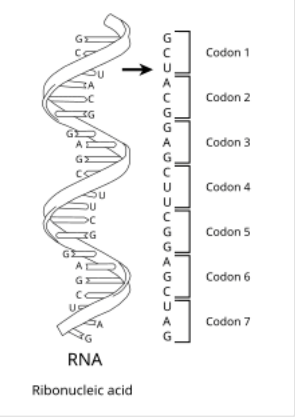
Genetic code

The **genetic code** is the set of rules used by living cells to translate information encoded within genetic material (DNA or RNA sequences of nucleotide triplets or codons) into proteins. Translation is accomplished by the ribosome, which links proteinogenic amino acids in an order specified by messenger RNA (mRNA), using transfer RNA (tRNA) molecules to carry amino acids and to read the mRNA three nucleotides at a time. The genetic code is highly similar among all organisms and can be expressed in a simple table with 64 entries.

The codons specify which amino acid will be added next during protein biosynthesis. With some exceptions,[1] a three-nucleotide codon in a nucleic acid sequence specifies a single amino acid. The vast majority of genes are encoded with a single scheme (see the RNA codon table). That scheme is often called the canonical or standard genetic code, or simply *the* genetic code, though variant codes (such as in mitochondria) exist.



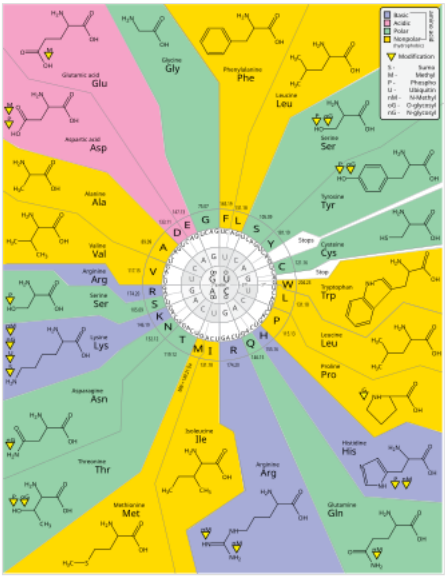
A series of codons in part of a messenger RNA (mRNA) molecule. Each codon consists of three nucleotides, usually corresponding to a single amino acid. The nucleotides are abbreviated with the letters A, U, G and C. This is mRNA, which uses U (uracil). DNA uses T (thymine) instead. This mRNA molecule will instruct a ribosome to synthesize a protein according to this code.

History

Efforts to understand how proteins are encoded began after DNA's structure was discovered in 1953. The key discoverers, English biophysicist Francis Crick and American biologist James Watson, working together at the Cavendish Laboratory of the University of Cambridge, hypothesied that information flows from DNA and that there is a link between DNA and proteins.[2] Soviet-American physicist George Gamow was the first to give a workable scheme for protein synthesis from DNA.[3] He postulated that sets of three bases (triplets) must be employed to encode the 20 standard amino acids used by living cells to build proteins, which would allow a maximum of $4^3 = 64$ amino acids.[4] He named this DNA–protein interaction (the original genetic code) as the "diamond code".[5]

In 1954, Gamow created an informal scientific organisation the RNA Tie Club, as suggested by Watson, for scientists of different persuasions who were interested in how proteins were synthesised from genes. However, the club could have only 20 permanent members to represent each of the 20 amino acids; and four additional honorary members to represent the four nucleotides of DNA.[6]

The first scientific contribution of the club, later recorded as "one of the most important unpublished articles in the history of science"[7] and "the most famous unpublished paper in the annals of molecular biology",[8] was made by Crick. Crick presented a type-written paper titled "On Degenerate Templates and the Adaptor Hypothesis: A Note for the RNA Tie Club"[9] to the members of the club in January 1955, which "totally changed the way we thought about protein synthesis", as Watson recalled.[10] The hypothesis states that the triplet code was not passed on to amino acids as Gamow thought, but carried by a different molecule, an adaptor, that interacts with amino acids.[8] The adaptor was later identified as tRNA.[11]



The genetic code

Codons

The Crick, Brenner, Barnett and Watts-Tobin experiment first demonstrated that **codons** consist of three DNA bases.

Marshall Nirenberg and J. Heinrich Matthaei were the first to reveal the nature of a codon in 1961.[12] They used a cell-free system to translate a poly-uracil RNA sequence (i.e., UUUUU...) and discovered that the polypeptide that they had synthesized consisted of only the amino acid phenylalanine.[13] They thereby deduced that the codon UUU specified the amino acid phenylalanine.

This was followed by experiments in Severo Ochoa's laboratory that demonstrated that the poly-adenine RNA sequence (AAAAA...) coded for the polypeptide poly-lysine[14] and that the poly-cytosine RNA sequence (CCCCC...) coded for the polypeptide poly-proline.[15] Therefore, the codon AAA specified the amino acid lysine, and the codon CCC specified the amino acid proline. Using various copolymers most of the remaining codons were then determined.

Subsequent work by Har Gobind Khorana identified the rest of the genetic code. Shortly thereafter, Robert W. Holley determined the structure of transfer RNA (tRNA), the adapter molecule that facilitates the process of translating RNA into protein. This work was based upon Ochoa's earlier studies, yielding the latter the Nobel Prize in Physiology or Medicine in 1959 for work on the enzymology of RNA synthesis.[16]

Extending this work, Nirenberg and Philip Leder revealed the code's triplet nature and deciphered its codons. In these experiments, various combinations of mRNA were passed through a filter that contained ribosomes, the components of cells that translate RNA into protein. Unique triplets promoted the binding of specific tRNAs to the ribosome. Leder and Nirenberg were able to determine the sequences of 54 out of 64 codons in their experiments.[17] Khorana, Holley and Nirenberg received the Nobel Prize (1968) for their work.[18]

The three stop codons were named by discoverers Richard Epstein and Charles Steinberg. "Amber" was named after their friend Harris Bernstein, whose last name means "amber" in German.[19] The other two stop codons were named "ochre" and "opal" in order to keep the "color names" theme.

Expanded genetic codes (synthetic biology)

In a broad academic audience, the concept of the evolution of the genetic code from the original and ambiguous genetic code to a well-defined ("frozen") code with the repertoire of 20 (+2) canonical amino acids is widely accepted.[20] However, there are different opinions, concepts, approaches and ideas, which is the best way to change it experimentally. Even models are proposed that predict "entry points" for synthetic amino acid invasion of the genetic code.[21]

Since 2001, 40 non-natural amino acids have been added into proteins by creating a unique codon (recoding) and a corresponding transfer-RNA:aminoacyl – tRNA-synthetase pair to encode it with diverse physicochemical and biological properties in order to be used as a tool to exploring protein structure and function or to create novel or enhanced proteins.[22][23]

H. Murakami and M. Sisido extended some codons to have four and five bases. Steven A. Benner constructed a functional 65th (*in vivo*) codon.[24]

In 2015 N. Budisa, D. Söll and co-workers reported the full substitution of all 20,899 tryptophan residues (UGG codons) with unnatural thienopyrrole-alanine in the genetic code of the bacterium *Escherichia coli*.[25]

In 2016 the first stable semisynthetic organism was created. It was a (single cell) bacterium with two synthetic bases (called X and Y). The bases survived cell division.[26][27]

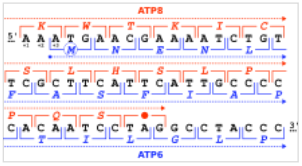
In 2017, researchers in South Korea reported that they had engineered a mouse with an extended genetic code that can produce proteins with unnatural amino acids.[28]

In May 2019, researchers reported the creation of a new "Syn61" strain of the bacterium *Escherichia coli*. This strain has a fully synthetic genome that is refactored (all overlaps expanded), recoded (removing the use of three out of 64 codons completely), and further modified to remove the now unnecessary tRNAs and release factors. It is fully viable and grows 1.6× slower than its wild-type counterpart "MDS42".[29][30]

Features

Reading frame

A reading frame is defined by the initial triplet of nucleotides from which translation starts. It sets the frame for a run of successive, non-overlapping codons, which is known as an "open reading frame" (ORF). For example, the string 5'-AAATGAACG-3' (see figure), if read from the first position, contains the codons AAA, TGA, and ACG ; if read from the second position, it contains the codons AAT and GAA ; and if read from the third position, it contains the codons ATG and AAC. Every sequence can, thus, be read in its 5' → 3' direction in three reading frames, each producing a possibly distinct amino acid sequence: in the given example, Lys (K)-Trp (W)-Thr (T), Asn (N)-Glu (E), or Met (M)-Asn (N), respectively (when translating with the vertebrate mitochondrial code). When DNA is double-stranded, six possible reading frames are defined, three in the forward orientation on one strand and three reverse on the opposite strand.[32]:330 Protein-coding frames are defined by a start codon, usually the first AUG (ATG) codon in the RNA (DNA) sequence.



Reading frames in the DNA sequence of a region of the human mitochondrial genome coding for the genes *MT-ATP8* and *MT-ATP6* (in black: positions 8,525 to 8,580 in the sequence accession NC_012920[31]). There are three possible reading frames in the 5' → 3' forward direction, starting on the first (+1), second (+2) and third position (+3). For each codon (square brackets), the amino acid is given by the vertebrate mitochondrial code, either in the +1 frame for *MT-ATP8* (in red) or in the +3 frame for *MT-ATP6* (in blue). The *MT-ATP8* genes terminates with the TAG stop codon (red dot) in the +1 frame. The *MT-ATP6* gene starts with the ATG codon (blue circle for the M amino acid) in the +3 frame.

In eukaryotes, ORFs in exons are often interrupted by introns.

