Augmented genetic decoding: Global, local and temporal alterations of decoding processes and codon meaning

Article *in* Nature Reviews Genetics · August 2015

DOI: 10.1038/nrg3963 · Source: PubMed

3 authors, including:

Pavel V Baranov
University College Cork
144 PUBLICATIONS **3,779** CITATIONS

SEE PROFILE

Martina Yordanova
University College Cork
6 PUBLICATIONS **115** CITATIONS

SEE PROFILE

Augmented genetic decoding: global, local and temporal alterations of decoding processes and codon meaning

Pavel V. Baranov[1], John F. Atkins[1,2] and Martina M. Yordanova[1]

Abstract | The non-universality of the genetic code is now widely appreciated. Codes differ between organisms, and certain genes are known to alter the decoding rules in a site-specific manner. Recently discovered examples of decoding plasticity are particularly spectacular. These examples include organisms and organelles with disruptions of triplet continuity during the translation of many genes, viruses that alter the entire genetic code of their hosts and organisms that adjust their genetic code in response to changing environments. In this Review, we outline various modes of alternative genetic decoding and expand existing terminology to accommodate recently discovered manifestations of this seemingly sophisticated phenomenon.

Genetic code
A correspondence between 64 triplet combinations of four nucleotides and their standard amino acid or stop meanings.

Variant genetic codes
Genetic codes that differ from the standard genetic code shown in Figure 1a.

Proteinogenic amino acids
Amino acids that are incorporated into proteins co-translationally.

Ribosomal frameshifting
A process in which a ribosome changes its reading frame.

[1]School of Biochemistry and Cell Biology, University College Cork, Ireland.
[2]Department of Human Genetics, University of Utah, 15 N 2030 E Rm. 7410, Salt Lake City, Utah 84112–5330, USA.
Correspondence to P.V.B.
e-mail: p.baranov@ucc.ie
doi:10.1038/nrg3963
Published online
11 August 2015

The famous Jacob Monod phrase "Anything found to be true of *E. coli* must also be true of elephants" reflects the prevalent mind-set of biochemists and molecular biologists at the time when the genetic code was being deciphered. Indeed, the chemistry and molecular mechanisms of genetic information inheritance and its decoding from nucleic acids into proteins seemed to be the same for all forms of life. This culminated in Francis Crick's 'frozen accident' hypothesis for the origin of the genetic code[1], according to which the genetic code (FIG. 1a) is not only universal but also unchangeable and unevolvable. Ironically, the time of the hypothesis formulation also marked the beginning of a series of experimental observations of various exceptions from what are known as the standard rules of the genetic decoding, leading to a 'melting' in perceptions of the universality of the genetic code[2–19] (FIG. 1b).

We are now aware of more than 20 natural variant genetic codes[20] and it is very likely that there are more to discover. Two non-universal proteinogenic amino acids (selenocysteine (Sec)[11] and pyrrolysine (Pyl)[16]) have been discovered in addition to the 20 amino acids in the standard genetic code, and the possibility that more exist cannot be dismissed[21]. Numerous mRNAs have evolved special sequence elements to alter the meaning of specific codons or to result in ribosomes shifting reading frame or even bypassing long untranslated regions[22]. Alterations in genetic decoding are expected to exist in almost all organisms, and bioinformatics screens of bacterial genomes are revealing thousands of genes the decoding of which requires ribosomal frameshifting[23–25], and metagenome analyses of environmental samples point to many organisms with variant genetic codes[26].

Nature also continues to surprise us with novel types of alternative genetic decoding. The most recent discoveries pose provocative questions and suggest heretical ideas: could the genetic code of an organism change during the lifetime of the organism? Could a parasite hijack the genetic code of its host? Could there be genetic codes with non-triplet features? In this Review, we provide a brief overview of the known classes of alternative genetic decoding and their implications for entire genomes and individual mRNAs (TABLE 1). We begin with global changes of codon meaning on the scale of the entire genetic code (codon reassignments) and continue with local changes that are specific to particular sites in certain mRNAs (codon redefinition). We also describe processes that disrupt the continuity of triplet decoding, such as ribosomal frameshifting and translational bypassing. Along the way we expand existing terminology to accommodate the growing complexity of the phenomenon of alternative genetic decoding.

Global codon reassignments

Variant genetic codes resulting from fixed codon reassignments. The NCBI Genetic Code database currently provides a list of 21 distinct genetic code tables. Examination of these tables reveals two patterns: the high abundance of mitochondrial variant codes and

a

First letter	Second letter: U	Second letter: C	Second letter: A	Second letter: G	Third letter
U	UUU Phe (F)	UCU Ser (S)	UAU Tyr (Y)	UGU Cys (C)	U
	UUC	UCC	UAC	UGC	C
	UUA Leu (L)	UCA	UAA Stop (*)	UGA Stop (*)	A
	UUG	UCG	UAG Stop (*)	UGG Trp (W)	G
C	CUU Leu (L)	CCU Pro (P)	CAU His (H)	CGU Arg (R)	U
	CUC	CCC	CAC	CGC	C
	CUA	CCA	CAA Gln (Q)	CGA	A
	CUG	CCG	CAG	CGG	G
A	AUU Ile (I)	ACU Thr (T)	AAU Asn (N)	AGU Ser (S)	U
	AUC	ACC	AAC	AGC	C
	AUA	ACA	AAA Lys (K)	AGA Arg (R)	A
	AUG Met (M)	ACG	AAG	AGG	G
G	GUU Val (V)	GCU Ala (A)	GAU Asp (D)	GGU Gly (G)	U
	GUC	GCC	GAC	GGC	C
	GUA	GCA	GAA Glu (E)	GGA	A
	GUG	GCG	GAG	GGG	G

b

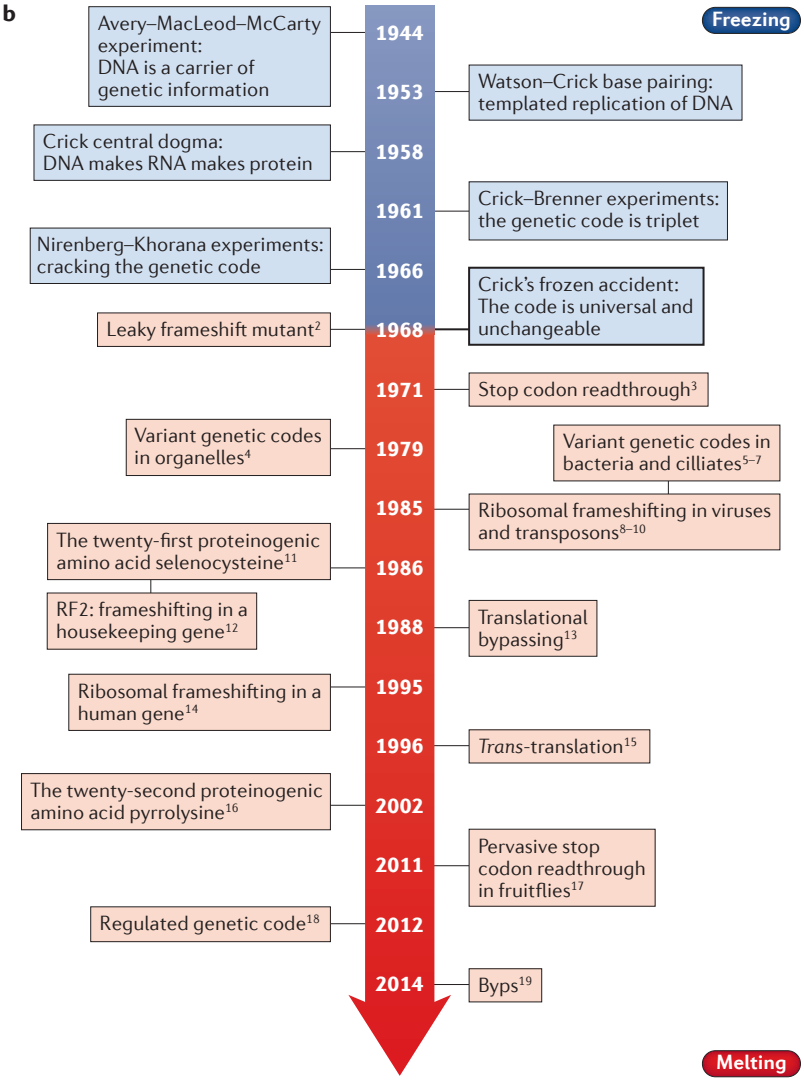


Figure 1 | **Freezing and melting of genetic decoding.** **a** | The standard genetic code. **b** | A timeline illustrating major events that shaped our current understanding of genetic decoding. Early discoveries that led to the formation of the universal principles of genetic decoding are shown in blue. More recent findings that have revealed evolvability and the flexibility of genetic decoding are shown in red.

fixed codon reassignment that involves stop codons. The only known non-mitochondrial sense codon reassignment is a CUG codon change from Leu to Ser in some yeast of the *Candida* genus[27]. The prevalence of mitochondrial reassignments is probably due to the small number of potentially affected genes, but could also be due to the properties of mitochondria and chloroplasts that are discussed in BOX 1. The prevalence of stop codon reassignment is due to the naturally infrequent occurrence of stop codons and their comparatively high evolutionary flexibility (BOX 2).

Alterations to the molecular machineries required for codon reassignment. The global rules of genetic decoding are defined by the molecular components of the translation machinery. Recognition of sense codons and stop codons differ. Whether a particular sense codon is decoded as a specific amino acid depends on two molecules, a tRNA that recognizes that codon and an aminoacyl-tRNA synthetase (aaRS) that charges the tRNA with the amino acid (FIG. 2). Therefore, a change of sense codon standard meaning should involve either modification of existing tRNAs and/or aaRSs or their loss or gain through gene duplications or horizontal gene transfer.

Stop codons are recognized by class I release factor (RF) proteins that directly interact with mRNA inside the ribosome and trigger hydrolysis and the release of the nascent peptide. Thus, which codon is recognized as a stop codon primarily depends on the structure of the RF proteins and lack of competition from cognate tRNAs. In most bacteria and organelles, there are two release factors that recognize stop codons semi-specifically: UGA and UAG are each recognized by only one specialized factor and UAA is recognized by both[28]. All known cases of codon reassignment in bacteria involve a loss of RF2, which recognizes UGA codons. This may not be accidental: a recent study of a large number of metagenomic sequences revealed many bacterial genes with reassigned stop codons that are exclusively UGA codons, suggesting that RF1, unlike RF2, may be indispensable for bacteria in the wild[26]. In the majority of characterized cases, UGA is reassigned to code for Trp[5,29,30], but in uncultured SR1 bacteria found in marine and fresh-water environments, as well as in human microbiota, UGA was recently shown to encode Gly[26,31].

Evolutionary routes to codon reassignment. There are several non-mutually exclusive models for the process of codon reassignment[20]. The codon capture model[32] involves the disappearance of a rare codon accompanied by a loss or change of function of its decoder (for example, UGA and RF2). The next step is the reappearance of the codon and its preferential decoding by a different molecule (for example, near-cognate reading of UGA by tryptophan tRNA, for which the cognate codon is UGG), followed by optimization of the new decoding. This model is supported by the existence of organisms with exceptionally rare codons in which the expression of the corresponding tRNAs is inessential for growth, for example, UUA in some *Streptomyces* species[33]. An

Table 1 | **Classification of different types of alternative genetic decoding events**

What is affected?* (individual codons, protein length or both)	Competes with standard translation in a concurrent or temporal manner?‡	Examples
***Global genomic distribution*§**		
Sense codon	No	Codon reassignments that lead to variant genetic codes, see http://www.ncbi.nlm.nih.gov/Taxonomy/Utils/wprintgc.cgi
	Yes	Decoding of CUG as serine and leucine in some *Candida* species[27]
Stop codon (protein length is affected)	No	Codon reassignments that lead to variant genetic codes, see http://www.ncbi.nlm.nih.gov/Taxonomy/Utils/wprintgc.cgi
	Yes	The meaning of UAG as a stop or Pyl in the genetic code of *Acetohalobium arabaticum* is dependent on the energy source[18]. During infection a phage-encoded RF2 alters the genetic code of its host[26]
Protein length	Yes or No	Ribosomal frameshifting in ciliates of the *Euplotes* genus[158] and translation bypassing in mitochondria of *Magnusiomyces capitatus*[19] are pervasive. However, it is not yet clear whether these processes are in competition with standard translation
***Local genomic distribution*§**		
Sense codon	No	No known natural examples
	Yes	Sec insertion in *Euplotes* species[81]
Stop codon (protein length is affected)	No	Sec insertion at mRNAs encoding selenoproteins with multiple Sec residues is likely to be highly efficient[76,77]
	Yes	Sec insertion; stop-codon readthrough
Protein length	No	No known natural examples
	Yes	Ribosomal frameshifting; translational bypassing

Pyl, pyrrolysine; RF2, release factor 2; Sec, selenocysteine.*An alternative decoding event may alter the meaning of a single codon with no effect on downstream codons. It can also alter the length of the protein sequence by combining protein sequences that are encoded in overlapping or disjointed open reading frames (ORFs). A change of codon meaning from stop to sense ultimately creates an elongated proteoform. ‡The alternative decoding event could occur in competition with standard translation (concurrent or temporal) or it could be a principal process. §Alternative decoding may affect the expression of all or most genes within a genome (globally distributed) or only a subset of individual mRNAs that bear specific stimulatory elements (local events).

