Medical Research Archives

European Society of Medicine

OPEN ACCESS

Published: August 31, 2023

Citation: Wong T and Xue H, 2023. Potential Impact of Genetic-Code Mutations on Medicine and Health, Medical Research Archives, [online] 11(8). https://doi.org/10.18103/mra.v11i8.4331

DOI
https://doi.org/10.18103/mra.v11i8.4331

ISSN: 2375-1924

RESEARCH ARTICLE

Potential Impact of Genetic-Code Mutations on Medicine and Health

Tze-Fei Wong* and Hong Xue

Division of Life Science, Hong Kong University of Science and Technology, Hong Kong, China

* **Correspondence:** bcjtw@ust.hk

ABSTRACT

The genetic code encoding amino acid sequences in ribosomal translation consists of an alphabet of 61 triplet codons for 20 amino acids and three chain termination signals. Basically the same universal code is employed by all organisms from the Last Universal Common Ancestor (LUCA)-proximal *Methanobacter kandleri* (Mka) to humans. This universal code, which has remained invariant for all living species, enables the transplantation of protein-coding genes between different species without loss of function, and constrains the chemical diversity of the encoded amino acids. Over the initial decades following the discovery of the code, its invariance coupled with the lack of any information regarding its origin have led to the view that the code might represent an inexplicable 'frozen accident' in the history of life. However, with the formulation of the *coevolution theory* of the genetic code and its multifaceted supporting evidence, this view has become untenable. Instead, the encoded amino acids are known to comprise two different classes: ten Class 1 amino acids available on prebiotic Earth were incorporated into the protocells as they evolved into life forms, while the ten Class 2 amino acids were produced by early life through biosynthesis. Thus, the later entry of the Class 2 amino acids identified them as end products of cellular evolution, which suggests the plausibility of continuing alterations of the encoded amino acids after an eons-long pause. Accordingly, attempts were made by our group to replace Trp by 4-fluroTrp (4FTrp) from the proteome of *Bacillus subtilis*. The targeted replacement obtained proved the inherent mutability of the code, and this has stimulated the development of a wide range of mutated codes through a variety of approaches. Hundreds of genetic code mutants have now been successfully isolated from microbes to animals, transforming the code from an immutable construct to a highly malleable molecular device. The effects of such new codes on medicine and health range from treatments for a variety of diseases to the alleviation of food crisis arising from the degradation of the environment and devastation due to natural disasters.

Introduction

The origin of the Last Universal Common Ancestor (LUCA) as the root of life has long been sought by science. Ever since their discovery, the submarine hydrothermal vents have been regarded as attractive sites to house prebiotic development[1,2]. These vents emitted hydrogen, carbon dioxide and little oxygen, thereby favoring chemical reduction over oxidation. At the vent temperatures (~350°C), energetics favored the formation of organic compounds from inorganic ones[3], which would later become intermediates of the reductive tricarboxylate, tricarboxylate and glyoxylate cycles[4,5]. Given the dissimilar pHs of vent fluid and the surrounding sea water, bioenergy may be generated as well through chemiosmosis in the vents[6].

In terms of genomics, the average divergences between alloacceptor tRNA sequences that serve as carriers of amino acids in translation differ between organisms, with a display of minimum divergence by the anaerobic hyperthermophilic archaeon *Methanopyrus kandleri* (Mka) isolated from the vents[7], indicating that Mka is phylogenetically closest to LUCA[8]. This conclusion has received widely based support from the usage of the simplest wobble rules for translation by Mka, its closely related tRNA sequences for reading the two separate codon domains of Ser[9], and the display of the smallest difference between the paralogous protein sequences of Val-tRNA synthetase (VARS) and Ile-tRNA synthetase (IARS) among 5000-plus species of organisms[10]. The Mka node was also the oldest among archaea, dating back to 2.8 Gya[11].The objective of the present study is to examine some of the major landmarks in the development of the genetic code from *Methanopyrus* to the present.

Formation of RNA Genes

A living cell has to fulfill two basic functions: to store and replicate its organization blueprint so that it can be passed on to the next generation, and to produce gene products to perform the catalytic and structural functions required by the act of living. Proteins cannot store information, whereas nucleic acids excel in information storage and also serve as catalysts, thus leading to the concept of an RNA World in the prebiotic age[12,13]. Cyclic and acyclic analogues of ribose and nucleobases were obtainable from thermosynthesis and meorites[14,15], suggesting that RNA-like structures preceded the formation of present day RNA through the evolution of RNA polymerase ribozymes[16]. The polymerization of nucleotides to yield a product RNA strand would be guided by a complementary template RNA strand via A-T and G-C basepairings, so that the product strand was bound to the template strand in a double helix[17]. Since only one out of 10^{24} random 40-mer RNA templates weighing one kilogram would include a functional 40-mer RNA with aptamer or ribozyme activity, it would need random RNA templates equal to the mass of the Earth to generate two functional RNA molecules[18]. Furthermore, because the vents did not undergo the temperature cycles of a polymerase-chain-reaction (PCR) incubator, the double helices formed by the old-new RNA strands would remain mostly as dead-end duplexes that can no longer replicate themselves. Although primer extension could bring about limited strand displacement, the problem posed by the rarity of functional RNA strands amidst a flood of random non-functional ones was an intractable obstacle to the development of a functional RNAome. To resolve this problem, it became essential to single out the minute fraction of duplexes that included an RNA strand with aptamer or ribozyme function, and induce their melting through the binding of a ligand of the aptamer or ribozyme to the functional strand, thus freeing up the other strand for continual replication. This *replicator induction by metabolites* (REIM) mechanism brought about the selective replication of duplexes containing functional RNAs while the mass of non-functional random duplexes remained unmelted and underwent decay in time, allowing the recycling of their constituents to suport the proliferation of the functional duplexes into clones of RNA genes[19,9].

Coevolution Theory

The ribozymes at the vents consisted of the four nucleotidyl letters U, C, A and G, and relied on their sidechains to catalyze different types of biochemical reactions, which were barely adequate for the task[20]. Accordingly, they had to acquire extra catalysis-competent sidechains post-transcriptionally, as in the case of modern tRNAs and rRNAs. Where the extra sidechains included amino acids and peptides from the prebiotic surroundings, the latter's outstanding catalytic properties soon became evident, and they were positioned at the catalytic centers of ribozymes, while the RNA moieties themselves functioned as mRNAs for the ribozyme sequences[21].

Many organisms contained amino acid biosynthetic pathways that converted precursor amino acids into product amino acids. The early cells selected some of these products for encoding by the code as additional letters of the protein alphabet. Comparison of the codon locations for the 20 finalized amino acids in the code with the

precursor-product biosynthetic relationships between them led to the coevolution theory which proposed that biosynthesis played an important role in determining codon locations: the Phase 1 amino acids Gly, Ala, Ser, Asp, Glu, Val, Leu, Ile, Pro and Thr in the code were selected from the prebiotic environment, whereas the Phase 2 amino acids Arg, His, Trp, Lys, Cys, Met, thermal unstable Gln and Asn, and UV-labile Phe and Tyr were selected subsequently from the products of biosynthetic pathways that coevolved with the code[22,23]. That the GRA 95229 Antarctic meteorite[24] and bombardment of high-energy particles[25] yielded only Phase 1 but no Phase 2 amino acids, together with the extremely low maximum concentrations of Gln and Asn, respectively 3.7 pM Gln and 24 nM Asn, in the prebiotic soup on account of their thermoinstability[26], strongly confirmed the coevolution view and the two-stage development of the 20-amino acid code.

Mutability of the Code

Despite the invariance of the 20-amino acid code among living species from archaea to bacteria and eukaryotes, there is no evidence for true invariance. On the contrary, that the Class 2 amino acids had to be evolved to complement the Class 1 amino acids could imply that code evolution entered into a prolonged pause due to the existence of oligogenic barriers consisting of conservative amino acid residues in the proteome firmly opposed to the loss of particular canonical amino acids (cAAs), in which case the frozen code may resume its evolution if the strength of these opposing residues were suppressed by mutations.

Accordingly, our group carried out rounds of iterative random mutations on the Trp-auxotrophic *Bacillus subtilis* QB928 strain with deleted Trp-biosynthesis, and looked for detectible growth on 4-fluoroTrp (4FTrp) on agar. This resulted in the isolation of an intermediate LC33 strain that grew well on Trp and 4FTrp, and subsequently the HR23 strain which grew well on 4FTrp but not on Trp. This HR23 mutant thus became the first organism in the history of life with a mutated code, where Trp has been displaced from the code by 4FTrp to become an inhibitor of growth on 4FTrp[27-29].

