SUPPLEMENTAL MATERIAL

Suppression of the *Escherichia coli rnpA49* conditionally lethal phenotype by different compensatory mutations

Arianne M. Babina[1,2]*, Leif A. Kirsebom[3], and Dan I. Andersson[1]

[1]Department of Medical Biochemistry and Microbiology, Uppsala University, Uppsala, Sweden

[2]Present Address: School of Infection and Immunity, University of Glasgow, Glasgow, United Kingdom

[3]Department of Cell and Molecular Biology, Uppsala University, Uppsala, Sweden

*For correspondence: arianne.babina@imbim.uu.se; Arianne.Babina@glasgow.ac.uk

Running head: Suppressors of *E. coli rnpA49*

Key words:

RNase P

Suppressor mutation

Lon

Gene amplification

RNase R

Table S1. *Escherichia coli* strains used in this study. All strains are derivatives of DA5438 (MG1655) or DA28696 (BW25113), as noted. All suppressor strains isolated are derivatives of DA61546 (MG1655 *rnpA49)*. Genomic coordinates are relative to the *E. coli* MG1655 reference genome (GenBank: U00096.3).

Strain	Genotype	Description	WGS?	Reference or Source
DA5438	MG1655 *fadK*::IS*1*, *yohK*[L34S]	MG1655 laboratory reference strain (*rnpA*(WT))	Yes; reference strain	Our collection
DA61319	F-, *lacZ8*(Am), *trpA36*, *glyA34*, *argA52*, *rpsL999*(*strR*), *rnpA49*(ts), *ilvG866*(Act)	Original A49 strain from CGSC; Strain #: 6465, Strain designation: N2020	No; local only	CGSC; Apirion et al. 1980
DA61546	DA5438 *yceM*/*murJ*::1127043::DELG, *malQ*[G95fs], *rnpA*[R46H]	MG1655 temperature-sensitive A49 parental strain (*rnpA49*)	Yes	This study
DA64107	DA61546 DUP(3737267..3928428)	3x *rnpA49* amplification	Yes	This study
DA64110	DA61546 DUP(3844972..3918184), *lsrA*::1600867::A>T (synonymous)	2x *rnpA49* amplification	Yes	This study
DA65198	DA61546 DUP(3868817..389802)	5x *rnpA49* amplification	Yes	This study
DA65199	DA61546 DUP(3804071..3899637)	2x *rnpA49* amplification	Yes	This study
DA56187	DA61546 DUP(3251561..3286079), *insA*::290499::A>G (synonymous)	16x *rnpB* amplification	Yes	This study
DA65194	DA61546 DUP(3267203..3284424)	7x *rnpB* amplification	Yes	This study
DA64106	DA61546 *rnr*::DEL(4404082..4404091)	*rnr*::Δ10bp	Yes	This study
DA65197	DA61546 *rnr*::440924::IS*1*, *trmA*::4160292::IS*186*	*rnr*::IS*1*	Yes	This study
DA64111	DA61546 *lon*::458019::IS*186*	*lon*::IS*186*	Yes	This study
DA64113	DA61546 *lon*::458020::IS*186*	*lon*::IS*186*	Yes	This study
DA65192	DA61546 *lon*[C39F] (TGT>TTT)	Lon C39F	No; local only	This study
DA65193	DA61546 *lon*[C39Y] (TGT>TAT)	Lon C39Y	No; local only	This study
DA65200	DA61546 *lon*[P63S] (CCG>TCG)	Lon P63S	No; local only	This study
DA64108	DA61546 *lon*[Q225P] (CAG>CCG)	Lon Q225P	Yes	This study
DA64112	DA61546 *lon*[230R] (CTG>CGG)	Lon L230R	Yes	This study
DA64109	DA61546 *lon*[E240G] (GAA>GGA)	Lon E240G	Yes	This study
DA65195	DA61546 *lon*[E240K] (GAA>AAA)	Lon E240K	No; local only	This study
DA74911	DA61546 DEL*rnr*::*kanR*	*rnr* deletion control strain	No; local only	This study
DA74912	DA61546 DEL*trmA*::*kanR*	*trmA* deletion control strain	No; local only	This study
DA74851	DA61546 DEL*pcnB*::*kanR*	*pcnB* deletion control strain	No; local only	This study
DA73974	DA61546 DEL*lon*::*kanR*	*lon* deletion control strain	No; local only	This study

