Cell Reports, Volume *37*

Supplemental information

Phage-delivered CRISPR-Cas9 for strain-specific depletion and genomic deletions in the gut microbiome

Kathy N. Lam, Peter Spanogiannopoulos, Paola Soto-Perez, Margaret Alexander, Matthew J. Nalley, Jordan E. Bisanz, Renuka R. Nayak, Allison M. Weakley, Feiqiao B. Yu, and Peter J. Turnbaugh

Supplemental Figures

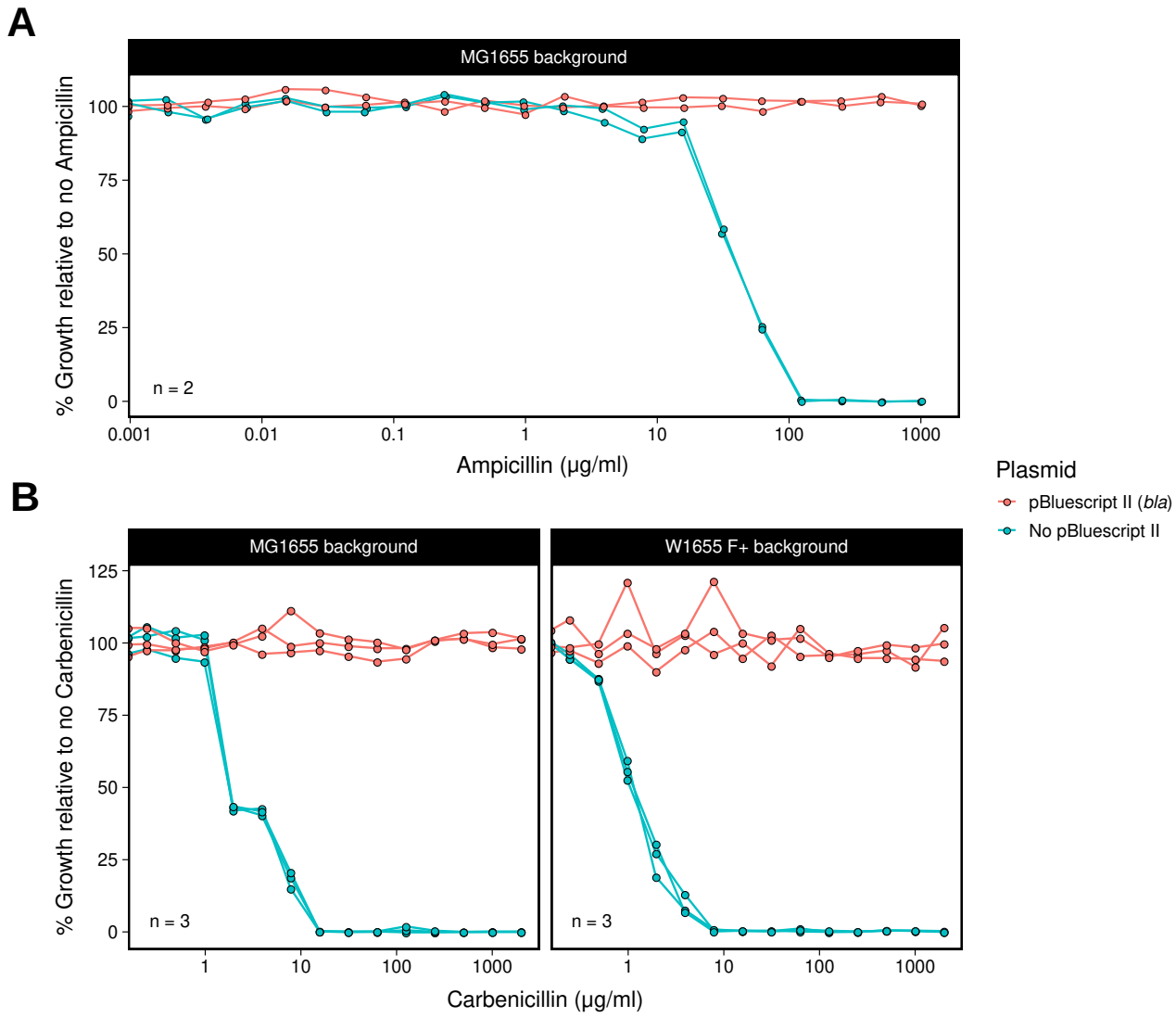

Figure S1. The plasmid pBluescript II confers resistance to beta-lactam antibiotics exceeding 1 mg/ml. **(A)** Minimum inhibitory concentration (MIC) assay for ampicillin in the *E. coli* MG1655 background. Harbouring pBluescript II, the strain exhibits an MIC of >1 mg/ml; in the absence of the plasmid, the MIC is approximately 100 µg/ml (n=2 biological replicates). **(B)** pBluescript II confers resistance to carbenicillin exceeding 2 mg/ml in both the *E. coli* MG1655 and W1655 F+ backgrounds; in the absence of the plasmid, the same strains have an MIC of approximately 10 µg/ml (n=3 biological replicates per strain). *bla*, beta-lactamase gene. Related to Figure 1.

