promoters, and a 3'-CCA that is also transcribed. The sequence surrounding the site of catalysis for RNase P, at the 5' mature end of the tRNA, is quite different from one RNA species to another.

In contrast to the tRNA sequence diversity, the structure of all bacterial tRNAs is quite similar. The commonly portrayed secondary structure of tRNAs resembles a cloverleaf, with four stems: the acceptor , TψC, anticodon, and D arms (Figure 1.1). Long-range nucleotide interactions define the 3-dimensional shape of tRNAs with two helices made up of co-axial arrangements of the D and anticodon stems and the TψC and aminoacyl stems (Figure 1.2). There is some degree of structural variability in tRNAs, as the "variable loop" (between the anticodon and D arms) can vary in length. Consequently, *B. subtilis* tRNAs vary in length from 70-92 nucleotides (nt). The minimal substrate structure for RNase P cleavage is the coaxial stems consisting of the acceptor stem and the T-stem with the TψC-loop [1]. However, bacterial RNase P recognition is also affected by several other tRNA structures, including the D-, anticodon and variable arms, as well as the 5' leader sequence and a Watson-Crick interaction between the 3'-CCA and an internal loop in the bacterial RNase P RNA subunit [8-12].

In addition to pre-tRNAs, bacterial RNase P is known to process several substrates that are proposed to contain tRNA-like structures. These bacterial substrates include 4.5S RNA, tmRNA, viral RNAs, mRNAs, riboswitches, ColE1 replication origin control RNAs, and C4 antisense RNA from phages P1 and P7 [13-21]. The presence of the

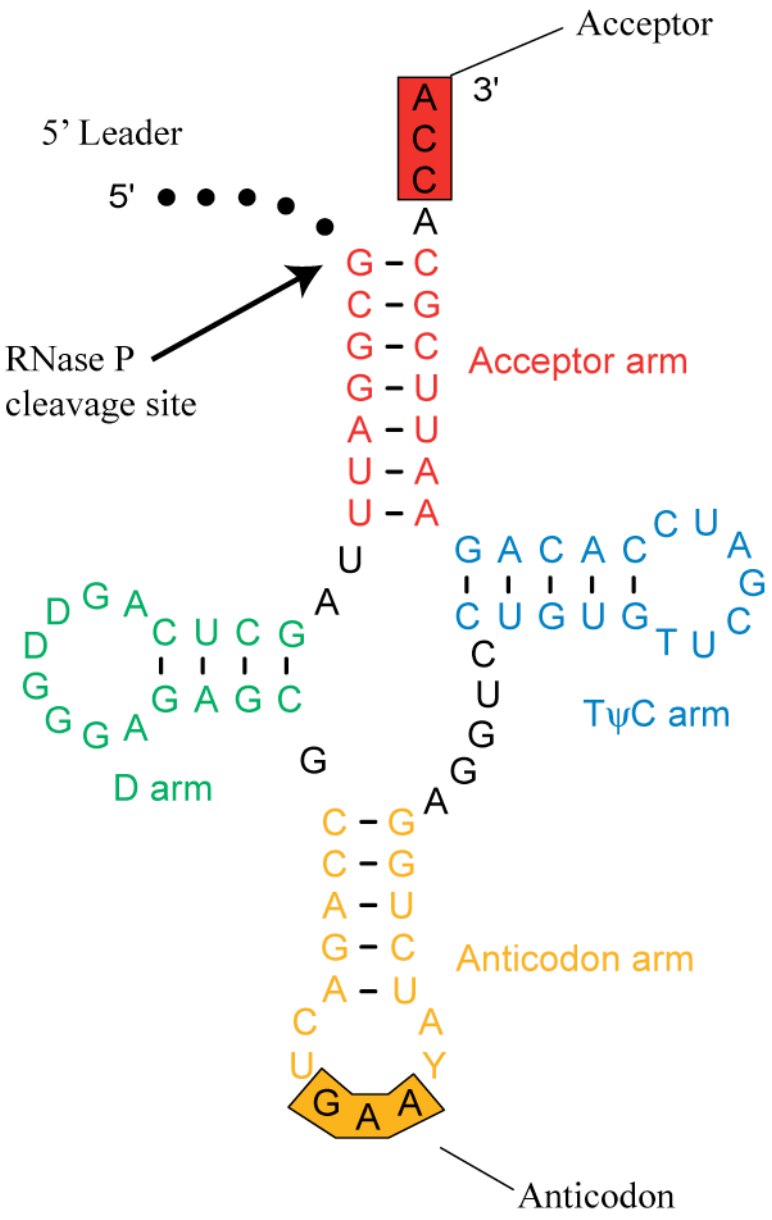


Figure 1.1. Representative secondary structure of a pre-tRNA

The cloverleaf secondary structure is one of the defining factors of a tRNA. Shown here is yeast pre-tRNA^{Phe}. The variable loop exists between the Anticodon and TψC arms.

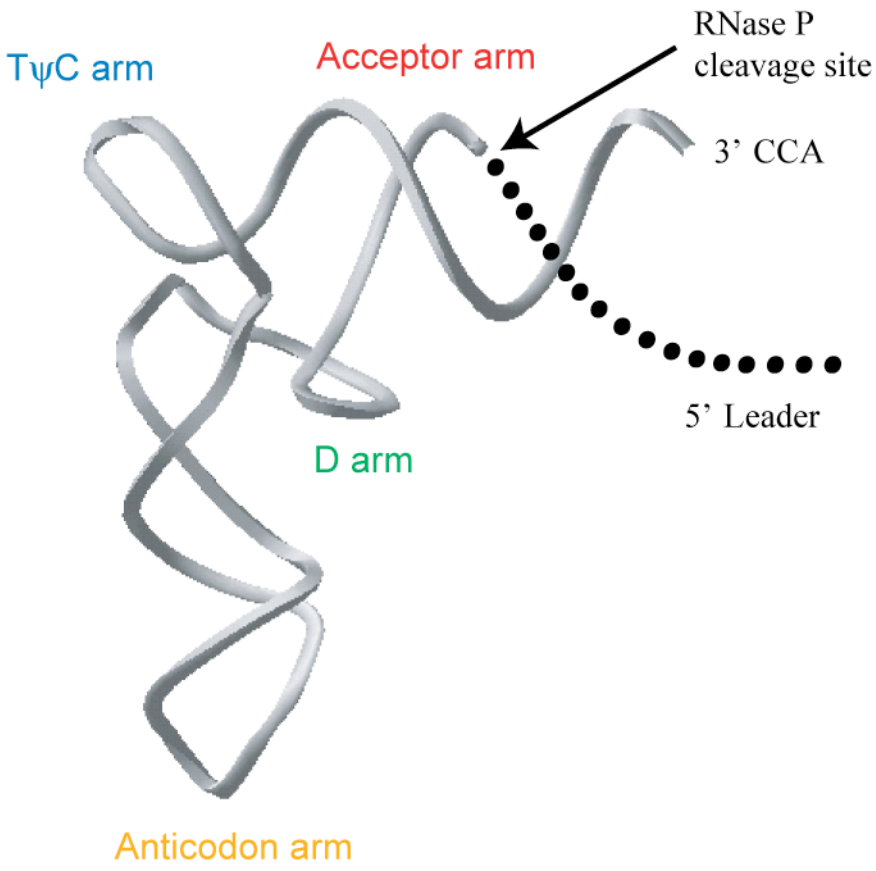


Figure 1.2. Three dimensional structure of a pre-tRNA

The cloverleaf structure illustrates the Watson-Crick basepair interactions, however the three-dimensional folding of a pre-tRNA also involves long range nucleotide interactions to form the co-axial stem that makes up the minimal substrate for RNase P.

protein subunit in the RNase P holoenzyme increases the substrate versatility of the enzyme over the RNA enzyme alone [7].

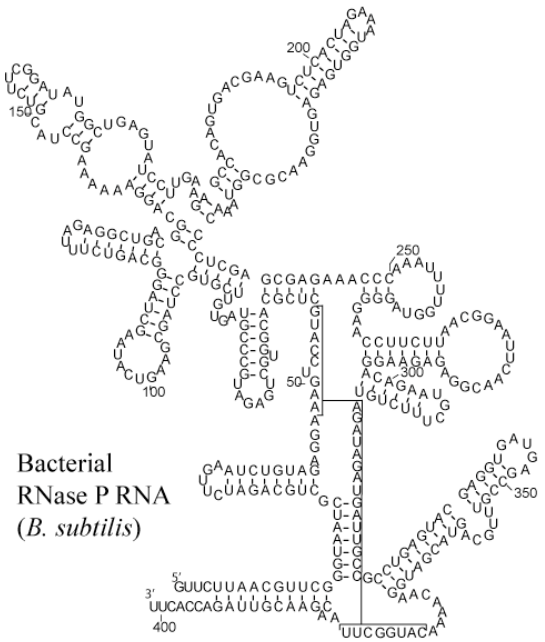
Eukaryotic RNase P

The function of nuclear RNase P in eukaryotic systems is much more complicated and less understood. First, there are two very similar enzymes in the nucleus that are related to bacterial RNase P, named RNase P and RNase MRP. Yeast RNase P is known to be responsible for processing pre-tRNAs, and RNase MRP is needed for the processing of a pre-rRNA (5.8S) and the regulated turnover of a cell cycle mRNA [22, 23]. Second, both RNase P and RNase MRP from nuclei are far more complex enzymes than bacterial RNase P [24]. Each enzyme still employs a distinctive, but related, RNA subunit and contains multiple protein subunits required for function *in vivo*. Yeast RNase P contains nine protein subunits: Rpr2p, Rpp1p*, Pop1p*, Pop3p*, Pop4p*, Pop5p*, Pop6p*, Pop7p*, Pop8p* [25]. Yeast RNase MRP, while a physically distinct enzyme, also contains the protein subunits with asterisks next to them as well as Snm1p and Rmp1p [26-29].

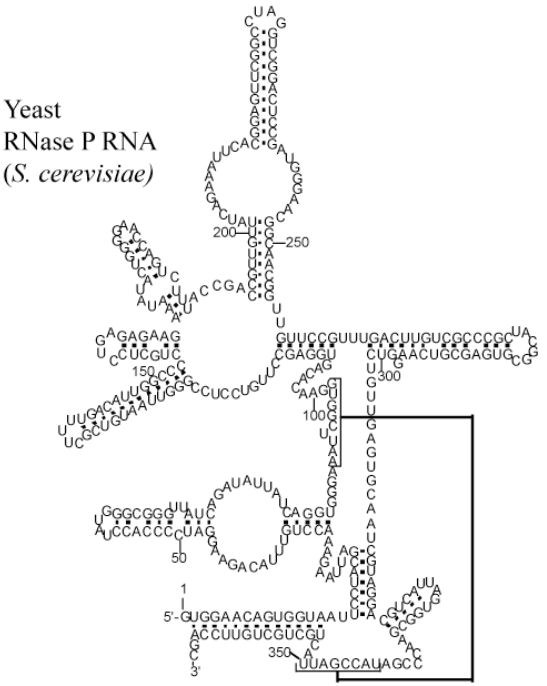
The eukaryotic RNase P RNA subunit alone is not generally considered to be catalytically active, although recent experiments have demonstrated the RNA subunit might contain a remnant of catalytic activity with an observed catalytic rate five to six orders of magnitude lower than the bacterial RNA-only enzyme [30]. This suggests that the eukaryotic enzyme relies heavily on its protein constituent. One known role for the

Figure 1.3. Comparison of RNase P holoenzyme from bacteria and yeast

RNase P from eukaryotes is a much more complex enzyme. While the RNAs are approximately the same size in bacteria and eukaryotes, the protein content changes dramatically. The bacterial RNase P protein is 10% (by mass) of the holoenzyme. In yeast, proteins make up 70% of the mass of the holoenzyme. Given the bacterial proteins role in substrate binding, specifically with the non-tRNA substrates, this massive increase in protein might facilitate the binding of multiple types of RNA substrates for cleavage by the catalytic RNA subunit.

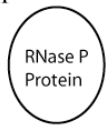


Bacterial
RNase P RNA
(*B. subtilis*)

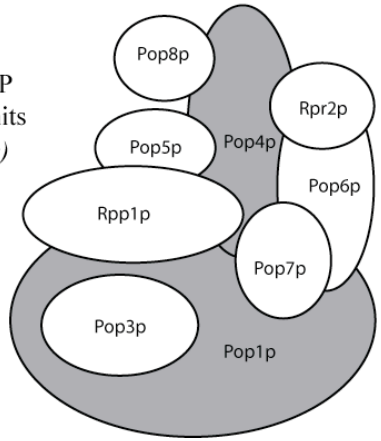


Yeast
RNase P RNA
(*S. cerevisiae*)

Bacterial RNase P
protein subunit



Yeast RNase P
protein subunits
(*S. cerevisiae*)



eukaryotic RNase P proteins is to ensure proper folding of the RNA subunit [31], but other functions remain unclear.

Yeast RNase P substrates

Although even the simple bacterial enzyme is known to cleave multiple non-pre-tRNA substrates, little is known about the diversity of substrates for eukaryotic RNase P. The huge increase in the protein content of the holoenzyme compared to the bacterial enzyme suggests the potential for a large increase in the variety of substrates that may be recognized. This would be akin to the increased diversity of promoters recognized by the structurally complex eukaryotic RNA polymerases, compared to the bacterial RNA polymerase, which contains a similar catalytic core. In yeast, the only defined set of RNase P substrates is the pre-tRNAs; no additional non-pre-tRNA substrates have yet been identified. In mammalian systems, even the set of pre-tRNA substrates has never previously been defined.

There are 274 tRNA genes in the yeast genome, encoding a set of pre-tRNAs that are more diverse than in bacteria. In addition to the diversity of tRNA sequence, in yeast the 3'-CCA sequences are not encoded as they are in bacteria, and some yeast tRNAs also contain introns. Introns are found in 61 yeast tRNA genes, corresponding to nine yeast tRNA families: Ile1, Leu1, Leu3, Lys2, Phe, Pro1, Ser3, Ser5, Trp, Tyr. The size of yeast introns ranges from 14-60nt long. In most cases RNase P removes the 5' leader before the intron and 3' trailing sequences are removed, indicating RNase P must be able to accommodate both intron-containing and intron-less tRNAs. Indeed, it has been

suggested that the presence of introns in some tRNAs facilitates their correct tertiary folds, allowing increased primary sequence diversity without compromising recognition by processing enzymes that depend on tertiary structure [32].

Although no physiological non-pre-tRNA substrate has yet been identified, a recent report has indicated that the non-coding RNA HRA1 is an *in vitro* substrate for the glycerol gradient fraction that includes RNase P [33]. Another recent study used microarray analysis to examine the RNAs affected when Rpp1p, one of the subunits common to both RNase P and RNase MRP, is depleted. There was an effect on the general mRNA population, but no specific RNA substrates were determined [34]. The authors did identify 74 transcripts, all from intergenic and antisense regions of the genome, that accumulate with Rpp1p depletion suggesting either RNases P or MRP might regulate these RNAs.

Multiple characteristics of the eukaryotic enzyme suggest that nuclear RNase P has additional substrates. First, seven of the nine protein subunits in yeast nuclear RNase P are highly positively charged (pI 9.3-10.0), which could provide multiple binding sites for negatively charged RNAs. This is consistent with previous studies on substrate binding kinetics that suggest eukaryotic RNase P has at least two RNA binding sites [35]. Second, eukaryotic RNase P is 1,000-times more susceptible to inhibition by single stranded homoribopolymers than bacterial RNase P [35], suggesting that the holoenzyme strongly binds single stranded RNAs in ways that inhibit pre-tRNA recognition. Experiments have shown potent, sequence-specific inhibition of yeast nuclear RNase P

by poly-U and poly-G RNAs, even greater than inhibition by pre-tRNAs (K_i <10 nM for poly-A and poly-G, compared to >20 nM for pre-tRNA) [35]. This demonstrates RNase P's ability to bind single stranded RNAs is at least as good as its ability to bind pre-tRNAs. Finally, a temperature sensitive mutation ($S_{827}S_{829}$) in one of the shared subunits between RNase P and RNase MRP, Pop1p, leads to cell death without affecting either pre-tRNA or 5.8S rRNA processing [36]. This indicates that the mutation has a lethal effect on an as yet undefined target. These three points, in addition to the precedence of the multiple substrates of bacterial RNase P and yeast RNase MRP, strongly suggests the possibility of non-tRNA substrates for eukaryotic RNase P.

The work reported in chapter II details our search for additional tRNA substrates for the nuclear form of yeast RNase P. First, we examine which RNAs physically associate with RNase P by identifying RNAs that copurify with the enzyme, using whole genome microarray analysis. We then ask which RNAs increase in abundance in yeast strains with mutations in different subunits of RNase P, also utilizing microarrays. Finally, we use northern blot analysis to examine multiple potential substrates for processing defects in the temperature sensitive mutant strains. Through this comprehensive approach, we identify numerous mRNAs that are both physically associated with RNase P and accumulate in the mutant strains. Furthermore, we identify the family of box C/D intron-encoded small nucleolar RNAs (snoRNAs) as physically associating with RNase P and accumulating a processing intermediate due to RNase P mutations. Focused examination of this group of snoRNAs has shown that RNase P is likely to participate in this processing pathway.

Mammalian RNase P substrates

In mammalian systems, it is surprisingly not known which tRNAs are actually expressed. There simply is no comprehensive analysis of expressed tRNAs of the type that has been performed for yeast. Presently, there are only 11 mouse tRNA sequences verified as expressed by direct sequencing of tRNAs [37]. Although there is a database of tRNA genes predicted by tRNAscan-SE available for all completed genomes [38], so far there has been little confirmation of the predictive power of tRNA scanning programs in the mammalian genome. In order to study mammalian RNase P and, more broadly, tRNA biogenesis, the set of actual tRNA genes is necessary.

The most commonly used tRNA scanning program is tRNAscan-SE, which employs both heuristic algorithms and covariance models. Initial tRNA gene candidates are first identified by either tRNAscan, identifying the A and B box promoters [39] and cloverleaf structure, or the Pavesi algorithm, identifying tRNAs on promoter and terminator sequences independent of a predicted secondary structure [39, 40]. tRNAscan-SE feeds the initial predictions into a third program that ranks the prediction based on a covariance model, a probabilistic model that describes both the primary sequence and secondary structure of tRNAs [41].

Recently a second program known as ARAGORN [38, 42, 43] was developed, which also scans genomes predicting probable tRNA genes. ARAGORN identifies candidate tRNA genes with a heuristic algorithm exclusively, identifying portions of the B box

sequence and then attempting to construct a cloverleaf with the neighboring sequences. In each case, the features common to all tRNAs are the characteristic "cloverleaf" secondary structure (Figure 1) and very limited patches of sequence conservation, termed the A box and B box, used as common recognition elements in both the transcription of the genes and structural and recognition elements in the tRNA transcripts.

There are many complicating factors that make identifying tRNA genes in mammalian genomes an especially difficult process. First, tRNA genes are short (<100nt) with little sequence homology, with the exception of the 11 nucleotide A box and 12 nucleotide B box, as described above. Second, mammalian genomes contain many highly repetetive tRNA-derived elements, known as Short Interspersed Elements (SINEs). Four out of the five abundant SINEs in mouse are derived from tRNA genes [44]. The most abundant family of SINEs in the mouse genome is the B2 family, with almost 100,000 elements [44]. Not only are these SINEs tRNA-derived, but they still contain the A and B box promoters, the two largest patches of sequence conservation in authentic tRNA genes. For unknown reasons, these SINEs are not expressed under normal conditions even though the tRNA-derived internal promoters allow transcription in vitro. These factors add to the difficulty of identifying true tRNA genes.

Chapter III describes our effort to define the full set of tRNA genes in the mouse genome, through a combination of computer prediction and experimental verification. As expected, over 80% of the original output of the tRNA scanning programs, both ARAGORN and tRNAscan-SE, were actually SINE elements. However, after removing

the SINEs, we predicted and verified the expression of tRNAs corresponding to 446 genes, of which 423 were sorted into 35 tRNA gene families based on sequence homology. The expression of all 35 tRNA gene families was confirmed by both microarray and northern blot analysis, and represents the first comprehensive index of expressed mouse tRNA genes.

Acknowledgements

Figures 1.1, 1.2, and 1.3 were based on figures prepared by Scott Walker and Shaohua Xiao.

CHAPTER II

Genome-wide search reveals nuclear RNase P is involved in maturation of intron-encoded box C/D small nucleolar RNAs

Introduction

Ribonuclease P (RNase P) is a conserved endoribonuclease responsible for removing the 5' leader sequence from precursor transfer RNAs (pre-tRNAs) found in Bacteria, Archaea, Eukarya [1, 45]. In all cases, with the possible exception of some organelles, RNase P is composed of both RNA and protein subunits. Bacterial RNase P is the simplest form of the holoenzyme, with one large RNA subunit and a single small protein subunit [1]. Although the RNA subunit of bacterial RNase P is sufficient for catalysis *in vitro* at high salt concentrations [3], both the RNA and protein subunits are required *in vivo*. The protein subunit appears to stabilize the catalytically active conformation of RNase P RNA, and assist with substrate binding [4-6, 46]. In addition to pre-tRNAs, bacterial RNase P is known to process several substrates that are proposed to contain tRNA-like structures. These bacterial substrates include 4.5S RNA, tmRNA, viral RNAs, mRNAs, riboswitches, ColE1 replication origin control RNAs, and C4 antisense RNA from phages P1 and P7 [13-21]. The presence of the protein subunit in the RNase P holoenzyme increases the substrate versatility of the enzyme over the RNA enzyme alone [7].

The eukaryotic nuclear RNase P is much more complicated. First, there are two very similar enzymes that are related to bacterial RNase P, termed RNase P and RNase MRP.

Yeast RNase P is responsible for processing pre-tRNAs, and RNase MRP processes pre-rRNA, mitochondrial RNA primers and is required for the regulated turnover of a cell cycle mRNA [22, 23, 47, 48]. Both the eukaryotic RNase P and RNase MRP from nuclei are far more complex enzymes than bacterial RNase P [24]. Each enzyme still employs a distinctive, but related RNA subunit, and contains multiple required protein subunits for function *in vivo*. In yeast the two enzymes have eight identical proteins subunits, with RNase P having one unique protein and RNase MRP having two unique proteins [26, 49]. Seven of the nine RNase P proteins are highly positively charged (pI 9.3-10.0), which could provide multiple substrate RNA binding sites in addition to the ones for pre-tRNAs that would be analogous to the bacterial enzyme. This might explain why yeast RNase P is much more susceptible to inhibition by single-stranded RNAs than bacterial RNase P [35] – the additional protein components might provide the ability to hold other types of RNA in position to occupy the active cleavage site provided by the conserved, catalytic RNA subunit. Thus, given the number of non-pre-tRNA substrates cleaved by even the bacterial enzymes, it seems likely that nuclear RNase P has been incorporated into the processing pathways for a number of different RNAs. Previous studies of eukaryotic enzymes have suggested this [33, 34], and there is substantial evidence that the closely related RNase MRP participates in regulated turnover of specific mRNAs [22, 23].

To search for physiologically relevant, novel substrates for nuclear RNase P, we used three different approaches in *Saccharomyces cerevisiae* (Figure 1). In the first, the RNase P holoenzyme was affinity purified and RNAs that copurify with the enzyme were identified using a whole genome microarray. The second and third approach utilizes

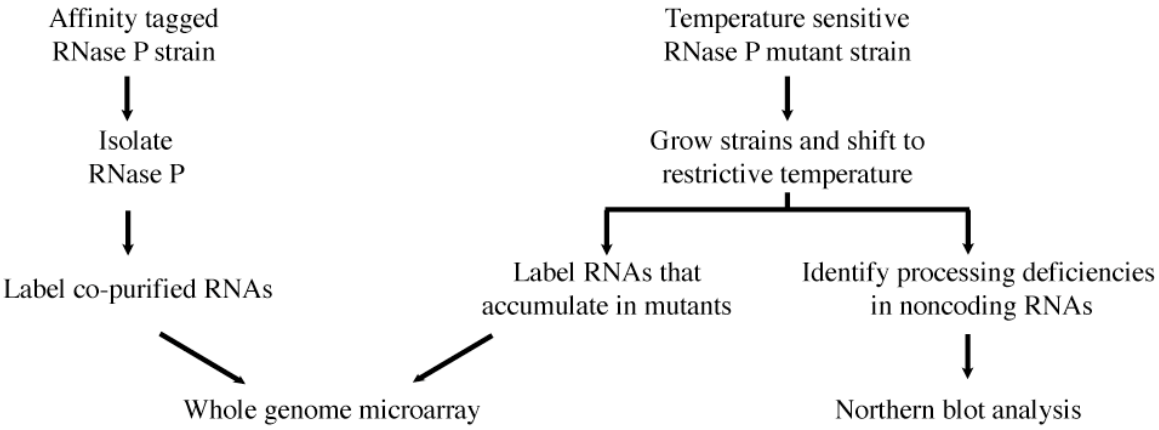


Figure 2.1. Multipronged approach to identify additional RNase P substrates

Three distinct approaches were taken to discover novel *in vivo* substrates for yeast RNase P. RNAs that physically associate with RNase P were identified by copurification with affinity-tagged holoenzyme. Functional relationships were identified in temperature sensitive mutant strains by examining changes in abundance by microarray or accumulation of aberrant-size processing products by northern blot.

temperature sensitive (ts) RNase P mutant strains. In the second approach, multiple temperature-sensitive (ts) mutant strains in multiple RNase P subunits were grown at the restrictive temperature and changes in the abundance of individual RNAs were measured using a whole genome microarray. In the third approach, we examined the processing of possible small RNA substrates in ts mutant strains by northern blot analysis to detect RNAs of altered size that accumulate in the absence of RNase P activity, even though they might not change in abundance. Here we report that this multipronged approach identified numerous potential substrates, and we focus on characterization of a particular class of RNAs that both copurify with RNase P and accumulate larger forms in the RNase P temperatue-sensitive mutants. This class is the set of box C/D small nucleolar RNAs (snoRNAs) that are encoded in the introns of six pre-mRNA introns. It was previously known that two pathways existed for excising these snoRNAs, one using the pre-mRNA splicing path and an other that was independent of splicing [50]. RNase P appears to participate in the splicing-independent path.

Results

Identifying RNAs that copurify with RNase P

Potential RNase P substrates were determined by identifying RNAs that co-purify with RNase P. RNase P was affinity purified using either a small RNA affinity tag (aptamer) incorporated into the RNA subunit (Rpr1r) that binds to streptavidin [51, 52], or a tandem affinity purification (TAP) tag [53] on the protein subunit that is unique to RNase P, Rpr2p. Strains expressing wild type RNase P, untagged, were subjected to the same

purification steps in order to establish a background for RNA contaminants in the purification process. The co-isolated RNA was then reverse transcribed into fluorescently labeled cDNA. The labeled cDNA was used to probe a microarray containing oligos to the entire yeast genome: open reading frames (ORFs), known non-coding RNAs, and intergenic regions [54, 55].

Comparison of results from independent purifications indicated that the enrichment values from the RNA subunit tag (streptavidin aptamer) purification (R^2=0.708) were much more consistent than the protein subunit (Rpr2p) TAP purification (R^2=0.196). The RNA aptamer and TAP purifications are single and dual column purifications, respectively. It might be expected that more transient interactions would be lost during the more protracted dual column purification (Rpr1p TAP tag). Additionally, the RNA aptamer purification is done under physiologic buffer conditions and eluted with only the addition of a small molecule (biotin). Therefore, we focused on data obtained from the RNA subunit tag purification.

Numerous RNAs were detected as co-purifying with RNase P. The 250 most abundantly co-purified RNAs were predominantly mRNAs involved in translation, although the microarray probes are double-stranded so the possibility of an antisense transcripts can not be ruled out (Table 2.1, full listing in Appendix A). The prevalence of these mRNAs in the co-purification does not correlate to their abundance in the cell (R^2=0.263), referenced to the yeast transcriptome [56], consistent with selective association with RNase P. The correlation drops even further (R^2=0.125) when limited to the 250 most

abundantly co-purified RNAs. It is interesting to note that tRNAs are not identified in this isolation, possibly due to the relatively transient binding of tRNA substrates and products to RNase P.

Identifying RNAs that accumulate in temperature sensitive mutants

The next approach to identifying novel RNase P substrates was to identify RNAs that change in abundance in the ts RNase P mutant strains. Temperature sensitive RNase P mutations were available in two subunits of yeast RNase P: the unique RNA subunit, Rpr1r [57, 58], and the largest protein subunit, Pop1p, that is also a component of RNase MRP [36].

The RNAs affected in the temperature sensitive strains are vastly different between *RPR1* ts and the two *POP1* ts strains. This could be due to the dual role of Pop1p in RNases P and MRP, or the different time courses of the temperature shift (2 hours to see growth inhibition in *RPR1* ts compared to 6 hours for *POP1* ts). However, there is an interesting general preference for the RNAs that co-isolate with the RNase P and accumulate in response to the ts mutations, in that they tend to be components of the translation machinery out of proportion to the abundance of the RNAs in the cell. This carries through when considering only the RNAs that both co-isolate and accumulate in response to ts mutation. Of these RNAs, 16 are mRNAs encoding protein subunits of the ribosome (Table 2). The remaining RNAs that both copurify with RNase P and accumulate in the mutant strain include mRNAs for two translation initiation factors (*TIF11*, *SUI13*), a box C/D snoRNA binding protein (*SNU13*), a common subunit in RNA polymerases I, II, and

III (*RPO26*), the CUP1-1 / RUF5 locus, and 6 intergenic regions. Signal from three of the intergenic regions neighboring ribosomal protein genes (*RPL42B. RPL41A, RPL38*) was also identified, although the signal from the coding regions of the genes themselves was not found and no characterized RNA is made from these regions. We note that pre-tRNAs do not accumulate substantially in this microarray analysis, but this is not unexpected in that the amount of uncut pre-tRNAs that accumulates before the cells stop growing is small compared to the stable population of mature tRNAs.

Intron-encoded snoRNAs

In the top 250 co-purified RNAs, we observed 7 intron-encoded small nucleolar RNAs (snoRNAs, Table 2.1). There are 75 known snoRNAs in the yeast genome. The majority of the snoRNAs in yeast are independently transcribed, however there are also examples of polycistronic transcript and intron encoded snoRNAs [59]. There are 8 total intron-encoded snoRNAs, six of which are box C/D. All six box C/D snoRNAs were in the top 250 RNAs co-purifying with RNase P. Of the two box H/ACA snoRNAs, only the snR44/*RPS22B* locus was in the top 250 co-purifying RNAs, while snR191/*NOG2* was the 319th ranked RNA.

Box C/D intron encoded snoRNAs accumulate known processing intermediate in an RNase P temperature sensitive mutant

As part of our screen for possible substrates for RNase P, we performed northern blot analysis on RNA from ts strains to see if some processing intermediates of some small

Table 2.1. Nuclear-encoded RNAs that co-purify with RNase P[a]

RNAs that copurify with RNase P RNA affinity tag (from top 250)	
Ribosomal small subunit mRNAs	***ASC1***[b], *RPS0A, RPS0B, RPS1A, RPS1B, RPS2, RPS3, RPS4A, RPS4B, RPS5, RPS6A, RPS7A, RPS7B, RPS8A, RPS9A, RPS10A, RPS10B, RPS11A, RPS11B, RPS12, RPS13, RPS14A, RPS15, RPS16B, RPS17A, RPS17B, RPS18A, RPS18B, RPS19A, RPS19B, RPS20, RPS21A, RPS21B, RPS22B, RPS23A, RPS23B, RPS24A, RPS24B, RPS25A, RPS25B, RPS27A, RPS27B, RPS28A, RPS28B, RPS29A, RPS29B, RPS30A, RPS31*
Ribosomal large subunit mRNAs	*RPL1A, RPL1B, RPP1B, RPL2B, RPP2A, RPP2B, RPL3, RPL4B, RPL5, RPL6A, RPL6B,* ***RPL7A***[b], ***RPL7B***[b], *RPL8A, RPL9A, RPL9B, RPL11A, RPL11B, RPL12A, RPL13A, RPL13B, RPL14A, RPL14B, RPL15B, RPL16B, RPL17B, RPL18A, RPL18B, RPL19A, RPL19B, RPL20A, RPL20B, RPL21B, RPL22A, RPL23A, RPL23B, RPL24A, RPL24B, RPL26A, RPL26B, RPL27A, RPL27B, RPL28, RPL29, RPL30, RPL31A, RPL33A, RPL33B, RPL34A, RPL34B, RPL35B, RPL36A, RPL36B, RPL37B, RPL38, RPL40A, RPL40B, RPL41A, RPL41B, RPL42A, RPL42B, RPL43A*
RNA polymerase I subunit mRNAs	*RPB8, RPA135, RPC40, RPO26*[c]
Translation Initiation mRNAs	*PAB1, TIF1, TIF11, NIP1, TIF3*
Translation elongation mRNAs	***EFB1***[b], ***TEF4***[b], *EFT1*
Non-coding RNAs	*RUF5*
RNAs that accumulate in *RPR1* ts strain (from top 250)	
Ribosomal small subunit mRNAs	*RPS30A*[c], *RPS17A*[c], *RPS18B*[c], *RPS16A*, *RPS10B*[c], *RPS28A*[c], *RPS10A*[c]
Ribosomal large subunit mRNAs	*RPL31A*[c], *RPL41B*[c], *RPL41A*[c], *RPL37B*[c], *RPL34A*[c], *RPL30*[c], *RPL34B*[c], *RPL38*[c], *RPL36A*[c]
RNA polymerase I subunit mRNAs	*RPC10, RPA34, RPC19, RPO26*[c]
RNA polymerase III subunit mRNAs	*RPC10, RPC19, RPO26*
Mitochondrial Ribosome	*MRPL37, MRPL49, MRP49, MRPL3, MRPL24, MRPL40*
Non-coding RNAs	*RUF5*[c]
RNAs that accumulate in *POP1660D6* ts strain (from top 250)	
Ribosomal large subunit mRNAs	*RPL41A*[c], *RLP24, RPL38*[c]
RNA polymerase I subunit mRNAs	*RPC10, RPA34, RPC19*
Non-coding RNAs	*RUF5*[c]
RNAs that accumulate in *POP1233E11* ts strain (from top 250)	
Ribosomal large subunit mRNAs	*RPL41A*[c], *RPL27B*[c], *RPL37B*[c], *RPL40B*[c], *RPL38*[c], *RPL36A*[c], *RPL33B*[c]
Ribosomal small subunit mRNAs	*RPS28A*[c], *RPS31*[c], *RPS26A, RPS20*[c], *RPS27A*[c]
Non-coding RNAs	*RUF5*[c]
RNAs that accumulate larger forms (possible precursors) in RNase P ts strain *RPR1* ts	
Intron-encoded snoRNAs	*ASC1*[c], *EFB1*[c], *IMD4*[c], *RPL7A*[c], *RPL7B*[c], *TEF4*[c]

[a] Full listing of individual genes found in Appendix A, [b] Bold genes accumulate a larger form in RNase P ts strain, [c] Underlined gene products copurify with RNase P RNA affinity tag.

RNAs might accumulate, even though the overall amount of RNA from that transcription unit did not accumulate significantly. 79 different non-coding RNAs were examined by northern blot, representing all classes of small nuclear RNAs, small cytoplasmic RNA (*SCR1*) and box H/ACA and C/D small nucleolar RNAs from independently transcribed, poly-cistronic, and intron-encoded genes (Appendix C, D). Accumulation of pre-tRNAs for these ts mutations had previously been demonstrated [36, 58]. In most cases RNAs other than tRNAs were not observed to accumulate larger (or smaller) forms in the RNase P mutant. Due to the observed physical interaction between RNase P and the intron-encoded snoRNAs, we specifically examined the processing of all of the intron-encoded snoRNAs in an RNase P temperature sensitive (ts) mutants [36, 58]. The ts mutation in the RNA subunit of RNase P was used in this study, since it (unlike Pop1p) is unique to RNase P. Without exception, the box C/D intron-encoded snoRNAs accumulate a processing intermediate in the RNase P ts mutant that is larger than the mature snoRNA (Figure 2.2). No aberrant forms of the two box H/ACA intron encoded snoRNAs were observed.

Probing northern blots with oligonucleotides spaced at the indicated positions (Figure 2) showed that for each gene the accumulated RNA is a 5' extended pre-snoRNA, which contains the 5' exon, the intron on the 5' side of the snoRNA, and the snoRNA itself. The intron sequence on the 3' side of the snoRNA and 3' exon are not contained in the accumulated RNA. The 5' extended pre-snoRNA is an expected processing intermediate in the splicing independent intron-encoded snoRNA maturation pathway (Figure 3) [50, 60]. Intron-encoded snoRNAs have two maturation pathways. The primary pathway in

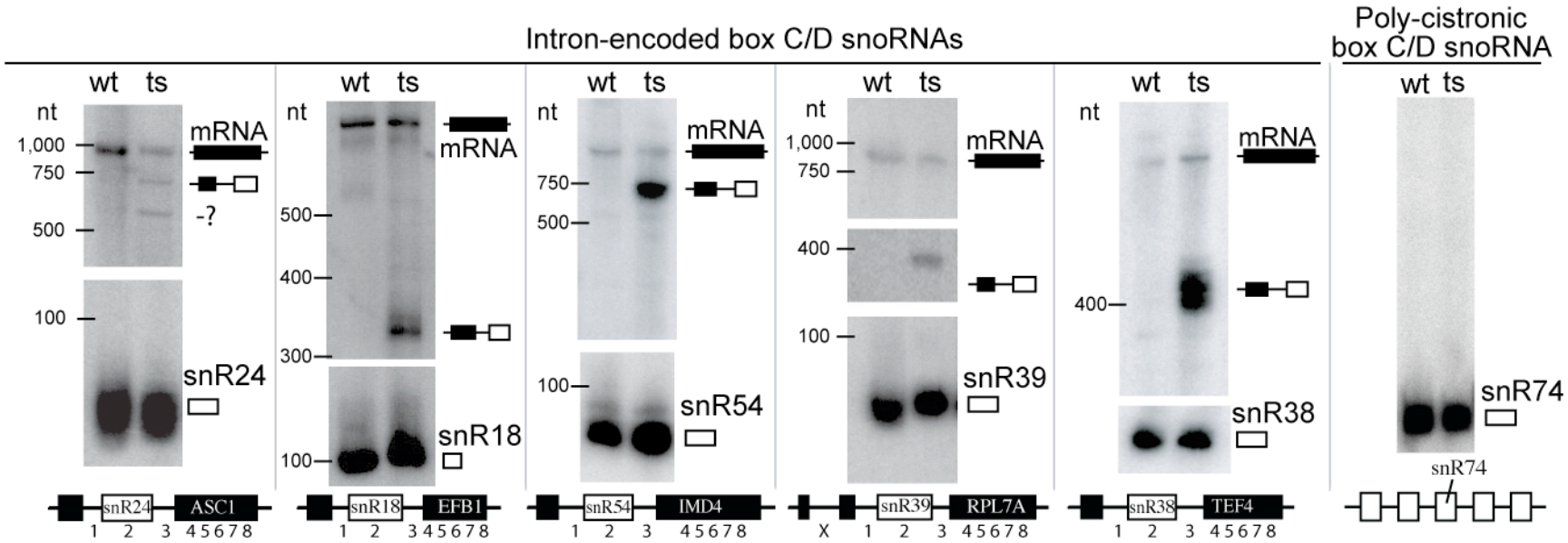


Figure 2.2. Box C/D intron encoded snoRNAs accumulate processing intermediate in temperature sensitive RNase P strain

Northern blots analysis of RNA from wild type (wt) and temperature sensitive (ts) RNase P mutants reveals an accumulation of an unusual processing form of the box C/D intron-encoded snoRNAs only in the RNase P deficient strain. Since RPL7A and RPL7B are highly homologous, only 7A and its snoRNA are shown. The identity of the accumulated transcripts was determined by which oligo probes detected them (Probes 1 and 2 hybridize to the processing intermediate, 3 and X did not give a signal, 4-8 hybridize to the mRNA, 2 hybridizes to the mature snoRNA), size, and primer extension to determine 5' termini. 79 noncoding RNAs were examined for altered forms by northern blot (Appendix C, D); the blot of a snoRNA from a polycistronic transcript is shown for contrast.

yeast involves splicing of the entire intron, followed by linearization by Dbr1p, and then release of the snoRNA by endonucleases and exonucleases. The minor maturation pathway in yeast is a splicing-independent pathway, where 3' and 5' endonucleases cut the pre-mRNA directly, leading to the destruction of the mRNA. The unusual 5'-extended pre-snoRNAs that accumulate in the RNase P mutant strain have already undergone 3' maturation of the snoRNA and require one or more 5' cleavages. It is worth noting that multiple attempts to delete *DBR1* in the presence of ts *RPR1* mutation were all unsuccessful (data not shown). This may be due to the elimination of the splicing-dependent snoRNA pathway in a strain where the RNase P-dependent pathway is weakened, which causes yeast to be inviable.

Although computer folding analysis of the pre-snoRNAs does not predict tRNA-like structures, we have previously shown that highly purified, yeast nuclear RNase P binds tightly to single stranded RNAs (Ziehler et al., 2000) and cuts at highly preferred sites in pre-rRNA that do not have obvious tRNA-like structures. We therefore tested to see whether the pre-snoRNA alone (in the absence of the snoRNA associated protein complex) was a highly selective substrate for yeast nuclear RNase P. Purified RNase P cuts T7-transcribed 5' extended pre-snoRNAs in multiple places *in vitro* (Figure 2.3). The major cleavage site in pre-snR14 is at the 5' end of the mature snoRNA, with a minor cut site 38nt upstream in the intron. In pre-snR18, the major cleavage site corresponds to ~43nt upstream of the 5' end of the mature snoRNA, which is 7nt upstream of a proposed stem essential for splicing independent snoRNA maturation [60]. RNase P makes multiple cuts into pre-snR38, the strongest is 58nt upstream of the mature

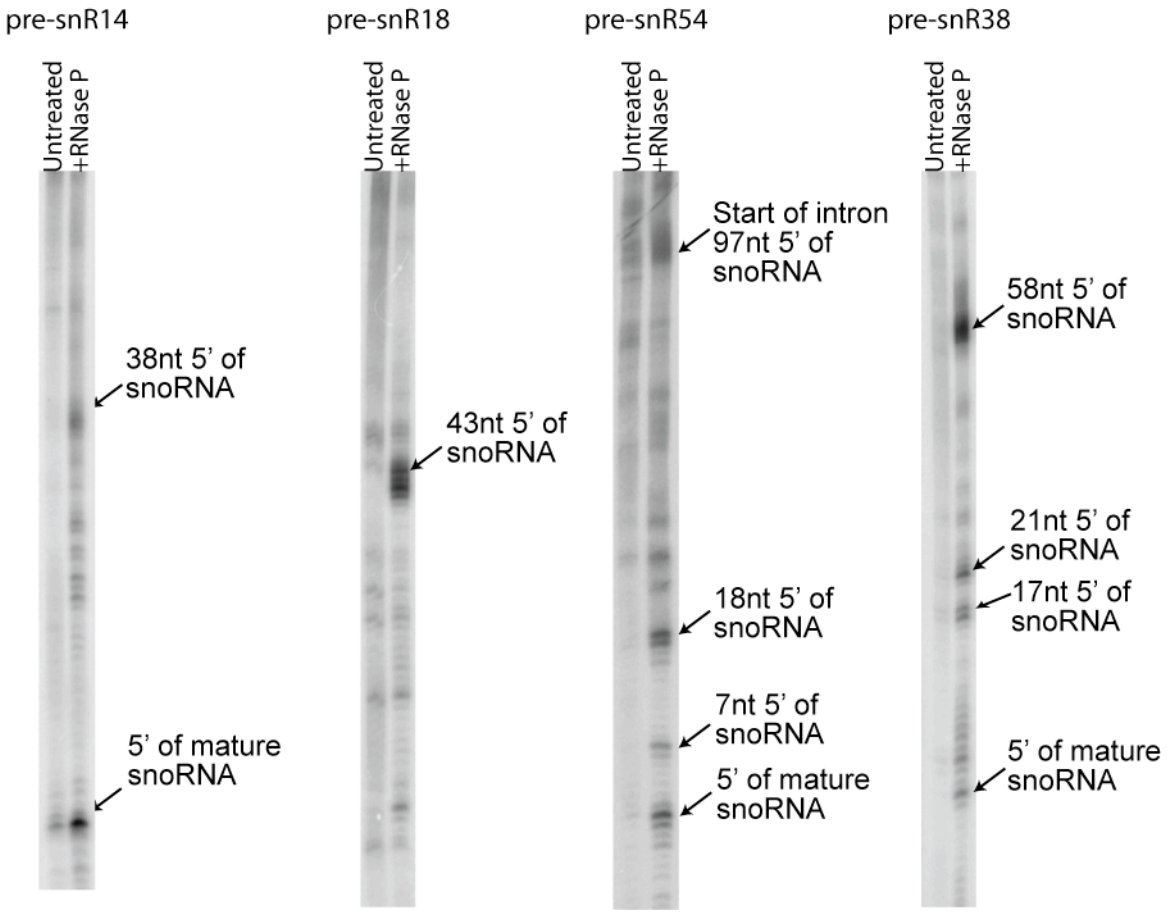


Figure 2.3. Primer extension of *in vitro* cleavage of 5' extended pre-snoRNAs by RNase P

5' extended pre-snoRNA transcripts were cut with affinity purified RNase P. Oligonucleotide primers complementary to the mature snoRNA were extended, which allows identification of a cleavage located on the 5' side of the snoRNA. The major cleavage sites were primarily in the intron on the 5' side of the mature snoRNA, although there was significant cleavage at the mature 5' end of pre-snR14.

snoRNA, with additional cuts at the 5' end of the mature snoRNA and 17nt and 21nt upstream of the 5' snoRNA site. Since the *in vivo* substrate is likely a ribonucleoprotein complex, we also tried to cleave accumulated RNPs in soluble extracts of the RNase P ts mutation after temperature shift by adding a large excess of purified RNase P to the extract under conditions where the longer form was not cleaved by any endogenous activity (4º or 37º). However, no preferred cleavages by RNase P were seen beyond slow degradation in the cellular extract (Figure 2.4).

Discussion

Numerous RNAs co-purify with nuclear RNase P and / or change in abundance in the temperature sensitive mutants. Recent studies have suggested various substrates for eukaryotic RNase P [33, 34]. While it has been shown that RNase P can cut the noncoding RNA *HRA1 in vitro*, we see no evidence of an *in vivo* association or function as HRA1 neither copurifies with RNase P nor does it accumulate in any of the three temperature sensitive mutants. Another recent study depleted one of the protein subunits found in both RNases P and MRP, and specifically identified 74 RNAs that accumulate with the Rpp1p depletion. Of the 74 noncoding RNAs identified, two copurify with RNase P, *MAN7* and *TLN1*, and another one, *TLN20*, accumulates in the RPR1 ts mutant strain (Appendix B). Interestingly, MAN7 and TLN1 are both antisense sequences from genes encoding ribosomal proteins.

We find that messenger RNAs for ribosomal proteins and other components of the translational apparatus are overwhelmingly the most abundant in co-purifying with both

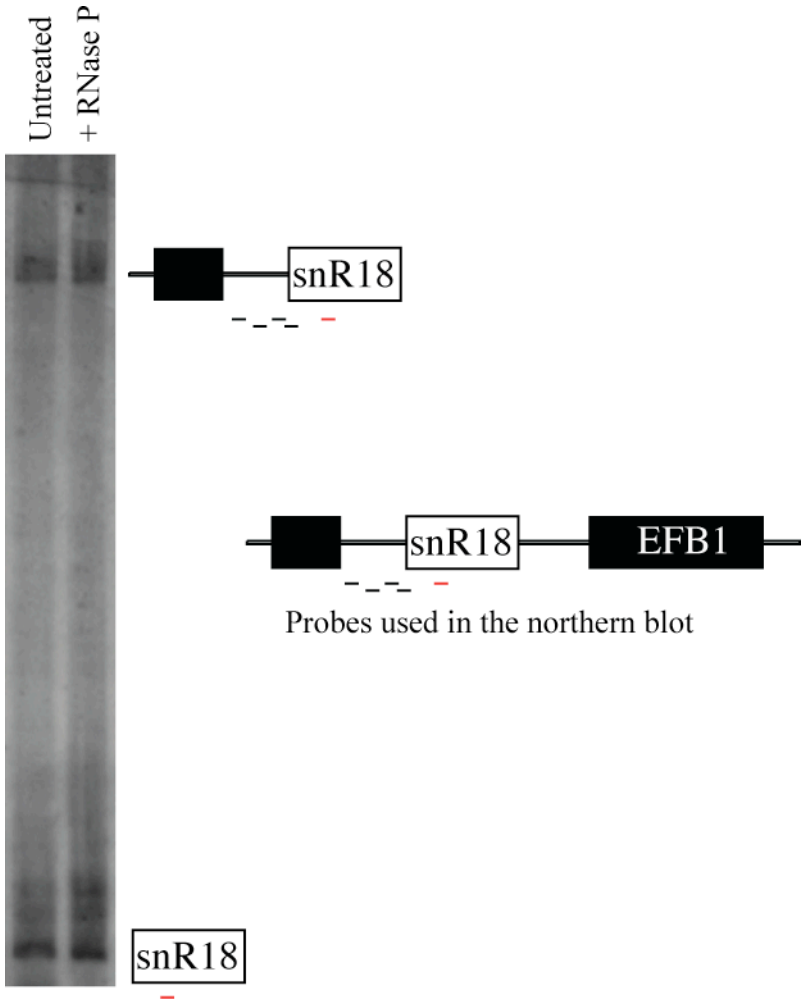


Figure 2.4. Adding wild type RNase P to extract made from RNase P mutant strain does not cleave 5' extended pre-snoRNA

In an attempt to accumulate physiological 5' extended pre-snoRNA with the appropriate snoRNP complex, cell extracts were made from temperature sensitive mutant *RPR1* strains grown at the restrictive temperature, 37°C. Purified wild type RNase P was added to the extract for 30 minutes at 4 °C, but no 5' extended pre-snoRNA appeared to be converted to a smaller form.

the RNA aptamer and the TAP tagged protein purifications; however, since the microarray probes are double stranded the signal could come from an unidentified transcript from the antisense strand as the case may be with *MAN7* and *TLN1*. Since the role of bacterial RNase P in pre-rRNA processing provides precedence for a role in mRNA turnover, it will be interesting to explore such a possible link to this set of mRNAs in the future. The existence of a possible link to mRNA turnover is also supported by the demonstrated participation of the highly similar enzyme, RNase MRP, in cell cycle-regulated turnover of specific mRNAs [22, 23]. Although the candidate mRNAs for RNase P are different, it is not surprising that the two enzymes, which differ by 1-2 protein subunits and have only related RNA subunits, would have developed differing substrate preferences.

Further investigations of the intron-encoded snoRNAs were pursued here, since they all copurify with RNase P and 5' extended pre-snoRNAs for each of the box C/D intron-encoded snoRNAs accumulated in the ts RNase P strains. Although the abundance of the pre-mRNAs from this pathway did not increase significantly in the RNase P ts mutants, this is not unexpected for a maturation, rather than turnover, defect. The 5' extended pre-snoRNA is a known processing intermediate in the splicing independent intron-encoded snoRNA maturation pathway (Figure 2.5). This splicing-independent pathway requires endonucleolytic cuts at both 5' and 3' ends of the snoRNA, and leads to the destruction of the mRNA [50]. The pre-snoRNA that accumulates in the RNase P mutant strain has already been trimmed at the 3' of the snoRNA, but it still contains the full transcript at the 5' end of the snoRNA, including both the exon and intron. It is possible that RNase P

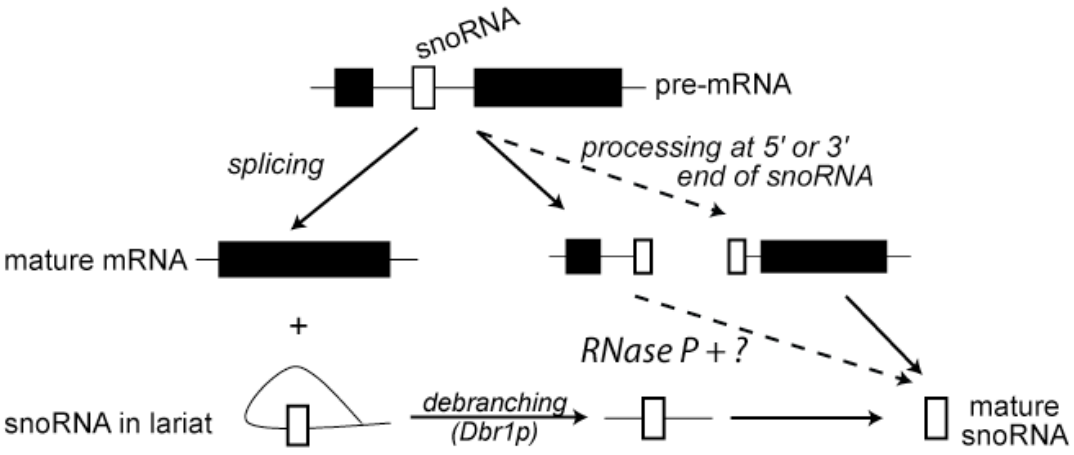


Figure 2.5. Intron-encoded snoRNA processing pathways

Two distinct processing pathways exist for intron-encoded snoRNAs. The splicing-dependent pathway produces the mature mRNA and snoRNA after the intron lariat form has been opened by Dbr1p and further processing. The splicing-independent pathway produces only the mature snoRNA. The dashed lines indicate the step affected by RNase P.

cuts at the 5' end of these snoRNAs in vivo, but it is also possible that it cuts somewhere upstream of the snoRNA (or at multiple places upstream) and the 5' maturation is subsequently performed by exonucleases. This would be similar to the case for 5' maturation of 5.8S rRNA by RNase MRP cleavage followed by further trimming.

In vitro cleavage assays with purified RNase P and pre-snoRNAs did not allow for a consistent model for the nature of the RNase P cleavage site in all examined intron-encoded snoRNAs. This leaves open the possibility that the RNase P effects on the intronic pre-snoRNAs is indirect. However, direct participation by RNase P is strongly suggested by the combination of 1) the physical interaction demonstrated by copurification of snoRNAs with RNase P, 2) the functional relationship seen by the accumulation of known processing intermediates in the RNase P mutant strain, and 3) the robust *in vitro* cleavage of pre-snoRNAs. It is quite possible that *in vitro* cleavage by RNase P is not sufficiently specific without an appropriate ribonucleoprotein (RNP) structure that is absent from the naked RNA and not preserved in the cellular extracts that have been tested so far from RNase P mutant strains. The analysis of the sequences and RNP structures that lead to RNase P recognition will presumably be an extended undertaking, especially since the highly complex 10-subunit RNase P RNP structure could hypothetically recognize a relatively large number of signals. It is also possible that RNase P makes the initial endonucleolytic cleavage in the intron on the 5' side of the snoRNA and then an exonuclease trims the remaining nucleotides back to the mature 5' end of the snoRNA. However, since *in vitro* reconstitution of snoRNPs is not possible at this time, we are unable to determine which of these is the case.

It is especially interesting that one of the intron-encoded snoRNA processing pathyways is compromised, since this suggests the pathway might be nucleolar. Not only is the final destination of the snoRNPs the nucleolus, but RNase P is also found primarily in the nucleolus in yeast [61]. Thus, RNase P might provide a link between production of both tRNAs and ribosomes, the two most abundant RNA components of the translational machinery. This link is strengthened by the identities of the host mRNAs of the intron-encoded snoRNA: seven of eight host mRNAs encoded proteins involved in translation.

Methods

Yeast Strains

Affinity purification: In a yeast strain containing a C-terminal tandem affinity purification (TAP) tag on RPR2 (Open Biosystems YSC1178-7501110), chromosomal RPR1 was disrupted with HIS3 and replaced with a plasmid, pRS315, containing RPR1 with RNA affinity tags for streptavidin and sephadex [51, 52].

Temperature sensitive mutations: The following TS, and respective WT, strains were used in this study: $G_{207}G_{211}$ RPR1 [62], R233K POP1, and R626L/P628K POP1[36].

Yeast Growth

Yeast were grown in standard synthetic media containing dextrose and lacking histidine (SDC-H). For temperature sensitive (TS) assays, yeast were grown at 30°C into log phase (OD_{600} of 0.6-0.8) and then diluted into SDC-H media pre-warmed to 37°C. The

strain-specific time period was determined from growth curves for wild type (WT) and TS RNase P strains, using the earliest time after the growth curve of the WT and TS strains diverged. $G_{207}G_{211}$ RPR1 was grown at 37°C for two hours and the POP1 strains were grown at 37°C for 6 hours.

RNase P Purification

Two yeast strains were subjected to RNase P purification: 1) Control strain that expresses wt RPR1 and wt RPR2 and 2) Tagged strain that expresses TAP tagged RPR2 and aptamer tagged RPR1.

RNase P was purified using either a single column aptamer affinity purification or a two column TAP purification. Sequential TAP then aptamer purifications were attempted but did not yield sufficient amounts of RNase P for analysis. Briefly, 8L of yeast were grown in YPD media to an OD600 of 0.8-1.0. Yeast were lysed in Lysis Buffer (50mM Hepes pH 7.5, 10% Glycerol, 0.5mM EDTA, 150mM NaCl, 1mM DTT, 0.1% NP40, cOmplete, EDTA-free protease inhibitor (Roche Diagnostics Corporation)) with a Microfluidizer, using five passes through a 200μm chamber and then five passes through a 100μm chamber. Cell extract was then cleared with an initial spin, 10 minutes at 10,000rpm, followed by a 1 hour 40 minute spin at 30,000rpm. 25μl of a 50% slurry of IgG Sepharose was added per 1ml of cell extract and incubated on a rotating drum at 4°C for 2 hours. The IgG Sepharose was then washed with 25x the column volume. RNase P was eluted overnight with the addition of 0.04μl (per ml of starting extract) tobacco etch virus (TEV) protease. The elution was adjusted to 2mM CaCl and bound for 2 hours to

25μl of a 50% slurry of Calmodulin Affinity Resin (CAR) per milliliter of starting extract. The CAR was washed five times with five volumes of Lysis Buffer (adjusted to 2mM CaCl). RNase P was eluted by adding five volumes of Lysis Buffer + 10mM EGTA.

Microarray Preparation

Associated RNAs were then reverse transcribed to cDNAs and fluorescently labeled with either Cy3 or Cy5 dyes. Labeled cDNAs were then hybridized to a yeast whole-genome microarray, which contains over 13,000 features corresponding to both known open reading frames (ORFs) and intergenic regions. The ratio of Cy3:Cy5 fluorescence indicates the relative amounts of RNA coming through the purification of the tagged to untagged yeast strains. The resulting data was analyzed as previously described [54].

Microarray Detection

RNAs were detected by microarray analysis as previously described [54]. Briefly, RNAs were reverse transcribed into cDNA in the presence of aminoallyl-dUTP using random nonamers as primers. The cDNA was then labeled with either Cy3 or Cy5 (Amersham Biosciences, Piscataway, NJ). Labeled cDNA was then hybridized to a yeast whole-genome microarray [63].

Northern Blotting of RNAs

Hot acid (pH 4.3) phenol [64] was used to extract total RNA from yeast cells harvested at 30°C and at 37°C and concentrations determined by UV absorbance. 10μg of total yeast

RNA per lane was electrophoresed on denaturing 8% polyacrylamide gels. The RNA was then electotransferred to a Nytran SuperCharge membrane (Schleicher & Schuell Bioscience).

Specific Oligodeoxynucleotide probes were designed to the majority of yeast small nuclear RNAs (snRNAs) and small nucleolar RNAs (snoRNAs). Probes were radiolabeled with γ-^{32}P-ATP. Labeled probes were hybridized and washed according to instructions accompanying the Nytran SuperCharge membrane. Signal on the Northern blots were determined with a PhosphorImager (Molecular Dynamics 445 SI).

In vitro Cleavage Reaction and Primer Extensions

PCR templates for T7 transcription were made using primers to the region 100nt 5' of the mRNA (in order to include any essential 5' untranslated region that may be structurally significant) and a primer complementary to the 3' end of the intron-encoded snoRNA. These templates were then used for T7 transcription. The *in vitro* transcribed RNA was gel purified on an 8% polyacrylamide gel. RNase P purified using the tandem affinity purification (TAP) tag on the unique protein subunit Rpr1p was added to the gel purified 5' extended pre-snoRNA in 1x RNase P Assay Buffer (10 mM HEPES, pH 7.9; 100 mM KCl; and 10 mM $MgCl_2$) and incubated for 15 minutes at 37°C. The reaction was then treated with proteinase K and the RNA was extracted using phenol-chloroform. Primers complementary to the snoRNA were kinased with γ-^{32}P-ATP, which was then gel isolated on a 12% polyacrylamide gel. The labeled primers were hybridized with the treated and untreated pre-snoRNA for 1 hour at 42°C and then extended using Superscript II reverse

transcriptase according to manufacturers instructions (Invitrogen). The cDNA was then electrophoresed onto an 8% polyacrylamide gel along with respective dideoxy sequencing ladders [65].

Acknowledgements

I would like to thank Felicia H Scott for her advice and assistance with the RNase P purification. Scott Walker also provided purified RNase P for the cleavage reaction and contributed the majority of work behind figure 2.4. The microarray analysis was done in collaboration with Jeff Pleiss in the laboratory of Christine Guthrie.

CHAPTER III

Prediction and verification of mouse tRNA gene families identifies intron-containing families with structures analogous those in yeast

Abstract

Computer algorithms are often used to identify tRNA genes in newly sequenced genomes. However, tRNA gene predictions are complicated by challenges such as structural variation, limited sequence conservation and the presence of highly reiterated short interspersed sequences (SINEs) that originally derived from tRNA genes or tRNA-like transcription units. To overcome this, we have employed two programs, tRNAScan-SE and ARAGORN, to predict the tRNA genes in the mouse nuclear genome, resulting in diverse but overlapping predicted gene sets. From these, we removed known SINE repeats and sorted the genes into predicted families and single-copy genes. In particular, four families of intron-containing tRNA genes were predicted for the first time in mouse, with introns in positions and structures analogous to the well characterized intron-containing tRNA genes in yeast. We verified the expression of the predicted tRNA genes by microarray analysis. We then confirmed the expression of appropriately sized RNA for the four intron-containing tRNA gene families, as well as the other 31 tRNA gene families creating an index of expression-verified mouse tRNAs. These represent all anticodons and all known mammalian tRNA structural groups, as well as a variety of tRNAs within families with altered ("rogue") anticodon identities.

Introduction

Transfer RNAs (tRNAs) are essential molecules responsible for decoding messenger RNAs (mRNAs) by delivering the proper amino acid into a growing peptide chain at the ribosome. Since a tRNA is required for each amino acid incorporated into every protein, tRNAs are one of the most abundant molecules in all living organisms. In order to make the large quantities of tRNAs needed, many tRNA genes appear to have been replicated in eukaryotic genomes through retrotransposition-like mechanisms. In some cases multiple copies of certain tRNA genes have been shown to be essential for a normal growth rate [66]. The tRNAs are duplicated by creating cDNA copies of the primary transcript, which include the internal promoter sequences, and the copies then re-insert at distant locations in the genome [37, 38]. Thus, the duplicated tRNA genes in yeast retain limited (<20 base pairs) conservation of upstream and downstream flanking sequences, as well as their intron sequences, when present. Introns are found in tRNAs in bacteria, archea, and eukarya, although the structure of the intron and the splicing process is specific to the domain of life [67]. In yeast, where the expression of individual gene copies has been verified, the tRNA genes flanking sequences and introns are not as tightly conserved as the mature coding regions, consistent with greater selection pressure for retaining the mature domains of the tRNAs intact.

The expression of tRNAs has only been thoroughly studied in a few bacterial, archaeal, and eukaryotic species. In eukaryotes, the synthesis, processing and utilization of tRNAs has been most extensively studied in the budding yeast, *Saccharomyces cerevisiae*.

These studies have yielded important information on tRNA synthesis by RNA polymerase III (pol III), post-transcriptional modifications, RNA transport, and gene organization, but there is a general lack of information on tRNA genes in mammalian genomes. For example, in mouse there are only 11 tRNA sequences verified as expressed in the tRNA database [37]. Although there is a database of tRNA genes predicted by tRNAscan-SE available [38], so far there has been little confirmation of the predictive power of tRNA scanning programs in the mammalian genome. This is an especially difficult analysis in most vertebrate genomes since they contain many tRNA-derived short interspersed elements, or SINEs [68]. Recent work from the Pan lab using microarray analysis confirmed the expression tRNAs corresponding to 374 human tRNA genes that were predicted by tRNAscan-SE [69, 70].

In an effort to comprehensively identify mouse tRNAs, we used two tRNA scanning programs, the commonly used tRNAscan-SE and more recently developed ARAGORN [38, 42, 43], to predict the probable tRNA genes in the mouse genome. tRNAscan-SE employs both heuristic algorithms and covariance models and is used extensively as the definitive tRNA gene identification program. Initial tRNA gene candidates are first identified by either tRNAscan, identifying the A and B box promoters and cloverleaf structure, or the Pavesi algorithm, identifying tRNAs on promoter and terminator sequences independent of a predicted secondary structure [39, 40]. tRNAscan-SE feeds the initial predictions into a third program that ranks the prediction based on a covariance model [41]. ARAGORN identifies candidate tRNA genes with a heuristic algorithm exclusively, identifying portions of the B box sequence and then attempting to construct a

cloverleaf with the neighboring sequences. In each case, the features common to all tRNAs are the characteristic "cloverleaf" secondary structures and very limited patches of sequence conservation, termed the A box and B box, used as common recognition elements in both the transcription of the genes and structural and recognition elements in the tRNA transcripts.

The predictions from scanning the *Mus musculus* genome [43] genome with tRNAscan-SE and ARAGORN were strikingly different, although there was significant overlap. After combining the predictions and removing known SINE sequences, we sorted the genes into families and singly-occurring ("orphan") genes based on sequence homology, and we experimentally verified expression of the predicted gene families. Mouse RNA from embryos and several tissues was hybridized onto a custom microarray with oligonucleotide probes tiled against each of the tRNA gene families and orphan tRNA genes to test expression. We also confirmed the expression of the tRNA families by northern blot, especially focusing on whether pre-tRNAs the size of those predicted for intron-containing genes were expressed. The results show that all of the predicted families and several orphan tRNA genes are expressed in all tissues tested.

Methods

Identifying potential tRNA genes

tRNA genes were predicted from the May 2004 release of the mouse genome using two publicly available computer programs: tRNAscan-SE [38] and ARAGORN [42]. The default settings were used with tRNAscan-SE, but ARAGORN was run with intron detection enabled. The resulting list of tRNA genes were merged based on genomic location. The tRNA genes were aligned by sequence homology using Clustal X [71] and then assigned to families based on high degrees of homology. A Clustal X alignment of the 5'-end of the predicted tRNA genes revealed many (1804) predicted tRNAs were homologous to the tRNA region of B2 SINEs. Comparison of predicted mouse tRNAs with the SINEs predicted by RepeatMasker, as annotated on the UCSC Genome Browser [72], identified additional B2 SINEs as well as Alu, B4, and ID SINEs. tRNA genes that did not have at least one similar sequence were not assigned to a family and are designated 'orphan' tRNA genes. The structure of intron containing tRNA gene sequences were predicted using Mfold [73] and then refined by hand.

Microarray design

Twenty probe sequences were allotted for each ncRNA prediction. Complementary DNA probes were designed to maximize spatial coverage of each predicted sequence while avoiding probes that have high self-folding potential as described previously [74] and were normalized by length (i.e. probe lengths were adjusted) to a uniform DNA-

RNA melting temperature of 70ºC. Probe sequences were on average 25.5 nt and were concatenated to 60 nucleotides. Probe sequences were submitted to Agilent Technologies for microarray production (Palo Alto, California). The designs included 1000 60-mer probes of random sequence, which were used as negative controls, and 696 positive control probes tiled across U4 and U5 snRNAs and 18S and 28S rRNAs. The design is accessible at NCBIs Gene Expression Omnibus (GEO, http://www.ncbi.nlm.nih.gov/geo/) database under platform accession GPL5420 [75].

RNA extraction, labeling, and hybridizations

Total RNAs from various mouse tissues were purchased from Ambion and Clontech. Integrity of rRNA was confirmed on 1% agarose-formaldehyde gels. 7 mg of total RNA was chemically labeled with Ulysis Alexa Fluor 546 or Ulysis Alexa Fluor 647 (Invitrogen) according to manufacturer's instructions. This protocol labels G residues [76], and there were no predicted RNAs that lacked G residues. Samples were resuspended in 0.5 mL of hybridization buffer (1 M NaCl, 0.5% sodium sarcosine, 50 mM N-morpholino ethane sulfonate, pH 6.5, 33% formamide and 40 mg salmon sperm DNA), denatured by heating at 65°C for 5 minutes, and snap-cooled on ice prior to hybridization. Hybridizations were carried out for 16-24 h at 42°C in a rotating hyb oven. Slides were then washed (rocking ~30 seconds in 6x SSPE, 0.005% sarcosine, then rocking ~30 seconds in 0.06x SSPE) and scanned with a 4000A microarray scanner (Axon Instruments, Union City, CA).

Microarray data processing and normalization

TIFF images were quantified with GenePix 3.0 (Axon Instruments, Union City, CA). Individual channels were spatially detrended (i.e. overall correlations between spot intensity and position on the slide removed) by high-pass filtering using 5% outliers. The individual channels were then normalized using Variance Stabilization [77] that allows for comparison across channels. All data is accessible at the GEO [75] database under series accession GSE8224.

Northern blots

10μg of RNA from each mouse tissue type was electrophoresed on an 8% polyacrlamide gel (SequaGel) and then electroblotted onto a Nytran SuperCharge membrane (Schleicher & Schuell Bioscience). Blots were probed with oligonucleotide probes to mature regions of the predicted tRNAs. In the case of genes predicted to have introns, intron probes were also used. Blots were reprobed for 5.8S ribosomal RNA (rRNA) as a loading control. The probe sequences can be found in online supplemental S_1. Blots were exposed to a phosphocapture screen, detected on a PhosphorImager (Molecular Dynamics 445 SI), and quantified with IPlab Gel software (Signal Analytics).

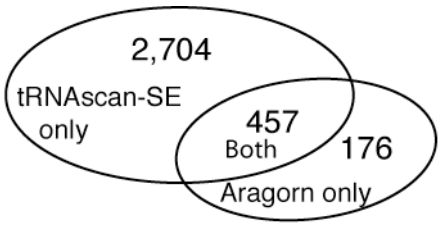
Results

Scanning the mouse genome for predicted tRNA genes

In an attempt to predict the functional tRNA genes in the mouse, both tRNAscan-SE and ARAGORN were used to scan the May 2004 release of the *Mus musculus* genome (Figure 3.1). tRNAscan-SE identified 3,161 putative tRNA genes and ARAGORN predicted 633 genes. Comparing the two sets, 457 of the putative tRNA genes identified by ARAGORN were also identified by tRNAscan-SE. Both programs identified all 11 of the verified mouse tRNA gene sequences and tRNA genes corresponding to all anticodons (Table 1). However, a large number of genes were predicted by only one of the programs, either tRNAscan-SE or ARAGORN, so we focused on removing potential artifacts from the set of tRNA gene predictions.

Removing SINEs from the predicted tRNA genes

One of the anticipated problems in this study was the abundance and variety of tRNA-derived SINE elements. tRNAscan-SE employs a scoring system to identify tRNA genes with predefined cutoff levels to distinguish a 'real gene' versus a pseudogene. tRNAscan-SE identified 22,027 sequences as pseudogenes, sequences that shared some features with tRNA genes but scored too low to be considered an actual tRNA gene. ARAGORN does not identify weak scoring sequences as pseudogenes. Examination of the pseudogene sequences identified by tRNAscan-SE revealed many to be homologous



	tRNAscan-SE only	Both	Aragorn only
tRNA genes identified	2,704	457	176
SINEs removed from prediction			
Alu	0	0	1
B2	1,978	31	95
B4	392	35	26
ID	265	15	25
Predicted tRNA genes	69	376	29
Total number of predicted tRNA genes		474	

Figure 3.1 – Different tRNAs and SINEs are identified by both scanning programs

tRNAscan-SE and ARAGORN identify different sequences as tRNA genes. However, the majority of the genes identified by a single program are tRNA-derived SINE elements.

to the B2 SINE consensus sequence. These pseudogenes were not considered in the rest of the analysis.

