Plasmid vectors for *in vivo* selection-free use with the probiotic *E. coli* Nissle 1917

SUPPLEMENTARY INFORMATION

Anton Kan[1]†, Ilia Gelfat[1,2]†, Sivaram Emani[1,3], Pichet Praveschotinunt[1,2], Neel S. Joshi[1,2,4]*

† equal contribution
1 Wyss Institute for Biologically Inspired Engineering, Harvard University, Boston, MA, USA
2 John A. Paulson School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, USA
3 Harvard College, Harvard University, Boston, MA, USA
4 Department of Chemistry and Chemical Biology, Northeastern University, Boston, MA, USA
***Email:** ne.joshi@northeastern.edu

Primer Name	Sequence	Target
Muta 5	AACTGTGAAGCGATGAACCC	pMUT1
Muta 6	GGACTGTTCAGAGAGCTATC	pMUT1
Muta 7	GACCAAGCGATAACCGGATG	pMUT2
Muta 8	GTGAGATGATGGCCACGATT	pMUT2
pM1S2chk_F	GAATAGGGTGACACTGGCGCC	pMUT1
pM1S2chk_R	CCAGATGGCATTGTAACAGACTTCTC	pMUT1
pM1S3chk_F	CGCATCCTTCCTGTTTTTCCGG	pMUT1
pM1S3chk_R	GTCTTGGTAGCCCTGCTTCTGG	pMUT1
pM2S2chk_F	AGTTTCGCACCCAAAGTGCG	pMUT2
pM2S2chk_R	GACAAAACAACCTATATCAGATAACAGC	pMUT2
pM2S3chk_F	GATAAAACTATCAACTCACCGTCTTG	pMUT2
pM2S3chk_R	GCCGTTGGTCTTTACTGATTTAAG	pMUT2
M1_1F	aaaccttgcttcttcgcggtGAATAGGGTGACACTGGCG	pMUT1
M1_1R	accgcattctagatttagggATATGAATATACCATATAATATATACTTTAAATATTTTGG	pMUT1
M1_1F2	aaaccttgcttcttcgcggtGAATAGGGTGACACTGGCGCCATTATTGTG	pMUT1
M1_1R2	accgcattctagatttagggATATGAATATACCATATAATATATACTTTAAATATTTTGGGGCTTAG	pMUT1
M1_2F	aaaccttgcttcttcgcggtGGAGTTAGCGATATGAAAACCGAACAACG	pMUT1
M1_2R	accgcattctagatttagggGTATCTAATTCAGGCAGGAAAAAATCTTTTCC	pMUT1
M1_3F	aaaccttgcttcttcgcggtGTTTCAGTGGTGCGTACAATTAAG	pMUT1
M1_3R	accgcattctagatttagggGCGCTGAACGCGATTCTG	pMUT1
M2_1F	aaaccttgcttcttcgcggtCCACTAAGTTACACCTCAACAACG	pMUT2
M2_1R	accgcattctagatttagggCAGAAAAAACAAAGCCCCG	pMUT2
M2_1F2	aaaccttgcttcttcgcggtCCACTAAGTTACACCTCAACAACGG	pMUT2
M2_1R2	accgcattctagatttagggCAGAAAAAACAAAGCCCCGAAATCATGC	pMUT2
M2_2F	aaaccttgcttcttcgcggtCCACTAAGTTACACCTCAACAACGG	pMUT2
M2_2R	accgcattctagatttagggCAGAAAAAACAAAGCCCCGAAATCATGC	pMUT2
M2_3F	aaaccttgcttcttcgcggtATTAAATAATGACAATGTTGGGTTG	pMUT2
M2_3R	accgcattctagatttagggGTTTCTGCCTATAAGATTACTTACAGTG	pMUT2

SI table 1. Primers used to assess pMUTs in colony PCR, as well as the primers used to insert the recombinant cassettes onto the pMUT plasmids to make the engineered versions. In all cases, the capitalized sequences indicate homology to the pMUT target. Primer pairs in italic could not produce an amplicon with EcN DNA as a template.

Name	Sequence
UNS0	GTTCCTTATCATCTGGCGAATCGGACCCACAAGAGCACTG
UNS1	CATTACTCGCATCCATTCTCAGGCTGTCTCGTCTCGTCTC
UNS2	GCTGGGAGTTCGTAGACGGAAACAAACGCAGAATCCAAGC
UNS3	GCACTGAAGGTCCTCAATCGCACTGGAAACATCAAGGTCG
UNS4	CTGACCTCCTGCCAGCAATAGTAAGACAACACGCAAAGTC
UNS5	GAGCCAACTCCCTTTACAACCTCACTCAAGTCCGTTAGAG
UNS6	CTCGTTCGCTGCCACCTAAGAATACTCTACGGTCACATAC
UNSX	CCAGGATACATAGATTACCACAACTCCGAGCCCTTCCACC

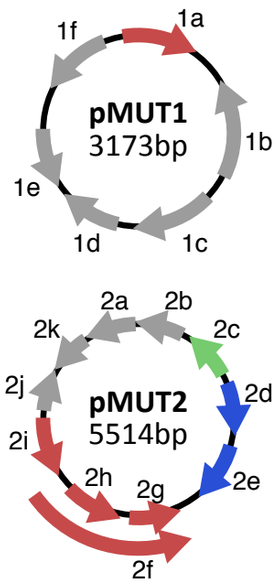
SI table 2. UNS sequences used in the engineered pMUT plasmid designs.

gRNA Name	Sequence	Target
gRNA1*	atgaactagcgattagtcgctatgacttaa	Targets pSC101 origin
gRNA2*	aaccacactagagaacatactggctaaata	Targets pSC101 origin
gRNA3*	ggttggactcaagacgatagttaccggata	Targets ColE1-like origins except colA
gRNA4*	ggcgaaacccgacaggactataaagatacc	Targets ColE1-like origins including colA
gRNA5	ccgatttgatggctatcgttcgggatcgtc	Targets pMUT2 at RelE
gRNA6	aactgcaccctcttcgataaaacccgcaag	Targets pMUT2 at hypothetical protein
gRNA7	gctctcttttcaggagagtgatttacccga	Targets pMUT2 at relaxase
gRNA8	ttgattttgtagcagtcatgcagctctcgc	Targets pMUT2 at RelB
gRNA9	cttgaatttgatccccgagccctgaaggaa	Targets pMUT2 at RelB
gRNA10	gcccacactcaccatcaaaaaccccgagaa	Targets pMUT2 at relaxase

SI table 3. gRNA sequences used in the pFREE and pCryptDel plasmids. gRNAs 1-4 (starred) were found in the original pFREE sequence from Lauritsen *et al.* [1]. In each case, the gRNA name, sequence and target are shown.