Codon redefinition
A local change of codon meaning that is dependent on the context in which it occurs.

Translational bypassing
A process in which ribosomes skip three or more nucleotides without decoding.

Fixed codon reassignment
A complete unconditional change of the standard meaning of a codon.

Standard meaning
The way the translational machinery interprets a codon (coding for a proteinogenic amino acid or a signal for translation termination) unless it occurs in a specific context.

Codon capture
An evolutionary event in which a codon that disappears from a genome reappears in its descendant and acquires a different standard meaning, thus leading to a variant genetic code.

example of extreme codon scarcity is the CGG codon that occurs only once in all protein-coding genes of *Candidatus* Carsonella ruddii[34].

The ambiguous intermediate hypothesis[35] proposes the existence of an intermediate ambiguous state, in which a codon has two meanings. A classic example of an ambiguous state is in *Candida albicans* in which the CUG-decoding tRNA can be aminoacylated by either seryl-tRNA synthetase or leucyl-tRNA synthetase. As a result, Leu and Ser are stochastically distributed in the *Candida albicans* proteome at positions corresponding to CUG codons[36,37]. Ambiguous decoding can be beneficial. In response to oxidative stress in mammalian cells specific phosphorylation of methionyl-tRNA synthetase leads to increased methionylation of non-cognate tRNAs, which increases the amount of methionine in the proteome for protection against reactive oxygen species[38,39]. A recent example of a species of bacteria that uses two different genetic codes depending on the carbon source is described below. This suggests that a transition from one code to another does not require a 'walk through a valley of low fitness', as can be imagined in the case of ambiguous decoding that leads to the synthesis of presumably harmful aberrant proteins. However, the organisms could have increased fitness in an environment with oscillating conditions by altering their proteomes to suit the changes. A scenario in which an organism evolves a regulated genetic code in response to a changing condition and fixes the new variant once the new condition becomes stable is plausible.

Regulated codon reassignment. Pyl is one of the two non-universal proteinogenic amino acids (the other is Sec) that are not specified by the standard genetic code. Pyl is found in methanogenic bacteria and archaea[16] and is incorporated at UAG codons. Unlike Sec incorporation (see below), Pyl incorporation uses standard elongation factors and does not require a specialized RNA structure[40], although it has been reported that in certain contexts RNA structure affects the efficiency of Pyl incorporation[41].

The extreme scarcity of UAG codons and alterations in the mRNA recognition domain of release factors in methanogenic archaea suggest that UAG is not an efficient stop codon and that it is mostly used for the constitutive incorporation of Pyl[42]. The frequency of UAG codons in Pyl-utilizing bacteria is much higher and similar to that of other stop codons. Moreover, as both release factors are present, UAG is expected to be recognized as a stop codon.

A recent discovery[18] provides a clue as to why certain bacteria maintain UAG as a codon for Pyl and as an efficient stop codon. *Acetohalobium arabaticum*, which lives in a saline marine environment, can use several different sources of energy, including trimethylamine[43]. When trimethylamine is available, *A. arabaticum* expresses an

Ambiguous intermediate
An evolutionary state in the history of an organism evolving a variant genetic code in which a particular codon has two standard meanings.

Regulated codon reassignment
A conditional change of the standard meaning of a codon.

Recoding
A process of context- or condition-specific alteration of genetic decoding.

Stop codon readthrough
A redefinition of a stop codon to a sense codon irrespective of functional implications of the identity of the incorporated amino acid.

Proteoforms
Groups of sequence-related proteins arising from the same mRNA.

SECIS element
(Sec insertion sequence element). An mRNA secondary structure that functions as a stimulatory element for selenocysteine (Sec) incorporation.

operon that is required for Pyl incorporation, but in the absence of trimethylamine the expression of this operon is turned off. Thus, *A. arabaticum* is capable of regulating the meaning of the UAG codon in response to environmental conditions[18]. This marvellous example illustrates the ability of an organism to change its genetic code by using regulated codon reassignment without rewiring its genome.

Another startling example of a regulated genetic code has been proposed in a large metagenomic study of stop codon reassignments, which identified a bacteriophage that encodes RF2 in its genome, as well as a tRNA that recognizes UAG codons[26]. These bacteriophages infect bacteria lacking RF2 (with UGA being a sense codon). During the early stage of infection the virus expresses genes using the host genetic code in which UGA is a sense codon and UAG specifies stop. During the later stages of infection when the virus needs to shut off host protein synthesis and redistribute cellular resources to viral particle production, the virus expresses its RF2 and UAG-recognizing tRNA genes, thus changing the genetic code to one in which UGA specifies stop and UAG is a sense codon[26].

Local codon redefinition

The meaning of a codon can be changed in the context of a specific mRNA or at a specific location within the mRNA. To distinguish it from codon reassignment, this phenomenon is often termed codon redefinition and is considered to be a class of recoding events[44] (see FIG. 2). Naturally, because codon redefinition takes place in the context of a single or a subset of mRNAs, these mRNAs should have specific properties or sequence elements that distinguish them from other mRNAs.

Box 1 | Decoding plasticity in mitochondria

Global alterations of genetic decoding are frequent in mitochondria. Most known codon reassignments are found in mitochondria[20], and pervasive translational bypassing of byps elements was also discovered in mitochondria[19]. Why are mitochondria so prone to changes to their genetic decoding? There could be several explanations, which are not mutually exclusive.

First, mitochondrial genomes are very small: the transfer of genetic material between mitochondrial and nuclear genomes predominantly occurs in one direction[174], leading to extreme mitochondrial genome reduction. The genetic code used for decoding a small genome is easier to modify because a change in a codon meaning would affect fewer codons and fewer genes. Second, endosymbiosis isolated mitochondria from the outer world, making horizontal gene transfer (HGT) to mitochondria extremely unlikely to occur. Free-living microorganisms could use HGT to improve their genomes to suit particular environments by obtaining genetic material from other inhabitants of the same environment. However, to use the foreign genetic material, the recipient needs to use the same genetic language as the donor. This may create evolutionary pressure to maintain a universal genetic code. This pressure is unlikely to exist for organelles that cannot use HGT. Third, mitochondrial genome evolution is asexual, thus there is no mechanism for eliminating mildly harmful mutations, in contrast to evolution that involves sexual reproduction[175]. All these features allow for an increased diversity of molecular processes among the mitochondria of different organisms. However, some specific codon reassignments in mitochondria may also be beneficial. AUA reassignment from isoleucine to methionine occurred independently at least twice during eukaryotic evolution[19]. The resultant increase in methionine content in the mitochondrial proteome may protect mitochondria from their highly oxidative environment[176].

Stop codon readthrough. Stop codon readthrough is useful when there is a need to synthesize proteoforms with variant carboxy-terminal ends from the same mRNA (FIG. 3). Such a situation is highly beneficial for viruses in which stop codon readthrough can be used to economically encode a second product with a C-terminal extension (see REFS 45,46 for reviews). Until recently, stop codon readthrough had been reported only in a very small number of cellular chromosomal genes[47–51]. However, emerging evidence indicates that stop codon readthrough can be abundant during cellular gene translation. Phylogenetic analysis of protein-coding genes in 12 *Drosophila* species revealed that sequences downstream of annotated stop codons have evolved under constraints of protein-coding selection in almost 300 genes[17]. This indicates that the encoded protein rather than the nucleotide sequences of these regions are important for fruit-fly fitness and therefore that these sequences are likely to be translated during certain stages of the *Drosophila* spp. life cycle. These predictions were confirmed by ribosome-profiling experiments that revealed an even greater number of *Drosophila* spp. genes with detectable stop codon readthrough[52]. Although frequent stop codon readthrough in these species may be an exception, the number of human genes with documented stop codon readthrough is also increasing[53–56].

Translational termination is slower and less accurate than elongation. Moreover, its efficiency and accuracy are context dependent[57–59]. Thus, low-efficiency stop codon readthrough can be achieved in the absence of any sophisticated stimulatory structures. Sequence constraints as short as six nucleotides downstream of a stop codon are sufficient to achieve a readthrough efficiency that is significantly higher than background levels[60]. Selection for a weak termination context downstream of stop codons is evident among many stop codon readthrough genes; however, a higher efficiency of readthrough often involves additional elements such as RNA secondary structures[61].

Selenocysteine incorporation. Sec is one of the two non-universal proteinogenic amino acids and its incorporation into proteins is another type of codon redefinition. Sec is specified by UGA codons and, as with stop codon readthrough, this results in the synthesis of a protein with a C-terminal extension relative to the product of termination (FIG. 3). However, synthesis of different lengths of alternative proteoforms is not the main purpose of Sec incorporation, as only the longest proteoforms are believed to be functional. Rather, Sec is often incorporated into the catalytic centres of certain enzymes to improve their biochemical properties[62,63].

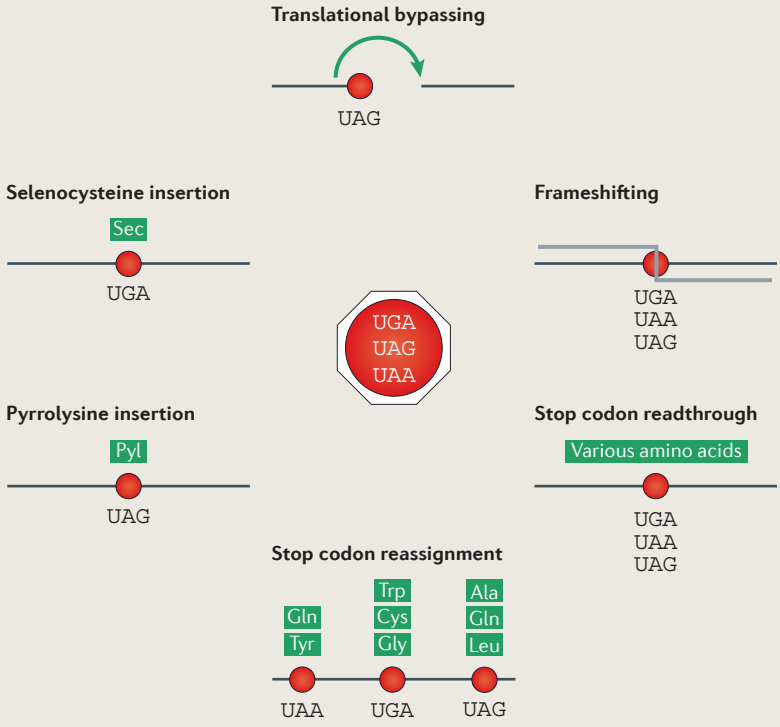
The mRNAs of selenoproteins are specified by a special RNA secondary structure known as a SECIS element (Sec insertion sequence element). The structures of SECIS elements and their locations differ between bacteria and eukaryotes. They occur within the coding region of mRNAs in bacteria[64] but are present in the 3′ untranslated regions (UTRs) in eukaryotes[65], although yet another auxiliary structure (the Sec codon redefinition element (SRE)) occurs within the coding regions of some eukaryotic selenoprotein-encoding mRNAs[66,67].

Box 2 | **The multitude of stop codon meanings**

Stop codons are clearly the most versatile codons. Nearly all codon reassignments in non-mitochondrial genomes involve stop codons. Stop codons are frequently found in +1 frameshifting sites, a stop codon is required for translational bypassing in T4 bacteriophage gene *60* and all known codon redefinitions involve stop codons (with the conspicuous exception of selenocysteine (Sec) insertion in *Euplotes* species[81]). See the figure for various examples of altered stop codon meanings. Stop codons are shown as red circles, and the different types of altered meanings are shown in green (for amino acid incorporations or translational bypassing) and grey (for frameshifting).

What makes stop codons so flexible? First, stop codons are the rarest codons in coding sequences. Genes that do not use alternative decoding in their expression need only a single stop codon. As a result, stop codons occur less frequently than most sense codons by an order of a magnitude. Furthermore, stop codon usage is often skewed[177] and, as a result, a particular stop codon can be exceptionally rare. Thus, a codon reassignment of a stop codon may affect only a few codons in the genome. In addition, alteration of stop codon meaning is unlikely to dramatically change the properties of protein products. Stop codons are as frequent outside of coding regions as any other triplet, thus if a stop codon is decoded as a sense codon a corresponding protein will be extended by only a few amino acids, as another stop codon is likely to be found downstream.

