ABSTRACT

tRNA are post-transcriptionally modified by chemical modifications that affect all aspects of tRNA biology. An increasing number of mutations underlying human genetic diseases map to genes encoding for tRNA modification enzymes. However, our knowledge on human tRNA-modification genes remains fragmentary and the most comprehensive RNA modification database currently contains information on approximately 20% of human cytosolic tRNAs, primarily based on biochemical studies. Recent high-throughput methods such as DM-tRNA-seq now allow annotation of a majority of tRNAs for six specific base modifications. Furthermore, we identified large gaps in knowledge when we predicted all cytosolic and mitochondrial human tRNA modification genes. Only 48% of the candidate cytosolic tRNA modification enzymes have been experimentally validated in mammals (either directly or in a heterologous system). Approximately 23% of the modification genes (cytosolic and mitochondrial combined) remain unknown. We discuss these ‘unidentified enzymes’ cases in detail and propose candidates whenever possible. Finally, tissue-specific expression analysis shows that modification genes are highly expressed in proliferative tissues like testis and transformed cells, but scarcely in differentiated tissues, with the exception of the cerebellum. Our work provides a comprehensive up to date compilation of human tRNA modifications and their enzymes that can be used as a resource for further studies.

INTRODUCTION

The acquisition of post-transcriptional chemical modifications is an essential part of the maturation process required to generate functional tRNA molecules (1). Modifications have different roles in controlling stability, folding and decoding properties of tRNAs and can be determinants or anti-determinants for other components of the translation apparatus like e.g. aminoacyl-tRNA synthetases (2,3). In addition, tRNA modifications can be recognition elements of ribonucleases (4), leading to the generation of tRNA fragments that affect multiple cellular processes (5).

However, very few modifications such as m5G37, Ψ55 or t6A37 are present at a specific position of a particular tRNA
in (almost) all known organisms. Most of them are specific to particular taxons, from species to kingdoms. For example, lysidine (k²C34) is a hallmark of bacteria (6), while 1-adenosine (G*) is only found in archaea (7). Depending on the organism, the total number of genes encoding tRNA modification enzymes varies between as little as eleven in some obligate symbionts (8) to around an estimated hundred in humans of which 50 are currently represented in MODOMICS (9).

The near complete sets of tRNA modification genes are currently available for only one organism per domain of life: Saccharomyces cerevisiae for eukarya, where only one gene required for the formation of cm²U out of cm³U is missing (10), Escherichia coli for bacteria where only the genes for ho³U34 and Acp³U47-formation remain unidentified and Haloferax volcanii for archaea where a handful of genes are missing (1). Beyond these three organisms, the annotation of tRNA modification genes remains scarce, because of several issues: First, tRNA modification enzymes are often part of large multifunctional protein families such as the Rossmann Fold Methyltransferase (RFM) superfamily, which can act on other substrates than RNA (11). For example, some RNA methyltransferases are closely related to protein methyltransferases or DNA methyltransferases (11). Second, closely related members of orthologous families often introduce a similar chemical modification, but in different RNAs and at different positions. For example, members of the TrmFO family methylate tRNA or rRNA depending on the organism (12). Third, related enzymes can generate chemically distinct modifications. For example closely related radical-SAM enzymes introduce methyl groups at different positions of nucleosides like in m²A or m³A (13). Finally, the same chemical modification, in particular methyl groups, can be introduced by proteins that are dissimilar (14) or even evolutionarily unrelated having arisen through non-orthologous gene displacements (15). For example, the formation of the universal m¹G37 is catalyzed by TrmD and Trm5, two enzymes of completely different evolutionary origins in bacteria and in eukarya/archaea (16). The combination of these factors has made it difficult to identify enzymes responsible for many tRNA modifications and hence to determine the function of these tRNA modifications in many species including humans.

Recently, an increasing number of mutations causing genetic diseases have been mapped to human genes encoding tRNA modification enzymes (see [17–22] and Table 1), making a comprehensive list of these genes highly desirable. However, to our knowledge, no complete compilation of modifications found in both cytosolic and mitochondrial human tRNAs with their corresponding predicted or validated modification enzymes is available. For mitochondria, the best approximation is a recent list of modifications of bovine tRNAs and the predicted enzymes (23), which has been extrapolated for human tRNAs (23,24). A prediction of human tRNA methyltransferases, based on the known yeast enzymes was performed more than five years ago (25) and was recently extended to homologs of the other yeast RNA modification genes (26). Surveys of specific enzyme families such as the human m¹C methyltransferases (27) or pseudouridine synthases (28) that target tRNA molecules have listed the known and missing genes for these specific modifications. The goal of our analysis was to compile a comprehensive list of known and predicted tRNA modifications in Homo sapiens with genes implicated in their biosynthesis. This analysis allowed for the identification of the remaining gaps of knowledge in the field of human tRNA modifications and will help to guide future experiments. Furthermore, we have used publicly available datasets in order to determine the expression profiles and proteomic evidence of known and predicted modification enzymes. Our work will facilitate access to the current knowledge on human tRNA modification enzymes for a wider community of biologists.

MATERIALS AND METHODS

The set of human isoacceptor tRNAs (i.e. tRNAs that are acylated with the same amino acid regardless of the anticodon sequence) was extracted from the Genomic tRNA Database (GtRNAdb): http://grtnadb.ucsc.edu/ (29) and is summarized here: http://grtnadb.ucsc.edu/genomes/eukaryota/Hsapi19/

All modifications present in the sequences of cytosolic and mitochondrial tRNA of human (H. sapiens), cow (Bos taurus), rat (Rattus norvegicus), and mouse (Mus musculus) were extracted from the MODOMICS database of RNA modification pathways (http://modomics.genesilico.pl) (9). This provided a first list that was then updated with one modification from the literature (m²C34 in Leu-CAA-tRNA) and several human modifications detected with novel tRNAseq methods (30) (m²C20 in Met-CAU-tRNA and m³C47 in Leu-CAG-tRNA and most Ser-tRNAs, m¹Al6 in mito-Arg-TCG-tRNA and m³C32 in mito-Thr-UGU-tRNA and mito-Ser-UGA) that were missing from MODOMICS. The MODOMICS database was updated accordingly.

High-throughput tRNA-seq modification data was derived from published study data sets (30–32). The protein and literature mining tools of NCBI (33) as well as the Uniprot resource and Id/Mapping tools (34) were used to gather data. Gene names were gathered from the HUGO Gene Nomenclature Committee (https://www.genenames.org) (35). Protein interaction data was derived from BioGrid (36) and the predicted mitochondrial localization from MitoCarta (https://www.broadinstitute.org/files/shared/metadata/mitocarta/human.mitocarta2.0.html) (37).

Human co-expression data was extracted from the Search-based Exploration of Expression Compendium (SEEK) database (http://seek.princeton.edu/index.jsp) (38). Phylogenetic trees for specific protein families were extracted from PhylomeDB (http://phylomadb.org) (39). For gene expression analyses, RNAseq data was obtained from the GTEx portal (www.gtexportal.org; GTExAnalysis_2016-01-15_v7_RNASeqCv1.1.8.gene TPM.gct.gz) on 30/04/2018. For each gene the transcript with the highest expression levels was selected for each tissue. Subsequently, relative expression levels were calculated and plotted as a heatmap using the heatmap.2 function in R. Tissues included in the analysis were selected to provide a general
physiological overview. Hierarchical clustering of genes was performed according to similarity of expression profile using Ward’s method (40). For tissue-specific proteomics evidence, we used the human proteome map (http://www.humanprotememap.org/) using the default settings (41). Proteins that were not detected in any tissue were manually removed.

RESULTS AND DISCUSSION

Compiling all mammalian cytosolic tRNA modifications

As a first step to predict the complete set of modification enzymes, we sought to list the nature and positions of all chemical modifications that have been identified in human cytosolic tRNAs. This task is not trivial as the set of human known and predicted tRNA modification genes that have been linked to human diseases