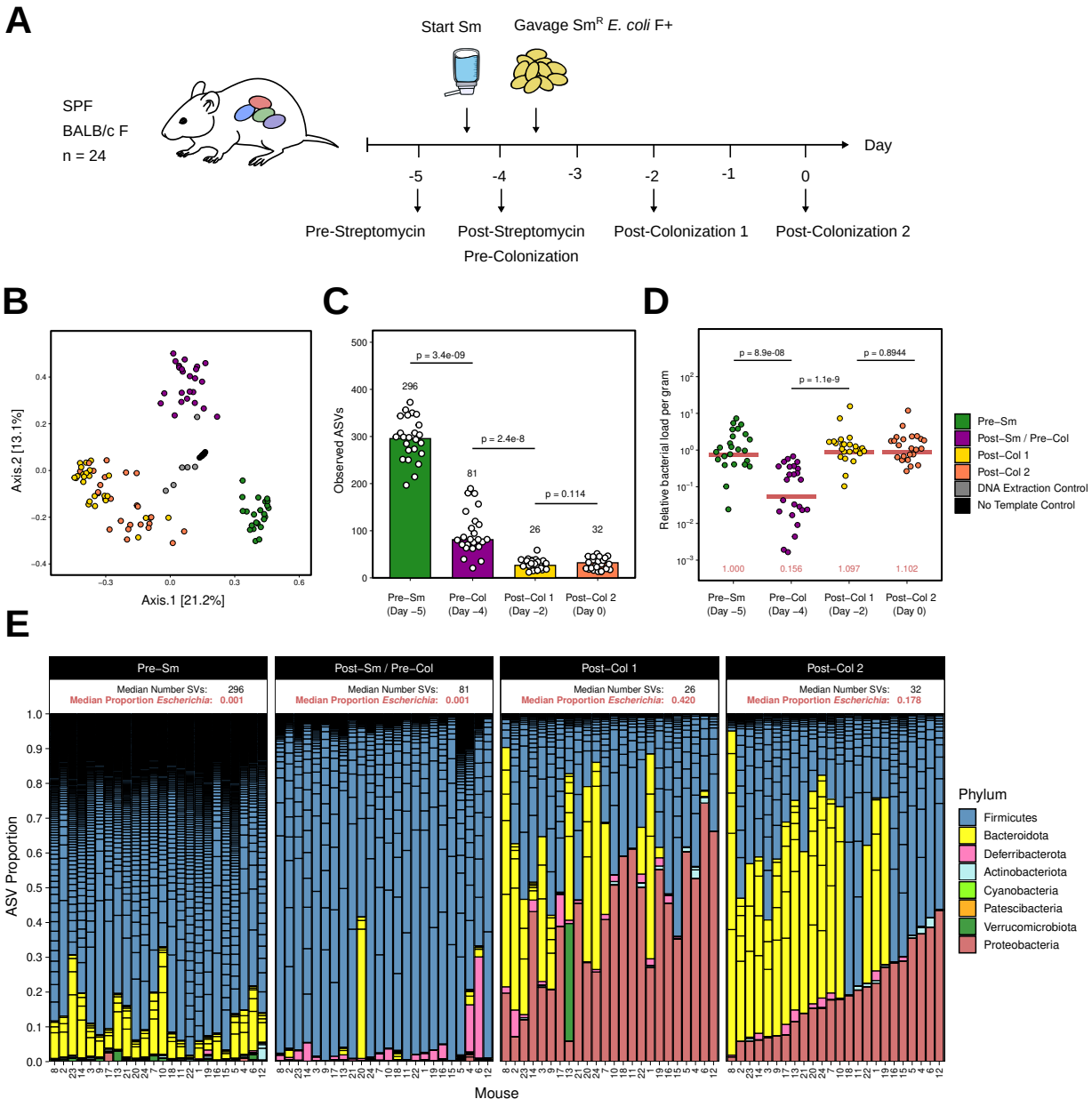

Figure S2. Streptomycin treatment reduces bacterial diversity and allows Sm^R *E. coli* to colonize at a high proportion in conventionally raised mice. **(A)** Typical timeline for a mouse experiment using water containing streptomycin (Sm) to allow colonization by Sm^R *E. coli*. Fecal samples of individually caged mice (n=24) were collected before streptomycin (Day -5), after streptomycin but before gavage of *E. coli* (Day -4), and two timepoints after *E. coli* (Day -2 and 0). **(B)** Principle coordinate analysis using Bray-Curtis dissimilarity indicates that fecal samples from Pre-Sm, Post-Sm/Pre-Colonization (Pre-Col), and Post-Colonization (Post-Col) are distinct. **(C)** Number of observed 16S rRNA gene amplicon sequence variants (ASVs) was lower in Post-Sm relative to Pre-Sm timepoint, and lower still in both Post-Col timepoints than Pre-Sm and Post-Sm/Pre-Col. Numbers above bars indicate median. **(D)** The bacterial load in fecal samples (based on qPCR of 16S rRNA gene) transiently decreased during the course of treatment. Red numbers and horizontal bar indicate median for each timepoint. **(E)** Proportion of individual ASVs detected in each mouse at the different timepoints, coloured by phylum. Where ASV count is high, e.g., in Pre-Sm timepoint, stacked bars appear to fade to black due to a large number of low abundance ASVs. For each group, the median number of ASVs as well as the median proportion of ASVs classified as *Escherichia-Shigella* is indicated. p-value, Mann-Whitney test. Related to Figure 1.

