Supplementary information

A ProQ/FinO family protein involved in plasmid copy number control favours fitness of bacteria carrying *mcr-1*-bearing IncI2 plasmids

Jun Yang[1,2], Hai-Hong Wang[2,3], Yaoyao Lu[1,2], Ling-Xian Yi[1,2], Yinyue Deng[4], Luchao Lv[1,2], Vincent Burrus[5], Jian-Hua Liu[1,2*]

[1] College of Veterinary Medicine, National Risk Assessment Laboratory for Antimicrobial Resistant of Microorganisms in Animals, Guangdong Provincial Key Laboratory of Veterinary Pharmaceutics Development and Safety Evaluation, Key Laboratory of Zoonosis of Ministry of Agricultural and Rural Affairs, South China Agricultural University, Guangzhou, China. [2] Guangdong Laboratory for Lingnan Modern Agriculture, Guangzhou, China. [3] Guangdong Provincial Key Laboratory of Protein Function and Regulation in Agricultural Organisms, College of Life Sciences, South China Agricultural University, Guangzhou, China. [4] School of Pharmaceutical Sciences (Shenzhen), Sun Yat-sen University, Guangzhou 510275, China. [5] Département de biologie, Université de Sherbrooke, Sherbrooke J1K 2R1, Québec, Canada.

*E-mail: jhliu@scau.edu.cn.

Table S1. *E. coli* strains used in this study

Strains	Genotype or phenotype	**References**
BW25113	Δ(*araD-araB*)*567,* Δ*lacZ4787*(::*rrnB-3*)*, λ⁻, rph-1,* Δ(*rhaD-rhaB*)*568, hsdR514*	(1)
BL21(DE3)	F⁻, *ompT, hsdSB* (*rB⁻mB⁻*)*, gal, dcm*	Novagen
SM10λpir	*thi, thr, leu, tonA, lacY, supE, recA*::*RP4-2-Tc*::*Mu*	(2)
BW25113Δ*gadW*	BW25113 with deletion of *gadW* (AraC family transcriptional regulator)	This study
BW25113Δ*gapdh*	BW25113 with deletion of *gapdh* (glyceraldehyde-3-phosphate dehydrogenase pseudogene)	This study
BW25113Δ*loip*	BW25113 with deletion of *loip* (metalloprotease LoiP)	This study
BW25113Δ*ack*	BW25113 with deletion of *ack* (acetate kinase)	This study
BW25113Δ*hf*	BW25113 with deletion of *hf* (hypothetical protein, locus on BW25113 genome: 4181949-4182155)	This study
BW25113Δ*arnT*	BW25113 with deletion of *arnT* (4-amino-4-deoxy-L-arabinose lipid A transferase)	This study
BW25113Δ*tf*	BW25113 with deletion of *tf* (transcriptional regulator, locus on BW25113 genome: 623991-625367)	This study

Table S2. Plasmids used in this study

Plasmids	Genotype or phenotype	References
pHNSHP45	IncI2 replicon, colistin^R	(3)
pHN1122-1	IncI2 replicon, ceftazidime^R	(4)
pHNSHP45△*pcnR*	pHNSHP45 with deletion of *pcnR*	This study
pHNSHP45△*orf00050*	pHNSHP45 with deletion of *orf00050*	This study
pHNSHP45△*orf00039*	pHNSHP45 with deletion of *orf00039*	This study
pHNSHP45△*pcnR*△*mcr-1*	pHNSHP45 with deletion both of *pcnR* and *mcr-1*	This study
pHN1122-1△*pcnR*	pHN1122-1 with deletion of *yaeC* (*pcnR*)	This study
pHSG575	$oriV_{pSC101}$(Cm^R)	(5)
pHSG575-*pcnR*	pHSG575 carrying *pcnR* with native promoter (Cm^R)	This study
pBAD24M	$oriV_{p15A}$; P_{ara} (Amp^R)；NcoI site of expression vector pBAD24 changed to an NdeI site	(6)
pBAD24M-*pcnR*	pBAD24 carrying *pcnR* under the control of P_{ara} (Amp^R)	This study
pHSG575-*mcr-1*	pHSG575 carrying *mcr-1* with native promoter (Cm^R)	This study
pHSG575-*repA*	pHSG575 carrying *repA*	This study
pHSG575-*rpoB*	pHSG575 carrying *rpoB*	This study
pHN7A8	IncFII plasmid and used as template for cloning *finO* gene (Amp^R)	(7)
pHSG575-*finO*	pHSG575 carrying *finO* with native promoter (Cm^R)	This study

pSC189	ori$V_{R6K\gamma}$; TnSC189 (Kan^R)	(8)
pTns1-pTns7	Plasmids in growth-restored transposon mutants	This study
pHGR01	Containing promoterless *lacZ* (Kan^R)	(9)
P_{repA}-*lacZ*	The upstream region (-450 to +24) of *repA* was fused with promoterless *lacZ*	This study
P_{repA}-*lacZ*-repR(T22A)	P_{repA}-*lacZ* with mutation (T-A) in +22 position of *repR*	This study
P_{repA}-*lacZ*-△AS RNA	P_{repA}-*lacZ* with mutation (CA-TT) in the -35 region of antisense RNA promoter	This study
pUC19-AS RNA	pUC19 vector carrying antisense RNA with native promoter	This study
P_{repA}-*lacZ* (GAAA-CTCT)	P_{repA}-*lacZ* with mutation (GAAA-CTCT) in +9 to +12 position of *repR*	This study
P_{repA}-*lacZ* (GA-CT)	P_{repA}-*lacZ* with mutation (GA-CT) in +5 and +6 position of *repR*	This study
P_{repA}-*lacZ*-repR(9)fusion	The upstream region (-450 to +9) of *repR* was fused with promoterless *lacZ*	This study
P_{repA}-*lacZ*-repR(24)fusion	The upstream region (-450 to +24) of *repR* was fused with promoterless *lacZ*	This study
pET28b	$oriV_{pBR322}$; P_{T7}; expression vector (Kan^R)	Novagen
pET28b-*pcnR*	pET28b carrying *pcnR* under the control of T7 promoter	This study
pKD46	$oriV_{psc101}$(Ts); λRed recombinase expression vector	(1)
pCP20	Flp recombinase expression vector (Cm^RAmp^R) (Ts)	(1)
pKD4	PCR template for λRed recombination system (Kan^R)	(1)

