Evolutionary outcomes of plasmid-CRISPR conflicts in an opportunistic pathogen

Wenwen Huo[ab], Valerie J. Price[ab], Ardalan Sharifi[a], Michael Q. Zhang[a], and Kelli L. Palmer[a]*

Department of Biological Sciences, The University of Texas at Dallas, Richardson, Texas, USA[a]

[b]These authors contributed equally

*Corresponding author

E-mail: kelli.palmer@utdallas.edu (KLP)

Abstract

The persistence of antibiotic resistance plasmids in pathogens is a global health concern. Plasmid persistence results from host-plasmid co-evolution that enhances plasmid stability, where the role of CRISPR-Cas is not well understood. *Enterococcus faecalis* is an opportunistic pathogen that disseminates antibiotic resistance via conjugative plasmids. Some *E. faecalis* possess CRISPR-Cas that limit acquisition of resistance plasmids; however, transconjugants arise despite CRISPR-Cas activity. We utilized *in vitro* evolution to investigate how the conflict between CRISPR-Cas and plasmid targets is resolved. We observed a cost to maintain both the plasmid and functional CRISPR-Cas. Under antibiotic selection, heterogeneous populations with compromised CRISPR-Cas emerged, which benefited acquisition of other plasmids. Using targeted sequencing, we demonstrate RecA-independent allelic heterogeneity provides an evolutionary basis for the emergence of compromised CRISPR-Cas. Overall, antibiotic selection for plasmids targeted by CRISPR-Cas results in host mutations that stabilize plasmid maintenance and reduce the barrier to future horizontal gene transfer events.

Introduction

Plasmid-mediated dissemination of antibiotic resistance contributes to the rapid emergence of multidrug-resistant (MDR) bacterial pathogens. The emergence and persistence of antibiotic resistance is one of the most challenging problems facing health care today. The opportunistic pathogen *Enterococcus faecalis* is a paradigm for the emergence of antibiotic resistance by horizontal gene transfer (HGT) (1-5). Conjugative and mobilizable plasmids belonging to both narrow and broad host range classes disseminate antibiotic resistance genes in *E. faecalis* (6-10). Hospital-adapted strains of *E. faecalis* have an enhanced propensity to engage in HGT, making them reservoirs for antibiotic resistance genes that can be transferred to other gram-positive pathogens including *Enterococcus faecium*, staphylococci, streptococci and *Clostridium difficile* (3, 11-14).

Much research on antibiotic resistance focuses on the mechanisms that facilitate persistence of antibiotic resistance plasmids in bacterial populations when antibiotic selection is absent (15-18). Persistent plasmids are reservoirs for accessory genes that can be readily shared and utilized for rapid adaptation to new environments (15, 17, 19). Overall, this speeds up the evolution of bacteria by negating the reliance on adaptation by mutation (20, 21). There is growing consensus in the field that the persistence of resistance plasmids in the absence of antibiotic selection is due to compensatory mutations made in the host, plasmid, or both (17, 21, 22). These mutations reduce the metabolic burden of plasmid carriage, and also reduce the rate at which the plasmid can engage in horizontal transfer while stabilizing the vertical inheritance of the plasmid, ultimately leading to plasmid persistence (23).

A facet of host-plasmid co-evolution that is currently poorly understood is the impact of host-encoded CRISPR-Cas systems, which can block the acquisition of potentially beneficial mobile genetic elements (MGEs) and thus significantly impact host fitness in different environments. CRISPR-Cas systems confer programmable genome defense against plasmids and phage. CRISPR-Cas systems consist of *cas* genes and a CRISPR array composed of spacers interspersed by direct and partially palindromic repeats (24, 25). Each spacer is a molecular memory of a previously encountered MGE (26-28). *E. faecalis* possesses Type II CRISPR-Cas systems that encode the endonuclease Cas9 (4, 29). CRISPR-Cas defense is afforded in three stages, adaptation, expression and interference, that ultimately result in sequence-specific cleavage of MGEs by a *cas*-encoded endonuclease. A Cas protein complex recognizes a protospacer from a newly encountered MGE in a PAM (Protospacer Adjacent Motif)-dependent manner, after which the protospacer is incorporated into the leader end of the CRISPR array (30, 31). During expression, the CRISPR array is transcribed and processed by Cas9, RNase III, and tracrRNA (trans-activating crRNA) generating a mature crRNA (32, 33). Each mature crRNA is bound by Cas9 and tracrRNA to form an active targeting complex. When the bacterial host is invaded by a MGE that has complementarity to a crRNA, the active targeting complex recognizes the target in a PAM-dependent manner and creates a double-stranded DNA break, which prevents MGE invasion (34-36).

We reported in previous studies that Type II CRISPR-Cas systems reduce plasmid dissemination in *E. faecalis* colony biofilms by 80-fold (37). Although CRISPR-Cas had a significant impact on plasmid transfer, we still observed a high number (10^5) transconjugants (37). This observation suggests that unique interactions occur under these mating conditions that allow plasmids to escape genome defense. Moreover, in these transconjugants a conflict is established between a CRISPR-Cas system and one

of its targets. These transconjugants present a unique opportunity to study the role of CRISPR-Cas systems in plasmid-host interactions. In this study, we used a combination of *in vitro* evolution and deep sequencing analysis to investigate how *E. faecalis* resolves conflicts between CRISPR-Cas and antibiotic resistance plasmids, and the role that antibiotic selection plays in this process. We conclude that antibiotic-driven PRP maintenance in *E. faecalis* can lead to compromised genome defense and enhanced susceptibility to other MGEs, ultimately transforming these strains into reservoirs for antibiotic resistance and other virulence traits. These findings demonstrate that antibiotics can alter pathogen evolution by accelerating the host-adaptation process that results in antibiotic resistance plasmid persistence.

Results

Design of *in vitro* evolution assay to study CRISPR-Cas and plasmid dynamics. *E. faecalis* T11RF possesses a Type II CRISPR-Cas system, CRISPR3-Cas, that has 21 unique spacers (Fig 1) (37). Spacer 6 has 100% sequence identity to the *repB* gene of the PRP pAD1 (38, 39). Previous research from our lab demonstrated that T11RF CRISPR3-Cas significantly reduces the conjugation frequency of pAM714, a derivative of pAD1 conferring erythromycin resistance via *ermB* (37). CRISPR3-Cas genome defense against pAM714 required both *cas9* and spacer 6 sequences. However, despite the activity of CRISPR3-Cas, a large number (~10^5) of T11RF pAM714 transconjugants were obtained from these conjugation reactions. We hypothesized that the T11RF pAM714 transconjugants were subject to intracellular conflict between the endogenous CRISPR3-Cas and its pAM714 target, and that antibiotic selection for pAM714 could impact the outcome of this conflict.

To understand how T11RF transconjugants resolve conflicts between active CRISPR3-Cas defense and a CRISPR-Cas target, we utilized an *in vitro* evolution assay. We randomly selected transconjugant colonies from two mating schemes, T11RF pAM714 and T11RF Δ*cas9* pAM714; the Δ*cas9* strain was included as a control for the condition where CRISPR-Cas is inactive. Next, the colonies were split equally into two growth media, BHI medium and BHI medium with erythromycin to maintain selection for pAM714; see Fig 2a (a detailed explanation of the assay conditions can be found in Materials and Methods). These populations were then passaged daily for 14 days. We performed *in vitro* evolution on a total of six T11RF pAM714 (referred to as WT1-WT6) and six T11RF Δ*cas9* pAM714 (referred to as Δ1- Δ6) transconjugants originating from two independent conjugation experiments each. Every 24 h during the course of the passage, the proportion of cells within the population that maintained pAM714 was enumerated by determining the percentage of erythromycin-resistant cells relative to the total viable population.

We established the frequency of pAM714 carriage in our transconjugant colonies at Day 0, prior to serial passage (Fig 2b). As expected, pAM714 was detected at ~100% frequency for the T11RF Δ*cas9* pAM714 transconjugant colonies. The frequency of plasmid carriage in the T11RF pAM714 transconjugant colonies varied greatly and was <25% for five of the six transconjugant colonies evaluated. We attribute the variability of plasmid carriage in the T11RF pAM714 transconjugants to the biofilm-like mode of colony growth, where different cells may be exposed to different antibiotic concentrations as a result of spatial heterogeneity.

In addition to determining the frequency of plasmid carriage over the course of the *in vitro* evolution experiments, we also selected a genetic locus to assay for variation. We

chose the CRISPR3 array, which is required to produce an active Cas9-crRNA targeting complex and houses a molecular memory of previous interactions with MGEs. To assess our transconjugants for pre-existing deletions in the CRISPR3 array, we performed PCR on the individual colonies, prior to initiating the passage experiments. We observed no fixed, pre-existing CRISPR3 array deletions that would allow stable maintenance of pAM714 prior to the evolution assay (Fig 2c), an observation confirmed by Sanger sequencing of CRISPR3 amplicons.

CRISPR3-Cas eliminates its target during passage in non-selective medium. For passage without erythromycin selection, a gradual decrease in frequency of pAM714-containing cells was observed for five out of the six WT transconjugants (Fig 3a); a discussion of WT4 will be provided later. These data are consistent with CRISPR3-Cas eliminating its target, pAM714, via Cas9 programmed with a crRNA derived from the spacer 6 (S_6) sequence. In contrast, pAM714 was stably maintained at high frequencies in all of the T11RFΔ*cas9* pAM714 transconjugant populations (Fig 3a). CRISPR3 array integrity was maintained over the course of serial passage for both T11RF pAM714 and T11RFΔ*cas9* pAM714 transconjugant populations (Fig 3c). Overall, these data demonstrate that *cas9*-dependent pAM714 loss occurs in *E. faecalis* when passaged in the absence of antibiotic selection for pAM714. By extension, for 5 of 6 WT transconjugants, the progenitor recipient cells for these lineages must have had functional CRISPR-Cas defense.

As stated above, the WT4 population did not exhibit plasmid loss in the absence of antibiotic selection. We reasoned that this transconjugant may have been CRISPR-Cas-deficient prior to serial passage. We sequenced the *cas9* coding region from passage day 1 of the WT4 population and identified a mutation resulting in an Ala749Thr

159 substitution. Ala749 occurs within the RuvC nuclease domain in T11RF Cas9 and is
160 conserved in the model *Streptococcus pyogenes* Cas9 (37). Due to the critical catalytic
161 function of the RuvC domain, we hypothesize that the Ala749Thr substitution confers a
162 loss of Cas9 function. We describe the inability of the WT4 population to interfere with
163 plasmid targets in a later section.
164
165 **Under continuous antibiotic selection, conflicts can be resolved by CRISPR**
166 **memory loss**. Although pAM714 initially escapes CRISPR-Cas defense in some cells,
167 when antibiotic selection for the plasmid is absent, CRISPR-Cas depletes pAM714 from
168 transconjugant populations over time (Fig 3a). When passaging the same original
169 transconjugant populations with erythromycin selection, we observed stable
170 maintenance of pAM714 in both WT and Δ*cas9* transconjugants (Fig 3b). Knowing that
171 the 5 of the 6 WT passage experiments each initiated with at least some cells in the
172 population having active CRISPR-Cas defense, we investigated how the conflict
173 between CRISPR-Cas and pAM714 was resolved in these populations under antibiotic
174 selection.
175
176 We amplified the CRISPR3 region of erythromycin-passaged transconjugants and
177 observed significant heterogeneity in the CRISPR3 array for only the WT transconjugant
178 populations (Fig 3c). By Day 14, four of the six T11RF pAM714 transconjugants had
179 visibly reduced CRISPR3 arrays (Fig 2c); the variation in array size initiated sporadically
180 over the 14 days and was unique in pattern of emergence for each transconjugant (SFig
181 1). We utilized Sanger sequencing as a first-line assessment of CRISPR3 allele
182 composition present in Day 1 and Day 14 erythromycin-passaged populations. The
183 results showed that, for transconjugant populations where CRISPR array reduction was
184 observed, S_6 was either deleted from the array or had low sequencing quality (Table 2).

Low Sanger sequencing quality likely resulted from mixed populations with different deletion events arising stochastically, resulting in S_6 deletion. In contrast to the WT populations, CRISPR3 arrays for the T11RF Δ*cas9* pAM714 transconjugants were unchanged (Fig 3c and Table 2). We chose the Δ4 population as a representative of the T11RFΔ*cas9* pAM714 transconjugant populations for future analyses as all data indicated that the six populations were equivalent. In summary, S_6 was poorly tolerated in four WT transconjugants whereas no diversification of the CRISPR3 array occurred in Δ*cas9* transconjugants. This suggests that under antibiotic selection, the conflict between CRISPR3-Cas and its target can be resolved by compromising the CRISPR3-Cas system, in the form of either S_6 loss (populations WT2, WT3, WT5, WT6) or *cas9* mutation (population WT4).

The WT1 population is discussed further here. PCR analysis of transconjugant WT1 indicated that the wild-type CRISPR3 allele was present after 14 days of passage with erythromycin (Fig 3c). However, Sanger sequencing detected a mixed population in the region of S_6 and S_7 after passage day 1, which was not detected after passage day 14 (Table 2). Therefore, a S_6 deletion arose during the passage experiment but did not become fixed. Mutations in other CRISPR-associated factors could have arisen in the WT1 population.