DA22599	DA5438 DEL*lon*(partial)	MG1655 *rnpA*(WT) *lon* deletion strain; the deletion in *lon* leaves intact sRNA gene *sraA* and the promoter for the downstream *hupB* gene.	No	Our collection; Nicoloff & Andersson 2013
DA31331	MG1655 /pCA24N ASKA clone library	MG1655 transformed with pooled ASKA clones. Estimated 460,000 clones resuspended from plates after transformation and pooled.	No	Our collection; Kitigawa et al. 2005
DA67027	DA61546 /pCA24N (-gfp) empty	A49 strain carrying the empty ASKA collection plasmid	No	This study; Kitigawa et al. 2005
DA7016	DA61546 /pCA24N-*rnpA* (A3200)	A49 strain carrying the *rnpA*(WT) ORF cloned into the ASKA collection plasmid; recovered from the ASKA library screen	No	This study; Kitigawa et al. 2005
DA29686	MG1655 /pCA24N (-gfp) empty	MG1655 (*rnpA*(WT)) carrying the empty ASKA collection plasmid	No	Our collection; Kitigawa et al. 2005
K3055	ME9062 BW25113 *rrnB*, DEL*lacZ*4787, *hsdR*514, DEL(*araBAD*)567, DEL(*rhaBAD*)568, *rph*-1, DEL*rnr*::*kanR*	KEIO collection DEL*rnr*::*kanR* knockout strain	No	Baba et al. 2006
K3848	ME9062 BW25113 *rrnB*, DEL*lacZ*4787, *hsdR*514, DEL(*araBAD*)567, DEL(*rhaBAD*)568, *rph*-1, DEL*trmA*::*kanR*	KEIO collection DEL*trmA*::*kanR* knockout strain	No	Baba et al. 2006
K3738	ME9062 BW25113 *rrnB*, DEL*lacZ*4787, *hsdR*514, DEL(*araBAD*)567, DEL(*rhaBAD*)568, *rph*-1, DEL*pcnB*::*kanR*	KEIO collection DEL*pcnB*::*kanR* knockout strain	No	Baba et al. 2006
K309	ME9062 BW25113 *rrnB*, DEL*lacZ*4787, *hsdR*514, DEL(*araBAD*)567, DEL(*rhaBAD*)568, *rph*-1, DEL*lon*::*kanR*	KEIO collection DEL*lon*::*kanR* knockout strain	No	Baba et al. 2006

Table S2. Summary of *Escherichia coli* A49 suppressor strains containing genomic amplifications. In 'Copy Number', the estimated copy number and relevant gene within the amplified units are given. In 'Relevant Genotype', the genomic coordinates for the boundaries of the amplified units are included. 'Characteristics of Amplifications' provides information on sequence homology flanking the boundaries/break-points of each amplified unit; 'ND,' not determined. Genomic coordinates are relative to the *E. coli* MG1655 reference genome (GenBank: U00096.3).

Strain	Copy Number	Relevant Genotype	Size of Amplified Unit (kbp)	Characteristics of Amplifications
DA64107	3x *rnpA49*	*rnpA49*, DUP(3737267..3928428)	192.5	IS*4*
DA64110	2x *rnpA49*	*rnpA49*, DUP(3844972..3918184)	73	ND
DA65198	5x *rnpA49*	*rnpA49*, DUP(3868817..389802)	29	IS*1*
DA65199	2x *rnpA49*	*rnpA49*, DUP(3804071..3899637)	95.6	IS*1*
DA65187	16x *rnpB*	*rnpA49*, DUP(3251561..3286079)	34.5	IS*1*
DA65194	7x *rnpB*	*rnpA49*, DUP(3267203..3284424)	17.2	IS*1*

Table S3. Summary of *Escherichia coli* A49 suppressor strains containing insertions.

