Review

Speeding with control: codon usage, tRNAs, and ribosomes

Eva Maria Novoa[1] **and Lluís Ribas de Pouplana**[1,2]

[1] Institute for Research in Biomedicine (IRB), c/Baldiri Reixac 15-21 08028, Barcelona, Catalonia, Spain
[2] Catalan Institution for Research and Advanced Studies (ICREA), Passeig Lluís Companys 23, 08010 Barcelona, Catalonia, Spain

Codon usage and tRNA abundance are critical parameters for gene synthesis. However, the forces determining codon usage bias within genomes and between organisms, as well as the functional roles of biased codon compositions, remain poorly understood. Similarly, the composition and dynamics of mature tRNA populations in cells in terms of isoacceptor abundances, and the prevalence and function of base modifications are not well understood. As we begin to decipher some of the rules that govern codon usage and tRNA abundances, it is becoming clear that these parameters are a way to not only increase gene expression, but also regulate the speed of ribosomal translation, the efficiency of protein folding, and the coordinated expression of functionally related gene families. Here, we discuss the importance of codon–anticodon interactions in translation regulation and highlight the contribution of non-random codon distributions and post-transcriptional base modifications to this regulation.

Codon usage bias

What is codon usage bias?

Due to the degeneracy of the genetic code, several codons ('synonymous' codons; see Glossary) are translated into the same amino acid. Synonymous codons are used with different frequencies, a phenomenon known as codon bias. Codon bias is a defining characteristic of each genome and is maintained by a balance between selection, mutation, and genetic drift [1–3]. Despite the relative universality of the genetic code and the conservation of the translation machinery across species, codon biases vary dramatically between organisms. Thus, the most frequent or most rare codon in a gene varies both between and within species depending on the gene [1,4].

It is generally accepted that the speed at which ribosomes decode a codon depends on the cellular concentration of the tRNA that recognize it [5–8], although there is some debate about this assumption [9]. Nevertheless, the most abundant codons pair with the most abundant tRNAs and vice versa. As a result, gene codon bias strongly correlates with gene expression levels in organisms as diverse as *Escherichia coli*, *Saccharomyces cerevisiae*, *Caenorhabditis elegans*, *Arabidopsis thaliana*, and *Drosophila melanogaster* [10–15]. It has been shown that the use of particular codons can increase the expression of a gene by more than 1000-fold [16].

Corresponding author: Ribas de Pouplana, L. (lluis.ribas@irbbarcelona.org).
Keywords: codon usage; tRNA; translation efficiency; tRNA modifications.

Why does codon usage bias exist?

The existence of codon bias is explained by two different lines of thought [1]. According to 'selectionists', codon bias contributes to the efficiency and accuracy of amino acid sequence, and this bias is maintained by selection [2,17]. By contrast, 'mutationalists' suggest that codon bias exists because of non-randomness in the mutational patterns, whereby some codons would be more mutable and, therefore, would have lower equilibrium frequencies [18,19]. According to this latter theory, genomic G+C composition is thought to be a major factor affecting codon usage variation [20], given that G+C frequencies can range from <20% to >90% in the third position of codons. These two explanations are not mutually exclusive, and both are supported by several studies (*vide supra*).

A clear association exists between the expression level of a gene and its codon composition, an observation that holds for organisms ranging from bacteria to mammals. It is generally accepted that the variation of codon usage

Glossary

Anticodon: sequence of three nucleotides of a tRNA that is complementary to a given codon.
Codon: sequence of three nucleotides of an mRNA that specifies the amino acid that will be added next during protein synthesis.
Mistranslation: phenomenon that occurs when an amino acid is attached to the wrong tRNA and subsequently misincorporated into the nascent protein.
Preferred codons: subsets of rapidly translated codons that are expected to increase translation efficiency and, therefore, to be over-represented in highly expressed proteins.
Ribosome stalling: ribosome pausing, which is thought to happen for several reasons, including the presence of rare codons, which are decoded more slowly.
Shine-Dalgarno sequence: ribosomal binding site that exists in the mRNA of Bacteria and Archaea, generally located eight base pairs upstream of the start codon AUG.
Translation efficiency: rate of mRNA translation into proteins within cells.
tRNA channeling: direct transfer of tRNAs from the aminoacyl-tRNA synthetases to the elongation factor and ribosomes without dissociation. It also includes the transfer of tRNAs leaving the ribosome to their cognate aminoacyl-tRNA synthetases, which will regenerate newly charged tRNAs ready to use again in protein synthesis.
tRNA decoding capacity: ability of a tRNA to recognize more than one codon from a subset of codons that encode the same amino acid, including both Watson-Crick and wobble base pairings.
tRNA isoacceptors: tRNA molecules that bind to alternate codons encoding the same amino acid residue.
tRNA microarray: specific microarray method to quantify tRNAs based on a fluorescent dye-labeling technique.
tRNA modifications: nucleotide modifications that alter the biophysical and biochemical properties of a tRNA, causing changes in the structure and dynamics of the tRNA to fine-tune its function.
Wobble base-pair: non-Watson-Crick base-pairing between two nucleotides in RNA molecules, but the thermodynamic stability is comparable to that of Watson-Crick base pairs.
Wobble position: third position in the codon, or first position of the tRNA anticodon (base 34).

0168-9525/$ – see front matter http://dx.doi.org/10.1016/j.tig.2012.07.006 Trends in Genetics, November 2012, Vol. 28, No. 11

between genes of the same genome is a product of selection, based on the observation that codon bias is more extreme in highly expressed genes, which are enriched in those codons that match the most abundant cognate tRNAs [15,18,21]. However, whether the codon bias found in highly expressed genes serves to optimize translational efficiency or improve codon reading accuracy has been a topic of active debate [18,19,22,23].

A priori, both translation efficiency and accuracy should be under positive selection. On the one hand, efficient elongation of a transcript might increase its protein yield [16] or provide a global benefit to the cell by freeing up ribosomes that can then translate other messages [24]. On the other hand, accurate elongation would benefit the cell by reducing the costs of useless mistranslation products [22].

Beyond this direct relation between codon composition and translation speed lies a more complex set of parameters that link codon usage and tRNA abundance to gene expression regulation. Factors such as codon autocorrelation [25], clustering of rare codons [26], mRNA secondary structure [24], ribosomal density [6], relative abundance of wobble base pairs [27], presence of Shine-Dalgarno-like features in coding sequences [9], or interactions with modified tRNAs [28] can further contribute to the regulation of gene expression through the phenomenon of synonymous codon bias and tRNA dynamics (Figure 1).

