BMC Evolutionary Biology

Research article

The cost of wobble translation in fungal mitochondrial genomes: integration of two traditional hypotheses

Xuhua Xia[1,2]

Address: [1]Department of Biology and Center for Advanced Research in Environmental Genomics, University of Ottawa 30 Marie Curie, Ottawa, K1N 6N5, Canada and [2]Ottawa Institute of Systems Biology, University of Ottawa, Ottawa, Canada

Email: Xuhua Xia - xxia@uottawa.ca

Abstract

Background: Fungal and animal mitochondrial genomes typically have one tRNA for each synonymous codon family. The codon-anticodon adaptation hypothesis predicts that the wobble nucleotide of a tRNA anticodon should evolve towards maximizing Watson-Crick base pairing with the most frequently used codon within each synonymous codon family, whereas the wobble versatility hypothesis argues that the nucleotide at the wobble site should be occupied by a nucleotide most versatile in wobble pairing, i.e., the tRNA wobble nucleotide should be G for NNY codon families, and U for NNR and NNN codon families (where Y stands for C or U, R for A or G and N for any nucleotide).

Results: We here integrate these two traditional hypotheses on tRNA anticodons into a unified model based on an analysis of the wobble costs associated with different wobble base pairs. This novel approach allows the relative cost of wobble pairing to be qualitatively evaluated. A comprehensive study of 36 fungal genomes suggests very different costs between two kinds of U:G wobble pairs, i.e., (1) between a G at the wobble site of a tRNA anticodon and a U at the third codon position (designated $M_{U3:G}$) and (2) between a U at the wobble site of a tRNA anticodon and a G at the third codon position (designated $M_{G3:U}$).

Conclusion: In general, $M_{U3:G}$ is much smaller than $M_{G3:U}$, suggesting no selection against U-ending codons in NNY codon families with a wobble G in the tRNA anticodon but strong selection against G-ending codons in NNR codon families with a wobble U at the tRNA anticodon. This finding resolves several puzzling observations in fungal genomics and corroborates previous studies showing that U3:G wobble is energetically more favorable than G3:U wobble.

Background

The wobble versatility hypothesis [1-6], abbreviated as WVH, states that the wobble site of tRNA anticodon should have G for NNY codons (where Y stands for C or U and N for any nucleotide) because G can pair with both C and U in RNA, and should have U for NNR to pair with both A and G. For NNN codon families, the wobble site should be U because U is known to be the most versatile in wobble-pairing [7-12]. In contrast, the codon-anticodon adaptation hypothesis, or CAAH for short, invokes the codon usage bias as a determining factor, i.e., the wobble site of tRNA anticodon should co-evolve with codon usage so that the nucleotide in the wobble site of tRNA anticodon should match the most abundant codon in a synonymous codon family [6,13-15]. The association between the major codon and the anticodon of the most

abundant tRNA has been documented in *Escherichia coli* [16,17], *Saccharomyces cerevisiae* [18], and other species and organelles [15,19-22].

Here we develop a general hypothesis of codon-anticodon adaptation based on an analysis of wobble costs, and derive its predictions that can be tested by genomic data. The wobble cost may be viewed as reduction in decoding efficiency and accuracy because such reduction would be selected against over evolutionary time. We will refer to this new general hypothesis based on wobble cost as WCH (for wobble cost hypothesis). The two traditional hypotheses, CAAH and WVH, will be shown to be special forms of WCH.

Following the shorthand notation of Ogle et al. [23], I designate the translation cost through wobble base-pairing between nucleotide i at the third codon position of a codon and the nucleotide j at the wobble site of tRNA anticodon as $M_{i3:j}$ (where M is for wobble cost. The letter C would be more suitable to represent cost but it may confuse with the nucleotide C). We assume $M_{i3:j} = 0$ if nucleotides i and j form Watson-Crick base pairing. The reason for this assumption is that C:G and A:U pairs have not been found to contribute to ribosomal stalling (which reduces translation efficiency) or amino acid mis-incorporation (which reduces translation accuracy) although almost all non-Watson-Crick base pairings have been shown to reduce translation efficiency and accuracy. We define $M_{Y3:U}$ as the wobble cost between a wobble U at the tRNA anticodon and a C or U at the third codon position. For simplicity, we also assume $M_{A3:A} = M_{C3:A} = M_{G3:A} = M_{A3:C} = M_{C3:C} = M_{U3:C} = M_{A3:G} = M_{G3:G} = M_O$ (where the subscript O is for "other", i.e., M_O is the wobble cost between base pairs that are not Watson-Crick base pairs, not U3:G or G3:U base pairs and not Y3:U pairs). In general, $M_{U3:G}$ and $M_{G3:U}$ are expected to be smaller than $M_{C3:U}$, $M_{U3:U}$, and M_O because G and U can base-pair in RNA, $M_{Y3:U}$ is expected to be smaller than M_O because a wobble U at tRNA anticodon is known to be the most versatile in wobble-pairing [7-12].

A classical study of U:G wobble pairs [24] suggests a preference for the U being at the 3' end rather than at the 5' end, which implies that U3:G wobble pair is energetically more favorable than G3:U wobble pair [25]. Subsequent studies have shown that, while the U3:G wobble pair occurs on the ribosome, the unmodified G3:U wobble pair does not [2,23,26]. These findings suggest that $M_{U3:G}$ may be smaller than $M_{G3:U}$, although its generality is unknown.

Two-fold NNY and NNR codon families
First consider the NNY codon family where Y is either C or U. Designate the number of C-ending and U-ending syn-

onymous codons by N_C and N_U, respectively, and the total cost of wobble pairing as M_{wG} when the wobble site of the anticodon is G, and as M_{wA} when the wobble nucleotide is A (we do not need to consider the case when the wobble site is U or C for NNY codon families because such cases have never been observed and because a tRNA with a wobble U or C to translate NNC and NNU codons is against physiochemical reasons). We now express the total cost M_{wG} and M_{wA} as

$$M_{wG} = N_U M_{U3:G}$$
$$M_{wA} = N_C M_{C3:A} = N_C M_O \tag{1}$$

Note that $M_{C3:G} = M_{U3:A} = 0$ according to our definition. The dependence of M_{wG} and M_{wA} on the relative frequencies of N_C, expressed as proportion of C-ending codons in the NNY codon family (P_C), is graphically shown in Fig. 1, with M_O assumed to equal $2 \bullet M_{U3:G}$. The condition for $M_{wG} = M_{wA}$, i.e., when the wobble site of the tRNA anticodon can take either G or A without a fitness differential, is

$$\frac{M_{U3:G}}{M_O} = \frac{N_C}{N_U} = \frac{P_C}{1-P_C} \tag{2}$$

In the scenario in Fig. 1 with $M_O = 2M_{U3:G}$, the condition for $M_{wG} = M_{wA}$ is $P_C = 1/3$. Thus, when P_C drifts above 1/3, it would become more beneficial (less costly) to have a G at the wobble nucleotide of tRNA anticodon. However, when PC drifts below 1/3, then a wobble A would be more advantageous and should be favored by natural selection. Naturally, if $M_{U3:G} = 0$, then we should always have a wobble G at the tRNA anticodon regardless of codon usage.

Now we consider three special cases with reference to Fig. 1 and the two traditional hypotheses, CAAH and WVH. First, if $N_C = N_U = N/2$, then CAAH, which predicts that the wobble nucleotide of tRNA anticodon should form Watson-Crick base pair with the most abundant codon, has no prediction at all because the two codons are equally abundant. In contrast, if we assume $M_{U3:G} < M_O$ as is depicted in Fig. 1, we have $M_{wG} < M_{wA}$, and WCH predicts that the anticodon wobble site should be a G. This is the same prediction as WVH.

Second, if $N_C > N_U$, then $M_{wG} < M_{wA}$ assuming $M_{U3:G} \leq M_O$, and WCH predicts that the anticodon wobble site should be occupied by a G. This prediction is shared by both CAAH and WVH.

Third, when $N_C < N_U$ but $N_C > N_U \cdot M_{U3:G}/M_O$, then WCH will still predict a G at the wobble nucleotide of tRNA anticodon because M_{wG} is still smaller than M_{wA}. Take the scenario in Fig. 1 for example, when $N_C < N_U$ but $P_C > 1/3$
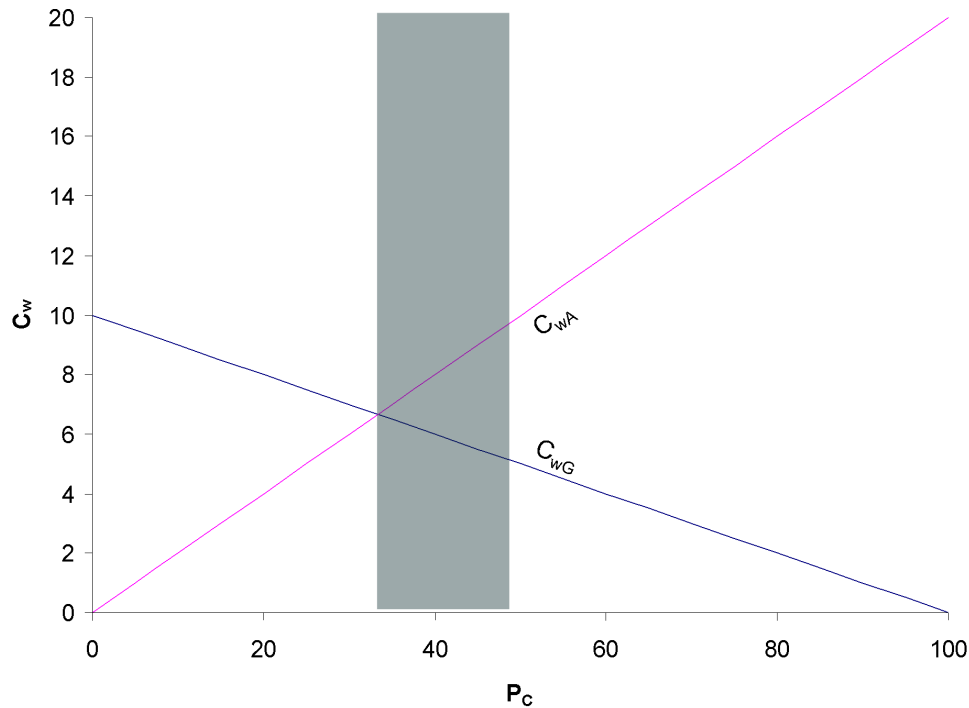
Figure 1
Conceptual illustration of the dependence of wobbling cost involving a G or an A at the wobble site of tRNA anticodon (M_{wG} and M_{wA}) on the proportion of C-ending codons (P_C) in an NNY codon family, with $M_{C3:A} = 2M_{U3:G}$. The shaded area corresponds to P_C smaller than 1/2 but larger than 1/3.

(which correspond to the shaded area in Fig. 1), we have $M_{wG} < M_{wA}$ (Fig. 1), so natural selection should favor a G at the wobble site of tRNA anticodon. WVH happens to have the same prediction. However, CAAH will predict an A at the wobble site of tRNA anticodon if U-ending codons are more abundant than C-ending codons. This is in contrast to WVH, i.e. CAAH is inapplicable when P_C is within the shaded range in Fig. 1.

Fourth, when $N_C \ll N_U$, especially in the extreme case when $N_C = 0$, then Eq. (1) is reduced to $M_{wG} = N_U \cdot M_{U3:G}$ and $M_{wA} = 0$. Because now $M_{wG} > M_{wA}$, WCH predicts an A at the anticodon wobble site. In this case, WVH would still predict a G at the wobble site because it ignores the codon frequencies (i.e., it ignores the relative magnitude of N_C and N_U), but CAAH would predict an A at the wobble site, which is the same prediction as WCH. Only in this particular case when $N_C \ll N_U$ can CAAH and WVH be clearly differentiated. As is depicted in Fig. 1, the chance for CAAH to be supported decreases as $M_{U3:G}$ decreases relative to M_O. CAAH would have little chance to be unambiguously supported if $M_{U3:G}$ is close to 0. Thus, as long as $M_{U3:G}$ is small relative to M_O, we should expect CAAH to be less supported than WVH. A recent study shows that CAAH indeed receives much less support

than WVH when both were tested against fungal mitochondrial genomic data [27].

The simple Eq. (1) offers a way to assess relative costs of different wobble pairs. For example, if nature has chosen a G at the wobble site of the tRNA anticodon, then we may infer that $M_{wG} < M_{wA}$. So, from Eq. (1), we have

$$N_U M_{U3:G} < N_C M_O$$
$$\frac{M_{U3:G}}{M_O} < \frac{N_C}{N_U} \tag{3}$$

Suppose we have an NNY codon family translated by a tRNA with a G at the wobble site of the anticodon. If $M_{U3:G}$ is large, then there would be strong selection favoring C-ending codons against U ending codons. So the ratio of N_C/N_U may be very large. In such a case the estimated $M_{U3:G}/M_O$ ratio, being smaller than a very large value, is not informative. However, in different codon families, mutation pressure may allow N_U to drift up relative to N_C. Suppose we have three NNY codon families each translated by a tRNA with a G at the anticodon wobble site. If the N_C/N_U ratio for the three NNY codon families are 100/2, 80/100 and 200/20, respectively, we may infer that $M_{U3:G}/M_O < 80/100$ (because it is the most pre-

cise). Alternatively, with many N_C/N_U ratios from many NNY codon families translated by tRNA with a G at its anticodon wobble site, we may compute the lower 95% confidence limit of the N_C/N_U ratio (LCL$_{95.G}$ where the subscript G indicates tRNA with a G at the anticodon wobble site) and infer that $M_{U3:G}/M_{C3:A} < $ LCL$_{95.G}$.

If we always have very large N_C/N_U ratios, we may infer that selection against U-ending codons must be strong, with little chance for mutation to elevate N_U. This is a strong indication of a large $M_{U3:G}$. Along the same line of reasoning, we may infer that $M_{U3:G}$ is very small if N_U can often as large as, or even larger than, N_C.

Similarly, if nature has chosen an A at the wobble site of tRNA anticodon, then we may infer $M_{wG} > M_{wA}$, so

$$
\begin{aligned}
N_U M_{U3:G} &> N_C M_O \\
\frac{M_{U3:G}}{M_O} &> \frac{N_C}{N_U}
\end{aligned}
\tag{4}
$$

We can apply exactly the same rationale for the NNR codon family leading to parallel conclusions. For example, if nature has chosen a U at the wobble site of the tRNA anticodon, then we may infer that $M_{wU} < M_{wC}$, so that

$$
\frac{M_{G3:U}}{M_O} < \frac{N_A}{N_G}
\tag{5}
$$

Similarly, if $M_{G3:U}$ is very large, then there should be strong selection against G-ending codons in favor of A-ending codons. This will produce large N_A/N_G ratios. In contrast, a large N_G comparable to N_A indicates a very small $M_{G3:U}$ cost.

Given previous studies indicating that U3:G is energetically much more favorable than G3:U [2,23-26], we should expect $M_{U3:G} < M_{G3:U}$. The reasoning above paves the way for us to test whether this is generally true among the genetically diverse fungal species.

Three-fold AUH (Isoleucine) codons
Designate codons ending with A, C, and U as N_A, N_C and N_U, respectively. The wobble cost of having an A, G, C, or U at the wobble site of tRNA anticodon is

$$
\begin{aligned}
M_{wA} &= N_A M_{A3:A} + N_C M_{C3:A} = (N_A + N_C) M_O \\
M_{wG} &= N_A M_{A3:G} + N_U M_{U3:G} = N_A M_O + N_U M_{U3:G} \\
M_{wC} &= N_A M_{A3:C} + N_C M_{C3:C} + N_U M_{U3:C} = (N_A + N_C) M_O + N_U M_{U3:C} \\
M_{wU} &= N_C M_{C3:U} + N_U M_{U3:U} = (N_C + N_U) M_{Y3:U}
\end{aligned}
\tag{6}
$$

It is obvious that M_{wC} is always greater than M_{wA}, so we should never find a C at the wobble site of a tRNA^{Ile} anticodon, i.e., we can disregard M_{wC}. If nature has chosen G

at the wobble site of tRNA^{Ile} anticodon, then we may infer that M_{wG} is the smallest. From $M_{wG} < M_{wA}$ and $M_{wG} < M_{wU}$, we have

$$
\begin{aligned}
\frac{M_{U3:G}}{M_O} &< \frac{N_C}{N_U} \\
N_U M_{U3:G} &< (N_C + N_U) M_{X3:U} - N_A M_O \\
\frac{M_{U3:G}}{M_{Y3:U}} &< \frac{N_C + N_U - N_A}{N_U}, \text{assuming } M_{Y3:U} < M_O
\end{aligned}
\tag{7}
$$

where the assumption is made on the basis of previous observations that U is generally the most versatile in wobble-pairing among the four nucleotides [7-12].

Four-fold NNN codons
Designate the number of codons ending with A, C, G, and U as N_A, N_C, N_G, and N_U, respectively. The wobble costs involving an A, C, G or U at the wobble site of tRNA anticodon are, respectively,

$$
\begin{aligned}
M_{wA} &= N_A M_{A3:A} + N_C M_{C3:A} + N_G M_{G3:A} = (N_A + N_C + N_G) M_O \\
M_{wC} &= N_A M_{A3:C} + N_C M_{C3:C} + N_U M_{U3:C} = (N_A + N_C) M_O + N_U M_{U3:C} \\
M_{wG} &= N_A M_{A3:G} + N_G M_{G3:G} + N_U M_{U3:G} = (N_A + N_G) M_O + N_U M_{U3:G} \\
M_{wU} &= N_C M_{C3:U} + N_G M_{G3:U} + N_U M_{U3:U} = (N_C + N_U) M_{Y3:U} + N_G M_{G3:U}
\end{aligned}
\tag{8}
$$

If nature has chosen a U at the wobble site of the tRNA anticodon, then Eq. (8) does not give us any simple inequality to estimate the cost ratios. However, if nature has chosen an A at the wobble site of the tRNA anticodon, then from $M_{wA} < M_{wG}$ and $M_{wA} < M_{wC}$, we can infer

$$
\begin{aligned}
\frac{M_{U3:G}}{M_O} &> \frac{N_C}{N_U} \\
\frac{M_{U3:C}}{M_O} &> \frac{N_G}{N_U}
\end{aligned}
\tag{9}
$$

In what follows, we estimate $M_{U3:G}/M_O$, $M_{G3:U}/M_O$, and $M_{U3:C}/M_O$ by using fungal mitochondrial genomic data. Cells in fungal species are generally rapid-replicating which necessitates efficient translation. Rapidly replicating unicellular organisms are theoretically expected to be under strong selection to increase the rate of biosynthesis [15,28] and they typically exhibit strong codon-anticodon adaptation [29]. Thus, fungal species should be ideal for evaluating evolutionary hypothesis on codon-anticodon adaptation.

Methods
We retrieved 36 fungal mitochondrial genomes (Table 1) by using NCBI Entrez. Three different genetic codes are used in different fungal genomes. Among the 36 fungal mitochondrial genomes, seven genomes use translation table 3, 27 genomes use translation table 4, and two

genomes use translation table 16 (Table 1). When results are similar among genomes using the same translation table, only results from a representative genome are presented. The number of codon families supporting CAAH (N_{CAAH}) and WVH (N_{WVH}) is compiled following the following rationale [27]. Suppose a lysine (Lys) codon family has 20 AAA and 60 AAG codons. WVH would ignore the codon usage bias and predict a wobble U in the tRNA-Lys anticodon because U can pair with both A and G, whereas CAAH would predict a wobble C in the tRNA^{Lys} anticodon to maximize the Watson-Crick match with the more frequent G-ending codons. If the tRNA^{Lys} anticodon is found to have a wobble U, then WVH is supported; if a wobble C is found, then CAAH is supported. If we have 60 AAA codons and 20 AAG codons and if tRNA^{Lys} anticodon

has a wobble U, then both hypotheses are supported, i.e., they are indistinguishable and are not included in Table 1. The methionine codon families are not included in Table 1 but discussed in detail elsewhere [27,30].

