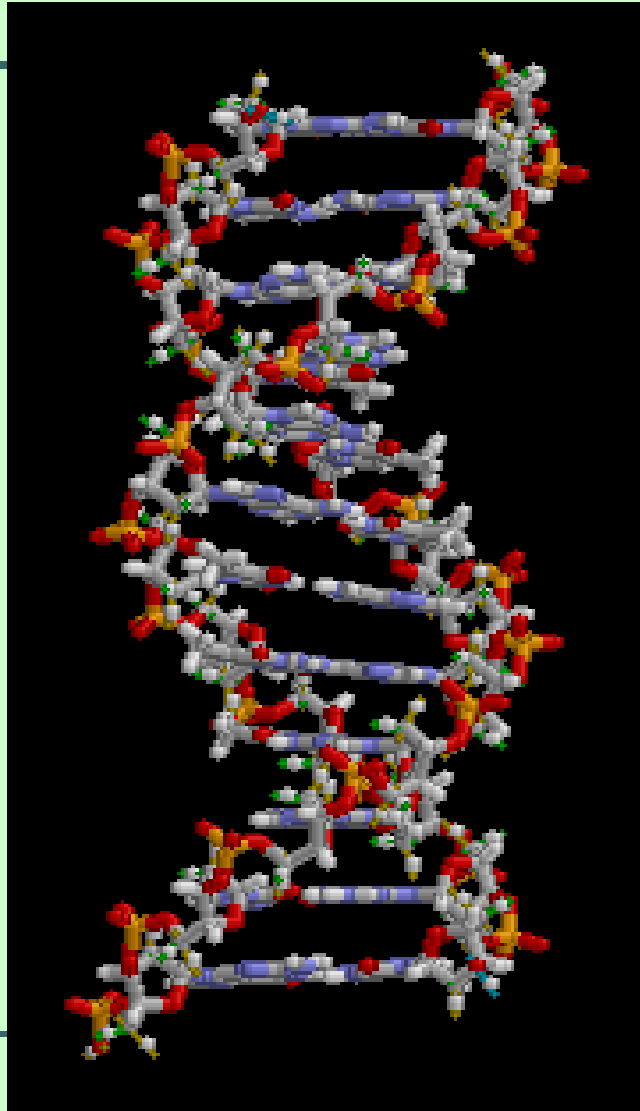


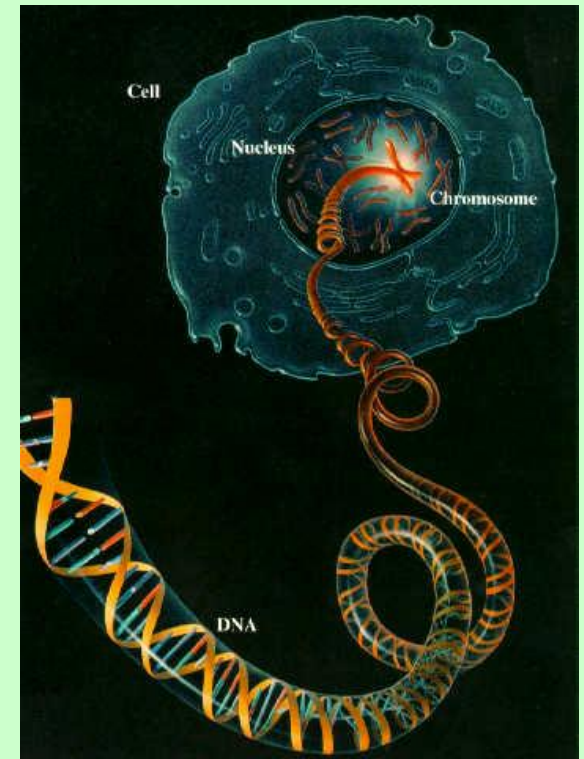
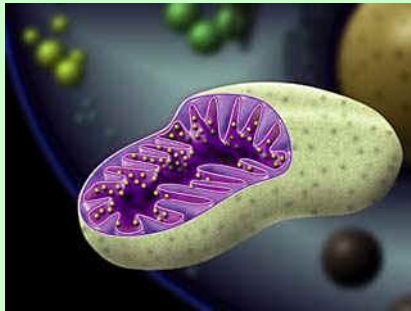
Molecular diagnostics



The human genome

the total genetic information (DNA content) in human cells

- nuclear
- mitochondrial - double-stranded DNA is organized into one circular molecule. Exclusively maternal inheritance

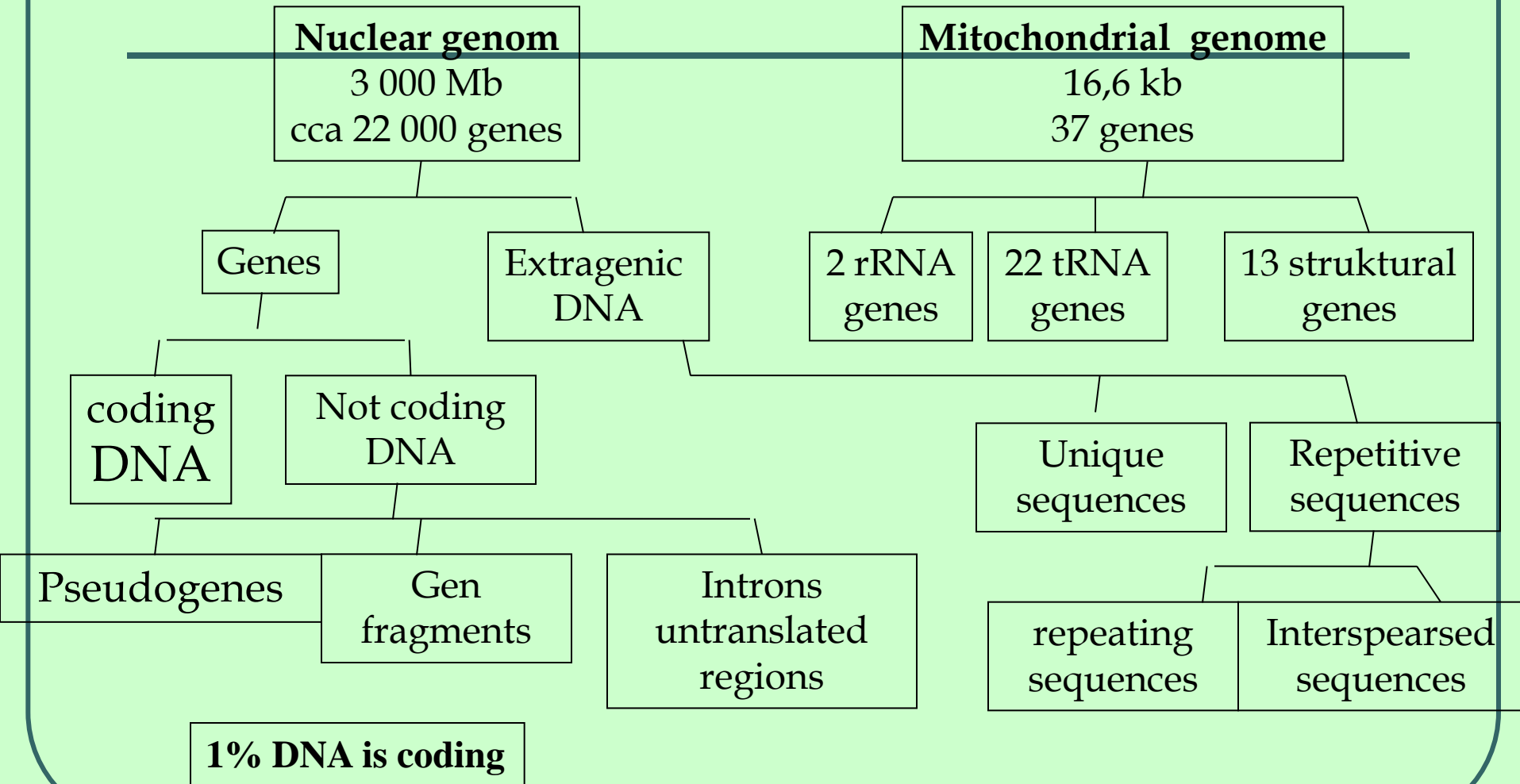


The human nuclear and mitochondrial genomes

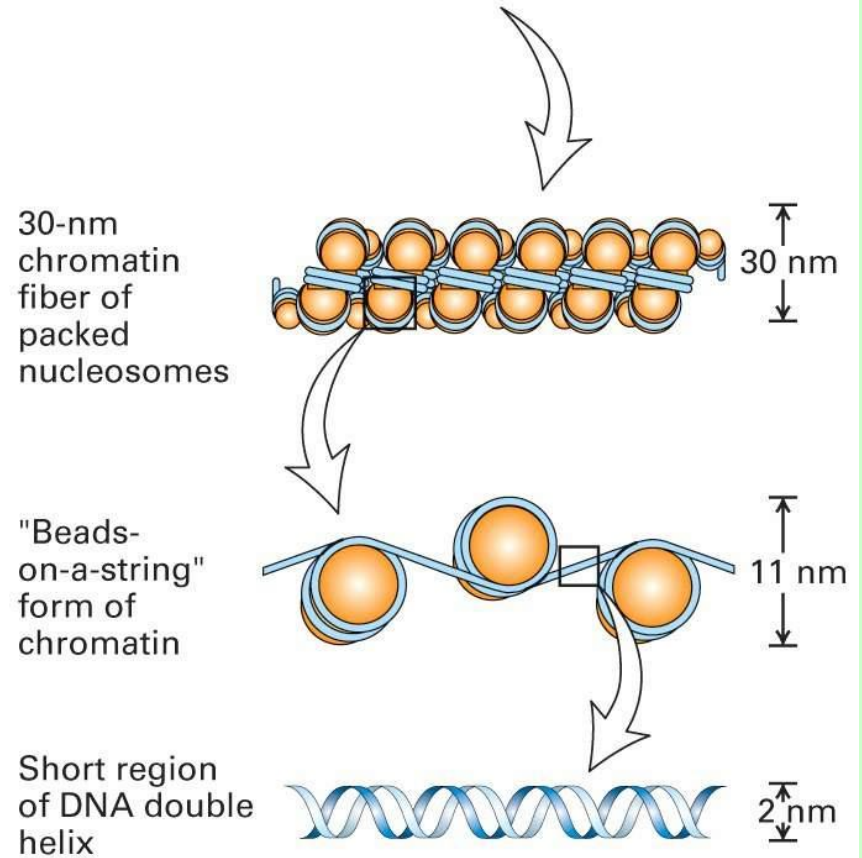
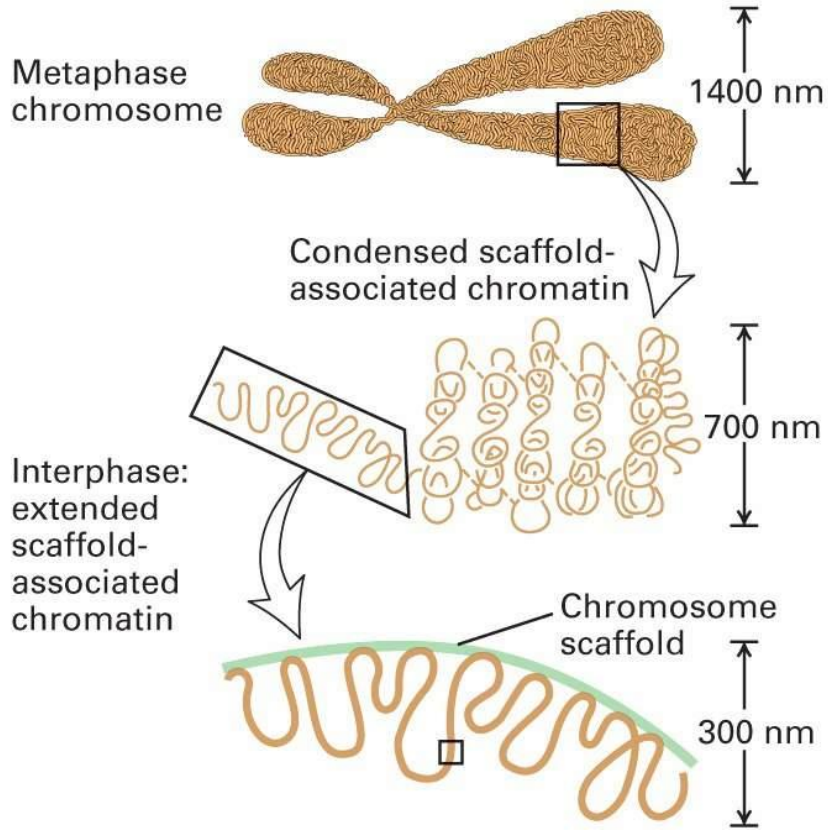
| | Nuclear genome | Mitochondrial genome |
|-------------------------------------|---|--|
| Size | 3000 Mb | 16.6 kb |
| No. of different DNA molecules | 23 (in XX) or 24 (in XY) cells, all linear | One circular DNA molecule |
| Total no. of DNA molecules per cell | 23 in haploid cells; 46 in diploid cells | Several $\times 10^3$ |
| Associated protein | Several classes of histone and nonhistone protein | Largely free of protein |
| No. of genes | 22 000 | 37 |
| Gene density | ~1/40 kb | 1/0.45 kb |
| Repetitive DNA | Large fraction, see <i>Figure 7.1</i> . | Very little |
| Transcription | The great bulk of genes are transcribed individually | Continuous transcription of multiple genes |
| Introns | Found in most genes | Absent |
| % of coding DNA | 1 – 2 % | ~93% |
| Codon usage | See <i>Figure 1.22</i> | See <i>Figure 1.22</i> |
| Recombination | At least once for each pair of homologs at meiosis | None |
| Inheritance | Mendelian for sequences on X and autosomes; paternal for sequences on Y | Exclusively maternal |

The human genome

It comprise two genomes:







Superstructure



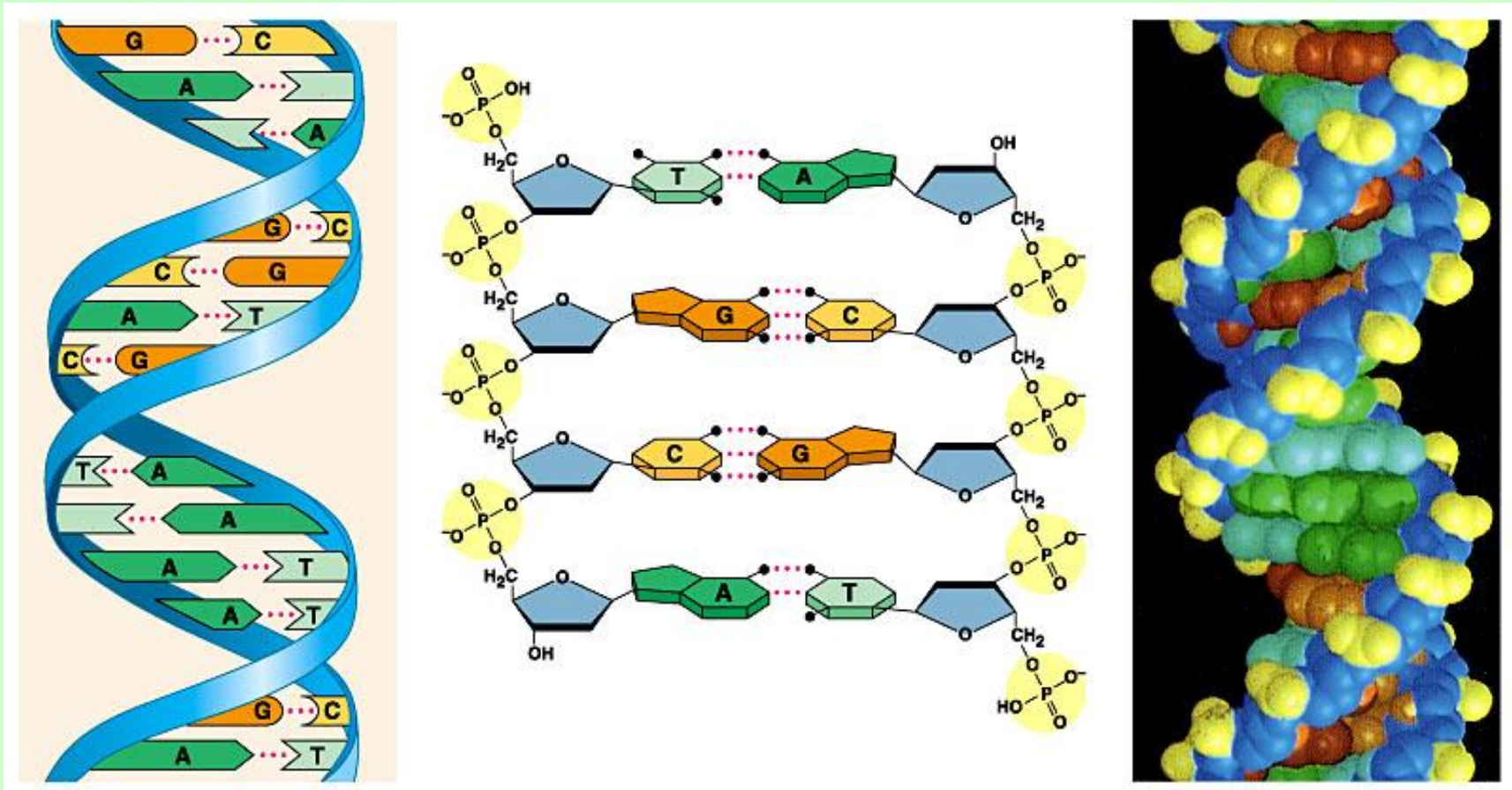


**Johann
Gregor Mendel**
published the results
of his investigations
of the inheritance of
"factors" in pea
plants

| | | pollen ♂ | |
|-------------|---|---|---|
| | | B | b |
| pistil ♀ | B |  BB |  Bb |
| | b |  Bb |  bb |

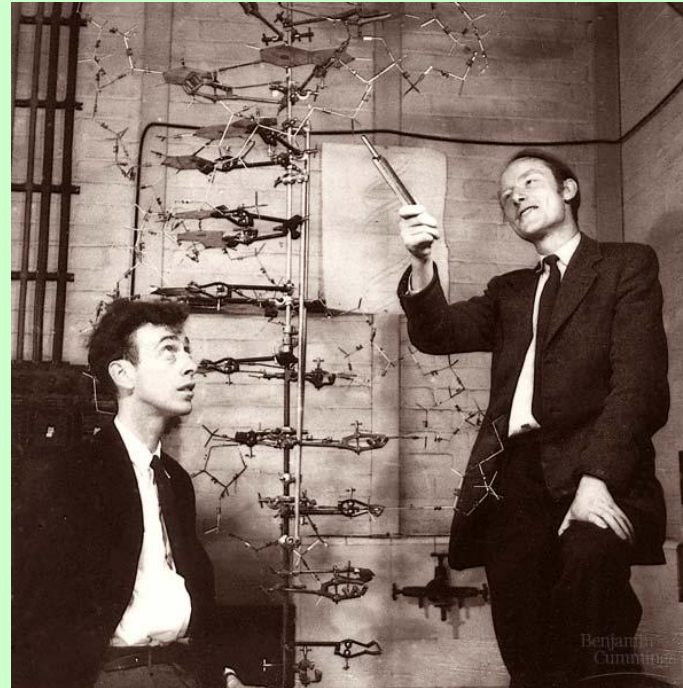
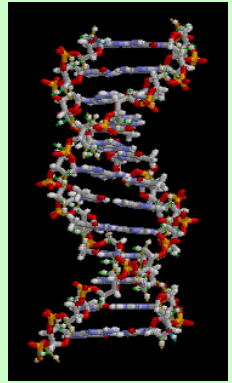
DNA - deoxyribonucleic acid

contains the complete genetic information that defines the structure and function of an organism. Proteins are formed using the genetic code of the DNA.