To determine if any B2 SINEs were classified as tRNA genes we used the B2 consensus sequence [44] and ran a BLAST search against tRNA genes predicted by both tRNAscan-SE and ARAGORN. We used the entire B2 consensus sequence, both upstream and downstream of the tRNA-like domain, to avoid identifying the functional tRNA family from which B2 SINEs are derived. The tRNA genes identified by both predictive programs were the least likely to be homologous to B2 SINEs, as only 29 sequences (6.3%) predicted by both tRNAscan-SE and ARAGORN were >70% homologous to the B2 SINE consensus sequence. 83 tRNA genes (47%) predicted only by ARAGORN and 1,806 tRNA genes (67%) predicted only by tRNAscan-SE were homologous to the B2 SINE consensus sequence. The remaining predicted tRNA genes were compared to SINEs annotated on the UCSC Genome Browser by RepeatMasker, which identified B1, B3, and ID repeat elements (Figure 3.1). Removing the SINEs from the predicted tRNA genes eliminated 97% of genes predicted by just tRNAscan-SE, 83% of genes predicted by just ARAGORN, and only17% of the tRNA genes predicted by both programs.

Assigning mouse tRNA genes into families

After removing B2 SINEs and merging the predicted tRNAs from tRNAscan-SE and ARAGORN, there are a total of 474 predicted tRNAs (Appendix E). These sequences were aligned using ClustalW and then manually sorted into families based on sequence similarity. 452 tRNAs were highly homologous to at least one other tRNA and were

sorted into 35 tRNA families (Table 1) representing all 20 essential amino acid anticodons. Genes encoding the 11 known unique mouse tRNAs from the tRNA database [37] are all included in the tRNA families. The number of genes per family ranged from 2 to 38. In the *Saccharomyces cerevisiae* genome the number of copies of a single tRNA gene range from 1-16 copies. Most of these yeast tRNA gene copies have identical sequences in the mature tRNA coding regions, but this is not true with the predicted mouse genes; in fact, the sequence similarity between members of mouse tRNA families is strikingly less than the similarity between yeast tRNA family members. Most of the tRNA gene copies in the mouse genome have multiple nucleotides different among family members, even though the majority of the tRNA gene sequence is conserved (Appendix E, G). However, there must still be pressure to maintain sequence identity in these mouse tRNA genes, since the introns and flanking sequences of the intron-containing tRNA families diverge much more rapidly.

Interestingly, there are substantial differences in the number of tRNAs that are charged with different amino acids. Since tRNA abundance has been shown to correlate with gene copy number [66], we asked whether the skewed distribution of tRNA genes within a family matched a bias in amino acid usage throughout the mouse genome. We compared the relative amino acid utilization (# specific amino acid in proteome / # total amino acids in proteome with tRNA type distribution (# specific tRNA types / total # tRNAs) but found no significant correlation between the two (data not shown).

Over half (19) of the tRNA gene families contain genes with a single anticodon sequence (Table 3.1). The remaining 15 tRNA families contain genes with different anticodon sequences, eight of which include either one or two 'rogue' tRNA genes that have an anticodon for a different amino acid than the majority of the family members (Appendix E - Ala1, Ala2, Glu1, Leu3, Lys1, Pro, Thr1, Val1). This was of concern, since "improper" charging of a tRNA relative to its anticodon would lead to the improper incorporation of an amino acid into a protein, which should be counter selected against. We confirmed the sequence of several of these genes by amplifying the genomic DNA locus by PCR and directly sequencing the PCR products (data not shown).

Comparisons between mouse and human tRNA gene families

We examined the predicted tRNA genes identified by tRNAscan-SE in the human genome to see if rogue tRNA genes existed within human tRNA gene families. There are 36 tRNA families in the human genome, based on sequence homology of the predicted human tRNA genes (Appendix F). 32 of predicted human tRNA families have a high degree of sequence homology with tRNA families in the mouse genome. The human Hs_Arg2, Hs_Leu3, Hs_Thr2, and Hs_Gln2 tRNA gene families do not have an identifiable homologous tRNA gene in the mouse genome. Eight of the predicted human tRNA gene families contain at least one rogue tRNA gene (Hs_Ala1, Hs_Arg1, Hs_Arg2, Hs_Cys, Hs_Glu, Hs_Gly1, Hs_Met2, Hs_Lys2), showing that the existence of rogue tRNAs is not unique to mouse. Neither the family that contains the rogue tRNA gene nor the rogue anticodons are conserved between the human and mouse genomes, indicating that the specific anticodon variants are not conserved in mammals.

Table 3.1. Description of mouse tRNA gene families

Gene family	# in family	Identified in Sprinzl Database[1]	# containing introns	Anticodons
Ala_1	27	X		TGC^{Ala}, CGC^{Ala}, AGC^{Ala}, TAC^{Val}
Ala_2	25			AGC^{Ala}, AAC^{Val}, ACC^{Gly}
Ala_3	8			AGC^{Ala}
Arg_1	14			ACG^{Arg}, TCG^{Arg}, CCG^{Arg}
Arg_2	11		5	TCT^{Arg}, CCT^{Arg}
Asn	14			GTT^{Asn}
Asp	14	X		GTC^{Asn}
Cys	38	X		GCA^{Cys}, ACA^{Cys}
Gln	18			CTG^{Gln}, TTG^{Gln}
Glu_1	16			CTC^{Glu}, TTC^{Glu}, CTT^{Lys}, CGC^{Ala}
Glu_2	6			TTC^{Glu}
Gly_1	17	X		GCC^{Gly}, CCC^{Gly}
Gly_2	6			TCC^{Gly}
Gly_3	2			CCC^{Gly}
His	9	X		GTG^{His}
Ile_1	12			AAT^{Ile}
Ile_2	4		3	TAT^{Ile}
Leu_1	12		4	CAG^{Leu}, CAA^{Leu}
Leu_2	9	X		AAG^{Leu}, TAG^{Leu}
Leu_3	5			TAA^{Leu}, TTA^{Gln}
Lys_1	38	X		CTT^{Lys}, TTT^{Lys}, TAA^{Lys}, TCA^{Sec}
Lys_2	11	X		TTT^{Lys}
Met_1	10	X		CAT^{Met}
Met_2	7			CAT^{Met}
Phe	8			GAA^{Phe}, GGA^{Phe}
Pro	4			AGG^{Pro}, GGG^{Pro}, TGG^{Pro}, AGG^{Leu}
Ser_1	14	X		AGA^{Ser}, CGA^{Ser}, TGA^{Ser}
Ser_2	8			GCT^{Ser}
Thr_1	10			AGT^{Thr}, CGT^{Thr}, TGT^{Thr}, CAT^{Met}
Thr_2	2			CGT^{Thr}
Trp	8			CCA^{Trp}
Tyr	12			GTA^{Tyr}
Val_1	21		12	CAC^{Val}, AAC^{Val}, ACC^{Gly}
Val_2	3			TAC^{Val}

Expression of predicted tRNAs

Custom microarrays were designed to test for the expression of the predicted tRNA genes. RNA from embryos and several different tissues were tested in case there was some substantially different pattern of expression for some families. Although most tRNAs are expressed constitutively, there are also examples where special tRNA families are transcribed in response to high demand for protein. In the silkworm *Bombyx mori*, an alanine tRNA is exclusively expressed in the silk gland and a glycine tRNA is overexpressed in the silk gland [78]. This tissue-specific tRNA expression allows production of the glycine and alanine-rich silk protein Fibroin. In *Xenopus*, an entire set of highly reiterated tRNA genes are transcribed during oogenesis as part of a process to store large quantities of translational machinery for the upcoming high protein production in developing embryos [79]. However, since the majority of tRNA studies have been performed on single cellular organisms, the possibility of tissue or developmentally regulated tRNAs in mammals is largely unexplored.

Total RNA from different developmental stages (7 day embryo, 10-12 day embryo) and different tissues (muscle, spleen, mammary gland, brain, ovary, thymus, liver, heart) was directly fluorescently labeled and used to probe the microarrays (see Methods). Probes unique to the predicted tRNA gene or gene family whose signal was greater than 99% of the negative controls were considered to be expressed. Using these criteria, all of the predicted tRNA gene families were detected as expressed by the microarray. However, because of the sequence homology between family members we can only conclude that some subset of the tRNA gene copies are being expressed. While no families are

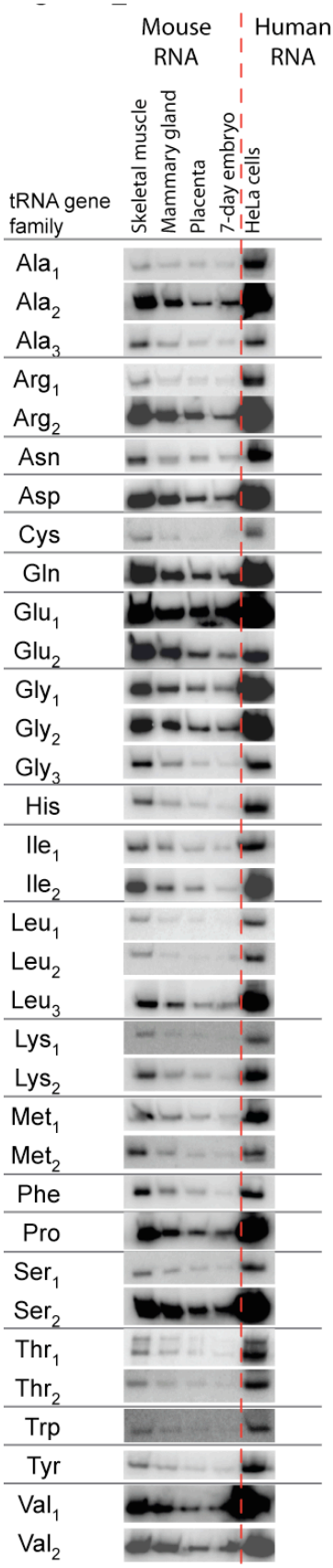
exclusively expressed in a particular developmental stage or tissue, there is a five-fold greater expression of tRNAs in brain and ovary relative to muscle and liver. Overall tRNA expression ranges from Brain > Ovary > Heart > 10-12 day embryo > Mammary Gland > Testis > Placenta > Thymus > 7 day embryo > Spleen > Skeletal Muscle > Liver.

Each tRNA family was probed by northern blot in a subset of mouse RNA samples to complement the microarray results, ensuring the signal detected was not due to cross hybridization of another type (size) RNA transcript. In particular, we focused on expression of intron-containing families in this work (see below), but confirmed the size and expression of all of the tRNA gene families in skeletal muscle, mammary gland, placenta, and 7-day embryo (Figure 3.2) as well as the presence of the homologous human tRNA family in RNA from HELA cells.

The microarray analysis identified 29 'orphan' tRNA genes that are also expressed in the various RNA samples. One of the 29 orphans tRNAs, tSeC(TCA)G, is a known selenocysteine tRNA, *Trsp [80]*. The microarray analysis detected the selenocysteine tRNA is expressed in all tissues types as well as in the 7 day and 10-12 day embryo, which is consistent with *Trsp* expression being essential for mouse embryogenesis [81]. In addition to selenocysteine tRNA, which has two homologues in the human genome, the tyrosine orphan tRNA, tY(GTA)B, has at least 20 homologues (>90% sequence identity) in the human genome. This tRNA gene was only identified by ARAGORN in the mouse genome and none of the human homologues are identified b tRNAscan-SE. The predicted structure of tY(GTA)B is very tRNA-like and the sequences

Figure 3.2. Northern blot confirmation of predicted tRNA gene families.

We confirmed the expression of tRNA-sized RNAs in RNA samples from four mouse tissues and RNA from actively growing human tissue culture cells (HeLa). The probe sequences are listed in Appendix H.


Mouse RNA
Human RNA
tRNA gene family
Skeletal muscle
Mammary gland
Placenta
7-day embryo
HeLa cells
Ala1
Ala2
Ala3
Arg1
Arg2
Asn
Asp
Cys
Gln
Glu1
Glu2
Gly1
Gly2
Gly3
His
Ile1
Ile2
Leu1
Leu2
Leu3
Lys1
Lys2
Met1
Met2
Phe
Pro
Ser1
Ser2
Thr1
Thr2
Trp
Tyr
Val1
Val2

are not tagged as a repetitive element by RepeatMasker in the mouse or human genomes. However, there is an oligo-T sequence (T_5) found in both mouse and human copies of the gene in the aminoacyl acceptor stem that would be predicted to result in early pol III transcription termination. However, all of the probes on the microarray that targeted the mature region of tY(GTA)B gave signal well above background (data available on GEO GSE8224).

Intron-containing tRNA genes confirmed by northern blot

Introns are found in four previously uncharacterized mouse tRNA gene families: Arg3, Tyr, Leu1, and Ile2 (Figure 3.3). Each of the homologous families in the human tRNA set also contains introns. There is also a single tRNA gene in the highly expanded human Pro family (4 genes in mouse *vs*. 21 genes in human) that contains an intron, which is not seen in the mouse Pro family. In *S. cerevisiae* there are introns in eight tRNA types: Phe, Ile, Lys, Leu, Pro, Ser, Trp, and Tyr. The presence of introns in tyrosine, leucine, and isoleucine tRNAs in organisms as divergent as yeast, mouse, and humans might indicate that some tRNAs are more tolerant of introns than other tRNA types. Alternatively, the introns might be functional, such as the yeast tyrosine tRNA where the presence of the intron is required for proper folding of the tRNA [32].

All of the genes in the mouse families Arg3, Ile2, and Tyr contain introns, while only four of the 12 tRNA genes in the Leu1 family contain introns. There appears to be little selective pressure to maintain the sequences of the introns in duplicated genes relative to the greater conservation of mature tRNA sequence (Figure 3.4), although the

Figure 3.3. Northern blot confirmation of intron-containing tRNA genes

Northern blot analysis confirms the expression of intron-containing tRNA gene families in all mouse tissues tested. Separate panels for precursor tRNAs and mature tRNAs allow for sufficient contrast of the precursors, since precursor tRNAs exist at a small fraction of the mature levels relative to more rapidly growing organisms. Precursor sizes were consistent with the predicted intron sizes. It is not known whether the multiple bands in the mature tRNA are due to variations in "mature" domain length of family members, incomplete removal of amino acids from the 3' termini, or trimming of the 3' CCA residues. While the expression levels vary between tissue types, the intron-containing tRNA gene families are ubiquitously expressed in all tissues tested.

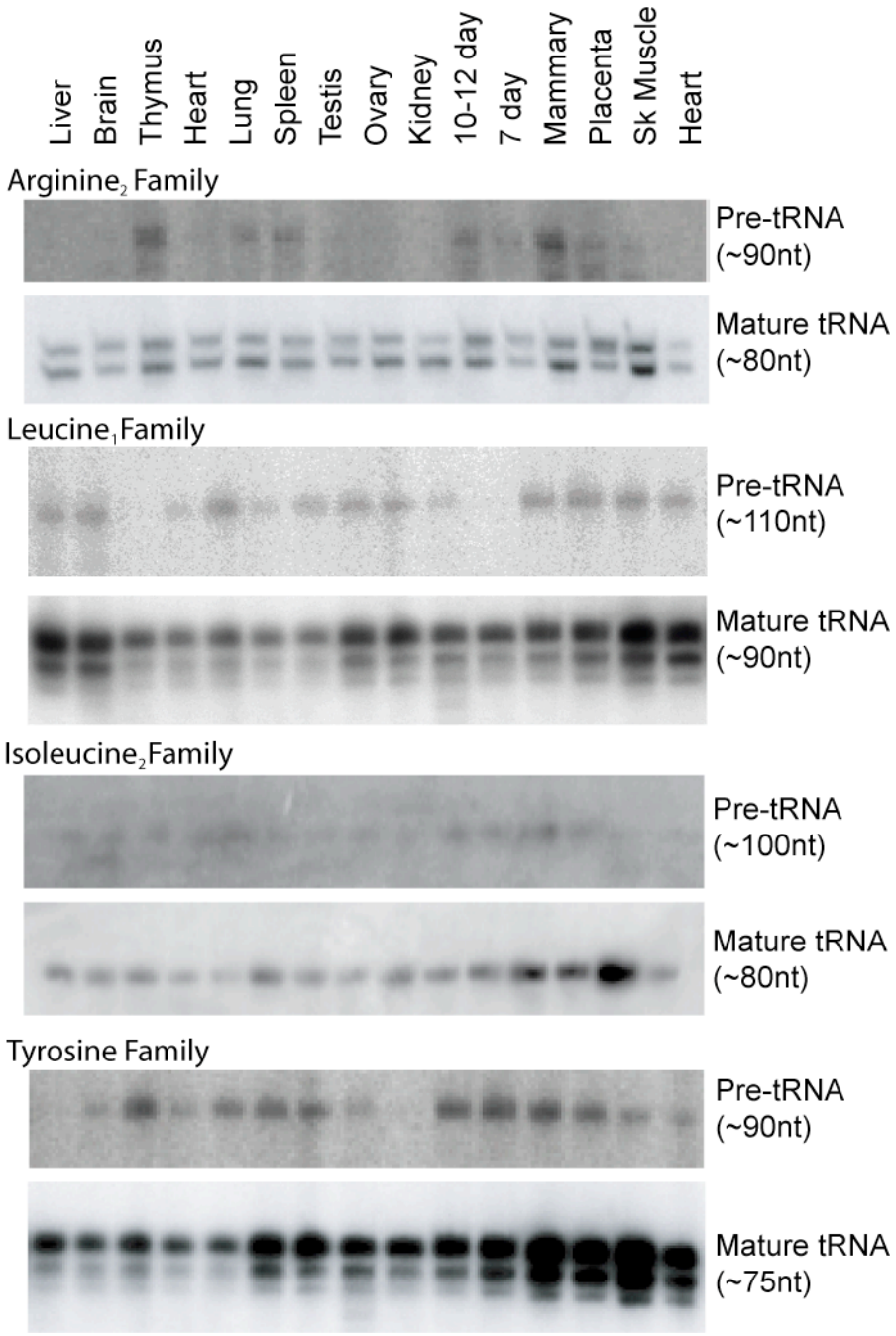
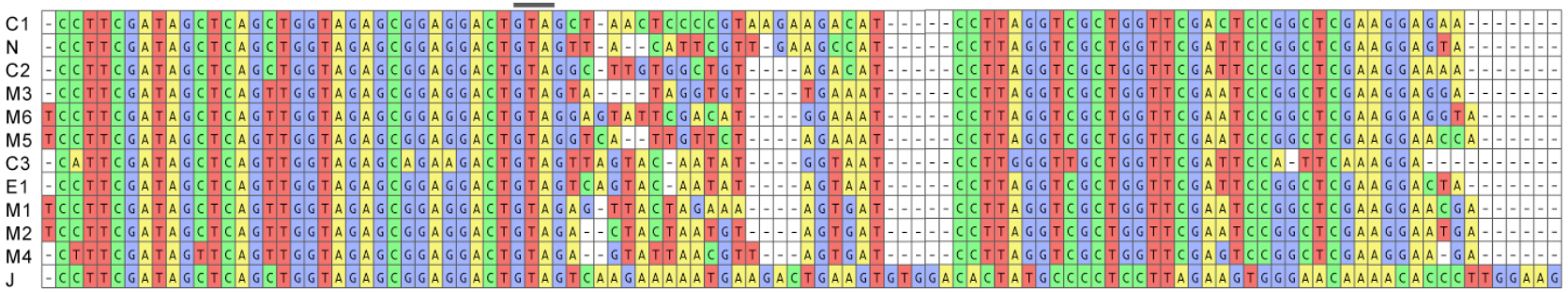

Liver
Brain
Thymus
Heart
Lung
Spleen
Testis
Ovary
Kidney
10-12 day
7 day
Mammary
Placenta
Sk Muscle
Heart
Arginine2 Family
Pre-tRNA (~90nt)
Mature tRNA (~80nt)
Leucine1 Family
Pre-tRNA (~110nt)
Mature tRNA (~90nt)
Isoleucine2 Family
Pre-tRNA (~100nt)
Mature tRNA (~80nt)
Tyrosine Family
Pre-tRNA (~90nt)
Mature tRNA (~75nt)

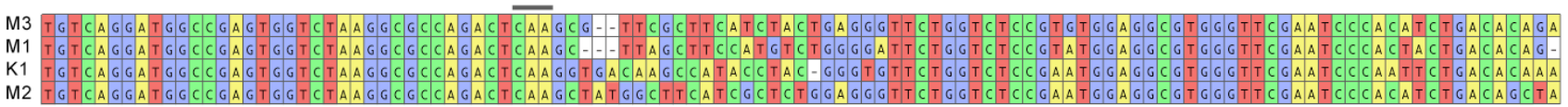
Figure 3.4. Sequence alignment of intron-containing tRNA genes

The four families of intron-containing genes are aligned. The solid line above the sequence indicates anticodon position and the introns are indicated with a dashed line. The letters and numbers (e.g., C1, N, C2 for Tyrosine family) correspond to the tRNA gene names in Appendix E.

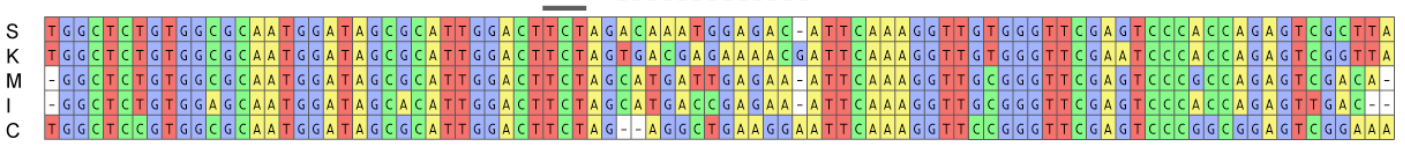
Tyrosine Family tRNAs (12 of 12 genes in family)



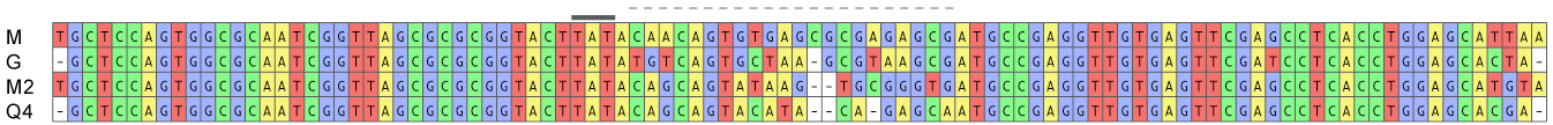
Leucine$_1$ Family intron containing tRNAs (4 of 12 genes in family)



Arginine$_2$ Family intron containing tRNAs (5 of 11 genes in family)



Isoleucine$_2$ Family tRNAs (4 of 4 genes in family)



intron sequences are more conserved than the sequences immediately upstream or downstream of the mature domain, predicted to be in the pre-tRNA primary transcript. The predicted structures of the intron containing tRNA genes are consistent with the intron location and structure in yeast (Figure 3.5) [82].

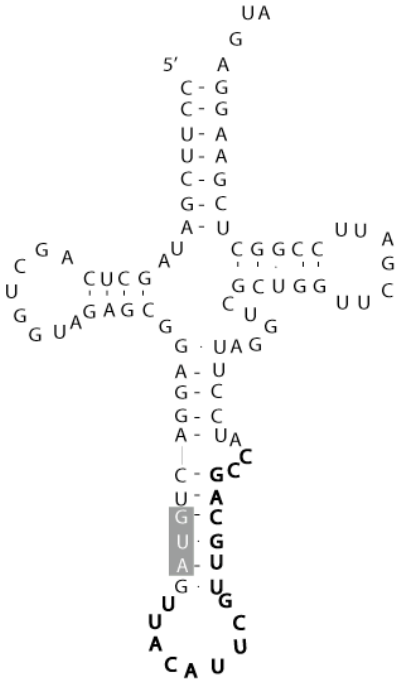
There are also 10 intron-containing orphan tRNAs that were detected as expressed by the microarray analysis. Multiple probes unique to the predicted tRNA sequence gave signal above the 99% confidence level based on the negative control probes. However, the predicted intron containing phenylalanine tRNA, tF(GAA)O, has a 478nt long intron and tA(GGC)O1 has a 2nt intron. Neither of these intron lengths are consistent with intron lengths or splicing mechanism in yeast (Figure 3.6). The intron in yeast isoleucine tRNA genes is 58nt, but the remaining yeast introns are all between 13-32nt long. The remaining intron-containing mouse orphans include: tA(AGC)Q4, tI(TAT)G, tI(TAT)M, *tL(CAA)K*, *tP(AGG)P*, *tT(TGT)E1*, tT(TGT)M2, and *tV(TAC)G*, the four genes that are predicted by only ARAGORN are indicated with italics (Appendix E).

Discussion

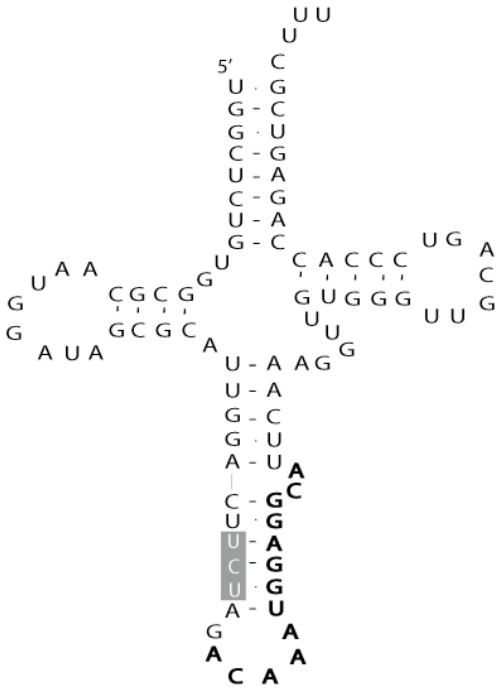
This work provides comparative analysis of *in silico* tRNA gene predictions using different algorithms, and experimental confirmation of the predicted tRNAs. It was necessary to first screen the predicted mouse tRNA genes for known SINE elements, since most of these highly repetitive sequences in most non-human vertebrates derive originally from tRNA genes. This eliminated 85% of the predictions by tRNAscan-SE and 36% from ARAGORN although it should be noted that most of the ~900,000 tRNA-

Figure 3.5. Predicted structure of intron-containing tRNA genes

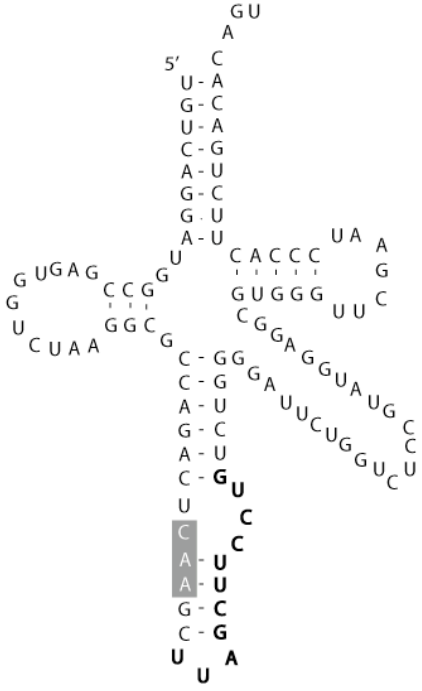
One representative structure is shown for each of the intron-containing tRNA gene families. The anticodon position and general helix-bulge-helix structure is consistent with yeast intron-containing tRNA structures



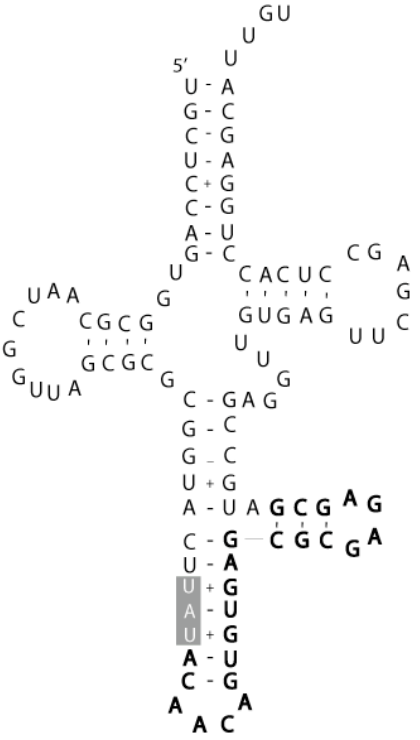
tY(GTA)N
Tyr Family



tR(TCT)S
Arg$_2$ Family

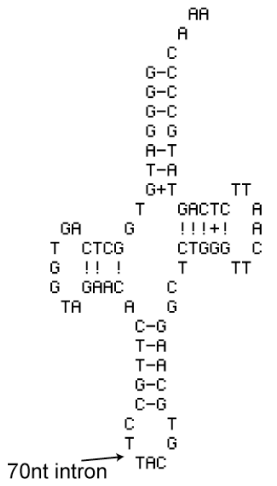


tL(CAA)M
Leu$_1$ Family

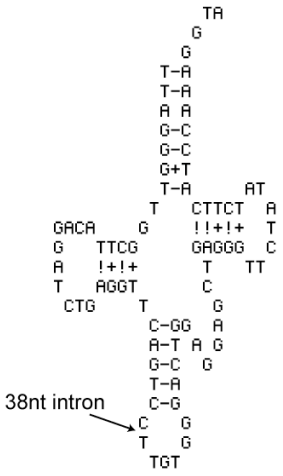


tI(TAT)M
Ile$_2$ Family

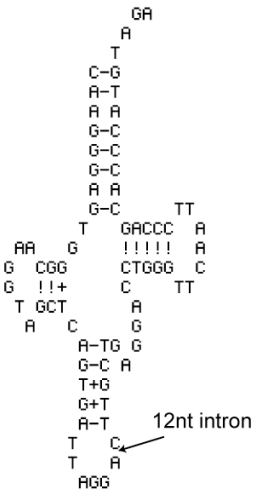
Figure 3.6. Predicted structures of the single-copy "orphan" pre-tRNAs that contain introns.
There are six intron-containing orphan tRNAs that are detected as expressed by the microarray analysis. However, the intron location of five of seven of the tRNA genes is not consistent with intron locations in yeast, only tF(GAA)O and tP(AGG)P are consistent with yeast. The intron size and insertion location is indicated with an arrow.



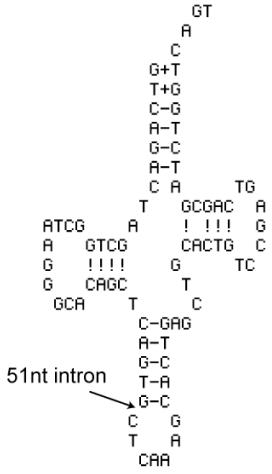
tV(TAC)G



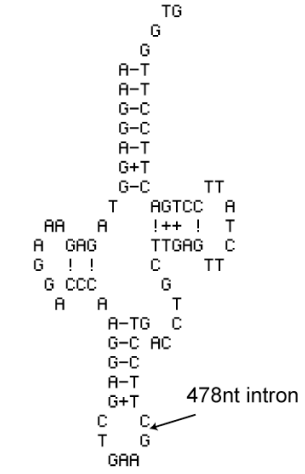
tT(TGT)E1



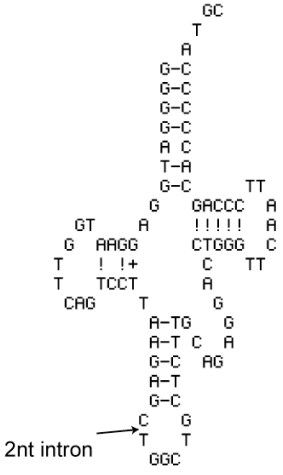
tP(AGG)P



tL(CAA)K



tF(GAA)O



tA(GGC)O1

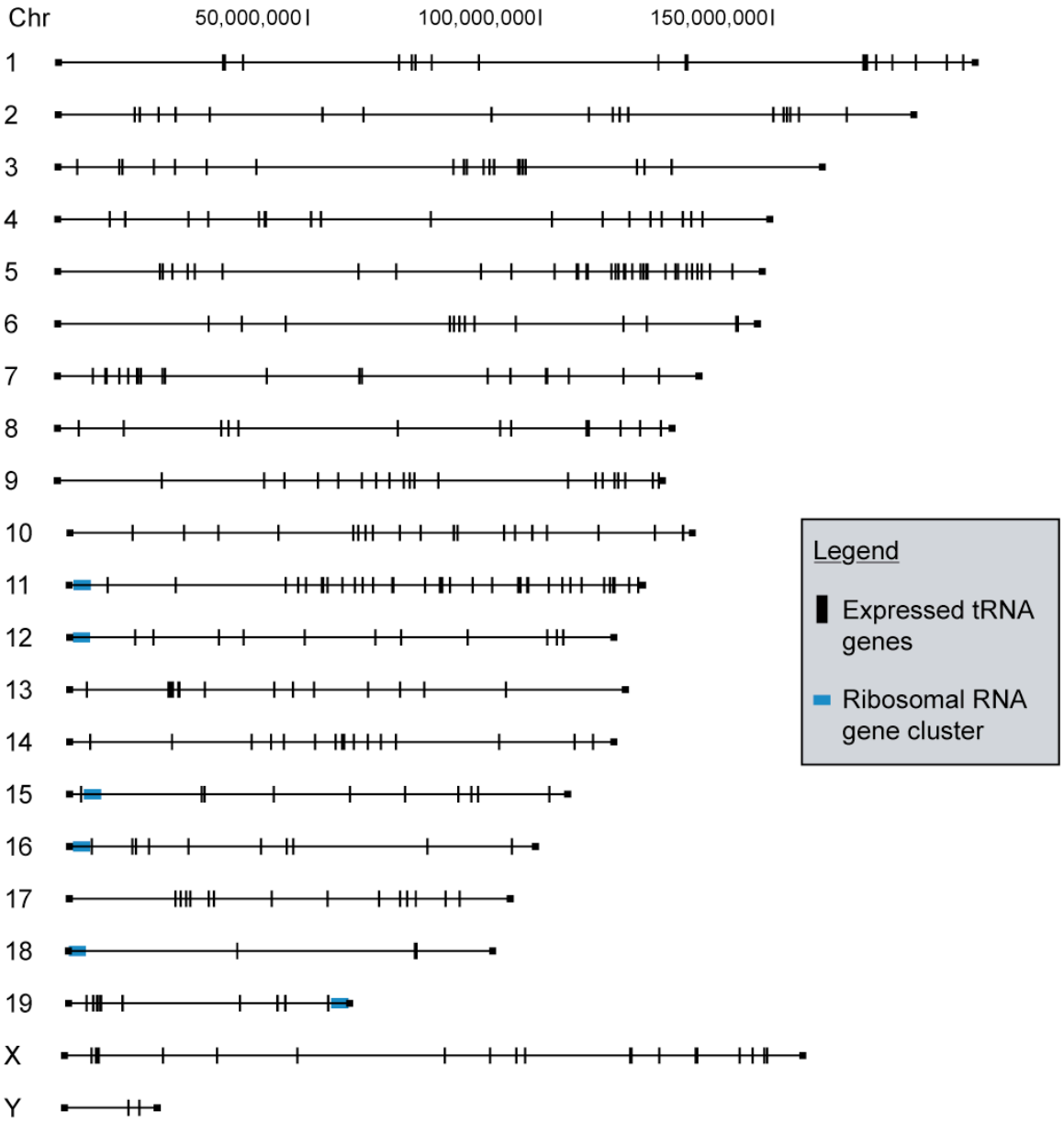
derived SINE elements are not identified by either program. Both of these programs are susceptible to identifying tRNA-derived SINE elements as functional tRNA genes, however overlapping predictions by both tRNAscan-SE and ARAGORN had the lowest likelihood of being SINEs. The predicted tRNA genes were then sorted into families based on sequence homology. The sequence variations of mouse tRNA gene copies within families are much greater than in yeast or bacteria, but similar to human tRNA genes (Goodenbour & Pan, 2006). As in yeast, tRNA genes were found dispersed throughout the mouse genome (Figure 3.7).

Analysis of the gene families identified several "rogue" tRNA genes, defined as having an anticodon for a different amino acid than the majority of the family members. In both mouse and human there are eight "rogue" tRNA genes. However, neither the anticodon nor the gene family of the rogue tRNA are conserved, the presence of rogue tRNAs in both the mouse and human genomes suggests the possibility that these tRNA genes are functional. These rogue tRNAs might facilitate anticodon variations, similar to the ambiguous intermediate hypothesis [83, 84]. By weakening a particular codon / anticodon fidelity, the codon is now more free for the incorporation of alternative amino acids such as selenocysteine.

This is the first extensive examination of tRNAs are found in mouse. Northern blot analysis of the 35 tRNA gene families confirms that tRNA-sized RNAs from all predicted families are expressed in mouse skeletal muscle, mammary gland, placenta, and

Figure 3.7. Map of tRNA gene locations in the mouse genome

Vertical lines indicate the location of verified tRNA genes throughout the mouse genome. The map illustrates the dispersed nature of tRNA genes with genes on every chromosome, compared with the ribosomal RNA gene clusters that are found on only a few chromosomes. However, there are instances of tRNA gene clustering, which appear as a thick line. One striking example of clustering is 26 of the 38 genes in the Cysteine family located within 400,000nt on Chromosome 6 (Appendix E).

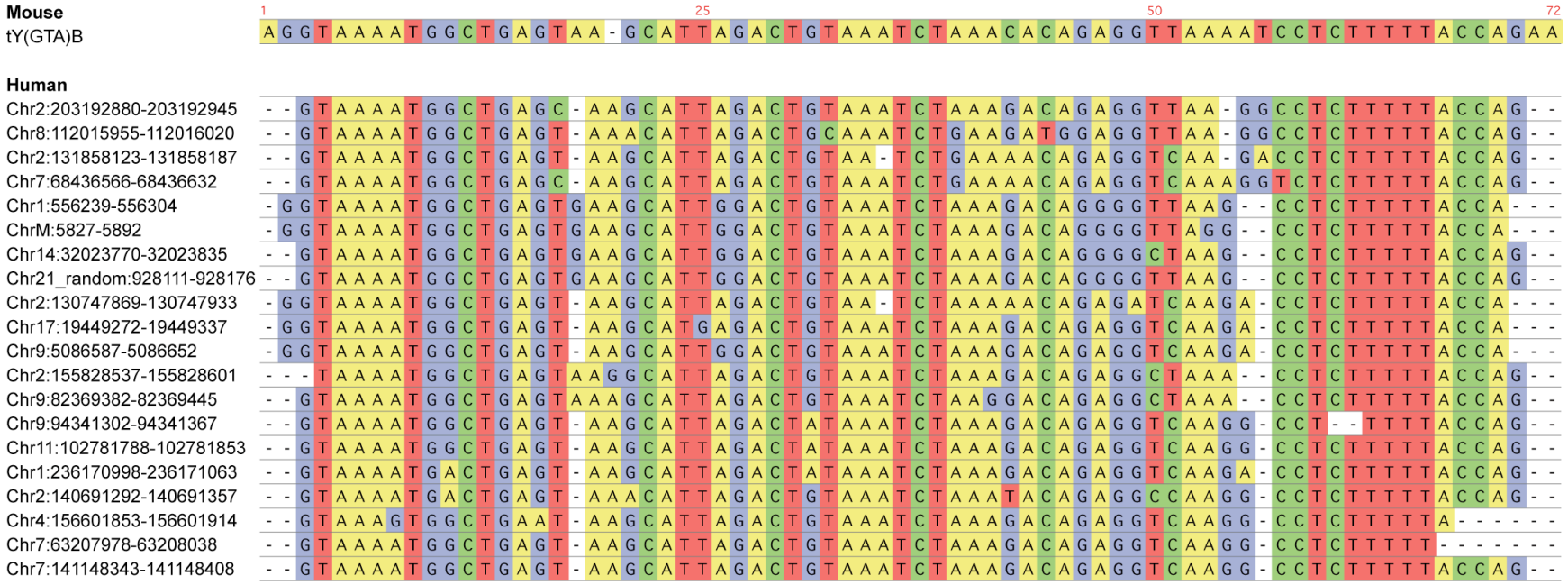

Chr
50,000,000
100,000,000
150,000,000
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
X
Y
Legend
Expressed tRNA genes
Ribosomal RNA gene cluster

7-day embryo. There are four intron-containing tRNA gene families that are expressed in mouse. The intron-containing tRNA families are conserved between mouse and humans and similar to the intron-containing tRNA types found in yeast. It is noteworthy that the intron-containing precursors can be arranged into structures similar to the intron precursor structures found in yeast [82]. Since the structure is implicated in recognition by the tRNA splicing endonuclease, this conservation would be consistent with the conservation of the splicing machinery, *SEN2* in yeast and *tSEN2* in mouse. The conservation of intron structures would also be consistent with any functions contributed by the introns. For example, it has been suggested that introns in certain tRNAs provide an additional driver for folding of the precursors to allow greater latitude in the permitted mature domain sequences. Consistent with this, the intron is required for proper folding by the yeast $tRNA^{Tyr}$ [32].

In addition to the 423 mouse tRNA genes found in families, there are 23 expressed orphan tRNAs, which were found as a single gene copy. One of the orphans is the well studied selenocysteine tRNA, *TRSP*, while another, tY(GTA)B, has 20 homologues in the human genome (Figure 3.8). This tyrosine tRNA was only identified by ARAGORN and none of the eight homologues in the human genome have been identified in the tRNAscan-SE database for the human genome. The sequence conservation and gene copy expansion in the human genome strongly argue that this is a functional RNA. Assignment of the remaining orphan tRNA genes, including the six with predicted non-cannonical introns, remains tentative at this time.

Figure 3.8. Mouse orphan $tRNA^{Tyr}$ gene corresponds to a multi-gene $tRNA^{Tyr}$ family in humans.

An orphan mouse tRNA gene, tY(GTA)B, which exists as a single copy in the mouse genome, corresponds to a probable human tRNA gene with 20 copies. This mouse tRNA gene was detected by only ARAGORN and the 20 human homologues were not detected by tRNAscan-SE. This alignment shows the homology between the mouse gene (shown above) and the 20 human homologues (>90% homology).


Mouse
tY(GTA)B
1
25
50
72
A G G T A A A A T G G C T G A G T A A - G C A T T A G A C T G T A A A T C T A A A C A C A G A G G T T A A A A T C C T C T T T T T A C C A G A A
Human
Chr2:203192880-203192945
Chr8:112015955-112016020
Chr2:131858123-131858187
Chr7:68436566-68436632
Chr1:556239-556304
ChrM:5827-5892
Chr14:32023770-32023835
Chr21_random:928111-928176
Chr2:130747869-130747933
Chr17:19449272-19449337
Chr9:5086587-5086652
Chr2:155828537-155828601
Chr9:82369382-82369445
Chr9:94341302-94341367
Chr11:102781788-102781853
Chr1:236170998-236171063
Chr2:140691292-140691357
Chr4:156601853-156601914
Chr7:63207978-63208038
Chr7:141148343-141148408

All of the mouse tRNA gene families are also found in multiple copies in the human genome, however while the number of gene copies per gene family is often similar it can vary significantly. The asparagine tRNA family has 29 gene copies in the human genome compared with only 14 in mouse. The proline family in humans has 26 gene copies, one of which contains an intron, compared with only 4 copies and no introns in the mouse genome. However it is not a general trend that humans have more tRNA gene copies than mouse, as mouse families Ala2 and Lys1 contain twice as many gene copies as in humans. We detected appropriate sized transcripts for each of the tRNA gene families in both mouse and human RNA samples, which indicates at least some of the gene copies are active in both organisms. The rogue tRNA genes and orphan tRNA genes in both mouse and human genomes require further study to determine whether they are active and producing functional tRNAs.

Acknowledgements

We thank Paul Good for his essential contributions to Figure 3.2, Tom Glover for providing mouse DNA, and Dan Bochar for providing HeLa cell extract. The custom microarrays were performed in collaboration with Tomas Babak in the laboratory of Tim Hughes. Chad Nihranz constructed the structures for the intron-containing tRNA families (figure 2.5) and assisted with the numerous northern blots.

CHAPTER IV

Conclusion

Discussion of yeast non-tRNA RNase P substrates

The comprehensive approach to identify non-tRNA substrates for yeast RNase P identified numerous potential substrates. Messenger RNAs for ribosomal proteins and other proteins involved in translational were overwhelmingly the most abundant in the co-purification with RNase P. Since the role of bacterial RNase P in pre-rRNA provides precedence for a role in ribosome biogenesis, it will be interesting to explore such a possible link to this set of mRNAs in the future. The existence of a possible link to mRNA turnover is also supported by the role of the highly similar enzyme, RNase MRP, in cell cycle-regulated turnover of specific mRNAs. Although the candidate mRNAs for RNase P are different, it is not surprising that the two enzymes, which differ by 1-2 protein subunits and have related RNA subunits, would have developed differing substrate preferences.

However, it is worth noting that the microarray results are not strand specific. The results could be due to unidentified transcripts originating from the antisense strand. This alternative possibility is strengthened when the RNAs that copurify with RNase P are compared with the results of recent work in which Rpp1p, a subunit of RNases P and MRP, was depleted and then RNAs were examined in a strand specific manner [34]. In response to Rpp1p depletion, the authors identified 74 transcripts arising from intergenic

and antisense regions of the genome. Two of the antisense transcripts, *MAN7* and *TLN1*, also copurify with RNase P and, interestingly, those are the only two antisense to ribosomal protein genes (Appendix B). In addition to the *MAN7* / *RPS14A* and *RLN1* / *RPL19B* loci, examining the strand-specific transcripts from the *RUF5* / *CUP1-1* locus in the RNase P temperature sensitive mutants would be particularly interesting. *RUF5* is a noncoding RNA expressed from the strand opposite of the copper binding protein Cup1p [85]. A transcript from this locus both copurifies with both Rpr1r and Rpr2p affinity tag purified RNase P and accumulates in the *RPR1* and both *POP1* RNase P temperature sensitive mutant strain (Appendix A – see *iYHR054C* and *YHR053C*).

Another class of non-coding RNAs was also identified in the multipronged approach, the intron-encoded box C/D snoRNAs. In yeast, there are eight intron-encoded snoRNAs, seven of which are found in mRNAs for proteins involved in translation. Further investigations into the processing of the intron-encoded snoRNAs were pursued here, since they all copurify with RNase P RNA affinity purifications. Northern blots revealed that a 5' extended pre-snoRNAs for each of the box C/D intron-encoded snoRNAs accumulated in the ts RNase P strains. The abundance of the pre-mRNAs from this pathway did not increase significantly in the RNase P ts mutants, but this is not unexpected for a maturation, rather than turnover defect.

The 5' extended pre-snoRNA is a known processing intermediate in the splicing independent intron-encoded snoRNA maturation pathway [50]. This splicing-independent pathway requires endonucleolytic cuts both 5' and 3' of the snoRNA, and leads to the

destruction of the mRNA [50]. The 5' extended pre-snoRNA already has the 3' end matured, but still contains the full transcript 5' of the snoRNA including intron, exon, and 5' untranslated region. It is possible that RNase P cuts at the 5' end of the snoRNAs *in vivo*, but it seems likely that RNase P cuts somewhere upstream of the snoRNA and 5' maturation is subsequently performed by an exonuclease. This would be similar to the case for 5' maturation of 5.8S rRNA by RNase MRP cleavage followed by exonuclease trimming.

An RNase P cut site upstream of the snoRNA is consistent with *in vitro* cleavage assays. RNase P made multiple cuts in the pre-snoRNAs, the strongest of which were in the intron upstream of the snoRNA. Analysis of the sequences in the introns revealed multiple poly-U stretches (Figure 4.1). The presence of multiple poly-U sequences 5' of the mature snoRNA site in all six box C/D intron-encoded snoRNAs could provide a binding site for RNase P. Homoribopolymers have been shown to be potent inhibitors of eukaryotic nuclear RNase P, with polyU inhibiting better than pre-tRNA substrate (Ki <10 nM, compared to >20 nM for pre-tRNA) [35]. Previous *in vitro* work with yeast pre-rRNA has identified similar, but more specific, sequence preferences [57]. The two strongest cleavage sites both occurred just 5' of the sequence: 5'-ANNAANAAUUUUN$_{9\text{-}12}$AAAUUUU-3'.

The involvement of RNase P in the processing of intron-encoded snoRNAs is especially important when considering the prevalence of snoRNAs encoded within introns in vertebrate systems. In yeast, snoRNAs are primarily individually transcribed and only 8

snoRNAs are found in introns. However, the majority of snoRNAs in vertebrate systems are intron encoded. This suggests a more substantial role for RNase P in these higher eukaryotes.

It is also interesting that seven of the eight host mRNAs in yeast are involved in translation and that the pathway that RNase P is involved in leads to a mature snoRNA and the destruction of the host messenger RNA. In addition to the snoRNA host mRNAs, the majority of RNAs that copurify with RNase P (from the top 250) are involved in translation, either directly part of the ribosome or involved in translation elongation. This might suggest that RNase P is involved in the regulation of translation through mRNA processing. This interesting role would be consistent with the involvement of bacterial RNase P and eukaryotic RNase MRP in pre-rRNA biogenesis. This suggestion is strengthened by the fact that the two noncoding RNAs identified in the Rpp1p depletion study that also copurify with RNase P are antisense to ribosomal protein subunits. The possibilities of RNase P being involved with direct mRNA turnover or the regulation of antisense transcripts raise many intriguing possibilities for translation regulation.

Discussion on mouse tRNA genes

The work on mouse tRNA genes identifies the set of pre-tRNA substrates for RNase P and provides a context for interpreting the results of *in silico* tRNA gene prediction in mammalian genomes. Over 80% of the original gene predictions were found to be tRNA-derived SINEs (although it should be noted that the ~2,900 SINE elements identified as tRNAs is much less then the ~900,000 tRNA-derived SINEs that exist in the

mouse genome). However, only 18% of the tRNA genes predicted by both tRNAscan-SE and ARAGORN were SINEs. This demonstrates the usefulness of two independent search algorithms, since each was susceptible to different SINE families.

The tRNA genes that exist in multiple copies throughout the mouse genome are the best candidates for functional tRNAs. It is interesting to note that the sequence variation of mouse tRNA gene copies is substantially greater than what is seen in bacteria or yeast. The sequence flexibility between gene copies in a tRNA family extends into the anticodon producing rogue tRNA genes. These rogue tRNA genes have the same sequence as the rest of the gene family, but has an anticodon for a different amino acid. Rogue tRNA genes appear in both mouse and human, although the gene family that contains the rogue tRNA is not conserved. The presence of rogue tRNAs has been hypothesized before being identified in the mouse or human genomes, described as the 'ambiguous intermediate hypothesis'. Rogue tRNAs could function to weaken the particular codon / anticodon fidelity, allowing for the incorporation of alternative amino acids, such as selenocysteine or pyrrolysine.

In addition to the tRNA gene families, 23 orphan tRNAs were identified and their expression was confirmed by microarray analysis. One of the orphan genes is the well studied selenocysteine tRNA, *TRSP*. BLAST searches of the mouse orphan tRNAs to the human genome revealed that another orphan tRNA gene, tY(GTA)B, has 20 homologous sequences (similarity > 90%). This orphan tyrosine tRNA is only identified by ARAGORN, and so has not been included in recent work on the human tRNAscan-SE

predictions. The sequence conservation between mouse and human and the gene copy expansion in the human genome strongly suggests that this is a functional tRNA.

Future Directions

There is an abundance of potential directions for the yeast substrate identification experiments. Examining the expression of translation machinery mRNAs and antisense RNAs in the RNase P temperature sensitive mutant could reveal a role for RNase P and the possibility of identifying novel non-coding RNAs. Further analysis of intron-encoded snoRNA biogenesis could include: competition assays with intron-encoded snoRNAs and pre-tRNAs, identifying the recognition requirements for intron-encoded snoRNAs, and exploring the involvement of RNase P in intron-encoded snoRNA biogenesis in mammalians.

The mouse / human tRNA project also leaves a lot of open doors for exploration. Since the northern blot probes hit all the members of the tRNA family, analysis of the expression of individual tRNA gene members would confirm the individual genes. This would be particularly beneficial to understanding how the rogue tRNAs are working, since the first question regarding the rogue tRNAs is whether or not they are expressed. The next question would be what amino acid they are charged with. Finally, are these charged rogue tRNAs used in translation.

Appendices

Appendix A

Top 250 most enriched RNAs for each RNase P copurification and temperature sensitive mutant accumulation experiment are highlighted in black. The values are the fold-enrichment over the non-tagged strain for the copurification and fold-enrichment over the wildtype strain grown at the restrictive temperature for the temperature sensitive (ts) mutants.

Name	ORF	Rpr1r-Aptamer co-purification	Rpr2p-TAP co-purification	RPR1 ts	POP1 233 ts	POP1 660 ts
RPL38	YLR325C	4.56	5.23	3.92	2.34	1.97
	iYHR054C	3.43	11.71	2.08	2.10	2.92
CUP1-1	YHR053C	3.36	10.52	1.87	2.16	2.65
	iYHR052W	3.16	8.42	1.98	2.08	2.84
RPS28A	YOR167C	4.33	3.88	2.19	6.51	1.18
RPL41A	YDL184C	3.57	2.26	3.05	2.04	1.98
Sui3	YPL238C	2.74	4.26	2.17	1.80	2.05
YGL231C	YGL231C	1.27	3.36	3.16	2.04	3.10
RPL34B	YIL052C	5.72	3.10	1.89	1.64	1.65
RPL34A	YER056C-A	5.11	2.99	2.02	1.66	1.77
	iYDL185W	3.87	1.39	2.88	2.93	1.40
RPL41B	YDL133C-A	3.73	3.48	2.79	1.83	1.79
	iYHR140W	3.56	1.21	2.60	2.76	1.41
RPS31	YLR167W	3.23	3.30	1.17	2.12	1.55
TIF11	YMR260C	3.20	2.59	1.96	1.75	1.95
RPL36A	YMR194W	2.99	1.40	2.17	2.39	1.36
GIS2	YNL255C	2.96	2.88	1.99	1.74	1.51
CUP1-2	YHR055C	2.93	8.91	1.79	1.99	2.45
RPS20	YHL015W	2.80	5.04	1.17	2.05	1.86
HCR1	YLR192C	2.73	6.47	1.64	3.18	1.74
RPL37B	YDR500C	2.67	1.29	2.60	2.93	1.64
	iYKL097C	2.56	1.53	2.18	2.51	1.68
	iYNLCdelta10	1.06	5.09	1.18	2.34	4.75
	YBLWTy21B	1.14	4.12	0.96	2.27	2.88
	YBLWdelta10	1.21	4.09	1.16	2.09	2.64
YOL109W	YOL109W	1.72	3.93	2.52	2.20	1.58
YLR022C	YLR022C	1.62	3.78	2.60	3.25	1.70
HSP10	YOR020C	1.30	3.41	2.45	1.53	2.79
	YKRCdelta11	1.14	3.07	1.60	2.36	2.10
YHR138C	YHR138C	0.38	1.57	3.70	2.32	2.48
YGR081C	YGR081C	1.13	1.34	3.44	4.25	2.02
YBL107C	YBL107C	1.62	1.50	3.25	2.63	2.82
YFR011C	YFR011C	1.06	1.01	2.79	2.57	2.63
RSM18	YER050c	1.12	1.75	2.76	2.51	2.73
	SNR7L	1.15	0.96	2.62	2.21	4.85
IES5	YER092w	1.18	1.31	2.55	3.89	2.04
BUD20	YLR074C	1.86	1.17	2.54	2.75	2.13
RPC10	YHR143W-A	1.25	1.54	2.53	3.11	1.93
LOC1	YFR001W	1.84	1.49	2.50	2.40	2.75
YPL071C	YPL071C	0.96	1.50	2.49	3.27	2.11
YDR339C	YDR339C	1.39	2.38	2.35	3.31	2.41
GIM5	YML094W	1.52	1.50	2.24	2.21	1.93
NOP16	YER002w	1.56	2.18	2.16	2.04	2.35
RPF2	YKR081C	1.94		2.16	2.50	2.12
MOT3	YMR070W	1.39	0.75	2.07	3.80	2.66
PHO2	YDL106C	0.90	0.70	2.05	2.75	2.57
YPL013C	YPL013C	1.47	1.87	2.04	2.05	3.08
	iYKL219W0	0.72	0.97	2.04	2.23	3.85
RPC19	YNL113W	2.01	1.28	2.00	3.13	1.93
	YORWdelta19	0.86	1.90	2.00	2.46	2.56
PRE7	YBL041W	1.45	0.81	1.99	2.05	1.91

Name	ORF	Rpr1r-Aptamer co-purification	Rpr2p-TAP co-purification	RPR1 ts	POP1 233 ts	POP1 660 ts
YPR143W	YPR143W	1.85	1.76	1.98	2.11	2.18
YHR081W	YHR081W	1.37	0.99	1.97	2.21	1.94
YBR113W	YBR113W	1.40	1.54	1.94	2.20	2.02
YIL105C	YIL105C	1.47	1.02	1.91	3.13	2.47
LTV1	YKL143W	1.17	1.18	1.91	2.84	2.49
MAK16	YAL025C	1.83	1.97	1.90	2.60	2.32
RDI1	YDL135C	1.02	1.25	1.88	2.26	2.00
	YILCdelta3	1.28	1.21	1.86	2.25	2.24
YLR407W	YLR407W	1.62	1.03	1.82	2.30	2.34
	YKLCdelta6	0.98	1.74	1.81	2.20	2.31
ABP140	YOR239W	8.65	2.47	1.85	1.71	1.43
	YOR240W	7.42	3.14	1.41	1.44	1.36
YGL102C	YGL102C	6.20	3.27	0.92	1.48	1.34
RPL5	YPL131W	6.05	4.98	1.05	1.26	0.84
YNL119W	YNL119W	5.14	3.51	0.98	1.02	0.91
RPL15B	YMR121C	4.94	3.67	1.12	1.81	1.21
RPS4A	YJR145C	4.74	3.56	1.15	1.24	0.90
YBT1	YLL048C	4.66	3.73	1.05	0.95	0.98
RPL30	YGL030W	4.65	1.24	2.01	1.98	1.46
RHR2	YIL053W	4.58	10.63	1.00	1.29	1.30
RPL12A	YEL054c	4.47	5.43	1.45	0.98	0.96
RPL27B	YDR471W	4.45	1.61	1.51	2.23	1.57
KRS1	YDR037W	4.35	3.58	1.12	1.18	1.06
RPL11A	YPR102C	4.32	4.08	1.16	1.84	1.52
YLR413W	YLR413W	4.27	3.33	1.10	1.09	0.90
RPS1B	YML063W	4.25	3.73	1.37	1.11	1.14
RPS11A	YDR025W	4.17	3.21	1.21	1.35	1.27
RPL2B	YIL018W	3.97	5.00	1.24	1.19	1.03
RPS8A	YBL072C	3.95	5.23	1.16	1.41	1.44
YLR198C	YLR198C	3.94	2.22	2.04	1.78	1.64
RPS5	YJR123W	3.88	3.48	1.13	1.17	1.13
RPS17A	YML024W	3.87	0.99	1.87	1.58	1.44
RPL20A	YMR242C	3.81	4.33	1.07	1.43	1.21
CBR1	YIL043C	3.81	2.94	0.90	1.09	0.91
RPL31A	YDL075W	3.74	2.43	1.98	2.01	1.74
RPS10A	YOR293W	3.72	1.73	2.17	1.92	1.80
RPL40B	YKR094C	3.69	1.07	1.49	2.10	1.71
ASC1	YMR116C	3.67	2.84	1.01	0.57	0.52
RPO26	YPR187W	3.66	1.33	2.04	1.07	1.71
RPS27A	YKL156W	3.61	0.69	1.44	2.08	1.30
RPS6A	YPL090C	3.59	3.12	1.02	1.20	1.38
RPL36B	YPL249C-A	3.57	4.33	1.22	1.80	1.28
RPL29	YFR032C-A	3.55	13.72	1.15	1.06	1.18
RPL33B	YOR234C	3.55	1.39	1.36	2.34	1.26
ASN1	YPR145W	3.47	5.36	1.01	1.44	0.83
RPS21B	YJL136C	3.44	4.54	1.19	1.97	1.59
RPL8A	YHL033C	3.34	17.16	1.29	1.24	1.01
RPS2	YGL123W	3.31	5.58	1.14	0.90	0.77
HIS1	YER055c	3.22	3.00	1.10	0.95	0.75
RPS18B	YML026C	3.19	2.01	2.05	1.94	1.47
SNU13	YEL026w	3.15	0.58	2.72	1.24	0.95
DIM1	YPL266W	3.14	0.93	1.77	1.98	2.02
PMA1	YGL008C	3.13	4.17	1.14	1.09	0.81
RPL17B	YJL177W	3.13	3.67	1.05	1.15	0.80
HEM13	YDR044W	3.09	3.63	1.02	0.75	0.66
EGD2	YHR193C	3.09	4.80	1.17	1.74	1.66
GSP1	YLR293C	3.06	15.55	1.10	0.97	
RPL24A	YGL031C	3.05	4.30	1.00	1.42	1.21
EFB1	YAL003W	3.01	2.83	1.35	1.21	1.22

Name	ORF	Rpr1r-Aptamer co-purification	Rpr2p-TAP co-purification	RPR1 ts	POP1 233 ts	POP1 660 ts
YLR193C	YLR193C	3.01	0.90	1.41	2.22	1.35
	iYLR324W	2.99	1.68	4.42	1.90	1.55
RPL4B	YDR012W	2.98	5.55	1.14	0.99	0.71
KRE30	YER036c	2.93	2.88	1.30	1.16	1.19
FPR4	YLR449W	2.93	3.42	1.71	1.62	1.48
RPL16B	YNL069C	2.91	3.23	0.94	0.72	0.68
RPL24B	YGR148C	2.90	7.74	1.22	1.71	1.15
RPL3	YOR063W	2.90	3.06	0.85	1.07	0.79
RLP24	YLR009W	2.88	1.61	1.77	1.94	2.03
ZUO1	YGR285C	2.87	5.37	1.23	1.01	0.82
YLL012W	YLL012W	2.86	3.11	1.49	1.82	1.09
YHB1	YGR234W	2.83	3.51	1.23	1.05	0.96
NIP1	YMR309C	2.82	3.47	1.68	1.16	1.20
	iYGL009C	2.79	3.82	1.20	1.02	0.90
PRO2	YOR323C	2.78	3.08	1.01	1.47	0.75
RPS30A	YLR287C-A	2.76	1.85	2.05	1.94	1.42
SEC53	YFL045C	2.73	2.93	1.18	0.93	1.05
TAH18	YPR048W	2.72	0.98	0.99	2.49	1.05
RPS10B	YMR230W	2.71		1.93	1.65	1.48
PSA1	YDL055C	2.70	3.44	1.00	1.41	1.40
RPL43A	YPR043W	2.70	4.82	1.58	1.33	1.38
RPB8	YOR224C	2.70	4.79	1.54	1.54	1.01
TIF1	YKR059W	2.67	4.30	1.05	0.93	0.76
ADH2	YMR303C	2.64	4.95	1.06	0.95	1.10
RPS15	YOL040C	2.63	5.72	1.10	1.29	0.61
CYS4	YGR155W	2.61	2.99	1.32	1.27	1.26
MMF1	YIL051C	2.61	3.13	0.90	1.24	0.89
ILV2	YMR108W	2.56	3.60	1.10	1.41	0.92
	irpr1	1.39	28.33	3.59	1.77	1.05
	YERWdelta21	1.08	7.78	1.14	1.75	42.13
	iYLL039C	1.10	6.32	1.13	1.36	2.35
	YCL019W	0.95	6.05	1.17	2.52	1.40
	YMR050C	1.08	5.87	1.15	1.56	2.32
	iYNLCdelta11	1.08	5.77	1.12	1.40	2.66
RPS26A	YGL189C	2.44	5.61	1.26	2.51	1.64
	YER138c	0.99	5.57	1.27	1.73	2.50
YJR028W	YJR028W	1.08	5.55	1.14	1.37	2.05
	YHRCTy11A	1.08	5.27	1.06	1.63	2.92
	YMR045C	1.07	5.23	1.14	1.41	2.09
	YJR027W	1.06	5.14	1.02	1.30	2.33
YJR029W	YJR029W	0.95	5.03	1.14	1.62	2.15
	YBL005WB	1.01	4.91	1.30	1.85	2.10
	YFLTyB	1.08	4.82	1.07	1.58	6.53
	YBL101W-A	1.08	4.81	1.14	1.51	2.41
	YGRCTy12A	1.08	4.61	1.14	1.56	1.91
RPT1	YKL145W	1.38	4.60	2.02	1.43	0.88
	YHR214CB	0.90	4.60	1.27	1.57	1.94
	YFLTyA	1.08	4.56	1.14	1.39	2.37
	iYLR035C-A0	1.08	4.38	1.14	1.44	2.29
	YAR009C	1.12	4.30	1.14	1.32	1.96
YDR366C	YDR366C	0.81	4.29	1.64	1.92	2.28
	iYAR009C	1.08	3.80	1.14	1.45	2.48
	YHRCTy11D	0.94	3.75	1.05	1.90	2.08
MFA2	YNL145W	2.36	3.72	1.95	1.36	0.98
	YBR012WB	0.82	3.62	1.61	1.44	2.24
MLC1	YGL106W	2.42	3.48	1.40	1.58	2.01
	iYDR170W-A0	1.19	3.30	1.15	2.01	2.65
	iYERWdelta211	1.09	3.24	1.09	1.32	2.75
PRE8	YML092C	1.85	3.22	1.98	1.01	0.80

Name	ORF	Rpr1r-Aptamer co-purification	Rpr2p-TAP co-purification	RPR1 ts	POP1 233 ts	POP1 660 ts
	iYKL177W	0.87	3.22	1.45	2.01	2.59
KAR2	YJL034W	1.29	3.18	0.76	2.22	1.06
VPS21	YOR089C	0.93	3.12	1.85	1.78	1.47
	LAMBDA32		3.05	2.33		
	YHRCTy11C	0.97	3.05	1.12	1.36	2.31
RLP7	YNL002C	2.45	2.95	1.70	2.64	1.68
ACB1	YGR037C	2.13	2.85	0.90	2.48	1.49
RPL-36A	YMR193C-A	1.50	0.67	4.13	3.99	1.34
YMR158W	YMR158W	1.57	1.44	3.44	1.25	2.16
UMP1	YBR173C	1.06	1.13	3.02	2.21	1.64
	YDL134C-A	2.10	0.85	3.01	2.16	1.06
	SNR4	1.22	1.58	2.83	1.62	4.25
BUR6	YER159c	1.32	1.59	2.51	3.21	1.91
SLF1	YDR515W	0.73		2.48	2.01	2.09
RRP8	YDR083w	1.61	0.93	2.43	3.16	1.77
YML025C	YML025C	1.10	1.67	2.39	1.35	1.94
YDR071C	YDR071c	1.48	2.01	2.35	1.89	2.31
SNF7	YLR025w	1.24	1.27	2.30	2.28	1.22
MIA1	YJL104W	1.90	0.75	2.29	1.65	2.44
	iYHR143W-A	0.76		2.26	2.38	1.78
NHA1	YLR138W	0.88	1.47	2.24	2.05	1.49
SRP14	YDL092W	1.71	1.43	2.22	3.22	1.64
YLR202C	YLR202C	1.59	1.29	2.22	3.35	1.67
AAD6	YFL056C	0.93		2.21	1.63	2.12
	SNR189	1.44	2.38	2.19	1.41	5.71
SED5	YLR026c	1.36	0.85	2.18	2.27	1.85
ATC1	YDR184C	1.02		2.15	1.57	2.40
	iYDRCsigma2	2.06	0.99	2.15	2.36	0.74
YOR252W	YOR252W	1.55		2.15	2.26	1.40
CDC42	YLR229C	1.75		2.13	2.02	2.14
YML108W	YML108W	0.88		2.11	2.15	1.67
YNL050C	YNL050C	1.10	1.08	2.10	2.21	1.32
COX17	YLL009c	1.00		2.07	0.87	2.82
YPR099C	YPR099C	1.09	2.04	2.04	1.44	2.50
BI3	Q0115	1.43		2.01	1.14	2.18
	iYML010CB	0.95		2.00	1.36	1.92
YDL152W	YDL152w	1.14		1.98	2.64	1.13
RTT102	YGR275W	1.00		1.98	2.46	1.51
YSA1	YBR111C	0.97	1.54	1.98	2.05	
	SNR33	1.80	0.75	1.97	0.97	2.23
YPL146C	YPL146C	1.39		1.97	2.06	1.74
SSS1	YDR086C	2.45	2.05	1.95	2.20	0.99
MRT4	YKL009W	0.88		1.94	0.91	2.82
YNR054C	YNR054C	1.55	1.08	1.93	2.70	1.44
EMI5	YOL071W	0.78	2.81	1.93	1.35	2.11
TOA1	YOR194C	1.09	0.87	1.92	2.23	1.24
YOR345C	YOR345C	1.55		1.92	2.25	1.75
YPT31	YER031c	1.90	1.63	1.91	2.17	1.29
UTP14	YML093W	1.39	1.38	1.89	2.00	1.91
	iYPL180W	1.19		1.89	2.60	1.66
TFG2	YGR005C	1.20	1.43	1.88	2.30	1.85
	SNR63	1.64	1.16	1.88	0.93	2.78
RPA34	YJL148W	1.93	1.27	1.88	1.32	2.43
NOP15	YNL110C	2.31		1.88	2.13	1.32
YSY6	YBR162W-A	1.97	1.59	1.87	2.87	1.62
YNR024W	YNR024W		1.28	1.87	3.21	0.74
YLR168C	YLR168C	0.76		1.86	1.86	2.22
SEM1	YDR363W-A	1.73	1.62	1.86	2.07	1.21
SVS1	YPL163C	1.33		1.85	3.03	

Name	ORF	Rpr1r-Aptamer co-purification	Rpr2p-TAP co-purification	RPR1 ts	POP1 233 ts	POP1 660 ts
YPR158W	YPR158W	0.91	0.85	1.84	2.36	1.78
SSF2	YDR312W	0.98	1.30	1.84	2.49	1.71
YMR290W-A	YMR290W-A	1.07	0.65	1.84	2.19	1.51
BET3	YKR068C	1.16	1.36	1.84	2.44	1.90
MRP49	YKL167C	0.93	0.81	1.82	1.06	2.61
TLG1	YDR468C	1.24	0.79	1.81	1.78	2.13
ZDS1	YMR273C	1.32		1.79	3.60	2.72
YOL031C	YOL031C	1.30	1.57	0.93	3.57	2.33
SEC28	YIL076W	1.51	0.85	1.76	3.19	2.89
YKR075C	YKR075C	0.65	1.07	1.29	2.79	3.52
	YELCdelta4	1.15	1.72	1.25	2.79	3.48
DRE2	YKR071C	1.62	0.94	1.62	2.73	2.04
RRP7	YCL031C	1.36	1.68	1.61	2.69	2.00
YER048W-A	YER048W-A	0.88	1.54	1.77	2.69	2.21
	YBL101WB	0.94	2.66	1.37	2.64	2.85
	YILCdelta2	1.57	1.45	1.70	2.63	2.59
VID24	YBR105C	0.88		0.85	2.57	1.93
TAF12	YDR145W	1.54	1.17	1.71	2.53	2.02
PUB1	YNL016W	0.88		1.53	2.52	2.40
	iYJL104W	1.36	0.46	1.60	2.51	2.05
EBP2	YKL172W	2.31	1.82	1.76	2.50	2.24
YDR210W	YDR210W	1.55	1.14	1.43	2.48	2.64
BFR2	YDR299W	1.35	0.74	1.76	2.48	1.97
	iYDR034C-A	0.82	2.21	1.51	2.47	3.36
CYC8	YBR112C	1.33		1.65	2.43	2.42
	YDRWdelta7	1.35	0.76	1.38	2.43	2.56
PCF11	YDR228C	1.06	0.59	1.76	2.42	2.02
YMR158C-B	YMR158CB	1.28	1.47	1.57	2.39	2.85
YFR008W	YFR008W	1.22	0.77	1.68	2.37	2.50
YNL114C	YNL114C	1.76	0.74	1.72	2.32	2.72
SEC72	YLR292C	1.94	1.56	1.29	2.28	2.21
YIL060W	YIL060W	0.71	1.28	1.44	2.25	2.19
UBP10	YNL186W	1.05	2.80	1.14	2.24	2.97
SSF1	YHR066W	2.16	1.42	1.62	2.24	2.08
	YJLWdelta9	0.91		1.61	2.18	2.04
	iYIL024C	2.25		1.68	2.18	4.40
POP2	YNR052C	1.85	1.09	1.46	2.16	1.94
YCR101C	YCR101C	0.61		1.35	2.16	2.94
	iYHL029C	0.82		1.68	2.15	2.09
ATP6	Q0085	1.57	0.99	1.26	2.13	2.88
YPR148C	YPR148C	1.50	2.36	1.35	2.13	2.30
YBL036C	YBL036C	1.33	0.97	1.81	2.12	2.23
YDR034W-B	YDR034WB	0.38		1.38	2.12	3.59
TAF3	YPL011C	1.16	1.30	1.37	2.10	2.16
PLP2	YOR281C	1.46	1.25	1.77	2.09	2.30
	iYERCdelta20	1.27	2.23	1.74	2.08	2.14
LCP5	YER127w	1.17	0.45	1.71	2.05	2.11
	YCLCdelta1	1.14	2.71	1.48	2.04	2.25
RPL13A	YDL082w	7.24	1.84	1.10		
RPL9A	YGL147C	5.61	1.87	1.36	1.94	1.61
DBP2	YNL112W	5.20	1.49	1.11	1.00	1.14
RPL19A	YBR084C-A	5.01	1.59	1.36	1.46	1.49
RPL19B	YBL027W	5.01		1.04	1.18	1.49
RPS12	YOR369C	4.90	1.30	1.24	0.87	1.08
RPS22B	YLR367W	4.84	1.48	0.84	1.37	1.13
RPL18A	YOL120C	4.73	2.13	0.85	1.03	0.93
RPS25A	YGR027C	4.71	1.51	1.03	1.34	1.29
RPS7A	YOR096W	4.62	1.21	0.74	0.67	1.08
RPL11B	YGR085C	4.62	2.79	0.91	1.45	1.39

