

1 **Supplementary information to:**

2

3 **Ultradeep characterisation of translational sequence determinants**
4 **refutes rare-codon hypothesis and unveils quadruplet base pairing**
5 **of initiator tRNA and transcript**

6 Simon Höllerer¹ and Markus Jeschek^{1,2*}

7

8 ¹Department of Biosystems Science and Engineering, Swiss Federal Institute of Technology – ETH
9 Zurich, Basel, CH-4058, Switzerland

10 ²Institute of Microbiology, Synthetic Microbiology Group, University of Regensburg, Regensburg,
11 D-93053, Germany

12

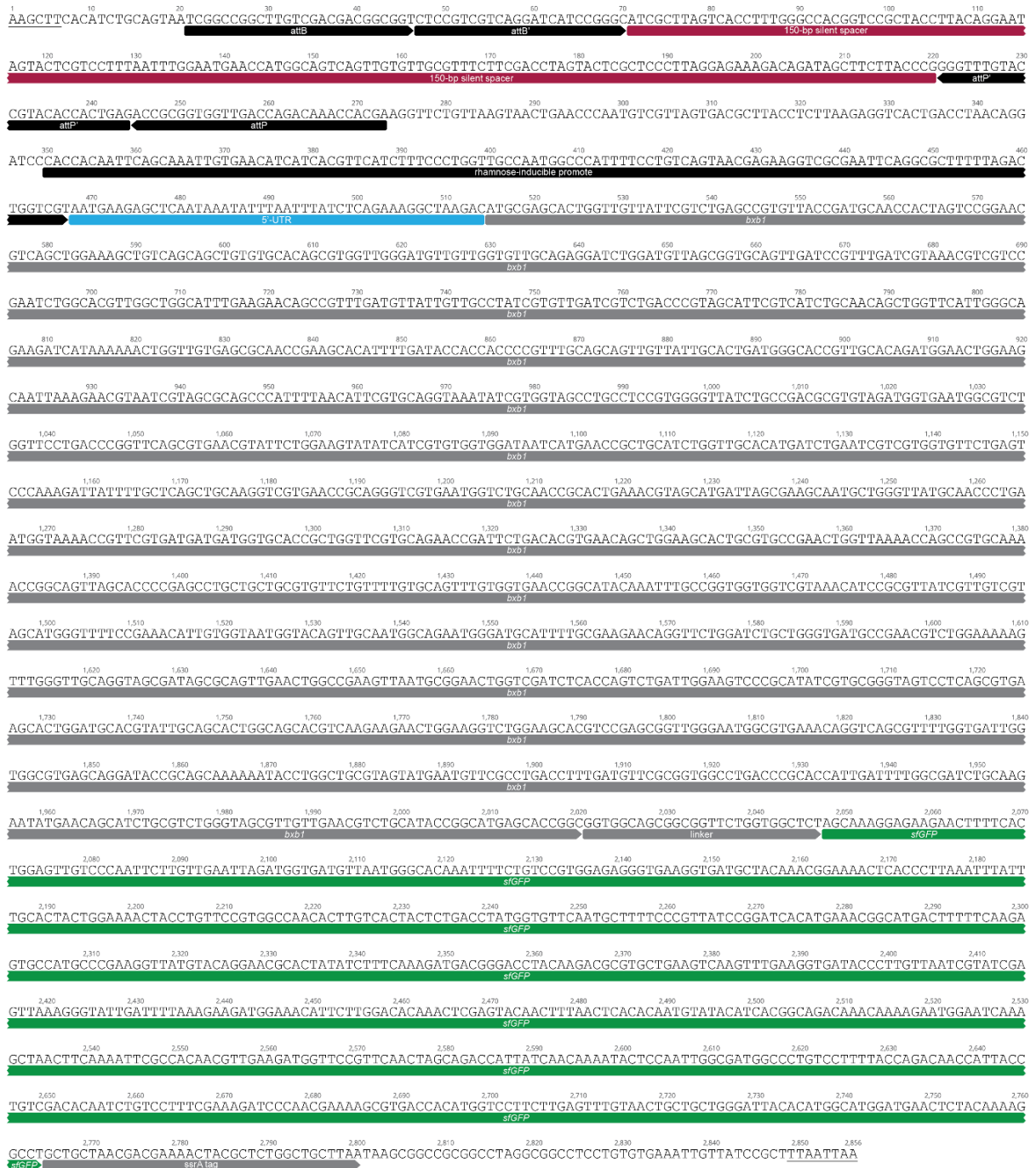
13 *To whom correspondence should be addressed: E-mail: markus.jeschek@ur.de

14 **Supplementary figures**

15

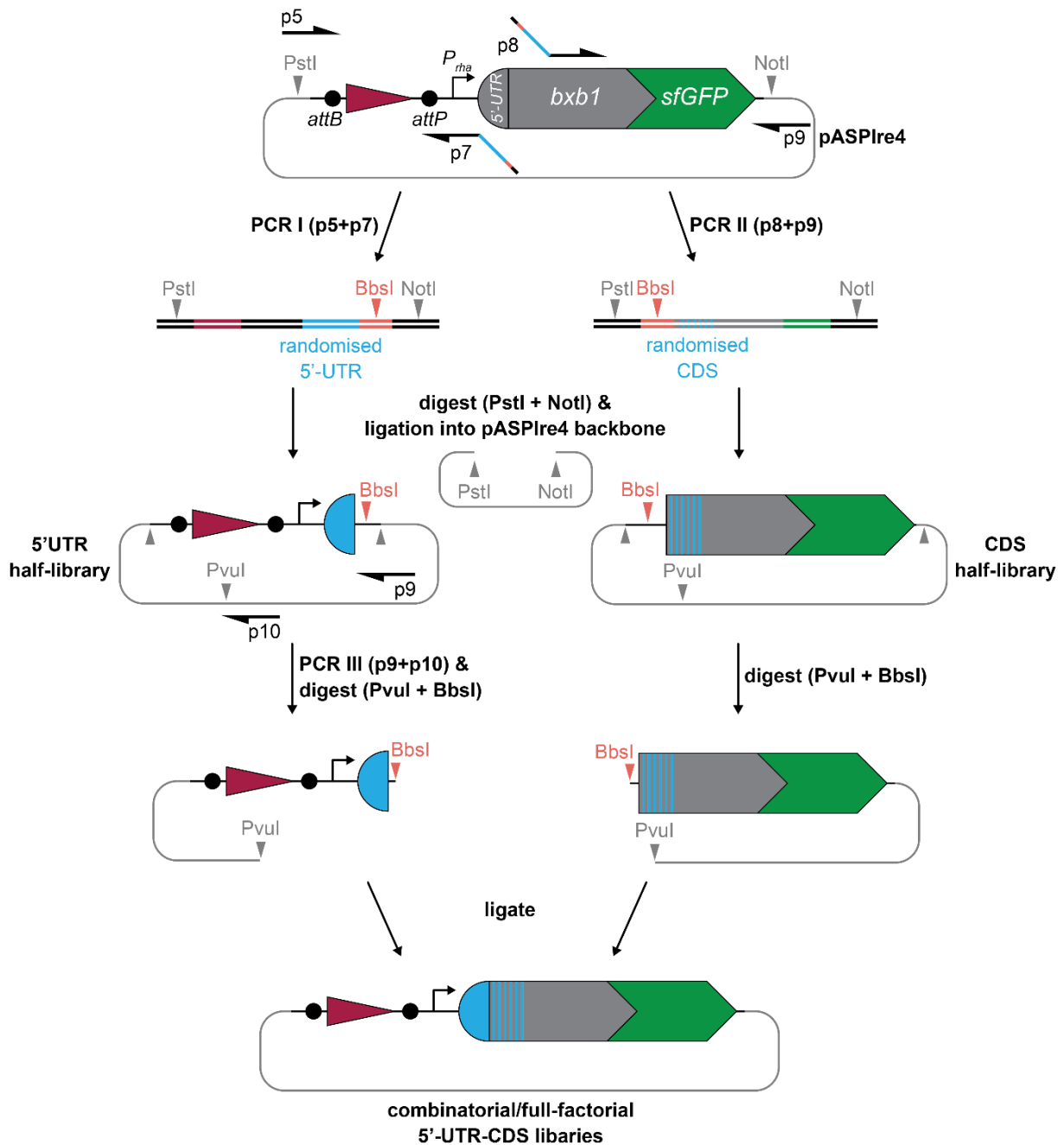
16 All DNA sequences follow the IUPAC nucleotide code; N: A/C/G/T; H: A/C/T; Y: C/T.

17



18

19 **Supplementary Figure 1: Graphical representation of pASPIre4.** The displayed sequence
20 corresponds to the insert between HindIII and PstI restriction sites (underlined) in pSEVA291. srfA:
21 proteolytic degradation tag. The full sequence is available in text format in **Supplementary Note 2.**



26

27 **Supplementary Figure 3: Cloning scheme for the generation of combinatorial and full-factorial**

28 **5'-UTR-CDS libraries.** Libraries were generated from 5'-UTR and CDS half-libraries as described in

29 the **Methods** section. The 5'-UTR (left) half-library was generated by PCR with primers p5 and p7 with

30 pASPIre4 as template. Primer p7 introduces degeneracy in the 5'-UTR and a BbsI site between the

31 randomised 5'-UTR and the NotI site. The CDS half-library (right) was generated by PCR with primers

32 p8 and p9 on pASPIre4 as template. Primer p8 introduces degeneracy in the CDS and a BbsI site

33 between the CDS and the PstI site. The resulting PCR products were then sub-cloned into pASPIre4

34 to retrieve half libraries. In a second step, 5'-UTR and CDS half-libraries were combined to generate

35 full libraries Lib_{comb1}, Lib_{comb2} and Lib_{fact}. To achieve this, plasmid DNA of the 5'-UTR half-library was

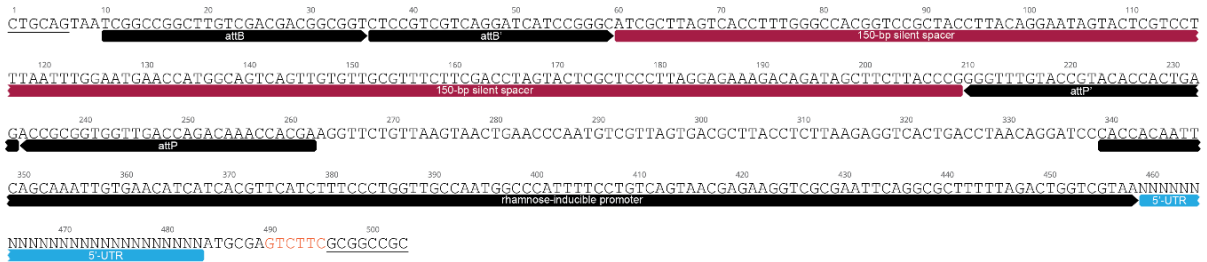
36 PCR-amplified with primers p9 and p10, and the product was digested with BbsI and PvuI and ligated

37 into the backbone prepared from the CDS half-library via digestion with PvuI and BbsI. Note that the

38 BbsI type IIS restriction site enables scarless joining of 5'-UTR and CDS half-libraries using ATGC (start
39 codon ATG + first downstream base) as sticky ends for ligation.

40

41



42

43 **Supplementary Figure 4: Graphical representation of 5'-UTR half-libraries.** The displayed
44 sequence corresponds to the insert between PstI and NotI restriction sites (underlined) in pASPIre4.
45 The introduced BbsI restriction site (orange) is used for scarless joining of 5'-UTR and CDS half-
46 libraries. The full sequence is available in text format in **Supplementary Note 4**.

