

iii. Watson-Crick model, DNA, RNAs, gene and genome, prokaryotes and eukaryotic genes, C-value paradox

iii. Watson-Crick model and forms of DNA; types of RNAs, Concept of gene and genome, difference between prokaryotes and eukaryotic genes, C-value paradox, Triplexes, quadruplexes and aptamers

Watson-Crick Model and Forms of DNA

The Watson-Crick Model (1953)

Historical Context

- Proposed by James Watson and Francis Crick, building on data from Rosalind Franklin's X-ray diffraction and Erwin Chargaff's base composition rules.

Key Features

- Double Helix:** Two polynucleotide strands coiled around a common axis.
- Antiparallel Strands:** One strand runs 5' → 3', the other 3' → 5'.
- Base Pairing:** A pairs with T (via 2 hydrogen bonds), and G pairs with C (via 3 hydrogen bonds).
- Major and Minor Grooves:** Regular indentations formed by the helix geometry, important for protein-DNA interactions (e.g., transcription factors).
- Stability Factors:** Hydrogen bonding between bases, base stacking (hydrophobic interactions), and the sugar-phosphate backbone's covalent bonds.

Forms (Conformations) of DNA

B-DNA (Watson-Crick Form)

- The most common form under physiological conditions (aqueous, moderate salt).
- Right-handed helix with ~10.5 base pairs per turn.
- Wide major groove, narrower minor groove.

A-DNA

- Formed under low hydration or high salt conditions.
- Right-handed helix with ~11 base pairs per turn, more compact and broader than B-DNA.
- DNA-RNA hybrids and double-stranded RNA often adopt an A-like conformation.

Z-DNA

- Left-handed helix with a zigzag phosphodiester backbone.
- Formed by alternating purine-pyrimidine sequences (e.g., GCGCGC) under high salt or specific protein-binding conditions.
- May play roles in transcription regulation and genome stability.

Others

- Minor variations include **C-DNA**, **D-DNA**, and more, but these are rarely observed in vivo.

Types of RNAs

While DNA serves primarily as a stable information repository, **RNA** participates in diverse cellular processes. Major classes of RNAs include:

- Messenger RNA (mRNA)**
 - Carries the genetic code from DNA to ribosomes for protein synthesis.

- In **prokaryotes**, often polycistronic (one mRNA can encode multiple proteins).
- In **eukaryotes**, generally monocistronic (each mRNA encodes a single protein).

2. Ribosomal RNA (rRNA)

- Major structural and catalytic components of ribosomes.
- In eukaryotes, 28S, 18S, 5.8S, and 5S rRNAs assemble with ribosomal proteins.
- rRNA ensures correct tRNA positioning and catalyzes peptide bond formation.

3. Transfer RNA (tRNA)

- Adapter molecules that bring amino acids to the ribosome during translation.
- Each tRNA recognizes specific codons on mRNA via its anticodon loop and carries the corresponding amino acid.

4. Small Nuclear RNA (snRNA)

- Primarily involved in **splicing** (removal of introns from pre-mRNA) within the spliceosome complex in eukaryotic cells.

5. MicroRNA (miRNA) and Small Interfering RNA (siRNA)

- Short (~20-24 nt) regulatory RNAs that modulate gene expression by binding to complementary mRNA sequences, leading to translational repression or mRNA degradation.

6. Long Noncoding RNA (lncRNA)

- Longer than 200 nt, often function in transcriptional and epigenetic regulation, serving as scaffolds for chromatin modifiers or mediators of gene expression.

7. Other Specialized RNAs

- **Piwi-interacting RNAs (piRNAs)**: Maintain genomic integrity in germline cells.
- **CRISPR RNAs (crRNAs)** in prokaryotes: Part of adaptive immune defense against phages.

Concept of Gene and Genome

Defining a Gene

Classical Definition

- A gene was historically considered a unit of heredity encoding a single polypeptide or trait.

Modern Perspective

- A gene is a **functional DNA (or RNA) region** that can produce a functional product (protein or functional RNA).
- Includes exons (coding regions), introns (noncoding) in eukaryotes, and regulatory sequences (promoters, enhancers, UTRs).

Genome

- The **genome** is the entire set of genetic material within an organism.
- In **prokaryotes**, typically a single circular chromosome (plus possible plasmids).
- In **eukaryotes**, multiple linear chromosomes (nuclear genome) plus organellar genomes (mitochondria, chloroplasts).

