Essentiality of threonylcarbamoyladenosine (t^6A), a universal tRNA modification, in bacteria

Patrick C. Thiaville[1,2,3,4$], Basma El Yacoubi[1$], Caroline Köhrer[5], Jennifer J. Thiaville[1], Chris Deutsch[6], Dirk Iwata-Reuyl[6], Jo Marie Bacusmo[1], Jean Armengaud[7], Yoshitaka Bessho[8,9], Collin Wetzel[10], Xiaoyu Cao[10], Patrick A. Limbach[10], Uttam L. RajBhandary[5], and Valérie de Crécy-Lagard[1,3]*

[1]Department of Microbiology and Cell Science, University of Florida, Gainesville, FL 32611, USA

[2]Genetics and Genomics Graduate Program, University of Florida, Gainesville, FL 32610, USA

[3]University of Florida Genetics Institute, University of Florida, Gainesville, FL 32610, USA

[4]Institut de Génétique et Microbiologie, Université of Paris-Sud, Orsay, France

[5]Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

[6]Department of Chemistry, Portland State University, Portland, Oregon, 97297, USA

[7]CEA, DSV, IBiTec-S, SPI, Li2D, Laboratory "Innovative technologies for Detection and Diagnostics", Bagnols-sur-Cèze, F-30200, France

[8]RIKEN SPring-8 Center, 1-1-1 Kouto, Sayo, Hyogo 679-5148, Japan

[9]Institute of Physics, Academia Sinica, 128 Sec. 2, Academia Rd., Nankang, Taipei 11529, Taiwan

[10]Rieveschl Laboratories for Mass Spectrometry, Department of Chemistry, University of Cincinnati, Cincinnati, Ohio 45221, United States

*To whom correspondence should be addressed. Valérie de Crécy-Lagard, Department of Microbiology and Cell Science, University of Florida, P.O. Box 110700, Gainesville, FL 32611-0700; E-mail : vcrecy@ufl.edu, Tel: (352) 392 9416; Fax: (352) 392 5922

[$] Contributed equally.

Running title: t^{6}A essentiality

Supplemental Results

Sequencing of E. coli derivatives

To further validate our genetic dissection experiments, the P_{TET}::*tsaD* (VDC7062) and P_{TET}::*tsaC* (VDC7066) strains as well as their parental strain *E. coli* B were fully sequenced to identify point mutations as described below. The reference genome closest to *E. coli* B is *E. coli* B Strain REL606. *E. coli* B is a direct descendant of *Bacillus coli* used by d'Herelle in 1918 and was first described by Luria in 1945. REL606 is derived from B but with 45 years of additional manipulations and was first described by Lenski in 1985 (Daegelen *et al.*, 2009). This would explain the 528 variations (single nucleotide polymorphisms (SNPs), insertions or deletions), we found between B and REL606 (Dataset 1). Although the P_{TET}::*tsaD* (VDC7062) and P_{TET}::*tsaC* (VDC7066) have mostly the same sequence as the B strain, a few SNPs found in REL606 compared to B, are also found in the these two strains, all in phage or hypothetical proteins. In addition, P_{TET}::*tsaD* (VDC7062) and P_{TET}::*tsaC* (VDC7066) each contained additional mutations (54 and 305, respectively) and most of these could be traced back to the 90 kb window from P1 transduction from the donor BW25113 strains.

A few mutations unique to the two strains were found. As predicted, the P_{TET}::*tsaC* (VDC7066) *tetR* allele of the P_{TET} cassette contains a T1173G change, creating a non-synonymous Q38P mutation, in the TetR DNA binding site. In addition, a C84G mutation that leads to a D28E change of the cold shock protein CspC protein (ECB_01793) was identified. Loss of CspC function has been shown to confer a competitive fitness advantage over isogenic wild-type strains (Rath & Jawali, 2006). The D28E change occurs adjacent to the RNA binding domain of CspC.

P_{TET}::*tsaD* (VDC7062) contains three unique variations: a deletion (T11-) in a hypothetical protein not predicted in the RefSeq, a silent A2838G mutation in RshA (ECB_03448), and a A1810C mutation leading to the T605P mutation in YhfK (ECB_03209, conserved hypothetical inner membrane protein).

t^6A essentiality suppression tests

The P_{TET}::*tsaC* and P_{TET}::*tsaD* strains (VDC5684 and VDC5801, respectively) were transformed with plasmids carrying *ileS* from yeast (pBY324.10) or *Bacillus subtilis* (pBY326.6) alone or in combination with a compatible plasmid carrying *Mycoplasma mobile* $tRNA^{Ile3}_{UAU}$ and *ileS* gene pair (pPCT080) (*Mm-Pair).* Transformants were selected in absence of aTc and in the presence or absence of inducer for expression of IleRS transgenes. None of the transformants displayed growth in the absence of aTc (Figure S3). As in Figure 4B, the complementing P_{BAD}::*tsaC* plasmid allows for growth of the P_{TET}::*tsaC* strain when neither inducer is used (Figure S3B), but, in repeated attempts, the addition of arabinose inducing expression of P_{BAD}::*tsaC* was toxic in combination with the constitutive expression of the *Mm-Pair* (Figure S3B). As noted by Bohlke and Budisa, expression of the *Mm-Pair* causes a growth defect in *E. coli*, possibly due to the misreading of AUG codons by the *Mm-Pair* (Bohlke & Budisa, 2014). It seems that high levels of t^6A modified tRNAs is incompatible with the *Mm-Pair*.

Analyses of codon choice and codon stretches in proteins that are downregulated in the D. radiodurans t^6A^- strain

Only two amino acids are encoded by both t^6A dependent and t^6A independent codons, Ser and Arg. $tRNA^{Arg}_{CCU}$ and $tRNA^{Arg}_{UCU}$ read AGG and AGA codons, respectively, and contain t^6A, whereas $tRNA^{Arg}_{ACG}$ and $tRNA^{Arg}_{CCG}$ that decode CGN codons do not (Jühling *et al.*, 2009). For $tRNA^{Ser}_{GCU}$ that decodes AGY, it depends on the organism: t^6A is found at position 37 in *E.*

coli, whereas i^6A is found in yeast (Jühling *et al.*, 2009). $tRNA^{Ser}_{GGA}$, $tRNA^{Ser}_{CGA}$, and $tRNA^{Ser}_{UGA}$ that decode UCN never contain t^6A (Jühling *et al.*, 2009). MS/MS analysis of *D. radiodurans* and *T. thermophilus* bulk tRNA confirmed that t^6A is present at position 37 of $tRNA^{Ser}_{GCU}$ (data not shown). Proteins that exclusively use the t^6A dependent Ser codons (AGY) or Arg codons (AGR) were compiled (Table S6 and S7 respectively). Four of these proteins were significantly reduced in the t^6A^- strain: DR_0265 (a regulator of N-acetyl glucosamine utilization), DR_0548 (Type IV pilin PilA), DR_0312 [(LSU ribosomal protein L4p (L1e)] and the tyrosine degradation protein DR_1609 (predicted 2-keto-4-pentenoate hydratase/2-oxohepta-3-ene-1,7-dioic acid hydratase) (Table S4 and S5).

The percentage of t^6A codons in stretches of two or more is in average 2.0% per protein for the whole theoretical proteome (Table S8) and 2.8% for both the mid and late-log detected proteome (Table S4 and S5). The average for the proteins that are down regulated is 3.4% for mid-log and 3.0% for late-log experiments (Table S4 and S5). The averages for proteins that are up-regulated are 3.0% for both mid and late-log experiments (Table S4 and S5). This suggests that at least for the mid-log experiments, down-regulated proteins show an enrichment of t^6A dependent codon stretches. Finally, proteins with the longest stretches of t^6A codons were ranked and 23 proteins were found to have stretches of t^6A codons longer or equal to seven (Table S8), of these two have significantly lower expression levels in the t^6A deficient strain (Table S4 and S5): DR_1046 (ClpB protein) and DR_2263 [(non-specific DNA-binding protein Dps/Iron-binding ferritin-like antioxidant protein/Ferroxidase (EC 1.16.3.1)].

Supplemental methods

Genomic DNA preparation

E. coli DNA was purified using the Qiagen DNeasy Blood and Tissue Kit following the manufacturer's instruction for isolation of DNA from gram negative bacteria. DNA from *S. mutans* was purified using the following protocol: five mL of cultures grown overnight, standing at 37 ˚C with 5% CO_2 was washed once with TE (pH 8.0) and suspended in 250 µL of TE and then heated at 60 ˚C for 20 minutes. Cells were chilled on ice, and then five µL of mutanolysin (Sigma, 5000 U/mL) and 10 µL of lysozyme (100 mg/mL) were added and incubated at 37 ˚C for 60 minutes. 2.5 µL of proteinase K (10 mg/mL stock solution) was added and incubation continued for 60 minutes at 37 ˚C with rotation. Next, 1.0 µL of RNaseA (Qiagen 100 mg/mL) and 75 µL of 10% SDS were added, and the tubes were gently inverted. 7.3 µL of 4M $NaCl_2$ was added and mixed, followed by 41 µL of 0.1M $NaCl_2$ / 10% CTAB. Samples were vortexed vigorously, and incubated at 65 ˚C for 20 minutes, prior to extraction with 450 µL chloroform:isoamyl alcohol (24:1). Tubes were centrifuged for seven minutes at top speed in microcentrifuge. The aqueous layer was extracted two more times with an equal volume of phenol:chloroform:isoamyl (1:1) and then with a volume of chloroform:isoamyl. The aqueous layer was subjected to ethanol precipitation with two volumes of ice-cold 95% ethanol and centrifuge for 10 minutes at top speed in microcentrifuge. Pellets were washed with 70% ETOH. The dried DNA were suspended in 100 µL TE. DNA quality was analyzed using Picogreen (Thermo) and Nanodrop.

Whole genome sequencing

Sequencing of genomic DNA from *S. mutans* and *E. coli* was performed by Beckman Genomics (Danvers, MA). Libraries were constructed using the Illumina TruSeq Paired End adaptors and Beckman Coulter SPRI works Library Construction using SPRI-TE instrumentation. All libraries passed QC for library size and concentration (>2nM). Libraries were sequenced on an Illumina HiSeq 2500 instrument at 2 x 100 bp using V4 chemistry. Quality of the sequencing run was determined using FastQC (www.bioinformatics.babraham.ac.uk/projects/fastqc/). Reads were processed using Illumina's bcl2fastq version 1.8.4 and fastq_illumina_filter-0.1. Fastq format validation and statistic reporting was done by the fastq_illumina_filter. Read alignment and SNP calling were performed using CLC Genomics Workbench 6.5 (http://www.clcbio.com) using the default settings to *E. coli* Strain B REL606 (NCBI TaxID: 37762, 413997.3.PATRIC.ffn) or *S. mutants* UA159 (NCBI TaxID: 21007, 210007.7.PATRIC.ffn) downloaded from the PatricBRC database (www.patricbrc.org). An average coverage of 1800X was obtained for the *E. coli* genomes and of 5000X for *S. mutans*. Sequences were deposited in the NCBI Short read Archive (SRA) under SRX1144572.

Plasmid construction

The IleRS gene from *S. cerevisiae* was PCR amplified from a cDNA clone (Openbiosystems) as *Pci*I/*Xho*I fragment, introducing a C-terminal His6 tag via the reverse primer, and inserted into pET15b cut with *Nco*I/*Xho*I. The original His6 tag from the vector was removed using this cloning strategy The IleRS genes from both *Saccharomyces cerevisiae* and *Bacillus subtilis* were subcloned from plasmids pET.ScIleRS and pET.BsIleRS. pET.ScIleRS was digested with *Sph*I and *Hin*dIII and the fragment containing the IleRS gene (*ileS*) was ligated into pBAD18 cut with *Sph*I and *Hin*dIII creating pBY324.10. pET.BsIleRS was digested with *Xba*I and *Xho*I and the fragment containing *ileS* was ligated into pBAD18 cut with *Xba*I and

*Sal*I creating pBY326.6. *Mycobacterium tuberculosis ileS* was amplified by PCR from *M. tuberculosis* H37Ra (ATCC 2577D-5) genomic DNA using oligos BAD33IleRSMtB.ol1 and BAD33IleRSMtB.ol2. The PCR fragment and pBAD33 were digested with *Xba*I and *Hin*dIII before ligation creating pBY227.1. All three inserts were verified by sequencing using plasmid specific primers BADOl5 and BADRev2. pPCT080 was created by disrupting the kanamycin resistance cassette of p*Mm*Pair with chloramphenicol resistance cassette from pKD3.

References

Allali-Hassani, A., T.L. Campbell, A. Ho, J.W. Schertzer & E.D. Brown, (2004) Probing the active site of YjeE: a vital Escherichia coli protein of unknown function. *Biochem J* **384**: 577-584.

Baba, T., T. Ara, M. Hasegawa, Y. Takai, Y. Okumura, M. Baba, K.A. Datsenko, M. Tomita, B.L. Wanner & H. Mori, (2006) Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol* **2**: 2006 0008.

Bitoun, J.P., S. Liao, G.G. Xie, W.L. Beatty & Z.T. Wen, (2014) Deficiency of BrpB causes major defects in cell division, stress responses and biofilm formation by *Streptococcus mutans*. *Microbiology* **160**: 67-78.

Bohlke, N. & N. Budisa, (2014) Sense codon emancipation for proteome-wide incorporation of noncanonical amino acids: rare isoleucine codon AUA as a target for genetic code expansion. *FEMS Microbiol Lett* **351**: 133-144.

Brosius, J. & A. Holy, (1984) Regulation of ribosomal RNA promoters with a synthetic lac operator. *Proc Natl Acad Sci U S A* **81**: 6929-6933.

Campbell, T.L., C.S. Ederer, A. Allali-Hassani & E.D. Brown, (2007) Isolation of the *rstA* gene as a multicopy suppressor of YjeE, an essential ATPase of unknown function in *Escherichia coli*. *J Bacteriol* **189**: 3318-3321.

Daegelen, P., F.W. Studier, R.E. Lenski, S. Cure & J.F. Kim, (2009) Tracing ancestors and relatives of *Escherichia col*i B, and the derivation of B strains REL606 and BL21(DE3). *J Mol Biol* **394**: 634-643.

El Yacoubi, B., I. Hatin, C. Deutsch, T. Kahveci, J.P. Rousset, D. Iwata-Reuyl, A.G. Murzin & V. de Crécy Lagard, (2011) A role for the universal Kae1/Qri7/YgjD (COG0533) family in tRNA modification. *Embo J* **30**: 882-893.

El Yacoubi, B., B. Lyons, Y. Cruz, R. Reddy, B. Nordin, F. Agnelli, J.R. Williamson, P. Schimmel, M.A. Swairjo & V. de Crecy-Lagard, (2009) The universal YrdC/Sua5 family is required for the formation of threonylcarbamoyladenosine in tRNA. *Nucleic Acids Res* **37**: 2894-2909.

Gerdes, S., B. El Yacoubi, M. Bailly, I.K. Blaby, C.E. Blaby-Haas, L. Jeanguenin, A. Lara-Núñez, A. Pribat, J.C. Waller, A. Wilke, R. Overbeek, A.D. Hanson & V. de Crécy

Lagard, (2011) Synergistic use of plant-prokaryote comparative genomics for functional annotations. *BMC Genomics* **12**: S2-S2.

Guzman, L.M., D. Belin, M.J. Carson & J. Beckwith, (1995) Tight regulation, modulation, and high-level expression by vectors containing the arabinose P_{BAD} promoter. *J Bacteriol* **177**: 4121-4130.

Handford, J.I., B. Ize, G. Buchanan, G.P. Butland, J. Greenblatt, A. Emili & T. Palmer, (2009) Conserved network of proteins essential for bacterial viability. *J Bacteriol* **191**: 4732-4749.

Jühling, F., M. Mörl, R.K. Hartmann, M. Sprinzl, P.F. Stadler & J. Pütz, (2009) tRNAdb 2009: compilation of tRNA sequences and tRNA genes. *Nucleic Acids Res* **37**: D159-162.

Kaczanowska, M. & M. Ryden-Aulin, (2004) Temperature sensitivity caused by mutant Release Factor 1 is suppressed by mutations that affect 16S rRNA Maturation. *J Bacteriol* **186**: 3046-3055.

Kitagawa, M., T. Ara, M. Arifuzzaman, T. Ioka-Nakamichi, E. Inamoto, H. Toyonaga & H. Mori, (2006) 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.

Onodera, T., K. Satoh, T. Ohta & I. Narumi, (2013) *Deinococcus radiodurans* YgjD and YeaZ are involved in the repair of DNA cross-links. *Extremophiles* **17**: 171-179.

Oshima, T. & K. Imahori, (1974) Physiochemical properties of deoxyribonucleic acid from an extreme thermophile. *J Biochem* **75**: 179-183.

Rath, D. & N. Jawali, (2006) Loss of expression of *cspC*, a cold shock family gene, confers a gain of fitness in *Escherichia coli* K-12 strains. *J Bacteriol* **188**: 6780-6785.

Zuther, E., H. Schubert & M. Hagemann, (1998) Mutation of a gene encoding a putative glycoprotease leads to reduced salt tolerance, altered pigmentation, and cyanophycin accumulation in the cyanobacterium *Synechocystis* sp. strain PCC 6803. *J Bacteriol* **180**: 1715-1722.

