Supplementary Information

[1]Medical Research Council Laboratory of Molecular Biology, Francis Crick Avenue, Cambridge, England, UK

Daniel L. Dunkelmann[1,2], Julian C. W. Willis[1,2], Jason W. Chin[1,*]

[1]Medical Research Council Laboratory of Molecular Biology, Francis Crick Avenue, Cambridge, England, UK

[2]These authors contributed equally.

[*]Correspondence: chin@mrc-lmb.cam.ac.uk

Methods

Identification of ΔNPylRS sequences

We identified PylRS protein sequences homologous to the C-terminal region of *Mm*PylRS or to *Desulfitobacterium hafniense* (*Dh)* PylSn by protein HMMER search[1] against the UniProtKB database[2] using *Mm*PylRSd184 or *Dh*PylSn as the query sequence respectively and filtering for expect values below 1×10^{-30}. From the identified protein sequences, which contain homology to the C-terminal region of *Mm*PylRS, we eliminated those for which sequence homology to *Dh*PylSn could be found within the same genome. From the remaining protein sequences, which correspond to ΔNPylRSs, we identified those that had not been previously reported.

Identification of ^{Pyl}tRNA sequences

Using the NCBI Nucleotide Database we used existing genome annotations to identify the DNA sequence for each PylRS gene within its host genome, and were thus able to identify the genomic region corresponding to the pyrrolysine gene cluster. In the sequences 40 kb upstream and downstream of the PylRS gene, $^{\Delta NPyl}$tRNA sequences were manually identified by searching for sequence similarity to known $^{\Delta NPyl}$tRNA sequences; $^{+NPyl}$tRNA sequences were manually identified by searching for sequence similarity to *Mm*^{Pyl}tRNA. tRNA secondary structure prediction was initially performed using *RNA structure*[3] and manually curated by inspection and comparison to *Mm*^{Pyl}tRNA.

DNA constructs

PylRS and ^{Pyl}tRNA genes were synthesised by IDT as gBlock double-stranded DNA fragments. We cloned the genes into pKW vectors by Gibson assembly. PylRS was expressed from a *glmS* promoter and ^{Pyl}tRNA was expressed from an *lpp* promoter.

PylRS genes were coded for expression in *E. coli* using the *IDT Codon Optimization Tool*. We appended the gene for *Mm*PylRS with a sequence encoding a C-terminal $Ser(Gly_4Ser)_4$FLAG-tag, while all other PylRS genes were appended with a sequence encoding a C-terminal $Ser(Gly_4Ser)_4His_6$SerGlyStrep-tag II. We used these plasmids together with pBAD GFP(150TAG)His_6 (in which sfGFP containing an amber stop codon at position 150 and a C-terminal His_6 tag is expressed from the arabinose promoter of pBAD; GFP refers to sfGFP throughout). We used Gibson cloning to insert each PylRS cassette under constitutive expression from the *glnS* promoter into pBAD CAT(111TAG) GFP(150TAG)His_6 vectors, in which a chloramphenicol acetyl transferase gene containing an amber stop codon at position 111 is under constitutive expression.

To create the plasmid pKW1-Triple for triple ncAA incorporation, PylRS genes were designed to be expressed as a single polycistronic mRNA transcript under the control of the glnS promoter, with RBS binding strengths of approximately 10,000 RBS units rationally designed using the RBS Calculator (https://www.denovodna.com/software/)[4-8] specifying Escherichia coli K-12 as the host organism. tRNA genes were designed to be expressed as a single polycistronic mRNA transcript under the control of the lpp promoter. Sequences between tRNAs were designed by manual examination of the E. coli K-12 MG1655 genome using EcoCyc[9] and identifying spacer sequences between tRNAs from the same isoacceptor class which are expressed as adjacent tRNAs in the same operon. Spacer sequences originating between AlaX and AlaW, and ValU and ValX genes were selected for use. Cassettes containing PylRS and PyltRNA genes were

synthesised by IDT as gBlock double-stranded DNA fragments. We cloned the genes into pKW vectors by Gibson assembly.

Library Generation

Libraries of *Int*$^{\Delta NPyl}$tRNA with randomised variable loop or acceptor stem sequences were constructed by Golden Gate cloning from a pKW *Int*$^{\Delta NPyl}$tRNA vector using PCR primers listed in **Supplementary Table 11** together with restriction enzyme BbsI and T4 DNA ligase.

We transformed each library separately into competent *E. coli* DH10B cells to give library diversities of more than 1×10^8, exceeding the theoretical diversity of 6×10^7 required for complete library coverage.

Selection to identify orthogonal Class A $^{\Delta NPyl}$tRNAs

For the variable loop library, we transformed each *Int*$^{\Delta NPyl}$tRNA variable loop library into competent *E. coli* DH10B cells bearing pBAD *Int*PylRS CAT(111TAG) GFP(150TAG)His_6. We recovered the transformed cells for 1 h at 37°C in 0.5 mL SOB medium supplemented with 8 mM BocK. The transformation was plated on LB agar containing 37.5 μg/mL spectinomycin, 12.5 μg/mL tetracycline and 100 μg/mL chloramphenicol. The plates were incubated at 37 °C for 40 h. After incubation, colonies on the plates were washed off and collected in 2XTY buffer and the plasmids were extracted by DNA midiprep (Qiagen kit). To remove the pBAD *Int*PylRS CAT(111TAG) GFP(150TAG)His_6 plasmid, the extracted DNA was digested with both NcoI restriction endonuclease and T5 exonuclease and re-purified using a PCR purification column. The

remaining pKW plasmids were transformed into competent *E. coli* DH10B cells bearing either pBAD *Alv*PylRS CAT(111TAG) GFP(150TAG)His_6 or pBAD *Mm*PylRS CAT(111TAG) GFP(150TAG)His_6. The transformed cells were recovered for 1 h at 37°C in 0.5 mL SOB medium. The transformation was plated on LB agar containing 37.5 μg /mL spectinomycin and 12.5 μg /mL tetracycline. The plates were incubated at 37°C for 20 h. For each library, 1,528 colonies were picked from the plates using a QPix 420 Colony Picking System and inoculated into 190 μl 2XTY-STA (2XTY medium with 75 μg/mL spectinomycin, 25 μg/mL tetracycline and 0.5% *L*-arabinose) in 96-well microtitre plate format supplemented with 8 mM BocK. The plates were incubated at 37°C and 220 rpm, and OD_{600} and GFP fluorescence (λ_{ex} 485 nm, λ_{em} 520 nm) measurements were recorded after 20 h using a SpectraMax i3. Cells from wells with the lowest GFP/OD_{600} ratios were used to inoculate 2XTY medium with 75 μg/mL spectinomycin, and the pKW plasmids containing $Int^{\Delta NPyl}$tRNA variants were extracted by DNA miniprep and then sequenced. Each hit corresponding to a distinct $Int^{\Delta NPyl}$tRNA sequence was cloned into a pKW *Int*PylRS vector, a pKW *Alv*PylRS vector, and a pKW *Mm*PylRS vector and re-phenotyped with pBAD GFP(150TAG)His_6.

For the acceptor stem library, we transformed each $Int^{\Delta NPyl}$tRNA acceptor stem library into competent *E. coli* DH10B cells bearing pBAD *Int*PylRS CAT(111TAG) GFP(150TAG)His_6. We recovered the transformed cells for 1 h at 37°C in 5 mL super optimal broth with catabolite repression (SOC) medium supplemented with 8 mM BocK. The transformation was plated on LB agar containing 75 μg/mL spectinomycin, 25 μg/mL tetracycline and 100 μg/mL chloramphenicol. The plates were incubated at 37°C for 24 h. After incubation, 192 colonies were picked into 1.7 mL 2XTY-STA in a 96-well

microtiter plate format supplemented with 8 mM BocK and grown overnight. Plasmids from all fluorescent cultures were extracted by DNA miniprep (Qiagen) and the extracted DNA was digested with both NcoI restriction endonuclease and T5 exonuclease. 1μL of the digestion products was transformed into chemically competent *E. coli* DH10B cells bearing either pBAD *Alv*PylRS CAT(111TAG) GFP(150TAG)His_6 or pBAD *Mm*PylRS CAT(111TAG) GFP(150TAG)His_6 by heat shock. The transformed cells were recovered for 1 h at 37°C in 180 μL ml SOC medium, and 10 μL was used to inoculate 180 μL 2XTY-STA in a 96-well microtiter plate format supplemented with 8 mM BocK and grown overnight. Cells from wells with the lowest GFP/OD_{600} ratios were used to inoculate 2XTY medium with 75 μg/mL spectinomycin, and the pKW plasmids containing $Int^{\Delta NPyl}$tRNA variants were extracted by DNA miniprep and then sequenced. Each hit corresponding to a distinct $Int^{\Delta NPyl}$tRNA sequence was cloned into a pKW *Int*PylRS vector, a pKW *Alv*PylRS vector, and a pKW *Mm*PylRS vector and re-phenotyped with pBAD GFP(150TAG)His_6.

Measuring the activity and specificity of PylRS/^{Pyl}tRNA$_{CUA}$ pairs with synthetase and tRNA expressed from different plasmids

To measure the activity and specificity of cognate and non-cognate PylRS/^{Pyl}tRNA combinations we transformed 0.4 μL of pKW ^{Pyl}tRNA plasmids into 8 μL chemically competent *E. coli* DH10B cells bearing either pBAD GFP(150TAG)His_6 or pBAD PylRS GFP(150TAG)His_6. We recovered the transformed cells for 1 h at 37°C and 750 rpm in 180 μL SOC medium in 96-well microtiter plate format. 10 μL of the transformed cells was used to inoculate 180 μL 2XTY-STA in 96-well microtiter plate format, supplemented with or without 8 mM BocK. OD_{600} and GFP fluorescence (λ_{ex} 485 nm,

λ_{em} 520 nm) measurements were recorded after 22-28 h incubation at 37°C and 700 rpm using a Tecan Infinite M200 Pro.

Measuring the activity and specificity of PylRS/^{Pyl}tRNA$_{CUA}$ pairs with synthetase and tRNA expressed from the same plasmid

The same procedure was followed as described above. However, for this expression system both PylRS and ^{Pyl}tRNA were encoded on the same pKW plasmid which was transformed into chemically competent *E. coli* DH10B cells bearing pBAD GFP(150TAG)His$_6$. For GFP expression, 25 µL of transformed cells was inoculated into 500 µL 2XTY-STA in 96-well microtiter plate format, in the presence or absence of 8 mM BocK, or 2 mM CbzK or 8 mM NmH. Cells were grown for 22-28 h at 750 rpm and 37˚C before OD_{600} 180 µL of each well were transferred to a 96 well plate and GFP fluorescence measurements were recorded as described above.

GFP(TAG)$_{His6}$ Expression for Mass Spectrometry

To express GFP incorporating BocK for mass spectrometry analysis we transformed pKW PylRS/^{Pyl}tRNA plasmids into competent *E. coli* DH10B cells bearing pBAD GFP(150TAG)His$_6$. We recovered the transformed cells for 1 h at 37°C in 1 mL SOC medium. The transformation was used to inoculate 20 mL 2XTY-STA supplemented with 8 mM BocK and incubated overnight at 37°C and 220 rpm for 20 h.

20 ml culture was pelleted by centrifugation and washed with 2 mL PBS. The cell pellets were resuspended in 1 mL lysis buffer (1X BugBuster Protein Extraction Reagent supplemented with 1X complete protease inhibitor cocktail, 1 mg /mL lysozyme and 1 mg /mL DNase I) and lysed for 1 h at 25°C with head-over-tail circular rotation. The lysate was clarified by centrifugation (21,000 g, 30 min, 4°C). GFP was purified by its C-

terminal His_6 tag using 75 µL Ni-NTA agarose beads and left to bind for 30 min at room temperature. The beads were washed five times with 1 mL PBS supplemented with 10 mM imidazole and eluted in 40 µL PBS supplemented with 250 mM imidazole.

The same procedure was used to assess the active site orthogonality of *1R26*PylRS(Cbz)*/Ma*PyltRNA(11)$_{CUA}$, *Lum1*PylRS(NmH)/IntPyltRNA(^{A}13,^{V}C10)$_{CUA}$ or *Mm*PylRS/*Spe*PyltRNA$_{CUA}$ but all three amino acids were added to the medium (8mM BocK, 2 mM CbzK, 8 mM NmH) simultaneously.

The eluting fraction was diluted and analysed by time of flight mass spectrometry.

O-GST-CAM(1XXX)$_{His6}$ Expression for SDS PAGE

To express O-GST-CaM(1XXX) proteins (X=TAG, AGGA or AGTA) we co-transformed competent *E. coli* DH10B cells with pKW-Triple *Mm*PylRS/*Spe*^{Pyl}tRNA$_{CUA}$, *Lum1*PylRS(NmH)/*Int*$^{\Delta NPyl}$tRNA(^{A}17,^{V}C10)$_{UCCU}$ and *1R26*PylRS(CbzK)/*Alv*$^{\Delta NPyl}$tRNA(8)$_{UACU}$, pRSF ribo-Q1 and pCUN O-GST-CaM(1XXX). We recovered the transformed cells for 1 h at 37°C in 1 mL SOC medium. The transformation was used to inoculate 5 mL 2XTY-KST (2XTY medium with 25 µg mL^{-1} kanamycin, 75 µg mL^{-1} spectinomycin, and 12.5 µg mL^{-1} tetracycline) and incubated overnight (37°C, 16 h, 220 rpm). 50 µL of the overnight culture was diluted in 5 mL 2XTY-KST containing a combination of the indicated ncAAs (8 mM BocK **1**, 8 mM NmH **2** and 2 mM CbzK **3**) or none of them and incubated at 37°C, 220 rpm. At OD_{600} 0.6, 50 µL 1M IPTG was added to a final concentration of 1 mM. After 16 h incubation at 37°C, 220 rpm, the cultures were pelleted and washed with 800 µL PBS. The cell pellets were resuspended in 1 mL of lysis buffer (1X BugBuster Protein Extraction Reagent supplemented with 1X cOmplete protease inhibitor cocktail) and

lysed for 1 h at 25 °C with head over tail rotation. The lysate was clarified by centrifugation (21,000 g, 30 min, 4°C). GST-containing proteins from the lysate supernatant were left to bind to 60 µL

glutathione sepharose beads for 1 h at 25 ˚C. The beads were washed five times with

800 µL PBS before eluting in 60 µL 20 mM reduced glutathione in PBS pH 8. Samples were analysed on 4-12% Bis-Tris SDS-PAGE gels, visualised with InstantBlue Coomassie stain and imaged using a ChemiDoc Touch Imaging System.

O-$_{Strep}$GFP(40TAG, 136AGGA, 150AGTA)$_{His6}$ Expression for Western Blot Analysis and Mass Spectrometry

To express O-$_{Strep}$GFP(40TAG, 136AGGA, 150AGTA or wt)$_{His6}$ we co-transformed competent *E. coli* DH10B cells with pKW-Triple *Mm*PylRS/*Spe*^{Pyl}tRNA$_{CUA}$, *Lum1*PylRS(NmH)/*Int*$^{\Delta NPyl}$tRNA(^{A}17,^{V}C10)$_{UCCU}$ and *1R26*PylRS(CbzK)/*Alv*$^{\Delta NPyl}$tRNA(8)$_{UACU}$, pRSF ribo-Q1 and pCUN O-$_{Strep}$GFP(40TAG, 136AGGA, 150AGTA)$_{His6}$.We recovered the transformed cells for 1 h at 37°C in 1 mL SOC medium. The transformation was used to inoculate 20 mL 2XTY-KST and incubated overnight (37°C, 16 h, 220 rpm). 1 mL of the overnight culture was diluted in 50 mL 2XTY-KST containing a combination of the indicated ncAAs (8 mM BocK **1**, 8 mM NmH **2** and 2 mM CbzK **3**) or none of them and incubated at 37°C, 220 rpm. At OD_{600} 0.6, 500 µL 1M IPTG was added to a final concentration of 1 mM. After 16 h incubation at 37°C, 220 rpm, the cultures were pelleted and washed with 5 mL PBS. The cell pellets were resuspended in 5 mL of lysis buffer (1X BugBuster Protein Extraction Reagent supplemented with 1X cOmplete protease inhibitor cocktail) and lysed for 1 h at 25 °C with head over tail rotation. The lysate was clarified by centrifugation (21,000 g,

30 min, 4°C). GFP-containing proteins from the lysate supernatant were left to bind to 80 µL

Ni-NTA beads for 1 h at 25 ˚C. The beads were washed five times with 800 µL PBS containing 25 mM imidazole before eluting in 80 µL PBS containing 250 mM imidazole. Samples were analysed by western blot using 4-12% Bis-Tris SDS-PAGE gels, primary antibody rabbit anti-Strep ab76949 (Abcam) and secondary antibody goat anti-rabbit IRDye 800CW (LI-COR).

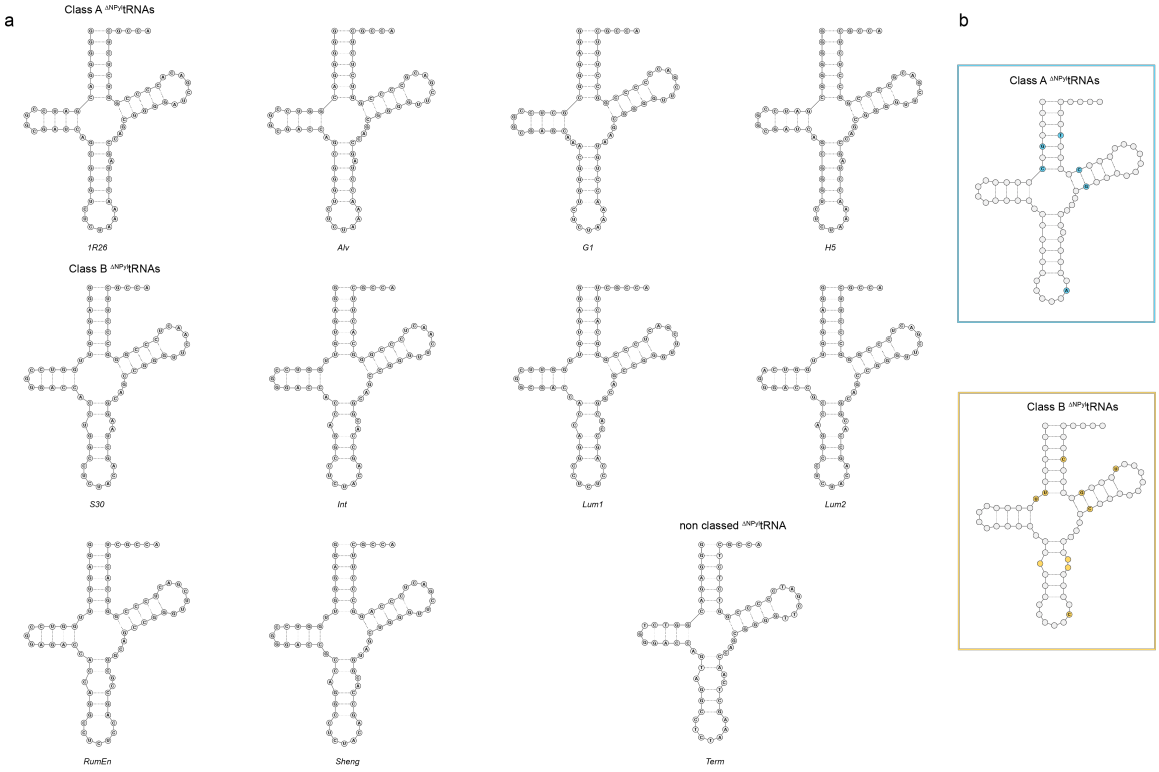
To obtain the mass spectrum of O-$_{Strep}$GFP(40TAG, 136AGGA, 150AGTA)$_{His6}$ we combined the purified protein fractions from three biological replicates expressed in the presence of all three ncAAs and incubated them overnight with 100 µL StrepTactin sepharose beads. Beads were washed five times with 800 µL PBS pH 8 and eluted three times in 100 µL 10 mM desthiobiotin pH 8. Fractions were combined, concentrated and analysed by time of flight mass spectrometry.

Electrospray ionization mass spectrometry

Denatured protein samples (~10µM) were subjected to LC-MS analysis. Briefly, proteins were separated on a C4 BEH 1.7µm, 1.0 x 100mm UPLC column (Waters, UK) using a modified nanoAcquity (Waters, UK) to deliver a flow of approximately 50 µl/min. The column was developed over 20 minutes with a gradient of acetonitrile (2% v/v to 80% v/v) in 0.1% v/v formic acid. The analytical column outlet was directly interfaced via an electrospray ionisation source, with a hybrid quadrupole time-of-flight mass spectrometer (Xevo G2, Waters, UK). Data was acquired over a m/z range of 300–2000, in positive ion mode with a cone voltage of 30V. Scans were summed together manually and deconvoluted using MaxEnt1 (Masslynx, Waters, UK). The theoretical molecular weights

of proteins with ncAAs was calculated by first computing the theoretical molecular weight of wild-type protein using an online tool (http://web.expasy.org/protparam/) and then manually correcting for the theoretical molecular weight of ncAAs.

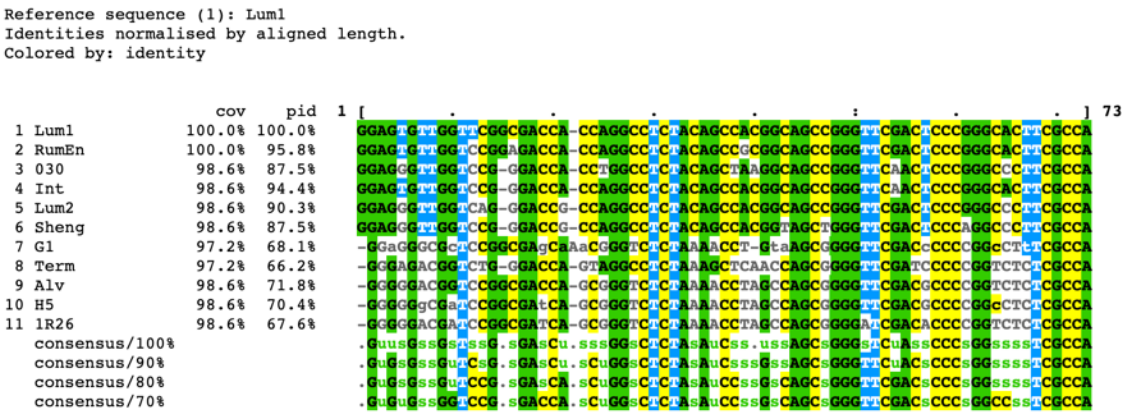
Supplementary Figures



Supplementary Figure 1

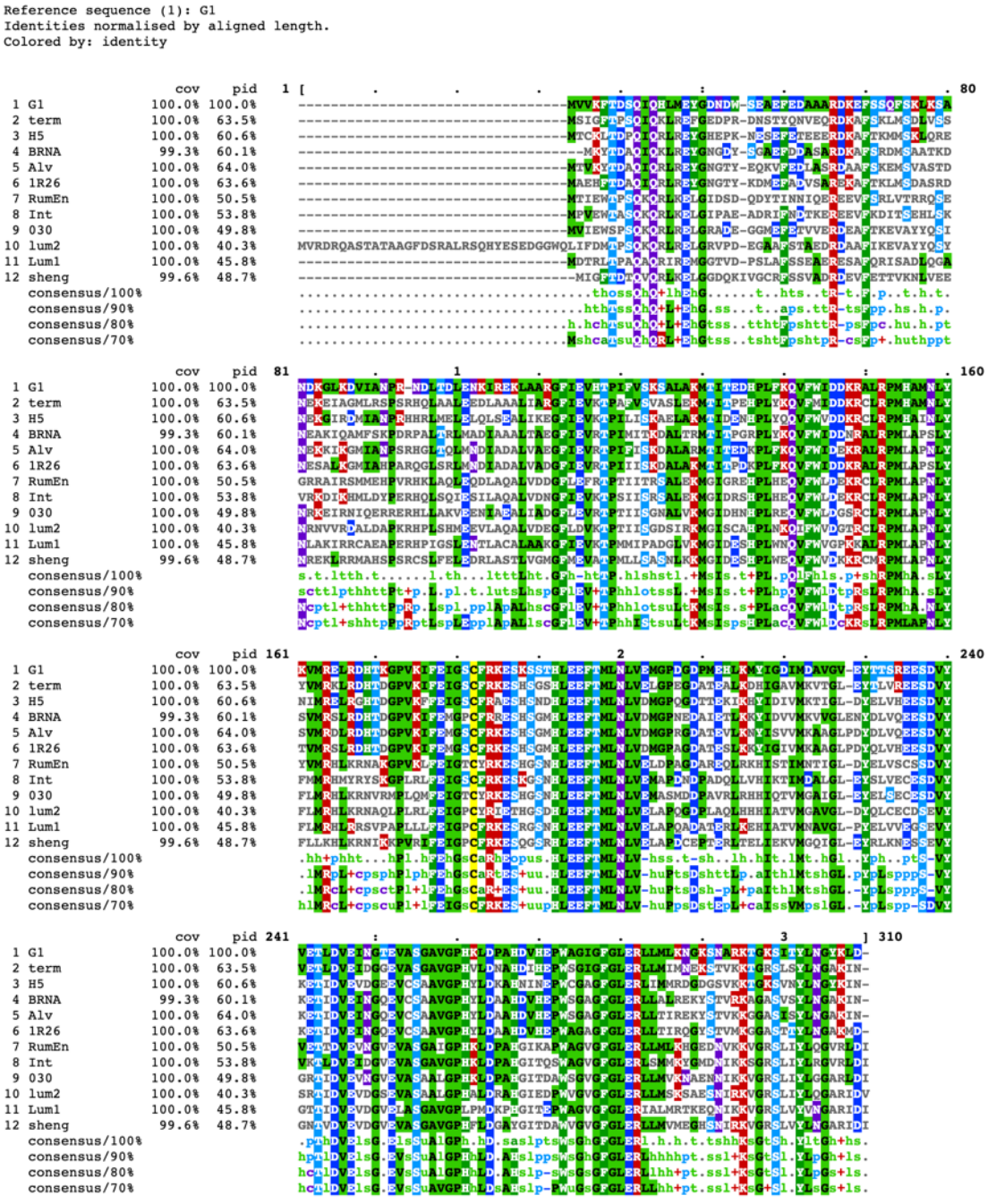
a, Predicted clover leaf structure of all $^{\Delta NPyl}$tRNAs. $^{\Delta NPyl}$tRNAs are grouped into Class A, Class B and non classed $^{\Delta NPyl}$tRNAs. **b**, Predicted clover leaf structure with all nucleotides that are conserved within each class of $^{\Delta NPyl}$tRNAs but differ between sequence Class A and sequence Class B $^{\Delta NPyl}$tRNAs highlighted in blue (Class A) and yellow (Class B). Sequence Class A $^{\Delta NPyl}$tRNAs (except *G1* and *Term*) were defined by a nucleotide bulge in the anticodon stem whereas sequence Class B $^{\Delta NPyl}$tRNAs and

Term$^{\Delta NPyl}$tRNA instead exhibit a nucleotide loop at this location. Sequence Class B $^{\Delta NPyl}$tRNAs were also defined by the insertion of an additional uracil nucleotide between the acceptor stem and the D-stem. *GI*$^{\Delta NPyl}$tRNA uniquely harbors an additional adenine nucleotide between the D arm and the anticodon stem and has no bulge in the anticodon stem, which differentiates it from all other $^{\Delta NPyl}$tRNAs in this study.



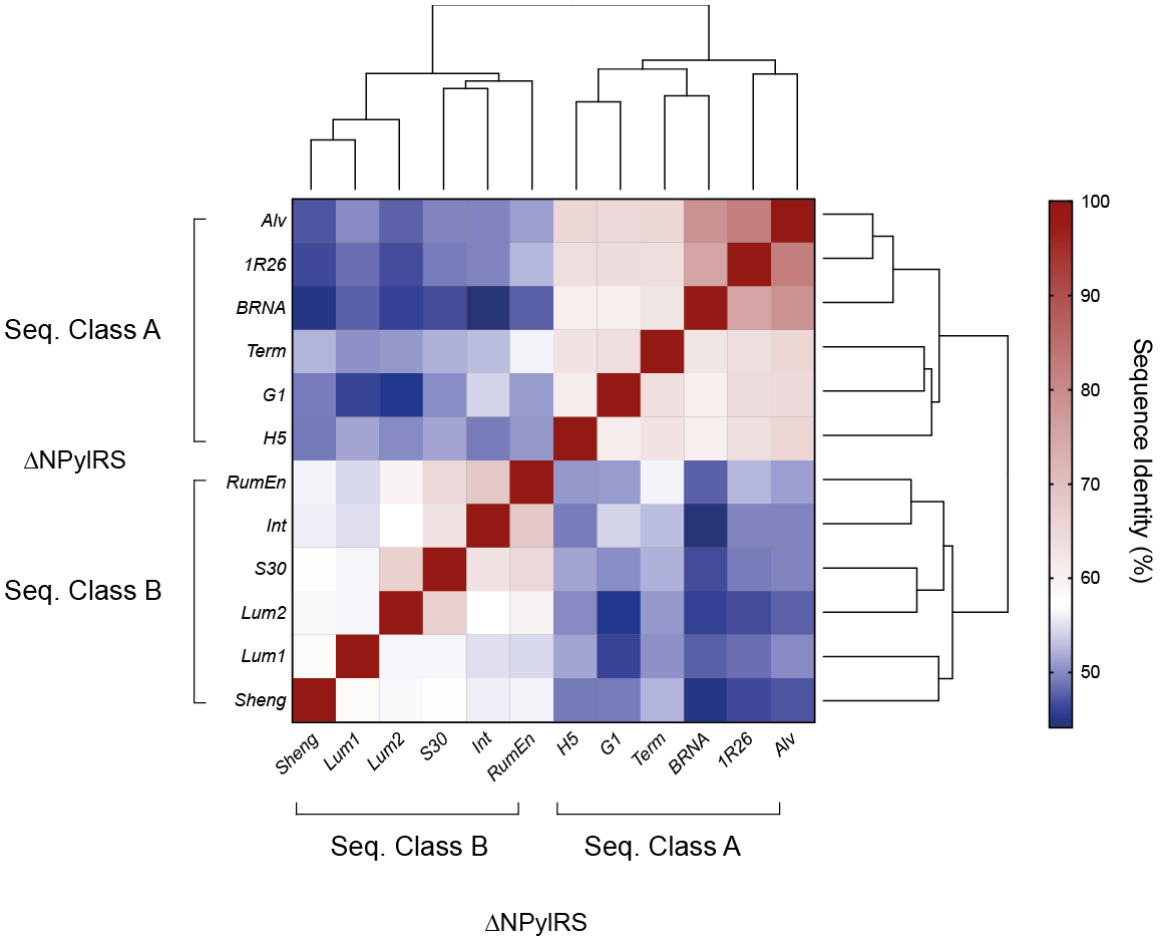
Supplementary Figure 2

Sequence alignment of all $^{\Delta NPyl}tRNA_{CUA}$s generated with *Clustal Omega*.



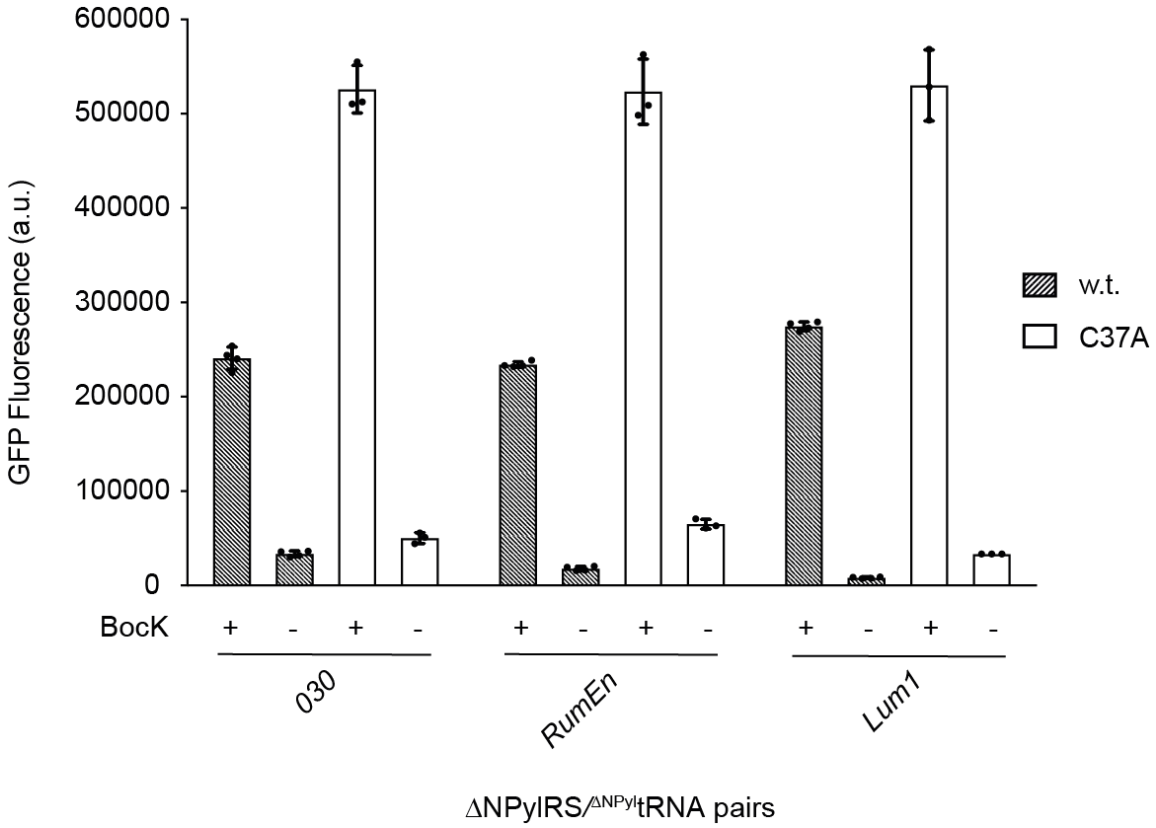
Supplementary Figure 3

Sequence alignment of all ΔNPylRSs generated with *Clustal Omega*.



