# The Human Nuclear Genome

**Dr. Sami Mamand**

Molecular Biology

Semester I

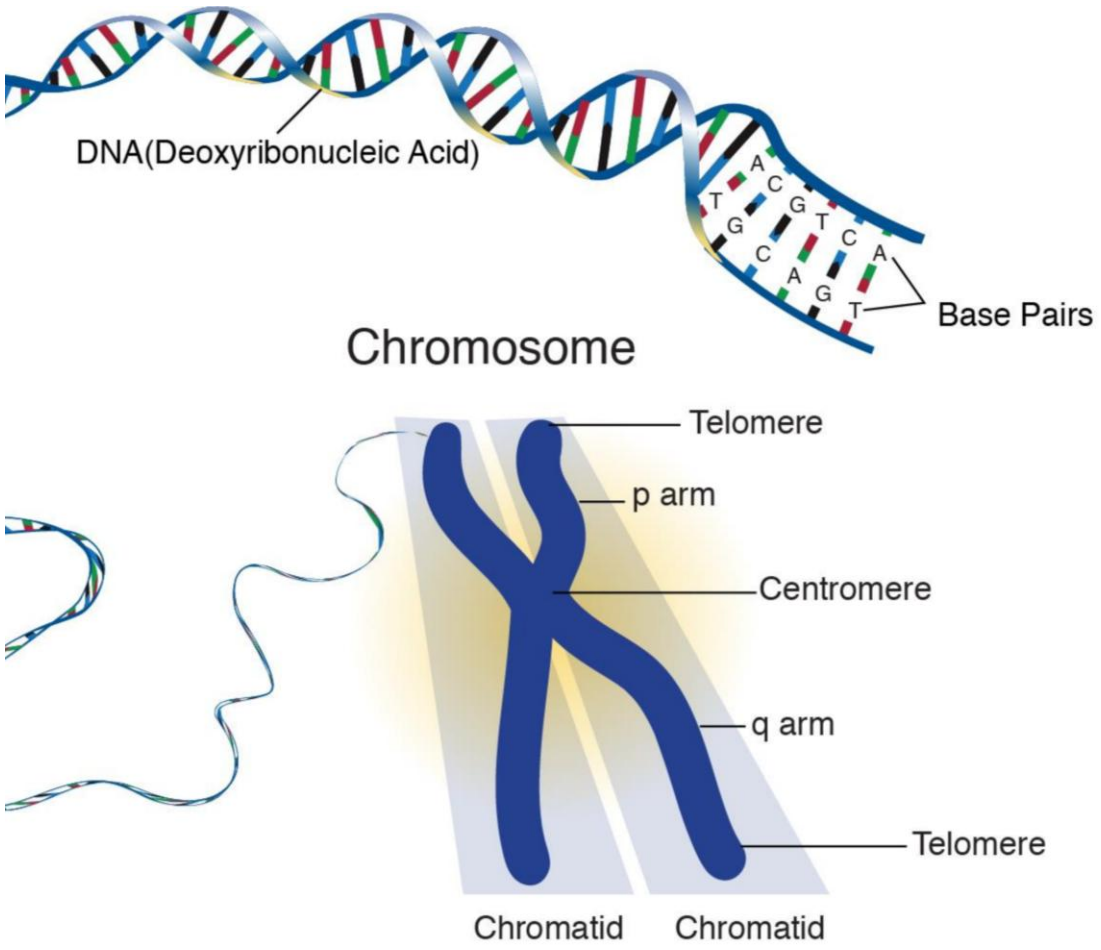
Week 3

10-21-2025

# Objectives

- What the human genome is.
- The number and types of genes in the human genome.
- How nucleic acids (DNA/RNA) control gene expression.
- The difference between coding and non-coding DNA.

- **Chromosome:** A DNA molecule with proteins and RNA that help organize and read genetic information.
- **Gene:** A DNA region that encodes a single polypeptide or functional RNA.
- **Protein:** One or more polypeptides folded and sometimes bound with cofactors.
- **Chromatin:** The complex of DNA + proteins (mostly histones) that form chromosomes in eukaryotes.



