Metabolism of RNA


European Bioinformatics Institute, New York University Langone Medical Center, Ontario Institute for Cancer Research, Oregon Health and Science University.

The contents of this document may be freely copied and distributed in any media, provided the authors, plus the institutions, are credited, as stated under the terms of Creative Commons Attribution 4.0 International (CC BY 4.0) License. For more information see our license.

22/09/2021
Introduction

Reactome is open-source, open access, manually curated and peer-reviewed pathway database. Pathway annotations are authored by expert biologists, in collaboration with Reactome editorial staff and cross-referenced to many bioinformatics databases. A system of evidence tracking ensures that all assertions are backed up by the primary literature. Reactome is used by clinicians, geneticists, genomics researchers, and molecular biologists to interpret the results of high-throughput experimental studies, by bioinformaticians seeking to develop novel algorithms for mining knowledge from genomic studies, and by systems biologists building predictive models of normal and disease variant pathways.

The development of Reactome is supported by grants from the US National Institutes of Health (P41 HG003751), University of Toronto (CFREF Medicine by Design), European Union (EU STRP, EMI-CD), and the European Molecular Biology Laboratory (EBI Industry program).

Literature references


Reactome database release: 77

This document contains 12 pathways (see Table of Contents)
Metabolism of RNA

Stable identifier: R-HSA-8953854

This superpathway encompasses the processes by which RNA transcription products are further modified covalently and non-covalently to yield their mature forms, and the regulation of these processes. Annotated pathways include ones for capping, splicing, and 3'-cleavage and polyadenylation to yield mature mRNA molecules that are exported from the nucleus (Hocine et al. 2010). mRNA editing and nonsense-mediated decay are also annotated. Processes leading to mRNA breakdown are described: deadenylation-dependent mRNA decay, microRNA-mediated RNA cleavage, and regulation of mRNA stability by proteins that bind AU-rich elements. psRNAP assembly is also annotated here.

The aminoacylation of mature tRNAs is annotated in the "Metabolism of proteins" superpathway, as a part of "Translation".

Literature references


Editions

<table>
<thead>
<tr>
<th>Date</th>
<th>Author/Editor</th>
<th>Person</th>
</tr>
</thead>
<tbody>
<tr>
<td>2016-12-29</td>
<td>Authored, Edited</td>
<td>D'Eustachio, P.</td>
</tr>
<tr>
<td>2016-12-30</td>
<td>Revised</td>
<td>D'Eustachio, P.</td>
</tr>
<tr>
<td>2017-02-28</td>
<td>Reviewed</td>
<td>Gillespie, ME.</td>
</tr>
</tbody>
</table>
The 5' ends of all eukaryotic pre-mRNAs studied thus far are converted to cap structures. The cap is thought to influence splicing of the first intron, and is bound by 'cap-binding' proteins, CBP80 and CBP20, in the nucleus. The cap is important for translation initiation, and it also interacts with the poly(A)terminus, via proteins, resulting in circularization of the mRNA to facilitate multiple rounds of translation. The cap is also important for mRNA stability, protecting it from 5' to 3' nucleases, and is required for mRNA export to the cytoplasm.

The capping reaction usually occurs very rapidly on nascent transcripts; after the synthesis of only a few nucleotides by RNA polymerase II. The capping reaction involves the conversion of the 5'-end of the nascent transcript from a triphosphate to a diphosphate by a RNA 5'-triphosphatase, followed by the addition of a guanosine monophosphate by the mRNA guanylyltransferase, to form a 5'-5'-triphosphate linkage. This cap is then methylated by 2'-O-methyltransferases.

**Literature references**


Processing of Capped Intron-Containing Pre-mRNA

Location: Metabolism of RNA

Stable identifier: R-HSA-72203

Compartments: nucleoplasm, nuclear envelope, cytosol

Co-transcriptional pre-mRNA splicing is not obligatory. Pre-mRNA splicing begins co-transcriptionally and often continues post-transcriptionally. Human genes contain an average of nine introns per gene, which cannot serve as splicing substrates until both 5' and 3' ends of each intron are synthesized. Thus the time that it takes for pol II to synthesize each intron defines a minimal time and distance along the gene in which splicing factors can be recruited. The time that it takes for pol II to reach the end of the gene defines the maximal time in which splicing could occur co-transcriptionally. Thus, the kinetics of transcription can affect the kinetics of splicing. Any covalent change in a primary (nascent) mRNA transcript is mRNA Processing. For successful gene expression, the primary mRNA transcript needs to be converted to a mature mRNA prior to its translation into polypeptide. Eucaryotic mRNAs undergo a series of complex processing reactions; these begin on nascent transcripts as soon as a few ribonucleotides have been synthesized during transcription by RNA Polymerase II, through the export of the mature mRNA to the cytoplasm, and culminate with mRNA turnover in the cytoplasm.

