Evolution of Methyltransferase-Like (METTL) Proteins in Metazoa: A Complex Gene Family Involved in Epitranscriptomic Regulation and Other Epigenetic Processes

Juliet M. Wong and Jose M. Eirin-Lopez

Environmental Epigenetics Laboratory, Institute of Environment, Florida International University, Miami, FL, USA

*Corresponding author: E-mail: jeirinlo@fiu.edu.
Associate editor: Amanda Larracuente

Abstract

The methyltransferase-like (METTL) proteins constitute a family of seven-beta-strand methyltransferases with S-adenosyl methionine-binding domains that modify DNA, RNA, and proteins. Methylation by METTL proteins contributes to the epigenetic, and in the case of RNA modifications, epitranscriptomic regulation of a variety of biological processes. Despite their functional importance, most investigations of the substrates and functions of METTLs within metazoaos have been restricted to model vertebrate taxa. In the present work, we explore the evolutionary mechanisms driving the diversification and functional differentiation of 33 individual METTL proteins across Metazoa. Our results show that METTLs are nearly ubiquitous across the animal kingdom, with most having arisen early in metazoan evolution (i.e., occur in basal metazoan phyla). Individual METTL lineages each originated from single independent ancestors, constituting monophyletic clades, which suggests that each METTL was subject to strong selective constraints driving its structural and/or functional specialization. Interestingly, a similar process did not extend to the differentiation of nucleoside-modifying and protein-modifying METTLs (i.e., each METTL type did not form a unique monophyletic clade). The members of these two types of METTLs also exhibited differences in their rates of evolution. Overall, we provide evidence that the long-term evolution of METTL family members was driven by strong purifying selection, which in combination with adaptive selection episodes, led to the functional specialization of individual METTL lineages. This work contributes useful information regarding the evolution of a gene family that fulfills a variety of epigenetic functions, and can have profound influences on molecular processes and phenotypic traits.

Key words: methyltransferase, METTL, epigenetics, phylogenetics, selection, metazoan.

Introduction

Methyltransferase enzymes catalyze the transfer of methyl groups to DNA, RNA, proteins, and other biomolecules (Cheng and Blumenthal 1999). An increasing interest in these proteins has been driven by epigenetics, formally defined as, “the study of phenomena and mechanisms that cause chromosome-bound, heritable changes to gene expression that are not dependent on changes to the DNA sequence” (Deans and Maggert 2015). In particular, work examining DNA methylation, as well as the enzymes responsible for this modification, has been conducted for decades across a broad range of contexts and model systems (Holliday 2006). In addition, other studies have recognized the structural and functional importance of methyltransferases that modify specific residues in proteins such as histones (Couture and Trievel 2006; Ng et al. 2009). More recently, there has been a growing interest in the post-transcriptional modification of RNA molecules, a concept known as RNA epigenetics or epitranscriptomics (He 2010; Saletore et al. 2012). Although over 170 different RNA modifications have been recorded (Machnicka et al. 2013), relatively little is known about the enzymes responsible for these modifications.

Depending on their protein structure, most methyltransferases are categorized into three large superfamilies: seven-beta-strand methyltransferases, SET methyltransferases, and SPOUT methyltransferases (Petrossian and Clarke 2009). Among the seven-beta-strand methyltransferases, the methyltransferase-like (METTL) gene family (table 1) encodes proteins characterized by a conserved S-adenosyl methionine (SAM or AdoMet) binding domain that is formed by part of the seven-beta-strand structure (Martin and McMillan 2002; Petrossian and Clarke 2009). In this study, we examine the METTL family because these enzymes have been demonstrated to modify DNA/RNA nucleosides as well as protein residues (Ignatova et al. 2019), leading to changes in gene expression and phenotype that can have profound effects on an organism’s condition. For example, METTL3 and METTL14 are responsible for the formation of N6-methyladenosine (m6A) in RNA (Liu et al. 2014). In eukaryotes, m6A is a very common RNA modification, and has been shown to play
a role in mammalian temperature stress response by promoting translation initiation of heat shock response genes (Zhou et al. 2015). As another example, it has been demonstrated that METTL21D methylates a lysine residue of valosin-containing protein (VCP), and may be linked to diseases, including cancer, in humans (Thiele et al. 2011; Kernstock et al. 2012). However, the precise targets and functions of several METTL proteins remain unresolved. In addition to structure, methyltransferases can be characterized by the type of biomolecule that they modify. Here, METTLs are characterized into three types: those that modify DNA/RNA, those that modify proteins, and those for which the biomolecule type they modify is currently unknown.

We also chose to examine the METTL family because it has been studied considerably less than other families of methyltransferases, such as the DNA methyltransferase (DNMT)

| Gene Name | NCBI ID | Aliases | Preferred Name | Target |
|-----------|---------|---------|----------------|--------|
| DNA/RNA-modifying METTLs | | | | |
| METTL1 | 4234 | C12orf1, TRM8, TRMT8, YDL201w | tRNA (guanine-N(7)-)-methyltransferase | m7G (tRNA, mRNA, miRNA) |
| METTL2 (METTL2A/B) | 339175/55798 | METL, PSNIP1 | tRNA N(3)-methylcytidine methyltransferase-like protein 2 | m3C (tRNA) |
| METTL3 | 56339 | IME4, M6A, MT-A70, Spo8, hMETTL3 | N6-adenosine-methyltransferase catalytic subunit | m6A (mRNA) |
| METTL4 | 64863 | Hst661 | N(6)-adenosine-specific methyltransferase METTL4 | m6A (DNA)/m6Am (snRNA) |
| METTL5 | 29081 | HSPC133, MRT72 | RNA N6-adenosine-methyltransferase METTL5 | m6A (rRNA) |
| METTL6 | 131965 | None | tRNA N(3)-methylcytidine methyltransferase METTL6 | m3C (tRNA) |
| METTL8 | 79828 | TIP | mRNA N(3)-methylcytidine methyltransferase METTL8 | m3C (mRNA) |
| METTL14 | 57721 | None | N6-adenosine-methyltransferase noncatalytic subunit | m6A (mRNA) |
| METTL15 | 196074 | METT5D1 | 125 rRNA N4-methylcytidine (m4C) methyltransferase | m4C (rRNA) |
| METTL16 | 79066 | METT10D | RNA N6-adenosine-methyltransferase METTL16 | m6A (ncRNA, pre-mRNA, mRNA) |
| METTL19 | 152992 | TRMT44, C4orf23, TRM44 | Probable tRNA (uracil-O(2))- methyltransferase | 2'-O-methyluridine (tRNA) |
| METTL25B | 51093 | RRNAD1, C1orf66, CGI-41 | Ribosomal RNA adenine dimethylate domain containing 1 | m6A (rRNA) |

Protein-modifying METTLs

| Gene Name | NCBI ID | Aliases | Preferred Name | Target |
|-----------|---------|---------|----------------|--------|
| METTL9 | 51108 | CGI-81, DREV, DREV1, PAP1 | Methyltransferase-like protein 9 | Unknown |
| METTL10 | 399818 | EEF1AKMT2, C10orf138, Elm4 | EEF1A lysine methyltransferase 2 | Lysine (EEF1A) |
| METTL11A | 28989 | NTMT1, AD-003, C9orf52, HOMT1A, NRMT, NRMT1, NTM1A | N-terminal Xaa-Pro-Lys N-methyltransferase 1 | N-terminal |
| METTL11B | 149281 | C1orf184, HOMT1B, NTM1B | Alpha N-terminal protein methyltransferase 1B | Lysine (citrate synthase) |
| METTL12 | 75107 | CSKMT, CS-KMT, U99HG | Citrate synthase-lysine N-methyltransferase CSKMT, mitochondrial | Lysine (EEF1A) |
| METTL13 | 51603 | EEF1AKMT, 5630401D24Rik, CGI-01, DFNB26, DFNB26M, DFNM1, KIAA0859, feat | eEF1A lysine and N-terminal methyltransferase | Lysine (EEF1A) |
| METTL18 | 92342 | AsTP2, C1orf156, HPM1 | Histidine protein methyltransferase 1 homolog | Histidine |
| METTL20 | 254013 | ETFBKMT, C12orf72, ETFB-KMT | Electron transfer flavoprotein beta subunit lysine methyltransferase | Lysine (ETFβ) |
| METTL21A | 151194 | FAM119A, HCA557b, HSPA-KMT | Protein N-lysine methyltransferase METTL21A | Lysine (HSPA) |
| METTL21B | 25895 | EEF1AKMT3, FAM119B | EEF1A lysine methyltransferase 3 | Lysine (EEF1A) |
| METTL21C | 19654 | C13orf53 | Protein-lysine methyltransferase METTL21C | Lysine (VCP) |
| METTL21D | 79609 | VCPKMT, C14orf138, VCP-KMT | Protein-lysine methyltransferase METTL21D | Lysine (VCP) |
| METTL21E | 403183 | 4832428D23Rik, Gm991 | Protein-lysine methyltransferase METTL21E | Lysine (VCP) |
The structures, activities, and functions of DNMTs have been studied across all domains of life (Lyko 2018; Bhattacharyya et al. 2020). In contrast, the examination of METTL proteins has remained extremely limited in nonmodel organisms despite the potential epigenetic functions of METTLs and their resulting influence on biological processes and phenotypic characteristics. Within Metazoa, studies of METTL proteins are often restricted to humans, mice, and rats. Thus, there remains a knowledge gap regarding the function of METTLs within other metazoan taxa, particularly within nonvertebrate phyla. Exploring how METTL proteins have evolved throughout Metazoa is important for understanding: 1) the biological functions of different METTLs, 2) the contribution of METTL functional specialization to the process of diversification among metazoan groups, and 3) the relative importance of different types of METTLs and their gene regions that vary in evolutionary conservation. The present work uses extensive data mining to collect available METTL sequences across metazoan taxa. From these, a select number of sequences from nine representative taxa across eight different metazoan phyla are used for multiple phylogenetic and evolutionary analyses. Lastly, evidence of diversifying selection episodes is assessed using vertebrate METTL sequences. By investigating METTL proteins across Metazoa, this work provides insight into the evolution of this epigenetically relevant gene family.