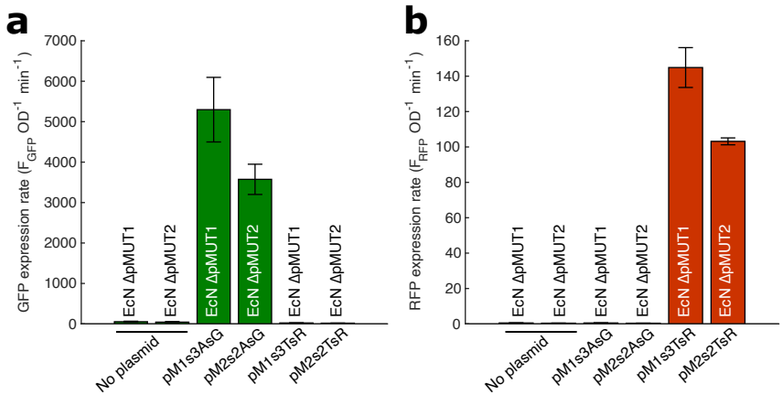
Plasmid	Promoter variant	Promoter sequence	Promoter strength (GFP)		CsgA-FE36 production (CR)		CsgA-NbGFP production (CR)		CsgA-FE36 production (ELISA)		CsgA-NbGFP production (ELISA)	
			37°C	30°C	37°C	30°C	37°C	30°C	37°C	30°C	37°C	30°C
No plasmid control			65.5	58.0	1.00	1.00	1.00	1.00	0.10	0.10	0.08	0.09
pMUT1	**A8**	tttaatttgtttgttagttagtttatttgttggtttgtttgtgttataGGGaagc	588.2	123.2	1.02	1.07	0.95	1.01	0.24	0.09	0.11	0.07
pMUT1	**A9**	tttaatttgtttgttagttagtttatttgttggtttgtttgtgttataTGCaagc	721.4	137.9	1.06	1.11	1.11	1.08	0.27	0.08	0.29	0.09
pMUT1	**C2**	tttaatttgtttgttagttagtttatttgttggtttgtttgtgttataGGCaagc	372.7	121.3	1.00	1.07	0.95	1.00	0.11	0.08	0.07	0.08
pMUT1	**C7**	tttaatttgtttgttagttagtttatttgttggtttgtttgtgttataCAAaagc	2402.1	179.3	1.44	1.12	0.95	1.12	0.49	0.10	0.11	0.10
pMUT1	**C8**	tttaatttgtttgttagttagtttatttgttggtttgtttgtgttataGCTaagc	1478.6	158.1	1.03	1.08	1.27	1.08	0.09	0.08	0.51	0.10
pMUT1	**D6**	tttaatttgtttgttagttagtttatttgttggtttgtttgtgttataCGAaagc	1355.9	152.6	1.32	1.17			0.47	0.08		
pMUT1	**D8**	tttaatttgtttgttagttagtttatttgttggtttgtttgtgttataGAAaagc	618.1	128.0	1.03	1.14	1.43	1.11	0.13	0.09	0.84	0.11
pMUT2	**E8**	tttaatttgtttgttagttagtttatttgttggtttgtttgtgttataGAGaagc	963.6	172.2	1.41	1.10	1.32	1.08	0.35	0.08	0.54	0.10
pMUT2	**E10**	tttaatttgtttgttagttagtttatttgttggtttgtttgtgttataAAAaagc	1016.5	196.7	1.30	1.08	1.31	1.09	0.68	0.08	0.52	0.12
pMUT2	**F7**	tttaatttgtttgttagttagtttatttgttggtttgtttgtgttataGCAaagc	718.1	64.4	1.25	1.03	1.00	1.00	0.19	0.09	0.29	0.09

SI table 4. Variant TlpA* promoter sequences, and their promoter strengths, CsgA-FE36 production capacity, and CsgA-NbGFP production capacity at when bacteria were grown on 37°C and at 30°C. The leftmost column indicates the host engineered plasmid. Promoter strength was measured using constructs pM#s#AsR_TS*. Curli material was produced with plasmids containing the TlpA* variant driving cassette 'csg-Etag' (columns CsgA-FE36), or 'csg-Etag-NbGFP' (columns CsgA-NbGFP), and curli production was measured using the Congo Red method (CR) or by immunostaining with anti-Etag on a filter plate (ELISA).

