Supplementary Information for:

Processing generates 3' ends of RNA masking transcription termination events in prokaryotes

Xun Wang[a], Monford Paul Abishek N[b], Heung Jin Jeon[b], Yonho Lee[b], Jin He[a], Sankar Adhya[c, 1], Heon M. Lim[b, 1]

[a] State Key Laboratory of Agricultural Microbiology, College of Life Science and Technology, Huazhong Agricultural University, Wuhan, Hubei 430070, PR China;

[b] Department of Biological Sciences, College of Biological Sciences and Biotechnology, Chungnam National University, Daejeon 305-764, Republic of Korea

[c] Laboratory of Molecular Biology, National Institutes of Health, National Cancer Institute, Bethesda, MD 20892-4264, USA

[1] Correspondence: adhyas@mail.nih.gov or hmlim@cnu.ac.kr.

This PDF file includes:

Supplementary Materials and Methods

www.pnas.org/cgi/doi/10.1073/pnas.1813181116

Fig. S1 and S2

Table S1

References for SI reference citations

Supplementary Materials and Methods

Bacterial strains and growth conditions

Strain MG1655 Δ*gal* and the RNase E temperature-sensitive mutant GW20 Δ*gal* (W3110 *zce*-726::Tn10 *rne*-1 Δ*galETKM*) were used in this study. Chromosomal deletion strains were generated by deleting the corresponding gene(s) from MG1655 using lambda red-mediated recombination [1-3]. The primers used are listed in Table S1. Cells harboring the described plasmids were grown at 37 °C in LB medium (10 g tryptone, 5 g yeast extract, and 10 g NaCl per liter of water) supplemented with 0.5% (wt/vol) galactose and chloramphenicol (15 μg/ml). The cells were grown to an OD_{600} of 0.6 before RNA isolation. Unless indicated, all the chemicals were purchased from Sigma-Aldrich (USA).

RNA preparation

Total RNAs were purified from clarified cell lysates using the Direct-zol™ RNA MiniPrep kit (Zymo Research, USA). To generate the cell lysates,

equal numbers of cells (2×10^8) were harvested at an OD_{600} of 0.6 and resuspended in 50 μl protoplasting buffer (15 mM Tris-HCl, pH 8.0; 0.45 M sucrose; and 8 mM EDTA). Five microliters of lysozyme (50 mg/ml) was added, and then the samples were incubated for 5 min at 25 °C. A phenolic detergent (1 ml TRI Reagent; Molecular Research Center, USA) was added, and the mixtures were vortexed for 10 s before incubation for 5 min at 25 °C. The RNAs were then isolated according to the manufacturer's recommendations. and the RNAs were dissolved in 30 μll of RNA storage buffer (Ambion, USA). RNA concentrations were determined by measuring the absorbance at 260 nm using a NanoDrop spectrophotometer (Thermo Fisher Scientific, USA).

3' RACE assay

Total RNAs were extracted as described above. RNA ligation was performed at 37 °C for 3 h in a 15 μl reaction volume containing 2.5 g of total RNA, 2 nM synthetic RNA oligomer possessing a 5'-phosphate and 3' -inverted deoxythymidine (27-mer; Dharmacon, USA), 5 U of T4 RNA ligase (Ambion, USA), and 10 U of rRNasin (Promega, USA). The RNA ligation reaction was applied to a G-50 column (GE healthcare, USA). One microgram of RNA (eluted from the G-50 column) was reverse transcribed at 37 °C for 2 h in a 20-μl reaction volume containing 4 U of Omniscript

reverse transcriptase (QIAGEN, Germany), 0.5 mM each dNTP, 0.4 M 3RP primer complementary to the RNA oligomer (see Table S1), and 10 U of rRNasin. A 2-μl sample of this reaction was used as the template for PCR amplification of the *gal* cDNA with gene-specific primers and the 3RP primer (see Table S1) using HotStar Taq DNA polymerase (Qiagen, Germany). To assay the 3' ends of the *gal* mRNAs, the amplified cDNAs were purified and used as templates for primer extension reactions performed in a volume of 20 μl with ^{32}P-labeled primers (complementary to *galM* of the *gal* operon mRNA; see Table S1) and 1 U of HotStar Taq polymerase. The reaction products were separated on an 8% sequencing gel for 2 h at 60 w, and the radioactive bands were visualized after exposure of X-ray film.