Table S3. Primers used in this study

Primers	Sequence[a]
knock-pcnR-F	TCAGGATATTAACTGGTGAATATGGTATAATTAGTCAATTGTAGGCTGGAGCTGCTTCG
knock-pcnR-R	ACAGTCAGCCAGGCTGAAATGATAAAGCCCTTCGGGGCTATGGGAATTAGCCATGGTCC
knock-orf50-F	ACAATTAAAACGTAAGGGATAAAGCCCCGAAGGGCTTTATGTAGGCTGGAGCTGCTTCG
knock-orf50-R	TACACAGTAAAGCTAATGTAACACACGAAAGGTGATATAATGGGAATTAGCCATGGTCC
knock-orf39-F	AGCCTGTAGTGAGTGGCTGCAACTTTTCATAATTAGCTTTGTAGGCTGGAGCTGCTTCG
knock-orf39-R	CCACCACCTGGTTTTTTGATGGTTGCGTTTATTTGGTACATGGGAATTAGCCATGGTCC
knock-mcr-1-F	CCGTAATTATCCCACCGTTTATTTTTTGAGTAGTTTCTCTGTAGGCTGGAGCTGCTTCG
knock-mcr-1-R	TGGAGTGTGCGGTGGGTTTGGAAAAAATACAGGGAGAAAATGGGAATTAGCCATGGTCC
knock-gadW-F	AGGTGTGAGATCCTGACCAATATTCAAATGCGAAATATG TGTAGGCTGGAGCTGCTTCG
knock-gadW-R	AAATTGAAATAATCGCAGTATGAAATATAAGGGATAATCATGGGAATTAGCCATGGTCC
knock-gapdh-F	AGTGAAAGAGGCGGAGGTTTTTTCCTCCGCCTGTGCGCGTGTAGGCTGGAGCTGCTTCG
knock-gapdh-R	GATCAATTTTCATCCGAACGTTCCTGACAGGAGAAAACCATGGGAATTAGCCATGGTCC
knock-loip-F	ATTTTTAGAATAATCCTGACCTTGTGCGGAAGAGAAAACTGTAGGCTGGAGCTGCTTCG
knock-loip-R	TGAACGCCTTATCTGACCTACGTTCGACACCACCAGGCTATGGGAATTAGCCATGGTCC
knock-ack-F	CAGGGAGCCATAGAGCGTAGCGCATGATGACACCGACATTGTAGGCTGGAGCTGCTTCG
knock-ack-R	CATGCAAAAAAAATTTGCAGTGCATGATGTTAATCATAAATGGGAATTAGCCATGGTCC
knock-hf-F	AATATAAGGCTCGGCAGAGAAGCGGTATTCAACGTCAACTGTAGGCTGGAGCTGCTTCG
knock-hf-R	AAAATACGCCAGTTTAAGTATCTGCCTGAACTGGCAAGGATGGGAATTAGCCATGGTCC
knock-arnT-F	AGGCTGGCTGGGTTGCCAACAAATTGCGGGTAGTCGCTGTGTAGGCTGGAGCTGCTTCG
knock-arnT-R	CAACGCTAAGCAAGCTGGCAAAGACTAATGTTAGCCAGAATGGGAATTAGCCATGGTCC
knock-tf-F	CAGGATGCAAGAAACCAATTTTTTCATAGAGGTTAACTATGTAGGCTGGAGCTGCTTCG
knock-tf-R	TTGCATACATCTAAATAATAAAATTGCGTCAATAAAAATATGGGAATTAGCCATGGTCC
pro-pcnR-F	CGGAATTCGTAAGTAGCAACCCGGAG
pro-pcnR-R	ACGCGTCGACTTCGGGGCTTTATCCCTT
pro-mcr-1-F	CGGAATTCCAAGATACAAATTATAAATACTCT
pro-mcr-1-R	ACGCGTCGACTCAGCGGATGAATGCGGTGCGGTC
pro-finO-F	CGGAATTCCGGAACGTCTGGCGCTGGTC
pro-finO-R	ACGCGTCGACTTATTTCTCATCAAGCACGGCCTGAA
pcnR-F	GGAATTCCATATGACAGCAATTATTGAGGCGAG
pcnR-R	ACGCGTCGACTTATCCCTTACGTTTTAATTGTTT
AS-RNA-F	GCTAATCACAGCAAGCTGCAC
AS-RNA-R	ACTTCAGGTAGTAAGCGATT

repA-F	CGGAATTCCTGTTTGTGTGGTTTATTCGATCGTG
repA-R	CCCAAGCTTGCCTGCTCTTCCGGTGATAGTTCAGAC
repR-stop-R	TTAGCTTGCACTGTTAGAGTTGATATTT
repR-stop-F	AAATATCAACTCTAACAGTGCAAGCTAA
CA-TT-R	ACTTAGCTTGCACTGTAAGAGTAAATATTTCCTTCTCAT
CA-TT-F	ATGAGAAGGAAATATTTACTCTTACAGTGCAAGCTAAGT
GAAA-CTCT-R	CACTGTAAGAGTTGATAAGAGCTTCTCATGGAATTAAACTCCA
GAAA-CTCT-F	TGGAGTTTAATTCCATGAGAAGCTCTTATCAACTCTTACAGTG
GA-TT-R	TGTAAGAGTTGATATTTCCTAATCATGGAATTAAACTCCA
GA-TT-F	TGGAGTTTAATTCCATGATTAGGAAATATCAACTCTTACA
repR9-R	CCCAAGCTTGCCCTTCTCATGGAATTAAACTCCACT
repR24-R	CCCAAGCTTGCAAACTCCACTTCAGGTAGTAAG
SP1	CTTCTTGACGAGTTCTTCTGAGC
SP2	TTCTGAGCGGGACTCTGGGGTACG
AD1	GGCCACGCGTCGACTAGTACNNNNNNNNNACGCC
AD2	GGCCACGCGTCGACTAGTACNNNNNNNNNCCTGG
AD3	GGCCACGCGTCGACTAGTACNNNNNNNNNCCTCG
RepA-TSS	ATGTTATCAGCTCAAGGCTC
AS RNA-TSS	CGTCTTTTCCAGGGAGCGCAGA
Race-TSS	GATTCGCGTCGACTAGTACGGGGGGGGGG
qrepA-F	TTCTGGCAGGACTACGAA
qrepA-R	GCTACCAGTTGCTCTAATCT
qrpoB-F	TTCACCACCATCCACATTC
qrpoB-R	TTCAGCCAGACGCATAAC
qmcr-1-F	AGCCGCTGCGTAGCTATGTCA
qmcr-1-R	TCACCGACGACGAACACCACTA
GAPDH-F	TCACGCTACTACCGCTACTCAG
GAPDH-R	CGGAACGCCATACCAGTCAGT

[a]Restriction sites are underlined.