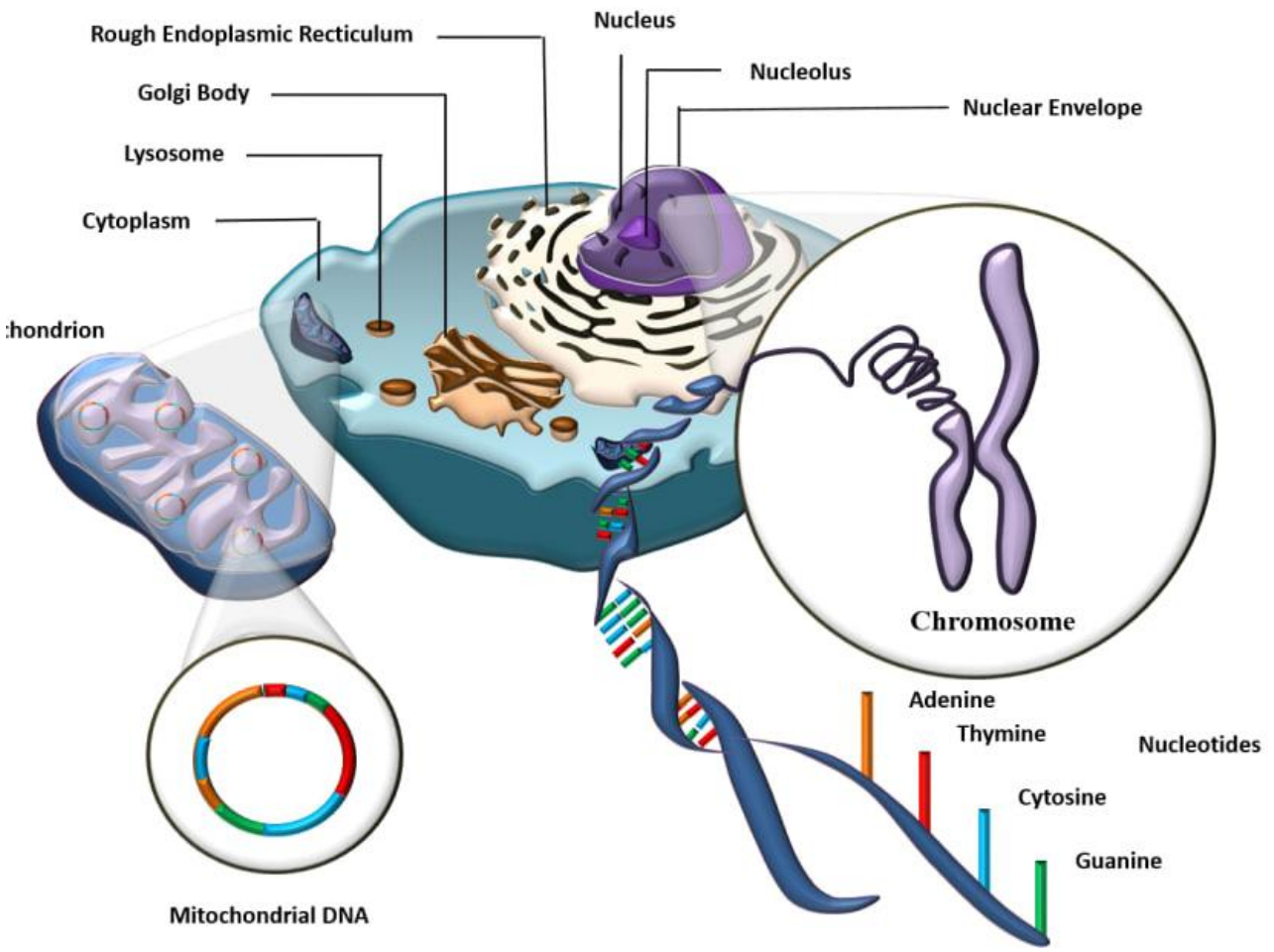
# The human genome

The **genome** is the complete set of an organism's hereditary information encoded in DNA (deoxyribonucleic acid).

In humans, the genome exists mainly in two compartments:

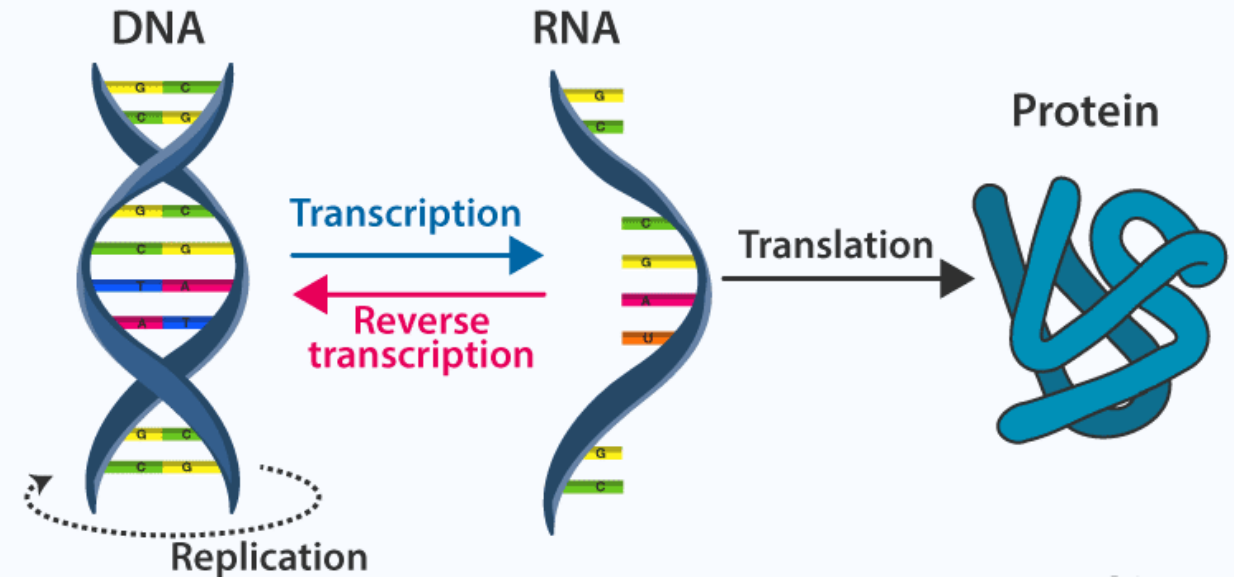
- The nucleus (**nuclear genome**)
- The mitochondria (**mitochondrial genome**)

• The human nuclear genome is the master repository of genetic information, controlling almost every cellular function through transcriptional and post-transcriptional regulation.



# The human genome

- The human nuclear genome contains roughly 19,000–20,000 protein-coding genes.
- The actual regions encoding proteins (exons) occupy only ~1–2 % of the genome's base pairs.
- In addition to protein-coding genes, there are many thousands of non-protein-coding RNA genes which are transcribed into RNA but not translated into protein.