Isolation of Mutated Codes

The isolation of the HR23 strain proved the intrinsic mutability of the code. This opened wide the floodgate for mutated codes, which led to the successful isolations of hundreds of mutant codes. The mutation technologies employed have included three general approaches[30]:

Selective pressure incorporation (SPI). Trp was changed to 4FTrp in the HR23 code through two rounds of selection: weakening the oligogenic barrier of conservative amino acids in the QB928 wildtype against the removal of some key Trp residues to yield LC33, followed by the nurturing of an opposite oligogenic barrier of conservative amino acids against the removal of some key 4FTrp residues. This method for adding or switching codon-amino acid affiliations has been employed to prepare 'fluorous' Teflon-like proteins by means of thorough fluorination of Pro, Phe and Trp residues with little change in the activity or structure of the proteins[31]. In another instance, genetic code expansion involved the amino acid selenocysteine (Sec) found in some species. Following a series of mutations, Sec became encodable by erstwhile amber stop codons, expediting the replacement of labile cystine bridges in the proteome by Sec-Sec bridges with improved resistance to reducing conditions[32].

Stop codon suppression (SCI). The translation apparati of living organisms consist of amino acids, tRNAs, ARSs, mRNAs, and ribosomes. Their mutual interactions are highly specific. Consequently, when an exogenous ARS/tRNA pair is placed into a host cell, they would form independent, i.e. orthogonal, components for incorporating a noncanonical amino acid (ncAA) into proteins in response to codons designated for them. Furthermore, we found that archaeal tRNAs shared with yeast, rat liver and wheat germ tRNAs a distinct preference for aminoacylation by nonbacterial ARS over bacterial ARS from *E. coli* and *Rhodopseudomonas spheroides*[33]. On this basis, a yeast suppressor tRNA(Phe)/ARS(Phe) pair was employed to incorporate the ncAA p-fluoroPhe into the DHFR marker protein in response to an erstwhile amber stop codon with two-thirds of the expression level of wildtype DHFR[34]. This usage of orthogonal ("O") components to encode ncAAs was soon adopted for single-site and multi-site ncAA incorporations in *E. coli*[35,30,36,37], and at a Pro residue in Xenopus oocytes[38]. The search for appropriate O-ARSs is challenging, but it can be facilitated by the application of phage-assisted continuous evolution (PACE) of ARS[39]. Spare codons can be derived by subdividing codon boxes that have been allocated to a single cAA[40]. O-tRNA/O-mRNA pairs are valuable for making quadruplet codons useful[41], and UAG stop codons can be converted to sense codons by the removal of release factor F1 to restrict competition against ncAA incorporation by the chain termination process[42].

Unnatural Basepairs (UBP). DNA contains the four nucleobases T, C, A and G, whereas RNA contains the four nucleobases U, C, A and G. However, it was found that thymine in *E. coli* DNA could be replaced by 5-chlorouracil[43]. This led to the usage of unnatural basepairs (UBP) to expand the genetic code. In the case of clonal semisynthetic organisms (SSOs), O-DNA-carrying plasmids were used to transform *E. coli* host cells harboring an accessory plasmid that encoded a chimeric pyrrolysyl-tRNA synthetase variant. The cells were grown with a dNaM-d5SICS UBP pair, followed by addition of the corresponding mRNA UBP pair and the ncAA AzK to ensure the permanent presence of AzK at the targeted protein sites[44]. DNA- and RNA-like systems were also built from eight nucleotide letters that formed four orthogonal pairs, and the synthetic biopolymer obtained met the structural requirements of Darwinism, including a polyelectrolyte backbone and predictable thermodynamic stability. Transcription of the nc-DNA into nc-RNA gave rise to aptamers with eight-letters[45].

Medical Applications

In recent years, much of the effort directed to genetic code mutation and expansion has focused on technology development, but a series of findings have demonstrated the far-reaching results that may be accomplished as illustrated by the following examples.

a. Genetic codes have been constructed that consist of 67 codons[44] or 68 codons[46], which provided more spare codons for ncAAs.
b. Enhancements of immunotherapeutics for treating disorders including tumors were exemplified by the synthesis of site-specific antibody-drug conjugates[47], use of a bispecific antibody against acute myeloid leukemia[48] and cancer treatment with IL-2[49],
c. Pegylations of proteins at designated sites turned them into more effective therapeutic agents, as exemplified by Fgf21 for overcoming insulin resistance in diabetes[50], conjugation of peptides to albumin to prolong their half-lives[51], diversification of antimicrobial peptides[52], control of high-density lipoprotein biogenesis[53], generation of a low-immunogenic and stable form of human growth hormone[54], and site-specific immobilization of Bone Morphogenetic Protein 2 to solid surfaces[55].
d. Extension of ncAA usages to whole genomes of organisms or organs. Examples included *E. coli*[56], yeast[57], animals[58], mouse brain[59], and compartmentalized genetic parts and circuits[60].
e. Application of ncAAs to whole organisms. Examples included using sense codon reassignments to induce viral resistance[61,62], and control of cell signaling in zebra fish embros[63].
f. After an ncAA has been introduced into the proteome of an organism, a supply of the ncAA has to be secured by the addition of the ncAA as a nutrient in order to maintain the presence of the ncAA. However, external ncAA supply would become unnecessary when the host organism can biosynthesize the ncAA. Examples included the ncAAs p-aminophenylalanine[64], selenocyteine[32], phosphothreonine[65]. 5-hydroxytryptophan[66] and Met-derivatives[67].

