RESEARCH ARTICLE

SYNTHETIC BIOLOGY

Refactored genetic codes enable bidirectional genetic isolation

Jérôme F. Zürcher[1], Wesley E. Robertson[1], Tomás Kappes[2], Gianluca Petris[1], Thomas S. Elliott[1], George P. C. Salmond[2], Jason W. Chin[1]*

The near-universal genetic code defines the correspondence between codons in genes and amino acids in proteins. We refactored the structure of the genetic code in *Escherichia coli* and created orthogonal genetic codes that restrict the escape of synthetic genetic information into natural life. We developed orthogonal and mutually orthogonal horizontal gene transfer systems, which permit the transfer of genetic information between organisms that use the same genetic code but restrict the transfer of genetic information between organisms that use different genetic codes. Moreover, we showed that locking refactored codes into synthetic organisms completely blocks invasion by mobile genetic elements, including viruses, which carry their own translation factors and successfully invade organisms with canonical and compressed genetic codes.

The near-universal genetic code defines the correspondence between codons in genes and amino acids in proteins (*1*, *2*). Because all forms of life use essentially the same genetic code, evolutionary innovation can be shared through horizontal gene transfer (HGT) between organisms (*3*, *4*), and this is a major driver of evolution (*5*).

However, the near-universal genetic code is also a liability for organisms; mobile genetic elements (or selfish genetic elements)—including transposons, viruses, and plasmids—exploit the universality of the code and co-opt the host cell's machinery to read their genes and propagate at the expense of host organisms. There is a clear tension between maintaining a common genetic code, to allow the acquisition of beneficial innovation through HGT, and excluding selfish genetic elements that exploit the common code for their own ends (*3*, *6*).

Several deviations from the standard genetic code have been documented in mitochondria and chloroplasts, and most characterized code reassignments involve stop codons (*7*–*9*). Known sense codon reassignments in the nuclear genome are rare. The "CTG yeast" decodes the CUG codon (which encodes leucine in the standard code) primarily as serine (97%, with the remaining 3% still assigned to leucine) (*10*). Viruses for the CTG yeasts are essentially unknown, which suggests that sense codon reassignment may protect against viruses (*11*). There are no experimentally validated examples of sense codon reassignment in bacteria, although computational evidence supports arginine codon reassignment in bacilli (*12*).

Genome synthesis (*13*, *14*) and editing provide the opportunity to rewrite the genetic code and create organisms with new properties (*14*–*19*). We synthesized a 4-Mb *Escherichia coli* genome in which we replaced all annotated occurrences of the TCG and TCA serine codons with the synonymous AGC and AGT codons using defined recoding rules (*20*); we also replaced the TAG stop codon with TAA. This created Syn61, an organism with a compressed genetic code (*14*). We further evolved the strain and deleted the genes for the tRNAs that decode TCG and TCA codons (*serU*, $tRNA_{CGA}^{Ser}$; *serT*, $tRNA_{UGA}^{Ser}$) and the gene for RF-1 (*prfA*) that terminates protein synthesis at the TAG stop codon. The resulting organism, Syn61Δ3, cannot read all the codons in the near-universal genetic code and therefore cannot read horizontally transferred genes that contain the codons deleted from its genome, as exemplified by resistance to a range of bacteriophage (*18*).

It has been widely hypothesized that refactoring the structure of the genetic code, through the reassignment of sense codons to distinct canonical amino acids, would create organisms with new properties and could create a genetic firewall to limit the escape of genetic information from synthetic organisms to natural organisms (*4*, *6*, *21*–*23*). In this work, we tested these hypotheses.

Compressed codes are nonorthogonal

A spectinomycin resistance gene written in the canonical genetic code [SpecR wild type (WT)] was correctly read in, and conferred spectinomycin resistance to, cells that contain the full complement of tRNAs to read the canonical code. However, consistent with previous observations (*18*), SpecR WT did not confer spectinomycin resistance to Syn61Δ3 cells (Fig. 1).

We created a recoded spectinomycin resistance gene [recSpecR (ΔTCG, TCA)], with the compressed genetic code used in the Syn61 genome. recSpecR (ΔTCG, TCA) conferred spectinomycin resistance to Syn61Δ3 cells (Fig. 1). The recSpecR (ΔTCG, TCA) gene also conferred spectinomycin resistance to cells that read the canonical genetic code; this was expected because the compressed genetic code uses a subset of the codons used in the canonical genetic code. We made similar observations with hygromycin resistance genes written in canonical and compressed codes (fig. S1).

These experiments demonstrated that genetic information written in the canonical code can be read in cells that decode the canonical code, but not in cells with genome-wide code compression and cognate tRNA deletion. However, code-compressed genes can be read in both cells with cognate tRNA deletion and cells that decode the canonical code. Therefore, there is no barrier limiting the flow of genetic information from engineered organisms with compressed genetic codes to natural forms of life. Creating orthogonal genetic codes that actively restrict the transfer of genetic information from engineered biological systems to natural systems is an important and unaddressed challenge.

tRNAs enable invasion of codon-compressed organisms

A WT F plasmid [F (WT)], written in the canonical genetic code, was efficiently transferred to cells that read the canonical code. By contrast, F (WT) was not transferred to Syn61Δ3 (Fig. 1 and data file S1), as expected. However, upon selecting for the conjugation of F (WT) from cells that read the canonical code into Syn61Δ2 cells (Syn61 cells deleted for *serU* and *serT* but containing *prfA*), we obtained two viable colonies in which recipient cells had received F (WT) (fig. S2). These colonies corresponded to rare events, which appeared at a frequency 10^6-fold less than the colonies resulting from conjugation of F (WT) into cells that read the canonical code. Sequencing the two clones revealed that they had acquired sequences that contained *serT* from the donor cell. This provided direct experimental evidence that selection for transfer of a mobile genetic element that uses the canonical code to recipients that cannot read the entire canonical code can enable selection for recipients that acquire the tRNA genes necessary to read the canonical genetic code.