Name	ORF	Rpr1r-Aptamer co-purification	Rpr2p-TAP co-purification	RPR1 ts	POP1 233 ts	POP1 660 ts
RPL1B	YGL135W	4.54	2.79	1.02	0.93	0.68
	iYHL015W	4.53	2.00	1.41	1.44	1.40
RPS25B	YLR333C	4.51	1.28	1.13	1.28	1.54
YDR417C	YDR417C	4.51	2.21	0.91	0.74	0.84
YPL197C	YPL197C	4.49	1.50	0.71	1.15	1.65
RPL6B	YLR448W	4.48	2.27	1.08	0.76	1.00
RPS3	YNL178W	4.38	1.12	0.67	0.83	0.55
	iSNR59	4.37	1.92	0.94	1.17	1.37
RPL7A	YGL076C	4.36	1.08	1.80	1.08	1.16
	iYLRCdelta8	4.32	0.32	1.37	0.79	1.25
RPS7B	YNL096C	4.24		0.92	1.35	0.94
RPS19A	YOL121C	4.20	2.81	1.33	1.97	1.41
FYV13	YGR160W	4.20	1.68	1.25	1.03	0.84
RPP2A	YOL039W	4.18	2.36	1.32	0.90	0.83
RPS19B	YNL302C	4.18	2.28	1.44	1.92	1.54
RPS17B	YDR447C	4.15	1.76	1.35	1.32	1.25
RPL18B	YNL301C	4.11	2.23	0.82	1.00	0.74
RPS1A	YLR441C	4.04	2.30	1.10	0.93	0.90
BUD19	YJL188C	4.03	0.81	1.15	1.51	1.58
RPS23B	YPR132W	4.00	1.33	1.11	1.34	1.29
RPL21B	YPL079W	3.94	2.66	1.39	1.69	1.33
RPS18A	YDR450W	3.94	2.77	1.76	1.94	1.74
RPL28	YGL103W	3.86	1.55	0.88	1.51	1.10
RPS29B	YDL061C	3.84	1.18	1.08	1.09	1.33
YLR076C	YLR076C	3.84	2.65	0.94	1.60	1.27
RPL33A	YPL143W	3.80	2.19	1.03	1.42	1.00
TIF3	YPR163C	3.79	2.11	1.20	1.10	1.26
RPL42A	YNL162W	3.78	2.48	1.48	1.13	1.08
RPL14B	YHL001W	3.78	2.00	1.22	1.13	0.75
RPL23A	YBL087C	3.78	1.57	1.46	1.45	1.42
RPL23B	YER117w	3.77	1.33	1.33	1.35	1.26
RPS14A	YCR031c	3.75	2.05	1.16	0.77	0.75
NSR1	YGR159C	3.75	1.37	1.18	1.15	1.04
BAR1	YIL015W	3.74	2.02	1.09	0.87	0.78
RPS16B	YDL083C	3.72	1.37	1.45	1.37	1.43
	iYLR159W	3.70	0.26	1.06	0.82	1.26
YOR309C	YOR309C	3.70	2.15	1.49	0.98	1.19
	iSNR65	3.69	2.29	1.57	0.79	0.89
RPS29A	YLR388W	3.67	2.33	1.08	1.44	1.30
RPS11B	YBR048W	3.67	2.15	1.23	1.36	1.34
RPS0A	YGR214W	3.66	2.57	1.07	1.01	0.86
RPL14A	YKL006W	3.62	1.85	1.25	1.16	0.97
RPL9B	YNL067W	3.60	2.02	1.03	1.49	1.22
SSB2	YNL209W	3.54	2.57	1.09	0.75	0.81
RPL20B	YOR312C	3.54	2.58	1.20	1.67	1.40
	tG(UCC)O	3.49		1.18	0.53	0.63
RPL26A	YLR344W	3.47	2.17	1.07	1.47	1.48
YBR025C	YBR025C	3.44	1.95	1.29	0.95	0.76
ARO2	YGL148W	3.43	1.51	0.62	0.91	0.83
RPL40A	YIL148W	3.43	1.04	1.10	0.97	0.99
SQT1	YIR012W	3.41	1.24	0.93	1.25	1.19
RPL42B	YHR141C	3.39	2.65	1.64	1.19	1.44
RPS4B	YHR203C	3.36	2.37	1.27	1.23	0.94
YLR339C	YLR339C	3.35	2.18	1.13	0.88	0.71
RPS24B	YIL069C	3.28	2.72	1.22	0.88	1.18
EFT1	YOR133W	3.27	2.33	1.24	0.50	0.72
KAP123	YER110c	3.24	1.90	0.93	0.86	0.80
RPP2B	YDR382W	3.24	2.27	1.21	0.92	0.87
ILV1	YER086w	3.22	2.01	1.08	0.94	0.78

Name	ORF	Rpr1r-Aptamer co-purification	Rpr2p-TAP co-purification	RPR1 ts	POP1 233 ts	POP1 660 ts
NHP2	YDL208W	3.21		1.39	1.63	1.91
IMD3	YLR432W	3.20	2.08	1.04	1.28	0.81
RPL7B	YPL198W	3.18	1.03	1.10	1.20	1.44
RPL27A	YHR010W	3.17	2.18	1.28	1.61	1.32
IMD4	YML056C	3.14	0.84	1.21	0.79	0.87
RPS13	YDR064W	3.10	1.62	1.28	1.34	1.27
RPA135	YPR010C	3.10	1.28	0.92	0.99	0.99
ILS1	YBL076C	3.10	2.80	0.90	1.11	0.93
MET17	YLR303W	3.07	2.04	0.72	1.21	1.19
RRB1	YMR131C	3.06	1.48	1.35	1.86	1.51
RPS24A	YER074w	3.06	1.83	1.33	0.95	1.22
YKL056C	YKL056C	3.06	1.77	1.53	1.66	1.27
RPC40	YPR110C	3.06	1.19	1.47	1.29	1.29
ARO4	YBR249C	3.05	2.46	1.07	1.17	0.91
IMD2	YHR216W	3.05	2.15	1.15	0.94	0.92
RPS0B	YLR048w	3.05	1.97	1.19	1.10	0.94
	iYKL156W	3.04		1.21	1.43	
RPL1A	YPL220W	3.04	1.08	0.92	0.82	0.91
MDL1	YLR188W	3.03	2.11	1.01	0.71	0.71
RPP1B	YDL130W	3.01	2.03	1.45	1.32	0.99
YER156C	YER156c	2.99	2.42	1.01	1.69	0.91
GUA1	YMR217W	2.98	2.03	1.22	0.96	0.87
GRS1	YBR121C	2.97	2.03	1.27	1.15	1.05
SLI15	YBR156C	2.96	1.88	1.03	0.85	0.69
LYS21	YDL131w	2.94	2.57	0.71	1.05	0.73
RPS21A	YKR057W	2.91	1.96	1.11	1.64	1.28
TEF4	YKL081W	2.90	1.79	1.42	0.52	0.80
PTR2	YKR093W	2.90	0.44	0.78	0.65	0.58
RPL13B	YMR142C	2.89		0.79	0.73	0.85
ALA1	YOR335C	2.88	1.65	0.87	1.01	0.97
SES1	YDR023W	2.87	2.53	1.56	1.17	1.18
	iYKL081W	2.87	1.38	0.85	0.69	0.68
YTM1	YOR272W	2.87		1.31	1.56	1.41
RPS27B	YHR021C	2.86		1.16	0.80	0.89
CCT5	YJR064W	2.86		1.25	1.44	0.78
PAB1	YER165w	2.86	2.44	1.11	1.13	0.83
YIL041W	YIL041W	2.86	2.34	0.91	1.00	1.03
RPL35B	YDL136w	2.86		1.70	1.80	1.88
RPL6A	YML073C	2.84	0.76	1.26	0.89	0.93
ARF1	YDL192W	2.83	1.31	0.71	0.87	1.79
PRP43	YGL120C	2.81	1.52	1.16	1.30	1.17
BRX1	YOL077C	2.80	2.81	1.74	1.73	1.48
	iYNL006W	2.80	1.16	1.16	0.46	0.84
YJR070C	YJR070C	2.80	0.79	1.36	1.18	1.00
LYS20	YDL182w	2.79	1.95	0.54	0.84	0.66
NOP58	YOR310C	2.79	1.35	1.28	1.27	1.18
RLI1	YDR091C	2.77	1.28	1.08	1.05	0.92
SUN4	YNL066W	2.76	1.60	1.03	0.90	0.55
FCY2	YER056c	2.76	1.58	0.93	1.00	0.73
IMD1	YAR073W	2.75	1.96	1.16	1.07	0.88
TIM44	YIL022W	2.74	2.12	1.23	1.57	1.20
RPS9A	YPL081W	2.74		1.78	0.82	0.90
TYS1	YGR185C	2.73	1.80	0.87	0.94	0.83
ARX1	YDR101C	2.72	1.65	1.64	1.56	1.37
YRB1	YDR002W	2.72	1.51	1.62	1.17	1.40
ADO1	YJR105W	2.71	1.72	1.05	0.74	0.56
RPS28B	YLR264W	2.71	0.89	1.50	1.43	1.17
RNA1	YMR235C	2.70		1.07	0.86	0.83
PHO84	YML123C	2.69	2.05	0.43	1.40	0.26

Name	ORF	Rpr1r-Aptamer co-purification	Rpr2p-TAP co-purification	RPR1 ts	POP1 233 ts	POP1 660 ts
HOR2	YER062c	2.69	1.90	0.90	1.81	1.67
SSZ1	YHR064C	2.68	1.53	1.09	1.20	0.82
DED81	YHR019C	2.68	2.29	0.93	0.87	0.58
BAT1	YHR208W	2.67	2.82	1.06	1.06	0.74
RPL26B	YGR034W	2.67	2.22	1.08	1.25	1.25
APT1	YML022W	2.66	0.76	1.50	0.98	0.80
CPA1	YOR303W	2.65	0.92	0.84	1.17	0.94
YKR043C	YKR043C	2.65	1.45	1.07	1.71	1.37
	iSNR44	2.65	1.04	0.68	0.69	0.96
UTR2	YEL040w	2.63	2.14	1.13	0.73	0.50
VTC3	YPL019C	2.62	1.65	0.96	1.03	1.24
GAR1	YHR089C	2.62	1.07	1.39	1.67	1.15
	iYEL018W	2.61	2.20	1.38	0.95	0.68
YCR051W	YCR051W	2.61	1.11	1.06	1.71	1.17
SCP160	YJL080C	2.60	2.25	1.02	0.88	1.11
LYS4	YDR234W	2.60	1.97	0.77	1.09	0.79
	iYGL008C	2.59	0.74	1.58	1.14	0.75
MES1	YGR264C	2.59	0.65	0.95	1.17	0.86
YLR194C	YLR194C	2.58	1.14	1.19	1.60	1.27
RPS23A	YGR118W	2.58	0.55	1.04	1.11	1.00
RPL22A	YLR061W	2.57	0.63	1.13	0.86	1.40
TAL1	YLR354C	2.55	2.46	0.84	1.18	0.87
TKL1	YPR074C	2.55	2.68	0.98	0.97	0.71
RPL8B	YLL045c	1.81	13.26	1.21	1.27	1.05
RPS26B	YER131w	2.34	13.17	1.22	1.86	1.07
YLL044W	YLL044W	2.24	11.42	1.24	1.43	1.17
CCW12	YLR110C	1.22	11.29	1.12	1.44	1.08
RPP1A	YDL081C	2.39	10.02	1.13	1.04	0.89
YDR134C	YDR134C	1.96	9.91	1.19	1.09	0.95
YDR233C	YDR233C	2.09	8.80	1.03	1.06	0.85
STM1	YLR150w	2.15	8.16	1.32	1.53	1.24
ERG2	YMR202W	1.88	7.98	1.45	1.26	1.05
	iYOR377W1		7.90	1.24		
	YKL097W-A	1.42	6.86	1.11	1.28	1.03
YEF3	YLR249W	1.83	6.71	1.14	1.07	0.97
HTB2	YBL002W	1.66	6.63	1.60	1.37	1.26
	iYERWdelta210	1.08	6.35	1.18	1.21	1.81
RPL4A	YBR031W	1.66	6.17	1.14	1.08	0.94
	iYERWdelta212	1.08	6.15	1.14	1.33	1.60
TEF2	YBR118W	1.41	6.03	1.14	1.19	1.18
EFT2	YDR385W	1.80	6.02	1.14	1.03	0.88
RPL17A	YKL180W	2.12	5.61	1.14	1.21	0.97
	iYPL266W	0.65	5.59	0.69		
	YER160c	1.03	5.43	1.26	1.84	1.78
	iYNLCdelta13	1.07	5.43	1.16	1.31	1.80
YDR533C	YDR533C	0.46	5.41	1.18	1.09	0.83
YDR170W-A	YDR170W-A	1.08	5.34	1.14	1.31	1.22
YDR492W	YDR492W	1.14	5.29	1.03	1.11	0.78
COX6	YHR051W	0.77	5.27	1.39	1.38	0.78
	YML039W	1.03	5.23	1.19	1.50	1.87
PDC1	YLR044c	1.14	5.18	1.12	1.04	0.91
	iCEN14	1.85	5.09	1.42	1.72	1.30
RNR4	YGR180C	1.43	5.08	1.05	0.61	0.93
YOR285W	YOR285W	0.65	5.07	1.11	1.49	1.48
	iYPR149W	0.70	5.02	1.02	0.78	0.58
	YAR010C	0.94	4.97	1.17	1.51	1.89
	iYLR035C-A1	1.07	4.95	1.14	1.21	1.21
	YDRCTy12A		4.91			
	iYNLCdelta12	1.06	4.87	1.18	1.41	1.51

Name	ORF	Rpr1r-Aptamer co-purification	Rpr2p-TAP co-purification	RPR1 ts	POP1 233 ts	POP1 660 ts
OLE1	YGL055W	1.34	4.84	1.14	1.06	0.78
	YBL005W-A	1.01	4.81	1.14	1.59	1.73
RPL2A	YFR031C-A	2.39	4.74	1.16	1.15	0.91
PMP1	YCR024C-A	2.06	4.72	1.20	0.95	0.88
TFP1	YDL185W	2.03	4.59	0.81	0.87	0.72
TDH3	YGR192C	1.18	4.50	1.13	1.16	0.95
SCW4	YGR279C	2.09	4.47	1.12	1.04	0.95
SEC14	YMR079W	2.00	4.43	1.80	1.82	1.50
ADH1	YOL086C	1.71	4.43	0.96	0.79	1.01
MSS18	YPR134W	1.13	4.38	1.00	1.21	1.21
HYP2	YEL034w	1.64	4.30	1.30	0.95	0.86
	YMR046C	1.08	4.28	1.14	1.37	1.35
CKB2	YOR039W	1.71	4.27	1.33	1.27	1.26
AAT2	YLR027c	1.79	4.25	0.89	0.91	0.71
	YHRCTy11B	1.08	4.23	1.13	1.31	1.70
TEF1	YPR080W	1.40	4.15	1.08	1.06	0.80
IDP1	YDL066W	1.67	4.08	0.71	1.02	0.71
	iYMR251W-A	0.60	4.07	1.02	0.77	0.65
HHT1	YBR010W	1.14	4.05	1.21	1.61	1.21
	IntQ0185A	1.35	4.04	0.86	1.03	1.05
	YDRCTy12D	1.07	4.00	1.10	1.19	1.18
	iYNL190W	1.47	3.95	1.24	0.70	0.68
SUR4	YLR372W	2.30	3.95	1.17	0.86	0.91
	YBR012W-A	1.08	3.94	1.14	1.30	1.24
ARC15	YIL062C	2.09	3.91	1.07	1.67	1.41
YDL228C	YDL228c	2.13	3.88	1.14	1.08	0.86
UTH1	YKR042W	1.45	3.87	1.19	1.26	1.18
YBR056W	YBR056W	0.91	3.85	1.11	1.05	0.84
	LSR1	1.00	3.82	1.19	1.61	1.74
FAA4	YMR246W	2.45	3.77	1.16	1.25	0.93
TDH2	YJR009C	2.49	3.73	1.03	0.63	0.72
	iYLR086W	0.78	3.68	1.26	1.22	1.19
ASN2	YGR124W	2.33	3.66	0.95	1.09	0.94
YEL033W	YEL033w	2.21	3.60	1.25	1.05	1.00
	YDRCTy12B	1.08	3.58	1.14	1.19	1.55
TIF4632	YGL049C	1.49	3.57	1.49	1.42	1.75
ENO2	YHR174W	1.43	3.55	1.14	0.99	1.04
	iYGR089W1	0.90	3.51	1.35	1.39	1.86
	YDRWdelta12	0.90	3.50	1.07	1.69	1.28
TDH1	YJL052W	1.08	3.49	1.14	1.07	1.18
TYE7	YOR344C	1.58	3.49	1.10	1.50	1.69
RPL32	YBL092W	2.53	3.48	1.73	1.35	1.56
ERV14	YGL054C	2.01	3.46	1.30	1.31	1.07
	iYPR009W	2.08	3.45	1.44	1.58	0.99
CPS1	YJL172W	1.23	3.42	1.02	1.05	0.74
FBA1	YKL060C	2.03	3.36	0.93	1.37	1.69
RVS167	YDR388W	1.21	3.35	1.60	1.24	1.14
INO2	YDR123C	1.21	3.32	1.02	1.51	0.97
YOL111C	YOL111C	1.28	3.30	1.14	1.41	1.03
YAH1	YPL252C	1.88	3.27	0.64	0.65	0.84
VTC1	YER072w	1.78	3.27	1.28	1.68	1.22
ARC1	YGL105W	2.05	3.26	1.14	1.10	1.08
	YGLWdelta4	0.92	3.25	1.41	1.65	1.88
GUK1	YDR454C	2.46	3.24	1.20	1.11	1.09
HHT2	YNL031C	1.19	3.24	1.16	1.48	1.09
TUB2	YFL037W	2.28	3.24	1.02	0.62	0.83
	iYBR103W	1.15	3.24	1.00	1.03	1.00
EGD1	YPL037C	2.02	3.24	1.54	1.67	1.28
SET2	YJL168C	0.78	3.19	1.30	1.40	1.47

Name	ORF	Rpr1r-Aptamer co-purification	Rpr2p-TAP co-purification	RPR1 ts	POP1 233 ts	POP1 660 ts
RPL15A	YLR029c	2.43	3.18	1.14	1.96	1.67
	iYLR380W	1.02	3.17	1.52	1.56	1.63
CCT7	YJL111W	2.43	3.16	1.31	1.37	1.19
RPP0	YLR340W	2.51	3.15	1.14	0.93	0.78
RPS8B	YER102w	1.98	3.15	1.14	1.47	1.34
	iYLR171W	2.50	3.13	1.20	1.40	1.04
CDC19	YAL038W	1.27	3.13	1.17	1.09	0.82
	tN(GUU)P	0.84	3.12	1.52	1.56	1.80
CKA1	YIL035C	2.45	3.06	1.66	1.48	1.89
YBR053C	YBR053C	1.03	3.05	1.17	1.22	0.80
SCS2	YER120w	2.17	3.05	1.03	0.91	0.80
URA5	YML106W	2.27	3.02	1.21	0.93	0.74
	YCR070w	1.18	3.02	0.95	1.19	0.83
	iYDL186W	1.96	3.00	1.25	0.83	0.63
YAL004W	YAL004W	1.42	3.00	0.93	1.24	0.85
SUI2	YJR007W	1.54	2.99	1.30	1.14	1.36
	iYJR120W	0.80	2.95	0.87	0.92	0.64
ERG6	YML008C	1.63	2.94	1.32	1.52	1.38
SSE1	YPL106C	1.66	2.94	1.09	1.07	0.76
BUB2	YMR055C	1.12	2.92	1.54	1.33	1.35
FYV9	YDR140W	1.28	2.92	1.31	1.95	1.55
GPM1	YKL152C	1.17	2.92	0.94	0.73	0.84
PDC5	YLR134w	1.39	2.91	1.17	0.96	0.87
	iYDR275W	0.89	2.91	1.15	0.74	0.67
YJR026W	YJR026W	1.08	2.91	1.14	1.19	1.35
SEC24	YIL109C	1.15	2.91	0.93	1.13	0.98
	iYJR064W	0.85	2.90	1.16	1.07	0.96
YBR246W	YBR246W	0.89	2.89	1.22	1.71	0.98
PFK2	YMR205C	2.20	2.88	1.15	0.72	0.71
	iYER106W	1.00	2.87	1.26	1.67	1.73
DIG1	YPL049C	1.81	2.85	1.01	1.29	0.93
STF2	YGR008C	0.94	2.85	0.88	0.96	1.27
SBP1	YHL034C	1.23	2.84	1.36	1.12	0.87
SMT3	YDR510W	1.19	2.83	1.71	1.69	1.44
YFL066C	YFL066C	1.35	2.83	1.06	0.96	0.97
CRP1	YHR146W	1.35	2.83	1.30	1.42	0.99
CAP2	YIL034C	2.30	2.82	1.26	1.14	0.98
	iYIL123W	1.14	2.82	1.11	0.76	1.17
	IntYBL087C	0.97		7.05	1.81	0.98
YOR235W	YOR235W	0.67		5.41	0.73	0.58
	iYMR194W	1.47	0.54	5.10	0.69	0.76
YOR235W	YOR235W	0.61		4.46	0.57	0.88
	IntYPL081W	1.13		3.40	1.05	1.02
	IntYHR010W	1.25		3.20	1.23	
	tL(C-AA)C	0.81		3.02	1.05	0.91
PHO5	YBR093C			2.93		
	iYHR216W	1.18		2.92	1.00	0.71
NOG2	YNR053C	2.33	0.79	2.89	1.19	1.70
	SNR34	1.93	0.74	2.72	0.86	1.45
	iYPR201W0	0.95	0.50	2.68	1.17	1.19
YDL110C	YDL110c	0.64		2.62		
CAP1	YKL007W	1.19	1.32	2.56	1.40	1.39
	SNR7S	0.94	0.63	2.54	0.96	1.18
	IntYBR048W	1.22	0.70	2.53	1.75	0.95
YMR188C	YMR188C	0.69	0.62	2.53	1.33	1.85
YMR002W	YMR002W	1.01	0.49	2.50	1.32	1.75
MUD1	YBR119W	1.20		2.48	1.14	1.04
	iYNR050C	2.38	1.02	2.47	0.50	1.82
LSM3	YLR438C-A			2.42		0.99

Name	ORF	Rpr1r-Aptamer co-purification	Rpr2p-TAP co-purification	RPR1 ts	POP1 233 ts	POP1 660 ts
ASP3-1	YLR155C	1.19	1.56	2.40	1.93	1.32
YOR385W	YOR385W	2.15	2.78	2.37	1.50	0.63
	iYLR146C	1.12	0.56	2.31	1.39	1.18
BSD2	YBR290W			2.29	1.79	
YML058C-A	YML058C-A	1.09	1.79	2.28	1.56	1.10
YML125C	YML125C	1.64	0.91	2.27	1.72	1.40
YHR097C	YHR097C	1.14	1.54	2.24	1.61	1.08
	iYKR090W	0.72	0.90	2.22	1.94	1.80
ARO10	YDR380W	0.78		2.22	0.71	0.83
	iYER159C	0.85		2.20	1.79	1.18
KTR2	YKR061W	0.99	0.71	2.18	1.97	1.60
YMR157C	YMR157C	0.76		2.18	1.15	1.89
YJL122W	YJL122W	1.21	0.66	2.17	1.98	1.61
MRPL40	YPL173W	1.22	1.57	2.16	1.76	0.74
CDC31	YOR257W	1.28	1.34	2.16	1.95	1.35
	tL(C-AA)N	0.73		2.16	1.13	1.00
YOL106W	YOL106W	1.07	1.46	2.16	1.68	1.71
	itL(C-AA)C	0.79		2.16	0.95	0.73
	iYDR373W			2.16		
RPN7	YPR108W	1.51	2.23	2.16	1.96	1.87
	YCRX16C	2.35		2.16	1.73	1.01
	iYOR060C	1.36		2.12	1.90	1.08
	lambda37			2.11	0.78	
PRM2	YIL037C			2.10	1.67	
FRE7	YOL152W	0.93		2.10	0.98	1.42
MRPL49	YJL096W	0.71	0.41	2.09	1.31	1.31
MRPL3	YMR024W	1.12	1.03	2.09	1.30	1.87
YKE2	YLR200W	1.48		2.08	2.02	1.52
LSM7	YNL147W	1.34		2.07	1.42	1.17
POM152	YMR129W	1.09		2.07	1.11	0.80
YML095C-A	YML095C-A	1.25		2.07	2.02	0.71
YDR279W	YDR279W	1.21	0.94	2.07	1.99	1.38
COX19	YLL018C-A			2.06	0.91	1.15
YJR014W	YJR014W	1.63	1.51	2.04	2.02	0.96
	iYHR215W	0.52		2.03	0.91	0.90
HMRA2	YCR096c	1.06	0.89	2.02	1.62	1.12
	iYML048W	1.17		2.01	1.43	1.25
YOR021C	YOR021C	1.35		2.01	1.39	1.47
	IntYPL143W			2.01	1.04	1.04
YDR365C	YDR365C	1.35	0.91	2.00	1.95	1.45
	iYGL170C			2.00	1.31	
YHR209W	YHR209W	0.64		1.99	1.30	1.89
CDC34	YDR054c	1.17	1.66	1.99	1.69	1.67
	iYMR246W	1.15	1.09	1.98	0.90	1.01
	YHRCdelta12	0.64		1.97	1.09	1.13
YAR1	YPL239W	1.09		1.97	1.65	1.66
	IntYNR053C	1.52	0.78	1.95	0.77	1.23
SNL1	YIL016W	2.42	1.77	1.95	1.01	1.03
NOC3	YLR002C	1.45		1.95	1.77	1.11
GLC8	YMR311C	0.96	1.50	1.95	1.77	1.48
	LAMBDA5	0.46		1.94	0.57	
	iYOL124C	1.09	0.62	1.94	1.39	
YBR242W	YBR242W	0.74		1.94	1.76	1.42
TOM7	YNL070W	1.44	1.91	1.93	1.86	1.44
SRB6	YBR253W			1.93	1.29	0.67
	iYNR053C	0.95		1.93	1.01	1.00
MCM1	YMR043W	1.60	0.99	1.91	1.67	1.63
ENP2	YGR145W	0.97		1.91	1.85	1.59
	snR3	1.81	1.11	1.91	0.60	1.61

Name	ORF	Rpr1r-Aptamer co-purification	Rpr2p-TAP co-purification	RPR1 ts	POP1 233 ts	POP1 660 ts
CTR1	YPR124W	0.69	0.63	1.91	0.75	1.10
YGL041C	YGL041C	0.99	1.67	1.90	1.73	1.28
SBA1	YKL117W	1.21	1.60	1.90	1.38	1.31
MCD1	YDL003W	1.16		1.90	1.33	1.30
	iYIL083C			1.90	0.93	
	iYOR051C	0.64		1.90	0.87	1.38
MRPL39	YML009C	0.84	0.69	1.89	1.30	1.76
MED11	YMR112C	0.98	1.44	1.89	1.63	1.48
	IntYNL302C			1.89	1.37	1.01
	iYOR384W			1.89		
	iYLR007W	0.76	0.90	1.88	0.59	1.58
YMR299C	YMR299C	0.92		1.88	1.78	1.81
URM1	YIL008W	1.81	1.07	1.88	1.43	1.35
	iYOL139C			1.88	1.58	0.76
YJR083C	YJR083C			1.88	1.13	0.75
RPS16A	YMR143W	1.53		1.88	0.85	0.87
HOF1	YMR032W	1.64	1.31	1.87	1.41	1.27
SUB1	YMR039C	1.33	0.74	1.87	2.01	1.63
	iYCR024C-A	1.33	2.70	1.86	0.93	
LSM1	YJL124C	1.17	0.93	1.86	1.91	1.48
MRPL37	YBR268W	1.55		1.86	1.13	1.56
DIA1	YMR316W	1.37	0.99	1.86	1.67	0.90
PRP3	YDR473C	0.88	0.55	1.86	1.89	1.29
NAB2	YGL122C	1.06		1.85	1.84	1.55
YML119W	YML119W	1.29	1.35	1.85	1.73	1.26
SNX4	YJL036W	0.81	0.99	1.85	1.70	1.49
NIP7	YPL211W	1.47		1.85	1.68	1.41
MRPL24	YMR193W	0.83	0.52	1.85	1.33	1.33
	iYGL099W	1.47	0.88	1.84	1.32	1.42
CTL1	YMR180C	0.82		1.84	1.92	1.89
YNL081C	YNL081C	0.87		1.84	1.40	0.62
	iYDL048C1			1.84		
AUT7	YBL078C	0.60	0.54	1.84	1.12	1.56
YPT32	YGL210W	1.54		1.84	1.66	1.42
YLR173W	YLR173W	1.01	0.86	1.83	1.90	1.07
QRI5	YLR204W	2.25	2.15	1.83	1.98	1.56
IES2	YNL215W	0.93		1.83	1.58	1.13
BGL2	YGR282C	1.25	0.99	1.83	1.04	0.90
COX19	YLL018C-A	0.79	0.80	1.82	1.62	0.97
YLL065W	YLL065W	0.86		1.82	1.05	0.89
TOM37	YMR060C	1.09		1.81	1.28	1.57
	iYNL174W	0.73	0.92	1.35	6.08	1.85
LTP1	YPR073C	1.19	1.21	1.74	4.09	1.42
YFR026C	YFR026C	0.62		0.66	3.60	1.40
YGR272C	YGR272C	0.77		1.15	3.46	1.24
THI3	YDL080c			1.59	3.24	
	itG(GCC)G1	1.01	0.65	0.93	3.14	1.29
RPR2	YIR015W	0.98	0.50	0.93	3.12	1.01
IMP3	YHR148W	0.92		1.69	3.04	0.95
YER030W	YER030w	1.36	0.99	1.65	2.90	1.39
FCY1	YPR062W	1.93		1.45	2.78	1.47
	15S_rRNA1	1.10	1.40	1.47	2.64	0.56
YFL046W	YFL046W	1.25	0.91	1.74	2.60	1.37
YIL127C	YIL127C	1.11		1.72	2.56	0.63
YPL044C	YPL044C	2.40	2.23	1.50	2.55	1.48
YGL242C	YGL242C	1.54		1.55	2.50	1.39
NOP4	YPL043W	2.54	1.76	1.55	2.49	1.61
YDR020C	YDR020c	1.19		1.51	2.49	1.18
ERO1	YML130C	1.45	1.26	1.23	2.47	1.76

Name	ORF	Rpr1r-Aptamer co-purification	Rpr2p-TAP co-purification	RPR1 ts	POP1 233 ts	POP1 660 ts
GLO3	YER122c	1.82	1.18	1.48	2.46	1.68
TCI1	YDR161W	1.20	0.62	1.46	2.46	1.33
RFM1	YOR279C	0.95		1.65	2.45	1.31
ECM13	YBL043W	0.61		0.54	2.45	1.05
FPR3	YML074C	2.09	2.09	1.79	2.42	1.44
URA3	YEL021w	1.02	1.35	0.99	2.41	1.36
DST1	YGL043W	1.05		1.72	2.40	1.81
RNA15	YGL044C	1.24	1.34	1.58	2.40	1.67
	iYGL044C			1.36	2.40	
YKL063C	YKL063C	1.03	1.25	1.36	2.39	1.66
IBD2	YNL164C	0.74		0.89	2.39	1.43
HCA4	YJL033W	2.19	0.98	1.59	2.38	1.64
CNS1	YBR155W	1.81		1.51	2.37	1.75
SED1	YDR077W	1.13	2.30	1.22	2.37	1.43
YML053C	YML053C	1.55		1.77	2.36	1.75
YBR271W	YBR271W			1.25	2.35	1.37
SSK1	YLR006c	0.77		1.22	2.35	0.89
ERV1	YGR029W	1.42		0.75	2.35	1.11
GRR1	YJR090C	0.84	1.16	1.47	2.34	1.85
YPR050C	YPR050C	1.20		1.40	2.34	1.37
GYP7	YDL234C	1.35	0.51	1.58	2.33	1.52
ECM1	YAL059W	1.26	0.86	1.61	2.32	1.59
CWC27	YPL064C	1.26	1.09	1.78	2.31	1.23
BUD22	YMR014W	0.99	1.08	1.67	2.31	1.53
CTF8	YHR191C	0.99	0.85	1.56	2.31	0.73
YNL260C	YNL260C	1.05	0.87	1.65	2.30	1.09
YNL174W	YNL174W	1.87	0.99	1.72	2.30	1.42
	iYMR078C	1.27	0.97	1.61	2.29	1.08
YDR026C	YDR026c	1.27	1.20	1.53	2.29	1.56
	iYPR164W	1.40		1.46	2.29	1.40
YDR426C	YDR426C	0.61		1.16	2.27	0.93
AIR1	YIL079C	1.82	1.17	1.23	2.26	1.28
AGE1	YDR524C	1.28	0.85	1.50	2.25	1.77
APS3	YJL024C	1.02	0.99	1.65	2.25	1.24
SLU7	YDR088C	0.92	1.05	1.51	2.25	1.28
PBI2	YNL015W	0.61	0.70	1.23	2.25	1.55
MUM2	YBR057C	1.19	0.93	1.30	2.24	1.62
CMK2	YOL016C	1.54		1.20	2.24	0.81
MRD1	YPR112C	1.78		1.40	2.22	1.52
DBP5	YOR046C	1.65	1.36	1.35	2.22	1.15
RRS1	YOR294W	1.83		1.68	2.21	1.24
UAF30	YOR295W	1.27		1.55	2.20	1.53
	iYDL217C	0.98		1.16	2.20	
YDR153C	YDR153C	1.16	1.19	1.77	2.20	1.44
NOP4	YPL043W	2.25	0.65	1.48	2.20	1.59
YBL081W	YBL081W	1.60	0.70	1.48	2.19	1.39
IMH1	YLR309C	1.01	0.89	1.45	2.19	1.85
	iYER164W	1.20		1.41	2.19	0.96
YPL157W	YPL157W	1.33		1.63	2.18	1.49
CBP6	YBR120C	0.65		1.41	2.18	1.19
SIW14	YNL032W	1.18	1.14	1.56	2.18	1.85
YJR003C	YJR003C	1.23	1.00	1.47	2.18	1.53
SRP101	YDR292C	1.31	0.78	1.46	2.18	1.48
HAS1	YMR290C	1.86		1.43	2.16	1.41
SAS10	YDL153c	1.42	1.31	1.73	2.15	1.62
RPB4	YJL140W	1.42	1.61	1.75	2.14	1.65
YOR277C	YOR277C	2.19		1.43	2.14	1.60
	iYNR012W	0.74		1.24	2.14	
YFH1	YDL120w	1.14	1.07	1.35	2.14	1.46

Name	ORF	Rpr1r-Aptamer co-purification	Rpr2p-TAP co-purification	RPR1 ts	POP1 233 ts	POP1 660 ts
YKR012C	YKR012C			1.28	2.13	1.38
WRS1	YOL097C	2.02	1.12	1.19	2.13	1.27
YGR046W	YGR046W	0.78	1.38	1.16	2.13	1.45
VPS29	YHR012W	0.95		1.31	2.12	1.12
	iYIR014W	0.84	0.40	1.12	2.12	0.90
MCA1	YOR197W	1.88	1.84	1.43	2.11	0.96
UTP18	YJL069C	2.00	0.86	1.38	2.11	1.82
RPC37	YKR025W	1.30		1.23	2.11	1.58
	YILWTy31D	1.02	1.55	1.75	2.11	1.52
YJL184W	YJL184W		1.07	1.70	2.11	1.24
UTP4	YDR324C	1.66	1.03	1.51	2.10	1.36
PXR1	YGR280C	1.23		1.44	2.10	1.01
PSH1	YOL054W	1.33		0.98	2.09	1.89
AHP1	YLR109W	0.77	1.31	1.55	2.09	1.69
SSP120	YLR250W	1.08	1.07	1.31	2.09	1.47
YGR251W	YGR251W	1.60	1.01	1.41	2.08	1.10
	YCRX04W	1.28	0.92	1.17	2.08	1.26
	iYERWdelta12	0.67		1.17	2.08	0.92
APM1	YPL259C	1.29	1.48	1.28	2.08	1.28
RER2	YBR002C	1.61	0.94	1.15	2.08	1.32
	iYDR419W	1.19		1.14	2.08	1.64
	iYEL043W	1.53	1.57	1.50	2.07	0.97
MIS1	YBR084W	2.42	1.81	1.37	2.07	1.02
YDR288W	YDR288W	0.97	0.67	1.33	2.07	1.14
GZF3	YJL110C	1.00	1.84	1.23	2.06	1.79
RXT2	YBR095C	0.85		1.04	2.06	0.76
CSL4	YNL232W	1.35	1.31	1.54	2.06	1.48
HIS6	YIL020C	1.50	0.97	1.50	2.06	1.24
	iYIL063C	1.38		1.80	2.05	1.50
TRM1	YDR120C	2.29	1.50	1.46	2.05	1.21
GCD14	YJL125C	1.59	0.94	1.74	2.05	1.78
PET18	YCR020c	0.51		1.42	2.04	0.87
	21S_rRNA2	1.08	0.84	1.14	1.51	93.13
	21S_rRNA0	1.07	0.51	1.15	1.59	13.38
AI3	Q0060	0.51		1.60	0.56	7.57
	21S_rRNA0	1.08	0.45	1.45	1.48	6.71
	Q0035	0.91		1.76	0.76	6.48
AI5_ALPHA	Q0070	0.64		0.89	0.94	5.24
	IntQ0280A	1.51	1.88	1.72	1.15	4.84
	IntQ0280A	1.24	0.97	1.35	1.11	4.08
	IntQ0280D	0.57	0.14	1.63	0.68	3.98
AI5_BETA	Q0075	0.63		1.05	0.65	3.91
	iYMR046W-A	0.94	1.19	1.28	1.93	3.70
	Q0280F	1.62	0.49	1.67	0.77	3.34
YER181C	YER181c	0.44		1.76	1.74	3.32
	Q0270	0.66	1.94	1.03	1.12	3.29
	Q0315	1.10		1.27	0.80	3.26
	Q0283	1.12		1.78	1.03	3.25
	Q0295	1.43	2.01	1.42	0.56	3.18
	Q0283	1.08		1.31	0.74	3.16
	Q0280F	1.41	0.60	1.18	0.57	3.08
	Q0030	0.85		1.19	0.83	3.07
	lambda25			1.14		2.96
	IntQ0280D	0.85		1.71	0.82	2.95
HMG1	YML075C					2.93
HXT2	YMR011W	1.44	0.61	1.07	1.90	2.80
AI3	Q0060	0.52		1.07	0.73	2.73
YHR033W	YHR033W		0.76			2.72
	Q0005	1.16		1.00	0.66	2.70

Name	ORF	Rpr1r-Aptamer co-purification	Rpr2p-TAP co-purification	RPR1 ts	POP1 233 ts	POP1 660 ts
	Q0270	0.84	2.05	1.11	1.24	2.68
	Q0295	1.46	1.65	1.55	0.58	2.67
AI4	Q0065	0.76		1.16	0.76	2.65
	iYCR100C	0.62		1.17	1.61	2.64
	21S_rRNA1	1.08	0.84	1.14	1.32	2.63
ATP6	Q0085	1.42	1.26	1.45	1.69	2.62
YJR162C	YJR162C	1.25		1.79	1.24	2.58
	itM(C-AU)O1	0.76		0.89	1.29	2.53
	IntQ0280E	1.13	1.36	1.48	0.99	2.53
	YCL021W	0.71		1.31	1.87	2.49
MAM33	YIL070C	1.09	1.05	1.46	1.16	2.44
	YCL020W	0.66	2.47	1.13	1.70	2.43
	SNR190	1.09	0.63	1.56	1.32	2.42
CYT2	YKL087C	1.05	0.99	1.77	1.74	2.41
GLN3	YER040w	1.01	1.43	1.33	1.71	2.41
MRPL7	YDR237W	1.18	0.91	1.79	1.00	2.41
RRN10	YBL025W	0.95	0.74	1.36	1.67	2.37
STE18	YJR086W	1.03		1.58	1.55	2.34
YHR040W	YHR040W					2.34
COB	Q0105	1.25		1.71	0.99	2.31
	YHLCsigma1	1.13	1.29	1.49	1.89	2.30
YLR435W	YLR435W	1.91	0.97	1.71	1.63	2.29
YHR145C	YHR145C	0.91		1.20	1.80	2.28
TIF35	YDR429C	2.24	1.58	1.52	1.66	2.27
MRP51	YPL118W	1.30	0.69	1.58	1.28	2.26
RSM10	YDR041W	0.86	1.09	1.29	1.21	2.26
YBR210W	YBR210W	1.41		1.06	1.52	2.25
STF1	YDL130W-A	0.45	0.99	1.18	1.12	2.24
SWD3	YBR175W			1.62	1.92	2.24
	YDRWdelta23	1.12	2.29	1.57	1.92	2.23
	YCRCdelta6	1.06	2.74	1.54	1.99	2.23
MRPL17	YNL252C	1.22	1.02	1.46	1.17	2.22
	itH(GUG)G1	0.53		1.04	1.51	2.20
YKL169C	YKL169C	1.20	0.65	1.53	0.78	2.20
YLR008C	YLR008C	1.46	0.60	1.80	1.79	2.19
	iYBR293W	0.64	0.89	1.75	1.36	2.19
	iYMR191W	0.96		1.27	2.02	2.18
	SNR70	2.03	0.69	1.11	0.78	2.18
	iYOR235W	0.97	0.83	1.15	1.11	2.16
YOR193W	YOR193W	0.59		0.98	1.83	2.14
MGE1	YOR232W	2.11	1.73	1.33	1.70	2.12
YOR135C	YOR135C	0.82	2.16	1.12	1.61	2.11
	iYPL144W	0.87	0.92	1.79	1.24	2.11
	IntQ0280E	0.81	1.03	1.39	0.71	2.10
YPL141C	YPL141C	1.21		1.60	1.39	2.08
LSM6	YDR378C	1.48		1.63	1.74	2.05
HXT4	YHR092C					2.05
	YBLCsigma1	0.99	1.62	1.02	2.03	2.05
AUT1	YNR007C	0.71	0.76	1.21	1.55	2.04
UTP7	YER082c	1.26	0.60	1.48	1.59	2.04
	Q0290	0.77		0.93	0.76	2.04
YGR021W	YGR021W	1.06		1.71	1.34	2.03
SRL3	YKR091W	0.86	1.50	1.63	1.81	2.03
YHR049C-A	YHR049C-A	1.16	1.19	1.51	1.78	2.03
RMD9	YGL107C	1.17	0.59	1.38	1.17	2.03
BRE4	YDL231c	1.01	1.20	1.24	1.76	2.02
	Q0350	1.30	1.04	1.31	0.66	2.02
YER080W	YER080w	1.01	0.97	1.19	1.96	2.01
YLR003C	YLR003C	1.63	0.80	1.54	1.86	2.01

Name	ORF	Rpr1r-Aptamer co-purification	Rpr2p-TAP co-purification	RPR1 ts	POP1 233 ts	POP1 660 ts
RSM19	YNR037C	1.38	0.82	1.56	1.46	2.01
MHR1	YDR296W	1.19	1.55	1.38	1.22	2.01
MHT1	YLL062C	0.93	0.39	1.24	0.94	2.01
YJL010C	YJL010C	0.82	2.24	1.65	2.00	1.99
YDR149C	YDR149C	1.17		1.00	1.60	1.99
YDL036C	YDL036c	1.11	2.04	1.34	1.59	1.98
YOR146W	YOR146W	2.05	2.12	1.64	1.90	1.97
YPT1	YFL038C	1.30	1.78	1.24	1.71	1.97
SWH1	YAR042W	1.04	0.99	1.04	1.20	1.97
	iYLL012W	0.63		1.13	0.93	1.96
	SNR13	1.72	0.88	1.48	1.21	1.95
YNL228W	YNL228W	1.30	0.89	1.55	2.00	1.94
YNR066C	YNR066C	1.49	0.53	1.15	1.37	1.94
YNL056W	YNL056W	1.40	1.05	1.41	1.81	1.94
	iYBL081W	0.62		1.06	1.99	1.93
	iYOR283W	0.97	1.65	1.58	1.68	1.93
	iYIR001C	0.91		1.44	1.32	1.93
YGR205W	YGR205W	0.75	1.04	1.01	1.10	1.93
	iYOL015W	0.83	1.49	0.95	0.91	1.92
PEX19	YDL065c	1.20	0.73	1.73	1.68	1.92
MET14	YKL001C	1.69		1.37	1.51	1.92
	iYLR228C0	0.76	1.32	1.09	1.75	1.91
YUH1	YJR099W	1.29		0.81	1.30	1.91
YPR090W	YPR090W	1.28	0.91	1.14	1.21	1.91

Appendix B

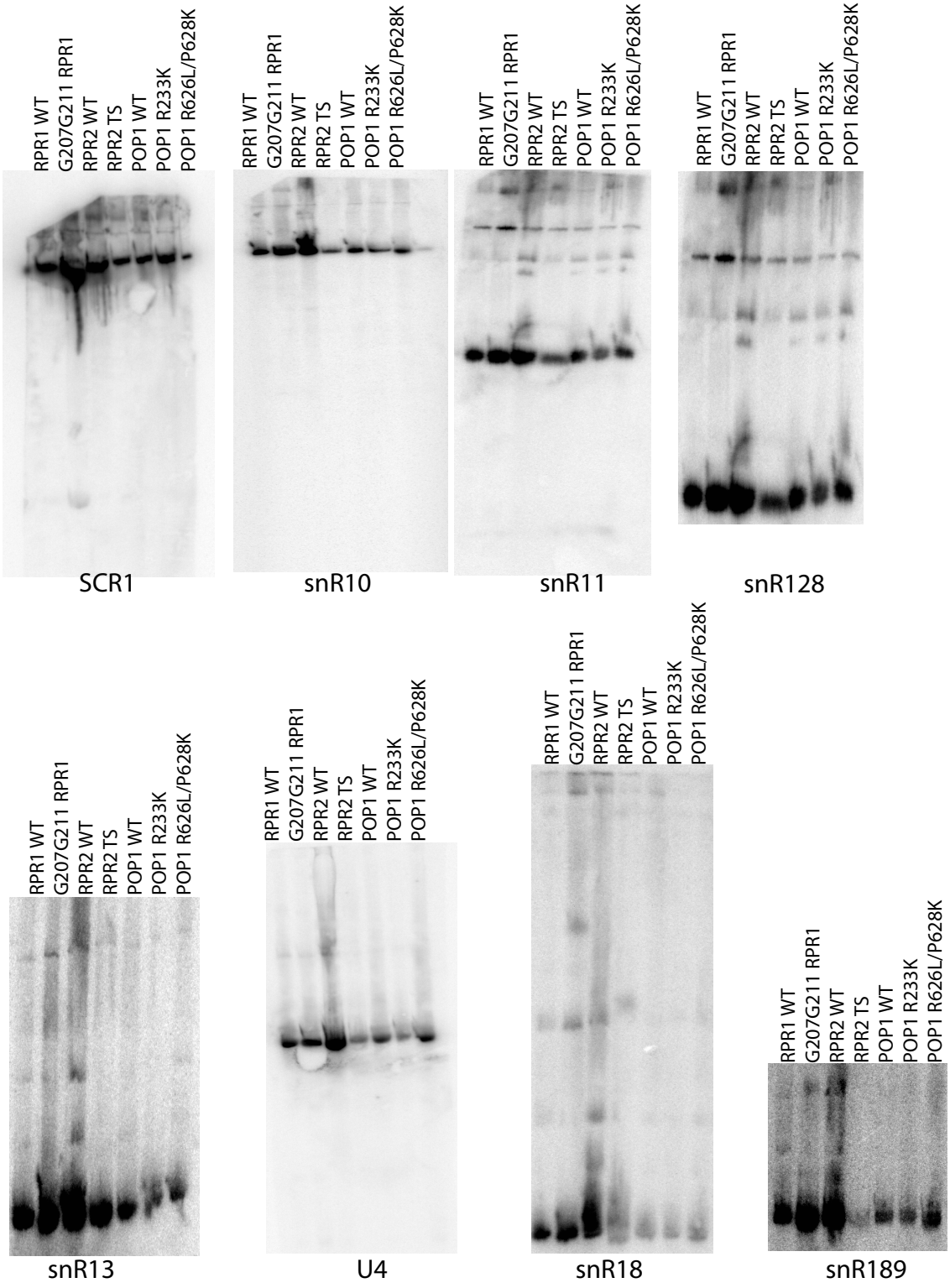
Some non-coding RNAs affected by Rpp1p depletion [34] also copurify with RNase P or are affected by ts mutant strains. Shown here is the list of RNAs identified in the Rpp1p depletion study and how well each RNA copurifies with RNase P or how much it is affected by the temperature sensitive mutations in the RNase P subunits. If the RNA was in the top 250 most copurifying or increased in abundance, the fold-enrichment value is highlighted in black.

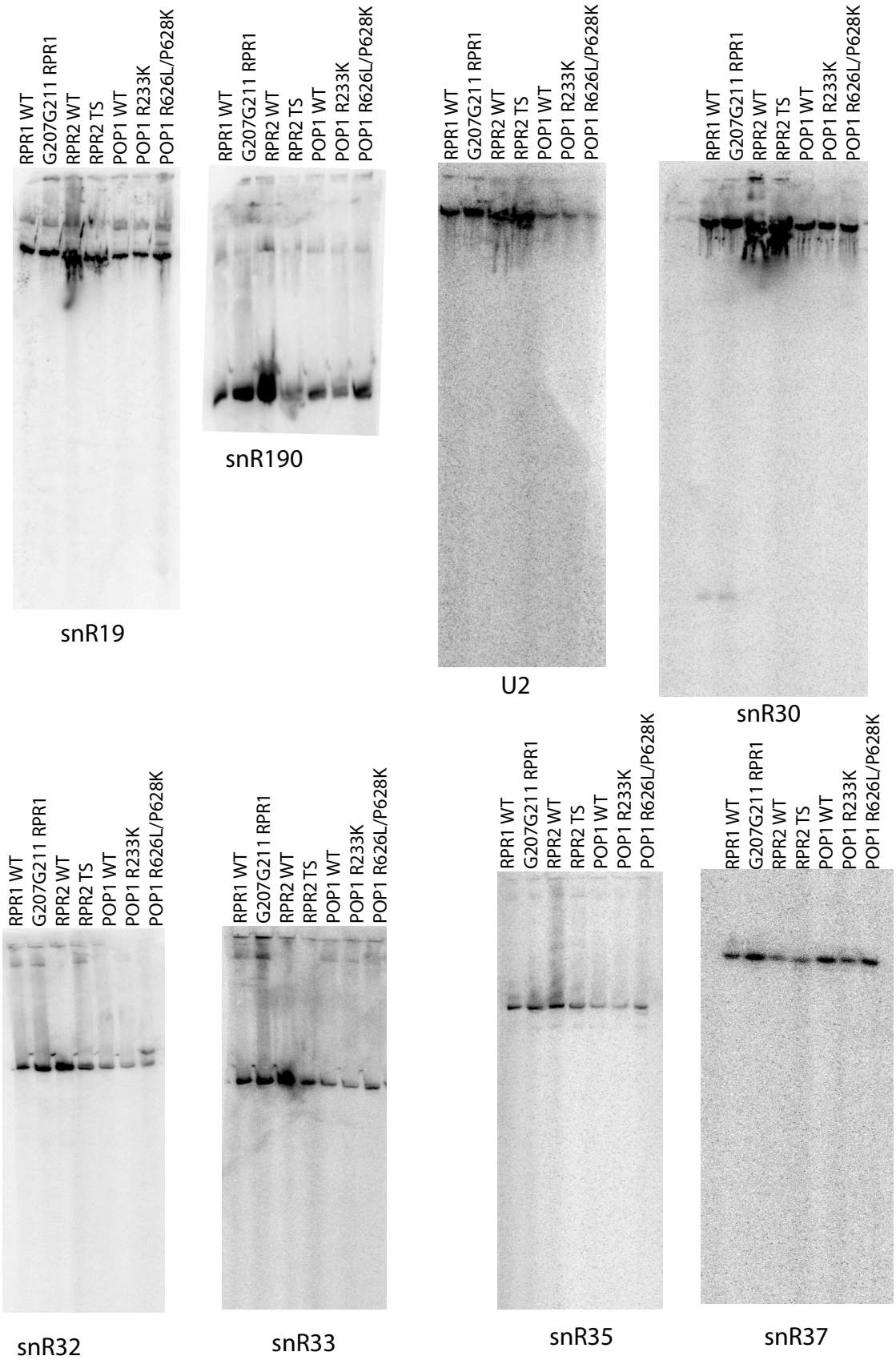
	Rpr1r RNA copurification	Rpr2p TAP copurification	RPR1 ts	R233K POP1 ts	R626L/P628K POP1 ts
HRA1	1.03	1.09	1.33	1.48	1.61
HRA5	1.09		1.12	1.32	1.11
HRA9	0.85	0.69	0.79	1.48	0.91
HRA10	0.85		1.00	1.33	0.81
HRA11	0.85		1.07	0.94	1.21
HRA13	1.25	0.69	1.24	1.32	1.46
HRA14	1.29	0.82	0.80	0.81	0.87
HRA16	0.94	0.73	0.92	0.97	0.92
MAN2	0.54		0.78	0.91	0.63
MAN3	0.73		0.97	0.93	0.82
MAN4	1.03	1.00	0.91	1.15	1.08
MAN5	1.06		0.92	1.31	0.71
MAN6	0.52		0.79	0.68	0.82
MAN7	*3.75*	2.05	1.16	0.77	0.75
MAN8	0.56	0.56	0.84	1.03	0.89
MAN9	1.05		1.00	1.76	0.45
MAN10	0.55		0.92	1.16	0.84
MAN11	0.58		1.01	0.89	0.91
MAN12	0.64		1.25	0.98	0.95
MAN13	0.63		1.01	0.97	0.94
MAN14	0.89	0.36	1.08	0.92	1.03
MAN15	0.93	0.89	1.09	1.00	0.94
MAN17	0.66		1.04	0.91	0.82
MAN18	0.52		0.78	0.69	0.69
MAN19	0.87	0.62	1.23	1.48	1.13
MAN20	0.89	1.09	0.99	1.18	1.02
MAN21	0.79		0.82	0.91	1.03
MAN22	1.01	0.73	1.00	0.87	0.81
MAN23	0.82		0.83	1.10	1.01
MAN25	0.93		1.56	1.34	0.84
MAN26	0.98		0.67	0.79	0.89
MAN27	0.96	0.27	1.09	1.14	1.54
MAN28	1.09	0.78	1.01	1.04	1.18
MAN29	0.67	0.42	1.23	0.92	1.40
MAN30	1.10	0.62	1.02	0.92	1.02

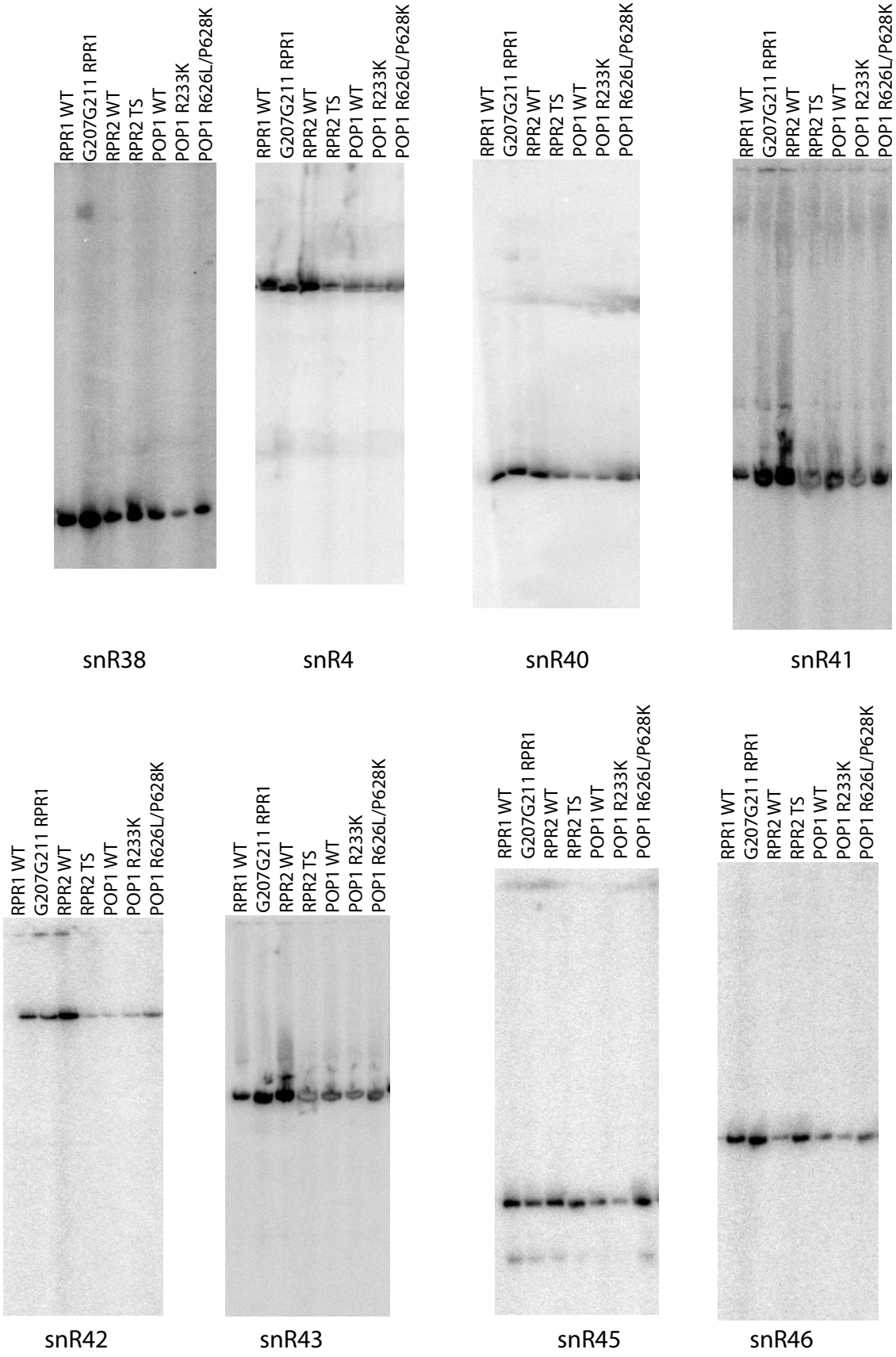
	Rpr1r RNA copurification	Rpr2p TAP copurification	RPR1 ts	R233K POP1 ts	R626L/P628K POP1 ts
MAN31	1.07		0.80	0.82	0.90
MAN33	1.12	0.56	2.31	1.39	1.18
MAN34	1.14	0.35	1.07	1.66	1.02
MAN35	0.60		0.95	0.78	0.86
MAN36	0.55		1.16	1.30	0.59
MAN37			1.05	1.17	0.70
MAN38			1.06	1.18	1.07
MAN39	0.75	0.88	1.00	1.10	1.17
MAN40	0.77	0.35	0.79	0.90	0.71
MAN41	0.64		0.53	0.79	0.62
MAN42	0.69		0.84	0.80	0.99
MAN43	0.84		0.57	0.57	0.73
MAN44	0.87		0.76	0.83	0.62
MAN45	0.70		0.86	0.93	
MAN47	1.32		0.75	0.94	0.93
MAN48	0.34				
MAN49	0.94	1.11	1.09	1.18	1.11
TLN1	*5.01*		1.04	1.18	1.49
TLN2	0.76	1.61	0.89	1.02	0.98
TLN3	0.79		0.97	0.79	0.93
TLN4	0.77	0.32	1.09	0.90	0.97
TLN5	1.01		1.23	1.24	1.03
TLN6	0.41		1.02	0.96	0.95
TLN7	1.24	0.37	0.97	1.45	
TLN8	0.71	0.52	0.85	0.91	0.83
TLN9	0.93	0.91	0.94	0.93	0.82
TLN10	0.64		0.79	0.60	
TLN11	0.67		0.51	0.82	0.73
TLN12	0.62		0.95	0.72	0.81
TLN13	1.63	1.47	1.12	1.13	1.12
TLN14	0.99	0.63	1.23	1.45	
TLN16	0.63		0.84	0.81	0.96
TLN17	0.95	0.59	0.93	1.23	1.06
TLN18	0.65		0.96	1.00	1.30
TLN19	1.17		1.08	0.54	0.82
TLN20	0.76		1.52	1.25	1.17
TLN22	0.71		0.81	0.79	0.79
TLN23	0.77		0.98	0.97	0.64
TLN26	0.60	1.04	1.26	0.94	0.79

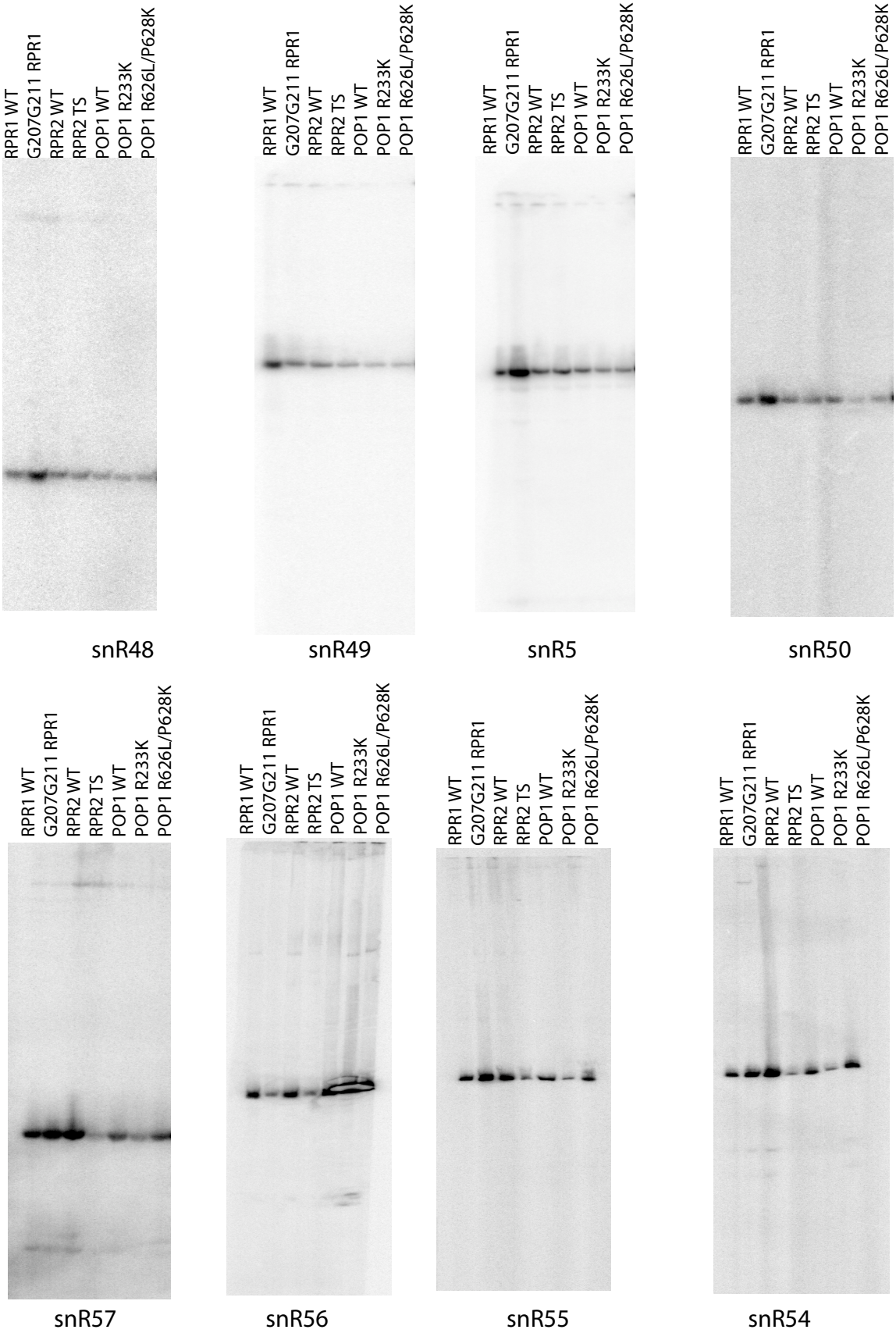
Appendix C

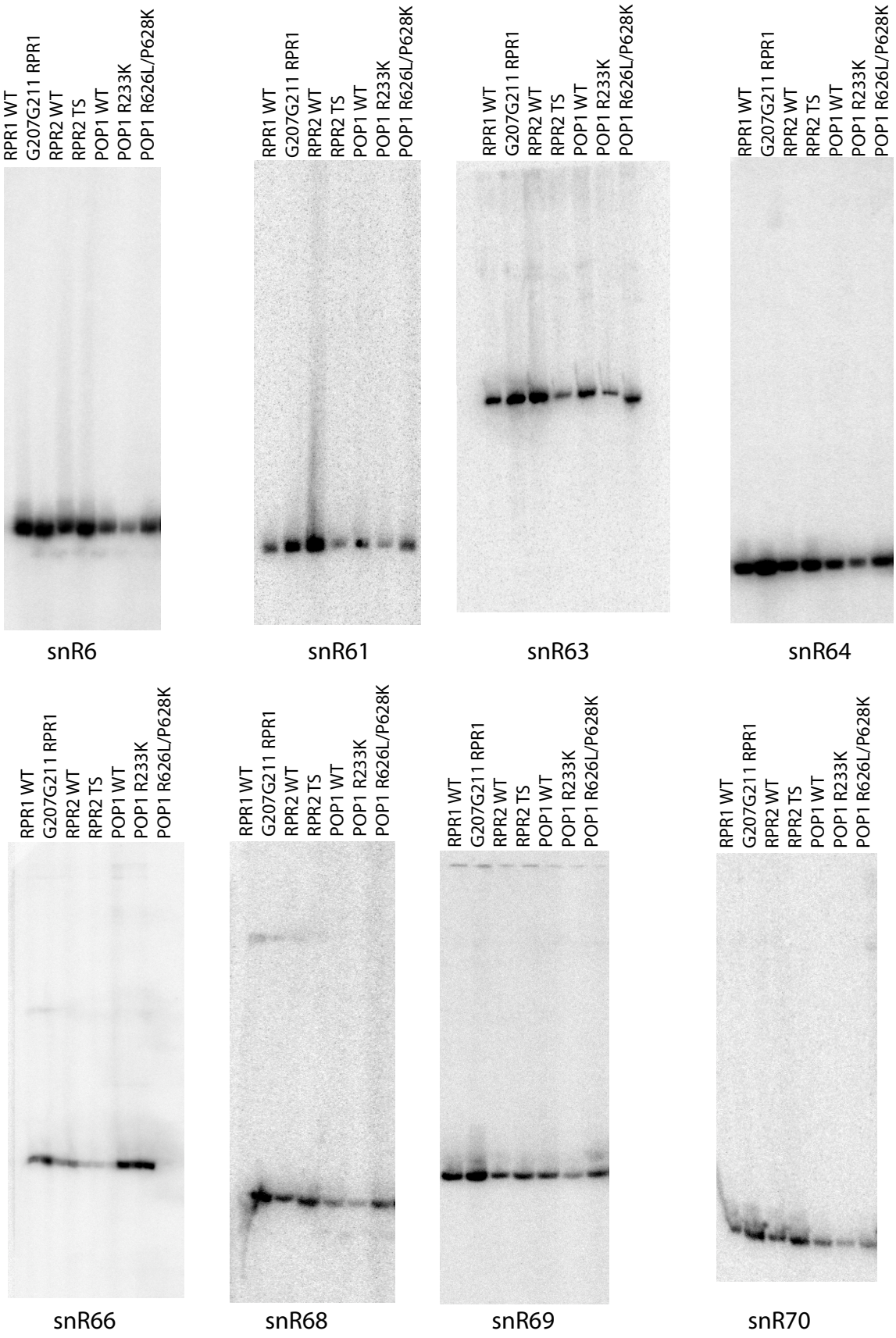
Northern blots of snoRNAs in RNase P temperature sensitive mutants were performed in order to identify any aberrant processing. Total RNA from yeast grown at the restrictive temperature (37ºC) was probed for various snoRNAs, one at a time. These are the snoRNAs that did not have aberrant processing in the ts mutants.