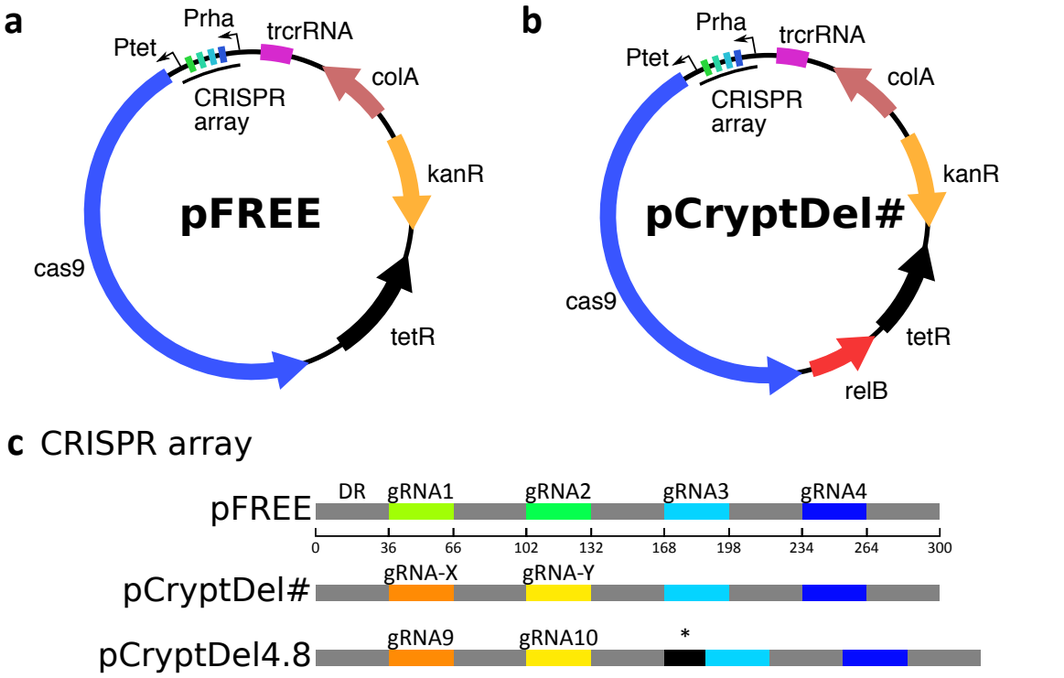


Predicted ORF	Gene name	Description	Notes
1a	*mobA*	Part of the relaxosome complex responsible for plasmid transfer during conjugation.	Part of plasmid conjugation machinery
1b	N/A	helix-turn-helix domain-containing protein	Putitative DNA-binding protein (MarR-like)
1c	N/A	Hypothetical protein	Multispecies, unknown function
1d	N/A	Hypothetical protein	Part of ColE1 plasmid replication origin, untranslated RNA
1e	N/A	Hypothetical protein	Part of ColE1 plasmid replication origin, untranslated RNA
1f	N/A	Rop family plasmid primer RNA-binding protein	Likely involved in plasmid copy number control
2a	N/A	Hypothetical protein	
2b	N/A	Hypothetical protein	
2c	*rep*	replicase	
2d	*RelB/Din*	Type II toxin-antitoxin system antitoxin	
2e	*RelE/StbE*	Addiction module toxin	
2f	*MobA*	Part of the relaxosome complex responsible for plasmid transfer during conjugation.	Part of plasmid conjugation machinery
2g	*MobD*	Plasmid mobilization protein	Part of plasmid conjugation machinery
2h	*MobB*	Essential to promote the specific transfer of the plasmid in the presence of conjugative plasmids.	Part of plasmid conjugation machinery
2i	*MobC*	Relaxosome protein	Part of plasmid conjugation machinery
2j	N/A	Hypothetical protein	
2k	N/A	DUF4868 domain-containing protein	Putitative bacteriocin?

