Negative selection and stringency modulation in phage-assisted continuous evolution

*Jacob C. Carlson[1], Ahmed H. Badran[1], Drago A. Guggiana-Nilo[2], and David R. Liu[1]**

[1]Department of Chemistry & Chemical Biology and Howard Hughes Medical Institute, Harvard University, 12 Oxford St, Cambridge, MA 02138 USA

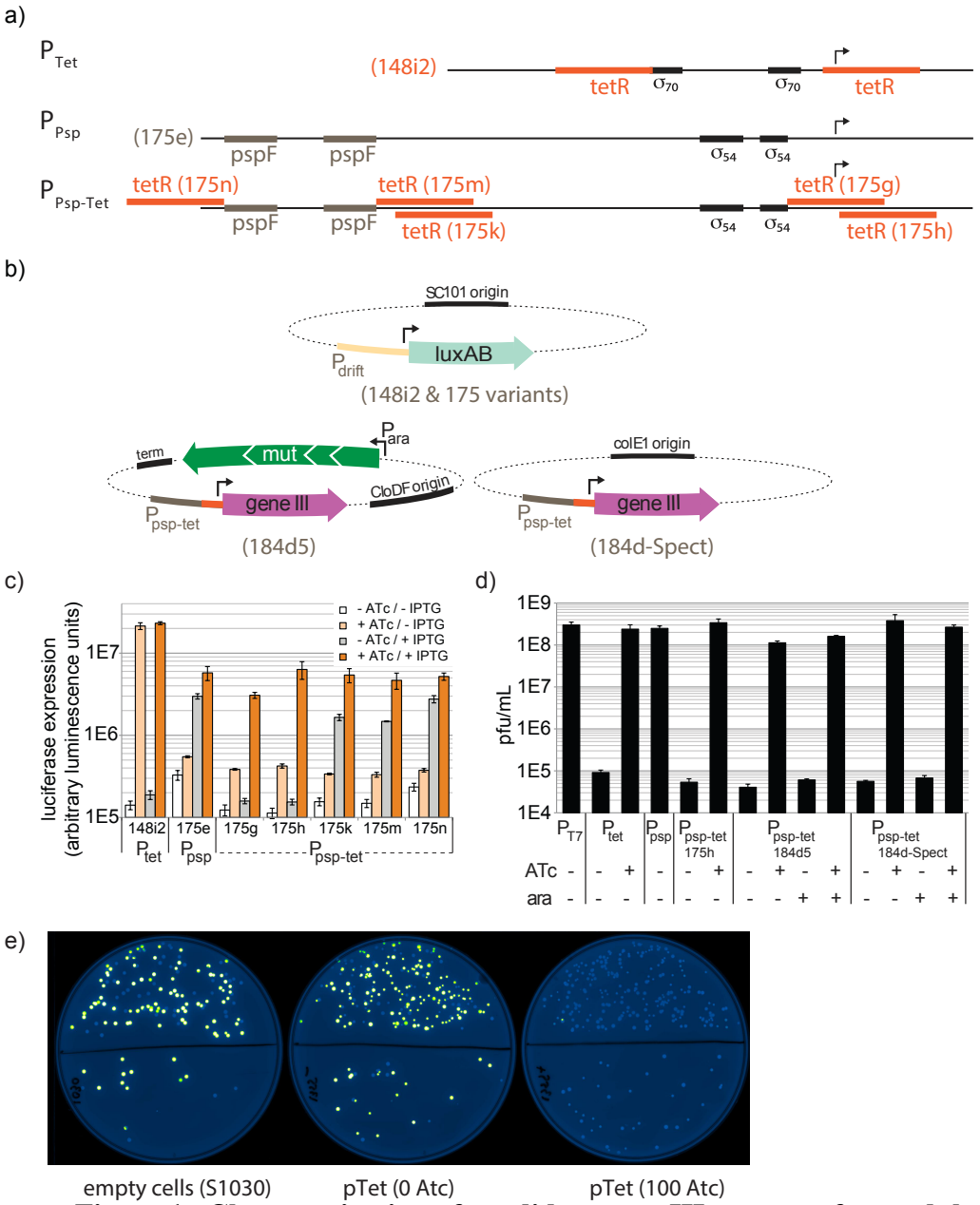
[2]Harvard Medical School, Building C-2, Room 122, 240 Longwood Avenue, Boston, MA 02115

