nature portfolio

Peer Review File

Engineering a genomically recoded organism with one stop codon

Corresponding Author: Professor Farren Isaacs

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Referee #1

(Remarks to the Author)
In this work, the authors describe their successful efforts to engineer a strain of E. coli that can efficiently and accurately synthesize a functional proteome using a translation system reliant on a single-stop codon (TAA). This is a significant advance as it shows that three-stop codons are not a requirement for viability and robust growth. This, in turn, allows the repurposing of the TAG and TGA stop codons for uses such as genetic code expansion. The implications for understanding the genetic code, its plasticity, and its potential for directed evolution are immense. In practical terms, the tools provided by this advance will considerably advance existing, and offer new, strategies for synthetic protein production.

The methods and results described are interesting and innovative and integrate whole genome recoding with directed modification of the translation machinery. The rationale for each step is also described in a way that will benefit other investigators.

The only aspect of the paper that sometimes falls short is the broader discussion beyond the utilitarian aspects of the work. For example, this work raises questions about whether three-stop codons reflect functional redundancy or a poorly understood aspect of translation robustness. If stop codons can be "compressed" from 3 to 1, how easily can sense codons be compressed and how many unique codons might be readily attainable?

Referee #2

(Remarks to the Author)
In their manuscript titled “Design and construction of a genomically recoded organism with a single non-degenerate stop codon” the authors report the construction of an E. coli strain (termed Ochre) in which both the amber (TAG) and the opal codon (TAA) have been recoded to ochre stop codons (TAA). This was achieved based on a previously reported strain, in which all amber mutants had been recoded. Starting from this strain, the authors recoded all instances of opal codons in the genome, and engineered release factor 2 and tRNATrp to preclude UGA recognition, thus fully removing the degeneracy of the three stop codons, as well as UGG. This enables repurposing of the two liberated codons (TAG, TGA) for the incorporation of unnatural amino acids (N-Boc-L-lysine, and p-Acetyl Phenylalanine) with high efficiencies (99%).

The authors motivate the interest of recoding strains to liberate redundant codons by the prospect of utilizing these codons for the incorporation of unnatural amino acids. In this respect, this work falls short of a previous report (Chin et al, Nature 2019), in which three codons have been liberated (TAG, TCG, and TCA). The authors highlight that their work represents the first case of a genetically recoded organism in which the degeneracy of codons is removed, but fall short to explain why that is of particular interest both fundamentally (methionine codon in the genetic code is not degenerate) or practically (not beneficial for unnatural amino acid incorporation and/or resistance to phages). In contrast to previous work on recoded organisms, liberation of the opal codon required additional changes to both host RF2 (to attenuate suppression of TGA) and tRNATrp (wobble pairs with TGA). While these measures strongly attenuated translational cross-talk, they did not completely abolish it.

In addition to reservations with respect to the novelty/utility of the work, there are some additional issues that need to be addressed:

1. The growth rates and optical densities reported in this work are very slow and low, respectively. For wild type Mg1655, a doubling time of 35 minutes at 37 °C is too slow, and a final OD600 of 0.8 is too low. The optical densities given in this work (Fig 4, Extended Data Fig. 3a, supplementary materials) appear to be relative optical densities obtained directly from the Biotek plate reader software without correction for path length (OD600 defined as 1 cm path length). This needs to be recalculated, or clearly indicated as relative ODs (and not OD600 or ODmax). The slow growth rates might be due to calculation errors caused by calculating growth rates based on these OD values.

2. Fig. 6: promoter in front of AraC should be PC not PBAD

3. Page 20 line 686: "…150 ml LB and 150 ml M9…": µl not ml

4. The authors should report yields of purified protein from this system (single and multiple suppression) to provide convincing evidence that this is useful technology. Previous work on recoded stains reported yields based on readouts of fluorescent protein- but the methodology did not provide useful yields of protein in practice.

Referee #3

(Remarks to the Author)
Review of Isaacs single stop codon paper

In this paper the Rinehart/Isaacs team describe the construction of an E. coli strain that is powered by a single stop codon TAA. Beginning with their previously described "no TAG" strain recognizing both TAA and TGA they first proceeded to remove as many TGA stop codons as possible using "MAGE and CAGE" methods described earlier. They then described methods to engineer the RF2 release factor which recognizes TAA and TGA to make it able to only terminate at TAA. This required the use of various methods (and literature mining) and ultimately gave rise to a TAA specific RF2. They also presented data showing that the tRNATrp was extremely effective at reading TGA codons, inserting tryptophan. This "misreading" would interfere with efficient insertion of noncanonical amino acids (ncAAs) and thus was unacceptable. In this case the solution was much easier and simply relied on a 1993 paper showing that changing one base adjacent to the anticodon of yeast tRNATrp sufficed to make it specific to TGG. By doing so, they constructed a strain that functions with a single stop codon. Performance of this strain with and without ncAAs was evaluated with reporter plasmids, and they showed successful in corporation of two ncAAs at TAG and TGA codons. Additional validation of strain performance was provided by phage sensitivity studies.

Overall this is a "First" and a well-designed and well -described study.

1) For practical deployment, it is important to know more about how sick this strain is compared to healthy E. coli strains. The current strain was compared to its immediate ancestor rather than a healthy E. coli strain. It would be good to do more extensive benchmarking against a few wild-type E. coli strains, not just MG1655, for example the Mach1 strain. Competitive growth would be a good addition. The experiments that were done indicate that the RF1 parental strain is very sick (despite efforts to improve it). How will these problems be solved in the long run? Has adaptive lab evolution been carried out on the parent prior to this engineering, or after?
2) Figure 2 d/e. I found these bar graphs exceedingly difficult to absorb/understand. There is way too much going on here. For starters, in c it shows cherry "X" YFP. That's good. But then we see in d/e separate panels for cherry and YFP. Are these reporters the same ones as in c… or a separate set? In the first d panel. Why is TAG nearly zero? Perhaps it would be best to dumb these down a bit and focu on the important point which is the % readthrough in ech configuration, and save some of the other data for the supplement. A final confusion is that I cant remember what the difference is between the first two strains in the key … can you add a scrap on info on what distinguishes them?
3) Figure 3. There two problems here. 1) There is no such number as zero in experiments of this type. The correct number to report is "< the minimum detectable number". 2) The statistics seem way off. The comparison of e.g. MG1655 to Ochre shows 5*10e8 vs. "Zero" plaques done in triplicate. It is very hard to believe that statistically this is a P<.05. I would think it was P <.many zeros one. My intuition is there is some major problem with calculating the stats here.

Other issues
4) Line 111 please insert ref to C321.ΔA
5) Line 135 "In the latter two categories of TGA overlap, we designed specialized oligos 135 to recode the TGA-containing gene to minimize polar effects". Cryptic!! Please provide more detail on how this done, what are these specialized oligos and exactly how are they designed?
6) Line 152 Please indicate the strategy used to identify the genes and pseudogenes that were deleted. Sounds as if this was all rather ad hoc, but some explanation of the considerations used would be important
7) Lines 166-175. Yikes!! Please specify how many mutations accumulated overall and include in a supplemental table. Why were auxotrophic mutations the only ones prioritized for fixing? What about missense mutations in essential genes or those required for fast growth?
8) Line 178 reword to e.g. "Engineering UAA-specific translation termination is required to establish a functionally recoded ΔTAG-178 ΔTGA GRO capable of efficiently inserting ncAAs at TAG and TGA codons".
9) Reword to e.g. "Our inability to delete RF2 in the presence of RF1, despite an absence of genomic TGA codons, supports the essentiality of alternative RF2 UGA termination-indepdent functions such as translational quality control35,36"
10) Line 203 "Previous …recognition" is not a sentence!
11) 204-211. The authors should confirm that the "PPF" RF2 fails to complement an RF2 deletion in their parent strain to "complete the genetic proof" here.