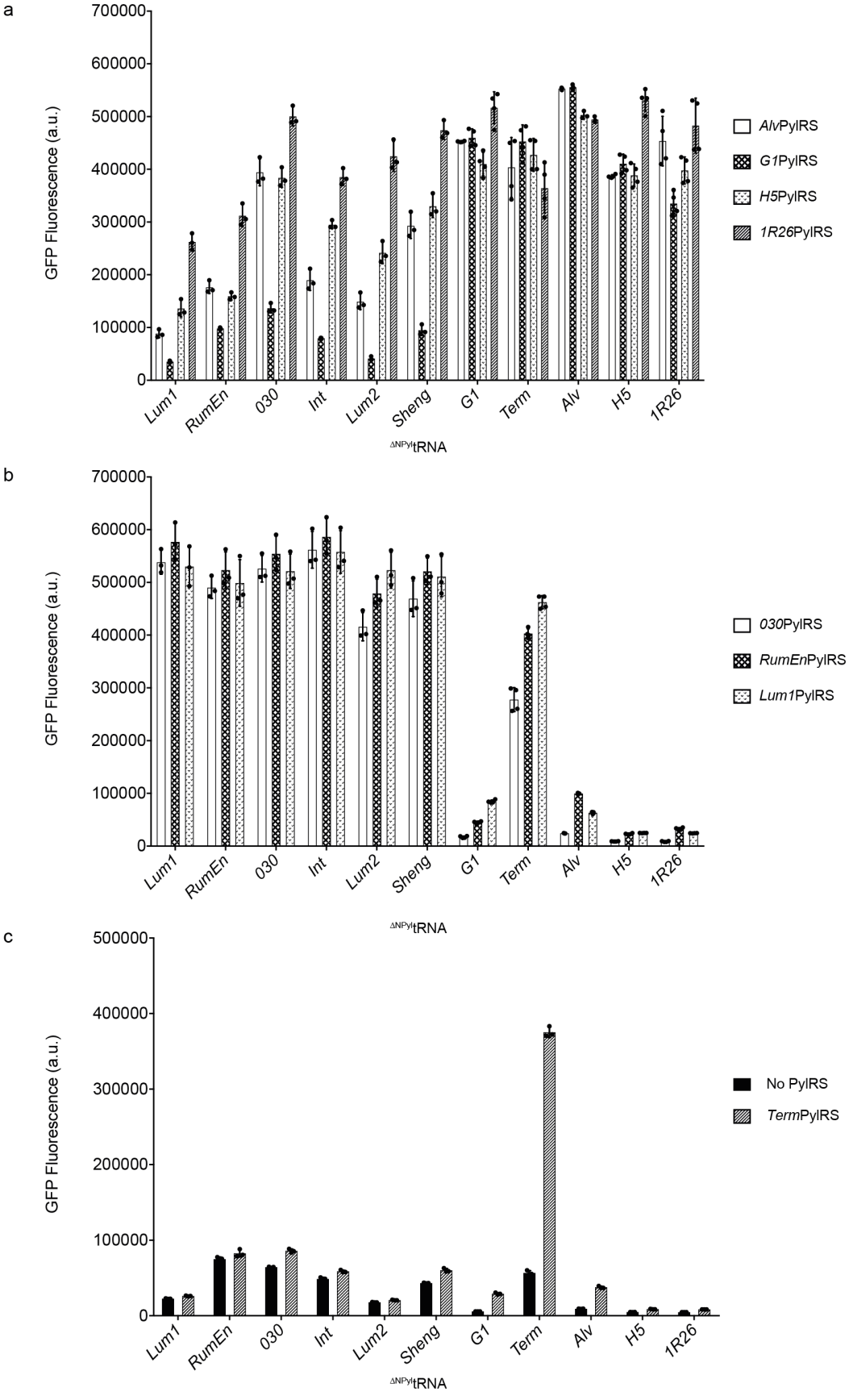
Supplementary Figure 4

Hierarchical clustering of ΔNPylRS with sequence similarity scores converted to Euclidean distance measures and complete linkage clustering in the program *R Studios*. Percentage sequence identity scores are displayed as a heatmap. The dendrograms resulting from the clustering demonstrate the grouping of the ΔNPylRSs in two sequence-dependent classes.



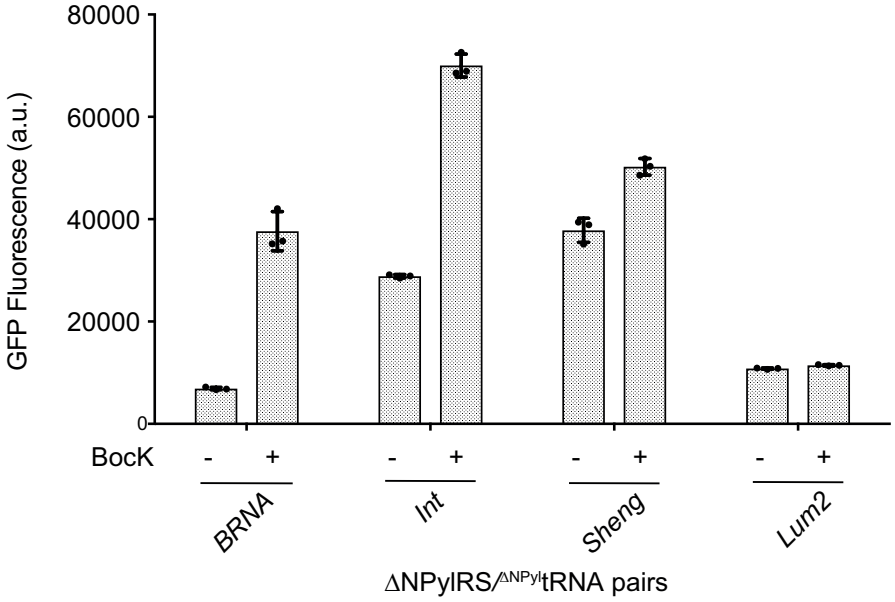
Supplementary Figure 5

Comparison of Class B ΔNPylRS/$^{\Delta NPyl}$tRNA$_{CUA}$ pairs with each tRNA having either the wild type (w.t.) anticodon loop sequence or harbouring a C37A mutation. *In vivo* suppression activity was assayed in *E. coli* DH10B bearing pBAD GFP(150TAG)His$_6$ ΔNPylRS and pKW $^{\Delta NPyl}$tRNA$_{CUA}$ in the presence and absence of BocK. Each bar chart represents four (for wild type (w.t.) anticodon loop sequence $^{\Delta NPyl}$tRNA$_{CUA}$) or three (C37A mutation $^{\Delta NPyl}$tRNA$_{CUA}$) biological replicates with error bars showing std..


a
GFP Fluorescence (a.u.)
700000
600000
500000
400000
300000
200000
100000
0
Lum1
RumEn
030
Int
Lum2
Sheng
G1
Term
Alv
H5
1R26
ΔNPyltRNA
AlvPylRS
G1PylRS
H5PylRS
1R26PylRS
b
GFP Fluorescence (a.u.)
700000
600000
500000
400000
300000
200000
100000
0
Lum1
RumEn
030
Int
Lum2
Sheng
G1
Term
Alv
H5
1R26
ΔNPyltRNA
030PylRS
RumEnPylRS
Lum1PylRS
c
GFP Fluorescence (a.u.)
500000
400000
300000
200000
100000
0
Lum1
RumEn
030
Int
Lum2
Sheng
G1
Term
Alv
H5
1R26
ΔNPyltRNA
No PylRS
TermPylRS

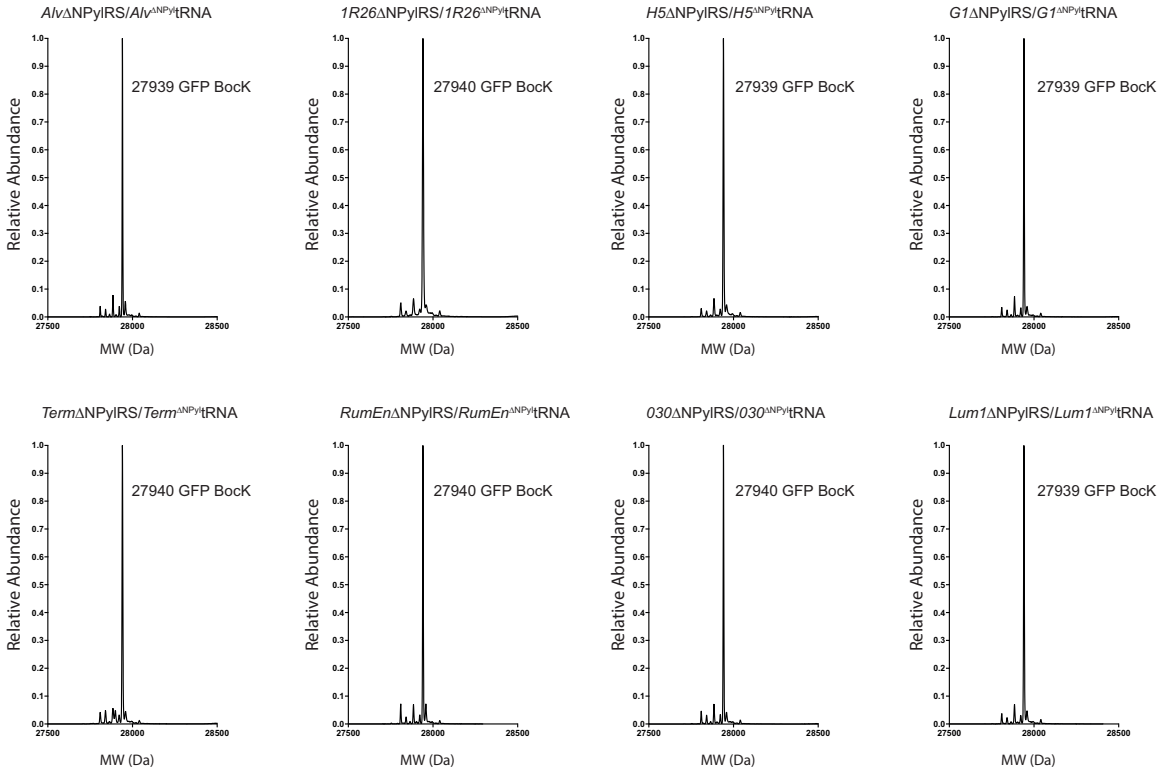
Supplementary Figure 6

In vivo amber suppression activity of homologous and heterologous combinations of ΔNPylRS/$^{\Delta NPyl}$tRNA$_{CUA}$ pairs. *In vivo* suppression activity was assayed in *E. coli* DH10B bearing pBAD GFP(150TAG)His$_6$ ΔNPylRS and pKW $^{\Delta NPyl}$tRNA$_{CUA}$ in the presence of BocK. Each bar represents an average of three (Class B $^{\Delta NPyl}$tRNA$_{CUA}$s, *Alv*$^{\Delta NPyl}$tRNA$_{CUA}$, *Term*$^{\Delta NPyl}$tRNA$_{CUA}$ and no PylRS) or four (Class A $^{\Delta NPyl}$tRNA$_{CUA}$s except *Alv*$^{\Delta NPyl}$tRNA$_{CUA}$s) biological replicates and error bars show the std. **a**, All combinations of Class A ΔNPylRSs with each $^{\Delta NPyl}$tRNA$_{CUA}$. **b**, All combinations of Class B ΔNPylRSs with each $^{\Delta NPyl}$tRNA$_{CUA}$. **c**, All combinations of *Term*PylRS with each $^{\Delta NPyl}$tRNA$_{CUA}$, and each $^{\Delta NPyl}$tRNA$_{CUA}$ in the absence of any ΔNPylRS.



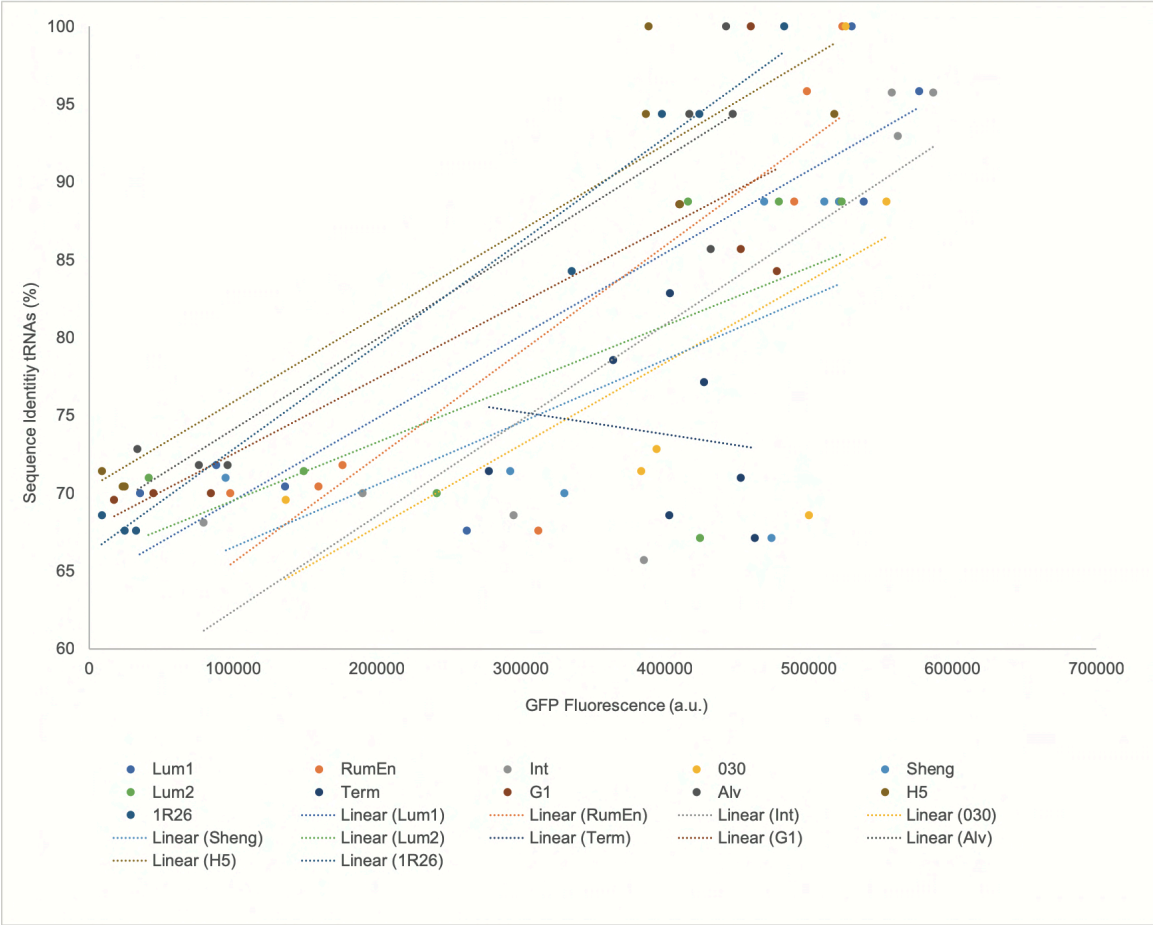
Supplementary Figure 7

In vivo amber suppression activity of homologous combinations of ΔNPylRS/$^{\Delta NPyl}$tRNA$_{CUA}$ pairs, which were deemed not active and or orthogonal enough for further experiments. *In vivo* suppression activity was assayed in *E. coli* DH10B bearing pBAD GFP(150TAG)His$_6$ ΔNPylRS and pKW $^{\Delta NPyl}$tRNA$_{CUA}$ in the presence of BocK. Each bar represents an average of three biological replicates and error bars show the std.



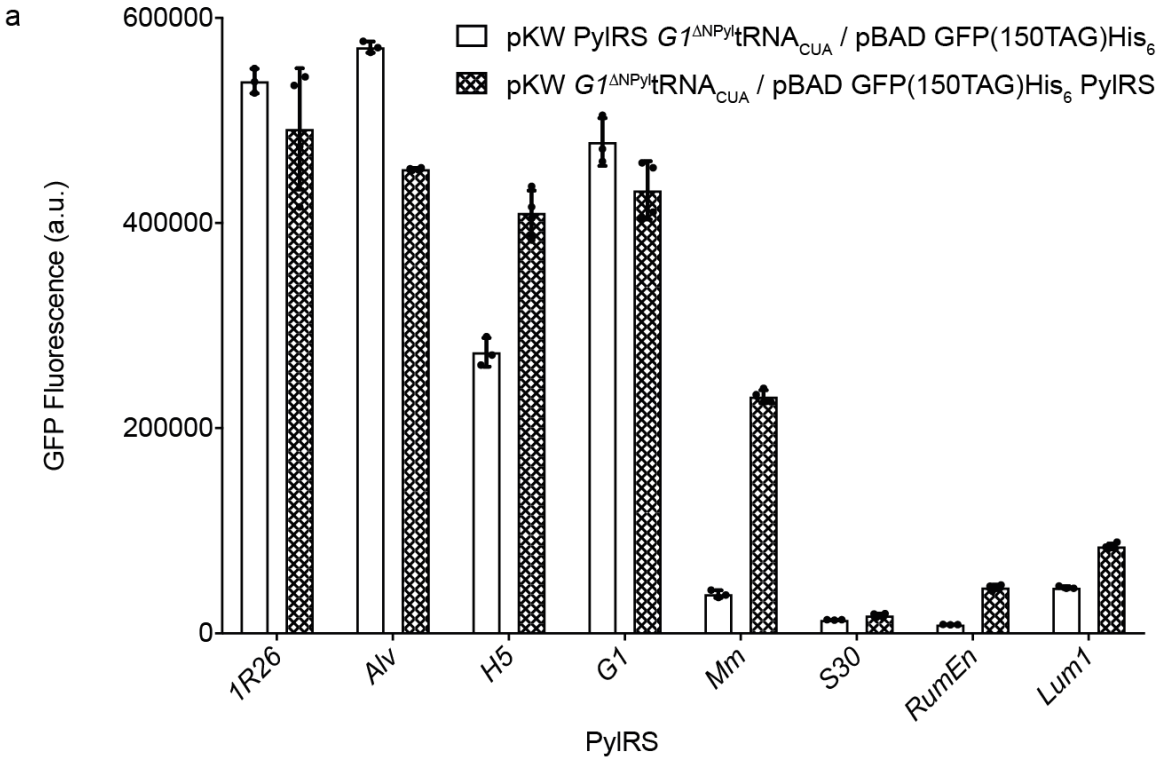
Supplementary Figure 8

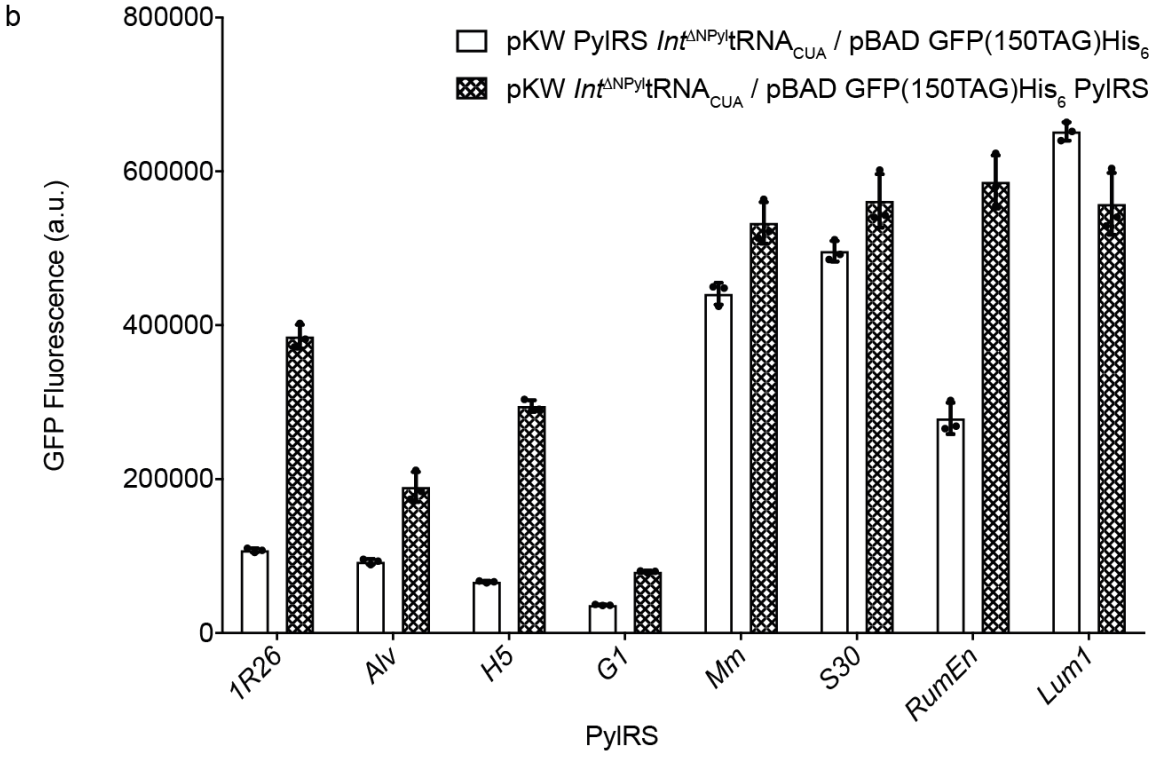
TOP-MS data of GFPHis$_6$ purified from *E. coli* DH10B bearing pBAD GFP(150TAG)His$_6$ ΔNPylRS and pKW $^{\Delta NPyl}$tRNA$_{CUA}$ in the presence of BocK. The main peaks correspond to site-specific BocK incorporation (predicted mass 27,941 Da, observed masses 27,938 - 27,940 Da). No peak corresponding to the misincorporation of BocK in place of any canonical amino acid was observed.



Supplementary Figure 9

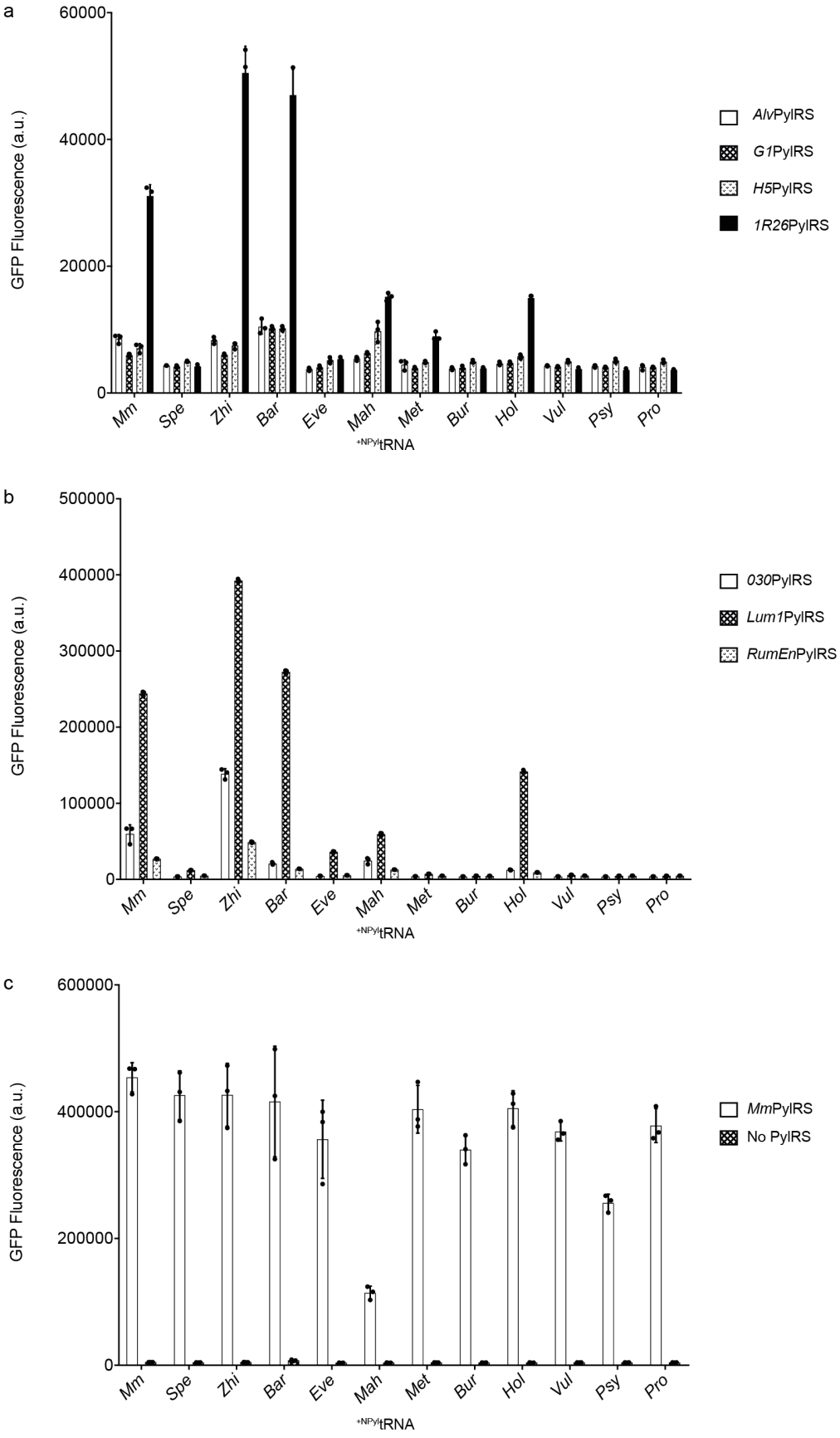
Linear correlation between the sequence identity scores of homologous and heterologous $^{\Delta NPyl}$tRNA combinations, and the functional activity of homologous and heterologous ΔNPylRS/$^{\Delta NPyl}$tRNA$_{CUA}$ pairs as determined by *in vivo* suppression activity. The same sequence identity scores and *in vivo* amber suppression data are displayed as described in **Fig 1**. All ΔNPylRSs apart from *Term*PylRS show a positive correlation between sequence identity and amber suppression activity.


a
600000
400000
200000
0
GFP Fluorescence (a.u.)
pKW PylRS G1ΔNPyltRNACUA / pBAD GFP(150TAG)His6
pKW G1ΔNPyltRNACUA / pBAD GFP(150TAG)His6 PylRS
1R26
Aiv
H5
G1
Mm
S30
RumEn
Lum1
PylRS


b
800000
600000
400000
200000
0
GFP Fluorescence (a.u.)
pKW PylRS IntΔNPyltRNACUA / pBAD GFP(150TAG)His6
pKW IntΔNPyltRNACUA / pBAD GFP(150TAG)His6 PylRS
1R26
Aiv
H5
G1
Mm
S30
RumEn
Lum1
PylRS

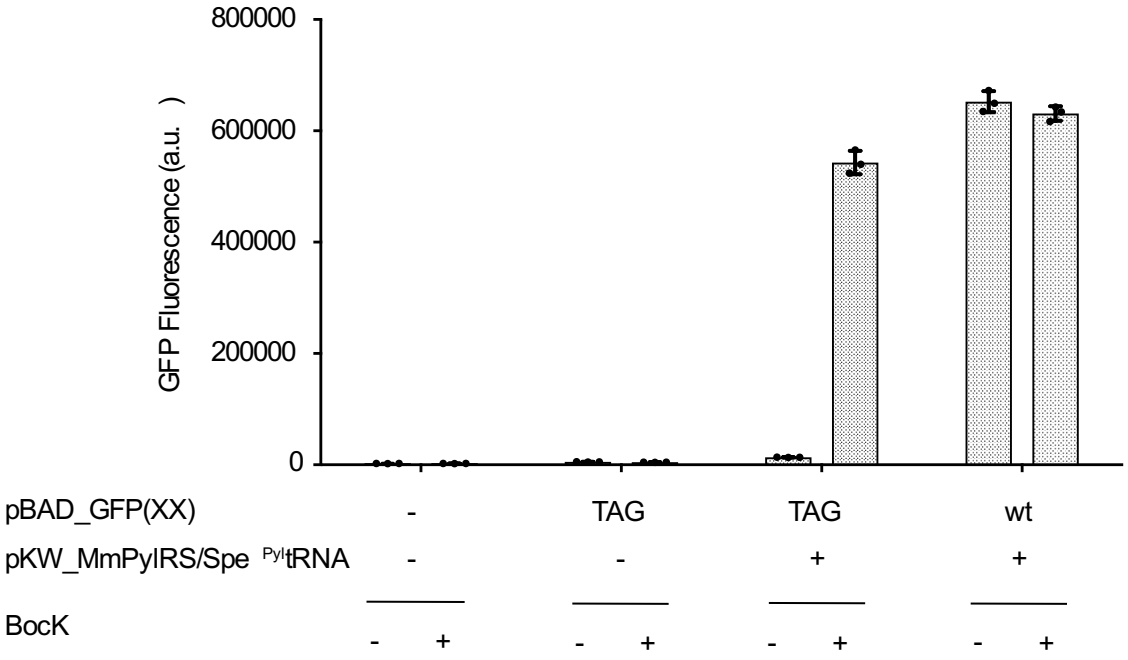
Supplementary Figure 10

a, Comparison of *in vivo* amber suppression activity for *GI*$^{\Delta NPyl}$tRNA$_{CUA}$ between the two expression systems used in this study. The same amber suppression data is displayed as described in **Fig. 1, 2** and **Supplementary Fig. 6, 16, 18**. In one expression system (pKW $^{\Delta NPyl}$tRNA$_{CUA}$, pBAD GFP(150TAG)His$_6$ PylRS) the tRNA and synthetase are expressed from different plasmids, while in the other expression system (pKW PylRS $^{\Delta NPyl}$tRNA$_{CUA}$ pBAD GFP(150TAG)His$_6$) the tRNA and synthetase are both expressed from the same plasmid. In both cases GFP(150TAG)His$_6$ is expressed from a pBAD plasmid. Both data sets follow the same trends, however, when PylRSs are expressed from pBAD we note a general increase in activity for the less active PylRS/$^{\Delta NPyl}$tRNA$_{CUA}$ pairs. **b**, Comparison of *in vivo* amber suppression activity for *Int*$^{\Delta NPyl}$tRNA$_{CUA}$ between the two expression systems used in this study, as in **a.**


a
60000
40000
20000
0
GFP Fluorescence (a.u.)
Mm
Spe
Zhi
Bar
Eve
Mah
Met
Bur
Hol
Vul
Psy
Pro
+NPyltRNA
AlvPylRS
G1PylRS
H5PylRS
1R26PylRS
b
500000
400000
300000
200000
100000
0
GFP Fluorescence (a.u.)
Mm
Spe
Zhi
Bar
Eve
Mah
Met
Bur
Hol
Vul
Psy
Pro
+NPyltRNA
030PylRS
Lum1PylRS
RumEnPylRS
c
600000
400000
200000
0
GFP Fluorescence (a.u.)
Mm
Spe
Zhi
Bar
Eve
Mah
Met
Bur
Hol
Vul
Psy
Pro
+NPyltRNA
MmPylRS
No PylRS

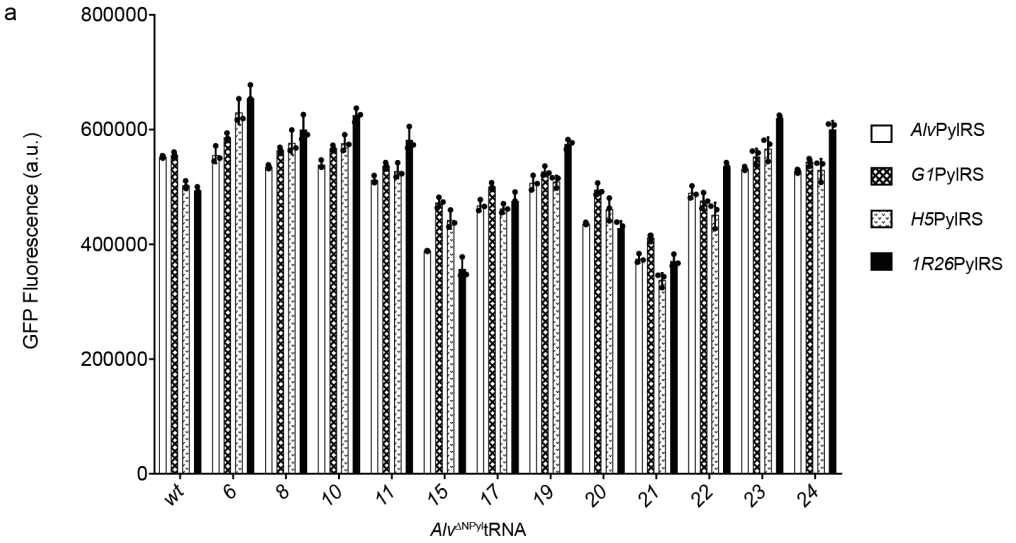
Supplementary Figure 11

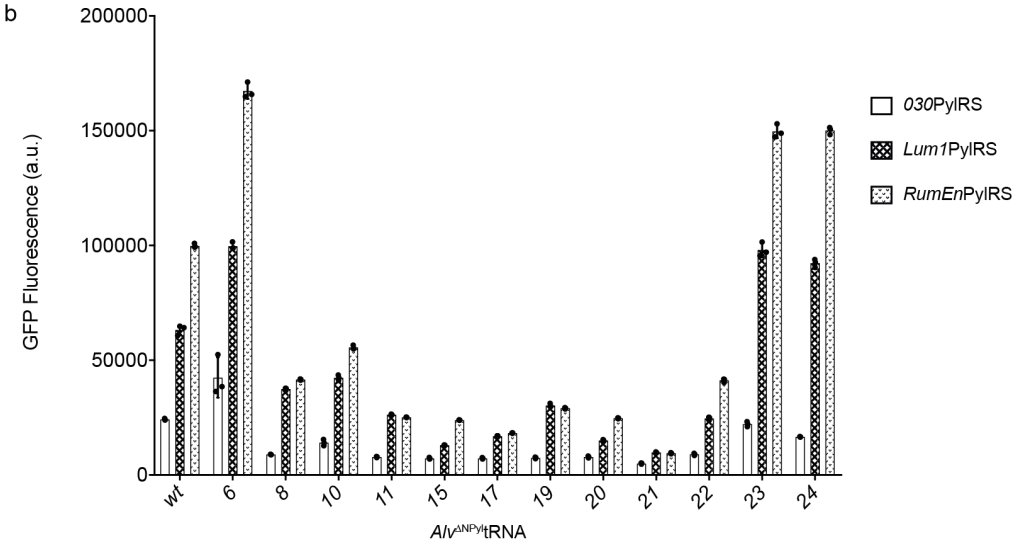
In vivo amber suppression activity of each $^{+NPyl}tRNA_{CUA}$ with each PylRS. Measurements were made in *E. coli* DH10B bearing pBAD GFP(150TAG)His_6 PylRS and pKW $^{+NPyl}tRNA_{CUA}$ in the presence of BocK. Each bar represents an average of three biological replicates and error bars show the std.. **a**, All combinations of Class A ΔNPylRSs with each $^{+NPyl}tRNA_{CUA}$. **b**, All combinations of Class B ΔNPylRSs with each $^{+NPyl}tRNA_{CUA}$. **c**, All combinations of *Mm*PylRS with each $^{+NPyl}tRNA_{CUA}$, and each $^{+NPyl}tRNA_{CUA}$ in the absence of any PylRS.