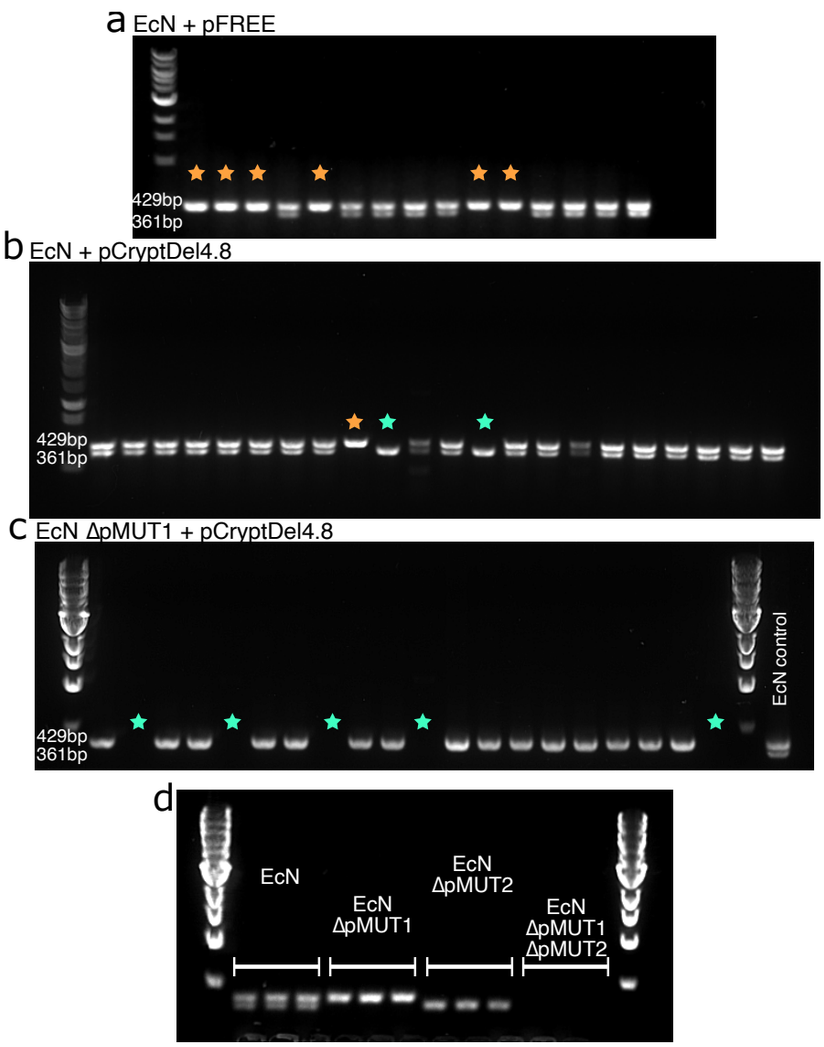
SI figure 1. Plasmid maps of pMUT1 and pMUT2 (left), alongside a table of annotations (right) for predicted ORFs in the native pMUT plasmids, with the 'Predicted ORF' referring to the labels on the plasmid maps