12) 212-220. Indicate somewhere that prfB encodes RF2, as most readers wont know this.
13) 221. The Chlorobium prfB factoid is interesting but raises the question – why does it have this sequence???
14) 703 change “a sense codon” to “a GCG codon”

Referee #4

(Remarks to the Author)
Isaacs et al. report the construction of an E. coli strain in which all natural TAG and TGA stop codons have been removed and Release Factor 2 (RF2) and tRNATrp further engineered such that the four codons TAG, TGA, TAA, and TGG are rendered nondegenerate and the TAG and TGA codons can be efficiently translated by unnatural suppressor tRNAs charged with non-standard amino acids (nsAAs).

This achievement is heroic and has both practical signficance for the incorporation of nsAAs and genome synthesis and engineering and basic science significance for our understanding of the genetic code.

I will consider the Figures in reverse order.

Figure 6 is important because it measures the utility of the genome engineering for the incorporation of nsAAs.

As I will come back to, the nomenclature used for the four different strains is non-intuitive. The figure would be much easier to digest, especially for the broad readership of Nature, if intuitive names for the strains were used and the strains composition was defined in the Figure Legend. For clarity, rEcΔ1.ΔA = their original ΔTAG strain; rEcΔ2.ΔA = their fully recoded ΔTAG, ΔTGA strain; Ochre = fully recoded ΔTAG, ΔTGA, mRF2-oKP strain; and Ochre.tW* = fully recoded ΔTAG, ΔTGA, mRF2-oKP, mutated tRNATrp strain. To emphasize, these strain names are non-intuitive; for example Ochre is both ΔTAG and ΔTGA, which is a non-standard meaning of Ochre.

The data in Figure 6b, c, and d are all reasonable. As the authors point out it is significant that the suppression of UAG and hence fluorescence goes up in both Ochre and Ochre.tW* where the TAG, TGA, TAA, and TGG are rendered nondegenerate. It is particularly significant in Figure 6d that they can suppress both UAG and UGA with two different amino acids 3 times each in a single construct. And that the normalized fluorescence is similar for 6d with 2 nsAAs and 6b and 6c with 1 nsAA each. That said they are using nsAAs and aminoacyl-tRNA synthetase/tRNA pairs that are known to be particularly robust in these kind of systems. The authors could do further experiments to demonstrate how well their engineered strain performs for incorporation of nsAAs. For example, what is the fluorescence of their reporter with no stop codons and why are the fluorescent values in Figures b, c, and d not shown relative to that? What is the fluorescence for single site suppression versus suppression of 3 codons? What is the suppression yield for proteins/positions that historically do not suppress well? In addition to relative fluorescence another good measure of the robustness of the system is protein yield. What is the relative yield of the GFP fusion protein with and without suppression? The authors should decide on the appropriate experiments to test the robustness of their system relative to amber suppression in wt E. coli.

Figure 5 measures the doubling time of the different strains. The authors should show the full log plots in the Extended Data.

Figure 4 clearly states where the tRNATrp mutants are and demonstrates that they reduce background suppression of UGA. The authors provide a concise and clear explanation of their design of their mutant tRNATrp in the manuscript. They are straightforward that there is still some readthrough of the UGA codon even in the tW* strain. They note the intriguing observation that there is tangible incorporation of Cys – although here they don’t offer an explanation. Again the Figure Legend could be easier to read. What is Ara inducing? What is TyrOTS and how is it turned on/off?

Figure 3 shows the phage infectivity data which is as expected but also interesting from the vantage point of genome recoding.

Figure 2 is very important for showing the impact of RF2 engineering on rendering the TAG, TGA, TAA, and TGG codons nondegenerate. The data is fine. However, it does not display the breadth of the RF2 engineering laid out in the text. The Figure and the text should match more. And related the text needs to tell a story. Now it reads as though the authors are telling us everything they did and in the order they did it. The paper would be much clearer to a broad audience if they picked a few key points to addressed and showed that data in Figure 2 as well. Finally, in the text describing the strain engineering the authors should use a standard set of names for the strains that are intuitive and evoke what changes were made to that strain. A small point I would not use purple in the Figure 2 because it has a different meeting than in Figure 6.

Figure 1 is a nice way to represent the authors genome engineering strategy, and they have used it successfully previously.

Version 1:

Reviewer comments:

Referee #1

(Remarks to the Author)
The authors have done a stellar job addressing reviewer concerns and questions, and the manuscript is now much improved.

Referee #3

(Remarks to the Author)
The authors have addressed the points raised well and I am generally very positive about this very interesting study.

In general readability has improved substantially, and some important control experiments have been added. The one change I'm less enthused about is the switch to the far more verbose (and thus accurate/systematic) strain names. The authors may want t o go back to using "ochre" as a nickname for the unmemorable "rEcΔ2.ΔA.B3.tW*".

In the prior review, point 10, rewording was suggested. The new sentence does not parse well at all. Suggest changing "Interrogating RF codon recognition in E.
coli, previous studies performed in vivo and in vitro characterizations of the S205P
mutation." to "Previous studies performed in vivo and in vitro characterizations of the S205P
mutation in order to evaluate RF codon recognition."

Referee #4

(Remarks to the Author)
The authors are very responsive to the reviewer's comments and the result is a manuscript which no question should be published in Nature. The authors emphasize the uniqueness and significance of their result that they can compress translation to a single stop codon and that they can eliminate translational crosstalk between four codons, UAA, UGA, UAG, and UGG. This result is of clear fundamental and practical interest across diverse fields. The manuscript and figures are extensively revised improving the clarity of the manuscript for a broad audience. The data are accurate with statistical significance clearly indicated. Enthusiasm for this manuscript is extremely high.

Below please find a point-by-point response (blue) to the reviewers' comments (*italicized*), with manuscript and figure references (red).

REVIEWERS' COMMENTS

Referee #1 (Remarks to the Author):

In this work, the authors describe their successful efforts to engineer a strain of E. coli that can efficiently and accurately synthesize a functional proteome using a translation system reliant on a single-stop codon (TAA). This is a significant advance as it shows that three-stop codons are not a requirement for viability and robust growth. This, in turn, allows the repurposing of the TAG and TGA stop codons for uses such as genetic code expansion. The implications for understanding the genetic code, its plasticity, and its potential for directed evolution are immense. In practical terms, the tools provided by this advance will considerably advance existing, and offer new, strategies for synthetic protein production.

1. We thank the reviewer for this positive assessment and for recognizing the significance of this study.

The methods and results described are interesting and innovative and integrate whole genome recoding with directed modification of the translation machinery. The rationale for each step is also described in a way that will benefit other investigators.

2. We thank the reviewer for this comment and recognizing the close integration of genome recoding and engineering translational compatibility to functionalize a recoded organism, which we hope provides a blueprint for future efforts by the community that seek to use or engineer organisms with recoded genomes.

The only aspect of the paper that sometimes falls short is the broader discussion beyond the utilitarian aspects of the work. For example, this work raises questions about whether three-stop codons reflect functional redundancy or a poorly understood aspect of translation robustness. If stop codons can be "compressed" from 3 to 1, how easily can sense codons be compressed and how many unique codons might be readily attainable?

3. We thank the reviewer for this perspective and agree that this study has a unique opportunity to enhance our understanding of functional codon redundancy and more fundamental aspects of translation. Our work demonstrates engineering of proteinaceous and nucleic acid translation factors that govern recognition of both stop (UAG, UGA, UAA) and sense (UGG) codons, thus serving as a blueprint for future successful stop and sense recoding. Inspired by the reviewer's comments, we have revised the manuscript abstract, introduction (e.g. page 3, lines 111-115), and discussion (e.g. pages 16-18, lines 565-599) to strengthen these points.

   Regarding compression of stop codons from 3 to 1, protein elongation from increased TGA stop codon suppression via $tRNA^{Trp}$ under stress conditions (Romero *et al. Nat. Comm.* 2024; 15, 4446) reveals additional cellular functions unique to TGA

besides termination. This suggests a possible evolutionary role for stop codon degeneracy in regulating stress responses and demonstrates that while the stop function can be compressed, additional functions specific to each stop codon could be impacted with compression. In other words, synonymous codons may not be wholly functionally synonymous as alternative or cryptic functions may yet be discovered. This insight can also be applied to sense codon recoding, illustrated by efforts in *E. coli* uncovering expression-regulating functions unique to AGG/AGA codons (Napolitano, PNAS 2016), and by work in yeast revealing that codon pairs can serve roles in regulating rates of protein translation at the mRNA level (Gamble *et al*. *Cell* 2016; 166, 679-690). We modified the discussion of our manuscript to reflect this recognition (page 18, lines 630-640).

