So far, we have looked at the mechanism by which the information in genes (DNA) is transcribed into RNA. The newly made RNA, also known as the primary transcript (the product of transcription is known as a transcript) is further processed before it is functional. Both prokaryotes and eukaryotes process their ribosomal and transfer RNAs.

Figure 5.5.1: RNA splicing

The major difference in RNA processing, however, between prokaryotes and eukaryotes, is in the processing of messenger RNAs. We will focus on the processing of mRNAs in this discussion. You will recall that in bacterial cells, the mRNA is translated directly as it comes off the DNA template. In eukaryotic cells, RNA synthesis, which occurs in the nucleus, is separated from the protein synthesis machinery, which is in the cytoplasm. In addition, eukaryotic genes have introns, noncoding regions that interrupt the gene’s coding sequence. The mRNA copied from genes containing introns will also therefore have regions that interrupt the information in the gene. These regions must be removed before the mRNA is sent out of the nucleus to be used to direct protein synthesis. The process of removing the introns and rejoining the coding sections or exons, of the mRNA, is called splicing. Once the mRNA has been capped, spliced and had a polyA tail added, it is sent from the nucleus into the cytoplasm for translation.

The initial product of transcription of a protein coding gene is called the pre-mRNA (or primary transcript). After it has been processed and is ready to be exported from the nucleus, it is called the mature mRNA or processed mRNA.

Figure 5.5.2: Steps in processing eukaryotic messenger RNAs

What are the processing steps for messenger RNAs?
In eukaryotic cells, pre-mRNAs undergo three main processing steps:

- Capping at the 5’ end
- Addition of a polyA tail at the 3’ end.
- Splicing to remove introns

In the capping step of mRNA processing, a 7-methyl guanosine (shown at left) is added at the 5’ end of the mRNA. The cap protects the 5’ end of the mRNA from degradation by nucleases and also helps to position the mRNA correctly on the ribosomes during protein synthesis.
The 3' end of a eukaryotic mRNA is first trimmed, then an enzyme called PolyA Polymerase adds a "tail" of about 200 ‘A’ nucleotides to the 3' end. There is evidence that the polyA tail plays a role in efficient translation of the mRNA, as well as in the stability of the mRNA. The cap and the polyA tail on an mRNA are also indications that the mRNA is complete (i.e., not defective). Introns are removed from the pre-mRNA by the activity of a complex called the spliceosome. The spliceosome is made up of proteins and small RNAs that are associated to form protein-RNA enzymes called small nuclear ribonucleoproteins or snRNPs (pronounced SNURPS). The splicing machinery must be able to recognize splice junctions (i.e., the end of each exon and the start of the next) in order to correctly cut out the introns and join the exons to make the mature, spliced mRNA.

What signals indicate where an intron starts and ends? The base sequence at the start (5' or left end, also called the donor site) of an intron is GU while the sequence at the 3' or right end (a.k.a. acceptor site) is AG. There is also a third important sequence within the intron, called a branch point, that is important for splicing.
Figure 5.5.4: Splicing

There are two main steps in splicing:

- In the first step, the pre-mRNA is cut at the 5’ splice site (the junction of the 5’ exon and the intron). The 5’ end of the intron then is joined to the branch point within the intron. This generates the lariat-shaped molecule characteristic of the splicing process.
- In the second step, the 3’ splice site is cut, and the two exons are joined together, and the intron is released.

Many pre-mRNAs have a large number of exons that can be spliced together in different combinations to generate different mature mRNAs. This is called alternative splicing, and allows the production of many different proteins using relatively few genes, since a single RNA can, by combining different exons during splicing, create many different protein coding messages. Because of alternative splicing, each gene in our DNA gives rise, on average, to three different proteins. Once protein coding messages have been processed by capping, splicing and addition of a poly A tail, they are transported out of the nucleus to be translated in the cytoplasm.

Figure 5.5.5: Splicing and protein diversity

Contributors

- Dr. Kevin Ahern and Dr. Indira Rajagopal (Oregon State University)