| Modification | Gene | Disease | Cyto. Pheno. | Mito. Pheno. | Article |
|--------------|------|---------|--------------|--------------|---------|
| xG           | THG1L| Microcephaly, developmental delay, nephrotic defect | + | + | (110,175,176) |
| m3G          | TRMT18A| Diabetes, intellectual disabilities, microcephaly, developmental defects | + | + | (111,143,177-180) |
| m4C          | NAT10| Cancer | + | | (112,181,182) |
| ac6C         | THUMP1| Cancer | + | | (133) |
| Gm           | TARBP1| Cancer | + | | (114,115) |
| D            | DUS2| Cancer | + | | (136) |
| Y            | PUS1| Mitochondrial myopathy and sideroblastic anemia (MLASA) | + | + | (117,183,184) |
| m1C          | METTL16| Cancer | + | | (18,119) |
| I            | ADAT3| Intellectual disabilities, microcephaly | + | | (120,185-187) |
| m3C          | NSUN2| Intellectual disabilities, developmental delay, reduced fertility, cancer | + | | (121,170-172,189,229-232) |
| cm5U, Um, Gm, f5Cm, hm4Cm | FTSP1| Intellectual disabilities | + | | (122,123,188) |
| cm5U, Um, Gm, f5Cm, hm4Cm | ELP1| Familial dysautonomia, Charcot-Marie-Tooth disease (CMT), cancer, | + | | (126,127,190) |
| cm5U, Um, Gm, f5Cm, hm4Cm | ELP3| Familial dysautonomia, Charcot-Marie-Tooth disease (CMT), cancer, | + | | (127,130,191,192) |
| cm5U, Um, Gm, f5Cm, hm4Cm | ELP4| Autism spectrum disorder, intellectual disabilities | + | | (128) |
| cm5U, Um, Gm, f5Cm, hm4Cm | ELP5| Cancer, diabetes | + | | (129,193,194) |
| e2U, Um, Gm, f5Cm, hm4Cm | CTU1| Cancer | + | | (127,130,131) |
| e2U, Um, Gm, f5Cm, hm4Cm | CTU2| Microcephaly, nephrotic defect, cancer | + | | (127,130,132,133) |
| e2U, Um, Gm, f5Cm, hm4Cm | MOCS3*| Molybdenum cofactor deficiency | + | | |
| e2U, Um, Gm, f5Cm, hm4Cm | MPS1*| Mammalian lysosomal ATPase (mLAP1), intellectual disabilities | + | | |
| e2U, Um, Gm, f5Cm, hm4Cm | NFSI*| Friedreich ataxia | + | | |
| e2U, Um, Gm, f5Cm, hm4Cm | SERGEF*| Hereditary deafness, artherosclerosis | + | | |
| e2U, Um, Gm, f5Cm, hm4Cm | CIAO1*| Hereditary paraganglioma-pheochromocytoma syndrome, retinitis pigmentosa | + | | |
| e2U, Um, Gm, f5Cm, hm4Cm | NUBP1*| Cancer | + | | |
| e2U, Um, Gm, f5Cm, hm4Cm | ISCU*| Myopathy with lactic acidosis, Friedreich ataxia | + | | |
| I            | ADAT1| Corneal artery disease | + | | (134) |
| m3G, m1I     | TRMT5| Failure to thrive, hypertrophic cardiomyopathy, exercise intolerance | + | + | (135,136) |
| o2Y, yW      | TRMT12| Cancer | + | + | (137,138) |
| o2Y, yW      | LCMT2| Cancer | + | + | (139) |
| e4A          | YRDC| Cancer | + | | (140) |
| e4A          | OSGEP| Galloway-Mowat syndrome, microcephaly, nephrotic defects | + | | (18,141,195-197) |
| e4A          | TP53RK| Galloway-Mowat syndrome, microcephaly, nephrotic defects, cancer | + | | (141,142,195,196) |
| e4A          | TPRKB| Galloway-Mowat syndrome, microcephaly, nephrotic defects | + | | (141,195,196) |
| e4A          | LAGE3| Galloway-Mowat syndrome, microcephaly, nephrotic defects | + | | (141,195,196) |
| m5s5A        | CDKAL1| Diabetes, microcephaly, cancer | + | | (144,198,199) |
| m3C          | TDMDT1| Metabolism, cancer | + | | (145,146) |
| Y            | PUS3| Intellectual disabilities | + | | (147,148) |
| U            | TRMT44| Partial Epilepsy with Pericentral Spikes (PEPS) | + | | (149) |
| m3G          | METTL1| Multiple sclerosis, cancer | + | | (150,200,201) |
| m3G          | WDR4| Microcephaly, cancer, nephrotic defects, developmental defects | + | | (151,202-204) |
| m3U          | TRMT7A| Cancer | + | | (152) |
| Y            | PUS10| Autoimmune diseases, intellectual disabilities | + | | (153,205,206) |
| m3A          | TRMT6| Cancer | + | | (159,154,155) |
| m3A          | TRMT6A| Cancer | + | | (159,154,155) |
| m3C          | NSUN6*| Cancer | + | | (156) |
| m3G,m1A      | TRMT10C| Lactic acidosis, hypotonia, feeding difficulties, deafness | + | | (157,158) |
| m3G,m1A      | HSDD1B10| Neurodegeneration, cardiomyopathy | + | | (158,207,208) |
| m3G          | TRMT1| Intellectual disabilities, microcephaly, + | + | | (159,209,210) |
| f3Cm         | NSUN3| Cancer | + | | (160,161) |
| tm3U         | GTTPB3| Mitochondrial encephalopathy, lactic acidosis, and stroke-like episodes (MELAS), non-syndromic hearing loss | + | | (162,183-213) |
| tm3U         | MTO1| Lactic acidosis, cardiomyopathy, encephalopathy, non-syndromic hearing loss, cancer, myoclonus epilepsy associated with ragged-red fibers (MERRF) | + | | (162,164,214-217) |
| tm3s2U       | TRMU| Leigh syndrome, hepatopathy associated with hypercalcemia, non-syndromic hearing loss | + | | (165,218-222) |
| t6A          | OSGEP1| Cancer, MERRF | + | | (166,167) |
| t6A          | TRT1| Microcephaly, developmental delay, epilepsy, cancer | + | | (168,223-225) |
| ms2t6A       | CDK5RAP1| Cancer, type II diabetes, vitiligo | + | | (169,226-228) |
| m3A          | TRMT61B| Cancer, Alzheimer’s disease | + | | (173,174) |

* Disease likely caused by defects other than loss of tRNA modification.
man tRNAs used in decoding is very complex (see (42) for a recent review). Indeed, not all tRNA sequences encoded in the human genome are expressed in common cell-lines (43,44). Based on the loss of canonical secondary structure, mutations at highly conserved positions, or positioning in transcriptional silent chromosomal regions, some candidate tRNA genes are likely tRNA-derived Short Interspersed Nuclear Elements or pseudogenes, and others may have non-canonical functions outside of translation (5,45,46). Therefore, additional filtering criteria are needed to select a list of tRNAs that most likely decode mRNAs in the human cytosol. An updated set of ‘high confidence’ human tRNAs has been generated by tRNAscan-SE 2.0 (Chan, Lin and Lowe, unpublished data) and is available in the GtRNAdb (29). This list of over 400 tRNA genes contains 47 distinct isodecoders families (including tRNA^Sec, the tRNA for selenocysteine insertion).

A first set of biochemically-determined tRNA isocacceptor sequences that include chemical modifications in at least one mammal was extracted from the MODOMICS database at the time of the initial analysis (30 May 2017). Furthermore, partial information is available for modifications at specific positions such as wobble uridine (U34). It is known that 5-carbamoylmethyluridine (ncm5U) is found in Val-UAC-tRNA, 5-methoxycarbonylmethyl-2-thiouridine (mc5U)m and 5-methoxycarbonylhydroxymethyluridine (mchm5U) in Gly-UCC-tRNA (47). Since the modifications of U34 in human Arg-UCU-tRNA and Gly-UCU-tRNA differ from those in the corresponding yeast tRNA, it is difficult to predict the nature of most of the U34 modifications in humans (48). Finally, a large fraction of the RNA sequence data stems from the 60s, 70s and 80s, so based on paper and thin-layer chromatography (TLC) (49), photometric characterization of nucleosides following chromatography (50), and mass-spectrometry using low-resolution, low-sensitivity instruments (49). These methods failed to identify or distinguish some of the modifications, which are therefore listed as N and xN in Supplementary Table S1A but can now be detected with high-resolution mass spectrometry (51–54).

To add to the complexity of the task, the human genome (in contrast to yeast) encodes isocacceptor families that include many unique isodecoders, which are tRNAs with the same anticodon but contains variations in the tRNA body (42). Different isodecoders can be expressed under specific conditions as shown for the neuron-specific tRNA-Arg-UCU (55) or in the case of cancer (56). New high-throughput sequencing methods have been developed and optimized to facilitate detection of full length tRNAs such as DM-tRNA-seq (31) or tRNA-HydroSeq (57) or AlkAniline-Seq (58) and of tRNA-derived small RNAs (ARM-seq (32)). While these high-throughput RNA modification mapping methods reviewed in (59) and (30) are not yet as precise or quantitative as mass spectrometry, they do offer a practical, inexpensive method to survey a subset of modifications across all expressed tRNAs for many cell types. Using these methods also offers a first glimpse of the diversity of modification states across different isodecoders. Some isodecoder families in human such as Ala-AGC can be highly complex, contrasting the relatively simple view previously seen in budding yeast. In human, there are 22 high confidence Ala-AGC tRNA genes detected in the genome, which encode 16 unique (by sequence) Ala-AGC tRNA transcripts; in yeast, there are 11 Ala-AGC genes, which all encode identical Ala-AGC tRNA transcripts. This variation in human tRNA sequences also leads to an apparent complexity in tRNA modifications that is only now being appreciated. For example, RNA modification data collected with traditional methods exists for just 2 out of 16 Ala-AGC isodecoders (Ala-AGC-8 and Ala-AGC-11). ARM-seq and DM-tRNA-seq, however, both detect transcripts and modifications for many more isodecoders (Supplementary Table S5). These high-throughput methods allowed to detect four human modifications that were missing in MODOMICS at the time of our first analysis (see methods section). The final count of tRNA isocacceptors with modification information is 27 in humans and 38 in mammals (Figure 1 and Supplementary Tables S1A and S5).