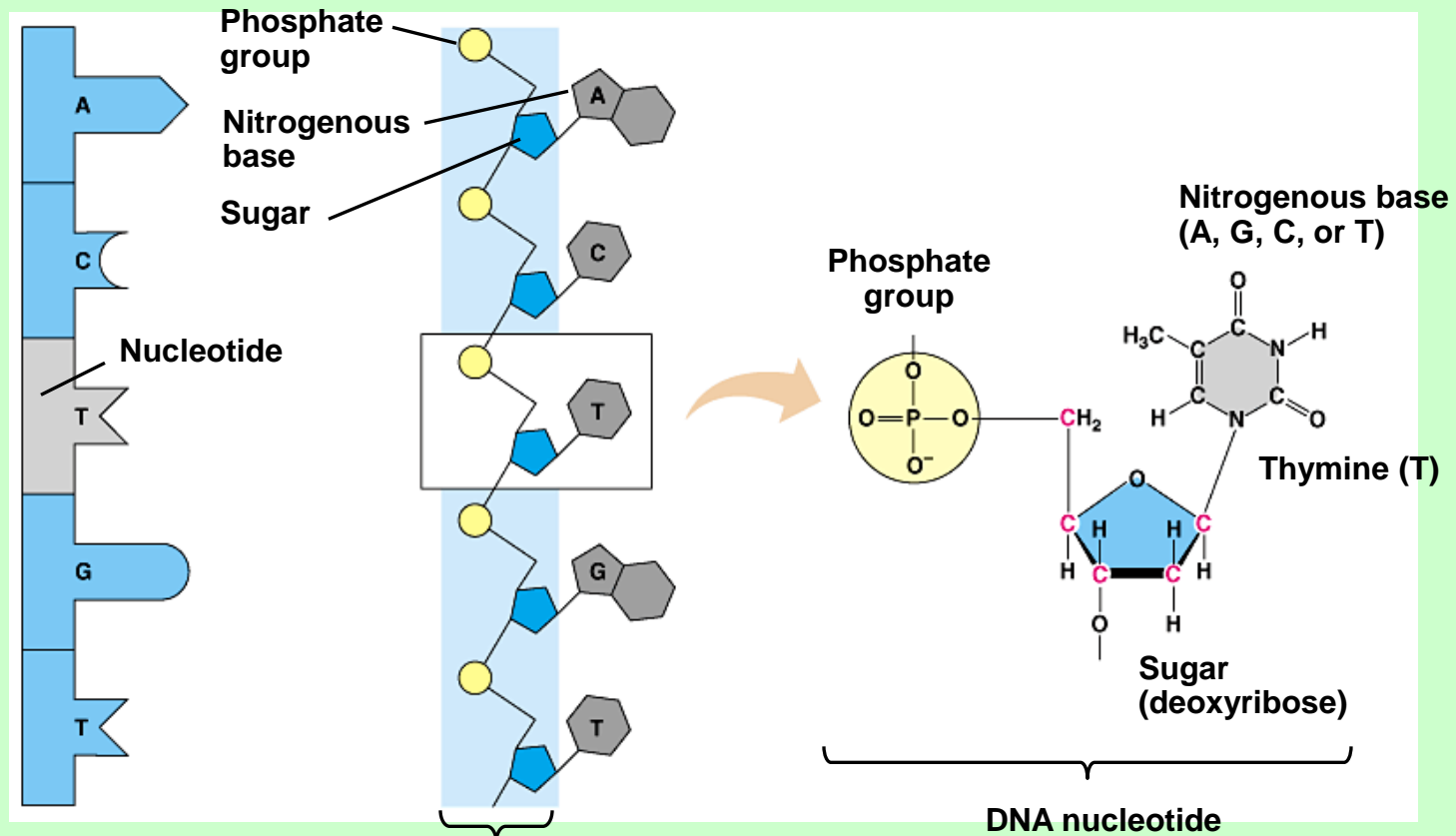


DNA is a double-stranded helix (1953)

James Watson and Francis Crick worked out the three-dimensional structure of DNA, based on work by Rosalind Franklin



DNA is a nucleic acid, made of long chains of nucleotides

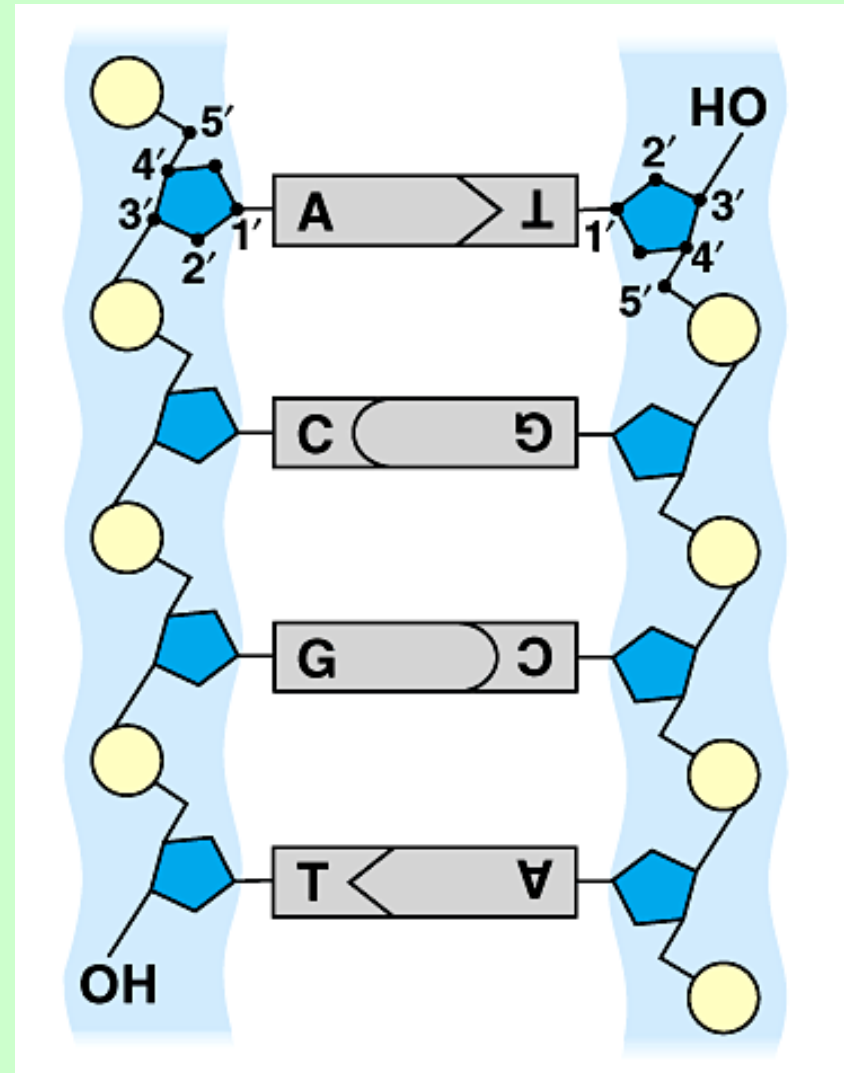


Polynucleotide

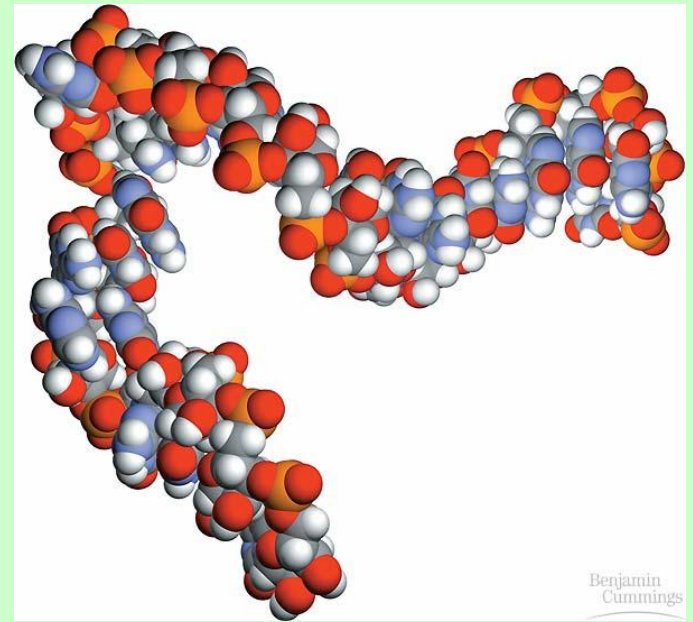
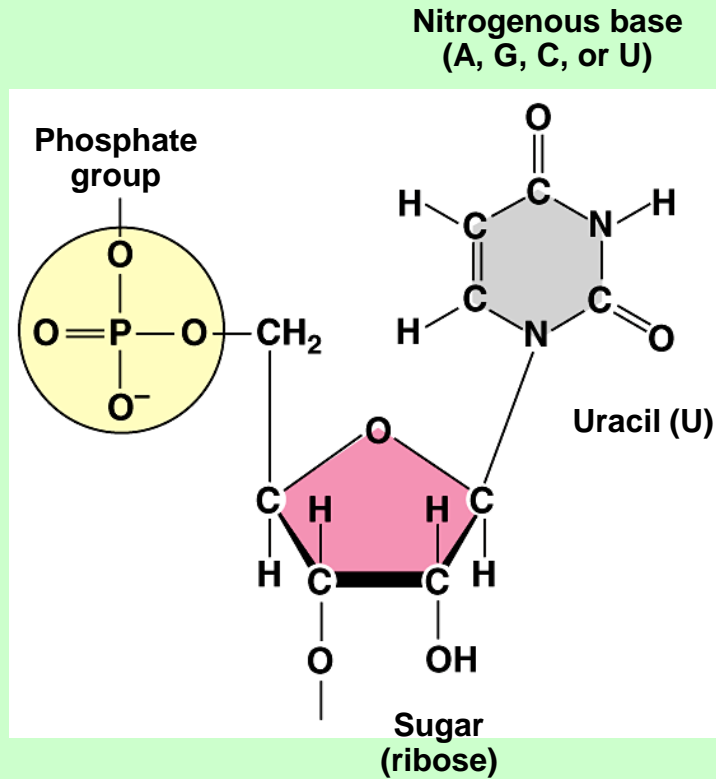
Sugar-phosphate backbone

DNA nucleotide

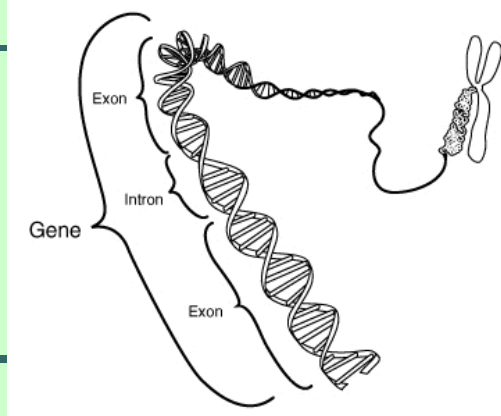
- These two strands run in opposite directions to each other and are therefore **anti-parallel**
- The DNA double helix is stabilized by **hydrogen bonds** between the bases attached to the two strands.
 - Each base pairs with a complementary partner
 - A pairs with T
 - G pairs with C
- The sugars are joined together by phosphate groups that form **phosphodiester bonds**



- RNA is also a nucleic acid
 - RNA has a slightly different sugar
 - RNA has U instead of T



Gene

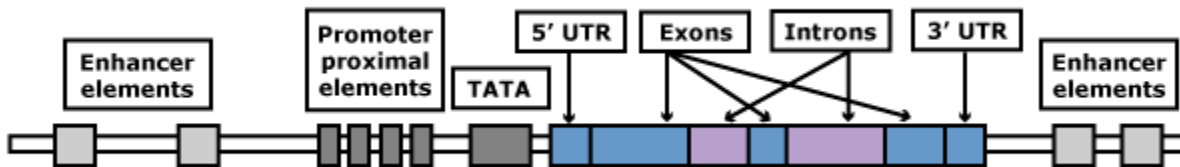


- is the basic unit of heredity in a living organism
- segment within a very long strand of DNA with specific instruction for the production of one specific protein
- contains both
 - „coding" sequences that determine what the gene does
 - „non-coding" sequences that determine when the gene is active (expressed)
- The 'gene coding region' (about 1.5 % of our DNA) codes for a polypeptide (around 25, 000 proteins).
- The non-coding region function remains unclear but can be as much as 5-45% of the total genome

Gene Structure

- **Exons:** Protein-coding DNA sequences of a gene
- **Introns:** Non-coding DNA sequence located in between exons

Transcription control regions of a gene, binding sites for transcription factors and RNA polymerase



One Gene One Polypeptide

- Theory- One gene is transcribed and translated to produce one polypeptide
- Some proteins are composed of a number of polypeptides and in this theory each polypeptide has its own gene.
- e.g. Hemoglobin is composed of 4 polypeptides (2 of each type) and there is a gene for each type of polypeptide.

One Gene One Polypeptide

This theory, like so many in biology has exceptions

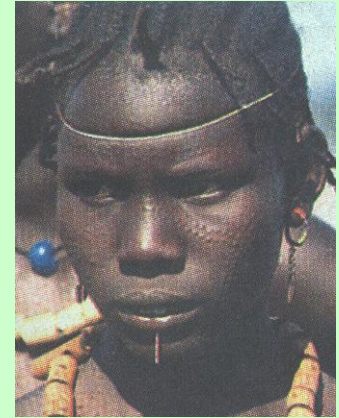
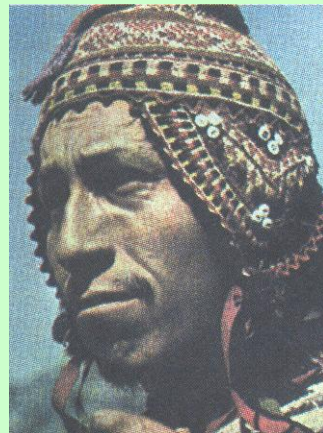
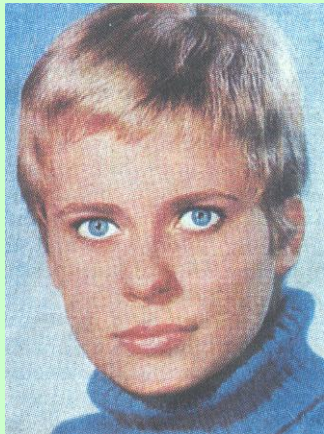
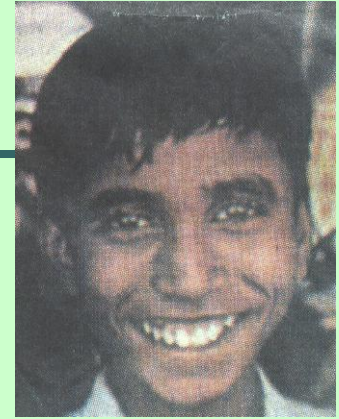
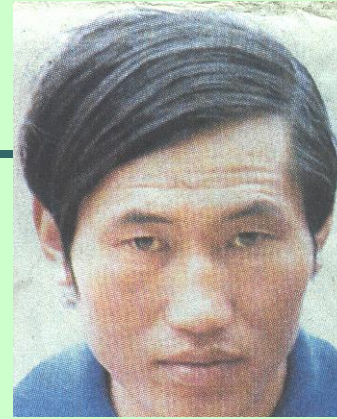
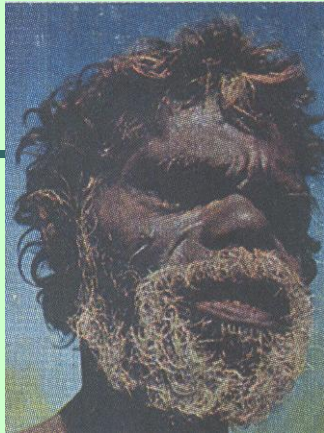
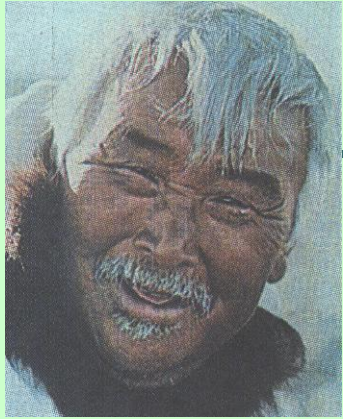
1. Some genes code for types of RNA which do not produce polypeptides
2. Some genes control the expression of other genes

DNA the Genetics Makeup



- Genes are inherited and are expressed
 - **genotype** (genetic makeup)
 - **phenotype** (physical expression)
- On the left, is the eye's phenotypes of green and black eye genes.

Genome x Genotype

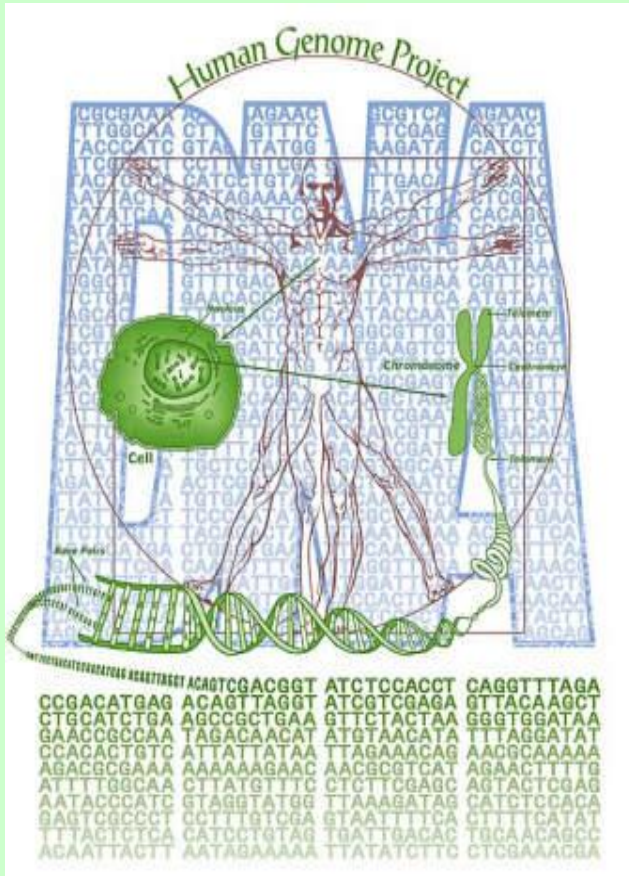


Individuals of the same species have the same genome.
Individuals of the same species have a different genotypes.

-
- **Dominant Gene:** When present this gene will express itself in the phenotype.
 - **Recessive Gene:** When present with a dominant gene this trait will not be shown in the phenotype

- **Allele**: different versions of the same gene. There are two alleles for every gene. One from the mother and one from the father. e.g. the gene for eye color, or blood type, etc.
- **Heterozygous**: A combination of genes for the same trait in the organism that are different from each other.
- **Homozygous**: A combination of genes for the same trait in the organism that are the same as each other.

Human genome project (HUGO)



- Identify all of the genes in human DNA
- Determine the sequence of the 3 billion chemical nucleotide bases that make up human DNA
- Store this information in data bases
- Develop faster, more efficient sequencing technologies
- Develop tools for data analysis
- Address the ethical, legal, and social issues (ELSI) that may arise from the project

Human genome project (HUGO)

- \$3-billion project founded in 1990 by the United States Department of Energy and the U.S. National Institutes of Health. The international consortium comprised also geneticists in the United Kingdom, France, Germany, Japan, China and India.