Besides such attained applications of ncAAs to medicine, it would be useful to consider a couple of more exploratory objectives over the longer term:

Aging-related tissue degeneration. All human tissues undergo aging, but often not at the same rates. As the world populations grow older, the disorders associated with instances of premature degeneration such as Alzheimer's disease, Parkinson's disease, type 2 diabetes are increasingly important but remain difficult to treat. For instance, osteoarthritis of the knee often incurs collagen degradation. In a recent study, subjects were treated with acetaminophen and native type 2 collagen each day. While the walking scores of the test subjects showed about 50% improvement, surpassing the effects of glucosamine and chondroitin, the collagen degradation markers in urine were undiminished[68], suggesting that replacement of degraded collagen might require reinforcement from 'fluorous' proteins containing fluorinated Pro and other cAA or ncAA residues[31].

For other types of degeneration, frequently no underlying defective proteins are clearly identified. Nonetheless, one may postulate that fragile cAA residues could be a contributing factor, and chemically unstable cAAs such as Cys and cystine in a model animal can be replaced by stable ncAA analogues such as Sec and diselenide[32] in order to determine if that would relieve some of the symptoms of the degeneration.

Alleviation of food crises. During the past decades, extreme weather has caused widespread floods and draught, compounding the food crises due to population pressure and conflicts. Therefore it would be useful if genetic code expansion can be deployed to increase food production. Possible approaches in this regard would include the following:

i. While vascular plants are used to normal sunlight with photosynthetically active radiation (PAR) of 100-2,000 micromole photons $m^{-2}s^{-1}$, algae at 300 m below sea surface can utilize dim light of 0.02 $m^{-2}s^{-1}$ for photosynthesis[69]. Such dim-light photosynthetic systems can be introduced into edible macro- and micro-algae grown rapidly on optimized cAA-ncAA mixtures to yield biomass.

ii. Wind power, tidal power, fusion power and power from solar panels covering all the world's highlands will be converted into photons for algal photosynthesis. The performance of the photosynthetic apparatus can be enhanced by ncAAs, as suggested by the kinetic alteration caused by the replacement of a single Trp residue in the Orange Carotenoid Protein of *Synechocystis* by benzothienyl-L-Ala[70].

iii. 'Autotrophic centers' can be established to culture archaeal methanogens on the hydrogen and carbon dioxide emissions from hydrothermal vents and boreholes, supported by cAA-ncAA mixtures for rapid growth.

iv. Synthesize starch chemoenzymatically from carbon dioxide, which can produce starch 8.5-fold faster than maize[71], using enzymes bearing cAA-ncAA mixtures to accelerate the rate of synthesis.

v. Water-compatible artificial spherical chromatophore nanomicelles comparable to the chromatophore of *Rhotobacter sphaeroides* has enabled the reduction of carbon dioxide to methane with 15% solar-to-fuel efficiency[72]. Admixture with protein and membrane components of *R. spheroides*, grown on optimized cAA-ncAA mixtures, may generate even higher efficiency.

The biomass and carbohydrates derived from these approaches can serve directly as food for humans, or fed to edible yeasts so that the ncAAs and archaeal lipids may be removed prior to human consumption.

Moreover, while collisions between medium-sized meteorites and the Earth can be avoided by diverting them from their collision courses, collisions with too-big-to-divert meteorites will bring about dust clouds that cut agricultural output. It is also predicted that the sun will turn into a red giant with decreased temperatures in about 5 Gyrs, but unforeseen factors might cause an extremely unlikely but not impossible premature occurrence of red-giant formation, massively reducing agricultural output. Under these circumstances, measures such as (i) to (v) can be activated to improve food supply. On the other hand, in the likely event that red-giant formation occurs only 5 Gyrs from now, these measures could eliminate or at least decrease the severity of the food shortages that are as old as human societies. Continuation of the bane of these shortages for another 5 Gyrs is not a tolerable outcome. In this regard, the ncAAs offer a rare opportunity to upgrade the productivity of food production processes.