In vitro transcription

The *in vitro* transcription was performed as previous described [4]. The pHL1703 plasmid (p*gal-gpmA*) was used as a DNA template for *in vitro* transcription (Supplementary Fig. 1). The *in vitro* transcription reactions were performed using *E. coli* σ70 (Epicentre, Sweden) according to the manufacturer's instructions. Briefly, DNA templates (2 nM) were incubated at 37 °C for 5 min in reaction buffer (20 mM Tris acetate, pH 7.8; 10 mM magnesium acetate; 200 mM potassium glutamate; 1 mM ATP;

and 1 mM dithiothreitol [DTT]) containing 2 U σ 70 and 40 U rRNasin in a 47.5-μl reaction. Rho, NusA and NusG were added at the amount indicated in the figure legends. The reactions were initiated by the addition of 2.5 μl NTP mix (final concentration, 0.1 mM each NTP) to the mixtures. After 30 min, the reactions were terminated by phenol-chloroform extraction, and then 30 μl of supernatants were purified using a G-50 column. The purified RNAs were then utilized for 3' RACE assays as described.

Northern blot analysis

The northern blot analysis was performed as previous described with some modifications.[4] Total RNA was isolated as described above. Total RNAs (10 μg with 10 μg/ml ethidium bromide) were electrophoresed in 1.2% (wt/vol) formaldehyde-agarose gel electrophoresis at 8 V/cm for 2 h. After electrophoresis, RNA integrities were assessed under UV light, and the RNAs were transferred overnight to a positively charged nylon membrane (Ambion, USA) using a downward transfer system (TurboBlotter; Whatman, UK). The RNAs were then fixed to the nylon membrane by baking at 80 °C for 1 h. The northern blot probe was prepared as follows. First, a 500-bp DNA fragment in the *galE* region (from +27 to +527) was prepared by PCR using primers indicated in Table S1. ^{32}P-labeled DNA

probes were then produced. Briefly, the template DNA (0.15 pmol) was mixed with random hexamer (4 μl of 1 mM) in a total volume of 28 μl and then heated at 95 °C for 3 min. Then, the reaction was rapidly cooled for 5 min on ice, and then 5 μl 10× Klenow fragment buffer, 5 μl dNTP mix (0.2 mM dATP, dGTP, and dTTP), 10 μl α^{32}P-dCTP, and 2 μl Klenow fragment (2 U/μl) (TAKARA, Japan) were added to the mixture. The reaction was performed by incubation at 37 °C for 1 h, and then the Klenow fragment was inactivated at 65 °C for 5 min (TAKARA, Japan). The product was purified via passage through the G-50 column. The blot was pre-hybridized in 7 ml ULTRAhyb® Hybridization Buffer (Ambion, USA) at 65° C for 30 min. The DNA probe was denatured at 95 °C for 5 min, and then 5 μl of the probe was added to the hybridization buffer. The hybridization was performed at 42 °C overnight. Then, the blot was washed twice in low-stringency wash buffer (2× SSC, 0.1% SDS) for 5 min each at room temperature and then twice in high-stringency wash buffer (0.2× SSC, 0.1% SDS) for 15 min each at 42 °C. Finally, the radioactive bands were visualized after exposure to X-ray film.

gal **mutants construction**

The plasmid p*gal-gpmA* was constructed via insertion of the galactose operon and the downstream gene, *gpmA* (from-75 to+5397), between the

*Eco*RI and *Bam*HI sites of pCC1BAC (Epicentre, Sweden). The DNA fragment was obtained via PCR amplification of the genomic DNA with the corresponding primer pairs as listed in Table S1. For the construction of the *gal* mutants, primers containing the desired mutations were synthesized. DNA fragments containing the mutations were obtained through PCR amplification. The DNA fragments were used as a "mega primer" for the next round of PCR[5]. The resulting PCR fragments were digested with *Mlu*I and *Bam*HI and then ligated into the plasmid p*gal-gpmA* to generate RIT^{o}1, RIT^{o}2, RIT^{o}3, RIT^{o}4, RDT^{o}, RDT^{o}*, and the *galM* stop^{o}. For the construction of RIT[2], the DNA fragment containing the *gal* intrinsic terminator sequence was ligated between *Bam*HI and *Sph*I of pHL1277 [4], generating an intermediate plasmid which was then single-cut with *Sph*I, treated with alkaline phosphatase, and ligated with DNA fragments containing sequence from 4334th to the end of *gpmA* obtained through PCR amplification.

