Article

A swapped genetic code prevents viral infections and gene transfer

https://doi.org/10.1038/s41586-023-05824-z

Received: 8 July 2022

Accepted: 10 February 2023

Published online: 15 March 2023

Check for updates

Akos Nyerges[1]✉, Svenja Vinke[1], Regan Flynn[1], Siân V. Owen[2], Eleanor A. Rand[2], Bogdan Budnik[3], Eric Keen[4,5], Kamesh Narasimhan[1], Jorge A. Marchand[1,7], Maximilien Baas-Thomas[1], Min Liu[6], Kangming Chen[6], Anush Chiappino-Pepe[1], Fangxiang Hu[6], Michael Baym[2] & George M. Church[1,3]✉

Engineering the genetic code of an organism has been proposed to provide a firewall from natural ecosystems by preventing viral infections and gene transfer[1–6]. However, numerous viruses and mobile genetic elements encode parts of the translational apparatus[7–9], potentially rendering a genetic-code-based firewall ineffective. Here we show that such mobile transfer RNAs (tRNAs) enable gene transfer and allow viral replication in *Escherichia coli* despite the genome-wide removal of 3 of the 64 codons and the previously essential cognate tRNA and release factor genes. We then establish a genetic firewall by discovering viral tRNAs that provide exceptionally efficient codon reassignment allowing us to develop cells bearing an amino acid-swapped genetic code that reassigns two of the six serine codons to leucine during translation. This amino acid-swapped genetic code renders cells resistant to viral infections by mistranslating viral proteomes and prevents the escape of synthetic genetic information by engineered reliance on serine codons to produce leucine-requiring proteins. As these cells may have a selective advantage over wild organisms due to virus resistance, we also repurpose a third codon to biocontain this virus-resistant host through dependence on an amino acid not found in nature[10]. Our results may provide the basis for a general strategy to make any organism safely resistant to all natural viruses and prevent genetic information flow into and out of genetically modified organisms.

The genetic code allows organisms to exchange functions through horizontal gene transfer (HGT) and enables recombinant gene expression in heterologous hosts. However, the shared language of the same code permits the undesired spread of antibiotic-, herbicide- and pesticide-resistance genes and allows viruses to cause diseases. By exploiting the shared nature of the genetic code, recombinant DNA technologies revolutionized our ability to produce small molecules, peptides, biologics and enzymes in vast quantities; however, the production cell cultures remained susceptible to viral contamination. Viral contamination in cell cultures remains a real risk with severe consequences: over the past four decades, dozens of viral contamination cases were documented in industry[11–13]. HGT also threatens the safe use of genetically modified organisms (GMOs) by enabling the spread of their engineered genetic information into natural ecosystems. Despite the impact of viral infections and HGT and the growing economic and societal role of GMOs and recombinant DNA, so far, no technology exists that could prevent viral infections and the escape of engineered genetic information from genetically modified biological systems. It is widely believed that genomically recoded organisms, whose genomes have been systematically redesigned to confer an alternative genetic code, would offer genetic isolation from natural ecosystems by obstructing the translation of horizontally transferred genetic material[1–5,14], including resistance to both viral infections and HGT. Indeed, the genome-wide removal of TAG stop codons and release factor 1 (RF1) from *E. coli*, which abolishes the ability of cells to terminate translation at TAG stop codons, provides substantial but not complete resistance to bacteriophages[2,3]. Most recently, a strain of *E. coli*, Syn61Δ3, was created with a synthetic recoded genome in which all annotated instances of two serine codons, TCG and TCA (together TCR), and the TAG stop codon were replaced with synonymous alternatives, and the corresponding serine tRNA genes (*serU* and *serT*) and RF1 (*prfA*) have been deleted[4,15]. Although the compressed genetic code of Syn61Δ3 provided resistance to five viruses[4], it could not prevent the escape of its engineered genetic information.

However, despite the potentially broad virus and gene transfer resistance of compressed genetic codes and widespread industrial applicability of these organisms, how natural genetic material could breach genetic-code-based resistance remained unanswered. Numerous viruses and mobile genetic elements encode parts of the translational apparatus, ranging from single tRNA genes and release factors up to lacking only ribosomal genes[7–9]. These genes allow mobile genetic elements to reduce their dependency on host translational processes[16–21].

[1]Department of Genetics, Harvard Medical School, Boston, MA, USA. [2]Department of Biomedical Informatics and Laboratory of Systems Pharmacology, Harvard Medical School, Boston, MA, USA. [3]Wyss Institute for Biologically Inspired Engineering, Harvard University, Boston, MA, USA. [4]Department of Pathology and Immunology, Washington University School of Medicine in St. Louis, St Louis, MO, USA. [5]The Edison Family Center for Genome Sciences and Systems Biology, Washington University School of Medicine in St. Louis, St Louis, MO, USA. [6]GenScript USA Inc., Piscataway, NJ, USA. [7]Present address: Department of Chemical Engineering, University of Washington, Seattle, WA, USA. ✉e-mail: akos_nyerges@hms.harvard.edu; gchurch@genetics.med.harvard.edu

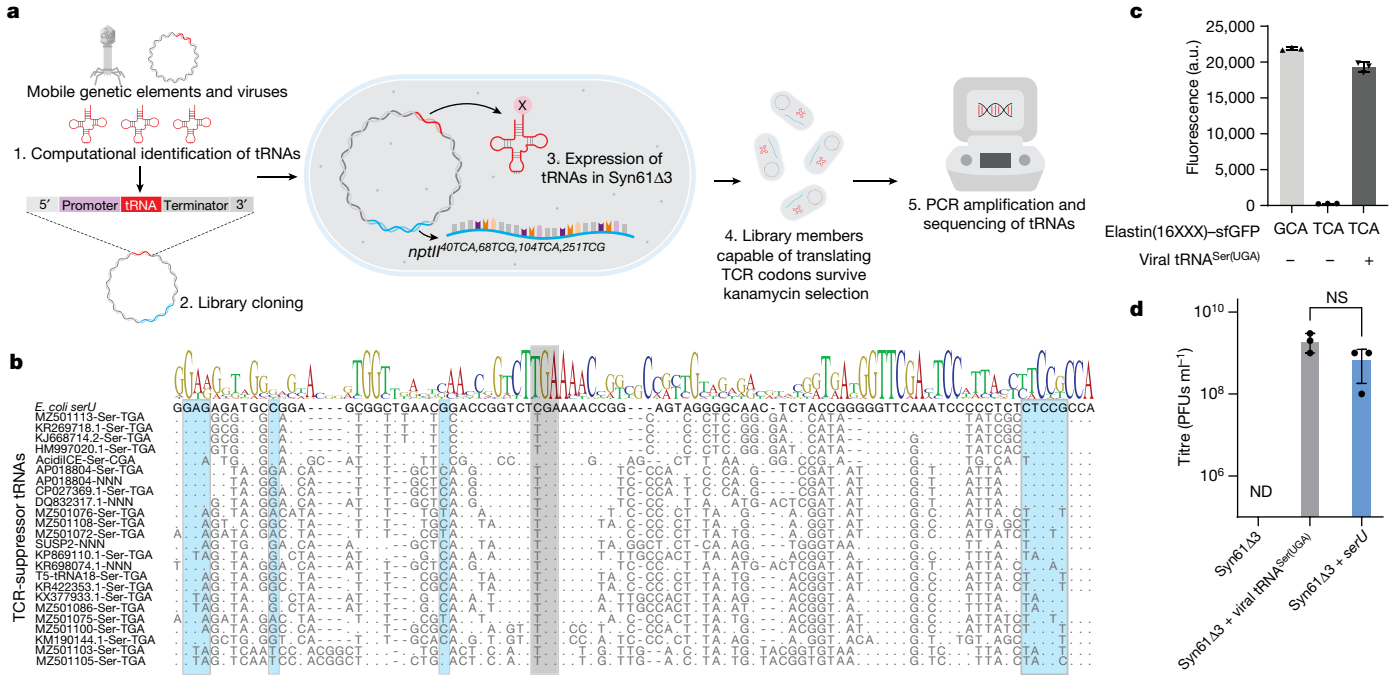


Fig. 1 | Discovery of mobile TCR-codon-translating tRNAs in *E. coli* Syn61Δ3. **a**, We screened the mobile tRNAome for tRNAs that can simultaneously translate TCA and TCG (together TCR) codons by computationally identifying tRNA genes in mobile genetic elements (1) and then synthesizing select candidates as an oligonucleotide library and cloning these variants into a plasmid vector carrying a *nptII*^{40TCA,68TCG,104TCA,251TCG} marker (conferring kanamycin resistance) (2). Following the transformation of this library into Syn61Δ3 (3), in which the deletion of *serU* (encoding tRNA$^{Ser(CGA)}$) and *serT* (encoding tRNA$^{Ser(UGA)}$) makes TCG and TCA codons unreadable, only variants carrying functional TCR-suppressor tRNAs survive kanamycin selection (4). Finally, high-throughput sequencing of the tRNA inserts from kanamycin-resistant clones identified suppressor tRNAs (5). **b**, Multiple sequence alignment of mobile TCR-codon-translating tRNAs. Grey shading indicates the anticodon region; the host's serine–tRNA ligase identity elements are shown in blue. **c**, Viral serine tRNA^{UGA} translates the TCA codon. Syn61Δ3 expressing elastin(16GCA(alanine))–sfGFP–His6 served as a wild-type expression control, and the elastin(16TCR)–sfGFP–His6 expression was compared with and without the coexpression of the tRNA$^{Ser(UGA)}$ of the *Escherichia* phage IrisVonRoten[24]. XXX represents the analysed codon, TCA or GCA. Bar graph represents the mean; error bars represent s.d. based on $n = 3$ independent experiments; a.u., arbitrary fluorescence units. **d**, The expression of viral TCR-suppressor tRNAs and *serU* (tRNA$^{Ser(CGA)}$) restores the replication of the T6 bacteriophage in Syn61Δ3. Circles represent data from $n = 3$ independent experiments, error bars represent s.d., and the bar graph represents the mean. ND, below the detection limit (that is, $<10^3$ plaque-forming units (PFUs) ml^{-1}); NS, not significant ($P = 0.116$) based on unpaired two-sided Student's *t*-test.

Therefore, the selection pressure posed by the compressed genetic code of genomically recoded organisms might facilitate the rapid evolution of viruses and mobile genetic elements capable of crossing a genetic-code-based barrier. Here we show that horizontally transferred tRNA genes can readily substitute cellular tRNAs and thus abolish genetic-code-based resistance to viral infections and HGT. We provide an example of a virus-resistant, biocontained bacterial host that prevents both incoming and outgoing HGT, which paves the way towards engineering multi-virus-resistant cell lines from any organisms and the safe use of GMOs in natural environments by eliminating HGT.

Mobile tRNAs abolish virus resistance

We first investigated whether tRNAs of mobile genetic elements can substitute cellular tRNAs and support viral infection in cells with a compressed genetic code. We sampled the mobile tRNAome, tRNA genes encoded by horizontally transferred genetic elements, by computationally screening thousands of viral genomes for the presence of tRNA genes, and then synthesizing 1,192 tRNA genes from phylogenetically diverse plasmids, transposable elements and bacteriophages infecting members of the Enterobacteriaceae family (Supplementary Data 1). Next we assayed these tRNAs for their ability to produce functional tRNAs in an *E. coli* host and substitute genomic tRNA genes to translate TCR codons. As depicted in Fig. 1a, this high-throughput assay is based on an *E. coli* strain with a synthetic recoded genome in which all annotated instances of two sense serine codons (TCG, TCA) and a stop codon (TAG) were replaced with synonymous alternatives, and the corresponding *serU*, *serT* tRNA genes and RF1 (*prfA*) have been deleted. This strain, *E. coli* Syn61Δ3, thereby relies on a 61-codon genetic code and prevents the expression of protein-coding genes containing TCR codons. Candidate tRNAs have been synthesized and cloned into a plasmid carrying each tRNA under a strong constitutive promoter together with an *nptII*^{40TCA,68TCG,104TCA,251TCG} aminoglycoside *O*-phosphotransferase antibiotic-resistance gene containing TCA codons at positions 40 and 104 and TCG codons at positions 68 and 251. In wild-type *E. coli* cells bearing the canonical genetic code, *nptII*^{40TCA,68TCG,104TCA,251TCG} confers resistance to kanamycin through serine incorporation at positions 40, 68, 104 and 251, and the production of full-length APH(3′)-II aminoglycoside *O*-phosphotransferase. In Syn61Δ3, however, the production of this resistance-conferring gene product is inhibited owing to the lack of *serU*- and *serT*-encoded tRNA$^{Ser(UGA)}$ and tRNA$^{Ser(CGA)}$ needed for TCR-codon decoding. Therefore, in our screen, only plasmid variants that are expressing tRNAs capable of decoding TCR codons will survive kanamycin selection. The transformation of this plasmid library into Syn61Δ3 and subsequent selection in the presence of kanamycin yielded thousands of colonies, indicating the presence of TCR-translating tRNAs in our library. Pooled extraction of plasmid variants from kanamycin-resistant colonies followed by amplicon sequencing of their tRNA insert identified 62 tRNA sequences capable of promoting *nptII*^{40TCA,68TCG,104TCA,251TCG} expression (Fig. 1b and Supplementary Data 1). These tRNAs represent 89% of all predicted TCR-codon-recognizing tRNAs in our library and share 33.7–61.1% (median = 46.2%) similarity to

the endogenous *serU* tRNA of *E. coli*. In agreement with the anticodon composition of mobile serine tRNAs, most tRNA hits contained a UGA anticodon and carried the identity elements necessary for recognition by the host's SerS serine–tRNA ligase (Fig. 1b).

Notable examples include the UAG-anticodon-containing serine tRNA of the laboratory model coliphage T5, tRNAs from plasmids of multidrug-resistant *E. coli* isolates (GenBank IDs AP018804 and CP023851) and the Ser-tRNA^{CGA} of the integrative conjugative element of *Acidithiobacillus ferrooxidans*. The presence of mobile tRNAs in integrative conjugative elements is especially concerning as these mobile genetic elements can carry up to 38 tRNAs corresponding to all 20 amino acids in a single operon and are capable of excision and transfer into neighbouring bacterial cells[20,22]. In agreement with the findings of previous studies[7,8,23], our computational screen also showed that mobile tRNA genes are not limited to mobile genetic elements of bacteria. Computational analysis of viruses infecting vertebrates and archaea revealed the presence of sense- and stop-codon-suppressor tRNAs in both groups, suggesting that mobile tRNAs are prevalent across viruses infecting prokaryotic, archaeal and eukaryotic hosts (Supplementary Data 2).

We confirmed the predicted serine amino acid identity of the TCR-codon-recognizing tRNAs by coexpressing a selected tRNA hit with an elastin(16TCA)–sfGFP–His6 construct harbouring a single TCA codon at position 16. The coexpression of the tRNA$^{Ser(UGA)}$ of the *Escherichia* phage IrisVonRoten[24] together with the elastin(16TCA)–sfGFP–His6 construct conferred near wild-type level expression (Fig. 1c), and tryptic digestion followed by reverse-phase liquid chromatography and tandem mass spectrometry (LC–MS/MS) confirmed serine incorporation at the TCA position (Extended Data Fig. 1a).

Next, we investigated whether mobile tRNAome-derived tRNAs could promote viral replication. A previous study demonstrated that Syn61Δ3 resists infection by multiple bacteriophages, including the Enterobacteria phage T6 (ref. [4]). Infecting Syn61Δ3 with the T6 phage recapitulated these results. By contrast, the infection of Syn61Δ3 harbouring a bacteriophage-derived Ser-tRNA^{UGA} gene with T6 resulted in rapid lysis, indicating that tRNA genes that reside in viral genomes can substitute cellular tRNAs and promote phage infection (Fig. 1d).

The discovery of diverse TCR-codon-translating tRNAs on horizontally transferred genetic elements indicates that mobile tRNA genes are widespread and can readily complement the lack of cellular tRNAs to promote viral replication and HGT.

Viruses infecting a recoded organism

We next investigated whether lytic viruses of Syn61Δ3 exist. We infected Syn61Δ3 cells with 11 coliphages whose genomes harbour TCR-translating tRNA genes on the basis of our plasmid-based screen (Fig. 1b). Notably, none of these eleven phages could overcome the recoded host's genetic isolation, indicating that the presence of tRNA genes on viral genomes does not directly rescue viral replication in recoded organisms (Extended Data Fig. 1b).

We next attempted to isolate lytic viruses from diverse environmental samples by carrying out a standard two-step enrichment-based phage isolation protocol and using Syn61Δ3 as the host. First, bacteria-free filtrates of environmental and wastewater samples (n = 13, from Massachusetts (USA), Extended Data Table 1a) were mixed with Syn61Δ3 and grown until stationary phase. Next, bacterial cells were removed, and we analysed the presence of lytic phages by mixing sample supernatants with Syn61Δ3 in soft-agar overlays. Five samples produced visible lysis. Viral plaque isolation from these samples followed by DNA sequencing and de novo genome assembly identified 12 new phage strains. All identified phages belong to the Caudovirales order and the Myoviridae family, taxa rich in tRNA-encoding bacteriophages[8] (Extended Data Table 1b). Computational identification of tRNA genes revealed the presence of tRNA operons in all phage isolates, with 10 to 27 tRNA genes in each genome (Supplementary Data 1). Notably, all isolates harboured TCR-codon-translating serine tRNAs with a UGA anticodon that we identified in our earlier *nptII*^{40TCA,68TCG,104TCA,251TCG} suppressor screen (Fig. 1b). One isolate, REP1, also harboured a predicted homing endonuclease (Fig. 2c). Homing endonucleases encoded in tRNA operons have been shown to be responsible for the horizontal transfer of tRNA gene clusters[25]. Phage isolates showed more than two orders of magnitude difference in viral titres after replication on recoded cells (Fig. 2a). One of the most virulent isolates, REP12, required 60 min to complete a replication cycle at 37 °C in Syn61Δ3 (Fig. 2b).

The isolated viral strains infecting Syn61Δ3 show that bacteriophages that can overcome viral resistance based on sense codon recoding exist and are widespread in environmental samples. However, the presence of TCR-codon-translating serine tRNAs on viral genomes does not directly rescue viral replication in the recoded strain.

Viral tRNAs substitute cellular tRNAs

We next investigated how tRNA-encoding viruses evade genetic-code-based resistance. Time-course transcriptome analysis of REP12-phage-infected Syn61Δ3 cells during the viral replication cycle revealed early and high-level expression of the viral tRNA operon (Supplementary Fig. 1). In agreement with this observation, the computational prediction of bacterial promoters driving the tRNA array indicated the presence of multiple strong constitutive promoters upstream of the tRNA operon region. We then investigated the time-course kinetics of tRNA expression in Syn61Δ3 cells that were infected with our REP12 phage by carrying out tRNA sequencing (tRNA-seq). Time-course tRNA-seq experiments revealed remarkably high-level expression of the viral tRNA$^{Ser(UGA)}$ immediately after phage attachment (that is, a relative viral tRNA$^{Ser(UGA)}$ abundance of 56.1% (±5%) compared to the host *serV* tRNA; Extended Data Fig. 2). Throughout the entire phage replication cycle, the phage tRNA$^{Ser(UGA)}$ remained one of the most abundant viral tRNA species inside infected Syn61Δ3 cells (Extended Data Fig. 2). We next investigated whether phage tRNA$^{Ser(UGA)}$ participates in translation by analysing the presence of their mature form. The gene encoding the tRNA$^{Ser(UGA)}$ in the genome of REP12 does not encode the universal 5′-CCA tRNA tail, which allows for amino acid attachment as well as for interaction with the ribosome. Therefore, CCA tail addition must happen before these tRNAs can participate in translational processes. The sequencing-based analysis of phage tRNA$^{Ser(UGA)}$ ends detected CCA tail addition in 62.9% (±1.9%) of all tRNA-seq reads immediately after phage attachment, indicating that mature tRNA$^{Ser(UGA)}$ is instantly produced after host infection (Extended Data Fig. 3).

We also investigated transcriptomic changes in Syn61Δ3 during phage replication. Analysis of the host transcriptome after phage infection revealed upregulation in genes responsible for tRNA maturation and modification. Upregulated genes include *queG*, encoding epoxyqueuosine reductase (which catalyses the final step in the de novo synthesis of queuosine in tRNAs)[26], and *trmJ*, encoding tRNA Cm32/Um32 methyltransferase[27] (which introduces methyl groups at the 2′-*O* position of U32 of several tRNAs, including tRNA$^{Ser(UGA)}$), suggesting the potential post-transcriptional modification of phage-derived tRNAs (Extended Data Fig. 4).