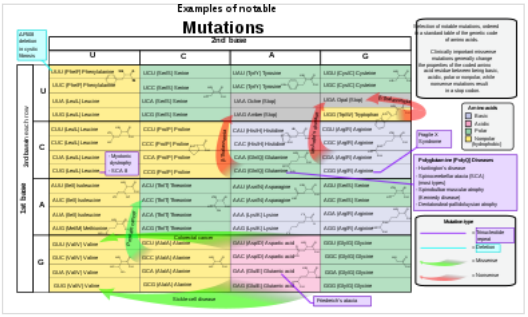
Start and stop codons

Translation starts with a chain-initiation codon or start codon. The start codon alone is not sufficient to begin the process. Nearby sequences such as the Shine-Dalgarno sequence in *E. coli* and initiation factors are also required to start translation. The most common start codon is AUG, which is read as methionine or as formylmethionine (in bacteria, mitochondria, and plastids). Alternative start codons depending on the organism include "GUG" or "UUG"; these codons normally represent valine and leucine, respectively, but as start codons they are translated as methionine or formylmethionine.[33]

The three stop codons have names: UAG is *amber*, UGA is *opal* (sometimes also called *umber*), and UAA is *ochre*. Stop codons are also called "termination" or "nonsense" codons. They signal release of the nascent polypeptide from the ribosome because no cognate tRNA has anticodons complementary to these stop signals, allowing a release factor to bind to the ribosome instead.[34]

Effect of mutations

During the process of DNA replication, errors occasionally occur in the polymerization of the second strand. These errors, mutations, can affect an organism's phenotype, especially if they occur within the protein coding sequence of a gene. Error rates are typically 1 error in every 10–100 million bases—due to the "proofreading" ability of DNA polymerases.[36][37]



Examples of notable mutations that can occur in humans[35]

Missense mutations and nonsense mutations are examples of point mutations that can cause genetic diseases such as sickle-cell disease and thalassemia respectively.[38][39][40] Clinically important missense mutations generally change the properties of the coded amino acid residue among basic, acidic, polar or non-polar states, whereas nonsense mutations result in a stop codon.[32]

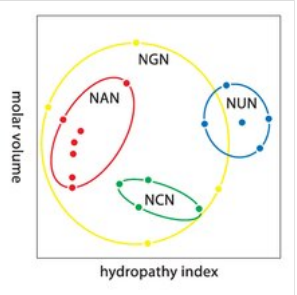
Mutations that disrupt the reading frame sequence by indels (insertions or deletions) of a non-multiple of 3 nucleotide bases are known as frameshift mutations. These mutations usually result in a completely different translation from the original, and likely cause a stop codon to be read, which truncates the protein.[41] These mutations may impair the protein's function and are thus rare in *in vivo* protein-coding sequences. One reason

inheritance of frameshift mutations is rare is that, if the protein being translated is essential for growth under the selective pressures the organism faces, absence of a functional protein may cause death before the organism becomes viable.[42] Frameshift mutations may result in severe genetic diseases such as Tay–Sachs disease.[43]

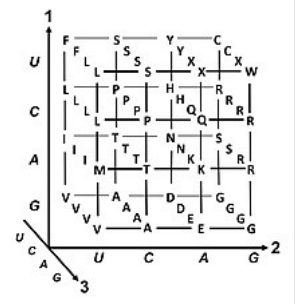
Although most mutations that change protein sequences are harmful or neutral, some mutations have benefits.[44] These mutations may enable the mutant organism to withstand particular environmental stresses better than wild type organisms, or reproduce more quickly. In these cases a mutation will tend to become more common in a population through natural selection.[45] Viruses that use RNA as their genetic material have rapid mutation rates,[46] which can be an advantage, since these viruses thereby evolve rapidly, and thus evade the immune system defensive responses.[47] In large populations of asexually reproducing organisms, for example, *E. coli*, multiple beneficial mutations may co-occur. This phenomenon is called clonal interference and causes competition among the mutations.[48]

Degeneracy

Degeneracy is the redundancy of the genetic code. This term was given by Bernfield and Nirenberg. The genetic code has redundancy but no ambiguity (see the codon tables below for the full correlation). For example, although codons GAA and GAG both specify glutamic acid (redundancy), neither specifies another amino acid (no ambiguity). The codons encoding one amino acid may differ in any of their three positions. For example, the amino acid leucine is specified by **Y**U**R** or CU**N** (UUA, UUG, CUU, CUC, CUA, or CUG) codons (difference in the first or third position indicated using IUPAC notation), while the amino acid serine is specified by UC**N** or AG**Y** (UCA, UCG, UCC, UCU, AGU, or AGC) codons (difference in the first, second, or third position).[49] A practical consequence of redundancy is that errors in the third position of the triplet codon cause only a silent mutation or an error that would not affect the protein because the hydrophilicity or hydrophobicity is maintained by equivalent substitution of amino acids; for example, a codon of NUN (where N = any nucleotide) tends to code for hydrophobic amino acids. NCN yields amino acid residues that are small in size and moderate in hydropathicity; NAN encodes average size hydrophilic residues. The genetic code is so well-structured for hydropathicity that a mathematical analysis (Singular Value Decomposition) of 12 variables (4 nucleotides x 3 positions) yields a remarkable correlation (C = 0.95) for predicting the hydropathicity of the encoded amino acid directly from the triplet nucleotide sequence, *without translation*.[50][51] Note in the table, below, eight amino acids are not affected at all by mutations at the third position of the codon, whereas in the figure above, a mutation at the second position is likely to cause a radical change in the physicochemical properties of the encoded amino acid. Nevertheless, changes in the first position of the codons are more important than changes in the second position on a global scale.[52] The reason may be that charge reversal (from a positive to a negative charge or vice versa) can only occur upon mutations in the first position of certain codons, but not upon changes in the second position of any codon. Such charge reversal may have dramatic consequences for the structure or function of a protein. This aspect may have been largely underestimated by previous studies.[52]



Grouping of codons by amino acid residue molar volume and hydropathicity. A more detailed version is available.



Axes 1, 2, 3 are the first, second, and third positions in the codon. The 20 amino acids and stop codons (X) are shown in single letter code.

Codon usage bias

The frequency of codons, also known as codon usage bias, can vary from species to species with functional implications for the control of translation. The codon varies by organism; for example, most common proline codon in E. coli is CCG, whereas in humans this is the least used proline codon.[53]

Human genome codon frequency table[54]

Codon	AA[C]	Fraction[D]	Freq ‰[E]	Number[F]	Codon	AA	Fraction	Freq ‰	Number	Codon	AA	Fraction	Freq ‰	Number	Codon	A
UUU	F	0.46	17.6	714,298	UCU	S	0.19	15.2	618,711	UAU	Y	0.44	12.2	495,699	UGU	C
UUC	F	0.54	20.3	824,692	UCC	S	0.22	17.7	718,892	UAC	Y	0.56	15.3	622,407	UGC	C
UUA	L	0.08	7.7	311,881	UCA	S	0.15	12.2	496,448	UAA	*	0.30	1.0	40,285	UGA	*
UUG	L	0.13	12.9	525,688	UCG	S	0.05	4.4	179,419	UAG	*	0.24	0.8	32,109	UGG	W
CUU	L	0.13	13.2	536,515	CCU	P	0.29	17.5	713,233	CAU	H	0.42	10.9	441,711	CGU	R
CUC	L	0.20	19.6	796,638	CCC	P	0.32	19.8	804,620	CAC	H	0.58	15.1	613,713	CGC	R
CUA	L	0.07	7.2	290,751	CCA	P	0.28	16.9	688,038	CAA	Q	0.27	12.3	501,911	CGA	R
CUG	L	0.40	39.6	1,611,801	CCG	P	0.11	6.9	281,570	CAG	Q	0.73	34.2	1,391,973	CGG	R
AUU	I	0.36	16.0	650,473	ACU	T	0.25	13.1	533,609	AAU	N	0.47	17.0	689,701	AGU	S
AUC	I	0.47	20.8	846,466	ACC	T	0.36	18.9	768,147	AAC	N	0.53	19.1	776,603	AGC	S
AUA	I	0.17	7.5	304,565	ACA	T	0.28	15.1	614,523	AAA	K	0.43	24.4	993,621	AGA	R
AUG	M	1.00	22.0	896,005	ACG	T	0.11	6.1	246,105	AAG	K	0.57	31.9	1,295,568	AGG	R
GUU	V	0.18	11.0	448,607	GCU	A	0.27	18.4	750,096	GAU	D	0.46	21.8	885,429	GGU	G
GUC	V	0.24	14.5	588,138	GCC	A	0.40	27.7	1,127,679	GAC	D	0.54	25.1	1,020,595	GGC	G
GUA	V	0.12	7.1	287,712	GCA	A	0.23	15.8	643,471	GAA	E	0.42	29.0	1,177,632	GGA	G
GUG	V	0.46	28.1	1,143,534	GCG	A	0.11	7.4	299,495	GAG	E	0.58	39.6	1,609,975	GGG	G

Alternative genetic codes

Non-standard amino acids

In some proteins, non-standard amino acids are substituted for standard stop codons, depending on associated signal sequences in the messenger RNA. For example, UGA can code for selenocysteine and UAG can code for pyrrolysine. Selenocysteine came to be seen as the 21st amino acid, and pyrrolysine as the 22nd.[55] Both selenocysteine and pyrrolysine may be present in the same organism.[55] Although the genetic code is normally fixed in an organism, the achaeal prokaryote *Acetohalobium arabaticum* can expand its genetic code from 20 to 21 amino acids (by including pyrrolysine) under different conditions of growth.[56]

Variations

Genetic code logo of the *Globobulimina pseudospinescens* mitochondrial genome by FACIL. The program is able to correctly infer that the Protozoan Mitochondrial Code is in use.[57] The logo shows the 64 codons from left to right, predicted alternatives in red (relative to the standard genetic code). Red line: stop codons. The height of each amino acid in the stack shows how often it is aligned to the codon in homologous protein domains. The stack height indicates the support for the prediction.

There was originally a simple and widely accepted argument that the genetic code should be universal: namely, that any variation in the genetic code would be lethal to the organism (although Crick had stated that viruses were an exception). This is known as the "frozen accident" argument for the universality of the genetic code. However, in his seminal paper on the origins of the genetic code in 1968, Francis Crick still stated that the universality of the genetic code in all organisms was an unproven assumption, and was probably not true in some instances. He predicted that "The code is universal (the same in all organisms) or nearly so".[58] The first variation was discovered in 1979, by researchers studying human mitochondrial genes.[59] Many slight variants were discovered thereafter,[60] including various alternative mitochondrial codes.[61] These minor variants for example involve translation of the codon UGA as tryptophan in *Mycoplasma* species, and translation of CUG as a serine rather than leucine in yeasts of the "CTG clade" (such as *Candida albicans*).[62][63][64] Because viruses must use the same genetic code as their hosts, modifications to the standard genetic code could interfere with viral protein synthesis or functioning. However, viruses such as totiviruses have adapted to the host's genetic code modification.[65] In bacteria and archaea, GUG and UUG are common start codons. In rare cases, certain proteins may use alternative start codons.[60] Surprisingly, variations in the interpretation of the genetic code exist also in human nuclear-encoded genes: In 2016, researchers studying the translation of malate dehydrogenase found that in about 4% of the mRNAs encoding this enzyme the stop codon is naturally used to encode the amino acids tryptophan and arginine.[66] This type of recoding is induced by a high-readthrough stop codon context[67] and it is referred to as *functional translational readthrough*.[68]

Despite these differences, all known naturally occurring codes are very similar. The coding mechanism is the same for all organisms: three-base codons, tRNA, ribosomes, single direction reading and translating single codons into single amino acids.[69] The most extreme variations occur in certain ciliates where the meaning of stop codons depends on their position within mRNA. When close to the 3' end they act as terminators while in internal positions they either code for amino acids as in *Condylostoma magnum*[70] or trigger ribosomal frameshifting as in *Euplotes*.[71]

The origins and variation of the genetic code, including the mechanisms behind the evolvability of the genetic code, have been widely studied,[72][73] and some studies have been done experimentally evolving the genetic code of some organisms.[74][75][76][77]

Inference

Variant genetic codes used by an organism can be inferred by identifying highly conserved genes encoded in that genome, and comparing its codon usage to the amino acids in homologous proteins of other organisms. For example, the program FACIL infers a genetic code by searching which amino acids in homologous protein domains are most often aligned to every codon. The resulting amino acid (or stop codon) probabilities for each codon are displayed in a genetic code

logo.[57]

As of January 2022, the most complete survey of genetic codes is done by Shulgina and Eddy, who screened 250,000 prokaryotic genomes using their Codetta tool. This tool uses a similar approach to FACIL with a larger Pfam database. Despite the NCBI already providing 27 translation tables, the authors were able to find new 5 genetic code variations (corroborated by tRNA mutations) and correct several misattributions.[78] Codetta was later used to analyze genetic code change in ciliates.[79]

Origin

The genetic code is a key part of the history of life, according to one version of which self-replicating RNA molecules preceded life as we know it. This is the RNA world hypothesis. Under this hypothesis, any model for the emergence of the genetic code is intimately related to a model of the transfer from ribozymes (RNA enzymes) to proteins as the principal enzymes in cells. In line with the RNA world hypothesis, transfer RNA molecules appear to have evolved before modern aminoacyl-tRNA synthetases, so the latter cannot be part of the explanation of its patterns.[80]

A hypothetical randomly evolved genetic code further motivates a biochemical or evolutionary model for its origin. If amino acids were randomly assigned to triplet codons, there would be 1.5×10^{84} possible genetic codes.[81]:163 (https://books.google.com/books?id=-YLBMmJE1WwC&pg=PA163) This number is found by calculating the number of ways that 21 items (20 amino acids plus one stop) can be placed in 64 bins, wherein each item is used at least once.[82] However, the distribution of codon assignments in the genetic code is nonrandom.[83] In particular, the genetic code clusters certain amino acid assignments.

Amino acids that share the same biosynthetic pathway tend to have the same first base in their codons. This could be an evolutionary relic of an early, simpler genetic code with fewer amino acids that later evolved to code a larger set of amino acids.[84] It could also reflect steric and chemical properties that had another effect on the codon during its evolution. Amino acids with similar physical properties also tend to have similar codons,[85][86] reducing the problems caused by point mutations and mistranslations.[83]

Given the non-random genetic triplet coding scheme, a tenable hypothesis for the origin of genetic code could address multiple aspects of the codon table, such as absence of codons for D-amino acids, secondary codon patterns for some amino acids, confinement of synonymous positions to third position, the small set of only 20 amino acids (instead of a number approaching 64), and the relation of stop codon patterns to amino acid coding patterns.[87]

Three main hypotheses address the origin of the genetic code. Many models belong to one of them or to a hybrid:[88]

- Random freeze: the genetic code was randomly created. For example, early tRNA-like ribozymes may have had different affinities for amino acids, with codons emerging from another part of the ribozyme that exhibited random variability. Once enough peptides were coded for, any major random change in the genetic code would have been lethal; hence it became "frozen".[89]
- Stereochemical affinity: the genetic code is a result of a high affinity between each amino acid and its codon or anti-codon; the latter option implies that pre-tRNA molecules matched their corresponding amino acids by this affinity. Later during evolution, this matching was gradually replaced with matching by aminoacyl-tRNA synthetases.[87][90][91]
- Optimality: the genetic code continued to evolve after its initial creation, so that the current code maximizes some fitness function, usually some kind of error minimization.[87][88][92]

Hypotheses have addressed a variety of scenarios:[93]

- Chemical principles govern specific RNA interaction with amino acids. Experiments with aptamers showed that some amino acids have a selective chemical affinity for their codons.[94] Experiments showed that of 8 amino acids tested, 6 show some RNA triplet-amino acid association.[81][91]
- Biosynthetic expansion. The genetic code grew from a simpler earlier code through a process of "biosynthetic expansion". Primordial life "discovered" new amino acids (for example, as by-products of metabolism) and later incorporated some of these into the machinery of genetic coding.[95] Although much circumstantial evidence has been found to suggest that fewer amino acid types were used in the past,[96] precise and detailed hypotheses about which amino acids entered the code in what order are controversial.[97][98] However, several studies have suggested that Gly, Ala, Asp, Val, Ser, Pro, Glu, Leu, Thr may belong to a group of early-addition amino acids, whereas Cys, Met, Tyr, Trp, His, Phe may belong to a group of later-addition amino acids.[99][100][101][102]
- Natural selection has led to codon assignments of the genetic code that minimize the effects of mutations.[103] A recent hypothesis[104] suggests that the triplet code was derived from codes that used longer than triplet codons (such as quadruplet codons). Longer than triplet decoding would increase codon redundancy and would be more error resistant. This feature could allow accurate decoding absent complex translational machinery such as the ribosome, such as before cells began making ribosomes.
- Information channels: Information-theoretic approaches model the process of translating the genetic code into corresponding amino acids as an error-prone information channel.[105] The inherent noise (that is, the error) in the channel poses the organism with a fundamental question: how can a genetic code be constructed to withstand noise[106] while accurately and efficiently translating information? These "rate-distortion" models[107] suggest that the genetic code originated as a result of the interplay of the three conflicting evolutionary forces: the needs for diverse amino acids,[108] for error-tolerance[103] and for minimal resource cost. The code emerges at a transition when the mapping of codons to amino acids becomes nonrandom. The code's emergence is governed by the topology defined by the probable errors and is related to the map coloring problem.[109]
- Game theory: Models based on signaling games combine elements of game theory, natural selection and information channels. Such models have been used to suggest that the first polypeptides were likely short and had non-enzymatic function. Game theoretic models suggested that the organization of RNA strings into cells may have been necessary to prevent "deceptive" use of the genetic code, i.e. preventing the ancient equivalent of viruses from overwhelming the RNA world.[110]
- Stop codons: Codons for translational stops are also an interesting aspect to the problem of the origin of the genetic code. As an example for addressing stop codon evolution, it has been suggested that the stop codons are such that they are most likely to terminate translation early in the case of a frame shift error.[111] In contrast, some stereochemical molecular models explain the origin of stop codons as "unassignable".[87]