To investigate this possibility, we performed whole genome Illumina deep sequencing on five Day 14 erythromycin-passaged T11RF pAM714 transconjugant populations and a control population, Δ4 (see Table 2). We observed variation in *cas9* sequence in the WT1, WT2, and WT3 populations (Table 3). All of the mutations led to nonsynonymous changes and are predicted to result in Cas9 loss of function (Table 3). In addition to *cas9* mutations, we observed variation in six other genes in some of the populations

211 (STable2). No variations were identified in the S_6 protospacer or PAM region of pAM714,
212 although one variation was identified elsewhere in *repB* in the WT2 population
213 (STable2).
214
215 **Reduced tolerance of S_6 in T11RF pAM714 transconjugant populations.** To attain
216 greater resolution of CRISPR3 alleles beyond what Sanger sequencing could achieve,
217 we deep-sequenced CRISPR3 amplicons from populations of interest, beginning with
218 BHI-passaged T11RF (lacking pAM714) (Fig. 4a) and a representative Δ*cas9* pAM714
219 transconjugant passaged in erythromycin (Fig. 4b) as controls. We first mapped
220 CRISPR3 amplicon reads to the T11 reference sequence and calculated coverage depth
221 to analyze mapping efficiency. Serial passaging for 14 days slightly altered the
222 distribution of reads across the amplicon but did not result in a strong preference for the
223 abundance or absence of any spacer.
224
225 We then expanded this analysis to the T11RF pAM714 transconjugants, excepting WT4.
226 As expected, depletion of S_6 was detected for WT2, WT3, WT5, and WT6 populations
227 after 14 days of passage with antibiotic selection (Fig 4d-g). For WT3, WT5, and WT6
228 populations, S_6 depletion was evident after one day of passage with selection (Fig 4e-g).
229 For WT1, depletion of S_6 was not detected after 14 days passage with selection (Fig 4c),
230 consistent with our Sanger sequencing results (Table 2).
231
232 To identify specific mutant CRISPR alleles in the amplicon deep sequencing, we
233 manually constructed artificial CRISPR reference sequences for every possible spacer
234 deletion event (see Materials and Methods for more information). In total, 484 references
235 were constructed, where wild type CRISPR alleles were represented by the wild type
236 references: 5'-$S_xRS_{(x+1)}$-3' ($0 \leq x < 21$) and 5'-$S_{21}TRS_T$-3'. Mutant alleles were

represented by 5'-S_xRS_y-3' ($y \neq x+1$). For control T11RF passaged for 1 or 14 days in plain BHI medium, all but 28 (day 1) and 4 (day 14) alleles out of 484 possible alleles were detected. We conclude that CRISPR3 heterogeneity naturally occurs in T11RF populations, possibly as a result of slippage during DNA replication and/or recombination between CRISPR repeat sequences. This is consistent with previous research that proposed that heterogeneity exists within CRISPR arrays in bacterial populations (40-42).

The electrophoresis analysis shown in Fig 3 revealed that some T11RF pAM714 transconjugant populations passaged in erythromycin possessed multiple CRISPR3 alleles. We resolved the most abundant mutant CRISPR3 alleles in transconjugant populations (Fig 5) by mapping amplicon reads to wild-type and mutant CRISPR3 references. The amplicon sequencing provided greater resolution than Sanger sequencing (Table 2). All T11RF transconjugant populations, other than WT1, possessed multiple co-existing CRISPR3 alleles after 14 days of passage with antibiotic selection, and each of those alleles lacked S_6. Overall, we conclude that natural heterogeneity in CRISPR3 arises at a low frequency in T11RF populations, and this heterogeneity provides a genetic basis for CRISPR (host)-plasmid conflict resolution in erythromycin-passaged WT transconjugant populations.

Preference for forward spacer deletion events. We categorized mutant CRISPR3 alleles into two groups of events: forward spacer deletion and backward spacer rearrangement. The forward mutant group represents the mutant alleles with spacer deletion (5'-S_xRS_y-3' where $y > x+1$ and 5'-S_xTRS_T-3' where $x < 21$), while the backward mutant group represents the mutant alleles with a terminal spacer becoming more leader-proximal (5'-S_xRS_y-3' where $y < x$). We calculated forward spacer deletion and

backward spacer rearrangement rates for each 5' spacer as described in Materials and Methods.

We observed slightly higher forward spacer deletion rates than backward spacer rearrangement rates for leader end spacers in Day 1 and Day 14 BHI-passaged T11RF, suggesting that spacers at the leader end are more readily deleted than flipped (Fig 6a). The forward deletion and backward rearrangement rates are similar at spacers S_9-S_{15} for Day 1 and Day 14 T11RF populations, indicating an equal chance of spacer deletion and flip (Fig 6a). As the 5' spacer reaches the terminal end of the array, the forward deletion and backward rearrangement rates decrease, indicating a dormant activity of spacer rearrangement near the terminal end. On average, Day 14 T11RF showed slightly lower rates for forward deletion and backward rearrangement than Day 1 T11RF.

The population diversity of the CRISPR3 array was evaluated in the same manner by analyzing the forward deletion and backward rearrangement rates in five T11RF pAM714 transconjugants and the Δ4 populations from Day 1 and Day 14 erythromycin passages. The distribution of forward deletion and backward rearrangement rates in the Δ4 and WT1 populations (Fig 6b and c) were similar to T11RF (Fig 6a), except that we observed an elevated forward deletion rate at S_5 for WT1 at Day 1, consistent with deep sequencing results which detected S_6 deletion alleles occurring in this population. The forward spacer deletion events in the other T11RF pAM714 erythromycin-passaged transconjugants have unique positional preferences based on an increase in the average number of reads mapped to the mutant allele references containing spacer deletions (Fig 6 c-g; red or black dots). The elevated forward deletion rates were often observed for spacers upstream of S_6, indicating a positional preference for forward deletion events upstream of S_6. We speculate that this is because internal spacer

deletions upstream of S_6 provide a selective advantage under these conditions. Finally, we did not observe significant fluctuation of backward rearrangements in erythromycin-passaged transconjugants (Fig 6 c-g), indicating that spacers are more readily deleted than flipped.

Compromised CRISPR3-Cas resulting from pAM714 conflict benefits other MGEs. Under antibiotic selection for pAM714, the CRISPR3-Cas system was compromised by S_6 deletion or *cas9* mutation. Considering that each spacer bears a unique memory of a previously encountered MGE, the loss of spacers surrounding S_6 could lead to compromised defense against multiple MGEs. To investigate this, we engineered pCF10, a PRP conferring tetracycline resistance (43), to encode different T11RF CRISPR3 protospacer targets along with the consensus CRISPR3 PAM sequence (37) (Table 1). We generated three pCF10 derivatives that would be targets for CRISPR3 S_1, S_6, or S_7, generating plasmids pWH107.S1, pWH107.S6, and pWH107.S7, respectively. Mutant CRISPR3 alleles with S_6 and S_7 deletions arose in all transconjugant populations that experienced array degeneracy after antibiotic passage (Table 2) and would therefore allow us to make conclusions about compromised defense. S_1 was maintained in Day 14 transconjugant populations (Table 2) and serves as a test for intact CRISPR-Cas function in passaged populations. We used wild type pCF10 as a control for baseline conjugation frequency as pCF10 is not targeted by T11RF CRISPR3-Cas (37).

We performed conjugation using Day 14 BHI-passaged T11RF as a recipient and *E. faecalis* OG1SSp bearing pCF10 and its derivatives as donors to ascertain the impact of CRISPR3-Cas on conjugation frequency of the plasmid constructs. All protospacers were targeted, resulting in significant reductions in conjugation frequencies relative to wild-type pCF10 (Fig 7). However, the degree of interference with plasmid transfer was different for each target;

315 CRISPR-Cas defense against a MGE bearing a target for S_7 was weak compared to S_1 and
316 S_6.
317
318 To determine the impact of serial passage on CRISPR-Cas defense in our transconjugant
319 populations, we assessed conjugation of pCF10 and its derivatives to Day 14 BHI- and
320 erythromycin-passaged populations of WT5. The erythromycin-passaged WT5
321 transconjugant population experienced the greatest diversification within the CRISPR3 array
322 (Table 2), and we expected it to be compromised in terms of its ability to defend against
323 other CRISPR targets. Day 14 BHI-passaged WT5 had conjugation frequencies very similar
324 to BHI-passaged T11RF for all pCF10 plasmids (Fig 7). This was expected because the
325 BHI-passaged WT5 population became depleted for pAM714 over time (Fig 3a), and did not
326 have any changes in the CRISPR3 array (Table 2), therefore it was expected to maintain
327 CRISPR-Cas activity against all CRISPR3 targets. In contrast, Day 14 erythromycin-
328 passaged WT5 exhibited defense only against pCF10 bearing a target for S_1 (Fig 7). This is
329 consistent with the amplicon analysis that identified multiple CRISPR3 alleles with deletions
330 of S_6 and S_7 in the erythromycin-passaged WT5 population (Table 2).
331
332 We also tested the WT4 transconjugant populations for CRISPR-Cas activity. We detected a
333 mutation within the RuvC catalytic domain coding region of *cas9* in WT4 after passage day
334 1, and WT4 failed to deplete pAM714 when passaged without erythromycin selection (Fig
335 3a). We expected both the BHI- and erythromycin-passaged populations of WT4 to be
336 completely deficient for CRISPR-Cas activity if the observed mutation conferred loss of
337 Cas9 function. CRISPR-Cas activity against S_1, S_6, and S_7 targets was in fact absent in
338 these populations (Fig 7).
339

We observed that the transfer frequencies of pCF10 and its derivatives were higher for all populations containing pAM714 (WT5-Erm, WT4-BHI, and WT4-Erm in Fig 7). We infer that pAM714 enhances pCF10 conjugation frequency via an unknown mechanism.

Spacer deletion is not exclusively RecA-dependent. Under antibiotic selection, the T11RF transconjugants lost S_6 to resolve the conflict between CRISPR-Cas and its target. The loss of S_6 was often coupled with the loss of surrounding spacers, ranging from S_1 to S_{18} (Table 2). The rearrangements associated with shortened CRISPR3-Cas arrays occurred between repeat-spacer junctions leaving behind perfectly intact repeat-spacer-repeat sequences that are still of use as guides for CRISPR interference. This phenomenon led us to hypothesize that either homologous recombination or DNA replication slippage plays a role in eliminating S_6 from the array. To study if homologous recombination had an impact on spacer loss, we constructed an in-frame deletion of *recA* in T11RF, generating strain T11RF Δ*recA*. The pAM714 plasmid was introduced into T11RF *ΔrecA* through the same conjugation procedures described previously and two select transconjugants (recA.TC1 and recA.TC2) were serially passaged for 14 days with continuous erythromycin selection. The PCR analysis for the select transconjugant colonies indicated that recA.TC1 had a wild type CRISPR3-Cas array size while recA.TC2 had a shortened CRISPR3-Cas array (Fig 8). Using Sanger sequencing, we observed that recA.TC1 lost S_6 after one day of passage in erythromycin, while the initial recA.TC2 colony had a deletion of S_6-S_7. The same CRISPR3 alleles were detected by Sanger sequencing from day 14 erythromycin-passaged. These data demonstrate that spacer deletion can occur in the absence of *recA*, and implicates DNA replication slippage in the emergence of mutant CRISPR alleles.

365 **Mechanism of conflict resolution is not specific to *E. faecalis* T11RF or pAM714**.
366 We wanted to determine if our observations were limited to one host-plasmid pair.
367 Therefore, we expanded our analysis to include *E. faecalis* OG1RF, which possesses a
368 Type II CRISPR-Cas system, CRISPR1-Cas, that is related to but distinct from
369 CRISPR3-Cas of T11RF (29, 44). Like all sequenced *E. faecalis* strains, OG1RF also
370 possesses the orphan CRISPR2 array (Fig 1). Previous research demonstrated that the
371 T11RF CRISPR2 array is active for genome defense in the presence of CRISPR1 *cas9*
372 (37). We interpret this to mean that the orphan CRISPR2 locus can be used as a native
373 genome defense system in *E. faecalis* OG1RF, due to the presence of endogenous
374 CRISPR1-Cas which was recently demonstrated to provide genome defense (45).
375
376 We utilized the shuttle vector pLZ12, which confers chloramphenicol resistance, as a
377 backbone for the generation of artificial OG1RF CRISPR1-Cas and CRISPR2
378 protospacer targets (Table 1). The pKH12 plasmid does not natively contain a
379 protospacer that would be targeted by either CRISPR1-Cas or CRISPR2 spacers (45).
380 pKHS96 is a pKH12 derivative with an engineered CRISPR1-Cas protospacer that is
381 targeted by OG1RF CRISPR1-Cas S_4. pKHS5 is a pKH12 derivative with an engineered
382 CRISPR2 protospacer that is targeted by OG1RF CRISPR2 S_6 The consensus PAM
383 sequence for both CRISPR1-Cas and CRISPR2 is NGG (37) and was included adjacent
384 to the engineered protospacers. pKH12, pKHS96 and pKHS5 were each transformed
385 into electrocompetent OG1RF. Twenty random transformants for each plasmid were
386 selected as templates for PCR to determine the initial integrity of the CRISPR1-Cas and
387 CRISPR2 arrays using Sanger sequencing. We determined that the CRISPR1-Cas and
388 CRISPR2 arrays were intact in all selected transformants, regardless of the plasmid that
389 was transformed.
390

We randomly selected three transformants for each plasmid to be used for *in vitro* evolution experiments. Each transformant was passaged in plain BHI medium and BHI medium supplemented with chloramphenicol for a period of 14 days. Similar to our observations for T11RF pAM714 transconjugants, we observed loss of pKHS5 and pKHS96 over the course of passaging without antibiotic selection (Fig 9a).

