Molecular Biology

2025-2024

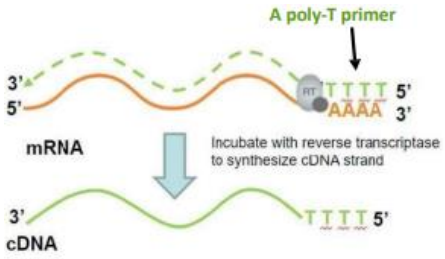
DR.Ahmad Al Qawasmi

ARKAN ACADEMY

Recombinant Proteins

- Production of a protein outside our bodies in a biological system requires cDNA for the gene of interest, expression vector and a biological system

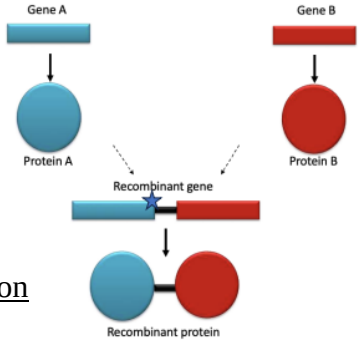
- A double stranded ***cDNA*** is produced by *reverse transcriptase* using a mature **mRNA** as a template
  - mRNA has a *poly A tail* which is complementary to the *poly T primer* used by the enzyme
  - It contains *only exons* and coding region



- Expression vectors are usually plasmids having the same features of the cloning vectors (origin of replication, antibiotic resistance gene, restriction site) with 3 other features:
  - *Promoter sequences* upstream the gene to be inserted to initiate transcription
  - *Ribosomal binding sequences* (*Shine Dalgarno [SD]sequences*) to initiate translation
  - A *transcription termination sequence* to stop transcription

- The recombinant DNA (vector + gene of interest) is inserted into the cell (bacterial or eukaryotic)
  - Transcription and translation occur producing the needed protein
  - Using ***bacterial*** cells to produce human proteins faces many challenges including:
    - ✓ *No internal disulfide bonds*
    - ✓ *No post-translational modification* (example: glycosylation)
    - ✓ *Protein misfolding*
    - ✓ *Protein degradation*
  - So, we use a ***eukaryotic*** system such as yeast

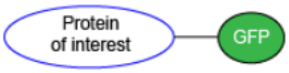
- **Recombinant protein** is a protein produced from the fusion of the genes of ***2 different proteins*** and can do the functions of the 2 original proteins
  - Recombination between ***domains*** of different proteins is also valid



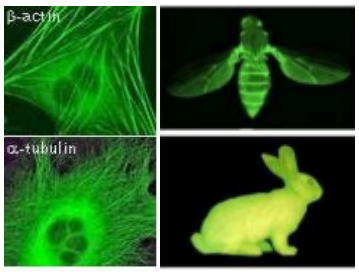
- It can utilized for the <u>detection of proteins</u> and study <u>protein-protein interaction</u>

GFP-tagged proteins

- Green Fluorescent Protein (GFP) allows for ***protein detection*** rather than for purification purposes

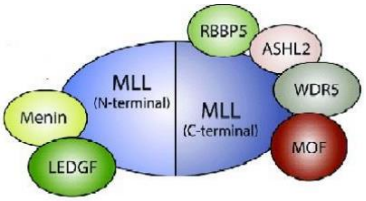


- Recombination between GFP gene and the gene of protein of interest causes the produced protein to give a *green light* which can be used to <u>track the protein in a living cell</u> or organism
  - It is also used to <u>study the morphology and organization</u> of molecules in the cells



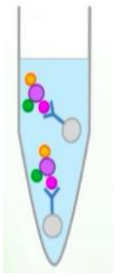
Protein-protein interaction

- Proteins interact with each other in order to do their functions
- It can be studied by 2 techniques:
  - Co-immunoprecipitation
  - Yeast two-hybrid system



(Co)-Immunoprecipitation

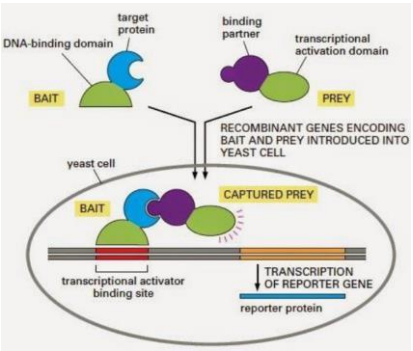
- This technique depends on the antibodies which can bind specifically to the protein of interest
  - The ***antibodies*** are bound to beads and then added to a mixture of proteins
  - Antibodies will bind to the protein of interest and ***precipitate*** it with all proteins that interacts with it (co-precipitation)