Finally, we also validated the role of phage tRNA$^{Ser(UGA)}$ tRNAs in decoding TCR codons. We first cloned the REP12 viral tRNA operon containing the hypothetical tRNA$^{Ser(UGA)}$ and its predicted promoter into a plasmid vector. Coexpression of this tRNA operon with an elastin(16TCA)–sfGFP–His6 and elastin(16TCG)–sfGFP–His6 construct, harbouring either a single TCA or TCG codon at position 16, respectively, resulted in high-level elastin–sfGFP–His6 expression (Fig. 2d). Next, tryptic digestion followed by LC/MS–MS analysis confirmed serine incorporation in response to both the TCA and TCG codon in these elastin(16TCR)–sfGFP–His6 samples (Fig. 2e and Supplementary Fig. 2). As expected, the coexpression of the same elastin(16TCA)–sfGFP–His6 construct

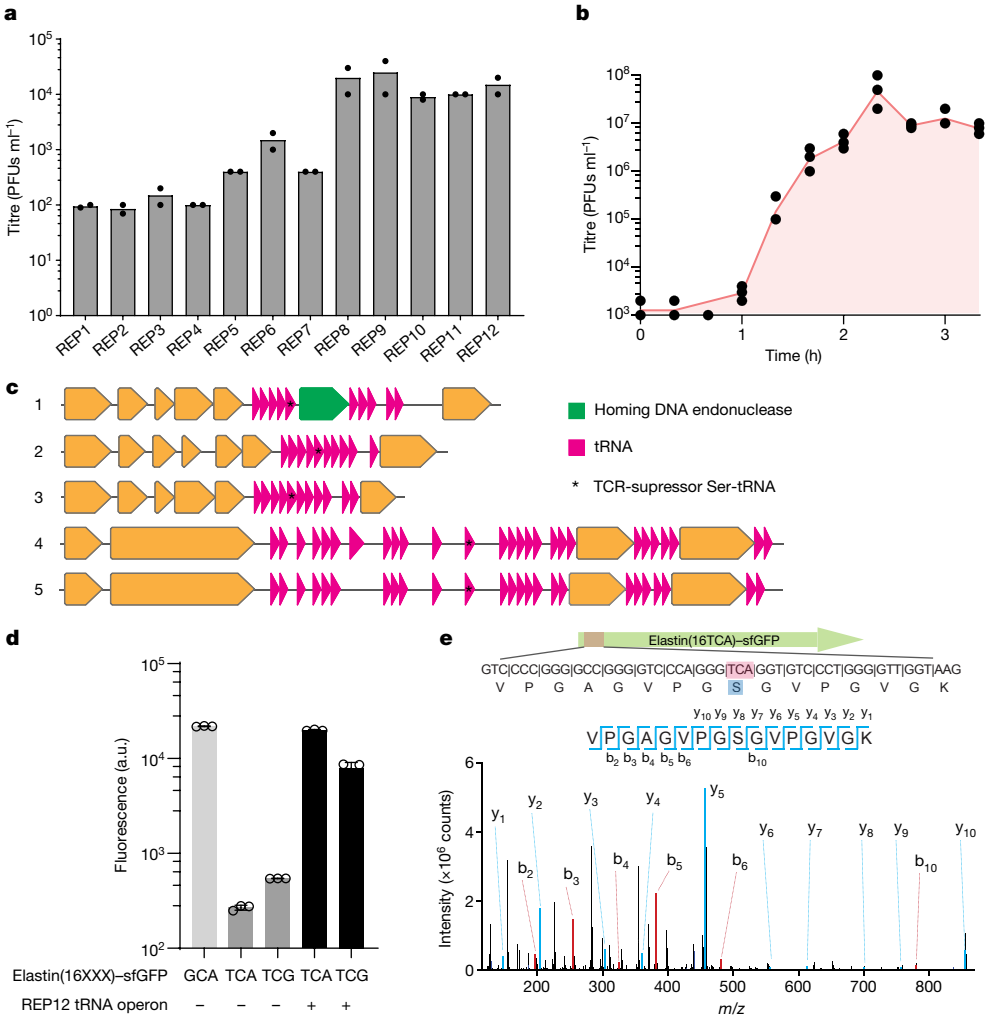


Fig. 2 | Lytic phages of Syn61Δ3. a, Titre of Syn61Δ3 phage isolates after replication on Syn61Δ3. Circles represent data from $n = 2$ independent experiments; bar graphs represent the mean. **b**, Single-step growth curve of the REP12 lytic Syn61Δ3 phage. Single-step growth was carried out in $n = 3$ independent experiments; red line represents the mean, and circles represent the total viral titre. **c**, Genomic maps of tRNA operons in lytic Syn61Δ3 phages. Magenta arrows represent predicted tRNA genes; asterisks denote tRNA genes identified in our earlier TCR codon suppressor screen (Fig. 1b); the green arrow represents a homing endonuclease gene, and orange arrows represent protein-coding genes. Phage operon numbers correspond to the following REP phages—1, REP1; 2, REP2; 3, REP4; 4, REP6; 5, REP12. **d**, tRNAs expressed by the viral tRNA operon translate TCR codons. Syn61Δ3 expressing elastin(16GCA(alanine))–sfGFP–His6 served as a wild-type expression control, and the elastin(16TCR)–sfGFP–His6 expression was compared with and without the coexpression of the REP12 viral tRNA operon. XXX represents the analysed codon, TCA, TCG or GCA. a.u., arbitrary fluorescence units; bar graphs represent the mean. Error bars represent s.d. based on $n = 3$ independent experiments. **e**, tRNAs expressed by the viral tRNA operon decode TCR codons as serine. The amino acid identity of the translated TCA codon within elastin(16TCA)–sfGFP–His6 was confirmed by MS/MS from Syn61Δ3 cells containing the REP12 tRNA operon and its cognate promoter. The figure shows the amino acid sequence and MS/MS spectrum of the analysed elastin(16TCR) peptide. MS/MS data were collected once.

with the only $tRNA^{Ser(UGA)}$ of the viral tRNA operon conferred a similar effect, and LC/MS–MS analysis confirmed the role of this tRNA in decoding viral TCR codons as serine (Supplementary Fig. 2).

Together these results show that lytic phages of Syn61Δ3 overcome genetic-code-based viral resistance by rapidly complementing the cellular tRNA pool with virus-encoded tRNAs.

Swapped genetic code resists viruses

We predicted that establishing an artificial genetic code, in which TCR codons encode an amino acid different from their natural serine identity, would create a genetic firewall that safeguards cells from HGT and infection by tRNA-encoding viruses. In an amino acid-swapped genetic code, viral tRNAs would compete with host-expressed tRNAs that decode TCR codons as a non-serine amino acid resulting in the mistranslation of viral proteins. As the genetic isolation provided by amino acid-swapped genetic codes is expected to correlate with the hydrophobicity and polarity distance between exchanged amino acids[28], we sought to reassign TCR serine codons to leucine, the amino acid most distant from the natural serine meaning of the TCR codon. The more distant tryptophan, phenylalanine, tyrosine and isoleucine exchanges were not investigated as their cognate aminoacyl-tRNA synthetases are nonpermissive to anticodon mutations and expected to require extensive enzyme and/or tRNA engineering to function[29–31]. Although alternative forms of swapped codes emerged in the past, these codes either targeted the exchange of chemically more similar amino acids (that is, serine to alanine, histidine and proline) or remained tested only in vitro[6,28,32,33].

To establish a serine(TCR)-to-leucine swapped genetic code (Fig. 3a), we utilized Syn61Δ3, which genome-wide lacks annotated instances of

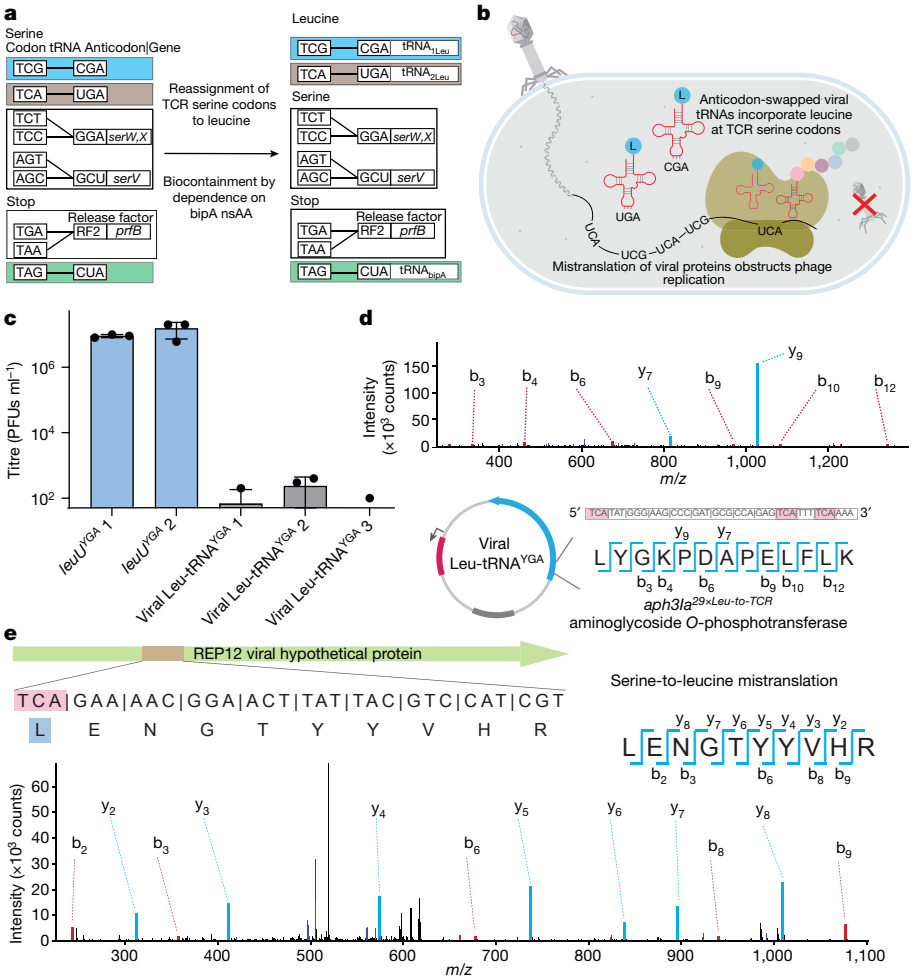


Fig. 3 | An amino acid-swapped genetic code provides multi-virus resistance. **a**, The creation of a genomically recoded organism, Ec_Syn61Δ3-SL, from *E. coli*, in which both TCA and TCG—naturally serine-meaning—codons are translated as leucine. The introduction of bacteriophage-derived Leu-tRNA^{UGA} and Leu-tRNA^{CGA} to Syn61Δ3 reassigns TCA and TCG codons to leucine, and the reassignment of the TAG stop codon to encode bipA in an essential gene of the host ensures the biocontainment of Ec_Syn61Δ3-SL. **b**, Schematic of viral infection in Ec_Syn61Δ3-SL. The reassignment of the sense codons TCA and TCG to leucine in Ec_Syn61Δ3-SL provides multi-virus resistance by mistranslating the viral proteome. **c**, Bacteriophage-derived Leu-tRNA^{UGA} and Leu-tRNA^{CGA} expression in Syn61Δ3 provides multi-virus resistance. The figure shows the titre of lytic Syn61Δ3 phages following the infection of the corresponding Leu-tRNA^{YGA}-expressing Syn61Δ3 strain with a mixture of 12 distinct REP Syn61Δ3 phages (Extended Data Table 1 and Supplementary Data 3). All experiments were carried out in three independent replicates; circles represent data from $n = 3$ independent experiments; bar graphs represent the mean; the error bars represent s.d. **d**, The reassignment of TCR codons to leucine within the coding sequence of *aph3Ia*$^{29\times Leu\text{-}to\text{-}TCR}$ in Syn61Δ3-LS was confirmed by MS/MS. The figure shows the amino acid sequence and MS/MS spectrum of the detected *aph3Ia*$^{29\times Leu\text{-}to\text{-}TCR}$ peptide and its coding sequence. MS/MS data were collected once. **e**, Mistranslated viral protein synthesis in Ec_Syn61Δ3-SL. The figure shows the amino acid sequence and MS/MS spectrum of a bacteriophage-expressed protein, together with its viral genomic sequence, in which the naturally serine-coding TCA codon is mistranslated as leucine. The experiment was carried out by infecting Ec_Syn61Δ3-SL cells, expressing Leu9-tRNA^{YGA} from the *Escherichia* phage OSYSP, with the REP12 phage, and the proteome of infected cells was analysed by MS/MS. MS/MS data were collected once.

TCR codons and their corresponding tRNA genes, and sought to identify tRNAs capable of efficiently translating TCR codons as leucine. We modified our high-throughput tRNA library screen (Fig. 1a) to evolve TCR suppressors from the endogenous *E. coli leuU* tRNA carrying a TCA and TCG decoding anticodon. We coexpressed a 65,536-member mutagenized library of the anticodon-swapped *leuU* tRNA gene in which the anticodon loop of both tRNAs has been fully randomized, together with *aph3Ia*$^{29\times Leu\text{-}to\text{-}TCR}$, a kanamycin-resistance-conferring gene in which all 29 instances of leucine codons were replaced with TCR serine codons. In this system, only anticodon-swapped *leuU* variants capable of simultaneously translating all 29 TCR codons as leucine would confer resistance to kanamycin. High kanamycin selection pressure in combination with 29 instances of TCR codons in *aph3Ia* was expected to select tRNA variants that provide wild-type-level translation efficiency for TCR codons. We identified two distinct *leuU* variants by applying 'high' kanamycin concentration (that is, 200 μg ml^{-1}) as selection pressure to Syn61Δ3 cells carrying the anticodon-swapped tRNA library. These variants, carrying tRNAs containing distinct anticodon loop mutations (Extended Data Fig. 5a), were then infected with a cocktail of all 12 phage isolates (Extended Data Table 1) that are capable of lysing Syn61Δ3 at a 10:1 cell-to-phage ratio (that is, a multiplicity of infection of 0.1). Instead of plaque and efficiency-of-plating assays[34,35] with selected phage isolates and assaying viral titre on the resistant strain, we quantified virus resistance using a mixture of all 12 Syn61Δ3-infecting viral isolates, and measured virion release from infected cells after 24 h by plating culture supernatants on a virus-susceptible wild-type *E. coli* host. This can reveal even slow or non-plaque-forming viral replication in the target cells[35,36] (Supplementary Note).

Notably, all selected *leuU* library members allowed robust phage replication, with phage titres reaching about 10^7 plaque-forming units

per millilitre after 24 h (Fig. 3c). We reasoned that viral replication in the presence of TCR-suppressing *leuU* variants is due to the lower suppression efficiency of their tRNAs compared to that of competing phage-carried serine tRNAs, which leads to rapid viral takeover. Viral $tRNA^{Ser(YGA)}$ species, that are $tRNA^{Ser(UGA)}$ and $tRNA^{Ser(CGA)}$, might: have higher aminoacylation efficiency by their corresponding *E. coli* aminoacyl-tRNA ligase than our selected *leuU* variants; have higher affinity towards the bacterial ribosome; and/or better evade phage- and host-carried tRNA-degrading effector proteins[16,37,38].

On the basis of this observation, we reasoned that bacteriophage-encoded tRNAs might provide much higher suppression efficiencies for their cognate codons than their native *E. coli* counterpart. Therefore, and to exploit this superior translation efficiency of bacteriophage-encoded tRNAs and their resistance to phage-carried tRNA-degrading enzymes, we next constructed a small, focused library that coexpressed YGA-anticodon-swapped mutants of 13 phage-encoded leucine tRNAs, together with the *aph3Ia*$^{29\times Leu\text{-}to\text{-}TCR}$ aminoglycoside *O*-phosphotransferase gene. The transformation of this library into Syn61Δ3 cells and subsequent 'high'-concentration (that is, 200 μg ml^{-1}) kanamycin selection identified three distinct tRNAs enabling robust growth. Identified tRNAs showed only 48.3–37.9% similarity to *E. coli leuU* but carried most of the canonical *E. coli* leucine–tRNA ligase identity elements (Extended Data Fig. 5b). Furthermore, the analysis of the total tRNA content of these cells by tRNA-seq confirmed the presence of synthetic phage Leu-$tRNA^{YGA}$ tRNAs with similar abundances to those of the cellular endogenous serine tRNAs (that is, a relative expression level of 172% and 140% for Leu-$tRNA^{UGA}$ and Leu-$tRNA^{CGA}$, respectively, compared to *serV* (Extended Data Fig. 6)).

Next, as in our previous infection assay, phage-$tRNA^{YGA}$-expressing cells were infected with a mixture of 12 distinct, lytic phages of Syn61Δ3 at a multiplicity of infection of 0.1. The analysis of phage titre in culture supernatants after 24 h showed a marked drop compared to the input phage inoculum, suggesting that anticodon-swapped viral leucine tRNAs block phage replication (Fig. 3c).

We then investigated the mechanism of phage resistance in *E. coli* cells carrying virus-derived $tRNA^{Leu(YGA)}$ tRNAs (Ec_Syn61Δ3 serine-to-leucine swap, or Ec_Syn61Δ3-SL in short; Fig. 3b) by carrying out total proteome analysis. Untargeted deep proteome analysis of uninfected cells by MS/MS validated the translation of TCR codons as leucine in Ec_Syn61Δ3-SL (Fig. 3d). Time-course untargeted proteome analysis after bacteriophage infection revealed extensive mistranslation at TCR codons in newly synthesized phage proteins (Fig. 3e and Supplementary Fig. 3), indicating that an amino acid-swapped genetic code broadly obstructs viral protein synthesis. In agreement with earlier reports that showed the partial recognition of TCT codons by $tRNA^{UGA}$ (refs. [39,40]) and owing to the extreme sensitivity of our untargeted proteomics assay, serine-to-leucine mistranslation was also detected at TCT codon positions in Ec_Syn61Δ3-SL cells (Extended Data Fig. 7). The recognition of TCT codons by phage $tRNA^{Leu(YGA)}$ tRNAs might also be responsible for the fitness decrease of Ec_Syn61Δ3-SL cells compared to those of its ancestor strain (that is, a doubling time of 69.3 min, compared to 44.29 min for the parental Syn61Δ3 strain in rich 2×YT medium (Extended Data Fig. 8)). Alternatively, the fitness decrease of Ec_Syn61Δ3-SL might also be attributable to the presence of TCR codons in essential genes of Syn61Δ3. According to our genome analysis, at least four essential genes of Syn61Δ3 (*mukE*, *ykfM*, *yjbS* and *safA*) contain TCR codons and become mistranslated in Ec_Syn61Δ3-SL (Supplementary Data 4).

Finally, we also sought to develop a tightly biocontained version of Ec_Syn61Δ3-SL because a virus-resistant strain might have a competitive advantage in natural ecosystems due to the lack of predating bacteriophages. Synthetic auxotrophy based on the engineered reliance of essential proteins on human-provided nonstandard amino acids (nsAAs; for example, L-4,4′-biphenylalanine (bipA)) offers tight biocontainment that remains stable under long-term evolution[10,41,42]. Therefore, we generated a recombination-deficient (that is, Δ*recA*), biocontained version of Ec_Syn61Δ3-SL bearing a bipA-dependent essential *adk* gene and the bipA aminoacyl-tRNA synthetase–$tRNA^{bipA(CUA)}$ system by first carrying out adaptive laboratory evolution on a *recA*-knockout Syn61Δ3, and then replacing the genomic *adk* copy with its bipA-dependent variant[43] (Methods). This strain maintained the low escape frequency of previously reported singe-gene synthetic auxotrophs[10] (that is, 2.9×10^{-6} ($\pm 5.9 \times 10^{-7}$) escapees per colony-forming units (CFUs) escape frequency) and provided robust growth. We also tested the viral resistance of Ec_Syn61Δ3-SL under mock environmental conditions by repeating our phage enrichment and isolation process with a mixture of 12 environmental samples, including fresh sewage (Extended Data Table 1 and Supplementary Note), but could not detect plaque-forming phages in culture supernatants (Extended Data Fig. 9).