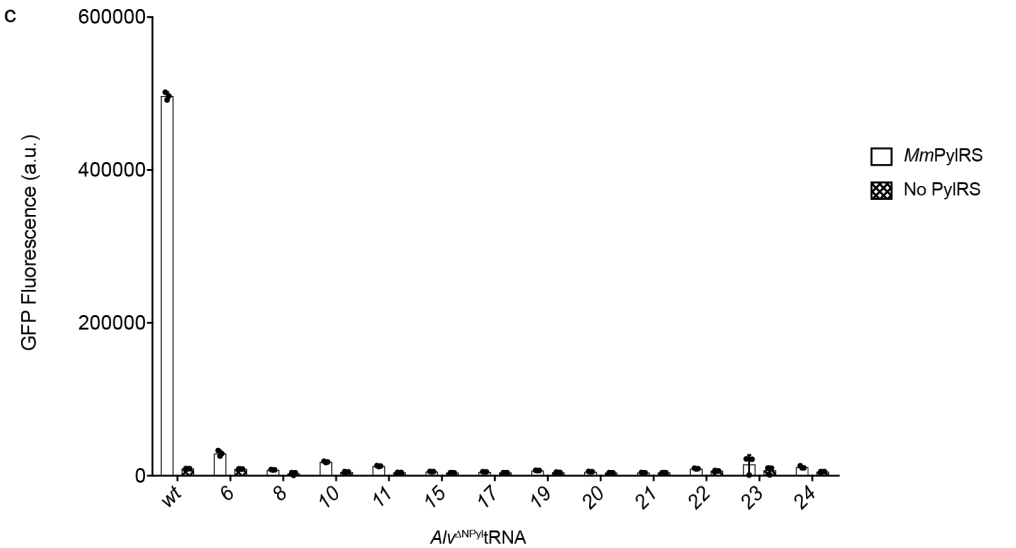


Supplementary Figure 12

Comparison of *in vivo* amber suppression activity of *Mm*PylRS/*Spe*^{Pyl}tRNA with fluorescence of the empty DH10B cells, the reporter GFP(150TAG)His_6 in absence of synthetase and tRNA and GFP(wt)His_6. *In vivo* fluorescence was assayed in *E. coli* DH10B without any plasmid, with just pBAD GFP(150TAG)His_6 or with a combination of pBAD GFP(150TAG)His_6/pBAD GFP(wt)His_6 and pKW *Spe*^{Pyl}$tRNA_{CUA}$. Each experiment was run in presence or absence of BocK. Each bar represents an average of three biological replicates and error bars show the std. Amber suppression activity reached 80% of wt. level.

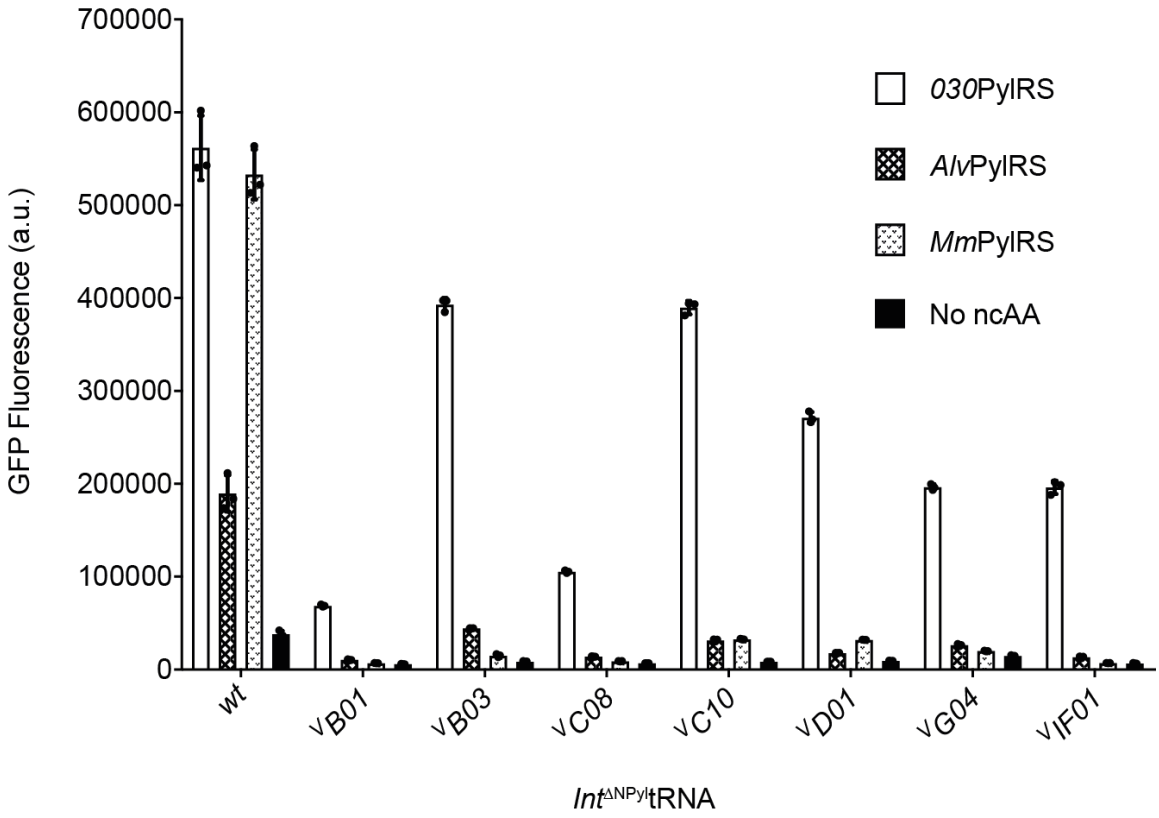

a
800000
600000
400000
200000
0
GFP Fluorescence (a.u.)
wt
6
8
10
11
15
17
19
20
21
22
23
24
AlvΔNPyltRNA
AlvPylRS
G1PylRS
H5PylRS
1R26PylRS


b
200000
150000
100000
50000
0
GFP Fluorescence (a.u.)
wt
6
8
10
11
15
17
19
20
21
22
23
24
AlvΔNPyltRNA
030PylRS
Lum1PylRS
RumEnPylRS


c
600000
400000
200000
0
GFP Fluorescence (a.u.)
wt
6
8
10
11
15
17
19
20
21
22
23
24
AlvΔNPyltRNA
MmPylRS
No PylRS

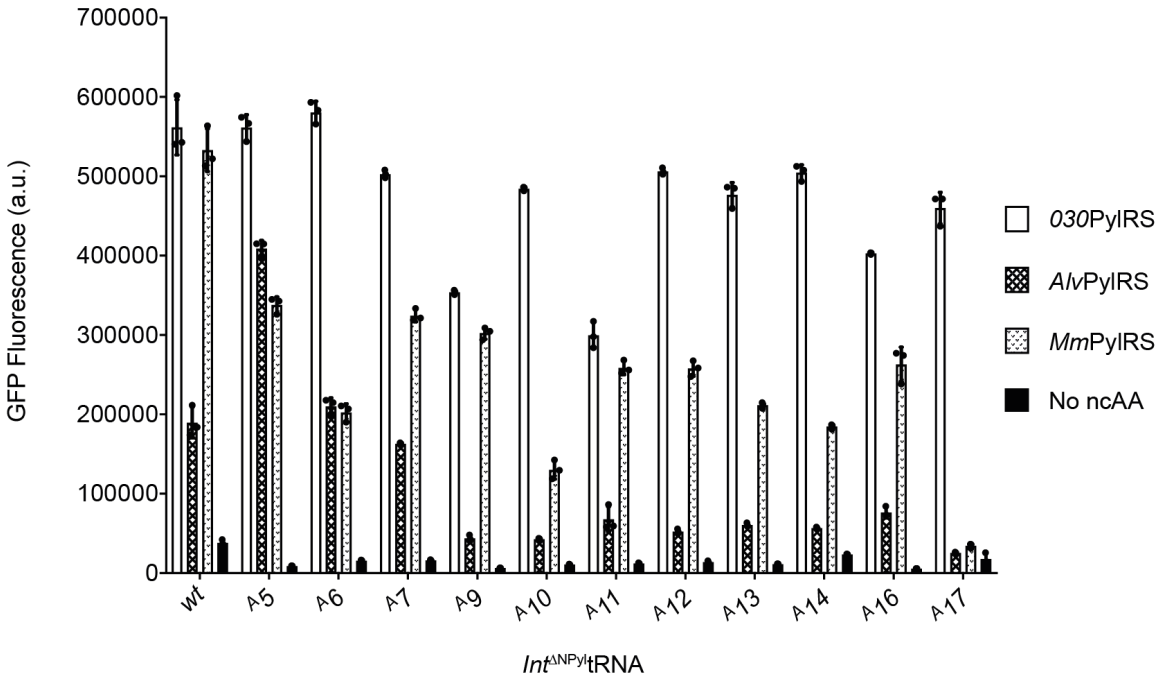
Supplementary Figure 13

In vivo amber suppression activity of wild type $Alv^{\Delta NPyl}tRNA_{CUA}$ or each $Alv^{\Delta NPyl}tRNA_{CUA}$ variant with each indicated PylRS. Measurements were made in *E. coli* DH10B bearing pBAD GFP(150TAG)His_6 PylRS and pKW $Alv^{\Delta NPyl}tRNA_{CUA}$ in the presence of BocK. Each bar represents an average of three biological replicates and error bars show the std.. **a**, All combinations of Class A ΔNPylRSs with $Alv^{\Delta NPyl}tRNA_{CUA}$ or each $Alv^{\Delta NPyl}tRNA_{CUA}$ variant. **b**, All combinations of Class B ΔNPylRSs with $Alv^{\Delta NPyl}tRNA_{CUA}$ or each $Alv^{\Delta NPyl}tRNA_{CUA}$ variant. **c**, All combinations of *Mm*PylRS with $Alv^{\Delta NPyl}tRNA_{CUA}$ or each $Alv^{\Delta NPyl}tRNA_{CUA}$ variant, and $Alv^{\Delta NPyl}tRNA_{CUA}$ or each $Alv^{\Delta NPyl}tRNA_{CUA}$ variant in the absence of any PylRS.



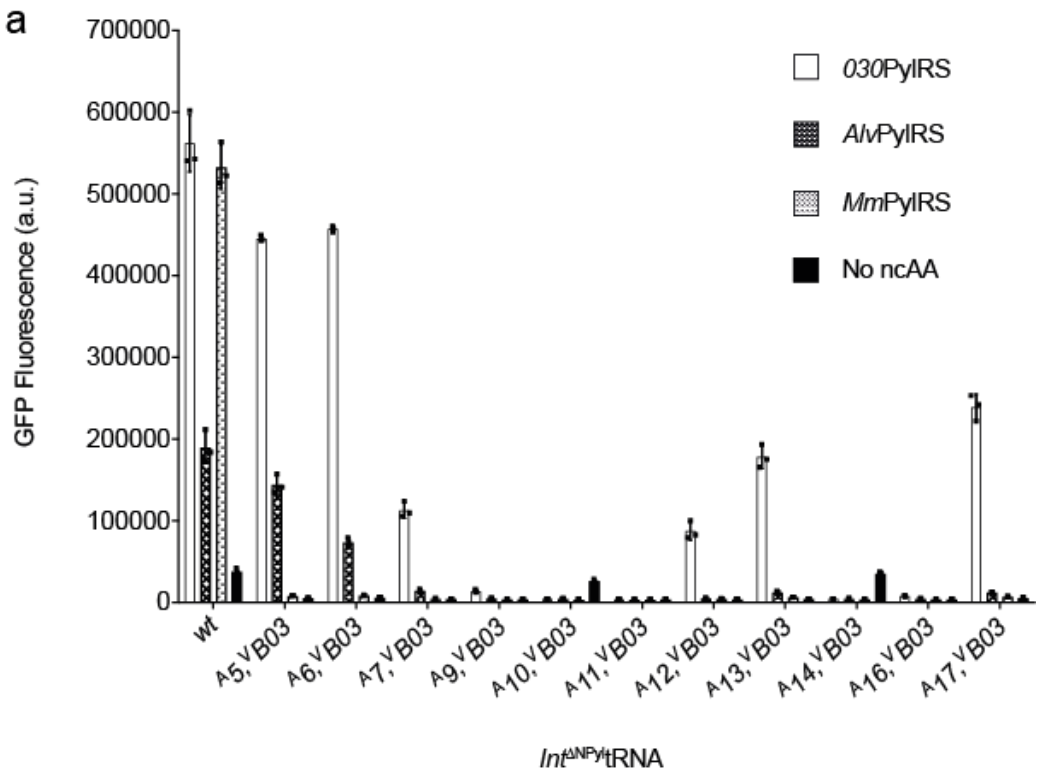
Supplementary Figure 14

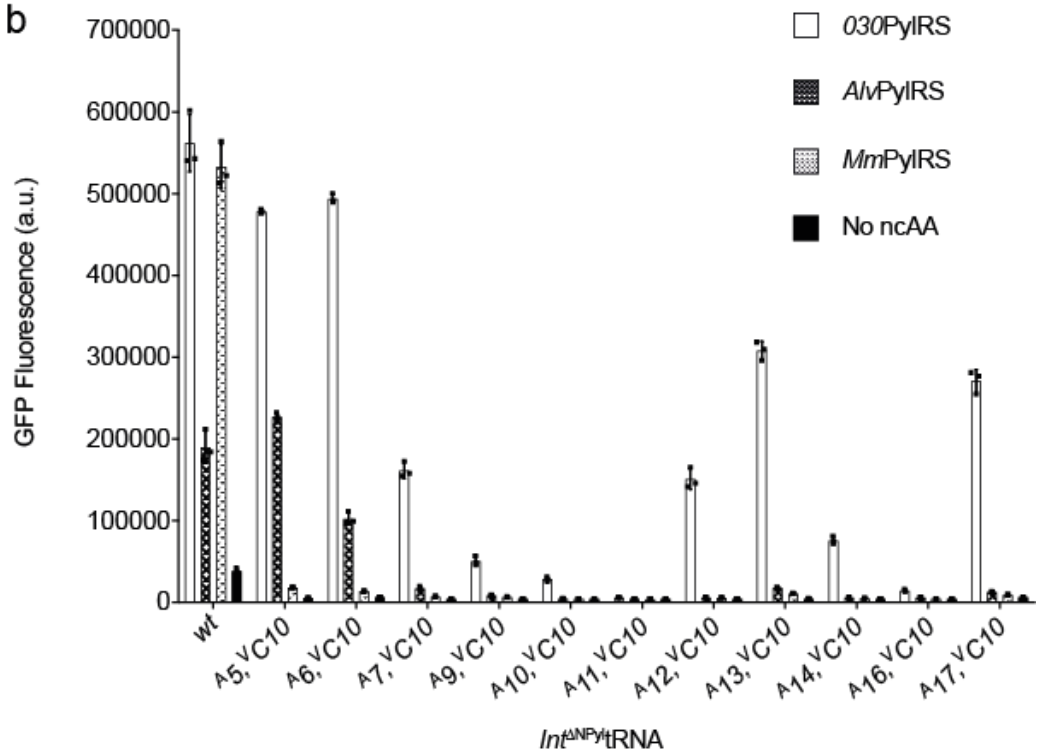
In vivo amber suppression activity of each $Int^{\Delta NPyl}tRNA_{CUA}$ variable loop library hit with *Mm*PylRS (Class +N), *Alv*PylRS (Class A) or *030*PylRS (Class B). Measurements were made in *E. coli* DH10B bearing pBAD GFP(150TAG)His_6 and pKW PylRS $Int^{\Delta NPyl}tRNA_{CUA}$ in the presence of BocK. Measurements for the background level of amber suppression activity in the absence of BocK were recorded for each $Int^{\Delta NPyl}tRNA_{CUA}$ variable loop library hit with *030*PylRS present in cells. Each bar chart represents three biological replicates with error bars showing std.. All hits showed improved orthogonality with respect to *Mm*PylRS (Class +N) and *Alv*ΔNPylRS (Class A). Only $Int^{\Delta NPyl}tRNA_{CUA}(^{V}B03)$ and $Int^{\Delta NPyl}tRNA_{CUA}(^{V}C10)$ were deemed sufficiently active for further studies.



Supplementary Figure 15

In vivo amber suppression activity of each $Int^{\Delta NPyl}tRNA_{CUA}$ acceptor stem library hit with *Mm*PylRS (Class +N), *Alv*PylRS (Class A) or *030*ΔNPylRS (Class B). Measurements were made in *E. coli* DH10B bearing pBAD GFP(150TAG)His_6 and pKW PylRS $Int^{\Delta NPyl}tRNA_{CUA}$ in the presence of BocK. Measurements for the background level of amber suppression activity in the absence of BocK were made for each $Int^{\Delta NPyl}tRNA_{CUA}$ acceptor stem library hit with *030*PylRS present in cells. Each bar chart represents three biological replicates with error bars showing std.. All hits showed improved orthogonality with respect to *Mm*PylRS (Class +N) and eight out of 11 hits showed improved orthogonality with respect to *Alv*PylRS (Class A).

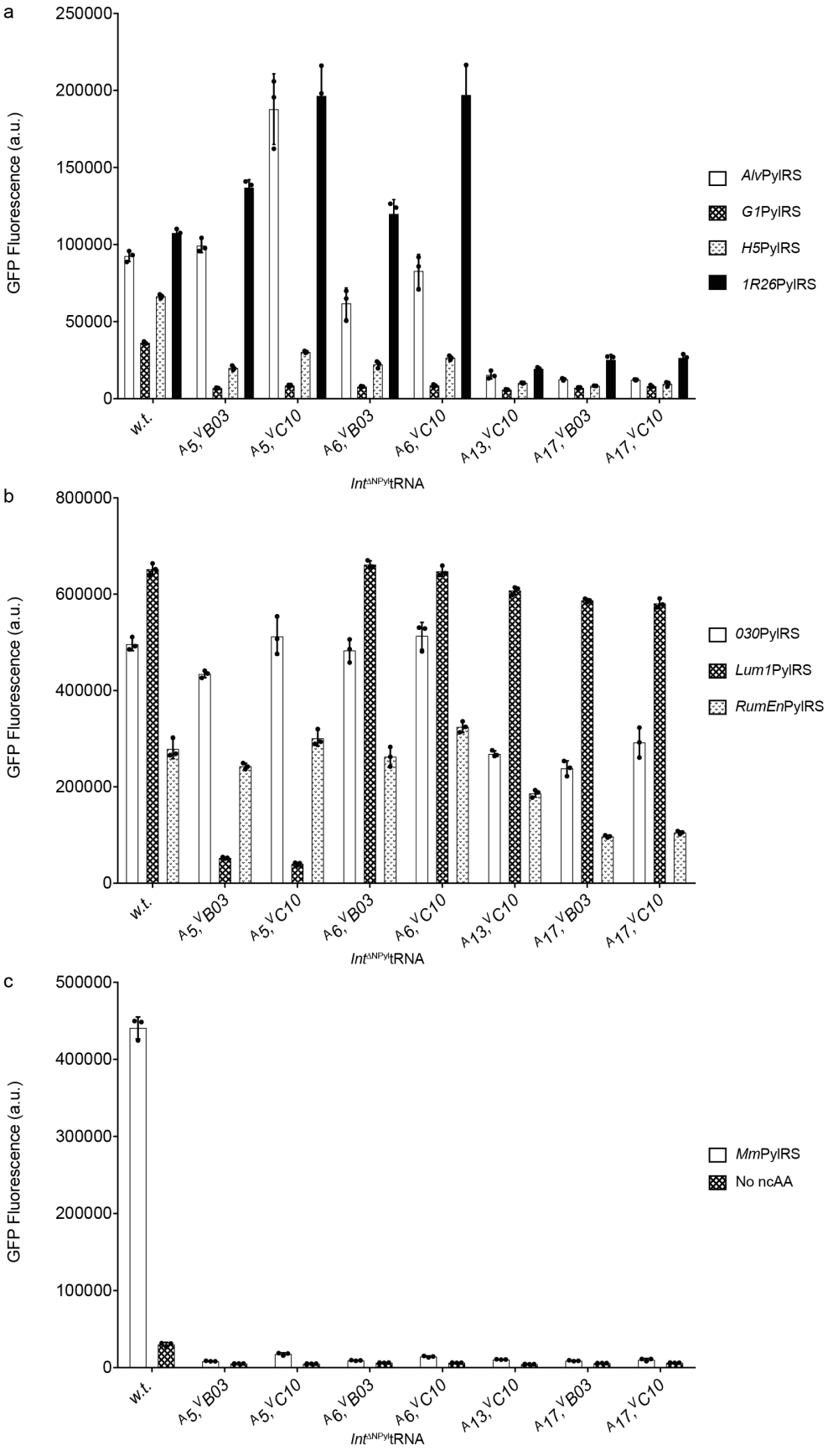

a
GFP Fluorescence (a.u.)
700000
600000
500000
400000
300000
200000
100000
0
030PylRS
AlvPylRS
MmPylRS
No ncAA
wt
A5,VB03
A6,VB03
A7,VB03
A9,VB03
A10,VB03
A11,VB03
A12,VB03
A13,VB03
A14,VB03
A16,VB03
A17,VB03
IntΔNPyltRNA


b
GFP Fluorescence (a.u.)
700000
600000
500000
400000
300000
200000
100000
0
030PylRS
AlvPylRS
MmPylRS
No ncAA
wt
A5,VC10
A6,VC10
A7,VC10
A9,VC10
A10,VC10
A11,VC10
A12,VC10
A13,VC10
A14,VC10
A16,VC10
A17,VC10
IntΔNPyltRNA

Supplementary Figure 16

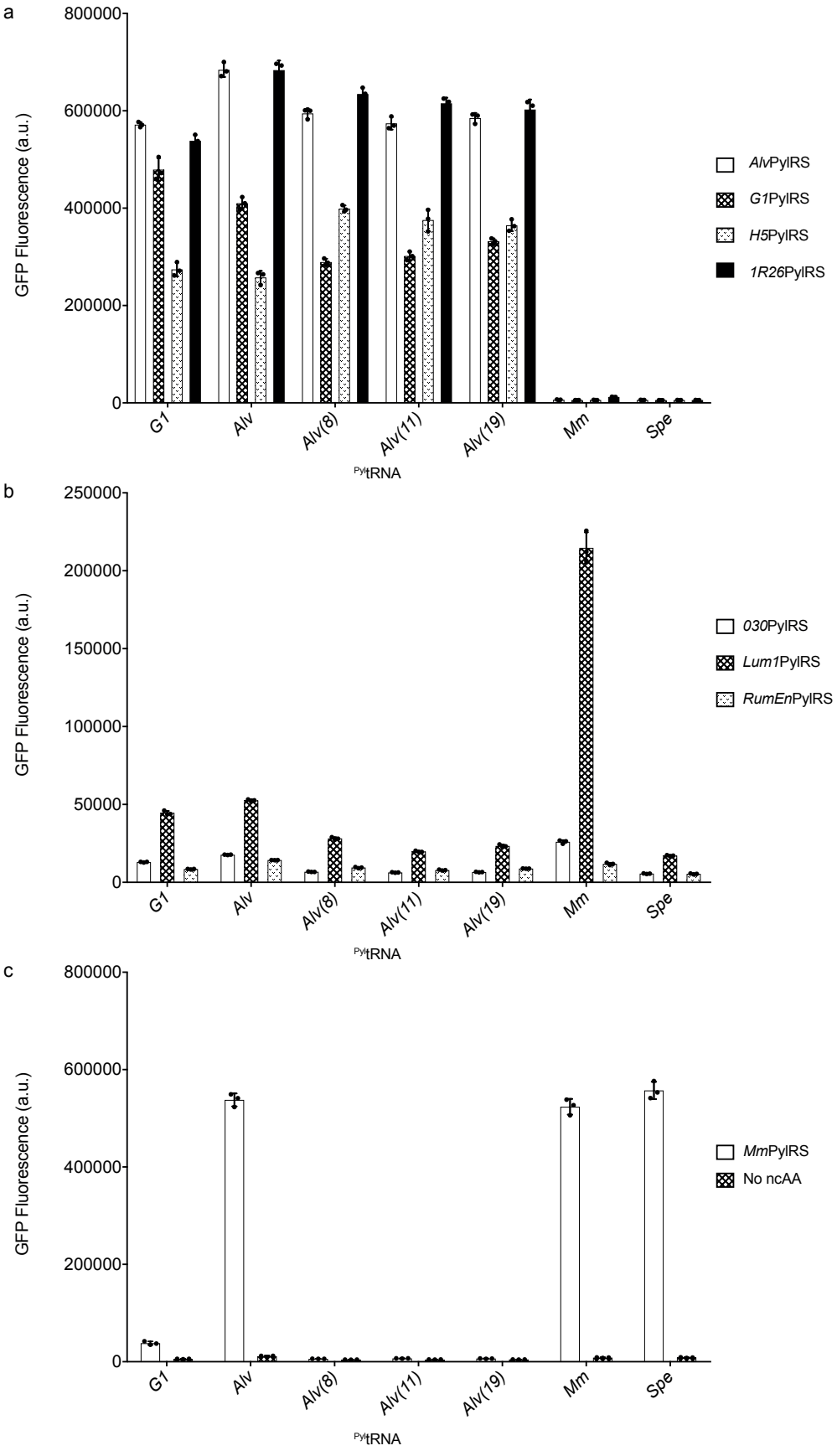
In vivo amber suppression activity of each *Int*$^{\Delta NPyl}$tRNA$_{CUA}$ hybrid with *Mm*PylRS (Class +N), *Alv*PylRS (Class A) or *030*PylRS (Class B). Measurements were made in *E. coli* DH10B bearing pBAD GFP(150TAG)His$_6$ and pKW PylRS *Int*$^{\Delta NPyl}$tRNA$_{CUA}$ in the presence or absence of BocK. Measurements for the background level of amber suppression activity in the absence of BocK were made for each *Int*$^{\Delta NPyl}$tRNA$_{CUA}$ hybrid with *030*PylRS present in cells. Each bar chart represents three biological replicates with error bars showing std.. All hybrids showed improved orthogonality with respect to *Mm*PylRS (Class +N) and *Alv*PylRS (Class A). Seven of the 22 tested hybrids were deemed active enough to be used in further studies.

In this expression system both the synthetase and tRNA are expressed from the same plasmid, which differs from the expression system used for library selection and screening in which the synthetase and tRNA are expressed from separate plasmids. While the general trends are consistent between the two systems, we observed some qualitative differences (discussed in **Supplementary Fig. 7**).


a
GFP Fluorescence (a.u.)
250000
200000
150000
100000
50000
0
w.t.
A5, VB03
A5, VC10
A6, VB03
A6, VC10
A13, VC10
A17, VB03
A17, VC10
IntΔNPyltRNA
AlvPylRS
G1PylRS
H5PylRS
1R26PylRS
b
GFP Fluorescence (a.u.)
800000
600000
400000
200000
0
w.t.
A5, VB03
A5, VC10
A6, VB03
A6, VC10
A13, VC10
A17, VB03
A17, VC10
IntΔNPyltRNA
030PylRS
Lum1PylRS
RumEnPylRS
c
GFP Fluorescence (a.u.)
500000
400000
300000
200000
100000
0
w.t.
A5, VB03
A5, VC10
A6, VB03
A6, VC10
A13, VC10
A17, VB03
A17, VC10
IntΔNPyltRNA
MmPylRS
No ncAA

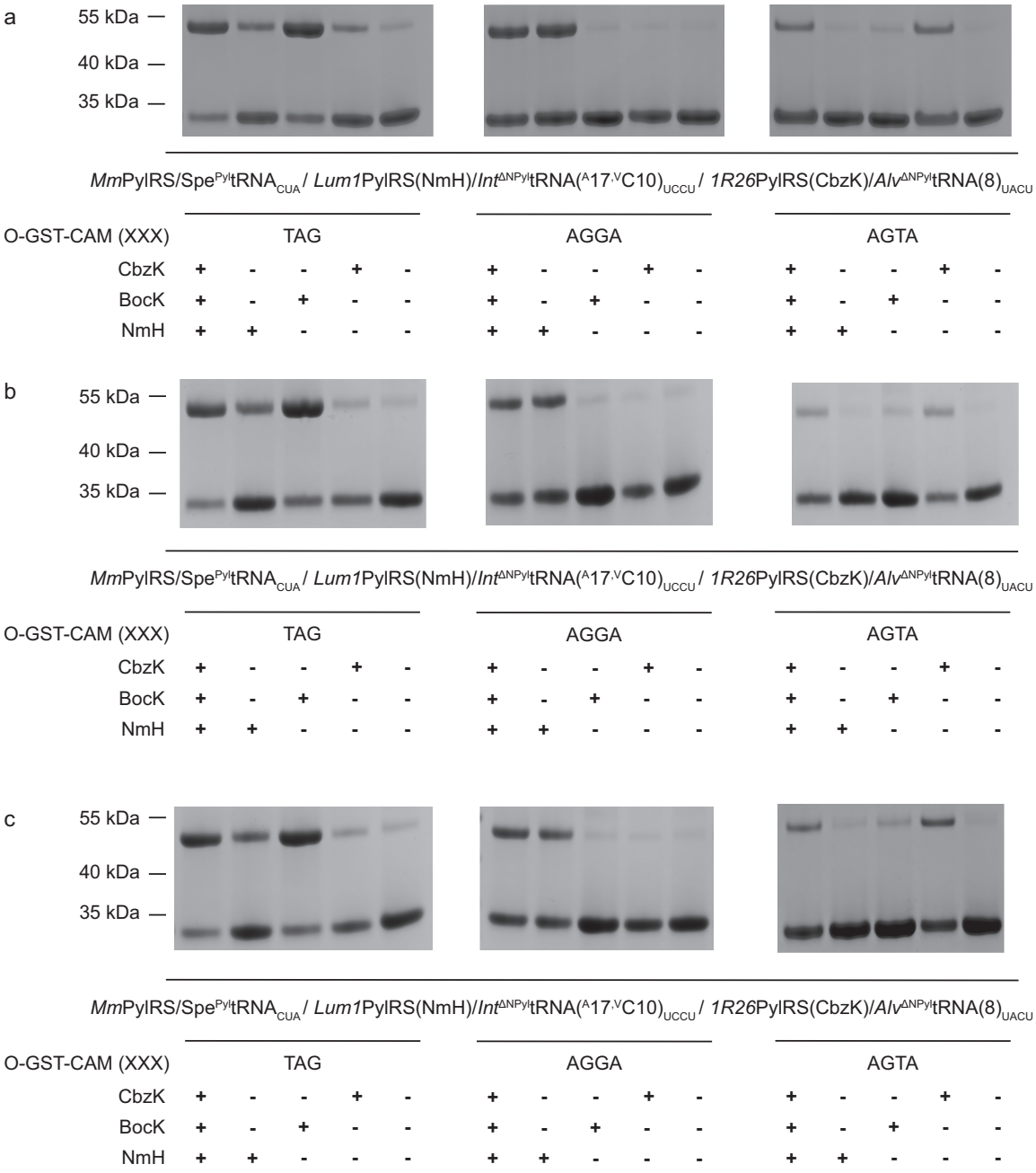
Supplementary Figure 17

In vivo amber suppression activity of wild type $Int^{\Delta NPyl}tRNA_{CUA}$ or each $Int^{\Delta NPyl}tRNA_{CUA}$ hybrid with each PylRS. Measurements were made in *E. coli* DH10B bearing pBAD GFP(150TAG)His_6 and pKW PylRS $Int^{\Delta NPyl}tRNA_{CUA}$ in the presence or absence of BocK. Each bar represents three biological replicates and errors are given as std.. **a**, All combinations of Class A ΔNPylRSs with $Int^{\Delta NPyl}tRNA_{CUA}$ or each $Int^{\Delta NPyl}tRNA_{CUA}$ hybrid in the presence of BocK. **b**, All combinations of Class B ΔNPylRSs with $Int^{\Delta NPyl}tRNA_{CUA}$ or each Int^{Δ}NPyltRNA$_{CUA}$ hybrid in the presence of BocK. **c**, All combinations of *Mm*ΔNPylRS with $Int^{\Delta NPyl}tRNA_{CUA}$ or hybrid $Int^{\Delta NPyl}tRNA_{CUA}$ hybrid in the presence of BocK, and $Int^{\Delta NPyl}tRNA_{CUA}$ or each $Int^{\Delta NPyl}tRNA_{CUA}$ hybrid in the absence of BocK and presence of *030*PylRS. We identified several $Int^{\Delta NPyl}tRNA_{CUA}$ hybrids which were orthogonal to *Mm*PylRS and some Class A ΔNPylRSs, whilst being highly active with specific Class B ΔNPylRSs.


