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

Kathy N. Lam[1], Peter Spanogiannopoulos[1], Paola Soto-Perez[1], Margaret Alexander[1], Matthew J. Nalley[1], Jordan E. Bisanz[1], Renuka R. Nayak[1], Allison M. Weakley[2], Feiqiao B. Yu[3], Peter J. Turnbaugh[1,3,*]

[1]Department of Microbiology and Immunology, University of California San Francisco, USA
[2]Stanford Microbiome Therapies Initiative, Stanford University, Stanford, USA
[3]Chan-Zuckerberg BioHub, San Francisco, USA

*Correspondence to: Peter.Turnbaugh@ucsf.edu

Mechanistic insights into the role of the human microbiome in the predisposition to and treatment of disease are limited by the lack of methods to precisely add or remove microbial strains or genes from complex communities. Here, we demonstrate that engineered bacteriophage M13 can be used to deliver DNA to *Escherichia coli* within the mouse gastrointestinal (GI) tract. Delivery of a programmable exogenous CRISPR-Cas9 system enabled the strain-specific depletion of fluorescently marked isogenic strains during competitive colonization and genomic deletions that encompass the target gene in mice colonized with a single strain. Multiple mechanisms enabled *E. coli* to escape targeting, including loss of the CRISPR array or even the entire CRISPR-Cas9 system. These results provide a robust and experimentally tractable platform for microbiome editing, a foundation for the refinement of this approach to increase targeting efficiency, and a *proof-of-concept* for the extension to other phage-bacterial pairs of interest.

Current strategies for manipulating the microbiome either lack species- or strain-level precision[1,2] or require the introduction of an exogenous bacterium into the host[3,4]. Pioneering studies of pathogenic bacteria that colonize the skin[5–7] and gut[8] support the potential for the use of engineered bacteriophage carrying an exogenous CRISPR-Cas system that could be directed to any target of interest; however, these methods have yet to be applied to the human or mouse microbiome. Given the tremendous diversity within both the bacterial[9] and viral[10] components of the human gut microbiota, we sought to establish a tripartite model system that

30 builds upon tools for the genetic manipulation of a bacteriophage and its bacterial target coupled
31 to an experimentally tractable mammalian host.

32 We focused on M13, a single-stranded DNA (ssDNA) filamentous inovirus[11,12] able to
33 replicate and release virions without causing cell lysis[13]. M13 infects *E. coli* and related
34 Enterobacteriaceae carrying the F sex factor necessary to form the conjugative F pilus[14,15]. M13
35 phagemid vectors combine the advantages of plasmid DNA manipulation with the ability to
36 easily package recombinant DNA into virions[16]. M13 has been used to target *E. coli*[17,18] and
37 *Helicobacter pylori*[19], including engineered M13 carrying CRISPR-Cas9 in an insect model of
38 bacterial infection[5]. However, the use of M13 to deliver genetic constructs (including
39 CRISPR-Cas systems) to cells within the mammalian gastrointestinal tract had not been
40 previously demonstrated. We also leveraged the streptomycin (Sm)-treated mouse model for
41 *E. coli* colonization within the mouse gut microbiota[20], providing a robust and accessible model
42 that could be used by any group with access to a mouse colony.

43 Results

44 **Bacteriophage M13 enables the delivery of DNA to the gut microbiome.** We utilized
45 phagemid pBluescript II[21] carrying the *bla* (β-lactamase) gene and a β-lactam antibiotic in the
46 drinking water to select for successfully infected *E. coli*. pBluescript II conferred *in vitro*
47 resistance to ampicillin and the semi-synthetic analogue carbenicillin at concentrations
48 exceeding 1 mg/ml (**Supplementary Fig. 1**). We used streptomycin-resistant (Sm^R) *E. coli* to
49 colonize the GI tract of Sm-treated mice. As expected, streptomycin altered gut microbial
50 community structure while decreasing diversity and colonization level (**Supplementary Figs.**
51 **2a-d**). Sm^R *E. coli* colonized at a high proportion (median 18% of the gut microbiota; range
52 1.4-43%) four days after gavage (**Supplementary Fig. 2e**). We introduced a Sm^R *E. coli*
53 population that was a mixture of 99.9% Amp^S (no plasmid) and 0.1% Amp^R cells (pBluescript II),
54 split the mice into two groups with access to water containing only streptomycin or both
55 streptomycin and ampicillin, and tracked both total *E. coli* and Amp^R *E. coli* in mouse feces. At
56 6 h post-*E. coli* introduction, the percentage of Amp^R *E. coli* in the feces of all mice was at or
57 close to 0.1%, consistent with the gavaged mixture transiting through the GI tract. Within 1-2
58 days, mice on water containing ampicillin exhibited an increase in the percent of Amp^R *E. coli* by
59 3 orders of magnitude, reaching complete or near complete colonization (**Fig. 1a**). In contrast,

the Amp^R subpopulation was lost in mice on water without ampicillin. These results demonstrate that β-lactam antibiotics can be used to select for resistant *E. coli.*

Antibiotics were capable of eradicating a sensitive population of *E. coli* that had established stable colonization in the mouse gut. We colonized Sm-treated mice with Sm^R *E. coli* MG1655 or W1655 F+ and tracked colonization levels during treatment with the β-lactam antibiotic carbenicillin. Carbenicillin decreased the median *E. coli* colonization level from 9.6×10^9 to 2.0×10^3 CFU/gram feces in the first day, and levels decreased to below our limit of detection (~10^2 CFU/g) in all mice over the course of treatment (**Fig. 1b**). When selection was lifted on Day 7, recolonization was observed for 5/6 mice. When carbenicillin was reintroduced on Day 13, colonization again dropped below our limit of detection. The low background of *E. coli* in the gut during carbenicillin treatment, as well as the lack of spontaneous resistant cells able to expand, supports the utility of this model for assessing the phage-mediated delivery of a resistance gene.

Next, we tested our ability to deliver an antibiotic resistance gene to *E. coli* within the gut. We colonized Sm-treated mice with either Sm^R *E. coli* W1655 F+ ($M13^S$) or W1655 F− ($M13^R$ as a control), and dosed each animal with either live or heat-inactivated M13 carrying pBluescript II (**Fig. 1c**). After dosing the mice with 1×10^{14} M13(pBluescript II), we immediately transferred them to water containing carbenicillin and tracked both total *E. coli* and $Carb^R$ *E. coli* in the feces. *E. coli* colonization fell rapidly and stayed near or below the limit of detection in control mice that were either colonized with F− and given live phage or colonized with F+ but given heat-inactivated phage. In contrast, mice colonized with F+ and dosed with live phage had a transient drop in colonization on the first day, during which the rise of $Carb^R$ cells occurred, and colonization was re-established within one day by an *E. coli* population resistant to carbenicillin (**Fig. 1c**). These results suggest that orally dosed M13 phage were able to infect *E. coli* in the gut and deliver a plasmid conferring resistance to carbenicillin.

We replicated M13-mediated pBluescript II delivery to *E. coli* in the gut in an independent animal experiment. Sm-treated mice were colonized with Sm^R *E. coli* W1655 F+ and orally dosed with ten-fold serial dilutions of M13(pBluescript II). Colonization by $Carb^R$ *E. coli* was consistent at high doses but variable at lower doses representing a significantly higher probability of successful colonization with increasing phage dose (**Fig. 1d;** P=0.009, odds ratio=2.5, logistic regression). Plasmid DNA of the expected size was detected in fecal $Carb^R$ *E. coli* isolates from all 11 mice that were successfully colonized (**Supplementary Fig. 3a**).

92 Genome sequencing confirmed the presence of pBluescript II in these 11 isolates, which was
93 undetectable in the parent strain (**Supplementary Fig. 3b**). These results indicate that plasmid
94 DNA was transferred from M13 phage into recipient *E. coli* colonizing the GI tract.

95 Finally, we repeated this experiment in the absence of carbenicillin selection. We
96 colonized mice with Sm^R *E. coli*, gavaged each mouse with M13(pBluescript II), and tracked
97 both infected ($Carb^R$) and total (Sm^R) *E. coli* in feces. The fraction of phage-infected $Carb^R$
98 *E. coli* was low, reaching a maximum of 0.1% of the total population (**Supplementary Fig. 4**),
99 potentially indicative of poor phage survival during GI transit. We gavaged mice with
100 M13(pBluescript II) and assayed for viable phage in the feces. The median output of viable
101 M13(pBluescript II) was reduced to 1×10^6 relative to an input of 6×10^{13} (**Supplementary Figs.**
102 **5a,b**). M13(pBluescript II) is acid tolerant *in vitro* (**Supplementary Fig. 5c**), suggesting that
103 additional factors may be responsible for the low *in vivo* viability and emphasizing the benefits of
104 pairing gene delivery with antibiotic selection.

105

106 **M13 carrying CRISPR-Cas9 can target *E. coli in vitro.*** We generated two fluorescently
107 marked isogenic derivatives of Sm^R W1655 F+ using the *mcherry* (red fluorescence) or the *sfgfp*
108 (green fluorescence) marker gene. Next, we constructed M13-compatible non-targeting (NT)
109 and GFP-targeting (GFPT) CRISPR-Cas9 vectors by cloning the spacers sequences, *bla* gene,
110 and f1 origin of replication into the previously described low-copy vector pCas9[22], generating
111 pCas9-NT-f1A/B and pCas9-GFPT-f1A/B (**Supplementary Fig. 6**). The *bla* and f1 *ori* were
112 cloned as a fragment from pBluescript II in both possible orientations (A or B) to make possible
113 M13 ssDNA packaging of either strand of vector DNA. We packaged these phagemids into M13
114 using a helper strain and called the resulting phage NT-M13 or GFPT-M13. The two phage were
115 used to infect the GFP+ or mCherry+ strains and cells were diluted and spotted on solid media
116 containing carbenicillin to select for the transferred phagemid. GFP+ *E. coli* infected with
117 GFPT-M13 exhibited impaired colony growth relative to the NT-M13 control (**Fig. 2a** and
118 **Supplementary Fig. 7**). Total CFUs were not markedly affected, indicating that cells can
119 recover from M13-delivered CRISPR-Cas9 targeting.

120 Analysis of the surviving cells provided mechanistic insights. Colonies arising from
121 infection with NT-M13 or GFPT-M13 were streak purified, allowing us to pick a mixture of bright
122 and dim colonies. Of 16 GFPT clones analyzed, 11 were non-fluorescent (**Fig. 2b**). PCR
123 amplification of *sfgfp* confirmed the intact gene in 4 NT controls and all 5 GFPT clones that

124 retained fluorescence (**Fig. 2c**). Sanger sequencing revealed that 1 GFPT clone had a point
125 mutation in the *sfgfp* target (**Supplementary Fig. 8**) while the 4 others had lost the spacer in the
126 CRISPR-Cas9 phagemid that leads to targeting (**Supplementary Fig. 9**). All 11 non-fluorescent
127 GFPT clones retained the spacer (**Supplementary Fig. 9**) and had chromosomal deletions at
128 the target locus: 10 were PCR negative (**Fig. 2c**) and 1 had a small deletion within *sfgfp*
129 (**Supplementary Fig. 8**). Finally, we used whole genome sequencing to define the size of each
130 deletion, which ranged from 45 bp to 82.6 kb (**Fig. 2d**), consistent with prior work demonstrating
131 that *E. coli* can repair Cas9-induced double-stranded breaks through homologous
132 recombination[23].

133 These results led us to hypothesize that targeted cells would be less able to recover
134 during competitive growth. We co-cultured GFP+ and mCherry+ *E. coli*, adding either NT-M13
135 or GFPT-M13 followed by carbenicillin to select for phage infection. GFPT-M13 decreased the
136 frequency of GFP+ colonies by 4 h, relative to the NT-M13 control (**Fig. 3a**). At later timepoints
137 (16-24 h), healthy GFP+ colonies increased in abundance, consistent with low levels of
138 carbenicillin after 4 h in cultures expressing the β-lactamase resistance gene (**Fig. 3b**). We
139 confirmed the loss of selection for the phagemid by re-analyzing our colonies on selective media
140 (**Supplementary Fig. 10**). Next, we used flow cytometry to better quantify the two strains in an
141 independent experiment. Compared to the NT-M13 control, GFPT-M13 co-cultures exhibited
142 fewer GFP+ events (**Fig. 3c**) and a bimodal distribution of fluorescence (**Fig. 3d**). Counts of
143 GFP+ cells were higher by flow cytometry than on solid media for the same co-cultures (**Fig. 3c**
144 inset), consistent with an impaired growth of these cells. GFP+ events further decreased at 24 h
145 in the GFPT-M13 group (**Supplementary Fig. 11**). Taken together, these results suggest that
146 competitive growth can increase the efficiency of targeting a strain for depletion due to the
147 resulting growth impairments in the targeted strain.

148

149 **Sequence-specific depletion of *E. coli* within the mouse gut microbiota.** We co-colonized
150 Sm-treated mice with both Sm^R F+ *sfgfp* and Sm^R F+ *mcherry* strains, orally dosed them with
151 either 10^{11} NT-M13 or GFPT-M13, and added carbenicillin in the water to select for phage
152 infection. After one week of treatment, carbenicillin was removed from the water and mice were
153 followed for an additional week to determine whether phage-induced changes would persist in
154 the absence of maintaining selection (**Fig. 4a**). Flow cytometry on mouse stool samples
155 revealed that the GFP+ strain outcompeted the mCherry+ strain in the NT-M13 group (**Figs.**

156 **4b,c** and **Supplementary Figs. 12,13**). In contrast, GFP+ events in the GFPT-M13 group
157 exhibited a sharp decrease on Day 2, followed by a recovery on Days 7 and 14 to levels below
158 the NT-M13 group (**Figs. 4b,c**). Culturing from mouse stool confirmed the decreased GFP+
159 events on Day 2 (**Supplementary Fig. 14**). In 4 mice that received GFPT-M13, the mCherry+
160 strain fixed in the population (GFP+ events were below background), an outcome that was not
161 observed for any mouse in the NT-M13 group (**Fig. 4d**). Despite lifting the carbenicillin selection
162 for 1 week, endpoint GFP+ events remained significantly lower in the GFPT-M13 group relative
163 to NT-M13 controls (**Fig. 4e**; P=0.0002, Mann-Whitney test). These data support the utility of
164 M13-delivered CRISPR-Cas9 for sequence-specific depletion of an otherwise isogenic bacterial
165 strain in the mouse gut.
166
167 **M13-delivered CRISPR-Cas9 induces chromosomal deletions in the gut microbiome.** We
168 constructed a double-marked Sm^R F+ *sfgfp mcherry* strain to quantify the efficiency of gene
169 deletion. We introduced this strain into Sm-treated mice, orally dosed each mouse with either
170 10^{11} NT-M13 or GFPT-M13, and added carbenicillin in the water. After one week, we removed
171 carbenicillin and followed the mice for a final week (**Fig. 5a**). GFP- mCherry+ events were
172 detectable in GFPT-M13 but not NT-M13 mice, indicative of successful CRISPR-Cas9 delivery
173 and gene deletion (**Fig. 5b** and **Supplementary Fig. 15**). By the final timepoint, GFP- mCherry+
174 events were detected in 3/8 mice (**Fig. 5c** and **Supplementary Fig. 16a**). The relative
175 abundance of GFP- mCherry+ cells varied from 12-96% (**Fig. 5c**). Culturing on solid media
176 confirmed the presence of viable red fluorescent colonies in proportions consistent with flow
177 cytometry results (**Fig. 5d** and **Supplementary Fig. 16b**).
178 To more definitively assess the presence or absence of the targeted genomic locus and
179 the CRISPR-Cas system, we isolated GFP- mCherry+ and GFP+ mCherry+ *E. coli* from Day 2
180 mouse stool. All of the GFP+ mCherry+ isolates from the NT-M13 group and the GFP-
181 mCherry+ isolates from the GFPT-M13 group had an intact spacer sequence (**Supplementary**
182 **Fig. 17a**). In contrast, 4/5 GFP+ mCherry+ isolates from the GFPT-M13 group had lost the
183 spacer (**Supplementary Fig. 17a**). Of note, the remaining isolate lost *cas9*, and parts of the
184 CRISPR array and tracrRNA (**Supplementary Figs. 17b,c**). Whole genome sequencing was
185 used to confirm putative chromosomal deletions and to quantify their size. Two representative
186 colonies were analyzed from each of the 3 mice with detectable GFP- mCherry+ cells, revealing
187 a wide range in deletion sizes that were not observed in a control GFP+ mCherry+ isolate from

each animal (**Fig. 5e**). These results indicate that while it is possible for CRISPR-Cas9-induced genomic deletion events to occur *in vivo*, resultant deletion strains may or may not outcompete the parent strain due to the potential to evade targeting through loss of some or all of the exogenous CRISPR-Cas system.

Discussion

Our results emphasize that foundational, reductionist, and highly controlled studies will be necessary to assess the feasibility, utility, and limitations of phage-based gene delivery as a tool for microbiome editing. While our results provide a *proof-of-principle* for strain-specific targeting within the GI tract, the full eradication of the targeted strain was difficult to achieve due to the ability of bacterial cells to survive Cas9-induced double-stranded breaks by homologous recombination[24]. We propose that CRISPR-Cas9 may be better suited to induce targeted genomic deletions, leveraging the conserved DNA repair pathways present in bacteria. An advantage of this approach is that the deletion of a single genomic locus is unlikely to have as large an impact on the rest of the gut microbiota than if the strain were to be removed entirely. Remarkably, we detected a wide range of deletion sizes (379–68,321 bp), highlighting the ability of bacteria to survive large deletions and opening up the potential for the *in vivo* removal of entire biosynthetic gene clusters or pathogenicity islands. In turn, our data suggests that it may also be feasible to deliver more complex genetic circuits to *E. coli*, with the goal of boosting metabolic pathways beneficial to its mammalian host. Furthermore, the size of resulting chromosomal deletions could be controlled by providing a DNA repair template alongside CRISPR-Cas9[22].

There are several limitations of our current approach that could be refined through iterative “design-build-test” cycles. Perhaps most importantly, we identified multiple mechanisms through which cells can escape CRISPR-Cas9 targeting during *in vitro* growth and/or within the GI tract, including the loss of the spacer in the CRISPR array, target site mutations, and even one case in which the entire CRISPR-Cas9 system was deleted. Spacer loss could be reduced through the use of single guide RNAs (sgRNAs)[25] or through the mutation of the repeat sequences that flank each spacer[26]. Target site mutation could be countered by targeting multiple sites simultaneously and by prioritizing conserved regions of the target genes essential for activity.

Another important caveat is that antibiotics were used to select for successful gene delivery, which is not ideal due to potential disruptions to the gut microbiota[1] and/or selection for antibiotic resistance. Removing this selection resulted in a low penetrance of cells carrying the delivered cargo, similar to that observed for delivery to bacteria in a soil community in the absence of selection[27]. Low rates of gene delivery are likely driven by the observed loss of viable M13 bacteriophage during transit through the GI tract. We confirmed that M13 is acid tolerant; however, it is sensitive to artificial gastric juice[28], suggesting that gastric proteases may be responsible for low phage survival. The aggregative behaviour of filamentous phage in response to microbial and/or host polymers[29] may also be important to consider. Encapsulating phage for oral transit may be able to circumvent these limitations or alternatively, it may be possible to use non-antibiotic selection strategies such as dietary supplementation of an exclusive substrate[3] that can only be used by phage-infected strains.

In conclusion, this work provides a valuable step towards establishing a modular toolkit for microbiome editing. The extension of these approaches to enable the genetic manipulation of a more diverse panel of bacteria found within the human microbiota will require a renewed effort to isolate and characterize bacteriophage that target strains of interest[30]. Robust *in vitro* methods to study and genetically modify novel bacteriophage are also needed, given that most of their host bacteria remain genetically intractable[31]. Finally, our work emphasizes the value of model systems for understanding the rules of engagement, complementing ongoing efforts to advance candidate phage-based therapies into the clinic.