### Linking the modifications of human cytosolic tRNAs to their corresponding modification enzymes

We generated a current list of chemical modifications found in human cytosolic tRNAs (Figure 2A, Supplementary Table S2) by combining the modification information from the tRNA sequences compiled in Supplementary Table S1A. Subsequently, we used this list as a starting point to generate the set of predicted human tRNA modification genes.

Once the list of human cytosolic tRNA modifications had been generated (Figure 2A, Supplementary Table S2), we linked the modifications to their corresponding modification enzymes whenever possible. This was done by using the advanced query tools of Uniprot for a first pass and then surveying the literature. By default, the reference linking the gene to the function is found by accessing the Uniprot entry for a given gene. Only when the reference had not yet been captured in Uniprot (~10 cases), did we add a PMID entry in Supplementary Table S2.

Not all predictions reach the same level of credibility. For example, in some cases, experimental validation is available for the human ortholog, while in other cases only the function of the yeast ortholog is validated. Therefore, we used the following code to classify the evidence of our functional annotation: [5] in vivo data in mammals; [4] in vivo data of the human or a related mammalian enzyme in a heterologous host; [3] in vitro data using the human enzyme; [2] similarity to an experimentally validated gene in a non-mammalian species; [1] candidates that have not been verified in any organism; [0] no clear candidate. These predictions are available in Supplementary Table S2, and we summarized all enzymes with evidence codes 2–5 in Figure 2A, using the protein names recommended by the HUGO Gene Nomenclature committee (35). According to this assessment, we predicted at least 76 proteins to be required for the modification of cytosolic tRNAs. Clearly, this is an underestimation, as more than 24 enzymes are still unknown (evidence code 0 or 1). Furthermore, for approximately 26 proteins there is no direct in vivo or in vitro experimental data using a mammalian homolog (evidence code 2). Thus, our analysis emphasizes that extensive experimental validations and research will be required to verify specific gene predic-
tRNA isoacceptors that have been biochemically characterized at the RNA level by traditional methods. Three additional tRNA isoacceptors (Rn-Val-NAC, Hs-Leu-NAA, Hs-Val-NAC) listed in Supplementary Table S1A weren’t placed in this figure due to their unknown nucleotide.

Identifying candidates for the ‘missing’ genes

To identify candidates of unidentified tRNA modifications enzymes, we compiled an initial list of ∼40 human proteins that are members of families known to be implicated in RNA modifications. These were mainly methyltransferases, pseudouridine synthases or THUMP-domain-containing proteins that have not yet been linked to a specific function. When we surveyed the recent literature, we found that ∼2/3 of these candidates had been reported to modify tRNA or proteins. For the remaining twelve proteins/gens, we gathered localization data from Mitocarta, and analyzed co-expression using the SEEK expression database profiles to identify the candidates that are coexpressed with RNA processing or translation related genes (Supplementary Table S4). This list is far from exhaustive as new methylase folds implicated in RNA modification are still being discovered (60).

Missing genes coding for cytosolic tRNA modification enzymes

In general, when the gene is missing, the function of the corresponding modification is very difficult to infer, as no genetic study can be conducted. In some cases, such as acp’tU20, the gene is not known in any organism, and almost no functional information is available. The only functional inference that can be done is if the gene encoding the enzyme responsible for the same modification is known in another organism. This is the case for a few modifications such as m1G6 and mcm5’s2U34, functional information is available. However, we feel it is a far stretch to transfer functional inferences made from prokaryotes or unicellular eukaryotes to human. Even if a related enzyme is known in another species, it is very difficult to predict how the unknown human enzyme discriminates substrate tRNAs from non-substrates. Thus, in the absence of information about the gene and enzyme responsible, very little information can be inferred about the function of the modification itself. Below we list modifications of human cytosolic tRNAs, for which the genes remain to be discovered and characterized, and this list also indicates the areas where functional information is missing.

- m1G6/7: This nucleotide is modified in multiple cytosolic and in at least one mitochondrial tRNA (Supplementary Table S1A and B). Trm14/TrmN are members of the COG0116 family of methyltransferases and target this position in several thermophilic bacteria and archaea (61). However other members of the same family, such as RmlL have been shown to methylate guanine residues in 23S RNA (62). THUMPD2 and THUMPD3 are two barely characterized members of this family in humans (Supplementary Table S4), and previous analyses suggested that these enzymes might be required for the formation of both cytosolic and mitochondrial modifications (23). THUMPD2 was found to form a complex with the m2,2G26-methylase TRMT1, while THUMPD3 was shown to interact with the methylase-activator protein TRM112 (63) in two high-throughput interactome studies (36), strengthening their role as tRNA methyltransferase candidates. However, experimental verification will be required to evaluate whether these two proteins are essential for the formation of m1G, whether they exhibit...
Figure 2. Cloverleaf representation of tRNA, with modified positions indicated for (A) cytoplasmic and (B) mitochondrial tRNAs, respectively, indicating genes/proteins experimentally validated in human, predicted with high confidence in other species, unknown with predictions, and unknown with no predictions.
different substrates specificities towards G6 or G7 and whether they act in mitochondria or in the cytoplasm.

- m¹A14: Enzymes responsible for this modification were identified in *S. cerevisiae* and belong to the pfam01746 family (64). The human genome encodes three members of this family: TRMT10A is required for the generation of m¹G9 in cytosolic tRNAs (65). TRMT10C as part of the RNase P complex, forms m¹G9 in mitochondrial tRNAs (62) and like some family members from other species, can also methylate adenosine to form m¹A9 (66). Hence, TRMT10B (Supplementary Table S4) is a candidate for the elusive m¹A14 methyltransferase, even if a recent report could not detect any tRNA methylation activity *in vitro* (67). As expected for a cytosolic enzyme, TRMT10B is not part of the predicted human mitoproteome (37). However, multiple reports of interactions with 25 mitochondrial ribosomal proteins (https://thebiogrid.org/127659) suggest this protein localizes to the mitochondria. Further experiments will be needed to determine whether TRMT10B is the missing m¹A14 methyltransferase or whether TRMT10A methylates both G9 and A14 or whether a yet unknown enzyme catalyzes this reaction.

- Acp³U20,20a: Only very few enzymes have been characterized that modify RNA by transferring the aminocarboxypropyl (acp) group of SAM, which is the methyl donor in most RNA-methylation reactions. However, acp-transporting enzymes belong to three unrelated superfamilies, which also contain methyltransferases. The only human enzyme currently known to introduce the acp³ modification is TSR3, a member of the COG2042 family (68), which is required for the biosynthesis of the hypermodified nucleotide m¹acp³Ψ in 18S rRNA (69). The crystal structures of its archaeal homologs revealed that TSR3 belongs to the SPOUT class of methyltransferases (69). The second structurally characterized acp-transferase Tyw2 belongs to the unrelated RFM superfamily (70). A different acp modification has been described in the diphthamide-biosynthesis pathway, where an acp group is transferred from SAM to the carbon atom of a histidine residue of eukaryotic translation elongation factor 2 (eEF2) by an enzyme that belongs to the Radical-SAM superfamily (71). acp³U is found in several positions in tRNA of different organisms like for example acp³U47 in *E. coli* tRNA, but the corresponding enzymes have not been identified in any of these species. Since all known acp transferases most likely arose independently from methyltransferases, the acp³-U-forming enzyme might currently be annotated as a hypothetical methyltransferase of unknown function (Supplementary Table S4) but it is difficult to select a plausible candidate in light of the diversity of known acp transferases.

- Ψ: The list of pseudouridine synthases modifying human tRNAs is far from complete. Several candidates have been proposed to be required for the modification of positions 30–32, 50, 72 or e11,12,13 (23), but several can likely be excluded as they were found to be required for the modification of mitochondrial rRNA and mito-tRNA at positions 27, 29, 39 and 50 (RPUSD4) (65,72) or mitochondrial mRNA (like RPUSD3) (73). RPUSD1 and RPUSD2 (Supplementary Table S4) have not been tested experimentally and are hence still valid candidates. Pus7/TruD, the enzyme that introduces Ψ13 is highly conserved in all three kingdoms (74) and is a member of the COG0585 family. The yeast Pus7 enzyme further modifies position 35 (75). PUS7 and PUS7L, two members of the COG0585 family in humans are products of a gene duplication that occurred most certainly in the common ancestors of metazoans (see http://phylomedb.org/?q=search_tree&seqid=Q9H0K6). Experiments will be required to determine whether these two enzymes have identical, overlapping or different substrates specificities. For example, one of the two enzymes might modify position 13, while the second enzyme might target position 35. Another possibility is that PUS7 and PUS7L target both positions 13 and 35, but in different tRNA isoacceptors. PUS7 is implicated in pseudouridylation of w8 in tRF derived from Ala-tRNA, Cys-tRNA and Val-tRNA but whether PUS7 acts directly on tRNA has formally not been shown (76). PUS1 is multisite specific so it is a plausible candidate for the positions 30 to 32, even though it has been found that the mouse homolog modifies positions 27, 28, 34 and 36 (77). Finally, based on experimental evidence from Archaea (78), it had been postulated that the human Pus10 is required for the formation of Ψ54 (28) and this was recently experimental validated in human (79).

- Q34: Queuosine in position 34 (Q34) is highly conserved in bacteria and eukarya. Humans like all eukaryotes are unable to synthesize Q but instead salvage the queuine (q) base from their diet and gut microflora as a micronutrient (80). Recent studies have shown that nutritionally determined Q-tRNA levels promote Dnm2-mediated methylation of tRNA-Asp and control translational speed of Q-decoded codons as well as at near-cognate codons (81). The heterodimeric human TGT enzyme formed by the QRT1 and QRT2 (previously called QTRTD1) subunits is the only fully characterized enzyme of the Q salvage pathway (80). A second human salvage-enzyme member of the DUF2419 family has been identified but its molecular function is unknown (82). Finally, the transporter for the q base or the precursor nucleoside Q is still elusive as well as the enzyme(s) that further modify the Q residue by attaching galactosyl or mannosyl moieties.

- mcמ³s¹U34: Wobble uridine is generally modified in all known organisms (see (83) for a recent review). The combination of modifications at positions 2 and 5 of the nucleobase results in an intricate tuning of codon-anticodon interactions, thus allowing the translation apparatus to distinguish codons in split-codon boxes and to introduce additional amino acids (83,84). 5-carboxymethyluridine (cm⁵U), the first step of the 5-modification is introduced by the action of the Elongator complex, a heteromeric complex consisting of two copies of Elp1–Elp6 that is activated by several auxiliary proteins (85). Orthologs of all yeast Elongator complex subunits are known and described in humans. However, human orthologs of the yeast regulatory components (the kinase Kti14, the phosphatase Sit4 and its regulatory subunits Sap185 and Sap190) could not be identified. Here, functional screens will be required to determine the counterparts of these components in humans. The conversion of cm⁵U to
mcm3U is catalyzed by the c-terminal Trm9 domain of ALKBH8 (86–88). mcm3U in some tRNA can be further hydroxylated to mcm3Um by the AlkB Domain of ALKBH8 (87) or 2'-O-methylated to cm3U by an unknown enzyme. The enzyme required for ncm3U formation from cm3U is not known in any organism and remains to be identified. 2-thiolation is achieved through the action of the URM1 pathway that shares features of bacterial sulfur-carrier proteins (SCP) and ubiquitin-like proteins (UBL) (89). The URM1 pathway components are straight forward to identify. Urm1 has two homologs in humans: URM1 and MOCS2A. However, MOCS2A is required for the synthesis of the molybdopterin cofactor while URM1 is required for tRNA thiolation and MOCS3 activates the SCP of both pathways (90). The final step of the thiolation reaction is performed by a complex consisting of CTU1 and CTU2.

- **Missing methyltransferases.** Methyltransferases are the biggest group of RNA modifying enzymes. While many tRNA methyltransferases have been discovered and characterized, a few of them remain to be identified (Supplementary Table S2). Members of the NSUN family (PF01189) usually introduce m3C modifications (91) and some such as NSUN2 are multi-site specific (92). However, NSUN2 is not required for the formation of m3C40 or m3C72 (92). NSUN7 is the only member of the NSUN family without a known substrate (Supplementary Table S4). Hence, it is a strong candidate for methylating one or both these positions. However, indirect data links it to methylation of enhancer RNAs (93). Three enzymes (METTL1A, METTL2B and METTL6) have been found to be involved in m3C32 modification potentially on different tRNA targets (94). Any of these three might be required for introducing m3C at position e2 and/or 47 as the biochemical assays have been inconclusive to date (See (94), Supplementary Table S5). It is unclear, which protein synthesizes m2U54 since two human homologs of yeast Trm2 were identified: TRMT2A and TRMT2B (Supplementary Table S2). It is not known whether these two proteins catalyze the same reaction or whether they differ in substrate specificity or sub-cellular localization. For example, TRMT2B is predicted to localize to mitochondria and might be required for modifying mitochondrial tRNAs (Supplementary Table S3). Finally, no candidate can easily be proposed for the formation of m1Ψ39, Ψm39 and Gm39. The pool of methyltransferase candidates among proteins with uncharacterized functions is large (~8, Supplementary Table S4), and we did not find evidence to favor a specific candidate.

### Identification of the genes encoding for mitochondrial tRNA modifications enzymes

The Suzuki laboratory published a thorough compilation of tRNA modification enzymes for the full set of 22 bovine mitochondrial tRNAs (23) and most of their functional annotations can be transferred to orthologous human enzymes (Figure 2B, Supplementary Table S3). Furthermore, some open cases have been solved since. Notably, ALKBH1 and NSUN3 are required for the formation of tC in initiator tRNA (95–97). The same ALKBH1 enzyme is further required for hm3C and tC formation in cytosolic tRNA (95).

A more complete compilation of the predicted human mitochondrial tRNA modification enzymes was published recently with extensive added functional information (24). We compiled these predictions and added evidence codes resulting in a list of 35 enzymes required to modify the full set of mitochondrial tRNAs (Figure 2B and Supplementary Table S3). An additional evidence code to classify enzymes that have been experimentally validated in the cytoplasm but not in mitochondria was added (evidence code 6). We will discuss here the remaining open questions.

The Q base is found in mitochondrial tRNAs and the catalytic subunit QTRT1 of the human transglycosylase complex is found in the mitoproteome (Supplementary Table S3). In the cytoplasm, QTRT1 forms a complex with QTRT2 (98) but it is not known whether this interaction also occurs in mitochondria. It has been shown that QTRT1 and QTRT2 are associated with the mitochondria with QTRT2 more loosely bound than QTRT1 (99). Is it possible that QTRT2 facilitates the transport of q, as the mitochondrial queuine transporter is missing?

Similar to cytosolic pseudouridine synthases, the set of enzymes introducing ψ residues in mitochondrial tRNAs is far from complete, in particular since different enzymes can introduce the same modification at a given position in different tRNAs, implying that many more might be missing. RPSU5D4 was recently shown to modify 16s rRNA from mitochondria and introduce ψ in tRNA(Phe) but not in mito-tRNA(Gly) (72). PUS3 was predicted to modify other mitochondrial tRNAs such as mito-tRNA(Glu) at position 39 (23). However, experimental data on PUS3 is available only for cytosolic tRNAs, requiring additional confirmation of its mitochondrial targets. Two pseudouridine synthases without known substrates (RPSU5D3 and PUSU1L) localize to the mitochondria (Supplementary Table S4). RUPSD3 modifies mitochondrial mRNAs (73), leaving PUSU1L as a strong candidate for an enzyme that modifies positions 30, 31, 50 and/or 57 (Supplementary Table S4). Nevertheless, we cannot exclude that a pseudouridine synthase not predicted to be mitochondrial such as RPSU5D1 or RPSU5D2 is actually dually-targeted as it has been recently shown for Pus10 that is translocated to the mitochondria only under specific physiological conditions (100).

In general, the situation is more complex when one gene encodes for two proteins that localize to different subcellular compartments, since the mitochondria-targeted isoform is often not identified as a mitochondrial protein. Thirty- seven proteins are predicted to be required for mitochondrial tRNA modifications with nine of these currently unknown, and eleven modify only mitochondrial tRNAs (Supplementary Table S3). In the Mitocarta analysis that integrates 14 different sources of predictions and experimental data to compile a list 1158 human mitochondrial protein (37), ten of these proteins were correctly identified as mitochondrial (Supplementary Table S3). The only exception is CDK5RAP1, an enzyme required for the thiolation reaction during ms2i6A formation (101,102). Of the dually targeted proteins, ten were correctly assigned as mitochondrial in Mitocarta (Supplementary Table S3) while seven others were not: TRM5, YRDC, PUS3, NSUN2,
TRM112, METTL2B and QTRT2. TRM112 and QTRT2 are non-catalytic subunits and it cannot be excluded that they are dispensable in mitochondria, as tRNA modification machineries can be simpler in mitochondria. For example, only two proteins are required for the synthesis of t6A in mitochondria while six proteins are required in the cytosol (103). The case of PUS3 has already been discussed above. For the remaining three cases, dual localization data in yeast for Sua5p (104) and for Trm5p (105) as well as predicted isoforms (see http://www.uniprot.org/uniprot/Q08J23 for NSUN2) suggest that they can be similarly found in both compartments in humans. Hence, the MitoCarta set of mitochondrial proteins may be incomplete.

Finally, the enzymes required for the formation of m\(^3\)G6, m\(^5\)C72 and of several \(\Psi\) residues are experimentally uncharacterized (Supplementary Table S3). As discussed above, the two candidates for the formation of cytosolic m\(^3\)G6 are THUMPD2 and THUMPD3. These two proteins are not predicted to localize to mitochondria (37) but this prediction might not be correct. The same is true for the candidate for the cytosolic m\(^5\)C72 methylase, NSUN7.

### Tissue expression and proteomics data of tRNA modification enzymes

Disease phenotypes of aberrant tRNA modification enzymes are often linked to neuronal phenotypes, metabolic disorders and cancer (17–21). However, the tissue-specific expression profiles of tRNA modification genes have not been systematically explored. Therefore, we used expression data available through the Genotype-Tissue-Expression (GTEx) project and compared expression levels of all modification genes in representative tissues of all organs (Figure 3). We also compiled tissue-specific proteomics evidence from the human proteome map (41) (Supplementary Figure S1). Overall, there are several general trends: First, expression levels of modification genes are quite uniform between different tissues. Second, expression levels of most genes are relatively low, in particular in whole blood. Third, generally high expression levels are observed in testis and transformed cells. Fourth, expression levels in brain are below average of the tissues with the exception of the cerebellum, where expression of a small number of genes reaches levels that are similar to expression observed in testis. This is surprising given the typical neuronal phenotypes observed upon defects in human modification enzymes. There are, however, clusters of genes that are upregulated in several brain tissues (Figure 3). These findings may point to a more crucial function during early steps of differentiation and proliferation of stem cells. Thus, it is likely that many of the observed defects in humans are either developmental phenotypes or relatively subtle. Finally, even though most genes of different pathways are found in similar clusters, this is not true for all pathways, e.g. the ELP pathway. A similar trend is seen when analyzing the proteomics data (Supplementary Figure S1). With the highly expressed genes, a correlation was observed between the transcriptomic and proteomics data (for example DUS2, PUS3 or TRMT5 in testis), however, the proteomic data is less complete due to the low expression levels of most enzymes (Supplementary Table S2).

![Figure 3. Expression of tRNA modification genes and candidates in a representative set of healthy human tissues. (A) Empirically selected tissues with physiological relevance are included. (B) Like (A), but only representing a set of brain tissues. Genes are clustered according to similarity in the expression profile. Source: gtxportal.org. Genes included are from Supplementary Table S2.](https://academic.oup.com/nar/article-abstract/47/5/2143/5304329)
CONCLUSIONS

This inventory of human tRNA modifications and the corresponding enzymes surprisingly reveals that despite the fact that the field of RNA modifications has dramatically expanded in recent years with 50 human modifications enzymes identified only in the last 10 years, the picture is far from complete. We estimate that between ~135 genes are required to modify cytosolic and mitochondrial tRNAs and that 23% of these genes still need to be identified and that another 22% require further experimental validations. Approximately 50% of the human modification genes have been linked to a number of human diseases (Table 1). Furthermore, all genes required for the formation of iron-sulfur clusters affect tRNA modification indirectly and are linked to diseases that are likely not mediated by tRNA modification defects. Like described before (17–22) the phenotypes are most often neurodegenerative or neurodevelopmental diseases like microcephaly and intellectual disabilities, but also renal and metabolic defects. Finally, roughly 50% of the disease genes have been linked to cancer (Table 1), suggesting that tRNA modification enzymes may provide interesting targets for cancer therapies. Also, given the wide diversity of tRNA transcript sequences in humans, the preference of different members of the modification enzyme families for different tRNA isodecoders remains an open question. An in-depth analysis of tRNA modification dynamics in various stress conditions and cell types will reveal the intimate relationship between tRNAs and their modifying partners in more detail. This compilation can act as a guide for future experiments to completely the characterization of the set of human tRNA modification enzymes.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Yuri Motorin, Jane Jackman, Marc Graille, Vincent P. Kelly, Ramesh Gupta, Peter C. Dedon and Michaela Frye for inputs and discussions, and Michaela Frye for inspiring us for Figure 1. The Genotype-Tissue Expression (GTEx) Project was supported by the Common Fund of the Office of the Director of the National Institutes of Health, and by NCI, NHGRI, NHLBI, NIDA, NIMH and NINDS.

FUNDING

National Institutes of Health [GM70641 to V. dC.-L.; HG006753 to T.M.L.]; Max Planck Society and the European Research Council [310489-tRNAmodi to S.A.L.]; German Research Foundation [SPP 1784 to S.A.L.]; IIMCB statutory funds (to P.B. and J.M.B.); Polish National Science Centre [2012/04/A/NZ2/00455 to J.M.B.]. The open access publication charge for this paper has been waived by Oxford University Press.

Conflict of interest statement. Janusz M. Bujnicki is an Executive Editor of Nucleic Acids Research.

REFERENCES

1. El Yacoubi, B., Bailly, M. and de Crécy-Lagard, V. (2012) Biosynthesis and function of posttranscriptional modifications of transfer RNAs. Annu. Rev. Genet., 46, 69–95.

2. Agris, P.F., Narendran, A., Sarachan, K., Väre, Y.Y.P. and Erusalim, E. (2017) The importance of being modified: The role of RNA modifications in translational fidelity. Enzymes, 41, 1–50.

3. Lorenz, C., Lünse, C.E. and Mörk, M. (2017) tRNA modifications: impact on structure and thermal adaptation. Biomolecules, 7, 35.

4. Chauvetreau, G.F. (2017) Impact of RNA modifications and RNA-modifying enzymes on eukaryotic ribonucleases. Enzymes, 41, 299–329.

5. Zhang, X., Cozen, A.E., Liu, Y., Chen, Q. and Lowe, T.M. (2016) Small RNA modifications: integral to function and disease. Trends Mol. Med., 22, 1025–1035.

6. Suzuki, T. and Numata, T. (2014) Convergent evolution of AUA decoding in bacteria and archaea. RNA Biol., 11, 1586–1596.

7. Gregson, J.M., Crain, P.F., Edmonds, C.G., Gupta, R., Hashizume, T., Phillipson, D.W. and McClosey, J.A. (1993) Structure of the archaenal transfer RNA nucleoside G*-15 (2-amino-4,7-dihydro-4-oxo-7-beta-D-ribofuranosyl-1H-pyrrol[2,3-d]pyrimidine-5-carboxi) amide (archaeosine). J. Biol. Chem., 268, 10076–10086.

8. Gallo, L.C., Penedo, F.J., Carnethon, M., Issasi, C.R., Sotres-Alvarez, D., Malcarne, V.L., Roesch, S.C., Youngblood, M.E., Davilus, M.L., Gonzalez, P. et al. (2014) The Hispanic Community Health Study/Study of Latinos Socio-cultural Ancillary Study: sample, design, and procedures. Ethn. Dis., 24, 77–83.

9. Boccialetto, P., Machnicka, M.A., Parta, E., Pitkowska, B., Baginski, B., Wreckci, T.K., de Crécy-Lagard, V., Ross, R., Limbach, P.A., Poter, A. et al. (2018) MODOMICS: a database of RNA modification pathways. 2017 update. Nucleic Acids Res., 46, D303–D307.

10. Chen, C., Huang, B., Anderson, J.T. and Byström, A.S. (2011) Unexpected accumulation of ncm(5)U and ncm(5)S(2)U (U) in a trm9 mutant suggests an additional step in the synthesis of ncm(5)U and ncm(5)S(2)U. PLoS One, 6, e20783.

11. Bujnicki, J.M. (1999) Comparison of protein structures reveals monophyletic origin of the AdoMet-dependent methyltransferase family and mechanistic convergence rather than recent differentiation of N4-cytosine and N6-adenine DNA methyltransferase. Silico Biol., 1, 175–182.

12. Larigüe, C., Lebaudy, A., Blanchard, A., Yacoubi, B., El, Rose, S., Grosjean, H. and Douthwaite, S. (2014) The flavoprotein McaP04076 (RlmF0) catalyzes m5U1939 modification in Mycoplasma Genus Silvaticum. Appl. Environ. Microb., 8073–8082.

13. Van, P. and Fujimori, D.G. (2011) RNA methylation by Radical SAM enzymes RlmN and Cfr proceeds via methylene transfer and hydride shift. Proc. Natl. Acad. Sci. U.S.A., 108, 3930–3934.

14. Husain, N., Tkaczuk, K.L., Tulsidas, S.R., Kaminska, K.H., Cubrilo, S., Maravic-Vlahovič, G., Bujnicki, J.M. and Sivaraman, J. (2010) Structural basis for the methylation of G1405 in 16s rRNA by aminoglycoside resistance methyltransferase Sgm from an antibiotic producer: a diversity of active sites in m7G methyltransferases. Nucleic Acids Res., 38, 4120–4132.