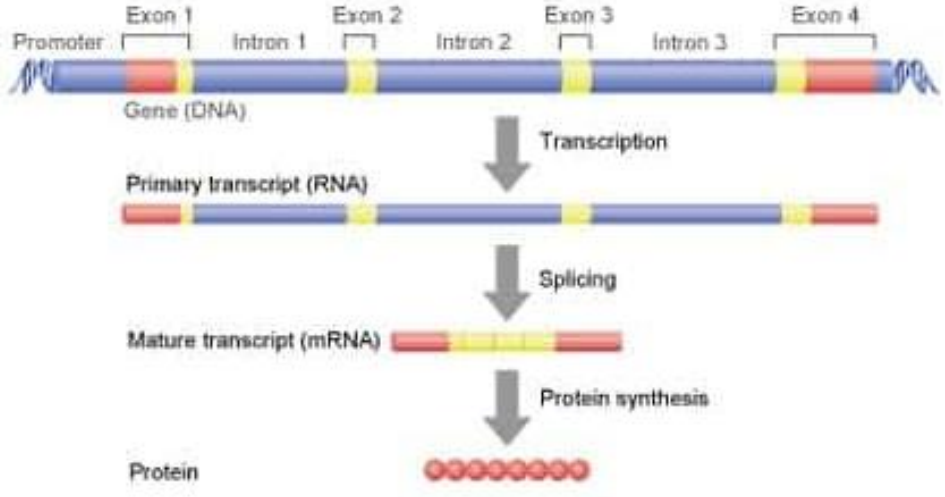


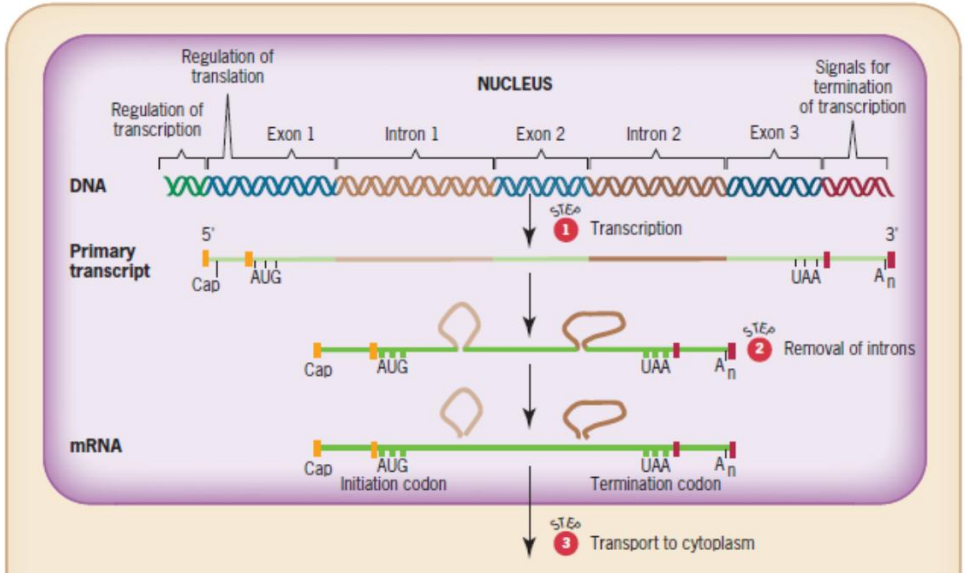
# The human genome

- The fact that the regions encoding proteins in the human nuclear genome occupy only ~1–2 % of the ~3.1 billion base-pairs means that ~98–99 % of the nuclear genome is non-protein-coding (though some of that may still have regulatory or RNA functions).
- The human nuclear genome contains both protein-coding genes and RNA-coding genes (which are transcribed into RNA but not translated into protein), and gene expression is regulated by epigenetic mechanisms.
- The sizes of individual protein-coding genes vary widely: for example, the human insulin gene (INS) is ~1.4 kb, the human insulin-receptor gene (INSR) spans >120 kb, and the human dystrophin gene (DMD) is ~2.4 Mb in size.

# The human genome

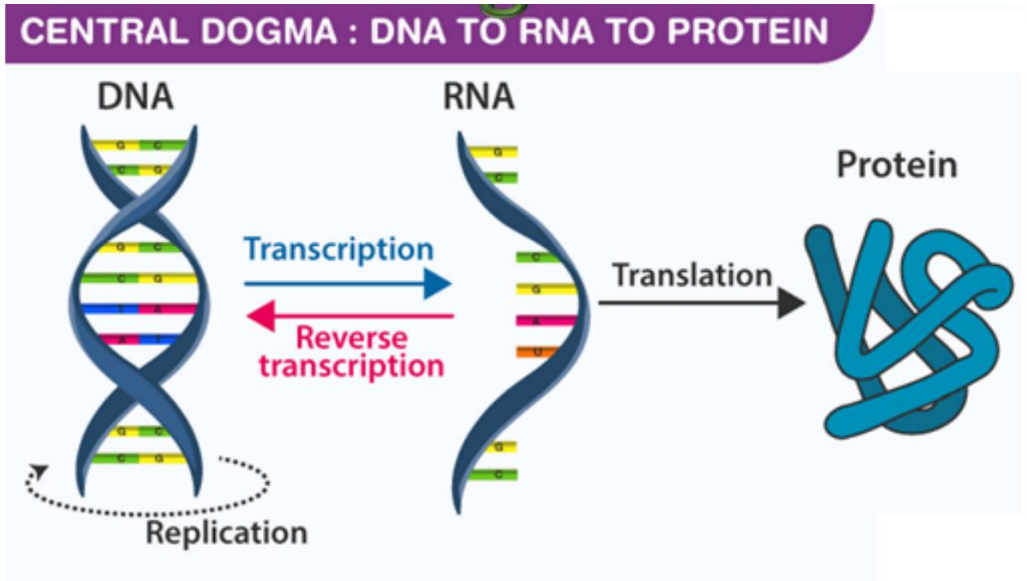
- **Exons** (expressed sequences) are the segments of a gene that remain in the mature RNA after splicing. They typically include the protein-coding regions (and sometimes UTRs) and are relatively short: in human genes, about 80 % of exons are less than ~200 bp in length.
- **Introns** (intervening sequences) are the non-coding segments interspersed between exons in the primary transcript and are removed by splicing. Introns vary widely in size in the human genome — many are several kilobases long (mean on the order of ~3,000–4,000 bp) and some span tens or hundreds of kilobases.