a
800000
600000
400000
200000
0
GFP Fluorescence (a.u.)
AlvPylRS
G1PylRS
H5PylRS
1R26PylRS
G1
Alv
Alv(8)
Alv(11)
Alv(19)
Mm
Spe
PyltRNA
b
250000
200000
150000
100000
50000
0
GFP Fluorescence (a.u.)
030PylRS
Lum1PylRS
RumEnPylRS
G1
Alv
Alv(8)
Alv(11)
Alv(19)
Mm
Spe
PyltRNA
c
800000
600000
400000
200000
0
GFP Fluorescence (a.u.)
MmPylRS
No ncAA
G1
Alv
Alv(8)
Alv(11)
Alv(19)
Mm
Spe
PyltRNA

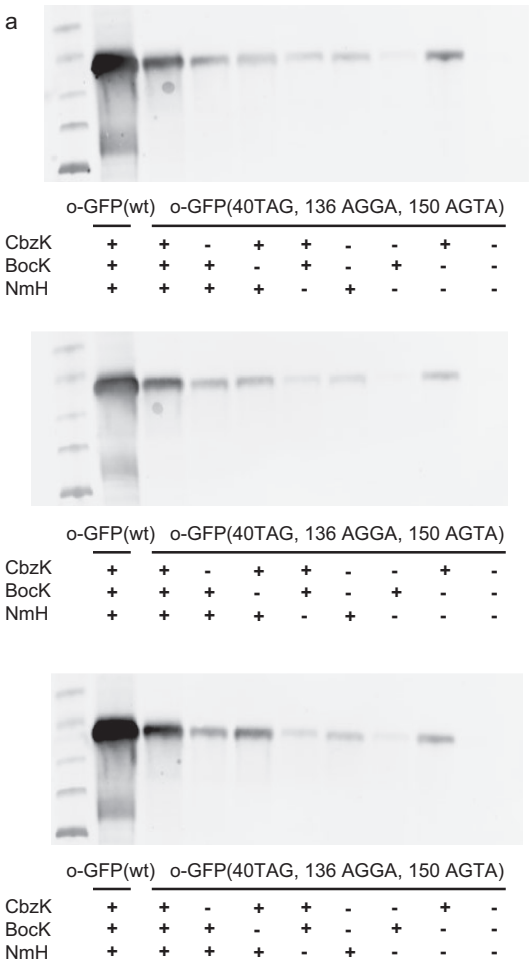
Supplementary Figure 18

In vivo amber suppression activity of discovered or evolved $^{Pyl}tRNA_{CUA}$s used as part of triply orthogonal PylRS/^{Pyl}tRNA pairs as compared against their starting wild type sequence. Measurements were made in *E. coli* DH10B bearing pBAD GFP(150TAG)His_6 and pKW PylRS $^{Pyl}tRNA_{CUA}$ in the presence or absence of BocK. Each bar represents three biological replicates and errors are given as std.. **a**, All combinations of Class A ΔNPylRSs with each $^{Pyl}tRNA_{CUA}$ in the presence of BocK. **b**, All combinations of Class B ΔNPylRSs with each $^{Pyl}tRNA_{CUA}$ in the presence of BocK. **c**, All combinations of *Mm*PylRS with each $^{Pyl}tRNA_{CUA}$ in the presence of BocK, and each $^{Pyl}tRNA_{CUA}$ in the absence of BocK and presence of *030*PylRS.


a
55 kDa
40 kDa
35 kDa
MmPylRS/SpePyltRNACUA / Lum1PylRS(NmH)/IntΔNPyltRNA(A17,VC10)UCCU / 1R26PylRS(CbzK)/AlvΔNPyltRNA(8)UACU
O-GST-CAM (XXX)
TAG
AGGA
AGTA
CbzK + - - + - + - - + - + - - + -
BocK + - + - - + - + - - + - + - -
NmH + + - - - + + - - - + + - - -
b
55 kDa
40 kDa
35 kDa
MmPylRS/SpePyltRNACUA / Lum1PylRS(NmH)/IntΔNPyltRNA(A17,VC10)UCCU / 1R26PylRS(CbzK)/AlvΔNPyltRNA(8)UACU
O-GST-CAM (XXX)
TAG
AGGA
AGTA
CbzK + - - + - + - - + - + - - + -
BocK + - + - - + - + - - + - + - -
NmH + + - - - + + - - - + + - - -
c
55 kDa
40 kDa
35 kDa
MmPylRS/SpePyltRNACUA / Lum1PylRS(NmH)/IntΔNPyltRNA(A17,VC10)UCCU / 1R26PylRS(CbzK)/AlvΔNPyltRNA(8)UACU
O-GST-CAM (XXX)
TAG
AGGA
AGTA
CbzK + - - + - + - - + - + - - + -
BocK + - + - - + - + - - + - + - -
NmH + + - - - + + - - - + + - - -

Supplementary Figure 19

Glutathione-S-Transferase-calmodulin (GST-CAM) purifications from *E. coli* containing ribo-Q1, O-GST-CAM(1XXX), where XXX stands for either TAG, AGGA or AGTA, *Mm*PylRS/*Spe*PyltRNA$_{CUA}$, *Lum1*PylRS(NMH) /*Int*PyltRNA(^{A}17,^{V}C10)$_{UCCU}$ and *1R26*PylRS(Cbz) /*Ma*PyltRNA(8)$_{UACU}$ grown in the presence and absence of the 8 mM BocK, 8mM NmH and 2 mM CbzK. Samples were analysed by SDS–PAGE. a, b, c are biological replicates and yielded similar results.



Supplementary Figure 20

Nickel NTA purification of $O_{strep}GFP(XXX)_{His6}$, where XXX stands for either wt or 40TAG, 136 AGGA and 150 AGTA, from *E. coli* containing ribo-Q1, O-strep-$GFP(XXX)_{His6}$, *Mm*PylRS/*Spe*$PyltRNA_{CUA}$, *Lum1*PylRS(NMH)/*Int*$PyltRNA(^{A}17,^{V}C10)_{UCCU}$ and *1R26*PylRS(CbzK)/*Ma*$PyltRNA(8)_{UACU}$ in absence and presence of indicated ncAAs (8 mM BocK, 8 mM NmH, 2 mM CbzK) analysed by western blot. Three biological replicates were performed yielding similar results.

Supplementary Tables

Supplementary Table 1

Species, abbreviations (abbrev.), original reporting and sequences of all $^{\Delta NPyl}$tRNAs used in this work.

Species	Abbrev.	Reported	Sequence
Methanomethylophilus sp. 1R26	*1R26*	This work	GGGGGACGATCCGGCGATCAGCGGGT CTCTAAAACCTAGCCAGCGGGGATCGA CACCCCGGTCTCTCGCCA
Thermoplasmatales archaeon BRNA1	*BRNA*	This work	GGGGGACGATCCGGCGATCAGCGGGT CTCTAAAACCTAGCCAGCGGGGATCGA CACCCCGGTCTCTCGCCA
Methanomassiliicoccales archaeon PtaU1.Bin030	*030*	This work	GGAGGGTTGGTCCGGGACCACCTGGCC TCTACAGCTAAGGCAGCCGGGTTCAAC TCCCGGGCCCTTCGCCA
Methanomassiliicoccales archaeon RumEn M1	*RumEn*	This work	GGAGTGTTGGTCCGGAGACCACCAGGC CTCTACAGCCGCGGCAGCCGGGTTCGA CTCCCGGGCACTTCGCCA
Methermicoccus shengliensis	*Sheng*	This work	GGAGGGTTGGTCCGGGACCGCCAGGC CTCTACAGCCACGGTAGCTGGGTTCGA CTCCCAGGCCCTTCGCCA
Methanogenic archaeon ISO4-G1	*G1*	(2018) Willis et al.[10]	GGAGGGCGCTCCGGCGAGCAAACGGG TCTCTAAAACCTGTAAGCGGGGTTCGA CCCCCCGGCCTTTCGCCA
Methanogenic archaeon ISO4-H5	*H5*	(2018) Willis et al.[10]	GGGGGGCGATCCGGCGATCAGCGGGT CTCTAAAACCTAGCCAGCGGGGTTCGA CGCCCCGGCCTCTCGCCA
Methanoplasma termitum	*Term*	(2018) Willis et al.[10]	GGGAGACGGTCTGGGACCAGTAGGCC TCTAAAGCTCAACCAGCGGGGTTCGAT CCCCCGGTCTCTCGCCA
Methanomethylophilus alvus	*Alv*	(2014) Borrel et al.[11]	GGGGGACGGTCCGGCGACCAGCGGGT CTCTAAAACCTAGCCAGCGGGGTTCGA CGCCCCGGTCTCTCGCCA
Methanomassiliicoccus luminyensis 1	*Lum1*	(2014) Borrel et al.[11]	GGAGTGTTGGTTCGGCGACCACCAGGC CTCTACAGCCACGGCAGCCGGGTTCGA CTCCCGGGCACTTCGCCA
Methanomassiliicoccus luminyensis 2	*Lum2*	(2014) Borrel et al.[11]	GGAGGGTTGGTCAGGGACCGCCAGGC CTCTACAGCCACGGCAGCCGGGTTCGA CTCCCGGGCCCTTCGCCA
Methanomassiliicoccus intestinalis	*Int*	(2014) Borrel et al.[11]	GGAGTGTTGGTCCGGGACCACCAGGCC TCTACAGCCACGGCAGCCGGGTTCAAC TCCCGGGCACTTCGCCA

Supplementary Table 2

Species, abbreviotions (abbrev.), original reporting, amino acid and nucleotide sequences of all PylRSs used in this work.

Species	Abbrev.	Reported	Amino acid sequence	Nucleotide sequence
Methanomassiliicoccales archaeon PtaU1.Bin030	*030*	This work	MVIEWSPSQKQRLRELGRADEGGMEFETVV ERDEAFTKEVAYYQSINRKEIRNIQERRERHL LAKVEENIAEALIADGFLEVRTPTIISGNALV KMGIDHNHPLREQVFWLDGSRCLRPMLAPN LYFLMRHLKRNVRMPLQMFEIGTCYRKESH GSNHLEEFTMLNLVEMASMDDPAVRLRHHI QTVMGAIGLEYELSECESDVYGRTIDVEVNG VEVASAALGPHKLDPAHGITDAWSGVGFGL ERLLMVKNAENNIKKVGRSLIYLGGARLDI	ATGGTCATTGAATGGAGCCCGTCTCAAAAGCAACGCTTGCGTGAAT TAGGCCGCGCAGACGAGGGGGGCATGGAGTTTGAAACCGTTGTTGA ACGCGACGAAGCGTTTACGAAAGAAGTAGCTTACTATCAAAGTATC AATCGTAAAGAAATCCGTAACATCCAGGAGCGCCGCGAACGCCATC TTCTGGCAAAAGTTGAGGAGAATATCGCGGAGGCATTGATCGCAGA TGGATTTTTAGAGGTTCGCACCCCAACTATTATCTCTGGAAATGCTC TGGTGAAGATGGGGATCGATCATAATCATCCATTACGCGAGCAGGT TTTTTGGCTGGACGGCTCCCGTTGTTTACGCCCCATGCTGGCTCCCA ACCTTTATTTCTTGATGCGTCACTTGAAACGCAATGTTCGTATGCCA CTGCAAATGTTTGAAATCGGCACATGCTACCGTAAAGAGTCCCACG GCTCCAACCATTTAGAAGAGTTTACGATGTTGAACTTAGTGGAAAT GGCTTCGATGGACGATCCTGCCGTCCGTTTGCGCCACCATATTCAAA CGGTAATGGGGGCAATCGGTCTGGAATATGAATTATCCGAATGCGA GTCCGACGTGTACGGGCGCACAATCGACGTTGAGGTCAATGGCGTG GAAGTTGCTTCAGCAGCTCTGGGACCTCACAAACTGGACCCGGCGC ATGGTATCACCGATGCTTGGAGCGGGGTTGGGTTTGGTCTGGAACG TTTATTAATGGTCAAAAACGCTGAAAATAATATTAAGAAAGTCGGG CGCTCATTAATCTACCTGGGAGGAGCCCGTCTGGACATTTAA
Methanomassiliicoccales archaeon PtaU1.Bin124	*124*	This work	MAITFSAAQNQRIRELDCESNLGECIFETESE REDVFRKVVNDLVDKNRSDLLSFARMPDTS GMHQLQCTLANRLAVSGFMQVHTPTMMSV ASLEKMGIGDDHPLRRQIFSLDQNRCLRPML APNLYAVMKRMARAVPGRFGIFEIGKCFRKE SKGAHHIEEFTMLNLVDVRPSDGPEARLKEL SGLLSRDLGLPIEIAQEGSDVYGTTLDLEVNG VELASAAYGPHPLDKAFHVDFPWAGIGMGL ERVLMVQAGSSNIHRHAASLVYQYGTRIDI	ATGGCAATTACTTTCTCCGCAGCACAGAACCAGCGCATCCGTGAGC TGGATTGCGAGTCAAACTTAGGTGAATGCATCTTCGAAACCGAGAG TGAGCGCGAAGATGTTTTTCGTAAGGTCGTAAATGATTTGGTTGAC AAGAACCGCTCCGATCTTTTATCTTTCGCACGCATGCCGGATACCAG CGGTATGCATCAGTTGCAGTGCACGCTTGCCAACCGTTTGGCCGTGT CTGGGTTCATGCAAGTGCATACTCCAACTATGATGAGCGTGGCTTCT TTGGAGAAGATGGGAATTGGGGACGACCACCCCCTTCGTCGTCAGA TTTTTTCGTTGGACCAAAATCGCTGTCTGCGTCCAATGCTGGCTCCG AACCTGTACGCAGTTATGAAACGTATGGCTCGTGCTGTACCTGGTC GTTTTGGTATCTTCGAGATCGGCAAATGTTTTCGTAAAGAGAGTAA AGGGGCTCACCACATCGAAGAGTTCACCATGTTGAACTTGGTAGAC GTCCGTCCCTCGGATGGTCCTGAGGCCCGTCTTAAAGAGTTATCGG GGCTTCTTTCCCGTGATCTGGGATTACCTATTGAAATTGCACAGGAG GGTTCAGATGTCTATGGTACGACCTTGGATCTTGAAGTGAATGGGG TGGAGCTTGCCTCTGCGGCTTATGGACCACACCCTCTTGATAAAGCG TTTCATGTAGACTTTCCATGGGCGGGTATCGGGATGGGTCTGGAGC GTGTCTTAATGGTCCAAGCGGGTAGTAGTAACATCCATCGCCATGC CGCTAGTTTAGTATATCAGTATGGGACGCGTATCGATATTTAA
Methanomethylophilus sp. 1R26	*1R26*	(2018) Willis et al.[10]	MAEHFTDAQIQRLREYGNGTYKDMEFADVS AREKAFTKLMSDASRDNESALKGMIAHPAR QGLSRLMNDIADALVADGFIEVRTPIIISKDA LAKMTITPDKPLFKQVFWIDDKRALRPMLAP SLYTVMRSLRDHTDGPVKIFEMGSCFRKESH SGMHLEEFTMLNLVDMGPAGDATESLKKYI GIVMKAAGLPDYQLVHEESDVYKETIDVEIN GQEVCSAAVGPHYLDAAHDVHEPWAGAGF GLERLLTIRQGYSTVMKGGASTTYLNGAKM D	ATGGCAGAACATTTTACAGATGCACAGATTCAGCGCTTACGCGAGT ATGGTAATGGGACATATAAGGATATGGAGTTCGCAGACGTGAGTGC GCGTGAAAAGGCGTTCACGAAGCTTATGTCTGATGCCAGTCGCGAT AATGAATCAGCCCTTAAGGGAATGATTGCGCACCCTGCACGCCAGG GTTTATCACGCTTAATGAATGACATTGCAGACGCTTTGGTGGCTGAT GGATTTATTGAAGTTCGCACACCCATTATTATCAGTAAAGATGCTTT AGCCAAAATGACAATCACTCCCGATAAACCATTATTCAAACAAGTA TTTTGGATTGATGATAAACGCGCTCTGCGTCCAATGTTAGCCCCGTC TCTGTACACGGTCATGCGTAGCCTGCGTGATCATACTGACGGCCCTG TCAAAATTTTCGAAATGGGCTCTTGTTTTCGCAAAGAATCGCACAGT GGGATGCACCTGGAAGAGTTTACCATGTTAAACCTGGTAGATATGG GCCCCGCCGGCGATGCCACTGAGTCCCTGAAGAAATATATTGGGAT CGTAATGAAGGCCGCCGGGCTGCCTGACTACCAATTAGTCCACGAG GAATCTGATGTGTATAAAGAAACGATTGATGTTGAAATCAACGGAC AAGAGGTTTGTTCGGCTGCTGTAGGTCCTCACTATTTGGATGCCGCC CATGACGTGCACGAGCCGTGGGCAGGGGCAGGGTTCGGCCTGGAG CGCCTGCTGACAATTCGCCAGGGATATAGCACAGTGATGAAAGGGG GAGCTTCCACAACCTATTTGAACGGGGCTAAGATGGACTAA
Thermoplasmatales archaeon BRNA1	*BRNA1*	(2018) Willis et al.[10]	MKYTDAQIQKLREYGNGDYSGAEFDDASAR DKAFSRDMSAATKDNEAKIQAMFSKPDRPA LTRLMADIAAALTAEGFIEVRTPIMITKDALT RMTITPGRPLYKQVFWIDDNRALRPMLAPSL YSVMRSLRDHTDGPVKIFEMGPCFRREHSG MHLEEFTMLNLVDMGPNEDAIETLKKYIDV VMKVVGLENYDLVQEESDVYKETIDVEING QEVCSAAVGPHYLDAAHDVHEPWSGAGFG LERLLALREKYSTVRKAGASVSYLNGAKIN	ATGAAATACACGGATGCACAAATTCAGAAGCTGCGCGAATACGGA AACGGAGATTACTCAGGAGCTGAGTTCGACGACGCTTCCGCTCGCG ATAAAGCATTTTCACGCGATATGTCAGCAGCCACCAAGGACAACGA GGCTAAGATCCAAGCGATGTTTAGCAAACCGGACCGTCCAGCACTT ACTCGTCTGATGGCAGACATCGCCGCAGCCCTTACAGCGGAGGGTT TTATCGAAGTACGTACCCCGATTATGATCACGAAGGATGCGTTAAC GCGCATGACTATCACTCCAGGACGTCCGCTGTATAAGCAGGTTTTCT GGATTGACGACAACCGCGCACTTCGTCCTATGCTGGCACCTAGCTT GTATTCAGTAATGCGTTCGCTGCGTGATCATACTGACGGACCCGTTA AGATCTTTGAAATGGGCCCTTGCTTTCGCCGCGAATCACATTCTGGG ATGCATTTAGAAGAGTTCACGATGCTTAATCTGGTCGATATGGGGC CGAATGAAGACGCGATTGAAACGTTGAAGAAATACATTGACGTGGT AATGAAGGTGGTTGGATTAGAGAATTATGATTTAGTTCAAGAAGAG AGCGACGTCTATAAAGAAACGATTGACGTGGAGATCAATGGTCAAG AAGTATGTTCCGCAGCAGTAGGCCCGCACTATTTAGACGCTGCTCA CGACGTCCATGAACCGTGGTCTGGGGCAGGCTTTGGACTGGAACGT CTGTTAGCGTTACGTGAGAAGTATAGCACGGTTCGCAAAGCGGGGG CATCTGTCTCATACTTAAATGGTGCAAAAATTAATTAA
Methanogenic archaeon ISO4-G1	*G1*	(2018) Willis et al.[10]	MVVKFTDSQIQHLMEYGDNDWSEAEFEDAA ARDKEFSSQFSKLKSANDKGLKDVIANPRND LTDLENKIREKLAARGFIEVHTPIFVSKSALA KMTITEDHPLFKQVFWIDDKRALRPMHAMN LYKVMRELRDHTKGPVKIFEIGSCFRKESKSS THLEEFTMLNLVEMGPDGDPMEHLKMYIGD IMDAVGVEYTTSREESDVYVETLDVEINGTE VASGAVGPHKLDPAHDVHEPWAGIGFGLER LLMLKNGKSNARKTGKSITYLNGYKLD	ATGGTGGTCAAGTTTACTGATTCCCAAATTCAACACCTTATGGAGTA TGGCGATAATGATTGGTCAGAGGCCGAATTCGAAGACGCTGCTGCA CGCGACAAAGAGTTCTCCAGTCAATTTTCCAAACTGAAGTCTGCAA ATGATAAAGGGCTGAAAGACGTTATTGCCAACCCTCGCAACGATTT AACCGACCTTGAAAACAAGATCCGTGAAAAACTTGCAGCGCGCGGT TTTATTGAGGTGCACACTCCCATCTTTGTTAGTAAGTCAGCCTTAGC GAAAATGACTATTACCGAGGACCACCCGCTTTTTAAGCAAGTTTTTT GGATCGATGATAAACGTGCGTTGCGTCCGATGCATGCTATGAATCT GTACAAAGTTATGCGCGAGTTACGCGATCACACCAAAGGCCCGGTT AAGATTTTTGAGATCGGTTCGTGCTTTCGCAAAGAGTCCAAGTCATC TACCCATTTGGAGGAGTTTACAATGTTAAATCTGGTCGAAATGGGG CCAGACGGGGATCCTATGGAGCATCTTAAAATGTACATTGGGGACA TCATGGACGCGGTAGGCGTTGAATACACGACCTCCCGCGAAGAAAG TGATGTATATGTTGAAACCTTGGATGTGGAAATCAATGGTACAGAG GTAGCCAGTGGCGCTGTCGGCCCCCACAAATTAGACCCCGCGCACG ACGTGCATGAGCCTTGGGCAGGTATCGGTTTCGGGCTTGAGCGCTT ACTTATGCTTAAAAATGGAAAATCAAATGCGCGTAAGACCGGGAAA TCTATCACTTATCTTAACGGATACAAGTTAGACTAA
Methanogenic archaeon ISO4-H5	*H5*	(2018) Willis et al.[10]	MTCKLTDPQIQRLREYGHEPKNESEFETEEE RDKAFTKMMSKLQRENEKGIRDMIANPRHH RLMELELQLSEALIKEGFIEVKTPILISKAELA KMTIDENHPLYQQVFWVDDKRCLRPMHAIN LYNIMRELRGHTDGPVKFFEIGSCFRAESHSN	ATGACTTGTAAATTGACAGACCCTCAGATTCAGCGTCTTCGTGAGTA TGGCCACGAACCCAAGAACGAGTCAGAATTTGAAACGGAAGAGGA ACGTGATAAAGCGTTTACGAAAATGATGTCAAAGTTACAACGCGAA AACGAAAAGGGGATTCGCGACATGATTGCTAATCCACGCCATCATC GCCTGATGGAATTGGAGTTACAATTATCCGAAGCACTTATCAAGGA

			DHLEEFTMLNLVDMGPQGDTTEKIKHYIDIV MKTIGLDYELVHEESDVYKETIDVEVDGEEV CSAAVGPHYLDKAHNINEPWCGAGFGLERLI MMRDGDGSVKKTGKSVNYLNGYKIN	GGGCTTCATTGAAGTAAAGACTCCTATCCTTATCTCGAAAGCCGAG CTTGCGAAAATGACTATCGATGAGAATCACCCCTTATACCAGCAAG TATTCTGGGTGGACGACAAACGCTGCCTTCGTCCTATGCACGCAATT AACCTGTATAACATCATGCGTGAACTGCGTGGCCATACCGATGGCC CCGTGAAATTCTTTGAAATTGGTTCGTGTTTCCGTGCGGAATCGCAC TCCAATGACCATTTAGAGGAGTTTACAATGCTTAATCTGGTTGACAT GGGCCCACAAGGTGACACAACTGAGAAAATCAAACACTACATTGAT ATTGTAATGAAAACTATTGGGTTGGATTATGAGCTGGTACATGAAG AATCTGATGTGTATAAGGAAACAATTGATGTGGAGGTAGACGGAGA AGAGGTCTGTTCCGCTGCTGTGGGCCCTCATTACTTGGATAAGGCTC ACAATATTAACGAACCTTGGTGTGGAGCAGGGTTTGGCTTAGAGCG CCTTATCATGATGCGTGACGGTGATGGTAGTGTAAAGAAAACTGGG AAGTCAGTTAACTATCTTAACGGTTACAAAATCAACTAA
Methanomassiliicoccales archaeon RumEn M1	*RumEn*	(2018) Willis et al.[10]	MTIEWTPSQKQRLKELGIDSDQDYTINNIQER EEVFSRLVTRRQSEGRRAIRSMMEHPVRHKL AQLEQDLAQALVDDGFLEFRTPTIITRSALEK MGIGREHPLHEQVFWLDEKRCLRPMLAPNL YYVMRHLKRNAKGPVKLFEIGTCYRKESHG SNHLEEFTMLNLVELDPAGDAREQLRKHISTI MNTIGLDYELVSCSSDVYVETTDVEVNGVE VASGAIGPHKLDPAHGIKAPWAGVGFGLERL LMLKHGEDNVKKVGRSLIYLQGVRLDI	ATGACTATTGAGTGGACGCCAAGCCAGAAACAGCGTCTTAAAGAAT TAGGGATCGATTCAGATCAAGATTATACCATCAATAACATTCAGGA GCGTGAGGAGGTTTTTTCGCGCCTGGTAACACGCCGCCAATCAGAA GGTCGTCGTGCAATCCGCTCCATGATGGAACACCCTGTCCGTCACA AACTGGCGCAGTTGGAACAGGACCTGGCGCAAGCTTTGGTCGATGA CGGCTTCCTGGAATTTCGCACGCCAACTATCATTACCCGCAGTGCTT TGGAAAAAATGGGCATTGGTCGTGAGCATCCCCTTCACGAACAAGT TTTTTGGTTAGACGAGAAGCGTTGCCTGCGTCCAATGCTTGCCCCCA ACTTGTATTACGTTATGCGCCACTTGAAGCGTAACGCGAAGGGTCC CGTCAAGTTATTTGAGATCGGCACCTGTTACCGCAAAGAAAGTCAC GGGTCGAACCATCTGGAAGAATTTACCATGTTGAATCTTGTAGAAT TGGACCCAGCTGGGGATGCCCGCGAACAATTACGCAAACACATTTC CACTATTATGAATACGATCGGTTTAGACTACGAGTTAGTGAGCTGCT CGTCGGACGTTTACGTTGAAACTACGGACGTAGAAGTCAATGGCGT AGAAGTAGCGAGCGGTGCTATTGGGCCCCATAAATTAGATCCTGCG CATGGCATTAAAGCTCCTTGGGCAGGTGTGGGTTTCGGTTTGGAGC GTTTATTAATGTTAAAGCACGGCGAGGATAACGTAAAAAAAGTAGG CCGCTCGTTAATTTACTTGCAAGGTGTCCGCCTGGACATCTAA
Methanoplasma termitum	*Term*	(2018) Willis et al.[10]	MSIGFTPSQIQKLREFGEDPRDNSTYQNVEQR DKAFSKLMSDLVSSNEKEIAGMLRSPSRHQL AALEEDLAAALIARGFIEVKTPAFVSVASLEK MTITPEHPLYKQVFMIDDKRCLRPMHAMNL YYVMRKLRDHTDGPVKIFEIGSCFRKESHSG SHLEEFTMLNLVELGPEGDATEALKDHIGAV MKVTGLEYTLVREESDVYVETLDVEIDGGE VASGAVGPHVLDNAHDIHEPWSGIGFGLERL LMIMNEKSTVKKTGRSLSYLNGAKIN	ATGAGTATTGGTTTTACTCCGTCACAAATCCAGAAGCTTCGTGAATT CGGCGAGGACCCCCGCGACAATAGTACCTACCAAAACGTAGAACA GCGCGATAAAGCATTTTCAAAACTGATGTCAGATTTAGTCTCTTCTA ATGAAAAGGAAATTGCTGGTATGTTGCGCTCCCCATCACGCCACCA GTTGGCAGCGCTTGAAGAGGATTTAGCGGCGGCACTGATTGCACGC GGGTTCATTGAGGTAAAAACACCCGCATTCGTCAGTGTGGCGTCCT TGGAGAAGATGACAATCACTCCCGAACATCCATTGTACAAACAGGT TTTTATGATCGACGACAAACGTTGTCTTCGTCCAATGCATGCTATGA ACCTGTATTATGTAATGCGCAAGTTGCGCGATCATACTGATGGACC CGTCAAGATTTTTGAGATTGGTTCGTGTTTTCGTAAAGAAAGCCATA GTGGTAGCCACTTGGAGGAGTTTACCATGCTTAACTTGGTCGAACTT GGGCCTGAGGGCGACGCTACGGAGGCTCTTAAAGACCACATCGGG GCTGTTATGAAAGTCACTGGACTGGAATACACGCTGGTTCGTGAGG AGAGTGACGTTTACGTAGAAACCCTGGACGTAGAGATCGACGGAG GCGAAGTCGCCAGCGGGGCGGTGGGTCCGCATGTTCTTGATAATGC ACATGACATCCATGAACCGTGGAGTGGGATCGGTTTTGGACTTGAA CGTTTATTAATGATTATGAACGAGAAGAGTACGGTGAAAAAGACTG GCCGTAGTCTGTCTTACCTTAATGGTGCTAAGATCAATTAA
Methanohalarchaeum thermophilum	*Therm 1*	(2018) Willis et al.[10]	MELTRSQSQRLRELGYQGEAPTFEDQEERDE FFERKETELQKKNRNKFKKLQRINEPDWKKT EQKLRKNLYESDFTEVQTPHIISMSVLKNKM NISEESNIYNQIYKLDEGNKCLRPMLAPNLYR QMKHFLRISKKDVVKLFELGTCFRKEQGKN HVREFKMLNAVEVGEIKDKEKRTREMIDEII GNLVDYKIEEEKSTVYGKTLDIEVNGLEIASS VIGPHPLDANFSINKPWIGIGIGVERLIQTKNE GNSIKSYARSLSYQDGIRLEIN	ATGGAATTGACACGCAGCCAAAGTCAACGTTTGCGTGAGTTGGGAT ACCAAGGAGAGGCGCCAACCTTCGAGGACCAGGAAGAGCGCGATG AGTTTTTTGAACGTAAAGAAACAGAATTACAAAAAAAAAACCGTAA TAAGTTTAAAAAACTTCAGCGCATCAATGAGCCTGACTGGAAGAAG ACGGAGCAAAAATTGCGCAAGAACTTATACGAGTCAGATTTCACGG AAGTTCAAACACCACATATTATCTCGATGAGTGTATTGAAAAATAA AATGAACATTTCGGAGGAATCGAACATTTATAATCAAATCTACAAA CTGGACGAAGGAAACAAGTGCCTGCGTCCCATGTTAGCTCCCAATC TTTATCGTCAGATGAAACACTTTCTTCGTATTTCAAAGAAAGATGTA GTTAAATTATTTGAGCTGGGAACTTGCTTCCGCAAGGAACAAGGGA AAAACCACGTTCGTGAATTTAAGATGTTAAATGCAGTCGAAGTTGG AGAGATCAAGGATAAGGAGAAGCGTACACGTGAGATGATTGATGA AATTATCGGCAATCTTGTAGACTATAAGATTGAGGAAGAAAAAAGT ACAGTCTATGGCAAAACCTTAGACATCGAGGTAAACGGACTGGAAA TCGCCTCCAGCGTCATCGGACCACATCCCCTGGATGCGAACTTTAGC ATCAACAAACCCTGGATCGGCATCGGCATCGGCGTTGAGCGCTTGA TTCAGACGAAGAACGAGGGGAATTCTATTAAAAGTTATGCTCGTTC GTTGAGTTACCAGGACGGTATTCGCCTTGAAATCAATTAA
Methanohalarchaeum thermophilum	*Therm 2*	(2018) Willis et al.[10]	MEFTETQKQRLRELGYKGEFPELDTKEEVNE AYSQLEKKLRKKHRKKLNDLFESKKPTWKN TVENIRQNLQDLGFIEVQTPLIISKNLLKKMKI DQKSDLMNQVYRINDNKVLRPMLAQNLYK ELENFSKLSNRDTIQLFEIGTCFRKEKGGKDH LNEFKMLNAVELGNFKDKEKRLKEVISTLFK DFDEYVLEKEKSTVYGETYDVLVNGTELAS CAIGPHQLDEKWDINRPWIGIGIGIERFTREL NNSDSTVKAYGRSFVYQDGIRLDIK	ATGGAGTTCACAGAAACCCAAAAGCAACGTTTGCGCGAACTGGGAT ATAAGGGGGAATTTCCGGAGTTGGACACTAAGGAAGAAGTGAATG AAGCCTACTCTCAGCTTGAGAAAAAGCTGCGTAAGAAGCACCGCAA GAAGTTGAACGACTTGTTTGAATCCAAAAAGCCGACATGGAAAAAT ACCGTTGAGAACATTCGCCAGAATTTGCAAGATCTTGGTTTCATTGA AGTACAAACTCCTTTAATTATCTCAAAGAACCTGTTGAAGAAGATG AAAATCGATCAGAAGTCCGATTTGATGAACCAGGTCTACCGCATCA ACGATAATAAGGTGCTGCGCCCAATGCTTGCACAGAACCTGTATAA GGAATTGGAAAACTTCAGCAAATTGTCGAACCGTGATACGATTCAA CTTTTTGAAATTGGGACTTGCTTCCGCAAGGAAAAGGGAGGGAAGG ACCACCTGAACGAGTTTAAGATGTTGAATGCGGTGGAATTGGGTAA CTTTAAAGATAAGGAGAAACGCTTAAAAGAAGTCATCTCGACATTG TTCAAGGACTTCGATGAATATGTTTTAGAAAAAGAGAAATCTACGG TCTATGGCGAGACATATGATGTGTTGGTAAATGGAACAGAGTTAGC CTCTTGTGCCATTGGTCCCCATCAATTAGATGAGAAATGGGACATTA ATCGTCCCTGGATCGGTATCGGCATTGGTATCGAGCGTTTCACGCGC GAGCTTAATAATTCGGACTCAACCGTCAAGGCGTATGGACGTTCGT TCGTGTACCAAGATGGCATTCGTCTTGACATCAAATAA
Methanonatronarchaeum thermophilum	*Tron*	(2018) Willis et al.[10]	MEFTVTQKQRLQELGFEGVFPSDFEDVDERN RFFEELVGRLRDRNRKRFERLVGNKIPFWRK VSSDLRNRFYELGFVEVRTPEIISYSLLEKMEI SDDLREQVYWLEEDNRCLRPMLAPNLYNEL RHFNRISNQSKVRIFEIGTCFRREKSSSEHLNE FTMLNAVEMGDIGDTEERLDRLIEEVFGEFT DYKKVGEESSLYGKTVDVLVDGVEVASCIA GPHPLDSNWSIDQPWVGIGLGVERLAMLLD DGSTAKAYGNSYIYQDGVRLDIK	ATGGAGTTCACGGTGACCCAGAAGCAACGCTTACAAGAGCTTGGTT TTGAAGGTGTATTCCCCTCGGACTTCGAGGATGTTGATGAGCGTAA CCGTTTCTTTGAAGAGTTAGTAGGCCGCTTACGTGACCGTAACCGCA AACGCTTTGAGCGCTTGGTAGGAAATAAAATCCCTTTCTGGCGTAA AGTGTCTTCGGACCTTCGTAATCGCTTCTATGAATTGGGTTTTGTTG AAGTTCGTACCCCGGAGATCATCTCATACAGCTTACTTGAGAAAAT GGAAATTTCAGACGATTTACGTGAACAAGTCTATTGGCTTGAAGAA GATAATCGTTGTCTGCGTCCAATGTTAGCGCCGAATCTTTACAACGA ATTGCGTCACTTTAACCGTATCTCAAATCAATCCAAGGTTCGTATCT TCGAAATCGGGACTTGTTTTCGCCGTGAGAAGAGCTCATCCGAGCA CCTGAACGAGTTTACGATGCTGAATGCCGTGGAAATGGGTGACATT GGCGACACGGAAGAACGCCTGGACCGCTTAATCGAAGAGGTCTTCG GTGAATTTACAGACTACAAGAAAGTTGGGGAAGAGTCAAGTTTGTA TGGAAAAACGGTTGACGTGTTGGTTGACGGCGTGGAAGTGGCTTCC TGTATTGCCGGACCCCACCCTCTTGACTCGAACTGGTCGATTGATCA ACCTTGGGTGGGGATCGGCCTGGGTGTGGAGCGTTTAGCAATGTTA CTGGACGATGGAAGTACCGCTAAGGCCTATGGCAATTCGTATATTT ATCAGGATGGTGTACGTCTTGACATCAAATAA
MSBL1 archaeon SCGC-AAA382A20	*SCGC*	(2018) Willis et al.[10]	MNLTSSQKQRLRELGWDGSIPDFDNKKERD QFFNKTATKLKNRNKERFLKLLENKVPSWR RVERKLRNIFYELGFVEVQTPSIIPSPLLEKMD IGEESKLYNQIYQIKGEKKSLRPMLAPNLYRE LRYFSRISDEEVIRLFELGSCFRKENGGERHL NEFKMLNAVEMGNIKDTKKRLDELISNVFSP FANYKVEKEKSTVYEETVDVNIKNTEVASC VIGPHFLDSNWHIDEPWVGLGIGVERLTRVIE GEPSVKPFGKSYVYQDGIRLDIE	ATGAACTTGACGAGCTCACAGAAACAACGCTTGCGCGAGTTAGGTT GGGACGGTTCAATCCCAGACTTCGACAATAAAAAAGAGCGCGACC AGTTCTTCAATAAAACTGCAACAAAATTAAAAAACCGTAATAAAGA ACGCTTTTTGAAGCTTCTGGAAAACAAGGTGCCAAGTTGGCGTCGC GTGGAACGCAAGTTGCGTAACATTTTCTACGAGCTGGGCTTCGTGG AGGTACAAACTCCATCGATTATTTCCCCGAGTTTACTGGAAAAAAT GGATATTGGTGAGGAATCTAAGTTATATAACCAGATTTACCAGATC AAAGGGGAAAAAAAGAGTCTGCGCCCCATGCTGGCCCCCAACTTAT ATCGTGAACTTCGTTATTTCAGTCGCATTTCGGACGAAGAGGTGATT