Conclusions

The genetic code, a centerpiece of the DNA-RNA-protein information system unique to life, began to evolve its familiar 20-amino acid alphabet prior to the appearance of *Methanopyrus*. This encoded alphabet has furnished an invaluable foundation for microbes, plants, animals and humans in their movement, flight, sensory signaling and memory. Now that the code has revealed its malleability, and begun to encode hundreds of ncAAs in addition to the cAAs, the effects of these new genetic codes and the ncAAs they can encode are not readily surmised. However, based on the new advances on a broad front of ncAA applications illustrated in the preceding section, it is evident that the chemical diversity of proteins will be steeply enhanced by the genetic encoding of an enlarged array of organic compounds. Known enzymic reactions will be accelerated, and enzymic reactions hitherto not observed in living organisms may be devised, so that medicine and health will benefit from novel approaches for the treatment of cancers, degenerative diseases and other disorders.

Furthermore, humans the space travelers are searching for habitable planets in our galaxy. If human hibernation can be achieved, the distance of travel to suitable planets will be increased, and attractive planets as goals of migration identified. Wherever the top choice might be, the local environment will not be identical to that on Earth, and the 20 amino acid alphabet and eight nucleobase building blocks may need to be revised to fit the local conditions and resources. Consequently, the ncAAs and UBPs currently investigated in genetic-code expansion research could pave the way for humans to re-optimize our molecular biology in adaptation to life on the new planet.

Acknowledgments

This research was funded by grant number 5112-703-0110D-42002 from the University Grants Council of Hong Kong to Applied Genomics Center of Hong Kong University of Science and Technology.

Author contributions

Conceptualization and writing, T.F.W. and H.X.; both authors have read and agreed to the published version of the manuscript.

References

1. Corliss JB, Dymond J, Gordon LI, et al. Submarine thermal springs on the Galapagos rift. *Science*. 1979;203(4385):1073-1083. doi:10.1126/science.203.4385.1073.
2. Baross JA, Hoffman SE. Submarine hydrothermal vents and associated gradient environments as sites for the origin and evolution of life. *Orig Life Evol Biosph*. 1985;15(4):327-345. doi:10.1007/bf01808177.
3. Shock EL. Constraints on the origins of organic compounds in hydrothermal vent, *Orig Life Evol Biosph*. 1990; 20:331-367.
4. Martin W, Russell MJ. On the origin of biochemistry at an alkaline hydrothermal vent. *Philos Trans R Soc B Biol Sci*. 2006;362(1486): 1887-1926. doi:10.1098/rstb.2006.1881.
5. Muchowska KB, Chevallot-Beroux E, Morana J. Recreating ancient metabolic pathways before enzymes. *Bioorganic Med Chem*. 2019;27(12):2292-2297. doi:10.1016/j.bmc.2019.03.012.