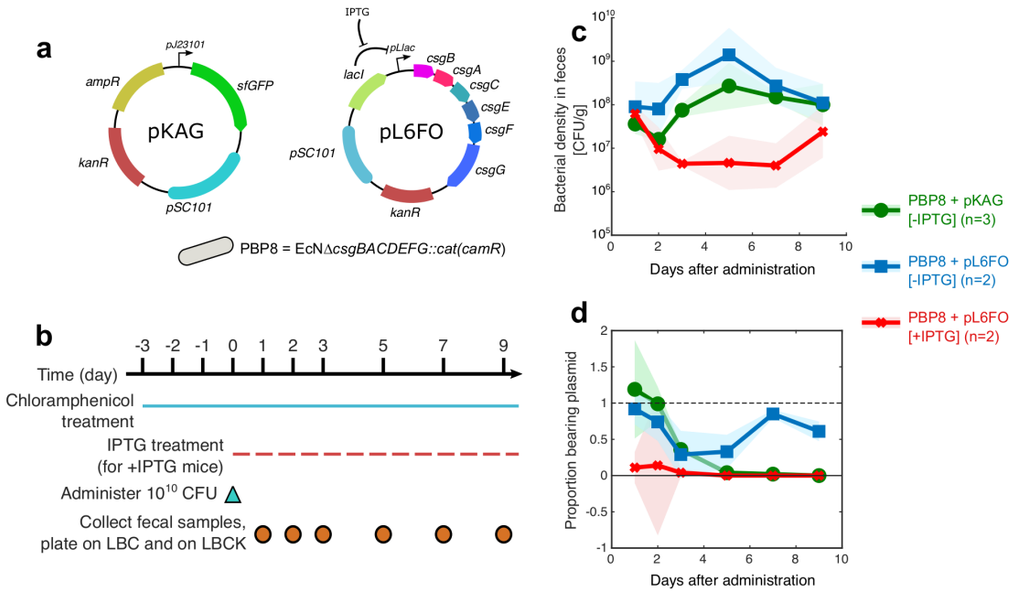
SI figure 2. Gene expression characterization from cassettes a) 'AsG' and b) 'TsR' on plasmids pM1s3 and pM2s2. In all cases, we used EcN without the relevant native pMUT plasmid. In both cases, pM1s3 plasmid backbone provides higher gene expression. The relative difference in recombinant protein expression strength between the engineered pMUT1 and pMUT2 plasmids is independent of the fluorescent protein used for characterization. Error bars show standard deviation from 8 replicates.



SI figure 3. Plasmid vectors to cure EcN cryptic plasmids. a) Plasmid map of pFREE, showing ATC inducible Cas9 and rhamnose inducible CRISPR gRNA array, b) plasmid map of the pCryptDel plasmid variants, which are based on pFREE with a modified CRISPR array and a *relB* anti-toxin gene. c) A detailed look at the gRNA arrays. gRNA-X and gRNA-Y refer to variants (shown in SI table 3), where X and Y pair were either 5 and 6, 7 and 8, or 9 and 10. The final construct, pCryptDel4.8, contained gRNA9 and 10, but had a 34bp insertion in the region just upstream of gRNA3.



SI figure 4. Representative TAE agarose gels of colony PCR results following a typical pMUT curing process, removing native plasmids from EcN with a) pFREE, b) pCryptDel4.8, and c) from EcN ΔpMUT1 with pCryptDel4.8. Panel d shows a gel for 3 replicate colonies each of unmodified EcN, and the pMUT knockout variants EcN ΔpMUT1, EcN ΔpMUT2, and EcN ΔpMUT1 ΔpMUT2. In all cases primers muta5, muta6, muta7 and muta8 were used, which result in a 429bp band in the presence of pMUT2, and a 361bp in the presence of pMUT1. In each case, an orange star shows a colony cured of pMUT1, and a blue star shows a colony cured of pMUT2.