				CGTTTGTTTGAATTGGGATCCTGTTTCCGTAAGGAGAATGGCGGTGA GCGCCATCTTAACGAGTTCAAGATGTTAAACGCGGTGGAAATGGGC AATATCAAGGATACTAAGAAGCGTTTGGATGAGTTAATCAGCAACG TCTTTTCTCCGTTTGCTAATTATAAGGTAGAAAAAGAAAAATCAAC AGTGTATGAGGAAACCGTAGACGTAAACATCAAAAACACAGAGGT CGCTTCCTGCGTCATTGGTCCGCACTTTCTGGATTCAAACTGGCATA TCGATGAACCTTGGGTAGGGCTTGGCATCGGTGTAGAGCGTTTGAC GCGCGTAATTGAAGGTGAACCATCGGTTAAACCATTCGGCAAGAGT TATGTTTACCAGGATGGGATCCGCCTTGACATTGAGTAA
Methermicoccus shengliensis	*Sheng*	(2018) Willis et al.[10]	MIGFTDTQVQRLKELGGDQKIVGCRFSSVAD RDEVFETTVKNLVEENREKLRRMAHSPSRCS LFELEDRLASTLVGMGFMEVATPMLLSASNL KKMGIDESHPLWEQVFWVDKKRCMRPMLA PNLYFLLKHLKRNIKKPVRIFEIGPCFRKESQ GSRHLEEFTMLNLVELAPDCEPTERLTELIEK VMGQIGLEYRLKNESSEVYGNTVDVEVDGV EVASGAVGPHFLDGAYGITDAWVGVGFGLE RLLMVMEGHSNIRKVGRSLVYLNGARIDI	ATGATTGGGTTCACCGACACCCAGGTACAGCGCCTGAAAGAGCTGG GCGGCGACCAAAAGATCGTGGGGTGCCGCTTCTCTTCAGTCGCTGA CCGTGACGAAGTTTTCGAGACTACAGTCAAAAACTTGGTCGAGGAG AACCGCGAAAAGCTGCGCCGCATGGCCCATTCGCCTTCTCGTTGCTC TTTGTTCGAATTGGAAGACCGCCTGGCCTCTACTCTTGTTGGGATGG GCTTCATGGAGGTCGCGACACCAATGTTATTAAGCGCGAGTAACCT TAAAAAGATGGGTATCGATGAATCACATCCCCTGTGGGAGCAAGTT TTTTGGGTCGATAAGAAACGCTGCATGCGTCCGATGCTGGCTCCCA ATTTGTACTTCTTATTAAAGCATCTTAAGCGCAACATTAAGAAACCA GTGCGTATTTTTGAGATCGGTCCATGTTTTCGCAAAGAGTCACAGGG CAGTCGCCACTTAGAGGAGTTTACAATGCTGAATCTGGTCGAGTTG GCTCCTGATTGCGAACCGACTGAACGCCTGACGGAACTGATTGAGA AGGTTATGGGCCAAATTGGATTGGAATACCGTCTGAAGAATGAGTC GTCCGAGGTATATGGAAACACGGTCGATGTAGAGGTTGATGGCGTT GAAGTCGCTTCGGGAGCCGTCGGTCCTCATTTCCTGGATGGAGCGT ATGGAATCACAGATGCCTGGGTGGGGGTTGGGTTTGGCTTGGAACG CTTACTTATGGTCATGGAGGGTCATTCTAACATTCGTAAAGTGGGTC GCAGCCTTGTATATTTGAATGGGGCTCGTATTGACATCTAA
Methanomethylophilus alvus	*Alv*	(2014) Borrel et al.[11]	MTVKYTDAQIQRLREYGNGTYEQKVFEDLA SRDAAFSKEMSVASTDNEKKIKGMIANPSRH GLTQLMNDIADALVAEGFIEVRTPIFISKDAL ARMTITEDKPLFKQVFWIDEKRALRPMLAPN LYSVMRDLRDHTDGPVKIFEMGSCFRKESHS GMHLEEFTMLNLVDMGPRGDATEVLKNYIS VVMKAAGLPDYDLVQEESDVYKETIDVEIN GQEVCSAAVGPHYLDAAHDVHEPWSGAGF GLERLLTIREKYSTVKKGGASISYLNGAKIN	ATGACGGTGAAATACACAGACGCGCAGATTCAGCGTTTACGCGAGT ATGGAAATGGGACTTACGAACAGAAGGTCTTTGAGGATCTGGCGAG TCGTGATGCAGCGTTTTCAAAAGAGATGTCCGTTGCGTCTACGGAT AATGAGAAGAAAATCAAGGGTATGATCGCCAACCCTAGTCGTCATG GGCTGACCCAATTAATGAATGACATTGCGGATGCCCTGGTCGCAGA GGGTTTTATCGAAGTGCGCACTCCAATCTTTATTTCTAAGGATGCTT TGGCACGCATGACCATCACCGAGGACAAACCATTATTTAAACAGGT TTTTTGGATCGACGAAAAACGTGCCTTACGCCCAATGCTGGCTCCA AACCTGTATAGTGTCATGCGTGATCTTCGCGATCACACAGATGGCC CTGTTAAGATTTTTGAGATGGGATCTTGTTTTCGTAAGGAGTCACAT TCTGGCATGCATCTGGAAGAGTTTACGATGCTTAATCTGGTCGATAT GGGGCCTCGCGGCGACGCGACTGAAGTATTGAAGAATTACATTAGC GTAGTTATGAAGGCAGCCGGGCTTCCTGACTATGACTTGGTCCAAG AAGAGTCGGATGTATATAAGGAAACTATCGATGTTGAGATCAACGG ACAAGAAGTTTGCAGCGCGGCGGTAGGTCCCCATTACCTGGATGCC GCTCACGATGTTCACGAGCCTTGGAGCGGCGCGGGTTTCGGCCTTG AACGCCTTTTAACTATCCGCGAGAAATATTCCACCGTGAAAAAGGG CGGCGCGTCAATCTCCTATCTGAATGGGGCCAAAATTAATTAA
Methanomassiliicoccus intestinalis	*Int*	(2014) Borrel et al.[11]	MPVEWTASQKQRLKELGIPAEADRIFNDTKE REEVFKDITSEHLSKVRKDIKHMLDYPERHQ LSQIESILAQALVDNGFIEVKTPSIISRSALEK MGIDRSHPLHEQVFWLDEKRCLRPMLAPNL YFMMRHMYRYSKGPLRLFEIGSCFRKESKGS NHLEEFTMLNLVEMAPDNDPADQLLVHIKTI MDALGLEYSLVECESDVYVKTLDVEIDGVE VASGAVGPHKLDPAHGITQSWAGVGFGLER LSMMKYGMDNIKKSGRSLIYLRGVRLDI	ATGCCGGTCGAGTGGACCGCATCTCAGAAACAGCGCCTTAAGGAAC TTGGGATTCCTGCGGAAGCTGACCGCATTTTCAATGATACGAAAGA ACGTGAAGAGGTCTTTAAAGATATTACATCGGAACACTTAAGTAAA GTGCGTAAGGATATTAAACATATGTTGGACTATCCTGAACGTCATC AGTTGTCCCAAATCGAGTCGATCCTGGCACAAGCGTTGGTAGATAA CGGTTTCATTGAGGTTAAGACACCCTCGATCATTAGCCGTAGCGCAT TGGAAAAGATGGGAATCGACCGTAGTCATCCATTGCACGAACAAGT TTTTTGGTTGGACGAGAAACGCTGTTTGCGTCCGATGTTAGCCCCGA ACCTGTACTTTATGATGCGTCACATGTACCGTTACAGTAAGGGCCCG TTGCGCCTGTTCGAGATCGGCTCTTGTTTCCGCAAAGAGAGTAAGG GATCGAATCATTTGGAGGAATTTACTATGTTGAACCTGGTGGAGAT GGCTCCAGACAACGATCCGGCAGACCAACTTTTAGTCCATATCAAA ACAATCATGGATGCCCTTGGATTAGAATACTCATTGGTGGAATGTG AATCAGATGTTTATGTCAAAACCTTGGATGTTGAAATCGACGGTGTT GAGGTCGCGTCTGGTGCGGTGGGCCCCCATAAACTGGACCCCGCAC ACGGAATCACACAAAGTTGGGCAGGAGTCGGATTTGGGCTTGAGCG CTTGTCCATGATGAAGTACGGTATGGACAACATCAAAAAGAGTGGC CGCTCTCTTATCTACTTGCGTGGAGTACGCCTTGATATTTAA
Methanomassiliicoccus luminyensis 1	*Lum1*	(2014) Borrel et al.[11]	MDTRLTPAQAQRIREMGGTVDPSLAFSSEAE RESAFQRISADLQGANLAKIRRCAEAPERHPI GSLENTLACALAAKGFIEVKTPMMIPADGLV KMGIDESHPLWNQVFWVGPKKALRPMLAP NLYFLMRHLRRSVPAPLLLFEIGPCFRKESRG SNHLEEFTMLNLVELAPQADATERLKEHIAT VMNAVGLPYELVVEGSEVYGTTIDVEVDGV ELASGAVGPLPMDKPHGITEPWAGVGFGLE RIALMRTKEQNIKKVGRSLVYVNGARIDI	ATGGATACACGCCTGACGCCCGCGCAAGCTCAGCGCATTCGTGAAA TGGGCGGGACGGTAGACCCGTCGTTAGCCTTTAGCAGCGAAGCCGA ACGTGAATCGGCATTTCAGCGTATTTCGGCAGATTTGCAGGGGGCG AACTTGGCCAAGATTCGTCGCTGTGCTGAAGCGCCAGAGCGCCATC CAATCGGTAGTCTGGAAAATACCCTGGCCTGCGCACTGGCGGCTAA GGGGTTTATTGAGGTTAAGACGCCGATGATGATTCCTGCAGACGGA CTGGTGAAAATGGGTATTGACGAGTCTCATCCATTATGGAATCAAG TTTTTTGGGTCGGCCCGAAAAAAGCACTGCGTCCGATGCTTGCTCCA AACCTGTACTTTTTGATGCGCCACTTACGTCGTTCTGTGCCCGCCCC ATTACTTCTTTTTGAAATTGGACCTTGTTTCCGCAAGGAGAGCCGCG GGTCCAATCATCTGGAGGAGTTTACTATGTTGAACTTAGTCGAGTTA GCTCCACAAGCTGACGCTACAGAACGCTTAAAAGAACACATTGCCA CAGTGATGAATGCTGTTGGACTGCCCTACGAGTTAGTAGTCGAGGG TTCTGAAGTTTATGGCACTACGATTGATGTGGAAGTAGACGGTGTT GAGCTGGCATCTGGCGCAGTAGGACCGTTACCTATGGATAAACCTC ACGGGATTACCGAGCCCTGGGCTGGTGTTGGGTTCGGATTGGAGCG TATTGCATTAATGCGCACGAAGGAGCAAAATATCAAAAAAGTAGG ACGCTCCTTGGTTTATGTTAATGGAGCCCGCATTGATATTTAA
Methanomassiliicoccus luminyensis 2	*Lum2*	(2014) Borrel et al.[11]	MVRDRQASTATAAGFDSRALRSQHYESEDG GWQLIFDMTPSQKQRLRELGRVPDEGAAFST AEDRDAAFIKEVAYYQSYNRNVVRDALDAP KRHPLSHMEEVLAQALVDEGFLDVKTPTIIS GDSIRKMGISCAHPLNKQIFWVDGTRCLRPM LAPNLYFLMRHLKRNAQLPLRLFEIGPCYRIE THGSDHLEEFTMLNLVELAPQGDPLAQLHH HIATVMGAVGLDYQLCECDSEVYSRTIDVEV DGSEVASAALGPHALDRAHGIEDPWVGVGF GLERLLMSKSAESNIRKVGRSLIYLQGARIDV	ATGGTGCGCGATCGTCAGGCATCAACCGCTACTGCTGCAGGATTTG ATAGTCGCGCATTGCGCAGCCAGCATTACGAATCTGAAGATGGAGG CTGGCAGCTTATTTTCGACATGACCCCTTCGCAAAAACAGCGTCTGC GTGAGCTGGGGCGTGTCCCAGACGAGGGCGCAGCCTTCAGTACGGC CGAAGACCGTGACGCAGCTTTTATCAAAGAGGTGGCCTACTATCAA TCGTACAACCGTAATGTTGTGCGTGACGCACTTGACGCCCCGAAGC GTCACCCCCTTTCACATATGGAGGAGGTATTGGCCCAGGCTTTAGTG GATGAAGGCTTTTTGGATGTAAAGACCCCGACAATTATCTCTGGGG ACAGTATCCGCAAGATGGGAATCAGTTGCGCACACCCCCTGAATAA GCAAATTTTTTGGGTTGATGGGACACGCTGTTTACGTCCGATGCTTG CTCCCAACTTGTACTTCTTGATGCGTCATTTGAAGCGCAACGCACAG TTACCGTTACGTCTGTTCGAGATTGGACCCTGTTACCGCATTGAGAC TCACGGTAGTGACCACTTGGAGGAATTTACGATGTTAAATTTGGTC GAGTTGGCCCCGCAAGGAGATCCATTAGCTCAACTTCATCACCACA TCGCCACTGTAATGGGGGCCGTCGGTCTTGATTATCAGCTTTGTGAG TGCGACAGCGAAGTGTACAGCCGCACCATTGACGTTGAAGTGGACG GGAGTGAGGTCGCCTCTGCGGCGCTTGGTCCGCACGCCTTAGATCG CGCACACGGTATTGAAGACCCTTGGGTGGGTGTCGGGTTCGGCCTG GAGCGCTTATTGATGTCCAAATCAGCTGAATCCAACATCCGCAAGG TAGGACGTAGTCTTATCTACTTGCAAGGGGCGCGTATCGACGTGTA A
Methanosarcina mazei	*Mm*	(2005) Krzycki[12]	MDKKPLNTLISATGLWMSRTGTIHKIKHHEV SRSKIYIEMACGDHLVVNNSRSSRTARALRH HKYRKTCKRCRVSDEDLNKFLTKANEDQTS VKVKVVSAPTRTKKAMPKSVARAPKPLENT EAAQAQPSGSKFSPAIPVSTQESVSVPASVST SISSISTGATASALVKGNTNPITSMSAPVQAS APALTKSQTDRLEVLLNPKDEISLNSGKPFRE LESELLSRRKKDLQQIYAEERENYLGKLEREI TRFFVDRGFLEIKSPILIPLEYIERMGIDNDTEL SKQIFRVDKNFCLRPMLAPNLYNYLRKLDRA	ATGGACAAAAAACCGCTGAACACCCTGATCTCTGCTACCGGTCTGT GGATGTCTCGTACCGGTACCATCCACAAAATCAAACACCACGAAGT TTCTCGTTCTAAAATCTACATCGAAATGGCTTGCGGTGACCACCTGG TTGTTAACAACTCTCGTTCTTCTCGTACCGCTCGTGCTCTGCGTCACC ACAAATACCGTAAAACCTGCAAACGTTGCCGTGTTTCTGACGAAGA CCTGAACAAATTCCTGACCAAAGCTAACGAAGACCAGACCTCTGTT AAAGTTAAAGTTGTTTCTGCTCCGACCCGTACCAAAAAAGCTATGC CGAAATCTGTTGCTCGTGCTCCGAAACCGCTGGAAAACACCGAAGC TGCTCAGGCTCAGCCGTCTGGTTCTAAATTCTCTCCGGCTATCCCGG TTTCTACCCAGGAATCTGTTTCTGTTCCGGCTTCTGTTTCTACCTCTA

LPDPIKIFEIGPCYRKESDGKEHLEEFTMLNFC
QMGSGCTRENLESIITDFLNHLGIDFKIVGDS
CMVYGDTLDVMHGDLELSSAVVGPIPLDRE
WGIDKPWIGAGFGLERLLKVKHDFKNIKRA
ARSESYYNGISTNL

TCTCTTCTATCTCTACCGGTGCTACCGCTTCTGCTCTGGTTAAAGGT
AACACCAACCCGATCACCTCTATGTCTGCTCCGGTTCAGGCTTCTGC
TCCGGCTCTGACCAAATCTCAGACCGACCGTCTGGAAGTTCTGCTG
AACCCGAAAGACGAAATCTCTCTGAACTCTGGTAAACCGTTCCGTG
AACTGGAATCTGAACTGCTGTCTCGTCGTAAAAAAGACTTACAACA
GATCTACGCTGAAGAACGTGAAAACTACCTGGGTAAACTGGAACGT
GAAATCACGCGTTTCTTCGTTGACCGTGGTTTCCTGGAAATCAAATC
TCCGATCCTGATCCCGCTGGAATACATCGAACGTATGGGTATCGAC
AACGACACCGAACTGTCTAAACAGATCTTCCGTGTTGACAAAAACT
TCTGCCTGCGTCCGATGCTGGCTCCGAACCTGTACAACTACCTGCGT
AAACTGGACCGTGCTCTGCCGGACCCGATCAAAATCTTCGAAATCG
GTCCGTGCTACCGTAAAGAATCTGACGGTAAAGAACACCTGGAAGA
ATTCACCATGCTGAACTTCTGCCAGATGGGTTCTGGTTGCACCCGTG
AAAACCTGGAATCTATCATCACCGACTTCCTGAACCACCTGGGTAT
CGACTTCAAAATCGTTGGTGACTCTTGCATGGTTTACGGTGACACCC
TGGACGTTATGCACGGTGACCTGGAACTGTCTTCTGCTGTTGTTGGT
CCGATCCCGCTGGACCGTGAATGGGGTATCGACAAACCGTGGATCG
GTGCTGGTTTCGGTCTGGAACGTCTGCTGAAAGTTAAACACGACTTC
AAAAACATCAAACGTGCTGCTCGTTCTGAATCTTACTACAACGGTA
TCTCTACCAACCTGTAA

Supplementary Table 3

Table of all fluorescence measurements taken during this work. Provided as *Excel* sheet.

Supplementary Table 4

Species, abbreviations (abbrev.), original reporting and sequences of all $^{\Delta NPyl}$tRNAs used in this work.

Species	Abbrev.	Reported	Sequence
Methanosarcina barkeri MS	*Bar*	(2007) Herring et al.[13]	GGGAACCTGATCATGTAGATCGAATG GACTCTAAATCCGTTTAGCCGGGTTA GATTCCCGGGGTTTCCGCCA
Methanococcoides burtonii	*Bur*	(2007) Herring et al.[13]	GGAGACTTGATCATGTAGATCGAACG GACTCTAAATCCTTTCAGCCGGGTTA GATTCCCGGAGTTTCCGCCA
Methanohalobium evestigatum	*Eve*	(2011) Gaston et al.[14]	GGAAACCCGATCAGGTAGATCGAAT GGACTCTAAATCCATTCAGCCGGGTT AGATTCCCGGGGTTTCCGCCA
Methanomethylovorans hollandica	*Hol*	(2014) Borrel et al.[11]	GGAAACCGGATCATGTTGATCAAATG GACTCTAAATCCGTTCAGCCGGGTTA AATTCCCGGGGTTTCCGCCA
Methanohalophilus mahii	*Mah*	(2011) Gaston et al.[14]	GGAAACCTGATCAGGTAGATCAAATG GACTCTAGATCCATTCAGCCGGGTTA GATTCCCGGGGTTTCCGCCA
Methanosarcina mazei	*Mm*	(2002) Srinivasan et al.[15]	GGAAACCTGATCATGTAGATCGAATG GACTCTAAATCCGTTCAGCCGGGTTA GATTCCCGGGGTTTCCGCCA
Methanococcoides methylutens	*Met*	This work	GGAGACTTGATCATGTAGATCGAACG GACTCTAAATCCGTTCAGCCGGGTTA GATTCCCGGAGTTTCCGCCA
Methanolobus profundi	*Pro*	This work	GGAAATCAGATCATGTTGATCGAATG GACTCTAAATCCGTTCAGTCGGGTTA AATTCCCGAGGTTTCCGCCA
Methanolobus psychrotolerans sp. YSF-03	*Psy*	This work	GGAAATCGGATCAGGTTGATCGAATG GACTCTAAATCCGTTCAGTCGGGTTA AATTCCCGGGGTTTCCGCCA
Methanosarcina spelaei	*Spe*	This work	GGAAATCTGATCATGTAGATCGAATG GACTCTAAATCCGTTCAGCCGGGTTA GATTCCCGGGGTTTCCGCCA
Methanolobus vulcani	*Vul*	This work	GGAAATCAGATCATGTTGATCAAATG GACTCTAAATCCGTTTAGCCGGGTTA AATTCCCGGGGTTTCCGCCA
Methanosalsum zhilinae	*Zhi*	This work	GGAAACCTGATCATGTAGATCAAATG GACTCTAAATCCGTTCAGCCGGGTTA GATTCCCGGGGTTTCCGCCA

Supplementary Table 5

Sequences of all *Alv*$^{\Delta NPyl}$tRNA variants used in this work. Mutated nucleotides are marked in lower case and bold.

Variant	Sequence
Alv$^{\Delta NPyl}$tRNA(6)	GGGGGACGGTCCGGCGACCAGCGGGTCTCTAAAACCTAGC**atag**CGGGGTTCGAC**a**CCCCGGTCTCTCGCCA
Alv$^{\Delta NPyl}$tRNA(8)	GGGGGACGGTCCGGCGACCAGCGGGTCTCTAAAACCTAGC**cttg**CGGGGTTCGAC**a**CCCCGGTCTCTCGCCA
Alv$^{\Delta NPyl}$tRNA(10)	GGGGGACGGTCCGGCGACCAGCGGGTCTCTAAAACCTAGC**ttag**CGGGGTTCGAC**a**CCCCGGTCTCTCGCCA
Alv$^{\Delta NPyl}$tRNA(11)	GGGGGACGGTCCGGCGACCAGCGGGTCTCTAAAACCTAGC**aaa**gCGGGGTTCGAC**c**CCCCGGTCTCTCGCCA
Alv$^{\Delta NPyl}$tRNA(15)	GGGGGACGGTCCGGCGACCAGCGGGTCTCTAAAACCTAGC**atca**gCGGGGTTCGAC**a**CCCCGGTCTCTCGCCA
Alv$^{\Delta NPyl}$tRNA(17)	GGGGGACGGTCCGGCGACCAGCGGGTCTCTAAAACCTAGC**gtaa**gCGGGGTTCGAC**a**CCCCGGTCTCTCGCCA
Alv$^{\Delta NPyl}$tRNA(19)	GGGGGACGGTCCGGCGACCAGCGGGTCTCTAAAACCTAGC**agg**agCGGGGTTCGAC**t**CCCCGGTCTCTCGCCA
Alv$^{\Delta NPyl}$tRNA(20)	GGGGGACGGTCCGGCGACCAGCGGGTCTCTAAAACCTAGC**aca**agCGGGGTTCGAC**a**CCCCGGTCTCTCGCCA
Alv$^{\Delta NPyl}$tRNA(21)	GGGGGACGGTCCGGCGACCAGCGGGTCTCTAAAACCTAGC**ataa**gCGGGGTTCGAC**c**CCCCGGTCTCTCGCCA
Alv$^{\Delta NPyl}$tRNA(22)	GGGGGACGGTCCGGCGACCAGCGGGTCTCTAAAACCTAGC**tcaa**ggCGGGGTTCGAC**t**CCCCGGTCTCTCGCCA
Alv$^{\Delta NPyl}$tRNA(23)	GGGGGACGGTCCGGCGACCAGCGGGTCTCTAAAACCTAGC**cttt**agCGGGGTTCGACGCCCCGGTCTCTCGCCA
Alv$^{\Delta NPyl}$tRNA(24)	GGGGGACGGTCCGGCGACCAGCGGGTCTCTAAAACCTAGC**atcg**tgCGGGGTTCGAC**t**CCCCGGTCTCTCGCCA

Supplementary Table 6

Sequences of *Int*$^{\Delta NPyl}$tRNA variable loop variants used in this work. Mutated nucleotides are marked in lower case and bold.

Variant	Sequence
Int$^{\Delta NPyl}$tRNA(^{V}B01)	GGAGTGTTGGTCCGGGACCACCAGGCCTCTAAAGCCACGGC**tt** GCCGGGTTCAACTCCCGGGCACTTCGCCA
Int$^{\Delta NPyl}$tRNA(^{V}B03)	GGAGTGTTGGTCCGGGACCACCAGGCCTCTAAAGCCACGGC**t** AGCCGGGTTCAACTCCCGGGCACTTCGCCA
Int$^{\Delta NPyl}$tRNA(^{V}C08)	GGAGTGTTGGTCCGGGACCACCAGGCCTCTAAAGCCACGG**gat** GCCGGGTTCAACTCCCGGGCACTTCGCCA
Int$^{\Delta NPyl}$tRNA(^{V}C10)	GGAGTGTTGGTCCGGGACCACCAGGCCTCTAAAGCCACGG**tt**A GCCGGGTTCAACTCCCGGGCACTTCGCCA
Int$^{\Delta NPyl}$tRNA(^{V}D01)	GGAGTGTTGGTCCGGGACCACCAGGCCTCTAAAGCCACGG**ga**C AGCCGGGTTCAACTCCCGGGCACTTCGCCA
Int$^{\Delta NPyl}$tRNA(^{V}G04)	GGAGTGTTGGT**t**CGG**t**GACCACCAGGCCTCTAAAGCCACGGCA**a** GCCGGGTTC**g**ACTCCCGGGCACTTCGCCA
Int$^{\Delta NPyl}$tRNA(^{V}IF01)	GGAGTGTTGGTCCGGGACCACCAGGCCTCTAAAGCCACGGCA**c** **a**GCCGGGTTCAACTCCCGGGCACTTCGCCA

Supplementary Table 7

Sequences of *Int*$^{\Delta NPyl}$tRNA acceptor stem variants used in this work. Mutated nucleotides are marked in lower case and bold.

Variant	Sequence
Int$^{\Delta NPyl}$tRNA(^{A}5)	GG**t**G**aca**TGGTCCGGGACCACCAGGCCTCTAAAGCCACGGCAGC CGGGTTCAACTCCCGG**tgt**C**a**TCGCCA
Int$^{\Delta NPyl}$tRNA(^{A}6)	GG**t**G**aac**TGGTCCGGGACCACCAGGCCTCTAAAGCCACGGCAGC CGGGTTCAACTCCCGGG**tt**C**a**TCGCCA
Int$^{\Delta NPyl}$tRNA(^{A}7)	GGA**aacc**TGGTCCGGGACCACCAGGCCTCTAAAGCCACGGCAGC CGGGTTCAACTCCCGGG**gtt**TTCGCCA
Int$^{\Delta NPyl}$tRNA(^{A}9)	GGA**cc**G**a**TGGTCCGGGACCACCAGGCCTCTAAAG**t**CACGGCAGC CGGGTTCAACTCCCGG**t**C**ac**TTCGCCA
Int$^{\Delta NPyl}$tRNA(^{A}10)	GGAGT**ta**TGGTCCGGGACCACCAGGCCTCTAAAG**t**CACGGCAGC CGGGTTCAACTCCCGG**ta**ACTTCGCCA
Int$^{\Delta NPyl}$tRNA(^{A}11)	GG**gc**T**aa**TGGTCCGGGACCACCAGGCCTCTAAAGCCACGGCAGC CGGGTTCAACTCCCGG**tt**A**gc**TCGCCA
Int$^{\Delta NPyl}$tRNA(^{A}12)	GGAGT**tc**TGGTCCGGGACCACCAGGCCTCTAAAGCCACGGCAGC CGGGTTCAACTCCCGGG**a**ACTTCGCCA
Int$^{\Delta NPyl}$tRNA(^{A}13)	GG**t**GT**tc**TGGTCCGGGACCACC**g**GGCCTCTAAAGCCACGGCAGCC GGGTTCAACTCCCGGG**a**AC**a**TCGCCA
Int$^{\Delta NPyl}$tRNA(^{A}14)	GGAGT**aa**TGGTCCGGGACCACCAGGCCTCTAAAGCCACGGCAG CCGGGTTCAACTCCCGG**tt**ACTTCGCCA
Int$^{\Delta NPyl}$tRNA(^{A}16)	GG**tc**T**ca**TGGTCCGGGACCACCAGGCCTCTAAAGCCACGGCAGC CGGGTTCAACTCCCGG**tg**A**ga**TCGCCA
Int$^{\Delta NPyl}$tRNA(^{A}17)	GG**c**G**aac**TGGTCCGGGACCACCAGGCCTCTAAAGCCACGGCAGC CGGGTTCAACTCCCGGG**tt**C**g**TCGCCA

Supplementary Table 8

Sequences of *Int*$^{\Delta NPyl}$tRNA hybrids used in this work. Mutated nucleotides are marked in lower case and bold.