Expression and purification of proteins

The *rho* gene was PCR amplified from genomic DNA and cloned between the *Nde*I and *Xho*I sites of pET-15b (Novagen, Germany) to generate pETRho. For the expression of Rho protein tagged with six histidine at the amino-terminus, BL21 (DE3)/pETRho cells were grown in LB medium

containing ampicillin (100 μg/ml) at 37 °C for 16 h. A 1:100 dilution of this culture was grown in fresh LB medium (50 ml), and cells were grown to an OD_{600} of 0.6 at 37 °C. IPTG (final concentration, 20 μM) was added, and the culture was grown for an additional 4 h. A total volume of 3 ml of the cells was pelleted and resuspended in 500 μl binding buffer (Bioneer, Korea). Cells were disrupted by sonication (25% amplitude, 5 s on and 5 s off for 20 cycles). The next steps of protein purification were performed according to the manufacturer's recommendations (Bioneer, Korea). Briefly, 50 μl of magnetic beads were equilibrated with 1 ml binding buffer with 5 min. The tube was placed on the magnetic stand. Cell pellets were centrifuged at 4°C, 12000 × *g* for 30 min, supernatants were collected and mixed with magnetic beads at room temperature (25°C) for 20 min. The tube was placed on the magnetic stand, and the supernatant was removed. 500 μl Binding & Washing solution was added to the beads and vortexed for 10 s. Washing was repeated for three time. Finally, 100 μl elution solution was added, mixed with tube rotator at room temperature (25°C) for 10 minutes. The tube was centrifuged for 5 s and the supernatant was collected. Protein concentration was determined by measuring A_{280} using NanoDrop spectrophotometer. NusA and NusG proteins were also purified as described for Rho.

In silico analysis of RIT and RDT distributions

The RITs were predicted using the WebGeSTer database [6]. The regions between +1 to +270, which are relative to the stop codons of the terminal cistrons among the operons were analyzed. RNA sequences and secondary structures containing the following parameters were recognized as RITs: stem length 4-12 bp, loop length 3-8 bp, max mismatch 3 bp, cut-off ΔG -12, hairpin followed by a 10 bp trail having >3 uridylates.

RDTs were analyzed in the regions between +1 to +700 relative to the stop codons of the terminal cistrons among the operons. The emboss freak program with a stepping value of 1 and an averaging window of 78 calculated the frequencies of cytidines and guanines residues (http://emboss.bioinformatics.nl/cgi-bin/emboss/freak). The regions which were longer than 78 bp, containing continuous C>G were identified as RDT.

Fig. S1

Schematic presentation of p*gal-gpmA* plasmid containing wild type *galETKM* and *gpmA* gene shown in blue and orange arrows, respectively.

