

SUPPLEMENTAL RESULTS AND DISCUSSION

The tRNA variety of *L. texanum* and *P. pacificus* is identical

The loss of introns from the *L. texanum* tRNA genes does not seem to correlate with other changes in the tRNA gene pool (Figure S1). We detected 48 different tRNA types in the HiFi *L. texanum* assembly and in the *P. pacificus* genome (<https://gtrnadb.ucsc.edu/genomes/eukaryota/Ppaci4/>) (1, 2). The same 48 tRNAs were present in the Illumina *L. texanum* assembly, although there were slight variations in the number of genes for some tRNA types.

The initial tRNAscan-SE analysis detected 46 tRNA types. The 47th type is an unusual Leu tRNA. One striking feature of the *C. elegans* tRNA pool is that it has two different tRNAs with UAU anticodons: it has a typical Ile tRNA with a UAU anticodon, but it also has an unusual Leu tRNA with a UAU anticodon (1, 2). This tL(UAU) tRNA (highlighted with a red outline in Supplemental Figure S1) is classified as tI(UAU) by tRNAscan-SE, but the two tRNAs are readily distinguished by their sequence as well as the longer variable arm and overall length of the Leu tRNA. Both *P. pacificus* and *L. texanum* (Supplemental Figure S1) also had these two different tRNAs.

The 48th type comes from separating two different Met tRNAs. The tRNAscan-SE output groups initiator Met tRNA and elongator Met tRNAs as one isodecoder even though they have distinct functions. We curated the Met tRNAs based on sequence similarity to *C. elegans* and *P. pacificus* tRNAs.

Overall, this analysis revealed a typical repertoire of tRNA genes for *L. texanum* (Supplemental Figure S1). The only difference in tRNA types that we detected is that *L. texanum* and *P. pacificus* have three different glycine tRNAs, while *C. elegans* only has two (Supplemental Figure S1). This indicates that in *C. elegans*, tG(UCC) decodes both GGA and GGG codons, while in *P. pacificus* and *L. texanum*, tG(CCC) may contribute to the decoding of the GGG codons. The third Glycine tRNA is highlighted with a green outline in Supplemental Figure S1. We do not think that this change is related to the loss of introns.

The total number of tRNA genes per eukaryotic genome is highly variable, and consistent with this, the HiFi *L. texanum* genome contains approximately three times as many tRNA genes as *C. elegans* and about 50% more than *P. pacificus* (Supplemental Figure S1). Strikingly, *L. texanum* has a drastically increased number of tC(GCA) genes (highlighted with an orange outline in Supplemental figure S1). The difference is explained by *L. texanum* contig 52, which contains 413 copies of tC(GCA) in a 84,056 bp stretch. These 413 tandem repeats of a single tRNA almost completely account for the 556 increase in *L. texanum* tRNA gene number compared to *P. pacificus*. The sequences of the tC(GCA) genes in the repeat were the same as the genes dispersed outside of the repeat. One notable difference between the previous Illumina based *L. texanum* genome assembly and our HiFi assembly is that the tC(GCA) repeat was missing from the Illumina assembly because of the difficulty in assembling such highly repeated regions from small reads. However, the *P. pacificus* assembly is also a high quality PacBio HiFi assembly and thus the difference in tC(GCA) gene number is likely a biological difference and not an assembly quality difference. We do not think that this change in tC(GCA) gene number is related to the loss of introns. Therefore, we conclude that the loss of tRNA introns has no obvious connection to the tRNA pool.

SUPPLEMENTAL FIGURES LEGENDS

Supplemental Figure S1: tRNA genes in the *L. texanum* genome compared to those in *P. pacificus* and *C. elegans*. Indicated is the number of genes for each of 48 different tRNAs detected in the HiFi assembly of the *L. texanum* genome (left) *P. pacificus* (middle) and *C. elegans* (right). tRNA genes that contain introns are highlighted in blue and those that have lost introns are in yellow (as in Figure 1). Red outline highlights an unusual leucine tRNA with a UAU anticodon that was previously described in *C. elegans* and *P. pacificus*. Green outline highlights tG(CCC) genes present in *L. texanum* and *D. magnus* but not in *C. elegans*. Orange outline highlights the dramatically increased copy number for tC(GCA) in *L. texanum*. Purple outline highlights ANN/GNN anticodon pairs where the GNN anticodon is commonly missing from eukaryotic tRNAs and the ANN anticodon tRNA is modified to INN by ADAT-2/ADTA-3. *L. texanum* has the typical 8 INN anticodon tRNAs, which does not explain the triplication of *adat-3*.

Supplemental Figure S2: TSEN-2, TSEN-34 and TSEN-54 are conserved in *P. pacificus* and *D. magnus*. Sequences for TSEN-2 (A), TSEN-34 (B) and TSEN-54 (C) were identified by BLAST and aligned with MUSCLE. Purple shading indicates identity to the consensus sequence. The TSEN subunits show low sequence similarity across eukaryotes but the catalytic residues (indicated with asterisks) are conserved.

Supplemental Figure S3: Sequence of tL(CAA), tY(GUA) and tI(UAU) tRNAs in *C. elegans*, *P. pacificus*, *D. magnus* and *L. texanum*. Aligned tRNA gene sequences identified by tRNAscanSE and with the introns removed are depicted as sequence logos. (A) The tL(CAA) sequence is highly conserved between *C. elegans*, *P. pacificus*, and *D. magnus* but more diverged in *L. texanum* as also depicted in Figure 3. The base pair between positions 31 and 39 is indicated with a double headed arrow. (B) The tY(GUA) sequence is highly conserved between the four species. (C) The tI(UAU) sequence is highly conserved between *C. elegans*, *P. pacificus*, and *D. magnus* but more variable in *L. texanum*. The divergence in *L. texanum* appears unconnected to the loss of the intron because intron loss is shared with *P. pacificus* and *D. magnus*. (D) The variability in tI(UAU) sequence in *L. texanum* reflects the presence of two gene families. It is unknown whether both families are functional.

Supplemental Figure S4: Sequence conservation of intron-dependent tRNA modification enzymes PUS7 (A), PUS1 (B), NSUN2 (C), TRMT61 (D), and TRMT6(E). Sequences were identified by BLAST and aligned with MUSCLE. Purple shading indicates identity to the consensus sequence. Blue and yellow shading reflects intron presence in the indicated tRNA substrates.

Supplemental Figure S5: Sequence conservation of *xbp-1* splicing enzymes IRE-1, RCTB-1, and Archease. Sequences for IRE-1 (A), RCTB-1 (B) and Archease (C) were identified by BLAST and aligned with MUSCLE. Purple shading indicates identity to the consensus sequence.

Levopalatum texanum tRNAs										
Four Box tRNA sets										Total
Ala	AGC	GGC	CGC	UGC						76
	40	0	16	20						
Gly	ACC	GCC	CCC	UCC						94
	0	43	3	48						
Pro	AGG	GGG	CGG	UGG						68
	34	0	14	20						
Thr	AGU	GGU	CGU	UGU						60
	27	0	20	13						
Val	AAC	GAC	CAC	UAC						61
	31	0	21	9						
Six Box tRNA Sets										
Ser	AGA	GGA	CGA	UGA	ACU	GCU				79
	33	0	17	8	0	21				
Arg	ACG	GCG	CCG	UCG	CCU	UCU				85
	38	0	4	21	12	10				
Leu	AAG	GAG	CAG	UAG	CAA	UAA	UAU*			109
	34	0	24	6	10	7	28			
Two Box tRNA Sets										
Phe	AAA	GAA								34
	0	34								
Asn	AUU	GUU								39
	0	39								
Lys			CUU	UUU						86
			58	28						
Asp	AUC	GUC								57
	0	57								
Glu			CUC	UUC						82
			51	31						
His	AUG	GUG								35
	0	35								
Gln			CUG	UUG						37
			26	11						
Other Box tRNA Sets										
Ile	AAU	GAU		UAU*						53
	38	0		15						
iMet/ Met			CAU							46
			24/22							
Tyr	AUA	GUA								28
	0	28								
Sup		UCA	CUA	UUA						0
		0	0	0						
Cys	ACA	GCA								436
	0	436								
Trp			CCA							21
			21							
SeiCys				UCA						1
				1						
Total										1587

Pristionchus pacificus tRNAs										
Four Box tRNA sets										Total
Ala	AGC	GGC	CGC	UGC						73
	47	0	12	14						
Gly	ACC	GCC	CCC	UCC						80
	0	35	1	44						
Pro	AGG	GGG	CGG	UGG						61
	35	0	12	14						
Thr	AGU	GGU	CGU	UGU						61
	34	0	16	11						
Val	AAC	GAC	CAC	UAC						51
	27	0	16	8						
Six Box tRNA Sets										
Ser	AGA	GGA	CGA	UGA	ACU	GCU				56
	23	0	13	9	0	11				
Arg	ACG	GCG	CCG	UCG	CCU	UCU				68
	38	0	1	12	5	12				
Leu	AAG	GAG	CAG	UAG	CAA	UAA	UAU*			76
	34	0	17	5	8	3	9			
Two Box tRNA Sets										
Phe	AAA	GAA								21
	0	21								
Asn	AUU	GUU								33
	0	33								
Lys			CUU	UUU						84
			58	26						
Asp	AUC	GUC								47
	0	47								
Glu			CUC	UUC						81
			50	31						
His	AUG	GUG								37
	0	37								
Gln			CUG	UUG						39
			26	13						
Other Box tRNA Sets										
Ile	AAU	GAU		UAU*						40
	32	0		8						
iMet/ Met			CAU							36
			17/19							
Tyr	AUA	GUA								29
	0	29								
Sup		UCA	CUA	UUA						0
		0	0	0						
Cys	ACA	GCA								34
	0	34								
Trp			CCA							22
			22							
SeiCys				UCA						2
				2						
Total										1031

Caenorhabditis elegans tRNAs										
Four Box tRNA sets										Total
Ala	AGC	GGC	CGC	UGC						32
	20	0	4	8						
Gly	ACC	GCC	CCC	UCC						49
	0	14	0	35						
Pro	AGG	GGG	CGG	UGG						40
	6	0	4	30						
Thr	AGU	GGU	CGU	UGU						32
	16	0	7	9						
Val	AAC	GAC	CAC	UAC						30
	19	0	6	5						
Six Box tRNA Sets										
Ser	AGA	GGA	CGA	UGA	ACU	GCU				39
	15	0	6	9	0	9				
Arg	ACG	GCG	CCG	UCG	CCU	UCU				38
	19	0	1	8	2	8				
Leu	AAG	GAG	CAG	UAG	CAA	UAA	UAU*			40
	18	0	5	3	7	3	4			
Two Box tRNA Sets										
Phe	AAA	GAA								14
	0	14								
Asn	AUU	GUU								19
	0	19								
Lys			CUU	UUU						43
			29	14						
Asp	AUC	GUC								27
	0	27								
Glu			CUC	UUC						41
			24	17						
His	AUG	GUG								16
	0	16								
Gln			CUG	UUG						26
			7	19						
Other Box tRNA Sets										
Ile	AAU	GAU		UAU*						24
	21	0		3						
iMet/ Met			CAU							18
			9/9							
Tyr	AUA	GUA								18
	0	18								
Sup		UCA	CUA	UUA						0
		0	0	0						
Cys	ACA	GCA								13
	0	13								
Trp			CCA							12
			12							
SeiCys				UCA						1
				1						
Total										572