The tRNA and CDS sequences were extracted and analyzed by using DAMBE [31,32]. The CDS-derived codon usage is also computed with DAMBE. The anticodon in almost all tRNA sequences from all species share the regular feature of being flanked by two nucleotides on either side to form a loop that is held together by a stem. For example, the anticodon loop (AC loop) of the tRNA^{Arg} genes translating CGN codons in *Epidermophyton floccosum* is 28CGUG**UU**ACG**G**CCACG42, where the starting and ending numbers indicate the position of the AC loop in

Table 1: Number of codon families unambiguously supporting the codon-anticodon adaptation hypothesis (N_{CAAH}) and the wobble versatility hypothesis (N_{WVH}) in each fungal species.

Species	Accession*	Code†	N_{CAAH}	N_{WVH}
Allomyces macrogynus	NC_001715	4	1	13
Ashbya gossypii ATCC 10895	NC_005789	3	0	10
Aspergillus niger	NC_007445	4	0	12
Aspergillus tubingensis	NC_007597	4	0	12
Candida albicans SC5314	NC_002653	4	1	11
Candida glabrata	NC_004691	3	0	12
Candida metapsilosis	NC_006971	4	1	15
Candida orthopsilosis	NC_006972	4	1	15
Candida parapsilosis	NC_005253	4	1	15
Candida stellata	NC_005972	4	1	13
Moniliophthora perniciosa	NC_005927	4	0	13
Epidermophyton floccosum	NC_007394	4	2	12
Harpochytrium sp. JEL94	NC_004760	4	1	3
Harpochytrium sp. JEL105	NC_004623	4	1	4
Hyaloraphidium curvatum	NC_003048	4	1	4
Hypocrea jecorina	NC_003388	4	1	13
Kluyveromyces lactis	NC_006077	4	0	13
Kluyveromyces thermotolerans	NC_006626	3	0	11
Lecanicillium muscarium	NC_004514	4	0	10
Monoblepharella sp. JEL 15	NC_004624	4	0	15
Mortierella verticillate	NC_006838	4	1	13
Penicillium marneffei	NC_005256	4	1	12
Pichia Canadensis	NC_001762	4	2	12
Podospora anserine	NC_001329	4	0	13
Rhizophydium sp.136	NC_003053	16	0	2
Rhizopus oryzae	NC_006836	4	1	14
Saccharomyces cerevisiae	NC_001224	3	1	12
Saccharomyces castellii	NC_003920	3	0	14
Saccharomyces servazzii	NC_004918	3	0	9
Schizophyllum commune	NC_003049	4	1	13
Schizosaccharomyces japonicus	NC_004332	4	1	16
Schizosaccharomyces octosporus	NC_004312	4	1	14
Schizosaccharomyces pombe	NC_001326	4	1	14
Smittium culisetae	NC_006837	4	1	15
Spizellomyces punctatus1	NC_003052	16	0	4
Yarrowia lipolytica	NC_002659	3	0	11
Sum			23	414

* GenBank accession number.
† Genetic code.

the tRNA sequence, with the anticodon 5'-ACG-3' (matching codon CGU) flanked by two nucleotides on either side (in bold) to form a loop that is held together by a stem made of the first and the last four nucleotides. Similarly, the other tRNA^{Arg} translating AGR codons is 25AAAAUA**CU**UCU**AA**UAUUUU43, with the AC loop held together by a six-base stem. However, some tRNA sequences have a suspicious AC loop and DAMBE will flag them out. The AC loop is then identified by aligning the tRNA sequences against other isoaccepting tRNA sequences with a regular AC loop [6]. Some tRNA anticodon loop has the anticodon flanked by three instead of two nucleotides. For example, the anticodon loop in tRNA^{Leu} in the mitochondrial genome of *Kluyveromyces thermotolerans* is GAUAC**UCU**UAA**GAU**GUAUU, with the anticodon UAA flanked by three nucleotides (in bold) on both sides. There are a few tRNA sequences in which anticodon loop cannot be identified.

Some mitochondrial genomes in GenBank are annotated incorrectly. For example, tRNA^{Pro} in the mitochondrial genome of *Ashbya gossypii* ATCC 10895 has an anticodon of UGG matching codon CCA (the most frequently used proline codon), but the GenBank file (NC_005789) annotated the anticodon to match codon CCU.

A few fungal mitochondrial genomes do not have a complete set of tRNA genes. For example, the mitochondrial genomes of *Hyaloraphidium curvatum* and *Harpochytrium sp*. JEL94 have seven and eight tRNA genes, respectively, and consequently will need tRNA import from the nuclear genome. This may cause complication in analyzing codon-anticodon adaptation. However, removing such genomes does not alter the conclusions.

Some species exhibit extreme avoidance of certain codon families. For example, *Ashbya gossypii* ATCC 10895 codes Arg with only AGR codons without using any CGN codons. In contrast, *Hyaloraphidium curvatum* codes Arg with only CGN codons without using any AGR codons. Such avoidance of certain codon families would facilitate the evolutionary loss of the associated tRNA [33-35], although it is not always clear whether the avoidance is the cause or the consequence of the loss of the associated tRNA.

We computed relative synonymous codon usage, or RSCU [36], as a measure of codon usage bias within a codon family by using DAMBE [31,32]. Some coding sequences are incomplete. For example, the cox1 CDS in *Aspergillus niger* is annotated as "join(<19768..20614,21640..22495)". The first two nucleotides (i.e., at positions 19768 and 19769) represent a partial codon and are discarded in computing codon frequencies.

Results and discussion
Wobble cost between G and U: MU3:G and MG3:U
Recall that the two inequalities

$$\frac{M_{U3:G}}{M_O} < \frac{N_C}{N_U} \text{ and}$$
$$\frac{M_{G3:U}}{M_O} < \frac{N_A}{N_G} \quad (10)$$

are, respectively, for NNY codons translated by tRNA with a G at the wobble site of tRNA anticodon, and for NNR codons translated by tRNA with a U at the wobble site of tRNA anticodon. The observed N_C/N_U ratios for *Allomyces macrogynus* (representing fungal mitochondrial genomes with translation table 4) are much smaller than N_A/N_G ratios (Table 2). The smallest N_C/N_U value is 0.279 whereas the smallest N_A/N_G value is 2.372 (Table 2). We have mentioned before that, if $M_{U3:G}$ is very small, then a wobble G at the tRNA anticodon will not impose strong selection against U-ending codons, and N_U may drift up and down with mutation relative to N_C. This will lead to relatively small N_C/N_U ratios. From the minimum N_C/N_U value of 0.279, we may infer that $M_{U3:G} < 0.279 \bullet M_O$, i.e., $M_{U3:G}$ is quite small relative to M_O.

The AUH codon family coding for amino acid Ile in *A. macrogynus* mitochondrial genome is translated by a tRNA with a GAU anticodon. According to Eq. (7), the $M_{U3:G}$/Co ratio should also be smaller than the N_C/N_U ratio. The observed N_C/N_U ratio is 0.3605 (= 159/441). This is similar to the N_C/N_U ratio in NNY codon families (Table 2). Thus, the wobble cost of $M_{U3:G}$ relative to M_O from the AUH codon family is similar to that derived from NNY codon families.

The CUN codon family coding for amino acid Leu in *A. macrogynus* mitochondrial genome is translated by a tRNA with an AAG anticodon. Note that no A→I conversion has been observed in mitochondria [37,38] so Eq. (9) is applicable. According to Eq. (9), the $M_{U3:G}$/Co ratio should be greater than the N_C/N_U ratio. The observed N_C/N_U ratio is 0.1905 (= 48/252). Thus we have $M_{U3:G} > 0.1905 \bullet M_O$. This inequality, together with the previous inequality of $M_{U3:G} < 0.279 \bullet M_O$, leads to $0.1905 \bullet M_O < M_{U3:G} < 0.279 \bullet M_O$.

Given that $M_{U3:G} > 0.1905 \bullet M_O$, we can infer that $M_{wG} > M_{wA}$ when $N_C/N_U < 0.1905$, i.e., when U-ending codons are more than five times as frequent as C-ending codons. In other words, when $N_C/N_U < 0.1905$, the wobble cost of having an A at the wobble site of the tRNA anticodon is smaller than that of having a G at the wobble site, and should be favored by natural selection. Among the 36 fungal mitochondrial genomes, only two genomes have a NNY codon family with a correspondent tRNA that has an

Table 2: N_C/N_U ratios for NNY codons and N_A/N_G ratios for NNR codons in *Allomyces macrogynus* (representing fungal mitochondrial genomes with translation table 4).

CF*	AA†	AC‡	N_{Cod}§	N_C/N_U	CF*	AA†	AC‡	N_{Cod}§	N_A/N_G
AAY	N	GUU	360	0.295	AAR	K	UUU	411	2.543
AGY	S	GCU	230	0.314	AGR	R	UCU	202	6.769
CAY	H	GUG	213	0.357	CAR	Q	UUG	173	3.806
GAY	D	GUC	356	0.299	GAR	E	UUC	263	2.372
UAY	Y	GUA	390	0.279	UUR	L	UAA	742	5.870
UGY	C	GCA	94	0.382					
UUY	F	GAA	592	0.726					
Minimum				0.279					2.372
Mean				0.379					4.272
Std Dev				0.157					1.975

* CF: Codon family
† AA: One-letter code of amino acid
‡ AC: Anticodon
§ N_{Cod}: Number of codons in the codon family

A at the wobble site of its anticodon. In the mitochondrial genome of *Penicillium marneffei*, the tRNA translating the AAY codon family has a wobble A in its anticodon. The N_C/N_U ratio is 0.087 (= 41/473) which is much smaller than 0.1905. Similarly, in the mitochondrial genome of *Pichia Canadensis*, the tRNA translating the AGY codon family has a wobble A in its anticodon. The N_C/N_U ratio is 0.0083 (= 1/120) which is also much smaller than 0.1905. Thus, the prediction that a wobble A at the tRNA anticodon is advantageous over a wobble G when N_C/N_U < 0.1905 is consistent with empirical data.

In contrast to the small N_C/N_U ratios in NNY, AUH and CUN codon families, all N_A/N_G ratios in NNR codon families are substantially larger (Table 2). We have argued before that, if $M_{G3:U}$ is very small, then a U at the wobble site of tRNA anticodon would impose little selection against G-ending codons in NNR codon families, and mutation may allow N_G to drift up, leading to large N_G values relative to N_A. However, if $M_{G3:U}$ is large, then G-ending codons should be strongly selected against and N_G would be small relative to N_A, leading to large N_A/N_G ratios. The much larger N_A/N_G ratios than N_C/N_U ratios (t = 5.2967, DF = 10, p = 0.0003, two-tailed test) suggest that $M_{G3:U}$ is much greater than $M_{U3:G}$.

There is a caveat in evaluating the relative magnitude of $M_{U3:G}/M_O$, and $M_{G3:U}/M_O$ by the N_C/N_U and N_A/N_G ratios because these ratios can be affected by AT-biased mutations. The mitochondrial genome of *A. macrogynus* is 57472 bp, with the number of C+G being 22700 and that of A+U being 34772. If we exclude those nucleotides in coding sequences, then the numbers of C+G and A+U are 14136 and 17656, respectively. This may be considered as the background frequencies maintained by mutation bias, which leads to the expected N_C/N_U ratio of 0.8001

(=N_{C+G}/N_{A+U}) and that of N_A/N_G ratio of 1.2490 (=N_{A+U}/N_{C+G}). Thus, we expect N_C/N_U to be smaller than N_A/N_G even when there is no difference between $M_{U3:G}$ and $M_{G3:U}$. To establish the argument that $M_{U3:G}$ is indeed smaller than $M_{G3:U}$, we need to demonstrate that (1) there is no selection against U-ending codons in NNY codon families by showing that the observed N_C/N_U ratio is not greater than 0.8001, and (2) there is selection against G-ending codons in NNR codon families by showing that N_A/N_G is significantly greater than 1.2490. It is not enough to show that N_C/N_U < N_A/N_G.

We note that the observed N_C/N_U values for the seven NNY codon families in Table 2 are all smaller than the expected 0.8001, suggesting no selection against U-ending codons in NNY codon families (i.e., small $M_{U3:G}$). In contrast the observed N_A/N_G ratios for the five NNR codon families are all much greater than the expected 1.2490 (Table 2), consistent with the interpretation of selection against G-ending codons in NNR codon families (i.e., large $M_{G3:U}$). This is consistent with the interpretation that $M_{U3:G}$ < $M_{G3:U}$. One can perform a χ^2-test for each of the NNR codon families to see if G-ending codons are under-used. The tests are all highly significant, with p < 0.00001.

The results are similar for fungal genomes with translation table 3, with the result from a representative species (*Saccharomyces cerevisiae*) presented in Table 3. Again the N_C/N_U ratios in NNY codon families are much smaller than N_A/N_G ratios in NNR codon families. We should note that the *S. cerevisiae* mitochondrial genome is much more AT-biased than the *A. macrogynus* mitochondrial genome, with the proportion of (G+C) in non-coding sequences being only 0.1484. The reason for the GC deficiency in yeast is not clear, but it may be caused either by mutation bias or by the low abundance of C in living cells [39-41].

In any case, the expected N_C/N_U and N_A/N_G ratios, given the biased genomic AT content, are 0.1742 and 5.7405, respectively. We note that the observed N_C/N_U ratios among the NNY codon families are all smaller than the expected value of 0.1742 except for the UUY (Phe) codon family (Table 3), against suggesting little selection against U-ending codons (i.e., small $M_{U3:G}$). In contrast, the N_A/N_G ratios are all significantly greater than the expected 5.7405 except for the AUR (Met) codon family, suggesting selection against G-ending codons (i.e., large $M_{G3:U}$). The exceptional AUR (Met) codon family has a tRNA with a CAU anticodon which would favor G-ending codons and is expected to be different.