Correspondence: David R. Liu (drliu@chemistry.harvard.edu)

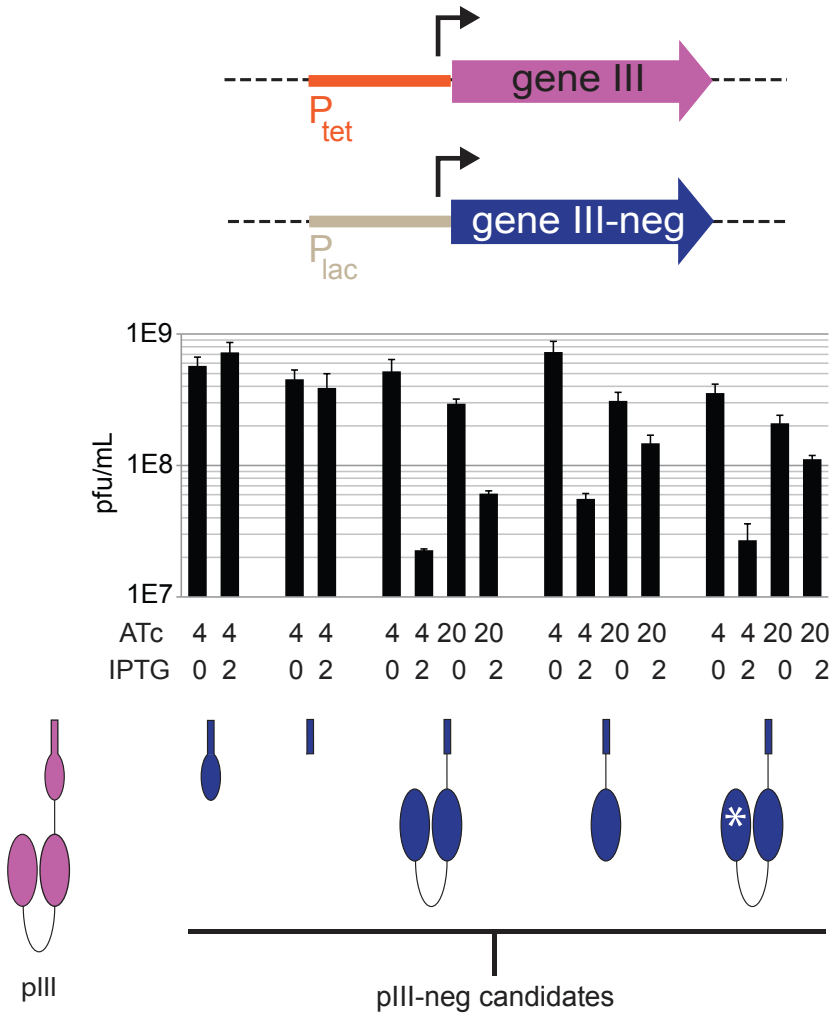
Supplementary Results

Supplementary Figure S1	Characterization of candidate drift cassettes
Supplementary Figure S2	Effect of pIII-neg candidate expression on infectious phage titer
Supplementary Figure S3	Dose-dependent effect of pIII-neg expression on phage production
Supplementary Figure S4	Genotypes of continuously evolved RNAP variants
Supplementary Figure S5	Sequence alignment of pIII-neg candidates and wild-type pIII
Supplementary Table S1	List of all plasmids used in this work
Supplementary References	