Although this study recoded stop codons at the genomic level, engineering compatibility at the translational level required the engineering of both release factors and tRNAs (essential translation factors for sense recoding). Specifically, our engineering of $tRNA^{Trp}$ demonstrates how targeted nucleotide changes in tRNAs can alter chemical modifications by RNA-modifying enzymes and consequently codon recognition of tRNAs. This finding is significant because it enabled us to eliminate the wobble effect of $tRNA^{Trp}$ at the UGA codon, thus eliminating translational crosstalk of translation factors and isolating codons and tRNAs to exclusive roles. These engineering results are unprecedented and establish a blueprint for recoding both stop and sense codons.

The question of how many sense codons can be compressed and how many unique codons can be attained is excellent and also challenging to provide a definitive answer. While examples of single codon functions, both natural (Met-AUG, Trp-UGG) and engineered (compression of three stop codons to one in this work), suggest a theoretical compression of 64 unique codons (20 natural AA, 1 stop, 43 nsAAs), evolutionary pressures, mutational malleability, and translational crosstalk will likely add constraints on the upper limit of a ‘maximal genetic code’. Pursuit of a fully nondegenerate code is expected to reveal functional and chemical limitations in codon recognition by translation factors. Future efforts ought to shed light on the possible evolutionary paths available to organisms through these limitations. In addition, as codon recognition by tRNAs also involves competition with cognate and near-cognate suppressors, deletion or alteration of tRNAs in future recoding efforts may better enable identification and mapping of hidden near-cognate suppression across translation. Such efforts may illuminate cryptic codon and translation factor functions, as well as elucidate the intricate web of translational crosstalk uniquely informed by synthetic recoding efforts.

Referee #2 (Remarks to the Author):

In their manuscript titled “Design and construction of a genomically recoded organism with a single non-degenerate stop codon” the authors report the construction of an E. coli strain (termed Ochre) in which both the amber (TAG) and the opal codon (TAA) have been recoded to ochre stop codons (TAA). This was achieved based on a previously reported

strain, in which all amber mutants had been recoded. Starting from this strain, the authors recoded all instances of opal codons in the genome, and engineered release factor 2 and tRNATrp to preclude UGA recognition, thus fully removing the degeneracy of the three stop codons, as well as UGG. This enables repurposing of the two liberated codons (TAG, TGA) for the incorporation of unnatural amino acids (N-Boc-L-lysine, and p-Acetyl Phenylalanine) with high efficiencies (99%).
The authors motivate the interest of recoding strains to liberate redundant codons by the prospect of utilizing these codons for the incorporation of unnatural amino acids. In this respect, this work falls short of a previous report (Chin et al, Nature 2019), in which three codons have been liberated (TAG, TCG, and TCA).

1. We thank the reviewer for this summary and recognize the points raised regarding significance to previous work. Inspired by this reviewer's comments, we revised the manuscript (pages 3-4, lines 140-145) to present the foundational differences that distinguish our work from prior work by Fredens *et al.* (*Nature* 2019; 569, 514). First, it is important to note that this study and ours are fundamentally different: Fredens *Nature* 2019 is a partial sense codon liberation while our study describes complete liberation of two stop codons to fully compress stop codon function to a single nondegenerate codon. In this context, it is first important to define a "liberated codon" as having three key properties: (1) a codon that has been reassigned to a synonymous codon throughout the genome, (2) deletion or engineering translation machinery that does not decode the reassigned codon, eliminating its function at the translation level, and (3) a codon that does not exhibit translational cross-talk with other codons or translation factors. Fredens *Nature* 2019 achieved (1) and only partly (2), in which not all Ser-tRNAs responsible for both TCG and TCA were removed. This process was completed in a subsequent study (Robertson *et al. Science* 2021; 372, 1057-1062). Notably, neither of these papers achieved (3) in that the tRNA that decodes TCA can wobble and decode TCG, resulting in translational crosstalk between TCG and TCA. Thus, this prior study did not liberate three codons and did not sufficiently mitigate translational crosstalk to precisely incorporate each nsAA at only their respective codons. Our current manuscript achieves (#1) and (#2), while significantly advancing (#3) to liberate both UAG and UGA.

   The key aspects of our study distinguishing us from Fredens *Nature* 2019 are: (i) the nature of the codons liberated (full compression of stop codon degeneracy), (ii) the methods of liberation (translation factor deletion vs rational engineering of essential factors to alter recognition and disentangle crosstalk), and (iii) the efficiency and purity with which multiple nonstandard amino acids are incorporated within single proteins with utility for making proteins with multiple instances of two distinct nsAAs. As previous recoding efforts simply relied on codon replacement and deletion of cognate translation factors to liberate codons (e.g., RF1, $tRNA^{Ser}$), the intertwining complexity of codon recognition (i.e., translational crosstalk) reveals the pervasive nature of crosstalk and the necessity to precisely engineer altered codon recognition in contrast to deletion. Thus, for most codons translational crosstalk has emerged as a major barrier to liberation. Our work overcomes this bottleneck in codon reassignment by introducing translation factor engineering as a necessary means of codon liberation with demonstrable results in reassignment efficacy.

Alongside deletion of RF1, our work employs rational engineering of two essential translation factors (RF2 and $tRNA^{Trp}$) to attenuate specific codon recognition and liberate stop codon TGA for sense codon reassignment. To achieve this, our work placed special emphasis on identifying and disentangling native translational crosstalk between 4 codons: UAG, UGA, UAA, and UGG. We illustrated this with a step-by-step breakdown of nsAA incorporation at single and multiple UAG alone, UGA alone, and both together. Fluorescent data, western blots (new data: **Extended Data Fig. 10h**), and MS-Read GFP data demonstrate precise and orthogonal incorporation of each nsAA at its respective codon with <1% misincorporation frequency. Nonsense suppression by OTSs has historically been inefficient without deletion or engineering of release factors. Our results surpass previous sense and nonsense incorporation efficiencies demonstrated within the field and reveal the practical value of identifying and mitigating complex translation competition for codons to achieve precision translation and high purity.

While our work emphasized the expression and purity of incorporation at each codon individually, the previous mentioned works demonstrated incorporation of nsAAs at multiple codons simultaneously. Incorporation at UAG, UCG, & UCA or incorporation at UAG & UCG were demonstrated through protein expression in the absence or presence of all corresponding nsAAs in media together. No codon was examined independently and no data on incorporation accuracies were shown. Given that the two recoded serine codons (TCG & TCA) are predicted to wobble pair with OTS tRNA anticodon UGA (Reis *et al. Nucl Acids Res.* 2004; 32(17), 5036-5044) employed in Robertson *et al.* (*Science* 2021; 372, 1057-1062), we expect misincorporation of nsAAs at UCG, thereby lowering translational precision, accuracy of nsAA incorporation, and protein purity, limiting practical utility. To our knowledge, the engineering of two tRNAs that exclusively decode TCG and TCA without crosstalk has not been reported. Our translation engineering efforts specifically sought to mitigate these limitations in precision codon targeting of UGA by mitigating $tRNA^{Trp}$ wobble (pages 11-12, lines 373-407). Looking forward, more cases exist of sense codons that share degenerate tRNAs where engineering codon recognition of factors may provide less deleterious outcomes than deletion. The lessons learned from our work in exploiting the plasticity of translation factor function, especially our work on $tRNA^{Trp}$, could inform engineering of $tRNA^{Ser}$, and more broadly, future sense recoding efforts. For further discussion on this, see our response #3 to Reviewer 1.

The authors highlight that their work represents the first case of a genetically recoded organism in which the degeneracy of codons is removed, but fall short to explain why that is of particular interest both fundamentally (methionine codon in the genetic code is not degenerate) or practically (not beneficial for unnatural amino acid incorporation and/or resistance to phages).

