DESIGNER GENES-MOLECULAR GENETICS

CENTRAL DOGMA OF MOLECULAR GENETICS



- Central dogma of molecular genetics is DNA \rightarrow RNA \rightarrow Protein.
- Exceptions among viruses RNA to DNA (retroviruses) Exception is in retroviruses where genetic storage vehicle is RNA. It then makes a DNA which replicates to form double stranded DNA and continues through dogma.

DNA Structure

- DNA structure ó double helix with sugar (deoxyribose), phosphate and nitrogen bases (Adenine, Thymine, Guanine, and Cytosine).
- Pairing is A with T and G with C



Nucleotide - basic unit of sugar, phosphate and nitrogen base - 4 kinds of nucleotides because of the 4 types of bases

Similarities between Prokaryotic and Eukaryotic DNA Replication

- Both prokaryotic and eukaryotic DNA replications occur before entering the nuclear division.
- Both prokaryotic and eukaryotic DNA replication works upon double-stranded DNA.
- The unwinding of both prokaryotic and eukaryotic DNA is done by DNA helicase.
- The unwound DNA strands are stabilized by single-stranded DNA-binding proteins (SSB).
- Both prokaryotic and eukaryotic DNA replication are multistep processes, which are carried out by an enzyme complex called DNA polymerases.
- Each type of DNA polymerases works in the 5øto 3ødirection.
- RNA primers are required for the initiation of both types of DNA replications.
- The synthesis of the RNA primer is done by the enzyme called primase.
- Both prokaryotic and eukaryotic DNA replications occur in a semi-conservative manner where one old strand of DNA and one new strand of DNA can be found in the daughter cell.
- Both prokaryotic and eukaryotic DNA replications are bi-directional since the replications progress in both ways.
- Leading and lagging strands are formed in both types of DNA replications.
- The lagging strand produces the small DNA fragments called Okazaki fragments, which are eventually joined together.
- The time taken for both types of replications are around one hour.

Prokaryotic DNA

Features of Plasmids

- Plasmids can be readily isolated from bacterial cells.
- They are self-replicative inside cells.
- They are composed of unique restriction sites for one or more restriction enzymes.
- The insertion of a foreign DNA fragment may not alter the replication properties of plasmids.
- Plasmids can be sequentially transformed into different types of cells and the transformants can be selected based on the antibiotic resistance properties of the transformed plasmids
- Plasmids can be used as vectors that carry foreign DNA molecules into both eukaryotic and prokaryotic cells.



Eukaryotic DNA



PROKARYOTIC vs.	EUKARYOTIC
Structure-Closed Circular double stranded DNA	Structure-Linear double-stranded DNA with 2 ends
In the central portion of the cytoplasm	Inside the nucleus (some in mitochondria and some
	in chloroplasts)
Do not form the typical chromosome	Form chromosome or chromatin in the nucleus
Do not interact with histone proteins	Associated with histone proteins
Forms loop-like structures by wrapping around	Form distinct structural repeats called nucleosomes
histone-like protein molecules but no nucleosome	and shows higher order packaging
formation	
Quantity of DNA is small	Quantity of DNA is 50 times more than Prokaryotic
Only very few proteins interact with prokaryotic	Large numbers of proteins interact with eukaryotic
DNA	DNA
Usually codes for 300 to 500 proteins	Codes for thousands of proteins
Majority of DNA is coding, only very little non-	Majority of DNA is non-coding. Size of coding
coding region occurs in prokaryotic DNA	region is less than the non-coding regions.
Replication takes place in cytoplasm	Replication takes place in nucleus
Continuous process	Occurs during S phase of the cell cycle
Single origin of replication	Multiple origins of replication (over 1000)
Carried out by DNA polymerase I and II	Carried out by DNA polymerase , , and .
The Okazaki fragments are comparatively large, 1000-	The Okazaki fragments are small, around 100-200
2000 nucleotides in length.	nucleotides in length
DNA gyrase is required	DNA gyrase is not required
Rapid process ó around 2000 nucleotides are added	Slow process ó around 100 nucleotides added per
per second	second
Final product is 2 circular chromosomes	Final product is 2 sister chromatids
Introns absent in the coding region of DNA	Introns occur in the coding region of DNA
Transcription produces-mRNA unit is polycistronic	Transcription produces- mRNA unit is
ó codes two or more proteins	monocistronic ó codes only one protein