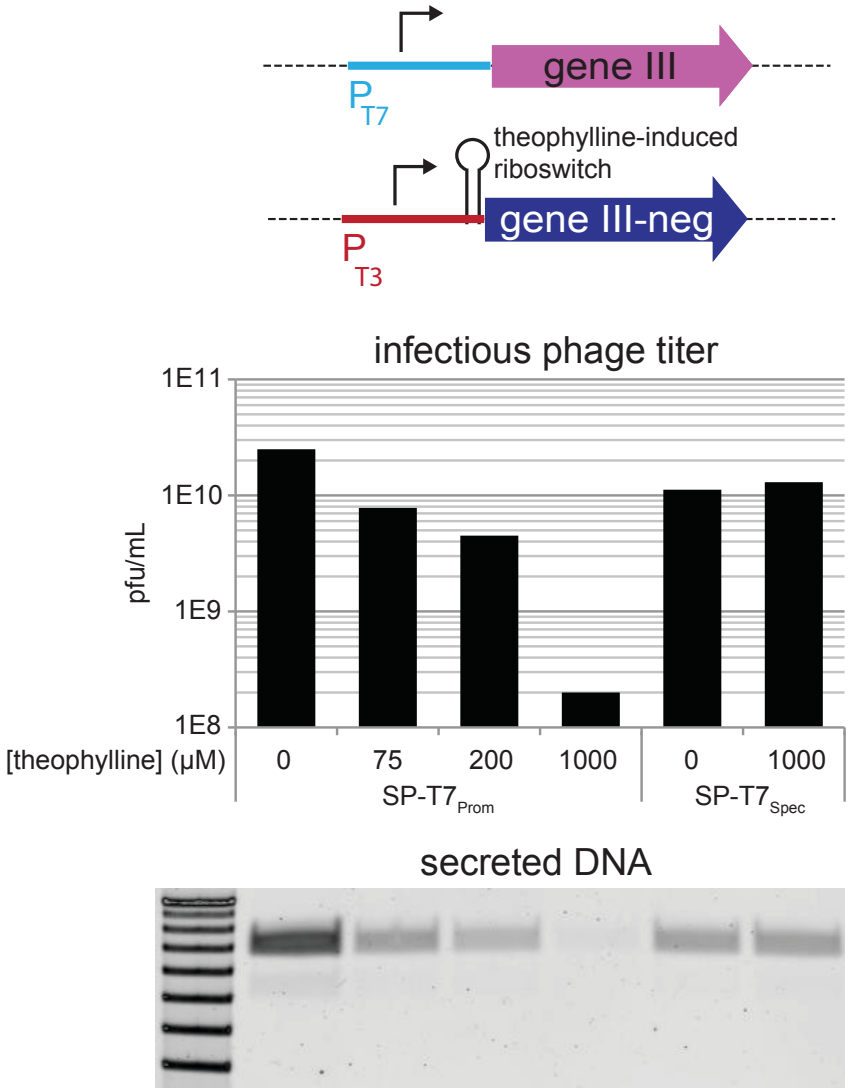
Nature Chemical Biology: doi:10.1038/nchembio.1453



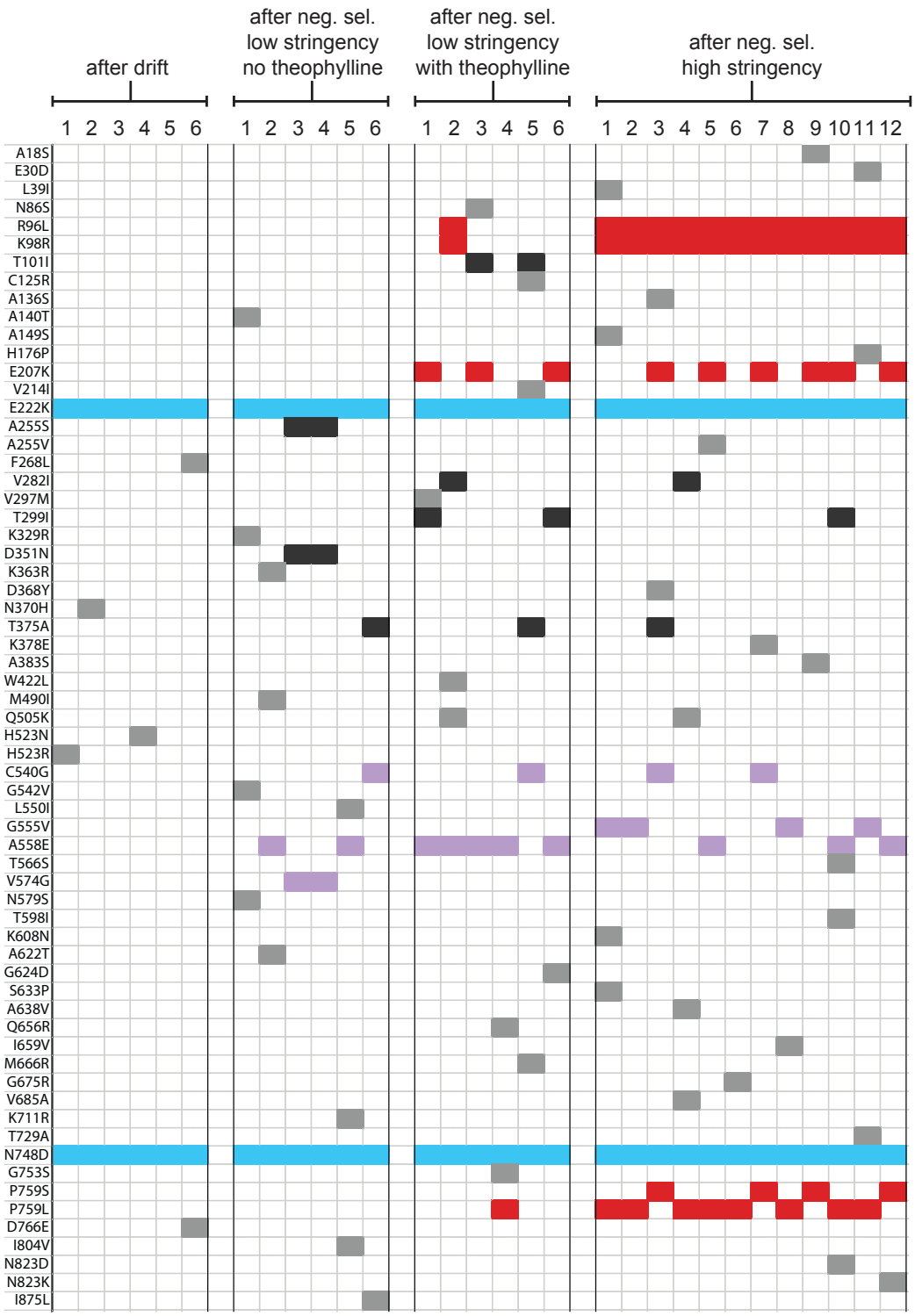
Supplementary Figure 1. Characterization of candidate gene III cassettes for modulating stringency and enabling drift. (a) Architecture of P_{tet} (top; 148i2), P_{psp} (middle; 175e), and $P_{psp\text{-}tet}$ candidates (bottom; 175g, h, k, m, and n variants). For the $P_{psp\text{-}tet}$ candidates, the numbers shown in parentheses refer to the constructs that contain the tetR-binding sites at the positions shown by the orange bar. tetR: tetR-binding sites, pspF: pspF-binding sites, σ70: *E. coli* sigma 70 RNAP promoter, σ54: *E. coli* sigma 54 RNAP promoter. (b) Architecture of plasmids highlighted in this figure. (c) Luciferase gene expression measurements of candidate drift cassettes in the presence and absence of ATc (which de-represses TetR) and IPTG (which de-represses lacI and induces pIV expression). pIV expression is used for these experiments to emulate filamentous phage infection and drive the psp response. (d) Phage production from discrete cultures containing various drift constructs and SP-$T7_{WT}$ in the presence and absence of ATc (to induce gene III expression) and arabinose (to induce mutagenesis). Robust phage titers are produced regardless of drift cassette copy number or mutagenesis induction. (e) Assay for infection of recipient cells carrying the pTet-pIII cassette. A low proportion of yellow colonies indicates resistance to infection, as seen for recipient cells carrying the induced pTet construct. Data represent mean values ± s.e. for two replicates.