Table S4. Open reading frames of pHNSHP45 coding for putative regulators

	Number of amino acids	Predicted function	Pfam-A matches	Sequence identify/similarity with bona-fide regulators	Conserved in IncI2 plasmids
ORF00039	168	CaiF/GrlA transcriptional regulator	CaiF_GrlA (PF07180)	(20.77% / 31%) [a]	Yes
ORF00049 (PcnR)	198	ProQ/FinO family	ProQ (PF04352)	(19.66% / 29%) [b]	Yes
ORF00050	94	Ribbon-helix-helix protein CopG family	RHH_1 (PF01402)	(20% / 29%) [c]	Yes

[a]Amino acid sequences alignment between ORF00039 and GrlA (accession number: EEQ5396029.1). [b]Amino acid sequences alignment between ORF00049 and FinO (WP_115440218.1). [c]Amino acid sequences alignment between ORF00050 and the DNA binding domain of CopASO (accession number: AAN54507.1).

Table S5. Genes disrupted by transposon insertion

Strains	Gene
Tns1	AraC family transcriptional regulator (*gadW*)
Tns2	Glyceraldehyde-3-phosphate dehydrogenase pseudogene (*gapdh*)
Tns3	Metalloprotease LoiP, Phe-Phe periplasmic metalloprotease (*loip*)
Tns4	Acetate kinase promoter (*ack*)
Tns5	Hypothetical protein (4181949-4182155) (*hf*)
Tns6	4-amino-4-deoxy-L-arabinose lipid A transferase (*arnT*)
Tns7	Transcriptional regulator (623991-625367) (*tf*)

Table S6. The putative ProQ/FinO family protein encoded by IncI1 (IncB/O) plasmid

GenBank Locus	Location (plasmid)	Number of amino acids	Predicted function	Pfam-A matches	Sequence identify/similarity with FinO	Sequence identify/similarity with PcnR
QIQ17060.1	79,585-80,187 (pIP72)	200	ProQ/FinO Family	ProQ (PF04352)	(21.37% / 29%)	（29.80% / 50%）
QOC74206.1	4344-4946 (pHNAHS65I-1)	200	ProQ/FinO Family	ProQ (PF04352)	(21.19% / 29%)	（29.80% / 50%）