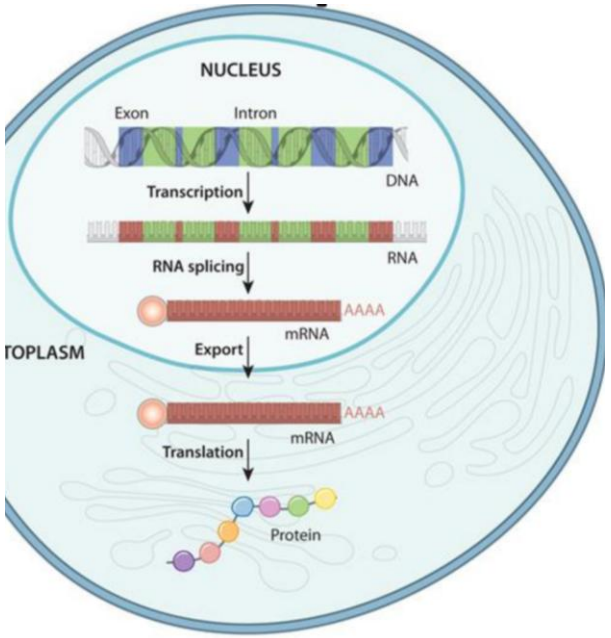


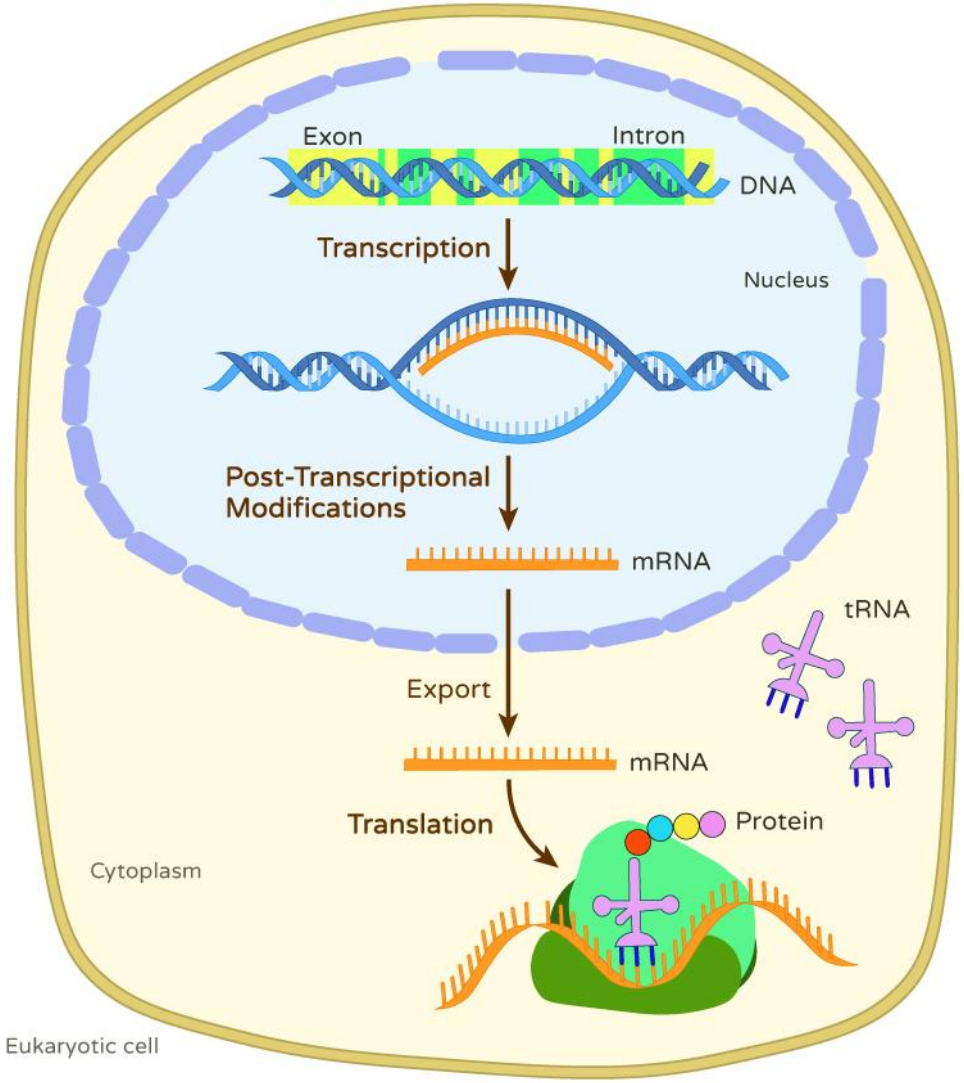
**Gene Expression:** is a fundamental biological process through which genes are transcribed and translated into proteins or RNA molecules.

- **Genes** are expressed in a 2-step process:
  - First, an RNA copy of a single gene is made (transcription).
  - Then, the nucleotide sequence of the RNA copy (messenger RNA) is translated into the amino acid sequence of the polypeptide.
- The **genetic code** is a list of which 3 base DNA or RNA sequence (codon) encodes which amino acid.
- The same genetic code is used in (almost) all organisms.



# Basic Steps of Gene Expression:




- **Transcription**
- **RNA Processing**
  - **Capping:** Addition of a 5′ cap, which protects the RNA and helps in ribosome binding during translation.
  - **Polyadenylation:** Addition of a poly-A tail at the 3′ end, which also protects the RNA and aids in export from the nucleus.
  - **Splicing:** Removal of non-coding regions (introns) and joining of coding regions (exons) to form mature mRNA.
- **Translation**
- **Post-Translational Modifications:**
- After translation, proteins may undergo various modifications (such as **phosphorylation** or **glycosylation**) that can affect their function and activity.

# RNA-Coding Genes:

Produce **active RNAs** that can profoundly alter **normal gene expression** and hence produce an **altered trait or disease**.

**a) 45S and 55 Ribosomal RNA (rRNA) Genes:**

The **rRNA genes** encode for **rRNAs** that are used in **protein synthesis**.

**b) Transfer RNA (tRNA) Genes:**

Transfers amino acids during translation.