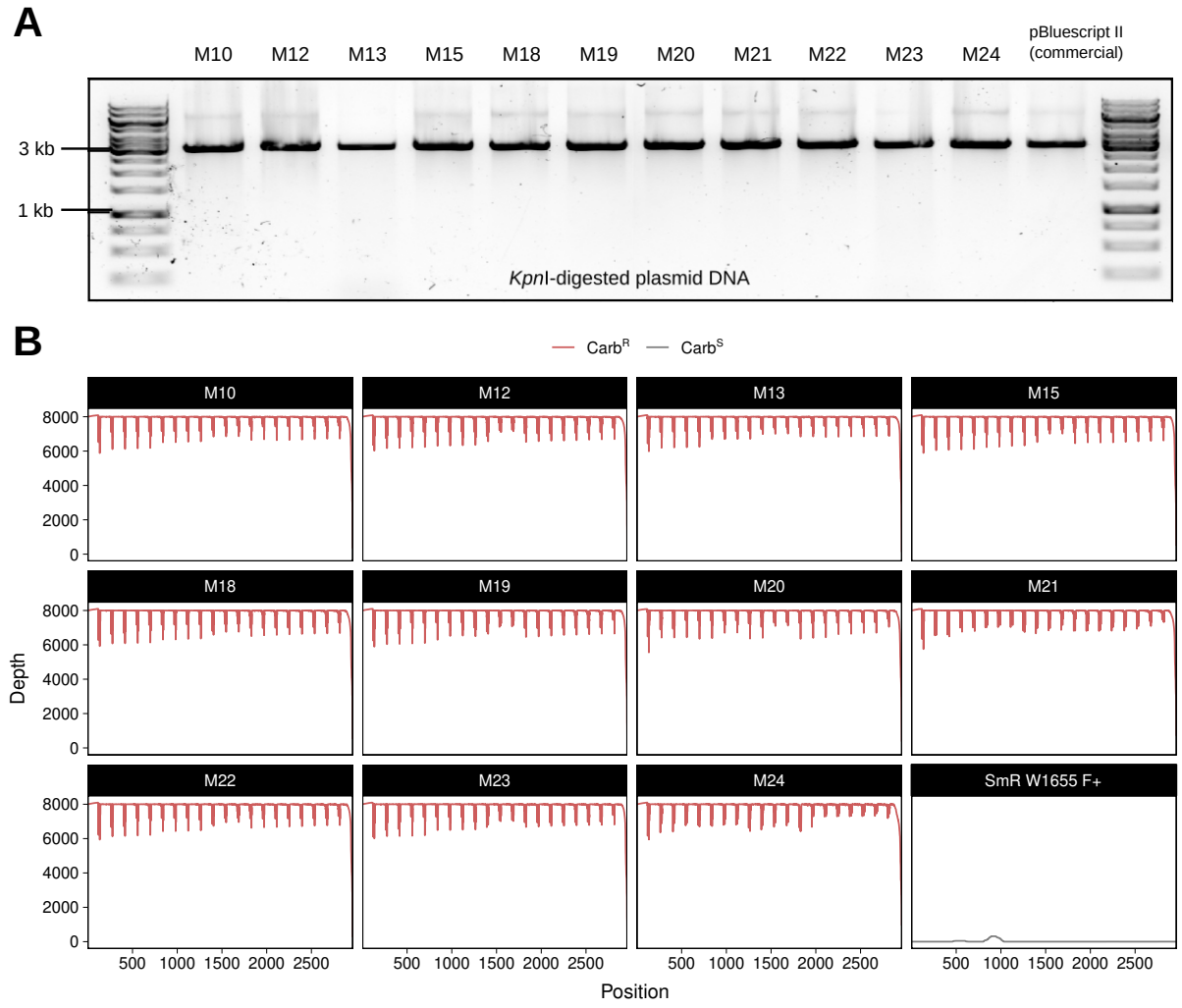

Figure S3. Successful transfer of pBluescript II plasmid from M13 phage to *E. coli* cells *in vivo*. **(A)** Diagnostic digest of extracted plasmid DNA from fecal isolates is consistent with pBluescript II. Plasmid DNA was recovered from carbenicillin-resistant ($Carb^R$) colonies isolated from the feces of the 11 mice that were successfully colonized during treatment with carbenicillin in the water (Fig. 1D); DNA was digested with restriction enzyme KpnI for comparison to linearized 3-kb pBluescript II. **(B)** Genome sequencing results from $Carb^R$ isolates confirms presence of pBluescript II. Genomic DNA was extracted from these isolates and sequenced; reads were mapped to the reference W1655 F+ genome and pBluescript II; read depth for the 2961-bp plasmid is shown. As a negative control, the Sm^R W1655 F+ strain used to colonize these mice at the start of the experiment [prior to treatment with M13(pBluescript II)] was sequenced in the same batch. Related to Figure 1.