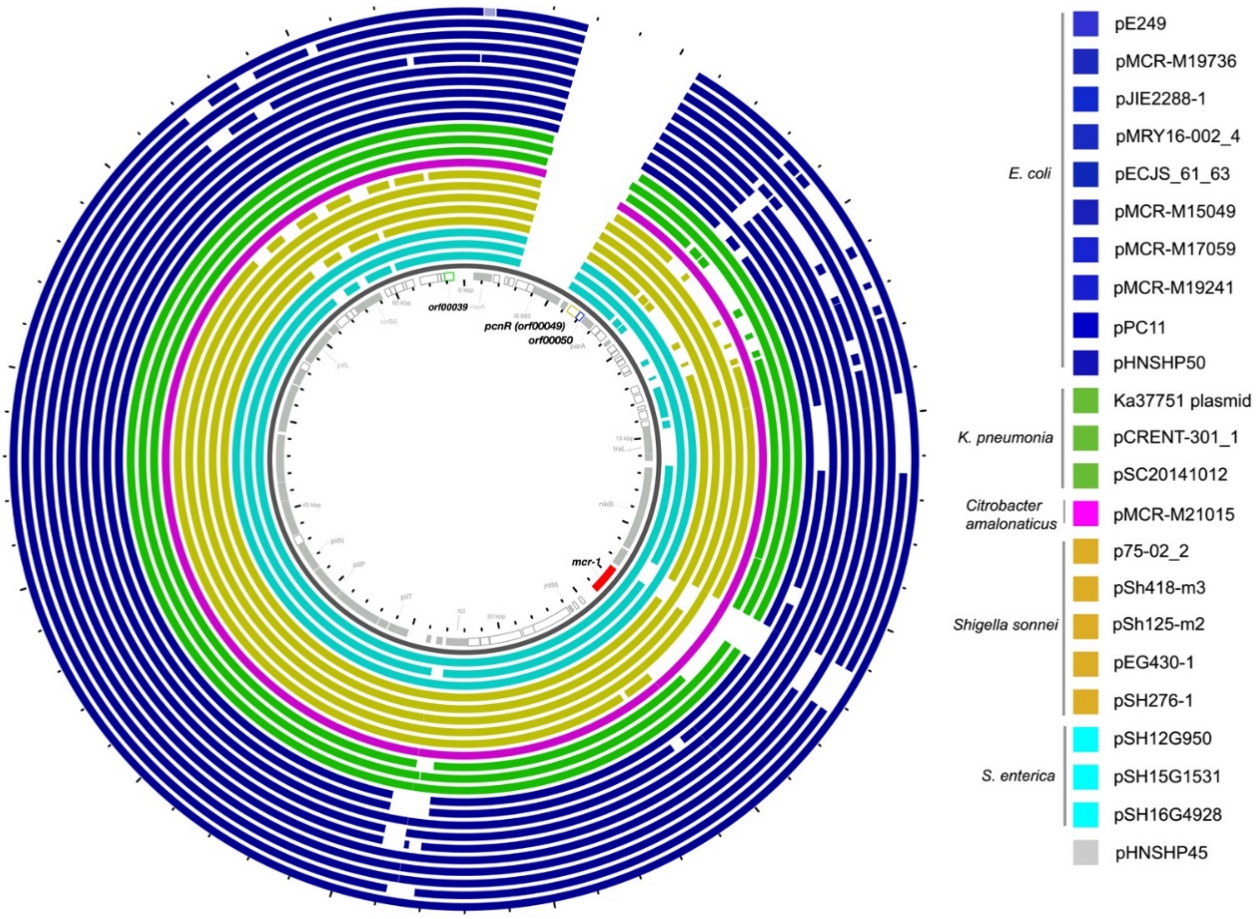

Figure S1. Genetic comparison of pHNSHP45 and other sequenced IncI2 plasmids carrying *mcr-1*. From inside to outside, coding sequences of pHNSHP45 appears on the inner most circle. Genes involved in conjugative transfer and replication are indicated in gray, and genes of unknown function are indicated in white. From the second to fourth circle in cyan are shown *S. enterica*-derived IncI2 plasmids. From the fifth to ninth circle in yellow are shown *Shigella sonnei*-derived IncI2 plasmids. The tenth circle in purple is represented a *Citrobacter amalonaticus*-derived IncI2 plasmid. From the eleventh to thirteenth in green are shown *K. pneumoniae*-derived IncI2 plasmids. The blue circles represent *E. coli* derived IncI2 plasmids. The BLAST Atlas was constructed by Gview server according to their sequence homology toward the reference plasmid pHNSHP45, and all plasmids were available in GenBank.

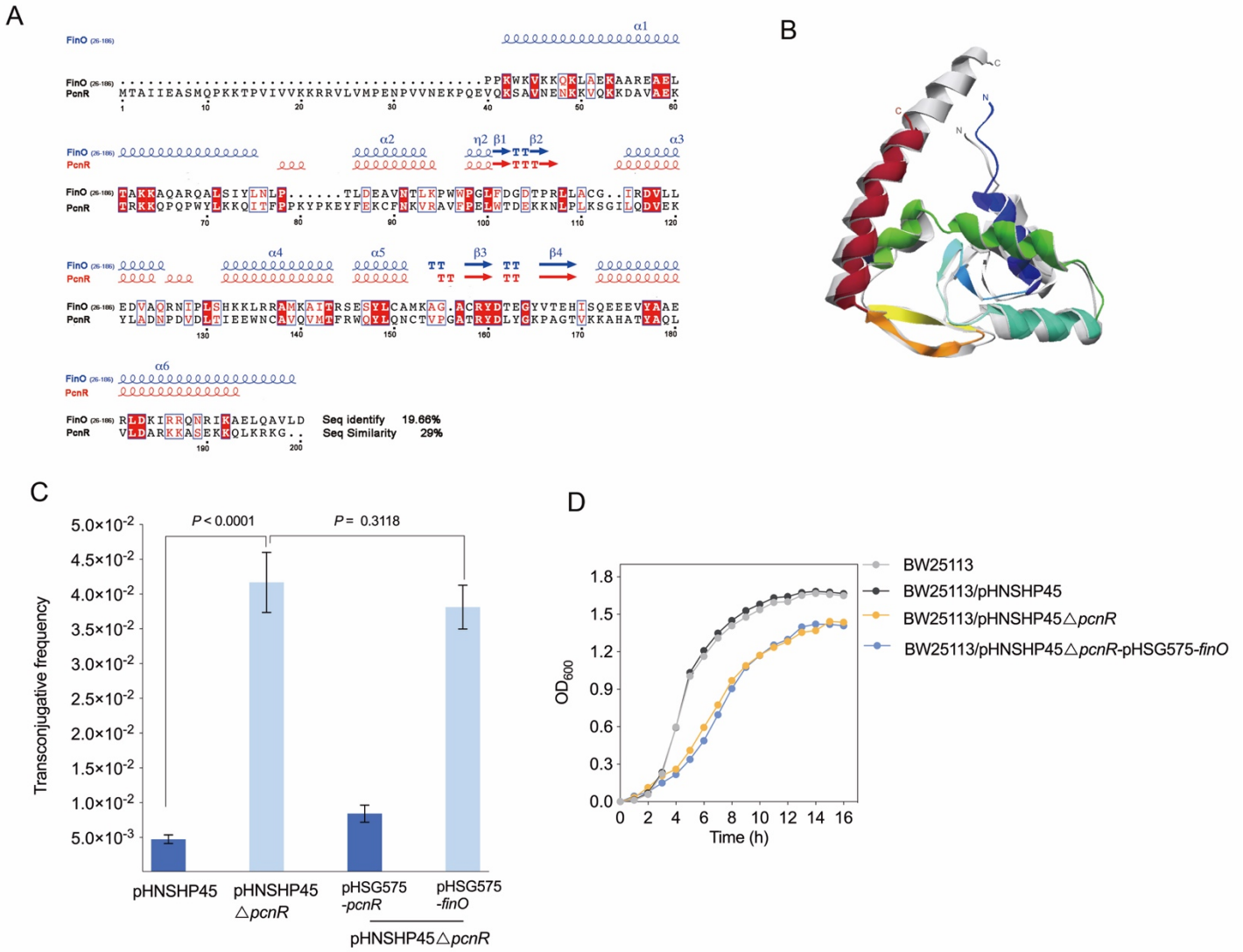

Figure S2. Comparison of the structure and function between PcnR and FinO protein. (**A**) Structure-based sequence alignment of FinO encoded by R6-5 plasmid (accession number: M38048) and PcnR with secondary structures indicated. (**B**) Structure superposition between template model ($FinO_{35-152}$ ID: 1DVO) and predicted structure of $PcnR_{79-190}$. The $PcnR_{79-190}$ structure was predicted by SWISS-MODEL server (https://swissmodel.expasy.org/). The predicted structure of $PcnR_{79-190}$ is shown in color, and the structure of $FinO_{35-152}$ is shown in light grey. The plasmid conjugation transfer frequency (**C**) and growth (**D**) of △*pcnR* mutant could not be complemented by *finO*. Complementation assays were performed by expressing *finO* from its native promoter on pHSG575-*finO*. Error bars represent the SD and *P*-values were calculated by two-tailed *t*-tests.

