In this sheet we are going to talk about four topics :

1-**SNP/single nucleotide polymorphism** (Not in the slides – I hope the sheet will be enough to understand it)

2- **DNA mutations** (slide 132 – slide 137)

3-**DNA rearrangement** (slide 138 – slide 146)

4-**DNA repair** (slide 147 – slide 159)

1st Topic: Single Nucleotide Polymorphism

Firstly, we will talk about the most common mutation in the human genome which in the **Point Mutation .**

When we say point mutation then we are talking about a single nucleotide involved, and usually these mutation is a **substitution mutation** (removing a nucleotide and put another one) . Today we refer to these mutations as a (SNPs) which is **S**ingle **N**ucleotide **P**olymorphisms .

***What is Polymorphism ? and what is the SNP ??**

**SNP : is a variation between the humans because of changing in a single nucleotide .

*Why we call them SNPs ? - Single nucleotide: because only one nucleotide involved.

- Polymorphisms: means a different forms of the sequence.

Polymorphism in genetics talks about genetic variant that appears in at least 1% of a population. But in molecular biology we don't concern about population, we are only interested in the **effect** of these polymorphism, will it cause a disease ?? or it will just make a change in a character like tall or short person ? or maybe there will be **NO** effect of these SNP .

***The percentage of these SNPs in our genome is 0.1 % , what does that mean ?**

If we took our genome which is consist of a 6.000.000.000 nucleotide, and we start screening, we will see that after every 1000 nucleotides there will be a nucleotide that differs from a person to another , which will result in a diversity (pathologically and physiologically) .So we are (human beings) similar in 99.9% of our genome and the differences between us are in the 0.1% .

***The discovery of the SNPs makes the studying of human genome much easier , how ?**

1) When the scientists finish the human genome map , they are interested in knowing the effect of each part of these genome , but the more clever persons think about studying the differences between the genomes , which is only 0.1% of our genome . (the 99.9% makes us normal humans but the 0.1% is what makes some of us predisposed to cancer , diabetes .. etc)

* Thinking about the 0.1% SNPs , how much it forms of our genome ?? we have 6 billion nucleotides and 0.1% of these 6 billion equals 6 million nucleotide , lets say around 70% of our genome is non-coding , then the effective SNPs approximately 2 million (in the genes) , we don't forget that there is an introns in the genes and **most** (not all) of the mutations in these introns have no effect , then the final number of the effective SNPs is around **200,000 – 300,000** and if we just study these points it will be much easier to know the differences than studying the whole genome .

2) There is a big project done by the scientists to make a **(SNP profiling) or (SNP Catalogue)** to identify each SNPs in the human genome separately and study it and its effect e.g SNP responsible for the differences in the color of the eye or increasing the recurrence of a cancer and so on .

*Just to make it more clear, after ending the project of the (SNPs Catalogue) we can take a cell from the embryo and indentify the SNPs not the whole genome , for example if we see a ATTAAC sequence in a specific place we will go to the catalogue and see what does these sequence mean (increasing the chance to have diabetes for example) .

One of the students asks a question : in genetics we said that the insertion and deletion is much more common than the substitution and here we said that the substitution is more common ?

The doctor answer was that the insertion and deletion is not more common , they are more serious and dangerous because there will be a frame shift in the whole amino acids sequence , and there is a big rule in genetics says that (**the least dangerous mutation is the most common mutation**) why ? simply because the more dangerous mutations are usually self-limited , in other words , the affected patient will not survive for a long time and thus he will not pass these mutation to his offspring .

2nd Topic : DNA mutations

After talking about polymorphism in general and SNPs we will return to our topic about the mutations .