CRISPR1 and CRISPR2 integrity was assessed for transformants on passage day 14 (Fig 9b). For passages with selection, all pKHS96 transformants had reduced CRISPR1 arrays, and two of three pKHS5 transformants had reduced CRISPR2 arrays, after 14 days. Using Sanger sequencing, we confirmed that three pKHS96 transformants lost CRISPR1 S_4 while CRISPR2 remained intact, and two pKHS5 transformants lost CRISPR2 S_6 while CRISPR1 remained intact. The chloramphenicol-passaged pKHS5 transformant without a visible reduction in the CRISPR2 amplicon size was confirmed to have a mixed spacer population. Sanger sequencing revealed mixed nucleotides with low sequencing quality overlapping S_5 and S_6. This confirms that both OG1RF CRISPR1-Cas and CRISPR2 can become compromised when selection for CRISPR-targeted MGEs is present, which is consistent with what we observed in T11RF CRISPR3-Cas. Overall, we conclude that regardless of the CRISPR subtype involved or the nature of the plasmid (naturally occurring PRP or shuttle vector), spacer loss events occur under antibiotic selection for CRISPR-Cas targets in *E. faecalis*.

Discussion

It has been well documented that antibiotic use contributes to the dissemination of antibiotic resistance plasmids and the emergence of MDR organisms. However, bacteria encode genome defense systems such as CRISPR-Cas to reduce plasmid acquisition.

417 The interactions of CRISPR-Cas systems and naturally occurring resistance plasmids
418 are poorly understood, as is the impact of selection (in this case, strong antibiotic
419 selection) on these evolutionary interactions. This is of particular concern in the
420 opportunistic pathogen *E. faecalis* due to its propensity to engage in intra- and
421 interspecies HGT. Compromised CRISPR-Cas systems have been observed in MDR
422 strains of *E. faecalis* (44), substantiating the hypothesis that compromised genome
423 defense contributes to the evolution of MDR *E. faecalis*.
424
425 In our study, we used *in vitro* passaging experiments and deep sequencing analysis of
426 CRISPR3 amplicons to study the dynamics of CRISPR-Cas and its plasmid target in
427 transconjugants where these systems are in conflict. We find that the CRISPR3 array of
428 T11RF populations is naturally heterogeneous in allelic structure, with most possible
429 spacer deletion alleles occurring at low frequencies. When a CRISPR target is present,
430 CRISPR-Cas eliminates its target from the population over time. However, when
431 antibiotic selection for the target is present, CRISPR-Cas mutants emerge that allow the
432 plasmid to be maintained. One would reason that the heterogeneity of a CRISPR array
433 could be a result of either homologous recombination or slippage during DNA replication.
434 Our results demonstrate that *recA* is not required for CRISPR compromisation by spacer
435 deletion. However, the fact that we observed flipped spacers, where x > y in 5'-S_xRS_y-3',
436 indicates that homologous recombination likely does play a role. It is likely that both
437 mechanisms contribute to the emergence of heterogeneous CRISPR alleles. We do not
438 have an estimate of which process has a greater effect, nor whether additional stresses
439 beyond antibiotic selection could influence rates for each. Moreover, we do not know
440 whether sub-inhibitory antibiotic concentrations, fluctuating selection, and spatial
441 heterogeneity could alter outcomes of these conflicts.
442

CRISPR arrays undergo dynamic evolution (46-51). The addition of new spacers into the leader end provides fresh immunity to newly evolved MGEs, while leader distal spacers act as molecular 'fossils' to track the evolutionary history of strains (50, 52). At the same time, CRISPR array expansion is not unlimited as internal spacers can be deleted, providing a basis for diversification and the emergence of heterogeneous bacterial populations with dynamic CRISPR array-allele variations (42, 45). Previously, researchers used mathematical modeling to estimate the rate of these deletion events to be around e^{-4} for a type III-A CRISPR-Cas system in *Staphylococcus epidermidis* (41). This study also concluded that the ability of a MGE to escape CRISPR-Cas defense was dependent on the existence of pre-existing CRISPR mutants in recipient populations. This is in contrast to our results, where functional CRISPR-Cas and its plasmid target can co-exist in conflict in *E. faecalis* cells, although over time, this conflict is resolved by either plasmid loss or the emergence of mutants with compromised CRISPR-Cas.

Our studies utilized pAM714, which encodes a toxin-antitoxin system (53-55). The system encodes a stable toxin that will kill daughter cells that have not inherited a plasmid copy; an unstable antitoxin is encoded from the same locus that blocks toxin translation in cells with proper plasmid segregation. However, in our study, we observed a gradual decrease of erythromycin-resistant cells when we passaged T11RF pAM714 in BHI for 14 days, indicating that the toxin-antitoxin system does not have robust activity in our experiments. The fact that pAM714, pKHS67 and pKHS5 were also eliminated gradually over the passage and not immediately suggests that *E. faecalis* CRISPR-Cas systems either act slowly or are poorly expressed under laboratory conditions. An initial lag in Cas9 activation would explain the ability of pAM714 to become established in a subpopulation of CRISPR-Cas active cells. It is of interest to study the efficiency of

plasmid elimination in T11RF that overexpresses Cas9. It is also of interest to identify mechanisms of *cas9* expression regulation, about which little is known.

The work presented here contributes to an understanding of host-plasmid co-evolution and the emergence of MDR bacteria. Our work underscores the short- and long-term effects antibiotic usage has on the evolutionary trajectory of opportunistic pathogens. We hypothesize that the evolutionary outcomes of CRISPR-plasmid conflicts described in this study occur in other bacterial pathogens with Type II CRISPR-Cas systems, and that the specific outcomes observed will be a function of Cas9 expression and kinetics.

Materials and Methods

Strains, reagents, and routine molecular biology procedures. Bacterial strains and plasmids used in this study are listed in Table 1. *E. faecalis* strains were grown in Brain Heart Infusion (BHI) broth or on agar plates at 37°C unless otherwise noted. Antibiotics were used for *E. faecalis* at the following concentrations: erythromycin, 50 µg/mL; chloramphenicol, 15 µg/mL; streptomycin, 500 µg/mL; spectinomycin, 500 µg/mL; rifampicin, 50 µg/mL; fusidic acid, 25 µg/mL. *Escherichia coli* strains used for plasmid propagation and were grown in lysogeny broth (LB) broth or on agar plates at 37°C. Chloramphenicol was used at 15 µg/mL for *E. coli*. PCR was performed using *Taq* (New England Biolabs) or Phusion (Fisher Scientific) polymerases. Primer sequences used are in STable 1. Routine DNA sequencing was carried out at the Massachusetts General Hospital DNA core facility (Boston, MA). *E. faecalis* electrocompetent cells were made using the lysozyme method as previously described (56).

493 **Generation of mutant *E. faecalis* strains and plasmids**. In-frame deletion of *recA* in
494 T11RF was generated using a previously established protocol (57). Briefly, ~750 bp
495 regions up- and downstream of *recA* in *E. faecalis* T11RF were amplified, digested, and
496 ligated into pLT06 (57) to generate pWH*recA*. The resulting plasmid was transformed
497 into competent T11RF cells via electroporation (56). Following transformation at 30°C, a
498 shift to the non-permissive temperature of 42°C and counterselection on p-chloro-
499 phenylalanine were performed to generate an in-frame, markerless deletion.
500
501 To insert the T11 CRISPR3 spacer 1 (S_1), S_6, and S_7 sequences and CRISPR3 PAM
502 (TTGTA) into pCF10, 47 bp and 39 bp single stranded DNA oligos were annealed to
503 each other to generate dsDNA with restriction enzyme overhangs for *BamHI* and *PstI*.
504 The annealed oligos were ligated into the pLT06 derivative pWH107 that includes
505 sequence from pCF10 *uvrB*, to insert these sequences into the *uvrB* gene of pCF10 by
506 homologous recombination. A knock-in protocol was performed as previously described
507 (37).
508
509 **Conjugation experiments.** *E. faecalis* donor and recipient strains were grown in BHI
510 overnight to stationary phase. A 1:10 dilution was made for both donor and recipient
511 cultures in fresh BHI broth and incubated for 1.5 hr to reach mid-exponential phase. A
512 mixture of 100 µL donor cells and 900 µL recipient cells was pelleted and plated on BHI
513 agar to allow conjugation. After 18 h incubation, the conjugation mixture was scraped
514 from the plate using 2 mL 1X PBS supplemented with 2 mM EDTA. Serial dilutions were
515 prepared from the conjugation mixture and plated on selective BHI agars. After 24-48 h
516 incubation, colony forming units per milliliter (CFU/mL) was determined using plates with
517 30 - 300 colonies. The conjugation frequency was calculated as the CFU/mL of
518 transconjugants divided by the CFU/mL of donors.

519

520 **Serial passage.** Transconjugant or transformant colonies were suspended in 50 µL BHI
521 broth. The 50 µL suspension was used as follows: 3 µL was used for PCR to confirm the
522 integrity of the CRISPR array, 10 µL was inoculated into plain BHI broth, another 10 µL
523 was inoculated into selective BHI broth for plasmid selection, and another 10 µL was
524 used for serial dilution and plating on selective medium to enumerate the initial number
525 of plasmid-containing cells in the transconjugant colonies. Broth cultures were incubated
526 for 24 h, followed by 1:1000 dilution into either fresh plain BHI or fresh selective BHI. At
527 each 24 h interval, 3 µL of each culture from the previous incubation was used for PCR
528 to check CRISPR array integrity, and 10 µL was used for serial dilution and plating on
529 agars to determine CFU/mL for total viable cells and plasmid-containing cells. The
530 cultures were passaged in this manner for 14 days; cryopreserved culture stocks were
531 made daily in glycerol. To use the Day 14 transconjugant populations in conjugation
532 reactions, the glycerol stocks were completely thawed on ice, and 20 µL was inoculated
533 into plain BHI broth. The cultures were incubated for 6-8 h to allow them to reach mid-
534 exponential phase (OD_{600nm} ≈ 0.5–0.7), and 900 µL was used as recipient in conjugation
535 reactions as described above.

536

537 **Deep sequencing of CRISPR3 amplicons and genomic DNA.** For CRISPR3 amplicon
538 sequencing, 3 µL from a broth culture was used as template in PCR using Phusion
539 Polymerase with CR3_seq_F/R primers (STable1). The PCR products were purified
540 using the Thermo Scientific PCR purification kit (Thermo Scientific). Genomic DNA was
541 isolated using the phenol-chloroform method (58). The purified PCR amplicons and
542 genomic DNA samples were sequenced using 2 x 150 bp paired end sequencing
543 chemistry by Molecular Research LP (MR DNA; Texas).

544

Whole genome sequencing analysis. T11 supercontig and pAD1 plasmid contig references were downloaded from NCBI (accession numbers: T11: NZ_GG688637.1-NZ_GG688649; pAD1: AB007844, AF394225, AH011360, L01794, L19532, L37110, M84374, M87836, U00681, X17214, X62657, X62658). Reads were aligned to these references using default parameters in CLC Genomics Workbench (Qiagen) where ≥50% of each mapped read has ≥80% sequence identity to the reference. Variations occurring with ≥35% frequency at positions with ≥10X coverage between our samples and the reference contigs were detected using the Basic Variant Detector. At the same time, local realignment was performed, followed by Fixed Ploidy variant detection using default parameters and variants probability ≥90% in CLC Genomics Workbench. The basic variants and fixed ploidy variants were combined for each sequencing sample and subjected to manual inspection. The variants that were detected in the T11 genome from all samples were inferred to be variants in our parent T11 stock and were manually removed. The variants that were detected in pAD1 genome from all transconjugant samples were inferred to be variants in our pAM714 stock, hence were also manually removed. Next, variants within the CRISPR3 array were removed as we analyzed CRISPR3 alleles using a different approach (amplicon deep sequencing; see below). All variants detected from all populations were manually checked for coverage depth to eliminate the detection bias. The variants detected in all samples are shown in STable2.

Analysis of CRISPR3 amplicon sequencing. Reads from the 1,763 bp CRISPR3 amplicon were mapped to the T11 CRISPR3 reference (NZ_GG688647.1, positions 646834 - 648596) using stringent mapping conditions in CLC Genomics Workbench. The stringent mapping conditions require 100% of each mapped read to have ≥95% identity to the reference. The percent mapped reads were calculated by dividing the number of reads mapped by the total number of reads, these percentages are listed in

STable3, step 1. The coverage depth was then calculated for each position within the PCR amplicon region using CLC Genomics Workbench, normalized using reads per million, and plotted against reference positions (Fig 4).