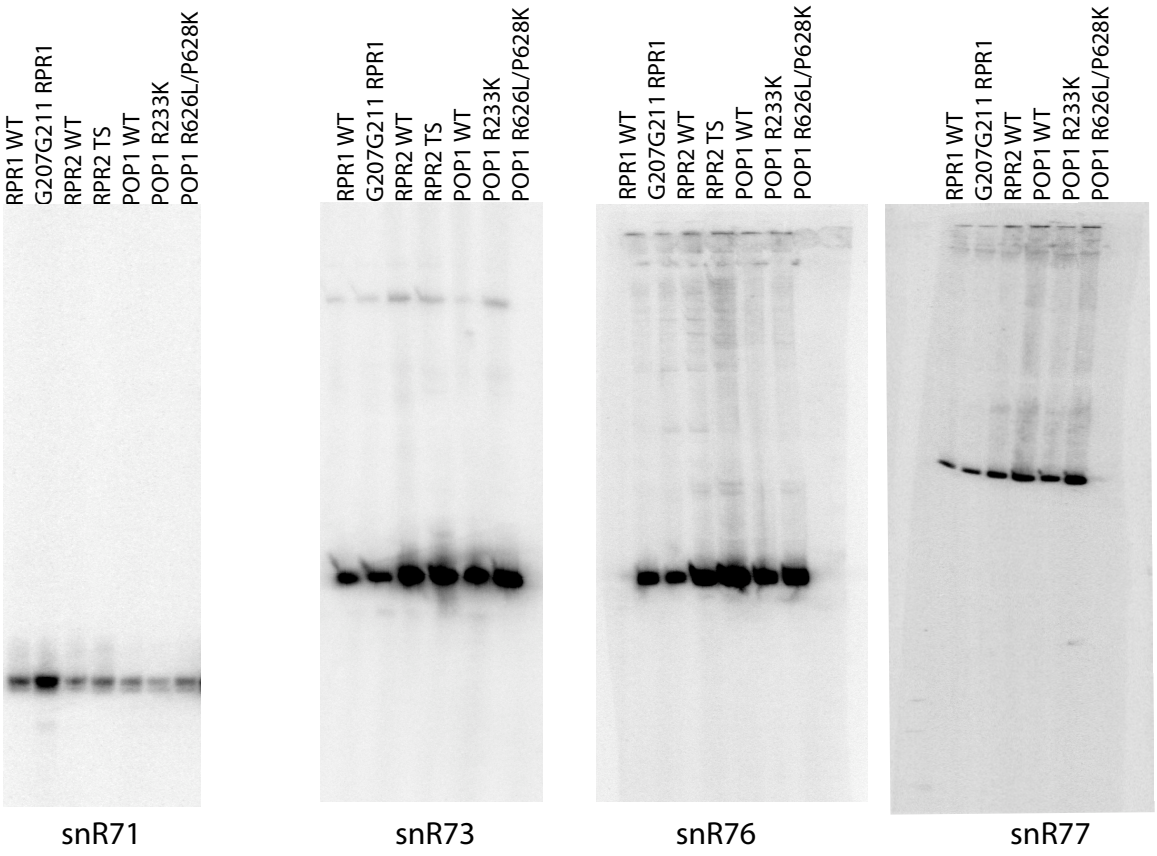


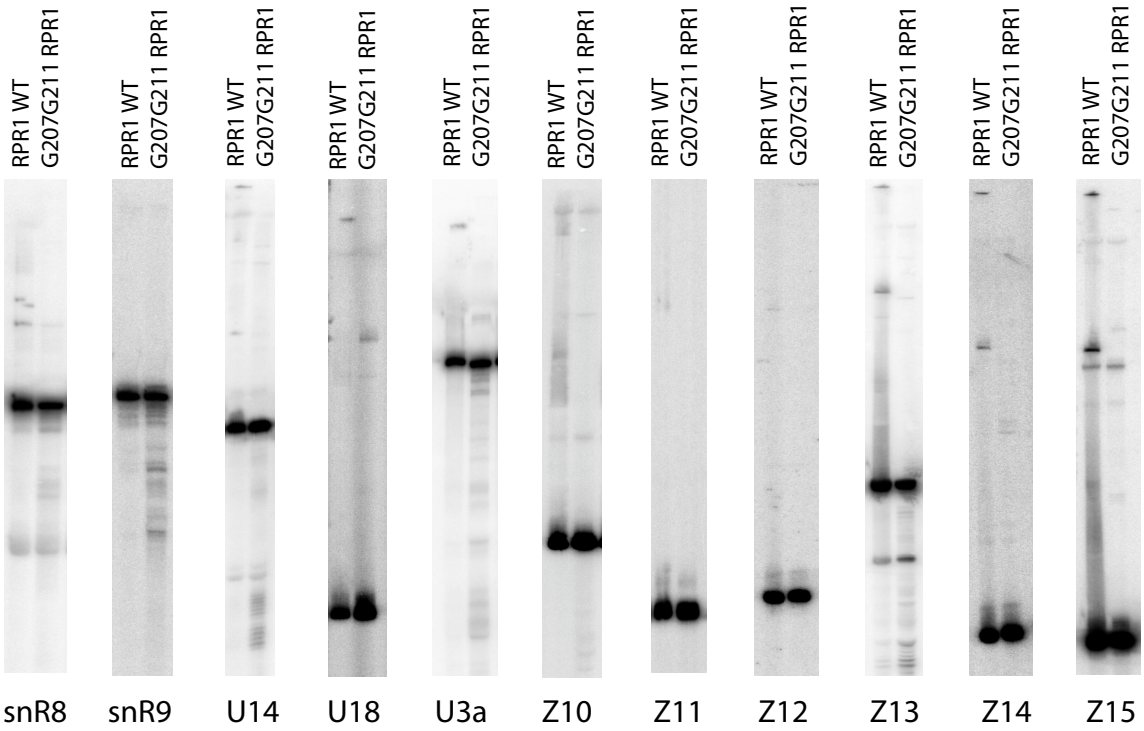

RPR1 WT
G207G211 RPR1
RPR2 WT
RPR2 TS
POP1 WT
POP1 R233K
POP1 R626L/P628K
snR19
snR190
U2
snR30
snR32
snR33
snR35
snR37

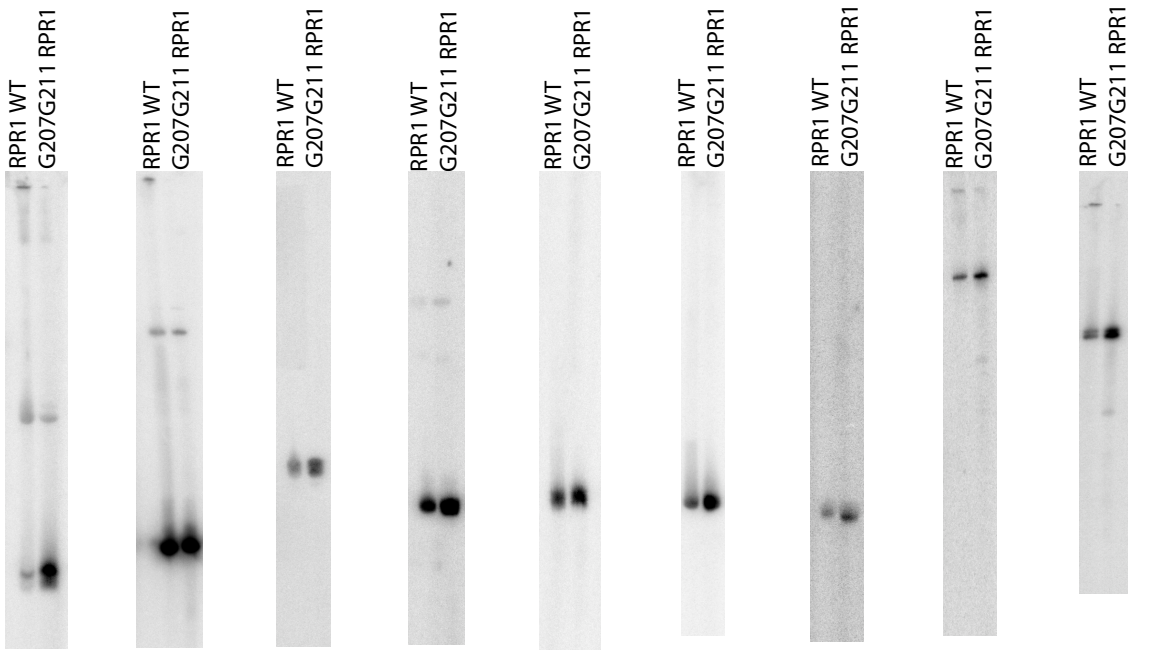

RPR1 WT
G207G211 RPR1
RPR2 WT
RPR2 TS
POP1 WT
POP1 R233K
POP1 R626L/P628K
snR38
snR4
snR40
snR41
snR42
snR43
snR45
snR46


RPR1 WT
G207G211 RPR1
RPR2 WT
RPR2 TS
POP1 WT
POP1 R233K
POP1 R626L/P628K
snR48
snR49
snR5
snR50
snR57
snR56
snR55
snR54


RPR1 WT
G207G211 RPR1
RPR2 WT
RPR2 TS
POP1 WT
POP1 R233K
POP1 R626L/P628K
snR6
snR61
snR63
snR64
snR66
snR68
snR69
snR70


RPR1 WT
G207G211 RPR1
RPR2 WT
RPR2 TS
POP1 WT
POP1 R233K
POP1 R626L/P628K
snR71
snR73
snR76
snR77


RPR1 WT
G207G211 RPR1
snR8
snR9
U14
U18
U3a
Z10
Z11
Z12
Z13
Z14
Z15


RPR1 WT
G207G211 RPR1
Z2
Z3
Z4
Z5
Z6
Z7
Z8
snR84
snR82

Appendix D

Oligonucleotide probe sequences used for northern blot analysis in Appendix C. These oligonucleotide sequences were designed to be complementary to the mature RNA sequence.

Noncoding RNA target	**Probe Sequence (5'-3' complementary to target RNA)**
7SL - SCR1	CGACTCGATATGTGCTATCCCGGCCGCCTCC
snR10	CTGTACGTGTTACGAATGGCTG
snR11	CTATATACGTCCACCGCCTT
snR128 - U14 snRNA	CGTAAGCGTACTCCTACCGTGG
snR13	CCACACCGTTACTGATTTGGC
snR14 - U4 snRNA-Like	TCTCGGACGAATCCTCACTG
snR18 - U18 snoRNA	TCCCATCATAAACACGGACC
snR189	CTCGGTTGTAGAGAGGACGTTGCC
snR19 - U1 snRNA	GATCAGTAGGACTTCTTGATCTCC
snR190	CCCTTGTCGTCATGGTCGAATCGG
snR20 - LSR1 - U2 snRNA	CCTGCGAGAAGAGCTCCTTCTCCTC
snR3	CTAGCAATCCACTCGAGTTC
snR30	GCATCTCTTATGTGATGCCGTTGTCC
snR32	CGTTCAATCTATCTACGCTTCAGTACTAC
snR33	GGCTTTCAATCTCTGCTCCTCC
snR34	CAGTCAACTGTGGCATCGTTTCCGTG
snR35	TGATGATCTCTCCGATGGACTTGACGC
snR36	GTCATCCAGCTCAAGATCGT
snR37	GCTATGGGTTATGATGAAATG
snR38	CCTATTATTACCCATTCAGACAGG
snR4	CGGCACAATCCACATCGACC
snR40	GTGGCATCCATGTTCAGACTG
snR41	CCACTATTCAGTCGGAACAATTG
snR42	CTCATTATCCTTTCTCTATCTCACC
snR43	CGAGACGCCGTCTACGGTTG
snR44	GGATTAAATATCCCGGACAC
snR45	GGTTGCGCAGGAACCGCTATCTCC
snR46	GCTTTGAATCCATAAACCACCGC
snR48	CTTCACATCCTAACATTAGAGATGCC
snR49	GGATTCGTTTACCATAGGCTACC
snR5	GACATATGGAGGCGTGATGTCTTAAGC
snR50	CTGCTGCAAATTGCTACCTC
snR51	GACCAATCTAGTACAGTGTG
snR52	GTATCAGAGATTGTTCACGC
snR54	CGTTTGATCACAGTCAGTAGAACG
snR55	CGATTGTGGTGTCTATTCATC

Noncoding RNA target	Probe Sequence (5'-3' complementary to target RNA)
snR56	GTTCAGTACAGGTCTGTGTT
snR57	GTCCTGCATATACTTCCTCAG
snR58	CTGAGGAAGTATATGCAGGAC
snR59	GACTAGTCGAGAATAAGGAATAG
snR6 - U6	
snRNA	GGAACTGCTGATCATCTCTG
snR60	TCAATCAGTTGAACTATGCATC
snR72	CGTTTTCTTCATTGATGTTCTC
snR61	CTTCCTATTATTTTGGTTCAG
snR63	CCGTGCGTCTGATTATGGTCC
snR64	CTGTTGTCCCTATCTGGTTCC
snR66	GCTTGACTCTGTTGCATTGG
snR68	CCGTCAATACGATAACGCAGT
snR69	GCTGGGTTTATAGCATTGTCACT
snR70	CCTTTAACAGATACTAATATGTCCG
snR71	AAGATCTGAGTGAGCTGAGA
snR75	CGAATGATCAGACTCGTCATC
snR73	ACGCAGTTGACCGTCGTGAA
snR74	GATCAGACATATGCTTGTCT
snR76	GCCCAGTGCTGTGGATCCTC
snR77	CGTTCAGCCAGTAATTCCAGC
snR78	TTATTTTGGTCATCAAGG
snR7-L - U5	
snRNA	CCACAGTTCTTGATGTTGACCTCC
snR8	GGAGTTGCTCTAGCTCTTCTT
snR9	CCACGCTTTCATAGCCATAGAGG
U14	GCGTACTCCTACCGTGGAAACTGCG
U18	CTTCCCATCATAAACACGGACC
U24	CTCAAAGTTCCATCTGAAGTAGC
U3a	GTTGGATTCAGTGGCTCTTTTG
z10 snoRNA	GACGATTGTGGTGTCTATTC
z11 snoRNA	GGTTCAGAAGCAGAACTGAATAG
z12 snoRNA	GGTCTAATCTCCTTCAGAAGTC
z13 snoRNA	GTATCAGAGATTGTTCACGCTA
z14 snoRNA	ATTGCTACCTCTTTCATCAT
z15 snoRNA	CAATCAGTTGAACTATGCATC
z2 snoRNA	CGTTTTCTTCATTGATGTTCTC
z3 snoRNA	AACGCAGTTGACCGTCGTGA
z4 snoRNA	AGATCAGACATATGCTTGTC
z5 snoRNA	CTTCACGAATGATCAGACTCGTCATC
z6 snoRNA	CCTCAGTGCCCAGTGCTGTGGATCC
z7 snoRNA	CAACATATACTCGTTCAGCC
z8 snoRNA	AGAATAAACGTTCTAATCAC
snR84 - RUF1	CCTCAATCATGCCTTTTCTCTCC
snR82 - RUF2	GACGGAAAAGCTAGCTTGGATCC

Appendix E

List of tRNA genes that were identified in the mouse genome and had the expression confirmed. The tRNA gene name is based on tRNA gene nomenclature in yeast. The gene coordinates are based on the May 2005 release of the mouse genome (mm5). The presence of an intron and rogue status are indicated as well.

tRNA gene name	tRNA family	Chromo-some	Sequence start (nt)	Sequence end (nt)	Amino acid type	Anticodon	Contains an intron	Rogue tRNA?	tRNA gene sequence (including introns)
tA(CGC)B	Ala1	2	57,135,446	57,135,520	Ala	CGC			GGGGATGTAGCTCAGTGGTAGAGCGCGCGC TTCGCATGTGTGAGGTCCCAGGTTCAATCCC CGGCATCTCCAAGA
tA(TGC)E1	Ala1	5	122,945,945	122,946,019	Ala	TGC			GGGGATGTAGCTCAGTGGTAGAGCGCATGC TTTGCATGTATGAGGCCCCAGGTTCGATCCC CGGCATCTCCACCA
tA(TGC)E2	Ala1	5	122,951,170	122,951,244	Ala	TGC			GGGGATGTAGCTCAGTGGTAGAGCGCATGC TTTGCATGTATGAGGCCCCAGGTTCAATCCC CGGCATCTCCAACA
tA(TGC)I	Ala1	9	66,457,904	66,457,976	Ala	TGC			GGGGATGTAGCTCAGTGGCAGAGTGCATAC TTTGCATGTATGAGTTACCTGGGTGAAAACT CCAGTATCTCCA
tA(TGC)K	Ala1	11	48,474,160	48,474,236	Ala	TGC			TGGGGATGTAGCTCAGTGGTAGAGCGCATG CTTTGCATGTATGAGGCCCAGGGTTCGATCC CCGGCATCTCCAAACA
tA(AGC)M3	Ala1	13	20,602,634	20,602,708	Ala	AGC			GGGGGTGTAGCTCAGTGGTAGAGCGCGTGC TTAGCATGCACGAGGCCCCAGGTTCAATCCC CGGCACCTCCAGTA
tA(AGC)M4	Ala1	13	20,614,093	20,614,167	Ala	AGC			GGGGGTGTAGCTCAGTGGTAGAGCGCGTGC TTAGCATGCACGAGGCCCCAGGTTCAATCCC CGGCACCTCCAGTA
tA(TGC)M	Ala1	13	20,620,993	20,621,067	Ala	TGC			GGGGGTGTAGCTCAGTGGTAGAGCGCATGC TTTGCATGCATGAGGCCCCAGGTTCGATCCC CGGCACCTCCACTA
tA(AGC)M5	Ala1	13	20,622,617	20,622,693	Ala	AGC			TGGGGGTGTAGCTCAGTGGTAGAGCGCGTG CTTAGCATGCACGAGGCCCAGGGTTCGATCC CCAGCACCTCCATTCA
tA(CGC)M1	Ala1	13	20,634,268	20,634,342	Ala	CGC			GGGGATGTAGCTCAGTGGTAGAGCGCATGC TTCGCATGTATGAGGCCCCAGGTTCGATCCC CGGCATCTCCAAGA
tA(AGC)M6	Ala1	13	20,639,815	20,639,891	Ala	AGC			TGGGGATGTAGCTCAGTGGTAGAGCGCATG CTTAGCATGCATGAGGTCCAGGGTTCGATCC CCAGCATCTCCAGGCA
tA(CGC)M2	Ala1	13	22,818,817	22,818,893	Ala	CGC			TGGGGATGTAGCTCAGTGGTAGAGCGCATG CTTCGCATGTATGAGGCCCAGGGTTCGATCC CCGGCATCTCCAGTTA
tV(TAC)X1	Ala1	X	126,337,960	126,338,036	Val	TAC		Rogue	TGGAGGTGTAGCTCAATGTCAGAGCTCTTGA TTTACATGTATGGGGTTCAGGGTTCGATTTCT GGCATTTCCAGATA
tA(TGC)X2	Ala1	X	126,364,763	126,364,839	Ala	TGC			TGGGGATGTAGCTCAGTGGTAGAGCACATG CTTTGCATGTATGGGGTCCAGGGTTCAATTC CCGGCATCTCCAGAGA
tA(CGC)X1	Ala1	X	126,374,882	126,374,958	Ala	CGC			TGGGGTTGTAGCTCAGAGGTAGAGCACATG CTTCGCATGTGTGTGGTCCAGGGTTCGATTC CCGGCATCTCCAGAGA
tA(CGC)X3	Ala1	X	126,386,524	126,386,600	Ala	CGC			TGGGGATGTAGCTCAGAGGTAGAGCACATG CTTCGCATGTGTGTGGTCCAGGGTTCGACTC CCGGCATCTCCAGAGA
tA(TGC)X4	Ala1	X	126,387,908	126,387,984	Ala	TGC			TGGGGATGTAGCTCAGCGGTAGAGCACATG CTTTGCATGTATGGGGTCCAGGGTTCGATTC CCGGCATCTCCACAGA
tA(CGC)X5	Ala1	X	126,402,680	126,402,756	Ala	CGC			TGGGGATGTAGCTCAGAGGTAGAGCACATG CTTCGCATGTGTGTGGTCCAGGGTTCGACTC CCGGCATCTCCAGAGA
tA(TGC)X6	Ala1	X	126,404,065	126,404,141	Ala	TGC			TGGGGATGTAGCTCAGCGGTAGAGCACATG CTTTGCATGTATGGGGTCCAGGGTTCGATTC CCGGCATCTCCACAGA
tA(CGC)X7	Ala1	X	126,409,029	126,409,105	Ala	CGC			TGGGGATGTAGCTCAGAGGTAGAGCACATG CTTCGCATGTGTGTGGTCCAGGGTTCGACTC CCGGCATCTCCAGAGA
tA(TGC)X8	Ala1	X	126,421,635	126,421,711	Ala	TGC			TGGAAGTGTAGTTCAATGGTAGAACCCTTGT TTTGCATATATGGGGTTCAGGGTTCGGTTCCT GGCACCACTAGGTA
tA(TGC)X10	Ala1	X	126,437,523	126,437,599	Ala	TGC			TGGGGATGTAGCTCAGTGGTAGAGCACATG CTTTGCATGTATGGGGTCCAGGGTTCGATTC CCGGCATCTCCAATGA
tA(TGC)X12	Ala1	X	126,560,886	126,560,962	Ala	TGC			TGGGGATGTAGCTCAGTGGTAGAGCACATG CTTTGCATGTATGGGGTCCAGGGTTCGATTC CCGGCATCTCCAATGA

tRNA gene name	tRNA family	Chromo-some	Sequence start (nt)	Sequence end (nt)	Amino acid type	Anticodon	Contains an intron	Rogue tRNA?	tRNA gene sequence (including introns)
tA(CGC)X9	Ala1	X	126,575,023	126,575,099	Ala	CGC			TGGGGATGTAGCTCAGAGGTAGAGCACATG CTTCGCATGTGTGTGGTCCAGGGTTCGACTC CCAGCATCTCCAGAGA
tA(TGC)X14	Ala1	X	126,576,407	126,576,483	Ala	TGC			TGGGGATGTAGCTCAGCGGTAGAGCACATG CTTTGCATGTATGGGGTCCAGGGTTCGATTC CCGGCATCTCCAGAGA
tA(CGC)X11	Ala1	X	126,576,924	126,577,001	Ala	CGC			TGGGGATGTAGCTCAGAGGTAGAGCACGTG CTTCGCATGTGTGTGGTCCAGGGGTTCGATT CCCGGCATCTCCAGGGA
tA(TGC)X16	Ala1	X	126,597,526	126,597,602	Ala	TGC			TGGGGATGTAGCTCAATGGTAGAGCACATGC TTTGCATGTATGAGGCACAGGGTTCGATTCC CGGCATCTCCAGGGA
tA(AGC)A3	Ala2	1	145,819,931	145,820,003	Ala	AGC			TTGGGGATTTAGCTCAGTGGTAGAGCGCTTG CCTAGCAAGCACAAGGCCCTGGGTTCGGTC CTCAGCTCTGAA
tA(AGC)B4	Ala2	2	120,194,117	120,194,189	Ala	AGC			TCAGGGATTTAGCTCAGTGGTAGAGCGCTTG CCTAGCAAGTGCAAGGCCCTGGGTTCGGTC CTCAGCTCTGAA
tA(AGC)E7	Ala2	5	91,844,561	91,844,635	Ala	AGC			TTGGGGATTTAGCTCAGAGGTAGAGAGCGC TTGCTTAGCAAGCACAAGGTCCTGGGTTCGA TCCTCAGCTCTGAA
tA(AGC)F4	Ala2	6	97,626,627	97,626,698	Ala	AGC			GGGATTTAGCTCAGTGGTAGAGCGCTGCTTG CCTAGCAAGCACAAGGCCCTGGGTTCGGTC CTCAGCTCTTG
tA(AGC)F5	Ala2	6	116,211,554	116,211,626	Ala	AGC			TTGGGGACTTAGCTCAGTGGTAGAGCAGTT GCCTAGCAAGCACAAGGCCCTGGGTTCAGT CCTCAGCTCTGAA
tA(AGC)G3	Ala2	7	32,467,780	32,467,852	Ala	AGC			TTGGGGATTTAGCTCAGTGGTAGAGTGCGTG CCTAGCAAGCACAAGGCCCTGGGTTCTGTC CTCAGCTCTGAA
tA(AGC)G4	Ala2	7	45,399,804	45,399,876	Ala	AGC			TTGGAGATTTAGCTCAGTGGTAGAGCACTTG CCTAGCAAGCCCAAGGCCCTGAGTTCGGTC CTCAGCTCTGAA
tV(AAC)H	Ala2	8	4,082,112	4,082,184	Val	AAC		Rogue	TTGGGGATTTAGCTCAGTGGTAGAGCGCTTG CCTAACAAAAGCAAGGCTCTGGGTTCGGTC CTCAGCTCTGAA
tA(AGC)H1	Ala2	8	32,175,313	32,175,379	Ala	AGC			GGATTTAGCTCAGTGGTAGAGCGCTTGCCTA GCAAGCACAAGGCCCTGGGTTCGGTCCTCA GCTCTG
tA(AGC)I2	Ala2	9	55,652,779	55,652,851	Ala	AGC			TTGGGGATTTAGCTCAGTGGTAGAGCGCTTG CCTAGCAAGCACAAGGCCCTGGGTTCAGTC CTCAGCTCTGAA
tA(AGC)I6	Ala2	9	95,869,025	95,869,097	Ala	AGC			TTGGGGATTTAGCTCGGTGGTAGAGCGCTTG CCTAGCAAGTGCAAGGCCCTGGGTTCGGTC CTCAGCTCTGAA
tA(AGC)J1	Ala2	10	43,319,775	43,319,847	Ala	AGC			TTGGGGATTTAGCTCAGTGGTAGAGCACTTG CCTAGCAAGTGCAAGGCCCTGGGTTCGGTC CTCAGCTCTGAA
tA(AGC)J4	Ala2	10	72,310,320	72,310,386	Ala	AGC			AGAAATAGCTCAGTGGTAGAGCGCTTGCCT AGCAAGCACAAGGCCCTGGGTTCGGTCCTC AGCTCTG
tA(AGC)J5	Ala2	10	80,239,235	80,239,307	Ala	AGC			TTGGGGATTTAGCTCAGTGGTAGAGCGCTCA CCCAGCAAGCAAAAGGCCCTGGGTTCAGTC CTCAGCTCTGAA
tA(AGC)K4	Ala2	11	61,980,219	61,980,293	Ala	AGC			TTGGGGATTTAGCTCAGTGGTAGAGCGCTTG CCTAGCAAACTCAAGGCTCTGAGTTCGGGC CTCAGTTCTTTAAA
tA(AGC)L2	Ala2	12	21,906,568	21,906,640	Ala	AGC			TTGGGGATTTAGCTCAGTGGTAGAGCACTTG CCTAGCAAGCACAAGGCCCTGGGTTCGGTT CTCAGCTCTGAA
tA(AGC)L3	Ala2	12	32,067,034	32,067,102	Ala	AGC			GGGGATTTAGCTAAGTGGTAGAGTGCTTGCT TAGCAAGTACAAGGCACTGAGTTCGATCCT CAGCCCCA
tA(AGC)O1	Ala2	15	41,259,553	41,259,624	Ala	AGC			TTGGGGATTTAGCTCAGTGGTAGAGCACTTG CCCAGCAAGCACAAGGCTTGGGTTCAGTCC TCAGCTCTGAA
tA(AGC)O2	Ala2	15	47,814,572	47,814,646	Ala	AGC			TTTGGGGATTTAGCTCAGTGGTAGAGCGTTT GCCTAGCAAGCACAAGGCCTTGGGTTCAGT CCTCAGCTCTAGAA
tA(AGC)O3	Ala2	15	63,206,629	63,206,701	Ala	AGC			TTGGGGATTTAGCTCAGTGGTAGAGCACTTG CCTAGCAAGCACAAGGCCCTGGGTTCAGTC CTCAGCTCTGAA
tA(AGC)P	Ala2	16	25,074,149	25,074,221	Ala	AGC			TTGGGGATTTAACTCAGTGGTAGAGCACTTG CCTAGCAAGCACAAGGCCCTGGGTTCAGTC CTCAGCTCTGAA
tA(AGC)Q1	Ala2	17	30,416,505	30,416,571	Ala	AGC			GGGGATTTAGCTCAGTGGTAGAGCGTGCCTA GCAAGCACAAGGCCCTGGGTTCTATCCTCA GCTCCA
tA(AGC)S4	Ala2	19	45,310,870	45,310,942	Ala	AGC			TTGGGGATTTAGCTCAGTGGTAGAGCACTTG TCTAGCAAGCACAAGGCCTTGGGTTCGGTC CTCAGCTCCAAA

tRNA gene name	tRNA family	Chromo-some	Sequence start (nt)	Sequence end (nt)	Amino acid type	Anticodon	Contains an intron	Rogue tRNA?	tRNA gene sequence (including introns)
tG(ACC)X1	Ala2	X	21,564,907	21,564,980	Gly	ACC		Rogue	TTGGGGATTTAGCTCAGTGGTAGAGCGTTGC CTACCAAGCACAAGGCCCTGGGTTCAGTCC TCAGCTCTGGAAA
tA(AGC)X	Ala2	X	134,259,185	134,259,257	Ala	AGC			TTGGGGATTTAGCTCAGTGGTAGAGCACTTG CCTAGCAAGAGCAAGGCCCTGGGTTTGGTC CTCAGCCCTGAA
tA(AGC)C1	Ala3	3	19,303,425	19,303,500	Ala	AGC			GGGGGATTAGCTCAAATGGTAGAGCGCTCG CTTAGCATGCGAGAGGTAGAGGGATCGATG CCCGCATCCTCCAGTA
tA(AGC)E3	Ala3	5	29,296,699	29,296,776	Ala	AGC			TGGGGGATTAGCTCAAATGGTAGAGCGCTC GCTTAGCATGCGAGAGGTAACGGGATCGAT GCCCGCATCCTCCACTTA
tA(AGC)M7	Ala3	13	22,665,773	22,665,848	Ala	AGC			GGGGGATTAGCTCAAATGGTAGAGCGCTCG CTTAGCATGCGAGAGGTAGAGGGATCGATG CCCACATCCTCCAGTA
tA(AGC)M8	Ala3	13	22,679,185	22,679,260	Ala	AGC			GGGGGATTAGCTCAAATGGTAGAGCGCTCG CTTAGCATGCGAGAGGTAGAGGGATCGATG CCCACATCCTCCAGTA
tA(AGC)M9	Ala3	13	22,780,916	22,780,991	Ala	AGC			GGGGGATTAGCTCAAATGGTAGAGCGCTCG CTTAGCATGCGAGAGGTAGAGGGATCGATG CCCACATCCTCCACAA
tA(AGC)M10	Ala3	13	22,793,061	22,793,138	Ala	AGC			TGGGGGATTAGCTCAAATGGTAGAGCGCTC GCTTAGCATGCAAGAGGTAATGGGATCGATG CCCACATCCTCCAGCTA
tA(AGC)M11	Ala3	13	22,811,536	22,811,611	Ala	AGC			GGGGAATTAGCTCAAATGGTAGAGCGCTCG CTTAGCATGCGAGAGGTAGAGGGATCGATG CCCGCATTCTCCAGTA
tA(AGC)S1	Ala3	19	3,364,636	3,364,713	Ala	AGC			TGGGGAATTAGCTCAAATGGTAGAGCGCTC GCTTAGCATGTGAGAGGTAACGGGATCGATG CCCGCATTCTCCATAGA
tR(ACG)C	Arg1	3	19,302,385	19,302,462	Arg	ACG			CGGGCCAGTGGCGCAATGGATAACGCGTCT GACTACGGATCAGAAGATTATAGGTTCGACT CCTACCTGGCTCGAGCA
tR(TCG)G	Arg1	7	66,546,217	66,546,294	Arg	TCG			CGGCCGCGTGGCCTAATGGATAAGGCGTCTG ACTTCGGATCAGAAGATTACAGGTTCGAGTC CTGCCGCGGTCGACTA
tR(ACG)I	Arg1	9	123,543,770	123,543,847	Arg	ACG			CGGGCCAGTGGCGCAATGGATAACGCGTCT GACTACGGATCAGAAGATTATAGGTTCGACT CCTGGCTGGCTCGACGA
tR(CCG)K	Arg1	11	106,683,990	106,684,067	Arg	CCG			TGACCCAGTGGCCTAATGGATAAGGCATCAG CCTCCGGAGCTGGGGATTATGGGTTCGAGTC CCATCTGGGTCGCATA
tR(TCG)K	Arg1	11	115,084,908	115,084,983	Arg	TCG			GACCGCGTGGCCTAATGGATAAGGCGTCTG ACTTCGGATCAGAAGATTGAGGGTTCGAGT CCCTTCGTGGTCGAAA
tR(CCG)M	Arg1	13	20,595,502	20,595,579	Arg	CCG			TGGCCGCGTGGCCTAATGGATAAGGCGTCTG ATTCCGGATCAGAAGATTAAGGGTTCGAGTC CCTTCGTGGTCGGTGA
tR(TCG)M1	Arg1	13	20,659,743	20,659,820	Arg	TCG			CGACCACGTGGCCTAATGGATAAGGCGTCTG ACTTCGGATCAGAAGATTAAGGGTTCGAATC CCTTCGTGGTTGGTAA
tR(ACG)M1	Arg1	13	21,401,264	21,401,341	Arg	ACG			CGGGCCAGTGGCGCAATGGATAACGCGTCT GACTACGGATCAGAAGATTATAGGTTCGACT CCTGGCTGGCTCGACCA
tR(ACG)M2	Arg1	13	22,826,913	22,826,988	Arg	ACG			GGGCCAGTGGCGCAATGGATAACGCGTCTG ACTACGGATCAGAAGATTCAAGGTTCGACT CCTGGCTGGCTCGGAA
tR(ACG)M3	Arg1	13	22,880,642	22,880,719	Arg	ACG			TGGGCCAGTGGCGCAATGGATAACGCGTCT GACTACGGATCAGAAGATTACAGGTTCGAC TCCTGGCTGGCTCGGTAA
tR(TCG)M2	Arg1	13	22,884,101	22,884,176	Arg	TCG			GACCACGTGGCCTAATGGATAAGGCGTCTG ACTTCGGATCAGAAGATTGAGGGTTCGAATC CCTTCGTGGTTGAGA
tR(TCG)M3	Arg1	13	22,906,576	22,906,653	Arg	TCG			TGACCACGTGGCCTAACGGATAAGGCGTCT GACTTCGGATCAGAAGATTAAGGGTTCGAAT CCCTTCGTGGTTACTTA
tR(ACG)N	Arg1	14	46,491,443	46,491,520	Arg	ACG			CGGGCCAGTGGCGCAATGGATAACGCGTCT GACTACGGATCAGAAGATTACAGGTTCGACT CCTGGCTGGCTCGTTGA
tR(CCG)Q	Arg1	17	22,146,189	22,146,264	Arg	CCG			GGCCGCGTGGCCTAATGGATAAGGCGTCTGA TTCCGGATCAGAAGATTGAGGGTTCGAGTC CCTTCGTGGTCGCTA
tR(TCT)A	Arg2	1	173,520,466	173,520,542	Arg	TCT			GTCTCTGTGGCGCAATGGACGAGCGCGCTG GACTTCTAATCCAGAGGTTATGGGTTCGAGT CCCGGCAGAGATGATA
tR(TCT)C	Arg2	3	122,675,382	122,675,471	Arg	TCT	Intron		TGGCTCCGTGGCGCAATGGATAGCGCATTGG ACTTCTAGAGGCTGAAGGAATTCAAAGGTT CCGGGTTCGAGTCCCGGCGGAGTCGGAAA

tRNA gene name	tRNA family	Chromo-some	Sequence start (nt)	Sequence end (nt)	Amino acid type	Anticodon	Contains an intron	Rogue tRNA?	tRNA gene sequence (including introns)
tR(CCT)F	Arg2	6	38,462,972	38,463,047	Arg	CCT			GCCCCAGTGGCCTAATGGATAAGGCATTGGC CTCCTAAGCCAGGGATTGAGGGTTCGAGTC CCATCTGGGGTGGTA
tR(TCT)I	Arg2	9	83,459,349	83,459,437	Arg	TCT	Intron		GGCTCTGTGGAGCAATGGATAGCACATTGG ACTTCTAGCATGACCGAGAAATTCAAAGGTT GCGGGTTCGAGTCCCACCAGAGTTGACA
tR(TCT)K	Arg2	11	68,736,034	68,736,125	Arg	TCT	Intron		TGGCTCTGTGGCGCAATGGATAGCGCATTGG ACTTCTAGTGACGAGAAAACGATTCAAAGG TTGTGGGTTCGAATCCCACCAGAGTCGGTTA
tR(CCT)K1	Arg2	11	115,084,110	115,084,185	Arg	CCT			GCCCCAGTGGCCTAATGGATAAGGCACTGG CCTCCTAAGCCAGGGATTGAGGGTTCGAGT CCCACCTGGGGTAAAA
tR(CCT)K2	Arg2	11	115,084,429	115,084,506	Arg	CCT			CGCCCCAGTGGCCTAATGGATAAGGCACTG GCCTCCTAAGCCAGGGATTATGGGTTCGAGT CCCACCTGGGGTGGTGA
tR(TCT)M	Arg2	13	21,297,283	21,297,371	Arg	TCT	Intron		GGCTCTGTGGCGCAATGGATAGCGCATTGGA CTTCTAGCATGATTGAGAAATTCAAAGGTTG CGGGTTCGAGTCCCGCCAGAGTCGACA
tR(CCT)Q	Arg2	17	22,144,694	22,144,769	Arg	CCT			GCCCCGGTGGCCTAATGGATAAGGCATTGGC CTCCTAAGCCAGGGATTGAGGGTTCGAGTC CCACCCGGGGTAATA
tR(TCT)S	Arg2	19	11,261,938	11,262,028	Arg	TCT	Intron		TGGCTCTGTGGCGCAATGGATAGCGCATTGG ACTTCTAGACAAATGGAGACATTCAAAGGT TGTGGGTTCGAGTCCCACCAGAGTCGCTTA
tR(CCT)X2	Arg2	X	102,704,669	102,704,744	Arg	CCT			GCCCCAGTGGCCTAATGGATAAGGCACTGG CCTCCTAAGCCAGGGATTGAGGGTTCGAGT CCCACCTGGGGTGGTA
tN(GTT)A1	Asn	1	171,111,504	171,111,580	Asn	GTT			GTCTCTGTGGCGCAATCGGTTAGCGCGTTCG GCTGTTAACCGAAAGGTTAGTGGTTCGAGC CCACCCAGGGACGGTA
tN(GTT)A2	Asn	1	188,611,697	188,611,775	Asn	GTT			TGCCTCTGTGGTGCAATTGGTTAGCATGTTC CGCTGTTAACCAAAAGGTAGGTGGTTCGAG CCCACCCAGGGACAACAA
tN(GTT)B	Asn	2	18,642,798	18,642,874	Asn	GTT			GTCTCTGTGGCGCAATCGGTTAGCGCGTTCG GCTGTTAACCGAAAGGTTAGTGGTTCGAGC CCACCCAGGGACGCCA
tN(GTT)C1	Asn	3	96,327,800	96,327,878	Asn	GTT			TGTCTCTGTGGCGCAATTGGTTAGCGCGTTC GGCTGTTAACCGAAAGGTAGGTGGTTCAAG CCCACCCAGGGACGTCGA
tN(GTT)C2	Asn	3	96,340,339	96,340,417	Asn	GTT			TGTCTCCGTGGCGCAATCGGTCAGCGCGTTC GGCTGTTAACCGAAAGGTAGGTGGTTCGAG CCCACCCGGGGACGTGTA
tN(GTT)C3	Asn	3	96,340,792	96,340,870	Asn	GTT			TGTCTCTGTGGTGCAATCGGTCAGCGCGTTT GGCTGTTAACTGGAAGGTAGGTGGTTTGAG CCCACCCGGGGACGTGTA
tN(GTT)C4	Asn	3	96,403,400	96,403,478	Asn	GTT			CGTCTCTGTGGCGCAATCGGTTAGCGCGTTC GGCTGTTAACCGAAAGGTAGGTGGTTCGAG CCCACCCAGGGACGAGCA
tN(GTT)C5	Asn	3	96,453,176	96,453,254	Asn	GTT			CGTCTCTGTGGCGCAATCGGTTAGCGCGTTC GGCTGTTAACCGAAAGGTAGGTGGTTCGAG CCCACCCAGGGACGACAA
tN(GTT)C6	Asn	3	96,501,560	96,501,638	Asn	GTT			TGTCTCTGTGGCGCAATCGGTTAGCGCGTTC GGCTGTTAACCGAAAGGTAGGTGGTTCGAG CCCACCCAGGGACGGTGA
tN(GTT)E3	Asn	5	146,564,060	146,564,138	Asn	GTT			TGTCTCTGTGGCGCAATCGGTTAGCGCGTTC GGCTGTTAACCGAAAGGTAGGTGGTTCGAG CCCACCCAGGGACGGCAA
tN(GTT)H	Asn	8	120,076,832	120,076,909	Asn	GTT			GTCTCTGTCATGCAGTCGGTTAGTGAGCGAG TTAGGCTGTTAACCGGAAGGTTGGTGGTTCG AGCCCACCCAAGGACA
tN(GTT)J	Asn	10	80,150,086	80,150,164	Asn	GTT			CGTCTCTGTGGCGCAATCGGTTAGCGCGTTC GGCTGTTAACCGAAAGGTAGGTGGTTCGAG CCCACCCAGGGACGACCA
tN(GTT)K	Asn	11	97,373,568	97,373,646	Asn	GTT			CGTCTCTGTGGCGCAATCGGTTAGCGCGTTC GGCTGTTAACCGAAAGGTAGGTGGTTCGAG CCCACCCAGGGACGGTCA
tN(GTT)O	Asn	15	16,185,170	16,185,246	Asn	GTT			GTCTCTGTGGCGCGATCGGTTAGTGCGTTTG GCTGTTAACCGAAAGGTTAGTGGTTCAAAC CCACCCGGGGACGACA
tD(GTC)A1	Asp	1	171,114,591	171,114,665	Asp	GTC			TCCTCGTTAGTATAGTGGTGAGTATCCCCGC CTGTCACGCGGGAGACCGAGGTTCGATTCC CCGACGGGGAGACA
tD(GTC)A2	Asp	1	171,143,364	171,143,438	Asp	GTC			TCCTCGTTAGTATAGTGGTGAGTATCCCCGC CTGTCACGCGGGAGACCGAGGTTCGATTCC CCGACGGGGAGAGA
tD(GTC)A3	Asp	1	171,173,626	171,173,700	Asp	GTC			TCCTCGTTAGTATAGTGGTGAGTATCCCCGC CTGTCACGCGGGAGACCGAGGTTCGATTCC CCGACGGGGAGATA

tRNA gene name	tRNA family	Chromo-some	Sequence start (nt)	Sequence end (nt)	Amino acid type	Anticodon	Contains an intron	Rogue tRNA?	tRNA gene sequence (including introns)
tD(GTC)B	Asp	2	113,884,942	113,885,018	Asp	GTC			TTCCTCGTTAGTATAGTGGTGAGTATCCCTGC CTGTCACGCAGGACACCAGGGTTCGATTTC CTGACGGGGAGGCAA
tD(GTC)E1	Asp	5	98,329,089	98,329,163	Asp	GTC			TCCTCGTTAGTATAGTGGTGAGTATCCCCGC CTGTCACACAGGAGACCGAGGTTCGATTCC CCGATGGGGAGAGA
tD(GTC)E2	Asp	5	122,947,301	122,947,377	Asp	GTC			TTCCTCGTTAGTATAGTGGTGAGTATCCCCGC CTGTCACGCGGGAGACCAGGGTTCGATTCC CCGACGGGGAGATTA
tD(GTC)E3	Asp	5	122,950,711	122,950,785	Asp	GTC			TCCTCGTTAGTATAGTGGTGAGTATCCCCGC CTGTCACGCGGGAGACCGAGGTTCGATTCC CCGACGGGGAGGCA
tD(GTC)J1	Asp	10	90,952,735	90,952,811	Asp	GTC			CTCCTCGTTAGTATAGTGGTTAGTATCCCCGC CTGTCACGCGGGAGACCAGGGTTCAATTCC CCGACGGGGAGGTTA
tD(GTC)J2	Asp	10	93,224,814	93,224,890	Asp	GTC			TTCCTCGTTAGTATAGTGGTGAGTATCCCCGC CTGTCACGCGGGAGACCAGGGTTCGATTCC CCGACGGGGAGGTAA
tD(GTC)K1	Asp	11	7,764,570	7,764,644	Asp	GTC			TCCTCGTTAGTATAGTGGTGAGTATCCCCGC CTGTCACGCGGGAGACCGAGGTTCGATTCC CCGACGGGGAGATA
tD(GTC)K2	Asp	11	68,653,539	68,653,615	Asp	GTC			CTCCTCGTTAGTATAGTGGTGAGTATCCCCG CCTGTCACGCGGGAGACCAGGGTTCGATTC CCCGACGGGGAGAAGA
tD(GTC)M1	Asp	13	21,291,436	21,291,512	Asp	GTC			CTCCTCGTTAGTATAGTGGTGAGTATCCCCG CCTGTCACGCGGGAGACCAGGGTTCGATTC CCCGACGGGGAGTCTA
tD(GTC)M2	Asp	13	21,315,617	21,315,693	Asp	GTC			TTCCTCGTTAGTATAGTGGTGAGTATCCCCGC CTGTCACGCGGGAGACCAGGGTTCGATTCC CCGACGGGGAGAAGA
tD(GTC)M3	Asp	13	21,320,440	21,320,514	Asp	GTC			TCCTCGTTAGTATAGTGGTGAGTATCCCCGC CTGTCACGCGGGAGACCGAGGTTCGATTCC CCGACGGGGAGAGA
tC(GCA)F3	Cys	6	47,887,487	47,887,563	Cys	GCA			AGGGGGTATAGCTCAGTGGTAGAGCATTTGA CTGCAGATCAAGAGGTCCACGGTTCAAATC CGGGTGCCCCCTTACA
tC(GCA)F4	Cys	6	48,042,818	48,042,892	Cys	GCA			GGGGGTATAGCTCAGGGGTAGAGCATTTGA CTGCAGATCAAGAGGTCCCAGGTTCAAATC CAGGTGCCCCCTTAA
tC(GCA)F5	Cys	6	48,070,603	48,070,677	Cys	GCA			GGGGGTATAGCTCAGGGGTAGAGCATTTGA CTGCAGATCAAGAGGTCCCAGGTTCAAATC CAGGTGCCCCCTACA
tC(GCA)F6	Cys	6	48,071,598	48,071,672	Cys	GCA			GCGGGTATAGCTCAGGGGTAGAATATTTGAC TGCAGATCAAGAGGTCCCATGTTCGAATTCA GGTGCCCTCTTTA
tC(GCA)F7	Cys	6	48,074,525	48,074,599	Cys	GCA			GGGGGTATAGCTCAGGGGTAGAGCATTTGA CTGCAGATCAAGAGGTCCCAGGTTCAAATC CAGGTGCCCCCTATA
tC(GCA)F8	Cys	6	48,096,655	48,096,731	Cys	GCA			GGGGGGTAAAGCTCAGGGGTAGAGCATTTG ACTGCAGATTAAGAGGTCCATGGTTCAAATC CAGGTACCCCCTGTTA
tC(GCA)F9	Cys	6	48,107,364	48,107,438	Cys	GCA			GGGGGTATAGCTCAGGGGTAGAGCATTTGA CTGCAGATCAAGAGGTCCCAGGTTCAAATC CAGGTGCCCCCTTGA
tC(GCA)F10	Cys	6	48,113,324	48,113,398	Cys	GCA			GGGTGTGTGGATCATGGGTAGAGCATTTAAC TGCACATCAAGAGGTGTCAGGTTCAAATCC AGATGTCCCCTATA
tC(GCA)F11	Cys	6	48,114,641	48,114,717	Cys	GCA			AGGGGGTATAGCTCAGGGGTGGAGCATTTG ACTGCAGATCAAGGGGTCCATGTTTCAAATC CAGGTGCCCCCTAATA
tC(GCA)F12	Cys	6	48,117,022	48,117,096	Cys	GCA			GGGGGTATAGCTCAGGGGTAGAGCATTTGA CTGCAGATCAAGAGGTCCCAGGTTCAAATC CAGGTGCCCCCTACA
tC(GCA)F13	Cys	6	48,119,409	48,119,483	Cys	GCA			GGGGGTATAGCTCAGGGGTAGAGCATTTGA CTGCAGATCAAGAGGTCCCAGGTTCAAATC CGGGTGCCCCCTACA
tC(ACA)F1	Cys	6	48,124,169	48,124,238	Cys	ACA			GGAGGCATAGCTCAGAGGTAGAGCATTTGA CTACAGATCAGCAGGACCCTGATTCAAATCA GGTGCCCTG
tC(GCA)F14	Cys	6	48,133,976	48,134,050	Cys	GCA			GGGGGCATAGCTCAGGGGTAGAGCATTTGA CTGCAGATCAAGAGGTCCCAGGTTCAAATC CGGGTGCTCCCTTCA
tC(GCA)F15	Cys	6	48,160,950	48,161,024	Cys	GCA			GGGGGTATAGCTCAGGGGTAGAGCATTTGA CTGCAGATCAAGAGGTCCCAGGTTCAAATC CAGGTGCCCCCTACA
tC(GCA)F16	Cys	6	48,168,737	48,168,811	Cys	GCA			GGGGGTATAGCTCAGGGGTAGAGCATTTGA CTGCAGATCAAGAGGTCCCAGGTTCAAATC CAGGTGCCCCCTTCA
tC(GCA)F17	Cys	6	48,169,714	48,169,788	Cys	GCA			GGGGGTATAGCTCAGGGGTAGAGCATTTGA CTGCAGATCAAGAGGTCCTAGGTTCAAATCC AGGTGTCCCCTACA

tRNA gene name	tRNA family	Chromo-some	Sequence start (nt)	Sequence end (nt)	Amino acid type	Anticodon	Contains an intron	Rogue tRNA?	tRNA gene sequence (including introns)
tC(GCA)F18	Cys	6	48,178,369	48,178,445	Cys	GCA			AGGGGGTATAGCTCAGGGGTAGAGCATTTG ACTGCAGATCAAGAGGTCCATGGTTCAAATC CAGGTGCCCCCTTTTA
tC(GCA)F19	Cys	6	48,186,762	48,186,838	Cys	GCA			AGGGGGTATAGCTCAGGGGTAGAGCATTTG ACTGCAGATCAAGAGGTCCATGGTTCAAATC CAGGTGCCCCCTTTTA
tC(GCA)F20	Cys	6	48,190,235	48,190,309	Cys	GCA			GGGGGTATAGCTCAGGGGTAGAGCATTTGA CTGCAGATCAAGAGGTCCCAGGTTCAAATC CAGGTGCCCCCTTAA
tC(GCA)F21	Cys	6	48,191,990	48,192,064	Cys	GCA			GGGGGTATAGCTCAGGGGTAGAGCATTTGA CTGCAGATCAAGAGGTCCCAGGTTCAAATC CAGGTGCCCCCTTAA
tC(GCA)F22	Cys	6	48,196,271	48,196,345	Cys	GCA			GGGGGAATAGCTTAGGGGTAGAGCATTTGA ATGCAGATCAAGAGGTCTCAGGTTCAAATCC AGGTGCCCCCTTTA
tC(GCA)F23	Cys	6	48,197,250	48,197,324	Cys	GCA			GGGGGTATAGCTCAGGGGTAGAGCATTTGA CTGCAGATCAAGAGGTCCCAGGTTCAAATC CAGGTGCCCCCTACA
tC(GCA)F24	Cys	6	48,206,355	48,206,429	Cys	GCA			GGGGGTATAGCTCAGGGGTAGAGTATTTGGC TGCAGATCAAGAGGTCCCAGGTTCAAATCC AGGTGCCCCCTTTA
tC(GCA)F25	Cys	6	48,210,487	48,210,561	Cys	GCA			GGGGGTATAGCTCAGGGGTAGAGCATTTGA CTGCAGATCAAGAGGTCCCAGGTTCAAATC CAGGTACCCCCTATA
tC(GCA)F26	Cys	6	48,214,916	48,214,990	Cys	GCA			GGGGGTATAGCTCAGGGGTAGAGCATTTGA CTGCAGATCAAGAGGTCCCAGGTTCAAATC CAGGTGCCCCCTTTA
tC(GCA)F27	Cys	6	48,216,430	48,216,504	Cys	GCA			GGGGGTATAGCTCAGGGGTAGAGCATTTGA CTGCAGATCAAGAGGTCCCAGGTTCAAATC CAGGTGCCCCCTATA
tC(GCA)F28	Cys	6	48,218,873	48,218,949	Cys	GCA			GGGGTTATAGCTCAGGTGTAGAGCATTTGAC TGCAGATCAAGAGGTTCCATGGTTCAAATCC AGGGTGTGCCCCCTA
tC(GCA)F29	Cys	6	48,219,885	48,219,959	Cys	GCA			GGGGGTATAGCTCAGGGGTAGAGCATTTGA CTGCAGATTAAGAGGTCCCAGGTTCAAATCC AGGTGCCCCCTTTA
tC(GCA)F30	Cys	6	48,277,546	48,277,620	Cys	GCA			GGGGGTATAGCTCAGGGGTAGAGCATTTGA CTGCAGATCAAGAGGTCCCAGGTTCAAATC CAGGTGCCCCCTTAA
tC(GCA)F31	Cys	6	48,278,462	48,278,533	Cys	GCA			AGGGATAGAGCTCAGGGGTAGAACACTTGA CTGCAGATCAAGAGGTCCCTGGTCCAAATC CTGGTACCCCTA
tC(GCA)I1	Cys	9	104,322,027	104,322,101	Cys	GCA			GGGGGTATAGCTCAGGGGTAGAGCATTTGA CTGCAGATCAAGAGGTCCCAGGTTCAAATC CAGGTGCCCCCTGTA
tC(GCA)I2	Cys	9	104,325,963	104,326,037	Cys	GCA			GGGGGTATAGCTCAGTGGTAGAGCATTTGAC TGCAGATCAAGAGGTCCCAGGTTCAAATCC AGGTGCCCCCTGGA
tC(GCA)K1	Cys	11	97,440,948	97,441,024	Cys	GCA			AGGGGGTATAGCTCAGGGGTAGAGCATTTG ACTGCAGATCAAGAGGTCCATGGTTCAAATC CGGGTGCCCCCTCCAA
tC(GCA)K2	Cys	11	97,468,194	97,468,270	Cys	GCA			AGGGGGTATAGCTCAGTGGTAGAGCATTTGA CTGCAGATCAAGAGGTCCACGGTTCAAATC CGGGTGCCCCCTTGGA
tC(GCA)K3	Cys	11	97,470,026	97,470,100	Cys	GCA			GGGGGTATAGCTCAGTGGTAGAGCATTTGAC TGCAGATCAAGAGGTCCCAGGTTCAAATCC GGGTGCCCCCTCAA
tC(GCA)K4	Cys	11	97,659,365	97,659,441	Cys	GCA			AGGGGGTATAGCTCAGTGGTAGAGCATTTGA CTGCAGATCAAGAGGTCCACGGTTCAAATC CGGGTGCCCCCTAGCA
tC(GCA)K5	Cys	11	97,660,043	97,660,117	Cys	GCA			GGGGGTATAGCTCAGTGGTAGAGCATTTGAC TGCAGATCAAGAGGTCCCAGGTTCAAATCC GGGTGCCCCCTCAA
tC(GCA)Q	Cys	17	70,031,560	70,031,634	Cys	GCA			GGGGGTATAGCTCAGTGGTAGAGCATTTGAC TGCAGATCAAGAGGTCCCAGGTTCAAATCC AGGTGCCCCCTTCA
tQ(CTG)B	Gln	2	112,100,587	112,100,663	Gln	CTG			GGGTTCCATGGTGTAATGGTTAGCACTCTGG ACTCTGAATCCAGCCATAAAAGTTCAAATCT CAGTGGAACCTTAAA
tQ(CTG)C1	Gln	3	96,346,978	96,347,054	Gln	CTG			AGGTTCCATGGTGTAATGGTTAGCACTCTGG ACTCTGAATCCAGCGATCAGAGTTCAAATCT CGGTGGGACCTCTCA
tQ(TTG)C	Gln	3	96,393,467	96,393,543	Gln	TTG			AGTTTCCATGGTGTAATGGTTGGCACTCTGG ACTTTGAATCCAGCAATCAAAGTTCAAGTCT CTGTGGGACCTCTCA
tQ(CTG)C2	Gln	3	96,400,238	96,400,314	Gln	CTG			GGGTTCCATGGTGTAATGGTTAGCACTCTGG ACTCTGAATCCAGCGATCAGAGTTCAAATCT CGGTGGGACCTTTGA
tQ(CTG)C3	Gln	3	96,483,170	96,483,246	Gln	CTG			GGGTTCCATGGTGTAATGGTTAGCACTCTGG ACTCTGAATCCAGCGATCAGAGTTCAAATCT CGGTGGGACCTTTTA

tRNA gene name	tRNA family	Chromo-some	Sequence start (nt)	Sequence end (nt)	Amino acid type	Anticodon	Contains an intron	Rogue tRNA?	tRNA gene sequence (including introns)
tQ(CTG)C4	Gln	3	97,681,246	97,681,322	Gln	CTG			GGGTTCCATGGTGTAATGGTGAGCACTCTGG ACTCTGAATCCAGCGATCAGAGTTCAAATCT CGGTGGGACCTTTCA
tQ(CTG)F	Gln	6	90,101,905	90,101,980	Gln	CTG			GGGTTCCTTGGTGTAAGATGAGCACTCTGGA TTCTGAATCCAGCGATCAAAGTTCAAATCTC GGTGGGACCTCCAA
tQ(CTG)I	Gln	9	64,977,091	64,977,165	Gln	CTG			GGTTCCATGGTGTAATGGTTAGCACTCTGGA CTCTGAATCCAGCGATCCAAGTTCAAATCTC GGTGGAACCTGCA
tQ(CTG)K	Gln	11	68,737,159	68,737,235	Gln	CTG			GGGTTCCATGGTGTAATGGTTAGCACTCTGG ACTCTGAATCCAGCGATCAGAGTTCAAATCT CGGTGGAACCTTAGA
tQ(TTG)K1	Gln	11	85,863,832	85,863,908	Gln	TTG			AGGACCCATGGTGTAATGGTTAGCACTCTGG ACTTTGAATCCAGCAATCAAAGTTCAAATCT CGGTGGGACCTCTTA
tQ(TTG)K2	Gln	11	95,530,291	95,530,367	Gln	TTG			AGGTCCCATGGTGTAATGGTTAGCACTCTGG ACTTTGAATCCAGCGATCAGAGTTCAAATCT CGGTGGGACCTCACA
tQ(TTG)M1	Gln	13	20,648,513	20,648,589	Gln	TTG			GGGTCCCATGGTGTAATGGTTAGCACTCTGG ACTTTGAATCCAGCAATCAGAGTTCAAATCT CGGTGGGACCTTTAA
tQ(CTG)M1	Gln	13	21,314,081	21,314,157	Gln	CTG			GGGTTCCATGGTGTAATGGTTAGCACTCTGG ACTCTGAATCCAGCGATCAGAGTTCAAATCT CGGTGGAACCTAGTA
tQ(CTG)M2	Gln	13	21,376,702	21,376,776	Gln	CTG			GGTTCCATGGTGTAATGGCTAGCACTCTGGA CTCTGAATCCAGCGATCCAAGTTCAAATCTC GGTGGGATTTATA
tQ(TTG)M2	Gln	13	22,899,173	22,899,249	Gln	TTG			AGGCCCCATGGTGTAATGGTTAGCACTCTGG ACTTTGAATCCAGCGATCAGAGTTCAAATCT CGGTGGGACCTCATA
tQ(TTG)M3	Gln	13	22,899,737	22,899,811	Gln	TTG			GGCCCCATGGTGTAATGGTTAGCACTCTGGA CTTTGAATCCAGCGATCCAAGTTCAAATCTC GGTGGGACCTTCA
tQ(TTG)M4	Gln	13	43,115,851	43,115,921	Gln	TTG			GACCCTGCAGTGTAATGGTTAGCACTCTGGA CTTTGGATTCAGTGGTCTGAGTTCAAACCTC AGTGGATCC
tQ(TTG)X1	Gln	X	12,190,883	12,190,959	Gln	TTG			TGGTCTCATGGTGTAATGGTTAGCACACTGG ACTTTGAGTCCAGCAATCAGAGTTCGAGTCT TGGTGAGACCACTCA
tE(CTC)A1	Glu1	1	156,493,583	156,493,657	Glu	CTC			TCCCTGGTGGTTTAGTGGTTAGGATTTGGCG CTCTCAACACCGAAGCCCAGGTTCAATTCCC AGTCAGGGAAGCA
tE(CTC)A2	Glu1	1	171,142,159	171,142,233	Glu	CTC			TCCCTGGTGGTCTAGTGGTTAGGATTCGGCG CTCTCACCGCCGCGGCCCAGGTTCGATTCCC GGTCAGGGAAGCA
tE(CTC)A3	Glu1	1	171,174,298	171,174,372	Glu	CTC			TCCCTGGTGGTCTAGTGGTTAGGATTCGGCG CTCTCACCGCCGCGGCCCAGGTTCGATTCCC GGTCAGGGAAGTA
tK(CTT)C4	Glu1	3	19,853,717	19,853,787	Lys	CTT		Rogue	TCCCTGGTGGTCTAGTGGTTAGGATTCAGTG CTCTTACCACCATGGCCTGGGGTTCGATTCC GTCAGGGAA
tE(TTC)C	Glu1	3	96,388,584	96,388,658	Glu	TTC			TCCCTGGTGGTCTAGTGGCTAGGATTCGGCG CTTTCACCGCCGCGGCCCAGGTTCGATTCCC GGTCAGGGAAGGA
tE(CTC)C1	Glu1	3	96,505,223	96,505,299	Glu	CTC			TTCCCTGGTGGTCTAGTGGTTAGGATTCGGC GCTCTCACCGCCGCGGCCAGGGTTCGATTCC CGGTCAGGGAAATAA
tE(CTC)C2	Glu1	3	97,158,903	97,158,979	Glu	CTC			TTCCCTGGTGGTCTAGTGGTTAGGATTCGGC GCTCTCACCGCCGCGGCCAGGGTTCGATTCC CGGTCAGGGAAATAA
tE(CTC)C3	Glu1	3	124,296,998	124,297,072	Glu	CTC			TCCCTGATGGTATAGTGGTTAGGACTCGGTG GTCTCACCAGCGCTGCCCAGGTTCAATTCCT GGTTAGGGAACCA
tA(CGC)D	Glu1	4	130,374,246	130,374,320	Ala	CGC		Rogue	TCCCTGGTAGTCTAGTGGTTAGGATTCGGTG CTCGCACCGCCGTGGCCCAGGTTTGAATCCT AGTCAGGGAAGTA
tE(CTC)G	Glu1	7	86,818,335	86,818,409	Glu	CTC			TCCCTGGTGGTCTAGTGGTTAGGCTTTGGTG CTCTCACCTCCATGGCCCAGGTTTGATTCCT GGTCAGGGAAGCA
tE(CTC)H	Glu1	8	35,497,246	35,497,317	Glu	CTC			TCCCTGGCGGCCTAGTGGTTAGGATTCAGTG CTCTCACAGCTGCAGCCCAGGTTTGATTCCT GGTCAGGGAC
tE(CTC)J	Glu1	10	30,599,981	30,600,055	Glu	CTC			TCCCTGGTGGTCTAGTGGTTAGGATTCGGCG CTCTCACCGCCGCGGCCCAGGTTCGATTCCC GGTCAGGGAAGCA
tE(CTC)K	Glu1	11	57,914,700	57,914,776	Glu	CTC			TTCCCTGGTGGTCTAGTGGTTAGGATTCGGC GCTCTCACCGCCGCGGCCAGGGTTCGATTCC CGGTCAGGGAAGTAA
tE(CTC)L	Glu1	12	37,488,909	37,488,985	Glu	CTC			TTCCCTGGTGGTCTAGTGGTTAGGAGTCATT GCTCTCACCACCGCGTCCAGGGTTCGATTCC CGGTCAGGGAAATAA

tRNA gene name	tRNA family	Chromo-some	Sequence start (nt)	Sequence end (nt)	Amino acid type	Anticodon	Contains an intron	Rogue tRNA?	tRNA gene sequence (including introns)
tE(CTC)M	Glu1	13	20,540,343	20,540,419	Glu	CTC			TTCCCTGGTGGTCTAGTGGTTAGGATTCGGC GCTCTCACCGCCGCGGCCAGGGTTCGATTCC CGGTCAGGGAAGTGA
tE(CTC)Q	Glu1	17	54,371,374	54,371,448	Glu	CTC			TCCCTGGTGGTCTAGTGGTTAGGATTTGGCG CTCTCACCGCCGCGGCCTAGGTTCGATTCCC GGTCAGGGAAGCA
tE(TTC)A1	Glu2	1	34,718,273	34,718,347	Glu	TTC			TCCCATATGGTCTAGCGGTTAGGATTCCTGG TTTTCACCCAGGCGGCCCAGGTTCGACTCCC GGTATGGGAACAA
tE(TTC)G	Glu2	7	46,001,323	46,001,399	Glu	TTC			TTCCCACATGGTCTAGCGGTTAGGATTCCTG GTTTTCACCCAGGCGGCCAGGGTTCGACTC CCGGTGTGGGAACAGA
tE(TTC)I	Glu2	9	104,330,525	104,330,601	Glu	TTC			TTCCCACATGGTCTAGCGGTTAGGATTCCTG GTTTTCACCCAGGCGGCCAGGGTTCGACTC CCGGTGTGGGAACACA
tE(TTC)M	Glu2	13	22,843,658	22,843,734	Glu	TTC			TTCCCACATGGTCTAGCGGTTAGGATTCCTG GTTTTCACCCAGGCGGCCAGGGTTCGACTC CCGGTGTGGGAAACTA
tE(TTC)N1	Glu2	14	67,916,770	67,916,844	Glu	TTC			TCCCACATGGTCTAGCGGTTAGGATTCCTGG TTTTCACCCAGGCGGCCCAGGTTCGACTCCC GGTGTGGGAACGA
tE(TTC)N2	Glu2	14	71,254,905	71,254,981	Glu	TTC			TTCCCATATGGTCTAGCGGTTAGGATTCCTG GTTTTCACCCAGGCGGCCAGGGTTCGACTC CCGGTATGGGAACAGA
tG(GCC)A1	Gly1	1	75,276,975	75,277,048	Gly	GCC			GCATTGGTGGTTCAGTGGTAGAATTCTCGCC TGCCACGCGGGAGGCCCGAGTTTGATTCCC GGCCAATGCATTA
tG(GCC)A2	Gly1	1	171,122,532	171,122,605	Gly	GCC			GCATTGGTGGTTCAGTGGTAGAATTCTCGCC TGCCACGCGGGAGGCCCGAGTTCGATTCCC GGCCAATGCACGA
tG(GCC)A3	Gly1	1	171,144,638	171,144,711	Gly	GCC			GCATGGGTGGTTCAGTGGTAGAATTCTCGCC TGCCACGCGGGAGGCCCGAGTTCGATTCCC GGCCCATGCAGAA
tG(GCC)B	Gly1	2	57,135,758	57,135,833	Gly	GCC			TGCATTGGTGGTTCAGTGGTAGAATTCTCGC CTGCCACGCGGGAGGCCCAGGTTCGATTCC CGGCCAATGCACTTA
tG(GCC)C	Gly1	3	84,628,966	84,629,039	Gly	GCC			GCATTGGTGGTTCAGTGGTAGAATTCTCGCC TGCCACGCGGGAGGCCCGAGTTCGATTCCC GGCCAATGCACAA
tG(CCC)C1	Gly1	3	96,316,992	96,317,065	Gly	CCC			GCATTGGTGGTTCAGTGGTAGAATTCTCGCC TCCCACATGGGGGACTTGAGCTCAATTCCCA GCCAATGCAAGA
tG(CCC)C2	Gly1	3	96,425,970	96,426,045	Gly	CCC			TGCATTGGTAGTTCAATGGTAGAATTCTCGC CTCCCACGCGGGTGACCCAGGTTCGATTCCC GGCCAATGCAGTAA
tG(CCC)D	Gly1	4	32,461,608	32,461,683	Gly	CCC			TGCATTGGTGGTTCAATGGTAGAATTCTCGC CTCCCACTCGGGTGACCCAGGTTCGATTCCC GGCCAATGCAATAA
tG(GCC)D2	Gly1	4	131,567,923	131,567,996	Gly	GCC			GCATTGGTGGTTCAGTGGTAGAATTCTTGCC TGCCACCCGGGAGGCCCAAGTTCAATTCCT GGCCAATGTACAA
tG(GCC)G	Gly1	7	15,381,733	15,381,808	Gly	GCC			GGCATGGGTGGTTCAGTGGTAGAATTCTCAC CTGCCATGAGGGAGGCCCAGGTTCAATTCC AGGCCCATTGCAGAA
tG(GCC)H1	Gly1	8	109,954,782	109,954,855	Gly	GCC			GCATTGGTGGTTCAGTGGTAGAATTCTCGCC TGCCACGCGGGAGGCCCGAGTTCGATTCCC GGCCAATGCAGTA
tG(GCC)H2	Gly1	8	109,955,475	109,955,548	Gly	GCC			GCATTGGTGGTTCAGTGGTAGAATTCTCGCC TGCCACGCGGGAGGCCCGAGTTCGATTCCC GGCCAATGCAGAA
tG(GCC)H3	Gly1	8	110,388,966	110,389,039	Gly	GCC			GCATTGGTGGTTCAGTGGTAGAATTCTCGCC TGCCACGCGGGAGGCCCGAGTTCGGTTCCC GGCCAATGCATAA
tG(GCC)J	Gly1	10	99,815,477	99,815,550	Gly	GCC			GTATTGGTGGTTCAGTGGTAGAATTCTCGCC TGCCACACCGGATGCCTGAGTTCCATTCCCG GCCAATGCACTA
tG(GCC)K1	Gly1	11	68,731,346	68,731,421	Gly	GCC			TGCATTGGTGGTTCAGTGGTAGAATTCTCGC CTGCCACGCGGGAGGCCCAGGTTCGATTCC CGGCCAATGCACAGA
tG(GCC)M1	Gly1	13	21,090,704	21,090,777	Gly	GCC			GCATTGGTGGTTCAGTGGTAGAATTCTCGCC TGCCACGCGGGAGGCCCGAGTTCGATTCCC GGCCAATGCAGGA
tG(GCC)M2	Gly1	13	22,897,327	22,897,400	Gly	GCC			GCATTGGTGGTTCAGTGGTAGAATTCTCGCC TGCCACGCGGGAGGCCCGAGTTCGATTCCC GGCCAATGCACTA
tG(TCC)A1	Gly2	1	171,115,558	171,115,634	Gly	TCC			TGCGTTGGTGGTATAGTGGTGAGCATAGCTG CCTTCCAAGCAGTTGACCAGGGTTCGATTCC CGGCCAACGCAAAGA
tG(TCC)A2	Gly2	1	171,142,701	171,142,775	Gly	TCC			GCGTTGGTGGTATAGTGGTGAGCATAGCTGC CTTCCAAGCAGTTGACCCAGGTTCGATTCCC GGCCAACGCAGCA

tRNA gene name	tRNA family	Chromo-some	Sequence start (nt)	Sequence end (nt)	Amino acid type	Anticodon	Contains an intron	Rogue tRNA?	tRNA gene sequence (including introns)
tG(TCC)A3	Gly2	1	171,174,010	171,174,086	Gly	TCC			TGCGTTGGTGGTATAGTGGTGAGCATAGCTG CCTTCCAAGCAGTTGACCAGGGTTCGATTCC CGGCCAACGCATAAA
tG(TCC)C	Gly2	3	96,504,785	96,504,861	Gly	TCC			TGCGTTGGTGGTATAGTGGTGAGCATAGCTG CCTTCCAAGCAGTTGACCAGGGTTCGATTCC CGGCCAACGCATAAA
tG(TCC)K	Gly2	11	68,653,830	68,653,906	Gly	TCC			TGCGTTGGTGGTATAGTGGTGAGCATAGCTG CCTTCCAAGCAGTTGACCAGGGTTCGATTCC CGGCCAACGCAACGA
tG(TCC)X1	Gly2	X	33,574,164	33,574,240	Gly	TCC			TGCGTTGGTGGTATAGTGGTGAGCATAGCTG CCTTCCAAGCAGTTGACCAAGGTTTGATACC CAGCCAACGCATGAA
tG(CCC)F	Gly3	6	86,775,971	86,776,044	Gly	CCC			GCGCCGCTGGTGTAGTGGTATCATGCAAGAT TCCCATTCTTGCGACCCGAGTTCGATTCCCG GGCGGCGCATGA
tG(CCC)Q	Gly3	17	24,474,352	24,474,425	Gly	CCC			GCGCCGCTGGTGTAGTGGTATCATGCAAGAT TCCCATTCTTGCGACCCGAGTTCGATTCCCG GGCGGCGCATGA
tH(GTG)B1	His	2	122,124,514	122,124,588	His	GTG			GCCGTGATCGTATAGTGGTTAGTACTCTGCG TTGTGGCCGCAGCAACCTAGGTTCGAATCC GAGTCACGGCATTA
tH(GTG)B2	His	2	122,126,383	122,126,457	His	GTG			GCCGTGATCGTATAGTGGTTAGTACTCTGCG TTGTGGCCGCAGCAACCTAGGTTCGAATCC GAGTCACGGCATTA
tH(GTG)B3	His	2	122,126,991	122,127,065	His	GTG			GCCGTGATCGTATAGTGGTTAGTACTCTGCG TTGTGGCCGCAGCAACCTAGGTTCGAATCC GAGTCACGGCATTA
tH(GTG)C1	His	3	96,411,252	96,411,326	His	GTG			GCCGTGATCGTATAGTGGTTAGTACTCTGCG TTGTGGCCGCAGCAACCTAGGTTCGAATCC GAGTCACGGCAGTA
tH(GTG)C2	His	3	96,437,878	96,437,952	His	GTG			GCCGTGATCGTATAGGGGTTAGTACTCTGCG TTGTGGCCGCAGCAACCTAGGTTCGAATCC GAGTCACGGCATTA
tH(GTG)C3	His	3	96,454,247	96,454,321	His	GTG			GCCGTGATCGTATAGTGGTTAGTACTCTGCG TTGTGGCCGCAGCAACCTAGGTTCGAATCC GAGTCACGGCATGA
tH(GTG)C4	His	3	96,467,748	96,467,824	His	GTG			TGCCGAGATCGTATAGTGGTTAGTACTCTGC ATTGTGGCTGCAGCAACCACGGTTCGAATCC GAGTCTCGGCACTAA
tH(GTG)C5	His	3	96,504,143	96,504,217	His	GTG			GCCGTGATCGTATAGTGGTTAGTACTCTGCG TTGTGGCCGCAGCAACCTAGGTTCGAATCC GAGTCACGGCATTA
tH(GTG)D	His	4	80,894,465	80,894,539	His	GTG			GCCGTGATCGTATAGTGGTTAGTACTCTGCG TTGTGGCCGCAGCAACCTAGGTTCGAATCC GAGTCACGGCAGGA
tI(AAT)D1	Ile	4	28,050,591	28,050,669	Ile	AAT			TGGCCGGTTAGCTCAGTTGGTTAGAGCATGG TGCTAATAACGCCAAGGTAACGGGTTCGATC CCCGTATGGGCCAGGGA
tI(AAT)H	Ile	8	46,264,143	46,264,215	Ile	AAT			GGCTAGTTAGCTCAGTTGGTTAGAGCATGGT GCTAATAATGCCAAGGTCGCAGTTTCAAACC TGTATGGGCTA
tI(AAT)K1	Ile	11	68,649,558	68,649,636	Ile	AAT			TGGCCGGTTAGCTCAGTTGGTTAGAGCGTG GTGCTAATAACGCCAAGGTAGCGGGTTCGAT CCCCGTACGGGCCAAAGA
tI(AAT)K2	Ile	11	68,674,872	68,674,950	Ile	AAT			TGGCCGGTTAGCTCAGTTGGTTAGAGCGTG GTGCTAATAACGCCAAGGTAGCGGGTTCGAT CCCCGTACGGGCCAGTGA
tI(AAT)L	Ile	12	105,591,749	105,591,827	Ile	AAT			TGGCCGGTTAGCTCAGTTGGTTAGAGCGTG GTGCTAATAACGCCAAGGTAGCGGGTTCGAT CCCCGTACGGGCCAGGGA
tI(AAT)M1	Ile	13	21,235,440	21,235,518	Ile	AAT			TGGCCGGTTAGCTCAGTTGGTTAGAGCGTG GTGCTAATAACGCCAAGGTAGCGGGTTCGAT CCCCGTACGGGCCAGTAA
tI(AAT)M2	Ile	13	21,262,153	21,262,231	Ile	AAT			TGGCCGGTTAGCTCAGTTGGTTAGAGCGTG GTGCTAATAACGCCAAGGTAGCGGGTTCGAT CCCCGTACGGGCCAGTGA
tI(AAT)M3	Ile	13	21,409,346	21,409,424	Ile	AAT			TGGCCGGTTAGCTCAGTTGGTTAGAGCGTG GTGCTAATAACGCCAAGGTAGCGGGTTCGAT CCCCGTACGGGCCAGTAA
tI(AAT)M4	Ile	13	22,689,172	22,689,250	Ile	AAT			TGGCCGGTTAGCTCAGTTGGTTAGAGCGTG GTGCTAATAACGCCAAGGTAGCGGGTTCGAT CCCCGTACGGGCCACACA
tI(AAT)M5	Ile	13	22,777,183	22,777,261	Ile	AAT			TGGCCGGTTAGCTCAGTTGGTTAGAGTGTGG TGCTAATAACGCCAAGGTAGCGGGTTCGATC CCCGTACGGGCCACATA
tI(AAT)M6	Ile	13	22,818,229	22,818,307	Ile	AAT			TGGCCGGTTAGCTCAGTTGGTTAGAGCGTG GTGCTAATAACGCCAAGGTAGCGGGTTCGAT CCCCGTACGGGCCAGCTA
tI(AAT)P	Ile	16	75,787,284	75,787,362	Ile	AAT			TGGCCGGTTAGCTCAGTTGGTTAGAGCATGG TGCTAATAACGCCAAGGTAGTGGGTTCGATG CCCGTATGGGCCAGCAA

tRNA gene name	tRNA family	Chromo-some	Sequence start (nt)	Sequence end (nt)	Amino acid type	Anticodon	Contains an intron	Rogue tRNA?	tRNA gene sequence (including introns)
tI(TAT)G	Ile2	7	17,337,157	17,337,254	Ile	TAT	Intron		GCTCCAGTGGCGCAATCGGTTAGCGCGCGG TACTTATATGTCAGTGCTAAGCGTAAGCGAT GCCGAGGTTGTGAGTTCGATCCTCACCTGG AGCACTA
tI(TAT)M2	Ile2	13	21,277,058	21,277,156	Ile	TAT	Intron		TGCTCCAGTGGCGCAATCGGTTAGCGCGCG GTACTTATACAGCAGTATAAGTGCGGGTGAT GCCGAGGTTGTGAGTTCGAGCCTCACCTGG AGCATGTA
tI(TAT)M	Ile2	13	22,650,541	22,650,641	Ile	TAT	Intron		TGCTCCAGTGGCGCAATCGGTTAGCGCGCG GTACTTATACAACAGTGTGAGCGCGAGAGC GATGCCGAGGTTGTGAGTTCGAGCCTCACC TGGAGCATTAA
tI(TAT)Q4	Ile2	17	82,434,678	82,434,773	Ile	TAT	Intron		GCTCCAGTGGCGCAATCGGTTAGCGCGCGG TACTTATACAGCAGTACATACAGAGCAATGC CGAGGTTGTGAGTTCGAGCCTCACCTGGAG CACGA
tL(CAG)A1	Leu1	1	171,116,178	171,116,265	Leu	CAG			TGTCAGGATGGCCGAGCGGTCTAAGGCGCT GCGTTCAGGTCGCAGTCTCACCTGGAGGCG TGGGTTCGAATCCCACTTCTGACAAATA
tL(CAG)A2	Leu1	1	171,143,994	171,144,079	Leu	CAG			GTCAGGATGGCCGAGCGGTCTAAGGCGCTG CGTTCAGGTCGCAGTCTCCACTGGAGGCGT GGGTTCGAATCCCACTCCTGACAGCA
tL(CAG)A3	Leu1	1	171,173,090	171,173,175	Leu	CAG			GTCAGGATGGCCGAGCGGTCTAAGGCGCTG CGTTCAGGTCGCAGTCTCCACTGGAGGCGT GGGTTCGAATCCCACTCCTGACAAAA
tL(CAG)C2	Leu1	3	23,926,279	23,926,364	Leu	CAG			GTCAGGATGGCCGAGCAGTCTAAGGCACTG CGTTCAGGTCGCAGTCTCCACTGGAGGCGT GGATTCGAATCCCACTCCTGACAACA
tL(CAG)H1	Leu1	8	37,257,589	37,257,674	Leu	CAG			GTCAGGATGGCCGAGTGGTCTAAGGAGCTG TGTTCAGGTCGCAGTCTCCACTGGAGGCGT GGGTTCGAATCCCACTCCTGACAGCA
tL(CAG)H2	Leu1	8	94,051,476	94,051,563	Leu	CAG			TGTCAGGATGGCCGAGCGGTCTAAGGCGCT GCGTTCAGGTCGCAGTCTCACCTGGAGGCG TGGGTTCGAATCCCACTTCTGACAGAAA
tL(CAG)H3	Leu1	8	94,051,846	94,051,933	Leu	CAG			TGTCAGGATGGCCGAGCGGTCTAAGGCGCT GCGTTCAGGTCGCAGTCTCACCTGGAGGCG TGGGTTCGAATCCCACTTCTGACAAGTA
tL(CAA)K1	Leu1	11	57,914,997	57,915,108	Leu	CAA	Intron		TGTCAGGATGGCCGAGTGGTCTAAGGCGCC AGACTCAAGGTGACAAGCCATACCTACGGG TGTTCTGGTCTCCGAATGGAGGCGTGGGTTC GAATCCCAATTCTGACACAAA
tL(CAA)M1	Leu1	13	20,552,608	20,552,718	Leu	CAA	Intron		TGTCAGGATGGCCGAGTGGTCTAAGGCGCC AGACTCAAGCTATGGCTTCATCGCTCTGAGG GTTCTGGTCTCCCCTGGAGGCGTGGGTTCG AATCCCACATCTGACAGCTA
tL(CAA)M	Leu1	13	20,580,242	20,580,351	Leu	CAA	Intron		TGTCAGGATGGCCGAGTGGTCTAAGGCGCC AGACTCAAGCTTAGCTTCCATGTCTGGGGAT TCTGGTCTCCGTATGGAGGCGTGGGTTCGAA TCCCACTACTGACACAGA
tL(CAA)M2	Leu1	13	21,294,961	21,295,071	Leu	CAA	Intron		TGTCAGGATGGCCGAGTGGTCTAAGGCGCC AGACTCAAGCGTTCGCTTCATCTACTGAGGG TTCTGGTCTCCGTGTGGAGGCGTGGGTTCG AATCCCACATCTGACACAGA
tL(CAG)M3	Leu1	13	22,833,553	22,833,638	Leu	CAG			GTCAGGATGGCCGAGCGGTCTAAGGCGCTG CGTTCAGGTCGCAGTCTCCACTGGAGGCGT GGGTTCGAATCCCACTCCTGACAACA
tL(TAG)B	Leu2	2	37,881,384	37,881,468	Leu	TAG			GGTAGCATGGCCAAGTGGTCTAAAGCACTG AATTTAGGCTCCAGTCATTACGATAGCATGG GTTCGAGTCCCACCACTGCCATAA
tL(TAG)G2	Leu2	7	108,130,874	108,130,958	Leu	TAG			GGTAGCGTGGCCGAGTGGTCTAAGGCGCTG GATTTAGGCTCCAGTCATTACGATGGCGTGG GTTCGAATCCCACCGCTGCCACAA
tL(AAG)G	Leu2	7	108,212,915	108,212,999	Leu	AAG			GGTAGTGTGGCCGAGCGGTCTAAGGCGCTG GATTAAGGCTCCAGTCTCTACGGGGGCGTG GGTTCGAATCCCACCGCTGCCAAGA
tL(AAG)K	Leu2	11	48,496,706	48,496,790	Leu	AAG			GGTAGCGTGGCCGAGCGGTCTAAGGCGCTG GATTAAGGCTCCAGTCTCTACGGGGGCGTG GGTTCGAATCCCACCGCTGCCAGTA
tL(TAG)K	Leu2	11	68,736,797	68,736,881	Leu	TAG			GGTAGCGTGGCCGAGCGGTCTAAGGCGCTG GATTTAGGCTCCAGTCTCTACGGAGGCGTGG GTTCGAATCCCACCGCTGCCAGGA
tL(AAG)M1	Leu2	13	20,549,598	20,549,682	Leu	AAG			GGTAGCGTGGCCGAGCGGTCTAAGGCGCTG GATTAAGGCTCCAGTCTCTACGGGGGCGTG GGTTCGAATCCCACCGCTGCCAACA
tL(AAG)M2	Leu2	13	20,549,768	20,549,852	Leu	AAG			GGTAGTGTGGCCGAGCGGTCTAAGGCGCTG GATTAAGGCTCCAGTCTCTACGGGGGCGTG GGTTCGAATCCCACCACTGCCAACA

tRNA gene name	tRNA family	Chromo-some	Sequence start (nt)	Sequence end (nt)	Amino acid type	Anticodon	Contains an intron	Rogue tRNA?	tRNA gene sequence (including introns)
tL(AAG)N	Leu2	14	43,757,328	43,757,414	Leu	AAG			TGGTAGCGTGGCCGAGCGGTCTAAGGCGCT GGATTAAGGCTCCAGTCTCATCGGGGGCGT GGGTTCGAATCCCACCGCTGCCATCTA
tL(TAG)N	Leu2	14	43,776,273	43,776,359	Leu	TAG			TGGTAGTGTGGCCGAGCGGTCTAAGGCGCT GGATTTAGGCTCCAGTCTCATCGGAGGCGTG GGTTCGAATCCCACCACTGCCAGTGA
tL(TAA)J	Leu3	10	12,633,339	12,633,424	Leu	TAA			ACCAGGATGGCCGAGTGGTTAAGGCGTTGG ACTTAAGATCCAATGGACAAATGTCTGCGTG GGTTCGAACCCCACTCCTGGTAAAA
tStop(TTA)M	Leu3	13	21,175,922	21,176,009	Stop	TTA		Rogue	TACTGGGATGGCTGAGTGGTTAAGGCCTTGG ACTTTAGATCCAATGGGCAGATGCCTGCGTG GGTTCAAACCCCACTCCCAGTATTAA
tL(TAA)M	Leu3	13	21,228,142	21,228,229	Leu	TAA			TACTGGGATGGCTGAGTGGTTAAGGCGTTG GACTTAAGATCCAATGGGCAGTTGCCTGCGT GGGTTCGAACCCCACTCCCAGTATGTA
tL(TAA)S	Leu3	19	11,261,436	11,261,523	Leu	TAA			TACCAGAATGGCCGAGTGGTTAAGGCGTTG GACTTAAGATCCAATGGATATATATCCGCGTG GGTTCGAACCCCACTTCTGGTAAATA
tL(TAA)X1	Leu3	X	151,104,987	151,105,072	Leu	TAA			GATGGGATGGCTGAGAGGTTAAGGCTTTGG ACTTAAGATCCAATGGGCAAATGCCTGCGTG GGTTTGAACCCCACTCCCAATATTA
tK(CTT)A	Lys1	1	110,913,332	110,913,409	Lys	CTT			TGCCTGGCTAGCTCAGTCCATAGAGCATGGG ACTCTTAATCCCAGGGTCATGGGTTCGAGCC CCATATTAGGCACCAA
tK(CTT)B	Lys1	2	86,035,043	86,035,120	Lys	CTT			AGCCTAGCTAGTTCAGTCGGTAGAGCATGA GACTCTTAATCTCAGGTTCATGAGTTTGAGC CCCATGTTGGTTTGGCA
tK(CTT)C1	Lys1	3	3,126,070	3,126,147	Lys	CTT			TGCCCAGCTAGCTCAGTCTGTAGAGCATGAG ACTCTTAGTCTCAGGGTCATGGGTTGGAGCC CCATGTTGTGCAAGAA
tK(CTT)C2	Lys1	3	3,151,823	3,151,900	Lys	CTT			TGCCCAGCTAGCTCAGTCTGTAGAGCATGAG ACTCTTAATCTCAGGGTCATGAGTTCGAGCC CCACGTTGGGTGAGAA
tK(CTT)C3	Lys1	3	3,201,314	3,201,380	Lys	CTT			GCCAGCTCAGTTGGTAGAGCATGAGACTCTT AATCTCAGGGTCGTGGGTTCGAGCCCCATAT TGGGG
tK(CTT)C5	Lys1	3	96,430,140	96,430,217	Lys	CTT			TGCCCGGCTAGCTCAGTCGGTAGAGCATGA GACTCTTAATCTCAGGGTCATGGGTTCGAGC CCCACGTTGGGCGCATA
tK(CTT)C6	Lys1	3	96,503,289	96,503,364	Lys	CTT			GCCCGGCTAGCTCAGTCGGTAGAGCATGAG ACTCTTAATCTCAGGGTCGAGGGTTCGAGCC CCACGTTGGGCGCTA
tSeC(TCA)E	Lys1	5	5,634,400	5,634,470	SeC	TCA		Rogue	ACCCAGCTAGCTCAGTGGGTAGAGCATGAG ACTCAATCTCAGGGTTGTGGGTTCGAGCCCC ACATTGGGCG
tK(CTT)E	Lys1	5	89,343,013	89,343,090	Lys	CTT			CACCCGACTGGGTTAGTCAATAGAACATGA GACTCTTAATCTCAGGGTTATGGGTTTGAGC CCCACCTTGGGTGGGTA
tL(TAA)E2	Lys1	5	90,894,680	90,894,754	Leu	TAA		Rogue	TCTGGCTGGCTCAGTTGGTAGAGCTTGAGA CTTAAATCTCAGGGTTGCAGGTTTGAGTTCT GTTGGGGGTCAGAT
tK(CTT)G	Lys1	7	114,653,411	114,653,486	Lys	CTT			GCCCAGCTACTTCAGTTGGTGGAGCAAGAG TCTCTTAATCTCAGGGTCAAGGGTTCAAGCC CCATGTTGGGTACCA
tK(TTT)H	Lys1	8	79,077,031	79,077,103	Lys	TTT			GCCTGGTTAGCTCAGTCAGTAGAGTATGAAA CTTTTAATCTCAGGGTTGTAGGTTTGAGCCT CACATTGGGCA
tK(CTT)I1	Lys1	9	14,794,790	14,794,865	Lys	CTT			GTCCTGCTGGCTCAGTCGGTACAGCATGGG ACTCTTAATCCCAGGGTCGAGGGTTCGAGCT CCACGTTGGGTACCA
tK(CTT)I2	Lys1	9	89,954,947	89,955,024	Lys	CTT			TGCCCGGCTAGCTCAGTCGGTAGAGCATGG GACTCTTAATCCCAGGGTCATGGGTTCGAGC CCCACGTTGGGCGGTTA
tK(CTT)K1	Lys1	11	22,386,936	22,387,011	Lys	CTT			ACCCAGCTAGTTCAGTCGGTAGAGCATGAG ACTCTTAATCTTAGAGTCGAGGGCTCAGGTC CCATGTTGGGTTCCA
tK(CTT)K2	Lys1	11	29,296,855	29,296,927	Lys	CTT			GCCTGGCTAGCTCAGTTGGTAGAGCATGGA ACTCTTAATCCCTGGGTTGTAAGTTTGAGCC TCATGTAGGGCA
tK(CTT)K3	Lys1	11	48,473,614	48,473,691	Lys	CTT			CGCCCGGCTAGCTCAGTCGGTAGAGCATGA GACTCTTAATCTCAGGGTCATGGGTTCGAGC CCCACGTTGGGCGCATA
tK(CTT)K4	Lys1	11	52,291,514	52,291,588	Lys	CTT			GCCCAAAACTAGCTCAGTCAGTAGAGAATG ATACTCTTAATCTCAGGGTCATGGGTTCAAA CCCCACGTTGGGTG
tK(CTT)K5	Lys1	11	73,749,866	73,749,943	Lys	CTT			AACCTGGCTAGGTCAGTTGGTAGATCATGAG ACTCTTAATCTCAGGGTCATGGGTTCAAGCC CCATGTTGGTTTGGCA

tRNA gene name	tRNA family	Chromo-some	Sequence start (nt)	Sequence end (nt)	Amino acid type	Anticodon	Contains an intron	Rogue tRNA?	tRNA gene sequence (including introns)
tK(CTT)L	Lys1	12	66,236,117	66,236,194	Lys	CTT			TGCCCGGCTAGCTCAGTCGGTAGAGCATGG GACTCTTAATCCCAGGGTCATGGGTTCGAGC CCCACGTTGGGCGGTGA
tK(CTT)M1	Lys1	13	3,994,182	3,994,254	Lys	CTT			ACACTGCTAGCTCAGTTGGTAGAGCATGAG ACTCTTAATCTCAGGGTCGTGGGTTCGAGCC TCACGTTGGGCG
tK(CTT)M2	Lys1	13	11,726,878	11,726,951	Lys	CTT			TGTTGTAAAGCTCAGTTGTTAGAGCATGAGA CTCTTATTCTTAGGGTTGTGAGTTCGAGCCC CACATTGGATAT
tK(CTT)M3	Lys1	13	22,816,463	22,816,540	Lys	CTT			TGCCCGGCTAGCTCAGTCGGTAGAGCATGA GACTCTTAATCTCAGGGTCATGGGTTCGAGC CCCACGTTGGGCGCAAA
tK(CTT)M4	Lys1	13	69,310,465	69,310,537	Lys	CTT			GACTCCAGCCCAGTCAGTAGAGCATGAGAC TCTTAATCTCAGGGTCGTGGGTTCGAGCCCC ACGTTGGGTGCA
tK(CTT)N	Lys1	14	53,538,938	53,539,015	Lys	CTT			TGCCCAGCTCACTCAGTCGGTAGAGCATGA GACTCTTAATCTCAGGGTCATAGGTTCGAGC CCCATGGTGGGCGAGAA
tK(CTT)P1	Lys1	16	3,026,421	3,026,498	Lys	CTT			TGCCCAGCTAGCTCAGTCTGTAGAGCATGAG ACTCTTAATCTCAGGGTTATAGGTTCGAGCC CCGCTATGGGTGAGAA
tK(TTT)P	Lys1	16	3,120,615	3,120,692	Lys	TTT			TGCCCAGCCAGCTCAGTAGGTAGAGTATGA GACTTTTAATCTCAGGGTGATGGGTTCGAGC CCCATGTTGGGGGAGAA
tK(CTT)P2	Lys1	16	16,630,954	16,631,029	Lys	CTT			GCCCAGCTAGCTTAGTTGGTAGAGCATGAG ACTCTTAATCTCAGAGTCAAGGGTTCAGGCC TCATGTTTGGCACCA
tK(CTT)P3	Lys1	16	47,201,463	47,201,540	Lys	CTT			TCCCCGGCTAGCTCAGTCAGTAGAGCTTGA GAATCTTAATCTCAGGGTCATGGGTTGGAGC CCCACGTTGGGCGAGAA
tK(CTT)Q1	Lys1	17	7,286,782	7,286,859	Lys	CTT			TGCCTGGCTAGCTTAGTTGGTAAAGCATAAG ACTCTTAATCCTAGGGTCATAAGTTTGAGCC CCACATTAGGCGCCAA
tK(CTT)Q2	Lys1	17	22,129,888	22,129,965	Lys	CTT			CGCCCGGCTAGCTCAGTCGGTAGAGCATGA GACTCTTAATCTCAGGGTCATGGGTTCGAGC CCCACGTTGGGCGAGAA
tK(CTT)Q3	Lys1	17	22,131,255	22,131,332	Lys	CTT			CGCCCGGCTAGCTCAGTCGGTAGAGCATGA GACTCTTAATCTCAGGGTCATGGGTTCGAGC CCCACGTTGGGCGAGAA
tK(CTT)Q4	Lys1	17	22,137,022	22,137,097	Lys	CTT			GCTCGGTTAGCTCAGTCGGTAGAGCATGAG ACTCTTATTCTCAGGGTTGAGGGTTTGAGTC TCCCGTTAGGCGTCA
tK(CTT)Q5	Lys1	17	22,143,288	22,143,365	Lys	CTT			CGCCCGGCTAGCTCAGTCGGTAGAGCATGA GACTCTTAATCTCAGGGTCATGGGTTCGAGC CCCACGTTGGGCGAGTA
tK(CTT)R	Lys1	18	3,136,757	3,136,834	Lys	CTT			TGCCCAGCTAGCTCAGTCGGTAGAGCATGA GACTCTTAATCTCAGGGTCATAGGTTTGAGC CCCGATATGGGTGAGAA
tK(CTT)X1	Lys1	X	89,076,047	89,076,120	Lys	CTT			GGTACTAGCTCAGTTGGTAGAGCTTGAGACT CTTAATCTCAGGAACATGGGTTCGAGCCCCA CATTGGGTGCCA
tK(CTT)X2	Lys1	X	126,202,890	126,202,967	Lys	CTT			TGCCCAGCTAGCTCAGTTGGTAGAGCGTGG GACTCTTAATCCTAGGGTCATGGGTTCGAAC CCCACGTTGGGCGGTTA
tK(CTT)Y	Lys1	Y	47,724,709	47,724,786	Lys	CTT			TGCCCAGCTAGCTCAGTCAGTAGAGCAGGA GACTCTTAGTCTCGGGGTTATAGGTTCGAGC CCCACAGTGGGCAAGGA
tK(TTT)A1	Lys2	1	132,915,912	132,915,987	Lys	TTT			GCCCGGATAGCTCAGTCGGTAGAGCATCAG ACTTTTAATCTGAGGGTCCAGGGTTCAAGTC CCTGTTCGGGCGCTA
tK(TTT)A2	Lys2	1	132,916,273	132,916,350	Lys	TTT			CGCCCGGATAGCTCAGTCGGTAGAGCATCA GACTTTTAATCTGAGGGTCAAGGGTTCAAGT CCCTGTTCGGGCGGCTA
tK(TTT)E	Lys2	5	30,095,683	30,095,762	Lys	TTT			AGCCTGGATGGCTCTCGGTCAGTAGGACATC AGACTTTTAATCTGAGGGACCAAGATTCAA GTCCCTGTTCAGGTTCTTA
tK(TTT)G	Lys2	7	17,762,839	17,762,914	Lys	TTT			GCCTGGATAGCTCAGTCGGTAGAGCATCAG ACTTTTAATCTGAGGGTCCAGGGTTCAAGTC CCTGTTCAGGCGGAA
tK(TTT)J	Lys2	10	87,577,726	87,577,801	Lys	TTT			ACCTGGATAGCTCAGTTGGTAGAGCATCAGA CTTTTAATCTGAGGCTCCAGGGTTCAAGTCC CTGTTCGAGCCCTA
tK(TTT)K	Lys2	11	68,737,729	68,737,804	Lys	TTT			GCCCGGATAGCTCAGTCGGTAGAGCATCAG ACTTTTAATCTGAGGGTCCAGGGTTCAAGTC CCTGTTCGGGCGCTA
tK(TTT)M1	Lys2	13	20,547,634	20,547,709	Lys	TTT			GCCCGGATAGCTCAGTCGGTAGAGCATCAG ACTTTTAATCTGAGGGTCCAGGGTTCAAGTC CCTGTTCGGGCGCTA
tK(TTT)M2	Lys2	13	21,288,064	21,288,139	Lys	TTT			GCCTGGATAGCTCAGTCGGTAGAGCATCAG ACTTTTAATCTGAGGGTCCAGGGTTCAAGTC CCTGTTCAGGCGGAA

tRNA gene name	tRNA family	Chromo-some	Sequence start (nt)	Sequence end (nt)	Amino acid type	Anticodon	Contains an intron	Rogue tRNA?	tRNA gene sequence (including introns)
tK(TTT)M3	Lys2	13	21,351,912	21,351,987	Lys	TTT			GCCTGGATAGCTCAATTGGTAGAGCATCAGA CTTTTAATCTGAGGGTTCAGGGTTCAAGTCC CTGTTCAGGCGCTA
tK(TTT)S1	Lys2	19	11,256,393	11,256,468	Lys	TTT			GCCCGGATAGCTCAGTCGGTAGAGCATCAG ACTTTTAATCTGAGGGTCCAGGGTTCAAGTC CCTGTTCGGGCGCTA
tK(TTT)S2	Lys2	19	11,260,047	11,260,122	Lys	TTT			GCCCGGATAGCTCAGTCGGTAGAGCATCAG ACTTTTAATCTGAGGGTCCAGGGTTCAAGTC CCTGTTCGGGCGGAA
tM(CAT)C	Met1	3	90,904,791	90,904,867	Met	CAT			TAGCAGAGTGGCGCAGCGGAAGCGTGCTGG GCCCATAACCCAGAGGTCGATGGATCGAAA CCATCCTCTGCTATCGA
tM(CAT)M4	Met1	13	21,091,117	21,091,193	Met	CAT			TAGCAGAGTGGCGCAGCGGAAGCGTGCTGG GCCCATAACCCAGAGGTCGATGGATCGAAA CCATCCTCTGCTAAGGA
tM(CAT)M5	Met1	13	21,287,234	21,287,310	Met	CAT			TAGCAGAGTGGCGCAGCGGAAGCGTGCTGG GCCCATAACCCAGAGGTCGATAGATCGAAA CCATCCTCTGCTAGTTA
tM(CAT)M6	Met1	13	21,311,290	21,311,364	Met	CAT			AGCAGAGTGGCGCAGCGGAAGCGTGCTGG GCCCATAACCCAGAGGTCGAAGGATCGAAA CCATCCTCTGCTACAA
tM(CAT)M7	Met1	13	21,355,155	21,355,231	Met	CAT			TAGCAGAGTGGCGCAGCGGAAGCGTGCTGG GCCCATAACCCAGAGGTCGATGGATCGAAA CCATCCTCTGCTATGAA
tM(CAT)M8	Met1	13	22,880,204	22,880,278	Met	CAT			AGCAGAGTGGCGCAGCGGAAGCGTGCTGG GCCCATAACCCAGAGGTCGAAGGATCGAAA CCATCCTCTGCTAGAA
tM(CAT)M9	Met1	13	22,897,816	22,897,890	Met	CAT			AGCAGAGTGGCGCAGCGGAAGCGTGCTGG GCCCATAACCCAGAGGTCGAAGGATCGAAA CCATCCTCTGCTAACA
tM(CAT)M10	Met1	13	22,910,468	22,910,544	Met	CAT			TAGCAGAGTGGCGCAGCGGAAGCGTGCTGG GCCCATAACCCAGAGGTCGATGGATCGAAA CCATCCTCTGCTATCTA
tM(CAT)O2	Met1	15	69,662,937	69,663,011	Met	CAT			AGCAGAGTGGCGCAGCGGAAGCATGCTGGG CCCATAACCCAGAGGTCGAAGGATCGAAAC CATCCTCTGCTAACA
tM(CAT)X	Met1	X	117,155,003	117,155,074	Met	CAT			AGCAGAGTGGCACAATGGAAGCGTGCTGGT CCCATAACCCAGAGGTCAATGGATTGAAAC CATCCTCTGCTT
tM(CAT)E	Met2	5	107,162,848	107,162,925	Met	CAT			TGCCTCCTTAGCATAGTAGGCAGCGCATCAG TCTCATAATCTGAAGGTCATGAGTTTGAACC TCAGAGGGGTCAACCA
tM(CAT)K	Met2	11	121,635,187	121,635,264	Met	CAT			TGCCTCCTTAGTGTAGTAGGCATTGCGTCAG TCTCATAATCTGAAGGTCATGAGTTCAAGCC TCAGAGTGGGCAAACA
tM(CAT)M1	Met2	13	20,547,165	20,547,242	Met	CAT			TGCCTCCTTAGCGCAGTAGGCAGCGCGTCA GTCTCATAATCTGAAGGTCATGAGTTCGAAC CTCAGAGGGGGCAGTTA
tM(CAT)M2	Met2	13	20,549,245	20,549,322	Met	CAT			TGCCTCCTTAGCGCAGTAGGCAGCGCGTCA GTCTCATAATCTGAAGGTCATGAGTTCGAAC CTCAGAGGGGGCAACCA
tM(CAT)M3	Met2	13	20,706,761	20,706,838	Met	CAT			TGCCTTCTTAGCGCAGTAGGCAGCGCGTCA GTCTCATAATCTGAAGGTCATGAGTTCGAAC CTCAGAGAGGGCAGATA
tM(CAT)H	Met3	8	121,054,586	121,054,661	Met	CAT			GCCTCGTTAGCGCAGTAGGTAGCGCGTCAG TCTCATAATCTGAAGGTCGAGAGTTCGATCC TCACACGGGGCATCA
tM(CAT)O1	Met3	15	58,172,200	58,172,277	Met	CAT			TGCCTCGTTAGCGCAGTAGGTAGCGCGTCA GTCTCATAATCTGAAGGTCATGAGTTCGATC CTCACACGGGGCACAAA
tF(GAA)E	Phe	5	122,947,750	122,947,825	Phe	GAA			GCCGAAATAGCTCAGTTGGGAGAGCGTTAG ACTGAAGATCTAAAGGTCCATGGTTCGATCC CGGGTTTCGGCAGCA
tF(GAA)J	Phe	10	80,149,889	80,149,964	Phe	GAA			GCCGAAATAGCTCAGTTGGGAGAGCGTTAG ACTGAAGATCTAAAGGTCCATGGTTCGATCC CGGGTTTCGGCAAGA
tS(GGA)K	Phe	11	100,209,281	100,209,356	Ser	GGA		Rogue	GCTGAAATAGCTCAGTTGGGAGAGCATTAG ACTGGAGATCTAAAGGTCCATGGTTTGATCC CGGGTTTCGGCAGTA
tF(GAA)M1	Phe	13	20,540,967	20,541,042	Phe	GAA			GCCGAAATAGCTCAGTTGGGAGAGCGTTAG ACTGAAGATCTAAAGGTCCATGGTTCGATCC CGGGTTTCGGCAACA
tF(GAA)M2	Phe	13	21,263,006	21,263,081	Phe	GAA			GCCGAAATAGCTCAGTTGGGAGAGCGTTAG ACTGAAGATCTAAAGGTCCATGGTTCAATCC CGGGTTTCGGCAAAA
tF(GAA)N	Phe	14	110,410,830	110,410,905	Phe	GAA			GCCGAAATAGCTCAGTTGGGAGAGCGTTAG ACTGAAGATCTAAAGGTCCATGGTTCGATCC CGGGTTTCGGCAGTA
tF(GAA)S1	Phe	19	11,252,590	11,252,665	Phe	GAA			GCCGAAATAGCTCAGTTGGGAGAGCGTTAG ACTGAAGATCTAAAGGTCCATGGTTCGATCC CGGGTTTCGGCAGTA

tRNA gene name	tRNA family	Chromo-some	Sequence start (nt)	Sequence end (nt)	Amino acid type	Anticodon	Contains an intron	Rogue tRNA?	tRNA gene sequence (including introns)
tF(GAA)S2	Phe	19	11,258,898	11,258,975	Phe	GAA			TGCTGAAATAGCTCAGTTGGGAGAGCGTTA GACTGAAGATCTAAAGGTCACTGGTTCGATC CCGGGTTTCAGCAAAGA
tP(AGG)A	Pro	1	78,768,123	78,768,197	Pro	AGG			GGCTTGTTGGTCTAGGGGTATGATTCTCACT TAGGGTGTGAGAGGTCCTAGGTTCAAATCTT GGACGAGTCCTCA
tL(AAG)P	Pro	16	3,121,043	3,121,117	Leu	AAG		Rogue	GGCTTGTTGGTCTAGGGGTATGATTCTCACT TAAGGTCTGAGAAGTCCTAGGTTCAAAGCT TGGACGAGTCCTCA
tP(GGG)Q	Pro	17	22,133,601	22,133,676	Pro	GGG			GGCTTGTTGGTCTGGGGGTATGGTTCTCGCT TGGGGTGTGAGGGGGTCCAGGGTTCAAGTC CCGGATAACCCCGCA
tP(TGG)Q	Pro	17	22,134,692	22,134,766	Pro	TGG			GGGTCATTGGTCTATGGGCATGATTCTCTCTT TGGGTGAGAGAGGTCCCAGGTTCAAATCCC GGATGAGCCCAGA
tS(AGA)D	Ser1	4	10,800,278	10,800,362	Ser	AGA			GTAGTCGTGGCCGAGTGGTTAAGGCGATGG ACTAGAAATCCATTGGGGTATCCCCGCGCAG GTTCGAATCCTGCCGACTACGGAA
tS(TGA)J	Ser1	10	63,194,756	63,194,840	Ser	TGA			GCAGCGATGGCCGAGTGGTTAAGGCGTTGG ACTTGAAATCCAATGGGGTATCCCCGCGCAG GTTCGAACCCTGCTCGCTGCGGAA
tS(CGA)J	Ser1	10	128,585,227	128,585,311	Ser	CGA			GTCACGGTGGCCGAGTGGTTAAGGCGTTGG ACTCGAAATCCAATGGGGTATCCCCGCACAG GTTCGAATCCTGTTCGTGACGGCA
tS(AGA)K	Ser1	11	68,650,032	68,650,118	Ser	AGA			TGTAGTCGTGGCCGAGTGGTTAAGGCGATG GACTAGAAATCCATTGGGGACTCCCCGCGC AGGTTCGAATCCTGCCGACTACGTCTA
tS(CGA)K	Ser1	11	68,723,475	68,723,559	Ser	CGA			GCTGTGATGGCCGAGTGGTTAAGGCGTTGG ACTCGAAATCCAATGGGGTATCCCCGCGCAG GTTCGAATCCTGCTCACAGCGCTA
tS(AGA)M1	Ser1	13	21,299,338	21,299,424	Ser	AGA			TGTAGTCGTGGCCGAGTGGTTAAGGCGATG GACTAGAAATCCATTGGGGACTCCCCGCGC AGGTTCGAATCCTGCCGACTACGTCAA
tS(TGA)M1	Ser1	13	21,305,056	21,305,140	Ser	TGA			GTAGTCGTGGCCGAGTGGTTAAGGCGATGG ACTTGAAATCCATTGGGGTATCCCCGCGCAG GTTCGAATCCTGCCGACTACGGTA
tS(AGA)M2	Ser1	13	21,307,980	21,308,066	Ser	AGA			TGTAGTCGTGGCCGAGTGGTTAAGGCGATG GACTAGAAATCCATTGGGGACTCCCCGCGC AGGTTCGAATCCTGCCGACTACGGTAA
tS(AGA)M3	Ser1	13	21,317,214	21,317,300	Ser	AGA			TGTAGTCGTGGCCGAGTGGTTAAGGCGATG GACTAGAAATCCATTGGGGATTCCCCGCGCA GGTTCGAATCCTGCCGACTACGGGTA
tS(AGA)M4	Ser1	13	21,321,034	21,321,120	Ser	AGA			TGTAGTCGTGGCCGAGTGGTTAAGGCGATG GACTAGAAATCCATTGGGGACTCCCCGCGC AGGTTCGAATCCTGCCGACTACGGGCA
tS(CGA)M	Ser1	13	21,402,540	21,402,624	Ser	CGA			GCTGTGATGGCCGAGTGGTTAAGGCGTTGG ACTCGAAATCCAATGGGTTATTCCCGCGCAG GTTCAAATCCTGCTCACAGCGTAA
tS(AGA)M5	Ser1	13	22,881,077	22,881,163	Ser	AGA			TGTAGTCGTGGCCGAGTGGTTAAGGCGATG GACTAGAAATCCATTGGGGACTCCCCGCGC AGGTTCGAATCCTGCCGACTACGATTA
tS(TGA)M2	Ser1	13	22,898,269	22,898,355	Ser	TGA			TGTAGTCGTGGCCGAGTGGTTAAGGCGATG GACTTGAAATCCATTGGGGATTCCCCGCGCA GGTTCGAATCCTGCCGACTACGTGAA
tS(AGA)X1	Ser1	X	157,022,395	157,022,479	Ser	AGA			GTAGTCGTGGCCAAGTGAGTAAGGCAATGG ACTAGAAATCCATTGGGGTATCCCAGCACAG GTTCAAATCCTGCTGACTATGGTA
tS(GCT)A	Ser2	1	182,141,342	182,141,428	Ser	GCT			GGACGAGGTGGCCGAGTGGTTAAAGCGATG GACTGCTAATCCACTGTGCACAGTATGCGTG GGTTCGAATCCCATCCTCGTCCGAAA
tS(GCT)B1	Ser2	2	118,828,323	118,828,407	Ser	GCT			GACGAGGTGGCCGAGTGGTTAAGGCGATGG ACTGCTAATCCATTGTGCTATGCACGCATGG GTTCGAATCCCATCCTCGTCGAAA
tS(GCT)K2	Ser2	11	68,675,595	68,675,679	Ser	GCT			GACGAGGTGGCCGAGTGGTTAAGGCGATGG ACTGCTAATCCATTGTGCTATGCACGCGTGG GTTCGAATCCCATCCTCGTCGTCA
tS(GCT)M1	Ser2	13	20,872,890	20,872,974	Ser	GCT			GATGAGGTGGCCGAGTGGTTAAGGCGATGG ACTGCTAATCCATTGTGCTATGCACGCATGG GTTCGAATCCCATCCTCATCGACA
tS(GCT)M2	Ser2	13	21,374,311	21,374,395	Ser	GCT			GACGAGGTGGCCGAGTGGTTAAGGCGATGG ACTGCTAATCCATTGTGCTATGCACGCGTGG GTTCGAATCCCACCTTCGTCGTCA
tS(GCT)M3	Ser2	13	21,430,806	21,430,892	Ser	GCT			TGACGAGGTGGCCGAGTGGTTAAGGCGATG GACTGCTAATCCATTGTGCACTGCACGCGTG GGTTCGAATCCCATCCTCGTCGGTCA
tS(GCT)M4	Ser2	13	22,901,993	22,902,079	Ser	GCT			TGACGAGGTGGCCGAGTGGTTAAGGCGATG GACTGCTAATCCATTGTGCACTGCACGCGTG GGTTCGAATCCCATCCTCGTCGTTCA
tS(GCT)S	Ser2	19	4,826,771	4,826,855	Ser	GCT			GACGAGGTGGCCGAGTGGTTAAGGCGATGG ACTGCTAATCCATTGTGCTATGCACGCGTGG GTTCGAATCCCATCCTCGTCGGTA

tRNA gene name	tRNA family	Chromo-some	Sequence start (nt)	Sequence end (nt)	Amino acid type	Anticodon	Contains an intron	Rogue tRNA?	tRNA gene sequence (including introns)
tT(AGT)G	Thr1	7	23,656,861	23,656,937	Thr	AGT			GGCGCCGTGGCTTAGTTGGTTAAAGCGCCT GTCTAGTAAACAGGAGATCATGGGTTCGAAT CCCAGCGGTGCCTGAA
tM(CAT)J	Thr1	10	62,958,645	62,958,740	Met	CAT		Rogue	GGCTCTGTGGCTTAGTTGGCTAAAGTGCCTG TCTCATAAACAGGAGATCATGTTGTAAACAG GAGATCGTGGGTTTGAATCCCAGTGGGGCC TGAA
tT(AGT)K1	Thr1	11	68,650,409	68,650,485	Thr	AGT			GGCGCCGTGGCTTAGTTGGTTAAAGCGCCT GTCTAGTAAACAGGAGATCATGGGTTCGAAT CCCAGCGGTGCCTTTA
tT(AGT)K2	Thr1	11	68,675,296	68,675,374	Thr	AGT			AGGCGCCGTGGCTTAGTTGGTTAAAGCGCC TGTCTAGTAAACAGGAGATACTGGGTTCGAA TCCCAGCGGTGCCTTTGA
tT(AGT)K3	Thr1	11	68,723,021	68,723,097	Thr	AGT			GGCGCCGTGGCTTAGCTGGTTAAAGCGCCT GTCTAGTAAACAGGAGATCATGGGTTCGAAT CCCAGCGGTGCCTGGA
tT(CGT)M1	Thr1	13	20,644,628	20,644,704	Thr	CGT			GGCTCCGTGGCTTAGTTGGCTAAAGCGCCTG TCTCGTAAACAGGAGATCATGGGTTCGAATC CCAGTGGGGCCTGGA
tT(TGT)M1	Thr1	13	20,656,004	20,656,080	Thr	TGT			GGCTCCATGGCTTAGTTGGTTAAAGCGCCTG TCTTGTAAACAGGAGATCATGGGTTCGAATC CCAGTGGGGCCTATA
tT(CGT)M2	Thr1	13	20,702,861	20,702,937	Thr	CGT			GGCTCCATGGCTTAGCTGGTTAAAGCGCCTG TCTCGTAAACAGGAGATCATGGGTTCGACTC CCAGTGGGGCCTTCA
tT(AGT)M	Thr1	13	22,830,746	22,830,822	Thr	AGT			GGCTCCGTGGCTTAGCTGGTTAAAGCGCCTG TCTAGTAAACAGGAGATCATGGGTTCGAATC CCAGCGGGGCCTTTA
tT(AGT)N2	Thr1	14	50,810,563	50,810,639	Thr	AGT			GGCACCGTGGCTTAGTTGGTTAAAGCGCCT GTCTAGTAAACAGGAGATCATGGGTTCGAAT TCCAGCGGTGCCTGAA
tT(CGT)K2	Thr2	11	79,317,205	79,317,281	Thr	CGT			AGGCGCGGTGGCCAAGTGGTAAGGCGTCGG TCTCGTAAACCGAAGATCGAGGGTTCGAAC CCCGTCCGTGCCTGCGA
tT(CGT)P	Thr2	16	13,180,952	13,181,028	Thr	CGT			AGGCGCGGTGGCCAAGTGGTAAGGCGTCGG TCTCGTAAACCGAAGATCAAGGGTTCGAAC CCCGTCCGTGCCTGCCA
tW(CCA)J1	Trp	10	23,539,187	23,539,263	Trp	CCA			TGACCTCGTGGCACAATGGTAGCACGTCTG ACTCCAGATCAGAAGGTTGAGTGTTCAAAT CACGTCGGGGTCATGAA
tW(CCA)J2	Trp	10	90,952,051	90,952,125	Trp	CCA			GACCTCGTGGCGCAACGGTAGCGCGTCTGA CTCCAGATCAGAAGGCTGCATGTTCGAATCA CGTCGGGGTCATAA
tW(CCA)K1	Trp	11	61,020,452	61,020,528	Trp	CCA			TGACCTCGTGGCGCAATGGTAGCGCGTCTG ACTCCAGATCAGAAGGTTGAGTGTTCAAGT CACGTCGGGGTCAAGTA
tW(CCA)K2	Trp	11	62,358,559	62,358,635	Trp	CCA			TGACCTCGTGGCGCAATGGTAGCGCGTCTG ACTCCAGATCAGAAGGTTGAGTGTTCAAAT CACGTCGGGGTCATGAA
tW(CCA)K3	Trp	11	68,654,382	68,654,456	Trp	CCA			GGCCTCGTGGCGCAACGGTAGCGCGTCTGA CTCCAGATCAGAAGGTTGCATGTTCAAATCA CGTCGGGGTCATCA
tW(CCA)K4	Trp	11	68,676,048	68,676,122	Trp	CCA			GGCCTCGTGGCGCAACGGTAGCGCGTCTGA CTCCAGATCAGAAGGTTGCATGTTCAAATCA CGTCGGGGTCAGCA
tW(CCA)M1	Trp	13	22,878,229	22,878,305	Trp	CCA			TGACCTCGTGGCGCAACGGTAGCGCGTCTG ACTCCAGATCAGAAGGTTGAGTGTTCAAAT CACGTCGGGGTCAGTGA
tW(CCA)M2	Trp	13	22,886,642	22,886,718	Trp	CCA			TGACCTCGTGGCGCAACGGTAGCGCGTCTG ACTCCAGATCAGAAGGTTGAGTGTTCAAAT CACGTCGGGGTCAAGTA
tY(GTA)C1	Tyr	3	19,302,785	19,302,880	Tyr	GTA	Intron		CCTTCGATAGCTCAGCTGGTAGAGCGGAGG ACTGTAGCTAACTCCCCGTAAGAAGACATCC TTAGGTCGCTGGTTCGACTCCGGCTCGAAG GAGAA
tY(GTA)C2	Tyr	3	19,303,212	19,303,303	Tyr	GTA	Intron		CCTTCGATAGCTCAGCTGGTAGAGCGGAGG ACTGTAGGCTTGTGGCTGTAGACATCCTTAG GTCGCTGGTTCGATTCCGGCTCGAAGGAAA A
tY(GTA)C3	Tyr	3	92,212,300	92,212,387	Tyr	GTA	Intron		CATTCGATAGCTCAGTTGGTAGAGCAGAAG ACTGTAGTTAGTACAATATGGTAATCCTTGG GTTGCTGGTTCGATTCCATTCAAAGGA
tY(GTA)E1	Tyr	5	29,296,268	29,296,359	Tyr	GTA	Intron		CCTTCGATAGCTCAGTTGGTAGAGCGGAGG ACTGTAGTCAGTACAATATAGTAATCCTTAGG TCGCTGGTTCGATTCCGGCTCGAAGGACTA
tY(GTA)J	Tyr	10	96,777,476	96,777,585	Tyr	GTA	Intron		CCTTCGATAGCTCAGCTGGTAGAGCGGAGG ACTGTAGTCAAGAAAAATGAAGACTGAAGT GTGGACACTATGCCCCTCCTTAGAAGTGGG AACAAAACACCCTTGGAAGG

tRNA gene name	tRNA family	Chromo-some	Sequence start (nt)	Sequence end (nt)	Amino acid type	Anticodon	Contains an intron	Rogue tRNA?	tRNA gene sequence (including introns)
tY(GTA)M1	Tyr	13	22,805,766	22,805,859	Tyr	GTA	Intron		TCCTTCGATAGCTCAGTTGGTAGAGCGGAGG ACTGTAGAGTTACTAGAAAAGTGATCCTTAG GTCGCTGGTTCGAATCCGGCTCGAAGGAAC GA
tY(GTA)M2	Tyr	13	22,806,628	22,806,720	Tyr	GTA	Intron		TCCTTCGATAGCTCAGTTGGTAGAGCGGAGG ACTGTAGACTACTAATGTAGTGATCCTTAGG TCGCTGGTTCGAATCCGGCTCGAAGGAATG A
tY(GTA)M3	Tyr	13	22,807,216	22,807,304	Tyr	GTA	Intron		CCTTCGATAGCTCAGTTGGTAGAGCGGAGG ACTGTAGTATAGGTGTTGAAATCCTTAGGTC GCTGGTTCGAATCCGGCTCGAAGGAGGA
tY(GTA)M4	Tyr	13	22,808,859	22,808,950	Tyr	GTA	Intron		CTTTCGATAGTTCAGTTGGTAGAGCGGAGGA CTGTAGAGTATTAACGTTAGTGATCCTTAGG TCGCTGGTTCGAGTCCGGCTCGAAGGAAGA
tY(GTA)M5	Tyr	13	22,813,110	22,813,202	Tyr	GTA	Intron		TCCTTCGATAGCTCAGTTGGTAGAGCGGAGG ACTGTAGGTCATTGTTCTAGAAATCCTTAGG TCGCTGGTTCGAATCCGGCTCGAAGGAACC A
tY(GTA)M6	Tyr	13	22,847,199	22,847,293	Tyr	GTA	Intron		TCCTTCGATAGCTCAGTTGGTAGAGCGGAGG ACTGTAGGAGTATTCGACATGGAAATCCTTA GGTCGCTGGTTCGAATCCGGCTCGAAGGAG GTA
tY(GTA)N	Tyr	14	43,782,936	43,783,028	Tyr	GTA	Intron		CCTTCGATAGCTCAGCTGGTAGAGCGGAGG ACTGTAGTTACATTCGTTGAAGCCATCCTTA GGTCGCTGGTTCGATTCCGGCTCGAAGGAGT A
tV(CAC)A	Val1	1	171,182,812	171,182,887	Val	CAC			GTTTCCGTAGTGTAGTGGTTATCACGTTCGC CTCACACGCGAAAGGTCCACGGTTCGAAAC CGGGCGGAAACAGCA
tV(AAC)C	Val1	3	30,365,837	30,365,912	Val	AAC			GTTTCCGTAGTGTAGTGGTTATCACGTTCGC CTAACACGCGAAAGGTCCACGGTTCGAAAC CGGGCGGAAACATAA
tV(CAC)C1	Val1	3	59,607,018	59,607,094	Val	CAC			TGTTTCTGTAGTGTAGTGGTTTCACATTTGCC TCACATGCAAAAGGTCCACGGTTCTCAACC GGGCAGAAACAACTA
tV(CAC)C3	Val1	3	96,338,468	96,338,543	Val	CAC			GTTTCCGTAGTGTAGTGGTTATCACGTTCGC CTCACACGCGAAAGGTCCACGGTTCGAAAC CGGGCGGAAACAAGA
tV(AAC)E	Val1	5	15,480,961	15,481,036	Val	AAC			GTTTCCGTAGTGTAGTGGTTATCATGTTTGTC TAACACGCGAAAGGTCCACAGTTTGAAACC GGGTGGAAAAAAAA
tV(CAC)F	Val1	6	9,831,301	9,831,376	Val	CAC			GTTTCTGTAGTGTAGTGGTTATCACGTTCGC CTCACACGCGAAAGGTCCACGGTTCGAAAC CGGGCAGAAACAAGA
tV(CAC)K1	Val1	11	48,458,277	48,458,352	Val	CAC			GTTTCCGTAGTGTAGTGGTTATCACGTTCGC CTCACACGCGAAAGGTCCACGGTTCGAAAC CGGGCGGAAACAACA
tG(ACC)K	Val1	11	48,463,463	48,463,539	Gly	ACC		Rogue	GTTTCCGTAGTGTAGTGGTTAGCGCGTTCGC CTACCAAAGCGAAAGGTCACCGGTTCGAAA CCGGGCGGAAACAAAA
tV(AAC)K1	Val1	11	48,496,372	48,496,447	Val	AAC			GTTTCCGTAGTGTAGTGGTTATCACGTTCGC CTAACACGCGAAAGGTCCACGGTTCGAAAC CGGGCGGAAACAAGA
tV(AAC)M1	Val1	13	21,395,911	21,395,988	Val	AAC			TGTTTCCGTAGTGTAGTGGTTATCACATTCGC CTAACACGCGAAAGGTCACCGGTTCGAAAC CGGGCGGAAACACGTA
tV(CAC)M1	Val1	13	21,406,671	21,406,748	Val	CAC			TGTTTCCGTAGTGTAGTGGTTATCACGTTCG CCTCACACGCGAAAGGTCACCGGTTCGAAA CCGGGCGGAAACAATGA
tV(AAC)M2	Val1	13	22,665,524	22,665,601	Val	AAC			TGTTTCCGTAGTGTAGTGGTCATCACGCTCG CCTAACACGCGAGAGGTCACCGGTTCGAAA CCGGGCGGAAACATTAA
tV(AAC)M3	Val1	13	22,678,904	22,678,981	Val	AAC			TGTTTCTGTAGTGTAGTGGTTATCACGCTCG CCTAACACGCGAGAGGTCACCGGTTCGAAA CCGGGCAGAAACAGTGA
tV(CAC)M2	Val1	13	22,680,245	22,680,322	Val	CAC			TGTTTCCGTAGTGTAGTGGTCATCACGCTCG CCTCACACGCGAGAGGTCACCGGTTCGAAA CCGGGCGGGAACAACAA
tV(CAC)M3	Val1	13	22,687,753	22,687,828	Val	CAC			GTTTCCGTAGTGTAGTGGTTATCACGCTCGC CTCACACGCGAGAGGTCCACGGTTCGAAAC CGGGCGGAAACAGTA
tV(AAC)M4	Val1	13	22,781,193	22,781,270	Val	AAC			TGTTTCCGTAGTGTAGTGGTCATCACGCTCG CCTAACACGCGAGAGGTCACCGGTTCGAAA CCGGGCGGGAACATTTA
tV(CAC)M4	Val1	13	22,792,197	22,792,274	Val	CAC			TGTTTTTGTAGTGTAGCGGTTATCACGCTCG CCTCACACGCGAGAGGTCATCGGTTCAAAA CCCAGTGGAAACATTTA

tRNA gene name	tRNA family	Chromo-some	Sequence start (nt)	Sequence end (nt)	Amino acid type	Anticodon	Contains an intron	Rogue tRNA?	tRNA gene sequence (including introns)
tV(AAC)M5	Val1	13	22,793,368	22,793,445	Val	AAC			TGTTTCCGTAGTGTAGTGGTCATCACGCTCG CCTAACACGCGAGAGGTCACCGGTTCGAAA CCGGGCGGAAACATGGA
tV(CAC)M5	Val1	13	22,826,263	22,826,338	Val	CAC			GTTTCCGTAGTGTAGTGGTTATCACGTTCGC CTCACACGCGAAAGGTCCACGGTTCGAAAC CGGGCGGAAACAGTA
tV(CAC)Q2	Val1	17	57,400,716	57,400,791	Val	CAC			GTTTCCCTAGTGTAGTGGTTATCACGTTCGC CTCACATGCAAAAGGTCCATGGTTCGAAAC CTGGCGGAAACAGTA
tV(CAC)X7	Val1	X	37,314,459	37,314,530	Val	CAC			GTTTCAGTAGTGTAGTGGTTATCACGTTTGC CTCACGCATGAAATGTCCCCGGTTGGAAAC CTGGGGAAACA
tV(TAC)S1	Val2	19	11,262,285	11,262,360	Val	TAC			GGTTCCATAGTGTAGCGGTTATCACGTCTGC TTTACACGCAGAAGGTCCAGGGTTCGAGCC CCAGTGGAACCACGA
tV(TAC)S2	Val2	19	11,262,619	11,262,696	Val	TAC			TGGTTCCATAGTGTAGCGGTTATCACGTCTG CTTTACACGCAGAAGGTCATGGGTTCGAGC CCCAGTGGAACCAGTGA
tV(TAC)X2	Val2	X	150,735,962	150,736,039	Val	TAC			TGGTTCCATAGTGTAGCGGTTATCACGTCTG CTTTACACGCAGAAGGTCATGGGTTCGAGC CCCAGTGGAACCATAGA
tH(ATG)A		1	171,740,670	171,740,747	His	ATG			CGGACCAGTGGTGCAATGGATAACACGTCT GACTATGGATCAGAAGATTATAGGTTCGACT CCTGTCTGGCTCGTTGA
tY(GTA)B		2	22,579,293	22,579,362	Tyr	GTA			GGTAAAATGGCTGAGTAAGCATTAGACTGTA AATCTAAACACAGAGGTTAAAATCCTCTTTT TACCAGAA
tT(TGT)E1		5	24,158,456	24,158,574	Thr	TGT	Intron		TTAGGGTTGGCTTACAGGATCTGAGGTTCAG TCCATTACCATCAAGGCGAGAGCATGAGAG TGTCCAGGTAATTGTGGGACTGGAGGAGCT GAGGGTTCAATATCTTCATCCAAAGGTA
tG(ACC)E		5	91,642,399	91,642,465	Gly	ACC			GGATGTAGCTCTGTGGTAGAACACTTGTCTA CCATACACAGGGTTCTGGGTTCGAATCCCAG CACCA
tSec(TCA)G		7	10,824,326	10,824,415	Sec	TCA			GCCCGGATGATCCTCAGTGGTCTGGGGTGC AGGCTTCAAACCTGTAGCTATTTAGCGACAG AGTGGTTCAATTCCACCTTTCGGGCGGGA
tV(TAC)G		7	23,020,210	23,020,352	Val	TAC	Intron		GGGGATGTGGCTCAGTGGTAGAACACTTGC CTAGAATCCCCCCCCCCCCAAGTGAGGGATT GGGATGGGGGGGTGGGGCTTAGTTCTAGAG CAATTGTCAACTACGTGCAAGGCTCTGGGTT CAATTCTCAGTATGCCCCAAA
tA(TGC)G		7	113,109,603	113,109,673	Ala	TGC			GAGATGATAGCTCAGTGGTAGAGGACACAC CCTGCACCTGTGAGGTCCTGGGTTCTAGCCG CAGCATCCCC
tI(AAT)G		7	125,033,749	125,033,823	Ile	AAT			GGGTAGGCGTGGCTTAGTGGCAAAGTAGGT ACTTAATATGCACAAGGCCCTGGGTTCAATC TCCAGCAGCACCCA
tA(GGC)H2		8	109,962,872	109,962,942	Ala	GGC			GCGTTGGTAGTTTAACGGGAGAATTCTCCAC TGGCACGTGGGAGCCCTGGGTTTGATTACC AGCCAGTACA
tG(CCC)I		9	57,369,946	57,370,021	Gly	CCC			GGGGATCTAGCCCAGTGGTAGAGCGCACAC TCTTCCCATGTGTGAGGTCATGGGTTCAATC CCAGCATCTCCAAGA
tT(GGT)J		10	126,722,048	126,722,119	Thr	GGT			TTAGGGATTGAGCCCAGTGGTAGGGCCTTGC CTGGTAAGCACAAGGTCTTAGGTTCAGTTCT CACCTCTGAG
tV(AAC)K2		11	64,344,654	64,344,725	Val	AAC			GCTGACCTTAGCTCAGTGGCAGAGCATGGG ACTAACACGCACAAAGCTCTGGGTTCTGTT CCCAGTGCAACA
tL(CAA)K		11	115,171,001	115,171,129	Leu	CAA	Intron		GTCAGACTAGCTGGCTAAGGGCACAGCTCA GTGGACAGGGTGGTGGGCAATGTGTGGGAG GGACTGGTCTAACAGTGCAGAGTTCTCAAA GCACTGAGCAGCACTGTCCGAGTCAGCGAT CTGGTCAGA
tV(CAC)N		14	39,549,215	39,549,298	Val	CAC			TCCGGGTGGGGTGTGCAGGGAGCACTTGCG GGTCACAGCCGCAGCAACAAGAAGCCGCG GGGTCTCCTCCCGCACCTGGGAAGA
tS(AGA)N		14	57,933,281	57,933,354	Ser	AGA			CTGGGGACTTAGCTCAGTGGTAGAGTGTTTA CTTAGAAAGGGCAAGACCCTGGGTTCCATC CTCAGCTCCTAGG

tRNA gene name	tRNA family	Chromo-some	Sequence start (nt)	Sequence end (nt)	Amino acid type	Anticodon	Contains an intron	Rogue tRNA?	tRNA gene sequence (including introns)
tF(GAA)O		15	3,687,891	3,688,442	Phe	GAA	Intron		AAGGAGGTAGAGAAAGGACCCAAGGAGCT GAAGGGGTTTGCAGCCCCATAAGAGGAACA ACAATATGAATGAACCAGTATCTCCAGAGCT CCCTGGGACAAAACTGCCAACCAAAGAAA ACACACGGTGGGACTCGTGGTTCTAGTGCA AATGTAGCAGAAGATGGCCTAGTCGGTCATC AATGGGAGGAGAGGCCCTTAATACTGTGAA GGTTAAATGCCCCAGTATAGGGGACTGCAA GGGCCAGGAAGAAGGAGTGTGTACTGGCTA GTTTTGTGTCAACTTGACACAGGTGGAGTAA TCACAGAGAAGGAGCTTCAGTTGAGGAAAT GCCTCCATGAGATCCAGCATTAAGGCATTTT CTCAATTAGTGATCAAGGGGGAAAGGCCCC TTGTGGGAGGGACCATCTCTGGGCTGGTAG TCTTGGGTTCTATAAGAGAGCAGGCTGAGC AAGCCAGGAGAAGCAAGCCAGTAAAGAAC ATCCCTCCATGGCCTCTGAATCAGCTCCTGC TTCCTGACCTGCTTGAGTTCTATTCCTGACTT CCTTGATA
tT(TGT)O		15	27,354,831	27,354,908	Thr	TGT			GGTGCGGTGGCTGAGCTGGTTAAAGCACCT GTCTTGTTAACAGGGCAGCCTGGGTTTGATT CCCAGAGCCCCCCACCC
tA(GGC)O1		15	42,435,453	42,435,534	Ala	GGC	Intron		GGGGATGGAGGAATGGTTCAGTCCTTAAGA GCACTGGCTGCTCTTGCAGAGGACCTGGGT TCAATTCCCAGCACCCCCATGA
tP(AGG)P		16	93,798,387	93,798,473	Pro	AGG	Intron		CAAGGGAGTGGGCAAGGTAGCTCAGTGATT AGGAACACTTGCTTTTCTTACTGAGGACCTG GGTTCAATTCCCAGCACCCATGTAGA
tS(TGA)X		X	126,591,980	126,592,051	Ser	TGA			TGAGACTTAGCTCAGTGGTAGAGCTGGTGTT CTGAAAGCATGTGGTCCCAGGTTCAATTCCT GGGGTCTTTA
tT(TGT)X2		X	126,614,305	126,614,379	Thr	TGT			GGGGAGGTAGCTCAATGGTAGAGCACATGC TTTGTGTGTATGAGGCACCAGGTTCTATTCC CGCCTGCTTTTTTA
tN(ATT)X		X	132,925,018	132,925,089	Asn	ATT			GGTTATATAGCTCAGTGGTAGAGTACGTGTT CATTATGCAGGAGGCCCTAGATTCCATCTCT AGTACAAAAA
tT(CGT)X		X	140,986,759	140,986,833	Thr	CGT			GGGGATGTAGCTCAGAGATAGAGTGTGTGC CTCGTATGTGTGAGGTCCCAGGTTCAATCCC CTGCATCTCCAAGA

Appendix F

List of tRNA genes that were identified in the human genome and had the expression confirmed. The tRNA gene name is based on tRNA gene nomenclature in yeast. The gene coordinates are based on the March 2006 release of the mouse genome (hg18). The presence of an intron is indicated as well.

