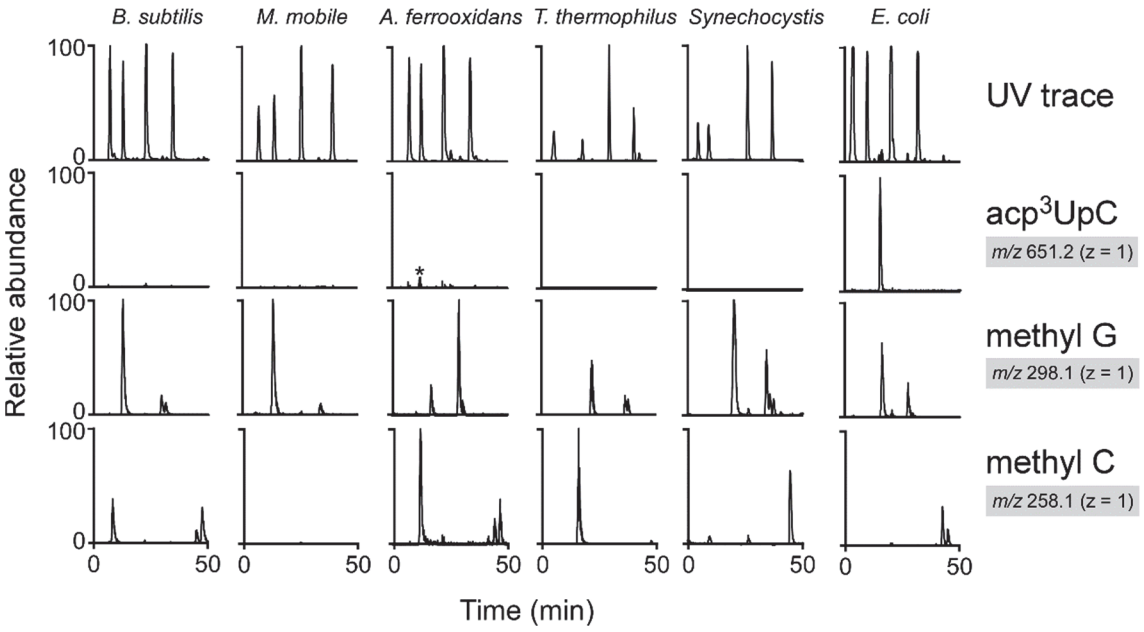
Supplementary information

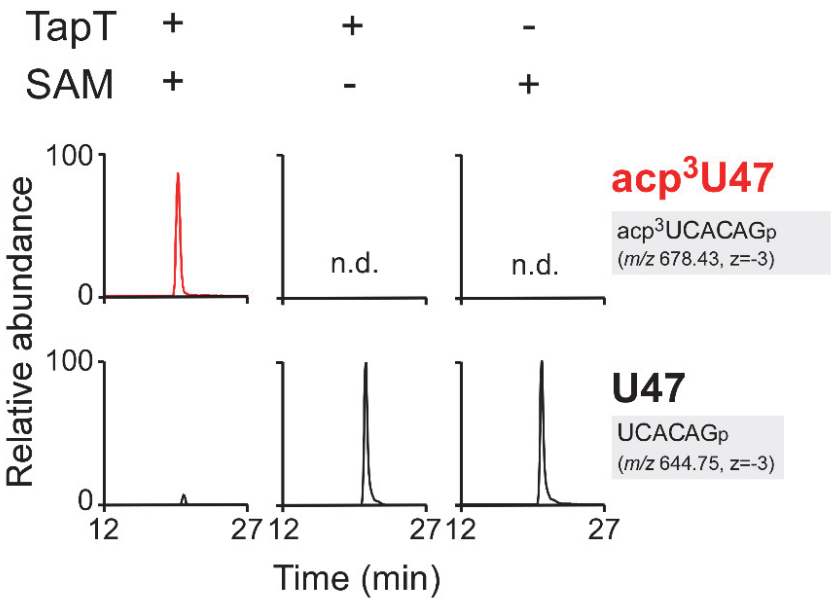
Biogenesis and functions of aminocarboxypropyluridine in tRNA

Takakura et al.



