DNA, Genes and their Regulation

Mette Voldby Larsen
PhD, Associate Professor
Learning Objectives
After this lecture, you should be able to

• Account for the structure of DNA and RNA including their similarities and differences, and how complementary base pairing can take place between two DNA strands or between a DNA and an RNA strand. Describe how DNA is packed into chromosomes.

• Describe the Central Dogma of Molecular Biology, in particular transcription. Deduce the sequence of a pre-mRNA molecule that has been transcribed from a given gene.

• Sketch the structure of a eukaryotic gene and explain the difference between exons and introns.

• Describe how gene expression is regulated in eukaryotes emphasizing the many different levels this can be achieved on.
DNA
Deoxyribose Nucleic Acid

- Deoxyribose, phosphate, base (Adenine, Thymine, Guanine, Cytosine)
- Double helix: Via hydrogen bonds, A pairs with T and C with G.
- The two strands are antiparallel.
RNA
RiboNucleic Acid

• **RNA** differs from DNA in three ways:

  – RNA is usually single-stranded

  – The sugar is *ribose*, not deoxyribose

  – RNA contains the base *uracil* (U) instead of *thymin* (T)

• **RNA** can basepair with single-stranded DNA (adenin pairs with uracil instead of with thymin)

• An **RNA-strand** can fold and basepair with itself (creating a secondary structure)
Genes are sequences of DNA

>gi|6007800:215-1075 Escherichia coli beta-lactamase variant TEM-1D (blaTEM-1D) gene, complete cds

5' ATGAGTATTCAACATTTCCGTGTCCCTATTTCCCTTTTTTCGCGGCATTTTGCCTTTTGGCTC
ACCCAGAAACGCCTGGTGAAGGATGAAAGATGCTGAAGATCAGTTGGGTGCACTGAGGTTACATCGAGCT
GGATCTCAACCAGCGTGAAGATCTCCCTTGGAGAGTTTTCGCCCGAGAAGAAGCTTTTCTCAATGATGAGCAACTTTT
AAAGTTCTGCTATGTGTTGTCGGTTATTATCCCTACGCCGGCAGCGAAGAGCACTCGTGCCGCATAC
ACTATTCTCAGAATGACTTTGTTGAGTACTCACCAGTCACAGAAAAAGCATCTTACGGATGCGATGACAG
AAGAGAAATTATGCGAGTCTGCCATAACCATGAGTGATAAACACTGCGGCAACTTTACTTCTGACAACGATC
GGAGGACCAGGACGAAGCATAACCGCTTTTTGCAACAACATGGGGGATCATGTATACCCGTTGGG
AACCAGAGCTGAAATGACGCAACATACCAAAACGACGACGCGTGAGCACCAGATGCCTGCAATGGCAACAAACAAC
GTTGCCAACCATTATTAACCTGGCCGAACTACTTACTCTAGCTTCCGCGGAAACAAATTAATAAGACTGGATGGAG
GCGGATATAAGTTGACGAGACCACCTTTGCGCTCGCCGCTTTCCCGGCTGGCTGTGATGTTATGTCAATAATCTG
GACCGGCTGAGCTCGGTGATCTCCTCCGCGATCATTGCAGCAGACTGGGCTGCGGATGTAAGCCCTCCCGTATCGT
AGTTATCTACACGACGGGAGTACGGCAACTATGGATGAACGAAATAGACAGATCGGCTGAGATAGGGTGG
TCAGTGATTACGACTTTGGTAA 3'
DNA is packed into chromosomes

1. A DNA molecule binds with histones, forming a vast number of nucleosomes.
2. Nucleosomes form "beads" on DNA "string."
3. Nucleosomes pack into a coil that twists into another larger coil, and so forth, producing condensed, supercoiled chromatin fibers.
4. The coils fold to form loops.
5. The loops coil even further, forming a chromosome.
Learning objective
Be able to account for the structure of DNA and RNA including their similarities and differences and how complementary base pairing can take place between two DNA strands or between a DNA and an RNA strand. Describe how DNA is packed into chromosomes.

Typical exam question

Q: Below is shown a single stranded DNA molecule. Which of the strands shown in a. - e. is the complementary DNA strand?