tRNA gene name	tRNA family	Chromo-some	Sequence Start	Sequence End	Amino acid type	Anti-codon	Intron	tRNA gene sequence (including introns)
tN(GTT)A1	Hs_Asn	1	16,592,459	16,592,386	Asn	GTT		GTCTCTGTGGCGCAATCGGTTAGCG CGTTCGGCTGTTAACTGAAAGGTTG GTGGTTCGAGCCCACCCAGGGACG
tE(TTC)A1	Hs_Glu1	1	16,607,151	16,607,080	Glu	TTC		TCCCTGGTGGTCTAGTGGCTAGGAT TCGGCGCTTTCACCGCCGCGGCCCG GGTTCGATTCCCGGTCAGGGAA
tG(CCC)A1	Hs_Gly1	1	16,617,810	16,617,740	Gly	CCC		GCATTGGTGGTTCAGTGGTAGAATT CTCGCCTCCCACGCGGGAGACCCG GGTTCAATTCCCGGCCAATGCA
tG(CCC)A2	Hs_Gly1	1	16,750,142	16,750,072	Gly	CCC		GCGTTGGTGGTTTAGTGGTAGAATT CTCGCCTCCCATGCGGGAGACCCGG GTTCAATTCCCGGCCACTGCA
tV(CAC)A1	Hs_Val	1	16,751,879	16,751,807	Val	CAC		GTTTCTGTGGTGTAGTGGTTATCATG TTCGCCTCACACGAGAAAAGTCCCT GATTCGAGACTGGGTGGGAACG
tG(CCC)A3	Hs_Gly1	1	16,799,086	16,799,156	Gly	CCC		GCCTTGGTGGTGCAGTGGTAGAATT CTCGCCTCCCACGTGGGAGACCCG GGTTCAATTCCCGGCCAATGCA
tG(CCC)A4	Hs_Gly1	1	16,933,722	16,933,792	Gly	CCC		GCATTGGTGGTTCAGTGGTAGAATT CTCGCCTCCCACGCGGGAGACCCG GGTTCAATTCCCGGCCAATGCA
tE(TTC)A2	Hs_Glu	1	16,944,384	16,944,455	Glu	TTC		TCCCTGGTGGTCTAGTGGCTAGGAT TCGGCGCTTTCACCGCCGCGGCCCG GGTTCGATTCCCGGCCAGGGAA
tN(GTT)A2	Hs_Asn	1	16,947,264	16,947,337	Asn	GTT		GTCTCTGTGGTGCAATCGGTTAGCG CGTTCGGCTGTTAACCATAAGGTTG GTGGTTAGAGACCACCCAGGGACG
tN(GTT)A3	Hs_Asn	1	16,961,478	16,961,551	Asn	GTT		GTCTCTGTGGCGCAATCGGTTAGCG CGTTCGGCTGTTAACCGAAAGATTG GTGGTTCGAGCCCACCCAGGGACG
tK(CTT)A1	Hs_Lys1	1	55,135,635	55,135,563	Lys	CTT		GCCCAGCTAGCTCAGTCGGTAGAGC ATGAGACTCTTAATCTCAGGGTCAT GGGTTTGAGCCCCACGTTTGGTG
tC(GCA)A	Hs_Cys	1	93,693,927	93,693,855	Cys	GCA		GGGGGTATAGCTCAGGTGGTAGAGC ATTTGACTGCAGATCAAGAGGTCCC CGGTTCAAATCCGGGTGCCCCCT
tR(TCT)A1	Hs_Arg4	1	94,025,150	94,025,234	Arg	TCT	Yes	GGCTCCGTGGCGCAATGGATAGCGC ATTGGACTTCTAGAGGCTGAAGGCA TTCAAAGGTTCCGGGTTCGAGTCCC GGCGGAGTCG
tN(GTT)A4	Hs_Asn	1	141,878,966	141,879,039	Asn	GTT		GTCTCTGTGGTGCAATCGGTTAGCG CGTTCCGCTGTTAACCGAAAGCTTG GTGGTTCGAGCCCACCCAGGGATG
tN(GTT)A5	Hs_Asn	1	141,886,042	141,885,969	Asn	GTT		GTCTCTGTGGCGCAATCGGTTAGCG CGTTTGACTGTTAACTGAAAGGTTG GTGGTGCAAGCCCATCCAGGGATG
tN(GTT)A6	Hs_Asn	1	142,059,195	142,059,268	Asn	GTT		GTCTCTGTGGTGCAATCGGTTAGCG CGTTCCGCTGTTAACCGAAAGCTTG GTGGTTTGAGCCCACCCAGGGATG

tN(GTT)A7	Hs_Asn	1	142,066,271	142,066,198	Asn	GTT	GTCTCTGTGGCGCAATCGGTTAGCG CGTTTGACTGTTAACTGAAAGGTTG GTGGTGCAAGCCCATCCAGGGATG
tK(CTT)A2	Hs_Lys1	1	142,884,638	142,884,566	Lys	CTT	GCCCGGCTAGCTCAGTCGGTAGAGC ATGAGACTCTTAATCTCAGGGTCGT GGGTTCGAGCCCCACGTTGGGCG
tH(GTG)A1	Hs_His	1	142,885,996	142,885,925	His	GTG	GCCGTGATCGTATAGTGGTTAGTACT CTGCGTTGTGGCCGCAGCAACCTCG GTTCGAATCCGAGTCACGGCA
tG(TCC)A1	Hs_Gly2	1	142,886,979	142,886,908	Gly	TCC	GCGTTGGTGGTATAGTGGTGAGCAT AGCTGCCTTCCAAGCAGTTGACCCG GGTTCGATTCCCGGCCAACGCA
tE(CTC)A1	Hs_Glu1	1	142,888,348	142,888,277	Glu	CTC	TCCCTGGTGGTCTAGTGGTTAGGAT TCGGCGCTCTCACCGCCGCGGCCCG GGTTCGATTCCCGGTCAGGGAA
tQ(CTG)A1	Hs_Gln	1	143,432,853	143,432,924	Gln	CTG	GGTTCCATGGTGTAATGGTGAGCAC TCTGGACTCTGAATCCAGCGATCCG AGTTCGAGTCTCGGTGGAACCT
tN(GTT)A8	Hs_Asn	1	143,448,656	143,448,583	Asn	GTT	GTCTCTGTGGCGCAATCGGTTAGCG CGTTCGGCTGTTAACTGAAAGGTTA GTGGTTCGAGCCCACCCGGGGACG
tH(GTG)A2	Hs_His	1	143,769,589	143,769,660	His	GTG	GCCGTGATCGTATAGTGGTTAGTACT CTGCGTTGTGGCCGCAGCAACCTCG GTTCGAATCCGAGTCACGGCA
tQ(CTG)A2	Hs_Gln	1	144,729,854	144,729,925	Gln	CTG	GGTTCCATGGTGTAATGGTGAGCAC TCTGGACTCTGAATCCAGCGATCCG AGTTCGAGTCTCGGTGGAACCT
tN(GTT)A9	Hs_Asn	1	144,745,656	144,745,583	Asn	GTT	GTCTCTGTGGCGCAATCGGTTAGCG CGTTCGGCTGTTAACTGAAAGGTTA GTGGTTCGAGCCCACCCGGGGACG
tQ(CTG)A3	Hs_Gln	1	144,852,365	144,852,294	Gln	CTG	GGTTCCATGGTGTAATGGTAAGCAC TCTGGACTCTGAATCCAGCGATCCG AGTTCGAGTCTCGGTGGAACCT
tH(GTG)A3	Hs_His	1	144,868,383	144,868,454	His	GTG	GCCGTGATCGTATAGTGGTTAGTACT CTGCGTTGTGGCCGCAGCAACCTCG GTTCGAATCCGAGTCACGGCA
tH(GTG)A4	Hs_His	1	144,889,828	144,889,757	His	GTG	GCCGTGATCGTATAGTGGTTAGTACT CTGCGTTGTGGCCGCAGCAACCTCG GTTCGAATCCGAGTCACGGCA
tQ(CTG)A4	Hs_Gln	1	144,915,849	144,915,920	Gln	CTG	GGTTCCATGGTGTAATGGTAAGCAC TCTGGACTCTGAATCCAGCCATCTG AGTTCGAGTCTCTGTGGAACCT
tN(GTT)A10	Hs_Asn	1	144,992,449	144,992,522	Asn	GTT	GTCTCTGTGGCGCAATCGGTTAGCG CGTTTGACTGTTAACTGAAAGGTTG GTGGTGCAAGCCCATCCAGGGATG
tN(GTT)A11	Hs_Asn	1	145,115,717	145,115,790	Asn	GTT	GTCTCTGTGGCGTAGTCGGTTAGCG CGTTCGGCTGTTAACCGAAAAGTTG GTGGTTCGAGCCCACCCAGGAACG
tN(GTT)A12	Hs_Asn	1	145,377,979	145,377,906	Asn	GTT	GTCTCTGTGGCGCAATCGGTTAGCG CATTCGGCTGTTAACCGAAAGGTTG GTGGTTCGAGCCCACCCAGGGACG
tN(GTT)A13	Hs_Asn	1	145,540,021	145,539,948	Asn	GTT	GTCTCTGTGGCGCAATCGGTTAGCG CATTCGGCTGTTAACCGAAAGGTTG GTGGTTCGAGCCCACCCAGGGACG
tH(GTG)A5	Hs_His	1	145,935,491	145,935,420	His	GTG	GCCATGATCGTATAGTGGTTAGTACT CTGCGCTGTGGCCGCAGCAACCTC GGTTCGAATCCGAGTCACGGCA
tQ(CTG)A5	Hs_Gln	1	145,965,717	145,965,788	Gln	CTG	GGTTCCATGGTGTAATGGTAAGCAC TCTGGACTCTGAATCCAGCGATCCG AGTTCGAGTCTCGGTGGAACCT

tN(GTT)A14	Hs_Asn	1	146,010,235	146,010,162	Asn	GTT	GTCTCTGTGGCGCAATGGGTTAGCG CGTTCGGCTGTTAACCGAAAGGTTG GTGGTTCGAGCCCATCCAGGGACG
tV(CAC)A2		1	146,074,328	146,074,258	Val	CAC	GCACTGGTGGTTCAGTGGTAGAATT CTCGCCTCACACGCGGGACACCCG GGTTCAATTCCCGGTCAAGGCA
tV(CAC)A3	Hs_Val	1	146,078,219	146,078,147	Val	CAC	GTTTCCGTAGTGTAGTGGTTATCAC GTTCGCCTCACACGCGAAAGGTCC CCGGTTCGAAACTGGGCGGAAACA
tN(GTT)A15	Hs_Asn	1	146,421,682	146,421,755	Asn	GTT	GTCTCTGTGGCGCAATCGGTTAGCG CGTTTGACTGTTAACTGAAAGGTTG GTGGTGCAAGCCCATCCAGGGATG
tN(GTT)A16	Hs_Asn	1	146,428,763	146,428,690	Asn	GTT	GTCTCTGTGGTGCAATCGGTTAGCG CGTTCCGCTGTTAACCGAAAGCTTG GTGGTTCGAGCCCACCCAGGGATG
tE(TTC)A3	Hs_Glu1	1	146,477,500	146,477,428	Glu	TTC	TCCCTGGTGGTCTAGTGGCTAGGAT TCGGCGCTTTCACCGCCTGCAGCTC GAGTTCGATTCCTGGTCAGGGAA
tV(CAC)A4	Hs_Val	1	146,497,234	146,497,161	Val	CAC	GTTTCCGTAGTGTAGTGGTTATCAC GTTCGCCTCACACGCGTAAAGGTCC CCGGTTCGAAACCGGGCGGAAACA
tN(GTT)A17	Hs_Asn	1	146,524,944	146,524,871	Asn	GTT	GTCTCTGTGGCGCAATCGGCTAGCG CGTTTGGCTGTTAACTAAAAGGTTG GTGGTTCGAACCCACCCAGAGGCG
tM(CAT)A	Hs_Met2	1	150,456,799	150,456,870	Met	CAT	AGCAGAGTGGCGCAGCGGAAGCGT GCTGGGCCCATAACCCAGAGGTCGA TGGATCGAAACCATCCTCTGCTA
tR(TCT)A2	Hs_Asn	1	155,924,547	155,924,474	Arg	TCT	GTCTCTGTGGCGCAATGGACGAGC GCGCTGGACTTCTAATCCAGAGGTT CCGGGTTCGAGTCCCGGCAGAGAT G
tV(CAC)A5	Hs_Val	1	158,182,635	158,182,563	Val	CAC	GTTTCCGTAGTGTAGTGGTTATCAC GTTCGCCTCACACGCGAAAGGTCC CCGGTTCGAAACCGGGCGGAAACA
tE(TTC)A4	Hs_Glu1	1	158,205,027	158,204,956	Glu	TTC	TCCCTGGTGGTCTAGTGGCTAGGAT TCGGCGCTTTCACCGCCGCGGCCCG GGTTCGATTCCCGGTCAGGGAA
tN(GTT)A18	Hs_Asn	1	158,211,013	158,210,940	Asn	GTT	GTCTCTGTGGCGCAATCGGTTAGCG CGTTCGGCTGTTAACCGAAAGGTTG GTGGTTCGAGCCCACCCAGGGACG
tG(TCC)A2	Hs_Gly2	1	158,223,105	158,223,034	Gly	TCC	GCGTTGGTGGTATAGTGGTGAGCAT AGTTGCCTTCCAAGCAGTTGACCCG GGCTCGATTCCCGCCCAACGCA
tD(GTC)A1	Hs_Asp	1	158,223,759	158,223,688	Asp	GTC	TCCTCGTTAGTATAGTGGTGAGTATC CCCGCCTGTCACGCGGGAGACCGG GGTTCGATTCCCCGACGGGGAG
tL(CAG)A1	Hs_Leu2	1	158,224,396	158,224,478	Leu	CAG	GTCAGGATGGCCGAGCGGTCTAAG GCGCTGCGTTCAGGTCGCAGTCTCC CCTGGAGGCGTGGGTTCGAATCCCA CTCCTGACA
tG(GCC)A1	Hs_Gly1	1	158,226,167	158,226,237	Gly	GCC	GCATGGGTGGTTCAGTGGTAGAATT CTCGCCTGCCACGCGGGAGGCCCG GGTTCGATTCCCGGCCCATGCA
tE(CTC)A2	Hs_Glu	1	158,230,144	158,230,073	Glu	CTC	TCCCTGGTGGTCTAGTGGTTAGGAT TCGGCGCTCTCACCGCCGCGGCCCG GGTTCGATTCCCGGTCAGGGAA
tG(TCC)A3	Hs_Gly2	1	158,230,501	158,230,430	Gly	TCC	GCGTTGGTGGTATAGTGGTGAGCAT AGCTGCCTTCCAAGCAGTTGACCCG GGTTCGATTCCCGGCCAACGCA
tD(GTC)A2	Hs_Asp	1	158,231,159	158,231,088	Asp	GTC	TCCTCGTTAGTATAGTGGTGAGTATC CCCGCCTGTCACGCGGGAGACCGG GGTTCGATTCCCCGACGGGGAG

tL(CAG)A2	Hs_Leu2	1	158,231,796	158,231,878	Leu	CAG	GTCAGGATGGCCGAGCGGTCTAAG GCGCTGCGTTCAGGTCGCAGTCTCC CCTGGAGGCGTGGGTTCGAATCCCA CTCCTGACA
tG(GCC)A2	Hs_Gly1	1	158,233,522	158,233,592	Gly	GCC	GCATGGGTGGTTCAGTGGTAGAATT CTCGCCTGCCACGCGGGAGGCCCG GGTTCGATTCCCGGCCCATGCA
tE(CTC)A3	Hs_Glu	1	158,237,524	158,237,453	Glu	CTC	TCCCTGGTGGTCTAGTGGTTAGGAT TCGGCGCTCTCACCGCCGCGGCCCG GGTTCGATTCCCGGTCAGGGAA
tG(TCC)A4	Hs_Gly2	1	158,237,882	158,237,811	Gly	TCC	GCGTTGGTGGTATAGTGGTGAGCAT AGCTGCCTTCCAAGCAGTTGACCCG GGTTCGATTCCCGGCCAACGCA
tD(GTC)A3	Hs_Asp	1	158,238,540	158,238,469	Asp	GTC	TCCTCGTTAGTATAGTGGTGAGTATC CCCGCCTGTCACGCGGGAGACCGG GGTTCGATTCCCCGACGGGGAG
tL(CAG)A3	Hs_Leu2	1	158,239,177	158,239,259	Leu	CAG	GTCAGGATGGCCGAGCGGTCTAAG GCGCTGCGTTCAGGTCGCAGTCTCC CCTGGAGGCGTGGGTTCGAATCCCA CTCCTGACA
tG(GCC)A3	Hs_Gly1	1	158,240,953	158,241,023	Gly	GCC	GCATGGGTGGTTCAGTGGTAGAATT CTCGCCTGCCACGCGGGAGGCCCG GGTTCGATTCCCGGCCCATGCA
tE(CTC)A4	Hs_Glu1	1	158,244,935	158,244,864	Glu	CTC	TCCCTGGTGGTCTAGTGGTTAGGAT TCGGCGCTCTCACCGCCGCGGCCCG GGTTCGATTCCCGGTCAGGGAA
tG(TCC)A5	Hs_Gly2	1	158,245,292	158,245,221	Gly	TCC	GCGTTGGTGGTATAGTGGTGAGCAT AGCTGCCTTCCAAGCAGTTGACCCG GGTTCGATTCCCGGCCAACGCA
tD(GTC)A4	Hs_Asp	1	158,245,950	158,245,879	Asp	GTC	TCCTCGTTAGTATAGTGGTGAGTATC CCCGCCTGTCACGCGGGAGACCGG GGTTCGATTCCCCGACGGGGAG
tL(CAG)A4	Hs_Leu2	1	158,246,587	158,246,669	Leu	CAG	GTCAGGATGGCCGAGCGGTCTAAG GCGCTGCGTTCAGGTCGCAGTCTCC CCTGGAGGCGTGGGTTCGAATCCCA CTCCTGACA
tG(GCC)A4	Hs_Gly1	1	158,248,313	158,248,383	Gly	GCC	GCATGGGTGGTTCAGTGGTAGAATT CTCGCCTGCCACGCGGGAGGCCCG GGTTCGATTCCCGGCCCATGCA
tE(CTC)A5	Hs_Glu1	1	158,252,315	158,252,244	Glu	CTC	TCCCTGGTGGTCTAGTGGTTAGGAT TCGGCGCTCTCACCGCCGCGGCCCG GGTTCGATTCCCGGTCAGGGAA
tG(TCC)A6	Hs_Gly2	1	158,252,673	158,252,602	Gly	TCC	GCGTTGGTGGTATAGTGGTGAGCAT AGCTGCCTTCCAAGCAGTTGACCCG GGTTCGATTCCCGGCCAACGCA
tD(GTC)A5	Hs_Asp	1	158,253,331	158,253,260	Asp	GTC	TCCTCGTTAGTATAGTGGTGAGTATC CCCGCCTGTCACGCGGGAGACCGG GGTTCGATTCCCCGACGGGGAG
tL(CAG)A5	Hs_Leu2	1	158,253,968	158,254,050	Leu	CAG	GTCAGGATGGCCGAGCGGTCTAAG GCGCTGCGTTCAGGTCGCAGTCTCC CCTGGAGGCGTGGGTTCGAATCCCA CTCCTGACA
tG(GCC)A5	Hs_Gly1	1	158,263,411	158,263,481	Gly	GCC	GCATAGGTGGTTCAGTGGTAGAATT CTTGCCTGCCACGCAGGAGGCCCA GGTTTGATTCCTGGCCCATGCA
tG(GCC)A6	Hs_Gly1	1	158,306,762	158,306,692	Gly	GCC	GCATTGGTGGTTCAGTGGTAGAATT CTCGCCTGCCACGCGGGAGGCCCG GGTTCGATTCCCGGCCAATGCA
tL(CAG)A6	Hs_Leu2	1	158,313,269	158,313,187	Leu	CAG	GTCAGGATGGCCGAGCGGTCTAAG GCGCTGCGTTCAGGTCGCAGTCTCC CCTGGAGGCGTGGGTTCGAATCCCA CTCCTGACA
tG(TCC)A7	Hs_Gly2	1	158,313,958	158,314,029	Gly	TCC	GCGTTGGTGGTATAGTGGTGAGCAT AGCTGCCTTCCAAGCAGTTGACCCG GGTTCGATTCCCGGCCAACGCA
tN(GTT)A19	Hs_Asn	1	158,323,086	158,323,159	Asn	GTT	GTCTCTGTGGCGCAATCGGTTAGCG CGTTCGGCTGTTAACCGAAAGGTTG GTGGTTCGATCCCACCCAGGGACG

tP(CGG)A	Hs_Pro	1	164,415,620	164,415,691	Pro	CGG		GGCTCGTTGGTCTAGGGGTATGATT CTCGCTTCGGGTGCGAGAGGTCCC GGGTTCAAATCCCGGACGAGCCC
tP(AGG)A	Hs_Pro	1	164,416,454	164,416,383	Pro	AGG		GGCTCGTTGGTCTAGGGGTATGATT CTCGCTTAGGGTGCGAGAGGTCCCG GGTTCAAATCCCGGACGAGCCC
tK(TTT)A1	Hs_Lys1	1	201,207,312	201,207,384	Lys	TTT		GCCCGGATAGCTCAGTCGGTAGAGC ATCAGACTTTTAATCTGAGGGTCCA GGGTTCAAGTCCCTGTTCGGGCG
tK(TTT)A2	Hs_Lys1	1	201,207,887	201,207,815	Lys	TTT		GCCCGGATAGCTCAGTCGGTAGAGC ATCAGACTTTTAATCTGAGGGTCCA GGGTTCAAGTCCCTGTTCGGGCG
tK(TTT)A3	Hs_Lys1	1	202,174,928	202,175,000	Lys	TTT		GCCCGGAGAGCTCAGTGGGTAGAG CATCAGACTTTTAATCTGAGGGTCC AGGGTTCAAGTCCTCGTTCGGGCA
tT(TGT)A	Hs_Thr2	1	219,026,742	219,026,814	Thr	TGT		GGCTCCATAGCTCAGTGGTTAGAGC ACTGGTCTTGTAAACCAGGGGTCGC GAGTTCGATCCTCGCTGGGGCCT
tL(CAA)A	Hs_Leu2	1	245,377,805	245,377,910	Leu	CAA	Yes	GTCAGGATGGCCGAGTGGTCTAAG GCGCCAGACTCAAGGTAAGCACCT TGCCTGCGGGCTTTCTGGTCTCCGG ATGGAGGCGTGGGTTCGAATCCCAC TTCTGACA
tE(CTC)A6	Hs_Glu1	1	245,378,198	245,378,269	Glu	CTC		TCCCTGGTGGTCTAGTGGTTAGGAT TCGGCGCTCTCACCGCCGCGGCCCG GGTTCGATTCCCGGTCAGGAAA
tY(GTA)B	Hs_Tyr	2	27,185,301	27,185,389	Tyr	GTA	Yes	CCTTCGATAGCTCAGTTGGTAGAGC GGAGGACTGTAGTGGATAGGGCGT GGCAATCCTTAGGTCGCTGGTTCGA TTCCGGCTCGAAGGA
tA(AGC)B	Hs_Ala2	2	27,185,733	27,185,805	Ala	AGC		GGGGGATTAGCTCAAATGGTAGAGC GCTCGCTTAGCATGCGAGAGGTAGC GGGATCGATGCCCGCATCCTCCA
tI(TAT)B	Hs_Ile2	2	42,949,327	42,949,419	Ile	TAT	Yes	GCTCCAGTGGCGCAATCGGTTAGCG CGCGGTACTTATACAGCAGTACATG CAGAGCAATGCCGAGGTTGTGAGT TCGAGCCTCACCTGGAGCA
tG(CCC)B	Hs_Gly3	2	70,387,844	70,387,774	Gly	CCC		GCGCCGCTGGTGTAGTGGTATCATG CAAGATTCCCATTCTTGCGACCCGG GTTCGATTCCCGGGCGGCGCA
tE(TTC)B	Hs_Glu	2	130,811,002	130,810,931	Glu	TTC		TCCCATATGGTCTAGCGGTTAGGATT CCTGGTTTTCACCCAGGTGGCCCGG GTTCGACTCCCGGTATGGGAA
tA(CGC)B	Hs_Ala1	2	157,082,789	157,082,860	Ala	CGC		GGGGATGTAGCTCAGTGGTAGAGC GCGCGCTTCGCATGTGTGAGGTCCC GGGTTCAATCCCCGGCATCTCCA
tG(GCC)B	Hs_Gly1	2	157,083,237	157,083,167	Gly	GCC		GCATTGGTGGTTCAGTGGTAGAATT CTCGCCTGCCACGCGGGAGGCCCG GGTTCGATTCCCGGCCAATGCA
tY(ATA)B	Hs_Tyr	2	218,936,055	218,936,147	Tyr	ATA	Yes	CCTTCAATAGTTCAGCTGGTAGAGC AGAGGACTATAGCTACTTCCTCAGT AGGAGACGTCCTTAGGTTGCTGGTT CGATTCCAGCTTGAAGGA
tR(ACG)C	Hs_Arg2	3	45,705,567	45,705,495	Arg	ACG		GGGCCAGTGGCGCAATGGATAACG CGTCTGACTACGGATCAGAAGATTC TAGGTTCGACTCCTGGCTGGCTCG
tC(GCA)C1	Hs_Cys	3	133,430,713	133,430,642	Cys	GCA		GGGGGTGTAGCTCAGTGGTAGAGC ATTTGACTGCAGATCAAGAGGTCCC TGGTTCAAATCCAGGTGCCCCCT
tC(GCA)C2	Hs_Cys	3	133,433,411	133,433,340	Cys	GCA		GGGGGTATAGCTCAGGGGTAGAGCA TTTGACTGCAGATCAAGAGGTCCCT GGTTCAAATCCAGGTGCCCCCT

tV(AAC)C	Hs_Val	3	170,972,720	170,972,792	Val	AAC	GTTTCCGTAGTGTAGTGGTTATCAC GTTCGCCTAACACGCGAAAGGTCCC CGGTTCGAAACCGGGCGGAAACA
tD(GTC)C	Hs_Asp	3	185,848,867	185,848,797	Asp	GTC	TTCTTGTTAATATAGTGGTGAGTATT CCCACCTGTCATGCGGGAGACGGG GTTCAATTCCCTGATGGGGAG
tQ(TTG)D	Hs_Arg1	4	40,749,743	40,749,671	Gln	TTG	GACCATGTGGCCTAAGGGAAAAGA CATCTCACTTTGGGTCAGAAGATTG AGGGTTCAAGTCCTTTCATGGTCA
tC(GCA)D	Hs_Cys	4	124,787,681	124,787,610	Cys	GCA	GGGGGTATAGCTCAGTGGTAGAGCA TTTGACTGCAGATCAAGAGGTCCCC GGTTCAAATCCGGGTGCCCCCT
tL(TAA)D	Hs_Ser2	4	156,742,657	156,742,583	Leu	TAA	GTTAAGATGGCAGAGCCTGGTAATT GCATAAAACTTAAAATTTTATAATCA GAGGTTCAACTCCTCTTCTTAACA
tV(CAC)E1	Hs_Val	5	180,456,676	180,456,748	Val	CAC	GTTTCCGTAGTGTAGTGGTTATCAC GTTCGCCTCACACGCGAAAGGTCC CCGGTTCGAAACCGGGCGGAAACA
tL(AAG)E1	Hs_Leu1	5	180,457,161	180,457,080	Leu	AAG	GGTAGCGTGGCCGAGCGGTCTAAG GCGCTGGATTAAGGCTCCAGTCTCT TCGGAGGCGTGGGTTCGAATCCCAC CGCTGCCA
tL(AAG)E2	Hs_Leu1	5	180,461,446	180,461,527	Leu	AAG	GGTAGCGTGGCCGAGCGGTCTAAG GCGCTGGATTAAGGCTCCAGTCTCT TCGGAGGCGTGGGTTCGAATCCCAC CGCTGCCA
tV(CAC)E2	Hs_Val	5	180,461,931	180,461,859	Val	CAC	GTTTCCGTAGTGTAGTGGTTATCAC GTTCGCCTCACACGCGAAAGGTCC CCGGTTCGAAACCGGGCGGAAACA
tV(AAC)E1	Hs_Val	5	180,523,760	180,523,832	Val	AAC	GTTTCCGTAGTGTAGTGGTTATCAC GTTCGCCTAACACGCGAAAGGTCCC CGGTTCGAAACCGGGCGGAAACA
tV(AAC)E2	Hs_Val	5	180,529,216	180,529,288	Val	AAC	GTTTCCGTAGTGTAGTGGTTATCAC GTTCGCCTAACACGCGAAAGGTCCC CGGTTCGAAACCGGGCGGAAACA
tV(CAC)E3	Hs_Val	5	180,533,256	180,533,328	Val	CAC	GTTTCCGTAGTGTAGTGGTTATCAC GTTCGCCTCACACGCGAAAGGTCC CCGGTTCGAAACCGGGCGGAAACA
tL(AAG)E3	Hs_Leu1	5	180,533,731	180,533,650	Leu	AAG	GGTAGCGTGGCCGAGCGGTCTAAG GCGCTGGATTAAGGCTCCAGTCTCT TCGGAGGCGTGGGTTCGAATCCCAC CGCTGCCA
tL(AAG)E4	Hs_Leu1	5	180,547,307	180,547,388	Leu	AAG	GGTAGCGTGGCCGAGCGGTCTAAG GCGCTGGATTAAGGCTCCAGTCTCT TCGGGGGCGTGGGTTCGAATCCCAC CGCTGCCA
tV(AAC)E3	Hs_Val	5	180,548,094	180,548,022	Val	AAC	GTTTCCGTAGTGTAGTGGTCATCAC GTTCGCCTAACACGCGAAAGGTCCC CGGTTCGAAACCGGGCGGAAACA
tP(TGG)E	Hs_Pro	5	180,548,531	180,548,460	Pro	TGG	GGCTCGTTGGTCTAGGGGTATGATT CTCGCTTTGGGTGCGAGAGGTCCCG GGTTCAAATCCCGGACGAGCCC
tT(TGT)E	Hs_Thr2	5	180,551,364	180,551,293	Thr	TGT	GGCTCCATAGCTCAGGGGTTAGAGC ACTGGTCTTGTAAACCAGGGTCGCG AGTTCAAATCTCGCTGGGGCCT
tA(TGC)E	Hs_Ala1	5	180,566,474	180,566,545	Ala	TGC	GGGGATGTAGCTCAGTGGTAGAGC GCATGCTTTGCATGTATGAGGCCCC GGGTTCGATCCCCGGCATCTCCA
tK(CTT)E1	Hs_Lys1	5	180,567,361	180,567,433	Lys	CTT	GCCCGGCTAGCTCAGTCGGTAGAGC ATGAGACTCTTAATCTCAGGGTCGT GGGTTCGAGCCCCACGTTGGGCG

tV(AAC)E4	Hs_Val	5	180,577,948	180,577,876	Val	AAC	GTTTCCGTAGTGTAGTGGTTATCAC GTTCGCCTAACACGCGAAAGGTCCC CGGTTCGAAACCGGGCGGAAACA
tK(CTT)E2	Hs_Lys1	5	180,581,657	180,581,585	Lys	CTT	GCCCGGCTAGCTCAGTCGGTAGAGC ATGAGACTCTTAATCTCAGGGTCGT GGGTTCGAGCCCCACGTTGGGCG
tV(CAC)E4	Hs_Val	5	180,582,073	180,582,001	Val	CAC	GTTTCCGTAGTGTAGTGGTTATCAC GTTCGCCTCACACGCGAAAGGTCC CCGGTTCGAAACCGGGCGGAAACA
tQ(CTG)F1	Hs_Gln1	6	18,944,381	18,944,452	Gln	CTG	GGTTCCATGGTGTAATGGTTAGCAC TCTGGACTCTGAATCCAGCGATCCG AGTTCAAATCTCGGTGGAACCT
tM(CAT)F1	Hs_Met2	6	26,394,733	26,394,804	Met	CAT	AGCAGAGTGGCGCAGCGGAAGCGT GCTGGGCCCATAACCCAGAGGTCGA TGGATCGAAACCATCCTCTGCTA
tR(TCG)F1	Hs_Arg1	6	26,407,884	26,407,956	Arg	TCG	GACCACGTGGCCTAATGGATAAGGC GTCTGACTTCGGATCAGAAGATTGA GGGTTCGAATCCCTTCGTGGTTA
tS(GCT)F1	Hs_Ser2	6	26,413,780	26,413,697	Ser	GCT	GGAGAGGCCTGGCCGAGTGGTTAA GGCGATGGACTGCTAATCCATTGTG CTCTGCACGCGTGGGTTCGAATCCC ATCCTCGTCG
tQ(TTG)F1	Hs_Gln1	6	26,419,474	26,419,403	Gln	TTG	GGCCCCATGGTGTAATGGTTAGCAC TCTGGACTTTGAATCCAGCGATCCG AGTTCAAATCTCGGTGGGACCT
tQ(TTG)F2	Hs_Gln1	6	26,420,025	26,419,954	Gln	TTG	GGCCCCATGGTGTAATGGTTAGCAC TCTGGACTTTGAATCCAGCGATCCG AGTTCAAATCTCGGTGGGACCT
tS(TGA)F1	Hs_Ser1	6	26,420,884	26,420,803	Ser	TGA	GTAGTCGTGGCCGAGTGGTTAAGGC GATGGACTTGAAATCCATTGGGGTC TCCCCGCGCAGGTTCGAATCCTGCC GACTACG
tM(CAT)F2	Hs_Met2	6	26,421,402	26,421,331	Met	CAT	AGCAGAGTGGCGCAGCGGAAGCGT GCTGGGCCCATAACCCAGAGGTCGA TGGATCGAAACCATCCTCTGCTA
tW(CCA)F1	Hs_Trp	6	26,427,380	26,427,309	Trp	CCA	GACCTCGTGGCGCAACGGTAGCGC GTCTGACTCCAGATCAGAAGGTTGC GTGTTCAAATCACGTCGGGGTCA
tR(TCG)F2	Hs_Arg1	6	26,431,025	26,431,097	Arg	TCG	GACCACGTGGCCTAATGGATAAGGC GTCTGACTTCGGATCAGAAGATTGA GGGTTCGAATCCCTCCGTGGTTA
tS(AGA)F1	Hs_Ser1	6	26,435,796	26,435,877	Ser	AGA	GTAGTCGTGGCCGAGTGGTTAAGGC GATGGACTAGAAATCCATTGGGGTC TCCCCGCGCAGGTTCGAATCCTGCC GACTACG
tR(ACG)F1	Hs_Arg2	6	26,436,347	26,436,419	Arg	ACG	GGGCCAGTGGCGCAATGGATAACG CGTCTGACTACGGATCAGAAGATTC CAGGTTCGACTCCTGGCTGGCTCG
tM(CAT)F3	Hs_Met2	6	26,438,579	26,438,508	Met	CAT	AGCAGAGTGGCGCAGCGGAAGCGT GCTGGGCCCATAACCCAGAGGTCGA TGGATCGAAACCATCCTCTGCTA
tW(CCA)F2	Hs_Trp	6	26,439,722	26,439,651	Trp	CCA	GACCTCGTGGCGCAACGGTAGCGC GTCTGACTCCAGATCAGAAGGTTGC GTGTTCAAATCACGTCGGGGTCA
tL(CAG)F	Hs_Leu2	6	26,629,415	26,629,497	Leu	CAG	GTCAGGATGGCCGAGCGGTCTAAG GCGCTGCGTTCAGGTCGCAGTCTCC CCTGGAGGCGTGGGTTCGAATCCCA CTCCTGACA
tT(AGT)F1	Hs_Thr1	6	26,641,197	26,641,124	Thr	AGT	GGCTCCGTGGCTTAGCTGGTTAAAG CGCCTGTCTAGTAAACAGGAGATCC TGGGTTCGAATCCCAGCGGGGCCT

tR(ACG)F2	Hs_Arg2	6	26,645,705	26,645,777	Arg	ACG		GGGCCAGTGGCGCAATGGATAACG CGTCTGACTACGGATCAGAAGATTC CAGGTTCGACTCCTGGCTGGCTCG
tV(CAC)F1	Hs_Val	6	26,646,261	26,646,333	Val	CAC		GTTTCCGTAGTGTAGTGGTTATCAC GTTCGCCTCACACGCGAAAGGTCC CCGGTTCGAAACCGGGCGGAAACA
tA(CGC)F1	Hs_Ala1	6	26,661,710	26,661,781	Ala	CGC		GGGGATGTAGCTCAGTGGTAGAGC GCATGCTTCGCATGTATGAGGTCCC GGGTTCGATCCCCGGCATCTCCA
tI(AAT)F1	Hs_Ile1	6	26,662,329	26,662,402	Ile	AAT		GGCCGGTTAGCTCAGTTGGTTAGAG CGTGGTGCTAATAACGCCAAGGTCG CGGGTTCGATCCCCGTACGGGCCA
tP(AGG)F	Hs_Pro	6	26,663,477	26,663,548	Pro	AGG		GGCTCGTTGGTCTAGGGGTATGATT CTCGCTTAGGGTGCGAGAGGTCCCG GGTTCAAATCCCGGACGAGCCC
tK(CTT)F	Hs_Lys1	6	26,664,753	26,664,825	Lys	CTT		GCCCGGCTAGCTCAGTCGGTAGAGC ATGAGACTCTTAATCTCAGGGTCGT GGGTTCGAGCCCCACGTTGGGCG
tY(GTA)F1	Hs_Tyr	6	26,677,065	26,677,155	Tyr	GTA	Yes	CCTTCGATAGCTCAGTTGGTAGAGC GGAGGACTGTAGTTGGCTGTGTCCT TAGACATCCTTAGGTCGCTGGTTCG AATCCGGCTCGAAGGA
tA(AGC)F1	Hs_Ala2	6	26,680,143	26,680,071	Ala	AGC		GGGGAATTAGCTCAAATGGTAGAGC GCTCGCTTAGCATGCGAGAGGTAGC GGGATCGATGCCCGCATTCTCCA
tY(GTA)F2	Hs_Tyr	6	26,683,777	26,683,866	Tyr	GTA	Yes	CTTTCGATAGCTCAGTTGGTAGAGC GGAGGACTGTAGGTTCATTAAACTA AGGCATCCTTAGGTCGCTGGTTCGA ATCCGGCTCGAAGGA
tY(GTA)F3	Hs_Tyr	6	26,685,311	26,685,399	Tyr	GTA	Yes	CCTTCGATAGCTCAGTTGGTAGAGC GGAGGACTGTAGGCTCATTAAGCAA GGTATCCTTAGGTCGCTGGTTCGAA TCCGGCTCGGAGGA
tY(GTA)F4	Hs_Tyr	6	26,703,081	26,703,169	Tyr	GTA	Yes	CCTTCGATAGCTCAGCTGGTAGAGC GGAGGACTGTAGGGGTTTGAATGTG GTCATCCTTAGGTCGCTGGTTCGAA TCCGGCTCGGAGGA
tA(AGC)F2	Hs_Ala3	6	26,781,569	26,781,641	Ala	AGC		GGGGAATTAGCTCAAGTGGTAGAG CGCTTGCTTAGCATGCAAGAGGTAG TGGGATCAATGCCCACATTCTCCA
tA(AGC)F3	Hs_Ala3	6	26,790,694	26,790,766	Ala	AGC		GGGGAATTAGCTCAAGTGGTAGAG CGCTTGCTTAGCATGCAAGAGGTAG TGGGATCGATGCCCACATTCTCCA
tA(AGC)F4	Hs_Ala3	6	26,795,464	26,795,536	Ala	AGC		GGGGAATTAGCTCAAGTGGTAGAG CGCTTGCTTAGCACGCAAGAGGTAG TGGGATCGATGCCCACATTCTCCA
tM(CAT)F4	Hs_Met1	6	26,809,691	26,809,763	Met	CAT		GCCCTCTTAGCGCAGCTGGCAGCGC GTCAGTCTCATAATCTGAAGGTCCT GAGTTCAAGCCTCAGAGAGGGCA
tA(AGC)F5	Hs_Ala3	6	26,813,585	26,813,657	Ala	AGC		GGGGAATTAGCTCAAGCGGTAGAG CGCTTGCTTAGCATGCAAGAGGTAG TGGGATCGATGCCCACATTCTCCA
tI(AAT)F2	Hs_Ile1	6	26,829,273	26,829,200	Ile	AAT		GGCCGGTTAGCTCAGTTGGTCAGAG CGTGGTGCTAATAACGCCAAGGTCG CGGGTTCGATCCCCGTACGGGCCA
tA(AGC)F6	Hs_Ala2	6	26,836,235	26,836,307	Ala	AGC		GGGGAATTGGCTCAAGCGGTAGAG CGCTTGCTTAGCATGCAAGAGGTAG CAGGATCGACGCCTGCACTCTCCA

tA(AGC)F7	Hs_Ala2	6	26,838,716	26,838,788	Ala	AGC		GGGGAATTAGCTCAGGCGGTAGAG CGCTCGCTTAGCATGCGAGAGGTAG CGGGATCGACGCCCGCATTCTCCA
tM(CAT)F5	Hs_Met1	6	26,843,625	26,843,553	Met	CAT		GCCCTCTTAGCGCAGCGGGCAGCG CGTCAGTCTCATAATCTGAAGGTCC TGAGTTCGAGCCTCAGAGAGGGCA
tI(AAT)F3	Hs_Ile1	6	26,853,307	26,853,234	Ile	AAT		GGCCGGTTAGCTCAGTTGGTTAGAG CGTGGTGCTAATAACGCTAAGGTCG CGGGTTCGATCCCCGTACTGGCCA
tA(AGC)F8	Hs_Ala2	6	26,859,897	26,859,969	Ala	AGC		GGGGGATTAGCTCAAGCGGTAGGG TGCCTGCTTAGCATGCAAGAGGTAG CAGGATCGACGCCTGCATTCTCCA
tM(CAT)F6	Hs_Met1	6	26,866,601	26,866,529	Met	CAT		GCCCTCTTAGCGCAGCGGGCAGCG CGTCAGTCTCATAATCTGAAGGTCC TGAGTTCGAGCCTCAGAGAGGGCA
tM(CAT)F7	Hs_Met1	6	26,874,423	26,874,495	Met	CAT		GCCCTCTTAGCGCAGCGGGCAGCG CGTCAGTCTCATAATCTGAAGGTCC TGAGTTCGAGCCTCAGAGAGGGCA
tA(AGC)F9	Hs_Ala2	6	26,879,341	26,879,269	Ala	AGC		GGGGAATTAGCTCAGGCGGTAGAG CGCTCGCTTAGCATGCGAGAGGTAG CGGGATCGACGCCCGCATTCTCCA
tA(AGC)F10	Hs_Ala2	6	26,881,822	26,881,750	Ala	AGC		GGGGAATTGGCTCAAGCGGTAGAG CGCTTGCTTAGCATGCAAGAGGTAG CAGGATCGACGCCTGCACTCTCCA
tI(AAT)F4	Hs_Ile1	6	26,888,811	26,888,884	Ile	AAT		GGCCGGTTAGCTCAGTTGGTCAGAG CGTGGTGCTAATAACGCCAAGGTCG CGGGTTCGATCCCCGTACGGGCCA
tA(AGC)F11	Hs_Ala3	6	26,904,057	26,903,985	Ala	AGC		GGGGAATTAGCTCAAGTGGTAGAG CGCTTGCTTAGCATGCAAGAGGTAG TGGGATCGATGCCCACATTCTCCA
tI(TAT)F1	Hs_Ile2	6	27,096,104	27,096,197	Ile	TAT	Yes	GCTCCAGTGGCGCAATCGGTTAGCG CGCGGTACTTATATGGCAGTATGTGT GCGAGTGATGCCGAGGTTGTGAGTT CGAGCCTCACCTGGAGCA
tP(CGG)F	Hs_Pro	6	27,167,500	27,167,571	Pro	CGG		GGCTCGTTGGTCTAGGGGTATGATT CTCGCTTCGGGTGTGAGAGGTCCCG GGTTCAAATCCCGGACGAGCCC
tS(GCT)F2	Hs_Ser2	6	27,173,064	27,173,145	Ser	GCT		GACGAGGTGGCCGAGTGGTTAAGG CGATGGACTGCTAATCCATTGTGCT CTGCACGCGTGGGTTCGAATCCCAC CCTCGTCG
tV(CAC)F2	Hs_Val	6	27,226,001	27,226,073	Val	CAC		GTTTCTGTAGTATGGTGGTTATCACG TTAGTCTCACACGTGAAAGGTCCCT GGTTCGAAACCAGGTGGAAACA
tH(GTG)F	Hs_His	6	27,233,885	27,233,956	His	GTG		GCCGTGATCGTATAGTGGTTAGTACT CTGCGTTGTGGCCGCAGCAACCTCG GTTCGAATCCGAGTCACGGCA
tT(AGT)F2	Hs_Thr1	6	27,238,029	27,238,102	Thr	AGT		GGCCCTGTGGCTTAGCTGGTCAAAG CGCCTGTCTAGTAAACAGGAGATCC TGGGTTCGAATCCCAGCGGGGCCT
tI(AAT)F5	Hs_Ile1	6	27,253,046	27,252,973	Ile	AAT		GGCCGGTTAGCTCAGTTGGTTAGAG CGTGGTGCTAATAACGCCAAGGTCG CGGGTTCGATCCCCGTACGGGCCA
tV(CAC)F3	Hs_Val	6	27,281,918	27,281,846	Val	CAC		GTTTCCGTAGTGGAGTGGTTATCAC GTTCGCCTCACACGCGAAAGGTCC CCGGTTTGAAACCAGGCGGAAACA

tS(CGA)F1	Hs_Ser1	6	27,285,607	27,285,688	Ser	CGA	GCTGTGATGGCCGAGTGGTTAAGGC GTTGGACTCGAAATCCAATGGGGTC TCCCCGCGCAGGTTCAAATCCTGCT CACAGCG
tR(ACG)F3	Hs_Arg2	6	27,289,674	27,289,602	Arg	ACG	GGGCCAGTGGCGCAATGGATAACG CGTCTGACTACGGATCAGAAGATTC TAGGTTCGACTCCTGGCTGGCTCG
tR(ACG)F4	Hs_Arg2	6	27,290,931	27,291,003	Arg	ACG	GGGCCAGTGGCGCAATGGATAACG CGTCTGACTACGGATCAGAAGATTC TAGGTTCGACTCCTGGCTGGCTCG
tL(TAA)F1	Hs_Leu3	6	27,306,395	27,306,313	Leu	TAA	ACCGGGATGGCTGAGTGGTTAAGG CGTTGGACTTAAGATCCAATGGACA GGTGTCCGCGTGGGTTCGAGCCCC ACTCCCGGTA
tV(AAC)F1	Hs_Val	6	27,311,267	27,311,339	Val	AAC	GTTTCCGTAGTGTAGTGGTTATCAC GTTTGCCTAACACGCGAAAGGTCCC CGGTTCGAAACCGGGCAGAAACA
tI(AAT)F6	Hs_Ile1	6	27,313,402	27,313,329	Ile	AAT	GGCCGGTTAGCTCAGTTGGTTAGAG CGTGGTGCTAATAACGCCAAGGTCG CGGGTTCGATCCCCGTACGGGCCA
tI(AAT)F7	Hs_Ile1	6	27,349,718	27,349,791	Ile	AAT	GGCTGGTTAGTTCAGTTGGTTAGAG CGTGGTGCTAATAACGCCAAGGTCG TGGGTTCGATCCCCATATCGGCCA
tI(AAT)F8	Hs_Ile1	6	27,351,042	27,350,969	Ile	AAT	GGCTGGTTAGCTCAGTTGGTTAGAG CGTGGTGCTAATAACGCCAAGGTCG CGGGTTCGATCCCCGTACTGGCCA
tV(CAC)F4	Hs_Val	6	27,356,100	27,356,028	Val	CAC	GCTTCTGTAGTGTAGTGGTTATCAC GTTCGCCTCACACGCGAAAGGTCC CCGGTTCGAAACCGGGCAGAAGCA
tV(TAC)F	Hs_Val	6	27,366,384	27,366,456	Val	TAC	GTTTCCGTGGTGTAGTGGTTATCAC ATTCGCCTTACACGCGAAAGGTCCT CGGGTCGAAACCGAGCGGAAACA
tQ(CTG)F2	Hs_Gln1	6	27,371,191	27,371,262	Gln	CTG	GGTTCCATGGTGTAATGGTTAGCAC TCTGGACTCTGAATCCGGTAATCCG AGTTCAAATCTCGGTGGAACCT
tS(GCT)F3	Hs_Ser2	6	27,373,754	27,373,835	Ser	GCT	GACGAGGTGGCCGAGTGGTTAAGG CGATGGACTGCTAATCCATTGTGCT CTGCACGCGTGGGTTCGAATCCCAC CTTCGTCG
tM(CAT)F8	Hs_Met2	6	27,408,814	27,408,743	Met	CAT	AGCAGAGTGGCGCAGCGGAAGCGT GCTGGGCCCATAACCCAGAGGTCGA TGGATCGAAACCATCCTCTGCTA
tK(TTT)F1	Hs_Lys2	6	27,410,820	27,410,748	Lys	TTT	GCCTGGGTAGCTCAGTCGGTAGAGC ATCAGACTTTTAATCTGAGGGTCCA GGGTTCAAGTCCCTGTCCAGGCG
tS(AGA)F2	Hs_Ser1	6	27,554,570	27,554,651	Ser	AGA	GTAGTCGTGGCCGAGTGGTTAAGGC GATGGACTAGAAATCCATTGGGGTC TCCCCGCGCAGGTTCGAATCCTGCC GACTACG
tD(GTC)F1	Hs_Asp	6	27,555,432	27,555,503	Asp	GTC	TCCTCGTTAGTATAGTGGTGAGTATC CCCGCCTGTCACGCGGGAGACCGG GGTTCGATTCCCCGACGGGGAG
tS(AGA)F3	Hs_Ser1	6	27,571,572	27,571,653	Ser	AGA	GTAGTCGTGGCCGAGTGGTTAAGGC GATGGACTAGAAATCCATTGGGGTC TCCCCGCGCAGGTTCGAATCCTGCC GACTACG
tS(AGA)F4	Hs_Ser1	6	27,578,797	27,578,878	Ser	AGA	GTAGTCGTGGCCGAGTGGTTAAGGC GATGGACTAGAAATCCATTGGGGTC TCCCCGCGCAGGTTCGAATCCTGCC GACTACG

tD(GTC)F2	Hs_Asp	6	27,579,502	27,579,573	Asp	GTC		TCCTCGTTAGTATAGTGGTGAGTATC CCCGCCTGTCACGCGGGAGACCGG GGTTCGATTCCCCGACGGGGAG
tS(TGA)F2	Hs_Ser1	6	27,581,667	27,581,586	Ser	TGA		GTAGTCGTGGCCGAGTGGTTAAGGC GATGGACTTGAAATCCATTGGGGTT TCCCCGCGCAGGTTCGAATCCTGTC GGCTACG
tQ(CTG)F3	Hs_Gln1	6	27,595,287	27,595,358	Gln	CTG		GGTTCCATGGTGTAATGGTTAGCAC TCTGGACTCTGAATCCAGCGATCCG AGTTCAAATCTCGGTGGAACCT
tS(AGA)F5	Hs_Ser1	6	27,607,966	27,608,047	Ser	AGA		GTAGTCGTGGCCGAGTGGTTAAGGC GATGGACTAGAAATCCATTGGGGTT TCCCCACGCAGGTTCGAATCCTGCC GACTACG
tS(AGA)F6	Hs_Ser1	6	27,617,614	27,617,533	Ser	AGA		GTAGTCGTGGCCGAGTGGTTAAGGC GATGGACTAGAAATCCATTGGGGTT TCCCCGCGCAGGTTCGAATCCTGCC GACTACG
tS(TGA)F3	Hs_Ser1	6	27,621,447	27,621,528	Ser	TGA		GTAGTCGTGGCCGAGTGGTTAAGGC GATGGACTTGAAATCCATTGGGGTT TCCCCGCGCAGGTTCGAATCCTGCC GACTACG
tQ(CTG)F4	Hs_Gln1	6	27,623,581	27,623,510	Gln	CTG		GGTTCCATGGTGTAATGGTTAGCAC TCTGGACTCTGAATCCAGCGATCCG AGTTCAAGTCTCGGTGGAACCT
tS(AGA)F7	Hs_Ser1	6	27,629,252	27,629,171	Ser	AGA		GTAGTCGTGGCCGAGTGGTTAAGGT GATGGACTAGAAACCCATTGGGGTC TCCCCGCGCAGGTTCGAATCCTGCC GACTACG
tR(TCT)F	Hs_Arg4	6	27,637,942	27,638,028	Arg	TCT	Yes	GGCTCTGTGGCGCAATGGATAGCGC ATTGGACTTCTAGCCTAAATCAAGA GATTCAAAGGTTGCGGGTTCGAGTC CCTCCAGAGTCG
tK(TTT)F2	Hs_Lys2	6	27,651,825	27,651,897	Lys	TTT		ACCTGGGTAGCTCAGTAGGTAGAAC ATCAGACTTTTAATCTGAGGGTCTA GGGTTCAAGTCCCTGTCCAGGCG
tD(GTC)F3	Hs_Asp	6	27,659,286	27,659,215	Asp	GTC		TCCTCGTTAGTATAGTGGTGAGTGT CCCCGTCTGTCACGCGGGAGACCG GGGTTCGATTCCCCGACGGGGAG
tK(TTT)F3	Hs_Lys2	6	27,667,644	27,667,572	Lys	TTT		GCCTGGATAGCTCAGTCGGTAGAGC ATCAGACTTTTAATCTGAGGGTCCA GGGTTCAAGTCCCTGTTCAGGCG
tM(CAT)F9	Hs_Met2	6	27,668,650	27,668,579	Met	CAT		AGCAGAGTGGCGCAGCGGAAGCGT GCTGGGCCCATAACCCAGAGGTCGA TGGATCGAAACCATCCTCTGCTA
tL(CAA)F1	Hs_Leu2	6	27,678,433	27,678,327	Leu	CAA	Yes	GTCAGGATGGCCGAGTGGTCTAAG GCGCCAGACTCAAGTTGCTACTTCC CAGGTTTGGGGCTTCTGGTCTCCGC ATGGAGGCGTGGGTTCGAATCCCAC TTCTGACA
tL(CAA)F2	Hs_Leu2	6	27,681,503	27,681,396	Leu	CAA	Yes	GTCAGGATGGCCGAGTGGTCTAAG GCGCCAGACTCAAGCTTACTGCTTC CTGTGTTCGGGTCTTCTGGTCTCCG TATGGAGGCGTGGGTTCGAATCCCA CTTCTGACA
tT(CGT)F1	Hs_Thr2	6	27,694,114	27,694,187	Thr	CGT		GGCCCTGTAGCTCAGCGGTTGGAGC GCTGGTCTCGTAAACCTAGGGGTCG TGAGTTCAAATCTCACCAGGGCCT
tI(TAT)F2	Hs_Ile2	6	27,707,179	27,707,272	Ile	TAT	Yes	GCTCCAGTGGCGCAATCGGTTAGCG CGCGGTACTTATACAACAGTATATGT GCGGGTGATGCCGAGGTTGTGAGTT CGAGCCTCACCTGGAGCA
tV(AAC)F2	Hs_Val	6	27,726,758	27,726,686	Val	AAC		GTTTCCGTAGTGTAGTGGTTATCAC GTTCGCCTAACACGCGAAAGGTCCC TGGATCAAAACCAGGCGGAAACA

tI(AAT)F9	Hs_Ile1	6	27,744,341	27,744,414	Ile	AAT	GGCCGGTTAGCTCAGTCGGCTAGAG CGTGGTGCTAATAACGCCAAGGTCG CGGGTTCGATCCCCGTACGGGCCA
tR(ACG)F5	Hs_Arg2	6	27,746,395	27,746,323	Arg	ACG	GGGCCAGTGGCGCAATGGATAACG CGTCTGACTACGGATCAGAAGATTC TAGGTTCGACTCCTGGCTGGCTCG
tS(CGA)F2	Hs_Ser1	6	27,748,289	27,748,208	Ser	CGA	GCTGTGATGGCCGAGTGGTTAAGGT GTTGGACTCGAAATCCAATGGGGGT TCCCCGCGCAGGTTCAAATCCTGCT CACAGCG
tV(AAC)F3	Hs_Val	6	27,756,936	27,756,864	Val	AAC	GTTTCCGTAGTGTAGTGGTTATCAC GTTCGCCTAACACGCGAAAGGTCC GCGGTTCGAAACCGGGCGGAAACA
tT(AGT)F3	Hs_Thr1	6	27,760,526	27,760,453	Thr	AGT	GGCTCCGTGGCTTAGCTGGTTAAAG CGCCTGTCTAGTAAACAGGAGATCC TGGGTTCGAATCCCAGCGGGGCCT
tI(AAT)F10	Hs_Ile1	6	27,763,946	27,764,019	Ile	AAT	GGCCGGTTAGCTCAGTTGGTTAGAG CGTGGTGCTAATAACGCCAAGGTCG CGGGTTCGATCCCCGTACTGGCCA
tL(TAA)F2	Hs_Leu3	6	27,796,959	27,796,877	Leu	TAA	ACCGGGATGGCCGAGTGGTTAAGG CGTTGGACTTAAGATCCAATGGGCT GGTGCCCGCGTGGGTTCGAACCCC ACTCTCGGTA
tT(AGT)F4	Hs_Thr1	6	27,802,452	27,802,525	Thr	AGT	GGCTTCGTGGCTTAGCTGGTTAAAG CGCCTGTCTAGTAAACAGGAGATCC TGGGTTCGAATCCCAGCGAGGCCT
tV(CAC)F5	Hs_Val	6	27,804,378	27,804,306	Val	CAC	GTTTCCGTAGTGTAGTGGTTATTATG TTCGCCTCACACGCGAAAAGTCCCC GGTTCGAAATCAGGCGGGAACA
tV(AAC)F4	Hs_Val	6	27,829,230	27,829,158	Val	AAC	GTTTCCGTAGTGTAGTGGTTATCAC GTTCGCCTAACACGCGAAAGGTCCC CGGTTCGAAACCGGGCGGAAACA
tM(CAT)F10	Hs_Met2	6	27,853,643	27,853,714	Met	CAT	AGCAGAGTGGCGCAGCGGAAGCGT GCTGGGCCCATAACCCAGAGGTCGA TGGATCTAAACCATCCTCTGCTA
tQ(CTG)F5	Hs_Gln1	6	27,867,185	27,867,114	Gln	CTG	GGCCCCATGGTGTAATGGTCAGCAC TCTGGACTCTGAATCCAGCGATCCG AGTTCAAATCTCGGTGGGACCC
tQ(TTG)F3	Hs_Gln1	6	27,871,690	27,871,619	Gln	TTG	GGCCCCATGGTGTAATGGTTAGCAC TCTGGACTTTGAATCCAGCGATCCG AGTTCAAATCTCGGTGGGACCT
tM(CAT)F11	Hs_Met2	6	27,978,321	27,978,250	Met	CAT	AGCAGAGTGGCGCAGCGGAAGCGT GCTGGGCCCATAACCCAGAGGTCGA TGGATCGAAACCATCCTCTGCTA
tG(GCC)F	Hs_Gly1	6	27,978,735	27,978,665	Gly	GCC	GCATTGGTGGTTCAGTGGTAGAATT CTCGCCTGCCACGCGGGAGGCCCG GGTTCGATTCCCGGCCAATGCA
tS(GCT)F4	Hs_Ser2	6	28,288,794	28,288,875	Ser	GCT	GACGAGGTGGCCGAGTGGTTAAGG CGATGGACTGCTAATCCATTGTGCT CTGCACACGTGGGTTCGAATCCCAT CCTCGTCG
tT(TGT)F	Hs_Thr1	6	28,550,381	28,550,308	Thr	TGT	GGCTCTATGGCTTAGTTGGTTAAAG CGCCTGTCTTGTAAACAGGAGATCC TGGGTTCGAATCCCAGTAGAGCCT
tL(AAG)F1	Hs_Leu1	6	28,554,460	28,554,379	Leu	AAG	GGTAGCGTGGCCGAGTGGTCTAAG ACGCTGGATTAAGGCTCCAGTCTCT TCGGGGGCGTGGGTTTGAATCCCAC CGCTGCCA
tT(CGT)F2	Hs_Thr1	6	28,564,822	28,564,749	Thr	CGT	GGCTCTATGGCTTAGTTGGTTAAAG CGCCTGTCTCGTAAACAGGAGATCC TGGGTTCGACTCCCAGTGGGGCCT

tI(TAT)F3	Hs_Ile2	6	28,613,346	28,613,439	Ile	TAT	Yes	GCTCCAGTGGCGCAATCGGTTAGCG CGCGGTACTTATAAGACAGTGCACC TGTGAGCAATGCCGAGGTTGTGAGT TCAAGCCTCACCTGGAGCA
tR(TCG)F3	Hs_Arg1	6	28,618,942	28,618,870	Arg	TCG		GACCACGTGGCCTAATGGATAAGGC GTCTGACTTCGGATCAGAAGATTGA GGGTTCGAATCCCTTCGTGGTTG
tQ(TTG)F4	Hs_Gln1	6	28,665,135	28,665,206	Gln	TTG		GGTCCCATGGTGTAATGGTTAGCAC TCTGGACTTTGAATCCAGCAATCCG AGTTCGAATCTCGGTGGGACCT
tS(GCT)F5	Hs_Ser2	6	28,673,177	28,673,096	Ser	GCT		GACGAGGTGGCCGAGTGGTTAAGG CGATGGACTGCTAATCCATTGTGCT CTGCACGCGTGGGTTCGAATCCCAT CCTCGTCG
tA(AGC)F12	Hs_Ala1	6	28,682,912	28,682,983	Ala	AGC		GGGGGTGTAGCTCAGTGGTAGAGC GCGTGCTTAGCATGTACGAGGTCCC GGGTTCAATCCCCGGCACCTCCA
tA(TGC)F1	Hs_Ala1	6	28,719,201	28,719,272	Ala	TGC		GGGGATGTAGCTCAGTGGTAGAGC GCATGCTTTGCATGTATGAGGTCCC GGGTTCGATCCCCGGCATCTCCA
tT(CGT)F3	Hs_Thr1	6	28,724,036	28,723,963	Thr	CGT		GGCTCTGTGGCTTAGTTGGCTAAAG CGCCTGTCTCGTAAACAGGAGATCC TGGGTTCGAATCCCAGCGGGGCCT
tA(AGC)F13	Hs_Ala3	6	28,734,064	28,733,993	Ala	AGC		GGGGATGTAGCTCAGTGGTAGAGC GCATGCTTAGCATGCATGAGGTCCC GGGTTCGATCCCCAGCATCTCCA
tA(CGC)F2	Hs_Ala1	6	28,749,663	28,749,592	Ala	CGC		GGGGATGTAGCTCAGTGGTAGAGC GCATGCTTCGCATGTATGAGGCCCC GGGTTCGATCCCCGGCATCTCCA
tA(CGC)F3	Hs_Ala1	6	28,771,759	28,771,688	Ala	CGC		GGGGGTGTAGATCAGTGGTAGAGC GCATGCTTCGCATGTACGAGGTCCC TGGTTCAATCCCTGGTACCTCCA
tA(AGC)F14	Hs_Ala1	6	28,786,345	28,786,416	Ala	AGC		GGGGGTGTAGCTCAGTGGTAGAGC GCGTGCTTAGCATGCACGAGGCCCT GGGTTCAATCCCCAGCACCTCCA
tA(AGC)F15	Hs_Ala1	6	28,795,460	28,795,531	Ala	AGC		GGGGGTGTAGCTCAGTGGTAGAGC GCGTGCTTAGCATGCACGAGGCCCC GGGTTCAATCCCTGGCACCTCCA
tT(AGT)F5	Hs_Thr1	6	28,801,774	28,801,847	Thr	AGT		GGCTCCGTAGCTTAGTTGGTTAAAG CGCCTGTCTAGTAAACAGGAGATCC TGGGTTCGACTCCCAGCGGGGCCT
tA(CGC)F	Hs_Ala1	6	28,805,071	28,805,142	Ala	CGC		GGGGGTGTAGCTCAGTGGTAGAGC GCGTGCTTCGCATGTACGAGGCCCC GGGTTCGACCCCCGGCTCCTCCA
tV(AAC)F5	Hs_Ala1	6	28,811,256	28,811,185	Val	AAC		GGGGGTGTAGCTCAGTGGTAGAGC GTATGCTTAACATTCATGAGGCTCTG GGTTCGATCCCCAGCACTTCCA
tR(CCG)F1	Hs_Arg1	6	28,818,780	28,818,708	Arg	CCG		GGCCGCGTGGCCTAATGGATAAGGC GTCTGATTCCGGATCAGAAGATTGA GGGTTCGAGTCCCTTCGTGGTCG
tK(TTT)F4	Hs_Lys2	6	28,823,500	28,823,572	Lys	TTT		GCCTGGATAGCTCAGTTGGTAGAAC ATCAGACTTTTAATCTGACGGTGCA GGGTTCAAGTCCCTGTTCAGGCG
tA(TGC)F2	Hs_Ala1	6	28,834,191	28,834,120	Ala	TGC		GGGGGTGTAGCTCAGTGGTAGAGC ACATGCTTTGCATGTGTGAGGCCCC GGGTTCGATCCCCGGCACCTCCA
tF(GAA)F1	Hs_Phe	6	28,839,426	28,839,353	Phe	GAA		GCTGAAATAGCTCAGTTGGGAGAG CGTTAGACTGAAGATCTTAAAGTTC CCTGGTTCAACCCTGGGTTTCAGCC
tF(GAA)F2	Hs_Phe	6	28,840,143	28,840,215	Phe	GAA		GCCAAAATTGCTCAGTTGGGAGAG CGTTAGACTGAAGATCTAAAGGTCC CTGGTTCGATCCCGGGTTTCACCA

tA(TGC)F3	Hs_Ala1	6	28,865,597	28,865,526	Ala	TGC		GGGGGTGTAGCTCAGTGGTAGAGC GCATGCTTTGCATGTATGAGGTCCC GGGTTCGATCCCCGGCACCTCCA
tF(GAA)F3	Hs_Phe	6	28,866,550	28,866,478	Phe	GAA		GCCGAAATAGCTCAGTTGGGAGAG CGTTAGACTGAAGATCTAAAGGTCC CTGGTTCGATCCCGGGTTTCGGCA
tA(AGC)F16	Hs_Ala1	6	28,871,791	28,871,720	Ala	AGC		GGGGGTATAGCTCAGTGGTAGAGCG CGTGCTTAGCATGCACGAGGTCCTG GGTTCGATCCCCAGTACCTCCA
tA(TGC)F4	Hs_Ala1	6	28,878,626	28,878,556	Ala	TGC		GGGGGTGTAGCTCAGTGGTAGAGC GCATGCTTTGCATGTATGAGGCCTC GGTTCGATCCCCGACACCTCCA
tF(GAA)F4	Hs_Phe	6	28,883,661	28,883,589	Phe	GAA		GCCGAGATAGCTCAGTTGGGAGAG CGTTAGACTGAAGATCTAAAGGTCC CTGGTTCAATCCCGGGTTTCGGCA
tA(AGC)F17	Hs_Ala1	6	28,887,899	28,887,828	Ala	AGC		GGGGGTATAGCTCAGCGGTAGAGCG CGTGCTTAGCATGCACGAGGTCCTG GGTTCAATCCCCAATACCTCCA
tA(TGC)F5	Hs_Ala1	6	28,893,062	28,892,991	Ala	TGC		GGGGGTGTAGCTCAGTGGTAGAGC GCATGCTTTGCATGTATGAGGCCTC GGGTTCGATCCCCGACACCTCCA
tF(GAA)F5	Hs_Phe	6	28,899,145	28,899,072	Phe	GAA		GCCGAAATAGCTCAGTTGGGAGAG CGTTAGACCGAAGATCTTAAAGGTC CCTGGTTCAATCCCGGGTTTCGGCA
tA(AGC)F18	Hs_Ala1	6	28,914,271	28,914,200	Ala	AGC		GGGGGTGTAGCTCAGTGGTAGAGC GCGTGCTTAGCATGCACGAGGCCCC GGGTTCAATCCCCGGCACCTCCA
tA(AGC)F19	Hs_Ala1	6	28,939,512	28,939,441	Ala	AGC		GGGGGTGTAGCTCAGTGGTAGAGC GCGTGCTTAGCATGCACGAGGCCCC GGGTTCAATCCCCGGCACCTCCA
tR(CCG)F2	Hs_Arg1	6	28,957,144	28,957,216	Arg	CCG		GGCCGCGTGGCCTAATGGATAAGGC GTCTGATTCCGGATCAGAAGATTGA GGGTTCGAGTCCCTTCGTGGTCG
tL(CAA)F3	Hs_Leu2	6	28,972,084	28,971,979	Leu	CAA	Yes	GTCAGGATGGCCGAGTGGTCTAAG GCGCCAGACTCAAGCTAAGCTTCCT CCGCGGTGGGGATTCTGGTCTCCAA TGGAGGCGTGGGTTCGAATCCCACT TCTGACA
tL(CAA)F4	Hs_Leu2	6	29,016,809	29,016,913	Leu	CAA	Yes	GTCAGGATGGCCGAGTGGTCTAAG GCGCCAGACTCAAGCTTGGCTTCCT CGTGTTGAGGATTCTGGTCTCCAAT GGAGGCGTGGGTTCGAATCCCACTT CTGACA
tQ(CTG)F6	Hs_Gln1	6	29,017,428	29,017,357	Gln	CTG		GGTTCCATGGTGTAATGGTTAGCAC TCTGGACTCTGAATCCAGCGATCCG AGTTCAAATCTCGGTGGAACCT
tL(AAG)F2	Hs_Leu1	6	29,019,459	29,019,378	Leu	AAG		GGTAGCGTGGCCGAGCGGTCTAAG GCGCTGGATTAAGGCTCCAGTCTCT TCGGGGGCGTGGGTTCGAATCCCAC CGCTGCCA
tM(CAT)F12	Hs_Met1	6	29,020,331	29,020,403	Met	CAT		GCCTCCTTAGCGCAGTAGGCAGCGC GTCAGTCTCATAATCTGAAGGTCCT GAGTTCGAACCTCAGAGGGGGCA
tK(TTT)F5	Hs_Lys2	6	29,026,785	29,026,857	Lys	TTT		GCCCGGATAGCTCAGTCGGTAGAGC ATCAGACTTTTAATCTGAGGGTCCA GGGTTCAAGTCCCTGTTCGGGCG
tM(CAT)F13	Hs_Met1	6	29,029,093	29,029,021	Met	CAT		GCCTCCTTAGCGCAGTAGGCAGCGC GTCAGTCTCATAATCTGAAGGTCCT GAGTTCGAACCTCAGAGGGGGCA
tF(GAA)F6	Hs_Phe	6	29,057,500	29,057,428	Phe	GAA		GCCGAAATAGCTCAGTTGGGAGAG CGTTAGACTGAAGATCTAAAGGTCC CTGGTTCGATCCCGGGTTTCGGCA

tE(CTC)F1	Hs_Glu	6	29,057,955	29,058,026	Glu	CTC	TCCCTGGTGGTCTAGTGGTTAGGAT TCGGCGCTCTCACCGCCGCGGCCCG GGTTCGATTCCCGGTCAGGGAA
tL(AAG)F3	Hs_Leu1	6	29,064,758	29,064,839	Leu	AAG	GGTAGCGTGGCCGAGCGGTCTAAG GCGCTGGATTAAGGCTCCAGTCTCT TCGGGGGCGTGGGTTCAAATCCCAC CGCTGCCA
tA(AGC)F20	Hs_Ala2	6	58,249,908	58,249,836	Ala	AGC	GGGGAATTAGCTCAAGCGGTAGAG CGCTCCCTTAGCATGCGAGAGGTAG CGGGATCGACGCCCCCATTCTCTA
tA(AGC)F21	Hs_Ala2	6	58,250,620	58,250,548	Ala	AGC	GGGGGATTAGCTCAAGCGGTAGAG CGCCTGCTTAGCATGCAAGAGGTAG CAGGATCGATGCCTGCATTCTCCA
tI(AAT)F11	Hs_Ile1	6	58,257,213	58,257,286	Ile	AAT	GGCCGGTTAGCTCAGTTGGTTAGAG CGTGGCGCTAATAACGCCAAGGTCG CGGGTTCGATCCCCGTACGGGCCA
tA(AGC)F22	Hs_Ala3	6	58,272,659	58,272,587	Ala	AGC	GGGGAATTAGCTCAAGCGGTAGAG CGCTTGCTTAGCATGCAAGAGGTAG TGGGATCGATGCCCACATTCTCCA
tM(CAT)F14	Hs_Met1	6	58,276,523	58,276,451	Met	CAT	GCCCTCTTAGTGCAGCTGGCAGCGC GTCAGTTTCATAATCTGAAAGTCCT GAGTTCAAGCCTCAGAGAGGGCA
tA(AGC)F23	Hs_Ala3	6	58,290,710	58,290,638	Ala	AGC	GGGGAATTAGCTCAAGTGGTAGAG CGCTTGCTTAGCATGCAAGAGGTAG TGGGATCGATGCCCACATTCTCCA
tA(AGC)F24	Hs_Ala3	6	58,295,475	58,295,403	Ala	AGC	GGGGAATTAGCGCAAGTGGTAGAG TGCTTGCTTAGCATGCAAGAGGTAG TGGGATCGATGCCCACATTCTCCA
tA(AGC)F25	Hs_Ala3	6	58,304,654	58,304,582	Ala	AGC	GGGGAATTAGCCCAAGTGGTAGAG CGCTTGCTTAGCATGCAAGAGGTAG TGGGATCGATGCCCACATTCTCCA
tE(CTC)F2	Hs_Glu	6	126,143,157	126,143,086	Glu	CTC	TCCCTGGTGGTCTAGTGGTTAGGAT TCGGCGCTCTCACCGCCGCGGCCCG GGTTCGATTCCCGGTCAGGGAA
tL(TAA)F3	Hs_Leu3	6	144,579,377	144,579,459	Leu	TAA	ACCAGGATGGCCGAGTGGTTAAGG CGTTGGACTTAAGATCCAATGGACA TATGTCCGCGTGGGTTCGAACCCCA CTCCTGGTA
tQ(TTG)F5	Hs_Gln1	6	145,545,552	145,545,623	Gln	TTG	GGTCCCATGGTGTAATGGTTAGCAC TCTGGGCTTTGAATCCAGCAATCCG AGTTCGAATCTTGGTGGGACCT
tW(CCA)G	Hs_Trp	7	98,711,958	98,712,029	Trp	CCA	GACCTCGTGGCGCAACGGCAGCGC GTCTGACTCCAGATCAGAAGGTTGC GTGTTCAAATCACGTCGGGGTCA
tP(AGG)G	Hs_Pro	7	128,017,455	128,017,526	Pro	AGG	GGCTCGTTGGTCTAGGGGTATGATT CTCGCTTAGGGTGCGAGAGGTCCCG GGTTCAAATCCCGGACGAGCCC
tR(CCT)G	Hs_Arg3	7	138,482,701	138,482,773	Arg	CCT	GCCCCAGTGGCCTAATGGATAAGGC ATTGGCCTCCTAAGCCAGGGATTGT GGGTTCGAGTCCCATCTGGGGTG
tC(GCA)G1	Hs_Cys	7	148,444,929	148,445,000	Cys	GCA	GGGGGCATAGCTCAGTGGTAGAGCA TTTGACTGCAGATCAAGAGGTCCCT GGTTCAAATCCAGGTGCCCCCT
tC(GCA)G2	Hs_Cys	7	148,465,868	148,465,939	Cys	GCA	GGGGGTATAGCTCAGGGGTAGAGCA TTTGACTGCAGATCAAGAGGTCCCT GGTTCAAATCCAGGTGCCCCCT
tC(GCA)G3	Hs_Cys	7	148,490,485	148,490,414	Cys	GCA	GGGGGTATAGCTCAGGGGTAGAGCA TTTGACTGCAGATCAAGAGGTCCCC AGTTCAAATCTGGGTGCCCCCT
tC(GCA)G4	Hs_Cys	7	148,510,569	148,510,498	Cys	GCA	GGGGGTATAGTTCAGGGGTAGAGCA TTTGACTGCAGATCAAGAGGTCCCT GGTTCAAATCCAGGTGCCCCCT

tC(GCA)G5	Hs_Cys	7	148,512,320	148,512,249	Cys	GCA		GGGGGTATAGCTCAGGGGTAGAGCA TTTGACTGCAGATCAAGAGGTCCCT GGTTCAAATCCAGGTGCCCCCC
tC(GCA)G6	Hs_Cys	7	148,549,948	148,549,877	Cys	GCA		GGGGGTATAGCTTAGCGGTAGAGCA TTTGACTGCAGATCAAGAGGTCCCC GGTTCAAATCCGGGTGCCCCCT
tC(GCA)G7	Hs_Cys	7	148,681,279	148,681,350	Cys	GCA		GGGGGTATAGCTCAGGGGTAGAGC ACTTGACTGCAGATCAAGAAGTCCT TGGTTCAAATCCAGGTGCCCCCT
tC(GCA)G8	Hs_Cys	7	148,691,450	148,691,521	Cys	GCA		GGGCGTATAGCTCAGGGGTAGAGCA TTTGACTGCAGATCAAGAGGTCCCC AGTTCAAATCTGGGTGCCCACT
tC(GCA)G9	Hs_Cys	7	148,719,464	148,719,535	Cys	GCA		GGGGGTATAGCTCAGGGGTAGAGCA TTTGACTGCAGATCAAGAGGTCTCT GGTTCAAATCCAGGTGCCCCCT
tC(GCA)G10	Hs_Cys	7	148,723,883	148,723,812	Cys	GCA		GGGGGTATAGCTCAGGGGTAGAGC ACTTGACTGCAGATCAAGAGGTCCC TGGTTCAAATCCAGGTGCCCCCT
tC(GCA)G11	Hs_Cys	7	148,730,024	148,729,953	Cys	GCA		GGGGGTATAGCTCACAGGTAGAGCA TTTGACTGCAGATCAAGAGGTCCCC GGTTCAAATCCGGTTACTCCCT
tC(GCA)G12	Hs_Cys	7	148,732,694	148,732,765	Cys	GCA		GGGCGTATAGCTCAGGGGTAGAGCA TTTGACTGCAGATCAAGAGGTCCCC AGTTCAAATCTGGGTGCCCCCT
tS(AGA)G	Hs_Cys	7	148,743,115	148,743,186	Ser	AGA		GGGTGTATGGCTCAGGGGTAGAGAA TTTGACTAGAGATCAAGAGGTCCCT GGTTCAAATCCAGGTGCCCCCT
tC(GCA)G13	Hs_Cys	7	148,747,875	148,747,804	Cys	GCA		GGGGGTATAGCTCAGGGGTAGAGCA TTTGACTGCAAATCAAGAGGTCCCT GATTCAAATCCAGGTGCCCCCT
tC(GCA)G14	Hs_Cys	7	148,770,426	148,770,497	Cys	GCA		GGGGGTATAGCTCAGGGGTAGAGCA TTTGACTGCAGATCAAGAGGTCCCT GGTTCAAATCCAGGTGCCCCCT
tC(GCA)G15	Hs_Cys	7	148,781,765	148,781,694	Cys	GCA		GGGGGTATAGCTTAGGGGTAGAGCA TTTGACTGCAGATCAAAAGGTCCCT GGTTCAAATCCAGGTGCCCCTT
tC(GCA)G16	Hs_Cys	7	148,799,563	148,799,634	Cys	GCA		GGGGGTATAGCTCACAGGTAGAGCA TTTGACTGCAGATCAAGAGGTCCCC GGTTCAAATCTGGGTGCCCCCT
tC(GCA)G17	Hs_Cys	7	148,825,991	148,825,920	Cys	GCA		GGGGATATAGCTCAGGGGTAGAGCA TTTGACTGCAGATCAAGAGGTCCCC GGTTCAAATCCGGGTGCCCCCC
tC(GCA)G18	Hs_Cys	7	148,842,408	148,842,479	Cys	GCA		GGGGGTATAGCTCAGGGGTAGAGCA TTTGACTGCAGATCAAGAGGTCCCT GGTTCAAATCCAGGTGCCCCCT
tY(GTA)H1	Hs_Tyr	8	66,772,173	66,772,086	Tyr	GTA	Yes	TCTTCAATAGCTCAGCTGGTAGAGC GGAGGACTGTAGGTGCACGCCCGT GGCCATTCTTAGGTGCTGGTTTGATT CCGACTTGGAGAG
tY(GTA)H2	Hs_Tyr	8	67,188,156	67,188,248	Tyr	GTA	Yes	CCTTCGATAGCTCAGCTGGTAGAGC GGAGGACTGTAGCTACTTCCTCAGC AGGAGACATCCTTAGGTCGCTGGTT CGATTCCGGCTCGAAGGA
tY(GTA)H3	Hs_Tyr	8	67,188,777	67,188,865	Tyr	GTA	Yes	CCTTCGATAGCTCAGCTGGTAGAGC GGAGGACTGTAGGCGCGCGCCCGT GGCCATCCTTAGGTCGCTGGTTCGA TTCCGGCTCGAAGGA
tA(AGC)H	Hs_Ala2	8	67,188,978	67,189,050	Ala	AGC		GGGGGATTAGCTCAAATGGTAGAGC GCTCGCTTAGCATGCGAGAGGTAGC GGGATCGATGCCCGCATCCTCCA
tS(AGA)H	Hs_Ser1	8	96,351,142	96,351,061	Ser	AGA		GTAGTCGTGGCCGAGTGGTTAAGGC GATGGACTAGAAATCCATTGGGGTC TCCCCGCGCAGGTTCGAATCCTGCC GACTACG
tM(CAT)H	Hs_Met1	8	124,238,723	124,238,651	Met	CAT		GCCTCGTTAGCGCAGTAGGTAGCGC GTCAGTCTCATAATCTGAAGGTCGT GAGTTCGATCCTCACACGGGGCA