The N_C/N_U ratio for the UUY (Phe) codon family is consistently greater than those of other NNY codon families (Tables 2, 3). One may suspect whether, for this particular codon family, there is a significant $M_{U3:G}$. The rate of tRNA[Phe] with anticodon 3'-AAG-5' dissociating from the UUU codon is about twice as high as that from the fully matched UUC codon [42]. Also, the tRNA[Phe] misread CUC codons more than twice as often as CUU codons [42]. These two lines of evidence suggest that C3:G pair is much more favorable than U3:G pair, i.e., $M_{U3:G}$ for the UUY (Phe) codon family may indeed be substantially greater than $M_{C3:G}$. Unfortunately, there has been no other similar studies on codon-anticodon pairing for other NNY codon families. One should also note that the tRNA[Phe] in this study comes from *Escherichia coli*, and the result may not be applicable to fungal species.

The two fungal species using genetic table 16, i.e., *Spizellomyces punctatus* and *Rhizophydium sp. 136* each have only a partial set of tRNA genes. Among NNY codon families in *S. punctatus*, only the GAY (coding for Asp) and UAY

codon family (coding for Tyr) have an identifiable tRNA, with anticodons being GUC and GUA, respectively. The expected N_C/N_U and N_A/N_G ratios are 0.4536 and 2.2048, respectively, based on the nucleotide frequencies of non-coding sequences in the mitochondrial genome. The observed N_C/N_U ratios for GAY and UAY codon families are 0.1667 (= 40/240) and 0.2411 (= 81/336), respectively, suggesting no selection against U-ending codons (i.e., small $M_{U3:G}$). In contrast, the N_A/N_G ratios are much larger, being 19.2308 (= 250/13) for the AAR codon family (coding for amino acid Lys) translated by a tRNA with a UUU anticodon and 11.1538 (= 145/13) for the CAR codon family (coding for amino acid Gln) translated by a tRNA with a UUG anticodon. These results suggest selection against G-ending codons (i.e., large $M_{G3:U}$) There is no other NNR codon family with identifiable tRNAs in *S. punctatus*. The other species using translation table 16, *Rhizophydium sp. 136*, exhibits a similar pattern.

In short, results from fungal mitochondrial genomes are consistent with no selection against U-ending codons in NNY codon families but significant selection against G-ending codons in NNR codon families, indicating that $M_{U3:G}$ is smaller than $M_{G3:U}$. These findings corroborate previous biochemical studies demonstrating that U3:G is energetically much more favorable than G3:U [2,23-26]. However, G3:U can be almost as good as A3:U when U is modified to xm[5]U [43].

The finding of a small $M_{U3:G}$ can explain puzzling observations in codon usage in fungal mitochondrial genomes. Take the tRNA[Ser] translating the AGY codon family in the mitochondrial genome of *Ashbya gossypii* ATCC 10895 for example. The genome contains 31 AGU codons and no AGC codon. CAAH would have predicted an ACU antico-

Table 3: N_C/N_U ratios for NNY codons and N_A/N_G ratios for NNR codons in *Saccharomyces cerevisiae* (representing fungal mitochondrial genomes with translation table 3).

CF*	AA†	AC‡	N_{Cod} §	N_C/N_U	CF*	AA†	AC‡	N_{Cod} §	N_A/N_G
AAY	N	GUU	786	0.065	AAR	K	UUU	599	16.114
AGY	S	GCU	119	0.035	AGR	R	UCU	225	36.500
AUY	I	GAU	728	0.093	AUR	M	CAU	428	1.326
CAY	H	GUG	183	0.070	CAR	Q	UUG	165	8.706
GAY	D	GUC	264	0.052	GAR	E	UUC	220	9.000
UAY	Y	GUA	495	0.076	UGR	W	UCA	130	20.667
UGY	C	GCA	92	0.070	UUR	L	UAA	911	64.071
UUY	F	GAA	445	0.369					
Minimum				0.035					1.326
Mean				0.104					22.341
Std Dev				0.109					21.560

* CF: Codon family
† AA: One-letter code of amino acid
‡ AC: Anticodon
§ N_{Cod}: Number of codons in the codon family

don with perfect base pair with AGU codons, but the observed anticodon is GCU consistent with WVH. Such an anticodon makes sense only if $M_{U3:G}$ is equal to $M_{C3:G}$, i.e., a wobble G in a tRNA anticodon can pair with U just as good as with C in the third codon position.

Can this finding be generalized to nuclear genomes and the translation machinery in the cytoplasm? If $M_{U3:G}$ is generally small, then we have an answer to a puzzling observation that has long baffled molecular evolutionary geneticists. The UGY codon family (coding for Cys) in the nuclear genome of *S. cerevisiae* is translated by tRNA coded in four tRNA^{Cys} genes all with a wobble G at the tRNA anticodon. One would have predicted that UGC codons should be strongly preferred over UGU codons to avoid the wobble cost $M_{U3:G}$. However, the observed numbers of UGC and UGU codons in *S. cerevisiae* CDSs is 13802 and 22873, respectively, opposite to the predicted trend. The bias is even stronger in highly expressed genes. For example, in the codon usage table of highly expressed genes distributed with the EMBOSS package [[44], Eyeast-cai.cut], the numbers of UGC and UGU codons are 3 and 39, respectively. The unexpected codon usage bias (unexpected because the bias is in the wrong direction) becomes easy to understand if $M_{U3:G}$ is equal to $M_{C3:G}$, i.e., there is no need to overuse the UGC codons to avoid $M_{U3:G}$ so AT-biased mutation in yeast (whose genome is highly AT-biased), which increases the frequency of U-ending codons, is not checked by counteracting selection.

If the finding of $M_{U3:G} << M_{G3:U}$ is applicable to nuclear genomes, then we predict that NNY codon families need only one type of tRNA (i.e., tRNA with a wobble G at the tRNA anticodon) for translation. In contrast, NNR codon families should ideally be translated by two different types of tRNAs, one with a wobble U for NNA codons and the other with a wobble C for NNG codons (to avoid the relatively high wobble cost of $M_{C3:U}$). A corollary is that, if a NNY codon family is translated by only tRNA with a wobble G and NNR codon family by tRNA with a wobble U, then codon usage bias should be smaller in the NNY codon family (in which selection against G-ending codon is weak because of the small $M_{U3:G}$) than in the NNR codon family (in which selection against U-ending codon is strong because of the relatively large $M_{G3:U}$). Below we test these predictions with genomic data.

NNY codon families are translated by one type of tRNA with a wobble G and NNR codon families are translated by two types of tRNA with a wobble U and a wobble C, respectively, in fungal nuclear genomes

We have inferred that a tRNA with a wobble G at its anticodon should be efficient not only in translating C-ending codons, but also in translating U-ending codons because of the small $M_{U3:G}$. For this reason, only one type of tRNA

with a wobble G should generally be sufficient in translating NNY codon families. In contrast, in NNR codon families, tRNA with a wobble U at its anticodon should be poor in translating G-ending codons because of the large $M_{G3:U}$. Thus, the presence of G-ending codons in NNR codon families should favour the use of two types of tRNAs, one with a wobble U for translating NNA codons and another with a wobble C for translating NNG codons. In mitochondrial genomes with limited gene content, each codon family is generally translated by a single tRNA species. So this prediction cannot be tested. However, this prediction can be tested with nuclear genomes where the number of tRNA genes is not so limited as in mitochondrial genomes.

