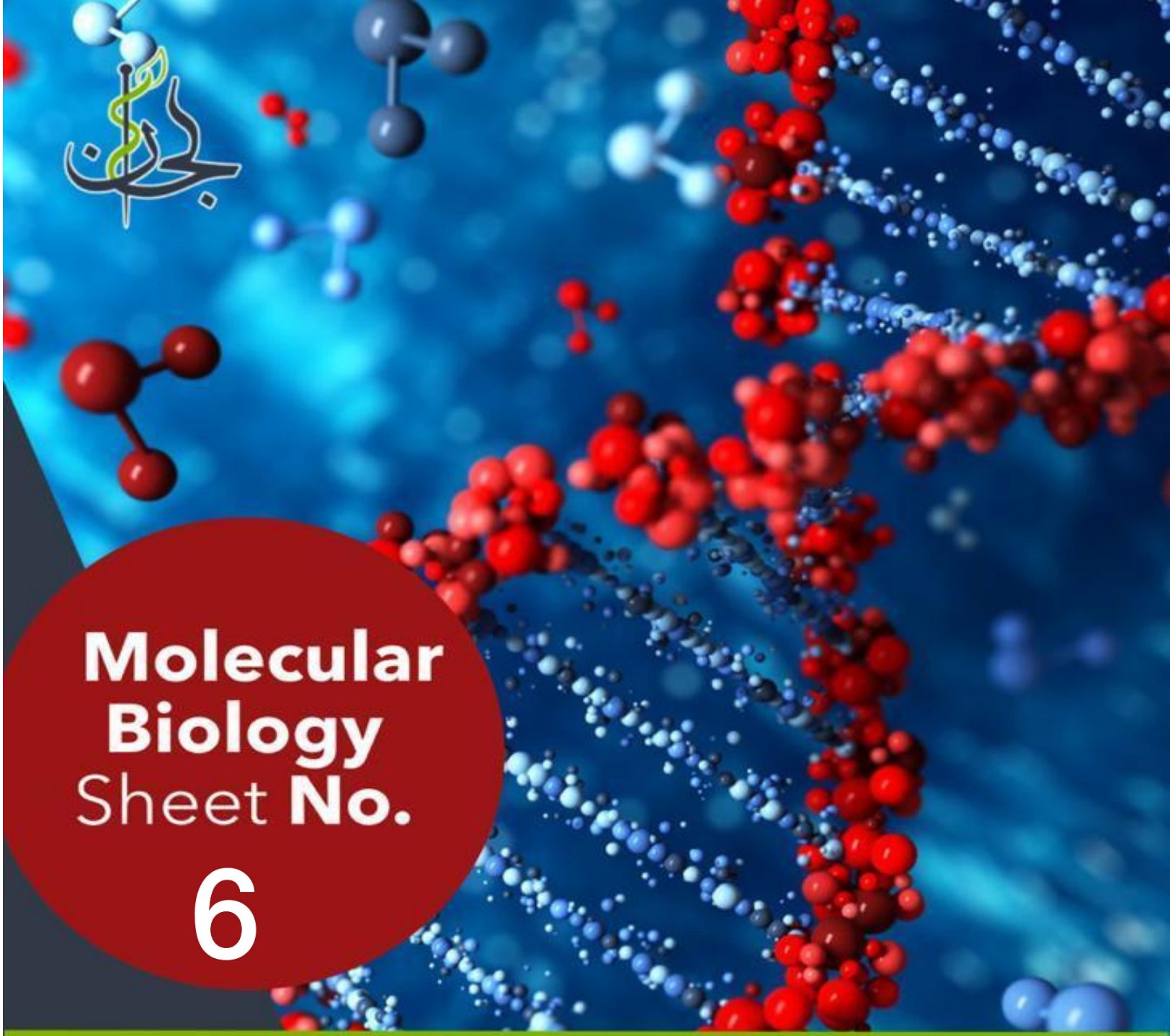


Molecular Biology Sheet No. 6

Writer Aya kotkot

Scientific correction Obada AL-limoon

Grammatical correction Obada Al-limoon

Doctor Dr.Maamoun Ahram

`

FIRST OF ALL , LETS TALK ABOUT THE PROJECT THAT WAS DONE TO GET INFORMATION RELATED TO HUMAN GENES

The human genome project

A $3 billion, 13-year, multi-national project launched in 1990, completed in 2004 and led by the US to sequence the human genome and map and identify the genes (a draft was published in 2001).

Major outcomes:

*IT WAS ESIMATED THAT THE HUMAN GENESAT FIRST NUMBERED 100.00 AND THEN DECREASED 50,000 BUT THEY KEEP DECREASING IN NUMBER ,AT THE END OF THIS PROJECT THE KNEW THAT THE NUMBER OF HUMAN GENES IS 22,300

THEY ALSO REALIZE THAT THE CODING REGION 2%(THE REGION THAT CODES PROTEINS MORE SPICIFCLY IT ENCODED TRANSCRIPTION FACTORS) REPRESENT AND THEY SAY THE 98% IS FOR CODE AND QOUT JUNK DNA

*MAJOR ROLE OF CONTROL REGIONS SUCH AS ENHANCERS WE WILL BE TALKING ABOUT LATER ON

*COMPLETING SEQUENCES OF OTHER GENOMES LIKE MITOCHONDRIA, DIFFIRENT TYPES OF BACTERIA ,YEAST ,DROSOPHILA THAT IS FRUIT FLY,MICE AND CHRIMPS.

*DEVELOPMENT OF MAJOR TECHNOLOGY,THAT IS WHY THEY WERE ABLE TO DO COMPLETING SEQUENCES

*OPEN DISCUSSION OF LEGAL AND ETHUCAQL ISSUES, THAT MEANS THAT THEY COULD TELL IF A PERSON HAS A MUTATION OR NOT ANOTHER BENEFIT IS PATENTING SO NO ONE ELSE CAN BENEFIT COMMERCIALLY.

THEY ALSO ESTIMATED THE HUMAN GENOME AND THAT IS COMPOSED OF ABOUT 3.2 BILLION BP ,BUT IT IS ACTUALLY 2.9 BILLION BP<AND SO FAR WE KNOW THAT WE HAVE (20,200-20,800)GENES,AND THEY ARE DISRIBUTED IN 46 CHROMOSOMES

`

SPECIES	BASE PAIRS (estimated)	GENES (estimated)	CHROMOSOMES
Human *(Homo sapiens)*	3.2 billion	~ 25,000	46
Mouse *(Mus musculus)*	2.6 billion	~ 25,000	40
Fruit Fly *(Drosophilia melanogaster)*	137 million	13,000	8
Roundworm *(Caenorhabditis elegans)*	97 million	19,000	12
Yeast *(Saccharomyces cerevisia)*	12.1 million	6,000	32
Bacteria *(Escherichia coli)*	4.6 million	3,200	1
Bacteria *(H. influenzae)*	1.8 million	1,700	1

DEAR VAGUS : U R NOT REQUIRED TO MEMORIZE THEM BUT NOTCE THE DIFFERENCE BETWEEN HUMAN AND ROUNDWORM ,AND HOW MOUSE IS CLOSE .ALSO ABOUT THE HUMAN GENOME.