A parallel project was conducted outside of government by the **Celera Corporation**



-
- June 6, 2000, the HGP and Celera Genomics held a joint press conference to announce that TOGETHER they had completed ~97% of the human genome

Human genome project

Key findings of Genome Project:

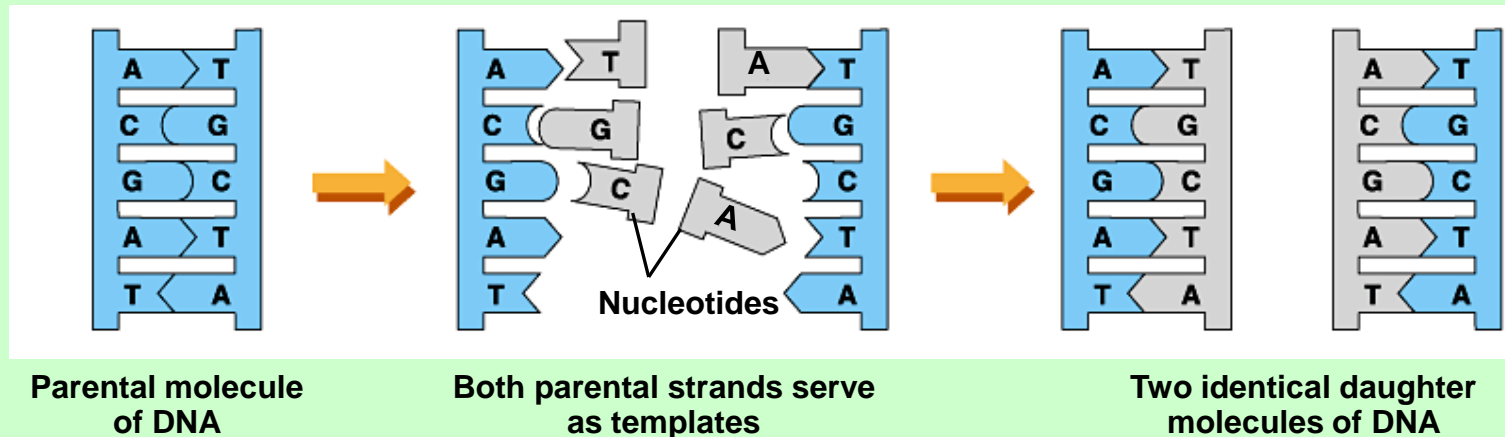
1. There are approx. 22,000 genes in human beings, the same range as in mice and twice that of roundworms.
Understanding how these genes express themselves will provide clues to how diseases are caused.
2. All human races are 99.99 % alike, so racial differences are genetically insignificant.
3. Most genetic mutation occurs in the male of the species and as such are agents of change.
They are also more likely to be responsible for genetic disorders.
4. Genomics has led to advances in genetic archaeology and has improved our understanding of how we evolved as humans and diverged from apes 25 million years ago.
It also tells how our body works, including the mystery behind how the sense of taste works.

Transfer of genetic information

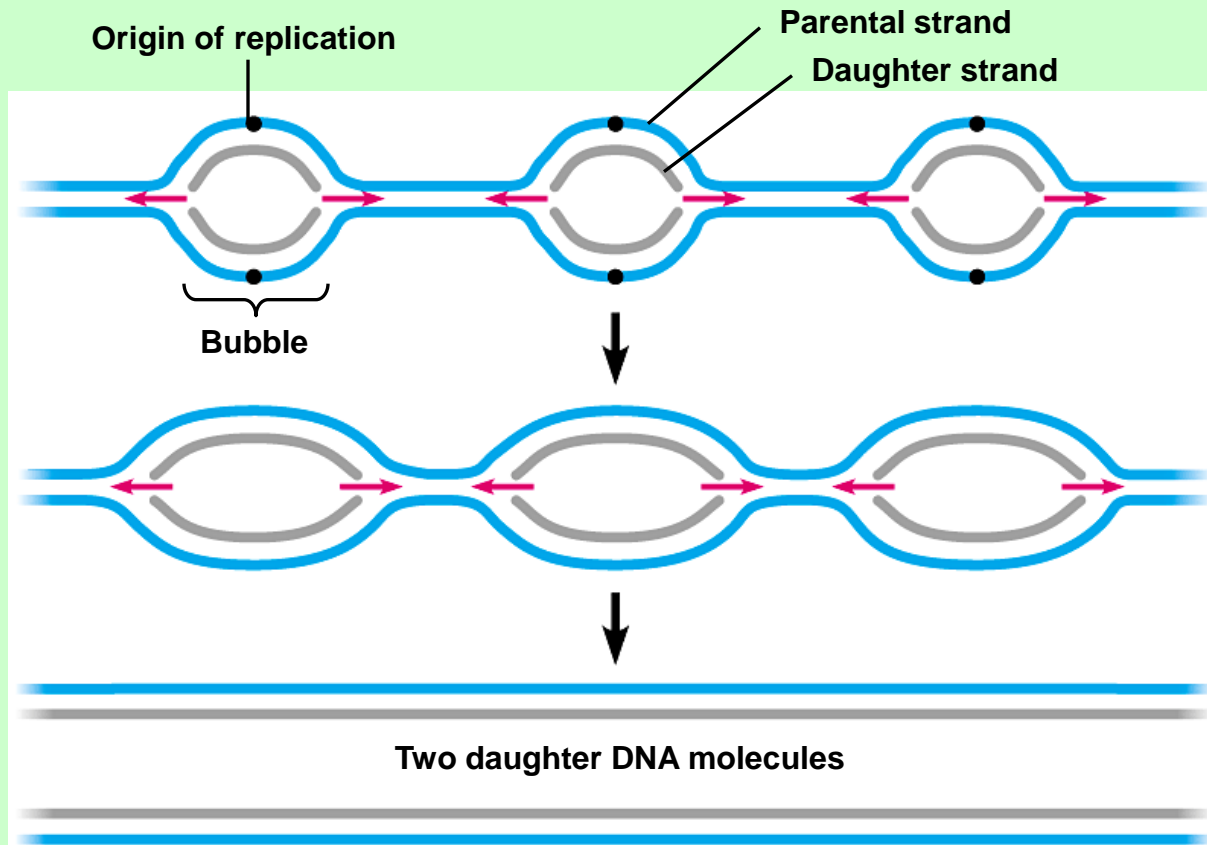
- Replication
- Transcription
- Translation

DNA replication

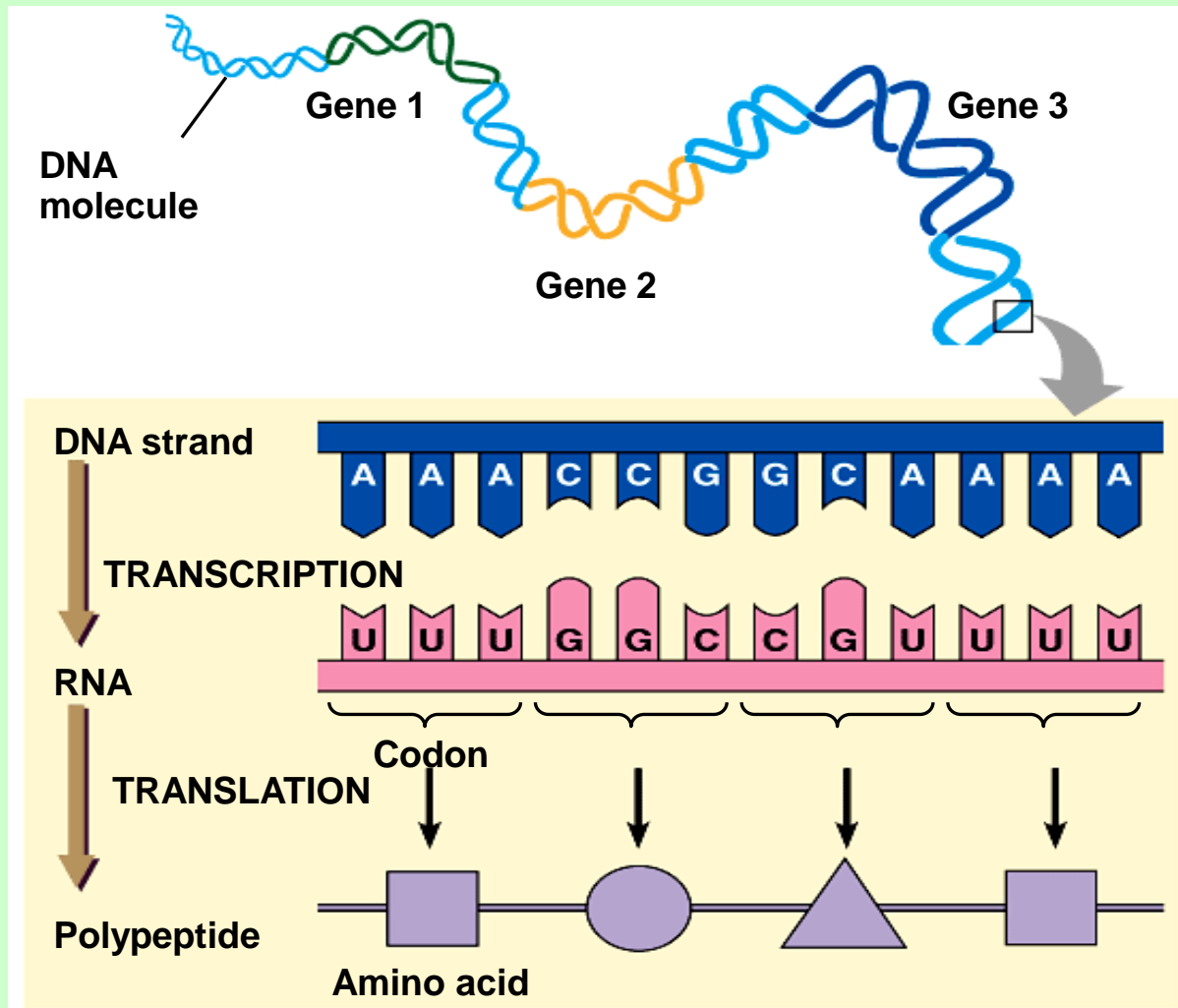
- In DNA **replication**, the strands separate
 - Enzymes assemble the new strands with each strand serving as a template
 - semiconservative



DNA replication begins at specific sites

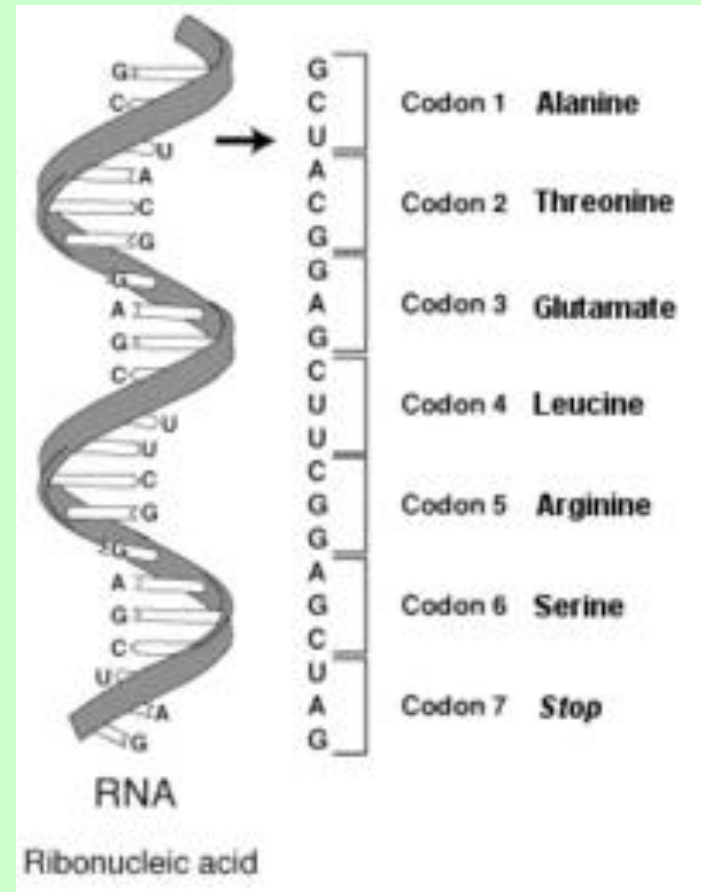


- The “words” of the DNA “language” are triplets of bases called **codons**
 - The **codons** in a gene specify the amino acid sequence of a polypeptide



Genetic code

- set of rules by which a gene is translated into a functional protein.
- a correspondence between nucleotides, the basic building blocks of genetic material, and amino acids, the basic building blocks of proteins
- Three codons are known as "stop codons"



Genetic code

There are 64 possible codons and only 20 standard amino acids

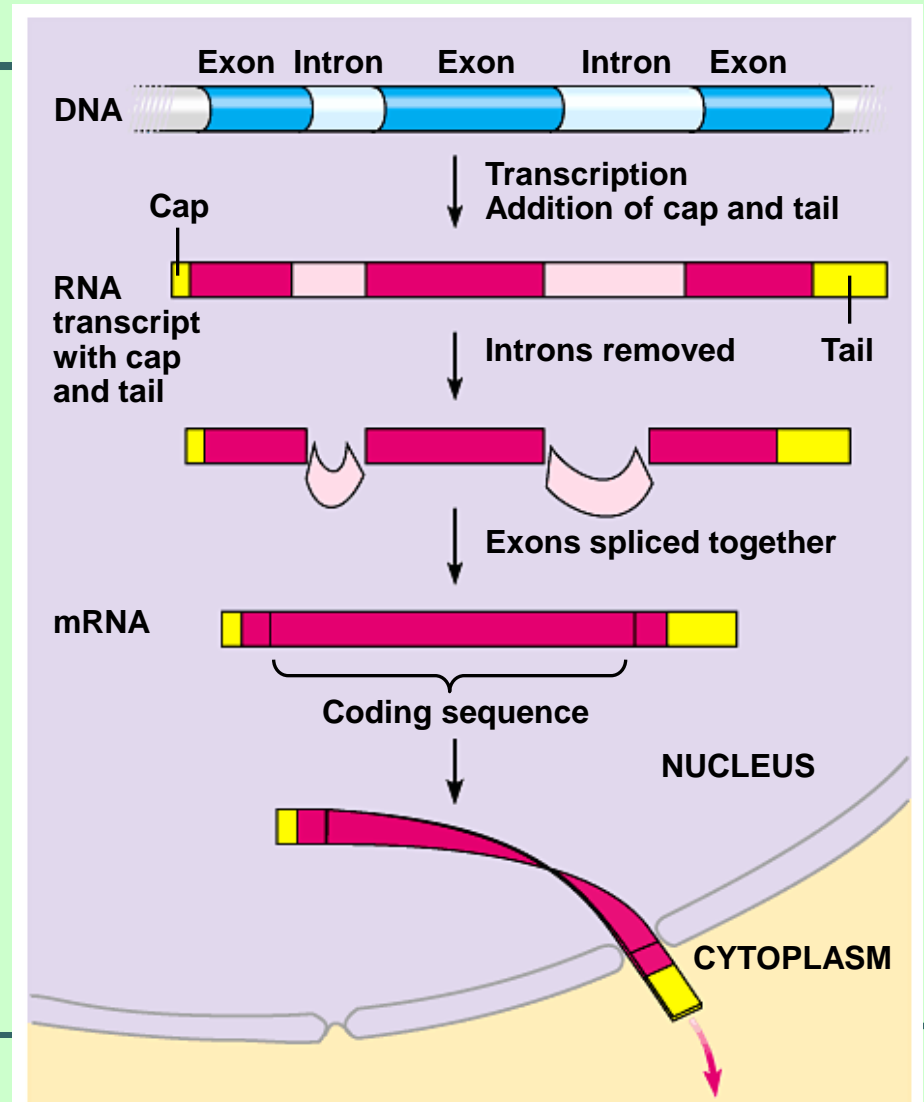
the code is redundant and multiple codons can specify the same amino acid

Virtually all organisms share the same genetic code

| | | Second Letter | | | | |
|--------------|---|--|--------------------------------------|--|---|------------------|
| | | T | C | A | G | |
| First Letter | T | TTT } Phe TTC } TTA } Leu TTG } | TCT } TCC } Ser TCA } TCG } | TAT } Tyr TAC } TAA } Stop TAG } Stop | TGT } Cys TGC } TGA } Stop TGG } Trp | T C A G |
| | C | CTT } CTC } Leu CTA } CTG } | CCT } CCC } Pro CCA } CCG } | CAT } His CAC } CAA } Gln CAG } | CGT } CGC } Arg CGA } CGG } | T C A G |
| | A | ATT } ATC } Ile ATA } ATG } Met | ACT } ACC } Thr ACA } ACG } | AAT } Asn AAC } AAA } Lys AAG } | AGT } Ser AGC } AGA } Arg AGG } | T C A G |
| | G | GTT } GTC } Val GTA } GTG } | GCT } GCC } Ala GCA } GCG } | GAT } Asp GAC } GAA } Glu GAG } | GGT } GGC } Gly GGA } GGG } | T C A G |

Eukaryotic RNA is processed before leaving the nucleus

- Noncoding segments called **introns** are cut out
- **Exons** are the coding regions
- A cap and a tail are added to the ends

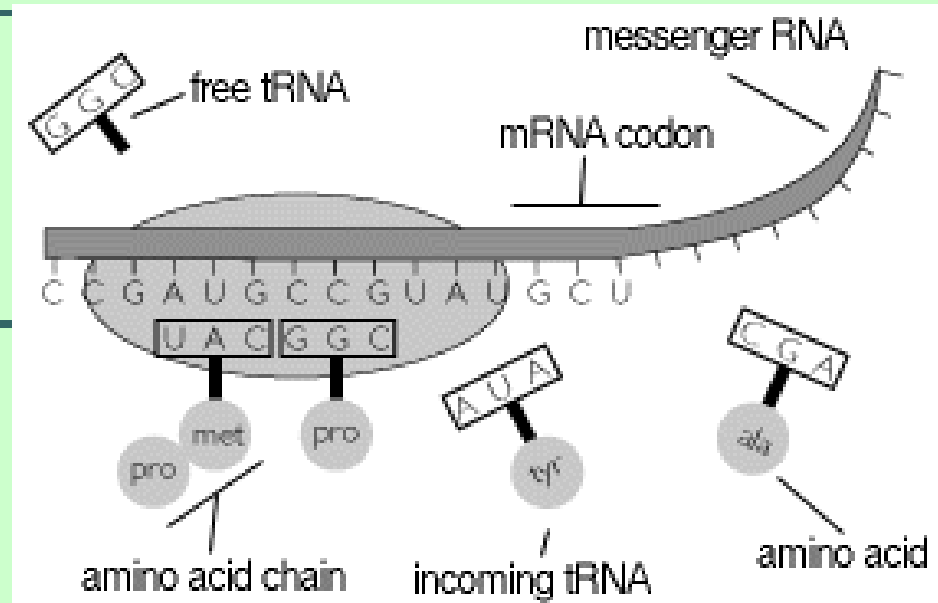


RNA

Several types exist, classified by function

- **mRNA** – this is what is usually being referred to when a Bioinformatician says “RNA”. This is used to carry a gene’s *message* out of the nucleus
- **tRNA** – *transfers* genetic information from mRNA to an amino acid sequence
- **rRNA** – *ribosomal* RNA. Part of the ribosome which is involved in translation

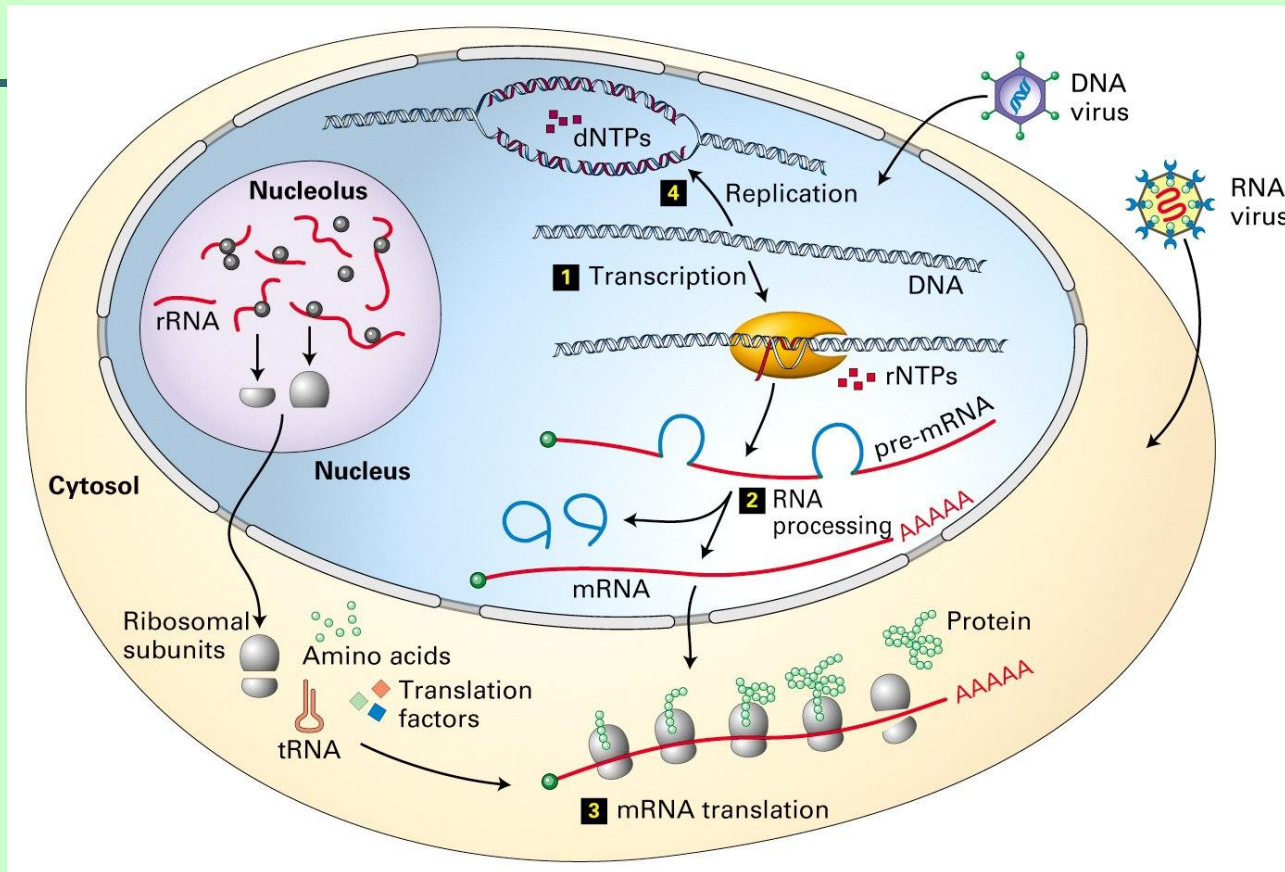
Translation



Include the roles of mRNA, tRNA, codons, anticodons, ribosomes and amino acids.