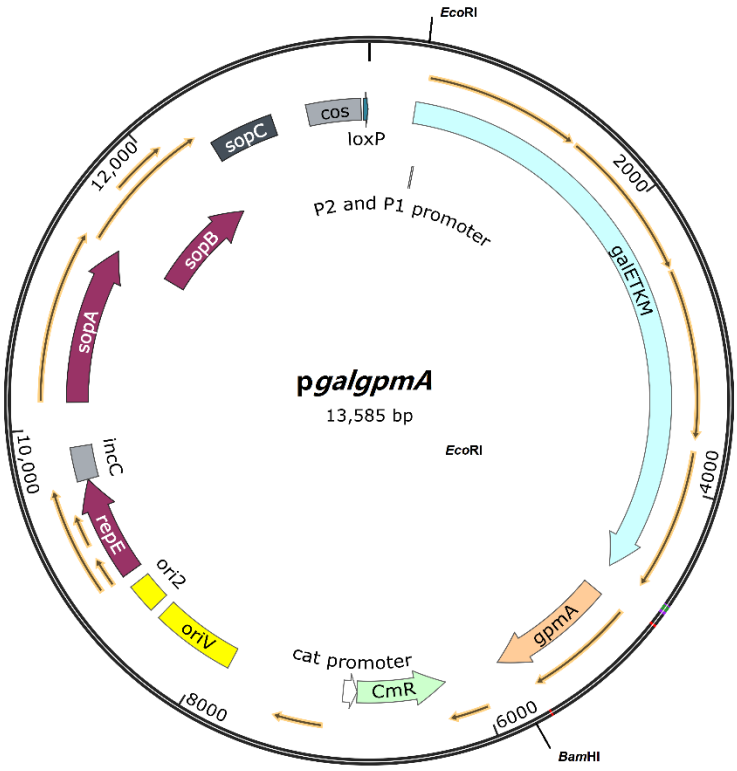
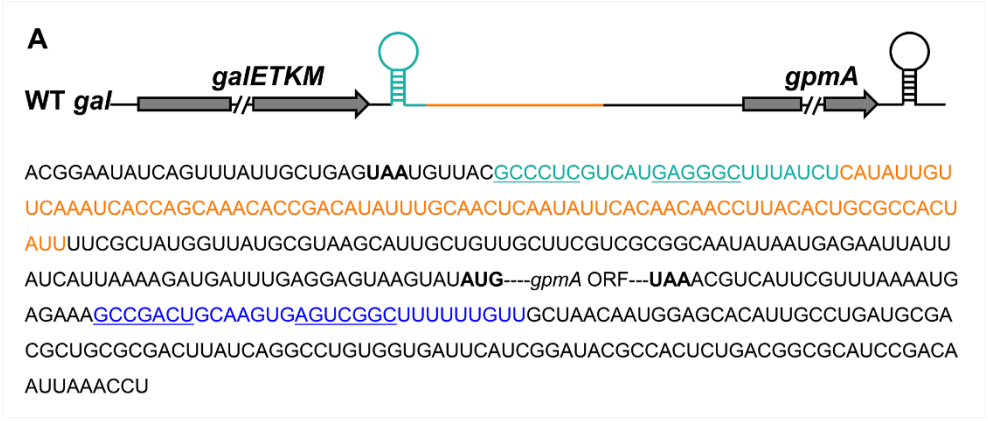


Fig. S2

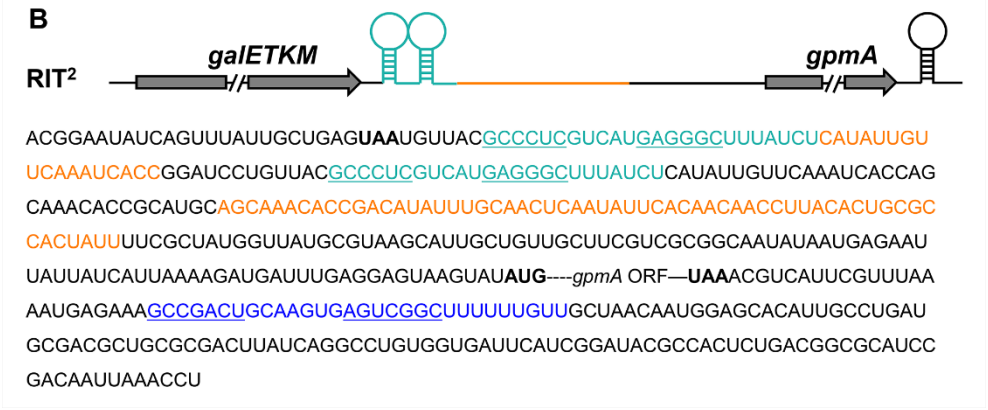
Top: A Schematic presentation of WT and mutant *gal* mRNAs and the neighboring transcript, *gpmA* mRNA. The two tandem terminators of *gal*; intrinsic termination signal (the terminator hairpin) presented as a cyan hairpin structure, and Rho-termination signal (C-rich region) depicted as an orange line. The intrinsic termination signal of *gpmA* is presented as a black hairpin. Deleted sequences are crossed in red.

Bottom: Sequence presentation of functional regions at the end of *gal* and *gpmA* mRNAs. The intrinsic termination signal sequences are marked in cyan. The stem sequences are underlined. The C-rich Region are shown in