The prediction is strongly supported by results from the nuclear genome of *Saccharomyces cerevisiae* (Table 4). All NNR codon families are translated by two types of tRNAs with anticodons matching NNA and NNG codons, respectively, whereas all NNY codon families are translated by one type of tRNA with a wobble G at the tRNA anticodon (Table 4). This is consistent with the interpretation that the inference of $M_{U3:G} << M_{G3:U}$ derived from fungal mitochondrial genome is also applicable to fungal nuclear genomes.

One may propose an alternative hypothesis for the observation that NNY codon families are translated by tRNAs with a wobble G whereas NNR codon families translated by tRNAs with a wobble C and a wobble U, respectively. The wobble A in some nuclear tRNAs is known to be converted to inosine or I for short [45-47] which may pair with A, C or U. If a tRNA translating an NNY codon family has its wobble G mutated to wobble A, then the wobble A may undergo the A→I conversion and misread NNA codons. For this reason, the wobble G→A mutation should be strongly selected against, which would explain the lack of wobble A in tRNA translating NNY codons.

This alternative hypothesis invoking the A→I conversion, while logically sound, was dismissed quite early after the discovery [45,47] that the A→I conversion is quite restrictive and occurs mainly at ACG or ARV anticodons (where R is the IUB code for A or G, and V is for A, G or C). Among tRNAs translating NNY codons, only tRNA^{Phe} has a middle A in the anticodon, and the rest do not have an R in the middle of the anticodon. This means that, for all NNY-translating tRNAs except for tRNA^{Phe}, even if their regular wobble G mutates to A, the resulting wobble A will NOT be converted to inosine. Thus, the alternative hypothesis can only explain the avoidance of a wobble A in tRNA^{Phe} but cannot explain the avoidance of a wobble A in other NNY-translating tRNAs.

Table 4: tRNAs translating two-fold codon families in the nuclear genome of *Saccharomyces cerevisiae*.

AA*	Codon†	T‡	F§
Arg	AGA	11	314
Arg	AGG	1	1
Gln	CAA	9	153
Gln	CAG	1	1
Glu	GAA	14	305
Glu	GAG	2	5
Leu	UUA	7	42
Leu	UUG	10	359
Lys	AAA	7	65
Lys	AAG	14	483
Asn	AAC	10	208
Asn	AAU	0	11
Asp	GAC	16	202
Asp	GAU	0	112
Cys	UGC	4	3
Cys	UGU	0	39
His	CAC	7	102
His	CAU	0	25
Phe	UUC	10	168
Phe	UUU	0	19
Ser	AGC	2	6
Ser	AGU	0	4
Tyr	UAC	8	141
Tyr	UAU	0	10

* Amion acid carried by tRNA
† Codons forming Watson-Crick base pair with the anticodon of tRNA
‡ Copy number of tRNA gene with anticodon matching the corresponding codon.
§ Codon frequencies of highly expressed yeast protein-coding genes compiled in the Eyeastcai.cut file distributed with EMBOSS [44].

Conclusion

In summary, our general hypothesis based on wobble costs allows the integration of the two conventional hypotheses (i.e., CAAH and WVH) on codon-anticodon wobble pairing. The integration leads to new ways of evaluating relative wobble cost of different wobble pairings. In particular, the finding that $M_{U3:G}$ is much smaller than $M_{G3:U}$ corroborates previous structural studies showing that U3:G is energetically more favourable than G3:U and leads to a better understanding of the translation efficiency mediated by codon and anticodon wobble pairing.

Acknowledgements

References

1. Bonitz SG, Berlani R, Coruzzi G, Li M, Macino G, Nobrega FG, Nobrega MP, Thalenfeld BE, Tzagoloff A: **Codon recognition rules in yeast mitochondria.** *Proc Natl Acad Sci U S A* 1980, **77(6):**3167-3170.
2. Agris PF: **Decoding the genome: a modified view.** *Nucleic Acids Res* 2004, **32(1):**223-238.
3. Tong KL, Wong JT: **Anticodon and wobble evolution.** *Gene* 2004, **333:**169-177.
4. Heckman JE, Sarnoff J, Alzner-DeWeerd B, Yin S, RajBhandary UL: **Novel features in the genetic code and codon reading patterns in *Neurospora crassa* mitochondria based on sequences of six mitochondrial tRNAs.** *Proc Natl Acad Sci U S A* 1980, **77(6):**3159-3163.
5. Martin RP, Sibler AP, Gehrke CW, Kuo K, Edmonds CG, McCloskey JA, Dirheimer G: **5-[[(carboxymethyl)amino]methyl]uridine is found in the anticodon of yeast mitochondrial tRNAs recognizing two-codon families ending in a purine.** *Biochemistry (Mosc)* 1990, **29(4):**956-959.
6. Xia X: **Mutation and selection on the anticodon of tRNA genes in vertebrate mitochondrial genomes.** *Gene* 2005, **345(1):**13-20.
7. Yokoyama S, Nishimura S: **Modified nucleotides and codon recognition.** In *tRNA: structure, biosynthesis and function* Edited by: Soll D, RajBhandary U. Washington , ASM Press; 1995:207-223.
8. Sibler AP, Dirheimer G, Martin RP: **Codon reading patterns in *Saccharomyces cerevisiae* mitochondria based on sequences of mitochondrial tRNAs.** *FEBS Lett* 1986, **194(1):**131-138.
9. Inagaki Y, Kojima A, Bessho Y, Hori H, Ohama T, Osawa S: **Translation of synonymous codons in family boxes by *Mycoplasma capricolum* tRNAs with unmodified uridine or adenosine at the first anticodon position.** *J Mol Biol* 1995, **251(4):**486-492.
10. Yokobori S, Suzuki T, Watanabe K: **Genetic code variations in mitochondria: tRNA as a major determinant of genetic code plasticity.** *J Mol Evol* 2001, **53(4-5):**314-326.
11. Andachi Y, Yamao F, Iwami M, Muto A, Osawa S: **Occurrence of unmodified adenine and uracil at the first position of anticodon in threonine tRNAs in *Mycoplasma capricolum*.** *Proc Natl Acad Sci U S A* 1987, **84(21):**7398-7402.
12. Barrell BG, Anderson S, Bankier AT, de Bruijn MH, Chen E, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Roe BA, Sanger F, Schreier PH, Smith AJ, Staden R, Young IG: **Different pattern of codon recognition by mammalian mitochondrial tRNAs.** *Proc Natl Acad Sci U S A* 1980, **77(6):**3164-3166.
13. Bulmer M: **Coevolution of codon usage and transfer RNA abundance.** *Nature* 1987, **325(6106):**728-730.
14. Bulmer M: **The selection-mutation-drift theory of synonymous codon usage.** *Genetics* 1991, **129:**897-907.
15. Xia X: **How optimized is the translational machinery in *Escherichia coli*, *Salmonella typhimurium* and *Saccharomyces cerevisiae*?** *Genetics* 1998, **149(1):**37-44.
16. Ikemura T: **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* 1981, **151:**389-409.
17. Gouy M, Gautier C: **Codon usage in bacteria: correlation with gene expressivity.** *Nucleic Acids Res* 1982, **10:** 7055-7064.
18. Bennetzen JL, Hall BD: **Codon selection in yeast.** *J Biol Chem* 1982, **257(6):**3026-3031.
19. Akashi H: **Translational selection and yeast proteome evolution.** *Genetics* 2003, **164(4):**1291-1303.
20. Eyre-Walker A: **Synonymous codon bias is related to gene length in *Escherichia coli*: selection for translational accuracy?** *Mol Biol Evol* 1996, **13(6):**864-872.
21. Xia X: **Maximizing transcription efficiency causes codon usage bias.** *Genetics* 1996, **144:**1309-1320.
22. Jia W, Higgs PG: **Codon usage in mitochondrial genomes: distinguishing context-dependent mutation from translational selection.** *Mol Biol Evol* 2008, **25(2):**339-351.
23. Ogle JM, Ramakrishnan V: **Structural insights into translational fidelity.** *Annu Rev Biochem* 2005, **74:**129-177.
24. Mizuno H, Sundaralingam M: **Stacking of Crick Wobble pair and Watson-Crick pair: stability rules of G-U pairs at ends of helical stems in tRNAs and the relation to codon-anticodon Wobble interaction.** *Nucleic Acids Res* 1978, **5(11):**4451-4461.
25. Weixlbaumer A, Murphy FV, Dziergowska A, Malkiewicz A, Vendeix FA, Agris PF, Ramakrishnan V: **Mechanism for expanding the decoding capacity of transfer RNAs by modification of uridines.** *Nat Struct Mol Biol* 2007, **14(6):**498-502.