- **c) Small Nuclear RNA (snRNA) Genes:**

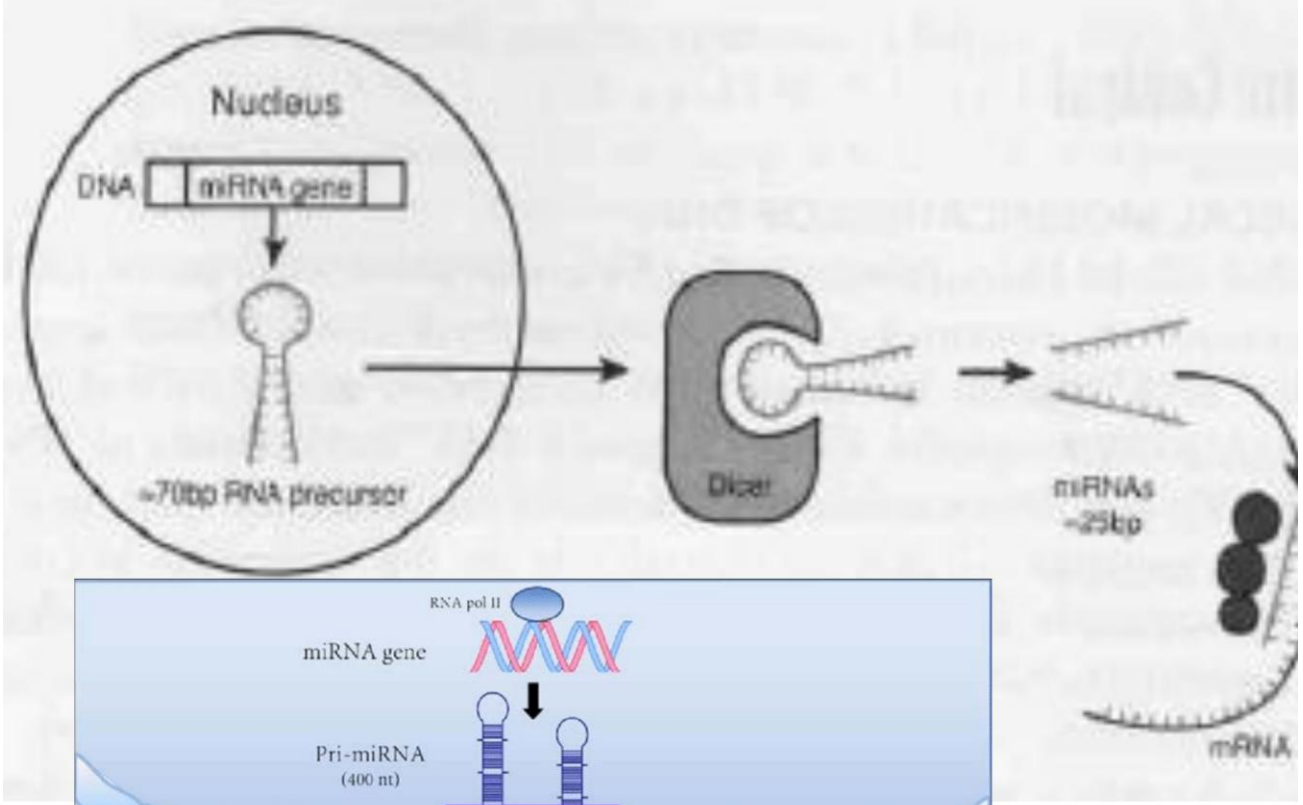
The **snRNA genes** encode for **snRNAs** that are components of the **spliceosome** used in **RNA splicing** during **protein synthesis**. (Involved in splicing pre-mRNA)

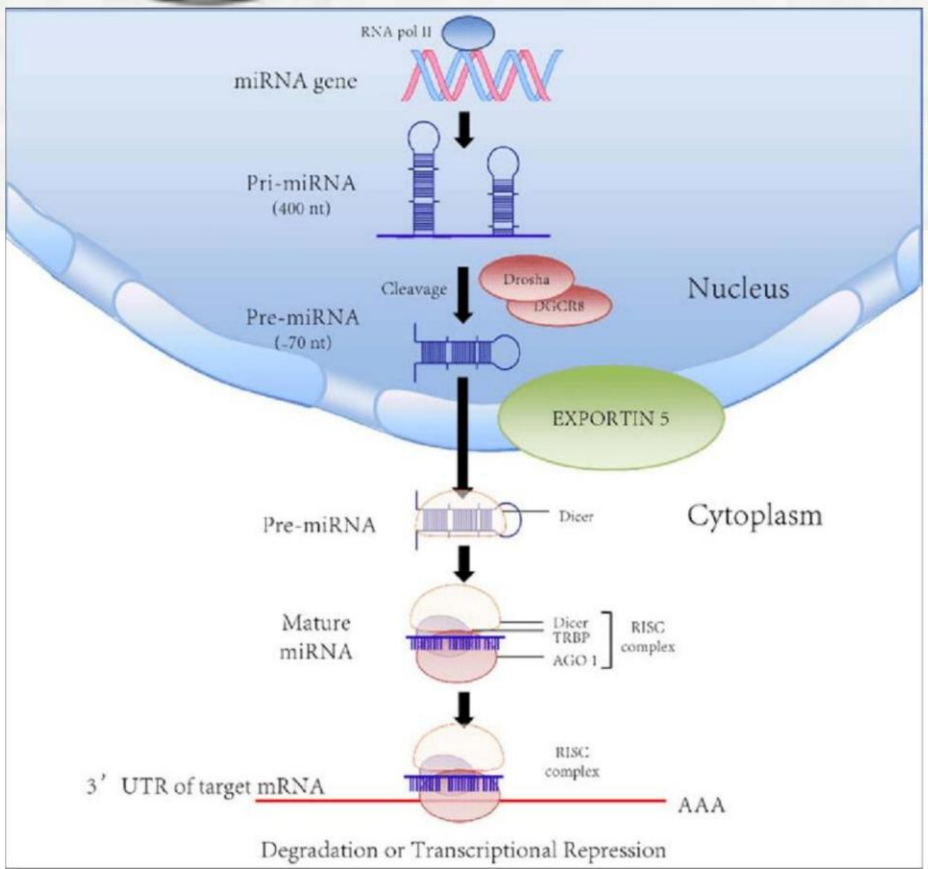
# RNA-Coding Genes:

Produce **active RNAs** that can profoundly alter **normal gene expression** and hence produce an **altered trait or disease**.

**d) REGULATORY RNA Genes:**

- Regulatory RNAs behave like mRNAs (transcribed by RNA Pol II, capped, polyadenylated).
- XIST gene: Produces RNA that coats and silences one X chromosome in females (Barr body formation).
- Function: RNA itself performs regulatory roles, not by coding for proteins but by altering gene expression.




## E. MicroRNA (miRNA) Genes:

**a) miRNA genes** encode for **miRNAs** that **block the expression of other genes**.

**b)** There are ≈ **250 miRNA genes** with the **lin-4 gene**, which encodes **lin-4 miRNA**, and the **let-7 gene**, which encodes **let-7 miRNA** and is the most studied.