To further analyze CRISPR3 spacer deletions and rearrangements, we manually created CRISPR3 references. The CRISPR3 amplicon references contain two spacers connected by a T11 CRISPR3-Cas repeat: 5'-spacer[x]-repeat-spacer[y]-3' (5'-SxRSy-3'), where spacer[x] and spacer[y] could be 30 bp upstream of the first repeat (leader end; or S_0 hereafter; Fig 1a), or any internal spacer within the CRISPR3 array (from spacer 1 to spacer 21; or S_1 to S_{21}; Fig 1a). Each manually generated CRISPR3 amplicon reference is 96 bp in length. The references where y=x+1 represent wild-type alleles. The terminal repeat following S_{21} in the CRISPR3 array is divergent from the regular direct repeat sequence, so references containing 5'-spacer[x]-TerminalRepeat-S_T-3' (5'-S_xTRS_T-3') were constructed, where spacer[x] ranges from S_0 to S_{21} and spacer S_T represents the sequence 30 bp downstream of the terminal repeat (Fig 1a). The 5'-S_{21}TRS_T-3' reference represents the wild-type. In total, 484 references with length of 96 bp were generated for the CRISPR3 amplicon. Considering that the read length is 150 bp, we manually split each read into two subsequences (one subsequence was 75 bp; with the remainder of the read being the second subsequence) to enhance mapping efficiency, allowing for retrieval of maximal sequence information. The split amplicon sequencing reads were mapped to the 5'-SxRSy-3' and 5'-SxTRS_T-3' references using stringent mapping parameters in CLC Genomics Workbench (Qiagen). The stringent mapping parameters require 100% of each mapped read to be ≥95% identical to one unique reference. Thus, the sequencing reads from different CRISPR alleles will be distinguished. These amplicon mapping results were applied to the calculation of forward spacer deletion and backward spacer rearrangement rates.

To further evaluate the mapping efficiency, the unmapped reads from initial mapping to the T11 CRISPR3 reference (STable3, step 1) were subjected to additional quality control analysis. The unmapped reads were mapped to the 484 manually created spacer[x]-repeat-spacer[y] references using the same mapping parameters in CLC as above (STable3, step 2 mapping; ignore unspecific mapping). The unmapped reads from step 2 were subjected to mapping to all possible references (CRISPR3 region plus manually created references) using default mapping parameters, ignoring unspecific mapping (80% of each mapped read has at least 50% identity to the reference sequence; STable3, step 3 mapping). The unmapped reads from step 3 were mapped to all possible references using the default mapping parameters and randomly map unspecific matching reads (STable3, step 4 mapping).

Forward spacer deletion and backward spacer rearrangement. We observed two categories of mutant CRISPR3 alleles: 5'-S_xRS_y-3' (y > x+1) and 5'-S_xRS_y-3' (y < x). The forward deletion mutants with 5'-S_xRS_y-3' (y > x+1) are the result of spacer deletions, with spacers from S_{x+1} to S_{y-1} deleted; while the backward rearrangement mutants with 5'-S_xRS_y-3' (y < x) are the result of spacer rearrangement, where a downstream spacer S_y flips to become upstream of an upstream spacer S_x. To study if there were positional preferences, the average forward spacer deletion rate and backward spacer rearrangement rate was calculated for each 5'-S_x (0 < x < 21) within the CRISPR3 array. For each 5'-S_x, the average forward deletion and backward rearrangement rate are calculated as:

$$P(\text{5'-Sx Forward}) = \frac{\#\text{ mapped reads to the reference of 5'-SxRSy-3'}}{\sum_{y=x+1}^{n} \#\text{ mapped reads to the references of 5'-SxRSy-3' and 5'-SnTRST-3'}}$$

$$\mathrm{P}(\text{5'-Sx Backward}) = \frac{\#\ \text{mapped reads to the reference of 5'-SxRSy-3'}}{\sum_{y=0}^{x-1} \#\ \text{mapped reads to the references of 5'-SxRSy-3'}}$$

where n is the total number of spacers within a CRISPR array, hence S_n represents terminal spacer, as described above (Fig 1a).

Acknowledgements

This work was supported by Public Health Service grant R01AI116610 to K.L.P and the Cecil H. and Ida Green Chairship to M.Q.Z. We thank Karthik Hullahalli for construction of pKH12 and its derivatives. We thank Dr. Chen Jia for consultation on data analysis methods.

Competing Interests

The authors declare no competing interests.

635 **References**

636 1. Hegstad K, Mikalsen T, Coque T, Werner G, & Sundsfjord A (2010) Mobile
637 genetic elements and their contribution to the emergence of antimicrobial
638 resistant *Enterococcus faecalis* and *Enterococcus faecium*. *Clinical microbiology*
639 *and infection* 16(6):541-554.

640 2. Huycke MM, Sahm DF, & Gilmore MS (1998) Multiple-drug resistant enterococci:
641 the nature of the problem and an agenda for the future. *Emerging infectious*
642 *diseases* 4(2):239.

643 3. Kristich CJ, Rice LB, & Arias C, A (2014) Enterococcal Infection-Treatment and
644 Antibiotic Resistance. *Enterococci: From Commensals to Leading Causes of*
645 *Drug Resistant Infection*.

646 4. Palmer KL, Kos VN, & Gilmore MS (2010) Horizontal gene transfer and the
647 genomics of enterococcal antibiotic resistance. *Current opinion in microbiology*
648 13(5):632-639.

649 5. Uttley A, George R, Naidoo J, Woodford N, Johnson A*, et al.* (1989) High-level
650 vancomycin-resistant enterococci causing hospital infections. *Epidemiology &*
651 *Infection* 103(1):173-181.

652 6. Clewell DB, Weaver KE, Dunny GM, Coque TM, Francia MV*, et al.* (2014)
653 Extrachromosomal and Mobile Elements in Enterococci: Transmission,
654 Matintenance, and Epidemiology. *Enterococci: From Commensals to Leading*
655 *Causes of Drug Resistant Infection*.

656 7. Dunny GM (2007) The peptide pheromone-inducible conjugation system of
657 *Enterococcus faecalis* plasmid pCF10: cell-cell signalling, gene transfer,
658 complexity and evolution. *Philosophical transactions of the Royal Society of*
659 *London. Series B, Biological sciences* 362(1483):1185-1193.

660 8. Gilmore MS, Segarra RA, Booth MC, Bogie CP, Hall LR*, et al.* (1994) Genetic
661 structure of the *Enterococcus faecalis* plasmid pAD1-encoded cytolytic toxin
662 system and its relationship to lantibiotic determinants. *J Bacteriol* 176(23):7335-
663 7344.

664 9. Clewell DB (2007) Properties of *Enterococcus faecalis* plasmid pAD1, a member
665 of a widely disseminated family of pheromone-responding, conjugative, virulence
666 elements encoding cytolysin. *Plasmid* 58(3):205-227.

10. Dunny GM (2013) Enterococcal sex pheromones: signaling, social behavior, and evolution. *Annual review of genetics* 47:457-482.

11. Jasni AS, Mullany P, Hussain H, & Roberts AP (2010) Demonstration of conjugative transposon (Tn*5397*)-mediated horizontal gene transfer between *Clostridium difficile* and *Enterococcus faecalis*. *Antimicrobial agents and chemotherapy* 54(11):4924-4926.

12. Showsh SA, De Boever EH, & Clewell DB (2001) Vancomycin resistance plasmid in *Enterococcus faecalis* that encodes sensitivity to a sex pheromone also produced by *Staphylococcus aureus*. *Antimicrobial agents and chemotherapy* 45(7):2177-2178.

13. Zhu W, Murray PR, Huskins WC, Jernigan JA, McDonald LC*, et al.* (2010) Dissemination of an *Enterococcus* Inc18-Like vanA plasmid associated with vancomycin-resistant *Staphylococcus aureus*. *Antimicrobial agents and chemotherapy* 54(10):4314-4320.

14. Tsvetkova K, Marvaud J-C, & Lambert T (2010) Analysis of the mobilization functions of the vancomycin resistance transposon Tn*1549*, a member of a new family of conjugative elements. *Journal of bacteriology* 192(3):702-713.

15. Andersson DI & Levin BR (1999) The biological cost of antibiotic resistance. *Current opinion in microbiology* 2:489-493.

16. De Gelder L, Ponciano JM, Joyce P, & Top EM (2007) Stability of a promiscuous plasmid in different hosts: no guarantee for a long-term relationship. *Microbiology* 153(2):452-463.

17. Harrison E & Brockhurst MA (2012) Plasmid-mediated horizontal gene transfer is a coevolutionary process. *Trends in microbiology* 20(6):262-267.

18. Sota M, Yano H, Hughes J, Daughdrill GW, Abdo Z*, et al.* (2010) Shifts in Host Range of a Promiscuous Plasmid through Parallel Evolution of its Replication Initiation Protein. *The ISME journal* 4(12):1568.

19. Hulter N, Ilhan J, Wein T, Kadibalban AS, Hammerschmidt K*, et al.* (2017) An evolutionary perspective on plasmid lifestyle modes. *Current opinion in microbiology* 38:74-80.

20. Gandon S & Vale PF (2014) The evolution of resistance against good and bad infections. *J Evol Biol* 27(2):303-312.

699 21. Stalder T, Rogers LM, Renfrow C, Yano H, Smith Z*, et al.* (2017) Emerging
700 patterns of plasmid-host coevolution that stabilize antibiotic resistance. *Scientific*
701 *reports* 7(1):4853.

702 22. Harrison E, Guymer D, Spiers AJ, Paterson S, & Brockhurst MA (2015) Parallel
703 compensatory evolution stabilizes plasmids across the parasitism-mutualism
704 continuum. *Current biology : CB* 25(15):2034-2039.

705 23. Hughes JM, Lohman BK, Deckert GE, Nichols EP, Settles M*, et al.* (2012) The
706 role of clonal interference in the evolutionary dynamics of plasmid-host
707 adaptation. *mBio* 3(4):e00077-00012.

708 24. Makarova KS, Wolf YI, Alkhnbashi OS, Costa F, Shah SA*, et al.* (2015) An
709 updated evolutionary classification of CRISPR-Cas systems. *Nature reviews.*
710 *Microbiology* 13(11):722-736.

711 25. Marraffini LA (2015) CRISPR-Cas immunity in prokaryotes. *Nature* 526(7571):55-
712 61.

713 26. Bolotin A, Quinquis B, Sorokin A, & Ehrlich SD (2005) Clustered regularly
714 interspaced short palindrome repeats (CRISPRs) have spacers of
715 extrachromosomal origin. *Microbiology* 151(Pt 8):2551-2561.

716 27. Mojica FJ, Diez-Villasenor C, Garcia-Martinez J, & Soria E (2005) Intervening
717 sequences of regularly spaced prokaryotic repeats derive from foreign genetic
718 elements. *Journal of molecular evolution* 60(2):174-182.

719 28. Pourcel C, Salvignol G, & Vergnaud G (2005) CRISPR elements in *Yersinia*
720 *pestis* acquire new repeats by preferential uptake of bacteriophage DNA, and
721 provide additional tools for evolutionary studies. *Microbiology* 151(Pt 3):653-663.

722 29. Bourgogne A, Garsin DA, Qin X, Singh KV, Sillanpaa J*, et al.* (2008) Large scale
723 variation in *Enterococcus faecalis* illustrated by the genome analysis of strain
724 OG1RF. *Genome biology* 9(7):R110.

725 30. Heler R, Samai P, Modell JW, Weiner C, Goldberg GW*, et al.* (2015) Cas9
726 specifies functional viral targets during CRISPR-Cas adaptation. *Nature*
727 519(7542):199-202.

728 31. Wei Y, Terns RM, & Terns MP (2015) Cas9 function and host genome sampling
729 in Type II-A CRISPR-Cas adaptation. *Genes & development* 29(4):356-361.

730 32. Chylinski K, Le Rhun A, & Charpentier E (2013) The tracrRNA and Cas9 families
731 of type II CRISPR-Cas immunity systems. *RNA biology* 10(5):726-737.

732 33. Deltcheva E, Chylinski K, Sharma CM, Gonzales K, Chao Y*, et al.* (2011)
733 CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III.
734 *Nature* 471(7340):602-607.

735 34. Anders C, Niewoehner O, Duerst A, & Jinek M (2014) Structural basis of PAM-
736 dependent target DNA recognition by the Cas9 endonuclease. *Nature*
737 513(7519):569-573.

738 35. Nishimasu H, Ran FA, Hsu PD, Konermann S, Shehata SI*, et al.* (2014) Crystal
739 structure of Cas9 in complex with guide RNA and target DNA. *Cell* 156(5):935-
740 949.

741 36. Sternberg SH, Redding S, Jinek M, Greene EC, & Doudna JA (2014) DNA
742 interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature*
743 507(7490):62-67.