Nucleotides per genomes

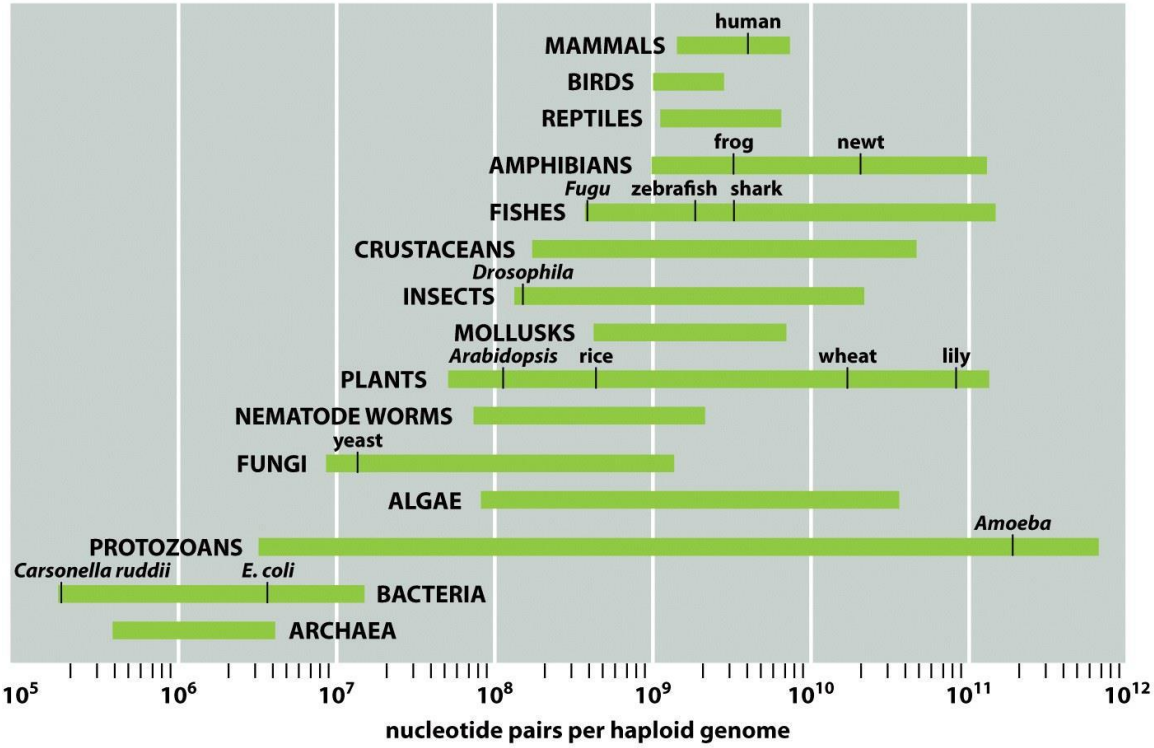
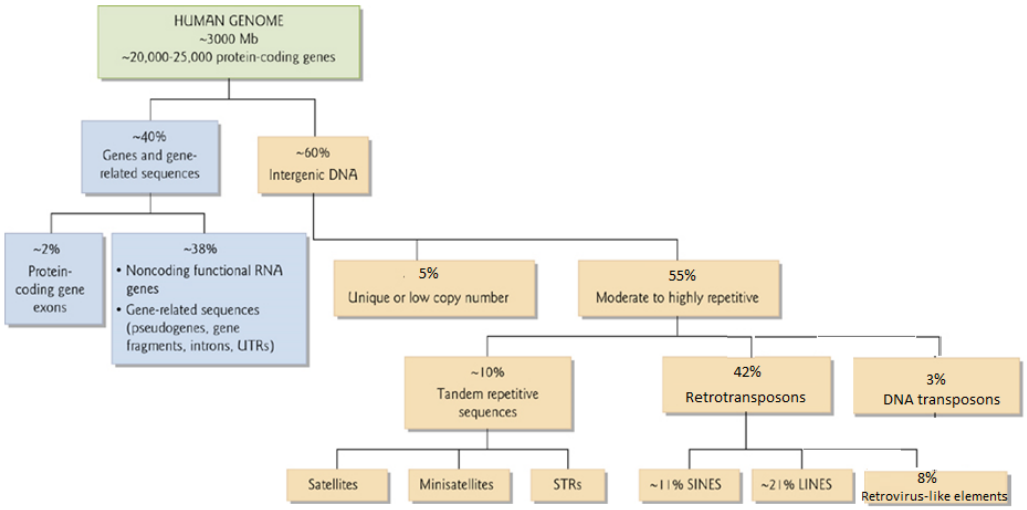


Figure 1-41 Essential Cell Biology 3/e (© Garland Science 2010)

DO NOT MEMORIZE THEM ,IT IS JUST A NUCLEOTIDES PER GENOME.



PLEASE GO BACK TO SLIDE 6

Components of the human genome

*40%=GENES AND GENE-RELATED SEQUENCE

*60%=INTERGENIC DNA(THEY ARE DNA SEQUENCES LOCATED BETWEEN GENES

THE 40% IS DEVIDED TO:

VID ABOUT EXONS AND INTRONS

2%=CODING FOR mRNA TO MAKE PROTEINS (THEY ARE MAINLY EXONS)

38%=NON-CODING (MAINLY INTRONS)(THESE 38% ARE 35% INTRONS AND THE REMAINING ARE PSEUDOGENES,GENEFRAGMENTS,UTRS)

NOTE:INTRON IS NUCLEOTIDE SEQUENSE WITHIN A GENE THAT IS REMOVED BY RNA SPLICING DURING MATURATION OF THE FINAL RNA PRODUCT

%60 IS DEVIDED TO

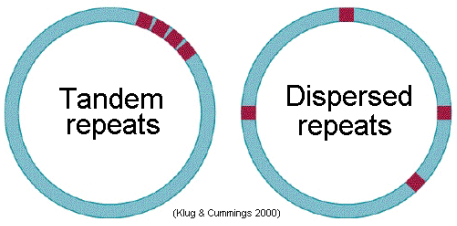
5% LOW COPY NO. ~5% of the genome contains sequences of noncoding DNA that are highly conserved indicating that they are critical to survival.

55% MODERATE HIGHLY REBETITVE THIS MEANS(SAME SEQUENCE BUT REPEATED IN DIFFERENT PLACES OR IN THE SAME PLACE)

THIS 55% IS DEVIDED TO

10%=TANDEM REPETITIVE TANDEM

TANDEM BASICALY MEANS SAME REPEAT BUT THEY



`

ARE VERY CLOSE TO EACH OTHER OR IT IS THE ONE

AFTER ANOTHER, SO THEY ARE CONSEQUENT TO EACH OTHER

AND THESE TANDEM ARE DEVIDED TO

1. SATELLITES
2. MINI SATELLITES
3. MICROSATELLITES OR STR OR SHORT TANDEM REPEATS

42%=RETROTRANSPOSONS WHICH ARE DEVIDED TO 3 THINGS WE WILL TALK ABOUT THEM.

3%=DNA TRANSPOSONS

NOTE: BOTH ARE CONSIDRED INTERSPERSED IT MEANS THAT THEIR LOCATION ARE DISPERSED NOT ONE AFTER ANOTHER LIKE TANDEM .

