Brief Report

Reconsideration of codon-anticodon wobble pairings in translation: decoding tables

Miki Wada *

Posted Date: 27 December 2023

doi: 10.20944/preprints202312.2074.v1

Keywords: decoding; tRNA; wobble pairing; genetic code; inosine; translation doi:10.20944/preprints202312.2074.v1

Brief Report

Reconsideration of Codon-Anticodon Wobble Pairings in Translation: Decoding Tables

Miki Wada

Laboratory of Molecular Genetics, Department of Computational Biology and Medical Sciences, Graduate School of Frontier Sciences, The University of Tokyo, Kashiwa-city, Chiba 277-8562, Japan.
* Correspondence: Miki Wada. Email: miki-wada@edu.k.u-tokyo.ac.jp, mikiwada01@gmail.com

Abstract: The gene translation system possesses inherent complexity in distinguishing codons with degeneracy through repertoires of aminoacyl-tRNAs in the ribosome. Implementation of inosine nucleoside at the anticodon wobble-position in certain tRNAs, especially in eukaryotes, might be arisen to establish complex decoding systems. This report introduces the novel concept of "decoding tables", which serve as concise summaries of codon-anticodon pairings provided in codon table-like format. Utilizing tRNA genome database information and a procedure based on compiled tRNA studies, the tRNA repertoire and the use of wobble pairings can be depicted for any organisms. The decoding table sheds light on the use of wobble pairings across diverse organisms.

Keywords: decoding; tRNA; wobble pairing; genetic code; inosine; translation

---

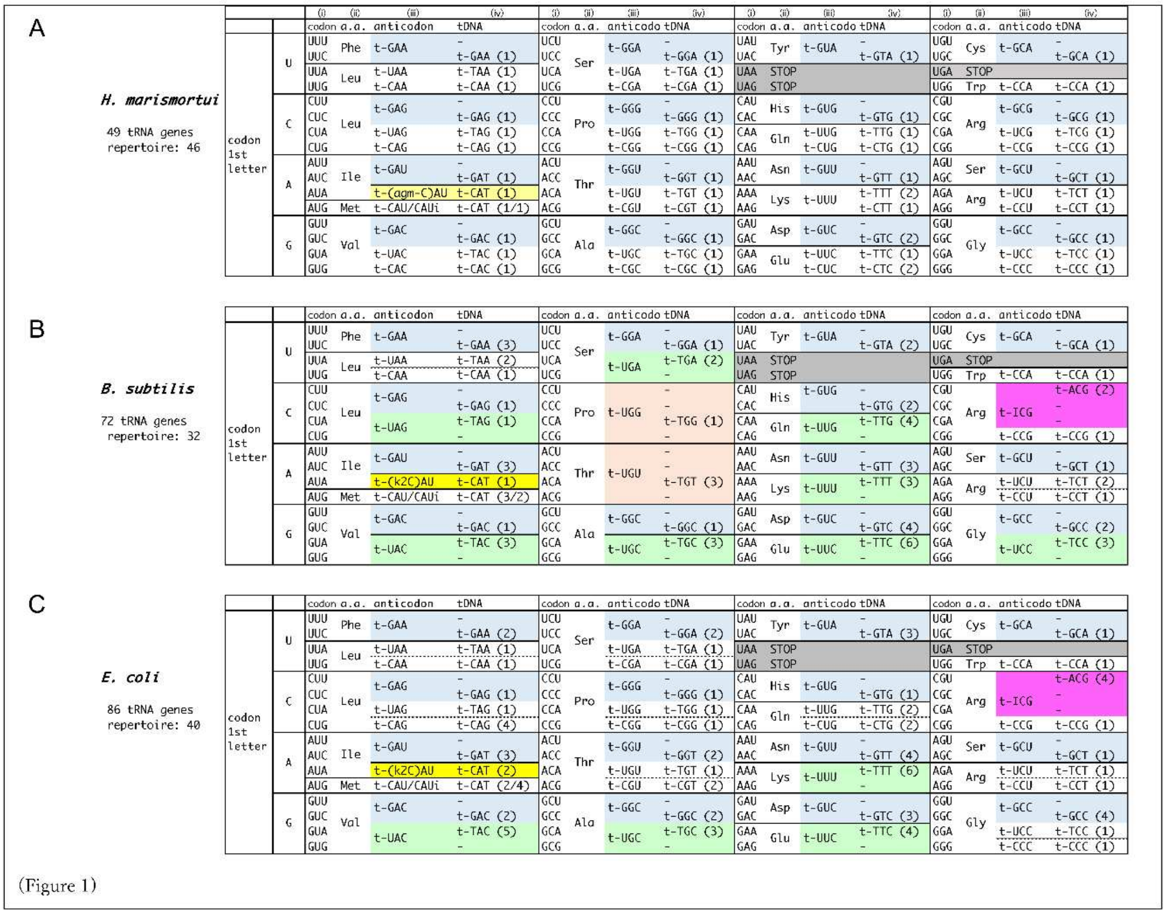
Introduction

Francis Click wrote "wobble hypothesis" in 1966 [1], and explained wobble base parings between inosine nucleoside (hypoxanthine as a base) of tRNA anticodon and codon bases, after the tRNA^{Ala} structure from brewery yeast was reported to have inosine [2] and was located in the first position of the anticodon, wobble position [3]. Subsequent studies demonstrated that utilization of inosine in the wobble position is ubiquitous, especially in eukaryotes, but not conserved in all applicable tRNAs of organisms. Since then, tRNA modifications, aminoacylation of tRNAs, the ribosomes and translation machinery were intensively studied revealing aspects of the translation systems.

As tRNA genome databases and analysis programs have been established in current genome era, starting from the tRNA genomic genes, reconsideration of codon-anticodon pairings should be useful to deepen understanding of translation of the genes. Here I propose a concept of "decoding tables" for genetic code as a conceptual framework. The decoding table depict codon-tRNA anticodon pairings, based on tRNA genomic gene information (tDNA)[4,5] and experimentally elucidated general features of tRNA. The use of wobble position pairings is specific to individual organisms. Among a lot of tRNA modifications, the decoding tables selectively include modifications at the anticodon site that affect selection of the pairing codon. The tables have limitations: 1) tRNA modifications that do not influence codon selection are not included, and 2) contextual effects are not considered. Despite these limitations, I believe that the decoding tables are crucial for understanding the codon-tRNA anticodon pairings in functional terms. Creatures often employ idiosyncratic system and possess complex compensatory mechanisms to survive and react to diverse situations and stresses. Given the intricate nature of those studies, achieving complete understanding may remain elusive, underscoring the importance of presenting current knowledge concisely.