In 'Duplication at Insertion Site', nucleotides in parentheses are not present in all sequencing reads for that strain. The genomic coordinates for the boundaries flanking the duplication and insertion sites are given, along with the estimated site of insertion. Genomic coordinates are relative to the *E. coli* MG1655 reference genome (GenBank: U00096.3).

Strain	**Relevant Genotype**	**Mobile Genetic Element/ Transposase**	**Duplication at Insertion Site (5'-3')**	**Boundaries of Duplication and Insertion Site**	**Approx. Position of Insertion**
DA65197	*rnpA49*, *rnr*::IS*1*,	IS*1*	CGGCCTGTG	4404920..4404928	440924
	trmA::IS*186*	IS*186*	GGGAAGCTATCC	4160286..4160297	4160292
DA64111	*rnpA49*, *lon*::IS*186*	IS*186*	GGGGAAACAT(C)CCCATA	458011..458027	458019
DA64113	*rnpA49*, *lon*::IS*186*	IS*186*	GGGAAACAT(C)CCCATA	458012..458027	458020

Table S4. Oligonucleotides used in this study.

Name	Sequence (5'-3')	Description	Source
37-rmpH-rnpAseqF	CCCGATTTATCCACAGGAC	for PCR amplifying and sequencing the *rmpH-rnpA* locus (*rnpA* promoters and upstream sequence)	This study
38-rnpAseqF	CAACCGTTCTCACGGCTTC	for PCR amplifying and sequencing the *rnpA* locus (*rnpA* ORF)	This study
39-rnpAseqR	CACCGTCAACCAACTGCC	for PCR amplifying and sequencing the *rnpA* locus	This study
244-EclonSeq-F	GTGGTTATCGACGAGTC	for PCR amplifying and sequencing the *lon*/*lon*::*kanR* locus	This study
245-EclonSeq-F2	CACCAGAAGTGCTGACG	for sequencing the *lon* locus	This study
246-EclonSeq-F3	CTTGGTCAGTCCATTGC	for sequencing the *lon* locus	This study
247-EclonSeq-R	CGATCCGCCATCTAAC	for PCR amplifying and sequencing the *lon*/*lon*::*kanR* locus	This study
765-pcnBseqF	GCAAATCCTTCAGTCAGCCG	for PCR amplifying and sequencing the *pcnB*/*pcnB*::*kanR* locus	This study
766-pcnBseqR	GGTAAGAATGTGGCTTTCAGG	for PCR amplifying and sequencing the *pcnB*/*pcnB*::*kanR* locus	This study
773-rnrseqF	CTTACGGAACTGGATAACTACACG	for PCR amplifying and sequencing the *rnr*/*rnr*::*kanR* locus	This study
774-rnrseqR	CCGGCCTACATGATCTCTGC	for PCR amplifying and sequencing the *rnr*/*rnr*::*kanR* locus	This study
204-ECrnpBqPCRF	AGTACGGGCCGTACCTTATGAA	for RT-qPCR targeting M1 RNA (*rnpB*)	Loveland et al. 2014
205-ECrnpBqPCRR	TCCGTGGCACGGTAAACTC	for RT-qPCR targeting M1 RNA (*rnpB*)	Loveland et al. 2014
212-ECcysGqPCRF	TTGTCGGCGGTGGTGATGTC	for RT-qPCR targeting *cysG* normalization control	Zhou et al. 2011
213-ECcysGqPCRR	ATGCGGTGAACTGTGGAATAAACG	for RT-qPCR targeting *cysG* normalization control	Zhou et al. 2011
214-EChcaTqPCRF	GCTGCTCGGCTTTCTCATCC	for RT-qPCR targeting *hcaT* normalization control	Zhou et al. 2011
215-EChcaTqPCRR	CCAACCACGCAGACCAACC	for RT-qPCR targeting *hcaT* normalization control	Zhou et al. 2011
pCA24N-F	GATTCAATTGTGAGCGGATAACAATTTC	for sequencing ORFs cloned into the ASKA collection pCA24N plasmid	Our collection; Omar Warsi
pCA24N-R	CATTACTGGATCTATCAACAGGAG	for sequencing ORFs cloned into the ASKA collection pCA24N plasmid	Our collection; Omar Warsi