The other important difference in the decoding of stop codons is that they are recognized by release factor proteins rather than by RNA molecules. In fact, the reason why they are recognized by proteins could also be due to their comparatively high evolvability. Release factors in eukaryotes and archaea do not share a common ancestor with those in bacteria, suggesting that the protein-based termination of translation is a relatively recent phenomenon and the last universal common ancestor (LUCA) used a different mechanism (perhaps RNA based) for the termination of protein synthesis. Because the termination of translation relies on protein factors, the mechanism of mRNA recognition is substantially different. Although release factors still recognize stop codons as triplets, they interact with larger regions of mRNA and thus the efficiency and accuracy of termination is more sensitive to the sequence downstream than in the case of strictly triplet tRNA decoding. As a result, the strength of stop codons varies substantially depending on the sequence downstream of stop codons[57–59]. Particularly weak codons are often used for stop codon readthrough or ribosomal frameshifting because the weak termination provides greater opportunities for competing processes.



The SECIS element is obviously insufficient for this process to take place in any organism. The organism also needs to be able to synthesize Sec and to incorporate it into proteins. These steps involve the expression of a number of genes, including those coding for a specialized Sec-tRNA, as well as specialized elongation factors that bring Sec-tRNA to the ribosome and auxiliary factors that allow the recruitment of the Sec-tRNA to only SECIS-containing mRNAs (see REFS 68–72 for reviews). Acquisition of the Sec machinery in bacteria could be conceivably due to frequent horizontal gene transfer from bacteria in which the required genes are organized on a single operon[73]. In eukaryotes, spontaneous simultaneous transfer of all genes required for Sec specification is unlikely; therefore, the predominant mode for the evolution of this trait is its loss, as is evident in certain yeast, nematodes[74] and insects[75].

The decoding of UGA as a Sec is not 100% efficient, although it is expected to be highly efficient in selenoprotein P mRNA in which Sec is incorporated at multiple locations[76]. For example, selenoprotein P in sea urchin has 28 Sec residues[77]. If we assume that the efficiency of Sec incorporation at each UGA is 90%, only approximately 5% (0.9^{28}) of the ribosomes that translate selenoprotein P mRNA would be able to synthesize the full-length protein, and the remainder would terminate on one of the 28 UGA codons. Therefore, it has been hypothesized that Sec incorporation is only inefficient at the first UGA codon. Sec incorporation at this codon might serve as a checkpoint for the availability of factors that are required for Sec incorporation. The efficiency of all subsequent Sec incorporations would be close to 100%[78].

The sensitivity of Sec incorporation to the availability of selenium varies among different selenoprotein mRNAs[66]. This allows the synthesis of particularly important selenoproteins at the expense of less important ones under conditions of selenium deficiency. Interestingly, the same Sec insertion machinery can be used to incorporate standard Cys at UGA codons, albeit inefficiently[79]. This mechanism provides the possibility of synthesizing low levels of a full-length but Sec-depleted variant of a 'selenoprotein' even in the absence of selenium. Although less active, enzymes with a cysteine instead of Sec retain some activity[62,63]: an inefficient enzyme is better than no enzyme at all.

Sec is not always specified by a stop codon. In ciliates of the genus *Euplotes*, which possess the Sec machinery, the UGA stop codon is reassigned to code for cysteine[80]. Despite this reassignment, the SECIS structure in the 3′ UTR of *Euplotes* spp. mRNA can, at UGA codons in specific locations in mRNAs, direct a proportion of the ribosomes to insert Sec instead of cysteine in the growing polypeptide[81].

Clearly, UGA did not evolve to encode Sec after being globally reassigned to a sense codon in *Euplotes* spp.. In an ancestor of *Euplotes* spp. with UGA specifying stop, UGA was redefined in some mRNAs to specify Sec and only then was UGA-stop reassigned to UGA-Cys in the entire genetic code. Therefore, even this sense-to-sense

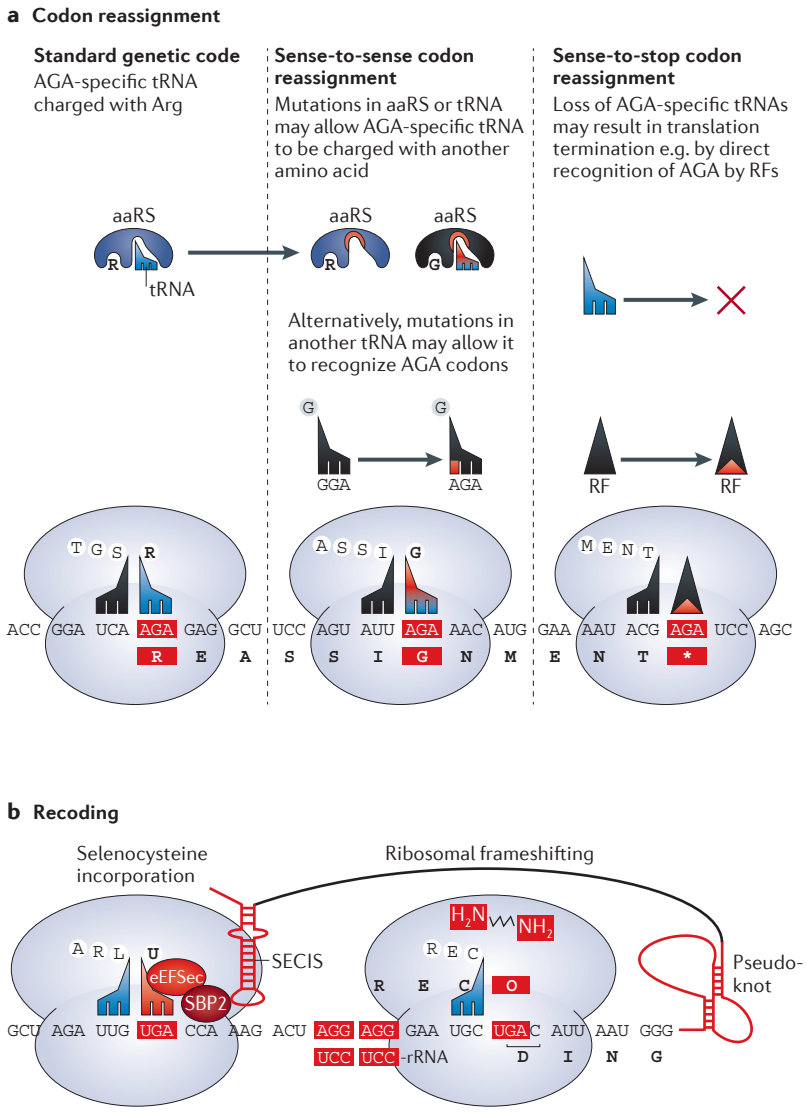


Figure 2 | **Components that shape alternative genetic decoding.** The figure illustrates a distinction between codon reassignment and recoding, which are sometimes confused in the literature. Collective representation (assembled from several examples from different organisms) of stop codon reassignment (part **a**) and recoding (part **b**) is shown. Different ways by which codon meaning can be reassigned, as exemplified by the AGA codon (part **a**), are shown. This codon is known to have four different meanings depending on the variant genetic code used in the corresponding organism. Codon reassignment can originate as a result of changes in tRNAs, aminoacyl-tRNA synthetases (aaRSs) or release factors (RFs), although it may also involve other components, as exemplified by AGA reassignment to stop codon in vertebrate mitochondria. Codon reassignment affects the expression of all genes in the organism that use the reassigned codon. Two recoding events are shown (part **b**). Codon redefinition (left) is exemplified with a schematic of selenocysteine (Sec; single letter code U) insertion at UGA codons in eukaryotes, which requires a Sec insertion sequence (SECIS) element in the 3′ UTR, a specialized tRNA, an elongation factor eEFSec and auxiliary protein SECIS-binding protein 2 (SBP2). Ribosomal frameshifting (right) is shown as a collective representation of several frameshifting events. Shine–Dalgarno interactions with ribosomal RNA stimulate frameshifting in bacterial release factor 2 mRNA decoding. A stimulatory downstream RNA pseudoknot structure is present in many eukaryotic antizyme mRNAs. Frameshifting in antizymes is also known to be sensitive to polyamine concentration. These examples illustrate how (in contrast to codon reassignment), recoding events are dependent on favourable sequence contexts that locally alter the interpretation of the codon sequences.

codon redefinition is a product of two evolutionary events, both involving a modification of stop codon meaning (BOX 2). However, this provokes the question of whether other sense-to-sense codon redefinitions may exist. It is comparatively simple to find codon redefinition events that involve stop codons owing to their dramatic effect on the decoded products (FIG. 3). It is easier to detect protein products for which size is substantially altered; it is also possible to use phylogenetic approaches for finding cases of stop codon redefinition, as the nucleotide sequence downstream of the redefined stop codons is likely to exhibit signatures of protein-coding evolution[17]. Sense-to-sense codon redefinition in *Euplotes* spp. was primarily discovered to satisfy scientific curiosity regarding how Sec is incorporated in *Euplotes* spp. in which UGA is not a stop codon, hence investigators knew exactly where to look[81]. If there are other cases of sense-to-sense codon redefinition, we do not know where to search for them. The benefit of redefining the meaning of a codon from one standard amino acid to another standard amino acid at some positions of specific mRNAs is unclear, as the same could often be accomplished with less than three point mutations in DNA. The advantage of sense codon redefinition is obvious only in the case of the incorporation of non-universal proteinogenic amino acids. Although we cannot exclude the possibility that there is a twenty-third proteinogenic amino acid[21], so far it has not been found[82,83].

Ribosomal frameshifting

Irrespective of the meaning of individual codons in a variant genetic code, most proteins in all modern organisms are expected to be decoded as uninterrupted sequences of nucleotide triplets, with no gaps or overlaps between the codons. Ribosomal frameshifting (FIG. 3) is often described as programmed ribosomal frameshifting (PRF) when it occurs at a specific location in the mRNA, implying that the sequence of mRNA 'programs' frameshifting to occur. There is a minority of genes the expression of which requires PRF for the synthesis of encoded proteins. Frameshifting may also affect the expression of other genes that encode proteins in the standard triplet manner. In this case, the synthesis of encoded full-length protein would be abolished. Because of the opposing effects of frameshifting on gene expression, we propose that productive PRF and abortive PRF need to be distinguished and separately defined, therefore we describe them in different subsections below.

Productive PRF. The requirement for ribosomal frameshifting is especially common in viral decoding, where it is sometimes used for the same purpose as stop codon readthrough; that is, to generate proteoforms with alternative C termini and different functional properties.

Owing to the sensitivity of ribosomal frameshifting to cellular conditions it is often used for regulatory purposes. In this case, the synthesis of a functional product depends on ribosomal frameshifting, whereby the product of standard translation is usually dysfunctional and the efficiency of frameshifting is regulated. One such example is the decoding of bacterial release factor 2 (RF2)

Programmed ribosomal frameshifting
(PRF). Ribosomal frameshifting that is programmed (by a sequence context) to occur at a specific mRNA location.

Productive PRF
Programmed ribosomal frameshifting (PRF) that is required for the production of a functional protein product.

Abortive PRF
Programmed ribosomal frameshifting (PRF) that results in the synthesis of dysfunctional protein products or in the downregulation of functional protein synthesis.

Purifying evolutionary selection
The removal of disadvantageous traits. In the case of protein-coding sequences it results in a higher rate of synonymous substitutions relative to non-synonymous substitutions.

mRNA, for which the product of its translation, the RF2 protein, inhibits frameshifting and thus downregulates its own expression through a negative feedback loop[12]. This mechanism is highly conserved and operates in the vast majority of bacteria[84,85]. In eukaryotes, frameshifting in antizyme mRNAs is dependent on the concentration of polyamines[14] the synthesis and uptake of which is regulated by antizyme proteins, and this mechanism operates in all sequenced antizyme genes with the exception of antizyme in *Tetrahymena* spp.[86]. It was recently discovered that the mechanism of response to antibiotics also involves frameshifting in some bacteria[87].

Thus far, numerous examples of productive PRF have been well documented and extensively studied and may be found in all or almost all organisms (see REFS 22,45,88–93 for selected comprehensive reviews). However, recent years have been particularly fruitful for discovering novel cases of productive PRF owing to the abundance of sequenced genomes, which have allowed powerful comparative sequence analyses. These comparative analyses can provide clear evidence of frameshifting functionality by detecting purifying evolutionary selection acting on either the sequence that is required for efficient ribosomal frameshifting or the protein-coding region that requires upstream frameshifting for its expression[94]. This approach permitted the identification of a number of novel 'hidden' genes in viruses that require a frameshifting event for their expression[95–101], including one in influenza A[102,103].

Similar analyses of bacterial genomes revealed numerous genes that are under the control of ribosomal frameshifting; these genes predominantly occur in transposable elements and prophages[23–25]. As in bacteria, the majority of eukaryotic nuclear genes that are known to use frameshifting in their expression are in transposable elements or their derivatives[104]. In humans, for example, only five genes that require frameshifting for their expression are currently known: three ornithine decarboxylase antizyme paralogues (*OAZ1*, *OAZ2* and *OAZ3*)[86] and two transposon-derived genes *PEG10* (REF. 105) and *PNMA3* (REF. 106). Although it is likely that many cases of productive PRF are still undiscovered, it is also likely that productive PRF is very rare among non-mobile chromosomal genes in most organisms (see below for a noticeable exception involving pervasive frameshifting in *Euplotes* spp.). For instance, comparative sequence analysis of insect genomic sequences revealed only four genes with overlapping coding regions that might be expressed through PRF[107]. Only one of these genes, which encodes a homologue of the human tumour suppressor adenomatous polyposis coli (APC), was confirmed to use ribosomal frameshifting for the expression

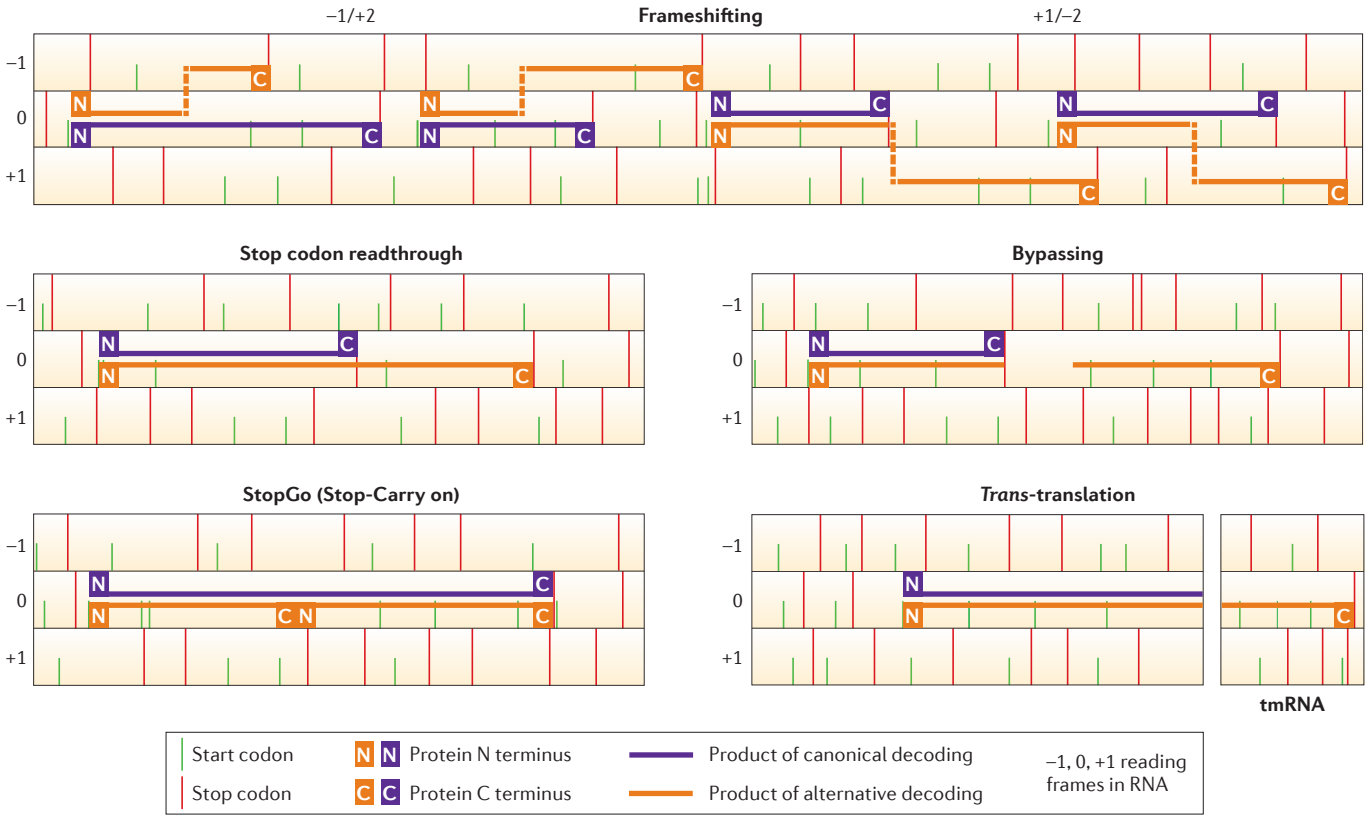


Figure 3 | **Relationship between nucleotide sequences and alternatively decoded proteins.** Nucleotide sequences are depicted as three horizontal boxes representing three different reading frames. Start (green) and stop (red) codons are shown as vertical lines. The sequences that are translated into proteins are shown as horizontal bars with amino- and carboxy-terminal ends indicated. The purple bars correspond to standard decoding and orange bars correspond to alternative events. Transitions between reading frames are denoted with broken lines. tmRNA, transfer and messenger RNA.

of an alternative proteoform[108]. The ribosome profiling (ribo-seq) technique [109] allows the determination of the positions of the translating ribosomes on mRNA *in vivo* and with sub-codon resolution. Thus, it can be used for detecting which reading frame is translated in mRNA. Ribo-seq data obtained in yeast and in human cultured cells did not reveal abundant ribosomal frameshifting, although it confirmed some of the previously identified cases and revealed a number of sequences that are translated in more than one frame[110]. Low sequence coverage in early ribo-seq data limited the predictive power of this technique, although perhaps future applications of similar methods will become more fruitful at identifying novel instances of ribosomal frameshifting.

Abortive PRF. Ribosomal frameshifting does not necessarily lead to the synthesis of functional products. When frameshifting occurs it leads to the synthesis of aberrant products owing to limitations of translation accuracy. Thus, it is expected that frameshift-prone sequences should be avoided in most coding regions. Indeed, a strong evolutionary selection is observed that eliminates sequences that are strongly prone to frameshifting in protein-coding genes[111]. Such selection is also observed for some[112], but not all, weakly frameshift-prone sequences[113–115], and in both cases they are frequently found among coding regions[112–115]. These patterns of occurrence suggest that abortive frameshifting at low frequency might also be beneficial for regulatory purposes under certain conditions. Such frameshifting has been shown to reduce the stability of those mRNAs where it occurs, thus downregulating the expression of the encoded proteins[116]. Moreover, such abortive frameshifting regulated by miRNAs was reported in a human gene, the HIV-1 co-repressor *CCR5* (REF. 117).

Frameshifting sites, stimulators and attenuators. Perhaps one of the reasons why ribosomal frameshifting is widespread as a local decoding alteration is its responsiveness to the sequence elements within an mRNA. Most cases of ribosomal frameshifting require a combination of a frameshifting site (a specific mRNA sequence in which frameshifting takes place) and a stimulatory element (a sequence in the same mRNA that increases the efficiency of frameshifting).

With the exception of the cases described below, frameshifting sites alone are insufficient to trigger efficient frameshifting and they require one or more stimulatory elements embedded in the mRNA. Stimulatory elements can be of diverse types. In bacteria, mRNA complementary to ribosomal RNA may facilitate frameshifting in the +1 as well as in the –1 directions[118–120]. Nascent peptides are also known to modulate ribosomal frameshifting[121,122]. The largest class of stimulators are RNA secondary structures: stem-loops[123–125], simple[126] and relatively complex RNA pseudoknots containing extra stems[127–129], or triple helices[130,131], kissing loops[132], G-quadruplexes[133,134] or long-range interactions[135,136] (see REFS 22,137–139 for reviews). mRNA interacts with various cellular components, and these interactions may alter the stimulatory properties of particular structures or sequences. Both protein molecules[140] and nucleic acid molecules[117,141,142] have been shown to modulate frameshifting in *trans*. The dependence of frameshifting on stimulatory elements, as well as the responsiveness to cellular conditions, provides translation with a powerful regulatory mechanism. Fine regulation can be achieved through the balance between positive and negative regulators; therefore, it is reasonable, though nonetheless surprising, that in addition to stimulators, there are also attenuators of frameshifting[129,143,144]. The remarkable way in which such an attenuator operates was described in the *Saccharomyces cerevisiae OAZ1* antizyme gene, where it has been proposed that a nascent peptide in the exit tunnel promotes ribosome stalling at the end of antizyme open reading frame 2 (ORF2) in the presence of high polyamine levels. A hypothesis is that the resultant 'pile up' of ribosomes promotes the termination of ribosomes at the ORF1 stop codon, thus reducing frameshifting efficiency[143].

Frameshifting at triplet repeats. The frameshifting observed during the translation of certain triplet repeats[145–148] cannot be described as a combination of a frameshifting site and its stimulator. In this case, it is difficult to clearly distinguish the stimulator from the frameshifting site because frameshifting can take place at more than one location in the sequence, which can also serve as a stimulator. Decoding a long sequence of consecutive codons corresponding to the same tRNA could lead to the depletion of the charged tRNA in the vicinity of the ribosome, meaning that the frameshifting on such repeats is sensitive to tRNA concentration[148]. This is consistent with a role for frameshifting as a sensor of cellular conditions.

Universal frameshifting. In some circumstances, stimulators are not needed for efficient frameshifting. In *S. cerevisiae*, up to 40% efficient +1 frameshifting can occur at C.UU_A.GG_C (underscores denote codon boundaries in the initial reading frame and dots denote codon boundaries in the shifted reading frame)[149]. The high efficiency of frameshifting can be explained by a simple kinetic model: upon tRNA slippage from CUU to UUA at the P-site, translation could quickly continue in the new frame owing to a high concentration of tRNA cognate for the +1 frame GGC, and translation in the original frame is slowed owing to the scarcity of tRNA cognate to zero frame AGG[150]. Frameshifting at C.UU_A.GG_C was initially found to be used in transposon Ty1 (REF. 151) and was also later found to be required for the synthesis of ABP140 (REF. 152). A few codons in addition to CUU were found to promote unusually efficient frameshifting when in the P-site[153]. For all these codons, tRNAs with optimal codon–anticodon base pairing are lacking in *S. cerevisiae*, and there is evidence that frameshifting is strongly promoted by the incorporation of near-cognate isoacceptor tRNAs at corresponding codons[154].

Therefore, +1 ribosomal frameshifting at these heptameric high-efficiency frameshift sequences is a specific

Frameshifting site
(Also known as frameshift site and shift site). A sequence in which ribosomal frameshifting takes place. It includes codons in the A- and P-sites of the ribosomes just before and after the frameshifting. It is useful to describe the sequence of the frameshifting site denoting codons in the original and new frames, for example, C.UU_A.GG_C. Such representation unambiguously reflects the direction (minus or plus) as well as the mechanism of frameshifting (+1, +2, and so on)

Stimulatory element
An mRNA element that is required for the efficient local alteration of genetic decoding.

P-site
The ribosomal site that accommodates the peptidyl-tRNA carrying the growing polypeptide chain.

Isoacceptor
One of a group of tRNA species carrying the same amino acids but with different anticodon sequences

feature of *S. cerevisiae* and could be regarded as a feature of its genetic code. This feature affects the evolution of all genes in the genome because the occurrence of strong frameshift-prone heptameric sequences would be highly detrimental to the accurate protein synthesis of genes that do not require frameshifting. As a result, these frameshift-prone heptamer sequences are the rarest heptamers in coding regions of *S. cerevisiae*[111].

Another case for frameshifting as a universal feature of the genetic code has been proposed for vertebrate mitochondria in which there are no tRNAs that recognize AGA and AGG codons. These were believed to be stop codons in the genetic code of vertebrate mitochondria. Temperley *et al.*[155] suggested that AGG and AGA codons are not recognized by termination factors, but that they instead promote –1 frameshifting at N_NN.U_AG.G and N_NN.U_AG.A (where N is any nucleotide). This places the A-site of the ribosome at the overlapping UAG stop codons. If this is true, the vertebrate mitochondrial genome could be described as non-triplet, with AGG and AGA codons as signals for –1 frameshifting rather than specifying amino acid or stops. However, evidence has also been provided that AGG and AGA may be recognized as stop codons by mitochondrial release factors[156].

More recently, the frameshifting-based model has been challenged by a study suggesting that the nascent peptide chain is released from the stalled ribosome with the aid of ICT1, a general rescue factor of stalled ribosomes in mammalian mitochondria[157]. Irrespective of whether Temperley's elegant hypothesis is right, it provides an intriguing scenario for the existence of genetic codes with non-triplet features.

