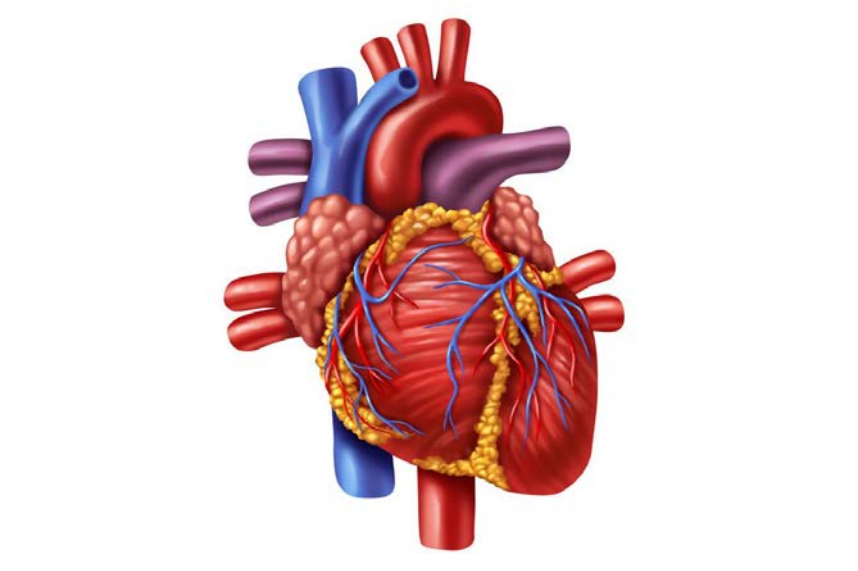
# HUMAN BIOLOGY (BIOL1003)



# Molecular Basis for Inheritance: Lecture 1

- Objectives:
  - Structure of DNA
  - How DNA is replicated
  - How info in DNA is transcribed into protein
  - Evidence that DNA encodes genetic info

- In 1940-50 (what was known):
  - **Genes**= hereditary 'factors' causing different traits, physical nature unknown
  - Famous experiments by Gregor **Mendel**, crossed peas with different appearances/traits, found that when you cross peas with different traits, get the traits of both parents in the resulting peas, even if the peas of the next gen look like just one of the parents, the info is there to encode info of other parent as well- evidence that there was info that parents gave offspring, but not clear what that information was made of or how it was encoded
  - Genes are carried on chromosomes (quite large, could be seen under microscope), without chromosomes there was no hereditary material, lead to idea that these were very important
  - Chromosomes consist of protein and DNA
  - **Proteins**= key components of organisms, 20 amino acids known, proteins= chains of amino acids
  - Other component of chromosomes= DNA, known to consist of 4 nucleotides, people at the time did not know what role it played or how it interacted with other molecules in organisms + structure not known at the time
- Linus **Pauling**, prominent molecular biologist in 1940-50, said:
  - DNA has only 4 nucleotides, therefore was too simple to encode proteins
  - Proteins contain 20 amino acids + make diverse and complex materials e.g. hair, skin
  - Protein likely to be a hereditary material

- <u>Experiments showing DNA was hereditary material:</u>
- **1928- Griffith's work on Streptococcus pneumoniae (bacteria)**→ this bacteria causes pneumonia in humans + some strains deadly for mice, S strain kills the mouse, R strain does not, found if S-strain heat killed-> mouse lives, surprising result was: when mouse was given R strain and dead S strain (both separately did not kill mouse in earlier tests, mouse died and found S strain live cells in dead mouse, looked like R strain was taking up some of the dead S strain information and using it to create live S strain cells that killed the mouse (today we know bacteria use transformation where they take in DNA from the environment)
- **1944- Avery et al. prove DNA is genetic info.** → did an experiment following Griffith's work, used same system, in this experiment gave the mouse an R strain and a dead S strain and each time they changed the material that was destroyed e.g. destroyed lipids, RNA, protein, polysaccharides then DNA, only when they destroyed DNA did the mouse survive as the R strains did not read anything that could allow them to produce more S strains, resulting info: became clear that it was DNA that coded for information
- ^proved genes were made of DNA

- Pauling published 'triple helix' model of DNA, problem with model= DNA ceased to be an acid (DNA= deoxyribonucleic acid/ deoxyribose nucleic acid), therefore model fundamentally flawed
- Watson and Crick used modelling, chem. + physics to try and understand DNA structure

- Info available to Watson and Crick:
    - Structure of nucleotide (units that DNA is made of): nitrogenous base= what makes DNA have 4 kinds (adenine, guanine, cytosine and thymine) + phosphate group + sugar (extra OH in RNA, extra H in DNA)
    - Every nucleotide has same type of sugar and same type of phosphate, different nitrogenous bases make different nucleotides
    - **Nucleotides**= carbon-based molecules, often name each of the carbon atoms, because each position can have several properties e.g. carbon-1 prime holds nitrogenous base and carbon-3 holds hydroxyl group, important in the case of DNA
    - **Chargaff's rules** of base compositions- amount of A is always the same as the amount of T and amount of C is always the same amount of G therefore, A+G = T+C, BUT A+T does not necessarily equal G+C
    - Rosaline Franklin did a key experiment (crucial but controversial) that helped determine model Watson and Crick made, looked at images of x-rays of DNA and say X shape typical of double helix structure