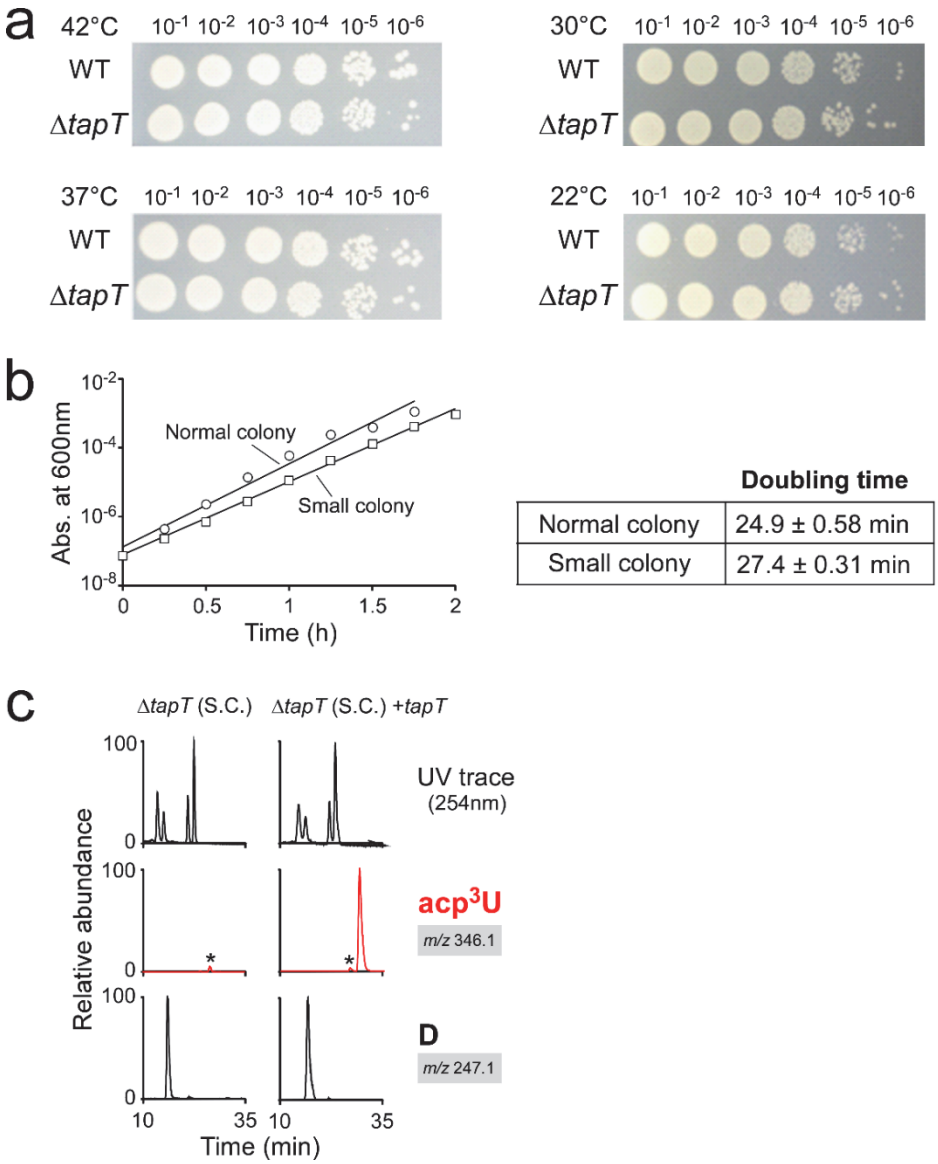
Supplementary Figure 1. Mass spectrometric nucleoside analyses of tRNA modifications in various bacteria.

Top panels represent UV trace. Second, third and fourth panels show XICs of the proton adducts of acp^{3}UpC, methylguanosines and methylcytidines, respectively. Asterisk indicates nonspecific peak with the same *m/z* value.



Supplementary Figure 2. *In vitro* reconstitution of acp^{3}U47 in tRNA^{Met} transcript by TapT.

In vitro acp^{3}U formation was performed in the presence or absence of TapT and SAM. XICs show negative ions of the RNase T_1-digested fragments with (upper panels) or without (lower panels) acp^{3}U47.



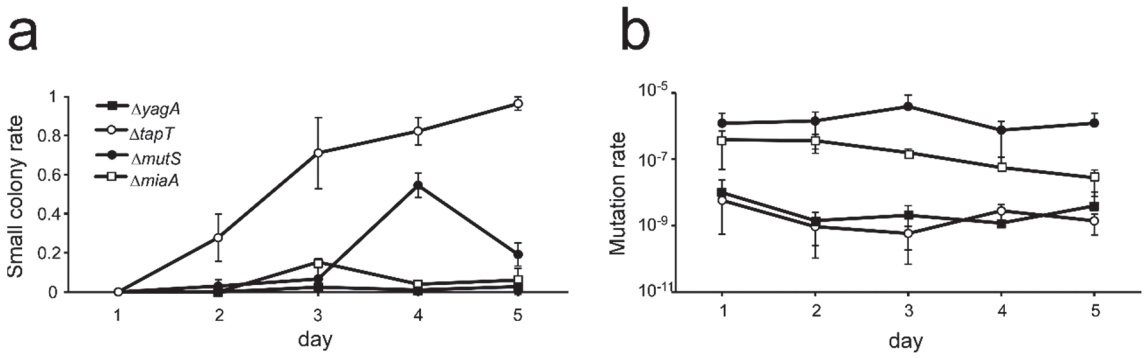
	Doubling time
Normal colony	24.9 ± 0.58 min
Small colony	27.4 ± 0.31 min

Supplementary Figure 3. Growth and morphological properties of Δ*tapT* strain.