DNA Replication

5° end Complementary 3' end Template strand strand DNA replication is semi-conservative. Hydrogen bond OH (et Phosphate Sugar Enzyme: Base DNA polymerase III "Growing" end: end Nucleotides 5' end added Deoxy-CTP OH 3' end **Events that occur:** 5' end

- **DNA polymerase** is the key enzyme
- DNA uncoils and splits
- template strand is read 3øto 5ø
- new complementary strand must add new nucleotides to the 3øend ó leading strand (continuous) while lagging strand is fragments (Okazaki fragments) latter attached with the enzyme ligase



Prokaryotic DNA Replication

- Prokaryotic DNA replication is the process by which prokaryotes such as bacteria and <u>archaea</u> duplicate their genome into a second copy, which can be transformed into a daughter cell.
- Prokaryotes consist of a double-stranded circular DNA molecule in their cytoplasm.
- Prokaryotic DN<u>A</u> comprises a single origin of replication. DNA helicase unwinds the DNA at the origin of replication by breaking the hydrogen bonds between the nitrogenous bases. The resultant Y-shaped structure is called the **replication fork**.
- Since prokaryotic DNA contains a single origin of replication, only two replication forks are formed during the replication process. These two replication forks process bi-directionally.
- The single-strand DNA-binding proteins (SSB) stabilizes the two unwound strands, which serve as the template strands for the replication.
- The enzyme, RNA primase synthesizes a five to ten base pairs long RNA primer, which is complementary to the template strand.



- Three types of DNA polymerases are involved in the prokaryotic DNA replication; DNA polymerase I, II, and III.
- Both initiation and elongation of the prokaryotic DNA replication are carried out by DNA polymerase III.
- The DNA polymerase III adds nucleotides in 5øto 3ødirection.
- Due to the antiparallel nature of the DNA double-helix, one strand runs from 5øto 3ødirection (leading strand). The other strand runs from 3øto 5ødirection (lagging strand).
- Since, the lagging strand requires RNA primers continuously in order to synthesize DNA in the 5øto 3ø direction, new fragments of DNA called Okazaki fragments are continuously formed.
- The gap filling and DNA repair are carried out by DNA polymerase I and II.
- The RNA primer is removed by the DNA polymerase I.

ENZYMES INVOLVED IN REPLICATION – Prokaryotic

The replication fork is the unwound helix, with each strand being synthesized into a new double helix

- *Topoisomerase* is responsible for initiation of the unwinding of the DNA.
- *Helicase* accomplishes unwinding of the original double strand, once supercoiling has been eliminated by the topoisomerase.
- DNA polymerase (III) proceeds along a single-stranded molecule of DNA, recruiting free
- **dNTP's** (<u>deoxy-nucleotide-triphosphates</u>) to hydrogen bond with their appropriate complementary dNTP on the single strand (A with T and G with C), and to form a covalent **phosphodiester bond** with the previous nucleotide of the same strand.

DNA polymerases **cannot** start synthesizing **de novo** on a bare single strand. It needs a primer with a **3'OH** group onto which it can attach a dNTP DNA polymerase also has proofreading activities, so that it can make sure that it inserted the right base, and nuclease (excision of nucleotides) activities so that it can cut away any mistakes it might have made.

- *Primase* attaches a small RNA primer to the single-stranded DNA to act as a substitute **3'OH** for DNA polymerase to begin synthesizing from. This RNA primer is eventually removed and the gap is filled in by **DNA polymerase (I)**.
- *Ligase* can catalyze the formation of a phosphodiester bond given an unattached but adjacent 3'OH and 5'phosphate. This can fill in the unattached gap left when the RNA primer is removed and filled in.
- *Single-stranded binding proteins* are important to maintain the stability of the replication fork. Single-stranded DNA is very labile, or unstable, so these proteins bind to it while it remains single stranded and keep it from being degraded.