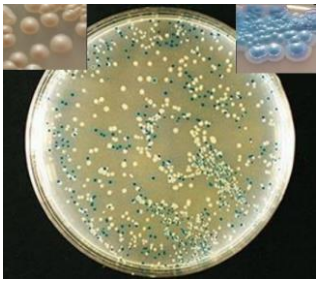
Yeast two-hybrid system

- In yeast, there is an enzyme called ***β-galactosidase*** that can cleave and convert X-gal (lactose analog) into a blue product
- β-galactosidase is encoded in ***Lac Z gene*** which requires a transcription facto called gal 4
  - *Gal4* must bind to *UAS* element upstream the promoter to activate the transcription
  - Gal4 consists of 2 domains (DNA-binding domain **BD**, Activation domain **AD**)

- Lac Z gene can be used as a reporter for protein-protein interaction
  - In order to discover unknown proteins (Y) that interact with a known protein (X)
  - X gene is cloned and produced recombined with the DNA binding (DB) domain
  - Y genes are separately cloned and produced recombined with the activation domain (AD)
  - Both recombinant plasmids are transferred into yeast cells so all of them express the known X gene-BD hybrid, but each one expresses a different unknown Y gene-AD hybrid



- If Y **interacts** with X: the clones will appear blue
- If Y **doesn't interact** with X: the clones appear white
- Also, if the plasmid inserted to the cell has only AD or BD without the other: the clone appears white



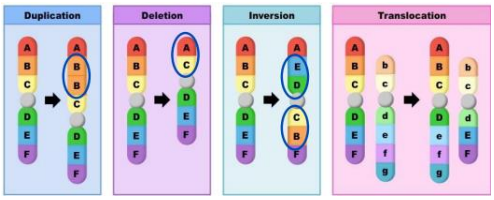
Mutations

Mutation: A rare change in DNA (<1% of population) may cause a disease

Polymorphism (SNP): a genetic variation (>1%) doesn't typically causes a disease

- Mutation is a <u>change in the genetic material</u> which can be:
  - ***Somatic mutations*** occur in somatic cells and are *not transmitted*
  - ***Germline mutations*** occur in gametes and are *heritable*

- The damaging effect of mutations has variable sizes from micro-mutations or macro-mutations
  - ***Macro-mutation:*** Changes in the *chromosome level*, includes:
    - ✓ *Translocation* (from one chromosome to another)
    - ✓ *Inversion* of DNA segment (on the same chromosome)
    - ✓ *Duplication*
    - ✓ *Deletion*



  - ***Micro-mutation:*** Changing the nucleotide sequence of a small regions of the DNA
    - ✓ It involves point mutations and deletion and insertion of long nucleotide stretches

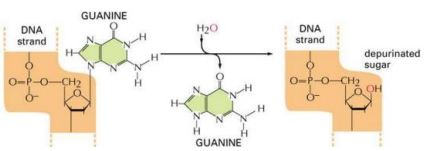
- ***Point mutations:*** The most common and include <u>substitutions, insertion, and deletion</u>
  - It involves a *single base pair* is addition, deletion, or change (substitutions)
  - *Silent mutation: No change* in the amino acid sequence
  - *Missense mutation:* Substitution of one amino acid for another (*incorrect amino acid*)
  - *Nonsense mutation:* Conversion of a codon into a *stop codon* (*premature termination* of translation)
  - *Frameshift mutation:* Insertion or deletion of nucleotides shifting the reading frame, *altering the entire downstream* amino acid sequence

Condition	Type of Mutation	Effect on Protein Synthesis
Duchenne muscular dystrophy	Frameshift (disrupts the reading frame of the dystrophin gene)	Complete loss of protein function: more severe form.
Becker muscular dystrophy	Frameshift (in less critical regions of the dystrophin gene)	Partial protein function retained: less severe but still significant.

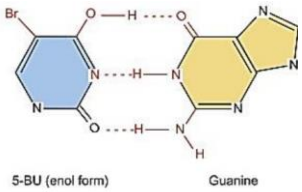
- At ***repeated sequences***, *strand slippage* occurs during DNA replication
  - Causing mutations (addition or loss of nucleotides, polymerase errors), and DNA looping

- ***Deamination*** (<u>*spontaneous*</u>) causes the change of a base into another:
  - Deamination of *cytosine* yields ***uracil*** (C into U)
  - Deamination of *methylated cytosine* yields ***thymine*** (mC into T)
  - Deamination of *adenine* yields ***hypoxanthine*** (A into hypoxanthine which is read as G)

- ***Depurination*** (<u>*spontaneous*</u>) is the *cleavage of the glycosidic bond* between the base (especially purine) and deoxyribose creates an apyrimidinic or apurinic site (***AP site***)
  - During replication, a random base can be inserted across from an AP site resulting in a mutation



- ***Incorporation of base analogs*** (*induced*) which have a similar structure to normal nucleotides and are incorporated into DNA during replication, and can cause mutations
  - *5-bromouracil (5-BU)*, an analog of ***thymine***, pairs with A, but when ionized, it pairs with ***G***
  - Its deoxyriboside derivative (***5-bromo-2- deoxy-uridine***) is used to treat neoplasms



Repair mechanisms

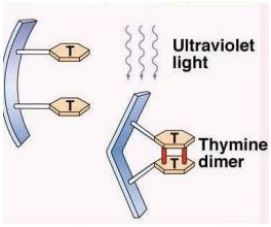
- It includes many mechanisms:

1. Prevention of errors before they happen

- Enzymes *neutralize potentially damaging* compounds before they even react with DNA
- *Detoxification* of reactive oxygen species and oxygen radicals by ***superoxide dismutase*** and ***catalase***

2. Direct reversal of damage

- ***Pyrimidine dimers:*** It is covalent interactions (50–100 reactions /sec) between 2 adjacent pyrimidines
  - Caused by *ultraviolet (UV)* wavelength of sunlight which is mutagenic
  - Commonly between two thymine
  - It causes distortion of DNA structure (*stops replication and transcription*)
  - They are reversed in bacteria by enzymes known as ***photolyases*** (not human)

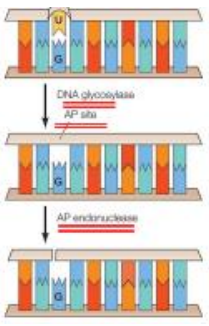


- ***Specific mispairing*** caused by the modification of some nucleotides:
  - *Alkylating agents* can transfer methyl group to guanine forming *6- methylguanine*, which pairs with ***thymine***
    - ✓ Repaired by *O6-methylguanine methyl**transferase***
  - Addition of large chemical adducts by carcinogens

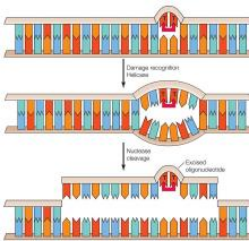
3. Excision repair pathways

- ***Base excision repair pathway:*** The cleavage of damaged base forming ***AP sites*** then it is repaired by an ***AP endonuclease repair*** pathway

- DNA glycosylases cleaves N-glycosidic bond (not the phosphodiester bond)
  - *Uracil-DNA glycosylase* removes **uracil from DNA** caused by:
    - ✓ Spontaneous deamination of cytosine
    - ✓ Incorporation of dUTP can lead to a C→T transition, if unrepaired
  - *AP endonucleases* cleave the phosphodiester bonds at AP sites and removes the deoxyribose then *DNA polymerase* fills the gap, and sealed by *DNA ligase*



- ***General excision repair (nucleotide excision repair)***, used for mutations such as ***thymine dimers***
  - It is done by XP proteins
  - Damaged DNA is recognized (by *XPC*) then recruits the helicases
  - DNA is unwound around the site of damage by a helicase (TFIIH)
    - ✓ ***Helicase (TFIIH)*** is *XPB* and *XPD* aided by XPG
    - ✓ *XPA*: ***replication protein A*** (RPA, SSB) which binds ssDNA



  - The DNA is then cleaved on both sides of a thymine dimer by ***endonuclease*** (*XPF/ERCC1*) resulting in the excision of oligonucleotide containing damaged bases
  - The gap is then filled by *DNA polymerase* and sealed by *ligase*

- Defects in nucleotide excision repair cause a condition known as ***Xeroderma pigmentosum (XP)*** and ***Cockayne's syndrome*** where individuals are extremely sensitive to UV light and develop multiple skin cancers on the regions of their bodies that are exposed to sunlight

4. Transcription-coupled repair

- There is a preferential repair of the transcribed strand of DNA for actively expressed genes
- RNA *polymerase pauses when encountering a lesion*, where TFIIH and other factors carry out the incision, excision, and repair reactions and then, transcription can continue normally
- ***Cockayne's syndrome:*** a condition caused by defects in XP proteins, but predominantly CSB
  - It is caused by a defect in preferential DNA repair of transcriptionally active DNA genes
  - Patients are characterized by *short statu*re, an abnormally *small head* (microcephaly), and *neurologic* abnormalities leading to intellectual disability and may have light-sensitive skin (*photosensitivity*)

5. Mismatch repair and replication-related repair

- During replication ***MSH and MLH*** proteins *recognizes* and bind the mismatched bases within the lagging and leading strand DNA, where they are excised and replication continues
- Mismatch repair is 3-4 times *more effective* on the **lagging** strand than the leading strand
  - DNA polymerase ***ε*** is ***more accurate*** than DNA polymerase ***δ***