See also

- List of genetic engineering software

- Codon tables

References

1. Turanov AA, Lobanov AV, Fomenko DE, Morrison HG, Sogin ML, Klobutcher LA, Hatfield DL, Gladyshev VN (January 2009). "Genetic code supports targeted insertion of two amino acids by one codon" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3088105). *Science*. **323** (5911): 259–61. doi:10.1126/science.1164748 (https://doi.org/10.1126%2Fscience.1164748). PMC 3088105 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3088105). PMID 19131629 (https://pubmed.ncbi.nlm.nih.gov/19131629).
2. Watson, J. D.; Crick, F. H. (30 May 1953). "Genetical implications of the structure of deoxyribonucleic acid" (https://pubmed.ncbi.nlm.nih.gov/13063483). *Nature*. **171** (4361): 964–967. Bibcode:1953Natur.171..964W (https://ui.adsabs.harvard.edu/abs/1953Natur.171..964W). doi:10.1038/171964b0 (https://doi.org/10.1038%2F171964b0). ISSN 0028-0836 (https://search.worldcat.org/issn/0028-0836). PMID 13063483 (https://pubmed.ncbi.nlm.nih.gov/13063483). S2CID 4256010 (https://api.semanticscholar.org/CorpusID:4256010).
3. Stegmann, Ulrich E. (1 September 2016). " 'Genetic Coding' Reconsidered: An Analysis of Actual Usage" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4990703). *The British Journal for the Philosophy of Science*. **67** (3): 707–730. doi:10.1093/bjps/axv007 (https://doi.org/10.1093%2Fbjps%2Faxv007). ISSN 0007-0882 (https://search.worldcat.org/issn/0007-0882). PMC 4990703 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4990703). PMID 27924115 (https://pubmed.ncbi.nlm.nih.gov/27924115).
4. Crick, Francis (10 July 1990). "Chapter 8: The Genetic Code" (https://books.google.com/books?id=awoXBQAAQBAJ&pg=PA89). *What Mad Pursuit: A Personal View of Scientific Discovery*. Basic Books. pp. 89–101. ISBN 9780465091386. OCLC 1020240407 (https://search.worldcat.org/oclc/1020240407).
5. Hayes, Brian (1998). "Computing Science: The Invention of the Genetic Code" (https://www.jstor.org/stable/27856930). *American Scientist*. **86** (1): 8–14. doi:10.1511/1998.17.3338 (https://doi.org/10.1511%2F1998.17.3338). ISSN 0003-0996 (https://search.worldcat.org/issn/0003-0996). JSTOR 27856930 (https://www.jstor.org/stable/27856930). S2CID 121907709 (https://api.semanticscholar.org/CorpusID:121907709).
6. Strauss, Bernard S (1 March 2019). "Martynas Yčas: The "Archivist" of the RNA Tie Club" (https://doi.org/10.1534/genetics.118.301754). *Genetics*. **211** (3): 789–795. doi:10.1534/genetics.118.301754 (https://doi.org/10.1534%2Fgenetics.118.301754). ISSN 1943-2631 (https://search.worldcat.org/issn/1943-2631). PMC 6404253 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6404253). PMID 30846543 (https://pubmed.ncbi.nlm.nih.gov/30846543).
7. "Francis Crick - Profiles in Science Search Results" (https://profiles.nlm.nih.gov/spotlight/sc/catalog?f%5breadonly_nlm-id_ssim%5d%5b%5d=101584582X73). *profiles.nlm.nih.gov*. Retrieved 21 July 2022.
8. Fry, Michael (2022). "Crick's Adaptor Hypothesis and the Discovery of Transfer RNA: Experiment Surpassing Theoretical Prediction" (https://journals.publishing.umich.edu/ptpbio/article/id/2628/). *Philosophy, Theory, and Practice in Biology*. **14**. doi:10.3998/ptpbio.2628 (https://doi.org/10.3998%2Fptpbio.2628). ISSN 2475-3025 (https://search.worldcat.org/issn/2475-3025). S2CID 249112573 (https://api.semanticscholar.org/CorpusID:249112573).
9. Crick, Francis (1955). "On Degenerate Templates and the Adaptor Hypothesis: A Note for the RNA Tie Club" (https://collections.nlm.nih.gov/catalog/nlm:nlmuid-101584582X73-doc). *National Library of Medicine*. Retrieved 21 July 2022.
10. Watson, James D. (2007). *Avoid Boring People: Lessons from a Life in Science* (https://books.google.com/books?id=mav7RvFfjDkC). Oxford University Press. p. 112. ISBN 978-0-19-280273-6. OCLC 47716375 (https://search.worldcat.org/oclc/47716375).
11. Barciszewska, Mirosława Z.; Perrigue, Patrick M.; Barciszewski, Jan (2016). "tRNA--the golden standard in molecular biology" (https://pubmed.ncbi.nlm.nih.gov/26549858). *Molecular BioSystems*. **12** (1): 12–17. doi:10.1039/c5mb00557d (https://doi.org/10.1039%2Fc5mb00557d). PMID 26549858 (https://pubmed.ncbi.nlm.nih.gov/26549858).
12. Yanofsky, Charles (9 March 2007). "Establishing the Triplet Nature of the Genetic Code" (https://doi.org/10.1016%2Fj.cell.2007.02.029). *Cell*. **128** (5): 815–818. doi:10.1016/j.cell.2007.02.029 (https://doi.org/10.1016%2Fj.cell.2007.02.029). PMID 17350564 (https://pubmed.ncbi.nlm.nih.gov/17350564). S2CID 14249277 (https://api.semanticscholar.org/CorpusID:14249277).
13. Nirenberg MW, Matthaei JH (October 1961). "The dependence of cell-free protein synthesis in E. coli upon naturally occurring or synthetic polyribonucleotides" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC223178). *Proceedings of the National Academy of Sciences of the United States of America*. **47** (10): 1588–602. Bibcode:1961PNAS...47.1588N (https://ui.adsabs.harvard.edu/abs/1961PNAS...47.1588N). doi:10.1073/pnas.47.10.1588 (https://doi.org/10.1073%2Fpnas.47.10.1588). PMC 223178 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC223178). PMID 14479932 (https://pubmed.ncbi.nlm.nih.gov/14479932).
14. Gardner RS, Wahba AJ, Basilio C, Miller RS, Lengyel P, Speyer JF (December 1962). "Synthetic polynucleotides and the amino acid code. VII" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC221128). *Proceedings of the National Academy of Sciences of the United States of America*. **48** (12): 2087–94. Bibcode:1962PNAS...48.2087G (https://ui.adsabs.harvard.edu/abs/1962PNAS...48.2087G). doi:10.1073/pnas.48.12.2087 (https://doi.org/10.1073%2Fpnas.48.12.2087). PMC 221128 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC221128). PMID 13946552 (https://pubmed.ncbi.nlm.nih.gov/13946552).
15. Wahba AJ, Gardner RS, Basilio C, Miller RS, Speyer JF, Lengyel P (January 1963). "Synthetic polynucleotides and the amino acid code. VIII" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC300638). *Proceedings of the National Academy of Sciences of the United States of America*. **49** (1): 116–22. Bibcode:1963PNAS...49..116W (https://ui.adsabs.harvard.edu/abs/1963PNAS...49..116W). doi:10.1073/pnas.49.1.116 (https://doi.org/10.1073%2Fpnas.49.1.116). PMC 300638 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC300638). PMID 13998282 (https://pubmed.ncbi.nlm.nih.gov/13998282).
16. "The Nobel Prize in Physiology or Medicine 1959" (http://nobelprize.org/nobel_prizes/medicine/laureates/1959/index.html) (Press release). The Royal Swedish Academy of Science. 1959. Retrieved 27 February 2010. "The Nobel Prize in Physiology or Medicine 1959 was awarded jointly to Severo Ochoa and Arthur Kornberg 'for their discovery of the mechanisms in the biological synthesis of ribonucleic acid and deoxyribonucleic acid'."
17. Nirenberg M, Leder P, Bernfield M, Brimacombe R, Trupin J, Rottman F, O'Neal C (May 1965). "RNA codewords and protein synthesis, VII. On the general nature of the RNA code" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC301388). *Proceedings of the National Academy of Sciences of the United States of America*. **53** (5): 1161–8. Bibcode:1965PNAS...53.1161N (https://ui.adsabs.harvard.edu/abs/1965PNAS...53.1161N). doi:10.1073/pnas.53.5.1161 (https://doi.org/10.1073%2Fpnas.53.5.1161). PMC 301388 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC301388). PMID 5330357 (https://pubmed.ncbi.nlm.nih.gov/5330357).
18. "The Nobel Prize in Physiology or Medicine 1968" (http://nobelprize.org/nobel_prizes/medicine/laureates/1968/index.html) (Press release). The Royal Swedish Academy of Science. 1968. Retrieved 27 February 2010. "The Nobel Prize in Physiology or Medicine 1968 was awarded jointly to Robert W. Holley, Har Gobind Khorana and Marshall W. Nirenberg 'for their interpretation of the genetic code and its function in protein synthesis'."
19. Edgar B (October 2004). "The genome of bacteriophage T4: an archeological dig" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1448817). *Genetics*. **168** (2): 575–82. doi:10.1093/genetics/168.2.575 (https://doi.org/10.1093%2Fgenetics%2F168.2.575). PMC 1448817 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1448817). PMID 15514035 (https://pubmed.ncbi.nlm.nih.gov/15514035).
20. Budisa, Nediljko (23 December 2005). *The book at the Wiley Online Library*. doi:10.1002/3527607188 (https://doi.org/10.1002%2F3527607188). ISBN 9783527312436.
21. Kubyshkin, V.; Budisa, N. (2018). "Synthetic alienation of microbial organisms by using genetic code engineering: Why and how?". *Biotechnology Journal*. **12** (8): 16000933. doi:10.1002/biot.201600097 (https://doi.org/10.1002%2Fbiot.201600097). PMID 28671771 (https://pubmed.ncbi.nlm.nih.gov/28671771).