Together, these results demonstrate that reassigning the sense codons TCA and TCG to leucine in vivo provides broad protection against viruses, including new mixtures of viruses directly from environmental samples, and the TAG stop codon can be simultaneously utilized to biocontain this virus-resistant strain through dependence on an amino acid not found in nature.

Genetic-code-based bidirectional firewall

Finally, we developed a set of plasmid vectors that we systematically addicted to an amino acid-swapped genetic code in which leucine is encoded as TCR codons. GMOs are increasingly deployed for large-scale use in agriculture, therapeutics, bioenergy and bioremediation. Consequently, it is critical to implement robust biocontainment strategies that prevent the unintended proliferation of GMOs and protect natural ecosystems from engineered genetic information. Although efficient biocontainment strategies exist (for example, bipA nsAA-based synthetic auxotrophy, as in Ec_Syn61Δ3-SL), current methods fail to prevent the HGT-based escape of engineered genetic information. Synthetic addiction to a swapped genetic code offers a solution to this problem. Using our Ec_Syn61Δ3-SL cells expressing virus-derived $tRNA^{Leu(YGA)}$, we therefore developed a set of plasmid vectors that depend on TCR codons to express leucine-containing proteins and thus can function only in cells that efficiently translate TCR codons as leucine (Fig. 4a). These plasmids, called the pLS plasmids, offer four widely used, orthogonal antibiotic-resistance markers in combination with four mutually orthogonal low- to high-copy-number origins of replication for stable maintenance in Ec_Syn61Δ3-SL cells (Fig. 4b and Supplementary Data 3). Antibiotic-resistance genes and proteins necessary for pLS plasmid replication encode leucine as TCR—naturally serine-meaning—codons and, therefore, fail to function in cells bearing the canonical genetic code. The addiction of resistance markers and replication proteins to an artificial genetic code ensures that pLS plasmids can stably and safely maintain synthetic genetic functions but restrict the functionality of these genes to Ec_Syn61Δ3-SL cells. Using our pLS plasmids, we achieved serine-to-leucine codon reassignment simultaneously at about 14,000 codon positions in a single cell, thus demonstrating the exceptional codon reassignment efficiency of our viral tRNA-based system.

We then tested the ability of our pLS vectors to function in cells bearing the standard genetic code by electroporating plasmid DNA from six variants (that is, pLS1–6) into wild-type *E. coli* K12 MG1655 cells but could not detect escapees carrying pLS plasmids (that is, an electroporation efficiency of <1 CFU per microgram, whereas the electroporation efficiency of the recipient cells was $3 \times 10^9 \pm 1.1 \times 10^9$ CFUs per microgram (±s.d., $n = 3$ independent experiments) based on a pUC-Kan^R control plasmid). The escape of pLS plasmids was similarly prevented when the expression cassette for phage $tRNA^{Leu(YGA)}$ was encoded within the plasmid backbone (that is, pLS1 and pLS2), indicating that anticodon-swapped viral tRNAs are severely toxic to wild-type cells. On the basis of these results, we also expect that, similarly to genes of pLS, any leucine-requiring protein, up to entire chromosomes, can

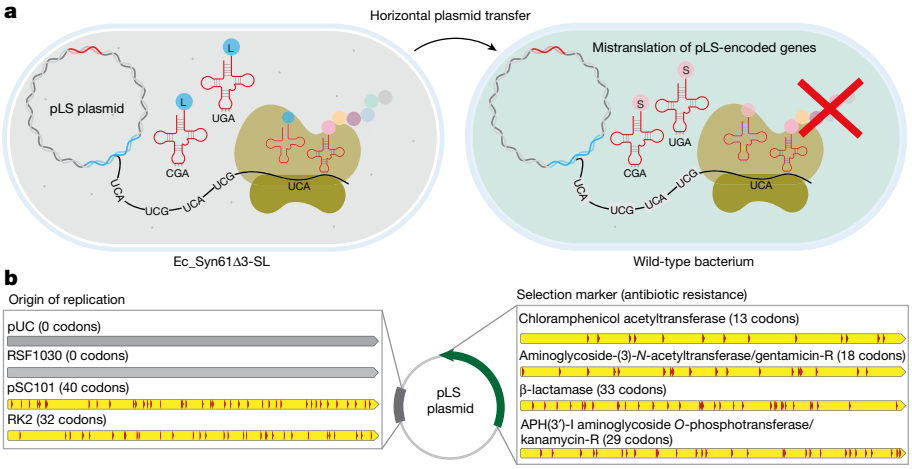


Fig. 4 | Addiction of synthetic genetic information to a genetic code in which TCR codons encode leucine prevents HGT. a, We developed a set of plasmid vectors, termed the pLS plasmids, that rely on TCR codons to express leucine-containing proteins. pLS plasmids function only in Ec_Syn61Δ3-SL expressing bacteriophage-derived synthetic $tRNA^{Leu(YGA)}$ tRNAs, and the encoded proteins of pLS plasmids become mistranslated in cells bearing the canonical genetic code. **b**, The pLS plasmids offer multiple mutually orthogonal antibiotic-resistance markers together with low- to high-copy-number origins of replication that are addicted to an artificial genetic code in which leucine is encoded as TCR codons. The number in parenthesis marks the number of Leu(TCR) codons in each gene. Detailed sequence information and a description of the pLS plasmids are available in Supplementary Data 3.

be addicted to Ec_Syn61Δ3-SL by recoding target genes to encode one or more leucine positions as TCR codons.

In sum, the addiction of pLS plasmids to an artificial genetic code in which leucine is encoded as TCR codons, in combination with nsAA-based synthetic auxotrophy, promises escape-free biocontainment for engineered genetic information.

Discussion

Previous works provided support for using rational genetic code engineering to isolate GMOs from natural ecosystems by preventing viral infections and HGT[2–4]; however, how natural viruses and transferred genes could breach genetic-code-based resistance remained unanswered. By systematically investigating HGT into *E. coli* Syn61Δ3, a synthetic organism with a compressed genetic code[4], we demonstrated that tRNAs expressed by bacteriophages, plasmids and integrative conjugative elements readily substitute cellular tRNAs and abolish genetic-code-based isolation. We discovered 12 viruses in environmental samples that can infect *E. coli* Syn61Δ3, despite its compressed genetic code (Fig. 2a). These bacteriophages express up to 27 tRNAs, including a functional $tRNA^{Ser(UGA)}$ needed to replace the host's deleted tRNAs. These findings suggest that the selection pressure posed by compressed genetic codes can facilitate the rapid evolution of viruses and mobile genetic elements capable of crossing a genetic-code-based barrier. This hypothesis is further supported by the co-localization of tRNAs and homing endonucleases (Fig. 2c) that catalyse copy-and-paste tRNA operon transfer[25]. We have also shown that mobile tRNAs are not limited to bacteria, as multiple archaeal and eukaryotic viruses also carry predicted tRNA genes.

Next, as a combined solution against HGT and viral infections, including tRNA-expressing viruses, we have created a new type of biocontained *E. coli* strain, Ec_Syn61Δ3-SL, carrying an amino acid-swapped genetic code. Ec_Syn61Δ3-SL achieves unprecedented gene-transfer resistance by the reassignment of TCR codons to leucine—an amino acid distant in physicochemical properties compared to their natural serine identity—using reprogrammed, virus-derived tRNAs. In contrast to previous studies[2–4,33], our study confirmed the virus resistance of this strain using a broad range of viruses. We found that viral tRNAs provide exceptional codon reassignment efficiency and are superior for establishing altered genetic codes compared to cellular tRNAs. We also showed that the swapped genetic code of Ec_Syn61Δ3-SL simultaneously prevents viral replication and the escape of synthetic genetic information into wild organisms. As a direct application, we addicted the most widely used plasmid vectors to express leucine-containing proteins with TCR codons and developed plasmids that cannot function in natural organisms. By demonstrating efficient serine-to-leucine reassignment, we made biological containment possible at the level of genes, to operons, to entire chromosomes. Adaptive laboratory evolution, which has substantially improved doubling times of organisms with altered genetic codes[4,44], will be applied to Ec_Syn61Δ3-SL to help move towards industrial use.

Potential limitations of our work include the sampled viral diversity and the inability of our viral resistance assays to sample multi-step, long-term evolutionary processes of viruses. Although our tests of the viral resistance of Ec_Syn61Δ3-SL included 12 lytic Syn61Δ3 phages and a complex mixture of environmental samples, including fresh sewage, we cannot rule out the existence of Ec_Syn61Δ3-SL-infecting phages in Earth's biome. Our assays provided an example of Ec_Syn61Δ3-SL culture contamination with soil, sewage and faecal material (Extended Data Table 1a); however, lysogenic viruses or jumbo phages and megaphages, with genomes of more than 200 and 500 kilobase pairs, respectively, could have remained undetected in our assays owing to their frequent inability to form visible plaques[9,36]. We reason that bacteriophages with few TCR codons in essential genes (for example, *Escherichia* phage EC6098 with only 33 TCR positions in its 6 protein-coding genes[45]) have the highest potential to overcome an amino acid-swapped code. Alternatively, virus-expressed tRNA-degrading proteins[37,38] could evolve to selectively destroy mistranslating tRNAs and thus promote viral escape; however, both mechanisms would require multiple simultaneous mutations and thus are unlikely to occur. Future work will explore whether such viruses can evade the amino acid-swapped genetic code of Ec_Syn61Δ3-SL.

After our manuscript was publicly submitted, another study[33] published findings consistent with some of ours, identifying two tRNAs that could overcome the virus resistance of Syn61Δ3 and noting the utility of swapped genetic codes to obstruct HGT. The use of weak *E. coli* tRNAs was insufficient to produce true resistance to a broad set of viruses in our experiments. The apparent success with such weak

tRNAs in ref. [33] may be due to testing only two (related) viral strains and using efficiency-of-plating[34,35] and conjugation assays (Supplementary Note). Furthermore, the lower expression level of mistranslating tRNAs used in ref. [33], in combination with the use of *E. coli* tRNAs still susceptible to virus-expressed tRNA-degrading enzymes[37,38], might result in compromised resistance in the presence of complex environmental viral communities. The authors of ref. [33] did not implement biocontainment strategies to limit the unwanted proliferation of their potentially virus-resistant host.

These findings fundamentally affect ongoing prokaryotic and eukaryotic genome recoding projects[5,46–48], including our aim to engineer a virus-resistant stop-codon-recoded human cell line and a 57-codon strain of *E. coli*, as some of the identified tRNAs are expected to enable viral replication in these recoded organisms. Therefore, we are now implementing the swapped genetic code of this work to ensure the virus and HGT resistance of these engineered hosts. We expect that our results will have implications for the safe use of GMOs in open environments by establishing a generalizable method for genetic code alteration that simultaneously prevents viral predation in natural ecosystems and blocks incoming and outgoing HGT with natural organisms. Follow-up works could utilize biocontained swapped genetic codes to prevent transgene release during bioremediation, offer containment for open cultures of engineered photoautotrophs for carbon sequestration, and enable safe microbiome engineering and vaccine production by addicting living treatments to use a swapped code. The combination of genome recoding and codon swap may provide a universal strategy to make any species resistant to many or all natural viruses.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-023-05824-z.

1. Church, G. M. & Regis, E. *Regenesis: How Synthetic Biology Will Reinvent Nature and Ourselves* (Basic Books, 2014).
2. Lajoie, M. J. et al. Genomically recoded organisms expand biological functions. *Science* **342**, 357–360 (2013).
3. Ma, N. J. & Isaacs, F. J. Genomic recoding broadly obstructs the propagation of horizontally transferred genetic elements. *Cell Syst.* **3**, 199–207 (2016).
4. Robertson, W. E. et al. Sense codon reassignment enables viral resistance and encoded polymer synthesis. *Science* **372**, 1057–1062 (2021).
5. Ostrov, N. et al. Design, synthesis, and testing toward a 57-codon genome. *Science* **353**, 819–822 (2016).
6. Fujino, T., Tozaki, M. & Murakami, H. An amino acid-swapped genetic code. *ACS Synth. Biol.* **9**, 2703–2713 (2020).
7. Abrahão, J. et al. Tailed giant Tupanvirus possesses the most complete translational apparatus of the known virosphere. *Nat. Commun.* **9**, 749 (2018).
8. Morgado, S. & Vicente, A. C. Global in-silico scenario of tRNA genes and their organization in virus genomes. *Viruses* **11**, 180 (2019).
9. Al-Shayeb, B. et al. Clades of huge phages from across Earth's ecosystems. *Nature* **578**, 425–431 (2020).
10. Mandell, D. J. et al. Biocontainment of genetically modified organisms by synthetic protein design. *Nature* **518**, 55–60 (2015).
11. Zou, X. et al. Systematic strategies for developing phage resistant *Escherichia coli* strains. *Nat. Commun.* **13**, 4491 (2022).
12. Barone, P. W. et al. Viral contamination in biologic manufacture and implications for emerging therapies. *Nat. Biotechnol.* https://doi.org/10.1038/s41587-020-0507-2 (2020).
13. Baltz, R. H. Bacteriophage-resistant industrial fermentation strains: from the cradle to CRISPR/Cas9. *J. Ind. Microbiol. Biotechnol.* **45**, 1003–1006 (2018).
14. Ostrov, N. et al. Synthetic genomes with altered genetic codes. *Curr. Opin. Syst. Biol.* https://doi.org/10.1016/j.coisb.2020.09.007 (2020).
15. Fredens, J. et al. Total synthesis of *Escherichia coli* with a recoded genome. *Nature* https://doi.org/10.1038/s41586-019-1192-5 (2019).
16. Yang, J. Y. et al. Degradation of host translational machinery drives tRNA acquisition in viruses. *Cell Syst.* **12**, 771–779 (2021).
17. Peters, S. L. et al. Experimental validation that human microbiome phages use alternative genetic coding. *Nat. Commun.* **13**, 5710 (2022).
18. Borges, A. L. et al. Widespread stop-codon recoding in bacteriophages may regulate translation of lytic genes. *Nat. Microbiol.* **7**, 918–927 (2022).
19. Abe, T. et al. tRNADB-CE 2011: tRNA gene database curated manually by experts. *Nucleic Acids Res.* **39**, D210–D213 (2011).
20. Alamos, P. et al. Functionality of tRNAs encoded in a mobile genetic element from an acidophilic bacterium. *RNA Biol.* **15**, 518–527 (2018).
21. Santamaría-Gómez, J. et al. Role of a cryptic tRNA gene operon in survival under translational stress. *Nucleic Acids Res.* **49**, 8757–8776 (2021).
22. Bustamante, P. et al. ICE*Afe*1, an actively excising genetic element from the biomining bacterium *Acidithiobacillus ferrooxidans*. *J. Mol. Microbiol. Biotechnol.* **22**, 399–407 (2012).
23. Bowden, R. J., Simas, J. P., Davis, A. J. & Efstathiou, S. Murine gammaherpesvirus 68 encodes tRNA-like sequences which are expressed during latency. *J. Gen. Virol.* **78**, 1675–1687 (1997).
24. Maffei, E. et al. Systematic exploration of *Escherichia coli* phage–host interactions with the BASEL phage collection. *PLOS Biol.* **19**, e3001424 (2021).
25. Brok-Volchanskaya, V. S. et al. Phage T4 SegB protein is a homing endonuclease required for the preferred inheritance of T4 tRNA gene region occurring in co-infection with a related phage. *Nucleic Acids Res.* **36**, 2094–2105 (2008).
26. Miles, Z. D., McCarty, R. M., Molnar, G. & Bandarian, V. Discovery of epoxyqueuosine (oQ) reductase reveals parallels between halorespiration and tRNA modification. *Proc. Natl Acad. Sci. USA* **108**, 7368–7372 (2011).
27. Liu, R.-J., Long, T., Zhou, M., Zhou, X.-L. & Wang, E.-D. tRNA recognition by a bacterial tRNA Xm32 modification enzyme from the SPOUT methyltransferase superfamily. *Nucleic Acids Res.* **43**, 7489–7503 (2015).
28. Schmidt, M. & Kubyshkin, V. How to quantify a genetic firewall? A polarity-based metric for genetic code engineering. *ChemBioChem* **22**, 1268–1284 (2021).
29. Yang, X.-L. et al. Two conformations of a crystalline human tRNA synthetase–tRNA complex: implications for protein synthesis. *EMBO J.* **25**, 2919–2929 (2006).
30. Kobayashi, T. et al. Structural basis for orthogonal tRNA specificities of tyrosyl-tRNA synthetases for genetic code expansion. *Nat. Struct. Mol. Biol.* **10**, 425–432 (2003).
31. Giege, R., Sissler, M. & Florentz, C. Universal rules and idiosyncratic features in tRNA identity. *Nucleic Acids Res.* **26**, 5017–5035 (1998).
32. Church, G., Baynes, B. & Pitcher, E. Hierarchical assembly methods for genome engineering. PCT/US2006/001427 Patent application (2007).
33. Zürcher, J. F. et al. Refactored genetic codes enable bidirectional genetic isolation. *Science* https://doi.org/10.1126/science.add8943 (2022).
34. Łoś, J. M., Golec, P., Węgrzyn, G., Węgrzyn, A. & Łoś, M. Simple method for plating *Escherichia coli* bacteriophages forming very small plaques or no plaques under standard conditions. *Appl. Env. Microbiol.* **74**, 5113–5120 (2008).
35. Abedon, S. T. & Yin, J. in *Bacteriophages: Methods and Protocols* Vol. 1 (eds Clokie, M. R. J. & Kropinski, A. M.) 161–174 https://doi.org/10.1007/978-1-60327-164-6_17 (Humana, 2009).
36. Serwer, P., Hayes, S. J., Thomas, J. A. & Hardies, S. C. Propagating the missing bacteriophages: a large bacteriophage in a new class. *Virology J.* **4**, 21 (2007).
37. Wang, J., Yashiro, Y., Sakaguchi, Y., Suzuki, T. & Tomita, K. Mechanistic insights into tRNA cleavage by a contact-dependent growth inhibitor protein and translation factors. *Nucleic Acids Res.* **50**, 4713–4731 (2022).
38. Tomita, K., Ogawa, T., Uozumi, T., Watanabe, K. & Masaki, H. A cytotoxic ribonuclease which specifically cleaves four isoaccepting arginine tRNAs at their anticodon loops. *Proc. Natl Acad. Sci. USA* **97**, 8278–8283 (2000).
39. Takai, K., Takaku, H. & Yokoyama, S. In vitro codon-reading specificities of unmodified tRNA molecules with different anticodons on the sequence background of *Escherichia coli* tRNASer1. *Biochem. Biophys. Res. Commun.* **257**, 662–667 (1999).
40. Takai, K., Okumura, S., Hosono, K., Yokoyama, S. & Takaku, H. A single uridine modification at the wobble position of an artificial tRNA enhances wobbling in an *Escherichia coli* cell-free translation system. *FEBS Lett.* https://doi.org/10.1016/S0014-5793(99)00255-0 (1999).
41. Kunjapur, A. M. et al. Synthetic auxotrophy remains stable after continuous evolution and in coculture with mammalian cells. *Sci. Adv.* **7**, eabf5851 (2021).
42. Rovner, A. J. et al. Recoded organisms engineered to depend on synthetic amino acids. *Nature* **518**, 89–93 (2015).
43. Nyerges, Á. et al. A highly precise and portable genome engineering method allows comparison of mutational effects across bacterial species. *Proc. Natl Acad. Sci. USA* https://doi.org/10.1073/pnas.1520040113 (2016).
44. Wannier, T. M. et al. Adaptive evolution of genomically recoded *Escherichia coli*. *Proc. Natl Acad. Sci. USA* https://doi.org/10.1073/pnas.1715530115 (2018).
45. Kirchberger, P. C. & Ochman, H. Resurrection of a global, metagenomically defined gokushovirus. *eLife* **9**, e51599 (2020).
46. Chen, Y. et al. Multiplex base editing to convert TAG into TAA codons in the human genome. *Nat. Commun.* **13**, 4482 (2022).
47. Boeke, J. D. et al. The Genome Project-Write. *Science* https://doi.org/10.1126/science.aaf6850 (2016).
48. Dai, J., Boeke, J. D., Luo, Z., Jiang, S. & Cai, Y. Sc3.0: revamping and minimizing the yeast genome. *Genome Biol.* **21**, 205 (2020).

Methods

Bacterial media and reagents

Lysogeny broth Lennox (LBL) was prepared by dissolving 10 g l^{-1} tryptone, 5 g l^{-1} yeast extract and 5 g l^{-1} sodium chloride in deionized H_2O and sterilized by autoclaving. Super optimal broth (SOB) was prepared by dissolving 20 g l^{-1} tryptone, 5 g l^{-1} yeast extract, 0.5 g l^{-1} sodium chloride, 2.4 g l^{-1} magnesium sulfate and 0.186 g l^{-1} potassium chloride in deionized H_2O and sterilized by autoclaving. 2×YT medium consisted of 16 g l^{-1} casein digest peptone, 10 g l^{-1} yeast extract, 5 g l^{-1} sodium chloride. LBL and 2×YT agar plates were prepared by supplementing LBL medium or 2×YT with agar at 1.6% w/v before autoclaving. Top agar for agar overlay assays was prepared by supplementing LBL medium with agarose at 0.7% w/v before autoclaving. SM buffer (50 mM Tris-HCl (pH 7.5), 100 mM NaCl, 8 mM $MgSO_4$, 0.01% gelatin) was used for storing and diluting bacteriophage stocks (Geno Technology). bipA was obtained from PepTech Corporation (USA).

Bacteriophage isolation

Bacteriophages were isolated from environmental samples from Massachusetts, USA, collected in the third quarter of 2021 (samples 1, 2 and 4–13) and first quarter of 2022 (samples 1–3; Extended Data Table 1a), and by using *E. coli* Syn61Δ3(ev5) (from the laboratory of Jason W. Chin (Addgene strain no. 174514)) as the host. For aqueous samples, including sewage, we directly used 50 ml filter-sterilized filtrates, whereas samples with mainly solid components, such as soil and animal faeces, were first resuspended to release phage particles and then sterilized by centrifugation and subsequent filtration. Environmental samples and the sterilized filtrates were stored at 4 °C in the dark. This protocol avoided the inactivation of chloroform-sensitive viruses. Sterilized samples were then mixed with exponentially growing cultures of Syn61Δ3(ev5) in SOB supplemented with 10 mM $CaCl_2$ and $MgCl_2$. Infected cultures were grown overnight at 37 °C aerobically and then filter sterilized by centrifugation at 4,000*g* for 15 min and filtered through a 0.45-µm PVDF Steriflip disposable vacuum filter unit (MilliporeSigma). Next, 1 ml from each sterilized enriched culture was mixed with 10 ml exponentially growing Syn61Δ3 (optical density at 600 nm (OD_{600nm}) = 0.2), supplemented with 10 mM $CaCl_2$ and $MgCl_2$, and mixed with 10 ml 0.7% LBL top agar. Top agar suspensions were then poured on top of LBL agar plates in 145 × 20-mm Petri dishes (Greiner Bio-One). Petri dishes were incubated overnight at 37 °C and inspected for phage plaques the next day. Areas with visible lysis or plaques were excised, resuspended in SM buffer, and diluted to single plaques on top agar lawns containing 99% Syn61Δ3 and 1% MDS42 cells. We note that adding trace amounts of MDS42 cells increased the visibility of plaques, and clear plaques, indicating phage replication on the recoded host, could be easily picked. Dilutions and single-plaque isolations were repeated four times for each plaque to purify isogenic phages. Finally, high-titre stocks were prepared by mixing sterilized suspensions from single plaques with exponentially growing MDS42 cells (OD_{600nm} = 0.3) in SOB supplemented with 10 mM $CaCl_2$ and $MgCl_2$. Phage-infected samples were grown at 37 °C until complete lysis (about 4 h) and then sterilized by filtration.

Bacteriophage culturing

Bacteriophage stocks were prepared by a modified liquid lysate Phage on Tap protocol in LBL medium[49]. High-titre lysates were prepared from single plaques by picking well-isolated phage plaques into SM buffer and then seeding 3–50 ml early-exponential-phase cultures of *E. coli* MDS42 cells with the resulting phage suspension in SOB supplemented with 10 mM $CaCl_2$ and $MgCl_2$. Phage-infected samples were grown at 37 °C until complete lysis and then sterilized by filtration. High-titre phage lysates were stored at 4 °C in the dark. Phages were archived as virocells and stored at −80 °C in the presence of 25% glycerol for long-term storage.

Phage replication assay

Phages containing genomic TCR-suppressor $tRNA^{Ser(UGA)}$ genes (based on Supplementary Data 1), corresponding to National Center for Biotechnology Information (NCBI) GenBank numbers MZ501046, MZ501058, MZ501065, MZ501066, MZ501067, MZ501074, MZ501075, MZ501089, MZ501096, MZ501098, MZ501105 and MZ501106 (ref. [24]), were obtained from DSMZ (Germany). Exponential-phase cultures (OD_{600nm} = 0.3) of MDS42 and Syn61Δ3(ev5) were grown in SOB supplemented with 10 mM $CaCl_2$ and $MgCl_2$ at 37 °C. Cultures were infected with phage at a multiplicity of infection (MOI) of approximately 0.001. Simultaneously, the same amount of each phage was added to sterile SOB supplemented with 10 mM $CaCl_2$ and $MgCl_2$ to act as a cell-free control for input phage calculation. Replication assays with the T6 bacteriophage in Syn61Δ3 in Fig. 1d were carried out by infecting exponential-phase cultures at a MOI of 0.01 with T6 phage. The figure shows total T6 titre after 24 h of incubation. Infected cultures were grown at 37 °C with shaking at 250 r.p.m. After 24 h, cultures were transferred to 1-ml tubes and centrifuged at 19,000*g* to remove cells and cellular debris, and the clarified supernatant was serially diluted in SM buffer to enumerate output phage concentration. A 1.5 µl volume of the diluted supernatants was applied to LBL 0.7% top agar seeded with MDS42 cells and 10 mM $CaCl_2$ and $MgCl_2$ using a 96-fixed-pin multi-blot replicator (VP407, V&P Scientific). Following 18 h of incubation at 37 °C, plaques were counted, and the number of plaques was multiplied by the dilution to calculate the phage titre of the original sample.

Single-step phage growth curve

An exponential-phase culture (OD_{600nm} = 0.3) of Syn61Δ3 was grown in 50 ml SOB supplemented with 10 mM $CaCl_2$ and $MgCl_2$ at 37 °C with shaking at 250 r.p.m. Cultures were then spun down and resuspended in 3 ml SOB supplemented with 10 mM $CaCl_2$ and $MgCl_2$, and 1-ml samples were infected with REP12 phage at a MOI of 0.01. Infected cultures were incubated at 37 °C for 10 min without shaking for phage attachment and then washed twice with 1 ml SOB by pelleting cells at 4,000*g* for 3 min. Infected cells were then diluted into 50 ml SOB supplemented with 10 mM $CaCl_2$ and $MgCl_2$ and incubated at 37 °C with shaking at 250 r.p.m. Every 20 min, a 1-ml sample was measured out into a sterile Eppendorf tube containing 100 µl chloroform, immediately vortexed, and then placed on ice. Phage titres were determined by centrifuging chloroform-containing cultures at 6,000*g* for 3 min and then serially diluting supernatants in SM buffer and spotting 1 µl dilutions to LBL 0.7% top agar plates seeded with MDS42 cells and 10 mM $CaCl_2$ and $MgCl_2$. Following 18 h of incubation at 37 °C, plaques were counted, and the number of plaques was multiplied by the dilution to calculate the phage titre of the original sample.

Bacteriophage genome sequencing, assembly and annotation

Genomic DNA of bacteriophages was prepared from high-titre (that is, $>10^{10}$ plaque-forming units (PFUs) ml^{-1}) stocks after DNase treatment using the Norgen Biotek Phage DNA Isolation Kit (catalogue no. 46800) according to the manufacturer's guidelines and sequenced at the Microbial Genome Sequencing Center (MiGS; Pittsburgh, PA, USA). Sequencing libraries were prepared using the Illumina DNA Prep kit and IDT 10-base pair (bp) UDI indices and sequenced on an Illumina NextSeq 2000, producing 150-bp paired-end reads. Demultiplexing, quality control and adapter trimming were carried out with bcl-convert (v3.9.3). Reads were trimmed to Q28 using BBDuk from BBTools. Phage genomes were then assembled de novo using SPAdes 3.15.2 in --careful mode with an average read coverage of 10–50×. Assembled genomes were then annotated using Prokka version 1.14.6 (ref. [50]) with default parameters, except that the PHROGs HMM database[51] was used as input to improve phage functional gene annotations.

Bacterial genome sequencing and annotation

Genomic DNA from overnight saturated cultures of isogenic bacterial clones was prepared using the MasterPure Complete DNA and RNA Purification Kit (Lucigen) according to the manufacturer's guidelines and sequenced at the MiGS (Pittsburgh, PA, USA). Sequencing libraries were prepared using the Illumina DNA Prep kit and IDT 10-bp UDI indices and sequenced on an Illumina NextSeq 2000, producing 150-bp paired-end reads. Demultiplexing, quality control and adapter trimming were carried out with bcl-convert (v3.9.3). Reads were then trimmed to Q28 using BBDuk from BBTools and aligned to their corresponding reference by using Bowtie2 2.3.0 (ref. [52]) in --sensitive-local mode. Single-nucleotide polymorphisms (SNPs) and indels were called using breseq (version 0.36.1)[53]. Only variants with a prevalence higher than 75% were voted as mutations. Following variant calling, mutations were also manually inspected within the aligned sequencing reads in all cases.

The de novo sequencing and genome assembly of Syn61Δ3(ev5) (from a single-colony isolate of Addgene strain no. 174514) was carried out by generating 84,136 Oxford Nanopore (ONT) long reads by PCR-free library generation (Oxford Nanopore) on a MinION Flow Cell (R9.4.1) and 4.5×10^6 150-bp paired-end reads on an Illumina NextSeq 2000. Quality control and adapter trimming were carried out with bcl2fastq 2.20.0.445 and porechop 0.2.3_seqan2.1.1 for Illumina and ONT sequencing, respectively. Next, we carried out hybrid assembly with Illumina and ONT reads by using Unicycler 0.4.8 by using the default parameters. Finally, the resulting single, circular contig representing the entire genome was manually inspected for errors in Geneious Prime 2022.1.1 and annotated on the basis of sequence homology by using the BLAST function (version 2.11.0) implemented in Geneious Prime 2022.1.1 based on *E. coli* K12 MG1655 (NCBI ID: U00096.3) as a reference. Gene essentiality was determined on the basis of ref. [54].

Transcriptome analysis of phage-infected cells

We explored transcriptomic changes and mRNA production in phage-infected Syn61Δ3 cells by carrying out a modified single-step growth experiment and collected samples at 20-min intervals. A 50 ml volume of early-exponential-phase (OD_{600nm} = 0.15) Syn61Δ3 cells (corresponding to 2×10^{10} CFUs) growing at 37 °C, 250 r.p.m. in SOB containing 10 mM $CaCl_2$ and $MgCl_2$ was spun down at room temperature and resuspended in 1 ml SOB. A 50 µl volume of this uninfected sample was immediately frozen in liquid N_2 and stored at −80 °C until RNA extraction. Next, 900 µl of this cell suspension was mixed with 10 ml prewarmed REP12 phage stock (that is, about 7×10^{10} PFUs to achieve a MOI of about 4) in SOB containing 10 mM $CaCl_2$ and $MgCl_2$, and then incubated at 37 °C for 10 min without shaking for phage absorption. Following phage attachment, samples were spun down, washed with 1 ml SOB twice to remove unabsorbed phages, and then resuspended in 10 ml SOB containing 10 mM $CaCl_2$ and $MgCl_2$. Samples were then incubated at 37 °C, 250 r.p.m. After 20- and 40-min post-infection, we spun down a 1-ml cell suspension from each sample, and the cell pellets were frozen in liquid N_2 and stored at −80 °C until RNA extraction. As expected, after 60 min post-infection, no cell pellet was visible. Phage infections were carried out in three independent replicates. Total RNA from frozen samples was extracted by using the RNeasy Mini Kit (Qiagen) according to the manufacturer's instructions and the extracted RNA was DNAse-treated with Invitrogen RNase-free DNAse (Thermo Fisher Scientific). Sequencing library preparation was then carried out using Stranded Total RNA Prep Ligation kit with Ribo-Zero Plus for rRNA depletion and by using 10-bp IDT for Illumina indices (all from Illumina). Sequencing was carried out on a NextSeq 2000 instrument in 2 × 50-bp paired-end mode. Demultiplexing, quality control and adapter trimming were carried out with bcl-convert (v3.9.3). cDNA reads were aligned to their corresponding reference by using Bowtie2 2.3.0 (ref. [52]) in --sensitive-local mode, and read count and expression metrics were determined by using Geneious Prime 2022.1.1 (Biomatters). Finally, differential expression analysis was carried out using DESeq2 (ref. [55]; version 1.38.3) with default settings.

tRNA-seq sample preparation

We explored tRNA expression levels and changes in phage-infected Syn61Δ3 cells by carrying out a modified single-step growth experiment with high MOI and cell mass. An early-exponential-phase culture (OD_{600nm} = 0.2) of Syn61Δ3 cells (corresponding to approximately 5×10^{10} CFUs) growing at 37 °C, 250 r.p.m. in SOB containing 10 mM $CaCl_2$ and $MgCl_2$ was spun down at room temperature and resuspended in 1.1 ml SOB. A 100 µl volume of this uninfected sample was immediately frozen in liquid N_2 and stored at −80 °C until tRNA extraction. Next, 1,000 µl of this cell suspension was mixed with 20 ml prewarmed REP12 phage stock (that is, about 10^{12} PFUs to achieve a MOI of about 20) in SOB containing 10 mM $CaCl_2$ and $MgCl_2$, and then incubated at 37 °C for 10 min without shaking for phage absorption. Following phage attachment, samples were spun down, the supernatant containing unabsorbed phages was removed, and the cell pellet was then resuspended in 7 ml SOB containing 10 mM $CaCl_2$ and $MgCl_2$. Samples were then incubated at 37 °C, 250 r.p.m. Immediately after phage attachment and after 20- and 40-min post-infection, 1-ml cell suspensions from each sample were spun down, and cell pellets were frozen in liquid N_2 and stored at −80 °C until total RNA extraction. Phage infections were carried out in two independent replicates.

We analysed the total tRNA content of Ec_Syn61Δ3-SL cells expressing KP869110.1 viral $tRNA24^{Leu(UGA)}$ and $tRNA24^{Leu(CGA)}$ by pelleting cells from 5 ml mid-exponential-phase (OD_{600nm} = 0.3) culture at 4,000*g* and flash-freezing the cell pellet in liquid N_2.

We extracted tRNAs by lysing samples at room temperature for 30 min in 150 µl lysis buffer containing 8 mg ml^{-1} lysozyme (from chicken egg white, no. 76346-678, VWR), 10 mM Tris-HCl pH 7.5 and 1 µl murine RNase inhibitor (New England Biolabs). Samples were then mixed with 700 µl Qiazol reagent (no. 79306, Qiagen) and incubated for 5 min at room temperature. Next, 150 µl chloroform was added, vortexed and incubated until phase separation. Samples were then spun at 15,000*g* for 15 min in a cooled centrifuge. The supernatant was transferred to a new Eppendorf tube and mixed with 350 µl 70% ethanol. Larger RNA molecules were then bound to an RNeasy MinElute spin column (no. 74204, Qiagen), and the flow-through was mixed with 450 µl of 100% ethanol, and tRNAs were bound to a new RNeasy MinElute spin column. The tRNA fraction was then washed first with 500 µl wash buffer (no. 74204, Qiagen), next with 80% ethanol, and then eluted in RNase-free water. The eluted tRNAs were deacylated in 60 mM pH 9.5 borate buffer (J62154-AK, Alfa Aesar, Thermo Fisher Scientific) for 30 min and then purified using a Micro Bio-Spin P-30 gel column (7326251, Bio-Rad).

tRNA-seq library preparation, sequencing and data analysis

We prepared tRNA cDNA libraries by reverse-transcribing tRNAs using the TGIRT-III template-switching reverse transcriptase (TGIRT50, InGex) according to the manufacturer's instructions. In brief, we prepared reaction mixtures containing 1 µl (about 100 ng) of the deacylated tRNAs, 2 µl of 1 µM TGIRT DNA–RNA heteroduplex (prepared by hybridizing equimolar amounts of rCrUrUrUrGrArGrCrCrUrArArUrGrCrCrUrGrArArArGrArUrCrGrGrArArGrArGrCrArCrArCrGrUrCrUrArGrUrUrCrUrArCrArGrUrCrCrGrArCrGrArU/3SpC3/ and ATCGTCGGACTGTAGAACTAGACGTGTGCTCTTCCGATCTTTCAGGCATTAGGCTCAAAGN oligonucleotides), 4 µl 5× TGIRT reaction buffer (2.25 M NaCl, 25 mM $MgCl_2$, 100 mM Tris-HCl, pH 7.5), 2 µl of 100 mM DTT, 9 µl RNase-free water and 1 µl TGIRT-III, and incubated at room temperature for 30 min to initiate template-switching. Next, 1 µl of 25 mM dNTPs (Thermo Fisher Scientific) was added to the reaction mixture, and samples were incubated at 60 °C for 30 min to carry out reverse transcription. RNA was then hydrolysed by NaOH, and then neutralized by HCl, and

the cDNA library was purified using the MinElute PCR purification kit. cDNAs were then ligated to a preadenylated DNA adapter, /5Phos/GAT CNNNAGATCGGAAGAGCGTCGTGT/3SpC3/ (in which NNN denotes an N, NN or NNN spacer), to increase library diversity during sequencing (preadenylated oligonucleotides were prepared by a 5′ DNA adenylation kit (E2610L) using thermostable 5′ App DNA/RNA ligase (M0319L, both from New England Biolabs) following the manufacturer's protocol). The cDNA library was purified using the MinElute PCR purification kit (Qiagen) and amplified using Q5 Host-Start High-Fidelity 2× Master Mix (New England Biolabs). PCR products were then size selected to remove adapter dimers below 200 bp using three subsequent size-selection rounds with a Select-a-Size DNA Clean & Concentrator Kit (D4080, Zymo Research). Finally, amplicon libraries were barcoded using the IDT 10-bp UDI indices (Illumina) and sequenced on an Illumina MiSeq to produce 250-bp paired-end reads. Read-demultiplexing was carried out with bcl-convert (v3.9.3). Paired-end reads were then aligned to their reference sequences by using Geneious assembler, implemented in Geneious Prime 2022.1.1, allowing a maximum of ten SNPs within tRNA reads compared to their reference. These settings allowed us to map lower-fidelity TGIRT-III-transcribed cDNA reads to their corresponding reference sequence without cross-mapping to tRNAs sharing sequence homology. tRNA reads from Ec_Syn61Δ3-SL cells expressing KP869110.1 viral $tRNA24^{Leu(UGA)}$ and $tRNA24^{Leu(CGA)}$ were mapped without allowing the presence of SNPs in sequencing reads to distinguish $tRNA24^{Leu(UGA)}$ and $tRNA24^{Leu(CGA)}$ that differs by only a single SNP within the anticodon region.