For example, codon usage bias has been linked to the control of cell cycle development [29] and stress-mediated specific responses [30]. Specific tRNAs and, consequently, certain codon compositions are a crucial component in the activation of some genetic programs [31], suggesting a novel layer of genomic regulation that is only now starting to be explored. Similarly, it has been recently shown that the emergence of certain anticodon modification enzymes during evolution has shaped the structure of genomes, contributing to the regulation of the speed of gene translation [28] (Figure 2). In this review, we discuss the latest advances leading to current understanding of how codon usage and tRNA populations evolved not only to optimize gene expression, but also to regulate it.

Codon usage and tRNA

Codon frequencies and tRNA abundances

tRNAs translate codons into amino acids during protein synthesis. Every organism has multiple tRNA species that read the codons for the same amino acid (tRNA isoacceptors). Several reports have shown that synonymous triplet variation across species is driven by the adaptation of codon usage to tRNA abundances or vice versa [15,16,28,32,33]. However, the search for a correlation between tRNA abundance and codon usage has been successful only in some organisms [5,34]. In several species, including many bacteria and eukaryotes, this search has failed [35,36], prompting the proposal that, in the latter organisms, translation efficiency might not be the primary factor influencing codon usage [36,37]. However, it was recently reported that two distinct modifications at the wobble position of certain anticodons are at the core of this apparent lack of correlation [28]. These modifications 'extend' the wobble pairing ability of anticodons and influence the codon usage bias in bacteria and eukaryotes, ultimately affecting codon usage and genomic tRNA compositions. The inclusion of these modifications corrects previously reported discrepancies between codon usage and tRNA abundance across all extant major phylogenetic groups, These results suggest not only that codon usage and tRNA abundances coevolve, but also that the diversification of the genetic code usage in evolution was at least partially driven by the appearance of certain tRNA modification enzymes [28].

Recently, *in vivo* translational speeds for all sense codons from *S. cerevisiae* were determined [38] using genome-wide ribosome profiling data. Surprisingly, similar translational speeds among synonymous codons were found, suggesting that preferentially used codons in highly expressed proteins are not translated faster than non-preferred ones. However, a correlation between codon usage bias and cognate tRNA abundances was indeed observed. These findings suggest that codon usage bias found in highly expressed genes is a product of natural selection for an overall cellular efficiency, rather than a product of stronger selection for translation efficiency in more highly expressed genes.

Variability in tRNA pools

tRNA gene copy number has often been used as a proxy for tRNA abundance in the cell [26,34,35,39]. This approximation has been validated for some unicellular organisms, such as yeast [6] and *E. coli* [5], but recent studies have demonstrated that tissue-specific differences in the expression of tRNA genes exist in more complex organisms [40]. Indeed, microarray-based quantification of cellular tRNAs shows significant variation in their levels among different tissues, both in terms of relative enrichments of specific tRNA isoacceptors and in total tRNA concentration [40]. Importantly, the correlation between relative tRNA abundances and the codon composition of highly expressed, tissue-specific genes was also observed in the different tissues analyzed.

Because mature tRNAs in humans are thought to be very stable, cellular tRNA levels are mostly determined by tRNA transcription rates [41]. tRNAs are transcribed by a multisubunit complex of RNA polymerase III (Pol III), TFIIIB, and TFIIIC [42,43] and are negatively regulated by Maf1, a protein under the control of the mammalian target of rapamycin (mTOR) pathway [44]. Thus, the regulation of tRNA transcript levels is closely linked to cellular conditions, such as nutrient availability and genome integrity.

To explore the evolutionary dynamics of tRNA gene transcription and the variation across different tissues in mammals, Pol III occupancy has been experimentally determined in several tissues from six mammalian species [45]. Pol III binding to different tRNA genes varies substantially in strength and genomic location for different species. However, there is a strong conservation of Pol III occupancy at the genes of grouped tRNA isoacceptor families [45]. These results suggest that, although the usage of individual tRNA genes has evolved rapidly, functional tRNA isoacceptor families have been maintained throughout evolution. This indicates that the major evolutionary