■ intron-containing tRNA genes

■ tRNA genes that lost intron

□ Leucine tRNA genes with a UAU anticodon

■ Glycine tRNA absent from *C. elegans*

■ Cysteine tRNA with increased copy number in *L. texanum*

■ ANN anticodon tRNAs require INN modifications by ADAT-2/3 to compensate for GNN anticodon absence

A

```

HsTSEN2 1 MAEAVFHAPKRRRVYETYESPLPIFG-QDHGPLEKFIFRAEMINNNVIVRNAEDIEQLYGKGYFGKILSRSRPSFTISDPKLVAKWKDMKTMPII
CeTSEN-2 1 --MDLLVFDKRNNGHRHVAGRVPVALNKEDELEVLGFRRIDAFINGGLVEIFDETASRIIHEMSGIGOWL-----KDE-----
PpTSEN-2 1 --MDVLKIERKKRGKMTVL-----DWDNE-GWTETEIVLRAGEYTPSTSHADILRWKAGYGEFP-----
DmTSEN-2 1 --MEV-----RHKRYSPRSI-----VWSQEVVEEGETEVIIRGGEFIVPSCFEFSAFLYEELHYGHHL-----WEE-----

HsTSEN2 100 TSKRYQHSVEWAAEIMLRQGDDESTVRRILKDYTKPLEHPPVKRNEEAQVHDKLNSGMVSNMEGTAGGERPSVNVNGDSKSGGVGDPREPLGCLQEGSGC
CeTSEN-2 72 -----RKESSRGVSRSCFSQS-----
PpTSEN-2 57 -----NEERRDRVNASLIFNG-----
DmTSEN-2 57 -----REEFQGRLEELFAT-----

HsTSEN2 200 HPTTESFEKSVREDASPLPHVCCCKQDALILQRLHHEDGSGHIGLLHPGDRGPDHEYVLVEEAEACAMSEREAAPNEELVQRNRLICRRNPYRIFEYLQL
CeTSEN-2 87 -----CGEPNGKLINEQAANEWAELSPIART-----LVL
PpTSEN-2 72 -----PEEK-----KQWSEFISSDNV-----LRL
DmTSEN-2 72 -----IDER-----TTWACKSLVDRR-----LRL

HsTSEN2 300 SLEEAFFLVYALGCLSIYKEP-LITIVLKWAFTVV---QPTFRTTYMAHYHFRSGKWVPKVGLKYGTDLILYRKGPFFYHSSVILVDDHFEGLS
CeTSEN-2 116 TPIETALLSIDWKIINVVKNGKISLDEEITWEMRNLYDNPKFEGKVSFVRFLKMGNGWIVTSGHTFGCDYLIYCLGAEFYHSSAGVLIADV
PpTSEN-2 92 STEEVYMLV-EERERAYLKRDGVR-LSIEETFHLSNW---GKRFFDSYATRYLRRAGWTPRCGISFGCDYVIYDGLPSEVHASAGVSIERE-----
DmTSEN-2 92 SAEVEFYLTIEDHRVMTDDGSI-LSPSNLWLSLSSL---SSPHFYRVRVYFLRHSGWIVRPGLPFGCDYLIYGGSPDETHASAGIVVGEK-----

HsTSEN2 396 RRPLSWKSLAALSRSVSVNSKELMLCYLIKPMSTMDKEMESPECMKRIKVQEVILSRWVSSRERSDQDDL-----
CeTSEN-2 207 ---IDPRRLTLTRILSHNKKALIV-----ATTSEVECFEDANAVDVKISTMKTYFLERDVAQMSNRNCNEVEYELNTV
PpTSEN-2 178 -RPLDNVEMEALSRSLWHVKKHVII-----LSITADSHEYGSVEGADVKTATLAPTTLNSLIG-----
DmTSEN-2 180 ---DEMHLDGLVRLTGTTKKSLIV-----TKVKSEGEEMDDVEKATVQTRIMGRVFITMTK-----

```

B

```

HsTSEN34 1 -----MLVVEVANGRSVLWGAEAVQAL-RERLGVGGRITVGALPRGPRQNSRLGLPLLLMPEEARLLAEIG
CeTSEN-34 1 MSHEILEV-----PVNSEKR-----KVEICHVTGRFFA-SKSDCPYLESAAHVCLNRRNSAEDDKNWHYNSNL-LP--ILDELVAIILVEFG
PpTSEN-34 1 MSVDSATN-----SVTCA-----PPRVHFNVSFVVFEEKADADLLSNCHRILPDFGSSLEQLGPNSSNSCYSC--LLPEQVAVALDNG
DmTSEN-34 1 MTGESAQDNGNEAAVEDCGPGPSVSCSVRFDYSTKIPIDVHLVNEFTFLFNNEVSVLLKEHRMVADRPCLLE--GLQSSSSCYAH--LMPEQVAVLVQH

HsTSEN34 65 AVTLVSAPRPSDRHSLALT--SFKRQOEESFOEQSALAAEARETRRQELLEKITEGQAACKQKLEQASGASSSQEAGSSQAAKEDETS DGOASGEQEEA
CeTSEN-34 77 FANVIRLE-----TTPDR-----
PpTSEN-34 77 VITIVRLKRPVTIGND-----EVPNEVEKPEIESLVLRKIAVGRKMKNLKRKAQEEGTVIQIKLRL-----DDVMVTAELN
DmTSEN-34 97 IARVQLKEDLTATEELFTPVPTTSTQEESEPNEEETIATKMAMVRKRTV---DRNGDGVPLSKARRSVA-----KDVSVSEEEIS

HsTSEN34 163 GPSSSQAGPSNGVAPLPRSALLVQLATARP RPVKARPLDWRVQSKDWPHAGRAPAHELRYSIYRDLWERCFFLSAAGKFGDGLVYPGDPLRFHAHYIAQC
CeTSEN-34 89 -----EIQRKLTLEISEI-----PFE-----NSKEF-----RTKRLVVRDFWKGYFIADGTRFGNGVLYVTSSPNTCHAEFVLIC
PpTSEN-34 150 SALAEISVPLQARNNAVQIRTNVSSKL-----AYESITFPEEFDLDF-----RIRKIVFRDLWSRGFYVTSGVRFGCHFVLYKNPPGSHVADFMVRC
DmTSEN-34 176 EAIABELGRGRQRENEIQRQRQIVLRLDLPENVEHITNTFSLNDAEY-----RRREIVLHDLWRRGYYITCGIRYGCQFMILSAPPGVSHAEFLVIC

HsTSEN34 263 WAPEDTIPLQDLVAAGRLGTSVRKTLTLLCSFPQPDG-KVYVTSQWASLQ-----
CeTSEN-34 154 -----APITDEERISAMRCCNQVKKTLTLLATTSSESTQPHYTQCTWFRPEFY-----
PpTSEN-34 237 VDAAEDQTSSTMSLFSARTANQVCKKSLIAIVS-GSIHPQYLTIQSFTPLSQDLD-SID-----
DmTSEN-34 268 IDADEGFSPRRLQTATRTAQVQRKILILAIVSHDSRQPFVLAERFRPWRTRTEDKSVDDLTFEFPNSI

