

Figure S1: (A) Sequencing spectra of RT-PCR amplicons derived from *O. oeni* tRNA^{Arg}_{ACG} (1), tRNA^{Ser}_{AGA} (2), tRNA^{Leu}_{AAG} (3), and tRNA^{Thr}_{AGU} (4). The anticodon sequence is highlighted, the 'wobble' position is indicated (arrow). Inosine at this position should be detected as a Guanosine, as opposed to the unmodified residue which is detected as Adenosine (tDNA sequences are shown for reference). Note that due to sequence similarities, sequencing of RT-PCR amplicons derived from tRNA^{Leu}_{AAG} are partially masked by amplicons derived from tRNA^{Leu}_{CAA}; and those derived from tRNA^{Thr}_{AGU} are partially masked by amplicons derived from tRNA^{Leu}_{CAA}; and those derived from *T. thermophila* tRNA^{Ala}_{AGC}, tRNA^{Arg}_{ACG}, tRNA^{Ie}_{AAU}, tRNA^{Leu}_{AAG}, tRNA^{Val}_{AAC}. Sequencing spectra using Forward and Reverse primers are shown. Inosine at the anticodon 'wobble' position should be detected as a Guanosine (Thymine in the reversed strand). **(C)** Evaluation of I34 presence/absence on *T. thermophila* tRNA^{Ala}_{AGC}, tRNA^{Thr}_{AGU}, and tRNA^{Pro}_{AGG} by SL-ID.



Figure S2: tRNA quantification boxplot for eukaryotic kingdoms (red), the eukaryotic superphylum *Heterokonta* (green) and different bacterial phyla (blue). The y-axis represents the fraction of all the isoacceptors for TAPS and LIVR amino acids among all the isoacceptors for all the amino acids. The width of each boxplot is proportional to the number of organisms analyzed. 'all' boxplot represents all the data together. Horizontal dashed lines are placed in Q1 = 0.41 and Q4 = 0.68 from 'all' boxplot.



Figure S3: Identification of TadA sequences scheme. NR: non-redundant NCBI database. e_x : BLAST e-value from X query. CDA: cytidine deaminase. pBLAST: protein BLAST. N: number of hits. The scheme is analogous to find the ADAT2 and ADAT3 protein sequences.



Organisms ranked according to TadA e-value

Figure S4: comparison of e-values between the Adenosine deaminase (black circles) and the cytidine deaminases (grey circles). The values are ranked for ADAT2 (a), ADAT3 (b) and TadA (c).