15. Kozbial, P.Z. and Mushegian, A.R. (2005) Natural history of S-adenosylmethionine-binding proteins. BMC Struct. Biol., 5, 19.

16. Goto-Ito, S., Ito, T. and Yokoyama, S. (2017) Trm5 and trm6: two enzymes from distinct origins catalyze the identical RNA modification, m1g37. Biol. Mol. Biotechnol., 7, 32.

17. Torres, A.G., Battle, E. and Ribas de Pouplana, L. (2014) Role of tRNA modifications in human diseases. Trends Mol. Med., 20, 306–314.
59. Limbach,P.A. and Paulines,M.J. (2017) Going global: the new era of mapping modifications in RNA. *Wiley Interdiscip. Rev. RNA*, 8, e1367.

60. Kimura,S., Miyauchi,K., Ikeuchi,Y., Thiaville,P.C., de Crécy-Lagard,V. and Suzuki,T. (2014) Discovery of the β-barrel-type RNA methyltransferase responsible for N6-methylation of N6-threonylcarbamoyladenosine in tRNAs. *Nucleic Acids Res.*, 42, 9350–9356.

61. Fislaier,R., Roovers,M., Tuszynska,I., Bujnicki,J.M., Droogmans,L. and Versées,W. (2012) Crystal structures of the tRNA:m2G methyltransferase Trm14/TrmN from two domains of life. *Nucleic Acids Res.*, 40, 5149–5161.

62. Seger,J.V., Blenow,A.A. and Dontsova,O.A. (2007) Ribosomal RNA guanine (N2)-methyltransferases and their targets. *Nucleic Acids Res.*, 35, 2295–2301.

63. Bourgeois,G., Létoquart,J., van Tran,N. and Graille,M. (2017) *EMBO Rep.* essential for mitochondrial protein synthesis and cell viability.

64. Jackman,J.E., Montange,R.K., Malik,H.S. and Phizicky,E.M. (2003) Identification of the yeast gene encoding the tRNA m1G methyltransferase responsible for modification at position 9. *RNA*, 9, 574–585.

65. Evans,D.G., Harkness,E., Laloo,F. and Howell,A. (2014) The diphthamide modification pathway from eukaryota. *Nucleic Acids Res.*, 42, 1213–1226.

66. Van Laer,B., Roovers,M., Wauters,L., Kasprzak,J.M., Dymza,M., Deyaert,E., Singh,R.K., Fella,A., Bujnicki,J.M., Droogmans,L. et al. (2016) Structural and functional insights into tRNA binding and adenosine N1-methylation by an archael Trm10 homologue. *Nucleic Acids Res.*, 44, 940–953.

67. Vilardo,E., Nachbagauer,C., Buzet,A., Taschner,A., Holzmann,J. and Rossmanith,W. (2018) A subcomplex of human mitochondrial RNase P is a bifunctional methyltransferase - extensive moonlighting in mitochondrial tRNA biogenesis. *Nucleic Acids Res.*, 46, 11126–11127.

68. Armengaud,J., Dedieu,A., Solques,O., Pellequer,J.L. and Quemeneur,E. (2005) Deciphering structure and topology of conserved COG2042 orphan proteins. *BMC Struct. Biol.*, 5, 3.

69. Meyer,B., Wurm,J.P., Sharma,S., Immer,C., Pogoryelov,D., Köttler,P., Lafontaine,D.L.J., Wöhnert,J. and Entian,K.D. (2016) Ribosome biogenesis factor Tsfr is the aminoarboxypropyl transferase responsible for 18S tRNA hypermodification in mammal and humans. *Nucleic Acids Res.*, 44, 4304–4316.

70. Umitsu,M., Nishimaru,H., Noma,A., Suzuki,T., Ishitani,R. and Nureki,O. (2009) Structural basis of AdoMet-dependent aminoarboxypropyl transfer reaction catalyzed by tRNA-wybutosine synthesizing enzyme, TYW2. *Proc. Natl. Acad. Sci. U.S.A.*, 106, 15616–15621.

71. Schaffrath,R., Abdel-Fattah,K.R., Malik,H.S. and Pfizicky,E.M. (2001) Identification of the yeast gene encoding the tRNA m1G methyltransferase. *J. Med. Genet.*, 38, 571–580.

72. Antonicka,H., Choquet,K., Lin,Z., Gingras,A., Kleinman,C.L. and Schaffrath,R., Abdel-Fattah,W., Klassen,R. and Stark,M.J.R. (2008) The human ortholog of archaeal Pus10 produces pseudouridine 54 in select tRNAs where its recognition sequence contains a modified residue. *RNA*, doi:10.1261/rna.068114.118.

73. Guzzi,N., Cie´sla,M., Ngoc,P.C.T., Lang,S., Arora,S., Dimitriou,M., Behm-Ansmant,I., Massenet,S., Immel,F., Patton,J.R., Motorin,Y. and Kwiatkowski,S. (2018) The queuosine microenvironment: charting a course from microbe to man. *Nutrients*, 7, 2897–2929.

74. Tuorto,F., Legrand,C., Cirzi,C., Federico,G., Liebers,R., Müller,M., Ehrenhofer-Murray,A.E., Dittmar,G., Gröne,H.-J. and Lyko,F. (2018) Queuosine-modified tRNAs confer nutritional control of yeast survival. *EMBO J.*, 37, e99777.

75. Zalot,R.P., Brocher-Armanet,C., Gaston,K.W., Foroughi,H. and Schaffrath,R., Hunt,J.F. and de Crécy-Lagard,V. (2014) Plant, animal, and fungal micronutrient queuosine is salvaged by members of the DUF2419 protein family. *ACS Chem. Biol.*, 9, 1812–1825.

76. Schaffrath,R. and Leidel,S.A. (2017) Wobble uridine modifications—a reason to live, a reason to die? *RNA Biol.*, 14, 1209–1222.

77. Grosse,J. and Westhof,E. (2016) An integrated, structure- and energy-based view of the genetic code. *Nucleic Acids Res.*, 44, 8200–8240.

78. Dauden,M.I., Jacmi,M., Müller,C.W. and Glatt,S. (2018) Structural asymmetry in the eukaryotic Elongator complex. *FEBS Lett.*, 592, 502–513.

79. Fu,Y., Drophy,J.A.N., Chan,C.T.Y., Atmore,K.A., Begley,U., Pauls,R.S., Dedon,P.C., Begley,T.J. and Samson,L.D. (2010) Human AlkB homolog ABH5 is a RNA methyltransferase required for wobble uridine modification and DNA damage survival. *Mol. Cell. Biol.*, 30, 2449–2459.

80. Fu,Y., Dai,Q., Zhang,W., Ren,J., Pan,T. and He,C. (2010) The AlkB domain of mammalian ABH5 catalyzes hydroxylation of 5-methoxycarbonyluridine at the wobble position of tRNA. *Angewem. Chem. Int. Ed. Engl.*, 49, 8885–8888.

81. Songe-Møller,L., van den Born,E., Linne,Stephen, F., Vågbo,C.B., Kristoffersen,T., Krokån,H.E., Kirpekar,F., Falnes,P.O. and Klanglund,A. (2010) Mammalian ALKBH8 possesses tRNA methyltransferase activity required for the biogenesis of multiple wobble uridine modifications implicated in translational decoding. *Mol. Cell. Biol.*, 30, 1814–1827.

82. Termathe,M. and Leidel,S.A. (2018) The Ubα4 domain interplay is mediated via a thioester that is critical for tRNA thiolation through Urn1 thio-ribosylation. *Nucleic Acids Res.*, 46, 5171–5181.

83. Chowdhury,M.M., Dosche,C., Löhmannsroben,H.-G. and Leinmüller,S. (2012) Dual role of the molybdenum cofactor biosynthesis protein Moco in RNA thiolation and molybdenum cofactor biosynthesis in humans. *J. Biol. Chem.*, 287, 17297–17307.

84. Bujnicki,J.M., Fedor,M., Ayres,C.L. and Redman,K.L. (2004) Sequence-structure-function studies of tRNA:m5C methyltransferase Misu is multisite-specific. *Mol. Cell. Biol.*, 34, 5920–5940.

85. Antonicka,H., Blessing,C., Hübner,B., Seikowski,J., Dennerlein,S., Rehling,P. et al. (2018) Pseudouridylate of tRNA-derived fragments steers translational control in stem cells. *Cell*, 173, 1204–1216.
131. Zinshuyen, B. and Gilbert, W.V. (2013) Loss of a conserved tRNA anticodon modification perturbs cellular signaling. PLoS Genet., 9, e1003675.

132. Shaheen, R., Patel, N., Shamsheldin, E., Alzahrani, F., Al-Yamany, R., AlMoisheer, A., Ewida, N., Anazi, S., Alnemer, M., Elsheikh, M. et al. (2016) Accelerating matchmaking of novel dysmorphology syndromes through clinical and genomic characterization of a large cohort. Genet. Med., 18, 688–695.

133. Shaheen, R., Al-Salam, Z., El-Hattab, A.W. and Alkuraya, F.S. (2016) The syndrome dysmorphic facies, renal agenesis, ambiguous genitalia, microcephaly, polydactyly and lissencephaly (DREAM-PL): Report of two additional patients. Am. J. Med. Genet. A., 170, 109–115.

134. Yamada, Y., Yasukouchi, Y., Kato, K., Oguri, M., Moribe, H., Fujimaki, T., Takeuchi, I. and Sakuma, J. (2018) Identification of 26 novel loci that confer susceptibility to early-onset coronary artery disease in a Japanese population. Biomed. Res., 9, 383–404.

135. Powell, C.A., Kopajtic, R.J., D’Souza, A.R., Rorbach, J.K., Lerner, S.S., Husain, R.A., Dallabona, C., Donmini, C., Alston, C.L., Griffin, H. et al. (2015) TRMT5 mutations cause a defect in post-transcriptional modification of mitochondrial tRNA associated with multiple Respiratory-Chain deficiencies. Am. J. Hum. Genet., 97, 319–328.

136. Zhou, M., Xue, L., Chen, Y., Li, H., He, Q., Wang, B., Meng, F., Wang, M. and Guan, M.-X. (2018) A hypertension-associated mitochondrial DNA mutation introduces an m1G37 modification into tRNAMet, altering its structure and function. J. Biol. Chem., 293, 1425–1438.

137. Rodriguez, V., Chen, Y., Elkahloun, A., Dutra, A., Pak, E. and Chandrasekharappa, S.C. (2007) Chromosome 8 BAC array amplification and overexpression of TRMT12 in breast cancer. Clin. Genet., 73, 1425–1438.

138. Ismail, N.A., Ragab, S., Abd El Dayem, S.M., Baky, A.N.A.E., Hamed, M., Ahmed Kamel, S. and Abd El Halim, D. (2018) Frameshift mutations in repeat sequences of ANK3, HACD4, TCP10L, TP53BP1, MFN1, LCMT2, RNMT, TRMT6, TRM6 and OLL1RAP1 identifies IL18RAP, PTPN2, TAGAP, and PUS10 as shared risk loci for Crohn’s disease and celiac disease. PLoS Genet., 7, e1002464.

139. Li, C., Wang, S., Xing, Z., Lin, A., Iqbal, M., Bi`eche, I., Legrand, N., et al. (2016) A meta-analysis of genome-wide association scans identifies L135RAF, PTNP2, TAGAP, and PUS10 as shared risk loci for Crohn’s disease and celiac disease. PLoS Genet., 7, e1002464.

140. Li, C., Wang, S., Xing, Z., Lin, A., Iqbal, M., Bi`eche, I., Legrand, N., et al. (2016) A meta-analysis of genome-wide association scans identifies L135RAF, PTNP2, TAGAP, and PUS10 as shared risk loci for Crohn’s disease and celiac disease. PLoS Genet., 7, e1002464.

141. Shaheen, R., Abdel-Salam, G.M.H., Guy, M.P., Alomar, R., Abdel-Hamid, M.S., Alfi, H.I., Ismail, S., Ewida, N., Rorbach, J.K., Pak, E. and Alkuraya, F.S. (2015) Mutation in WDR4 impairs tRNA m7G46 methylation and causes a distinct form of microphthalmic primordial dwarfism. Genome Biol., 16, 210.

142. Hicks, D.G., Janathanan, B.R., Vardarajan, R., Kulkarni, S.A., Khoury, T., Dim, D., Budd, G.T., Yoder, B.J., Bubbs, R. and Schreider, M.T. et al. (2018) The expression of TRMT2A, a novel cell cycle regulated protein, identifies a subset of breast cancer patients with HER2 over-expression that are at an increased risk of recurrence. BMC Cancer., 10, 108.

143. Festen, E.A.M., Goyette, P., Green, T., Boucher, G., Beauchamp, C., Trynka, G., Dubois, P.C., Lagace, C., Stokkers, P.C.F., Hommes, D.W. et al. (2011) A meta-analysis of genome-wide association scans identifies L135RAF, PTNP2, TAGAP, and PUS10 as shared risk loci for Crohn’s disease and celiac disease. PLoS Genet., 7, e1002464.

144. Li, C., Wang, S., Xing, Z., Lin, A., Iqbal, M., Bi`eche, I., Legrand, N., et al. (2016) A meta-analysis of genome-wide association scans identifies L135RAF, PTNP2, TAGAP, and PUS10 as shared risk loci for Crohn’s disease and celiac disease. PLoS Genet., 7, e1002464.

145. Macari, F., El-Houfi, Y., Boldina, G., Xu, H., Khoury-Hanna, S., Ollier, J., Yazdani, L., Zheng, G., Bi`eche, I., LeGrand, N., et al. (2016) TRM6/61 connects PCoA with translational control through tRNA(met) stabilization: impact on tumorigenesis. Oncogene., 35, 1785–1796.

146. Li, C., Wang, S., Xing, Z., Lin, A., Liang, K., Song, J., Hu, Q., Yao, J., Zhen, C., Park, P.K. et al. (2017) A ROR1-HER3-lncRNA signalling axis modulates the Hippo-YAP pathway to regulate bone metastasis. Nat. Cell Biol., 19, 106–119.

147. Metodiev, M.D., Thompson, K., Alston, C.L., Morris, A.A.M., He, L., Essouline, Z., Rö, M., Bahn-Buisson, N., Pyle, A., Griffin, H. et al. (2016) Recessive mutations in TRMT10C cause defects in mitochondrial RNA processing and multiple respiratory chain deficiencies. Am. J. Hum. Genet., 98, 993–1000.

148. Oerum, S., Roovers, M., Leichenring, M., Acquaviva-Bourdain, C., Beermann, F., Gemperle-Britschgi, C., Foulloux, A., Korwitz-Reichelt, A., Bailey, H.J., Droogmans, L. et al. (2017) Novel patient missense mutations in the HSD17B10 gene affect dehydrogenase and mitochondrial tRNA modification functions of the encoded protein. Biochim. Biophys. Acta Mol. Basis Dis., 1863, 3294–3302.

149. Bläuel, S., Al-Salam, Z., El-Hattab, A.W., Alkuraya, F.S. et al. (2016) Mutations in the tRNA methyltransferase 1 gene TRMT1 cause congenital microcephaly, isolated inferior vernal hypoplasia and cystic leukomalacia in addition to intellectual disability. Am. J. Med. Genet. A., 176, 2517–2521.

150. Job, B., Bernheim, A., Beau-Faller, M., Camilleri-Broelter, S., Girard, P., Hofman, P., Mazieres, J., Toujani, S., Lacroix, L., Laffaire, J., et al. (2010) Genomic aberrations in lung adenocarcinoma in never smokers. PLoS One., 5, e13145.

151. Cheng, J.X., Chen, L., Li, Y., Cloe, A., Yue, M., Wei, J., Watanabe, K.A., Shammao, J.M., Anastasi, J., Shen, Q.J. et al. (2018) RNA virus methylation and tRNA methyltransferases mediate
chromatin organization and 5-azacytidine response and resistance in leukemia. Nat. Commun., 9, 1163.

162. Kopajtich, R., Nicholls, T.J., Borbach, J., Metodiev, M.D., Freisenger, P., Mandel, H., Vanlander, A., Ghezzi, D., Carrozzo, R., Taylor, R.W. et al. (2014) Mutations in GTPBP3 cause a mitochondrial translation defect associated with hypertrophic cardiomyopathy, lactic acidosis, and encephalopathy. Am. J. Hum. Genet., 95, 708–720.

163. Bykhovskaya, Y., Mengesha, E., Wang, D., Yang, H., Estivill, X., Mendoza-Fandino, G.A., Nord, S., Lilyquist, J., Olswold, C., Hermand, D., Delaunay, S., Klevernic, I., Jiang, Z., Jacques, N. et al. (2018) Matchmaking facilitates the diagnosis of an intellectual disability. J. Med. Genet., 55, 811–817.

164. Zhang, X., Liu, J., Yan, S., Huang, K., Bai, Y. and Zheng, S. (2015) High expression of N-acetyltransferase 10: a novel independent prognostic marker of worse outcome in patients with hepatocellular carcinoma. Int. J. Clin. Exp. Pathol., 8, 14765–14771.

165. Ma, R., Chen, J., Jiang, S., Lin, S., Zhang, X. and Liang, X. (2016) Up regulation of NAT10 promotes metastasis of hepatocellular carcinoma cells through epithelial-to-mesenchymal transition. Am. J. Transl. Res., 8, 4215–4223.

166. Donner, I., Katainen, R., Kaasinen, E., Aavikko, M., Sipilä, L.J., Pukkala, E. and Aaltonen, L.A. (2018) Candidate susceptibility variants in angioimmunoblastic T-cell lymphoma. J. Med. Genet., 55, 705–715.