**Results and Discussion**

**METTLs Are Widespread across Metazoa**

In this study, we examined 33 genes that encode proteins within the METTL family (table 1). These were broadly characterized into three types: 1) those that target and modify the nucleosides of DNA or RNA molecules, 2) those that target and modify the residues of proteins, and 3) those with unknown modification targets. Overall, we assessed 12 METTLs that have been demonstrated to modify DNA or RNA nucleosides, 15 METTLs that have been demonstrated to modify protein residues, and 6 METTLs of currently unknown function (table 1). A detailed search of these methyltransferase-like genes was conducted within the NCBI GenBank database (https://www.ncbi.nlm.nih.gov/genbank). Importantly, this search was restricted to NCBI and may be limited by sequence data availability and annotation. Thus, any failure to detect a specific METTL gene within a particular phylum does not necessarily indicate that the gene is lost in that phylum. Nonetheless, the search found that METTL genes are present in most of the major metazoan phyla examined here including Porifera, Cnidaria, Brachiopoda, Mollusca, Platyhelminthes, Nematoda, Priapulida, Arthropoda, Echinodermata, Hemichordata, and Chordata (fig. 1). Chordates possess the largest number of METTL genes (i.e., 33 out of 33 examined in this current study), however, the diversification of METTL genes does not appear to be restricted to chordates and other deuterostomes. Indeed, among protostomes, our search detected that Mollusca possesses 28 unique METTL genes and Arthropoda has 26 unique METTLs. Basally to triploblastic animals, diploblastic cnidarians have 25 METTLs and parazoan sponges, which lack specialized tissues, have 19 METTLs. The widespread presence of METTLs in ancient metazoan lineages such as Cnidaria and Porifera suggests that they arose quite early in the evolutionary history of Metazoa. Furthermore, it appears that the early functional diversification of these various METTL types was necessary to enable critical cellular functionality, and therefore, the evolution of these phyla (Carroll et al. 2013).

There were several sequences that were determined to be probable METTL sequences, but could not be definitively identified as specific METTL genes based on the data available (fig. 1, represented as striped-gray boxes). These sequences were tentatively recognized as METTLs via orthologous gene
searches and Basic Local Alignment Search Tool (BLAST) searches but did not have any predicted gene identity or function recorded in NCBI. Gene searches in NCBI only identified putative METTL orthologs within one species of annelid, the leech *Helobdella robusta* (NCBI: taxid46124). For example, searches for orthologs to human METTL1 identified one gene in *H. robusta* (HELRODRAFT_91834), however, it is described as a “hypothetical protein” and the reference sequence is defined as, “*Helobdella robusta* hypothetical protein partial mRNA” (XM_009013229.1). NCBI BLAST results of this sequence include hits to METTL1 sequences in other metazoan phyla, although with relatively low query coverage and percent identity. All search results of putative METTL orthologs within Annelida were similar “hypothetical” proteins in *H. robusta*. Upon including putative METTL protein sequences from *H. robusta* in a separate phylogenetic analysis, each sequence clustered with its prospective METTL orthologs from other metazoan taxa (supplementary fig. S5, Supplementary Material online). Thus, it is probable that METTLs occur in Annelida, however, based on the current data available in NCBI, we hesitate to state with certainty the precise identity of these *H. robusta* sequences.

Our search in NCBI was not able to locate any methyltransferase-like genes within the phylum Ctenophora. This is not to say that the methylation of DNA, RNA, or protein molecules does not occur in ctenophores. DNA methylation has been detected in the promoter and gene body regions of ctenophore genomes, and ctenophores have been shown to possess DNMT1 (Dabe et al. 2015). DNMT1, a highly conserved member of the DNMT family of methyltransferases, acts to methylate cytosines in DNA and has been demonstrated to maintain methylation patterns following DNA replication (Kangaspeska et al. 2008; Lyko 2018). Ctenophore genomes also contain a DNMT2 gene (Dabe et al. 2015). Although a member of the DNMT family, DNMT2 is a RNA methyltransferase also known as tRNA aspartic acid methyltransferase 1 (TRDMT1) (Goll et al. 2006; Lyko 2018). Thus, methylation in ctenophores may be driven by methyltransferases outside of the METTL family, such as by DNMTs. Alternatively, METTLs may be present within ctenophores, but they were undetected in our search due to a limited availability of genomic resources. For instance, Ensembl Metazoa (https://metazoaensembl.org/index.html) detected at least one putative METTL in the ctenophore *Mnemiopsis leidyi*, which was identified as a potential ortholog of METTL1 (ML09559a). Upon including this putative METTL1 sequence from *M. leidyi* in a separate phylogenetic analysis, it clustered with METTL1 sequences from other metazoan taxa, although with low support (supplementary fig. S5, Supplementary Material online). This indicates that METTLs likely occur in Ctenophora. Nevertheless, evidence of this remains scant due to limited sequence data for multiple ctenophore species across available databases. Indeed, neither this nor similar METTL sequences were detected in Ctenophora by our NCBI search used for this study (fig. 1). As sequence data continue to increase in quality and quantity, METTLs may later be identified within Ctenophora and we expect additional METTLs will be found within other phyla as well (e.g., Annelida).

METTL2, a methyltransferase that forms N3-methylcytidine (m3C) in tRNAs (Xu et al. 2017), is found in the most phyla examined here (fig. 1). Additionally, humans possess two paralogs of METTL2 that share 99% amino acid sequence identity, METTL2A and METTL2B (Ambasseri et al. 2016; Lentini et al. 2020). Interestingly, the majority of METTLs found present in at least nine metazoan phyla are known to modify RNA (e.g., METTL1, METTL2, METTL3, METTL5, METTL6, METTL14, and METTL15). The ubiquity of RNA methyltransferases underscores the importance of RNA modifications in metazoan evolution. It is probable that RNA molecules required more than only four canonical nucleosides, and that the evolution of chemically distinct modified nucleosides enabled their multiple, necessary functions (Grosjean 2009).

**Phylogeny of METTLs Part I: Overview**

Following our data search, we opted to limit our analyses to nine preselected metazoan taxa (the sponge *Amphimedon queenslandica*, the stony coral *Acropora millepora*, the sea hare *Aplysia californica*, the priapulid *Priapulus caudatus*, the shrimp *Litopenaeus vannamei*, the sea star *Acanthaster planci*, the acorn worm *Saccoglossus kowalevskii*, the lancelet *Branchiostoma belcheri*, and humans *Homo sapiens*). This allowed us to analyze a reasonable quantity of data while including taxa that represented a diverse set of metazoan phyla. Furthermore, each of these taxa is the subject of numerous ecology and evolution studies and was selected due to its genome availability on NCBI. Thus, the phylogenetic and subsequent evolutionary analyses, with the exception of the episodic selection analysis that is limited to vertebrate METTLs, only include taxa from the phyla Porifera, Cnidaria, Mollusca, Priapulida, Arthropoda, Echinodermata, Hemichordata, and Chordata. Representatives from the phyla Ctenophora, Brachiopoda, Annelida, Platychelminthes, and Nematoda are not included.

Phylogenetic reconstructions were performed based on an alignment of METTL protein sequences (alignment available in Supplementary Material online) from the nine preselected metazoan taxa. *Escherichia coli* prokaryotic rRNA dimethyltransferase (KsgA) was selected as the outgroup. KsgA, which was first identified in *E. coli* (Helser et al. 1972), dimethylates adenosines in the terminal helix (helix 45) near the 3’ end of the small subunit rRNA (Van Knippenberg et al. 1984). *Escherichia coli* KsgA is a member of the KsgA/Dim1 methyltransferase family, and like METTL proteins, it contains a seven-beta-strand sheet structure (Tu et al. 2009). KsgA, as well as the modifications it catalyzes, are nearly universally conserved throughout evolution (Xu et al. 2008), and have been described in archaea (O’Farrell et al. 2006), eubacteria (Helser et al. 1972; Van Buul et al. 1983), eukaryotes (Lafontaine et al. 1994; Housen et al. 1997), and in eukaryotic organelles (Tokuhiisa et al. 1998; Seidel-Rogol et al. 2003). Using both in vivo and in vitro analyses, one study showed that archaeal and eukaryotic orthologs of KsgA were capable...
of complimenting for KsgA function in bacteria, demonstrating that the recognition elements and methyltransferase activity of the KsgA/Dim1 family have evolved little since the three domains diverged (O’Farrell et al. 2006). Thus, E. coli KsgA was selected as a methyltransferase outside of, but that shares structural similarities with, the METTL family, and has orthologs present throughout all domains of life but has demonstrated little evolutionary change. The selection of E. coli KsgA as the outgroup for our metazoan METTL phylogeny is further supported by a separate phylogenetic analysis that included several METTL orthologs from eukaryotic taxa outside of Metazoa (supplementary fig. S5, Supplementary Material online). This included METTL protein sequences from the choanoflagellate Salpingoea rosetta and the yeast Saccharomyces cerevisiae, all of which grouped with their respective METTL lineages rather than by taxa or with E. coli KsgA.

The obtained tree showed that overall, the amino acid sequences cluster together by METTL lineage rather than by the taxa to which they belong (fig. 2; supplementary fig. S1, Supplementary Material online). Nearly all METTL lineages were each represented by a distinct monophyletic clade supported with high confidence values (e.g., all METTL1 orthologs from different metazoan taxa formed a monophyletic clade with 100% bootstrap). This observation suggests that the evolution of METTLs has been largely driven by selective constraints associated with the particular functional identity of each METTL protein, likely involved in specific outcomes of structural or functional specialization. Prior studies examining the structures, interactions, and functions of various METTLs have provided insight into their probable targets for methylation (e.g., Cloutier et al. 2013; Xu et al. 2017; Ignatova et al. 2019). Accordingly, METTL proteins tend to cluster together by their modification targets (i.e., DNA/RNA nucleosides versus protein residues), although with several exceptions.

Despite generally grouping by type (i.e., METTLs that modify nucleic acids versus METTLs that modify proteins), it seems clear that not all of the DNA/RNA-modifying METTLs (fig. 2, labeled in blue) exclusively possess a single, unique monophyletic origin and therefore do not appear to have evolved from a shared common ancestor. All of the protein-modifying METTLs (fig. 2, labeled in red) do not appear to share a single, unique monophyletic origin with one another either. Indeed, METTLs that target nucleosides for methylation are distributed across multiple clades within the phylogeny, and are occasionally grouped with METTLs that modify protein residues. This pattern could be explained by convergent molecular evolution that resulted in the emergence of METTLs with similar targets due to comparable selective pressures and modification requirements (Losos 2011). Alternatively, the diversification of the METTL family may have occurred very early in, or possibly prior to, the evolution of metazoans; this may have been driven by different selective pressures and constraints or even stochastic processes (Stayton 2015). The phylogenetic relationships among the various METTL types that target nucleic acids, proteins, or whose targets remain undetermined is further elaborated below.

Phylogeny of METTLs Part II: DNA and RNA Modifiers DNA- and RNA-modifying METTLs target nucleosides for modification, resulting in the formation of methylcytidines, methylguanosines, methyladenosines, or methyluridines. Almost all of these METTLs appear to primarily target RNA molecules, although there is evidence that METTL4 may act to methylate DNA (Kweon et al. 2019; Zhang et al. 2020). METTL15, METTL6, METTL2, and METTL8 are all associated with the formation of methylcytidine. METTL15 falls most basally to the E. coli KsgA outgroup (fig. 2) and forms N^5- methylcytidine (m5C) in RNA (Van Haute et al. 2019). The remaining proteins that methylate cytidine, METTL2, METTL6, and METTL8, form a highly supported cluster (100% bootstrap). Both METTL2 and METTL6 form m5C in specific tRNA molecules, whereas METTL8 forms m5C in mRNA (Xu et al. 2017). Furthermore, of the METTL genes, our NCBI search detected that METTL2 and METTL6 are present across the greatest number of major metazoan phyla, whereas METTL8 was only found to be present in the phylum Chordata (fig. 1). METTL1, the only METTL protein known to contribute methylguanosine, forms N^7-methylguanosine (m7G) in tRNA, mRNA, and miRNA (Okamoto et al. 2014; Pandolfini et al. 2019; Zhang et al. 2019). In the phylogeny, METTL1 groups with methylcytidine-forming proteins as well as with METTL7A and METTL7B, whose precise function is unknown, although with very low confidence (bootstrap <10%).

Several other nucleic acid-modifying METTLs target adenosine to form methyladenosines. These include METTL3, METTL4, METTL5, METTL14, METTL16, and METTL25B, which all function to form m6A. However, despite this shared function, these METTLs do not form a single cluster in the METTL phylogeny (fig. 2). Perhaps the best characterized of these are METTL3 and METTL14, which together form a heterocomplex with WTAP (Wilms’ Tumor1-Associating Protein) to form m6A in mRNA (Fu et al. 2014; Liu et al. 2014; Meyer and Jaffrey 2017). METTL3, METTL14, and WTAP are considered to be m6A writers (i.e., factors that encode the chemical modification) (Lewis et al. 2017). Although METTL3 is the primary m6A-forming enzyme, METTL14 does not demonstrate enzymatic activity and instead appears to bind substrate RNA and augment METTL3 activity (Sédé and Jinek 2016; Wang P et al. 2016; Wang X et al. 2016; Meyer and Jaffrey 2017). Unsurprisingly, METTL3 and METTL14 cluster together in the METTL phylogeny (86% bootstrap) (fig. 2).

The precise function(s) of METTL4, which clusters closest with METTL14 (95% bootstrap), is less clear than that of the METTL3–METTL14 heterocomplex. There is evidence that METTL4 modifies RNA, DNA, or both nucleic acids. Studies have demonstrated that METTL4 forms m6A in DNA (Kweon et al. 2019; Zhang et al. 2020), whereas others have shown that it forms N^2′,O-dimethyladenosine (m5Am) in snRNA (Chen et al. 2020; Goh et al. 2020). Interestingly, a recent study demonstrated that the human METTL3–METTL14 complex was active in vitro as a DNA N^2′-adine methyltransferase, although it is unclear if it functions in vivo (Woodcock et al. 2019).
Evidence suggests that METTL5, METTL16, and METTL25B also function in the formation of m⁶A. METTL5 and METTL16 cluster near to each other in the METTL phylogeny (92% bootstrap). METTL5 forms a heterodimeric complex with TRMT112 (tRNA methyltransferase subunit 11-2) to provide metabolic stability in the formation of m⁶A in 18S rRNA (van Tran et al. 2019; Leismann et al. 2020). The METTL5/C0TRMT112 complex has structural similarities to that of a m⁶A DNMT, and has been hypothesized to possess an RNA-binding mode unique to other m⁶A methyltransferases (van Tran et al. 2019). METTL16, which forms m⁶A in ncRNA, pre-mRNAs (Warda et al. 2017), and mRNA (Nance et al. 2020) appears to be structurally distinct from METTL5 (van Tran et al. 2019) as well as the METTL3–METTL14 heterocomplex (Ruszkowska et al. 2018). Furthermore, because METTL16 has been documented as both a nuclear protein and as a cytoplasmic methyltransferase, it has been hypothesized to possess an RNA-binding mode unique to other m⁶A methyltransferases (van Tran et al. 2019). METTL16, which forms m⁶A in ncRNA, pre-mRNAs (Warda et al. 2017), and mRNA (Nance et al. 2020) appears to be structurally distinct from METTL5 (van Tran et al. 2019) as well as the METTL3–METTL14 heterocomplex (Ruszkowska et al. 2018). Furthermore, because METTL16 has been documented as both a nuclear protein and as a cytoplasmic methyltransferase, it has been suggested that its RNA binding targets may differ by the location of METTL16 within the cell (Nance et al. 2020). The function of METTL25B is less clear, although Gene Ontology (GO) annotations related to METTL25B infers it possesses rRNA (adenine-N6, N6-)-dimethyltransferase activity. Lastly, the function of METTL19, which does not cluster with other RNA-modifying METTLs in the phylogeny (fig. 2), remains uncertain, but it has been reported to be a likely tRNA uracil-O(2)-methyltransferase that forms 2′-O-methyluridine (Leschziner et al. 2011). Overall, the majority of METTLs that modify nucleic acids appear to target primarily RNA, rather than DNA, nucleosides. Although some of these DNA/RNA-modifying METTLs are closely related phylogenetically, they do not share a single monophyletic origin and are instead interspersed with METTLs that modify protein residues.

Phylogeny of METTLs Part III: Protein Modifiers

Multiple METTL proteins are responsible for the methylation of eukaryotic elongation factor 1 alpha (EEF1A). In addition to its role in protein synthesis as a translation factor that delivers aminoacyl-tRNA to the ribosome, EEF1A has been ascribed to a wide variety of other functions in eukaryotes beyond protein synthesis (Mateyak and Kinzy 2010). METTL10, METTL21B, and METTL13 are responsible for the methylation of lysine residues within EEF1A (Li et al. 2014; Shimazu et al. 2014; Hamey et al. 2017; Malecki et al. 2017; Jakobsson et al. 2018). Interestingly, these METTLs do not cluster together within the phylogeny (fig. 2). For example, METTL10 clusters closest with
METTL19 (86% bootstrap), which is predicted to form Z'-O-methyluridine in RNA. On the other hand, METTL21B appears to share a close common ancestor with METTL21A (100% bootstrap), which also methylates lysine, although its target is within the molecular chaperone heat shock protein 70 (Hsp70) (Shimazu et al. 2014). METTL13 is unique in that it possesses two distinct methyltransferase domains and methylates the N-terminal as well as a lysine residue (Lys55) within EEF1A (Jakobsson et al. 2018; Liu et al. 2019).

Other METTL proteins methylate lysine residues of VCP, also known as p97, an ATP-driven chaperone involved in numerous, independent cellular processes that is highly conserved across eukaryotes (Kernstock et al. 2012; Meyer et al. 2012). METTL21C and METTL21E, which form a highly supported cluster (100% bootstrap), are both responsible for lysine methylation in VCP (Wiederstein et al. 2018; Wang et al. 2019). Although METTL21D, which is also known to methylate lysine residues of VCP (Kernstock et al. 2012), forms a cluster with other METTL21 lineages (98% bootstrap), it is not as closely related to the other VCP methyltransferases as they are to each other. Our search detected METTL21C and METTL21E only in chordates, whereas METTL21D is present in at least eight different metazoan phyla (fig. 1). Thus, METTL21D may have evolved prior to METTL21C and METTL21E, although their functions appear to be conserved.

METTL22, which methylates a lysine residue of Kin17 (Cloutier et al. 2013), appears to share a common monophyletic origin with a handful of other lysine methyltransferases, METTL21A-E (fig. 2). Kin17 functions in DNA replication and repair as well as RNA metabolism (Kannouche et al. 2000; Despras et al. 2003; Masson et al. 2003; Angulo et al. 2005). METTL22 may impact these processes because methylation of Kin17 affects its distribution between chromatin and the cytoplasm (Cloutier et al. 2014). METTL20 targets and trimethylates lysines of the electron transfer flavoprotein b-subunit (ETFb) (Rhein et al. 2014; Malecki et al. 2015). Unlike most lysine methyltransferases, METTL20 is localized to the mitochondria where it appears to modulate the function of ETF. This difference in target and function may partially explain why METTL20 fails to cluster with other lysine protein methyltransferases (fig. 2). METTL20 is yet another lysine methyltransferase, although it targets citrine synthase (CS) (Malecki et al. 2017; Rhein et al. 2017). CS is localized within the mitochondrial matrix of eukaryotic cells and catalyzes the first step in the Krebs cycle (Wiegand and Remington 1986).