	#Δ		# #						
Consensus Identity		ALLDKY GXXE V F X B C	INVICERCINO		SCHNOREGG (GSN)	S <mark>IH</mark> KDSAIHRG	YPCTGGLLEEA		EN FN
H.sap_Meta_8 A.tha Plan_8	NATRA MVAID NATRA MOAID	Q <mark>ML</mark> EWCRQSSPSP <mark>M</mark> FEH Q <mark>LW</mark> GQWQKD-VAEKFSKO	IV IXMI VERCIMO IV IXMI CERCIMO	aalrimkipi. Salsfigikevym	SCONEREGEOGSME SCENEREGEOGSME	N <mark>I</mark> ASA DLP - N TGR B S <mark>IH</mark> LGS EEAQRGKG	EQCIPGYRAEEA YKCRGGIMAEEA	/EMIRTERKO	E PN
N.gad_Prot_4 F.gra_Fung_8	ng ir a evive Ng ir a emalg	E I L ARGH PPS V F Q E C A LL SY P PK DV DR S I I R E S	CI <mark>VYVT</mark> CERCIMO SI <mark>LYVT</mark> V <mark>EPCVMO</mark>	GALSILKIRRVVI ISILROUGIKKVYII	go (n dregg (gsi e Bav n dregg 10 g v e:	S <mark>VH</mark> GA PH LRHR S <mark>IH</mark> ANS LPVN I ESG	y va da <mark>g e</mark> men ea. Y ei eg g wgr <mark>dea</mark> i	ir lensemar /Gillinety/(ENRO
N.cra_Fung_8 S.pom_Fung_8	NGIR ALIMAIC SGIR ALITAIE	ALLEYSGDAVDRSVISE KILEHYPASVFKE OTKAMLGSRGVVIVFKD	TINNIVERCIMO	NSI IRQIRIKKAY Naalkoihika Ny N Isaikoihika Ny N	a vn dreige 16 gm 1 50 gn dreige ogsvæ	R eh kns efsn vefg S in k Dos I Des S mi h dicven saac	YEVEGGWGRDEA YPVYPGLFYSEA YFSTPGTIRKEA	AT LLER OBTA C AM LLER OBTA C AM LLER OBTA C	ENGR
D.mer Meta 8		ALLASCRER PARQLES EL QVLDWCR LRDCKEVCEQ		AALH ILGVKETIM AALRITRIPFVVM	GOENDREGGRTV V - Sokneregg(GSVI)	DNAAVVGHR DNSSDHLP-HTGTS	IEITGGVRADEA EKCIAGYRAEEA	ALLETEK ALLETEK	E PN
S.pur_Meta_8 P.pat_Plan_8	NATRA ILALE NATRA MVAID	EAMRWC DEK ER EE <mark>V</mark> F SR ' M <mark>V L</mark> SKWQQ LPS TEGF QQO	TK <mark>LEV T</mark> V EPCIMO? CI <mark>LXVI</mark> C EPCIMO?	GALRIMGIKKVVM AALSLSGERKIYM	EGRNEREGE GSUE EGINEREGE GSML	S <mark>VN</mark> S DELP-S IGE B D <mark>TH</mark> S DGCA PQK TKG	FECKAGLYADTA FECVGGFMADEA	AQ DIR F EM KG AN <mark>DIR</mark> G EM E Q	iq n en G n qn
O.luc_Plan_5 A.cas_Prot_8	ng tr al feavd Dat 1 alfvaig	ALLDAHGG-RAACRFEDA KIMDQLRQSRWTELFALS	TIEXMIC DECVMOR SELEVICE EBCIMOR	GANSALGVREMVIN GANLHAGIGKVY	egan dkeggagtmi Egsnereggogsmi	DA H ARGCGAG VGK E R IH DQSF PG LAG	Y DARG <mark>GI</mark> FER EA . Y PCVS <mark>GI</mark> FV <mark>EEA</mark> .	EMERENVE Co lle ofy li	GN PK XGN PN
A.lai_Prot_8 D.dis_Prot_8	NGIR ALLEAFD	YIIVSNYNN TRAILEQIE QIIFINKEINFKISILLER KIIIDDA TRGPPISSIAD	CINYNTY ERCIMOS VINNEY ERCIMOS	VALQ FCK IKRVE VALQFCK IKRVE VALUNR DAHVER	SCHINEKEGGNGSMA SCHINEKEGGNGSMA SCHNEKEGGNGSMA	S LH EPTANIPRI E IN FSPISNGRE A LH ASPROOEWWP3	ESCHPGILRIEA YNCITGLIKNQA WYSEGGHTEAFA		ENKK ENKK