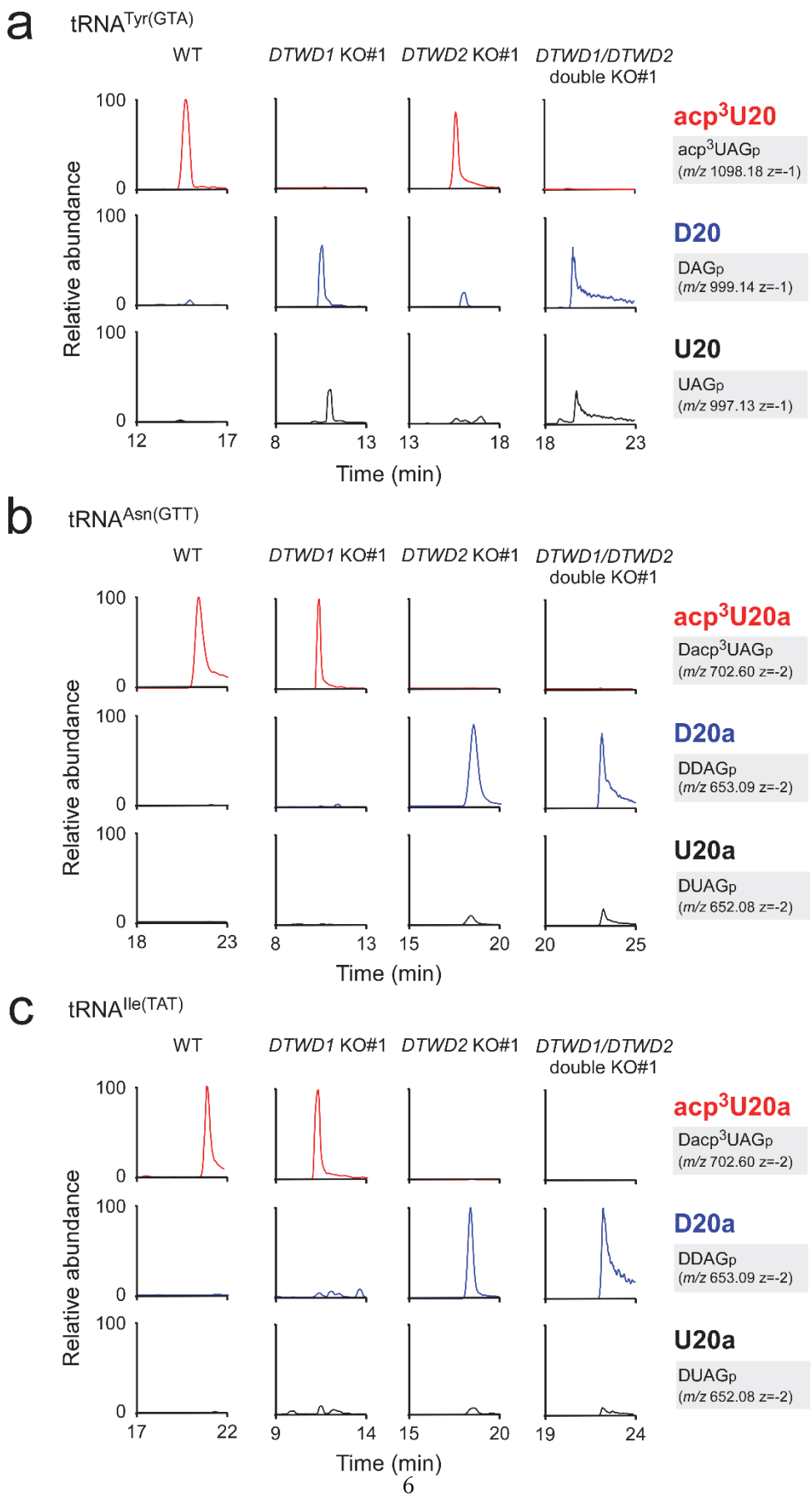
(a) Overnight full-growth cultures were serially diluted (A_{600}, 10^{-1} to 10^{-6}), spotted onto LB plates, and cultivated at 42°C, 37°C, 30°C, or 22°C.

(b) Growth curves of small colony-derived cells (square) and the normal colony-derived cells (circle) cultured at 37°C in LB medium. The doubling time of each strain represents average value of biological triplicates, with s.d. Source data are provided as a Source Data file.

(c) Mass spectrometric nucleoside analyses of tRNA modifications from Δ*tapT* (left panels), and Δ*tapT* rescued by plasmid-encoded *tapT* (right panels). Top panels are UV trace at 254 nm. XICs show the proton adducts of nucleosides for acp^3U (middle) and D (bottom). S.C. stands for small colony. Asterisks indicate nonspecific peaks with the same *m/z* value.

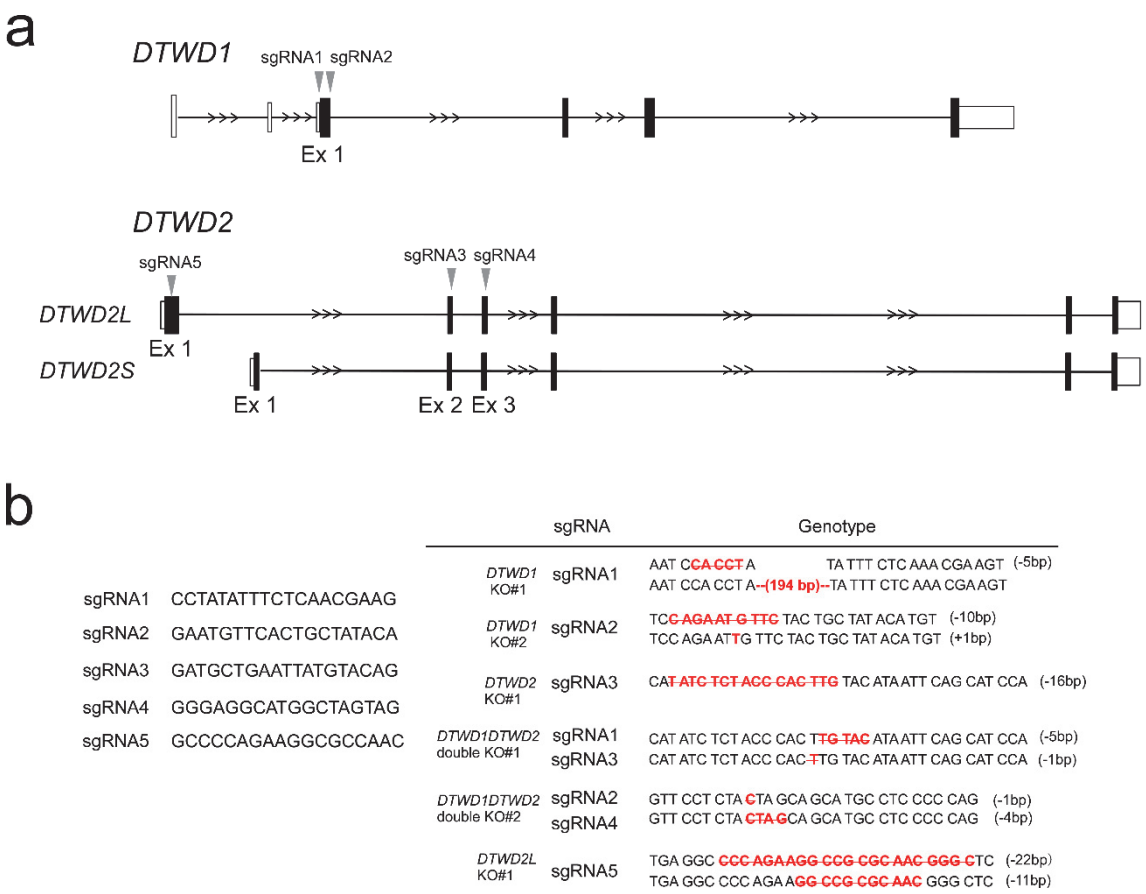


Supplementary Figure 4. Mutation rate does not increase in Δ*tapT* strain.
Frequencies of small colony appearance (a) and mutation ratio (b) in Δ*yagA* (closed square), Δ*tapT* (open circle), Δ*mutS* (closed circle), and Δ*miaA* (open square) cultivated continuously at 42°C. Data represent average values of biological triplicates, with s.d. Source data are provided as a Source Data file.


a
tRNA^Tyr(GTA)
WT
DTWD1 KO#1
DTWD2 KO#1
DTWD1/DTWD2 double KO#1
Relative abundance
100
0
acp3U20
acp3UAGp
(m/z 1098.18 z=-1)
D20
DAGp
(m/z 999.14 z=-1)
U20
UAGp
(m/z 997.13 z=-1)
12
17
8
13
13
18
18
23
Time (min)
b
tRNA^Asn(GTT)
acp3U20a
Dacp3UAGp
(m/z 702.60 z=-2)
D20a
DDAGp
(m/z 653.09 z=-2)
U20a
DUAGp
(m/z 652.08 z=-2)
18
23
8
13
15
20
20
25
Time (min)
c
tRNA^Ile(TAT)
acp3U20a
Dacp3UAGp
(m/z 702.60 z=-2)
D20a
DDAGp
(m/z 653.09 z=-2)
U20a
DUAGp
(m/z 652.08 z=-2)
17
22
9
14
15
20
19
24
Time (min)

Supplementary Figure 5. Human *DTWD1* and *DTWD2* are responsible for acp^{3}U formation.

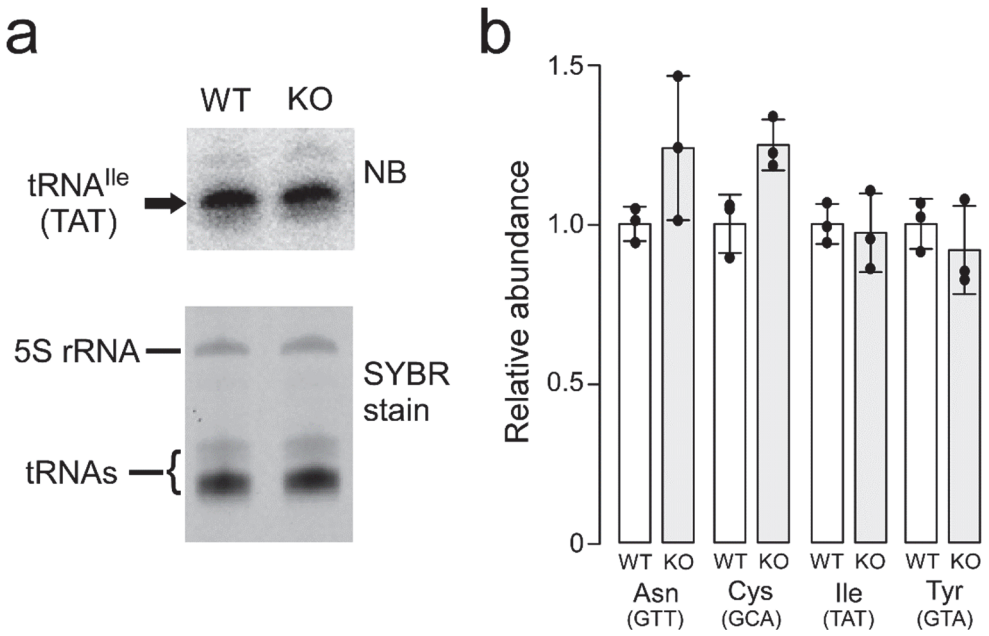
Mass spectrometric analyses of acp^{3}U20 in tRNA$^{Tyr(GTA)}$ (a), and acp^{3}U20a in tRNA$^{Asn(GTT)}$ (b) and tRNA$^{Ile(TAT)}$ (c) isolated from WT (left), *DTWD1* KO (left middle), *DTWD2* KO (right middle), *DTWD1*/*DTWD2* double-KO (right) strains. XICs show corresponding negative ions of the RNase T_1-digested fragments, as indicated on the right-hand side in each chromatogram. Mass spectra of D-containing fragments overlap with the isotopic peaks of U-containing fragments. The peak intensities of these fragments are normalized in consideration of those isotopes.



Supplementary Figure 6. Construction of human *DTWD1* and *DTWD2* knockout strains.

(a) Diagram of the human *DTWD1* and *DTWD2* genes, with the sites targeted by sgRNAs.

(b) Sequences of sgRNAs that target human *DTWD1* and *DTWD2* genes (left panel). Genotypes of human *DTWD1* and *DTWD2* KO strains (right panel). The indels (insertions and deletions) in each strain are shown in red letters.

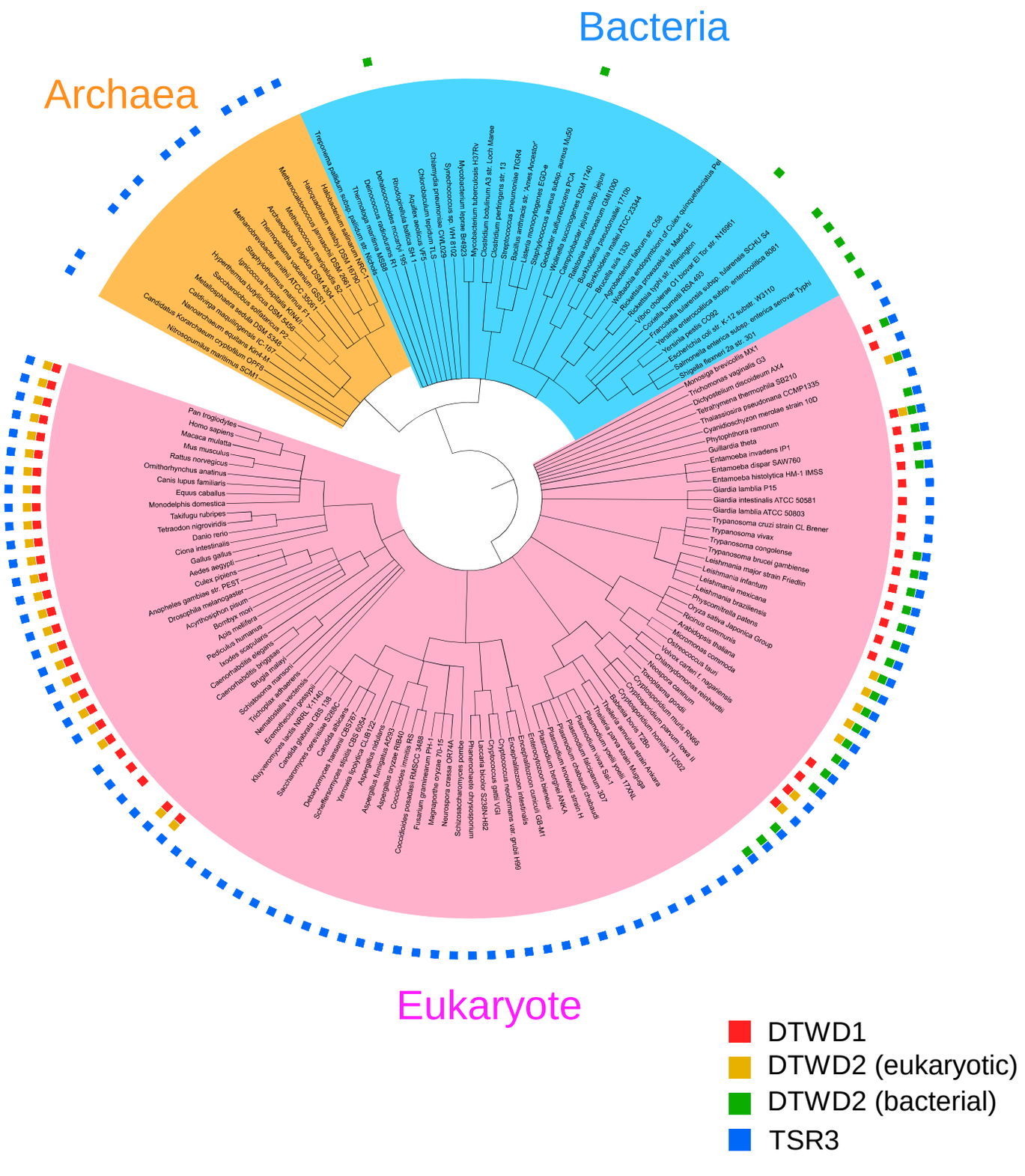


Supplementary Figure 7. Steady-state levels of cytoplasmic tRNAs in the *DTWD1*/*DTWD2* double-KO strain.

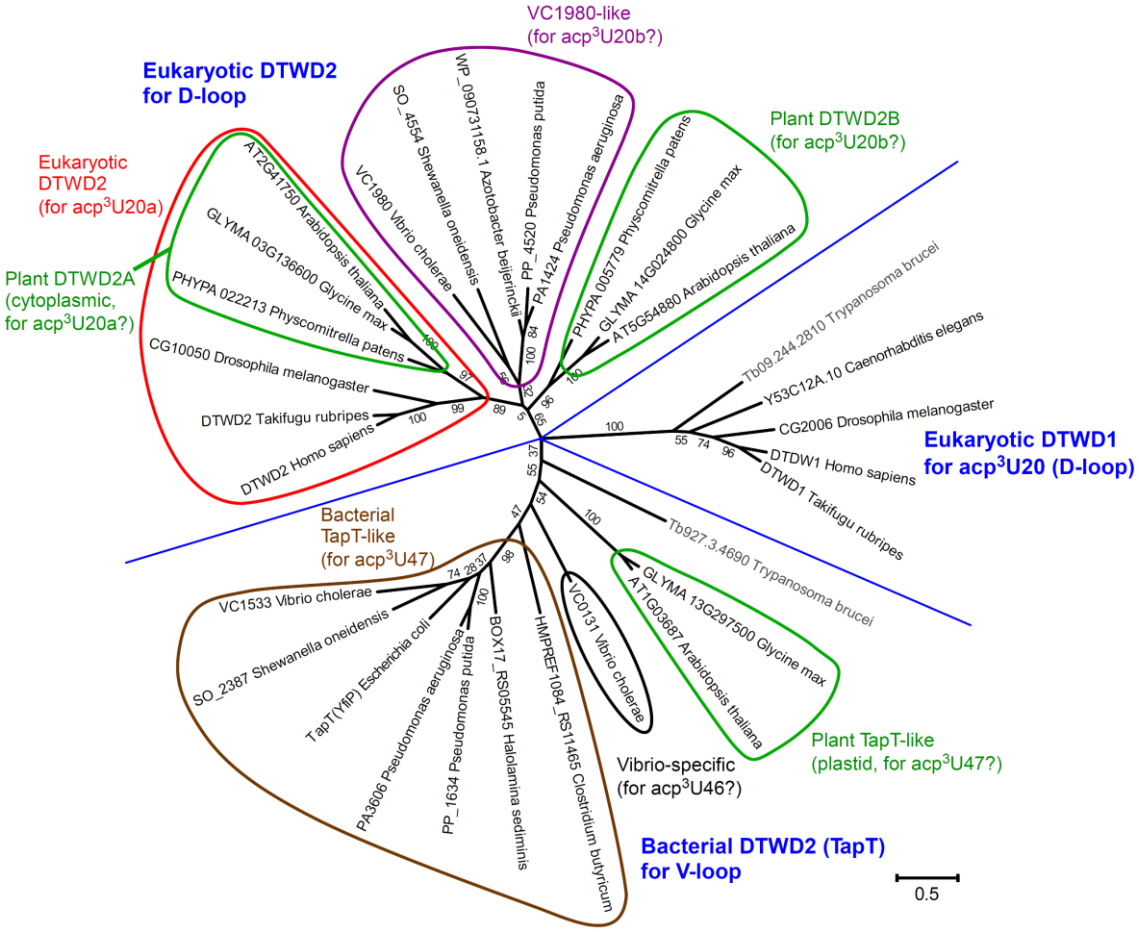
(a) Northern blotting (upper panel) of cytoplasmic $tRNA^{Ile(TAT)}$ in total RNAs of WT HEK293T and the double-KO strain. The lower panel shows the gel stained by SYBR Gold. Source data are provided as a Source Data file.

(b) Relative abundance of steady-state levels of tRNAs bearing acp^3U modification. The band intensity of the respective tRNA in the Northern blotting was normalized against the intensity of 5S rRNA stained in the gel. Data represent average values of biological triplicates, with s.d. Source data are provided as a Source Data file.

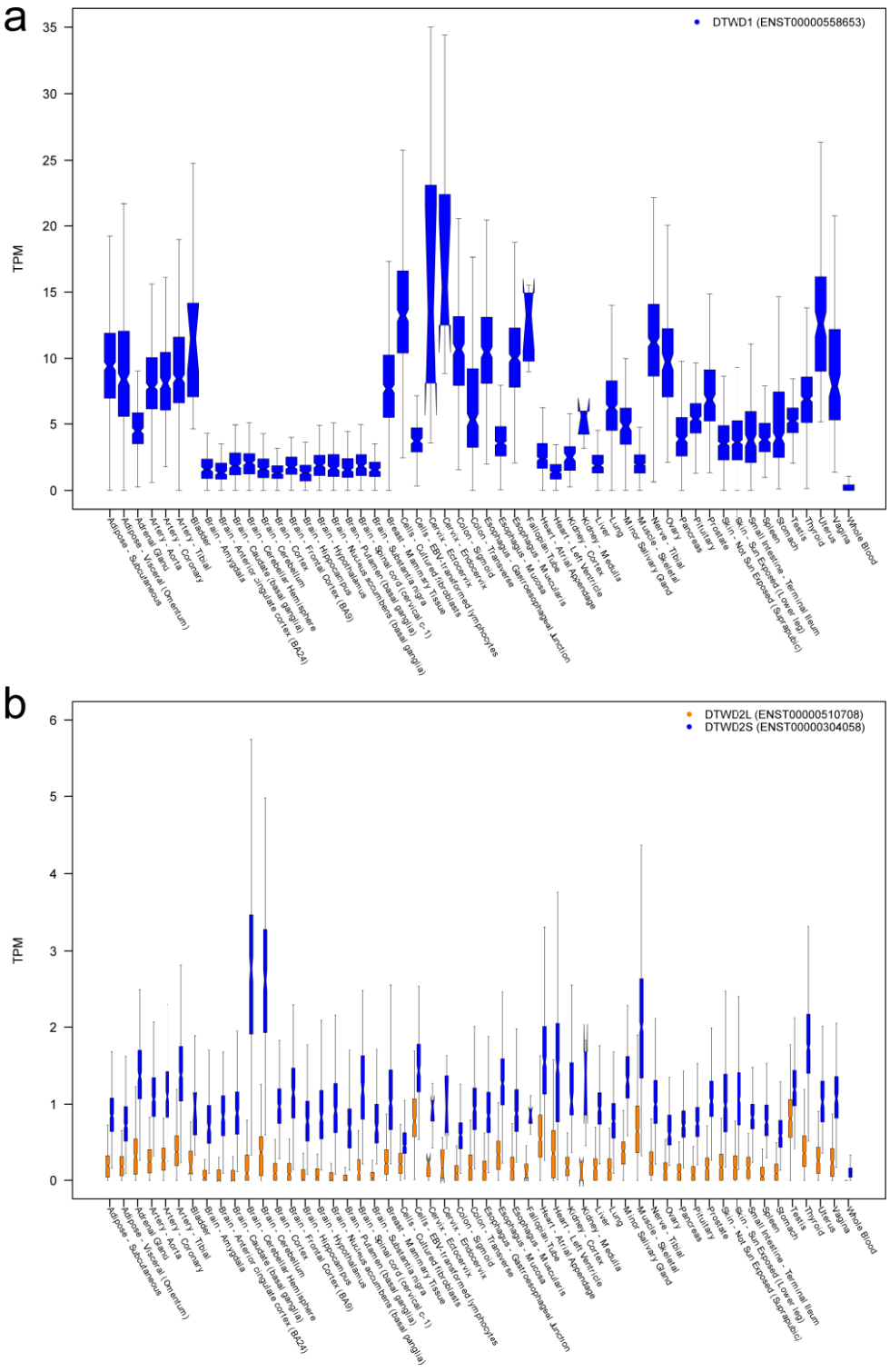
a


Bacteria
Archaea
Eukaryote
Treponema pallidum subsp. pallidum str. Nichols
Halobacterium salinarum NRC-1
Thermotoga maritima MSB8
Deinococcus radiodurans R1
Rhodopirellula baltica SH 1
Aquifex aeolicus VF5
Chlorobaculum tepidum TLS
Chlamydia pneumoniae CWL029
Synechococcus sp. WH 8102
Mycobacterium leprae Br4923
Mycobacterium tuberculosis H37Rv
Clostridium botulinum A3 str. Loch Maree
Clostridium perfringens str. 13
Streptococcus pneumoniae TIGR4
Bacillus anthracis str. 'Ames Ancestor'
Listeria monocytogenes EGD-e
Staphylococcus aureus subsp. aureus Mu50
Gloebacter sulfureducens PCA
Wolinella succinogenes DSM 1740
Campylobacter jejuni subsp. jejuni
Ralstonia solanacearum GMI1000
Burkholderia pseudomallei 1710b
Brucella suis 1330
Agrobacterium fabrum str. C58
Wolbachia endosymbiont of Culex quinquefasciatus Pel
Rickettsia prowazekii str. Madrid E
Rickettsia typhi str. Wilmington
Vibrio cholerae O1 biovar El Tor str. N16961
Coxiella burnetii RSA 493
Francisella tularensis subsp. tularensis SCHU S4
Yersinia enterocolitica subsp. enterocolitica 8081
Yersinia pestis CO92
Escherichia coli str. K-12 substr. W3110
Salmonella enterica subsp. enterica serovar Typhi
Escherichia fergusonii ATCC 35469
Shigella flexneri 2a str. 301
Monosiga brevicollis MX1
Monosiga ovata
Trichomonas vaginalis G3
Dictyostelium discoideum AX4
Tetrahymena thermophila SB210
Thalassiosira pseudonana CCMP1335
Phytophthora ramorum
Cyanidioschyzon merolae strain 10D
Guillardia theta
Entamoeba invadens IP1
Entamoeba dispar SAW760
Entamoeba histolytica HM-1 IMSS
Giardia lamblia P15
Giardia intestinalis ATCC 50581
Giardia lamblia ATCC 50803
Trypanosoma cruzi strain CL Brener
Trypanosoma vivax
Trypanosoma congolense
Trypanosoma brucei gambiense
Leishmania major strain Friedlin
Leishmania infantum
Leishmania mexicana
Leishmania braziliensis
Physcomitrella patens
Oryza sativa Japonica Group
Ricinus communis
Arabidopsis thaliana
Micromonas commoda
Ostreococcus tauri
Volvox carteri f. nagariensis
Chlamydomonas reinhardtii
Toxoplasma gondii
Neospora caninum
Cryptosporidium muris RN66
Cryptosporidium parvum Iowa II
Babesia bovis T2Bo
Theileria annulata strain Ankara
Theileria parva strain Muguga
Plasmodium yoelii yoelii 17XNL
Plasmodium falciparum 3D7
Plasmodium chabaudi chabaudi
Plasmodium knowlesi strain H
Plasmodium berghei ANKA
Plasmodium vivax Sal-1
Encephalitozoon cuniculi GB-M1
Encephalitozoon intestinalis
Cryptococcus neoformans var. grubii H99
Cryptococcus gattii WM276
Cryptococcus gattii VGI
Laccaria bicolor S238N-H82
Phanerochaete chrysosporium
Schizosaccharomyces pombe
Neurospora crassa OR74A
Magnaporthe oryzae 70-15
Fusarium graminearum PH-1
Coccidioides posadasii RMSCC 3488
Coccidioides immitis RS
Aspergillus oryzae RIB40
Aspergillus fumigatus A1163
Aspergillus nidulans
Aspergillus clavatus NRRL 1
Yarrowia lipolytica CLIB122
Scheffersomyces stipitis CBS 6054
Debaryomyces hansenii CBS767
Candida albicans
Candida glabrata CBS 138
Saccharomyces cerevisiae S288C
Kluyveromyces lactis NRRL Y-1140
Eremothecium gossypii
Nematostella vectensis
Trichoplax adhaerens
Schistosoma mansoni
Brugia malayi
Caenorhabditis briggsae
Caenorhabditis elegans
Ixodes scapularis
Pediculus humanus
Acyrthosiphon pisum
Bombyx mori
Apis mellifera
Drosophila melanogaster
Anopheles gambiae str. PEST
Culex pipiens
Aedes aegypti
Gallus gallus
Ciona intestinalis
Danio rerio
Tetraodon nigroviridis
Takifugu rubripes
Monodelphis domestica
Equus caballus
Canis lupus familiaris
Ornithorhynchus anatinus
Rattus norvegicus
Mus musculus
Macaca mulatta
Homo sapiens
Pan troglodytes
Nitrosopumilus maritimus SCM1
Candidatus Korarchaeum cryptofilum OPF8
Nanoarchaeum equitans Kin4-M
Caldivirga maquilingensis IC-167
Metallosphaera sedula DSM 5348
Sulfolobus solfataricus P2
Saccharolobus solfataricus 98/2
Hyperthermus butylicus DSM 5456
Staphylothermus marinus F1
Ignicoccus hospitalis KIN4/I
Thermofilum pendens Hrk 5
Methanobrevibacter smithii ATCC 35061
Archaeoglobus fulgidus DSM 4304
Methanococcus maripaludis S2
Methanococcus vannielii SB
Halogeometricum borinquense DSM 11551
Haloquadratum walsbyi DSM 16790
Halobacterium salinarum R1
DTWD1
DTWD2 (eukaryotic)
DTWD2 (bacterial)
TSR3

b



Supplementary Figure 8. Phylogenetic analyses of DTW-containing proteins.
(a) Phylogenetic distribution of TDD superfamily proteins. Organisms and TDD superfamily protein sequences were obtained from OrthoMCL database[1]. A phylogenetic tree of the organisms was built using the common tree tool in NCBI taxonomy database and drawn by the iTOL v4 server[2]. DTWD2 proteins are classified into two subfamilies, eukaryotic and bacterial (TapT), based on the phylogenetic tree of DTWD2 family proteins. The presence of TDD superfamily proteins in each organism was manually checked by NCBI BLAST.
(b) Phylogenetic tree of DTW domain-containing proteins. Multiple alignment of DTW superfamily proteins was constructed using MAFFT. The maximum likelihood tree was constructed using MEGA software[3] with a bootstrap test.



Supplementary Figure 9. Expression profiles of *DTWD1*, *DTWD2L*, and *DTWD2S* in human tissues.

Transcript per million (TPM) of *DTWD1* (a), *DTWD2L*, and *DTWD2S* (b) from various human tissues were obtained from GTEx Analysis V8 [4] and plotted as notched boxplot.

Supplementary References

1 Chen, F., Mackey, A. J., Stoeckert, C. J., Jr. & Roos, D. S. OrthoMCL-DB: querying a comprehensive multi-species collection of ortholog groups. *Nucleic Acids Res* 34, D363-368, doi:10.1093/nar/gkj123 (2006).

2 Letunic, I. & Bork, P. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res* 47, W256-W259, doi:10.1093/nar/gkz239 (2019).

3 Kumar, S., Stecher, G., Li, M., Knyaz, C. & Tamura, K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Molecular biology and evolution* 35, 1547-1549, doi:10.1093/molbev/msy096 (2018).

4 Carithers, L. J. *et al.* A Novel Approach to High-Quality Postmortem Tissue Procurement: The GTEx Project. *Biopreservation and biobanking* 13, 311-319, doi:10.1089/bio.2015.0032 (2015).

Supplementary Table 1. List of primers and probes

Primers

yfiPCDS_Fw	GTGATCTGCATATGACCGAAAACGCTGTTCT
yfiPCDS_Rv	GTGATCTGCTCGAGTTAAACGCTTTCTAACTGTTCTGC
Met_transcript_Fw	GCTAATACGACTCACTATA GGCUACGUAGCUC
Met_transcript_Rv	TGGTGGCTACGACGGGATTCGAACCCG
Met_transcript_body	GGCTACGTAGCTCAGTTGGTTAGAGCACATCACTCATAATGATGGGGTCACAGGTTCGAATC
DTWD1CDS_Fw	GTGATCTGGATATC ATGTCTCTCAATCCACCTATATTTC
DTWD1CDS_Rv	GTGATCTGGCGGCCGC ATGTGTAAGTTTTCCTGTTTCC
DTWD2Lnest_Fw	TTCAATGCCACGGCCTGAC
DTWD2Lnest_Rv	CCTGCTTTTCAGTGAGCCAGTG
DTWD2LCDS_Fw	GTGATCTGGATATC ATGGAGTCGCAGAAAGAG
DTWD2LCDS_Rv	GCGAGCTGGCGGCCGCAATTTTAACACTATTCATTAAC
DTWD2S_Fw	CGGCCTCAGAAAGTGTGTTTG
DTWD2S_Rv	GATATCTGCAGAATTCCACCACAC
sgRNA1_Fw	CACCGCTTCGTTTGAGAAATATAGG
sgRNA1_Rv	AAACCCTATATTTCTCAAACGAAGC
PCR1_Fw	AGTGTGCCAATTTGGTGGTT
PCR1_Rv	TCCCACTTTGCTGAGCTTTT
sgRNA2_Fw	CACCGTGTATAGCAGTAGAACATTC
sgRNA2_Rv	AAACGAATGTTCTACTGCTATACAC
PCR2_Fw	CTTCCATAGCTTCAGAAGATCCCC
PCR2_Rv	CAGAAGTAAATGTGCCCTGCTGAG
sgRNA3_Fw	CACCGATGCTGAATTATGTACAAG
sgRNA3_Rv	AAACCTTGTACATAATTCAGCATC
PCR3_Fw	CTTCGTGTTTCCTCTTGTGGACTC
PCR3_Rv	CCTCTTTGTATAAATGCAGTGGGATTAAGGTC
sgRNA4_Fw	CACCGGGAGGCATGCTGCTAGTAG
sgRNA4_Rv	AAACCTACTAGCAGCATGCCTCCC
PCR4_Fw	TGCCCTAACAGTTAATTTACCTTTCTTCACTGAAGC
PCR4_Rv	CTGTGGTCTTTAGTAAAAGACTTGTGAAAGTTATGC
sgRNA5_Fw	CACCGCCCCAGAAGGCCGCGCAAC
sgRNA5_Rv	AAACGTTGCGCGGCCTTCTGGGGC
PCR5_Fw	TGCACTCAGGCCTCCGCT
PCR5_Rv	TGCGGGCATTGCGCTACG

Probes for RCC and Northern blotting

E. coli Met	TGGTGGCTACGACGGGATTCGAACCTGTGA
human Asn-GTT-2	TGGCGTCCCTGGGTGGGCTCGAACCACCAACCTTTCGGTT
human Cys-GCA-2	TGGAGGGGGCACCCGGATTTGAACCGGGGACCTCTTGATC
human Ile-AAT-1	TGGTGGCCCGTACGGGGATCGAACCCGCGACCTTGGCGTT
human Ile-TAT-2	TGGTGCTCCAGGTGAGGCTCGAACTCACAACCTCGGCATT
human Tyr-GTA-2	TGGTCCTTCGAGCCGGAATCGAACCAGCGACCTAAGGATC