2. We thank the reviewer for raising these interesting points. The theoretical implications of removing codon degeneracy and ultimately pursuing nondegeneracy of the entire code involve: practically (a) significant improvements in protein purity arising from enhanced precision of nsAA incorporation at liberated codons (as elaborated in the

previous answer), (b) maximal expansion of available *in vivo* peptide chemistries (using canonical triplet-base codons) to the point where one could envision production of proteins consisting entirely of nonstandard amino acids, thereby granting properties like immunity to proteases and bio-orthogonal chemistries, and fundamentally (c) interrogation of the natural limitations and expansions of codon recognition by translation factors (e.g., tRNAs, release factors). Many tRNAs without modifications present wobble recognition of multiple codons, particularly with less discriminate recognition of the third base: e.g., G pairing with C and U, and Inosine (I) pairing A, C, or U. Anti-codon modifications can further restrict or expand codon wobble recognition (e.g., rendering $tRNA^{Met}$ exclusive for AUG, $tRNA^{Ile}$ exclusive to AUA, or $tRNA^{Trp}$ expanded to UGA from UGG). One such case was the focus of a biomolecular engineering effort in this study: engineering tRNA-Trp to exclusively decode UGG. Whether anti-codons and chemical modifications are available to ensure each canonical codon can receive an exclusive translation factor binding partner is an open question that can be addressed in pursuance of a nondegenerate code, as mechanical and chemical limitations in codon recognition are discovered. Additionally, with the liberation of each codon, unknown near-cognate suppressing factors and cryptic codon functions may be discovered as wildtype codon competition is relieved. The increased suppression of UGA by $tRNA^{Trp}$ under stress responses, elongating proteins and altering global gene expression, reveals additional cellular functions unique to the TGA stop codon besides termination (Romero, *Nat. Comm*. 2024). Once a codon is liberated, stress testing of cells may reveal codon functions lost to codon replacement.

Regarding further practical applications, our experiments modifying release factor codon recognition have opened opportunities to (i) further interrogate the complexity of finely-tuned interactions between RF1, RF2, and RF3 in controlling translation termination, (ii) apply RF engineering knowledge in other organisms (e.g., eukaryotes typically have a single omnipotent RF), and (iii) reassign RFs to terminate at liberated sense codons, creating non-canonical stop codons to serve as a defense against phage infection. Further optimization of UAA-specificity of RF2 should also enable broad resistance to UGA-dependent phage, as indicated by our phage data (**Fig. 3; Extended Data Fig. 7**) showing a shift in resistance from UAG-dependent Lambda phage to UGA-dependent Mu phage.

In contrast to previous work on recoded organisms, liberation of the opal codon required additional changes to both host RF2 (to attenuate suppression of TGA) and tRNATrp (wobble pairs with TGA). While these measures strongly attenuated translational cross-talk, they did not completely abolish it.

3. We thank the reviewer for this comment and opportunity to clarify. This study genetically isolates four codons: three stop codons and UGG. The mitigation of translation crosstalk is a key feature of this work that sets it apart as a major advance. With deletion of RF1, complete mitigation of UGA suppression by $tRNA^{Trp}$, and highly attenuated UGA recognition by mutant RF2, we achieved complete genomic translational isolation of 3 stop codons (UAA, UAG, and UGG) and the Trp codon. The minor remaining UGA cleavage by mutant RF2 modestly limits protein yield

(demonstrated in response #9 below) while maximizing purity (<1% misincorporation), resulting in pure synthetic protein at mg-scales. Further optimization of our system is achievable. Our phage results presenting immunity to UGA-dependent Mu phage *or* UAG-dependent Lambda phage suggest that complete abolition of UGA recognition by RFs is obtainable and may be achieved with a more comprehensive sampling of RF2 mutational space that we have begun investigating and plan to report in follow-up work.

In addition to reservations with respect to the novelty/utility of the work, there are some additional issues that need to be addressed:

1. The growth rates and optical densities reported in this work are very slow and low, respectively. For wild type Mg1655, a doubling time of 35 minutes at 37 °C is too slow, and a final OD600 of 0.8 is too low. The optical densities given in this work (Fig 4, Extended Data Fig. 3a, supplementary materials) appear to be relative optical densities obtained directly from the Biotek plate reader software without correction for path length (OD600 defined as 1 cm path length). This needs to be recalculated, or clearly indicated as relative ODs (and not OD600 or ODmax). The slow growth rates might be due to calculation errors caused by calculating growth rates based on these OD values.

4. We thank the reviewer for pointing out the inconsistency in growth data values relative to values typically reported in the literature. The reviewer is correct that optical densities given were raw reads without path length correction. We have generated a calibration curve (see **Methods,** page 26, lines 886-890) to account for the path length correction and used this to recalculate the doubling times and max OD_{600} consistent with prior reports (Kuznetsov *et al. Genome Biol* 18, 100; Lajoie *et al. Nucleic Acids Res* **40**, e170). This analysis allowed us to correct plate reader inconsistencies and path length, bringing the results more in-line with the field. **Figure 5** and **Extended Figures 4 and 8** have been updated to reflect this. These changes in OD and doubling time calculations are reflected in a revised Figure 5. The manuscript has been updated to account for changes in data and interpretations.

2. Fig. 6: promoter in front of AraC should be PC not PBAD

5. We thank the reviewer for pointing out this error and have corrected **Figure 6**.

3. Page 20 line 686: "…150 ml LB and 150 ml M9…": µl not ml

6. We thank the reviewer for pointing out this error and have made the necessary corresponding changes (page 26, line 871).

4. The authors should report yields of purified protein from this system (single and multiple suppression) to provide convincing evidence that this is useful technology. Previous work on recoded stains reported yields based on readouts of fluorescent protein- but the methodology did not provide useful yields of protein in practice.

7. The reviewer raises a valid point regarding the practical utility of this system for producing proteins with multiple instances of two distinct nsAAs at useful (mg-scale) yields. Notably, prior work in our lab and others have dedicated focused efforts toward this goal, incorporating 30 nsAAs in proteins or enhancing protein yield of C321 by 17-fold (e.g., Amiram *et al. Nat Biotechnol* 2015; 33, 1272-1279; Perez *et al. Biotechnol J.* 2021; 17(4):2100330). Inspired by this reviewer's comments to better demonstrate the usefulness of our technology, we conducted new experiments to demonstrate both single and triple dual incorporations of BocK and pAcF nsAAs at UAG and UGA, respectively, within ELP-GFPs, collecting protein yields (**Extended Data Fig. 10f**) following a validated quantification scheme published by Csibra & Stan, *Nat Commun* 2022; 13, 6600. Protein yields collected from rEc∆2.∆A.B3.tW* were compared to that of BL21, which is commonly used for protein expression, employing the same expression systems and represented in **Extended Data Fig. 10g**. Overall yields of purified dual-incorporated proteins were ~4.5 mg/L in TB for 1x(TGA-TAG) and ~2.3 mg/L in TB for 3x(TGA-TAG) (>99% accuracy), well over that of BL21 by ~20.8x and ~30.9x, respectively. Our rEc∆2.∆A.B3.tW* strain yields are significantly higher than that of BL21, demonstrating that recoding and RF engineering are required to enable single and multi-site nsAA incorporations of two distinct nsAAs. This is an important point as it counters a common misconception that simple stop codon suppression in non-recoded strains is largely sufficient for genetic code expansion. In addition, our results reach yields and translational fidelity generally sufficient for crystallography or other applications while employing the same constructs used in our ELP-GFP fluorescence experiments without additional optimization to maximize yield. We acknowledge that further optimization is needed to increase yield and anticipate that this will require increasing the plasmid copy number of proteins of interest, a systems level engineering of the orthogonal translation machinery, release factor engineering, and strain engineering. Such efforts are ongoing and will be the focus of follow-up studies. New data has been added to **Extended Data Figure 10f-g** and referenced in the manuscript (page 16, lines 542-551) to reflect these experimental results.