167. Salehi Chaleshtori, A.R., Miyake, N., Ahmadvand, M., Bashti, O., Al, Rabei, H., Abdelraouf, H., Hashem, M., Alazami, A.M., Patel, N. et al. (2016) ADAT3-related intellectual disability: Further delineation of the phenotype. Am. J. Med. Genet. A, 170A, 1142–1147.

168. Starka, R., Zalán, A., Jobareen-Masri, A., Zahalka, H. and Mahajnah, M. (2018) A new case confirming and expanding the phenotype spectrum of ADAT3-related intellectual disability syndrome. Eur. J. Med. Genet., 61, 2057–2075.

169. Anderson, S.L., Coli, R., Daly, I.W., Kichula, E., Rork, M.J., Volpi, S.A., Ekstein, J. and Rubin, B.Y. (2001) Familial dysautonomia is caused by mutations of the IKAP gene. Am. J. Hum. Genet., 68, 753–758.

170. Bento-Abreu, A., Jager, G., Swinnen, B., Rui, L., Hendrickx, S., Jones, A., Staats, K.A., Taes, I., Eykens, C., Nommensen, A. et al. (2018) Elp3 drives Wnt-dependent tumor initiation and regeneration in the intestine. J. Exp. Med., 215, 2057–2075.

171. Belalcazar, L.M., Papadonatous, G.D., McCaffery, J.M., Peter, I., Pajewski, N.M., Erat, B., Allred, N.D., Balasubramanyam, A., Bowden, D.W., Brautbar, A. et al. (2015) A common variant in the CLDN7/ELP5 locus predicts adiponectin change with lifestyle modifications. Nat. Genet. 47, 1234–1239.

172. Yew, T.W., McCr yeast, L., Colclough, K., Ellard, S. and Pearson, E.R. (2016) tRNA methyltransferase homologue gene TRMT10A mutation in young adult-onset diabetes with intellectual disability, microcephaly and epilepsy. Diabet. Med., 33, e21–e25.
intervention and improvedfitness in obese individuals with diabetes. Physiol. Genomics, 47, 215–224.
194. Close,P., Gillard,M., Ladang,A., Jiang,Z., Papuga,J., Hawkes,N., Nguyen,L., Chapelle,J.-P., Bouillene,F., Svejstrup,J. et al. (2012) DERP6 (ELP5) and C3ORF75 (ELP6) regulate tumorigenicity and migration of melanoma cells as subunits of Elongator. J. Biol. Chem., 287, 32535–32545.
195. Braun,D.A., Rao,J., Mollet,G., Schapiro,D., Daugeron,M.-C., Tan,W., Gribouval,O., Boyer,O., Revy,P., Jobst-Schwan,T. et al. (2017) Mutations in KEOPS-complex genes cause nephrotic syndrome with primary microcephaly. Nat. Genet., 49, 1529–1538.
196. Wang,P.Z.T., Prasad,C., Rodriguez Cuellar,C.I. and Filler,G. (2018) Recent topics: the mechanisms, molecular genetics, and treatment of mitochondrial diseases. J. Hum. Genet., 64, 113–125.
197. He,X.-Y., Isaacs,C. and Yang,S.-Y. (2018) Roles of mitochondrial WDR4 mutations.
198. Park,S., Liu,M. and Kang,S. (2018) Alcohol Intake Interacts with CARD15 (-308A>G) to increase the risk of ulcerative colitis and celiac disease. Am. J. Med. Genet. A, 2460–2465.
199. Ghezzi,D., Baruffini,E., Haack,T.B., Invernizzi,F., Melchionda,L., Ghezzi,D., Baruffini,E., Haack,T.B., Invernizzi,F., Melchionda,L., Umeda,N., Suzuki,T., Yukawa,M., Ohya,Y., Shindo,H., Tanaka,R., Yamane,Y., Yamamoto,T., Goto,T. et al. (2018) Recent advances in the diagnosis, molecular pathogenesis of human mitochondrial diseases. J. Biol. Chem., 280, 1613–1624.
200. Asano,K., Suzuki,T., Schara,U., von Kleist-Retzow,J.-C., Lainka,E., Gerner,P., Pyle,A., Holinski-Feder,E. et al. (2011) Mutations in WDR4 as a new cause of Galloway-Mowat syndrome with primary microcephaly. Hum. Genet., 130, 374–377.
201. Chen,S., Zheng,Z., Tang,J., Lin,X., Wang,X. and Lin,J. (2013) Functional characterization of the candidate tumor suppressor gene DDX6 in glioblastoma cell lines. PLoS One, 8, e50227.
202. Li,X. and Guan,M.-X. (2002) A human mitochondrial GTP binding protein related to tRNA modification may modulate phenotypic expression of the deafness-associated mitochondrial 12S rRNA modification. Mol. Cell. Biol., 22, 7701–7711.
203. Murayama,K., Shimura,M., Liu,Z., Okazaki,Y. and Ohtake,A. (1998) Recent topics: the mechanisms, molecular genetics, and treatment of mitochondrial diseases. J. Hum. Genet., 64, 113–125.
204. Asano,K., Suzuki,T., Schara,U., von Kleist-Retzow,J.-C., Lainka,E., Gerner,P., Pyle,A., Holinski-Feder,E. et al. (2011) Acute infantile liver failure due to mutations in the TRMU gene. J. Inherit. Metab. Dis., 34, 197–201.
205. Chen,W., Hosseini,M., Behjati,F., Haas,S., Jamali,P. et al. (2011) Deep sequencing reveals 50 novel genes for recessive cognitive disorders. Nature, 478, 57–63.
206. Meseguer,S., Martinez-Zamora,A., Garcia-Arumi,E., Andreu,A.L. and Armengol,M.-E. (2015) The ROS-sensitive microRNA-9/9* controls the expression of mitochondrial tRNA-modifying enzymes and is involved in the molecular mechanism of MELAS syndrome. Hum. Mol. Genet., 24, 1567–1584.
207. Ghezzi,D., Baruffini,E., Haack,T.B., Invernizzi,F., Melchionda,L., Ghezzi,D., Baruffini,E., Haack,T.B., Invernizzi,F., Melchionda,L., Umeda,N., Suzuki,T., Yukawa,M., Ohya,Y., Shindo,H., Tanaka,R., Yamane,Y., Yamamoto,T., Goto,T. et al. (2018) Recent advances in the diagnosis, molecular pathogenesis of human mitochondrial diseases. J. Biol. Chem., 280, 1613–1624.
208. Asano,K., Suzuki,T., Schara,U., von Kleist-Retzow,J.-C., Lainka,E., Gerner,P., Pyle,A., Holinski-Feder,E. et al. (2011) Acute infantile liver failure due to mutations in the TRMU gene. J. Inherit. Metab. Dis., 34, 197–201.
Cdkal1, a type 2 diabetes susceptibility gene, regulates mitochondrial function in adipose tissue. Mol. Metab., 6, 1212–1225.

227. Shin, M.-K., Uhm, Y.-K., Lee, J.-H., Kim, S.-K., Chung, J.-H. and Lee, M.-H. Association between CDK5RAP1 polymorphisms and susceptibility to vitiligo in the Korean population. Eur. J. Dermatol., 22, 495–499.

228. Steinthorsdottir, V., Thorleifsson, G., Reynisdottir, I., Benediktsson, R., Jonsdottir, T., Walters, G.B., Styrkarsdottir, U., Gretarsdottir, S., Emilsson, V., Ghosh, S. et al. (2007) A variant in CDKAL1 influences insulin response and risk of type 2 diabetes. Nat. Genet., 39, 770–775.

229. Okamoto, M., Hirata, S., Sato, S., Koga, S., Fujii, M., Qi, G., Ogawa, I., Takata, T., Shimamoto, F. and Tatsuka, M. (2012) Frequent increased gene copy number and high protein expression of tRNA (cytosine-5-) methyltransferase (NSUN2) in human cancers. DNA Cell Biol., 31, 660–671.

230. Martinez, F.J., Lee, J.H., Lee, J.E., Blanco, S., Nickerson, E., Gabriel, S., Frye, M., Al-Gazali, L. and Gleeson, J.G. (2012) Whole exome sequencing identifies a splicing mutation in NSUN2 as a cause of a Dubowitz-like syndrome. J. Med. Genet., 49, 380–385.

231. Komara, M., Al-Shamsi, A.M., Ben-Sulem, S., Ali, B.R. and Al-Gazali, L. (2015) A novel Single-Nucleotide deletion (c.1020delA) in NSUN2 causes intellectual disability in an emirati child. J. Mol. Neurosci., 57, 393–399.

232. Yi, J., Gao, R., Chen, Y., Yang, Z., Han, P., Zhang, H., Dou, Y., Liu, W., Wang, W., Du, G. et al. (2017) Overexpression of NSUN2 by DNA hypomethylation is associated with metastatic progression in human breast cancer. Oncotarget, 8, 20751–20765.