Interestingly, METTL12 was not phylogenetically close to any other lysine methyltransferases (fig. 2).

Not all protein-modifying METTL targets are restricted to lysine residues. For instance, METTL18 appears to monomethylate a histidine residue of the ribosomal protein RPL3, and potentially other proteins, to form a 3-methylhistidine (Webb, Zurita-Lopez, et al. 2010; Cloutier et al. 2013). METTL18 is currently the only METTL for which there is evidence that its primary target is histidine, although it shares a common phylogenetic origin with several lysine methyltransferases (fig. 2). Other protein-modifiers such as METTL11A and METTL11B appear to target N-terminal residues. METTL11A and METTL11B are homologs and cluster closely together (100% bootstrap) (fig. 2), and whereas our search found that METTL11A is present in several metazoan phyla, METTL11B is only found in chordates (fig. 1). METTL11A recognizes the N-terminal X-Pro-Lys sequence motif (Webb, Lipson, et al. 2010) and likely acts upon multiple protein targets (Petkowski et al. 2012; Faughn et al. 2018). Furthermore, METTL11A is capable of mono-, di-, or trimethylating N-terminal residues (Tooley et al. 2010). Although it recognizes the same sequence motifs, METTL11B appears to be an N-terminal monomethylase that increases the activity of METTL11A in chordates (Petkowski et al. 2013; Faughn et al. 2018). METTL11A and METTL11B appear to be closely related to METTL9 (94% bootstrap) (fig. 2). Although the precise identities of METTL9 target molecules is currently unknown, METTL9 has been demonstrated to interact with multiple protein partners, including the membrane protein Calnexin precursor (CANX) (Ignatova et al. 2019).

Lastly, METTL23 is localized in both the nucleus and cytoplasm, and is predicted to contain a domain for either an adenosine—methionine or lysine methyltransferase (Reiff et al. 2014). Although its precise target has not been definitely identified, METTL23 has been shown to associate with a transcription factor subunit, GABPA (GA-binding protein transcription factor, alpha subunit), and a molecular chaperone, GroEL (Bernkopf et al. 2014; Reiff et al. 2014). Here, METTL23 appears to be phylogenetically related to several lysine methyltransferases that are associated with various molecular chaperones and other stress responsive proteins (e.g., Hsp70, VCP, and possibly Kin17), including METTL21A-E and METTL22 (fig. 2). Future studies investigating molecular responses to stress may benefit from examining these METTLs and their potential contributions to post-translational modifications of stress-activated proteins. Overall, although the target substrates of many of the protein-modifying METTLs have been identified, much less is known regarding the specific functional purpose behind each modification. It is also worth noting that although there has been a sizable effort to examine enzymes that methylate histones (Clarke 2013), the majority of protein-modifying METTLs discussed here appear to act upon the lysine residues of non-histone proteins. Regardless, all of the protein-modifying METTLs do not share a single, unique monophyletic origin.

Phylogeny of METTLs Part IV: Unknown Modifiers

The functions and biomolecule target types of several METTLs are currently unknown. This includes METTL7A, METTL7B, METTL7C, METTL7D, METTL26, and METTL27. The phylogenetic reconstruction may provide some insight into the potential function and target of these METTLs (fig. 2, labeled in purple), although this should be interpreted with caution. METTL7A has been reported to be an integral membrane protein, and METTL7B has been shown to interact closely with TMEM126A, a mitochondrial membrane protein (Zehmer et al. 2009; Ignatova et al. 2019). Furthermore, both appear to function in lipid droplet formation in the endoplasmic reticulum (Turro et al. 2006; Zehmer et al. 2009). Here we found that METTL7A and METTL7B are closely related to...
one another (100% bootstrap), and appear to be most phylogenetically related to RNA-modifying METTLs that create m\(^{3}\)C, although with relatively low confidence (70% bootstrap, fig. 2).

METTL17 has been shown to physically and functionally interact with estrogen receptors (ERs), ER\(\alpha\) and ER\(\beta\), and acts as a coactivator, modulating their transcriptional activities (Du et al. 2015). This suggests that METTL17 is a protein modifier that targets ER proteins. However, another study proposed that METTL17 may function to form methylcystidine in RNA (Shi et al. 2019). This study found that METTL17 interacted with 12S mitochondrial ribosomal RNA (mt\(\mathrm{rRNA}\) and small subunits of mitochondrial ribosome (MSSU). In particular, METTL17 appeared to regulate m\(^{3}\)C and m\(^{4}\)C modifications in 12S mt\(\mathrm{rRNA}\), and its presence or absence altered mitochondrial ribosome function (Shi et al. 2019). Thus, it is unclear if METTL17 modifies protein residues or RNA nucleosides. In the phylogeny, METTL17 was most closely related to METTL18, a protein histidine methyltransferase, though with low support (fig. 2).

Lastly, very little is known about METTL25, METTL26, and METTL27, although all three are predicted to have the characteristic seven-beta-strand structure METTL motif and possess S-adenosylmethionine-dependent methyltransferase (SAM) activity. METTL25 possibly functions in RNA methylation as it is phylogenetically very similar to METTL25B (100% bootstrap, fig. 2), which forms m\(^{8}\)A in RNA. METTL26 forms a lowly supported cluster with METTL11A, METTL11B, and METTL9, which may suggest its function as a protein-modifying METTL. Based on the phylogeny, there is no clear indication of the potential function or target of METTL27, which forms a lowly supported cluster with a protein-modifying METTL (METTL12), several RNA-modifying METTLs (METTL2, METTL6, and METTL8), and other METTLs of unknown function (METTL2A and METTL7B) (fig. 2). The phylogenetic tree may provide some indication of the potential substrates for the METTLs of currently unknown functions. However, given that the DNA/RNA-modifying METTLs and protein-modifying METTLs are not separated into distinct clades, we must once again stress that these findings should be interpreted with caution. Additional studies will be required to identify the specific purpose of METTL7A, METTL7B, METTL17, METTL25, METTL26, and METTL27, but the phylogeny does provide some insight into their evolution.

**Phylogeny of METTLs Part V: Conserved Domains**

All conserved domains (CDs) that were identified within each METTL fell within one of two CD superfamilies: the AdoMet MTases superfamily or the MT-A70 superfamily (table 2; supplementary tables S1 and S2, Supplementary Material online). Only METTL3, METTL14, and METTL4 contain CDs within the MT-A70 superfamily (CDD accession number c101947), whereas all other METTLs contain CDs within the AdoMet MTases superfamily (CDD accession number c17173). S-adenosyl-L-methionine (SAM or AdoMet) methyltransferases (MTases) use SAM/AdoMet as a substrate for methyl transfer (Schubert et al. 2003). MT-A70 refers specifically to the 70-kDa SAM-binding subunit first identified in human mRNA (N6-adenosine)-methyltransferase (i.e., human METTL3 (Bokar et al. 1997; Bujnicki et al. 2002). Thus, the CD Search successfully identified the conserved SAM-binding domains formed by the central seven-beta-strand motif that is structurally conserved across all METTLs.

For several METTLs, no specific CD was recognized (i.e., only the CD superfamily was identified). A second phylogenetic reconstruction was performed using only the CD-containing regions of the protein sequences (alignment available in Supplementary Material online). Interestingly, although some METTLs that share CDs cluster together (e.g., METTL3, METTL14, and METTL4), this is not true for all METTLs with shared CDs (fig. 3; supplementary fig. S2, Supplementary Material online). For instance, eight different METTLs share the CD Methtyltransfer_25 (table 2), but these METTLs form separate clusters in the phylogenetic reconstruction (fig. 3). Thus, even within relatively well-conserved areas of the amino acid sequence, there exists enough sequence-level diversity to support the specialization of multiple METTLs with different and highly specific targets and functions.

Overall, the full sequence phylogeny and the CD phylogeny were similar, which indicates that METTL evolution was largely determined by the selective constraints operating on these CDs. However, there were some notable differences between the full sequence and the CD phylogenies. For instance, in the phylogeny represented by full protein sequences (fig. 2), METTL17 was most closely related to METTL18, a protein histidine methyltransferase, whereas in the CD phylogeny (fig. 3), METTL17 forms a lowly conserved cluster with several RNA-modifying METTLs: METTL6, METTL2, METTL8, and METTL1 (62% bootstrap). As previously mentioned, there is conflicting evidence regarding whether METTL17 acts on protein residues or RNA nucleosides. Another difference between the phylogenetic reconstructions worth noting is the relationship of METTL27 to other METTLs, which remained unclear in the full sequence phylogeny (fig. 2). In the CD phylogeny, however, METTL27 is most closely related to METTL11A and METTL11B (78% bootstrap, fig. 3), which modify protein N-terminal residues. Thus, restricting the protein sequences to only the CD regions provided some additional insight into the potential targets of METTLs of unknown function.

**Molecular Evolution and Selection in METTL Proteins**

Despite shared similarities in protein structure (i.e., an SAM-binding domain and seven-beta-strand motif), METTL evolution throughout Metazoa appears to have been largely driven by functional constraints that led to the specialization of each METTL lineage. To further examine the evolutionary mechanisms underlying their differentiation, variation within each METTL lineage was evaluated using two separate alignments (nucleotide and protein) of all METTLs from each of the nine preselected metazoan taxa (table 3) (alignments available in Supplementary Material online). Although variation in METTLs with only one representative taxa/sequence in the selected data set was not assessed (i.e., METTL8, METTL11B,
The obtained results showed that METTL27, a METTL of unknown function, displayed the greatest variation (mean amino acid distance, $d_{AA} = 0.69 \pm 0.03$) (Table 3). METTL27 also displayed slightly higher levels of codon bias ($ENC = 48.47$). However, codon bias was relatively low overall, with average ENC values per METTL ranging from 45.13 (METTL21B) to 54.92 (METTL7B).