1 10 20 30 40 50 60 70 80 90 100 110
 CTGCAGGAAGACCCATGCGNGCNCNTNGTNGTNAHCHGNCNTNCNCGNGTACNGAYGNCNACNACTAGTCCGGAACGTGAGCTGGAAAGCTGTGACGACGCTGTGTGCACAGC

120 130 140 150 160 170 180 190 200 210 220
 GTGGTTGGGATGTTGTTGGTGTGACAGGATCTGGATGTTAGCCGGTGCAGTTGATCCGTTTGTATCGTAAACGTGTCGGAATCTGGCACGTTGGCTGGCATTGAAGAAC

230 240 250 260 270 280 290 300 310 320 330
 AGCCGTTTGTATGTTATGTTGCCTATCGTGTGATCGTCTGACCCGTAGCATTTCGTATCTGCAACAGCTGGTTCATTGGGCAGAAGATCATAAAAAACGGTGTGTGAGCG

340 350 360 370 380 390 400 410 420 430 440
 CAACCGAAGCACATTTTGATACCACCACCCCGTTTGCAGCAGTTGTTATTCGACTGATGGGCACCGTTGCACAGATGGAACGGAAAGCAATTAAGAACGTAATCGTAGCG

450 460 470 480 490 500 510 520 530 540 550
 CAGCCCATTTAACATTCGTGCAGGTAATATCGTGTAGCCTGCCCTCCGTTGGGTTATCTGCCGACCGCTGATAGTGGTGAATGGGCTCTGTTCTGACCCGTTCCAGC

560 570 580 590 600 610 620 630 640 650 660
 GTGAACGTATTCTGGAAGTATATCATCGTGTGGTGGATAATCATGAACCCGCTGCATCTGGTTGCACATGATCTGAATCGTGTGGTGTCTGAGTCCCAAAGATTATTTTG

670 680 690 700 710 720 730 740 750 760 770
 CTCAGTGCAGGTCGTGAACCGCAGGGTCGTGAATGGTTCGAACCGCACTGAAACGTAGCATGATAGCGAAGCAATGCTGGTATGCAACCCGTAATGGTAAACCCG

780 790 800 810 820 830 840 850 860 870 880
 TTCGTGATGATGATGGTGCACCGCTGGTTCGTGCAGAACGATTTCTGACACGTGAACAGCTGGAAGCACTGCGTGCCGAACGTTAAACCCAGCCGTGCAAAACCGGCAG

890 900 910 920 930 940 950 960 970 980 990
 TTAGCACCCGACCCCTGCTGCTGCGTGTCTGTTTTGTCAGTTTGTGGTGAACCGGCATACAAATTTGCCGGTGGTGGTCTAAACATCCCGCTTATCGTTGCTGATGCA

1,000 1,010 1,020 1,030 1,040 1,050 1,060 1,070 1,080 1,090 1,100 1,110
 TGGGTTTTCCGAAACATTGTGGTAATGGTACAGTTGCAATGGCAGAAATGGGATGCATTTTGGCAAGAACAGGTTCTGGATCTGCTGGGTGATGCCAACGTCGGAAAAAG

1,120 1,130 1,140 1,150 1,160 1,170 1,180 1,190 1,200 1,210 1,220
 TTTGGGTTGACGATAGCGATAGCGCAGTTGAACGGCCGAAAGTTAATGCGGAACCTGGTCGATCTCACCAGTCTGATGGAAGTCCCACATATCGTGGGGTAGTCTCAGC

1,230 1,240 1,250 1,260 1,270 1,280 1,290 1,300 1,310 1,320 1,330
 GTGAAGCACTGGATGCACGTATTGCAGCCTGGCAGCAGCTCAAGAAGAACTGGAAGTCTGGAAGCACGTCGGAGCGGTTGGAAATGGCGTGAACAGGTCACGCTTTTG

1,340 1,350 1,360 1,370 1,380 1,390 1,400 1,410 1,420 1,430 1,440
 GTGATTGGTGGCGTAGCAGGATACCCGAGCAAAAAATACCTGGCTGCTAGTATGAATGTTGCGCTGACCTTTGATGTTGCGGTTGGCTGACCCCGACCATGATTTTG

1,450 1,460 1,470 1,480 1,490 1,500 1,510 1,520 1,530 1,540 1,550
 GCGATCTGAAGAATATGAACAGCATCTGCGCTGGGTAGCGTTGTTGAACGCTGCAATACCGGCATGAGCACCGGCGGTTGGCAGCGCGGTTCTGGTGGCTTAGCAAAAG

1,560 1,570 1,580 1,590 1,600 1,610 1,620 1,630 1,640 1,650 1,660
 GAGAAGAACTTTCACTGGAGTTGCCAAATCTTGTGAATAGATGGTGTGTTAATGGGCACAAAATTTCTGTCCTGGAGAGGGTGAAGGTGATGTACAAACGGAA

1,670 1,680 1,690 1,700 1,710 1,720 1,730 1,740 1,750 1,760 1,770
 AACTCACCCCTTAATTTATTTGCACTACTGGAAAACCTACCTGTTCCGTGGCCAACTGTGCACTACTGACCTATGGTGTCAATGCTTTTCCCGTTATCCGGATCACA

1,780 1,790 1,800 1,810 1,820 1,830 1,840 1,850 1,860 1,870 1,880
 TGAACGGCATGACTTTTCAAGAGTGCCTATGCCGAAAGGTTATGTACAGGAACCGCATAATCTTTCAAAGATGACGGGACCTACAAGACGCGTGTGAAGTCAAGTTTG

1,890 1,900 1,910 1,920 1,930 1,940 1,950 1,960 1,970 1,980 1,990
 AAGGTGATACCCCTTGTAAATCGTATCGAGTTAAAGGGTATTGATTTTAAAGAAGAAGTGAACAACTTCTGGACACAACTCGAGTACAACCTTAACTCACACAATGTATACA

2,000 2,010 2,020 2,030 2,040 2,050 2,060 2,070 2,080 2,090 2,100
 TCACGGCAGACAAACAAAAGAAATGGAATCAAAGCTAACTCAAATTCGCCAACAGCTGAAGATGGTTCGGTTCAACTAGCAGACCATTAACAACAAAATACTCCAATTG

2,110 2,120 2,130 2,140 2,150 2,160 2,170 2,180 2,190 2,200 2,210 2,220
 GCGATGGCCCTGTCTTTTACAGACAACCTTACCTGTGCACACAATCTGCTCTTCCGAAAGATCCCAACGAAAGCGTGACCACATGCTCTTCTTGTGACTTTGTAACCTC

2,230 2,240 2,250 2,260 2,270 2,280 2,290 2,300 2,311
 CTGCTGGGATTACACATGGCATGGATGAACCTACAAAAGGCTGCTGCTAACGACGAAAACCTACGCTCTGGCTGCTTAAATAAGCGGCCGC

Diagrammatic representation of the sequence above:
 - **Underlined**: PstI and NotI sites (1,450-1,550)
 - **Orange bar**: BbsI site (1,450-1,550)
 - **Green bars**: stGFP genes (1,560-1,660, 1,670-1,770, 1,780-1,880, 1,890-1,990, 2,000-2,100, 2,110-2,220)
 - **Grey bar**: ssrA tag (2,270-2,280)

47

48 **Supplementary Figure 5: Graphical representation of CDS half-libraries.** The displayed sequence
 49 corresponds to the insert between PstI and NotI restriction sites (underlined) in pASPIre4. The
 50 introduced BbsI restriction site (orange) is used for scarless joining of 5'-UTR and CDS half-libraries.
 51 The full sequence in text format is available in **Supplementary Note 5**.

```

1      10      20      30      40      50      60      70      80      90      100     110
GGTACCAAATAACACCCCTGCTTAATTAAGTATGATGAGCCTGGATTTCGGCTCTCACTGAATTTTATGCAAAATAAATGAGTTTTCATTTAATCATCTTTTATCGGAGACAGG

120     130     140     150     160     170     180     190     200     210     220     230
AAGAGTTTAGTGTGTTTTTTGTAAAATAATGCGCTTAAGGGAGAGCAGGAGAAGGCAAAAGTATTCACAAATGAAAGTGAAGTGGATATTCATTCACATGATTAGCAATAAACGT
                                     promoter metY2

240     250     260     270     280     290     300     310     320     330     340
TGACAAATGTGGCGTGGATCAGTATAATGCCIGCAGATTTTACCTCCCGTCTCGGTACACCAAAATCCCAGCAGTATTTGCATTTTACCCAAAACGAGTAGAATTTGCCACGTT
promoter metY2                                     promoter metY1

350     360     370     380     390     400     410     420     430     440     450     460
TCAGGCGGGGGTGGAGCAGCCTGGTAGCTCGTCCGGGCTCATNACCCGAAGGTCGTCGGTTCAAATCCGGCCCCCGCAACCACTTCCCTTAGAGTCCTTTTCAAATATACTGTG
promoter metY1                                     tRNAfMet variant

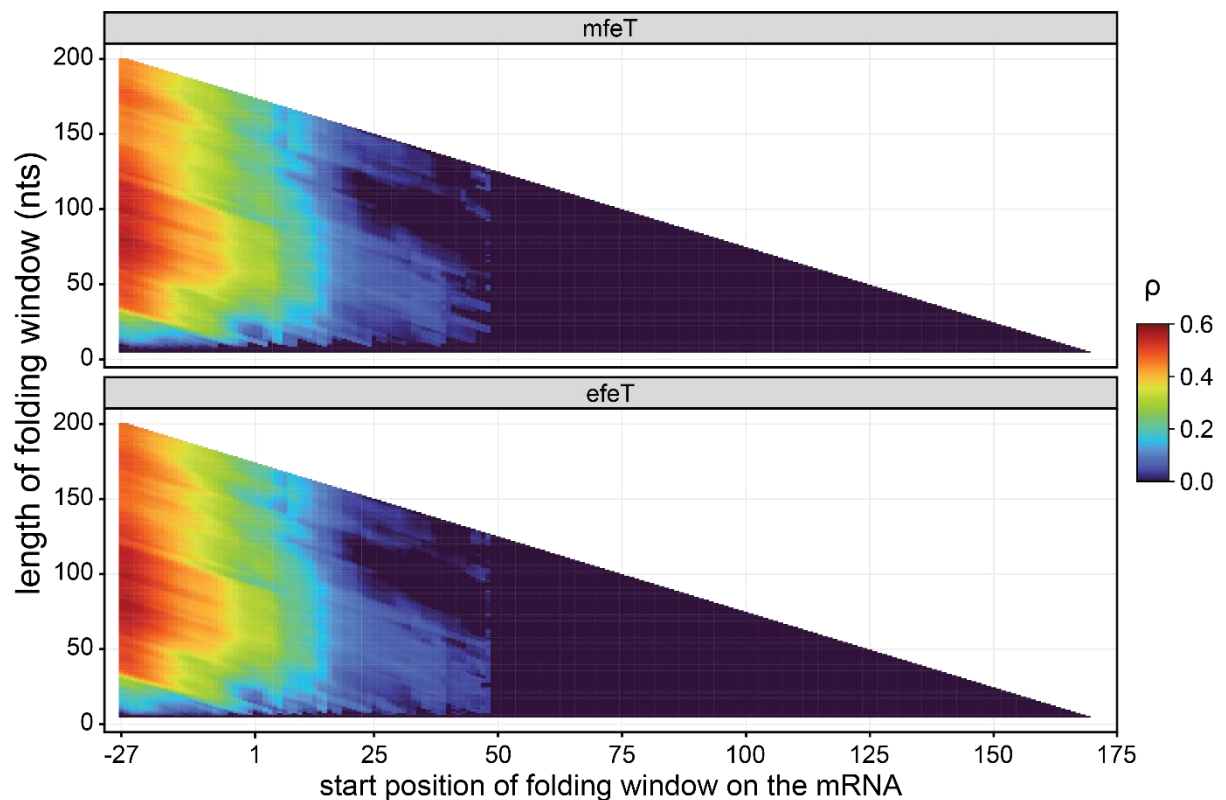
470     480     490     500     510     520     530     540     550     560     570     580
AAGACTTCGGCCTTCGTAGTGGGATTTGAAAAATCCTTCTGGAAAGTGCTCCAGACCCAGTTGCGGTTATAGGGTTCAGTTATATAAAGCCCGATTTATCGGGGTTTTTTTGTG
                                     terminator

590     600     610     620     630     640     642
ATCTGACTACAGAATAACTGGGCTTTAGGCCCTTTTTTTATGTCCTGGGGTGGGCACTAGT
                                     terminator

```

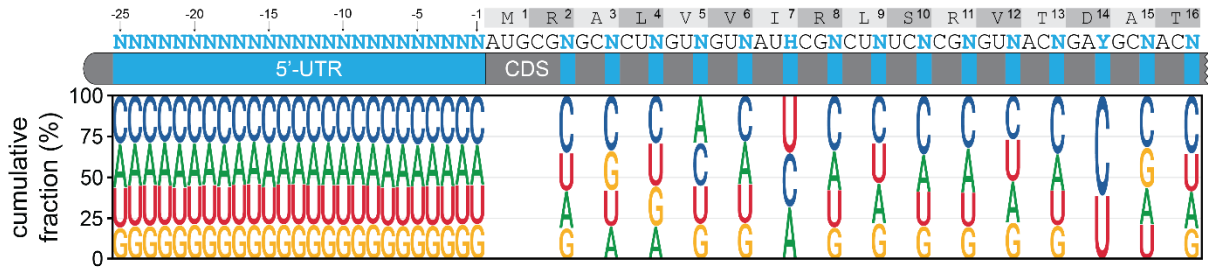
52
53
54
55
56
57
58
59

Supplementary Figure 6: Graphical representation of plasmids for the overexpression of tRNA^{fMet} variants. The displayed sequence corresponds to the native chromosomal locus of *metY* including regulatory sequences. The bold N nucleotide indicates the mutated position 37 (p37) in tRNA^{fMet}. The window between KpnI and SpeI restriction sites (underlined) in pSEVA361 is shown. The full sequence is available in text format in **Supplementary Note 6**.