22. Xie J, Schultz PG (December 2005). "Adding amino acids to the genetic repertoire". *Current Opinion in Chemical Biology*. **9** (6): 548–54. doi:10.1016/j.cbpa.2005.10.011 (https://doi.org/10.1016%2Fj.cbpa.2005.10.011). PMID 16260173 (https://pubmed.ncbi.nlm.nih.gov/16260173).
23. Wang Q, Parrish AR, Wang L (March 2009). "Expanding the genetic code for biological studies" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2696486). *Chemistry & Biology*. **16** (3): 323–36. doi:10.1016/j.chembiol.2009.03.001 (https://doi.org/10.1016%2Fj.chembiol.2009.03.001). PMC 2696486 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2696486). PMID 19318213 (https://pubmed.ncbi.nlm.nih.gov/19318213).
24. Simon M (7 January 2005). *Emergent Computation: Emphasizing Bioinformatics* (https://books.google.com/books?id=Uxg51oZNkIsC&pg=PA105). Springer Science & Business Media. pp. 105–106. ISBN 978-0-387-22046-8.
25. Hoesl, M. G.; Oehm, S.; Durkin, P.; Darmon, E.; Peil, L.; Aerni, H.-R.; Rappsilber, J.; Rinehart, J.; Leach, D.; Söll, D.; Budisa, N. (2015). "Chemical evolution of a bacterial proteome" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4782924). *Angewandte Chemie International Edition*. **54** (34): 10030–10034. doi:10.1002/anie.201502868 (https://doi.org/10.1002%2Fanie.201502868). PMC 4782924 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4782924). PMID 26136259 (https://pubmed.ncbi.nlm.nih.gov/26136259). NIHMSID: NIHMS711205
26. "First stable semisynthetic organism created | KurzweilAI" (http://www.kurzweilai.net/first-stable-semisynthetic-organism-created). *www.kurzweilai.net*. 3 February 2017. Retrieved 9 February 2017.
27. Zhang Y, Lamb BM, Feldman AW, Zhou AX, Lavergne T, Li L, Romesberg FE (February 2017). "A semisynthetic organism engineered for the stable expansion of the genetic alphabet" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5307467). *Proceedings of the National Academy of Sciences of the United States of America*. **114** (6): 1317–1322. Bibcode:2017PNAS..114.1317Z (https://ui.adsabs.harvard.edu/abs/2017PNAS..114.1317Z). doi:10.1073/pnas.1616443114 (https://doi.org/10.1073%2Fpnas.1616443114). PMC 5307467 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5307467). PMID 28115716 (https://pubmed.ncbi.nlm.nih.gov/28115716).
28. Han S, Yang A, Lee S, Lee HW, Park CB, Park HS (February 2017). "Expanding the genetic code of Mus musculus" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5321798). *Nature Communications*. **8**: 14568. Bibcode:2017NatCo...814568H (https://ui.adsabs.harvard.edu/abs/2017NatCo...814568H). doi:10.1038/ncomms14568 (https://doi.org/10.1038%2Fncomms14568). PMC 5321798 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5321798). PMID 28220771 (https://pubmed.ncbi.nlm.nih.gov/28220771).
29. Zimmer, Carl (15 May 2019). "Scientists Created Bacteria With a Synthetic Genome. Is This Artificial Life? - In a milestone for synthetic biology, colonies of E. coli thrive with DNA constructed from scratch by humans, not nature" (https://www.nytimes.com/2019/05/15/science/synthetic-genome-bacteria.html). *The New York Times*. Archived (https://ghostarchive.org/archive/20220102/https://www.nytimes.com/2019/05/15/science/synthetic-genome-bacteria.html) from the original on 2 January 2022. Retrieved 16 May 2019.
30. Fredens, Julius; et al. (15 May 2019). "Total synthesis of Escherichia coli with a recoded genome" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7039709). *Nature*. **569** (7757): 514–518. Bibcode:2019Natur.569..514F (https://ui.adsabs.harvard.edu/abs/2019Natur.569..514F). doi:10.1038/s41586-019-1192-5 (https://doi.org/10.1038%2Fs41586-019-1192-5). PMC 7039709 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7039709). PMID 31092918 (https://pubmed.ncbi.nlm.nih.gov/31092918). S2CID 205571025 (https://api.semanticscholar.org/CorpusID:205571025).
31. *Homo sapiens* mitochondrion, complete genome. "Revised Cambridge Reference Sequence (rCRS): accession NC_012920" (https://www.ncbi.nlm.nih.gov/nuccore/NC_012920.1), *National Center for Biotechnology Information*. Retrieved on 27 December 2017.
32. King RC, Mulligan P, Stansfield W (10 January 2013). *A Dictionary of Genetics* (https://books.google.com/books?id=5jhH0HTjEdkC). OUP USA. p. 608. ISBN 978-0-19-976644-4.
33. Touriol C, Bornes S, Bonnal S, Audigier S, Prats H, Prats AC, Vagner S (2003). "Generation of protein isoform diversity by alternative initiation of translation at non-AUG codons" (https://doi.org/10.1016%2FS0248-4900%2803%2900033-9). *Biology of the Cell*. **95** (3–4): 169–78. doi:10.1016/S0248-4900(03)00033-9 (https://doi.org/10.1016%2FS0248-4900%2803%2900033-9). PMID 12867081 (https://pubmed.ncbi.nlm.nih.gov/12867081).
34. Maloy S (29 November 2003). "How nonsense mutations got their names" (http://www.sci.sdsu.edu/~smaloy/MicrobialGenetics/topics/rev-sup/amber-name.html). *Microbial Genetics Course*. San Diego State University. Retrieved 10 March 2010.
35. References for the image are found in Wikimedia Commons page at: Commons:File:Notable mutations.svg#References.
36. Griffiths AJ, Miller JH, Suzuki DT, Lewontin RC, et al., eds. (2000). "Spontaneous mutations" (https://www.ncbi.nlm.nih.gov/books/bv.fcgi?rid=iga.section.2706). *An Introduction to Genetic Analysis* (7th ed.). New York: W. H. Freeman. ISBN 978-0-7167-3520-5.
37. Freisinger E, Grollman AP, Miller H, Kisker C (April 2004). "Lesion (in)tolerance reveals insights into DNA replication fidelity" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC391067). *The EMBO Journal*. **23** (7): 1494–505. doi:10.1038/sj.emboj.7600158 (https://doi.org/10.1038%2Fsj.emboj.7600158). PMC 391067 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC391067). PMID 15057282 (https://pubmed.ncbi.nlm.nih.gov/15057282).
38. Boillée, S; Vande Velde, C; Cleveland, D. W. (2006). "ALS: A disease of motor neurons and their nonneuronal neighbors" (https://doi.org/10.1016%2Fj.neuron.2006.09.018). *Neuron*. **52** (1): 39–59. doi:10.1016/j.neuron.2006.09.018 (https://doi.org/10.1016%2Fj.neuron.2006.09.018). PMID 17015226 (https://pubmed.ncbi.nlm.nih.gov/17015226).
39. Chang JC, Kan YW (June 1979). "beta 0 thalassemia, a nonsense mutation in man" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC383714). *Proceedings of the National Academy of Sciences of the United States of America*. **76** (6): 2886–9. Bibcode:1979PNAS...76.2886C (https://ui.adsabs.harvard.edu/abs/1979PNAS...76.2886C). doi:10.1073/pnas.76.6.2886 (https://doi.org/10.1073%2Fpnas.76.6.2886). PMC 383714 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC383714). PMID 88735 (https://pubmed.ncbi.nlm.nih.gov/88735).
40. Boillée S, Vande Velde C, Cleveland DW (October 2006). "ALS: a disease of motor neurons and their nonneuronal neighbors" (https://doi.org/10.1016%2Fj.neuron.2006.09.018). *Neuron*. **52** (1): 39–59. doi:10.1016/j.neuron.2006.09.018 (https://doi.org/10.1016%2Fj.neuron.2006.09.018). PMID 17015226 (https://pubmed.ncbi.nlm.nih.gov/17015226).
41. Isbrandt D, Hopwood JJ, von Figura K, Peters C (1996). "Two novel frameshift mutations causing premature stop codons in a patient with the severe form of Maroteaux-Lamy syndrome" (https://doi.org/10.1002%2F%28SICI%291098-1004%281996%297%3A4%3C361%3A%3AAID-HUMU12%3E3.0.CO%3B2-0). *Human Mutation*. **7** (4): 361–3. doi:10.1002/(SICI)1098-1004(1996)7:4<361::AID-HUMU12>3.0.CO;2-0 (https://doi.org/10.1002%2F%28SICI%291098-1004%281996%297%3A4%3C361%3A%3AAID-HUMU12%3E3.0.CO%3B2-0). PMID 8723688 (https://pubmed.ncbi.nlm.nih.gov/8723688). S2CID 22693748 (https://api.semanticscholar.org/CorpusID:22693748).
42. Crow JF (1993). "How much do we know about spontaneous human mutation rates?". *Environmental and Molecular Mutagenesis*. **21** (2): 122–9. Bibcode:1993EnvMM..21..122C (https://ui.adsabs.harvard.edu/abs/1993EnvMM..21..122C). doi:10.1002/em.2850210205 (https://doi.org/10.1002%2Fem.2850210205). PMID 8444142 (https://pubmed.ncbi.nlm.nih.gov/8444142). S2CID 32918971 (https://api.semanticscholar.org/CorpusID:32918971).
43. Lewis R (2005). *Human Genetics: Concepts and Applications* (6th ed.). Boston, Mass: McGraw Hill. pp. 227–228. ISBN 978-0-07-111156-0.
44. Sawyer SA, Parsch J, Zhang Z, Hartl DL (April 2007). "Prevalence of positive selection among nearly neutral amino acid replacements in Drosophila" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1871816). *Proceedings of the National Academy of Sciences of the United States of America*. **104** (16): 6504–10. Bibcode:2007PNAS..104.6504S (https://ui.adsabs.harvard.edu/abs/2007PNAS..104.6504S). doi:10.1073/pnas.0701572104 (https://doi.org/10.1073%2Fpnas.0701572104). PMC 1871816 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1871816). PMID 17409186 (https://pubmed.ncbi.nlm.nih.gov/17409186).
45. Bridges KR (2002). "Malaria and the Red Cell" (https://web.archive.org/web/20111127201806/http://sickle.bwh.harvard.edu/malaria_sickle.html). *Harvard*. Archived from the original (http://sickle.bwh.harvard.edu/malaria_sickle.html) on 27 November 2011.