Variant	Sequence
Int$^{\Delta NPyl}$tRNA(^{A}5, ^{V}B03)	GG**t**G**aca**TGGTCCGGGACCACCAGGCCTCTAAAGCCACGGC**t**AGC CGGGTTCAACTCCCGG**tgt**C**a**TCGCCA
Int$^{\Delta NPyl}$tRNA(^{A}6, ^{V}B03)	GG**t**G**aac**TGGTCCGGGACCACCAGGCCTCTAAAGCCACGGC**t**AGC CGGGTTCAACTCCCGGG**tt**C**a**TCGCCA
Int$^{\Delta NPyl}$tRNA(^{A}7, ^{V}B03)	GGA**aacc**TGGTCCGGGACCACCAGGCCTCTAAAGCCACGGC**t**AGC CGGGTTCAACTCCCGGG**gtt**TTCGCCA
Int$^{\Delta NPyl}$tRNA(^{A}9, ^{V}B03)	GGA**cc**G**a**TGGTCCGGGACCACCAGGCCTCTAAAG**t**CACGGC**t**AGC CGGGTTCAACTCCCGG**tCac**TTCGCCA
Int$^{\Delta NPyl}$tRNA(^{A}10, ^{V}B03)	GGAGT**ta**TGGTCCGGGACCACCAGGCCTCTAAAG**t**CACGGC**t**AGC CGGGTTCAACTCCCGG**ta**ACTTCGCCA
Int$^{\Delta NPyl}$tRNA(^{A}11, ^{V}B03)	GG**gc**T**aa**TGGTCCGGGACCACCAGGCCTCTAAAGCCACGGC**t**AGC CGGGTTCAACTCCCGG**tt**A**gc**TCGCCA
Int$^{\Delta NPyl}$tRNA(^{A}12, ^{V}B03)	GGAGT**tc**TGGTCCGGGACCACCAGGCCTCTAAAGCCACGGC**t**AG CCGGGTTCAACTCCCGGG**a**ACTTCGCCA
Int$^{\Delta NPyl}$tRNA(^{A}13, ^{V}B03)	GG**t**GT**tc**TGGTCCGGGACCACC**g**GGCCTCTAAAGCCACGGC**t**AGC CGGGTTCAACTCCCGGG**a**AC**a**TCGCCA
Int$^{\Delta NPyl}$tRNA(^{A}14, ^{V}B03)	GGAGT**aa**TGGTCCGGGACCACCAGGCCTCTAAAGCCACGGC**t**AG CCGGGTTCAACTCCCGG**tt**ACTTCGCCA
Int$^{\Delta NPyl}$tRNA(^{A}16, ^{V}B03)	GG**tc**T**ca**TGGTCCGGGACCACCAGGCCTCTAAAGCCACGGC**t**AGC CGGGTTCAACTCCCGG**tg**A**ga**TCGCCA
Int$^{\Delta NPyl}$tRNA(^{A}17, ^{V}B03)	GG**c**G**aac**TGGTCCGGGACCACCAGGCCTCTAAAGCCACGGC**t**AGC CGGGTTCAACTCCCGGG**tt**C**g**TCGCCA
Int$^{\Delta NPyl}$tRNA(^{A}5, ^{V}C10)	GG**t**G**aca**TGGTCCGGGACCACCAGGCCTCTAAAGCCACGG**tt**AGC CGGGTTCAACTCCCGG**tgt**C**a**TCGCCA
Int$^{\Delta NPyl}$tRNA(^{A}6, ^{V}C10)	GG**t**G**aac**TGGTCCGGGACCACCAGGCCTCTAAAGCCACGG**tt**AGC CGGGTTCAACTCCCGGG**tt**C**a**TCGCCA
Int$^{\Delta NPyl}$tRNA(^{A}7, ^{V}C10)	GGA**aacc**TGGTCCGGGACCACCAGGCCTCTAAAGCCACGG**tt**AGC CGGGTTCAACTCCCGGG**gtt**TTCGCCA
Int$^{\Delta NPyl}$tRNA(^{A}9, ^{V}C10)	GGA**cc**G**a**TGGTCCGGGACCACCAGGCCTCTAAAG**t**CACGG**tt**AGC CGGGTTCAACTCCCGG**tCac**TTCGCCA
Int$^{\Delta NPyl}$tRNA(^{A}10, ^{V}C10)	GGAGT**ta**TGGTCCGGGACCACCAGGCCTCTAAAG**t**CACGG**tt**AGC CGGGTTCAACTCCCGG**ta**ACTTCGCCA
Int$^{\Delta NPyl}$tRNA(^{A}11, ^{V}C10)	GG**gc**T**aa**TGGTCCGGGACCACCAGGCCTCTAAAGCCACGG**tt**AGC CGGGTTCAACTCCCGG**tt**A**gc**TCGCCA
Int$^{\Delta NPyl}$tRNA(^{A}12, ^{V}C10)	GGAGT**tc**TGGTCCGGGACCACCAGGCCTCTAAAGCCACGG**tt**AGC CGGGTTCAACTCCCGGG**a**ACTTCGCCA
Int$^{\Delta NPyl}$tRNA(^{A}13, ^{V}C10)	GG**t**GT**tc**TGGTCCGGGACCACC**g**GGCCTCTAAAGCCACGG**tt**AGCC GGGTTCAACTCCCGGG**a**AC**a**TCGCCA
Int$^{\Delta NPyl}$tRNA(^{A}14, ^{V}C10)	GGAGT**aa**TGGTCCGGGACCACCAGGCCTCTAAAGCCACGG**tt**AG CCGGGTTCAACTCCCGG**tt**ACTTCGCCA
Int$^{\Delta NPyl}$tRNA(^{A}16, ^{V}C10)	GG**tc**T**ca**TGGTCCGGGACCACCAGGCCTCTAAAGCCACGG**tt**AGC CGGGTTCAACTCCCGG**tg**A**ga**TCGCCA
Int$^{\Delta NPyl}$tRNA(^{A}17, ^{V}C10)	GG**c**G**aac**TGGTCCGGGACCACCAGGCCTCTAAAGCCACGG**tt**AGC CGGGTTCAACTCCCGGG**tt**C**g**TCGCCA

Supplementary Table 9

Sequence identity matrix of $^{\Delta NPyl}$tRNA combinations.

	RumEn	***Lum1***	***Int***	***Lum2***	***Sheng***	***030***	***Term***	***G1***	***1R26***	***H5***	***Alv***
RumEn	100	95.83	95.77	91.55	88.73	88.73	68.57	70	67.61	70.42	71.83
Lum1	95.83	100	95.77	91.55	88.73	88.73	67.14	70	67.61	70.42	71.83
Int	95.77	95.77	100	92.96	90.14	92.96	68.57	68.12	65.71	68.57	70
Lum2	91.55	91.55	92.96	100	94.37	91.55	70	69.57	67.14	70	71.43
Sheng	88.73	88.73	90.14	94.37	100	88.73	68.57	71.01	67.14	70	71.43
30	88.73	88.73	92.96	91.55	88.73	100	71.43	69.57	68.57	71.43	72.86
Term	68.57	67.14	68.57	70	68.57	71.43	100	71.01	78.57	77.14	82.86
G1	70	70	68.12	69.57	71.01	69.57	71.01	100	84.29	88.57	85.71
1R26	67.61	67.61	65.71	67.14	67.14	68.57	78.57	84.29	100	94.37	94.37
H5	70.42	70.42	68.57	70	70	71.43	77.14	88.57	94.37	100	94.37
Alv	71.83	71.83	70	71.43	71.43	72.86	82.86	85.71	94.37	94.37	100

Supplementary Table 10

Sequence identity matrix of ΔNPylRS combinations.

	Alv	*1R26*	*BRNA*	*Term*	*G1*	*H5*	*RumEn*	*Int*	*S30*	*Lum2*	*Lum1*	**Sheng**
Alv	100	82.18	78.75	65.33	64.47	65.33	51.09	49.64	49.64	47.81	50	47.25
1R26	82.18	100	75.46	63.87	64.1	63.5	52.55	49.64	49.27	46.72	48.54	46.52
BRNA	78.75	75.46	100	62.5	60.52	60.66	47.79	44.12	46.69	45.96	47.79	44.85
Term	65.33	63.87	62.5	100	63.74	62.77	56.2	52.92	52.19	50.73	50.36	52.38
G1	64.47	64.1	60.52	63.74	100	60.81	50.92	54.21	50.18	45.42	46.15	49.26
H5	65.33	63.5	60.66	62.77	60.81	100	50.73	49.27	51.46	50	51.46	49.08
RumEn	51.09	52.55	47.79	56.2	50.92	50.73	100	68.36	65.09	59.27	54.55	56.2
Int	49.64	49.64	44.12	52.92	54.21	49.27	68.36	100	62.91	57.09	54.91	55.84
S30	49.64	49.27	46.69	52.19	50.18	51.46	65.09	62.91	100	66.55	56.36	57.3
Lum2	47.81	46.72	45.96	50.73	45.42	50	59.27	57.09	66.55	100	56.36	56.57
Lum1	50	48.54	47.79	50.36	46.15	51.46	54.55	54.91	56.36	56.36	100	58.03
Sheng	47.25	46.52	44.85	52.38	49.26	49.08	56.2	55.84	57.3	56.57	58.03	100

Supplementary Table 11

Primers used for library generation.

Primer name	**Library**	**Sequence**
int-aaBsbIF	Variable Loop 4-6	CGGgaagacCGCCAactagtATCCTTAGCGAAAGCTAAG GATTTTTTTTaagcttGGCACTGGCCGTCGTTTTAC
int-aaL3-4-2BbsIR	Variable Loop 4	AAGgaagacAGTTGGCGAAGTGCCCGGGnnTTGAACC CGGnnnnCCGTGGCTTTAGAGGCCTGGTGGTCnnnG ACCAACACTCCagatctagcgttacaagtatTACACAAA
int-aaL3-5-2BbsIR	Variable Loop 5	AAGgaagacAGTTGGCGAAGTGCCCGGGnnTTGAACC CGGnnnnnCCGTGGCTTTAGAGGCCTGGTGGTCnnnG ACCAACACTCCagatctagcgttacaagtatTACACAAA
int-aaL3-6-2BbsIR	Variable Loop 6	AAGgaagacAGTTGGCGAAGTGCCCGGGnnTTGAACC CGGnnnnnnCCGTGGCTTTAGAGGCCTGGTGGTCnnn GACCAACACTCCagatctagcgttacaagtatTACACAAA
oDD299	Acceptor Stem	AAGGAAGACCCTTAGCTTTCGCTAAGGATACTAGT TGGCGAnnnnnCCGGGAGTTGAACCCGGCTGCCGTG GCTTTAGAGGCCTGGTGGTCCCGGACCAnnnnnCCA GATCTAGCGTTACAAGTATTACACAAAG
oDD301	Acceptor Stem	CCGgaagacAGCTAAGGATTTTTTTTaagcttGGCACTGG CCGTCGTTTTACAACGTCGTGACTGGGAAAACCCT GGCGTTACCCAAC

Plasmids

pKW1 *Alv*PylRS *Alv*$^{\Delta NPyl}$tRNA

Sequence feature	Nucleotide position
lpp promoter	534 - 558
Alv$^{\Delta NPyl}$tRNA	579 - 649
rrnC terminator	656 - 684
SpR antibiotic resistance	1369 - 2160
pMB1 replication origin	2316 - 2904

tgcagtttcaaacgctaaattgcctgatgcgctacgcttatcaggcctacatgatctctgcaatatattgagtttgcgtgcttttgtagg
ccggataaggcgttcacgccgcatccggcaagaaacagcaaacaatccaaaacgccgcgttcggcggtcgacacagatgtag
gtgttccacagggtagccagcagcatcctgcgatgcagatccggaacataatggtgcagggcgcttgtttcggcgtgggtatgg
tggcaggccccgtggccgggggactgttgggcgctgccggcacctgtcctacgagttgcatgataaagaatacagtcataagt
gcggcgacgatagtcatgccccgcgcccaccggaaggagctaccggcagcggtgcggactgttgtaactcagaataagaaat
gaggccgctcatggcgttctgttgcccgtctcactggtgaaaagaaaaacaaccctggcgccgcttctttgagcgaacgatcaaa
aataagtggcgccccatcaaaaaaataTTCTCAACATAAAAAACTTTGTGTAatacttgtaacgctagatc
tGGGGGACGGTCCGGCGACCAGCGGGTCTCTAAAACCTAGCCAGCGGGGTTC
GACGCCCCGGTCTCTCGCCAactagtATCCTTAGCGAAAGCTAAGGATTTTTTTTa
agcttggcactggccgtcgttttacaacgtcgtgactgggaaaaccctggcgttacccaacttaatcgccttgcagcacatccccc
tttcgccagacgctctcccttatgcgactcctgcattaggaagcagcccagtagtaggttgaggccgttgagcaccgccgccgc
aaggaatggtgcatgcaaggagcccgagatgcgccgcgtgcggctgctggagatggcggacgcgatggatatgttctggcg
gccgcataatgcttaagtcgaacagaaagtaatcgtattgtacacggccgcataatcgaaattaaattggatatcggcaggccac
gcgatcgctgacgtcggtaccctcgagtctggtaaagaaaccgctgctgcgaaatttgaacgccagcacatggactcgtctacta
gcgcagcttaattaacctaggctgctgccaccgctgagcaataactagcataacccctttggggcctctaaacgggtcttgaggg
gttttttgctgaaacctcaggcatttgagaagcacacggtcacactgcttccggtagtcaataaaccggtaaaccagcaatagaca
taagcggctatttaacgaccctgccctgaaccgacgaccgggtcatcgtggccggatcttgcggcccctcggcttgaacgaatt
gttagacaTTATTTGCCGACTACCTTGGTGATCTCGCCTTTCACGTAGTGGACAAA
TTCTTCCAACTGATCTGCGCGCGAGGCCAAGCGATCTTCTTCTTGTCCAAGAT
AAGCCTGTCTAGCTTCAAGTATGACGGGCTGATACTGGGCCGGCAGGCGCTC
CATTGCCCAGTCGGCAGCGACATCCTTCGGCGCGATTTTGCCGGTTACTGCGC
TGTACCAAATGCGGGACAACGTAAGCACTACATTTCGCTCATCGCCAGCCCA
GTCGGGCGGCGAGTTCCATAGCGTTAAGGTTTCATTTAGCGCCTCAAATAGAT
CCTGTTCAGGAACCGGATCAAAGAGTTCCTCCGCCGCTGGACCTACCAAGGC
AACGCTATGTTCTCTTGCTTTTGTCAGCAAGATAGCCAGATCAATGTCGATCG
TGGCTGGCTCGAAGATACCTGCAAGAATGTCATTGCGCTGCCATTCTCCAAAT
TGCAGTTCGCGCTTAGCTGGATAACGCCACGGAATGATGTCGTCGTGCACAA
CAATGGTGACTTCTACAGCGCGGAGAATCTCGCTCTCTCCAGGGGAAGCCGA
AGTTTCCAAAAGGTCGTTGATCAAAGCTCGCCGCGTTGTTTCATCAAGCCTTA
CGGTCACCGTAACCAGCAAATCAATATCACTGTGTGGCTTCAGGCCGCCATC
CACTGCGGAGCCGTACAAATGTACGGCCAGCAACGTCGGTTCGAGATGGCGC
TCGATGACGCCAACTACCTCTGATAGTTGAGTCGATACTTCGGCGATCACCGC
TTCCCTCATactcttccttttcaatattattgaagcatttatcagggttattgtctcatgagcggatacatatttgaatgtatttag

aaaaataaacaaatagctagcgagcaaaaggccagcaaaaggccaggaaccgtaaaaaggccgcgttgctggcgttTTTC
CATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGA
GGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAG
CTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCG
CCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTAT
CTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCC
CGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACC
CGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAG
CAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAAC
TACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAG
TTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGC
TGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAA
GGATCTCAAgaagatcctttgatcttttctacggggtctgacgctcagt

pKW1 *Alv*PylRS *Alv*$^{\Delta NPyl}$tRNA

Sequence feature	**Nucleotide position**
glnS promoter	14 - 39
*Alv*PylRS	81 - 908
lpp promoter	1443 - 1467
Alv$^{\Delta NPyl}$tRNA	1488 - 1558
rrnC terminator	1565 - 1593
SpR antibiotic resistance	2278 - 3069
pMB1 replication origin	3225 - 3813

ggatcctcgggagTTGTCAGCCTGTCCCGCTTATAAGATcatacgccgttatacgttgtttacgctttgag
gaatcccatATGACGGTGAAATACACAGACGCGCAGATTCAGCGTTTACGCGAGT
ATGGAAATGGGACTTACGAACAGAAGGTCTTTGAGGATCTGGCGAGTCGTGA
TGCAGCGTTTTCAAAAGAGATGTCCGTTGCGTCTACGGATAATGAGAAGAAA
ATCAAGGGTATGATCGCCAACCCTAGTCGTCATGGGCTGACCCAATTAATGA
ATGACATTGCGGATGCCCTGGTCGCAGAGGGTTTTATCGAAGTGCGCACTCC
AATCTTTATTTCTAAGGATGCTTTGGCACGCATGACCATCACCGAGGACAAAC
CATTATTTAAACAGGTTTTTTGGATCGACGAAAAACGTGCCTTACGCCCAATG
CTGGCTCCAAACCTGTATAGTGTCATGCGTGATCTTCGCGATCACACAGATGG
CCCTGTTAAGATTTTTGAGATGGGATCTTGTTTTCGTAAGGAGTCACATTCTG
GCATGCATCTGGAAGAGTTTACGATGCTTAATCTGGTCGATATGGGGCCTCGC
GGCGACGCGACTGAAGTATTGAAGAATTACATTAGCGTAGTTATGAAGGCAG
CCGGGCTTCCTGACTATGACTTGGTCCAAGAAGAGTCGGATGTATATAAGGA
AACTATCGATGTTGAGATCAACGGACAAGAAGTTTGCAGCGCGGCGGTAGGT
CCCCATTACCTGGATGCCGCTCACGATGTTCACGAGCCTTGGAGCGGCGCGG
GTTTCGGCCTTGAACGCCTTTTAACTATCCGCGAGAAATATTCCACCGTGAAA
AAGGGCGGCGCGTCAATCTCCTATCTGAATGGGGCCAAAATTAATTAActgcagt
ttcaaacgctaaattgcctgatgcgctacgcttatcaggcctacatgatctctgcaatatattgagtttgcgtgcttttgtaggccgga
taaggcgttcacgccgcatccggcaagaaacagcaaacaatccaaaacgccgcgttcggcggtcgacacagatgtaggtgttc

cacagggtagccagcagcatcctgcgatgcagatccggaacataatggtgcagggcgcttgtttcggcgtgggtatggtggca
ggccccgtggccgggggactgttgggcgctgccggcacctgtcctacgagttgcatgataaagaagacagtcataagtgcgg
cgacgatagtcatgccccgcgcccaccggaaggagctaccggcagcggtgcggactgttgtaactcagaataagaaatgagg
ccgctcatggcgttctgttgcccgtctcactggtgaaaagaaaaacaaccctggcgccgcttctttgagcgaacgatcaaaaata
agtggcgccccatcaaaaaaataTTCTCAACATAAAAAACTTTGTGTAatacttgtaacgctagatctG
GGGGACGGTCCGGCGACCAGCGGGTCTCTAAAACCTAGCCAGCGGGGTTCGA
CGCCCCGGTCTCTCGCCAactagtATCCTTAGCGAAAGCTAAGGATTTTTTTTaagct
tggcactggccgtcgttttacaacgtcgtgactgggaaaaccctggcgttacccaacttaatcgccttgcagcacatcccccttc
gccagacgctctcccttatgcgactcctgcattaggaagcagcccagtagtaggttgaggccgttgagcaccgccgccgcaag
gaatggtgcatgcaaggagcccgagatgcgccgcgtgcggctgctggagatggcggacgcgatggatatgttctggcggcc
gcataatgcttaagtcgaacagaaagtaatcgtattgtacacggccgcataatcgaaattaaattggatatcggcaggccacgcg
atcgctgacgtcggtaccctcgagtctggtaaagaaaccgctgctgcgaaatttgaacgccagcacatggactcgtctactagcg
cagcttaattaacctaggctgctgccaccgctgagcaataactagcataacccttggggcctctaaacgggtcttgaggggttttt
tgctgaaacctcaggcatttgagaagcacacggtcacactgcttccggtagtcaataaaccggtaaaccagcaatagacataag
cggctatttaacgaccctgccctgaaccgacgaccgggtcatcgtggccggatcttgcggcccctcggcttgaacgaattgttag
acaTTATTTGCCGACTACCTTGGTGATCTCGCCTTTCACGTAGTGGACAAATTC
TTCCAACTGATCTGCGCGCGAGGCCAAGCGATCTTCTTCTTGTCCAAGATAAG
CCTGTCTAGCTTCAAGTATGACGGGCTGATACTGGGCCGGCAGGCGCTCCATT
GCCCAGTCGGCAGCGACATCCTTCGGCGCGATTTTGCCGGTTACTGCGCTGTA
CCAAATGCGGGACAACGTAAGCACTACATTTCGCTCATCGCCAGCCCAGTCG
GGCGGCGAGTTCCATAGCGTTAAGGTTTCATTTAGCGCCTCAAATAGATCCTG
TTCAGGAACCGGATCAAAGAGTTCCTCCGCCGCTGGACCTACCAAGGCAACG
CTATGTTCTCTTGCTTTTGTCAGCAAGATAGCCAGATCAATGTCGATCGTGGC
TGGCTCGAAGATACCTGCAAGAATGTCATTGCGCTGCCATTCTCCAAATTGCA
GTTCGCGCTTAGCTGGATAACGCCACGGAATGATGTCGTCGTGCACAACAAT
GGTGACTTCTACAGCGCGGAGAATCTCGCTCTCTCCAGGGGAAGCCGAAGTT
TCCAAAAGGTCGTTGATCAAAGCTCGCCGCGTTGTTTCATCAAGCCTTACGGT
CACCGTAACCAGCAAATCAATATCACTGTGTGGCTTCAGGCCGCCATCCACT
GCGGAGCCGTACAAATGTACGGCCAGCAACGTCGGTTCGAGATGGCGCTCGA
TGACGCCAACTACCTCTGATAGTTGAGTCGATACTTCGGCGATCACCGCTTCC
CTCATactcttccttttcaatattattgaagcatttatcagggttattgtctcatgagcggatacatatttgaatgtatttagaaaaat
aaacaaatagctagcgagcaaaaggccagcaaaaggccaggaaccgtaaaaaggccgcgttgctggcgttTTTCCAT
AGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGT
GGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTC
CCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCT
TTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTC
AGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGT
TCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGG
TAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAG
AGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTAC
GGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTA
CCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGG
TAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGA
TCTCAAgaagatcctttgatcttttctacggggtctgacgctcagtggaacgaaaactcacgttaagggattttgg

pBAD GFP(150TAG)$_{His6}$

Sequence feature	Nucleotide position
p15A replication origin	349 - 1175
TetR antibiotic resistance	1347 - 2537
AraC arabinose transcriptional regulator	2845 - 3723
sfGFP(150TAG)$_{His6}$	4068 - 4811

gagagaagattttcagcctgatacagattaaatcagaacgcagaagcggtctgataaaacagaatttgcctggcggcagtagcg
cggtggtcccacctgaccccatgccgaactcagaagtgaaacgccgtagcgccgatggtagtgtggggtctccccatgcgaga
gtagggaactgccaggcatcaaataaaacgaaaggctcagtcgaaagactgggcctttcgttttatctgttgtttgtcggtgaacg
ctctcctgagtaggacaaatccgccgggagctgtccctcctgttcagctactgacggggtggtgcgtaacggcaaaagcaccg
ccggacatcagcGCTAGCGGAGTGTATACTGGCTTACTATGTTGGCACTGATGAGGG
TGTCAGTGAAGTGCTTCATGTGGCAGGAGAAAAAAGGCTGCACCGGTGCGTC
AGCAGAATATGTGATACAGGATATATTCCGCTTCCTCGCTCACTGACTCGCTA
CGCTCGGTCGTTCGACTGCGGCGAGCGGAAATGGCTTACGAACGGGGCGGAG
ATTTCCTGGAAGATGCCAGGAAGATACTTAACAGGGAAGTGAGAGGGCCGC
GGCAAAGCCGTTTTTCCATAGGCTCCGCCCCCCTGACAAGCATCACGAAATC
TGACGCTCAAATCAGTGGTGGCGAAACCCGACAGGACTATAAAGATACCAGG
CGTTTCCCCCTGGCGGCTCCCTCGTGCGCTCTCCTGTTCCTGCCTTTCGGTTTA
CCGGTGTCATTCCGCTGTTATGGCCGCGTTTGTCTCATTCCACGCCTGACACT
CAGTTCCGGGTAGGCAGTTCGCTCCAAGCTGGACTGTATGCACGAACCCCCC
GTTCAGTCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCC
GGAAAGACATGCAAAAGCACCACTGGCAGCAGCCACTGGTAATTGATTTAGA
GGAGTTAGTCTTGAAGTCATGCGCCGGTTAAGGCTAAACTGAAAGGACAAGT
TTTGGTGACTGCGCTCCTCCAAGCCAGTTACCTCGGTTCAAAGAGTTGGTAGC
TCAGAGAACCTTCGAAAAACCGCCCTGCAAGGCGGTTTTTTCGTTTTCAGAGC
AAGAGATTACGCGCAGACCAAAACGATCTCAAGAAGATCATCTTATTAatcagat
aaaatatttctagatttcagtgcaatttatctcttcaaatgtagcacctgaagtcagccccatacgatataagttgtaattctcatgtttg
acagcttatcatcgataagctttaatgcggtagtttatcacagttaaattgctaacgcagtcaggcaccgtgtATGAAATCT
AACAATGCGCTCATCGTCATCCTCGGCACCGTCACCCTGGATGCTGTAGGCAT
AGGCTTGGTTATGCCGGTACTGCCGGGCCTCTTGCGGGATATCGTCCATTCCG
ACAGCATCGCCAGTCACTATGGCGTGCTGCTAGCGCTATATGCGTTGATGCA
ATTTCTATGCGCACCCGTTCTCGGAGCACTGTCCGACCGCTTTGGCCGCCGCC
CAGTCCTGCTCGCTTCGCTACTTGGAGCCACTATCGACTACGCGATCATGGCG
ACCACACCCGTCCTGTGGATCCTCTACGCCGGACGCATCGTGGCCGGCATCA
CCGGCGCCACAGGTGCGGTTGCTGGCGCCTATATCGCCGACATCACCGATGG
GGAAGATCGGGCTCGCCACTTCGGGCTCATGAGCGCTTGTTTCGGCGTGGGT
ATGGTGGCAGGCCCCGTGGCCGGGGGACTGTTGGGCGCCATCTCCTTGCATG
CACCATTCCTTGCGGCGGCGGTGCTCAACGGCCTCAACCTACTACTGGGCTGC
TTCCTAATGCAGGAGTCGCATAAGGGAGAGCGTCGACCGATGCCCTTGAGAG
CCTTCAACCCAGTCAGCTCCTTCCGGTGGGCGCGGGGCATGACTATCGTCGCC
GCACTTATGACTGTCTTCTTTATCATGCAACTCGTAGGACAGGTGCCGGCAGC
GCTCTGGGTCATTTTCGGCGAGGACCGCTTTCGCTGGAGCGCGACGATGATC
GGCCTGTCGCTTGCGGTATTCGGAATCTTGCACGCCCTCGCTCAAGCCTTCGT

CACTGGTCCCGCCACCAAACGTTTCGGCGAGAAGCAGGCCATTATCGCCGGC
ATGGCGGCCGACGCGCTGGGCTACGTCTTGCTGGCGTTCGCGACGCGAGGCT
GGATGGCCTTCCCCATTATGATTCTTCTCGCTTCCGGCGGCATCGGGATGCCC
GCGTTGCAGGCCATGCTGTCCAGGCAGGTAGATGACGACCATCAGGGACAGC
TTCAAGGATCGCTCGCGGCTCTTACCAGCCTAACTTCGATCATTGGACCGCTG
ATCGTCACGGCGATTTATGCCGCCTCGGCGAGCACATGGAACGGGTTGGCAT
GGATTGTAGGCGCCGCCCTATACCTTGTCTGCCTCCCCGCGTTGCGTCGCGGT
GCATGGAGCCGGGCCACCTCGACCTGAatggaagccggcggcacctcgctaacggattcaccactcc
aagaattggagccaatcaattcttgcggagaactgtgaatgcgcaaaccaacccttggcagaacatatccatcgcgtgcggccg
cgtctccgggagctgcatgtgtcagaggttttcaccgtcatcaccgaaacgcgcgaggcagcagatcaattcgcgcgcgaagg
cgaagcggcatgcataatgtgcctgtcaaatggacgaagcagggattctgcaaaccctatgctactccgtcaagccgtcaattgt
ctgattcgttaccaaTTATGACAACTTGACGGCTACATCATTCACTTTTTCTTCACAACC
GGCACGGAACTCGCTCGGGCTGGCCCCGGTGCATTTTTTAAATACCCGCGAG
AAATAGAGTTGATCGTCAAAACCAACATTGCGACCGACGGTGGCGATAGGCA
TCCGGGTGGTGCTCAAAAGCAGCTTCGCCTGGCTGATACGTTGGTCCTCGCGC
CAGCTTAAGACGCTAATCCCTAACTGCTGGCGGAAAAGATGTGACAGACGCG
ACGGCGACAAGCAAACATGCTGTGCGACGCTGGCGATATCAAAATTGCTGTC
TGCCAGGTGATCGCTGATGTACTGACAAGCCTCGCGTACCCGATTATCCATCG
GTGGATGGAGCGACTCGTTAATCGCTTCCATGCGCCGCAGTAACAATTGCTC
AAGCAGATTTATCGCCAGCAGCTCCGAATAGCGCCCTTCCCCTTGCCCGGCGT
TAATGATTTGCCCAAACAGGTCGCTGAAATGCGGCTGGTGCGCTTCATCCGG
GCGAAAGAACCCCGTATTGGCAAATATTGACGGCCAGTTAAGCCATTCATGC
CAGTAGGCGCGCGGACGAAAGTAAACCCACTGGTGATACCATTCGCGAGCCT
CCGGATGACGACCGTAGTGATGAATCTCTCCTGGCGGGAACAGCAAAATATC
ACCCGGTCGGCAAACAAATTCTCGTCCCTGATTTTTCACCACCCCCTGACCGC
GAATGGTGAGATTGAGAATATAACCTTTCATTCCCAGCGGTCGGTCGATAAA
AAAATCGAGATAACCGTTGGCCTCAATCGGCGTTAAACCCGCCACCAGATGG
GCATTAAACGAGTATCCCGGCAGCAGGGGATCATTTTGCGCTTCAGCCATacttt
tcatactcccgccattcagagaagaaaccaattgtccatattgcatcagacattgccgtcactgcgtcttttactggctcttctcgcta
accaaaccggtaaccccgcttattaaaagcattctgtaacaaagcgggaccaaagccatgacaaaaacgcgtaacaaaagtgtc
tataatcacggcagaaaagtccacattgattatttgcacggcgtcacactttgctatgccatagcatttttatccataagattagcgga
tcctacctgacgctttttatcgcaactctctactgtttctccatacccgtttttgggctaacaggaggaattaaccATGGTTAG
CAAAGGTGAAGAACTGTTTACCGGCGTTGTGCCGATTCTGGTGGAACTGGAT
GGTGATGTGAATGGCCATAAATTTAGCGTTCGTGGCGAAGGCGAAGGTGATG
CGACCAACGGTAAACTGACCCTGAAATTTATTTGCACCACCGGTAAACTGCC
GGTTCCGTGGCCGACCCTGGTGACCACCCTGACCTATGGCGTTCAGTGCTTTA
GCCGCTATCCGGATCATATGAAACGCCATGATTTCTTTAAAAGCGCGATGCC
GGAAGGCTATGTGCAGGAACGTACCATTAGCTTCAAAGATGATGGCACCTAT
AAAACCCGTGCGGAAGTTAAATTTGAAGGCGATACCCTGGTGAACCGCATTG
AACTGAAAGGTATTGATTTTAAAGAAGATGGCAACATTCTGGGTCATAAACT
GGAATATAATTTCAACAGCCATTAGGTGTATATTACCGCCGATAAACAGAAA
AATGGCATCAAAGCGAACTTTAAAATCCGTCACAACGTGGAAGATGGTAGCG
TGCAGCTGGCGGATCATTATCAGCAGAATACCCCGATTGGTGATGGCCCGGT
GCTGCTGCCGGATAATCATTATCTGAGCACCCAGAGCGTTCTGAGCAAAGAT
CCGAATGAAAAACGTGATCATATGGTGCTGCTGGAATTTGTTACCGCCGCGG
GCATTACCCACGGTATGGATGAACTGTATAAAGGCAGCCACCATCATCATCA