744 37. Price VJ, Huo W, Sharifi A, Palmer KL, & Fey PD (2016) CRISPR-Cas and
745 Restriction-Modification Act Additively against Conjugative Antibiotic Resistance
746 Plasmid Transfer in *Enterococcus faecalis*. *mSphere* 1(3):e00064-00016.

747 38. Clewell D, Tomich P, Gawron-Burke M, Franke A, Yagi Y*, et al.* (1982) Mapping
748 of *Streptococcus faecalis* plasmids pAD1 and pAD2 and studies relating to
749 transposition of Tn*917*. *Journal of Bacteriology* 152(3):1220-1230.

750 39. Ike Y, Clewell D, Segarra R, & Gilmore M (1990) Genetic analysis of the pAD1
751 hemolysin/bacteriocin determinant in *Enterococcus faecalis*: Tn*917* insertional
752 mutagenesis and cloning. *Journal of bacteriology* 172(1):155-163.

753 40. Gudbergsdottir S, Deng L, Chen Z, Jensen JV, Jensen LR*, et al.* (2011) Dynamic
754 properties of the *Sulfolobus* CRISPR/Cas and CRISPR/Cmr systems when
755 challenged with vector‐borne viral and plasmid genes and protospacers.
756 *Molecular microbiology* 79(1):35-49.

757 41. Jiang W, Maniv I, Arain F, Wang Y, Levin BR*, et al.* (2013) Dealing with the
758 evolutionary downside of CRISPR immunity: bacteria and beneficial plasmids.
759 *PLoS genetics* 9(9):e1003844.

760 42. Lopez‐Sanchez MJ, Sauvage E, Da Cunha V, Clermont D, Ratsima Hariniaina
761 E*, et al.* (2012) The highly dynamic CRISPR1 system of *Streptococcus*

agalactiae controls the diversity of its mobilome. *Molecular microbiology* 85(6):1057-1071.

43. Dunny G, Yuhasz M, & Ehrenfeld E (1982) Genetic and physiological analysis of conjugation in *Streptococcus faecalis Journal of Bacteriology* 151(2):855-859.

44. Palmer KL & Gilmore MS (2010) Multidrug-resistant enterococci lack CRISPR-cas. *mBio* 1(4).

45. Hullahalli K, Rodrigues M, & Palmer KL (2017) Exploiting CRISPR-Cas to manipulate *Enterococcus faecalis* populations. *eLife* 6.

46. Andersson AF & Banfield JF (2008) Virus population dynamics and acquired virus resistance in natural microbial communities. *Science* 320(5879):1047-1050.

47. Bzymek M & Lovett ST (2001) Instability of repetitive DNA sequences: the role of replication in multiple mechanisms. *Proceedings of the National Academy of Sciences* 98(15):8319-8325.

48. Erdmann S & Garrett RA (2012) Selective and hyperactive uptake of foreign DNA by adaptive immune systems of an archaeon via two distinct mechanisms. *Molecular microbiology* 85(6):1044-1056.

49. Iranzo J, Lobkovsky AE, Wolf YI, & Koonin EV (2013) Evolutionary dynamics of the prokaryotic adaptive immunity system CRISPR-Cas in an explicit ecological context. *Journal of bacteriology* 195(17):3834-3844.

50. Louwen R, Horst-Kreft D, de Boer AG, van der Graaf L, de Knegt G*, et al.* (2013) A novel link between *Campylobacter jejuni* bacteriophage defence, virulence and Guillain-Barre syndrome. *European journal of clinical microbiology & infectious diseases : official publication of the European Society of Clinical Microbiology* 32(2):207-226.

51. Tyson GW & Banfield JF (2008) Rapidly evolving CRISPRs implicated in acquired resistance of microorganisms to viruses. *Environmental microbiology* 10(1):200-207.

52. Hullahalli K, Rodrigues M, Schmidt BD, Li X, Bhardwaj P*, et al.* (2015) Comparative Analysis of the Orphan CRISPR2 Locus in 242 *Enterococcus faecalis* Strains. *PloS one* 10(9):e0138890.

53. Greenfield TJ & Weaver KE (2000) Antisense RNA regulation of the pAD1 *par* post-segregational killing system requires interaction at the 5' and 3' ends of the RNAs. *Molecular microbiology* 37(3):661-670.

54. Weaver KE & Clewell DB (1989) Construction of *Enterococcus faecalis* pAD1 Miniplasmids: Identification of a Minimal Pheromone Response Regulatory Region and Evaluation of a Novel Pheromone-Dependent Growth Inhibition. *Plasmid* 22:106-119.

55. Weaver KE & Tritle DJ (1994) Identification and Characterization of an *Enterococcus faecalis* Plasmid pAD1-Encoded Stability Determinant Which Produces Two Small RNA Molecules Necessary for Its Function. *Plasmid* 32:168-181.

56. Bae T, Kozlowicz B, & Dunny GM (2002) Two targets in pCF10 DNA for PrgX binding: their role in production of Qa and prgX mRNA and in regulation of pheromone-inducible conjugation. *Journal of molecular biology* 315(5):995-1007.

57. Thurlow LR, Thomas VC, & Hancock LE (2009) Capsular polysaccharide production in *Enterococcus faecalis* and contribution of CpsF to capsule serospecificity. *Journal of bacteriology* 191(20):6203-6210.

58. Manson JM, Keis S, Smith JM, & Cook GM (2003) A clonal lineage of VanA-type *Enterococcus faecalis* predominates in vancomycin-resistant enterococci isolated in New Zealand. *Antimicrobial agents and chemotherapy* 47(1):204-210.

59. Gold OG, Jordan H, & Van Houte J (1975) The prevalence of enterococci in the human mouth and their pathogenicity in animal models. *Archives of Oral Biology* 20(7):473IN415-477.

60. Clewell DB, Tomich PK, Gawron-Burke MC, Franke AE, Yagi Y*, et al.* (1982) Mapping of *Streptococcus faecalis* plasmids pAD1 and pAD2 and studies relating to transposition of Tn*917 Journal of Bacteriology* 152(3):1220-1230.

61. Perez-Casal J, Caparon M, & Scott J (1991) Mry, a trans-acting positive regulator of the M protein gene of *Streptococcus pyogenes* with similarity to the receptor proteins of two-component regulatory systems. *Journal of Bacteriology* 173(8):2617-2624.

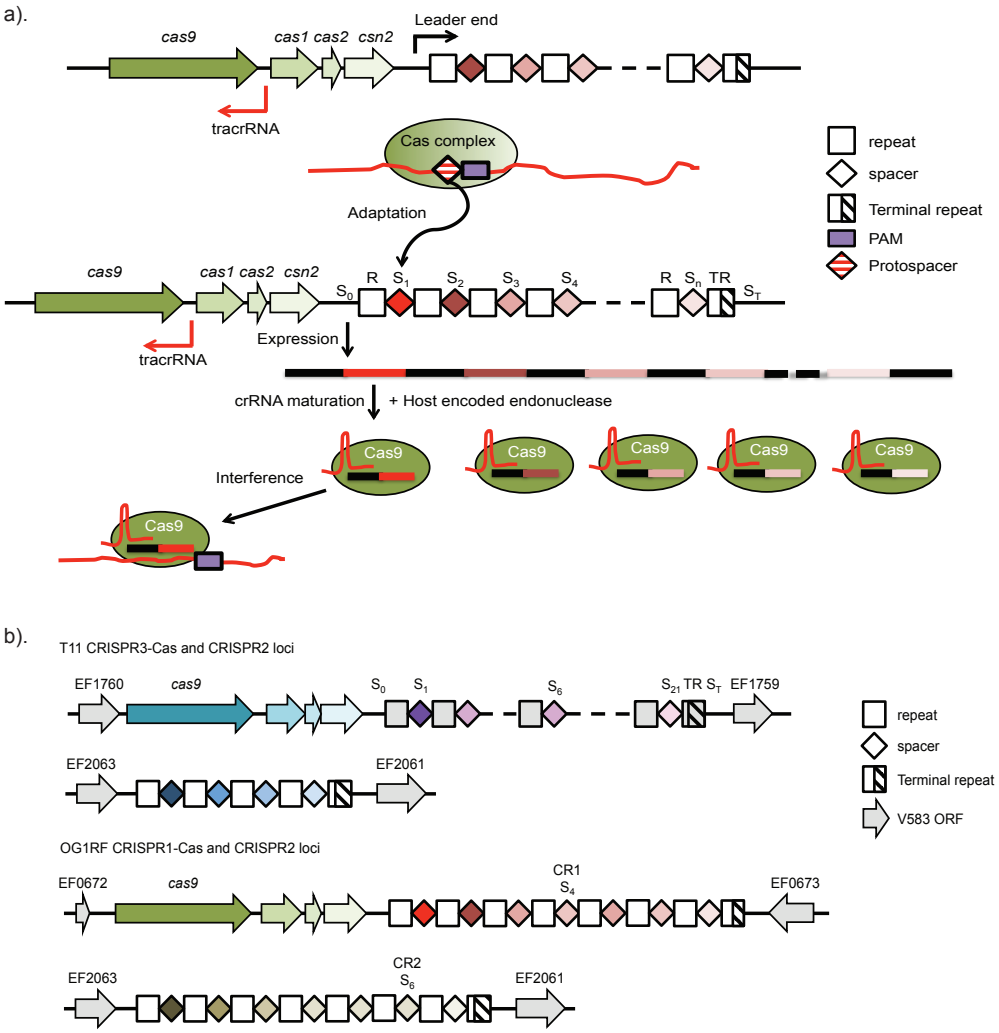


Figure 1. *E. faecalis* possesses two Type II CRISPR-Cas systems and one orphan CRISPR. a) Schematic mechanism of Type II CRISPR-Cas defense in bacteria. Upon MGE invasion, CRISPR-Cas acts as a genome defense system. When a new MGE is encountered, the protospacer is recognized based on Protospacer Adjacent Motif (PAM). A complex of Cas proteins incorporates the protospacer into the leader end of CRISPR array to form a new spacer (Adaptation). During the expression stage, the CRISPR array is transcribed into pre-crRNA, which is further processed into mature crRNA by Cas9, tracrRNA and a host-encoded endonuclease. The mature crRNA consists of part of a repeat and part of a spacer, which is bound to a Cas9:tracrRNA complex to form an effector complex. When the previously encountered MGE invades again, the effector complex recognizes the target by sequence complementarity and the presence of a PAM. Upon recognition, the target is cleaved and thus invasion by the MGE is blocked. The definition of R, TR and S_n is described in Material and Methods. b). CRISPR-Cas loci occurring in *E. faecalis* T11RF and OG1RF. *E. faecalis* T11RF encodes a CRISPR3-Cas system and a CRISPR2 array. S_6 within the CRISPR3 array targets pAM714 and pWH107.S6, while S_1 and S_7 within CRISPR3 array target pWH107.S1 and pWH107.S7, respectively. *E. faecalis* OG1RF encodes a CRISPR1-Cas system and a CRISPR2 array. S_4 within CRISPR1 array targets pKHS96, while S_6 within CRISPR2 array targets pKHS5.

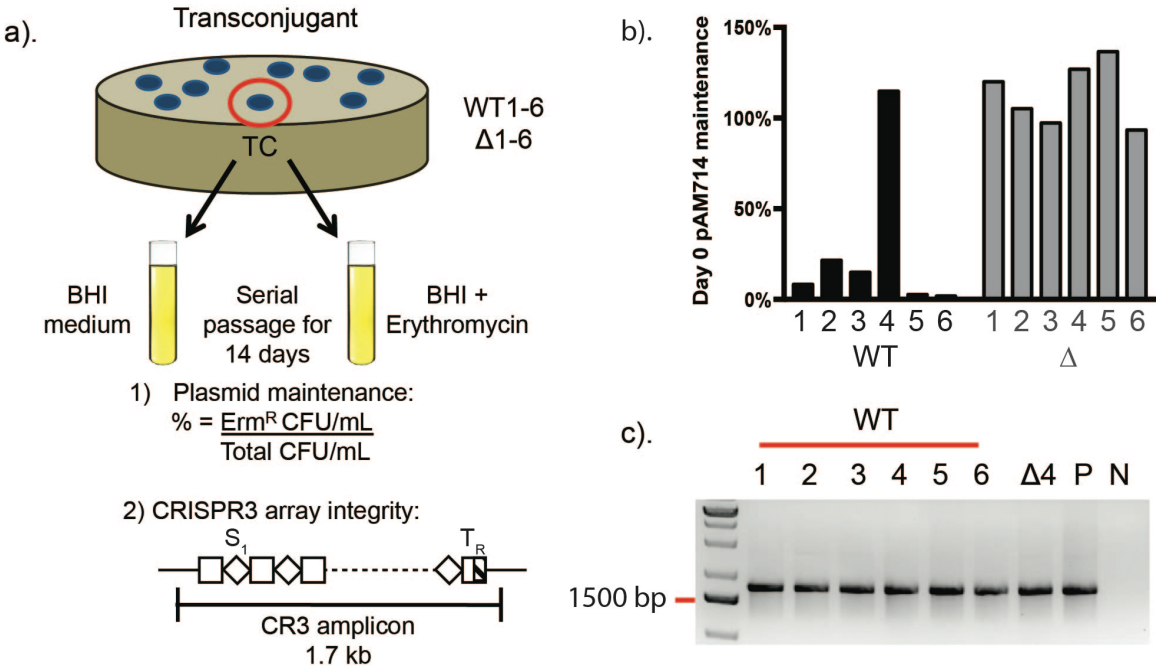


Figure 2. Design of *in vitro* evolution experiment and initial plasmid carriage and CRISPR-Cas phenotypes of select transconjugants. a) Design of *in vitro* evolution assay. Randomly selected T11RF pAM714 and T11RFΔ*cas9* pAM714 transconjugants were passaged for 14 days in the presence and absence of antibiotic selection for pAM714. These populations were monitored daily for: 1) pAM714 maintenance by determining the percentage of the population that was erythromycin-resistant, and 2) deviations in the CRISPR3 array by amplifying the 1.7 kb region encompassing the CRISPR3 array. b) Frequency of pAM714 carriage in transconjugant colonies used to initiate serial passage experiments. c) CRISPR3 amplicon PCR results for transconjugant colonies used to initiate serial passage experiments. Shown are CRISPR3 amplicon sizes for six T11RF pAM714 transconjugants (WT 1-6), a representative T11RFΔ*cas9* pAM714 transconjugant (Δ4), T11RF genomic DNA as a positive control (P), and a reagent control (N).

