Organisation of Prokaryotic and Eukaryotic Genome

Eukaryotes vs Prokaryotes
Differences

Feature	Eukaryotic Genome	Prokaryotic Genome
Size	Large (10^7-10^{11} base pairs)	Small (10^5-10^7 base pairs)
Appearance of genetic material	Multiple linear molecules	Single circular molecule
Molecule making up genetic material	Double helix DNA	Double helix DNA
Location of genetic material	Nucleus	Nucleoid region (not membrane bound)
Origin of replication	Multiple	Single
Non-coding regions	Common (about 98%) Introns present Enhancers/silencers present Many tandem repeat sequences (centromeres and telomeres)	Uncommon (Less than 15%) Introns not present Enhancers/silencers rarely present Few tandem repeat sequences (centromeres and telomeres)
Gene organisation	Monocistronic	Organised as operons
Extrachromosomal DNA	Present in the form of plasmids	No plasmids but mitochondria and chloroplast have their own DNA
Number of genes	Many (~25,000)	Fewer (~4500)
Association of genetic material with proteins	Large amounts with histone proteins and scaffold proteins	Relatively less with histone-like proteins

Level of DNA packing	High where DNA is associated with equivalent mass of histone and there are different levels of packing present DNA associated with octamers of 8 histone proteins forming nucleosomes → Coils around itself to form chromatin fibre/solenoid → Looped domains formed with scaffold proteins → Supercoiling into condensed chromosome	Relatively low where DNA is associated with small amounts of protein and there is only some looping present Looped domains formed with histone-like proteins → Supercoiling → Condensed DNA

Why only about 1% of genetic information is transcribed into functional RNA sequence in most eukaryotic cells?

- Genome is of large size, comprising both coding and non-coding regions
- Most of the DNA sequences represent non-coding DNA sequences which are not transcribed
    - Regulatory sequences e.g. promoter, enhancer, silencer
    - Repeated sequences e.g. telomeres, centromeres
- Many intron sequences in genes which are transcribed and removed during mRNA splicing
- In specialised cells, only a small number of genes that produce proteins required for the cells are transcribed
- Regulation of gene expression e.g. DNA methylation, histone deacetylation, chromatin remodelling complex

Why only 1-5% of genes expressed in a differentiated cell?

- Other 95% of proteins coded for are not needed in the particular cell type ⇒ Only genes that produce proteins that are required are transcribed
- Regulation of gene expression e.g. DNA methylation, histone deacetylation, chromatin remodelling complex

Why is eukaryotic genome coiled around histone proteins?

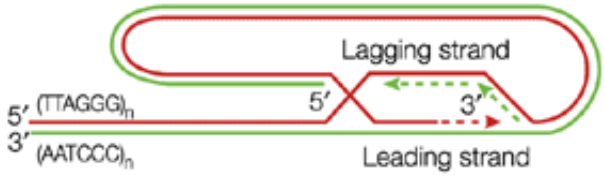
- Make long DNA molecule more compact to fit in nucleus
- Prevent entanglement, hence preventing DNA breakage/damage
- Regulation of gene expression as DNA wound tightly around histones prevents transcription factors and RNA polymerase from accessing promoter regions of genes

that do not need to be expressed

Non-coding DNA

- Introns (Eukaryotes only)
  - Structure:
    - Non-coding DNA sequences found within the gene between exons
  - Function:
    - Enables **alternative RNA splicing** of a single pre-mRNA
      - One pre-mRNA able to produce many different mature mRNAs depending on the combination of exons spliced together
      - One gene can code for more than one type of polypeptide ⇒ Larger number of proteins produced relative to the number of genes
- Promoters (Proximal control element)
  - Structure:
    - Non-coding DNA sequence located upstream of the transcription start site
    - Contains critical elements
      - e.g. TATA, CAAT and GC box in eukaryotes
      - e.g. -10 and -35 sequences in prokaryotes
  - Function:
    - Recognition site for the binding of transcription factors and RNA polymerase to initiate transcription
    - The more the **critical elements** in a given promoter resemble the **consensus sequences**, the greater the binding efficiency between the RNA polymerase and the promoter and hence the frequency of transcription
- Enhancers and Silencers (Distal control element; Eukaryotes only)
  - Structure:
    - Non-coding DNA sequences usually located far away from the promoter (upstream/downstream)
  - Function:
    - Regulate transcription
    - Enhancers
      - Increase efficiency of transcription by promoting assembly of transcription initiation complex when bound with **activators** by:
        - Bending of spacer DNA to allow direct interaction of activators with RNA polymerase/general transcription factors at promoter
        - Recruiting chromatin remodelling complex/histone acetylase to increase accessibility of promoter DNA to RNA polymerase/general transcription factors
    - Silencers
      - Reduce efficiency of transcription by inhibiting assembly of transcription initiation complex when bound with **repressors** by:
        - Interfering with action of activator
          - Competitive DNA binding - Enhancer region overlaps

with silencer region
        - Masking activation surface - Repressor binds to activator
        - Direct interaction with general transcription factors
    - Changing chromatin structure by recruiting histone deacetylase/chromatin remodelling complex
- Telomeres (Eukaryotes only)
    - Structure
        - **Non-coding** DNA made up of a series of **tandem repeat sequences** found at both ends of linear eukaryotic chromosomes
        - Single stranded region known as 3' overhang
            - Loops back and displaces the same sequence in the upstream region by binding to complementary sequence of the other strand



        -
    - Function:
        - Ensures that vital genetic information is not lost with each round of DNA replication due to end replication problem
            - Each round of DNA replication will result in the shortening of daughter molecules at the telomeres because DNA polymerase is unable to replace the last RNA primer with DNA nucleotides
        - Protects and stabilises terminal ends of chromosomes by forming a loop with the 3' overhang
            - Prevent fusion of the ends with other chromosomes
            - Prevents DNA repair machinery from recognising the ends of chromosomes as DNA breaks, hence preventing apoptosis
        - 3' overhang of telomeres allow their own extension by providing an attachment point for the correct positioning of **telomerase** in certain cells
- Centromere (Eukaryotes only)
    - Structure:
        - Non-coding tandem repeat DNA sequences
        - Found at constricted regions on condensed chromosomes
    - Function:
        - Allow **sister chromatids** to adhere to each other
        - Allow **kinetochores** and subsequently **spindle fibres** to attach to centromere so that sister chromatids can be separated to opposite poles when spindle fibres shorten during mitosis
        - Centromeres divide during anaphase of mitosis

Telomerase

1. Telomerase has an **active site** that is complementary in conformation and charge to

the specific telomere DNA sequence

2. **Telomerase RNA** anneals to single-stranded 3' overhang on telomere by **complementary base pairing**, aligning the telomerase reverse transcriptase to the DNA
3. Telomerase reverse transcriptase uses telomerase RNA as a **template** to form complementary DNA sequence where free incoming DNA nucleotides complementary base pair with telomerase RNA
4. Telomerase reverse transcriptase catalyses the formation of phosphodiester bonds between adjacent DNA nucleotides, elongating the 3' overhang
5. Telomerase translocates 6 nucleotides in the 5' to 3' direction to synthesise another series of tandem repeats
6. Primase synthesises RNA primer close to the end of the 3' overhang of the telomere
7. DNA polymerase adds DNA nucleotides to the 3' OH end of RNA primer and synthesises a complementary DNA strand
8. Nick is sealed by DNA ligase