CCATTAAagctcgagcgaagcttgggcccgaacaaaaactcatctcagaagaggatctgaatagcgccgtcgaccatca
tcatcatcatcattgagtttaaacggtctccagcttggctgttttggcggat

pBAD *Alv*PylRS CAT(111TAG) GFP(150TAG)$_{His6}$

Sequence feature	Nucleotide position
p15A replication origin	349 - 1175
TetR antibiotic resistance	1347 - 2537
glnS promoter	2658 - 2683
*Alv*PylRS-Ser(Gly$_4$Ser)$_4$His$_6$SerGlyStrep-tag II	2725 - 3663
CAT(111TAG)	4140 - 4799
AraC arabinose transcriptional regulator	5278 - 6156
sfGFP(150TAG)$_{His6}$	6501 - 7244

gagagaagattttcagcctgatacagattaaatcagaacgcagaagcggtctgataaaacagaatttgcctggcggcagtagcg
cggtggtcccacctgaccccatgccgaactcagaagtgaaacgccgtagcgccgatggtagtgtggggtctccccatgcgaga
gtagggaactgccaggcatcaaataaaacgaaaggctcagtcgaaagactgggcctttcgttttatctgttgtttgtcggtgaacg
ctctcctgagtaggacaaatccgccgggagctgtccctcctgttcagctactgacggggtggtgcgtaacggcaaaagcaccg
ccggacatcagcGCTAGCGGAGTGTATACTGGCTTACTATGTTGGCACTGATGAGGG
TGTCAGTGAAGTGCTTCATGTGGCAGGAGAAAAAAGGCTGCACCGGTGCGTC
AGCAGAATATGTGATACAGGATATATTCCGCTTCCTCGCTCACTGACTCGCTA
CGCTCGGTCGTTCGACTGCGGCGAGCGGAAATGGCTTACGAACGGGGCGGAG
ATTTCCTGGAAGATGCCAGGAAGATACTTAACAGGGAAGTGAGAGGGCCGC
GGCAAAGCCGTTTTTCCATAGGCTCCGCCCCCCTGACAAGCATCACGAAATC
TGACGCTCAAATCAGTGGTGGCGAAACCCGACAGGACTATAAAGATACCAGG
CGTTTCCCCCTGGCGGCTCCCTCGTGCGCTCTCCTGTTCCTGCCTTTCGGTTTA
CCGGTGTCATTCCGCTGTTATGGCCGCGTTTGTCTCATTCCACGCCTGACACT
CAGTTCCGGGTAGGCAGTTCGCTCCAAGCTGGACTGTATGCACGAACCCCCC
GTTCAGTCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCC
GGAAAGACATGCAAAAGCACCACTGGCAGCAGCCACTGGTAATTGATTTAGA
GGAGTTAGTCTTGAAGTCATGCGCCGGTTAAGGCTAAACTGAAAGGACAAGT
TTTGGTGACTGCGCTCCTCCAAGCCAGTTACCTCGGTTCAAAGAGTTGGTAGC
TCAGAGAACCTTCGAAAAACCGCCCTGCAAGGCGGTTTTTTCGTTTTCAGAGC
AAGAGTTTACGCGCAGACCAAAACGATCTCAAGAAGATCATCTTATTAatcagat
aaaatatttctagatttcagtgcaatttatctcttcaaatgtagcacctgaagtcagccccatacgatataagttgtaattctcatgtttg
acagcttatcatcgataagctttaatgcggtagtttatcacagttaaattgctaacgcagtcaggcaccgtgtATGAAATCT
AACAATGCGCTCATCGTCATCCTCGGCACCGTCACCCTGGATGCTGTAGGCAT
AGGCTTGGTTATGCCGGTACTGCCGGGCCTCTTGCGGGATATCGTCCATTCCG
ACAGCATCGCCAGTCACTATGGCGTGCTGCTAGCGCTATATGCGTTGATGCA
ATTTCTATGCGCACCCGTTCTCGGAGCACTGTCCGACCGCTTTGGCCGCCGCC
CAGTCCTGCTCGCTTCGCTACTTGGAGCCACTATCGACTACGCGATCATGGCG
ACCACACCCGTCCTGTGGATCCTCTACGCCGGACGCATCGTGGCCGGCATCA
CCGGCGCCACAGGTGCGGTTGCTGGCGCCTATATCGCCGACATCACCGATGG
GGAAGATCGGGCTCGCCACTTCGGGCTCATGAGCGCTTGTTTCGGCGTGGGT
ATGGTGGCAGGCCCCGTGGCCGGGGGACTGTTGGGCGCCATCTCCTTGCATG
CACCATTCCTTGCGGCGGCGGTGCTCAACGGCCTCAACCTACTACTGGGCTGC
TTCCTAATGCAGGAGTCGCATAAGGGAGAGCGTCGACCGATGCCCTTGAGAG

CCTTCAACCCAGTCAGCTCCTTCCGGTGGGCGCGGGGCATGACTATCGTCGCC
GCACTTATGACTGTCTTCTTTATCATGCAACTCGTAGGACAGGTGCCGGCAGC
GCTCTGGGTCATTTTCGGCGAGGACCGCTTTCGCTGGAGCGCGACGATGATC
GGCCTGTCGCTTGCGGTATTCGGAATCTTGCACGCCCTCGCTCAAGCCTTCGT
CACTGGTCCCGCCACCAAACGTTTCGGCGAGAAGCAGGCCATTATCGCCGGC
ATGGCGGCCGACGCGCTGGGCTACGTCTTGCTGGCGTTCGCGACGCGAGGCT
GGATGGCCTTCCCCATTATGATTCTTCTCGCTTCCGGCGGCATCGGGATGCCC
GCGTTGCAGGCCATGCTGTCCAGGCAGGTAGATGACGACCATCAGGGACAGC
TTCAAGGATCGCTCGCGGCTCTTACCAGCCTAACTTCGATCATTGGACCGCTG
ATCGTCACGGCGATTTATGCCGCCTCGGCGAGCACATGGAACGGGTTGGCAT
GGATTGTAGGCGCCGCCCTATACCTTGTCTGCCTCCCCGCGTTGCGTCGCGGT
GCATGGAGCCGGGCCACCTCGACCTGAatggaagccggcggcacctcgctaacggattcaccactcc
aagaattggagccaatcaattcttgcggagaactgtgaatgcgcaaaccaacccttggcagaacataggatcctcgggagTT
GTCAGCCTGTCCCGCTTATAAGATcatacgccgttatacgttgtttacgctttgaggaatcccatATGAC
GGTGAAATACACAGACGCGCAGATTCAGCGTTTACGCGAGTATGGAAATGGG
ACTTACGAACAGAAGGTCTTTGAGGATCTGGCGAGTCGTGATGCAGCGTTTT
CAAAAGAGATGTCCGTTGCGTCTACGGATAATGAGAAGAAAATCAAGGGTAT
GATCGCCAACCCTAGTCGTCATGGGCTGACCCAATTAATGAATGACATTGCG
GATGCCCTGGTCGCAGAGGGTTTTATCGAAGTGCGCACTCCAATCTTTATTTC
TAAGGATGCTTTGGCACGCATGACCATCACCGAGGACAAACCATTATTTAAA
CAGGTTTTTTGGATCGACGAAAAACGTGCCTTACGCCCAATGCTGGCTCCAA
ACCTGTATAGTGTCATGCGTGATCTTCGCGATCACACAGATGGCCCTGTTAAG
ATTTTTGAGATGGGATCTTGTTTTCGTAAGGAGTCACATTCTGGCATGCATCT
GGAAGAGTTTACGATGCTTAATCTGGTCGATATGGGGCCTCGCGGCGACGCG
ACTGAAGTATTGAAGAATTACATTAGCGTAGTTATGAAGGCAGCCGGGCTTC
CTGACTATGACTTGGTCCAAGAAGAGTCGGATGTATATAAGGAAACTATCGA
TGTTGAGATCAACGGACAAGAAGTTTGCAGCGCGGCGGTAGGTCCCCATTAC
CTGGATGCCGCTCACGATGTTCACGAGCCTTGGAGCGGCGCGGGTTTCGGCC
TTGAACGCCTTTTAACTATCCGCGAGAAATATTCCACCGTGAAAAAGGGCGG
CGCGTCAATCTCCTATCTGAATGGGGCCAAAATTAATAGCGGAGGGGGAGGG
TCTGGAGGGGGGGGTAGTGGAGGAGGCGGCTCCGGAGGCGGTGGTTCACAT
CACCATCACCATCACTCTGGATGGAGCCACCCGCAGTTCGAAAAATAActgcagt
ttcaaacgctaaattgcctgatgcgctacgcttatcaggcctacatgatctctgcaatatattgagtttgcgtgcttttgtaggccgga
taaggcgttcacgccgcatccggcaagaaacagcaaacaatccaaaacgccgcgttcggcggtcgacattaacgaagcgcta
accgtttttatcatgctctgggaggcagaataaatgatcatatcgtcaattattacctccacggggagagcctgagcaaactggcct
caggcatttgagaagcacacggtcacactgcttccggtagtcaataaaccggtaaaccagcaatagacataagcggctatttaac
gaccctgccctgaaccgacgaccgggtcgaatttgctttcgaatttctgccattcatccgcttattatcacttattcaggcgtagcaa
ccaggcgtttaagggcaccaataactgccttaaaaaaaTTACGCCCCGCCCTGCCACTCATCGCAGT
ACTGTTGTAATTCATTAAGCATTCTGCCGACATGGAAGCCATCACAAACGGC
ATGATGAACCTGAATCGCCAGCGGCATCAGCACCTTGTCGCCTTGCGTATAAT
ATTTGCCCATGGTGAAAACGGGGGCGAAGAAGTTGTCCATATTGGCCACGTT
TAAATCAAAACTGGTGAAACTCACCCAGGGATTGGCTGAGACGAAAAACATA
TTCTCAATAAACCCTTTAGGGAAATAGGCCAGGTTTTCACCGTAACACGCCAC
ATCTTGCGAATATATGTGTAGAAACTGCCGGAAATCCTAGTGGTATTCACTCC
AGAGCGATGAAAACGTTTCAGTTTGCTCATGGAAAACGGTGTAACAAGGGTG
AACACTATCCCATATCACCAGCTCACCGTCTTTCATTGCCATACGGAATTCCG

GATGAGCATTCATCAGGCGGGCAAGAATGTGAATAAAGGCCGGATAAAACTT
GTGCTTATTTTTCTTTACGGTCTTTAAAAAGGCCGTAATATCCAGCTGAACGG
TCTGGTTATAGGTACATTGAGCAACTGACTGAAATGCCTCAAAATGTTCTTTA
CGATGCCATTGGGATATATCAACGGTGGTATATCCAGTGATTTTTTTCTCCATt
ttagcttccttagctcctgaaaatctcgataactcaaaaaatacgcccggtagtgatcttatttcattatggtgaaagttggaacctctt
acgtgccgatcaacgtctcattttcgccaaaagttggcccagggcttcccggtatcaacagggacaccaggatttatttattctgcg
aagtgatcttccgtcacaggtatttattcggcgcaaagtgcgtcgggtgatgctgccaacttactgatttagtgtatgatggtgttttt
gaggtgctccagtggcttctgtttctatcgtctccgggagctgcatgtgtcagaggttttcaccgtcatcaccgaaacgcgcgagg
cagcagatcaattcgcgcgcgaaggcgaagcggcatgcataatgtgcctgtcaaatggacgaagcagggattctgcaaaccct
atgctactccgtcaagccgtcaattgtctgattcgttaccaaTTATGACAACTTGACGGCTACATCATTC
ACTTTTTCTTCACAACCGGCACGGAACTCGCTCGGGCTGGCCCCGGTGCATTT
TTTAAATACCCGCGAGAAATAGAGTTGATCGTCAAAACCAACATTGCGACCG
ACGGTGGCGATAGGCATCCGGGTGGTGCTCAAAAGCAGCTTCGCCTGGCTGA
TACGTTGGTCCTCGCGCCAGCTTAAGACGCTAATCCCTAACTGCTGGCGGAA
AAGATGTGACAGACGCGACGGCGACAAGCAAACATGCTGTGCGACGCTGGC
GATATCAAAATTGCTGTCTGCCAGGTGATCGCTGATGTACTGACAAGCCTCGC
GTACCCGATTATCCATCGGTGGATGGAGCGACTCGTTAATCGCTTCCATGCGC
CGCAGTAACAATTGCTCAAGCAGATTTATCGCCAGCAGCTCCGAATAGCGCC
CTTCCCCTTGCCCGGCGTTAATGATTTGCCCAAACAGGTCGCTGAAATGCGGC
TGGTGCGCTTCATCCGGGCGAAAGAACCCCGTATTGGCAAATATTGACGGCC
AGTTAAGCCATTCATGCCAGTAGGCGCGCGGACGAAAGTAAACCCACTGGTG
ATACCATTCGCGAGCCTCCGGATGACGACCGTAGTGATGAATCTCTCCTGGC
GGGAACAGCAAAATATCACCCGGTCGGCAAACAAATTCTCGTCCCTGATTTT
TCACCACCCCCTGACCGCGAATGGTGAGATTGAGAATATAACCTTTCATTCCC
AGCGGTCGGTCGATAAAAAAATCGAGATAACCGTTGGCCTCAATCGGCGTTA
AACCCGCCACCAGATGGGCATTAAACGAGTATCCCGGCAGCAGGGGATCATT
TTGCGCTTCAGCCATactttcatactcccgccattcagagaagaaaccaattgtccatattgcatcagacattgccgt
cactgcgtcttttactggctcttctcgctaaccaaaccggtaaccccgcttattaaaagcattctgtaacaaagcgggaccaaagcc
atgacaaaaacgcgtaacaaaagtgtctataatcacggcagaaaagtccacattgattatttgcacggcgtcacactttgctatgc
catagcatttttatccataagattagcggatcctacctgacgctttttatcgcaactctctactgtttctccatacccgtttttttgggctaa
caggaggaattaaccATGGTTAGCAAAGGTGAAGAACTGTTTACCGGCGTTGTGCCGA
TTCTGGTGGAACTGGATGGTGATGTGAATGGCCATAAATTTAGCGTTCGTGGC
GAAGGCGAAGGTGATGCGACCAACGGTAAACTGACCCTGAAATTTATTTGCA
CCACCGGTAAACTGCCGGTTCCGTGGCCGACCCTGGTGACCACCCTGACCTA
TGGCGTTCAGTGCTTTAGCCGCTATCCGGATCATATGAAACGCCATGATTTCT
TTAAAAGCGCGATGCCGGAAGGCTATGTGCAGGAACGTACCATTAGCTTCAA
AGATGATGGCACCTATAAAACCCGTGCGGAAGTTAAATTTGAAGGCGATACC
CTGGTGAACCGCATTGAACTGAAAGGTATTGATTTTAAAGAAGATGGCAACA
TTCTGGGTCATAAACTGGAATATAATTTCAACAGCCATTAGGTGTATATTACC
GCCGATAAACAGAAAAATGGCATCAAAGCGAACTTTAAAATCCGTCACAACG
TGGAAGATGGTAGCGTGCAGCTGGCGGATCATTATCAGCAGAATACCCCGAT
TGGTGATGGCCCGGTGCTGCTGCCGGATAATCATTATCTGAGCACCCAGAGC
GTTCTGAGCAAAGATCCGAATGAAAAACGTGATCATATGGTGCTGCTGGAAT
TTGTTACCGCCGCGGGCATTACCCACGGTATGGATGAACTGTATAAAGGCAG
CCACCATCATCATCACCATTAAagctcgagcgaagcttgggcccgaacaaaaactcatctcagaagagga
tctgaatagcgccgtcgaccatcatcatcatcattgagtttaaacggtctccagcttggctgttttggcggat

pKW1 TRIPLE

Sequence feature	Nucleotide position
glnS promoter	14 - 39
*Mm*PylRS-S(G_4S)$_4$linker-FLAG	81 - 1455
*1R26*PylRS(CbzK)	1585 - 2412
*Lum1*PylRS(NmH)	2423 - 3250
lpp promoter	3785 - 3809
Spe^{Pyl}tRNA$_{CUA}$	3830 - 3901
Int$^{\Delta NPyl}$tRNA(^{A}17,^{V}C10)$_{UCCU}$	3941 - 4013
Alv$^{\Delta NPyl}$tRNA(8)$_{UACU}$	4058 - 4130
rrnC terminator	4137 - 4165
SpR antibiotic resistance	4850 - 5711
pMB1 replication origin	5797 - 6385

ggatcctcgggagTTGTCAGCCTGTCCCGCTTATAAGATcatacgccgttatacgttgtttacgctttgag
gaatcccatATGATGGACAAAAAACCGCTGAACACCCTGATCTCTGCTACCGGTCT
GTGGATGTCTCGTACCGGTACCATCCACAAAATCAAACACCACGAAGTTTCT
CGTTCTAAAATCTACATCGAAATGGCTTGCGGTGACCACCTGGTTGTTAACAA
CTCTCGTTCTTCTCGTACCGCTCGTGCTCTGCGTCACCACAAATACCGTAAAA
CCTGCAAACGTTGCCGTGTTTCTGACGAAGACCTGAACAAATTCCTGACCAA
AGCTAACGAAGACCAGACCTCTGTTAAAGTTAAAGTTGTTTCTGCTCCGACCC
GTACCAAAAAAGCTATGCCGAAATCTGTTGCTCGTGCTCCGAAACCGCTGGA
AAACACCGAAGCTGCTCAGGCTCAGCCGTCTGGTTCTAAATTCTCTCCGGCTA
TCCCGGTTTCTACCCAGGAATCTGTTTCTGTTCCGGCTTCTGTTTCTACCTCTA
TCTCTTCTATCTCTACCGGTGCTACCGCTTCTGCTCTGGTTAAAGGTAACACC
AACCCGATCACCTCTATGTCTGCTCCGGTTCAGGCTTCTGCTCCGGCTCTGAC
CAAATCTCAGACCGACCGTCTGGAAGTTCTGCTGAACCCGAAAGACGAAATC
TCTCTGAACTCTGGTAAACCGTTCCGTGAACTGGAATCTGAACTGCTGTCTCG
TCGTAAAAAAGACTTACAACAGATCTACGCTGAAGAACGTGAAAACTACCTG
GGTAAACTGGAACGTGAAATCACGCGTTTCTTCGTTGACCGTGGTTTCCTGGA
AATCAAATCTCCGATCCTGATCCCGCTGGAATACATCGAACGTATGGGTATC
GACAACGACACCGAACTGTCTAAACAGATCTTCCGTGTTGACAAAAACTTCT
GCCTGCGTCCGATGCTGGCTCCGAACCTGTACAACTACCTGCGTAAACTGGA
CCGTGCTCTGCCGGACCCGATCAAAATCTTCGAAATCGGTCCGTGCTACCGTA
AAGAATCTGACGGTAAAGAACACCTGGAAGAATTCACCATGCTGAACTTCTG
CCAGATGGGTTCTGGTTGCACCCGTGAAAACCTGGAATCTATCATCACCGACT
TCCTGAACCACCTGGGTATCGACTTCAAAATCGTTGGTGACTCTTGCATGGTT
TACGGTGACACCCTGGACGTTATGCACGGTGACCTGGAACTGTCTTCTGCTGT
TGTTGGTCCGATCCCGCTGGACCGTGAATGGGGTATCGACAAACCGTGGATC
GGTGCTGGTTTCGGTCTGGAACGTCTGCTGAAAGTTAAACACGACTTCAAAA
ACATCAAACGTGCTGCTCGTTCTGAATCTTACTACAACGGTATCTCTACCAAC
CTGTCTGGTGGCGGTGGCTCTGGTGGCGGTGGCTCTGGTGGCGGTGGCTCTGG
TGGCGGTGGCTCTGACTACAAGGACGACGACGACAAGTAAcctattgaagggttgctaag
aataacaagatcaaataaggaggttttttATGGCAGAACATTTTACAGATGCACAGATTCAGCG

CTTACGCGAGTATGGTAATGGGACATATAAGGATATGGAGTTCGCAGACGTG
AGTGCGCGTGAAAAGGCGTTCACGAAGCTTATGTCTGATGCCAGTCGCGATA
ATGAATCAGCCCTTAAGGGAATGATTGCGCACCCTGCACGCCAGGGTTTATC
ACGCTTAATGAATGACATTGCAGACGCTTTGGTGGCTGATGGATTTATTGAAG
TTCGCACACCCATTATTATCAGTAAAGATGCTTTAGCCAAAATGACAATCACT
CCCGATAAACCATTATTCAAACAAGTATTTTGGATTGATGATAAACGCGCTCT
GCGTCCAATGTTAGCCCCGTCTCTGGGCACGGTCCTGCGTAGCCTGCGTGATC
ATACTGACGGCCCTGTCAAAATTTTCGAAATGGGCTCTTGTTTTCGCAAAGAA
TCGCACAGTGGGATGCACCTGGAAGAGTTTACCATGTTAAACCTGGTAGATA
TGGGCCCCGCCGGCGATGCCACTGAGTCCCTGAAGAAATATATTGGGATCGT
AATGAAGGCCGCCGGGCTGCCTGACTACCAATTAGTCCACGAGGAATCTGAT
GTGTATAAAGAAACGATTGATGTTGAAATCAACGGACAAGAGGTTTGTTCGG
CTGCTGTAGGTCCTCACTATTTGGATGCCGCCCATGACGTGCACGAGCCGTGG
GCAGGGGCAGGGTTCGGCCTGGAGCGCCTGCTGACAATTCGCCAGGGATATA
GCACAGTGATGAAAGGGGGAGCTTCCACAACCTATTTGAACGGGGCTAAGAT
GGACTAAggaggacatcATGGATACACGCCTGACGCCCGCGCAAGCTCAGCGCAT
TCGTGAAATGGGCGGGACGGTAGACCCGTCGTTAGCCTTTAGCAGCGAAGCC
GAACGTGAATCGGCATTTCAGCGTATTTCGGCAGATTTGCAGGGGGCGAACT
TGGCCAAGATTCGTCGCTGTGCTGAAGCGCCAGAGCGCCATCCAATCGGTAG
TCTGGAAAATACCCTGGCCTGCGCACTGGCGGCTAAGGGGTTTATTGAGGTT
AAGACGCCGATGATGATTCCTGCAGACGGACTGGTGAAAATGGGTATTGACG
AGTCTCATCCATTATGGAATCAAGTTTTTTGGGTCGGCCCGAAAAAAGCACTG
CGTCCGATGATGGCTCCAAACATTTTCTTTTTGGCGCGCCACTTACGTCGTTCT
GTGCCCGCCCCATTACTTCTTTTTGAAATTGGACCTTGTTTCCGCAAGGAGAG
CCGCGGGTCCAATCATCTGGAGGAGTTTACTATGTTGAACTTATTTGAGTTAG
CTCCACAAGCTGACGCTACAGAACGCTTAAAAGAACACATTGCCACAGTGAT
GAATGCTGTTGGACTGCCCTACGAGTTAGTAGTCGAGGGTTCTGAAGTTTATG
GCACTACGATTGATGTGGAAGTAGACGGTGTTGAGCTGGCATCTGGCGCAGT
AGGACCGTTACCTATGGATAAACCTCACGGGATTACCGAGCCCTGGGCTGGT
GTTGGGTTCGGATTGGAGCGTATTGCATTAATGCGCACGAAGGAGCAAAATA
TCAAAAAAGTAGGACGCTCCTTGGTTTATGTTAATGGAGCCCGCATTGATATT
TAActgcagtttcaaacgctaaattgcctgatgcgctacgcttatcaggcctacatgatctctgcaatatattgagtttgcgtgcttt
tgtaggccggataaggcgttcacgccgcatccggcaagaaacagcaaacaatccaaaacgccgcgttcggcggtcgacaca
gatgtaggtgttccacagggtagccagcagcatcctgcgatgcagatccggaacataatggtgcagggcgcttgtttcggcgtg
ggtatggtggcaggccccgtggccgggggactgttgggcgctgccggcacctgtcctacgagttgcatgataaagaagacagt
cataagtgcggcgacgatagtcatgccccgcgcccaccggaaggagctaccggcagcggtgcggactgttgtaactcagaat
aagaaatgaggccgctcatggcgttctgttgcccgtctcactggtgaaaagaaaaacaaccctggcgccgcttctttgagcgaac
gatcaaaaataagtggcgccccatcaaaaaaataTTCTCAACATAAAAAACTTTGTGTAatacttgtaac
gctagatctGGAAATCTGATCATGTAGATCGAATGGACTCTAAATCCGTTCAGCCG
GGTTAGATTCCCGGGGTTTCCGCCAaattttgcacccagcaaacttggtacgtaaacgcatcgtGGCG
AACTGGTCCGGGACCACCAGGCCTTCCTAAGCCACGGTTAGCCGGGTTCAAC
TCCCGGGTTCGTCGCCAactactttatgtagtctccgccgtgtagcaagaaattgagaagtGGGGGACGG
TCCGGCGACCAGCGGGTCTTACTAAACCTAGCCTTGCGGGGTTCGACACCCC
GGTCTCTCGCCAactagtATCCTTAGCGAAAGCTAAGGATTTTTTTTaagcttggcactgg
ccgtcgttttacaacgtcgtgactgggaaaaccctggcgttacccaacttaatcgccttgcagcacatccccctttcgccagacgc
tctcccttatgcgactcctgcattaggaagcagcccagtagtaggttgaggccgttgagcaccgccgccgcaaggaatggtgca

tgcaaggagcccgagatgcgccgcgtgcggctgctggagatggcggacgcgatggatatgttctggcggccgcataatgctta
agtcgaacagaaagtaatcgtattgtacacggccgcataatcgaaattaaattggatatcggcAggccacgcgatcgctgacgt
cggtaccctcgagtctggtaaagaaaccgctgctgcgaaatttgaacgccagcacatggactcgtctactagcgcagcttaatta
acctaggctgctgccaccgctgagcaataactagcataacccccttggggcctctaaacgggtcttgaggggttttttgctgaaacc
tcaggcatttgagaagcacacggtcacactgcttccggtagtcaataaaccggtaaaccagcaatagacataagcggctatttaa
cgaccctgccctgaaccgacgaccgggtcatcgtggccggatcttgcggcccctcggcttgaacgaattgttagacaTTAT
TTGCCGACTACCTTGGTGATCTCGCCTTTCACGTAGTGGACAAATTCTTCCAA
CTGATCTGCGCGCGAGGCCAAGCGATCTTCTTCTTGTCCAAGATAAGCCTGTC
TAGCTTCAAGTATGACGGGCTGATACTGGGCCGGCAGGCGCTCCATTGCCCA
GTCGGCAGCGACATCCTTCGGCGCGATTTTGCCGGTTACTGCGCTGTACCAAA
TGCGGGACAACGTAAGCACTACATTTCGCTCATCGCCAGCCCAGTCGGGCGG
CGAGTTCCATAGCGTTAAGGTTTCATTTAGCGCCTCAAATAGATCCTGTTCAG
GAACCGGATCAAAGAGTTCCTCCGCCGCTGGACCTACCAAGGCAACGCTATG
TTCTCTTGCTTTTGTCAGCAAGATAGCCAGATCAATGTCGATCGTGGCTGGCT
CGAAGATACCTGCAAGAATGTCATTGCGCTGCCATTCTCCAAATTGCAGTTCG
CGCTTAGCTGGATAACGCCACGGAATGATGTCGTCGTGCACAACAATGGTGA
CTTCTACAGCGCGGAGAATCTCGCTCTCTCCAGGGGAAGCCGAAGTTTCCAA
AAGGTCGTTGATCAAAGCTCGCCGCGTTGTTTCATCAAGCCTTACGGTCACCG
TAACCAGCAAATCAATATCACTGTGTGGCTTCAGGCCGCCATCCACTGCGGA
GCCGTACAAATGTACGGCCAGCAACGTCGGTTCGAGATGGCGCTCGATGACG
CCAACTACCTCTGATAGTTGAGTCGATACTTCGGCGATCACCGCTTCCCTCAT
ACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAG
CGGATACATATTTGAAtgtatttagaaaaataaacaaatagctagcgagcaaaaggccagcaaaaggccagga
accgtaaaaaggccgcgttgctggcgttTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACA
AAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGAT
ACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTG
CCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTC
TCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGC
TGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGT
AACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAG
CAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGA
GTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGT
ATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTG
ATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGC
AGATTACGCGCAGAAAAAAAGGATCTCAAgaagatcctttgatcttttctacggggtctgacgctcag
tggaacgaaaactcacgttaagggattttgg

pCUN_O-Strep-sfGFP-40TAG-136AGGA-150AGTA

Sequence feature	**Nucleotide position**
TetR antibiotic resistance	419 - 1603
p15A replication origin	1778 - 2604
Promoter	2957 - 2986
O-RBS	3007 - 3014
O-Strep-	3021 - 3796

sfGFP(40TAG,136AGGA,150AGTA)

ggtaccggttccactctctgttgcgggcaacttcagcagcacgtaggggacttccgcgtttccagactttacgaaacacggaaac
cgaagaccattcatgttgttgctcaggtcgcagacgttttgcagcagcagtcgcttcacgttcgctcgcgtatcggtgattcattctg
ctaaccagtaaggcaaccccgccagcctagccgggtcctcaacgacaggagcacgatcatgcgcacccgtggccaggaccc
aacgctgcccgagatgcgccgcgtgcggctgctggagatggcggacgcgatggatatgttctgccaagggttggtttgcgcatt
cacagttctccgcaagaattgattggctccaattcttggagtggtgaatccgttagcgaggtgccgccggcttccattcaGGTC
GAGGTGGCCCGGCTCCATGCACCGCGACGCAACGCGGGGAGGCAGACAAGG
TATAGGGCGGCGCCTACAATCCATGCCAACCCGTTCCATGTGCTCGCCGAGG
CGGCATAAATCGCCGTGACGATCAGCGGTCCAGTGATCGAAGTTAGGCTGGT
AAGAGCCGCGAGCGATCCTTGAAGCTGTCCCTGATGGTCGTCATCTACCTGC
CTGGACAGCATGGCCTGCAACGCGGGCATCCCGATGCCGCCGGAAGCGAGA
AGAATCATAATGGGGAAGGCCATCCAGCCTCGCGTCGCGAACGCCAGCAAG
ACGTAGCCCAGCGCGTCGGCCGCCATGCCGGCGATAATGGCCTGCTTCTCGC
CGAAACGTTTGGTGGCGGGACCAGTGACGAAGGCTTGAGCGAGGGCGTGCA
AGATTCCGAATACCGCAAGCGACAGGCCGATCATCGTCGCGCTCCAGCGAAA
GCGGTCCTCGCCGAAAATGACCCAGAGCGCTGCCGGCACCTGTCCTACGAGT
TGCATGATAAAGAAGACAGTCATAAGTGCGGCGACGATAGTCATGCCCCGCG
CCCACCGGAAGGAGCTGACTGGGTTGAAGGCTCTCAAGGGCATCGGTCGACG
CTCTCCCTTATGCGACTCCTGCATTAGGAAGCAGCCCAGTAGTAGGTTGAGGC
CGTTGAGCACCGCCGCCGCAAGGAATGGTGCATGCAAGGAGATGGCGCCCA
ACAGTCCCCCGGCCACGGGGCCTGCCACCATACCCACGCCGAAACAAGCGCT
CATGAGCCCGAAGTGGCGAGCCCGATCTTCCCCATCGGTGATGTCGGCGATA
TAGGCGCCAGCAACCGCACCTGTGGCGCCGGTGATGCCGGCCACGATGCGTC
CGGCGTAGAGGATCCACAGGACGGGTGTGGTCGCCATGATCGCGTAGTCGAT
AGTGGCTCCAAGTAGCGAAGCGAGCAGGACTGGGCGGCGGCCAAAGCGGTC
GGACAGTGCTCCGAGAACGGGTGCGCATAGAAATTGCATCAACGCATATAGC
GCTAGCAGCACGCCATAGTGACTGGCGATGCTGTCGGAATGGACGATATCCC
GCAAGAGGCCCGGCAGTACCGGCATAACCAAGCCTATGCCTACAGCATCCAG
GGTGACGGTGCCGAGGATGACGATGAGCGCATTGTTAGATTTcatacacggtgcctga
ctgcgttagcaatttaactgtgataaactaccgcattaaagcttatcgatgataagctgtcaaacatgagaattacaacttatatcgta
tggggctgacttcaggtgctacatttgaagagataaattgcactgaaatctagaaatattttatctgatTAATAAGATGAT
CTTCTTGAGATCGTTTTGGTCTGCGCGTAATCTCTTGCTCTGAAAACGAAAAA
ACCGCCTTGCAGGGCGGTTTTTCGAAGGTTCTCTGAGCTACCAACTCTTTGAA
CCGAGGTAACTGGCTTGGAGGAGCGCAGTCACCAAAACTTGTCCTTTCAGTTT
AGCCTTAACCGGCGCATGACTTCAAGACTAACTCCTCTAAATCAATTACCAGT
GGCTGCTGCCAGTGGTGCTTTTGCATGTCTTTCCGGGTTGGACTCAAGACGAT
AGTTACCGGATAAGGCGCAGCGGTCGGACTGAACGGGGGGTTCGTGCATACA
GTCCAGCTTGGAGCGAACTGCCTACCCGGAACTGAGTGTCAGGCGTGGAATG
AGACAAACGCGGCCATAACAGCGGAATGACACCGGTAAACCGAAAGGCAGG
AACAGGAGAGCGCACGAGGGAGCCGCCAGGGGGAAACGCCTGGTATCTTTA
TAGTCCTGTCGGGTTTCGCCACCACTGATTTGAGCGTCAGATTTCGTGATGCT
TGTCAGGGGGGCGGAGCCTATGGAAAAACGGCTTTGCCGCGGCCCTCTCACT
TCCCTGTTAAGTATCTTCCTGGCATCTTCCAGGAAATCTCCGCCCCGTTCGTA
AGCCATTTCCGCTCGCCGCAGTCGAACGACCGAGCGTAGCGAGTCAGTGAGC
GAGGAAGCGGAATATATCCTGTATCACATATTCTGCTGACGCACCGGTGCAG