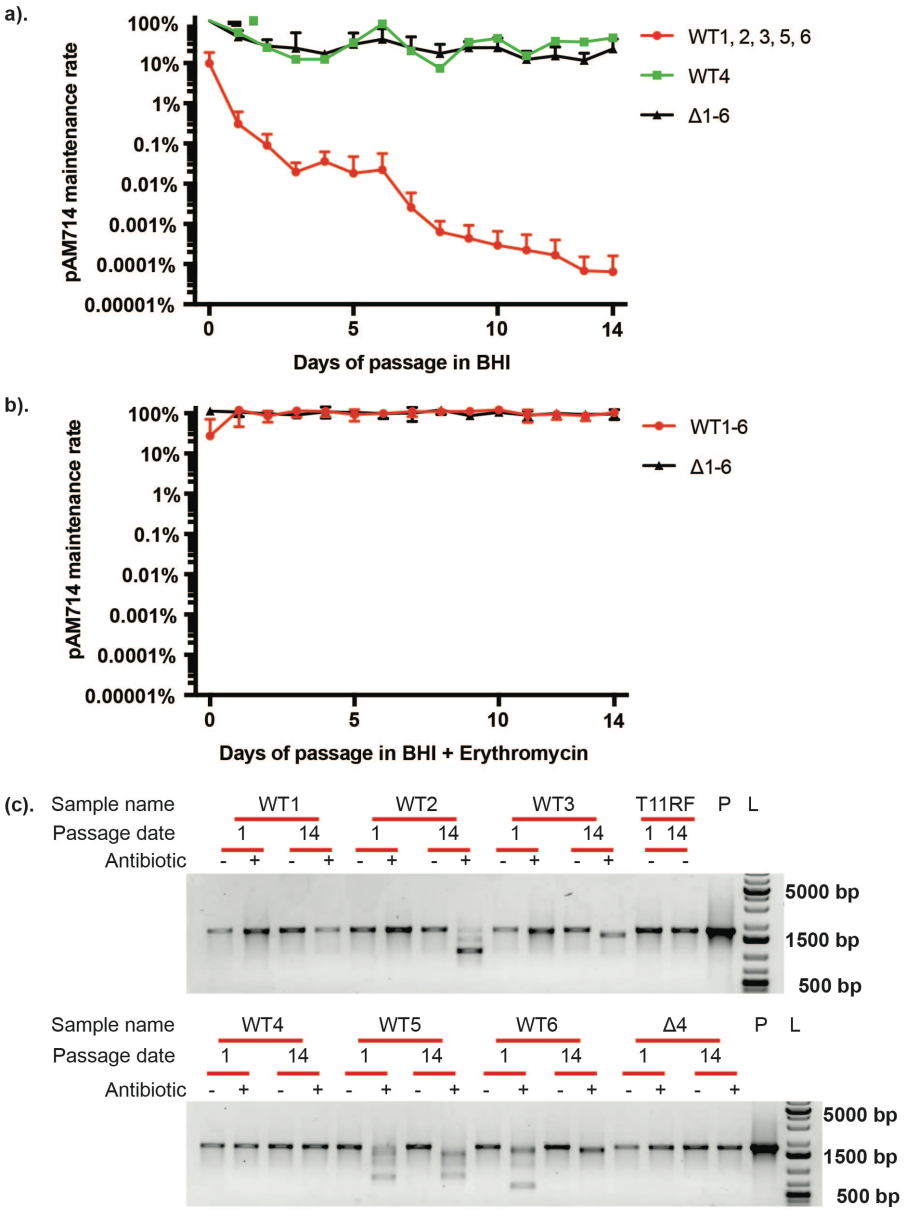
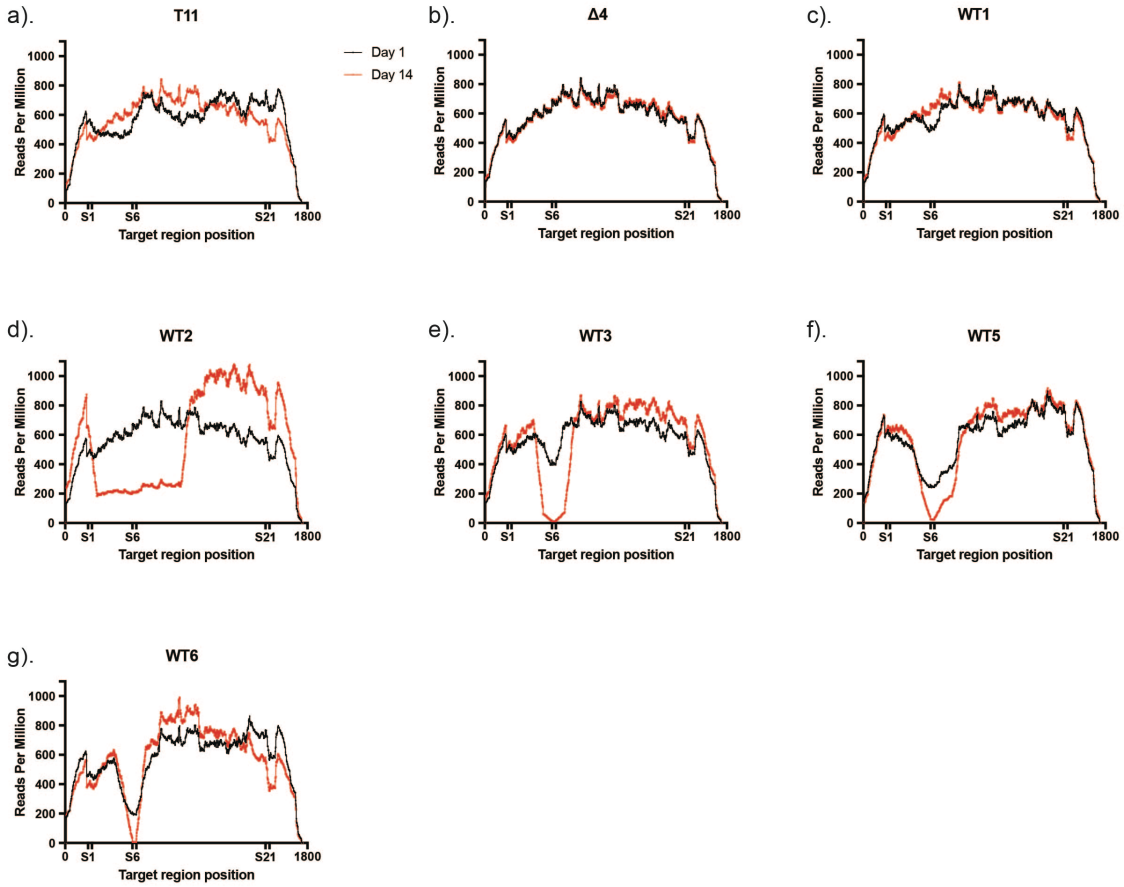


Figure 3. Antibiotic selection-specific phenotypes reflect outcomes of plasmid-host interactions. a-b) pAM714 maintenance over the course of passage without (a) and with (b) antibiotic selection. Plasmid maintenance is expressed as percentage of bacterial cells conferring erythromycin resistance. WT populations are shown in green or red and Δ*cas9* populations are shown in black. c) CRISPR3 amplicon size from early (Day 1) and late (Day 14) passage dates for six WT transconjugant populations and a representative Δ*cas9* transconjugant population (Δ4). As a control, T11RF without pAM714 was passaged for 14 days and the CRISPR3 locus was queried (T11RF). P: positive control, T11RF genomic DNA. L: DNA ladder.

The following figure supplement is available for Figure 3:

Figure supplement 1. Gel electrophoresis of CRISPR3 amplicons in T11RF pAM714 transconjugants over 14 days passaged with and without antibiotic.

877



878

879 **Figure 4. Targeted sequencing revealed stochastic spacer loss after passage in**
880 **antibiotic.** The coverage depth is calculated for each position within the amplicon and
881 normalized using reads per million, which is then plotted against the genomic position.
882 For each sample, the results for Day 1 and Day 14 are represented in black and red
883 lines, respectively. The beginning and end of the regions along the amplicon
884 corresponding to S_1, S_6 and S_{21} within the CRISPR3 array are labeled with vertical hash
885 marks on the x-axis. a) BHI passaged T11RF parent strain. b) Erythromycin passaged
886 Δ4 transconjugant. c-g) Erythromycin passaged WT transconjugants. Here, the WT4
887 population is not included due to the inactivating *cas9* mutation, as discussed in the
888 main text.
889
890 The following source data is available for Figure 4:
891 **Figure 4 - Source data 1**. Coverage depth mapping report for each position within the
892 CRISPR3 amplicon.
893

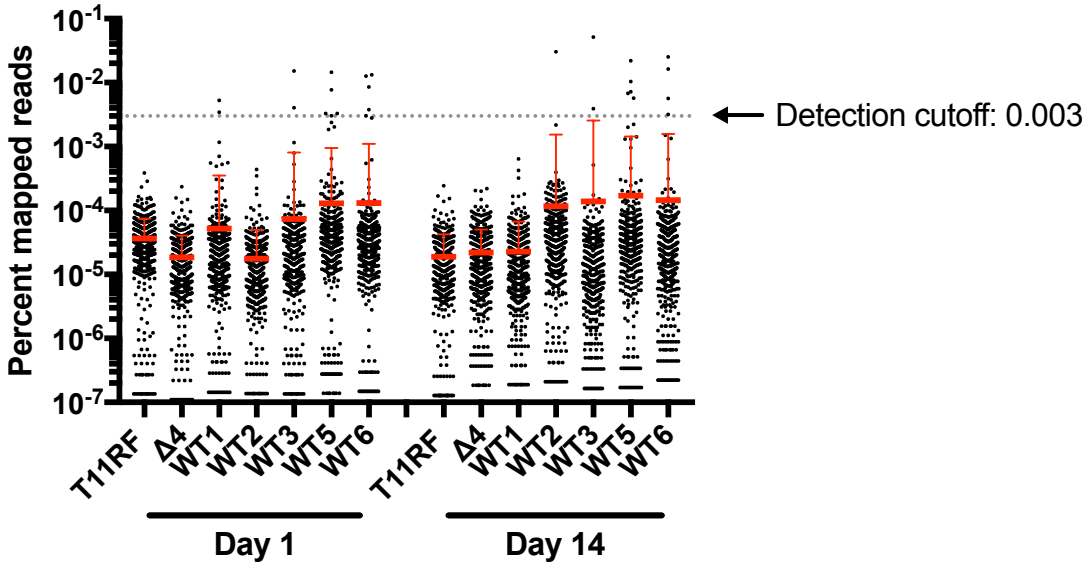


Figure 5. Distribution of mutant CRISPR3 array alleles among antibiotic-passaged wild-type transconjugants. Percent mapped reads were calculated for each artificial reference by dividing mapped reads at each position to the total number of mapped reads. The percent mapped reads to mutant alleles (dots) are shown here with average (thick red bar) and standard deviation (thin red bar). A detection cutoff value was applied so that mutant alleles with high abundances can be detected.

The following source data is available for Figure 5:

Figure 5 - Source data 1. Percent reads mapped to artificial CRISPR references.