Pervasive frameshifting. Ribosomal frameshifting is usually very infrequent. This is true even in organisms with universally strong frameshifting patterns, such as C.UU_A.GG_C in *S. cerevisiae*, because despite their high frameshifting efficiency these heptamers occur only rarely in the genome. By contrast, ribosomal frameshifting in *Euplotes* spp. ciliates is widespread. Analysis of the sequence and expression of several *Euplotes* genes revealed that the frequent presence of frameshifting was strongly associated with A.AA_U.AA_N and its minor variant A.AA_U.AG_N. It has been estimated that frameshifting is likely to occur in the decoding of 10% of all *Euplotes* genes[158]. As described above, the *Euplotes* genetic code is a variant one with the UGA stop codon reassigned to cysteine. The current explanation for the high frequency of frameshifting in *Euplotes* is that the alteration of the mRNA-recognition properties of the *Euplotes* release factors required for the reassignment also weakens the recognition of UAG and UAA codons. Inefficient termination at these codons could favour an alternative event, ribosomal frameshifting at AAA codons[159]. It remains unclear, however, whether all occurrences of A.AA_U.AA_N and A.AA_U.AG_N sequences lead to frameshifting in *Euplotes*, and the efficiency of this process is also unknown. Whether these sequences are the only sequences to promote frameshifting in *Euplotes* is not known.

Unusual disruptions of triplet continuity

Translational bypassing. What happens during the translation of bacteriophage T4 gene *60* could be described as a +50 frameshift, as two codons that encode a pair of adjacent amino acids in its protein product are separated by a 50-nucleotide-long non-coding gap (FIG. 3). Ribosomes suspend translation at a specific glycine codon and resume translation at another glycine codon that is 50 nucleotides downstream. This case was discovered in 1988 (REF. 13) (FIG. 1b), and until recently was the only known example of translational bypassing (also known as ribosomal hopping). Intensive research into the molecular mechanism of translational bypassing revealed a complex nexus of stimulatory elements within the mRNA, including an enigmatic requirement for a stop codon at the start of the non-coding gap to be dynamically folded into an RNA secondary structure and a role for the specific sequence of the nascent peptide[160–164].

The rarity of this case, combined with the complexity of the sequence elements involved, seeded doubt about the existence of other examples of translational bypassing. However, likely candidates have recently been identified in bacteriophages infecting *Streptomyces* species[165], with the most striking examples of translational bypassing found in the mitochondria of the yeast *Magnusiomyces capitatus*[19]. The transcriptome of these mitochondria contains dozens of untranslated regions of various sizes, which were termed byps to indicate their relationship to the translational bypassing in decoding bacteriophage T4 gene *60*. The exact sequences of byps vary, although they share certain features with the primary as well as the secondary RNA structure[19]. Bypassing at *M. capitatus* byps does not occur when the corresponding sequences are expressed in *Escherichia coli* and hence is likely to require specific features of the *M. capitatus* translational machinery[19]. Byps are probably mobile genetic elements and it is likely that the *M. capitatus* translational machinery evolved the ability to avoid them during translation so that the insertion of byps into a coding region would not be deleterious[19].

Trans-*translation.* The term *trans*-translation describes the translation of two messenger RNAs into a single polypeptide (FIG. 3). Most bacteria use *trans*-translation to manage truncated mRNAs that lack stop codons. Ribosomes that stall at the end of truncated mRNAs are recognized by elongation factor EF-Tu in a complex with a molecule known as tmRNA (transfer and messenger RNA). tmRNA functions as a tRNA, as it is charged with an amino acid that is then transferred onto the nascent peptide in the stalled ribosome. It also functions as an mRNA, as it contains a short ORF that is decoded after its incorporation into the ribosome. The tmRNA ORF encodes a signal for protein degradation, thus allowing the cell to destroy potentially toxic products of truncated mRNAs and also to recycle the stalled ribosomes (see REFS 166,167 for reviews and the database of tmRNA sequences[168]). This system is remarkably conserved in bacteria and has been lost only in a few highly reduced genomes[169]. Therefore,

A-site
The ribosomal site that accommodates either the aminoacyl-tRNA carrying the next amino acid to be added to the growing polypeptide chain or a release factor.

Byps
Non-coding gaps in mRNAs of mitochondria (in *Magnusiomyces capitatus* and related species) that escape decoding through frequent translational bypassing.

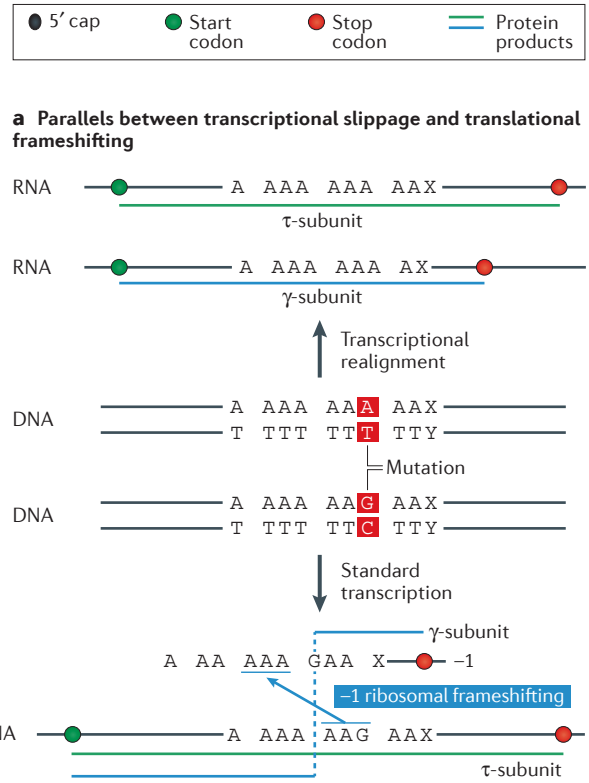
Trans-translation
A process in which a single protein is translated from two mRNA molecules as templates.

Box 3 | **Transcriptional recoding and alternative initiation of translation**

The focus of this Review is on alternative decoding events that take place during translation elongation or termination. However, there are many pre-translational events (occurring co-transcriptionally or post-transcriptionally) that alter genetic decoding in similar ways. C-to-U and A-to-I RNA editing may change the meaning of a codon template in DNA (see REFS 178–180 for reviews). Pseudouridinilization of mRNA[181] affects decoding in a complex way and could result in the readthrough of stop codons containing a pseudouridine[182].

Programmed transcriptional realignment (PTR) affects decoding similarly to programmed ribosomal frameshifting (PRF), and these two mechanisms are easily interchangeable, potentially via a single nucleotide mutation (see the figure, part **a**). The use of PRF and PTR in orthologous genes was first revealed in *dnaX*[183].

Alternative initiation of translation is analogous to stop codon readthrough in the sense that both generate multiple proteoforms, but with variant amino termini in the case of alternative initiation and carboxyl termini in the case of stop codon readthrough (see the figure, part **b**). Alternative translational initiation is a widespread phenomenon with important implications for gene expression. Many human genes contain evolutionarily conserved protein-coding non-AUG-initiated extensions in their 5′ leaders[184], and repeat-associated non-AUG initiation is implicated in human neurodegenerative disorders[185–187].



b Parallels between alternative initiation and stop codon readthrough

trans-translation may be described as a global feature of genetic decoding in organisms with tmRNA, as it universally adds a particular peptide sequence to any protein product of a translated ORF that lacks a stop codon.

StopGo. StopGo (also known as Stop-Carry on) allows two separate peptides to be produced from the same ORF (FIG. 3). It was first characterized in an aphthovirus, foot-and-mouth disease virus (FMDV), where it occurs during the decoding of a polyprotein-encoding mRNA[170,171]. In this case, ribosomes decode codons according to the standard genetic code and in a triplet manner. However, the continuity of decoding is disrupted by a break at a specific location. The function of this event is parallel to the proteolytic cleavage at specific sites in other viral polyproteins, but with StopGo there is no requirement for a protease.

Conclusions and future perspectives

In this Review, we have focused on alternative genetic decoding that occurs during translation elongation. Parallel mechanisms can also be used during transcription or translation initiation, and we discuss these briefly in BOX 3. However, the remarkable flexibility of genetic decoding is evident from the examples discussed. Genetic decoding can be globally altered as a result of modifications to the translation machinery or locally altered in specific mRNAs that evolve special sequence elements to alter their decoding and to regulate their own translation. A crucial aspect of alternative genetic decoding is its relevance to synthetic biologists who can take advantage of genetic decoding plasticity to construct regulatory and sensory genetic modules. Synthetic organisms with genetic codes that are not found in nature are also now being generated[172,173].

It is unlikely that we are aware of all the ways in which genetic decoding can be altered in nature, and future discoveries will undoubtedly continue to surprise and inspire us. Comparative sequence analysis may help us to reveal alternatively decoded genes that are currently hidden from our sight because they do not fit the standard gene models that are used for the identification of protein-coding genes. However, it is unlikely to reveal the entire spectrum of recoded genes. The power of comparative sequence analysis is limited by the range of species. With higher eukaryotes a problem lies in deconvoluting the signatures of evolution in protein-coding regions in the presence of alternative splicing. Complementary experimental techniques for the genome-wide analysis of protein synthesis, such as ribosome profiling, have also emerged. Although for many recoded genes it may be difficult to find conditions in which they are expressed, combinations of genome-wide phylogenetic and biochemical approaches are already accelerating the discovery of new cases. The growing repertoire of alternative genetic decoding events will ultimately challenge the way in which we annotate genes and genomes and how we represent protein-coding information in sequence databases.

StopGo
(Also known as Stop-Carry on). A process in which the production of a polypeptide chain is interrupted at a specific place while triplet mRNA decoding continues. This results in the production of two protein products from a single open reading frame.