60
61
62
63
64

Supplementary Figure 7: Correlation between predicted folding energy and rTR for different mRNA windows. Folding energies (mfeT and efeT) of all possible mRNA sequence windows with lengths between 5 and 200 bases were predicted using *RNAfold* of the *Vienna RNA* package and the resulting values were correlated with rTR using Spearman's correlation (**Methods**).

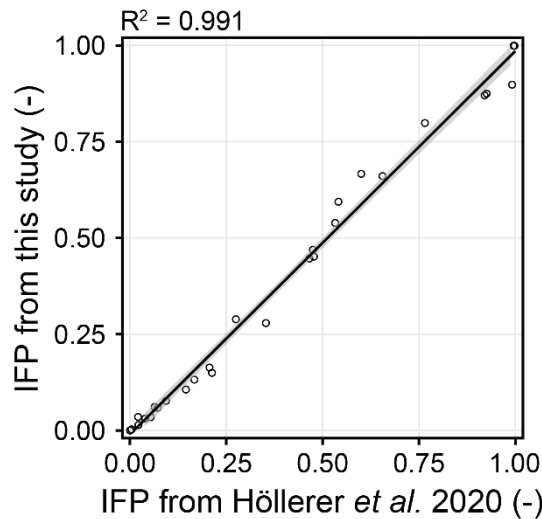


65

66 **Supplementary Figure 8: Base distribution in Lib_{random} .** The base distribution amongst all 198,174
 67 variants of Lib_{random} above high quality read count threshold is displayed for each randomised position.
 68 Bases are ordered from the most (top) to the least (bottom) frequent for each position. Note that in
 69 codons 7 (Ile) and 14 (Asp) only three (A, C, U) and two (C, U) bases were allowed, respectively, to
 70 avoid non-synonymous mutations.

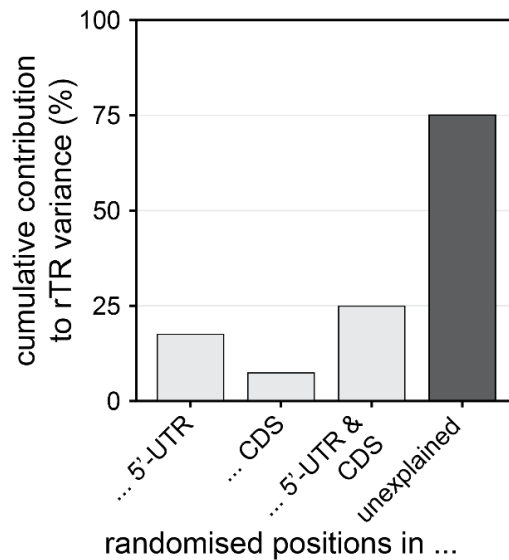
71

72



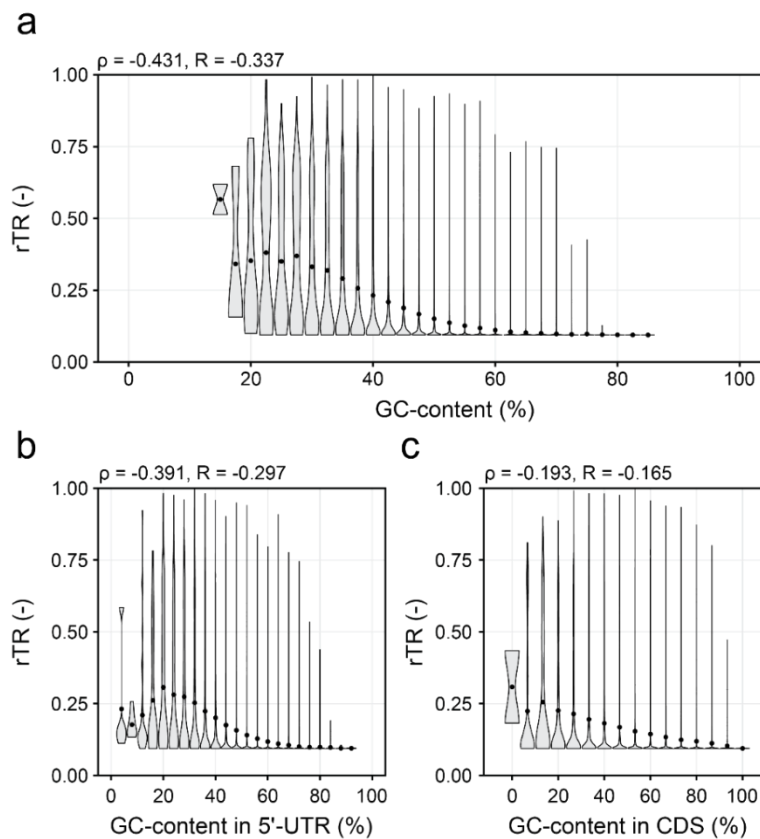
73

74 **Supplementary Figure 9: Cross-study reproducibility of uASPIre.** The IFP of 31 standard RBSs as
 75 measured in this study is plotted over the IFP of the same standard RBSs as determined in a previous
 76 study (1). The black line corresponds to a linear regression with 95% confidence interval (grey area).



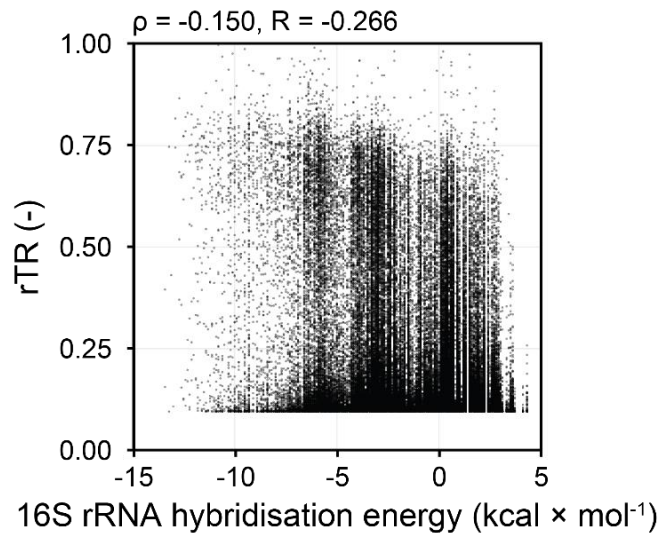
77
78
79
80
81
82

Supplementary Figure 10: Cumulative impact of randomised positions in 5'-UTR and CDS on the rTR. Bars are the cumulative contribution of all randomised positions in the mRNA part(s) indicated on the horizontal axis except for “unexplained” which corresponds to the residuals. The contribution of individual positions (**Fig. 2a**) was determined by ANOVA (**Methods**).



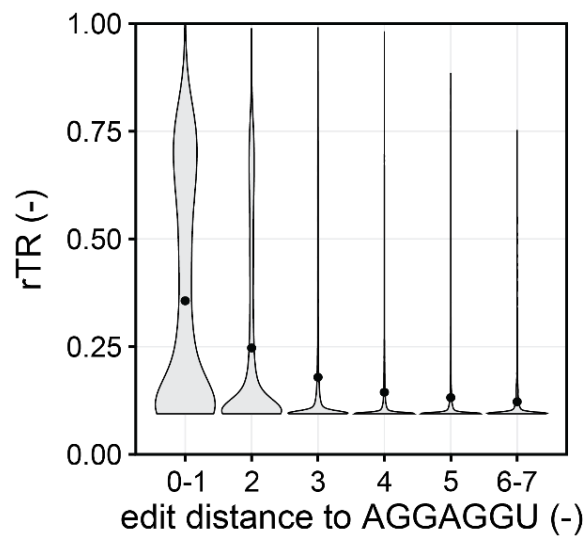
83
84
85
86
87

Supplementary Figure 11: Impact of GC-content on rTR. The correlation of rTR and GC-content in all randomised positions (**a**) as well as in the randomised positions in 5'-UTR (**b**) and CDS (**c**) individually was assessed. Variants were binned according to their GC-content in the respective mRNA part, and the rTR distribution of bins is displayed as violin with each bin's mean rTR (black circles).



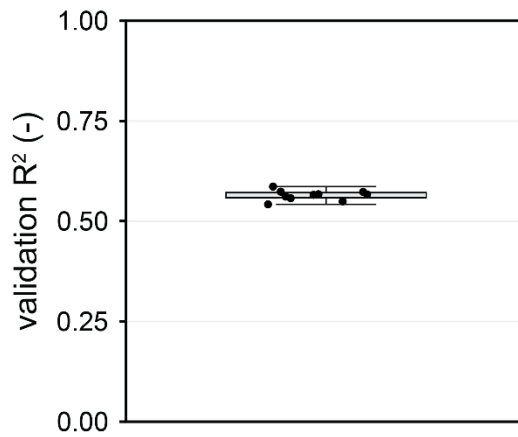
88
89
90
91
92
93
94

Supplementary Figure 12: Correlation between predicted 16S rRNA hybridisation energy and rTR. The hybridisation energy of the 3'-end of *E. coli*'s 16S rRNA (sequence: 5'-ACCUCCUUA-3') and the 5'-UTR between position -18 and -4 was predicted using *RNA duplex* of the *Vienna RNA* package and the resulting energy values were correlated with rTR.



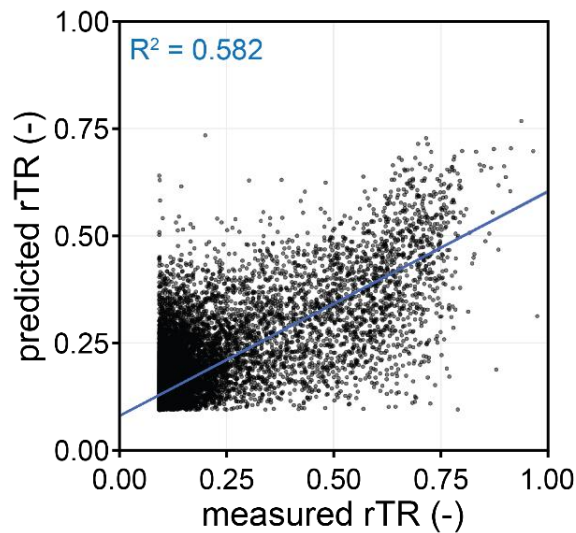
95
96
97
98
99

Supplementary Figure 13: Impact of similarity to the SD motif on rTR. The minimum edit distance between the canonical SD motif AGGAGGU and all 7-nt windows between 5'-UTR positions -18 and -4 was calculated for each variant (**Methods**). Violins are rTR distributions amongst variants grouped by minimum edit distance with group means (black circles).



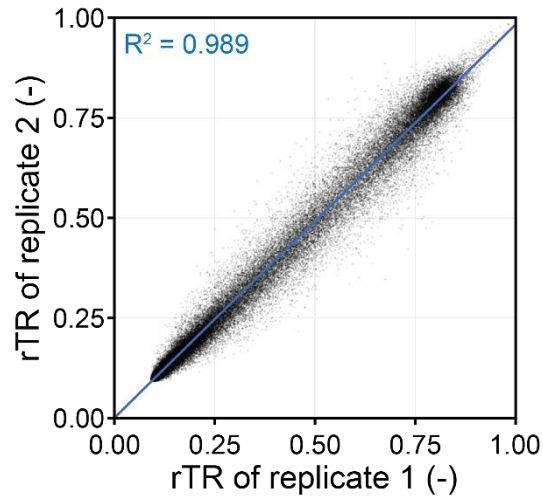
100
101
102
103
104
105

Supplementary Figure 14: Cross-validation of random forest model (Fig. 3e). Validation R^2 values are shown for 10 independent validation runs as circles. Box represents interquartile range with whiskers marking the 1.5-fold interquartile range.



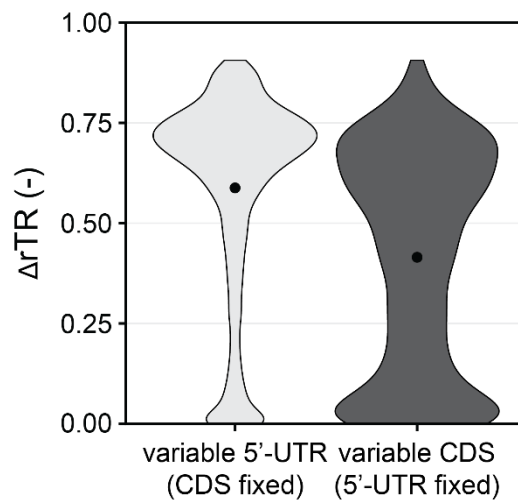
106
107
108
109
110

Supplementary Figure 15: Random forest performance. The rTR values predicted by the random forest model trained on $\text{Lib}_{\text{random}}$ (compare Fig. 3e) are plotted over the experimentally measured rTRs for 19,816 randomly selected test set variants strictly held out during training. The blue line corresponds to a linear regression with corresponding Pearson's R^2 .