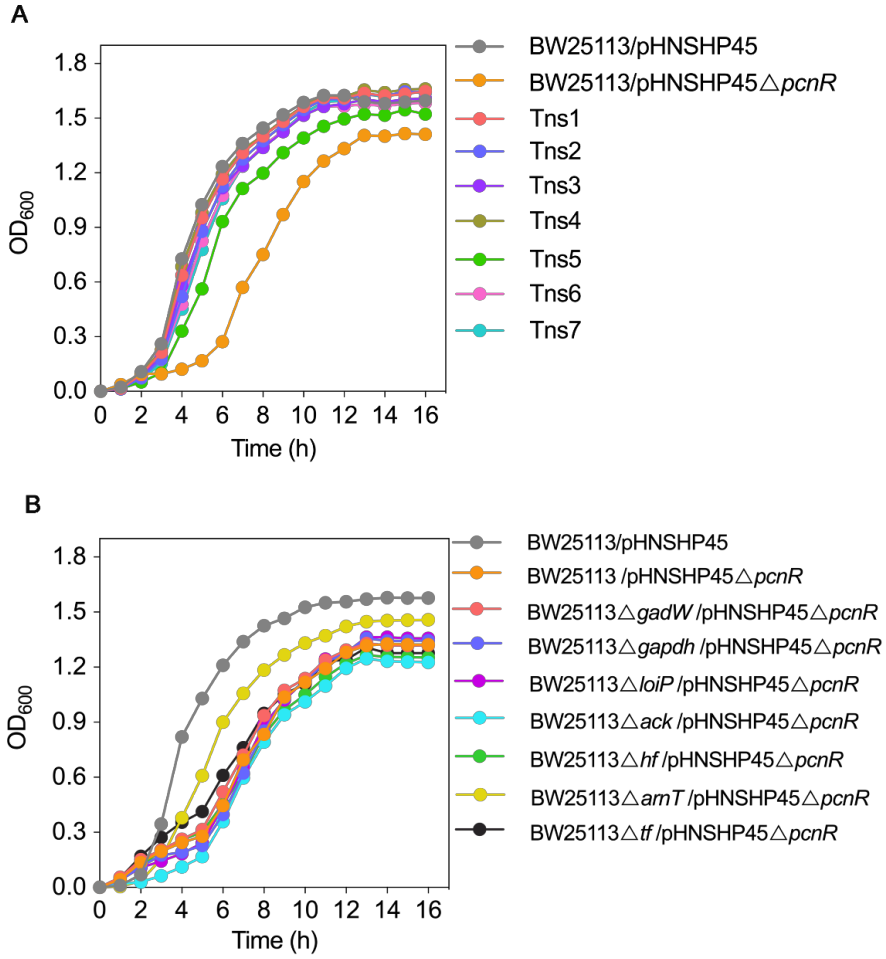

Figure S3. Growth curve of transposon mutants (**A**) and the effect of deletion of genes inserted by transposon on the growth of BW25113/pHNSHP45Δ*pcnR* (**B**). These data represent the mean of three independent experiments.

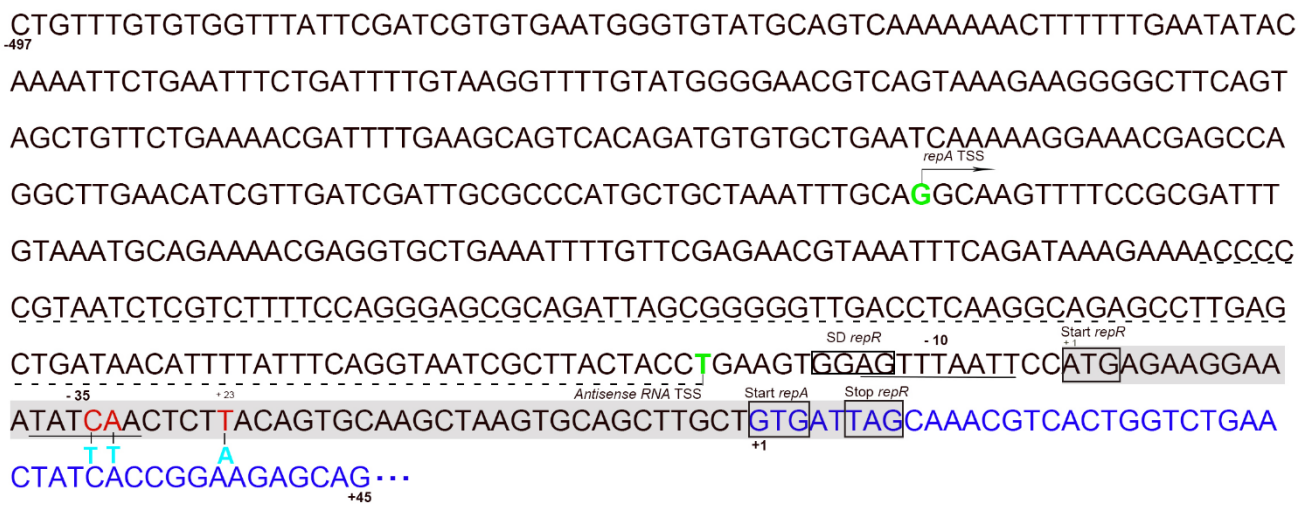

Figure S4. The upstream nucleotide sequence (-497 to +45) of *repA*. This fragment was fused with *lacZ* to monitor the expression of *repA*, and the resulting fusion was named P_{repA}-*lacZ*. The *repA* transcripts as a polycistronic mRNA that contains ORFs encoding RepR and RepA, and the transcription starts site of *repA* is shown in green. The -10 and -35 region of AS RNA are underlined, and the transcription starts site of AS RNA is shown in green. The coding sequence of *repR* was indicated by grey shade. The antisense RNA was indicated by dotted line. Constructing T-A mutation at +23 position of *repR* to introduce a stop codon in *repR* open reading frame, and the resulting mutation was named *repR* (T23A). The -35 region of the promoter of antisense RNA (AS RNA) was mutated (CA-TT) to construct ΔAS RNA mutation.

