

## Swapped genetic code blocks viral infections and gene transfer

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

<sup>1</sup>Department of Genetics, Harvard Medical School, Boston, MA 02115, USA

<sup>2</sup>Department of Biomedical Informatics and Laboratory of Systems Pharmacology, Harvard Medical School, Boston, MA 02115, USA

<sup>3</sup>Wyss Institute for Biologically Inspired Engineering, Harvard University, Boston, MA 02115, USA

<sup>4</sup>Department of Pathology and Immunology, Washington University School of Medicine in St. Louis, St. Louis, MO, USA

<sup>5</sup>The Edison Family Center for Genome Sciences and Systems Biology, Washington University School of Medicine in St. Louis, St. Louis, MO, USA

<sup>6</sup>GenScript USA Inc., Piscataway, NJ 08854, USA

\*Corresponding author. E-mail: [gchurch@genetics.med.harvard.edu](mailto:gchurch@genetics.med.harvard.edu) (G.M.C.); [akos.nyerges@hms.harvard.edu](mailto:akos.nyerges@hms.harvard.edu) or [nyergesakos@gmail.com](mailto:nyergesakos@gmail.com) (A.N.)

### Abstract

Removing cellular transfer RNAs (tRNAs), making their cognate codons unreadable, creates a genetic firewall that prevents viral replication and horizontal gene transfer. However, numerous viruses and mobile genetic elements encode parts of the translational apparatus, including tRNAs, potentially rendering a genetic-code-based firewall ineffective. In this paper, we show that such horizontally transferred tRNA genes can enable viral replication in *Escherichia coli* cells despite the genome-wide lack of three codons and the previously essential cognate tRNAs and release factor 1. By repurposing viral tRNAs, we then develop recoded 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 completely 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. Finally, we also repurpose the third free codon to biocontain this virus-resistant host via dependence on an amino acid not found in nature.

## Introduction

The universal 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 resistance genes and allows viruses to replicate, to kill both pro- and eukaryotic cells, and to cause diseases. Horizontal gene transfer also threatens the safe use of Genetically Modified Organisms (GMOs) by enabling the release and spread of their engineered genetic information into natural ecosystems. It is widely hypothesized that Genomically Recoded Organisms (GROs), whose genomes have been systematically redesigned to confer an alternate genetic code, would offer genetic isolation from natural ecosystems by obstructing the translation of horizontally transferred genetic material<sup>1-6</sup>, including both resistance to viral infections and horizontal gene transfer. Indeed, the genome-wide removal of TAG stop codons and release factor 1 (RF1) from *Escherichia coli*, which abolishes cells' ability to terminate translation at TAG stop codons, provides substantial resistance to bacteriophages and horizontal plasmid transfer<sup>2,3</sup>. 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 TGA stop codon were replaced with synonymous alternatives, and the corresponding serine transfer RNA (tRNA) genes (*serU* and *serT*) and RF1 (*prfA*) have been deleted<sup>4,7</sup>. Syn61Δ3 resists a broad range of phages without detectable viral replication due to its inability to translate TCR and TAG codons, including those phages that could overcome RF1-deletion-based resistance<sup>4</sup>.

However, 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 for a fully host-independent translation<sup>8-10</sup>. These translational elements allow viruses to reduce their dependency on host translational processes by substituting elements of the translational apparatus or, in more extreme cases, even alter the host's genetic code during viral replication<sup>11-13</sup>. Similarly, mobile genetic elements that encode transfer RNAs are widespread in nature. Recent studies highlighted the presence of mobile tRNA genes in diverse species, ranging from plasmids to actively spreading conjugative elements capable of decoding all twenty amino acids with their encoded tRNAs<sup>14-16</sup>. Therefore, the selection pressure posed by the altered genetic codes of GROs might facilitate the rapid evolution of viruses and mobile genetic elements capable of crossing a genetic-code-based barrier.

In this paper, we show that horizontally transferred tRNA genes can readily substitute cellular tRNAs in GROs and thus abolish genetic-code-based resistance to viral infections and HGT. Next, by repurposing virus-encoded tRNAs, we develop an amino-acid-swapped genetic

code that—by reassigning the amino acid identity of two sense codons—provides complete virus resistance and enables the tight biocontainment of engineered genetic information. These developments provide a fundamental advance toward engineering multi-virus-resistant cell lines and the safer use of GMOs in natural environments.

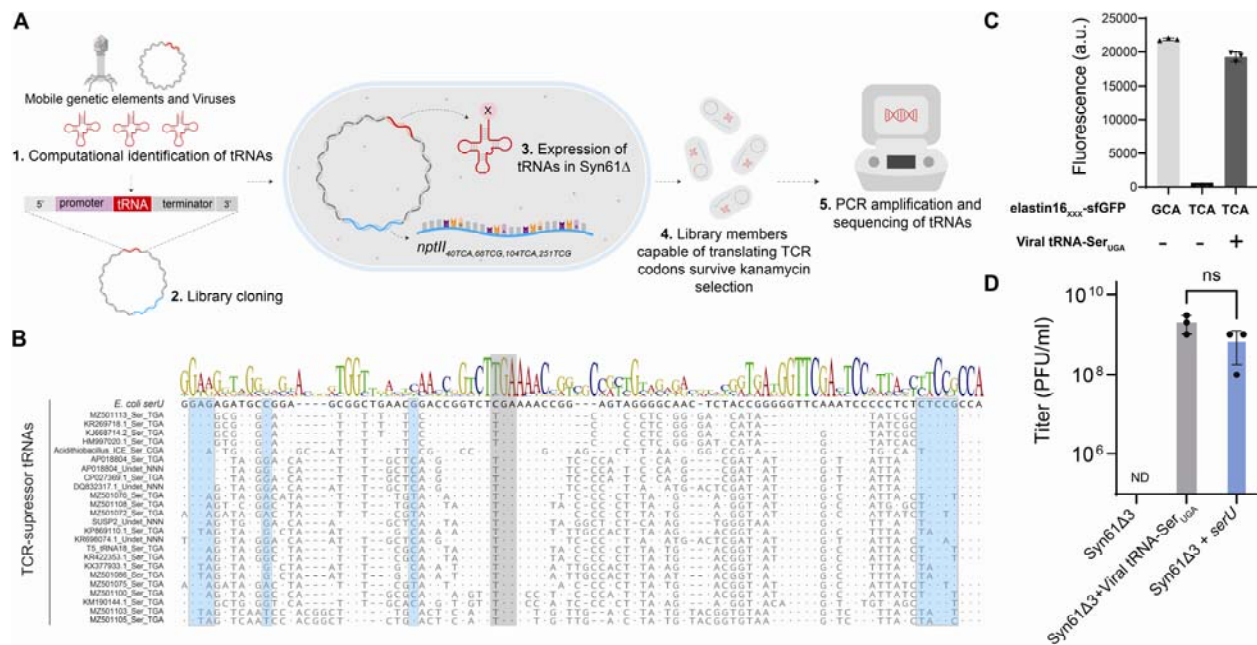
## Results

### Mobile tRNA genes participate in translation and facilitate horizontal gene transfer

We first investigated whether mobile genetic element-encoded tRNAs can complement 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 selecting and synthesizing 1192 tRNA genes from phylogenetically diverse plasmids, transposable elements, and bacteriophages infecting members of the *Enterobacteriaceae* family (Supplementary Table 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 **Figure 1A**, this 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 release factor 1 (*prfA*) have been deleted. This strain, *E. coli* Syn61 $\Delta$ 3, 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*<sub>40TCA,68TCG,104TCA,251TCG</sub> aminoglycoside-O-phosphotransferase antibiotic resistance gene containing TCA codons at positions 40, 104 and TCG codons at positions 68 and 251. In wild-type *E. coli* cells bearing the canonical genetic code, *nptII*<sub>40TCA,68TCG,104TCA,251TCG</sub> 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 Syn61 $\Delta$ 3, however, the production of this resistance-conferring gene product is inhibited due to the lack of *serU* and *serT*-encoded tRNA-Ser<sub>UGA</sub> and -Ser<sub>CGA</sub> 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 Syn61 $\Delta$ 3 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 61 tRNA sequences capable of promoting

*nptII*<sub>40TCA,68TCG,104TCA,251TCG</sub> expression (**Figure 1B, Supplementary Table 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 Ser tRNAs, most tRNA hits contained a UGA anticodon and carried the identity elements necessary for recognition by the host's SerS serine-tRNA-ligase (**Figure 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<sub>CGA</sub> 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 neighboring bacterial cells<sup>15,17</sup>. In agreement with prior studies<sup>8,9,18</sup>, 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* highlighted the presence of sense and stop codon suppressor tRNA encoding genes in both groups, suggesting that mobile tRNAs are prevalent across viruses infecting prokaryotic, archaeal, and eukaryotic hosts (**Supplementary Table 2**).



**Figure 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 (*i.e.*, 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 under the control of a constitutive bacterial promoter into a plasmid vector carrying a *nptII*<sub>40TCA,68TCG,104TCA,251TCG</sub> marker (conferring kanamycin resistance) (2.). Following the transformation of this library into Syn61Δ3 (3.), in which the deletion of *serU* (encoding tRNA-Ser<sub>CGA</sub>) and *serT* (encoding tRNA-Ser<sub>UGA</sub>) makes TCG and TCA codons unreadable, only variants carrying functional TCR suppressor tRNAs survive kanamycin selection (4.). Finally, PCR amplification and 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, while the host's serine-tRNA-ligase identity elements are shown in blue. (C) Viral serine tRNA<sub>UGA</sub> translates the TCA codon. Syn61Δ3 expressing elastin<sub>16</sub> GCA (alanine)-sfGFP-His6 served as wild-type expression control, and the elastin<sub>16</sub> TCR-sfGFP-His6 expression was compared with and without the coexpression of the tRNA-Ser<sub>UGA</sub> of *Escherichia* phage IrisVonRoten<sup>19</sup>. xxx marks the analyzed codon, TCA or GCA. Error bars represent SD based on three replicates. a.u. denotes arbitrary fluorescence units. (D) The expression of viral TCR suppressor tRNAs and *serU* (tRNA-Ser<sub>CGA</sub>) restores the replication of T6 bacteriophage in Syn61Δ3. Cultures were infected at an MOI of 0.01 with T6 phage, and the figure shows total T6 titer after 24 hours of incubation. All experiments were performed in three independent replicates, the dots represent data from independent replicates, and the bar graph represents the mean. ND represents below the detection limit (*i.e.*, <10<sup>3</sup> PFU/ml); ns indicates lack of significance (*p* ≥ 0.5) based on Student's t-test.

We confirmed the TCR codon-recognizing tRNAs' predicted serine amino acid identity by coexpressing a selected tRNA hit with an elastin<sub>16TCA</sub>-sfGFP-His6 construct harboring a single TCA codon at position 16. The coexpression of the tRNA-Ser<sub>UGA</sub> of *Escherichia* phage IrisVonRoten<sup>19</sup> together with the elastin<sub>16TCA</sub>-sfGFP-His6 construct conferred near wild-type level expression (Figure 1C) and tryptic digest followed by reverse-phase liquid chromatography and tandem mass spectrometry (LC/MS-MS) confirmed serine incorporation at the TCA position (Supplementary Figure 1).

Next, we investigated whether mobile tRNA<sub>ome</sub>-derived tRNAs could promote viral replication. A previous study demonstrated that Syn61Δ3 resists infection by multiple bacteriophages, including Enterobacteria phage T6<sup>4</sup>. Infecting Syn61Δ3 with T6 phage recapitulated these results. In contrast, the infection of Syn61Δ3 harboring a bacteriophage-derived Ser-tRNA<sub>UGA</sub> gene with T6 resulted in rapid lysis, indicating that tRNA genes that reside in viral genomes can substitute cellular tRNAs and promote phage infection (Figure 1D).

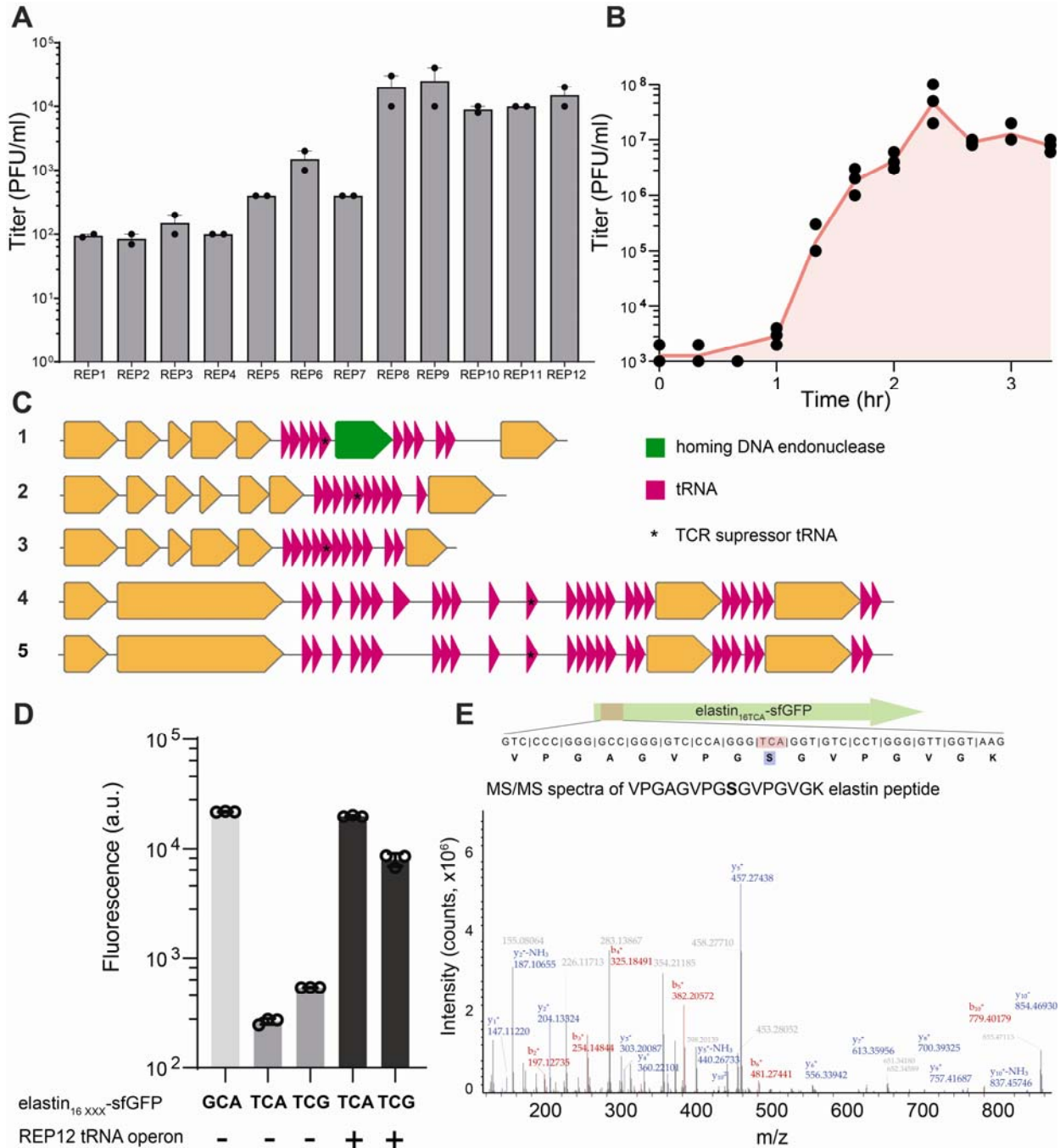
The discovery of diverse TCR codon translating tRNAs on horizontally transferred genetic elements indicates that mobile tRNA genes are widespread and can readily complement the lack of cellular tRNAs to promote viral replication and horizontal gene transfer.

### Isolation of lytic viruses infecting Syn61Δ3

We next investigated whether lytic viruses of Syn61Δ3 exist. We infected Syn61Δ3 cells with eleven coliphages whose genome harbor TCR translating tRNA genes based on our plasmid-based screen (**Figure 1B**). Surprisingly, 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 (**Supplementary Figure 2**).

We next attempted to isolate lytic viruses from diverse environmental samples by performing a standard two-step enrichment-based phage isolation protocol and using Syn61Δ3 as host. First, bacteria-free filtrates of environmental and wastewater samples ( $n=13$ , from Massachusetts (USA), **Table S1**) were mixed with Syn61Δ3 and grown until stationary phase. Next, bacterial cells were removed, and we analyzed 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 novel phage strains. All identified phages belong to the *Caudovirales* order and the *Myoviridae* family, taxa rich in tRNA-encoding bacteriophages<sup>9</sup> (**Supplementary Table 1**). 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 Table 1**). Surprisingly, all isolates harbored TCR suppressor serine tRNAs with a UGA anticodon that we identified in our earlier *nptII*<sub>40TCA,68TCG,104TCA,251TCG</sub> suppressor screen (**Figure 1B**). One isolate, REP1, also harbored a predicted homing endonuclease within its tRNA operon (**Figure 2**). Homing endonucleases encoded in tRNA operons have been shown to be responsible for the horizontal transfer of tRNA gene clusters within phage genomes<sup>20</sup>. Phage isolates showed more than two orders of magnitude difference in viral titers after replication on recoded cells (**Figure 2A**). One of the most virulent isolates, REP12, required 60 minutes to complete a replication cycle at 37 °C in Syn61Δ3 (**Figure 2B**).

The isolated viral strains infecting Syn61Δ3 show that bacteriophages that can overcome sense codon recoding-based viral resistance exist and are widespread in environmental samples.



**Figure 2. Lytic phages of Syn61 $\Delta$ 3.** (A) Titer of Syn61 $\Delta$ 3 phage isolates. Early exponential phase cultures of Syn61 $\Delta$ 3 were infected at an MOI of 0.001 with the corresponding REP phages, and free phage titers were determined after 24 hours of incubation. Dots represent data from independent measurements; bar graphs represent the mean; the error bars represent the SEM. (B) Single-step growth curve of REP12 lytic Syn61 $\Delta$ 3 phage. Growth experiment was performed in triplicates, and dots represent the total viral titer. (C) Genomic maps of tRNA operons in lytic Syn61 $\Delta$ 3 phages. Magenta arrows represent predicted tRNA genes; star denotes tRNA genes identified in our earlier TCR codon suppressor screen (Figure 1B); green arrow represents homing endonuclease gene, while 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)** Viral tRNA operon-expressed tRNAs translate TCR codons. Syn61 $\Delta$ 3 expressing elastin<sub>16</sub> GCA (alanine)-sfGFP-His6 served as wild-type expression control, and the elastin<sub>16</sub> TCR-sfGFP-His6 expression was compared with and without the coexpression of the REP12 viral tRNA operon. xxx marks the analyzed codon, TCA, TCG, or GCA. A.u. denotes arbitrary fluorescence units. Error bars represent SD based on three independent replicates. **(E)** Viral tRNA operon-expressed tRNAs decode TCR codons as serine. The amino acid identity of the translated TCA codon within elastin<sub>16</sub> TCA-sfGFP-His6 was confirmed by tandem mass spectrometry from Syn61 $\Delta$ 3 cells containing the REP12 tRNA operon and its cognate promoter. The figure shows the amino acid sequence and MS/MS spectrum of the analyzed elastin<sub>16</sub> TCR peptide. MS/MS data was collected once.

### **Viral tRNAs substitute cellular tRNAs to support translation**

We next investigated how tRNA-encoding viruses evade genetic-code-based resistance. Time-course transcriptome analysis of REP12 phage-infected Syn61 $\Delta$ 3 cells during the viral replication cycle revealed early and high-level expression of the viral tRNA operon (**Supplementary Figure 3**, Supplementary Table 3). 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 (**Supplementary Figure 3**). We then investigated the time-course kinetics of tRNA expression in Syn61 $\Delta$ 3 cells that were infected with our REP12 phage by performing tRNA sequencing (tRNAseq). Time-course tRNAseq experiments revealed remarkably high-level expression of the viral tRNA-Ser<sub>UGA</sub> immediately after phage attachment (*i.e.*, a relative viral tRNA-Ser<sub>UGA</sub> abundance of 56.1% ( $\pm$ 5%) compared to the host *serV* tRNA). Throughout the entire phage replication cycle, the phage tRNA-Ser<sub>UGA</sub> remained one of the most abundant viral tRNA species inside infected Syn61 $\Delta$ 3 cells (Supplementary Table 3). We next investigated whether phage tRNA-Ser<sub>UGA</sub> participates in translation by analyzing the presence of their mature form. The gene encoding the tRNA-Ser<sub>UGA</sub> in REP12's genome 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<sub>UGA</sub> ends detected CCA tail addition in 62.9% ( $\pm$ 1.9%) of all tRNA sequencing reads immediately after phage attachment, indicating that mature tRNA-Ser<sub>UGA</sub>s are instantly being produced after host infection (**Supplementary Figure 4**).

We also investigated transcriptomic changes in Syn61 $\Delta$ 3 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 that catalyzes the final step in the *de novo* synthesis of queuosine in tRNAs<sup>21</sup>, and *trmJ*, tRNA C<sup>m32</sup>/U<sup>m32</sup> methyltransferase<sup>22</sup>, which introduces methyl groups at the 2'-O position of U<sup>32</sup> of several tRNAs, including tRNA-Ser<sub>UGA</sub>, suggesting the potential posttranscriptional modification of phage-derived tRNAs (Supplementary Table 3).

Finally, we also validated the role of phage tRNA-Ser<sub>UGA</sub> tRNAs in decoding TCR codons. We first cloned the REP12 viral tRNA operon containing the hypothetical tRNA-Ser<sub>UGA</sub> and its predicted promoter into a plasmid vector. Coexpression of this tRNA operon with an elastin<sub>16 TCA</sub>-sfGFP-His6 and elastin<sub>16 TCG</sub>-sfGFP-His6 construct, harboring either a single TCA or TCG codon at position 16, respectively, resulted in high-level elastin-sfGFP-His6 expression (**Figure 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<sub>16 TCR</sub>-sfGFP-His6 samples (**Figure 2E, Supplementary Figure 5A**). As expected, the coexpression of the same elastin<sub>16 TCA</sub>-sfGFP-His6 construct with the only tRNA-Ser<sub>UGA</sub> 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 Figure 5B**).

Together these results show that lytic phages of Syn61 $\Delta$ 3 overcome genetic-code-based viral resistance by rapidly complementing the cellular tRNA pool with virus-encoded tRNAs.

### Creation of an amino-acid-swapped genetic code

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 horizontal gene transfer 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. Although swapping the amino acid identity of sense codons presents a possible way to prevent horizontal gene transfer<sup>23-25</sup>, it was impossible to test this hypothesis *in vivo* until now. To establish a serine<sub>TCR</sub>-to-leucine swapped genetic code (**Figure 3A**), we utilized Syn61 $\Delta$ 3, which genome-wide lacks annotated instances of TCR codons and their corresponding tRNA genes, and sought to identify tRNAs capable of efficiently translating TCR codons as leucine. To this aim, we modified our previous tRNA library selection screen (**Figure 1A**) to evolve efficient 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 an *aph3la*<sub>29xLeu→TCR</sub>, 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 translating all 29 TCR codons as leucine would confer resistance to kanamycin. We identified two distinct *leuU* variants by applying “high” kanamycin concentration (*i.e.*, 200 µg/ml) as selection pressure to Syn61Δ3 cells carrying the anticodon-swapped tRNA library. These variants, carrying tRNAs containing distinct anticodon loop mutations (**Supplementary Figure 6**), were then infected with a cocktail of all twelve phage isolates (Table S2) that are capable of lysing Syn61Δ3 at a 10:1 cell-to-phage ratio (*i.e.*, a Multiplicity of Infection (MOI) of 0.1). Surprisingly, all selected *leuU* library members allowed robust phage replication, with phage titers reaching ~10<sup>7</sup> PFU/ml after 24 hours (**Figure 3C**). We hypothesized that viral replication in the presence of TCR suppressing *leuU* variants is due to these tRNAs lower suppression efficiency compared to phage-carried serine tRNAs, which leads to rapid viral takeover. Viral tRNA-Ser<sub>YGA</sub>, that are tRNA-Ser<sub>UGA</sub> and tRNA-Ser<sub>CGA</sub>, might have i.) higher aminoacylation efficiency by their corresponding *E. coli* aminoacyl-tRNA-ligase than our selected *leuU* variants, ii.) higher affinity towards the bacterial ribosome, and/or iii.) better evade phage- and host-carried tRNA-degrading effector proteins<sup>11,26,27</sup>.

Based on this observation, we hypothesized that bacteriophage-encoded tRNAs might provide higher suppression efficiencies for their cognate codons than their native *E. coli* counterpart. Therefore, we next constructed a small, focused library that coexpressed YGA anticodon-swapped mutants of 13 phage-encoded leucine tRNAs, together with an *aph3la*<sub>29xLeu→TCR</sub> aminoglycoside O-phosphotransferase gene in which all 29 instances of leucine codons were replaced with TCR serine codons. The transformation of this library into Syn61Δ3 cells and subsequent “high” concentration (*i.e.*, 200 µg/ml) kanamycin selection identified three distinct variants displaying 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 (**Supplementary Figure 7**). Furthermore, the analysis of the total tRNA content of these cells by tRNAseq confirmed the presence of synthetic phage Leu-tRNA<sub>YGA</sub> tRNAs with similar abundances as the cellular endogenous serine tRNAs (*i.e.*, a relative expression level of 172% and 140% for Leu-tRNA<sub>UGA</sub> and Leu-tRNA<sub>CGA</sub> respectively, compared to *serV* (Supplementary table 3)).

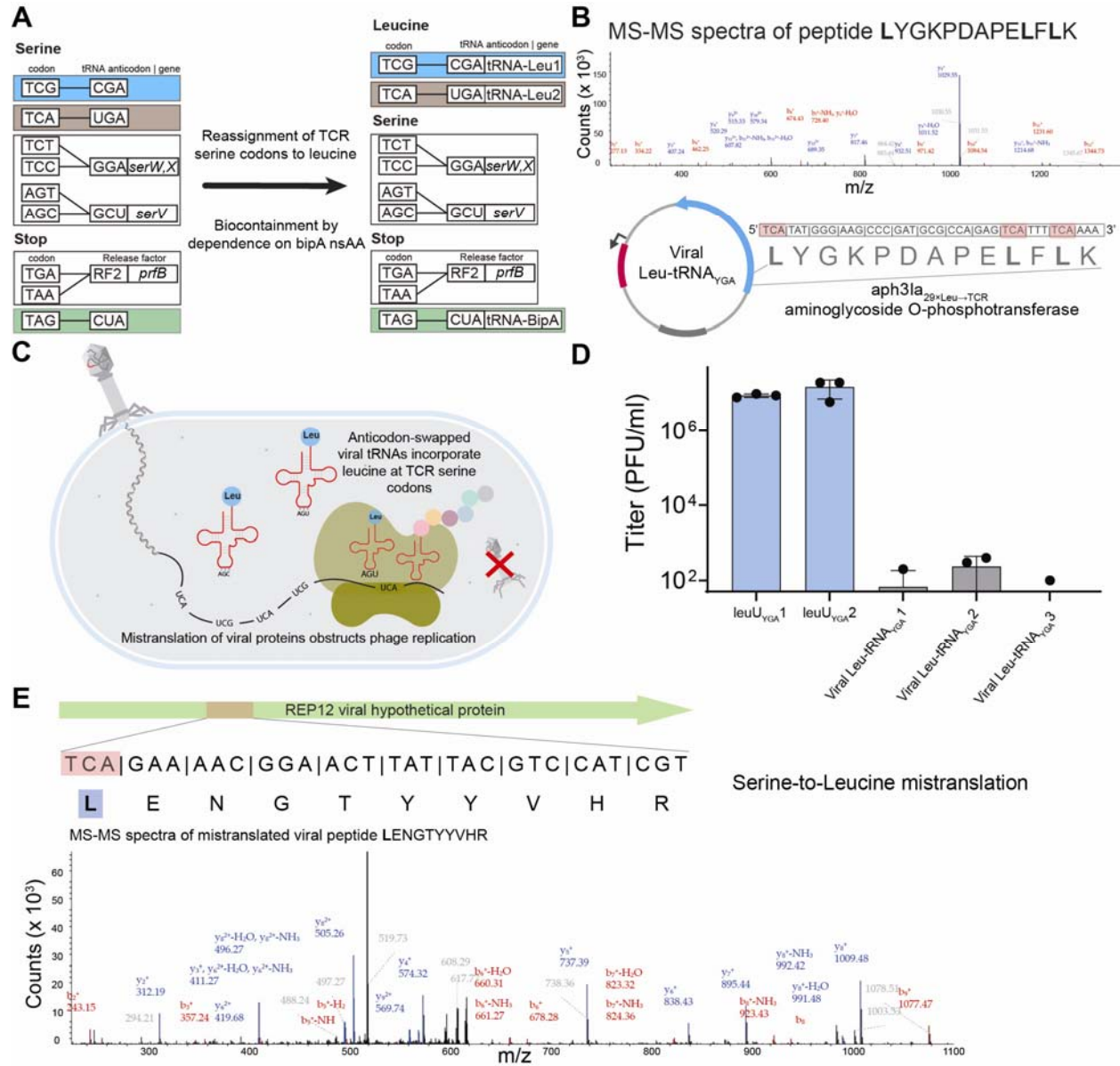
Next, similarly to our previous infection assay, phage tRNA<sub>YGA</sub> expressing cells were infected with a mixture of twelve distinct, lytic phages of Syn61Δ3 at a MOI = 0.1. The analysis of phage titer in culture supernatants after 24 hours showed a marked drop compared to the

input phage inoculum, suggesting that anticodon-swapped viral leucine tRNAs block phage replication (**Figure 3C**).

We then investigated the mechanism of phage resistance in *E. coli* cells carrying phage-derived tRNA-Leu<sub>YGA</sub> tRNAs (Ec\_Syn61Δ3 Ser→Leu Swap, or Ec\_Syn61Δ3-SL in short) by performing total proteome analysis. Untargeted proteome analysis of uninfected cells by tandem mass spectrometry validated the translation of TCR codons as leucine in Ec\_Syn61Δ3-SL (**Figure 3B**). Time-course untargeted proteome analysis after bacteriophage infection revealed extensive mistranslation at TCR codons in newly synthesized phage proteins (**Figure 3E**, **Supplementary Figure 8**), 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<sub>UGA</sub><sup>28,29</sup>, we also detected serine-to-leucine mistranslation at TCT codon positions in Ec\_Syn61Δ3-SL cells (**Supplementary Figure 9**). The recognition of TCT codons by phage tRNA-Leu<sub>YGA</sub> tRNAs might also be responsible for the slight fitness decrease of Ec\_Syn61Δ3-SL cells compared to its ancestor strain (*i.e.*, a doubling time of 69.3 minutes, compared to 44.29 minutes for the parental Syn61Δ3 strain in rich 2xYT media (**Supplementary Figure 10**)). 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*, *safA*, contain TCR codons and become mistranslated in Ec\_Syn61Δ3-SL (Supplementary Data).

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), *e.g.*, L-4,4'-biphenylalanine (*bipA*), offers tight, likely escape-free biocontainment that remains stable under long-term evolution<sup>30-32</sup>. Therefore, we generated a recombination deficient (*i.e.*, Δ*recA*), biocontained version of Ec\_Syn61Δ3-SL bearing a *bipA*-dependent essential *adk* gene and the *bipA* aminoacyl-tRNA synthetase/tRNA-*bipA*<sub>CUA</sub> system, by first performing adaptive laboratory evolution on a *recA* knock-out Syn61Δ3, and then replacing the genomic *adk* copy with its *bipA*-dependent variant<sup>33,34</sup> (Methods). This strain maintained the low escape frequency of previously reported single-gene synthetic auxotrophs<sup>30</sup> (*i.e.*,  $2.9 \times 10^{-6}$  ( $\pm 5.9 \times 10^{-7}$ ) 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 sewage (**Table S1**), but could not detect phages in culture supernatants (**Supplementary Figure 11**).

Together, these results demonstrate that reassigning sense codons TCA and TCG to leucine *in vivo* provides multivirus resistance, and the TAG stop codon can be simultaneously utilized to biocontain this virus-resistant strain via dependence on an amino acid not found in nature.



**Figure 3. An amino-acid-swapped genetic code provides multi-virus resistance. (A)** The creation of an *E. coli* GRO, Ec\_Syn61 $\Delta$ 3-SL, in which both TCA and TCG—naturally serine-meaning—codons are translated as leucine. The introduction of bacteriophage-derived Leu-tRNA<sub>UGA</sub> and Leu-tRNA<sub>CGA</sub> to Syn61 $\Delta$ 3 reassigns TCA and TCG codons to leucine, while the reassignment of the TAG stop codon to encode L-4,4'-biphenylalanine (bipA) in an essential gene of the host ensures the biocontainment of Ec\_Syn61 $\Delta$ 3-SL. **(B)** The reassignment of TCR codons to leucine within the coding sequence of *aph3la*<sub>29xLeu $\rightarrow$ TCR</sub> in Syn61 $\Delta$ 3-LS was confirmed by tandem mass spectrometry. The figure shows the amino acid sequence and MS/MS spectrum of the detected *aph3la*<sub>29xLeu $\rightarrow$ TCR</sub> peptide and its coding sequence. **(C)** Schematic of

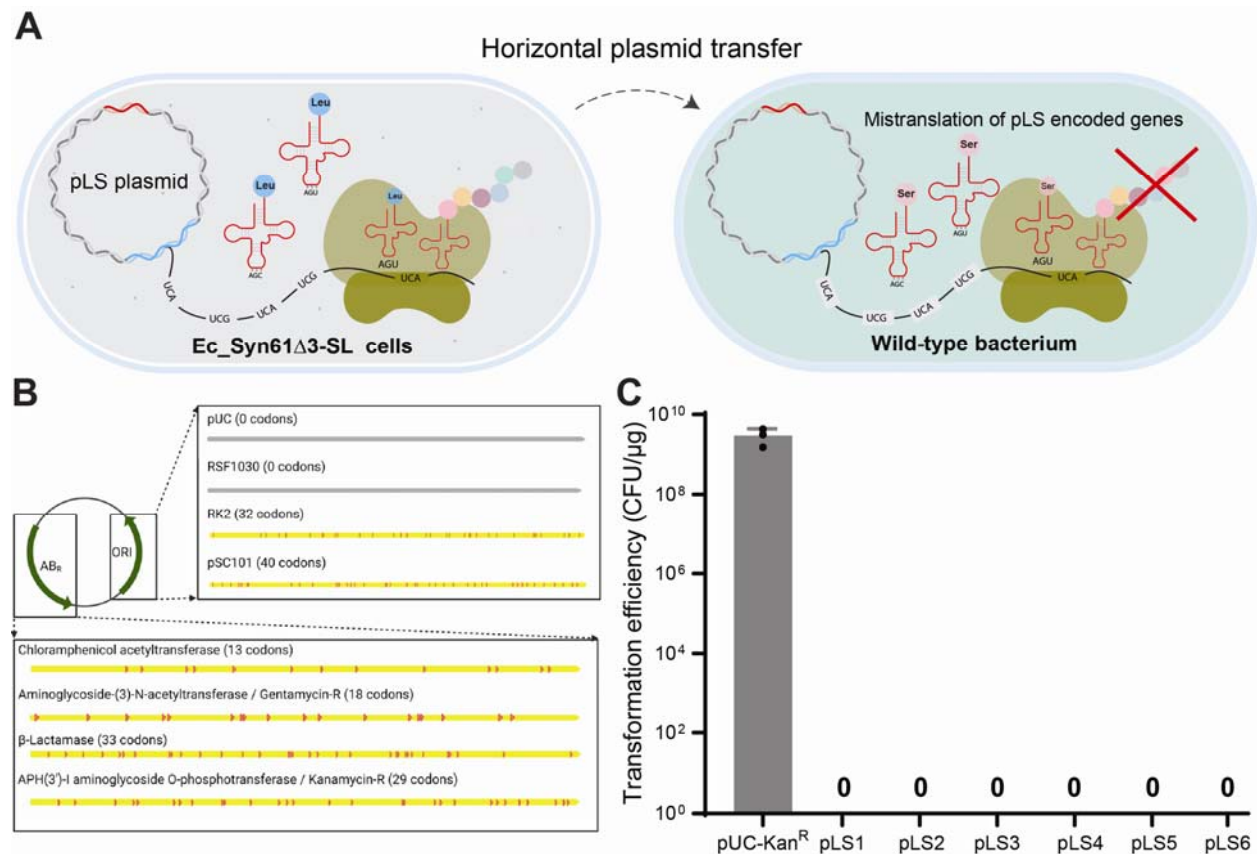
viral infection in *Ec\_Syn61Δ3-SL*. The reassignment of sense codons TCA and TCG to leucine in *Ec\_Syn61Δ3-SL* provides multivirus resistance by mistranslating the viral proteome. **(D)** Bacteriophage-derived Leu-tRNA<sub>UGA</sub> and Leu-tRNA<sub>CGA</sub> expression in *Syn61Δ3* provides multivirus resistance. The figure shows the titer of lytic *Syn61Δ3* phages following the infection of the corresponding Leu-tRNA<sub>YGA</sub>-expressing *Syn61Δ3* strain with a mixture of twelve distinct REP *Syn61Δ3* phages (Table S1). Early exponential phase cultures of *Syn61Δ3* carrying the corresponding Leu-tRNA<sub>YGA</sub> expression construct were infected at an MOI of 0.1 with the corresponding phage mixture, free phages were removed, and phage titers were determined after 24 hours of incubation (Methods, Supplementary Table 4). All experiments were performed in three independent replicates; dots represent data from independent replicates; bar graphs represent the mean; the error bars represent SD. **(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 performed by infecting *Ec\_Syn61Δ3-SL* cells, expressing Leu9-tRNA<sub>YGA</sub> from *Escherichia* phage OSYSP, with the REP12 phage at an MOI = 12, and the proteome of infected cells was analyzed by tandem mass spectrometry (Methods).

### **Addiction to an amino-acid-swapped genetic code provides a bidirectional firewall for synthetic genetic information**

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. Genetically Modified Organisms (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 for GMOs exist (e.g., *bipA* nsAA-based synthetic auxotrophy, as in *Ec\_Syn61Δ3-SL*), current methods fail to prevent the horizontal gene transfer (HGT)-based escape of engineered genetic information. Synthetic addiction to an artificial genetic code offers a solution to this problem. Using our phage-derived tRNA-Leu<sub>YGA</sub> expressing *Ec\_Syn61Δ3-SL* cells, we, therefore, developed a set of plasmid vectors that depend on TCR codons to express leucine-containing proteins and thus can only function in cells that efficiently translate TCR codons as leucine (**Figure 4A**). These plasmids, called the pLS plasmids, offer four 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 (**Figure 4B**, Supplementary Table 4). 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 these genes' functionality to *Ec\_Syn61Δ3-SL* cells. We tested the ability of our pLS vectors to function in cells bearing the standard genetic code cells by transforming six variants into wild-type *E. coli* K-12 MG1655 cells but could not detect escapees carrying pLS plasmids. The escape of pLS plasmids was similarly prevented when the phage tRNA-Leu<sub>YGA</sub> expression cassette was encoded within the plasmid backbone (*i.e.*, pLS1 and pLS2), indicating that anticodon-swapped tRNAs are severely toxic to wild-type cells (**Figure 4C**). Based on these results, we also expect that, similarly to pLS' genes, any leucine-requiring protein can be addicted to *Ec\_Syn61Δ3-SL* by recoding target genes to encode one or more leucine positions as TCR codons.

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



**Figure 4. Addition of synthetic genetic information to a genetic code in which TCR codons encode leucine prevents horizontal gene transfer.** (A) We developed a set of plasmid vectors, termed the pLS plasmids, that rely on TCR codons to express leucine-containing proteins. pLS plasmids only function in *Ec\_Syn61Δ3-SL* expressing bacteriophage-derived synthetic tRNA-Leu<sub>YGA</sub> 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 ( $AB_R$ ) together with low to high copy-number origins-of-replication (ORI) that are addicted to an artificial genetic code in which leucine is encoded as TCR codons. Number in parenthesis marks the number of  $Leu_{TCR}$  codons in each gene. Detailed sequence information and the description of pLS plasmids are available in Supplementary Table 4. (C) The transformation of pLS plasmids into wild-type *E. coli* K-12 MG1655 cells resulted in no escapees after 7 days of incubation. In each case, 1  $\mu$ g pLS plasmid from Ec\_Syn61 $\Delta$ 3-SL was electroporated into *E. coli* K-12 MG1655 (electroporation efficiency =  $\sim 3 \times 10^9$  transformants/ $\mu$ g, based on the transformation of a plasmid carrying a pUC origin-of-replication and kanamycin resistance (pUC-Kan<sup>R</sup>)) and then plated on antibiotic-containing agar plates. Besides an amino-acid-swapped antibiotic resistance gene, pLS1 and pLS2 carry the phage tRNA- $Leu_{YGA}$  expression cassette within the plasmid backbone (Supplementary Table 4). Experiments were performed in triplicates; data points represent data from independent experiments.

## Discussion

We have shown that tRNAs expressed by horizontally transferred genetic elements, including bacteriophages, plasmids, and integrative conjugative elements—the mobile tRNA<sub>ome</sub>—readily substitute cellular tRNAs and can abolish the genetic-code-based isolation of Genomically Recoded Organisms (GROs). By screening more than a thousand mobile tRNA<sub>ome</sub>-derived tRNAs, we discovered tRNA species capable of restoring viral replication in a recently developed *E. coli* GRO, Syn61 $\Delta$ 3, lacking sense TCR serine codons and the TAG stop codon together with their cognate serine tRNAs and release factor 1. We have also shown that the mobile tRNA<sub>ome</sub> is not limited to bacteria, as multiple archaeal and eukaryotic viruses also carry predicted tRNA genes. We hypothesize that in future studies, our general, multiplexed tRNA suppressor screen described herein (**Figure 1A**) will facilitate the analysis of mobile tRNA genes in these organisms as well.

We then discovered twelve lytic viruses in environmental samples that can lyse *E. coli* Syn61 $\Delta$ 3 (**Figure 2A-B, Table S2**). These bacteriophages harbor and express up to 27 tRNA genes, including a functional tRNA- $Ser_{UGA}$  needed to overcome the host's genetic-code-based virus resistance. Using tRNA sequencing and tandem mass spectrometry, we also showed that viral tRNA- $Ser_{UGA}$  becomes highly abundant in infected bacterial cells directly after phage entry, and it inserts serine in response to TCR codons. These findings impact ongoing recoding projects, including our aim to engineer a TAG stop codon recoded human GRO and a 57-codon recoded strain of *E. coli*<sup>35–37</sup>, as some of the identified viral tRNAs might enable viral replication in these recoded organisms (**Supplementary Table 1**). Therefore, we are now implementing additional genetic firewalls to ensure the complete virus resistance of these engineered hosts.

Finally, we have created a biocontained *E. coli* GRO, Ec\_Syn61 $\Delta$ 3-SL, with an artificial genetic code that resists infection by a wide range of bacteriophages, including phages in sewage and all lytic *E. coli* Syn61 $\Delta$ 3 phages from this study. Ec\_Syn61 $\Delta$ 3-SL achieves this remarkable virus resistance by the engineered reassignment of TCR codons to leucine—an amino acid different from their natural serine identity—thus mistranslating viral proteomes and mobile genetic elements that rely on the standard genetic code. Consequently, the genetic code of Ec\_Syn61 $\Delta$ 3-SL poses a bidirectional genetic firewall that simultaneously prevents viral replication and the escape of synthetic genetic information from Ec\_Syn61 $\Delta$ 3-SL into natural ecosystems. Next, by adding plasmids to express leucine-containing proteins with TCR codons, we developed a set of vectors that cannot function in cells bearing the canonical genetic code. These plasmids, called the pLS plasmid series (**Figure 4A-B**, Supplementary Table 4), restrict any synthetic constructs' functionality to Ec\_Syn61 $\Delta$ 3-SL cells and thus provide escape-free containment for engineered genetic information.

Future work will explore how viruses evade the amino-acid-swapped genetic code of Ec\_Syn61 $\Delta$ 3-SL. We hypothesize that bacteriophages that rely on the fewest number of TCR codons to express essential proteins, e.g., *Escherichia phage EC6098* with only 33 TCR positions in its six protein-coding genes<sup>38</sup>, have the highest potential to overcome an amino-acid-swapped code. Alternatively, viral tRNA-degrading proteins<sup>26,27</sup> could evolve to selectively destruct mistranslating tRNAs and thus promote viral escape.

We expect that these results will have broad implications on the safe use of Genetically Modified Organisms in open environments by establishing a generalizable method for genetic code alteration in GROs that simultaneously prevents viral predation in natural ecosystems and blocks incoming and outgoing HGT with natural organisms. The combination of genome recoding and codon reassignment might provide a universal strategy to make any species resistant to all natural viruses.

## Acknowledgments

We thank György Pósfai (Biological Research Centre, Hungary) for sharing MDS42 and Jason W. Chin's team (Medical Research Council Laboratory of Molecular Biology, UK) for sharing Syn61 $\Delta$ 3 via Addgene. Funding for this research was provided by the US Department of Energy (DOE) under grant DE-FG02-02ER63445 and by the National Science Foundation (NSF) Award number: 2123243 (both to G.M.C.). A.N. was supported by the EMBO LTF 160-2019 Long-Term fellowship. The authors thank Andrew Millard's laboratory for making the PHROG HMM



database available for bacteriophage annotation, GenScript USA Inc. for their DNA synthesis support, and Dan Snyder, Katrina Harris, and all members of the Microbial Genome Sequencing Center (MiGS), Pittsburgh, PA for their support with DNA and RNA sequencing. We are thankful to Ting Wu for her support, Yue Shen and Shirui Yan (Institute of Biochemistry, Beijing Genomics Institute) for our collaboration on genome recoding, and Behnoush Hajian for graphical design and her help with illustrations.

### **Author contribution**

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. performed tRNA-Leu suppressor screens, sfGFP expression assays, assisted in the construction of the pLS plasmids and biocontainment experiments. R.F. assisted in experiments, performed adaptive laboratory evolution and growth rate measurements. S.V.O., E.R., and M.B. provided environmental samples for phage isolation, performed replication assays, and provided support for phage experiments and genome analyses. M.L., K.C., F.H. performed DNA synthesis, while M.B.T., A.C.P., and E.K. supported the project. B.B. performed tandem mass spectrometry analyses. K.N. and J.A.M. provided reagents for tRNAseq experiments.

### **Conflict of interest statement**

The authors declare competing financial 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, and 64x Bio. Other potentially relevant financial interests of G.M.C. are listed at <http://arep.med.harvard.edu/gmc/tech.html>.

### **Data availability**

Raw data from whole-genome sequencing, transcriptome, and tRNA sequencing experiments have been deposited to Sequence Read Archive (SRA) under BioProject ID PRJNA856259. Mass spectra and proteome measurements will be deposited to public databases prior to publication. All materials used in this study are available from the corresponding authors upon request. Assembled bacteriophage genomes and the annotated genome of Syn61 $\Delta$ 3(ev5) (Addgene #174514) are available in the Supplementary Material of this paper.



## Methods

### Bacterial media and reagents

Lysogeny Broth Lennox (LBL) was prepared by dissolving 10 g/l tryptone, 5g/l yeast extract, and 5 g/l sodium chloride in deionized H<sub>2</sub>O and sterilized by autoclaving. Super Optimal Broth (SOB) was prepared by dissolving 20 g/l tryptone, 5 g/l yeast extract, 0.5 g/l sodium chloride, 2.4 g/l magnesium sulfate, and 0.186 g/l potassium chloride in deionized H<sub>2</sub>O and sterilized by autoclaving. 2×YT media consisted of 16 g/l casein digest peptone, 10 g/l yeast extract, 5 g/l 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<sub>4</sub>, 0.01% gelatin, was used for storing and diluting bacteriophage stocks (Geno Technology, Inc., St. Louis, MO, USA). L-4,4'-biphenylalanine (bipA) was obtained from PepTech Corporation (USA).

### Bacteriophage isolation

Bacteriophages were isolated from environmental samples from Massachusetts, United States (Table S1) by using *E. coli* Syn61Δ3(ev5) (from the laboratory of Jason W. Chin (Addgene strain #174514)) as host. For aqueous samples, including sewage, we directly used 50 ml filter-sterilized filtrates, while samples with mainly solid components, like soil and animal feces, were first resuspended to release phage particles and then sterilized by centrifugation and subsequent filtration. 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<sub>2</sub> and MgCl<sub>2</sub>. Infected cultures were grown overnight at 37 °C aerobically and then filter sterilized by centrifugation at 4000× g for 15 minutes 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 (OD<sub>600</sub> = 0.2), supplemented with 10 mM CaCl<sub>2</sub> and MgCl<sub>2</sub>, 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 on 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-titer stocks were prepared by mixing sterilized suspensions from single plaques with exponentially growing MDS42 cells ( $OD_{600} = 0.3$ ) in SOB supplemented with 10 mM  $CaCl_2$  and  $MgCl_2$ . Phage-infected samples were grown at 37 °C until complete lysis (~4 hrs) and then sterilized by filtration.

### **Bacteriophage culturing**

Bacteriophage stocks were prepared by a modified liquid lysate Phages on Tap protocol in LBL medium<sup>39</sup>. High-titer 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 resulted 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-titer 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**

Genomic TCR-suppressor tRNA-Ser<sub>UGA</sub> gene containing phages (based on Supplementary Table 1), corresponding to NCBI GenBank numbers MZ501046, MZ501058, MZ501065, MZ501066, MZ501067, MZ501074, MZ501075, MZ501089, MZ501096, MZ501098, MZ501105, MZ501106<sup>19</sup>, were obtained from DSMZ (Germany). Exponential phase cultures ( $OD_{600} = 0.3$ ) of MDS42 and Syn61 $\Delta$ 3(ev5) were grown in SOB supplemented with 10 mM  $CaCl_2$  and  $MgCl_2$  at 37 °C. Cultures were infected with phage at an 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. Infected cultures were grown at 37 °C with shaking at 250 rpm. After 24 hours, cultures were transferred to 1 ml tubes and centrifuged at 19,000x g to remove cells and cellular debris, and the clarified supernatant was serially diluted in SM buffer to enumerate output phage concentration. 1.5  $\mu$ l of the diluted supernatants were 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 hours of incubation at 37 °C, plaques were counted, and the number of plaques was multiplied by the dilution to calculate the phage titer of the original sample.

### **Single-step phage growth curve**

An exponential phase culture ( $OD_{600} = 0.3$ ) of Syn61 $\Delta$ 3 was grown in 50 ml SOB supplemented with 10 mM  $CaCl_2$  and  $MgCl_2$  at 37 °C with shaking at 250 rpm. Cultures were then spun down

and resuspended in 3 ml SOB supplemented with 10 mM CaCl<sub>2</sub> and MgCl<sub>2</sub>, and 1 ml samples were infected with REP12 phage at an MOI of 0.01. Infected cultures were incubated at 37 °C for 10 minutes without shaking for phage attachment and then washed twice with 1 ml SOB by pelleting cells at 4000× g for 3 minutes. Infected cells were then diluted into 50 ml SOB supplemented with 10 mM CaCl<sub>2</sub> and MgCl<sub>2</sub> and incubated at 37 °C with shaking at 250 rpm. At every 20 minutes, 1 ml sample was measured out into a sterile Eppendorf tube containing 100 µl chloroform, immediately vortexed, and then placed on ice. Phage titers were determined by centrifuging chloroformed cultures at 6000× g for 3 minutes 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<sub>2</sub> and MgCl<sub>2</sub>. Following 18 hours of incubation at 37 °C, plaques were counted, and the number of plaques was multiplied by the dilution to calculate the phage titer of the original sample.

### **Bacteriophage genome sequencing, assembly, and annotation**

Genomic DNA of bacteriophages was prepared from high-titer (*i.e.*, >10<sup>10</sup> PFU/mL) stocks after DNase treatment using the Norgen Biotek Phage DNA Isolation Kit (Cat# 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 bp UDI indices and sequenced on an Illumina NextSeq 2000, producing 150 bp paired-end reads. Demultiplexing, quality control, and adapter trimming were performed 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<sup>40</sup> with default parameters, except that the PHROGs HMM database<sup>41</sup> 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 Microbial Genome Sequencing Center (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 performed 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<sup>42</sup> in --sensitive-local mode. Single-nucleotide polymorphisms (SNPs) and indels were called using breseq (version 0.36.1)<sup>43</sup>. 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 Syn61 $\Delta$ 3(ev5) (from a single-colony isolate of Addgene strain #174514) was performed by generating 84,136 Oxford Nanopore (ONT) long-reads by PCR-free library generation (Oxford Nanopore, UK) on a MinION Flow Cell (R9.4.1) and  $4.5 \times 10^6$  150 bp paired-end reads on an Illumina NextSeq 2000. Quality control and adapter trimming were performed with bcl2fastq 2.20.0.445 and porechop 0.2.3\_seqan2.1.1 for Illumina and ONT sequencing, respectively. Next, we performed hybrid assembly with Illumina and ONT reads by using Unicycler 0.4.8 by using the default parameters. Finally, the resulted single, circular contig representing the entire genome was manually inspected for errors in Geneious Prime® 2022.1.1. and annotated based on sequence homology by using the BLAST function implemented in Geneious Prime® 2022.1.1. based on *E. coli* K-12 MG1655 (NCBI ID: U00096.3) as reference. Gene essentiality was determined based on Ref<sup>44</sup>.

### **Transcriptome analysis of phage-infected cells**

We explored transcriptomic changes and mRNA production in phage-infected Syn61 $\Delta$ 3 cells by performing a modified single-step growth experiment and collected samples at 20 minutes intervals. 50 ml of early-exponential ( $OD_{600} = 0.15$ ) Syn61 $\Delta$ 3 cells (corresponding to  $2 \times 10^{10}$  CFU) growing at 37 °C, 250 rpm in SOB containing 10 mM CaCl<sub>2</sub> and MgCl<sub>2</sub> were spun down at room temperature and resuspended in 1 ml of SOB. 50  $\mu$ l of this uninfected sample was immediately frozen in liquid N<sub>2</sub> and stored at -80 °C until RNA extraction. Next, 900  $\mu$ l of this cell suspension was mixed with 10 ml prewarmed REP12 phage stock (*i.e.*,  $\sim 7 \times 10^{10}$  PFU to achieve a MOI of  $\sim 4$ ) in SOB containing 10 mM CaCl<sub>2</sub> and MgCl<sub>2</sub>, and then incubated at 37 °C for 10 minutes 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<sub>2</sub> and MgCl<sub>2</sub>. Samples were then incubated at 37 °C, 250 rpm. After 20- and 40-minutes post-infection, we spun down 1 ml cell suspension from each sample, and the cell pellets were frozen in liquid nitrogen and stored at -80 °C until RNA extraction. As expected, after 60 minutes post-infection, no cell pellet was visible. Phage infections were performed in three independent replicates. Total RNA from frozen samples was extracted by using the RNeasy Mini Kit (Qiagen, USA) according to the manufacturer's instructions and the extracted RNA was DNase treated with Invitrogen RNase-free DNase

(Thermo Fisher Scientific, USA). Sequencing library preparation was then performed 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, USA). Sequencing was done on a NextSeq2000 instrument in 2x50 bp paired-end mode. Demultiplexing, quality control, and adapter trimming were performed with bcl-convert (v3.9.3). cDNA reads were aligned to their corresponding reference by using Bowtie2 2.3.0<sup>42</sup> in --sensitive-local mode, and read count and expression metrics were determined by using Geneious Prime® 2022.1.1. (Biomatters Ltd.). Finally, differential expression analysis was performed by using DESeq2<sup>45</sup> with standard settings.

### **tRNA sequencing sample preparation**

We explored tRNA expression levels and changes in phage-infected Syn61Δ3 cells by performing a modified single-step growth experiment with high MOI and cell mass. An early-exponential phase culture ( $OD_{600} = 0.2$ ) of Syn61Δ3 cells (corresponding to approximately  $5 \times 10^{10}$  CFU) growing at 37 °C, 250 rpm in SOB containing 10 mM CaCl<sub>2</sub> and MgCl<sub>2</sub> were spun down at room temperature and resuspended in 1.1 ml of SOB. 100 μl of this uninfected sample was immediately frozen in liquid N<sub>2</sub> and stored at -80 °C until tRNA extraction. Next, 1000 μl of this cell suspension was mixed with 20 ml prewarmed REP12 phage stock (*i.e.*,  $\sim 10^{12}$  PFU to achieve a MOI of  $\sim 20$ ) in SOB containing 10 mM CaCl<sub>2</sub> and MgCl<sub>2</sub>, and then incubated at 37 °C for 10 minutes 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<sub>2</sub> and MgCl<sub>2</sub>. Samples were then incubated at 37 °C, 250 rpm. Immediately after phage attachment and after 20- and 40-minutes post-infection, 1 ml cell suspensions from each sample were spun down, and cell pellets were frozen in liquid N<sub>2</sub> and stored at -80 °C until total RNA extraction. Phage infections were performed in two independent replicates.

We analyzed the total tRNA content of Ec\_Syn61Δ3-SL cells expressing KP869110.1 viral tRNA<sub>24-Leu<sub>UGA</sub></sub> and tRNA<sub>24-Leu<sub>CGA</sub></sub> by pelleting cells from 5 ml mid-exponential ( $OD_{600} = 0.3$ ) culture at 4000x g and flash-freezing the cell pellet in liquid nitrogen.

We extracted tRNAs by lysing samples at room temperature (RT) for 30 mins in 150 μl lysis buffer containing 8 mg/mL lysozyme (from chicken egg white, #76346-678, VWR, USA), 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 (#79306, Qiagen) and incubated for 5 minutes at RT. Next, 150 μl chloroform was added, vortexed, and incubated until phase separation. Samples were then spun at 15,000x g for 15 min at in a cooled centrifuge. The supernatant was

transferred into a new Eppendorf tube and mixed with 350  $\mu$ l 70% ethanol. Larger RNA molecules were then bound to an RNeasy MinElute spin column (#74204, Qiagen), and the flow-through was mixed with 450  $\mu$ l of 100% ethanol, and tRNAs were bound to a new RNeasy MinElute spin column. The tRNA fraction was then washed first with 500  $\mu$ l wash buffer (#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 minutes and then purified using a Micro Bio-Spin P-30 Gel Column (7326251, from Bio-Rad).

### **tRNA sequencing library preparation, sequencing, and data analysis**

We prepared tRNA cDNA libraries by reverse-transcribing tRNAs using the TGIRT<sup>TM</sup>-III template-switching reverse-transcriptase (TGIRT50, InGex, USA) according to the manufacturer's instructions. In brief, we prepared reaction mixtures containing 1  $\mu$ l (~100 ng) of the deacylated tRNAs, 2  $\mu$ l of 1  $\mu$ M TGIRT DNA/RNA heteroduplex (prepared by hybridizing equimolar amounts of rCrUrUrUrGrArGrCrCrUrArArUrGrCrCrUrGrArArArGrArUrCrGrGrArArGrArGrCrArCrArCrGrUrCrUrArGrUrUrCrUrArCrArGrUrCrCrGrArCrGrArU/3SpC3/ and ATCGTCGGACTGTAGAACTAGACGTGTGCTCTTCCGATCTTTCAGGCATTAGGCTCAAAGN oligos), 4  $\mu$ l 5 $\times$  TGIRT<sup>TM</sup> reaction buffer (2.25 M NaCl, 25 mM MgCl<sub>2</sub>, 100 mM Tris-HCl, pH 7.5), 2  $\mu$ l of 100 mM DTT, 9  $\mu$ l RNase-free water, and 1  $\mu$ l TGIRT-III, and incubated at room temperature for 30 minutes to initiate template-switching. Next, 1  $\mu$ l of 25 mM dNTPs (Thermo Fisher Scientific, USA) was added to the reaction mixture, and samples were incubated at 60  $^{\circ}$ C for 30 minutes to perform reverse transcription. RNA was then hydrolyzed by NaOH, neutralized by HCl, and the cDNA library was purified using MinElute PCR purification kit. cDNAs were then ligated to a preadenylated DNA adapter /5Phos/GATCNNNAGATCGGAAGAGCGTCGTGT/3SpC3/, in which NNN denotes an N, NN, NNN spacer to increase library diversity during sequencing (preadenylated oligos were prepared by 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 2x Master Mix (New England Biolabs). PCR products were then size selected to remove adaptor-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 performed 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 tRNA<sup>24-Leu<sub>UGA</sub></sup> and tRNA<sup>24-Leu<sub>CGA</sub></sup> were mapped without allowing the presence of SNPs in sequencing reads to distinguish tRNA<sup>24-Leu<sub>UGA</sub></sup> and tRNA<sup>24-Leu<sub>CGA</sub></sup> 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 performed 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<sup>46,47</sup>, Addgene #42876), and the λRed operon, consisting of *gam*, *exo*, and *bet* (from pORTMAGE311B<sup>48</sup>, Addgene #120418). Next, cells were made electrocompetent using a standard protocol<sup>46,47</sup> for Cas9-assisted recombineering and transformed with 2 μl of 100 μM 90 nucleotide-long ssDNA oligonucleotide inserting a stop codon and a frameshift mutation into *recA* (Supplementary Table 4). Successful edits were selected by cotransforming 1 μg from a variant of the pCRISPR plasmid<sup>46,47</sup> carrying a 5'-AGTTGATACCTTCGCCGTAG 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 resulted Syn61Δ3(ev5) Δ*recA* strain was validated by whole genome sequencing and then evolved for increased fitness (available in section “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<sup>30</sup> was performed by first transforming cells with a plasmid carrying a constitutively expressed MjTyrRS-derived bipA aaRS (variant 10, based on Ref<sup>49</sup>) 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 Table 4). 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 Table 4). Cells were grown in the presence of 200  $\mu$ M bipA in 2 $\times$ YT media throughout the entire procedure. Successful edits were selected using a dual-targeting crRNA expression construct, carrying 5'-GCAATGCGTATCATTCTGCT and 5'-GCCGTCAACTTTCGCGTATT guide sequences (from GenScript USA Inc.). Positive colonies were selected by screening colonies with allele-specific PCR (Supplementary Table 4) and validated by whole genome sequencing. Finally, the escape rate of the resulted Syn61 $\Delta$ 3(ev5)  $\Delta$ recA (ev1) adk.d6 strain was determined as described earlier<sup>30</sup>, but instead of chloramphenicol, cells were grown in the presence of 10  $\mu$ g/ml gentamycin in 2 $\times$ YT. Plates were incubated for seven days at 37 °C. Escape rate measurements were performed in triplicates;  $\pm$  indicates standard deviation.

### **Adaptive laboratory evolution of Syn61 $\Delta$ 3**

We performed standard adaptive laboratory evolution in rich bacterial media for 30 days (~270 cell generations) on Syn61 $\Delta$ 3(ev5)  $\Delta$ recA cells to increase fitness. At each transfer step, 10<sup>9</sup>-10<sup>10</sup> bacterial cells were transferred into 500 ml LBL medium containing 1.5 g/l Tris/Tris-HCl and incubated aerobically for 24 hours at 37 °C, 250 rpm in a 2000 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 resulted evolved variant, Syn61 $\Delta$ 3(ev5)  $\Delta$ recA (ev1), are available in Supplementary Table 4.

### **Doubling time measurements**

To determine growth parameters under standard laboratory conditions, saturated overnight cultures of *E. coli* Syn61 $\Delta$ 3(ev5), Syn61 $\Delta$ 3(ev5)  $\Delta$ recA, and its evolved variant, Syn61 $\Delta$ 3(ev5)  $\Delta$ recA (ev1) were diluted 1:200 into 50 ml of 2 $\times$ YT and LBL in a 300 ml Erlenmeyer flask with vented cap and incubated aerobically at 37 °C, 250 rpm. Ec\_Syn61 $\Delta$ 3-SL cells were characterized similarly, but by using 2 $\times$ YT containing 50  $\mu$ g/ml kanamycin. All growth measurements were performed in triplicates. Optical density at 600 nm (OD<sub>600</sub>) measurements were taken every 20 minutes for 8 hours or until stationary phase was reached on a CO8000 Cell Density Meter, WPA. The doubling time was calculated for each independent replicate by log<sub>2</sub>-transforming OD<sub>600</sub> 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<sub>2</sub>-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 performed with Prism9 (GraphPad). Error bars show  $\pm$  standard deviation.

### **tRNA annotation**

We detected tRNA genes in the Viral genomic NCBI Reference Sequence Database (Accessed: January 2, 2022) and in individual phage isolates' genomes by using tRNAscan-SE 2.0.9 in bacterial (-B), archaeal (-A), or eukaryotic (-E) maximum sensitivity mode (-I --max)<sup>50</sup>. 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 Table 1), driven by a strong bacterial *proK* tRNA promoter and followed by two transcriptional terminators as 10 pmol ssDNA oligonucleotide libraries (10 pmol oPool, from Integrated DNA Technologies, USA). Oligonucleotides were resuspended in 1x 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 hrs 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). Prior to electrotransformation, bacterial cells were made electrocompetent by growing cells after a 1:100 dilution in SOB until mid-log phase (OD=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 kanamycin in 145x20 mm Petri dishes (Greiner Bio-One). Plates were incubated at 37 °C until colony formation. Approximately 1000-5000 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 performed with bcl-convert (v3.9.3). Paired-end reads were then trimmed using BBDuk from BBTools (in Geneious Prime® 2022.1.1., Biomatters Ltd.), 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<sub>YGA</sub> library generation and selection**

We identified leucine tRNAs that can translate TCR codons as leucine by performing 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<sub>YGA</sub> expression library consisting of bacteriophage derived Leu 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<sub>UGA</sub>-spacer-leuU<sub>CGA</sub>-*proK* terminator sequence, in which the anticodon loop of both *leuU* tRNAs has been fully randomized, as an oPool library (Supplementary Table 4) (Integrated DNA Technologies, USA). 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*<sub>29xLeu→TCR</sub>) 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, USA). The ligation was performed at a 3:1 insert-to-vector ratio and by using T4 DNA ligase (New England Biolabs, USA) for 16 hours at 16 °C according to the manufacturer's instructions. Finally, the ligation product was purified using the GeneJet PCR purification kit (Thermo Fischer Scientific, USA). We constructed the second, 13-member tRNA-Leu<sub>YGA</sub> expression library (Supplementary Table 4) consisting of bacteriophage-derived Leu tRNA variants bearing a UGA and CGA anticodon by using the same method as for our *leuU*<sub>YGA</sub> 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 hours, 250 rpm. After recovery, the cells were plated to 2xYT agar plates containing kanamycin at 200 μg/ml 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, USA) and subjected to whole plasmid sequencing (SNPsaurus, Eugene, Oregon, US).

### **Virus resistance analysis of Ec\_Syn61Δ3-SL cells**

An exponential phase culture (OD<sub>600</sub> = 0.3) of the corresponding strain was grown in 3 ml SOB supplemented with 10 mM CaCl<sub>2</sub> and MgCl<sub>2</sub> and 75 μg/ml kanamycin at 37 °C with shaking. Cultures were then spun down and resuspended in 1 ml SOB supplemented with 10 mM CaCl<sub>2</sub> and MgCl<sub>2</sub> and infected with a 1:1 mixture of all 12 Syn61Δ3-lytic phage isolates from this study

(Table S2) at an MOI of 0.1. Infected cultures were incubated at 37 °C without shaking for 10 minutes for phage attachment and then washed three times with 1 ml SOB supplemented with 10 mM CaCl<sub>2</sub> and MgCl<sub>2</sub> by pelleting cells at 4000× g for 3 minutes. Infected cells were then diluted into 4 ml SOB supplemented with 10 mM CaCl<sub>2</sub> and MgCl<sub>2</sub> and 75 µg/ml kanamycin and incubated at 37 °C with shaking at 250 rpm. After 24 hours 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 performed in three independent replicates. Phage titers were determined by centrifuging chloroformed cultures at 6000× g for 3 minutes 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<sub>2</sub> and MgCl<sub>2</sub>. Following 18 hours of incubation at 37 °C, plaques were counted, and the number of plaques was multiplied by the dilution to calculate the phage titer of the original sample.

Phage enrichment experiments were performed by mixing 50 ml early exponential phase cultures (OD<sub>600</sub> = 0.2) of bipA-biocontained Ec\_Syn61Δ3-SL carrying pLS1 and pLS2 plasmids with 10 ml environmental sample mix, containing the mixture of Sample 2-13 from our study (Table S1). Infected Ec\_Syn61Δ3-SL cells with the corresponding plasmid were grown overnight in SOB supplemented with 200 mM bipA, 10 mM CaCl<sub>2</sub> and MgCl<sub>2</sub>, and 75 µg/ml kanamycin at 37 °C with shaking at 250 rpm. On the next day, cells were removed by centrifugation at 4000× g for 20 minutes, 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<sub>600</sub> = 0.2) of the corresponding strain, incubated for 20 minutes at 37 °C for phage absorption, pelleted by centrifugation at 4000× g for 15 minutes, and then resuspended in 50 ml SOB supplemented with 200 mM bipA, 10 mM CaCl<sub>2</sub> and MgCl<sub>2</sub>, and 75 µg/ml kanamycin. Infected cultures were then incubated at 37 °C with shaking at 250 rpm. Cultures were grown overnight and then sterilized by centrifugation at 4000× g for 15 minutes and filtered through a 0.45 µm PVDF Steriflip™ disposable vacuum filter unit (MilliporeSigma). Finally, phage titers were determined by using MDS42 cells as above. Phage enrichment experiments were performed in two independent replicates. The lytic phage titer of the unenriched sample mix was determined by diluting 100 µl of the input environmental sample mix, containing a mixture of Sample 2-13 (Table S1), 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 Table 4 were synthesized as gBlocks by IDT and circularized either by ligation with T4 DNA ligase (New England Biolabs, USA), or, in the case of pSC101 and RK2 plasmid-derived variants, by isothermal assembly using the HiFi DNA Assembly Master Mix (New England Biolabs, USA). Following assembly, purified plasmid assemblies were electroporated into Ec\_Syn61 $\Delta$ 3-SL cells carrying pLS1. pLS1 and pLS2 were designed to express two distinct combinations of the previously identified phage tRNA-Leu<sub>YGA</sub> tRNAs in antiparallel orientation to avoid repeat-mediated instability<sup>51</sup>, together with a pUC origin-of-replication, the *aph3Ia*<sub>29xLeu $\rightarrow$ TCR</sub> and *aminoglycoside-(3)-N-acetyltransferase*<sub>18xLeu $\rightarrow$ TCR</sub> marker genes. Transformants carrying either pLS1 or pLS2 were identified by transforming assemblies into Syn61 $\Delta$ 3(ev5)  $\Delta$ recA (ev1) and selecting for kanamycin resistance. Finally, plasmids from antibiotic-resistant clones were purified using a Monarch plasmid miniprep kit (New England Biolabs, USA) and subjected to whole plasmid sequencing (SNPsaurus, Eugene, Oregon, US).

### Escape rate analysis of viral Leu-tRNA<sub>YGA</sub>s and pLS plasmids

We analyzed the ability of pLS plasmids to function outside Ec\_Syn61 $\Delta$ 3-SL cells by transforming extracted plasmids into *E. coli* K-12 MG1655. Plasmids were purified from biocontained Ec\_Syn61 $\Delta$ 3-SL cells, carrying either pLS1 or pLS2 to express tRNA-Leu<sub>YGA</sub>, or pLS1 together with pLS3-5, by using the PureLink™ Fast Low-Endotoxin Midi Plasmid Purification Kit (Thermo Fisher Scientific, USA) according to the manufacturer's instructions. Next, we electroporated 1  $\mu$ g from each plasmid prep into freshly made electrocompetent cells of *E. coli* K-12 MG1655. Cells were made electrocompetent by diluting an overnight SOB culture of MG1655 1:100 into 500 ml SOB in a 2000 ml flask and growing cells aerobically at 32 °C with shaking at 250 rpm. At OD<sub>600</sub> = 0.3, cells were cooled on ice and then pelleted by centrifugation and resuspended in 10% glycerol-in-water. Cells were washed 4-times with 10% glycerol-in-water and then resuspended in 400  $\mu$ l 20% glycerol-in-water. 1000 ng from each plasmid sample was then mixed with 80  $\mu$ l 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  $\mu$ F). Electroporations were performed in three 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 rpm. Finally, 500  $\mu$ l from each recovery culture was plated to LBL agar plates containing antibiotics corresponding to the given pLS plasmid's resistance marker (15  $\mu$ g/ml gentamycin plus 50  $\mu$ g/ml kanamycin in the case of pLS1 and 2; 100  $\mu$ g/ml carbenicillin for pLS4, 30  $\mu$ g/ml chloramphenicol for pLS3 and pLS6, and 20  $\mu$ g/ml gentamycin for pLS5) in

145x20 mm Petri dishes (Greiner Bio-One). Plates were incubated at 37 °C for seven days and inspected for growth. Electroporation efficiency measurements were performed by electroporating a plasmid carrying a pUC origin-of-replication and kanamycin resistance into MG1655 electrocompetent cells under identical conditions.

### **Cloning of REP12 tRNA operon**

We analyzed the incorporated amino acid in Syn61 $\Delta$ 3 cells bearing the tRNA operon and its native promoter from the REP12 phage by subcloning the genomic tRNA operon into a low-copy 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, USA) from extracted phage gDNA and purified using the GeneJet PCR purification kit (Thermo Fischer Scientific, USA). 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, USA). After incubation for 60 mins at 50 °C, the assembly was purified using the DNA Concentrator & Clean kit (Zymo Research, USA) and transformed into Syn61 $\Delta$ 3(ev5)  $\Delta$ *recA* (ev1) cells expressing an MSKGPGKVPGAGVPGxGVPGVGKGGGT- elastin peptide fused to sfGFP with a terminal 6xHis tag (in which **x** denotes the analyzed codon, TCA or TCG) on a plasmid containing a kanamycin resistance gene and a pUC origin of replication<sup>52</sup>. Following an overnight recovery at 32 °C, cells were plated to 2xYT agar plates containing kanamycin and chloramphenicol. Finally, plasmid sequences in outgrowing colonies were validated by whole-plasmid sequencing. Elastin<sub>16TCR</sub>-sfGFP-6xHIS expression measurements have been performed as described below.

### **Elastin<sub>16TCA</sub>-sfGFP-6xHIS expression measurements**

We assayed the amino acid identity of the serine tRNA<sub>UGA</sub> tRNAs of MZ501075 and REP12 (see Supplementary Table 4 for sequence information) by coexpressing selected tRNAs and a constitutively expressed MSKGPGKVPGAGVPGxGVPGVGKGGGT- elastin peptide fused to sfGFP with a terminal 6xHis tag (in which **x** denotes the analyzed codon, TCA or TCG)<sup>52</sup> on a plasmid containing a kanamycin resistance gene and a pUC origin-of-replication in Syn61 $\Delta$ 3(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 utilized the same elastin-sfGFP-6xHIS 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 hours at 37 °C, 200 rpm aerobically. We then determined sfGFP expression levels in samples by pelleting and washing 1 ml of the culture with PBS and resuspending cell pellets in 110 µL BugBuster Protein Extraction Reagent (MilliporeSigma). Reactions were incubated for 5 minutes and then spun down at 13,000× rpm for 10 minutes. The fluorescence of the BugBuster-treated supernatants and the OD<sub>600</sub> of the original culture was measured using the Synergy H1 Hybrid Reader (BioTek) plate reader using the bottom mode analysis with an excitation at 480 nm and emission measurement at 515 nm, with the gain set to 50. Fluorescence values were normalized based on OD<sub>600</sub> 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 minutes. The lysed cell mixture was spun down at 13,000× rpm for 10 minutes, and the supernatant was mixed in a 1:1 ratio with HIS-Binding/Wash Buffer (G-Biosciences, USA) and 50 µl HisTag Dynabeads (Thermo Fischer Scientific, USA). Following an incubation period of 5 minutes, the beads were separated on a magnetic rack and washed with 300 µl HIS-Binding/Wash Buffer and PBS (Phosphate Buffered Saline) three times. After the last wash step, the bead pellets containing the bound elastin-sfGFP-6×HIS protein samples were frozen at -80 °C until MS/MS sample preparation. Protein production experiments were performed in three independent replicates.

### **Tandem liquid chromatography and mass spectrometry (LC/MS-MS) analysis of tryptic elastin-sfGFP-6×HIS**

Samples from elastin-sfGFP-6×HIS expression experiments were digested directly on HisTag Dynabeads according to the FASP digest procedure<sup>53</sup>. In brief, samples were washed with 50 mM TEAB (triethylammonium bicarbonate buffer) and then rehydrated with 50 mM TEAB-trypsin solution, followed by a three-hour digest at 50 °C. Digested peptides were then separated from HisTag Dynabeads and concentrated by spinning and drying samples at 3,000× rpm using a SpeedVac concentrator. Samples were then solubilized in 0.1% formic acid-in-water for subsequent analysis by tandem mass spectrometry. LC-MS/MS analysis of digested samples was performed on a Lumos Tribrid Orbitrap Mass Spectrometer equipped with an Ultimate 3000 nano-HPLC (both from Thermo Fisher Scientific, USA). 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 GmbH, Germany) followed by a 50 cm analytical column (PharmaFluidics, Belgium). Separation was achieved by applying a gradient from 4% to 30%



acetonitrile in 0.1% formic acid over 60 mins at 200 nl/min. Electrospray ionization was performed 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, Denmark). The mass spectrometry survey scan was performed in the Orbitrap in the range of 400-1,800 m/z at a resolution of  $6 \times 10^4$ , followed by the selection of the twenty 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, AGC (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 seconds.

The raw data were analyzed using Proteome Discoverer 2.4 (Thermo Fisher Scientific, USA). Assignment of MS/MS spectra was performed using the Sequest HT algorithm by searching the data against a protein sequence database, including all protein entries from *E. coli* K-12 MG1655, all proteins sequences of interest (including the elastin-sfGFP fusion protein), as well as other known contaminants such as human keratins and common lab contaminants. Quantitative analysis between samples was performed by LFQ (label free quantitation) between different samples. Sequest HT searches were performed using a 10-ppm precursor ion tolerance and requiring each peptides N-/C termini to adhere 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-terminus acetylation (+42.01057 Da) was set as variable modifications. We then determined the amino acid incorporated at position **x** in our elastin-sfGFP-6xHis construct by analyzing changes compared to Phe. To cover all 20 possible amino acid exchange cases at the **x** position, we performed five separate searches with four different amino acids as possible variable modifications in each search. All cysteines were set to permanent no modification due to no alkylation procedure. An overall false discovery rate of 1% on both protein and peptide level was achieved by performing target-decoy database search using Percolator<sup>54</sup>.

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

We analyzed the translation of viral proteins in Ec\_Syn61Δ3-SL cells (Syn61Δ3(ev5) ΔrecA (ev1) expressing a *proK* promoter-driven Leu9-tRNA<sub>YGA</sub> construct from Escherichia phage OSYSP (GenBank ID MF402939.1) and APH(3')-I aminoglycoside O-phosphotransferase

(*aph3la*<sub>29xLeu→TCR</sub>), on a high copy-number pUC plasmid), by performing a modified single-step growth experiment and subsequent time-course tandem mass spectrometry-based proteome analysis. An early-exponential phase culture (OD<sub>600</sub> = 0.2) of Ec\_Syn61Δ3-SL cells (corresponding to approximately 4×10<sup>10</sup> CFU) growing at 37 °C, 250 rpm in SOB containing 10 mM CaCl<sub>2</sub>, MgCl<sub>2</sub>, and 75 μg/ml kanamycin were spun down at room temperature and resuspended in 1.1 ml SOB containing 10 mM CaCl<sub>2</sub>, MgCl<sub>2</sub>, and 75 μg/ml kanamycin. 100 μl of this uninfected sample was immediately frozen in liquid N<sub>2</sub> and stored at -80 °C until proteome analysis. Next, 1000 μl of this cell suspension was mixed with 10 ml prewarmed REP12 phage stock (*i.e.*, ~5×10<sup>11</sup> PFU to achieve a MOI of ~12) in SOB containing 10 mM CaCl<sub>2</sub>, MgCl<sub>2</sub>, and 75 μg/ml kanamycin, and then incubated at 37 °C for 10 minutes 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<sub>2</sub> and MgCl<sub>2</sub>. Samples were then incubated at 37 °C, 250 rpm. After 20- and 40-minutes post-infection, 1 ml cell suspensions were spun down, and cell pellets were frozen in liquid N<sub>2</sub> 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<sup>53</sup>. In brief, samples were washed with 50 mM TEAB buffer on a 10 kDa cutoff filter (Pall Corp, CA) and then rehydrated with 50 mM TEAB-trypsin solution, followed by a three-hour digest at 37 °C. Digested peptides were then extracted and separated into 10 fractions by using the Pierce™ High pH Reversed-Phase Peptide Fractionation Kit according to the manufacturer's protocol (Thermo Fisher Scientific, USA). Following fractionation, peptides were concentrated and dried by spinning samples at 3.000× rpm using a SpeedVac concentrator. Samples were then solubilized in 0.1% formic acid-in-water for subsequent analysis by tandem mass spectrometry. LC-MS/MS analysis of digested samples was performed on a Lumos Tribrid Orbitrap Mass Spectrometer equipped with an Ultimate 3000 nano-HPLC (both from Thermo Fisher Scientific, USA). 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 GmbH, Germany) followed by a 50 cm analytical column (PharmaFluidics, Belgium). Separation was achieved by applying a gradient from 5% to 27% acetonitrile in 0.1% formic acid over 90 mins at 200 nl/min. Electrospray ionization was performed 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, Denmark). The mass spectrometry survey scan was performed in the Orbitrap in the range of 400-1,800 m/z at a resolution of 6×10<sup>4</sup>, followed by the selection of the twenty 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, AGC (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 seconds.

The raw data was analyzed using Proteome Discoverer 2.4 (Thermo Fisher Scientific, USA). Assignment of MS/MS spectra was performed using the Sequest HT algorithm by searching the data against a protein sequence database, including all protein entries from *E. coli* K-12 MG1655, all protein sequences of the corresponding REP12 bacteriophage and the Aph3la APH(3')-I aminoglycoside O-phosphotransferase, as well as other known contaminants such as human keratins and common lab contaminants. Quantitative analysis between samples was performed by LFQ (label free quantitation) between different samples. Sequest HT searches were performed using a 10-ppm precursor ion tolerance and requiring each peptides N-/C termini to adhere 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-terminus acetylation (+42.01057 Da) was 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 variable modification. All cysteines were set to permanent no modification due to no alkylation procedure. An overall false discovery rate of 1% on both protein and peptide levels was achieved by performing target-decoy database search using Percolator<sup>54</sup>.

## References

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. *cebs* **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. Ostrov, N. *et al.* Synthetic genomes with altered genetic codes. *Current Opinion in Systems Biology* (2020) doi:10.1016/j.coisb.2020.09.007.
7. Fredens, J. *et al.* Total synthesis of Escherichia coli with a recoded genome. *Nature* **1** (2019) doi:10.1038/s41586-019-1192-5.
8. Abrahão, J. *et al.* Tailed giant Tupanvirus possesses the most complete translational apparatus of the known virosphere. *Nature Communications* **9**, 1–12 (2018).
9. Morgado, S. & Vicente, A. C. Global In-Silico Scenario of tRNA Genes and Their Organization in Virus Genomes. *Viruses* **11**, 180 (2019).
10. Al-Shayeb, B. *et al.* Clades of huge phages from across Earth's ecosystems. *Nature* **578**, 425–431 (2020).
11. Yang, J. Y. *et al.* Degradation of host translational machinery drives tRNA acquisition in viruses. *Cell Systems* **12**, 771-779.e5 (2021).
12. Peters, S. *et al.* Validation that human microbiome phages use alternative genetic coding with TAG stop read as Q. <http://biorxiv.org/lookup/doi/10.1101/2022.01.06.475225> (2022) doi:10.1101/2022.01.06.475225.

13. Borges, A. L. *et al.* Stop codon recoding is widespread in diverse phage lineages and has the potential to regulate translation of late stage and lytic genes. 35.
14. Abe, T. *et al.* tRNADB-CE 2011: tRNA gene database curated manually by experts. *Nucleic Acids Research* **39**, D210–D213 (2011).
15. Alamos, P. *et al.* Functionality of tRNAs encoded in a mobile genetic element from an acidophilic bacterium. *RNA Biology* **15**, 518–527 (2018).
16. Santamaría-Gómez, J. *et al.* Role of a cryptic tRNA gene operon in survival under translational stress. *Nucleic Acids Research* (2021) doi:10.1093/nar/gkab661.
17. Bustamante, P. *et al.* ICEAfe1, an Actively Excising Genetic Element from the Biominer Bacterium *Acidithiobacillus ferrooxidans*. *MMB* **22**, 399–407 (2012).
18. Bowden, R. J., Simas, J. P., Davis, A. J. & Efstathiou, S. 1997. Murine gammaherpesvirus 68 encodes tRNA-like sequences which are expressed during latency. *Journal of General Virology* **78**, 1675–1687.
19. Maffei, E. *et al.* Systematic exploration of *Escherichia coli* phage–host interactions with the BASEL phage collection. *PLOS Biology* **19**, e3001424 (2021).
20. 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).
21. Miles, Z. D., McCarty, R. M., Molnar, G. & Bandarian, V. Discovery of epoxyqueuosine (oQ) reductase reveals parallels between halorespiration and tRNA modification. *Proceedings of the National Academy of Sciences* **108**, 7368–7372 (2011).
22. 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 Research* **43**, 7489–7503 (2015).
23. Fujino, T., Tozaki, M. & Murakami, H. An Amino Acid-Swapped Genetic Code. *ACS Synth. Biol.* **9**, 2703–2713 (2020).

24. Church, G., Baynes, B. & Pitcher, E. Hierarchical Assembly Methods for Genome Engineering. (2007).
25. Schmidt, M. & Kubyskin, V. How To Quantify a Genetic Firewall? A Polarity-Based Metric for Genetic Code Engineering. *ChemBioChem* **22**, 1268–1284 (2021).
26. 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 Research* **50**, 4713–4731 (2022).
27. Tomita, K., Ogawa, T., Uozumi, T., Watanabe, K. & Masaki, H. A cytotoxic ribonuclease which specifically cleaves four isoaccepting arginine tRNAs at their anticodon loops. *Proceedings of the National Academy of Sciences* **97**, 8278–8283 (2000).
28. 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* tRNA<sup>Ser1</sup>. *Biochemical and Biophysical Research Communications* **257**, 662–667 (1999).
29. 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 Letters* **447**, 1–4 (1999).
30. Mandell, D. J. *et al.* Biocontainment of genetically modified organisms by synthetic protein design. *Nature* **518**, 55–60 (2015).
31. Kunjapur, A. M. *et al.* Synthetic auxotrophy remains stable after continuous evolution and in coculture with mammalian cells. *Science Advances* **7**, eabf5851 (2021).
32. Rovner, A. J. *et al.* Recoded organisms engineered to depend on synthetic amino acids. *Nature* **518**, 89–93 (2015).
33. Nyerges, Á. *et al.* A highly precise and portable genome engineering method allows comparison of mutational effects across bacterial species. *PNAS* 201520040 (2016) doi:10.1073/pnas.1520040113.

34. Wannier, T. M. *et al.* Recombineering and MAGE. *Nature Reviews Methods Primers* **1**, 1–24 (2021).
35. Ostrov, N. *et al.* Design, synthesis, and testing toward a 57-codon genome. *Science* **353**, 819–822 (2016).
36. Chen, Y. *et al.* Multiplex base editing to convert TAG into TAA codons in the human genome. 2021.07.13.452007 Preprint at <https://doi.org/10.1101/2021.07.13.452007> (2021).
37. Boeke, J. D. *et al.* The Genome Project–Write. *Science* aaf6850 (2016) doi:10.1126/science.aaf6850.
38. Kirchberger, P. C. & Ochman, H. Resurrection of a global, metagenomically defined gokushovirus. *eLife* **9**, e51599 (2020).
39. Bonilla, N. *et al.* Phage on tap—a quick and efficient protocol for the preparation of bacteriophage laboratory stocks. *PeerJ* **4**, e2261 (2016).
40. Seemann, T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* **30**, 2068–2069 (2014).
41. Terzian, P. *et al.* PHROG: families of prokaryotic virus proteins clustered using remote homology. *NAR Genomics and Bioinformatics* **3**, lqab067 (2021).
42. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**, 357–359 (2012).
43. 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).
44. Goodall, E. C. A. *et al.* The Essential Genome of Escherichia coli K-12. *mBio* **9**, e02096-17 (2018).
45. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15**, 550 (2014).

46. 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).
47. 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).
48. Szili, P. *et al.* Rapid evolution of reduced susceptibility against a balanced dual-targeting antibiotic through stepping-stone mutations. *Antimicrobial Agents and Chemotherapy* AAC.00207-19 (2019) doi:10.1128/AAC.00207-19.
49. Kunjapur, A. M. *et al.* Engineering posttranslational proofreading to discriminate nonstandard amino acids. *PNAS* **115**, 619–624 (2018).
50. 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 Research* **49**, 9077–9096 (2021).
51. Hossain, A. *et al.* Automated design of thousands of nonrepetitive parts for engineering stable genetic systems. *Nat Biotechnol* **38**, 1466–1475 (2020).
52. Mohler, K. *et al.* MS-READ: Quantitative measurement of amino acid incorporation. *Biochimica et Biophysica Acta (BBA) - General Subjects* **1861**, 3081–3088 (2017).
53. Wiśniewski, J. R., Zougman, A., Nagaraj, N. & Mann, M. Universal sample preparation method for proteome analysis. *Nat Methods* **6**, 359–362 (2009).
54. 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).