**7.013 Recitation 7 – Spring 2018**

*(**Note:** The recitation summary should NOT be regarded as the substitute for lectures)*

## Summary of Lectures 11 (3/5) and 12 (3/7):

**Modification of the two ends of the nascent mRNA:** Transcription results in the formation of the nascent/pre-mature mRNA that contains both introns and exons (coding sequences). Once the pre-mature mRNA is transcribed, both its 5' and 3' ends get modified.

- **5'Cap:** A cap consisting of methylated guanosine is added to the 5′ end of the RNA in opposite polarity to the RNA itself. Thus, whereas all the bases in the RNA precursor are linked 5′ to 3′, the cap structure is linked 5′ to 5′ leaving no free 5′ phosphate group on the nuclear RNA. The 5′ cap is necessary for the stability and translocation of the mRNA from the nucleus into the cytosol; for the binding of mRNA to the ribosome and for its subsequent translation.

- **PolyA tail:** The 3′ terminus of the mRNA is usually modified in the nucleus by having roughly 200 adenylate residues (Poly A tail) added on as a tail. First, the mRNA is cleaved at the 5'AAUAA3" poly A signal and then the Poly A-tail that has been put together enzymatically by the poly A polymerase (PAP) is added to the 3'end of the mRNA.

**Splicing and its significance:** Splicing (in eukaryotes is the process by which pieces of the "nascent" premature mRNA or the introns are removed to form the final, shorter mature mRNA transcript that gets translated into proteins. Splicing helps…

1. By allowing the cell to make multiple mature mRNA transcripts (and hence the proteins) from the same gene through alternative splicing of introns.

2. By providing defense against the transposable elements (viruses that has been stripped of its coat) which can integrate anywhere within the genome to make more of themselves.

What is most remarkable about splicing is that it is the RNA, instead of proteins, that facilitates the enzymatic reaction of cleaving and joining the exons of the mRNA transcripts. The machinery that facilitates splicing is called the spliceosome, which is an RNA-protein complex.

Below are some helpful links:

https://www.youtube.com/watch?v=FVuAwBGw_pQ
http://media.hhmi.org/biointeractive/media/DNAi_transcription_vo1-lg.mov

**Steps of Splicing:** Splicing involves two sequential trans-esterification reactions (**Note:** The details outlined in the steps below are for your own interest)

- **Step 1:** Each intron has a 5'splice donor site (that has a specific 5'GU3' sequence) and a 3'splice acceptor site (that has a specific 5'AG3' sequence) and a branching point that has an A. First, the 2' –OH on the branch point A attacks the phosphoryl group of the G in the 5' splice site. The attack leads to the breaking of the phosphodiester bond between the 3' end of the exon and the 5' end of the intron and the formation of a new phosphodiester bond between the branch point A and the 5' end of the intron. As is illustrated in the figure above, the first splicing step breaks the primary transcript into two pieces. One piece contains the 5' end of an exon, with a new free 3' –OH. The other piece is a strange looking structure that resembles a lasso and as a result is called a lariat structure. The lariat contains the intron that is still connected to a second exon.

- **Step 2:** In the second step of primary transcript splicing, the newly released 3' –OH of exon 1 attacks the phosphoryl group on the 5'-most nucleotide of exon 2. The resulting phosphodiester bond is the boundary between the intron and the second exon. The final products of the reaction are the two joined exons and an intron, which is still in the lariat structure.

Since the spliceosome is an RNA-protein complex, its RNA component base pairs with the 5'end of the intron and the branch point and facilitate the first trans-esterification reaction and then with the 3' end of the intron to facilitate the 2nd reaction. The intron is then released as a lariat structure, which then gets degraded.

**Translation:** Translation begins when the ribosome orients itself at the start codon 5'-AUG-3' on an mRNA. A tRNA that recognizes and base pairs with that codon (i.e. a tRNA with the anticodon 3'-UAC-5') fits into the ribosome and donates the amino acid to which it is bound (methionine) to the new protein. From then on, three nucleotides in the mRNA are read at a time by another tRNA, which is modified with a single amino acid (an amino acyl tRNA). Each time, the next amino acyl-tRNA comes and fits into the ribosome, donates its amino acid to the new protein, waits for the next amino acyl-tRNA to come in, and then exits. This proceeds until a stop codon is reached, and the ribosome dissociates from both the mRNA and the newly made protein. Every protein has an N (amino) and a C (carboxyl) terminus, and proteins are synthesized in the N -> C direction, such that the 5' end of the mRNA corresponds to the N-terminus of the protein and the 3' end of the original template DNA strand.

**Mutations:** Most mutations that geneticists' study are single nucleotide mutations that cause phenotypic changes. The four types of single nucleotide mutations are silent, missense, nonsense, and frame shift. A silent mutation changes a codon but does not change the amino acid encoded by that codon. A missense mutation changes the identity of the amino acid at one position. A nonsense mutation causes a protein to be truncated because a codon is changed to a stop codon. A single nucleotide frame shift is a change that either inserts or deletes a single nucleotide from the coding region of a gene, leading to a change in the reading frame of that gene.

**Process of Gene regulation:** The production of the final, functional protein from a gene can be regulated at many steps.

- The gene may or may not begin to be transcribed, depending on whether the regulatory proteins (activator or repressor) are bound to the regulatory sites of the gene.
- If the gene is transcribed, RNA polymerase may or may not transcribe all the way through till the end of the gene.
- The introns may or may not be spliced out and the message (mRNA) may or may not be transported from the nucleus to the cytoplasm.
- The mRNA may or may not be stable.
- Translation of the message by the ribosome may or may not occur.
- If the protein is made from the message, that protein may or may not be active. (Many proteins require different kinds of covalent modifications, such as phosphorylation, in order to be active.)
- The protein may or may not be stable.
- The protein may or may not need to be transported to a specific subcellular location in order to access its substrate and perform its function.
- The transcription may not happen due to chromatin modification such as altered methylation of bases or de-acetylation of histone. Methylation of DNA inhibits transcription, and the pattern of methylation of DNA varies depending on cell type and cell age.