Supplementary Figure 2. Effect of pIII-neg candidate expression on infectious phage titer. Gene III expression is driven by P_{tet} (induced by ATc) and the expression of a gene III-neg candidate is driven by pLac (induced by IPTG). Constructs from left to right are pJC156a2 (C-domain), pJC156c2 (C83), pJC156j2 (N-C83), pJC156m2 (N2-C83), pJC156o2 (N*-C83). Amino acid sequences of the pIII-neg candidates are provided in Supplementary Figure S5. ATc was used at 4 or 20 ng/mL and IPTG was used at 0 or 2 mM. Data represent mean values ± s.e. for two replicates.



Supplementary Figure 3. Dose-dependent effect of pIII-neg expression on phage production. An SP encoding a promiscuous T7 RNAP variant (SP-T7$_{Prom}$) or an SP encoding a P_{T7}-specific RNAP variant (SP-T7$_{Spec}$) were used to infect cells harboring an AP in which P_{T7} drives gene III expression and an AP-neg in which P_{T3} drives gene III-neg expression in a theophylline-dependent manner. Cells were infected using excess phage for 10 min at 37 °C, centrifuged and washed to remove residual phage, and resuspended in fresh Davis rich media. Infected cells were grown in the presence of the indicated concentrations of theophylline. The resulting titers of progeny phage after reaching mid-log phase are shown in the graph. SP-T7$_{Prom}$ results in fewer progeny phage only when theophylline is added. The total secreted DNA was also measured (bottom) and correlated with phage production. Data are from representative single (n =1) measurements of phage concentration.