P.fal Prot 8 P.tet Prot 8	NGSR SELIAID	KY DY DGNY EERINN LKKO RMNRDLED		/YANKINGIKONY SQANYWATHSMYMO	G NERECCOSMI Schins Reconstruct	S IH KKYEDINVH S <mark>IN</mark> K	Y -IEH PECS <mark>EKS</mark> YPSFG G HLEY DO	IN IMPARYKA IK I LO I FYRE	IGN PS
T.bru Prot 8 T.the Prot 8	HALANAN FVAVE NATIONIN	E <mark>LL</mark> RQA TAGAVSQ DLA D I E <mark>L L</mark> DKY GQN DEQQ <mark>L</mark> F KQI	IN LYN VV DECIMON I 1 LEV 19 DECIMON	AMLLYNRVRKVY 10 Y Sls imgikevy 10	SC TN FREGGNGTML SCYNEKEGGNGTIL	S <mark>VH</mark> NSYKGCAALIG D <mark>IN</mark> TANYGAN T	YESCGGYRA EEA YPSYGGFLE EE CI	MELOPYRE Kovincever	EN TN GN IN
E.his_Prot_5 G.lam_Prot_6	DG TOLALIM CIN NSTV ALIV ALA	Q LV EKHVHLS R Q L PSGINLSNO	CI IMMICERCIMO CV IMMI IERCIMO	EALKQCGITKIIM AALSIMGLTNIIM	EGSNAR EGGOGSVM Farink Eggogsvi	TMIKSSLYKN-HEI D <mark>VN</mark> NHIES PWTK	NYCYHS MELEG - INAKYTHIERA	LE MIR CE ESM NY MIR ESPER	IEN EK
T.vag_Prot_7	DATR ALLY TFK	elloghkkon lhe Hekesrs Dykqe lkes Amlegdoy taatlre		VAR SII FIQAVI ISA KMMGI FRVV VAR AIVOVKRVVI	SC IN DR FIGG (GSM) SC IN DR FIGG (GSM) SC IN DR FIGG (GSM)	N TH THEIM-PS LPS S IH DPTAF PDHHIG	Y DVISGVLGEEA. Y DVISGVLGEEA.		E P PK
A.inv_Hete_5 B.hom Hete 7	NATK ALIVALE NATI LA MVAL S	A <mark>T L</mark> EEEA Y TSS TF R <mark>B</mark> C R T <mark>L</mark> 3AN GN DPT <mark>T</mark> F EGS	CT <mark>IMVI</mark> CERCIMO ST <mark>IMVI</mark> I ERCIMO	AALAILHVKRVV SALAQIGVSRVV	ECHN DREECNEST L Basn dreec ocsmil	S <mark>IL</mark> Q DANAF PS HH TG S ILH DQN EF SA GHKH	Y PI TH G L MKN EA YEV TRHIQKERS	DA LIRIEY DR De lingrey Nr	GN FR
B.sp. Hete 7 F.cyl_Hete_8	NATI A IVAIS DATRIS IVAID	EALHSNGNDPSLF LQS R <mark>BL</mark> INKQQQRSKD <mark>I</mark> FRHC	ST LYVT I E POLMOF CI <mark>LFVT</mark> O <mark>DECIMO</mark> F	NSALAQEGIKRAV VAALAMVGIRRVVII	BASINDREGGOGSME Box(NDREGGOGSTE	S <mark>IN</mark> LEK GKVV H LH NN DNESS YN SG	YEVTR GV FQEKT YKIKS <mark>GV</mark> LEKEA	CH <mark>LLLOCY</mark> CR CT <mark>LLL</mark> S <mark>SY</mark> CR	ENQY EN LH
P.par Hete 3 P.hal_Hete_7	NAT ALIVAIE	ADAAKY GD-KAGEV LADO STAAKY SD-QASEV LADO ADD DAGHSSEV SDIV		GALAHMFIKRVY GALAIISIKHVY MALAINSKHVY	SCHNER FRANCSSME SCHNER FRANCSSME	N LH ERSALEKAHRG	FPCVSGILKDEA FPCVSGILKDEA Net TTTSVMKDEA	ITLIKKEMIN SILENEMIS	ENPR
T.oce_Hete_7	DATRISTIVATE	R <mark>LL</mark> TSGRSSY EKE <mark>V</mark> LRKC		VAALSINGINRVE	SCRIPTICS CSSII	THE EPNALPDKHLG	YPIHGGV LEEDA		ENFH

