tRNA Metabolism and Neurodevelopmental Disorders

Ashleigh E. Schaffer, Otis Pinkard, and Jeffery M. Coller

Department of Genetics and Genome Sciences and Center for RNA Science and Therapeutics, Case Western Reserve University, Cleveland, Ohio 44106, USA; email: ashleigh.schaffer@case.edu

Keywords
tRNA, neurodevelopment, neurodegeneration, translation, sequencing, oligodendrocyte

Abstract
tRNAs are short noncoding RNAs required for protein translation. The human genome includes more than 600 putative tRNA genes, many of which are considered redundant. tRNA transcripts are subject to tightly controlled, multistep maturation processes that lead to the removal of flanking sequences and the addition of nontemplated nucleotides. Furthermore, tRNAs are highly structured and posttranscriptionally modified. Together, these unique features have impeded the adoption of modern genomics and transcriptomics technologies for tRNA studies. Nevertheless, it has become apparent from human neurogenetic research that many tRNA biogenesis proteins cause brain abnormalities and other neurological disorders when mutated. The cerebral cortex, cerebellum, and peripheral nervous system show defects, impairment, and degeneration upon tRNA misregulation, suggesting that they are particularly sensitive to changes in tRNA expression or function. An integrated approach to identify tRNA species and contextually characterize tRNA function will be imperative to drive future tool development and novel therapeutic design for tRNA-associated disorders.
1. INTRODUCTION

tRNAs are small noncoding RNAs that function as adapters to decode mRNAs during protein translation by mediating aminoacylation reactions. Following transcription, each tRNA must undergo an extensive, multistep maturation process to become functional. Each step in this highly regulated and complex progression offers an opportunity to fine-tune tRNA activity and abundance based on cellular cues. At any given time within a cell, there may be subtle, single-base variances that differentiate between tRNA molecules; thus, discerning the composition of the tRNAs that are present and available for protein synthesis can be challenging. Moreover, new roles for tRNAs are still being uncovered even though they were first discovered eight decades ago. Our understanding of tRNA biology has been limited primarily by the tools available to interrogate tRNA identity and function in vivo. With the inception of RNA-sequencing and bioinformatics methods optimized for tRNA-sequencing reads, combined with recent advances in proteomics that offer protein readout correlations, a multifaceted view of tRNA complexity and function has begun to emerge (85, 154, 208).

Neurodevelopmental disorders arise in utero or shortly after birth and are caused by impaired growth of the brain or nervous system. They can result from abnormal development or early degeneration of neurons within the central nervous system caused by genetic or environmental insults. For genetic cases, clinical genomic sequencing has enabled the discovery of new causative genes and mutations as well as improved disease classification. A subset of patients with pediatric neurological disorders harbor variants in genes that are known to regulate tRNA-processing events (Table 1). Perhaps surprisingly, mutations in these genes lead to a wide array of phenotypes and diseases. Histological analysis of tissue from affected patients has shown that some cell types and tissues are affected by these mutations, while others are spared, which is also known as selective variability (50). However, some clinical features are shared, consistent with a common disease mechanism.

In this review, we describe the process by which tRNAs are created, regulated, and degraded within the cell and discuss ongoing challenges and recent solutions in the study of tRNAs in research and the clinic. We also identify shared mechanisms and phenotypes of tRNA-mediated neurological diseases and explore future therapeutic avenues to target tRNA processes in disease contexts.

2. THE tRNA LIFE CYCLE

2.1. Regulation of Transcription

In contrast to mRNA transcription, which is mediated by RNA polymerase II, transcription of tRNA gene elements is mediated by RNA polymerase III. RNA polymerase III is recruited to tRNA genes by the transcription factors TFIIIB and TFIIIC, which recognize two internal promoter elements, box A and box B, respectively (47, 68, 107, 155) (Figure 1). These transcription factors consist of multiple subunits under independent transcriptional and posttranslational modification by proteins associated with nutrient and cell stress responses as well as normal cell proliferation. Upregulation of transcription and activating posttranslational modifications of TFIIIB subunits are typically associated with increased tRNA transcription, while activation of TFIIIC represses transcription (188, 195).

One of the most well-characterized regulators of RNA polymerase III is Maf1. Maf1 acts as an integration center for nutrient and cellular stress signaling pathways and functions as a repressor of RNA polymerase III activity, with orthologs present in all phylogenetic domains. While cells are in a proliferative state, Maf1 is phosphorylated and confined to the cytoplasm.
Table 1  Childhood neurological disorders and associated genes, grouped by category of tRNA expression or function