Supplementary Figure 4. Genotypes of continuously evolved RNAP variants. Mutations present in clones isolated following the drift stage (left, t = 28 hrs), low-stringency negative selection without theophylline (center left, t = 32 hrs), low-stringency negative selection with theophylline (center right, t = 52 hrs), and high-stringency negative selection (right, t = 70.5 hrs). Mutations in blue are conserved in all clones and were found to confer activity on P_{T3}. Mutations in red are conserved in clones with high specificity for P_{T3} over P_{T7}. Mutations in magenta are, as a group, conserved and predicted based on the structure of T7 RNAP[1] to be physically clustered, but mutually exclusive of each other (negatively epistatic). Mutations in grey or black are isolated or modestly conserved, respectively.

```
wt pIII         MKKLLFAIPLVVPFYSHSAETVESCLAKPHTENSFTNVWKDDKTLDRYANYEGCLWNATG 60
N-C83           MKKLLFAIPLVVPFYSHSAETVESCLAKPHTENSFTNVWKDDKTLDRYANYEGCLWNATG 60
N*-C83          MKKLLFAIPLVVPFYSHSAETVESCLAKPHTENSFTNVWKDDKTLDRYANYEGCLWNATG 60
N2-C83          MKKLLFAIPLVVPFYSHSAETVG------------------------------------ 23
C-domain        MKKLLFAIPLVVPFYSHSAETVD------------------------------------ 23
C83             MKKLLFAIPLVVPFYSHSAETVS------------------------------------ 23
                ********************

wt pIII         VVVCTGDETQCYGTWVPIGLAIPENEGGGSEGGGSEGGGSEGGGTKPPEYGDTPIPGYTY 120
N-C83           VVVCTGDETQCYGTWVPIGLAIPENEGGGSEGGGSEGGGSEGGGTKPPEYGDTPIPGYTY 120
N*-C83          VVVCTGDETQCW---VPIGLAIPENEGGGSEGGGSEGGGSEGGGTKPPEYGDTPIPGYTY 117
N2-C83          ------------------------------------GGSEGGGTKPPEYGDTPIPGYTY 46
C-domain        ------------------------------------------------------------
C83             ------------------------------------------------------------

wt pIII         INPLDGTYPPGTEQNPANPNPSLEESQPLNTFMFQNNRFRNRQGALTVYTGTVTQGTDPV 180
N-C83           INPLDGTYPPGTEQNPANPNPSLEESQPLNTFMFQNNRFRNRQGALTVYTGTVTQGTDPV 180
N*-C83          INPLDGTYPPGTEQNPANPNPSLEESQPLNTFMFQNNRFRNRQGALTVYTGTVTQGTDPV 177
N2-C83          INPLDGTYPPGTEQNPANPNPSLEESQPLNTFMFQNNRFRNRQGALTVYTGTVTQGTDPV 106
C-domain        ------------------------------------------------------------
C83             ------------------------------------------------------------

wt pIII         KTYYQYTPVSSKAMYDAYWNGKFRDCAFHSGFNEDPFVCEYQGQSSDLPQPPVNAGGGSG 240
N-C83           KTYYQYTPVSSKAMYDAYWNGKFRDCAFHSGFNEDPFVCEYQGQSSDLPQPPVNAGGGSG 240
N*-C83          KTYYQYTPVSSKAMYDAYWNGKFRDCAFHSGFNEDPFVCEYQGQSSDLPQPPVNAGGGSG 237
N2-C83          KTYYQYTPVSSKAMYDAYWNGKFRDCAFHSGFNEDPFVCEYQGQSSDLPQPPVNAGGGSG 166
C-domain        ------------------------------------------------------------
C83             ------------------------------------------------------------

wt pIII         GGSGGGSEGGGSEGGGSEGGGSEGGGSGGGSGSGDFDYEKMANANKGAMTENADENALQS 300
N-C83           GGSGGGSEGGGSEGGGSEGGGSEGGGSGGGS---------------------------- 271
N*-C83          GGSGGGSEGGGSEGGGSEGGGSEGGGSGGGS---------------------------- 268
N2-C83          GGSGGGSEGGGSEGGGSEGGGSEGGGSGGGS---------------------------- 197
C-domain        ------------------------------------FDYEKMANANKGAMTENADENALQS 48
C83             ------------------------------------------------------------

wt pIII         DAKGKLDSVATDYGAAIDGFIGDVSGLANGNGATGDFAGSNSQMAQVGDGDNSPLMNNFR 360
N-C83           ----------------------------------------SQMAQVGDGDNSPLMNNFR 290
N*-C83          ----------------------------------------SQMAQVGDGDNSPLMNNFR 287
N2-C83          ----------------------------------------SQMAQVGDGDNSPLMNNFR 216
C-domain        DAKGKLDSVATDYGAAIDGFIGDVSGLANGNGATGDFAGSNSQMAQVGDGDNSPLMNNFR 108
C83             ----------------------------------------QMAQVGDGDNSPLMNNFR 41
                                                        ******************

wt pIII         QYLPSLPQSVECRPFVFGAGKPYEFSIDCDKINLFRGVFAFLLYVATFMYVFSTFANILR 420
N-C83           QYLPSLPQSVECRPFVFGAGKPYEFSIDCDKINLFRGVFAFLLYVATFMYVFSTFANILR 350
N*-C83          QYLPSLPQSVECRPFVFGAGKPYEFSIDCDKINLFRGVFAFLLYVATFMYVFSTFANILR 347
N2-C83          QYLPSLPQSVECRPFVFGAGKPYEFSIDCDKINLFRGVFAFLLYVATFMYVFSTFANILR 276
C-domain        QYLPSLPQSVECRPFVFGAGKPYEFSIDCDKINLFRGVFAFLLYVATFMYVFSTFANILR 168
C83             QYLPSLPQSVECRPFVFGAGKPYEFSIDCDKINLFRGVFAFLLYVATFMYVFSTFANILR 101
                ************************************************************

wt pIII         NKES 424
N-C83           NKES 354
N*-C83          NKES 351
N2-C83          NKES 280
C-domain        NKES 172
C83             NKES 105
                ****
```

Supplementary Figure 5. Sequence alignment of pIII-neg candidates and wild-type pIII. ClustalW2 (http://www.ebi.ac.uk/Tools/msa/clustalw2/) was used to align all pIII-neg candidates against full-length pIII.