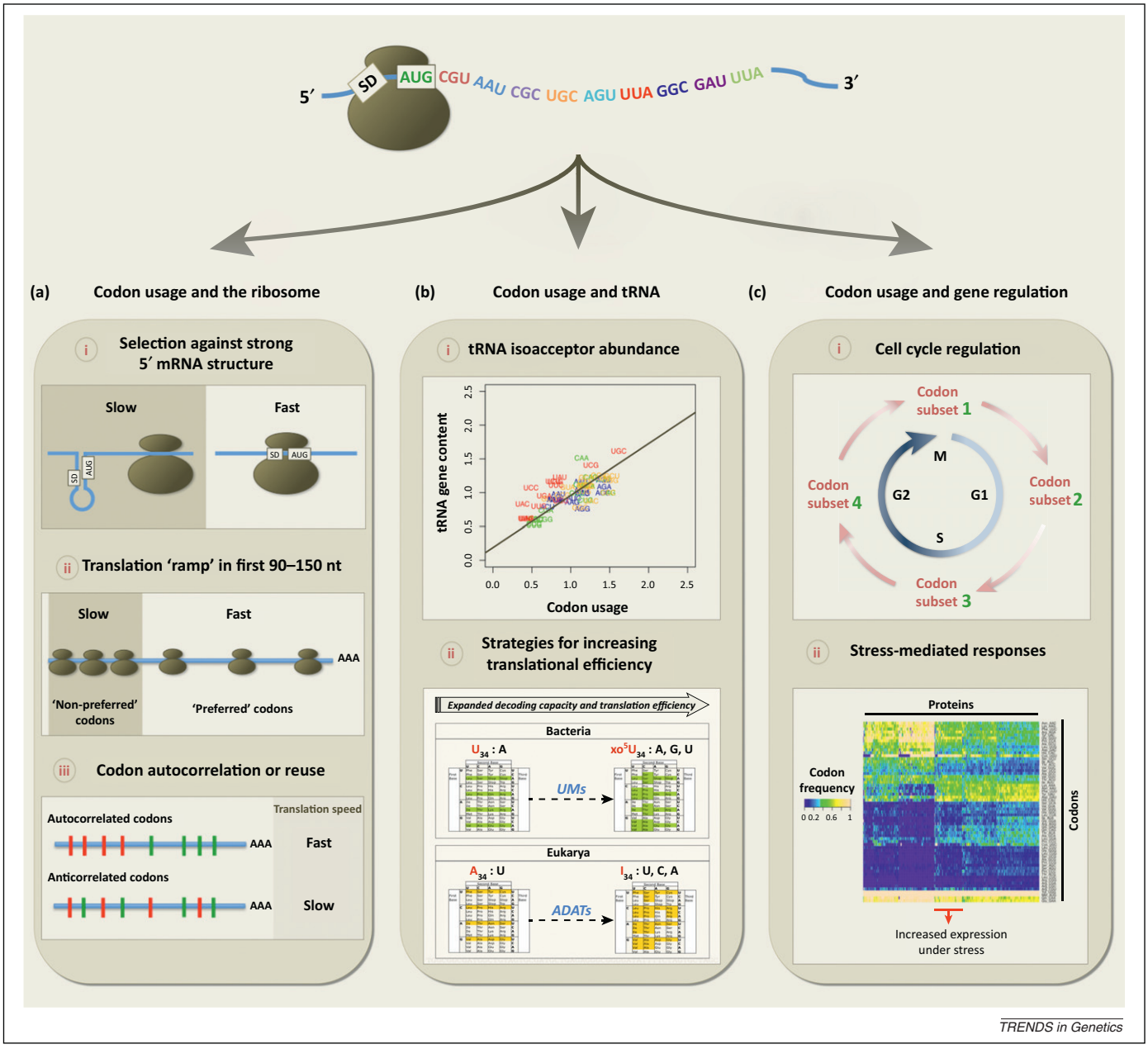

Figure 1. Relevant mechanisms involving the unequal use of synonymous codons and their effect on translation efficiency. **(a)** The distribution of synonymous codons along the gene affects the speed of the ribosome and, consequently, the translation efficiency. As general rules, mRNA transcripts lack strong 5′ secondary structure (i) [24], 'non-preferred' codons cluster at the beginning of the transcript (ii) [6], and autocorrelated codons, which allow tRNA recycling, increase the speed of translation (iii) [25]. **(b)** tRNA gene content tends to correlate with the codon usage bias of highly expressed genes (i) [34]. tRNA gene content biases appear to increase protein translation efficiency by increasing the number of tRNA isoacceptors that are capable of being modified by tRNA modification enzymes, which expands their wobbling capacity. These tRNA modification enzymes differ between bacterial [uridine methyltransferases (UMs)] and eukaryal [adenosine deaminases (ADATs)] species, which in turn have caused differential increases of specific tRNA isoacceptors between kingdoms (ii) [28]. **(c)** The sets of genes that are expressed in each stage of the cell cycle present similar codon covariations, and these differ from those found in other stages, suggesting that the codon preferences change during the cell cycle (i) [29]. Codon preferences may change due to the activity of specific tRNA modification enzymes [e.g., tRNA methyltransferase 9 (Trm9) in *Saccharomyces cerevisiae*]. Under stress conditions, Trm9 modifies a subset of tRNAs and, consequently, their decoding capacities and codon preferences, thus enhancing the expression of a subset of codons that is enriched in proteins that respond to stress (ii) [30].

forces driving relative tRNA abundance and codon composition are conserved across mammals.

Codon usage and the ribosome

Codon distribution and local enrichment

Traditionally, analyses of codon usage for individual genes have only considered the overall codon composition of transcripts. However, patterns of unequal codon distribution along genes exist, and these are thought to be important for the control of ribosome speed and translation stability [6,25]. Indeed, the notion that translation rates can change across different regions of an mRNA transcript has been known for some time [46,47] and has recently gained additional experimental support [48,49].

Recently, a study of the translation efficiency of codons as a function of their location on the transcript [6] reported that, for most genes, the speed of translation is reduced during the first 30–50 codons (known as the 'translation ramp') and then increases for the remainder of the gene. This ramp of poorly adapted codons (i.e., those read by low

