other
1550 bacteria. *Methods in enzymology* **204**:63–113. doi:10.1016/0076-6879(91)04006-a

1551 Håvarstein LS, Coomaraswamy G, Morrison DA. 1995. An unmodified heptadecapeptide
1552 pheromone induces competence for genetic transformation in Streptococcus pneumoniae.
1553 *Proceedings of the National Academy of Sciences of the United States of America*
1554 **92**:11140–11144. doi:10.1073/pnas.92.24.11140

1555 Håvarstein LS, Gaustad P, Nes IF, Morrison DA. 1996. Identification of the streptococcal
1556 competence-pheromone receptor. *Molecular microbiology* **21**:863–869.
1557 doi:10.1046/j.1365-2958.1996.521416.x

1558 Hiller NL, Sá-Leão R. 2018. Puzzling Over the Pneumococcal Pangenome. *Frontiers in*
1559 *microbiology* **9**:2580. doi:10.3389/fmicb.2018.02580

1560 Hui FM, Zhou L, Morrison DA. 1995. Competence for genetic transformation in
1561 Streptococcus pneumoniae: organization of a regulatory locus with homology to two
1562 lactococcin A secretion genes. *Gene* **153**:25–31. doi:10.1016/0378-1119(94)00841-f

1563 Humbert O, Prudhomme M, Hakenbeck R, Dowson CG, Claverys JP. 1995. Homeologous
1564 recombination and mismatch repair during transformation in Streptococcus pneumoniae:
1565 saturation of the Hex mismatch repair system. *Proceedings of the National Academy of*
1566 *Sciences of the United States of America* **92**:9052–9056. doi:10.1073/pnas.92.20.9052

1567 Janoir C, Podglajen I, Kitzis MD, Poyart C, Gutmann L. 1999. In vitro exchange of
1568 fluoroquinolone resistance determinants between Streptococcus pneumoniae and viridans
1569 streptococci and genomic organization of the parE-parC region in S. mitis. *Journal of*
1570 *Infectious Diseases* **180**:555–558. doi:10.1086/314888

1571 Johnston C, Martin B, Fichant G, Polard P, Claverys J-P. 2014. Bacterial transformation:
1572 distribution, shared mechanisms and divergent control. *Nature Reviews Microbiology*
1573 **12**:181–196. doi:10.1038/nrmicro3199

1574 Johnston C, Soulet A-L, Berge M, Prudhomme M, Lemos DD, Polard P. 2020. The
1575 alternative sigma factor σX mediates competence shut-off at the cell pole in Streptococcus
1576 pneumoniae. *Biorxiv* 2020.06.12.147637. doi:10.1101/2020.06.12.147637

1577 Jong IG de, Beilharz K, Kuipers OP, Veening J-W. 2011. Live Cell Imaging of Bacillus
1578 subtilis and Streptococcus pneumoniae using Automated Time-lapse Microscopy. *Journal*
1579 *of visualized experiments : JoVE*. doi:10.3791/3145

1580 Keller L, Rueff A, Kurushima J, Veening J. 2019. Three New Integration Vectors and
1581 Fluorescent Proteins for Use in the Opportunistic Human Pathogen Streptococcus
1582 pneumoniae. *Genes-basel* **10**:394. doi:10.3390/genes10050394

1583 Konforti BB, Davis RW. 1990. The preference for a 3' homologous end is intrinsic to RecA-
1584 promoted strand exchange. *J Biological Chem* **265**:6916–20.

1585 LACKS S. 1962. Molecular fate of DNA in genetic transformation of Pneumococcus.
1586 *Journal of Molecular Biology* **5**:119–131. doi:10.1016/s0022-2836(62)80067-9

1587 Lam T, Maienschein-Cline M, Eddington DT, Morrison DA. 2020. Multiplex gene transfer
1588 by genetic transformation between isolated S. pneumoniae cells confined in microfluidic
1589 droplets. *Integr Biology Quantitative Biosci Nano Macro*. doi:10.1093/intbio/zyz036

1590 Laurenceau R, Péhau-Arnaudet G, Baconnais S, Gault J, Malosse C, Dujeancourt A, Campo
1591 N, Chamot-Rooke J, Cam EL, Claverys J-P, Fronzes R. 2013. A type IV pilus mediates
1592 DNA binding during natural transformation in Streptococcus pneumoniae. *PLoS*
1593 *Pathogens* **9**:e1003473. doi:10.1371/journal.ppat.1003473

1594 Lee MS, Seok C, Morrison DA. 1998. Insertion-duplication mutagenesis in Streptococcus
1595 pneumoniae: targeting fragment length is a critical parameter in use as a random insertion
1596 tool. *Applied and Environmental Microbiology* **64**:4796–4802.

1597 Lefevre JC, Claverys JP, Sicard AM. 1979. Donor deoxyribonucleic acid length and marker
1598 effect in pneumococcal transformation. *J Bacteriol* **138**:80–6.

1599 Levy C, Ouldali N, Caeymaex L, Angoulvant F, Varon E, Cohen R. 2019. Diversity of
1600 Serotype Replacement After Pneumococcal Conjugate Vaccine Implementation in Europe.
1601 *The Journal of pediatrics* **213**:252-253.e3. doi:10.1016/j.jpeds.2019.07.057

1602 Lin M, Kussell E. 2017. Correlated Mutations and Homologous Recombination Within
1603 Bacterial Populations. *Genetics* **205**:891–917. doi:10.1534/genetics.116.189621

1604 Litt M, Marmur J, Ephrussi-Taylor H, Doty P. 1958. The Dependence of Peumococcal
1605 Transformation on The Molecular Weight of Deoxyribose Nucleic Acid. *Proceedings of*
1606 *the National Academy of Sciences of the United States of America* **44**:144–152.
1607 doi:10.1073/pnas.44.2.144

1608 Liu X, Gallay C, Kjos M, Domenech A, Slager J, Kessel SP van, Knoops K, Sorg RA, Zhang
1609 J-R, Veening J-W. 2017. High-throughput CRISPRi phenotyping identifies new essential
1610 genes in Streptococcus pneumoniae. *Molecular systems biology* **13**:931.
1611 doi:10.15252/msb.20167449

1612 Liu Y, Zeng Y, Huang Y, Gu L, Wang S, Li C, Morrison DA, Deng H, Zhang J-R. 2019.
1613 HtrA-mediated selective degradation of DNA uptake apparatus accelerates termination of
1614 pneumococcal transformation. *Molecular microbiology* **112**:1308–1325.
1615 doi:10.1111/mmi.14364

1616 Lo SW, Gladstone RA, Tonder AJ van, Plessis MD, Cornick JE, Hawkins PA, Madhi SA,
1617 Nzenze SA, Kandasamy R, Ravikumar KL, Elmdaghri N, Kwambana-Adams B, Almeida
1618 SCG, Skoczynska A, Egorova E, Titov L, Saha SK, Paragi M, Everett DB, Antonio M,
1619 Klugman KP, Li Y, Metcalf BJ, Beall B, McGee L, Breiman RF, Bentley SD, Gottberg A
1620 von, Global Pneumococcal Sequencing Consortium. 2019. A mosaic tetracycline
1621 resistance gene tet(S/M) detected in an MDR pneumococcal CC230 lineage that
1622 underwent capsular switching in South Africa. *The Journal of antimicrobial*
1623 *chemotherapy* **19**:759. doi:10.1093/jac/dkz477

1624 Luo P, Li H, Morrison DA. 2003. ComX is a unique link between multiple quorum sensing
1625 outputs and competence in Streptococcus pneumoniae. *Molecular microbiology* **50**:623–
1626 633. doi:10.1046/j.1365-2958.2003.03714.x

Maamar H, Dubnau D. 2005. Bistability in the Bacillus subtilis K-state (competence) system requires a positive feedback loop. *Molecular microbiology* **56**:615–624. doi:10.1111/j.1365-2958.2005.04592.x

Marie L, Rapisarda C, Morales V, Berge M, Perry T, Soulet A-L, Gruget C, Remaut H, Fronzes R, Polard P. 2017. Bacterial RadA is a DnaB-type helicase interacting with RecA to promote bidirectional D-loop extension. *Nature communications* **8**:15638–14. doi:10.1038/ncomms15638

Martin B, Granadel C, Campo N, Hénard V, Prudhomme M, Claverys J-P. 2010. Expression and maintenance of ComD-ComE, the two-component signal-transduction system that controls competence of Streptococcus pneumoniae. *Molecular microbiology* **75**:1513–1528. doi:10.1111/j.1365-2958.2010.07071.x

Martin B, Soulet A-L, Mirouze N, Prudhomme M, Mortier-Barrière I, Granadel C, Noirot-Gros M-F, Noirot P, Polard P, Claverys J-P. 2013. ComE/ComE~P interplay dictates activation or extinction status of pneumococcal X-state (competence). *Molecular microbiology* **87**:394–411. doi:10.1111/mmi.12104

Méjean V, Claverys JP. 1984. Use of a cloned DNA fragment to analyze the fate of donor DNA in transformation of Streptococcus pneumoniae. *Journal of bacteriology* **158**:1175–1178.

Mell JC, Redfield RJ. 2014. Natural competence and the evolution of DNA uptake specificity. *Journal of bacteriology* **196**:1471–1483. doi:10.1128/jb.01293-13

Mirouze N, Bergé MA, Soulet A-L, Mortier-Barrière I, Quentin Y, Fichant G, Granadel C, Noirot-Gros M-F, Noirot P, Polard P, Martin B, Claverys J-P. 2013. Direct involvement of DprA, the transformation-dedicated RecA loader, in the shut-off of pneumococcal competence. *Proceedings of the National Academy of Sciences of the United States of America* **110**:E1035-44. doi:10.1073/pnas.1219868110

Moreno-Gámez S, Sorg RA, Domenech A, Kjos M, Weissing FJ, Doorn GS van, Veening J-W. 2017. Quorum sensing integrates environmental cues, cell density and cell history to control bacterial competence. *Nature communications* **8**:854. doi:10.1038/s41467-017-00903-y

Mortier-Barrière I, Saizieu A de, Claverys JP, Martin B. 1998. Competence-specific induction of recA is required for full recombination proficiency during transformation in Streptococcus pneumoniae. *Molecular microbiology* **27**:159–170. doi:10.1046/j.1365-2958.1998.00668.x

Mortier-Barrière I, Velten M, Dupaigne P, Mirouze N, Piétrement O, McGovern S, Fichant G, Martin B, Noirot P, Cam EL, Polard P, Claverys J-P. 2007. A key presynaptic role in transformation for a widespread bacterial protein: DprA conveys incoming ssDNA to RecA. *Cell* **130**:824–836. doi:10.1016/j.cell.2007.07.038

O'Brien KL, Wolfson LJ, Watt JP, Henkle E, Deloria-Knoll M, McCall N, Lee E, Mulholland K, Levine OS, Cherian T, Team H and PGB of DS. 2009. Burden of disease

caused by Streptococcus pneumoniae in children younger than 5 years: global estimates. *Lancet (London, England)* **374**:893–902. doi:10.1016/s0140-6736(09)61204-6

Ouldali N, Levy C, Varon E, Bonacorsi S, Béchet S, Cohen R, Angoulvant F, Network FPM. 2018. Incidence of paediatric pneumococcal meningitis and emergence of new serotypes: a time-series analysis of a 16-year French national survey. *The Lancet Infectious diseases* **18**:983–991. doi:10.1016/s1473-3099(18)30349-9

Paintdakhi A, Parry B, Campos M, Irnov I, Elf J, Surovtsev I, Jacobs-Wagner C. 2016. Oufti: an integrated software package for high-accuracy, high-throughput quantitative microscopy analysis. *Molecular microbiology* **99**:767–777. doi:10.1111/mmi.13264

Pestova EV, Håvarstein LS, Morrison DA. 1996. Regulation of competence for genetic transformation in Streptococcus pneumoniae by an auto-induced peptide pheromone and a two-component regulatory system. *Molecular microbiology* **21**:853–862. doi:10.1046/j.1365-2958.1996.501417.x

Pestova EV, Morrison DA. 1998. Isolation and characterization of three Streptococcus pneumoniae transformation-specific loci by use of a lacZ reporter insertion vector. *Journal of bacteriology* **180**:2701–2710.