Genome editing and biocontainment of Syn61Δ3

We first generated a deficient recombination variant of Syn61Δ3(ev5) by eliminating the expression of the genomic *recA* gene using Cas9-assisted recombineering. *recA*-deletion experiments were carried out by first transforming Syn61Δ3(ev5) cells with a plasmid carrying a pSC101 origin of replication, a constitutively expressed chloramphenicol resistance marker, SpCas9 and tracrRNA (from pCas9 (refs. [56,57]), Addgene no. 42876), and the λRed operon, consisting of *gam*, *exo* and *bet* (from pORTMAGE311B (ref. [58]), Addgene no. 120418). Next, cells were made electrocompetent using a standard protocol[56,57] for Cas9-assisted recombineering and transformed with 2 µl of 100 µM 90-nucleotide-long single-stranded DNA oligonucleotide inserting a stop codon and a frameshift mutation into *recA* (Supplementary Data 3). Successful edits were selected by cotransforming 1 µg from a variant of the pCRISPR plasmid[56,57] carrying a 5′-AGTTGATACCTTCGCCGTAG-3′ guide sequence to cleave the genomic *recA* sequence in unedited cells. All plasmids were recoded to lack TCR and TAG codons in protein-coding genes, and synthesized by GenScript USA Inc. The resulting Syn61Δ3(ev5) Δ*recA* strain was validated by whole-genome sequencing and then evolved for increased fitness (see the below section entitled Adaptive laboratory evolution of Syn61Δ3). Finally, the replacement of the genomic *adk* gene of Syn61Δ3(ev5) Δ*recA* (ev1) with the bipA-dependent adk.d6 variant[10] was carried out by first transforming cells with a plasmid carrying a constitutively expressed MjTyrRS-derived bipA aaRS (variant 10, based on ref. [59]) together with its associated tRNA under the control of a *proK* tRNA promoter and an aminoglycoside-(3)-*N*-acetyltransferase gene, conferring gentamycin resistance, all on a plasmid containing a p15A origin of replication (Supplementary Data 3). Next, we integrated the adk.d6 variant by Cas9-assisted recombineering as described above, but instead of oligonucleotide-mediated recombineering, we transformed 4 µg of a dsDNA cassette carrying the full-length adk.d6 variant with 400-bp flanking genomic homology (constructed by GenScript USA Inc.; Supplementary Data 3). Cells were grown in the presence of 200 µM bipA in 2×YT medium throughout the entire procedure. Successful edits were selected using a dual-targeting crRNA expression construct, carrying 5′-GCAATGCGTATCATTCTGCT-3′ and 5′-GCCGTCAACTTTCGCGTATT-3′ guide sequences (from GenScript). Positive colonies were selected by screening colonies with allele-specific PCR (Supplementary Data 3) and validated by whole-genome sequencing. Finally, the escape rate of the resulting Syn61Δ3(ev5) Δ*recA* (ev1) adk.d6 strain was determined as described earlier[10], but instead of chloramphenicol, cells were grown in the presence of 10 µg ml^{-1} gentamycin in 2×YT. Plates were incubated for 7 days at 37 °C. Escape rate measurements were carried out in triplicate; ± indicates standard deviation.

Adaptive laboratory evolution of Syn61Δ3

We carried out standard adaptive laboratory evolution in rich bacterial medium for 30 days (about 270 cell generations) on Syn61Δ3(ev5) Δ*recA* cells to increase fitness. At each transfer step, 10^9–10^{10} bacterial cells were transferred into 500 ml LBL medium containing 1.5 g l^{-1} Tris/Tris-HCl and incubated aerobically for 24 h at 37 °C, 250 r.p.m. in a 2,000-ml Erlenmeyer flask with a vented cap. Following 30 transfers, bacterial cells were spread onto LBL agar plates, and an individual colony was isolated and subjected to whole-genome sequencing. The identified mutations in the resulting evolved variant, Syn61Δ3(ev5) Δ*recA* (ev1), are listed in Supplementary Data 3.

Doubling-time measurements

To determine growth parameters under standard laboratory conditions, saturated overnight cultures of *E. coli* Syn61Δ3(ev5), Syn61Δ3(ev5) Δ*recA* and its evolved variant Syn61Δ3(ev5) Δ*recA* (ev1) were diluted 1:200 into 50 ml of 2×YT and LBL in a 300-ml Erlenmeyer flask with vented cap and incubated aerobically at 37 °C, 250 r.p.m. Ec_Syn61Δ3-SL cells were characterized similarly, but by using 2×YT containing 50 µg ml^{-1} kanamycin. All growth measurements were carried out in triplicate. OD_{600nm} measurements were taken every 20 min for 8 h or until stationary phase was reached on a CO8000 Cell Density Meter (WPA). The doubling time was calculated for each independent replicate by log_2-transforming OD_{600nm} values and calculating the doubling time based on every six consecutive data points during the exponential growth phase. We calculated the doubling time (1/slope) from a linear fit to log_2 derivatives of the six data points within this window and reported the shortest doubling time for each independent culture. Curve fitting, linear regression and doubling-time calculations were carried out with Prism9 (GraphPad). Error bars show ±s.d.

tRNA annotation

We detected tRNA genes in the Viral genomic NCBI Reference Sequence Database (accessed 2 January 2022) and in the genomes of individual phage isolates by using tRNAscan-SE 2.0.9 in bacterial (-B), archaeal (-A) or eukaryotic (-E) maximum sensitivity mode (-I -max)[60]. tRNAscan-SE detection parameters were chosen according to the predicted host of the corresponding viral strain.

Mobile tRNAome tRNA library generation and selection

We generated our mobile tRNAome expression library by synthesizing tRNAscan-SE-predicted tRNAs from diverse sources (Supplementary Data 1), driven by a strong bacterial *proK* tRNA promoter and followed by two transcriptional terminators as 10-pmol single-stranded DNA oligonucleotide libraries (10 pmol oPool, from Integrated DNA Technologies). Oligonucleotides were resuspended in 1× TE buffer and then amplified using 5′ phosphorylated primers. Amplicons were then blunt-end ligated into pCR4Blunt-TOPO (Invitrogen, Zero Blunt TOPO PCR Cloning Kit) for 18 h at 16 °C and then purified by using the Thermo Scientific GeneJET PCR Purification Kit. We then electroporated 50 ng purified plasmid in five parallel electroporations into 5 × 40 µl freshly made electrocompetent cells of MDS42 and Syn61Δ3(ev5). Before electrotransformation, bacterial cells were made electrocompetent by growing cells after a 1:100 dilution in SOB until mid-log phase (OD_{600nm} = 0.3) at 32 °C and then washing cells three times using ice-cold water. Electroporated cultures were allowed to recover overnight at 37 °C and then plated to LBL agar plates containing

50 μg ml^{-1} kanamycin in 145 × 20 mm Petri dishes (Greiner Bio-One). Plates were incubated at 37 °C until colony formation. Approximately 1,000–5,000 colonies were then washed off from selection plates, and plasmids were extracted by using the Monarch Plasmid Miniprep Kit (New England Biolabs). The tRNA insert from isolated plasmids was then amplified with primers bearing the standard Nextera Illumina Read 1 and Read 2 primer binding sites, barcoded using the IDT 10-bp UDI indices, and sequenced on an Illumina NextSeq 2000, producing 150-bp paired-end reads. Demultiplexing was carried out with bcl-convert (v3.9.3). Paired-end reads were then trimmed using BBDuk from BBTools (in Geneious Prime 2022.1.1, Biomatters), merged and aligned to their reference sequences by using Geneious assembler, implemented in Geneious Prime 2022.1.1, allowing maximum a single SNV within the tRNA read.

tRNA$^{Leu(YGA)}$ library generation and selection

We identified leucine tRNAs that can translate TCR codons as leucine by carrying out two consecutive screens with plasmid libraries expressing an anticodon-loop-mutagenized 65,536-member library of *leuU* tRNA variants and a smaller, 13-member tRNA$^{Leu(YGA)}$ expression library consisting of bacteriophage-derived leucine tRNA variants, both bearing two tRNAs under the control of a *proK* promoter and with anticodons swapped to UGA and CGA. To construct a 65,536-member *leuU* tRNA library, we synthesized an expression construct consisting of a *proK* promoter–leuU_{UGA}–spacer–leuU_{CGA}–*proK* terminator sequence, in which the anticodon loop of both *leuU* tRNAs has been fully randomized, as an oPool library (Supplementary Data 3; Integrated DNA Technologies). Next, we amplified these *leuU* variants by using Q5 Hot Start High-Fidelity Master mix using 5′-phosphorylated primers and then ligated the library into a plasmid backbone containing a high-copy-number pUC origin of replication and an APH(3′)-I aminoglycoside *O*-phosphotransferase (*aph3Ia*$^{29\times Leu\text{-}to\text{-}TCR}$) gene in which all 29 instances of leucine-coding codons were replaced with TCR serine codons (synthesized as a gBlock dsDNA fragment by Integrated DNA Technologies). The ligation was carried out at a 3:1 insert-to-vector ratio and by using T4 DNA ligase (New England Biolabs) for 16 h at 16 °C according to the manufacturer's instructions. Finally, the ligation product was purified using the GeneJet PCR purification kit (Thermo Fischer Scientific). We constructed the second, 13-member tRNA$^{Leu(YGA)}$ expression library (Supplementary Data 3) consisting of bacteriophage-derived leucine tRNA variants bearing a UGA and CGA anticodon by using the same method as for our *leuU*^{YGA} library. Following library generation, 100 ng from each library was electroporated into freshly made electrocompetent cells of Syn61Δ3 (ev5) Δ*recA* (ev1) and recovered in SOB at 37 °C for 16 h, 250 r.p.m. After recovery, the cells were plated to 2×YT agar plates containing kanamycin at 200 μg ml^{-1} concentration, and selection plates were incubated at 37 °C until colony formation. Finally, plasmids from clones were purified using a Monarch plasmid miniprep kit (New England Biolabs) and subjected to whole-plasmid sequencing (SNPsaurus, Eugene, OR, USA).

Virus resistance analysis of Ec_Syn61Δ3-SL cells

An exponential-phase culture (OD_{600nm} = 0.3) of the corresponding strain was grown in 3 ml SOB supplemented with 10 mM $CaCl_2$ and $MgCl_2$ and 75 μg ml^{-1} kanamycin at 37 °C with shaking. Cultures were then spun down and resuspended in 1 ml SOB supplemented with 10 mM $CaCl_2$ and $MgCl_2$ and infected with a 1:1 mixture of all 12 Syn61Δ3-lytic phage isolates from this study (Extended Data Table 1b) at a MOI of 0.1. Infected cultures were incubated at 37 °C without shaking for 10 min for phage attachment and then washed three times with 1 ml SOB supplemented with 10 mM $CaCl_2$ and $MgCl_2$ by pelleting cells at 4,000*g* for 3 min. Infected cells were then diluted into 4 ml SOB supplemented with 10 mM $CaCl_2$ and $MgCl_2$ and 75 μg ml^{-1} kanamycin and incubated at 37 °C with shaking at 250 r.p.m. After 24 h of incubation, 500-μl samples were measured out into a sterile Eppendorf tube containing 50 μl chloroform, immediately vortexed and then placed on ice. Phage infection experiments were carried out in three independent replicates. Phage titres were determined by centrifuging chloroformed cultures at 6,000*g* for 3 min and then plating 5 μl of the supernatant directly or its appropriate dilutions mixed with 300 μl MDS42 cells in LBL 0.7% top agar with 10 mM $CaCl_2$ and $MgCl_2$. Following 18 h of incubation at 37 °C, plaques were counted, and the number of plaques was multiplied by the dilution to calculate the phage titre of the original sample.

Phage enrichment experiments were carried out by mixing 50 ml early-exponential-phase cultures (OD_{600nm} = 0.2) of bipA-biocontained Ec_Syn61Δ3-SL carrying pLS1 and pLS2 plasmids with 10 ml environmental sample mix, containing the mixture of samples 2–13 from our study (Extended Data Table 1b). We maximized phage diversity in the environmental sample mix by using freshly collected and filter-sterilized sewage. Infected Ec_Syn61Δ3-SL cells with the corresponding plasmid were grown overnight in SOB supplemented with 200 mM bipA, 10 mM $CaCl_2$ and $MgCl_2$, and 75 μg ml^{-1} kanamycin at 37 °C with shaking at 250 r.p.m. On the next day, cells were removed by centrifugation at 4,000*g* for 20 min, and the supernatant was filter-sterilized using a 0.45-μm filter. Next, 5 ml of the sterilized sample was mixed again with 50 ml early-exponential-phase cultures (OD_{600nm} = 0.2) of the corresponding strain, incubated for 20 min at 37 °C for phage absorption, pelleted by centrifugation at 4,000*g* for 15 min, and then resuspended in 50 ml SOB supplemented with 200 mM bipA, 10 mM $CaCl_2$ and $MgCl_2$, and 75 μg ml^{-1} kanamycin. Infected cultures were then incubated at 37 °C with shaking at 250 r.p.m. Cultures were grown overnight and then sterilized by centrifugation at 4,000*g* for 15 min and filtered through a 0.45-μm PVDF Steriflip disposable vacuum filter unit (MilliporeSigma). Finally, phage titres were determined by using MDS42 cells as above. Phage enrichment experiments were carried out in two independent replicates. The lytic phage titre of the unenriched sample mix was determined by diluting 100 μl of the input environmental sample mix, containing a mixture of samples 2–13 (Extended Data Table 1), in SM buffer and 10 μl samples from each dilution steps were mixed with late-exponential-phase MDS42 cells and 4 ml 0.7% top agar, and then poured on top of LBL agar plates. Plates were incubated until plaque formation at 37 °C.

Construction of pLS plasmids

All pLS plasmids listed in Supplementary Data 3 were synthesized as gBlocks by IDT and circularized either by ligation with T4 DNA ligase (New England Biolabs), or, in the case of pSC101 and RK2 plasmid-derived variants, by isothermal assembly using the HiFi DNA Assembly Master Mix (New England Biolabs). Following assembly, purified plasmid assemblies were electroporated into Ec_Syn61Δ3-SL cells carrying pLS1. pLS1 and pLS2 were designed to express two distinct combinations of the previously identified phage tRNA$^{Leu(YGA)}$ tRNAs in antiparallel orientation to avoid repeat-mediated instability[61], driven by the strong bacterial *proK* and SLP2018-2-101 promoters, together with a high-copy-number pUC origin of replication, the *aph3Ia*$^{29\times Leu\text{-}to\text{-}TCR}$ and aminoglycoside-(3)-*N*-acetyltransferase$_{18\times Leu\text{-}to\text{-}TCR}$ marker genes. Transformants carrying either pLS1 or pLS2 were identified by transforming assemblies into Syn61Δ3(ev5) Δ*recA* (ev1) and selecting for kanamycin resistance. Finally, plasmids from antibiotic-resistant clones were purified using a Monarch plasmid miniprep kit (New England Biolabs) and subjected to whole-plasmid sequencing (SNPsaurus, Eugene, OR, USA).

Escape rate analysis of viral Leu-tRNA^{YGA} and pLS plasmids

We analysed the ability of pLS plasmids to function outside Ec_Syn61Δ3-SL cells by transforming extracted plasmids into *E. coli* K12 MG1655. Plasmids were purified from biocontained Ec_Syn61Δ3-SL cells, carrying either pLS1 or pLS2 to express tRNA$^{Leu(YGA)}$, or pLS1 together with pLS3-5, by using the PureLink Fast Low-Endotoxin Midi Plasmid Purification Kit (Thermo Fisher Scientific) according to the manufacturer's instructions. Next, we electroporated 1 μg from each

plasmid preparation into freshly made electrocompetent cells of *E. coli* K12 MG1655. Cells were made electrocompetent by diluting an overnight SOB culture of MG1655 1:100 into 500 ml SOB in a 2,000-ml flask and growing cells aerobically at 32 °C with shaking at 250 r.p.m. At OD_{600nm} = 0.3, cells were cooled on ice and then pelleted by centrifugation and resuspended in 10% glycerol-in-water. Cells were washed four times with 10% glycerol-in-water and then resuspended in 400 µl 20% glycerol-in-water. A 1,000 ng quantity from each plasmid sample was then mixed with 80 µl electrocompetent cells and electroporated by using standard settings in two 1-mm electroporation cuvettes by using standard electroporator settings (1.8 kV, 200 Ohm, 25 µF). Electroporations were carried out in three independent replicates. Electroporated cells were then resuspended in 1 ml SOB, and the culture was allowed to recover overnight at 37 °C with shaking at 250 r.p.m. Finally, 500 µl from each recovery culture was plated to LBL agar plates containing antibiotics corresponding to the given pLS plasmid's resistance marker (15 µg ml^{-1} gentamycin plus 50 µg ml^{-1} kanamycin in the case of pLS1 and 2; 100 µg ml^{-1} carbenicillin for pLS4, 30 µg ml^{-1} chloramphenicol for pLS3 and pLS6, and 20 µg ml^{-1} gentamycin for pLS5) in 145 × 20-mm Petri dishes (Greiner Bio-One). Plates were incubated at 37 °C for 7 days and inspected for growth. Electroporation efficiency measurements were carried out by electroporating a plasmid carrying a pUC origin of replication and kanamycin resistance (pUC-Kan^R) into MG1655 electrocompetent cells under identical conditions.

Cloning of REP12 tRNA operon

We analysed the incorporated amino acid in Syn61Δ3 cells bearing the tRNA operon and its native promoter from the REP12 phage by subcloning the genomic tRNA operon into a low-copy-number plasmid containing an RK2 origin of replication and a chloramphenicol-acetyltransferase marker, both recoded to contain no TCR or TAG codons. The genomic tRNA operon of REP12 was PCR amplified using the Q5 Hot Start Master Mix (New England Biolabs) from extracted phage gDNA and purified using the GeneJet PCR purification kit (Thermo Fischer Scientific). Next, 100 ng of the amplified tRNA operon was assembled into the linearized pRK2-cat backbone using the HiFi DNA Assembly Master Mix (New England Biolabs). After incubation for 60 min at 50 °C, the assembly was purified using the DNA Concentrator & Clean kit (Zymo Research) and transformed into Syn61Δ3(ev5) Δ*recA* (ev1) cells expressing an MSKGPGKVPGAGVPGXGVPGVGKGGGT-elastin peptide fused to sfGFP with a terminal 6×His tag (in which X denotes the analysed codon, TCA or TCG) on a plasmid containing a kanamycin resistance gene and a pUC origin of replication[62]. Following an overnight recovery at 32 °C, cells were plated to 2×YT agar plates containing kanamycin and chloramphenicol. Finally, plasmid sequences in outgrowing colonies were validated by whole-plasmid sequencing. Elastin(16TCR)–sfGFP–His6 expression measurements have been carried out as described below.

Elastin(16TCR)–sfGFP–His6 expression measurements

We assayed the amino acid identity of the serine $tRNA^{UGA}$ tRNAs of MZ501075 and REP12 (see Supplementary Data 3 for sequence information) by coexpressing selected tRNAs and a constitutively expressed MSKGPGKVPGAGVPGXGVPGVGKGGGT-elastin peptide fused to sfGFP with a terminal 6×His tag (in which X denotes the analysed codon, TCA or TCG)[62] on a plasmid containing a kanamycin resistance gene and a pUC origin of replication in Syn61Δ3(ev5). Similarly, the pRK2-REP12 plasmid carrying the REP12 phage tRNA operon under the control of its native promoter was coexpressed with the same pUC plasmid carrying no tRNA genes. As a control, we used the same elastin–sfGFP–His6 expression construct in which position X has been replaced with an alanine GCA codon. For fluorescence and MS/MS measurements, we diluted cultures 1:100 from overnight starters into 50 ml 2×YT in 300-ml shake flasks containing the corresponding antibiotics and cultivated for 48 h at 37 °C, 200 r.p.m. aerobically. We then determined sfGFP expression levels in samples by pelleting and washing 1 ml of the culture with phosphate-buffered saline and resuspending cell pellets in 110 µl BugBuster Protein Extraction Reagent (MilliporeSigma). Reactions were incubated for 5 min and then spun down at 13,000 r.p.m. for 10 min. The fluorescence of the BugBuster-treated supernatants and the OD_{600nm} of the original culture were measured using the Synergy H1 Hybrid Reader (BioTek) plate reader using bottom-mode analysis with excitation at 480 nm and emission measurement at 515 nm, with the gain set to 50. Fluorescence values were normalized on the basis of OD_{600nm} data. The remaining 49 ml culture was spun down, and the cell pellet was resuspended in 2 ml BugBuster Protein Extraction Reagent (MilliporeSigma) and incubated at room temperature for 5 min. The lysed cell mixture was spun down at 13,000 r.p.m. for 10 min, and the supernatant was mixed in a 1:1 ratio with His-Binding/Wash Buffer (G-Biosciences) and 50 µl HisTag Dynabeads (Thermo Fischer Scientific). Following an incubation period of 5 min, the beads were separated on a magnetic rack and washed with 300 µl His-Binding/Wash Buffer and phosphate-buffered saline three times. After the last wash step, the bead pellets containing the bound elastin–sfGFP–His6 protein samples were frozen at −80 °C until MS/MS sample preparation. Protein production experiments were carried out in three independent replicates.