Editions

2003-06-05 Author: Wahle, E., Carmichael, GG., Hammarskjold, ML., Hastings, ML., Marzluff, WF., Krainer, AR. et al.
2021-05-18 Edited: Gillespie, ME., Gopinathrao, G., Joshi-Tope, G.
2021-05-18 Reviewed: Nilsen, TW.
Co-transcriptional pre-mRNA splicing is not obligatory. Pre-mRNA splicing begins co-transcriptionally and often continues post-transcriptionally. Human genes contain an average of nine introns per gene, which cannot serve as splicing substrates until both 5' and 3' ends of each intron are synthesized. Thus the time that it takes for pol II to synthesize each intron defines a minimal time and distance along the gene in which splicing factors can be recruited. The time that it takes for pol II to reach the end of the gene defines the maximal time in which splicing could occur co-transcriptionally. Thus, the kinetics of transcription can affect the kinetics of splicing.

There are two classes of intronless pre-mRNAs (mRNAs expressed from genes that lack introns). The first class encodes the replication dependent histone mRNAs. These mRNAs have unique 3' ends that do not have a polyA tail. The replication dependent histone mRNAs in all metazoans, as well as Chlamydomonas and Volvox fall into this class.

The second class of mRNAs end in polyA tails, which are formed by a mechanism similar to that which poly-adenylate pre-mRNAs containing introns. In the intronless genes there is a different method of replacing the 3' splice site that activates polyadenylation.

**Literature references**


**Editions**

<table>
<thead>
<tr>
<th>Date</th>
<th>Activity</th>
<th>By</th>
</tr>
</thead>
<tbody>
<tr>
<td>2003-08-22</td>
<td>Authored</td>
<td>Marzluff, WF.</td>
</tr>
<tr>
<td>2021-05-18</td>
<td>Edited</td>
<td>Joshi-Tope, G.</td>
</tr>
</tbody>
</table>
mRNA Editing

Location: Metabolism of RNA

Stable identifier: R-HSA-75072

Compartments: nucleoplasm

After transcription, some RNA molecules are altered to contain bases not encoded in the genome. Most often this involves the editing or modification of one base to another, but in some organisms can involve the insertion or deletion of a base. Such editing events alter the coding properties of mRNA.

RNA editing can be generally defined as the co- or post transcriptional modification of the primary sequence of RNA from that encoded in the genome through nucleotide deletion, insertion, or base modification mechanisms.

There are two pathways of RNA editing: the substitution/conversion pathway and the insertion/deletion pathway. The insertion/deletion editing occurs in protozoans like Trypanosoma, Leishmania; in slime molds like Physarum spp., and in some viral categories like paramyxoviruses, Ebola virus etc. To date, the substitution/conversion pathway has been observed in human along with other mammals, Drosophila, and some plants. The RNA editing processes are known to create diversity in proteins involved in various pathways like lipid transport, metabolism etc. and may act as potential targets for therapeutic intervention (Smith et al., 1997).

The reaction mechanisms of cytidine and adenosine deaminases is represented below. In both these reactions, NH3 is presumed to be released:

**Literature references**


<table>
<thead>
<tr>
<th>Date</th>
<th>Action</th>
<th>Author</th>
</tr>
</thead>
<tbody>
<tr>
<td>2003-08-22</td>
<td>Authored</td>
<td>Carmichael, GG.</td>
</tr>
<tr>
<td>2021-05-18</td>
<td>Edited</td>
<td>Gopinathrao, G.</td>
</tr>
</tbody>
</table>
Regulation of mRNA stability by proteins that bind AU-rich elements

**Location:** Metabolism of RNA

**Stable identifier:** R-HSA-450531

**Compartments:** cytosol, nucleoplasm

RNA elements rich in adenine and uracil residues (AU-rich elements) bind specific proteins which either target the RNA for degradation or, more rarely, stabilize the RNA. The activity of the AU-element binding proteins is regulated, usually by phosphorylation but also by subcellular localization.