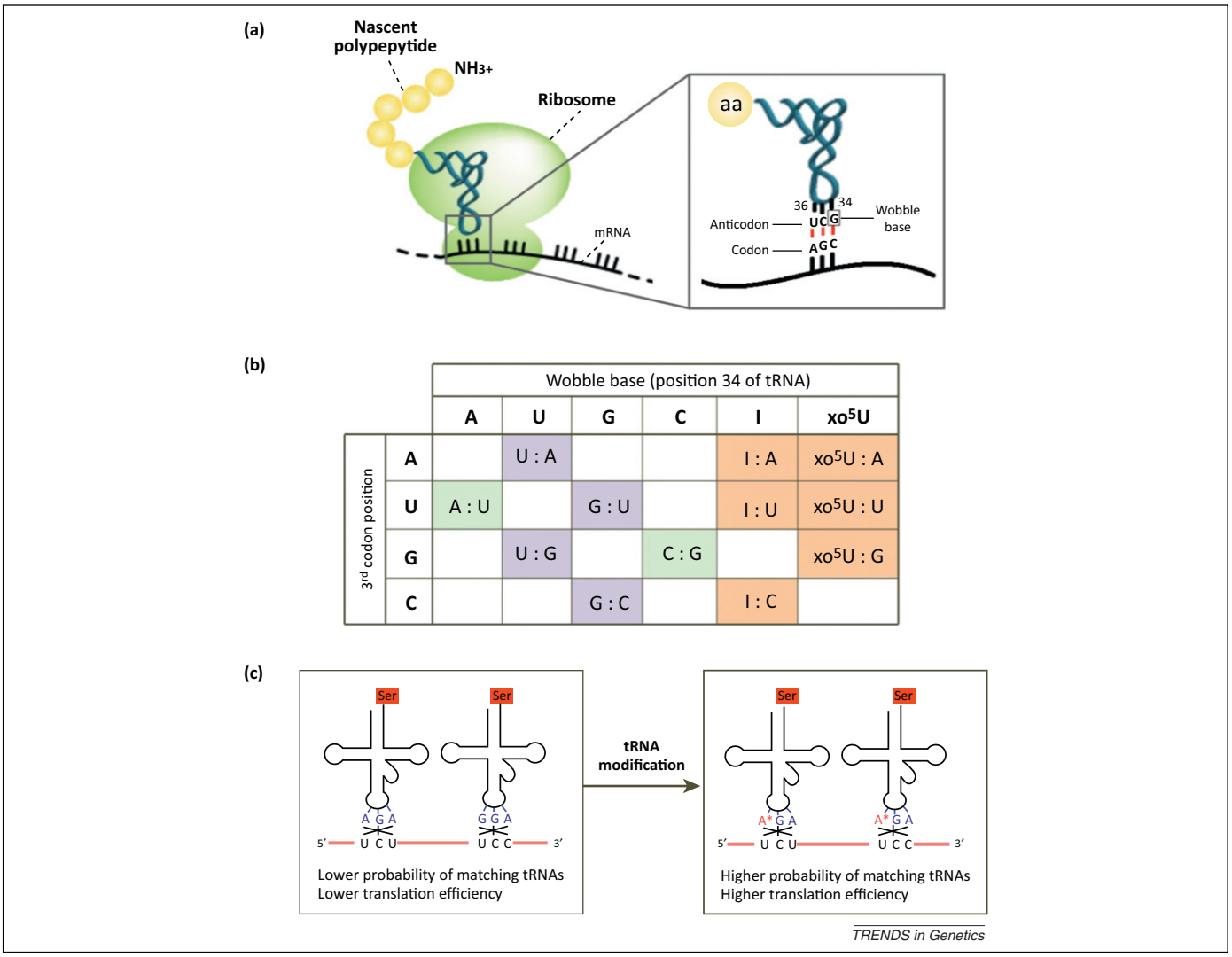

Figure 2. Increased decoding capacity of modified tRNAs at the wobble position. **(a)** The aminoacylated tRNA enters the ribosome and is selected based on the correct pairing of its anticodon bases (positions 34, 35, and 36 of the tRNA) with the respective codon. The wobble base (position 34) recognizes the third position of the mRNA codon, the degenerate codon position. **(b)** Representation of all possible codon–anticodon pairings according to the extended wobble base-pairing rules. A and C in position 34 can only recognize one base (shown in green), whereas G and U can recognize two different bases (shown in purple): one through Watson-Crick pairing and the other through the G:U wobbling. The activities of adenosine deaminases (ADATs; A-to-I conversion) and uridine methyltransferases (UMs; U-to-xo^{5}U conversion) expand the wobbling capacities of base 34, allowing them to pair with three different codon bases (shown in orange). **(c)** Proposed model for the effect of tRNA modification enzymes upon translation efficiency.

abundant tRNAs) would presumably slow elongation at the beginning of the gene, reducing the frequency of ribosomal stalling. This model is further supported by experimentally determined ribosome profile densities along mRNAs [50] and has now been proposed as a general feature of gene translation in both prokaryotic and eukaryotic species.

It has also been shown that once a particular synonymous codon has been used in a transcript, other codons recognized by the same tRNA isoacceptor will be favored in that gene [25]. This observation holds for both frequent and rare codons, and the observed enrichment diminishes as a function of the distance between subsequent synonymous codons. This indicates that sequences optimized for tRNA reuse are expressed more efficiently than are sequences that require different tRNA isoacceptors. In accordance with this model, previous studies have proposed that tRNA diffusion away from the ribosome is slower than translation, and that some tRNA channeling takes place to optimize ribosome function [51]. Specifically, it has been suggested that, after release from the ribosome, tRNAs remain bound to the multi tRNA–synthetase complex [52] or to elongation factors [53], which might themselves be associated with the ribosome. Such a mechanism would effectively raise the local concentration of tRNAs that recognize codons that have already appeared in given transcript. Thus, genes that reuse the same codons, even rare codons, may be more efficiently transcribed.

Codon composition and RNA secondary structure

At the gene level, it is well known that mRNA structure influences translational efficiency. In bacteria, the formation of strong hairpin loops around the Shine-Dalgarno ribosomal binding site and the initiation codon can significantly reduce expression levels [54]. Therefore, strong mRNA structure near the 5′ end of a transcript is generally

thought of as disadvantageous and can inhibit ribosomal translation initiation [55,56]. This is further supported by a study that looked at the expression levels of a synthetic library of GFPs with random synonymous changes in *E. coli* and found that upstream sequence composition influenced mRNA folding near the ribosomal binding site. In fact, it was estimated that the mRNA sequence composition in this region explained more than half of the variation that was found in protein levels [24].

Protein folding regulated by codon composition

Numerous reports indicate that the speed of translation along some transcripts may be critical to the formation of the native structure of a protein. Pausing has been identified during the translation of certain proteins [57–59] and, in many cases, it appears to be caused by local mRNA structures [60], which may be required for the correct folding of the nascent polypeptide [49,61]. These translation effects support the theory of co-translational protein folding and highlight the importance of mRNA sequence and codon usage bias in protein structure formation [62,63]. Indeed, synonymous mutations can have significant consequences in the folding process of the nascent protein and even change the substrate specificity of enzymes [64].

Codon usage and gene regulation

Codon usage and cell cycle control

It is becoming increasingly evident that the use of specific subsets of codons can be a strategy to optimize parameters other than protein synthesis efficiency. For example, previous works in bacteria and fungi demonstrated that functionally related genes that probably need to be expressed at similar levels tend to have similar patterns of codon bias [65,66]. In a recent study, it was shown that certain non-optimal codon compositions were related to cell cycle-dependent oscillations in protein levels [29]. Indeed, cell cycle-regulated genes display different codon preferences, suggesting that codon usage has a role in cell cycle regulation.

The same study also concluded that cell cycle-regulated genes have a strong preference for codons with low codon–anticodon binding affinity [29], based on published thermodynamic data for binding affinities of several possible base pairings [67]. If subsets of functionally related genes exhibit specific biases towards particular codons, then the regulation of the expression of these genes may also be linked to specific codon usage patterns. Thus, it appears that subsets of 'preferentially expressed' genes form coherent groups in terms of codon usage, and that these codon composition 'preferences' change throughout the life cycle of a cell. Similarly, these codon preferences might also be capable of responding to a variety of external stimuli, such as stress [30].

Proteome regulation through modulation of codon-anticodon pairings

As stated above, genes that need to be expressed at similar levels tend to have similar codon biases [65,66]. Importantly, anticodon bases can be customized by tRNA modification enzymes to alter their translation decoding capacity, potentially impacting the subset of 'preferred' codons in the genome. This potential variability in the sets of 'preferred' codons implies that modulating the activity of modification enzymes may be an avenue for regulating the composition of the proteome when needed (Figure 3).