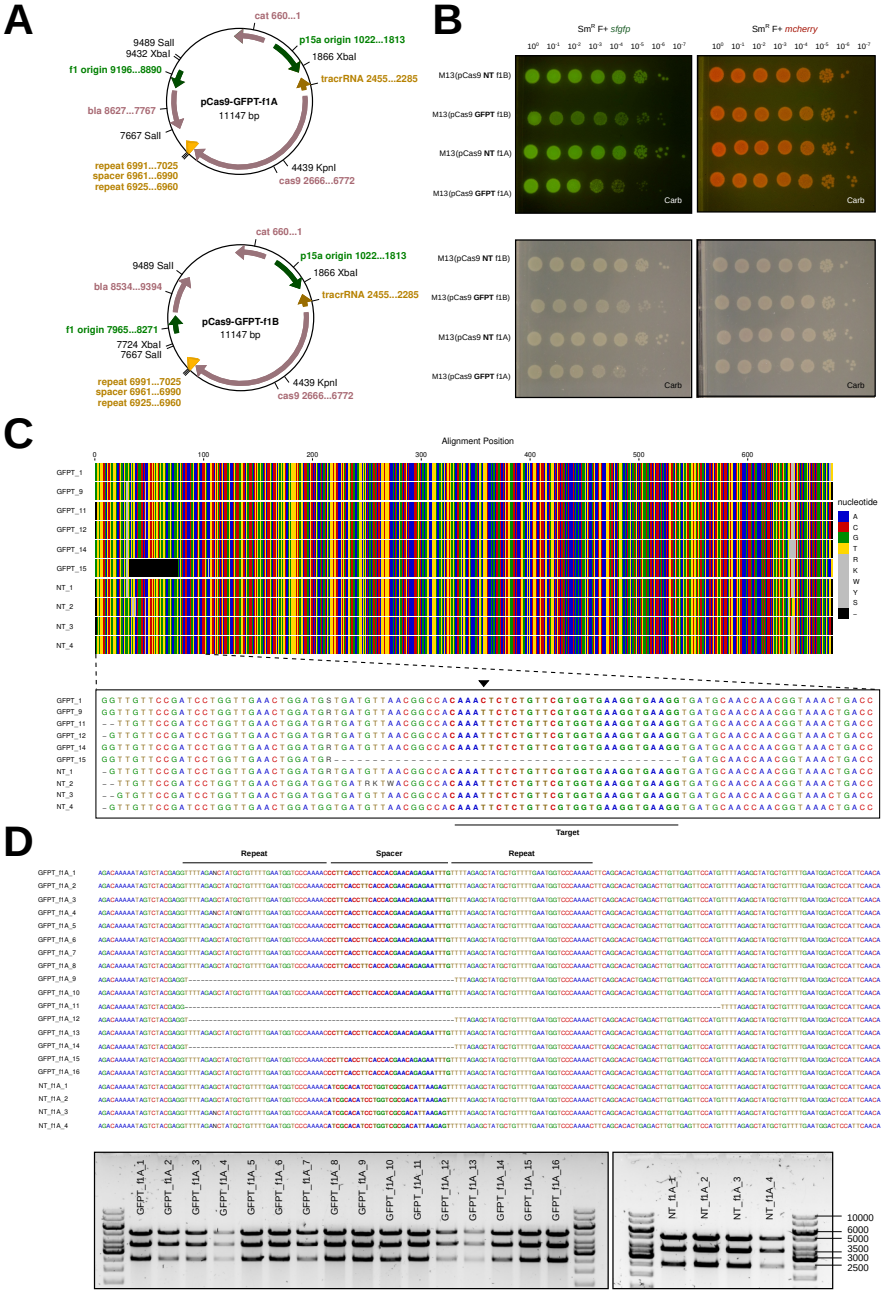

Figure S4. Delivery of GFP-targeting (GFPT) CRISPR-Cas9 phagemids by M13 and mechanisms of escape from CRISPR-Cas9 targeting in vitro. (A) GFPT phagemids. The non-targeting (NT) versions of these vectors (not shown) are identical to the GFPT vectors except in the spacer sequence. The f1-*bla* fragment was cloned as a SalI fragment in both possible orientations, designated f1A (top) or f1B (bottom), for either strand of DNA to be packaged into M13. *cat*, chloramphenicol acetyltransferase (Cm^R); *bla*, beta-lactamase ($Carb^R$). **(B)** GFP-marked *E. coli* exhibits impaired colony growth after infection with GFPT-M13. Growth impairment of the strain under GFPT conditions was evident under blue light (top) and colonies exhibited a translucent quality that was more pronounced under white light (bottom). **(C)** Escape by mutation in the target sequence. Sanger sequencing of *sfgfp* PCR amplicons from streak-purified clones after treatment with GFPT-M13 or NT-M13 *in vitro* (Fig. 3C) confirmed the partial loss observed by gel electrophoresis for clone GFPT 15. Pullout: the lost region of the *sfgfp* coding sequence from clone GFPT 15 encompasses the target site. Closer examination of clone GFPT 1 revealed the presence of a single nucleotide change in the target site (black arrow) that allowed this clone to escape targeting yet remain fluorescent. **(D)** Spacer loss from GFPT phagemids. Sanger sequencing for spacer presence in phagemid DNA isolated from clones after treatment with GFPT-M13 or NT-M13. All 4 clones isolated after infection with NT-M13 retained the spacer. Of 16 clones isolated after infection with GFPT-M13, 4 had lost the spacer (clones 9, 11, 12, and 14). Diagnostic digest of plasmid DNA isolated from clones using KpnI and XbaI revealed phagemid DNA of the expected size. Expected fragments: 4993, 3581, and 2573 bp. Related to Figure 3.