**Literature references**


**Editions**

<table>
<thead>
<tr>
<th>Date</th>
<th>Action</th>
<th>Author</th>
</tr>
</thead>
<tbody>
<tr>
<td>2009-12-29</td>
<td>Authored, Edited</td>
<td>May, B.</td>
</tr>
<tr>
<td>2010-06-29</td>
<td>Reviewed</td>
<td>Wilusz, J.</td>
</tr>
</tbody>
</table>

https://reactome.org
Insulin-like Growth Factor-2 mRNA Binding Proteins (IGF2BPs/IMPs/VICKZs) bind RNA

**Location:** Metabolism of RNA

**Stable identifier:** R-HSA-428359

**Compartments:** cytosol

Insulin-like Growth Factor-2 mRNA Binding Proteins (IGF2BPs) bind specific sets of RNA and regulate their translation, stability, and subcellular localization. IGF2BP1, IGF2BP2, and IGF2BP3 bind about 8400 protein-coding transcripts. The target RNAs contain the sequence motif CAUH (where H is A, U, or, C) and binding of IGFBPs increases the stability of the target RNAs.

**Literature references**


**Editions**

<table>
<thead>
<tr>
<th>Date</th>
<th>Action</th>
<th>Author/Reviewer</th>
</tr>
</thead>
<tbody>
<tr>
<td>2009-07-05</td>
<td>Authored, Edited</td>
<td>May, B.</td>
</tr>
<tr>
<td>2010-05-30</td>
<td>Reviewed</td>
<td>Singer, RH., Chao, JA.</td>
</tr>
</tbody>
</table>
Deadenylation-dependent mRNA decay

Location: Metabolism of RNA

Stable identifier: R-HSA-429914

Compartments: cytosol

After undergoing rounds of translation, mRNA is normally destroyed by the deadenylation-dependent pathway. Though the trigger is unclear, deadenylation likely proceeds in two steps: one catalyzed by the PAN2-PAN3 complex that shortens the poly(A) tail from about 200 adenosine residues to about 80 residues and one catalyzed by the CCR4-NOT complex or by the PARN enzyme that shortens the tail to about 10-15 residues.

After deadenylation the mRNA is then hydrolyzed by either the 5' to 3' pathway or the 3' to 5' pathway. It is unknown what determinants target a mRNA to one pathway or the other.

The 5' to 3' pathway is initiated by binding of the Lsm1-7 complex to the 3' oligoadenylate tail followed by decapping by the DCP1-DCP2 complex. The 5' to 3' exoribonuclease XRN1 then hydrolyzes the remaining RNA.

The 3' to 5' pathway is initiated by the exosome complex at the 3' end of the mRNA. The exosome progressively hydrolyzes the mRNA from 3' to 5', leaving only a capped oligoribonucleotide. The cap is then removed by the scavenging decapping enzyme DCPS.

Literature references


https://reactome.org

**Editions**

<table>
<thead>
<tr>
<th>Date</th>
<th>Activity</th>
<th>Person</th>
</tr>
</thead>
<tbody>
<tr>
<td>2009-07-22</td>
<td>Authored, Edited</td>
<td>May, B.</td>
</tr>
<tr>
<td>2009-09-17</td>
<td>Reviewed</td>
<td>Wilusz, J.</td>
</tr>
</tbody>
</table>
The Nonsense-Mediated Decay (NMD) pathway activates the destruction of mRNAs containing premature termination codons (PTCs) (reviewed in Isken and Maquat 2007, Chang et al. 2007, Behm-Ansmant et al. 2007, Neu-Yilik and Kulozik 2008, Rebbapragada and Lykke-Andersen 2009, Bhuvanagiri et al. 2010, Nicholson et al. 2010, Durand and Lykke-Andersen 2011). In mammalian cells a termination codon can be recognized as premature if it precedes an exon-exon junction by at least 50-55 nucleotides or if it is followed by an abnormal 3' untranslated region (UTR). While length of the UTR may play a part, the qualifications for being "abnormal" have not been fully elucidated. Also, some termination codons preceding exon junctions are not degraded by NMD so the criteria for triggering NMD are not yet fully known (reviewed in Rebbapragada and Lykke-Andersen 2009). While about 30% of disease-associated mutations in humans activate NMD, about 10% of normal human transcripts are also degraded by NMD (reviewed in Stalder and Muhlemann 2008, Neu-Yilik and Kulozik 2008, Bhuvanagiri et al. 2010, Nicholson et al. 2010). Thus NMD is a normal physiological process controlling mRNA stability in unmutated cells.

Exon junction complexes (EJCs) are deposited on an mRNA during splicing in the nucleus and are displaced by ribosomes during the first round of translation. When a ribosome terminates translation the A site encounters the termination codon and the eRF1 factor enters the empty A site and recruits eRF3. Normally, eRF1 cleaves the translated polypeptide from the tRNA in the P site and eRF3 interacts with Polyadenylate-binding protein (PABP) bound to the polyadenylated tail of the mRNA.

During activation of NMD eRF3 interacts with UPF1 which is contained in a complex with SMG1, SMG8, and SMG9. NMD can arbitrarily be divided into EJC-enhanced and EJC-independent pathways. In EJC-enhanced NMD, an exon junction is located downstream of the PTC and the EJC remains on the mRNA after
termination of the pioneer round of translation. The core EJC is associated with UPF2 and UPF3, which interact with UPF1 and stimulate NMD. Once bound near the PTC, UPF1 is phosphorylated by SMG1. The phosphorylation is the rate-limiting step in NMD and causes UPF1 to recruit either SMG6, which is an endoribonuclease, or SMG5 and SMG7, which recruit ribonucleases. SMG6 and SMG5:SMG7 recruit phosphatase PP2A to dephosphorylate UPF1 and allow further rounds of degradation. How EJC-independent NMD is activated remains enigmatic but may involve competition between PABP and UPF1 for eRF3.

**Literature references**


**Editions**

<table>
<thead>
<tr>
<th>Date</th>
<th>Action</th>
<th>Author(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2010-08-06</td>
<td>Authored, Edited</td>
<td>May, B.</td>
</tr>
<tr>
<td>2011-05-19</td>
<td>Reviewed</td>
<td>Neu-Yilik, G.</td>
</tr>
</tbody>
</table>
Each eukaryotic cytosolic ribosome contains 4 molecules of RNA: 28S rRNA (25S rRNA in yeast), 5.8S rRNA, and 5S rRNA in the 60S subunit and 18S rRNA in the 40S subunit. The 18S rRNA, 5.8S rRNA, and 28S rRNA are produced by endonucleolytic and exonucleolytic processing of a single 47S precursor (pre-rRNA) (reviewed in Henras et al. 2015). Transcription of ribosomal RNA genes, processing of pre-rRNA, and assembly of precursor 60S and 40S subunits occur in the nucleolus (reviewed in Hernandez-Verdun et al. 2010), with a few late reactions occurring in the cytosol. Within the nucleolus non-transcribed DNA and inactive polymerase complexes are located in the fibrillar center, active DNA polymerase I transcription occurs at the interface between the fibrillar center and the dense fibrillar component, early processing of pre-rRNA occurs in the dense fibrillar component, and late processing of pre-rRNA occurs in the granular component (Stanek et al. 2001).

Processed ribosomal RNA contains many modified nucleotides which are generated by enzymes acting on encoded nucleotides contained in the precursor rRNA (reviewed in Boschi-Muller and Motorin 2013). The most numerous modifications are pseudouridine residues and 2’-O-methylribonucleotides. Pseudouridylation is guided by base pairing between the precursor rRNA and a small nucleolar RNA (snoRNA) in a Box C/D snoRNP (reviewed in Henras et al 2004, Yu and Meier 2014). Similarly, 2’-O-methylation is guided by base pairing between the precursor rRNA and a snoRNA in a Box H/ACA snoRNP (reviewed in Henras et al. 2004, Hamma and Ferre-D’Amare 2010). Other modifications include N(1)-methylpseudouridine, 5-methylcytosine, 7-methylguanosine, 6-dimethyladenosine, and 4-acetylcytidine. Modification of nucleotides occur as the pre-rRNA is being cleaved. However, the order of cleavage and modification steps is not clear so these two processes are presented separately here. Defects in ribosome biogenesis factors can cause disease (reviewed in Freed et al. 2010).
Mitochondrial ribosomes are completely distinct from cytoplasmic ribosomes, having different protein subunits and 12S rRNA and 16S rRNA. The mitochondrial rRNAs are encoded in the mitochondrial genome and are produced by processing of a long H strand transcript. Specific residues in the rRNAs are modified by enzymes to yield 5 different types of modified nucleotides:

**Literature references**


**Editions**

<table>
<thead>
<tr>
<th>Date</th>
<th>Authorship</th>
<th>Person</th>
</tr>
</thead>
<tbody>
<tr>
<td>2003-07-15</td>
<td>Authored</td>
<td>Joshi-Tope, G.</td>
</tr>
<tr>
<td>2015-08-12</td>
<td>Authored, Edited</td>
<td>May, B.</td>
</tr>
<tr>
<td>2016-01-30</td>
<td>Reviewed</td>
<td>Vincent, NG.</td>
</tr>
</tbody>
</table>
**tRNA processing**

**Location:** Metabolism of RNA

**Stable identifier:** R-HSA-72306

Genes encoding transfer RNAs (tRNAs) are transcribed by RNA polymerase III in the nucleus and by mitochondrial RNA polymerase in the mitochondrion.