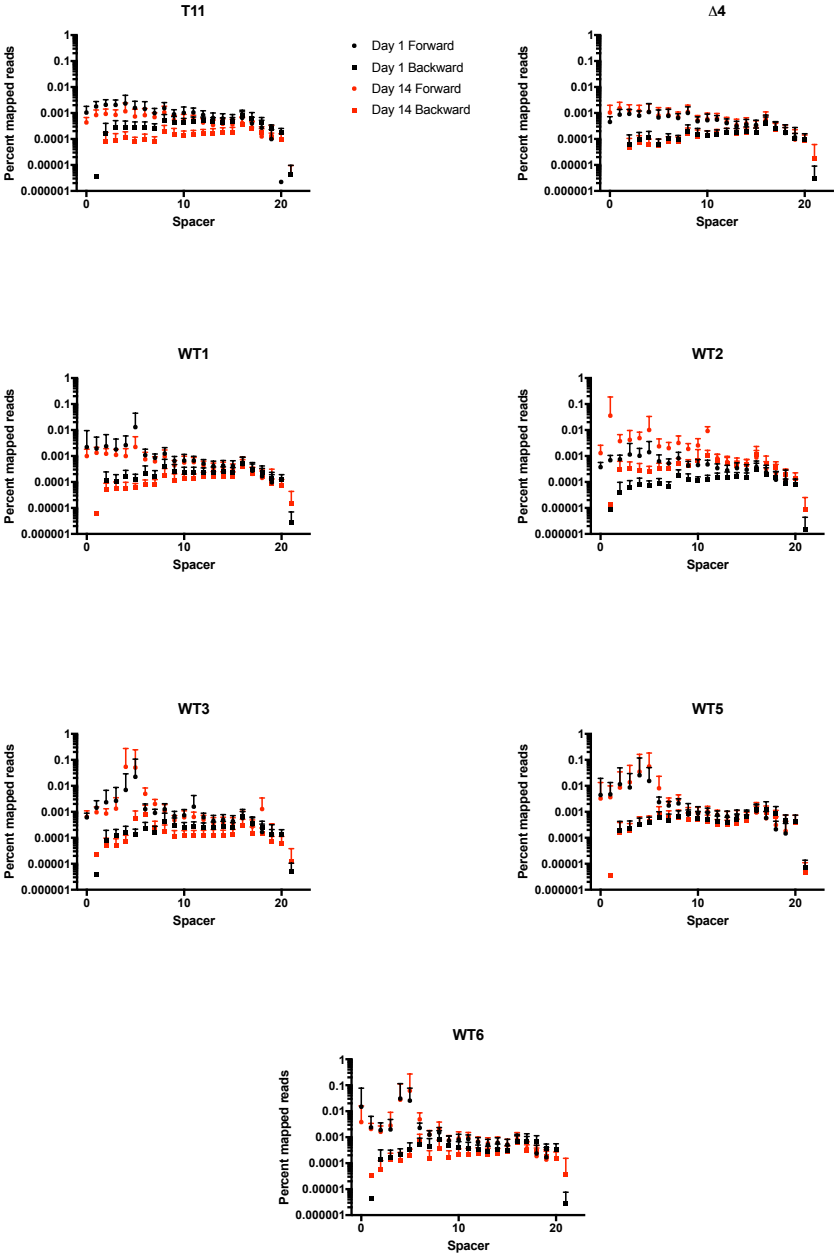


Figure 6. Mutant CRISPR alleles arise predominately through forward spacer deletion events. The forward spacer deletion and backward rearrangement rates (y-axis) were calculated for the CRISPR3 amplicon of each passaged population and are plotted against each spacer occurring in the CRISPR3 array shown on the x-axis. For each sample, the forward deletion (dots) and backward rearrangement (squares) rates for Day 1 and Day 14 of the passage are shown in black and red, respectively. a) BHI passaged T11RF parent strain. b) Erythromycin passaged Δ4 transconjugant. c-g) Erythromycin passaged WT transconjugants.

The following source data is available for Figure 6:

Figure 6 - Source data 1. Percent reads mapped to all possible CRISPR alleles.

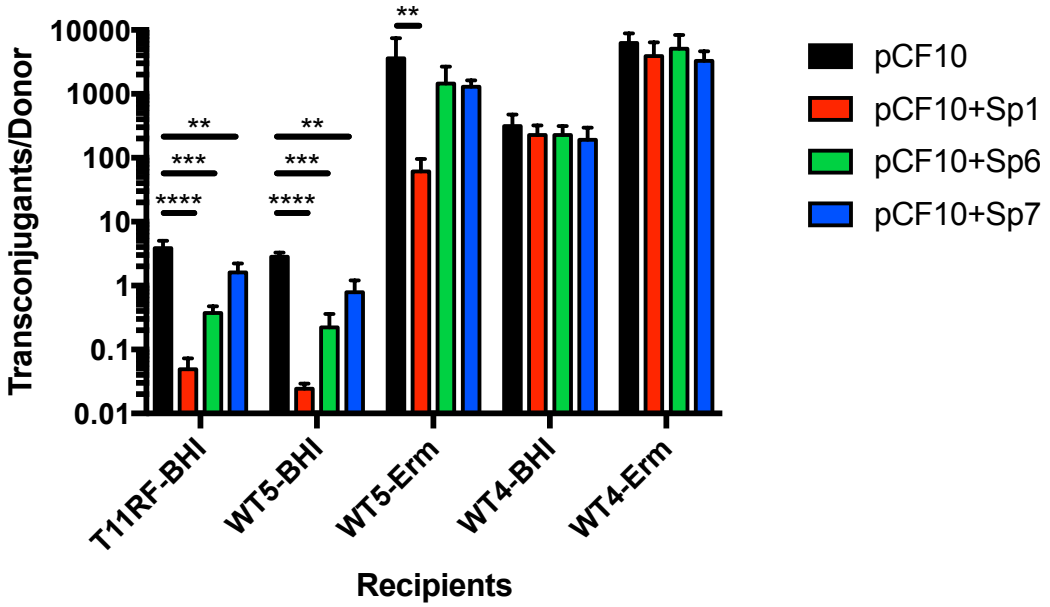


Figure 7. Compromised CRISPR-Cas primes populations for MGE acquisition. Day 14 transconjugant populations passaged in BHI and erythromycin were used as recipients in conjugation with OG1SSp pCF10 and derivatives with protospacers corresponding to spacers 1, 6 and 7 of the T11RF CRISPR3 array. The Day 14 BHI-passaged T11RF control population was used as recipient in conjugation, serving as positive control. The graph shows the conjugation frequency or ratio of transconjugants to donors from mating reactions. Statistical significance was determined using a student's t-test; P-values: ** ≤ 0.01; *** ≤ 0.001; **** ≤ 0.0001.

The following source data is available for Figure 7:

Figure 7 - Source data 1. CFU/mL of transconjugant and donor cell populations from conjugation reactions.

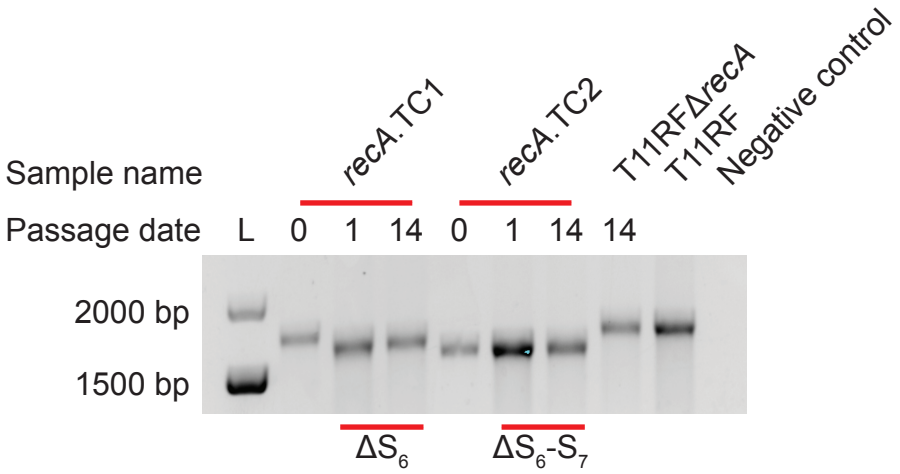


Figure 8. CRISPR3 array reduction is not dependent on RecA. Two randomly selected transconjugants were passaged *in vitro* with erythromycin selection and the CRISPR3 amplicon sizes were monitored using PCR and gel electrophoresis. As a control, the T11RFΔ*recA* parent strain was passaged in BHI and used as a control in PCR analysis. L: DNA Ladder.

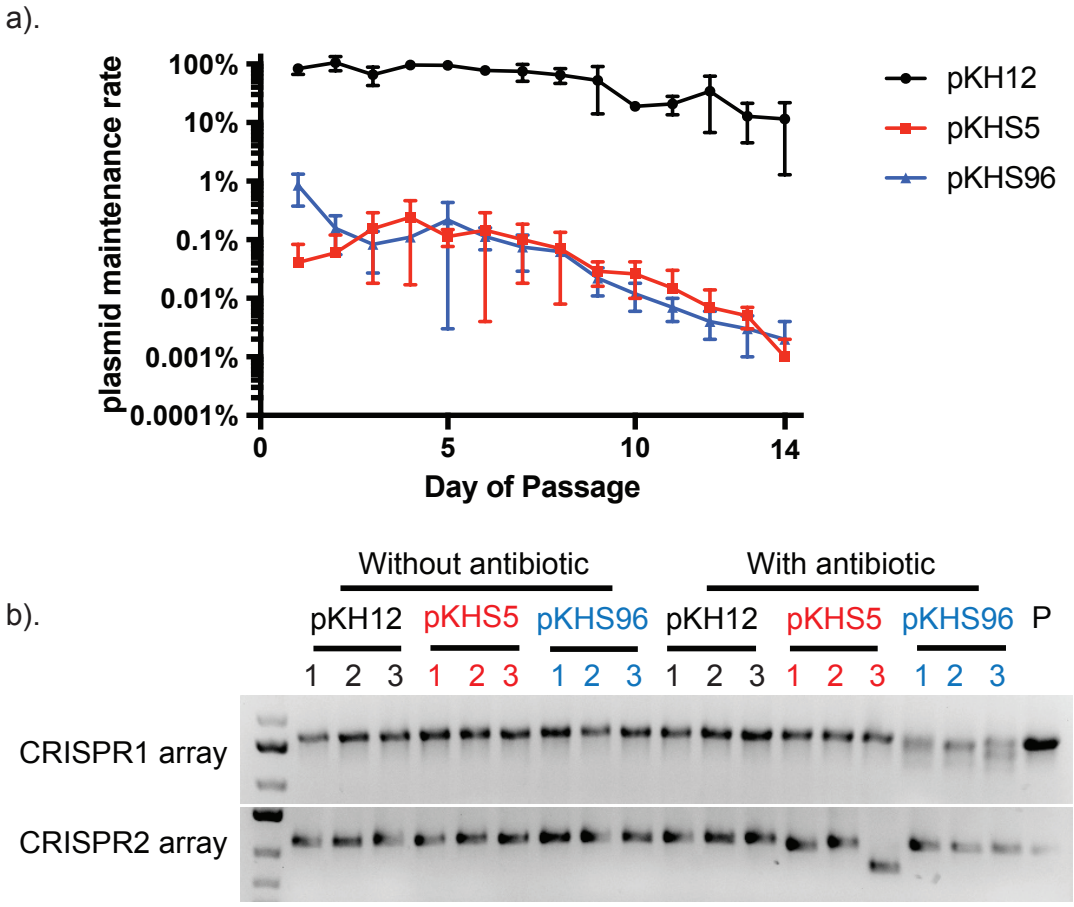


Figure 9. Antibiotic-driven CRISPR compromisation is conserved in all type II CRISPR-Cas systems in *E. faecalis*. a) Plasmid maintenance rates of OG1RF transformants passaged in the absence of chloramphenicol. Each dot represents the average rate from three transformants with the standard deviation. b) CRISPR1 and CRISPR2 amplicon PCR results from Day 14 transformant populations passaged without antibiotic (left) and with antibiotic (right). P: positive control.

951 **Table 1. Bacterial strains and plasmids used.**

Name	Description	Reference
***E. faecalis* strains**		
T11RF	Rifampicin- and fusidic acid-resistant derivative of the human urine isolate T11	(4, 37)
T11RFΔ*cas9*	Derivative of T11RF with *cas9* deleted	(37)
T11RFΔ*recA*	Derivative of T11RF with *recA* deleted	This study
OG1RF	Rifampicin- and fusidic acid-resistant derivative of the human oral isolate OG1	(29, 59)
OG1SSp	Spectinomycin- and streptomycin-resistant derivative of OG1; donor strain for conjugation assays	(39, 43, 60)
Plasmids		
pAM714	65 kb PRP encoding erythromycin on Tn*917*, derivative of pAD1	(39)
pCF10	67 kb PRP encoding tetracycline resistance on Tn*925*	(43)
pLZ12	Broad host range shuttle vector encoding chloramphenicol resistance	(61)
pKH12	pLZ12 with *oriT*	(45)
pKHS5	pKH12 with CRISPR2 protospacer S5 and CRISPR1/2 PAM	This study
pKHS96	pKH12 with CRISPR1 protospacer S96 and CRISPR1/2 PAM	(45)
pWH*recA*	pLT06 with ~750 bp up- and downstream of T11RF *recA*	
pVP107	pLT06 with T11CR2 protospacer and CRISPR1/2 PAM	(37)
pWH107	pVP107 digested with XbaI/SphI and re-ligated with primers pVP107_ XbaI _For/pVP107_SphI_Rev to remove PstI enzyme site	This study
pWH107.S1	pWH107 with T11CR3 protospacer S1 and CRISPR3 PAM inserted between BamHI/PstI	This study
pWH107.S6	pWH107 with T11CR3 protospacer S6 and CRISPR3 PAM inserted between BamHI/PstI	This study
pWH107.S7	pWH107 with T11CR3 protospacer S10 and CRISPR3 PAM inserted between BamHI/PstI	This study

952

Table 2. CRISPR alleles.