I WANT U TO PUT THIS IN UR MIND THAT WHENEVER U HEAR RETRO IT MEANS THAT THE SOURCE OF THE SEQUENCE IS VIRUS.

*PAY ATTENTION THAT THE NUMBERS ARE APPROXIMATED SO THE DOC IS NOT GONNA TELL THE SAME NUMBER BUT CLOSE TO IT, U DON'T HAVE TO KNOW WHY CUZ IT IS NON OF UR BUSINESS.

ANOTHER PROJECT IS DONE TO UNDERSTAND THE DIFFERENT COMPONENETS OF THE HUMAN GENOME, IT IS CALLED THE ENCODE PROJECT)

IT IS BASICALY TO KNOW ABOUT THE ELEMNTS OF DNA (COMPONENTS OR REGION OF DNA).AS A CONCLUSION OF THIS PROJECT IS THAT THE KNEW

*80% OF THE ENTIRE HUMAN GENOME IS RELEVANT
(either transcribed, binds to regulatory proteins, or is associated with some other biochemical activity).

`

The ENCODE project (2003-on)

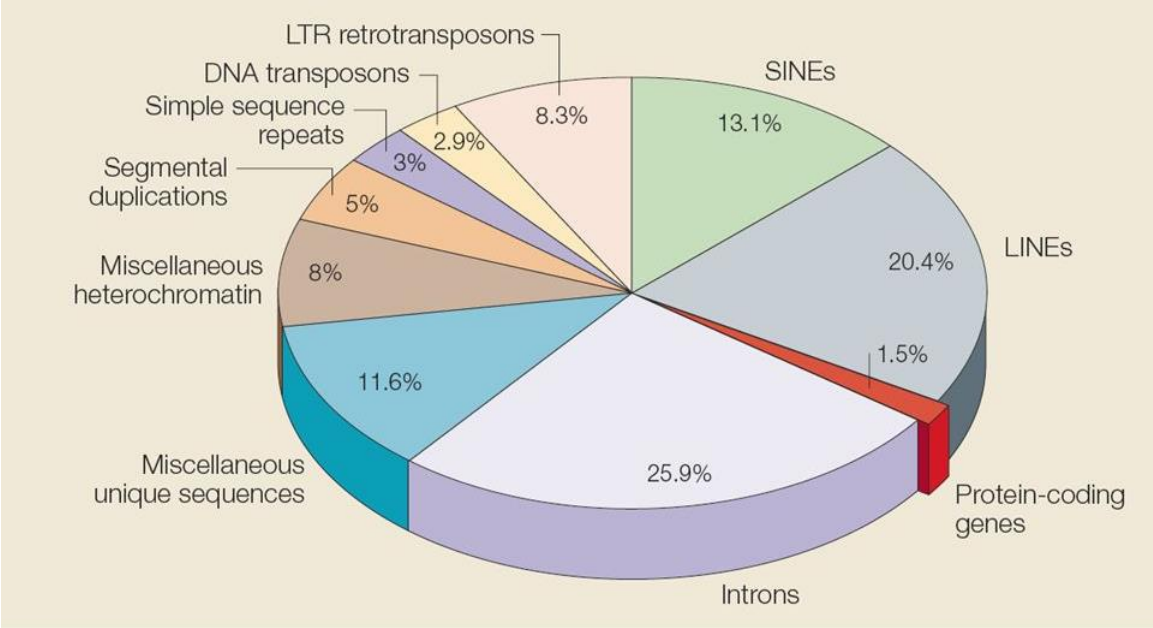
IN THE FIRST PROJECT THEY THOUGHT THAT ONLY 2% IS BEING USED FOR TRANSCRIPTION, BUT HERE IT IS 75 TO 80 PERCEN IS USED FOR TRANSCRIPTION

Summary of ENCODE Results	
Protein-coding genes	20,687
Short noncoding RNAs	8801
Long noncoding RNAs	
Pseudogenes	11,224
Percentage of genome transcribed into RNA	74.7%
Percentage of genome-binding transcription factors	8.1%

PSEUDOGENES: THEY LOOK LIKE GENES BUT THEY ARE NOT AND THEY CAN BE HOMOLOGOUS TO OTHER GENES,THEY ARE DUPLICATED BUT CANNOT ENCODE PROTEINS

EX: GLOBIN WHICH IS RESPONSIBLE FOR MAKING HEMOGLOBIN AND MYOGLOBIN.

PI CHART SHOWING THE DIFFERENT COMPONENTS OF THE HUMAN GENOM



*EVERYTHING IN THIS FIGURE WILL BE EXPLAINED IN DETAILES

`

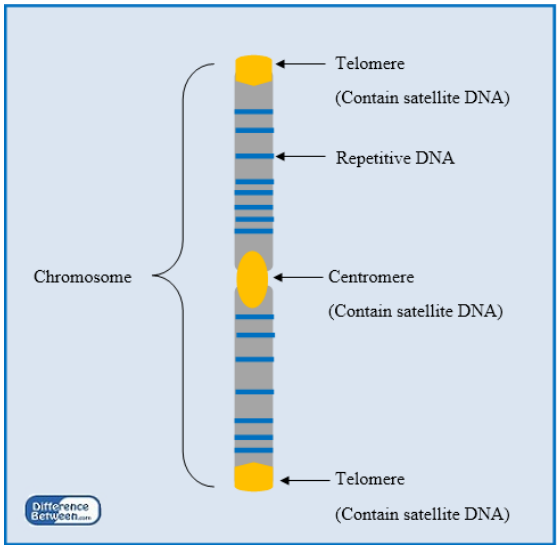
NOW LET US BEGIN WITH TANDEM REPEATS

*SATELLITES
Regions of 5-300 bp repeated 106-107times (10% of genome)

Centromeric A/T-rich repeats (171 bp) called α-satellite unique to each chromosome (you make chromosome-specific probes) by **fluorescence in situ hybridization**.

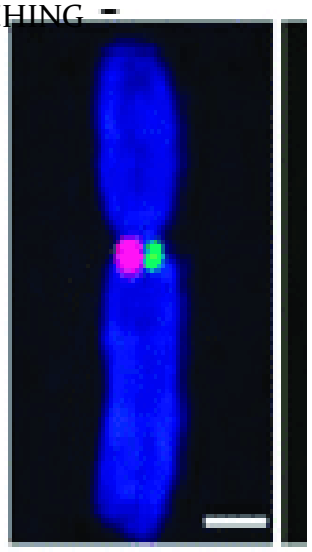
SITU: MEANS IN PLACE, SO **fluorescence in situ hybridization** MEANS THAT U R STAINING A CERTAIN THING IN ITS PLACE USING FLUORESCNCE METHODS