- Translation- the process of assembling polypeptides from information encoded for in the mRNA
- Each codon specifies the addition of a particular amino acid to the growing polypeptide

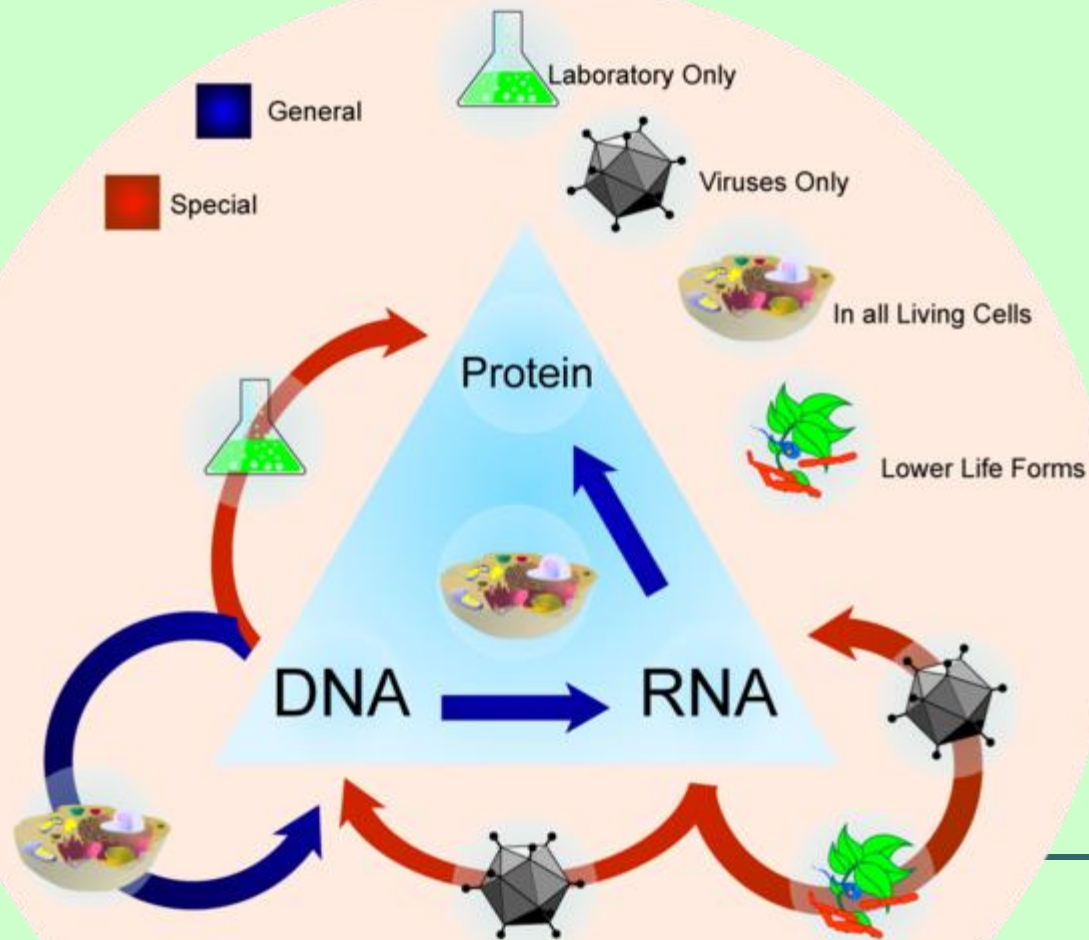
The flow of genetic information in the cell is DNA→RNA→protein



A gene is expressed in two steps

- Transcription: RNA synthesis
- Translation: Protein synthesis

The central dogma of molecular biology

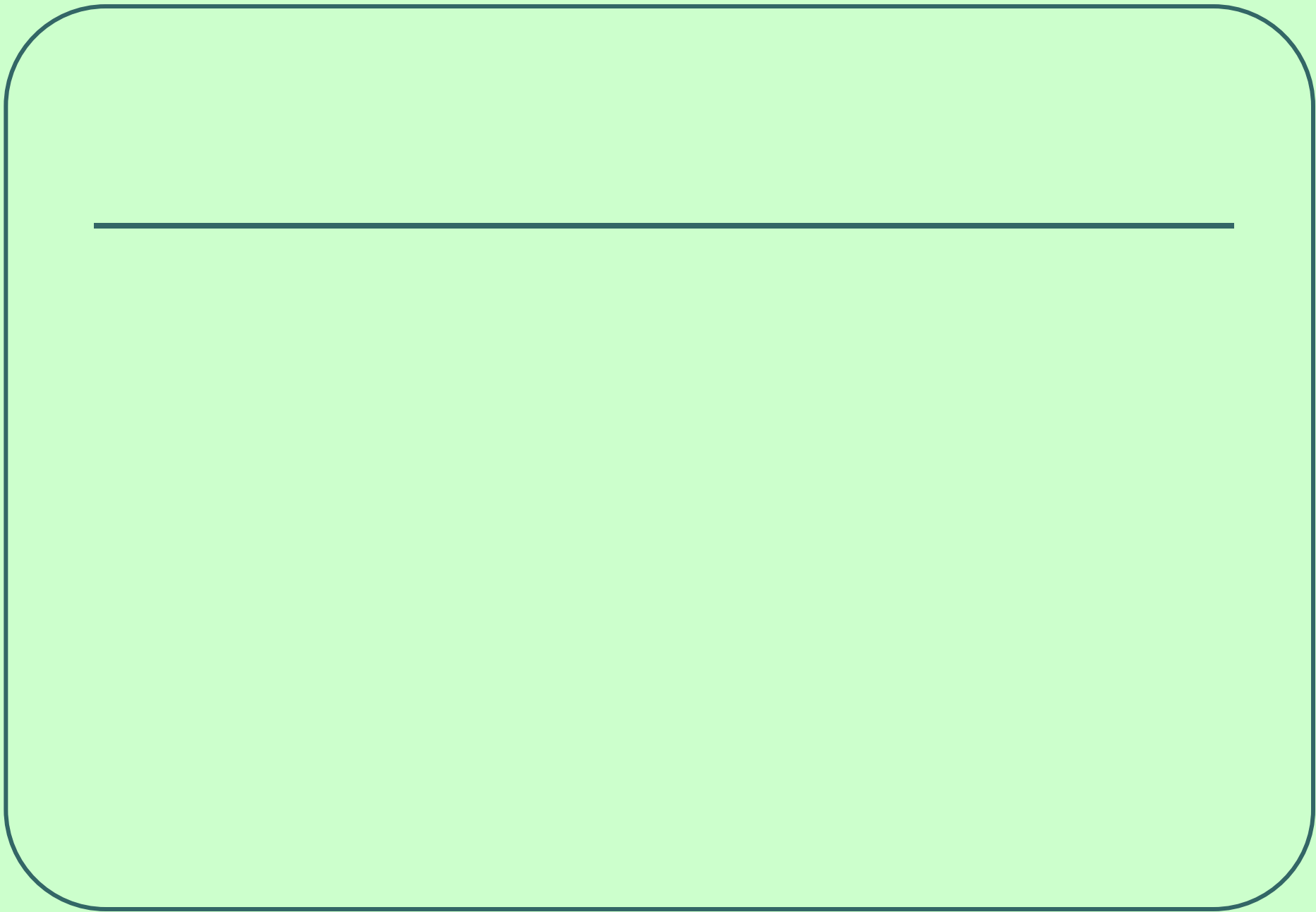


The central dogma of molecular biology

the transfer of sequence information between sequential information-carrying biopolymers - DNA and RNA (both nucleic acids), and protein

The general transfers describe the normal flow of biological information:

- DNA can be copied to DNA (DNA replication),
- DNA information can be copied into mRNA, (transcription),
- proteins can be synthesized using the information in mRNA as a template (translation)



Mutations

Any alteration in a gene from its natural state; may be disease causing or a benign, normal variant
Frequency less than 1 %

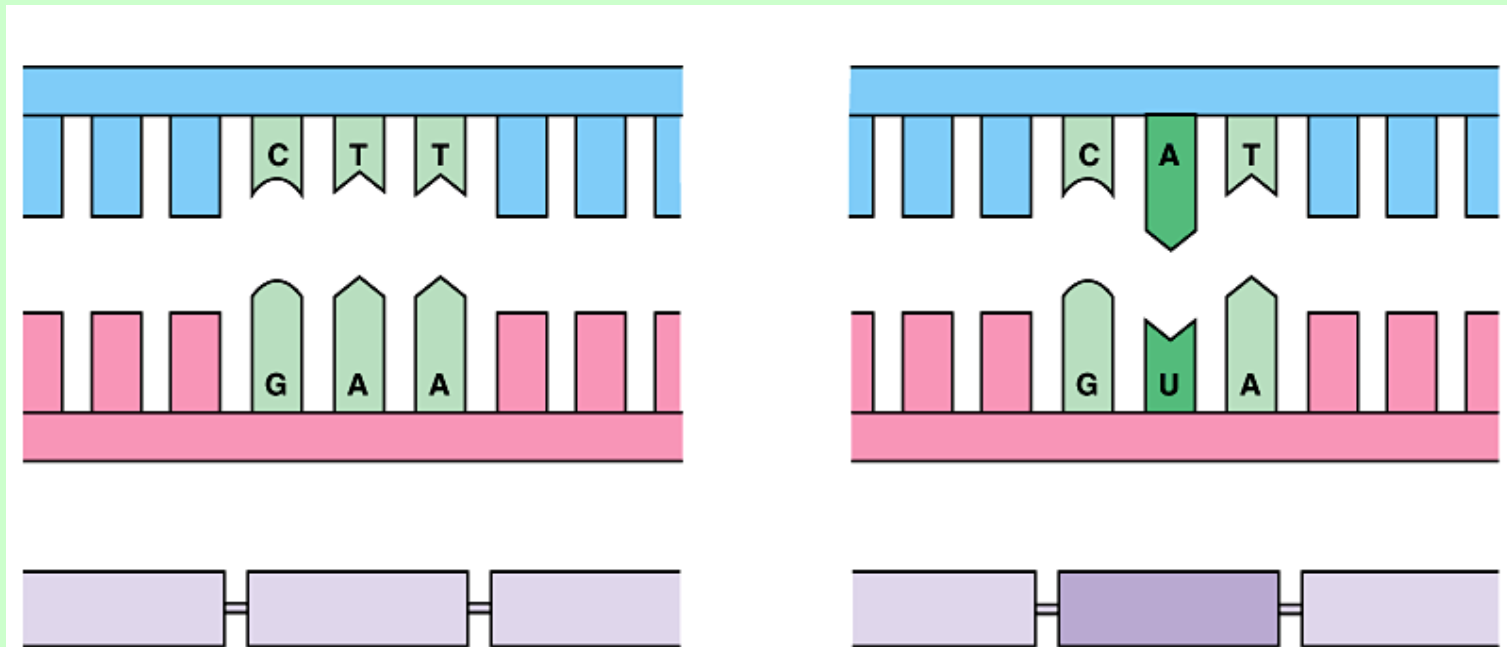
Mutations - positive (variability, selection)

- negative (4500 monogenic diseases, ageing)
- neutral

Each human: 5 – 10 pathologic mutations

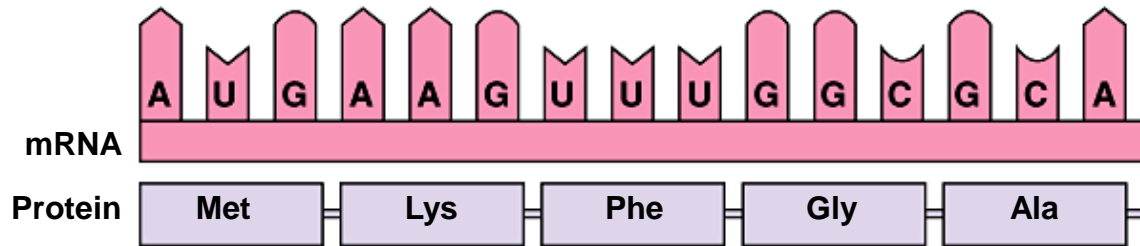
Mutations are changes in the DNA base sequence

These are caused by errors in DNA replication or by mutagens

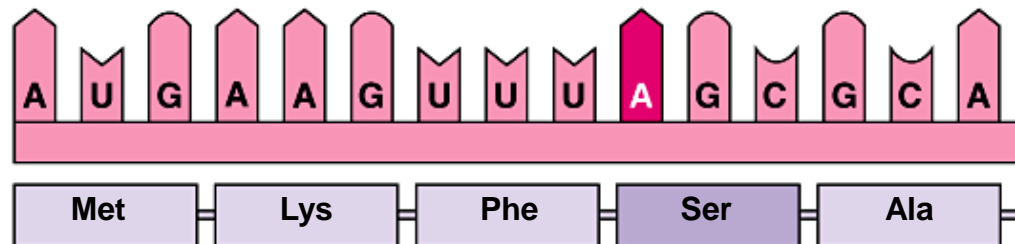


Types of mutations

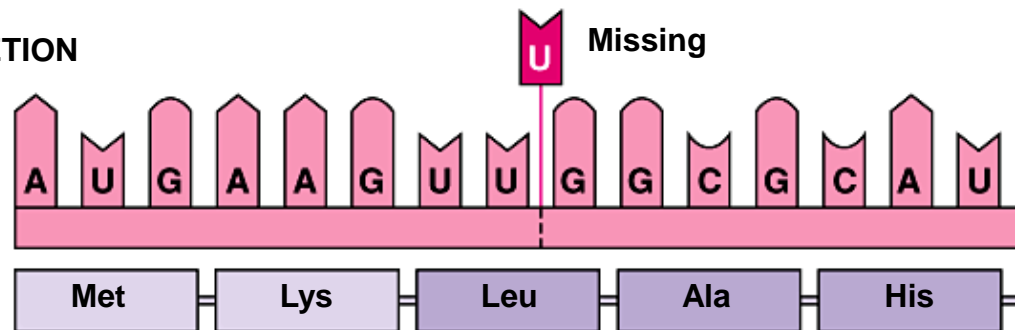
NORMAL GENE



BASE SUBSTITUTION



BASE DELETION



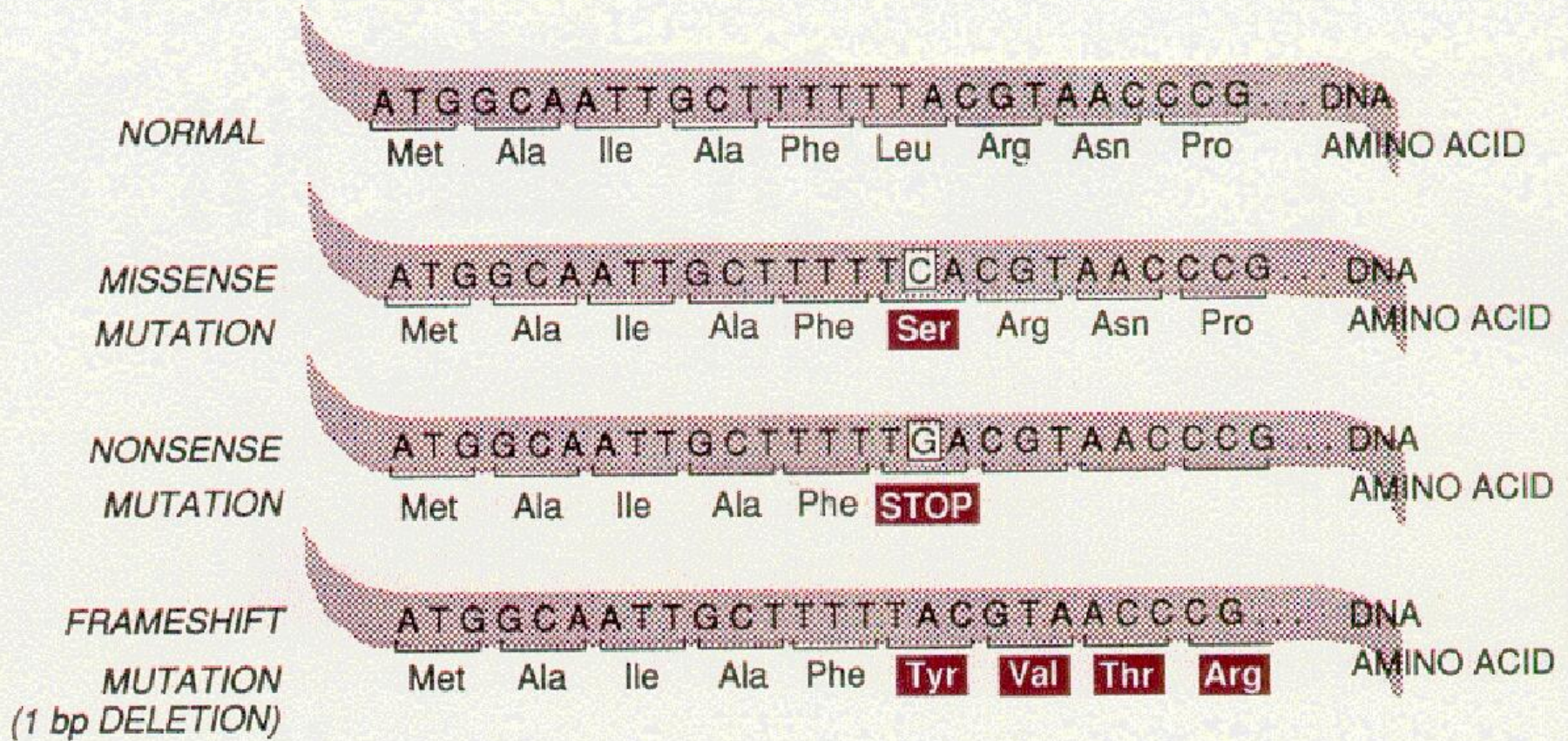


Figure 2.14. Examples of mutation. The “sense” strand DNA sequence of a coding region is shown, together with the encoded amino acid sequence. Three different mutations affecting the second nucleotide of a leucine codon are shown.