Eukaryotic DNA Replication

- Eukaryotic DNA replication is the process by which the eukaryotic genome duplicates prior to cell division. Though the basic mechanism of the eukaryotic DNA replication is similar to prokaryotic DNA replication, there are some differences due to the size and the structure of eukaryotic DNA.
- Eukaryotic DNA is double-stranded linear molecules.
- The amount of the eukaryotic DNA is around 50 times more than the prokaryotic DNA.
- Moreover, eukaryotic DNA is tightly packed with histones inside the nucleus of the cell.
- Therefore, DNA replication occurs in three steps; initiation, elongation, and termination.

Initiation

- The eukaryotic DNA replication occurs through multiple replication origins.
- The multiple replication origins form several replication bubbles per chromosome.
- DNA helicase and SSBs are involved in the unwinding and stabilization of the two templates at each origin of replication.

Elongation

- During elongation, DNA polymerases add new nucleotides to the existing 3øends.
- The three types of DNA polymerases which are involved in eukaryotic DNA replication are DNA polymerase , , and .
- The DNA polymerase initiates the DNA replication whereas the DNA polymerase and are involved in the elongation.
- DNA polymerase also requires an RNA primer to synthesize the new DNA strand and the primer is removed by the DNA polymerase .
- The leading and lagging strands are formed in the same manner as in prokaryotic DNA replication. Eukaryotic DNA replication elongation is shown in *figure 2*.



Termination

- Once the leading strand of a one replication bubble meets a lagging strand of a second replication bubble, the replication process is halted.
- Then, the RNA primer is removed, and the gap is filled by the freely-floating DNA polymerases. The nicks are joined by the DNA ligase.
- The multiple replication bubbles are shown in *figure 3*.



Figure 3: Multiple Replication Bubbles

DNA Repair

- Genes encode proteins that correct mistakes in DNA caused by incorrect copying during replication and environmental factors such as by-products of metabolism, exposure to ultraviolet light or mutagens.
- The DNA repair process must operate constantly to correct any damage to the DNA as soon as it occurs.

Differences between RNA & DNA

- RNA is single stranded DNA is double stranded
- RNA has Ribose ó DNA has Deoxyribose
- RNA has Uracil ó DNA has Thymine

GENE EXPRESSION

Transcription and Translation utilize the DNA template code to ultimately produce proteins:

- Transcription ó DNA is template for making RNA (in nucleus). There are 3 types of RNA.
- **Translation (protein synthesis)** in cytoplasm at the ribosome. m-RNA has blueprint, t-RNA transfers amino acids, and Ribosome (r-RNA) allows t-RNA to attach to m-RNA at appropriate site.
- many factors control gene expression including:
 - o factors affecting DNA structure,
 - \circ gene expression,
 - o factors affecting assembly of proteins after
 - \circ translation,
 - \circ hormones,
 - o environmental factors such as viruses.

Types of RNA

Kinds of RNA ó three kinds of RNA are produced in the nucleus using DNA coding strands

- Messenger RNA (m-RNA) ó carries genetic code from DNA into cytoplasm
- **Transfer RNA (t-RNA)** ó brings the amino acids for building of protein to be attached according to the genetic code of the m-RNA
- **Ribosomal RNA (r-RNA)** ó make up the ribosome with a protein unit and reads the code of m-RNA and allow t-RNA to attach and connect amino acids

MicroRNAs(miRNAs)

- miRNAs are RNA genes (20-25 nucleotides long) which are transcribed from DNA, but are **not translated** into protein (<u>non-coding RNA</u>)
- Small non-coding **RNA** molecule which functions in transcriptional and post-transcriptional regulation of gene expression
- MicroRNAs are a class of post-transcriptional regulators
- They have the ability to regulate gene expression.
- *MicroRNAs* are a type of regulatory RNA that can inhibit gene expression by halting translation.
- They do so by binding to a specific location on mRNA, preventing the molecule from being translated.
- MicroRNAs have also been linked to the development of some types of cancers and a particular chromosome mutation called a translocation.