In the nucleus transcription reactions produce precursor tRNAs (pre-tRNAs) that have extra 5' leaders, 3' trailers, and, in some cases, introns which are removed by enzymes and enzyme complexes: RNase P cleaves the 5' leader, RNase Z cleaves the 3' trailer, TRNT1 polymerizes CCA onto the resulting 3' end, the TSEN complex cleaves at each end of the intron, and the tRNA ligase complex ligates the resulting exons (reviewed in Rossmanith et al. 1995, Phizicky and Hopper 2010, Suzuki et al. 2011, Abbott et al. 2014, Li and Mason 2014). The nucleotides within tRNAs undergo further chemical modifications such as methylation and deamination by a diverse set of enzymes (reviewed in Helm and Alfonzo 2014, Boschi-Muller and Motorin 2013). The order of events for each tRNA is not fully known and the understanding of the overall process is complicated by the retrograde (cytosol to nucleus) transport of tRNAs.

In the mitochondrial matrix transcription produces long precursor RNAs, H strand transcripts and an L strand transcript, that are cleaved by mitochondrial RNase P (an entirely proteinaceous complex), ELAC2, and other nucleases to yield 12S rRNA, 16S rRNA, mRNAs, and pre-tRNAs lacking 3' CCA sequences (reviewed in Van Haute et al. 2015). TRNT1 polymerizes an untemplated CCA sequence onto the 3' ends of the pre-tRNAs and chemical modifications are made to several nucleotides in the tRNAs.

**Literature references**


[https://reactome.org](https://reactome.org)


**Editions**

<table>
<thead>
<tr>
<th>Date</th>
<th>Action</th>
<th>Author</th>
</tr>
</thead>
<tbody>
<tr>
<td>2003-07-15</td>
<td>Authored</td>
<td>Joshi-Tope, G.</td>
</tr>
<tr>
<td>2015-05-29</td>
<td>Authored, Edited</td>
<td>May, B.</td>
</tr>
<tr>
<td>2015-08-11</td>
<td>Reviewed</td>
<td>Levinger, L.</td>
</tr>
<tr>
<td>2015-08-25</td>
<td>Reviewed</td>
<td>Motorin, Y.</td>
</tr>
<tr>
<td>2015-10-24</td>
<td>Reviewed</td>
<td>Jarrous, N.</td>
</tr>
</tbody>
</table>
The term non-coding is commonly employed for RNA that does not encode a protein, but this does not mean that such RNAs do not contain information nor have function. There is considerable evidence that the majority of mammalian and other complex organism's genomes is transcribed into non-coding RNAs, many of which are alternatively spliced and/or processed into smaller products. Around 98% of all transcriptional output in humans is non-coding RNA. RNA-mediated gene regulation is widespread in higher eukaryotes and complex genetic phenomena like RNA interference are mediated by such RNAs. These non-coding RNAs are a growing list and include rRNAs, tRNAs, snRNAs, snoRNAs siRNAs, 7SL RNA, 7SK RNA, the RNA component of RNase P RNA, the RNA component of RNase MRP, and the RNA component of telomerase.
# Table of Contents

- **Introduction** 1
- **Metabolism of RNA** 2
  - mRNA Capping 3
  - Processing of Capped Intron-Containing Pre-mRNA 4
  - Processing of Capped Intronless Pre-mRNA 5
- **mRNA Editing** 7
- **Regulation of mRNA stability by proteins that bind AU-rich elements** 9
- **Insulin-like Growth Factor-2 mRNA Binding Proteins (IGF2BPs/IMPs/VICKZs) bind RNA** 10
- **Deadenylation-dependent mRNA decay** 11
- **Nonsense-Mediated Decay (NMD)** 13
- **rRNA processing** 15
- **tRNA processing** 17
- **Metabolism of non-coding RNA** 19

Table of Contents 20