## Questions:

**1.** The following is a partial sequence from the hypothetical gene, gene X. The boxed region is the promoter, and the arrow indicates the direction of transcription. Transcription begins at and includes the first G/C base pair after the box.

```
5′ ATCAGACAACGTCTCATGGGAGTACTTGGATGGAAGAGTAGAAGGTCATGACCAACCTCTTCCAATCCAACCACAAACAG
   +---------+---------+---------+---------+---------+---------+---------+---------
3′ TAGTCTGTTGCAGAGTACCCTCATGAACCTACCTTCTCATCTTCCAGTACTGGTTGGAGAAGGTTAGGTTGGTGTTTGTC

5′ AAAATCAGCCAATATGTCCGACTTCGAGAACAAGAACCCCAACAACGTCCTTGGCGGACACAAGGCCACCCTTCACAACC
   +---------+---------+---------+---------+---------+---------+---------+---------
3′ TTTTAGTCGGTTATACAGGCTGAAGCTCTTGTTCTTGGGGTTGTTGCAGGAACCGCCTGTGTTCCGGTGGGAAGTGTTGG
```

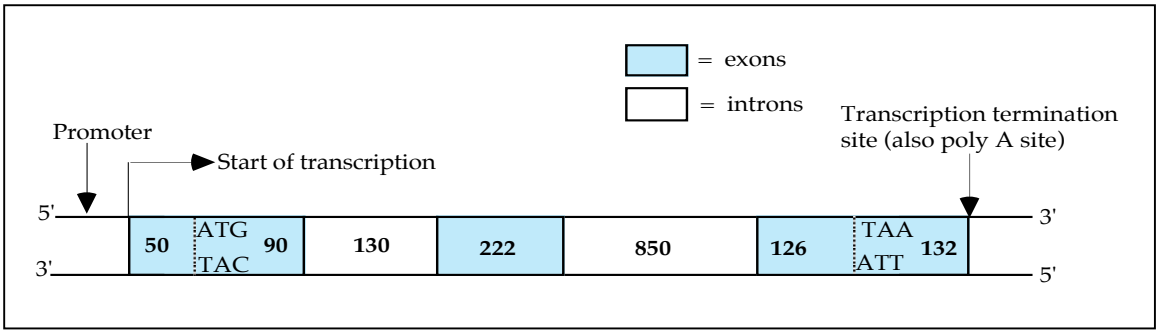
Give the first 10 nucleotides of the mRNA produced from Gene X. _____

**2.** The mature mRNA sequence has some nucleotides that are not represented in the template strand of the genomic DNA sequence.

**a)** What nucleotide(s) is added to the 5’ end of the mature mRNA? Why is it important?

**b)** What nucleotides are added to the 3’ end of the mature mRNA? Why are they important?

**3.** Shown below is the genomic structure of Gene A. The numbers within the boxes indicate the length (in nucleotides) of each region. The DNA sequences corresponding to the start codon and the stop codon are indicated.




**a)** What is the length (in nucleotides) of the nascent mRNA? _____

**b)** What is the length (in nucleotides) of the mature or processed mRNA? _____

**c)** If you were to consider alternative splicing how many mature mRNA transcripts could be produced from this gene? Give the length (in nucleotides) of each. _____

**4.** Assume that the sequence of DNA below is a short protein-encoding gene; the sequences in between the transcription start and stop sites are shown. The entire DNA sequence of the very short gene is:

```
5' CGCTTATAGAACCCAATCTCTCATAGGC  3'
3' GCGAATATCTTGGGTTAGAGAGTATCCG  5'
```

**a)** What would the resulting mRNA be if the top strand of this DNA molecule were used as a template in transcription? Label the 5’ and 3’ ends of your molecule. **5’**_____**3’**

**b)** What is the full sequence of the protein that would be translated from this RNA? Label the N and C termini of your molecule. **N**_____**C**

**c)** What would happen to the encoded protein if the underlined nucleotide **C** were mutated to a **T**?
**N**_____**C**

**5.** Complete the table below.

| | **Translation** |
|---|---|
| Subcellular organelle (s) in eukaryotic cell where translation occurs | |
| Monomer used to form proteins | |
| Rule for adding a type of incoming monomer? | |
| Covalent bond formed between two adjacent monomers in a growing polypeptide chain? | |
| Number and type of template (RNA or DNA) for protein synthesis | |
| In what direction is the template read? | |
| In what direction is the polypeptide chain synthesized? | |

**6.** Drawn below is part of a wild-type gene. The DNA sequence shown encodes the last amino acids of a protein that is normally 380 amino acids long. The **bold** & **underlined** codon indicates the correct reading frame of this gene. The lower strand of the gene is used as the template during the transcription of mRNA from this gene.

…**GCT**AAGTATTGCTCAAGATTAGGATGATAAATAACTGG–3′
…CGATTCATAACGAGTTCTAATCCTACTATTTATTGACC–5′

**a)** In the copy of the sequence drawn above, circle one base pair that you could change to make a mutant form of the gene that produces a protein that is now 381 amino acids long. Indicate the identity of one new base pair that could take its place.

**b)** In the copy of the sequence drawn above, draw a slash between two base pairs where you could add one extra base pair in order to make a single mutant form of the gene that produces a protein that is 373 amino acids long. Indicate the identity of the one new base pair you are adding.

# 7.013 Introductory Biology

Spring 2018