First / The substitution mutation may be a :

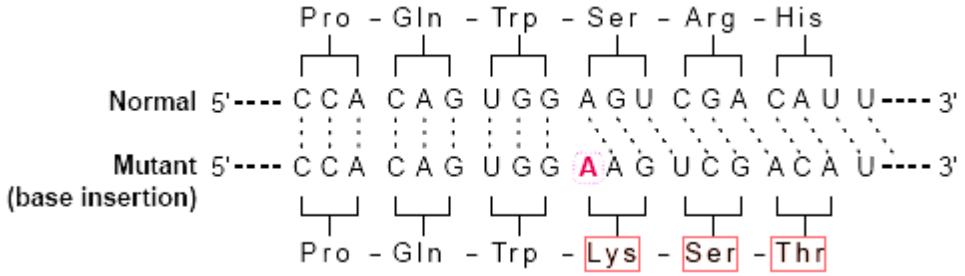
1- **Silent mutation** : it has no effect in the resulted amino acid e.g (CG**A**) and (CG**G**) , both of them codes for Arginine as a resulted amino acid . Note : usually the two codons that code for the same amino acid are different in the **THIRD** nucleotide .

2- **Missense mutation** : there is a change in the formed amino acid e.g changing from (C**G**A) to (C**C**A) will cause a change in the amino acid because :
CGA = Argenine CCA = Proline
What about the effect of these mutation ? actually it depends :

The amino acid that comes after a substitution mutation may has a similar properties to the first one so it will do its function (no change in the protein structure)and there will be **no effect .** Or the properties may be different then there **will be an effect.**

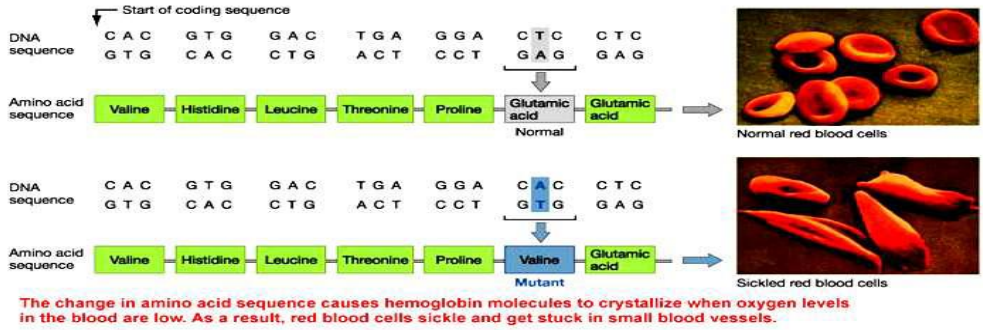
3- **Non-sense mutation :** the new codon that formed from the mutation is a **Stop codon** it will results in a premature termination of the translation or what we call **Truncated protein .**

Second / The insertion and deletion mutations are adding or removing a single nucleotide (or more) to the sequence and the result will be a frame shift to the whole nucleotides .



(These figure is an example of the frame shift that caused by the insertion of an (A) nucleotide)

Mutation causing sickle cell anemia



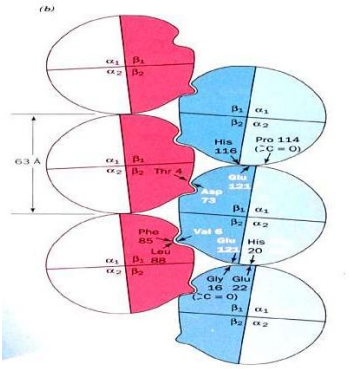
Sickle cell anemia is an example of a disorder that caused by a SNP (remember that not all SNPs have a pathological effect) .

As we see in the figure above , the substitution will result in a formation of a **Valine** instead of **Glutamic acid** , the properties of these two amino acids is greatly different , how ?

Glutamic acid is * polar *charged (acidic) *found on the surface forming ionic bonds

Valine is *non-polar *hydrophobic ***found hidden inside the protein**

When the Valine hides inside the cell it will form a cavity on the surface of the protein , and by a chance, these cavity will be occupied by a protrusion of another **globin protein** (like what happens in the figure below).