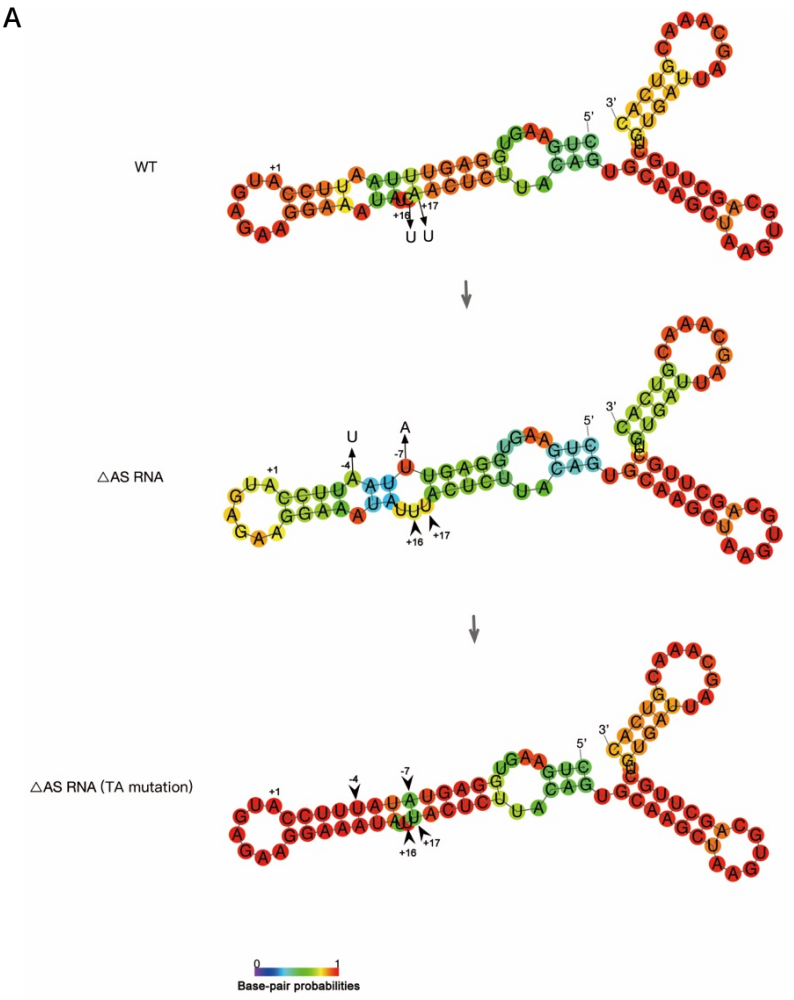

Figure S5. The effect of the first stem-loop structure of *repR* mRNA and AS RNA on the expression of *repA*. (A) Predicted RNA structure of the *repR* mRNA with

mutations. The mutational changes introduced for the present study are shown by arrows. **(B)** Effect of mutations in *repR* mRNA and antisense RNA on the expression of *repA*. The stem structure of *repR* mRNA is predicted to be impaired by the △AS RNA mutation (CA-TT). The TA mutation (T-A/A-T) was reintroduced in the stem structure of *repR* mRNA with the ΔAS RNA mutation, and the resulting mutation was named ΔAS RNA (TA mutation). The leader region (-497 to +45) of *repA* containing these mutations was fused with *lacZ* to monitor the expression of *repA*. Complementation assays were performed by pUC19-AS RNA. Error bars represent the SD and *P*-values were calculated by two-tailed *t*-tests.

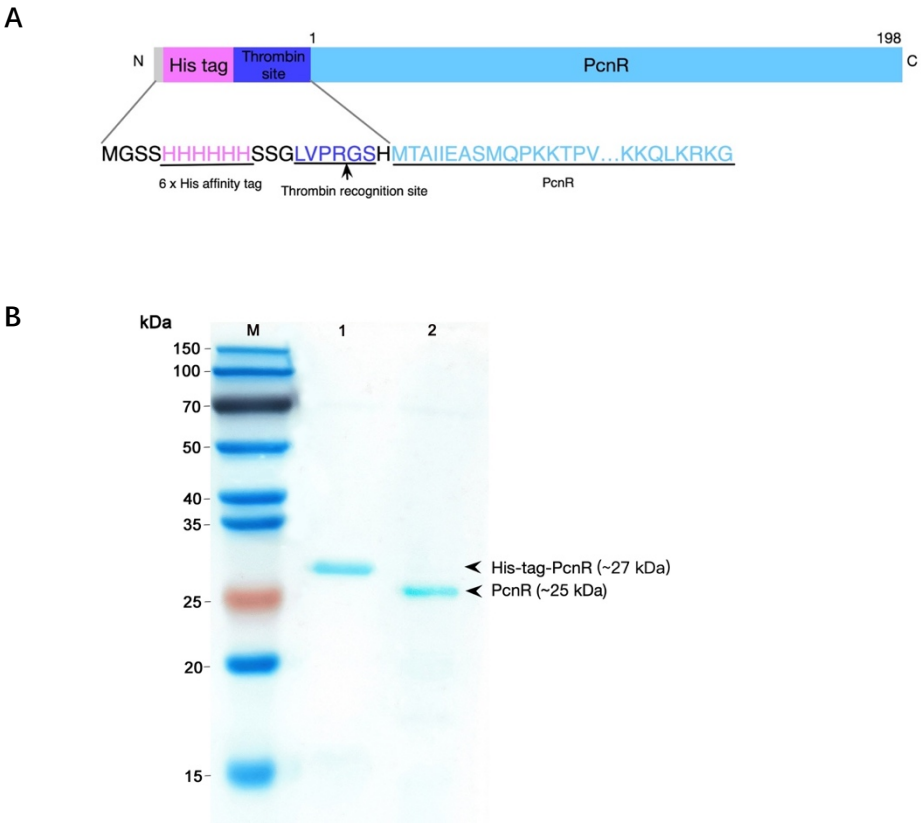

Figure S6. SDS-PAGE analysis of purified PcnR protein. (A) PcnR was fused with a thrombin recognition site and his-tag at N-terminal. (B) SDS-PAGE analysis of the fusion protein His-tag-PcnR before and after thrombin digestion. Lane M: protein maker; Lane 1: Purified his-tag-PcnR; Lane 2: His-tag was cut from PcnR by thrombin.