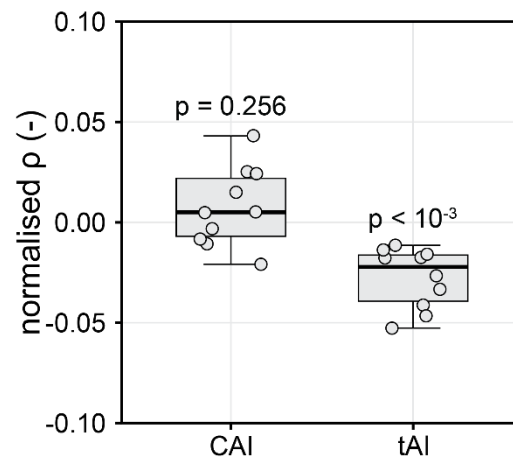
111
 112
 113
 114
 115
 116

Supplementary Figure 16: Biological reproducibility within Lib_{fact}. rTR values of two independently cultured, biological replicates of Lib_{fact} are shown (112,296 variants). The blue line corresponds to a linear regression with corresponding Pearson's R².



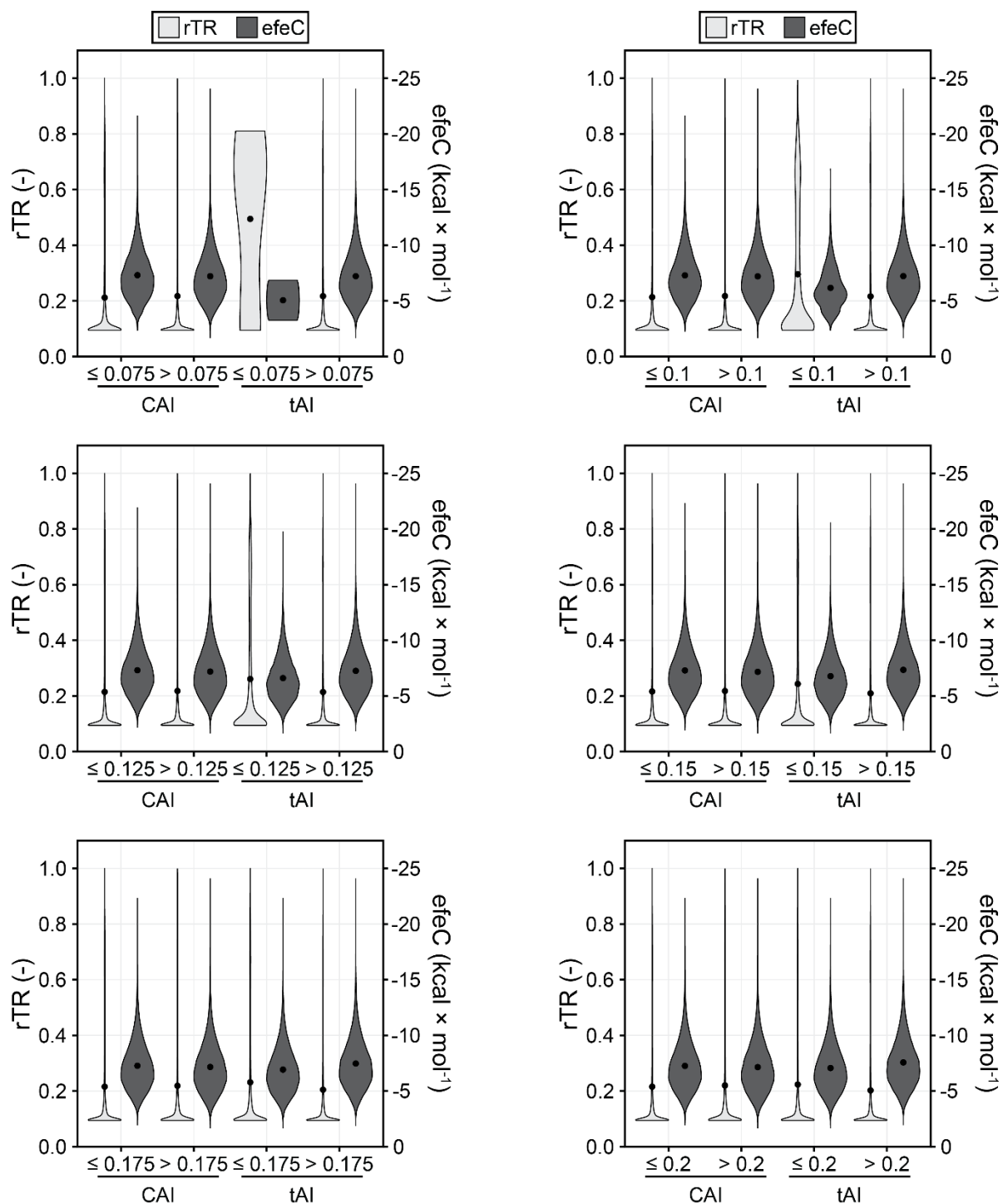
117
 118
 119
 120
 121

Supplementary Figure 17: Coverage of rTR range upon exchange of 5'-UTR (left) and CDS (right). For each CDS (or 5'-UTR) that appeared at least twice in all libraries, the absolute difference in rTR between the strongest and weakest 5'-UTR (or CDS) combined with this CDS (or 5'-UTR) is displayed (ΔrTR). Violins indicate the distribution of ΔrTR with black circles representing the mean value.



123

124 **Supplementary Figure 18: Correlation between codon indices and the CDS-derived contribution**
125 **to rTR variance.** Pearson's correlation between CAI/tAI and the mean rTR of each CDS (over all 5'-
126 UTRs combined with that CDS) was calculated for each of the ten batches of Lib_{fact} (grey circles) and
127 normalised to the mean effect of the CDS on rTR (Fig. 4c). Boxes represents interquartile range with
128 whiskers marking the 1.5-fold interquartile range and medians as solid black lines. P-values of one
129 sample t-tests are indicated above each box.

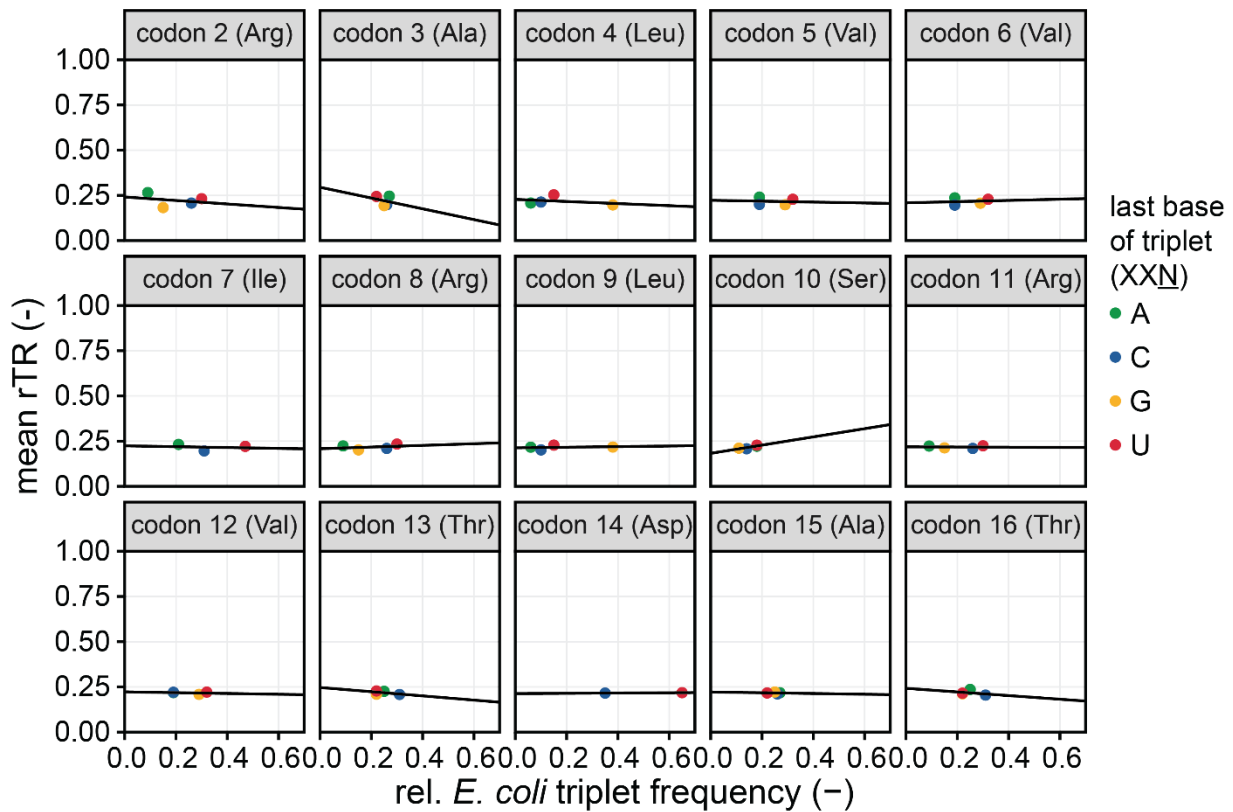


130

131 **Supplementary Figure 19: Comparison of rTRs and predicted folding energies (efeC) of variants**

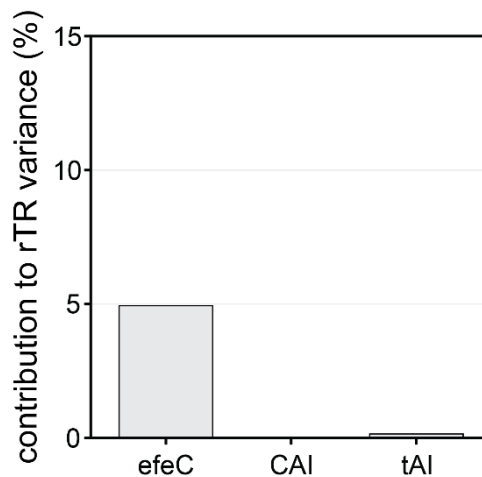
132 **with low and high CAI/tAI. Different thresholds were tested for CAI and tAI. Black circles within violins**

133 **are mean rTR/efeC values (compare Fig. 4e).**



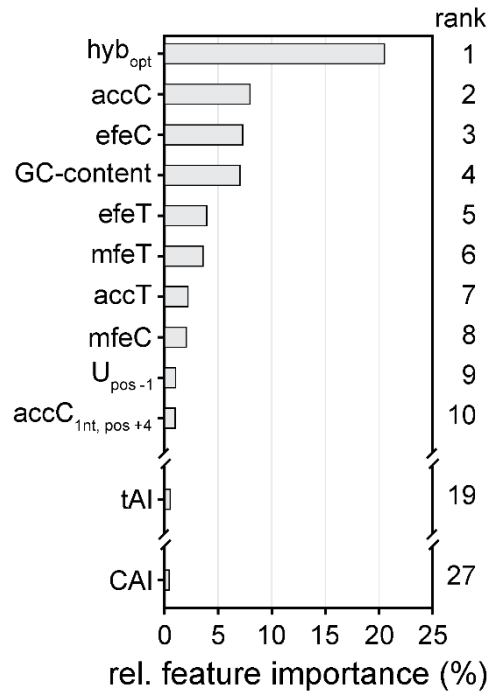
134
 135
 136
 137
 138
 139
 140
 141

Supplementary Figure 20: Dependence between triplet frequency in *E. coli* and rTR. Mean rTRs of variants from all libraries with the same triplet in the randomised codons are plotted over the relative codon frequency of the respective triplet in *E. coli*. Triplet frequencies can be found in **Supplementary Table 6**. Black lines are best linear fits of data points. Note that no consistent up- or downward trend is observed across all codons as well as amongst triplets coding for the same amino acid but in different positions.