<table>
<thead>
<tr>
<th>Disorder</th>
<th>OMIMa</th>
<th>Inheritance patternb</th>
<th>Gene(s)</th>
<th>Reference(s)c</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mitochondrial tRNA gene mutations or deletions</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kearns–Sayre syndrome</td>
<td>530000</td>
<td>Mi</td>
<td>Deletion of MT-TL2, MT-TS2, MT-TH, MT-TR, and/or MT-TG</td>
<td>93</td>
</tr>
<tr>
<td>Leigh syndrome</td>
<td>256000</td>
<td>Mi</td>
<td>MT-TV, MT-TK, MT-TW, MT-TL1, MT-TI, MT-TL2</td>
<td>94, 149, 184, 187</td>
</tr>
<tr>
<td>Mitochondrial complex I deficiency</td>
<td>252010</td>
<td>Mi</td>
<td>MT-TL1, MT-TW, MT-TN</td>
<td>105, 172</td>
</tr>
<tr>
<td>Mitochondrial complex IV deficiency</td>
<td>220110</td>
<td>Mi</td>
<td>MT-TS1, MT-TL1</td>
<td>16, 74</td>
</tr>
<tr>
<td>Mitochondrial DNA depletion syndrome 1 (MNGIE type)</td>
<td>603041</td>
<td>AR</td>
<td>TYMP</td>
<td>51</td>
</tr>
<tr>
<td>Mitochondrial myopathy with diabetes</td>
<td>500002</td>
<td>Mi</td>
<td>MT-TE</td>
<td>61</td>
</tr>
<tr>
<td>Mitochondrial myopathy, encephalopathy, lactic acidosis, and stroke-like episodes (MELAS)</td>
<td>540000</td>
<td>Mi</td>
<td>MT-TL1, MT-TQ, MT-TH, MT-TK, MT-TC, MT-TS1, MT-TS2</td>
<td>9, 58, 103, 111, 207</td>
</tr>
<tr>
<td>Mitochondrial myopathy, infantile, transient (MMIT)</td>
<td>500009</td>
<td>Mi</td>
<td>MT-TE</td>
<td>64</td>
</tr>
<tr>
<td>Mitochondrial myopathy, lethal, infantile (LIMM)</td>
<td>551000</td>
<td>Mi</td>
<td>MT-TT</td>
<td>205</td>
</tr>
<tr>
<td>Myoclonic epilepsy with ragged red fibers (MERRF)</td>
<td>545000</td>
<td>Mi</td>
<td>MT-TL1, MT-TK, MT-TH, MT-TS1, MT-TS2, MT-TF</td>
<td>15, 101, 103, 111, 162</td>
</tr>
<tr>
<td>Wolfram syndrome, mitochondrial form</td>
<td>598500</td>
<td>Mi</td>
<td>Deletion of MT-TL2, MT-TS2, MT-TH, MT-TR, MT-TG, MT-TK, MT-TF, MT-TD, and/or MT-TS1</td>
<td>145</td>
</tr>
<tr>
<td><strong>tRNA biogenesis</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cerebellofaciodental syndrome</td>
<td>616202</td>
<td>AR</td>
<td>BRF1</td>
<td>18</td>
</tr>
<tr>
<td>Combined oxidative phosphorylation deficiency 17</td>
<td>615440</td>
<td>AR</td>
<td>ELAC2</td>
<td>60</td>
</tr>
<tr>
<td>HSD10 mitochondrial disease</td>
<td>300438</td>
<td>XLD</td>
<td>HSD17B10</td>
<td>212</td>
</tr>
<tr>
<td>Leukodystrophy, hypomyelinating, 7, with or without oligodentia and/or hypogonadotropic hypogonadism</td>
<td>607694</td>
<td>AR</td>
<td>POLR3A</td>
<td>14</td>
</tr>
<tr>
<td>Leukodystrophy, hypomyelinating, 8, with or without oligodentia and/or hypogonadotropic hypogonadism</td>
<td>614381</td>
<td>AR</td>
<td>POLR3B</td>
<td>148</td>
</tr>
<tr>
<td>Pontocerebellar hypoplasia, type 10</td>
<td>615803</td>
<td>AR</td>
<td>CLP1</td>
<td>81, 151</td>
</tr>
<tr>
<td>Pontocerebellar hypoplasia, type 2B</td>
<td>612389</td>
<td>AR</td>
<td>TSEN2</td>
<td>23</td>
</tr>
<tr>
<td>Pontocerebellar hypoplasia, type 2C</td>
<td>612390</td>
<td>AR</td>
<td>TSEN34</td>
<td>23</td>
</tr>
<tr>
<td>Pontocerebellar hypoplasia, type 2F</td>
<td>617026</td>
<td>AR</td>
<td>TSEN15</td>
<td>3, 21</td>
</tr>
<tr>
<td>Pontocerebellar hypoplasia, type 4 or 5</td>
<td>225753</td>
<td>AR</td>
<td>TSEN34</td>
<td>23</td>
</tr>
<tr>
<td>Sideroblastic anemia with B cell immunodeficiency, periodic fevers, and developmental delay</td>
<td>616084</td>
<td>AR</td>
<td>TRNT1</td>
<td>26</td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>Disorder</th>
<th>OMIM</th>
<th>Inheritance pattern</th>
<th>Gene(s)</th>
<th>Reference(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>tRNA charging</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cataracts, growth hormone deficiency, sensory neuropathy, sensorineural hearing loss, and skeletal dysplasia</td>
<td>616007</td>
<td>AR</td>
<td>LARS2</td>
<td>157</td>
</tr>
<tr>
<td>Charcot–Marie–Tooth disease, axonal, type 2N</td>
<td>613287</td>
<td>AD</td>
<td>AARS</td>
<td>91</td>
</tr>
<tr>
<td>Charcot–Marie–Tooth disease, axonal, type 2U</td>
<td>616280</td>
<td>AD</td>
<td>MARS</td>
<td>56</td>
</tr>
<tr>
<td>Charcot–Marie–Tooth disease, axonal, type 2W</td>
<td>616625</td>
<td>AD</td>
<td>HARS</td>
<td>186</td>
</tr>
<tr>
<td>Charcot–Marie–Tooth disease, dominant intermediate C</td>
<td>608323</td>
<td>AD</td>
<td>YARS</td>
<td>77</td>
</tr>
<tr>
<td>Charcot–Marie–Tooth disease, recessive intermediate, B</td>
<td>613641</td>
<td>AR</td>
<td>KARS</td>
<td>102</td>
</tr>
<tr>
<td>Charcot–Marie–Tooth disease, type 2D</td>
<td>601472</td>
<td>AD</td>
<td>GARS</td>
<td>70</td>
</tr>
<tr>
<td>Combined oxidative phosphorylation deficiency 12</td>
<td>614924</td>
<td>AR</td>
<td>EARS2</td>
<td>170</td>
</tr>
<tr>
<td>Combined oxidative phosphorylation deficiency 14</td>
<td>614946</td>
<td>AR</td>
<td>EARS2</td>
<td>160</td>
</tr>
<tr>
<td>Combined oxidative phosphorylation deficiency 15</td>
<td>614947</td>
<td>AR</td>
<td>MTFMT</td>
<td>180</td>
</tr>
<tr>
<td>Combined oxidative phosphorylation deficiency 20</td>
<td>615917</td>
<td>AR</td>
<td>VARS2</td>
<td>175</td>
</tr>
<tr>
<td>Combined oxidative phosphorylation deficiency 21</td>
<td>615918</td>
<td>AR</td>
<td>TARS2</td>
<td>38</td>
</tr>
<tr>
<td>Combined oxidative phosphorylation deficiency 24</td>
<td>616239</td>
<td>AR</td>
<td>NARS2</td>
<td>168</td>
</tr>
<tr>
<td>Combined oxidative phosphorylation deficiency 25</td>
<td>616430</td>
<td>AR</td>
<td>MARS2</td>
<td>192</td>
</tr>
<tr>
<td>Combined oxidative phosphorylation deficiency 8</td>
<td>614096</td>
<td>AR</td>
<td>AARS2</td>
<td>59</td>
</tr>
<tr>
<td>Epileptic encephalopathy, early infantile, 29</td>
<td>616339</td>
<td>AR</td>
<td>AARS</td>
<td>164</td>
</tr>
<tr>
<td>Growth retardation, intellectual developmental disorder, hypotonia, and hepatopathy</td>
<td>617093</td>
<td>AR</td>
<td>LARS2</td>
<td>86</td>
</tr>
<tr>
<td>Hydrops, lactic acidosis, and sideroblastic anemia</td>
<td>617021</td>
<td>AR</td>
<td>LARS2</td>
<td>136</td>
</tr>
<tr>
<td>Hyperuricemia, pulmonary hypertension, renal failure, and alkalosis</td>
<td>613845</td>
<td>AR</td>
<td>SARS2</td>
<td>12</td>
</tr>
<tr>
<td>Hypomyelination with brain stem and spinal cord involvement and leg spasticity</td>
<td>615281</td>
<td>AR</td>
<td>DARS</td>
<td>173</td>
</tr>
<tr>
<td>Infantile-onset multisystem neurologic, endocrine, and pancreatic disease</td>
<td>616263</td>
<td>AR</td>
<td>PTRH2</td>
<td>65</td>
</tr>
<tr>
<td>Leukodystrophy, hypomyelinating, 17</td>
<td>618006</td>
<td>AR</td>
<td>AIMP2</td>
<td>163</td>
</tr>
<tr>
<td>Leukodystrophy, hypomyelinating, 3</td>
<td>260600</td>
<td>AR</td>
<td>AIMP1</td>
<td>46</td>
</tr>
<tr>
<td>Leukodystrophy, hypomyelinating, 9</td>
<td>616140</td>
<td>AR</td>
<td>DARS2</td>
<td>197</td>
</tr>
<tr>
<td>Leukoencephalopathy with brain stem and spinal cord involvement and lactate elevation</td>
<td>611105</td>
<td>AR</td>
<td>DARS2</td>
<td>152</td>
</tr>
<tr>
<td>Leukoencephalopathy, progressive, with ovarian failure</td>
<td>615889</td>
<td>AR</td>
<td>AARS2</td>
<td>33</td>
</tr>
<tr>
<td>Microcephaly, progressive, seizures, and cerebellar and cerebellar atrophy</td>
<td>615760</td>
<td>AR</td>
<td>QARS</td>
<td>209</td>
</tr>
<tr>
<td>Neurodevelopmental disorder with brain, liver, and lung abnormalities</td>
<td>618007</td>
<td>AR</td>
<td>FARSB</td>
<td>6</td>
</tr>
<tr>
<td>Neurodevelopmental disorder with microcephaly, ataxia, and seizures</td>
<td>617709</td>
<td>AR</td>
<td>SARS</td>
<td>109</td>
</tr>
</tbody>
</table>

(Continued)
Table 1 (Continued)

<table>
<thead>
<tr>
<th>Disorder</th>
<th>OMIMa</th>
<th>Inheritance patternb</th>
<th>Gene(s)</th>
<th>Reference(s)c</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neurodevelopmental disorder with microcephaly, seizures, and cortical atrophy</td>
<td>617802</td>
<td>AR</td>
<td>VARs</td>
<td>80</td>
</tr>
<tr>
<td>Neurodevelopmental disorder, mitochondrial, with abnormal movements and lactic acidosis, with or without seizures</td>
<td>617710</td>
<td>AR</td>
<td>WARs2</td>
<td>109</td>
</tr>
<tr>
<td>Neuropathy, distal hereditary motor, type VA</td>
<td>600794</td>
<td>AD</td>
<td>GARs</td>
<td>5</td>
</tr>
<tr>
<td>Pontocerebellar hypoplasia, type 6</td>
<td>611523</td>
<td>AR</td>
<td>RARs2</td>
<td>42</td>
</tr>
<tr>
<td>Spastic ataxia 3</td>
<td>611390</td>
<td>AR</td>
<td>MARS2</td>
<td>10</td>
</tr>
<tr>
<td>Spastic paraplegia 77</td>
<td>617046</td>
<td>AR</td>
<td>FARS2</td>
<td>185</td>
</tr>
<tr>
<td>Usher syndrome type 3B</td>
<td>614504</td>
<td>AR</td>
<td>HARS</td>
<td>131</td>
</tr>
</tbody>
</table>

**tRNA modification**

<table>
<thead>
<tr>
<th>Disorder</th>
<th>OMIMa</th>
<th>Inheritance patternb</th>
<th>Gene(s)</th>
<th>Reference(s)c</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mental retardation, autosomal recessive 5</td>
<td>611091</td>
<td>AR</td>
<td>NSUN2</td>
<td>1, 84</td>
</tr>
<tr>
<td>Myopathy, lactic acidosis, and sideroblastic anemia 1</td>
<td>600462</td>
<td>AR</td>
<td>PUS1</td>
<td>24</td>
</tr>
<tr>
<td>Combined oxidative phosphorylation deficiency 26</td>
<td>616539</td>
<td>AR</td>
<td>TRMT5</td>
<td>129</td>
</tr>
<tr>
<td>Combined oxidative phosphorylation deficiency 35</td>
<td>617873</td>
<td>AR</td>
<td>TRIT1</td>
<td>201</td>
</tr>
<tr>
<td>Combined oxidative phosphorylation deficiency 23</td>
<td>616198</td>
<td>AR</td>
<td>GTPBP3</td>
<td>86</td>
</tr>
<tr>
<td>Mental retardation, autosomal recessive 55</td>
<td>617051</td>
<td>AR</td>
<td>PUS3</td>
<td>159</td>
</tr>
<tr>
<td>Galloway–Mowat syndrome 3</td>
<td>617729</td>
<td>AR</td>
<td>OSGEP</td>
<td>41</td>
</tr>
<tr>
<td>Combined oxidative phosphorylation deficiency 10</td>
<td>614702</td>
<td>AR</td>
<td>MT01</td>
<td>52</td>
</tr>
<tr>
<td>Mental retardation, autosomal recessive 36</td>
<td>615286</td>
<td>AR</td>
<td>ADHT3</td>
<td>2</td>
</tr>
</tbody>
</table>

aPhenotype OMIM (Online Mendelian Inheritance in Man) number.
bAbbreviations: AD, autosomal dominant; AR, autosomal recessive; Mi, mitochondrial; XLD, X-linked disease.
cDue to space constraints, references are limited to the seminal discovery or discoveries.