2. tRNA repertoires

Genome analyses have revealed variations in the number of tRNA genes and distinct tRNA repertoires among individual organisms. The tRNA repertoire refers to the number of tRNA genes defined by anticodons. For example, the *Saccharomyces cerevisiae* genome has 275 tRNA genes, constituting a 41 tRNA repertoire with different anticodons, to decode 61 sense codons, except for termination codons (GtRNAdb, http://gtrnadb.ucsc.edu) (SGD, https://www.yeastgenome.org)[4,6]. The number of tRNA genes varies substantially across organisms, ranging from fewer than 35 (e,g., *Ureplasma urealyticum*, and *Borrelia. burgdorferi*) to more than 3000 (e,g., *Xenopus tropicalis*, and *Danio rerio*). The number of tRNA repertoire typically falls within 32 to 54. No organism has a 61 tRNA repertoire corresponding to 61 sense codons, i.e., any of tRNA repertoire should decipher more than one codon. Considering the different tRNA repertoires among organisms, we propose the concept of "decoding tables" that describe the codon-anticodon pairings in each organism, comparatively showing the features of codon-anticodon pairings among organisms (Figure 1). The heart of the decoding table is wobble base-pairings [1], that is allowed non-Watson-Crick pairing between the 3rd base of the mRNA codon and the 1st base of the tRNA anticodon (position 34 of tRNA). Although tRNAs are heavily modified[7–9], the tables include only major modifications at anticodons known to change codon specificity. The wobble position inosine nucleoside (hypoxanthine as the base) edited from adenine is essential in the decoding, having pairing ability with C-ending codons in addition to U-ending codons [10], in eukaryotes. Lysidine and agmatidine modified from cytidine, in bacteria and in archaea respectively, play a specific role in deciphering AUA codon as isoleucine, distinguishing from AUG methionine codon and AUG initiation codon[11,12]. Eukaryotic mitochondrial tRNAs are not included in this report.



A

H. marismortui

49 tRNA genes
repertoire: 46

		(i)	(ii)	(iii)	(iv)	(i)	(ii)	(iii)	(iv)	(i)	(ii)	(iii)	(iv)	(i)	(ii)	(iii)	(iv)
		codon	a.a.	anticodon	tDNA	codon	a.a.	anticodo	tDNA	codon	a.a.	anticodo	tDNA	codon	a.a.	anticodo	tDNA
codon 1st letter	U	UUU	Phe	t-GAA	-	UCU	Ser	t-GGA	-	UAU	Tyr	t-GUA	-	UGU	Cys	t-GCA	-
		UUC			t-GAA (1)	UCC			t-GGA (1)	UAC			t-GTA (1)	UGC			t-GCA (1)
		UUA	Leu	t-UAA	t-TAA (1)	UCA		t-UGA	t-TGA (1)	UAA	STOP			UGA	STOP		
		UUG		t-CAA	t-CAA (1)	UCG		t-CGA	t-CGA (1)	UAG	STOP			UGG	Trp	t-CCA	t-CCA (1)
	C	CUU	Leu	t-GAG	-	CCU	Pro	t-GGG	-	CAU	His	t-GUG	-	CGU	Arg	t-GCG	-
		CUC			t-GAG (1)	CCC			t-GGG (1)	CAC			t-GTG (1)	CGC			t-GCG (1)
		CUA		t-UAG	t-TAG (1)	CCA		t-UGG	t-TGG (1)	CAA	Gln	t-UUG	t-TTG (1)	CGA		t-UCG	t-TCG (1)
		CUG		t-CAG	t-CAG (1)	CCG		t-CGG	t-CGG (1)	CAG		t-CUG	t-CTG (1)	CGG		t-CCG	t-CCG (1)
	A	AUU	Ile	t-GAU	-	ACU	Thr	t-GGU	-	AAU	Asn	t-GUU	-	AGU	Ser	t-GCU	-
		AUC			t-GAT (1)	ACC			t-GGT (1)	AAC			t-GTT (1)	AGC			t-GCT (1)
		AUA		t-(agm-C)AU	t-CAT (1)	ACA		t-UGU	t-TGT (1)	AAA	Lys	t-UUU	t-TTT (2)	AGA	Arg	t-UCU	t-TCT (1)
		AUG	Met	t-CAU/CAUi	t-CAT (1/1)	ACG		t-CGU	t-CGT (1)	AAG			t-CTT (1)	AGG		t-CCU	t-CCT (1)
	G	GUU	Val	t-GAC	-	GCU	Ala	t-GGC	-	GAU	Asp	t-GUC	-	GGU	Gly	t-GCC	-
		GUC			t-GAC (1)	GCC			t-GGC (1)	GAC			t-GTC (2)	GGC			t-GCC (1)
		GUA		t-UAC	t-TAC (1)	GCA		t-UGC	t-TGC (1)	GAA	Glu	t-UUC	t-TTC (1)	GGA		t-UCC	t-TCC (1)
		GUG		t-CAC	t-CAC (1)	GCG		t-CGC	t-CGC (1)	GAG		t-CUC	t-CTC (2)	GGG		t-CCC	t-CCC (1)