The relation between codon frequency and tRNA abundance is further confounded by the existence of post-transcriptional modifications in tRNA nucleotides, over 100 of which have been recognized and described to date (http://rna-mdb.cas.albany.edu/RNAmods/). Modifications contribute to tRNA folding, structure, and stability, as well as to translation efficiency and amino acid substitution rates [68–70]. Many of the known base modifications are not essential for life and have often been characterized as a mere expansion of the repertoire of the nucleotide bases. Nevertheless, increasing evidence indicates that tRNA modifications can have regulatory roles in cells, especially in response to stress conditions [30,71].

The function of many tRNA modifications, particularly with regards to gene expression regulation, remains unclear. To approach this problem, novel mass spectrometric methods to quantify tRNA modifications with high precision are being used [72]. One such study exposed *S. cerevisiae* to various environmental stresses and analyzed the resulting changes in tRNA modification levels. Interestingly, the prevalence of several tRNA modifications changed as a function of the stress response being activated, suggesting

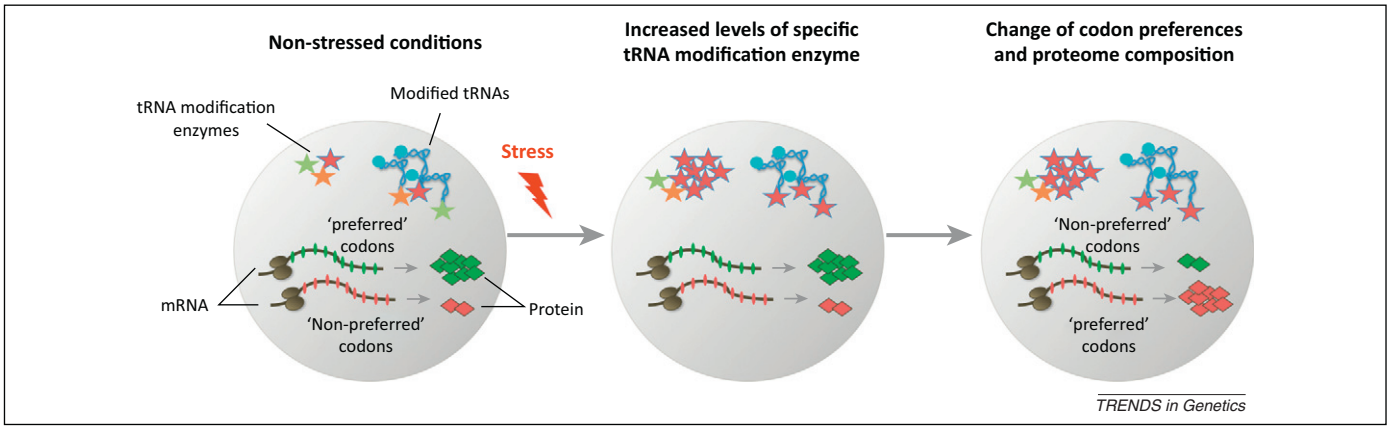

Figure 3. Codon usage bias as a mechanism for tuning the proteome. In response to a particular signal, such as an environmental stress, the levels of a given tRNA modification enzyme change, altering the codon preferences of the tRNA. These changes in turn cause an increase in the protein expression levels of those mRNAs that are found to be enriched in that specific subset of newly 'preferred' codons. Such a mechanism may operate on a set of proteins involved in the specific response to the signal or stress.

Table 1. Characterization of tRNA modifications in *Saccharomyces cerevisiae*[a]

Modification enzyme	tRNA modification	Target tRNAs[c]	tRNA position[c]	Changes in modifications in enzyme-lacking mutants[b]		Function[c]
				Decreased[d]	Increased	
Trm1	m_2^2G	Several	G26	m_2^2G	–	
Trm2	m^5U	All	U54	m^5U	–	Suggested role in tRNA stabilization and maturation
Trm3	Gm	Several	G18	–	–	
Trm4	m^5C	Several	C34, C40, C48, C49	m^5C	–	Suggested role in ribosome biogenesis
Trm5	m^1G/yW	Several	G37	yW	Y, Gm, Um, Am, m_2^2G	Required for yW modification
Trm7	Cm	Several	C32, G34	ncm^5U, yW	–	
Trm8	m^7G	Several	G46	m^7G, (yW)	–	Required to maintain tRNA stability
Trm9	mcm^5U/mcm^5s^2U	Arg(UCU), Glu(UUC)	U34	mcm^5U, mcm^5s^2U	–	Role in stress response; interacts with Trm112
Trm10	m^1G	Several	G9	m^1G	ncm^5U	
Trm11	m^2G	Several	G10	m^2G, (yW)	–	Interacts with Trm112
Trm12	yW	Phe	G14	yW	–	Not methyltransferase
Trm13	Cm	Gly, His, Pro	G4	(Cm)	–	
Trm44	Um	Ser	U44	(Um)	––	
Trm82	m^7G	Several	G46	m^7G, (yW)	m^1G, m^3C, t^6A, m_2^2G, m^2G, m^1I, mcm^5U, mcm^5s^2U	Required to maintain tRNA stability; complexes with Trm8
Tad1	I	Ala	A37	m^1I, yW, (D), (Y)	–	
Mod5	i^6A	Several	A37	i^6A, yW, (D), (Y)	–	
Tan1	ac^4C	Leu, Ser	C12	ac^4C, (yW)	m^1G, m^3C,m^1A, m^2G, m^1I, mcm^5U, m^7G	Acetyltransferase

[a]Abbreviations: A, adenosine; ac, acetyl; C, cytidine; D, dihydrouridine; G, guanosine; I, inosine; m, methyl; mcm, methoxycarbonylmethyl; s, thio; t, threonyl; U, uridine; Y, pseudouridine; yW, wybutosine.

[b]Data from [72].

[c]Data from public databases.

[d]Changes shown in parenthesis correspond to more subtle changes in modification levels compared to those shown without parenthesis.

that the control of tRNA modifications in cellular response pathways is a dynamic process.

Similarly, it has been shown that certain clusters of yeast mRNAs enriched in AGA codons are differentially translated under stress due to an increase in anticodon modifications mediated by tRNA methyltransferase 9 (Trm9) [30]. Therefore, similar transcriptomes may result in different proteome compositions as a consequence of changes in the activity of anticodon modification enzymes [30]. This mechanism is probably not limited to Trm9, and it is possible that other responses are linked to the activity of other tRNA modification enzymes (Figure 3, Table 1).

The redundancy of the genetic code offers an opportunity to fine-tune gene expression levels depending on the usage of synonymous codons. In this regard, functionally related genes (i.e., cell cycle-related genes or stress-response genes) seem to have similar codon usage profiles, suggesting that their translation is somehow favored under certain conditions. Whether this regulation is achieved through changes in tRNA abundance or through the regulation of modifications is something that must be further studied, although supporting evidence for both regulatory mechanisms exists [30,40,72,73].

Two tRNA modification enzymes are known to increase codon-pairing ability: tRNA-dependent adenosine deaminases (ADATs) and tRNA-dependent uridine methyltransferases (UMs) [74,75]. These enzymes expand the wobbling capacity of tRNAs and increase the translation efficiency of the codons recognized by the modified tRNAs. Indeed, highly expressed genes (e.g., ribosomal genes) are found to be most enriched in 'preferred' codons (in this case, those read by tRNAs with modified anticodons), again supporting the possibility that the activity of tRNA modification enzymes constitutes a novel mechanism for post-transcriptional regulation of protein abundance [28]. However, it is an open question whether the activity of these enzymes is regulated in response to specific conditions and, if so, by what means this regulation is accomplished.

Future directions

mRNA sequences contain far more information than just the encoded amino acids. Although the multiple regulatory layers that result from modification of DNA and proteins have been extensively studied, RNA modifications still remain an unexplored territory [76]. In this regard, the complexity of cellular tRNA populations holds great potential for the discovery of new cellular regulatory mechanisms.

It has been shown that certain post-transcriptional RNA modifications can be dynamic and reversible, suggesting that some modifications have functions beyond fine-tuning the structure and function of the RNA [77]. Similarly, tRNA modifications can also be regulated and maintained in distinct cell types and physiological states [76,78].

Information about the range of biological functions of tRNA modifications has only recently begun to emerge [30,72]. Due to their complex nature, post-transcriptional RNA modifications are difficult to study, and understanding of these modifications is sorely lacking compared with

other areas in cell biology. New technologies capable of systems-level analyses of RNA modification changes under diverse cellular conditions will surely bring novel insights into the biosynthesis of tRNA modifications and their role in cellular responses [72,79]. We expect that, during the coming years, these types of approach will shed light on the roles of tRNA modification enzymes in the proteomic changes that accompany transitions in the cell cycle, stress responses, and cell differentiation.

In addition to affecting translation efficiency, codon choice has also been shown to govern translation fidelity by influencing the rate of mistranslation [18,19,22]. Indeed, in the search for the right tRNA, the ribosome might incorrectly bind to a near-cognate tRNA (i.e., a tRNA with one base mismatch relative to the codon), causing the incorporation of a different amino acid. The frequency of this type of mistranslation error has been estimated *in vivo* to range from 10^{-2} in *Bacillus subtilis* [80] to 10^{-5} in yeast cells [81]. The fact that there is a significant correlation between codon conservation and conserved amino acid position suggests that translation accuracy has been under positive selection [22]. Importantly, however, mistranslation might also be beneficial. Recent work has shown that, under certain stress conditions, mistranslation rates increase, leading to increased misincorporation of methionine residues into the mammalian proteome [82]. Moreover, in certain organisms, proteome-wide mistranslation has been shown to increase their fitness under particular environmental conditions [83]. These novel observations suggest that mistranslation evolved as a cellular strategy to adapt to environmental changes and that the codon choice have evolved such that errors can be introduced in nonessential positions of proteins. Exploring the biological significance of mistranslation represents one of the most exciting new directions in this field.

Concluding remarks

Although any given amino acid can be encoded by multiple codons, these 'synonymous' codons are not equally used across genes or genomes. Codon usage has been shown to influence gene expression levels, but the precise rules that govern codon composition remain unclear. Recent efforts have started to uncover specific parameters that affect codon choice, such as codon autocorrelation, codon order, tRNA isoacceptor abundance, or gene coregulation.

In this already complex scenario, tRNA modifications emerge as novel players that can modulate the translation efficiency of codons and, consequently, the expression levels of specific subsets of genes. Indeed, recent studies suggest that tRNA modifications have an important role in genome regulation by specifically enhancing the expression levels of those genes involved in a cellular response. This indicates that the complexity of mature cellular tRNA populations, which has only recently started to be appreciated, holds great potential for the discovery of new cellular regulatory mechanisms.

Acknowledgments

This work has been supported by grant BIO2009-09776 from the Spanish Ministry of Education and Science, and by grant MEPHITIS-223024 from the European Union. EMN is supported by a La Caixa/IRB International PhD Programme Fellowship.

References

1 Hershberg, R. and Petrov, D.A. (2008) Selection on codon bias. *Annu. Rev. Genet.* 42, 287–299
2 Bulmer, M. (1991) The selection–mutation–drift theory of synonymous codon usage. *Genetics* 129, 897–907
3 Rocha, E.P. (2004) Codon usage bias from tRNA's point of view: redundancy, specialization, and efficient decoding for translation optimization. *Genome Res.* 14, 2279–2286
4 Plotkin, J.B. and Kudla, G. (2011) Synonymous but not the same: the causes and consequences of codon bias. *Nat. Rev. Genet.* 12, 32–42
5 Dong, H. *et al.* (1996) Co-variation of tRNA abundance and codon usage in *Escherichia coli* at different growth rates. *J. Mol. Biol.* 260, 649–663
6 Tuller, T. *et al.* (2010) An evolutionarily conserved mechanism for controlling the efficiency of protein translation. *Cell* 141, 344–354
7 Sorensen, M.A. *et al.* (1989) Codon usage determines translation rate in *Escherichia coli*. *J. Mol. Biol.* 207, 365–377
8 Varenne, S. *et al.* (1984) Translation is a non-uniform process. Effect of tRNA availability on the rate of elongation of nascent polypeptide chains. *J. Mol. Biol.* 180, 549–576
9 Li, G.W. *et al.* (2012) The anti-Shine-Dalgarno sequence drives translational pausing and codon choice in bacteria. *Nature* 484, 538–541
10 Goetz, R.M. and Fuglsang, A. (2005) Correlation of codon bias measures with mRNA levels: analysis of transcriptome data from *Escherichia coli*. *Biochem. Biophys. Res. Commun.* 327, 4–7
11 Ghaemmaghami, S. *et al.* (2003) Global analysis of protein expression in yeast. *Nature* 425, 737–741
12 Castillo-Davis, C.I. and Hartl, D.L. (2002) Genome evolution and developmental constraint in *Caenorhabditis elegans*. *Mol. Biol. Evol.* 19, 728–735
13 Duret, L. (2002) Evolution of synonymous codon usage in metazoans. *Curr. Opin. Genet. Dev.* 12, 640–649
14 Gouy, M. and Gautier, C. (1982) Codon usage in bacteria: correlation with gene expressivity. *Nucleic Acids Res.* 10, 7055–7074
15 Ikemura, T. (1985) Codon usage and tRNA content in unicellular and multicellular organisms. *Mol. Biol. Evol.* 2, 13–34
16 Gustafsson, C. *et al.* (2004) Codon bias and heterologous protein expression. *Trends Biotechnol.* 22, 346–353
17 Shields, D.C. and Sharp, P.M. (1987) Synonymous codon usage in *Bacillus subtilis* reflects both translational selection and mutational biases. *Nucleic Acids Res.* 15, 8023–8040
18 Akashi, H. (1994) Synonymous codon usage in *Drosophila melanogaster*: natural selection and translational accuracy. *Genetics* 136, 927–935
19 Stoletzki, N. and Eyre-Walker, A. (2007) Synonymous codon usage in *Escherichia coli*: selection for translational accuracy. *Mol. Biol. Evol.* 24, 374–381
20 Chen, S.L. *et al.* (2004) Codon usage between genomes is constrained by genome-wide mutational processes. *Proc. Natl. Acad. Sci. U.S.A.* 101, 3480–3485
21 Sharp, P.M. and Li, W.H. (1987) The rate of synonymous substitution in enterobacterial genes is inversely related to codon usage bias. *Mol. Biol. Evol.* 4, 222–230
22 Drummond, D.A. and Wilke, C.O. (2008) Mistranslation-induced protein misfolding as a dominant constraint on coding-sequence evolution. *Cell* 134, 341–352
23 Zhou, T. *et al.* (2009) Translationally optimal codons associate with structurally sensitive sites in proteins. *Mol. Biol. Evol.* 26, 1571–1580
24 Kudla, G. *et al.* (2009) Coding-sequence determinants of gene expression in *Escherichia coli*. *Science* 324, 255–258
25 Cannarozzi, G. *et al.* (2010) A role for codon order in translation dynamics. *Cell* 141, 355–367
26 Parmley, J.L. and Huynen, M.A. (2009) Clustering of codons with rare cognate tRNAs in human genes suggests an extra level of expression regulation. *PLoS Genet.* 5, e1000548
27 Stadler, M. and Fire, A. (2011) Wobble base-pairing slows *in vivo* translation elongation in metazoans. *RNA* 17, 2063–2073
28 Novoa, E.M. *et al.* (2012) A role for tRNA modifications in genome structure and codon usage. *Cell* 149, 202–213
29 Frenkel-Morgenstern, M. *et al.* (2012) Genes adopt non-optimal codon usage to generate cell cycle-dependent oscillations in protein levels. *Mol. Syst. Biol.* 8, 572
30 Begley, U. *et al.* (2007) Trm9-catalyzed tRNA modifications link translation to the DNA damage response. *Mol. Cell* 28, 860–870