In times of stress, Maf1 is dephosphorylated and localizes to the nucleus, where its interactions with RNA polymerase III repress tRNA transcription (29, 79), indicating the presence of a dynamic mechanism that modulates the global tRNA abundance associated with cell state. One interesting finding suggests that changes occur in individual tRNA gene expression in addition to changes in global levels: Gingold et al. (53) showed that tRNA isoacceptors (tRNAs that recognize a codon of the same amino acid) and isodecoders (tRNAs that recognize the same codon) have discrete expression profiles in different phases of proliferation. However, the mechanisms modulating the abundance of specific tRNA isoacceptor classes have yet to be elucidated.

After transcription initiation and subsequent synthesis of the nascent tRNA transcript, RNA polymerase III activity is terminated. The consensus termination signal of four or more thymidine residues on the sense strand produces a heterogeneous population of pre-tRNA transcripts with varying lengths of 3’ oligouridylation (114). This oligouridylation provides necessary protection from 3’-to-5’ degradation mediated by nuclear surveillance machinery. The 3’ oligouridylated transcripts act as preferential substrates for competitive La protein binding (Figure 1), while nascent transcripts that lack 3’ oligouridylation preferentially bind to the Trf4–Air2 subcomplex of the TRAMP4 nuclear exosome complex, resulting in polyadenylation and 3’-to-5’ degradation (89, 204). Recruitment of La protein to the newly formed immature tRNA transcript marks one of many posttranscriptional regulatory events governing the levels of a mature tRNA abundance.
2.2. Processing of Transcript Sequence

Early enzymatic processing of pre-tRNA transcripts involves the removal of the 5′ leader, 3′ trailer, and introns of variable length in cistronic tRNA genes. The order and cellular localization of these events can vary across species and tRNA isoacceptor families. However, these maturation steps remain highly conserved in eukaryotes. Late enzymatic processing includes the addition of CCA to the 3′ terminus, export of the mature tRNA transcript out of the nucleus to the cytoplasm, and proposed localization to subcellular compartments (Figure 1). Throughout the maturation process, pre-tRNA transcripts gain posttranscriptional modifications, with varying effects specific to individual mature tRNA species. The posttranscriptional modifications found ubiquitously among mature tRNA species are discussed in Section 2.6.

(Caption appears on following page)
Figure 1 (Figure appears on preceding page)

Processing and maturation of transfer RNA. (Step 1) RNA polymerase III (Pol III) is recruited to tRNA gene elements by the transcription factor TFIIB. The immature transcript contains 5′ leader (green), intronic (blue), and 3′ trailer (red) sequences. Mature sequences are shown in black. (Step 2) The La protein binds to the 3′ oligouridylate acid [oligo(U)] sequence that protects against degradation by the nuclear exosome complex. TSEN and CLP1 form a complex to remove introns and ligate exonic sequences from intron-containing tRNA species. (Step 3) RNase P interacts with phosphorylated La protein to remove 5′ leader sequences. (Step 4) ELAC isoforms remove 3′ trailer sequences, leaving a single-nucleotide overhang termed the discriminator base. (Step 5) TRNT1 mediates the addition of CCA to the 3′ terminus of immature species. (Step 6) tRNA species with proper addition of CCA as well as secondary and tertiary structure are bound by exportin-1 or exportin-5. After translocation through the nuclear pore complex (light blue), the mature tRNAs are able to participate in aminoacylation. (Step 7) tRNA species are recognized by anticodon loop and acceptor stem sequences, and aminoacylation is added by cognate aminoacyl-tRNA synthetases (aaRSs). (Step 8) Aminoacylated tRNA species participate in translation through incorporation into the ribosomal A-site. (Step 9) In times of cellular stress, tRNA species may be cleaved by angiogenin into 5′ and 3′ tRNA halves (tiRNAs). (Step 10) The 5′ tiRNA species inhibit translation initiation through competitive binding of the elf4F complex. (Step 11) The 5′ tiRNA species inhibit the formation of the apoptosome through competitive binding of cytochrome c (Cyt c).

Precise removal of 5′ leader sequences from pre-tRNA transcripts is enzymatically mediated by the holoenzyme RNase P (Figure 1). In humans, the RNase P structure contains the catalytic ribonucleotide sequence H1 and 10 identified protein subunits that act as cofactors (8, 75, 76). The ribosezyme sequence and complexity of protein cofactors vary greatly across the phylogenetic tree; however, the role of RNase P is functionally represented across all domains of life.

The binding of RNase P to the pre-tRNA substrate is regulated in part by phosphorylation of La protein at S366 (69). The interaction of La protein with the oligouridine 3′ trailer sequence enables binding of the basic C terminus to the triphosphate present on the 5′ residue of the leader sequence. Phosphorylation at S366 relieves this interaction and facilitates binding of RNase P to execute enzymatic cleavage of the pre-tRNA transcript (Figure 1). Acting as a partial substrate in the crystal structure of RNase P isolated from Thermotoga maritima in complex with mature tRNA^Phe, the intact TψC loop sequence seemingly plays an important role, along with contributions from the D-loop in RNase P recruitment to the nascent pre-tRNA transcript and subsequent processing of the 5′ leader sequence (123, 135).

Enzymatic processing of the 3′ trailer sequence is dependent on the level of 3′ oligouridylation. This process is mediated by RNase Z in bacteria and the eukaryotic homolog ELAC2 in humans (153, 174) (Figure 1). Two isoforms of ELAC2 exist in eukaryotic cells secondary to alternative translation initiation sites (141). ELAC2 is the full-length gene product that retains the mitochondrial localization signal found in the mitochondria, while ELAC1 initiates at a start codon downstream of this localization signal, resulting in a truncated protein confined to the cytosol. Also implicated in normal 3′ processing of nucleus-encoded tRNA are two exonucleases in competition with La protein binding for the 3′ terminal nucleotides: Rxp1p (a 3′-to-5′ exonuclease) and Rrp6p (part of the TRAMP4 complex associated with nuclear surveillance of small RNAs) (31, 118). The 3′ processing can be divided into species of pre-tRNA whose 3′ trailer sequences contain oligouridylation sequences suitable for La protein recruitment. In species with sufficient 3′ oligouridylation and subsequent recruitment of La protein, La-dependent processing is thought to be the dominant form of 3′ trailer removal. In pre-tRNA species that are not suitable substrates for La protein, it is thought that Rxp1p trims the 3′ trailer. Following 5′ leader excision, both La-dependent and La-independent pathways result in a single-nucleotide overhang. This overhang acts as a suitable substrate for further enzymatic reactions in the maturation process and a less suitable substrate for the nuclear exosome pathway.
In cistronic species of tRNA genes, excision of the intronic sequences is mediated by the tRNA-specific endonuclease (TSEN) complex, resulting in 5′ and 3′ exons (179, 198) (Figure 1). In mammalian cells, this process occurs in the nucleus. In Saccharomyces cerevisiae, the TSEN complex is localized to the surface of the mitochondria, necessitating nuclear export for final processing (206). Intron sequences of tRNA have been identified exclusively between the anticodon and TψC loops with heterogeneous lengths of 14–60 nucleotides. The 5′ exonic and intronic sequences of the nascent transcript are recognized by the TSEN complex and cleaved, resulting in a 5′ hydroxyl group on the 3′ exonic sequence, as well as a 2′,3′-cyclic phosphodiester on the 3′ end of the 5′ exonic sequence following cyclization. The 3′ exonic fragment is phosphorylated by cleavage and polyadenylation factor 1 subunit 1 protein (CLP1) (134). The 5′ and 3′ exons serve as substrates for ligation by the human tRNA ligase complex. RNA interference studies targeting CLP1 and RTCB, which are subunits of the ligase complex, resulted in decreased maturation and accumulation of intron-containing pre-tRNA halves, suggesting the importance of these enzymatic processes in the normal processing of cistronic tRNA species (128).