Polycistronic mRNA vs. Monocistronic mRNA

Polycistronic mRNA

- contains for more than one cistron
- codes for more than one protein
- is transcribed from more than one gene (cistron) and has as many initiation and termination codes
- is present in **prokaryotes**



Monocistronic mRNA

- contains codons of a single cistron
- codes for a single protein
- is transcribed from a single gene (cistron) and has one initiation and termination codon
- is present in **eukaryotes**



Control of Gene Expression in Prokaryotes

Gene expressions levels are strictly controlled at many levels to ensure the organism having the appropriate response to its environment or internal changes. This is important for prokaryotes because they are usually single-cell organisms, and they largely depend on their environment for all of their activities

In bacteria transcription often occur as polycistrons, i.e., many functional-related genes are clustered and transcribed under the same types of regulation. These are called **operons.** An operon usually contains regulatory genes and structure genes. The gene expression can be induced under certain circumstances or be constitutive.

Lac & Trp Operons - examples of prokaryotic gene regulation

- Many of the prokaryotic genes as in *E.coli* are expressed or are always turned "on".
- Others are active only when their products are needed by the cell, so their expression must be regulated.
- Examples of Operons in E. coli
 - The genes for the five enzymes in the Trp synthesis pathway are clustered on the same chromosome in what is called the Trp Operon ó
 - If the amino acid tryptophan (Trp) is added to a culture of *E coli*, the bacteria soon stop producing the five enzymes needed to synthesize Trp from intermediates produced during the respiration of glucose so the presence of the products of enzyme action represses enzyme synthesis
 - This is a **repressable operon** where genes are expressed in the absence of a substance and the presence of the substance shuts off the gene
 - The genes that code for the enzymes needed for lactose catabolism are clustered on the same chromosome in what is called **the Lac Operon**
 - Prokaryotes, such as E. coli have a mechanism for metabolizing lactose ó a sugar used for energy.
 - Three proteins or enzymes are needed in lactose metabolism and they are encoded in a single expressible unit of DNA called the *lac* operon.
 - *E. coli* only express the genes and make these enzymes when lactose is available to be metabolized.
 - This is an **inducible operon** where genes are expressed in the presence of a substance





Prokaryotic gene organization

Promoter typically within 500bp of translational start site (ATG start codon of ORF)

- Often (not always) polycistronic, containing more than one ORF per promoter. Co-transcribed genes are said to be in an operon.
- No introns are present
- Multiple ribosome binding site (RBS) present in polycistronic mRNA



- Extended promoter with enhancers/silencers often more than 1kb away
- RNA is processed (post-transcription modifications)
 - Introns are excised and exons spliced together
 - o 5ø cap and poly-A tail added to transcripts
- Only one open reading frame (no operons/polycistronic mRNAs)
- No ribosome binding site on mRNA

 \circ ribosome binds 5ø end of mRNA and scans for translational start codon

Prokaryotic promoter - region of DNA that initiates transcription of a particular gene

- The -10 and -35 regions of a canonical prokaryotic promoter are recognized by the bacterial RNA
- polymerase sigma subunit.
 The site of transcription initiation occurs 10bp downstream of the -10 regionøs center, hence its name. The consensus sequences for the -35 and -10 sequences are shown.
- These sequences lead to strong transcription of a promoter, but variations of the sequences are common and can still interact strongly with the RNA polymerase sigma subunit.



TRANSCRIPTION UNIT

Control of Gene Expression in Eukaryotes

Eukaryotic genes usually contain three basic regulatory components:

- Enhancers short regions of DNA that can be bound with proteins to *promote expression* of a distal or a proximal gene.
- **Promoters** proximal DNA sequences that binds to RNA polymerase *for regulating gene expression*.
 - region of DNA that initiates transcription of a particular gene
 - located near the genes they transcribe, on the same strand and upstream on the DNA (towards the 3' region of the anti-sense strand also called template strand and non-coding strand
- **TATA Box** binds to transcription *factor for regulating gene expression*, usually within 30bp of the transcription start site.



Components of a Genes and promoters

- Promoter: regulatory sequence upstream of gene to which RNA polymerase binds to initiate transcription
- **Transcriptional start site**: nucleotide within promoter where transcription begins. It is upstream of the ATG start codon where translation begins. Difficult to identify with genome scanning software, so often needs to be determined experimentally.
- 5' Untranslated Region (UTR): stretch of mRNA between the transcriptional start site and the translational start site, often controls rate of translation initiation
- **Open Reading Frame (ORF):** the part of a gene that codes for protein. Begins with an ATG start codon and ends with a stop codon that are in frame with each other. Very easy to recognize with genome scanning software.
- **3' Untranslated Region (UTR):** stretch of mRNA between the final stop codon and the end of the transcript. Like 5øUTR, 3øUTR can also be regulatory.

Eukaryotic promoter

- Eukaryotic promoters also contain a õTATAö box upstream of the transcriptional start site, and this sequence is bound by a series of transcription factors that recruit RNA polymerase to initiate basal levels of transcription.
- Like in prokaryotes, the rate of transcription is then further influenced by other transcriptional activators/repressors that bind other locations in the promoter and/or at enhancer sequences.
- Unlike prokaryotes, these enhancer sequences can often be very distant from the gene they regulate. DNA with the enhancer sequence can loop back to allow a transcriptional regulator to interact with RNA polymerase at the promoter.



Controls include:

- Transcriptional Control
- Post transcriptional Control ó assembling proteins
- Cell differentiation and specialization
- Turning genes õonö and õoffö
- Chemical Signals ó Hormones
- Chemical Modifications
- Relocation of DNA ó transposons
- Abnormal Expression of Genes

Transcription

Transcription - production of RNA in the nucleus using a DNA segment as a template and **RNA polymerase** as the key enzyme.



Post-transcription Modifications in Eukaryotic Cells

RNAøs are modified in eukaryotes before leaving the nucleus.

- Pre-m-RNA has exons (coding segments) and introns (noncoding segments between exons)
- introns (the noncoding segments) are removed
- a cap is added to the 5° end
- a poly A tail is added to the 3^{ϕ} end before it leaves the nucleus



		U	С	А	G			
First position (5' end)	U	UUU UUC UUA UUG	UCU UCC UCA UCG	-	UGU Cys UGC Cys UGA Stop UGG Trp	U C A G	Third	
	с	CUU CUC CUA CUG	CCU CCC CCA CCG	CAU His CAC His CAA GIn CAG GIn	CGU CGC CGA CGG	U C A G		
	(5 end)	А	AUU AUC AUA AUG	ACU ACC ACA ACG	AAU Asn AAC Asn AAA Lys AAG	AGU Ser AGC Ser AGA Arg AGG	U C A G	(3' end)
	G	GUU GUC GUA GUG	GCU GCC GCA GCG	GAU GAC GAA GAG GAG	GGU GGC GGA GGG	U C A G		

Universal Code (Codon = Amino Acid)

Amino acid name

Ala = alanine	GIn = glutamine	Leu = leucine	Ser = serine
Arg = arginine	Glu = glutamate	Lys = lysine	Thr = threonine
Asn = asparagine	Gly = glycine	Met = methionine	Trp = tryptophan
Asp = aspartate	His = histidine	Phe = phenylalanine	Tyr = Tyrosine
Cys = cysteine	Ile = Isolevcine	Pro = proline	Val = valine

- Each three base codon on the messenger RNA (m-RNA) is a code for an amino acid
- There are 64 possible three base codons ó 61 are codes for one of the 20 amino acids
- The three remaining codons are termed **stop codons** because the signal the end of a peptide segment
- Notice that many of the amino acids have more than one codon
- A three base **code** on the **DNA** produces the **mRNA codon**
- The three base code on the **t RNA** is termed an **anticodon** because it will bond to a m-RNA codon during translation or protein synthesis

Translation (Protein Synthesis)

- Translation is the universal process of synthesizing proteins as the second step in gene expression.
- Both prokaryotic and eukaryotic ribosomes decode mRNAs in fundamentally similar methods
- The main difference between prokaryotic and eukaryotic translation is that prokaryotic translation is a simultaneous process with transcription whereas eukaryotic translation is a separate process from its transcription.
- **Translation** ó m-RNA template containing genetic code is used to form amino acid sequence using m-RNA, t-RNA, and r-RNA (ribosomes) occurs in the cytoplasm at the ribosome. Many key enzymes (proteins) are involved.