46. Drake JW, Holland JJ (November 1999). "Mutation rates among RNA viruses" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC24164). *Proceedings of the National Academy of Sciences of the United States of America*. **96** (24): 13910–3. Bibcode:1999PNAS...9613910D (https://ui.adsabs.harvard.edu/abs/1999PNAS...9613910D). doi:10.1073/pnas.96.24.13910 (https://doi.org/10.1073%2Fpnas.96.24.13910). PMC 24164 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC24164). PMID 10570172 (https://pubmed.ncbi.nlm.nih.gov/10570172).
47. Holland J, Spindler K, Horodyski F, Grabau E, Nichol S, VandePol S (March 1982). "Rapid evolution of RNA genomes". *Science*. **215** (4540): 1577–85. Bibcode:1982Sci...215.1577H (https://ui.adsabs.harvard.edu/abs/1982Sci...215.1577H). doi:10.1126/science.7041255 (https://doi.org/10.1126%2Fscience.7041255). PMID 7041255 (https://pubmed.ncbi.nlm.nih.gov/7041255).
48. de Visser JA, Rozen DE (April 2006). "Clonal interference and the periodic selection of new beneficial mutations in Escherichia coli" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1456385). *Genetics*. **172** (4): 2093–100. doi:10.1534/genetics.105.052373 (https://doi.org/10.1534%2Fgenetics.105.052373). PMC 1456385 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1456385). PMID 16489229 (https://pubmed.ncbi.nlm.nih.gov/16489229).
49. Watson, James D. (2008). *Molecular Biology of the Gene* (https://books.google.com/books?id=MByWPwAACAAJ). Pearson/Benjamin Cummings. ISBN 978-0-8053-9592-1. :102–117 (https://books.google.com/books?id=MByWPwAACAAJ&pg=PA102) :521–522 (https://books.google.com/books?id=MByWPwAACAAJ&pg=PA521)
50. Michel-Beyerle, Maria Elisabeth (1990). *Reaction centers of photosynthetic bacteria: Feldafing-II-Meeting* (https://books.google.com/books?id=xD5OAQAAIAAJ). Springer-Verlag. ISBN 978-3-540-53420-4.
51. Füllen G, Youvan DC (1994). "Genetic Algorithms and Recursive Ensemble Mutagenesis in Protein Engineering". Complexity International 1.
52. Fricke, Markus (2019). "Global importance of RNA secondary structures in protein coding sequences" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7109657). *Bioinformatics*. **35** (4): 579–583. doi:10.1093/bioinformatics/bty678 (https://doi.org/10.1093%2Fbioinformatics%2Fbty678). PMC 7109657 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7109657). PMID 30101307 (https://pubmed.ncbi.nlm.nih.gov/30101307). S2CID 51968530 (https://api.semanticscholar.org/CorpusID:51968530).
53. "Codon Usage Frequency Table(chart)-Genscript" (https://www.genscript.com/tools/codon-frequency-table). *www.genscript.com*. Retrieved 4 February 2022.
54. "Codon usage table" (http://www.kazusa.or.jp/codon/cgi-bin/showcodon.cgi?species=9606&aa=1&style=N). *www.kazusa.or.jp*.
55. Zhang Y, Baranov PV, Atkins JF, Gladyshev VN (May 2005). "Pyrrolysine and selenocysteine use dissimilar decoding strategies" (http://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=1071&context=biochemgladyshev). *The Journal of Biological Chemistry*. **280** (21): 20740–51. doi:10.1074/jbc.M501458200 (https://doi.org/10.1074%2Fjbc.M501458200). PMID 15788401 (https://pubmed.ncbi.nlm.nih.gov/15788401).
56. Prat L, Heinemann IU, Aerni HR, Rinehart J, O'Donoghue P, Söll D (December 2012). "Carbon source-dependent expansion of the genetic code in bacteria" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3529041). *Proceedings of the National Academy of Sciences of the United States of America*. **109** (51): 21070–5. Bibcode:2012PNAS..10921070P (https://ui.adsabs.harvard.edu/abs/2012PNAS..10921070P). doi:10.1073/pnas.1218613110 (https://doi.org/10.1073%2Fpnas.1218613110). PMC 3529041 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3529041). PMID 23185002 (https://pubmed.ncbi.nlm.nih.gov/23185002).
57. Dutilh BE, Jurgelenaite R, Szklarczyk R, van Hijum SA, Harhangi HR, Schmid M, de Wild B, Françoijs KJ, Stunnenberg HG, Strous M, Jetten MS, Op den Camp HJ, Huynen MA (July 2011). "FACIL: Fast and Accurate Genetic Code Inference and Logo" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3129529). *Bioinformatics*. **27** (14): 1929–33. doi:10.1093/bioinformatics/btr316 (https://doi.org/10.1093%2Fbioinformatics%2Fbtr316). PMC 3129529 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3129529). PMID 21653513 (https://pubmed.ncbi.nlm.nih.gov/21653513).
58. Crick, F.H.C. (28 December 1968). "The origin of the genetic code" (https://linkinghub.elsevier.com/retrieve/pii/0022283668903926). *Journal of Molecular Biology*. **38** (3): 367–379. doi:10.1016/0022-2836(68)90392-6 (https://doi.org/10.1016%2F0022-2836%2868%2990392-6).
59. Barrell BG, Bankier AT, Drouin J (1979). "A different genetic code in human mitochondria". *Nature*. **282** (5735): 189–194. Bibcode:1979Natur.282..189B (https://ui.adsabs.harvard.edu/abs/1979Natur.282..189B). doi:10.1038/282189a0 (https://doi.org/10.1038%2F282189a0). PMID 226894 (https://pubmed.ncbi.nlm.nih.gov/226894). S2CID 4335828 (https://api.semanticscholar.org/CorpusID:4335828). ([1] (https://www.ncbi.nlm.nih.gov/pubmed/226894))
60. Elzanowski A, Ostell J (7 April 2008). "The Genetic Codes" (https://www.ncbi.nlm.nih.gov/Taxonomy/Utils/wprintgc.cgi?mode=c). National Center for Biotechnology Information (NCBI). Retrieved 10 March 2010.
61. Jukes TH, Osawa S (December 1990). "The genetic code in mitochondria and chloroplasts". *Experientia*. **46** (11–12): 1117–26. doi:10.1007/BF01936921 (https://doi.org/10.1007%2FBF01936921). PMID 2253709 (https://pubmed.ncbi.nlm.nih.gov/2253709). S2CID 19264964 (https://api.semanticscholar.org/CorpusID:19264964).
62. Fitzpatrick DA, Logue ME, Stajich JE, Butler G (1 January 2006). "A fungal phylogeny based on 42 complete genomes derived from supertree and combined gene analysis" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1679813). *BMC Evolutionary Biology*. **6**: 99. doi:10.1186/1471-2148-6-99 (https://doi.org/10.1186%2F1471-2148-6-99). PMC 1679813 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1679813). PMID 17121679 (https://pubmed.ncbi.nlm.nih.gov/17121679).
63. Santos MA, Tuite MF (May 1995). "The CUG codon is decoded in vivo as serine and not leucine in Candida albicans" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC306886). *Nucleic Acids Research*. **23** (9): 1481–6. doi:10.1093/nar/23.9.1481 (https://doi.org/10.1093%2Fnar%2F23.9.1481). PMC 306886 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC306886). PMID 7784200 (https://pubmed.ncbi.nlm.nih.gov/7784200).
64. Butler G, Rasmussen MD, Lin MF, et al. (June 2009). "Evolution of pathogenicity and sexual reproduction in eight Candida genomes" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2834264). *Nature*. **459** (7247): 657–62. Bibcode:2009Natur.459..657B (https://ui.adsabs.harvard.edu/abs/2009Natur.459..657B). doi:10.1038/nature08064 (https://doi.org/10.1038%2Fnature08064). PMC 2834264 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2834264). PMID 19465905 (https://pubmed.ncbi.nlm.nih.gov/19465905).
65. Taylor DJ, Ballinger MJ, Bowman SM, Bruenn JA (2013). "Virus-host co-evolution under a modified nuclear genetic code" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3628385). *PeerJ*. **1**: e50. doi:10.7717/peerj.50 (https://doi.org/10.7717%2Fpeerj.50). PMC 3628385 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3628385). PMID 23638388 (https://pubmed.ncbi.nlm.nih.gov/23638388).
66. Hofhuis J, Schueren F, Nötzel C, Lingner T, Gärtner J, Jahn O, Thoms S (2016). "The functional readthrough extension of malate dehydrogenase reveals a modification of the genetic code" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5133446). *Open Biol*. **6** (11): 160246. doi:10.1098/rsob.160246 (https://doi.org/10.1098%2Frsob.160246). PMC 5133446 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5133446). PMID 27881739 (https://pubmed.ncbi.nlm.nih.gov/27881739).
67. Schueren F, Lingner T, George R, Hofhuis J, Gartner J, Thoms S (2014). "Peroxisomal lactate dehydrogenase is generated by translational readthrough in mammals" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4359377). *eLife*. **3**: e03640. doi:10.7554/eLife.03640 (https://doi.org/10.7554%2FeLife.03640). PMC 4359377 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4359377). PMID 25247702 (https://pubmed.ncbi.nlm.nih.gov/25247702).
68. F. Schueren und S. Thoms (2016). "Functional Translational Readthrough: A Systems Biology Perspective" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4973966). *PLOS Genetics*. **12** (8): e1006196. doi:10.1371/journal.pgen.1006196 (https://doi.org/10.1371%2Fjournal.pgen.1006196). PMC 4973966 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4973966). PMID 27490485 (https://pubmed.ncbi.nlm.nih.gov/27490485).
69. Kubyshkin V, Acevedo-Rocha CG, Budisa N (February 2018). "On universal coding events in protein biogenesis" (https://doi.org/10.1016%2Fj.biosystems.2017.10.004). *Bio Systems*. **164**: 16–25. Bibcode:2018BiSys.164...16K (https://ui.adsabs.harvard.edu/abs/2018BiSys.164...16K). doi:10.1016/j.biosystems.2017.10.004 (https://doi.org/10.1016%2Fj.biosystems.2017.10.004). PMID 29030023 (https://pubmed.ncbi.nlm.nih.gov/29030023).