The final alteration of both cytoplasmic tRNA and mitochondrial tRNA (mt-tRNA) transcript sequences is the addition of CCA nucleotides to the 3′ terminus by tRNA nucleotidyl transferase 1 (TRNT1) in humans and CCA1 in S. cerevisiae (95, 110) (Figure 1). As the only known step in tRNA maturation consisting of polymerase-mediated addition, these enzymes modify pre-tRNA species with high fidelity, which is attributed to the specificity of the binding pocket for pre-tRNA, with properly trimmed 5′ leader and 3′ trailer sequences protecting mature tRNA species from multiple additions of CCA. The presence of 3′ CCA is required for nuclear export and serves as a substrate for tRNA-aminoacyltransferases in the cytoplasm to function in translation.

### 2.3. Localization and Aminoacylation

Export of nucleus-encoded mature tRNA transcripts to the cytoplasm and other extranuclear targets, including mitochondria and the endoplasmic reticulum, represents a final shared regulatory step among known tRNA genes. In mammalian cells, exportin-1 is critical to the proper export of mature tRNA species, with minor contributions from exportin-5 (7, 17) (Figure 1). Inhibition of function results in nuclear accumulation of mature tRNA transcripts. Pathways of tRNA export consist of ternary complex formation with and without Ran family GTPases. Recognition of the acceptor stem; proper processing of 5′ leader, 3′ trailer, and intron excision; and correct secondary structure, including TψC and D-loop formation, are required for exportin-T- and exportin-5-mediated nuclear export. By contrast, cistronic tRNA genes in yeast, which have not undergone intron excision in the nucleus, may be exported by homologs of exportin-T/5, Los1, and Msn5, for further processing by the TSEN complex on the cytosolic leaflet of the outer mitochondrial membrane (67).

Mature tRNA species containing 3′ CCA nucleotides act as substrates for aminoacylation by aminoacyl-tRNA synthetases (aaRSs), recognition by eEF2a, and subsequent incorporation in the ribosome complex to participate in translation elongation (36) (Figure 1). Proper decoding of the genome requires highly accurate pairing of amino acids to cognate isoacceptor tRNA families by aaRSs, which is achieved by 20 different aaRSs that are specialized for each of the 20 canonical amino acids. These aaRSs can be subdivided into two general classes, class I and class II, based on the primary and secondary structures of the catalytic sites mediating tRNA aminoacylation. Coupled with this action, all three phylogenetic domains express a subset of aaRSs with domains that can survey and correct misincorporation of amino acids to their cognate tRNA (122). Loss of these proofreading and editing functions results in decreased fidelity of polypeptide primary structure and is implicated in human disease (96).
The proper recognition and charging of tRNA isoacceptor classes depend on interactions between the anticodon loops and acceptor stem elements with the catalytic site of the cognate aaRS (20, 25, 104, 137). The discriminating base (located proximal to the 3′ terminal CCA) and the so-called wobble base (a base modification at position 34 in the anticodon loop) represent two critical and highly conserved structural components present in all isodecoders of a specific isoacceptor family. Studies in *Escherichia coli* have shown that tRNA species with alterations in these sites have a reduced charging capacity and increased rates of degradation, reducing the available mature species for participation in translation (20).

2.4. Mitochondrial tRNA Processing

The mitochondrial genome in mammals is limited, encoding two ribosomal subunits, 22 mt-tRNA genes, and 11 of the 13 proteins necessary for oxidative phosphorylation. Genes associated with mt-tRNA processing and maturation are encoded in the nucleus and require localization to the mitochondria (142, 176). After the necessary mt-tRNA maturation machinery is localized to the mitochondria, the maturation processes of nucleus- and mitochondrion-encoded tRNA genes have many parallels.

Processing of the 5′ leader and 3′ trailer sequences is mediated by nucleus-encoded RNase P and ELAC, respectively. The RNase P isoform responsible for processing of pre-mt-tRNAs consists of protein subunits and lacks any structural or catalytic ribozyme (189). While low levels of the nucleus-encoded H1 RNA have been isolated from mitochondrial extracts, it has been suggested that the catalytic activity of the mitochondrial RNase P dominates mt-RNA 5′ leader removal (143). As mentioned above, ELAC2, the long isoform of ELAC, appears to be the only reported pathway mediating enzymatic removal of the 3′ trailer sequences of mt-tRNA. This is in contrast to nucleus-encoded tRNA, which has both ELAC1 and Rsp1p available in the La-dependent and La-independent pathways (22). Addition of the terminal 3′ CCA to mitochondrial RNA is mediated by the same isoform of TRNT1 as cytosolic tRNA species (110).

Aminoacylation of the 22 human mt-tRNAs is achieved by 20 nucleus-encoded mitochondrial aaRSs (mt-aaRSs) with varying sequence and length homology to their cytosolic counterparts (127). A hallmark difference between these two sets of aaRSs is the ability of the majority of mt-aaRSs to charge cytosolic tRNAs with their cognate amino acids, offering a redundancy of aminoacylation machinery, while cytosolic aaRSs do not have the ability to charge mt-tRNAs. This difference is thought to result from the noncanonical secondary structures that mt-tRNAs can adopt, reducing the binding efficiency to the specific catalytic domain in the cognate cytoplasmic aaRSs (166). Due to the lack of redundancy, most mutations characterized as imparting a negative effect on the aminoacylation of an mt-aaRS to a cognate tRNA are linked to human disease (165, 171).

2.5. tRNA Degradation Pathways

The extensive secondary and tertiary structures of functional tRNA molecules confer tremendous stability, making tRNA one of the most stable species of RNA. Studies in bacteria have demonstrated that tRNAs are markedly more stable than mRNAs (35). The reported mRNA half-lives range from minutes to an average of several hours in different cellular contexts, whereas the reported half-lives for species of mature tRNA range from many hours to several days (82, 156). This stability can be partly attributed to the extensive secondary and tertiary structures inherent to functional tRNA species. Conformations altered via changes in primary sequence or modifications subject the aberrant transcript to turnover through normal degradation pathways, safeguarding functional tRNA pools (4, 194).
As mentioned above, in yeast, the presence of an oligouridylic acid [oligo(U)] suitable for La protein binding or methylation at the A58 site (m\(^1\)A58) reduces the rate of nuclear polyadenylation and 3′-to-5′ degradation by the TRAMP4 complex (4). Nuclear surveillance represents a well-characterized pathway for removing malformed tRNA species from the available tRNA pools. Studies have suggested that the rapid tRNA decay pathway, which is mediated by Rat1 and Xrn1 and exhibits 5′-to-3′ exonucleolytic activity, is a second and redundant program to further protect tRNA pools and decoding fidelity (4, 194). This pathway selectively degrades hypomodified nascent pre-tRNA transcripts or species with mutated acceptor arm sequences, resulting in a noncanonical tRNA structure. In yeast, a failure of this pathway to properly clear mutated tRNA transcripts can result in tRNA accumulation, failure of normal decoding, and amino acid misincorporation.

The above pathways are responsible for clearing tRNA in the nucleus; however, the mechanisms regulating the turnover of damaged cytosolic mature tRNAs are not well defined. Current working hypotheses include the ideas that species are marked with damaged 3′ CCA sequences by TRNT1 and that retrograde transport to the nucleus occurs for degradation in the above-mentioned nuclear surveillance pathways (196). If mature tRNA loses the stability of its secondary or tertiary structure through hypomodification or is unable to participate in aminoacylation, these transcripts may act as a substrate for cytosolic TRNT1, which adds a second CCA to the 3′ end of the damaged transcript, resulting in a 5′-NCCACCA-3′ or 5′-NCCACC3′ following isomerization of the acceptor arm. This process has been observed in a subset of tRNA species with a 5′ GG and a slightly unstable acceptor stem caused by a G-U wobble (196).