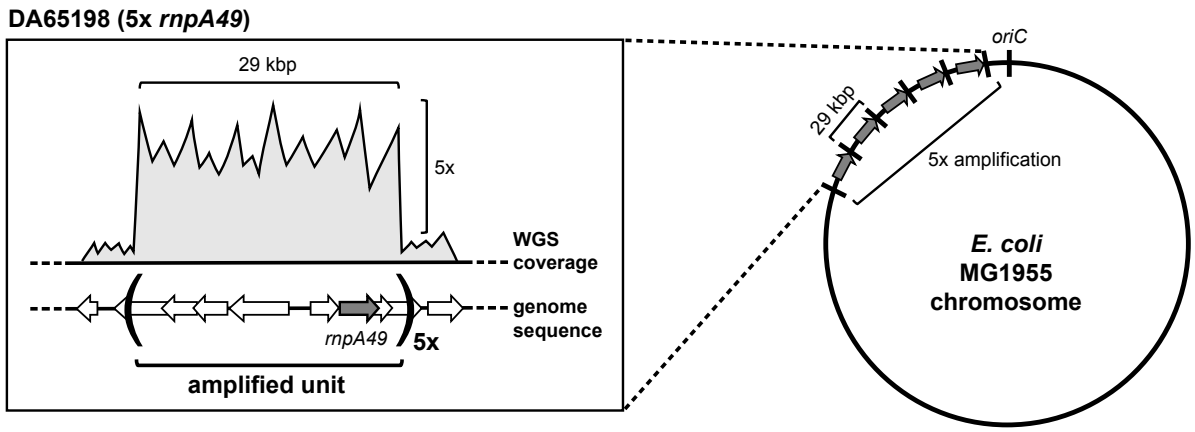
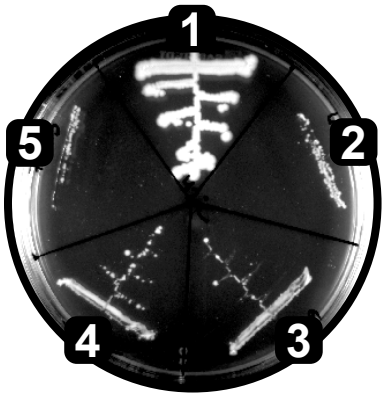


Figure S1. Suppression of the *E. coli rnpA449* ts phenotype via gene amplifications. Six suppressor strains were found to contain tandem amplifications of chromosomal regions containing either *rnpA49* or *rnpB* that increased strain fitness at the sub-lethal and non-permissive temperatures. The amplified units encompass large regions of the chromosome (ranging from 17.2 to 192.5 kbp) and contain many other genes in addition to the genes of interest (in this case, *rnpA49* or *rnpB*). Amplifications commonly form through homologous recombination between directly repeated sequences, such as insertion sequence elements, transposons, or rRNA operons (Table S2) (Tlsty et al. 1984; Reams and Roth 2015; Nicoloff et al. 2019). Genome amplifications are typically identified as regions with increased coverage in whole genome sequencing (WGS) data, and the increase in copy number can be estimated by dividing the average sequence coverage of the amplifications by the average sequence coverage of the rest of the chromosome (Nicoloff et al. 2019). A schematic for suppressor strain DA65198 is shown, which carries a 5x amplification of a 29 kbp chromosomal region containing the *rnpA49* allele, among other genes. Schematic not drawn to scale.