1. Crick, F. H. The origin of the genetic code. *J. Mol. Biol.* **38**, 367–379 (1968).
This is a seminal paper in which the frozen accident hypothesis of the origin of the genetic code was formulated.
2. Riyasaty, S. & Atkins, J. F. External suppression of a frameshift mutant in salmonella. *J. Mol. Biol.* **34**, 541–557 (1968).
3. Weiner, A. M. & Weber, K. Natural read-through at the UGA termination signal of Q-β coat protein cistron. *Nat. New Biol.* **234**, 206–209 (1971).
4. Barrell, B. G., Bankier, A. T. & Drouin, J. A different genetic code in human mitochondria. *Nature* **282**, 189–194 (1979).
5. Yamao, F. *et al.* UGA is read as tryptophan in *Mycoplasma capricolum*. *Proc. Natl Acad. Sci. USA* **82**, 2306–2309 (1985).
6. Horowitz, S. & Gorovsky, M. A. An unusual genetic code in nuclear genes of Tetrahymena. *Proc. Natl Acad. Sci. USA* **82**, 2452–2455 (1985).
7. Preer, J. R. Jr., Preer, L. B., Rudman, B. M. & Barnett, A. J. Deviation from the universal code shown by the gene for surface protein 51A in Paramecium. *Nature* **314**, 188–190 (1985).
8. Clare, J. & Farabaugh, P. Nucleotide sequence of a yeast Ty element: evidence for an unusual mechanism of gene expression. *Proc. Natl Acad. Sci. USA* **82**, 2829–2833 (1985).
9. Jacks, T. & Varmus, H. E. Expression of the Rous sarcoma virus pol gene by ribosomal frameshifting. *Science* **230**, 1237–1242 (1985).
10. Mellor, J. *et al.* A retrovirus-like strategy for expression of a fusion protein encoded by yeast transposon Ty*1*. *Nature* **313**, 243–246 (1985).
11. Chambers, I. *et al.* The structure of the mouse glutathione peroxidase gene: the selenocysteine in the active site is encoded by the 'termination' codon, TGA. *EMBO J.* **5**, 1221–1227 (1986).
12. Craigen, W. J. & Caskey, C. T. Expression of peptide chain release factor 2 requires high-efficiency frameshift. *Nature* **322**, 273–275 (1986).
13. Huang, W. M. *et al.* A persistent untranslated sequence within bacteriophage T4 DNA topoisomerase gene 60. *Science* **239**, 1005–1012 (1988).
14. Matsufuji, S. *et al.* Autoregulatory frameshifting in decoding mammalian ornithine decarboxylase antizyme. *Cell* **80**, 51–60 (1995).
15. Keiler, K. C., Waller, P. R. & Sauer, R. T. Role of a peptide tagging system in degradation of proteins synthesized from damaged messenger RNA. *Science* **271**, 990–993 (1996).
This work reveals the function of tmRNA that rescues ribosomes from truncated mRNA lacking stop codons through the process of *trans*-translation.
16. Srinivasan, G., James, C. M. & Krzycki, J. A. Pyrrolysine encoded by UAG in Archaea: charging of a UAG-decoding specialized tRNA. *Science* **296**, 1459–1462 (2002).
This work describes the discovery of twenty-second proteinogenic amino acid.
17. Jungreis, I. *et al.* Evidence of abundant stop codon readthrough in Drosophila and other metazoa. *Genome Res.* **21**, 2096–2113 (2011).
This work reveals signatures of protein-coding evolution in the sequences downstream of stop codons in many genes in *Drosophila* species. This is strong evidence that stop codon readthrough has important functional roles in these organisms.
18. Prat, L. *et al.* Carbon source-dependent expansion of the genetic code in bacteria. *Proc. Natl Acad. Sci. USA* **109**, 21070–21075 (2012).
This paper describes a microorganism that uses alternative genetic codes depending on its environment.
19. Lang, B. F. *et al.* Massive programmed translational jumping in mitochondria. *Proc. Natl Acad. Sci. USA* **111**, 5926–5931 (2014).
From 1988 until the publication of this paper in 2014, only a single naturally occurring case of translational bypassing was known. In this work the authors describe numerous cases of such bypassing that occur during the translation of mRNA in the mitochondria of the yeast *Magnusiomyces capitatus*.
20. Knight, R. D., Freeland, S. J. & Landweber, L. F. Rewiring the keyboard: evolvability of the genetic code. *Nat. Rev. Genet.* **2**, 49–58 (2001).
21. Ambrogelly, A., Palioura, S. & Soll, D. Natural expansion of the genetic code. *Nat. Chem. Biol.* **3**, 29–35 (2007).
22. Atkins, J. F. & Gesteland, R. F. *Recoding: Expansion of Decoding Rules Enriches Gene Expression* (Springer, 2010).
23. Sharma, V. *et al.* A pilot study of bacterial genes with disrupted ORFs reveals a surprising profusion of protein sequence recoding mediated by ribosomal frameshifting and transcriptional realignment. *Mol. Biol. Evol.* **28**, 3195–3211 (2011).
This work introduces a scheme for the large-scale identification of evolutionary conserved recoded genes in bacterial genomes.
24. Antonov, I., Coakley, A., Atkins, J. F., Baranov, P. V. & Borodovsky, M. Identification of the nature of reading frame transitions observed in prokaryotic genomes. *Nucleic Acids Res.* **41**, 6514–6530 (2013).
25. Antonov, I., Baranov, P. & Borodovsky, M. GeneTack database: genes with frameshifts in prokaryotic genomes and eukaryotic mRNA sequences. *Nucleic Acids Res.* **41**, D152–D156 (2013).
26. Ivanova, N. N. *et al.* Stop codon reassignments in the wild. *Science* **344**, 909–913 (2014).
In this work the authors used metagenome sequences to analyse the occurrence of stop codon reassignment in organisms from different environments. Among several intriguing findings was a bacteriophage that modifies the genetic code of its host during infection.
27. Ohama, T. *et al.* Non-universal decoding of the leucine codon CUG in several Candida species. *Nucleic Acids Res.* **21**, 4039–4045 (1993).
28. Duarte, I., Nabuurs, S. B., Magno, R. & Huynen, M. Evolution and diversification of the organellar release factor family. *Mol. Biol. Evol.* **29**, 3497–3512 (2012).
29. Bove, J. M. Molecular features of mollicutes. *Clin. Infect. Dis.* **17** (Suppl. 1), S10–S31 (1993).
30. McCutcheon, J. P., McDonald, B. R. & Moran, N. A. Origin of an alternative genetic code in the extremely small and GC-rich genome of a bacterial symbiont. *PLoS Genet.* **5**, e1000565 (2009).
31. Campbell, J. H. *et al.* UGA is an additional glycine codon in uncultured SR1 bacteria from the human microbiota. *Proc. Natl Acad. Sci. USA* **110**, 5540–5545 (2013).
32. Osawa, S. & Jukes, T. H. Codon reassignment (codon capture) in evolution. *J. Mol. Evol.* **28**, 271–278 (1989).
33. Chater, K. F. & Chandra, G. The use of the rare UUA codon to define "expression space" for genes involved in secondary metabolism, development and environmental adaptation in streptomyces. *J. Microbiol.* **46**, 1–11 (2008).
34. Nakabachi, A. *et al.* The 160-kilobase genome of the bacterial endosymbiont *Carsonella*. *Science* **314**, 267 (2006).
35. Schultz, D. W. & Yarus, M. On malleability in the genetic code. *J. Mol. Evol.* **42**, 597–601 (1996).
36. Suzuki, T., Ueda, T. & Watanabe, K. The 'polysemous' codon — a codon with multiple amino acid assignment caused by dual specificity of tRNA identity. *EMBO J.* **16**, 1122–1134 (1997).
37. Santos, M. A., Ueda, T., Watanabe, K. & Tuite, M. F. The non-standard genetic code of *Candida* spp.: an evolving genetic code or a novel mechanism for adaptation? *Mol. Microbiol.* **26**, 423–431 (1997).
38. Netzer, N. *et al.* Innate immune and chemically triggered oxidative stress modifies translational fidelity. *Nature* **462**, 522–526 (2009).
39. Lee, J. Y. *et al.* Promiscuous methionyl-tRNA synthetase mediates adaptive mistranslation to protect cells against oxidative stress. *J. Cell Sci.* **127**, 4234–4245 (2014).
40. Namy, O. *et al.* Adding pyrrolysine to the *Escherichia coli* genetic code. *FEBS Lett.* **581**, 5282–5288 (2007).
41. Longstaff, D. G., Blight, S. K., Zhang, L., Green-Church, K. B. & Krzycki, J. A. *In vivo* contextual requirements for UAG translation as pyrrolysine. *Mol. Microbiol.* **63**, 229–241 (2007).
42. Zhang, Y., Baranov, P. V., Atkins, J. F. & Gladyshev, V. N. Pyrrolysine and selenocysteine use dissimilar decoding strategies. *J. Biol. Chem.* **280**, 20740–20751 (2005).
43. King, G. M. Methanogenesis from methylated amines in a hypersaline algal mat. *Appl. Environ. Microbiol.* **54**, 130–136 (1988).
44. Atkins, J. F. & Baranov, P. V. The distinction between recoding and codon reassignment. *Genetics* **185**, 1535–1536 (2010).
45. Firth, A. E. & Brierley, I. Non-canonical translation in RNA viruses. *J. Gen. Virol.* **93**, 1385–1409 (2012).
46. Dreher, T. W. & Miller, W. A. Translational control in positive strand RNA plant viruses. *Virology* **344**, 185–197 (2006).
47. Steneberg, P. & Samakovlis, C. A novel stop codon readthrough mechanism produces functional Headcase protein in *Drosophila* trachea. *EMBO Rep.* **2**, 593–597 (2001).
48. Namy, O. *et al.* Identification of stop codon readthrough genes in *Saccharomyces cerevisiae*. *Nucleic Acids Res.* **31**, 2289–2296 (2003).
49. Namy, O., Duchateau-Nguyen, G. & Rousset, J. P. Translational readthrough of the PDE2 stop codon modulates cAMP levels in *Saccharomyces cerevisiae*. *Mol. Microbiol.* **43**, 641–652 (2002).
50. Robinson, D. N. & Cooley, L. Examination of the function of two kelch proteins generated by stop codon suppression. *Development* **124**, 1405–1417 (1997).
51. Klagges, B. R. *et al.* Invertebrate synapsins: a single gene codes for several isoforms in *Drosophila*. *J. Neurosci.* **16**, 3154–3165 (1996).
52. Dunn, J. G., Foo, C. K., Belletier, N. G., Gavis, E. R. & Weissman, J. S. Ribosome profiling reveals pervasive and regulated stop codon readthrough in Drosophila melanogaster. *eLife* **2**, e01179 (2013).
In this work the authors used ribosome profiling to analyse the protein-coding potential of the *Drosophila melanogaster* genome and confirmed the widespread occurrence of stop codon readthrough in this organism.
53. Loughran, G. *et al.* Evidence of efficient stop codon readthrough in four mammalian genes. *Nucleic Acids Res.* **42**, 8928–8938 (2014).
54. Eswarappa, S. M. *et al.* Programmed translational readthrough generates antiangiogenic VEGF-Ax. *Cell* **157**, 1605–1618 (2014).
55. Schueren, F. *et al.* Peroxisomal lactate dehydrogenase is generated by translational readthrough in mammals. *eLife* **3**, e03640 (2014).
56. Stiebler, A. C. *et al.* Ribosomal readthrough at a short UGA stop codon context triggers dual localization of metabolic enzymes in fungi and animals. *PLoS Genet.* **10**, e1004685 (2014).
57. Pavlov, M. Y. *et al.* A direct estimation of the context effect on the efficiency of termination. *J. Mol. Biol.* **284**, 579–590 (1998).
58. Bonetti, B., Fu, L., Moon, J. & Bedwell, D. M. The efficiency of translation termination is determined by a synergistic interplay between upstream and downstream sequences in *Saccharomyces cerevisiae*. *J. Mol. Biol.* **251**, 334–345 (1995).
59. Namy, O., Hatin, I. & Rousset, J. P. Impact of the six nucleotides downstream of the stop codon on translation termination. *EMBO Rep.* **2**, 787–793 (2001).
60. Skuzeski, J. M., Nichols, L. M., Gesteland, R. F. & Atkins, J. F. The signal for a leaky UAG stop codon in several plant viruses includes the two downstream codons. *J. Mol. Biol.* **218**, 365–373 (1991).
61. Firth, A. E., Wills, N. M., Gesteland, R. F. & Atkins, J. F. Stimulation of stop codon readthrough: frequent presence of an extended 3' RNA structural element. *Nucleic Acids Res.* **39**, 6679–6691 (2011).
62. Lee, S. R. *et al.* Mammalian thioredoxin reductase: oxidation of the C-terminal cysteine/selenocysteine active site forms a thioselenide, and replacement of selenium with sulfur markedly reduces catalytic activity. *Proc. Natl Acad. Sci. USA* **97**, 2521–2526 (2000).
63. Zhong, L., Arner, E. S. & Holmgren, A. Structure and mechanism of mammalian thioredoxin reductase: the active site is a redox-active selenolthiol/ selenenylsulfide formed from the conserved cysteine–selenocysteine sequence. *Proc. Natl Acad. Sci. USA* **97**, 5854–5859 (2000).
64. Heider, J., Baron, C. & Bock, A. Coding from a distance: dissection of the mRNA determinants required for the incorporation of selenocysteine into protein. *EMBO J.* **11**, 3759–3766 (1992).
65. Berry, M. J., Banu, L., Harney, J. W. & Larsen, P. R. Functional characterization of the eukaryotic SECIS elements which direct selenocysteine insertion at UGA codons. *EMBO J.* **12**, 3315–3322 (1993).
This work describes discovery of the SECIS element, an RNA structure in 3' UTRs of eukaryotic mRNAs that is required for Sec incorporation.
66. Howard, M. T. *et al.* Recoding elements located adjacent to a subset of eukaryal selenocysteine-specifying UGA codons. *EMBO J.* **24**, 1596–1607 (2005).