In response to external stimuli and environmental stress, eukaryotic cells have evolved discrete pathways for rapid tRNA degradation. Angiogenin, a member of the RNase A family, was one of the first endonucleases shown to cleave mature tRNA at the anticodon loop in response to environmental stress (49, 177) (Figure 1). It appears that, due to the broad binding specificity of angiogenin, species of mature tRNA from all isoacceptor families act as substrates for cleavage, resulting in the formation of distinct 5′ and 3′ tRNA halves (tiRNAs). These species of tiRNA have exhibited various effects on cellular metabolism, proliferation, and cell death. In several studies, 5′ and 3′ tiRNAs isolated from mammalian cells demonstrated phospho-eIF2α-independent repression of translation in both human and mouse cells in vitro (72, 200) (Figure 1). This repression could be explained by a global reduction in mature tRNA levels; however, the tiRNA levels did not appear to rise above 1% of the levels of their full-length counterparts. The proposed mechanism for translation repression includes binding of eIF4F, inhibiting interactions with the 5′ m\(^7\)G cap on mRNA. This inhibition lowers the recruitment of the 40S ribosomal subunit and translation initiation globally. It is apparent that mammalian cells have evolved the production of tiRNA species to conserve resources during times of cellular stress; however, a second study suggesting inhibition of the intrinsic apoptotic pathway via association with cytochrome c reinforced the notion that tiRNAs act as pleiotropic bioactive species responsible for maintaining cell viability (147) (Figure 1). Importantly, various studies have suggested that tiRNAs have more targeted microRNA effects and undergo retrograde transport to the nucleus, with the potential to regulate gene expression through mechanisms yet to be elucidated (88).

### 2.6. tRNA Modifications

Mature tRNA is one of the most abundantly modified species of RNA, with more than 100 unique modifications identified that have varying effects on structure and function (99, 199). These modifications are enzymatically introduced throughout the transcript’s life cycle, from nascent pre-tRNA to mature aminoacylated forms. The functional importance of these modifications ranges...
from normal processing and proper formation of secondary and tertiary structures to the reinforcement of base–base interactions, increasing the fidelity of decoding. Hypomodification of tRNA has increased the rate of ribosomal pausing during translation, effectively altering gene expression (108, 113).

Early in pre-tRNA processing, m^1A58 results in different effects across species. In *S. cerevisiae*, strains deficient in TRM6 displayed decreased levels of m^1A58-modified tRNA^Met^ (78, 190). As mentioned above, hypomodification at this site increased the degradation of this initiator tRNA by the TRAMP4 complex as part of nuclear surveillance, resulting in slowed growth patterns compared with wild-type strains. m^1A58 hypomodification studies in human cells did not recapitulate the finding of decreased tRNA^Met^ stability; rather, hypomodification resulted in decreased charging efficiency by the cognate aminoacyltransferase (146). This finding suggests that the function of the highly conserved m^1A58 modification changed from altering transcript stability in yeast to acting as a rheostat for translation efficiency in humans. The studies in human cells demonstrated an association of m^1A58 with polysomes in select isoacceptor families (146), providing an example of a singular tRNA modification exhibiting a modulation of functionality under different cellular contexts.

The formation and maintenance of the proper secondary cloverleaf and tertiary L-shaped structure of mature tRNA are achieved in part through base modification. A subset of base modifications have ostensibly evolved so that they can be incorporated at specific base sites to increase the strength of intrastrand interactions that favor these canonical structures and reduce the avidity of intrastrand interactions that favor conformations with lower functionality (73, 97). Pseudouridine (ψ) modification at position 55 in the TψC-loop represents one of the most ubiquitous RNA modifications across all phylogenetic domains and is found consistently on the mature functional tRNA of virtually all known isoacceptor families (62, 120, 139). The intrastrand interactions with ψ-modified uracil are critical to the formation of the L-shaped tertiary structure of mature tRNA and are thought to be mediated through the increased hydrophobic properties of ψ-modified uracil over its unmodified counterpart (48). These properties increase the base-stacking characteristics of nucleotides and help stabilize the RNA–RNA homoduplex present in the distal region of mature tRNA species.

Nucleotide modifications in the anticodon loop influence base–base interactions and are critical to proper decoding in some isoacceptor families. Bases 34 and 37 represent the most frequent locations for these modifications. In yeast, queuosine (Q) modification of the guanosine in position 34—the wobble base of the anticodon sequence—in tRNA^Ψ^QUA bolsters base–base interactions with the UAC codon in mRNA, effectively reducing the frequency of frameshifting of the P-site (182, 183). Reducing the frameshifting frequency may increase the efficiency of translation over this codon, altering gene expression. Q34, in particular, serves a second purpose of inhibiting degradation by angiogenin in vitro (191), potentially increasing tRNA abundance and reducing the formation of 5′ tRNA in the tRNA^Ψ^QUA isoacceptor family. The effects of increased decoding fidelity of transcripts containing this codon in certain codon contexts illustrate the pleiotropic effects of tRNA hypomodification at a single nucleotide position.

3. tRNA-MEDIATED REGULATION OF mRNA AND TRANSLATION

In addition to more canonical mechanisms regulating gene expression, studies in *S. cerevisiae* have demonstrated that the specific codon used to decode a given amino acid can affect gene expression levels (13, 130). Differential usage of codons that decode the same amino acid (called synonymous codons) affects the rate of mRNA degradation and stability. More stable mRNA transcripts have longer half-lives and can participate in translation for longer durations, increasing gene expression
Effects of codon optimality on mRNA stability. (a) mRNA transcripts enriched in optimal codons have an increased rate of translation elongation and increased transcript stability, which results in increased PABP association with the 3′ poly(A) tail and reduced Dhh1 recruitment to ribosomes. (b) mRNA transcripts enriched in nonoptimal codons are associated with increased Dhh1 recruitment. Dissociation of PABP from the 3′ poly(A) tail exposes the transcript to deadenylation mediated by the PAN2/PAN3 and Ccr4/NOT complexes. Deadenylation is associated with the recruitment of the decapping proteins Dcp1 and Dcp2, which remove the m^7^-guanosine (m^7^G) cap from the 5′ end of the transcript. The exposed 5′ end is then degraded by the 5′-to-3′ exonuclease Xrn1.

Figure 2

Studies have linked the stability of a given transcript to the rate of translation elongation sensed by the highly conserved eukaryotic DEAD-box helicases Dhh1 in yeast and DDX6 in humans (30, 66, 100, 133). Usage of synonymous codons that increase the rate of translation elongation sensed by Dhh1 (called optimal codons) stabilizes mRNA transcripts, resulting in longer half-lives; usage of codons that decrease the rate of translation elongation sensed by Dhh1 (called nonoptimal codons) destabilizes mRNA transcripts, resulting in shorter half-lives.

Further studies in yeast have demonstrated that mRNA transcripts enriched in nonoptimal codons are subject to elevated rates of degradation, primarily through a 5′-to-3′ exonucleolytic pathway (126). Initially, transcripts enriched in nonoptimal codons undergo enzymatic shortening of the 3′ poly(A) tail (deadenylation) (193). The precise mechanisms regulating deadenylation remain ambiguous; however, an incomplete and simplified model suggests that dissociation of poly(A)-binding proteins (PABPs) from the 3′ poly(A) tail exposes transcripts to 3′-to-5′ deadenylation, mediated primarily by the exonucleolytic activity of the PAN2/PAN3 and Ccr4/NOT complexes (119, 203). Dcp1 and Dcp2 are recruited to deadenylated transcripts and mediate removal of the 5′ m^7^G cap. This decapping event inhibits the recruitment of the eIF4G complex and 40S small ribosomal subunit necessary for translation initiation and subjects these transcripts to degradation by the 5′-to-3′ exonuclease Xrn1. Degradation occurs cotranslationally because the decay machinery, including Dhh1 and Dcp1/2, is associated with the accumulation of multiple ribosomes (polysomes) on a given transcript (66, 126).

While the effects of nonoptimal codon usage on mRNA stability and gene expression are well understood, the mechanisms that govern codon optimality in the context of translation elongation are unclear. Mature tRNA abundance is one of many proposed key regulators of translation elongation (and, therefore, codon optimality) (90, 132, 158, 211). Studies in mammalian cells have demonstrated that altered levels of tRNA abundance are associated with normal and abnormal...
cellular states. In human tissue samples, Gingold et al. (53) observed discrete mature tRNA abundance profiles during states of cellular proliferation or differentiation. The abundance of a particular tRNA isoacceptor species, which decodes a single codon–anticodon interaction, directly correlated with the synonymous codon usage biases observed in matched RNA-sequencing experiments, suggesting that mRNA and tRNA transcription work in concert as defined programs controlling cell fate.

These findings also reinforce studies that directly link the effects of codon optimality to mRNA stability in human cells. Because mutations that alter mRNA transcription affect cellular function and human disease pathogenesis, alterations in mature tRNA abundance should presumably also affect cellular proliferation and differentiation. This concept of tRNA expression affecting cellular identity is further supported by studies showing that elevated levels of tRNA\textsubscript{Glu,UUC} increase the proliferative and metastatic potential of benign human breast tissue cell lines through altered gene expression in vitro, a finding that was recapitulated in fixed malignant breast tumor samples (57). In addition to the effects of altered tRNA expression in adult breast cancer, germline mutations affecting tRNA maturation and function frequently manifest in neural tissue (Table 1). Mutations in the central-nervous-system-specific isodecoder for (AA), n-Tr20, are associated with increased ribosomal pausing and widespread neurodegeneration in the absence of GTPB2 in mice (71). The normal function of GTPB2 is to dissociate stalled ribosomes with a vacant A-site from mRNA, allowing the recovery of translation elongation. These studies suggest that functional mature tRNA abundance may participate in determining a codon’s optimality by acting as a regulator of ribosomal A-site vacancy.