Supplemental Tables

Table S1: List of strains used in this study

Strain	Description	Reference
E. coli		
C600	F-, *thr-1*, *leuB6*(Am), *fhuA21*, *cyn-101*?, *lacY1*, *glnX44*(AS), λ^-, *rfbC1*, *glpR200*(glp^c), *thiE1*	CGSC
E. coli B	F-, *Ion-*, *dcm-*, *malB*	CGSC
BW25113	F-, Δ*(araD-araB)567*, Δ*lacZ4787*(::rrnB-3), λ^-, *rph-1*, Δ*(rhaD-rhaB)568*, *hsdR514*	(Baba *et al.*, 2006)
MG1655	F- lambda- *ilvG- rfb-50 rph-1*	CGSC
VDC5684	BW25113 P_{TET}::*tsaC*::*aph*	(Gerdes *et al.*)
VDC5801	BW25113 P_{TET}::*tsaD*::*aph*	(El Yacoubi *et al.*, 2011)
VDC7061/62	*E. coli* B P_{TET}::*tsaD*::*aph*	This study
VDC7065/66	*E. coli* B P_{TET}::*tsaC*::*aph*	This study
VDC7069/70	C600 P_{TET}::*tsaD*::*aph*	This study
VDC7073/74	C600 P_{TET}::*tsaC*::*aph*	This study
US475	*ara* Δ*(gpt-lac)5, zcg-174::Tn10, gyrA(nalA), arg*(UAG)*, rpoB(rif), thi*	(Kaczanowska & Ryden-Aulin, 2004).
US477	US475, *prfA1 zcg-174::Tn10*	(Kaczanowska & Ryden-Aulin, 2004).
MRA100	US477, Δ*yrdC*	(Kaczanowska & Ryden-Aulin, 2004).
JH19	MC4100-A Δ*tsaBattB*::(*araC*+P_{BAD}:*tsaB*+); Kn^r	(Handford *et al.*, 2009)
JH22	MC4100-A *ΔtsaDattB*::(*araC*+P_{BAD}:*tsaD*+); Kn^r	(Handford *et al.*, 2009)
EB444.6	MG1655 *araBAD*::*tsaE*::*aph* Δ*tsaE::cat;* Kn^r*,* Cm^r	(Allali-Hassani *et al.*, 2004, Campbell *et al.*, 2007)
VDC9602	MG1655 (pKK223-3); Ap^r	This study
VDC9603	MG1655 (pRstA); Ap^r	This study
VDC9604	JH19 (pKK223-3); Kn^r, Ap^r	This study
VDC9605	JH19 (pRstA); Kn^r, Ap^r	This study
VDC9606	JH22 (pKK223-3); Kn^r, Ap^r	This study
VDC9607	JH22 (pRstA); Kn^r, Ap^r	This study
VDC9608	EB444.6 (pKK223-3); Kn^r, Cm^r, Ap^r	This study
VDC9609	EB444.6 (pRstA); Kn^r, Cm^r, Ap^r	This study
D. radiodurans		

R1	Wild-type	(Onodera *et al.*, 2013)
XYD	R1, Δ*tsaD*	(Onodera *et al.*, 2013)
XYZ	R1, Δ*tsaB*	(Onodera *et al.*, 2013)
WDZ	R1, Δ*tsaD*, Δ*tsaD*	(Onodera *et al.*, 2013)
S. mutans		
UA159	Wild-type	(Bitoun *et al.*, 2014)
JB409	UA159, Δ*tsaE*	(Bitoun *et al.*, 2014)
JB409c	JB409(pTsaE)	(Bitoun *et al.*, 2014)
Synechocystis sp.		
PCC 6803	Wild-type	(Zuther *et al.*, 1998)
gcp	PCC 6803, *tsaD*::*aph*II	(Zuther *et al.*, 1998)
T. thermophilus		
HB8	Wild-type	(Oshima & Imahori, 1974)
ΔTtTsaC	HB8, Δ*tsaC*::*HTK*	This study

Table S2: Plasmids used in this study

Plasmid Name	Description	Reference
pBAD24	Ap^R, ColE1, arabinose inducible promoter	(Guzman *et al.*)
pBAD33	Cm^R, p15A, arabinose inducible promoter	(Guzman *et al.*)
pBAD18	Ap^R, ColE1, arabinose inducible promoter	(Guzman *et al.*)
pCA24N	Cm^R; *lacI* ^{q}	(Kitagawa *et al.*, 2006)
$pRstA_{ASKA}$	Cm^R; *lacI* ^{q}, $P_{T5\text{-}lac}$::*rstA*	(Kitagawa *et al.*, 2006)
pKK223-3	P_{tac}; Ap^r	(Brosius & Holy, 1984)
pBF9-*rstA*	*rstA* gene from *E. coli* cloned in pBF9 (derivative of pKK223-3); rstA driven by P_{tac}; Ap^r	(Campbell *et al.*, 2007)
p*ywlC*	pBAD24::*ywlC* from *B. subtilis*	(El Yacoubi *et al.*, 2009)
pTsaD	pBAD::*tsaD* from *E. coli*	(El Yacoubi *et al.*, 2011)
pNB26´2_*Mm*Pair	*Mycoplasma mobile* □□□□$^{□□□}_{□□□}$^{3} and *ileS* into pNB26´2 (Kan^R, p15A, P_{LAC} promoter)	(Bohlke & Budisa, 2014)
pBY227.1	pBAD33 carrying *ileS* from *Mycobacterium tuberculosis*	This study
pET.ScIleRS	*S. cerevisiae* was PCR amplified from a cDNA clone (Openbiosystems) as *Pci*I/*Xho*I fragment, introducing a C-terminal His6 tag via the reverse primer, and inserted into pET15b cut with *Nco*I/*Xho*I.	This study

pET.BsIleRS	*ileS* from *B. subtilis* and ligated into pET15b	This study
pBY324.10	pBAD18 carrying *ileS* from *Saccharomyces cerevisiae*	This study
pBY326.6	pBAD18 carrying *ileS* from *Bacillus subtilis*	This study
pPCT080	pNB26’2_*Mm*Pair Klenow treated *Pvu*I digest ligated with Klenow treated *Xba*I digested *cat* from pKD3; Cm^{r}	This study

Table S3: Oligonucleotides used in the study

Primer name	Sequence
Dr_tRNA31-met_fwd	GATATTGATGAACGTGCCGGACAG
Dr_tRNA31-met_rev	GAAACCTACCGCGAGGTCAT
Dr_tRNA11-ile_fwd	TAAGGGGTCAAGGCGAAACTTAG
Dr_tRNA11-ile_rev	GACGGAGTGGAAGACAGCTC
Dr_tRNA41-met_fwd	CAAGCTGATGGACGGCTAAAAAG
Dr_tRNA41-met_rev	GAACGTCCTGAGTCTGGCTTTG
Dr_tRNA1-met_fwd	ATTCAAGCAAATCTGGGGTAAAGC
Dr_tRNA1-met_rev	ACCCAGAAGAAAGTGGTGGTTC
D-rad_TsaB_out-5'	TTTGGCAATGTTCGACATAGTGAG
D-rad_TsaB_out-3'	CAAGATTTCGTTCGGTGAAGGTG
D-rad_TsaB_in-5'	CCCTCTTTCTGTGACCCTCG
D-rad_TsaB_in-3'	CCTCATAGACCGCGCTATACAC
D-rad_TsaD_out-5'	GCGAGATCAGCGAGATTATCTTTG
D-rad_TsaD_out-3'	CTACCAGGACTTCAACACCCAG
D-rad_TsaD_in-5'	TCCTGGTCGCAGACGTTCAC
D-rad_TsaD_in-3'	GTAAGGTGCTTGCAACTGCTC
TTHA0793_Fw	TCATGGCGCTCCGGATCTTGCCCCTTA
TTHA0793_Rv	GTTCGGGAGAACGAGCTTTATAGC
TTHA0793_in_Fw	GAAAGAGTATACCAGGAAGCCCTG
TTHA0793_in_Rv	AAGGGCCTCCTTGGGGATCGCCCCTT
FtRNA^{Ile}(cat)	AATTCCTGCAGTAATACGACTCACTATAGGGCCCCTTAGCTCAGTGGTTAGAGCAGGCG
RtRNA^{Ile}(cat)	TGGTGGCCCCTGCTGGACTTGAACCAGCGACCAAGCGATTATGAGTCGCCTGCTCT
RtRNA^{Ile}(cat)	AATTCCTGCAGTAATACGACTCACTATAAGGCTTGTAGCTCAGGTGGTTAGAGCGC
RtRNA^{Ile}(cat)	TGGTAGGCCTGAGTGGACTTGAACCACCGACCTCACCCTTATCAGGGGTGCGCTCTAAC
BAD33IleRSMtB.ol1	ATATATTATTCTAGAGGAGTGAAACGATGACCGATAACGCATATCCAAAG
BAD33IleRSMtB.ol2	AATTAATTAATTAAGCTTTTAGGTCTTTTCGATGCTTACCC
BADOl5	AAGATTAGCGGATCCTACCTG
BADRev2	TTCTGATTTATTCTGTATCAGGC
Ptet2771-fwd	ATAGGCGTATCACGAGGCCC
ygjD661-rev	GGTGCCGTTGTCACGAATGGT
yrdC573-rev	TTACCCCTGTCGAAACAGTTCA

Table S4. List of proteins detected from mid-log phase growth, predicted fitness, and frequency of run of two or more ANN codons. (see Excel sheet)

Table S5. List of proteins detected from late-log phase growth, predicted fitness, and frequency of run of two or more ANN codons. (see Excel sheet)

Table S6. D. radiodurans ORFs with all Ser codons coded by AGY codons.