```

C

```

HsTSEN54 1 MEPEPEPAAVEVPAGRVLSARELFAARSRSQKLPQRSHGPKDFLPDGSAAQAEERLRRCREELWQLLAEQRVERLGSLSVAEWRPEEGFVELKSPAGKFWQ
CeTSEN-54 1 -----MAEWNRAKR-----KRKTIEIQYTGLASNSTLKVTKVTGKHLE
PpTSEN-54 1 -----MDESIDDAGPFTVAKVDDRI RL-----ENSLTLITYHEECQNSRFVDRFRSHQS
DmTSEN-54 1 -----MTEGQSSSTVEVISQMLTQDEGRVRM-----KDTLTSITVSTRGP IAFVRVKKFVSGHER

HsTSEN54 101 TMGFSEQG---RQRLHPEEALYILE-CGSIHLFHQDLPLSLIQEAYQLLTDHTVTFLQVQVFSHLKRLGYVVRRFQ--PSSVLSPYERQNLNDASVQHLE
CeTSEN-54 39 SMGIPGKF---GYEVYPEEAVYLMETGSATISTSSGRDLSLEAF-SILEAMSVPMKYEAHYKVKRTGLVVLRRP---KPTIDFERLRHVEPKIRQKF
PpTSEN-54 52 TMGEPQSK-GAGHILFPEEATYLVIEITQGVIRTPEGRYMTLEELF-MVLGDFSIPQOVYSAACKRAGFVIRPGILEHSPASP-
DmTSEN-54 55 ILGHKEGGIGKWNHLHPEEAI FLVESQCGMAVNEGQRPLCLEELL-NFAGRGGLSMHAYAVYAAAKRAGFVCRPS--DCRPVQPVKKESGDG-----

HsTSEN54 195 DGDGKRKRSSSSPRSKKAKALDNSLQPKSLAASSPPPCSQSCPEEKPOESSPMKGGPGGFQLLGLSLGSPGPGAREGVGCSWES-GRAENGVTGAGK
CeTSEN-54 131 AEKS-----FEMLVNRHTVPEERVPN--PDSFSPFENRNGQISMKMVLTHT-SPALVNF----LPLEMLESIQTSTILQNL
PpTSEN-54 135 EGAV-----PPAVDCALPPFRVRPRVILDMFNMGSDTN-----VLRTK-KPELFPE----GCRLOK-GNEKEKSVEEKN
DmTSEN-54 145 GGDG-----RKEKSLSLPFRPRPS-CFSSLESMDKPYG-----ILLSSIHSLNLPN----GCESSITGAVTTGSGVDEK

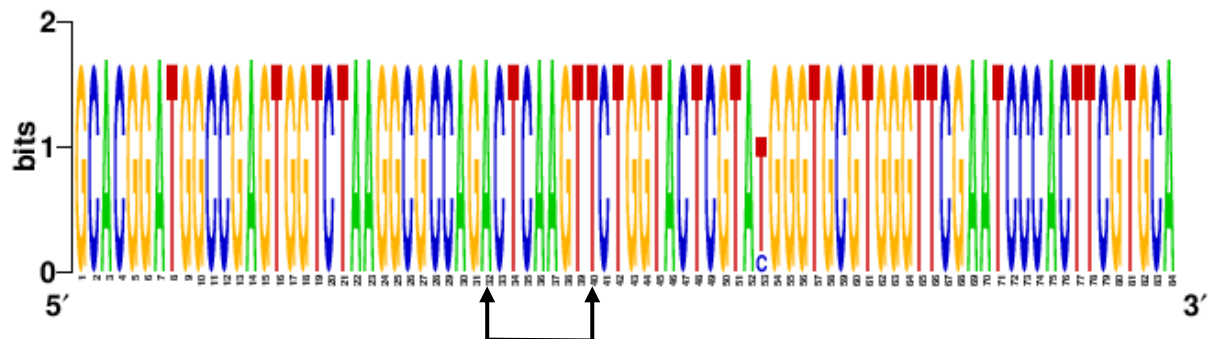
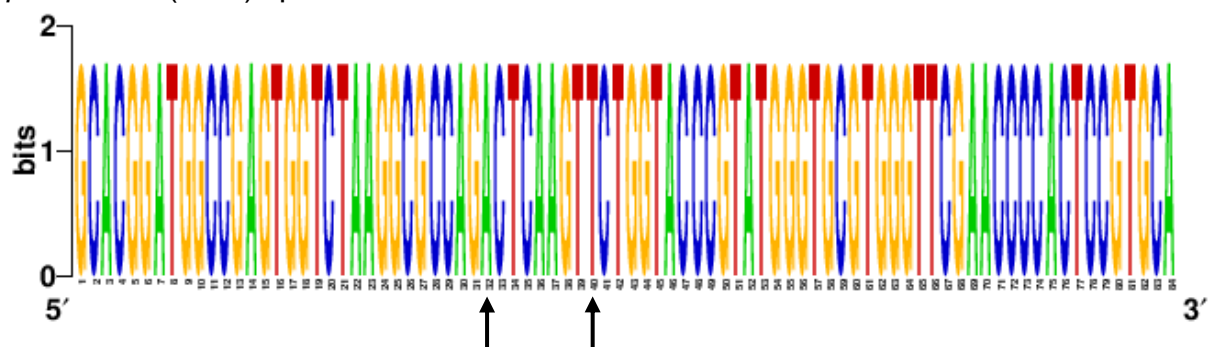
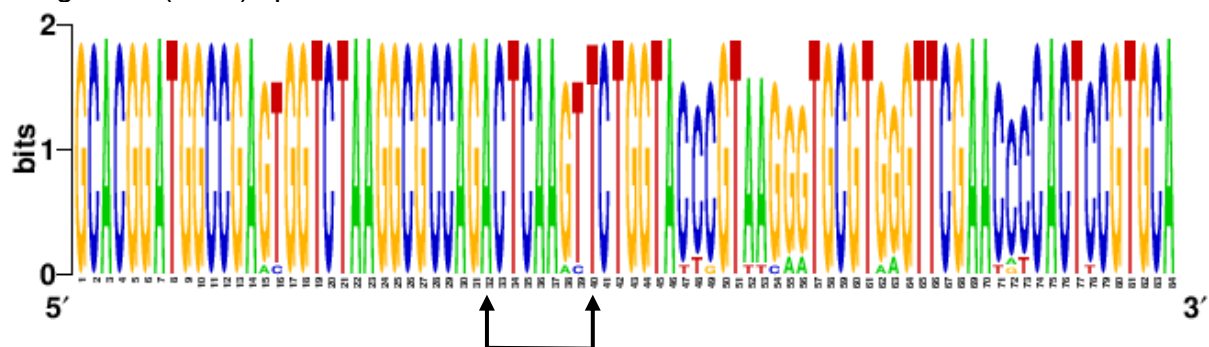
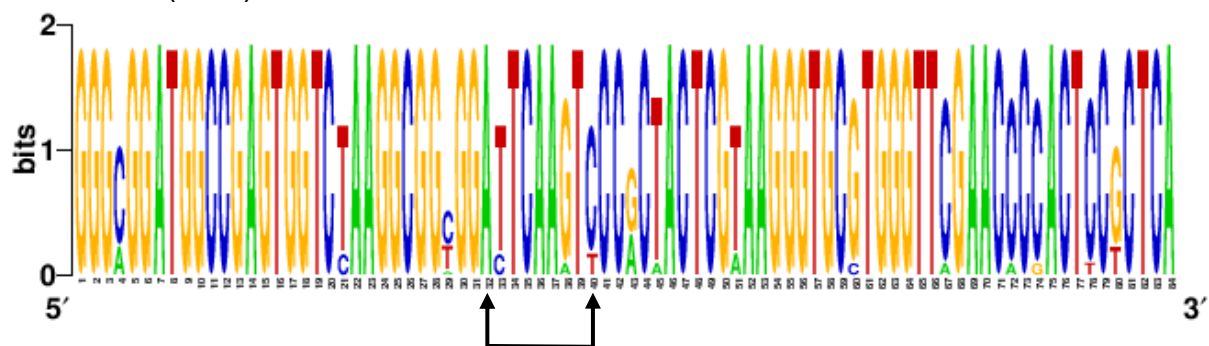
HsTSEN54 294 RRWNFEQISFPNMASDSRHTLLRAPAPPELLPANVAGRETDAESWCQKLNQRKEKLSRREREHHAEEAQFQEDVNADPEVQRCSSWREYKELLQRRQVQRS
CeTSEN-54 200 KPP-----TPNRPPCRPPSYWPFVDH-----HVTNWQEFRF-----QAEQA
PpTSEN-54 199 VAI-----VG-----EQGGK-----KRRDPPSIRPRAWPSLAY-----ASKNAGDWKEYAE-----
DmTSEN-54 210 KIW-----AKSTAEPLPETSQKGPQKGGK-----KKEWKGSLRPHWPPLRN-----ATKGVKTKWKEFSE-----

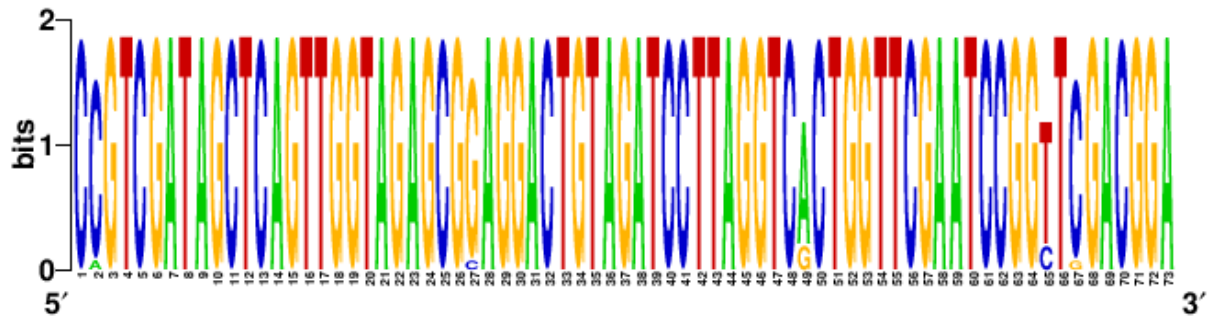
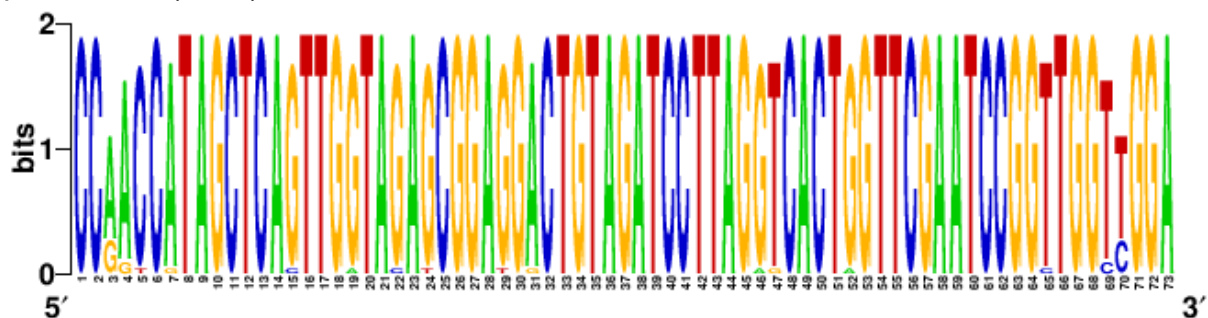
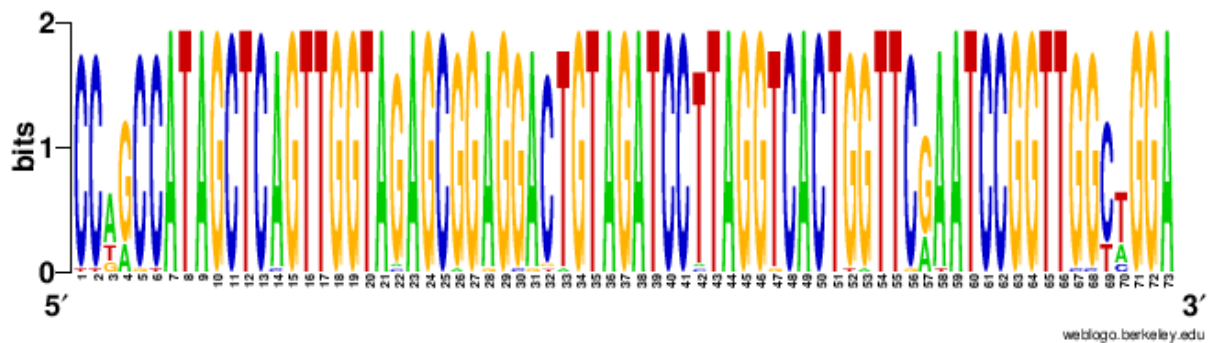
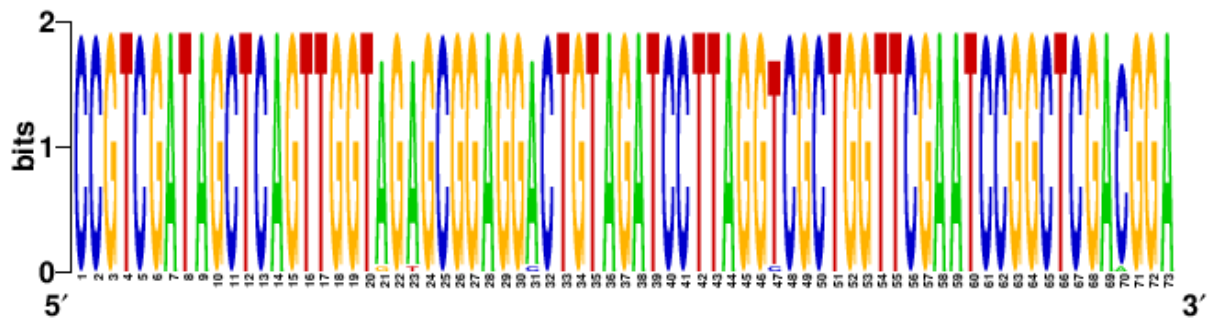
HsTSEN54 394 QRRAPHLWGQVPTPLSPGQASSPAVVLQHISVLQTTHLPDGGARLLEKSGGLEIIFDVYQADAVATFRKNNPG-KPYARMCISGF-DE-----P-----
CeTSEN-54 236 RQETVLR---LKKLRPPSHAE LRERITVENVD-----FHVFSPASFSH---RLPA-RPLFSLICYSATG-----P-----
PpTSEN-54 241 LRRNLLGK-----SRYSSPLDRTH-----FLLYSAKEYSHAARTGEIRPIARVAVDFYSGRACPSHP-----
DmTSEN-54 265 RRRKTIIGRGLPVRLSPSGSPSLPSPIVH-----FLLYSVNGYSHYKRDASVIREAYRVQVHYDRTG-----PLKAF

HsTSEN54 482 VPDCLSLKRLSYQSGDVPILIFALVDHGDISFYFSFRDFTLPQDVGH-----
CeTSEN-54 296 PLKISELAELN-----VVFAMEEAGKVN---VAMSGASIDLNNYL---
PpTSEN-54 300 PCREDSTRFRNDLSSVPLLKAFVSLPRISF---TKESGAPIDLRSIEQSEGAKKKQSEGTKKNKKKK
DmTSEN-54 332 PSKNDVCPSPNSLDEGIPLEAVVEGPRVAF---RKTSGEWIELGTLLQK-----

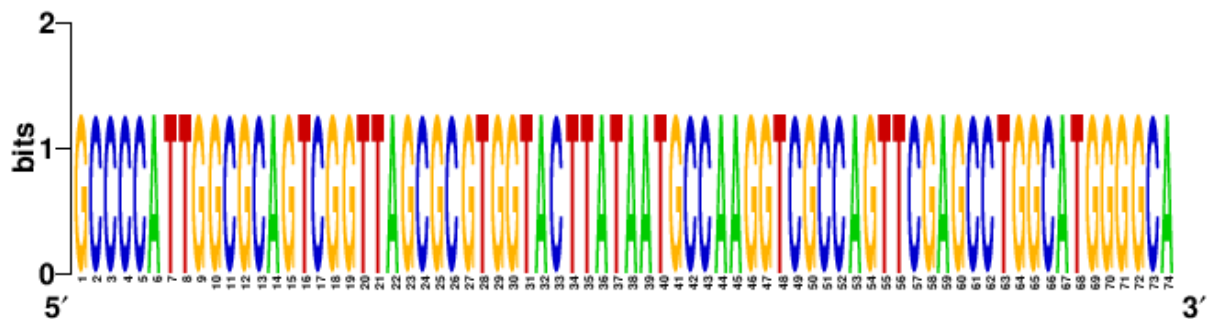
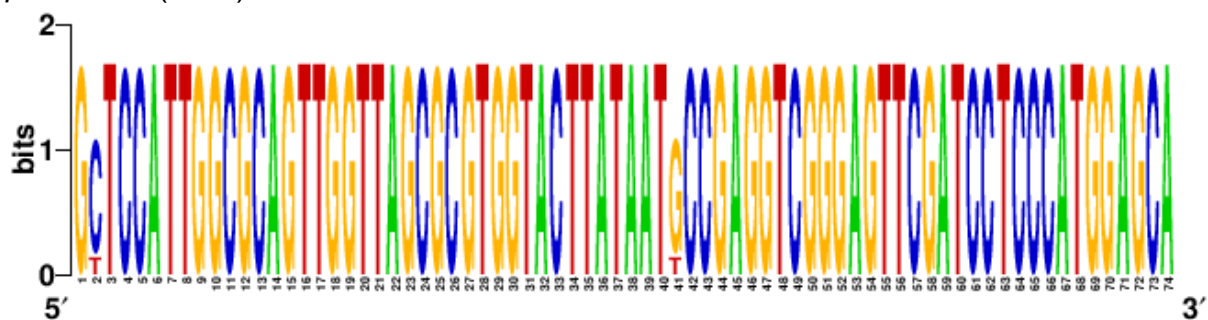
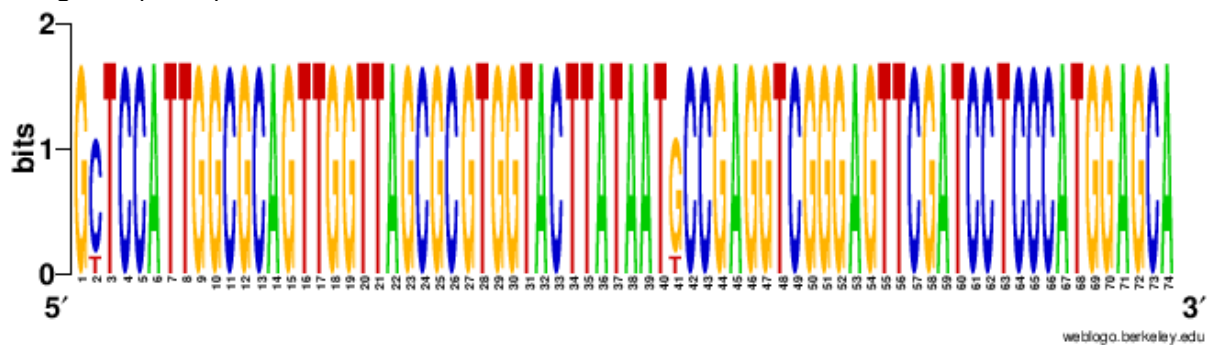
```

A

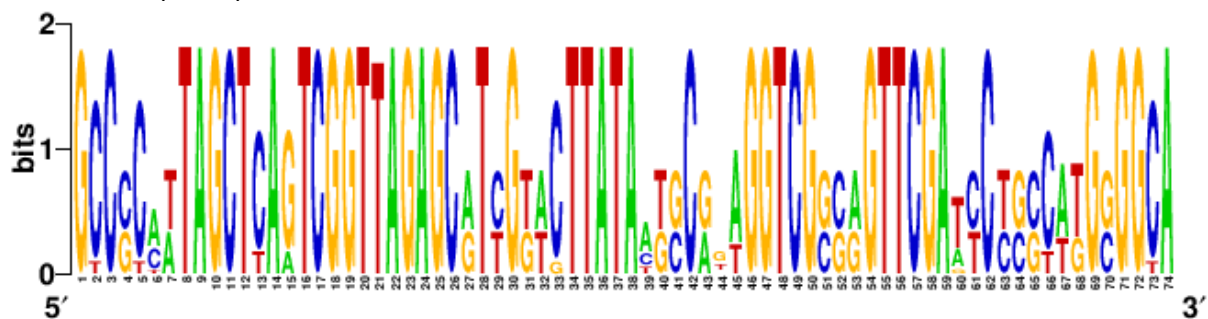
C. elegans tL(CAA) spliced tRNAs:*P. pacificus* tL(CAA) spliced tRNAs:*D. magnus* tL(CAA) spliced tRNAs:*L. texanum* tL(CAA) intronless tRNAs:

B*C. elegans* tY(GUA) spliced tRNAs:*P. pacificus* tY(GUA) intronless tRNAs:*D. magnus* tY(GUA) intronless tRNAs:*L. texanum* tY(GUA) intronless tRNAs:

C

C. elegans tI(UAU) spliced tRNAs:*P. pacificus* tI(UAU) intronless tRNAs:*D. magus* tI(UAU) intronless tRNAs:

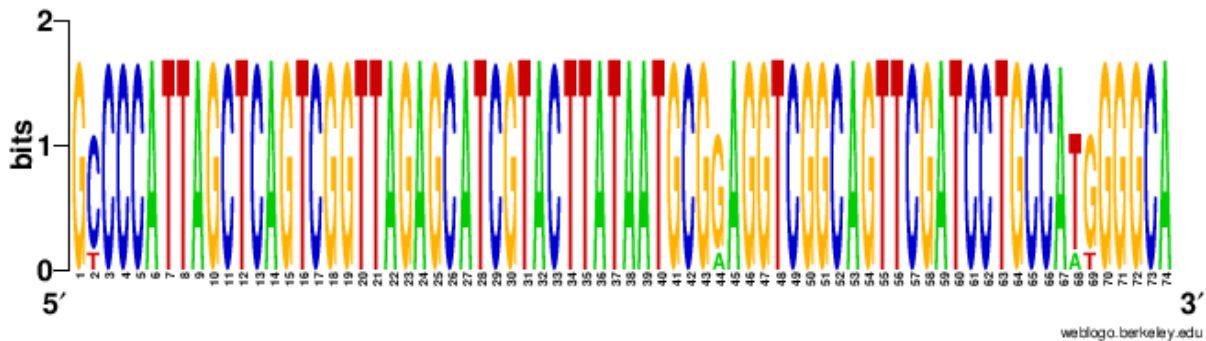
weblogo.berkeley.edu

L. texanum tI(UAU) intronless tRNAs:

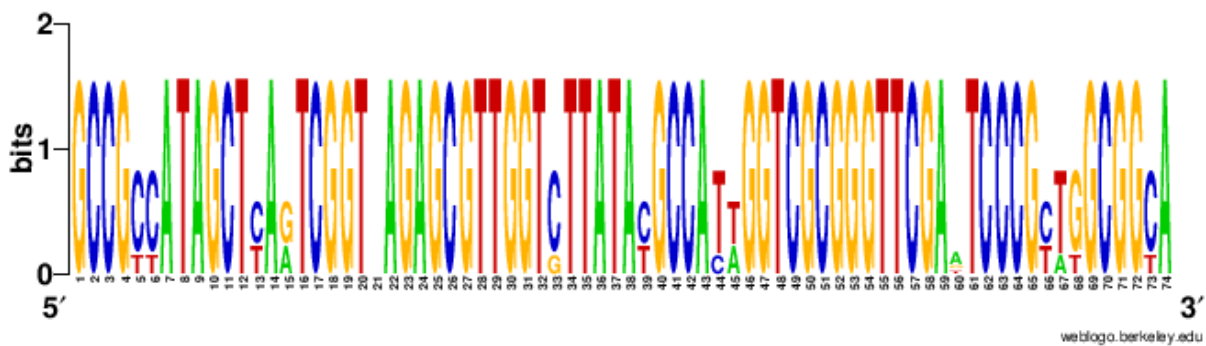
D

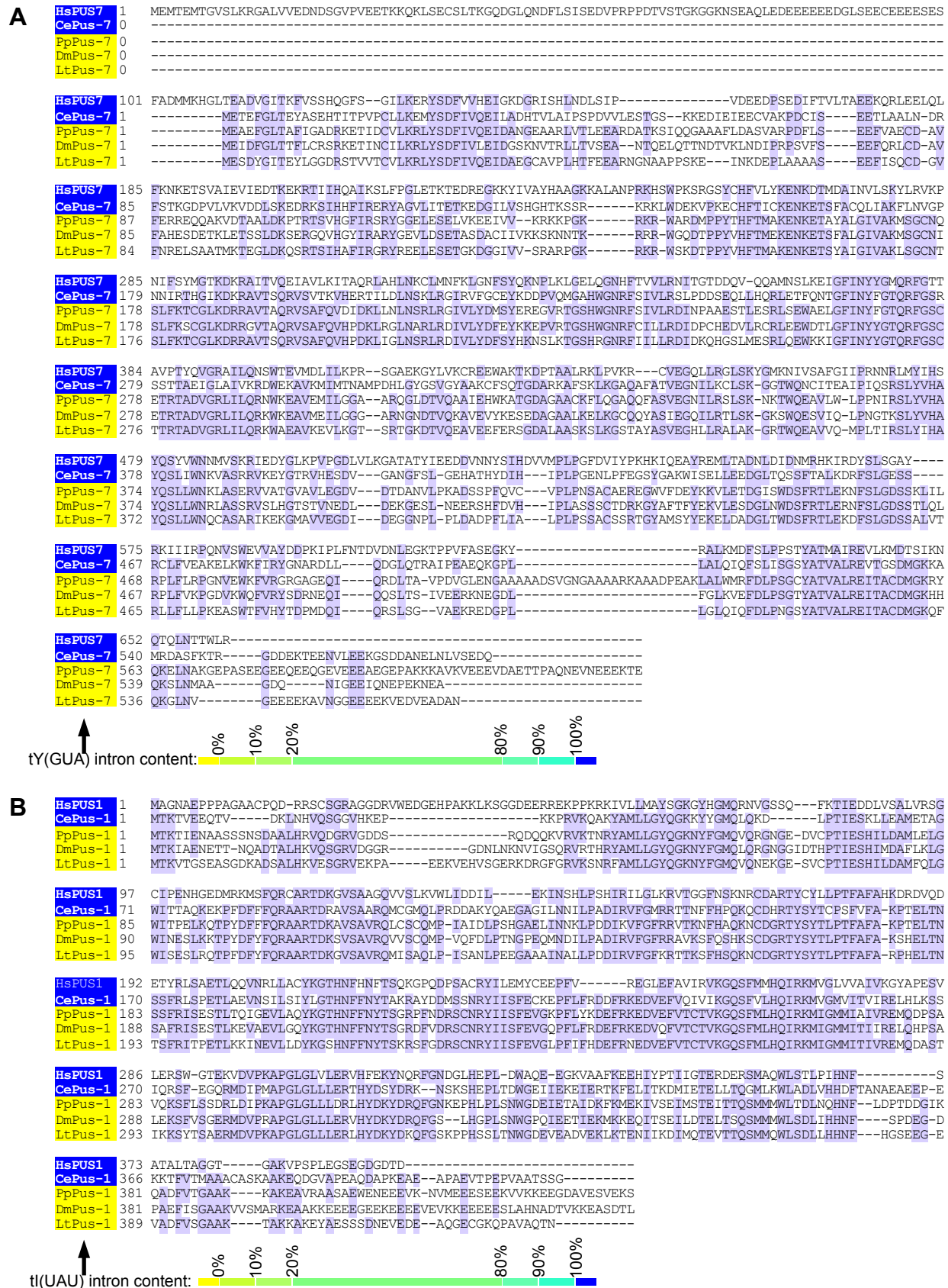
L. texanum tI(UAU) intronless tRNAs:

Sequence logo for of 7 of 12 genes:



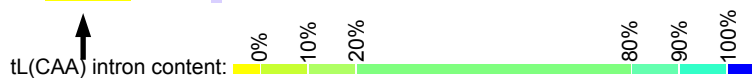
Sequence logo for of 5 of 12 genes:





C

HsNSUN-2	1	MGRSRGRRLQQQORPEDAEDGAEGGGKRGAGWEGGYEIVKENKLFHYYQELKIVPEGEWQFMDALREPLPATLRITGYKSHAKEILHCLKNKYFK
CeNSUN-2	1	MVH---SSRRKFKSKNGHEPRENVWRDDRDPAGNQGYRMVEKENEKYQYQYKEQNVFPLEQFDEFRAALQRDLPSVSRFQGGCHKDREQLIHENRFFG
PpNSUN-2	1	MGK--FRNKKNFKNRSKSNQ---WTQDRVVGF-----KTNDKLFDFYKKQGLIPEGEWDQFVETLKKDLPVSFRVQGGCHKDREILQELKERFIA
DmNSUN-2	1	MGKGNRNKHKHNNRSKSNQ---WSQDRVVGF-----RNNDKFFEFYKKQGLIPEDEWTFEAAALKRDLPSVSRVQGGCHKDREILRDELEEKYFK
LtNSUN-2	1	MGK---NKNKKHKNRSKSNQ---WTQDRDRVANT-----KTSEFFFEYQKQALIPGEWDDFVAALKRDLPSVSRVQGGCHKDREILRKELEDKFFA
HsNSUN2	101	ELEDLEVDGQKVEVPQPLSWYPEELAWHTNLSRKILRSKPHLEKFHQLVSETESGNISRQEAIVSMIPLLLNVRPHHKIILDMCAAPGSKTTQLIEMLHA
CeNSUN-2	98	KIMEAGDTA--VCVPKSLPWYKE--AYQTPMSRTAVRSHPILAQLHNFLVTEAELGNLSRQEAIVSMIPLLLPATSEHFVLDLCAAPGSKTTQLLEMIHE
PpNSUN-2	88	PIAASGDAG--AFEPKPLPWYYG--AYQTPMSRSEVRSHPITLALHNFLVTEAELGNVSRQEAIVSMIPLLLPKPAHHHVLVCAAPGSKTQMQLIEMQHE
DmNSUN-2	90	PIEESTETG--VFVPKKLPWHPY--AYQTKMSRSEVRAHPSLAKLSFLVTEAELGNVSRQEAIVSMIPLLLAPEPHHHVLVCAAPGSKTQMQLIEMMHE
LtNSUN-2	89	PIAASEEAG--VFVPKPLPWYKG--AYQTPMSRSEVRAHPTLSALHNFLVTEAELGHISRQEAIVSMIPLLLPAPQPHHHVLVCAAPGSKTQMQLIEMLHE
HsNSUN2	201	DMNVFPPEGFVIANDVDNKRCLLVHQA-KRLSSPCIMVNHDAISSIRLQIDVDGRKEILFYDRILCDVPCSGDGTMRKNIDVWKWTTLNLSQLHGLQ
CeNSUN-2	193	--NEENPKGMVIANVDNKRCLLVHQA-KRLSSPCIMVNHDAISSIRLQIDVDGRKEILFYDRILCDVPCSGDGTMRKNIDVWKWTTLNLSQLHGLQ
PpNSUN-2	183	--TDPNPSGFIANDVDNKRCLLVHQA-KRLSSPCIMVNHDAISSIRLQIDVDGRKEILFYDRILCDVPCSGDGTMRKNIDVWKWTTLNLSQLHGLQ
DmNSUN-2	185	--NDPNPSGFVANDVDNKRCLLVHQA-KRLSSPCIMVNHDAISSIRLQIDVDGRKEILFYDRILCDVPCSGDGTMRKNIDVWKWTTLNLSQLHGLQ
LtNSUN-2	184	--GDANPTGFVANDVDNKRCLLVHQA-KRLSSPCIMVNHDAISSIRLQIDVDGRKEILFYDRILCDVPCSGDGTMRKNIDVWKWTTLNLSQLHGLQ
HsNSUN2	300	LRIATRGAEQLAEGGRMVYSTCSLNPIEDEAVIASLLEKSEGALELADVSNELPGLKMPGITQWKVMTKDGQWFTDWDVAVPHSR--HTQIRPTMFPFKD
CeNSUN-2	289	IAIARKGAQQLKVGGRMVYSTCSMNPIEDEAVVAQLLRDSKGSKLVDTSLLPELKRGTGVQWKVDRDMKLYKSLEDISEEK-MKKVIVSSLPFPTTE
PpNSUN-2	279	LSIARRSPQQLKVGGLLVYSTCSMNPIEDEAVVAQLLRDSKGSKLVDTSLLPELKRGTGVQWKVDRDMKLYKSLEDISEEK-MKKVIVSSLPFPTTE
DmNSUN-2	281	LSIARRSALQKVGGLLVYSTCSMNPIEDEAVVAQLLRDSKGSKLVDTSLLPELKRGTGVQWKVDRDMKLYKSLEDISEEK-MKKVIVSSLPFPTTE
LtNSUN-2	280	LSIARRSIQQLKVGGLLVYSTCSMNPIEDEAVVAQLLRDSKGSKLVDTSLLPELKRGTGVQWKVDRDMKLYKSLEDISEEK-MKKVIVSSLPFPTTE
HsNSUN2	398	PEKLQAMHLERCLRIPLHHQNTGGFFVAVLVKSSMPWNRKQPKLQKGSSETRESTQLSPADLTGKPTDPSKLESPTSFTGTGDTIEAHATEDLENNGSK
CeNSUN-2	387	-QEAKEMLNLFHTRITPHAQDTGGFFVALIEK-----VS-----ETSF
PpNSUN-2	378	-EEAKEMLNLFHTRITPHAQDTGGFFVALIEK-----VS-----ETSF
DmNSUN-2	380	-EEANTFNLHVMRVVPHHQTGGFFVALIER-----VAED--DDEE
LtNSUN-2	378	-EEAQFNLEYVLRVVPHHQTGGFFVALMER-----VAED--ETTK
HsNSUN2	498	KDVGCGPPSKMKMLFGKEDPFVFIPEDDPLFPPIEKFYALDPSPFRMNLTRTTE---GKKRQLYVMVSKELNRVLLNNSEKMKVINTGIKVVCRNNSG
CeNSUN-2	425	DPSAANGPAWKKKMM--FKDEPFTFLKLDDEWINDIRNHYGVDESEFLYQNLFSRRLESNDANSRQLFYANDAVKNEFKENMSKVSIQNAQMKMFGRNENK
PpNSUN-2	419	RRVGARAPAHKQNM--FKDEPFTFLQKDDERWDDLSYGIADSFYENLFRNLIE--GDNAQLFYVNEGKQFILKNMKSLSIQNAQMKMFGRNENK
DmNSUN-2	420	ISNATKAPGYKQKMM--FKDDPFTFLQKDDPRWFDLKPYYGIDDAFKYENLFRNMQE--GDNAQLFYVNDGVKEFILLNMKSLSIQNAQMKMFGRNENK
LtNSUN-2	421	KEKTAKAPGYKQKMM--FKDDPFTFLQKDDPRWFDLKPYYGIDDAFKYENLFRNMQE--GDNAQLFYVNDGVKEFILLNMKSLSIQNAQMKMFGRNENK
HsNSUN2	595	EEFDCAFLAQEGYITLYPFINSRIITVSMEDVKILLTQE-----NPFKRLSSETYSQAKDLAKGSIVLKYEPDSANPDALQCPVILCGWRGKASI
CeNSUN-2	523	VE-NTRFRLSQEGIRYLFKFMQKQVKIGTEDMLFMLKSD-----ENLVPLEKLECKGDVRKQNGTVVVYCDDEDEP-----VCTWVGYHTI
PpNSUN-2	515	VE-KVRYRISQEGVDYLHFRMSNQKMRIPKEDMLILLRAAGKQDEKGVKALVDMELLQSKTELRTLKSGSVVLYVDETRP-----ICAWIGQRTA
DmNSUN-2	516	VE-KVRYRISQEGVEHVYFPMFSKQKMQIPVEDMLTMLKAAAMREGG-KALVDMELLKSKDALRTLKSGSCVLYVDESERP-----VCAWIGQRTA
LtNSUN-2	517	IE-KARFRISQEGVEYLYKFMFSKQKMQIPKEDLLTLRAAS--KGEK-KALVDMELLASAEKRLRTLKSGSVVLYVDEERP-----VCAWIGQRTA
HsNSUN2	687	RTFVPKNERLHYLRMMG----LEVLGEKKKEGVILTNEASAAGTQPDNDVTEGQRAQEPNSPDAAE---ANSPDVTAGCDPAGVHPPR-----
CeNSUN-2	604	APYISKEERLHLLRMMGVDCAEIELLMKSKRKEKAAQDREAARKEAAKAAEPSSSSDAPESAETLEIPEENVQEMTE-----DELKA
PpNSUN-2	604	APYIGKEERIHLRLGEDVDSMEAIEKTKRKQKAMEAREAAWAEKAKEDEATG--AAIKGEDIIVE--GGDEAEKQVKRIKLDESA-----DELKA
DmNSUN-2	604	APYVGKEERIHLRLYGEDVHTLEALQTKRKKEKAMKREEKAEQNA-EKAAGDAEIEGEMKEEENGEGEDTESNGEENVEEAEKAPRLEEKEE
LtNSUN-2	603	APYVGKEERIHLRLYGEDVDTMEALMKTREKAMRDREAWEAKAQDKAAAEATNGDEGTNEEGEEDKTEEDDSSPLPKKSRLEGGD---QAEAMEG
HsNSUN2	767	----
CeNSUN-2	684	----
PpNSUN-2	693	EDEE
DmNSUN-2	703	EVKE
LtNSUN-2	699	ATEE