Petit MA, Dimpfl J, Radman M, Echols H. 1991. Control of large chromosomal duplications in Escherichia coli by the mismatch repair system. *Genetics* **129**:327–332.

Piechowska M, Fox MS. 1971. Fate of transforming deoxyribonucleate in Bacillus subtilis. *Journal of bacteriology* **108**:680–689.

Porter RD, Guild WR. 1969. Number of transformable units per cell in Diplococcus pneumoniae. *Journal of bacteriology* **97**:1033–1035.

Prina E, Ranzani OT, Torres A. 2015. Community-acquired pneumonia. *Lancet (London, England)* **386**:1097–1108. doi:10.1016/s0140-6736(15)60733-4

Puyet A, Greenberg B, Lacks SA. 1990. Genetic and structural characterization of endA. A membrane-bound nuclease required for transformation of Streptococcus pneumoniae. *Journal of Molecular Biology* **213**:727–738. doi:10.1016/s0022-2836(05)80259-1

Raaphorst R, Kjos M, Veening J. 2017. Chromosome segregation drives division site selection in Streptococcus pneumoniae. *P Natl Acad Sci Usa* **114**:E5959–E5968. doi:10.1073/pnas.1620608114

Raaphorst R van, Kjos M, Veening J-W. 2020. BactMAP: An R package for integrating, analyzing and visualizing bacterial microscopy data. *Molecular microbiology* **11**:2699. doi:10.1111/mmi.14417

Salvadori G, Junges R, Morrison DA, Petersen FC. 2019. Competence in Streptococcus pneumoniae and Close Commensal Relatives: Mechanisms and Implications. *Frontiers in cellular and infection microbiology* **9**:94. doi:10.3389/fcimb.2019.00094

1702 Sanchez D, Boudes M, Tilbeurgh H van, Durand D, Quevillon-Cheruel S. 2015. Modeling
1703 the ComD/ComE/comcde interaction network using small angle X-ray scattering. *The*
1704 *FEBS journal* **282**:1538–1553. doi:10.1111/febs.13240

1705 Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S,
1706 Rueden C, Saalfeld S, Schmid B, Tinevez J-Y, White DJ, Hartenstein V, Eliceiri K,
1707 Tomancak P, Cardona A. 2012. Fiji: an open-source platform for biological-image
1708 analysis. *Nature methods* **9**:676–682. doi:10.1038/nmeth.2019

1709 Shanker E, Federle MJ. 2017. Quorum Sensing Regulation of Competence and Bacteriocins
1710 in Streptococcus pneumoniae and mutans. *Genes* **8**:15. doi:10.3390/genes8010015

1711 Slager J, Aprianto R, Veening J-W. 2019. Refining the pneumococcal competence regulon by
1712 RNA-sequencing. *Journal of bacteriology* JB.00780-18. doi:10.1128/jb.00780-18

1713 Slager J, Aprianto R, Veening J-W. 2018. Deep genome annotation of the opportunistic
1714 human pathogen Streptococcus pneumoniae D39. *Nucleic acids research* **46**:9971–9989.
1715 doi:10.1093/nar/gky725

1716 Slager J, Kjos M, Attaiech L, Veening J-W. 2014. Antibiotic-induced replication stress
1717 triggers bacterial competence by increasing gene dosage near the origin. *Cell* **157**:395–406.
1718 doi:10.1016/j.cell.2014.01.068

1719 Smits WK, Eschevins CC, Susanna KA, Bron S, Kuipers OP, Hamoen LW. 2005. Stripping
1720 Bacillus: ComK auto-stimulation is responsible for the bistable response in competence
1721 development. *Molecular microbiology* **56**:604–614. doi:10.1111/j.1365-
1722 2958.2005.04488.x

1723 Sorg RA, Kuipers OP, Veening J-W. 2015. Gene expression platform for synthetic biology in
1724 the human pathogen Streptococcus pneumoniae. *ACS synthetic biology* **4**:228–239.
1725 doi:10.1021/sb500229s

1726 Straume D, Stamsås GA, Håvarstein LS. 2015. Natural transformation and genome evolution
1727 in Streptococcus pneumoniae. *Infection, genetics and evolution : journal of molecular*
1728 *epidemiology and evolutionary genetics in infectious diseases* **33**:371–380.
1729 doi:10.1016/j.meegid.2014.10.020

1730 Stylianidou S, Brennan C, Nissen SB, Kuwada NJ, Wiggins PA. 2016. SuperSegger: robust
1731 image segmentation, analysis and lineage tracking of bacterial cells. *Molecular*
1732 *microbiology* **102**:690–700. doi:10.1111/mmi.13486

1733 Tomasz A. 1966. Model for the mechanism controlling the expression of competent state in
1734 Pneumococcus cultures. *Journal of bacteriology* **91**:1050–1061.

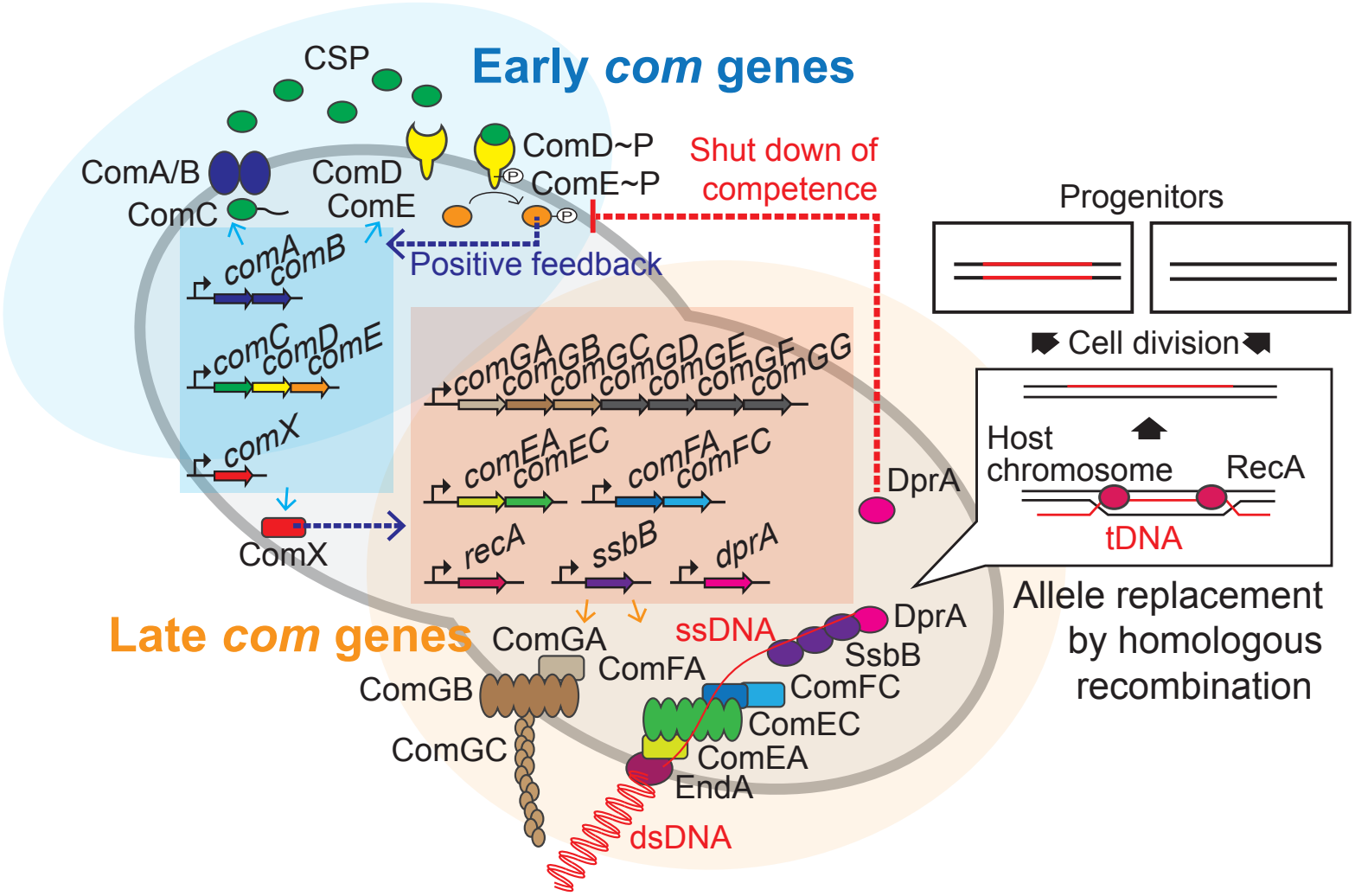
1735 Uptain SM, Chamberlin MJ. 1997. Escherichia coli RNA polymerase terminates transcription
1736 efficiently at rho-independent terminators on single-stranded DNA templates. *Proc*
1737 *National Acad Sci* **94**:13548–13553. doi:10.1073/pnas.94.25.13548

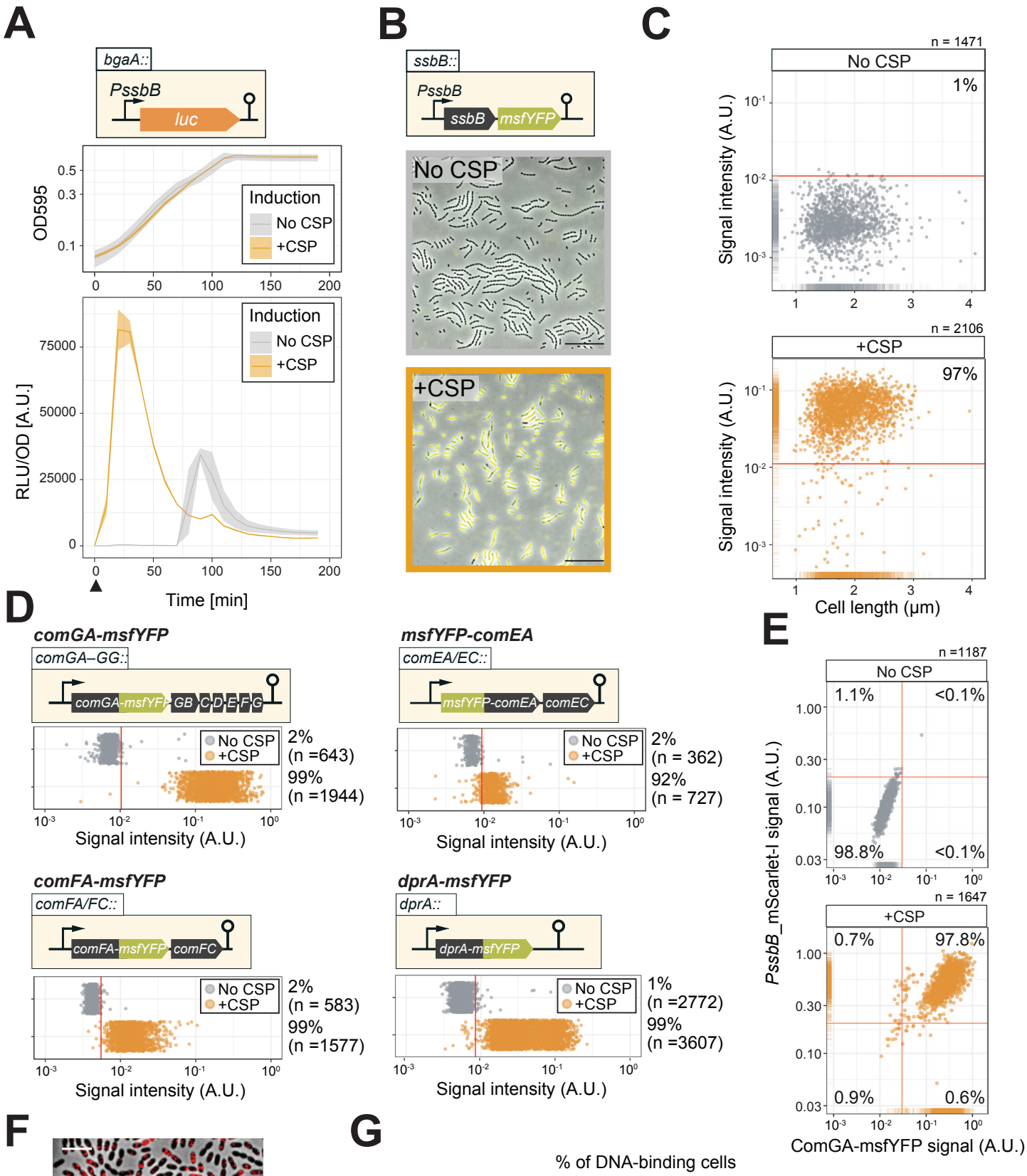
1738 Veening J-W, Blokesch M. 2017. Interbacterial predation as a strategy for DNA acquisition
1739 in naturally competent bacteria. *Nature Reviews Microbiology* **15**:629–629.
1740 doi:10.1038/nrmicro.2017.89

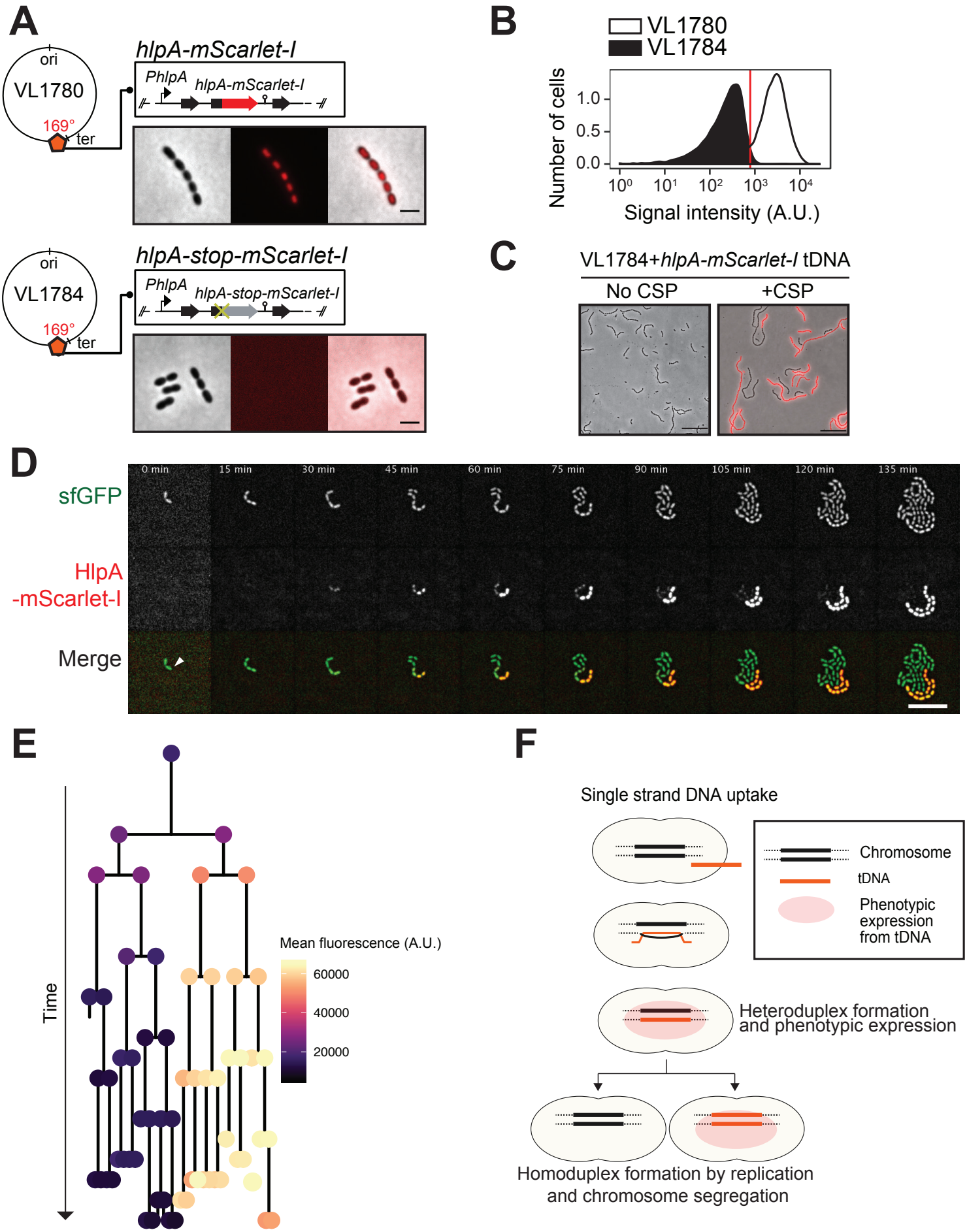
1741 Ween O, Gaustad P, Håvarstein LS. 1999. Identification of DNA binding sites for ComE, a
1742 key regulator of natural competence in Streptococcus pneumoniae. *Molecular*
1743 *microbiology* **33**:817–827. doi:10.1046/j.1365-2958.1999.01528.x

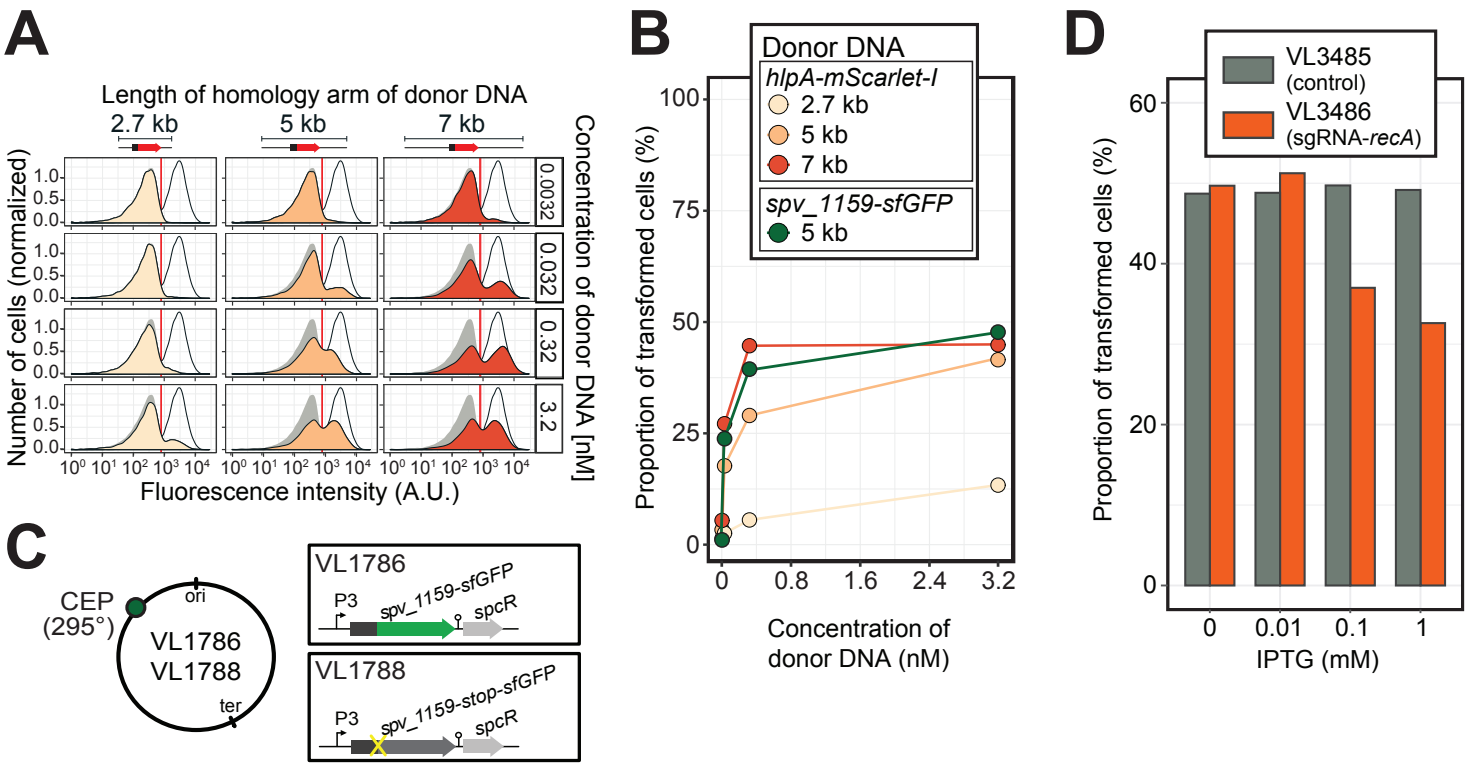
1744 Weng L, Piotrowski A, Morrison DA. 2013. Exit from competence for genetic transformation
1745 in Streptococcus pneumoniae is regulated at multiple levels. *PloS one* **8**:e64197.
1746 doi:10.1371/journal.pone.0064197

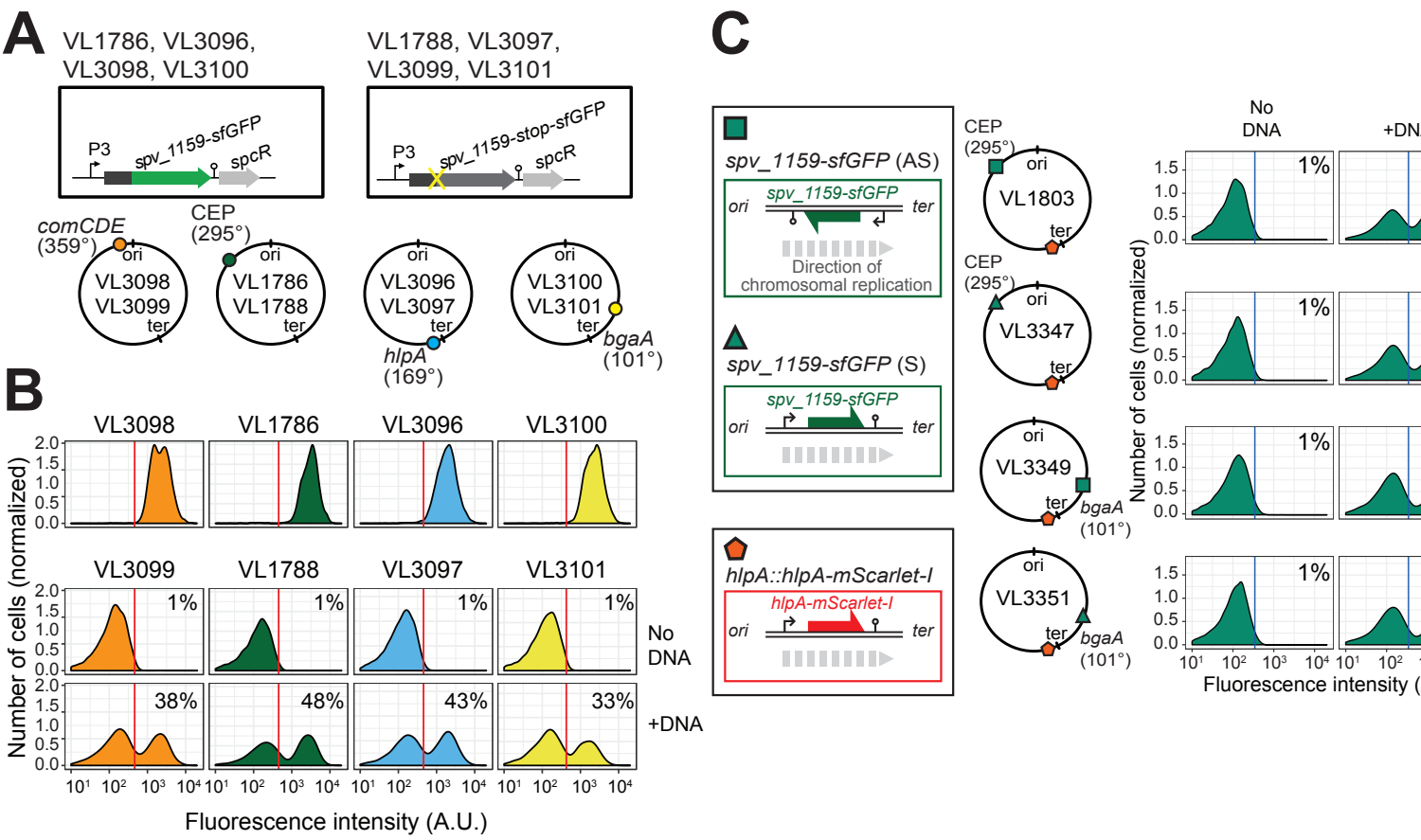
1747

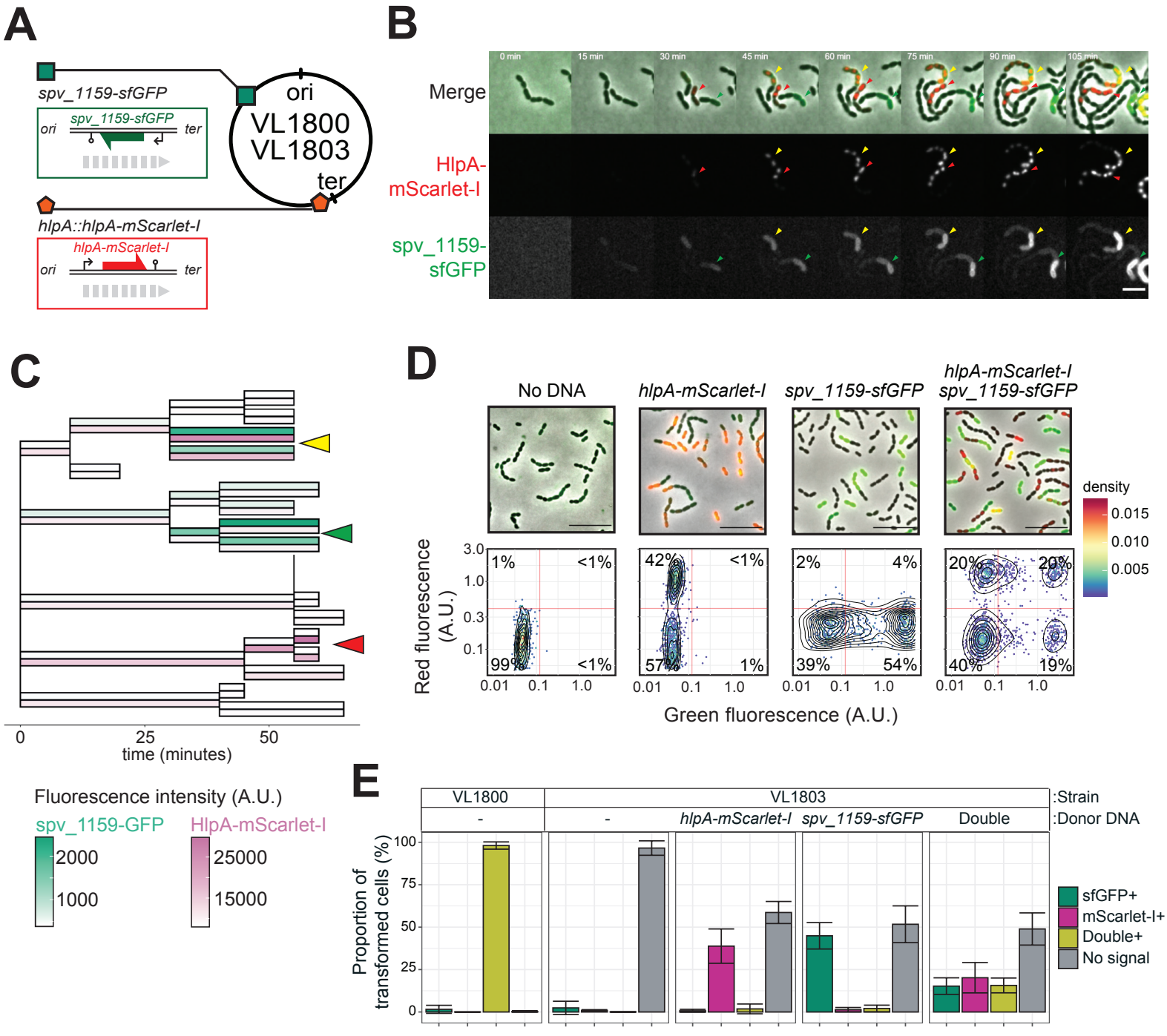


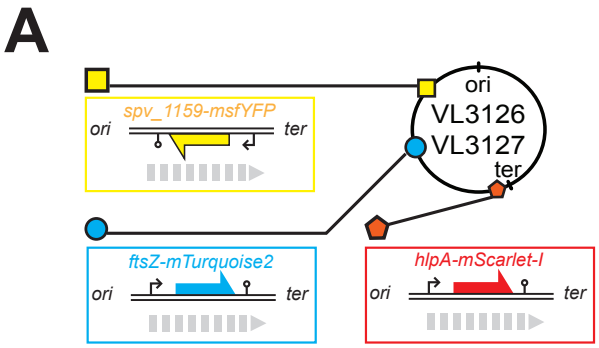


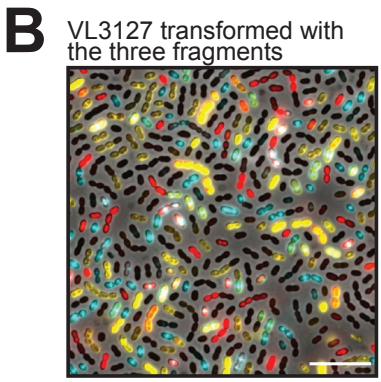


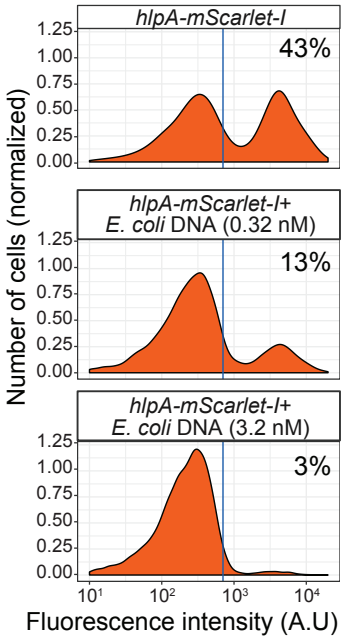








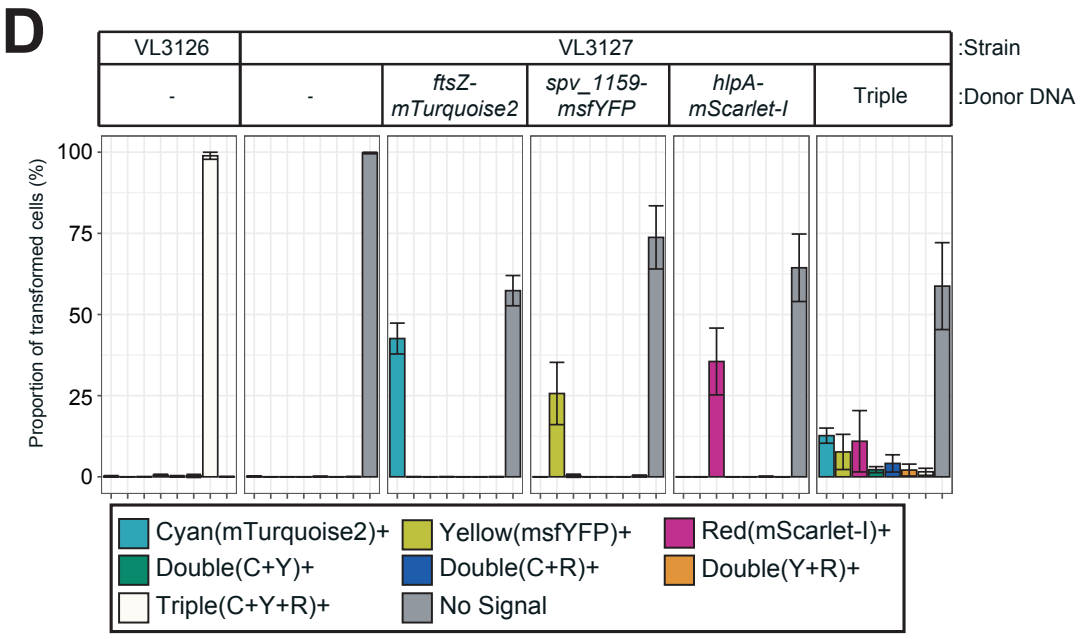


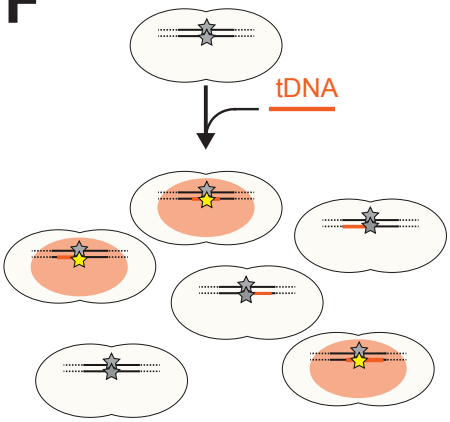


C

Proportion of transformed cells (%)