4. tRNA INVESTIGATIVE CHALLENGES

Despite being one of the earliest known classes of noncoding RNA molecules, tRNAs are still extraordinarily difficult to study. The challenge is due largely to the RNA secondary structure, chemical modifications to constituent nucleotides, and copy number and sequence conservation of tRNA genes within the genome (27, 98). As the number of RNA-sequencing data sets increases, it is becoming more important to annotate and quantify tRNA genes in order to have an accurate representation of the full transcriptome.

Prior attempts at tRNA quantitation were based on northern blot and microarray analysis. However, limited depth, coverage, and throughput have hindered the application of these approaches. For example, microarrays require the development of custom arrays with probes that are specific for each organism and uniquely detectable tRNA species (39, 181). Furthermore, the utility of microarrays for quantitative purposes is constrained by their sensitivity. The need for high-throughput, quantitative approaches to measure tRNA dynamics genome-wide has led to a recent investment in the development of methods based on next-generation sequencing (45, 55, 121, 161, 210).

Most next-generation sequencing technologies necessitate the conversion of RNA to cDNA, the ligation of adapter sequences to facilitate binding and multiplexing, and PCR amplification. However, tRNAs are heavily structured, with more than half of the bases participating in stem-loop base-pairing, impeding linearization and efficient cDNA conversion. This structure is thought to render tRNAs exceptionally stable, although individual tRNA half-lives can be highly variable and dependent on both cellular context and posttranscriptional processing events (4, 19). tRNA molecules are among the most chemically modified RNAs, but little is known about the abundance and role of each modification. The modifications are extensive and contribute to tRNA maturation, function, and stability (83). Unfortunately, tRNA chemical modifications block reverse-transcriptase-mediated cDNA synthesis, precluding the efficient preparation of
next-generation sequencing libraries and leading to truncation products (121). Together, these features prevent tRNAs from accurately being represented in next-generation sequencing libraries.

To overcome these challenges, tRNA fragmentation, novel adapters, and engineered reverse-transcriptase and demethylase enzymes have been employed to characterize tRNA gene expression, pre-tRNA and mature tRNA abundance, chemical modification events, and the order of addition from bulk cell populations (45, 55, 121, 161, 210). In a landmark study, Zheng et al. (210) utilized an engineered AlkB bacterial methylase to remove the majority of m1-adenine (m1A), m3-cytosine (m3C), and m1-guanosine (m1G) methylation marks from tRNAs to improve reverse-transcriptase processivity during cDNA synthesis. Furthermore, they used a thermostable group II intron reverse transcriptase (TGIRT) to add next-generation sequencing adapters via template switching rather than ligation, enhancing the efficiency of this step over traditional approaches. Their method led to substantial improvement in the full-length cDNA conversion of mature tRNA transcripts from HEK293T cells and high correlation of tRNA levels detected by high-throughput sequencing compared with those detected with a microarray. However, some limitations remained: The demethylase mixture was not effective at removing m2G or modifications other than methylation, and the adapter sequences promoted a template-switching reaction in mature tRNAs containing a 3′ CCA, instead of in pre-tRNA transcripts. Furthermore, because large quantities of RNA from cell cultures were collected, whether this protocol can be adapted for low amounts of RNA and/or tissue is unclear. Evans et al. (45) adapted this approach to quantify the levels of charged and uncharged mature tRNAs by adding a chemical step to remove the 3′ A nucleotide from uncharged tRNA prior to deacylation and library preparation. The levels of charged tRNAs can be inferred computationally by comparing the ratio of A+ to A−-ending tRNAs and the quantity of each species.

To enable accurate tRNA quantification using conventional RNA-sequencing methods from low levels of purified RNA, Shigematsu et al. (161) developed Y-shaped adapter-ligated mature tRNA sequencing (YAMAT-seq). In this technique, a unique Y-shaped adapter is annealed and ligated to deacylated total RNA by Rnl2, significantly increasing efficiency over the Rnl1-mediated ligation of linear adapters often used in RNA-sequencing protocols. cDNA synthesis and amplification can then be performed using a standard small-RNA sample preparation kit. This approach is more convenient than methods that require multiple gel purification steps or polyadenylation, detects both cytoplasmic tRNAs and mt-tRNAs, and correlates sequencing results with abundance via northern blot analysis. It can also be used with as little as 300 ng of total RNA. However, it does not contain nucleotide modification removal steps to address the effect of mature tRNA modifications on cDNA synthesis, and it cannot detect specific tRNAs. As in other approaches, RNA is collected from cell cultures as opposed to tissues, and the ligation reaction preferentially targets mature tRNAs containing 3′ CCA over pre-tRNA transcripts.

To annotate tRNA genes, quantitatively measure pre-tRNA gene expression and mature tRNA abundance, and identify modified nucleotides, Gogakos et al. (55) performed partial alkaline hydrolysis and sequencing of size-selected RNAs containing full-length tRNAs (hydro-tRNAseq). The resulting tRNA fragments were less structured and had fewer modifications per fragment, enabling efficient cDNA conversion using conventional small-RNA-sequencing library preparation approaches. After combining this data set with photoactivatable ribonucleoside-enhanced cross-linking and immunoprecipitation (PAR-CLIP) of SSB/La, the first protein to interact with pre-tRNAs following transcription, they were able to define pre-tRNA leader, trailer, and intron sequences in addition to the order of nucleotide modifications at higher resolution than prior studies (55, 178). Importantly, their data showed that pre-tRNA and mature tRNA abundance were not correlated and that alternative approaches to assessing POLR3 occupancy or tRNA gene count as a proxy for expression or abundance are inaccurate, at least in humans. Furthermore, the
study curated a reference database of sequencing mismatch signatures for modified nucleotides following cDNA synthesis that serves as an important resource for defining the location, spatial addition, and temporal addition of tRNA modifications (55). As with other approaches, this technique requires large quantities of purified RNA and multiple rounds of gel purification for sequencing library construction, and because the study isolated RNA from immortalized human cells, its utility for tissue samples or low amounts of RNA remains untested. Furthermore, because the detection of a modified tRNA position was tested at a mismatch frequency threshold of greater than 10%, rare or infrequent modification events are likely absent from this data set.

As mentioned above, efficient next-generation sequencing of full-length tRNA transcripts requires removing chemical modifications, which masks the position, identity, and abundance of these important posttranscriptional regulatory events. Because of this limitation, it has not been possible to map or measure the abundance of tRNA modifications on a global scale. The most accepted and accurate approach to mapping tRNA modifications is liquid chromatography tandem mass spectrometry (87, 140). RNase enzymes have been engineered to enable complete coverage with a single enzymatic digestion (169). A recent alternative approach to measuring tRNA modification levels at nucleotide resolution is comparative methylation analysis through AlkB-facilitated RNA methylation sequencing (ARM-seq) (32), although this technique is currently limited to methylation at tRNA positions accessible to the AlkB enzyme. Although preliminary, nanopore-based sequencing methods have also been adapted for tRNA sequencing (167). With the distinct advantages of direct RNA sequencing and detection of modifications at base resolution on individual full-length transcripts, this approach may be promising after further optimization of base calling, adapter design, and RNA sample requirements.

There are as many as 614 predicted tRNA genes and 176 tRNA pseudogenes or tRNA-derived repetitive elements encoded in the nuclear genome for the 61 codons (27). Because most of these genes are multicopy genes that evolve together at specific loci (138), tRNA paralogs and pseudogenes have nearly identical genomic sequences. Furthermore, the pre-tRNA-processing steps of leader, trailer, and intron removal and the steps of 3' CCA addition and chemical modification can affect mapping by introducing gaps and mismatches, respectively. Computational tools designed for differential expression analysis commonly discard tRNA reads due to multimapping and mismatch threshold limits. These features complicate tRNA annotation and preclude accurate mapping of tRNA-sequencing reads, and specialized mapping strategies are therefore needed to analyze tRNA-sequencing data. Novel computational approaches have been developed to overcome this challenge by parsing tRNA-sequencing data into informative fractions for interpretation using iterative processes (55, 63). Best-practice pipelines have been designed that use simulated and real tRNA-sequencing data to measure each desired tRNA read fraction (pre-tRNAs and mature, charged, and modified tRNAs) (45, 55, 63). As an example, initial mapping of tRNA-sequencing reads can be performed on sequences from a curated tRNA database (27). tRNA reads containing shared mismatches that encompass a significant percentage (e.g., 10%) of the population can be compiled and added to the annotation set as a predicted modified tRNA. This process can be run iteratively to create a reference of sequencing reads corresponding to modified tRNAs (55). Alternatively, pre-tRNAs can be mapped and discarded (or separately analyzed) before mature tRNAs are uniquely mapped to clustered mature tRNA reference sequences with a 3' CCA tail (63). Mismatches from this alignment are annotated as tRNAs likely to contain modifications (63). Both of these approaches identify tRNA modifications with high confidence (55, 63) but are limited to cDNA synthesis events that cause nucleotide transversions or transitions. For instance, ribose methylation and pseudouridylation modifications would require specific chemical treatment of RNA before they could be detected by next-generation sequencing techniques. These computational approaches work best for full-length tRNAs, since
the probability of multimapping is anticorrelated with the length of the fragment, and analysis of short tRNA fragments would therefore likely have limited success.