D

```

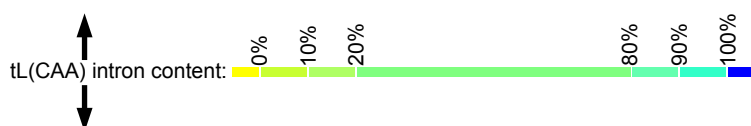
HsTRMT61A 1 -----MSFVAYEELIKEGDTAILSLGHGAMVAVRVQGAQTQTRHGVLRLHSVDLIGRPFPGSK
CeTRMT-61 1 -----MTSSPPSFMSEYEDRIQEGDVTIVVYVYTGQMVPTIVKRGQTLMMKYGALRHEF-IIGKRWGQR
PpTRMT-61 1 -----MLKTESTDSTMSQKGVVVGSPFVCLSDTIEEGDVTIVVVSFGSMYPIVVKRGMSTMKFGQLRHEF-LIGKQWGSR
DmTRMT-61 1 MTSGEENMGLGEEGRKAEDIRAMVKAESTDSTVSDKGST--GPSFVRWMDKIEEGDVTIVVYVTFGSMYPTVVRRLSLTMMKYGQLRHEF-LIGKFWGSR
LtTRMT-61 1 MASGDTLI----EGIK--DARLAMMKAESTDSTMSDRGST--GPSFARWSDKIEEGDVTIVVYVTFGSMYPIVVKRGLSLTMMKYGQLRHEF-LIGKQWGSR

HsTRMT61A 58 VTCGRGWVYVILHPTPELWTLNLPHTQILYSTDIALITMMLLRPGSVVCESGTSGSVSHAIIRTIAPTGHLLHTVEFHQQRAEKAREEFQEHVRGRWV
CeTRMT-61 62 LSA-TAGYVYILRPTSDLWTRALPRRTQILYTADCAQILSLLDKAPGSVICSSTGTSGSLSHAIALTVAPTGHLYTHDIDETRTHKVLVEEFKVHGLGSVT
PpTRMT-61 77 VTT-TAGYIHALRPTSDLWTRCLPKRTQILYTPDVSFILSLLDVHGGSVVAESGTSGSLSHALAIIVLPDGHVYTHDIDETRTRKIEAEFEKHLGSHVT
DmTRMT-61 98 ITT-TAGYIHALRPSSDLWTRCLPKRTQILYTPDISLILYLLDVGGSVVAESGTSGSLSHAMACAVLPDGHVYTHDIDESRTRKVEFEFHEGLAHVV
LtTRMT-61 92 VTT-TAGYIHALRPTSDLWTRCLPKRTQILYTPDVSFILYLLDVKGCVIAESGTSGSLSHATATAVLPDGHLYTHDIDESRTRKVEFEFHEGLAHVT

HsTRMT61A 158 TVRTQDVCRSGFGVSHVADAVFLDIPSPWEAVGHWADAL-KVEGGRFCFSFPCIEQVQRTQALAAARGFSELSLEVLQV--VYNVRTVSLPP-PDLGT
CeTRMT-61 161 TAVVRNVCTSGFHVENAADGVFLDVPAPWEAIPFAAKSIFRARGRLVSFSPCLQVQVQACLAMAAAGFVSIEITIELVPEQKKIVSQHRQTLEEFEEELGD
PpTRMT-61 176 TAVVQNVCIIEGFFVENS CDGVFLDVPSPWEAVENAAARALS PRGGRLVSFS PCIEQVQRTATMMRESGFVQVECIETVPRQLLVVDERKQSLSEFNLYGD
DmTRMT-61 197 TAVVQNVCTEGFFVENS CDGVFLDVPSPWEAIDAAAIALS PRGGRLVSFS PCIEQVQRTCVAMKEKGFVQIETVEIVPRQLVVVDERKQSLRDFAAANGD
LtTRMT-61 191 TAINQNICVEGFFVENACDGVFLDVPSPWEAVGHAARALS PRGGRLVSFS PCIEQVQRTALAMREEGFVQIETIEIVPRSLLVVDQRKQLTAAAFNEFGD

HsTRMT61A 253 GTDGPAGSDTSPFRSGT-----PMKEAVG-----HTGYLTFATKTPG--
CeTRMT-61 261 AYP----KERKRNVDGTIVEQSSSAIPRITIAIVYVYSPQPTHTGYLTHATMLPVAE
PpTRMT-61 276 GEECAPGKGRKRKIPSGKDDDKMKELQST-TIICPSTQPTHTGYLISATMLPRL-
DmTRMT-61 297 YEECAPGGRKRKATGG---REGLEDNNLT-TIITPSTQPTHTGYLISATMLPRL-
LtTRMT-61 291 AEECAPGKGRKRKAPGG---DGPDEEAVT-TIISPTQPTHTGYLISATMLPRM-

```

**E**

```

HsTRMT6 1 -----MEGSGEQPGPOPHPGDHRIRDGDFVLKR--EDVFKAQVQRRKKVTFEKQWFYLDNVIGHSYGTAFEVTSGGSLQ-----PKKKREEPT
CeTRMT-6 1 -----METEPAETSOKLIKSEGYLVQKIDGQSRIVRFTPQKILIEKLKFVADSAPGKPHGL-FEVSNNOCLPMSVDRLEELQLOEAS
PpTRMT-6 1 MHSAMATIESGTSLESTTPLDASTPIITEGCVAVVRKVGGEHQVRVRLTANSNVLEKLRFEAESAIHGHPFGL-FEVIGKRLVPATV----EDLKRKRDGG
DmTRMT-6 1 ---MATIESGTSLESATPVDASTPIITEGCVAVVRKVGGEHTRVVRLLTNTNVLIEKLRFEAESAIHNPFGI-FEVIGKRLVPASV----EDLKRKRDGG
LtTRMT-6 1 ---MQMATIESGTSLESAPLDASTPMITEGCAIITQKVGGEHTRVVRLLTANTNVLIEKLRFEAESAIHGHPFI-FEVIGKRLVPATV----EDLKRKRDGG

HsTRMT6 85 AETKEAGTDNRNIVDDG-----KSQKLTQDDIKALKDKGIKGEIEVQQLIENSTTFRDKTEFAQDKYIKKKKKYEAIIITVVKPSTR
CeTRMT-6 86 VATSSIDPEVTEPSEPSLDPRVVIAPSALKLEPEQKQKLEMDAVLEMKKQGVSGQEMVAKLVEGSASFQTRTVFSQSKYIKRKAHSDRVLILRPTIR
PpTRMT-6 92 -DIEMAADVADNVGLNGNEQGEVIAPSAIDAE---TRQOLTEEQISEMKKGGAQANTLVSKLVSGSASFASRTTHAKSKYIRRKTKKHSRDFVILKPTIR
DmTRMT-6 95 EDIEMAENVGPMNGKDE-----IMAPSALNVD---TRQKMTAEDITEMKDGGTAGDVVSKLVAGNASFAGRTVYAKSKYIRRKTKKHSRDFVILRPTIR
LtTRMT-6 94 DDIEMATLHDETLPNGGEKP---VAPSAVVAE---TRQAITQEQITEMKEGGKASDVVSKLIEGSASFAGRTAYSKSKYIRRKTKKHSRDFVILKPTIR

HsTRMT6 167 ILSIMYAREPGKINHMRYDTLAQMLTLGNIRAGNKMIVMETCAGLVLAGMMERMGGFGSI IQLYPGGGPVRAATACF---GFPKSFLSGLYEFPLNKVD
CeTRMT-6 186 LLAKEYLKDPRDLAMLADQLALILQAGLHHGKNVVFETLGLITS AVIERLGGTGACIHLR--GAVAQSI PCVHSMNYDDNTLSFTLPV---RIK
PpTRMT-6 192 LLCEAYLRKDYDRAGCLRLDQLSLIIHQGAVHTGRKVLVFDQVLGLITAAVTERLAGAGACIHLR--GTVPQSI PCFNMSMEFGPEILSAFYFV---RID
DmTRMT-6 183 LVCEAYLRKDHDKAGSLRVDQLGQIIHLGAIHTGRKVVFDQVLGIVTASTVERLAGNGACIHLR--GTIAQSI PCLNMSMNFKAELISTFYFV---RIE
LtTRMT-6 188 LLSDTYLRKDHDRVGCRLRIDQLAAVMHQGAIHTGRKVI VFDQVLGVFTA AAVVERLAGSGACIHLR--HTVAQGIPCFNAMNFKPEVLSTFYFV---RID

HsTRMT6 264 SLLHGTFSAKMLSSSEPKDSALVEESNGTLEEKQASEQENEDSMAEAPESNHPEDQETMETISQDPEHKGPKERGSKKDYIQEQRRQEEQQRKRLHEAAAL
CeTRMT-6 281 CVLAGKHLP-----YDNNRAVNGDDDNLEEVNGDKEPEK-----IEDQDLEALARRNDRLDREKRGIDL
PpTRMT-6 287 CLNGFAQP-----SEEGERMEVEEASAEADDD---NSQPAEKEKQVQVQKQSKDKEGWGRKH--DEAWQAQKQKADRLQREADGLAA
DmTRMT-6 278 SVLNGCKQV-----EEEKGKEKDEEMNGGKT VNNMNGEHP-----GIVKKAERLAREVEGLNI
LtTRMT-6 283 SVLHGWQLP-----PDQEGTGDRKEEKGKNGRPSDECDEW-----KEAKEAASAAKKNERLQRESEGLAF

HsTRMT6 364 LSERNADGLIVASR-FHPTPLLLSLDFVAPSRPFVYCYQYKEPLLECYTKLRERGGVINRLSETWL----RNYQVLPDRSHPKLLMSGGGGYLLSGF
CeTRMT-6 341 INDEQIHSLIIGSRTVDPI SVLELLYPKLAPSGTIVIIYSPHONILISAYDWLSKR-QTINLVLTDMC-----RVMQVLPDRTHPLMAQHVAGGHILTGI
PpTRMT-6 370 I-EQGLDILIGSRVDPCLSLLELLYDKLRSGNVVYSPTIQHVTRAQQWLRER-GAIHIVLTDMQYRVQVQVQVLPDRTHPLMSQMVGGVVLAIAI
DmTRMT-6 330 V-VQGIETILIIATRSVCPISILEVLWPLKAAAGTVVYSPTVEHCLRAQQWIRKE-GAVNVTLSDQ-Y-----RVQVQLPDRTHPLMSQMVGGVVLAIAI
LtTRMT-6 346 LNEGGEVTVLIATRSVPSLLEECWGLKPAAGNVVYSPTLQHCARAQQWLQK-GAIHITLTDQLY-----RVHQLVLPDRTHPLMSQMVGGVVLAIAI

HsTRMT6 458 TVAMDNLKADTSLKSNASTLESHEETEPAKKRKPESDS
CeTRMT-6 435 KVIDQ-----
PpTRMT-6 468 KVLPAAPAEKKE-----
DmTRMT-6 424 KVLPAAPPTN-----
LtTRMT-6 440 KVIIIPVSAV-----