70. Heaphy SM, Mariotti M, Gladyshev VN, Atkins JF, Baranov PV (November 2016). "Novel Ciliate Genetic Code Variants Including the Reassignment of All Three Stop Codons to Sense Codons in *Condylostoma magnum*" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5062323). *Molecular Biology and Evolution*. **33** (11): 2885–2889. doi:10.1093/molbev/msw166 (https://doi.org/10.1093%2Fmolbev%2Fmsw166). PMC 5062323 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5062323). PMID 27501944 (https://pubmed.ncbi.nlm.nih.gov/27501944).
71. Lobanov AV, Heaphy SM, Turanov AA, Gerashchenko MV, Pucciarelli S, Devaraj RR, et al. (January 2017). "Position-dependent termination and widespread obligatory frameshifting in *Euplotes* translation" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5295771). *Nature Structural & Molecular Biology*. **24** (1): 61–68. doi:10.1038/nsmb.3330 (https://doi.org/10.1038%2Fnsmb.3330). PMC 5295771 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5295771). PMID 27870834 (https://pubmed.ncbi.nlm.nih.gov/27870834).
72. Koonin EV, Novozhilov AS (February 2009). "Origin and Evolution of the Genetic Code: The Universal Enigma" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3293468). *IUBMB Life*. **61** (2): 91–111. doi:10.1002/iub.146 (https://doi.org/10.1002%2Fiub.146). PMC 3293468 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3293468). PMID 19117371 (https://pubmed.ncbi.nlm.nih.gov/19117371).
73. Sengupta S, Higgs PG (June 2015). "Pathways of Genetic Code Evolution in Ancient and Modern Organisms". *Journal of Molecular Evolution*. **80** (5–6): 229–243. Bibcode:2015JMolE..80..229S (https://ui.adsabs.harvard.edu/abs/2015JMolE..80..229S). doi:10.1007/s00239-015-9686-8 (https://doi.org/10.1007%2Fs00239-015-9686-8). PMID 26054480 (https://pubmed.ncbi.nlm.nih.gov/26054480). S2CID 15542587 (https://api.semanticscholar.org/CorpusID:15542587).
74. Xie J, Schultz PG (August 2006). "A chemical toolkit for proteins--an expanded genetic code". *Nature Reviews Molecular Cell Biology*. **7** (10): 775–782. doi:10.1038/nrm2005 (https://doi.org/10.1038%2Fnrm2005). PMID 16926858 (https://pubmed.ncbi.nlm.nih.gov/16926858). S2CID 19385756 (https://api.semanticscholar.org/CorpusID:19385756).
75. Neumann H, Wang K, Davis L, Garcia-Alai M, Chin JW (March 2010). "Encoding multiple unnatural amino acids via evolution of a quadruplet-decoding ribosome". *Nature*. **18** (464): 441–444. doi:10.1038/nrm2005 (https://doi.org/10.1038%2Fnrm2005). PMID 16926858 (https://pubmed.ncbi.nlm.nih.gov/16926858). S2CID 19385756 (https://api.semanticscholar.org/CorpusID:19385756).
76. Liu CC, Schultz PG (2010). "Adding new chemistries to the genetic code". *Annual Review of Biochemistry*. **79**: 413–444. doi:10.1146/annurev.biochem.052308.105824 (https://doi.org/10.1146%2Fannurev.biochem.052308.105824). PMID 20307192 (https://pubmed.ncbi.nlm.nih.gov/20307192).
77. Chin JW (February 2014). "Expanding and reprogramming the genetic code of cells and animals". *Annual Review of Biochemistry*. **83**: 379–408. doi:10.1146/annurev-biochem-060713-035737 (https://doi.org/10.1146%2Fannurev-biochem-060713-035737). PMID 24555827 (https://pubmed.ncbi.nlm.nih.gov/24555827).
78. Shulgina, Y; Eddy, SR (9 November 2021). "A computational screen for alternative genetic codes in over 250,000 genomes" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8629427). *eLife*. **10**. doi:10.7554/eLife.71402 (https://doi.org/10.7554%2FeLife.71402). PMC 8629427 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8629427). PMID 34751130 (https://pubmed.ncbi.nlm.nih.gov/34751130).
79. Chen, W; Geng, Y; Zhang, B; Yan, Y; Zhao, F; Miao, M (4 April 2023). "Stop or Not: Genome-Wide Profiling of Reassigned Stop Codons in Ciliates" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1008964). *Molecular Biology and Evolution*. **40** (4). doi:10.1093/molbev/msad064 (https://doi.org/10.1093%2Fmolbev%2Fmsad064). PMC 1008964 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1008964). PMID 36952281 (https://pubmed.ncbi.nlm.nih.gov/36952281).
80. Ribas de Pouplana L, Turner RJ, Steer BA, Schimmel P (September 1998). "Genetic code origins: tRNAs older than their synthetases?" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC21636). *Proceedings of the National Academy of Sciences of the United States of America*. **95** (19): 11295–300. Bibcode:1998PNAS...9511295D (https://ui.adsabs.harvard.edu/abs/1998PNAS...9511295D). doi:10.1073/pnas.95.19.11295 (https://doi.org/10.1073%2Fpnas.95.19.11295). PMC 21636 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC21636). PMID 9736730 (https://pubmed.ncbi.nlm.nih.gov/9736730).
81. Yarus, Michael (2010). *Life from an RNA World: The Ancestor Within* (https://books.google.com/books?id=-YLBMmJE1WwC). Harvard University Press. ISBN 978-0-674-05075-4.
82. "Mathematica function for # possible arrangements of items in bins? – Online Technical Discussion Groups—Wolfram Community" (http://community.wolfram.com/groups/-/m/t/319970). *community.wolfram.com*. Retrieved 3 February 2017.
83. Freeland SJ, Hurst LD (September 1998). "The genetic code is one in a million". *Journal of Molecular Evolution*. **47** (3): 238–48. Bibcode:1998JMolE..47..238F (https://ui.adsabs.harvard.edu/abs/1998JMolE..47..238F). doi:10.1007/PL00006381 (https://doi.org/10.1007%2FPL00006381). PMID 9732450 (https://pubmed.ncbi.nlm.nih.gov/9732450). S2CID 20130470 (https://api.semanticscholar.org/CorpusID:20130470).
84. Taylor FJ, Coates D (1989). "The code within the codons". *Bio Systems*. **22** (3): 177–87. Bibcode:1989BiSys..22..177T (https://ui.adsabs.harvard.edu/abs/1989BiSys..22..177T). doi:10.1016/0303-2647(89)90059-2 (https://doi.org/10.1016%2F0303-2647%2889%2990059-2). PMID 2650752 (https://pubmed.ncbi.nlm.nih.gov/2650752).
85. Di Giulio M (October 1989). "The extension reached by the minimization of the polarity distances during the evolution of the genetic code". *Journal of Molecular Evolution*. **29** (4): 288–93. Bibcode:1989JMolE..29..288D (https://ui.adsabs.harvard.edu/abs/1989JMolE..29..288D). doi:10.1007/BF02103616 (https://doi.org/10.1007%2FBF02103616). PMID 2514270 (https://pubmed.ncbi.nlm.nih.gov/2514270). S2CID 20803686 (https://api.semanticscholar.org/CorpusID:20803686).
86. Wong JT (February 1980). "Role of minimization of chemical distances between amino acids in the evolution of the genetic code" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC348428). *Proceedings of the National Academy of Sciences of the United States of America*. **77** (2): 1083–6. Bibcode:1980PNAS...77.1083W (https://ui.adsabs.harvard.edu/abs/1980PNAS...77.1083W). doi:10.1073/pnas.77.2.1083 (https://doi.org/10.1073%2Fpnas.77.2.1083). PMC 348428 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC348428). PMID 6928661 (https://pubmed.ncbi.nlm.nih.gov/6928661).
87. Erives A (August 2011). "A model of proto-anti-codon RNA enzymes requiring L-amino acid homochirality" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3223571). *Journal of Molecular Evolution*. **73** (1–2): 10–22. Bibcode:2011JMolE..73...10E (https://ui.adsabs.harvard.edu/abs/2011JMolE..73...10E). doi:10.1007/s00239-011-9453-4 (https://doi.org/10.1007%2Fs00239-011-9453-4). PMC 3223571 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3223571). PMID 21779963 (https://pubmed.ncbi.nlm.nih.gov/21779963).
88. Freeland SJ, Knight RD, Landweber LF, Hurst LD (April 2000). "Early fixation of an optimal genetic code" (https://doi.org/10.1093%2Foxfordjournals.molbev.a026331). *Molecular Biology and Evolution*. **17** (4): 511–18. doi:10.1093/oxfordjournals.molbev.a026331 (https://doi.org/10.1093%2Foxfordjournals.molbev.a026331). PMID 10742043 (https://pubmed.ncbi.nlm.nih.gov/10742043).
89. Crick FH (December 1968). "The origin of the genetic code". *Journal of Molecular Evolution*. **38** (3): 367–79. doi:10.1016/0022-2836(68)90392-6 (https://doi.org/10.1016%2F0022-2836%2868%2990392-6). PMID 4887876 (https://pubmed.ncbi.nlm.nih.gov/4887876). S2CID 4144681 (https://api.semanticscholar.org/CorpusID:4144681).
90. Hopfield JJ (1978). "Origin of the genetic code: a testable hypothesis based on tRNA structure, sequence, and kinetic proofreading" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC336109). *PNAS*. **75** (9): 4334–4338. Bibcode:1978PNAS...75.4334H (https://ui.adsabs.harvard.edu/abs/1978PNAS...75.4334H). doi:10.1073/pnas.75.9.4334 (https://doi.org/10.1073%2Fpnas.75.9.4334). PMC 336109 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC336109). PMID 279919 (https://pubmed.ncbi.nlm.nih.gov/279919).
91. Yarus M, Widmann JJ, Knight R (November 2009). "RNA-amino acid binding: a stereochemical era for the genetic code" (https://doi.org/10.1007%2Fs00239-009-9270-1). *Journal of Molecular Evolution*. **69** (5): 406–29. Bibcode:2009JMolE..69..406Y (https://ui.adsabs.harvard.edu/abs/2009JMolE..69..406Y). doi:10.1007/s00239-009-9270-1 (https://doi.org/10.1007%2Fs00239-009-9270-1). PMID 19795157 (https://pubmed.ncbi.nlm.nih.gov/19795157).
92. Brown, Sean M.; Voráček, Václav; Freeland, Stephen (5 April 2023). "What Would an Alien Amino Acid Alphabet Look Like and Why?". *Astrobiology*. **23** (5): 536–549. Bibcode:2023AsBio..23..536B (https://ui.adsabs.harvard.edu/abs/2023AsBio..23..536B). doi:10.1089/ast.2022.0107 (https://doi.org/10.1089%2Fast.2022.0107). PMID 37022727 (https://pubmed.ncbi.nlm.nih.gov/37022727). S2CID 257983174 (https://api.semanticscholar.org/CorpusID:257983174).