Gene	**Annotation**	**Length (AA)**
DR_C0012	transcriptional regulator, GerE family	59
DR_C0025	chromosome partitioning protein, putative, ParA family	74
DR_C0030	hypothetical protein	68
DR_0031	FIG00577642: hypothetical protein	80
DR_0175	SSU ribosomal protein S9p (S16e)	133
DR_0207	hypothetical protein	47
DR_0208	hypothetical protein	130
DR_0216	Sulfur acceptor protein SufE for iron-sulfur cluster assembly	147
DR_0221	DNA methylase, putative	71
DR_0234	Transcriptional regulator	86
DR_0265	Predicted transcriptional regulator of N-Acetylglucosamine utilization, GntR family	190
DR_0269	FIG00577917: hypothetical protein	322
DR_0292	Twin-arginine translocation protein TatA	117
DR_0312	LSU ribosomal protein L4p (L1e)	205
DR_0318	LSU ribosomal protein L16p (L10e)	141
DR_0320	SSU ribosomal protein S17p (S11e)	96
DR_0322	LSU ribosomal protein L24p (L26e)	115
DR_0328	FIG01121178: hypothetical protein	140
DR_0331	hypothetical protein	66
DR_0338	hypothetical protein	57
DR_0341	SSU ribosomal protein S15p (S13e)	91
DR_0360	hypothetical protein	143
DR_0381	NADPH dependent preQ0 reductase-like	134
DR_0476	FIG013268: hypothetical protein	103
DR_0545	hypothetical protein	91
DR_0548	Type IV pilin PilA	158
DR_0586	ATP-dependent Clp protease adaptor protein ClpS	113
DR_0637	hypothetical protein	77
DR_0661	Programmed cell death antitoxin YdcD	84
DR_0691	hypothetical protein	202
DR_0734	hypothetical protein	72
DR_0780	hypothetical protein	92
DR_0800	hypothetical protein	60

DR_0825	LSU ribosomal protein L31p @ LSU ribosomal protein L31p, zinc-independent	73
DR_0865	Peroxide stress regulator; Ferric uptake regulation protein; Fe2+/Zn2+ uptake regulation proteins	132
DR_0877	hypothetical protein	114
DR_0909	hypothetical protein	160
DR_0955	hypothetical protein	77
DR_0993	hypothetical protein	116
DR_0994	hypothetical protein	112
DR_0995	FIG00579650: hypothetical protein	115
DR_1004	Quaternary ammonium compound-resistance protein SugE	96
DR_1047	hypothetical protein	68
DR_1071	hypothetical protein	105
DR_1108	hypothetical protein	146
DR_1111	transcriptional regulator, MerR family	137
DR_1134	hypothetical protein	53
DR_1162	Transcription elongation factor GreA	151
DR_1204	4-carboxymuconolactone decarboxylase (EC 4.1.1.44)	132
DR_1216	hypothetical protein	156
DR_1241	hypothetical protein	71
DR_1332	hypothetical protein	132
DR_1404	hypothetical protein	73
DR_1409	hypothetical protein	382
DR_1429	FIG00579912: hypothetical protein	57
DR_1609	Predicted 2-keto-4-pentenoate hydratase/2-oxohepta-3-ene-1,7-dioic acid hydratase	250
DR_1663	hypothetical protein	150
DR_1676	FIG00579720: hypothetical protein	305
DR_1686	FIG00580334: hypothetical protein	79
DR_1737	Signal peptidase I (EC 3.4.21.89)	196
DR_1739	Transcriptional regulator, PadR family	105
DR_1776	Nudix dNTPase DR1776 (EC 3.6.1.-)	157
DR_1806	Nucleotidyltransferase (EC 2.7.7.-)	74
DR_1807	Nucleotidyltransferase (EC 2.7.7.-)	134
DR_1816	hypothetical protein	142
DR_1844	hypothetical protein	93
DR_1847	hypothetical protein	72
DR_1904	hypothetical protein	51
DR_1942	Acyl carrier protein	76
DR_1954	Transcriptional regulator, PadR family	107
DR_1957	hypothetical protein	148
DR_1968	Oxygen-insensitive NAD(P)H nitroreductase (EC 1.-.-.-) / Dihydropteridine reductase (EC 1.5.1.34)	212
DR_2005	LSU ribosomal protein L35p	66
DR_2009	KH domain RNA binding protein YlqC	82
DR_2030	hypothetical protein	159

DR_2043	LSU ribosomal protein L7/L12 (P1/P2)	122
DR_2044	LSU ribosomal protein L10p (P0)	169
DR_2049	LSU ribosomal protein L33p	55
DR_2061	hypothetical protein	145
DR_2079	Transcriptional regulator, PadR family	116
DR_2085	glutaredoxin, putative	81
DR_2109	SSU ribosomal protein S14p (S29e)	89
DR_2114	LSU ribosomal protein L30p (L7e)	55
DR_2161	protein-tyrosine phosphatase-related protein	177
DR_2165	PaaD-like protein (DUF59) involved in Fe-S cluster assembly	106
DR_2230	hypothetical protein	58
DR_2247	carbon monoxide dehydrogenase G protein	153
DR_2266	hypothetical protein	109
DR_2273	DNA topology modulation protein FlaR-related protein	162
DR_2292	hypothetical protein	70
DR_2355	GCN5-related N-acetyltransferase	154
DR_2369	hypothetical protein	110
DR_2378	hypothetical protein	64
DR_2380	FIG01230208: hypothetical protein	73
DR_2456	hypothetical protein	57
DR_2512	Endoribonuclease L-PSP	125
DR_2529	hypothetical protein	145
DR_2534	grpB protein	111
DR_2573	hypothetical protein	103
DR_2580	Iojap protein	116
DR_A0059	Transcriptional regulator, PadR family	182
DR_A0077	hypothetical protein	142
DR_A0095	hypothetical protein	130
DR_A0102	hypothetical protein	106
DR_A0165	FIG00514740: hypothetical protein	134
DR_A0167	conserved protein	183
DR_A0267	Formylmethanofuran dehydrogenase subunit E	198
DR_A0300	hypothetical protein	138
DR_A0314	Urease accessory protein UreE	114
DR_A0319	Urease beta subunit (EC 3.5.1.5) / Urease gamma subunit (EC 3.5.1.5)	228

Table S7. D. radiodurans ORFs with all Arg codons coded by AGR codons.

PatricBRC Locus Tag	RefSeq Locus Tag	Annotation
VBIDeiRad64572_1241	DR_1045	hypothetical protein
VBIDeiRad64572_1529		hypothetical protein
VBIDeiRad64572_2039		hypothetical protein
VBIDeiRad64572_2919		hypothetical protein

Table S8. t^6A dependent codon usage of D. radiodurans. (see Excel sheet)

Table S9. D. radiodurans ORFs containing consecutive AUA codons.

Gene	**Annotation**	**Total number of repeated AUA**	**% of repeated AUA in protein**	**Fitness score[#]**
DR_0664	hypothetical protein	2	1.2	1
DR_0666	transposase	2	0.5	1
DR_0709	snoG protein	2	0.7	0.922
DR_0913	hypothetical protein	2	1.7	1
DR_0978	transposase	2	0.5	1
DR_1381	transposase	2	0.5	1
DR_1593	transposase	2	0.5	1
DR_1651	transposase	2	0.5	1
DR_1655	ABC transporter periplasmic substrate-binding protein	2	0.6	1
DR_1933	transposase	2	0.5	1
DR_2087	translation initiation factor IF-3	2	0.9	0.043
DR_2250	methoxyneurosporene dehydrogenase	2	0.4	1
DR_2301	hypothetical protein	2	0.9	1
DR_2324	transposase	2	0.5	1
DR_A0036	hypothetical protein	2	0.5	1
DR_A0039	mannosyltransferase	2	0.6	1
DR_A0172	thiazole synthase	2	0.7	0.775
DR_B0084	hypothetical protein	2	3.0	ND

http://cefg.uestc.edu.cn/ifim/data.php?id=NC_001263&strain=Deinococcus_radiodurans_R1&thread=1
and
http://cefg.uestc.edu.cn/ifim/data.php?id=NC_001264&strain=Deinococcus_radiodurans_R1&thread=1
A score < 0.65
ND: not determined

Table S10. Essential genes in E. coli with runs of two or more AUA codons

RefSeq Locus Tag	**Function**	**Gene**	**Length of AUA Repeat**	**Number of Occurrences of Repeat**
b4261	LPS export ABC transporter permease	LptF	2	1
b1740	NAD synthase ammonia dependent	NadE	2	1
b3973	Bifunctional biotin protein ligase biotin operon repressor	BirA	2	1

Table S11. Essential genes in E. coli with runs of two or more AUU codons

RefSeq Locus Tag	**Function**	**Gene**	**Length of AUU Repeat**	**Number of Occurrences of repeat**
b3250	Cell division and growth	MreC	3	1
b1912	Phosphatidylglycerophosphate synthase	PgsA	3	1
b3176	Phosphoglucosamine mutase	GlmM	3	1
b0886	Glutathione/cysteine ABC transporter permease/ATPase	CydC	2	1
b4261	LPS export ABC transporter permease	LptF	2	1
b1876	Glutathione/cysteine ABC transporter permease/ATPase	CydC	2	1
b0657	Arginine--tRNA ligase	ArgS	2	1
b0526	Apolipoprotein N-acyltransferase	Lnt	2	1
b4052	Cysteine--tRNA ligase	CysS	2	1
b2563	Replicative DNA helicase	DnaB	2	1
b2315	ACP-CoA phosphopantetheinyltransferase	AcpS	2	1
b0529	Dihydrofolate:folylpolyglutamate synthase	FolC	2	1
b0172	Methenyltetrahydrofolate dehydrogenase/cyclohydrolase	FolD	2	1
b3463	Ribosome recycling factor (RRF)	Frr	2	1
b3805	Cell division ATP-binding protein	FtsE	2	1
b0154	Porphobilinogen deaminase	HemC	2	1
b0918	Glutamate-1-semialdehyde aminomutase	HemL	2	1
b1175	3-deoxy-manno-octulosonate cytidylyltransferase	KdsB	2	1
b1174	Inhibitor of FtsZ ring polymerization	MinD	2	2
b0087	Blocks MinCD inhibition of FtsZ polymerization at cell center	MinE	2	1
b3249	UDP-N-acetylmuramoyl-pentapeptide:undecaprenyl-PO4 phosphatase	MraY	2	1
b0085	Mecillinam resistance	MreD	2	1
b0086	meso-diaminopimelate adding enzyme	MurE	2	2
b3300	D-alanyl:D-alanine adding enzyme	MurF	2	1
b2585	SecYEG inner membrane translocon core subunit	SecY	2	1
b3783	Phosphatidylserine synthase	PssA	2	1
b3320	Transcription termination factor Rho	Rho	2	1
b2606	50S ribosomal subunit protein L3	RplC	2	1
b3318	50S ribosomal subunit protein L19	RplS	2	1
b0098	50S ribosomal subunit protein L23	RplW	2	1