B

B. subtilis

72 tRNA genes
repertoire: 32

		codon	a.a.	anticodon	tDNA	codon	a.a.	anticodo	tDNA	codon	a.a.	anticodo	tDNA	codon	a.a.	anticodo	tDNA
codon 1st letter	U	UUU	Phe	t-GAA	-	UCU	Ser	t-GGA	-	UAU	Tyr	t-GUA	-	UGU	Cys	t-GCA	-
		UUC			t-GAA (3)	UCC			t-GGA (1)	UAC			t-GTA (2)	UGC			t-GCA (1)
		UUA	Leu	t-UAA	t-TAA (2)	UCA		t-UGA	t-TGA (2)	UAA	STOP			UGA	STOP		
		UUG		t-CAA	t-CAA (1)	UCG			-	UAG	STOP			UGG	Trp	t-CCA	t-CCA (1)
	C	CUU	Leu	t-GAG	-	CCU	Pro	t-UGG	-	CAU	His	t-GUG	-	CGU	Arg	t-ICG	t-ACG (2)
		CUC			t-GAG (1)	CCC			-	CAC			t-GTG (2)	CGC			-
		CUA		t-UAG	t-TAG (1)	CCA			t-TGG (1)	CAA	Gln	t-UUG	t-TTG (4)	CGA			-
		CUG			-	CCG			-	CAG			-	CGG		t-CCG	t-CCG (1)
	A	AUU	Ile	t-GAU	-	ACU	Thr	t-UGU	-	AAU	Asn	t-GUU	-	AGU	Ser	t-GCU	-
		AUC			t-GAT (3)	ACC			-	AAC			t-GTT (3)	AGC			t-GCT (1)
		AUA		t-(k2C)AU	t-CAT (1)	ACA			t-TGT (3)	AAA	Lys	t-UUU	t-TTT (3)	AGA	Arg	t-UCU	t-TCT (2)
		AUG	Met	t-CAU/CAUi	t-CAT (3/2)	ACG			-	AAG			-	AGG		t-CCU	t-CCT (1)
	G	GUU	Val	t-GAC	-	GCU	Ala	t-GGC	-	GAU	Asp	t-GUC	-	GGU	Gly	t-GCC	-
		GUC			t-GAC (1)	GCC			t-GGC (1)	GAC			t-GTC (4)	GGC			t-GCC (2)
		GUA		t-UAC	t-TAC (3)	GCA		t-UGC	t-TGC (3)	GAA	Glu	t-UUC	t-TTC (6)	GGA		t-UCC	t-TCC (3)
		GUG			-	GCG			-	GAG			-	GGG			-

