Molecular Biology and Genomics MBLG2071/2971 Study Notes

Composition of DNA

Functions of DNA:

- Phenotypic = controls phenotype
- Genotype = transmission of genetic information
- Evolutionary = variation through mutation

DNA composition:

- Biopolymer of nucleotides (base, sugar and phosphate) – double stranded
- Sugar:
    - 1st Carbon: N-glycosidic bond – joins base to sugar
    - 2nd Carbon: Deoxy sugar (OH missing)
    - 3rd Carbon: 3' OH – joins to 5' phosphate of sugar below
    - 4th Carbon: joins to first carbon
    - 5th Carbon: 5' Phosphate attached – joins sugar above phosphate bond
- Bases:
    - Purine = 2 rings:
        - Adenine – NH2/amine group (top)
        - Guanine – Oxygen/ketone (top) and HN2 group (side)
    - Pyramidine = 1 ring:
        - Thymine – Methyl group and two oxygens
        - Cytosine – NH2 group (top) and Oxygen (side)
        - Uracil – No methyl group
- Base pairs: hydrogen bonds
    - A+T – Amine group of A gives H to ketone group to T and ring N of T gives H to ring N of A
    - G+C – Amine groups give H to matching ketone groups and G ring N gives to C ring N

Features of Chromosomes

- Origin of replication: genomic DNA sequences that initiate replication
    - Bacteria only has 1
    - Eukaryotes have many
- Centromere
    - Where spindle attaches during chromosome separation
    - Defines the position of markers on a chromosome – as it is different for every chromosome
    - Divides chromosome in 2 arms
        - Short arm = p
        - Long arm = q
- Telomeres
    - Ends of chromosomes
    - Simple repeating structure
    - Allows ends to be replicated
    - Distinguishable from break in chromosome
- Stains:
    - Giemsa – forms G bands – each chromosome has different bands
    - Feulgen:
        - Dark = heterochromatin = dense = repeats
        - Weak = euchromatin = not condensed = actively transcribed genes

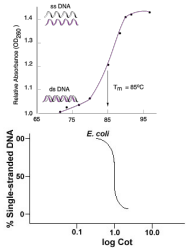
Compaction of DNA

- Prokaryotes – Supercoiled loops (folded genome)
- Eukaryotes – DNA winding around histone proteins (Chromatin)
    - Core histones H2A, H2B, H3 and H4
    - Each of these form a dimer and then they all join together to create core octamer
    - Nucleosome = histone core with core DNA wrapped around it
    - Nucleosomes joined by linker DNA and linker histone (H1) = tighter wrapping- pulls it together 6X
    - Forms 30nm fibre

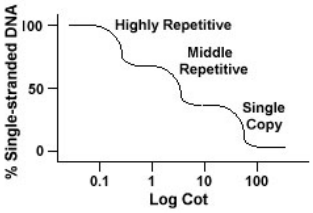
    - Loops of fibres held together by nuclear scaffold proteins
    - Many loops make up chromosome (DNA, histone proteins and non-histone proteins) 40X compaction
- Packaging and expression
    - DNA releases and rebinds to histone – allowing transcription factors and other regulatory proteins to gain access when needed
    - Chromatin remodelling (moving nucleosomes) – nucleosome remodelling complexes use ATP to move nucleosomes and allow access for proteins e.g. polymerases to specific sites – two types:
        - Sliding – slide nucleosome over
        - Displacement – remove nucleosome (histone core)
    - Chromatin modification (changing chemical groups on histone proteins)
        - Each histone proteins have a different N-terminal tail that contains amino acid residues that can be altered by adding or removing chemical groups
        - E.g. Histone acetyl transferases (HATs) add acetyl groups to lysine residues and histone deacetlyases remove them (HDACs)
        - This modification changes charges or recruits other proteins to change its interaction with DNA/packaging and hence alter expression

C-Value

- C-value = genome size of an organism (pictograms in a haploid set of DNA)
- Humans = 3.5, Shrimp = 15
- Paradox = disparity between C value and relative complexity of an organism (c value not related to complexity)
- Genome is made out of:
    - Genes and related sequences (introns etc.)
    - Intergenic DNA (genome wide repeats, regulatory regions, miRNA, microsatellites)
- Studying genome composition:
    - Melting curve: DNA melted and monitored using UV spectrometry at 260nm (shows melting temperature by vertical part of line – large jump in absorption from double stranded to single stranded)
    - Reassociation curve:
        - Simple DNA: DNA cut into manageable lengths, melted and reannealed – measured by UV absorbance decrease from single stranded to double stranded



        - Complex DNA: Triphasic (3 kinds of sequences) show 3 different steps on reannealing curve:
            - Highly repetitive sequences reanneal first due to more sequences to pair with (highest step):
                - Microsatellites (around centromere and telomeres)
                - Tandem arrays of short repeated units
                - Dinucleotide repeats
            - Moderately repetitive sequences reanneal next (Middle step):
                - Multiple copy protein-coding genes
                - Ribosomal and transfer RNAs
                - Genome wide repeats:
                    - LINEs (long) – Autonomous poly-A retrotransposons – contains all information needed to transpose by itself (move to other areas of the genome while leaving a copy behind) e.g. L1
                    - SINEs (short) – nonautonomous – need to hijack activities from LINEs to transfer successfully e.g. Alu



            - Non-repetitive sequences reanneal last (bottom step):
                - Single copy protein coding genes

G-Value

- Paradox = number of protein coding genes (gene number) relates to complexity of organism
- However, some organisms make more use of alternative splicing
- Also, Non-protein coding sequences = also transcribed and also important but not counted in G value:
    - Infrastructural RNAs:
        - Ribosomal and Transfer RNA affect translation
        - Small nuclear used in spliceosome and affect splicing
    - Other non-coding RNAs: Small regulatory RNAs: