Supplementary Material

A swapped genetic code prevents viral infections and gene transfer

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

[1]Department of Genetics, Harvard Medical School, Boston, MA 02115, USA

[2]Department of Biomedical Informatics and Laboratory of Systems Pharmacology, Harvard Medical School, Boston, MA 02115, USA

[3]Wyss Institute for Biologically Inspired Engineering, Harvard University, Boston, MA 02115, USA

[4]Department of Pathology and Immunology, Washington University School of Medicine in St. Louis, St. Louis, MO, USA

[5]The Edison Family Center for Genome Sciences and Systems Biology, Washington University School of Medicine in St. Louis, St. Louis, MO, USA

[6]GenScript USA Inc., Piscataway, NJ 08854, USA

*Corresponding author. E-mail: gchurch@genetics.med.harvard.edu (G.M.C.); akos_nyerges@hms.harvard.edu (A.N.)

Supplementary Note

Comparison of tRNA expression levels and phage resistance assays between Zürcher, J. F. et al.[33] and this work.

Comparison of mistranslating tRNA expression:

	Plasmid copy/cell	tRNA promoter	Relative tRNA expression level
This work	pUC-based, ~400	*proK* promoter (Relative promoter strength = 2.4);	~192 for *proK*-based expression, as in tRNA screen of Fig. 3c*;
	pUC-based, ~400	SLP2018-2-101 promoter[62] (Relative promoter strength = 775)	~62,000 for SLP2018-2-101-based expression, as in pLS1 and pLS2 in Ec_Syn61Δ3-SL.
Zürcher, J. F. et al.	pSC101-based, ~5	*lpp* promoter (Relative promoter strength = 1.0)	1.0

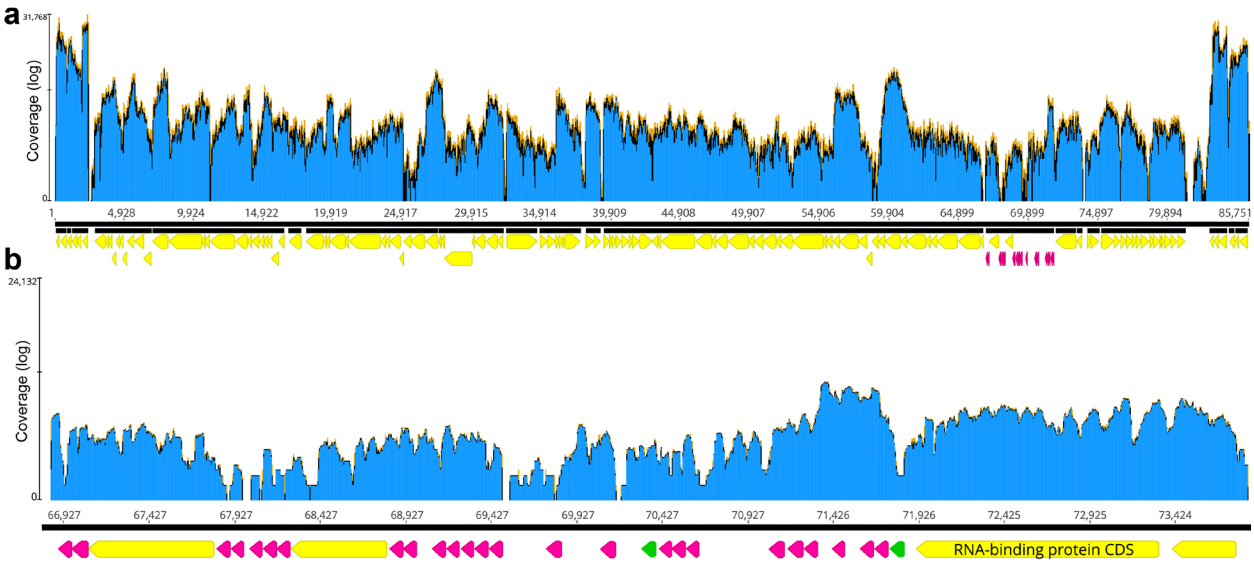
*According to our quantitative tRNAseq measurements (**Extended Data Figure 6**), the pUC plasmid + *proK* promoter-based expression system results in an expression level of the mistranslating viral Leu-tRNA$_{YGA}$ tRNAs that is similar to the cellular serine tRNAs, *i.e.,* a relative expression level of 172% and 140% for Leu-tRNA$_{UGA}$ and Leu-tRNA$_{CGA}$ respectively, compared to *serV*. This *proK*-based system provides broad virus resistance, as demonstrated on **Fig. 3c**.

Viral resistance assays and their sensitivity:

Assay(s) to determine viral resistance		
This work	Modified single-step growth curve with virion counting on phage-susceptible MDS42 *E. coli* strain (estimated detection limit: 200 PFU/ml, **detects slowly-replicating viruses**); Experiments **performed with 12 viruses**, including REP8-12 phages ~100× more virulent than T4-like phages (*i.e.,* REP1-4), Fig. 2a & Fig. 3c.	Enrichment-based resistance **analysis using environmental samples** (estimated detection limit: 200 PFU/ml post-enrichment, ~1 infecting virion / ml pre-enrichment**). Analyzed environmental samples contained ~3×10^7 coliphages.
Zürcher, J. F. et al.	Efficiency-of-plating assay directly on resistant strain (detection limit not reported; **previously noted to not detect slowly-replicating or non-plaque-forming viruses**[34,35]), Experiments **performed with two T4-like phages** (Tequatrovirus)**,** Φ6 and Φ12.	**Not tested using environmental samples.**

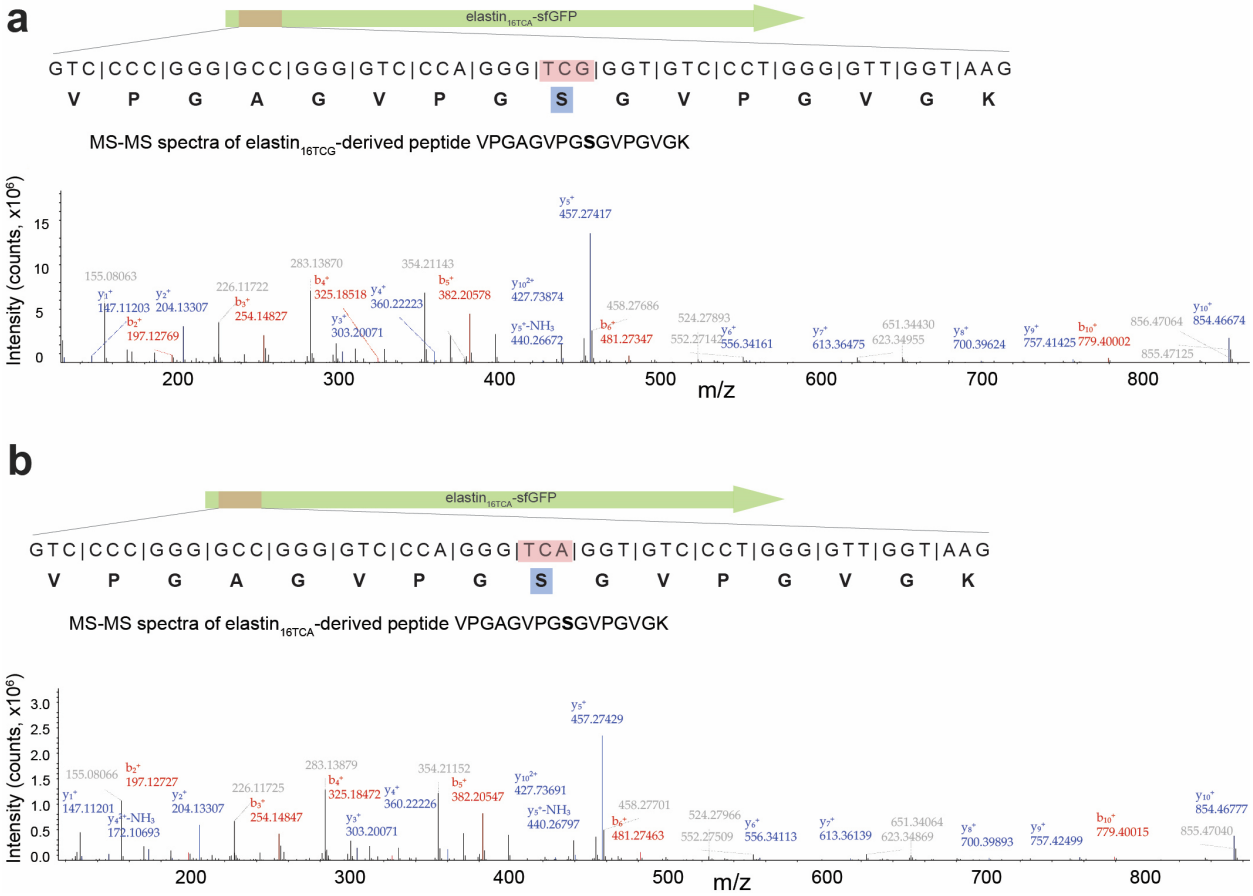
**Based on the dilution factors in our enrichment-based analysis (Methods), if one phage produces ten progenies (burst size = 10) and the phage replication cycle is ~24 hours, the overall detection limit is 100 PFU/ml. If any virus replicates faster than ~12 hours, the detection limit is 1 PFU/ml.

Supplementary Figure 1.



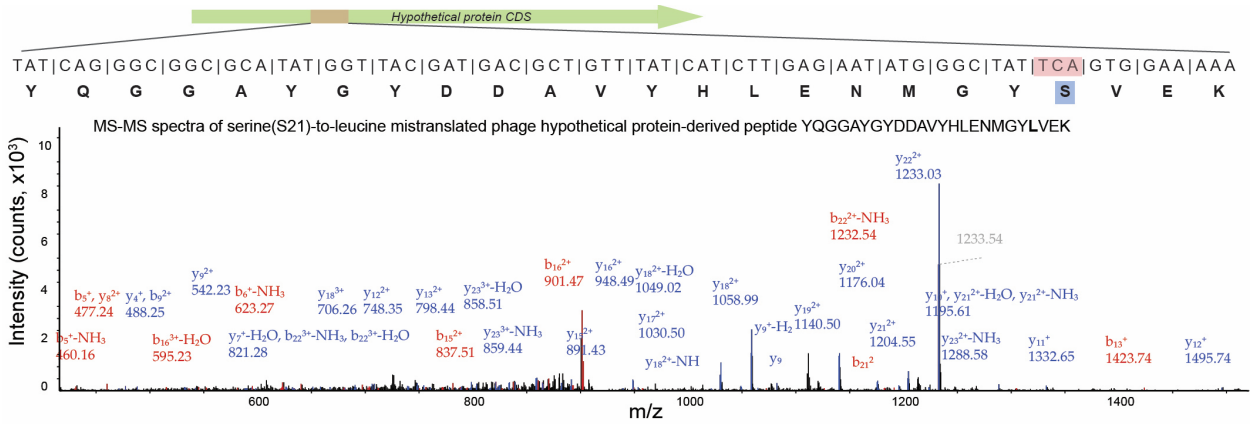
Supplementary Figure 1. Gene expression of REP12. (**a**) Transcriptomic analysis of the REP12 genome at 20 minutes post-infection in Syn61Δ3 cells. Blue shading indicates RNA sequencing read coverage. Numbers indicate genomic positions on the REP12 genome. (**b**) Predicted promoters (marked in green) and transcriptomic analysis of phage RNA expression at the genomic region of REP12 phage containing tRNA genes. tRNAs are marked as magenta arrows, and protein-coding sequences are marked as yellow arrows. Blue shading indicates RNA sequencing read coverage at 20 minutes post-infection based on a single experiment. *E. coli* σ^{70} promoters were predicted using the statistical thermodynamic model of transcriptional initiation described in Ref[67]. Source data is available within this paper.

Supplementary Figure 2.



Supplementary Figure 2. REP12 viral tRNA operon-expressed tRNA-Ser$_{UGA}$ decodes TCR codons as serine. The amino acid identity of the translated TCG and TCA codon within elastin$_{16TCR}$-sfGFP-His6 was confirmed by tandem mass spectrometry from Syn61Δ3 cells containing the REP12 viral tRNA operon together with its native promoter (**a**) or the REP12 viral tRNA operon-derived tRNA-Ser$_{UGA}$, driven by a *proK* promoter (**b**). The figure shows the amino acid sequence and MS/MS spectrum of the analyzed elastin$_{16TCR}$ peptide. MS/MS data was collected once.

Supplementary Figure 3.



Supplementary Figure 3. Serine-to-leucine mistranslation of viral TCR serine codons in Ec_Syn61Δ3-SL cells (Example 2). The figure shows the amino acid sequence and MS/MS spectrum of a REP12 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$_{YGA}$ from *Escherichia* phage OSYSP (GenBank ID MF402939.1), with the REP12 lytic phage at an MOI = 12, and the proteome of infected cells was analyzed by tandem mass spectrometry (Methods). MS/MS data was collected once.

References:

67. Fleur, T. L., Hossain, A. & Salis, H. M. Automated Model-Predictive Design of Synthetic Promoters to Control Transcriptional Profiles in Bacteria. BioRxiv, 2021.09.01.458561 Preprint at https://doi.org/10.1101/2021.09.01.458561 (2021).