**c) microRNAs (miRNAs)** are approximately **22 nucleotides long**, **small noncoding RNAs** that function in **posttranscriptional gene silencing** and play **fundamental roles** in biology.

**d)** The **human genome** contains over **1,800 miRNAs**, which target about **60% of human mRNAs**.

**e) Mutations** in **miRNA genes** and the **miRNA pathway genes** are associated with various human diseases, including **cancers**.

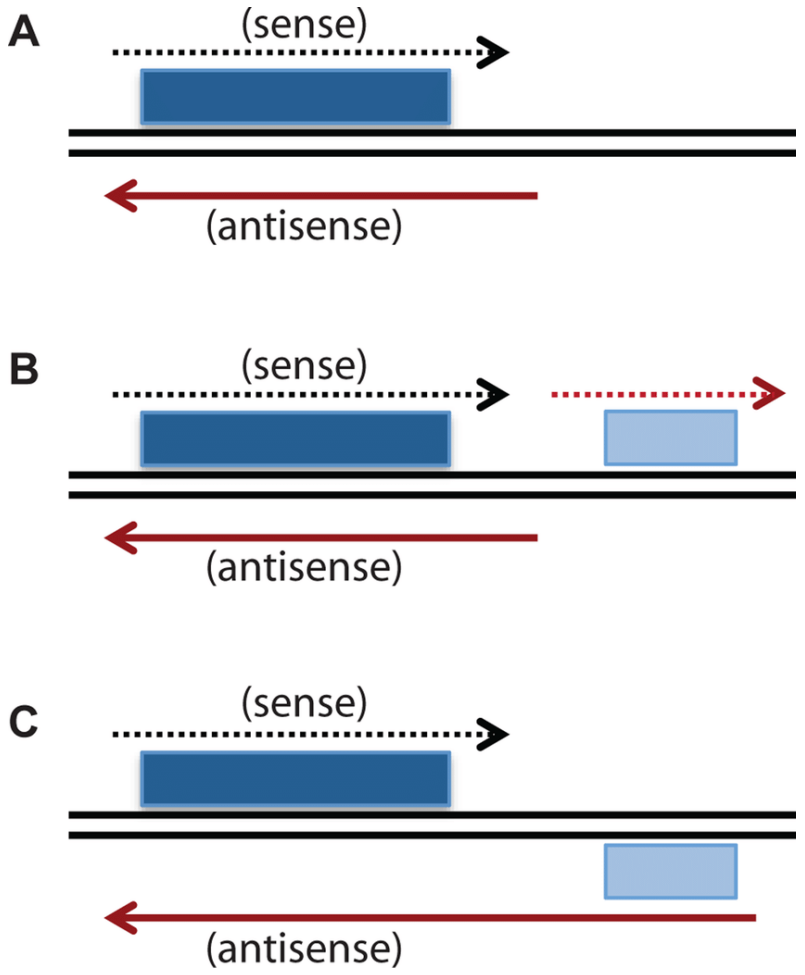
# G. ANTISENSE RNA Genes:

1 **The antisense RNA genes** encode for **antisense RNA** that binds to **mRNA** and physically **blocks translation**.

2 **Antisense RNAs** are conventionally **DNA-encoded transcripts**.

3 **Antisense RNA** is a type of **RNA molecule** that is **complementary** to a specific **mRNA (messenger RNA)** sequence.

4 While the **sense strand of DNA** is used to create **mRNA** that codes for **proteins**, the **antisense RNA** can bind to the corresponding **mRNA**, often leading to **inhibition of translation** or **destabilization of the mRNA**.

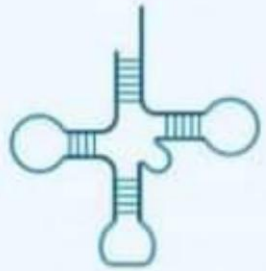


# Types of RNA Produced in Cells

## mRNA

Encodes proteins

## tRNA

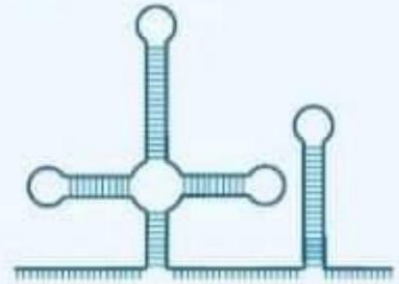


Acts as adaptor between mRNA and amino acids

## rRNA

Forms the ribosome

## snRNA



Functions in various nuclear processes (e.g. splicing)