6. Lane N, Allen JF Martin W. How did LUCA make a living? Chemiosmosis in the origin of life. *BioEssays*. 2010;32(4):271-280. doi:10.1002/bies.200900131.
7. Huber R, Kurr M, Jannasch HW, Stetter KO. A novel group of abyssal methanogenic archaebacteria (Methanopyrus) growing at 110°C. *Nature*. 1989;342(6251):833-834. doi:10.1038/342833a0.
8. Xue H, Tong KL, Marck C, Grosjean H, Wong JTF. Transfer RNA paralogs: evidence for genetic code-amino acid biosynthesis coevolution and an archaeal root of life. *Gene*. 2003;310:59-66. doi:10.1016/s0378-1119(03)00552-3.
9. Wong JTF, Ng SK, Mat WK, Hu T, Xue H. Coevolution theory of the genetic code at age forty: pathway to translation and synthetic life. *Life*. 2016;6(1):12. doi:10.3390/life6010012.
10. Long X, Xue H, Wong JTF. Descent of bacteria and eukarya from an archaeal root of life. Evol Bioinfor 2020; 16: 1-11.
11. Blank CE. Low rates of lateral gene transfer among metabolic genes define the evolving biogeochemical niches of archaea through deep time. *Archaea*. 2012;2012:1-23. doi:10.1155/2012/843539.
12. Gilbert W. Origin of life: The RNA world. *Nature*. 1986;319(6055):618-618. doi:10.1038/319618a0.
13. Cech TR. Self-splicing of group I introns. *Annu Rev Biochem*. 1990;59(1):543-568. doi:10.1146/annurev.bi.59.070190.002551
14. Stairs S, Nikmal A, Bučar DK, Zheng SL, Szostak JW, Powner MW. Divergent prebiotic synthesis of pyrimidine and 8-oxo-purine ribonucleotides. *Nat Commun s*. 2017;8(1). doi:10.1038/ncomms15270.
15. Callahan MP, Smith KE, Cleaves HJ II, et al. Carbonaceous meteorites contain a wide range of extraterrestrial nucleobases. *Proc Natl Acad Sci U S A* 2011;108(34):13995-13998. doi:10.1073/pnas.1106493108.
16. Tjhung KF, Shokhirev MN, Horning DP, Joyce GF. An RNA polymerase ribozyme that synthesizes its own ancestor. *Proc Natl Acad Sci USA*. 2020;117(6):2906-2913. doi:10.1073/pnas.1914282117.
17. Orgel, L. E.Prebiotic chemistry and the origin of the RNA world. *Crit Rev Biochem Mol Biol*. 2004;39(2):99-123. doi:10.1080/10409230490460765.
18. Joyce GF, Orgel LE. Prospects for understanding the origin of the RNA World. In: Gesteland RF, Cech TR, Atkins JF, eds. *The RNA World, Second Edition*. Cold Spring Harbor Laboratory Press; 1999:49-77. doi:10.1017/S1355838299001259.
19. Wong JTF. Emergence of life: from functional RNA selection to natural selection and beyond. *Front Biosci - Landmark*. 2014;19(7):1117. doi:10.2741/4271.
20. Cech TR. Crawling out of the RNA World. Cell 2009; 136(4): 599-602. Doi.org/10.1016/jcell2009.02.0.
21. Wong JTF. Origin of genetically encoded protein synthesis: A model based on selection for RNA peptidation. *Orig Life Evol Biosph*. 1991;21(3):165-176. doi:10.1007/bf01809445.
22. Wong JTF. A Co-Evolution theory of the genetic code. *Proc Natl Acad Sci U S A*. 1975;72(5):1909-1912. doi:10.1073/pnas.72.5.1909.
23. Wong JTF. Coevolution of genetic code and amino acid biosynthesis. *Trends Biochem Sci*. 1981;6:33-36. doi:10.1016/0968-0004(81)90013-x.