Name	Class	Antibiotic Resistance	Origin of Replication	Promoter	Gene
pJC148i2	AP	Carb	SC101	P_{tet}	gIII-luxAB
pJC156a2	AP-neg	Spect	colEI	P_{lac}	C-domain-Venus
pJC156c2	AP-neg	Spect	colEI	P_{lac}	C83-Venus
pJC156j2	AP-neg	Spect	colEI	P_{lac}	N1-N2-C83-Venus
pJC156m2	AP-neg	Spect	colEI	P_{lac}	N2-C83-Venus
pJC156o2	AP-neg	Spect	colEI	P_{lac}	N1*-N2-C83-Venus
pJC173b	AP	Carb	SC101	P_{T7}	gIII-luxAB
pJC173c	AP	Carb	SC101	P_{T7}	6xHisTag-gIII-luxAB
pJC173e4	AP	Carb	SC101	P_{T7}	luxAB
pJC173f-R5	AP	Spect	colEI	P_{T7}	TheoRibo-6xHisTag-gIII-Venus
pJC173g-SD8	AP-neg	Spect	pUC	P_{T7}	N1*-N2-C83-Venus (strong RBS)
pJC174c-R5	AP	Spect	colEI	P_{T3}	TheoRibo-6xHisTag-gIII-Venus
pJC174e4	AP	Carb	SC101	P_{T3}	luxAB
pJC174f	AP	Carb	SC101	P_{T3}	6xHisTag-gIII-luxAB
pJC174k	AP	Carb	SC101	P_{T3}	6xHisTag-gIII-luxAB (weak RBS)
pJC175e	AP	Carb	SC101	P_{psp}	gIII-luxAB
pJC175g	AP	Carb	SC101	$P_{psp\text{-}tetO}$	gIII-luxAB
pJC175h	AP	Carb	SC101	$P_{psp\text{-}tet}$	gIII-luxAB
pJC175k	AP	Carb	SC101	$P_{psp\text{-}tetO3}$	gIII-luxAB
pJC175m	AP	Carb	SC101	$P_{psp\text{-}tetO5}$	gIII-luxAB
pJC175n	AP	Carb	SC101	$P_{psp\text{-}tetO6}$	gIII-luxAB
pJC184	MP	Chlor	CloDF13	P_{BAD}	dnaQ926-umuD'-umuC-recA730
pJC184c-rrnB	AP	Chlor	CloDF13	P_{psp}	gIII
pJC184d5	DP	Chlor	CloDF13	P_{BAD} / $P_{psp\text{-}tet}$	dnaQ926-umuD'-umuC-recA730 / gIII
pJC184d-Spect	AP	Chlor	CloDF13	$P_{psp\text{-}tet}$	gIII
SP-$T7_{WT}$	SP	Kan	f1	P_{gIII}	wt T7 RNAP
SP-$T7_{SP}$	SP	Kan	f1	P_{gIII}	T7 RNAP mutant L2-48.3
SP-$T7_{Pr}$	SP	Kan	f1	P_{gIII}	T7 RNAP mutant L1-192.2

Supplementary Table 1. List of plasmids used in this work.

Supplementary References

1 Cheetham, G. M., Jeruzalmi, D. & Steitz, T. A. Structural basis for initiation of transcription from an RNA polymerase-promoter complex. *Nature* **399**, 80-83, doi:10.1038/19999 (1999).