1. DA5438 (*rnpA*^{WT})
2. DA61546 (*rnpA49*)
3. DA65197 (*rnpA49, rnr*::IS*1, trmA*::IS*186*)
4. DA74911 (*rnpA49,* Δ*rnr*::*kanR*)
5. DA74913 (*rnpA49,* Δ*trmA*::*kanR*)

Figure S2. Disruption of *trmA* does not contribute to the rescue phenotype of suppressor strain DA65197. Growth of select *E. coli* strains on LB-agar after 24 h incubation at the non-permissive temperature (42°C). DA5438 is *E. coli* MG1655 carrying the wild-type *rnpA* allele; DA61546 is the *E. coli* A49 parental strain carrying the ts *rnpA49* allele; DA65197 carries the *rnpA49* allele along with the *rnr*::IS*1* suppressor mutation and the additional *trmA*::IS*186* mutation. Control strains containing *kanR* deletion cassettes from the KEIO collection in the parental DA61546 A49 background are included for comparison.

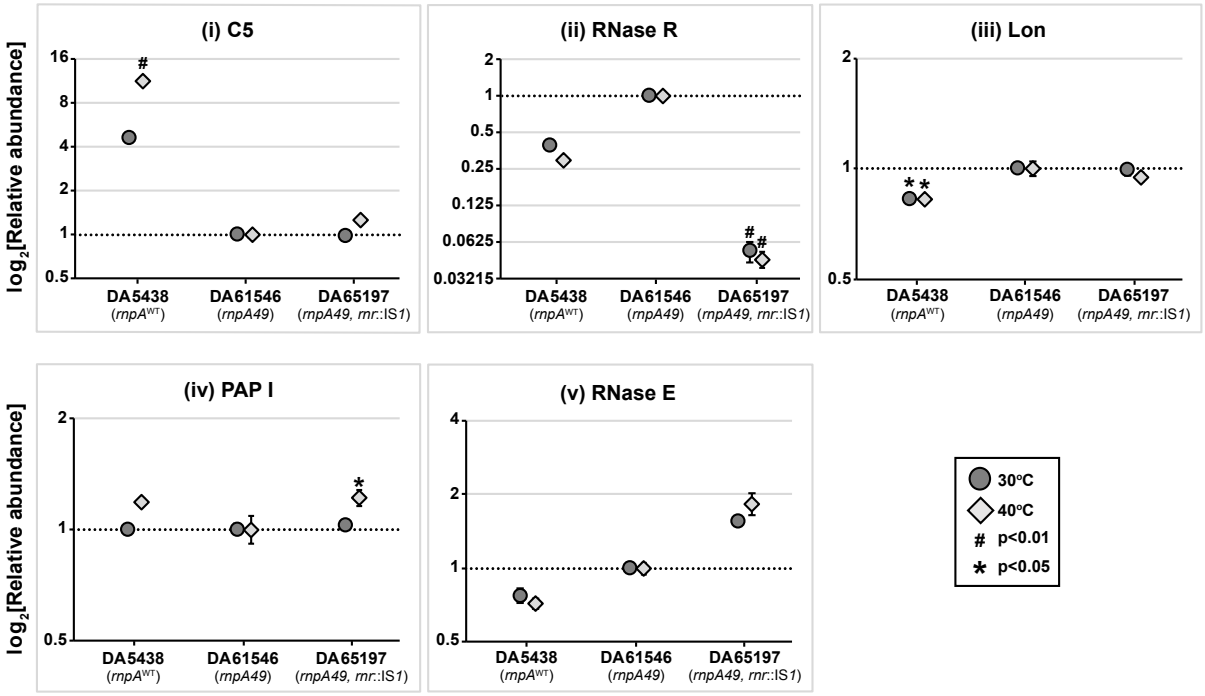


Figure S3. Relative abundance of select proteins in strain DA65197 carrying the *rnr*::IS*1* suppressor mutation. Relative abundance of select proteins during early-to-mid exponential phase growth at the permissive (30°C) or sub-lethal temperature (40°C), as compared to the original A49 parental strain (DA61546, set to 1). DA5438 is *E. coli* MG1655 carrying the wild-type *rnpA* allele; DA65197 carries the *rnpA49* mutant allele in addition to the *rnr*::IS*1* suppressor mutation. The values reported represent the mean of three independent biological replicates; error bars represent standard deviation of the mean. To determine statistical significance, the relative abundance of select proteins from DA5438 and DA65197 were compared to that of the DA61546 A49 parental strain at the corresponding temperature using a two-tailed Mann-Whitney test; values were considered significantly different if $p<0.05$ (*) and very significantly different if $p<0.01$ (#). Of note: (ii) RNase R is depleted in DA65197 on account of the IS*1* insertion. (iii) Lon abundance is elevated in strains DA61546 and DA65197; this is likely due to the activation of stress response pathways on account of the *rnpA49* ts allele (Gur 2013).