- 1953- Watson and Crick solve DNA structure= double helix
- **Double helix**= two strands/backbones that are twisted together, not directly facing each other, they are just off such that we get a minor groove and major groove in structure, has info about twists in DNA (differences between molecules for e.g.)
- Two strands are anti-parallel
- Backbone is made by carbon-3 on one side and carbon-5 of this sugar, then phosphate which is attached to carbon-5, attaches to the next nucleotide on carbon-3 (3, 5, phosphate, 3, 5, phosphate), on opposite strand this sequence is reversed
- Connecting the two strands= the nitrogenous bases, A's only connect with T's and G's only connect with C's (Chargaff's rules apply)
- Due to these connections, DNA can replicate (has particular sequence that can be copied), can hold information in the specific sequence of nucleotides (forming words or codons), it is also stable (could self-repair), BUT could also mutate occasionally, e.g. when new nucleotides are being added to the sequence- mechanism for how evolution might occur
- DNA replication= semi-conservative, we have original/template helix, when replicated new strands end up with old strands such that this 'unzips' and creates a new strand for each of the 2 original templates, another model= conservative, where original strands stay with each other

- DNA replication has leading and lagging strands, one strand of DNA is leading and other is lagging, when start 'unzipping' two strands are left loose, then start writing/adding new nucleotides, HOWEVER DNA can only be **written in a carbon-3-> carbon-5 direction** (this is the key property that defines leading and lagging strands), therefore can only add new nucleotides to the carbon-3 (one with hydroxyl group) cannot add new units to side where the phosphate is, the carbon-5, molecule therefore grows from carbon-3

- Therefore, in replication, one strand is being grown in direction of 'unzipping', but the other is growing in the opposite direction, because this strand is growing in the opposite direction, will run out of template, so will have to stop and begin again, causes gaps to form on one of the strands, while other in other strand nucleotides are being added continuously
- One growing continuously is called leading strand, gaps is called lagging strand, gaps will eventually be joined by enzymes
- How DNA is translated to protein: important piece of matter in this process= RNA
- RNA and DNA are similar but:
    - In RNA uracil replace thymine as a nitrogenous base, A, C, U, G, therefore wherever there normally would be a T there is a U instead
    - RNA exists as a single strand, not double helix, RNA is the one that makes complex large molecules
    - In RNA Sugar backbone is made of ribonucleotides (contains oxygen) not deoxyribonucleotides (does not contain oxygen)

- DNA (info. store genes + chain of sugars)-> transcription-> mRNA (mRNA= messenger RNA, intermediate carrier of info.)-> translation-> protein (product of info. + chain of amino acids), called central dogma of gene expression

- Overall: DNA is **transcribed** into RNA then **translated** into protein
- RNA can also make sequences such that they contain the same sequence and same info as DNA
- Need 3 nucleotides to code for one amino acid
- Genetic code works with 3 nucleotide 'words' (called codons) e.g. UUG, GUG, UGA (always 3 letters)
- There is some redundancy in the genetic code as some codons will code to the same amino acid e.g. UUU and UUC, also some do not code for any amino acids but instead code for stopping a process of replication
- AUG is a universal start codon
- Genetic code is almost universal across all phyla (taxonomic term for large group of organisms) suggests common origin of all life
- Some minor codon usage changes, evidence that there is a common origin of all life, same code= same source

- Inside cells= nucleus that contains DNA, in nucleus, RNA polymerase (enzyme) transcribes RNA from DNA= Eukaryotic Protein Synthesis
- Introns are parts of primary RNA transcript that are not used/not translated into protein, they are excised from transcription remaining exons (parts that do code for information) are spliced together, producing mRNA
- After introns are excised, mRNA is transported out of the nucleus, in the cytoplasm, ribosomal subunits bind to the mRNA, ribosome engulfs mRNA and it is in here that the process of translation occurs
- DNA never leaves the nucleus
- In ribosome, tRNA (contains anti-codon, complementary to codon, and contain corresponding amino acid to the codon) molecules become attached to specific amino acids with the help of activating enzymes, amino acids are brought to the ribosome in the order directed by the mRNA to create chain

- tRNAs bring their amino acids in at one side of the ribosome, peptide bonds form between amino acids at top and tRNAs exit the ribosome on the other side
- Finally, polypeptide chain grows until the protein is completed
- Different possible ways to read mRNA called reading frames, 3 possible reading frames
- Amino acid is named by the letter it starts with
- Open Reading Frame (ORF)= sequence that makes a sensible transcript, contains start codon, then usually a few hundred or more codons, then has a stop codon

What you need to know from this:

- DNA is genetic material (important to know evidence for why it is genetic material)
- Structure of DNA (e.g. properties/components of DNA, that backbone is 3->5, also why the 3->5 is important (leading and lagging strands))
- How DNA encodes info (codons-triplets of nucleotides, idea of transcription and translation, the central dogma, and how we go from DNA to complex molecules made of proteins/amino acids)
- Transcription and translation^