The acquisition of clinically relevant amoxicillin resistance in *Streptococcus pneumoniae* requires ordered horizontal gene transfer of four loci

Paddy S. Gibson[1], Evan Bexkens[1], Sylvia Zuber[1], Lauren A. Cowley[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] Department of Biology & Biochemistry, Milner Centre for Evolution, University of Bath, Claverton Down Campus, Bath BA2 7AY, United Kingdom

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

Abstract

Understanding how antimicrobial resistance spreads is critical for optimal application of new treatments. In the naturally competent human pathogen *Streptococcus pneumoniae*, resistance to β-lactam antibiotics is mediated by recombination events in genes encoding the target proteins, resulting in reduced drug binding affinity. However, for the front-line antibiotic amoxicillin, the exact mechanism of resistance still needs to be elucidated. Through successive rounds of transformation with genomic DNA from a clinically resistant isolate, we followed amoxicillin resistance development. Using whole genome sequencing, we showed that multiple recombination events occurred at different loci during one round of transformation. We found examples of non-contiguous recombination, and demonstrated that this could occur either through multiple D-loop formation from one donor DNA molecule, or by the integration of multiple DNA fragments. We also show that the final minimum inhibitory concentration (MIC) differs depending on recipient genome, explained by differences in the extent of recombination at key loci. Finally, through back transformations of mutant alleles and fluorescently labelled penicillin (bocillin-FL) binding assays, we confirm that *pbp1a*, *pbp2b*, *pbp2x*, and *murM* are the main resistance determinants for amoxicillin resistance, and that the order of allele uptake is important for successful resistance evolution. We conclude that recombination events are complex, and that this complexity contributes to the highly diverse genotypes of amoxicillin-resistant pneumococcal isolates.

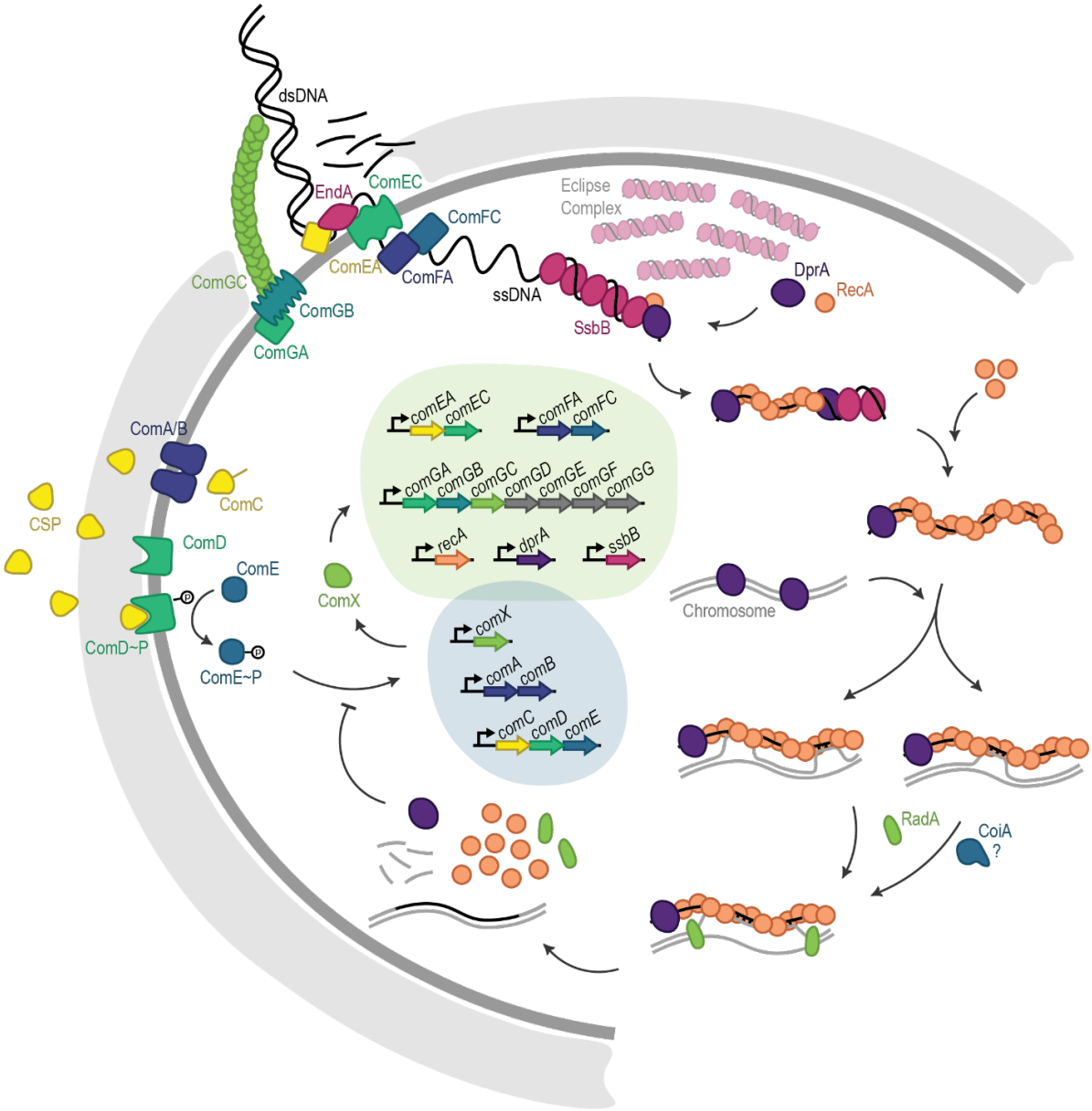
Introduction

Genomic plasticity through frequent and large-scale recombination drives the evolution of antibiotic resistance in *Streptococcus pneumoniae* (the pneumococcus) (Salvadori et al., 2019). This naturally competent member of the human nasopharyngeal commensal microbiota opportunistically causes otitis media as well as severe invasive diseases such as pneumonia, bacteremia, and meningitis (O'Brien et al., 2009; Shak et al., 2013). Despite conjugate vaccine introduction, the species remains an important human pathogen, as vaccine-escape and antibiotic resistant variants constantly arise (Croucher et al., 2011; Wyres et al., 2013).

Variants occur primarily through natural transformation and homologous recombination of DNA from strains or closely related species that occupy the same niche, such as *S. mitis* and *S. oralis* (Croucher et al., 2012; Dowson et al., 1993; Sibold et al., 1994). One case study showed more than 7% of the genome had been transferred between two pneumococcal strains during the polyclonal infection of a pediatric patient over three months (Hiller et al., 2010). Early studies of transformation in the pneumococcus showed uptake of fragments ranging from 2-6 Kb (Gurney and Fox, 1968), but cell-to-cell contact was found to facilitate transfer of long DNA fragments (Cowley et al., 2018), and recombinant regions as large as 30 – 50 Kb in length have been observed in clinically relevant lineages such as PMEN-1 (Croucher et al., 2011; Wyres et al., 2013, 2012). Events such as these have played a significant role in shaping the pneumococcal pangenome (Donati et al., 2010).

Natural competence is activated when the competence stimulating peptide (CSP) interacts with the ComDE two-component system (Straume et al., 2015) (Figure 1), where phosphorylated ComE activates transcription of the genes encoding the alternative sigma factor *comX* (Martin et al., 2013; Slager et al., 2019; Ween et al., 1999). ComX then activates transcription of the late *com* genes necessary for DNA uptake and recombination. This includes the type IV-like pilus (ComGC), endonuclease EndA, and transport proteins ComFA, and ComEC (Bergé et al., 2013; Lacks and Greenberg, 1976; Lacks and Neuberger, 1975; Laurenceau et al., 2013; Rosenthal and Lacks, 1980). RecA and the transformation-dedicated DNA processing protein A (DprA) are essential for single-stranded DNA (ssDNA) uptake without degradation (Attaiech et al., 2011). Although not fully understood, DprA and RecA are thought to polymerize on ssDNA upon cell entry, initiating the homology search, while

64 additional DNA molecules are then bound by multiple SsbB proteins protecting them from
65 degradation by endogenous nucleases (Attaiech et al., 2011; Morrison et al., 2007). It has
66 been hypothesized that the accumulation of stable ssDNA-SsbB eclipse complexes in the
67 cytoplasm enables successive recombination events during one competence window,
68 increasing genetic plasticity (Attaiech et al., 2011). Sequestered ssDNA can then be accessed
69 by the transformation-dedicated DNA processing protein A (DprA), which mediates the
70 loading of RecA onto the pre-synaptic filament (Lisboa et al., 2014; Mortier-Barrière et al.,
71 2007; Quevillon-Cheruel et al., 2012). DprA is also involved in competence shut-off, along with
72 ComX and RpoD sigma factor competition, and CSP degradation by HtrA (Cassone et al., 2012;
73 Martin et al., 2013; Mirouze et al., 2013; Piotrowski et al., 2009; Weng et al., 2013).



74

Figure 1: DNA uptake and recombination during natural competence in *Streptococcus pneumoniae*. Schematic overview of the regulatory steps required for the pneumococcus to uptake and recombine extracellular DNA. Competence is activated when CSP interacts with the ComDE two-component system, where phosphorylated ComE activates transcription of ComX and other early competence genes (darker grey bubble). This results in a positive feedback loop, where ComC is produced, then modified and exported from the cell by ComA/B as CSP. Meanwhile, ComX activates transcription of late competence genes, including the type IV-like pilus (ComGC), which brings DNA within reach of endonuclease EndA. EndA performs random nicks, then degrades one strand as the other is pulled into the cell and associated with DprA, RecA, and SsbB, to protect from degradation (exact mechanism not fully understood). DprA mediates loading of RecA onto SsbB-ssDNA eclipse complexes to form the presynaptic filament which then searches the chromosome for homology. RecA-ssDNA filaments may interact with the chromosome at multiple points, while successful base pairing between incoming and chromosomal DNA facilitates strand invasion and exchange. The helicase RadA facilitates D-loop extension, and CoiA is essential for successful recombination during transformation, although its exact role is not known. DprA is also involved in competence shut off, through blocking phosphorylation of ComE.

75

Competence is tightly controlled, with expression of most genes turned off within 20 minutes of CSP induction, providing a small window for DNA uptake and recombination (Alloing et al., 1998). To accelerate homology detection, the RecA-ssDNA presynaptic filament has been shown to interact with the chromosome in 3-dimensional space at multiple points of contact (Forget and Kowalczykowski, 2012; Wiktor et al., 2021; Yang et al., 2020). Displacement-loop (D-loop) formation is facilitated by base pairing between the chromosome and the ssDNA molecule, within the two DNA-binding sites of RecA (Bell and Kowalczykowski, 2016; Torres et al., 2019). Eight consecutive bases must be successfully paired before the DNA-DNA interaction is stabilized and strand exchange can occur (Hsieh et al., 1992; Mazin and Kowalczykowski, 1996). In *S. pneumoniae*, D-loop extension is promoted by RadA (DnaB-type helicase) which helps to unwind chromosomal DNA, extending recombination over long genomic distances (Marie et al., 2017), and while its function is not known, CoiA is essential for successful recombination (Desai and Morrison, 2007; Mortier-Barrière et al., 2007). Although initial testing is highly stringent, once crossover has started and the region of complementation is extended, stringency decreases, allowing some mismatches and thereby increasing desirable genetic variation (Carrasco et al., 2016; Danilowicz et al., 2015). This is aided by the generalized mismatch repair system in *S. pneumoniae* (hex) which is not induced during competence (Claverys and Lacks, 1986), is only able to identify mismatches in the donor-recipient heteroduplex before D-loop resolution (Claverys and Lacks, 1986; Gasc et al., 1989), (Gasc et al., 1989) and is easily overwhelmed by excess mismatches (Humbert et al., 1995).

Penicillin binding proteins (PBPs) are key β-lactam resistance determinants and are essential for bacterial cell wall synthesis by catalyzing the transglycosylation (TG) and transpeptidation (TP) reactions responsible for peptidoglycan (PGN) formation (Scheffers and Pinho, 2005). While highly conserved in penicillin (PEN) sensitive strains, PBPs of resistant isolates are highly variable, with up to 10% of amino acid changes (Hakenbeck et al., 2012). Significant amounts of this variation was recently confirmed to be acquired from *S. mitis* (Kalizang'oma et al., 2021). These large-scale changes alter the structure of the active sites, reducing β-lactam binding affinity (Hakenbeck et al., 2012). Amoxicillin (AMX) is the first line oral antibiotic prescribed for bacterial lower respiratory tract infections. AMX resistant strains first appeared in the mid-90s in the United States and France (Doern et al., 1996; Doit et al., 1999) and are

107 currently on the rise in Spain, a phenomenon which correlates with increased usage of oral
108 AMX/clavulanic acid (Càmara et al., 2018). Early AMX resistant lineages were closely related
109 to known PEN resistance clones already circulating in the population (Doit et al., 1999;
110 Stanhope et al., 2007). Sequencing of *pbp* alleles from susceptible and resistant strains found
111 high variation in *pbp2x* and *pbp1a*, while SNPs shared between AMX resistant isolates were
112 only found in *pbp2b* (Cafini et al., 2006; Chesnel et al., 2005; du Plessis et al., 2002; Kosowska
113 et al., 2004). However, no SNP or block of SNPs in any one *pbp*, when transferred to a
114 susceptible strain, has been sufficient to confer resistance in a susceptible strain (Hakenbeck
115 et al., 1998; Kosowska et al., 2004). This is not entirely surprising, as the transpeptidase
116 activity of PBPs is an essential function, and modifications to the active site are likely to have
117 deleterious effects which need to be compensated for (Albarracín Orio et al., 2011)(Albarracín
118 Orio et al., 2011). One mechanism of fitness compensation appears to be through
119 substitutions in MurM, which results in abnormally high proportions of branched
120 muropeptides in the cell wall (Filipe et al., 2002; Filipe and Tomasz, 2000; Garcia-Bustos and
121 Tomasz, 1990; Smith and Klugman, 2001). Branched muropeptides were shown to have
122 increased importance in strains depleted for Pbp2b, indicating a link between muropeptide
123 composition and cell elongation (Berg et al., 2013). Other non-target proteins which have
124 been implicated in resistance include CpoA (LafB) (Grebe et al., 1997) and CiaH (Guenzi et al.,
125 1994).

126 Although multiple residues in Pbp1a, Pbp2x, and Pbp2b were found to be under positive
127 selection in AMX resistant isolates (Stanhope et al., 2008), a specific a block of ten
128 substitutions in the 590-641 region of the Pbp2b TP domain is repeatedly associated with
129 AMX resistance (Cafini et al., 2006; du Plessis et al., 2002; Kosowska et al., 2004). However,
130 transformation of *pbp2x*, *pbp2b*, and *pbp1a* alleles to AMX susceptible strain R6 did not
131 recreate the AMX resistance of the donor strain, and the addition of *murM* had no effect.
132 Only when genomic DNA was used could the original resistance level be achieved, indicating
133 the role of a non-*pbp* or *murM* resistance determinant (Chesnel et al., 2005; du Plessis et al.,
134 2002), or unknown epistatic interactions (Arnold et al., 2018). Interestingly, other studies
135 have found a correlation between *murM* mutations and AMX resistance, although their
136 presence depended on the *pbp2b* allele (Cafini et al., 2006; Càmara et al., 2018). Taken

137 together, these studies strongly suggest the existence of multiple routes to AMX resistance
138 and imply a tight link in the roles of Pbp2b and MurM in the mechanisms.

139 To study recombination and the development of AMX resistance in the pneumococcus in
140 more detail and identify possible evolutionary routes towards AMX resistance, we
141 sequentially transformed genomic DNA from a resistant clinical isolate (serotype 11A, ST6521,
142 German National Reference Centre for Streptococci 2017, SN75752) into two different
143 recipient strains coming from different clonal complexes (serotype 2 D39V and serotype 4
144 TIGR4). Whole genome sequencing showed recipient-dependent, genome-wide mutation
145 uptake. In addition to identifying the necessary determinants for AMX resistance, we found a
146 strong tendency towards the order of *pbp* and *murM* allele uptake. Long read Pacbio
147 sequencing excluded a role for DNA methylation or genomic rearrangements or movement
148 of genetic mobile elements as resistance determinants. We also investigated non-contiguous
149 recombination, previously observed in *S. pneumoniae*, *Haemophilus influenzae*, and
150 *Helicobacter pylori* (Croucher et al., 2012; Kulick et al., 2008; Mell et al., 2011), and propose
151 a model for how these complex events occur. To conclude, we show that the uptake of four
152 alleles from a resistant strain (*pbp2X*, *pbp2B*, *pbp1A* and *murM*), in a specific order, is required
153 and sufficient for clinically relevant resistance development towards amoxicillin in sensitive
154 *S. pneumoniae*.
155

156 **Results**

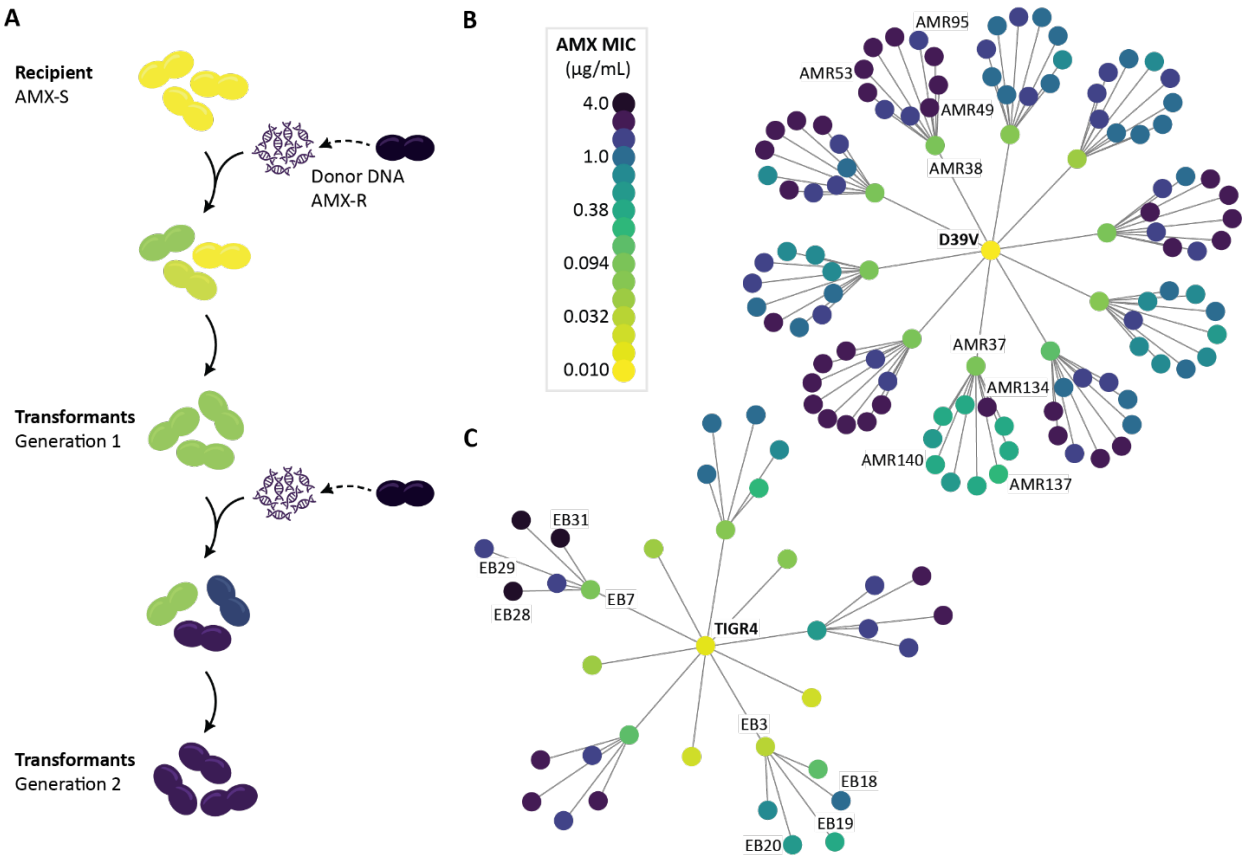
157 ***Serial transformation with genomic DNA from an AMX resistant strain increased AMX MIC***

158 In order to understand the sequence of recombination events leading to the development of
159 AMX resistance, a susceptible *S. pneumoniae* strain D39V (minimum inhibitory concentration
160 (MIC) 0.01 µg/mL) was transformed in two successive rounds with genomic DNA originating
161 from AMX resistant clinical isolate 11A (MIC 4 µg/mL) followed by selection on a range of AMX
162 concentrations above MIC (Figure 2A, Table 1). In the first round, ten colonies were picked at
163 random from the highest AMX concentration with growth (0.03 µg/mL AMX). The MICs of
164 these strains ranged from 0.047 – 0.125 µg/mL, 4 – 10 times that of wild-type D39V (Figure
165 2B, Table S1).

166 The second round of transformation resulted in 100 strains with MICs from 0.25 to 2 µg/mL
167 (Figure 2B), lower than the donor strain, and a third round of transformation did not increase

168 the MIC further. Interestingly, five of the ten lineages had seven or more strains with MICs of
169 1 µg/mL or higher. In contrast, the AMR37 lineage resulted in MICs from 0.25 – 0.5 µg/mL,
170 with one exception at 2 µg/mL. These lineage-dependent differences in AMX MIC range (p =
171 5 x 10^{-4}, Fisher's Exact Test) suggested that mutations acquired in the first round contributed
172 to the recombination events which occurred in the second round, and thus to the MICs of the
173 resulting strains.

174



175

Figure 2: **Transformation as an experimental tool to track the evolutionary path to AMX resistance**. **(A)** Experimental design. Susceptible D39V and TIGR4 were transformed with genomic DNA from an AMX resistant clinical isolate then selected on a range of AMX concentrations. Ten colonies were picked at random from the highest concentration with growth (first round on 0.03 µg/mL and second round on 0.17 – 0.3 µg/mL, depending on MIC of the recipient strain). Isolated transformants with increased MICs were then subjected to a second round of transformation. **(B)** Pedigree chart of recombinant strains derived from D39V, and **(C)** from TIGR4. Color represents AMX MIC. Strains which were sequenced and used throughout the study are labelled.

176 Similar lineage-dependent MIC patterns were seen when TIGR4 was used as the recipient
177 strain (p = 1.00x10^{-3}, Fisher's Exact Test) (Figure 2B). Here, the AMX MICs of the first-
178 generation ranged from 0.023 – 0.5 µg/mL, with one strain showing a more than 30-fold
179 increase from wild-type TIGR4. In the second-generation, three recombinant strains reached

180 the same MIC as the donor strain (4 μg/mL), suggesting potential key differences in the
181 recipient strain genomes, or deleterious epistatic effects of AMX resistance mutations which
182 affected resistance development through recombination.
183
184 ***Cell lysis upon AMX treatment correlated with PBP affinities***
185 To test whether transformed strains phenocopy the donor with regards to AMX resistance,
186 phase contrast microscopy with an AMX concentration above the recipient MICs (1 μg/mL)
187 was performed. A complete stall in growth was observed for both recipient strains and first-
188 generation recombinant AMR38, with mild to severe lysis after 6 hrs of treatment (Figure 3A).
189 In contrast, donor strain 11A grew to form a microcolony in this AMX concentration, while
190 second-generation AMR53 grew more slowly and with abnormally elongated cells (Figure 3A).
191 To test whether the increased resistance of AMR53 is related to alterations in the affinity of
192 the PBPs towards AMX, we performed Bocillin-Fl labelling. Bocillin-Fl is a fluorescently labelled
193 penicillin derivative that binds all six PBPs of *S. pneumoniae* (Zhao et al., 1999), and the
194 intensity of labelling can be used as a proxy for the affinity of the PBP for β-lactam antibiotics
195 (Kocaoglu et al., 2015). All six PBPs of D39V and TIGR4 had a high affinity for the fluorescently
196 labelled penicillin (Figure3B). Interestingly, TIGR4 showed some evidence of Pbp1A
197 degradation products that bound Bocillin-Fl with high affinity even though it is 99.7 % identical
198 to D39V Pbp1A. In contrast, 11A showed an altered PBP pattern, where PBP2b and Ppb1a
199 showed complete loss of affinity, and Pbp2x was labelled more faintly and migrated more
200 slowly (Figure 3B).
201 Both first- and second-generation recombinants demonstrated a loss of affinity of Pbp2b and
202 Pbp2x for Bocillin-FL, with PBP1a in AMR53 also showing reduced affinity, and potentially
203 some degradation (Figure 3B). Together, these characterizations demonstrate that
204 transformed recipients phenocopy the donor AMX resistant strain and have acquired PBPs
205 with reduced affinity towards bocillin-FL.

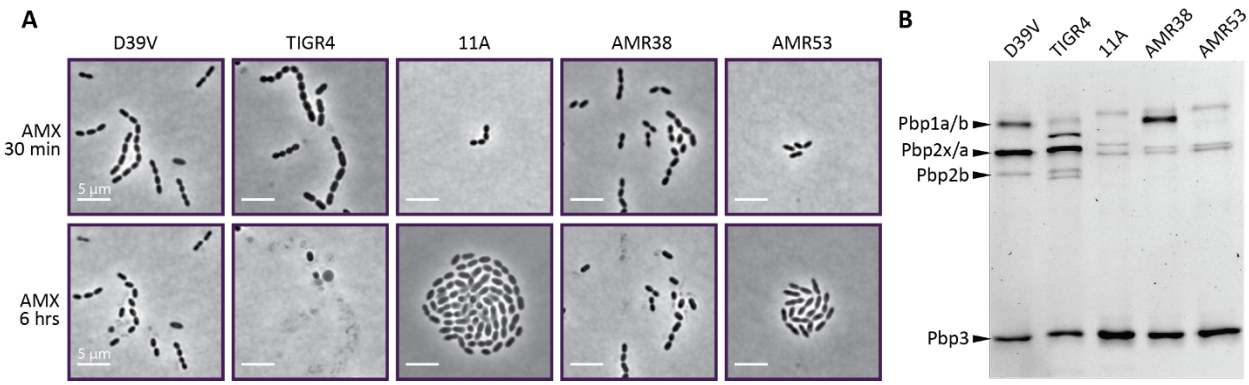


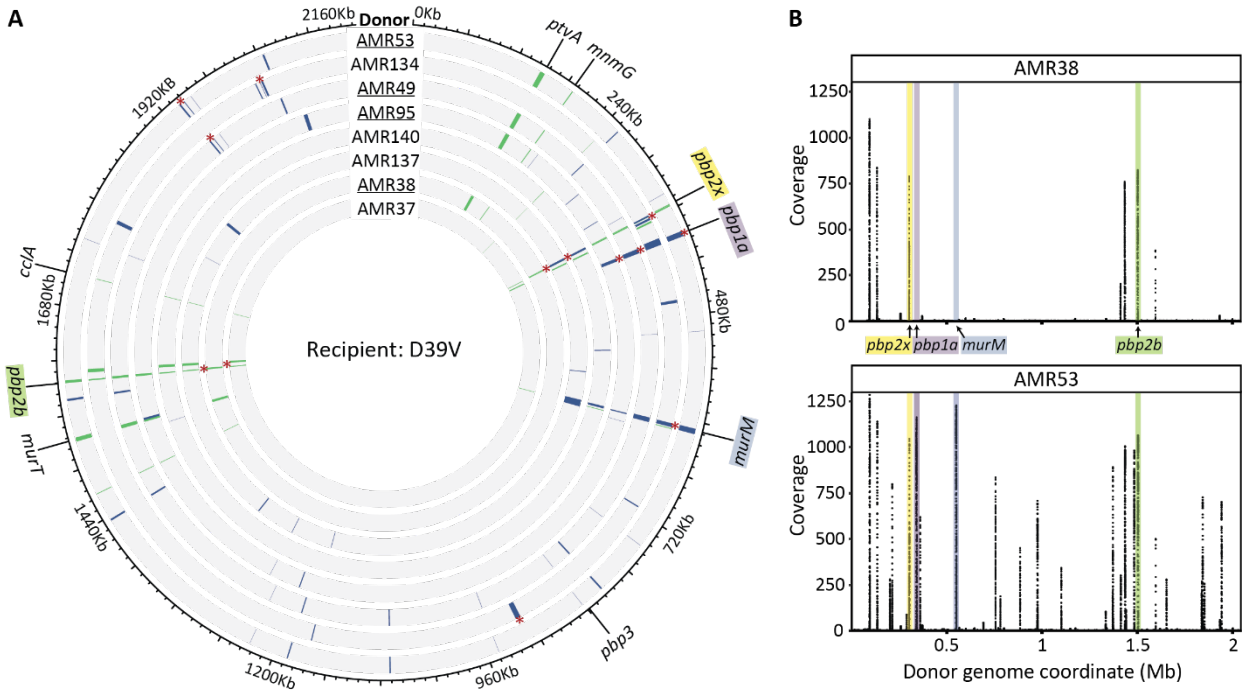
Figure 3: Phenotypic characterization of key strains used in this study. (**A**) Phase-contrast snapshots of recipient strains D39V and TIGR4, donor strain 11A, and recombinant strains AMR38 (generation 1) and AMR53 (generation 2) on C+Y agarose pads after 30 minutes and 6 hours of treatment with 1 µg/mL AMX. This concentration was above MIC for D39V, TIGR4, and AMR38, and lead to a complete stall in growth and partial lysis. Donor strain 11A grew well in this AMX concentration, although some heterogeneity in cell shape could be observed. Although lower than the AMX MIC for AMR53, this strain showed decreased growth and abnormally elongated cells, indicating cells were stressed. Scale bar is 5 µm. (B) Bocillin-FL binding patterns for the same five strains. Expected PBP band sizes for D39V are shown on the left. Image of Coomassie stained gel can be found in Figure S1)

Recombination events clustered around resistance-associated loci

In order to examine the genetic differences in recombinant strains that may explain the differences in MIC, we selected eight recombinant strains from each experiment, and Illumina sequenced to ~600-fold mean coverage. For the D39V experiment, we chose AMR38 and three of its descent strains, as this lineage had the highest MICs. To contrast, we also chose AMR37 and three descendent strains, which had the lowest and most variable MICs in the second generation (Fig 2B). We used a similar selection criteria for the TIGR4 experiment, taking strains from the high MIC EB7 lineage, and from the low MIC EB3 lineage (Fig. 2C).

To be able to accurately map recipient to donor, we also performed long-read PacBio sequencing followed by short-read polishing on the 11A donor and TIGR4 recipient (an accurate genome map was already available for D39V (Slager et al., 2018)). Recombination was detected using SNPs, as described in Cowley et al. (2018). On average 0.38 % ± 0.22 % of the donor genome was transferred during a single transformation event, and the total percentage transferred after two rounds was 0.80 ± 0.05%. Although higher numbers of recombination events were detected in the D39V dataset, the average event length was larger in the TIGR4 dataset (D39V 0.79 Kb and TIGR4 2.29 Kb, Table S3 and S4), resulting in similar total amounts of donor genome transfer after two rounds of transformation (D39V 0.79 ±

224 0.23%, TIGR4 0.80 ± 0.16%, Table S3 and S4). Of note, the longest recorded recombination
225 fragment was 4.00 Kb for D39V and 12.52 Kb for TIGR4.
226 Recombination events clustered around key loci for pneumococcal β-lactam resistance in
227 both datasets; *pbp2x*, *pbp2b*, *pbp1a* and *murM* (Figure 4A and 5C). SNPs were also found at
228 closely located loci, likely due to linkage. In the D39V strains, there was extensive
229 recombination at distantly located sites. Examples include *ptvABC*, associated with
230 vancomycin tolerance (Liu et al., 2017), and *murT*, which performs an amidation step essential
231 for peptidoglycan crosslinking (Figueiredo et al., 2012; Münch et al., 2012).
232



233

Figure 4: **Widespread transformation and uptake of 11A donor DNA in recipient D39V upon AMX selection**. **(A)** Circos plot with detected recombination events in D39V derived strains, arranged from lowest to highest MIC from the center. AMR38 and all derived strains are underlined, AMR37 and all derived strains are not. Grey indicates the sequences matches that of the recipient D39V, while recombination events are colored. Events acquired in the first round of transformation are shown in green, in the second round in blue, and lengths were increased to aid in visualization (see Table S3 for a list of all detected events). Recombination events determined to be non-contiguous are marked with (*). Non-contiguous events detected in the first-generation were removed from the data set on the second-generation strains, so are not marked. **(B)** Coverage of reads mapped competitively to the donor genome and plotted by coordinate beginning at the origin. Recombination events cluster at known β-lactam resistance determinants, *pbp2x*, *pbp2b*, *murM*, and *pbp1a*.

234 In contrast, of the six total recombination loci in the TIGR4 dataset, only three were located
235 distally from the *pbp* and *murM* genes, and all were detected based on the presence of very
236 few SNPs.

237 We hypothesised that differences in recombination patterns could have arisen from reduced
238 similarity with the donor genome, however comparing total shared and unique SNPs did not
239 suggest major differences in variability between the two recipient genomes (Figure 5A). When
240 considering variation around key loci, we saw that the TIGR4 *pbp2b* region was more
241 divergent from the donor than D39V. The quality of donor DNA used may have also played a
242 role. Together, these results demonstrate that regardless of the recipient strain, four loci are
243 consistently acquired in high AMX resistant recombinants, namely *pbp2x*, *pbp2b*, *pbp1a* and
244 *murM* (Figure 5B).
245 Structural variation and base modification detection with PacBio sequencing revealed no
246 major structural changes or differences in methylated motifs in the D39V recombinant strains.
247 However, a local rearrangement was identified at the *hsdS-creX* locus in EB7, EB28, and EB29.
248 In this region, spontaneous reshuffling of *hsdS* genes by CreX results in different
249 methyltransferase specificity of the type I restriction modification system SpnD39III (Feng et
250 al., 2014; Manso et al., 2014; Slager et al., 2018). TIGR4 was found to be predominantly in the
251 B-configuration, EB7 and EB29 in the D-configuration, and EB28 in the C-configuration, while
252 all D39V-derived strains and 11A were mostly in the F-configuration. Differences in
253 methylated motifs have transcriptomic consequences (Manso et al., 2014), however MIC
254 differences among these recombinants can largely be explained by SNPs in key loci for β-
255 lactam resistance.

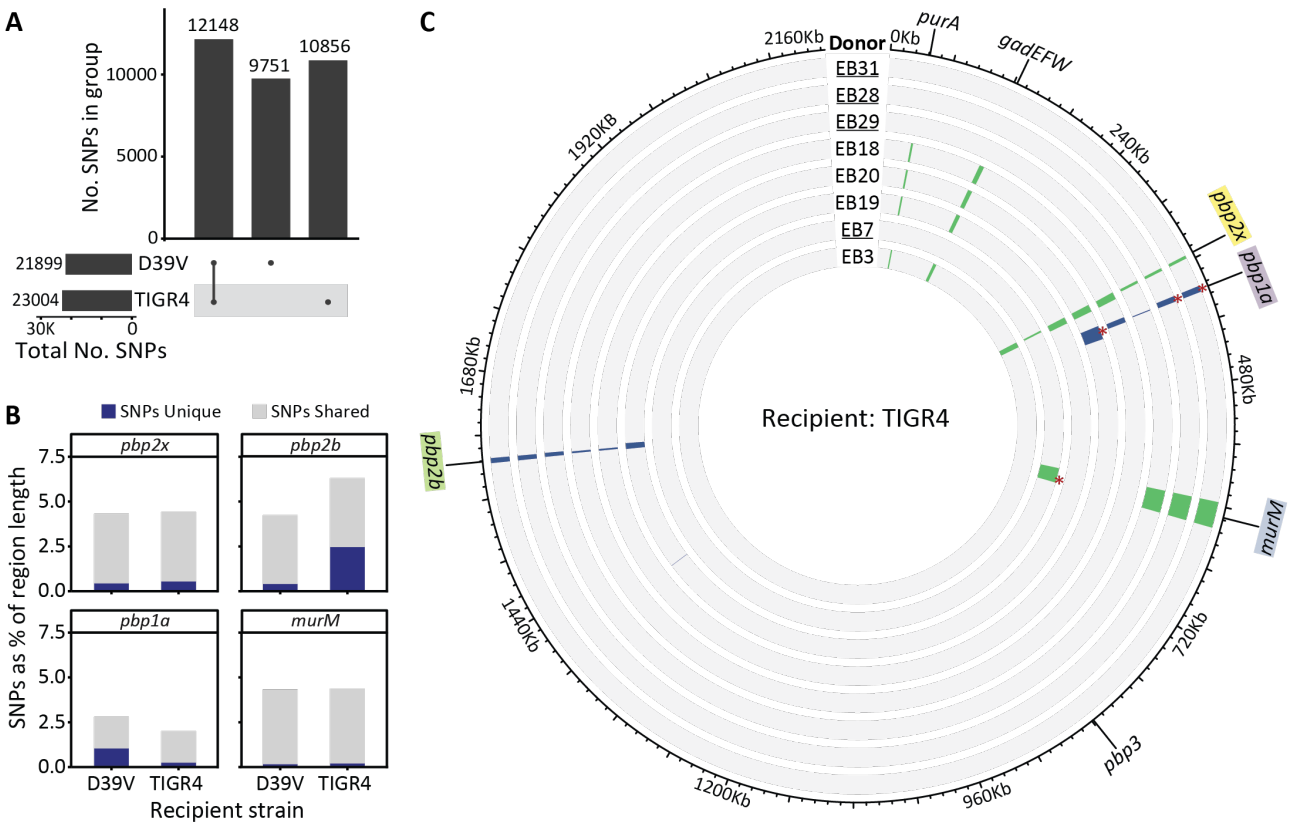


Figure 5: Uptake of four donor fragments is sufficient for AMX resistance in TIGR4. (A) Upset plot of shared and unique SNPs between each recipient and the donor. Bars on the left show the total number of SNPs for each recipient genome compared to the donor. In the central bar plot, from left to right, SNPs shared between D39V and TIGR4, SNPs found only in D39V, and SNPs found only in TIGR4. Although the two recipients differ from the donor strain by similar amounts of SNPs, only about half are unique. **(B)** SNP counts in pbp/murM genes, including 5 Kb up- and downstream. Number of SNPs unique to each recipient are shown in blue, and shared in light grey. **(C)** Circos plot showing detected recombination events in TIGR4 derived strains, arranged from lowest to highest MIC from inside to out. EB7 and all derived strains are underlined, EB3 and all derived strains are not. Grey indicates the sequences matches that of the recipient D39V, while recombination events are colored. Events acquired in the first round of transformation are shown in green, and in the second round in blue, and lengths were increased to aid in visualization (see Table S3 for a list of all detected events). Only six loci with recombination were identified in the entire data set, four of which are known β-lactam resistance determinants, *pbp2x*, *pbp2b*, *murM*, *pbp1a*.

Non-contiguous recombination events likely result from a single donor molecule

Recombination is often thought of as a donor molecule forming a single D-loop at a homologous site on the recipient chromosome, which is then expanded in both directions until the requirement for homology is no longer met, at which point the interaction ends. However, we detected recombination events from one round of transformation located very close together on the chromosome, sometimes within the same gene but with separating recipient SNPs, a motif reminiscent of clinically resistant mosaic *pbp* alleles. This has also been noted previously in both *in vitro* pneumococcal and *H. influenzae* recombination studies and

265 was hypothesized to result from a single donor molecule (Croucher et al., 2012; Mell et al.,
266 2011; Mortier-Barriere et al., 1997). To determine if events were significantly close together,
267 forming a single non-contiguous recombination event, a bootstrapping approach was applied
268 to the distances between novel recombination events within the same strain (Croucher et al.,
269 2012). In this approach, the shortest distances between all recombination events for all
270 strains were pooled to form the test population. Bootstrap tests were performed 1000 times,
271 where the test population was sampled with replacement, and the resulting distribution
272 compared to the shortest recombination distance for each strain (d_{test}). If d_{test} was below the
273 0.05 quantile value in at least 95% of bootstrap tests, then the events were considered to be
274 linked, and the next closest pairing was tested. We found 37 events across all strains that
275 could be amalgamated into single events under the assumptions of this approach (Figure S2,
276 Table S4).

277 In addition, using PCR amplified fragments of the *pbp2x-mraY* locus from 11A, we were able
278 to test, for the first time, whether non-contiguous recombination events occurred more
279 frequently when the template was provided as a single molecule, or as two overlapping
280 fragments. By ligating additional regions of non-homologous DNA, we also tested the effect
281 of molecule length on recombination.

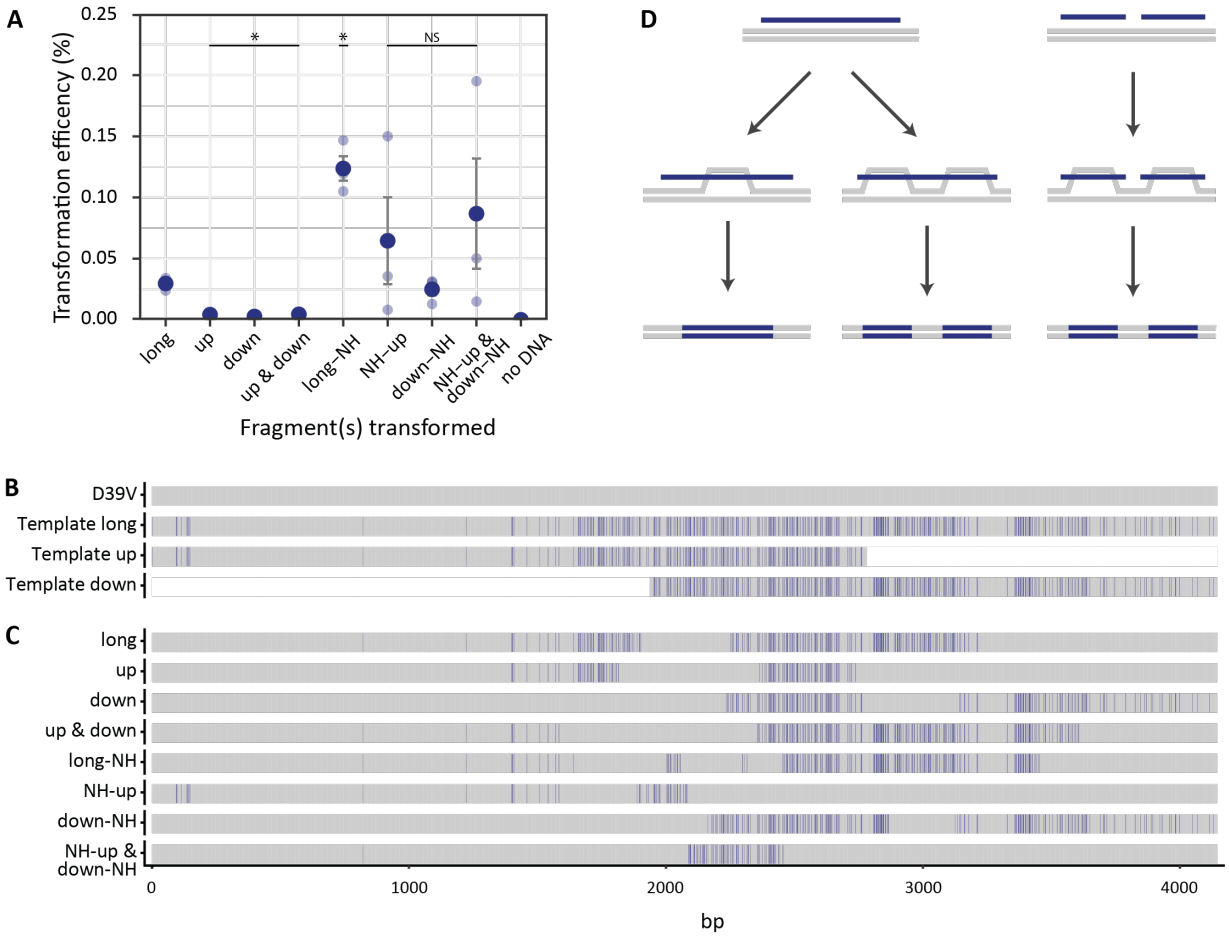
282 Transformation efficiency increased with length of both the overall molecule, and the
283 homologous region (Figure 6A). This is in line with previous studies showing the importance
284 of molecule length and homology in pneumococcal transformation (Kurushima et al., 2020;
285 Morrison and Guild, 1972). We hypothesize that increased length may increase the likelihood
286 of spontaneous contact with the ComGC pilus, while also decreasing the likelihood of
287 homologous region fragmentation by EndA, prior to entry (Morrison and Guild, 1973). The
288 efficiencies of the short overlapping fragments were lower than that of the long fragments,
289 although increased slightly when transformed in combination.

290 Recombination occurred in all colonies sequenced. When two overlapping fragments were
291 provided, uptake was often unequal, with either the upstream or downstream fragment
292 incorporated more often. Recombination of both fragments was observed in 1/20 cases and
293 was not observed when fragments were lengthened by non-homologous DNA (Figure S3A-H,
294 Table S7).

295 Non-contiguous recombination was detected in all conditions, although the frequency of
296 occurrence increased with length of homology (Figure 6C, Figure S3A-H, Table S7). It was
297 observed both when the template was given as a single long fragment, and as two overlapping
298 fragments.

299 Overall, the results suggested that the formation of multiple D-loops from either a single
300 template molecule, or from multiple molecules, could lead to non-contiguous recombination
301 events (Figure 6D). However, given that longer homology and total molecule length increased
302 the efficiency of transformation partly by reducing strand degradation (Morrison and Guild,
303 1972), there may have been more opportunity for complex recombination patterns to occur
304 when longer fragments were taken up.

305



306

Figure 6: **Fragmented transformation of donor DNA predominantly originates from non-contiguous recombination events.** Three fragments were PCR amplified from the 11A strain *pbp2x* locus, one long 4.2 Kb fragment, and two overlapping short fragments (up 2837 bp, and down 2260 bp). In addition, the same fragments were ligated to stretches of non-homologous DNA, taking all total lengths to 6 Kb (long-NH, NH-up, and down-NH). **(A)** Efficiencies of transformation of fragments into D39V with selection on AMX (0.015 µg/mL).

Statistically significant differences in efficiency compared to the long fragment was determined by *t* test with BH correction for multiple testing. NS not significant, * indicates $p < 0.05$. **(B)** Alignment of the homologous region of template fragments used in the experiment. Donor SNPs are shown in blue, while bases matching the recipient (D39V) are shown in grey. **(C)** Alignment of 4.2 Kb locus from colonies transformed with different fragments. Bases shared with donor but not with recipient are shown in blue. **(D)** Schematic showing hypothesized molecular mechanisms for non-contiguous and contiguous recombination. As shown in panel B, both events can take place, but based on transformation efficiencies, the process on the left with long donor DNA is more efficient.

307

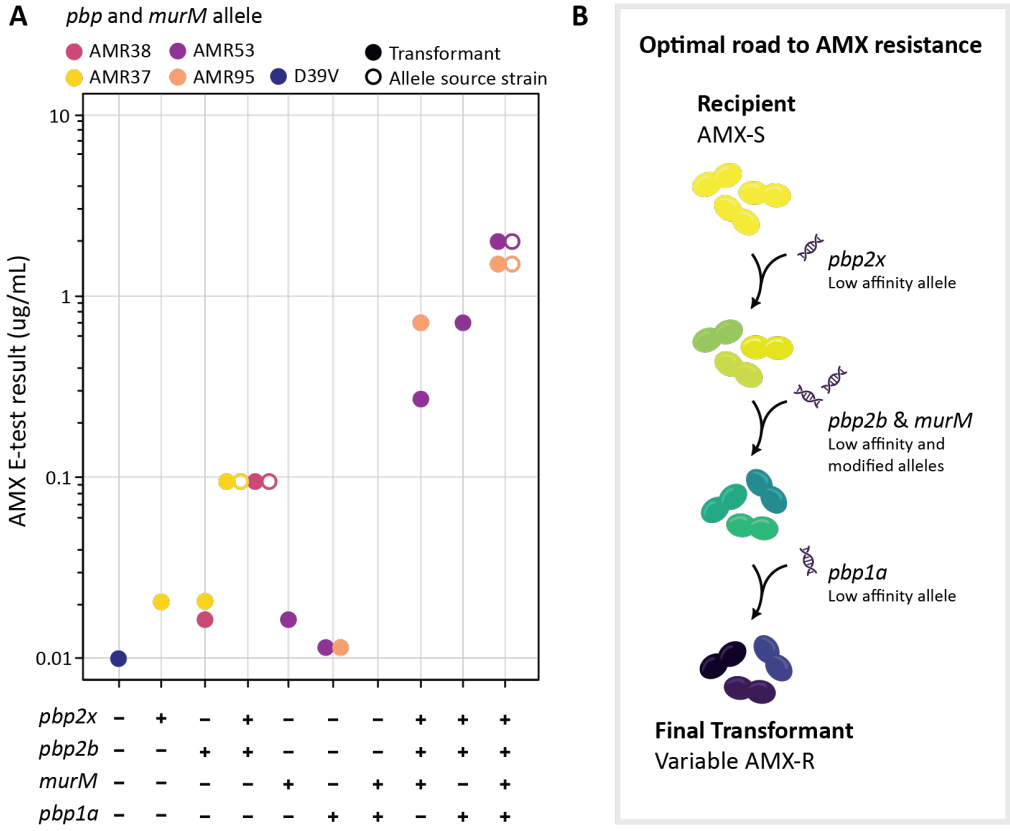
308 ***AMX resistant* pbp *and* murM *alleles are sufficient to explain AMX MICs***

309 Substitutions in all four proteins (Pbp1a, Pbp2b, Pbp2x and MurM) were required to reach an
310 MIC close or equal to that of the donor strain (Figure S4, and Table S5), but the order of uptake
311 varied both stochastically, and depending on the recipient genome. The correlation between
312 allele uptake and MIC increase was consistent in the D39V background. First-generation
313 strains had mutations in *pbp2b* and *pbp2x*, the two essential PBPs for cell growth (Berg et al.,
314 2013). In the second-generation, all strains acquired mutations in *murM*, and those with MICs
315 higher than 1 µg/mL also had mutations in *pbp1a*. In TIGR4, *pbp2x* alleles were also acquired
316 in the first-generation, however *pbp2b* mutations only appeared after the second round of
317 transformation. Mutations in *murM* were recombined in EB7 in the first round of
318 transformation, then were inherited vertically in that lineage, but were not acquired by EB3
319 or any descendant strains (Figure 5C). Although *murM* mutations were not present in all
320 strains with *pbp1a* modifications, strains with recombination events at all four loci had higher
321 MICs (Figure 2C). This indicated that *murM* mutations may be less important for successful
322 integration and selection for *pbp1a* mutations but are key for expression of the AMX resistant
323 phenotype of the donor (Filipe et al., 2002; Filipe and Tomasz, 2000; Garcia-Bustos and
324 Tomasz, 1990; Smith and Klugman, 2001).

325 An experiment was designed to test whether the order of allele uptake in D39V recombinant
326 strains was crucial for AMX resistance development, and if PBP and MurM substitutions alone
327 were sufficient to explain AMX MICs. Alleles of *pbp2x* and *pbp2b* from AMR37 (*pbp2x37*,
328 *pbp2b37*) and AMR38 (*pbp2x38*, *pbp2b38*) were transformed into D39V and selected on AMX.
329 Strains carrying either a recombinant *pbp2x* or *pbp2b* allele had E-test MICs up to 0.064
330 µg/mL. However, when both *pbp* alleles were integrated ($D39V^{2x38\text{-}2b38}$), the E-test result
331 equalled that of the recombinant allele donor (Figure 7), suggesting that mutations outside

332 these two loci were not involved in the observed reduction in AMX susceptibility in the first-
333 generation strains.
334 To assess whether mutated *pbp2x* and *pbp2b* were critical for *murM* mutation uptake, D39V
335 was transformed with *murM* from both AMR53 (*murM53*) and AMR95 (*murM95*) and MICs
336 determined. Surprisingly, we were able to select for colonies with integration of the AMR53
337 *murM* allele (D39V^{M53}) resulting in an E-test result of 0.023 µg/mL, despite this protein not
338 being a direct target of AMX (Figure 7 and Figure 8A.B). The genome was sequenced, and no
339 spontaneous point mutations were identified in previously described resistance determinants
340 outside *murM* (Table S8). When transformed into the D39V$^{2x38\text{-}2b38}$, the E-test MICs reached
341 0.5 µg/mL and 0.38 µg/mL for AMR95 and AMR53 *murM* alleles (D39V$^{2x38\text{-}2b38\text{-}M95}$, D39V$^{2x38\text{-}}$
342 $^{2b38\text{-}M53}$), respectively. AMX susceptibility decreased substantially when all three genes were
343 mutated, however failed to reach the final MIC of the recombinant allele donor, indicating
344 the necessity of another resistance determinant (Figure 7).
345 *pbp1a* alleles from AMR53 and AMR95 (*pbp1a53*, *pbp1a95*) could not be successfully selected
346 on AMX when transformed into D39V, either alone or in conjunction with a *murM* allele from
347 the same donor. Indeed, Pbp1a is not a key target of AMX, and this result was not unexpected.
348 We instead transformed a construct where a kanamycin resistance cassette was cloned
349 behind *pbp1a*, with 1 Kb of homology up- and downstream. Growth in liquid culture of the
350 resulting strains at AMX 0.01 µg/mL was similar to D39V (Figure 8A.B). Interestingly,
351 successful integration occurred for both *pbp1a53* and *pbp1a95* when transformed into
352 D39V$^{2x38\text{-}2b38}$ (D39V$^{2x38\text{-}2b38\text{-}1a53}$, D39V$^{2x38\text{-}2b38\text{-}1a95}$), suggesting that a mutated *murM* allele was
353 not required for mutations in *pbp1a* to cause a selectable increase in AMX MIC. However,
354 substitutions in Pbp2x, Pbp2b, and Pbp1a alone were insufficient to replicate the E-test
355 results of the allele source strains (Figure 7).
356 We therefore transformed *murM* and *pbp1a* alleles from AMR53 and AMR95 into D39V$^{2x38\text{-}}$
357 ^{2b38} simultaneously, and found that selection was possible on higher concentrations of AMX
358 (1 µg/mL) than for either gene alone (0.2 µg/mL). The resulting strains (D39V$^{2x38\text{-}2b38\text{-}M53\text{-}1a53}$,
359 D39V$^{2x38\text{-}2b38\text{-}M95\text{-}1a95}$) had AMX resistance phenotypes matching those of AMR53 and AMR95,
360 with E-test results of 2 µg/mL and 1.5 µg/mL respectively (Figure 7).
361 The strains containing all four genes from AMR53 and AMR95 were whole genome sequenced
362 to check for spontaneous point mutations outside these loci. One non-synonymous mutation
363 in *glnA* (glutamine synthase type I) was found in both strains. Downregulation of *glnA* has

364 been linked to the pneumococcal penicillin stress response (el Khoury et al., 2017), but no
365 differences in AMX survival were observed in these strains compared to the allele donors
366 lacking the SNP.
367

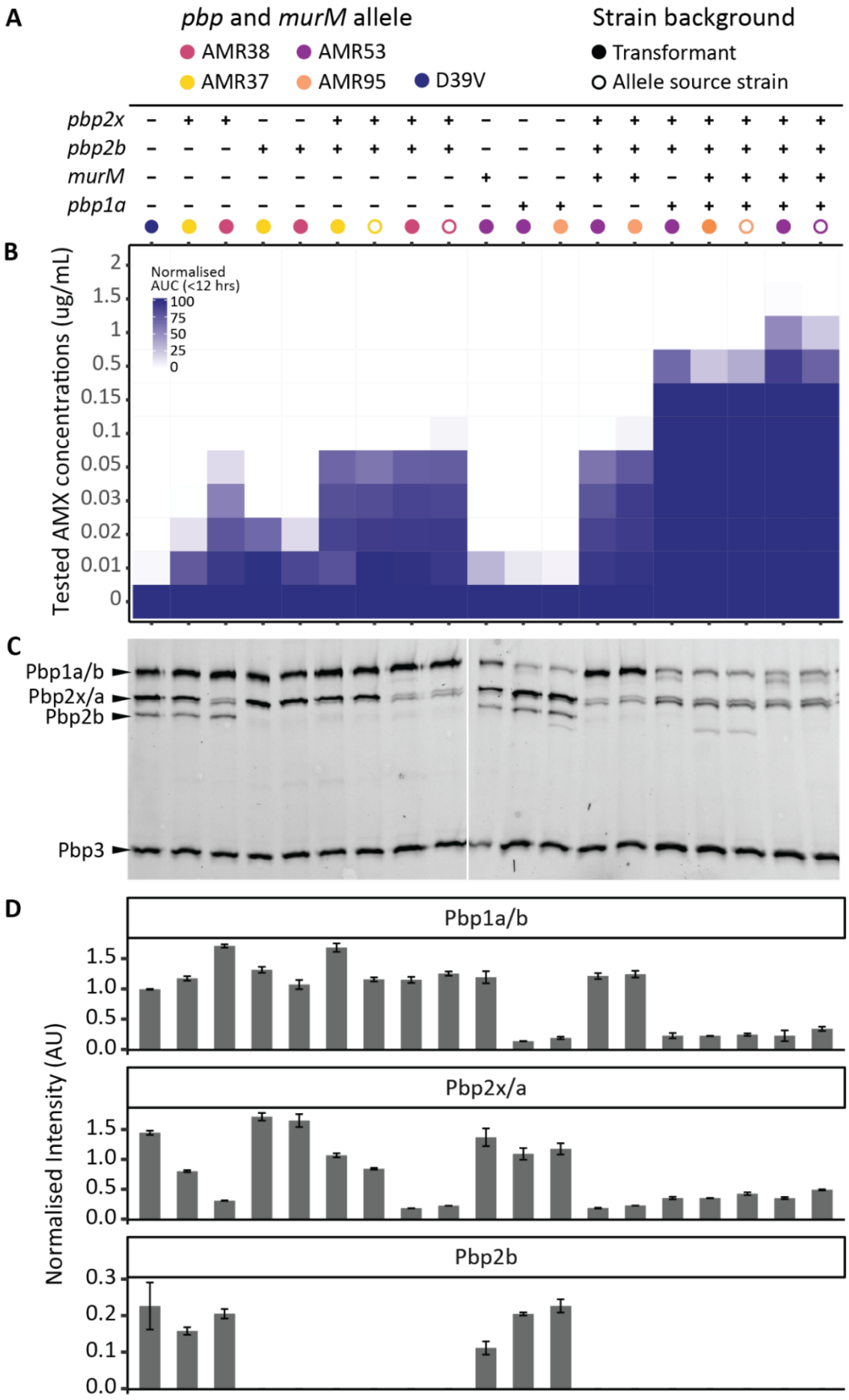


368

Figure 7: Uptake of mosaic *pbp1a*, *pbp2b*, *pbp2x* and *murM* is required and sufficient for clinically relevant amoxicillin resistance in *S. pneumoniae*. (A) AMX MICs determined by E-test for D39V transformant strains (solid circles) where *pbp* and *murM* alleles from different source strains (empty circles) were transformed in combination. Color shows allele source. (B) Schematic overview of the optimal road to AMX resistance development through horizontal gene transfer. Low affinity *pbp2x* was uptaken in the first round of transformation in both datasets, whereas donor *murM* and *pbp2b* alleles were acquired in the first or second round, depending on recipient and lineage. Low affinity *pbp1a* was recombined in the second round, and only found in strains with AMX MICs above 0.75 µg/mL. In principle uptake of low affinity and modified alleles could be achieved in a single (unlikely) or double step.

369 To quantify the affinities of resistant PBP alleles to β-lactam antibiotics, Bocillin-FL binding
370 patterns were determined for allele swap strains and the source strains (Figure 8C.D). This
371 confirmed that the recombinant PBPs had reduced penicillin binding affinities. For Pbp2b,
372 both the AMR37 and AMR38 alleles resulted in almost complete loss of binding, and a band
373 could not be quantified (Figure 8D). A large reduction in affinity was observed for Pbp2x38,
374 larger than for Pbp2x37, which correlated to a small difference in growth at 0.1 µg/mL AMX
375 between the two source strains (Figure 8A.B). This could be explained by 13 amino acid

376 substitutions including S389L and T338A which were not present in the AMR37 allele but have
377 previously been linked to reduced susceptibility to β-lactams (Kocaoglu et al., 2015; Kosowska
378 et al., 2004).
379 Perhaps most interesting were the two low affinity Pbp1a alleles tested. In all strains
380 containing *pbp1a95* (AMR95 derived), reduced band intensity was observed in the
381 Pbp1a/Pbp1b size range, however, a faster migrating species was also detected. The *pbp1a95*
382 allele contained a spontaneous point mutation (not present in the 11A donor) causing a
383 premature stop codon. This resulted in the loss of 65 amino acids (of a total of 720) and put
384 the predicted size at 73.1 kDa instead of 79.6 kDa. Reduced intensity at the expected
385 Pbp1a/Pbp1b size range was also observed for the AMR53 allele (Pbp1a53) however, an
386 additional Bocillin-FL-tagged species which migrated slightly faster than expected was also
387 observed. The predicted protein length of Pbp1a53 is identical to D39V and the donor, with
388 no premature stop codons identified. The additional band may indicate some degradation of
389 this protein species, although the TP domain must still be available for Bocillin-FL binding.
390 Together, these experiments demonstrate that transfer of the *pbp2X*, *pbp2B*, *pbp1a* and
391 *murM* alleles of 11A into strain D39V is sufficient to explain the observed AMX resistance in
392 transformants obtained using chromosomal DNA.



393

Figure 8: Recombinant PBP proteins visualized by PBP-bocillin binding affinities. Transformed AMX resistant strains produce mosaic PBPs with reduced affinity for Bocillin-FL. Each column shows the phenotypic

characterization of one *pbp* and *murM* allele swap strain. **(A)** Table showing the presence (+) and absence (-) of alleles from different source strains (colour). Allele source strains are also included and are indicated by the unfilled circles. **(B)** Survival in different AMX concentrations quantified by determining the area under the growth curve (AUC) after 12 hours growth in AMX normalized by AUC in the absence of AMX. **(C)** Bocillin-FL binding affinity gel. **(D)** Intensity of bands corresponding to Pbp1a/b, Pbp2x/a, and Pbp2b, normalized by the intensity of the Pbp3 band of the same lane, as this allele remained unmodified in all strains.

394 ***High AMX resistance in the D39V comes at a significant fitness loss***

395 Of the 16 recombinant strains sequenced, only EB31 and EB28 from the TIGR4 collection had
396 the same AMX MIC as the donor strain, and a third round of transformation with gDNA into
397 AMR53 failed to increase the MIC. The only region of mutations that EB28 and EB31 had in
398 common that was not present in the highest MIC AMR strains, was a block of 10 mutations in
399 the 590-641 region of Pbp2b (Figure 9C, Table S5). To test whether these substitutions were
400 sufficient to explain the MIC difference, the *pbp2b* allele from EB31 (*pbp2b31*) was
401 transformed into AMR53 and AMR95. The AMX E-test scores for the resulting strains were
402 both 4-6 µg/mL. However, growth curves revealed a significant fitness loss, with a mean
403 maximum specific growth rate of 0.11 µ/h for AMR95 decreasing almost 100-fold to 0.0036
404 µ/h for AMR95^{2b31}, which was exacerbated by the presence of AMX (Figure 9B). In contrast,
405 AMR53^{2b31} grew almost 3-fold slower than AMR53 (AMR53 0.12 µ/h, AMR53^{2b31} 0.046 µ/h,
406 Figure 9B), but showed a lytic phenotype in stationary phase. Interestingly, this strain outgrew
407 the 11A donor in sublethal AMX concentrations (Figure 9A). The difference in growth rate
408 between AMR95^{2b31} and AMR53^{2b31} was likely explained by the truncated Pbp1a95 protein.
409 Significant modifications to the active sites of essential proteins often need to be
410 compensated for, which may be impacted by the sub-optimal performance of other proteins
411 in the system. Although the TP and TG domains of Pbp1a95 were intact, it is possible that the
412 loss of 65 amino acids from the C-terminal resulted in an overall reduction in function,
413 unrelated to its AMX binding affinity. This correlates with previous studies showing that
414 compensation by substitutions in Pbp1A are required to mitigate fitness loss resulting from
415 modifications to the Pbp2b active site (Albarracín Orio et al., 2011). Together, these
416 experiments show that D39V is capable of reaching high AMX resistance, but that the absence
417 of compensation by Pbp1A substitutions may result in significant fitness costs.

418

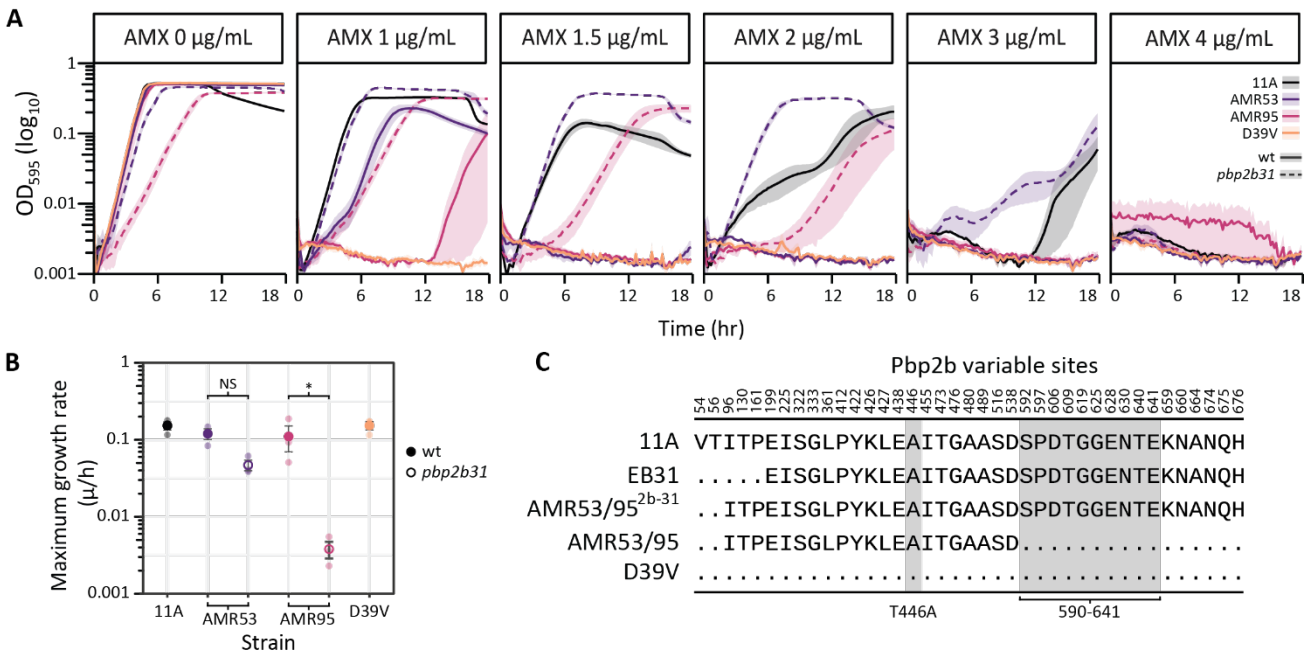


Figure 9: Ten substitutions in Pbp2b 590-641 region essential for high AMX resistance. (A) Growth of AMR53^{2b31} and AMR95^{2b31} (dotted lines) in different AMX concentrations. 11A, D39V, AMR53, and AMR95 (solid lines) are shown as controls. **(B)** Dot plot of maximum specific growth rates (µ/h) calculated from the AMX 0 µg/mL condition in triplicate (transparent circles). The mean for each strain is shown by an opaque circle, solid for wild-type (wt) strains and empty for AMR53^{2b31} and AMR95^{2b31}. Color denotes strain background. Significance was determined by one-way ANOVA with a Tukey correction for multiple testing. **(C)** Amino acid sites in pbp2b which differ from D39V for strains 11A, EB31, AMR53 (identical to AMR95), and AMR53^{2b31} (the result of transforming AMR53 or AMR95 with the EB31 pbp2b allele). Dots indicate the amino acid matches D39V, and key substitutions, T446A and a block of 10 between 590-641 (Cafini et al., 2006; du Plessis et al., 2002; Kosowska et al., 2004), are highlighted in grey.

Discussion

Natural competence and homologous recombination in *S. pneumoniae* allow for enormous genomic plasticity which contributes to the rapid development of vaccine-escape and antibiotic resistant isolates. In one round of transformation, more than 29 recombination events can occur in a single cell (Cowley et al., 2018), while events up to 50 Kb in length have been identified in clinical isolates (Wyres et al., 2013), allowing for significant remodelling at multiple loci simultaneously. Indeed, here we show a two-step sequential transformation experiment with up to 20 individual recombination events around the genome and lengths up to 13 Kb, resulting in an average donor genome transfer per step of 0.38%. This allows for huge genomic diversity, which, when placed under selective pressure, can lead to the rapid spread of antibiotic resistance-associated genotypes (Chewapreecha et al., 2014).

β-lactam resistance in the pneumococcus relies on DNA movement between closely related strains and species within the same niche, allowing tens of substitutions to accumulate in the

433 target proteins (Hakenbeck et al., 2012). It has previously been shown through *in vitro*
434 transformation experiments into a laboratory strain that the main determinants of β-lactam
435 resistance are modified alleles of *pbp2x*, *pbp2b*, *pbp1a*, and *murM* (Fani et al., 2014; Smith
436 and Klugman, 2001; Trzciński et al., 2004). For the front-line antibiotic AMX, resistance has
437 also been previously associated with mutations in these alleles although the relative
438 importance of the changes in each locus is not well understood (Cafini et al., 2006; Chesnel et
439 al., 2005; du Plessis et al., 2002; Garcia-Bustos and Tomasz, 1990; Kosowska et al., 2004; Smith
440 and Klugman, 2001). We confirmed the requirement for mutations in all four genes and
441 showed that subtle differences between alleles could cause noticeable changes in both AMX-
442 dependent fitness and the penicillin-binding affinity of PBPs.

443 In addition, a specific order of mutation uptake for penicillin resistance has been
444 demonstrated in both *Staphylococcus aureus* and the pneumococcus (Demerec, 1945;
445 Hotchkiss, 1951; Shockley and Hotchkiss, 1970), and for AMX this order was found to be
446 *pbp2x*, *pbp2b*, then *pb1a*, with the role of *murM* unable to be confirmed in that experimental
447 system (du Plessis et al., 2002). Here, we confirm that substitutions in Pbp2x are a critical first
448 step for AMX resistance, perhaps reflecting the protein's high affinity for the antibiotic
449 (Kocaoglu et al., 2015)(Figure 7C), and that the T338A substitution is required for resistance
450 to fully develop. We also show that mutations in both *pbp2b* and *murM* are essential
451 intermediary steps, and that a modified *pbp1a* allele is crucial for high-level AMX resistance.
452 It is however important to note that these experiments were performed using high
453 concentrations of naked DNA as a model, while in the nasopharynx genetic exchange likely
454 occurs via fratricide or from small quantities of free environmental DNA (Eldholm et al., 2009;
455 Johnsborg et al., 2008). It may be that lower donor DNA concentrations cause slower
456 resistance allele acquisition, and result in lower fitness costs. In addition, previous studies
457 have strongly implied the existence of multiple mechanisms of AMX resistance in *S.*
458 *pneumoniae*, associated with different *pbp2b* and *murM* alleles (Cafini et al., 2006; Càmara et
459 al., 2018; du Plessis et al., 2002), and that by using a single donor strain we were limited to
460 the study of only one. However, we confirmed that a previously described block of ten
461 substitutions in the 590-641 region of the Pbp2b TP domain (Cafini et al., 2006; du Plessis et
462 al., 2002; Kosowska et al., 2004) was essential to reconstruct the AMX MIC of the serotype
463 11A ST6521 donor strain in D39V and TIGR4. Perhaps most interestingly, we also show that

464 the loss of 65 amino acids from the C-terminus of Pbp1a results in a significant fitness cost in
465 the presence of these additional Pbp2b substitutions, highlighting the importance of
466 interactions and fitness loss compensation between essential cell-wall synthesis proteins
467 (Albarracín Orio et al., 2011).

468 In addition to these definite experiments demonstrating the evolutionary route to clinical
469 AMX resistance, we show that non-contiguous recombination could result from either a
470 single, or multiple, donor DNA molecules. It is easy to imagine a long ssDNA-RecA presynaptic
471 filament interacting with the chromosome at multiple points in the 3-dimensional cellular
472 space (Forget and Kowalczykowski, 2012; Yang et al., 2020)(Figure 1 model), which then
473 initiates the formation of several D-loops simultaneously as homologous DNA-DNA
474 interactions stabilize (Bell and Kowalczykowski, 2016; Hsieh et al., 1992; Mazin and
475 Kowalczykowski, 1996). On the other hand, SsbB coated ssDNA sequestered in the cytoplasm
476 could provide a pool from which multiple donor molecules homologous for the same
477 chromosomal region could be sourced (Attaiech et al., 2011). However, we also observed that
478 both transformation efficiency and non-contiguous event occurrence increased with length
479 of homology and total length of the donor molecule, suggesting that single donor molecule
480 events may simply be more likely. This concurs with the previous observation of transforming
481 DNA concentration-independent differences in recombination event density, which would
482 not have been expected under a multiple donor molecule hypothesis (Croucher et al., 2012).
483 Recombination structures such as these have also been observed in *H. influenzae*, with the
484 authors noting that localized clustering of long tracts of donor sequence appeared too close
485 to have occurred by chance (Mell et al., 2011). They hypothesized incoming DNA degradation
486 via cytosolic or translocation endonucleases resulted in these clusters. In addition, mismatch
487 repair via the Hex system could in principle be responsible for the reversion of small numbers
488 of point mutations between non-contiguous recombination segments, although the system
489 was likely saturated in our model (Humbert et al., 1995). To untangle the different possible
490 mechanisms of non-contiguous recombination from other effects on transformation
491 efficiency, testing mixtures of different donor fragment lengths would be necessary.
492 Altering the active sites of essential proteins in cell division is risky and requires a delicate
493 balance between antibiotic avoidance and fitness loss. Accumulation of sufficient mutations
494 in the appropriate order is aided both by the simultaneous recombination of multiple donor

495 fragments distally around the genome, and by the occurrence of complex non-contiguous
496 events. The experimental transformation experiments combined with whole genome
497 sequencing as performed here provide valuable insights into viable and non-viable
498 evolutionary paths toward AMX resistance development in the pneumococcus. The strong co-
499 dependence among different proteins and their alleles creates significant challenges for
500 complete understanding of resistance mechanisms, and emphasizes the importance of
501 combining data from phenotypic, molecular, genomic, and genome-wide association studies.

502 # Methods

503 ***Bacterial strains, growth conditions, and antibiotics***

504 *S. pneumoniae* strains are shown in Table 1 and Table S1. Bacteria were cultured in liquid C+Y
505 medium with no shaking at 37°C. C+Y media was adapted from Adams and Roe (Martin et al.,
506 1995).

507 For transformation, *S. pneumoniae* was grown in C+Y (pH 6.8) at 37°C to OD_{595} 0.1.
508 Competence was induced by addition of 100 ng/ml synthetic CSP-1 (D39V) or CSP-2 (TIGR4)
509 and 12 min incubation at 37°C. DNA uptake occurred during 20 min at 30°C, followed by
510 dilution and recovery for 1.5 hr. Transformants were selected by plating inside Columbia agar
511 supplemented with 3% defibrinated sheep blood (CBA, Thermo Scientific) containing the
512 appropriate antibiotic and incubated at 37°C in 5% CO_2 overnight. Successful transformants
513 were confirmed by PCR and Sanger sequencing (Microsynth).

514 Strains were stocked at OD_{595} 0.3 in C + Y with 15% glycerol at -80°C.

515 AMX (Sigma Aldrich) powder was dissolved initially in 100% DMSO. This solution was diluted
516 in molecular grade water with 4% DMSO to two concentrations, 100 µg/mL and 1 mg/mL, and
517 stored at -80°C.

518

519

Table 1: Key strains used in this study. The full collection of recombinant strains and their AMX MIC can be found in Table S1

Strain Name	Strain ID	Genotype and relevant characteristics	AMX MIC (µg/mL)	reference
D39V	VL1	Serotype 2 laboratory strain, AMX susceptible	0.001	Slager et al., 2018
TIGR4	VL2177	Serotype 4 laboratory strain, AMX susceptible	0.016	Tettelin et al., 2001
11A	SN75752 VL1313	Serotype 11A ST6521 clinical isolate, AMX resistant	4	German National Reference Center for Streptococci
AMR37	VL4481	D39V; recombinant strain, generation 1	0.094	This study
AMR38	VL4482	D39V; recombinant strain, generation 1	0.094	This study
AMR53	VL3196	D39V; recombinant strain, generation 2	2	This study
AMR95	VL4484	D39V; recombinant strain, generation 2	1.5	This study
EB31	VL4495	TIGR4; recombinant strain, generation 2	4	This study
D39V^{2x37}	VL4496	D39V; *pbp2x37*	0.023	This study
D39V^{2x38}	VL4302	D39V; *pbp2x38*	0.064	This study
D39V^{2b37}	VL4497	D39V; *pbp2b37*	0.023	This study
D39V^{2b38}	VL4498	D39V; *pbp2b38*	0.023	This study
D39V$^{2x37\text{-}2b37}$	VL4499	D39V; *pbp2x37*; *pbp2b37*	0.094	This study
D39V$^{2x38\text{-}2b38}$	VL3895	D39V; *pbp2x38*; *pbp2b38*	0.094	This study
D39V^{M53}	VL3960	D39V; *murM53*	0.023	This study
D39V^{1a53}	VL4159	D39V; *pbp1a53*	0.016	This study
D39V^{1a95}	VL4160	D39V; *pbp1a95*	0.016	This study
D39V$^{2x38\text{-}2b38\text{-}M53}$	VL3961	D39V; *pbp2x38*; *pbp2b38*; *murM53*	0.38	This study
D39V$^{2x38\text{-}2b38\text{-}M95}$	VL3962	D39V; *pbp2x38*; *pbp2b38*; *murM95*	0.5	This study
D39V$^{2x38\text{-}2b38\text{-}1a53}$	VL3896	D39V; *pbp2x38*; *pbp2b38*; *pbp1a53*	0.75	This study
D39V$^{2x38\text{-}2b38\text{-}M53\text{-}1a53}$	VL3963	D39V; *pbp2x38*; *pbp2b38*; *murM53*; *pbp1a53*	2	This study
D39V$^{2x38\text{-}2b38\text{-}M95\text{-}1a95}$	VL4500	D39V; *pbp2x38*; *pbp2b38*; *murM95*; *pbp1a95*	1.5	This study
AMR53^{2b31}	VL4174	AMR53; *pbp2b31*	6	This study
AMR95^{2b31}	VL4184	AMR95; *pbp2b31*	6	This study

520

521 ***Minimum Inhibitory Concentrations (MICs)***

522 Bacteria were grown in C+Y broth to OD_{595} 0.1, then AMX MIC was determined using E-tests
523 (Biomeriaux, Lyon) on Mueller Hinton agar containing 5% defibrinated sheep's blood. MICs
524 were read after 18 h incubation at 37°C and 5% CO_2.

525 Strains with MICs below the limit of the E-tests (0.016 µg/mL) were tested in Mueller Hinton
526 cation-adjusted broth with 5% defibrinated sheep's blood using the broth micro-dilution
527 method (Charles et al., 2016). Briefly, bacteria were streaked onto MH2-blood plates and
528 grown overnight. Colonies were picked from plates and ressupended in PBS (pH 7.4) to
529 McFarlane standard 0.5. This inoculum was diluted 100-fold into 96-well plates containing
530 MH2-blood broth and AMX (in 2-fold concentration steps) then plated on agar plates as a
531 control for contamination and cell density. Plates were incubated for 18 hours at 37°C and 5%
532 CO_2. MIC was defined as the lowest concentration of antibiotic where visible α-haemolysis
533 occupied approximately less than 10% of the surface area of the well. Due to a high level of
534 subjectivity, MIC determination was repeated twice on different days, and MIC calls were
535 confirmed by a second person in each case.

536

537 ***Growth curves and analysis***

538 Strains were precultured in C+Y broth to OD 0.1, then diluted 100-fold into 96-well microtiter
539 plates, containing fresh C+Y and AMX. The OD_{595} was measured every 10 min for 12 hr in a
540 plate reader (Infinite F200, Tecan) with incubation at 37°C. All strains were tested in triplicate
541 Area under the curve was determined using TecanExtracter (Dénéréaz, unpublished). The
542 AUC of each replicate was divided by that of the no antibiotic control for each strain, to
543 normalize for strain-dependent differences in growth. The mean and standard error of the
544 mean (SEM) of the normalized AUC was determined for three biological replicates.

545 Statistical significance of maximum specific growth rate (µ/h) was determined using one-way
546 ANOVA with a Tukey correction for multiple testing.

547

548 ***Bocillin-FL binding assay***

549 Bocillin-FL is a fluorescently labelled penicillin molecule which can be used to visualise and
550 quantify PBP-penicillin binding affinity *in vitro*. Strains were grown to OD 0.2 in 4 mL C+Y,
551 washed and resuspended in phosphate buffered saline (PBS), then incubated for 30 min at
552 37C in 5 µg/mL bocillin-FL (Invitrogen). Cells were washed in ice cold PBS, then lysed by

553 sonication (10 x 1s pulses at 30% power). The cell lysate was spun down and supernatant
554 removed to collect the membrane fraction, which was homogenized by sonication (1s pulse
555 at 10%). Total protein concentration was determined by nanodrop, then samples were diluted
556 in NuPAGE LDS sample buffer (Life Technologies) with 10 mM DTT and incubated at 95°C for
557 5 min. Samples were loaded into two 15-well Novex™ WedgeWell™ 10% Tris-Glycine 1.0 mm
558 Mini Protein Gels (Invitrogen), to a total protein amount of approximately 1.8 mg in a
559 maximum volume of 30 µL. Where necessary, gels were imaged simultaneously on an
560 Amersham Typhoon (GE Healthcare) with Cy2 DIGE filter setup. Bands were quantified using
561 ImageQuant (GE Healthcare). As all strains tested had identical PBP3 proteins, the intensities
562 of the PBP1a/b, PBP2x/a, and PBP2b bands were divided by the PBP3 band intensity of the
563 corresponding sample for normalisation. Imaging and quantification were performed in
564 duplicate. Coomassie Blue staining was also performed, to confirm equal sample loading.
565
566 ***Phase-contrast time-lapse microscopy***
567 For time-lapse microscopy, cells were pre-cultured in liquid C + Y medium at 37°C and spotted
568 on a low-melting 1.2% C + Y agarose patch, with or without AMX at 1 µg/mL. Phase-contrast
569 pictures were taken every 10 min as described (de Jong et al., 2011) using a Leica DMi8 with
570 a 100x phase contrast objective.
571
572 ***Isolation of genomic DNA***
573 10 mL of OD_{595} 0.2 culture was pelleted by centrifugation, then resuspended in Nuclei Lysis
574 solution (Promega) supplemented with 0.05% SDS, 0.025% deoxycholate (DOC), and 200
575 µg/mL RNase A. This was incubated at 37°C for 20 minutes, 80°C for 5 minutes, then 37°C for
576 10 minutes. Protein Precipitation Solution (Promega) was added to the lysate, vortexed
577 vigorously, then incubated on ice for 10 minutes. The precipitated protein was pelleted by
578 centrifugation and the supernatant transferred into isopropanol to precipitate the DNA,
579 which was then collected by centrifugation. DNA was washed once in 70% ethanol before air
580 drying and resuspension in the appropriate buffer (either TE buffer or molecular grade water).
581 DNA was run on a 1% agarose gel to confirm no significant degradation had occurred, and
582 stored at 4°C.
583
584 ***Serial Transformation of* S. pneumoniae *with genomic DNA***

585 AMX susceptible recipient strains were transformed with 100 µg/mL genomic DNA isolated
586 from AMX resistant strain 11A (Table 1) as above, with the following modifications. DNA
587 uptake occurred during 30 min, followed by recovery for 2 hrs. Transformants were selected
588 by plating inside Columbia agar supplemented with 5% defibrinated sheep blood and a range
589 of AMX (0.03 µg/mL). Ten colonies were then picked at random and streaked onto CBA plates
590 containing AMX. The next day, single colonies were grown up in C+Y with AMX and stocked.
591 MICs were determined for these 10 strains, which were then subjected to another round of
592 transformation and selection on a range of AMX concentrations (0.17 - 0.5 µg/mL), and 5-10
593 single colonies were picked for each of the 10 strains transformed. These were isolated, and
594 MICs determined, as above. To test for significant differences between MICs of second-
595 generation strains in different lineages, a Fisher's Exact Test with Monte Carlo p-value
596 simulation was used.
597 To control for spontaneous mutations arising during transformation and recombination, cells
598 transformed with either no DNA, or with D39V genomic DNA, were also plated in CBA with
599 AMX (0.03). Neither resulted in visible colonies.

600

601 ***Whole genome sequencing and quality control***

602 Genomic DNA from recombinant strains from each recipient background, as well as 11A and
603 TIGR4 were Illumina sequenced (PE150, Novogene and Eurofins Genomics), with a mean
604 coverage of 600-fold. Reads were trimmed using Trimmomatic (Bolger et al., 2014), then
605 assembled using SPAdes (Nurk et al., 2013). Assembly quality was assessed with Quast
606 (Gurevich et al., 2013), and read depth was determined in samtools (Li et al., 2009). Previously
607 published D39V Illumina reads were filtered as above and used where necessary (GEO
608 accessions GSE54199 and GSE69729) (Slager et al., 2014). Allelic exchange strains ($D39V^{M53}$,
609 $D39V^{2x38\text{-}2b38\text{-}M53\text{-}1a53}$, $D39V^{1a53}$, $D39V^{2x38\text{-}2b38\text{-}M95\text{-}1a95}$) were treated in the same way, then
610 variants were called using bbmap (Bushnell, 2013), Freebayes (Garrison and Marth, 2012),
611 and vcftools (Danecek et al., 2011), in order to check for random mutations arising elsewhere
612 in the genome.
613 11A, TIGR4, and a subset of recombinant strains were sent for PacBio sequencing (Sequel I
614 and II, Lausanne Genomic Technologies Facility). Demultiplexing and quality control were
615 performed using SMRTlink (PacBio), then reads shorter than 1000 bp were removed with

616 filtlong (Wick, 2017). Recombinant strain reads were mapped to the recipient reference
617 genome using NGMLR and structural variants (SVs) called with sniffles with default settings,
618 except with read mapping quality > 40 and read depth > 50 (Sedlazeck et al., 2018). Imprecise
619 SVs were then filtered out. DNA methylation detection was performed using motifMaker
620 (Clark et al., 2012).
621 Raw read files are deposited with NCBI Sequence Read Archive under BioProject
622 PRJNA789167 (Individual SRA accession numbers for recombinant strains can be found in
623 Table S4).
624
625 ***Hybrid assemblies of reference genomes***
626 As PacBio sequencing depth was greater than 200x, genomes were assembled in a long-read
627 only manner using Trycycler (Wick et al., 2021), followed by short-read polishing.
628 Briefly, filtered reads were subsampled 12 times and *de novo* assembled in four different
629 assemblers, flye (Kolmogorov et al., 2019), raven (Wang et al., 2018), wtdbg2 (Ruan and Li,
630 2020), and miniasm plus minipolish (Li, 2018, 2016). Some assemblies were removed during
631 the Trycycler cluster and reconcile steps, leaving the final consensus sequence to be built from
632 seven to eight assemblies. The consensus assembly was subjected to rounds of polishing with
633 PacBio reads using pbmm2 and pbgcpp (PacBio Toolkit), followed by one round of short-read
634 polishing with Pilon (Walker et al., 2014).
635 Genomes were annotated in Prokka (Seemann, 2014) using Barrnap and Aragorn (Laslett and
636 Canback, 2004; Seemann, 2019) and are deposited at NCBI under BioProject PRJNA789167
637 (accession numbers currently in process).
638
639 ***Recombination detection with NGS***
640 Competitive mapping of filtered illumina reads onto a pseudogenome containing both
641 recipient and donor chromosomes was used for crude visualization of recombination loci
642 (bbmap) (Matthey et al., 2019). In order to characterize these loci more accurately, filtered
643 reads of recombinant and recipient strains were mapped onto the 11A hybrid genome
644 assembly using smalt, then variants were called with bcftools (Li, 2011). SNPs were filtered as
645 described previously (Cowley et al. 2018), and used to build a consensus alignment of donor,
646 recipient, and recombinant strains. The consensus alignment was run through a python script
647 designed to identify SNPs between recombinant and donor which are not shared with the

recipient genome. Once identified, the match was extended until the donor and recombinant sequences are no longer identical, then the start, end, and length of detected recombination events are extracted (Cowley et al 2018).

Identifying non-contiguous recombination events

Many recombination events were located very close together, with sometimes only a few recipient allele SNPs separating the two, suggesting a single donor DNA molecule as the source. To determine if events were significantly close together, a bootstrapping approach was applied (Croucher et al., 2012). Firstly, the distances between one recombination event and each of the other recombination events in a single genome were calculated. The circular chromosome gives two values, and the shortest distance between any two events was added to the test population. This was calculated for every sequenced strain in the D39V and TIGR4 recombinants, and for second-generation strains any prior events carried through from the first-generation were removed from the dataset. The shortest recombination distance within a strain became the test value (d_{test}), and then the test population was sampled (with replacement).

The distribution of a bootstrapped sample (same size as the test population) was used to test the hypothesis that d_{test} was significantly shorter than expected under the null hypothesis that recombination events are located at random around the chromosome. A Holm correction was applied to account for multiple testing, based on the number of recombination events in the strain. Bootstrap tests were performed one thousand times for each d_{test} value and considered significantly short if the null hypothesis was rejected at least 95% of the time. If a d_{test} value was significant, then the distance between the next shortest pairing within the strain became d_{test}, and was tested in the same fashion. This continued until a d_{test} where the null hypothesis could not be rejected.

When recombination events were considered significantly close together, they were linked together into non-contiguous recombination events.

Confirming the origins of non-contiguous recombination events in vivo

A 4.25 Kb region of the 11A genome spanning *pbp2x* was used to test the origin of non-contiguous recombination. This region included an 846 bp stretch with SNPs conferring a small decrease in AMX susceptibility sufficient for selection. One long fragment spanning the

680 entire region was amplified with OVL5375 and OVL5378. Two shorter fragments were then
681 amplified, an upstream fragment (Up) with OVL5379 and OVL5375 (2.84 Kb) and a
682 downstream fragment (Down) with OVL5376 and OVL5378 (2.26 Kb), which overlapped at the
683 SNPs determined to be necessary for selection (Table S2, S5). The three donor DNA templates
684 used in the experiment and their SNPs are shown in Figure 6C (Table S6).
685 D39V was transformed as above, with equimolar quantities of each fragment separately, as
686 well as a combination of the two short fragments, and a negative control with no DNA.
687 Transformants were selected on AMX (0.015 µg/mL), and the transformation efficiency was
688 determined by dividing by the total viable count. Individual colonies were restreaked on AMX,
689 then single colonies were picked and the 4.25 Kb region of interest was amplified and Sanger
690 sequenced (Microsynth, Table S2 oligos). Nucleotide sequences were aligned in Mega (Kumar
691 et al., 2018).
692 As molecule length affects transformation efficiency (Keller et al., 2019; Kurushima et al.,
693 2020; Lee et al., 1998), an additional experiment was performed, where the amplified
694 fragments were all extended to 6 Kb by the addition of non-homologous DNA amplified from
695 *E. coli* and ligated to one end using Golden Gate assembly with BsmBI and T4 DNA ligase (NEB,
696 Vazyme). All oligos and the corresponding templates used in the cloning for these
697 experiments are outlined in Table S6, and a comprehensive oligo list can be found in Table S2.
698 The assemblies were then amplified to ensure high quality of transforming fragments,
699 resulting in three fragments of equal size, with differing lengths of homologous regions to the
700 *pbp2x* locus. These were then transformed, amplified, and Sanger sequenced as above, and
701 the transformation efficiency determined.
702
703 ***Gene and protein alignments***
704 Nucleotide sequences of *pbp* and *murM* genes were translated using Emboss:TransSeq
705 (Madeira et al., 2019). Alignments were performed with Muscle (Edgar, 2004). Phylogenetic
706 trees were produced using concatenated protein alignments of Pbp2x, Pbp2b, Pbp1a, and
707 MurM as input for FastTree (default settings), and visualised in ITOL (Letunic and Bork, 2021;
708 Price et al., 2009).
709
710 ***Amplification and transformation of* pbp *and* murM *loci for allelic exchange experiments***

711 Genes for *pbps* were amplified with Phanta Max Super-Fidelity DNA Polymerase (Vazyme)
712 and primers ordered from Sigma (Table S2). Equimolar quantities of PCR products were used
713 for transformation. To avoid transforming mutations in *murN* alongside those in *murM*, alleles
714 at this locus were cloned into a construct with D39V flanking sequences. The upstream and
715 downstream fragments were amplified from D39V with OVL5540/OVL5779 and
716 OVL5780/OVL5541 respectively, and the *murM* allele from AMR53 or AMR95 with
717 OVL5777/OVL5778. Fragments were then assembled using BsmBI and T4 DNA ligase.
718 *pbp2x* and *pbp2b* alleles from AMR37 and AMR38 were transformed into D39V as above,
719 selected on 0.012 µg/mL, 0.015 µg/mL, 0.02 µg/mL AMX, and confirmed by Sanger
720 sequencing. D39V$^{2x38\text{-}2b38}$ (Table 1) was then used as the recipient for *murM* and *pbp1a* alleles
721 from AMR53 and AMR95, with selection on 0.2 µg/mL, 0.5 µg/mL and 1 µg/mL AMX.
722

723 **Acknowledgements**

724 We would like to thank Mark van der Linden at the German National Reference Center for
725 Streptococci for the 11A (SN75752) clinical isolate, and the Lausanne Genomic Technologies
726 Facility for their valuable work in PacBio sequencing and data quality control. We would also
727 like to thank Arnau Domenech Pena, Alice Wallef, Vincent de Bakker, Julien Denereaz, and
728 the rest of the Veening Lab for their help, contributions, and feedback.
729 P.G. was supported by the University of Lausanne Faculty of Biology and Medicine PhD
730 fellowship. Work in the Veening lab is supported by the Swiss National Science Foundation
731 (SNSF) (project grant 310030_192517), SNSF JPIAMR grant (40AR40_185533), SNSF NCCR
732 'AntiResist' (51NF40_180541) and ERC consolidator grant 771534-PneumoCaTChER.
733

734 **References**

735 Albarracín Orio AG, Piñas GE, Cortes PR, Cian MB, Echenique J. 2011. Compensatory evolution of pbp mutations
736 restores the fitness cost imposed by β-lactam resistance in streptococcus pneumoniae. *PLoS Pathogens*
737 **7**:e1002000. doi:10.1371/journal.ppat.1002000

738 Alloing G, Martin B, Granadel C, Claverys JP. 1998. Development of competence in Streptococcus pneumoniae:
739 Pheromone autoinduction and control of quorum sensing by the oligopeptide permease. *Molecular*
740 *Microbiology* **29**:75–83. doi:10.1046/J.1365-2958.1998.00904.X

741 Arnold BJ, Gutmann MU, Grad YH, Sheppard SK, Corander J, Lipsitch M, Hanage WP. 2018. Weak epistasis may
742 drive adaptation in recombining bacteria. *Genetics* **208**:1247–1260. doi:10.1534/genetics.117.300662

743 Attaiech L, Olivier A, Mortier-Barrière I, Soulet AL, Granadel C, Martin B, Polard P, Claverys JP. 2011. Role of the
744 single-stranded dna-binding protein Ssbb in pneumococcal transformation: Maintenance of a reservoir for
745 genetic plasticity. *PLoS Genetics* **7**:e1002156. doi:10.1371/journal.pgen.1002156

746 Bell JC, Kowalczykowski SC. 2016. RecA: Regulation and Mechanism of a Molecular Search Engine. *Trends in*
747 *Biochemical Sciences*. doi:10.1016/j.tibs.2016.04.002

748 Berg KH, Stamsås GA, Straume D, Håvarstein LS. 2013. Effects of low PBP2b levels on cell morphology and
749 peptidoglycan composition in streptococcus pneumoniae R6. *Journal of Bacteriology* **195**:4342–4354.
750 doi:10.1128/JB.00184-13

751 Bergé MJ, Kamgoué A, Martin B, Polard P, Campo N, Claverys J-P. 2013. Midcell Recruitment of the DNA Uptake
752 and Virulence Nuclease, EndA, for Pneumococcal Transformation. *PLOS Pathogens* **9**:e1003596.
753 doi:10.1371/JOURNAL.PPAT.1003596

754 Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*
755 **30**:2114–20. doi:10.1093/bioinformatics/btu170

756 Bushnell B. 2013. BBMap download | SourceForge.net. *SourceForge*. https://sourceforge.net/projects/bbmap/

757 Cafini F, del Campo R, Alou L, Sevillano D, Morosini MI, Baquero F, Prieto J, García E, Casal J, Fenoll A, de la Campa
758 AG, Bouza E, Baquero F, Soriano F, Pallarés R, Liñares J, Garau J, Lacasa JM, Latorre C, Pérez-Trallero E,
759 García de Lomas J, Fleites A. 2006. Alterations of the penicillin-binding proteins and murM alleles of clinical
760 Streptococcus pneumoniae isolates with high-level resistance to amoxicillin in Spain. *Journal of*
761 *Antimicrobial Chemotherapy* **57**:224–229. doi:10.1093/jac/dki442

762 Càmara J, Cubero M, Martín-Galiano AJ, García E, Grau I, Nielsen JB, Worning P, Tubau F, Pallarés R, Domínguez
763 MÁ, Kilian M, Liñares J, Westh H, Ardanuy C. 2018. Evolution of the β-lactam-resistant Streptococcus
764 pneumoniae PMEN3 clone over a 30 year period in Barcelona, Spain. *Journal of Antimicrobial*
765 *Chemotherapy* **73**:2941–2951. doi:10.1093/jac/dky305

766 Carrasco B, Serrano E, Sánchez H, Wyman C, Alonso JC. 2016. Chromosomal transformation in Bacillus subtilis is
767 a non-polar recombination reaction. *Nucleic Acids Research* **44**:2754–2768. doi:10.1093/nar/gkv1546

768 Cassone M, Gagne AL, Spruce LA, Seeholzer SH, Sebert ME. 2012. The HtrA protease from Streptococcus
769 pneumoniae digests both denatured proteins and the competence-stimulating peptide. *J Biol Chem*
770 **287**:38449–59. doi:10.1074/jbc.M112.391482

771 Charles MK, Berenger BM, Turnbull L, Rennie R, Fuller J. 2016. Variability of β-lactam susceptibility testing for
772 Streptococcus pneumoniae using 4 commercial test methods and broth microdilution. *Diagnostic*
773 *Microbiology and Infectious Disease* **84**:240–245. doi:10.1016/J.DIAGMICROBIO.2015.11.014

774 Chesnel L, Carapito R, Croizé J, Dideberg O, Vernet T, Zapun A. 2005. Identical penicillin-binding domains in
775 penicillin-binding proteins of Streptococcus pneumoniae clinical isolates with different levels of beta-
776 lactam resistance. *Antimicrob Agents Chemother* **49**:2895–2902. doi:10.1128/AAC.49.7.2895

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

781 Clark TA, Murray IA, Morgan RD, Kislyuk AO, Spittle KE, Boitano M, Fomenkov A, Roberts RJ, Korlach J. 2012.
782 Characterization of DNA methyltransferase specificities using single-molecule, real-time DNA sequencing.
783 *Nucleic Acids Research* **40**. doi:10.1093/nar/gkr1146

784 Claverys JP, Lacks SA. 1986. Heteroduplex deoxyribonucleic acid base mismatch repair in bacteria.
785 *Microbiological Reviews*. doi:10.1128/mmbr.50.2.133-165.1986

786 Cowley LA, Petersen FC, Junges R, Jimson D. Jimenez M, Morrison DA, Hanage WP. 2018. Evolution via
787 recombination: Cell-to-cell contact facilitates larger recombination events in Streptococcus pneumoniae.
788 *PLOS Genetics* **14**:e1007410. doi:10.1371/journal.pgen.1007410

789 Croucher NJ, Harris SR, Barquist L, Parkhill J, Bentley SD. 2012. A High-Resolution View of Genome-Wide
790 Pneumococcal Transformation. *PLoS Pathogens* **8**:e1002745. doi:10.1371/journal.ppat.1002745

791 Croucher NJ, Harris SR, Fraser C, Quail MA, Burton J, van der Linden M, McGee L, von Gottberg A, Song JH, Ko
792 KS, Pichon B, Baker S, Parry CM, Lambertsen LM, Shahinas D, Pillai DR, Mitchell TJ, Dougan G, Tomasz A,
793 Klugman KP, Parkhill J, Hanage WP, Bentley SD. 2011. Rapid pneumococcal evolution in response to clinical
794 interventions. *Science (1979)* **331**:430–434. doi:10.1126/SCIENCE.1198545

795 Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST,
796 McVean G, Durbin R. 2011. The variant call format and VCFtools. *Bioinformatics* **27**:2156–2158.
797 doi:10.1093/bioinformatics/btr330

798 Danilowicz C, Yang D, Kelley C, Prévost C, Prévost P, Prentiss M. 2015. The poor homology stringency in the
799 heteroduplex allows strand exchange to incorporate desirable mismatches without sacrificing recognition
800 in vivo. *Nucleic Acids Research* **43**:6473–6485. doi:10.1093/nar/gkv610

801 Demerec M. 1945. Production of Staphylococcus strains resistant to various concentrations of penicillin.
802 *Proceedings of the National Academy of Sciences* **31**:16–24.

803 Desai B v, Morrison DA. 2007. Transformation in Streptococcus pneumoniae: Formation of eclipse complex in a
804 coiA mutant implicates CoiA in genetic recombination. *Molecular Microbiology* **63**:1107–1117.
805 doi:10.1111/j.1365-2958.2006.05558.x

806 Doern G v, Brueggemann A, Holley HP, Rauch AM. 1996. Antimicrobial resistance of Streptococcus pneumoniae
807 recovered from outpatients in the United States during the winter months of 1994 to 1995: results of a
808 30-center national surveillance study. *Antimicrob Agents Chemother* **40**:1208–13.
809 doi:10.1128/AAC.40.5.1208

810 Doit C, Loukil C, Fitoussi F, Geslin P, Bingen E. 1999. Emergence in France of Multiple Clones of
811 ClinicalStreptococcus pneumoniae Isolates with High-Level Resistance to Amoxicillin. *Antimicrobial Agents*
812 *and Chemotherapy* **43**:1480–1483. doi:10.1128/AAC.43.6.1480

813 Donati C, Hiller NL, Tettelin H, Muzzi A, Croucher NJ, Angiuoli S v., Oggioni M, Dunning Hotopp JC, Hu FZ, Riley
814 DR, Covacci A, Mitchell TJ, Bentley SD, Kilian M, Ehrlich GD, Rappuoli R, Moxon ER, Masignani V. 2010.
815 Structure and dynamics of the pan-genome of Streptococcus pneumoniae and closely related species.
816 *Genome Biology* **11**:R107. doi:10.1186/gb-2010-11-10-r107

817 Dowson CG, Coffey TJ, Kell C, Whiley RA. 1993. Evolution of penicillin resistance in Streptococcus pneumoniae;
818 the role of Streptococcus mitis in the formation of a low affinity PBP2B in S. pneumoniae. *Molecular*
819 *Microbiology* **9**:635–643. doi:10.1111/j.1365-2958.1993.tb01723.x

820 du Plessis M, Bingen E, Klugman KP, Plessis M du, Bingen E, Klugman KP, du Plessis M, Bingen E, Klugman KP.
821 2002. Analysis of Penicillin-Binding Protein Genes of Clinical Isolates of Streptococcus pneumoniae with
822 Reduced Susceptibility to Amoxicillin. *Antimicrobial Agents and Chemotherapy* **46**:2349–2357.
823 doi:10.1128/AAC.46.8.2349

824 Edgar RC. 2004. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids*
825 *Research* **32**:1792–1797. doi:10.1093/nar/gkh340

826 el Khoury JY, Boucher N, Bergeron MG, Leprohon P, Ouellette M. 2017. Penicillin induces alterations in glutamine
827 metabolism in Streptococcus pneumoniae. *Scientific Reports* **7**. doi:10.1038/s41598-017-15035-y

828 Eldholm V, Johnsborg O, Haugen K, Ohnstad HS, Havastein LS. 2009. Fratricide in Streptococcus pneumoniae:
829 Contributions and role of the cell wall hydrolases CbpD, LytA and LytC. *Microbiology (N Y)* **155**:2223–2234.
830 doi:10.1099/MIC.0.026328-0

831 Fani F, Leprohon P, Zhanel GG, Bergeron MG, Ouellette M. 2014. Genomic analyses of DNA transformation and
832 penicillin resistance in Streptococcus pneumoniae clinical isolates. *Antimicrobial Agents and*
833 *Chemotherapy* **58**:1397–1403. doi:10.1128/AAC.01311-13

834 Feng Z, Li J, Zhang JR, Zhang X. 2014. Qdnamod: A statistical model-based tool to reveal intercellular
835 heterogeneity of DNA modification from SMRT sequencing data. *Nucleic Acids Research* **42**:13488–13499.
836 doi:10.1093/nar/gku1097

837 Figueiredo TA, Sobral RG, Ludovice AM, de Almeida JMF, Bui NK, Vollmer W, de Lencastre H, Tomasz A. 2012.
838 Identification of Genetic Determinants and Enzymes Involved with the Amidation of Glutamic Acid
839 Residues in the Peptidoglycan of Staphylococcus aureus. *PLOS Pathogens* **8**:e1002508.
840 doi:10.1371/JOURNAL.PPAT.1002508

841 Filipe SR, Severina E, Tomasz A. 2002. The murMN operon: A functional link between antibiotic resistance and
842 antibiotic tolerance in Streptococcus pneumoniae. *Proceedings of the National Academy of Sciences*
843 **99**:1550–1555. doi:10.1073/PNAS.032671699

844 Filipe SR, Tomasz A. 2000. Inhibition of the expression of penicillin resistance in Streptococcus pneumoniae by
845 inactivation of cell wall muropeptide branching genes. *Proceedings of the National Academy of Sciences*
846 **97**:4891–4896. doi:https://doi.org/10.1073/pnas.080067697

847 Forget AL, Kowalczykowski SC. 2012. Single-molecule imaging of DNA pairing by RecA reveals a three-
848 dimensional homology search. *Nature 2012 482:7385* **482**:423–427. doi:10.1038/nature10782

849 Garcia-Bustos J, Tomasz A. 1990. A biological price of antibiotic resistance: major changes in the peptidoglycan
850 structure of penicillin-resistant pneumococci. *Proc Natl Acad Sci U S A* **87**:5415–9.
851 doi:10.1073/pnas.87.14.5415

852 Garrison E, Marth G. 2012. Haplotype-based variant detection from short-read sequencing. *arXiv.org*.

853 Gasc AM, Sicard AM, Claverys JP. 1989. Repair of single- and multiple-substitution mismatches during
854 recombination in Streptococcus pneumoniae. *Genetics* **121**:29–36. doi:10.1093/genetics/121.1.29

855 Grebe T, Paik J, Hakenbeck R. 1997. A novel resistance mechanism against β-lactams in Streptococcus
856 pneumoniae involves CpoA, a putative glycosyltransferase. *Journal of Bacteriology* **179**:3342–3349.
857 doi:10.1128/jb.179.10.3342-3349.1997

858 Guenzi E, Gasc A-M, Sicard MA, Hakenbeck R. 1994. A two-component signal-transducing system is involved in
859 competence and penicillin susceptibility in laboratory mutants of Streptococcus pneumoniae. *Molecular*
860 *Microbiology* **12**:505–515. doi:10.1111/j.1365-2958.1994.tb01038.x

861 Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: quality assessment tool for genome assemblies.
862 *Bioinformatics* **29**:1072–1075. doi:10.1093/bioinformatics/btt086

863 Gurney T, Fox MS. 1968. Physical and genetic hybrids formed in bacterial transformation. *Journal of Molecular*
864 *Biology* **32**:83–100. doi:10.1016/0022-2836(68)90147-2

865 Hakenbeck R, Brückner R, Denapaite D, Maurer P. 2012. Molecular mechanisms of β-lactam resistance in
866 Streptococcus pneumoniae. *Future Microbiology*. doi:10.2217/fmb.12.2

867 Hakenbeck R, Nig AK, Kern I, van der Linden M, Keck W, Billot-Klein D, Legrand R, Schoot B, Gutmann L. 1998.
868 Acquisition of Five High-M r Penicillin-Binding Protein Variants during Transfer of High-Level beta-Lactam
869 Resistance from Streptococcus mitis to Streptococcus pneumoniae. *Journal of Bacteriology* **180**:1831–
870 1840.

871 Hiller NL, Ahmed A, Powell E, Martin DP, Eutsey R, Earl J, Janto B, Boissy RJ, Hogg J, Barbadora K, Sampath R,
872 Lonergan S, Post JC, Hu FZ, Ehrlich GD. 2010. Generation of Genic Diversity among Streptococcus
873 pneumoniae Strains via Horizontal Gene Transfer during a Chronic Polyclonal Pediatric Infection. *PLOS*
874 *Pathogens* **6**:e1001108. doi:10.1371/JOURNAL.PPAT.1001108

875 Hotchkiss RD. 1951. Transfer of penicillin resistance in pneumococci by the desoxyribonucleate derived from
876 resistant cultures. *Cold Spring Harb Symp Quant Biol* **16**:457–461. doi:10.1101/SQB.1951.016.01.032

877 Hsieh P, Camerini-Otero CS, Camerini-Otero RD. 1992. The synapsis event in the homologous pairing of DNAs:
878 RecA recognizes and pairs less than one helical repeat of DNA. *Proc Natl Acad Sci U S A* **89**:6492–6496.
879 doi:10.1073/pnas.89.14.6492

880 Humbert O, Prudhomme M, Hakenbeck R, Dowson CG, Claverys JP. 1995. Homeologous recombination and
881 mismatch repair during transformation in Streptococcus pneumoniae: Saturation of the Hex mismatch
882 repair system. *Proc Natl Acad Sci U S A* **92**:9052–9056. doi:10.1073/pnas.92.20.9052

883 Johnsborg O, Eldholm V, Bjørnstad ML, Håvarstein LS. 2008. A predatory mechanism dramatically increases the
884 efficiency of lateral gene transfer in *Streptococcus pneumoniae* and related commensal species. *Molecular*
885 *Microbiology* **69**:245–253. doi:10.1111/j.1365-2958.2008.06288.x

886 Kalizang'oma A, Chaguza C, Gori A, Davison C, Beleza S, Antonio M, Beall B, Goldblatt D, Kwambana-Adams B,
887 Bentley SD, Heyderman RS. 2021. Streptococcus pneumoniae serotypes that frequently colonise the
888 human nasopharynx are common recipients of penicillin-binding protein gene fragments from
889 Streptococcus mitis. *Microbial Genomics* **7**:000622. doi:10.1099/mgen.0.000622

890 Keller LE, Rueff A-SS, Kurushima J, Veening J-WW. 2019. Three new integration vectors and fluorescent proteins
891 for use in the opportunistic human pathogen Streptococcus pneumoniae. *Genes (Basel)* **10**:394.
892 doi:10.3390/genes10050394

893 Kocaoglu O, Tsui H-CCT, Winkler ME, Carlson EE. 2015. Profiling of β-lactam selectivity for penicillin-binding
894 proteins in Streptococcus pneumoniae D39. *Antimicrobial Agents and Chemotherapy* **59**:3548–3555.
895 doi:10.1128/AAC.05142-14

896 Kolmogorov M, Yuan J, Lin Y, Pevzner PA. 2019. Assembly of long, error-prone reads using repeat graphs. *Nature*
897 *Biotechnology* **37**:540–546. doi:10.1038/s41587-019-0072-8

898 Kosowska K, Jacobs MR, Bajaksouzian S, Koeth L, Appelbaum PC. 2004. Alterations of Penicillin-Binding Proteins
899 1A, 2X, and 2B in Streptococcus pneumoniae Isolates for Which Amoxicillin MICs Are Higher than Penicillin
900 MICs **48**. doi:10.1128/AAC.48.10.4020-4022.2004

901 Kulick S, Moccia C, Didelot X, Falush D, Kraft C, Suerbaum S. 2008. Mosaic DNA imports with interspersions of
902 recipient sequence after natural transformation of Helicobacter pylori. *PLoS ONE* **3**:e3797.
903 doi:10.1371/journal.pone.0003797

904 Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: Molecular evolutionary genetics analysis across
905 computing platforms. *Molecular Biology and Evolution* **35**:1547–1549. doi:10.1093/molbev/msy096

906 Kurushima J, Campo N, van Raaphorst R, Cerckel G, Polard P, Veening JW. 2020. Unbiased homeologous
907 recombination during pneumococcal transformation allows for multiple chromosomal integration events.
908 *Elife* **9**:1–70. doi:10.7554/ELIFE.58771

909 Lacks S, Greenberg B. 1976. Single-strand breakage on binding of DNA to cells in the genetic transformation of
910 Diplococcus pneumoniae. *Journal of Molecular Biology* **101**:255–275. doi:10.1016/0022-2836(76)90376-4

911 Lacks S, Neuberger M. 1975. Membrane location of a deoxyribonuclease implicated in the genetic
912 transformation of Diplococcus pneumoniae. *Journal of Bacteriology* **124**:1321–1329.
913 doi:10.1128/jb.124.3.1321-1329.1975

914 Laslett D, Canback B. 2004. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide
915 sequences. *Nucleic Acids Research* **32**:11–16. doi:10.1093/nar/gkh152

916 Laurenceau R, Péhau-Arnaudet G, Baconnais S, Gault J, Malosse C, Dujeancourt A, Campo N, Chamot-Rooke J, le
917 Cam E, Claverys JP, Fronzes R. 2013. A Type IV Pilus Mediates DNA Binding during Natural Transformation
918 in Streptococcus pneumoniae. *PLoS Pathogens* **9**:e1003473. doi:10.1371/journal.ppat.1003473

919 Lee MS, Seok C, Morrison DA. 1998. Insertion-duplication mutagenesis in Streptococcus pneumoniae: Targeting
920 fragment length is a critical parameter in use as a random insertion tool. *Applied and Environmental*
921 *Microbiology* **64**:4796–4802. doi:10.1128/aem.64.12.4796-4802.1998

922 Letunic I, Bork P. 2021. Interactive tree of life (iTOL) v5: An online tool for phylogenetic tree display and
923 annotation. *Nucleic Acids Research* **49**:W293–W296. doi:10.1093/nar/gkab301

924 Li H. 2018. Minimap2: Pairwise alignment for nucleotide sequences. *Bioinformatics* **34**:3094–3100.
925 doi:10.1093/bioinformatics/bty191

926 Li H. 2016. Minimap and miniasm: Fast mapping and de novo assembly for noisy long sequences. *Bioinformatics*
927 **32**:2103–2110. doi:10.1093/bioinformatics/btw152

928 Li H. 2011. A statistical framework for SNP calling, mutation discovery, association mapping and population
929 genetical parameter estimation from sequencing data. *Bioinformatics* **27**:2987–2993.
930 doi:10.1093/bioinformatics/btr509

931 Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project
932 Data Processing Subgroup. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*
933 **25**:2078–2079. doi:10.1093/bioinformatics/btp352

934 Lisboa J, Andreani J, Sanchez D, Boudes M, Collinet B, Liger D, Tilbeurgh H van, Guérois R, Quevillon-Cheruel S.
935 2014. Molecular determinants of the DprA-RecA interaction for nucleation on ssDNA. *Nucleic Acids*
936 *Research* **42**:7395–7408. doi:10.1093/nar/gku349

937 Liu X, Li JW, Feng Z, Luo Y, Veening JW, Zhang JR. 2017. Transcriptional repressor PtvR regulates phenotypic
938 tolerance to vancomycin in Streptococcus pneumoniae. *Journal of Bacteriology* **199**:54–71.
939 doi:10.1128/JB.00054-17

940 Madeira F, Park Y mi, Lee J, Buso N, Gur T, Madhusoodanan N, Basutkar P, Tivey ARN, Potter SC, Finn RD, Lopez
941 R. 2019. The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Research* **47**:W636–
942 W641. doi:10.1093/nar/gkz268

943 Manso AS, Chai MH, Atack JM, Furi L, de Ste Croix M, Haigh R, Trappetti C, Ogunniyi AD, Shewell LK, Boitano M,
944 Clark TA, Korlach J, Blades M, Mirkes E, Gorban AN, Paton JC, Jennings MP, Oggioni MR. 2014. A random
945 six-phase switch regulates pneumococcal virulence via global epigenetic changes. *Nature Communications*
946 **5**:1–9. doi:10.1038/ncomms6055

947 Marie L, Rapisarda C, Morales V, Bergé M, Perry T, Soulet AL, Gruget C, Remaut H, Fronzes R, Polard P. 2017.
948 Bacterial RadA is a DnaB-Type helicase interacting with RecA to promote bidirectional D-loop extension.
949 *Nature Communications* **8**:1–14. doi:10.1038/ncomms15638

950 Martin B, Garcia P, Castanié M -P, Claverys J -P. 1995. The recA gene of Streptococcus pneumoniae is part of a
951 competence-induced operon and controls lysogenic induction. *Molecular Microbiology* **15**:367–379.
952 doi:10.1111/j.1365-2958.1995.tb02250.x

953 Martin B, Soulet AL, Mirouze N, Prudhomme M, Mortier-Barrière I, Granadel C, Noirot-Gros MF, Noirot P, Polard
954 P, Claverys JP. 2013. ComE/ComE∼P interplay dictates activation or extinction status of pneumococcal X-
955 state (competence). *Molecular Microbiology* **87**:394–411. doi:10.1111/mmi.12104

956 Matthey N, Stutzmann S, Stoudmann C, Guex N, Iseli C, Blokesch M. 2019. Neighbor predation linked to natural
957 competence fosters the transfer of large genomic regions in vibrio cholerae. *Elife* **8**.
958 doi:10.7554/eLife.48212