To follow the effects of introducing *serT* into recipient cells in a reproducible system, we created the mobile genetic element F (WT + *serT*), a variant of F (WT) that contains *serT*. We demonstrated that F (WT + *serT*)

[1]Medical Research Council Laboratory of Molecular Biology, Cambridge, UK. [2]Department of Biochemistry, University of Cambridge, Cambridge, UK.
*Corresponding author. Email: chin@mrc-lmb.cam.ac.uk

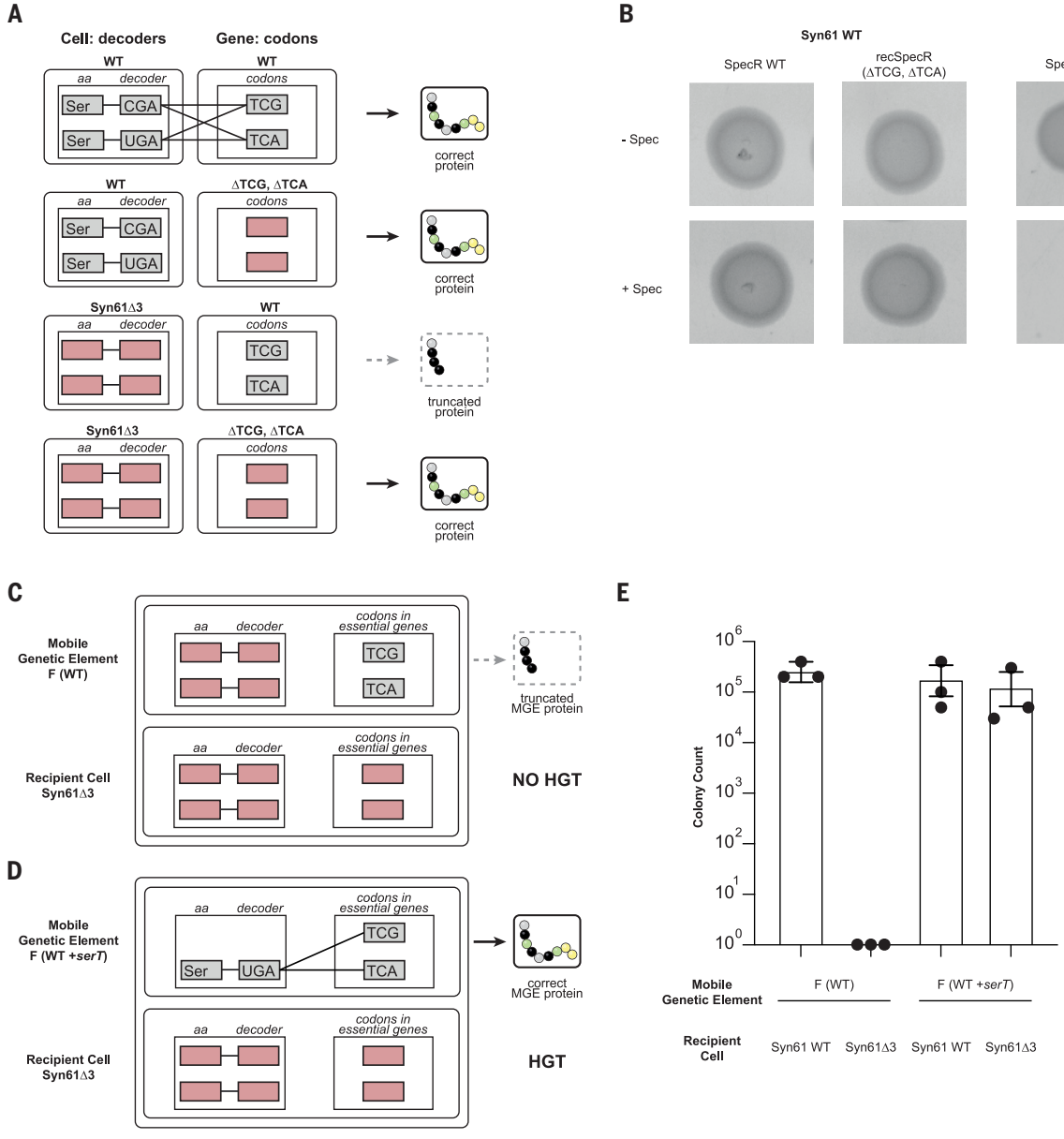

Fig. 1. Compressed genetic codes are nonorthogonal. (**A**) The relationship between the TCG and TCA codons in genes, the decoders for these codons in cells with canonical (WT) decoding and Syn61Δ3 decoding (ΔTCG, ΔTCA), and the corresponding protein sequence synthesized. The anticodon of the tRNAs that read TCG or TCA codons is indicated (decoder). The amino acid (aa) used by the tRNA is indicated. Gray for a decoder indicates that the tRNA is loaded with serine. Gray for a codon indicates that the codon is within a non–codon-compressed gene, and its decoding as serine will make the correct protein sequence. Pink for a decoder–amino acid pair indicates that the tRNA is deleted. Pink for a codon indicates the codon is absent from the gene because the gene has been designed with codon compression. (**B**) Functional assessment of SpecR WT and codon-compressed spectinomycin resistance [recSpecR (ΔTCG, TCA)] genes in (left) cells that use the full complement of tRNAs to decode all the codons in the reading frame (Syn61 WT) and (right) cells in which the tRNAs that decode TCG and TCA codons have been deleted (Syn61Δ3). Cells were spotted on agar plates in the presence or absence of spectinomycin and incubated overnight. The growth of cells in the presence of spectinomycin indicates that the indicated SpecR gene is functional in the indicated strain. (**C** and **D**) Predicted protein synthesis and HGT outcomes from mobile genetic elements (MGEs) and recipient cells with the indicated decoders and codons in essential genes. (C) A mobile genetic element encoding its genes according to the canonical genetic code, in which TCG and TCA encode serine, cannot be horizontally transferred to Syn61Δ3 cells that have no decoders for TCG and TCA codons. Translation will stall at TCG and TCA codons, and no full-length protein will be synthesized from the essential genes within the mobile genetic element that contain TCG and TCA codons. (D) A mobile genetic element encoding its genes according to the canonical genetic code that also carries a gene for a tRNA decoding TCG and TCA codons can be horizontally transferred to Syn61Δ3 cells. The tRNA encoded on the mobile genetic element can rescue decoding of TCG and TCA codons within essential genes in the mobile genetic element to make the correct MGE proteins. (**E**) Transfer of mobile genetic elements through conjugation. Colony count indicates successful transconjugants received from ~10^6 cells. A WT mobile genetic element (F WT) can be transferred into cells that read the canonical code (Syn61 WT) but not into cells that lack tRNAs decoding TCG and TCA (Syn61Δ3). A WT mobile genetic element that encodes a tRNA decoding TCG and TCA codons as serine [F (WT + *serT*)] can be transferred to both Syn61 WT and Syn61Δ3.

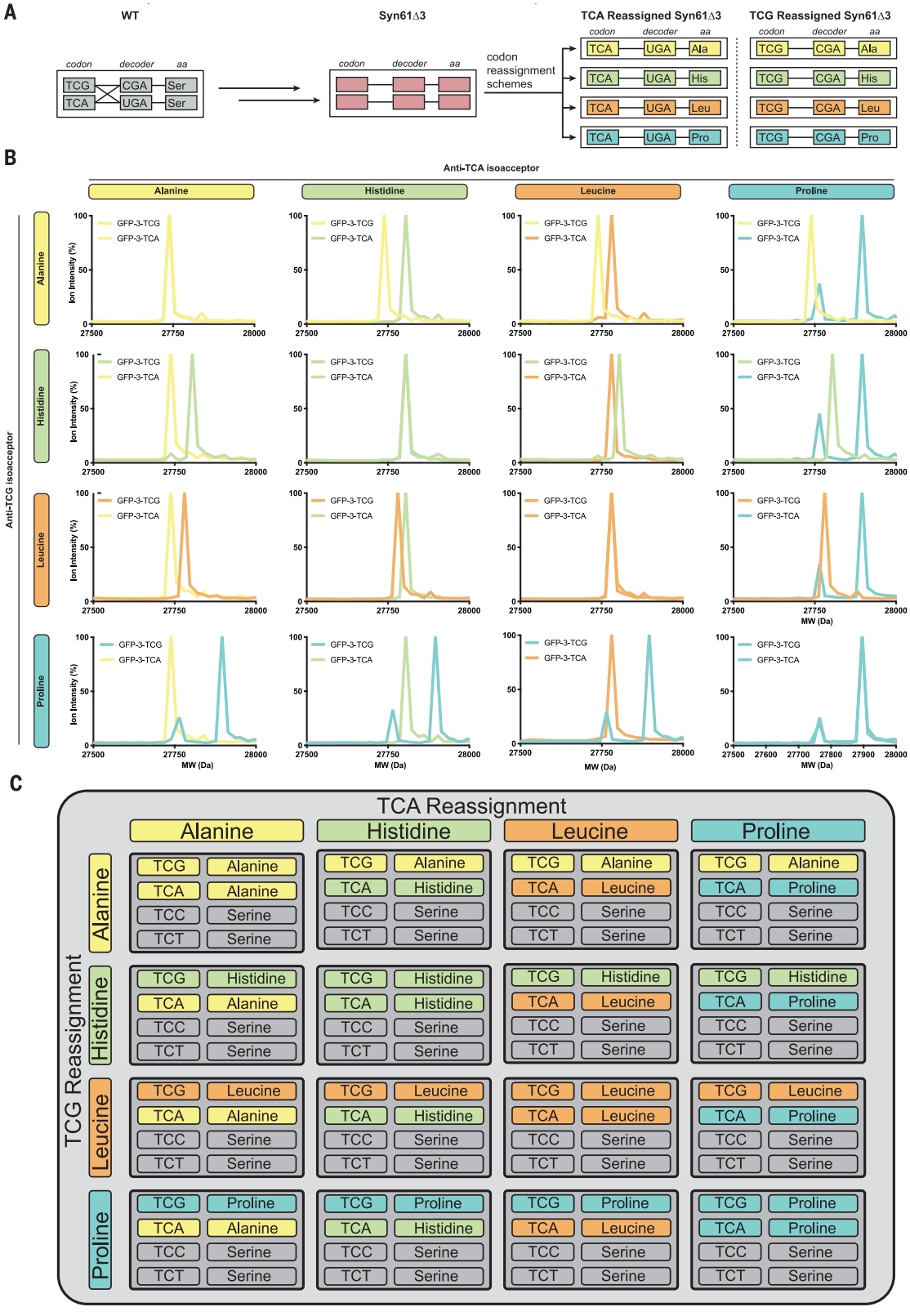

Fig. 2. Sense codon reassignment generates genetic codes. (**A**) Total synthesis of a codon-compressed genome followed by tRNA and release factor deletion yielded Syn61Δ3. The discovery of tRNAs that direct the incorporation of distinct canonical amino acids in response to TCG or TCA codons enables sense codon reassignment to create new genetic codes. (**B**) Isoacceptor tRNAs for the indicated amino acids with anticodons altered to the Watson-Crick complement of TCG or TCA codons were introduced into Syn61Δ3 in the indicated pairwise combinations. We read out the identity of the amino acid incorporated into each codon using GFP genes with TCG or TCA codons at position 3 and electrospray ionization mass spectrometry. When pairs of isoacceptors for distinct amino acids were used, each codon led to the specific incorporation of the amino acid attached to the Watson-Crick–paired isoacceptor. The secondary peak measured in the proline incorporations results from incomplete methionine cleave at the N terminus. A complete list of found and expected masses are provided in data file S2. MW, molecular weight. (**C**) Sixteen genetic codes in which TCG and TCA codons are reassigned to alanine, histidine, leucine, and proline.

can be transferred to Syn61Δ3 cells and that this transfer is dependent on *serT* (Fig. 1). We conclude that acquisition of *serT* is sufficient to circumvent the genetic isolation that is provided by code compression and cognate tRNA deletion in Syn61Δ3. These experiments highlight that creating systems that actively obstruct invasion by mobile genetic elements that carry their own decoders is an important challenge.

Refactoring code structure

We found that chimeric tRNAs for alanine ($tRNA_{CGA}^{Ala}$, $tRNA_{UGA}^{Ala}$), histidine ($tRNA_{CGA}^{His}$, $tRNA_{UGA}^{His}$), leucine ($tRNA_{CGA}^{Leu}$, $tRNA_{UGA}^{Leu}$), and proline ($tRNA_{CGA}^{Pro}$, $tRNA_{UGA}^{Pro}$) specifically

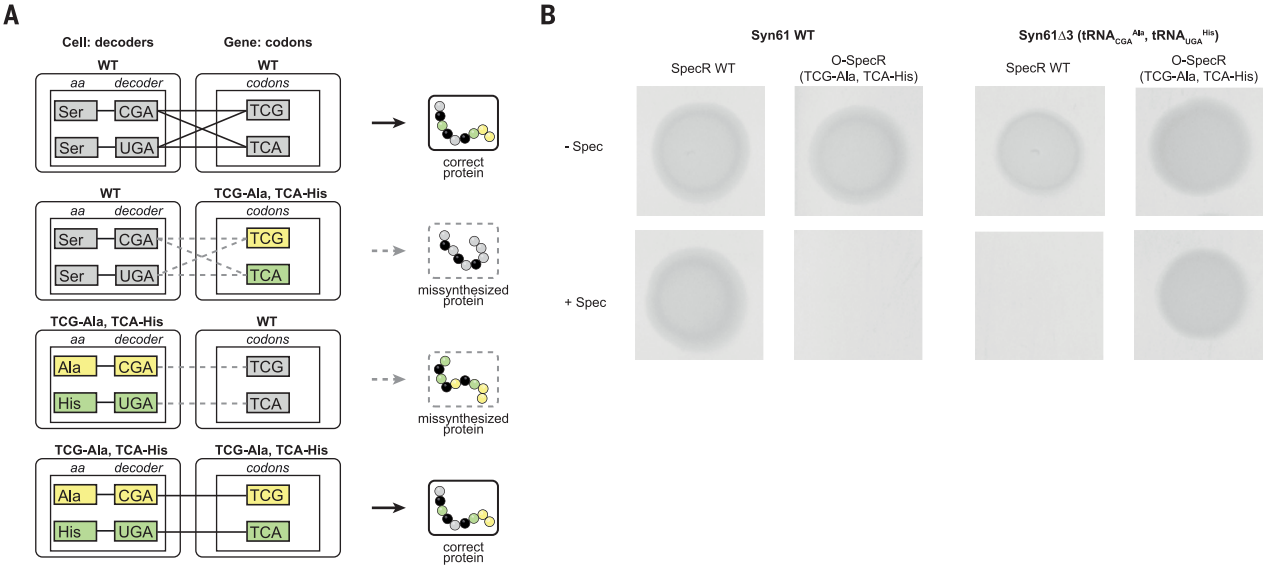

Fig. 3. Orthogonal genetic systems. (**A**) Relationship between the TCG and TCA codons in genes; the decoders for these codons in cells with canonical (WT) decoding and decoding by $tRNA_{CGA}^{Ala}$, $tRNA_{UGA}^{His}$ in Syn61Δ3; and the corresponding protein sequence synthesized. The anticodon of the tRNAs that read TCG or TCA codons is indicated (decoder). The amino acid (aa) used by the tRNA is indicated. Gray for a codon indicates that its decoding as serine will make the correct protein sequence. Yellow for a codon indicates that its decoding as alanine will make the correct protein sequence. Green for a codon indicates that its decoding as histidine will make the correct protein sequence. (**B**) Functional assessment of SpecR WT (written in the canonical genetic code) and O-SpecR (TCG-Ala, TCA-His), which is codon-compressed according to the Syn61 recoding scheme and has alanine codons replaced with TCG and histidine codons replaced with TCA. The genes are read in cells that read the canonical code (Syn61 WT), and cells where TCG is decoded as alanine and TCA is decoded as histidine [Syn61Δ3 ($tRNA_{CGA}^{Ala}$, $tRNA_{UGA}^{His}$)]. Cells were spotted on agar plates in the presence or absence of spectinomycin and incubated overnight. The growth of cells in the presence of spectinomycin indicates that the indicated SpecR gene is functional in the indicated strain.

direct the incorporation of the amino acid defined by the parent isoacceptor tRNA in response to their cognate codon (TGC or TCA) at position 3 in superfolder green fluorescent protein (sfGFP) or position 11 in ubiquitin in Syn61Δ3 (Fig. 2, figs. S3 to S13, and data files S2 and S3) and produce good yields of protein (fig. S3). The fidelity of $tRNA_{UGA}^{Leu}$ was lower than that of other tRNAs (data file S3 and fig. S5). We investigated alanyl and leucyl-tRNAs because their anticodons are not identity elements for their cognate aminoacyl-tRNA synthetases, and they were therefore expected to be permissive to anticodon mutation; the other tRNAs were identified through a screen (fig. S14). We found that unlike $tRNA_{CGA}^{Ser}$ and $tRNA_{UGA}^{Ser}$, our chimeric tRNAs specifically decode the Watson-Crick complement of their anticodon sequence; for example, $tRNA_{CGA}^{Ala}$ decodes TCG codons in preference to TCA codons, and $tRNA_{UGA}^{Ala}$ decodes TCA codons in preference to TCG codons (figs. S3 and S5). These tRNAs are also specific with respect to other TCN codons (figs. S6 and S7), and most reassigned strains grew comparably with parental strains (fig. S15). Our data show that we can independently reassign the TCA and TCG codons to alanine, histidine, leucine, or proline in Syn61Δ3 and thereby create 16 new genetic codes (Fig. 2B, figs. S3 and S5, and data files S2 and S3). In each of these genetic codes, we changed the identity of the canonical amino acids encoded at specific sense codons with respect to both the canonical code and the other 15 codes we have created (Fig. 2C). Our reassignment strategy is analogous to models that have been proposed for codon capture in natural evolution (*24*).

Overall, we have refactored the structure of the genetic code. Our genetic codes expand the number of codons used to encode alanine and proline (from four to six), double the number of codons used to encode histidine from two to four, and increase the number of codons used to primarily encode leucine from six to eight. These experiments also show that the UCN codon box, which encodes serine in the canonical code, can be split to encode additional canonical amino acids.

Orthogonal code–orthogonal decoder pairs

Genes that are written by using the canonical genetic code, in which TCG and TCA codons encode serine, will make the correct protein product in natural cells that read these codons as serine. However, these genes will yield the incorrect, likely nonfunctional protein product in cells that decode these codons to incorporate amino acids other than serine.

Similarly, synthetic genes—in which we compress the genetic code using the Syn61 recoding scheme and replace codons for specific natural amino acids with TCG and TCA codons—will make the correct protein product in cells that decode the TCG and TCA codons to incorporate the correct amino acid. However, these synthetic genes will yield an incorrect, likely nonfunctional protein product in cells that read the canonical genetic code (Fig. 3).

We converted all 27 GCN codons (which encode alanine in the canonical code) to TCG codons and all six CAT/C codons (which encode histidine in the canonical code) to TCA codons in recSpecR (ΔTCG, TCA). This created the orthogonal resistance gene O-SpecR (TCG-Ala, TCA-His). We demonstrated that O-SpecR (TCG-Ala, TCA-His) can be decoded in, and confer spectinomycin resistance to, Syn61Δ3 ($tRNA_{CGA}^{Ala}$, $tRNA_{UGA}^{His}$) cells, in which TCG is read as alanine and TCA is read as histidine. We further demonstrated that O-SpecR (TCG-Ala, TCA-His) did not confer spectinomycin resistance to cells that read the canonical genetic code. Last, we demonstrated that SpecR WT, in which serine is encoded by using TCG and TCA codons, cannot confer spectinomycin resistance to Syn61Δ3 ($tRNA_{CGA}^{Ala}$, $tRNA_{UGA}^{His}$) cells (Fig. 3). We extended this approach to six other reassignment schemes, as well as to other genes (fig. S16).

These experiments demonstrated that we can create a genetic code–decoder pair for synthetic genes that is functionally orthogonal with respect to the canonical genetic code–decoder pair for natural genes. The orthogonal

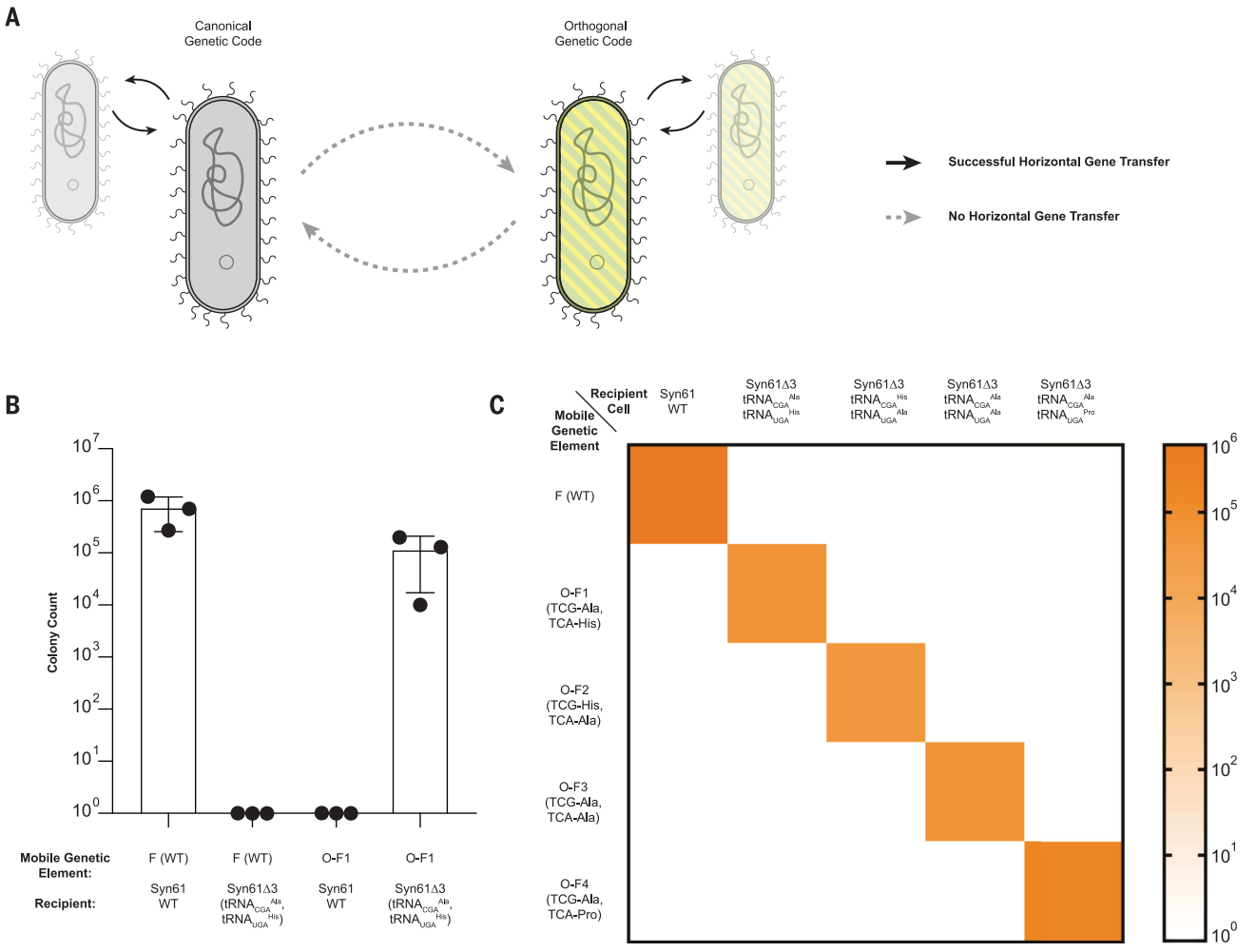

Fig. 4. Orthogonal and mutually orthogonal HGT systems. (**A**) HGT between organisms that use distinct, orthogonal genetic codes is prohibited (dashed gray arrows), whereas HGT can occur between cells that share a common genetic code (solid arrows). (**B**) Orthogonal horizontal transfer of mobile genetic elements. Colony count indicates the number of transconjugants received from ~10^6 donor cells that bear the indicated mobile genetic element. A WT mobile genetic element [F (WT)] was transferred into cells that read the canonical genetic code (Syn61 WT) but not into cells where TCG is reassigned to alanine and TCA is reassigned to histidine [Syn61Δ3 ($tRNA_{CGA}^{Ala}$, $tRNA_{UGA}^{His}$)]. An orthogonal mobile genetic element (O-F1), in which alanine codons were replaced with TCG and histidine codons were replaced with TCA, was transferred into Syn61Δ3 ($tRNA_{CGA}^{Ala}$, $tRNA_{UGA}^{His}$) but not into Syn61 WT. (**C**) Mutually orthogonal HGT systems. Colony count (orange heat map) indicates successful transconjugants received from ~10^6 donor cells bearing the indicated mobile genetic element.

code, written in synthetic genes, is correctly read by the cognate orthogonal decoder but not by the canonical decoder. The canonical code, written in natural genes, is correctly read by the canonical decoder but not by the orthogonal decoder.

The functional orthogonality of genes in cells with altered decoders will depend on the frequency of reassigned codons and the functional consequences of codon reassignments. The consequences of amino acid substitutions—a result of codon reassignment—may globally and crudely correlate with differences in amino acid polarity and hydrophobicity (*6*, *25*). The consequences of amino acid substitutions at particular sites in proteins may be predicted by using computational approaches that leverage evolutionary sequence- and/or structural information (*26*–*29*). Although the composition of natural genes is fixed, the codon usage in synthetic genes—written in the standard code or any orthogonal code—can be simply designed to maximize the number of codons that are subject to reassignment, which may maximize the functional orthogonality of synthetic genes.

Orthogonal HGT

Next, we created orthogonal HGT (O-HGT) systems composed of an orthogonal decoder and a mobile genetic element that uses an orthogonal genetic code. Cells that read the canonical code can transfer a WT mobile genetic element between themselves but cannot transfer the WT mobile genetic element to cells that contain orthogonal decoders. Cells that contain O-HGT systems can transfer their mobile genetic element to cells that contain a compatible orthogonal decoder but cannot transfer their mobile genetic element to cells that contain an incompatible orthogonal decoder or to cells that read the canonical code.

A mobile genetic element [F (WT)] written in the canonical code was transferred to cells that read the canonical code, as expected. We also showed that F (WT) could not be transferred to Syn61Δ3 ($tRNA_{CGA}^{Ala}$, $tRNA_{UGA}^{His}$)

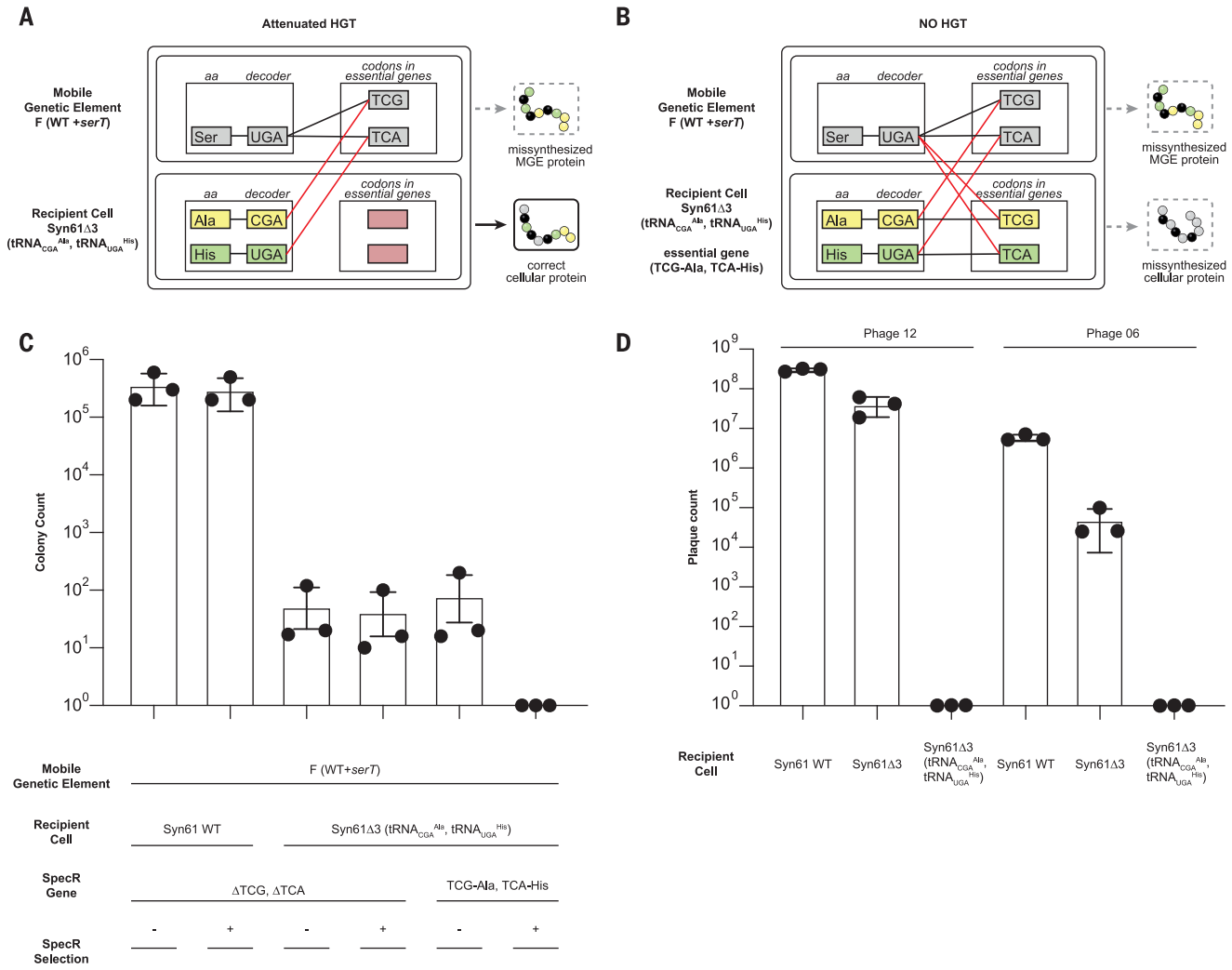

Fig. 5. Orthogonal code locking blocks invading codes. (**A** and **B**) Predicted protein synthesis and HGT outcomes from mobile genetic elements (MGEs) and recipient cells with the indicated decoders, and codons in essential genes. (A) Transfer of a WT mobile genetic element that encodes for a tRNA decoding TCG and TCA codons as serine into a cell where TCG is reassigned to alanine and TCA is reassigned to histidine. Essential genes in the WT mobile genetic element, which contain TCG and TCA codons, will be missynthesized, with each TCG and TCA codon in the gene being stochastically decoded as serine or alanine/histidine. This is predicted to attenuate HGT. (B) Transfer of a WT mobile genetic element that encodes for a tRNA decoding TCG and TCA codons as serine into a cell where TCG is reassigned to alanine and TCA is reassigned to histidine. Essential genes in the WT mobile genetic element, which contain TCG and TCA codons, will be missynthesized, with each TCG and TCA codon in the gene being stochastically decoded as serine or alanine/histidine. In addition, essential genes in the host cell, in which TCG is used to encode alanine and TCA is used to encode histidine, will be missynthesized. This is predicted to further restrict HGT. (**C**) HGT of a WT mobile genetic element [F (WT + *serT*)] is ablated in cells that use a refactored genetic code in essential genes. Colony count indicates successful transconjugants received from ~10^6 cells. Recipient cells and spectinomycin resistance gene variant (SpecR gene) in the recipient cell are indicated. Correctly reading the indicated SpecR gene in the recipient cell is made essential by addition of spectinomycin. (**D**) T4-like phage encoding a seryl-tRNA$_{UGA}$ infect Syn61Δ3 but not cells that bear orthogonal genetic codes. Plaque count indicates the number of successfully replicating phage obtained from infection with 1.1×10^{10} plaque-forming units (PFU)/ml (phage 12) and 7.5×10^9 PFU/ml (phage 6). Cells contain cognate spectinomycin resistance genes, as in (C); all experiments were performed in the presence of spectinomycin.

cells, in which TCG codons are read as alanine and TCA codons are read as histidine (Fig. 4).

Next, we investigated HGT for mobile genetic elements written in altered genetic codes. We synthesized the mobile genetic element O-F1 (TCG-Ala, TCA-His). The genetic code in all annotated open reading frames of this F plasmid was compressed by using the Syn61 scheme, and GCN codons (which encode alanine in the canonical code) and CAT/C codons (which encode histidine in the canonical code) were converted to TCG and TCA codons, respectively, within the *trfA* gene; this gene is essential for the replication of the mobile genetic element (*30*).

O-F1 (TCG-Ala, TCA-His) was horizontally transferred to Syn61Δ3 (tRNA$_{CGA}^{Ala}$, tRNA$_{UGA}^{His}$) cells. We further demonstrated that O-F1 (TCG-Ala, TCA-His) was not horizontally transferred to cells that read the canonical genetic code (Fig. 4). These experiments demonstrated that we can create O-HGT systems. We created additional HGT systems that are orthogonal to the natural genetic system and mutually orthogonal to each other (Fig. 4 and fig. S17). Overall, we demonstrated the scalability of our approach through the creation of five mutually orthogonal HGT systems.

Blocking invading codes

We hypothesized that reassigning TCA and TCG codons to specific natural amino acids and replacing codons for specific natural amino acids in essential genes with TCA and TCG codons in Syn61Δ3 (Fig. 5) would obstruct the *serT*-mediated HGT we observed in this strain (Fig. 1).

Transfer of F (WT + *serT*) to Syn61Δ3 ($tRNA_{CGA}^{Ala}$, $tRNA_{UGA}^{His}$, O-SpecR [TCG-Ala, TCA-His]) was obstructed (10^4-fold) in the absence of spectinomycin because $tRNA_{CGA}^{Ala}$ and $tRNA_{UGA}^{His}$ compete with $tRNA_{UGA}^{Ser}$ in the recipient cell to decrease the production of functional proteins from the mobile genetic element. However, this obstruction was not sufficient to completely ablate transfer of F (WT + *serT*). Upon the addition of spectinomycin, which makes O-SpecR (TCG-Ala, TCA-His) an essential gene in the cell, the decoding of TCG codons as alanine and the decoding of TCA codons as histidine become essential and "locked in." Under these conditions, the transfer of F (WT + *serT*) was completely ablated (Fig. 5). Similar results were obtained with other refactored codes and other essential genes (figs. S18 and S19).

To extend our approach to viral infection, we identified pools of phage from the River Cam that can infect Syn61Δ3 (data file S4 and supplementary materials, materials and methods). From these pools, we isolated two individual phage (12 and 06, both T4-like phage), which carry an identical $tRNA_{UGA}^{Ser}$ gene and infect Syn61Δ3 (fig. S20); some viruses are known to carry their own tRNAs and other translation factors to augment the cellular pool of translation factors and assist in the translation of codons within their own genes (*9*, *31*, *32*). As expected, expression of this tRNA in Syn61Δ3 is sufficient to confer susceptibility to infection by (otherwise noninfectious) T4 phage (fig. S21). We demonstrated that unlike Syn61Δ3, several refactored, code-locked strains were completely resistant to infection with phage 06 and phage 12 (Fig. 5 and fig. S22).

Our results demonstrate that cells with refactored genetic codes resist invasion by mobile genetic elements that use competing codes. Locking the refactored codes into essential genes in the cell enhances this resistance and can ablate transfer of mobile genetic elements with competing codes.

Discussion

Previous work has shown that the choice of synonymous codons in individual genes and viruses can alter their robustness and evolvability (*33*–*35*), but such approaches are limited to exploring subsets of the canonical code. Although a large body of theoretical work and a limited number of in vitro experiments have considered the relationship between the structure of the genetic code and its robustness and evolvability (*21*, *22*), it has been impossible to investigate the resulting hypotheses through experiments in living cells. Refactoring the structure of the genetic code directly alters the number and types of amino acids that can be accessed by point mutations in live cells (figs. S23 and S24) and provides opportunities to experimentally test how altered codes affect the robustness and evolvability of protein and cellular function. In future work, we will aim to leverage genetic code refactoring to accelerate directed evolution.

Competition between pools of genotypes written in different genetic codes may have led to a universal code being fixed in extant life. Future work may exploit organisms with refactored and mutually orthogonal codes to experimentally investigate competition between codes and the role of HGT in fixing and maintaining a universal genetic code (*4*, *36*).

Shielding synthetic organisms from environmental genetic elements may be valuable for biotechnological applications on an industrial scale, for which contamination with mobile genetic elements, including viruses, can cause financial losses and disrupt vital supply chains (*37*). Synthetic organisms with decoders for an orthogonal code and essential genes that lock the cognate orthogonal code into the organism exhibit resistance to mobile genetic elements that are written in the canonical code. This resistance extends to mobile genetic elements that carry tRNA genes that otherwise allow correct reading of the canonical code in their genes. Our work defines a paradigm for creating organisms that actively resist invasion by foreign codes.

Our refactored genetic codes limit the transfer of genetic information from synthetic organisms to natural organisms and may form the basis of genetic firewalls that isolate synthetic genetic systems from the environment. Such genetic firewalls complement strategies for controlling the survival and growth of synthetic organisms, to realize biocontainment (*16*, *38*, *39*). This is especially important when considering applications of engineered organisms outside the laboratory.

The strategies that we have described should be generally applicable to any gene or genetic system added to the synthetic organism. Because the canonical genetic code is near universally conserved, we anticipate that the principles we have established may be applied to a broad range of organisms.

REFERENCES AND NOTES

1. F. H. Crick, L. Barnett, S. Brenner, R. J. Watts-Tobin, *Nature* **192**, 1227–1232 (1961).
2. M. W. Nirenberg, J. H. Matthaei, *Proc. Natl. Acad. Sci. U.S.A.* **47**, 1588–1602 (1961).
3. R. J. Hall, F. J. Whelan, J. O. McInerney, Y. Ou, M. R. Domingo-Sananes, *Front. Microbiol.* **11**, 1569 (2020).
4. K. Vetsigian, C. Woese, N. Goldenfeld, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 10696–10701 (2006).
5. S. M. Soucy, J. Huang, J. P. Gogarten, *Nat. Rev. Genet.* **16**, 472–482 (2015).
6. E. V. Koonin, A. S. Novozhilov, *IUBMB Life* **61**, 99–111 (2009).
7. M. Kollmar, S. Mühlhausen, *BioEssays* **39**, 1600221 (2017).
8. J. Ling *et al*., *Nucleic Acids Res.* **42**, 499–508 (2014).
9. A. L. Borges *et al*., *Nat. Microbiol.* **7**, 918–927 (2022).
10. M. A. Santos, A. C. Gomes, M. C. Santos, L. C. Carreto, G. R. Moura, *C. R. Biol.* **334**, 607–611 (2011).
11. D. J. Taylor, M. J. Ballinger, S. M. Bowman, J. A. Bruenn, *PeerJ* **1**, e50 (2013).
12. Y. Shulgina, S. R. Eddy, *eLife* **10**, e71402 (2021).
13. D. G. Gibson *et al*., *Science* **319**, 1215–1220 (2008).
14. J. Fredens *et al*., *Nature* **569**, 514–518 (2019).
15. M. J. Lajoie *et al*., *Science* **342**, 357–360 (2013).
16. D. J. Mandell *et al*., *Nature* **527**, 264–264 (2015).
17. A. J. Rovner *et al*., *Nature* **527**, 264 (2015).
18. W. E. Robertson *et al*., *Science* **372**, 1057–1062 (2021).
19. D. de la Torre, J. W. Chin, *Nat. Rev. Genet.* **22**, 169–184 (2021).
20. K. Wang *et al*., *Nature* **539**, 59–64 (2016).
21. G. Pines, J. D. Winkler, A. Pines, R. T. Gill, *mBio* **8**, e01654-17 (2017).
22. J. Calles, I. Justice, D. Brinkley, A. Garcia, D. Endy, *Nucleic Acids Res.* **47**, 10439–10451 (2019).
23. P. Marliere, *Syst. Synth. Biol.* **3**, 77–84 (2009).
24. S. Osawa, T. H. Jukes, *J. Mol. Evol.* **28**, 271–278 (1989).
25. M. Schmidt, V. Kubyshkin, *ChemBioChem* **22**, 1268–1284 (2021).
26. J. Frazer *et al*., *Nature* **599**, 91–95 (2021).
27. S. Teng, A. K. Srivastava, C. E. Schwartz, E. Alexov, L. Wang, *Int. J. Comput. Biol. Drug Des.* **3**, 334–349 (2010).
28. V. Parthiban, M. M. Gromiha, D. Schomburg, *Nucleic Acids Res.* **34**, W239–W242 (2006).
29. P. C. Ng, S. Henikoff, *Annu. Rev. Genomics Hum. Genet.* **7**, 61–80 (2006).
30. I. Konieczny, K. S. Doran, D. R. Helinski, A. Blasina, *J. Biol. Chem.* **272**, 20173–20178 (1997).
31. V. A. Delesalle, N. T. Tanke, A. C. Vill, G. P. Krukonis, *Bacteriophage* **6**, e1219441 (2016).
32. M. Bailly-Bechet, M. Vergassola, E. Rocha, *Genome Res.* **17**, 1486–1495 (2007).
33. B. A. Renda, M. J. Hammerling, J. E. Barrick, *Mol. Biosyst.* **10**, 1668–1678 (2014).
34. G. Moratorio *et al*., *Nat. Microbiol.* **2**, 17088 (2017).
35. J. R. Coleman *et al*., *Science* **320**, 1784–1787 (2008).
36. V. Kubyshkin, C. G. Acevedo-Rocha, N. Budisa, *Biosystems* **164**, 16–25 (2018).
37. P. W. Barone *et al*., *Nat. Biotechnol.* **38**, 563–572 (2020).
38. P. Marlière *et al*., *Angew. Chem. Int. Ed.* **50**, 7109–7114 (2011).
39. J. W. Lee, C. T. Y. Chan, S. Slomovic, J. J. Collins, *Nat. Chem. Biol.* **14**, 530–537 (2018).

ACKNOWLEDGMENTS

We thank J. Fredens (National University of Singapore) for advice on conjugation assays; S. Grazioli and A. Kleefeldt for help with de novo assembly of phage genomes; and the LMB mass spectrometry facility, especially S.-Y. Peak-Chew, for performing mass spectrometry experiments. **Funding:** This work was supported by the Medical Research Council (MRC), UK (MC_U105181009, MC_UP_A024_1008 to J.W.C.); G.P. was supported by a Marie Skłodowska-Curie European Postdoctoral Fellowship (897663). Work in the G.P.C.S. laboratory was supported by award BB/W000105/1 from the Biotechnology and Biological Sciences Research Council, UK (UKRI). **Author contributions:** J.F.Z. and J.W.C. conceptualized and planned the project, with input from W.E.R.; J.F.Z. demonstrated unidirectional genetic isolation and loss of isolation through tRNA import. J.F.Z., T.S.E., G.P., and W.E.R. designed and characterized tRNAs for genetic code refactoring. J.F.Z. demonstrated bidirectional genetic isolation and (mutually) orthogonal HGT systems. J.F.Z. demonstrated improved genetic isolation through code refactoring and genetic code locking. J.F.Z. and T.K. performed phage experiments under supervision of G.P.C.S. J.W.C. and J.F.Z. wrote the manuscript, with input from all authors. J.W.C. supervised the project. **Competing interests:** The MRC has filed a provisional patent application related to this work on which J.F.Z., W.E.R., and J.W.C. are listed as inventors. J.W.C. has a commercial interest in Constructive Bio. **Data and materials**

availability: The GenBank accession nos. for all the plasmids described in the text are provided in data file S5, and the authors agree to provide any materials and strains used in this study upon request. All data are available in the main text or supplementary materials.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.add8943
Materials and Methods
Figs. S1 to S24
References (*40*–*44*)
Data Files S1 to S5
MDAR Reproducibility Checklist

Submitted 12 July 2022; accepted 7 October 2022
Published online 20 October 2022
10.1126/science.add8943