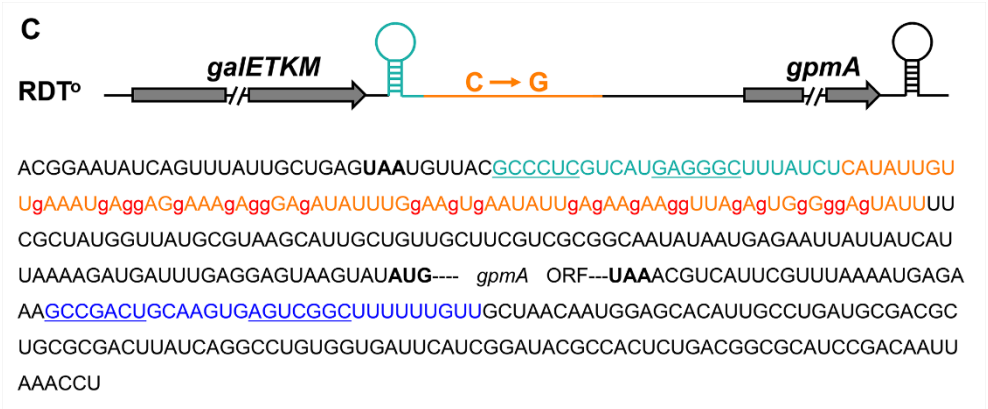
orange colors. Changed nucleotides are shown in red. The intrinsic terminator of *gpmA* is shown in blue, other functional sequences, such as translation stop codon of *galM*, translation start and stop codon of *gpmA* are shown in bold text.



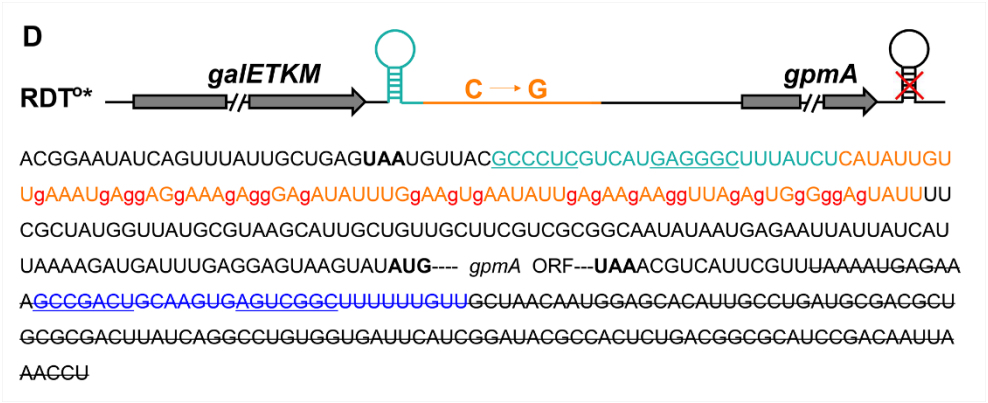
ACGGAAUAUCAGUUUAUUGCUGAG**UAA**UGUUACGCCCUCGUCAUGAGGGCUUUAUCUCAUAUUGU
UCAAAUCACCAGCAAACACCGACAUAUUUGCAACUCAAUAUUCACAACAACCUUACACUGCGCCACU
AUUUUCGCUAUGGUUAUGCGUAAGCAUUGCUGUUGCUUCGUCGCGGCAAUAUAAUGAGAAUUAUU
AUCAUUAAAAGAUGAUUUGAGGAGUAAGUAU**AUG**----*gpmA* ORF---**UAA**ACGUCAUUCGUUUAAAAUG
AGAAAGCCGACUGCAAGUGAGUCGGCUUUUUUGUUGCUAACAAUGGAGCACAUUGCCUGAUGCGA
CGCUGCGCGACUUAUCAGGCCUGUGGUGAUUCAUCGGAUACGCCACUCUGACGGCGCAUCCGACA
AUUAAACCU