```

A

```

HsIRE1 1 -----MPARRLLLTLLLPGLGIF-----GSTSTVTLPE---TLLFVSTLDGSLHAVSKRTGSIKWTLKEDPVLQV--PT-HVEEP-AFLPD
CeIRE-1 1 -----MRATFHLFTFFILLLFSSVICISTPGFRNDHESIGDDEETSTILVSTIDGRRLALDSSETGEIKWTLQEEFVLRS--PS-AVQGGFTFLPN
PpIRE-1 1 -----MMALRVIFITLLISIGISSPSAVESNKNGETVGADPD--VSMILVSTLDGRRLALDVATGKEKWEQEDPVIRA--PS-TVTQGGFTFLPN
DmIRE-1 -160 -----MKRTFLLLLFLFGIGRGSNVHHKDGDRVTNTED--I--VLVSTLDGRRLQALDVLTKQRWQMQEAAVLRA--PN-SITRGFTFLPN
LtIRE-1 1 MRIRECSDRSQMRLQAAVLFLFLCFLVPGTHIE-QDSRSKPDWLDQ--I--IIVSTVDRRLGLDIASGKIRWQFSDEFPQAQSTPNGSFDPGFTFFPS

HsIRE1 77 PNDGSLYTLGSKNNEGLTKLPFTTPELVQASPCRSSDGILYMGKKQDIWYVIDLLTGK-----QOTLSSAFADSLCPST--SLLYLGRTEYT
CeIRE-1 90 PLDGSLYVL---KNSSLKKLPFNIPQLVHASPCCKGNDGILYAGSKKDVWFGIDPKTGLK-----VETLSSASADRCPCANQQTIFLGRTEYR
PpIRE-1 87 PQDGSLSYL---KEGVLRLKLPFSIPQVVSASPVVGNMGVYFAGSKKDVWIGIDPISGEK-----LETSSASYHRVCPANGVHTVFIGRREYR
DmIRE-1 -79 PQDGSLSYL---RNGALERLPFTIPQVVAASPTRGNMGVYFAGSKKDWLIGDPTTGKVLGNCLHFVQIIVETLSTAASTRMCPANHPNLVYIGRREYR
LtIRE-1 96 YQDGSLYQM---RKGE LRKMPFSIPQVVSAPCKNGMGLYAGSKKDEWIGLNPFTGER-----VERHGTGKPSSTPCVASHDDVVFATTEFA

HsIRE1 163 ITMYDTKTRELWNATYFDYAASLPED-DVDYKMSHFVSNGDGLVTVDSSEGDVLWIQNYASPVVAFYVWQR-EGLRKVMHINVAVELTRYLTFMGSEV
CeIRE-1 175 VSMFDEKNRGKTWNATFNDYSAHLLE-VNTWPFKHYASSSHGYILTFDRETGEMRWEQDLKQPVVALYLLRD-DGLHKLPPFVVGKETME--NVA-KNI
PpIRE-1 172 VTMYPDPTNRTSWNTTFVDYSTHLQPE-DISYPLRHLVSSSSGDVLAHEKTSGLLWQTDVGGQPVVAMYLRS-NGLHKLAYTPVGRETME--GIVEKSL
DmIRE-1 19 LTMYPDENVTRSWNTTFDYSAYLQPA-DSKYPRHLVSTSTGDVLAHKKDGSLLWQADVGGQPIVAMYLRS-NGLHRLLEYTPVGRETME--SIV-KSF
LtIRE-1 181 LTMVDPADQRSWNTTFVDYSLVDNARSAAFLHLHLVSSSTGDILAIDKKSGSFVWKISVGGQVAAAMYLKKPEGLLRIDAMPVGSVDLN--NTAENS-

HsIRE1 261 GRITWKYYPFPKETEAKSK-----LPTLYVGKYSTSLYASPMV-HEGVAVVPRGSTL-PLLEGPQTDGVTIGDKGECVITPSTDVKFDPGLKSKNKL
CeIRE-1 270 FTVDQWPTVLGVNAADPQTSLTNQFFPALFVGESSFGLYAEALVDHQTITYSFKLLGP-PLLEGPAPIALTEMEKEEYLPRRPIIRNIPPSITHKTS
PpIRE-1 268 ADKSIWSEVLSAAVSPSAKTG---FYPALYVGESVHGFASHAFVDEKTRTFAPRYLGP-NLIEGPTPIAMPSPSERP--IPRRRAVDIDLNTIIHKTE
DmIRE-1 114 ADASIWAELSLAVSPSAKTG---FFPALYVGATSNGLFAANAFVDDKTFTLAPKYLGRLLLEGPTPIGSADGGT---IPRRPIVITIDSDTIHKTD
LtIRE-1 277 -DPSVWRDLTLTASSPSKGTG---YYPAVYIGESTHGLFASGVYVDGSTPTIG-RSTAP-LLLDGPDGTSGTGDSG-----RPI--FDAEQIVTTKD

HsIRE1 353 NYLRNYWLLIGHHETPLSASTKMLERFPNNLPKHRENVIPADSEKKSFEVNLVDQTSNAPTIVSRDVEEKPAHAPARPEAPVDLSMLKDMATIILSTF
CeIRE-1 369 D---GEYLLGYHHRPMMTMATI-----IPT-R-YYPV-----GPHKAIKGSTIERPPQLLGEVPEKQHD-TSFIILLNNH
PpIRE-1 361 D---GEFLILGYHVPQLNMAAV-----IPT-D-NLPHSSSFIPFPYPTIEYV-KPKNDQKVGIPD---QPTLTI-----DDESMLALLITMYNRH
DmIRE-1 207 D---GEFLILGYHEPPRINLGAIV-----IPNHA-MPQFPAQAMHPYP-HIEYVIGTKDDDKKIIGTD---PLPQLTSDTSGDGEHDTLVLFWLKEQH
LtIRE-1 361 -----VILGYHEPPPEATEAA-----VPDVR-FVSPSGGGGPGGPKRIDYI--PEDDGPKVIPHVLDSQRESVLL-----QDETIFILGRIYERY

HsIRE1 453 LLIQWVAFIITYPLSMHQQQQLQHQQFQKELEKIQLLQQQQQLLPFHPPGDTAQDGEILLDTSGPYSESSGTSSPSTSPRASNHS-----LCSGSSASKAG
CeIRE-1 436 PIPFYATLVTFMALLLTVTI-----WQCGRQWDQKSTSRMDSFEIYNNPG-----ESRSAQTSKQS
PpIRE-1 439 PQQVISAATVAVILLIATI-----WMCGRQMGVAEG-ARSAQTSRVVRDE-----EEKDEV
DmIRE-1 292 FYALYSSAVTILILIISTI-----FMCGRQAGVAAGEARSQOTSRSVSTTM-----DGREES
LtIRE-1 441 P-TFMALCGTIVVLFGTTV-----WLCGRQSTLADLEARSQAASRVSVASSRDGIEEDGREGSRGG

HsIRE1 548 SSPSLEQDDGDEE--TSVVIVG-KISFCPKDVLGHGAEGTIVYRGMDNRDVAVKRILPECFSFADREVQLLRESDEHPNVIRYFCTEKDRQFYIAIEL
CeIRE-1 492 NRGSGFWANRKIEPEGWMAVGSKLMYSPSIDLGTGCEGTVVYRGTFDGREVAVKRVVSEFVKFAHREADLLRESDTHPHVIRYFCMESDSQFRYLAL
PpIRE-1 489 TFPSSGKNRARIQLTGWMQVG-KVHYDASSVLGTGCGQTVVYKGFDFGREVAVKRVLTETFLPLVEREVELLRSDDPHPNVIRYFCMESDSQFRYLAL
DmIRE-1 343 GSGNVGWKRGRQP-VTGWIQVG-KVHYDTSIDLGTGCGQTVVYKGFDFGREVAVKRVLTETFLPLVREVELLRESDSHAHVIRYFCMESDSQFRYLAL
LtIRE-1 501 GGGGGGKNRNRG-LQGWIPVG-KLEYNASEVLGTGCEGTVVYKGFDFGREVAVKRVLTETFLPLVREVEYLRSDSHPNVIRYFCESDSQFRYLAL

HsIRE1 645 CAATLQEQYVEQK-----D-FAHLGLEPITLLOQTTSGLAHLHSLNIVHRDLKPHNILISMPNAHGKIKAMISDFGLCKKLAVGRHSFSR--RSGVPG
CeIRE-1 592 CIASINDYVEQK-----EVQONVTIALRDMQQTADGLAHLHASKIVHRDMKPNQNLITMASORGEMRAVISDFGLCKRVQPKNSISRGIASGLAG
PpIRE-1 588 CSASLYDYVEKKRRKNDEQEFDFLEISPEEILRQSTEGLAHLHSLNIVHRDLKPNVLIS-TRGRG-SRALISDFGLCKRLHPGRHSLSK--RSGVAG
DmIRE-1 441 CVASLYDYVERKK-----EEVRQFVEMEPCEILRQATDGLAHLHSLNIVHRDLKPNVLISNTGGRK-TRALISDFGLCKRLHPGRHSLSK--RSGVAG
LtIRE-1 599 CAGSLYDYVERKK-----DLRGRLPITTEFELLKQATDGLKHLHEMKIVHRDLKPNVLINVSOG-GRASAVISDFGLCKRLNPGMRSLSR--RSLGAG

HsIRE1 734 TEGWIAPEMLSEDCKENPTYTVDFISAGCVFYVYISEGSHFFPKSLQQRANILLGACSLDCLHPEKHEDVIARELIEKMIAMPQKRPSAKHVLPKHPFFW
CeIRE-1 684 TDGWIAPEVLI---SASTSYFVDIFSLGCIFYVYLTSGTHPFPGKSLHRQANIVNGEYTLNKL-ADLDDWSLADDLSSMLNVEPLHRLTADAVLNHPFFW
PpIRE-1 684 TDGWIAPEALT---QGSTSFPMDFISLGCIFYVYLTGSHPHYGEALHRTSNILNGAHSLTGL-SAEET-RLSTHLISLAISPSPTRRPSAQAILTHPFFW
DmIRE-1 532 TDGWIAPEALK---AQSTSFPMDFISLGCIFYVYLTSGRHPFGESLHRTSNILNGMYTLTSAV-SEGECDRRSTHLIDQMLRPVPHRRPTASATLTHPFFW
LtIRE-1 689 TDGWIAPEALK---RLSTSFPMDFISLGCIFYVYLTGKHPHYGEPLHRNGNIFNGSHRLDDL-PELES-RRFGHLIEAMLAASPTRRPTAAAVLHPFFW

HsIRE1 834 SLEKQLQFFQDVSDRIEKESLDGPIVKQLERGGRAVVKMDWRENITVPLQTDLRKFRITYKGSVRDLLRAMRNKKHHYRELPAEVRETLSGLPDDFVCYF
CeIRE-1 780 TSEKRLAYFSDVSDRVEKEEDNSPVVRRITDARIIVCGGWREKICDALKEDLRKFRITYKFSVRDLLRAMRNKKHHYRELPEDEVQSLGDPDQFLHYF
PpIRE-1 779 GTEKQLAFFQDVSDRVEKEVDMSFVLRLEKNGFNVVKKNWRDRICDALREDLRKFRITYKAHSVRDLLRAMRNKKHHYRELPEDEVNSLGEVPHEFVYVF
DmIRE-1 628 DKGRQLQFFSDVSDRVEKEMDMSPIRLRLEKNGFAVVKKNWREKICEALQEDLRKFRITYKGHSVRDLLRAMRNKKHHYRELPEDEVNSLGEVPHEFVYVF
LtIRE-1 784 DTAQQLQFFSDVSDRVEKEPDLSPIITRLEKSGRFVVKMNWRERICHELQEDLRKFRITYKGHSVRDLLRAMRNKKHHYRELPEEVRLSGLPIPHDFVAYF

HsIRE1 934 TSFRFPPLLAHTYRAMELCSHERLFQPYFHE-----PEEPQ-----
CeIRE-1 880 TSFRFPPLLLHYVYKATEYCSGEAVFKRYSDDVARMYPIVEEEERVKKIKEEMANEVWARAP-----KPVEQRTPLKLDKRNK-----
PpIRE-1 879 TSFRFPDLLHYVEAMWCSEEAIFISYYSDVRLRMKEVREEME--KRMAEENSSETWHRGRPSVSTDCLNASVPERRRLELLPRSKPHITSESSI
DmIRE-1 728 TSFRFPPELLMHVYDAMEWCAEEAVFMGYYSNEVRDRIGEVRAETE--KRMAEENSSEVWLRGRVSTSSNQD---SPQBERKKLELLPRSFP-----I
LtIRE-1 884 TTRFPPELLMHVYEAIEWCAEETVFVHYYSSEVVRIRDRVRAETE--KRMEESSDNEVWQRGVAPPPSDSP---SPQTERKRIVLLPRSKE-----

HsIRE1 969 -----PPVTPDAL-----
CeIRE-1 959 -----KKSNNPTD-----
PpIRE-1 977 TSEDQSAIVDEVTPMEEKDT-PTLVPRSELIVATTTTNTVVLNMN-GPPPGLEDV-----VPIP-----PEGMOT-NGGNKKKKKKKK
DmIRE-1 815 TTPPTDSATQATPSEEEETPTLVPPSELKSAESTASMTVLDLIGPPPGLEGM-----APIP-----PDVSDPFQKTERTKKKKKRR
LtIRE-1 970 PSESTALIETLVPPSELLPKNPPIAP--SMKADVSIEAEVEIMM--GPPPGLEGIRREEAVPNQAPANGEEATTPTGGAAPQEATTKKKKKRR

```

B

HsRTCB	1	MSRSYNDELQFLEKINKNCWRIKKGFVPNMQVEGVFYVNDALEKLMFEELRNACRGGGVGGFLPAMKQIGNVAALPGIVHRSIGLPDVHSGYGFAGNMA
CeRTCB-1	1	MPRTFEEECDFIDRLTDTKFRIKKGFVPNMNVEGRFYVNNLEQLMFDELKFCSDGQIGGFLPAVRQIANVASLPGIVGHSIGLPDIHSGYGFSGIGNIA
PpRTCB-1	1	MPRTFAEEKSYIERISPTCFRIKKGFVPNMVVEGRFYVNNLEKLMFNELEVFTNNPGYGGFLPAVRQMANVAALPGIVGASIGLPDIHSGYGFSGIGNIA
DmRTCB-1	1	MPRTFEEECQFIERIDPTCFRIKKGFVPNMKVEGKFYVNSELEKLMFKELEVFTSSPAYGGFLPGVRQIANVAALPGIVGHSIGLPDIHSGYGFSGIGNIA
LtRTCB-1	1	MPRTFKEECQFIQRVDDTSFRIKKGFVPNMNVEGKFYVNSELEKLMFDELEKYSNNPGYGGFLPAVRQIGNVAALPGIVGOSIGLPDIHSGYGFSGIGNIA
HsRTCB	101	AFDMNDPEAVVSPGGVGFDINCGVRLRLTNLDES DVQPVKEQLAQAMFDHIPVGVSGKGVIPMNAKDLEEAELEMGVDWSLREGYAWAEDKEHCEEYGRML
CeRTCB-1	101	AFDVGNPESVISPGVGFDINCGVRLRLTNLFEENVKPLKEQLTQSLFDHIPVGVSGKGAIPMLASDLVECLEMGMDWTLREGYSWAEDKEHCEEYGRML
PpRTCB-1	101	AFDVSNPESVISPGVGFDINCGVRLMRTNLSLEDVLPVKEKLTQNLFDHIPVGVSGKGAIPMTRNALEEAMEMGMDWTLREGYSWPEDQEHCEEDGRML
DmRTCB-1	101	AFDTSDEPESVISPGVGFDINCGVRLIRTNLSLEDVLPVKEKLTQSLFDHIPVGVSGKGAIPMTAKNLIECEMEMGMDWTLREGYSWPEDQEHCEEDGRML
LtRTCB-1	101	AFDASDPQSIIVSPGGVGFDINCGVRLIRTNLTLEDVAPVKEKLTQSLFDHIPVGVSGKGAIPMSARNLVECEMEMGMDWTLREGYSWPEDQEHCEEDGRML
HsRTCB	201	QADPNKVSARAKKRGLPQLGLTGAGNHYAEIQVVDIEIFNEYAAKKMGIDHKGQVCVMIHSGSRGLGHQVATDALVAMEKAMKRDKIIVNDRQLACARIAS
CeRTCB-1	201	QADASKVSLRAKKRGLPQLGLTGAGNHYAEVQVVDIEIDYDKHAASTMGIDEQGQVVMHLHCGSRGLGHQVATDSLVEMEKAMARDGIVVNDKQLACARINS
PpRTCB-1	201	QADHTKVSDKAKKRGLPQLGLTGAGNHYAEVQVQIEIDYDKHAAAAMGIDRKDQIVMLHCGSRGLGHQVATDSLVMHEKAMNRDSIIVNDKQLACARIHS
DmRTCB-1	201	QADASKVSTKAKKRGLPQLGLTGAGNHYAEVQVVEEIIDYDKHAAAAMGIDRKDQVVMHLHCGSRGLGHQVATDSLVMVERAMGRDAINVNDKQLACARINS
LtRTCB-1	201	QADASKVSSRAKKRGLPQLGLTGAGNHYAEVQVVEEIDFKHAAAAMGIDRIDQVVMHLHCGSRGLGHQVATDSLVEMEKAMGRDQIEVNDKQLACARINS
HsRTCB	301	PEGQDYLKGMMAAGNYAWVNRSSMTFLTRQAFKVFNTTDDDLDLHYIYDVSHNIAKVEQHVVVGKERTLLVHRKGSTRAFFPPHPLIAVDYQLTGQGPVL
CeRTCB-1	301	VEGKNYFSGMAAAANFAWVNRSCITFCVRNFAQKTFGMSADDMQVIYDVSHNVAKMEEHMDGRPRQLCVHRKGATRAFPAAHPLIPVDYQLIGQGPVL
PpRTCB-1	301	KEGQEYLSGMAAAANFAWVNRSCITYCIRKAFSRTFECQPDLDLQVVDVSHNVAKMEEHLLNGRPHQLCVHRKGATRAFPHPHPLIPVDYQLCGQGPVL
DmRTCB-1	301	QEGKEYLAGMAAAANFAWVNRSCITYCIRKAFSRTFECQPDLDLQVVDVSHNVAKMEEHMDGRPKELCVHRKGATRAFPHPHPLIPVDYQLCGQGPVL
LtRTCB-1	301	QEGKDYLAGMAAAANFAWVNRSCITYCIRKAFSRTFECQPDLDLQVVDVSHNVAKMEEHMDGRPRQLCVHRKGATRAFPHPHPLIPVDYQLCGQGPVL
HsRTCB	401	IGGTMGTCYSVLTGTGTEQGMTEFTFGTTCAGRALSRASRRNLFDQVLDKLDADMGIARVASPKLVMEEAPESYKNVTDVNTCHDAGISKAIKLRPI
CeRTCB-1	401	IGGSMGTCSYSVLTGTGTEQLVETFTFGTTCAGRALSRASRRITWDSVIDDLKKKEISIRIASPKLIMEEAPESYKNVTDVVDTCDAAGISKKAVKLRPI
PpRTCB-1	401	VGGSMGTYSYSVLTGTGTEKGMTETFTFGTTCAGRALSRASRREITYDQVLNDLEAKGISIRIASPKLIMEEAPDSYKDVDRVDVTDCHAAAGISNKAIKLRPI
DmRTCB-1	401	VGGSMGTYSYSVLTGTGTEGLVETFTFGTTCAGRALSRASRREITYEQVLGELEAKGISIRIASPKLIMEEAPGSYKDVDRVDVTDCHAGISNKAIKLRPI
LtRTCB-1	401	VGGSMGTHYSVLTGTGESGLVQTFGTTCAGRALSRASRREITYEEVVLQTLAKGISIRIASPKLIMEEAPGSYKDVTDVVDTCCHAGISNKAIKLRPI
HsRTCB	501	AVIKG
CeRTCB-1	501	AVIKG
PpRTCB-1	501	AVVKG
DmRTCB-1	501	AVVKG
LtRTCB-1	501	AVVKG

C

HsArchease	1	-----MKGGSRVSNPAVMAQEEDVDRDY-----NLTEEQKAIKAKYPPVNRKYEYLDHT
CeARCH-1	1	-----MPSTSMIEDRSEI-----ERRRFEYLDHP
PpARCH-1	1	MPIGKRASSTEDKDDGKSTTSKRSRRSESDNE SDGRDESANEVVRWTVNVDLRMELESIAEDEEGVFDVSNNGVEVEKEDADMSPSPVSNESRGFEYLEHP
DmARCH-1	1	-----MRESEKRSVDLEDNYSSK-----AKSETGAEMEDVLTNDSKGFYLEHP
LtARCH-1	1	-----MCEVSLKDMVVGADKE-----GSRGFEYLEHP
HsArchease	50	ADVQLHAWGDTLEAEFQCAMAMFGYMTDTGTVEPLQTVETQGDLLQSLLFHFLDEWLYKFSADEFFIPREVKVLSIDQRNF-----KLRSIGWGEEF
CeARCH-1	25	ADIQLHWSWGSTMEEAEACLVSMFGYMTDLAKVDEMEFYFKASGDSLGLLQFQLDEALNSFHAEPCFVAKRVEILRFDKKKF-----EIEFRGWGESF
PpARCH-1	101	ADILLHAWGTTDGEAMEALIESLYGLTELESIEEVVSYEMEAQGETEDVLINFTLQEALFAFQAEFFVGRRVVVNI EKDTKNGTMKGINIRAWGESF
DmARCH-1	47	ADILLHWSGDCDERALESLLSEMYGALTDLDSLKEMYSYELIAEGEHEEALLFNILQEALYAFQAEFFVGRRVVVQLERDGE-GKLKTAHVRAWGESF
LtARCH-1	29	ADVLLHAWGKDDEEAVALLLESYLALLTERNGIREVYSFEMEADGEHEHALFNILQEALYNFQTEPFVGRRVAVTKVEREA--GELRAIQVRWGESF
HsArchease	145	SLSKHPGTEVKAITYSAMQVYNE-ENPEVFVVIDI
CeARCH-1	120	DTSKHETEADIKSPTYSNMQINEKPERCDIYVIVDI
PpARCH-1	201	EMGKHVQKNDVKAPTYSNMQIKREKDRVDLYVVVDV
DmARCH-1	146	EFGKHMQKSDVKAPTYSNMQIKRELGRVEYFVVVDV
LtARCH-1	127	ETGTHVQKSDVKAPTYSNMQVKRENGRVEIFVVVDV

SUPPLEMENTAL REFERENCES

1. K. Hamashima *et al.*, Nematode-specific tRNAs that decode an alternative genetic code for leucine. *Nucleic Acids Res* **40**, 3653-3662 (2012).
2. K. Hamashima *et al.*, Analysis of genetic code ambiguity arising from nematode-specific misacylated tRNAs. *PLoS One* **10**, e0116981 (2015).