Strain		Single: Cyan	Single: Yellow	Single: Red	Double: Cyan +Yellow	Double: Cyan +Red	Double: Yellow +Red	Triple	No Signal
VL3126	No	0.1 ± 0.2	0.0 ± 0.0	0.0 ± 0.0	0.5 ± 0.3	0.1 ± 0.2	0.3 ± 0.5	98.9 ± 1.1	0.0 ± 0.1
VL3127	No	0.1 ± 0.1	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.1 ± 0.1	0.0 ± 0.0	0.0 ± 0.1	99.7 ± 0.2
	ftsZ-mTq	42.6 ± 4.8	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	57.3 ± 4.7
	1159-msfYFP	0.0 ± 0.0	25.7 ± 9.6	0.3 ± 0.5	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.3 ± 0.3	73.8 ± 9.7
	hlpA-mSc	0.0 ± 0.0	0.0 ± 0.0	35.5 ± 10.3	0.0 ± 0.0	0.0 ± 0.0	0.1 ± 0.1	0.0 ± 0.0	64.4 ± 10.4
	triple	12.7 ± 2.3	7.7 ± 5.4	11.0 ± 9.4	2.2 ± 0.9	4.1 ± 2.7	2.1 ± 1.8	1.5 ± 1.1	58.7 ± 13.4





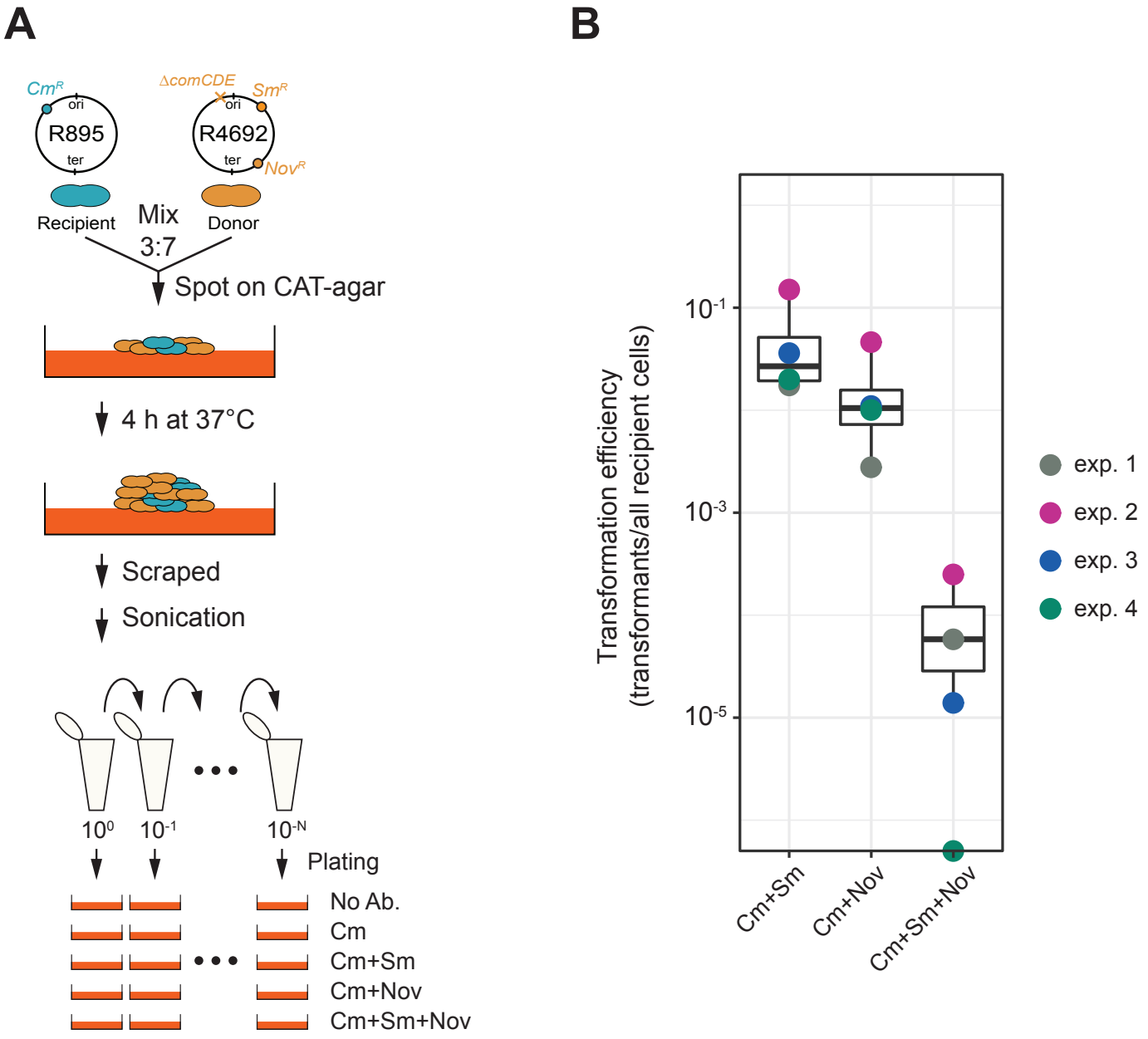


Figure 2–figure supplement 1

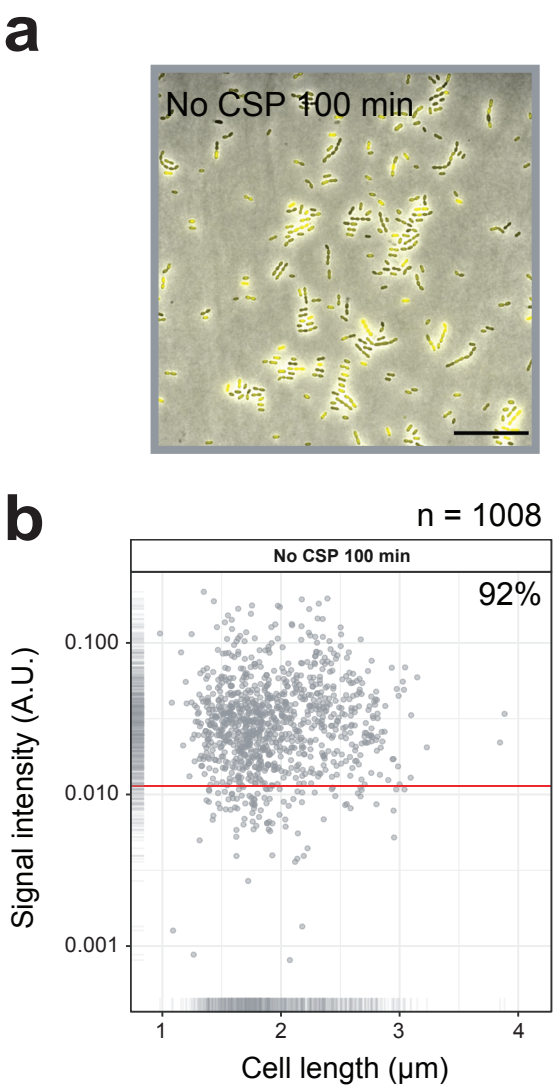
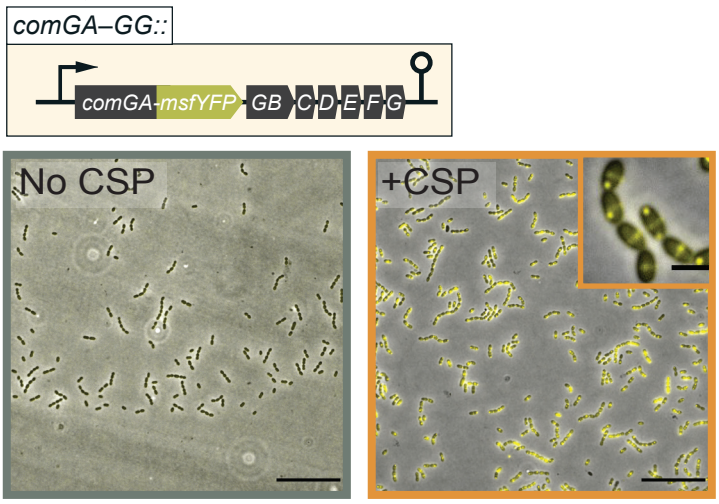
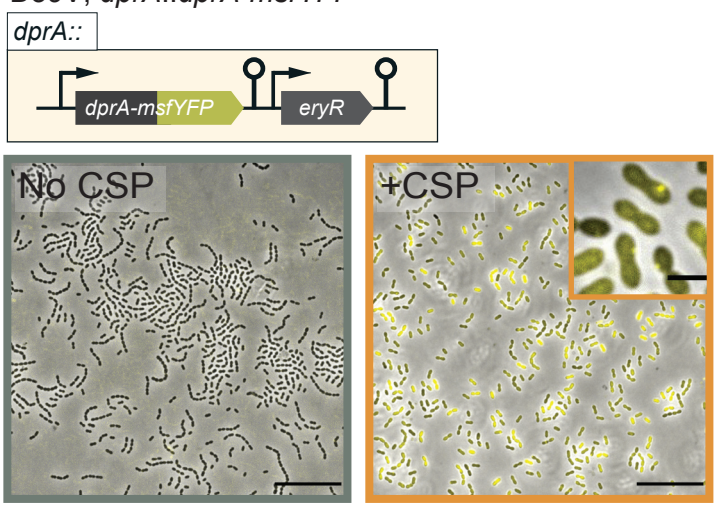