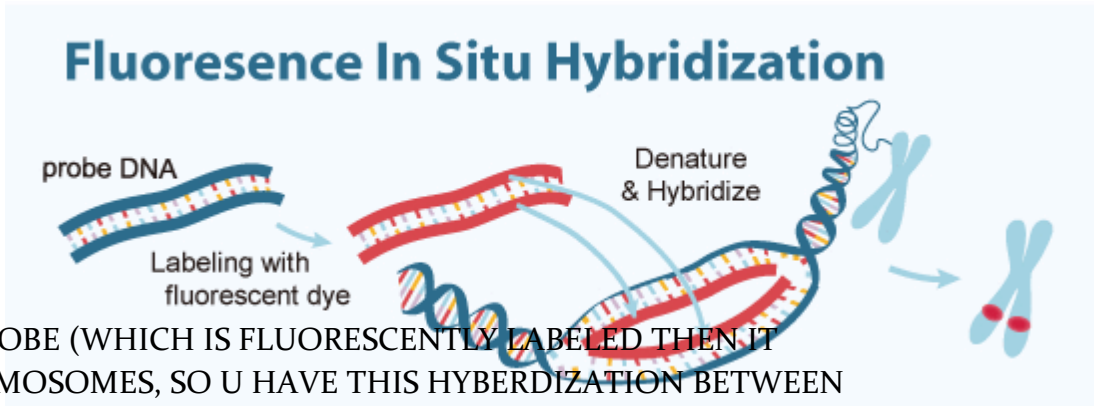
RICH IN A AND T : IMPORTANT FOR SEPARATION OF CHROMOSOMES ,AND PULLING THEM TO EITHER A POL OF THE CELL IN MITOSIS, BY ATTACHING TO A LONG FILAMENTS



RIGHT HERE IS CHROMOSOME

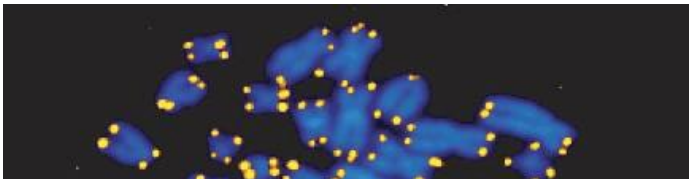
AND FLUOROSCENTLY LABELED





WE USE A DNA PROBE (WHICH IS FLUORESCENTLY LABELED THEN IT TARGETS A CHROMOSOMES, SO U HAVE THIS HYBERDIZATION BETWEEN THE PROBE AND CHROMOSOME

NOTE: THIS TEQHNIUQ IS VERY COMMON FOR STAINING CHROMOSOMES



`

1 THESE ARE CALLED TELOMERES AND THEY EXIST AT THE END OF THE OF
1 CHROMOSOMES

*TELO MEANS END

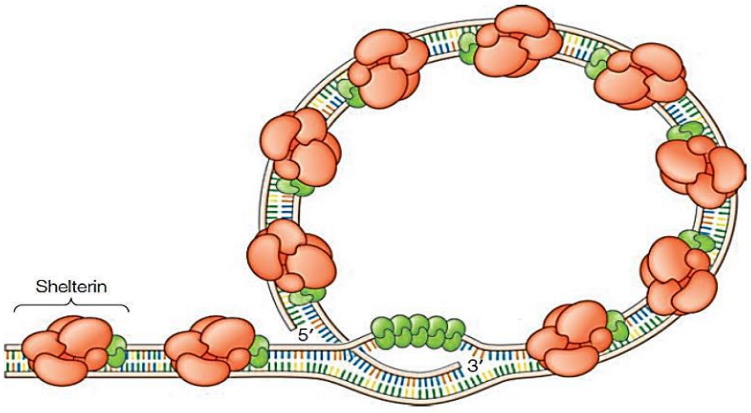
1 THE FIGURE SHOWS DNA WITH LEABELED TELOMERES.

(TTAGGGT) is repeated hundreds to thousands of times at the termini of human chromosomes with a 3′ overhang of single-stranded DNA.

The repeated sequences form loops that bind a protein complex called **shelterin** , which protects the chromosome termini from degradation and it shelters chromosomes.

*REMEMBE WHEN WE SAID THAT THE TELOMERES ARE TIED UP

SO AT THE END OF THE CHROMOSOME IS BEING TIED UP WITH THE HELP OF PROTEIN CALLED SHELTERIN.



IT IS KNOWN THAT TELOMERES AND CENTROMERES CANNOT BE TRANSCRIPED , BUT WITH THIS DEVELOPMENT OF TECHNIQUES WE FOUN OUT THAT THEY COULD BE TRANSCRIPED.

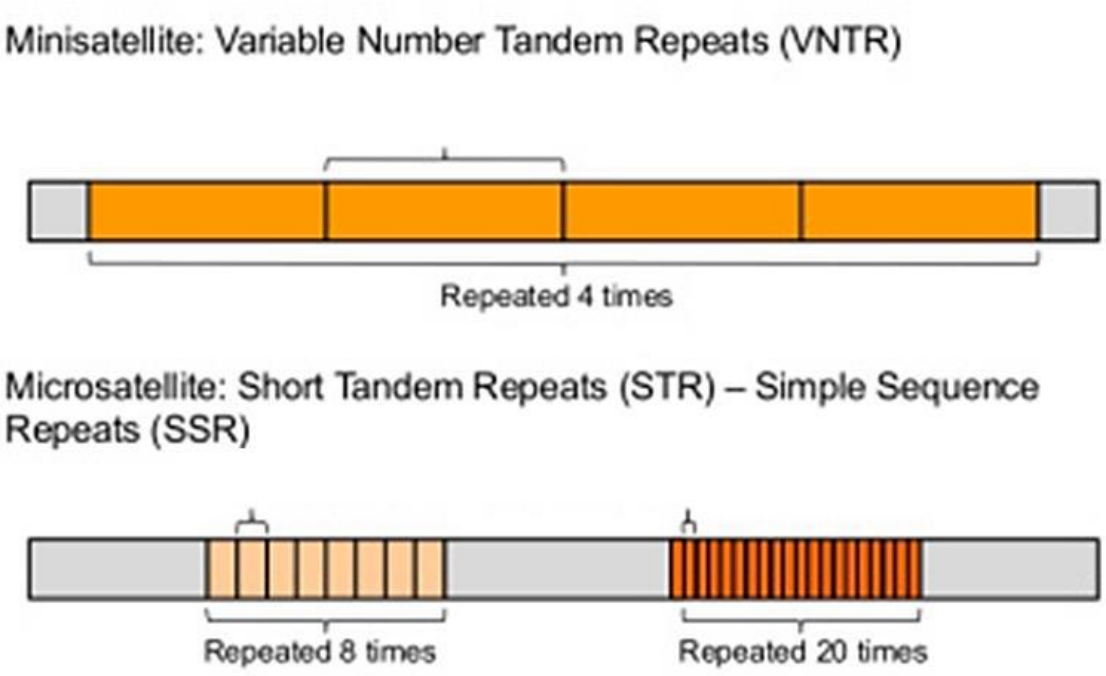
NOW LET'S GO A LITTLE BIT DEEPER IN TELOMERES COMPONENTS WE WILL SEE

Telomeric repeat-containing RNA(TERRA):BELONGS TO THE CLASS OF long non-coding RNA transcribed from telomeres and functions in:

- maintaining the integrity of chromosome termini, LIKE SHELTERIN
- regulating telomerase activity,
- maintaining the heterochromatic state of telomeres, REMEMBER THA HETEROCHROMATIN ARE HIGHLY PACKED
- protecting DNA from deterioration or fusion(GENETIC ABNORMALITY) with neighboring chromosomes

THE LAST TYPES OF TANDEM IS (MINI AND MICRO DNA)

HUGHLY REPEATED BUT DIFFERENT IN SIZE

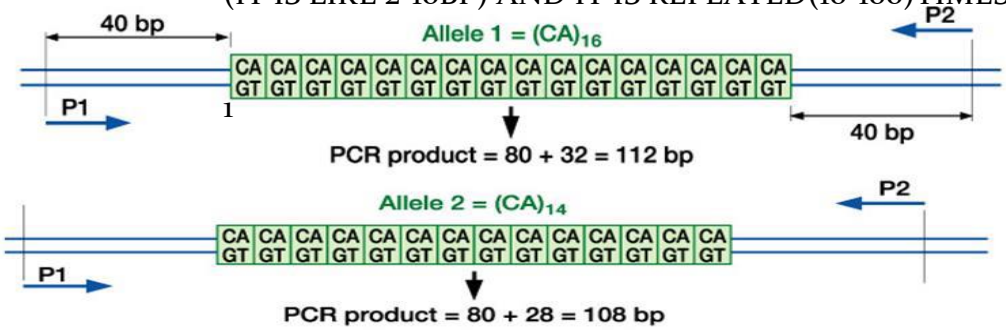


THE FIRST ONE IS TANDEMLY REPEATED, ONE AFTER THE OTHER

(IT IS LIKE 20-100BP)AND IT IS REPEATD(20-50)TIMES

THE SECOND ONE, THE BAIS PAIRS ARE TANDEMLY REPEATED

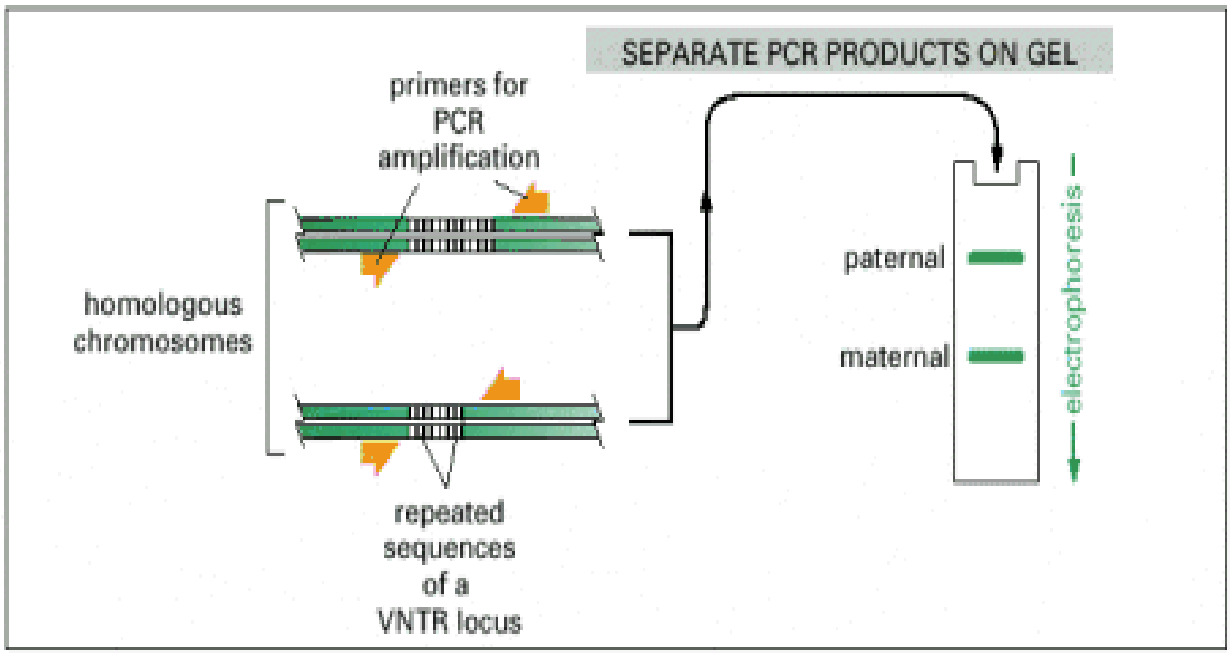
(IT IS LIKE 2-10BP) AND IT IS REPEATED(10-100)TIMES



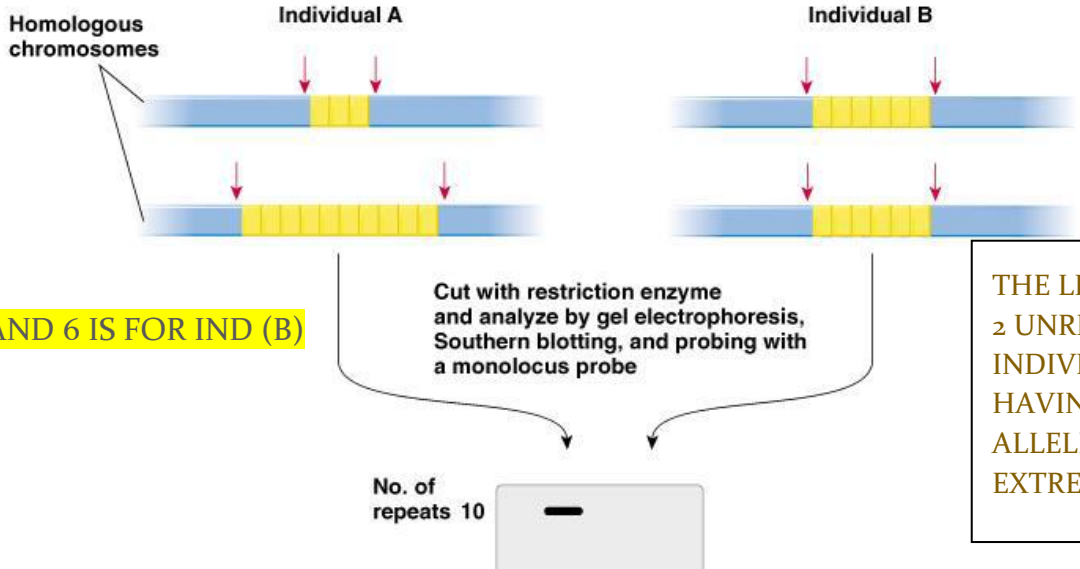
* HOMOLOGOUS CHROMOSOMES

WE HAVE A LOT OF VARIATION IN OUR CHROMOSOMES AND THIS DUE TO DIFFERENT REPETITION IN TH BP,SINCE WE ARE DIPLOID IT IS (CA).

POLYMORPHISMS :WHICH IS A DIFFERENT SHAPES OF DNA SEQUENCES in homologous chromosomes



*DUE TO POLYMORPHOSIS, WE HAVE DIFFERENT SEQUENCES OF VNTR AND STR IN HOMOLOGOUS CHROMOSOMES, AND THIS HAVE BEEN HILPFUL IN FORENSIC TESTING,CUZ WE CAN DIIFERENTIAT SAMPLES OF INDIVIDUAL FROM EACH OTHER BASED ON THIS POLYMORPHISM.