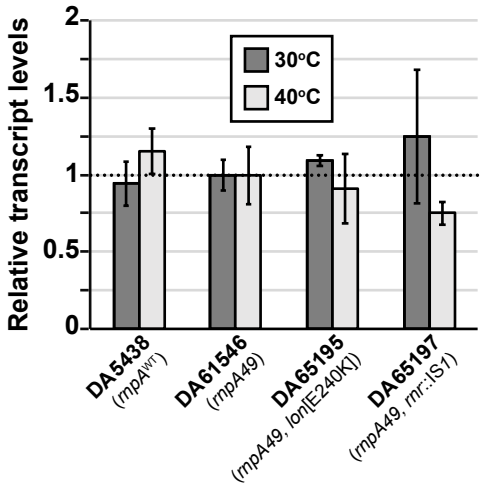


Figure S4. Select *E. coli* A49 suppressor mutations do not significantly affect M1 RNA levels. Relative transcript levels of M1 RNA (encoded by *rnpB*) in select *E. coli* strains grown to early-to-mid exponential phase at the permissive (30°C) or sub-lethal temperature (40°C), as measured by RT-qPCR and compared to the original A49 parental strain (DA61546, set to 1). DA5438 is *E. coli* MG1655 carrying the wild-type *rnpA* allele; strains DA65195 and DA65197 carry the *rnpA49* mutant allele in addition to the suppressor mutations indicated. The values reported represent the mean of three independent biological replicates (with three technical replicates each); error bars represent standard deviation of the mean. No statistically significant differences were found between the M1 RNA transcript levels measured in DA61546 and those measured in the other strains at the corresponding temperatures, as determined using a two-tailed Mann-Whitney test ($p<0.05$).

1. DA61546 **-** (*rnpA49*)
2. DA65200 **-** (*rnpA49, lon*[P63S])
3. DA65195 **+** (*rnpA49, lon*[E240K])
4. DA64109 **++** (*rnpA49, lon*[E240G])
5. DA73974 **+++** (*rnpA49,* Δ*lon*::*kanR*)

Figure S5. Select A49 *lon* mutant suppressor strains are mucoid at 30°C. Growth of select *E. coli* strains on LB-agar after 24 h incubation at the permissive temperature (30°C); not mucoid at 30°C (-) or mucoid at 30°C (+ < ++ < +++). DA61546 is the *E. coli* A49 parental strain carrying the ts *rnpA49* and wild-type *lon* alleles; DA65200, DA65195, and DA64109 carry the *rnpA49* allele along with the indicated nonsynonymous *lon* mutations. DA73974 is a control strain containing the Δ*lon*::*kanR* deletion cassette from the KEIO collection in the parental DA61546 A49 background.