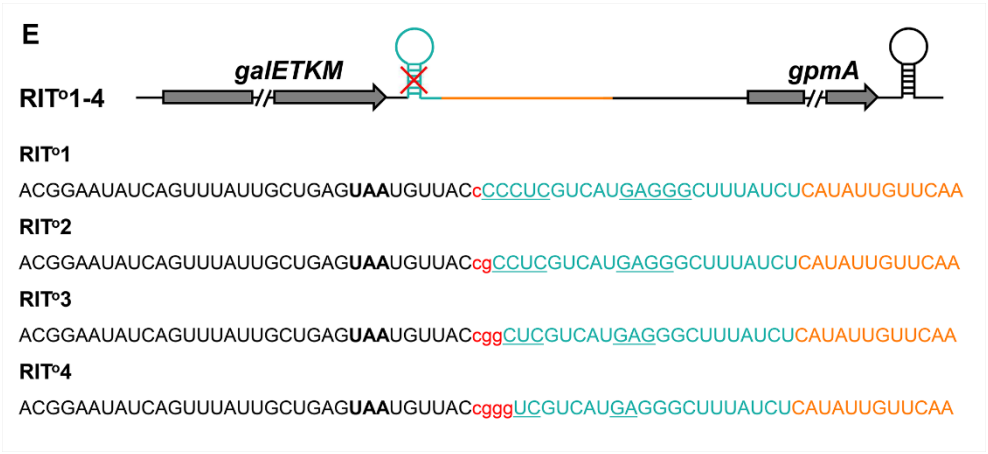


ACGGAAUAUCAGUUUAUUGCUGAG**UAA**UGUUACGCCCUCGUCAUGAGGGCUUUAUCUCAUAUUGU
UCAAAUCACCGGAUCCUGUUACGCCCUCGUCAUGAGGGCUUUAUCUCAUAUUGUUCAAAUCACCAG
CAAACACCGCAUGCAGCAAACACCGACAUAUUUGCAACUCAAUAUUCACAACAACCUUACACUGCGC
CACUAUUUUCGCUAUGGUUAUGCGUAAGCAUUGCUGUUGCUUCGUCGCGGCAAUAUAAUGAGAAU
UAUUAUCAUUAAAAGAUGAUUUGAGGAGUAAGUAU**AUG**----*gpmA* ORF—**UAA**ACGUCAUUCGUUUAA
AAUGAGAAAGCCGACUGCAAGUGAGUCGGCUUUUUUGUUGCUAACAAUGGAGCACAUUGCCUGAU
GCGACGCUGCGCGACUUAUCAGGCCUGUGGUGAUUCAUCGGAUACGCCACUCUGACGGCGCAUCC
GACAAUUAAACCU



ACGGAAUAUCAGUUUAUUGCUGAG**UAA**UGUUACGCCCUCGUCAUGAGGGCUUUAUCUCAUAUUGU
UgAAAUgAggAGgAAAgAggGAgAUAUUUGgAAgUgAAUAUUgAgAAgAAggUUAgAgUGgGggAgUAUUUU
CGCUAUGGUUAUGCGUAAGCAUUGCUGUUGCUUCGUCGCGGCAAUAUAAUGAGAAUUAUUAUCAU
UAAAAGAUGAUUUGAGGAGUAAGUAU**AUG**---- *gpmA* ORF---**UAA**ACGUCAUUCGUUUAAAAUGAGA
AAGCCGACUGCAAGUGAGUCGGCUUUUUUGUUGCUAACAAUGGAGCACAUUGCCUGAUGCGACGC
UGCGCGACUUAUCAGGCCUGUGGUGAUUCAUCGGAUACGCCACUCUGACGGCGCAUCCGACAAUU
AAACCU





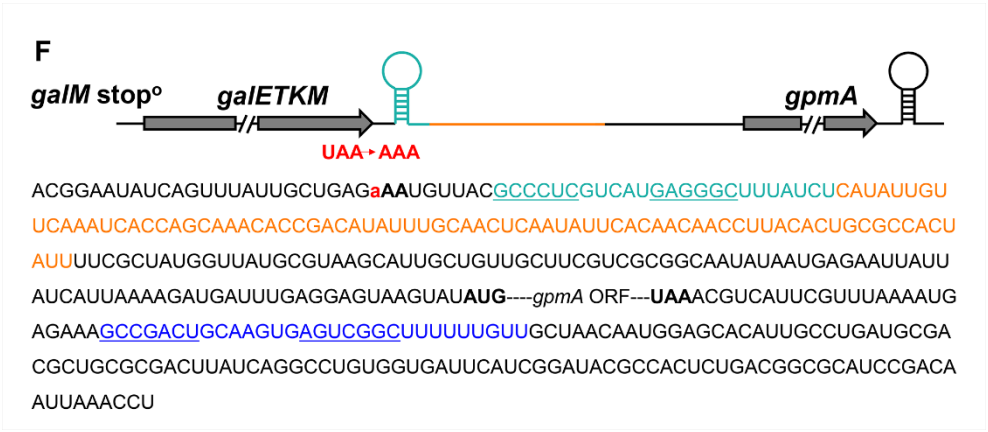


Table S1. Primers list.