BAND 6 IS FOR IND (B)

THE LIKELIHOOD OF 2 UNRELATED INDIVIDUALS HAVING THE SAME ALLELIC PATTERN IS EXTREY IMPOSSIBLE

`

IN THE PREVIOUS FIGUIRE U R LOOKING AT STRs OR VNTRs

INDIVIDUAL A

CHROMOSOME HAS 3 REPEATS

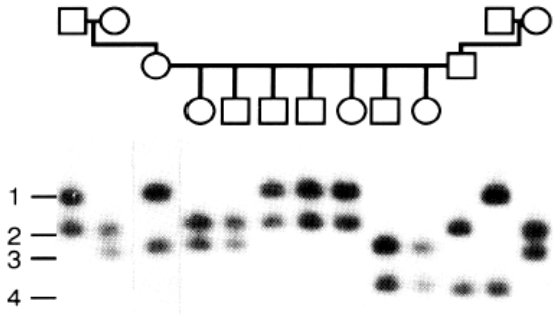
CHROMOSOME HAS 10 REPEATS

INDIVIDUAL B

BOTH CHROMOSOMES HAVE(6)

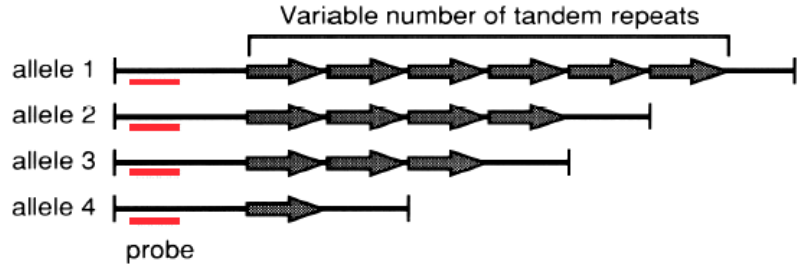
THE EXACT SAME NO. OF REPEATS

THIS CALLED PEDIGREE



IT IS A REAL EXAMPLE

single-locus probe but multiple alleles



Thompson & Thompson Genetics in Medicine, p. 130, 1991

WE CAN USE VNTR OR STR FOR

- PATERNITY TESTING
- FORENSIC MEDICINE
- DIAGNOSTIC MEDICINE

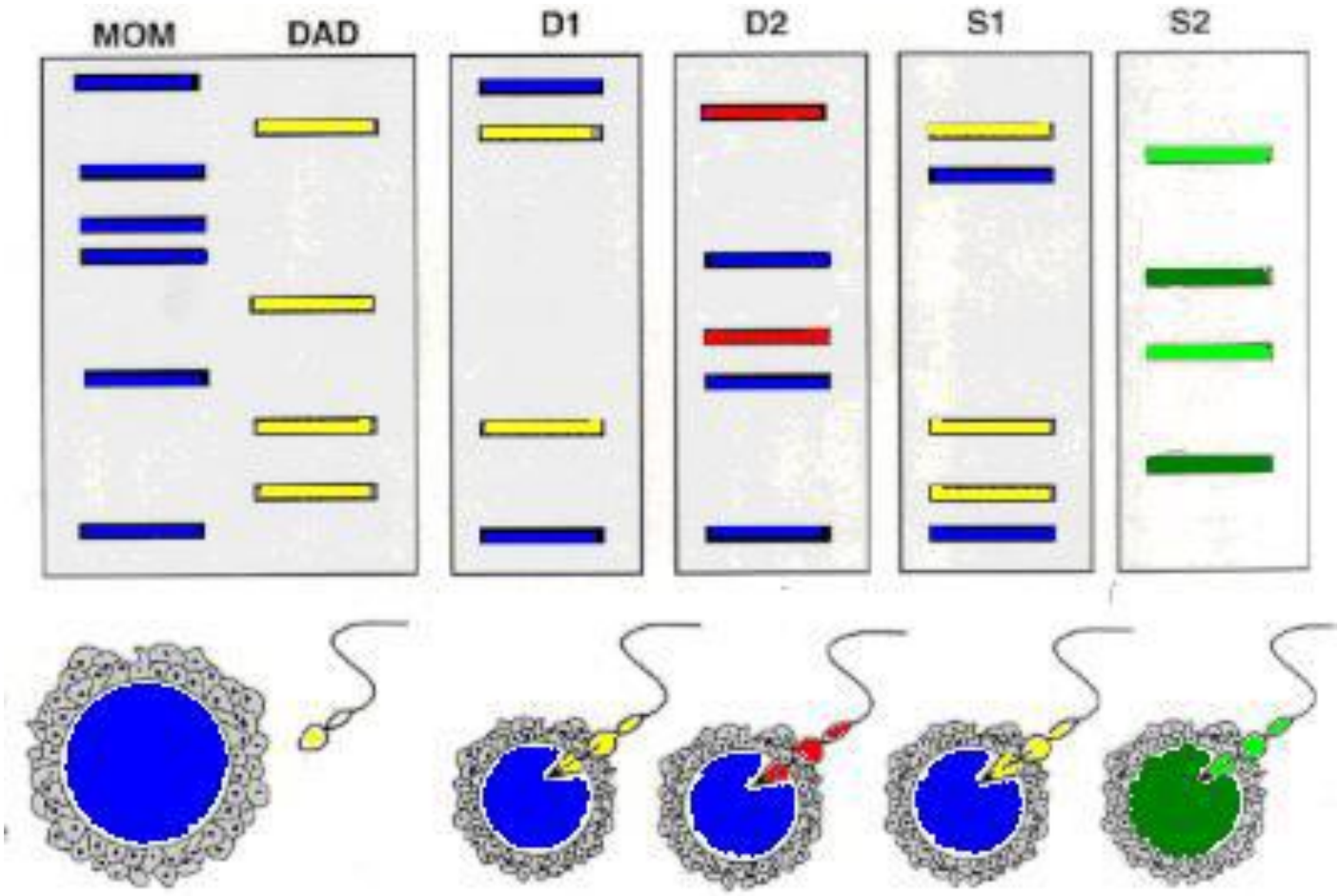
SO ABOVE WE HAVE AN EXAMPLE OF PATERNITY TESTING
THE NUMBERED CIRCLES REPRESENT THE ALLELES IN THE PERSON ABOVE,
ON THE LEFT SIDE OF THE PEDIGREE
*WE HAVE GRANDFATHER THAT HAS AN ALLELE 1 AND ALLELE 2
*GRANDMOTHER HAS ALLELE 2 AND ALLELE 3
*SO THE DAUGHTER WILL HAVE ONE FROM THE GRANDFATHER(HER FATHER) WHICH IS ALLELE 1 AND OTHER FROM THE GRANDMOTHER(HER MOTHER) WHICH IS ALLELE 3
NOTE: يا رب ما اكون فوتكم بالحيط

`

ON THE RIGHT SIDE OF THE PEDIGREE

*IN THE SAME WAY, THE FIRST MALE WILL HAVE ONE FROM HIS FATHER (ALLELE 4)AND THE OTHER FROM HIS MOTHER(ALLELE 2)

SO IN THIS WAY U CAN FIND OUT THE PATERNITY FROM THE ALLELES BY LOOKING NOT ONLY ON ONE STR/VNTR WE RATHER LOOK AT MULTIPLE STRs/VNTRs.