CCTTTTTTCTCCTGCCACATGAAGCACTTCACTGACACCCTCATCAGTGCCAA
CATAGTAAGCCAGTATACACTCCGCTAGCgctgatgtccggcggtgcttttgccgttacgcaccacc
ccgtcagtagctgaacaggagggacagctgatagaaacagaagccactggagcacctcaaaaacaccatcatacactaaatca
gtaagttggcagcatcacccgacgcactttgcgccgaataaatacctgtgacggaagatcacttcgcagaataaataaatcctgg
tgtccctgttgataccgggaagccctgggccaacttttggcgaaaatgagacgttgatcggcacgtagggccctctagatgcatg
ctcgagcggcggccagtgtgatggatatctgcagaattcgcccttcaaaaatcgatgagctGTTGACAATTAATCA
TCGAACTAGTTTAATgtgtggaagcggccgctttcATATCCCTccgcaaATGTGGAGCCACC
CGCAGTTCGAAAAAtctggaGTTAGCAAAGGTGAAGAACTGTTTACCGGCGTTG
TGCCGATTCTGGTGGAACTGGATGGTGATGTGAATGGCCATAAATTTAGCGTT
CGTGGCGAAGGCGAAGGTGATGCGACCTAGGGTAAACTGACCCTGAAATTTA
TTTGCACCACCGGTAAACTGCCGGTTCCGTGGCCGACCCTGGTGACCACCCTG
ACCTATGGCGTTCAGTGCTTTAGCCGCTATCCGGATCATATGAAACGCCATGA
TTTCTTTAAAAGCGCGATGCCGGAAGGCTATGTGCAGGAACGTACCATTAGC
TTCAAAGATGATGGCACCTATAAAACCCGTGCGGAAGTTAAATTTGAAGGCG
ATACCCTGGTGAACCGCATTGAACTGAAAGGTATTGATTTTAAAGAAGATGG
CAGGAATTCTGGGTCATAAACTGGAATATAATTTCAACAGCCATAGTAGTGT
ATATTACCGCCGATAAACAGAAAAATGGCATCAAAGCGAACTTTAAAATCCG
TCACAACGTGGAAGATGGTAGCGTGCAGCTGGCGGATCATTATCAGCAGAAT
ACCCCGATTGGTGATGGCCCGGTGCTGCTGCCGGATAATCATTATCTGAGCAC
CCAGAGCGTTCTGAGCAAAGATCCGAATGAAAAACGTGATCATATGGTGCTG
CTGGAATTTGTTACCGCCGCGGGCATTACCCACGGTATGGATGAACTGTATA
AAGGCAGCCACCATCATCATCACCATTAAtaatgataaagaattcgaagcttgggcccgaacaaaaa
ctcatctcagaagaggatctgaatagcgccgtcgaccatcatcatcatcattgagtttaaacgacgtccagcttggctgttttgg
cggatgagagaagattttcagcctgatacagattaaatcagaacgcagaagcggtctgataaaacagaatttgcctggcggcag
tagcgcggtggtcccacctgaccccatgccgaactcagaagtgaaacgccgtagcgccgatggtagtgtgaggcctcccatgc
gagagtagggaactgccaggcatcaaataaaacgaaaggctcagtcgaaagactgggcctttcgttttatctgttgtttgtcggtg
aacgatatctgcttttcttcgcgaatgagctc

pCUN_O-GST-CAM(1TAG)

Sequence feature	Nucleotide position
TetR antibiotic resistance	419 - 1603
p15A replication origin	1778 - 2604
lac promoter	2969 - 2999
O-RBS	3033 - 3039
O-GST-CAM(1TAG)-Strep	3046 - 4212
rrnB T1 terminator	4433 - 4479

ggtaccggttccactctctgttgcgggcaacttcagcagcacgtaggggacttccgcgtttccagactttacgaaacacggaaac
cgaagaccattcatgttgttgctcaggtcgcagacgttttgcagcagcagtcgcttcacgttcgctcgcgtatcggtgattcattctg
ctaaccagtaaggcaaccccgccagcctagccgggtcctcaacgacaggagcacgatcatgcgcacccgtggccaggaccc
aacgctgcccgagatgcgccgcgtgcggctgctggagatggcggacgcgatggatatgttctgccaagggttggtttgcgcatt
cacagttctccgcaagaattgattggctctaattcttggagtggtgaatccgttagcgaggtgccgccggcttccattcaGGTC
GAGGTGGCCCGGCTCCATGCACCGCGACGCAACGCGGGGAGGCAGACAAGG
TATAGGGCGGCGCCTACAATCCATGCCAACCCGTTCCATGTGCTCGCCGAGG

CGGCATAAATCGCCGTGACGATCAGCGGTCCAGTGATCGAAGTTAGGCTGGT
AAGAGCCGCGAGCGATCCTTGAAGCTGTCCCTGATGGTCGTCATCTACCTGC
CTGGACAGCATGGCCTGCAACGCGGGCATCCCGATGCCGCCGGAAGCGAGA
AGAATCATAATGGGGAAGGCCATCCAGCCTCGCGTCGCGAACGCCAGCAAG
ACGTAGCCCAGCGCGTCGGCCGCCATGCCGGCGATAATGGCCTGCTTCTCGC
CGAAACGTTTGGTGGCGGGACCAGTGACGAAGGCTTGAGCGAGGGCGTGCA
AGATTCCGAATACCGCAAGCGACAGGCCGATCATCGTCGCGCTCCAGCGAAA
GCGGTCCTCGCCGAAAATGACCCAGAGCGCTGCCGGCACCTGTCCTACGAGT
TGCATGATAAAGAAGACAGTCATAAGTGCGGCGACGATAGTCATGCCCCGCG
CCCACCGGAAGGAGCTGACTGGGTTGAAGGCTCTCAAGGGCATCGGTCGACG
CTCTCCCTTATGCGACTCCTGCATTAGGAAGCAGCCCAGTAGTAGGTTGAGGC
CGTTGAGCACCGCCGCCGCAAGGAATGGTGCATGCAAGGAGATGGCGCCCA
ACAGTCCCCCGGCCACGGGGCCTGCCACCATACCCACGCCGAAACAAGCGCT
CATGAGCCCGAAGTGGCGAGCCCGATCTTCCCCATCGGTGATGTCGGCGATA
TAGGCGCCAGCAACCGCACCTGTGGCGCCGGTGATGCCGGCCACGATGCGTC
CGGCGTAGAGGATCCACAGGACGGGTGTGGTCGCCATGATCGCGTAGTCGAT
AGTGGCTCCAAGTAGCGAAGCGAGCAGGACTGGGCGGCGGCCAAAGCGGTC
GGACAGTGCTCCGAGAACGGGTGCGCATAGAAATTGCATCAACGCATATAGC
GCTAGCAGCACGCCATAGTGACTGGCGATGCTGTCGGAATGGACGATATCCC
GCAAGAGGCCCGGCAGTACCGGCATAACCAAGCCTATGCCTACAGCATCCAG
GGTGACGGTGCCGAGGATGACGATGAGCGCATTGTTAGATTTcatacacggtgcctga
ctgcgttagcaatttaactgtgataaactaccgcattaaagcttatcgatgataagctgtcaaacatgagaattacaacttatatcgta
tggggctgacttcaggtgctacatttgaagagataaattgcactgaaatctagaaatattttatctgatTAATAAGATGAT
CTTCTTGAGATCGTTTTGGTCTGCGCGTAATCTCTTGCTCTGAAAACGAAAAA
ACCGCCTTGCAGGGCGGTTTTTCGAAGGTTCTCTGAGCTACCAACTCTTTGAA
CCGAGGTAACTGGCTTGGAGGAGCGCAGTCACCAAAACTTGTCCTTTCAGTTT
AGCCTTAACCGGCGCATGACTTCAAGACTAACTCCTCTAAATCAATTACCAGT
GGCTGCTGCCAGTGGTGCTTTTGCATGTCTTTCCGGGTTGGACTCAAGACGAT
AGTTACCGGATAAGGCGCAGCGGTCGGACTGAACGGGGGGTTCGTGCATACA
GTCCAGCTTGGAGCGAACTGCCTACCCGGAACTGAGTGTCAGGCGTGGAATG
AGACAAACGCGGCCATAACAGCGGAATGACACCGGTAAACCGAAAGGCAGG
AACAGGAGAGCGCACGAGGGAGCCGCCAGGGGGAAACGCCTGGTATCTTTA
TAGTCCTGTCGGGTTTCGCCACCACTGATTTGAGCGTCAGATTTCGTGATGCT
TGTCAGGGGGGCGGAGCCTATGGAAAAACGGCTTTGCCGCGGCCCTCTCACT
TCCCTGTTAAGTATCTTCCTGGCATCTTCCAGGAAATCTCCGCCCCGTTCGTA
AGCCATTTCCGCTCGCCGCAGTCGAACGACCGAGCGTAGCGAGTCAGTGAGC
GAGGAAGCGGAATATATCCTGTATCACATATTCTGCTGACGCACCGGTGCAG
CCTTTTTTCTCCTGCCACATGAAGCACTTCACTGACACCCTCATCAGTGCCAA
CATAGTAAGCCAGTATACACTCCGCTAGCgctgatgtccggcggtgcttttgccgttacgcaccacc
ccgtcagtagctgaacaggagggacagctgatagaaacagaagccactggagcacctcaaaaacaccatcatacactaaatca
gtaagttggcagcatcacccgacgcactttgcgccgaataaatacctgtgacggaagatcacttcggtgacttaactcgtctggg
atccagacctggtaccagcgcccaatacgcaaaccgcctctccccgcgcgttggccgattcattaatgcagctggcacgacag
gtttcccgactggaaagcgggcagtgagcgcaacgcaattaatgtgagttagctcactcattaggcaccccaggcTTTACA
CTTTATGCTTCCGGCTCGTATGTTGtgtggaattgtgagcggctcgagacaattttcaTATCCCTccg
caaATGTCCCCTATACTAGGTTATTGGAAAATTAAGGGCCTTGTGCAACCCACT
CGACTTCTTTTGGAATATCTTGAAGAAAAATATGAAGAGCATTTGTATGAGC

GCGATGAAGGTGATAAATGGCGAAACAAAAAGTTTGAATTGGGTTTGGAGTT
TCCCAATCTTCCTTATTATATTGATGGTGATGTTAAATTAACACAGTCTATGG
CCATCATACGTTATATAGCTGACAAGCACAACATGTTGGGTGGTTGTCCAAA
GGAGCGTGCAGAGATTTCAATGCTTGAAGGAGCGGTTTTGGATATTAGATAC
GGTGTTTCGAGAATTGCATATAGTAAAGACTTTGAAACTCTCAAAGTTGATTT
TCTTAGCAAGCTACCTGAAATGCTGAAAATGTTCGAAGATCGTTTATGTCATA
AAACATATTTAAATGGTGATCATGTAACCCATCCTGACTTCATGTTGTATGAC
GCTCTTGATGTTGTTTTATACATGGACCCAATGTGCCTGGATGCGTTCCCAAA
ATTAGTTTGTTTTAAAAAACGTATTGAAGCTATCCCACAAATTGATAAGTACT
TGAAATCCAGCAAGTATATAGCATGGCCTTTGCAGGGCTGGCAAGCCACGTT
TGGTGGTGGCGACCATCCTCCAAAAGGCTCGGGCGAAAACCTGTACTTCCAA
GGCGGCTCGGGCTAGGCTGACCAACTGACAGAAGAGCAGATTGCAGAGTTCA
AAGAAGCCTTCTCATTATTCGACAAGGATGGGGACGGCACCATCACCACAAA
GGAACTTGGCACCGTTATGAGGTCGCTTGGACAAAACCCAACGGAAGCAGAA
TTGCAGGATATGATCAATGAAGTCGATGCTGATGGCAATGGAACGATTGACT
TTCCTGAATTTCTTACTATGATGGCTAGAAAAATGAAGGACACAGACAGCGA
AGAGGAAATCCGAGAAGCATTCCGTGTTTTTGACAAGGATGGGAACGGCTAC
ATCAGCGCTGCTGAATTACGTCACGTCATGACAAACCTCGGGGAGAAGTTAA
CAGACGAAGAAGTTGACGAAATGATAAGGGAAGCAGATATTGATGGTGACG
GCCAAGTAAACTACGAAGAGTTTGTACAAATGATGACAGCAAAATGGAGCC
ACCCGCAGTTCGAAAAATAAgcttaattagctgacctactagtcggccggcggatgagagaagattttcagc
ctgatacagattaaatcagaacgcagaagcggtctgataaaacagaatttgcctggcggcagtagcgcggtggtcccacctgac
cccatgccgaactcagaagtgaaacgccgtagcgccgatggtagtgtgaggcctcccatgcgagagtagggaactgccaggc
atCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATctgttgttt
gtcggtgaacgatatctgcttttcttcgcgaatgagctc

pRSF_D11-riboX-G9

Sequence feature	**Nucleotide position**
Tac promoter	193 - 220
16S rRNA	504 - 2071
23S rRNA	2486 - 5389
KanR antibiotic resistance	6769 - 7584
pRSF replication origin	7692 - 8443

gtttgacagcttatcatcgactgcacggtgcaccaatgcttctggcgtcaggcagccatcggaagctgtggtatggctgtgcagg
tcgtaaatcactgcataattcgtgtcgctcaaggcgcactcccgttctggataatgttttttgcgccgacatcataacggttctggca
aatattctgaaatgagctgTTGACAATTAATCATCCGGCTCGTATAAtgtgtggaattgtgagcggata
acaatttcacacaggaaacagcgccgctgagaaaaagcgaagcggcactgctctttaacaatttatcagacaatctgtgtgggca
ctcgaggccgctgagaaaaagcgaagcggcactgctctttaacaatttatcagacaatctgtgtgggcactcgaagatacggatt
cttaacgtcgcaagacgaaaaatgaataccaagtctcaagagtgaacacgtaattcattacgaagtttaattctttgagcgtcaaac
ttttAAATTGAAGAGTTTGATCATGGCTCAGATTGAACGCTGGCGGCAGGCCTA
ACACATGCAAGTCGAACGGTAACAGGAAGAAGCTTGCTTCTTTGCTGACGAG
TGGCGGACGGGTGAGTAATGTCTGGGAAACTGCCTGATGGAGGGGGATAACT
ACTGGAAACGGTAGCTAATACCGCATAACGTCGCAAGACCAAAGAGGGGGA

CCTTCGGGCCTCTTGCCATCGGATGTGCCCAGATGGGATTAGCTAGTAGGTGG
GGTAACGGCTCACCTAGGCGACGATCCCTAGCTGGTCTGAGAGGATGACCAG
CCACACTGGAACTGAGACACGGTCCAGACTCCTACGGGAGGCAGCAGTGGG
GAATATTGCACAATGGGCGCAAGCCTGATGCAGCCATGCCGCGTGTATGAAG
AAGGCCTTCGGGTTGTAAAGTACTTTCAGCGGGGAGGAAGGGAGTAAAGTTA
ATACCTTTGCTCATTGACGTTACCCGCAGAAGAAGCACCGGCTAACTCCGTG
CCAGCAGCCGCGGTAATACGGAGGGTGCAAGCGTTAATCGGAATTACTGGG
CGTAAAGCGCACGCAGGCGGTTTGTTAAGTCAGATGTGAAATCCCCGGGCTC
AACCTGGGAACTGCATCTGATACTGGCAAGCTTGAGTCTCGTAGAGGGGGGT
AGAATTCCAGGTGTAGCGGTGAAATGCGTAGAGATCTGGAGGAATACCGGTG
GCGAAGGCGGCCCCCTGGACGAAGACTGACGCTCAGGTGCGAAAGCGTGGG
GAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGTCGACT
TGGAGGTTGTGCCCTTGAGGCGTGGCTTCCGGAGCTAACGCGTTAAGTCGAC
CGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAATGAATTGACGGGGGCC
CGCACAAGCGGTGGAGCATGTGGTTTAATTCGATGCAACGCGAAGAACCTTA
CCTGGTCTTGACATCCACGGAAGTTTTCAGAGATGAGAATGTGCCTTCGGGA
ACCGTGAGACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTTGTGAAATGTTG
GGTTAAGTCCCGCAACGAGCGCAACCCTTATCCTTTGTTGCCAGCGGTCCGGC
CGGGAACTCAAAGGAGACTGCCAGTGATAAACTGGAGGAAGGTGGGGATGA
CGTCGGGTCATCATGGCCCTTACGACCAGGGCTACACACGTGCTACAATGGC
GCATACAAAGAGAAGCGACCTCGCGAGAGCAAGCGGACCTCATAAAGTGCG
TCGTAGTCCGGATTGGAGTCTGCAACTCGACTCCATGAAGTCGGAATCGCTA
GTAATCGTGGATCAGAATGCCACGGTGAATACGTTCCCGGGCCTTGTACACA
CCGCCCGTCACACCATGGGAGTGGGTTGCAAAAGAAGTAGGTAGCTTAACCT
TCGGGAGGGCGCTTACCACTTTGTGATTCATGACTGGGGTGAAGTCGTAACA
AGGTAACCGTAGGGGAACCTGCGGTTGGATCAAGTGGTTACCTTAAAGAAGC
GTACTTTGTAGTGCtcacacagattgtctgatagaaagtgaaaagcaaggcgtttacgcgttgggagtgaggctga
agagaataaggccgttcgctttctattaatgaaagctcaccctacacgaaaatatcacgcaacgcgtgataagcaattttcgtgtcc
ccttcgtctagaggcccaggacaccgccctttcacggcggtaacaggggttcgaatcccctaggggacgccacttgctggtttgt
gagtgaaagtcgccgaccttaatatctcaaaactcatcttcgggtgatgtttgagatttttgctctttaaaaatctggatcaagctgaa
aattgaaacactgaacaacgagagttgttcgtgagtctctcaaattttcgcaacacgatgatgaatcgaaagaaacatcttcgggtt
gtgaGGTTAAGCGACTAAGCGTACACGGTGGATGCCCTGGCAGTCAGAGGCGA
TGAAGGACGTGCTAATCTGCGATAAGCGTCGGTAAGGTGATATGAACCGTTA
TAACCGGCGATTTCCGAATGGGGAAACCCAGTGTGTTTCGACACACTATCATT
AACTGAATCCATAGGTTAATGAGGCGAACCGGGGGAACTGAAACATCTAAGT
ACCCCGAGGAAAAGAAATCAACCGAGATTCCCCCAGTAGCGGCGAGCGAAC
GGGGAGCAGCCCAGAGCCTGAATCAGTGTGTGTGTTAGTGGAAGCGTCTGGA
AAGGCGCGCGATACAGGGTGACAGCCCCGTACACAAAAATGCACATGCTGT
GAGCTCGATGAGTAGGGCGGGACACGTGGTATCCTGTCTGAATATGGGGGGA
CCATCCTCCAAGGCTAAATACTCCTGACTGACCGATAGTGAACCAGTACCGT
GAGGGAAAGGCGAAAAGAACCCCGGCGAGGGGAGTGAAAAAGAACCTGAA
ACCGTGTACGTACAAGCAGTGGGAGCACGCTTAGGCGTGTGACTGCGTACCT
TTTGTATAATGGGTCAGCGACTTATATTCTGTAGCAAGGTTAACCGAATAGGG
GAGCCGAAGGGAAACCGAGTCTTAACTGGGCGTTAAGTTGCAGGGTATAGAC
CCGAAACCCGGTGATCTAGCCATGGGCAGGTTGAAGGTTGGGTAACACTAAC
TGGAGGACCGAACCGACTAATGTTGAAAAATTAGCGGATGACTTGTGGCTGG

GGGTGAAAGGCCAATCAAACCGGGAGATAGCTGGTTCTCCCCGAAAGCTATT
TAGGTAGCGCCTCGTGAATTCATCTCCGGGGGGTAGAGCACTGTTTCGGCAAG
GGGGTCATCCCGACTTACCAACCCGATGCAAACTGCGAATACCGGAGAATGT
TATCACGGGAGACACACGGCGGGTGCTAACGTCCGTCGTGAAGAGGGAAAC
AACCCAGACCGCCAGCTAAGGTCCCAAAGTCATGGTTAAGTGGGAAACGATG
TGGGAAGGCCCAGACAGCCAGGATGTTGGCTTAGAAGCAGCCATCATTTAAA
GAAAGCGTAATAGCTCACTGGTCGAGTCGGCCTGCGCGGAAGATGTAACGGG
GCTAAACCATGCACCGAAGCTGCGGCAGCGACGCTTATGCGTTGTTGGGTAG
GGGAGCGTTCTGTAAGCCTGCGAAGGTGTGCTGTGAGGCATGCTGGAGGTAT
CAGAAGTGCGAATGCTGACATAAGTAACGATAAAGCGGGTGAAAAGCCCGC
TCGCCGGAAGACCAAGGGTTCCTGTCCAACGTTAATCGGGGCAGGGTGAGTC
GACCCCTAAGGCGAGGCCGAAAGGCGTAGTCGATGGGAAACAGGTTAATATT
CCTGTACTTGGTGTTACTGCGAAGGGGGGACGGAGAAGGCTATGTTGGCCGG
GCGACGGTTGTCCCGGTTTAAGCGTGTAGGCTGGTTTTCCAGGCAAATCCGG
AAAATCAAGGCTGAGGCGTGATGACGAGGCACTACGGTGCTGAAGCAACAA
ATGCCCTGCTTCCAGGAAAAGCCTCTAAGCATCAGGTAACATCAAATCGTAC
CCCAAACCGACACAGGTGGTCAGGTAGAGAATACCAAGGCGCTTGAGAGAA
CTCGGGTGAAGGAACTAGGCAAAATGGTGCCGTAACTTCGGGAGAAGGCAC
GCTGATATGTAGGTGAGGTCCCTCGCGGATGGAGCTGAAATCAGTCGAAGAT
ACCAGCTGGCTGCAACTGTTTATTAAAAACACAGCACTGTGCAAACACGAAA
GTGGACGTATACGGTGTGACGCCTGCCCGGTGCCGGAAGGTTAATTGATGGG
GTTAGCGCAAGCGAAGCTCTTGATCGAAGCCCCGGTAAACGGCGGCCGTAAC
TATAACGGTCCTAAGGTAGCGAAATTCCTTGTCGGGTAAGTTCCGACCTGCAC
GAATGGCGTAATGATGGCCAGGCTGTCTCCACCCGAGACTCAGTGAAATTGA
ACTCGCTGTGAAGATGCAGTGTACCCGCGGCAAGACGGAAAGACCCCGTGA
ACCTTTACTATAGCTTGACACTGAACATTGAGCCTTGATGTGTAGGATAGGTG
GGAGGCTTTGAAGTGTGGACGCCAGTCTGCATGGAGCCGACCTTGAAATACC
ACCCTTTAATGTTTGATGTTCTAACGTTGACCCGTAATCCGGGTTGCGGACAG
TGTCTGGTGGGTAGTTTGACTGGGGCGGTCTCCTCCTAAAGAGTAACGGAGG
AGCACGAAGGTTGGCTAATCCTGGTCGGACATCAGGAGGTTAGTGCAATGGC
ATAAGCCAGCTTGACTGCGAGCGTGACGGCGCGAGCAGGTGCGAAAGCAGG
TCATAGTGATCCGGTGGTTCTGAATGGAAGGGCCATCGCTCAACGGATAAAA
GGTACTCCGGGGATAACAGGCTGATACCGCCCAAGAGTTCATATCGACGGCG
GTGTTTGGCACCTCGATGTCGGCTCATCACATCCTGGGGCTGAAGTAGGTCCC
AAGGGTATGGCTGTTCGCCATTTAAAGTGGTACGCGAGCTGGGTTTAGAACG
TCGTGAGACAGTTCGGTCCCTATCTGCCGTGGGCGCTGGAGAACTGAGGGGG
GCTGCTCCTAGTACGAGAGGACCGGAGTGGACGCATCACTGGTGTTCGGGTT
GTCATGCCAATGGCACTGCCCGGTAGCTAAATGCGGAAGAGATAAGTGCTGA
AAGCATCTAAGCACGAAACTTGCCCCGAGATGAGTTCTCCCTGACCCTTTAA
GGGTCCTGAAGGAACGTTGAAGACGACGACGTTGATAGGCCGGGTGTGTAAG
CGCAGCGATGCGTTGAGCTAACCGGTACTAATGAACCGTGAGGCTTAACCTTa
caacgccgaagctgttttggcggatgagagaagattttcagcctgatacagattaaatcagaacgcagaagcggtctgataaaac
agaatttgcctggcggcagtagcgcggtggtcccacctgaccccatgccgaactcagaagtgaaacgccgtagcgccgatggt
agtgtggggtctccccatgcgagagtagggaactgccaggcatcaaataaaacgaaaggctcagtcgaaagactgggcctttc
gttttatctgttgtttgtcggtgaacgctctcctgagtaggacaaatccgccgggagcggatttgaacgttgcgaagcaacggccc
ggagggtggcgggcaggacgcccgccataaactgccaggcatcaaattaagcagaaggccatcctgacggatggccttttg

cgtttctacaaactcttcctgtcgtcatatctacaagccatccccccacagatacggtaaactagcctcgtttttgcatcaggaaagc
agctatgaaccactccttaaaaccctggaacacatttggcattgatcataatgctcagcacattgtatgtgccgaagacgaacaac
aattactcaatgcctggcagtatgcaaccgcagaaggacaacccgttcttattctgggtgaaggaagtaatgtacttttttctggagg
actatcgcggcacggtgatcatcaaccggatcaaaggtatcgaaattcatgatgaacctgatgcgtggtatttacatgtaggagcc
ggagaaaactggcatcgtctggtaaaatacactttgcaggaaggtatgcctggtctggaaaatctggcattaattcctggttgtgtc
ggctcatcacctatccagaatattggtgcttatggcgtagaattacagcgagtttgcgcttatgttgattctgttgaactggcgacag
gcaagcaagtgcgcttaactgccaaagagtgccgttttggctatcgcgacagtatttttaaacatgaataccaggaccgcttcgct
attgtagccgtaggtctgcgtctgccaaaagagtggcaacctgtactaacgtatggtgacttaactcgtctgggatccggtaccga
attcgggcctctaaacgggtcttgaggggttttttgctgaaacctcaggcatttgagaagcacacggtcacactgcttccggtagtc
aataaaccggtaaaccagcaatagacataagcggctatttaacgaccctgccctgaaccgacgacaagctgacgaccgggtct
ccgcaagtggcacttttcggggaaatgtgcgcggaacccctatttgtttatttttctaaatacattcaaatatgtatccgctcatgaatt
aattcTTAGAAAAACTCATCGAGCATCAAATGAAACTGCAATTTATTCATATCAG
GATTATCAATACCATATTTTTGAAAAAGCCGTTTCTGTAATGAAGGAGAAAA
CTCACCGAGGCAGTTCCATAGGATGGCAAGATCCTGGTATCGGTCTGCGATT
CCGACTCGTCCAACATCAATACAACCTATTAATTTCCCCTCGTCAAAAATAAG
GTTATCAAGTGAGAAATCACCATGAGTGACGACTGAATCCGGTGAGAATGGC
AAAAGTTTATGCATTTCTTTCCAGACTTGTTCAACAGGCCAGCCATTACGCTC
GTCATCAAAATCACTCGCATCAACCAAACCGTTATTCATTCGTGATTGCGCCT
GAGCGAGACGAAATACGCGGTCGCTGTTAAAAGGACAATTACAAACAGGAA
TCGAATGCAACCGGCGCAGGAACACTGCCAGCGCATCAACAATATTTTCACC
TGAATCAGGATATTCTTCTAATACCTGGAATGCTGTTTTCCCGGGGATCGCAG
TGGTGAGTAACCATGCATCATCAGGAGTACGGATAAAATGCTTGATGGTCGG
AAGAGGCATAAATTCCGTCAGCCAGTTTAGTCTGACCATCTCATCTGTAACAT
CATTGGCAACGCTACCTTTGCCATGTTTCAGAAACAACTCTGGCGCATCGGGC
TTCCCATACAATCGATAGATTGTCGCACCTGATTGCCCGACATTATCGCGAGC
CCATTTATACCCATATAAATCAGCATCCATGTTGGAATTTAATCGCGGCCTAG
AGCAAGACGTTTCCCGTTGAATATGGCTCATactcttccttttcaatattattgaagcatttatcaggg
ttattgtctcatgagcggatacatatttgaatgtatttagaaaaataaacaaataggcatgcagcgctCTTCCGCTTCCTC
GCTCACTGACTCGCTACGCTCGGTCGTTCGACTGCGGCGAGCGGTGTCAGCTC
ACTCAAAAGCGGTAATACGGTTATCCACAGAATCAGGGGATAAAGCCGGAA
AGAACATGTGAGCAAAAAGCAAAGCACCGGAAGAAGCCAACGCCGCAGGCG
TTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAA
GCCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCC
TGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACC
TGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTT
GGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAA
CCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTC
CAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCATTGGTAACTGA
TTTAGAGGACTTTGTCTTGAAGTTATGCACCTGTTAAGGCTAAACTGAAAGAA
CAGATTTTGGTGAGTGCGGTCCTCCAACCCACTTACCTTGGTTCAAAGAGTTG
GTAGCTCAGCGAACCTTGAGAAAACCACCGTTGGTAGCGGTGGTTTTTCTTTA
TTTATGAGATGATGAATCAATCGGTCTATCAAGTCAACGAACAGCTATTCCGT
TACcatatggtgcactctcagtacaatctgctctgatgccgcatagttaagccagtatacactccgctatcgctacgtgactggg
tcatggctgcgccccgacacccgccaacacccgctgacgcgccctgacgggcttgtctgctcccggcatccgcttacagacaa
gctgtgaccgtctccgggagctgcatgtgtcagaggttttcaccgtcatcaccgaaacgcgcgaggcagcagatcaattcgcgc
gcgaaggcgaagcggcatgcatttacgttgacaccatcgaatggtgcaaaacctttcgcggtatggcatgatagcgcccggaa

gagagtcaattcagggtggtgaatgtgaaaccagtaacgttatacgatgtcgcagagtatgccggtgtctcttatcagaccgtttcc
cgcgtggtgaaccaggccagccacgtttctgcgaaaacgcgggaaaaagtggaagcggcgatggcggagctgaattacattc
ccaaccgcgtggcacaacaactggcgggcaaacagtcgttgctgattggcgttgccacctccagtctggccctgcacgcgccg
tcgcaaattgtcgcggcgattaaatctcgcgccgatcaactgggtgccagcgtggtggtgtcgatggtagaacgaagcggcgtc
gaagcctgtaaagcggcggtgcacaatcttctcgcgcaacgcgtcagtgggctgatcattaactatccgctggatgaccaggat
gccattgctgtggaagctgcctgcactaatgttccggcgttatttcttgatgtctctgaccagacacccatcaacagtattattttctcc
catgaagacggtacgcgactgggcgtggagcatctggtcgcattgggtcaccagcaaatcgcgctgttagcgggcccattaag
ttctgtctcggcgcgtctgcgtctggctggctggcataaatatctcactcgcaatcaaattcagccgatagcggaacgggaaggc
gactggagtgccatgtccggttttcaacaaaccatgcaaatgctgaatgagggcatcgttcccactgcgatgctggttgccaacg
atcagatggcgctgggcgcaatgcgcgccattaccgagtccgggctgcgcgttggtgcggatatctcggtagtgggatacgac
gataccgaagacagctcatgttatatcccgccgtcaaccaccatcaaacaggattttcgcctgctggggcaaaccagcgtggac
cgcttgctgcaactctctcagggccaggcggtgaagggcaatcagctgttgcccgtctcactggtgaaaagaaaaaccaccctg
gcgcccaatacgcaaaccgcctctccccgcgcgttggccgattcattaatgcagctggcacgacaggtttcccgactggaaagc
gggcagtgagcgcaacgcaattaatgtgagttagcgcgaattgatctg

Data availability statement

Source data for all figures are available from the corresponding author upon reasonable request.

References

1 Potter, S. C. *et al.* HMMER web server: 2018 update. *Nucleic Acids Res.* **46**, W200--W204, doi:10.1093/nar/gky448 (2018).

2 Consortium, T. U. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res.* **47**, D506--D515, doi:10.1093/nar/gky1049 (2018).

3 Bellaousov, S., Reuter, J. S., Seetin, M. G. & Mathews, D. H. RNAstructure: web servers for RNA secondary structure prediction and analysis. *Nucleic Acids Res.* **41**, W471--W474, doi:10.1093/nar/gkt290 (2013).

4 Salis, H. M., Mirsky, E. A. & Voigt, C. A. Automated design of synthetic ribosome binding sites to control protein expression. *Nat. Biotechnol.* **27**, 946–950, doi:10.1038/nbt.1568 (2009).

5 Salis, H. M. Chapter two - The Ribosome Binding Site Calculator. *Methods Enzymol.* **498**, 19–42, doi:10.1016/b978-0-12-385120-8.00002-4 (2011).

6 Espah Borujeni, A., Channarasappa, A. S. & Salis, H. M. Translation rate is controlled by coupled trade-offs between site accessibility, selective RNA unfolding and sliding at upstream standby sites. *Nucleic Acids Res.* **42**, 2646–2659, doi:10.1093/nar/gkt1139 (2013).

7 Espah Borujeni, A. & Salis, H. M. Translation Initiation is Controlled by RNA Folding Kinetics via a Ribosome Drafting Mechanism. *J. Am. Chem. Soc.* **138**, 7016–7023, doi:10.1021/jacs.6b01453 (2016).

8 Espah Borujeni, A. *et al.* Precise quantification of translation inhibition by mRNA structures that overlap with the ribosomal footprint in N-terminal coding sequences. *Nucleic Acids Res.* **45**, 5437–5448, doi:10.1093/nar/gkx061 (2017).

9 Keseler, I. M. *et al.* The EcoCyc database: reflecting new knowledge about Escherichia coli K-12. *Nucleic Acids Res.* **45**, D543–D550, doi:10.1093/nar/gkw1003 (2016).

10 Willis, J. C. W. & Chin, J. W. Mutually orthogonal pyrrolysyl-tRNA synthetase/tRNA pairs. *Nature Chemistry* **10**, 831-837, doi:10.1038/s41557-018-0052-5 (2018).

11 Borrel, G. *et al.* Unique Characteristics of the Pyrrolysine System in the 7th Order of Methanogens: Implications for the Evolution of a Genetic Code Expansion Cassette. *Archaea* **2014**, 11, doi:10.1155/2014/374146 (2014).

12 Krzycki, J. A. The direct genetic encoding of pyrrolysine. *Curr. Opin. Microbiol.* **8**, 706--712, doi:10.1016/j.mib.2005.10.009 (2005).

13 Herring, S. *et al.* The amino-terminal domain of pyrrolysyl-tRNA synthetase is dispensable in vitro but required for in vivo activity. *FEBS Lett.* **581**, 3197--3203, doi:10.1016/j.febslet.2007.06.004 (2007).
14 Gaston, M. A., Jiang, R. & Krzycki, J. A. Functional context, biosynthesis, and genetic encoding of pyrrolysine. *Curr. Opin. Microbiol.* **14**, 342–349, doi:10.1016/j.mib.2011.04.001 (2011).
15 Srinivasan, G., James, C. M. & Krzycki, J. A. Pyrrolysine Encoded by UAG in Archaea: Charging of a UAG-Decoding Specialized tRNA. *Science* **296**, 1459--1462, doi:10.1126/science.1069588 (2002).