- ***Hereditary nonpolyposis colon cancer (HNPCC)*** affects one in 300 people and 15% of colon cancer cases which is mainly caused by mutations in MSH and MLH
  - FAP (familial adenomatous polyposis), numerous polyps develop throughout the colon
  - The most common cancers are colorectal, breast and lung

6. Translesion DNA synthesis

- *Specialized DNA polymerases* (not the typical replicative enzymes) can synthesize DNA over the lesions, but they have low fidelity and lack proofreading mechanisms, hence, are error-prone
  - More likely to introduce additional mutations during DNA synthesis
  - But they are *selective toward the introduction of A nucleotides*, so that TT dimers are often replicated correctly (repair the ***thymine dimers***)

7. *Recombinational repair*

- Two repair mechanisms: Non-homologous end joining (NHEJ) and Homologous repair

- *Non-homologous end joining (NHEJ) repair*
  - DSB ends are recognized & bound by *Ku70-Ku80 complex*, which recruits additional factors including a DNA ligase
    - ✓ If ligation is possible, *ligase* will immediately ligate the 2 DNA strands and the original sequence can be retained
    - ✓ If direct ligation is not possible, additional proteins are needed but ***insertions or deletions mutations (INDELS)*** are introduced

Ionizing radiation can cause different type of DNA damage:
- Creation of AP sites
- Base damage
- ***Strand breaks*** either single (SSB) or double (DSB) which can result in blunt or staggered ends

- *Homology-Directed Repair (HDR, homologous recombination)*
  - Uses DNA sequences within the undamaged ***homologous chromosome*** from the other parent
  - More accurate than NHEJ
  - In germline cells, HDR generates ***genetic diversity*** by producing different combinations of alleles

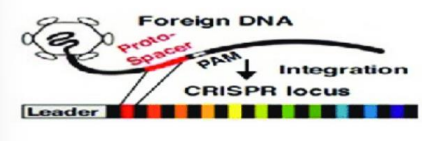
- Mutations in *BRCA1 and 2* genes are responsible for a portion of hereditary ***breast and ovarian cancers***
  - *BRCA1* activates homologous recombination repair of DNA **double**-stranded breaks
    - ✓ BRCA1 is also involved in transcription and transcription-coupled DNA repair
    - ✓ It is activated by ATM kinase
  - *BRCA2* can recruit Rad51 to the **ssDNA**

Type of DNA repair	Mechanism	Genes/proteins
Base excision repair	Removal of abnormal bases	DNA glycosylases
Nucleotide excision repair	Removal of thymine dimers and large chemical adducts	XP proteins, CSB
Mismatch repair	Correction of mismatched bases caused by DNA replication	MLH1, MSH2
Post-replication repair	Removal of double-strand breaks by HR or NHEJ	BRCA1, BRCA2

CRISPR/Cas9

- *CRISPR:* clustered regularly interspaced short palindromic repeats in the bacterial genetic system
  - It constitutes the immune system of bacteria against phages
- *Cas9* is *RNA-guided nuclease* that can introduce ***double-strand breaks*** creating blunt-ended fragments
  - ***guide RNA (gRNA) or single guide RNA (sgRNA)*** short RNA fragment complementary to the ***target*** segment of the genome

- When a phage infects a bacterial cell, the cell degrades the ***phage DNA*** into smaller pieces using restriction endonucleases and integrates one of these fragments into the CRISPR cluster
  - Bacteriophage DNA fragments are ***separated by the CRISPR*** repeats, when the phage infects the cell again, the cell transcribes the DNA into gRNA, which is integrated into the Cas9 nuclease and guides it to the phage DNA to degrade it



- Plasmids carrying genes for Cas9 and a gene expressing gRNA sequences homologous to the target gene, are introduced into the cells
  - The gRNA directs Cas9 to the target gene and Cas9 creates a double-stranded break.
  - In the presence of a homologous copy of the gene, *homology-directed repair* replaces the broken target gene with the mutated copy
  - In the absence of a homologous copy of the gene, *non-homologous end joining* reseals the broken DNA introducing Insertion/deletion mutations (INDELS) that make the gene nonfunctional

- Using CRISPR-Cas9 for Gene Editing
  - It can be used for either *replacing a mutant gene* by a normal one or *turning off its expression*
- *Transcriptional regulatory factors* can be added to an enzymatically ***inactive or "dead" Cas9 (dCas9)*** enabling these factors to ***turn genes on or off***
- *GFP* can be added to dCas9 to find a particular stretch of DNA in the cell or even ***visualize*** the three-dimensional architecture of a chromosome

- Cas9 makes a double stranded DNA cut, and some engineered Cas9 can make single-stranded cuts
- ***Cas12a:*** Can make *staggered* cuts
- ***Cas13a:*** *RNA* endonuclease

ARKAN

ACADEMY

علم في كل مكان

Arkan academy

www.arkan-academy.com

Arkanacademy

+962 790408805