LC/MS–MS analysis of tryptic elastin–sfGFP–His6

Samples from elastin–sfGFP–His6 expression experiments were digested directly on HisTag Dynabeads according to the FASP digest procedure[63]. In brief, samples were washed with 50 mM triethylammonium bicarbonate (TEAB) buffer and then rehydrated with 50 mM TEAB–trypsin solution, followed by a 3-h digest at 50 °C. Digested peptides were then separated from HisTag Dynabeads and concentrated by spinning and drying samples at 3,000 r.p.m. using a SpeedVac concentrator. Samples were then solubilized in 0.1% formic acid-in-water for subsequent analysis by MS/MS. LC–MS/MS analysis of digested samples was carried out on a Lumos Tribrid Orbitrap mass spectrometer equipped with an Ultimate 3000 nano-HPLC (both from Thermo Fisher Scientific). Peptides were separated on a 150-µm-inner-diameter microcapillary trapping column packed first with 2 cm of C18 Reprosil resin (5 µm, 100 Å, from Dr. Maisch) followed by a 50-cm analytical column (PharmaFluidics). Separation was achieved by applying a gradient from 4% to 30% acetonitrile in 0.1% formic acid over 60 min at 200 nl min^{-1}. Electrospray ionization was carried out by applying a voltage of 2 kV using a custom electrode junction at the end of the microcapillary column and sprayed from metal tips (PepSep). The MS survey scan was carried out in the Orbitrap in the range of 400–1,800 m/z at a resolution of 6×10^4, followed by the selection of the 20 most intense ions for fragmentation using collision-induced dissociation in the second MS step (CID-MS2 fragmentation) in the ion trap using a precursor isolation width window of 2 m/z, automatic gain control setting of 10,000 and a maximum ion accumulation of 100 ms. Singly charged ion species were excluded from CID fragmentation. The normalized collision energy was set to 35 V and an activation time of 10 ms. Ions in a 10-ppm m/z window around ions selected for MS/MS were excluded from further selection for fragmentation for 60 s.

The raw data were analysed using Proteome Discoverer 2.4 (Thermo Fisher Scientific). Assignment of MS/MS spectra was carried out using the Sequest HT algorithm by searching the data against a protein sequence database, including all protein entries from *E. coli* K12 MG1655, all proteins sequences of interest (including the elastin–sfGFP fusion protein), and other known contaminants such as human keratins and common laboratory contaminants. Quantitative analysis between samples was carried out by label-free quantification between different samples. Sequest HT searches were carried out using a 10-ppm precursor ion tolerance and requiring the N and C termini of each peptide to conform with trypsin protease specificity while allowing up to two missed cleavages. Methionine oxidation (+15.99492 Da), deamidation (+0.98402 Da) of asparagine and glutamine amino acids, phosphorylation at serine, threonine and tyrosine amino acids (+79.96633 Da) and

N-terminal acetylation (+42.01057 Da) were set as variable modifications. We then determined the amino acid incorporated at position X in our elastin–sfGFP–6×His construct by analysing changes compared to phenylalanine. To cover all 20 possible amino acid exchange cases at the X position, we carried out five separate searches with four different amino acids as possible variable modifications in each search. All cysteines were analysed without modification as no alkylation procedure was included in our workflow. An overall false discovery rate of 1% on both protein and peptide level was achieved by carrying out target–decoy database search using Percolator[64].

Total proteome analysis and the detection of serine-to-leucine mistranslation events

We analysed the translation of viral proteins in Ec_Syn61Δ3-SL cells (Syn61Δ3(ev5) ΔrecA (ev1) expressing a *proK* promoter-driven Leu9-tRNA^{YGA} construct from the *Escherichia* phage OSYSP (GenBank ID MF402939.1) and APH(3′)-I aminoglycoside *O*-phosphotransferase (*aph3Ia*$^{29\times Leu\text{-}to\text{-}TCR}$), on a high-copy-number pUC plasmid), by carrying out a modified single-step growth experiment and subsequent time-course MS/MS-based proteome analysis. An early-exponential-phase culture (OD_{600nm} = 0.2) of Ec_Syn61Δ3-SL cells (corresponding to approximately 4×10^{10} CFUs) growing at 37 °C, 250 r.p.m. in SOB containing 10 mM $CaCl_2$, $MgCl_2$ and 75 μg ml^{-1} kanamycin was spun down at room temperature and resuspended in 1.1 ml SOB containing 10 mM $CaCl_2$, $MgCl_2$ and 75 μg ml^{-1} kanamycin. A 100 μl volume of this uninfected sample was immediately frozen in liquid N_2 and stored at −80 °C until proteome analysis. Next, 1,000 μl of this cell suspension was mixed with 10 ml prewarmed REP12 phage stock (that is, about 5×10^{11} PFUs to achieve a MOI of about 12) in SOB containing 10 mM $CaCl_2$, $MgCl_2$ and 75 μg ml^{-1} kanamycin, and then incubated at 37 °C for 10 min without shaking for phage absorption. Following phage attachment, samples were spun down, the supernatant containing unabsorbed phages was removed, and the cell pellet was resuspended in 5 ml SOB containing 10 mM $CaCl_2$ and $MgCl_2$. Samples were then incubated at 37 °C, 250 r.p.m. After 20- and 40-min post-infection, 1-ml cell suspensions were spun down, and cell pellets were frozen in liquid N_2 and stored at −80 °C until total protein extraction. Samples from control and phage-infected Syn61Δ3 cells were then digested by using the FASP digest procedure[63]. In brief, samples were washed with 50 mM TEAB buffer on a 10-kDa-cutoff filter (Pall Corp) and then rehydrated with 50 mM TEAB–trypsin solution, followed by a 3-h digestion at 37 °C. Digested peptides were then extracted and separated into ten fractions by using the Pierce High pH Reversed-Phase Peptide Fractionation Kit according to the manufacturer's protocol (Thermo Fisher Scientific). Following fractionation, peptides were concentrated and dried by spinning samples at 3,000 r.p.m. using a SpeedVac concentrator. Samples were then solubilized in 0.1% formic acid-in-water for subsequent analysis by MS/MS. LC–MS/MS analysis of digested samples was carried out on a Lumos Tribrid Orbitrap Mass Spectrometer equipped with an Ultimate 3000 nano-HPLC (both from Thermo Fisher Scientific). Peptides were separated on a 150-μm-inner-diameter microcapillary trapping column packed first with 2 cm of C18 Reprosil resin (5 μm, 100 Å, from Dr. Maisch) followed by a 50-cm analytical column (PharmaFluidics). Separation was achieved by applying a gradient from 5% to 27% acetonitrile in 0.1% formic acid over 90 min at 200 nl min^{-1}. Electrospray ionization was carried out by applying a voltage of 2 kV using a custom electrode junction at the end of the microcapillary column and sprayed from metal tips (PepSep). The MS survey scan was carried out in the Orbitrap in the range of 400–1,800 m/z at a resolution of 6×10^4, followed by the selection of the 20 most intense ions for fragmentation using collision-induced dissociation in the second MS step (CID-MS2 fragmentation) in the ion trap using a precursor isolation width window of 2 m/z, automatic gain control setting of 10,000 and a maximum ion accumulation of 100 ms. Singly charged ion species were excluded from CID fragmentation. The normalized collision energy was set to 35 V and an activation time of 10 ms. Ions in a 10-ppm m/z window around ions selected for MS/MS were excluded from further selection for fragmentation for 60 s.

The raw data were analysed using Proteome Discoverer 2.4 (Thermo Fisher Scientific). Assignment of MS/MS spectra was carried out using the Sequest HT algorithm by searching the data against a protein sequence database, including all protein entries from *E. coli* K12 MG1655, all protein sequences of the corresponding REP12 bacteriophage and the Aph3Ia APH(3′)-I aminoglycoside *O*-phosphotransferase, as well as other known contaminants such as human keratins and common laboratory contaminants. Quantitative analysis between samples was carried out by label-free quantification between different samples. Sequest HT searches were carried out using a 10-ppm precursor ion tolerance and requiring the N and C termini of each peptide to conform with trypsin protease specificity while allowing up to two missed cleavages. Methionine oxidation (+15.99492 Da), deamidation (+0.98402 Da) of asparagine and glutamine amino acids, phosphorylation at serine, threonine and tyrosine amino acids (+79.96633 Da) and N-terminal acetylation (+42.01057 Da) were set as variable modifications. Special modification of serine-to-leucine amino acid exchange (+26.052036 Da) on all serine amino acid positions was used as a variable modification. All cysteines were analysed withouth modification as no alkylation procedure was included in our workflow. An overall false discovery rate of 1% on both protein and peptide levels was achieved by carrying out target–decoy database search using Percolator[64].

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Raw data from whole-genome sequencing, transcriptome and tRNA-seq experiments have been deposited to the Sequence Read Archive under the BioProject ID PRJNA856259. tRNA data and the generated sequences in this study are included in the Supplementary Data files. Mass spectra and proteome measurements have been deposited to MassIVE (MSV000089854; https://doi.org/10.25345/C5FF3M41W)[65]. The Syn61Δ3(ev5) Δ*recA* (ev1) strain is available from Addgene (bacterial strain no. 189857). We cannot deposit Ec_Syn61Δ3-SL and Ec_Syn61Δ3 adk.d6 at Addgene owing to the incompatibility of Addgene's methods of strain distribution and growth medium requirements, but all materials used in this study are freely available from the corresponding authors for academic research use upon request. The PHROG HMM database is available at https://phrogs.lmge.uca.fr and from ref. [51]. The assembled annotated genome of *E. coli* Syn61 substrain Syn61Δ3(ev5) is available in Supplementary Data 4, and the annotated genomes of REP phages have been deposited to NCBI GenBank under the accession numbers OQ174500, OQ174501, OQ174502, OQ174503, OQ174504, OQ174505, OQ174506, OQ174507, OQ174508, OQ174509, OQ174510 and OQ174511. Source data are provided with this paper.

49. Bonilla, N. et al. Phage on tap–a quick and efficient protocol for the preparation of bacteriophage laboratory stocks. *PeerJ* **4**, e2261 (2016).
50. Seemann, T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* **30**, 2068–2069 (2014).
51. Terzian, P. et al. PHROG: families of prokaryotic virus proteins clustered using remote homology. *NAR Genomics Bioinform.* **3**, lqab067 (2021).
52. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
53. Deatherage, D. E. & Barrick, J. E. Identification of mutations in laboratory-evolved microbes from next-generation sequencing data using breseq. *Methods Mol. Biol.* **1151**, 165–188 (2014).
54. Goodall, E. C. A. et al. The essential genome of *Escherichia coli* K-12. *mBio* **9**, e02096-17 (2018).
55. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
56. Jiang, W., Bikard, D., Cox, D., Zhang, F. & Marraffini, L. A. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat. Biotech.* **31**, 233–239 (2013).

57. Umenhoffer, K. et al. Genome-wide abolishment of mobile genetic elements using genome shuffling and CRISPR/Cas-assisted MAGE allows the efficient stabilization of a bacterial chassis. *ACS Synth. Biol.* **6**, 1471–1483 (2017).
58. Szili, P. et al. Rapid evolution of reduced susceptibility against a balanced dual-targeting antibiotic through stepping-stone mutations. *Antimicrob. Agents and Chemother.* **63**, e00207-19 (2019).
59. Kunjapur, A. M. et al. Engineering posttranslational proofreading to discriminate nonstandard amino acids. *Proc. Natl Acad. Sci. USA* **115**, 619–624 (2018).
60. Chan, P. P., Lin, B. Y., Mak, A. J. & Lowe, T. M. tRNAscan-SE 2.0: improved detection and functional classification of transfer RNA genes. *Nucleic Acids Res.* **49**, 9077–9096 (2021).
61. Hossain, A. et al. Automated design of thousands of nonrepetitive parts for engineering stable genetic systems. *Nat. Biotechnol.* **38**, 1466–1475 (2020).
62. Mohler, K. et al. MS-READ: quantitative measurement of amino acid incorporation. *Biochim. Biophys. Acta Gen. Subj.* **1861**, 3081–3088 (2017).
63. Wiśniewski, J. R., Zougman, A., Nagaraj, N. & Mann, M. Universal sample preparation method for proteome analysis. *Nat. Methods* **6**, 359–362 (2009).
64. Käll, L., Storey, J. D. & Noble, W. S. Non-parametric estimation of posterior error probabilities associated with peptides identified by tandem mass spectrometry. *Bioinformatics* **24**, i42–i48 (2008).
65. Nyerges, A. et al. Swapped genetic code block viral infections and gene transfer (MSV000089854). MassIVE https://doi.org/10.25345/C5FF3M41W (2023).

Acknowledgements We thank G. Pósfai (Biological Research Centre, Hungary) for sharing MDS42 and J. W. Chin's team (Medical Research Council Laboratory of Molecular Biology, UK) for sharing Syn61Δ3 through Addgene. Financial support for this research was provided by the US Department of Energy under grant DE-FG02-02ER63445 and by National Science Foundation award number 2123243 (both to G.M.C.). M.B. acknowledges support from the NIGMS of the National Institutes of Health (R35GM133700), the David and Lucile Packard Foundation, the Pew Charitable Trusts and the Alfred P. Sloan Foundation. A.N. was supported by the EMBO LTF 160–2019 Long-Term Fellowship. We thank A. Millard's laboratory for making the PHROG HMM database available for bacteriophage annotation, GenScript USA Inc. for their DNA synthesis support, and D. Snyder, K. Harris and all members of the MiGS, Pittsburgh, PA for their support with DNA and RNA sequencing. We are grateful to T. Wu for support, Y. Shen and S. Yan (Institute of Biochemistry, Beijing Genomics Institute) for collaboration on genome recoding, and B. Hajian for graphical design and help with illustrations.

Author contributions A.N. developed the project, led analyses and wrote the manuscript with input from all authors. A.N. and G.M.C. supervised research. S.V. carried out $tRNA^{Leu}$ suppressor screens and sfGFP expression assays, and assisted in the construction of the pLS plasmids and biocontainment experiments. R.F. assisted in experiments, and carried out adaptive laboratory evolution and growth rate measurements. S.V.O., E.A.R. and M.B. provided environmental samples for phage isolation, carried out replication assays and provided support for phage experiments and genome analyses. M.L., K.C. and F.H. carried out DNA synthesis; M.B.-T., A.C.-P. and E.K. supported the project. B.B. carried out MS/MS analyses. K.N. and J.A.M. provided reagents for tRNA-seq experiments.

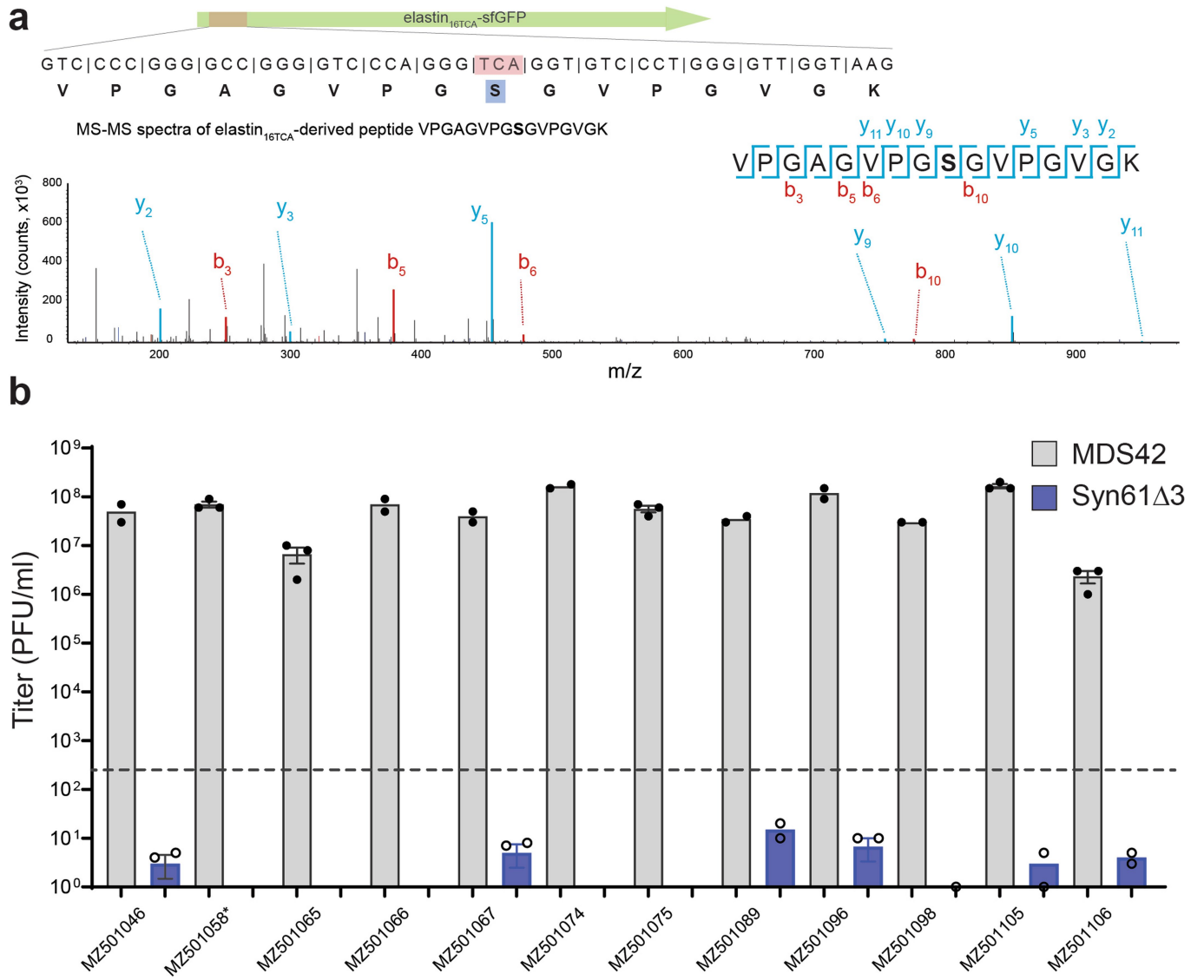
Competing interests Harvard Medical School has filed a provisional patent application related to this work on which A.N., S.V. and G.M.C. are listed as inventors. M.L., K.C. and F.H. are employed by GenScript USA Inc., but the company had no role in designing or executing experiments. G.M.C. is a founder of the following companies in which he has related financial interests: GRO Biosciences, EnEvolv (Ginkgo Bioworks) and 64x Bio. Other potentially relevant financial interests of G.M.C. are listed at http://arep.med.harvard.edu/gmc/tech.html.

Additional information

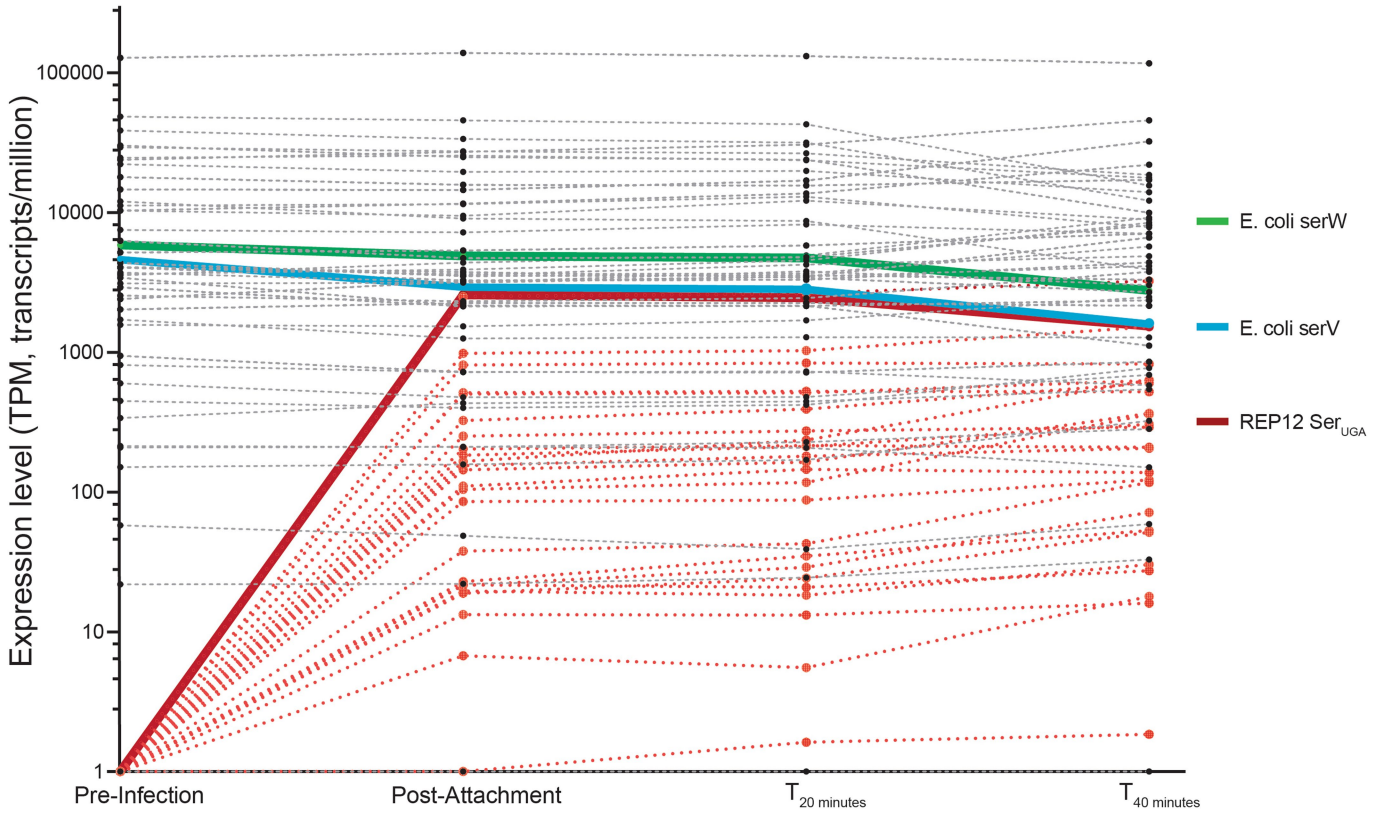
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41586-023-05824-z.