Differences Between Prokaryotic and Eukaryotic Genes

Feature	Prokaryotic Genes	Eukaryotic Genes
Genome Organization	Often a single circular chromosome; may have plasmids	Multiple linear chromosomes; organelle genomes
Operons	Common (e.g., lac operon); one promoter for multiple genes	Rare or absent; typically 1 gene per promoter
mRNA Structure	Polycistronic mRNA common	Usually monocistronic
Introns	Typically absent or extremely rare	Introns common; removed by splicing
Regulatory Elements	Simple promoters, operators	Complex promoters, enhancers, silencers

Chromatin	No histones (except in some Archaea), no true nucleosomes	Histone-bound DNA forming nucleosomes, chromatin
Post-Transcriptional Modifications	Minimal (mRNA rapidly translated or degraded)	Capping, polyadenylation, splicing, extensive regulation

C-Value Paradox

Definition

- The **C-value** is the amount of DNA (haploid genome size) in a species.
- The **paradox**: Genome size does not correlate strictly with organismal complexity. Some simple organisms have enormous genomes, while more complex organisms can have relatively smaller genomes.

Explanations for the Paradox

1. **Noncoding DNA**: Large quantities of repetitive elements, transposons, and other noncoding sequences can inflate genome size.
2. **Gene Duplication and Pseudogenes**: Gene families, partial or nonfunctional duplicates can expand genome content without necessarily increasing complexity.
3. **Polypliody**: Whole-genome duplications can lead to large increases in DNA content, common in plants.
4. **Regulatory Complexity**: Complexity may be governed more by how genes are regulated than by the raw number of genes or overall genome size.

Non-Canonical DNA Structures: Triplexes, Quadruplexes

DNA Triplexes (H-DNA)

- **Formation**: Occur when a single DNA strand binds to the major groove of a double-stranded DNA helix, often in polypurine-polypyrimidine sequences.
- **Mechanism**: A third strand can form Hoogsteen or reverse Hoogsteen hydrogen bonds with base pairs in the duplex.
- **Biological Relevance**: Implicated in transcription regulation, recombination, and can be targets for therapeutic interventions (e.g., triplex-forming oligonucleotides to modulate gene expression).

DNA Quadruplexes (G-Quadruplexes)

- **Structure**: Stacked units of G-tetrads (planar arrangements of four guanines held by Hoogsteen bonds).
- **Common Sites**: Found in G-rich regions such as telomeres and certain gene promoters (e.g., *c-MYC*).
- **Functional Importance**:
 - **Telomere Maintenance**: May regulate telomerase activity, implicating them in aging and cancer.
 - **Gene Regulation**: Quadruplexes near promoters can alter transcription efficiency.
 - **Therapeutic Targets**: Small molecules stabilizing G-quadruplexes are being explored for anticancer therapeutics.

Aptamers

Definition

- **Aptamers** are short, single-stranded DNA or RNA molecules that fold into unique 3D shapes to bind specific target molecules (proteins, small molecules, or even whole cells) with high affinity and specificity.

Selection Process (SELEX)

- **Systematic Evolution of Ligands by Exponential Enrichment (SELEX)**: Iterative process where large random oligonucleotide libraries undergo multiple rounds of binding to a target, partitioning bound vs. unbound molecules, and amplifying the bound population.

Applications

- **Biosensors:** Aptamers can detect specific analytes (e.g., pathogens, toxins) in diagnostic assays.
- **Therapeutics:** Some aptamers (e.g., pegaptanib for macular degeneration) are used clinically to inhibit disease-related targets.
- **Drug Delivery:** Aptamers can serve as targeting ligands to guide nanoparticles or therapeutic agents to specific cells or tissues.

Concluding Remarks

From the **Watson-Crick double helix** to the multifaceted **types of RNA**, our understanding of genetic material has expanded to encompass not only classical structures and functions but also intricate **noncanonical DNA conformations** (triplexes, quadruplexes) and functional nucleic acid molecules like **aptamers**. Concurrently, the **concept of a gene** has evolved to include introns, regulatory sequences, and the realization that vast stretches of noncoding DNA can influence phenotype and evolution, exemplified in the **C-value paradox**. Furthermore, comparing **prokaryotic vs. eukaryotic genes** underscores how genome complexity is managed by vastly different structural and regulatory architectures.

Collectively, these discoveries illustrate the richness of genomic architecture and regulation, as well as the continued expansion of molecular biology's frontiers—paving the way for innovative applications in biotechnology, medicine, and synthetic biology.