tH(GTG)I	Hs_His	9	14,424,009	14,423,938	His	GTG		GCCGTGATCGTATAGTGGTTAGTACT CTGCGTTGTGGCCGCAGCAACCTCG GTTCGAATCCGAGTCACGGCA
tR(TCG)I	Hs_Arg1	9	110,040,358	110,040,430	Arg	TCG		GGCCGTGTGGCCTAATGGATAAGGC GTCTGACTTCGGATCAAAAGATTGC AGGTTTGAGTTCTGCCACGGTCG
tR(TCT)I	Hs_Arg4	9	128,181,999	128,181,909	Arg	TCT	Yes	GGCTCTGTGGCGCAATGGATAGCGC ATTGGACTTCTAGCTGAGCCTAGTG TGGTCATTCAAAGGTTGTGGGTTCG AGTCCCACCAGAGTCG
tV(TAC)J	Hs_Val	10	5,935,752	5,935,680	Val	TAC		GGTTCCATAGTGTAGTGGTTATCACA TCTGCTTTACACGCAGAAGGTCCTG GGTTCAAGCCCCAGTGGAACCA
tN(GTT)J	Hs_Asn	10	22,558,517	22,558,444	Asn	GTT		GTCTCTGTGGCGCAATCGGTTAGCG CGTTCGGCTGTTAACCGAAAGGTTG GTGGTTCGAGCCCACCCAGGGACG
tS(TGA)J	Hs_Ser1	10	69,194,267	69,194,348	Ser	TGA		GCAGCGATGGCCGAGTGGTTAAGG CGTTGGACTTGAAATCCAATGGGGT CTCCCCGCGCAGGTTCGAACCCTGC TCGCTGCG
tL(CAA)K	Hs_Met1	11	9,253,366	9,253,439	Leu	CAA		GCCTCCTTAGTGCAGTAGGTAGCGC ATCAGTCTCAAAATCTGAATGGTCC TGAGTTCAAGCCTCAGAGGGGGCA
tA(TGC)K	Hs_Ala1	11	50,190,526	50,190,455	Ala	TGC		GGGGGTGTAGCTCAGTGGTAGAGC GGATGCTTTGCATGTATGAGACTTT GGGTTGGATCCCCAGCACCTCCA
tV(TAC)K1	Hs_Val	11	59,074,750	59,074,678	Val	TAC		GGTTCCATAGTGTAGTGGTTATCAC GTCTGCTTTACACGCAGAAGGTCCT GGGTTCGAGCCCCAGTGGAACCA
tV(TAC)K2	Hs_Val	11	59,075,108	59,075,036	Val	TAC		GGTTCCATAGTGTAGCGGTTATCAC GTCTGCTTTACACGCAGAAGGTCCT GGGTTCGAGCCCCAGTGGAACCA
tR(TCT)K	Hs_Arg4	11	59,075,343	59,075,428	Arg	TCT	Yes	GGCTCTGTGGCGCAATGGATAGCGC ATTGGACTTCTAGATAGTTAGAGAA ATTCAAAGGTTGTGGGTTCGAGTCC CACCAGAGTCG
tL(TAA)K	Hs_Leu3	11	59,075,804	59,075,886	Leu	TAA		ACCAGAATGGCCGAGTGGTTAAGG CGTTGGACTTAAGATCCAATGGATT CATATCCGCGTGGGTTCGAACCCCA CTTCTGGTA
tK(TTT)K1	Hs_Lys2	11	59,080,478	59,080,550	Lys	TTT		GCCCGGATAGCTCAGTCGGTAGAGC ATCAGACTTTTAATCTGAGGGTCCG GGGTTCAAGTCCCTGTTCGGGCG
tF(GAA)K1	Hs_Phe	11	59,081,618	59,081,546	Phe	GAA		GCCGAAATAGCTCAGTTGGGAGAG CGTTAGACTGAAGATCTAAAGGTCC CTGGTTCGATCCCGGGTTTCGGCA
tK(TTT)K2	Hs_Lys2	11	59,084,456	59,084,384	Lys	TTT		GCCCGGATAGCTCAGTCGGTAGAGC ATCAGACTTTTAATCTGAGGGTCCA GGGTTCAAGTCCCTGTTCGGGCG
tF(GAA)K2	Hs_Phe	11	59,090,501	59,090,429	Phe	GAA		GCCGAAATAGCTCAGTTGGGAGAG CGTTAGACTGAAGATCTAAAGGTCC CTGGTTCAATCCCGGGTTTCGGCA
tS(GCT)K	Hs_Ser2	11	65,872,167	65,872,248	Ser	GCT		GACGAGGTGGCCGAGTGGTTAAGG CGATGGACTGCTAATCCATTGTGCTT TGCACGCGTGGGTTCGAATCCCATC CTCGTCG
tP(AGG)K	Hs_Pro	11	75,624,205	75,624,276	Pro	AGG		GGCTCGTTGGTCTAGGGGTATGATT CTCGCTTAGGGTGCGAGAGGTCCCG GGTTCAAATCCCGGACGAGCCC

tP(TGG)K	Hs_Pro	11	75,624,588	75,624,517	Pro	TGG	GGCTCGTTGGTCTAGGGGTATGATT CTCGGTTTGGGTCCGAGAGGTCCCG GGTTCAAATCCCGGACGAGCCC
tK(TTT)K3	Hs_Lys2	11	121,935,865	121,935,937	Lys	TTT	GCCTGGATAGCTCAGTTGGTAGAGC ATCAGACTTTTAATCTGAGGGTCCA GGGTTCAAGTCCCTGTTCAGGCG
tS(CGA)L	Hs_Ser1	12	54,870,415	54,870,496	Ser	CGA	GTCACGGTGGCCGAGTGGTTAAGG CGTTGGACTCGAAATCCAATGGGGT TTCCCCGCACAGGTTCGAATCCTGT TCGTGACG
tD(GTC)L1	Hs_Asp	12	94,932,267	94,932,338	Asp	GTC	TCCTCGTTAGTATAGTGGTGAGTATC CCCGCCTGTCACGCGGGAGACCGG GGTTCGATTCCCCGACGGGGAG
tD(GTC)L2	Hs_Asp	12	97,399,749	97,399,820	Asp	GTC	TCCTCGTTAGTATAGTGGTTAGTATC CCCGCCTGTCACGCGGGAGACCGG GGTTCAATTCCCCGACGGGGAG
tW(CCA)L	Hs_Trp	12	97,400,498	97,400,569	Trp	CCA	GACCTCGTGGCGCAACGGTAGCGC GTCTGACTCCAGATCAGAAGGCTGC GTGTTCGAATCACGTCGGGGTCA
tD(GTC)L3	Hs_Asp	12	121,385,804	121,385,874	Asp	GTC	TCCTTGTTAGTATAGTGGTGAGTGTT TCTGCCTGTCATGTGGAGACTGGAG TTTGAGTCCCCAACAGGGAG
tA(TGC)L1	Hs_Ala1	12	123,931,252	123,931,181	Ala	TGC	GGGGATGTAGCTCAGTGGTAGAGC GCATGCTTTGCATGTATGAGGCCCC GGGTTCGATCCCCGGCATCTCCA
tD(GTC)L4	Hs_Asp	12	123,936,842	123,936,771	Asp	GTC	TCCTCGTTAGTATAGTGGTGAGTATC CCCGCCTGTCACGCGGGAGACCGG GGTTCGATTCCCCGACGGGGAG
tF(GAA)L	Hs_Phe	12	123,937,341	123,937,269	Phe	GAA	GCCGAAATAGCTCAGTTGGGAGAG CGTTAGACTGAAGATCTAAAGGTCC CTGGTTCGATCCCGGGTTTCGGCA
tD(GTC)L5	Hs_Asp	12	123,949,144	123,949,073	Asp	GTC	TCCTCGTTAGTATAGTGGTGAGTATC CCCGCCTGTCACGCGGGAGACCGG GGTTCGATTCCCCGACGGGGAG
tA(TGC)L2	Hs_Ala1	12	123,949,392	123,949,463	Ala	TGC	GGGGATGTAGCTCAGTGGTAGAGC GCATGCTTTGCACGTATGAGGCCCC GGGTTCAATCCCCGGCATCTCCA
tN(GTT)M	Hs_Asn	13	30,146,174	30,146,101	Asn	GTT	GTCTCTGTGGCGCAATCGGTTAGCG CGTTCGGCTGTTAACCGAAAGGTTG GTGGTTCGAGCCCACCCAGGGACG
tE(TTC)M1	Hs_Glu	13	40,532,945	40,532,874	Glu	TTC	TCCCATATGGTCTAGCGGTTAGGATT CCTGGTTTTCACCCAGGTGGCCCGG GTTCGACTCCCGGTATGGGAA
tE(TTC)M2	Hs_Glu	13	44,390,133	44,390,062	Glu	TTC	TCCCACATGGTCTAGCGGTTAGGAT TCCTGGTTTTCACCCAGGCGGCCCG GGTTCGACTCCCGGTGTGGGAA
tF(GAA)M	Hs_Phe	13	93,999,977	93,999,905	Phe	GAA	GCCGAAATAGCTCAGTTGGGAGAG CGTTAGACTGAAGATCTAAAGGTCC CTGGTTCGATCCCGGGTTTCGGCA
tP(AGG)N1	Hs_Pro	14	20,147,406	20,147,335	Pro	AGG	GGCTCGTTGGTCTAGGGGTATGATT CTCGCTTAGGGTGCGAGAGGTCCCG GGTTCAAATCCCGGACGAGCCC
tL(AAG)N	Hs_Leu1	14	20,148,131	20,148,212	Leu	AAG	GGTAGCGTGGCCGAGCGGTCTAAG GCGCTGGATTAAGGCTCCAGTCTCT TCGGGGGCGTGGGTTCGAATCCCAC CGCTGCCA
tP(AGG)N2	Hs_Pro	14	20,151,471	20,151,400	Pro	AGG	GGCTCGTTGGTCTAGGGGTATGATT CTCGCTTAGGGTGCGAGAGGTCCCG GGTTCAAATCCCGGACGAGCCC
tT(TGT)N1	Hs_Thr2	14	20,151,861	20,151,789	Thr	TGT	GGCTCCATAGCTCAGGGGTTAGAGC GCTGGTCTTGTAAACCAGGGGTCGC GAGTTCAATTCTCGCTGGGGCCT
tL(TAG)N	Hs_Leu1	14	20,163,369	20,163,450	Leu	TAG	GGTAGTGTGGCCGAGCGGTCTAAG GCGCTGGATTTAGGCTCCAGTCTCT TCGGGGGCGTGGGTTCGAATCCCAC CACTGCCA

tT(TGT)N2	Hs_Thr2	14	20,169,231	20,169,159	Thr	TGT		GGCTCCATAGCTCAGGGGTTAGAGC ACTGGTCTTGTAAACCAGGGGTCGC GAGTTCAAATCTCGCTGGGGCCT
tP(TGG)N1	Hs_Pro	14	20,171,005	20,171,076	Pro	TGG		GGCTCGTTGGTCTAGTGGTATGATT CTCGCTTTGGGTGCGAGAGGTCCCG GGTTCAAATCCCGGACGAGCCC
tY(GTA)N1	Hs_Tyr	14	20,191,191	20,191,098	Tyr	GTA	Yes	CCTTCGATAGCTCAGCTGGTAGAGC GGAGGACTGTAGCCTGTAGAAACAT TTGTGGACATCCTTAGGTCGCTGGT TCGATTCCGGCTCGAAGGA
tY(GTA)N2	Hs_Tyr	14	20,195,556	20,195,463	Tyr	GTA	Yes	CCTTCGATAGCTCAGCTGGTAGAGC GGAGGACTGTAGATTGTATAGACAT TTGCGGACATCCTTAGGTCGCTGGT TCGATTCCAGCTCGAAGGA
tY(GTA)N3	Hs_Tyr	14	20,198,050	20,197,957	Tyr	GTA	Yes	CCTTCGATAGCTCAGCTGGTAGAGC GGAGGACTGTAGACTGCGGAAACG TTTGTGGACATCCTTAGGTCGCTGG TTCAATTCCGGCTCGAAGGA
tY(GTA)N4	Hs_Tyr	14	20,201,284	20,201,191	Tyr	GTA	Yes	CCTTCGATAGCTCAGCTGGTAGAGC GGAGGACTGTAGATTGTACAGACAT TTGCGGACATCCTTAGGTCGCTGGT TCGATTCCGGCTCGAAGGA
tT(TGT)N3	Hs_Thr2	14	20,219,689	20,219,761	Thr	TGT		GGCCCTATAGCTCAGGGGTTAGAGC ACTGGTCTTGTAAACCAGGGGTCGC GAGTTCAAATCTCGCTGGGGCCT
tY(GTA)N5	Hs_Tyr	14	20,221,272	20,221,360	Tyr	GTA	Yes	CCTTCGATAGCTCAGCTGGTAGAGC GGAGGACTGTAGTACTTAATGTGTG GTCATCCTTAGGTCGCTGGTTCGATT CCGGCTCGAAGGA
tP(TGG)N2	Hs_Pro	14	20,222,015	20,222,086	Pro	TGG		GGCTCGTTGGTCTAGGGGTATGATT CTCGCTTTGGGTGCGAGAGGTCCCG GGTTCAAATCCCGGACGAGCCC
tR(ACG)N	Hs_Arg2	14	22,468,750	22,468,822	Arg	ACG		GGGCCAGTGGCGCAATGGATAACG CGTCTGACTACGGATCAGAAGATTC CAGGTTCGACTCCTGGCTGGCTCG
tK(CTT)N	Hs_Lys1	14	57,776,438	57,776,366	Lys	CTT		GCCCGGCTAGCTCAGTCGGTAGAGC ATGGGACTCTTAATCCCAGGGTCGT GGGTTCGAGCCCCACGTTGGGCG
tC(GCA)N	Hs_Cys	14	72,499,432	72,499,503	Cys	GCA		GGGGGTATAGCTCAGGGGTAGAGCA TTTGACTGCAGATCAAGAGGTCCCC GGTTCAAATCCGGGTGCCCCCT
tA(AGC)N	Hs_Ala2	14	88,515,195	88,515,267	Ala	AGC		GGGGAATTAGCTCAAGTGGTAGAG CGCTCGCTTAGCATGCGAGAGGTAG TGGGATCGATGCCCGCATTCTCCA
tI(AAT)N	Hs_Ile1	14	101,853,182	101,853,255	Ile	AAT		GGCCGGTTAGCTCAGTTGGTTAGAG CGTGGTGCTAATAACGCCAAGGTCG CGGGTTCGATCCCCGTACGGGCCA
tE(TTC)O	Hs_Glu	15	23,878,545	23,878,474	Glu	TTC		TCCCACATGGTCTAGCGGTTAGGAT TCCTGGTTTTCACCCAGGCGGCCCG GGTTCGACTCCCGGTGTGGGAA
tS(GCT)O	Hs_Ser2	15	38,673,396	38,673,315	Ser	GCT		GACGAGGTGGCCGAGTGGTTAAGG CGATGGACTGCTAATCCATTGTGCT CTGCACGCGTGGGTTCGAATCCCAT CCTCGTCG
tH(GTG)O1	Hs_His	15	43,278,167	43,278,096	His	GTG		GCCGTGATCGTATAGTGGTTAGTACT CTGCGTTGTGGCCGCAGCAACCTCG GTTCGAATCCGAGTCACGGCA
tH(GTG)O2	Hs_His	15	43,279,974	43,279,903	His	GTG		GCCGTGATCGTATAGTGGTTAGTACT CTGCGTTGTGGCCGCAGCAACCTCG GTTCGAATCCGAGTCACGGCA
tH(GTG)O3	Hs_His	15	43,280,641	43,280,712	His	GTG		GCCGTGATCGTATAGTGGTTAGTACT CTGCGTTGTGGCCGCAGCAACCTCG GTTCGAATCCGAGTCACGGCA

tQ(CTG)O	Hs_Gln1	15	63,948,525	63,948,454	Gln	CTG		GGTTCCATGGTGTAATGGTTAGCAC TCTGGACTCTGAATCCAGCGATCCG AGTTCAAATCTCGGTGGAACCT
tK(CTT)O	Hs_Lys1	15	76,939,959	76,940,031	Lys	CTT		GCCCGGCTAGCTCAGTCGGTAGAGC ATGGGACTCTTAATCCCAGGGTCGT GGGTTCGAGCCCCACGTTGGGCG
tC(GCA)O	Hs_Cys	15	77,824,052	77,824,124	Cys	GCA		GGGGGTATAGCTCAGTGGGTAGAGC ATTTGACTGCAGATCAAGAGGTCCC CGGTTCAAATCCGGGTGCCCCCT
tR(TCG)O	Hs_Arg1	15	87,679,308	87,679,380	Arg	TCG		GGCCGCGTGGCCTAATGGATAAGGC GTCTGACTTCGGATCAGAAGATTGC AGGTTCGAGTCCTGCCGCGGTCG
tG(CCC)P	Hs_Gly3	16	626,807	626,737	Gly	CCC		GCGCCGCTGGTGTAGTGGTATCATG CAAGATTCCCATTCTTGCGACCCGG GTTCGATTCCCGGGCGGCGCA
tR(CCG)P	Hs_Arg1	16	3,140,676	3,140,748	Arg	CCG		GGCCGCGTGGCCTAATGGATAAGGC GTCTGATTCCGGATCAGAAGATTGA GGGTTCGAGTCCCTTCGTGGTCG
tR(CCT)P1	Hs_Arg3	16	3,142,902	3,142,974	Arg	CCT		GCCCCGGTGGCCTAATGGATAAGGC ATTGGCCTCCTAAGCCAGGGATTGT GGGTTCGAGTCCCACCCGGGGTA
tK(CTT)P1	Hs_Lys1	16	3,147,479	3,147,407	Lys	CTT		GCCCGGCTAGCTCAGTCGGTAGAGC ATGAGACCCTTAATCTCAGGGTCGT GGGTTCGAGCCCCACGTTGGGCG
tP(TGG)P1	Hs_Pro	16	3,148,924	3,148,995	Pro	TGG		GGCTCGTTGGTCTAGGGGTATGATT CTCGCTTTGGGTGCGAGAGGTCCCG GGTTCAAATCCCGGACGAGCCC
tP(AGG)P1	Hs_Pro	16	3,150,387	3,150,481	Pro	AGG	Yes	GGCTCGTTGGTCTAGGGGTGTGGTT CTCGCTTAGGGACCACAGGGACAA GCCCGGGAGACCCAAGAGGTCCCG GGTTCAAATCCCGGACGAGCCC
tP(CGG)P	Hs_Pro	16	3,162,050	3,162,121	Pro	CGG		GGCTCGTTGGTCTAGGGGTATGATT CTCGCTTCGGGTGCGAGAGGTCCC GGGTTCAAATCCCGGACGAGCCC
tK(CTT)P2	Hs_Lys1	16	3,165,693	3,165,765	Lys	CTT		GCCCGGCTAGCTCAGTCGGTAGAGC ATGAGACTCTTAATCTCAGGGTCGT GGGTTCGAGCCCCACGTTGGGCG
tK(CTT)P3	Hs_Lys1	16	3,170,628	3,170,556	Lys	CTT		GCCCGGCTAGCTCAGTCGATAGAGC ATGAGACTCTTAATCTCAGGGTCGT GGGTTCGAGCCGCACGTTGGGCG
tP(AGG)P2	Hs_Pro	16	3,172,707	3,172,636	Pro	AGG		GGCTCGTTGGTCTAGGGGTATGATT CTCGCTTAGGGTGCGAGAGGTCCCG GGTTCAAATCCCGGACGAGCCC
tP(TGG)P2	Hs_Pro	16	3,174,205	3,174,134	Pro	TGG		GGCTCGTTGGTCTAGGGGTATGATT CTCGCTTTGGGTGCGAGAGGTCCCG GGTTCAAATCCCGGACGAGCCC
tP(TGG)P3	Hs_Pro	16	3,178,095	3,178,166	Pro	TGG		GGCTCGTTGGTCTAGGGGTATGATT CTCGCTTTGGGTGCGAGAGGTCCCG GGTTCAAATCCCGGACGAGCCC
tP(AGG)P3	Hs_Pro	16	3,179,635	3,179,706	Pro	AGG		GGCTCGTTGGTCTAGGGGTATGATT CTCGCTTAGGGTGCGAGAGGTCCCG GGTTCAAATCCCGGACGAGCCC
tK(CTT)P4	Hs_Lys1	16	3,181,502	3,181,574	Lys	CTT		GCCCGGCTAGCTCAGTCGGTAGAGC ATGGGACTCTTAATCTCAGGGTCGT GGGTTCGAGCCCCACGTTGGGCG
tP(AGG)P4	Hs_Pro	16	3,181,990	3,182,061	Pro	AGG		GGCTCGTTGGTCTAGGGGTATGATT CTCGCTTAGGATGCGAGAGGTCCCG GGTTCAAATCCCGGACGAGCCC
tR(CCT)P2	Hs_Arg3	16	3,183,919	3,183,991	Arg	CCT		GCCCCAGTGGCCTGATGGATAAGGT ACTGGCCTCCTAAGCCAGGGATTGT GGGTTCGAGTTCCACCTGGGGTA

tT(CGT)P	Hs_Thr1	16	14,287,251	14,287,322	Thr	CGT		GGCGCGGTGGCCAAGTGGTAAGGC GTCGGTCTCGTAAACCGAAGATCAC GGGTTCGAACCCCGTCCGTGCCT
tL(TAG)P	Hs_Leu1	16	22,114,614	22,114,533	Leu	TAG		GGTAGCGTGGCCGAGTGGTCTAAG GCGCTGGATTTAGGCTCCAGTCATT TCGATGGCGTGGGTTCGAATCCCAC CGCTGCCA
tL(AAG)P	Hs_Leu1	16	22,215,962	22,216,043	Leu	AAG		GGTAGCGTGGCCGAGCGGTCTAAG GCGCTGGATTAAGGCTCCAGTCTCT TCGGGGGCGTGGGTTCGAATCCCAC CGCTGCCA
tL(CAG)P1	Hs_Leu2	16	55,891,364	55,891,446	Leu	CAG		GTCAGGATGGCCGAGCGGTCTAAG GCGCTGCGTTCAGGTCGCAGTCTCC CCTGGAGGCGTGGGTTCGAATCCCA CTTCTGACA
tL(CAG)P2	Hs_Leu2	16	55,891,975	55,891,893	Leu	CAG		GTCAGGATGGCCGAGCGGTCTAAG GCGCTGCGTTCAGGTCGCAGTCTCC CCTGGAGGCGTGGGTTCGAATCCCA CTTCTGACA
tG(GCC)P1	Hs_Gly1	16	69,369,685	69,369,615	Gly	GCC		GCATTGGTGGTTCAGTGGTAGAATT CTCGCCTGCCACGCGGGAGGCCCG GGTTTGATTCCCGGCCAGTGCA
tG(GCC)P2	Hs_Gly1	16	69,370,513	69,370,443	Gly	GCC		GCATTGGTGGTTCAGTGGTAGAATT CTCGCCTGCCACGCGGGAGGCCCG GGTTCGATTCCCGGCCAATGCA
tG(GCC)P3	Hs_Gly1	16	69,380,098	69,380,168	Gly	GCC		GCATTGGTGGTTCAGTGGTAGAATT CTCGCCTGCCATGCGGGCGGCCGG GCTTCGATTCCTGGCCAATGCA
tG(GCC)P4	Hs_Gly1	16	69,380,911	69,380,981	Gly	GCC		GCATTGGTGGTTCAGTGGTAGAATT CTCGCCTGCCACGCGGGAGGCCCG GGTTCGATTCCCGGCCAATGCA
tM(CAT)P1	Hs_Met1	16	70,017,897	70,017,969	Met	CAT		GCCCTCTTAGCGCAGTGGGCAGCGC GTCAGTCTCATAATCTGAAGGTCCT GAGTTCGAGCCTCAGAGAGGGCA
tK(TTT)P	Hs_Lys2	16	72,069,789	72,069,717	Lys	TTT		GCCTGGATAGCTCAGTTGGTAGAGC ATCAGACTTTTAATCTGAGGGTCCA GGGTTCAAGTCCCTGTTCAGGCA
tM(CAT)P2	Hs_Met1	16	85,975,201	85,975,129	Met	CAT		GCCTCGTTAGCGCAGTAGGCAGCGC GTCAGTCTCATAATCTGAAGGTCGT GAGTTCGAGCCTCACACGGGGCA
tK(TTT)Q	Hs_Lys2	17	7,963,198	7,963,270	Lys	TTT		GCCCGGATAGCTCAGTCGGTAGAGC ATCAGACTTTTAATCTGAGGGTCCA GGGTTCAAGTCCCTGTTCGGGCG
tQ(CTG)Q	Hs_Gln1	17	7,963,795	7,963,866	Gln	CTG		GGTTCCATGGTGTAATGGTTAGCAC TCTGGACTCTGAATCCAGCGATCCG AGTTCAAATCTCGGTGGAACCT
tL(TAG)Q	Hs_Leu1	17	7,964,438	7,964,357	Leu	TAG		GGTAGCGTGGCCGAGCGGTCTAAG GCGCTGGATTTAGGCTCCAGTCTCT TCGGAGGCGTGGGTTCGAATCCCAC CGCTGCCA
tR(TCT)Q	Hs_Arg4	17	7,964,968	7,965,055	Arg	TCT	Yes	GGCTCTGTGGCGCAATGGATAGCGC ATTGGACTTCTAGTGACGAATAGAG CAATTCAAAGGTTGTGGGTTCGAAT CCCACCAGAGTCG
tG(GCC)Q	Hs_Gly1	17	7,969,789	7,969,859	Gly	GCC		GCATTGGTGGTTCAGTGGTAGAATT CTCGCCTGCCACGCGGGAGGCCCG GGTTCGATTCCCGGCCAATGCA
tS(CGA)Q	Hs_Ser1	17	7,983,005	7,982,924	Ser	CGA		GCTGTGATGGCCGAGTGGTTAAGGC GTTGGACTCGAAATCCAATGGGGTC TCCCCGCGCAGGTTCGAATCCTGCT CACAGCG
tT(AGT)Q1	Hs_Thr1	17	7,983,568	7,983,495	Thr	AGT		GGCGCCGTGGCTTAGCTGGTTAAAG CGCCTGTCTAGTAAACAGGAGATCC TGGGTTCGAATCCCAGCGGTGCCT

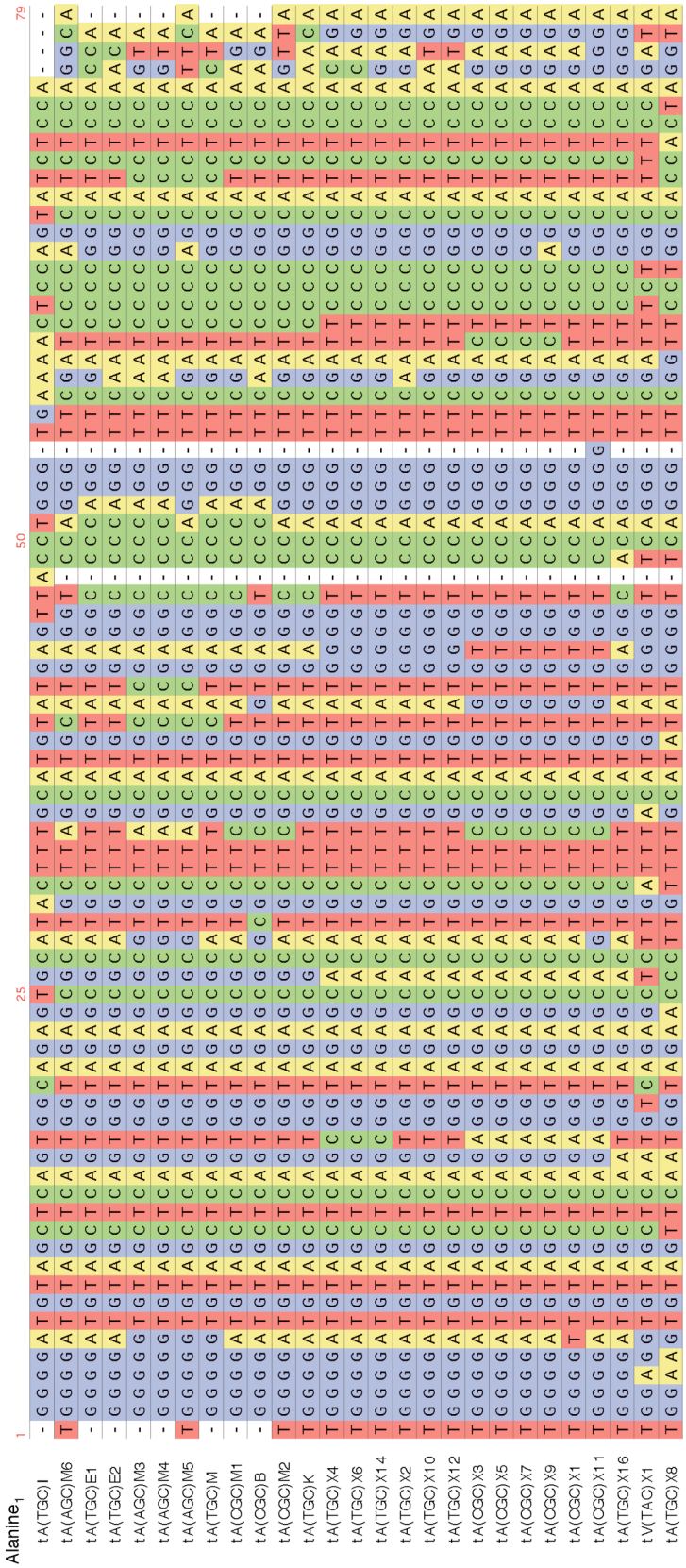
tW(CCA)Q1	Hs_Trp	17	8,030,401	8,030,472	Trp	CCA	GACCTCGTGGCGCAACGGTAGCGC GTCTGACTCCAGATCAGAAGGTTGC GTGTTCAAATCACGTCGGGGTCA
tS(GCT)Q	Hs_Ser2	17	8,030,909	8,030,990	Ser	GCT	GACGAGGTGGCCGAGTGGTTAAGG CGATGGACTGCTAATCCATTGTGCT CTGCACGCGTGGGTTCGAATCCCAT CCTCGTCG
tT(AGT)Q2	Hs_Thr1	17	8,031,203	8,031,276	Thr	AGT	GGCGCCGTGGCTTAGTTGGTTAAAG CGCCTGTCTAGTAAACAGGAGATCC TGGGTTCGAATCCCAGCGGTGCCT
tI(AAT)Q1	Hs_Ile1	17	8,031,636	8,031,709	Ile	AAT	GGCCGGTTAGCTCAGTTGGTTAGAG CGTGGTGCTAATAACGCCAAGGTCG CGGGTTCGATCCCCGTACGGGCCA
tW(CCA)Q2	Hs_Trp	17	8,064,983	8,064,912	Trp	CCA	GGCCTCGTGGCGCAACGGTAGCGC GTCTGACTCCAGATCAGAAGGTTGC GTGTTCAAATCACGTCGGGGTCA
tG(TCC)Q	Hs_Gly2	17	8,065,591	8,065,662	Gly	TCC	GCGTTGGTGGTATAGTGGTAAGCAT AGCTGCCTTCCAAGCAGTTGACCCG GGTTCGATTCCCGGCCAACGCA
tD(GTC)Q	Hs_Asp	17	8,066,352	8,066,281	Asp	GTC	TCCTCGTTAGTATAGTGGTGAGTATC CCCGCCTGTCACGCGGGAGACCGG GGTTCGATTCCCCGACGGGGAG
tP(CGG)Q	Hs_Pro	17	8,066,947	8,066,876	Pro	CGG	GGCTCGTTGGTCTAGGGGTATGATT CTCGCTTCGGGTGCGAGAGGTCCC GGGTTCAAATCCCGGACGAGCCC
tT(AGT)Q3	Hs_Thr1	17	8,070,351	8,070,278	Thr	AGT	GGCGCCGTGGCTTAGTTGGTTAAAG CGCCTGTCTAGTAAACAGGAGATCC TGGGTTCGAATCCCAGCGGTGCCT
tS(AGA)Q	Hs_Ser1	17	8,070,734	8,070,653	Ser	AGA	GTAGTCGTGGCCGAGTGGTTAAGGC GATGGACTAGAAATCCATTGGGGTC TCCCCGCGCAGGTTCGAATCCTGCC GACTACG
tI(AAT)Q2	Hs_Ile1	17	8,071,107	8,071,034	Ile	AAT	GGCCGGTTAGCTCAGTTGGTTAGAG CGTGGTGCTAATAACGCCAAGGTCG CGGGTTCGAACCCCGTACGGGCCA
tW(CCA)Q3	Hs_Trp	17	19,352,086	19,352,157	Trp	CCA	GACCTCGTGGCGCAATGGTAGCGCG TCTGACTCCAGATCAGAAGGTTGCG TGTTCAAGTCACGTCGGGGTCA
tG(CCC)Q	Hs_Gly1	17	19,704,767	19,704,837	Gly	CCC	GCATTGGTGGTTCAATGGTAGAATT CTCGCCTCCCACGCAGGAGACCCA GGTTCGATTCCTGGCCAATGCA
tT(CGT)Q	Hs_Thr1	17	26,901,213	26,901,284	Thr	CGT	GGCGCGGTGGCCAAGTGGTAAGGC GTCGGTCTCGTAAACCGAAGATCGC GGGTTCGAACCCCGTCCGTGCCT
tN(GTT)Q	Hs_Asn	17	34,161,633	34,161,560	Asn	GTT	GTCTCTGTGGCGCAATCGGTTAGCG CGTTCGGCTGTTAACCGAAAGGTTG GTGGTTCGAGCCCACCCAGGGACG
tC(GCA)Q1	Hs_Cys	17	34,271,534	34,271,463	Cys	GCA	GGGGGTATAGCTCAGGGGTAGAGCA TTTGACTGCAGATCAAGAAGTCCCC GGTTCAAATCCGGGTGCCCCCT
tC(GCA)Q2	Hs_Cys	17	34,277,424	34,277,495	Cys	GCA	GGGGGTATAGCTCAGTGGTAGAGCA TTTGACTGCAGATCAAGAGGTCCCC GGTTCAAATCCGGGTGCCCCCT
tC(GCA)Q3	Hs_Cys	17	34,279,142	34,279,071	Cys	GCA	GGGGGTATAGCTCAGTGGTAGAGCA TTTGACTGCAGATCAAGAGGTCCCT GGTTCAAATCCGGGTGCCCCCT
tC(GCA)Q4	Hs_Cys	17	34,563,584	34,563,513	Cys	GCA	GGGGGTATAGCTCAGTGGTAGAGCA TTTGACTGCAGATCAAGAGGTCCCC GGTTCAAATCCGGGTGCCCCCT
tC(GCA)Q	Hs_Cys	17	34,564,341	34,564,270	Cys	GCA	GGGGGTATAGCTCAGTGGTAGAGCA TTTGACTGCAGATCAAGAGGTCCCC GGTTCAAATCCGGGTGCCCCCT
tQ(TTG)Q	Hs_Gln1	17	44,624,889	44,624,960	Gln	TTG	GGTCCCATGGTGTAATGGTTAGCAC TCTGGACTTTGAATCCAGCGATCCG AGTTCAAATCTCGGTGGGACCT

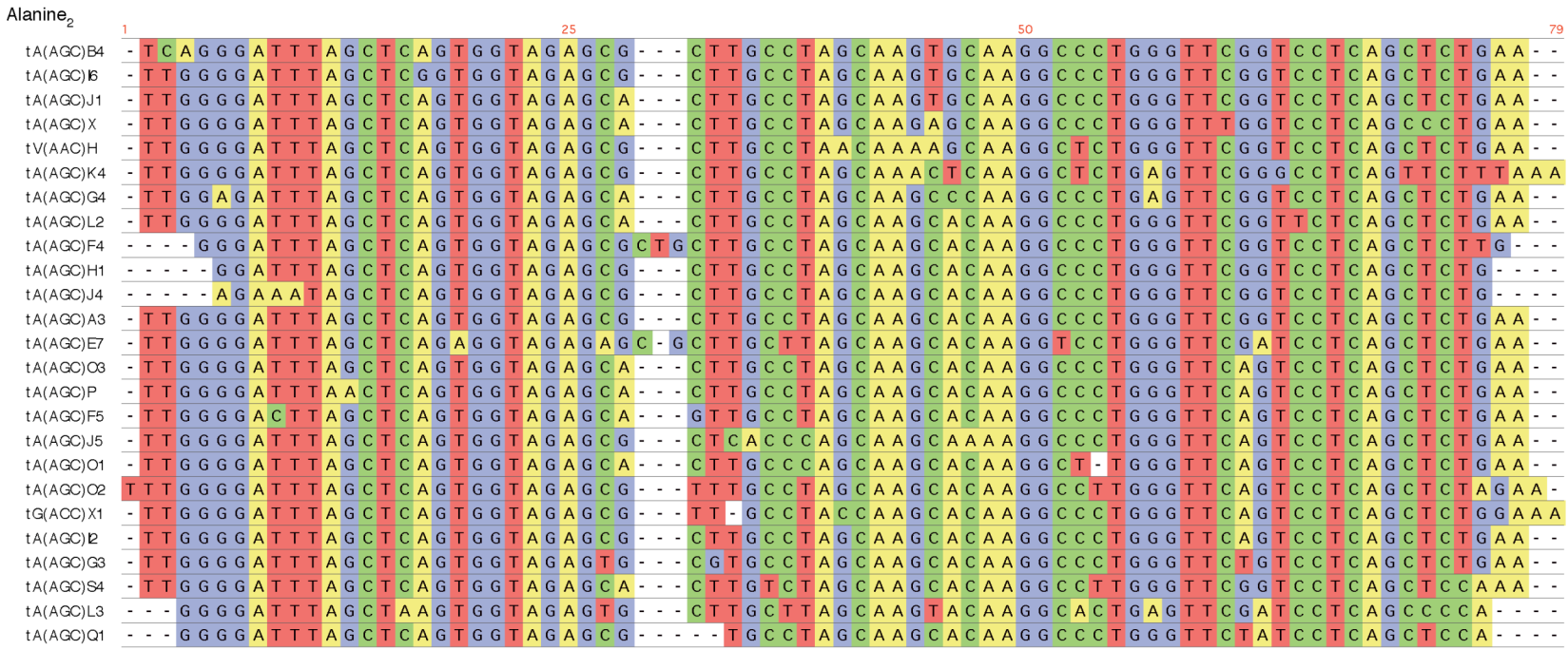
tSUP(TTA)Q	Hs_Lys2	17	56,218,375	56,218,445	Sup	TTA		GCCCGGATAGTTCAGTTGGTAGAGC ATCAGACTTAATCAGAGGGTCCAGG GTTCAAGTCCCTGTTTGGGTG
tR(CCG)Q	Hs_Arg3	17	63,446,547	63,446,475	Arg	CCG		GACCCAGTGGCCTAATGGATAAGGC ATCAGCCTCCGGAGCTGGGGATTGT GGGTTCGAGTCCCATCTGGGTCG
tR(CCT)Q1	Hs_Arg3	17	70,541,596	70,541,668	Arg	CCT		GCCCCAGTGGCCTAATGGATAAGGC ACTGGCCTCCTAAGCCAGGGATTGT GGGTTCGAGTCCCACCTGGGGTA
tR(CCT)Q2	Hs_Arg3	17	70,542,193	70,542,121	Arg	CCT		GCCCCAGTGGCCTAATGGATAAGGC ACTGGCCTCCTAAGCCAGGGATTGT GGGTTCGAGTCCCACCTGGGGTG
tR(TCG)Q	Hs_Arg1	17	70,542,803	70,542,875	Arg	TCG		GACCGCGTGGCCTAATGGATAAGGC GTCTGACTTCGGATCAGAAGATTGA GGGTTCGAGTCCCTTCGTGGTCG
tM(CAT)Q	Hs_Met2	17	78,045,957	78,045,886	Met	CAT		AGCAGAGTGGCGCAGCGGAAGCGT GCTGGGCCCATAACCCAGAGGTCGA TGGATCGAAACCATCCTCTGCTA
tK(CTT)R	Hs_Lys1	18	41,923,341	41,923,269	Lys	CTT		GACGAGCTAGCTCAGTCGGTAGAG CATGGGACTCTTAATCCCAGGGTCG TGGGTTTGAGCCCCATGTTGGGCA
tF(GAA)S	Hs_Phe	19	1,334,433	1,334,361	Phe	GAA		GCCGAAATAGCTCAGTTGGGAGAG CGTTAGACTGAAGATCTAAAGGTCC CTGGTTCGATCCCGGGTTTCGGCA
tN(GTT)S	Hs_Asn	19	1,334,562	1,334,635	Asn	GTT		GTCTCTGTGGCGCAATCGGTTAGCG CGTTCGGCTGTTAACCGAAAGGTTG GTGGTTCGAGCCCACCCAGGGACG
tG(TCC)S	Hs_Gly2	19	4,675,082	4,675,153	Gly	TCC		GCGTTGGTGGTATAGTGGTTAGCAT AGCTGCCTTCCAAGCAGTTGACCCG GGTTCGATTCCCGGCCAACGCA
tV(CAC)S	Hs_Val	19	4,675,719	4,675,647	Val	CAC		GTTTCCGTAGTGTAGCGGTTATCACA TTCGCCTCACACGCGAAAGGTCCCC GGTTCGATCCCGGGCGGAAACA
tT(AGT)S	Hs_Thr1	19	38,359,803	38,359,876	Thr	AGT		GGCGCCGTGGCTTAGTTGGTTAAAG CGCCTGTCTAGTAAACAGGAGATCC TGGGTTCGAATCCCAGCGGTGCCT
tI(TAT)S	Hs_Ile2	19	44,594,740	44,594,648	Ile	TAT	Yes	GCTCCAGTGGCGCAATCGGTTAGCG CGCGGTACTTATATGACAGTGCGAG CGGAGCAATGCCGAGGTTGTGAGT TCGATCCTCACCTGGAGCA
tSeC(TCA)S	Hs_SeC	19	50,673,785	50,673,700	SeC(e)	TCA		GCCCGGATGATCCTCAGTGGTCTGG GGTGCAGGCTTCAAACCTGTAGCTG TCTAGCGACAGAGTGGTTCAATTCC ACCTTTCGGGC
tK(TTT)S	Hs_Lys2	19	54,729,817	54,729,745	Lys	TTT		ACCTGGGTAGCTTAGTTGGTAGAGC ATTGGACTTTTAATTTGAGGGCCCA GGTTTCAAGTCCCTGTTTGGGTG
tG(GCC)U	Hs_Gly1	21	17,749,048	17,748,978	Gly	GCC		GCATGGGTGGTTCAGTGGTAGAATT CTCGCCTGCCACGCGGGAGGCCCG GGTTCGATTCCCGGCCCATGCA
tSeC(TCA)V	Hs_SeC	22	42,871,438	42,871,523	SeC(e)	TCA		GCTCGGATGATCCTCAGTGGTCTGG GGTGCAGGCTTCAAACCTGTAGCTG TCTAGTGACAGAGTGGTTCAATTCC ACCTTTGTAGG
tN(GTT)A	Hs_Asn	1_random	906,435	906,508	Asn	GTT		GTCTCTGTGGCGCAATCGGCTAGCG CGTTTGGCTGTTAACTAAAAGGTTG GCGGTTCGAACCCACCCAGAGGCG

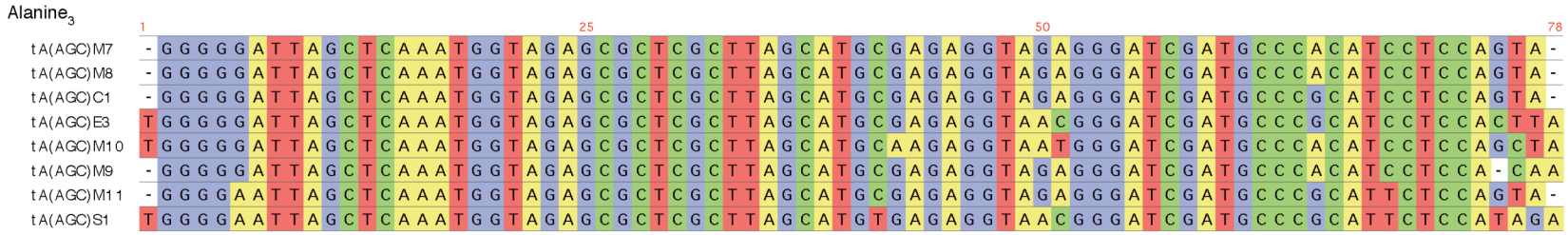
tT(TGT)A1	Hs_Thr2	1_random	1,654,722	1,654,794	Thr	TGT	GGCTCCATAGCTCAGTGGTTAGAGC ACTGGTCTTGTAAACCAGGGGTCGC GAGTTCGATCCTCGCTGGGGCCT
tT(TGT)A2	Hs_Thr2	1_random	2,030,046	2,030,118	Thr	TGT	GGCTCCATAGCTCAGTGGTTAGAGC ACTGGTCTTGTAAACCAGGGGTCGC GAGTTCGATCCTCGCTGGGGCCT
tR(CCG)Q	Hs_Arg1	17_random	1,279,261	1,279,333	Arg	CCG	GACCCAGTGGCCTAATGGATAAGGC ATCAGCCTCCGGAGCTGGGGATTGT GGGTTCGAGTCCCATCTGGGTCG
tK(TTT)G	Hs_Lys2	7_random	626,926	627,002	Lys	TTT	GCCCACGTAGCTCAATGGTCAGAGC GTGCGGCTTTTAACCGCAAGGAAG GCTGCGAGTTCGACCCTCGCCGTGG GCT
tV(TAC)X	Hs_Val	X	18,452,758	18,452,686	Val	TAC	GGTTCCATAGTGTAGTGGTTATCAC GTCTGCTTTACACGCAGAAGGTCCT GGGTTCGAGCCCCAGTGGAACCA
tI(GAT)X1	Hs_Ile1	X_random	86,496	86,423	Ile	GAT	GGCCGGTTAGCTCAGTTGGTAAGAG CGTGGTGCTGATAACACCAAGGTCG CGGGCTCGACTCCCGCACCGGCCA
tI(GAT)X2	Hs_Ile1	X_random	118,398	118,471	Ile	GAT	GGCCGGTTAGCTCAGTTGGTAAGAG CGTGGTGCTGATAACACCAAGGTCG CGGGCTCGACTCCCGCACCGGCCA
tI(GAT)X3	Hs_Ile1	X_random	399,021	398,948	Ile	GAT	GGCCGGTTAGCTCAGTTGGTAAGAG CGTGGTGCTGATAACACCAAGGTCG CGGGCTCGACTCCCGCACCGGCCA
tI(GAT)X4	Hs_Ile1	X_random	406,943	407,016	Ile	GAT	GGCCGGTTAGCTCAGTTGGTAAGAG CGTGGTGCTGATAACACCAAGGTCG CGGGCTCGACTCCCGCACCGGCCA
tI(GAT)X5	Hs_Ile1	X_random	465,544	465,617	Ile	GAT	GGCCGGTTAGCTCAGTTGGTAAGAG CGTGGTGCTGATAACACCAAGGTCG CGGGCTCGACTCCCGCACCGGCCA

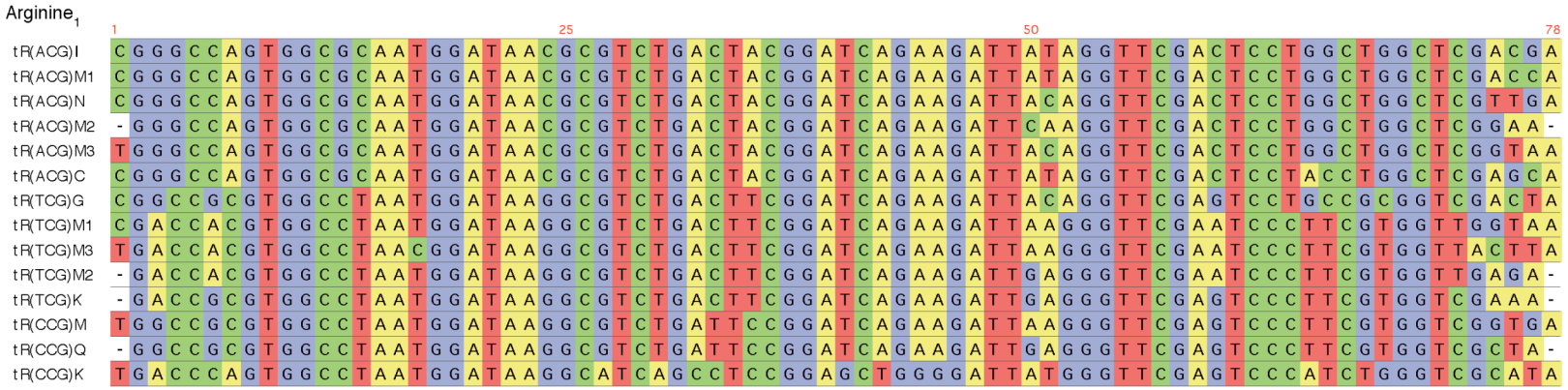
Appendix G

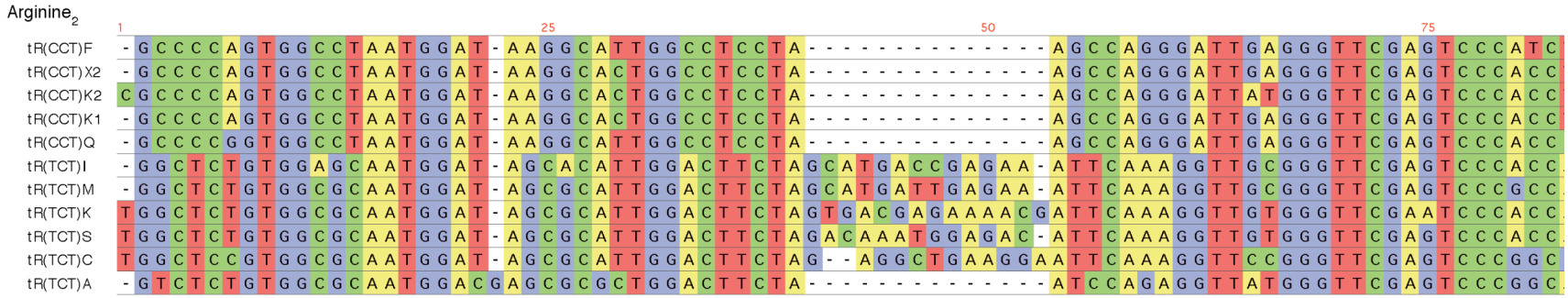
Shown here are sequence alignments for all the mouse tRNA gene families. The sequence variation is much greater in mouse than in yeast or bacteria. The gene names are on the left column. If a gene contains an intron it is included in the alignment.

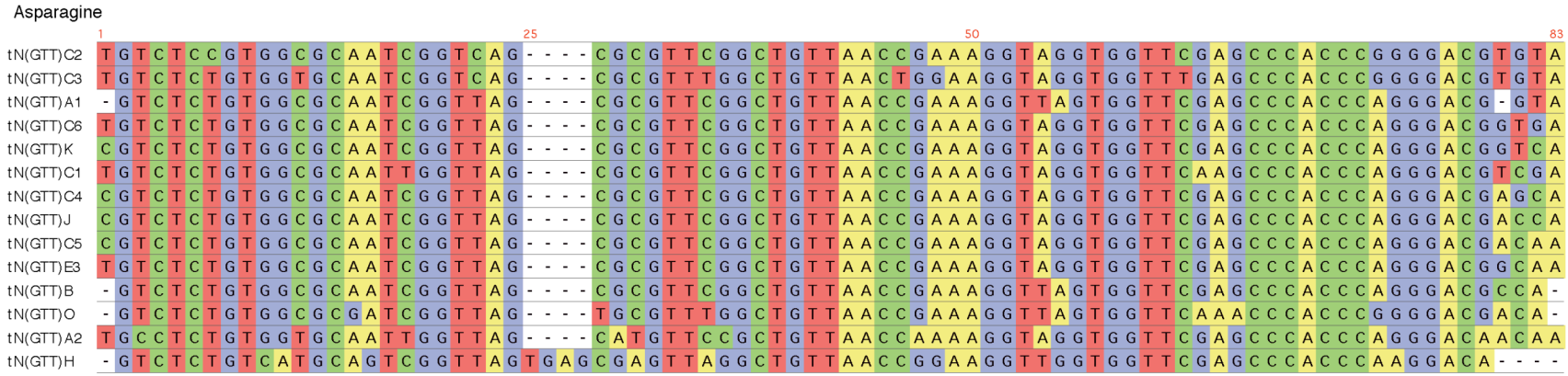


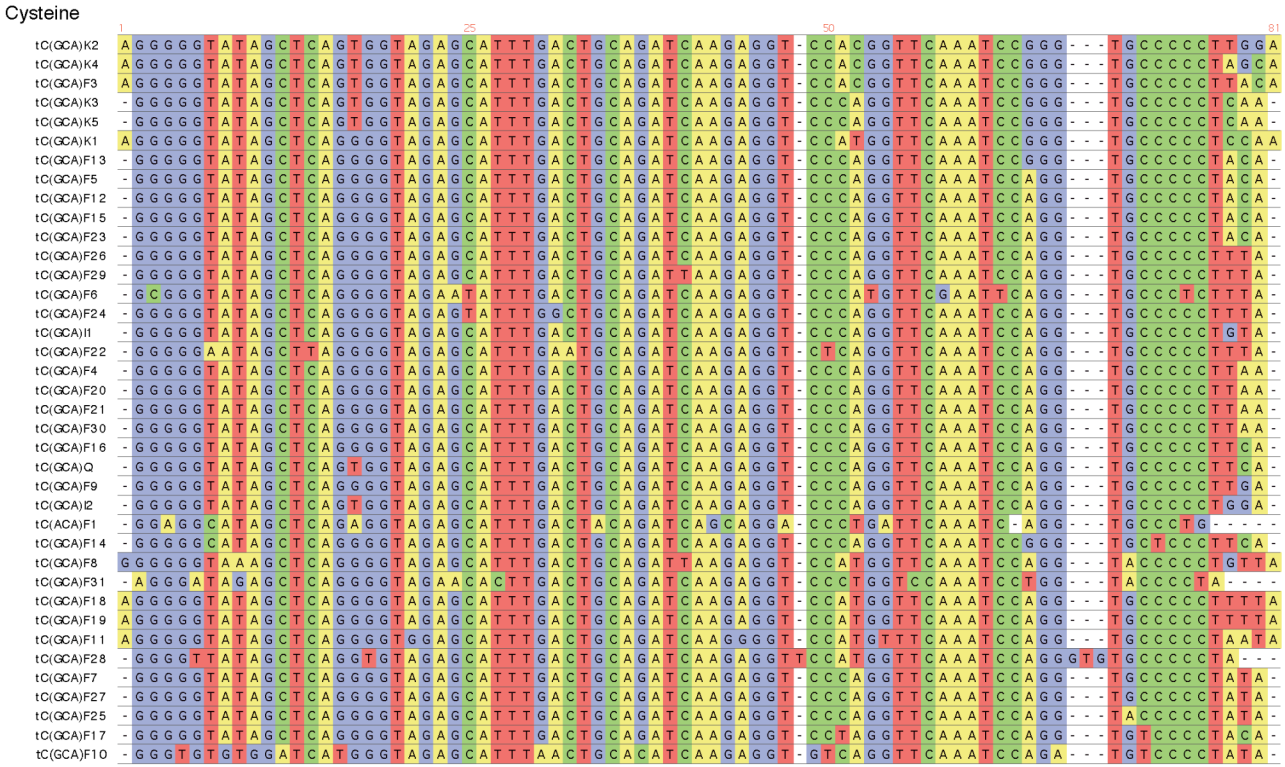

Alanine2
1
25
50
79
tA(AGC)B4
tA(AGC)I6
tA(AGC)J1
tA(AGC)X
tV(AAC)H
tA(AGC)K4
tA(AGC)G4
tA(AGC)L2
tA(AGC)F4
tA(AGC)H1
tA(AGC)J4
tA(AGC)A3
tA(AGC)E7
tA(AGC)O3
tA(AGC)P
tA(AGC)F5
tA(AGC)J5
tA(AGC)O1
tA(AGC)O2
tG(ACC)X1
tA(AGC)I2
tA(AGC)G3
tA(AGC)S4
tA(AGC)L3
tA(AGC)Q1

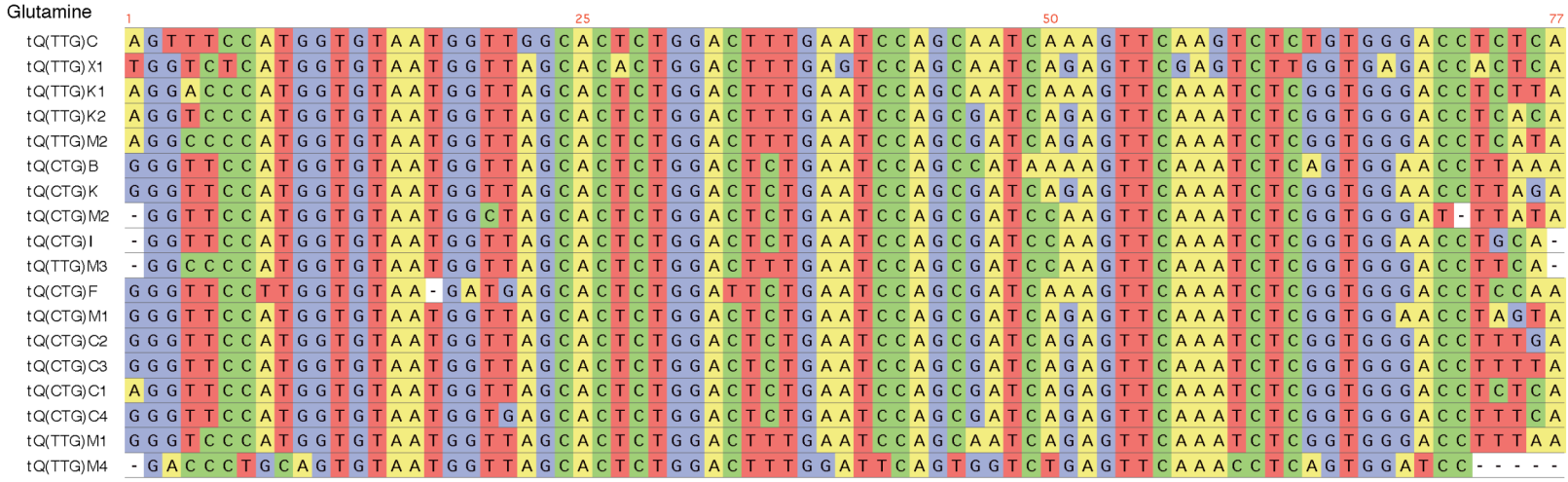

Alanine3
1
25
50
78
tA(AGC)M7
tA(AGC)M8
tA(AGC)C1
tA(AGC)E3
tA(AGC)M10
tA(AGC)M9
tA(AGC)M11
tA(AGC)S1

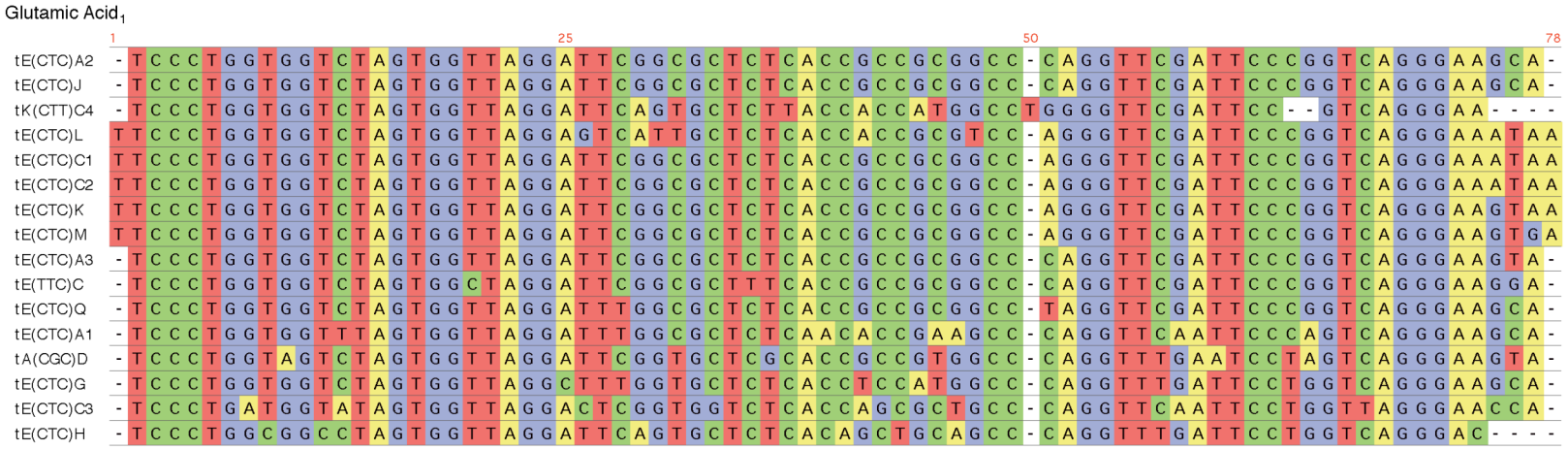

Arginine1
1
25
50
78
tR(ACG)I CGGGCCAGTGGCGCAATGGATAACGCGTCTGACTACGGATCAGAAGATTATAGGTTCGACTCCTGGCTGGCTCGACGA
tR(ACG)M1 CGGGCCAGTGGCGCAATGGATAACGCGTCTGACTACGGATCAGAAGATTATAGGTTCGACTCCTGGCTGGCTCGACCA
tR(ACG)N CGGGCCAGTGGCGCAATGGATAACGCGTCTGACTACGGATCAGAAGATTACAGGTTCGACTCCTGGCTGGCTCGTTGA
tR(ACG)M2 -GGGCCAGTGGCGCAATGGATAACGCGTCTGACTACGGATCAGAAGATTCAAGGTTCGACTCCTGGCTGGCTCGGAA-
tR(ACG)M3 TGGGCCAGTGGCGCAATGGATAACGCGTCTGACTACGGATCAGAAGATTACAGGTTCGACTCCTGGCTGGCTCGGTAA
tR(ACG)C CGGGCCAGTGGCGCAATGGATAACGCGTCTGACTACGGATCAGAAGATTATAGGTTCGACTCCTACCTGGCTCGAGCA
tR(TCG)G CGGCCGCGTGGCCTAATGGATAAGGCGTCTGACTTCGGATCAGAAGATTACAGGTTCGAGTCCTGCCGCGGTCGACTA
tR(TCG)M1 CGACCACGTGGCCTAATGGATAAGGCGTCTGACTTCGGATCAGAAGATTAAGGGTTCGAAATCCCTTCGTGGTTGGTAA
tR(TCG)M3 TGACCACGTGGCCTAACGGATAAGGCGTCTGACTTCGGATCAGAAGATTAAGGGTTCGAAATCCCTTCGTGGTTACTTA
tR(TCG)M2 -GACCACGTGGCCTAATGGATAAGGCGTCTGACTTCGGATCAGAAGATTGAGGGTTCGAAATCCCTTCGTGGTTGAGA-
tR(TCG)K -GACCGCGTGGCCTAATGGATAAGGCGTCTGACTTCGGATCAGAAGATTGAGGGTTCGAGTCCCTTCGTGGTCGAAAA-
tR(CCG)M TGGCCGCGTGGCCTAATGGATAAGGCGTCTGATTCCGGATCAGAAGATTAAGGGTTCGAGTCCCTTCGTGGTCGGTGA
tR(CCG)Q -GGCCGCGTGGCCTAATGGATAAGGCGTCTGATTCCGGATCAGAAGATTGAGGGTTCGAGTCCCTTCGTGGTCGCTA-
tR(CCG)K TGACCCAGTGGCCTAATGGATAAGGCATCAGCCTCCGGAGCTGGGGATTATGGGTTCGAGTCCCATCTGGGTCGCATA

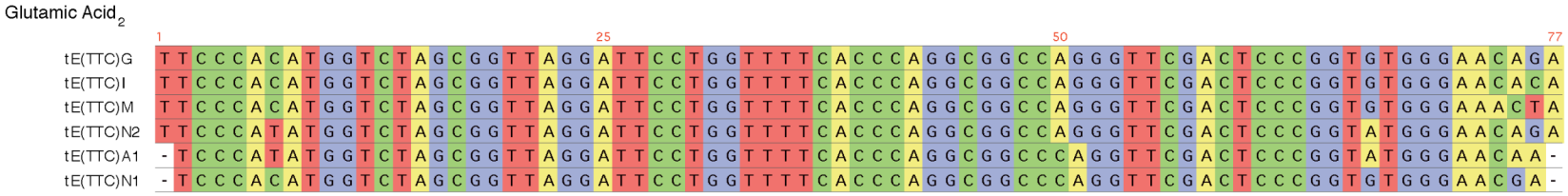

Arginine2
1
25
50
75
tR(CCT)F
tR(CCT)X2
tR(CCT)K2
tR(CCT)K1
tR(CCT)Q
tR(TCT)I
tR(TCT)M
tR(TCT)K
tR(TCT)S
tR(TCT)C
tR(TCT)A

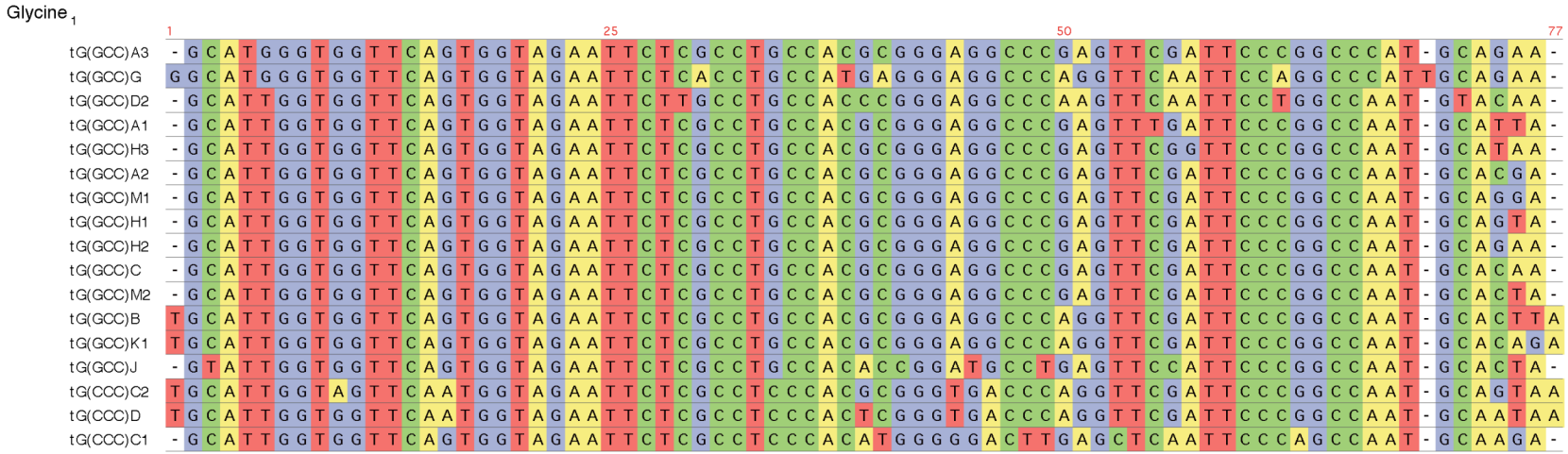

Asparagine
1
25
50
83
tN(GTT)C2
tN(GTT)C3
tN(GTT)A1
tN(GTT)C6
tN(GTT)K
tN(GTT)C1
tN(GTT)C4
tN(GTT)J
tN(GTT)C5
tN(GTT)E3
tN(GTT)B
tN(GTT)O
tN(GTT)A2
tN(GTT)H

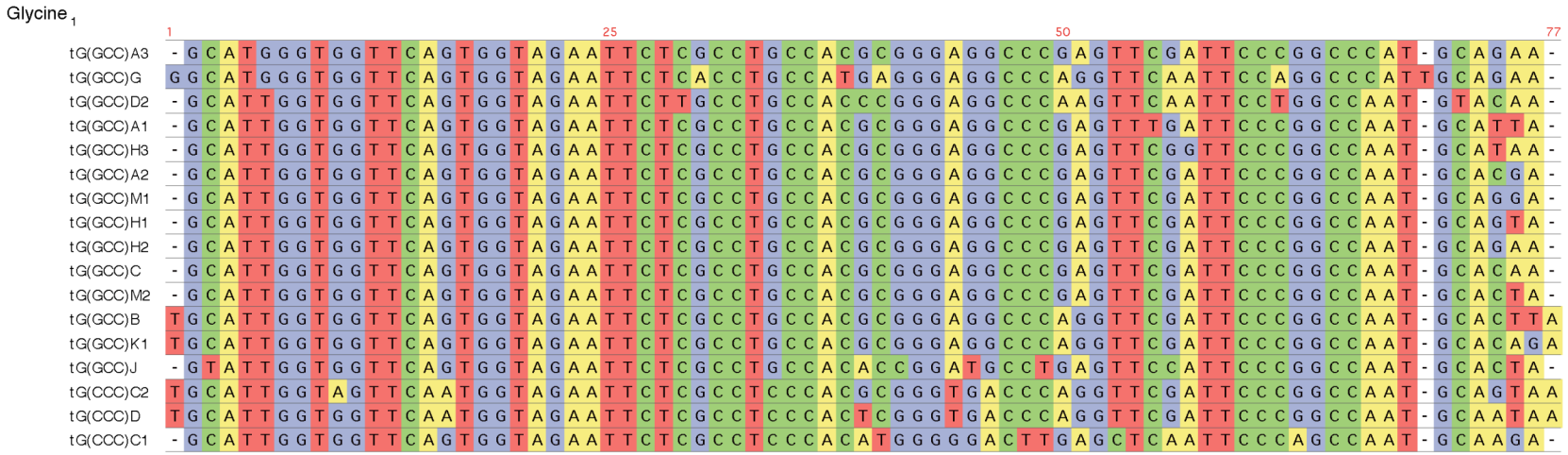

Cysteine
1
25
50
81
tC(GCA)K2
tC(GCA)K4
tC(GCA)F3
tC(GCA)K3
tC(GCA)K5
tC(GCA)K1
tC(GCA)F13
tC(GCA)F5
tC(GCA)F12
tC(GCA)F15
tC(GCA)F23
tC(GCA)F26
tC(GCA)F29
tC(GCA)F6
tC(GCA)F24
tC(GCA)I1
tC(GCA)F22
tC(GCA)F4
tC(GCA)F20
tC(GCA)F21
tC(GCA)F30
tC(GCA)F16
tC(GCA)Q
tC(GCA)F9
tC(GCA)I2
tC(ACA)F1
tC(GCA)F14
tC(GCA)F8
tC(GCA)F31
tC(GCA)F18
tC(GCA)F19
tC(GCA)F11
tC(GCA)F28
tC(GCA)F7
tC(GCA)F27
tC(GCA)F25
tC(GCA)F17
tC(GCA)F10