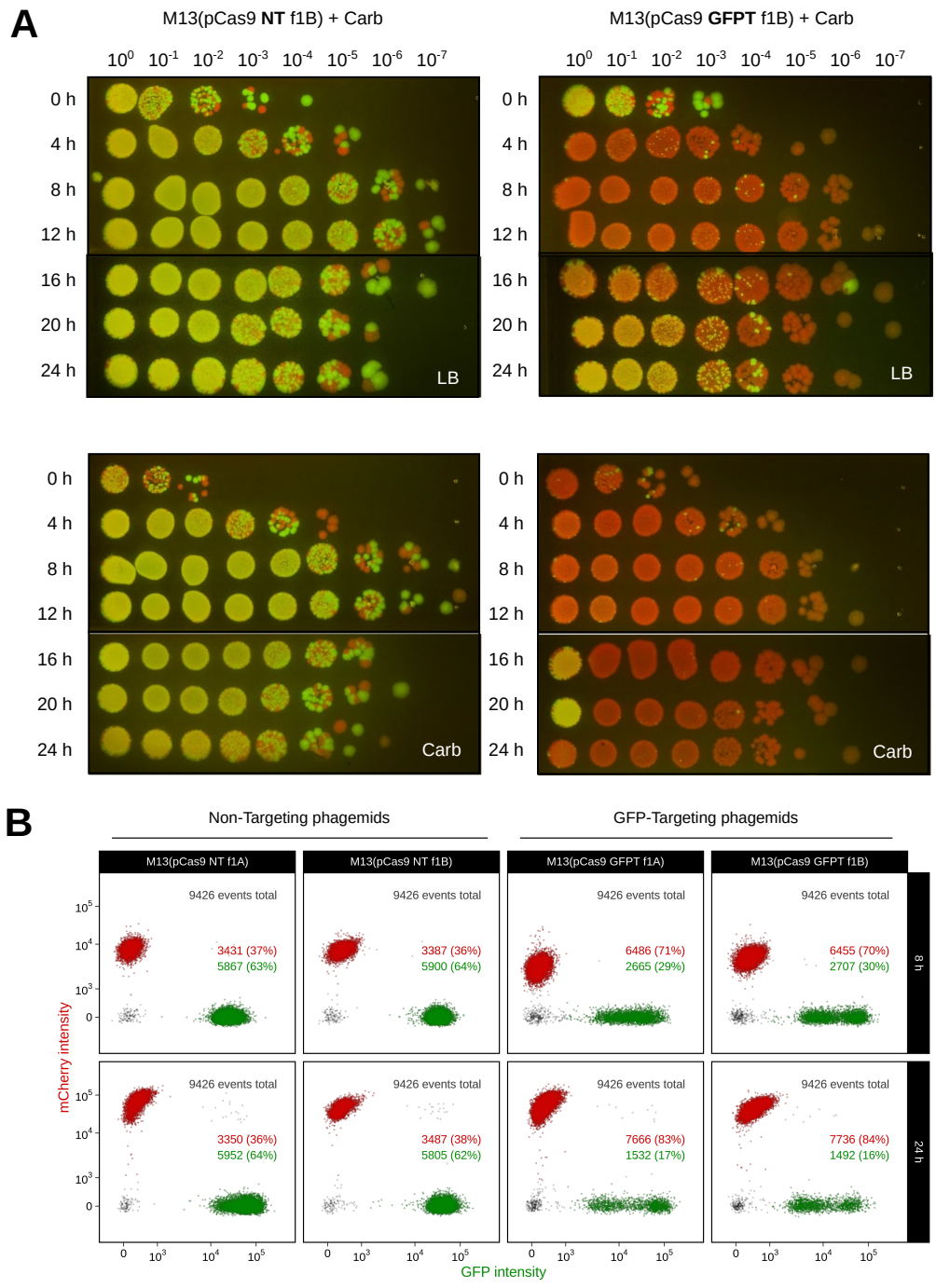

Figure S5. Culturing and flow cytometry from *in vitro* co-culture of GFP-marked and mCherry-marked *E. coli* in the presence of GFPT-M13 show that the GFP+ strain can be selectively targeted *in vitro*. **(A)** Recovery of GFP+ cells at later timepoints after infection with GFPT-M13 is likely due to lack of selection for the CRISPR-Cas9 phagemid. On non-selective media (top), GFP fluorescent colonies are detected at later timepoints of the co-culture infected with GFPT-M13. Lack of GFP fluorescent colonies after testing the same co-culture on media with carbenicillin (bottom) indicates that those GFP+ colonies at later timepoints derive from cells that are $Carb^S$ and suggests that they do not harbour the CRISPR-Cas9 phagemid. **(B)** The GFP+ strain shows a decrease in relative abundance at 24 h after treatment with GFPT-M13 when compared to the 8 h timepoint. Co-cultures of GFP-marked and mCherry-marked strains were infected with NT-M13 or GFPT-M13, carbenicillin was added to select for phage infection, and fluorescent populations were assayed by flow cytometry. The relative abundance of GFP+ events is decreased in GFPT conditions at 8 h and further decreased by 24 h. Non-targeting phagemids are pCas9-NT-f1A and pCas9-NT-f1B; GFP-targeting phagemids are pCas9-GFPT-f1A and pCas9-GFPT-f1B. Related to Figure 4.