31 Maraia, R.J. *et al.* (2008) It's a mod mod tRNA world. *Nat. Chem. Biol.* 4, 162–164
32 Kanaya, S. *et al.* (1999) Studies of codon usage and tRNA genes of 18 unicellular organisms and quantification of *Bacillus subtilis* tRNAs: gene expression level and species-specific diversity of codon usage based on multivariate analysis. *Gene* 238, 143–155
33 Percudani, R. *et al.* (1997) Transfer RNA gene redundancy and translational selection in *Saccharomyces cerevisiae*. *J. Mol. Biol.* 268, 322–330
34 Ikemura, T. (1981) Correlation between the abundance of *Escherichia coli* transfer RNAs and the occurrence of the respective codons in its protein genes: a proposal for a synonymous codon choice that is optimal for the *E. coli* translational system. *J. Mol. Biol.* 151, 389–409
35 Kanaya, S. *et al.* (2001) Codon usage and tRNA genes in eukaryotes: correlation of codon usage diversity with translation efficiency and with CG-dinucleotide usage as assessed by multivariate analysis. *J. Mol. Evol.* 53, 290–298
36 dos Reis, M. *et al.* (2004) Solving the riddle of codon usage preferences: a test for translational selection. *Nucleic Acids Res.* 32, 5036–5044
37 Duret, L. and Mouchiroud, D. (1999) Expression pattern and, surprisingly, gene length shape codon usage in *Caenorhabditis*, *Drosophila*, and *Arabidopsis*. *Proc. Natl. Acad. Sci. U.S.A.* 96, 4482–4487
38 Qian, W. *et al.* (2012) Balanced codon usage optimizes eukaryotic translational efficiency. *PLoS Genet.* 8, e1002603
39 Duret, L. (2000) tRNA gene number and codon usage in the *C. elegans* genome are co-adapted for optimal translation of highly expressed genes. *Trends Genet.* 16, 287–289
40 Dittmar, K.A. *et al.* (2006) Tissue-specific differences in human transfer RNA expression. *PLoS Genet.* 2, e221
41 Lin, K. *et al.* (2002) Conserved codon composition of ribosomal protein coding genes in *Escherichia coli*, *Mycobacterium tuberculosis* and *Saccharomyces cerevisiae*: lessons from supervised machine learning in functional genomics. *Nucleic Acids Res.* 30, 2599–2607
42 Paule, M.R. and White, R.J. (2000) Survey and summary: transcription by RNA polymerases I and III. *Nucleic Acids Res.* 28, 1283–1298
43 Geiduschek, E.P. and Kassavetis, G.A. (2001) The RNA polymerase III transcription apparatus. *J. Mol. Biol.* 310, 1–26
44 Phizicky, E.M. and Hopper, A.K. (2010) tRNA biology charges to the front. *Genes Dev.* 24, 1832–1860
45 Kutter, C. *et al.* (2011) Pol III binding in six mammals shows conservation among amino acid isotypes despite divergence among tRNA genes. *Nat. Genet.* 43, 948–955
46 Arava, Y. *et al.* (2003) Genome-wide analysis of mRNA translation profiles in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. U.S.A.* 100, 3889–3894
47 O'Brien, T. and Lis, J.T. (1991) RNA polymerase II pauses at the 5′ end of the transcriptionally induced *Drosophila hsp70* gene. *Mol. Cell. Biol.* 11, 5285–5290
48 Siller, E. *et al.* (2010) Slowing bacterial translation speed enhances eukaryotic protein folding efficiency. *J. Mol. Biol.* 396, 1310–1318
49 Zhang, G. *et al.* (2009) Transient ribosomal attenuation coordinates protein synthesis and co-translational folding. *Nat. Struct. Mol. Biol.* 16, 274–280
50 Ingolia, N.T. *et al.* (2009) Genome-wide analysis *in vivo* of translation with nucleotide resolution using ribosome profiling. *Science* 324, 218–223
51 Stapulionis, R. and Deutscher, M.P. (1995) A channeled tRNA cycle during mammalian protein synthesis. *Proc. Natl. Acad. Sci. U.S.A.* 92, 7158–7161
52 Petrushenko, Z.M. *et al.* (2002) Novel complexes of mammalian translation elongation factor eEF1A.GDP with uncharged tRNA and aminoacyl-tRNA synthetase. Implications for tRNA channeling. *Eur. J. Biochem.* 269, 4811–4818
53 Gaucher, E.A. *et al.* (2001) Function-structure analysis of proteins using covarion-based evolutionary approaches: elongation factors. *Proc. Natl. Acad. Sci. U.S.A.* 98, 548–552
54 Kubo, M. and Imanaka, T. (1989) mRNA secondary structure in an open reading frame reduces translation efficiency in *Bacillus subtilis*. *J. Bacteriol.* 171, 4080–4082
55 de Smit, M.H. and van Duin, J. (1990) Control of prokaryotic translational initiation by mRNA secondary structure. *Prog. Nucleic Acid Res. Mol. Biol.* 38, 1–35
56 Gu, W. *et al.* (2010) A universal trend of reduced mRNA stability near the translation-initiation site in prokaryotes and eukaryotes. *PLoS Comput. Biol.* 6, e1000664
57 Kim, J. *et al.* (1991) Ribosomes pause at specific sites during synthesis of membrane-bound chloroplast reaction center protein D1. *J. Biol. Chem.* 266, 14931–14938
58 Makhoul, C.H. and Trifonov, E.N. (2002) Distribution of rare triplets along mRNA and their relation to protein folding. *J. Biomol. Struct. Dyn.* 20, 413–420
59 Yanagitani, K. *et al.* (2011) Translational pausing ensures membrane targeting and cytoplasmic splicing of XBP1u mRNA. *Science* 331, 586–589
60 Zama, M. (1995) Discontinuous translation and mRNA secondary structure. *Nucleic Acids Symp. Ser.* 34, 97–98
61 Saunders, R. and Deane, C.M. (2010) Synonymous codon usage influences the local protein structure observed. *Nucleic Acids Res.* 38, 6719–6728
62 Cortazzo, P. *et al.* (2002) Silent mutations affect in vivo protein folding in *Escherichia coli*. *Biochem. Biophys. Res. Commun.* 293, 537–541
63 Komar, A.A. (2007) Genetics. SNPs, silent but not invisible. *Science* 315, 466–467
64 Kimchi-Sarfaty, C. *et al.* (2007) A 'silent' polymorphism in the *MDR1* gene changes substrate specificity. *Science* 315, 525–528
65 Fraser, H.B. *et al.* (2004) Coevolution of gene expression among interacting proteins. *Proc. Natl. Acad. Sci. U.S.A.* 101, 9033–9038
66 Lithwick, G. and Margalit, H. (2005) Relative predicted protein levels of functionally associated proteins are conserved across organisms. *Nucleic Acids Res.* 33, 1051–1057
67 Watkins, N.E. and SantaLucia, J. (2005) Nearest-neighbor thermodynamics of deoxyinosine pairs in DNA duplexes. *Nucleic Acids Res.* 33, 6258–6267
68 Agris, P.F. (2004) Decoding the genome: a modified view. *Nucleic Acids Res.* 32, 223–238
69 Alexandrov, A. *et al.* (2006) Rapid tRNA decay can result from lack of nonessential modifications. *Mol. Cell* 21, 87–96
70 Urbonavicius, J. *et al.* (2001) Improvement of reading frame maintenance is a common function for several tRNA modifications. *EMBO J.* 20, 4863–4873
71 Kramer, G.F. and Ames, B.N. (1988) Isolation and characterization of a selenium metabolism mutant of *Salmonella typhimurium*. *J. Bacteriol.* 170, 736–743
72 Chan, C.T. *et al.* (2010) A quantitative systems approach reveals dynamic control of tRNA modifications during cellular stress. *PLoS Genet.* 6, e1001247
73 Pavon-Eternod, M. *et al.* (2009) tRNA over-expression in breast cancer and functional consequences. *Nucleic Acids Res.* 37, 7268–7280
74 Gerber, A.P. and Keller, W. (1999) An adenosine deaminase that generates inosine at the wobble position of tRNAs. *Science* 286, 1146–1149
75 Nasvall, S.J. *et al.* (2004) The modified wobble nucleoside uridine-5-oxyacetic acid in tRNAPro(cmo5UGG) promotes reading of all four proline codons *in vivo*. *RNA* 10, 1662–1673
76 Yi, C. and Pan, T. (2011) Cellular dynamics of RNA modification. *Acc. Chem. Res.* 44, 1380–1388
77 Jia, G. *et al.* (2011) N6-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. *Nat. Chem. Biol.* 7, 885–887
78 He, C. (2010) Grand challenge commentary: RNA epigenetics? *Nat. Chem. Biol.* 6, 863–865
79 Globisch, D. *et al.* (2011) Systems-based analysis of modified tRNA bases. *Angew. Chem. Int. Ed. Engl.* 50, 9739–9742
80 Meyerovich, M. *et al.* (2010) Visualizing high error levels during gene expression in living bacterial cells. *Proc. Natl. Acad. Sci. U.S.A.* 107, 11543–11548
81 Stansfield, I. *et al.* (1998) Missense translation errors in *Saccharomyces cerevisiae*. *J. Mol. Biol.* 282, 13–24
82 Netzer, N. *et al.* (2009) Innate immune and chemically triggered oxidative stress modifies translational fidelity. *Nature* 462, 522–526
83 Moura, G.R. *et al.* (2009) Genetic code ambiguity: an unexpected source of proteome innovation and phenotypic diversity. *Curr. Opin. Microbiol.* 12, 631–637