Primer name	Primer sequences (5′→3′)	Usage
-73-EcoRI-for	GGCGAATTCAATTCTTGTGTAAACGATTCC	Cloning the *gal* operon and *gpmA* into pCC1BAC
gpmA-BamHI-	ATGCAGTAAGGATCCGTCGCGGGTAAG	Cloning the *gal* operon and

rev	G	*gpmA* into pCC1BAC
3RP	AGCATGCGGCCGCTAAGAAC	Revers transcription and PCR for 3' RACE assay
M3	TCCGCACGACGGCCTGAAAT	PCR for 3' RACE assay
M4240	CGAAGAGTATTCCAGCCTG	Primer extension for 3' RACE assay
RT-M-for2	GAACCTTACGCCGACTGGCAAG	Primer extension for RIT^{o}1-4 mutants 3' RACE assay
E1	ATGAGAGTTCTGGTTACCGGTGGTA	Northern blot analysis, the *galE* probe generation
E2	TGGGCTTTTTGCAGATCGGTGAGGA	Northern blot analysis, the *galE* probe generation
gal-2865-for	CATGAACTGGACCCGATCGTGG	Forward flanking primer for PCR mutagenesis
gal-5615-rev	CACCCCAGGCTTTACACTTTATGC	Reverse flanking primer for PCR mutagenesis
sphI-4334_For	CATGCATGCAGCAAACACCGACATATTTG	RIT^{2} plasmid construction
gmHMM1-for	ATGTTACCCCCTCGTCATGAGGGC	Mutagenic primer for RIT^{o}1 plasmid construction
gmHMM2-for	ATGTTACCGCCTCGTCATGAGGGC	Mutagenic primer for RIT^{o}2 plasmid construction
gmHMM3-for	ATGTTACCGGCTCGTCATGAGGGC	Mutagenic primer for RIT^{o}3 plasmid construction
gmHMM4-for	ATGTTACCGGGTCGTCATGAGGGC	Mutagenic primer for RIT^{o}4

		plasmid construction
gmCRR-for	GGAGATATTTGGAAGTGAATATTGAGAAGAAGGTTAGAGTGGGGGAGTATTTTCG	RDT^{o} plasmid construction
gmCRR-rev	CTTCCAAATATCTCCCTCTTTCCTCCTCATTTCAACAATATGAGATAAAGCCCTC	RDT^{o} plasmid construction
gmCRR*-rev	CGGGATCCAACGAATGACGTTTACTT	Mutagenic primer for RDT^{o}* plasmid construction
gmNS-for	GCTGAGAAATGTTACGCCCTCGTC	Mutagenic primer for *galM* sto p^{o} plasmid construction
rnbKO-for	TGTCAGCCGCTCTAATGGCCACCAAAATAGACAATTGTGTAGGCTGGAGCTGCTTC	Chromosomal *RNaseII (rnb)* deletion
rnbKO-rev	ACGGCCCATCCATGAGGAATGGGCCGTGAAAGGAGACTGTCAAACATGAGAATTAA	Chromosomal *RNaseII (rnb)* deletion
galE-EcoRI	GGATATCGTGGACTCTATCGCTATGTC	Chromosomal *gal* operon deletion
galM-HpaI	CGGTTAACAGCCATATACTTACTCCTC	Chromosomal *gal* operon deletion
rnbKO-con-for	CAGATTCGCGTAAAACTGTCAG	Chromosomal *RNaseII (rnb)* deletion
rnbKO-con-rev	GGAAATAAACGGCCCATCCATG	Chromosomal *RNaseII (rnb)* deletion

Reference

1. Datsenko, K.A. & Wanner, B.L. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products.

Proc Natl Acad Sci U S A **97**, 6640-5 (2000).

2. Yu, D. et al. An efficient recombination system for chromosome engineering in *Escherichia coli*. *Proc Natl Acad Sci U S A* **97**, 5978-83 (2000).
3. Sharan, S.K., Thomason, L.C., Kuznetsov, S.G. & Court, D.L. Recombineering: a homologous recombination-based method of genetic engineering. *Nat Protoc* **4**, 206-23 (2009).
4. Wang, X. et al. Expression of each cistron in the *gal* operon can be regulated by transcription termination and generation of a *galk*-specific mRNA, mK2. *J Bacteriol* **196**, 2598-606 (2014).
5. Tyagi, R., Lai, R. & Duggleby, R.G. A new approach to 'megaprimer' polymerase chain reaction mutagenesis without an intermediate gel purification step. *BMC Biotechnol* **4**, 2 (2004).
6. Mitra, A., Kesarwani, A.K., Pal, D. & Nagaraja, V. WebGeSTer DB--a transcription terminator database. *Nucleic Acids Res* **39**, D129-35 (2011).