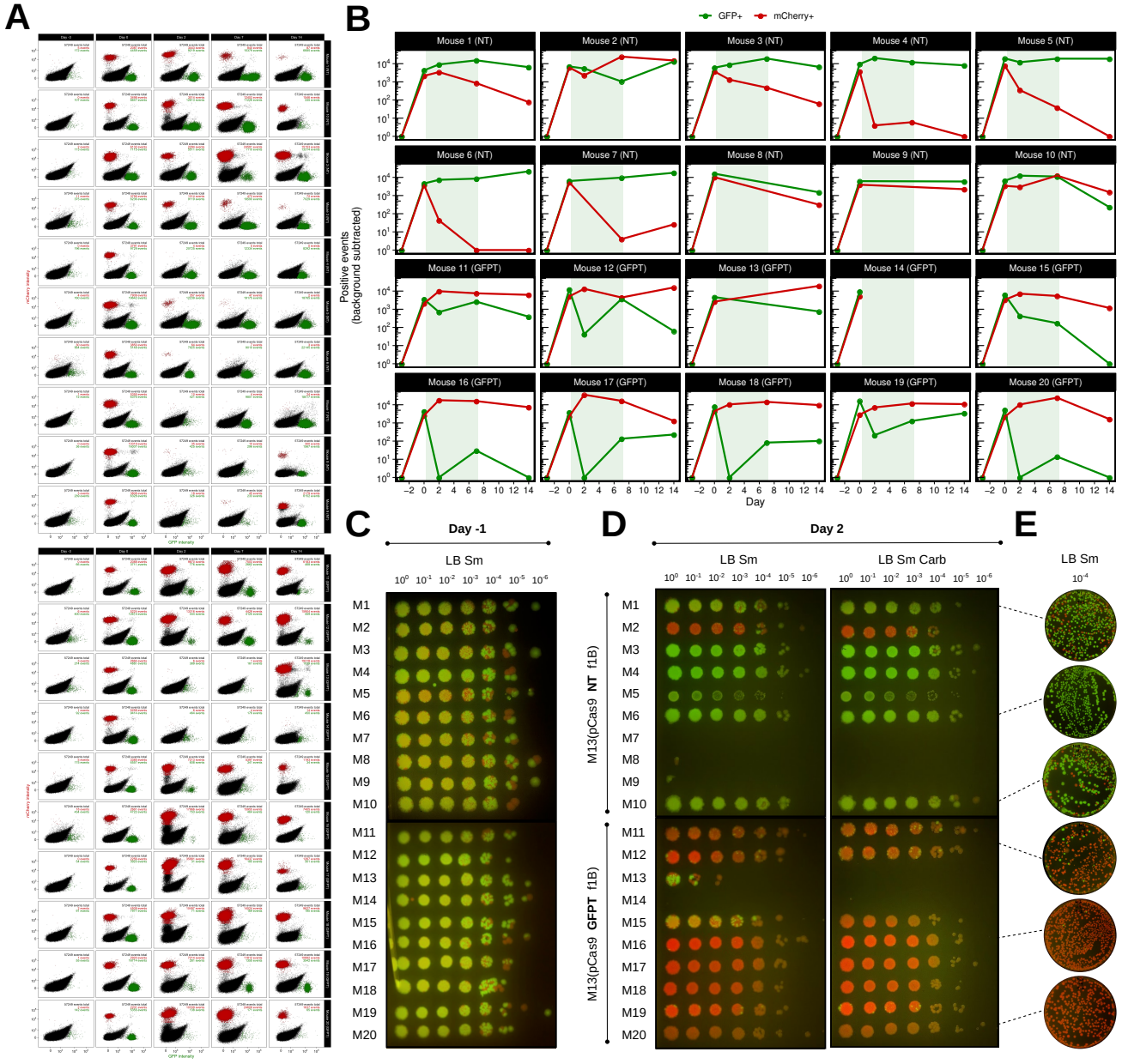

Figure S6. Flow cytometry and culturing from fecal samples of mice co-colonized with GFP-marked *E. coli* in competition with mCherry-marked *E. coli* and treated with GFPT-M13 show that the GFP+ strain can be specifically targeted in vivo. (A) Flow cytometry plots of fecal samples for all mice at all timepoints. Mice (n = 10 per group) were given either NT-M13 (top) or GFPT-M13 (bottom). Day -3, before colonization by *E. coli*; Day 0, after colonization by both GFP+ and mCherry+ strains; Day -2, post phage and carbenicillin treatment; Day 7, one week post-phage and carbenicillin; Day -14, one week after removing carbenicillin from drinking water. **(B)** Fluorescence in fecal samples for individual mice. Time course flow cytometry results for mice treated with NT-M13 (M1 to M10) or GFPT-M13 (M11 to M20). For each mouse, the number of positive events on Day -3 (before *E. coli* colonization) was used to subtract background for all subsequent timepoints. Shaded green area indicates duration of carbenicillin treatment. Timepoints were excluded when both mCherry+ and GFP+ events were below background thresholds. **(C)** Culturing on LB streptomycin from fecal samples of mice co-colonized with GFP-marked and mCherry-marked strains on Day -1 (before phage treatment) confirms colonization by both strains. **(D)** After treating with phage and carbenicillin to select for infection, culture from fecal samples on Day 2 shows that GFPT mice have decreased GFP fluorescence. Culturing from the same samples on LB with both streptomycin and carbenicillin suggests that for some mice, fluorescent colonies arising on LB streptomycin are $Carb^S$, i.e., that they do not carry the CRISPR-Cas9 phagemid. Lack of fluorescent *E. coli* in fecal samples indicates eradication by carbenicillin where phage infection leading to colonization by $Carb^R$ *E. coli* has not occurred. **(E)** Day 2 fecal samples from a subset of the mice (M1, M6, M10 for NT; M11, M16, M20 for GFPT) were cultured on larger plates for confirmation. Related to Figure 5.

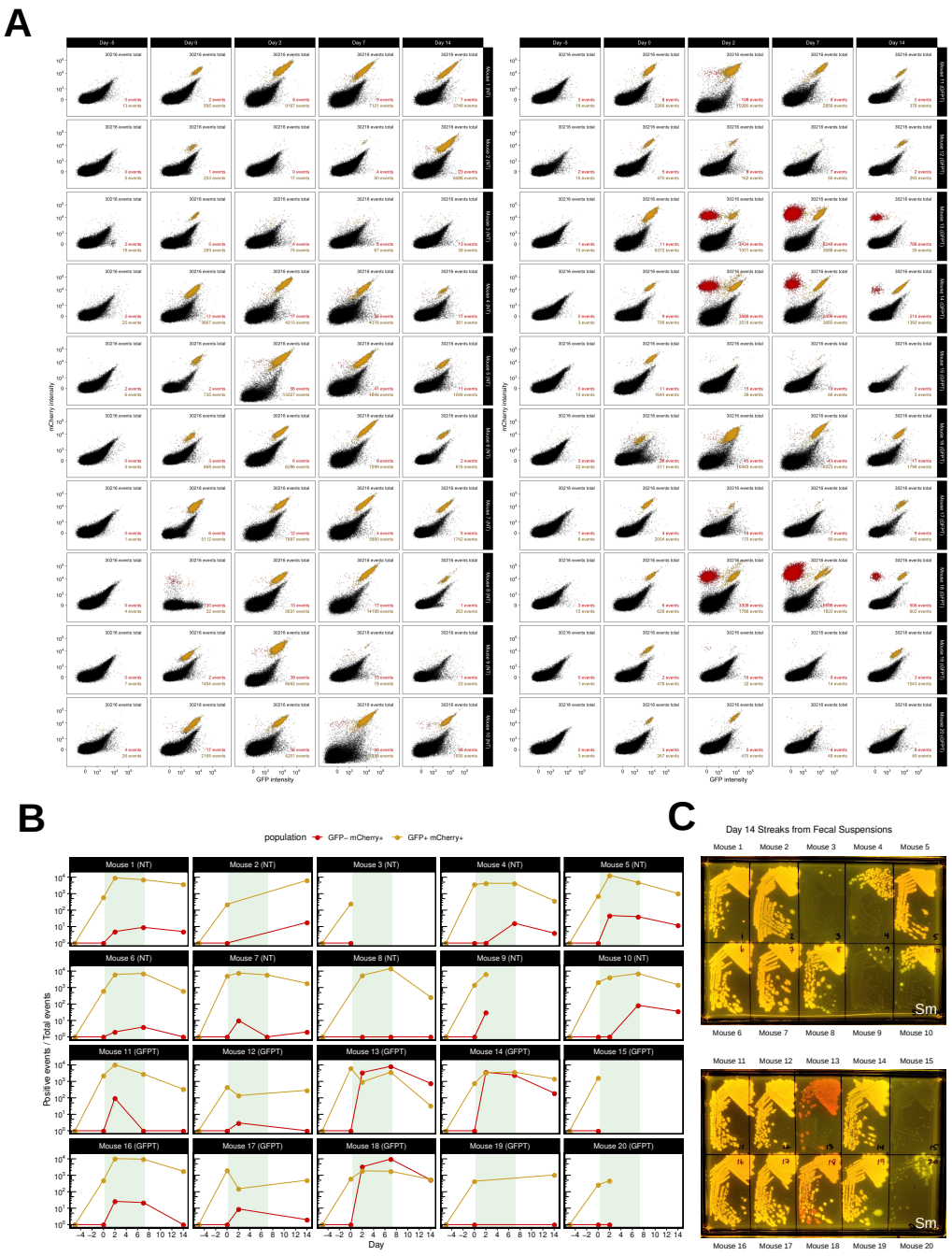

Figure S7. Flow cytometry and culturing from fecal samples of mice colonized with GFP+ mCherry+ *E. coli* and treated with GFPT-M13 show that loss of the target gene can occur *in vivo*. **(A)** Flow cytometry plots of fecal samples for all mice at all timepoints during *in vivo* targeting of double-marked *E. coli*. Mice (n=10/group) were given either NT-M13 (left) or GFPT-M13 (right). Day -5, before colonization by *E. coli*; Day 0, after colonization by double-marked GFP+ mCherry+ *E. coli*; Day -2, post phage and carbenicillin treatment; Day 7, one week post-phage and carbenicillin; Day -14, one week after removing carbenicillin from drinking water. Based on visual inspection, the sample from Mouse 8 Day 0 was omitted from analyses. **(B)** Fluorescence in fecal samples for individual mice. For each mouse, the number of GFP+ mCherry+ events on Day -5 (before *E. coli* colonization) was used to subtract GFP+ mCherry+ background for all subsequent timepoints, and the number of GFP− mCherry+ events on Day 0 (before phage treatment) was used to subtract mCherry+ background from all subsequent timepoints. Shaded green area indicates duration of carbenicillin treatment. Timepoints were excluded when both GFP+ mCherry+ and GFP−mCherry+ events were below background thresholds. **(C)** Culture on LB streptomycin from Day 14 fecal suspensions of mice treated with NT-M13 (M1 to M10, top) and GFPT-M13 (M11 to M20, bottom). Lack of fluorescent *E. coli* in fecal samples indicates eradication by carbenicillin where phage infection leading to colonization by Carb^{R} *E. coli* did not occur. Related to Figure 6.

Supplemental Tables

Table S1. Oligonucleotides, related to the STAR Methods section.

Sequence (5'–3')	Purpose*	Source	Identifier