HERE IS ANOTHER EXAMPLE, NO NEED TO EXPLAIN HOW, IT IS

A PIECE OF CAKE

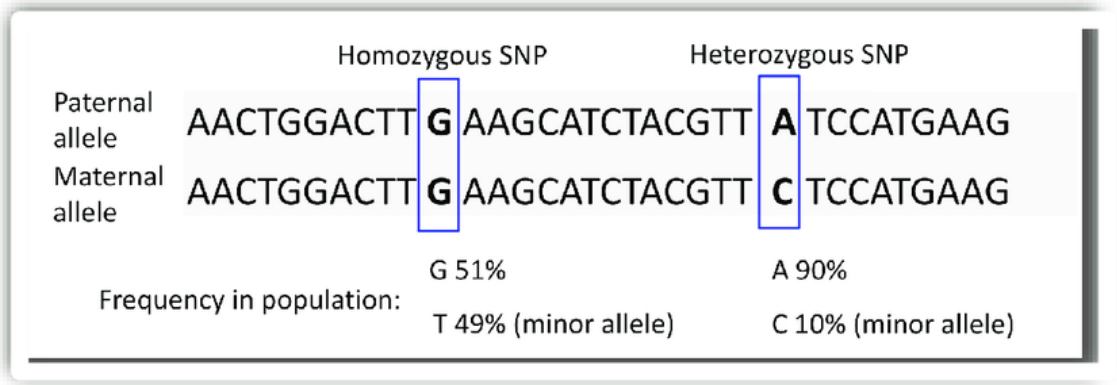
`

WHEN DOES A CHANGE CAN BE CALLED MUTATION OR POLYMORPHISM!!

Two or more versions of a sequence must each be present in at least one percent of the general population WE CALL THAT POLYMORPHISM
IF IT IS LESS THAN ONE PERCENT WE CALL IT MUTATION

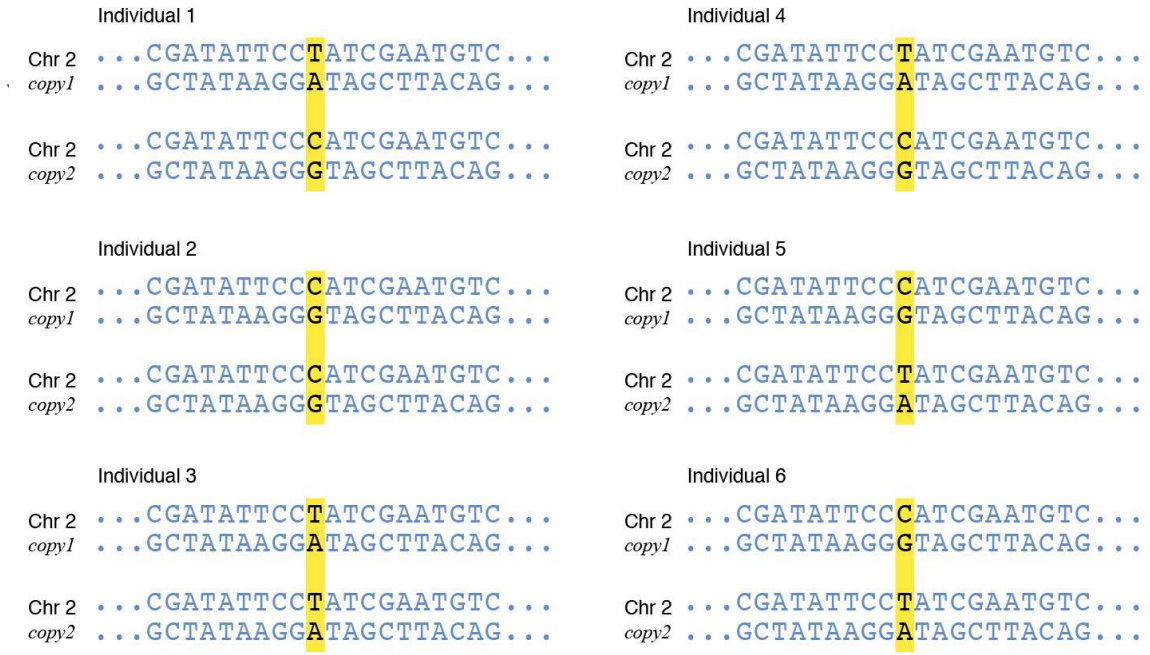
SNPs occur throughout the human genome -about one in every 300 nucleotide base pairs.
*~10 million SNPs within the 3-billion-nucleotide human genome

Only 500,000 SNPs are thought to be relevant

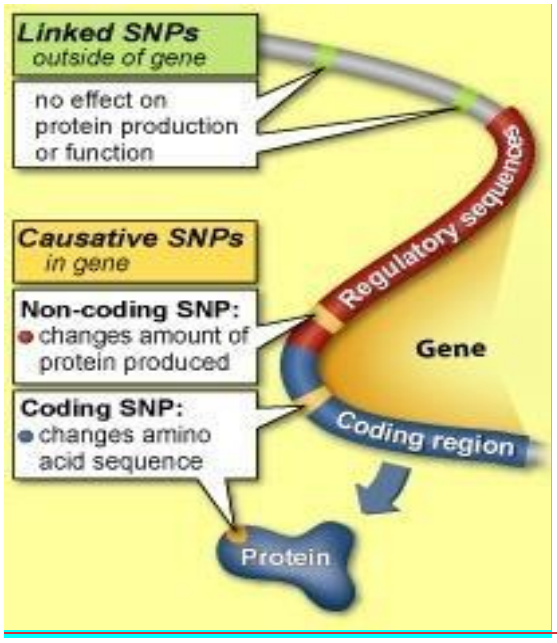


IN THE FIRST PARTICULAR SNP THEY ARE HOMOZYGOUS, THE SECOND SNP IS HETEROZYGOUS.

ACCORDING TO THE PERCENTAGES OF POPULATION FOR THE NITROGENUOS BASES U COULD TELL WHICH ONE IS MINOR ALLELE (LESS POPULATION)



6 DIFF INDIVIDUALS WITH DIFF SNPs(BUT MAINLY THEY ARE T OR C)

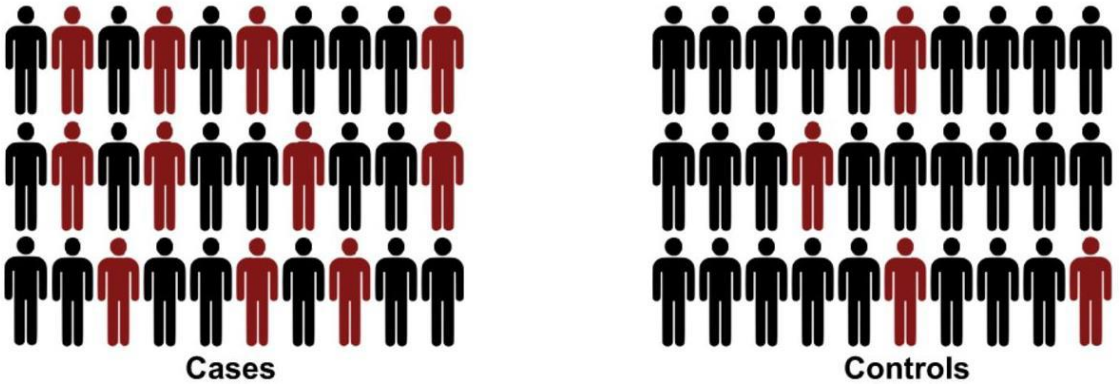


CAUSATIVE: CAUSES PHNOTYPE

PHENOTYPE U CAN SEE LIKE SKIN COLOR OR,IT APEARES IN CERTAIN CONDITIO EX: 2 DIFF PERSONS HAVING DIFF RESPONSE FOR THE SAME DOSE OF (PARACETAMOL FOR EXAMPLE).LOOK AT THE FIGUIR TO KNOW WHERE DOES CAUSATIVE SNP LOCATED AND WHAT THEY DO ON THEIR LOCATIONS