93. Knight RD, Freeland SJ, Landweber LF (June 1999). "Selection, history and chemistry: the three faces of the genetic code" (https://www.sciencedirect.com/science/article/abs/pii/S0968000499013924). *Trends in Biochemical Sciences*. **24** (6): 241–7. doi:10.1016/S0968-0004(99)01392-4 (https://doi.org/10.1016%2FS0968-0004%2899%2901392-4). PMID 10366854 (https://pubmed.ncbi.nlm.nih.gov/10366854).
94. Knight RD, Landweber LF (September 1998). "Rhyme or reason: RNA-arginine interactions and the genetic code" (https://doi.org/10.1016%2FS1074-5521%2898%2990001-1). *Chemistry & Biology*. **5** (9): R215–20. doi:10.1016/S1074-5521(98)90001-1 (https://doi.org/10.1016%2FS1074-5521%2898%2990001-1). PMID 9751648 (https://pubmed.ncbi.nlm.nih.gov/9751648).
95. Sengupta S, Higgs PG (2015). "Pathways of genetic code evolution in ancient and modern organisms". *Journal of Molecular Evolution*. **80** (5–6): 229–243. Bibcode:2015JMolE..80..229S (https://ui.adsabs.harvard.edu/abs/2015JMolE..80..229S). doi:10.1007/s00239-015-9686-8 (https://doi.org/10.1007%2Fs00239-015-9686-8). PMID 26054480 (https://pubmed.ncbi.nlm.nih.gov/26054480). S2CID 15542587 (https://api.semanticscholar.org/CorpusID:15542587).
96. Brooks DJ, Fresco JR, Lesk AM, Singh M (October 2002). "Evolution of amino acid frequencies in proteins over deep time: inferred order of introduction of amino acids into the genetic code" (https://doi.org/10.1093%2Foxfordjournals.molbev.a003988). *Molecular Biology and Evolution*. **19** (10): 1645–55. doi:10.1093/oxfordjournals.molbev.a003988 (https://doi.org/10.1093%2Foxfordjournals.molbev.a003988). PMID 12270892 (https://pubmed.ncbi.nlm.nih.gov/12270892).
97. Amirnovin R (May 1997). "An analysis of the metabolic theory of the origin of the genetic code". *Journal of Molecular Evolution*. **44** (5): 473–6. Bibcode:1997JMolE..44..473A (https://ui.adsabs.harvard.edu/abs/1997JMolE..44..473A). doi:10.1007/PL00006170 (https://doi.org/10.1007%2FPL00006170). PMID 9115171 (https://pubmed.ncbi.nlm.nih.gov/9115171). S2CID 23334860 (https://api.semanticscholar.org/CorpusID:23334860).
98. Ronneberg TA, Landweber LF, Freeland SJ (December 2000). "Testing a biosynthetic theory of the genetic code: fact or artifact?" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC17637). *Proceedings of the National Academy of Sciences of the United States of America*. **97** (25): 13690–5. Bibcode:2000PNAS...9713690R (https://ui.adsabs.harvard.edu/abs/2000PNAS...9713690R). doi:10.1073/pnas.250403097 (https://doi.org/10.1073%2Fpnas.250403097). PMC 17637 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC17637). PMID 11087835 (https://pubmed.ncbi.nlm.nih.gov/11087835).
99. Trifonov, Edward N. (September 2009). "The origin of the genetic code and of the earliest oligopeptides" (https://linkinghub.elsevier.com/retrieve/pii/S0923250809000576). *Research in Microbiology*. **160** (7): 481–486. doi:10.1016/j.resmic.2009.05.004 (https://doi.org/10.1016%2Fj.resmic.2009.05.004). PMID 19524038 (https://pubmed.ncbi.nlm.nih.gov/19524038).
100. Higgs, Paul G.; Pudritz, Ralph E. (June 2009). "A Thermodynamic Basis for Prebiotic Amino Acid Synthesis and the Nature of the First Genetic Code" (http://www.liebertpub.com/doi/10.1089/ast.2008.0280). *Astrobiology*. **9** (5): 483–490. arXiv:0904.0402 (https://arxiv.org/abs/0904.0402). Bibcode:2009AsBio...9..483H (https://ui.adsabs.harvard.edu/abs/2009AsBio...9..483H). doi:10.1089/ast.2008.0280 (https://doi.org/10.1089%2Fast.2008.0280). ISSN 1531-1074 (https://search.worldcat.org/issn/1531-1074). PMID 19566427 (https://pubmed.ncbi.nlm.nih.gov/19566427). S2CID 9039622 (https://api.semanticscholar.org/CorpusID:9039622).
101. Chaliotis, Anargyros; Vlastaridis, Panayotis; Mossialos, Dimitris; Ibba, Michael; Becker, Hubert D.; Stathopoulos, Constantinos; Amoutzias, Grigorios D. (17 February 2017). "The complex evolutionary history of aminoacyl-tRNA synthetases" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5388404). *Nucleic Acids Research*. **45** (3): 1059–1068. doi:10.1093/nar/gkw1182 (https://doi.org/10.1093%2Fnar%2Fgkw1182). ISSN 0305-1048 (https://search.worldcat.org/issn/0305-1048). PMC 5388404 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5388404). PMID 28180287 (https://pubmed.ncbi.nlm.nih.gov/28180287).
102. Ntountoumi, Chrysa; Vlastaridis, Panayotis; Mossialos, Dimitris; Stathopoulos, Constantinos; Iliopoulos, Ioannis; Promponas, Vasilios; Oliver, Stephen G; Amoutzias, Grigoris D (4 November 2019). "Low complexity regions in the proteins of prokaryotes perform important functional roles and are highly conserved" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6821194). *Nucleic Acids Research*. **47** (19): 9998–10009. doi:10.1093/nar/gkz730 (https://doi.org/10.1093%2Fnar%2Fgkz730). ISSN 0305-1048 (https://search.worldcat.org/issn/0305-1048). PMC 6821194 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6821194). PMID 31504783 (https://pubmed.ncbi.nlm.nih.gov/31504783).
103. Freeland SJ, Wu T, Keulmann N (October 2003). "The case for an error minimizing standard genetic code". *Origins of Life and Evolution of the Biosphere*. **33** (4–5): 457–77. Bibcode:2003OLEB...33..457F (https://ui.adsabs.harvard.edu/abs/2003OLEB...33..457F). doi:10.1023/A:1025771327614 (https://doi.org/10.1023%2FA%3A1025771327614). PMID 14604186 (https://pubmed.ncbi.nlm.nih.gov/14604186). S2CID 18823745 (https://api.semanticscholar.org/CorpusID:18823745).
104. Baranov PV, Venin M, Provan G (2009). Gemmell NJ (ed.). "Codon size reduction as the origin of the triplet genetic code" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2682656). *PLOS ONE*. **4** (5): e5708. Bibcode:2009PLoSO...4.5708B (https://ui.adsabs.harvard.edu/abs/2009PLoSO...4.5708B). doi:10.1371/journal.pone.0005708 (https://doi.org/10.1371%2Fjournal.pone.0005708). PMC 2682656 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2682656). PMID 19479032 (https://pubmed.ncbi.nlm.nih.gov/19479032).
105. Tlusty T (November 2007). "A model for the emergence of the genetic code as a transition in a noisy information channel". *Journal of Theoretical Biology*. **249** (2): 331–42. arXiv:1007.4122 (https://arxiv.org/abs/1007.4122). Bibcode:2007JThBi.249..331T (https://ui.adsabs.harvard.edu/abs/2007JThBi.249..331T). doi:10.1016/j.jtbi.2007.07.029 (https://doi.org/10.1016%2Fj.jtbi.2007.07.029). PMID 17826800 (https://pubmed.ncbi.nlm.nih.gov/17826800). S2CID 12206140 (https://api.semanticscholar.org/CorpusID:12206140).
106. Sonneborn TM (1965). Bryson V, Vogel H (eds.). *Evolving genes and proteins*. New York: Academic Press. pp. 377–397.
107. Tlusty T (February 2008). "Rate-distortion scenario for the emergence and evolution of noisy molecular codes". *Physical Review Letters*. **100** (4): 048101. arXiv:1007.4149 (https://arxiv.org/abs/1007.4149). Bibcode:2008PhRvL.100d8101T (https://ui.adsabs.harvard.edu/abs/2008PhRvL.100d8101T). doi:10.1103/PhysRevLett.100.048101 (https://doi.org/10.1103%2FPhysRevLett.100.048101). PMID 18352335 (https://pubmed.ncbi.nlm.nih.gov/18352335). S2CID 12246664 (https://api.semanticscholar.org/CorpusID:12246664).
108. Sella G, Ardell DH (September 2006). "The coevolution of genes and genetic codes: Crick's frozen accident revisited". *Journal of Molecular Evolution*. **63** (3): 297–313. Bibcode:2006JMolE..63..297S (https://ui.adsabs.harvard.edu/abs/2006JMolE..63..297S). doi:10.1007/s00239-004-0176-7 (https://doi.org/10.1007%2Fs00239-004-0176-7). PMID 16838217 (https://pubmed.ncbi.nlm.nih.gov/16838217). S2CID 1260806 (https://api.semanticscholar.org/CorpusID:1260806).
109. Tlusty T (September 2010). "A colorful origin for the genetic code: information theory, statistical mechanics and the emergence of molecular codes". *Physics of Life Reviews*. **7** (3): 362–76. arXiv:1007.3906 (https://arxiv.org/abs/1007.3906). Bibcode:2010PhLRv...7..362T (https://ui.adsabs.harvard.edu/abs/2010PhLRv...7..362T). doi:10.1016/j.plrev.2010.06.002 (https://doi.org/10.1016%2Fj.plrev.2010.06.002). PMID 20558115 (https://pubmed.ncbi.nlm.nih.gov/20558115). S2CID 1845965 (https://api.semanticscholar.org/CorpusID:1845965).
110. Jee J, Sundstrom A, Massey SE, Mishra B (November 2013). "What can information-asymmetric games tell us about the context of Crick's 'frozen accident'?" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3785830). *Journal of the Royal Society, Interface*. **10** (88): 20130614. doi:10.1098/rsif.2013.0614 (https://doi.org/10.1098%2Frsif.2013.0614). PMC 3785830 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3785830). PMID 23985735 (https://pubmed.ncbi.nlm.nih.gov/23985735).
111. Itzkovitz S, Alon U (2007). "The genetic code is nearly optimal for allowing additional information within protein-coding sequences" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1832087). *Genome Research*. **17** (4): 405–412. doi:10.1101/gr.5987307 (https://doi.org/10.1101%2Fgr.5987307). PMC 1832087 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1832087). PMID 17293451 (https://pubmed.ncbi.nlm.nih.gov/17293451).

Further reading

- Griffiths AJ, Miller JH, Suzuki DT, Lewontin RC, Gilbert WM (1999). *An Introduction to genetic analysis* (https://www.ncbi.nlm.nih.gov/books/bv.fcgi?call=bv.View..ShowTOC&rid=iga.TOC) (7th ed.). San Francisco: W.H. Freeman. ISBN 978-0-7167-3771-1.
- Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P (2002). *Molecular biology of the cell* (https://www.ncbi.nlm.nih.gov/books/bv.fcgi?call=bv.View..ShowTOC&rid=mboc4.TOC&depth=2) (4th ed.). New York: Garland Science. ISBN 978-0-8153-3218-3.
- Lodish HF, Berk A, Zipursky SL, Matsudaira P, Baltimore D, Darnell JE (2000). *Molecular cell biology* (https://archive.org/details/molecularcellbi000lodi) (4th ed.). San Francisco: W.H. Freeman. ISBN 9780716737063.
- Caskey CT, Leder P (April 2014). "The RNA code: nature's Rosetta Stone" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4000803). *Proceedings of the National Academy of Sciences of the United States of America*. **111** (16): 5758–9. Bibcode:2014PNAS..111.5758C (https://ui.adsabs.harvard.edu/abs/2014PNAS..111.5758C). doi:10.1073/pnas.1404819111 (https://doi.org/10.1073%2Fpnas.1404819111). PMC 4000803 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4000803). PMID 24756939 (https://pubmed.ncbi.nlm.nih.gov/24756939).

External links

- The Genetic Codes: Genetic Code Tables (https://www.ncbi.nlm.nih.gov/Taxonomy/taxonomyhome.html/index.cgi?chapter=cgencodes)
- The Codon Usage Database (http://www.kazusa.or.jp/codon/) — Codon frequency tables for many organisms
- History of deciphering the genetic code (http://history.nih.gov/exhibits/nirenberg/)