CGTGGCATGGAAATACTCCG	F primer to amplify *rpsL* for recombineering	This study	PS-rpsL1
GCATCGCCCTAAAATTCGGC	R primer to amplify *rpsL* for recombineering	This study	PS-rpsL2
AAACCCTTCACCTTCACCACGAACAGAGAATTTG	Oligo 1 to generate GFP-targeting (GFPT) spacer	This study	PSP116
AAAACAAATTCTCTGTTCGTGGTGAAGGTGAAGG	Oligo 2 to generate GFP-targeting (GFPT) spacer	This study	PSP117
AAACATCGCACATCCTGGTCGCGACATTAAGAGT	Oligo 1 to generate Non-targeting (NT) spacer	This study	PSP120
AAAAACTCTTAATGTCGCGACCAGGATGTGCGAT	Oligo 2 to generate Non-targeting (NT) spacer	This study	PSP121
TTAATAAATGCAGTAATACAGG	Primer to Sanger sequence spacer in CRISPR array	This study	PSP108
CCTGTCGACGGTATCGATAAGCTTGATATCG	F primer to clone f1-*bla* segment from pBluescript II as SalI fragment	This study	KL215
CCTGTCGACGATTATCAAAAAGGATCTTCACCTAGATCC	R primer to clone f1-*bla* segment from pBluescript II as SalI fragment	This study	KL216
CTGTTCACCGGTGTTGTTCC	F primer to amplify *sfgfp* fragment	This study	KL207
TTATTTGTAGAGTTCATCCATGCCG	R primer to amplify *sfgfp* fragment	This study	KL200
ACTCCTACGGGAGGCAG	F primer to amplify 16S rRNA gene fragment	Yu et al., 2005	BAC338F
GACTACCAGGGTATCTAATCC	R primer to amplify 16S rRNA gene fragment	Yu et al., 2005	BAC805R

TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGTGCCAGCMGCCGCGGTAA	F primer for 16S rRNA gene sequencing primary PCR	Gohl et al., 2016	V4 515F Nextera
GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGGACTACHVGGGTWTCTAAT	R primer for 16S rRNA gene sequencing primary PCR	Gohl et al., 2016	V4 806R Nextera
AATGATACGGCGACCACCGAGATCTACACNNNNNNNNTCGTCGGCAGCGTC	F primer for 16S rRNA gene sequencing indexing PCR	Gohl et al., 2016	Forward indexing primer
CAAGCAGAAGACGGCATACGAGATNNNNNNNNGTCTCGTGGGCTCGG	R primer for 16S rRNA gene sequencing indexing PCR	Gohl et al., 2016	Reverse indexing primer

**bla*, beta-lactamase