NOTE: CODING SNP—ON PROTEIN---MAYBE LESS EFFICINCE THAN OTHERS THAT DON NOT HAVE IT

`



TTGGCCAGCTGGACGAGGGCGATGAC

TTGGCCAGCTGGATGAGGGCGATGAC

CHANGING THE C FOR T DOES NOT CHANGE ANY PHENOTYPE ,IF SOMENOE HAS DISEASE HE STILL HAVE IT AFTER CHANGING

JUST A PROPABILITY THAT HIGH FREQUENCY OF THOSE WHO HAVE DISEASE HAVE Ts , AND ALSO IT IS NOT FOR SURE THAT HE HAS DISEASE

LINKED MEANS THA IT IS LINKED (BEING SO CLOSE)TO A GENE AS A RESULT IF THIS WILL BE INHERITED THE SNP ISINHERITED ALONG WITH IT

ITERSPERSED REPEATS

LAST BUT NO LEAST WE CAN TALK ABOUT INTERSPERSED(MAJORITY OF HUMAN GENOME)THEY ARE ALSO KNOWN AS TRANSPOSONS(CALLED JUMPING GENES CUZ IT CAN MOVE FROM ONE PLACE TO ANOTHER)

***MAJORITY OF TRANSPOSONES DO NOT MOVE,BUT SOME CAN LIKE IN MICE BUT NOT IN HUMAN**

`

TWO TYPES OF TRANSPOSONS

- **DNA transposons (3% of human genome)**
- **RNA transposons or retrotransposons (42% of human genome).**

*REMEMBER : I ASKED TO PUT IN UR HEAD THAR WHENEVER U SEE RETEROTRANSPOSONS,THEY COME FROM VIRUSES

*Long interspersed elements (LINEs, 21%)

*Short interspersed elements (SINEs, 13%)
An example of SINEs is Alu (300 bp)
Retrovirus-like elements (8%)

- Over 99% of the transposons in the human genome lost their ability to move, but we still have some active transposable elements that can sometimes cause disease.

Hemophilia A and B (BLOOD DISEASE THAT RESULTS IN EXCESSIVE BLEEDING), severe combined immunodeficiency, porphyria (RELATED TO HEME METABOLISM), predisposition to cancer, and Duchenne muscular dystrophy.

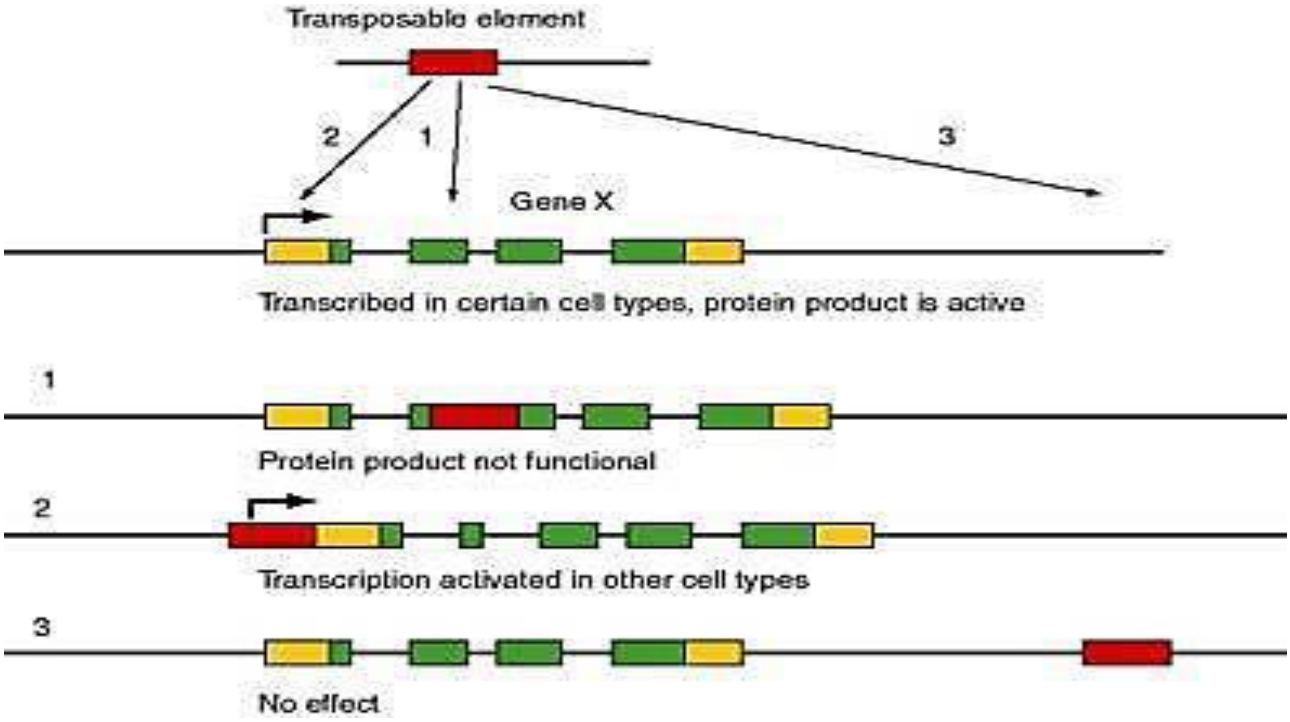
THE QUESTION IS WHY DO THEY CAUSE DISEASES!!

- TRANSPOSON CAN INSERTS ITSELF WITHIN A GENE SO IT DISRUPTS THE GENE.AS A RESULT THE GENE IS NOT FUNCTIONING ANYMORE AND IF IT DOES THE PROTEIN WILL NOT FINCTION .

- IT CAN GET CLOSE TO A GENE AND MAKE THIS GENE CAN BE OVER EXPRESSED AND THEN PRODUCING MORE RNA AND PROTEINS

OR IT CAN JUST BE UNDEREXPRESSED

- IT CAN BE INSERTED FARAWAY FROM THE GENE SO THERE WOULD BE NO EFFECT



“God has a purpose for your pain, a reason for your struggle, and a reward for your faithfulness. Trust Him and don’t give up.”

ALL THE BEST, WISH ME LUCK

`

WHY PASS UR EXAM WHEN U CAN JUST PASS AWAY!

IT IS A JOKE

IT IS NOT