24. Pizzarello S. Meteorites and the chemistry that preceded life's origin. In: Wong JTF, Lazcano A, eds. *Prebiotic Evolution and Astrobiology*. CRC Press; 2009:46-51. http://dx.doi.org/10.1201/9781498713986.
25. Kobayashi K, Kaneko T, Saito T, Oshima T. Amino acid formation in gas mixtures by high energy particle irradiation. *Orig Life Evol Biosph*. 1998;28(2):155-165. doi:10.1023/a:1006561217063.
26. Wong JTF, Bronskill PM. Inadequancy of prebiotic synthesis as origin of proteinous amino acids. J Mol Evol 1979; 13: 115-125.
27. Wong JTF. Membership mutation of the genetic code: loss of fitness by tryptophan. *Proc Natl Acad Sci U S A*. 1983; 80(20):6303-06. doi:10.1073/pnas.80.20.6303.
28. Mat WK, Xue H, Wong JTF. Genetic code mutations: The breaking of a three billion year invariance. *PLoS ONE*. 2010;5(8):e12206. doi:10.1371/journal.pone.0012206.
29. Wong JTF, Xue H. Synthetic genetic codes as the basis of synthetic life. In *Chemical Synthetic Biology*, Luisi PL, Chiarabelli C.eds. Wiley, New York2010; pp.178-199.
30. Hoesl MG, Budisa N. In vivo incorporation of multiple noncanonical amino acids into proteins. *Angew Chemie Intern Ed* 2011; doi.org/10.1002/anie.201005680.
31. Merkel L, Schauer M, Budisa N. et al. Parallel incorporation of different fluorinated amino acids: On the way to"Teflon" like proteins. *ChemBioChem*. 2010;11(11):1505-1507. doi:10.1002/cbic.201000295.
32. Thyer R, Shroff R, Klein DR, et al. Custom selenoprotein production enabled by laboratory evolution of recoded bacterial strains. *Nat Biotechnol*. 2018;36(7):624-631. doi:10.1038/nbt.4154.
33. Kwok Y, Wong JTF. Evolutionary relationship between Halobacterium cutirubrum and eukaryotes determined by use of aminoacyl-tRNA synthetases as phylogenetic probes. *Can J Biochem*. 1980;58(3):213-218. doi:10.1139/080-029.
34. Furter R. Expansion of the genetic code: Site-directed p-fluoro-phenylalanine incorporation in Escherichia coli. *Protein Sci*. 1998;7(2):419-426. doi:10.1002/pro.5560070223
35. Santoro SW, Anderson JC, Schultz PG et al. An archaebacteria-derived glutamyl-tRNA synthetase and tRNA pair for unnatural amino acid mutagenesis of proteins in Escherichia coli. *Nucleic Acids Res*. 2003;31(23):6700-6709. doi:10.1093/nar/gkg903.
36. Italia JS, Addy PS, Erickson SB et al. Mutually orthogonal nonsense -suppression systems for precise protein labelling at up to three distinct sites. *J Am Chem Soc* 2019; 141: 6204-6212.
37. Isaacs FJ, Carr PA, Wang HH, Lajoie MJ, Church GM et al. Precise manipulation of chromosomes in vivo enables genome-wide codon replacement. Science 2011; 333: 348-353.
38. Lummis SCR, Beene DL, Lee LW, Lester HA, Broadhurst RW, Dougherty DA. Cis-trans isomerization at a proline opens the pore of a neurotransmitter-gated ion channel, Nature 2005; 438: 248-252. Doi: 10.1038/nature04130.
39. Bryson DI, Fan C, Guo LT, Miller C, Söll D, Liu DR. Continuous directed evolution of aminoacyl-tRNA synthetases. *Nat Chem Biol*. 2017;13(12):1253-1260. doi:10.1038/nchembio.
40. Xue H, Wong JTF. Future of the genetic code. *Life*. 2017;7(1):10. doi:10.3390/life7010010.
41. Xi Z, Davis L, Baxter K, Tynan A, Goutou A, Greiss S. Using a quadruplet codon to expand the genetic code of an animal. *Nucleic Acids Res*. 2021;50(9):4801-4812. doi:10.1093/nar/gkab1168.
42. Mukai T, Hayashi A, Iraha F, et al. Codon reassignment in the Escherichia coli genetic code. *Nucleic Acids Res*. 2010;38(22):8188-8195. doi:10.1093/nar/gkq707
43. Marlière P, Patrouix J, Döring V, et al. Chemical evolution of a bacterium's genome. *Angew Chemie Int Ed*. 2011;50(31):7109-7114. doi:10.1002/anie.201100535.
44. Zhang Y, Lamb BM, Romesberg FE et al. A semisynthetic organism engineered for the stable expansion of the genetic alphabet. *Proc Natl Acad Sci U S A*. 2017;114(6):1317-1322. doi:10.1073/pnas.1616443114.
45. Hoshika S, Leal NA, Benner S.A. et al. Hachimoji DNA and RNA: A genetic system with eight building blocks. *Science*. 2019;363(6429):884-887. doi:10.1126/science.aat0971.
46. Dunkelmann DL, Oehm SB, Chin JW. A 68-codon genetic code to incorporate four distinct non-canonical amino acids enabled by automated orthogonal mRNA design. *Nat Chem*. 2021;13(11):1110-1117. doi:10.1038/s41557-021-00764-5
47. Axup JY, Bajjuri KM, Scultz PG et al. Synthesis of site-specific antibody-drug conjugates using unnatural amino acids. *Proc Nat Acad Sci USA* 2012; 109(40): 16101-16106.
48. Lu H, Zhou Q, Deshmukh V, et al. Targeting human C-type lectin-like molecule-1 (CLL1) with a bispecific antibody for immunotherapy of acute myeloid leukemia. *Angew Chem Int Ed*

