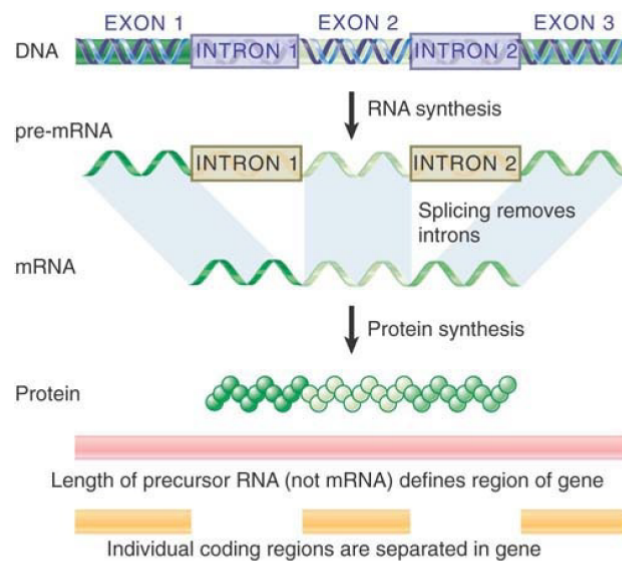


# Chapter 3

## Key Concepts:

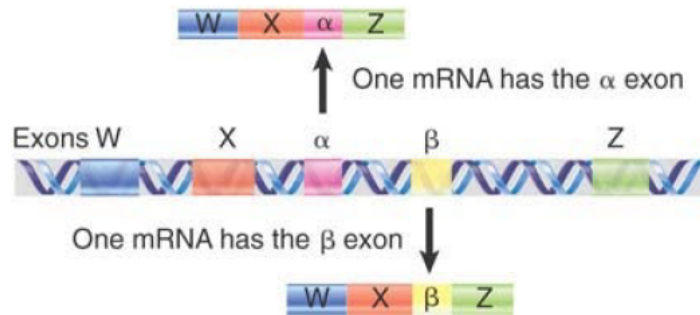
- **Pre-mRNA** (前mRNA) or **primary transcript** (初始转录物): exactly complementary to the original DNA sequence, including exon and intron
  - RNA splicing: delete the introns from pre-mRNA and then join the exons together to form a **mature transcript** (mRNA)
    - Mutations that affect splicing are usually deleterious.
  - The length of a gene is defined by the length of the primary mRNA transcript
    - The length of a gene is determined primarily by the lengths of introns.



**FIGURE 3.1** Interrupted genes are expressed via a precursor RNA. Introns are removed when the exons are spliced together. The mature mRNA has only the sequences of the exons.

- Base in exons and introns: four rules by Erwin Chargaff in 1940s
  - **first parity rule** (第一均等规则): for duplex DNA, A-T and G-C
  - **second parity rule** (第二均等规则): amounts of A and T in each single strand of the duplex are almost equal. So, amounts of C and G in each single strand are also almost equal.
  - **cluster rule** (成簇原则): purines (A and G) tended to be clustered in one strand (usually the nontemplate strand), and pyrimidines (T and C) cluster in the template strand.
  - **GC rule**: GC content tends to be a species-specific character
    - GC content tends to be greater in exons than in introns.
- **Exons and Introns**:
  - homologies (同源性) between genes: exons may be related, but introns are not related at all.
  - negative selection: exons are conserved, but introns vary
  - positive selection: exons vary but introns are conserved

- **Alternative splicing** (可变剪切): a single sequence of DNA encodes more than one polypeptide.



**FIGURE 3.13** Alternative splicing generates the a and b variants of troponin T.

- **Gene family**: a group of genes that encode related or identical products due to mutations in gene duplication
  - Example: globin (珠蛋白) family includes two subfamily ( $\alpha$  globin and  $\beta$  globin)
  - **superfamily** (超家族): genes that are more distantly related but that still can be recognized as having common ancestry.
    - Myoglobin (肌红蛋白) in animals and Leghemoglobins (豆血红蛋白) in legume (豆类) plants are superfamily with  $\alpha$  globin and  $\beta$  globin
  - **Orthologous genes** (orthologs, 种间同源基因) : genes that are homologous (homologs, 同源基因) due to speciation (物种形成); in other words, they are related genes in different species.