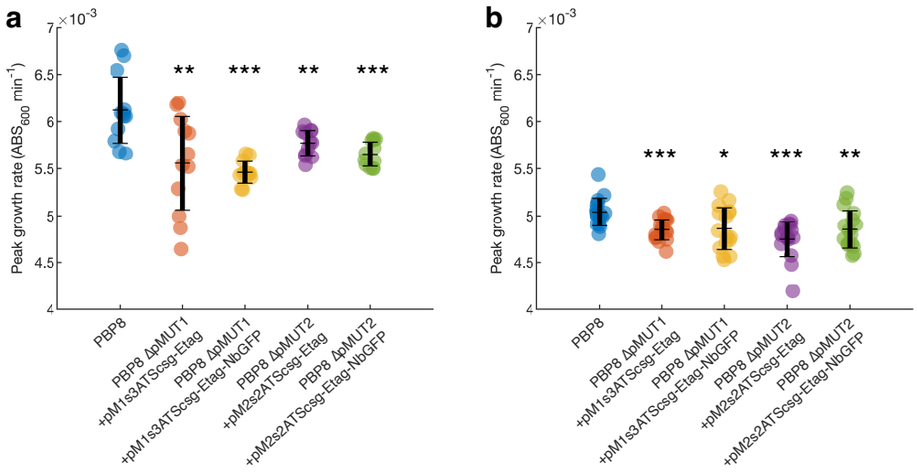


SI figure 5. a) Synthetic plasmids pKAG, which constitutively expresses sfGFP, and pL6FO, which expresses the synthetic curli operon *csgBACEFG* with IPTG induction, were transformed into *E. coli* Nissle strain PBP8. b) Administration and sampling schedule, with mice in all groups were treated with chloramphenicol to select for PBP8 cells from day -3 to the end of the experiment. On day 0, mice were administered with 10^{10} CFU of PBP8 transformed with either pKAG (n=3), or pL6FO (n=4), and half of the PBP8+pL6FO mice were given the IPTG inducer in their water. Fecal samples were collected regularly to detect PBP8 (chloramphenicol resistant) or PBP8 with plasmid (chloramphenicol and kanamycin resistant) by plating assays. c) After administration, PBP8 cells were maintained in the mice for all conditions throughout the experiment, but d) all plasmids suffered significant plasmid loss, particularly after day 2. Shaded areas show relative standard error for panel c and standard deviation for panel d. Detailed assay methodology can be found in supplementary methods below.

Supplementary methods: Synthetic plasmid retention in the mouse gut

7 female 8- to 9-week-old C57BL/6NCrl mice, obtained from Charles River Laboratories, were randomly split into 3 experimental cohorts: PBP8+pKAG, PBP8+pL6FO [-IPTG] and PBP8+pL6FO [+IPTG]. Housing and feed were the same as described in the main methods section. 3 days prior to initial administration of bacteria, the drinking water was supplemented with 0.5 g/L chloramphenicol, and all subsequent water contained chloramphenicol. For the +IPTG cohort, the water was supplemented with a further 10mM IPTG from day 0 onwards. Bacterial suspensions were prepared in advance by growing to mid-exponential phase (OD600 of 0.5) at 37°C (shaking at 225 RPM), pelleting the cells, resuspending to OD600 of 10 in PBS supplemented with 20% sucrose and 10% glycerol. All mice were gavaged with 10^{10} CFU of the relevant strain on day 0 of the experiment.

Fecal pellets were collected and weighed on day 1, 2, 3, 5, 7, and 9. Immediately following daily collection of fecal pellets, each sample was homogenized in 1 mL of PBS, serially diluted, and plated in quadruplicate to enumerate colony forming units (CFU). Samples were plated on two types of LB agar plates - 25 µg/mL chloramphenicol-only plates (LBC) and 25 µg/mL chloramphenicol + 50 µg/mL kanamycin plates (LBCK). While all PBP8-derived strains carried a chromosomal *camR* resistance gene, only the synthetic plasmid bearing population harbored the *kanR* resistance gene. Total PBP8 bacterial density was found by counting colonies on LBC plates and normalizing by the weight of fecal matter sampled. Plasmid retention rate was estimated by calculating the cell density from LBCK plates and dividing by the density from LBC plates.



SI figure 6. *In vitro* growth rates of bacterial strains used in the *in vivo* experiments grown at a) 37°C and b) 30°C. Growth rates were measured by measuring the absorbance at 600 nm of the bacterial cultures in a plate reader every 10 minutes for 16 hours, the fitting this curve to a Gompertz model and extracting the peak growth rate from the model. In all cases the bacteria were grown in LB media, with carbenicillin added when the engineered plasmids were present. When compared to PBP8, the bacteria with harbouring engineered plasmids grew significantly slower, and this difference was most pronounced in the 37°C condition, where the temperature sensitive promoter would be active and expressing the modified curli material. The error bars show the standard deviation of the samples (n>=12), and each plasmid bearing sample was compared to the PBP8 control with a two sample t-test assuming unequal variances, * p<0.05, ** p<0.01, *** p<0.001.