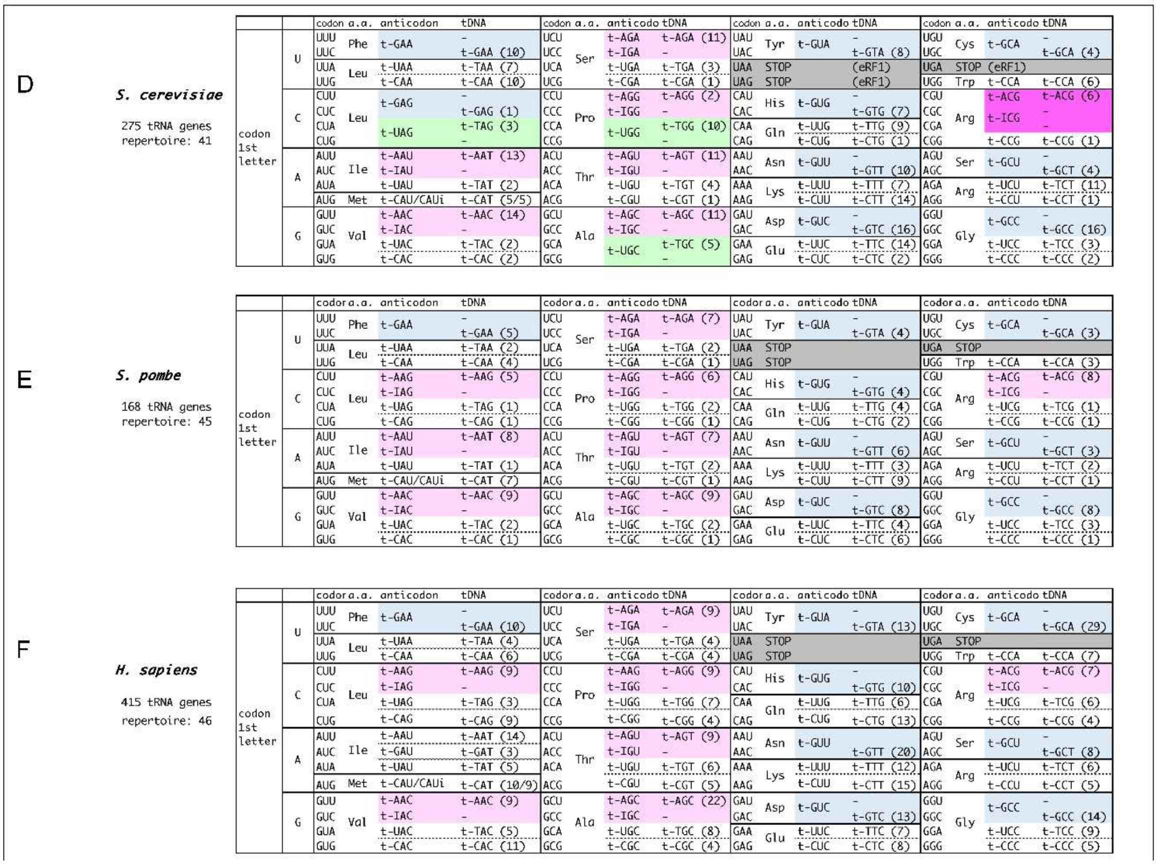
C

E. coli

86 tRNA genes
repertoire: 40

		codon	a.a.	anticodon	tDNA	codon	a.a.	anticodo	tDNA	codon	a.a.	anticodo	tDNA	codon	a.a.	anticodo	tDNA
codon 1st letter	U	UUU	Phe	t-GAA	-	UCU	Ser	t-GGA	-	UAU	Tyr	t-GUA	-	UGU	Cys	t-GCA	-
		UUC			t-GAA (2)	UCC			t-GGA (2)	UAC			t-GTA (3)	UGC			t-GCA (1)
		UUA	Leu	t-UAA	t-TAA (1)	UCA		t-UGA	t-TGA (1)	UAA	STOP			UGA	STOP		
		UUG		t-CAA	t-CAA (1)	UCG		t-CGA	t-CGA (1)	UAG	STOP			UGG	Trp	t-CCA	t-CCA (1)
	C	CUU	Leu	t-GAG	-	CCU	Pro	t-GGG	-	CAU	His	t-GUG	-	CGU	Arg	t-ICG	t-ACG (4)
		CUC			t-GAG (1)	CCC			t-GGG (1)	CAC			t-GTG (1)	CGC			-
		CUA		t-UAG	t-TAG (1)	CCA		t-UGG	t-TGG (1)	CAA	Gln	t-UUG	t-TTG (2)	CGA			-
		CUG		t-CAG	t-CAG (4)	CCG		t-CGG	t-CGG (1)	CAG		t-CUG	t-CTG (2)	CGG		t-CCG	t-CCG (1)
	A	AUU	Ile	t-GAU	-	ACU	Thr	t-GGU	-	AAU	Asn	t-GUU	-	AGU	Ser	t-GCU	-
		AUC			t-GAT (3)	ACC			t-GGT (2)	AAC			t-GTT (4)	AGC			t-GCT (1)
		AUA		t-(k2C)AU	t-CAT (2)	ACA		t-UGU	t-TGT (1)	AAA	Lys	t-UUU	t-TTT (6)	AGA	Arg	t-UCU	t-TCT (1)
		AUG	Met	t-CAU/CAUi	t-CAT (2/4)	ACG		t-CGU	t-CGT (2)	AAG			-	AGG		t-CCU	t-CCT (1)
	G	GUU	Val	t-GAC	-	GCU	Ala	t-GGC	-	GAU	Asp	t-GUC	-	GGU	Gly	t-GCC	-
		GUC			t-GAC (2)	GCC			t-GGC (2)	GAC			t-GTC (3)	GGC			t-GCC (4)
		GUA		t-UAC	t-TAC (5)	GCA		t-UGC	t-TGC (3)	GAA	Glu	t-UUC	t-TTC (4)	GGA		t-UCC	t-TCC (1)
		GUG			-	GCG			-	GAG			-	GGG		t-CCC	t-CCC (1)

(Figure 1)



D

S. cerevisiae

275 tRNA genes
repertoire: 41

		codon	a.a.	anticodon	tDNA	codon	a.a.	anticodo	tDNA	codon	a.a.	anticodo	tDNA	codon	a.a.	anticodo	tDNA
codon 1st letter	U	UUU	Phe	t-GAA	-	UCU	Ser	t-AGA	t-AGA (11)	UAU	Tyr	t-GUA	-	UGU	Cys	t-GCA	-
		UUC			t-GAA (10)	UCC		t-IGA	-	UAC			t-GTA (8)	UGC			t-GCA (4)
		UUA	Leu	t-UAA	t-TAA (7)	UCA		t-UGA	t-TGA (3)	UAA	STOP		(eRF1)	UGA	STOP	(eRF1)	
		UUG		t-CAA	t-CAA (10)	UCG		t-CGA	t-CGA (1)	UAG	STOP		(eRF1)	UGG	Trp	t-CCA	t-CCA (6)
	C	CUU	Leu	t-GAG	-	CCU	Pro	t-AGG	t-AGG (2)	CAU	His	t-GUG	-	CGU	Arg	t-ACG	t-ACG (6)
		CUC			t-GAG (1)	CCC		t-IGG	-	CAC			t-GTG (7)	CGC		t-ICG	-
		CUA		t-UAG	t-TAG (3)	CCA		t-UGG	t-TGG (10)	CAA	Gln	t-UUG	t-TTG (9)	CGA			-
		CUG			-	CCG			-	CAG		t-CUG	t-CTG (1)	CGG		t-CCG	t-CCG (1)
	A	AUU	Ile	t-AAU	t-AAT (13)	ACU	Thr	t-AGU	t-AGT (11)	AAU	Asn	t-GUU	-	AGU	Ser	t-GCU	-
		AUC		t-IAU	-	ACC		t-IGU	-	AAC			t-GTT (10)	AGC			t-GCT (4)
		AUA		t-UAU	t-TAT (2)	ACA		t-UGU	t-TGT (4)	AAA	Lys	t-UUU	t-TTT (7)	AGA	Arg	t-UCU	t-TCT (11)
		AUG	Met	t-CAU/CAUi	t-CAT (5/5)	ACG		t-CGU	t-CGT (1)	AAG		t-CUU	t-CTT (14)	AGG		t-CCU	t-CCT (1)
	G	GUU	Val	t-AAC	t-AAC (14)	GCU	Ala	t-AGC	t-AGC (11)	GAU	Asp	t-GUC	-	GGU	Gly	t-GCC	-
		GUC		t-IAC	-	GCC		t-IGC	-	GAC			t-GTC (16)	GGC			t-GCC (16)
		GUA		t-UAC	t-TAC (2)	GCA		t-UGC	t-TGC (5)	GAA	Glu	t-UUC	t-TTC (14)	GGA		t-UCC	t-TCC (3)
		GUG		t-CAC	t-CAC (2)	GCG			-	GAG		t-CUC	t-CTC (2)	GGG		t-CCC	t-CCC (2)

E

S. pombe

168 tRNA genes
repertoire: 45

		codor	a.a.	anticodon	tDNA	codor	a.a.	anticodo	tDNA	codor	a.a.	anticodo	tDNA	codor	a.a.	anticodo	tDNA
codon 1st letter	U	UUU	Phe	t-GAA	-	UCU	Ser	t-AGA	t-AGA (7)	UAU	Tyr	t-GUA	-	UGU	Cys	t-GCA	-
		UUC			t-GAA (5)	UCC		t-IGA	-	UAC			t-GTA (4)	UGC			t-GCA (3)
		UUA	Leu	t-UAA	t-TAA (2)	UCA		t-UGA	t-TGA (2)	UAA	STOP			UGA	STOP		
		UUG		t-CAA	t-CAA (4)	UCG		t-CGA	t-CGA (1)	UAG	STOP			UGG	Trp	t-CCA	t-CCA (3)
	C	CUU	Leu	t-AAG	t-AAG (5)	CCU	Pro	t-AGG	t-AGG (6)	CAU	His	t-GUG	-	CGU	Arg	t-ACG	t-ACG (8)
		CUC		t-IAG	-	CCC		t-IGG	-	CAC			t-GTG (4)	CGC		t-ICG	-
		CUA		t-UAG	t-TAG (1)	CCA		t-UGG	t-TGG (2)	CAA	Gln	t-UUG	t-TTG (4)	CGA		t-UCG	t-TCG (1)
		CUG		t-CAG	t-CAG (1)	CCG		t-CGG	t-CGG (1)	CAG		t-CUG	t-CTG (2)	CGG		t-CCG	t-CCG (1)
	A	AUU	Ile	t-AAU	t-AAT (8)	ACU	Thr	t-AGU	t-AGT (7)	AAU	Asn	t-GUU	-	AGU	Ser	t-GCU	-
		AUC		t-IAU	-	ACC		t-IGU	-	AAC			t-GTT (6)	AGC			t-GCT (3)
		AUA		t-UAU	t-TAT (1)	ACA		t-UGU	t-TGT (2)	AAA	Lys	t-UUU	t-TTT (3)	AGA	Arg	t-UCU	t-TCT (2)
		AUG	Met	t-CAU/CAUi	t-CAT (7)	ACG		t-CGU	t-CGT (1)	AAG		t-CUU	t-CTT (9)	AGG		t-CCU	t-CCT (1)
	G	GUU	Val	t-AAC	t-AAC (9)	GCU	Ala	t-AGC	t-AGC (9)	GAU	Asp	t-GUC	-	GGU	Gly	t-GCC	-
		GUC		t-IAC	-	GCC		t-IGC	-	GAC			t-GTC (8)	GGC			t-GCC (8)
		GUA		t-UAC	t-TAC (2)	GCA		t-UGC	t-TGC (2)	GAA	Glu	t-UUC	t-TTC (4)	GGA		t-UCC	t-TCC (3)
		GUG		t-CAC	t-CAC (1)	GCG		t-CGC	t-CGC (1)	GAG		t-CUC	t-CTC (6)	GGG		t-CCC	t-CCC (1)

F

H. sapiens

415 tRNA genes
repertoire: 46

		codor	a.a.	anticodon	tDNA	codor	a.a.	anticodo	tDNA	codor	a.a.	anticodo	tDNA	codor	a.a.	anticodo	tDNA
codon 1st letter	U	UUU	Phe	t-GAA	-	UCU	Ser	t-AGA	t-AGA (9)	UAU	Tyr	t-GUA	-	UGU	Cys	t-GCA	-
		UUC			t-GAA (10)	UCC		t-IGA	-	UAC			t-GTA (13)	UGC			t-GCA (29)
		UUA	Leu	t-UAA	t-TAA (4)	UCA		t-UGA	t-TGA (4)	UAA	STOP			UGA	STOP		
		UUG		t-CAA	t-CAA (6)	UCG		t-CGA	t-CGA (4)	UAG	STOP			UGG	Trp	t-CCA	t-CCA (7)
	C	CUU	Leu	t-AAG	t-AAG (9)	CCU	Pro	t-AAG	t-AGG (9)	CAU	His	t-GUG	-	CGU	Arg	t-ACG	t-ACG (7)
		CUC		t-IAG	-	CCC		t-IGG	-	CAC			t-GTG (10)	CGC		t-ICG	-
		CUA		t-UAG	t-TAG (3)	CCA		t-UGG	t-TGG (7)	CAA	Gln	t-UUG	t-TTG (6)	CGA		t-UCG	t-TCG (6)
		CUG		t-CAG	t-CAG (9)	CCG		t-CGG	t-CGG (4)	CAG		t-CUG	t-CTG (13)	CGG		t-CCG	t-CCG (4)
	A	AUU	Ile	t-AAU	t-AAT (14)	ACU	Thr	t-AGU	t-AGT (9)	AAU	Asn	t-GUU		AGU	Ser	t-GCU	-
		AUC		t-GAU	t-GAT (3)	ACC		t-IGU	-	AAC			t-GTT (20)	AGC			t-GCT (8)
		AUA		t-UAU	t-TAT (5)	ACA		t-UGU	t-TGT (6)	AAA	Lys	t-UUU	t-TTT (12)	AGA	Arg	t-UCU	t-TCT (6)
		AUG	Met	t-CAU/CAUi	t-CAT (10/9)	ACG		t-CGU	t-CGT (5)	AAG		t-CUU	t-CTT (15)	AGG		t-CCU	t-CCT (5)
	G	GUU	Val	t-AAC	t-AAC (9)	GCU	Ala	t-AGC	t-AGC (22)	GAU	Asp	t-GUC	-	GGU	Gly	t-GCC	-
		GUC		t-IAC	-	GCC		t-IGC	-	GAC			t-GTC (13)	GGC			t-GCC (14)
		GUA		t-UAC	t-TAC (5)	GCA		t-UGC	t-TGC (8)	GAA	Glu	t-UUC	t-TTC (7)	GGA		t-UCC	t-TCC (9)
		GUG		t-CAC	t-CAC (11)	GCG		t-CGC	t-CGC (4)	GAG		t-CUC	t-CTC (8)	GGG		t-CCC	t-CCC (5)

Figure 1. A-F. Organism-specific decoding tables. The decoding tables are depicted according to the rules described in "basis of the decoding tables". The total number of tRNA genes and tRNA repertoire are shown below the species name. (i) Codon, (ii) amino acid (aa), (iii) tRNA anticodon, and (iv) tDNA anticodon (tDNA) are shown within the boxes, in this order. "I" denotes the base component of the inosine nucleoside, hypoxanthine, in the tables. In parentheses within (iv), the tRNA gene numbers with the tDNA anticodon (not necessarily identical) are indicated. The anticodon shows inferred tRNA anticodon residues corresponding to the codon. Only the anticodon editing and modifications that change (or expand) codon pairing are included.

3. Basis of the decoding tables

The decoding tables for six organisms from three domains of life are depicted in Figure 1. The rules implemented to depict the decoding tables are as follows:

1. A set of decoding information comprises (i) codons, (ii) amino acids, (iii) tRNA anticodons, and (iv) tDNA anticodons, described in this order within the boxes. The number of corresponding tRNA genes with the anticodon is shown in the parentheses within (iv). In short, (i) and (iii) represent the codon-anticodon pairing.
2. tDNA anticodons (iv) are essentially derived from the tRNA genome database [4], and placed by appropriate codon-anticodon pairings. When a codon (i) forms complete Watson-Crick base pairs with a tDNA anticodon (iv) (either adenine (A)- thymine (T) or guanine (G)- cytosine (C)), the tRNA anticodon (iii) is corresponding RNA anticodon to the tDNA anticodon.
3. When a codon does not make perfect match pairing with any of tDNA anticodons, potential wobble-pairing is introduced into the 1st base of the anticodon (position 34 of tRNA) (iii) from the tDNA anticodon (iv).
    3-1) The anticodon 1st G is assumed to pair with the codon 3rd U, in addition to 3rd C (shown in light blue in Figure 1), as G-U pairing in secondary structure is often observed in RNA (shown in pale blue). [13–15]
    3-2) The G-U pairing allows the anticodon 1st U to pair with the codon 3rd G, in addition to 3rd A (shown in light green).

3-3) In some codon boxes (codons that share 1st and 2nd bases), only one tRNA repertoire with anticodon 1st U is assigned. In these cases, superwobbling, which describes the anticodons with 1st U can pair with any of codon 3rd base, is implemented (shown in orange). Superwobbling has been evidenced in some organisms [16,17].

3-4) Where the anticodon 1st G is not assigned but 1st A in codon boxes, inosine (I) is implemented to pair with both the codon 3rd C and 3rd U (shown in pale pink). In special cases of tRNA^{Arg} (ICG), the anticodon 1st I pairs with the codon 3rd A in addition to 3rd C and U, and the type of pairing is called A-I wobble, as is observed in *Escherichia coli* and *S. cerevisiae* (shown in pink) [18]. Intriguingly, the anticodon 1st inosine does not necessarily pair with U, allowing the possibility of the precursor tRNA with anticodon 1st A to be utilized. In this meaning, the extent of usage of the anticodon 1st inosine containing tRNAs (i-tRNAs) is yet to be determined, though i-tRNAs are definitely necessary to decipher C-ending codons in the boxes.

4. In prokaryotes and archaea, lysidine(k2) and agmatidine(agm) modifications of the anticodon 1st C of CAU are used to decipher AUA codon, respectively, according to previous reports. These k2-CAU and amg-CAU are deciphered by tRNA^{Ile} as isoleucine, distinguishing AUG codon for methionine (shown yellow) [11,12].
5. Variant genetic codes should be considered for organisms in use [19,20], however, these organisms are not included in Figure 1.

Following the scheme described above, a decoding table can be constructed for any organisms using the genome tRNA database (or genome sequence data and a tRNA gene finding program), as a concise summary. The decoding tables described in Figure 1 are illustrative models. The anticodon defining procedure ensures no inconsistency with previous reports in relation to codon- anticodon pairings. The accuracy of the decoding tables largely relies on tRNA genome database, i.e., the tRNA-finding algorithm, which may not be entirely precise, especially in higher eukaryotes with complex genome structures.

4. Comparative wobble pairing usage

The comparative decoding tables for the organisms, archaeal *Haloarcula maristortui*, bacterial *Bacillus subtilis* and *Escherichia coli*, eukaryotic *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, and *Homo sapiens,* are depicted in Figure 1. Wobble usage is denoted by colour, revealing domain-specific tendencies and additional organism-specificities. Inosine-containing i-tRNAs are one in *E. coli*, seven in *S. cerevisiae*, eight in *S. pombe*, and seven in *H. sapiens*. The i-tRNAs are predominantly utilized in eukaryotes [21,22]. The inosines in i-tRNAs are edited from adenines by the oxidative deamination via the Tad2/Tad3 complex in *S. cerevisiae*, ADAT2/ADAT3 in *H. sapiens*, and TadA in *E. coli* [10,23,24]. The inosine adapts a guanine-like structure for base pairing, and read as G in canonical nucleotide sequencing [10,18]. The tRNA^{Arg}(ICG) pairs with U, C and A, that may be contributed tautomerism in the decoding. The type of pairings is restricted to bacteria and a few yeasts, and is the only one i-tRNA in bacteria. The other i-tRNAs are used only in eukaryotes.

Uracil plays a substantial role in wobble pairing, demonstrating the widest pairing capacity [16,21]. In *B subtilis*, tRNA^{Pro} and tRNA^{Thr} with wobble position U correspond to four codons (shown in pale orange), which are all utilized according to the codon usage of *B. subtilis* (Codon Usage Database: https://www.kazusa.or.jp/codon). The type of 4-codon-decoding tRNAs are restricted to the codon boxes in which amino acids are not defined by the codon's 3rd base, and are observed in certain bacteria. In the pairings, probably unmodified or lightly modified tRNA wobble U likely pairs with any of four bases with varying levels of efficiency. On the other hand, anticodon 1st U is variously modified depending on tRNAs (modomics, https://iimcb.genesilico.pl/modomics)[8]. Examples of such modifications for *E. coli* and *S. cerevisiae* are shown in Figures S1 and Figure S2. tRNA anticodon modifications, other than inosine, that affect decoding appears to have been introduced to distinguish codons, thereby increasing codon-anticodon specificity.

Although the decoding tables do not indicate decoding efficiency in the ribosome, they reveal intriguing deviations in translation systems among species and the evolution of tRNA repertoires.

The decoding tables may offer valuable insights into the significance of codon degeneracy, as the codon usage in mRNAs and the repertoire of tRNAs directly impact the GC content (%) of organisms, especially those with a small genome size. Based on the decoding tables, codon degeneracy appears to be defined by the supplied aminoacyl-tRNA repertoires rather than wobble codon acceptability of tRNAs, suggesting a presumed co-evolution. Furthermore, the decoding tables may provide important information for heterologous expression in genetic engineering.