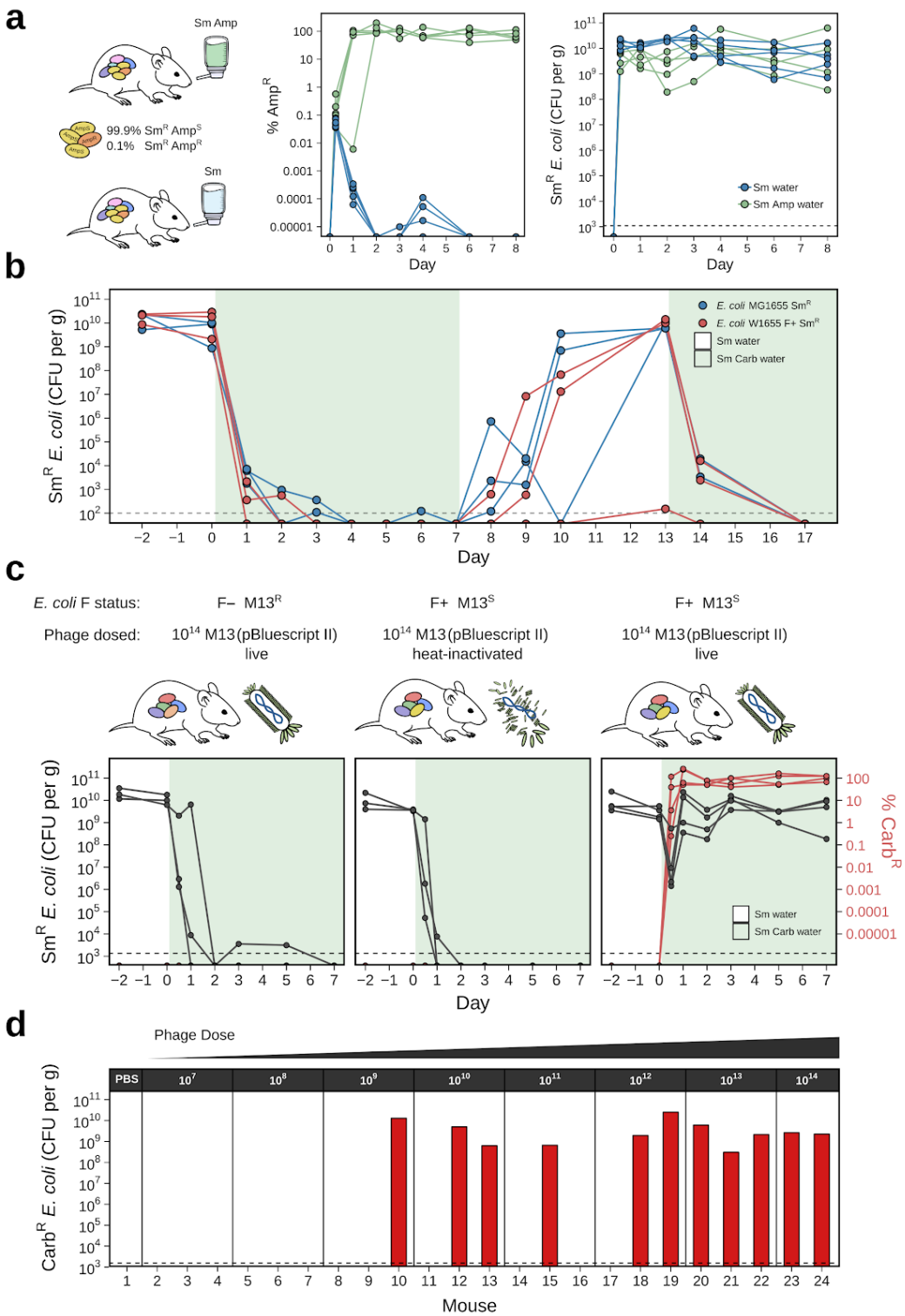
References

1. Basolo, A. *et al.* Effects of underfeeding and oral vancomycin on gut microbiome and nutrient absorption in humans. *Nat. Med.* **26**, 589–598 (2020).
2. Smillie, C. S. *et al.* Strain Tracking Reveals the Determinants of Bacterial Engraftment in the Human Gut Following Fecal Microbiota Transplantation. *Cell Host Microbe* **23**, 229–240.e5 (2018).
3. Shepherd, E. S., DeLoache, W. C., Pruss, K. M., Whitaker, W. R. & Sonnenburg, J. L. An exclusive metabolic niche enables strain engraftment in the gut microbiota. *Nature* **557**, 434–438 (2018).
4. Isabella, V. M. *et al.* Development of a synthetic live bacterial therapeutic for the human metabolic disease phenylketonuria. *Nat. Biotechnol.* **36**, 857–864 (2018).
5. Citorik, R. J., Mimee, M. & Lu, T. K. Sequence-specific antimicrobials using efficiently delivered RNA-guided nucleases. *Nat. Biotechnol.* **32**, 1141–1145 (2014).
6. Bikard, D. *et al.* Exploiting CRISPR-Cas nucleases to produce sequence-specific antimicrobials. *Nat. Biotechnol.* **32**, 1146–1150 (2014).
7. Park, J. Y. *et al.* Genetic engineering of a temperate phage-based delivery system for CRISPR/Cas9 antimicrobials against *Staphylococcus aureus*. *Sci. Rep.* **7**, 44929 (2017).
8. Selle, K. *et al.* In Vivo Targeting of *Clostridioides difficile* Using Phage-Delivered CRISPR-Cas3 Antimicrobials. *mBio* **11**, (2020).
9. Human Microbiome Project Consortium. Structure, function and diversity of the healthy human microbiome. *Nature* **486**, 207–214 (2012).
10. Camarillo-Guerrero, L. F., Almeida, A., Rangel-Pineros, G., Finn, R. D. & Lawley, T. D. Massive expansion of human gut bacteriophage diversity. *Cell* **184**, 1098–1109.e9 (2021).
11. Hofschneider, P. H. Untersuchungen über „kleine“ *E. coli* K 12 Bakteriophagen. *Zeitschrift für Naturforschung B* **18**, 203–210 (1963).
12. Ackermann, H.-W. Phage classification and characterization. *Methods Mol. Biol.* **501**, 127–140 (2009).
13. Salivar, W. O., Tzagoloff, H. & Pratt, D. Some Physical-Chemical and Biological Properties of the Rod-Shaped Coliphage M13. *Virology* **24**, 359–371 (1964).
14. Lee, G. S. & Ames, G. F. Analysis of promoter mutations in the histidine transport operon of *Salmonella typhimurium*: use of hybrid M13 bacteriophages for cloning, transformation, and

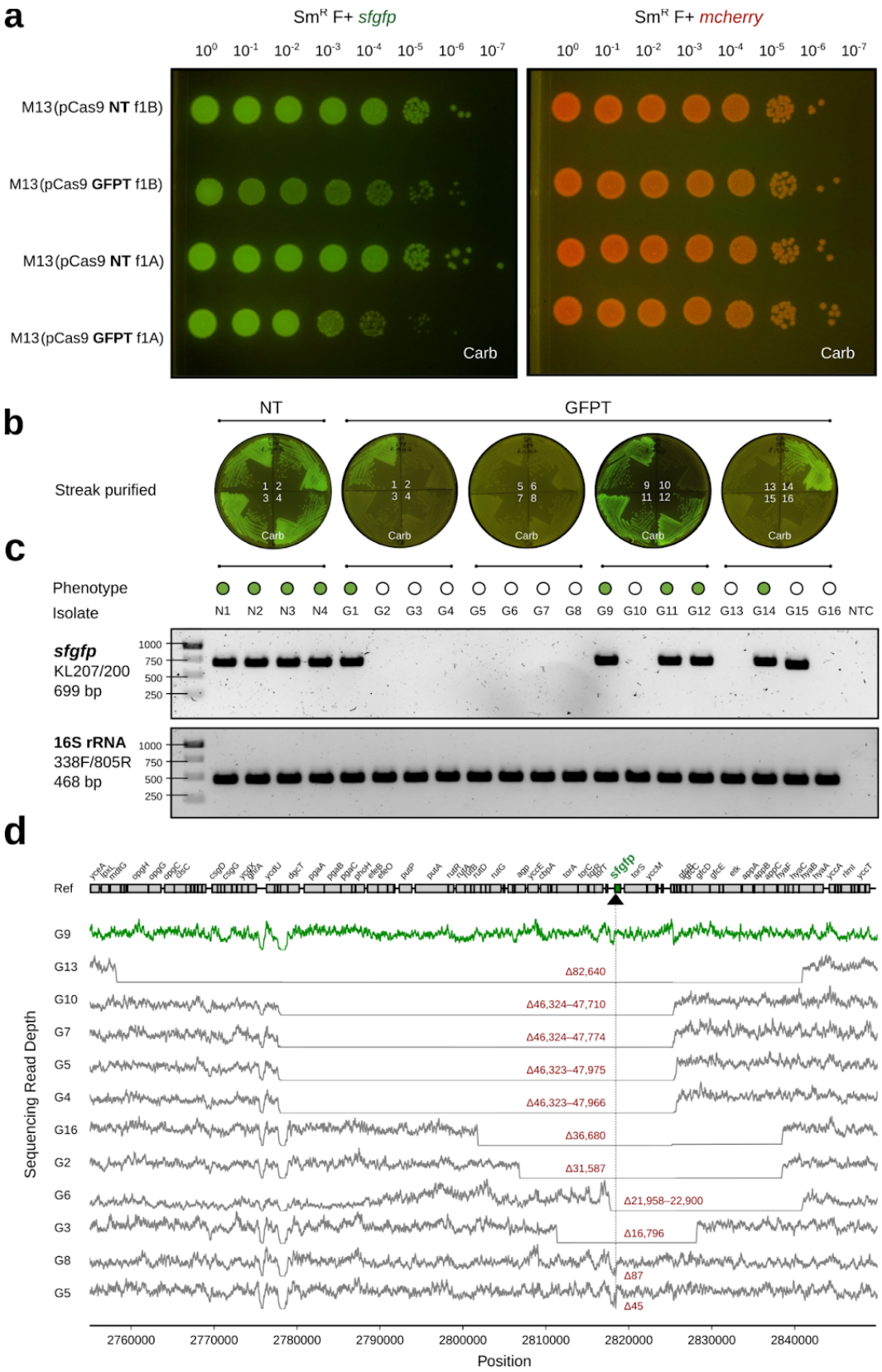
269 sequencing. *J. Bacteriol.* **159**, 1000–1005 (1984).
270 15. Guiney, D. G. Host range of conjugation and replication functions of the *Escherichia coli* sex
271 plasmid F*lac*. Comparison with the broad host-range plasmid RK2. *J. Mol. Biol.* **162**,
272 699–703 (1982).
273 16. Zinder, N. D. & Boeke, J. D. The filamentous phage (Ff) as vectors for recombinant DNA--a
274 review. *Gene* **19**, 1–10 (1982).
275 17. Westwater, C. *et al.* Use of Genetically Engineered Phage To Deliver Antimicrobial Agents
276 to Bacteria: an Alternative Therapy for Treatment of Bacterial Infections. *Antimicrob. Agents*
277 *Chemother.* **47**, 1301–1307 (2003).
278 18. Lu, T. K. & Collins, J. J. Engineered bacteriophage targeting gene networks as adjuvants
279 for antibiotic therapy. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 4629–4634 (2009).
280 19. Cao, J. *et al. Helicobacter pylori*-antigen-binding fragments expressed on the filamentous
281 M13 phage prevent bacterial growth. *Biochim. Biophys. Acta* **1474**, 107–113 (2000).
282 20. Myhal, M. L., Laux, D. C. & Cohen, P. S. Relative colonizing abilities of human fecal and K
283 12 strains of *Escherichia coli* in the large intestines of streptomycin-treated mice. *Eur. J.*
284 *Clin. Microbiol.* **1**, 186–192 (1982).
285 21. Alting-Mees, M. A. & Short, J. M. pBluescript II: gene mapping vectors. *Nucleic Acids Res.*
286 **17**, 9494 (1989).
287 22. Jiang, W., Bikard, D., Cox, D., Zhang, F. & Marraffini, L. A. RNA-guided editing of bacterial
288 genomes using CRISPR-Cas systems. *Nat. Biotechnol.* **31**, 233–239 (2013).
289 23. Cui, L. & Bikard, D. Consequences of Cas9 cleavage in the chromosome of *Escherichia*
290 *coli*. *Nucleic Acids Res.* **44**, 4243–4251 (2016).
291 24. Dillingham, M. S. & Kowalczykowski, S. C. RecBCD enzyme and the repair of
292 double-stranded DNA breaks. *Microbiol. Mol. Biol. Rev.* **72**, 642–71, Table of Contents
293 (2008).
294 25. Guo, C.-J. *et al.* Depletion of microbiome-derived molecules in the host using *Clostridium*
295 genetics. *Science* **366**, (2019).
296 26. Csörgő, B. *et al.* A compact Cascade–Cas3 system for targeted genome engineering. *Nat.*
297 *Methods* **17**, 1183–1190 (2020).
298 27. Rubin, B. E. *et al.* Targeted Genome Editing of Bacteria Within Microbial Communities.
299 doi:10.1101/2020.07.17.209189.
300 28. Tóthová, L. ’ubomíra, Bábíčková, J. & Celec, P. Phage survival: the biodegradability of M13

301 phage display library *in vitro*. *Biotechnol. Appl. Biochem.* **59**, 490–494 (2012).
302 29. Secor, P. R. *et al.* Filamentous Bacteriophage Promote Biofilm Assembly and Function. *Cell*
303 *Host Microbe* **18**, 549–559 (2015).
304 30. Soto-Perez, P. *et al.* CRISPR-Cas System of a Prevalent Human Gut Bacterium Reveals
305 Hyper-targeting against Phages in a Human Virome Catalog. *Cell Host Microbe* **26**,
306 325–335.e5 (2019).
307 31. Bisanz, J. E. *et al.* A Genomic Toolkit for the Mechanistic Dissection of Intractable Human
308 Gut Bacteria. *Cell Host Microbe* **27**, 1001–1013.e9 (2020).

Figures and legends



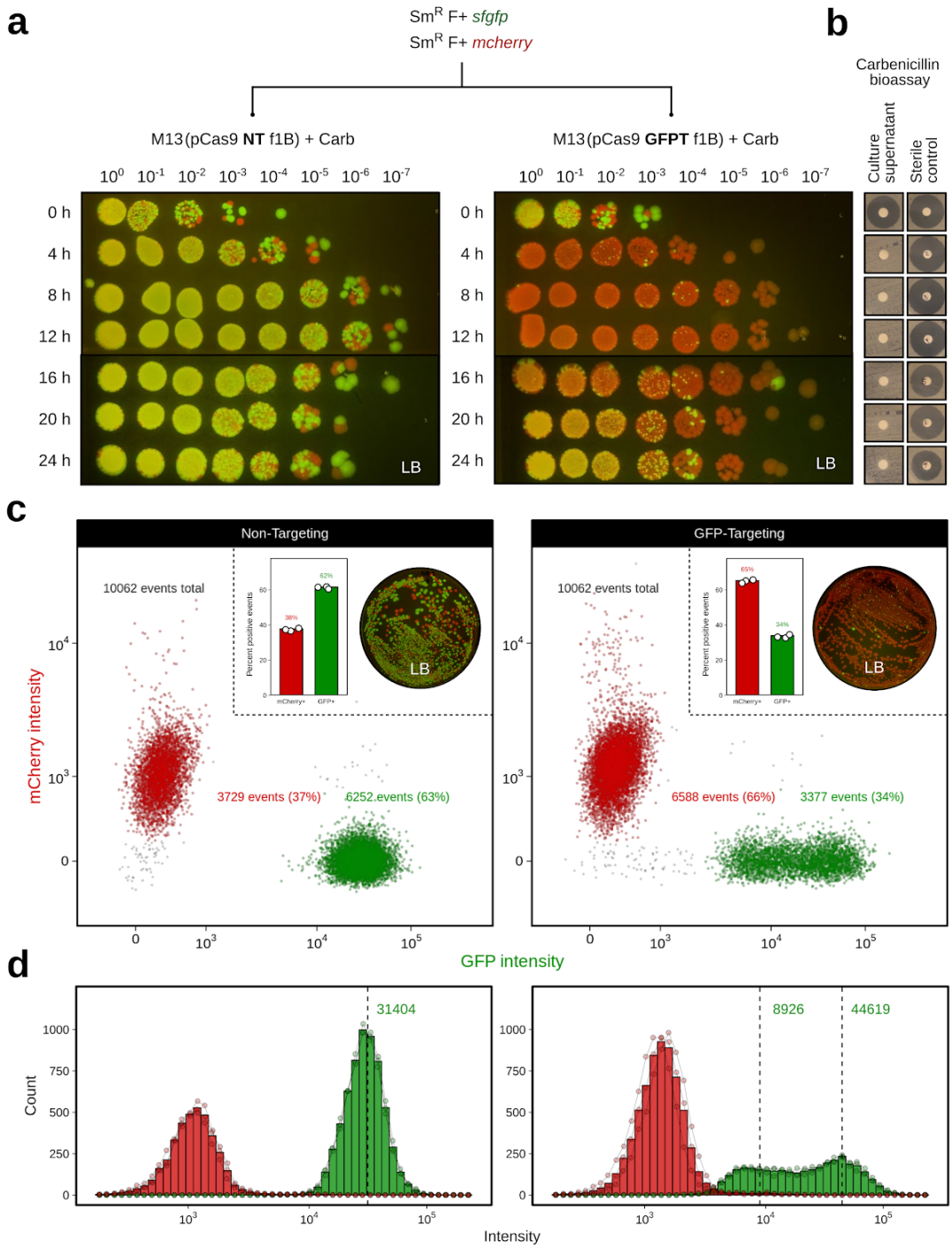
312 **Figure 1. M13 bacteriophage can deliver a plasmid-borne antibiotic resistance gene to *E.***
313 ***coli* in the mouse gut. (a)** A resistant subpopulation of *E. coli* can be selected in the gut when
314 a β-lactam antibiotic is provided in the water. Streptomycin(Sm)-treated mice were gavaged with
315 Sm^R MG1655 containing 99.9% ampicillin-sensitive (Amp^S) and 0.1% ampicillin-resistant (Amp^R)
316 cells, and provided water containing Sm (n=5) or Sm+Amp (n=6). **(b)** A sensitive *E. coli*
317 population is unable to maintain colonization in the gut when carbenicillin (Carb) is provided in
318 the water. Mice were colonized with either Sm^R MG1655 or Sm^R W1655 F+ (n = 3 each) and
319 Carb was provided in the water (shaded timepoints). **(c)** M13(pBluescript II) can infect F+ *E. coli*
320 in the gut. Mice were split into three groups based on colonization and phage treatment: (1) Sm^R
321 W1655 F− and live phage (n=3); (2) Sm^R W1655 F+ and heat-inactivated phage (n=3); (3) Sm^R
322 W1655 F+ and live phage (n=4). 10^{14} phage were dosed and Carb was provided in the water.
323 Total *E. coli* (black) and %Carb resistant colonies ($Carb^R$; red) in mouse fecal pellets are shown.
324 **(d)** M13-based delivery of an antibiotic resistance gene is dose-dependent. Varying doses
325 (10^7-10^{14}) of M13(pBluescript II) were given to mice (n=2-3/dose) and Carb was provided in the
326 water. After 2 days, $Carb^R$ CFU/gram feces was determined. Dashed line indicates our limit of
327 detection.


a
SmR F+ sfgfp
SmR F+ mcherry
10^0 10^-1 10^-2 10^-3 10^-4 10^-5 10^-6 10^-7
M13(pCas9 NT f1B)
M13(pCas9 GFPT f1B)
M13(pCas9 NT f1A)
M13(pCas9 GFPT f1A)
Carb
b
NT
GFPT
Streak purified
c
Phenotype
Isolate
N1 N2 N3 N4 G1 G2 G3 G4 G5 G6 G7 G8 G9 G10 G11 G12 G13 G14 G15 G16 NTC
sfgfp
KL207/200
699 bp
16S rRNA
338F/805R
468 bp
1000
750
500
250
d
Ref
sfgfp
G9
G13
G10
G7
G5
G4
G16
G2
G6
G3
G8
G5
Δ82,640
Δ46,324–47,710
Δ46,324–47,774
Δ46,323–47,975
Δ46,323–47,966
Δ36,680
Δ31,587
Δ21,958–22,900
Δ16,796
Δ87
Δ45
Sequencing Read Depth
2760000 2770000 2780000 2790000 2800000 2810000 2820000 2830000 2840000
Position