b3648	Preprotein translocase secAYEG receptor/ATPase subunit	SecA	2	1
b3705	Guanylate kinase	Gmk	2	1
b3288	Membrane protein insertase	YidC	2	1
b3189	Methionyl-tRNA formyltransferase	Fmt	2	1
b4040	UDP-N-acetylglucosamine enoylpyruvyl transferase	MurA	2	1
b3018	4-Hydroxybenzoate polyprenyltransferase	UbiA	2	1
b3623	1-Acyl-n-glycerol-3-phosphate acyltransferase	PlsC	2	2
b3793	Adds terminal GlcNac side branch to the lipopolysaccharide core prior to attachment of the O antigen	WaaU	2	2
b3201	Wzy protein involved in ECA polysaccharide chain elongation	WzyE	2	1
b4084	LPS export ABC transporter ATPase	LptB	2	1
b2615	D-allose kinase	AlsK	2	1
b0103	ATP-NAD kinase	NadK	2	1
b0176	Dephospho-CoA kinase	CoaE	2	1
b4162	Inner membrane zinc RIP metalloprotease	RseP	2	2
b3183	3' to 5' oligoribonuclease	Orn	2	1
b0890	DNA-binding GTPase involved in cell partioning and DNA repair	ObgE	2	1
b2949	DNA translocase at septal ring sorting daughter chromsomes	FtsK	2	1
b0174	Putative anti-termination factor for Rho-dependent terminators	YqgF	2	1
b1116	Undecaprenyl pyrophosphate synthase	IspU	2	1
b1118	LolA-dependent release of lipoproteins from inner membrane	LolC	2	1

Table S12. Essential genes in E. coli with runs of two or more AUC codons

RefSeq Locus Tag	**Function**	**Gene**	**Length of AUC Repeat**	**Number of occurrences of repeat**	**Length of AUC Repeat**	**Number of occurrences of repeat**
b3793	Wzy protein involved in ECA polysaccharide chain elongation	WzyE	2	1	3	1
b3705	Membrane protein insertase	YidC	3	1		
b1876	Arginine--tRNA ligase	ArgS	2	1		
b3433	Aspartate semialdehyde dehydrogenase	Asd	2	1		
b1866	Aspartate--tRNA ligase	AspS	2	2		

b0657	Apolipoprotein N-acyltransferase	Lnt	2	1
b0733	Cytochrome d (bd-I) ubiquinol oxidase subunit 1	CydA	2	1
b0526	Cysteine--tRNA ligase	CysS	2	1
b4052	Replicative DNA helicase	DnaB	2	1
b2779	Enolase	Eno	2	3
b2566	Ribosome-associated GTPase essential for growth	Era	2	1
b0048	Dihydrofolate reductase	FolA	2	1
b0093	Divisome assembly protein	FtsQ	2	1
b0089	Putative lipid II flippase	FtsW	2	1
b3340	Elongation Factor EF-G	FusA	2	4
b2515	1-hydroxy-2-methyl-2-(E)-butenyl 4-diphosphate synthase	IspG	2	1
b3729	Glucosamine-6-phosphate synthase	GlmS	2	2
b0680	Glutamine--tRNA ligase	GlnS	2	1
b0154	Glutamate-1-semialdehyde aminomutase	HemL	2	2
b0026	Isoleucine--tRNA ligase	IleS	2	5
b3168	Translation initiation factor IF-2	InfB	2	2
b0918	3-deoxy-manno-octulosonate cytidylyltransferase	KdsB	2	1
b2942	S-adenosylmethionine synthase	MetK	2	1
b4143	Chaperonin Cpn60	GroL	2	3
b3251	Cell wall structural actin-like protein in MreBCD complex	MreB	2	1
b0085	meso-diaminopimelate adding enzyme	MurE	2	1
b3169	Transcription termination/antitermination L factor	NusA	2	1
b4041	Glycerol-3-phosphate	PlsB	2	1

	acyltransferase			
b3863	DNA polymerase I	PolA	2	1
b3300	SecYEG inner membrane translocon core subunit	SecY	2	1
b1207	Phosphoribosylpyrophosphate synthase	Prs	2	3
b2585	Phosphatidylserine synthase	PssA	2	1
b1204	Peptidyl-tRNA hydrolase	Pth	2	1
b2780	CTP synthase	PyrG	2	1
b3319	50S ribosomal subunit protein L4	RplD	2	1
b3308	50S ribosomal subunit protein L5	RplE	2	1
b3231	50S ribosomal subunit protein L13	RplM	2	1
b3310	50S ribosomal subunit protein L14	RplN	2	2
b3294	50S ribosomal subunit protein L17	RplQ	2	1
b0169	30S ribosomal subunit protein S2	RpsB	2	1
b3303	30S ribosomal subunit protein S5	RpsE	2	1
b3307	30S ribosomal subunit protein S14	RpsN	2	1
b0098	Preprotein translocase secAYEG receptor/ATPase subunit	SecA	2	3
b0408	SecDFyajC inner membrane secretion protein complex subunit	SecD	2	2
b0893	Serine--tRNA ligase	SerS	2	2
b2533	Inositol-1-monophosphatase	SuhB	2	1
b4258	Valine--tRNA ligase	ValS	2	1
b1092	Malonyl-CoA-acyl carrier protein transacylase	FabD	2	1
b4040	4-Hydroxybenzoate polyprenyltransferase	UbiA	2	1
b0915	Lipid A 4' kinase	LpxK	2	2
b3176	Phosphoglucosamine	GlmM	2	1

	mutase			
b4262	LPS export ABC transporter permease	LptG	2	1
b3398	Putative RcsCDB-response attenuator	YrfF	2	1
b2949	Putative anti-termination factor for Rho-dependent terminators	YqgF	2	1
b1118	LolA-dependent release of lipoproteins from inner membrane	LolE	2	1
b1069	Putative lipid II flippase	MurJ	2	1
b2412	FtsZ stabilizer	ZipA	2	1
b2595	TPR-repeat lipoprotein required for OM biogenesis	BamD	2	1
b3185	50S ribosomal subunit protein L27	RpmA	2	1

Table S13. Synechocystis sp. PCC 6803 ORFs containing consecutive AUA codons.

RefSeq Locus Tag	**Function**	**Length of AUA repeat**	**Number of Occurrences of Repeat**
sll0092	Mobile element protein	2	3
slr5029	Mobile element protein	2	3
sll7002	Mobile element protein	2	3
slr2036	RecD-like DNA helicase YrrC (Transposase)	2	2
slr1210	hypothetical protein	2	2
sll1780	RecD-like DNA helicase YrrC (Transposase)	2	2
sll0842	Maltodextrin glucosidase (EC 3.2.1.20)	2	2
slr1243	hypothetical protein	2	2
sll1997	RecD-like DNA helicase YrrC (Transposase)	2	2
sll1255	RecD-like DNA helicase YrrC (Transposase)	2	2
slr1682	Mobile element protein	2	2
sll1560	RecD-like DNA helicase YrrC (Transposase)	2	2
slr1635	RecD-like DNA helicase YrrC (Transposase)	2	2
slr0180	RecD-like DNA helicase YrrC (Transposase)	2	2
sll0315	RecD-like DNA helicase YrrC (Transposase)	2	2
slr0468	hypothetical protein	2	2
slr0913	hypothetical protein	2	2
sll1474	RecD-like DNA helicase YrrC (Transposase)	2	2

sll5131	RecD-like DNA helicase YrrC (Transposase)	2	2
slr7008	RecD-like DNA helicase YrrC (Transposase)	2	2
slr6105	RecD-like DNA helicase YrrC (Transposase)	2	2
sll1408	Transcriptional regulator, AraC family	2	1
sll1401	hypothetical protein	2	1
slr0727	hypothetical protein	2	1
slr1063	Dolichol-phosphate mannosyltransferase (EC 2.4.1.83) in lipid-linked oligosaccharide synthesis cluster	2	1
slr1070	Glycosyl transferase, family 2	2	1
slr1042	two-component system, regulatory protein	2	1
ssl1920	Mobile element protein	2	1
sll0986	Mobile element protein	2	1
sll1582	hypothetical protein	2	1
sll1571	hypothetical protein	2	1
sll1820	tRNA pseudouridine synthase A (EC 4.2.1.70)	2	1
ssr2009	hypothetical protein	2	1
	hypothetical protein	2	1
sll1716	Mobile element protein	2	1
	hypothetical protein	2	1
sll1765	hypothetical protein	2	1
sll1710	Mobile element protein	2	1
slr0265	Mobile element protein	2	1
ssl0438	50S ribosomal protein L12 homologue	2	1
slr2080	hypothetical protein	3	1
sll1429	putative	2	1
slr2111	hypothetical protein	2	1
slr2116	hypothetical protein	2	1
sll1241	hypothetical protein	2	1
sll1094	hypothetical protein	2	1
slr1683	Mobile element protein	2	1
sll1531	FIG00560114: hypothetical protein	2	1
sll0428	Prolyl 4-hydroxylase, alpha subunit	2	1
sll1861	Mobile element protein	2	1
sll1860	Mobile element protein	2	1
slr0350	Mobile element protein	2	1
sll0012	Mobile element protein	2	1
slr0222	hypothetical protein	2	1
slr0574	cytochrome P450	2	1
slr0511	Mobile element protein	2	1
sll0677	Mobile element protein	2	1
slr0703	Mobile element protein	2	1
sll0665	Mobile element protein	2	1
sll0664	hypothetical protein	2	1

slr0666	hypothetical protein	2	1
sll0441	hypothetical protein	2	1
sll5083	hypothetical protein	2	1
sll5104	Arsenate reductase (EC 1.20.4.1)	2	1
sll7031	FIG032766: hypothetical protein	2	1
sll7047	hypothetical protein	2	1
slr6013	hypothetical protein	2	1
slr6016	hypothetical protein	2	1
slr6037	Arsenate reductase (EC 1.20.4.1)	2	1
slr6072	hypothetical protein	2	1
slr6075	hypothetical protein	2	1
slr6106	hypothetical protein	2	1

Table S14. S. mutans ORFs containing consecutive AUA codons.

RefSeq Locus Tag	Function	Gene	Length of ANN repeat	Number of Occurrences of Repeat	Length of ANN repeat	Number of Occurrences of Repeat
SMU.2033c	FIG01116025: hypothetical protein		2	2	3	1
SMU.02	DNA polymerase III beta subunit (EC 2.7.7.7)	dnaN	2	1		
SMU.09	Ribosome-associated heat shock protein implicated in the recycling of the 50S subunit (S4 paralog)		2	1		
SMU.46	hypothetical protein		2	1		
SMU.52	hypothetical protein		2	1		
SMU.117c	FIG01118983: hypothetical protein		2	1		
SMU.156	FIG01114029: hypothetical protein		2	1		
SMU.158	Cysteinyl-tRNA synthetase (EC 6.1.1.16)	cysS	2	1		
SMU.191c	Integrase		2	1		
SMU.221c	putative integrase		2	1		
SMU.224c	hypothetical protein		2	1		
SMU.239c	FIG01115616: hypothetical protein		2	2		
SMU.246	Undecaprenyl-phosphate N-acetylglucosaminyl 1-phosphate transferase (EC 2.7.8.-)	rgpG	2	1		
SMU.255	Oligopeptide ABC transporter, periplasmic oligopeptide-binding protein OppA (TC 3.A.1.5.1)	oppA	2	1		
SMU.286	Competence-stimulating peptide ABC transporter ATP-binding protein ComA		2	1		
SMU.343	FIG01118487: hypothetical		2	1		

	protein			
SMU.387	YgjD/Kae1/Qri7 family, required for N6-threonylcarbamoyl adenosine t(6)A37 modification in tRNA	tsaD	2	1
SMU.411c	Hypothetical protein in cluster with Ecs transporter (in Streptococci)		2	1
SMU.431	Zinc ABC transporter, ATP-binding protein ZnuC		2	1
SMU.482	16S rRNA (cytosine(967)-C(5))-methyltransferase (EC 2.1.1.176)	sunL	2	1
SMU.572	Methylenetetrahydrofolate dehydrogenase (NADP+) (EC 1.5.1.5) / Methenyltetrahydrofolate cyclohydrolase (EC 3.5.4.9)	folD	2	1
SMU.582	Octaprenyl diphosphate synthase (EC 2.5.1.90) / Dimethylallyltransferase (EC 2.5.1.1) / (2E,6E)-farnesyl diphosphate synthase (EC 2.5.1.10) / Geranylgeranyl diphosphate synthase (EC 2.5.1.29)		2	2
SMU.640c	Transcriptional regulator, GntR family domain / Aspartate aminotransferase (EC 2.6.1.1)		2	1
SMU.654	Lantibiotic transport ATP-binding protein srtF	mutF	2	1
SMU.728	putative oxidoreductase		2	1
SMU.735	FIG01118003: hypothetical protein		2	1
SMU.796	FIG01119187: hypothetical protein		2	1
SMU.809	Excinuclease ABC subunit B	uvrB	2	1
SMU.820	FIG01118636: hypothetical protein		2	1
SMU.832	Glucosyltransferase (side chain biosynthesis) (EC 2.4.1.-)		2	1
SMU.851	NADH pyrophosphatase (EC 3.6.1.22)		2	1
SMU.864	ABC transporter permease protein		2	1
SMU.876	MSM (multiple sugar metabolism) operon regulatory protein	msmR	2	1
SMU.905	ABC transporter, ATP-binding/permease protein		2	1
SMU.939	Isopentenyl-diphosphate delta-isomerase, FMN-dependent (EC 5.3.3.2)		2	1
SMU.947	Dihydrofolate reductase (EC 1.5.1.3)	dfrA	2	1
SMU.988	Cardiolipin synthetase (EC 2.7.8.-)		2	1
SMU.995	Iron(III) dicitrate transport system		2	1

	permease protein FecD (TC 3.A.1.14.1)			
SMU.996	Uncharacterized iron compound ABC uptake transporter, permease protein		2	1
SMU.1017	Oxaloacetate decarboxylase beta chain (EC 4.1.1.3)	oadB	2	1
SMU.1027	Transcriptional regulator, TetR family		2	1
SMU.1028	alpha/beta hydrolase fold		2	1
SMU.1042	ABC transporter permease protein		2	1
SMU.1044c	Similar to ribosomal large subunit pseudouridine synthase D, Bacillus subtilis YjbO type		2	1
SMU.1067c	putative ABC transporter, permease protein		2	1
SMU.1102	6-phospho-beta-glucosidase (EC 3.2.1.86)	arb	2	1
SMU.1129	Two component system response regulator CiaR	ciaR	2	1
SMU.1150	putative transporter, trans-membrane domain bacteriocin immunity protein		3	1
SMU.1160c	FIG01119470: hypothetical protein		2	1
SMU.1174	ATP-dependent DNA helicase UvrD/PcrA	pcrA	2	1
SMU.1235	GTPase and tRNA-U34 5-formylation enzyme TrmE	trmE	2	1
SMU.1260c	Cell division initiation protein DivIVA		2	1
SMU.1262c	hypothetical protein		2	2
SMU.1339	Long-chain-fatty-acid--CoA ligase (EC 6.2.1.3)	bacD	2	2
SMU.1340	putative bacitracin synthetase	bacA2	2	1
SMU.1342	Long-chain-fatty-acid--CoA ligase (EC 6.2.1.3)	bacA1	2	2
SMU.1346	Thioesterase	bacT	2	1
SMU.1456c	hypothetical protein		2	1
SMU.1486c	Histidinol-phosphatase (EC 3.1.3.15)		2	1
SMU.1541	Glycogen debranching enzyme (EC 3.2.1.-) / Pullulanase (EC 3.2.1.41)	pulA	2	1
SMU.1550c	ABC transporter permease protein		2	1
SMU.1558c	Ribosomal-protein-S5p-alanine acetyltransferase		2	1
SMU.1577c	hypothetical protein		2	1
SMU.1611c	Multidrug resistance efflux pump PmrA		2	1
SMU.1661c	Signal peptidase-like protein		2	1
SMU.1679c	Phospholipid-binding protein		2	1
SMU.1692	Pyruvate formate-lyase activating	pflA	2	1

	enzyme (EC 1.97.1.4)			
SMU.1714c	Tyrosine recombinase XerD	xerD	2	1
	hypothetical protein		2	1
SMU.1810	FIG01116767: hypothetical protein	scnE	2	1
SMU.1823	Nicotinamidase (EC 3.5.1.19)	pncA	2	1
SMU.1835	Holo-[acyl-carrier protein] synthase (EC 2.7.8.7)	acpS	3	1
SMU.1876	Putative membrane protein precursor SPs0273		2	1
SMU.1881c	Competence-stimulating peptide ABC transporter ATP-binding protein ComA		2	2
	Transposase		2	1
SMU.1916	Histidine kinase of the competence regulon ComD	comD	2	1
SMU.1926	Transcriptional regulator, TetR family	psaR	2	1
SMU.1940c	Catalyzes the cleavage of p-aminobenzoyl-glutamate to p-aminobenzoate and glutamate, subunit A		2	1
SMU.2001	DNA-directed RNA polymerase alpha subunit (EC 2.7.7.6)	rpoA	2	1
SMU.2078c	Putative Holliday junction resolvase YqgF		2	1

Table S15. E. coli ORFs containing consecutive AUA codons.

RefSeq Locus Tag	Function	Gene	Length of AUA repeat	Number of Occurrences of Repeat
b2213	ADA regulatory protein / Methylated-DNA--protein-cysteine methyltransferase (EC 2.1.1.63)	ada	2	1
b4494	FIG00638086: hypothetical protein	arpB	2	1
b3739	FIG048548: ATP synthase protein I2	atpI	2	1
b2901	6-phospho-beta-glucosidase (EC 3.2.1.86)	bglA	2	1
b3973	Biotin operon repressor / Biotin-protein ligase (EC 6.3.4.15)	birA	2	1
b3911	Copper sensory histidine kinase CpxA	cpxA	2	1
b0886	Transport ATP-binding protein CydC	cydC	2	1
b4123	C4-dicarboxylate transporter DcuB	dcuB	2	1
b1201	Phosphoenolpyruvate-dihydroxyacetone phosphotransferase operon regulatory protein DhaR	dhaR	2	1
b2269	ElaD protein	elaD	2	1
b2367	Inner membrane component of tripartite multidrug resistance system	emrY	3	1

b2370	Hybrid sensory histidine kinase in two-component regulatory system with EvgA	evgS	2	1
b4317	type 1 fimbriae anchoring protein FimD	fimD	2	1
b1072	Flagellar basal-body P-ring formation protein FlgA	flgA	2	1
b2321	Cell division protein	flk	2	1
b3447	Gamma-glutamyltranspeptidase (EC 2.3.2.2)	ggt	2	1
b2975	Glycolate permease	glcA	2	1
b3322	PioO protein	gspB	2	1
b2352	putative ligase	gtrS	2	1
b3805	Porphobilinogen deaminase (EC 2.5.1.61)	hemC	2	1
b1579	Integrase	intQ	2	1
b1345	Phage integrase	intR	2	1
b2289	LysR family transcriptional regulator lrhA	lrhA	2	1
b1159	5-methylcytosine-specific restriction enzyme A (EC 3.1.21.-)	mcrA	2	1
b4345	enzyme; Degradation of DNA	mcrC	2	1
b1686	Methionine ABC transporter ATP-binding protein	metN	2	1
b0199	Molybdenum transport system permease protein ModB (TC 3.A.1.8.1)	modB	2	1
b0764	RelF inactive antibacterial toxin protein	mokC	2	1
b0018	NAD synthetase (EC 6.3.1.5)	nadE	2	1
b1740	Nitrate/nitrite sensor protein (EC 2.7.3.-)	narX	2	1
b1222	Phenylacetate-coenzyme A ligase (EC 6.2.1.30) PaaF	paaK	2	1
b1398	O-antigen ligase	rfaL	2	1
b3593	Lipopolysaccharide core biosynthesis protein RfaS	rfaS	2	1
b0497	core protein	rhsA	2	1
b1822	core protein	rhsD	2	1
b1861	Ribosomal RNA large subunit methyltransferase A (EC 2.1.1.51)	rlmA	2	1
b2702	Holliday junction DNA helicase RuvA	ruvA	2	1
b3622	PTS system, glucitol/sorbitol-specific IIC component (EC 2.7.1.69)	srlA	2	1
b3629	Colanic acid polymerase WcaD	wcaD	2	1
b2056	FIG00638728: hypothetical protein	yahL	2	1
b0326	Putative uncharacterized protein YaiO	yaiO	2	1
b0358	Transcriptional regulator, AraC/XylS family	ybcM	2	1
b0546	Mobile element protein	ybfL	2	1
b0705	FIG00638277: hypothetical protein	ybhM	2	1
b0787	Sensory box/GGDEF domain protein	ycdT	2	1
b1025	Putative hydrolase YcdX (EC 3.1.-.-)	ycdX	2	1

b1034	FIG00638073: hypothetical protein	ycjD	2	1
b1289	Putative oxidoreductase YcjS (EC 1.-.-.-), NADH-binding	ycjS	2	1
b1315	FIG00638289: hypothetical protein	ydbD	2	1
b1407	Mobile element protein	ydcC	2	1
b1460	internalin, putative	yddK	2	1
b1471	FIG002649: ydiI hotdog fold superfamily	ydiI	2	1
b1690	Putative transport system permease protein	ydiM	2	1
b1697	Electron transfer flavoprotein, beta subunit	ydiQ	2	1
b1770	Putative HTH-type transcriptional regulator YdjF	ydjF	2	1
b1730	FIG00638395: hypothetical protein	ydjO	2	1
b1834	Paraquat-inducible protein B	yebT	2	1
b2273	FIG00640159: hypothetical protein	yfbN	2	1
b2362	Eae protein	yfdS	2	1
b2625	Hyphotheical protein	yfjI	2	1
b2642	FIG00638879: hypothetical protein	yfjW	3	2
b4462	Alpha-amylase (EC 3.2.1.1) (1,4-alpha-D-glucan glucanohydrolase) (GLYCOGENASE)	ygaQ	2	1
b2735	Putative transcriptional regulator YgbI, DeoR family	ygbI	2	1
b2858	ORF_f143	ygeN	2	1
b2930	Fructose-1,6-bisphosphatase, GlpX type (EC 3.1.3.11)	yggF	2	1
b3087	Possible oxidoreductase	ygjR	2	1
b3173	Outer-membrane protein yhbX precursor	yhbX	2	1
b3484	Mobile element protein	yhhI	2	1
b3442	hypothetical protein	yhhZ	2	2
b3504	FIG00641106: hypothetical protein	yhiS	2	1
b3638	DNA repair protein RadC	yicR	2	1
b3818	FIG00640330: hypothetical protein	yigG	2	1
b3861	Putative GTP-binding protein	yihF	2	1
b4047	FIG00639441: hypothetical protein	yjbL	2	1
b4231	Putative sugar ABC transport system, permease protein YjfF	yjfF	2	1
b4365	Putative regulator	yjjQ	2	1
b4567	FIG00638078: hypothetical protein	yjjZ	2	1
b1527	FIG00638496: hypothetical protein	yneK	2	1
b3013	Uncharacterized protein yqhG precursor	yqhG	2	1