67. Howard, M. T., Moyle, M. W., Aggarwal, G., Carlson, B. A. & Anderson, C. B. A recoding element that stimulates decoding of UGA codons by Sec tRNA$^{[Ser]Sec}$. *RNA* **13**, 912–920 (2007).
68. Squires, J. E. & Berry, M. J. Eukaryotic selenoprotein synthesis: mechanistic insight incorporating new factors and new functions for old factors. *IUBMB Life* **60**, 232–235 (2008).
69. Driscoll, D. M. & Copeland, P. R. Mechanism and regulation of selenoprotein synthesis. *Annu. Rev. Nutr.* **23**, 17–40 (2003).
70. Labunskyy, V. M., Hatfield, D. L. & Gladyshev, V. N. Selenoproteins: molecular pathways and physiological roles. *Physiol. Rev.* **94**, 739–777 (2014).
71. Yoshizawa, S. & Bock, A. The many levels of control on bacterial selenoprotein synthesis. *Biochim. Biophys. Acta* **1790**, 1404–1414 (2009).
72. Rother, M., Resch, A., Wilting, R. & Bock, A. Selenoprotein synthesis in archaea. *Biofactors* **14**, 75–83 (2001).
73. Gursinsky, T., Jager, J., Andreesen, J. R. & Sohling, B. A selDABC cluster for selenocysteine incorporation in Eubacterium acidaminophilum. *Arch. Microbiol.* **174**, 200–212 (2000).
74. Otero, L. *et al.* Adjustments, extinction, and remains of selenocysteine incorporation machinery in the nematode lineage. *RNA* **20**, 1023–1034 (2014).
75. Chapple, C. E. & Guigo, R. Relaxation of selective constraints causes independent selenoprotein extinction in insect genomes. *PLoS ONE* **3**, e2968 (2008).
76. Hill, K. E., Lloyd, R. S., Yang, J. G., Read, R. & Burk, R. F. The cDNA for rat selenoprotein P contains 10 TGA codons in the open reading frame. *J. Biol. Chem.* **266**, 10050–10053 (1991).
77. Lobanov, A. V., Hatfield, D. L. & Gladyshev, V. N. Reduced reliance on the trace element selenium during evolution of mammals. *Genome Biol.* **9**, R62 (2008).
78. Berry, M. & Howard, M. in *Recoding: Expansion of Decoding Rules Enriches Gene Expression* (eds Atkins, J. F. & Gesteland, R. F.) 29–52 (Springer New York, 2010).
79. Xu, X. M. *et al.* Targeted insertion of cysteine by decoding UGA codons with mammalian selenocysteine machinery. *Proc. Natl Acad. Sci. USA* **107**, 21430–21434 (2010).
80. Hoffman, D. C., Anderson, R. C., DuBois, M. L. & Prescott, D. M. Macronuclear gene-sized molecules of hypotrichs. *Nucleic Acids Res.* **23**, 1279–1283 (1995).
81. Turanov, A. A. *et al.* Genetic code supports targeted insertion of two amino acids by one codon. *Science* **323**, 259–261 (2009).
This work describes a unique case of codon redefinition that involves a sense codon.
82. Lobanov, A. V., Kryukov, G. V., Hatfield, D. L. & Gladyshev, V. N. Is there a twenty third amino acid in the genetic code? *Trends Genet.* **22**, 357–360 (2006).
83. Fujita, M., Mihara, H., Goto, S., Esaki, N. & Kanehisa, M. Mining prokaryotic genomes for unknown amino acids: a stop-codon-based approach. *BMC Bioinformatics* **8**, 225 (2007).
84. Baranov, P. V., Gesteland, R. F. & Atkins, J. F. Release factor 2 frameshifting sites in different bacteria. *EMBO Rep.* **3**, 373–377 (2002).
85. Bekaert, M., Atkins, J. F. & Baranov, P. V. ARFA: a program for annotating bacterial release factor genes, including prediction of programmed ribosomal frameshifting. *Bioinformatics* **22**, 2463–2465 (2006).
86. Ivanov, I. P. & Atkins, J. F. Ribosomal frameshifting in decoding antizyme mRNAs from yeast and protists to humans: close to 300 cases reveal remarkable diversity despite underlying conservation. *Nucleic Acids Res.* **35**, 1842–1858 (2007).
This is a comprehensive survey of ribosomal frameshifting sites and stimulatory elements in antizyme mRNAs occurring in species from a large phylogenetic spectrum of organisms.
87. Gupta, P., Kannan, K., Mankin, A. S. & Vazquez-Laslop, N. Regulation of gene expression by macrolide-induced ribosomal frameshifting. *Mol. Cell* **52**, 629–642 (2013).
88. Cobucci-Ponzano, B., Rossi, M. & Moracci, M. Translational recoding in archaea. *Extremophiles* **16**, 793–803 (2012).
89. Dinman, J. D. Control of gene expression by translational recoding. *Adv. Protein Chem. Struct. Biol.* **86**, 129–149 (2012).
90. Baranov, P. V., Gesteland, R. F. & Atkins, J. F. Recoding: translational bifurcations in gene expression. *Gene* **286**, 187–201 (2002).
91. Baranov, P. V., Fayet, O., Hendrix, R. W. & Atkins, J. F. Recoding in bacteriophages and bacterial IS elements. *Trends Genet.* **22**, 174–181 (2006).
92. Farabaugh, P. J. *Programmed Alternative Reading of the Genetic Code* (R. G. Landes, 1997).
93. Namy, O., Rousset, J. P., Napthine, S. & Brierley, I. Reprogrammed genetic decoding in cellular gene expression. *Mol. Cell* **13**, 157–168 (2004).
94. Firth, A. E., Bekaert, M. & Baranov, P. V. in *Recoding: Expansion of Decoding Rules Enriches Gene Expression* 435–461 (Springer New York, 2010).
95. Jiang, H. *et al.* Orsay virus utilizes ribosomal frameshifting to express a novel protein that is incorporated into virions. *Virology* **450–451**, 213–221 (2014).
96. Fang, Y. *et al.* Efficient -2 frameshifting by mammalian ribosomes to synthesize an additional arterivirus protein. *Proc. Natl Acad. Sci. USA* **109**, E2920–E2928 (2012).
97. Loughran, G., Firth, A. E. & Atkins, J. F. Ribosomal frameshifting into an overlapping gene in the 2B-encoding region of the cardiovirus genome. *Proc. Natl Acad. Sci. USA* **108**, E1111–E1119 (2011).
98. Firth, A. E., Blitvich, B. J., Wills, N. M., Miller, C. L. & Atkins, J. F. Evidence for ribosomal frameshifting and a novel overlapping gene in the genomes of insect-specific flaviviruses. *Virology* **399**, 153–166 (2010).
99. Melian, E. B. *et al.* NS1′ of flaviviruses in the Japanese encephalitis virus serogroup is a product of ribosomal frameshifting and plays a role in viral neuroinvasiveness. *J. Virol.* **84**, 1641–1647 (2010).
100. Firth, A. E. & Atkins, J. F. A conserved predicted pseudoknot in the NS2A-encoding sequence of West Nile and Japanese encephalitis flaviviruses suggests NS1′ may derive from ribosomal frameshifting. *Virol. J.* **6**, 14 (2009).
101. Firth, A. E., Chung, B. Y., Fleeton, M. N. & Atkins, J. F. Discovery of frameshifting in Alphavirus 6K resolves a 20-year enigma. *Virol. J.* **5**, 108 (2008).
102. Firth, A. E. *et al.* Ribosomal frameshifting used in influenza A virus expression occurs within the sequence UCC_UUU_CGU and is in the +1 direction. *Open Biol.* **2**, 120109 (2012).
103. Jagger, B. W. *et al.* An overlapping protein-coding region in influenza A virus segment 3 modulates the host response. *Science* **337**, 199–204 (2012).
104. Gao, X., Havecker, E. R., Baranov, P. V., Atkins, J. F. & Voytas, D. F. Translational recoding signals between gag and pol in diverse LTR retrotransposons. *RNA* **9**, 1422–1430 (2003).
105. Shigemoto, K. *et al.* Identification and characterisation of a developmentally regulated mammalian gene that utilises -1 programmed ribosomal frameshifting. *Nucleic Acids Res.* **29**, 4079–4088 (2001).
106. Wills, N. M., Moore, B., Hammer, A., Gesteland, R. F. & Atkins, J. F. A functional -1 ribosomal frameshift signal in the human paraneoplastic *Ma3* gene. *J. Biol. Chem.* **281**, 7082–7088 (2006).
107. Lin, M. F. *et al.* Revisiting the protein-coding gene catalog of *Drosophila melanogaster* using 12 fly genomes. *Genome Res.* **17**, 1823–1836 (2007).
108. Baranov, P. V. *et al.* Programmed ribosomal frameshifting in the expression of the regulator of intestinal stem cell proliferation, adenomatous polyposis coli (APC). *RNA Biol.* **8**, 637–647 (2011).
109. Ingolia, N. T. Ribosome profiling: new views of translation, from single codons to genome scale. *Nat. Rev. Genet.* **15**, 205–213 (2014).
110. Michel, A. M. *et al.* Observation of dually decoded regions of the human genome using ribosome profiling data. *Genome Res.* **22**, 2219–2229 (2012).
111. Shah, A. A. *et al.* Computational identification of putative programmed translational frameshift sites. *Bioinformatics* **18**, 1046–1053 (2002).
112. Gurvich, O. L. *et al.* Sequences that direct significant levels of frameshifting are frequent in coding regions of *Escherichia coli*. *EMBO J.* **22**, 5941–5950 (2003).
113. Jacobs, J. L., Belew, A. T., Rakauskaite, R. & Dinman, J. D. Identification of functional, endogenous programmed -1 ribosomal frameshift signals in the genome of *Saccharomyces cerevisiae*. *Nucleic Acids Res.* **35**, 165–174 (2007).
114. Sharma, V. *et al.* Analysis of tetra- and hepta-nucleotides motifs promoting -1 ribosomal frameshifting in *Escherichia coli*. *Nucleic Acids Res.* **42**, 7210–7225 (2014).
115. Belew, A. T., Hepler, N. L., Jacobs, J. L. & Dinman, J. D. PRFdb: a database of computationally predicted eukaryotic programmed -1 ribosomal frameshift signals. *BMC Genomics* **9**, 339 (2008).
116. Belew, A. T., Advani, V. M. & Dinman, J. D. Endogenous ribosomal frameshift signals operate as mRNA destabilizing elements through at least two molecular pathways in yeast. *Nucleic Acids Res.* **39**, 2799–2808 (2011).
117. Belew, A. T. *et al.* Ribosomal frameshifting in the CCR5 mRNA is regulated by miRNAs and the NMD pathway. *Nature* **512**, 265–269 (2014).
118. Weiss, R. B., Dunn, D. M., Dahlberg, A. E., Atkins, J. F. & Gesteland, R. F. Reading frame switch caused by base-pair formation between the 3′ end of 16S rRNA and the mRNA during elongation of protein synthesis in *Escherichia coli*. *EMBO J.* **7**, 1503–1507 (1988).
119. Larsen, B., Wills, N. M., Gesteland, R. F. & Atkins, J. F. rRNA-mRNA base pairing stimulates a programmed -1 ribosomal frameshift. *J. Bacteriol.* **176**, 6842–6851 (1994).
120. Prere, M. F., Canal, I., Wills, N. M., Atkins, J. F. & Fayet, O. The interplay of mRNA stimulatory signals required for AUU-mediated initiation and programmed -1 ribosomal frameshifting in decoding of transposable element IS911. *J. Bacteriol.* **193**, 2735–2744 (2011).
121. Gurvich, O. L., Nasvall, S. J., Baranov, P. V., Bjork, G. R. & Atkins, J. F. Two groups of phenylalanine biosynthetic operon leader peptides genes: a high level of apparently incidental frameshifting in decoding *Escherichia coli* pheL. *Nucleic Acids Res.* **39**, 3079–3092 (2011).
122. Yordanova, M. M., Wu, C., Andreev, D. E., Sachs, M. S. & Atkins, J. F. A nascent peptide signal responsive to endogenous levels of polyamines acts to stimulate regulatory frameshifting on antizyme mRNA. *J. Biol. Chem.* http://dx.doi.org/10.1074/jbc.M115.647065 (2015).
123. Kim, H. K. *et al.* A frameshifting stimulatory stem loop destabilizes the hybrid state and impedes ribosomal translocation. *Proc. Natl Acad. Sci. USA* **111**, 5538–5543 (2014).
124. Yu, C. H., Noteborn, M. H., Pleij, C. W. & Olsthoorn, R. C. Stem–loop structures can effectively substitute for an RNA pseudoknot in -1 ribosomal frameshifting. *Nucleic Acids Res.* **39**, 8952–8959 (2011).
125. Mazauric, M. H., Licznar, P., Prere, M. F., Canal, I. & Fayet, O. Apical loop–internal loop RNA pseudoknots: a new type of stimulator of -1 translational frameshifting in bacteria. *J. Biol. Chem.* **283**, 20421–20432 (2008).
126. Brierley, I., Digard, P. & Inglis, S. C. Characterization of an efficient coronavirus ribosomal frameshifting signal: requirement for an RNA pseudoknot. *Cell* **57**, 537–547 (1989).
This is the first work that demonstrated the ability of RNA pseudoknot to promote −1 ribosomal frameshifting at 'slippery' patterns.
127. Baranov, P. V. *et al.* Programmed ribosomal frameshifting in decoding the SARS-CoV genome. *Virology* **332**, 498–510 (2005).
128. Plant, E. P. *et al.* A three-stemmed mRNA pseudoknot in the SARS coronavirus frameshift signal. *PLoS Biol.* **3**, e172 (2005).
129. Su, M. C., Chang, C. T., Chu, C. H., Tsai, C. H. & Chang, K. Y. An atypical RNA pseudoknot stimulator and an upstream attenuation signal for -1 ribosomal frameshifting of SARS coronavirus. *Nucleic Acids Res.* **33**, 4265–4275 (2005).
130. Chou, M. Y. & Chang, K. Y. An intermolecular RNA triplex provides insight into structural determinants for the pseudoknot stimulator of -1 ribosomal frameshifting. *Nucleic Acids Res.* **38**, 1676–1685 (2010).
131. Chen, G., Chang, K. Y., Chou, M. Y., Bustamante, C. & Tinoco, I. Jr. Triplex structures in an RNA pseudoknot enhance mechanical stability and increase efficiency of -1 ribosomal frameshifting. *Proc. Natl Acad. Sci. USA* **106**, 12706–12711 (2009).
132. Herold, J. & Siddell, S. G. An 'elaborated' pseudoknot is required for high frequency frameshifting during translation of HCV 229E polymerase mRNA. *Nucleic Acids Res.* **21**, 5838–5842 (1993).
133. Endoh, T. & Sugimoto, N. Unusual -1 ribosomal frameshift caused by stable RNA G-quadruplex in open reading frame. *Anal. Chem.* **85**, 11435–11439 (2013).