Table 2. Conserved Domains within Each METTL Gene Based on Results from CD-Search and Batch CD-Search (NCBI).

| Superfamily          | CD                    | Gene            |
|----------------------|-----------------------|-----------------|
| AdoMet MTases superfamily | COG2263               | METTL5          |
|                      | DREV                  | METTL9          |
|                      | DUF938                | METTL26         |
|                      | Methyltransf. 11      | METTL7A, METTL7B|
|                      | Methyltransf. 12      | METTL2, METTL6  |
|                      | Methyltransf. 16      | METTL21A, METTL21D |
|                      | Methyltransf. 25      | METTL2, METTL6, METTL7A, METTL8, METTL10, METTL12, METTL13, METTL27 |
|                      | Methyltransf. 31      | METTL10         |
|                      | Methyltransf. 32      | METTL25B        |
|                      | Methyltransf. 4       | METTL1          |
|                      | Methyltransf. PK      | METTL11A        |
|                      | Nnt1                  | METTL20         |
|                      | PRK00050              | METTL15         |
|                      | No specific CD (superfamily only) | METTL11B, METTL16, METTL17, METTL18, METTL19, METTL21B, METTL21C, METTL21E, METTL22, METTL23, METTL25 |
| MT-A70 superfamily   | MT-A70 (MT-A70)       | METTL3, METTL14 |
|                      | No specific CD (superfamily only) | METTL4 |

**FIG. 3.** Maximum likelihood phylogeny reconstructed based on the CDs of each METTL protein across Metazoa. Bootstrap values (1,000 replicates) are displayed at each node. Branch label colors are identical to figure 2.
in METTL12 (\(d_{AA} = 0.63 \pm 0.03\)), a lysine methyltransferase found in several different metazoan phyla (fig. 1) that was not closely related to other lysine methyltransferases in the phylogenetic reconstruction (fig. 2). The METTL that exhibited the least diversity was METTL14 (\(d_{AA} = 0.32 \pm 0.02\)), an m^2ATP writer that forms a heterocomplex with METTLC3 and WTAP. The next lowest levels of variation were in METTL1, METTL2, METTLC3, METTLC5, and METTLC6 (average \(d_{AA} = 0.38 \pm 0.02\)). Thus, the six METTLs that showed the least amount of variation are each associated with the methylation of RNA nucleosides, indicating that high sequence conservation may be necessary to preserve their function.

Many conserved multigene families are thought to be subject to concerted evolution, in which gene family members evolve as a unit (i.e., in concert) via gene conversion or unequal crossing-over (Eirin-López et al. 2012). In the case of concerted evolution, we might expect to observe roughly the same level of synonymous variation (\(d_d\)) and non-synonymous variation (\(d_{N}\)) because gene conversion and crossover processes should affect synonymous and non-synonymous sites equally (Nei and Rooney 2005). However, the nucleotide variation underlying the diversity observed among METTLs is primarily synonymous. For all METTLs, \(d_d\) was significantly greater than \(d_{N}\) and all Z-tests for selection (i.e., HA: \(d_N < d_d\)) were highly significant (\(P < 0.001\)) (table 3). Therefore, strong purifying selection appears to have operated on the different METTLs. This finding was additionally supported by the results from Fisher’s exact tests of selection (supplementary tables S3 and S4, Supplementary Material online). Indications of strong purifying selection have been reported in many other protein coding gene families, including highly conserved histone and ubiquitin gene families (Nei et al. 2000; Piontkivska et al. 2002; Rooney et al. 2002; Eirin-López et al. 2004; González-Romero et al. 2008, 2010, 2012). These studies argued that, rather than concerted evolution, these gene families are subject to the birth-and-death model of evolution in which gene family members evolve independently via gene duplication followed by their maintenance or loss (Nei and Hughes 1992). Indeed, the birth-and-death concept may be the primary mechanism directing the long-term evolution of most multigene families (Eirin-López et al. 2012). Although detecting birth-and-death was not a goal of the present work, future studies of this gene
family might investigate intra- and interspecific gene duplication and the presence of pseudogenes, which would further support that METTLs are subject to this model of evolution.

Rates of METTL Evolution
In order to further explore functional constraints across METTLs that vary in target and function, the present study conducted estimations of the rates of evolution in METTL proteins. Results show that several RNA-modifying METTLs exhibited a relatively slow rate of evolution (fig. 4). In particular, METTL1, the only METTL known to form methylguanosine, had the slowest rate of protein evolution at 3.91 \times 10^{-4} amino acid substitutions/site/million years (My) (supplementary table S5, Supplementary Material online). METTL14 and METTL3, which together form a heterocomplex to form m^6A in RNA, also had relatively slow rates of evolution. In fact, the six METTLs with the slowest evolution rates (i.e., METTL1, METTL14, METTL2, METTL6, METTL3, and METTL5) all function to methylate RNA molecules (fig. 4; supplementary table S5, Supplementary Material online). These same METTLs are also found in multiple metazoan phyla, including basal metazoans (fig. 1). Thus, the formation of m^7G, m^3C, and m^6A also found in multiple metazoan phyla, including basal metazoans (fig. 1). Thus, the formation of m^7G, m^3C, and m^6A modifications in RNA appears to be a fundamental epitranscriptomic mechanism that has been preserved throughout metazoan evolution. In contrast, METTL12, otherwise known as citrate synthase-lysine N-methyltransferase (CSKMT), exhibited the fastest rate of evolution with 7.95 \times 10^{-4} amino acid substitutions/site/My (fig. 4; supplementary table S5, Supplementary Material online). Methylation modifications have been identified in numerous diverse, structurally complex, and specialized proteins (Paik and Kim 1971; Clarke 2013). Therefore, a comparatively rapid evolution and diversification of protein-modifying methyltransferases may have been necessary to target these proteins and fulfill a wide variety of highly specific roles. Protein-modifying METTLs being subjected to different selective pressures than those that modify DNA/RNA nucleosides may explain their different rates of evolution. Also exhibiting relatively rapid rates of evolution were METTL7A and METTL17, although the precise functions of these proteins are unknown. All remaining METTLs that exhibited a relatively intermediate rate of evolution were a combination of RNA-modifying METTLs, protein-modifying METTLs, and METTLs of unknown function (fig. 4). These METTLs exhibited highly similar rates of evolution at an average of 6.64 \times 10^{-4} \pm 3.48 \times 10^{-5} amino acid substitutions/site/My (supplementary table S5, Supplementary Material online).

Episodic Selection within METTL Lineages
We wanted to explore what drives the variation in evolutionary rates that we calculated across the METTL family, and we searched for indications of adaptive selective episodes that may have occurred during METTL evolution. Since METTL lineages each have a monophyletic origin (fig. 2), additional evolutionary analyses using only vertebrate METTLs were implemented to further explore how selection drove the differentiation of METTL types (alignment available in Supplementary Material online). Vertebrates were selected because all members of the METTL family are present within this subphylum of metazoans. Only two representative orthologous sequences per METTL were included in the analysis due to computational requirements and constraints. Accordingly, we examined ortholog sequences of all METTLs from each of two classes of vertebrates: Mammalia and Amphibia. METTL7B was not included in this analysis as orthologous gene searches in NCBI failed to clearly identify METTL7B orthologs in amphibians. The obtained results revealed that METTL proteins diverged from a homogeneous evolution pattern as expected by the global molecular clock hypothesis (In L with clock = -10239.1, ln L without clock = -9715.854, P value < 0.001). This result was also supported when testing the molecular clock hypothesis using nucleotide sequences of the vertebrate METTLs (In L with clock = -109637.787, In L without clock = -108792.321, P value < 0.001). These findings, along with the estimated rates of evolution reported above, reaffirm that various METTL proteins appear to have evolved at different rates from one another.

Given the observed heterogeneous rates of evolution across the METTL family, the presence of diversifying (i.e., adaptive) selection episodes was examined across vertebrate METTLs, uncovering traces of diversifying selection (\(\omega > 1\)) on specific branches of the phylogeny (P value < 0.05) (fig. SA). The terminal branch leading to METTL6, which forms m^3C in tRNA (Xu et al. 2017), was significant (P value < 0.01) (fig. SA). The terminal branch of METTL9, a probable protein methyltransferase (Ignatova et al. 2019), also showed significant traces of diversifying selection (P value < 0.001) (fig. SA). Several subtrees that included closely related METTLs (i.e., METTL11A and METTL11B; METTL21A, METTL21B, METTL21C, METTL21D, and METTL21E; METTL25 and
METTL25B) were collapsed into a single branch for figure readability.

Several individual sites subject to diversifying selection in METTLs were also identified using the mixed effects model of evolution (MEME) model (Murrell et al. 2012) (fig. 5B and C), a method we have previously used to explore the evolution of high mobility group (HMG) proteins (González-Romero et al. 2015) and sex-determining proteins (Eirín-López and Sánchez 2015). As expected, the number and identities of these sites varied depending on whether analyses included all METTL types, or if analyses discriminated among METTLs based on their target molecule type (i.e., those that target DNA or RNA nucleosides versus those that target protein residues). Upon analyzing all METTL sequences available for vertebrates

Fig. 5. Episodes of diversifying selection shaping the evolution of METTL proteins in vertebrates. For figure readability, "METTL" has been removed from the names of each tip label (e.g., "METTL1" is simply displayed as "1"). (A) The strength of selection at significant branches is represented in red ($\omega > 5$), gray ($\omega = 1$), and blue ($\omega = 0$), with the proportion of sites within each class represented by the color width. Thicker branches have been classified as undergoing episodic diversifying selection at corrected $P$ value < 0.001 (thicker branches) and $P$ value < 0.01 (thinner branches). (B) The physical positions of adaptive selection episodes involved in the diversification of METTL genes. Numbers of synonymous (blue bars) and nonsynonymous (red bars) substitutions at codon positions that are subject to significant episodes of diversifying selection in vertebrates ($P$ value < 0.05). (C) Phylogenetic location of the mutations involved in such episodes. Branches in red account for higher numbers of nonsynonymous mutations, branches in blue indicate higher numbers of synonymous mutations, and branches in green represent cases with the same numbers of nonsynonymous and synonymous mutations. Light blue, light red, and light purple backgrounds behind each branch indicate METTLs that modify DNA/RNA, proteins, or unknown targets, respectively. Subtrees that include several closely related METTLs (i.e., METTL11A and METTL11B; METTL21A, METTL21B, METTL21C, METTL21D, and METTL21E; METTL25 and METTL25B) were collapsed into a single branch for figure readability.
regardless of type (alignment available in Supplementary Material online), 19 codon sites were identified, most of which were subjected to episodic positive selection (fig. 5B). Four sites of predominantly positive selection (codons 89, 114, 747, and 785) were common to the majority of METTLs, spanning a variety of different targets and functions (fig. 5C). Analysis of these codon positions within the context of the METTL phylogeny suggests that mutations at these sites led to the differentiation of multiple METTL lineages (fig. 5C).

Evidence of positive selection at codon 89 was indicated at several internal nodes of the phylogeny, as well as the terminal branches leading to METTL1, METTL3, and METTL14 (fig. 5C). Codon 89 was also identified as a site of positive selection when the MEME analysis was restricted to only METTLs that modify DNA or RNA nucleosides (fig. 6A). Analysis of codon 114 indicated instances of both positive and negative (purifying) selection (fig. 5C). Specifically, although the majority of METTL lineages were subject to positive selection at codon 114, a few terminal branches (e.g., METTL3 and METTL14) indicate the presence of more synonymous than nonsynonymous mutations. Codon 114 was additionally identified by both MEME analyses that separately analyzed DNA/RNA-modifying METTLs and protein-modifying METTLs (fig. 6). Thus, episodic selection at codon 114 may have been particularly important to the diversification of METTLs, both within and across different METTL types. Mutations at codon 747 were also common to many different METTL lineages within the phylogeny. Interestingly, this position is located within the CD regions of all METTLs that possess CDs belonging to the AdoMet MTases superfamily (supplemental table 2, Supplementary Material online). Codon 747 is not, however, located within the CDs belonging to the MT-A70 superfamily, which are found in METTL3, METTL14, and METTL4. Codon 785 had the most substitutions of the sites identified from the MEME analysis of all vertebrate METTLs (fig. 5B), and appears to have contributed to the differentiation of many of the METTL lineages. This site was also identified as a site of positive selection for the MEME analysis of only DNA/RNA-modifying METTLs, and therefore, appears to have played a significant role in the diversification of METTLs that bind to nucleosides. Although only a few prominent sites identified from the MEME analysis of all METTLs are discussed here,

![Fig. 6](https://academic.oup.com/molbev/advance-article/doi/10.1093/molbev/msab267/6364192)

**Fig. 6.** Episodes of adaptive selection identified upon separately analyzing vertebrate DNA/RNA-modifying METTLs and vertebrate protein-modifying METTLs. (A) The physical positions of adaptive selection episodes involved in the diversification of METTLs that modify DNA/RNA nucleosides (top) and METTLs that modify protein residues (bottom). Numbers of synonymous (blue bars) and nonsynonymous (red bars) substitutions at codon positions that are subject to significant episodes of diversifying selection in vertebrates (P value < 0.05). (B) Phylogenetic location of the mutations involved in such episodes for DNA/RNA-modifying METTLs, indicated by a light blue background, and (C) protein-modifying METTLs, indicated by a light red background. For figure readability, “METTL” has been removed from the names of each tip label (e.g., “METTL1” is simply displayed as “1”). Branches in red account for higher numbers of nonsynonymous mutations, branches in blue indicate higher numbers of synonymous mutations, and branches in green represent cases with the same numbers of nonsynonymous and synonymous mutations.
additional substitutions at over a dozen more positions were identified (fig. 5B and C).

Several additional sites subject to diversifying selection that were not recognized in the indiscriminate MEME analysis of all vertebrate METTLs were identified when performing MEME analyses that were separated by METTL type (fig. 6A). The analysis of only METTLs that modify DNA or RNA nucleosides identified additional sites of predominately positive selection at codons 33, 53, 365, 1295, 1520, 1793, and 1825 (fig. 6A). Sites at codons 89, 114, 414, and 785 were previously identified in the METTL analysis that included all vertebrate METTLs (fig. 5B). Upon examining these sites within the context of the phylogeny, a mutation at codon 89 shows evidence of positive selection on the internal node leading to METTL2, METTL6, METTL8, and METTL15 (fig. 6B). However, additional substitutions at codons 33, 53, 114, 785, 1520, and 1825 were necessary for the differentiation of these METTL lineages at a few sites (i.e., codon 114, 1793, and 1825) (fig. 6B).

Although sites at codons 114 and 541 had been identified in the METTL analysis that included all vertebrate METTLs (fig. 5B), by separately analyzing only METTLs that modify protein residues, MEME identified additional sites of predominately positive selection at codons 69, 72, 91, 243, 719, 768, and 997 (fig. 6A and C). Nearly all of the site mutations identified using MEME in protein-modifying METTLs were subjected to positive selection. However, there was evidence of purifying selection at codon 69 (METTL20), codon 719 (METTL23), and codon 768 (METTL10 and METTL22) (fig. 6C). There were also evidence of equal synonymous and non-synonymous substitutions (i.e., neutral evolution) for METTL22 at codons 72 and 997 when analyzed within the context of the phylogeny (fig. 6C).

**Conclusions**

METTLs are widespread throughout the animal kingdom, with some having emerged quite early in the evolution of Metazoa. Individual METTL lineages formed independent monophyletic clades, and whereas various METTLs grouped together roughly by methylation target type, those with similar targets did not appear to have evolved from the same, recent common ancestor (e.g., all METTLs that modify nucleosides did not share an exclusive, single monophyletic origin). Evidence indicates that the long-term evolution of the METTL family is primarily driven by strong purifying selection and exhibits heterogeneous rates of evolution across the different METTL lineages. Functional specialization of the various METTLs seems to have occurred via episodes of adaptive selection at specific evolutionary times and codon sites. Lastly, given their presence in basal metazoans, comparatively low sequence-level variation, and slow estimated rates of evolution, several METTLs that target and methylate RNA nucleosides seem to be more conserved than those that modify protein residues, possibly due to their biological necessity and particular functional constraints. In view of the ubiquity of METTLs throughout metazoans, it is clear that these proteins fulfill a diverse set of essential biological functions across an expansive range of animal taxa, and that the epigenetic mechanisms associated with this gene family contributed to metazoan evolution and complexity. We therefore stress the importance of investigating these methyltransferases in species beyond the traditional metazoan model systems. This present work provides important insight into the evolution of this gene family and constitutes a valuable resource for studying the epigenetic functions of METTL proteins, and their impacts on biological processes and phenotype.

**Materials and Methods**

**Molecular Data Mining**

Extensive data mining was performed using the GenBank database (https://www.ncbi.nlm.nih.gov/genbank) to collect METTL sequences (available as of June 2020). METTLs across metazoan taxa were located using orthologous gene searches as well as the BLAST in NCBI. Each METTL gene (NCBI IDs provided in table 1) was used to search for NCBI orthologs and similar genes. These are calculated using protein sequence similarity, local synteny information, and similarity of protein architectures (CDD domains defined by NCBI SPARCLE) by NCBI’s Eukaryotic Genome Annotation pipeline and Gene database. Each METTL gene listed in table 1 was searched in BLAST using the somewhat similar sequences (BlastN) algorithm and limiting the organism search set to Porifera (NCBI: txid6040), Cnidaria (NCBI: txid6073), Brachiopoda (NCBI: txid6578), Mollusca (NCBI: txid6447), Annelida (NCBI: txid6340), Platyhelminthes (NCBI: txid6157), Priapulida (NCBI: txid33467), Nematoda (NCBI: txid6231), Arthropoda (NCBI: txid66565), Hemichordata (NCBI: txid10219), and Echinodermata (NCBI: txid75858). For any phyla for which search results only located sequences that were similar to a METTL, but did not have any predicted gene identity or function recorded in NCBI, they were noted as probable METTL sequences for that phylum.

Ultimately, sequences were selected from nine representative taxa across eight different metazoan phyla for phylogenetic and evolution analyses. These taxa included the sponge (phylum Porifera) Amphimedon queenslandica (NCBI: txid400682), the stony coral (phylum Cnidaria) Acropora millepora (NCBI: txid45264), the sea hare (phylum Mollusca) Aplysia californica (NCBI: txid6500), the priapidul (phylum Priapulida) Priapulus caudatus (NCBI: txid33762), the shrimp (phylum Arthropoda) Litopenaeus vannamei (NCBI: txid6689), the sea star (phylum Echinodermata) Acanthaster planci (NCBI: txid133343), the acorn worm (phylum Hemichordata) Saccoglossus kowalevskii (NCBI: txid10224), the lancelet (phylum Chordata) Branchiostoma belcheri (NCBI: txid7741), and humans (phylum Chordata) Homo sapiens (NCBI: txid9606). A total of 207 nucleotide coding sequences (CDS) and their corresponding amino

---

**Table 1**: Representative Metazoan Taxa and Targeted METTL Gene IDs

| Taxon | NCBI ID |
|-------|---------|
| Acanthaster planci | txid133343 |
| Amphimedon queenslandica | txid400682 |
| Aplysia californica | txid6500 |
| Branchiostoma belcheri | txid7741 |
| Litopenaeus vannamei | txid6689 |
| Priapulus caudatus | txid33762 |
| Saccoglossus kowalevskii | txid10224 |
| Homo sapiens | txid9606 |
acid sequences were collected across 33 METTLs (supplementary tables S6 and S7, Supplementary Material online). Sequence alignments for both CDS nucleotide sequences and amino acid sequences were performed using MAFFT version 7.309 (Katoh 2002; Katoh and Standley 2013) in Geneious version 9.1.8 (https://www.geneious.com). All METTL CDS nucleotide sequences, using a translation alignment, and all METTL amino acid sequences were aligned. Both the CDS nucleotide and amino acid alignments were aligned using the E-INS-I strategy algorithm and the BLOSUM62 scoring matrix with a 1.53 gap open penalty. CDs were identified within each of 207 CDS nucleotide sequences and 207 protein sequences using the CD-Search and Batch CD-Search tools (Marchler-Bauer and Bryant 2004; Marchler-Bauer et al. 2011) in NCBI. All sequences were searched across the CD Database (CDD) (Lu et al. 2020) using an expect value (E-value) threshold of 0.01.