26. Masquida B, Westhof E: **On the wobble GoU and related pairs.** *RNA* 2000, **6(1):**9-15.
27. Carullo M, Xia X: **An Extensive Study of Mutation and Selection on the Wobble Nucleotide in tRNA Anticodons in Fungal Mitochondrial Genomes.** *J Mol Evol* 2008, **66(5):**484-493.
28. Xia X: **Body temperature, rate of biosynthesis and evolution of genome size.** *Mol Biol Evol* 1995, **12:**834-842.
29. Rocha EP: **Codon usage bias from tRNA's point of view: Redundancy, specialization, and efficient decoding for translation optimization.** *Genome Res* 2004, **12:**12.
30. Xia X, Huang H, Carullo M, Betran E, Moriyama EN: **Conflict between Translation Initiation and Elongation in Vertebrate Mitochondrial Genomes.** *PLoS ONE* 2007, **2:**e227.
31. Xia X, Xie Z: **DAMBE: Software package for data analysis in molecular biology and evolution.** *J Hered* 2001, **92(4):**371-373.
32. Xia X: **Data analysis in molecular biology and evolution.** Boston , Kluwer Academic Publishers; 2001:277.
33. Higgs PG, Jameson D, Jow H, Rattray M: **The evolution of tRNA-Leu genes in animal mitochondrial genomes.** *J Mol Evol* 2003, **57(4):**435-445.
34. Sengupta S, Higgs PG: **A unified model of codon reassignment in alternative genetic codes.** *Genetics* 2005, **170(2):**831-840.
35. Sengupta S, Yang X, Higgs PG: **The mechanisms of codon reassignments in mitochondrial genetic codes.** *J Mol Evol* 2007, **64(6):**662-688.
36. Sharp PM, Tuohy TM, Mosurski KR: **Codon usage in yeast: cluster analysis clearly differentiates highly and lowly expressed genes.** *Nucleic Acids Res* 1986, **14(13):**5125-5143.
37. Ling J, Roy H, Qin D, Rubio MA, Alfonzo JD, Fredrick K, Ibba M: **Pathogenic mechanism of a human mitochondrial tRNAPhe mutation associated with myoclonic epilepsy with ragged red fibers syndrome.** *Proc Natl Acad Sci U S A* 2007, **104(39):**15299-15304.
38. Grosjean H, Sprinzl M, Steinberg S: **Posttranscriptionally modified nucleosides in transfer RNA: their locations and frequencies.** *Biochimie* 1995, **77(1-2):**139-141.
39. Rocha EP, Danchin A: **Base composition bias might result from competition for metabolic resources.** *Trends Genet* 2002, **18(6):**291-294.
40. Xia X, Palidwor G: **Genomic Adaptation to Acidic Environment: Evidence from Helicobacter pylori.** *Am Nat* 2005, **166(6):**776-784.
41. Xia X, Wang HC, Xie Z, Carullo M, Huang H, Hickey DA: **Cytosine usage modulates the correlation between CDS length and CG content in prokaryotic genomes.** *Mol Biol Evol* 2006, **23(7):**1450-1454.
42. Gromadski KB, Daviter T, Rodnina MV: **A uniform response to mismatches in codon-anticodon complexes ensures ribosomal fidelity.** *Mol Cell* 2006, **21(3):**369-377.
43. Kurata S, Weixlbaumer A, Ohtsuki T, Shimazaki T, Wada T, Kirino Y, Takai K, Watanabe K, Ramakrishnan V, Suzuki T: **Modified Uridines with C5-methylene Substituents at the First Position of the tRNA Anticodon Stabilize U{middle dot}G Wobble Pairing during Decoding.** *J Biol Chem* 2008, **283(27):**18801-18811.
44. Rice P, Longden I, Bleasby A: **EMBOSS: the European Molecular Biology Open Software Suite.** *Trends Genet* 2000, **16(6):**276-277.
45. Haumont E, Fournier M, de Henau S, Grosjean H: **Enzymatic conversion of adenosine to inosine in the wobble position of yeast tRNAAsp: the dependence on the anticodon sequence.** *Nucleic Acids Res* 1984, **12(6):**2705-2715.
46. Grosjean H, Auxilien S, Constantinesco F, Simon C, Corda Y, Becker HF, Foiret D, Morin A, Jin YX, Fournier M, Fourrey JL: **Enzymatic conversion of adenosine to inosine and to N1-methylinosine in transfer RNAs: a review.** *Biochimie* 1996, **78(6):**488-501.
47. Auxilien S, Crain PF, Trewyn RW, Grosjean H: **Mechanism, specificity and general properties of the yeast enzyme catalysing the formation of inosine 34 in the anticodon of transfer RNA.** *J Mol Biol* 1996, **262(4):**437-458.