Where is the problem in these structure ? the problem that the protein (hemoglobin) will not move freely in the surface of the Oxygen-exchanging RBCs , instead , they will accumulate as a **fibers of hemoglobin** , then these will cause a change in the circular shape of RBCs . The cell will elongates forming the sickle-shaped cell, and these shape will make two problems :

1-It will reduce the surface area that will bind to the oxygen , reducing the efficiency in transporting the oxygen .
2-These shape is not a smooth one , so it's able to block the small capillaries , the result will by a hypoxia , coagulation and other vascular diseases.

3rd **Topic :** DNA rearrangement

Now we will talk about the **Genetic Rearrangement** and its types , which are :

1- **Homologous recombination** or **Crossing over** and these type is related to the chromosomal level at the process of **Meiosis .**
During meiosis when the maternal and paternal chromosomes take their places there might be an exchange of an opposite fragments of the homologous chromosome (like exchange between paternal Ch9 and maternal Ch9) , and usually these fragment are similar so these type of rearrangement is usually not pathological , instead it is actually helpful in the **diversity .**

2- **Translocation (Similar to crossing over)** like the crossing over there will be a chromosomal-fragments exchange but with two differences :
1-between **non-homologous chromosomes** (Ch9 with Ch22 for example)
2-**Not** during meiosis , in somatic cells normal life.

Why these translocation occurs ?? usually the reason is a **mutagen , radiation** for example , the radiation is a waves with energy that can cross the cell , during the crossing it might crosses the **weak points** in the chromosome , breaking the chromosomes in these points , and the fragments that formed from these breaking may be rearranged in a wrong way in another chromosome causing the translocation mutation **. Chemicals** are also an example of a mutagen that can cause translocation by affecting the weak points .

What is the dangerous about the translocation since there is no loss or gain of an nucleotide ?

One of the cases is translocation of part of chromosome from inactive region to highly active region (the upcoming example will make it more clear) .

the oncogene (cancer-related gene) must always be beside a weak promoter to prevent it from being highly activated causing a cancer. During the translocation mutation , the breaking point of the chromosome may be between the oncogene and the weak promoter , then the oncogene will be translocated to a region near a strong promoter (by a chance) , the result will be an **over-expression** of the oncogene forming a cancer .

One of the classical examples of the previous case is the **B-cell** which secrete the antibodies that are formed by a genes found on **chromosome 14**, this chromosome has a well-known weak point , located downstream to the genes that code the antibodies , the **region downstream to these genes in Ch14** must not be occupied by an oncogene because the result will by a cancer. Two examples caused by these type of translocation are :

1-Burkitt Lymphoma caused of translocation between (8, 14) , part of Ch8 has an oncogene called (MYC) , which will be translocated to Ch14 (below immunoglobulin region) causing Burkitt's Lymphoma.

2- Follicular lymphoma caused by translocation between (14,18) , part of Ch18 has an oncogene called (BCL 2) , which will be translocated to Ch14 (below immunoglobulin region) causing Follicular Lymphoma.

Just to remember : we talked about two types of genetic rearrangement :

1-homolgous recombination

2-translocation

3-The third type of genetic rearrangement is **Stem cells ,** these type is special case for **B-cells** , (B-cells have a huge area of genes called **immunoglobulin region** that produce a specific antibody for each antigen enters our body , all these numerous antibody formed by these area , how ?? by mixing and joining different segments) . In this case of DNA rearrangement , the mother B-cell , stem B-cell , Naïve B-cell – all are the same - , when it wants to divide and form a clone of B-cells that secrete a specific Ab , it will **NOT** gives her daughter cells the whole genome , it gives her everything but when it comes to the **Immunoglobulin region** it gives her daughter **ONLY** the needed genes to produce these specific antibody (so actually it inherits 99.99% of her mother's genome **not everything**) .