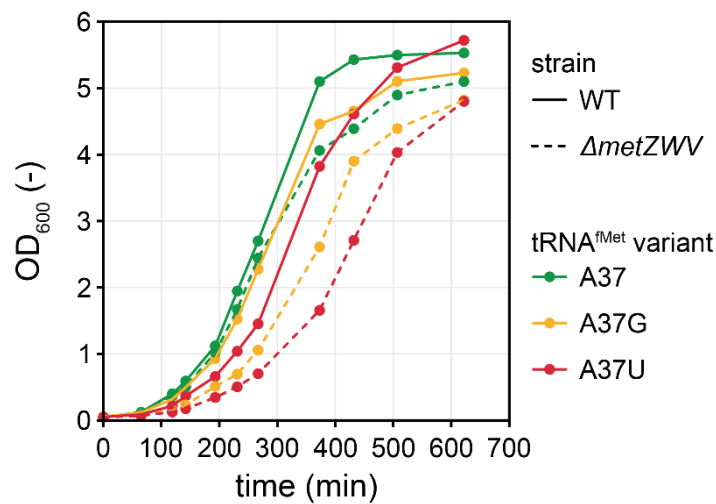
142
 143
 144
 145
 146

Supplementary Figure 21: Contribution of efeC, CAI and tAI to the rTR variance on weakly folding mRNAs. An ANOVA with efeC, CAI and tAI as only covariates was performed to assess their impact on the rTR amongst the 727,570 variants with low tendency for mRNA folding (i.e. efeC \geq -7.5 kcal mol⁻¹, compare **Fig. 4f**).



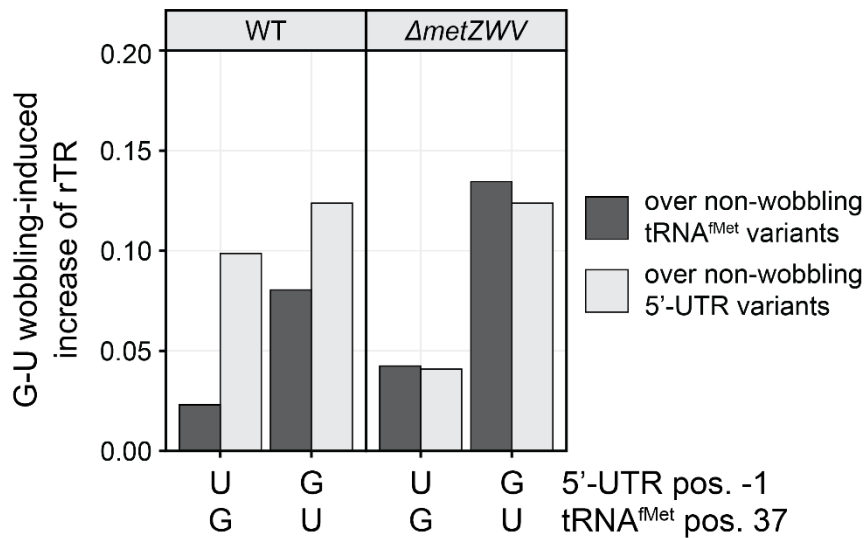
147
148
149
150
151
152
153

Supplementary Figure 22: Feature importance of random forest model trained with codon usage and folding metrics. Relative importance of the ten most important features as well as tAI and CAI for the model displayed in **Figure 5g** (second bar from left) is shown. Ranking of features is indicated to the right. AccC_{1nt, pos +4}: AccC score for position +4 of the CDS.



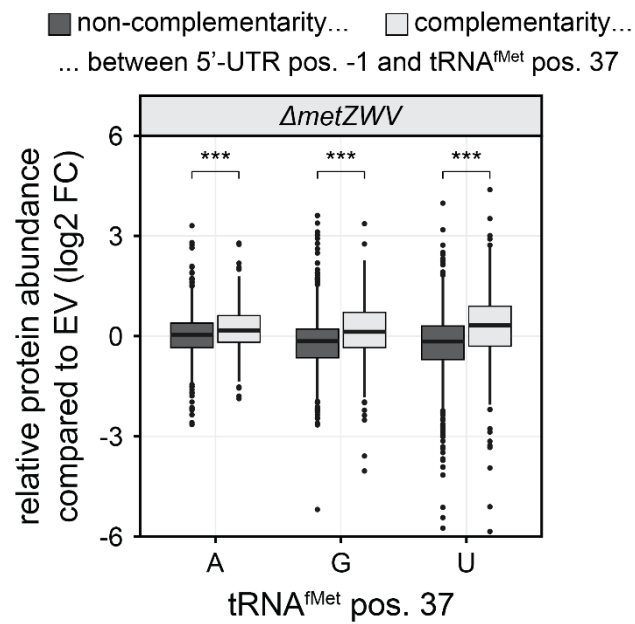
154
155
156
157
158

Supplementary Figure 23: Growth of *E. coli* WT and $\Delta metZ WV$ expressing different tRNA^{fMet} variants. Approximately 50,000 variants of Lib_{random} were co-cultivated in shake flasks (**Methods**) for each of the six combinations of *E. coli* strains (WT or $\Delta metZ WV$) and overexpressed tRNA^{fMet} variants (A37, A37G or A37U).



159

160 **Supplementary Figure 24: Effect of G-U wobbling on rTR.** The mean increase of the rTR is shown
 161 for different combinations of G-U wobbling base pairs between 5'-UTR position -1 and tRNA^{fMet} position
 162 37 as indicated on the horizontal axis. The rTR increases of the respective wobbling pair over all
 163 non-wobbling and non-complementary tRNA^{fMet} variants (dark grey) or over all non-wobbling and non-
 164 complementary 5'-UTR variants (light grey) are displayed. In order to facilitate comparison between the
 165 different experimental groups (i.e. six combinations of strain and tRNA^{fMet} variant), rTR values were
 166 normalised to the mean rTR of each group prior to determination of the G-U wobbling-induced increase
 167 of rTR.



180

181 **Supplementary Figure 26: Impact of overexpression of tRNA^{fMet} variants on endogenous protein**
 182 **expression.** Label-free proteomics were performed on strain *ΔmetZWV* cells overexpressing tRNA^{fMet}
 183 variants with an A, G or U nucleotide at tRNA^{fMet} position 37 (see **Methods**). Relative protein
 184 abundances compared to an empty vector control (EV) are shown as log₂ fold-changes (log₂ FC) for
 185 the 1,098 endogenous proteins detectable in all strains and replicates. Proteins are grouped according
 186 to their complementarity or non-complementarity between 5'-UTR position -1 and tRNA^{fMet} position 37.
 187 *** denote p-values < 10⁻⁴ in a one-sided Welch two sample t-test.

188 **Supplementary tables**

189

190 **Supplementary Table 1: Primers used in this study.** Restriction sites (BbsI/NotI/PstI/SpeI) are
191 underlined. The introduced BbsI site is marked in orange. Primer binding regions are highlighted in bold.

Name	Sequence (5'-3')	Description
p1	TTTGTTCAAAATCATGCCAAATCCGTGATCGGGGTAAAAA TGATCGGCAC GTAAGAGGTTCC	Forward primer for the generation of the knockout cassette $\Delta metZ WV::specR$
p2	GAGAAGGGGATGATAAAAAGGCGCTGAATGGCGCTTTTTTAT TTATTGGCTG GCACCAAGCAG	Forward primer for the generation of the knockout cassette $\Delta metZ WV::specR$
p3	GCGGCAAGCATTGCCACAACCGTGC	Forward genotyping primer for $\Delta metZ WV$
p4	CGTATTTTGCCGATGGGGCGACGCTGG	Reverse genotyping primer for $\Delta metZ WV$
p5	TCTATCAACAGGAGTCCAAG	Forward primer for the generation of all libraries
p6	TATATACTAGTNGTNGCRTCNGTNACNCGNAGNAGNCGDATNACNACNAG NGCNCGCATNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN TTACGACCAGTCTAAAA AGCGCC	Reverse primer for the generation of Lib _{random}
p7	ATATATGCGGCCCGC GAAGACT TCGCATNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN TTACGACCAGTCTAAAAAGCGCC	Reverse primer for the generation of 5'-UTR half-library
p8	ATATATCTGCAG GAAGAC CCCATGCGNGCNCNTNGTNGTNATHCGNCTNTCNC GNGTNACNGAYGCNACN ACTAGTCCGGAACGTCACTGG	Forward primer for the generation of CDS half-library
p9	GCCTTTCGTTTTATTGATGCC	Reverse primer for the generation of CDS half-library and reverse primer for amplification of 5'-UTR half-library insert
p10	GGTTATCCAGGCTAAAATCG	Forward primer for the amplification of 5'-UTR half-library insert

192

193

194 **Supplementary Table 2: Strains used in this study.**

Name	Genotype	Description	Source/Reference
<i>E. coli</i> TOP10 $\Delta rhaA$	<i>F mcrA</i> $\Delta(mrr-hsdRMS-mcrBC)$ $\phi 80lacZ\Delta M15 \Delta lacX74 nupG$ <i>recA1 araD139</i> $\Delta(ara-leu)7697$ <i>galE15 galK16 rpsL(Str^R) endA1</i> $\lambda^- \Delta rhaA$	Rhamnose utilisation-deficient derivative of <i>E. coli</i> TOP10.	Hoellerer <i>et al.</i> (1)
<i>E. coli</i> TOP10 $\Delta rhaA \Delta metZ WV$	<i>F mcrA</i> $\Delta(mrr-hsdRMS-mcrBC)$ $\phi 80lacZ\Delta M15 \Delta lacX74 nupG$ <i>recA1 araD139</i> $\Delta(ara-leu)7697$ <i>galE15 galK16 rpsL(Str^R) endA1</i> $\lambda^- \Delta rhaA \Delta metZ WV::specR$	Derivative of <i>E. coli</i> TOP10 $\Delta rhaA$ with deleted <i>metZ WV</i> locus.	This study (Methods)

195

196

197 **Supplementary Table 3: Plasmids used in this study.**

Name	Description	Source
pSEVA291	contains a pBR322 replicon, a kanamycin resistance cassette, and a multiple cloning site	Silva-Rocha <i>et al.</i> (2)
pSEVA361	Backbone for tRNA ^{fMet} plasmids; contains a p15A replicon, a chloramphenicol resistance cassette, and a multiple cloning site	Silva-Rocha <i>et al.</i> (2)
pASPIre3	Derivative of pSEVA291; contains a <i>bxh1-sfGFP</i> gene (translational fusion) under control of <i>P_{tha}</i> and an <i>attB-attP</i> -flanked 150 bp stretch of silent DNA	Hoellerer <i>et al.</i> (1)
pASPIre4	Derivative of pASPIre3 with <i>SpeI</i> site in the CDS of <i>bxh1</i> . PCR template and backbone for all 5'-UTR and CDS libraries	This study
pASPIre4 _{lib}	Derivative of pASPIre4 with randomised 5'-UTR and CDS	This study
ptRNA ^{fMet-A37}	Derivative of pSEVA361 with expression cassette for tRNA ^{fMet} with the native A at position 37	This study
ptRNA ^{fMet-A37C}	Derivative of ptRNA ^{fMet-A37} with A37C substitution	This study
ptRNA ^{fMet-A37G}	Derivative of ptRNA ^{fMet-A37} with A37G substitution	This study
ptRNA ^{fMet-A37U}	Derivative of ptRNA ^{fMet-A37} with A37T/U substitution	This study

198

199

200 **Supplementary Table 4: DNA adapters for NGS.** Adapters contain *NcoI*- and *SpeI*-compatible
 201 overhangs (orange) for ligation. Binding regions for Illumina flow cell (blue) and sequencing primers
 202 (green) are highlighted. Sample-specific indices are shown in red. P: 5'-phosphorylation.