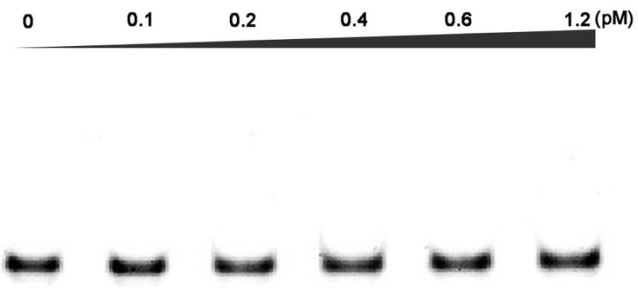

Figure S7. EMSA analysis of interaction between PcnR and antisense RNA. 6×His-tagged PcnR were expressed in *E. coli* BL21(DE3) from pET28b-*pcnR* and purified by Ni-NTA affinity chromatography. Purified His-tag-PcnR was cut by thrombin to remove the His-tag and then purified PcnR was used for RNA-EMSA. The antisense RNA was transcribed by T7 promoter *in vitro*.

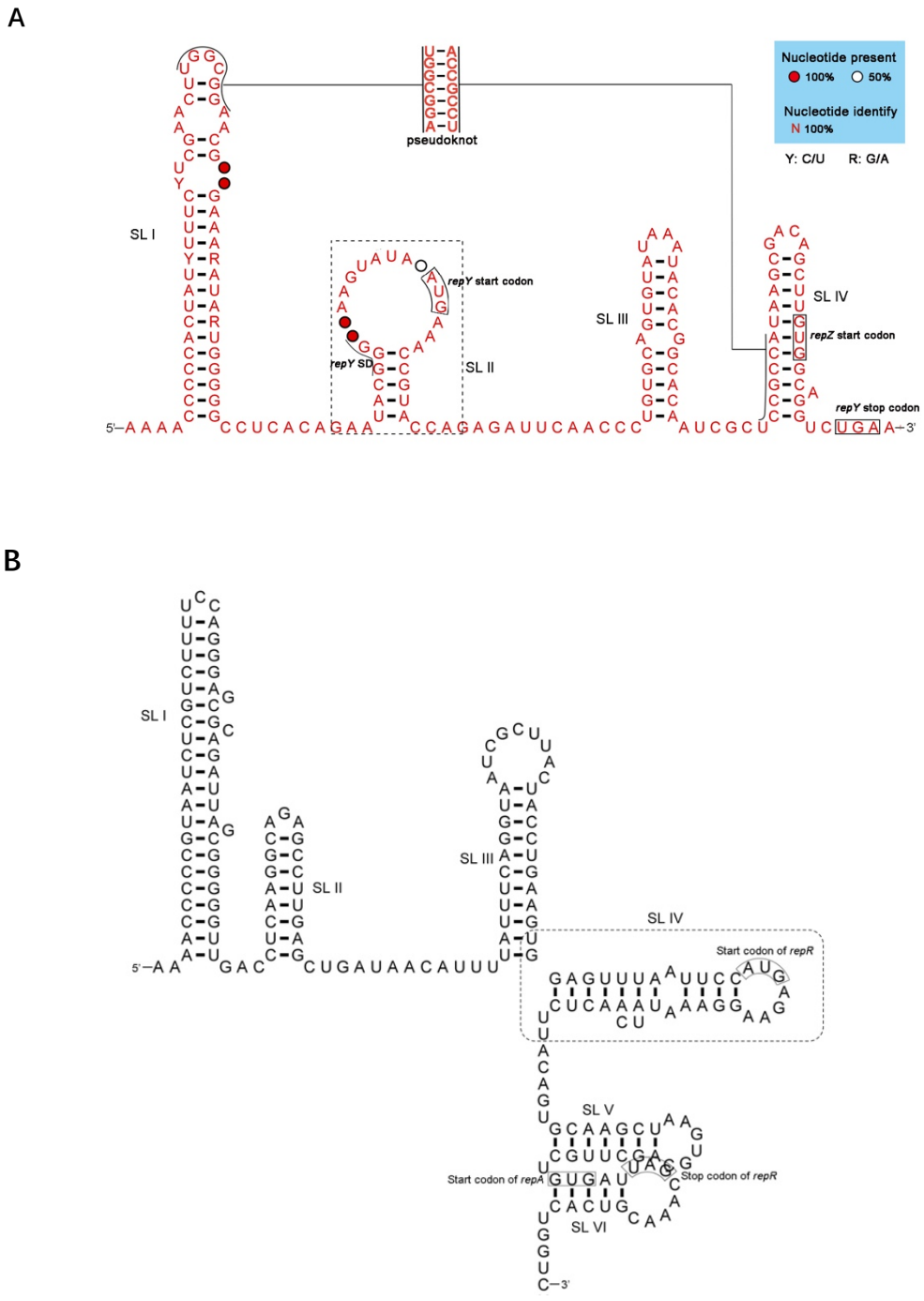

Figure S8. Comparison of the leader region of *repA/Z* mRNA between IncI1/IncB/O and IncI2 plasmids. (A) Consensus structure of the leader region of *repZ* mRNA derived from IncI1 (pHNAHS65I-1, accession number: MN219406) and IncB/O plasmids. (pIP72, accession number: MN612051). This region contains four major secondary structures SLI, SLII (in dashed box), SLIII and SLIV. The pseudoknot sequences are underlined, and the start and stop codons are boxed. The SLI is the target

of IncRNA and the pseudoknot sequence (5'-UGGCGGA-3') (10,11) is present in the loop structure. The SLII (in dashed box) contains the SD sequence and start codon of leader peptide RepY. The SLIII present in the ORF of *repY* and previous studies has shown that this SL is not essential for regulation of *repZ*. The SLIV contains the pseudoknot sequence (5'-ACCGCCU-3') (10,11) and the start codon of *repZ*. The *repY* mRNA codes for the leader peptide, and the *repZ* mRNA codes for replication initiation protein. The consensus structure of mRNA was predicted by LocARNA program (12) and visualized by R-scape program (13). The KineFold program (14) was used to analyze pseudoknots. (B) The structure of the leader region of *repA* mRNA of IncI2 plasmid (pHNSHP45, accession number: KP347127). This region contains six major secondary structures SLI, SLII, SLIII , SLIV(in dashed box), SLV, and SL VI. The SLI, SLII and SLIII are the target of antisense RNA. The *repR* mRNA codes for the leader peptide, and the *repA* mRNA codes for replication initiation protein. The start and stop codons are boxed. The mRNA structure was generated by RNAfold program.