The steps of translation:

- Initiation: mRNA enters the cytoplasm and becomes associated with ribosomes (rRNA + proteins) and tRNAs, each carrying a specific amino acid, pair up with the mRNA codons inside the ribosomes. The base pairing (A-U, G-C) between mRNA codons and tRNA anticodons determines the order of amino acids in a protein.
- **Elongation:** involves the **addition of amino acids one-by-one**: As the **ribosome** moves along the **mRNA**, each **tRNA** transfers its amino acid to the growing protein chain, producing the protein



• **Termination:** when the **ribosomes** hits a stop codon - UAA, UGA, or UAG ó no tRNA with its amino acid can be added so the ribosome falls apart and the process ends. The same **mRNA** may be used hundreds of times during translation by many **ribosomes** before it is degraded (broken down) by the cell.

Nuclear vs Cytoplasmic DNA in Eukaryotic Cells

- Nuclear DNA ó in chromosomes within the nucleus of the cell
- Cytoplasmic (or Organelle DNA) ó in chloroplasts and mitochondria
 - Mitochondria and Chloroplasts have DNA similar to Prokaryotic cells
 - It is believed that these organelles were once independent prokaryotes who took up residence and formed a mutualistic relationship
 - They are involved in energy transfer- photosynthesis & respiration

• Chloroplast DNA (cpDNA)

• Mitochondrial DNA (mtDNA)

Features:

- Maternal inheritance
- Resemble prokaryotic DNA
- Slow accumulation of mutations at the population level

Mitochondrial Inheritance -

- The inheritance of a trait encoded in the mitochondrial genome
- Mitochondrial DNA or mtDNA -genetic make-up of mitochondria, genetic code and patterns transmitted through mother.
- The mtDNA is circular and resembles prokaryotic DNA
- The mitochondria are responsible for energy production
- Mitochondria can reproduce independent of the rest of the cell ó an advantage in energy production
- Persons with a mitochondrial disease may be male or female but they are always related in the maternal line and no male with the disease can transmit it to his children
- Mitochondrial myopathies are a group of neuromuscular diseases caused by damage to the mitochondria-small, energy-producing structures that serve as the cells' "power plants."



Mutations

- Gene ó section of DNA which carries the blueprint for making a peptide strand or RNA.
- DNA in the living cell is subject to **many chemical alterations** If the genetic information encoded in the DNA is to remain uncorrupted, any chemical changes must be corrected.
- A failure to repair DNA produces a **mutation**
- Mutation ó changes in genetic code (DNA blueprint) of genes or chromosomes and causes changes in expression or in the directions for making protein or RNA
- Gene mutation
- Chromosomal mutation
- Agents causing mutations ó radiation, chemicals, excess heat, viruses

Mutagenesis is the process of making mutations. These mutations can be random or can target specific genes.

Forward genetics is the classical approach to genetics where mutations are randomly made in the genome and the experimenter screens for a desired phenotype. Mutations that lead to the desired phenotype are then mapped to determine which gene(s) were mutated.

Common examples of random mutagenesis: radiation, chemicals, transposon insertions

Reverse genetics is a more recent approach to genetics in which a specific gene of interest is mutated in order to determine the phenotype(s) that arise as a result of the mutation.

Common examples of targeted mutagenesis: targeted transposon insertion, deletion, allelic exchange, CRISPR-Cas modifications

Genetic Disorders

- Causes of mutations chemicals, radiation, temperature, viruses
- **Nondisjunction** ó chromatids do not separate properly during meiosis. Individual formed from such gametes have extra or missing chromosomes. as Downøs Syndrome
- **Trinucleotide repeats** sequences of 3 nucleotides is repeated, often several times in a gene. When too many repeats are formed it cause genetic disorders from the triplet nucleotides being repeated too often, such as Huntingtonøs
- **Defective genes** ó does not produce correct protein, such as sickle cell anemia (A & T traded places)
- Genetic disorders and their causes as nondisjunction (Down& syndrome), trinucleotide repeats (fragile X and Huntington), defective genes (sickle cell anemia, hemophilia)
- **Human genetic disorders** can be dominant, recessive, sex-linked, epistatic, variable expressed
- Crossover frequency during meiosis, pieces trade places ó determining frequency