Glutamine
1
25
50
77
tQ(TTG)C AGTTTCCATGGTGTAATGGTTGGCACTCTGGACTTTGAATCCAGCAATCAAAGTTCAAGTCTCTGTGGGACCTCTCA
tQ(TTG)X1 TGGTCTCATGGTGTAATGGTTAGCACACTGGACTTTGAGTCCAGCAATCAGAGTTCGAGTCTTGGTGAGACCACTCA
tQ(TTG)K1 AGGACCCATGGTGTAATGGTTAGCACTCTGGACTTTGAATCCAGCAATCAAAGTTCAAATCTCGGTGGGACCTCTTA
tQ(TTG)K2 AGGTCCCATGGTGTAATGGTTAGCACTCTGGACTTTGAATCCAGCGATCAGAGTTCAAATCTCGGTGGGACCTCACA
tQ(TTG)M2 AGGCCCCATGGTGTAATGGTTAGCACTCTGGACTTTGAATCCAGCGATCAGAGTTCAAATCTCGGTGGGACCTCATA
tQ(CTG)B GGGTTCCATGGTGTAATGGTTAGCACTCTGGACTCTGAATCCAGCCATAAAAGTTCAAATCTCAGTGGAACCTTAAA
tQ(CTG)K GGGTTCCATGGTGTAATGGTTAGCACTCTGGACTCTGAATCCAGCGATCAGAGTTCAAATCTCGGTGGAACCTTAGA
tQ(CTG)M2 -GGTTCCATGGTGTAATGGCTAGCACTCTGGACTCTGAATCCAGCGATCCAAGTTCAAATCTCGGTGGGAT-TTATA
tQ(CTG)I -GGTTCCATGGTGTAATGGTTAGCACTCTGGACTCTGAATCCAGCGATCCAAGTTCAAATCTCGGTGGAACCTGCA-
tQ(TTG)M3 -GGCCCCATGGTGTAATGGTTAGCACTCTGGACTTTGAATCCAGCGATCCAAGTTCAAATCTCGGTGGGACCTTCA-
tQ(CTG)F GGGTTCCTTGGTGTAA-GATGAGCACTCTGGATTCTGAATCCAGCGATCAAAGTTCAAATCTCGGTGGGACCTCCAA
tQ(CTG)M1 GGGTTCCATGGTGTAATGGTTAGCACTCTGGACTCTGAATCCAGCGATCAGAGTTCAAATCTCGGTGGAACCTAGTA
tQ(CTG)C2 GGGTTCCATGGTGTAATGGTTAGCACTCTGGACTCTGAATCCAGCGATCAGAGTTCAAATCTCGGTGGGACCTTTGA
tQ(CTG)C3 GGGTTCCATGGTGTAATGGTTAGCACTCTGGACTCTGAATCCAGCGATCAGAGTTCAAATCTCGGTGGGACCTTTTA
tQ(CTG)C1 AGGTTCCATGGTGTAATGGTTAGCACTCTGGACTCTGAATCCAGCGATCAGAGTTCAAATCTCGGTGGGACCTCTCA
tQ(CTG)C4 GGGTTCCATGGTGTAATGGTGAGCACTCTGGACTCTGAATCCAGCGATCAGAGTTCAAATCTCGGTGGGACCTTTCA
tQ(TTG)M1 GGGTCCCATGGTGTAATGGTTAGCACTCTGGACTTTGAATCCAGCAATCAGAGTTCAAATCTCGGTGGGACCTTTAA
tQ(TTG)M4 -GACCCTGCAGTGTAATGGTTAGCACTCTGGACTTTGGATTCAGTGGTCTGAGTTCAAACCTCAGTGGATCC-----

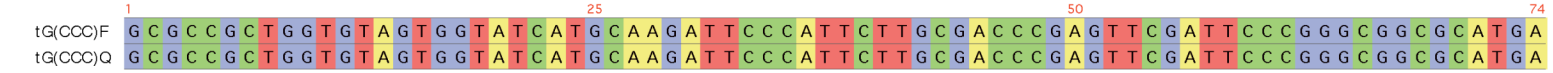

Glutamic Acid1
1
25
50
78
tE(CTC)A2
tE(CTC)J
tK(CTT)C4
tE(CTC)L
tE(CTC)C1
tE(CTC)C2
tE(CTC)K
tE(CTC)M
tE(CTC)A3
tE(TTC)C
tE(CTC)Q
tE(CTC)A1
tA(CGC)D
tE(CTC)G
tE(CTC)C3
tE(CTC)H

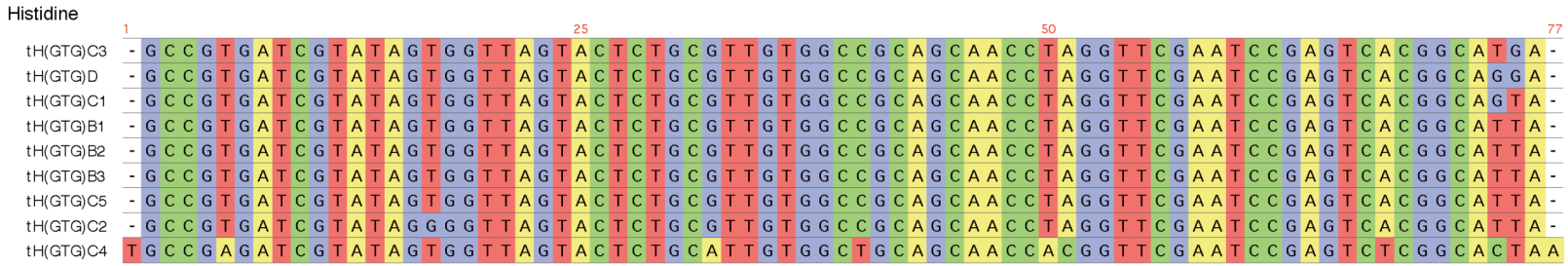

Glutamic Acid2
1
25
50
77
tE(TTC)G
tE(TTC)I
tE(TTC)M
tE(TTC)N2
tE(TTC)A1
tE(TTC)N1

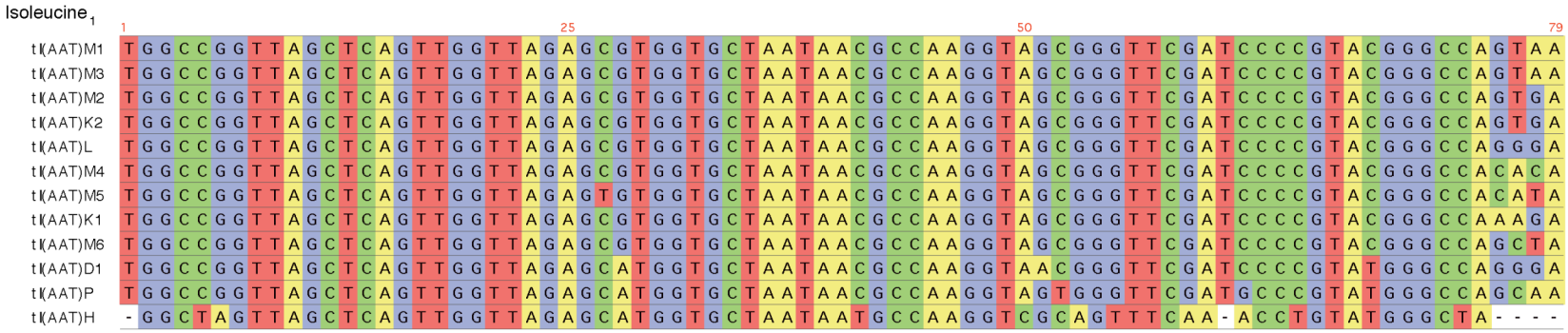

Glycine 1
1
25
50
77
tG(GCC)A3
tG(GCC)G
tG(GCC)D2
tG(GCC)A1
tG(GCC)H3
tG(GCC)A2
tG(GCC)M1
tG(GCC)H1
tG(GCC)H2
tG(GCC)C
tG(GCC)M2
tG(GCC)B
tG(GCC)K1
tG(GCC)J
tG(CCC)C2
tG(CCC)D
tG(CCC)C1


Glycine 1
1
25
50
77
tG(GCC)A3 - G C A T G G G T G G T T C A G T G G T A G A A T T C T C G C C T G C C A C G C G G G A G G C C C G A G T T C G A T T C C C G G C C C A T - G C A G A A -
tG(GCC)G G G C A T G G G T G G T T C A G T G G T A G A A T T C T C A C C T G C C A T G A G G G A G G C C C A G G T T C A A T T C C A G G C C C A T T G C A G A A -
tG(GCC)D2 - G C A T T G G T G G T T C A G T G G T A G A A T T C T T G C C T G C C A C C C G G G A G G C C C A A G T T C A A T T C C T G G C C A A T - G T A C A A -
tG(GCC)A1 - G C A T T G G T G G T T C A G T G G T A G A A T T C T C G C C T G C C A C G C G G G A G G C C C G A G T T T G A T T C C C G G C C A A T - G C A T T A -
tG(GCC)H3 - G C A T T G G T G G T T C A G T G G T A G A A T T C T C G C C T G C C A C G C G G G A G G C C C G A G T T C G G T T C C C G G C C A A T - G C A T A A -
tG(GCC)A2 - G C A T T G G T G G T T C A G T G G T A G A A T T C T C G C C T G C C A C G C G G G A G G C C C G A G T T C G A T T C C C G G C C A A T - G C A C G A -
tG(GCC)M1 - G C A T T G G T G G T T C A G T G G T A G A A T T C T C G C C T G C C A C G C G G G A G G C C C G A G T T C G A T T C C C G G C C A A T - G C A G G A -
tG(GCC)H1 - G C A T T G G T G G T T C A G T G G T A G A A T T C T C G C C T G C C A C G C G G G A G G C C C G A G T T C G A T T C C C G G C C A A T - G C A G T A -
tG(GCC)H2 - G C A T T G G T G G T T C A G T G G T A G A A T T C T C G C C T G C C A C G C G G G A G G C C C G A G T T C G A T T C C C G G C C A A T - G C A G A A -
tG(GCC)C - G C A T T G G T G G T T C A G T G G T A G A A T T C T C G C C T G C C A C G C G G G A G G C C C G A G T T C G A T T C C C G G C C A A T - G C A C A A -
tG(GCC)M2 - G C A T T G G T G G T T C A G T G G T A G A A T T C T C G C C T G C C A C G C G G G A G G C C C G A G T T C G A T T C C C G G C C A A T - G C A C T A -
tG(GCC)B T G C A T T G G T G G T T C A G T G G T A G A A T T C T C G C C T G C C A C G C G G G A G G C C C A G G T T C G A T T C C C G G C C A A T - G C A C T T A
tG(GCC)K1 T G C A T T G G T G G T T C A G T G G T A G A A T T C T C G C C T G C C A C G C G G G A G G C C C A G G T T C G A T T C C C G G C C A A T - G C A C A G A
tG(GCC)J - G T A T T G G T G G T T C A G T G G T A G A A T T C T C G C C T G C C A C A C C G G G A T G C C T G A G T T C C A T T C C C G G C C A A T - G C A C T A -
tG(CCC)C2 T G C A T T G G T A G T T C A A T G G T A G A A T T C T C G C C T C C C A C G C G G G T G A C C C A G G T T C G A T T C C C G G C C A A T - G C A G T A A
tG(CCC)D T G C A T T G G T G G T T C A A T G G T A G A A T T C T C G C C T C C C A C T C G G G T G A C C C A G G T T C G A T T C C C G G C C A A T - G C A A T A A
tG(CCC)C1 - G C A T T G G T G G T T C A G T G G T A G A A T T C T C G C C T C C C A C A T G G G G G A C T T G A G C T C A A T T C C C A G C C A A T - G C A A G A -

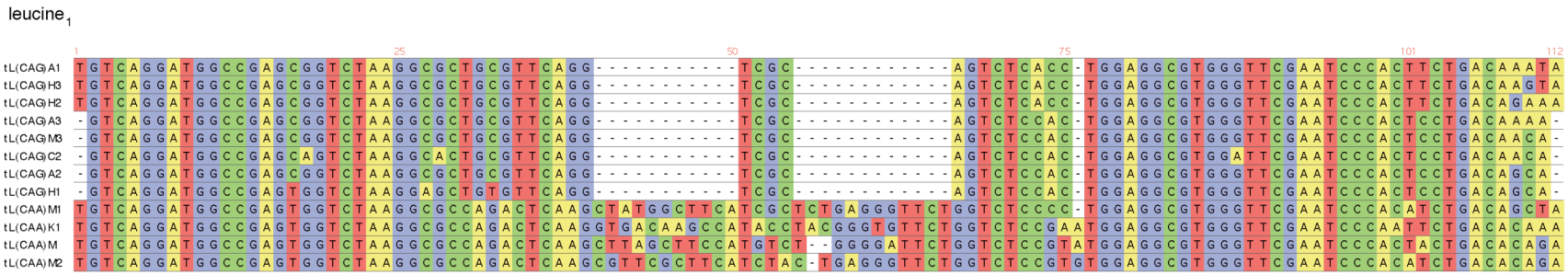
Glycine$_3$

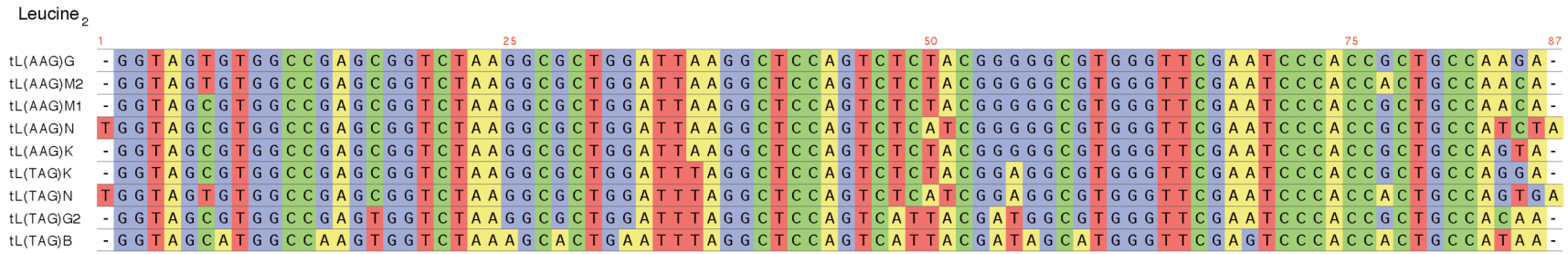

1
25
50
74
tG(CCC)F G C G C C G C T G G T G T A G T G G T A T C A T G C A A G A T T C C C A T T C T T G C G A C C C G A G T T C G A T T C C C G G G C G G C G C A T G A
tG(CCC)Q G C G C C G C T G G T G T A G T G G T A T C A T G C A A G A T T C C C A T T C T T G C G A C C C G A G T T C G A T T C C C G G G C G G C G C A T G A

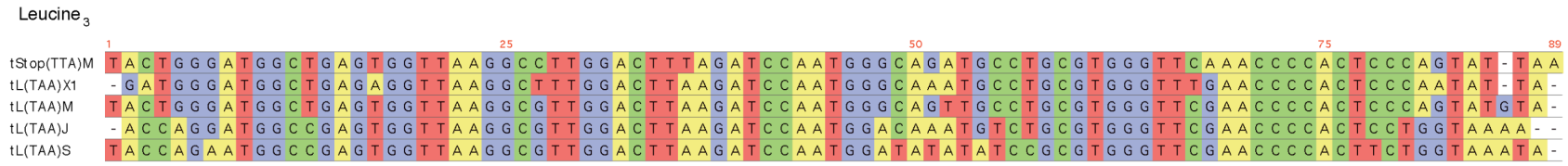

Histidine
1
25
50
77
tH(GTG)C3 -GCCGTGATCGTATAGTGGTTAGTACTCTGCGTTGTGGCCGCAGCAACCTAGGTTCGAATCCGAGTCACGGCATGA-
tH(GTG)D -GCCGTGATCGTATAGTGGTTAGTACTCTGCGTTGTGGCCGCAGCAACCTAGGTTCGAATCCGAGTCACGGCAGGA-
tH(GTG)C1 -GCCGTGATCGTATAGTGGTTAGTACTCTGCGTTGTGGCCGCAGCAACCTAGGTTCGAATCCGAGTCACGGCAGTA-
tH(GTG)B1 -GCCGTGATCGTATAGTGGTTAGTACTCTGCGTTGTGGCCGCAGCAACCTAGGTTCGAATCCGAGTCACGGCATTA-
tH(GTG)B2 -GCCGTGATCGTATAGTGGTTAGTACTCTGCGTTGTGGCCGCAGCAACCTAGGTTCGAATCCGAGTCACGGCATTA-
tH(GTG)B3 -GCCGTGATCGTATAGTGGTTAGTACTCTGCGTTGTGGCCGCAGCAACCTAGGTTCGAATCCGAGTCACGGCATTA-
tH(GTG)C5 -GCCGTGATCGTATAGTGGTTAGTACTCTGCGTTGTGGCCGCAGCAACCTAGGTTCGAATCCGAGTCACGGCATTA-
tH(GTG)C2 -GCCGTGATCGTATAGGGGTTAGTACTCTGCGTTGTGGCCGCAGCAACCTAGGTTCGAATCCGAGTCACGGCATTA-
tH(GTG)C4 TGCCGAGATCGTATAGTGGTTAGTACTCTGCATTGTGGCTGCAGCAACCACGGTTCGAATCCGAGTCTCGGCACTAA

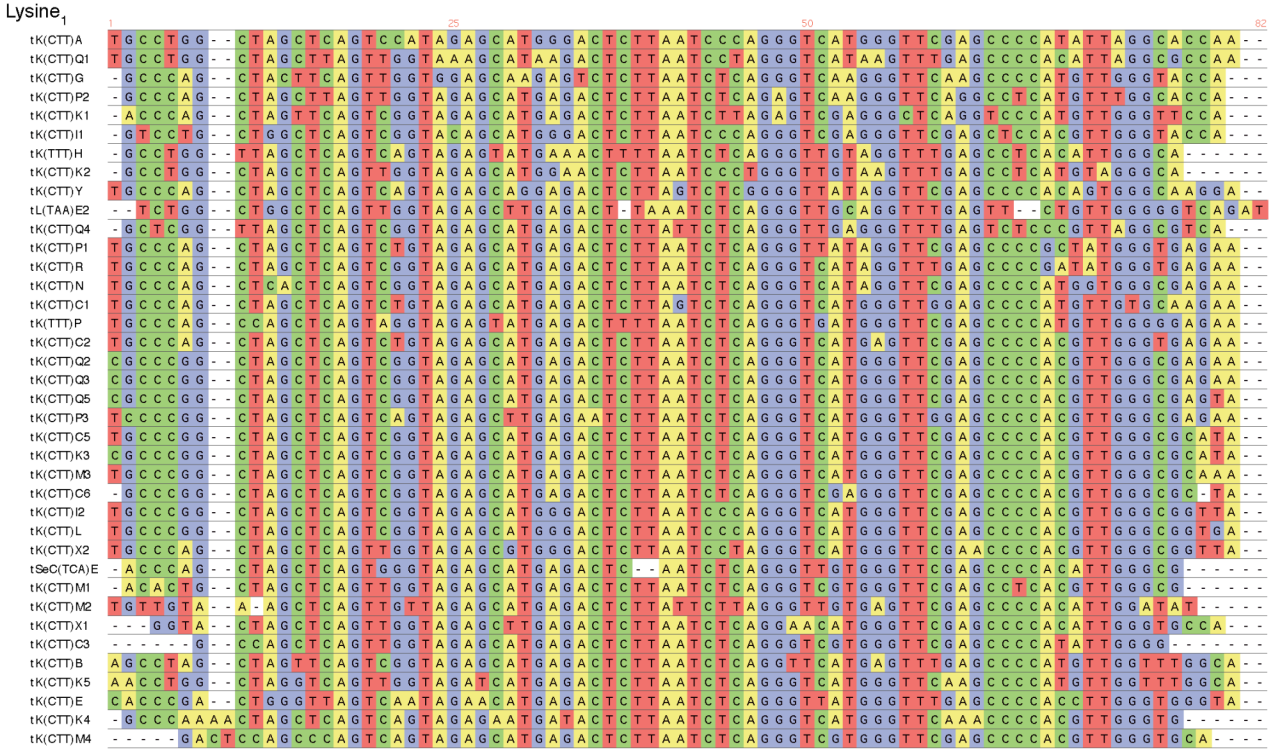

Isoleucine 1
1
25
50
79
tI(AAT)M1 TGGCCGGTTAGCTCAGTTGGTTAGAGCGTGGTGCTAATAACGCCAAGGTAGCGGGTTCGATCCCCGTACGGGCCAGTAA
tI(AAT)M3 TGGCCGGTTAGCTCAGTTGGTTAGAGCGTGGTGCTAATAACGCCAAGGTAGCGGGTTCGATCCCCGTACGGGCCAGTAA
tI(AAT)M2 TGGCCGGTTAGCTCAGTTGGTTAGAGCGTGGTGCTAATAACGCCAAGGTAGCGGGTTCGATCCCCGTACGGGCCAGTGA
tI(AAT)K2 TGGCCGGTTAGCTCAGTTGGTTAGAGCGTGGTGCTAATAACGCCAAGGTAGCGGGTTCGATCCCCGTACGGGCCAGTGA
tI(AAT)L TGGCCGGTTAGCTCAGTTGGTTAGAGCGTGGTGCTAATAACGCCAAGGTAGCGGGTTCGATCCCCGTACGGGCCAGGGA
tI(AAT)M4 TGGCCGGTTAGCTCAGTTGGTTAGAGCGTGGTGCTAATAACGCCAAGGTAGCGGGTTCGATCCCCGTACGGGCCACACA
tI(AAT)M5 TGGCCGGTTAGCTCAGTTGGTTAGAGTGTGGTGCTAATAACGCCAAGGTAGCGGGTTCGATCCCCGTACGGGCCACATA
tI(AAT)K1 TGGCCGGTTAGCTCAGTTGGTTAGAGCGTGGTGCTAATAACGCCAAGGTAGCGGGTTCGATCCCCGTACGGGCCAAAGA
tI(AAT)M6 TGGCCGGTTAGCTCAGTTGGTTAGAGCGTGGTGCTAATAACGCCAAGGTAGCGGGTTCGATCCCCGTACGGGCCAGCTA
tI(AAT)D1 TGGCCGGTTAGCTCAGTTGGTTAGAGCATGGTGCTAATAACGCCAAGGTAACGGGTTCGATCCCCGTATGGGCCAGGGA
tI(AAT)P TGGCCGGTTAGCTCAGTTGGTTAGAGCATGGTGCTAATAACGCCAAGGTAGTGGGTTCGATGCCCGTATGGGCCAGCAA
tI(AAT)H -GGCTAGTTAGCTCAGTTGGTTAGAGCATGGTGCTAATAATGCCAAGGTCGCAGTTTCAA-ACCTGTATGGGCTA----

Isoleucine$_2$

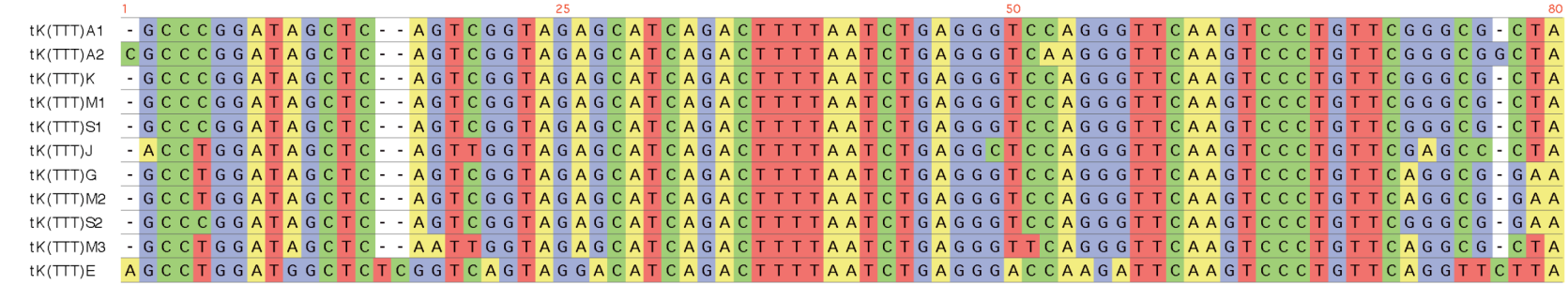

leucine1
1
25
50
75
101
112
tL(CAG)A1
tL(CAG)H3
tL(CAG)H2
tL(CAG)A3
tL(CAG)M3
tL(CAG)C2
tL(CAG)A2
tL(CAG)H1
tL(CAA)M1
tL(CAA)K1
tL(CAA)M
tL(CAA)M2

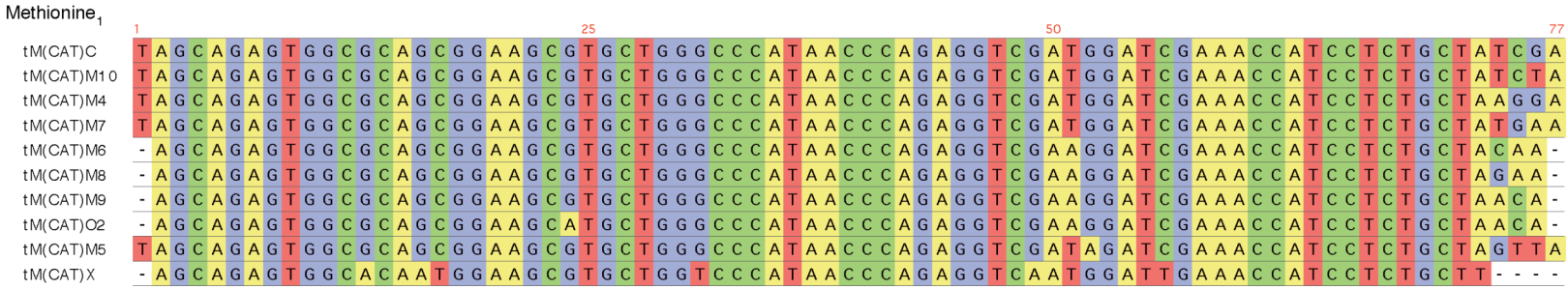

Leucine 2
1
25
50
75
87
tL(AAG)G
tL(AAG)M2
tL(AAG)M1
tL(AAG)N
tL(AAG)K
tL(TAG)K
tL(TAG)N
tL(TAG)G2
tL(TAG)B

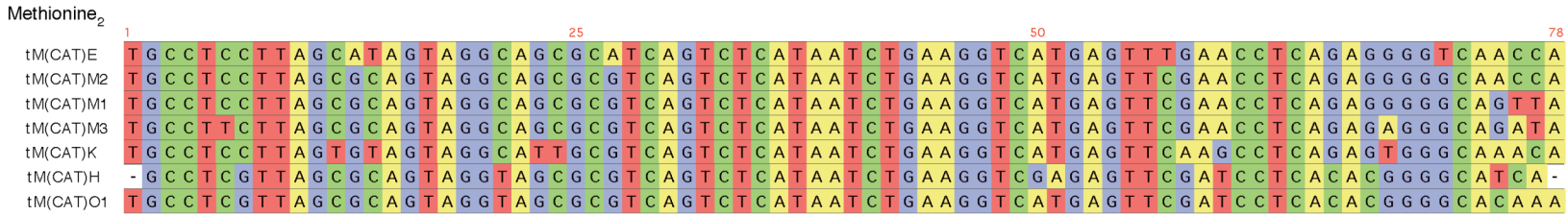

Leucine 3
1
25
50
75
89
tStop(TTA)M TACTGGGATGGCTGAGTGGTTAAGGCCTTGGACTTTAGATCCAATGGGCAGATGCCTGCGTGGGTTCAAACCCCACTCCCAGTAT-TAA
tL(TAA)X1 -GATGGGATGGCTGAGAGGTTAAGGCTTTGGACTTAAGATCCAATGGGCAAAATGCCTGCGTGGGTTTGAACCCCACTCCCAATAT-TA-
tL(TAA)M TACTGGGATGGCTGAGTGGTTAAGGCGTTGGACTTAAGATCCAATGGGCAGTTGCCTGCGTGGGTTCGAACCCCACTCCCAGTATGTA-
tL(TAA)J -ACCAGGATGGCCGAGTGGTTAAGGCGTTGGACTTAAGATCCAATGGACAAAATGTCTGCGTGGGTTCGAACCCCACTCCTGGTAAAAA--
tL(TAA)S TACCAGAATGGCCGAGTGGTTAAGGCGTTGGACTTAAGATCCAATGGATATATATCCGCGTGGGTTCGAACCCCACTTCTGGTAAATA-

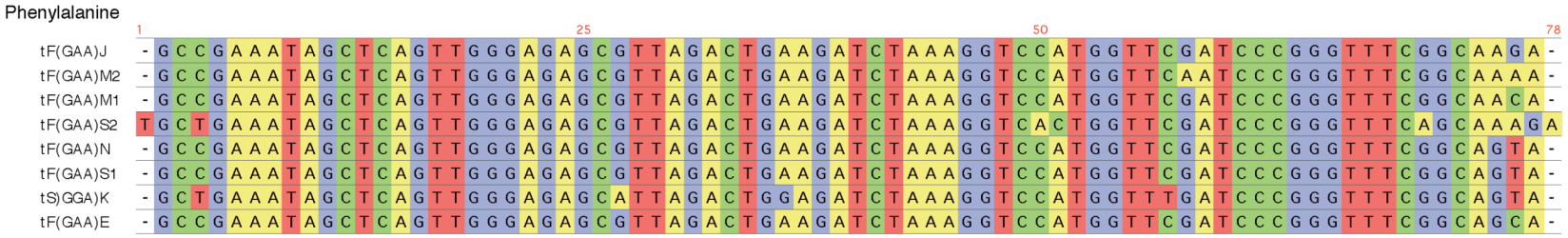

Lysine1
1
25
50
82
tK(CTT)A
tK(CTT)Q1
tK(CTT)G
tK(CTT)P2
tK(CTT)K1
tK(CTT)I1
tK(TTT)H
tK(CTT)K2
tK(CTT)Y
tL(TAA)E2
tK(CTT)Q4
tK(CTT)P1
tK(CTT)R
tK(CTT)N
tK(CTT)C1
tK(TTT)P
tK(CTT)C2
tK(CTT)Q2
tK(CTT)Q3
tK(CTT)Q5
tK(CTT)P3
tK(CTT)C5
tK(CTT)K3
tK(CTT)M3
tK(CTT)C6
tK(CTT)I2
tK(CTT)L
tK(CTT)X2
tSeC(TCA)E
tK(CTT)M1
tK(CTT)M2
tK(CTT)X1
tK(CTT)C3
tK(CTT)B
tK(CTT)K5
tK(CTT)E
tK(CTT)K4
tK(CTT)M4

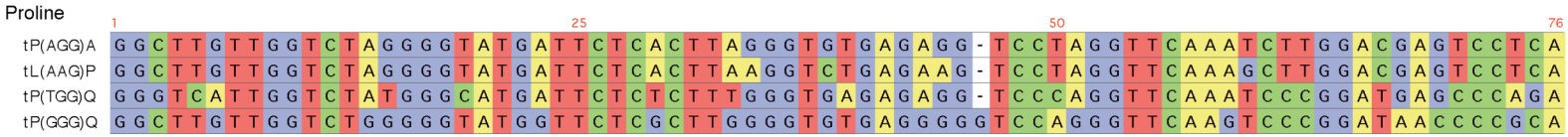
Lysine$_2$

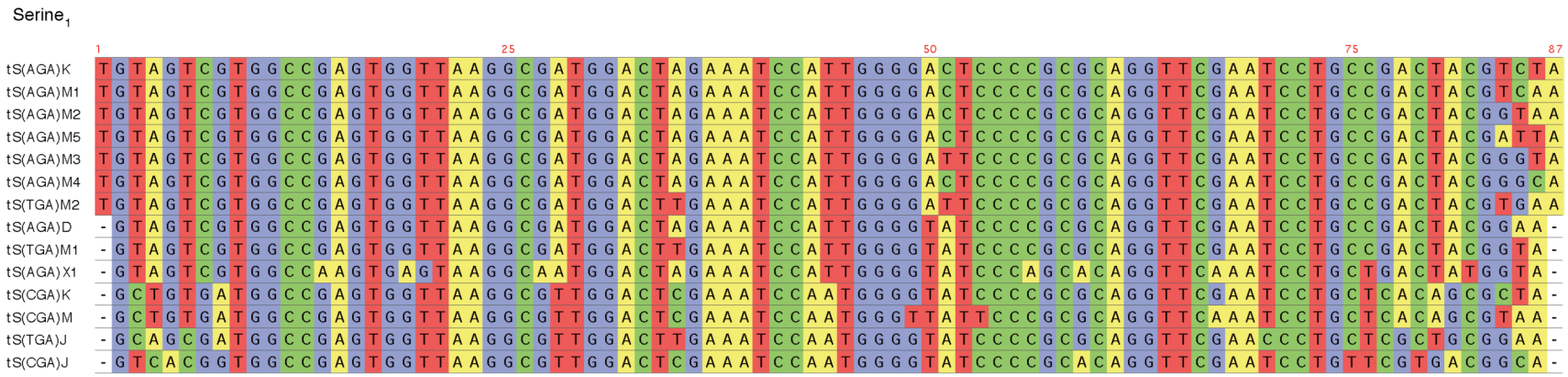

1
25
50
80
tK(TTT)A1 - G C C C G G A T A G C T C - - A G T C G G T A G A G C A T C A G A C T T T T A A T C T G A G G G T C C A G G G T T C A A G T C C C T G T T C G G G C G - C T A
tK(TTT)A2 C G C C C G G A T A G C T C - - A G T C G G T A G A G C A T C A G A C T T T T A A T C T G A G G G T C A A G G G T T C A A G T C C C T G T T C G G G C G G C T A
tK(TTT)K - G C C C G G A T A G C T C - - A G T C G G T A G A G C A T C A G A C T T T T A A T C T G A G G G T C C A G G G T T C A A G T C C C T G T T C G G G C G - C T A
tK(TTT)M1 - G C C C G G A T A G C T C - - A G T C G G T A G A G C A T C A G A C T T T T A A T C T G A G G G T C C A G G G T T C A A G T C C C T G T T C G G G C G - C T A
tK(TTT)S1 - G C C C G G A T A G C T C - - A G T C G G T A G A G C A T C A G A C T T T T A A T C T G A G G G T C C A G G G T T C A A G T C C C T G T T C G G G C G - C T A
tK(TTT)J - A C C T G G A T A G C T C - - A G T T G G T A G A G C A T C A G A C T T T T A A T C T G A G G C T C C A G G G T T C A A G T C C C T G T T C G A G C C - C T A
tK(TTT)G - G C C T G G A T A G C T C - - A G T C G G T A G A G C A T C A G A C T T T T A A T C T G A G G G T C C A G G G T T C A A G T C C C T G T T C A G G C G - G A A
tK(TTT)M2 - G C C T G G A T A G C T C - - A G T C G G T A G A G C A T C A G A C T T T T A A T C T G A G G G T C C A G G G T T C A A G T C C C T G T T C A G G C G - G A A
tK(TTT)S2 - G C C C G G A T A G C T C - - A G T C G G T A G A G C A T C A G A C T T T T A A T C T G A G G G T C C A G G G T T C A A G T C C C T G T T C G G G C G - G A A
tK(TTT)M3 - G C C T G G A T A G C T C - - A A T T G G T A G A G C A T C A G A C T T T T A A T C T G A G G G T T C A G G G T T C A A G T C C C T G T T C A G G C G - C T A
tK(TTT)E A G C C T G G A T G G C T C T C G G T C A G T A G G A C A T C A G A C T T T T A A T C T G A G G G A C C A A G A T T C A A G T C C C T G T T C A G G T T C T T A

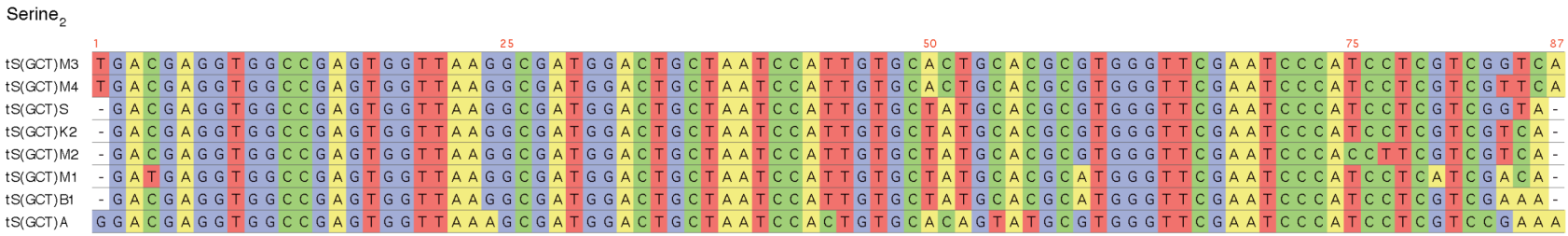

Methionine1
1
25
50
77
tM(CAT)C
tM(CAT)M10
tM(CAT)M4
tM(CAT)M7
tM(CAT)M6
tM(CAT)M8
tM(CAT)M9
tM(CAT)O2
tM(CAT)M5
tM(CAT)X


Methionine2
1
25
50
78
tM(CAT)E
tM(CAT)M2
tM(CAT)M1
tM(CAT)M3
tM(CAT)K
tM(CAT)H
tM(CAT)O1

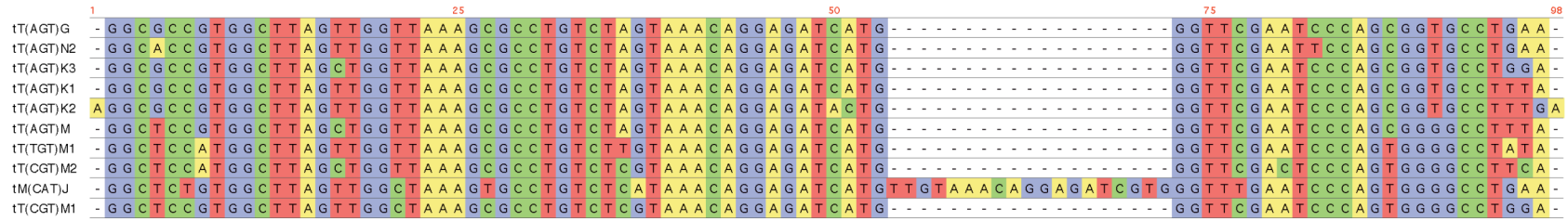

Phenylalanine
1
25
50
78
tF(GAA)J
tF(GAA)M2
tF(GAA)M1
tF(GAA)S2
tF(GAA)N
tF(GAA)S1
tS)GGA)K
tF(GAA)E


Proline
1
25
50
76
tP(AGG)A
tL(AAG)P
tP(TGG)Q
tP(GGG)Q

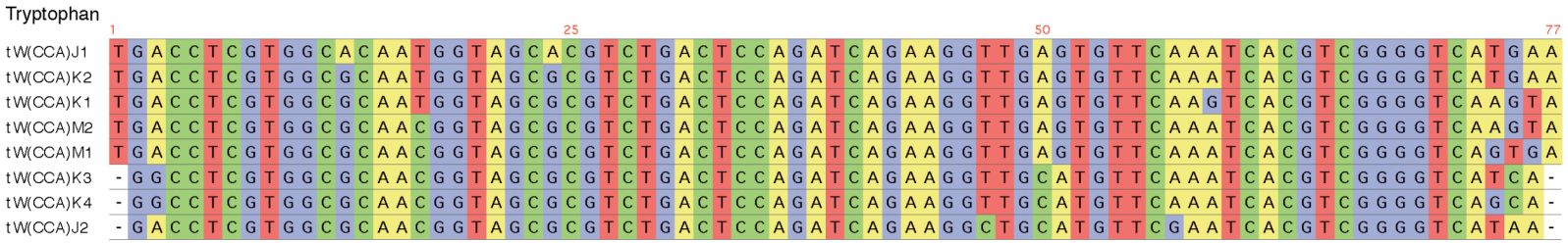

Serine1
1
25
50
75
87
tS(AGA)K TGTAGTCGTGGCCGAGTGGTTAAGGCGATGGACTAGAAATCCATTGGGGACTCCCCGCGCAGGTTCGAATCCTGCCGACTACGTCTA
tS(AGA)M1 TGTAGTCGTGGCCGAGTGGTTAAGGCGATGGACTAGAAATCCATTGGGGACTCCCCGCGCAGGTTCGAATCCTGCCGACTACGTCAA
tS(AGA)M2 TGTAGTCGTGGCCGAGTGGTTAAGGCGATGGACTAGAAATCCATTGGGGACTCCCCGCGCAGGTTCGAATCCTGCCGACTACGGTAA
tS(AGA)M5 TGTAGTCGTGGCCGAGTGGTTAAGGCGATGGACTAGAAATCCATTGGGGACTCCCCGCGCAGGTTCGAATCCTGCCGACTACGATTA
tS(AGA)M3 TGTAGTCGTGGCCGAGTGGTTAAGGCGATGGACTAGAAATCCATTGGGGATTCCCCGCGCAGGTTCGAATCCTGCCGACTACGGGTA
tS(AGA)M4 TGTAGTCGTGGCCGAGTGGTTAAGGCGATGGACTAGAAATCCATTGGGGACTCCCCGCGCAGGTTCGAATCCTGCCGACTACGGGCA
tS(TGA)M2 TGTAGTCGTGGCCGAGTGGTTAAGGCGATGGACTTGAAATCCATTGGGGATTCCCCGCGCAGGTTCGAATCCTGCCGACTACGTGAA
tS(AGA)D -GTAGTCGTGGCCGAGTGGTTAAGGCGATGGACTAGAAATCCATTGGGGTATCCCCGCGCAGGTTCGAATCCTGCCGACTACGGAA-
tS(TGA)M1 -GTAGTCGTGGCCGAGTGGTTAAGGCGATGGACTTGAAATCCATTGGGGTATCCCCGCGCAGGTTCGAATCCTGCCGACTACGGTA-
tS(AGA)X1 -GTAGTCGTGGCCAAGTGAGTAAGGCAATGGACTAGAAATCCATTGGGGTATCCCAGCACAGGTTCAAATCCTGCTGACTATGGTA-
tS(CGA)K -GCTGTGATGGCCGAGTGGTTAAGGCGTTGGACTCGAAATCCAATGGGGTATCCCCGCGCAGGTTCGAATCCTGCTCACAGCGCTA-
tS(CGA)M -GCTGTGATGGCCGAGTGGTTAAGGCGTTGGACTCGAAATCCAATGGGTTATTCCCCGCGCAGGTTCAAATCCTGCTCACAGCGTAA-
tS(TGA)J -GCAGCGATGGCCGAGTGGTTAAGGCGTTGGACTTGAAATCCAATGGGGTATCCCCGCGCAGGTTCGAACCCTGCTCGCTGCGGAA-
tS(CGA)J -GTCACGGTGGCCGAGTGGTTAAGGCGTTGGACTCGAAATCCAATGGGGTATCCCCGCACAGGTTCGAATCCTGTTCGTGACGGCA-


Serine2
1
25
50
75
87
tS(GCT)M3 TGACGAGGTGGCCGAGTGGTTAAGGCGATGGACTGCTAATCCATTGTGCACTGCACGCGTGGGTTCGAATCCCATCCTCGTCGGTCA
tS(GCT)M4 TGACGAGGTGGCCGAGTGGTTAAGGCGATGGACTGCTAATCCATTGTGCACTGCACGCGTGGGTTCGAATCCCATCCTCGTCGTTCA
tS(GCT)S -GACGAGGTGGCCGAGTGGTTAAGGCGATGGACTGCTAATCCATTGTGCTATGCACGCGTGGGTTCGAATCCCATCCTCGTCGGTA-
tS(GCT)K2 -GACGAGGTGGCCGAGTGGTTAAGGCGATGGACTGCTAATCCATTGTGCTATGCACGCGTGGGTTCGAATCCCATCCTCGTCGTCA-
tS(GCT)M2 -GACGAGGTGGCCGAGTGGTTAAGGCGATGGACTGCTAATCCATTGTGCTATGCACGCGTGGGTTCGAATCCCACCTTCGTCGTCA-
tS(GCT)M1 -GATGAGGTGGCCGAGTGGTTAAGGCGATGGACTGCTAATCCATTGTGCTATGCACGCATGGGTTCGAATCCCATCCTCATCGACA-
tS(GCT)B1 -GACGAGGTGGCCGAGTGGTTAAGGCGATGGACTGCTAATCCATTGTGCTATGCACGCATGGGTTCGAATCCCATCCTCGTCGAAA-
tS(GCT)A GGACGAGGTGGCCGAGTGGTTAAAGCGATGGACTGCTAATCCACTGTGCACAGTATGCGTGGGTTCGAATCCCATCCTCGTCCGAAA

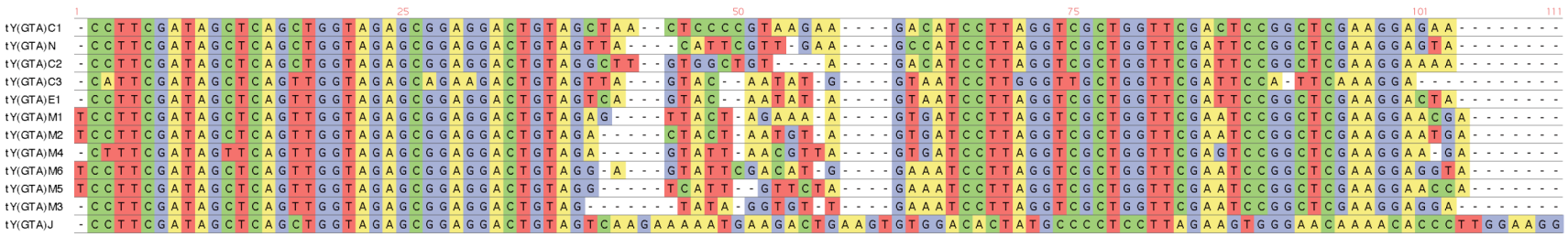
Threonine$_1$


1
25
50
75
98
tT(AGT)G
tT(AGT)N2
tT(AGT)K3
tT(AGT)K1
tT(AGT)K2
tT(AGT)M
tT(TGT)M1
tT(CGT)M2
tM(CAT)J
tT(CGT)M1

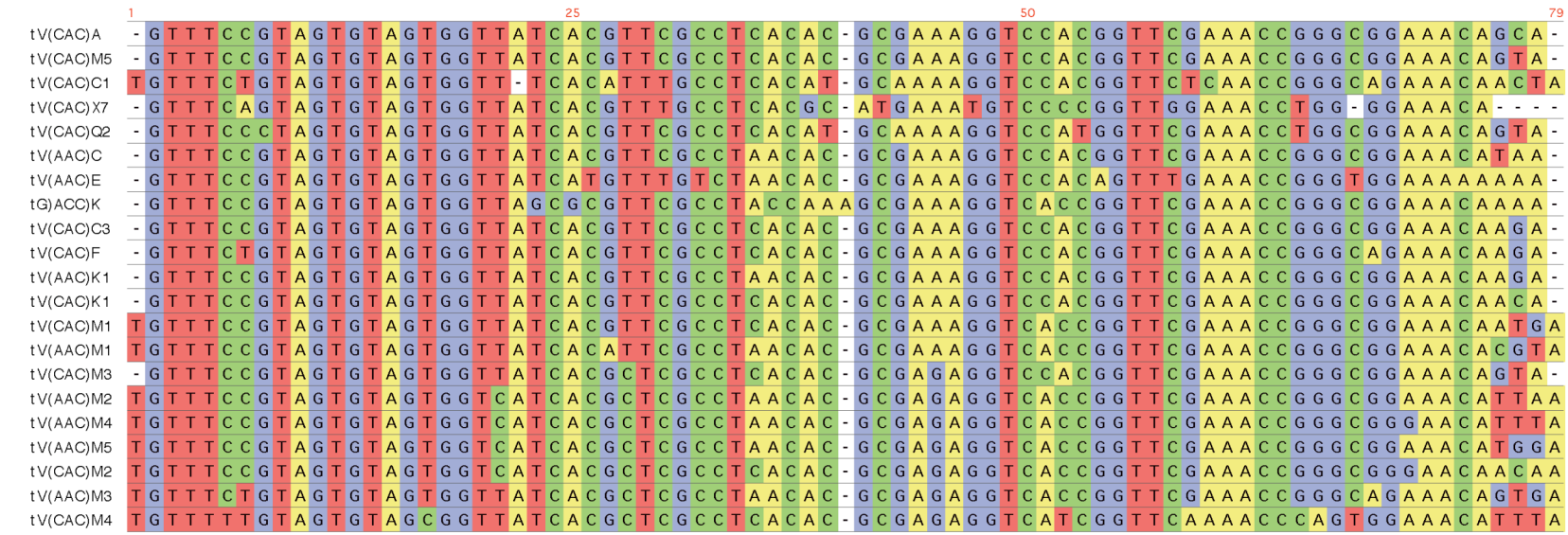

Threonine2
1
25
50
77
tT(CGT)K2
A G G C G C G G T G G C C A A G T G G T A A G G C G T C G G T C T C G T A A A C C G A A G A T C G A G G G T T C G A A C C C C G T C C G T G C C T G C G A
tT(CGT)P
A G G C G C G G T G G C C A A G T G G T A A G G C G T C G G T C T C G T A A A C C G A A G A T C A A G G G T T C G A A C C C C G T C C G T G C C T G C C A

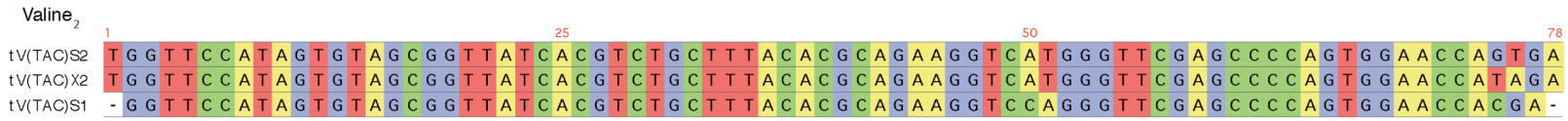

Tryptophan
1
25
50
77
tW(CCA)J1 TGACCTCGTGGCACAATGGTAGCACGTCTGACTCCAGATCAGAAGGTTGAGTGTTCAAATCACGTCGGGGGTCATGAA
tW(CCA)K2 TGACCTCGTGGCGCAATGGTAGCGCGTCTGACTCCAGATCAGAAGGTTGAGTGTTCAAATCACGTCGGGGGTCATGAA
tW(CCA)K1 TGACCTCGTGGCGCAATGGTAGCGCGTCTGACTCCAGATCAGAAGGTTGAGTGTTCAAGTCACGTCGGGGGTCAAGTA
tW(CCA)M2 TGACCTCGTGGCGCAACGGTAGCGCGTCTGACTCCAGATCAGAAGGTTGAGTGTTCAAATCACGTCGGGGGTCAAGTA
tW(CCA)M1 TGACCTCGTGGCGCAACGGTAGCGCGTCTGACTCCAGATCAGAAGGTTGAGTGTTCAAATCACGTCGGGGGTCAGTGA
tW(CCA)K3 -GGCCTCGTGGCGCAACGGTAGCGCGTCTGACTCCAGATCAGAAGGTTGCATGTTCAAATCACGTCGGGGGTCATCA-
tW(CCA)K4 -GGCCTCGTGGCGCAACGGTAGCGCGTCTGACTCCAGATCAGAAGGTTGCATGTTCAAATCACGTCGGGGGTCAGCA-
tW(CCA)J2 -GACCTCGTGGCGCAACGGTAGCGCGTCTGACTCCAGATCAGAAGGCTGCATGTTCGAATCACGTCGGGGGTCATAA-

Tyrosine


1
25
50
75
101
111
tY(GTA)C1
tY(GTA)N
tY(GTA)C2
tY(GTA)C3
tY(GTA)E1
tY(GTA)M1
tY(GTA)M2
tY(GTA)M4
tY(GTA)M6
tY(GTA)M5
tY(GTA)M3
tY(GTA)J

Valine$_1$


1
25
50
79
tV(CAC)A
tV(CAC)M5
tV(CAC)C1
tV(CAC)X7
tV(CAC)Q2
tV(AAC)C
tV(AAC)E
tG)ACC)K
tV(CAC)C3
tV(CAC)F
tV(AAC)K1
tV(CAC)K1
tV(CAC)M1
tV(AAC)M1
tV(CAC)M3
tV(AAC)M2
tV(AAC)M4
tV(AAC)M5
tV(CAC)M2
tV(AAC)M3
tV(CAC)M4


Valine2
1
25
50
78
tV(TAC)S2 TGGTTCCATAGTGTAGCGGTTATCACGTCTGCTTTACACGCAGAAGGTCATGGGTTCGAGCCCCAGTGGAACCAGTGA
tV(TAC)X2 TGGTTCCATAGTGTAGCGGTTATCACGTCTGCTTTACACGCAGAAGGTCATGGGTTCGAGCCCCAGTGGAACCATAGA
tV(TAC)S1 -GGTTCCATAGTGTAGCGGTTATCACGTCTGCTTTACACGCAGAAGGTCCAGGGTTCGAGCCCCAGTGGAACCACGA-

Appendix H

Oligonucleotide probe sequences used for northern blot analysis in Figure 3.2. These oligonucleotide sequences were designed to be complementary to the mature RNA sequence.

Mouse tRNA Family	Probe Sequence (5'-3')
Ala1	GCG CTC TAC CAC TGA GCT ACA CC
Ala2	GCT CTA CCA CTG AGC TAA ATC C
Ala3	GCG CTC TAC CAT TTG AGC TAA TCC
Arg1	CGA ACC CTT AAT CTT CTG ATC CG
Arg2	GGA GGC CAA TGC CTT ATC CAT TAG G
Asn	GGC TCG AAC CAC CTA CCT TTC GGT
Asp	GTG ACA GGC GGG GAT ACT CAC C
Cys	CCT GCT GAT CTG TAG TCA AAT GCT C
Gln	CGC TGG ATT CAG AGT CCA GAG TGC
Glu1	GCG CCG AAT CCT AAC CAC TAG ACC A
Glu2	GGC CGC CTG GGT GAA AAC CAG G
Gly1	GTG GGA GGC GAG AAT TCT ACC ACT GAA
Gly2	GAA GGC AGC TAT GCT TAC CAC TAT A
Gly3	CGC AAG AAT GGG AAT CTT GCA TGA T
His	CCT AGG TTG CTG CGG CCA CAA CG
Ile1	CCT TGG CGT TAT TAG CAC CAC GCT C
Ile2	GGT GAG GCT CGA ACT CAC AAC CTC GGC
Leu1	CCG TAG AGA CTG GAG CCT TAA TCC
Leu2	CCT CCA GTG GAG ACT GCG ACC TG
Leu3	CCA ACG CCT TAA CCA CTC AGC CAT CC
Lys1	CCT GAG ATT AAG AGA CTC TTG CTC
Lys2	CAG ATT AAA AGT CTG ATG CTC TAC C
Met1	CTG GGT TAT GGG ACC AGC ACG C
Met2	TGC GCT GCC TAC TAT GCT AAG G
Met3	CGC GCT ACC TAC TGC GCT AAC G
Phe	AGA TCT TCA GTC TAA CGC TCT CC
Pro	GTG AGA ATC ATA CCC CTA GAC CAA CAA GC
Ser1	GGG ATA CCC CAA TGG ATT TCT AG
Ser2	CCA TCG CCT TAA CCA CTC GGC CAC CTC G
Thr1	GCT GGG ATT CGA ACC CAT GAT CTC CTG
Thr2	CGA GAC CGA CGC CTT ACC ACT TGG
Trp	CTG GAG TCA GAC GTG CTA CCA TTG
Tyr	CAG TCC TCC GCT CTA CCA ACT GAG C
Val1	CGA ACG TGA TAA CCA CTA CAC TAC GG
Val1	CCG GTT TCG AAC CGG TGA CCT TTC GC
Val2	CCT GCA TGT GAG GCG AGC GAT CAC CAC

Bibliography

1. Frank, D.N. and N.R. Pace, *Ribonuclease P: unity and diversity in a tRNA processing ribozyme*. Annu Rev Biochem, 1998. **67**: p. 153-80.
2. Pace, N.R. and J.W. Brown, *Evolutionary perspective on the structure and function of ribonuclease P, a ribozyme*. J Bacteriol, 1995. **177**(8): p. 1919-28.
3. Guerrier-Takada, C., et al., *The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme*. Cell, 1983. **35**(3 Pt 2): p. 849-57.
4. Kurz, J.C., S. Niranjanakumari, and C.A. Fierke, *Protein component of Bacillus subtilis RNase P specifically enhances the affinity for precursor-tRNAAsp*. Biochemistry, 1998. **37**(8): p. 2393-400.
5. Fang, X.W., et al., *The Bacillus subtilis RNase P holoenzyme contains two RNase P RNA and two RNase P protein subunits*. Rna, 2001. **7**(2): p. 233-41.
6. Rueda, D., et al., *The 5' leader of precursor tRNAAsp bound to the Bacillus subtilis RNase P holoenzyme has an extended conformation*. Biochemistry, 2005. **44**(49): p. 16130-9.
7. Liu, F. and S. Altman, *Differential evolution of substrates for an RNA enzyme in the presence and absence of its protein cofactor*. Cell, 1994. **77**(7): p. 1093-100.
8. McClain, W.H., C. Guerrier-Takada, and S. Altman, *Model substrates for an RNA enzyme*. Science, 1987. **238**(4826): p. 527-30.
9. Kahle, D., U. Wehmeyer, and G. Krupp, *Substrate recognition by RNase P and by the catalytic M1 RNA: identification of possible contact points in pre-tRNAs*. Embo J, 1990. **9**(6): p. 1929-37.
10. Thurlow, D.L., D. Shilowski, and T.L. Marsh, *Nucleotides in precursor tRNAs that are required intact for catalysis by RNase P RNAs*. Nucleic Acids Res, 1991. **19**(4): p. 885-91.
11. Kirsebom, L.A. and S.G. Svard, *Base pairing between Escherichia coli RNase P RNA and its substrate*. Embo J, 1994. **13**(20): p. 4870-6.
12. Stams, T., et al., *Ribonuclease P protein structure: evolutionary origins in the translational apparatus*. Science, 1998. **280**(5364): p. 752-5.
13. Alifano, P., et al., *Ribonuclease E provides substrates for ribonuclease P-dependent processing of a polycistronic mRNA*. Genes Dev, 1994. **8**(24): p. 3021-31.
14. Altman, S., et al., *RNase P cleaves transient structures in some riboswitches*. Proc Natl Acad Sci U S A, 2005. **102**(32): p. 11284-9.
15. Giege, R., C. Florentz, and T.W. Dreher, *The TYMV tRNA-like structure*. Biochimie, 1993. **75**(7): p. 569-82.
16. Gimple, O. and A. Schon, *In vitro and in vivo processing of cyanelle tmRNA by RNase P*. Biol Chem, 2001. **382**(10): p. 1421-9.
17. Hartmann, R.K., et al., *Precursor of C4 antisense RNA of bacteriophages P1 and P7 is a substrate for RNase P of Escherichia coli*. Proc Natl Acad Sci U S A, 1995. **92**(13): p. 5822-6.
18. Komine, Y., et al., *A tRNA-like structure is present in 10Sa RNA, a small stable RNA from Escherichia coli*. Proc Natl Acad Sci U S A, 1994. **91**(20): p. 9223-7.

19. Li, Y. and S. Altman, *A specific endoribonuclease, RNase P, affects gene expression of polycistronic operon mRNAs*. Proc Natl Acad Sci U S A, 2003. **100**(23): p. 13213-8.
20. Li, Y., K. Cole, and S. Altman, *The effect of a single, temperature-sensitive mutation on global gene expression in Escherichia coli*. Rna, 2003. **9**(5): p. 518-32.
21. Peck-Miller, K.A. and S. Altman, *Kinetics of the processing of the precursor to 4.5 S RNA, a naturally occurring substrate for RNase P from Escherichia coli*. J Mol Biol, 1991. **221**(1): p. 1-5.
22. Cai, T., et al., *The Saccharomyces cerevisiae RNase mitochondrial RNA processing is critical for cell cycle progression at the end of mitosis*. Genetics, 2002. **161**(3): p. 1029-42.
23. Gill, T., et al., *RNase MRP cleaves the CLB2 mRNA to promote cell cycle progression: novel method of mRNA degradation*. Mol Cell Biol, 2004. **24**(3): p. 945-53.
24. Walker, S.C. and D.R. Engelke, *Ribonuclease P: the evolution of an ancient RNA enzyme*. Crit Rev Biochem Mol Biol, 2006. **41**(2): p. 77-102.
25. Chamberlain, J.R., et al., *Purification and characterization of the nuclear RNase P holoenzyme complex reveals extensive subunit overlap with RNase MRP*. Genes Dev, 1998. **12**(11): p. 1678-90.
26. Salinas, K., et al., *Characterization and purification of Saccharomyces cerevisiae RNase MRP reveals a new unique protein component*. J Biol Chem, 2005. **280**(12): p. 11352-60.
27. Dichtl, B. and D. Tollervey, *Pop3p is essential for the activity of the RNase MRP and RNase P ribonucleoproteins in vivo*. Embo J, 1997. **16**(2): p. 417-29.
28. Lygerou, Z., et al., *The POP1 gene encodes a protein component common to the RNase MRP and RNase P ribonucleoproteins*. Genes Dev, 1994. **8**(12): p. 1423-33.
29. Stolc, V. and S. Altman, *Rpp1, an essential protein subunit of nuclear RNase P required for processing of precursor tRNA and 35S precursor rRNA in Saccharomyces cerevisiae*. Genes Dev, 1997. **11**(21): p. 2926-37.
30. Kikovska, E., S.G. Svard, and L.A. Kirsebom, *Eukaryotic RNase P RNA mediates cleavage in the absence of protein*. Proc Natl Acad Sci U S A, 2007. **104**(7): p. 2062-7.
31. Tranguch, A.J., et al., *Structure-sensitive RNA footprinting of yeast nuclear ribonuclease P*. Biochemistry, 1994. **33**(7): p. 1778-87.
32. Leontis, N., et al., *Effects of tRNA-intron structure on cleavage of precursor tRNAs by RNase P from Saccharomyces cerevisiae*. Nucleic Acids Res, 1988. **16**(6): p. 2537-52.
33. Yang, L. and S. Altman, *A noncoding RNA in Saccharomyces cerevisiae is an RNase P substrate*. Rna, 2007. **13**(5): p. 682-90.
34. Samanta, M.P., et al., *Global identification of noncoding RNAs in Saccharomyces cerevisiae by modulating an essential RNA processing pathway*. Proc Natl Acad Sci U S A, 2006. **103**(11): p. 4192-7.
35. Ziehler, W.A., et al., *Effects of 5' leader and 3' trailer structures on pre-tRNA processing by nuclear RNase P*. Biochemistry, 2000. **39**(32): p. 9909-16.

36. Xiao, S., et al., *Functional characterization of the conserved amino acids in Pop1p, the largest common protein subunit of yeast RNases P and MRP*. Rna, 2006.
37. Sprinzl, M. and K.S. Vassilenko, *Compilation of tRNA sequences and sequences of tRNA genes*. Nucleic Acids Res, 2005. **33**(Database issue): p. D139-40.
38. Lowe, T.M. and S.R. Eddy, *tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence*. Nucleic Acids Res, 1997. **25**(5): p. 955-64.
39. Fichant, G.A. and C. Burks, *Identifying potential tRNA genes in genomic DNA sequences*. J Mol Biol, 1991. **220**(3): p. 659-71.
40. Pavesi, A., et al., *Identification of new eukaryotic tRNA genes in genomic DNA databases by a multistep weight matrix analysis of transcriptional control regions*. Nucleic Acids Res, 1994. **22**(7): p. 1247-56.
41. Eddy, S.R. and R. Durbin, *RNA sequence analysis using covariance models*. Nucleic Acids Res, 1994. **22**(11): p. 2079-88.
42. Laslett, D. and B. Canback, *ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences*. Nucleic Acids Res, 2004. **32**(1): p. 11-6.
43. Waterston, R.H., et al., *Initial sequencing and comparative analysis of the mouse genome*. Nature, 2002. **420**(6915): p. 520-62.
44. Krayev, A.S., et al., *Ubiquitous transposon-like repeats B1 and B2 of the mouse genome: B2 sequencing*. Nucleic Acids Res, 1982. **10**(23): p. 7461-75.
45. Xiao, S., et al., *Eukaryotic ribonuclease P: a plurality of ribonucleoprotein enzymes*. Annu Rev Biochem, 2002. **71**: p. 165-89.
46. Niranjanakumari, S., et al., *Protein component of the ribozyme ribonuclease P alters substrate recognition by directly contacting precursor tRNA*. Proc Natl Acad Sci U S A, 1998. **95**(26): p. 15212-7.
47. Chang, D.D. and D.A. Clayton, *A novel endoribonuclease cleaves at a priming site of mouse mitochondrial DNA replication*. Embo J, 1987. **6**(2): p. 409-17.
48. Lee, D.Y. and D.A. Clayton, *Initiation of mitochondrial DNA replication by transcription and R-loop processing*. J Biol Chem, 1998. **273**(46): p. 30614-21.
49. Schmitt, M.E. and D.A. Clayton, *Characterization of a unique protein component of yeast RNase MRP: an RNA-binding protein with a zinc-cluster domain*. Genes Dev, 1994. **8**(21): p. 2617-28.
50. Villa, T., et al., *Processing of the intron-encoded U18 small nucleolar RNA in the yeast Saccharomyces cerevisiae relies on both exo- and endonucleolytic activities*. Mol Cell Biol, 1998. **18**(6): p. 3376-83.
51. Srisawat, C. and D.R. Engelke, *Streptavidin aptamers: affinity tags for the study of RNAs and ribonucleoproteins*. Rna, 2001. **7**(4): p. 632-41.
52. Srisawat, C. and D.R. Engelke, *RNA affinity tags for purification of RNAs and ribonucleoprotein complexes*. Methods, 2002. **26**(2): p. 156-61.
53. Rigaut, G., et al., *A generic protein purification method for protein complex characterization and proteome exploration*. Nat Biotechnol, 1999. **17**(10): p. 1030-2.
54. Inada, M. and C. Guthrie, *Identification of Lhp1p-associated RNAs by microarray analysis in Saccharomyces cerevisiae reveals association with coding and noncoding RNAs*. Proc Natl Acad Sci U S A, 2004. **101**(2): p. 434-9.

55. Iyer, V.R., et al., *Genomic binding sites of the yeast cell-cycle transcription factors SBF and MBF*. Nature, 2001. **409**(6819): p. 533-8.
56. Holstege, F.C., et al., *Dissecting the regulatory circuitry of a eukaryotic genome*. Cell, 1998. **95**(5): p. 717-28.
57. Chamberlain, J.R., et al., *An RNase P RNA subunit mutation affects ribosomal RNA processing*. Nucleic Acids Res, 1996. **24**(16): p. 3158-66.
58. Pagan-Ramos, E., Y. Lee, and D.R. Engelke, *Mutational analysis of Saccharomyces cerevisiae nuclear RNase P: randomization of universally conserved positions in the RNA subunit*. Rna, 1996. **2**(5): p. 441-51.
59. Petfalski, E., et al., *Processing of the precursors to small nucleolar RNAs and rRNAs requires common components*. Mol Cell Biol, 1998. **18**(3): p. 1181-9.
60. Villa, T., F. Ceradini, and I. Bozzoni, *Identification of a novel element required for processing of intron-encoded box C/D small nucleolar RNAs in Saccharomyces cerevisiae*. Mol Cell Biol, 2000. **20**(4): p. 1311-20.
61. Bertrand, E., et al., *Nucleolar localization of early tRNA processing*. Genes Dev, 1998. **12**(16): p. 2463-8.
62. Pagan-Ramos, E., Y. Lee, and D.R. Engelke, *A conserved RNA motif involved in divalent cation utilization by nuclear RNase P*. Rna, 1996. **2**(11): p. 1100-9.
63. DeRisi, J.L., V.R. Iyer, and P.O. Brown, *Exploring the metabolic and genetic control of gene expression on a genomic scale*. Science, 1997. **278**(5338): p. 680-6.
64. Kohrer, K. and H. Domdey, *Preparation of high molecular weight RNA*. Methods Enzymol, 1991. **194**: p. 398-405.
65. Hull, M.W., et al., *Protein-DNA interactions in vivo--examining genes in Saccharomyces cerevisiae and Drosophila melanogaster by chromatin footprinting*. Methods Cell Biol, 1991. **35**: p. 383-415.
66. Bystrom, A.S. and G.R. Fink, *A functional analysis of the repeated methionine initiator tRNA genes (IMT) in yeast*. Mol Gen Genet, 1989. **216**(2-3): p. 276-86.
67. Abelson, J., C.R. Trotta, and H. Li, *tRNA splicing*. J Biol Chem, 1998. **273**(21): p. 12685-8.
68. Weiner, A.M., *SINEs and LINEs: the art of biting the hand that feeds you*. Curr Opin Cell Biol, 2002. **14**(3): p. 343-50.
69. Goodenbour, J.M. and T. Pan, *Diversity of tRNA genes in eukaryotes*. Nucleic Acids Res, 2006. **34**(21): p. 6137-46.
70. Dittmar, K.A., J.M. Goodenbour, and T. Pan, *Tissue-Specific Differences in Human Transfer RNA Expression*. PLoS Genet, 2006. **2**(12): p. e221.
71. Chenna, R., et al., *Multiple sequence alignment with the Clustal series of programs*. Nucleic Acids Res, 2003. **31**(13): p. 3497-500.
72. Karolchik, D., et al., *The UCSC Genome Browser Database*. Nucleic Acids Res, 2003. **31**(1): p. 51-4.
73. Zuker, M., *Mfold web server for nucleic acid folding and hybridization prediction*. Nucleic Acids Res, 2003. **31**(13): p. 3406-15.
74. Hughes, T.R., et al., *Microarray analysis of RNA processing and modification*. Methods Enzymol, 2006. **410**: p. 300-16.

75. Edgar, R., M. Domrachev, and A.E. Lash, *Gene Expression Omnibus: NCBI gene expression and hybridization array data repository*. Nucleic Acids Res, 2002. **30**(1): p. 207-10.
76. Wiegant, J.C., et al., *ULS: a versatile method of labeling nucleic acids for FISH based on a monofunctional reaction of cisplatin derivatives with guanine moieties*. Cytogenet Cell Genet, 1999. **87**(1-2): p. 47-52.
77. Huber, W., et al., *Variance stabilization applied to microarray data calibration and to the quantification of differential expression*. Bioinformatics, 2002. **18 Suppl 1**: p. S96-104.
78. Patel, C.V. and K.P. Gopinathan, *Development stage-specific expression of fibroin in the silk worm Bombyx mori is regulated translationally*. Indian J Biochem Biophys, 1991. **28**(5-6): p. 521-30.
79. Stutz, F., E. Gouilloud, and S.G. Clarkson, *Oocyte and somatic tyrosine tRNA genes in Xenopus laevis*. Genes Dev, 1989. **3**(8): p. 1190-8.
80. Ohama, T., et al., *Mouse selenocysteine tRNA([Ser]Sec) gene (Trsp) and its localization on chromosome 7*. Genomics, 1994. **19**(3): p. 595-6.
81. Kelly, V.P., et al., *The distal sequence element of the selenocysteine tRNA gene is a tissue-dependent enhancer essential for mouse embryogenesis*. Mol Cell Biol, 2005. **25**(9): p. 3658-69.
82. Hopper, A.K. and E.M. Phizicky, *tRNA transfers to the limelight*. Genes Dev, 2003. **17**(2): p. 162-80.
83. Soll, D. and U.L. RajBhandary, *The genetic code - thawing the '\frozen accident'*. J Biosci, 2006. **31**(4): p. 459-63.
84. Schultz, D.W. and M. Yarus, *Transfer RNA mutation and the malleability of the genetic code*. J Mol Biol, 1994. **235**(5): p. 1377-80.
85. McCutcheon, J.P. and S.R. Eddy, *Computational identification of non-coding RNAs in Saccharomyces cerevisiae by comparative genomics*. Nucleic Acids Res, 2003. **31**(14): p. 4119-28.