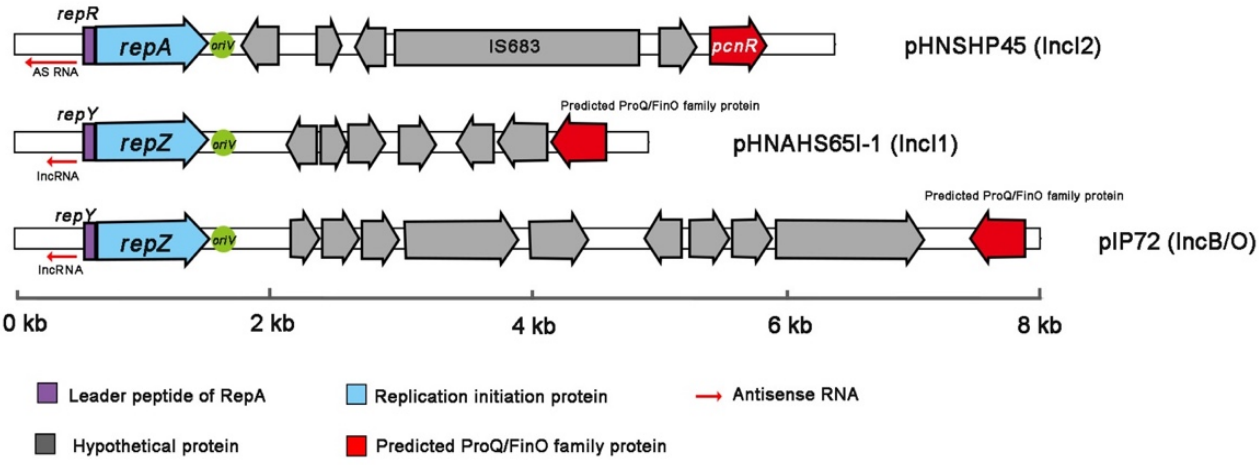

Figure S9. Comparison of IncI2 plasmid replication region with IncI1 and IncB/O plasmids. The ProQ/FinO family proteins were predicted by Pfam. Sequences with the following GenBank accession numbers were used: pHNSHP45, KP347127; pHNAHS65I-1, MN219406; pIP72, MN612051.

References:

1. Datsenko, K.A. and Wanner, B.L. (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A*, 97, 6640-6645.
2. Miller, V.L. and Mekalanos, J.J. (1988) A novel suicide vector and its use in construction of insertion mutations: osmoregulation of outer membrane proteins and virulence determinants in *Vibrio cholerae* requires toxR. *Journal of Bacteriology*, 170, 2575-2583.
3. Liu, Y.-Y., Wang, Y., Walsh, T.R., Yi, L.-X., Zhang, R., Spencer, J., Doi, Y., Tian, G., Dong, B., Huang, X. *et al.* (2016) Emergence of plasmid-mediated colistin resistance mechanism MCR-1 in animals and human beings in China: a microbiological and molecular biological study. *The Lancet Infectious Diseases*, 16, 161-168.
4. Lv, L., Partridge, S.R., He, L., Zeng, Z., He, D., Ye, J. and Liu, J.-H. (2013) Genetic Characterization of IncI2 Plasmids Carrying $bla_{\text{CTX-M-55}}$ Spreading in both Pets and Food Animals in China. *Antimicrobial Agents and Chemotherapy*, 57, 2824-2827.
5. Takeshita, S., Sato, M., Toba, M., Masahashi, W. and Hashimoto-Gotoh, T. (1987) High-copy-number and low-copy-number plasmid vectors for lacZα-complementation and chloramphenicol- or kanamycin-resistance selection. *Gene*, 61, 63-74.
6. Zhu, L., Lin, J., Ma, J., Cronan, J.E. and Wang, H. (2010) Triclosan resistance of *Pseudomonas aeruginosa* PAO1 is due to FabV, a triclosan-resistant enoyl-acyl carrier protein reductase. *Antimicrobial agents and chemotherapy*, 54, 689-698.
7. He, L., Partridge, S.R., Yang, X., Hou, J., Deng, Y., Yao, Q., Zeng, Z., Chen, Z. and Liu, J.-H. (2012) Complete nucleotide sequence of pHN7A8, an F33:A-:B-type epidemic plasmid carrying $bla_{\text{CTX-M-65}}$, *fosA3* and *rmtB* from China. *Journal of Antimicrobial Chemotherapy*, 68, 46-50.
8. Chiang, S.L. and Rubin, E.J. (2002) Construction of a mariner-based transposon for epitope-tagging and genomic targeting. *Gene*, 296, 179-185.
9. Fu, H., Jin, M., Ju, L., Mao, Y. and Gao, H. (2014) Evidence for function overlapping of CymA and the cytochromebc1complex in the *Shewanella oneidensis* nitrate and nitrite respiration. *Environmental Microbiology*, 16, 3181-3195.
10. Asano, K. and Mizobuchi, K. (1998) An RNA pseudoknot as the molecular. switch. for

translation of the *repZ* gene encoding the replication initiator of IncIα plasmid ColIb-P9. *The Journal of biological chemistry*, 273, 11815-11825.

11. Praszkier, J. and Pittard, A.J. (2005) Control of replication in I-complex plasmids. *Plasmid*, 53, 97-112.

12. Will, S., Joshi, T., Hofacker, I.L., Stadler, P.F. and Backofen, R. (2012) LocARNA-P: accurate boundary prediction and improved detection of structural RNAs. *RNA*, **18**, 900-914.

13. Rivas, E., Clements, J. and Eddy, S.R. (2017) A statistical test for conserved RNA structure shows lack of evidence for structure in lncRNAs. *Nat Methods*, **14**, 45-48.

14. Xayaphoummine, A., Bucher, T. and Isambert, H. (2005) Kinefold web server for RNA/DNA folding path and structure prediction including pseudoknots and knots. *Nucleic Acids Res*, **33**, W605-610.