Phylogenetic Analyses

The ribosomal RNA small subunit methyltransferase A (KsgA) from *Escherichia coli* O157: H7 strain Sakai (NCBI: txid386585) was included as an outgroup for the phylogenetic analyses. Best-fit substitution models for nucleotide and protein sequence alignments were calculated using a maximum likelihood approach in MEGA version 10.1.8 (Tamura et al. 2013), and included all used sites (i.e., no data were excluded). Maximum likelihood phylogenetic trees were reconstructed using IQ-TREE version 1.6.12 (Nguyen et al. 2015) using the best-fit substitution models determined from MEGA. Accordingly, the LG model (Le and Gascuel 2008) corrected for discrete Gamma model (Yang 1994) with four rate categories (LG + G) was used for the reconstruction of METTL protein phylogenies across the nine preselected metazoan taxa. Similarly, a protein tree was also reconstructed for all METTL proteins across the nine preselected metazoan taxa but using only CD amino acid sequences identified using the Batch CD-Search tools (Marchler-Bauer and Bryant 2004; Marchler-Bauer et al. 2011) in NCBI. The reliability of the reconstructed topologies was contrasted by an ultra-fast bootstrap approximation (UFBoot) (Hoang et al. 2018) with 1,000 replicates. Additional phylogenetic reconstructions were performed using the CDS nucleotide alignment as well as for individual METTL lineage protein sequences (supplementary figs. S3 and S4, Supplementary Material online).

Molecular Evolution and Selection Analyses

Several molecular evolutionary analyses were performed in MEGA version 10.1.8 (Tamura et al. 2013). Sequences from each METTL gene were analyzed based on CDS nucleotide and amino acid alignments of all METTL sequences from each of the nine preselected metazoan taxa (alignments available in Supplementary Materials online). The nucleotide alignment includes 9,684 nucleotide sites (3,228 codons) and the protein alignment includes 3,124 amino acid sites. The transition/transversion ratio (R) was calculated for all METTLs except for those in which only one sequence across the nine preselected taxa was available (i.e., METTL8, METTL11B, METTL21B, METTL21C, and METTL21E). Due to their smaller
material is also based upon work supported by the National Science Foundation under Grant No. HRD-1547798. This NSF Grant was awarded to Florida International University as part of the Centers of Research Excellence in Science and Technology (CREST) Program. This is contribution number 266 from the Coastlines and Oceans Division, and contribution number 1036 from the Freshwater Resources Division of the Institute of Environment at Florida International University.

Data Availability

The data underlying this article are available within the National Center for Biotechnology Information (NCBI) GenBank database (https://www.ncbi.nlm.nih.gov/genbank), and all GenBank accession numbers are listed in supplementary tables S6, S7, and S9, Supplementary Material online. Alignments analyzed in the present work are available in the article’s Supplementary Material online.

References

Angulo JF, Mauffrey P, Pinon-Lataillade G, Miccolli L, Biard DS. 2005. Putative roles of kin17, a mammalian protein binding curved DNA, in transcription. In: Ohyama T, editor. DNA conformation and transcription. Molecular Biology Intelligence Unit. Boston (MA): Springer. p. 75–89.

Armbraser AG, Iben J, Wei F-Y, Rijal K, Tomizawa K, Hafner M, Marais RJ. 2016. Evolving specificity of tRNA 3-methyl-cytidine-32 (m3C32) modification: a subset of tRNAser requires N6-isopentenylation at A37. RNA 22(9):1400–1410.

Bernkopf M, Webersinke G, Tongsook C, Koyani CN, Rafiq MA, Ayaz M, Muller D, Enzinger C, Aslam M, Naem F, et al. 2014. Disruption of the methyltransferase-like 23 gene METTL23 causes mild autosomal recessive intellectual disability. Hum Mol Genet. 23(15):4015–4023.

Bhattacharyya M, De S, Chakrabarti S. 2020. Origin and evolution of DNA methyltransferases (DNMT) along the tree of life: a multi-genome survey. bioRxiv 2020.04.09.033167.

Bokar JA, Shambaugh ME, Polayes D, Matera AG, Rottman FM. 1997. Purification and cDNA cloning of the AdoMet-binding subunit of the human mRNA (N6-adenosine)-methyltransferase. RNA 3(11):1233–1247.

Bujnicki JM, Fedor M, Radlinska M, Blumenthal RM. 2002. Structure prediction and phylogenetic analysis of a functionally diverse family of proteins homologous to the MT-A70 subunit of the human mRNA(m6)A methyltransferase. J Mol Evol. 55(4):431–444.

Carroll SB, Grenier JK, Weatherbee SD. 2013. From DNA to diversity: molecular genetics and the evolution of animal design. Hoboken (NJ): John Wiley & Sons.

Cheng X, Blumenthal R. 1999. S-adenosylmethionine-dependent methyltransferases: structures and functions. Singapore: World Scientific.

Chen H, Gu L, Orellana EA, Wang Y, Guo J, Liu Q, Wang L, Shen Z, Wu H, Gregory RI, et al. 2020. METTL4 is an snRNA m6Am methyltransferase that regulates RNA splicing. Cell Res. 30(6):544–547.

Clarke SG. 2013. Protein methylation at the surface and buried deep: thinking outside the histone box. Trends Biochem Sci. 38(5):243–252.

Cloutier P, Lavallée-Adam M, Faubert D, Blanchette M, Coulombe B. 2013. A newly uncovered group of distantly related lysine methyltransferases preferentially interact with molecular chaperones to regulate their activity. PLoS Genet. 9(1):e1003210.

Cloutier P, Lavallée-Adam M, Faubert D, Blanchette M, Coulombe B. 2014. Methylation of the DNA/RNA-binding protein Kin17 by METTL22 affects its association with chromatin. J Proteomics 100(11):15–124.
Liu J, Yue Y, Han D, Wang X, Fu Y, Zhang L, Jia G, Yu M, Lu Z, Deng X, et al. 2014. A METTL3-METTL14 complex mediates mammalian nuclear RNA N6-adenosine methylation. Nat Chem Biol. 10(2):93–95.

Liu S, Hausmann S, Carlson SM, Fuentes ME, Francis JW, Pillai R, Lofgren SM, Hulea L, Tandock K, Lu J, et al. 2019. METTL13 methylation of eEF1A increases translational output to promote tumorigenesis. Cell 176(3):491–504.e21.

Li Z, Gonzalez PA, Sasvari Z, Kinzy TG, Nagy PD. 2014. Methylation of translation elongation factor 1A by the METTL10-like See1 methyltransferase facilitates tombusvirus replication in yeast and plants. Virolology 448:43–54.

Losos JB. 2011. Convergence, adaptation, and constraint. Evolution 65(7):1827–1840.

Lu S, Wang J, Chitsaz F, Derbyshire MK, Geer RC, Gonzales NR, Gwadz M, Losos JB. 2011. Convergence, adaptation, and constraint. Cell 145(6):863–878.

Marchler-Bauer A, Bryant SH. 2004. CD-Search: protein domain annotation.

Marchler-Bauer A, Lu S, Anderson JB, Chitsaz F, Derbyshire MK, et al. 2009. CDD: a conserved domain database for the functional annotation of genomes. Nucleic Acids Res. 37(Database issue):D225–D229.

Marchler-Bauer A, Lu S, Anderson JB, Chitsaz F, Derbyshire MK, et al. 2010. CDD: a conserved domain database for the functional annotation of proteins. Nucleic Acids Res. 40(Database issue):D225–D229.

Marchler-Bauer A, Bryant SH. 2004. CD-Search: protein domain annotations on the fly. Nucleic Acids Res. 32(Web Server issue):W327–W331.

Marchler-Bauer A, Lu S, Anderson JB, Chitsaz F, Derbyshire MK, DeWeese-Scott C, Fong JH, Geer LY, Geer RC, et al. 2015. CDD: a conserved domain database for the functional annotation of proteins. Nucleic Acids Res. 43(Database issue):D225–D229.

Martin JL, McMillan FM. 2002. SAM (dependent) I AM: the S-adenosyl-L-methionine-dependent methyltransferase fold. Curr Opin Struct Biol. 12(6):783–793.

Masson C, Menaa F, Pinon-Lataillade G, Frobert Y, Chevillard S, Radicella JP, Sarasin A, Angulo JF. 2003. Global genome repair is required to activate KIN17, a UV-sensitive gene involved in DNA replication. Proc Natl Acad Sci U S A. 100(26):16162–6167.

Mateyak MK, Kinzy TG. 2010. eEF1A: thinking outside the ribosome. J Biol Chem. 285(28):21209–21213.

Meyer H, Bug M, Bremer S. 2012. Emerging functions of the VCP/p97 AAA-ATPase in the ubiquitin system. Nat Cell Biol. 14(5):117–123.

Meyer KD, Jaffe SR. 2017. Rethinking m6A readers, writers, and erasers. Annu Rev Cell Dev Biol. 33:319–342.

Murrell B, Werthoin JC, Moola S, Weighill T, Scheffler K, Kosakovski Pond SL. 2012. Detecting individual sites subject to episodic diverging selection. PLoS Genet. 8(7):e1002764.

Nance DJ, Satterwhite ER, Bhashar B, Misra S, Carraway KR, Mansfield KD. 2020. Characterization of METTL16 as a cytoplasmic RNA binding protein. PLoS One 15(11):e027647.

Nei M, Gojobori T. 1986. Simple methods for estimating the numbers of synonymous and non-synonymous nucleotide substitutions. Mol Biol Evol. 3(5):418–426.

Nei M, Hughes AL. 1992. Balanced polymorphism and evolution by the birth-and-death process in the MHC loci. In: Tsuji K, Aiizawa M, Sasazuki T, editors. 11th Histocompatibility Workshop and Conference. Oxford: Oxford University Press. Available from: https://cniiajc.np/jad/10015427478/.

Nei M, Kumar S. 2000. Molecular evolution and phylogenetics. Oxford: Oxford University Press.

Nei M, Rogozin IB, Piontkivska H. 2000. Purifying selection and birth-and-death evolution in the ubiquitin gene family. Proc Natl Acad Sci U S A. 97(20):10866–10871.

Nei M, Rooney AP. 2005. Concerted and birth-and-death evolution of multigene families. Annu Rev Genet. 39:121–152.

Ng SS, Yue WW, Oppermann U, Klose RJ. 2009. Dynamic protein methylation in chromatin biology. Cell Mol Life Sci. 66(3):407–422.

Pond SL, Frost SDW. 2004. A genetic algorithm approach to detecting lineage-specific variation in selection pressure. Mol Biol Evol. 22(3):478–485.

Pond SLK, Frost SDW, Muse SV. 2005. HyPhy: hypothesis testing using phylogenies. Bioinformatics. 21(5):676–679.

Poon AFY, Frost SDW, Kosakovsky Pond SL. 2009. Detecting signatures of selection from DNA Sequences Using Datamonkey. Methods Mol Biol. 537:163–183.

Reiff RE, Ali BR, Baron B, Yu TW, Ben-Salem S, Coulter ME, Schubert CR, Hill RS, Akawi NA, Al-Younes B, et al. 2014. METTL23, a transcriptional partner of GABPA, is essential for human cognition. Hum Mol Genet. 23(13):3456–3466.

Rhein VF, Carroll J, Ding S, Fearnley IM, Walker JE. 2017. Human METTL12 is a mitochondrial methyltransferase that modifies citrate synthase through inactivating and dynamic methylation of Lys-165 in human eukaryotic elongation factor 1 alpha (eEF1A). Nucleic Acids Res. 45(8):4370–4389.

Rhein VF, Prulcherla N, Desai PM, Rife JP. 2006. Recognition of a complex substrate by the KsgA/Dim1 family of enzymes has been conserved throughout evolution. RNA 12(5):725–733.

Rhein VF, Prulcherla N, Desai PM, Rife JP. 2006. Recognition of a complex substrate by the KsgA/Dim1 family of enzymes has been conserved throughout evolution. RNA 12(5):725–733.

Rhein VF, Prulcherla N, Desai PM, Rife JP. 2006. Recognition of a complex substrate by the KsgA/Dim1 family of enzymes has been conserved throughout evolution. RNA 12(5):725–733.

Rhein VF, Prulcherla N, Desai PM, Rife JP. 2006. Recognition of a complex substrate by the KsgA/Dim1 family of enzymes has been conserved throughout evolution. RNA 12(5):725–733.

Rhein VF, Prulcherla N, Desai PM, Rife JP. 2006. Recognition of a complex substrate by the KsgA/Dim1 family of enzymes has been conserved throughout evolution. RNA 12(5):725–733.

Rhein VF, Prulcherla N, Desai PM, Rife JP. 2006. Recognition of a complex substrate by the KsgA/Dim1 family of enzymes has been conserved throughout evolution. RNA 12(5):725–733.
Schubert HL, Blumenthal RM, Cheng X. 2003. Many paths to methyltransfer: a chronicle of convergence. Trends Biochem Sci. 28(6):329–335.

Seidel-Rogol BL, McCulloch V, Shadel GS. 2003. Human mitochondrial transcription factor B1 methylates ribosomal RNA at a conserved stem-loop. Nat Genet. 33(1):23–24.

Shimazu T, Tarjau J, Sohtome Y, Sodeoka M, Shinkai Y. 2014. Selenium-based S-adenosylmethionine analog reveals the mammalian seven-beta-strand methyltransferase METTL10 to be an EF1A1 lysine methyltransferase. PLoS One 9(8):e105394.

Shi Z, Xu S, Xing S, Yao K, Zhang L, Xue L, Zhou P, Wang M, Yan G, Yang P, et al. 2019. Mettl17, a regulator of mitochondrial ribosomal RNA modifications, is required for the translation of mitochondrial coding genes. FASEB J. 33(11):13040–13050.

Šledí P, Janík M. 2016. Structural insights into the molecular mechanism of the mbA writer complex. Elife 5:e18434.

Stafton CT. 2015. What does convergent evolution mean? The interpretation of convergence and its implications in the search for limits to evolution. Interface Focus 5(6):20150039.

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol. 30(12):2725–2729.

Taverne S. 2016. Some probabilistic and statistical problems in the analysis of DNA sequences. Lectures Math Life Sci 17:57–86.

Thiele W, Novac N, Mink S, Schreiber C, Plaumann D, Fritzmann J, Schindler N, Rothley M, Schwager C, Regiert T, et al. 2011. Discovery of a novel tumour metastasis-promoting gene, NVM-1. J Pathol. 225(1):96–105.

Tokuhisa JG, Vijayan P, Feldmann KA, Browse JA. 1998. Chloroplast subunits. Mol Gen Genet. 259(1):475–478.

Tooleyn CES, Petkowski JJ, Muratore-Schroeder TL, Balbaugh JL, Shabanowitz J, Sabat M, Minor W, Hunt DF, Macara IG. 2010. NRMT is an alpha-N-methyltransferase that methylates RCC1 and retinoblastoma protein. Nature 466(7310):1125–1128.

van Tran N, Ernst FGM, Van Knippenberg PH, Van Kimmenade JM, Heus HA. 1984. Phylogeny of the conserved 3’ terminal structure of the RNA of small ribosomal subunits. Nucleic Acids Res. 12(6):2595–2604.

Wang C, Zhang B, Ratliff AC, Arrington J, Chen J, Xiong Y, Yue F, Nie Y, Hu K, Jin W, et al. 2019. Methyltransferase-like 2e inhibits 26S proteasome activity to facilitate hypertrophy of type IIb myofibers. FASEB J. 33(9):9672–9684.

Wang P, Doxtader KA, Nam Y. 2016. Structural basis for cooperative function of Mettl3 and Mettl14 methyltransferases. Mol Cell 63(2):306–317.

Wang X, Feng J, Xue Y, Guan Z, Zhang D, Liu Z, Gong Z, Wang Q, Huang J, Tang C, et al. 2016. Structural basis of N(6)-adenosine methylation by the Mettl3-Mettt14 complex. Nature 534(7608):575–578.

Warda AS, Kretschmer J, Hackert P, Lenz C, Urlaub H, Höbartner C, Sloan KE, Bohnassack MT. 2017. Human METTL16 is a -methyladenosine (mA) methyltransferase that targets pre-mRNAs and various non-coding RNAs. EMBO Rep. 18(11):2004–2014.

Webb KJ, Lipson RS, Al-Hadid Q, Whitelegg JP, Clarke SG. 2010. Identification of protein N-terminal methyltransferases in yeast and humans. Biochemistry. 49(25):5225–5235.

Webb KJ, Zurita-Lopez CI, Al-Hadid Q, Laganowsky A, Young BD, Lipson RS, Souda P, Faulf KF, Whitelegg JP, Clarke SG. 2010. A novel 3-methylhistidine modification of yeast ribosomal protein Rpl3 is dependent upon the YIL110W methyltransferase. J Biol Chem. 285(48):37598–37606.

Wiederstein JL, Nolte H, Günther S, Piller T, Baraldo M, Kostis S, Bloch W, Schindler N, Sandrin M, Blauw B, et al. 2018. Skeletal muscle-specific methyltransferase METTL21C Trimethylates p97 and regulates autophagy-associated protein breakdown. Cell Rep. 23(5):1342–1356.

Wiegand G, Remington SJ. 1986. Citrate synthase: structure, control, and mechanism. Annu Rev Biophys Chem. 15:97–117.

Woodcock CB, Yu D, Hajan T, Li J, Huang Y, Dai N, Corréa IR, Wu T, Vedadi M, Zhang X, et al. 2019. Human Mettl3—Mettl14 complex is a sequence-specific DNA adenine methyltransferase active on single-strand and unpaired DNA in vitro. Cell Discov. 5:63.

Xu L, Liu X, Sheng N, Oo KS, Liang J, Chiong YH, Xu J, Ye F, Gao Y-G, Dedon PC, et al. 2017. Three distinct 3-methylcytidine (mC) methyltransferases modify tRNA and mRNA in mice and humans. J Biol Chem. 292(35):14695–14703.

Xu Z, O’Farrell HC, Rife JP, Culver GM. 2008. A conserved mRNA methyltransferase regulates ribosome biogenesis. Nat Struct Mol Biol. 15(5):533–536.

Yang Z. 1994. Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. J Mol Evol. 39(3):306–314.

Zehner JK, Bartz R, Bisel B, Liu P, Seemann J, Anderson RGW. 2009. Targeting sequences of UBXD8 and AAM-B reveal that the ER has a direct role in the emergence and regression of lipid droplets. J Cell Sci. 122(Pt 20):3694–3702.

Zhang L-S, Liu C, Ma H, Dai Q, Sun H-L, Luo G, Zhang Z, Zhang L, Hu L, Dong X, et al. 2019. Transcriptome-wide mapping of internal N7-methylguanosine methylome in Mammalian mRNA. Mol Cell 74(6):1304–1316.e8.

Zhang Z. 1994. Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. J Mol Evol. 39(3):306–314.