Over the past five years, significant technical and computational advances have improved our ability to detect and quantify tRNA transcripts by next-generation sequencing methods. However, much more work will be necessary to map and quantify all tRNA processing, modification, charging, and degradation events and to reveal their functional significance for normal and diseased cellular states.

5. tRNA-ASSOCIATED HUMAN NEURODEVELOPMENTAL DISORDERS

tRNA-associated neurodevelopmental disorders are caused by mutations in tRNA genes or genes that regulate tRNA biogenesis or function, and are often congenital and degenerative in nature. Because the genes that encode and regulate tRNAs are located within the nuclear and mitochondrial genomes, this broad class of disorders can be inherited in an autosomal recessive, autosomal dominant, X-linked, or maternal pattern.

The recent adoption of next-generation sequencing in the clinic has exponentially improved our ability to simultaneously diagnose patients with developmental brain disorders and discover new candidate variants and genes underlying these conditions. The standard of care for patients with suspected genetic disease is clinical genetic testing. Clinical testing for variants in causative genes established by Sanger sequencing is accurate and cost effective for disorders with a limited genetic spectrum. For genetically or clinically heterogeneous disorders, clinical exome sequencing with comparative genomic hybridization arrays or genome sequencing can identify candidate mutations across a large number of genes, discover structural variants, and diagnose patients (40). Primarily due to limitations of cost and variant interpretation, the majority of disease-causing variants have been found within the exome, the protein-coding region of the genome (28).

Likewise, our knowledge of pathogenic variants that affect tRNA biogenesis and function reflects our ability to detect and interpret variants within the nuclear and mitochondrial genomes. While there are undoubtedly other variants in tRNAs and small noncoding genes that lead to disease and remain to be discovered, the currently known pathogenic mutations affect four broad categories of tRNA expression and function: mt-tRNA gene deletions or mutations, tRNA biogenesis, tRNA charging, and tRNA modification (Table 1).

5.1. Mitochondrial Encephalomyopathies

Mitochondrial encephalomyopathies encompass a clinically heterogeneous group of disorders caused by defects in mitochondrial function, with an estimated prevalence of 1 in 5,000 individuals (34, 150). Mutations in mt-tRNAs or proteins that regulate mitochondrial DNA integrity often lead to a depletion in mt-tRNA genes and impaired protein synthesis necessary for ATP production. Patients can be diagnosed with Kearn’s–Sayre syndrome; myoclonic epilepsy with ragged red fibers (MERRF); Pearson syndrome; Leigh syndrome; mitochondrial complex deficiency; or mitochondrial myopathy, encephalopathy, lactic acidosis, and stroke-like episodes (MELAS). These disorders are often maternally inherited, and the phenotypic severity and age of onset are dictated by mitochondrial heteroplasmy levels. The shared clinical features of encephalopathy and myopathy are observed upon magnetic resonance imaging (MRI) and muscle biopsy, respectively. These features can be accompanied by additional clinical phenotypes, including short stature, eye involvement, diabetes mellitus, seizures, ataxia, heart block, high lactic acid levels, hearing loss, and cognitive impairment. Patients are diagnosed molecularly after genetic testing for known causative
gene mutations and subsequently treated with drugs to enhance mitochondrial respiratory chain function, such as CoQ10 and synthetic derivatives, with variable success (54).

5.2. Charcot–Marie–Tooth Disease

Charcot–Marie–Tooth disease is a prevalent neurological disorder (affecting 1 in 1,000 individuals) that was first described in 1886 by three physicians for whom the disease is named (144). The clinical subtypes are defined by the underlying cellular cause (axonal degeneration or loss of white matter) and genetic cause. Subtypes of Charcot–Marie–Tooth disease type 2 are characterized by peripheral axonal neuropathy caused by mutations in aaRS proteins, which regulate tRNA charging. Alternative causative mutations in proteins that have no function related to tRNA biogenesis have also been described for this clinical subtype. Most patients do not present with structural brain abnormalities; however, severe recessive and X-linked forms of the disease lead to cognitive impairment and loss of white matter within the central nervous system, which is observed as predominantly posterior lesions in T2-weighted MRI (92, 124). The disease typically first manifests in adolescence but can also occur shortly after birth. Patients are often treated with orthopedic devices and pain medication, but there are no cures or effective therapies to halt disease progression.

5.3. Leukodystrophy

Like some clinical subtypes of Charcot–Marie–Tooth disease, leukodystrophy is caused by the abnormal development or loss of oligodendrocytes, the specialized neuronal support cells that produce myelin sheaths. While most tRNA biogenesis genes affect neurons, leading to neurodevelopmental brain phenotypes, mutations in aaRS genes often lead to loss of oligodendrocytes and neurodegeneration. Patients are phenotypically similar to those with Pelizaeus–Merzbacher disease and typically present shortly after birth with progressive loss of vision, hearing, coordination, and cognitive abilities (116). Diagnostic MRI shows hypermyelination together with variable cerebellar atrophy and hypoplastic corpus callosum. The variable neurological features can include microcephaly and seizures. Leukodystrophies are caused by mutations in tRNA biogenesis genes that are inherited in an autosomal recessive manner. Standard therapies for patients are based on symptomatic treatment and support, but there are no therapies that slow progression or offer a cure. Mutations in RNA polymerase III subunits, which are required for tRNA transcription, can also cause leukodystrophy, suggesting that oligodendrocytes are particularly sensitive to levels of tRNA expression and function.

5.4. Pontocerebellar Hypoplasia

Pontocerebellar hypoplasia (PCH) is a rare, severe prenatal/perinatal brain disorder caused by the segmental degeneration of the cerebellum, pons, and brain stem. The 11 clinical subtypes (PCH1–11) are defined by the co-occurrence of the neurological abnormalities with additional unique features, such as ambiguous genitalia (PCH7), spinal anterior horn disease (PCH1), and chorea/dyskinesia (PCH2), in addition to the variable shared features of progressive microcephaly and seizures. PCH is predominantly caused by biallelic mutations in RNA-processing proteins and is inherited in an autosomal recessive manner. PCH1, -2, -4, -5, and -10 are caused by mutations in the TSEN subunits and the associated protein, CLP1, which is required for intron-containing pre-tRNA splicing (Table 1). Interestingly, genotype–phenotype correlations in patients harboring mutations in this complex show varying clinical features and phenotypes, even among patients...
with mutations in the same causative gene (112). These findings suggest that the variant severity and/or genetic modifiers of this molecular pathway contribute to the clinical presentation of patients and should be taken into consideration for diagnosis. Mechanistically, these results imply that the cerebellum and Purkinje cells are particularly vulnerable to reduced levels of functional tRNAs during development; however, it is not yet clear why other cell types are not equally affected. Individual cell types may have differential expression of intron-containing tRNAs (39), or compensatory mechanisms may be in place in unaffected cell types, but these hypotheses have not been tested.

In addition, PCH6 is caused by mutations in the arginyl-tRNA synthetase gene RARS (Table 1). PCH6 patients show overlapping features with those diagnosed with mitochondrial encephalomyopathy, as would be expected based on the shared cellular mechanism of mitochondrial dysfunction in both diseases. There are no effective therapies for PCH. Patients often receive support to manage their symptoms and control seizure frequency.

5.5. Intellectual Disability Syndromes

Intellectual disability is a common, genetically complex disorder with an estimated prevalence of 1–3 in 100 children born. Almost 900 causative genes and loci have been mapped, but these account for only 20% of the genetic cases, leaving many genetic causes yet to be discovered (44). A small subset of the identified genetic variants have been found in genes that regulate tRNA biogenesis and modification (Table 1). These variants cause the rare, clinically defined disorders of Galloway–Mowat and Usher syndromes, in addition to other, less well characterized intellectual disability (mental retardation) disorders (Table 1). These syndromes are caused by biallelic mutations and are inherited in a Mendelian, autosomal recessive manner. They share the diagnostic criteria of reduced cognitive ability (IQ < 70), diagnosed in childhood. Each disorder has additional unique clinical features, such as microcephaly and renal dysfunction in Galloway–Mowat syndrome and vision and hearing loss in Usher syndrome. Due to the severe impact of structural brain defects on cognitive abilities, children with structural brain abnormalities commonly present with intellectual disability. Interestingly, mutations in proteins that chemically modify tRNA transcripts cause intellectual disability syndromes, often without associated structural brain defects, suggesting that regulation of tRNA activity is particularly important for neuronal connectivity and function. Presently, intellectual disability patients are treated through behavioral modification programs, and there are no effective pharmacological therapies known to improve cognitive function.