Figure S5: Multiple sequence alignments of ADAT2 deaminase domains. Residues with 100%, 80-100%, 60-80% and <60% of similarity are respectively framed with green, olive, orange and white. The deaminase domains are overlined, where the residues that coordinates with Zn (#) and the glutamic acid that allows the nucleophilic attack (Δ) are depicted. The organisms names are displayed as *X.yyy_zzzz_n* where *X* is the first letter of the Genre, *yyy* the first 3 letters of the specie, *zzzz* the 4 first letters of the phylum and *n* the number of different A34 tRNAs.



Figure S6: (A) ADAT stretch distribution. Median values of ADAT-sensitive codon enrichment as a function of ADAT amino acid stretch length, plotted for different groups based on A34-tRNA diversity values (color-coded). Linear regressions (colored dashed lines) were calculated for all groups with more than 5 data points. Numbers on top of each data point refer to the number of samples for each data point. Data points with less than 15 samples were omitted. $\Delta_{1\%}$ represents the number of codons needed to increase ADAT enrichment by 1% based on the slope of the linear model. (B) Graph as in (A) showing the full ADAT stretch size distribution.

Α

Name	Sequence	Details
oTFW_94	gaatgtcataagcgCCAAGCGAGCGCTCTACCATTTG	Bridge oligo T. the/H. sapiens tRNA AlaAGC - Splinted Ligation
oAGT_230	gaatgtcataagcgCCAGGCGAATGCTCTAACCACTG	Bridge oligo T. the tRNA ThrAGT - Splinted Ligation
oAGT_231	gaatgtcataagcgCCAAACGAGAATCATGCCACTAG	Bridge oligo T. the tRNA ProAGG - Splinted Ligation
oAGT_312	gaatgtcataagcgGCAGTCAGACGCTCTATCCAATTG	Bridge oligo O. oeni tRNA ArgACG - Splinted Ligation
oAGT_313	gaatgtcataagcgTCAATCCTGCGCGTCTGCCAATTC	Bridge oligo O. oeni tRNA LeuAAG - Splinted Ligation
oAGT_314	gaatgtcataagcgCCAGACCGACCCCTTCAGCCAC	Bridge oligo O. oeni tRNA SerAGA - Splinted Ligation
oAGT_315	gaatgtcataagcgCCAGTGAAGTGCTCTAGCCAAC	Bridge oligo O. oeni tRNA ThrAGT - Splinted Ligation
oTFW-31	AGCTTAATACGACTCACTATAGGGGATCTAGCTCA	FWD primer T. thermophila PCR Ala
oTFW-36	GATCCACATGTTGGTGGAGAACCTGGGCATT	RVR primer T. thermophila PCR Ala
oTFW-37	AGCTTAATACGACTCACTATAGGGGTGATGG	FWD primer T. thermophila PCR Arg
oTFW-42	GATCCCCTGGCGAGATGAGCAGGACTCGAAC	RVR primer T. thermophila PCR Arg
oTFW-43	AGCTTAATACGACTCACTATAGCTCGGGTAGCTCAG	FWD primer T. thermophila PCR IIe
oTFW-48	GATCCCCTGGTGCTCCGGGAGGGGGCTTGAAC	RVR primer T. thermophila PCR Ile
oTFW-49	AGCTTAATACGACTCACTATAGATGAAGTGGCCGAG	FWD primer T. thermophila PCR Leu
oTFW-54	GATCCCCTGGTGATGAAGGCGAGATTCGAACT	RVR primer T. thermophila PCR Leu
oTFW-55	AGCTTAATACGACTCACTATAGGGTTGTTGGTC	FWD primer T. thermophila PCR Pro
oTFW-60	GATCCCCTGGGGGGGTCGTCCGAGAATCGA	RVR primer T. thermophila PCR Pro
oTFW-61	AGCTTAATACGACTCACTATAGACAATTTGTCCGAG	FWD primer T. thermophila PCR Ser
oTFW-66	GATCCCCTGGCGACAACTGCAGGATTCGA	RVR primer T. thermophila PCR Ser
oTFW-67	AGCTTAATACGACTCACTATAGCCGCTTTAGCTC	FWD primer T. thermophila PCR Thr
oTFW-72	GATCCCCTGGAGCCACTTGGCGGGATTG	RVR primer T. thermophila PCR Thr
oTFW-73	AGCTTAATACGACTCACTATAGATTCCTTAGTG	FWD primer T. thermophila PCR Val
oTFW-78	GATCCCCTGGTGATTCCTCCGAGGTTTGA	RVR primer T. thermophila PCR Val
		FWD primer tRNA Arg ACG Oenococcus oeni with M13-
oAGT_296	CAGGAAACAGCTATGACCCACCATTAGCGCAATTGG	RP adaptor for sequencing
oAGT 297	GCACCATGTAGGAGTCGAAC	RVR primer tRNA Arg ACG Oenococcus oeni
_		FWD primer tRNA Leu AAG Oenococcus oeni with M13-
oAGT_298	CAGGAAACAGCTATGACCGACGTGGCGGAATTGGCAG	RP adaptor for sequencing
oAGT_299	GGCGATGGGAGTCGAACCCATAC	RVR primer tRNA Leu AAG Oenococcus oeni
_		FWD primer tRNA Ser AGA Oenococcus oeni with M13-
oAGT_300	CAGGAAACAGCTATGACCGATGGATACCCAAGTGGC	RP adaptor for sequencing
oAGT_301	GAGAGATTCGAACTCTCG	RVR primer tRNA Ser AGA Oenococcus oeni
		FWD primer tRNA Thr AGT Oenococcus oeni with Sp6
oAGT_302	ATTTAGGTGACACTATAGAATAGCTCAGTTGGCTAGAGCAC	adaptor for sequencing
oAGT_303	TGCCGACTAGAGGATTCG	RVR primer tRNA Thr AGT Oenococcus oeni

 Table S2: The Universal Genetic Code. ADAT codons are colored in blue (dark and soft). ADAT-sensitive codons are colored in dark blue. For each codon, the amino acid that decodes is depicted in one-letter and three-letter convention.

 Second Coder Letter

		U		С		Α		G			
		F	Phe	S	Ser	Y	Tyr	С	Cys	U]
First Codon Letter	U	F	Phe	S	Ser	Y	Tyr	С	Cys	С	1
		L	Leu	S	Ser	Stop		Stop		Α	
		L	Leu	S	Ser	Stop		W	Trp	G	
	С	L	Leu	Р	Pro	Н	His	R	Arg	U	Thi
		L	Leu	Р	Pro	н	His	R	Arg	С	
		L	Leu	Р	Pro	Q	Gln	R	Arg	Α	d
		L	Leu	Р	Pro	Q	Gln	R	Arg	G	öd
	A	1	lle	Т	Thr	Ν	Asn	S	Ser	U	ß
		1	lle	Т	Thr	Ν	Asn	S	Ser	С	Et
		I.	lle	Т	Thr	К	Lys	R	Arg	Α	ter
		Μ	Met	Т	Thr	К	Lys	R	Arg	G	
	G	V	Val	А	Ala	D	Asp	G	Gly	U	
		V	Val	А	Ala	D	Asp	G	Gly	С	
		V	Val	А	Ala	E	Glu	G	Gly	Α	
		V	Val	А	Ala	E	Glu	G	Gly	G	

Second Codon Letter