Single stranded DNA molecule: 5’ ATGCCCGGG 3’

a. 3’ ATGCCCGGG 5’

b. 5’ CCCGGGCGA 3’

c. 3’ CCCGGGCGA 5’

d. 3’ UACGCCGCG 5’

e. 5’ TACGCCGCG 3’
The Central Dogma of Molecular Biology

- Genetic information is transferred from DNA to RNA to protein. (Almost) never in the opposite direction (Francis Crick).
>gi|6007800:215-1075 Escherichia coli beta-lactamase variant TEM-1D (blaTEM-1D) gene, complete cds
ATGAGTATTCAACATTTCCGTGTGCCCTTATTTCCCCCCCCCGATTTTGGCTCCCTCTCTGCTC
ACCCAGAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTACATCGAGCT
GGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTT
AAAGTTCTGCTATGTGGTGCGGTATTATCCCGTGTTGACGCCGGCAAGAACGCAACTCGGTCGCCGCATAC
ACTATTTCAGAATAGCTTGTTGAGTACTACCACTACAGGTGAAAGCAACATCTTACGGGATAGCATGACGT
AAGAGAATTATATGCAGTGCTGCCATAAACATGAGTAACACTGCGGCCAACTTACTTCTGGAACAGCATC
GGAGGACCGAAGAGACCTAACCCTTTTTGGACACAACTGGGGATACGTGAACCCGCCCTTGATCTGTGGG
AACCAGGAGTCTGAATAGCCCATACCCAAGAGAGCGCTGACACACAGATGCTGCAATGGGAACAAAC
GTTCGCAAAACTTAATTAAGGGGACTTACTTACTCTAGGCTTTCCGCCGCAACAATTAAGACTGGATGAG
GGCGGATAAAAGTTGCAAGAGCCACTTCTCGCTCGCCCTCTCCGGCTGCTGTGCTGTTTATTTGCTGAATAATCTG
GAGGCCTTGAGCGCTGATCTGCCGCTATCATGCGAAGCATTGGAAGTGTGAAACGCCTCTGCTGATCGT
AGTTATCTACACGCAGGGGAGTCAGGCAACTATGGGATGAAACGAAATAGCACGATCGCTGAGATAGGTG
CTCAGTATTAAGCATTGGTAA

>blaTEM-1D mRNA
AUGAGUAUUCAACAUUUCGGUUCGCCCUAUUUCUUUUUCUGCGCAUUUUGCUCUUUCUGUGCCAAACGGGUAACG
ACCCAGAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTACATCGAGCT
GGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTT
AAAGTTCTGCTATGTGGTGCGGTATTATCCCGTGTTGACGCCGGCAAGAACGCAACTCGGTCGCCGCATAC
ACTATTTCAGAATAGCTTGTTGAGTACTACCACTACAGGTGAAAGCAACATCTTACGGGATAGCATGACGT
AAGAGAATTATATGCAGTGCTGCCATAAACATGAGTAACACTGCGGCCAACTTACTTCTGGAACAGCATC
GGAGGACCGAAGAGACCTAACCCTTTTTGGACACAACTGGGGATACGTGAACCCGCCCTTGATCTGTGGG
AACCAGGAGTCTGAATAGCCCATACCCAAGAGAGCGCTGACACACAGATGCTGCAATGGGAACAAAC
GTTCGCAAAACTTAATTAAGGGGACTTACTTACTCTAGGCTTTCCGCCGCAACAATTAAGACTGGATGAG
GGCGGATAAAAGTTGCAAGAGCCACTTCTCGCTCGCCCTCTCCGGCTGCTGTGCTGTTTATTTGCTGAATAATCTG
GAGGCCTTGAGCGCTGATCTGCCGCTATCATGCGAAGCATTGGAAGTGTGAAACGCCTCTGCTGATCGT
AGTTATCTACACGCAGGGGAGTCAGGCAACTATGGGATGAAACGAAATAGCACGATCGCTGAGATAGGTG
CTCAGTATTAAGCATTGGTAA
Learning objective
Be able to describe the Central Dogma of Molecular Biology, in particular transcription. Deduce the sequence of a pre-mRNA molecule that have been transcribed from a given gene.

Typical exam questions
Q: The “central dogma” of molecular biology states that

a. information flow between DNA, RNA, and protein is reversible.
b. information flow in the cell is unidirectional, from protein to DNA.
c. information flow in the cell is unidirectional, from DNA to protein.
d. the DNA sequence of a gene can be predicted if we know the amino acid sequence of the protein it encodes.