6. DIAGNOSTIC AND THERAPEUTIC APPROACHES

6.1. tRNA Nonsense Suppressors

Germline mutations that convert an amino acid codon to a premature termination codon (PTC), also known as a nonsense mutation, drive the pathogenesis of a host of heritable diseases. Recurrent nonsense mutations have been identified in the chloride channel CFTR in cystic fibrosis, the structural sarcomere protein dystrophin in myotonic dystrophy, and the sodium channel SCN1a in Dravet syndrome. The polypeptide products translated from mRNA transcripts containing these PTCs are truncated and typically lack the functionality of the full-length gene product. In addition, eukaryotes have evolved the nonsense-mediated decay pathway to conserve translation machinery through the detection and degradation of transcripts containing PTCs. These compounding mechanisms result in an added decrease in the expression of these critical genes.
Rather than correcting nonsense mutations with gene-editing technologies, another proposed therapeutic strategy to mitigate the deleterious effects of nonsense mutations is increasing the ribosomal read-through of the PTC. This is achieved through the endogenous expression or exogenous introduction of tRNA species capable of decoding the PTC with an amino acid. Given that the mutated mRNA transcript contains all necessary information for the gene product, read-through of the codon should result in a full-length polypeptide with random incorporation of an amino acid at the PTC site. tRNAs capable of decoding PTCs are referred to as nonsense suppressors (i.e., these tRNA species suppress the premature termination of translation by the nonsense mutation).

The concept of an expanded genetic code through the introduction of exogenous tRNA species that can decode a stop codon has been utilized in laboratory studies for decades (117). This technique is utilized to incorporate noncanonical amino acids into a specified location of a polypeptide, typically by decoding the amber stop codon (UAG). Briefly, the requirements for this process include expression of the exogenous tRNA and a cognate aminocycltransferase capable of charging the above-mentioned tRNA species with the chosen noncanonical amino acid. Nonsense suppressor therapy is similar to noncanonical amino acid incorporation, but the suppressor tRNA has fewer restrictions, with the ability to utilize any endogenously expressed cytosolic or mitochondrial aminocycltransferases (11). Appropriate charging of the exogenous nonsense suppressor tRNA and successful decoding of the premature stop codon offers the potential to regain sufficient functional protein expression to ameliorate the disease phenotype.

6.2. Novel Antimicrobial Development

A potential avenue for therapeutics related to tRNA biogenesis is the development of new antimicrobials. As mentioned above, the machinery necessary for tRNA processing and maturation is essential and functionally represented in all known forms of life. While orthologous enzymes are functionally similar, the structures of these orthologs can vary greatly, offering novel targets for therapeutic development capable of treating drug-resistant bacterial and fungal infections. Previous studies demonstrated the specificity of the aminoglycoside neomycin B for bacterial RNase P (tRNA 5' leader removal), suggesting this holoenzyme as a potential target for the development of directed inhibitors (106). Recently, small molecules targeting the peptide subunits of the RNase P ortholog expressed in methicillin-resistant *Staphylococcus aureus* have demonstrated selective activity and the potential for use in implantable devices and as a stand-alone therapy (43).

6.3. Identification of Unique Biomarkers

The recent development of high-throughput tRNA-sequencing methodologies coupled with the increased stability conferred to tRNA by their structure and nucleotide modification offers the potential for the discovery of unique biomarkers related to human health and disease. Previous studies demonstrated increased expression of specific isodecoders and tRNAs or tRNA fragments in several human malignancies (57, 115, 125). Recently published sequencing studies have identified several classes of small noncoding RNAs in clinical samples from healthy patients and have isolated tRNA and tRNA fragments in human serum, urine, and saliva that map to more than 400 human tRNA genes (37, 202). Application of next-generation tRNA-sequencing methodologies has the potential to identify changes in mature tRNA species and tRNAs or tRNA fragments as biomarkers to monitor disease progression and response to therapy. Samples could be obtained with minimally invasive techniques, potentially reducing patient discomfort and minimizing direct patient expenses as the cost of sequencing applications decreases over time.
DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

LITERATURE CITED

17. Bohnsack MT, Czapinski K, Gorlich D. 2004. Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. RNA 10:185–91


79. Kantidakis T, Ramsbottom BA, Birch JL, Dowding SN, White RJ. 2010. mTOR associates with TFIIC, is found at tRNA and 5S rRNA genes, and targets their repressor Maf1. *PNAS* 107:11823–28


192. Whipple JM, Lane EA, Chernyakov I, D’Silva S, Phizicky EM. 2011. The yeast rapid tRNA decay pathway primarily monitors the structural integrity of the acceptor and T-stems of mature tRNA. *Genes Dev.* 25:1173–84


Contents

The Development of Human Genetics at the National Research Centre, Cairo, Egypt: A Story of 50 Years
Samia A. Temtamy ................................................................. 1

Epigenetic Regulation and Risk Factors During the Development of Human Gametes and Early Embryos
Yang Wang, Qiang Liu, Fuchou Tang, Liying Yan, and Jie Qiao ......................... 21

The Genetics of Human Skin and Hair Pigmentation
William J. Pavan and Richard A. Sturm .......................................................... 41

Pregnancy Immunogenetics and Genomics: Implications for Pregnancy-Related Complications and Autoimmune Disease
Hing-Yuen Yeung and Calliope A. Dendrou ................................................... 73

Massively Parallel Assays and Quantitative Sequence–Function Relationships
Justin B. Kinney and David M. McCandlish ..................................................... 99

Advances in the Genetic Basis and Pathogenesis of Sarcomere Cardiomyopathies
Raquel Yotti, Christine E. Seidman, and Jonathan G. Seidman ......................... 129

Consanguinity and Inbreeding in Health and Disease in North African Populations
Lilia Romdhane, Nessrine Mezzi, Yosr Hamdi, Ghada El-Kamah, Abdelhamid Barakat, and Sonia Abdelhak ......................................................... 155

The Future of Genomic Studies Must Be Globally Representative: Perspectives from PAGE

Gene and Induced Pluripotent Stem Cell Therapy for Retinal Diseases
Akiko Maeda, Michiko Mandai, and Masayo Takahashi ................................... 201

Genetic Etiologies, Diagnosis, and Treatment of Tuberous Sclerosis Complex
Catherine L. Salussolia, Katarzyna Klonowska, David J. Kwiatkowski, and Mustafa Sabin ................................................................. 217
Genetic Predisposition to Childhood Cancer in the Genomic Era
Sharon E. Plon and Philip J. Lupo ................................................................. 241

The Genetics and Epigenetics of Facioscapulohumeral
Muscular Dystrophy
Charis L. Himeda and Peter L. Jones ......................................................... 265

Lynch Syndrome: From Screening to Diagnosis to Treatment
in the Era of Modern Molecular Oncology
Stacey A. Cohen, Colin C. Pritchard, and Gail P. Jarvik ............................. 293

Measuring Clonal Evolution in Cancer with Genomics
Marc J. Williams, Andrea Sottoriva, and Trevor A. Graham ........................ 309

Roles of Extracellular Vesicles in High-Grade Gliomas:
Tiny Particles with Outsized Influence
Michael W. Graner .................................................................................. 331

tRNA Metabolism and Neurodevelopmental Disorders
Ashleigh E. Schaffer, Otis Pinkard, and Jeffery M. Coller ............................ 359

Early Lessons from the Implementation of Genomic
Medicine Programs
Marc S. Williams .................................................................................... 389

The Status and Impact of Clinical Tumor Genome Sequencing
Kenna R. Mills Shaw and Anirban Maitra ................................................ 413

The Causes and Consequences of Genetic Interactions (Epistasis)
Júlia Domingo, Pablo Baeza-Centurion, and Ben Lehner ............................ 433

Thinking About the Evolution of Complex Traits in the Era of
Genome-Wide Association Studies
Guy Sella and Nicholas H. Barton .............................................................. 461

Genomic Research Through an Indigenous Lens:
Understanding the Expectations
Nanibaa’ A. Garrison, Māui Hudson, Leah L. Ballantyne,
Ibrahim Garba, Andrew Martinez, Maile Taualii, Laura Arbour,
Adine R. Caron, and Stephanie Carroll Rainie .......................................... 495

International Divergence in Gene Patenting
Dianne Nicol, Rochelle C. Dreyfuss, E. Richard Gold, Wei Li,
John Liddicoat, and Geertrui Van Overwalle ............................................ 519

Errata
An online log of corrections to *Annual Review of Genomics and Human Genetics* articles
may be found at http://www.annualreviews.org/errata/genom