Correspondence and requests for materials should be addressed to Akos Nyerges or George M. Church.

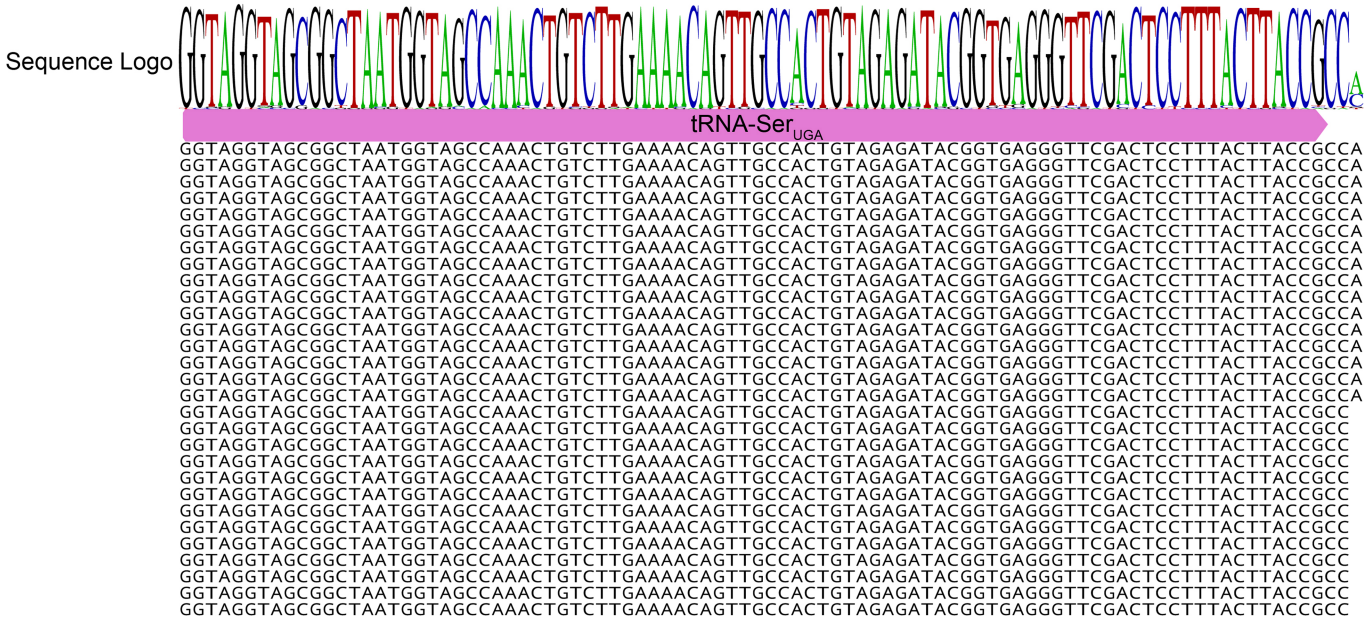
Peer review information *Nature* thanks Benjamin Blount and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.



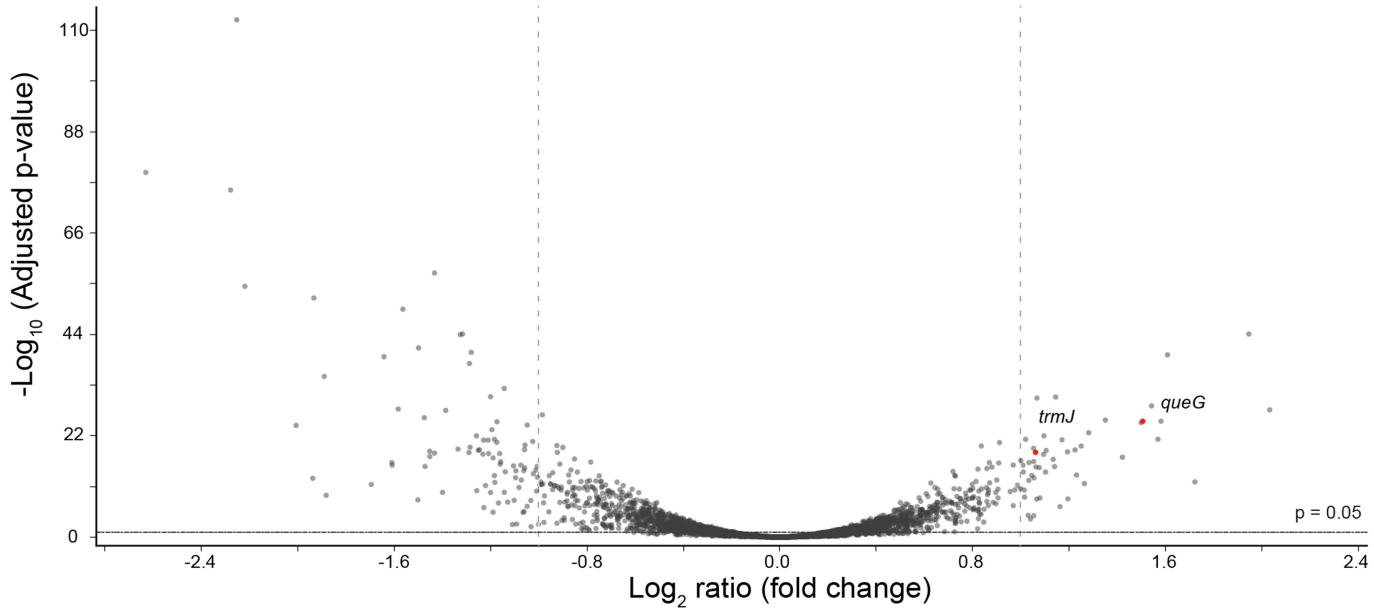
Extended Data Fig. 1 | Viral serine tRNAs decode TCA codons as serine, but Syn61Δ3 obstructs the replication of viruses containing genomic tRNA-Ser$_{UGA}$. a) Viral TCR suppressor tRNAs decode TCA codons as serine. The amino acid identity of the translated TCR codon within elastin$_{16\,TCA}$-sfGFP-His6 was confirmed by tandem mass spectrometry from Syn61Δ3 expressing the tRNA-Ser$_{UGA}$ of *Escherichia* phage IrisVonRoten (GenBank ID MZ501075)[24]. The figure shows the amino acid sequence and MS/MS spectrum of the analyzed elastin$_{16\,TCA}$ peptide. MS/MS data was collected once. **b)** Syn61Δ3 obstructs the replication of viruses containing genomic tRNA-Ser$_{UGA}$. Figure shows the titer of twelve tRNA gene-containing coliphages[24], after 24 h of growth on MDS42 and Syn61Δ3. All analyzed bacteriophages, except MZ501058, contain a genomic tRNA-Ser$_{UGA}$ tRNA that provides TCR suppressor activity based on our screen (Fig. 1b, Supplementary Data 1). Early exponential phase cultures of MDS42 and Syn61Δ3 were infected at an MOI of 0.001 with the corresponding phages, and free phage titers were determined after 24 h of incubation. Measurements were performed in n = 3 independent experiments (*i.e.*, MDS42 + MZ501058, MZ501065, MZ501075, MZ501105, MZ501106; Syn61Δ3 + MZ501046, MZ501067, MZ501066, MZ501096, MZ501074, MZ501098) or in n = 2 independent experiments (*i.e.*, MDS42 + MZ501046, MZ501066, MZ501067, MZ501074, MZ501089, MZ501096, MZ501098; Syn61Δ3 + MZ501058, MZ501065, MZ501075, MZ501089, MZ501105, MZ501106); dashed line represents input phage titer without bacterial cells (*i.e.*, a titer of 420 PFU/ml); dots represent data from independent experiments; bar graphs represent the mean; error bars represent the SEM based on n = 3 independent experiments.



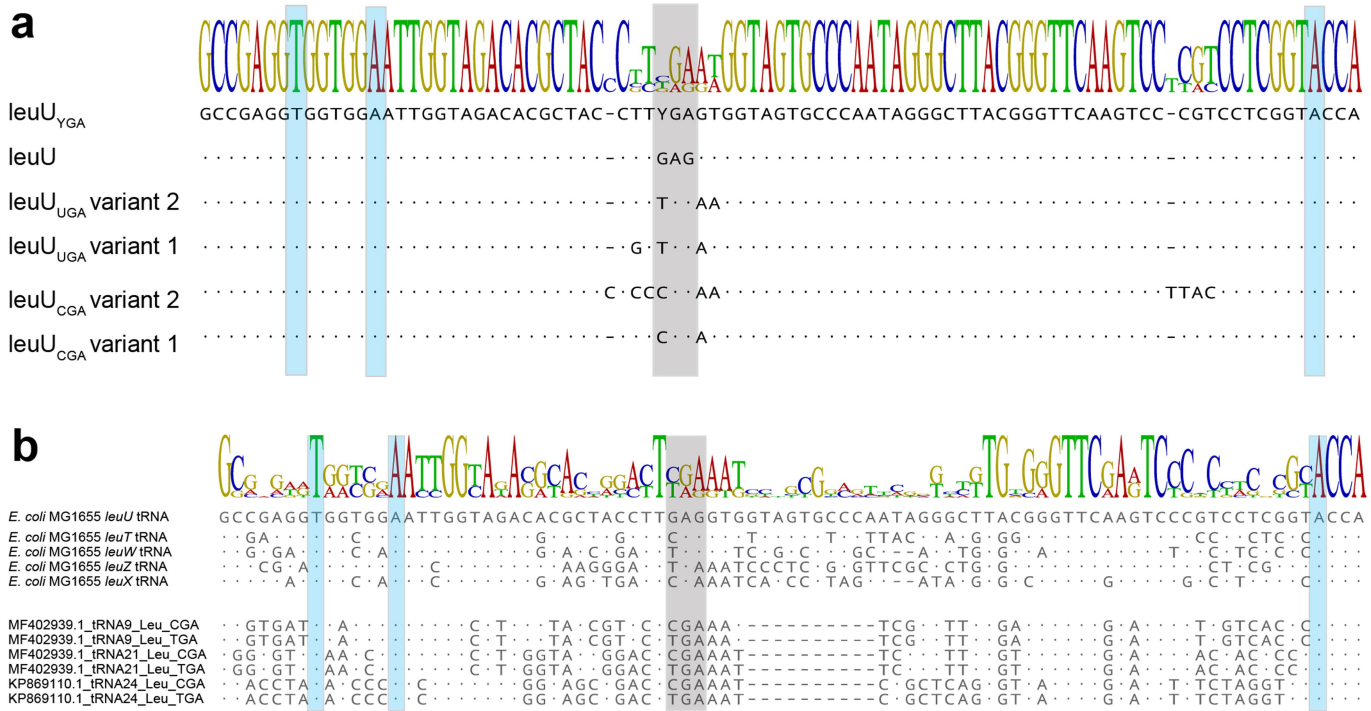
Extended Data Fig. 2 | Viruses overcome genetic-code-based resistance by rapidly complementing the cellular tRNA pool with virus-encoded tRNAs. The time-course kinetics of host and viral tRNA expression in Syn61Δ3 cells following REP12 phage infection was quantified using tRNAseq (Methods). The endogenous *serV* and *serW* tRNAs of the host Syn61Δ3 are highlighted in green and blue, respectively, while the tRNA-Ser_{UGA} of the REP12 virus is highlighted in red. REP12 viral tRNAs are shown in light red; endogenous tRNAs of Syn61Δ3 are shown in gray. Data represent mean TPM (transcript/million). Source data is available within this paper.



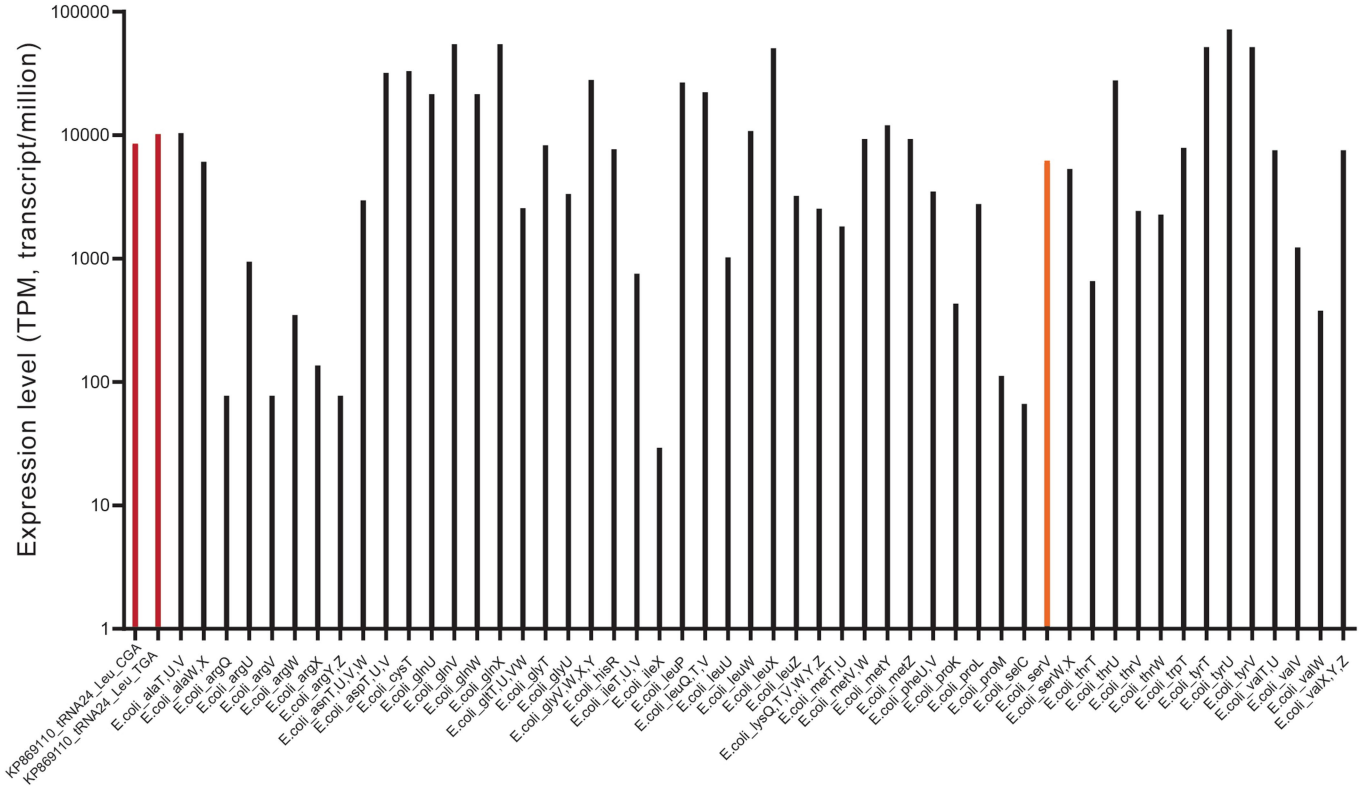
Extended Data Fig. 3 | tRNAseq-based detection of CCA tRNA tail addition to REP12 tRNA-Ser$_{UGA}$. The tRNA tail is modified into CCA even if the phage-encoded tRNA-Ser$_{UGA}$ does not encode a CCA end. The sequence of genomic phage-encoded tRNA-Ser$_{UGA}$ is highlighted in magenta. Black letters indicate example tRNAseq sequencing reads from REP12 infected Syn61Δ3 cells directly after phage attachment based on a single experiment.



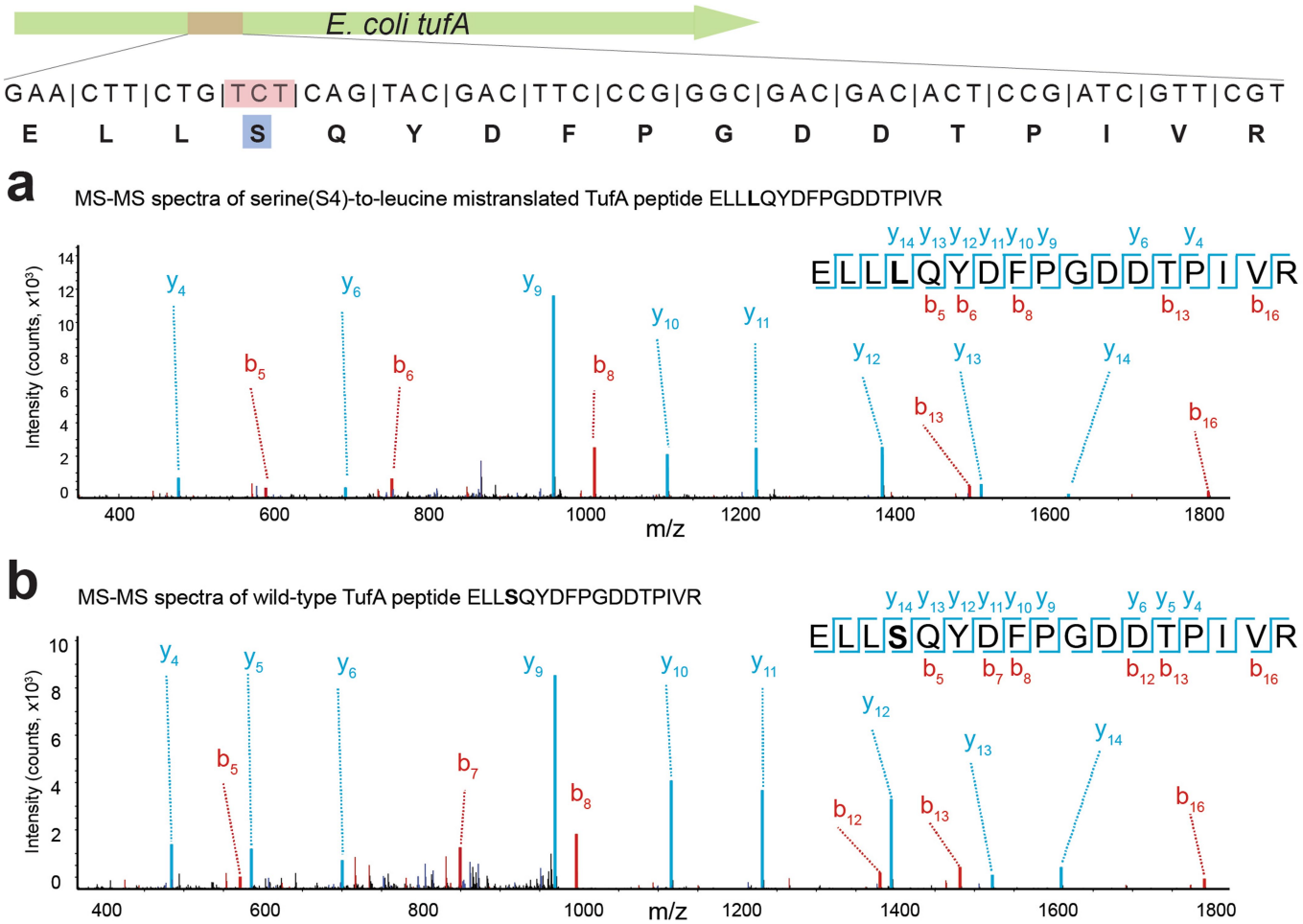
Extended Data Fig. 4 | Transcriptomic changes in Syn61Δ3 following phage infection. Volcano plot shows the differential expression between uninfected and REP12-infected Syn61Δ3 cells, 40 min post-infection, based on n = 3 independent experiments (Methods). *QueG* encodes epoxyqueuosine reductase that catalyzes the final step in the de novo synthesis of queuosine in tRNAs. *TrmJ* encodes tRNA Cm32/Um32 methyltransferase that introduces methyl groups at the 2′-O position of U32 of several tRNAs, including tRNA-Ser_{UGA}. Differential expression and -Log_{10} adjusted p-values were calculated using the DESeq2 algorithm[55]. Source data is available within this paper.