Class**Result****Single base-pair substitutions
(point mutations)**

Altered structure of gene product

Missense mutation

Nonsense mutation

RNA-splicing mutation

Altered quantity of gene product

Mutations in regulatory sequences

Mutations in RNA processing and translation

Single amino acid replacement in the protein

A termination codon in the middle of the gene results in premature termination of protein synthesis

The protein may be missing part or all of an exon sequence

Transcription of the gene is altered, which can reduce or eliminate the gene product

The stability of the messenger RNA is altered, which may reduce the amount of gene product

Insertions or deletions

One or two base pairs (frameshift mutations)

Large numbers of base pairs

The addition or deletion of one or two base pairs can affect the reading frame of the gene, resulting in a grossly altered or absent gene product

Large pieces of the DNA may be lost or large segments of DNA may insert into the middle of a gene, resulting in loss of function

Expansion of trinucleotide repeat sequences

Unstable trinucleotide repeats can suddenly expand in number, resulting in the alteration of production or structure of a particular gene product

Chromosomal alterations

Inversions, translocations, duplications, or gene amplification may result

Mutations in the coding sequence of a structural gene

- **Silent mutations** do not alter the amino acid sequence of the polypeptide
- **Missense mutations** - an amino acid change does occur
 - Example: Sickle-cell anemia
 - If the substituted amino acids have similar chemistry, the mutation is said to be neutral
- **Nonsense mutations** change a normal codon to a termination codon
- **Frameshift mutations** involve the addition or deletion of nucleotides in multiples of one or two
 - This shifts the reading frame so that a completely different amino acid sequence occurs downstream from the mutation

Clasification of mutations according to its effect on gene product

1. *Product with lower to zero function (loss-of-function)*

- typical product is enzyme
- type of mutation is frequently deletion

2. *Product with abnormal function (gain-of-function)*

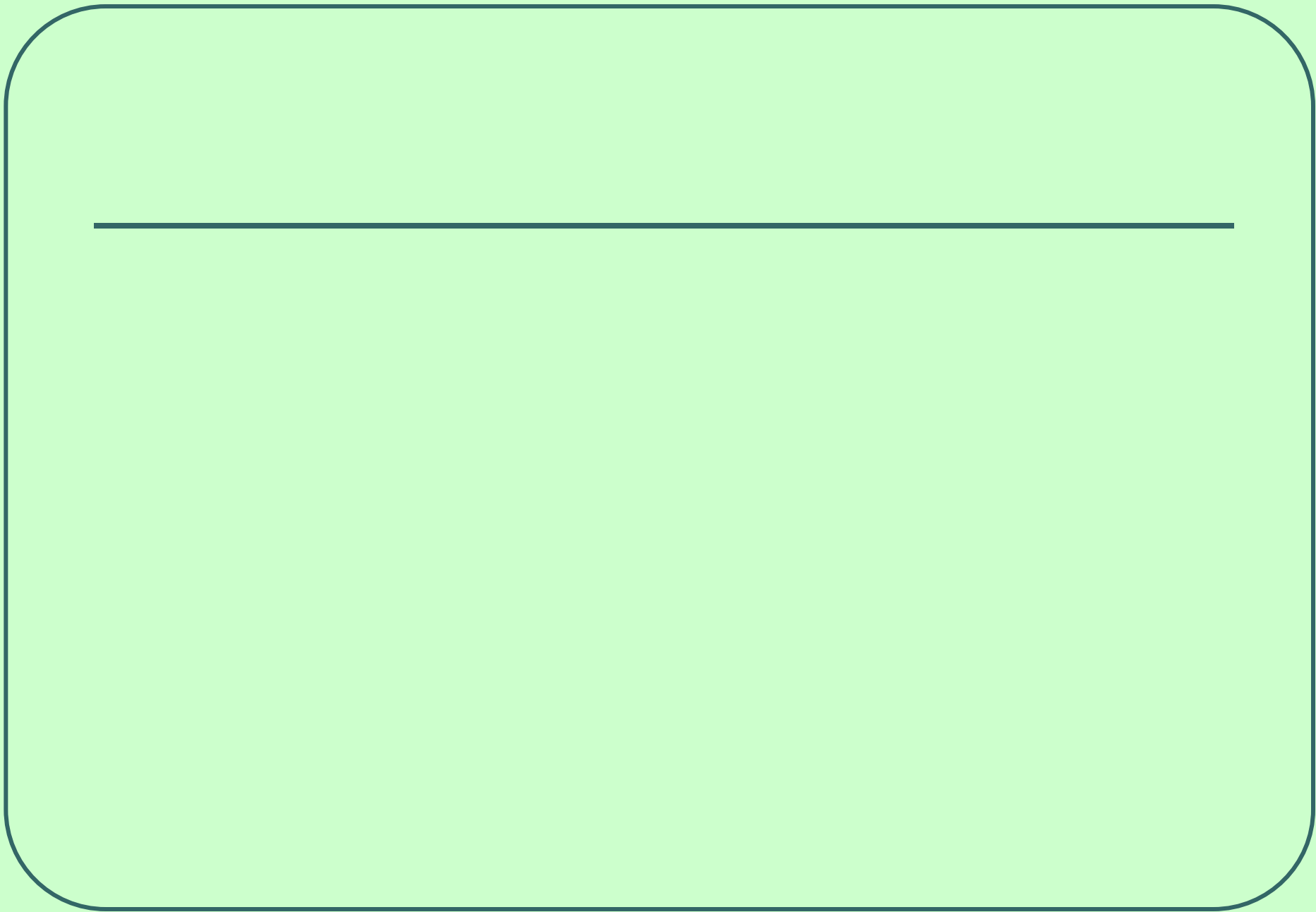
- typical product is nonenzymatic protein
- frequently in tumours (somat. mutation), rarely in monogenic diseases
- deletions do not lead to new function

Type 1 frequently recessive, type 2 dominant mutations
In some genes- both types of mutations

Type of mutation

Example

| | |
|--------------------|---|
| Normal | THE ONE BIG FLY HAD ONE RED EYE |
| Missense | THQ ONE BIG FLY HAD ONE RED EYE |
| Nonsense | THE ONE BIG |
| Frameshift | THE ONE QBI GFL YHA DON ERE DEY |
| Deletion | THE ONE BIG HAD ONE RED EYE |
| Insertion | THE ONE BIG WET FLY HAD ONE RED EYE |
| Duplication | THE ONE BIG FLY FLY HAD ONE RED EYE |
| Expanding mutation | |
| generation 1 | ● THE ONE BIG FLY HAD ONE RED EYE |
| generation 2 | ● THE ONE BIG FLY FLY FLY HAD ONE RED EYE |
| generation 3 | ● THE ONE BIG FLY FLY FLY FLY FLY HAD ONE RED EYE |



Major types of Genetic diseases

a.) chromosomal diseases

- are the result of the addition or deletion of entire chromosomes or part of chromosomes
- most major chromosome disorders are characterised by growth retardation, mental retardation and variety of somatic abnormalities
- typical examples of major chromosomal disease is Down syndrom (trisomy 21), Edwards sy(trisomy 18), Patau sy (trisomy13)

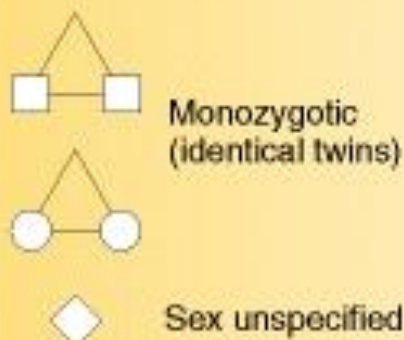
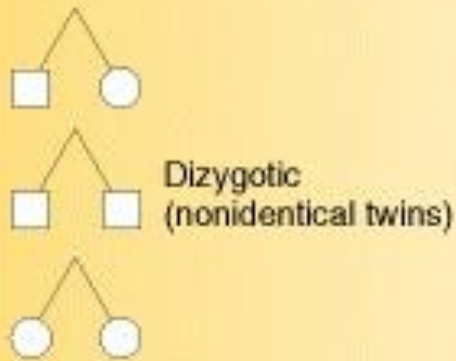
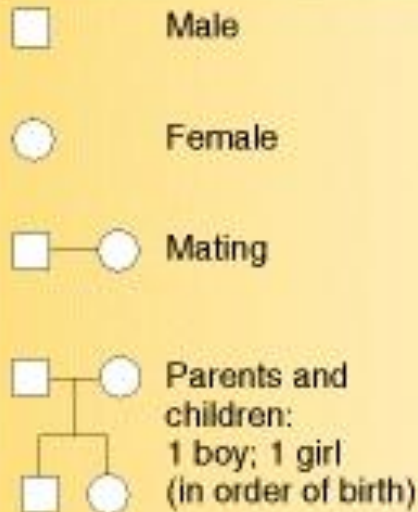
b.) monogenic diseases (single gene defects)

- only a single gene is altered (mutant) → flawed protein → manifestation (development) of a disease
- inherited in simple Mendelian fashion
- some 6000 distinct disorders are now known (sickle cell anemia, familial hypercholesterolemia, cystic fibrosis, Hemophilia A., Duchenne Muscular Dystrophy, Huntington Disease...)

c.) polygenic diseases

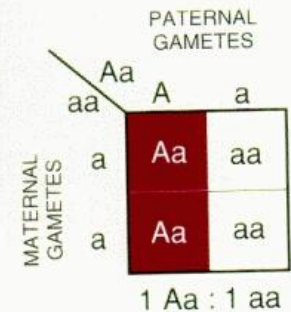
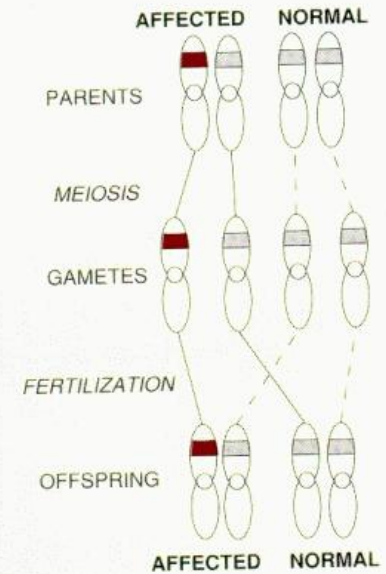
- result from the interaction of multiplex genes, each of which may have a relatively minor effect
- environmental factors contribute to the manifestation of these diseases (e.g. nutrition, exercise)
- for this group of illnesses, the contribution of the gene can be thought of as a “predisposition”
- examples: diabetes mellitus, hypertension, schizophrenia and congenital defects such as cleft lip, cleft palate and most congenital heart diseases
- very common in the population

Human pedigree



Autosomal dominant inheritance process

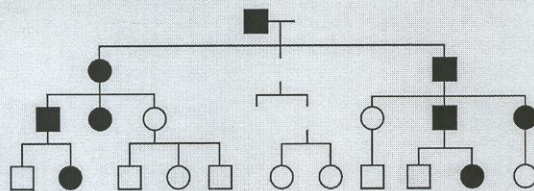
Only one of the two homologous genes is mutated and although another normal gene is present (**heterozygosity**), the illness still appears (**dominant gene effect**). If, therefore, one of the parents carries this gene, there is a 50% probability that it will be transmitted to each child. Both men and women can be affected by this. This inheritance pattern accounts for over 60% of monogenic diseases, representing by far the most common inheritance process. Obviously a mutated protein in just half the amount will have a pathological effect on the human organism in such cases. E.g. **achondroplasia**



A = MUTANT, a = NORMAL

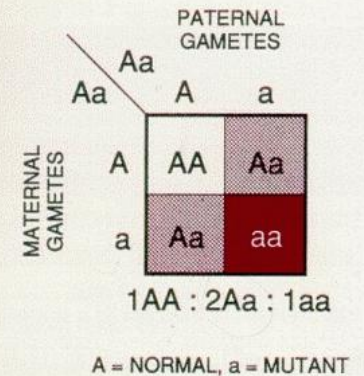
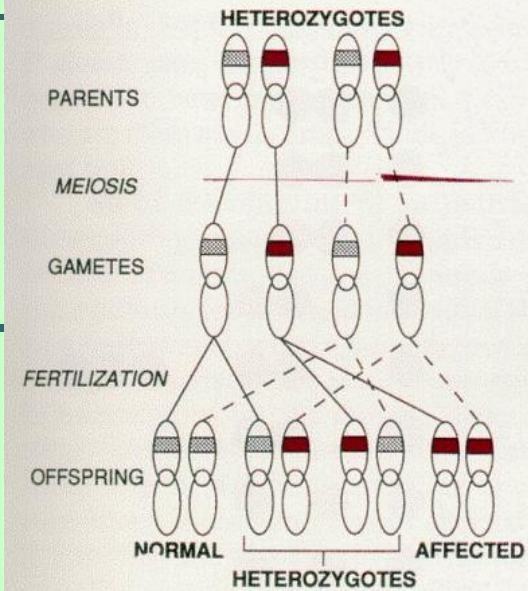
Figure 3.4. Autosomal dominant inheritance. The mating diagram (top) and Punnett square (bottom) show the outcome of a mating between a heterozygous affected individual and a homozygous normal individual. The mutant allele in the mating diagram is indicated in red, and affected offspring in the Punnett square are indicated in red.

- healthy man
- diseased man
- healthy woman
- diseased woman



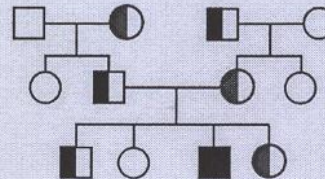
Autosomal recessive inheritance

In this inheritance pattern, both homologous genes must be mutated (**homozygosity**) in order to produce an illness in the affected person. Individuals, who only receive one version of the mutated gene are called **carriers**. Both sexes can be affected. If, for example, both parents are carriers, there is a 25% chance that the child will receive both mutated genes and so develop the illness. Many metabolic diseases fall into this category (e.g. **cystic fibrosis, phenylketonuria, adrenogenital syndrome, haemochromatosis**).



- healthy man
- healthy woman
- diseased man/woman

- ◻ male carrier
- ◐ female carrier

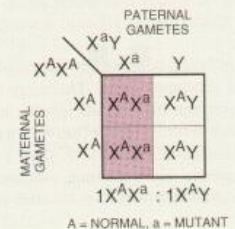
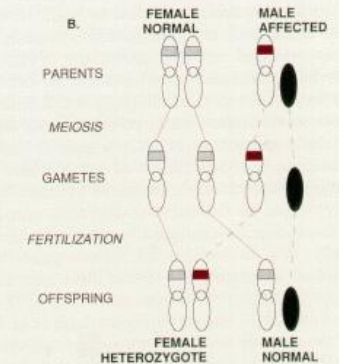
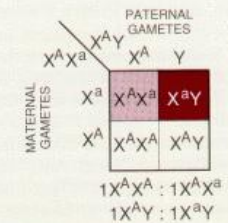
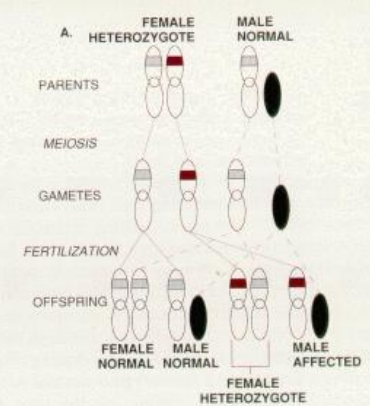
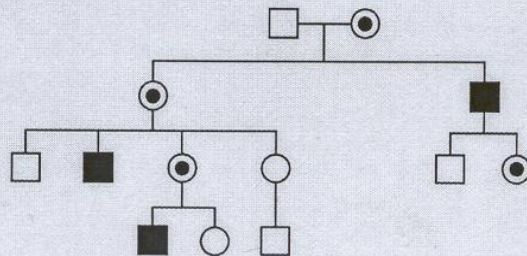


X chromosome inheritance (sex-linked inheritance)

Women have two X chromosomes. If they have a recessively acting mutated gene on one X chromosome, they are carriers for the corresponding illness. Men have only one X chromosome, since the other sex chromosome is a Y chromosome. If they have the mutated gene on the X chromosome, they will develop the illness as a rule.

If a woman is a carrier for the illness inherited by the X chromosome, there is a 50% chance that she will pass on this illness to her son. Her daughters have a 50% chance of becoming a carrier for this illness.

- healthy man
- diseased man
- healthy woman
- female carrier



Identification of inherited diseases

1.) Phenotype analysis

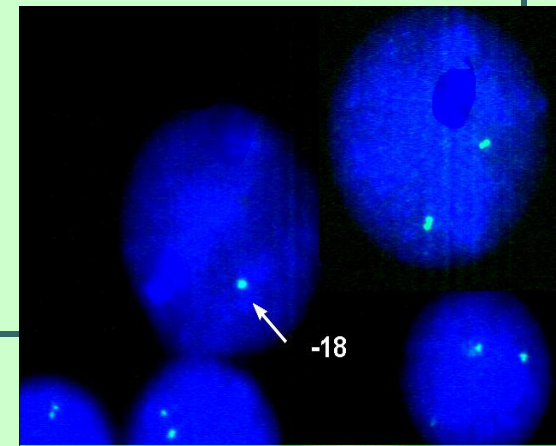
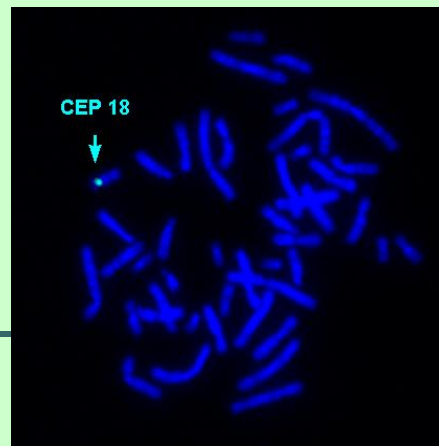
Genes are directly responsible for the production of hormones, enzymes and other proteins. Investigation procedure: Diagnostic measurement of altered or missing proteins using blood or urine analysis. This provides indirect evidence of a mutation of the gene responsible for this.

Examples: Phenylketonuria, alpha1-antitrypsin deficiency

2.) Chromosome analysis (cytogenetic investigations)

This includes microscope examinations to investigate chromosome alterations in terms of number (duplication or loss of individual chromosomes = **numeric chromosome aberration**) and in terms of structure (wrong composition, chromosome breaking = **structural chromosome aberration**). There is no detailed investigation of individual genes in such cases.

Indication: Anomalies in children (malformations, retarded development) in the context of prenatal diagnosis, tendency to miscarriages, infertility.



3.) Molecular genetics testing (DNA analysis, genome analysis DNA tests)

This provides evidence of a gene mutation responsible for producing the illness. Here it is determined whether the sequence of the DNA bases (**nucleotide sequence**) has changed within the affected

DNA/RNA diagnosis of genetic diseases

Not all mutation test use DNA. Testing RNA by **RT-PCR** has advantages when screening genes with many exons (NF1 gene, DMD gene...) or seeking splicing mutations.

Very important in molecular genetic testing is using a **protein-based functional assay**, which may classify the products into two simple groups: *functional* and *nonfunctional* – essential question in most diagnostics

Limitations of DNA analysis

monogenic and also polygenic diseases sometimes do not occur in both twins, even though the genetic information is the same in identical twins. This is due to several factors:

Penetrance: not every pathogenic mutation leads to the manifestation of a disease in the lifetime of a person.

Expressivity on the other hand describes quantitative differences in the manifestation of the disease/symptoms. Sometimes, the two concepts are difficult to separate, when, for example, a disease is so weakly manifested that it can no longer be diagnosed.

The age at which the disease manifests itself can vary strongly. An example of this is Huntington's chorea. Differences in the onset of diseases are sometimes explained by so-called **dynamic mutations**. In passing on to the next generation, the disease-inducing mutation can lead to an earlier onset of the illness (**anticipation**) involving the extension of a mutated sequence of bases.

In many cases, genetic information is manifested in a different way when it is inherited from the mother than when it is inherited from the father. Here one speaks of **imprinting**.

Molecular genetics testing (DNA analysis, genome analysis DNA tests)

A.) Direct testing

– DNA from a patient is tested to see whether or not it carries a given pathogenic mutation

B.) Indirect testing (gene tracking)

- linked markers are used in family studies to discover whether or not the consultand inherited the disease-carrying chromosome from a parent

A.) Direct testing

- provides evidence of a gene mutation responsible for producing the illness. It is determined whether the sequence of the DNA bases (**nucleotide sequence**) has changed
- to see whether the DNA of tested person has a gene normal or mutant

Detection of mutation in relevant gene **always confirms the clinical diagnosis**

we must know

which **gene** to examine

the relevant „**normal**“ (wild type) **sequence**

Mutation testing methods can be divided into two groups:

1. **Mutation detection methods (scoring)** – test the DNA for the presence or absence of one specific mutation. Searching for known mutations
2. **Mutation screening methods (scanning)** – screen a sample for any deviation from the standard sequence.

