The genetic code is not quite sacrosanct, given that there are some divergences from it in endosymbionts and a few organisms, but it is largely canonical with the standard 20 amino acids generally in play. Nonetheless, there have been a number of attempts to expand the genetic code based on both “top down” methods (in which organisms are forced to use noncanonical amino acids) and “bottom up” methods (in which organisms are engineered to use noncanonical amino acids). Most recently, there have been attempts to do both, together. In a recent article in *ACS Central Science*, Agostini and colleagues furthered a “top down” method for forced incorporation of 4- and 5-fluorotryptophan into the *Escherichia coli* proteome.

Similar feats had previously been attempted with 4-fluorotryptophan for both *E. coli* and for *Bacillus subtilis*. However, in the current instance, the authors did not feed *E. coli* the noncanonical amino acids, but rather the immediate precursors, 4- and 5-fluoroindole. The pathway for the production of the normal precursor, indole, was inactivated by knocking out the *trpLEDC* and *tnaA* genes leaving only the last step (tryptophanyl synthase, *trpBA*) for conversion of the modified indoles into tryptophan derivatives that were in turn incorporated into proteins following tRNA charging by tryptophanyl tRNA synthetase (*trpS*) (Figure 1).

Over a span of almost a hundred serial passages, the strains learned to incorporate solely fluorotryptophan into their proteomes at least as measured by a single protein, enhanced GFP. In consequence, the organism’s growth rates, i.e., doubling times, were slowed, by a factor of 2–8-fold relative to ancestral strain growth. In previous work, Bacher and Ellington had attempted to evolve a tryptophan auxotroph for the incorporation of 4-fluorotryptophan. This was similarly successful with proteins apparently almost completely incorporating tryptophan, but with great reductions in strain growth. A key difference between these results, though, was that the strains incorporating modified indoles could grow even in the absence of any supplied indole, whereas in the earlier experiment, contaminating tryptophan led to the evolution of efficient scavengers. While the current modified-indole utilizing strains still grew well in the presence of indole, the differences in growth were not nearly as stark as the previous tryptophan-scavenging strains, which in the end could not be fully weaned from their dependence on tryptophan.

The modified indole-adapted strains were analyzed by whole genome sequencing at various points throughout the evolution process, and these analyses revealed that few genomic changes were required for *E. coli* to adapt to forced...
incorporation. Mutations were noted in the repressor of the trp operon (trpR) and in tryptophanyl tRNA synthetase (trpS) that likely led to more facile production and incorporation of the modified amino acid. In addition, mutations in membrane transporters, efflux pumps, and permeases may have led to the accumulation of needed metabolites in the cell, including tryptophan precursors and fluorotryptophan. While some of these genes (trpS) were also observed to mutate in the tryptophan scavenger strain, others (the trp permease, mtr) showed very different outcomes between the two experiments, thereby implying that the two experiments set the strains on very different evolutionary paths as suggested below.

To deal with the concomitant proteome perturbations, protein chaperones and proteases involved in protein folding were upregulated in the early stages of evolution but decreased during the late stages. It is unclear whether this implied that the disruptions of proteostasis early on were due to inefficient protein synthesis or failure to fully accommodate the unnatural amino acid during folding. General improvements in incorporation efficiency would have solved the first problem, while mutations that impacted the folding of individual proteins might have remediated the second.

Overall, these results continue to suggest that it is difficult to move proteomes away from the 20 canonical amino acids, even when the substitution is as small as a fluorine for a hydrogen.

Overall, these results continue to suggest that it is difficult to move proteomes away from the 20 canonical amino acids, even when the substitution is as small as a fluorine for a hydrogen. Noncanonical amino acids are generally regarded by the organisms as toxins that slow growth, and the genetic substitutions observed do not change the preference for tryptophan, but allow the accommodation of the unnatural derivatives. Stress responses of various sorts are quickly engaged, even well beyond the maintenance of proteostasis, as described above. For example, in the case of a selenocysteine-utilizing strain, mutations were found that likely compensated...
for oxidative stresses caused by the new amino acid. Interestingly, in the modified indole-utilizing strains an increase in the concentration of biotin and lipid metabolites was noted by the authors, which in turn led to an increase in membrane fluidity.

Nonetheless, the authors set out to explore the extent to which an “oligogenic barrier” exists for the adoption of new genetic codes, in the form of a relatively few proteins that must mutate in order to incorporate a new, noncanonical amino acid (Figure 2). The original evolution of a B. subtilis strain that actually grew faster in 4-fluorotryptophan supplemented media than with the natural amino acid tryptophan was first shown nearly 40 years ago⁵ and still stands as the best proof of such an “oligogenic barrier.” More recent sequence analysis of B. subtilis strains that prefer modified tryptophans show a relatively small number of mutations, and some of which (RNA polymerase subunits rpoB and rpoC) may reflect changes in global gene expression that compensate for proteostasis disruption. In contrast, the tryptophan scavenger strain evolved by Bacher and Ellington⁴ either never leaped the barrier, or perhaps avoided it entirely during directed evolution (Figure 2). In this regard, the failure of the scavenger strain to “take the jump” may be similar to how the long-term evolution of an E. coli strain that can now utilize citrate as a carbon source was confounded by the presence of glucose in the media.⁷,⁸

The modified-indole utilizing strains now lie somewhere in between: they were properly forced to deal with a complete lack of tryptophan, and in consequence began the protein-by-protein modifications necessary to accommodate the new amino acid. It is possible that, were directed evolution to continue in the adapted E. coli strains, the authors might reach the new fitness optimum apparently identified in B. subtilis, where the noncanonical amino acid is favored. Alternatively, E. coli and B. subtilis (as well as plate-based versus solution-based selections) are not analogous as E. coli has more essential protective genes to attenuate or buffer against stress conditions, while B. subtilis has a smaller set of essential genes and is able to survive extreme conditions by endospore formation.⁹ In other words, different proteomes may be differentially malleable, and the “oligogenic barrier” may greatly vary between organisms.

Figure 2. The oligogenic race to fitness on a new genetic code. In the oligogenic hypothesis, there are various key proteins that must mutate to accommodate a new amino acid. Wong (1983) found that relatively few mutations could lead to a switch in preference for fluorotryptophan. Agostini (2020) found that relatively few mutations could lead to an accommodation of fluorotryptophan, with tryptophan still preferred. Bacher (2001) evolved a strain that required both and still greatly preferred tryptophan. The differences between these paths may have to do with the relative ability of the strains to accommodate change.

In other words, different proteomes may be differentially malleable, and the “oligogenic barrier” may greatly vary between organisms.
Author Information

Corresponding Author
Andrew D. Ellington — Center for Systems and Synthetic Biology, University of Texas at Austin, Austin, Texas 78712, United States; orcid.org/0000-0001-6246-5338; Email: ellingtonlab@gmail.com

Author
Fan Zhang — Center for Systems and Synthetic Biology, University of Texas at Austin, Austin, Texas 78712, United States

Complete contact information is available at:
https://pubs.acs.org/10.1021/acscentsci.0c01549

Notes
The authors declare no competing financial interest.

REFERENCES

(1) Bacher, J. M.; Hughes, R. A.; Tze-Fei Wong, J.; Ellington, A. D. Evolving New Genetic Codes. Trends Ecol. Evol. 2004, 19 (2), 69−75.
(2) Thyer, R.; Shroff, R.; Klein, D. R.; d’Oelsnitz, S.; Cotham, V. C.; Byrom, M.; Brodbelt, J. S.; Ellington, A. D. Custom Selenoprotein Production Enabled by Laboratory Evolution of Recoded Bacterial Strains. Nat. Biotechnol. 2018, 36 (7), 624−631.
(3) Agostini, F.; Sinn, L.; Petras, D.; Schipp, C. J.; Kubyshevkin, V.; Berger, A.; Dorrestein, P. C.; Rappsilber, J.; Budisa, N.; Koksch, B. Multomics Analysis Provides Insight into the Laboratory Evolution of Escherichia Coli towards the Metabolic Usage of Fluorinated Indoles. ACS Cent. Sci. 2020 DOI: 10.1021/acscentsci.0c00679.
(4) Bacher, J. M.; Ellington, A. D. Selection and Characterization of Escherichia Coli Variants Capable of Growth on an Otherwise Toxic Tryptophan Analogue. J. Bacteriol. 2001, 183 (18), 5414−5425.
(5) Wong, J. T. Membership Mutation of the Genetic Code: Loss of Fitness by Tryptophan. Proc. Natl. Acad. Sci. U. S. A. 1983, 80 (20), 6303−6306.
(6) Yu, A. C. S.; Yim, A. K. Y.; Mat, W. K.; Tong, A. H. Y.; Lok, S.; Xue, H.; Tsui, S. K. W.; Wong, J. T. F.; Chan, T. F. Mutations Enabling Displacement of Tryptophan by 4Fluorotryptophan as a Canonical Amino Acid of the Genetic Code. Genome Biol. Evol. 2014, 6 (3), 629−641.
(7) Lenski, R. E. Phenotypic and Genomic Evolution during a 20,000-Generation Experiment with the Bacterium Escherichia Coli. Plant Breed Rev. 2010, 24, 225−265.
(8) Quandt, E. M.; Deatherage, D. E.; Ellington, A. D.; Georgiou, G.; Barrick, J. E. Recursive Genomewide Recombination and Sequencing Reveals a Key Refinement Step in the Evolution of a Metabolic Innovation in Escherichia Coli. Proc. Natl. Acad. Sci. U. S. A. 2014, 111, 2217−2222.
(9) Juhas, M.; Reuß, D. R.; Zhu, B.; Comnichau, F. M. Bacillus Subtilis and Escherichia Coli Essential Genes and Minimal Cell Factories after One Decade of Genome Engineering. Microbiology 2014, 160 (11), 2341−2351.