134. Yu, C. H., Teulade-Fichou, M. P. & Olsthoorn, R. C. Stimulation of ribosomal frameshifting by RNA G-quadruplex structures. *Nucleic Acids Res.* **42**, 1887–1892 (2014).
135. Tajima, Y., Iwakawa, H. O., Kaido, M., Mise, K. & Okuno, T. A long-distance RNA–RNA interaction plays an important role in programmed -1 ribosomal frameshifting in the translation of p88 replicase protein of Red clover necrotic mosaic virus. *Virology* **417**, 169–178 (2011).
136. Barry, J. K. & Miller, W. A. A -1 ribosomal frameshift element that requires base pairing across four kilobases suggests a mechanism of regulating ribosome and replicase traffic on a viral RNA. *Proc. Natl Acad. Sci. USA* **99**, 11133–11138 (2002).
137. Giedroc, D. P. & Cornish, P. V. Frameshifting RNA pseudoknots: structure and mechanism. *Virus Res.* **139**, 193–208 (2009).
138. Chung, B. Y., Firth, A. E. & Atkins, J. F. Frameshifting in alphaviruses: a diversity of 3′ stimulatory structures. *J. Mol. Biol.* **397**, 448–456 (2010).
139. Brierley, I., Gilbert, R. J. & Pennell, S. RNA pseudoknots and the regulation of protein synthesis. *Biochem. Soc. Trans.* **36**, 684–689 (2008).
140. Li, Y. *et al.* Transactivation of programmed ribosomal frameshifting by a viral protein. *Proc. Natl Acad. Sci. USA* **111**, E2172–E2181 (2014).
141. Howard, M. T., Gesteland, R. F. & Atkins, J. F. Efficient stimulation of site-specific ribosome frameshifting by antisense oligonucleotides. *RNA* **10**, 1653–1661 (2004).
142. Olsthoorn, R. C. *et al.* Novel application of sRNA: stimulation of ribosomal frameshifting. *RNA* **10**, 1702–1703 (2004).
143. Kurian, L., Palanimurugan, R., Godderz, D. & Dohmen, R. J. Polyamine sensing by nascent ornithine decarboxylase antizyme stimulates decoding of its mRNA. *Nature* **477**, 490–494 (2011).
144. Plant, E. P., Rakauskaite, R., Taylor, D. R. & Dinman, J. D. Achieving a golden mean: mechanisms by which coronaviruses ensure synthesis of the correct stoichiometric ratios of viral proteins. *J. Virol.* **84**, 4330–4340 (2010).
145. Stochmanski, S. J. *et al.* Expanded *ATXN3* frameshifting events are toxic in *Drosophila* and mammalian neuron models. *Hum. Mol. Genet.* **21**, 2211–2218 (2012).
146. Wills, N. M. & Atkins, J. F. The potential role of ribosomal frameshifting in generating aberrant proteins implicated in neurodegenerative diseases. *RNA* **12**, 1149–1153 (2006).
147. Toulouse, A. *et al.* Ribosomal frameshifting on *MJD-1* transcripts with long CAG tracts. *Hum. Mol. Genet.* **14**, 2649–2660 (2005).
148. Girstmair, H. *et al.* Depletion of cognate charged transfer RNA causes translational frameshifting within the expanded CAG stretch in huntingtin. *Cell Rep.* **3**, 148–159 (2013).
149. Belcourt, M. F. & Farabaugh, P. J. Ribosomal frameshifting in the yeast retrotransposon Ty: tRNAs induce slippage on a 7 nucleotide minimal site. *Cell* **62**, 339–352 (1990).
This work demonstrated that ribosomal frameshifting in yeast requires a specific sequence of only seven nucleotides to occur at a high efficiency (~40%).
150. Baranov, P. V., Gesteland, R. F. & Atkins, J. F. P-site tRNA is a crucial initiator of ribosomal frameshifting. *RNA* **10**, 221–230 (2004).
151. Clare, J. J., Belcourt, M. & Farabaugh, P. J. Efficient translational frameshifting occurs within a conserved sequence of the overlap between the two genes of a yeast Ty *1* transposon. *Proc. Natl Acad. Sci. USA* **85**, 6816–6820 (1988).
152. Asakura, T. *et al.* Isolation and characterization of a novel actin filament-binding protein from *Saccharomyces cerevisiae*. *Oncogene* **16**, 121–130 (1998).
153. Vimaladithan, A. & Farabaugh, P. J. Special peptidyl-tRNA molecules can promote translational frameshifting without slippage. *Mol. Cell. Biol.* **14**, 8107–8116 (1994).
154. Sundararajan, A., Michaud, W. A., Qian, Q., Stahl, G. & Farabaugh, P. J. Near-cognate peptidyl-tRNAs promote +1 programmed translational frameshifting in yeast. *Mol. Cell* **4**, 1005–1015 (1999).
155. Temperley, R., Richter, R., Dennerlein, S., Lightowlers, R. N. & Chrzanowska-Lightowlers, Z. M. Hungry codons promote frameshifting in human mitochondrial ribosomes. *Science* **327**, 301 (2010).
156. Young, D. J. *et al.* Bioinformatic, structural, and functional analyses support release factor-like MTRF1 as a protein able to decode nonstandard stop codons beginning with adenine in vertebrate mitochondria. *RNA* **16**, 1146–1155 (2010).
157. Akabane, S., Ueda, T., Nierhaus, K. H. & Takeuchi, N. Ribosome rescue and translation termination at non-standard stop codons by ICT1 in mammalian mitochondria. *PLoS Genet.* **10**, e1004616 (2014).
158. Klobutcher, L. A. & Farabaugh, P. J. Shifty ciliates: frequent programmed translational frameshifting in euplotids. *Cell* **111**, 763–766 (2002).
This review discusses evidence supporting the frequent utilization of ribosomal frameshifting during protein synthesis in ciliates of the *Euplotes* genus.
159. Vallabhaneni, H., Fan-Minogue, H., Bedwell, D. M. & Farabaugh, P. J. Connection between stop codon reassignment and frequent use of shifty stop frameshifting. *RNA* **15**, 889–897 (2009).
160. Wills, N. M. *et al.* Translational bypassing without peptidyl-tRNA anticodon scanning of coding gap mRNA. *EMBO J.* **27**, 2533–2544 (2008).
161. Herr, A. J., Gesteland, R. F. & Atkins, J. F. One protein from two open reading frames: mechanism of a 50 nt translational bypass. *EMBO J.* **19**, 2671–2680 (2000).
162. Herr, A. J., Atkins, J. F. & Gesteland, R. F. Coupling of open reading frames by translational bypassing. *Annu. Rev. Biochem.* **69**, 343–372 (2000).
163. Weiss, R. B., Huang, W. M. & Dunn, D. M. A nascent peptide is required for ribosomal bypass of the coding gap in bacteriophage T4 gene *60*. *Cell* **62**, 117–126 (1990).
164. Samatova, E., Konevega, A. L., Wills, N. M., Atkins, J. F. & Rodnina, M. V. High-efficiency translational bypassing of non-coding nucleotides specified by mRNA structure and nascent peptide. *Nat. Commun.* **5**, 4459 (2014).
165. Smith, M. C. *et al.* Evolutionary relationships among actinophages and a putative adaptation for growth in *Streptomyces* spp. *J. Bacteriol.* **195**, 4924–4935 (2013).
166. Keiler, K. C. & Ramadoss, N. S. Bifunctional transfer–messenger RNA. *Biochimie* **93**, 1993–1997 (2011).
167. Himeno, H., Kurita, D. & Muto, A. tmRNA-mediated *trans*-translation as the major ribosome rescue system in a bacterial cell. *Front. Genet.* **5**, 66 (2014).
168. Hudson, C. M. & Williams, K. P. The tmRNA website. *Nucleic Acids Res.* **43**, D138–D140. (2015).
169. Hudson, C. M., Lau, B. Y. & Williams, K. P. Ends of the line for tmRNA-SmpB. *Front. Microbiol.* **5**, 421 (2014).
170. Donnelly, M. L. *et al.* Analysis of the aphthovirus 2A/2B polyprotein 'cleavage' mechanism indicates not a proteolytic reaction, but a novel translational effect: a putative ribosomal 'skip'. *J. Gen. Virol.* **82**, 1013–1025 (2001).
In this work the authors demonstrated that ribosomes can interrupt continuous elongation of polypeptide chains without interruption of mRNA decoding, thus producing two peptide products from a single ORF.
171. Donnelly, M. L. *et al.* The 'cleavage' activities of foot-and-mouth disease virus 2A site-directed mutants and naturally occurring '2A-like' sequences. *J. Gen. Virol.* **82**, 1027–1041 (2001).
172. Chin, J. W. Expanding and reprogramming the genetic code of cells and animals. *Annu. Rev. Biochem.* **83**, 379–408 (2014).
173. Haimovich, A. D., Muir, P. & Isaacs, F. J. Genomes by design. *Nat. Rev. Genet.* http://www.dx.doi.org/10.1038/nrg3956 (2015).
174. Timmis, J. N., Ayliffe, M. A., Huang, C. Y. & Martin, W. Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. *Nat. Rev. Genet.* **5**, 123–135 (2004).
175. Kondrashov, A. S. Deleterious mutations and the evolution of sexual reproduction. *Nature* **336**, 435–440 (1988).
176. Bender, A., Hajieva, P. & Moosmann, B. Adaptive antioxidant methionine accumulation in respiratory chain complexes explains the use of a deviant genetic code in mitochondria. *Proc. Natl Acad. Sci. USA* **105**, 16496–16501 (2008).
177. Korkmaz, G., Holm, M., Wiens, T. & Sanyal, S. Comprehensive analysis of stop codon usage in bacteria and its correlation with release factor abundance. *J. Biol. Chem.* **289**, 30334–30342 (2014).
178. Maas, S. Posttranscriptional recoding by RNA editing. *Adv. Protein Chem. Struct. Biol.* **86**, 193–224 (2012).
179. Kiran, A., Loughran, G., O'Mahony, J. J. & Baranov, P. V. Identification of A-to-I RNA editing: dotting the i's in the human transcriptome. *Biochem. (Mosc)* **76**, 915–923 (2011).
180. Mallela, A. & Nishikura, K. A-to-I editing of protein coding and noncoding RNAs. *Crit. Rev. Biochem. Mol. Biol.* **47**, 493–501 (2012).
181. Carlile, T. M. *et al.* Pseudouridine profiling reveals regulated mRNA pseudouridylation in yeast and human cells. *Nature* **515**, 143–146 (2014).
182. Karijolich, J. & Yu, Y. T. Converting nonsense codons into sense codons by targeted pseudouridylation. *Nature* **474**, 395–398 (2011).
183. Larsen, B., Wills, N. M., Nelson, C., Atkins, J. F. & Gesteland, R. F. Nonlinearity in genetic decoding: homologous DNA replicase genes use alternatives of transcriptional slippage or translational frameshifting. *Proc. Natl Acad. Sci. USA* **97**, 1683–1688 (2000).
184. Ivanov, I. P., Firth, A. E., Michel, A. M., Atkins, J. F. & Baranov, P. V. Identification of evolutionarily conserved non-AUG-initiated N-terminal extensions in human coding sequences. *Nucleic Acids Res.* **39**, 4220–4234 (2011).
185. Wojciechowska, M., Olejniczak, M., Galka-Marciniak, P., Jazurek, M. & Krzyzosiak, W. J. RAN translation and frameshifting as translational challenges at simple repeats of human neurodegenerative disorders. *Nucleic Acids Res.* **42**, 11849–11864 (2014).
186. Kearse, M. G. & Todd, P. K. Repeat-associated non-AUG translation and its impact in neurodegenerative disease. *Neurotherapeutics* **11**, 721–731 (2014).
187. Cleary, J. D. & Ranum, L. P. Repeat-associated non-ATG (RAN) translation in neurological disease. *Hum. Mol. Genet.* **22**, R45–R51 (2013).

Acknowledgements
This Review was inspired by the discussions during EMBO Workshop "Recoding: Reprogramming genetic decoding" that took place in Killarney, Ireland, on the 13–18 May 2014. Therefore, the authors are grateful to all participants who shared their ideas and experimental data during this meeting. They apologize to their colleagues whose relevant works were not cited as this article is not intended as a comprehensive review of the topic and the authors aimed to keep it concise. The authors wish to acknowledge support by SFI grants [12/IA/1335 to P.V.B. and 08/IN.I/B1889,12/IP/1492 to J.F.A], P.V.B. is also supported by Wellcome Trust grant [094423].

Competing interests statement
The authors declare no competing interests.

FURTHER INFORMATION
Genome-wide information on protein synthesis (GWIPS-viz): http://gwips.ucc.ie
Recode v. 2.0 Database of translational recoding events: http://recode.ucc.ie
SelenoDB: http://www.selenodb.org/
The Genetic Code Tables: http://www.ncbi.nlm.nih.gov/Taxonomy/Utils/wprintgc.cgi

ALL LINKS ARE ACTIVE IN THE ONLINE PDF