1. Mutation detection methods – test a DNA for the presence or absence of one specific mutation

searching for *known* sequence change is **possible** for:

- diseases where all affected people in the population have **one particular mutation**
- most affected people in the population have **one** of limited number of **specific mutations**
- diagnosis **within a family** - once mutation is characterized, other family members need to be tested for that particular mutation

2. Mutation screening methods - screen a sample for any deviation from the standard sequence

The mutation screening is possible for diseases where a good proportion of patients carry independent mutations.

Testing for unknown mutations in laboratory suffer **two limitations**:

methods are quite **laborious** and **expensive** for use in diagnostic service, which needs to produce answers quickly

detect differences between the patient's sequence and published normal sequence (not distinguish between pathogenic and nonpathogenic changes.)

Polymerase chain reaction (PCR)

To amplify a single or a few copies of a piece of DNA across several orders of magnitude, generating thousands to millions of copies of a particular DNA sequence. The method relies on thermal cycling, consisting of cycles of repeated heating and cooling of the reaction for DNA melting and enzymatic replication of the DNA

Kary Mullis – 1983 discovered the PCR procedure, for which he was awarded the Nobel prize



PCR

selective amplification of specific target DNA sequence within heterogeneous collection of DNA (total genomic DNA or complex cDNA) requires:

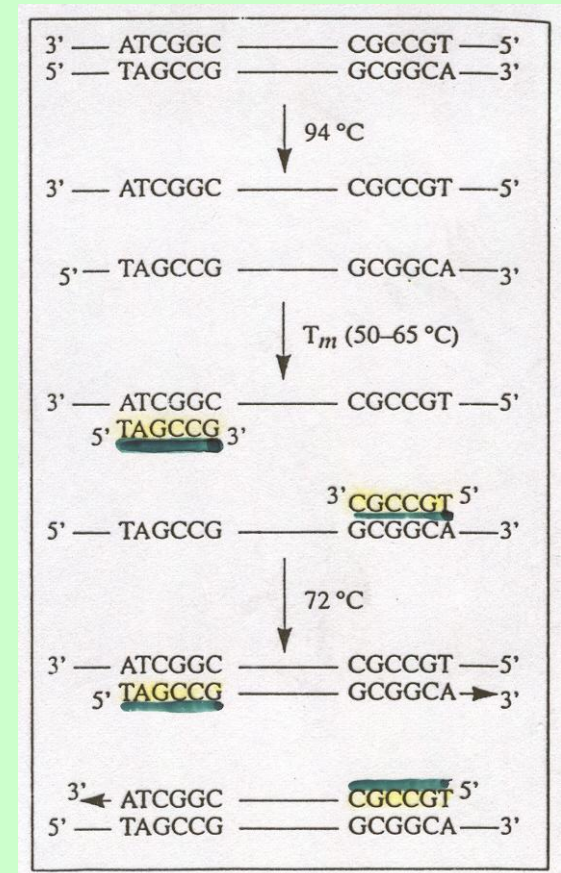
- sequence information from the target sequence for construction two oligonucleotide primer sequences (15 – 30 nucleotides long)
- denatured genomic DNA
- heat stable DNA polymerase
- DNA precursors (four deoxynucleotide triphosphates dATP, dCTP, dGTP and dTTP)

PCR involves sequential cycles composed of three steps:

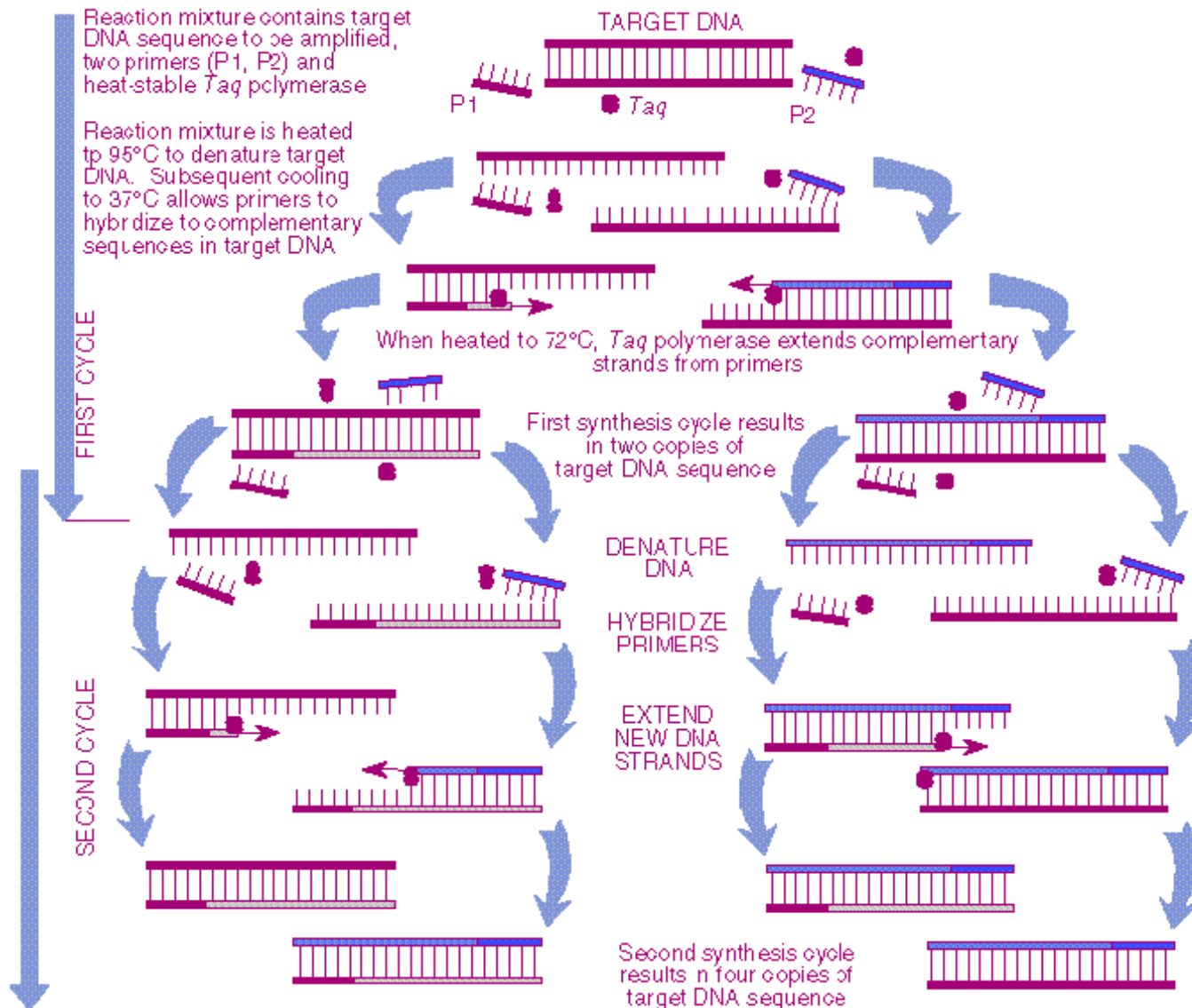
- **Denaturation** (typically at about 93 – 95° C)
- **Reannealing** (at temperatures usually from about 50° – 70°C, depending on T_m of the expected duplex
- **DNA synthesis** – typically at about 70 – 75°

Sensitivity of PCR allows us to use a wide range of samples:

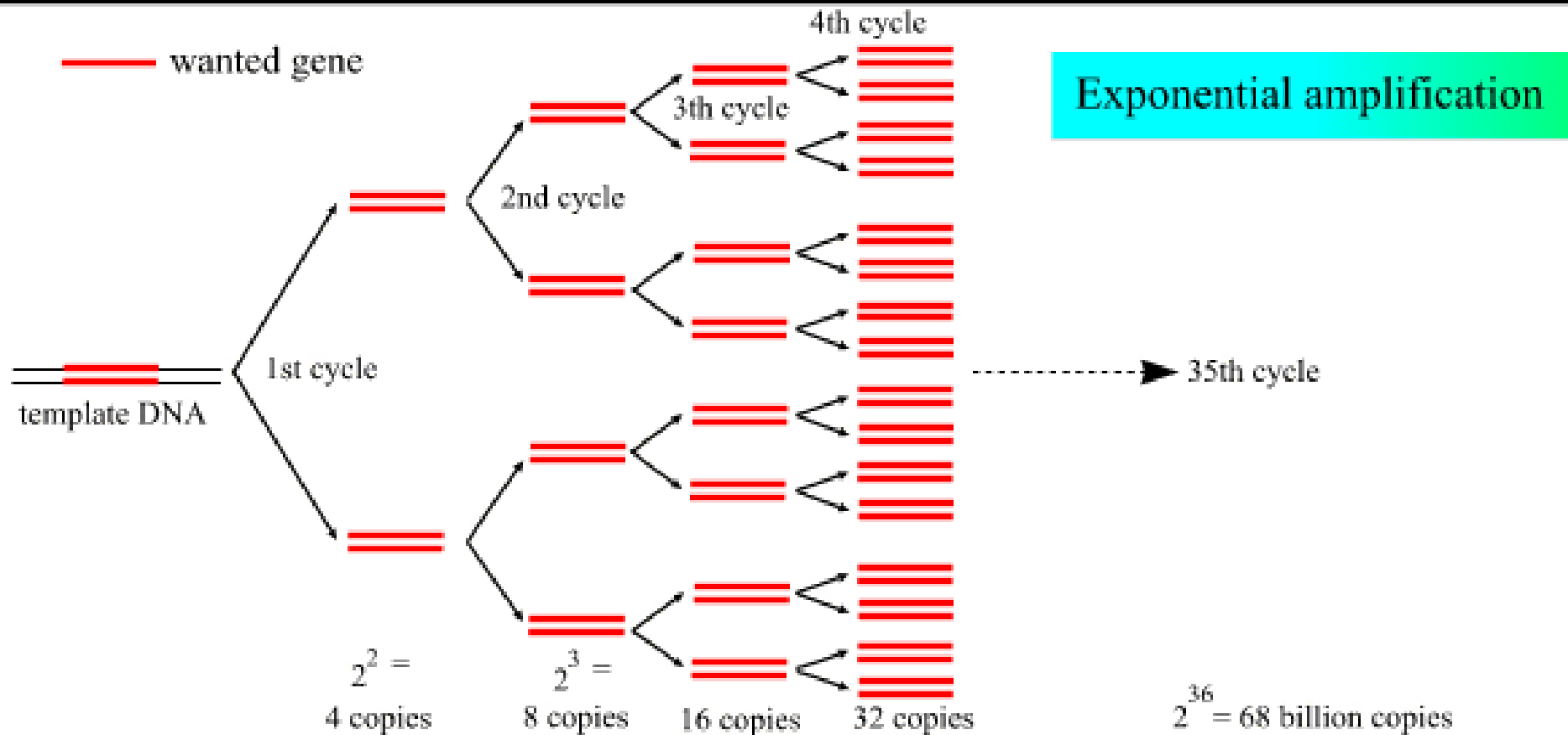
- blood samples
 - monthwashes or buccal scrapes
 - chorionic villus biopsy samples
 - amniocentesis specimens
 - one or two cells (removed from eight-cell stage embryos)
 - hair, semen
 - archived pathological specimens
- Guthrie cards (spot of dried blood)



DNA Amplification Using Polymerase Chain Reaction



Source: *DNA Science*, see Fig. 13.



(Andy Vierstraete 1999)

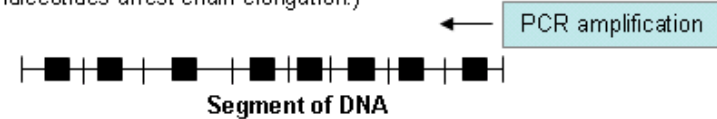
Table 16.1: Methods for scanning a gene for point mutations

| Method | Advantages | Disadvantages |
|---|--|---|
| Sequencing | Detects all changes Mutations fully characterized | Laborious Generates excessive information |
| Heteroduplex gel mobility | Very simple | Short sequences only (<200 bp) Limited sensitivity Does not reveal position of change |
| Single-strand conformation polymorphism (SSCP) analysis | Simple | Short sequences only (<200 bp) Does not reveal position of change |
| Denaturing gradient gel electrophoresis (DGGE) | High sensitivity | Choice of primers is critical GC-clamped primers expensive Does not reveal position of change |
| Mismatch cleavage (i) chemical (ii) enzymic | High sensitivity Shows position of change | Toxic chemicals Experimentally difficult |
| Protein truncation test (PTT) | High sensitivity for chain terminating mutations Shows position of change | Chain terminating mutations only Expensive Experimentally difficult Best with RNA |

- **sequence analysis:** (synonym: sequencing) Process by which the nucleotide sequence is determined for a segment of DNA

Step 1: Amplification.

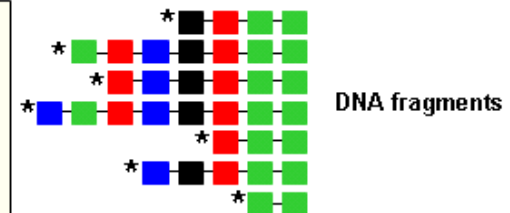
- A. The segment of DNA to be sequenced is PCR-amplified using normal nucleotides (i.e., dATP, dTTP, dGTP, and dCTP) and fluorescently labeled, dideoxy nucleotides (i.e., ddATP, ddTTP, ddGTP, ddCTP). (Dideoxy nucleotides arrest chain elongation.)



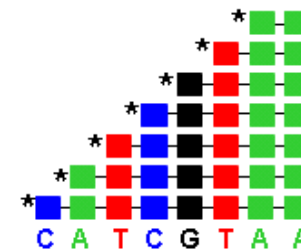
- B. The DNA segment is copied when normal nucleotides are incorporated. Copying ceases when a dideoxy nucleotide is incorporated. By this process, many different-sized fluorescently labeled DNA fragments are produced.

Key: Nucleotides used in PCR reaction

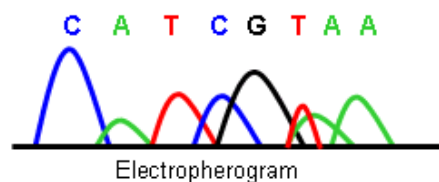
| Normal nucleotides | Fluorescently labeled, dideoxy nucleotides |
|---|--|
| ■ dATP | * ■ ddATP |
| ■ dTTP | * ■ ddTTP |
| ■ dGTP | * ■ ddGTP |
| ■ dCTP | * ■ ddCTP |



- Step 2: Sequence determination.** The fragments are sorted by length. A sequencing machine reads the fluorescent wavelengths to determine which nucleotide is at the end of each fragment.

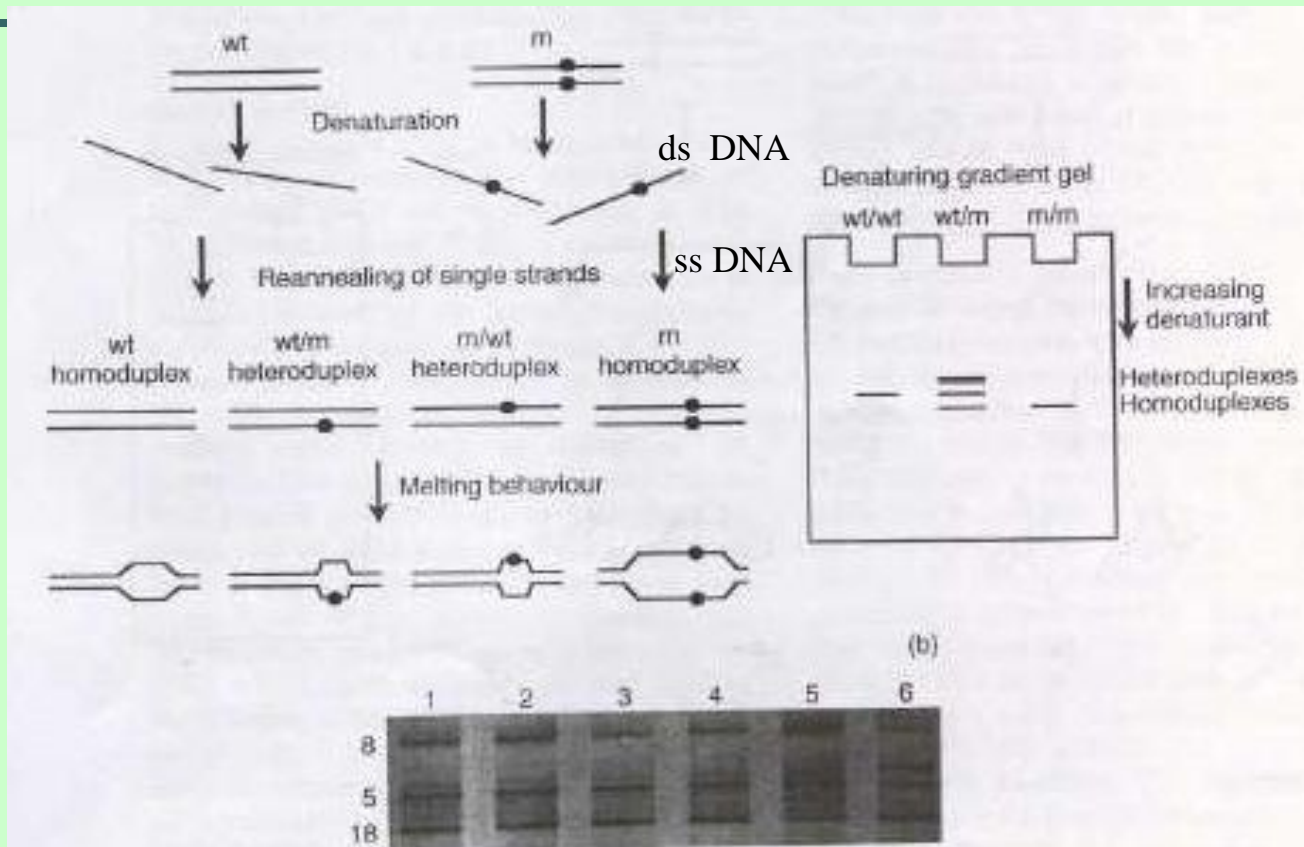


- Step 3: Sequence reporting.** Sequence data are typically displayed on an electropherogram as colored peaks. Each peak represents a nucleotide, corresponding to the letter above it.



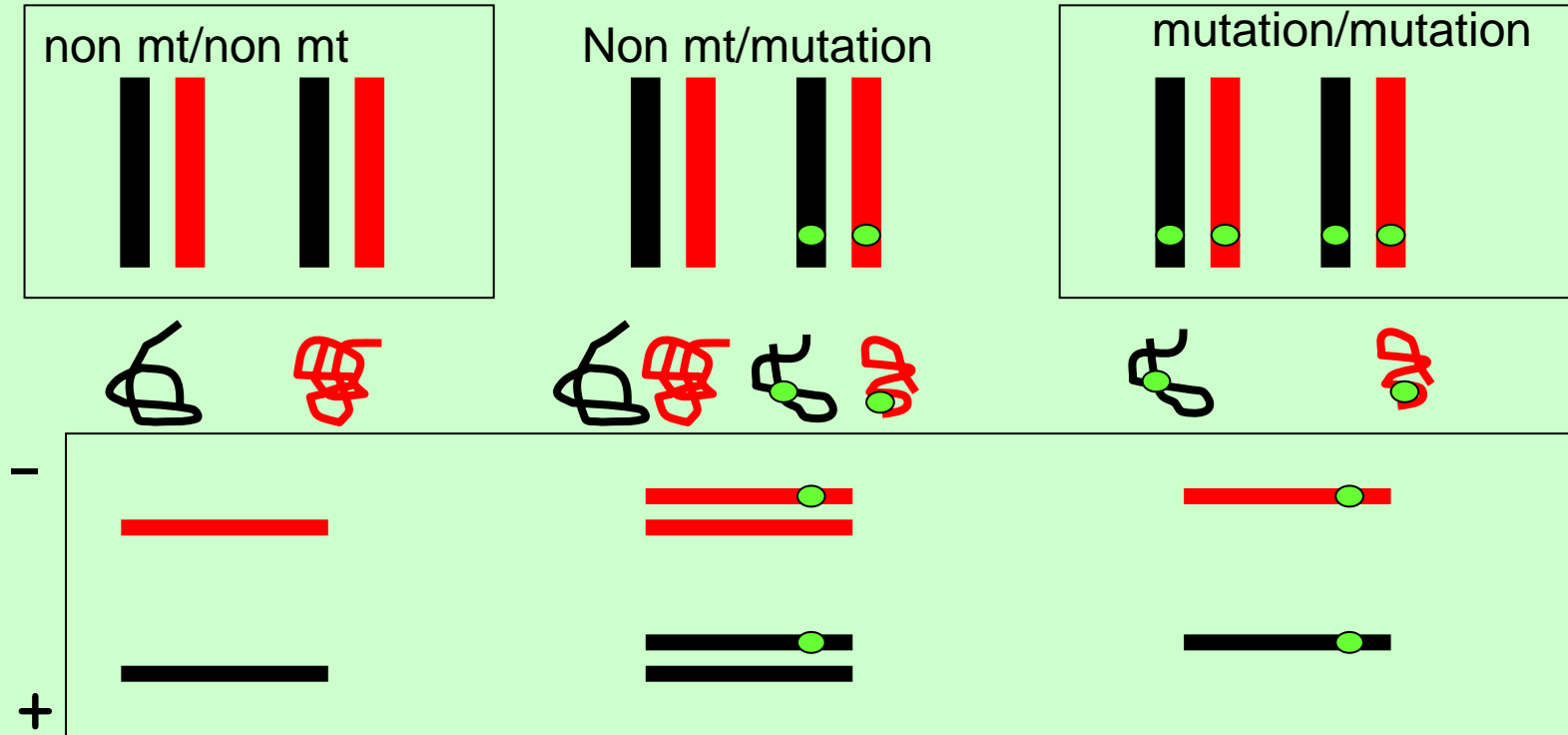
| Wavelength | Nucleotide |
|------------|------------|
| Red | T |
| Green | A |
| Blue | C |
| Black | G |

Princip metody DGGE



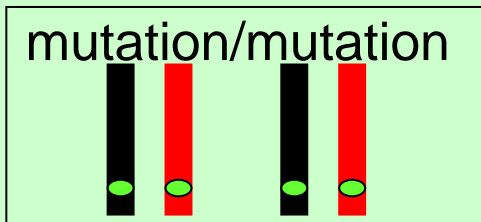
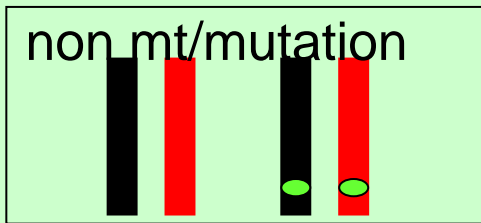
DGGE: the sequence-specific denaturation characteristics in a chemical gradient (in the gel) lead to partial separation of strands. This in turn leads to differential mobility and results in a single band per variant

SSCP in gel (Single-strand conformation polymorphism)

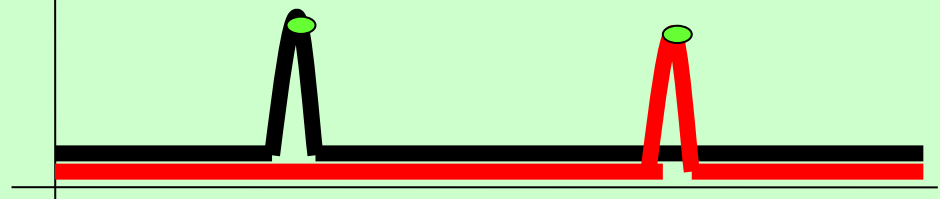
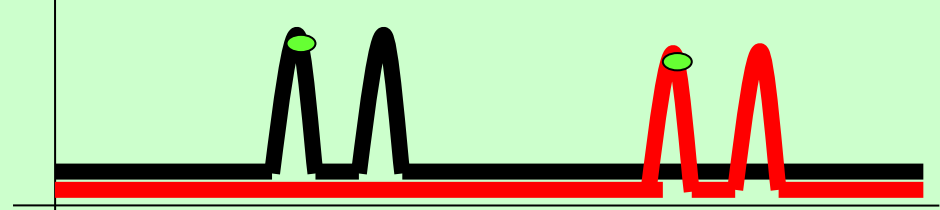
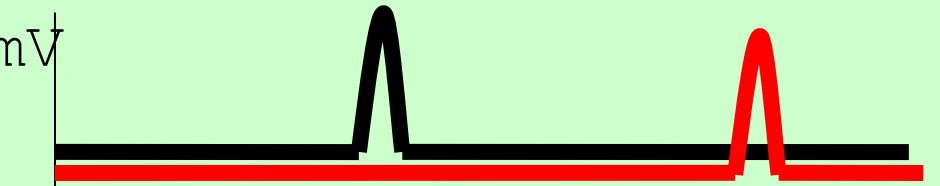


SSCP: after denaturation, single strands form a sequence-specific structure. This structure leads to differential mobility in a non-denaturing matrix and two bands per variant

SSCP in capillary



mV



time

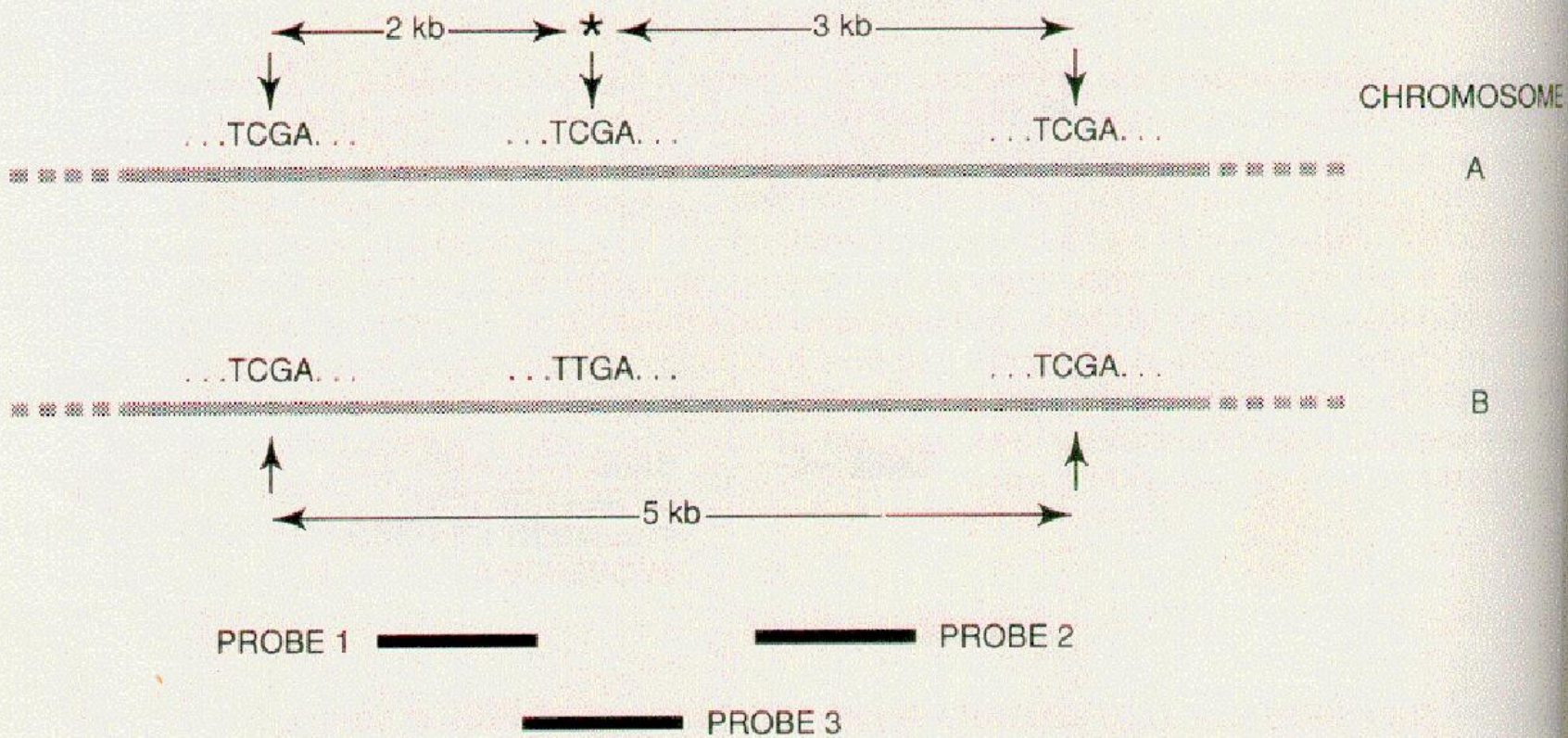
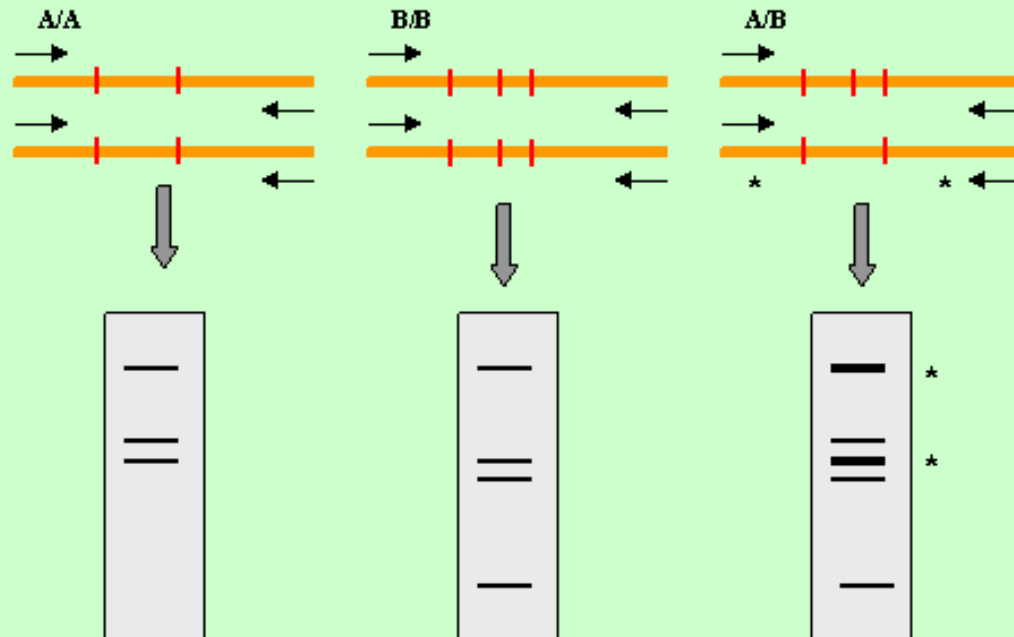


Figure 5.9. Restriction fragment length polymorphism (RFLP). Chromosomes A and B differ at a sequence marked by the *asterisk*, which results in the presence of a *TaqI* site for chromosome A but not for chromosome B; *TaqI* recognizes the sequence TCGA.

RFLP

assay: amplification – digestion – gel separation



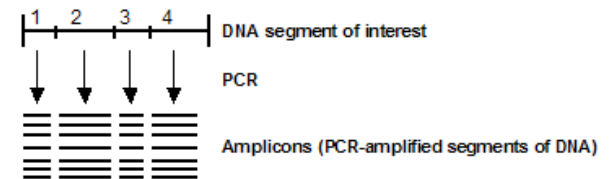
- Unique sequence primers are used to amplify a mapped DNA sequence from two related individuals, A/A and B/B, and from the heterozygote A/B. In the case of the heterozygote A/B, two different PCR products will be obtained, one which is cleaved three times and one which is cleaved twice.

mutation scanning (synonym: mutation screening):

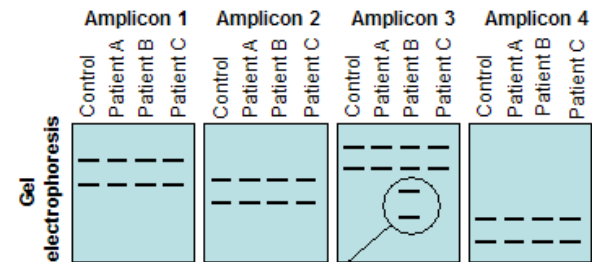
A process by which a segment of DNA is screened via one of a variety of methods to identify variant gene region(s). Variant regions are further analyzed (by sequence analysis or mutation analysis) to identify the sequence alteration

Step 1: Scanning.

A. PCR amplification of the DNA segment(s) of interest to create enough DNA for analysis



B. Visualization and comparison. Patient and control amplicons are compared using one of several different scanning methods (e.g., SSCP, CSGE, DGGE, DHPLC). In the SSCP example below, gel electrophoresis separates amplicons by mobility.



Interpretation: The additional bands and an abnormal migration pattern indicate that a DNA sequence alteration is present in amplicon 3 of Patient B.

Step 2: In mutation scanning, variant DNA segments (e.g., segments with altered mobility in the SSCP example) may be subjected to further testing, such as sequence analysis, to identify the sequence alteration (mutation).

Some Clinical Implications

- Mutation scanning is used when mutations are distributed throughout a gene, when most families have different mutations, and when sequence analysis would be excessively time-consuming due to the size of a given gene.
- Mutation scanning may cover the entire gene or select regions.
- The sequence alteration identified in a segment of DNA may be a benign variant (polymorphism), a disease-causing mutation, or an alteration of undetermined significance.

Types of sequence alterations that may be detected:

- Pathogenic sequence alteration reported in the literature
- Sequence alteration predicted to be pathogenic but not reported in the literature
- Unknown sequence alteration of unpredictable clinical significance
- Sequence alteration predicted to be benign but not reported in the literature
- Benign sequence alteration reported in the literature

Possibilities if a sequence alteration is not detected

Patient does not have a mutation in the tested gene (e.g., a sequence alteration exists in another gene at another locus)

Patient has a sequence alteration that cannot be detected by sequence analysis (e.g., a large deletion)

Patient has a sequence alteration in a region of the gene (e.g., an intron or regulatory region) not covered by the laboratory's test

B.) Indirect testing

historically first type of DNA diagnostic method
most of the mendelian diseases went through a phase of gene tracking
and moved on to direct test once the genes were cloned
with some diseases, even though the gene has been cloned, mutations are hard to find
mutations are scattered widely over a large gene
the existence of homologous pseudogenes
the lack of mutational hot spots

never confirm clinical diagnosis!

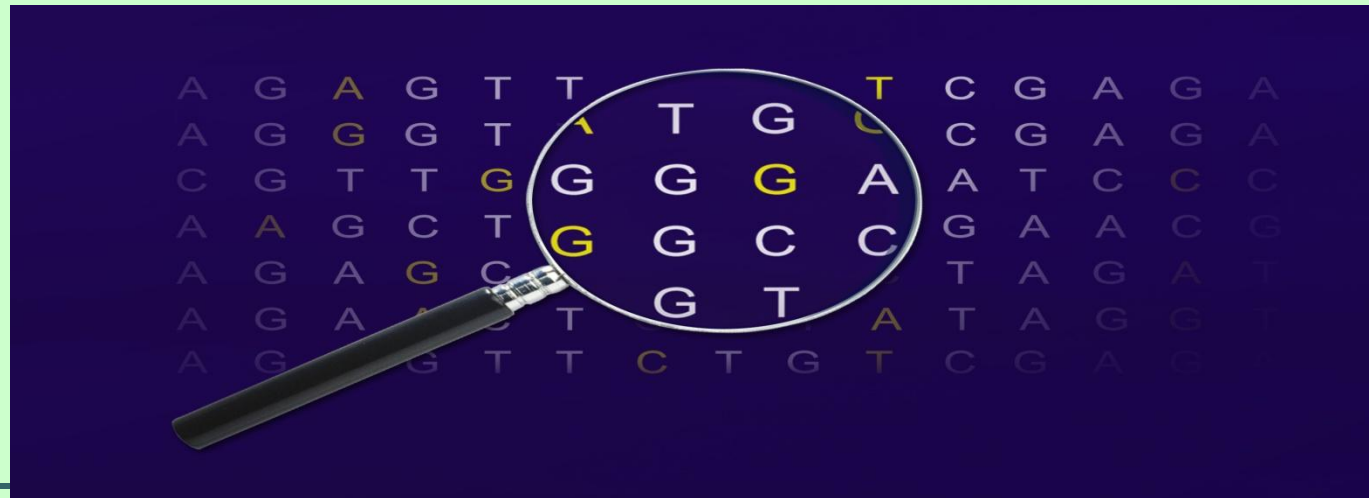
linkage analysis: (synonym: **indirect DNA analysis**) Testing DNA sequence polymorphisms (normal variants) that are near or within a gene of interest to track within a family the inheritance of a disease-causing mutation in a given gene

DNA sequence polymorphisms

- Single nucleotide polymorphism (SNP) – substitution of bases. In genome approx. 30 mil. SNP
- Minisatellite (VNTR) consist of repetitive, generally GC-rich, variant repeats (> 6bp) that range in length from 10 to over 100bp, these variant repeats are tandemly intermingled
- **Microsatelite – Short Tandem Repeats (STR)** consist of short sequence typically from 2 to 6 nucleotides long tandemly repeated several times (2 – 100x), and characterised by many alleles

Use of polymorphic regions

- ❑ Identification of persons/samples DNA
- ❑ paternity testing (VNTR, STR)
- ❑ Undirect diagnostics of monogenic diseases
- ❑ Searching of new genes
- ❑ SNP and multifactorial diseases

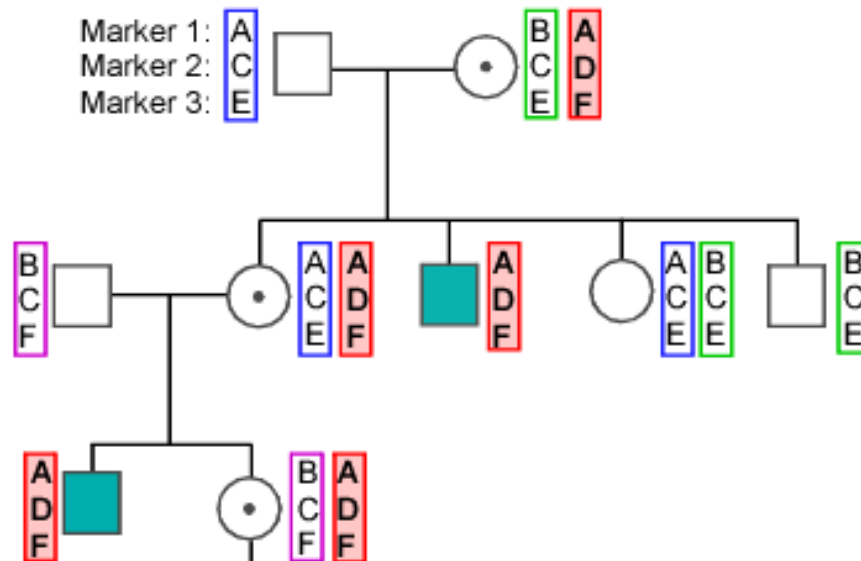


The three steps of linkage analysis

- **Establish haplotypes:** Multiple DNA markers lying on either side of (flanking) or within (intragenic) a gene-region of interest are tested to determine the set of markers (haplotypes) of each family member.
- **Establish phase:** The haplotypes are compared between family members whose genetic status is known (e.g., affected, unaffected) in order to establish the haplotype associated with the disease-causing allele.
- **Determine genetic status:** Once the disease-associated haplotype is established, it is possible to determine the genetic status of at-risk family members.

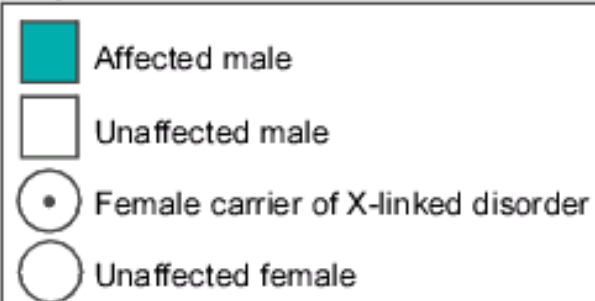
Example: Linkage analysis for prenatal diagnosis of an X-linked disorder.

- Establish haplotype:** In this example, three markers on the X chromosome are tested. At marker one, the two possible variations are: **A** and **B**; at marker two: **C** and **D**; and at marker three: **E** and **F**. The sets of letters (i.e., haplotypes) next to each individual represent the variants they have for each variant. Women have two X chromosomes and thus have two haplotypes of X-chromosome markers; men have a single X chromosome and thus have only one haplotype of X-chromosome markers.
- Establish phase:** It can be inferred that the haplotype **ADF** is linked to the mutated allele on the X chromosome.



- Determine genetic status:** Male fetus inherited from the mother the X-chromosome haplotype linked to the normal allele in this family and, therefore, is predicted to be unaffected.

Key



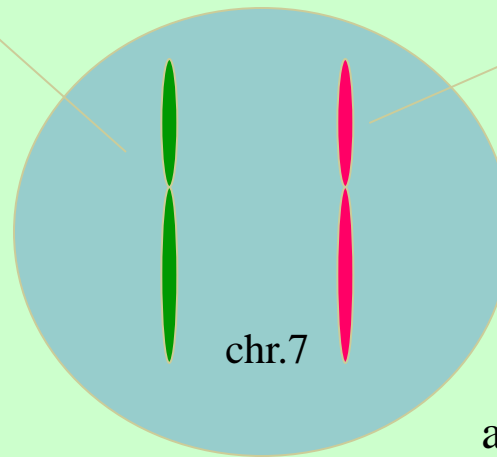
Indirect DNA analysis

gene CFTR - intron 8 - polymorphic site (CA)_n

GTATCACACACATTCGG

from father

from mother



allele A1:

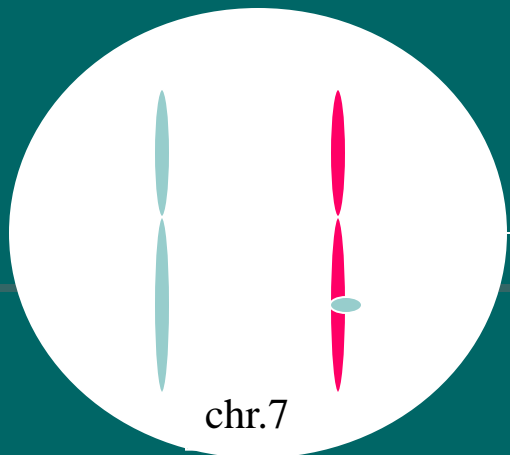
-----GTATCACATTCGG-----

the length of this allele is 130 bp

allele A2:

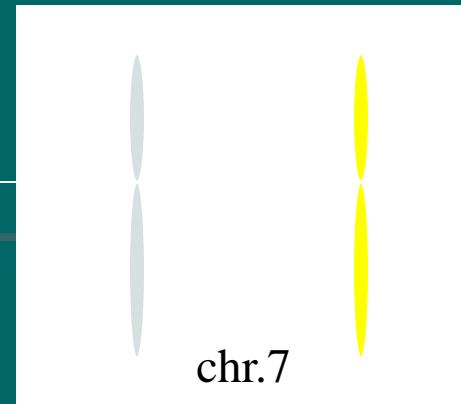
-----GTATCACACACATTCGG-----

the length of this allele is 134 bp



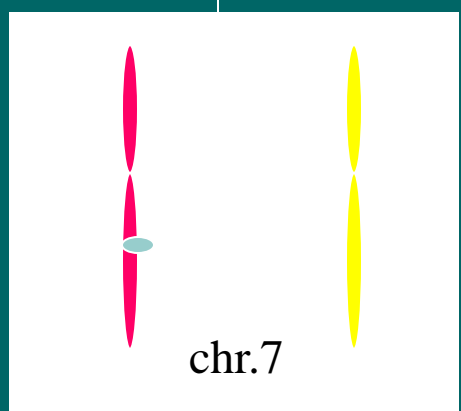
dF508 / non

A1 / A3



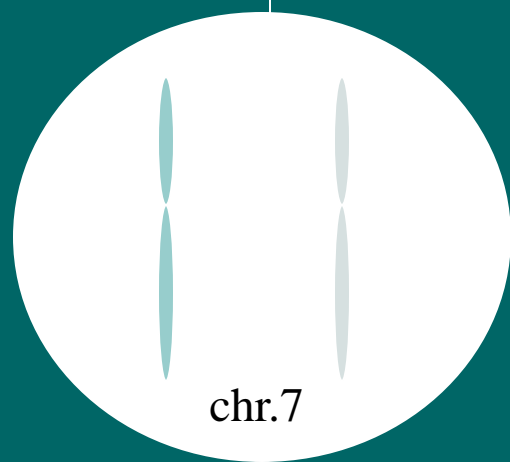
non / ?

A1 / A1



dF508 / ?

A1 / A1



non / non

A1 / A3 non-informative

○ mutation in CFTRgene



Linkage analysis is often used when *direct DNA analysis is not possible* because the gene of interest is unknown or a mutation within that gene cannot be detected in a specific family.

In most instances, the haplotype itself has no significance; *it has meaning only in the context of a family study.*

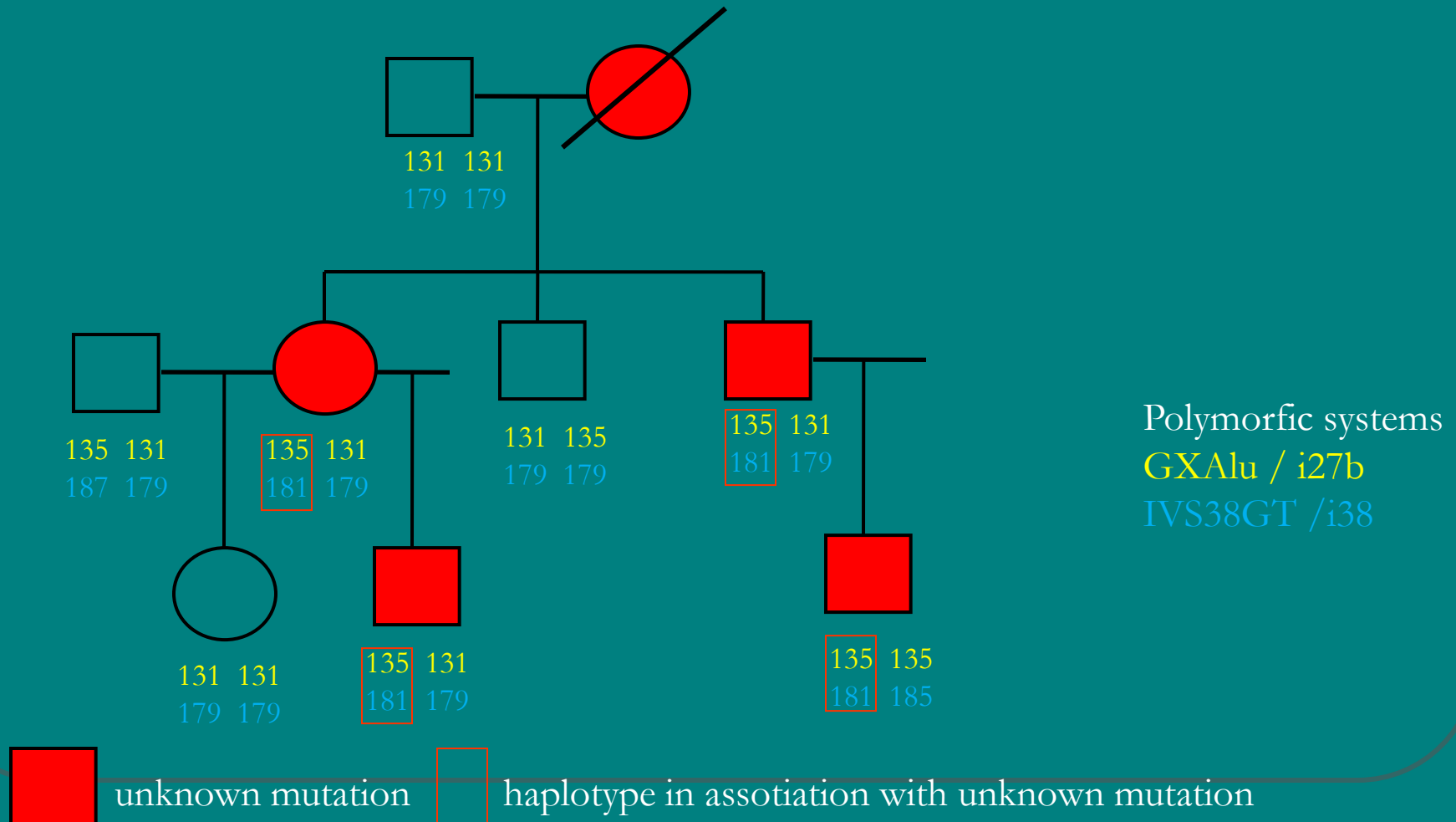
The accuracy of linkage analysis is dependent on:

- The accuracy of *the clinical diagnosis* in affected family member(s).
- The *distance between the disease-causing mutation and the markers*. Linkage analysis may yield false positive or false negative results if recombination of markers between maternally and paternally-inherited chromosomes occurs during gamete formation. The risk of recombination is proportional to the distance between the disease-causing mutation and the markers. The *risk of recombination* is lowest if *intragenic markers* are used.

The *informativeness of genetic markers* in the patient's family. If the DNA sequence for a given variant differs on the maternally-inherited and paternally-inherited chromosomes, that marker is **informative**. If the DNA sequence for a given variant does not differ on the two chromosomes, that marker is not **informative**.

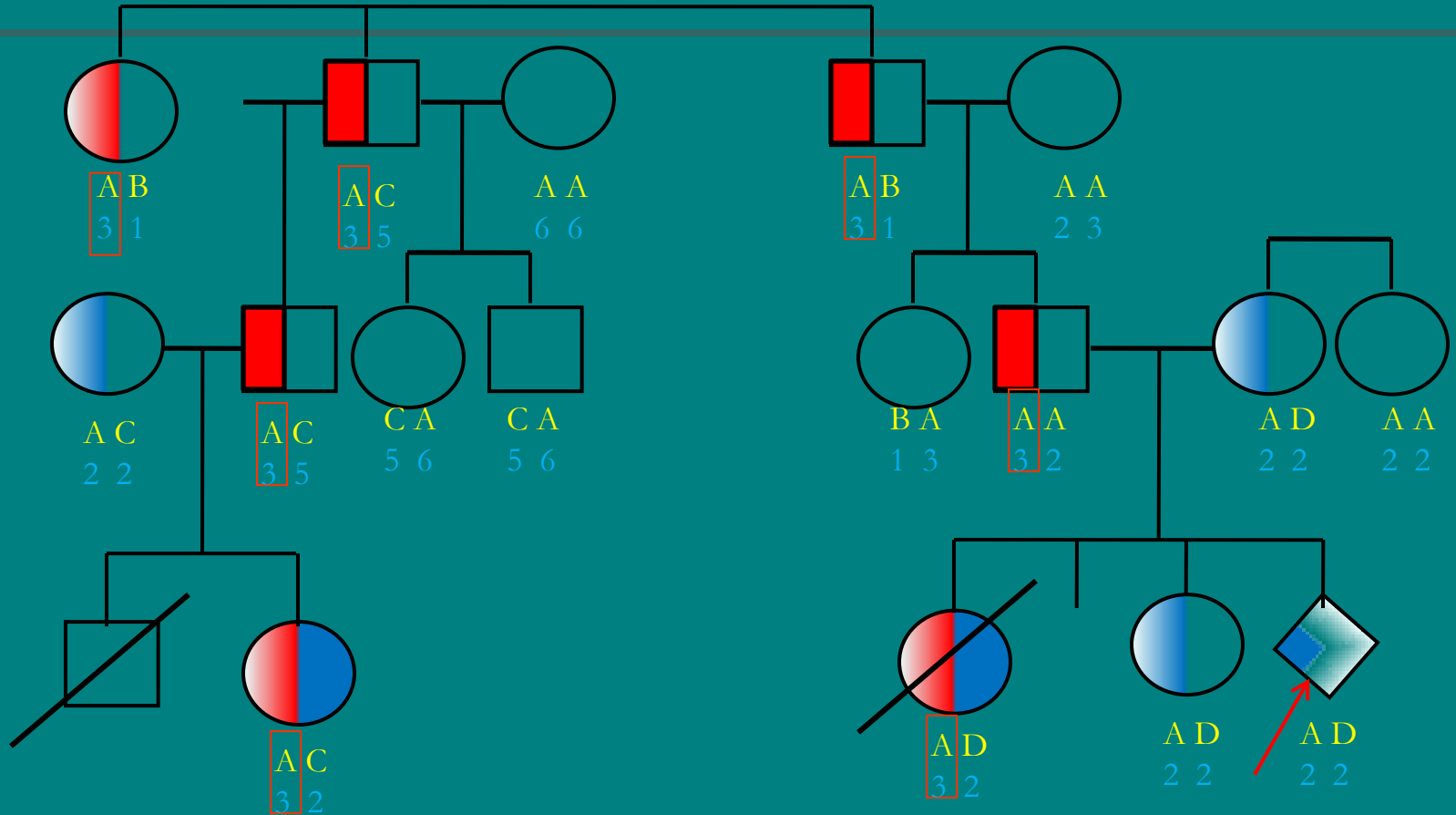
Indirect diagnosis – Neurofibromatosis type 1

Autosomal dominant



Indirect diagnosis – cystic fibrosis

Autosomal recessive

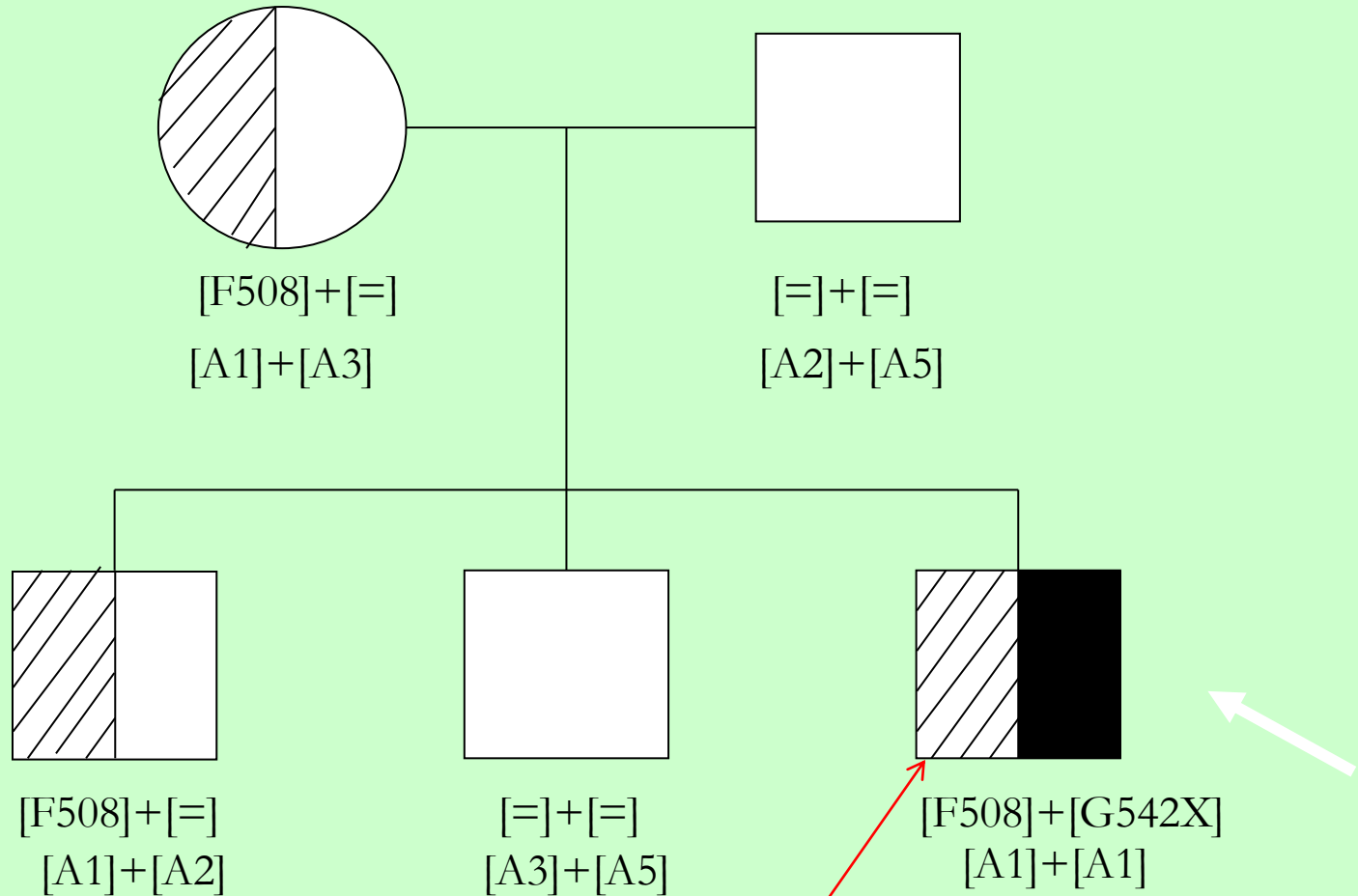


Polymorphic systems IVS17BTA aley 1 -6

IVS8BTA aley A - D

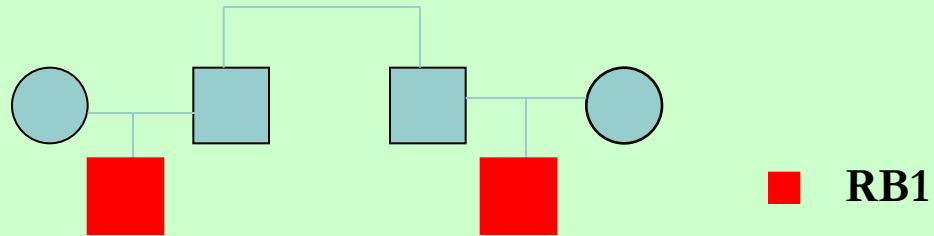
haplotype in association with unknown mutation

Indirect diagnosis – cystic fibrosis



de novo mutation

Retinoblastoma



Mutation analysis of Rb1 was done

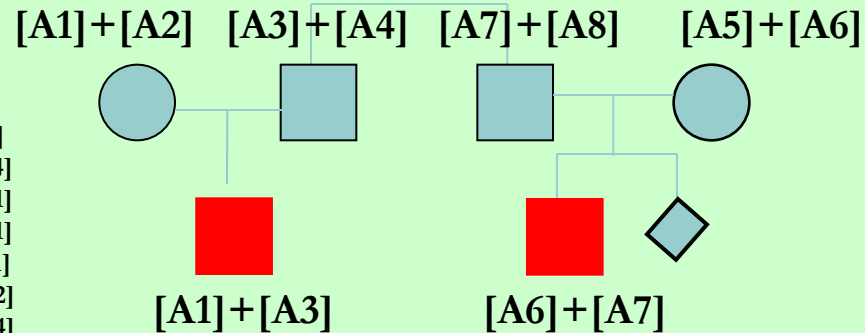


Pathology in Rb1 gene was not detected

Retinoblastoma - Indirect diagnostics

Polymorphic markers

- extragene (DS13S 1307, DS13S 272, DS13S 164)
- intragene (Rb1.20B)



A1: DS 13S 1307 [141] DS 13S 272 [133] DS13S 164 [179] Rb1.20B [3]
 A2: DS 13S 1307 [151] DS 13S 272 [133] DS13S 164 [188] Rb1.20B [4]
 A3: DS 13S 1307 [139] DS 13S 272 [127] DS13S 164 [179] Rb1.20B [1]
 A4: DS 13S 1307 [139] DS 13S 272 [133] DS13S 164 [186] Rb1.20B [1]
 A5: DS 13S 1307 [139] DS 13S 272 [131] DS13S 164 [188] Rb1.20B [1]
 A6: DS 13S 1307 [126] DS 13S 272 [133] DS13S 164 [188] Rb1.20B [2]
 A7: DS 13S 1307 [126] DS 13S 272 [129] DS13S 164 [188] Rb1.20B [4]
 A8: DS 13S 1307 [139] DS 13S 272 [127] DS13S 164 [178] Rb1.20B [5]

Haplotype with pathology cannot be established

Explanation:

- occurrence of mutation in another system of cell division and growth regulation
- nonhereditary form of retinoblastoma in both cousins