959 Mazin A v, Kowalczykowski SC. 1996. The specificity of the secondary DNA binding site of RecA protein defines
960 its role in DNA strand exchange. *Proc Natl Acad Sci U S A* **93**:10673–10678. doi:10.1073/pnas.93.20.10673

961 Mell JC, Shumilina S, Hall IM, Redfield RJ. 2011. Transformation of natural genetic variation into *Haemophilus*
962 *influenzae* genomes. *PLoS Pathogens* **7**:e1002151. doi:10.1371/journal.ppat.1002151

963 Mirouze N, Bergé MA, Soulet AL, Mortier-Barrière I, Quentin Y, Fichant G, Granadel C, Noirot-Gros MF, Noirot P,
964 Polard P, Martin B, Claverys JP. 2013. Direct involvement of DprA, the transformation-dedicated RecA

965 loader, in the shut-off of pneumococcal competence. *Proc Natl Acad Sci U S A* **110**.
966 doi:10.1073/pnas.1219868110

967 Morrison DA, Guild WR. 1973. Breakage prior to entry of donor DNA in Pneumococcus transformation. *BBA*
968 *Section Nucleic Acids And Protein Synthesis* **299**:545–556. doi:10.1016/0005-2787(73)90226-8

969 Morrison DA, Guild WR. 1972. Transformation and deoxyribonucleic acid size: extent of degradation on entry
970 varies with size of donor. *J Bacteriol* **112**:1157–1168. doi:10.1128/jb.112.3.1157-1168.1972

971 Morrison DA, Mortier-Barrière I, Attaiech L, Claverys JP. 2007. Identification of the major protein component of
972 the pneumococcal eclipse complex. *Journal of Bacteriology* **189**:6497–6500. doi:10.1128/JB.00687-07

973 Mortier-Barriere I, Humbert O, Martin B, Prudhomme M, Claverys JP. 1997. Control of recombination rate during
974 transformation of Streptococcus pneumoniae: An overview. *Microbial Drug Resistance* **3**:233–242.
975 doi:10.1089/mdr.1997.3.233

976 Mortier-Barrière I, Velten M, Dupaigne P, Mirouze N, Piétrement O, McGovern S, Fichant G, Martin B, Noirot P,
977 le Cam E, Polard P, Claverys JP. 2007. A Key Presynaptic Role in Transformation for a Widespread Bacterial
978 Protein: DprA Conveys Incoming ssDNA to RecA. *Cell* **130**:824–836. doi:10.1016/j.cell.2007.07.038

979 Münch D, Roemer T, Lee SH, Engeser M, Sahl HG, Schneider T. 2012. Identification and in vitro Analysis of the
980 GatD/MurT Enzyme-Complex Catalyzing Lipid II Amidation in Staphylococcus aureus. *PLOS Pathogens*
981 **8**:e1002509. doi:10.1371/JOURNAL.PPAT.1002509

982 Nurk S, Bankevich A, Antipov D, Gurevich A, Korobeynikov A, Lapidus A, Prjibelsky A, Pyshkin A, Sirotkin A,
983 Sirotkin Y, Stepanauskas R, McLean J, Lasken R, Clingenpeel SR, Woyke T, Tesler G, Alekseyev MA, Pevzner
984 PA. 2013. Assembling genomes and mini-metagenomes from highly chimeric reads.In: Deng M., Jiang R.,
985 Sun F., Zhang X. (Eds) Research in Computational Molecular Biology. Springer, Berlin, Heidelberg. pp. 158–
986 170. doi:10.1007/978-3-642-37195-0_13

987 O'Brien KL, Wolfson LJ, Watt JP, Henkle E, Deloria-Knoll M, McCall N, Lee E, Mulholland K, Levine OS, Cherian T.
988 2009. Burden of disease caused by Streptococcus pneumoniae in children younger than 5 years: global
989 estimates. *The Lancet* **374**:893–902. doi:10.1016/S0140-6736(09)61204-6

990 Piotrowski A, Luo P, Morrison DA. 2009. Competence for Genetic Transformation in Streptococcus pneumoniae:
991 Termination of Activity of the Alternative Sigma Factor ComX Is Independent of Proteolysis of ComX and
992 ComW. *Journal of Bacteriology* **191**:3359. doi:10.1128/JB.01750-08

993 Price MN, Dehal PS, Arkin AP. 2009. Fasttree: Computing large minimum evolution trees with profiles instead of
994 a distance matrix. *Molecular Biology and Evolution* **26**:1641–1650. doi:10.1093/molbev/msp077

995 Quevillon-Cheruel S, Campo N, Mirouze N, Mortier-Barrière I, Brooks MA, Boudes M, Durand D, Soulet AL, Lisboa
996 J, Noirot P, Martin B, van Tilbeurgh H, Noirot-Gros MF, Claverys JP, Polard P. 2012. Structure-function
997 analysis of pneumococcal DprA protein reveals that dimerization is crucial for loading RecA recombinase
998 onto DNA during transformation. *Proc Natl Acad Sci U S A* **109**:E2466–E2475.
999 doi:10.1073/pnas.1205638109

1000 Rosenthal AL, Lacks SA. 1980. Complex structure of the membrane nuclease of Streptococcus pneumoniae
1001 revealed by two-dimensional electrophoresis. *Journal of Molecular Biology* **141**:133–146.
1002 doi:10.1016/0022-2836(80)90381-2

1003 Ruan J, Li H. 2020. Fast and accurate long-read assembly with wtdbg2. *Nature Methods* **17**:155–158.
1004 doi:10.1038/s41592-019-0669-3

1005 Salvadori G, Junges R, Morrison DA, Petersen FC. 2019. Competence in Streptococcus pneumoniae and close
1006 commensal relatives: Mechanisms and implications. *Frontiers in Cellular and Infection Microbiology*.
1007 doi:10.3389/fcimb.2019.00094

1008 Scheffers D-J, Pinho MG. 2005. Bacterial Cell Wall Synthesis: New Insights from Localization Studies.
1009 *Microbiology and Molecular Biology Reviews* **69**:585–607. doi:10.1128/mmbr.69.4.585-607.2005

1010 Sedlazeck FJ, Rescheneder P, Smolka M, Fang H, Nattestad M, von Haeseler A, Schatz MC. 2018. Accurate
1011 detection of complex structural variations using single-molecule sequencing. *Nature Methods* **15**:461–468.
1012 doi:10.1038/s41592-018-0001-7

1013 Seemann T. 2019. GitHub - tseemann/barrnap: Bacterial ribosomal RNA predictor. *GitHub*.
1014 https://github.com/tseemann/barrnap

1015 Seemann T. 2014. Prokka: Rapid prokaryotic genome annotation. *Bioinformatics* **30**:2068–2069.
1016 doi:10.1093/bioinformatics/btu153

1017 Shak JR, Vidal JE, Klugman KP. 2013. Influence of bacterial interactions on pneumococcal colonization of the
1018 nasopharynx. *Trends Microbiol* **21**:129–35. doi:10.1016/j.tim.2012.11.005

1019 Shockley TE, Hotchkiss RD. 1970. Stepwise introduction of transformable penicillin resistance in Pneumococcus.
1020 *Genetics* **64**:397–408. doi:10.1093/genetics/64.3-4.397

1021 Sibold C, Henrichsen J, König A, Martin C, Chalkley L, Hakenbeck R. 1994. Mosaic pbpX genes of major clones of
1022 penicillin-resistant Streptococcus pneumoniae have evolved from pbpX genes of a penicillin-sensitive
1023 Streptococcus oralis. *Molecular Microbiology* **12**:1013–1023. doi:10.1111/j.1365-2958.1994.tb01089.x

1024 Slager J, Aprianto R, Veening JW. 2019. Refining the pneumococcal competence regulon by RNA sequencing.
1025 *Journal of Bacteriology* **201**:e00780-18. doi:10.1128/JB.00780-18

1026 Slager J, Aprianto R, Veening JW. 2018. Deep genome annotation of the opportunistic human pathogen
1027 Streptococcus pneumoniae D39. *Nucleic Acids Research* **46**:9971–9989. doi:10.1093/nar/gky725

1028 Slager J, Kjos M, Attaiech L, Veening JW. 2014. Antibiotic-induced replication stress triggers bacterial
1029 competence by increasing gene dosage near the origin. *Cell* **157**:395–406. doi:10.1016/j.cell.2014.01.068

1030 Smith AM, Klugman KP. 2001. Alterations in MurM, a cell wall muropeptide branching enzyme, increase high-
1031 level penicillin and cephalosporin resistance in Streptococcus pneumoniae. *Antimicrobial Agents and*
1032 *Chemotherapy* **45**:2393–2396. doi:10.1128/AAC.45.8.2393-2396.2001

1033 Stanhope MJ, Lefébure T, Walsh SL, Becker JA, Lang P, Pavinski Bitar PD, Miller LA, Italia MJ, Amrine-Madsen H.
1034 2008. Positive selection in penicillin-binding proteins 1a, 2b, and 2x from Streptococcus pneumoniae and
1035 its correlation with amoxicillin resistance development. *Infection, Genetics and Evolution* **8**:331–339.
1036 doi:10.1016/J.MEEGID.2008.02.001

1037 Stanhope MJ, Walsh SL, Becker JA, Miller LA, Lefébure T, Lang P, Bitar PDP, Amrine-Madsen H. 2007. The relative
1038 frequency of intraspecific lateral gene transfer of penicillin binding proteins 1a, 2b, and 2x, in amoxicillin
1039 resistant Streptococcus pneumoniae. *Infection, Genetics and Evolution* **7**:520–534.
1040 doi:10.1016/J.MEEGID.2007.03.004

1041 Straume D, Stamsås GA, Håvarstein LS. 2015. Natural transformation and genome evolution in Streptococcus
1042 pneumoniae. *Infection, Genetics and Evolution* **33**:371–380. doi:10.1016/j.meegid.2014.10.020

1043 Torres R, Serrano E, Alonso JC. 2019. Bacillus subtilis RecA interacts with and loads RadA/Sms to unwind
1044 recombination intermediates during natural chromosomal transformation. *Nucleic Acids Research*
1045 **47**:9198–9215. doi:10.1093/nar/gkz647

1046 Trzciński K, Thompson CM, Lipsitch M. 2004. Single-step capsular transformation and acquisition of penicillin
1047 resistance in Streptococcus pneumoniae. *Journal of Bacteriology* **186**:3447–3452.
1048 doi:10.1128/JB.186.11.3447-3452.2004

1049 Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo CA, Zeng Q, Wortman J, Young SK, Earl
1050 AM. 2014. Pilon: An integrated tool for comprehensive microbial variant detection and genome assembly
1051 improvement. *PLoS ONE* **9**:e112963. doi:10.1371/journal.pone.0112963

1052 Wang H, Marcišauskas S, Sánchez BJ, Domenzain I, Hermansson D, Agren R, Nielsen J, Kerkhoven EJ. 2018. RAVEN
1053 2.0: A versatile toolbox for metabolic network reconstruction and a case study on Streptomyces coelicolor.
1054 *PLoS Computational Biology* **14**:e1006541. doi:10.1371/journal.pcbi.1006541

1055 Ween O, Gaustad P, Håvarstein LS. 1999. Identification of DNA binding sites for ComE, a key regulator of natural
1056 competence in Streptococcus pneumoniae. *Molecular Microbiology* **33**:817–827. doi:10.1046/j.1365-
1057 2958.1999.01528.x

1058 Weng L, Piotrowski A, Morrison DA. 2013. Exit from Competence for Genetic Transformation in Streptococcus
1059 pneumoniae Is Regulated at Multiple Levels. *PLoS ONE* **8**:e64197. doi:10.1371/journal.pone.0064197

1060 Wick R. 2017. rrwick/Filtlong: quality filtering tool for long reads. *GitHub*. https://github.com/rrwick/Filtlong

1061 Wick RR, Judd LM, Cerdeira LT, Hawkey J, Méric G, Vezina B, Wyres KL, Holt KE. 2021. Trycycler: consensus long-
1062 read assemblies for bacterial genomes. *Genome Biology* **22**:1–17. doi:10.1186/s13059-021-02483-z

1063 Wiktor J, Gynnå AH, Leroy P, Larsson J, Coceano G, Testa I, Elf J. 2021. RecA finds homologous DNA by reduced
1064 dimensionality search. *Nature* **597**:426–429. doi:10.1038/s41586-021-03877-6

1065 Wyres KL, Lambertsen LM, Croucher NJ, McGee L, von Gottberg A, Liñares J, Jacobs MR, Kristinsson KG, Beall
1066 BW, Klugman KP, Parkhill J, Hakenbeck R, Bentley SD, Brueggemann AB. 2013. Pneumococcal capsular
1067 switching: A historical perspective. *Journal of Infectious Diseases* **207**:439–449. doi:10.1093/infdis/jis703

1068 Wyres KL, Lambertsen LM, Croucher NJ, McGee L, von Gottberg A, Liñares J, Jacobs MR, Kristinsson KG, Beall
1069 BW, Klugman KP, Parkhill J, Hakenbeck R, Bentley SD, Brueggemann AB. 2012. The multidrug-resistant
1070 PMEN1 pneumococcus is a paradigm for genetic success. *Genome Biology* **13**:R103. doi:10.1186/gb-2012-
1071 13-11-r103

1072 Yang H, Zhou C, Dhar A, Pavletich NP. 2020. Mechanism of strand exchange from RecA–DNA synaptic and D-loop
1073 structures. *Nature* **586**:801–806. doi:10.1038/s41586-020-2820-9

1074 Zhao G, Meier TI, Kahl SD, Gee KR, Blaszczak LC. 1999. BOCILLIN FL, a sensitive and commercially available
1075 reagent for detection of penicillin-binding proteins. *Antimicrobial Agents and Chemotherapy* **43**:1124–
1076 1128. doi:10.1128/aac.43.5.1124

1077

1078 ## Supplementary Tables

1079 **Table S1**: **Recombinant strains isolated in sequential rounds of transformation with AMX resistant DNA and**
1080 **their associated MICs**.

1081 **Table S2**: **Primers used in this study.**

1082 **Table S3**: **List of all recombination events identified in transformants using SNPs.**

1083 **Table S4**: **Summary statistics for the characterisation of SNP and recombination results for sequenced**
1084 **recombinant strains**. (A) All predicted events detected when compared to parent strain genome. Potential non-
1085 contiguity not accounted for. (B) Total Dna transferred during the final round of transformation for that strain.
1086 (*) All DNA or SNPs transferred into the original recipient (D39V or TIGR4), through one or two rounds of
1087 transformation (depending on the generation). SRA accession numbers for raw Illumina and PacBio reads used
1088 for analysis are also provided.

1089 **Table S5**: **Key amino acid substitutions in PBPs and MurM of recombinant strains**. Substitutions listed have
1090 been associated with β-lactam resistance in the literature previously. Cells coloured green have the same allele
1091 as D39V, those coloured blue have the same allele as 11A. Strain AMX MICs are also indicated for reference.

1092 **Table S6**: **Cloning design for fragments used to test for single-molecule non-contiguous recombination at the**
1093 ***pbp2x-mraY* locus**.

1094 **Table S7**: **Summary statistics from colonies isolated after transformation with different length and**
1095 **combinations of donor fragments to explain the occurrence of non-contiguous recombination.**

1096 **Table S8**: **SNPs detected in D39V^{M53}** .

1097 ## Supplementary Figures

1098 **Figure S1: Fluorescence and coomassie stained images of Bocillin-FL gels**. (A) Bocillin-FL gel from Figure 3 and
1099 (B) gels from Figure 8.

1100 **Figure S2: Identification of non-contiguous recombination from sequenced recombinant strains using a**
1101 **bootstrapping method** (Croucher et al. 2012). (A) Distances between recombination events which were found
1102 to shorter than expected under a null hypothesis where events occur randomly around the chromosome. (B)
1103 Schematic showing how events were treated if they were found to be closer to each other than expected. (C)
1104 Histogram of recombination events length across all sequenced recombinant strains if non-contiguous
1105 recombination is not taken into account. (D) Histogram of recombination events length across all sequenced
1106 recombinant strains if non-contiguous recombination is accounted for.

1107 **Figure S3: Alignments of pbp2x-mraY from colonies isolated after transformation with different length and**
1108 **combinations of donor fragments to explain the occurrence of non-contiguous recombination.** Facets show
1109 colonies selected from different donor templates which correspond to Figure 6. Bases which match the recipient
1110 (D39V) are coloured grey, those which match the donor (11A) are shown in blue. (A) Long, (B) Up, (C) Down, (D)
1111 Up & Down, (E) Long-NH, (F) NH-Up, (G) Down-NH, and (H) NH-Up and Down-NH.

1112 **Figure S4: Dendrogram of strain relationships built from amino acid sequences of Pbp and murM proteins**.
1113 Alignments of Pbp1a, Pbp2b, Pbp2x, and MurM amino acid sequences were used to construct a phylogenetic
1114 tree of donor, recipient, and recombinant strains. AMX MICs are shown on the right, colour scale represents MIC
1115 and corresponds to figure 2, where yellow is an AMX MIC of 0.01 µg/mL and dark purple is an MIC of 4 µg/mL.