Figures

Figure S1

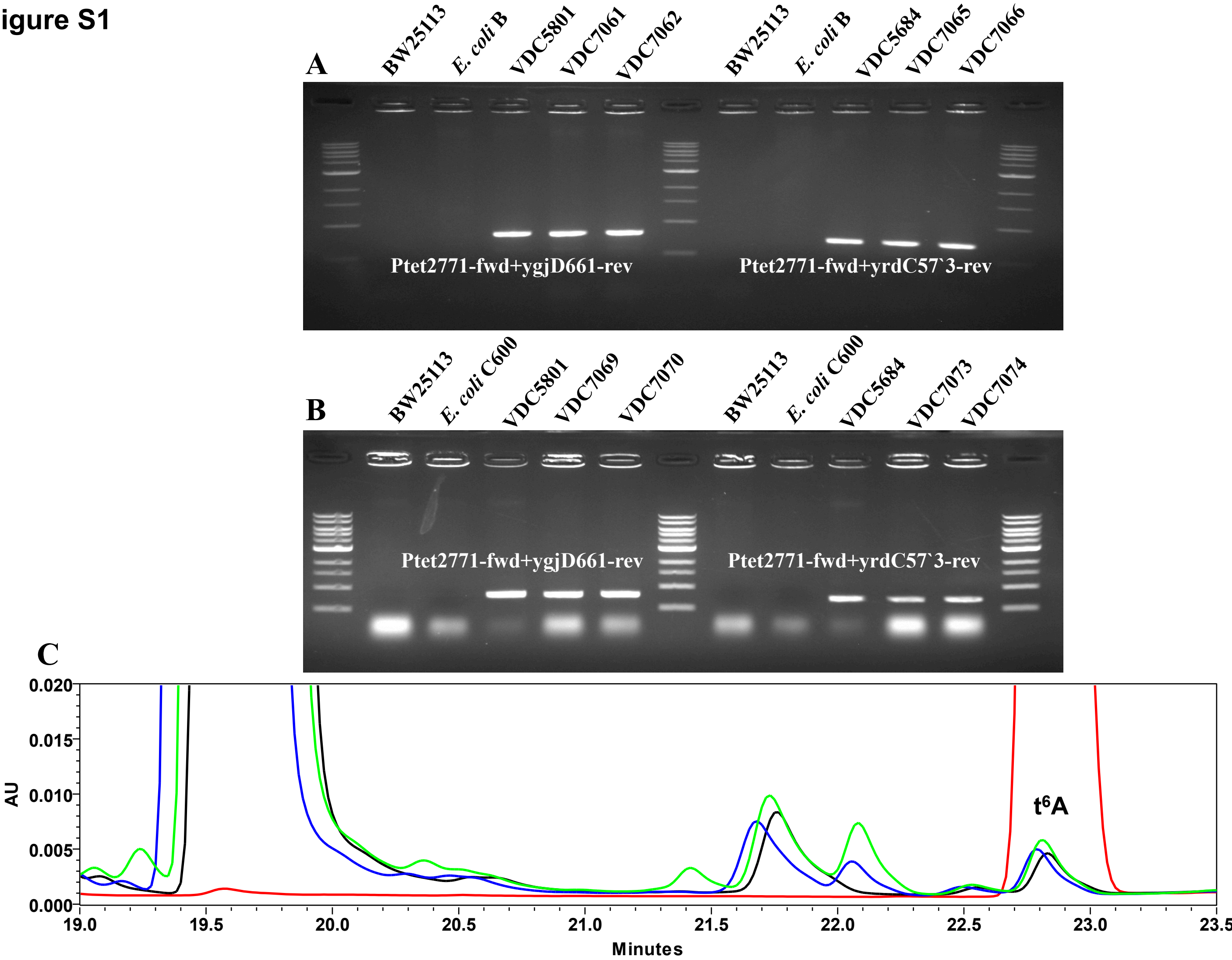

Figure S1. (A) PCR verification of P1 transduction from VDC5801 ($\mathrm{P}_{\mathrm{TET}}$::*tsaD*) and VDC5684 ($\mathrm{P}_{\mathrm{TET}}$::*tsaC*) into *E. coli* B. (B) PCR verification of P1 transduction from VDC5801 ($\mathrm{P}_{\mathrm{TET}}$::*tsaD*) and VDC5684 ($\mathrm{P}_{\mathrm{TET}}$::*tsaC*) into *E. coli* C600. The primer pairs used are indicated in the respective figure panels. (C) HPLC analysis of nucleosides in digests of tRNAs extracted from *E. coli* B (black line), $\mathrm{P}_{\mathrm{TET}}$::*tsaD* (VDC7062) (blue line) and $\mathrm{P}_{\mathrm{TET}}$::*tsaC* (VDC7066) (green line). Strains were grown in LB without the addition of the inducer aTc. Red line indicates t^{6}A synthetic standard.

Figure S2

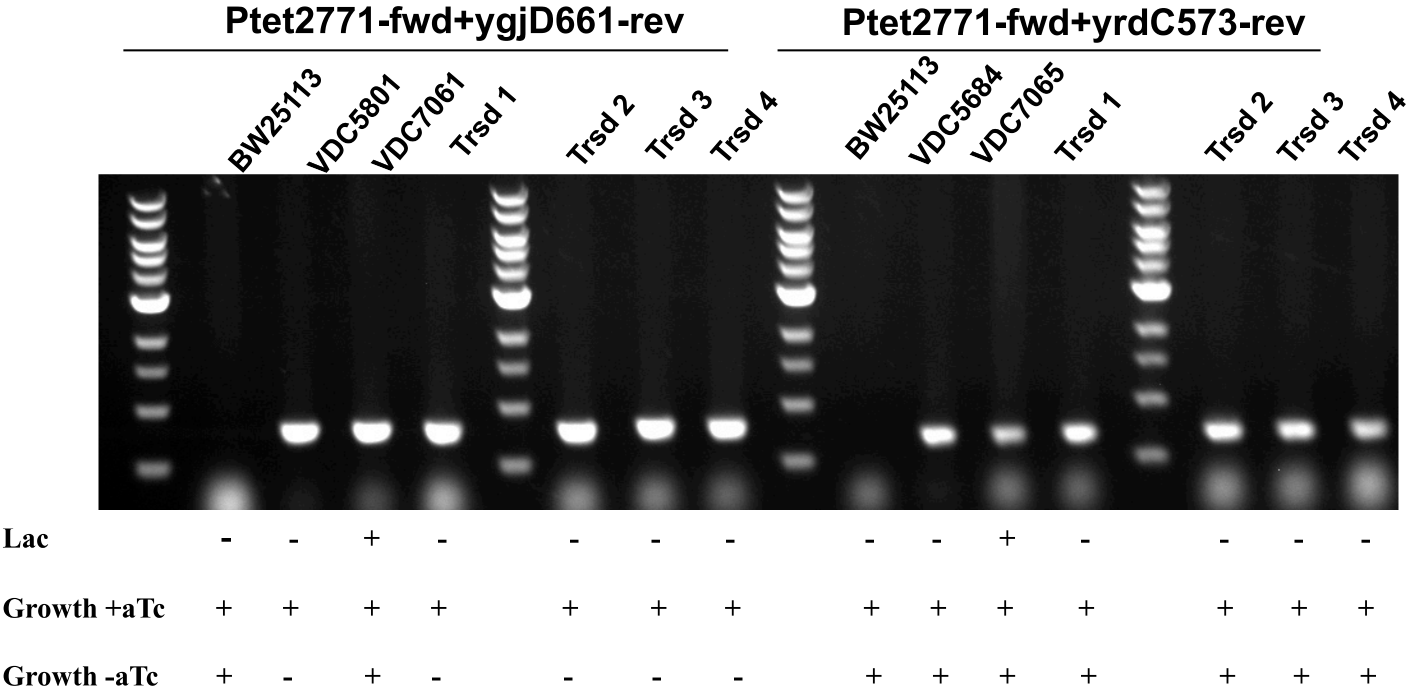

Figure S2. PCR verification of four independent clones (Trsd1-4) from P1 transduction from VDC7061 (P_{TET}::*tsaD* in *E. coli* B) and VDC7065 (P_{TET}::*tsaC* in *E. coli* B) back into *E. coli* BW25113. The primer pairs used are indicated on top of the respective figure panels. The growth phenotypes on LB +/- aTc and LB Xgal are reported below for each strain (donors are $LacZ^+$ and recipients and BW25113 are $LacZ^-$). Four transductants of P_{TET}::*tsaC* back into BW25113 remained aTc independent, suggesting that mutations in the P_{TET} promoter made it constitutive.

Figure S3

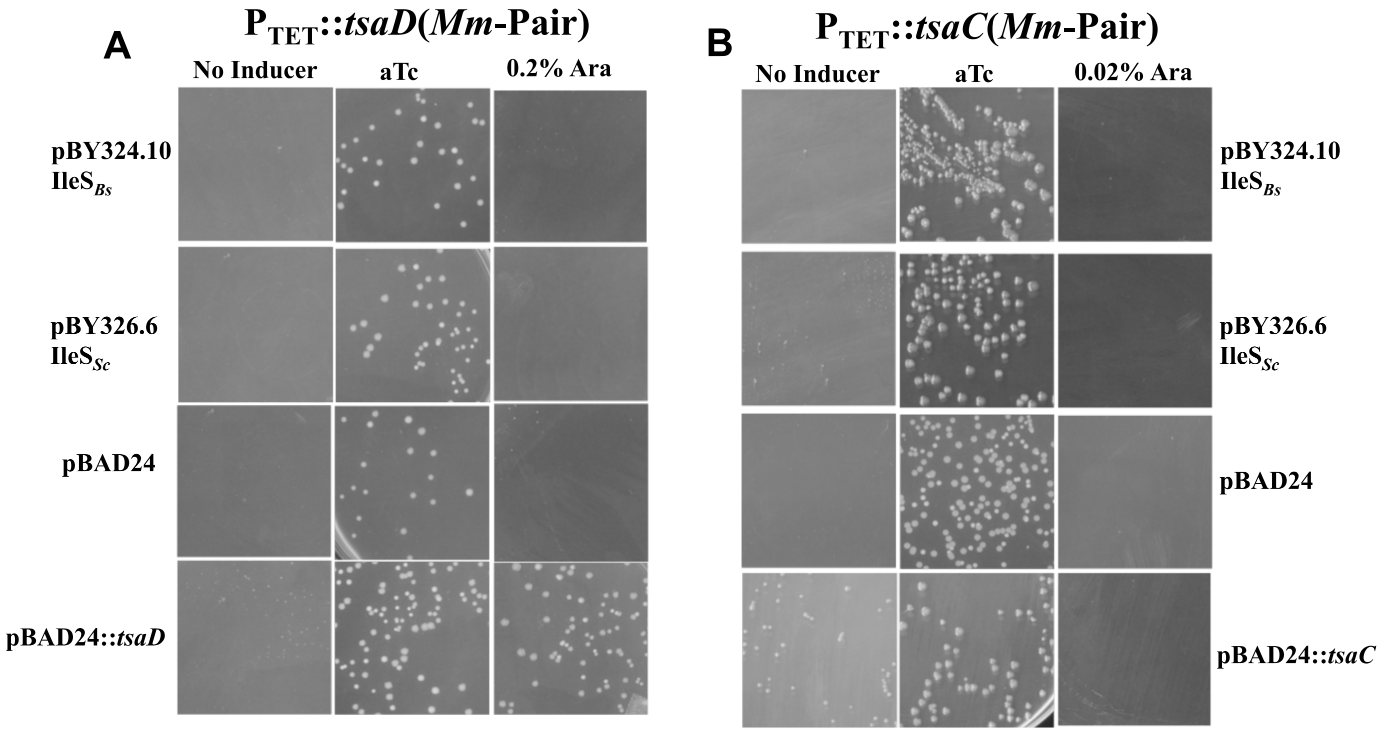

Figure S3. The t^6A essentiality phenotype of *E. coli* cannot be suppressed. *E. coli* strains P_{TET}::*tsaD* (A) and P_{TET}::*tsaC* (B) were equipped with a *Mycoplasma mobile* $tRNA^{Ile3}_{UAU}$::IleRS pair (*Mm*-Pair). For complementation assays, cells were transformed with 100 ng of plasmid (pBY324.10 containing the *B. subtilis* IleRS gene, pBY326.6 containing the *S. cerevisiae* IleRS gene, pBAD24 or pBAD::*tsaD*) via electroporation, then recovered with 1 mL of LB and placed at 37°C for 1 hour with shaking. 200 µL of cells were plated on LB supplemented with or without aTc (50 ng/ml) or arabinose.