Q: Below, the coding strand of a DNA molecule is shown. The arrow indicates where, and in which direction transcription occurs. Write the resulting mRNA molecule

\[ 5' \text{GGTCTATATAAGCAGAGCTTTATGAACCGTCAGATGAG} 3' \]

a. \[ 5' \text{GUUUAUGAACCAGUCAGAUGAG} 3' \]

b. \[ 3' \text{GUUUAUGAACCAGUCAGAUGAG} 5' \]

c. \[ 3' \text{CAAATACTTTGGCAGTCTACTC} 5' \]
BREAK
Structure of eukaryotic protein-coding genes

Exons: Code for amino acids

Introns: Code for nothing

Start codon

Promotor
Here the RNA polymerase binds

Stop codon

Terminator
Signals to the RNA polymerase to stop transcribing
Structure of eukaryotic protein-coding genes
Alternative splicing

Pre-mRNA

Mature mRNA

Exon1  Intron1  Exon2  Intron2  Exon3
**Learning objective**
Be able to sketch the structure of a eukaryotic gene and explain the difference between exons and introns

**Typical exam question**

The *RB* gene contains 27 exons and 26 introns. Below, exon no. 17-19 and intron no. 17-19 are shown schematically. Sketch the mature mRNA molecule with regards to these exons/introns (you may assume that there is no alternative splicing).

*Answer:*

*Only exons are present in the mature mRNA. Introns are spliced out*
Regulating gene expression in multicellular organisms

- Same genome – different expression at different stages
- Same genome – different expression in different places
**Chromatin** (DNA+histones)

Euchromatin: Lightly packed chromatin. The genes are transcribed.

Heterochromatin: Tightly packed chromatin. The genes are normally not transcribed.

[vsrp.uhnres.utoronto.ca/research1.htm](vsrp.uhnres.utoronto.ca/research1.htm)
Barr body – One of the X chromosomes of women are inactivated as heterochromatin

During the development of a female embryo, one of the two X chromosomes are inactivated.

The female body is a mosaic, where some areas contain cells that have the one X chromosome inactivated, while cells in other areas have the other X chromosome inactivated.
Remodelling chromatin structure before transcription

Gene expression can be regulated here
Initiating transcription

Gene expression can be regulated here
The same transcription factor can bind/regulate several genes!

Gene expression can be regulated here
MicroRNA – regulation mRNA longevity

• Small RNA molecules, ≈ 22 to 23 nucleotides

• Bind via basepairing to the 3’ end of the mRNA

• Binding inhibits translation and sometimes the mRNA is even degraded

• Computers predict that more than 1000 genes encode microRNA

• It is also predicted that microRNA molecules regulate the expression of more than 1/3 of all human genes
Post translational regulation of expression

Ubiquitination - or

Gene expression can be regulated here
Learning objective
Describe how gene expression is regulated in eukaryotes emphasizing the many different levels this can be achieved on.

Typical exam question

Q. A DNA sequence is several thousands of nucleotides distant from the promoter. When this sequence is bound by a protein, transcription rates increase greatly. This sequence is most likely a(n)
a. TATA box.
**b. enhancer.**
c. operon.
d. promoter.
e. consensus sequence.
DNA Microarray Technology

© Sinauer Associates, Inc.
Eksempel på undtagelse

• Visse RNA virus, kaldet retrovirus (herunder HIV), oversætter deres RNA-genom ”baglæns” til DNA efter at de har inficeret en celle
• Denne proces kaldes *revers transkription*, og foretages af enzymet *revers transkriptase*
• Retroviruses DNA bliver integreret i værtscellens genom og kan ligge der i lang tid, før det bliver brugt til at danne nye virale RNA genomer ved almindelig transkription

Forekommer også i forbindelse med *retrotransposons* (kap. 14)
**Initiering** begynder ved en **promotor**, en særlig sekvens på DNA

- I eukaryoter er der (mindst) én promotor for hvert gen
- I prokaryoter kan flere gener aflæses fra samme promotor
- RNA polymerasen binder til promotor området
- Promotor sekvensen dirigerer hvilken streng RNA polymerasen skal bruge som template og dermed i hvilken retning den kører
- DNA dobbelt helixen skal delvist vikles ud for at fungere som template – dette gøres af RNA polymerasen
Transkription: Elongering

- RNA polymerase vikler DNA’et ud ca. 10 basepar ad gangen og læser template i 3’ → 5’ retning
- Efterhånden som RNA transkriptet dannes, frigøres det fra baseparring med DNA template strengen, og DNA går tilbage til dobbelthelix-konformationen
- Det nye RNA vokser i 5’ → 3’ retning; dvs. RNA transkriptet er antiparallel til DNA template strengen
- RNA polymerasen kan ikke korrekturlæse: Transkriptionsfejl er relativt hyppige sammenlignet med DNA replikation