4-**Transposons (jumping genes) :** segments in our genomes , usually found in the non-coding region , huge in there number (thousands) , these genes make an enzyme called **transposase** and the only function of this enzyme is 1.going back to the genes that code him 2. remove them from their place 3.put them in other region in the genome .

In other words , the transposons **, a genes that encode the enzyme that move them from one place to another .**

Do they cause a problem ?? although most of their "jumping" have no effect because most of our genome are non-coding , some of them do , in some cases , and during their jumping , they might insert themselves **inside** an important gene and interrupt this gene , a process called **insertional mutagenesis .** Sometimes they might insert themselves in a **tumor suppressor gene** , interrupt this genes , allowing the cell to be cancerous .

The other type of transponsons called the **Retrotransposons , what is the difference between them ?**

Transposons (genes) → RNA →Protein (transposase) →the protein will move the transposons

Retrotransposons (genes) →RNA → reverse transcription → DNA → go to other places in genome

So actually the transposons move themselves , but the retrotransposons send a copy of them by a reverse transcription process (RNA intermediate) .

Now , the interesting thing about these retrotransposons , is the **relation between them and the retroviruses .**

Why we think about the relation between the retrotransposons and retroviruses?

1 - The most famous retrovirus is the **HIV ,** it's a RNA virus that enters the cell , then the reverse transcription process will start , forming DNA which will be inserted in the cell genome , actually like the retrotransposons , so **there is a relation .**

2 - At the both ends of the retrotransposons (as a genes) , there is a repeat sequence of nucleotides for several hundreds of times (300-500nucleotides)we call these repeats as **Direct repeats**

in the retrotransposon and **terminal repeats** in the retrovirus . (direct repeats and terminal repeats are almost similar)

But , who comes first ? the retrovirus or the retrotransposons ?

The two theories are acceptable :

first theory / The retrotansposon comes from retrovirus : if a virus infect our cell , first process that will take a place in reverse transcription , forming DNA from the viral RNA , then the DNA will be integrated in the cellular genome , and as the cell divides the viral genome will divides with it. Fortunately , the virus cannot be inherited to the offspring if the effected cell was a somatic cell , but , because we were in the earth since a millions of years , there must be a chance to integrate a retrovirus to a **germ cell** , then it will be inherited and become a part of our genome . So these theory says that **the transposons are scars of the surface of the genome caused by a retrovirus infection** (exactly like the scars in surface of the moon that caused by النيازك)

Note : the viruses that have the ability of integration is a source of variation in human genome .

second theory / Retrovirus comes from retrotransposon : while the transposon was jumping inside the cell , and by a chance , it jumped strongly , so it get out of the cell and take a part of the cell membrane forming the viral envelop , then we got the **retroviruses .**

4th **Topic** : DNA repair

If you put your hand in the sun for 30 mins , every cell in your skin could accumulate up to 50.000 to 70.000 mutation , put we repair them because we have a very efficient **DNA repair .** Every type of DNA damage repaired by a specific combination of an enzymes , put all the damages share the same steps for repairing :

1-recognition .
2-damaged area will be removed by endo- and exo- nucleases .
3-gap is filled by DNA polymerase .
4-joins of two fragment by ligase .

Then the doctor talked about **Benzopyrine as a mutagen** , (you can assume this part of the sheet as a separated part , it's not directly related to the DNA repair) .

Benzopyrine is one of the components of the cigarettes , in the structure of this molecule there is a five rings , which have a very high affinity to the **Guanine (G Nucleotide) ,** when it binds to the G , it will separated the G from the C nucleotide (that is coupled with it on the other strand) . The Benzopyrine stays in this place until the cell replication. When the cell wants to divide , the **DNA polymerase** may be prevented from seeing the **G that is hidden behind the mutagen ,** so it might confuse the(**G** nucleotide) , and replicate it as an (**A** nucleotide for example) , converting these part of the sequence from **GC pair** to **AT pair ,** the final result is a **SNP (Point mutation).**