Figure 2–figure supplement 2

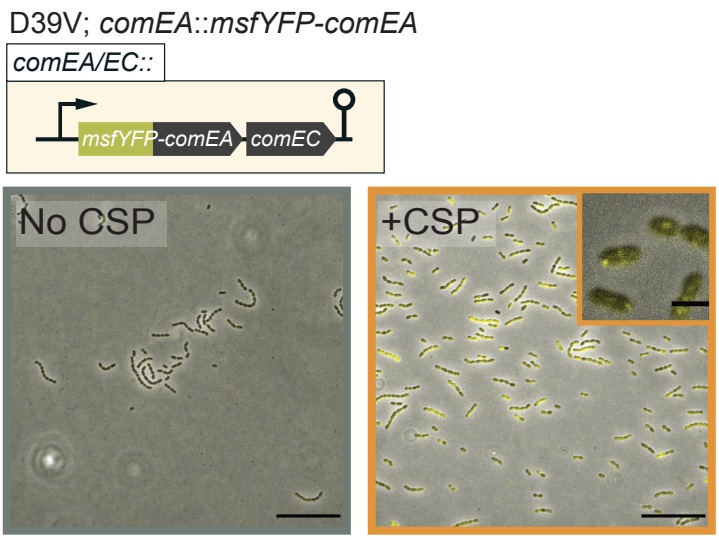
D39V; *comGA*::*comGA-msfYFP*



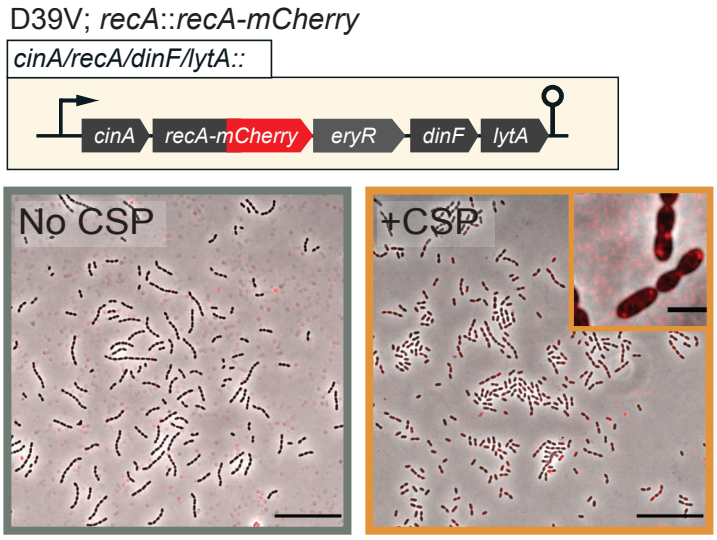
D39V; *dprA*::*dprA-msfYFP*



D39V; *comEA*::*msfYFP-comEA*



D39V; *recA*::*recA-mCherry*



D39V; *comFA*::*comFA-msfYFP*

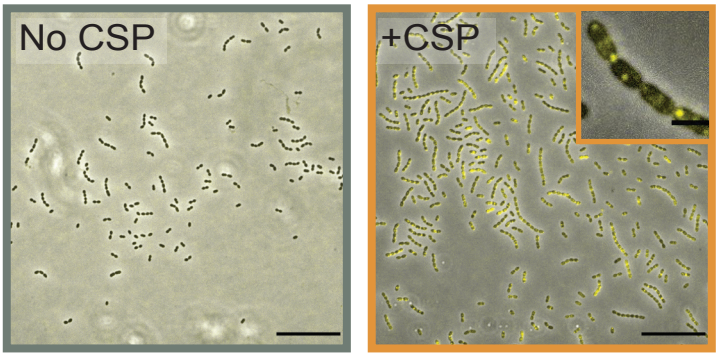


Figure 3–figure supplement 1

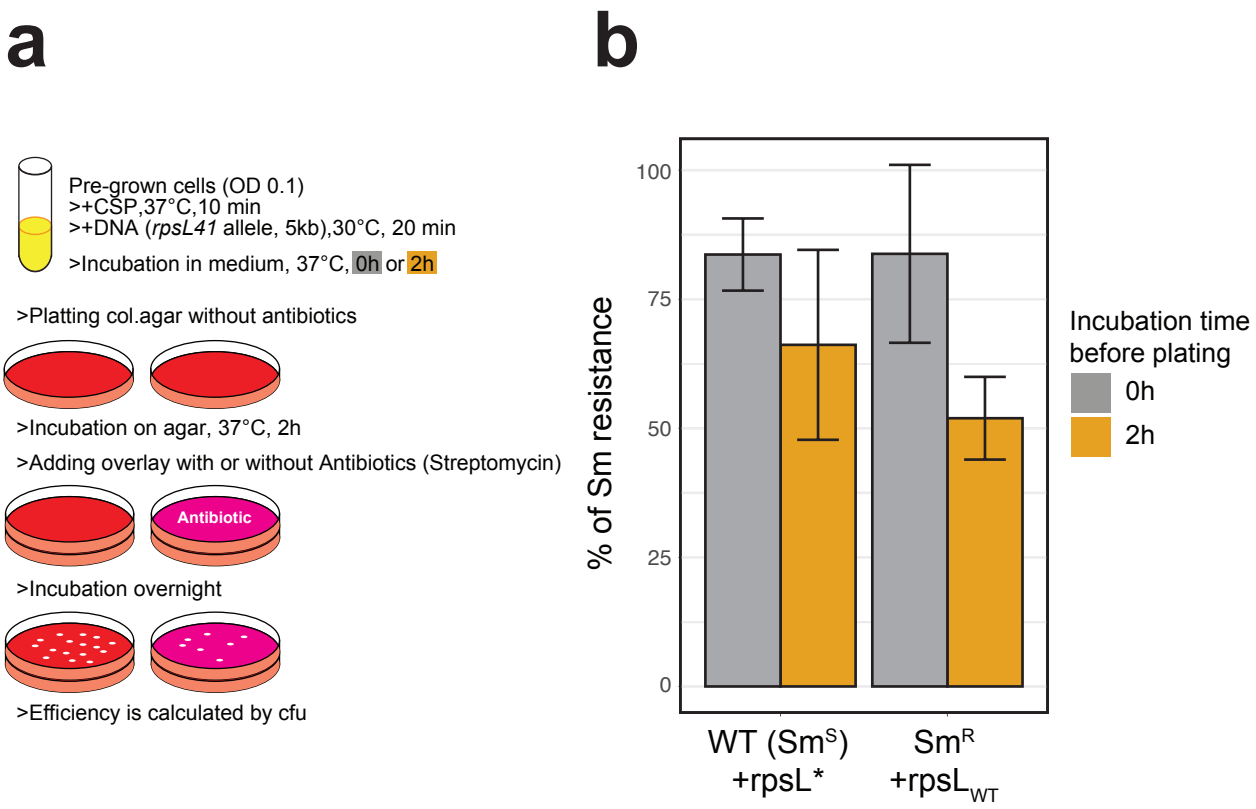


Figure 3–figure supplement 2

A

B

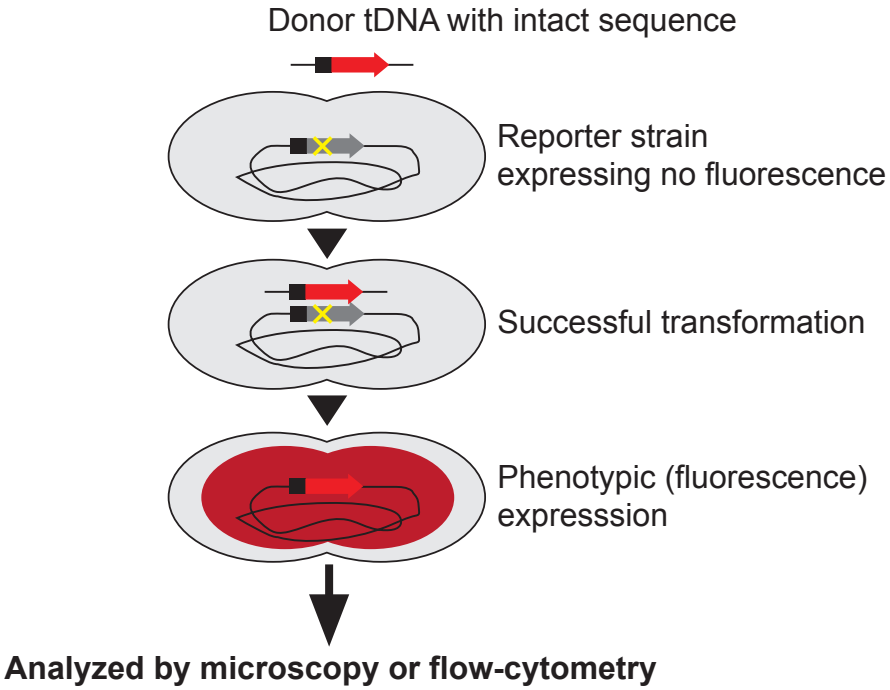


Figure 3–figure supplement 3

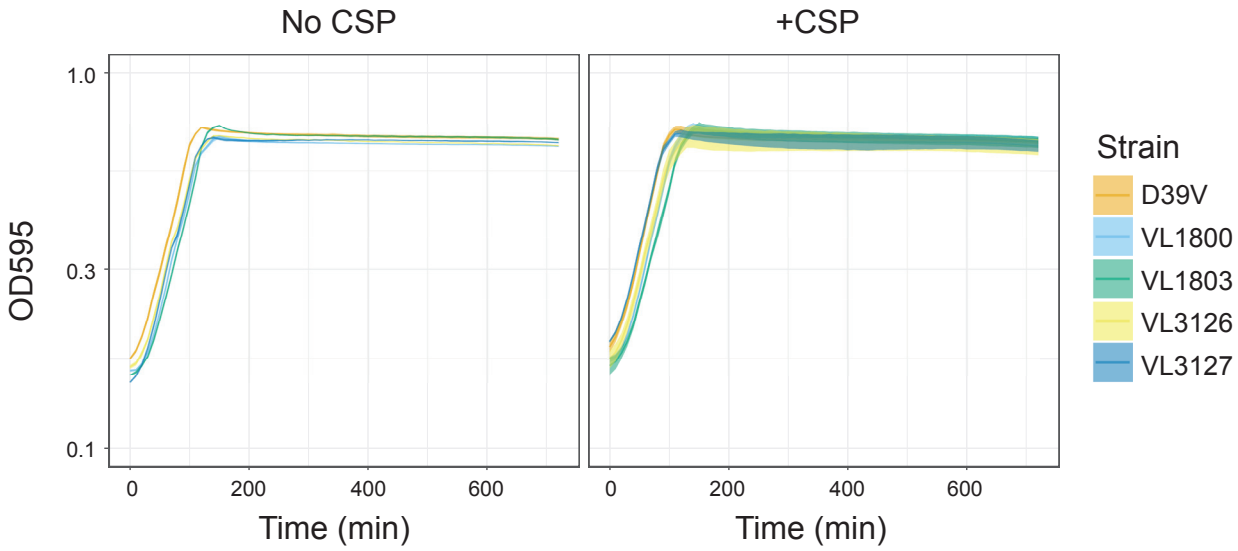


Figure 3–figure supplement 4

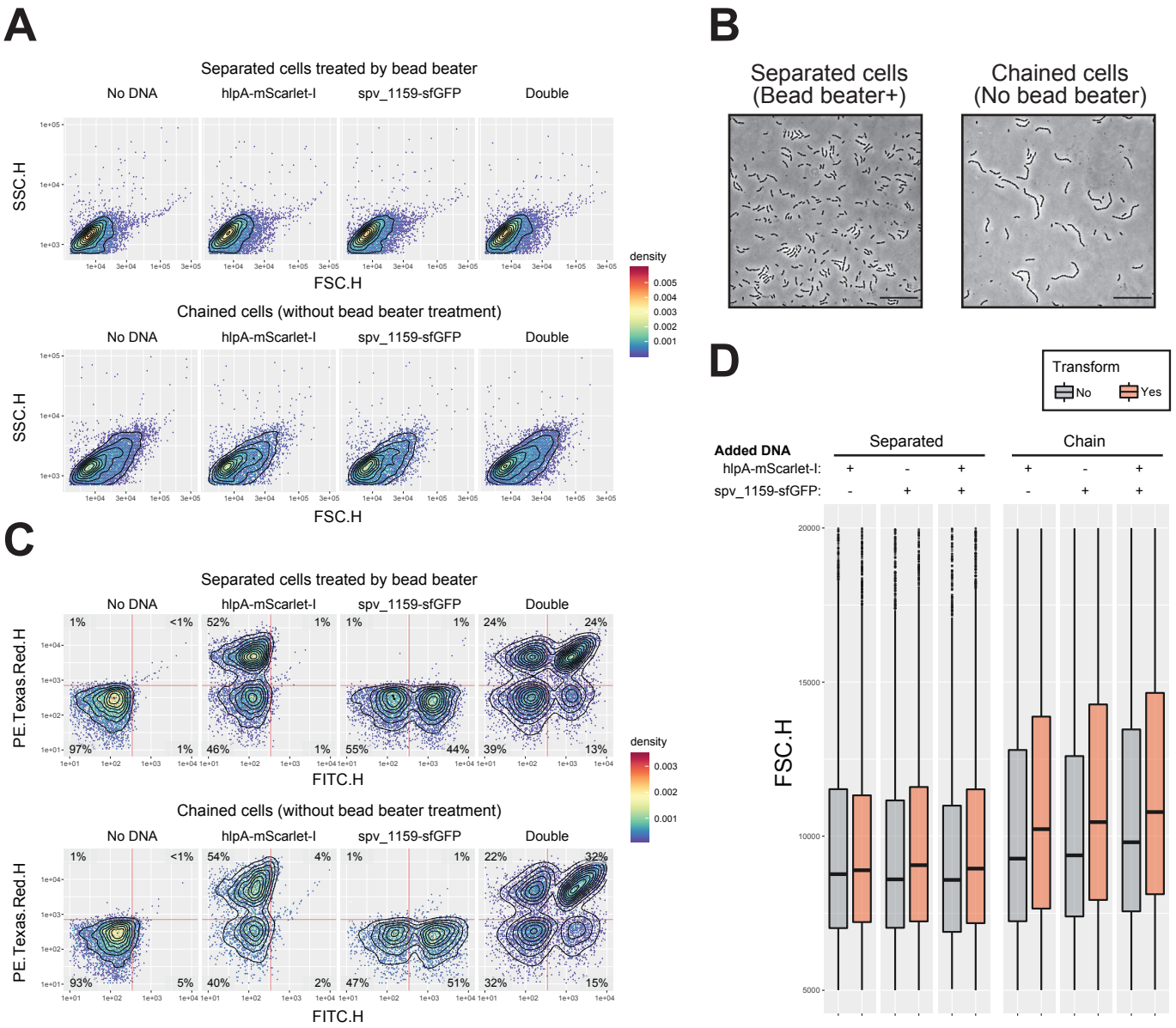


Figure 4–figure supplement 1

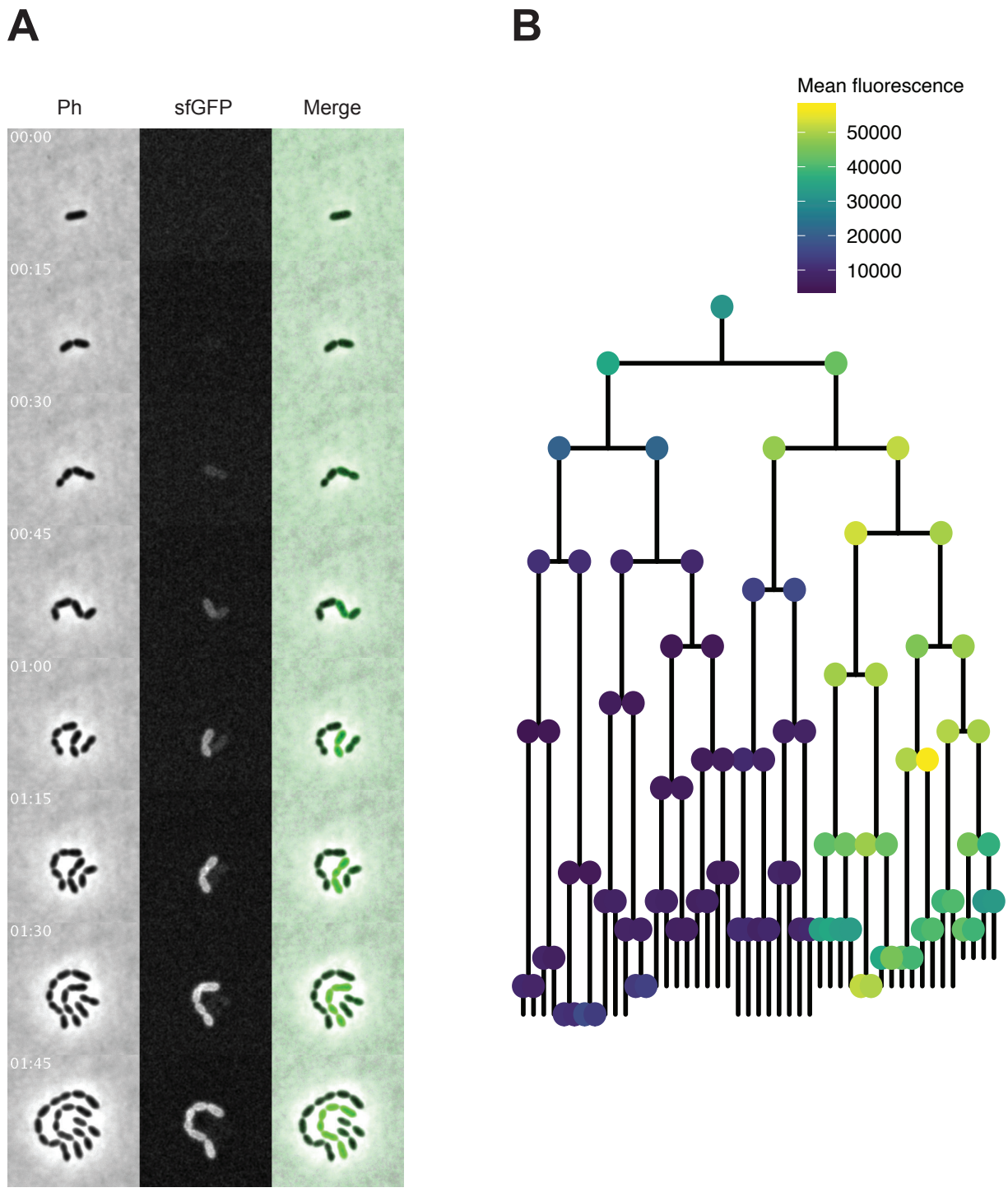


Figure 4—figure supplement 2

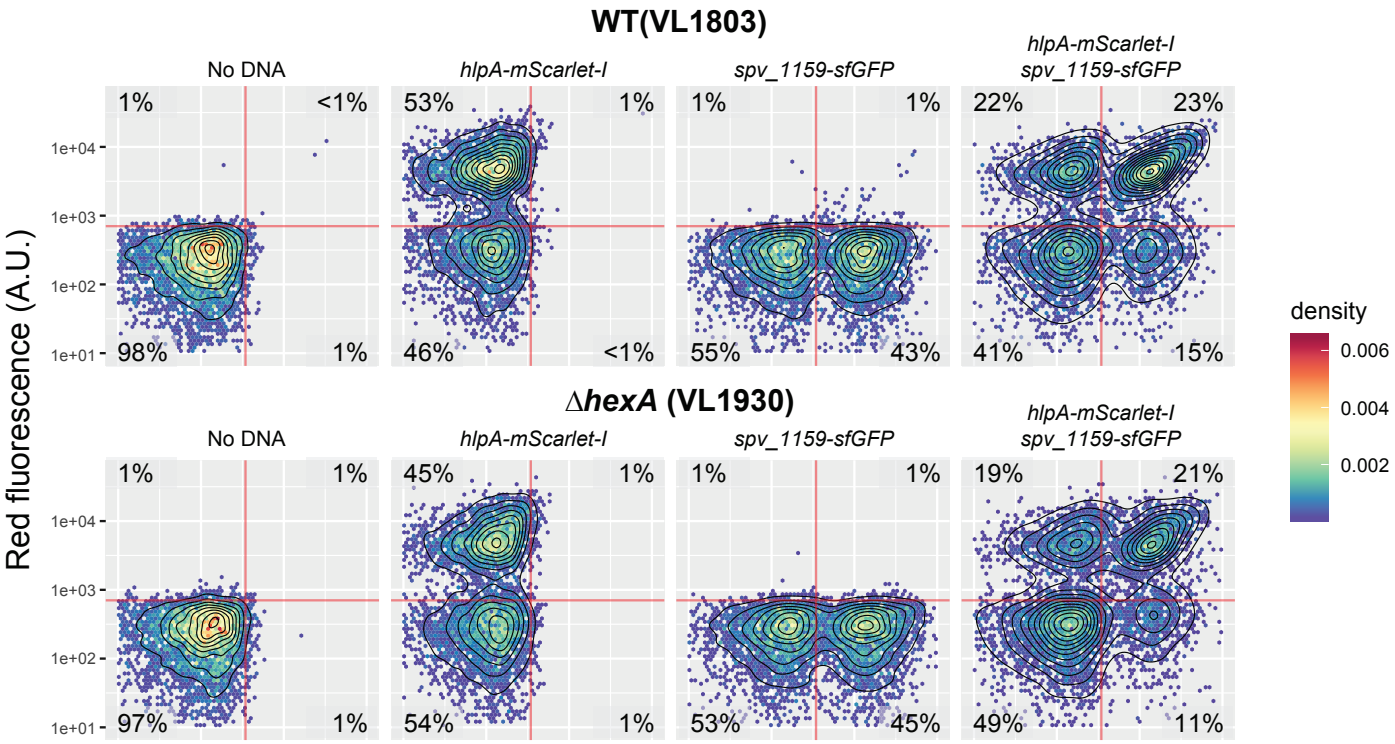


Figure 4–figure supplement 3

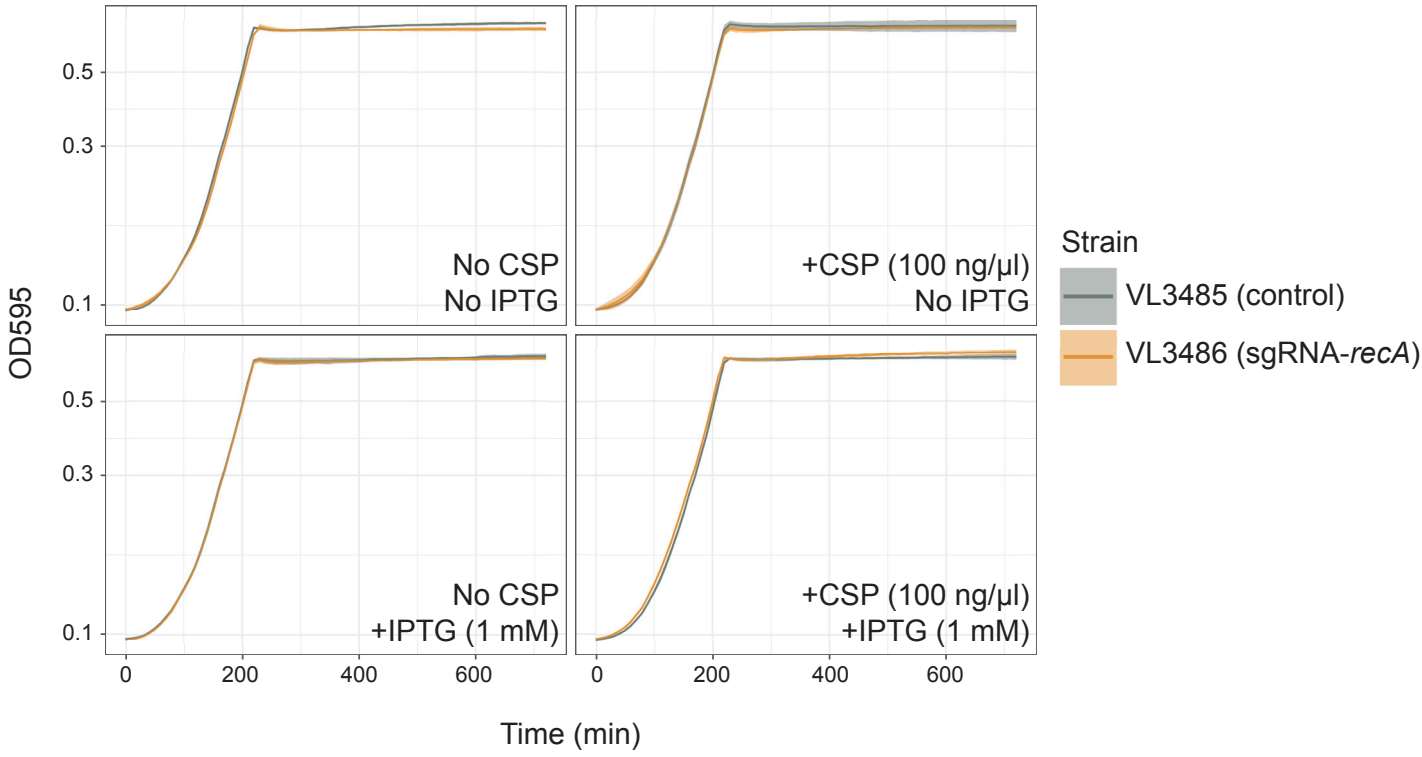


Figure 5–figure supplement 1