5. Around initiation and termination codons

Two distinct features in the decoding tables exist around AUG codon, and around UGA stop codon. The AUG codon and UGA stop codons must have been needed to establish orthogonality to be used as initiation and stop codon, respectively. To distinguish AUA codon from AUG codon, archaea and bacteria appear to have established different tRNA anticodon 1st C modifications, namely agmatidine and lysidine (depicted in Figure 1A for *H. marismortui*, Figure 1B for *B. subtilis*, and Figure 1C for *E. coli*). In contrast, eukaryotes somehow established specific tRNA(UAU) to read only AUA codon. In *S. cerevisiae*, two anticodon uracyl bases of tRNA(UAU) are modified to pseudouridines to be distinguished (Figure S2). The lysidine and agmatidine modifications are introduced into the pyrimidine-ring-C2 of position 34C base, conversely, most of modifications of position 34U base are introduced into the pyrimidine-ring-C5. These modifications are positioned in structurally opposite directions [25].

As for UGA stop codon, most of organisms have established orthogonality with a specific $tRNA^{Trp}$(CCA) to decipher UGG codon exclusively. However, the A-I wobble observed in certain bacteria and yeasts to decipher CGA codon might be a remnant of the evolutionary process of distinguishing UGA codon. Bacteria utilize two peptide release factors, RF1 for UAA and UAG stop codons and RF2 for UAA and UGA stop codons [26]. In eukaryotes, omnipotent eRF1 deciphers all three stop codons including UGA codon, which established orthogonality with UGG. Instead, CGA codon, which is one base difference with UGA at non-wobble position, is deciphered by eRF1/eRF3 complex in misdecoding level to release polypeptide under specific conditions [27]. It is of interesting to investigate whether the A-I wobble is strategically employed in relation to translation termination. The introduction of initiation codon and stop codons into the genetic code must have been a complex process.

6. Evaluation of the decoding tables

The decoding tables show the most suitable codon-anticodon pairings deduced from the databases and experimental reports. While reports describe the overlapping of decoding codons, the decoding tables do not consider multiple tRNA anticodons for a codon. It should be reminded in relation to stop codons, such as variant genetic code, and fidelity of the decoding of sense and stop codons. While the concise decoding tables provide a clear overview, detailed information on the extent of these effects is challenging to ascertain. Reconstitution experiments *in vitro* face difficulties in setting up conditions, and *in vivo* experiments often trigger compensatory reactions, even though they surely demonstrate aspects of the decoding of genetic codes. The decoding tables effectively illustrate codon-anticodon pairings, particularly highlighting tRNA repertoires, given their foundation on a compilation of reports.

From the perspective of aminoacylation, typically only one aminoacyl-tRNA synthetase (aa-RS) is responsible for adding each amino acid to the corresponding tRNA [28,29]. Evaluating the decoding tables in the context of aminoacylation is crucial. Each aa-RS recognizes specific elements in its substrate tRNAs, known as identity elements [29,30]. Among the anticodons, positions 35 and 36 bases (2nd and 3rd bases of anticodons) are included in identity elements of not a few of aminoacyl-tRNAs. Additionally, position 34G is included in the identity elements of some tRNAs, such as $tRNA^{Cys}$ $tRNA^{Tyr}$, confirming the anticodon 1st G containing tRNAs are the substrates of the aa-RSs. On the contrary, A or I at tRNA position 34 are not included in identity elements. This raises an important question regarding whether position 34A, the precursor of position 34I, is used for decoding or not. It should be clarified at a general level. Answers might have been within reported data hidden to be analyzed to a certain level. Some reports indicate position 34 I/A are involved in some regulation [31,32], but not intensively studied.

The recognition of tRNA^{Ile}(k2-CAU) and tRNA^{Ile}(agm-CAU) by isoleucyl-tRNA synthetases is rather exceptional, and reported that modified position 34 bases are included in the identity elements of their synthetases [29]. Thus, for these tRNAs, the indicated modifications are prerequisites for aminoacylation. It is an interesting contrast that the position 34I and 34A are not included in identity elements.

The decoding tables apply to organisms without major discrepancies. They shed light on codon-tRNA anticodon pairings in the ribosome which might be unconsciously overlooked under uniform translation system from genes to proteins. The decoding tables contribute to our understanding of tRNA repertoires and the decoding in translation in organisms, and evolution of translation systems. It is of interesting to consider the evolution of translation machinery that decipher genetic codes, essential to all organisms. The intricacies of decoding of the codons in the ribosome including related factors biogenesis, is what a contrast with resulting decoding products observed as polypeptides.

Supplementary Materials: The following supporting information can be downloaded at the website of this paper posted on Preprints.org; Figure S1 and S2.

Funding: This research received no external funding.

Acknowledge: I thank Prof. Koichi Ito for kind discussions and comments. I thank GSFS, the University of Tokyo, for treating me as a guest scientist.

Conflicts of Interest: The author declares no conflict of interest.

Abbreviation

i-tRNA: position 34 inosine containing tRNA

References

1. Crick:F.H.C. [1966] Codon—anticodon pairing: The wobble hypothesis. *Journal of Molecular Biology*, **19**, 548–555.
2. Holley,R. [1965] Structure of an alanine transfer ribonucleic acid. *science*, **194**, 868–871.
3. Madison,J., Eferett,G. and Kung,H. [1966] Nucleotide sequence of a yeast tyrosine transfer RNA. *science*, **153**, 531–534.
4. Lowe,T.M. and Chan,P.P. [2016] tRNAscan-SE On-line: integrating search and context for analysis of transfer RNA genes. *Nucleic Acids Research*, **44**, W54–W57.
5. Chan,P.P., Lin,B.Y., Mak,A.J. and Lowe,T.M. [2021] classification of transfer RNA genes. **49**, 9077–9096.
6. Cherry,J.M., Hong,E.L., Amundsen,C., Balakrishnan,R., Binkley,G., Chan,E.T., Christie,K.R., Costanzo,M.C., Dwight,S.S., Engel,S.R., *et al.* [2012] Saccharomyces Genome Database: The genomics resource of budding yeast. *Nucleic Acids Research*, **40**, 700–705.
7. Dunin-Horkawicz,S., Czerwoniec,A., Gajda,M.J., Feder,M., Grosjean,H. and Bujnicki,J.M. [2006] MODOMICS: a database of RNA modification pathways. *Nucleic acids research*, **34**, 145–149.
8. Cappannini,A., Ray,A. and Mukherjee,S. [2023] MODOMICS : a database of RNA modifications and related information . 2023 update.
9. Grosjean,H., de Crécy-Lagard,V. and Marck,C. [2010] Deciphering synonymous codons in the three domains of life: Co-evolution with specific tRNA modification enzymes. *FEBS Letters*, **584**, 252–264.
10. Gerber,A.P. and Keller,W. [1999] An adenosine deaminase that generates inosine at the wobble position of tRNAs. *Science*, **286**, 1146–1149.
11. Suzuki,T. and Miyauchi,K. [2010] Discovery and characterization of tRNAIle lysidine synthetase (TilS). *FEBS Letters*, **584**, 272–277.
12. Mandal,D., Köhrer,C., Su,D., Russell,S.P., Krivos,K., Castleberry,C.M., Blum,P., Limbach,P.A., Söll,D. and RajBhandary,U.L. [2010] Agmatidine, a modified cytidine in the anticodon of archaeal tRNA Ile, base pairs with adenosine but not with guanosine. *Proceedings of the National Academy of Sciences of the United States of America*, **107**, 2872–2877.
13. Volume 5 Number 11 November 1978 Nucleic Acids Research [1978] *Nucleic Acids Research*, **5**, 4451–4462.
14. Demeshkina,N., Jenner,L., Westhof,E., Yusupov,M. and Yusupova,G. [2013] New structural insights into the decoding mechanism: Translation infidelity via a G·U pair with Watson-Crick geometry. *FEBS Letters*, **587**, 1848–1857.

15. Ananth,P., Goldsmith,G. and Yathindra,N. [2013] An innate twist between Crick's wobble and Watson-Crick base pairs. *Rna*, **19**, 1038–1053.
16. Rogalski,M., Karcher,D. and Bock,R. [2008] Superwobbling facilitates translation with reduced tRNA sets. *Nature Structural and Molecular Biology*, **15**, 192–198.
17. Lei,L. and Burton,Z.F. [2022] "Superwobbling" and tRNA-34 Wobble and tRNA-37 Anticodon Loop Modifications in Evolution and Devolution of the Genetic Code. *Life*, **12**.
18. Wada,M. and Ito,K. [2023] The CGA codon decoding through tRNAArg(ICG) supply governed by Tad2/Tad3 in Saccharomyces cerevisiae. *FEBS Journal*, **290**, 3480–3489.
19. Lozupone,C.A., Knight,R.D. and Landweber,L.F. [2001] The molecular basis of nuclear genetic code change in ciliates. *Current Biology*, **11**, 65–74.
20. Kim,O.T.P., Sakurai,A., Saito,K., Ito,K., Ikehara,K. and Harumoto,T. [2008] Ciliates use both variant and universal genetic codes: Evidence of omnipotent eRF1s in the class Litostomatea. *Gene*, **417**, 51–58.
21. Novoa,E.M., Pavon-Eternod,M., Pan,T. and Ribas De Pouplana,L. [2012] A role for tRNA modifications in genome structure and codon usage. *Cell*, **149**, 202–213.
22. Rafels-Ybern,À., Torres,A.G., Camacho,N., Herencia-Ropero,A., Frigolé,H.R., Wulff,T.F., Raboteg,M., Bordons,A., Grau-Bove,X., Ruiz-Trillo,I., *et al.* [2019] The expansion of inosine at the wobble position of tRNAs, and its role in the evolution of proteomes. *Molecular Biology and Evolution*, **36**, 650–662.
23. Ramos-Morales,E., Bayam,E., Del-Pozo-Rodríguez,J., Salinas-Giegé,T., Marek,M., Tilly,P., Wolff,P., Troesch,E., Ennifar,E., Drouard,L., *et al.* [2021] The structure of the mouse ADAT2/ADAT3 complex reveals the molecular basis for mammalian tRNA wobble adenosine-to-inosine deamination. *Nucleic Acids Research*, **49**, 6529–6548.
24. Wolf,J., Gerber,A.P. and Keller,W. [2002] TadA, an essential tRNA-specific adenosine deaminase from Escherichia coli. *EMBO Journal*, **21**, 3841–3851.
25. Agris,P.F., Eruysal,E.R., Narendran,A., Väre,V.Y.P., Vangaveti,S. and Ranganathan,S. V. [2018] Celebrating wobble decoding: Half a century and still much is new. *RNA Biology*, **15**, 537–553.
26. Ito,K., Uno,M. and Nakamura,Y. [2000] A tripeptide 'anticodon' deciphers stop codons in messenger RNA. *Nature*, **403**, 680–684.
27. Wada,M. and Ito,K. [2019] Misdecoding of rare CGA codon by translation termination factors, eRF1/eRF3, suggests novel class of ribosome rescue pathway in S. cerevisiae. *FEBS Journal*, **286**, 788–802.
28. Gomez,M.A.R. and Ibba,M. [2020] Aminoacyl-tRNA synthetases. *Rna*, **26**, 910–936.
29. Gieg,R. and Eriani,G. [2023] The tRNA identity landscape for aminoacylation and beyond. *Nucleic Acids Research*, **51**, 1528–1570.
30. Abelson,J. [1989] tRNA IDENTITY Jennifer Normanl / and John Abelson. *Identity*.
31. Rubio,M.A.T., Ragone,F.L., Gaston,K.W., Ibba,M. and Alfonzo,J.D. [2006] C to U editing stimulates A to I editing in the anticodon loop of a cytoplasmic threonyl tRNA in Trypanosoma brucei. *Journal of Biological Chemistry*, **281**, 115–120.
32. Bertotti,S., Fleming,I., de los Milagros Cámara,M., Cameán,C.C., Carmona,S.J., Agüero,F., Balouz,V., Zahn,A., Di Noia,J.M., Alfonzo,J.D., *et al.* [2022] Characterization of ADAT2/3 molecules in Trypanosoma cruzi and regulation of mucin gene expression by tRNA editing. *Biochemical Journal*, **479**, 561–580.