Name	Sequence
<i>NcoI</i> ₁	5' AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTATCACC 5' 3' TTACTATGCCGCTGGTGGCTCTAGATGTGAGAAAGGATGTGCTGCGAGAAGGCTAGAAAGTAGTGGGTAC-P 3'
<i>NcoI</i> ₂	5' AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTATCGATGTC 5' 3' TTACTATGCCGCTGGTGGCTCTAGATGTGAGAAAGGATGTGCTGCGAGAAGGCTAGATAAGCTACAGGTAC-P 3'
<i>NcoI</i> ₃	5' AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTGATCTTGTAC 5' 3' TTACTATGCCGCTGGTGGCTCTAGATGTGAGAAAGGATGTGCTGCGAGAAGGCTAGACTAGAACATGGTAC-P 3'
<i>NcoI</i> ₄	5' AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTCGATGCCAATC 5' 3' TTACTATGCCGCTGGTGGCTCTAGATGTGAGAAAGGATGTGCTGCGAGAAGGCTAGAGCTACGGTITAGGTAC-P 3'
<i>NcoI</i> ₅	5' AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTTCGATACAGTGC 5' 3' TTACTATGCCGCTGGTGGCTCTAGATGTGAGAAAGGATGTGCTGCGAGAAGGCTAGAAAGCTATGTACAGGTAC-P 3'
<i>NcoI</i> ₆	5' AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTATCGATACTTGAC 5' 3' TTACTATGCCGCTGGTGGCTCTAGATGTGAGAAAGGATGTGCTGCGAGAAGGCTAGATAGCTATGAAGTGGTAC-P 3'
<i>SpeI</i> ₁	5' P-CTAGAATAACGTAAGATCGGAAGAGCACACGTCTGAACTCCAGTCACATCTCGTATGCCGCTTCTGCTTG 5' 3' TTATTGCATTTCTAGCCTTCTCGTGTGCAGACTTGAGGTCAGTGTAGAGCATACGGCAGAAGACGAAC 3'
<i>SpeI</i> ₂	5' P-CTAGATTCTTGAAATAGATCGGAAGAGCACACGTCTGAACTCCAGTCACATCTCGTATGCCGCTTCTGCTTG 5' 3' TAAGAACTTTATCTAGCCTTCTCGTGTGCAGACTTGAGGTCAGTGTAGAGCATACGGCAGAAGACGAAC 3'
<i>SpeI</i> ₃	5' P-CTAGAGGCAGATCATCAGATCGGAAGAGCACACGTCTGAACTCCAGTCACATCTCGTATGCCGCTTCTGCTTG 5' 3' TCCGTCTAGTAGTCTAGCCTTCTCGTGTGCAGACTTGAGGTCAGTGTAGAGCATACGGCAGAAGACGAAC 3'
<i>SpeI</i> ₄	5' P-CTAGACTATGTTAATCGAGATCGGAAGAGCACACGTCTGAACTCCAGTCACATCTCGTATGCCGCTTCTGCTTG 5' 3' TGATACAATTAGCTCTAGCCTTCTCGTGTGCAGACTTGAGGTCAGTGTAGAGCATACGGCAGAAGACGAAC 3'
<i>SpeI</i> ₅	5' P-CTAGAGTTGACGCATCGAAGATCGGAAGAGCACACGTCTGAACTCCAGTCACATCTCGTATGCCGCTTCTGCTTG 5' 3' TCAACTGCGTAGCTTCTAGCCTTCTCGTGTGCAGACTTGAGGTCAGTGTAGAGCATACGGCAGAAGACGAAC 3'
<i>SpeI</i> ₆	5' P-CTAGAAATCTAGAAATCGATAGATCGGAAGAGCACACGTCTGAACTCCAGTCACATCTCGTATGCCGCTTCTGCTTG 5' 3' TTAGATGCTTAGCTATCTAGCCTTCTCGTGTGCAGACTTGAGGTCAGTGTAGAGCATACGGCAGAAGACGAAC 3'

203

204

205

206

Supplementary Table 5: NGS runs and sample-specific adapter combinations from this study.

NGS run	SRA number	Library	Sample number	Time point (min)	Forward adapter	Reverse adapter
1	SAMN27867644	Lib _{random}	1	0	Ncol ₁	Spel ₁
			2	95	Ncol ₂	Spel ₂
			3	225	Ncol ₃	Spel ₃
			4	290	Ncol ₄	Spel ₄
			5	360	Ncol ₅	Spel ₅
			6	480	Ncol ₆	Spel ₆
		Lib _{comb1}	1	0	Ncol ₁	Spel ₂
			2	95	Ncol ₂	Spel ₃
			3	225	Ncol ₃	Spel ₄
			4	290	Ncol ₄	Spel ₅
			5	360	Ncol ₅	Spel ₆
			6	480	Ncol ₆	Spel ₁
		Lib _{comb2}	1	0	Ncol ₁	Spel ₃
			2	95	Ncol ₂	Spel ₄
			3	225	Ncol ₃	Spel ₅
			4	290	Ncol ₄	Spel ₆
			5	360	Ncol ₅	Spel ₁
			6	480	Ncol ₆	Spel ₂
Standard RBSs	1	0	Ncol ₁	Spel ₄		
	2	95	Ncol ₂	Spel ₅		
	3	225	Ncol ₃	Spel ₆		
	4	290	Ncol ₄	Spel ₁		
	5	360	Ncol ₅	Spel ₂		
	6	480	Ncol ₆	Spel ₃		
2	SAMN27867645	Lib _{fact} (replicate 1)	1	0	Ncol ₁	Spel ₁
			2	95	Ncol ₂	Spel ₂
			3	225	Ncol ₃	Spel ₃
			4	290	Ncol ₄	Spel ₄
			5	360	Ncol ₅	Spel ₅
			6	480	Ncol ₆	Spel ₆
		Lib _{fact} (replicate 2)	1	0	Ncol ₁	Spel ₆
			2	95	Ncol ₂	Spel ₅
			3	225	Ncol ₃	Spel ₄
			4	290	Ncol ₄	Spel ₃
			5	360	Ncol ₅	Spel ₂
			6	480	Ncol ₆	Spel ₁
3	SAMN27867646	Lib _{random} in WT strain with ptRNA ^{fMet-A37}	1	0	Ncol ₁	Spel ₁
			2	95	Ncol ₂	Spel ₂
			3	225	Ncol ₃	Spel ₃
			4	290	Ncol ₄	Spel ₄
			5	360	Ncol ₅	Spel ₅
			6	480	Ncol ₆	Spel ₆
		Lib _{random} in WT strain with ptRNA ^{fMet-A37G}	1	0	Ncol ₁	Spel ₂
			2	95	Ncol ₂	Spel ₃
			3	225	Ncol ₃	Spel ₄
			4	290	Ncol ₄	Spel ₅
			5	360	Ncol ₅	Spel ₆
			6	480	Ncol ₆	Spel ₁
		Lib _{random} in WT strain with ptRNA ^{fMet-A37U}	1	0	Ncol ₁	Spel ₃
			2	95	Ncol ₂	Spel ₄
			3	225	Ncol ₃	Spel ₅
			4	290	Ncol ₄	Spel ₆
			5	360	Ncol ₅	Spel ₁
			6	480	Ncol ₆	Spel ₂
			1	0	Ncol ₁	Spel ₄

3	SAMN27867646	Lib _{random} in $\Delta metZWV$ strain with ptRNA ^{fMet-A37}	2	95	Ncol ₂	Spel ₅
			3	225	Ncol ₃	Spel ₆
			4	290	Ncol ₄	Spel ₁
			5	360	Ncol ₅	Spel ₂
			6	480	Ncol ₆	Spel ₃
			1	0	Ncol ₁	Spel ₅
		Lib _{random} in $\Delta metZWV$ strain with ptRNA ^{fMet-A37G}	2	95	Ncol ₂	Spel ₆
			3	225	Ncol ₃	Spel ₁
			4	290	Ncol ₄	Spel ₂
			5	360	Ncol ₅	Spel ₃
			6	480	Ncol ₆	Spel ₄
			1	0	Ncol ₁	Spel ₆
		Lib _{random} in $\Delta metZWV$ strain with ptRNA ^{fMet-A37U}	2	95	Ncol ₂	Spel ₁
			3	225	Ncol ₃	Spel ₂
			4	290	Ncol ₄	Spel ₃
			5	360	Ncol ₅	Spel ₄
			6	480	Ncol ₆	Spel ₅

208

209

210 **Supplementary Table 6: Codon weight and frequencies for calculation of CAI and tAI.**211 w(CAI)/w(tAI): weight for the calculation of CAI (Sharp *et al.* (3))/tAI (dos Reis *et al.* (4)). *E. coli* fraction:212 fraction of each triplet amongst all triplets coding for the same amino acid in the *E. coli* genome213 (according to: <https://www.kazusa.or.jp/codon/>).

Triplet	Amino acid	w(CAI)	w(tAI)	Rel. codon frequency
GCA	Ala	0.586	0.375	0.270
GCC	Ala	0.122	0.250	0.260
GCG	Ala	0.424	0.120	0.250
GCU	Ala	1.000	0.110	0.220
AGA	Arg	0.004	0.125	0.130
AGG	Arg	0.002	0.165	0.070
CGA	Arg	0.004	0.000	0.090
CGC	Arg	0.356	0.360	0.260
CGG	Arg	0.004	0.125	0.150
CGU	Arg	1.000	0.500	0.300
AAC	Asn	1.000	0.500	0.410
AAU	Asn	0.051	0.220	0.590
GAC	Asp	1.000	0.375	0.350
GAU	Asp	0.434	0.165	0.650
UGC	Cys	1.000	0.125	0.480
UGU	Cys	0.500	0.055	0.520
CAA	Gln	0.124	0.250	0.350
CAG	Gln	1.000	0.330	0.650
GAA	Glu	1.000	0.500	0.640
GAG	Glu	0.259	0.160	0.360
GGA	Gly	0.010	0.125	0.190
GGC	Gly	0.724	0.500	0.290
GGG	Gly	0.019	0.165	0.180
GGU	Gly	1.000	0.220	0.340
CAC	His	1.000	0.125	0.370
CAU	His	0.291	0.055	0.630
AUA	Ile	0.003	0.163	0.210
AUC	Ile	1.000	0.375	0.310
AUU	Ile	0.185	0.165	0.470
CUA	Leu	0.007	0.125	0.060
CUC	Leu	0.037	0.125	0.100

CUG	Leu	1.000	0.540	0.380
CUU	Leu	0.042	0.055	0.150
UUA	Leu	0.020	0.125	0.180
UUG	Leu	0.020	0.165	0.130
AAA	Lys	1.000	0.750	0.710
AAG	Lys	0.253	0.240	0.290
AUG	Met	1.000	1.000	1.000
UUC	Phe	1.000	0.250	0.360
UUU	Phe	0.296	0.110	0.640
CCA	Pro	0.135	0.125	0.230
CCC	Pro	0.012	0.125	0.160
CCG	Pro	1.000	0.165	0.370
CCU	Pro	0.070	0.055	0.240
AGC	Ser	0.410	0.125	0.200
AGU	Ser	0.085	0.055	0.180
UCA	Ser	0.077	0.125	0.180
UCC	Ser	0.744	0.250	0.140
UCG	Ser	0.017	0.165	0.110
UCU	Ser	1.000	0.110	0.180
ACA	Thr	0.076	0.125	0.250
ACC	Thr	1.000	0.250	0.310
ACG	Thr	0.099	0.290	0.220
ACU	Thr	0.965	0.110	0.220
UGG	Trp	1.000	0.165	1.000
UAC	Tyr	1.000	0.375	0.350
UAU	Tyr	0.239	0.165	0.650
GUA	Val	0.495	0.625	0.190
GUC	Val	0.066	0.250	0.190
GUG	Val	0.221	0.200	0.290
GUU	Val	1.000	0.110	0.320

215 **Supplementary Table 7: Optimisation of hybridisation window.** Correlation between rTR and the
 216 mean hybridisation energy of different sequence windows with different lengths. The mean of the
 217 hybridisation energy at position -11 and position -10 (rank 1) was denoted hyb_{opt} . The data was ranked
 218 by the sum of R and ρ .