Sample name	Day 1 Sanger[c]	Day 14 Sanger[c]	Day 1 Amplicon[d]	Day 14 Amplicon[d]
T11RF control[a]	WT	WT	WT	WT
Δ4[b]	WT	WT	WT	WT
WT1[b]	Poor quality at S_6-S_7	WT	ΔS_6 ΔS_6-S_7	WT
WT2[b]	WT	ΔS_2-S_{11}	WT	ΔS_2-S_{11}
WT3[b]	Poor quality at S_6-S_7	ΔS_5-S_7	ΔS_6 ΔS_5-S_7	ΔS_5-S_7 ΔS_6
WT5[b]	Poor quality at S_3-S_8	Poor quality at S_3-S_9	ΔS_5-S_8 ΔS_3-S_{16} ΔS_6 ΔS_4-S_9 ΔS_1-S_{14}	ΔS_5-S_8 ΔS_4-S_9 ΔS_6 ΔS_6-S_{18} ΔS_3-S_{16}
WT6[b]	Poor quality at S_1-S_7	Poor quality at S_5-S_6	ΔS_5-S_6 ΔS_1-S_{17} ΔS_5-S_7 ΔS_6 ΔS_6-S_9	ΔS_6 ΔS_5-S_6 ΔS_5-S_7 ΔS_6-S_{10}

[a]T11RF without pAM714 passaged for 14 days in BHI medium.
[b]pAM714 transconjugants passaged for 14 days in BHI medium with erythromycin.
[c]CRISPR3 alleles detected by Sanger sequencing.
[d]CRISPR3 alleles detected by Illumina amplicon deep sequencing. Mutant alleles with >0.3% abundance are shown for each population and are listed from highest to lowest abundance. If no mutant alleles were detected above this threshold, "WT" is stated.

The following source data is available for Table 2:
Figure 5 - Source data 1. Percent reads mapped to artificial CRISPR references.

964 **Table 3. Nonsynonymous *cas9* mutations detected by whole genome sequencing.**

Position	Ref	Allele	Amino acid change	WT1[a]	WT2[a]	WT3[a]
652983	G	A	Gln506*	31.6% (689x)	ND (590x)	ND (577x)
653184	C	T	Glu439Lys	ND (640x)	ND (536x)	24.2% (594x)
653180	AG	T	Leu440fs	ND (640x)	ND (528x)	23.5% (590x)
654165	G	-	Arg112fs	ND (676x)	48.4% (659x)	ND (772x)

965 [a]Shown are variation frequency and coverage at the indicated nucleotide position on *E.*
966 *faecalis* T11 contig 1.11.
967 ND, not detected.
968
969 The following figure supplements and source data are available for Table 3:
970 **Table supplement 2.** SNPs detection in all gDNA sequencing samples.
971 Source data has been deposited in the sequencing reads archive under **BioProject ID:**
972 **PRJNA418345.**
973

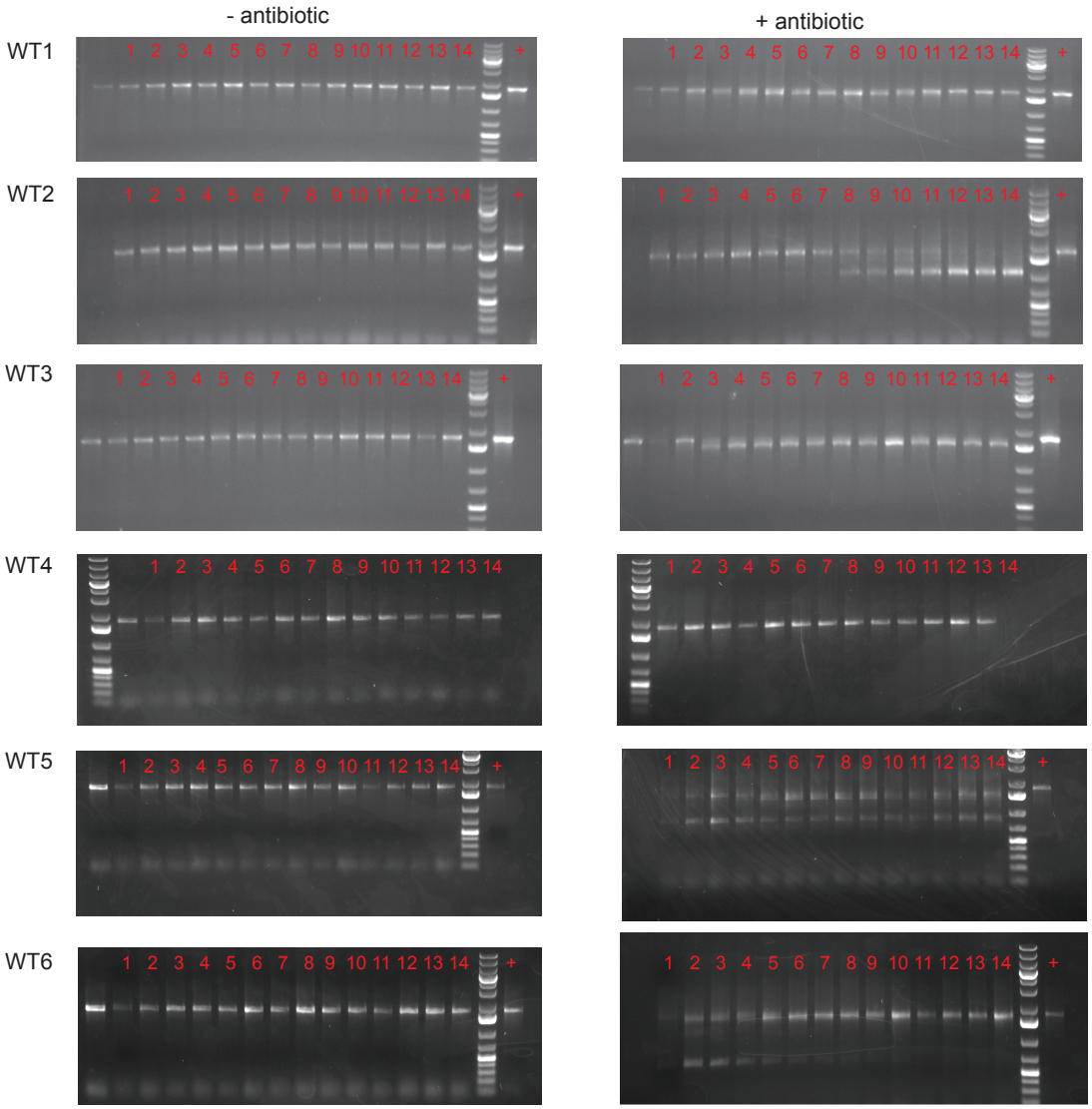


Figure supplement 1. Gel electrophoresis of CRISPR3 amplicons in T11RF pAM714 transconjugants over 14 days passaged with and without antibiotic. Six T11RF pAM714 transconjugants were serially passaged for 14 days in BHI (left panel) or BHI with erythromycin (right panel). The size of the CRISPR3 array was monitored using PCR and gel electrophoresis on each passage day.

983 **Table supplement 1. Primers used in this study.**

Name	Primer sequence
CRISPR seq primers	
CRISPR1 seq For	CGTATTTGACAGAGGATGAAG
CRISPR1 seq Rev	CGAATATGCCTGTGGTGAAA
CRISPR2 seq For	TGCTGTTACAGCTACTAAA
CRISPR2 seq Rev	GCCAATGTTACAATATCAAACA
CRISPR3 seq For	GCTCACTGTATTGGAAGAAC
CRISPR3 seq Rev	CATCGATTCATTATTCCTCCAA
pWH107 and derivatives	
pVP107_XbaI_For	CTAGAGATAATATATCTTTTATATAGAAGATGGGTACCATGGCATG
pVP107_SphI_Rev	CCATGGTACCCATCTTCTATATAAAAGATATATTATCT
T11CR3sp6_BamPst_F	GATCCTCCCGATACAGCTCTTTATTCTTCTAATTACATTGTACTGCA
T11CR3sp6_BamPst_R	GTACAATGTAATTAGAAGAATAAAGAGCTGTATCGGGAG
T11CR3sp1_BamPst_F	GATCCTCAAAAGTTGAATATGTTTCGCTTTGGTGTAATTGTACTGCA
T11CR3sp1_BamPst_R	GTACAATTACACCAAAGCGAAACATATTCAACTTTTGAG
T11CR3sp7_BamPst_F	GATCCTCTGCTTTTGGAGGAATACAAATGAGAAGATTTTGTACTGCA
T11CR3sp7_BamPst_R	GTACAAAATCTTCTCATTTGTATTCCTCCAAAAGCAGAG
T11CR2S1 seq arm1 F	CGAAATCAGCACATGGAACA
T11CR2S1 seq arm2 R	CCAGTAACTGTATCAACTAC

984

Table supplement 2. SNPs detection in all gDNA sequencing samples.

Contig	Position	Ref	Allele	Annotation	Amino acid change	T11RF[a]	WT3[b]	WT1[c]	WT2[c]	WT3[c]	WT5[c]	WT6[c]	Δ4[c]
AD1REPABC	1904	C	T	Replication protein	AAB00504.1: p.Pro100Leu				45.49				
1.11	51613	C	T	Collagen adhesin								25.00	
1.11	51674	G	A	Collagen adhesin	WP_002379313.1: p.Gly392Ser						33.90	33.39	
1.11	51745	A	G	Collagen adhesin							33.02		
1.11	337315	G	T	Helicase	WP_002382285.1: p.Glu425*							37.16	
1.11	652983	G	A	Cas9	WP_002379510.1: p.Gln506*			31.64					
1.11	653180	AG	T	Cas9	WP_002379510.1: p.Leu440fs		22.74			23.52			
1.11	653184	C	T	Cas9	WP_002379510.1: p.Glu439Lys		23.21			24.17			
1.11	654165	G	-	Cas9	WP_002379510.1: p.Arg112fs				48.41				
1.11	685523	A	T	Conserved hypothetical protein	WP_002382413.1: p.Glu5Val		23.24						
1.11	685938	A	T	Conserved hypothetical protein	WP_002382413.1: p.Glu143Asp		26.00						
1.11	974866	AT	GC	Cell wall surface anchor family protein	WP_002379620.1: p.Ile972Ala							23.41	
1.11	974991	A	G	Cell wall surface anchor family protein								22.67	
1.2	2903	C	T	Peptidylprolyl isomerase	WP_002355150.1: p.Val167Ile				46.03				

[a]Day 14 BHI-passaged

[b]Day 1 Erythromycin-passaged population WT3 to confirm presence of *cas9* mutations prior to Day 14

[c]Day 14 Erythromycin-passaged

Source data for Table supplement 2 has been deposited in the sequencing reads archive under **BioProject ID: PRJNA418345.**

Table supplement 3. Quality control of the amplicon sequencing reads.

Sample name	# of Total reads	Total mapped reads	# mapped to WT references	# mapped to mutant references	% reads mapped in Step 1	% reads mapped in Step 2	% reads mapped in Step 3	% reads mapped in Step 4	% reads left
Day 1									
T11RF	19,082,652	7398495	7273291	125204	97.543%	0.642%	0.543%	0.865%	0.408%
Δ4	20,808,404	9073461	8995910	77551	98.155%	0.362%	0.616%	0.702%	0.165%
WT1	16,736,162	6976811	6808760	168051	97.629%	0.985%	0.544%	0.709%	0.132%
WT2	16,859,152	7281948	7222339	59609	98.292%	0.344%	0.562%	0.664%	0.139%
WT3	17,357,970	7387447	7136664	250783	97.141%	1.416%	0.626%	0.672%	0.145%
WT5	17,744,230	7215381	6784881	430500	96.024%	2.393%	0.699%	0.747%	0.137%
WT6	16,179,400	6744053	6337706	406347	95.871%	2.488%	0.814%	0.672%	0.154%
Day 14									
T11RF	17,973,946	7829494	7760594	68900	98.233%	0.373%	0.601%	0.649%	0.143%
Δ4	14,403,868	5424100	5369063	55037	98.173%	0.445%	0.625%	0.529%	0.227%
WT1	14,158,158	5275288	5219558	55730	98.136%	0.454%	0.624%	0.545%	0.240%
WT2	14,538,836	4744257	4489642	254615	96.340%	2.166%	0.707%	0.527%	0.260%
WT3	17,174,946	6039118	5653212	385906	96.018%	2.538%	0.738%	0.514%	0.192%
WT5	15,200,646	5844417	5384356	460061	95.193%	2.998%	0.885%	0.738%	0.187%
WT6	12,325,042	4505777	4201804	303973	94.419%	2.735%	1.501%	1.005%	0.340%

Source data for Table supplement 3 has been deposited in the sequencing reads archive under **BioProject ID: PRJNA418345.**