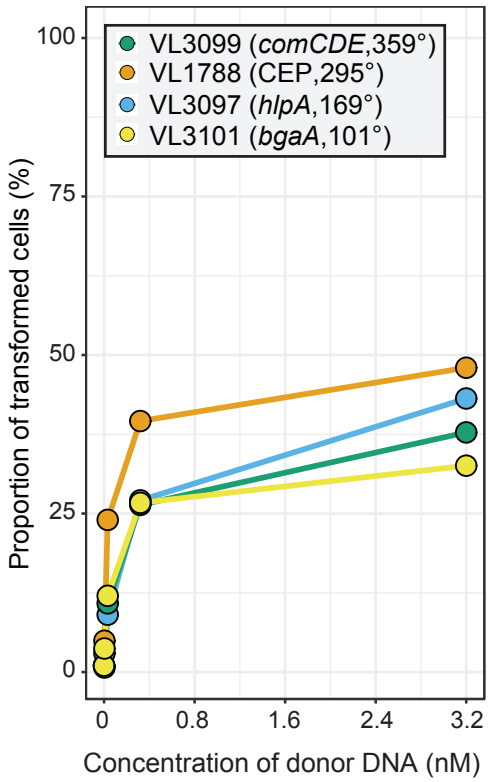
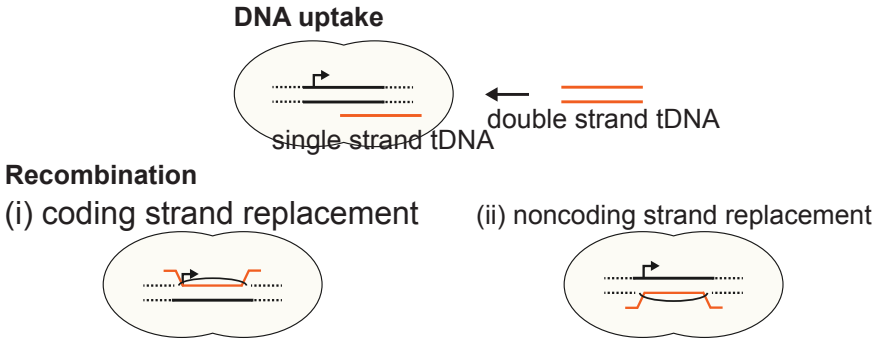


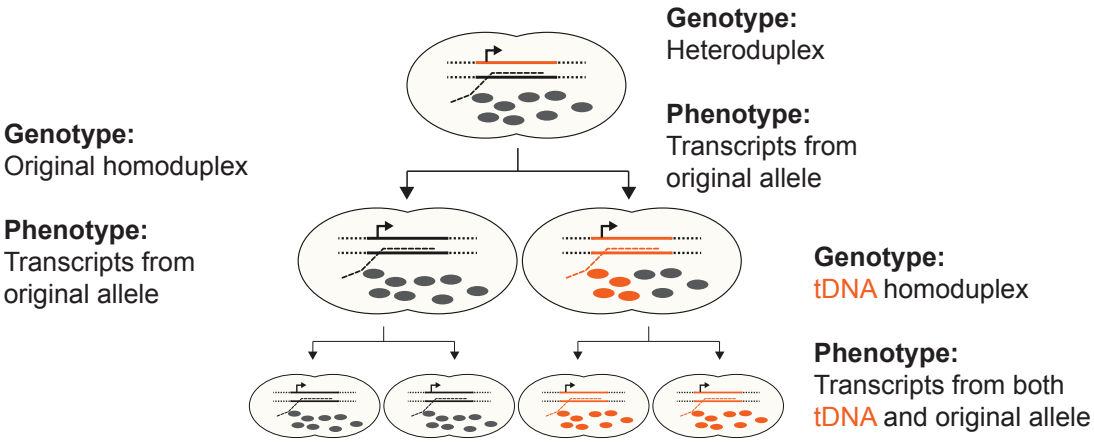
Figure 5–figure supplement 2

A

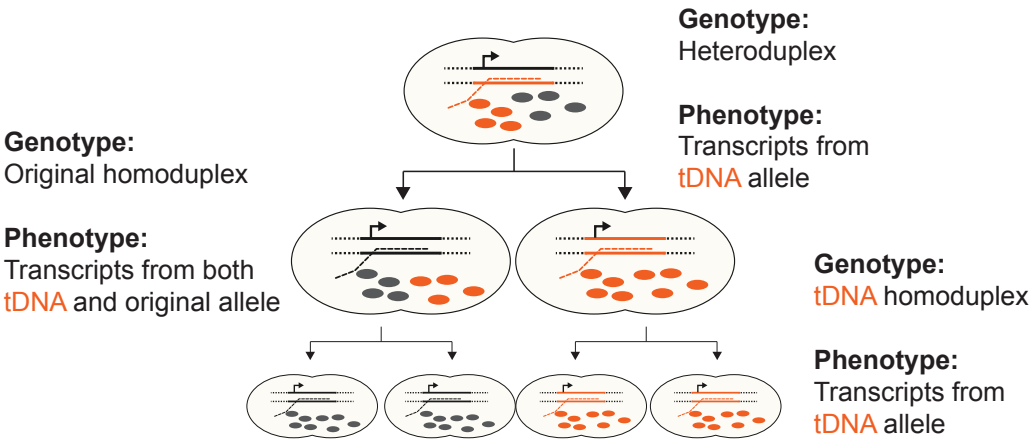


B

(i) Allele replacement on coding strand (late phenotypic expression)



(ii) Allele replacement on noncoding strand (early phenotypic expression)



C

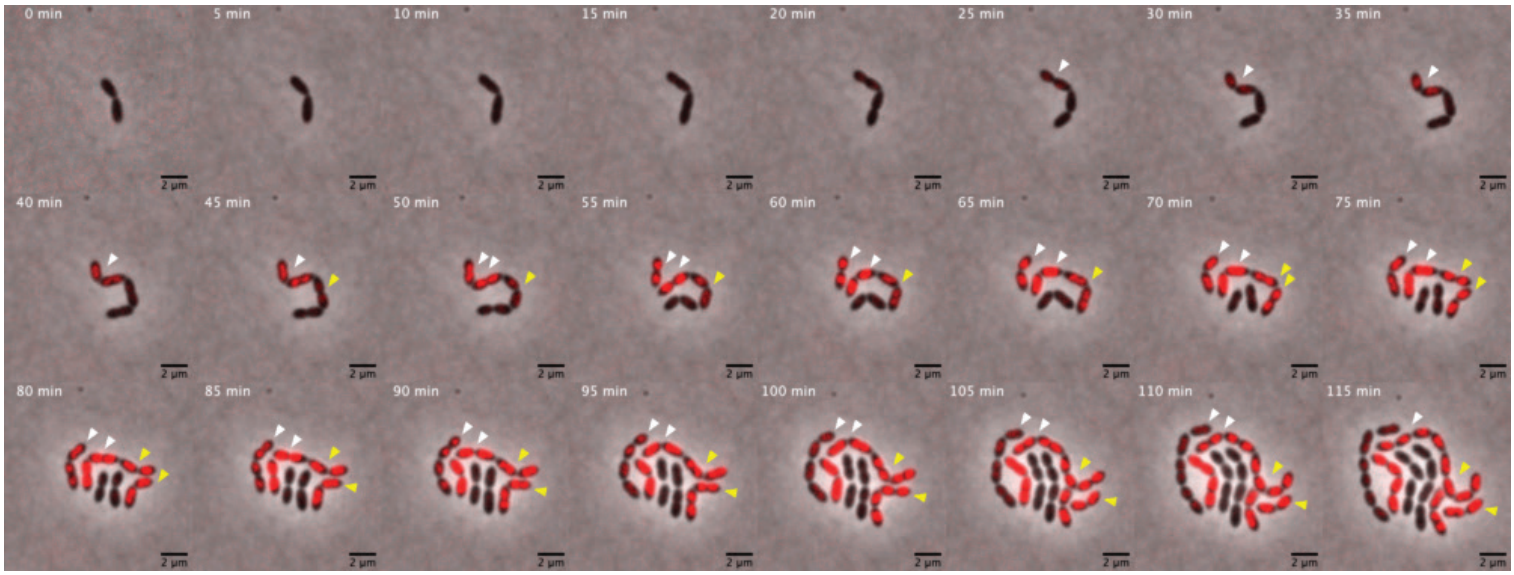


Figure 6–figure supplement 1