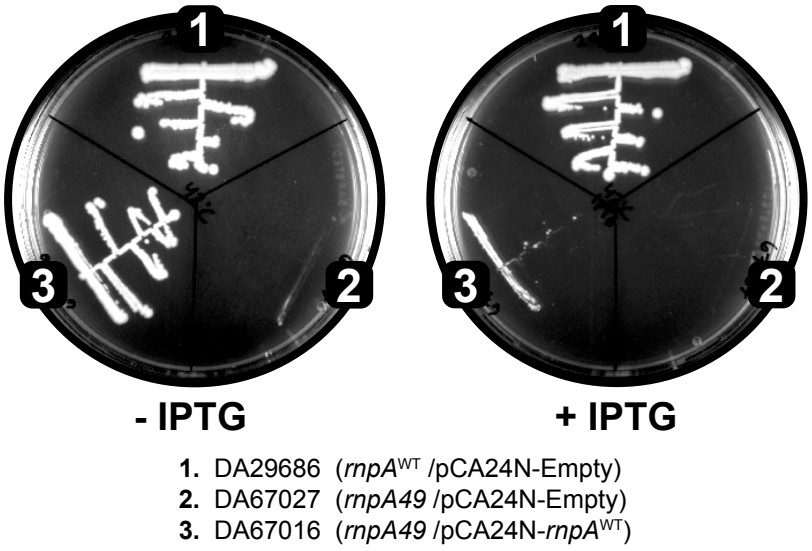


Figure S6. Only overexpression of *rnpA*^{WT} was found to rescue *E. coli* strain A49 following an ASKA library screen. Growth of select *E. coli* strains on LB-agar supplemented with 20 μg/ml chloramphenicol ± 1 mM IPTG after 24 h incubation at the non-permissive temperature (42°C). DA67027 carries the pCA24N-*rnpA*^{WT} plasmid (encoding the C5 protein) isolated from the only "positive hit" recovered from the ASKA library screen performed with our *E. coli* A49 strain to identify potential ORFs whose overexpression could rescue strain growth at the non-permissive temperature. ORFs cloned into plasmid pCA24N are under the control of the IPTG-inducible $P_{T5\text{-}lac}$ promoter, and expression can be "leaky" depending on the cloned ORF. High overexpression of C5 protein is toxic to *E. coli*, thus DA67016 exhibits reduced growth on the + IPTG plate (Jovanovic et al. 2002).

REFERENCES

Apirion D. 1980. Genetic mapping and some characterization of the *rnpA49* mutation of *Escherichia coli* that affects the RNA-processing enzyme Ribonuclease P. *Genetics* **94**: 291–299.

Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M, Wanner BL, Mori H. 2006. Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: The Keio collection. *Mol Syst Biol* **2**: 2006 0008.

Gur E. 2013. The Lon AAA+ protease. In *Regulated Proteolysis in Microorganisms* (ed. D.A. Dougan), Vol. 66, pp. 35–51, Springer, New York.

Jovanovic M, Sanchez R, Altman S, Gopalan V. 2002. Elucidation of structure-function relationships in the protein subunit of bacterial RNase P using a genetic complementation approach. *Nucleic Acids Res* **30**: 5065–5073.

Kitagawa M, Ara T, Arifuzzaman M, Ioka-Nakamichi T, Inamoto E, Toyonaga H, Mori H. 2005. Complete set of ORF clones of *Escherichia coli* ASKA library (A complete set of *E. coli* K-12 ORF archive): Unique resources for biological research. *DNA Res* **12**: 291–299.

Loveland JL, Rice J, Turrini PCG, Lizotte-Waniewski M, Dorit RL. 2014. Essential is not irreplaceable: Fitness dynamics of experimental *E. coli* RNase P RNA heterologous replacement. *J Mol Evol* **79**: 143–152.

Nicoloff H, Andersson DI. 2013. Lon protease inactivation, or translocation of the *lon* gene, potentiate bacterial evolution to antibiotic resistance. *Mol Microbiol* **90**: 1233–1248.

Nicoloff H, Hjort K, Levin BR, Andersson DI. 2019. The high prevalence of antibiotic heteroresistance in pathogenic bacteria is mainly caused by gene amplification. *Nat Microbiol* **4**: 504–514.

Reams A, Roth JR. 2015. Mechanisms of gene duplication and amplification. *Cold Spring Harb Perspect Biol* **7**: 1–25.

Tlsty TD, Albertini AM, Miller JH. 1984. Gene amplification in the *lac* region of *E. coli*. *Cell* **37**: 217–224.

Zhou K, Zhou L, Lim Q, Zou R, Stephanopoulos G, Too HP. 2011. Novel reference genes for quantifying transcriptional responses of *Escherichia coli* to protein overexpression by

quantitative PCR. *BMC Mol Biol* **12**: 18.