## snoRNA

Facilitates chemical modification of RNAs

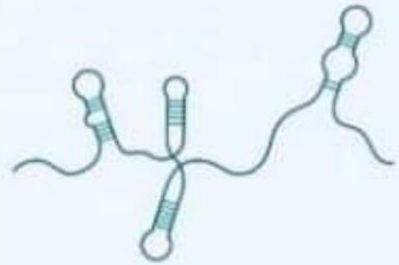
## miRNA

Regulates gene expression

## siRNA

Silences gene expression

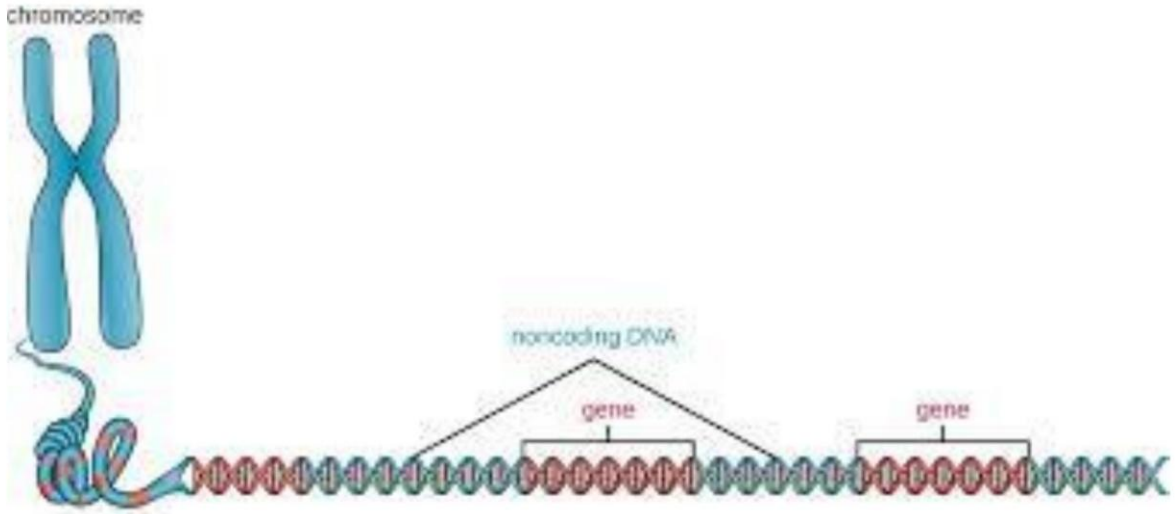
## lncRNA

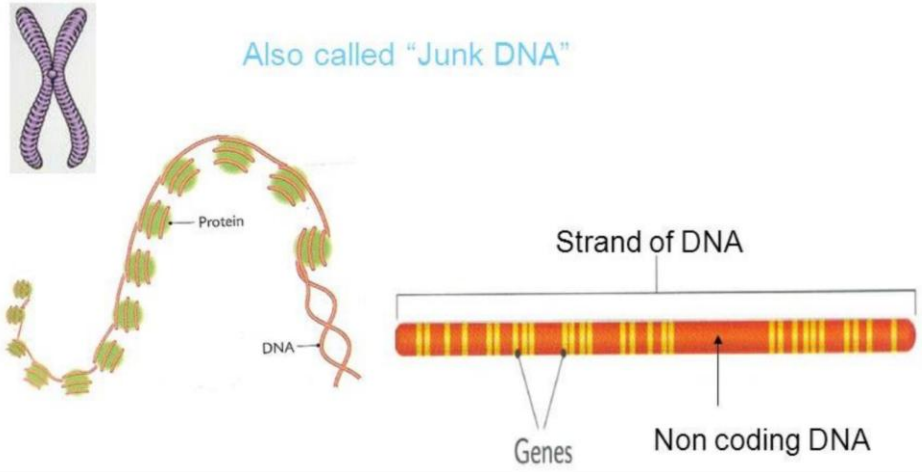


Regulates gene expression

# Noncoding DNA from the Human Nuclear Genome

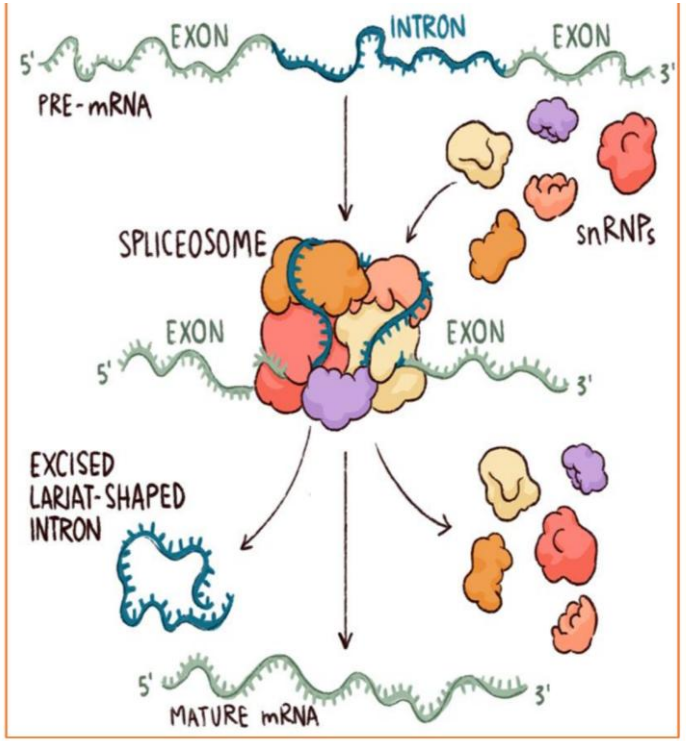
- **Introns** also qualify as **non-coding DNA**, making up about **98.5%** of the total DNA.
- An **intron** is a portion of a **gene** that is **transcribed** into **precursor mRNA (pre-mRNA)** but is **not translated** into a **protein**.
- **Introns** are the **sequences** between the **exons**, and the **intron messages** are **spliced out** by the **spliceosome** during the formation of **coding RNAs** (**mRNA**, **rRNA**, and **tRNA**).

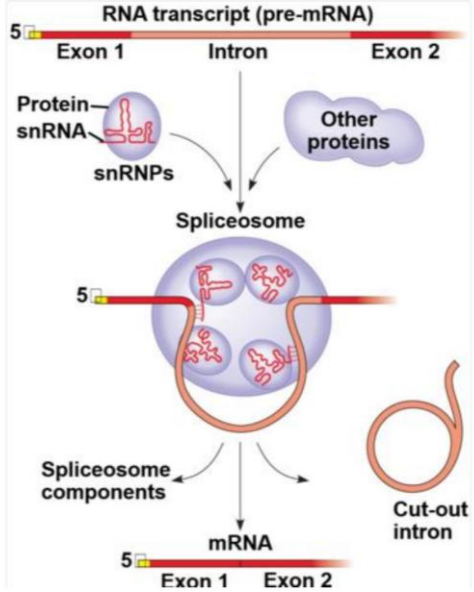


# Noncoding DNA from the Human Nuclear Genome

- After **transcription**, the **pre-mRNA** undergoes a process called **splicing**, where **introns** are removed, and **exons** are joined together to form **mature mRNA**.
- This **mature mRNA** is then **translated** into **protein**.
- The **splicing** is carried out by a **complex** called the **spliceosome**, which recognizes **specific sequences** at the **boundaries of introns and exons**.



# Noncoding DNA from the Human Nuclear Genome

**A. SATELLITE DNA:**
Satellite DNA is composed of very large-sized blocks (1.00 kb → several Mb) of tandemly repeated noncoding DNA.

**B. MINISATELLITE DNA:**
Minisatellite DNA is composed of moderately sized blocks (0.1 kb → >20.0 kb) of tandemly repeated noncoding DNA.

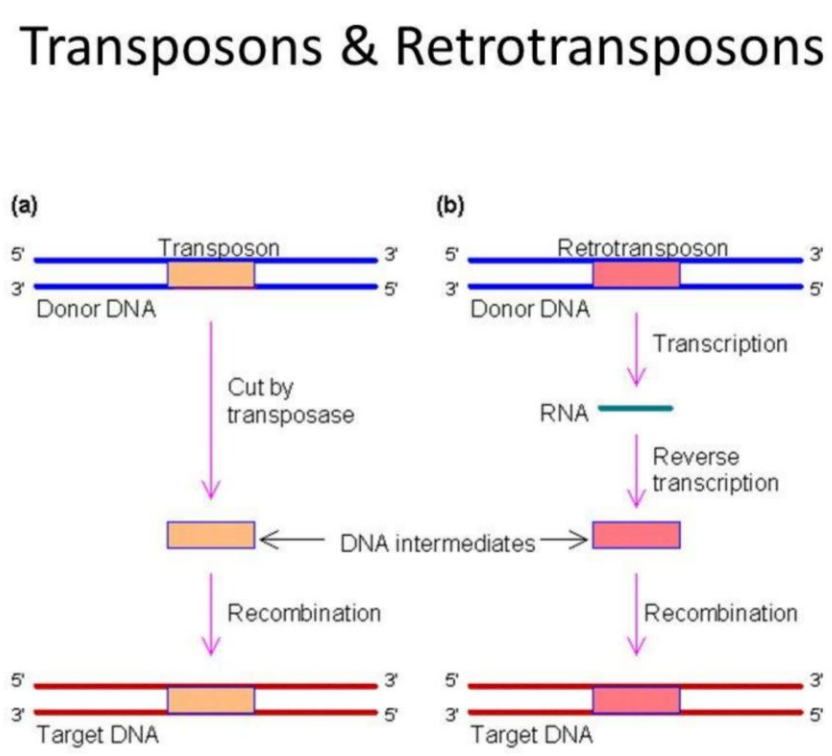
**C. MICROSATELLITE DNA (Simple Sequence Repeat [SSR]):**
Microsatellite DNA is composed of small-sized blocks (< 100 bp) of tandemly repeated noncoding DNA. Microsatellite DNA consists of a <10 bp repeat unit and is found dispersed throughout all chromosomes.

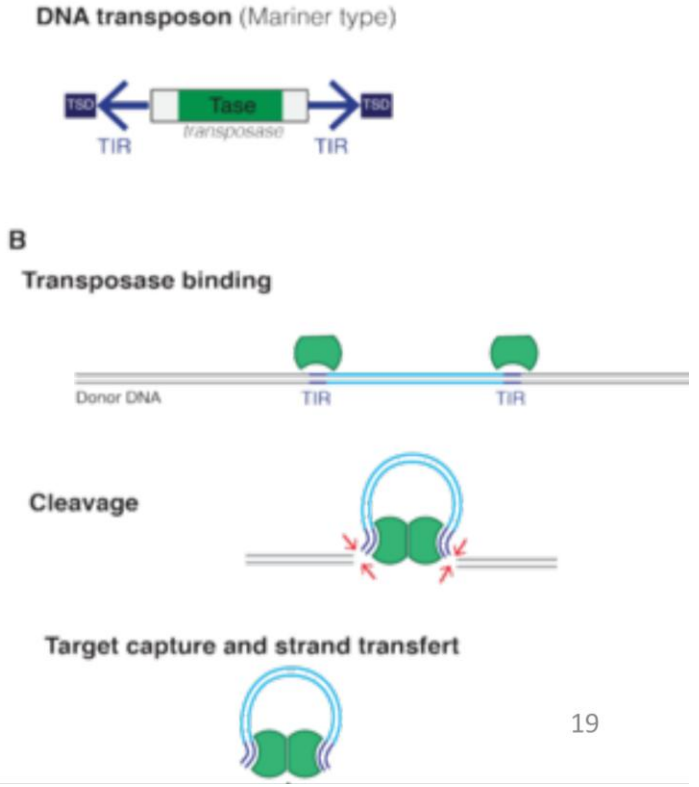
**D. TRANSPOSONS (Transposable Elements; "Jumping Genes"):**
**Transposons** are composed of interspersed repetitive noncoding DNA that make up an incredible 45% of the human nuclear genome.
Transposons are mobile DNA sequences that jump from one place in the genome to another (called transposition).

# Types of Transposons:

1. Short interspersed nuclear elements (SINEs)
2. Long interspersed nuclear elements (LINEs)
3. Long terminal repeat (LTR transposons)
4. DNA transposons — These are classified into two families: MER-I (Charlie) and MER-2 (Tigger).

**Roles:**
- Promote mutation and genome rearrangement.
- Facilitate exon shuffling and new gene formation.
- Drive evolutionary diversity.



# Lecture Summary:

- Only a small fraction of DNA directly codes for proteins.
- RNA-coding genes and noncoding sequences have crucial regulatory functions.
- Repetitive DNA and transposons maintain structural and evolutionary flexibility.
- Understanding genome organization helps explain gene expression, regulation, and genetic disease mechanisms.

# Reference


- Alberts, B., Johnson, A., Lewis, J., Morgan, D., Raff, M., Roberts, K., & Walter, P. (2022). Molecular Biology of the Cell (7th ed.). W. W. Norton & Company.→ Comprehensive overview of genome structure, chromatin, and gene regulation.
- Watson, J. D., Baker, T. A., Bell, S. P., Gann, A., Levine, M., & Losick, R. (2014). Molecular Biology of the Gene (7th ed.). Pearson Education.→ Covers chromosomes, transcription, translation, and genetic code fundamentals.
- Berg, J. M., Tymoczko, J. L., Gatto, G. J., & Stryer, L. (2019). Biochemistry (9th ed.). W. H. Freeman and Company.→ Useful for linking molecular structure (nucleic acids, proteins) with gene function.
- Lodish, H., Berk, A., Kaiser, C. A., Krieger, M., Bretscher, A., Amon, A., … & Martin, K. C. (2021). Molecular Cell Biology (9th ed.). W. H. Freeman.→ Excellent background on transcriptional control and RNA processing.