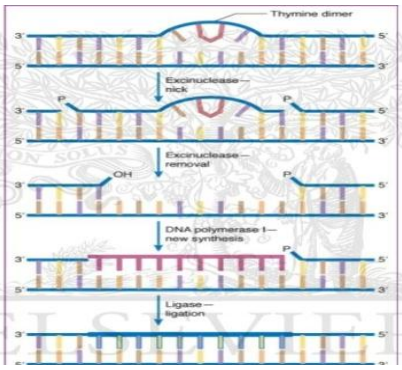
Returning back to our topic about the **DNA repair ,** we have four types of the DNA repair :

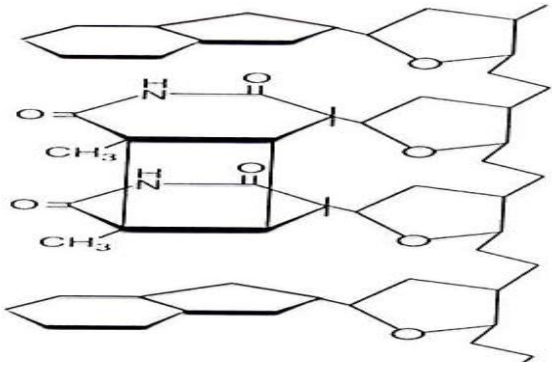
1-**Nucleotide excision repair (NER) :** these type of repair applied on the **big problems ,** that cause distortion in the double helix structure of the **DNA .**

In this case there is a distortion in the double helix , for example , the benzopyrine separates the (G) from the coupled(C) , causing a hump in the DNA structure , such a large distortion needs this type of repair (NER) , by removing the whole region around the damaged part (**NOT only the damaged one)** , then the action of **Polymerase** and **Ligase** will takes a place , completing the repair process .

Thymine dimer : one of the causes of the distortion in the DNA double helix structure , and it's the most important problem that can affect us when we exposed to the UV light (sun) .

The mechanism of the Thymine dimer : it might happens when there is a two (T) nucleotide beside each other , and both of them bind to an (A) nucleotide on the other strand . When we exposed to the sun light , the UV will cause a **cross linking** , the two neighbor (T) will bind to each other covalently , and also will cause separation of this (T)s from the coupled (A)s , forming a **Hump ,** but there is an enzymes that can resolve this problem frequently . (**first figure** : the repairing of a problem cause by thymine dimer (nucleotide excision repair) / **second figure** : the structure of thymine dimer.





2-**Base excision repair (BER) :** there is **No distortion in the DNA ,** it's only responsible of repairing a small problems , like removing a methyl group from an (eg:A nucleotide) , the enzymes will **only remove the damaged nucleotide** not the whole area and put another one there.

3-**Mismatch repair :** there is no structural problem , but the cell noticed that there is a **mismatch** problem , like having a (G) nucleotide coupled with an (A) nucleotide . but , **the cell will remove (A) and put (C) or it will remove the (G) and put (T) ?? In bacteria** , when the DNA formed , there will be a process called DNA methylation , so any new nucletiotide that is not methylated is a new nucleotide not original one , and the enzymes always remove the new one **. But in the human we don't know how the cell choose the wrong nucleotide !! it's still a problem to the scientists (NO methylation) .**

4-**Transcription coupled repair ,** this type is more complicated , the problem here is that there is an (A-T) base pair , but there must be a (G-C) base pair instead . These problem will be repaired in the **transcriptional level** , when the RNA Polymerase start coping the DNA , it will noticed that there is a problem according to its memory , so it will call the repair enzymes to go there and do their job ((it's really strange thing , because the RNA polymerase is working just like the copier of the papers , so it's unusual that you want to copy a paper , and the machine refuse to copy it because there is a spelling problem in the paper)) , so it's still not fully discovered by the scientists .

Done ..

I hope that the sheet was helpful to you and sorry for any mistakes =)