328

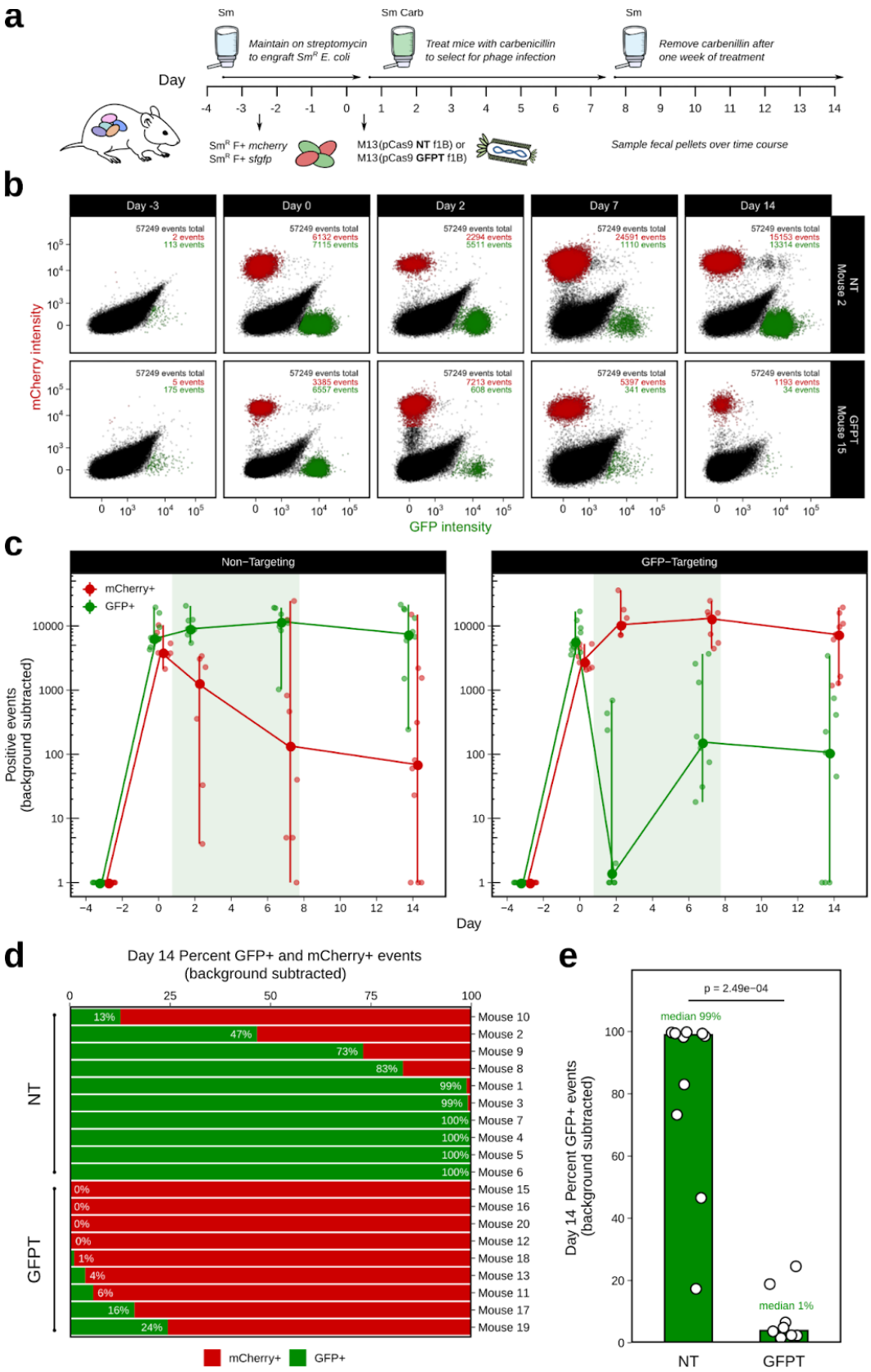
329 **Figure 2. M13-mediated delivery of CRISPR-Cas9 to *E. coli in vitro* causes impaired**
330 **colony growth and can induce chromosomal deletions that encompass the targeted**
331 **gene. (a)** GFP+ *E. coli* exhibit a sick colony morphology after infection with M13 phage carrying
332 GFP-targeting CRISPR-Cas9. NT (non-targeting) or GFPT (GFP-targeting) M13 were used to
333 infect Sm^R W1655 F+ *sfgfp* or Sm^R W1655 F+ *mcherry* as a control. Cells were infected, diluted,
334 and spotted onto media with selection for the vector; f1A or f1B indicates version of vector.
335 **(b)** CRISPR-Cas9 targeting the *sfgfp* gene can induce loss of fluorescence. Colonies arising
336 from infection with NT-M13 or GFPT-M13 were subjected to several rounds of streak purification
337 on selective media to ensure phenotypic homogeneity and clonality. The majority (11/16) of
338 GFPT clones exhibited a loss of fluorescence. **(c)** Clones exhibiting loss of fluorescence either
339 lack an *sfgfp* PCR amplicon or exhibit an amplicon of decreased size. Genomic DNA was
340 isolated from streak-purified clones and PCR was used to determine whether the *sfgfp* gene
341 was present; PCR for the 16S rRNA gene was performed as a positive control. **(d)** Genome
342 sequencing results confirm that nonfluorescent clones have chromosomal deletions
343 encompassing the targeted gene. Read depth surrounding *sfgfp* locus for G9 clone (green line,
344 fluorescent control) and all nonfluorescent clones (grey lines). Deletion size indicated in red;
345 range indicates a deletion flanked by repetitive sequences. Black arrow and vertical line denote
346 position of targeting. Carb, carbenicillin.

347


a
SmR F+ sfgfp
SmR F+ mcherry
M13(pCas9 NT f1B) + Carb
M13(pCas9 GFPT f1B) + Carb
10^0 10^-1 10^-2 10^-3 10^-4 10^-5 10^-6 10^-7
0 h
4 h
8 h
12 h
16 h
20 h
24 h
LB
b
Carbenicillin bioassay
Culture supernatant
Sterile control
c
Non-Targeting
GFP-Targeting
10062 events total
3729 events (37%)
6252 events (63%)
6588 events (66%)
3377 events (34%)
mCherry intensity
GFP intensity
d
Count
Intensity
31404
8926
44619

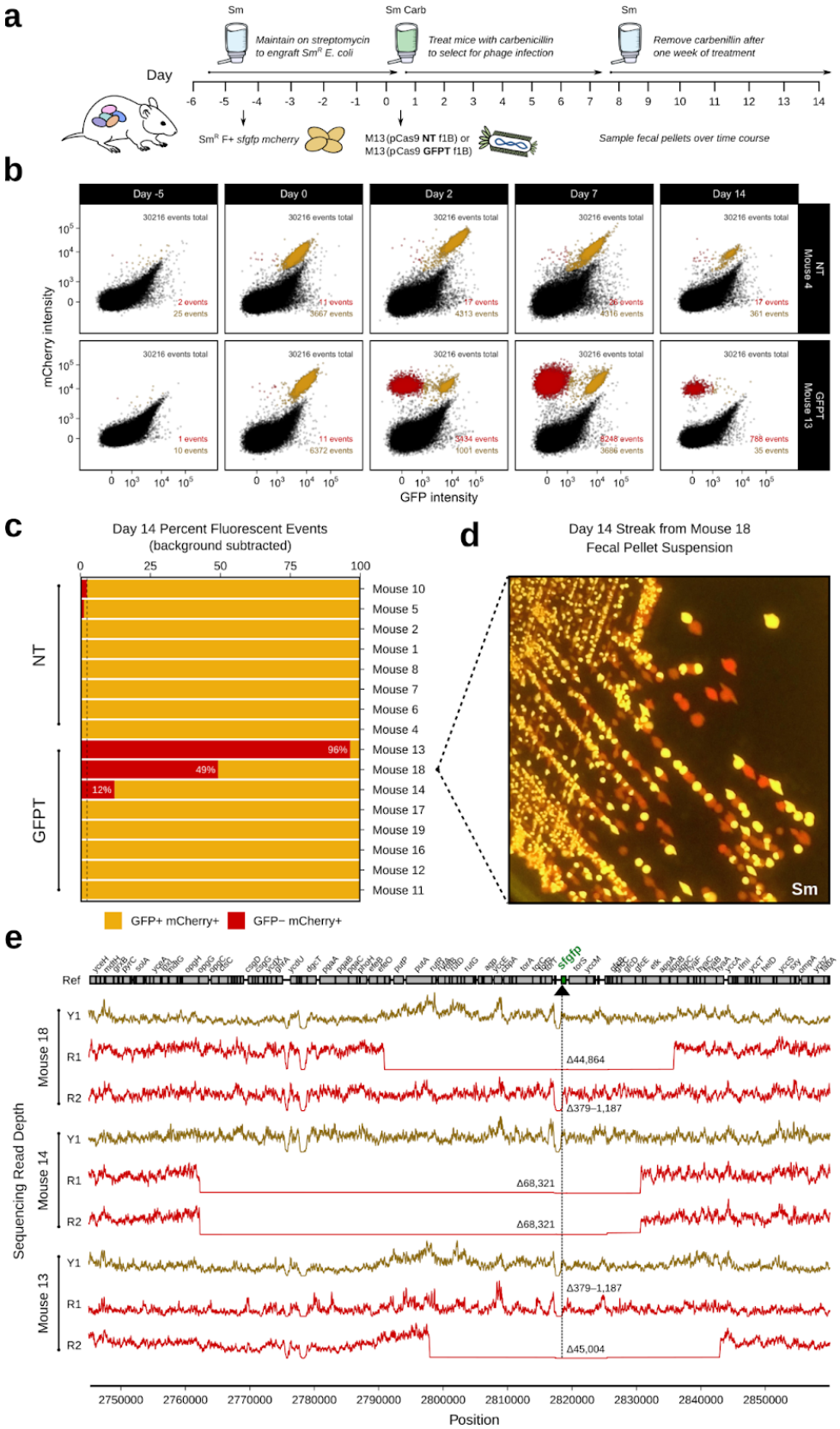
348

349 **Figure 3. M13-delivered CRISPR-Cas9 for sequence-specific targeting of *E. coli* in *in vitro***
350 **co-cultures of fluorescently marked isogenic strains. (a)** M13-delivered GFP-targeting
351 CRISPR-Cas9 leads to reduced competitive fitness of the GFP-marked strain. A co-culture of
352 Sm^R F+ *sfgfp* and Sm^R F+ *mcherry* was incubated with NT-M13 or GFPT-M13 at a starting MOI
353 of ~500. Carbenicillin (Carb) was added to a final concentration of 100 µg/ml to select for phage
354 infection. Co-cultures were sampled every 4 h over 24 h; cells were washed, serially diluted,
355 and spotted onto non-selective media to assess targeting of the GFP-marked strain.
356 **(b)** Carbenicillin in culture supernatants was not detectable within 4 h of growth, using a
357 carbenicillin bioassay against indicator strain *Bacillus subtilis* 168; bioassay detection limit
358 approximately 2.5 µg/ml. **(c)** Flow cytometry of co-cultures 8 h following the addition of phage
359 and carbenicillin show reduced GFP+ events in the GFPT versus NT condition. Representative
360 flow plot shows data from one of three replicates. Inset: bar graph quantifying percent GFP+
361 and mCherry+ events for three replicates (left); plating results for a single replicate on
362 non-selective media (right). **(d)** GFPT CRISPR-Cas9 changes the shape of the distribution of
363 GFP+ population. Histogram of mCherry+ and GFP+ events by intensity shows that a
364 proportion of GFP+ cells in the GFPT condition have shifted to a state of lower fluorescence.
365 Bars indicate the mean of three replicates; connected points are individual replicates.
366


a
Sm
Sm Carb
Sm
Maintain on streptomycin to engraft SmR E. coli
Treat mice with carbenicillin to select for phage infection
Remove carbenillin after one week of treatment
Day
-4 -3 -2 -1 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14
SmR F+ mcherry
SmR F+ sfgfp
M13(pCas9 NT f1B) or
M13(pCas9 GFPT f1B)
Sample fecal pellets over time course
b
Day -3
Day 0
Day 2
Day 7
Day 14
57249 events total
2 events
113 events
6132 events
7115 events
2294 events
5511 events
24591 events
1110 events
15153 events
13314 events
5 events
175 events
3385 events
6557 events
7213 events
608 events
5397 events
341 events
1193 events
34 events
NT Mouse 2
GFPT Mouse 15
mCherry intensity
GFP intensity
c
Non−Targeting
GFP−Targeting
mCherry+
GFP+
Positive events (background subtracted)
Day
d
Day 14 Percent GFP+ and mCherry+ events (background subtracted)
NT
GFPT
Mouse 10
Mouse 2
Mouse 9
Mouse 8
Mouse 1
Mouse 3
Mouse 7
Mouse 4
Mouse 5
Mouse 6
Mouse 15
Mouse 16
Mouse 20
Mouse 12
Mouse 18
Mouse 13
Mouse 11
Mouse 17
Mouse 19
13%
47%
73%
83%
99%
99%
100%
100%
100%
100%
0%
0%
0%
0%
1%
4%
6%
16%
24%
mCherry+
GFP+
e
p = 2.49e−04
median 99%
median 1%
Day 14 Percent GFP+ events (background subtracted)
NT
GFPT

367

368 **Figure 4. M13-delivered CRISPR-Cas9 for sequence-specific depletion of *E. coli* in the gut**
369 **of mice colonized by competing fluorescently marked isogenic strains. (a)** Timeline: Day
370 -3, colonize with 50/50 mixture of streptomycin(Sm)-resistant Sm^R F+ *sfgfp* and Sm^R F+ *mcherry*;
371 Day 0, dose with 10^{11} NT-M13 or GFPT-M13 (n=10/group) and provide carbenicillin (Carb) in the
372 water; Day 7, remove carbenicillin. **(b)** GFPT-M13 can lead to loss of the GFP-marked strain.
373 Time series flow plots of fecal samples for one mouse from each of NT and GFPT groups. Top
374 right: number of total, mCherry+, and GFP+ events. **(c)** Mice in GFPT group exhibited a
375 decrease in number of fecal GFP+ events in over time compared to NT group; timepoints were
376 excluded if both GFP+ and mCherry+ events were below background thresholds. Line graph:
377 points indicate median; vertical lines, range. **(d)** Mice in GFPT group exhibited depletion or loss
378 of the GFP-marked strain. Percent GFP+ and mCherry+ events for each mouse on Day 14.
379 Mice were excluded if both GFP+ and mCherry+ events were both below background
380 thresholds (final n=9 GFPT and n=10 NT). **(e)** A significant difference was observed in the
381 percent of GFP+ events in fecal samples at Day 14 in the GFPT group compared to NT. Bars
382 are medians; *p*-value, Mann-Whitney test.


a
Sm
Sm Carb
Sm
Maintain on streptomycin to engraft Sm^R E. coli
Treat mice with carbenicillin to select for phage infection
Remove carbenillin after one week of treatment
Day
-6 -5 -4 -3 -2 -1 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14
Sm^R F+ sfgfp mcherry
M13 (pCas9 NT f1B) or
M13 (pCas9 GFPT f1B)
Sample fecal pellets over time course
b
Day -5
Day 0
Day 2
Day 7
Day 14
30216 events total
NT Mouse 4
GFPT Mouse 13
mCherry intensity
GFP intensity
c
Day 14 Percent Fluorescent Events (background subtracted)
NT
GFPT
Mouse 10
Mouse 5
Mouse 2
Mouse 1
Mouse 8
Mouse 7
Mouse 6
Mouse 4
Mouse 13
Mouse 18
Mouse 14
Mouse 17
Mouse 19
Mouse 16
Mouse 12
Mouse 11
96%
49%
12%
GFP+ mCherry+
GFP− mCherry+
d
Day 14 Streak from Mouse 18 Fecal Pellet Suspension
Sm
e
Ref
sfgfp
Mouse 18
Mouse 14
Mouse 13
Y1
R1
R2
Δ44,864
Δ379–1,187
Δ68,321
Δ68,321
Δ379–1,187
Δ45,004
Sequencing Read Depth
2750000 2760000 2770000 2780000 2790000 2800000 2810000 2820000 2830000 2840000 2850000
Position

383

384 **Figure 5. M13-delivered CRISPR-Cas9 can induce chromosomal deletions encompassing**
385 **the targeted gene in *E. coli* colonizing the mouse gut. (a)** Timeline: Day -5, colonize with
386 double-marked streptomycin(Sm)-resistant Sm^R F+ *sfgfp mcherry*; Day 0, dose with 10^{11} NT-M13
387 or GFPT-M13 (n=10/group) and provide carbenicillin (Carb) in the water; Day 7, remove
388 carbenicillin. **(b)** GFPT-M13 can cause loss of GFP fluorescence in double-marked *E. coli*. Time
389 series flow plots of fecal samples for select mice, one from each of NT and GFPT groups. Top
390 right: total number of events; bottom right: GFP- mCherry+ events and GFP+ mCherry+ events.
391 **(c)** Day 14 fecal samples of three mice in GFPT group exhibited mCherry-only fluorescence.
392 Percent GFP+ mCherry+ and GFP- mCherry+ events for each mouse; mice were excluded if
393 both populations were below the background threshold (final n=8/group). Dashed line indicates
394 maximum mCherry fluorescence for NT group. **(d)** Colonies arising from culture of Mouse 18
395 Day 14 fecal sample confirmed presence of red-only fluorescence. **(e)** Genome sequencing
396 results confirm red fluorescent isolates from Mouse 13, 14, and 18 have chromosomal deletions
397 encompassing the targeted gene. Read depth surrounding *sfgfp* locus for GFP+ mCherry+ (Y1;
398 yellow lines) and GFP- mCherry+ fluorescent (R1 or R2; red lines) isolates from Day 2 fecal
399 samples. Deletion size indicated in red; range indicates a deletion flanked by repetitive
400 sequences. Black arrow and vertical line denote position of targeting.

401 # Methods

402 **Strains, plasmids, phage, and oligonucleotides.** Bacterial strains, plasmids, and phage used
403 in this study, including descriptions and sources, are provided in **Supplementary Table 1**.
404 Oligonucleotides used in this study are provided in **Supplementary Table 2**.

405 **Minimum inhibitory concentration (MIC) assay.** Cells were prepared by standardizing an
406 overnight culture to an OD_{600} of 0.1 using saline (0.85% NaCl), and further diluted ten-fold in
407 saline then ten-fold in LB. The drug was prepared by dissolving the antibiotic in vehicle (sterile
408 distilled water) and filter-sterilizing, then serially diluting two-fold in vehicle to prepare 100× stock
409 solutions, and finally diluting ten-fold in LB for 10× stock. To wells of a 96-well plate, 60 µl of LB,
410 15 µl of drug, and 75 µl of cells were added and mixed well. Final drug concentrations ranged
411 between 0.002 µg/ml to 1000 µg/ml for ampicillin and 0.24 µg/ml to 2000 µg/ml for carbenicillin.
412 The plate was incubated overnight at 37°C without shaking and OD_{600} was measured the
413 following morning after agitation.

414 **16S rRNA gene sequencing.** Mouse fecal pellets were stored at −80°C. DNA was extracted
415 from single pellets using a ZymoBIOMICS 96 MagBead DNA Kit and 16S rRNA gene
416 sequencing was performed using a dual indexing strategy[32]. Briefly, a 22-cycle primary PCR
417 was performed using KAPA HiFi Hot Start DNA polymerase (KAPA KK2502) and V4 515F/806R
418 Nextera primers. The reaction was diluted in UltraPure DNase/RNase-free water (Life Tech
419 0977-023) and used as template for a 10-cycle secondary (indexing) PCR using sample-specific
420 dual indexing primers. The reactions were normalized using a SequelPrep Normalization plate
421 (Life Tech A10510-01) and the DNA was eluted and pooled. To purify and concentrate the DNA,
422 5 volumes of PB Buffer (Qiagen 28004) were added, mixed, and purified using a QIAquick PCR
423 Purification Kit (Qiagen 28106). The DNA was gel extracted using a MinElute Gel Extraction Kit
424 (Qiagen 28604), quantified by qPCR using a KAPA Library Quantification Kit for Illumina
425 Platforms (KAPA KK4824), and paired-end sequenced on the Illumina MiSeq platform. Data
426 were processed using a 16S rRNA gene analysis pipeline
427 (https://github.com/jbisanz/AmpliconSeq) based on QIIME2[33] incorporating DADA2[34], and
428 analyzed using R packages qiime2R (v0.99.23; https://github.com/jbisanz/qiime2R), phyloseq
429 (v1.33.0)[35], and phylosmith (v1.0.4)[36]. See Data Availability for more information.

430 **Construction of streptomycin-resistant *E. coli* strains.** Strains resistant to the antibiotic
431 streptomycin were generated by either selection for spontaneous resistance or by lambda Red
432 recombineering[37,38]. Spontaneous resistant mutants were selected by plating overnight cultures
433 on LB supplemented with 500 µg/ml streptomycin. Lambda Red recombineering was later used
434 to introduce a specific allele for genetic consistency between strains as different mutations in
435 the *rpsL* gene can confer resistance to streptomycin[39]. Briefly, cells were transformed with the
436 $Carb^R$ temperature-sensitive plasmid pSIJ8[38], and electrocompetent cells were prepared from
437 cells grown in LB carbenicillin at 30°C to early exponential phase and lambda Red recombinase
438 genes were induced by addition of L-arabinose to 7.5 mM. Cells were electroporated with an
439 *rpsL*-Sm^R PCR product amplified from a spontaneous streptomycin-resistant mutant of MG1655
440 using primers PS-rpsL1 and PS-rpsL2, and recombinants were selected on LB supplemented
441 with 500 µg/ml streptomycin. The pSIJ8 plasmid was cured by culturing in liquid at 37°C in the
442 absence of carbenicillin, plating for single colonies, and confirming $Carb^S$. The *rpsL* gene of Sm^R
443 strains was confirmed by Sanger sequencing.

444 **Construction of fluorescently marked *E. coli* strains.** P1 lysates were generated of
445 AV01::pAV01 and AV01::pAV02 carrying clonetegrated *sfgfp* and *mcherry*, respectively[40]. Briefly,
446 150 µl of overnight culture in LB supplemented with 12.5 µg/ml kanamycin was mixed with 1 µl
447 to 25 µl P1 phage (initially propagated from ATCC on MG1655). The mixture was incubated for
448 10 min at 30°C to aid adsorption, added to 4 ml LB 0.7% agar, and overlaid on pre-warmed LB
449 agar supplemented with 25 µg/ml kanamycin 10 mM $MgSO_4$. Plates were incubated overnight at
450 30°C, and phage were harvested by adding 5 ml SM buffer, incubating at room temperature for
451 10 min, and breaking and scraping off the top agar into a conical tube. Phage suspensions were
452 centrifuged to pellet agar; the supernatant was passed through a 100 µm cell strainer, then
453 through a 0.45 µm syringe filter, and lysates were stored at 4°C. For transduction, 1-2 ml of
454 recipient overnight culture was pelleted and resuspended in 1/3 volume LB 10 mM $MgSO_4$ 5 mM
455 $CaCl_2$. 100 µl of cells was mixed with 1 µl to 10 µl P1 lysate and incubated at 30°C for 60
456 minutes. To minimize secondary infections, 200 µl 1 M sodium citrate was added, followed by 1
457 ml of LB. The mixture was incubated at 30°C for 2 h, then plated on LB 10 mM sodium citrate 25
458 µg/ml kanamycin to select for transductants. For excision of the vector backbone including the
459 kanamycin resistance gene and heat-inducible integrase, cells were electroporated with
460 pE-FLP[41]; transformants were selected on carbenicillin and confirmed for Km^S. pE-FLP was
461 cured by culturing in liquid at 37°C in the absence of carbenicillin, plating for single colonies,

462 and confirming $Carb^S$. Strains were subsequently grown routinely at 37°C. For imaging
463 fluorescent strains on agar, plates were typically incubated at 37°C overnight, transferred to
464 room temperature to allow fluorescence intensity to increase, and then imaged.

465 **Mouse experiments with *E. coli* colonization, antibiotic water, and phage treatment.**
466 Animal procedures were approved by the University of California, San Francisco (UCSF)
467 Institutional Animal Care and Use Committee (IACUC), and animal experiments performed were
468 in compliance with ethical regulations. Specific pathogen free female BALB/c mice from the
469 vendor Taconic were used for all mouse experiments. Streptomycin water was prepared by
470 dissolving USP grade streptomycin sulfate (VWR 0382) in autoclaved tap water to a final
471 concentration of 5 mg/ml and passing through 0.45 μm filtration units. Mice were provided
472 streptomycin water for 1 day, followed by oral gavage of 0.2 ml containing approximately 10^9
473 CFU of streptomycin-resistant *E. coli*. Mice were kept on streptomycin water thereafter to
474 maintain colonization. For selection with β-lactam antibiotics, USP grade ampicillin sodium salt
475 (Teknova A9510) or USP grade carbenicillin disodium salt (Teknova C2110) was also dissolved
476 in the water to a final concentration of 1 mg/ml; carbenicillin was preferred for its increased
477 stability over ampicillin[42]. Drinking water containing streptomycin was prepared fresh weekly;
478 with the addition of a β-lactam antibiotic, it was prepared fresh every 3-4 days. For phage
479 treatment, filtered phage solutions stored at −80°C were thawed and used directly for oral
480 gavage. Unfiltered phage solutions were precipitated by diluting approximately 5-fold in PBS,
481 adding 0.2 volumes phage precipitation solution (20% PEG-8000, 2.5 M NaCl), incubating for 15
482 min on ice, pelleting at 15,000–21,000g for 15 min at 4°C, resuspending in PBS, centrifuging to
483 pellet insoluble matter, and filtering through 0.45 μm. Heat-inactivated phage were prepared by
484 incubating 1 ml aliquots at 95°C in a water bath for 30 min. Streptomycin-treated mice colonized
485 with Sm^R *E. coli* were orally gavaged with 0.2 ml of phage and placed on drinking water
486 containing both streptomycin and carbenicillin.

487 **Enumeration and culture of *E. coli* from mouse feces.** Fecal pellets were collected from
488 individual mice and CFU counts were performed on the same day to determine CFU per gram
489 feces. Briefly, fecal samples (typically 10-40 mg) were weighed on an analytical balance and
490 250 μl to 500 μl PBS or saline was added. Samples were incubated for 5 min at room
491 temperature and suspended by manual mixing and vortexing. Large particulate matter was
492 pelleted by centrifuging at 100g, ten-fold serial dilutions were made in PBS, and 5 μl of each

493 dilution was spotted on Difco MacConkey agar (BD 212123) supplemented with the appropriate
494 antibiotics, i.e., streptomycin (100 μg/ml) or carbenicillin (50 μg/ml). For qualitative assessment
495 of the fluorescent strains in feces, samples were spotted onto LB supplemented with the
496 appropriate antibiotics. For isolating *E. coli* from fecal samples for genomic or plasmid DNA
497 analysis, the fecal suspension was streaked on agar, and single colonies were further
498 streak-purified.

499 **Construction of CRISPR-Cas9 phagemid vectors.** Cultures were grown in LB or TB media
500 supplemented with the appropriate antibiotics. Plasmid DNA was prepared by QIAprep Spin
501 Miniprep Kit (Qiagen 27106), eluted in TE buffer, and incubated at 60°C for 10 min. Samples
502 were quantified using a NanoDrop One spectrophotometer. The vector pCas9[22] was digested
503 with BsaI (NEB R0535) and gel extracted with a QIAquick Gel Extraction Kit (Qiagen 28706).
504 Spacers were generated by annealing and phosphorylating the two oligos (PSP116 and PSP117
505 for GFPT; PSP120 and PSP121 for NT[40]) at 10 μM each in T4 ligation buffer (NEB B0202S) with
506 T4 polynucleotide kinase (NEB M0201S) by incubating at 37°C for 2 h, 95°C for 5 min, and
507 ramping down to 20°C at 5°C/min. The annealed product was diluted 1 in 200 in sterile distilled
508 water and used for directional cloning by ligating (Thermo Scientific FEREL0011) to 60 ng of
509 BsaI-digested, gel extracted pCas9 overnight at room temperature. Ligations were used to
510 transform NEB 5-alpha competent cells (NEB C2987H) and the cloned spacer was verified by
511 Sanger sequencing using primer PSP108. The trailing repeat was later confirmed to lack the
512 starting 5'G, which did not interfere with GFP-targeting function. The 1.8-kb fragment carrying
513 the f1 origin of replication and β-lactamase gene (f1-*bla*) was amplified from pBluescript II with
514 SalI adapters using primers KL215 and KL216 and KOD Hot Start DNA polymerase (Millipore
515 71842-3). The PCR product was purified using a QIAquick PCR Purification Kit (Qiagen 28104),
516 digested with SalI (Thermo Fisher FD0644), gel extracted, and used to ligate to SalI-digested,
517 FastAP-dephosphorylated (Thermo Fisher FEREF0651) vector. Ligations were used to
518 transform DH5α and clones were screened by restriction digest for both possible insert
519 orientations (A or B) using XbaI (Thermo Scientific FD0684) and one of each orientation was
520 saved for both the GFPT and NT phagemids.

521 **Preparation of M13 carrying pBluescript II.** This protocol was adapted from those to generate
522 phage display libraries[43]. XL1-Blue MRF' was transformed with pBluescript II (Agilent 212208).
523 An overnight culture of this strain was prepared in 5 ml LB supplemented with tetracycline (5

524 μg/ml) and carbenicillin (50 μg/ml) and subcultured the following day 1-in-100 into 5 ml 2YT
525 supplemented with the same antibiotics. At an OD_{600} of 0.8, cells were infected with helper
526 phage M13KO7 (NEB N0315S) or VCSM13 (Agilent 200251) at a multiplicity of infection of
527 approximately 10-to-1 for 1 h at 37°C The infected cells were used to seed 2YT supplemented
528 with carbenicillin (100 μg/ml) and kanamycin (25 μg/ml) at 1-in-100, and the culture was grown
529 overnight to produce phage. Cells were pelleted at 10,000g for 15 min, and the supernatant
530 containing phage was transferred. Phage were precipitated by adding 0.2 volumes phage
531 precipitation solution, inverting to mix well, and incubating for 30 min on ice. Phage were
532 pelleted at 15,000g for 15 min at 4°C and the supernatant was discarded. The phage pellet was
533 resuspended in PBS at 1-4% of the culture volume. The resuspension was centrifuged to pellet
534 insoluble material and transferred to a new tube. Glycerol was added to a final concentration of
535 10-15%. Phage preparations were aliquoted into cryovials and stored at −80°C.

536 **Preparation of M13 carrying CRISPR-Cas9 phagemids.** DH5α(HP4_M13)[44] was transformed
537 with the GFPT phagemid (pCas9-GFPT-f1A or pCas9-GFPT-f1B) or the NT phagemid
538 (pCas9-GFPT-f1A or pCas9-GFPT-f1B) and plated on LB media containing carbenicillin and
539 kanamycin. Transformants were inoculated into 5 ml 2YT supplemented with 100 μg/ml
540 carbenicillin and 25 μg/ml kanamycin, incubated overnight, used 1-in-100 to seed 250 ml of the
541 same media, and incubated overnight. Cells were pelleted at 10,000g for 15 min, and the
542 supernatant containing phage was transferred. Phage were precipitated by adding 0.2 volumes
543 phage precipitation solution, inverting to mix well, and incubating for 30 min on ice. Phage were
544 pelleted at 20,000g for 20 min at 4°C with slow deceleration. The supernatant was completely
545 removed, phage were resuspended in PBS at 1% of the culture volume, and glycerol was added
546 to a final concentration of 10-15%. The phage solution was centrifuged at 21,000g to pellet
547 insoluble matter, filtered through 0.45 μm, and stored at −80°C.

548 **Titration of M13 phage carrying phagemid DNA.** Phage titer was determined using indicator
549 strain XL1-Blue MRF' or Sm^R W1655 F+. An overnight culture of the indicator strain in LB
550 supplemented with the appropriate antibiotics was subcultured 1-in-100 or 1-in-200 into fresh
551 media and grown to an OD_{600} of 0.8. To estimate titer, serial ten-fold dilutions of the phage
552 preparation were made in PBS, and 10 μl of each dilution was used to infect 90 μl of cells. After
553 incubating at 37°C for 30 min with shaking, 10 μl of the infection mix was spotted onto LB
554 supplemented with carbenicillin. For more accurate titration, 100 μl of phage dilutions were

555 mixed with 900 µl cells in culture tubes, incubated at 37°C for 30 min with shaking, and 100 µl
556 was plated on LB carbenicillin.

557 **Enumeration of viable M13 from mouse feces.** Mice were orally gavaged with 6×10^{13}
558 M13(pBluescript II) or as negative controls, heat-inactivated phage or PBS. Approximately 100
559 mg of feces were collected at 0, 3, 6, 9, and 24 h post-gavage, and samples at each timepoint
560 were processed immediately. 500 µl PBS was added, samples were incubated for 5 min at room
561 temperature, then suspended by manual mixing and vortexing. Samples were centrifuged at
562 21,000g for 1 min, the supernatant was transferred to a new tube, and phage titer was
563 determined against indicator strain XL1-Blue MRF' by diluting samples in PBS, incubating with
564 cells, and plating on LB supplemented with carbenicillin. For all dilutions and the undiluted
565 suspension, 10 µl was used to infect 90 µl cells; additionally, for the undiluted suspension, 100
566 µl was used to infect 900 µl cells to maximize the limit of detection.

567 **Assay for acid survival.** Phage M13(pBluescript II) stored in PBS was diluted 1-in-100 in
568 saline. Solutions varying in pH (1.2, 2, 3, 4, 5, 6, and 7) were prepared by mixing different ratios
569 of 0.2 M sodium phosphate dibasic and 0.1 M citric acid and adjusting with concentrated HCl.
570 200 µl of each pH solution was transferred to the wells of a microtiter plate, and 10 µl of phage
571 was added containing 1×10^{9} M13(pBluescript II). Phage were incubated in the solution, and 10
572 µl was sampled at 5, 15, and 60 min. Samples were diluted 1-in-100 in PBS to make acidic
573 samples neutral and phage titer was determined against indicator strain XL1-Blue MRF' by
574 plating on LB supplemented with carbenicillin. Solution-only controls were assayed
575 simultaneously and cells were plated on LB to confirm viability of the indicator strain in the
576 presence of samples originating from an acidic pH.

577 **Targeting experiments *in vitro* with M13 CRISPR-Cas9.** Overnight cultures of fluorescently
578 marked Sm^{R} W1655 F+ *sfgfp* and *mcherry* were prepared in LB supplemented with
579 streptomycin, subcultured 1 in 200 into fresh media, and grown to an OD_{600} of 0.8. 900 µl cells
580 (approximately 1×10^{9}) was transferred to a culture tube, 100 µl phage (approximately 1×10^{10} for
581 f1A vectors and approximately 5×10^{10} for f1B vectors) was added, and the tube was incubated
582 at 37°C for 30 min. The infection culture was transferred to a microfuge tube, cells were pelleted
583 at 21,000g for 1 min, and the supernatant was removed. Cells were washed twice by adding 1
584 ml PBS, vortexing, pelleting cells, and removing supernatant. Cells were resuspended in 1 ml

585 PBS, and ten-fold serially diluted in PBS. 10 µl of each dilution was spotted onto LB
586 supplemented with carbenicillin and 100 µl was plated on larger plates for isolating single
587 colonies for analysis. Colonies were picked and streak-purified four times to ensure phenotypic
588 homogeneity and clonality.

589 **Co-culture experiments with *sfgfp* and *mcherry*-marked strains infected with M13**
590 **CRISPR-Cas9.** Overnight cultures of fluorescently marked Sm^R W1655 F+ *sfgfp* and *mcherry*
591 were prepared in LB supplemented with streptomycin. For each culture, three serial ten-fold
592 dilutions were made in PBS, followed by a fourth ten-fold dilution into LB. Equal volumes of
593 each were combined and 5 ml aliquots were transferred to culture tubes. Using a CFU assay,
594 the input was determined to be 6×10^6 CFU of each strain or 1×10^7 CFU total. 10 µl (5×10^9) M13
595 carrying CRISPR-Cas9 was added, the co-culture was incubated at 37°C for 30 min, and
596 carbenicillin was added to a final concentration of 100 µg/ml. The co-culture was sampled for
597 the t = 0 timepoint and then incubated for 24 h with further sampling every 4 h. At each
598 timepoint, 200 µl was taken; 100 µl was used to assay carbenicillin in the media (see section:
599 Carbenicillin bioassay) and the remaining 100 µl was used for plating as follows. To the 100 µl
600 sample of culture, 900 µl was added and cells were washed by vortexing. Cells were pelleted by
601 centrifuging at 21,000g for 1 min, and 900 µl of the supernatant was removed. To remove
602 residual phage and antibiotic, the wash was repeated once more by adding 900 µl PBS,
603 vortexing, pelleting cells, and removing 900 µl. Cells were resuspended in the remaining 100 µl.
604 Serial ten-fold dilutions were made in PBS and 10 µl of each dilution was spotted onto LB or LB
605 carbenicillin.

606 **Carbenicillin bioassay.** Cultures were sampled over time, cells were pelleted at 21,000g for 1
607 min, and the supernatant was transferred to a new tube and frozen at −20°C until all timepoints
608 were collected. The supernatants were thawed and assayed using a Kirby-Bauer disk diffusion
609 test. An overnight culture of the indicator organism (*Bacillus subtilis* 168) was diluted in saline to
610 an OD_{600} of 0.1. A cotton swab was dipped into this dilution and spread across LB agar,
611 antibiotic sensitivity disks (Fisher Scientific S70150A) were overlaid using tweezers, and 20 µl of
612 the supernatant was applied to the disk. At the same time, carbenicillin standards were
613 prepared from 1 µg/ml to 100 µg/ml and also applied to discs. Plates were incubated overnight
614 at 37°C and imaged the following morning.

615 **Flow cytometry.** For turbid *in vitro* cultures, samples were diluted 1-in-10,000 in PBS. For
616 mouse fecal pellets, samples were used fresh or thawed from −80°C, and suspended in 500 µl
617 PBS by manual mixing and vortexing. Fecal suspensions were incubated aerobically at 4°C
618 overnight to improve fluorescence signal (**Supplementary Fig. 18**). Samples were vortexed to
619 mix, large particulate matter was pelleted by centrifuging at 100g for 30 seconds, and the
620 sample was diluted 1-in-100 in PBS. Samples were run on a BD LSRFortessa flow cytometer
621 using a 530/30 nm filter for GFP fluorescence and 610/20 nm for mCherry fluorescence, with
622 the following voltages: 750 V for FSC, 400 V for SSC, 700 V for mCherry, and 700-800 V (in
623 vivo) or 650 V (in vitro) for GFP. Flow cytometry data were analyzed in R using packages
624 flowCore (v1.52.1)[45], Phenoflow (v1.1.2)[46], and ggcyto (v1.14.0)[47]. Typically, between 10,000
625 and 100,000 events were collected per sample, and data were rarefied after gating on FSC and
626 SSC. Background events were accounted for on a per-mouse basis. For co-colonization with the
627 *sfgfp*-marked and *mcherry*-marked strains, GFP+ and mCherry+ events from Day -3 (pre-*E. coli*)
628 were used to subtract background at subsequent timepoints. For colonization with the
629 double-marked strain, GFP+ mCherry+ events from Day -5 (pre-*E. coli*) were used to subtract
630 background of double fluorescence at subsequent timepoints, and GFP- mCherry+ events from
631 Day 0 (pre-phage) were used to subtract background of red fluorescence at subsequent
632 timepoints. For exclusion of timepoints due to lack of colonization, the background threshold was
633 calculated as the maximum background observed for that population across all timepoints multiplied
634 by a factor of three. See Code Availability for more information.

635 **Quick extraction and PCR analysis of genomic DNA from *in vitro* or *in vivo* isolates.**
636 Genomic DNA was extracted crudely to use as template for PCR. Briefly, 1.5 ml to 3 ml of
637 culture was transferred to a microfuge tube, cells were pelleted by centrifuging, and the
638 supernatant was discarded. The pellet was frozen, allowed to thaw on ice, resuspended in 100
639 µl TE, and incubated at 100°C for 15 min in an Eppendorf ThermoMixer. Samples were cooled
640 on ice, cell debris was pelleted by centrifuging at 21,000g for 1 min, the supernatant was
641 transferred to a new tube, and diluted 1-in-100 in TE to use as template DNA. PCR was
642 performed using KOD Hot Start DNA polymerase (Millipore 71842-3) using primers
643 KL207/KL200 for the *sfgfp* gene and primers BAC338F/BAC805R for the 16S rRNA gene[48].

644 **Extraction of DNA for hybrid assembly.** *E. coli* strains KL68 (W1655 F+ or ATCC 23590),
645 KL114 (W1655 F+ *rpsL*-Sm^R *sfgfp*), and KL204 (W1655 F+ *rpsL*-Sm^R *sfgfp mcherry*) were

646 cultured in 50 ml LB supplemented with streptomycin. Cells were collected by centrifuging at
647 6,000g for 10 min at room temperature, washed in 10 ml 10 mM Tris 25 mM EDTA (pH 8.0), and
648 resuspended in 4 ml of the same buffer. 12.5 mg lysozyme (Sigma-Aldrich L6876), 100 µl 5 M
649 NaCl, and 50 µl 10 mg/ml RNase A (Thermo-Fisher EN0531) were added and the mixture was
650 incubated at 37°C for 15 min. To lyse cells, 350 µl 5 M NaCl, 20 µl 20 mg/ml Proteinase K
651 (Ambion AM2546), and 500 µl 10% SDS were added, and the mixture was incubated at 60°C
652 for 1 h with gentle inversions. 2.75 ml of 7.5 M ammonium acetate was added, and the mixture
653 was incubated on ice 20 min to precipitate proteins. Debris was removed by centrifuging
654 20,000g for 10 min and the supernatant was transferred to a new tube. To extract, an equal
655 volume of chloroform was added and mixed; phases were separated by centrifuging at 2,000g
656 for 10 min, and the aqueous phase was transferred to a new tube. To precipitate the DNA, 1
657 volume of isopropanol was added, and the tube was inverted until a white precipitate formed.
658 The DNA was pelleted by centrifuging at 2,000g for 10 min and the supernatant was removed.
659 The pellet was washed with 500 µl ice-cold 70% ethanol, allowed to dry, 1 ml TE was added,
660 and the pellet allowed to dissolve overnight at 4°C. To further remove RNA, 250 µl of the
661 genomic prep was transferred to a new tube, 12.5 µl 10 mg/ml RNase A was added, and the
662 mixture was incubated at 37°C for 2 h with mixing every 30 min. To precipitate the DNA, 0.1
663 volume of 3 M sodium acetate was added followed by 3 volumes of 100% ethanol, and the
664 mixture was inverted until a white precipitate formed. DNA was pelleted by centrifuging at
665 2,000g for 10 min, the supernatant was removed, the pellet washed with 100 µl 70% ethanol,
666 allowed to dry, and resuspended in 100 µl TE. Samples were quantified by Qubit dsDNA BR
667 Assay and DNA integrity was confirmed by 0.4% agarose gel electrophoresis using GeneRuler
668 High Range DNA Ladder (Thermo-Fisher FERSM1353). DNA was used for both Oxford
669 Nanopore sequencing and Illumina sequencing.

670 **Illumina whole genome sequencing.** DNA concentration was quantified using PicoGreen
671 (ThermoFisher). Genomic DNA was normalized to 0.18 ng/µl for library preparation. Nextera XT
672 libraries were constructed in 384-well plates using a custom, miniaturized version of the
673 standard Nextera XT protocol. Small volume liquid handlers such as the Mosquito HTS (TTP
674 LabTech) and Mantis (Formulatrix) were used to aliquot precise reagent volumes of <1.2 µl to
675 generate a total of 4 µl per library. Libraries were normalized and 1.2 µl of each normalized
676 library was pooled and sequenced on the Illumina NextSeq or MiSeq platform using 2x146 bp
677 configurations. 12 bp unique dual indices were used to avoid index hopping, a phenomenon

678 known to occur on ExAmp based Illumina technologies. See Data Availability for more
679 information.

680 **Oxford Nanopore sequencing and hybrid Nanopore/Illumina assembly.** PCR-free long read
681 libraries were prepared using the Ligation Sequencing Kit (SQK-LSK109), multiplexed using the
682 Native Barcoding Kit (EXP-NBD114), and sequenced on the MinION platform using flow cell
683 version MIN106 (Oxford Nanopore Technologies). Basecalling of MinION raw signals was done
684 using Guppy (v2.2.2, Oxford Nanopore Technologies). Reads were demultiplexed with qcat
685 (v1.1.0, Oxford Nanopore Technologies). Quality control was achieved using porechop (v0.2.3
686 seqan2.1.1) (https://github.com/rrwick/Porechop) using the discard middle option. Reads were
687 filtered using NanoFilt (v2.6.0)[49] with the following parameters: minimum average read quality
688 score of 10 (-q 10) and minimum read length of 100 (-l 100). Illumina reads were quality filtered
689 using fastp (v0.20.1)[50] with the following parameters: cut front, cut tail, cut window size 4, cut
690 mean quality 20, length required 60. Filtered MinION and Illumina reads were then provided to
691 Unicycler (v0.4.8)[51] for hybrid assembly; default parameters were used unless otherwise noted.
692 See Data Availability for more information.

693 **Analysis of isolates after *in vitro* or *in vivo* M13-mediated delivery of phagemid.** Isolates
694 were cultured on LB or Difco MacConkey agar plates supplemented with carbenicillin or both
695 carbenicillin and streptomycin. Isolates from *in vitro* GFP-targeting experiments were streak
696 purified 4 times on agar to ensure clonality. Isolates from *in vivo* experiments were obtained by
697 suspending a fecal pellet in 500 µl PBS, streaking the suspension onto agar, followed by streak
698 purification of single colonies. For DNA extraction, single colonies were inoculated into LB or TB
699 supplemented with the appropriate antibiotics. For analysis of the phagemid, plasmid DNA was
700 extracted using a QIAprep Spin Miniprep Kit (Qiagen 27106), eluted in TE buffer, and incubated
701 at 60°C for 10 min. DNA was quantified using a NanoDrop One spectrophotometer and 200-600
702 ng was digested with FastDigest restriction enzymes (KpnI, Thermo Scientific FD0524; XbaI,
703 Thermo Scientific FD0684) for 10 min at 37°C followed by gel electrophoresis. Spacer
704 sequences on phagemids were confirmed by Sanger sequencing using primer PSP108. For
705 genome sequencing, genomic DNA was either extracted using a DNeasy Blood & Tissue Kit
706 (Qiagen 69506) or an in-house protocol. Briefly, isolates were cultured in 3 ml TB supplemented
707 with streptomycin and carbenicillin. Cells were pelleted and resuspended in 460 ul of freshly
708 prepared buffer [per sample: 400 ul 10 mM Tris (pH 8.0) 25 mM ETDA, 50 µl 5 M NaCl, and 10

709 µl 10 mg/ml RNase A (Thermo-Fisher EN0531)]. 50 µl 10% SDS was added, mixed well, and
710 samples were incubated at 60°C for 1 h with periodic inversions. 260 µl of 7.5 M ammonium
711 acetate was added, and the mixture was incubated on ice 20 min to precipitate proteins.
712 Precipitate was removed by centrifuging 21,000g for 5 min and the supernatant was transferred
713 to a new tube. To extract, an equal volume of chloroform was added and mixed; phases were
714 separated by centrifuging at 21,000g for 2.5 min. The aqueous phase was transferred to a new
715 tube, centrifuged at 21,000g for 2.5 min, and 500 ul was transferred to a new tube. To
716 precipitate the DNA, 500 µl isopropanol was added, and the tube was inverted until a white
717 precipitate formed. Using a pipette tip, the clump was transferred to a new tube, washed with
718 100 µl cold 70% ethanol, and allowed to dry. 50 µl TE was added and the pellet was allowed to
719 dissolve at 4°C overnight. DNA integrity was confirmed by gel electrophoresis and used for
720 Illumina whole genome sequencing (see section: Illumina whole genome sequencing).
721 Sequence reads were quality filtered using fastp (v0.20.1)[50] and reads were aligned using
722 bowtie2 (v2.3.5.1)[52] to reference genomes and phagemid sequences; complete reference
723 genomes were generated using hybrid assembly (see section: Oxford Nanopore sequencing
724 and hybrid Nanopore/Illumina assembly). For $Carb^R$ isolates obtained after delivery of the
725 pBluescript II phagemid, reads were simultaneously aligned to the genome of strain KL68
726 (W1655 F+) and the pBluescript II sequence (NCBI accession X52329.1). For $Carb^R$ isolates
727 obtained after delivery of CRISPR-Cas9 phagemids, reads were simultaneously aligned to the
728 genome of the strain used for *in vitro* (KL114; *sfgfp*) or *in vivo* (KL204; *sfgfp mcherry*)
729 experiments and the sequence of the delivered phagemid. For isolates from GFP-targeting
730 experiments, deletions were visualized by using samtools (v1.9)[53] to filter multi-mapping and
731 low-quality read alignments with MAPQ<2 (view -q 2), and depth was calculated using a sliding
732 window of 20; breseq (v0.35.4)[54] was used to assess deletion size. See Data Availability for
733 more information.

734 Data availability

735 All raw sequence data have been deposited at the NCBI Sequence Read Archive under
736 BioProject PRJNA642411. For 16S rRNA gene sequencing, Illumina sequence data have been
737 deposited with accession numbers SRR12118792 to SRR12118959. For sequencing of isolates
738 from mouse fecal samples after delivery of pBluescript II, Illumina sequence data deposited with
739 accession numbers SRR14278062 to SRR14278073. For Nanopore/Illumina hybrid assembly of

740 reference strain KL68 (W1655 F+ or ATCC 23590) as well as Sm^R fluorescent derivatives KL114
741 (*sfgfp*) and KL204 (*sfgfp mcherry*), Illumina sequence data have been deposited with accession
742 numbers SRR14296642 to SRR14296644 and Oxford Nanopore MinION data with accession
743 numbers SRR14297452 to SRR14297454. For sequencing of isolates after targeting with
744 phage-delivered CRISPR-Cas9, Illumina sequence data deposited with accession numbers
745 SRR14289086 to SRR14289109. Supporting data have also have been made publicly available
746 on Github: https://github.com/turnbaughlab/2021_Lam_M13_CRISPRCas9

747 Code availability

748 Bash scripts and R Markdown documents for flow cytometry, 16S rRNA gene sequencing, and
749 genomic deletion analyses have been made publicly available on GitHub:
750 https://github.com/turnbaughlab/2021_Lam_M13_CRISPRCas9

751 ## References

752 32. Gohl, D. M. *et al.* Systematic improvement of amplicon marker gene methods for increased
753 accuracy in microbiome studies. *Nat. Biotechnol.* **34**, 942–949 (2016).
754 33. Bolyen, E. *et al.* Reproducible, interactive, scalable and extensible microbiome data
755 science using QIIME 2. *Nat. Biotechnol.* **37**, 852–857 (2019).
756 34. Callahan, B. J. *et al.* DADA2: High-resolution sample inference from Illumina amplicon data.
757 *Nat. Methods* **13**, 581–583 (2016).
758 35. McMurdie, P. J. & Holmes, S. phyloseq: an R package for reproducible interactive analysis
759 and graphics of microbiome census data. *PLoS One* **8**, e61217 (2013).
760 36. Smith, S. D. phylosmith: an R-package for reproducible and efficient microbiome analysis
761 with phyloseq-objects. *Journal of Open Source Software* **4**, 1442 (2019).
762 37. Datsenko, K. a. & Wanner, B. L. One-step inactivation of chromosomal genes in
763 *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 6640–6645
764 (2000).
765 38. Jensen, S. I., Lennen, R. M., Herrgård, M. J. & Nielsen, A. T. Seven gene deletions in
766 seven days: Fast generation of *Escherichia coli* strains tolerant to acetate and osmotic
767 stress. *Sci. Rep.* **5**, 17874 (2015).
768 39. Timms, A. R., Steingrimsdottir, H., Lehmann, A. R. & Bridges, B. A. Mutant sequences in
769 the *rpsL* gene of *Escherichia coli* B/r: mechanistic implications for spontaneous and
770 ultraviolet light mutagenesis. *Mol. Gen. Genet.* **232**, 89–96 (1992).
771 40. Vigouroux, A., Oldewurtel, E., Cui, L., Bikard, D. & van Teeffelen, S. Tuning dCas9's ability
772 to block transcription enables robust, noiseless knockdown of bacterial genes. *Mol. Syst.*
773 *Biol.* **14**, e7899 (2018).
774 41. St-Pierre, F. *et al.* One-step cloning and chromosomal integration of DNA. *ACS Synth. Biol.*
775 **2**, 537–541 (2013).
776 42. Bobrowski, M. & Borowski, E. Interaction between carbenicillin and β-lactamases from
777 Gram-negative bacteria. *Microbiology* **68**, 263–272 (1971).
778 43. Tonikian, R., Zhang, Y., Boone, C. & Sidhu, S. S. Identifying specificity profiles for peptide
779 recognition modules from phage-displayed peptide libraries. *Nat. Protoc.* **2**, 1368–1386
780 (2007).
781 44. Praetorius, F. *et al.* Biotechnological mass production of DNA origami. *Nature* **552**, 84–87

782 (2017).
783 45. Hahne, F. *et al.* flowCore: a Bioconductor package for high throughput flow cytometry. *BMC*
784 *Bioinformatics* **10**, 106 (2009).
785 46. Props, R., Monsieurs, P., Mysara, M., Clement, L. & Boon, N. Measuring the biodiversity of
786 microbial communities by flow cytometry. *Methods Ecol. Evol.* **7**, 1376–1385 (2016).
787 47. Van, P., Jiang, W., Gottardo, R. & Finak, G. ggCyto: next generation open-source
788 visualization software for cytometry. *Bioinformatics* **34**, 3951–3953 (2018).
789 48. Yu, Y., Lee, C., Kim, J. & Hwang, S. Group-specific primer and probe sets to detect
790 methanogenic communities using quantitative real-time polymerase chain reaction.
791 *Biotechnol. Bioeng.* **89**, 670–679 (2005).
792 49. De Coster, W., D'Hert, S., Schultz, D. T., Cruts, M. & Van Broeckhoven, C. NanoPack:
793 visualizing and processing long-read sequencing data. *Bioinformatics* **34**, 2666–2669
794 (2018).
795 50. Chen, S., Zhou, Y., Chen, Y. & Gu, J. fastp: an ultra-fast all-in-one FASTQ preprocessor.
796 *Bioinformatics* **34**, i884–i890 (2018).
797 51. Wick, R. R., Judd, L. M., Gorrie, C. L. & Holt, K. E. Unicycler: Resolving bacterial genome
798 assemblies from short and long sequencing reads. *PLoS Comput. Biol.* **13**, e1005595
799 (2017).
800 52. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods*
801 **9**, 357–359 (2012).
802 53. Li, H. *et al.* The sequence alignment/map format and SAMtools. *Bioinformatics* **25**,
803 2078–2079 (2009).
804 54. Deatherage, D. E. & Barrick, J. E. Identification of mutations in laboratory-evolved microbes
805 from next-generation sequencing data using breseq. *Methods Mol. Biol.* **1151**, 165–188
806 (2014).

Acknowledgments

We thank the UCSF Gnotobiotics Core staff (Jessie Turnbaugh, Kimberly Ly, and Jolie Ma) for animal care support and assistance with learning animal procedures. We are grateful to Daryll Gempis, Bernarda Lopez, and Ernesto Valencia of the UCSF G.W. Hooper Foundation for laboratory and administrative support. We thank Katja Engel (University of Waterloo) for help in translating the original publication on the isolation of M13 written in German. We are grateful to Antoine Vigouroux (Institut Pasteur) for generously sharing MG1655 derivatives carrying *sfgfp* and *mcherry* marker genes. We thank the Chan-Zuckerberg BioHub for sequencing through the Microbiome Initiative. We are grateful to Joseph Bondy-Denomy and Oren Rosenberg (UCSF) for constructive criticism of the manuscript. The graphic of a laboratory mouse was adapted from a Wikimedia Commons graphic distributed under CC-BY-SA-4.0 by Gwilz. KNL and PS were both supported by a postdoctoral fellowship from the Canadian Institutes of Health Research (CIHR), JEB by a postdoctoral fellowship from the Natural Sciences and Engineering Research Council of Canada (NSERC), MA by an F32 fellowship from the National Institutes of Health (F32AI147456-01), and PSP by a scholarship from the UCSF Discovery Fellows Program. This work was supported by the National Institutes of Health (PJT, R01HL122593, R01AT011117; RRN, K08AR073930). PJT held an Investigators in the Pathogenesis of Infectious Disease Award from the Burroughs Wellcome Fund, was a Chan Zuckerberg Biohub investigator, and was a Nadia's Gift Foundation Innovator supported, in part, by the Damon Runyon Cancer Research Foundation (DRR-42-16) and the Searle Scholars Program (SSP-2016-1352).

Author contributions

KNL and PJT conceived the ideas. KNL supervised laboratory work and analyzed the data. KNL, PS, PSP, and PJT designed the experiments. KNL and PSP constructed plasmids. KNL constructed *E. coli* strains with assistance from PS. KNL performed phage and animal experiments with assistance from PSP. KL prepared samples and PS performed antibiotic assays. KNL, PSP and MN prepared samples, FBY performed 16S rRNA gene sequencing, and KNL and JEB analyzed the data. MA performed flow cytometry and KNL analyzed the data. KNL analyzed plasmid DNA from mouse fecal isolates with assistance from MN. KNL prepared

samples, FBY and AMW performed Nanopore and Illumina whole genome sequencing, RRN performed the hybrid assembly of reference genomes, and KNL analyzed the genomic deletion data. PJT provided reagents and materials. KNL made the figures. KNL and PJT wrote the manuscript; PS, PSP, MA, JEB, and FBY assisted with editing.

Competing Interests

KNL, PS, and PJT are listed inventors on a U.S. provisional patent application related to this work (33167/55262P1). PJT is on the scientific advisory boards for Kaleido, Pendulum, Seres, and SNIPRbiome. All other authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to PJT.

SUPPLEMENTARY INFORMATION

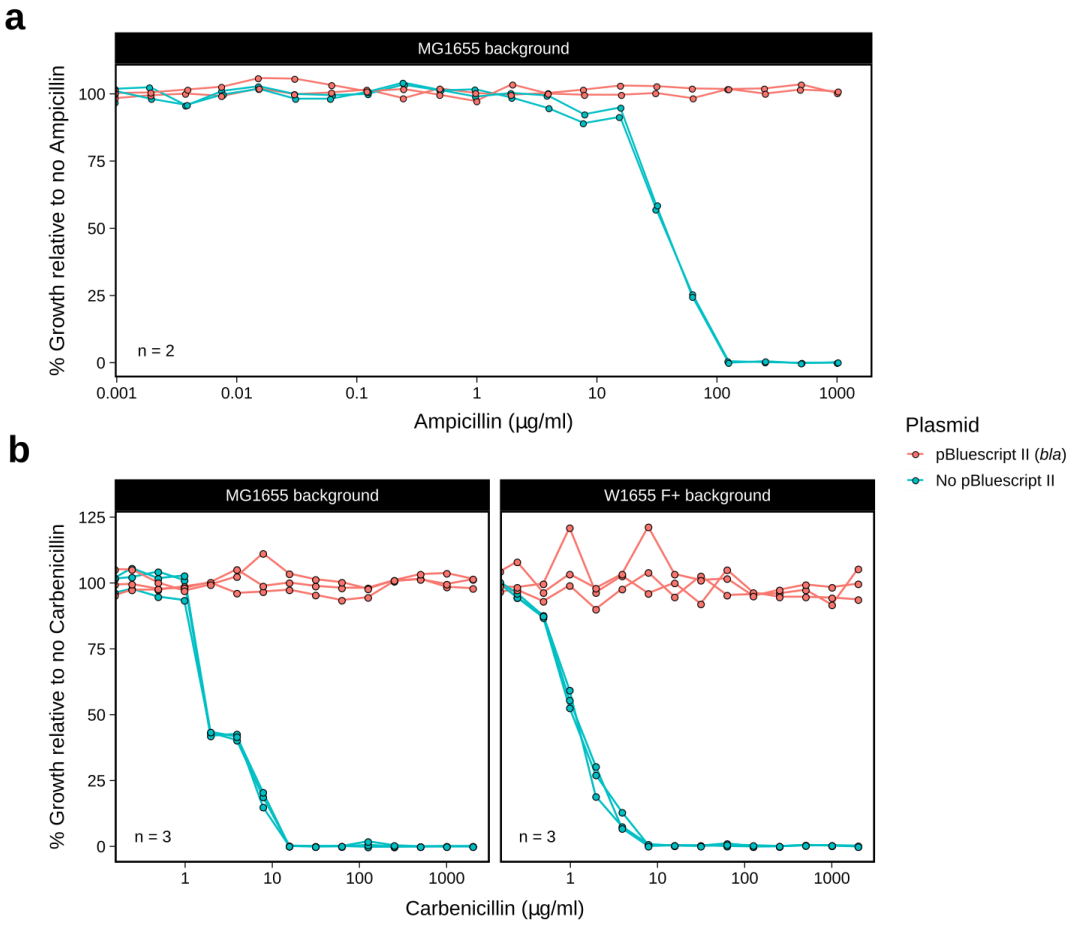
Phage-delivered CRISPR-Cas9 for strain-specific depletion and genomic deletions in the gut microbiota

Kathy N. Lam[1], Peter Spanogiannopoulos[1], Paola Soto-Perez[1], Margaret Alexander[1], Matthew J. Nalley[1], Jordan E. Bisanz[1], Renuka Nayak[1], Allison M. Weakley[2], Feiqiao B. Yu[3], Peter J. Turnbaugh[1,3,*]

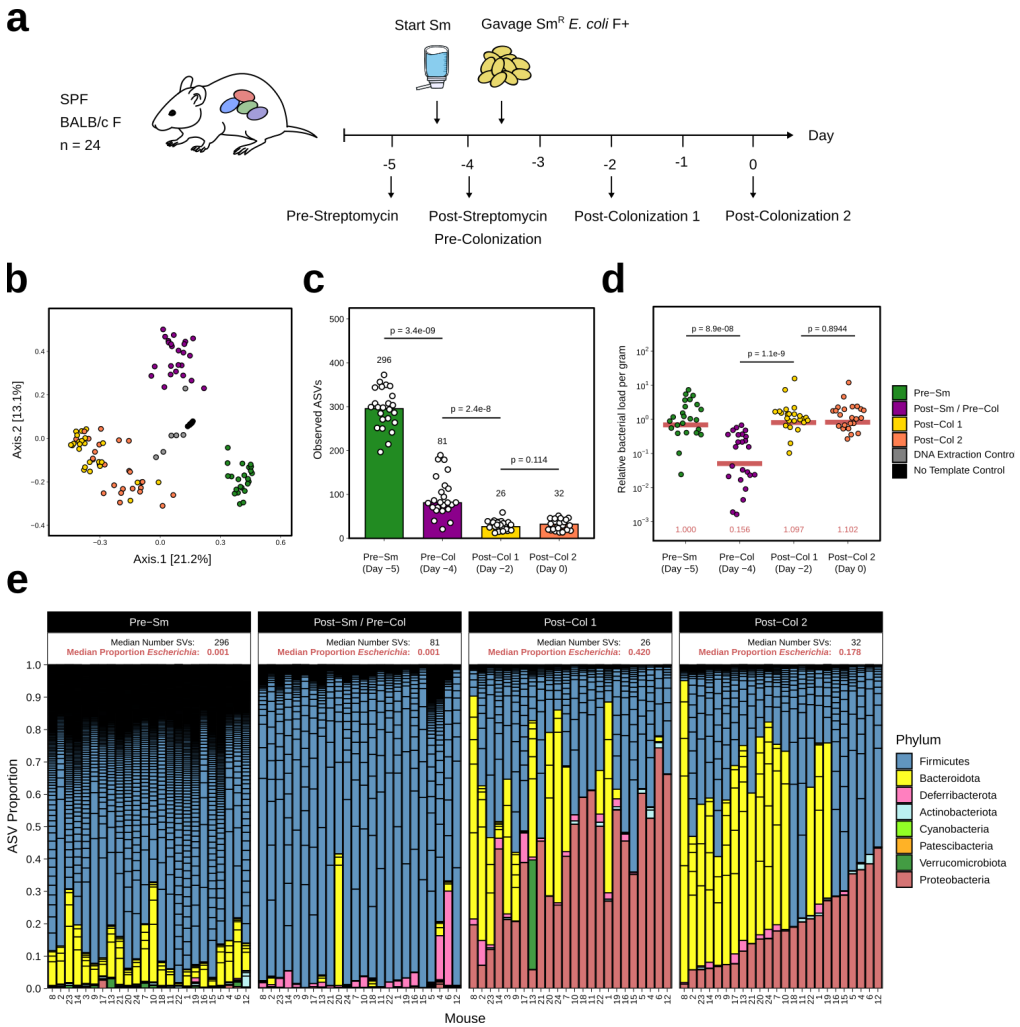
[1]Department of Microbiology and Immunology, University of California San Francisco, USA
[2]Stanford Microbiome Therapies Initiative, Stanford University, Stanford, USA
[3]Chan-Zuckerberg BioHub, San Francisco, USA

*Correspondence to: Peter.Turnbaugh@ucsf.edu

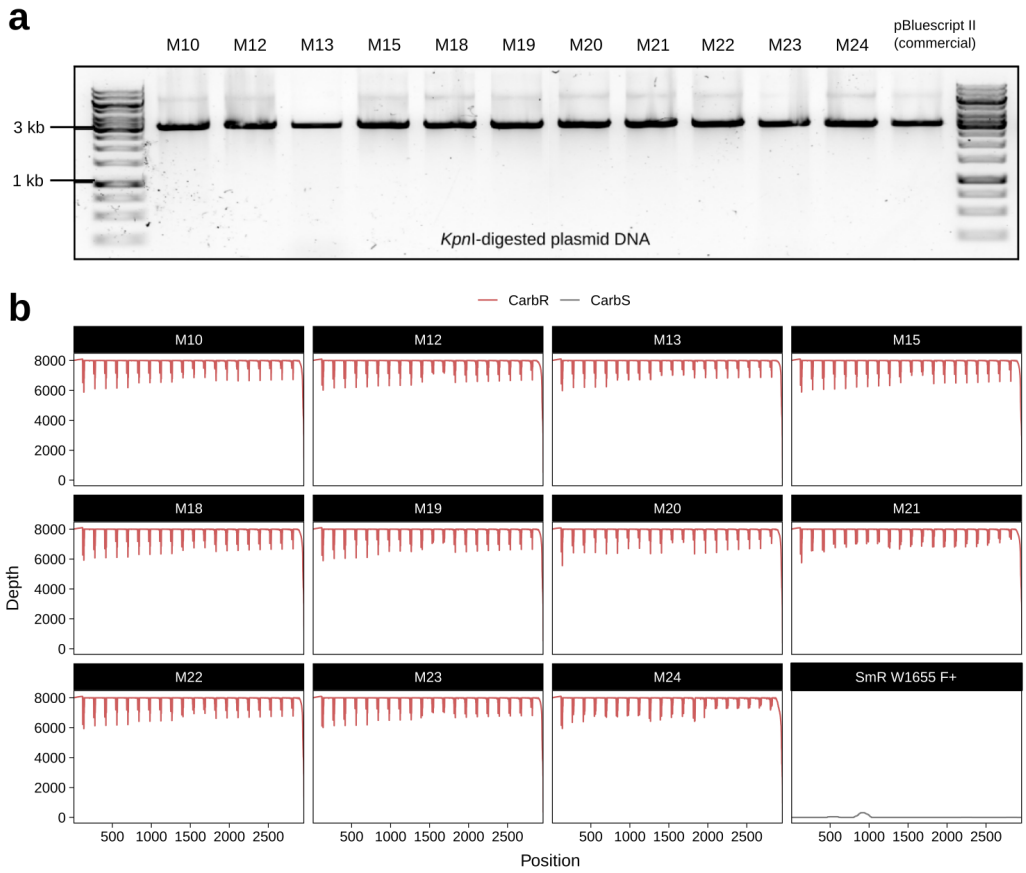
Supplementary Figures



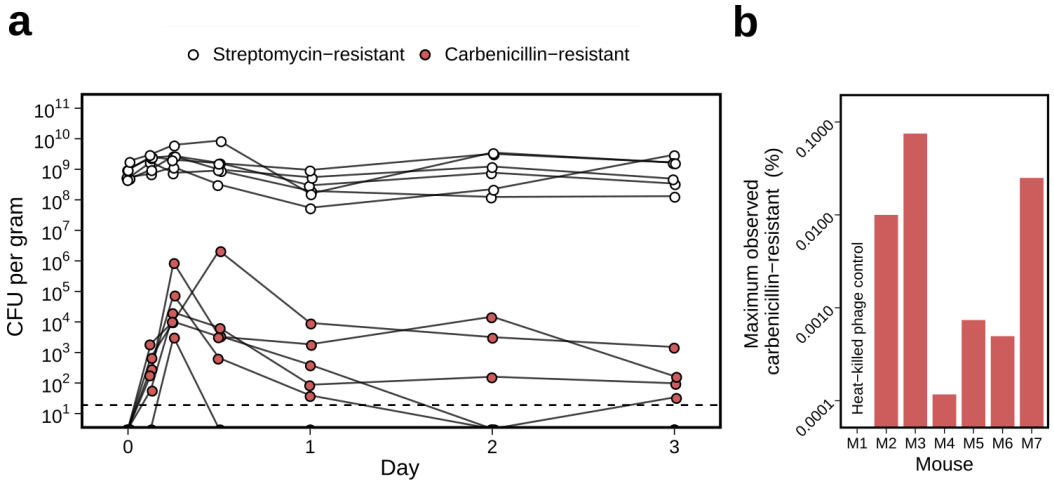
Supplementary Figure 1. The plasmid pBluescript II confers resistance to beta-lactam antibiotics exceeding 1 mg/ml. (a) Minimum inhibitory concentration (MIC) assay for ampicillin in the *E. coli* MG1655 background. Harbouring pBluescript II, the strain exhibits an MIC of >1 mg/ml; in the absence of 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.



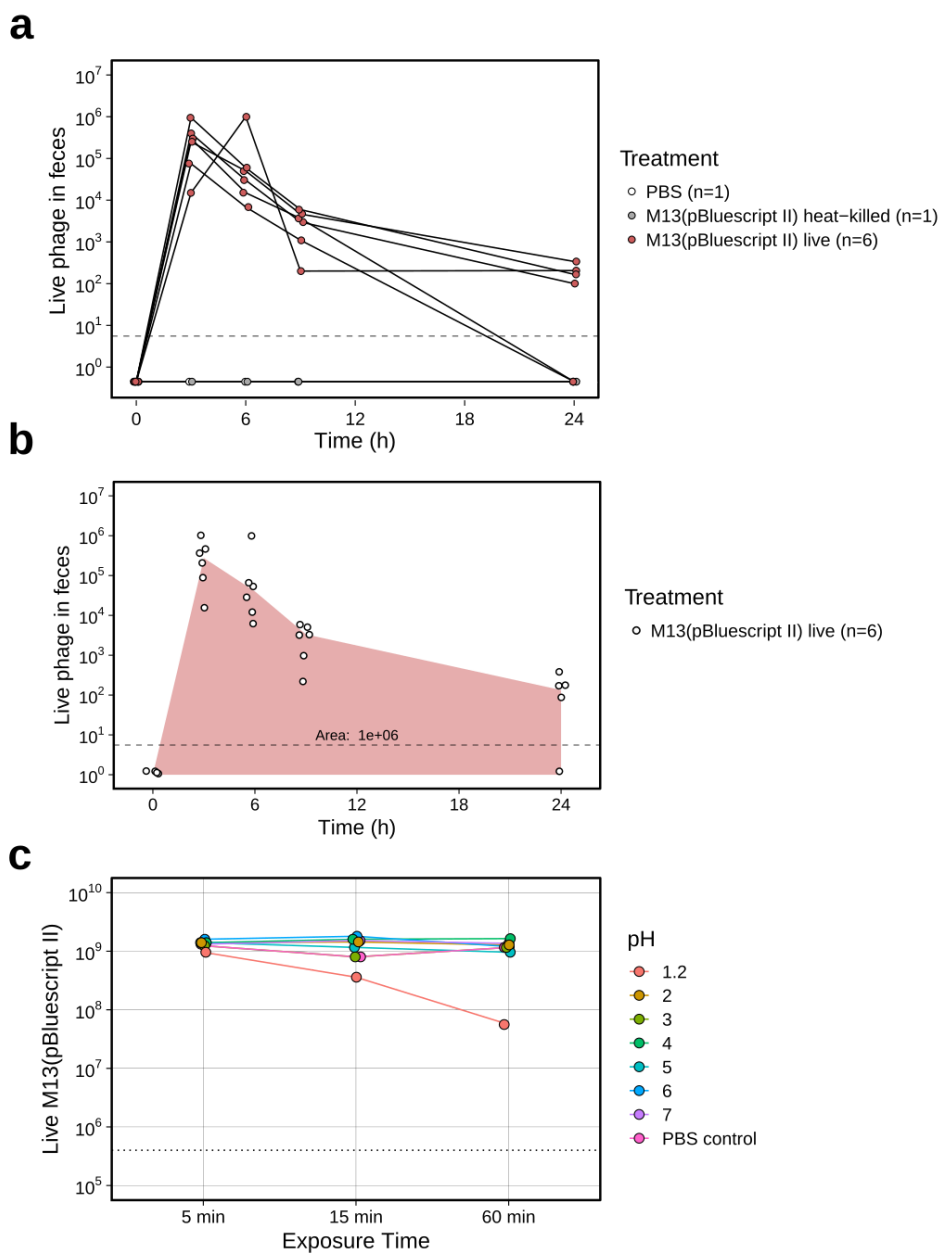
Supplementary Figure 2. Streptomycin treatment reduces bacterial diversity and allows Sm^R *E. coli* to colonize at a high proportion in conventionally raised mice. (a) Typical timeline for a mouse experiment using water containing streptomycin (Sm) to allow colonization by Sm^R *E. coli.* Fecal samples 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.



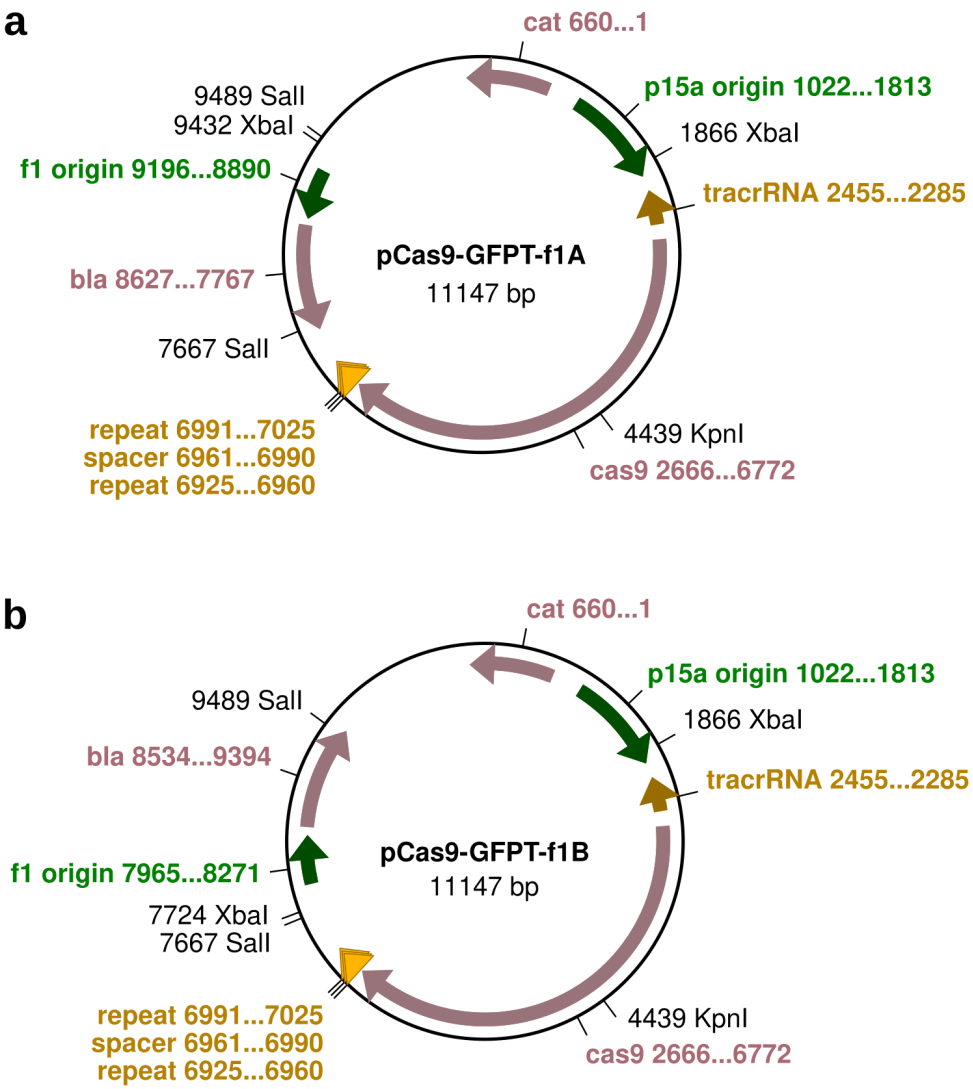
Supplementary Figure 3. Successful transfer of pBluescript II plasmid from M13 phage to *E. coli* cells *in vivo*. (a) Diagnostic digest of extracted plasmid DNA from fecal isolates is consistent with pBluescript II. Plasmid DNA was recovered from carbenicillin-resistant ($Carb^R$) colonies isolated 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.



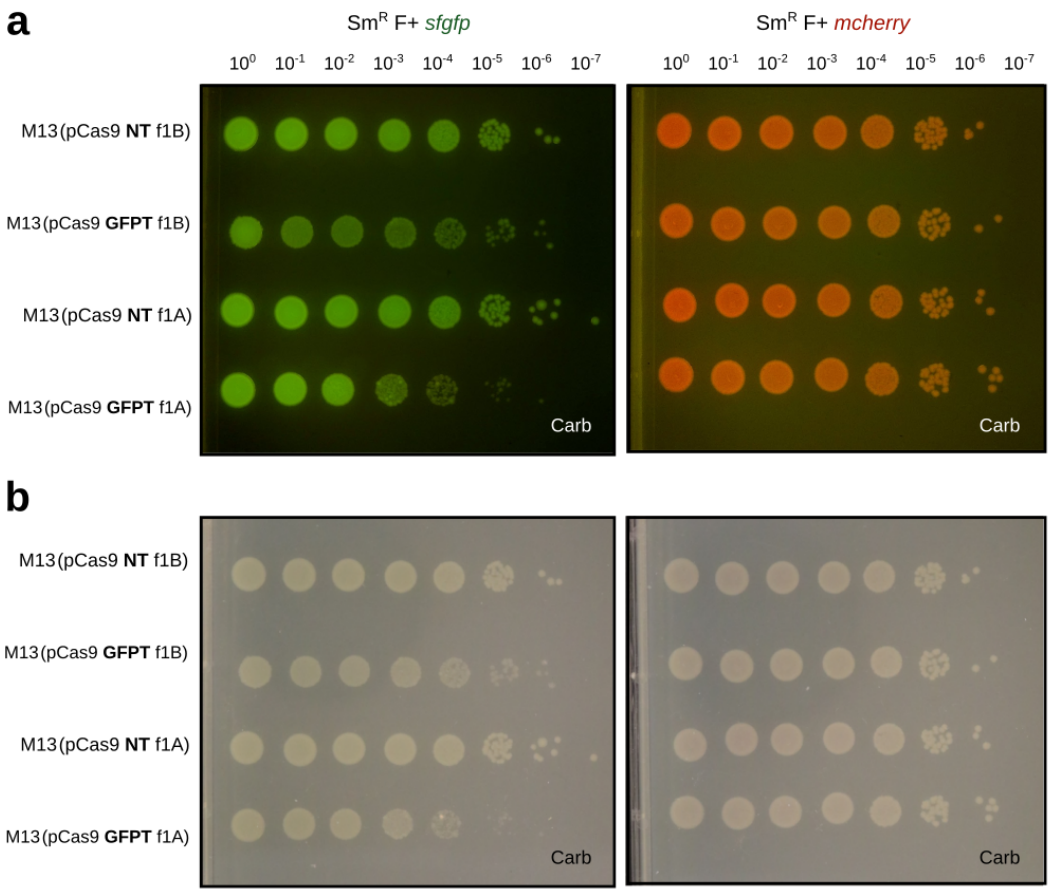
Supplementary Figure 4. Low frequency of carbenicillin-resistant *E. coli* following treatment with M13(pBluescript II) in the absence of carbenicillin selection. (a) Sm^R *E. coli* W1655 F+ was introduced and maintained in streptomycin-treated mice (n=6). Oral gavage of 10^{13} M13(pBluescript II) was performed on Day 0, but contrary to **Fig. 1c**, carbenicillin was not added to the drinking water to select for phage infection. Sm^R and $Carb^R$ CFU were enumerated up to 3 days post-gavage. Sm^R CFU indicative of total *E. coli* remained steady while $Carb^R$ CFU indicative of phage-infected cells were much lower in number, peaked between 6-12 h after gavage, and decreased over time. Dashed line indicates limit of detection. **(b)** The maximum observed percentage of $Carb^R$ CFU (phage-infected) over Sm^R CFU (total) was approximately 0.1%. For each of the six mice in panel a, the maximum percentage of $Carb^R/Sm^R$ was calculated, with values ranging over 3 orders of magnitude between ~0.0001% and ~0.1%. Time-series fecal samples from a mouse orally gavaged with heat-killed phage was also assayed as a negative control.



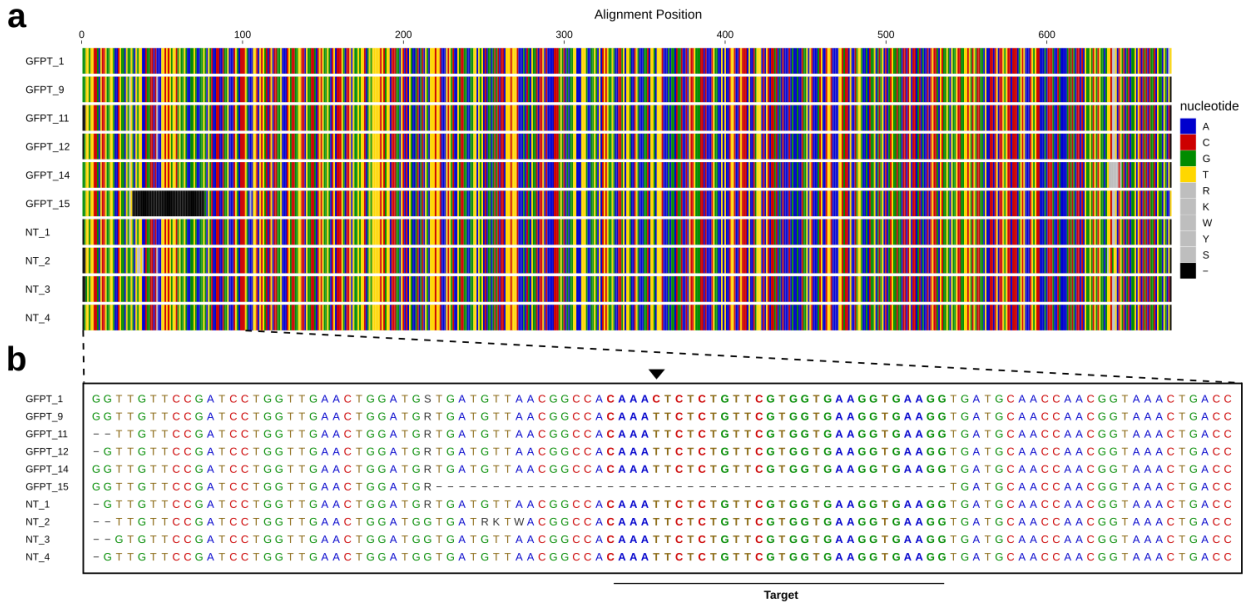
Supplementary Figure 5. Low recovery of M13 bacteriophage following GI transit despite high acid tolerance. (a) Enumeration of phage M13(pBluescript II) in feces of conventionally-raised mice after oral gavage. Mice were treated with 10^{13} M13(pBluescript II) (n=6) or as negative controls, heat-killed M13(pBluescript II) or PBS. Live phage in fecal samples were assayed using indicator strain XL1-Blue MRF' at t = 0, 3, 6, 9, and 24 h post-gavage. **(b)** Using the median live phage output in the feces at each timepoint for the six mice gavaged with live M13(pBluescript II), the area under-the-curve from 0 to 24 hours is 1×10^{6}. **(c)** Phage M13 displays resistance to acidic conditions as extreme as pH 2. 10^{9} M13(pBluescript II) were incubated in pH solutions 1.2 to 7, and sampled over the course of 60 min to assay for viability. Dashed line indicates limit of detection.



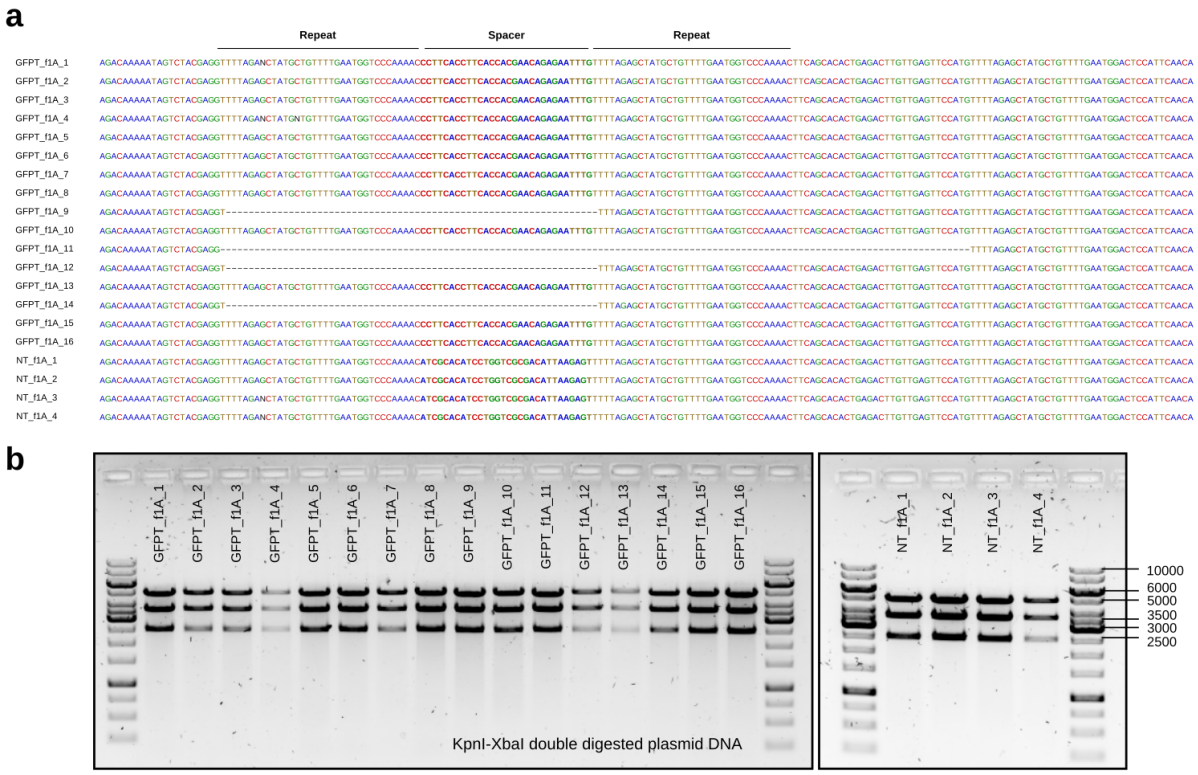
Supplementary Figure 6. GFP-targeting (GFPT) CRISPR-Cas9 phagemids. The non-targeting (NT) versions of these vectors (not shown here) are identical to the GFPT vectors except in the spacer sequence. The f1-*bla* fragment was cloned as a SalI fragment in both possible orientations for either strand of DNA to be packaged into M13 phage. **(a)** The first orientation is designated f1A. **(b)** The second orientation is designated f1B. *cat*, chloramphenicol acetyltransferase (Cm^R); *bla*, beta-lactamase ($Carb^R$).



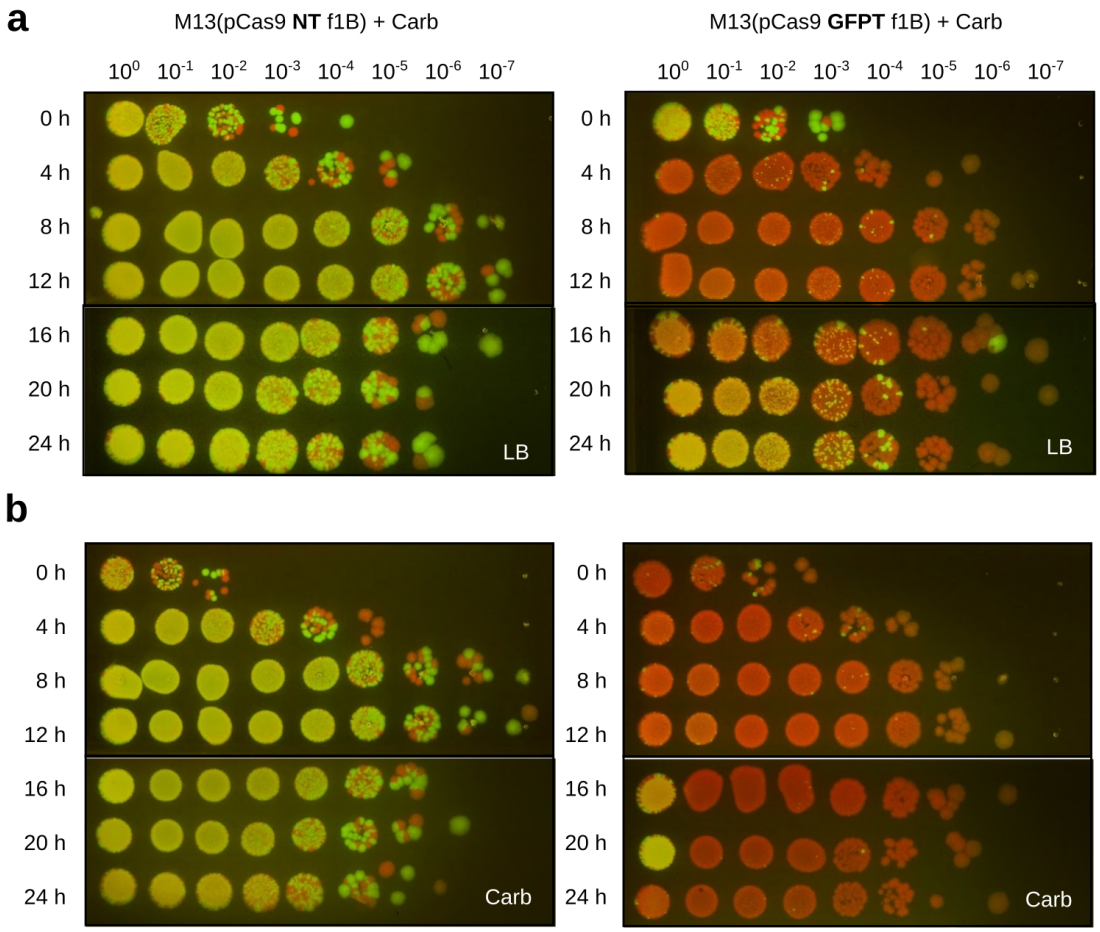
Supplementary Figure 7. GFP-marked *E. coli* exhibits impaired colony growth after infection with GFPT-M13 *in vitro*. NT-M13 or GFPT-M13 were used to infect Sm^R W1655 F+ *sfgfp* or Sm^R W1655 F+ *mcherry* (negative control). **(a)** Growth impairment of the GFP-marked strain under GFPT conditions was evident under blue light. **(b)** Impaired colonies exhibited a translucent quality that was more pronounced under white light.



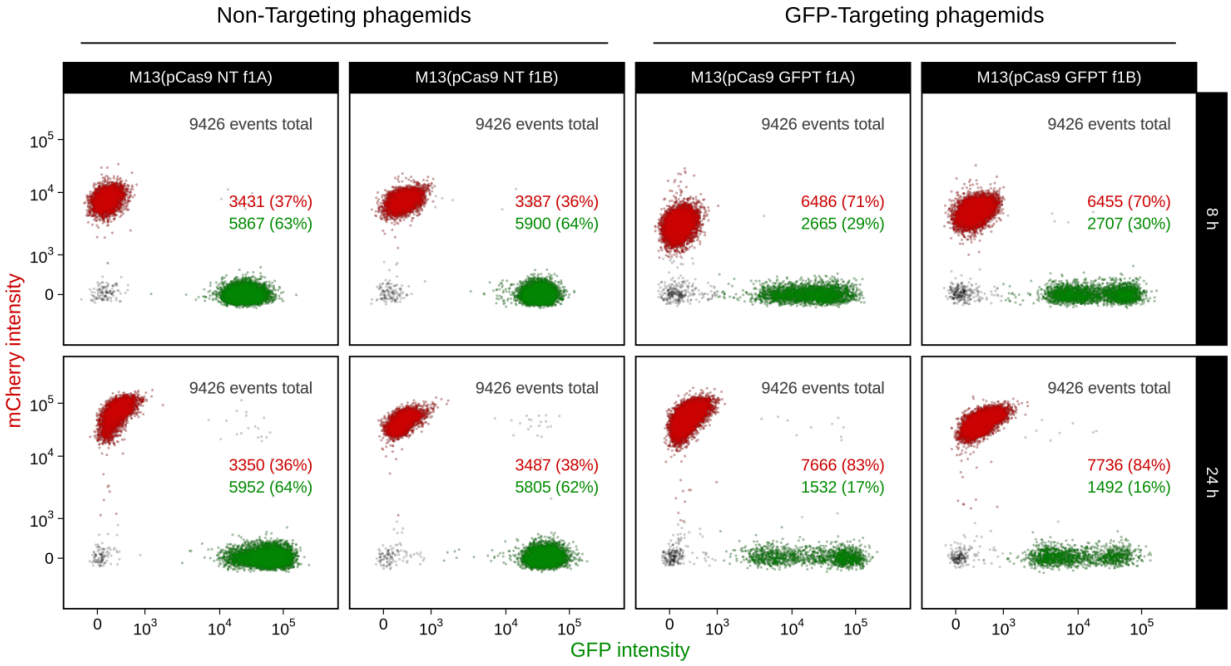
Supplementary Figure 8. Escape from GFP-targeting CRISPR-Cas9 by mutation in the target sequence. (a) Sanger sequencing of *sfgfp* PCR amplicons from streak-purified clones after treatment with GFPT-M13 or NT-M13 *in vitro* (**Fig. 2c**) confirmed the partial loss observed for clone GFPT 15 by gel electrophoresis. **(b)** Pullout showing the lost region of the *sfgfp* coding sequence from clone GFPT 15 encompasses the CRISPR-Cas9 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.



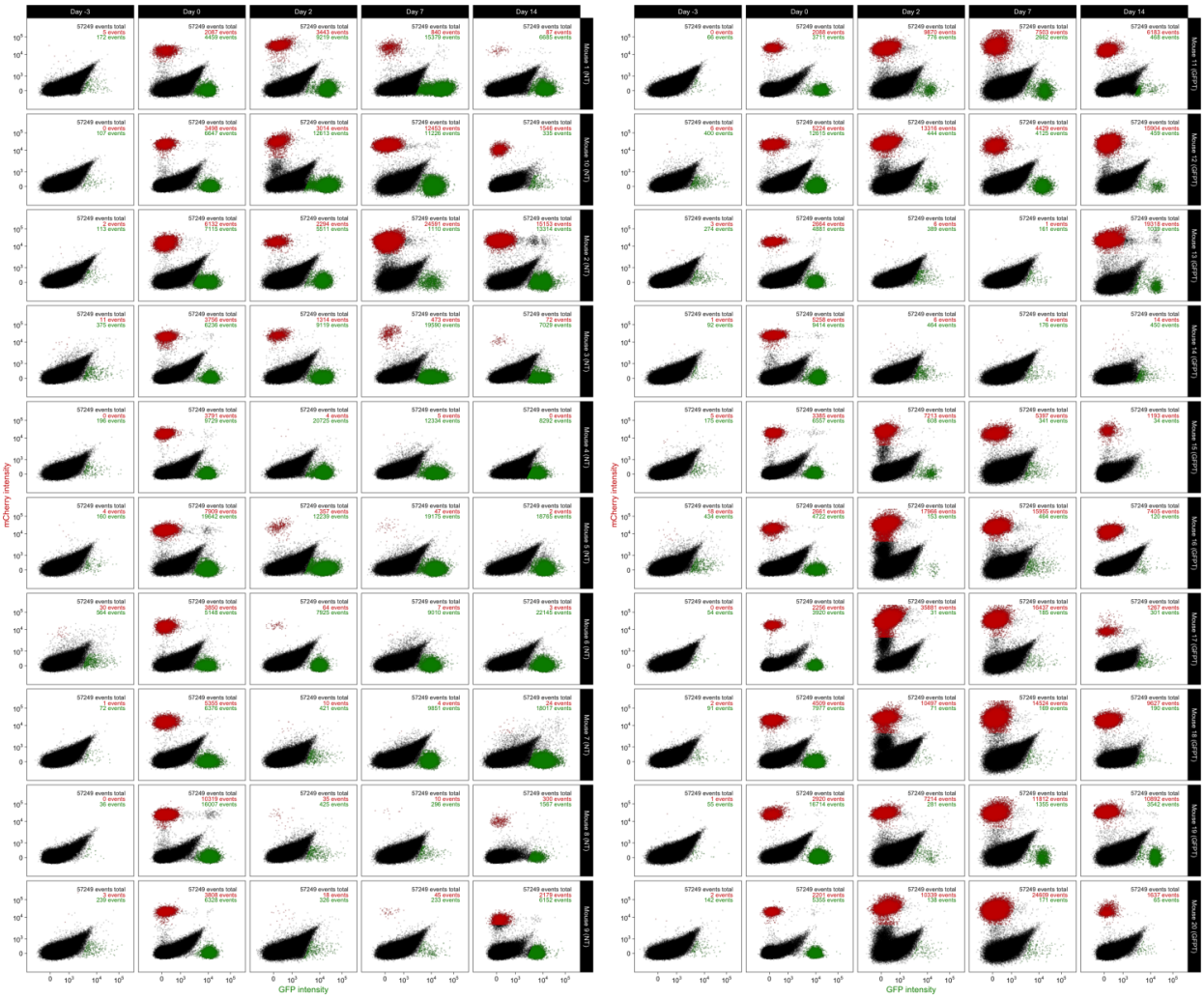
Supplementary Figure 9. Spacer loss from GFP-targeting CRISPR-Cas9 phagemids. (a) Sanger sequencing to check for spacer presence in phagemid DNA isolated from clones after treatment with GFPT-M13 or NT-M13 *in vitro*. All 4 clones isolated after infection with NT-M13 retained the spacer. Of 16 clones isolated after infection with GFPT-M13, 4 had lost the spacer (clones 9, 11, 12, and 14). **(b)** Diagnostic digest of plasmid DNA isolated from clones using KpnI and XbaI revealed phagemid of the expected size. Expected fragments: 4993, 3581, and 2573 bp.



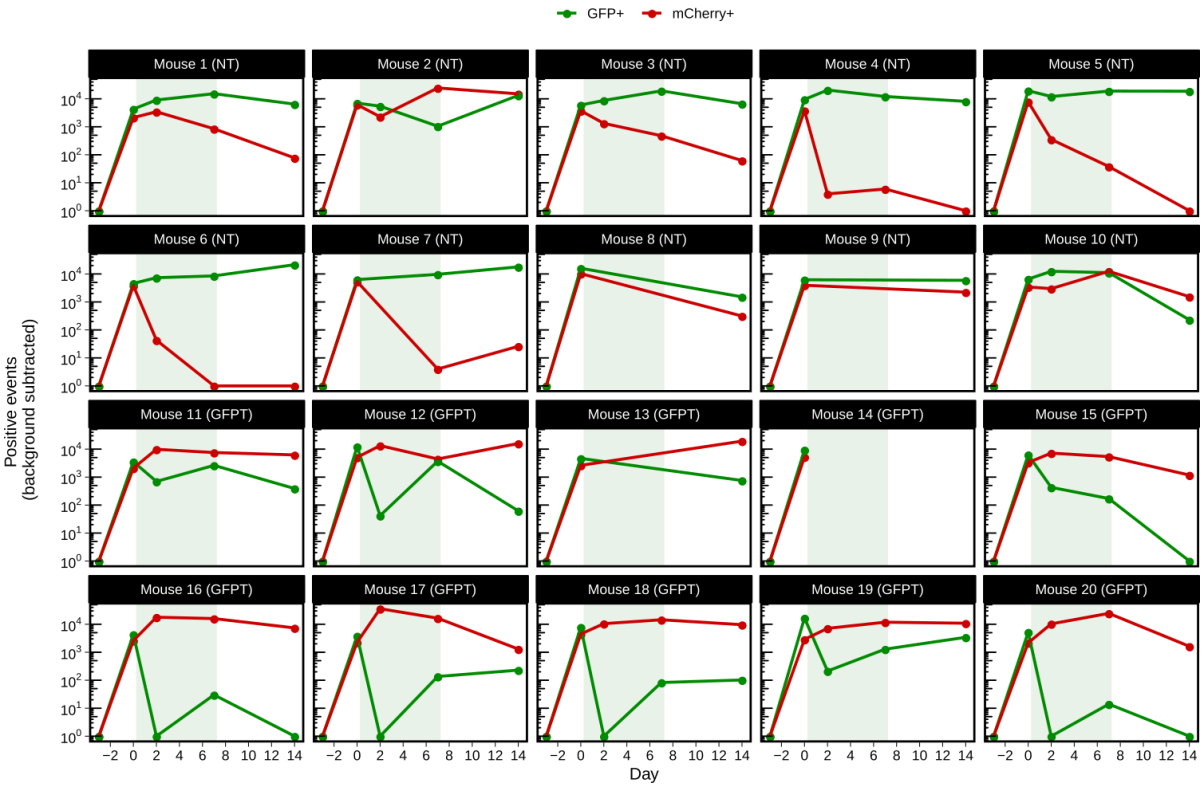
Supplementary Figure 10. Recovery of GFP+ cells at later timepoints from *in vitro* co-cultures after infection with GFPT-M13 is likely due to lack of selection for the CRISPR-Cas9 phagemid. (a) On non-selective media, GFP fluorescent colonies are detected at later timepoints of the co-culture infected with GFPT-M13. **(b)** Lack of GFP fluorescent colonies after testing the same co-culture on media with carbenicillin 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.



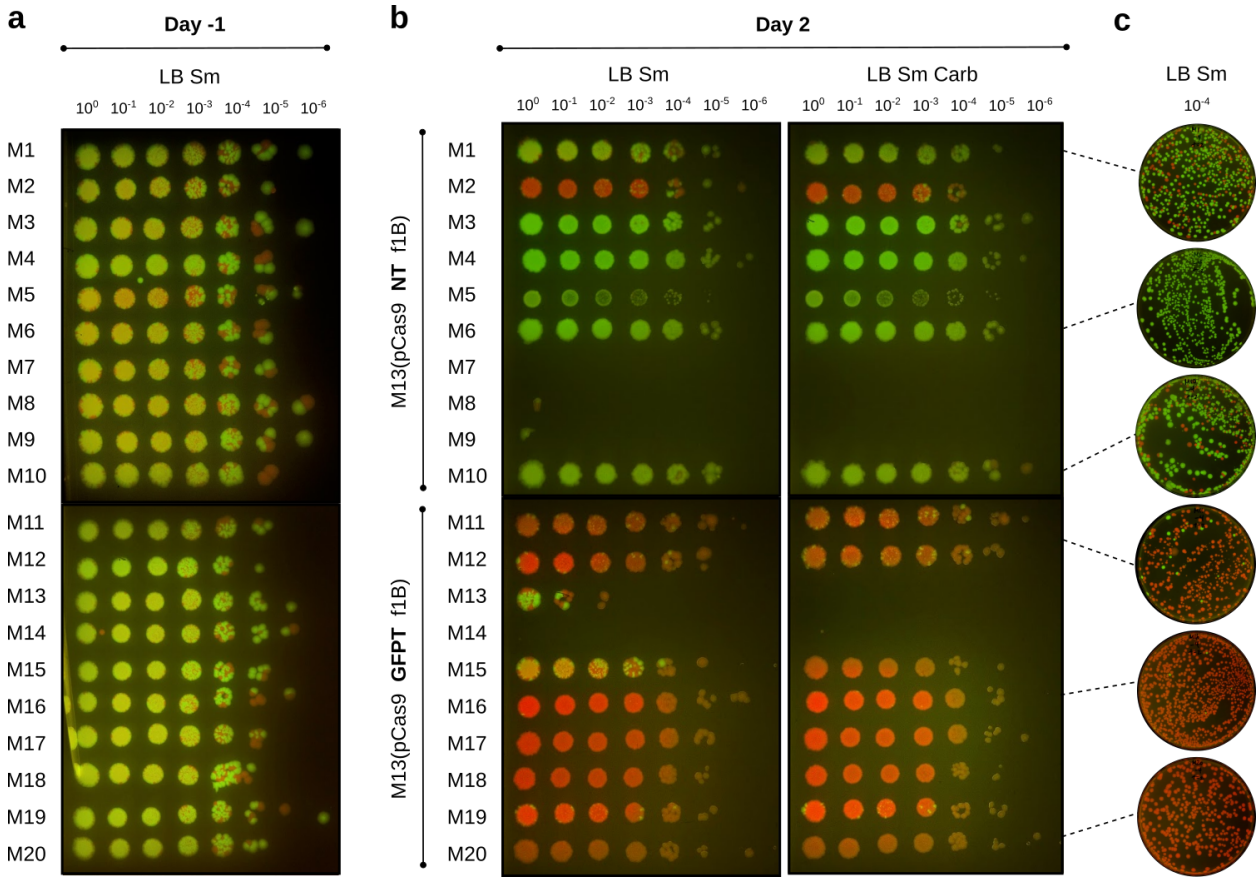
Supplementary Figure 11. The GFP+ strain shows a further decrease in relative abundance at 24 h after treatment with GFPT-M13 when compared to the 8 h timepoint. Co-cultures of GFP-marked and mCherry-marked *E. coli* F+ were infected with NT-M13 or GFPT-M13, carbenicillin was added to select for phage 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.



Supplementary Figure 12. Flow cytometry plots of fecal samples for all mice at all timepoints during in vivo targeting of GFP-marked *E. coli* in competition with mCherry-marked *E. coli*. Mice (n = 10 per group) were given either NT-M13 (left) or GFPT-M13 (right). 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.

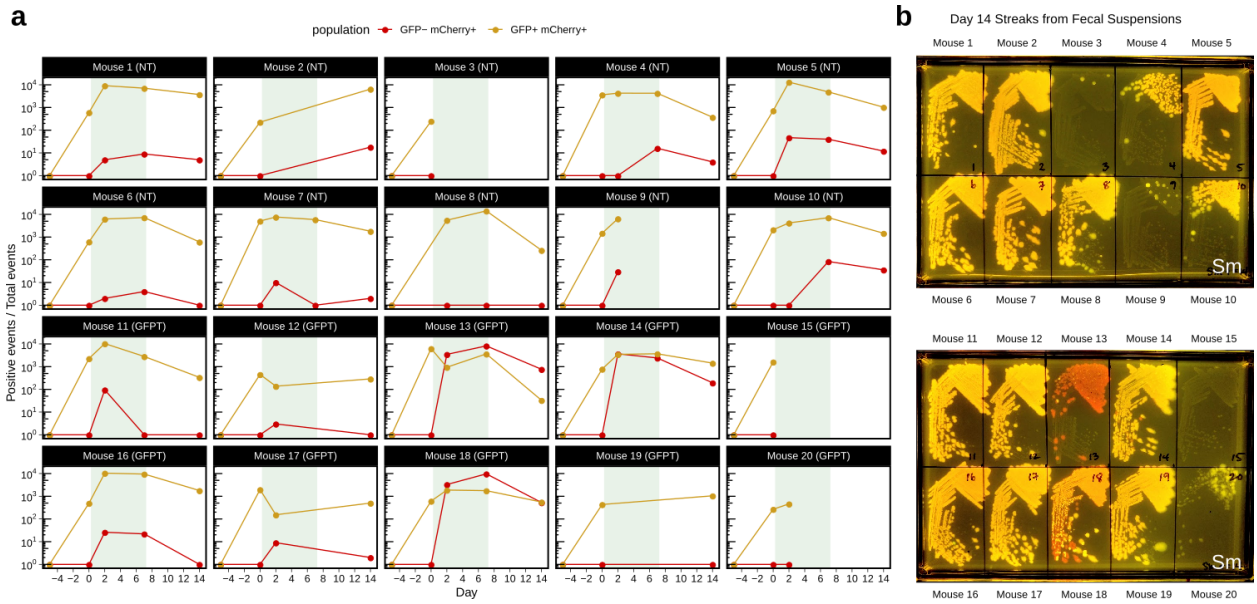


Supplementary Figure 13. GFP+ and mCherry+ events in fecal samples over time for individual mice during in vivo targeting of GFP-marked *E. coli* in competition with mCherry-marked *E. coli*. Mice were treated with either NT-M13 (M1 to M10) or GFPT-M13 (M11 to M20). For each mouse, the number of positive events by flow cytometry 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.

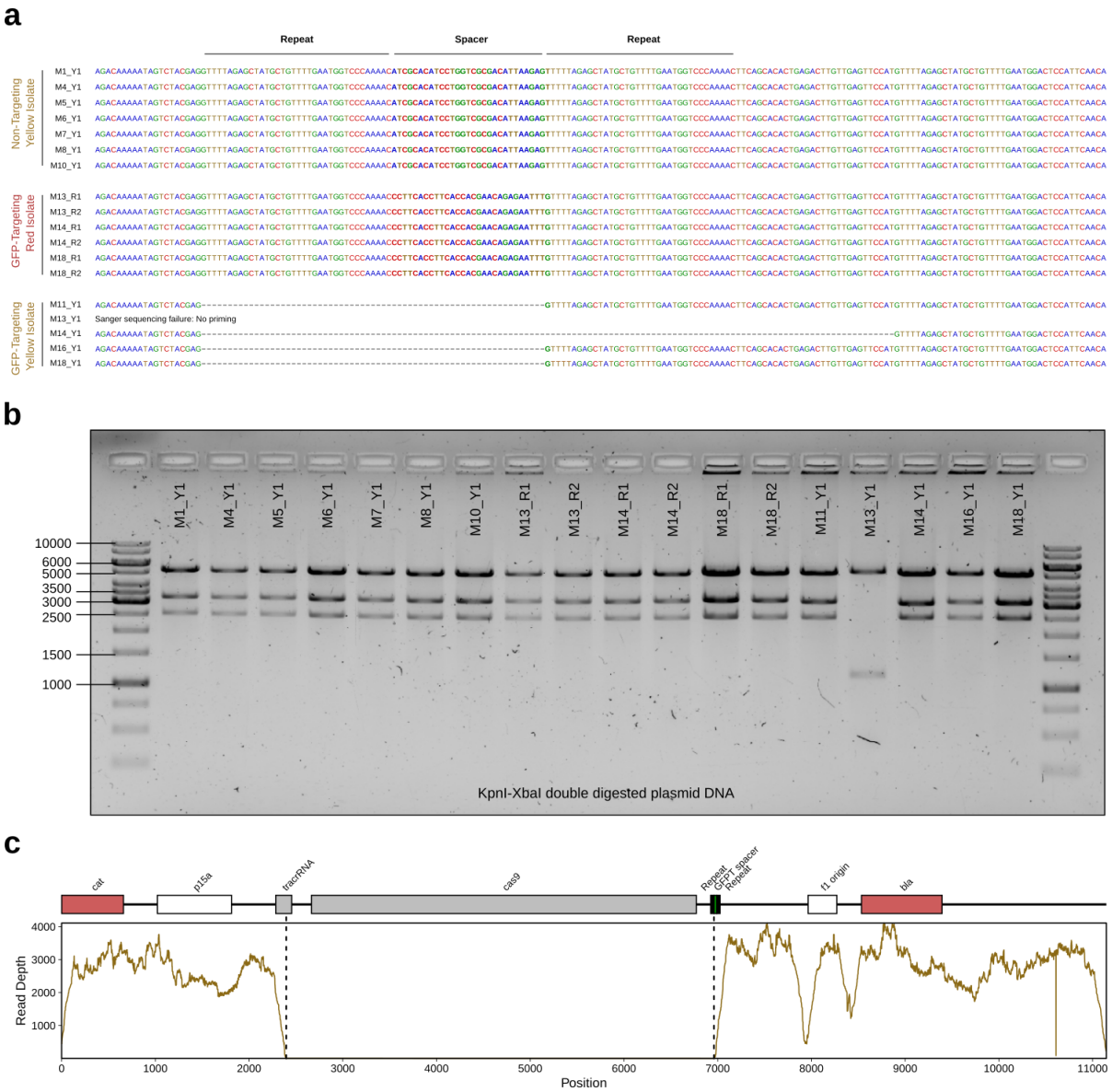


Supplementary Figure 14. Culturing from fecal samples of mice co-colonized with GFP-marked and mCherry-marked strains before and after treatment with phage confirms depletion of GFP+ strain under GFPT conditions on Day 2. (a) At Day -1, colonization by both the GFP+ and mCherry+ *E. coli* was confirmed in fecal samples of all mice by culturing on LB with streptomycin. **(b)** After treating with NT-M13 (M1 to M10) or GFPT-M13 (M11 to M20) and carbenicillin to select for phage infection, culture of *E. coli* on LB streptomycin (Sm) from fecal samples on Day 2 of GFPT mice exhibit decreased GFP fluorescence. Culturing from the same samples on LB with both streptomycin and carbenicillin (Carb) 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. **(c)** Day 2 fecal suspensions from a subset of the mice (M1, M6, M10 for NT; M11, M16, M20 for GFPT) were cultured on larger plates for confirmation.

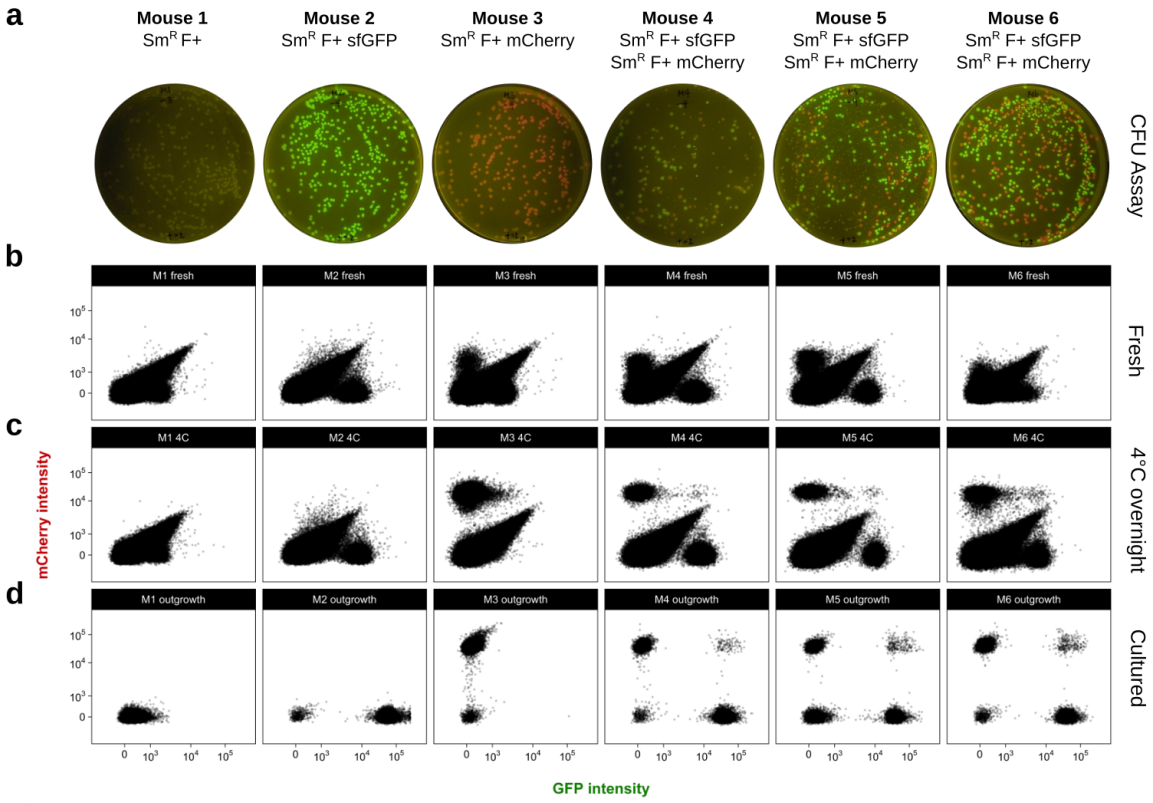
Supplementary Figure 15. Flow cytometry plots of fecal samples for all mice at all timepoints during *in vivo* targeting of double-marked *E. coli*. Mice (n=10/group) were given either NT-M13 (left) or GFPT-M13 (right). Day -5, before colonization by *E. coli*; Day 0, after colonization by double-marked 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.



Supplementary Figure 16. Fluorescence in fecal samples for individual mice colonized with double-marked GFP+ mCherry+ *E. coli* and dosed with NT-M13 or GFPT-M13. (a) 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. **(b)** Culture on LB with 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 during the treatment phase.



Supplementary Figure 17. Mechanisms enabling double-positive *E. coli* to escape CRISPR-Cas9 targeting. (a) Sanger sequencing results confirm the expected spacer present in phagemid DNA extracted from fluorescent yellow isolates (Y1) colonizing NT mice (M1, M4, M5, M6, M7, M8, M10) and fluorescent red isolates (R1 and R2) colonizing GFPT mice (M13, M14, M18). In contrast, 4/5 fluorescent yellow isolates colonizing GFPT mice (M11, M13, M14, M16, M18) were confirmed to have lost the spacer. No Sanger sequence data was obtained for the last isolate (M13) with report for failing being “No Priming”, suggesting loss of a larger fragment from the phagemid. **(b)** Diagnostic digest of CRISPR-Cas9 phagemid DNA confirms a loss of phagemid DNA for the phagemid extracted from M13 Y1. Expected fragment sizes from KpnI-XbaI double digest: 5289, 3285, and 2573 bp. **(c)** Genome sequencing data for M13_Y1 also confirms loss of DNA from phagemid. Sequence coverage across the GFPT phagemid reveals lack of reads corresponding to the *cas9* gene, and parts of the CRISPR array and tracrRNA.



Supplementary Figure 18. Quantification of fluorescent *E. coli* in mouse fecal pellets by flow cytometry improves with overnight incubation of fecal suspensions at 4°C. (a) Culture on LB streptomycin of fecal suspensions from streptomycin-treated mice colonized with non-fluorescent Sm^R *E. coli* W1655 F+ (Mouse 1), the GFP-marked (Mouse 2) or mCherry-marked derivative (Mouse 3), or both fluorescent strains (Mouse 4, 5, and 6). Flow cytometry was performed on fecal suspensions: **(b)** immediately after collecting, **(c)** overnight incubation at 4°C, or **(d)** after inoculation in media and overnight culture at 37°C.

Supplementary Tables

Supplementary Table 1. Strains, plasmids, and phage.

Supplementary Table 2. Oligonucleotides.