Rank	Window length (nt)	Window start (5'-UTR pos.)	Pearson's R	Spearman's ρ
1	2	-11	-0.298	-0.170
2	3	-11	-0.298	-0.167
3	4	-12	-0.299	-0.166
4	3	-12	-0.294	-0.165
5	5	-12	-0.295	-0.161
6	5	-13	-0.292	-0.160
7	6	-13	-0.293	-0.159
8	2	-10	-0.288	-0.162
9	4	-11	-0.289	-0.159
10	7	-13	-0.287	-0.152
11	2	-12	-0.280	-0.159
12	4	-13	-0.281	-0.156
13	6	-12	-0.285	-0.152
14	7	-14	-0.283	-0.151
15	8	-14	-0.281	-0.148
16	6	-14	-0.278	-0.150
17	3	-10	-0.275	-0.151
18	5	-11	-0.274	-0.147
19	8	-13	-0.274	-0.143
20	9	-14	-0.272	-0.140

219

220

221 **Supplementary Notes**

222

223 **Supplementary Note 1: Sequence of spectinomycin resistance gene (*specR*) including**
224 **constitutive promoter used for knockout of *metZ*WV. The start and stop codon of *specR* are**
225 **underlined.**

226 TGATCGGCACGTAAGAGGTTCCAAC^{TTT}CACCATAATGAAATAAGATCACTACCGGGCGTATTTTTTGGAGTTATC
227 GAGATTTTTCAGGAGCTAAGGAAGCTACATATGAGTGAAAAAGTGCCCGCCGAGATTTTCGGTGAAC^{ACT}TATCACAAG
228 CACTCAACGTCATCGGGGCCACTTGGAGTCGACGTTGCTGGCCGTGCATTTGTACGGCTCCGCAC^{TGG}ATGGCG
229 GATTGAAACCGTACAGTGATATTGATTTGCTGGTGA^{CTGT}AGCTGCACCGCTCAATGATGCCGTGCGGCAAGCCC
230 TGCTCGTCGATCTCTTGGAGTTTTAGCTTCCCCTGGCCAAAACAAGGCACTCCGCGCCTTGGAA^{GTG}ACCATCG
231 TCGTGCACAGTGACATCGTACCTTGGCGTTATCCGGCCAGGCGGAACTGCAGTTCGGAGAGTGGCAGCGCAAAG
232 ACATCCTTGCGGGCATCTTCGAGCCCCGCCACAACCGATTCTGACTTGGCGATTCTGCTAACAAAGGCAAAGCAAC
233 ATAGCGTCGTCTTGGCAGGTTTCCAGCAGCGAAGGATCTCTTCAGCTCAGTCCCAGAAAGCGATCTATTCAAGGCAC
234 TGGCCGATACTCTGAAGCTATGGA^{ACT}CGCCGCCAGATTGGGCGGGCGATGAGCGGAATGTAGTGC^TTACTTTGT
235 CTCGTATCTGGTACACCGCAGCAACCGGCAAGATCGCGCCAAAGGATGTTGCTGCCACTTGGGCAATGGCAGCT
236 TGCCAGCTCAACATCAGCCCATCTGTTGAATGCCAAGCGGGCTTATCTTGGGCAAGAAGAAGATTATTTGCCCG
237 CTCGTGCGGATCAGGTGGCGGCGCTCATTAAATTCGTGAAGTATGAAGCAGTTAAACTGC^{TTGGT}GCCAGCCAAT
238 AA
239

240 **Supplementary Note 2: Sequence of pASPIre4. The displayed sequence corresponds to the insert**
241 **between HindIII and PacI restriction sites (underlined) in pSEVA291. A graphical representation of**
242 **pASPIre4 is shown in Supplementary Figure 1.**

243 AAGCTTCACATCTGCAGTAATCGGCCGGCTTGTGACGACGGCGGTCTCCGTCGTCAGGATCATCCGGGCATCGC
244 TTAGTCACCTTTGGGCCACGGTCCGCTACCTTACAGGAATAGTACTCGTCCTTTAATTTGGAATGAACCATGGCA
245 GTCAGTTGTGTTGCGTTTTCTTCGACCTAGTACTCGCTCCCTTAGGAGAAAAGACAGATAGCTTCTTACCCGGGGTT
246 TGTACCGTACACCACTGAGACCGCGGTGGTTGACCAGACAAACCACGAAGTTCTGTTAAGTAACTGAACCCAAT
247 GTCGTTAGTGACGCTTACCTCTTAAGAGGTCACTGACCTAACAGGATCCCACCACAATTCAGCAAATTTGTGAACA
248 TCATCACGTTTCATCTTTCCCTGGTTGCCAATGGCCCATTTTCTGTGCTAAGGAGGTCGCGAATTCAGGCG
249 CTTTTTAGACTGGTCGTAATGAAGAGCTCAATAAATATTTAATTTATCTCAGAAAGGCTAAGACATGCGAGCACT
250 GGTGTTTATTTCGTCTGAGCCGTGTTACCGATGCAACCACTAGTCCGGAACGTCAGCTGGAAGCTGTCAGCAGCT
251 GTGTGCACAGCGTGGTTGGGATGTTGTTGGTGTGTCAGAGGATCTGGATGTTAGCGGTGCAGTTGATCCGTTTTGA
252 TCGTAAACGTCGTCCGAATCTGGCAGTTGGCTGGCATTTGAAGAACAGCCGTTTATGTTATGTTGCCTATCG
253 TGTTGATCGTCTGACCCGTAGCATTTCGTATCTGCAACAGCTGGTTCATTGGGCAGAAGATCATAAAAAACTGGT
254 TGTGAGCGCAACCGAACACATTTTATACCACCCCGTTTGCAGCAGTTGTTATTGCACTGATGGGCACCGT
255 TGCAGATGGAACCTGGAAGCAATTAAGAAGCTGATTCAGCGAGCCCAATTTAACATCTCGTGCAGTAAATA
256 TCGTGGTAGCCTGCCTCCGTGGGTTATCTGCCGACGCGTGTAGATGTTGAATGGCGTCTGGTTCCCTGACCCGTT
257 TCAGCGTGAACGTTATTCTGGAAGTATATCATCGTGTGGTGGATAATCATGAACCGCTGCATCTGGTTGCACATGA
258 TCTGAATCGTCTGGTGTCTGAGTCCCAAGATTATTTTTGCTCAGCTGCAAGGTCGTGAACCGCAGGGTTCGTGA
259 ATGGTCTGCAACCGCACTGAAACGTAGCATGATTAGCGAAGCAATGCTGGGTTATGCAACCCCTGAATGGTAAAAC
260 CGTTCGTGATGATGATGGTGCACCGCTGGTTCGTGCAGAACCAGATTCTGACACGTGAACAGCTGGAAGCACTGCG
261 TGCCGAACCTGGTTAAAACAGCCGTGCAAAACCGGCAGTTAGCACCCCGAGCCTGCTGCTGCGTGTCTGTTTTG
262 TGCAGTTTGTGGTGAACCGGCATACAAATTTGCCGGTGGTGGTTCGTAAACATCCGCGTTATCGTTGTCGTAGCAT
263 GGGTTTTCCGAAACATTGTGGTAATGGTACAGTTGCAATGGCAGAATGGGATGCATTTTGCGAAGAACAGGTTCT
264 GGATCTGCTGGGTGATGCCGAACGCTCTGGA^{AAA}AGTTTGGGTTGCAGGTAGCGATAGCGCAGTTGAACTGGCCGA
265 AGTTAATGCGGAACCTGGTCGATCTCACCAGTCTGATTGGAAGTCCCGCATATCGTGCGGGTAGTCCCTCAGCGTGA
266 AGCACTGGATGCACGATTTGCAGCACTGGCAGCACGTCGAAGAAGAACTGGAAGGTCGGAAGCACGTCAGGCGG
267 TTGGGAATGGCGTGAACAGGTCAGCGTTTTGGTGGTGGTGGCGTGAGCAGGATACCGCAGCAAAAAATACCTG
268 GCTGCGTAGTATGAATGTTTCGCTGACCTTTGATGTTTCGCGGTGGCCTGACCCGCACCATGATTTTGGCGATCT
269 GCAAGAATATGAACAGCATCTGCGTCTGGGTAGCGTTGTTGAACGCTCTGCATACCGGCATGAGCACCGGCGGTGG
270 CAGCGGCGGTTCTGGTGGCTCTAGCAAAGGAGAAGAACTTTTCACTGGAGTTGTCCCAATTTCTGTTGAATTAGA
271 TGGTGTATGTTAATGGGCACAAATTTTCTGTCCGTGGAGAGGGTGAAGGTGATGCTACAAACGGAAAAC^TCACCT
272 TAAATTTATTTGC^{ACT}ACTGGA^{AA}ACTAGCTGTTCCGTGGCCAACTGTTACTACTACTGACCTATGGTGTTC
273 ATGCTTTTCCCGTATCCGGATCACAACCGCATGACTTTTCAAGAGTGCCATGCCGAAGGTTATGTACA
274 GGAACGCACTATATCTTTCAAAGATGACGGGACCTACAAGACCGTGC^{TGA}AGTCAAGTTTGAAGGTGATACCCT
275 TGTTAATCGTATCGAGTTAAAGGTTATTGATTTTAAAGAAGATGGAAACATTTCTTGGACACAACTCGAGTACAA
276 CTTTAACTCACACAATGTATACATCACGGCAGACAAAACAAAAGAAATGGAATCAAAGCTAACTTCAAAATTCGCCA

277 CAACGTTGAAGATGGTTCCGTTCAACTAGCAGACCATTATCAACAAAATACTCCAATTGGCGATGGCCCTGTCCT
278 TTTACCAGACAACCATTACCTGTGCGACACAATCTGTCTTTTGAAGATCCCAACGAAAAGCGTGACCACATGGT
279 CCTTCTTGAGTTTGTAACTGCTGCTGGGATTACACATGGCATGGATGAACTTACAAAAGGCCGTGCTGCTAACGA
280 CGAAAACACTACGCTCTGGCTGCTTAATAAGCGGCCCGCGCCTAGGCGGCCCTCTGTGTGAAATTGTTATCCGCTTT
281 AATTAA
282

283 **Supplementary Note 3: Sequence of pASPIre4_{lib}.** The displayed sequence corresponds to the insert
284 between PstI and NotI restriction sites (underlined) in pASPIre4. A graphical representation of
285 pASPIre4_{lib} is shown in **Supplementary Figure 2.**

286 CTGCAGTAATCGGCCGGCTTGTGCGACGACGGCGGTCTCCGTCGTCAGGATCATCCGGGCATCGCTTAGTCACCTT
287 TGGGCCACGGTCCGCTACCTTACAGGAATAGTACTCGTCTTTAATTTGGAATGAACCATGGCAGTCAGTTGTGT
288 TGCGTTTCTTCGACCTAGTACTCGCTCCCTTAGGAGAAAAGACAGATAGCTTCTTACCCGGGTTTGTACCGTACA
289 CCACTGAGACCGCGGTGGTTGACCAGACAAACCAGAAAGTTCTGTAAAGTAACTGAACCCAATGTCGTTAGTGA
290 CGCTTACCTCTTAAGAGGTCAGTACCTAACAGGATCCACCACAATTCAGCAAATTTGTGAACATCATCACGTTT
291 ATCTTTCCCTGGTTGCCAATGGCCATTTTCTGTGAGTAAACGAGAAAGGTCGCGAATTCAGGCGCTTTTTAGACT
292 GGTGCGTAANNNNNNNNNNNNNNNNNNNNNNNNNNNATGCGNGCNCNTNGTNGTNATHCGNCNTNCNGNGTNACNGAY
293 GCNACNACTAGTCCGGAACGTCAGCTGGAAAGCTGTCAGCAGCTGTGTGCACAGCGTGGTTGGGATGTTGTTGGT
294 GTTGCGAGAGGATCTGGATGTTAGCGGTGCAGTTGATCCGTTTGTATCGTAAACGTCGTCGAAATCTGGCACGTTGG
295 CTGGCATTGGAAGAACAGCCGTTTGTATGTTATTGTTGCCTATCGTGTGATCGTCTGACCCGTAGCATTCGTCAT
296 CTGCAACAGCTGGTTCATTGGGCAGAAGATCATAAAAACTGGTTGTTGAGCGCAACCGAAGCACATTTTGATACC
297 ACCACCCCGTTTGCAGCAGTTGTTATTGCACTGATGGGCACCGTTGCACAGATGGAACGGAAGCAATTAAGAA
298 CGTAATCGTAGCGCAGCCATTTTAAACATTCGTGCAGGTAAATATCGTGGTAGCCTGCCCTCCGTGGGGTTATCTG
299 CCGACGCGTGTAGATGGTGAATGGCGTCTGGTTCCCTGACCCGGTTTCAGCGTGAACGTAATTCGGAAGTATATCAT
300 CGTGTGGTGGATAATCATGAACCGCTGCATCTGGTTGCACATGATCTGAATCGTGTGGTGTCTGAGTCCCAA
301 GATTATTTTGCTCAGCTGCAAGGTCGTGAACCGCAGGGTTCGTGAATGGTCTGCAACCGCACTGAAACGTAGCATG
302 ATTAGCGAAGCAATGCTGGGTTATGCAACCCGTAATGGTAAAACCGTTCGTGATGATGATGGTGCACCGCTGGTT
303 CGTGCAGAACCATTCTGACACGCTGAACAGCTGGAAGCAGTCCGTCGCGAATGTTAAAACCGCCGTGCAAAA
304 CCGGCAGTTAGCACCCCGAGCTGCTGCTGCTGTTCTGTTTGTGTCAGTTTGTGGTGAACCGGCATACAAATTT
305 GCCGGTGGTGGTCGTAAACATCCGCGTTATCGTTGTCGTAGCATGGGTTTTCCGAAACATTTGTGGTAATGGTACA
306 GTTGCAATGGCAGAATGGGATGCATTTTGCGAAGAACAGGTTCTGGATCTGCTGGGTGATGCCGAACGTCTGGAA
307 AAAGTTTGGGTTGCAGGTAGCGATAGCGCAGTTGAACTGGCCGAAGTTAATGCGGAACGGTTCGATCTCACCAGT
308 CTGATTGGAAGTCCCGCATATCGTGCGGGTAGTCTCAGCGTGAAGCACTGGATGCACGTATTGCAGCACTGGCA
309 GCACGTCAAGAAGAAGTGAAGGTTCTGGAAGCACGTCCGAGCGGTTGGGAATGGCGTGAACAGGTCAGCGTTTT
310 GGTGATTGGTGGCGTGAGCAGGATACCGCAGCAAAAAATACCTGGCTGCGTAGTATGAATGTTCCGCTGACCTTT
311 GATGTTCCGCGGTGGCCTGACCCGCACCATTTGATTTTGGCGATCTGCAAGAATATGAACAGCATCTGCGTCTGGGT
312 AGCGTTGTTGAACGTCGATACCGGCATGAGCACCGCGGTTGGCAGCGCGGTTCTGGTGGCTCTAGCAAAGGA
313 GAAGAACTTTTCACTGGAGTTGTCCCAATCTTGTGTAATTAGATGGTGTGTTAATGGGCACAAATTTTCTGTC
314 CGTGGAGAGGGTGAAGGTGATGCTACAAACGGAAAACCTCACCTTAAATTTATTTGCACTACTGGAACACTACCT
315 GTTCCGTGGCCAACACTTGTCACTACTCTGACCTATGGTGTTCATGCTTTTTCCCGTTATCCGGATCACATGAAA
316 CGGCATGACTTTTTCAAGAGTGCCATGCCCCGAAGGTTATGTACAGGAACGCACTATATCTTTCAAAGATGACGGG
317 ACCTACAAGACGCGTGTGAAGTCAAGTTTGAAGGTGATACCTTGTAAATCGTATCGAGTTAAAGGGTATTGAT
318 TTTAAAGAAGATGGAACATTTCTGGACACAACTCGAGTACAATTTAACTCACACAATGTATAACATCACGGCA
319 GACAAACAAAAGAATGGAATCAAAGCTAACTTCAAAATTCGCCACAACGTTGAAGATGGTTCCGTTCAACTAGCA
320 GACCATTATCAACAAAATACTCCAATTGGCGATGGCCCTGTCTTTTACCAGACAACCATTACCTGTCGACACAA
321 TCTGTCTTTCGAAAGATCCCAACGAAAAGCGTGACCACATGGTCTTCTTGAGTTTGTAACTGCTGCTGGGATT
322 ACACATGGCATGGATGAACTCTACAAAAGGCCGTGCTGCTAACGACGAAAACACTACGCTCTGGCTGCTTAATAAGCG
323 GCCGC
324

325 **Supplementary Note 4: Sequence of 5'-UTR half-library.** The displayed sequence corresponds to
326 the insert between PstI and NotI restriction sites (underlined) in pASPIre4. The BbsI restriction site for
327 scarless cloning of 5'-UTR-CDS combinations is marked in orange. A graphical representation is
328 available in **Supplementary Figure 4.**

329 CTGCAGTAATCGGCCGGCTTGTGCGACGACGGCGGTCTCCGTCGTCAGGATCATCCGGGCATCGCTTAGTCACCTT
330 TGGGCCACGGTCCGCTACCTTACAGGAATAGTACTCGTCTTTAATTTGGAATGAACCATGGCAGTCAGTTGTGT
331 TGCGTTTCTTCGACCTAGTACTCGCTCCCTTAGGAGAAAAGACAGATAGCTTCTTACCCGGGTTTGTACCGTACA
332 CCACTGAGACCGCGGTGGTTGACCAGACAAACCAGAAAGTTCTGTAAAGTAACTGAACCCAATGTCGTTAGTGA

333 CGCTTACCTCTTAAGAGGTCACCTGACCTAACAGGATCCCACCACAATTGAGCAAATTTGTGAACATCATCACGTT
334 ATCTTTCCCTGGTTGCCAATGGCCCATTTTCTGTGACGTAACGAGAAGGTCGCGAATTCAGGCGCTTTTTAGACT
335 GGTCGTAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNATGCGAGTCTTCGCGGCCG
336

337 **Supplementary Note 5: Sequence of CDS half-library.** The displayed sequence corresponds to the
338 insert between PstI and NotI restriction sites (underlined) in pASPIre4. The BbsI restriction site for
339 scarless cloning of 5'-UTR-CDS combinations is marked in orange. A graphical representation is
340 available in **Supplementary Figure 5**.

341 CTGCAGGAAGACCCATGCGNGCNCNTNGTNGTNATHCGNCTNTCNCNGTGNACNGAYGCNACNACTAGTCCGGAAC
342 GTCAGCTGGAAAGCTGTCAGCAGCTGTGTGCACAGCGTGGTTGGGATGTTGTTGGTGTGTCAGAGGATCTGGATG
343 TTAGCGGTGCAGTTGATCCGTTTGGATCGTAAACGTCGTCCGAATCTGGCACGTTGGCTGGCATTGGAAGAACAGC
344 CGTTTGATGTTATTGTTGCCTATCGTGTGATCGTCTGACCCGTAGCATTCGTCATCTGCAACAGCTGGTTCATT
345 GGGCAGAAGATCATAAAAAACTGGTTGTGAGCGCAACCGAAGCACATTTTGATACCACCACCCCGTTTGCAGCAG
346 TTGTTATTGCACTGATGGGCACCGTTGCACAGATGGAACCTGGAAGCAATTAAGAAGCGTAATCGTAGCGCAGCCC
347 ATTTTAAACATTCGTGCAGGTAATATCGTGGTAGCCTGCCTCCGTGGGGTTATCTGCCGACGCGTGTAGATGGTG
348 AATGGCGTCTGGTTCCTGACCCGGTTCAGCGTGAACGTATCTGGAAGTATATCATCGTGTGGTGGATAATCATG
349 AACCGCTGCATCTGGTTGCACATGATCTGAATCGTGTGGTGTCTGAGTCCCAAAGATTATTTGCTCAGCTGC
350 AAGGTGCGTAACCCGAGGGTCGTGAATGGTCTGCAACCGCATGAAACGTAGCATGATTAGCGAAGCAATGCTGG
351 GTTATGCAACCCGTAATGGTAAAACCGTTCGTGATGATGATGTTGACCGCTGGTTCGTGCGAAGCAATGCTGA
352 CACGTGAACAGCTGGAAGCACTGCGTGCCGAACCTGGTTAAAACCGCGTGCAAAACCGGCAGTTAGCACCCCGA
353 GCCTGCTGCTGCGTGTCTGTTTTGTGTCAGTTTTGTGGTGAACCGGCATACAAAATTTGCCGGTGGTGGTTCGTAAC
354 ATCCCGTTATCGTTGTGCTAGCATGGGTTTTCCGAAAACATTGTGGTAATGGTACAGTTGCAATGGCAGAATGGG
355 ATGCATTTTGCAGAAGAACAGGTTCTGGATCTGCTGGGTGATGCCGAACGCTCGGAAAAGTTTGGGTTGCAGGTA
356 GCGATAGCGCAGTTGAACTGGCCGAAGTTAATGCGGAACCTGGTCGATCTCACCAGTCTGATTGGAAGTCCCGCAT
357 ATCGTGCGGGTAGTCTCAGCGTGAAGCACTGGATGCACGTATTGCAGCACTGGCAGCACGTCAAGAAGAAGTGG
358 AAGGTCTGGAAGCACGTCAGCGGTTGGGAATGGCGTGAACAGGTCAGCGTTTTGGTATTGGTGGCGTGAGC
359 AGGATACCGCAGCAAAAATAACCTGGCTGCGTAGTATGAATGTTGCGCTGACCTTTGATGTTGCGGTTGGCCTGA
360 CCCGCACCATTGATTTTGGCGATCTGCAAGAATATGAACAGCATCTGCGTCTGGGTAGCGTTGTTGAACGCTGTC
361 ATACCGGCATGAGCACCGGCGGTGGCAGCGGCGGTTCTGGTGGCTCTAGCAAAGGAGAAGAACTTTTCACTGGAG
362 TTGTCCCAATTTCTGTTGAATTAGATGGTGTGTTAATGGGCACAAAATTTCTGTCCGTGGAGAGGGTGAAGGTG
363 ATGCTACAAACGGAAAACCTCACCTTAAATTTATTTGCACTACTGGAAAACCTACCTGTTCCGTGGCCAACACTTG
364 TCACTACTCTGACCTATGGTGTTCATGCTTTTCCCGTTATCCGGATCACATGAAACGGCATGACTTTTTTCAAGA
365 GTGCCATGCCCCAAGGTTATGTACAGGAACGCACTATATCTTTCAAAGATGACGGGACCTACAAGACGCGTGCTG
366 AAGTCAAGTTTGAAGGTGATACCCTTGTAAATCGTATCGAGTTAAAGGGTATTGATTTTAAAGAAGATGGAAACA
367 TTCTTGGACACAACTCGAGTACAACCTTAACTCACACAATGTATACATCACGGCAGACAAAACAAAAGAAATGGAA
368 TCAAAGCAACTTCAAATTCGCCACAACGTTGAAGATGGTTCCGTTCACTAGCAGACCATTATCAAACAAAATA
369 CTCCAATTGGCGATGGCCCTGTCTTTTACCAGACAACCAATTACCTGTGACACAATCTGTCTTTTGCAAAAGATC
370 CCAACGAAAAGCGTGACCACATGGTCTCTTGTAGTTTGTAACTGCTGCTGGGATTACACATGGCATGGATGAAC
371 TCTACAAAAGGCCTGCTGCTAACGACGAAAACCTACGCTCTGGCTGCTTAATAAGCGGCCG
372

373 **Supplementary Note 6: Sequence of ptRNA^{fMet} variants.** The displayed sequence corresponds to
374 the insert between KpnI and SpeI restriction sites (underlined) in pSEVA361. The mutated position 37
375 in tRNA^{fMet} is highlighted in bold. A graphical representation is available in **Supplementary Figure 6**.

376 GGTACCAAAATAACACCCTGCTTAATTAAGCGGATAACAATTTACACACAGGAGGCCGCCTAGGCCGCGGCCGCG
377 CGAATTCGAGCTCGGTACCAAAATAACACCCTGCTTAATTAAGCGGATAACAATTTACACACAGGAGGCCGCCTAGGCCGCGGCCGCG
378 ATTTTTATGCAAAATAAATGAGTTTTCATTTAATCATCTTTTATCGGAGACAGGAAGAGTTTAGTGTGTTTTTTG
379 TAAAATAATGCGCTTAAGGGAGAGCAGGAGAAGGCAAAAGTATTTCAACAAATGAAAGTGAACCTGGATATTCATTC
380 ACATGATTAGCAATAAACGTTGACAAAATGTGGCGTGGATCACTATAATGCCTGCAGATTTTACGTCCCGTCTCG
381 GTACACCAATCCCAGCAGTATTTGCATTTTTTACCAAAACGAGTAGAATTTGCCACGTTTCAGGCGCGGGGTG
382 GAGCAGCCTGGTAGCTCGTGGGCTCAT**N**ACCCGAAGGTCGTGCGTTCAAATCCGGCCCCCGCAACCACTTTCCC
383 TTAGAGTCCTTTTTCAAATATACTGTGAAGACTTCGGCCTTCGTAGTGGGATTTGAAAAATCCTTCTGGAAAGT
384 GCTCCAGACCGCAGTTGCGGTTATAGGGTTCAGTTATATAAAGCCCCGATTTATCGGGTTTTTTGTTATCTGAC
385 TACAGAATAACTGGGCTTTAGGCCCTTTTTTTATGTCTGGGGTGGGCACTAGT
386

387

388 **References (Supplementary Information)**

- 389 1. Hollerer, S., Papaxanthos, L., Gumpinger, A.C., Fischer, K., Beisel, C., Borgwardt, K.,
390 Benenson, Y. and Jeschek, M. (2020) Large-scale DNA-based phenotypic recording and deep
391 learning enable highly accurate sequence-function mapping. *Nat Commun*, **11**, 3551.
- 392 2. Silva-Rocha, R., Martinez-Garcia, E., Calles, B., Chavarria, M., Arce-Rodriguez, A., de Las
393 Heras, A., Paez-Espino, A.D., Durante-Rodriguez, G., Kim, J., Nickel, P.I. *et al.* (2013) The
394 Standard European Vector Architecture (SEVA): a coherent platform for the analysis and
395 deployment of complex prokaryotic phenotypes. *Nucleic Acids Res*, **41**, D666-675.
- 396 3. Sharp, P.M. and Li, W.H. (1987) The codon Adaptation Index--a measure of directional
397 synonymous codon usage bias, and its potential applications. *Nucleic Acids Res*, **15**, 1281-
398 1295.
- 399 4. dos Reis, M., Wernisch, L. and Savva, R. (2003) Unexpected correlations between gene
400 expression and codon usage bias from microarray data for the whole Escherichia coli K-12
401 genome. *Nucleic Acids Res*, **31**, 6976-6985.
- 402