Engl. 2014;53(37):9841-9845. doi:10.1002/anie.201405353.

49. Rosenberg SA. IL-2: the first effective immunotherapy for human cancer. *J Immunol*. 2014;192(12):5451-5458. doi:10.4049/jimmunol.1490019.
50. Camacho RC, Zafian PT, Achanfuo-Yeboah J et al. Pegylated Fgf21 rapidly normalizes insulin-stimulated glucose utilization in diet-induced insulin resistant mice. *Eur J Pharmacol*. 2013;715(1-3):41-45. doi:10.1016/j.ejphar.2013.06.023.
51. Bak M, Park J, Min K. et al. Recombinant peptide production platform coupled with site-specific albumin conjugation enables a convenient production of long-acting therapeutic peptide. *Pharmaceutics*. 2020;12 (4):364. doi:10.3390/pharmaceutics12040364.
52. Baumann T, Nickling JH, Bartholomae M, Budisa N et al. Prospects of *In vivo* incorporation of non-canonical amino acids for the chemical diversification of antimicrobial peptides. *Front Microbiol*. 2017;8:124. Published 2017 Feb 2. doi:10.3389/fmicb.2017.00124
53. Zamanian-Daryoush M, Gogonea V, DiDonato AJ, et al. Site-specific 5-hydroxytryptophan incorporation into apolipoprotein A-1 impairs cholesterol efflux activity and high-density lipoprotein biogenesis. *J Biol Chem* 2020;295(15):4836-4848. doi:10.1074/jbc.ra119.012092.s.
54. Wu L, Chen J, Wu Y, et al. Precise and combinatorial PEGylation generates a low-immunogenic and stable form of human growth hormone. *J Control Release*. 2017;249:84-93. doi:10.1016/j.jconrel.2017.01.029.
55. Siverino C, Tabisz B, Lühmann T, et al. Site-Directed Immobilization of bone morphogenetic protein 2 to solid surfaces by click chemistry. *J Vis Exp*. 2018;(133):56616. Published 2018 Mar 29. doi:10.3791/56616.
56. Fredens J, Wang K, de la Torre D, et al. Total synthesis of Escherichia coli with a recoded genome. *Nature*. doi:10.1038/s41586-019-1192-5.
57. Richardson SM, Mitchell LA, Stracquadanio G, et al. Design of a synthetic yeast genome. *Science*. 2017;355(6329):1040-1044. doi:10.1126/science.aaf4557.
58. Brown W, Liu J, Deiters A. Genetic code expansion in animals. *ACS Chem Biol*. 2018;13(9):2375-2386. doi:10.1021/acschembio.8b00520.
59. Ernst RJ, Krogager TP, Maywood ES, et al. Genetic code expansion in the mouse brain. *Nat Chem Biol*. 2016;12(10):776-778. doi:10.1038/nchembio.2160.
60. Ellefson JW, Meyer AJ, Hughes RA, Cannon JR, Brodbelt JS, Ellington AD. Directed evolution of genetic parts and circuits by compartmentalized partnered replication. *Nat Biotechnol*. 2014;32(1):97-101. doi:10.1038/nbt.2714.
61. Lajoie MJ, Rovner AJ, Goodman DB, Aerni HR et al. Genomically recoded organisms expand biological functions. *Science* 2013; 342:357-360.
62. Robertson WE, Funke LFH, de la Torre D, et al. Sense codon reassignment enables viral resistance and encoded polymer synthesis. *Science*. 2021;372(6546):1057-1062. doi:10.1126/science.abg3029.
63. Liu J, Hemphill J, Samanta S, Tsang M, Deiters A. Genetic code expansion in zebrafish embryos and its application to optical control of cell signaling. *J Am Chem Soc*. 2017;139(27):9100-9103. doi:10.1021/jacs.7b02145.
64. Mehl RA, Anderson JC, Santoro SW, et al. Generation of a bacterium with a 21 amino acid genetic code. *J Am Chem Soc*. 2003;125(4):935-939. doi:10.1021/ja0284153.
65. Zhang MS, Brunner SF, Hueguenin-Dezot N, Liang AD et al. Biosynthesis and encoding of phosphothreonine through parallel selection and deepmsequencing. *Nat Methods*. 2017; 14(7): 729-736. Doi:101038/nmeth4302.
66. Chen Y, Tang J, Wang L, et al. Creation of bacterial cells with 5-hydroxytryptophan as a 21st amino acid building block. *Chem*. 2020;6(10):2717-2727. doi:10.1016/j.chempr.2020.07.013.
67. Schipp CJ, Ma Y, Al-Shameri A, et al. An engineered Escherichia coli strain with synthetic metabolism for in-cell production of translationally active methionine derivatives. *Chembiochem*. 2020;21(24):3525-3538. doi:10.1002/cbic.202000257.
68. Bakilan F, Armagan O, Ozgen M, Tascioglu F, Bolluk O, Alatas O. Effects of native type II collagen treatment on knee osteoarthritis: A randomized controlled trial. *Eurasian J Med*. 2016;48(2):95-101. doi:10.5152/eurasianjmed.2015.15030.
69. Wang X, Wang Y, Ling A, Shi Y et al. Rationales: photosynthesis of vascular plants in dim light. *Front Plant Sci* 2020; 11: 573881, doi: 10.3389/fpls.2020.573881.
70. Moldenhauer M, Tseng H-W, Kraskov A, Tavraz NN et al Parameterization of a single H-bond in Orange Carotenoid Protein by

atomic mutation reveals principles of evolutionary design of complex chemical photosystems. *Front Mol Biosci* 2023; 10:1072606.
Doi: 10.3389/fmolb.2023.1072606.

71. Cai T, Sun H, Qiao J, et al. Cell-free chemoenzymatic starch synthesis from carbon dioxide. *Science*. 2021;373(6562):1523-1527. doi:10.1126/science.abh4049.
72. Yu J, Huang L, Tang Q, Yu SB, Qi QY et al. Artificial spherical chromatophore nanomicelles for selective CO_2 reduction in water. *Nature Catalysis* 2023; doi. org/10.1038/s41929-023-00962-z.