Referee #3 (Remarks to the Author):

Review of Isaacs single stop codon paper

In this paper the Rinehart/Isaacs team describe the construction of an *E. coli* strain that is powered by a single stop codon TAA. Beginning with their previously described "no TAG" strain recognizing both TAA and TGA they first proceeded to remove as many TGA stop codons as possible using "MAGE and CAGE" methods described earlier. They then described methods to engineer the RF2 release factor which recognizes TAA and TGA to make it able to only terminate at TAA. This required the use of various methods (and literature mining) and ultimately gave rise to a TAA specific RF2. They also presented data showing that the tRNATrp was extremely effective at reading TGA codons, inserting tryptophan. This "misreading" would interfere with efficient insertion of noncanonical amino

acids (ncAAs) and thus was unacceptable. In this case the solution was much easier and simply relied on a 1993 paper showing that changing one base adjacent to the anticodon of yeast tRNATrp sufficed to make it specific to TGG. By doing so, they constructed a strain that functions with a single stop codon. Performance of this strain with and without ncAAs was evaluated with reporter plasmids, and they showed successful incorporation of two ncAAs at TAG and TGA codons. Additional validation of strain performance was provided by phage sensitivity studies.

Overall this is a “First” and a well-designed and well -described study.

- We appreciate the reviewer’s summary and positive assessment of our work.

1) For practical deployment, it is important to know more about how sick this strain is compared to healthy E. coli strains. The current strain was compared to its immediate ancestor rather than a healthy E. coli strain. It would be good to do more extensive benchmarking against a few wild-type E. coli strains, not just MG1655, for example the Mach1 strain. Competitive growth would be a good addition.

- The original manuscript benchmarked growth to both the ancestor and a wild-type *E. coli* strain (MG1655). Inspired by the reviewer’s comment, we expanded this analysis across a panel of commonly used *E. coli* strains (TOP10, DH10B, Mach1, BL21) to grow alongside our recoded strains in both nutrient rich (LB) and minimal media (M9). Notably, the observed auxotrophies in TOP10 and DH10B are expected and reported previously (Durfee *et al. J Bacteriol.* 2008; 190(7):2597-2606). These data have been included in **Figure 4** with kinetic curves included in **Extended Data Figure 8d**.

The experiments that were done indicate that the RF1 parental strain is very sick (despite efforts to improve it). How will these problems be solved in the long run? Has adaptive lab evolution been carried out on the parent prior to this engineering, or after?

- Concurrent to this work, targeted strain engineering was conducted on the parent strain (C321.ΔA or rEcΔ1.ΔA) using both adaptive lab evolution (Wannier *et al. Proc Natl Acad Sci U S A* 115, 3090-3095) and genomic, metabolomic, and proteomic analyses (Hemez, *et al.* **doi:** https://doi.org/10.1101/2024.08.29.610322). These strain engineering efforts reverted previously validated mutations (*prfB*-A246T; Wannier *et al. Proc Natl Acad Sci U S A* 115, 3090-3095) and newly identified mutants (e.g., *ilvG*-Δ982T; Hemez *et al*. https://doi.org/10.1101/2024.08.29.610322) that collectively improve growth phenotype. These improved mutations were introduced into rEcΔ2.ΔA. Additional strain engineering experiments are in progress to further improve the health and fitness of our strains, including both adaptive evolutionary experiments alongside the rational targeting of key proteins within metabolic pathways, indels resulting in frameshifts, and nonsense mutations. We plan to isolate rEcΔ2.ΔA.B3.tW* strains with enhanced phenotypes and will report on these advances in future work. Included in our **Supplementary Table 14** are mutations identified and reverted during our recoding efforts to improve or restore

strain fitness (e.g. *purL* or *kefB*). Related growth data is included in **Supplementary Table 15**.

2) Figure 2 d/e. I found these bar graphs exceedingly difficult to absorb/understand. There is way too much going on here. For starters, in c it shows cherry "X" YFP. That's good. But then we see in d/e separate panels for cherry and YFP. Are these reporters the same ones as in c… or a separate set? In the first d panel. Why is TAG nearly zero? Perhaps it would be best to dumb these down a bit and focus on the important point which is the % readthrough in each configuration, and save some of the other data for the supplement. A final confusion is that I cant remember what the difference is between the first two strains in the key … can you add a scrap on info on what distinguishes them?

- The reviewer raises important questions regarding the clarity and interpretation of data in **Figure 2 d-e**. To clarify:

- *"we see in d/e separate panels for cherry and YFP. Are these reporters the same ones as in c… or a separate set?"*
  This is the same reporter as in c. The data from YFP and mCherry channels are plotted separately to quantify protein expression before and after each stop codon relative to the GCG codon control.

- *"In the first d panel. Why is TAG nearly zero? Perhaps it would be best to dumb these down a bit and focus on the important point which is the % readthrough in each configuration"*
  This is an important observation. We have previously characterized that in the absence of a release factor (RF1) and suppressor tRNA, ribosome stalling and product degradation results in loss of the protein and thus low mCherry expression at an open TAG codon (Ma et al. eLife 2018; 7:e34878. https://doi.org/10.7554/eLife.34878). Addition of a TAG suppressor tRNA (*supD_CUA*) alleviates stalling and promotes readthrough (Fig 2d-e, Middle). This effect also highlights why it is important to look at YFP and mCherry signals separately. High mCherry + Low YFP, High mCherry + High YFP and Low mCherry + Low YFP results reflect different translation outcomes. For instance, simply deriving %readthrough by dividing YFP by mCherry signal would result in artificially high values in cases where mCherry is degraded. Please see our next point on efforts we made to better communicate the subtleties of the data in **Figure 2d-e**.

- We recognize the reviewer's concerns in data presentation in **Figure 2d-e**. In response, we eliminated sense codon bars to simplify and minimize redundant information displayed (as bars are already normalized to sense codon fluorescence), updated our reporter schematics within each bar graph for clarity, created a guide (AKA "scrap") to the left of the bar graphs to aid readers in interpreting fluorescence results in relation to our reporter system, altered TAA bar graph colors to greyscale to mitigate color confusion with other figures (also in response to Reviewer 4's Question #10), and present absolute fluorescence data within a new **Supplementary Table 20**.

3) Figure 3. There two problems here.

1) There is no such number as zero in experiments of this type. The correct number to report is “< the minimum detectable number”.

a. This is a great catch by the reviewer and “zeros” have been changed to “N.D.” - Not Detected: lower than the minimal detectable number. **Figure 3** has been updated accordingly.

2) The statistics seem way off. The comparison of e.g. MG1655 to Ochre shows 5*10e8 vs. “Zero” plaques done in triplicate. It is very hard to believe that statistically this is a P<.05. I would think it was P <.many zeros one. My intuition is there is some major problem with calculating the stats here.

b. We agree with this intuition and consulted experts at Yale’s Statistics and Data Science Department and confirmed that the p-values of <0.05 (*) are mathematically valid when t-tests are conducted comparing population measurements (n=3) with populations consisting of all zero values (used to represent sample values N.D.).

Other issues
4) Line 111 please insert ref to C321.ΔA

- We thank the reviewer for this attention to detail and have added the reference (page 4, line 147).

5) Line 135 “In the latter two categories of TGA overlap, we designed specialized oligos 135 to recode the TGA-containing gene to minimize polar effects”. Cryptic!! Please provide more detail on how this done, what are these specialized oligos and exactly how are they designed?

- We thank the reviewer for pointing out the lack of clarity and have revised the manuscript accordingly. Oligos for this third category (i.e. TGA codons extensively overlapping genes) were designed to simply convert G to A within TGA, inevitably introducing overlapping gene mutations. This clarification was added to the main text (page 4, lines 158-173).
- “Although most oligos were designed to introduce a single-nucleotide substitution to convert TGA to TAA, 380 terminal TGA codons overlap a neighboring gene, and most cannot be converted in this way without impacting a neighboring gene. This overlap can fall into one of three categories (**Fig. 1c**). In the first category (26 instances), the A of TGA and ATG is shared, and thus mutating G to A in the TGA codon does not alter the ATG codon. In the second category (257 instances), both the T and G are shared between the ATG and TGA codons. Therefore, recoding TGA to TAA disrupts the start codon, which could preclude translation of the overlapped gene. To avoid this problem, the insertion of three bases – TAA – preserves both a new TAA codon and the ATG start codon, albeit with the ribosome binding site (RBS) shifted by three nucleotides from the start codon. The third

category of overlap (97 instances) encompasses all instances where two genes overlap extensively (3+ nucleotides shared), generally resulting in nonsynonymous mutations within overlapping genes (**Supplementary Table 2)**. In this latter category, oligos were designed to convert G to A within TGA, inevitably introducing an overlapping gene mutation. Since spacing of the RBS from the start codon or RBS-overlapping mutations can impact translation initiation rates[31], we used the RBS calculator[32] to predict translation initiation rates for both the wild type and recoded sequences (**Supplementary Table 3)** and found little predicted deviation in translation initiation upon recoding (**Fig. 1d**)."

6) Line 152 Please indicate the strategy used to identify the genes and pseudogenes that were deleted. Sounds as if this was all rather ad hoc, but some explanation of the considerations used would be important.

- The genes and pseudogenes were identified using the Genbank annotation. The approach used to delete these genes/pseudogenes was to identify genomic regions containing a high density of TGA-containing genes, screened against the Blattner paper (Posfai, *et al. Science* 2006; 312(5776):1044) to identify genetic segments that can be deleted without generating a deleterious phenotype. Based on these criteria, we performed 15 targeted deletions with counter selectable markers (tolC) to delete a large genome chunk and then remove the marker to leave a seamless deletion. We then narrowed down the number of regions by selecting those that contain numerous TGA-ending genes, while avoiding those that contain motility genes, resulting in 15 targeted genomic deletions. This clarification was added to the main text (page 5, lines 192-195). In response to remaining pseudogenes, we added oligos during the MAGE recoding process to revert other pseudogene containing TGA codons to passively convert these sites during the recoding process. Thus, a significant fraction of TGA-containing pseudogenes was reassigned to TAA and identified in the WGS data at the completion of the recoding process.

7) Lines 166-175. Yikes!! Please specify how many mutations accumulated overall and include in a supplemental table. Why were auxotrophic mutations the only ones prioritized for fixing? What about missense mutations in essential genes or those required for fast growth?

- We included additional details regarding mutations arising during MAGE mutagenesis and means of addressing these mutations to our manuscript (pages 5-6, lines 207-223). Auxotrophic mutations were not the only ones prioritized, but were easily integrated into growth profiling on minimal media during the MAGE recoding process and CAGE assembly steps. Auxotrophies are indicative of non-lethal mutations impacting core pathways in nucleic acid and amino acid metabolism. These mutations are fairly easy to screen against via growth characterization in minimal media (M9), and fairly easy to correct due to broad reporting on auxotrophy-inducing gene knockouts, the capacity to employ multiplexed oligo targeted mutations reversions coupled with robust M9 selection. Meanwhile, mutations impacting growth rates can be cryptic and complex, especially within

genetic backgrounds containing numerous uncharacterized mutations with possible synthetic effects. Thus, to minimize acquisition of such mutations, we characterize 48-96 colonies in both LB and M9 to select for fastest growing variants after multiple rounds of MAGE mutagenesis and CAGE events. We have accounted for and categorized all mutations within our strains. Systematic identification and targeted reversion of potentially high impact strains is a clear next step for achieving further strain optimization to enhance both growth fitness and protein yields that will be the focus of follow-on studies.

The updated manuscript passage (pages 6, lines 208-223) has been included below: "Throughout the recoding and deletion process, background mutations were acquired across strain lineages toward the final rEcΔ2.ΔA genome, comprising 199 indels (including130 frameshifts; **Supplementary Table 11**) and 526 single nucleotide polymorphisms (SNPs) (including 169 synonymous SNPs and 301 nonsynonymous SNPs resulting in 13 nonsense mutations; **Supplementary Table 12**). In addition, longread sequencing revealed various complex mutations, including the loss of 4 tRNAs (LeuP, LeuQ, MetZ, MetW), duplication of LysT, and instances of rRNA mutagenesis (**Supplementary Table 13**). To mitigate the emergence of detrimental growth phenotypes arising during strain construction, cell cultures were plated after successive rounds of MAGE or conjugative assemblies. 95 colonies were picked and characterized for cell fitness and growth, measuring maximum OD_{600} and doubling time in lysogeny broth (LB) and minimal media (M9). Fastest growing strains with highest max OD_{600} were selected to progress for further recoding. When auxotrophy arose, strains were initially backcrossed with ancestral rEcΔ1.ΔA to overwrite background mutations arising outside of recoded subdomains (**Supplementary Methods**). If backcrosses failed to restore prototrophy, implicated mutations within recoded regions were identified through WGS and reverted via MAGE (**Supplementary Methods, Supplementary Tables 14a and 15**)."

8) Line 178 reword to e.g. "Engineering UAA-specific translation termination is required to establish a functionally recoded ΔTAG-178 ΔTGA GRO capable of efficiently inserting ncAAs at TAG and TGA codons".

- Sentence rewritten (page 6, lines 226-228) to: "Engineering UAA-specific translation termination is the first step required to decouple UGA translational crosstalk and establish a functionally recoded ΔTAG-ΔTGA GRO capable of efficiently inserting nsAAs at TAG and TGA codons."

9) Reword to e.g. "Our inability to delete RF2 in the presence of RF1, despite an absence of genomic TGA codons, supports the essentiality of alternative RF2 UGA termination-independent functions such as translational quality control35,36"

- Sentence rewritten (page 6, lines 235-238) to: "Our inability to delete RF2 in the presence of RF1, despite an absence of genomic TGA codons, supports the essentiality of alternative RF2 UGA termination-independent functions such as translation quality control[35,36], or suggests an insufficiency of RF1 to act as a solitary release factor in this genetic context."

10) Line 203 “Previous …recognition” is not a sentence!

- Line changed (page 8, lines 280-282) to: “Interrogating RF codon recognition in *E. coli*, previous studies performed *in vivo* and *in vitro* characterizations of the S205P mutation.”

11) 204-211. The authors should confirm that the “PPF” RF2 fails to complement an RF2 deletion in their parent strain to “complete the genetic proof” here.

- In response to this reviewer’s request, we constructed pVanCC expression plasmids bearing two *prfB* variants, B1(wildtype S205) and B2(mutant P205), under the control of a *vanR* regulator for inducible expression. Under a range of expression levels, we attempted to replace the genomic *prfB* gene with a kanamycin resistance marker within parent strains MG1655 & rEcΔ1.ΔA alongside recoded strain rEcΔ2.ΔA (**Extended Data Fig. 5a-b)**. While B1(S205) was able to complement *prfB* deletion in all strains, B2(P205) failed to complement a *prfB* deletion in MG1655 (**Extended Data Fig. 5c**), yet was able to achieve viable complementation in both rEcΔ1.ΔA & rEcΔ2.ΔA. We confirmed genomic *prfB* knockouts with PCR screens (**Extended Data Fig. 5d**) as well as WGS to confirm no compensatory secondary mutations had arisen. Titrating B2(P205) expression with vanillic acid showed direct effects on doubling time in both strains (**Extended Data Figure 5e-f**), with a narrow window of optimal growth and protein production around 5 μM vanillic acid. Both higher and lower expression levels were toxic, dramatically increasing doubling times, supporting why B2(P205) was genomically unattainable. We also characterize TGA codon recognition ability as a factor of B2(P205) expression and find that TGA permits significantly greater readthrough at concentrations of vanillic acid below 5 μM (**Supplementary Discussion 1**).
- We believe that the higher level of B2(P205) plasmid expression relative to genomic expression enabled viability of the S205P mutation in the two recoded strains without the need for compensatory mutations (e.g. *prfB*-E170K) to improve enzyme kinetics. Additionally, we infer that the sensitivity of strain fitness (doubling time and fluorescent protein expression) to varying B2(P205) expression suggests a depletion in UAA termination ability that also may have compromised the internal translation regulation mechanism of this *prfB* variant, which relies on accurate “UAA” termination to regulate RF2 concentrations (**Figure 2a**). This disruption in B2(P205) autoregulation may lead to greater changes in cellular concentrations relative to autoregulation-intact B1(S205), which demonstrated little fitness change over broad induction levels across strains. We infer that in our final B3(P205) variant, the compensatory E170K mutation sufficiently restores RF2 termination activity at UAA to enable viability at chromosomal expression levels while still permitting reassignment of UGA.
- The manuscript has been updated to include both a **Supplemental Discussion 1** and an **Extended Data Figure 5** to present data and analyses we refer to in the main text (page 8, line 290-293), while moving details on compensatory mutations, acquired when trying to achieve viable B2(P05) clones, into the supplemental discussion.

12) 212-220. Indicate somewhere that prfB encodes RF2, as most readers won't know this.

- Our original introduction to *prfB* is as follows (page 6, lines 228-230): "As in most prokaryotes, *E. coli* RF1, encoded by *prfA*, terminates translation at UAG and UAA, while RF2, encoded by *prfB*, terminates translation at UGA and UAA." We added an additional clarification later in the text (page 7, line 258) to remind readers the connection between RF2 and *prfB*. "The *prfB* mRNA, encoding RF2, contains an internal RBS upstream of a "slippery" CUU-UGA…"

13) 221. The Chlorobium prfB factoid is interesting but raises the question – why does it have this sequence???

- We share the reviewer's sentiment. We hypothesize that the employment of UAA in place of UGA within this autoregulatory region may have evolved as a means of precise regulation of RF2 translation using both RF1 and RF2 recognition and termination (Baranov *et al. EMBO Rep* 2002; 3, 373-377). However, we do not understand whether UAA or UGA impact the efficiency of RF2 autoregulation and wish to conduct future *in vitro* assessments of *prfB* translation with variant autoregulatory sequences as a possible means of optimizing strain fitness and performance.

14) 703 change "a sense codon" to "a GCG codon"

- Corrected! Page 25, Line 836

Referee #4 (genetic code expansion)

Isaacs et al. report the construction of an E. coli strain in which all natural TAG and TGA stop codons have been removed and Release Factor 2 (RF2) and tRNATrp further engineered such that the four codons TAG, TGA, TAA, and TGG are rendered nondegenerate and the TAG and TGA codons can be efficiently translated by unnatural suppressor tRNAs charged with non-standard amino acids (nsAAs).

This achievement is heroic and has both practical significance for the incorporation of nsAAs and genome synthesis and engineering and basic science significance for our understanding of the genetic code.

1. We thank the reviewer for this positive assessment and for recognizing both the basic science and applied significance of this study.

I will consider the Figures in reverse order.

Figure 6 is important because it measures the utility of the genome engineering for the incorporation of nsAAs.

As I will come back to, the nomenclature used for the four different strains is non-intuitive. The figure would be much easier to digest, especially for the broad readership of Nature, if intuitive names for the strains were used and the strains composition was defined in the Figure Legend. For clarity, rEcΔ1.ΔA = their original ΔTAG strain; rEcΔ2.ΔA = their fully recoded ΔTAG, ΔTGA strain; Ochre = fully recoded ΔTAG, ΔTGA, mRF2-oKP strain; and Ochre.tW* = fully recoded ΔTAG, ΔTGA, mRF2-oKP, mutated tRNATrp strain. To emphasize, these strain names are non-intuitive; for example Ochre is both ΔTAG and ΔTGA, which is a non-standard meaning of Ochre.

2. We recognize the reviewer's critique on nomenclature. To simplify language and clarity, we have revised strain names as follows:

    rEcΔ1.ΔA = ancestral recoded ΔTAG, ΔRF1 strain (a.k.a. C321.ΔA)
    rEcΔ2.ΔA = fully recoded ΔTAG, ΔTGA, ΔRF1 strain
    rEcΔ2.ΔA.B0 = fully recoded ΔTAG, ΔTGA, ΔRF1, WT-RF2 (B0: T246A) strain
    rEcΔ2.ΔA.B1 = fully recoded ΔTAG, ΔTGA, ΔRF1, mRF2 (B1: T246A, D26N) strain
    rEcΔ2.A.B2 = fully recoded ΔTAG, ΔTGA, RF1-WT mRF2 (B2: T246A, D26N, S205P) strain
    rEcΔ2.ΔA.B3 = fully recoded ΔTAG, ΔTGA, ΔRF1, mRF2 (B3: T246A, D26N, S205P, E170K) strain
    rEcΔ2.ΔA.B4 = fully recoded ΔTAG, ΔTGA, ΔRF1, mRF2 (B4: T246A, D26N, E170K) strain
    rEcΔ2.ΔA.B3.tW* = fully recoded ΔTAG, ΔTGA, ΔRF1, mRF2 (B3: T246A, D26N, S205P, E170K), mutant $tRNA^{Trp}$(A37G) strain

The data in Figure 6b, c, and d are all reasonable. As the authors point out it is significant that the suppression of UAG and hence fluorescence goes up in both Ochre and Ochre.tW* where the TAG, TGA, TAA, and TGG are rendered nondegenerate. It is particularly significant in Figure 6d that they can suppress both UAG and UGA with two different amino acids 3 times each in a single construct. And that the normalized fluorescence is similar for 6d with 2 nsAAs and 6b and 6c with 1 nsAA each. That said they are using nsAAs and aminoacyl-tRNA synthetase/tRNA pairs that are known to be particularly robust in these kind of systems. The authors could do further experiments to demonstrate how well their engineered strain performs for incorporation of nsAAs. For example, what is the fluorescence of their reporter with no stop codons and why are the fluorescent values in Figures b, c, and d not shown relative to that? What is the fluorescence for single site suppression versus suppression of 3 codons? What is the suppression yield for proteins/positions that historically do not suppress well? In addition to relative fluorescence another good measure of the robustness of the system is protein yield. What is the relative yield of the GFP fusion protein with and without suppression? The authors should decide on the appropriate experiments to test the robustness of their system relative to amber suppression in wt E. coli.

3. We thank the reviewer for these insightful comments. In response, we have performed a series of experiments to better characterize the performance of our strains at nsAA incorporation. Below we outline our actions regarding each point raised.
    a. *"[W]hat is the fluorescence of their reporter with no stop codons[?]"* – To address this point, we modified **Figure 6** to normalize protein fluorescence to sense codon (no stop codon) ELP-GFP fluorescence and included raw data readings within a new **Extended Data Fig. 20**.

    b. *"What is the fluorescence for single site suppression versus suppression of 3 codons?"* – To address this point, we performed additional OTS experiments demonstrating ELP-GFP fluorescent expression with a single UAG, a single UGA, or no stop codon (instead employing sense codon UAC) included within the ELP peptide. Results have been added to **Extended Data Fig. 10** and the manuscript updated (pages 13-15, lines 468-520) to reflect these experimental results. These results demonstrate similarly robust orthogonal and site-specific incorporation of pAcF at UGA and BocK at UAG in comparison to previously demonstrated triple codon incorporation of single amino acids.

    c. "That said they are using nsAAs and aminoacyl-tRNA synthetase/tRNA pairs that are known to be particularly robust in these kinds of systems. The authors could do further experiments to demonstrate how well their engineered strain performs for incorporation of nsAAs… What is the suppression yield for proteins/positions that historically do not suppress well?" – We agree and intentionally selected these systems to allow us to assess the performance of the strain, which is the focus of this study. Characterization of a broader set of nsAAs and tRNA—aaRS pairs is an important effort that we believe is beyond the scope of this study, and will be investigated in follow-up work employing our recoded system for production and characterization of novel biopolymers with applications for the biosynthesis of enzymes or biomaterials.

    d. *"What is the relative yield of the GFP fusion protein with and without suppression?"* - To address this point, we conducted further experiments to demonstrate zero, single, and triple dual incorporations of BocK and pAcF nsAAs at UAG and UGA, respectfully, into ELP-GFPs within rEc∆2.∆A.B3.tW*, collecting batch protein yields (presented in **Extended Data Fig. 10f**) following a quantification scheme modified from published work by Csibra & Stan, *Nat Commun* 2022; 13, 6600 (**see Methods**). We then compared expression yields of nsAA-incorporated proteins collected from rEc∆2.∆A.B3.tW* and the wild-type BL21 employing the same expression systems (represented in **Extended Data Fig. 10g**). Overall yields of purified dual-incorporated proteins were ~4.5 mg/L in LB for 1x(TGA-TAG) and ~2.3 mg/L in LB for 3x(TGA-TAG) (>99% accuracy), well over that of BL21 by ~20.8x and ~30.9x, respectively. These results reach yields at mg-scale and translation fidelities (>99% accuracy) generally sufficient for crystallography or other applications, yet had employed the same constructs used in our ELP-GFP fluorescence experiments without additional optimization (e.g., increasing plasmid copy number of protein, strain engineering to increase titer) to maximize yield, leaving room for future enhancement. Results have been

added to **Extended Data Figure 10f-g** and the manuscript updated (page 16, lines 542-551) to reflect these experimental results.

Figure 5 measures the doubling time of the different strains. The authors should show the full log plots in the Extended Data.

4. In response to the reviewer's comment, we have added the requested data to a new **Extended Data Fig. 8**.

Figure 4 clearly states where the tRNATrp mutants are and demonstrates that they reduce background suppression of UGA. The authors provide a concise and clear explanation of their design of their mutant tRNATrp in the manuscript. They are straightforward that there is still some readthrough of the UGA codon even in the tW* strain. They note the intriguing observation that there is tangible incorporation of Cys – although here they don't offer an explanation.

5. We thank the reviewer for raising this insightful comment. Cysteine near-cognate incorporation at UGA had been previously identified in *S. cerevisiae* yeast resulting from an i6 A37 anti-codon stem-loop modification inducing an anticodon wobble at the third base (Blanchet *et al. Proc Natl Acad Sci USA* 2018; 115, 3018-3023), acting as a possible parallel to *E. coli*. This note has been added to the manuscript for clarity (page 11, lines 398-400). While Cys incorporation at UGA is a common suppression event in *E. coli*, we did not see instances of Cys incorporation at UGA under expression of a UGA suppressor in our recoded strain rEcΔ2.ΔA.B3.tW*. Thus, modification of this tRNA to examine mitigation of Cys incorporation is not necessary for application of our work. However, this tRNA could present another future direction for further fine-tuning and disentangling of translational crosstalk should Cys impurities arise at UGA with other applications.

Again the Figure Legend could be easier to read. What is Ara inducing? What is TyrOTS and how is it turned on/off?

6. We thank the reviewer for this attention to detail and have made changes to the **Figure 4** by adding a schematic representation of our expression plasmid to enhance clarity, defined terms, and adjusted our figure legend accordingly. This comment also enabled us to detect an error in our reporting on Tyrosine OTS (TyrOTS) expression (i.e., we induced *Mj*-TyrRS with an IPTG-inducible promoter, not Ara, and expressed the accompanying $tRNA^{Tyr}$(UCA) with a constitutive ProK promoter). Here we included an updated excerpt from our main text (page 11, lines 390-393): "We used our YFP:mCherry fluorescent assays with an isopropyl β-d-1-thiogalactopyranoside (IPTG)-inducible *Methanocaldococcus jannaschii* tyrosyl-tRNA synthetase and proK constitutively expressed $tRNA^{Tyr}$(UCA) pair (MjTyr-OTS) designed to encode Tyr at UGA to evaluate both OTS activity and native $tRNA^{Trp}$ suppression at UGA (**Fig. 4b**)."

Figure 3 shows the phage infectivity data which is as expected but also interesting from

the vantage point of genome recoding.

7. We thank the reviewer for this comment.

Figure 2 is very important for showing the impact of RF2 engineering on rendering the TAG, TGA, TAA, and TGG codons nondegenerate. The data is fine. However, it does not display the breadth of the RF2 engineering laid out in the text. The Figure and the text should match more. And related the text needs to tell a story. Now it reads as though the authors are telling us everything they did and in the order they did it. The paper would be much clearer to a broad audience if they picked a few key points to addressed and showed that data in Figure 2 as well.

8. We greatly appreciate the reviewer's attention to clarity and agree with the recommendation for revision. We have employed this feedback to readdress our written layout of section "Engineering and characterization of mutant release factor 2" with a focus on each key mutation in alignment with the figure, as opposed to a layout of how mutations were achieved and in what order. The rewrite of this section can be summarized by the following passage (page 7, lines 243-251): "Five primary mutations were introduced via MAGE to the *prfB* gene encoding RF2: (1) T246A codon change to restore wildtype (WT) RF methylation, (2) terminal TGA-to-TAA conversion for transcript termination, (3) internal TGA to TAA for autoregulatory maintenance, (4) S205P codon substitution for UAA-specific codon recognition, and (5) E170K codon substitution to convey a peptidyl charge-flip and enable strain viability (**Fig 2a-b; Supplementary Table 14**). Two *in vivo* assays were used to measure the functional impact of key RF2 mutations on release at UAG and UGA codons. These assays measured (1) the ability of the mutant release factor to outcompete near-cognate and cognate suppression permitting or occluding fluorescent reporter translation and (2) the ability of the mutant release factor to inhibit phage infection."

Finally, in the text describing the strain engineering the authors should use a standard set of names for the strains that are intuitive and evoke what changes were made to that strain.

9. Please see our second response to this reviewer addressing this point.

A small point I would not use purple in the Figure 2 because it has a different meeting than in Figure 6.

10. We thank the reviewer for this feedback and have changed purple to grey scale in **Figure 2**. In response to this reviewer and comment #2 from Reviewer #3, we have made changes to the presentation of data in Figure 2 that should clarify illustration of this data set and eliminate color confusion concerns between Figures 2 and 6.

Figure 1 is a nice way to represent the authors genome engineering strategy, and they have used it successfully previously.

11. We appreciate the reviewer's positive remark.

Referees' comments: 2024-03-04380A

Point-by-point response (in blue) to Referees' comments:

Referee #1 (Remarks to the Author):

The authors have done a stellar job addressing reviewer concerns and questions, and the manuscript is now much improved.

- We thank the reviewer for comments and feedback during this process, particularly regarding focus and scope. These changes we believe strengthened the impact and applicability of this work.

Referee #3 (Remarks to the Author):

The authors have addressed the points raised well and I am generally very positive about this very interesting study.

- We thank the reviewer for comments and feedback during this process. The attention to method and detail no doubt strengthened this work.

In general readability has improved substantially, and some important control experiments have been added. The one change I'm less enthused about is the switch to the far more verbose (and thus accurate/systematic) strain names. The authors may want to go back to using "ochre" as a nickname for the unmemorable "rEcΔ2.ΔA.B3.tW*".

- We have returned to "Ochre" as a descriptor for the strain, while retaining "rEcΔ2.ΔA.B3.tW*" and names of ancestral strains within the main text where appropriate for clarity.

In the prior review, point 10, rewording was suggested. The new sentence does not parse well at all. Suggest changing "Interrogating RF codon recognition in E.
coli, previous studies performed in vivo and in vitro characterizations of the S205P mutation." to "Previous studies performed in vivo and in vitro characterizations of the S205P mutation in order to evaluate RF codon recognition."

- We have made a variant of the suggested change: "Previous characterizations of S205P evaluated RF codon recognition *in vivo* and *in vitro*" (page 7, line 220-221).

Referee #4 (Remarks to the Author):

The authors are very responsive to the reviewer's comments and the result is a manuscript which no question should be published in Nature. The authors emphasize the uniqueness and significance of their result that they can compress translation to a single stop codon and that they can eliminate translational crosstalk between four codons, UAA, UGA, UAG, and UGG. This result is of clear fundamental and practical interest across diverse fields. The manuscript and figures are extensively revised improving the clarity of the manuscript for a broad audience. The

data are accurate with statistical significance clearly indicated. Enthusiasm for this manuscript is extremely high.

- We thank the reviewer for comments and feedback during this process. The experiments performed in response no doubt helped strengthened the clarity of our writing and results.