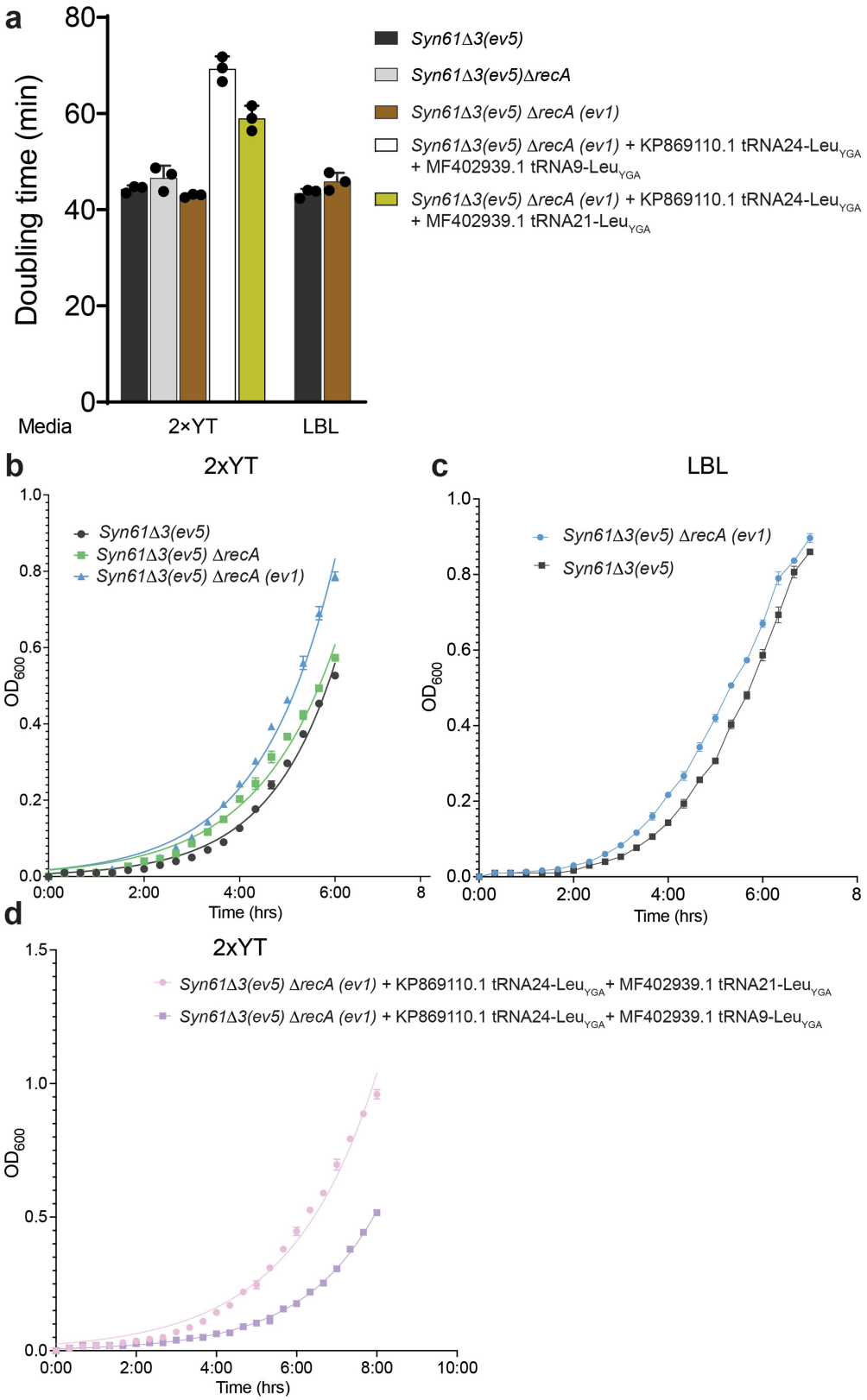
Extended Data Fig. 5 | Multiple sequence alignment of *leuU*$_{YGA}$ and phage-derived tRNA-Leu$_{YGA}$ variants. a) Multiple sequence alignment of *leuU*$_{YGA}$ variants selected in aminoglycoside O-phosphotransferase expression screen, compared to *E. coli leuU* and the YGA anticodon-swapped *E. coli leuU* tRNA variant. Grey shading indicates the anticodon region, and the host's LeuS leucine-tRNA-ligase identity elements[31] are shown in blue. Sequence information of the *leuU*$_{YGA}$ variants is available in Supplementary Data 3. **b**) Multiple sequence alignment of phage-derived tRNA-Leu$_{YGA}$ variants selected in the *aph3Ia*$_{29\times Leu\rightarrow TCR}$ aminoglycoside O-phosphotransferase expression screen, compared to endogenous *E. coli* leucine tRNAs. Grey shading indicates the anticodon region, while the host's LeuS leucine-tRNA-ligase identity elements[31] are shown in blue. Sequence information of the phage-derived tRNA-Leu$_{YGA}$ variants is available in Supplementary Data 3.



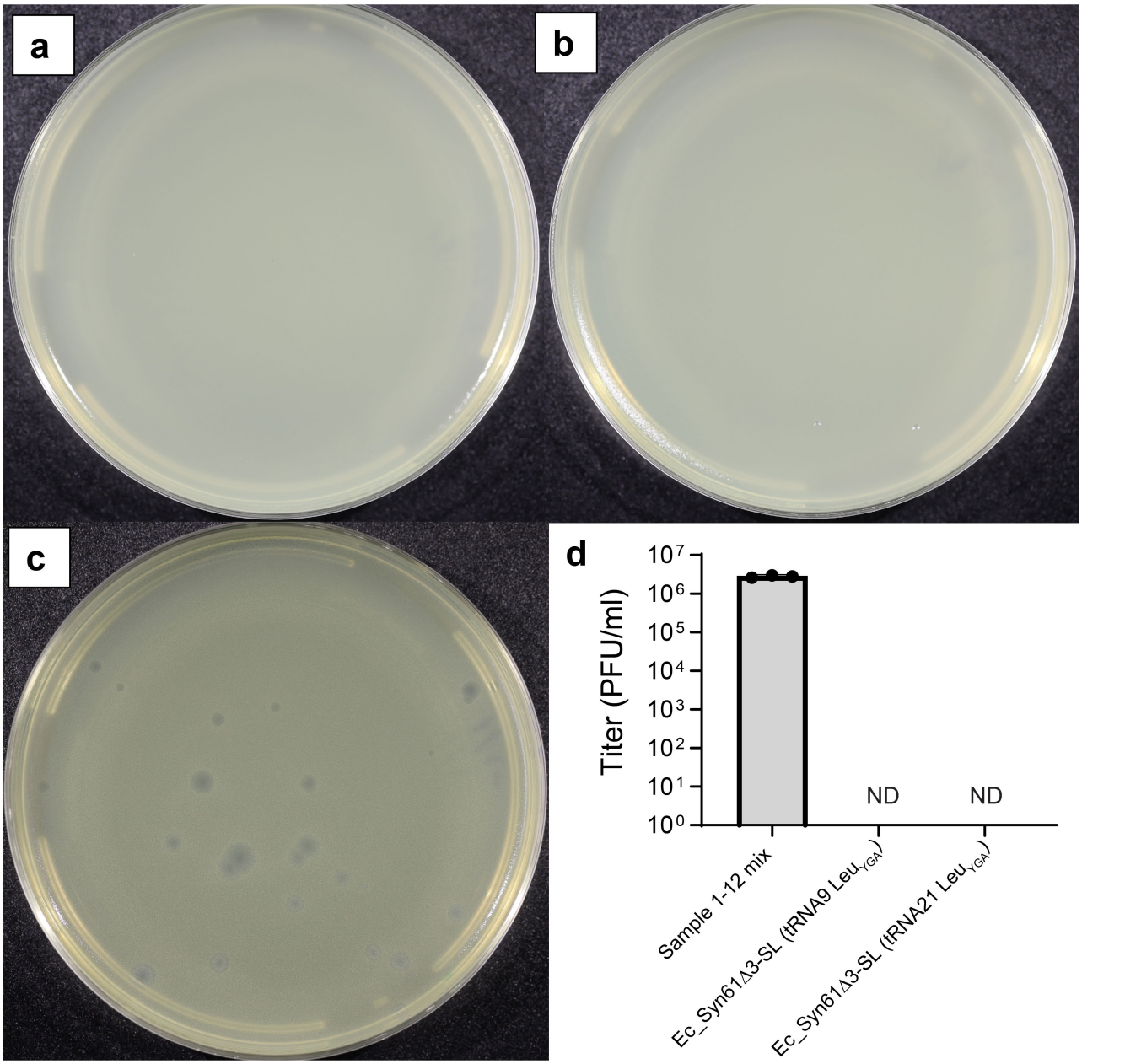
Extended Data Fig. 6 | tRNAseq-based quantification of viral Leu-tRNA$_{UGA}$ and Leu-tRNA$_{CGA}$ in the tRNAome of Ec_Syn61Δ3-SL. tRNA levels of Ec_Syn61Δ3-SL were quantified using tRNAseq (Methods). Viral Leu-tRNA$_{UGA}$ and Leu-tRNA$_{CGA}$ are highlighted in red, and the host's endogenous *E. coli serV* tRNA is highlighted in orange. tRNAseq data was collected once. Data represent TPM (transcript/million). Source data is available within this paper.



Extended Data Fig. 7 | Serine-to-leucine mistranslation of TCT codons in Ec_Syn61Δ3-SL cells. a) MS/MS spectrum of the serine-to-leucine mistranslated TufA peptide. **b**) MS/MS spectrum of the wild-type TufA peptide. The figure shows the amino acid sequence and MS/MS spectrum of the Ec_Syn61Δ3-SL-expressed TufA protein fragment, together with its genomic sequence, in which the serine-coding TCT codon (as shown in panel **b**) is partially mistranslated as leucine (as shown in panel **a**). The experiment was performed by analyzing the total proteome of Ec_Syn61Δ3-SL cells, expressing Leu9-tRNA$_{YGA}$ from *Escherichia* phage OSYSP (GenBank ID MF402939.1) by tandem mass spectrometry (Methods). MS/MS data was collected once.



Extended Data Fig. 8 | Doubling time and growth curves of Syn61Δ3(ev5), Syn61Δ3(ev5)Δ*recA*, Syn61Δ3(ev5)Δ*recA*(ev1), and Ec_Syn61Δ3-SL. **a**) Doubling times of Syn61Δ3(ev5), Syn61Δ3(ev5)Δ*recA*, Syn61Δ3(ev5) Δ*recA*(ev1), and Ec_Syn61Δ3-SL, calculated based on growth curves (shown in panels **b**, **c**, **d**) in rich bacterial media under standard laboratory conditions. **b**) Growth curves of Syn61Δ3(ev5), Syn61Δ3(ev5)Δ*recA*, and Syn61Δ3(ev5) Δ*recA*(ev1) in LBL broth. **c**) Growth curves of Syn61Δ3(ev5) and Syn61Δ3(ev5) Δ*recA*(ev1) in 2×YT broth. **d**) Growth curves of Ec_Syn61Δ3-SL in 2×YT broth containing 50 μg/ml kanamycin. Three independent cultures were grown aerobically in vented shake flasks at 37 °C, and OD_{600} measurements were taken during exponential growth (Methods). Data curves and bars represent the mean. Error bars show standard deviation based on n = 3 independent experiments.



Extended Data Fig. 9 | Ec_Syn61Δ3-SL resists viruses in environmental samples. (**a**) Phage enrichment experiment using Ec_Syn61Δ3-SL, expressing KP869110.1 tRNA24 Leu_{YGA} and MF402939.1 tRNA9 Leu_{YGA}, as host. (**b**) Phage enrichment experiment using Ec_Syn61Δ3-SL, expressing KP869110.1 tRNA24 Leu_{YGA} and MF402939.1 tRNA21 Leu_{YGA}, as host. Phage enrichment experiments were performed by mixing early exponential cultures of Ec_Syn61Δ3-SL with 10 ml environmental sample mix containing the mixture of Sample 2–13 from our study (Extended Data Table 1a). After two enrichment cycles (Methods, Supplementary Note), filter-sterilized culture supernatants were mixed with phage-susceptible *E. coli* MDS42 cells in top agar and plated on LBL agar plates to determine viral titer. Enrichment experiments were performed in n = 2 independent replicates with the same result. (**c**) Lytic *E. coli* MDS42 phage plaques after 10^3-fold dilution of the environmental sample mix. **d**) Lytic phage titer of the environmental sample mix, before and after enrichment on Ec_Syn61Δ3-SL. Dots represent the viral titer of the unenriched sample based on three independent experiments, measured on *E. coli* MDS42 cells. ND represents no plaques detected. Bar represents the mean. Error bar shows standard deviation based on n = 3 independent experiments.

Extended Data Table 1 | Isolation of lytic viruses infecting Syn61Δ3

a

Sample ID	Sample description	Sample treatment	Sample source location
1	Wastewater (primary effluent)	Chloroform treated	Massachusetts, USA
2	Wastewater (primary effluent)	0.22 μm filtered	Massachusetts, USA
3	River water	0.22 μm filtered	Massachusetts, USA
4	Soil from horse and alpaca enclosure	0.22 μm filtered	Massachusetts, USA
5	Water from porta-potty overflow	0.22 μm filtered	Massachusetts, USA
6	Irrigation water near wild rat nest	0.22 μm filtered	Massachusetts, USA
7	Soil from pasture-raised chicken enclosure	0.22 μm filtered	Massachusetts, USA
8	Soil from egg-laying chicken shed	0.22 μm filtered	Massachusetts, USA
9	Fresh pig feces	0.22 μm filtered	Massachusetts, USA
10	Mixed fresh horse and cow feces	0.22 μm filtered	Massachusetts, USA
11	Farm compost pile	0.22 μm filtered	Massachusetts, USA
12	Farm compost pile	0.22 μm filtered	Massachusetts, USA
13	Water from pig and goose water trough	0.22 μm filtered	Massachusetts, USA

b

Isolate	Genome size (kb)	Taxonomic classification
REP1	167.4	Viruses; Duplodnaviria; Heunggongvirae; Uroviricota; Caudoviricetes; Caudovirales; Myoviridae; Tevenvirinae; Tequatrovirus
REP2	167.9	Viruses; Duplodnaviria; Heunggongvirae; Uroviricota; Caudoviricetes; Caudovirales; Myoviridae; Tevenvirinae; Tequatrovirus
REP3	167,9	Viruses; Duplodnaviria; Heunggongvirae; Uroviricota; Caudoviricetes; Caudovirales; Myoviridae; Tevenvirinae; Tequatrovirus
REP4	166.8	Viruses; Duplodnaviria; Heunggongvirae; Uroviricota; Caudoviricetes; Caudovirales; Myoviridae; Tevenvirinae; Tequatrovirus
REP5	88.9	Viruses; Duplodnaviria; Heunggongvirae; Uroviricota; Caudoviricetes; Caudovirales; Myoviridae; Ounavirinae; Felixounavirus
REP6	87.2	Viruses; Duplodnaviria; Heunggongvirae; Uroviricota; Caudoviricetes; Caudovirales; Myoviridae; Ounavirinae; Felixounavirus
REP7	88.9	Viruses; Duplodnaviria; Heunggongvirae; Uroviricota; Caudoviricetes; Caudovirales; Myoviridae; Ounavirinae; Felixounavirus
REP8	85.8	Viruses; Duplodnaviria; Heunggongvirae; Uroviricota; Caudoviricetes; Caudovirales; Myoviridae; Ounavirinae; Felixounavirus
REP9	85.8	Viruses; Duplodnaviria; Heunggongvirae; Uroviricota; Caudoviricetes; Caudovirales; Myoviridae; Ounavirinae; Felixounavirus
REP10	85.7	Viruses; Duplodnaviria; Heunggongvirae; Uroviricota; Caudoviricetes; Caudovirales; Myoviridae; Ounavirinae; Felixounavirus
REP11	85.7	Viruses; Duplodnaviria; Heunggongvirae; Uroviricota; Caudoviricetes; Caudovirales; Myoviridae; Ounavirinae; Felixounavirus
REP12	85.8	Viruses; Duplodnaviria; Heunggongvirae; Uroviricota; Caudoviricetes; Caudovirales; Myoviridae; Ounavirinae; Felixounavirus

a) Environmental samples analyzed in this study for the presence of lytic viruses of Syn61Δ3. Samples containing lytic phages of Syn61Δ3 are highlighted in gray.
b) Genome size and taxonomy of lytic viruses infecting Syn61Δ3. Annotated genome sequences of all REP phage isolates are available in the Supplementary Data of this paper. REP is an abbreviation of **R**ecoded ***E***. *coli* **P**hage. The annotated genomes of REP phages have been deposited to NCBI GenBank under Accession numbers OQ174500, OQ174501, OQ174502, OQ174503, OQ174504, OQ174505, OQ174506, OQ174507, OQ174508, OQ174509, OQ174510, and OQ174511.

nature portfolio

Corresponding author(s): Akos Nyerges, George M. Church

Last updated by author(s): Jan 29, 2023

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	No software was used for data collection.
Data analysis	Data analysis was done using the following publicly available softwares and scripts: GraphPad Prism 9.3.1.471, bcl-convert (v3.9.3), BBDuk trimmer 1.0, SPAdes 3.15.2, Prokka version 1.14.6, Bowtie2 2.3.0, DESeq2 (version 1.38.3), BLAST (version 2.11.0), breseq version 0.36.1, bcl2fastq 2.20.0.445, porechop 0.2.3_seqan2.1.1, Unicycler 0.4.8, Geneious Prime® 2022.1.1. (Biomatters Ltd.), DESeq2, tRNAscan-SE 2.0.9, Proteome Discoverer 2.4 with Sequest HT (Thermo Fisher Scientific, USA) .

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

Raw data from whole-genome sequencing, transcriptome, and tRNA sequencing experiments have been deposited to Sequence Read Archive (SRA) under

BioProject ID PRJNA856259. tRNA data and the generated sequences in this study are included in the Supplementary Data files. Mass spectra and proteome measurements have been deposited to MassIVE (# MSV000089854; doi:10.25345/C5FF3M41W). The Syn61Δ3(ev5) ΔrecA (ev1) strain is available from Addgene (Bacterial strain #189857). All materials used in this study are freely available from the corresponding authors for academic research use upon request. The PHROG HMM database is available at https://phrogs.lmge.uca.fr/ and from reference 4. The assembled annotated genome of Escherichia coli Syn61 substr. Syn61Δ3(ev5) have is available in Supplementary Data 4, and the annotated genome of all REP phages have been deposited to NCBI GenBank under accession numbers OQ174500, OQ174501, OQ174502, OQ174503, OQ174504, OQ174505, OQ174506, OQ174507, OQ174508, OQ174509, OQ174510, and OQ174511. Source data is provided with this paper.

Human research participants

Policy information about studies involving human research participants and Sex and Gender in Research.

Reporting on sex and gender	Not applicable.
Population characteristics	Not applicable.
Recruitment	Not applicable.
Ethics oversight	Not applicable.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample sizes are described in the figure legends. No statistical calculations were done to predetermine the sample size. Sample sizes were chosen as large as possible while still feasible in terms of sample handling and data collection. Three replicates were performed in almost all cases with small standard deviation between measurements. We used this strategy and sample size because variation in the utilized assays is small and we were interested in large effects.
Data exclusions	No data was excluded.
Replication	Measurements were performed in independent replicates and only little variation was observed between replicates. The exact number of replicates is stated in the corrsponding figure legend and methods section. All measurements were reproducible.
Randomization	No randomization was performed because samples form defined groups.
Blinding	Sequencing and mass spectrometry raw data generation were blinded. All other forms of experiments and data generation was not blinded because experimental groups were defined and studied by the same investigator during experiments. Data analysis was performed with standard pipelines that did not include "subjective" criteria that could influence results.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging