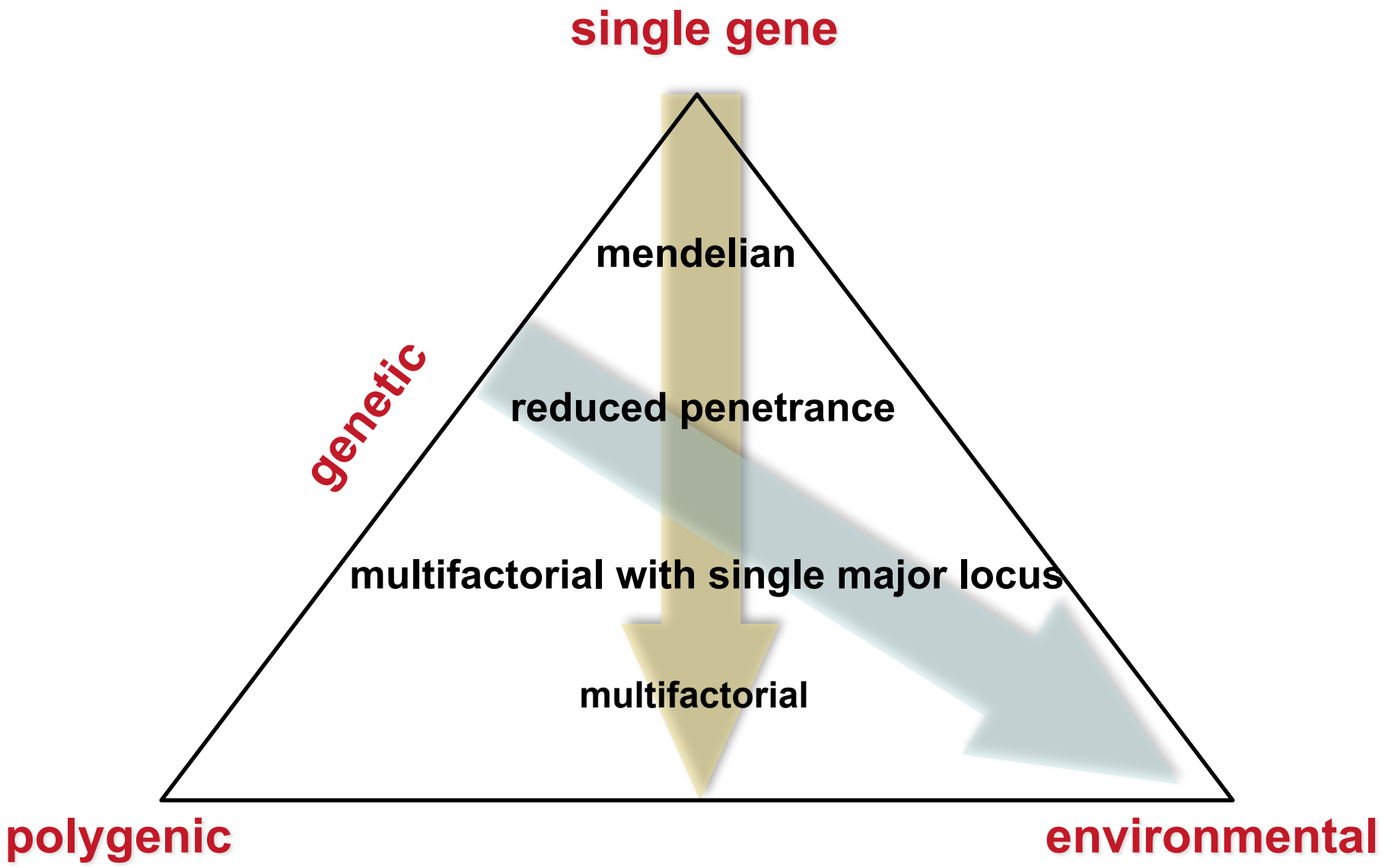
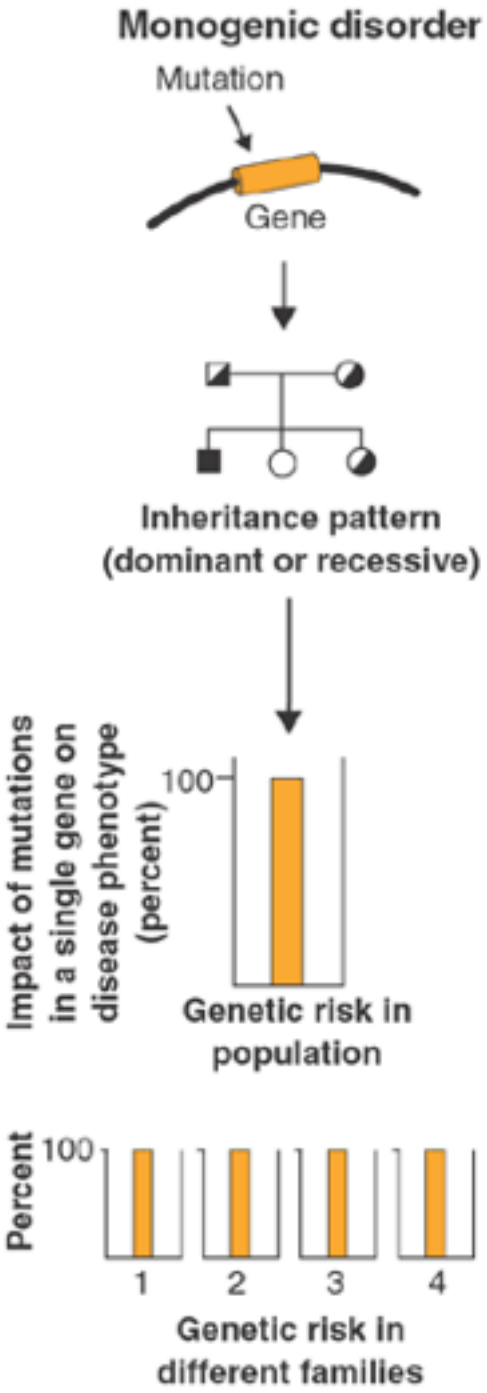
THE SPECTRUM OF HUMAN CHARACTERS

few characters are purely mendelian, purely polygenic or purely environmental





MONOGENIC DISEASES

In monogenic diseases, mutations in a single gene are both **necessary** and **sufficient** to produce the clinical phenotype and to cause the disease.

The impact of the gene on **genetic risk** for the disease is the **same** in all families.

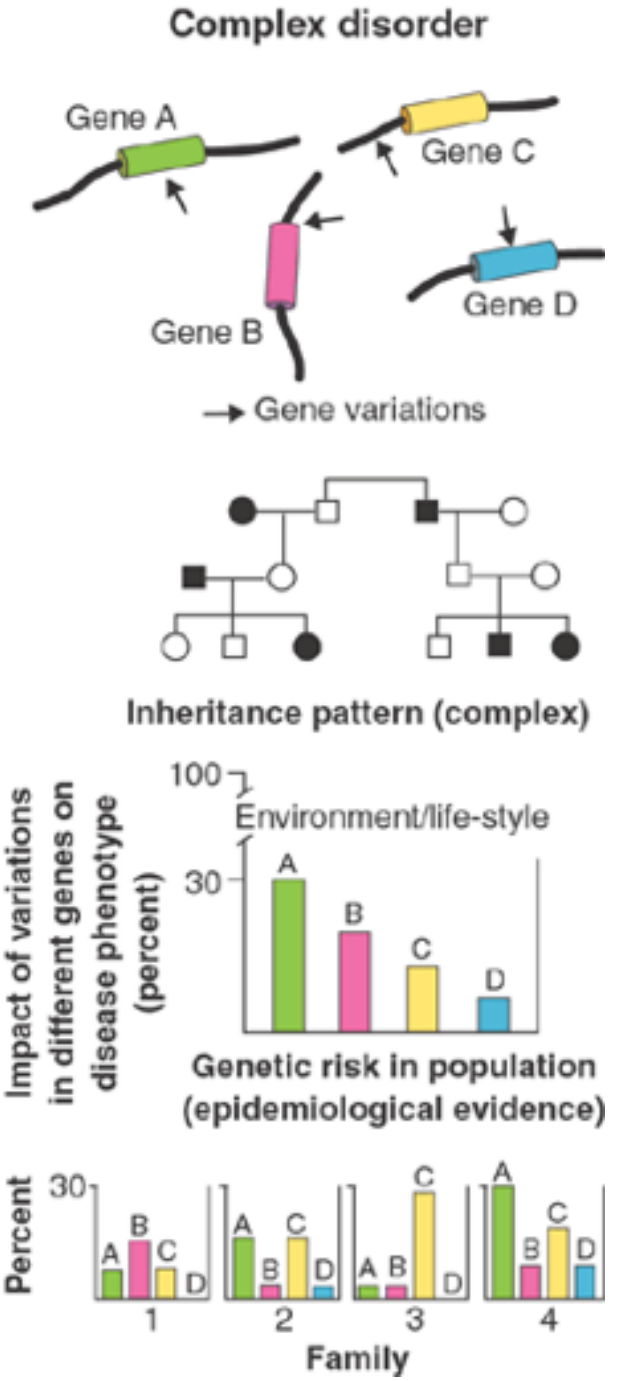
From Leena Peltonen and Victor A. McKusick; Science 2001

COMPLEX DISEASES

In complex disorders with **multiple causes**, variations in **a number of genes** result in a **genetic predisposition** to a clinical phenotype.

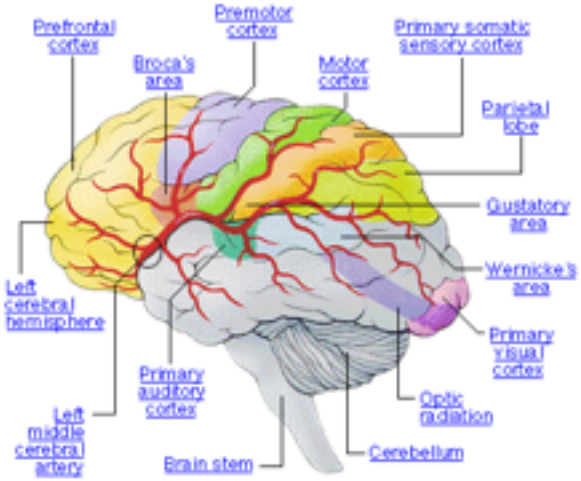
Pedigrees reveal **no Mendelian inheritance**, and gene mutations are often **neither sufficient nor necessary** to explain the disease phenotype.

From Leena Peltonen and Victor A. McKusick; Science 2001

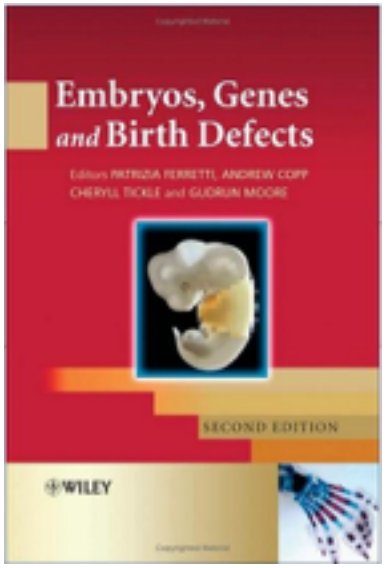


Complex Diseases

Cardio-vascular diseases

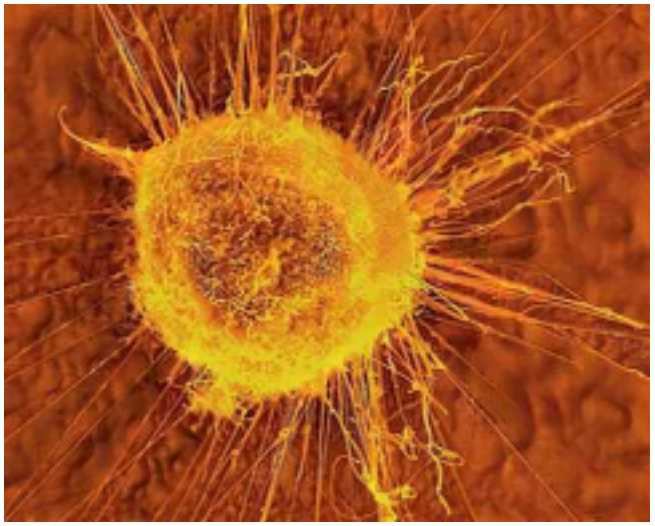


Psychiatric and degenerative disease of brain



Non-syndromic birth defects

Cancer

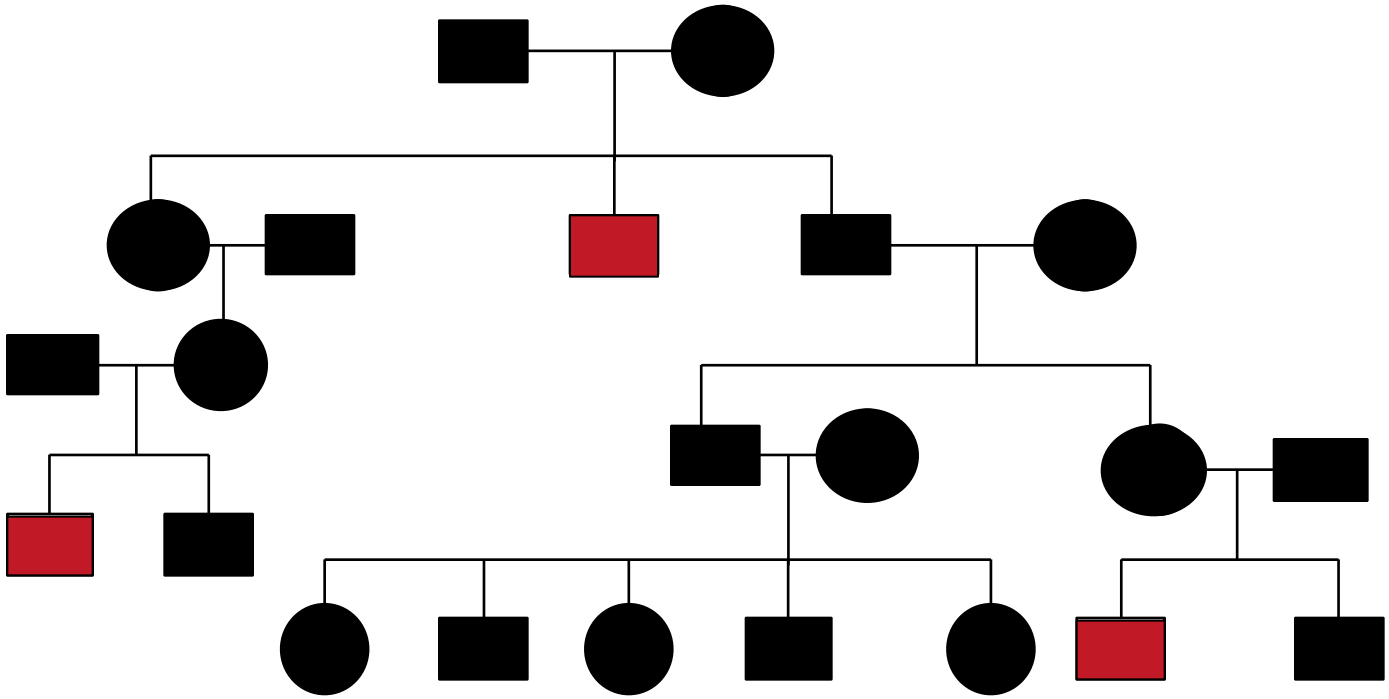


Complex Diseases

- Selected Characteristics
  - Can aggregate in families but **do not segregate** in mendelian fashion
  - **Multigenic**
  - Multiple **environmental factors**
  - Phenotypic and Genetic heterogeneity
  - Incomplete (variable) penetrance
  - **Complex interplay** between genetic and environmental factors

Complex Diseases

Can aggregate in families but do not segregate in Mendelian fashion



Penetrance

Probability of disease given the genotype

penetrance = p(D=1|genotype)

- Examples: assume 3 possible genotypes at a locus: (AA,Aa,aa)
  p(D|AA) = 1, p(D|Aa) = 1, p(D|aa) = 0
  p(D|AA) = 1, p(D|Aa) = 0, p(D|aa) = 0
  p(D|AA) < 1, p(D|Aa) <1, p(D|aa) = 0

- Incomplete penetrance: penetrance < 1

The frequency with which a gene manifests itself in the phenotype of the carriers

Complex Diseases

$$\text{Relative risk}: \lambda = \frac{\text{risks of disease among relatives of affected}}{\text{risks of disease in general population}}$$

- Complex diseases tend to have low λ
- Large λ suggest a major disease genes
- Small λ may indicate many genes, each contributing a small effect

λ_s = 500 Cystic Fibrosis

λ_s = 15 Type I Diabetes

λ_s = 1 – 4 Hypertension

Cancer = 2.3 millions deaths/year in industrial countries;

500.000 in the United States

Cancer is a genetic disease

Hereditary Tumors — One pathogenic variant in one gene in all cells (germinal mutations)

Familial Tumors — Some variants in some genes in all cells (germinal variants)

Sporadic Tumors — Some pathogenic variants in some genes in some cells (specific tissues)

Accumulation of a high number of tissue-specific mutations (progression)

Genes/environment interplay is necessary for the progression

Genes associated with cancers…

Tumor-suppressor genes

- "gatekeepers" — Cell-cycle and proliferation control, differentiation and apoptosis
- "caretakers" — DNA repair control
- "landscape" — Microenvironment control

Oncogenes

Code for growth factors, growth factor receptors, signal transduction proteins

High penetrant genes

"Strong" pathogenic phenotype

Variants with functional effect

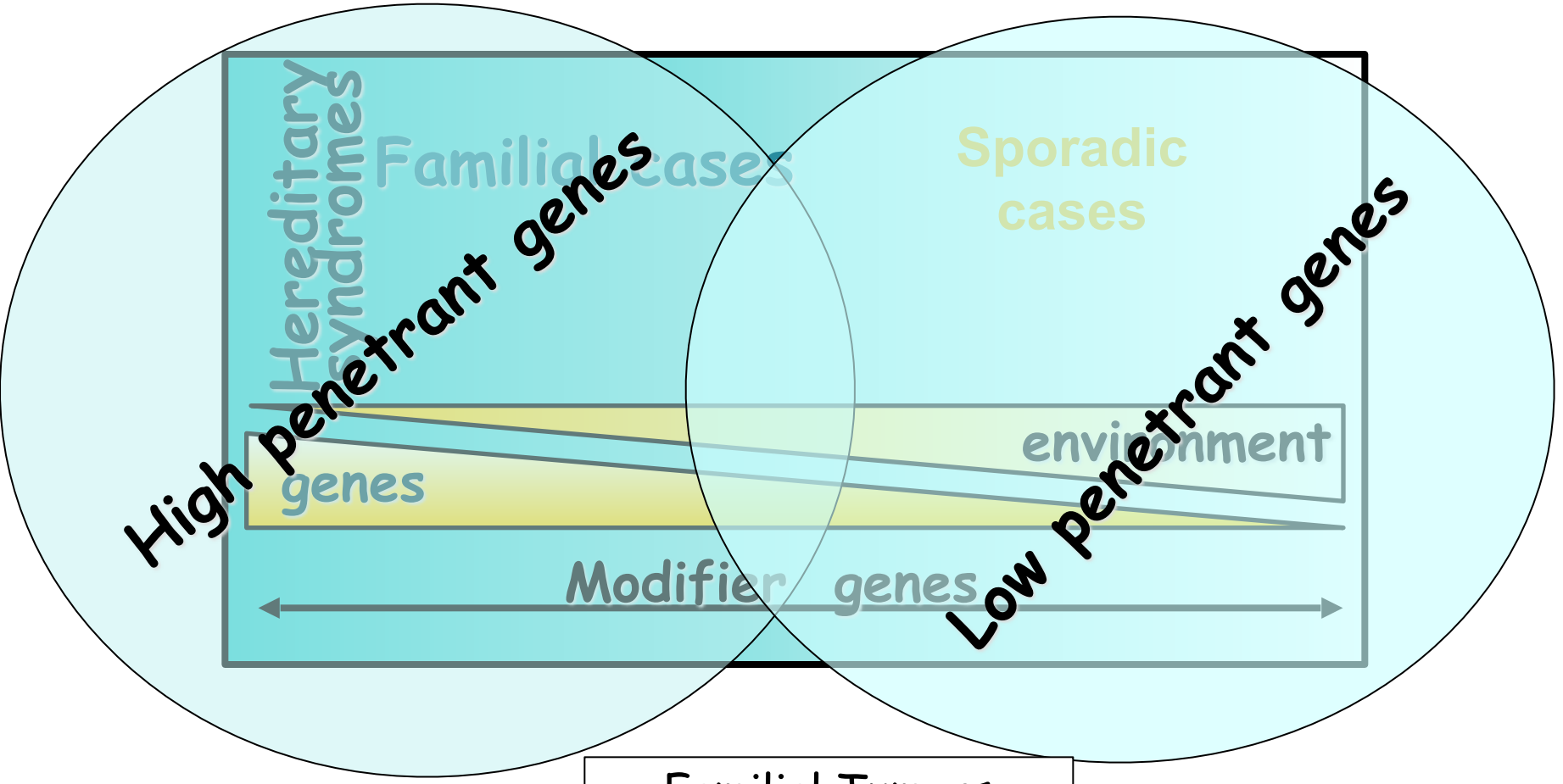
Low penetrant genes

"Mild or non evident" pathogenic phenotype

Variants with mild or non evident effect

Hereditary tumors or monogenic syndromes

5%



Familial Tumors

10-20% ?

Sporadic tumors

80-85% ?

Hereditary tumors

High penetrant monogenic syndromes

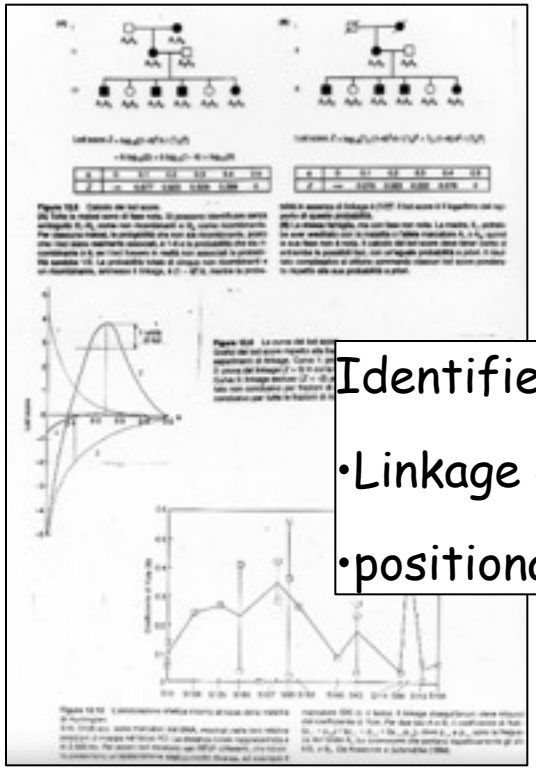
Cancer recognized as hereditary disease at the of the 19th century

High penetrant genes

APC
VHL
WT1
RB1
NF1
NF2
p53
p16
BRCA1
BRCA2

hMSH2
hMLH1
hPMS1
hPMS2
ATM
XPA BLM
MUTYH

RET
MET



Identified by:

- Linkage analysis
- positional cloning

Most of the susceptibility cancer syndromes are dominant in the families

Most of them follow Knudsen's "two hits" hypothesis

Hereditary tumors

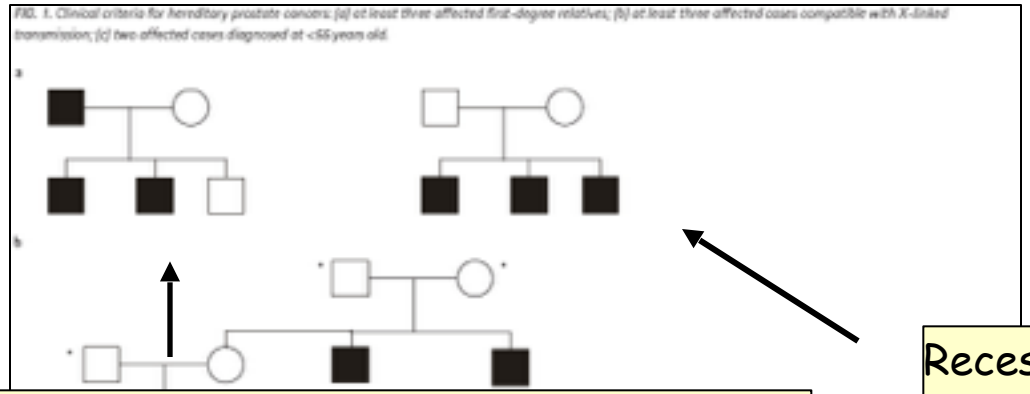
High penetrant monogenic syndromes

Several individuals with rare tumors

Individuals with different types of cancers

Families with several individuals affected by the same type of cancers

Families with cancer susceptibility



Dominant autosomic transmission

One affected allele è sufficient to acquire the susceptibility to develop the disease

FAP, HNPCC, breast and ovarian susceptibility genes,etc

Recessive autosomic transmission

Both affected alleles are necessary to develop the disease

Some syndromes associated with DNA repair genes, such as the recessive polyposis

Familial tumors

Polygenic syndromes

Predisposition associated with the interaction of different genes with the environment

High penetrant genes

Low penetrant genes

Variants in high penetrant genes

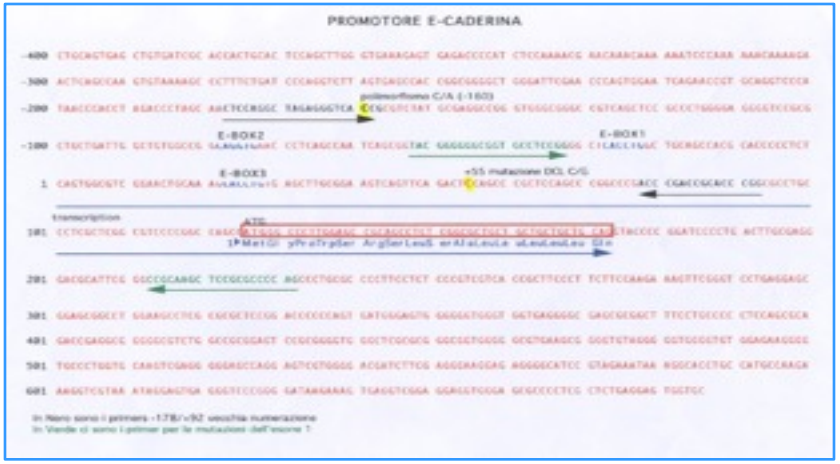
Variants in genes involved in common tumorigenic nolecular pathways (DNA repair, detoxification, immunity etc.)

Most variants are single nucleotide polymorphisms (SNPs)

Genomic sequence variants in coding and non-coding sequence

SNPs Frequency= 1/1000 bp

Total genomic SNPs = 3 x 10^6



Familial tumors

Polygenic syndromes

Families in which there is no evidence of a clear mendeleian inheritance

Polygenic risk

Low penetrant genes

To identify these genes linkage analysis is not useful

Use of inbred mice

Studies of homozygous mice in controlled environment

So far identified about 50 low penetrant genes in different tumor types

Homo sapiens ?

Use of association studies

Case-control studies

Identified only very few genes :

- DNA repair (XRCC1, XPD)

- detoxification (cyt. P450, GSST1)

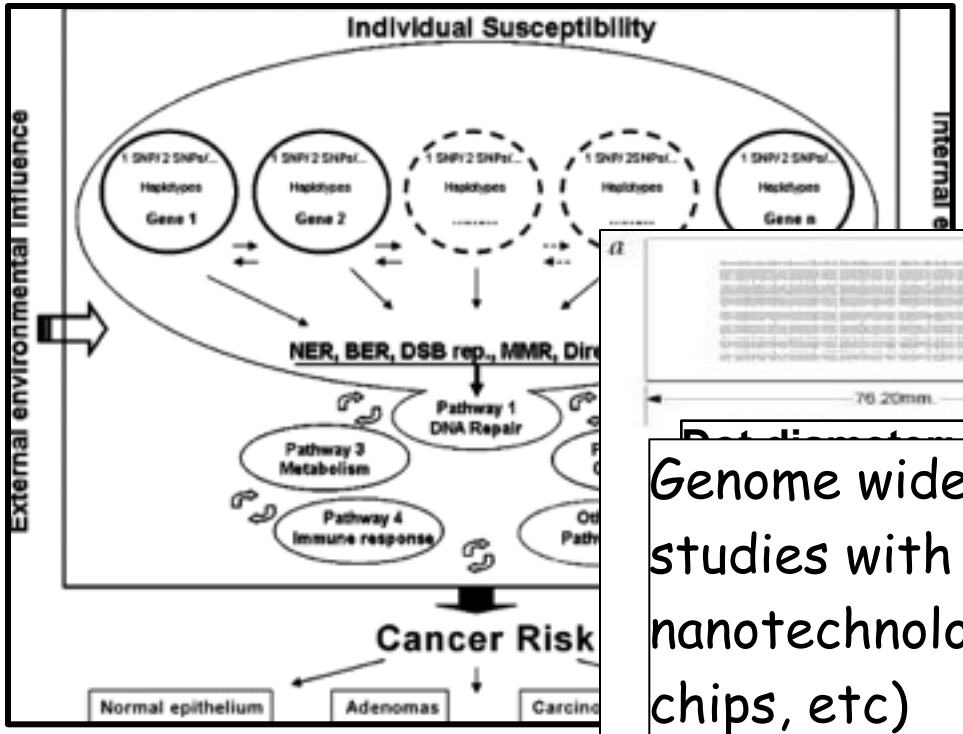
- Inflammation response (PLA2, COX1/2, TNF-a)

Sporadic Tumors

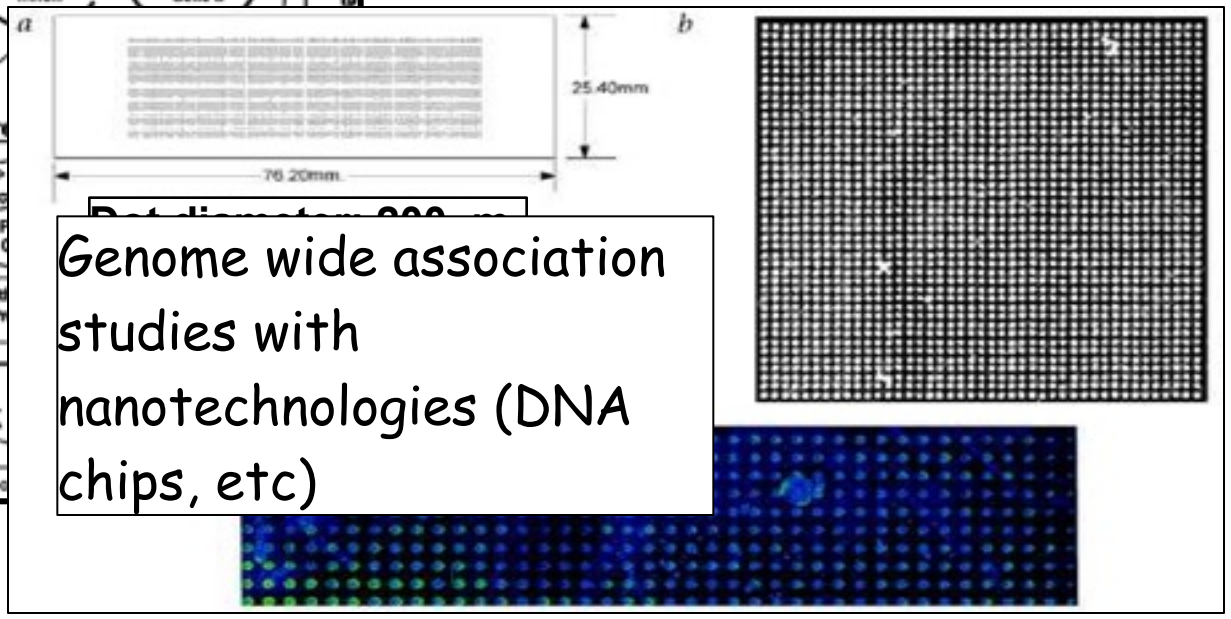
Interaction of the genetic background with environment

Relatives of cancer patients show 2/4 times higher risk of developing the same tumors

In industrial countries is difficult to point out the genetic background because enviroment background is high and confouding



Complex association studies



Genome wide association studies with nanotechnologies (DNA chips, etc)

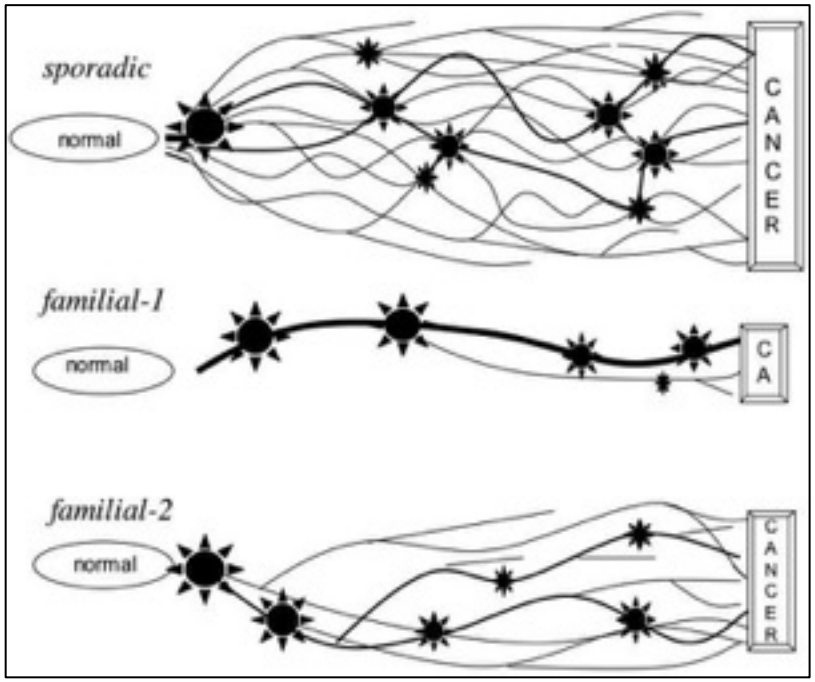
Growing debate over the nature of the genetic contribution to individual susceptibility to common complex diseases such as cancer

- **'Common Disease, Common Variant (CDCV)'**

Genetic variations with appreciable frequency in the population, but relatively low penetrance, are the major contributors to genetic susceptibility to common diseases

- **'Common Disease, Rare Variant (CDRV)'**

Multiple rare DNA sequence variations, each with relatively high penetrance, are the major contributors to genetic susceptibility to common diseases.

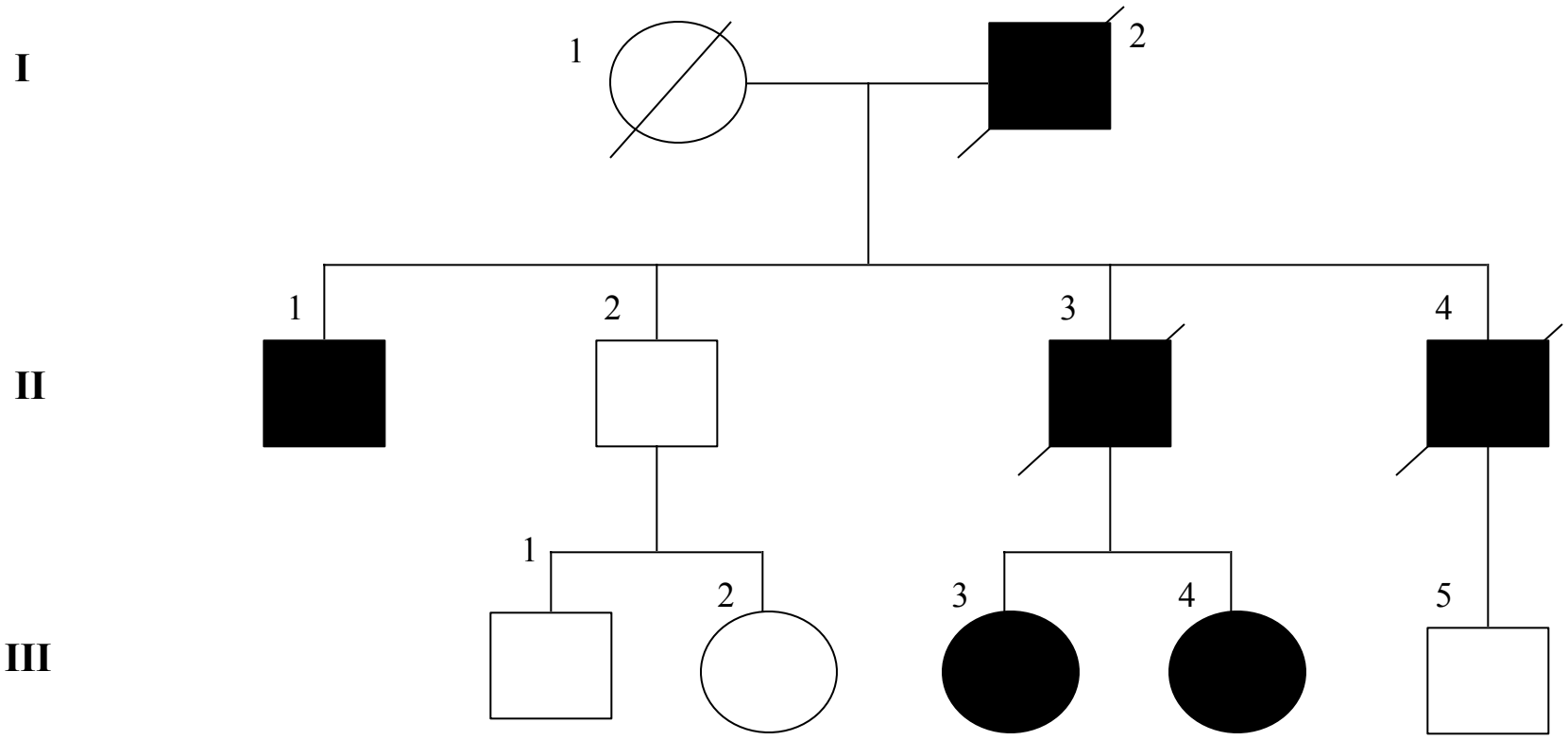


More than 90–95% of the heritable component of a disease has been left unexplained after extensive GWAS interrogation.

This suggests that individual common inherited variations are not likely to explain the majority of common chronic disease prevalence

GWA studies have reached their limits in the identification of common variations contributing to common diseases
This opens the door for the discovery of **multiple rare variations that contribute to common diseases** (or possibly other forms of genetic and epigenetic variation).

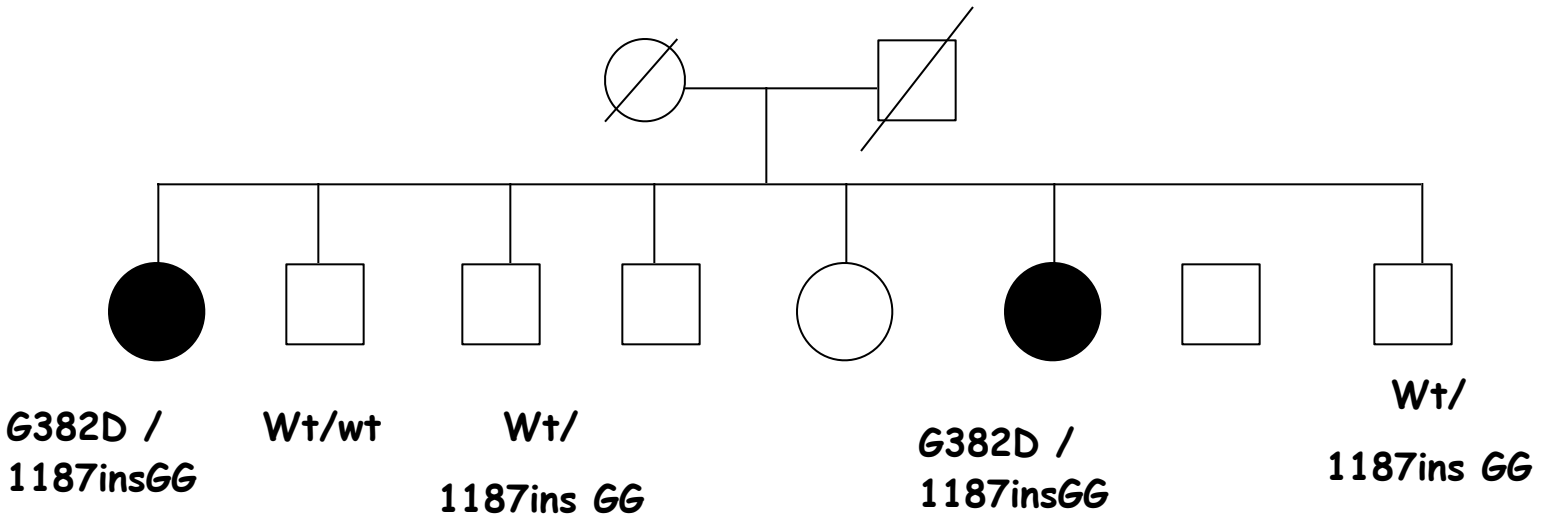
Autosomal dominant inheritance



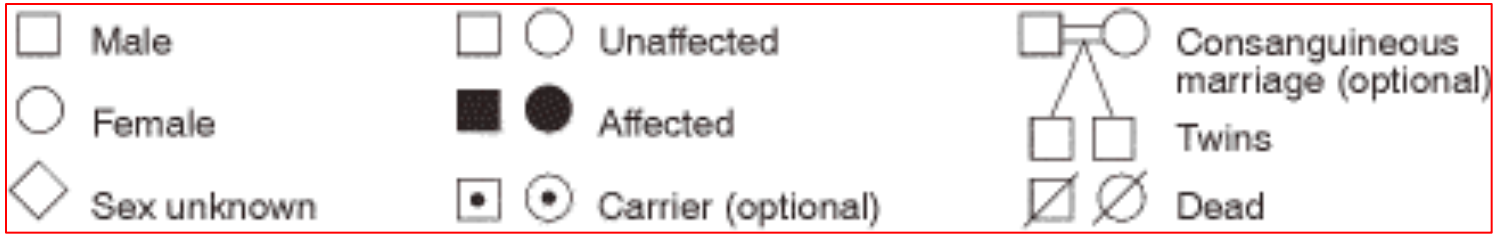
A character is **dominant** if it is manifested in heterozygous genotype (Aa)

Familial Adenomatous Polyposis Syndrome, Non-polyposis Hereditary Colorectal Cancer Syndrome, Susceptibility of the Breast and Ovarian Cancer Syndrome etc.

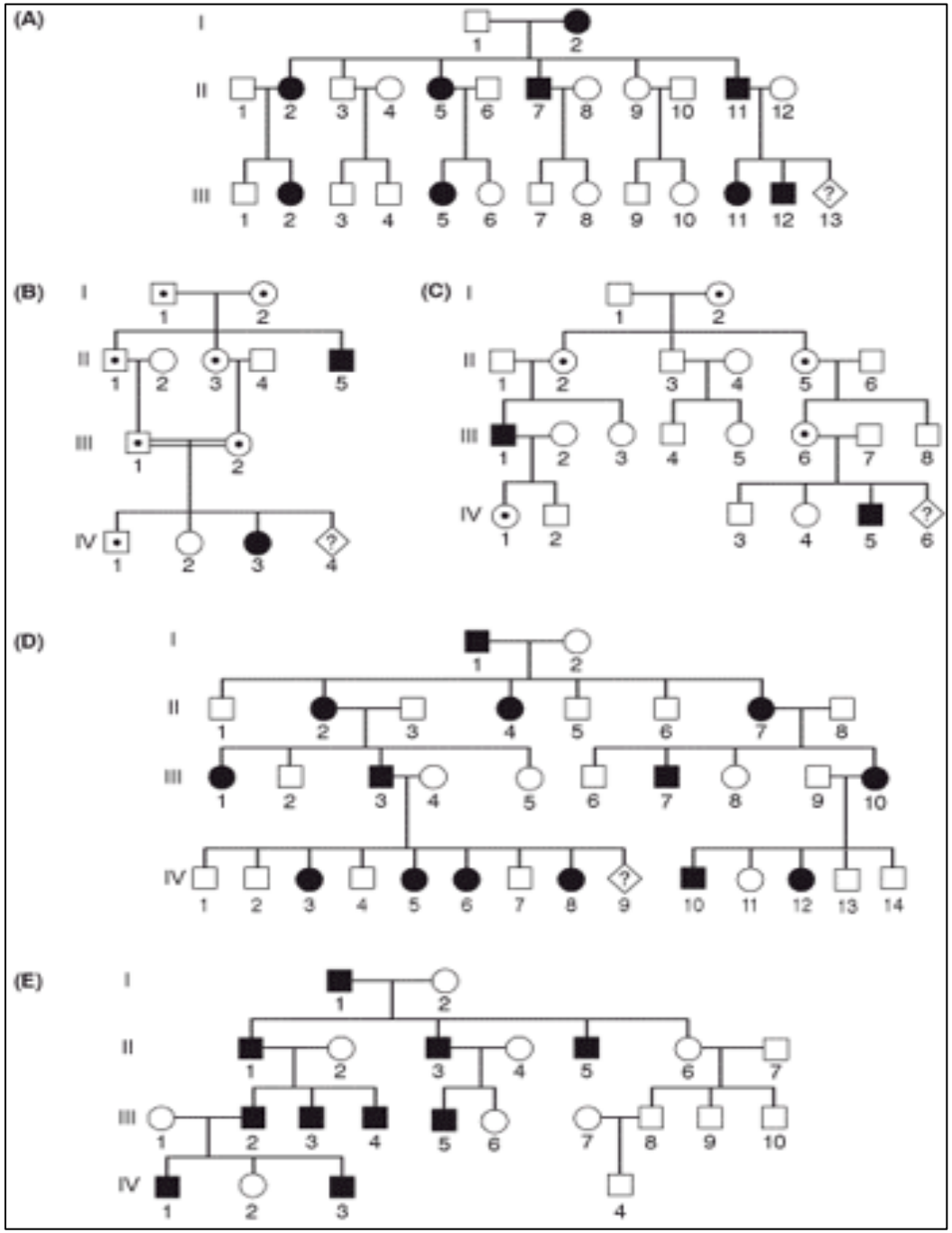
Autosomal recessive inheritance



A character is **recessive** when both the alleles are altered: some of the DNA repair syndromes, a type of recessive colorectal polyposis (MUTYH associated polyposis, etc).



- Autosomal dominant
- Autosomal recessive
- X-linked dominant
- X-linked recessive
- Y-linked
- Mitochondrial inheritance (non-medeleian matrilinear pedigree pattern)



DNA MUTATIONS

• LOSS OF FUNCTION

The protein product may have **reduced or no function**

- **Recessive phenotypes because for most genes the precise quantity is not crucial;**
- **This type of mutation generally affects tumor suppressor genes**
- **Sometimes the product of the mutated allele affects the product of the normal allele (p53, APC) dominant negative or haploinsufficiency**

• GAIN OF FUNCTION

The protein product may function in an **abnormal way**

- **This type of mutations generally affects oncogenes**
- **Dominant phenotype because mutations involve the escape from the normal control or doing something different**
- **The product of the mutant allele prevents the product of the normal allele of functionioning**

The distinction is sometimes difficult and the same gene can be affected by both loss or gain of function mutations !!!

Mutations can be quantitative or qualitative

SEQUENCE SITES

- Coding-sequence
- Non-coding sequence
- Gene-promoters
- Genomic repeated sequences

TYPES

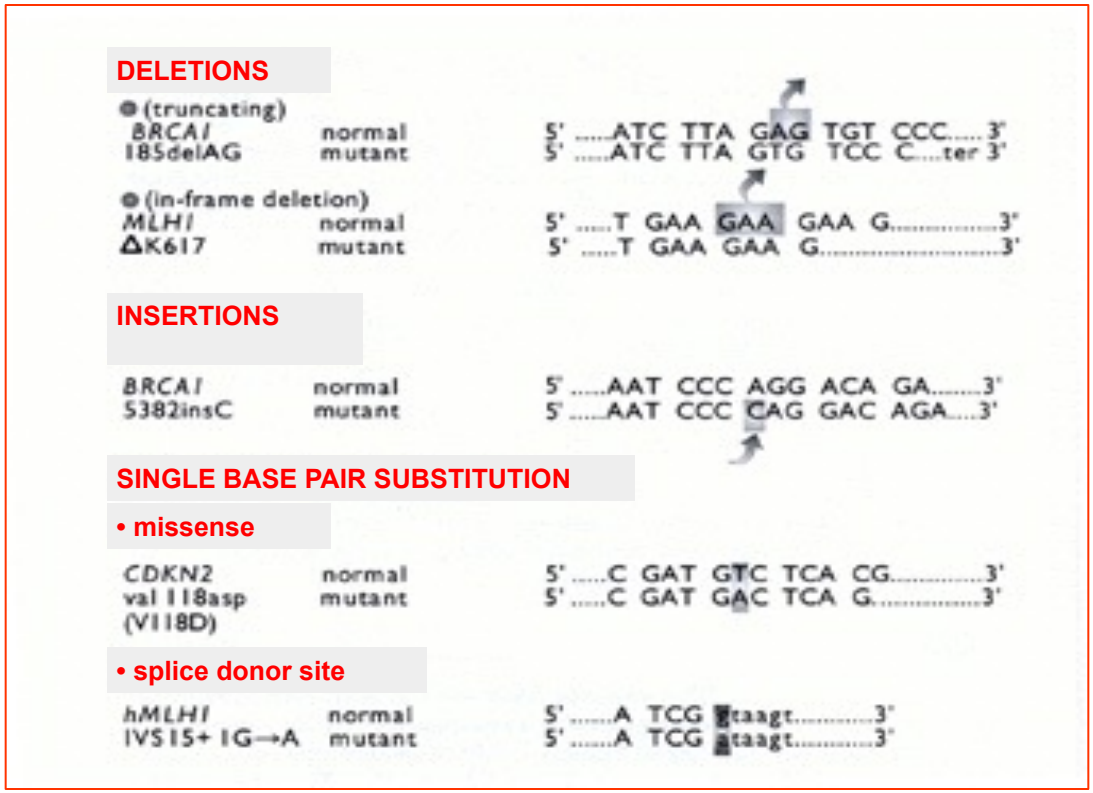
Structural chromosomal abnormalities:

deletion, amplification

translocation

inversion

insertion



MUTATION SCREENING

- Non-functional methods
  - SSCP
  - DGGE
  - DHPLC
  - DNA chips
- Functional methods
  - Protein Truncation Test (PTT)
  - In situ hybridization PCR

CHARACTERIZATION

- Sanger sequencing, pyrosequencing, mass array (MALDI-TOF), Next-Generation Sequencing(NGS)…

Mutational analysis based on the DNA ability of denaturation, separation and reassociation

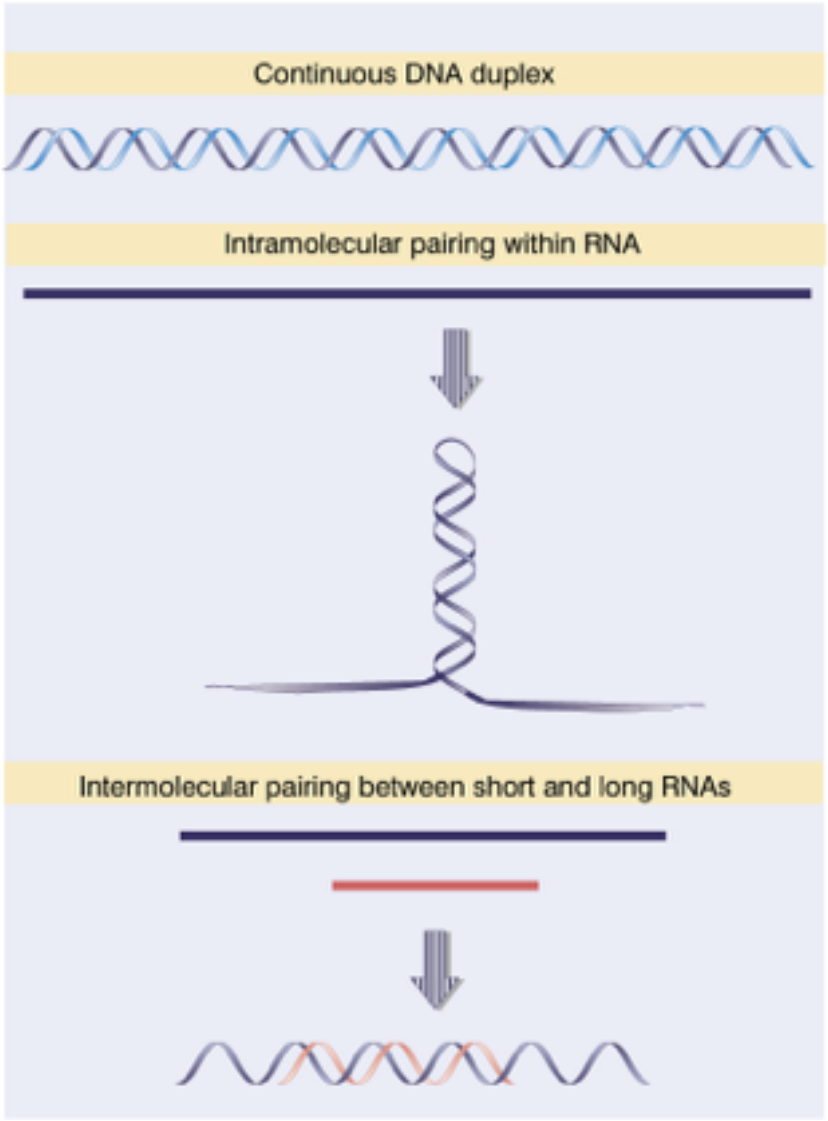
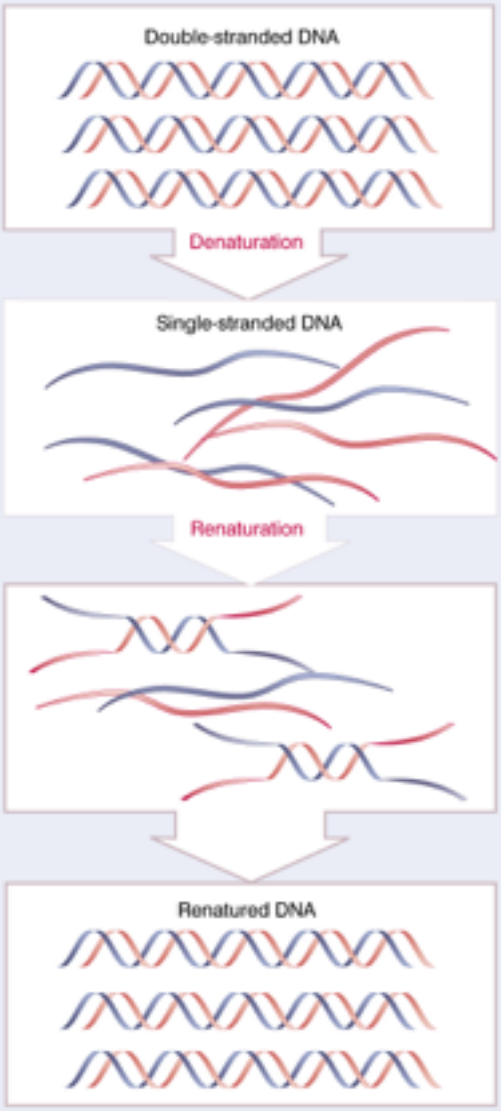
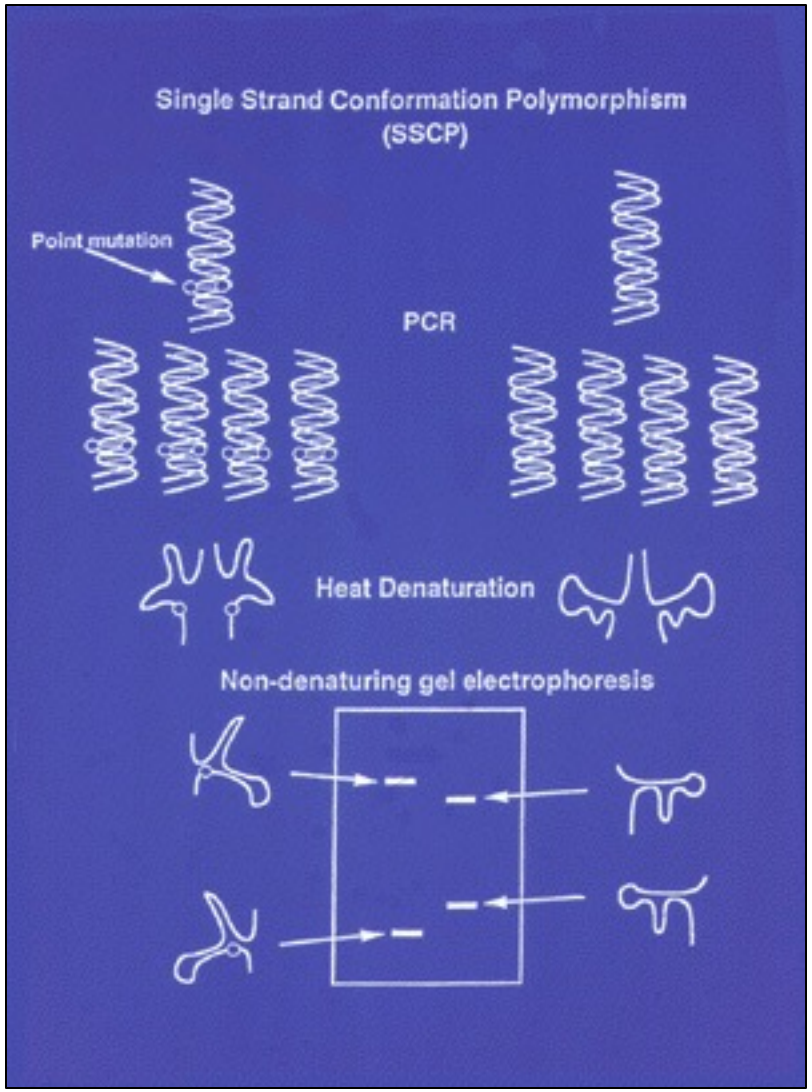


Figure 1.13 Denatured single strands of DNA can renature to give the duplex form.

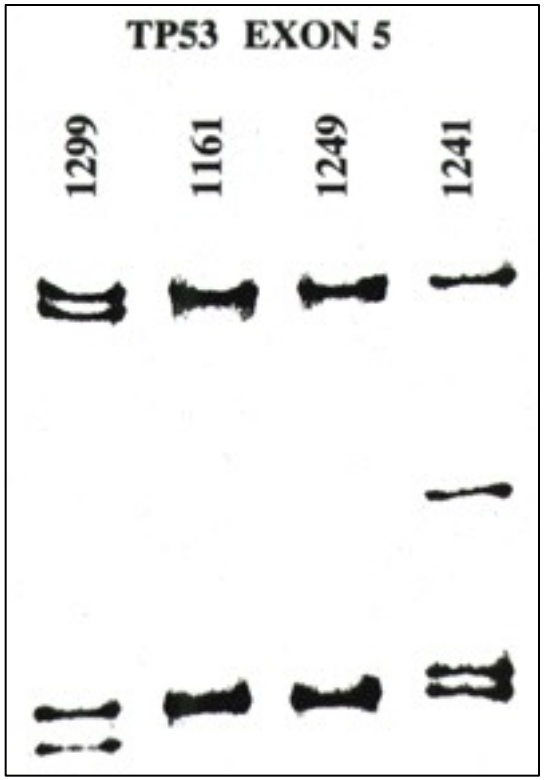


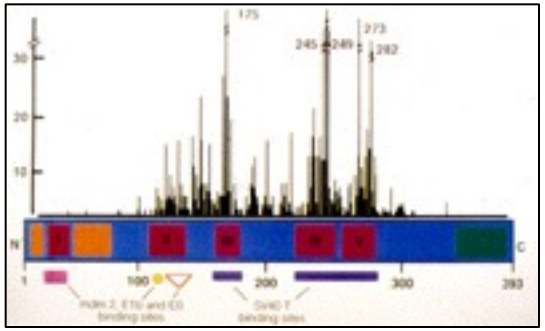
Single Strand Conformation Polymorphism

Methods



Results





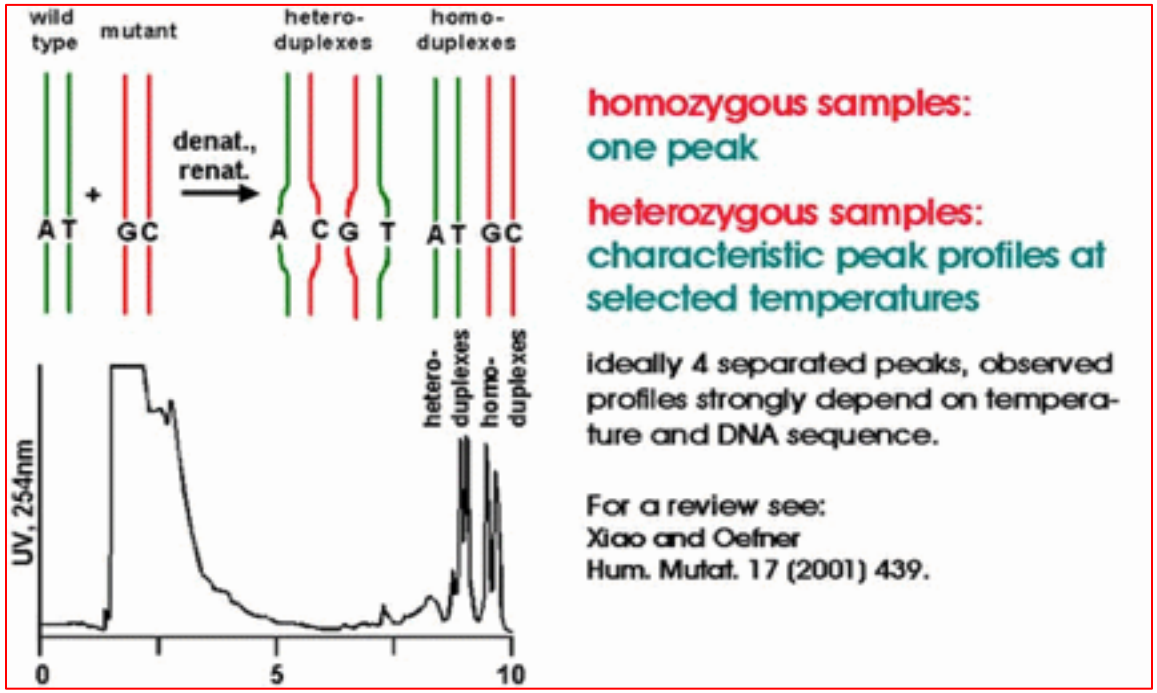
p53 mutation distribution

Denaturing High Performance Liquid Chromatography (DHPLC)

is a chromatographyc method for the detection of <u>DNA base substitutions, small deletions or insertions</u>

PCR amplification of a region with a hemizygous polymorphism gives two fragments corresponding to the **wild type and polymorphic alleles**

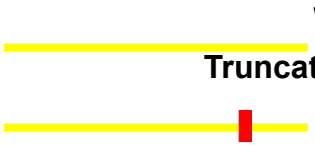
Denaturation – reassociation of this PCR product creates **hetero and homoduplexes** molecules



- **Heteroduplexes** are thermally <u>less stable</u> than homoduplexes and will be <u>resolved differently</u> by chromatography when subjected to a specific high temperature
- The mismatch will decrease the interaction with the column (matrix) and <u>reduce retention time</u> compared to the homoduplexes in chromatographic separation

PROTEIN TRUNCATION TEST

Transcript or gene



PCR amplification

⇓

RNA transcription

⇓

Protein transduction

⇓

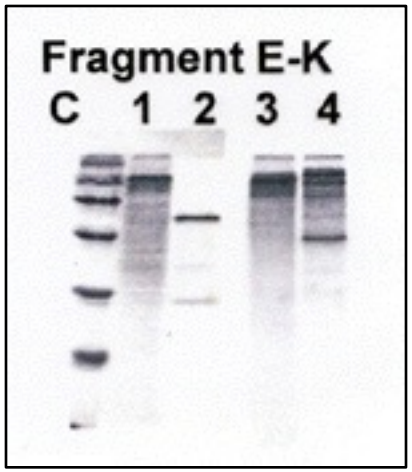
Protein resolution with SDS- PAGE

⇓

Autoradiographic detection

⇓

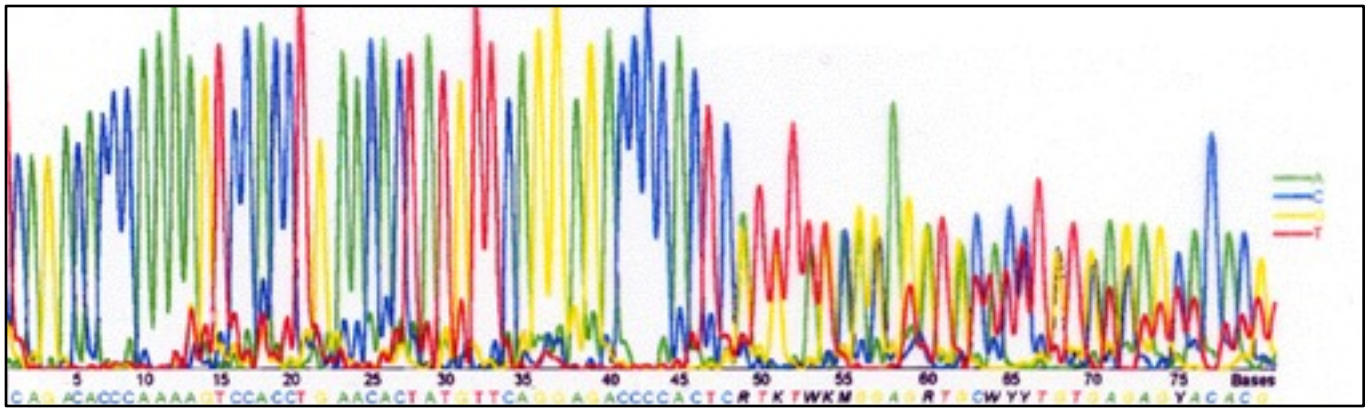
Sequencing



--- --- Normal product

---Truncated product

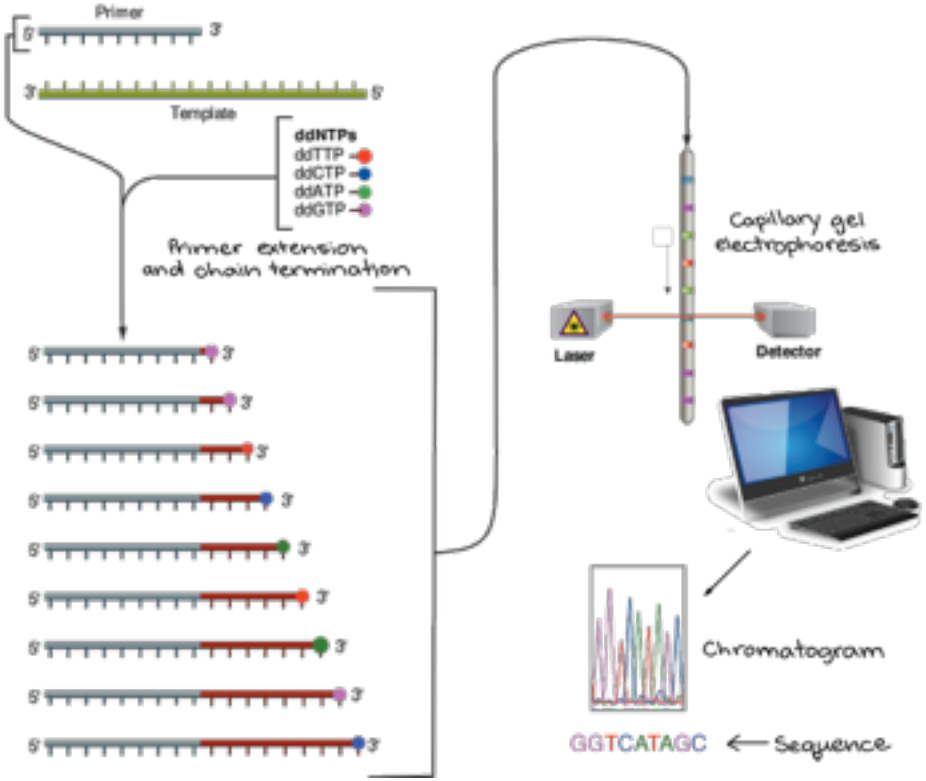
Wild-type Mutated



MUTATION CHARACTERIZATION

Sanger sequencing is a qualitative method of DNA sequencing based on the selective incorporation of chain-terminating dideoxynucleotides by DNA polymerase during in vitro DNA replication.

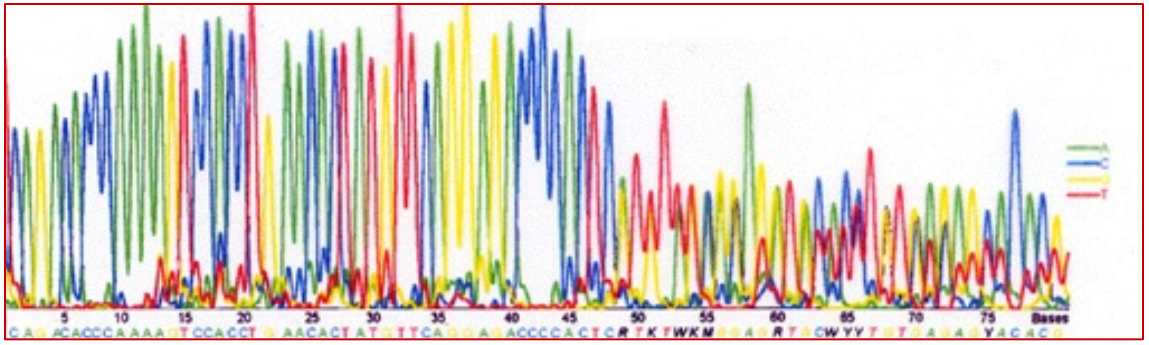
Developed by Frederick Sanger and colleagues in 1977, it was the most widely used sequencing method for approximately 40 years.



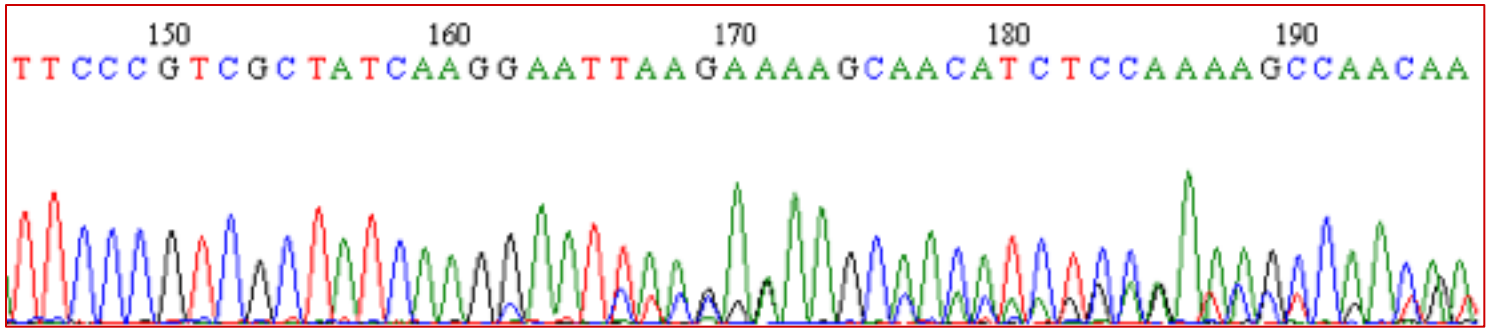
More recently, higher volume Sanger sequencing has been supplanted by "Next-Gen" sequencing methods, especially for large-scale, automated genome analyses.

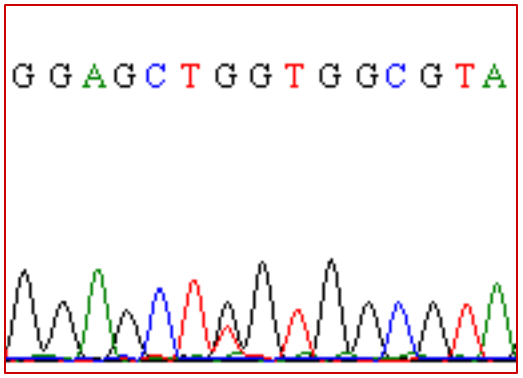
However, the Sanger method remains in wide use, for smaller-scale projects, validation of Next-Gen results and for obtaining especially long contiguous DNA sequence reads (> 500 nucleotides).

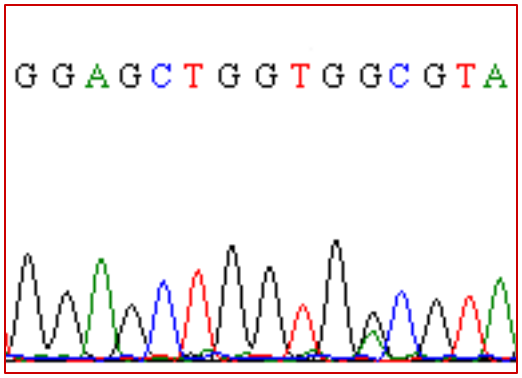
Sanger sequencing with automatic sequencer



Truncating mutation
frameshift



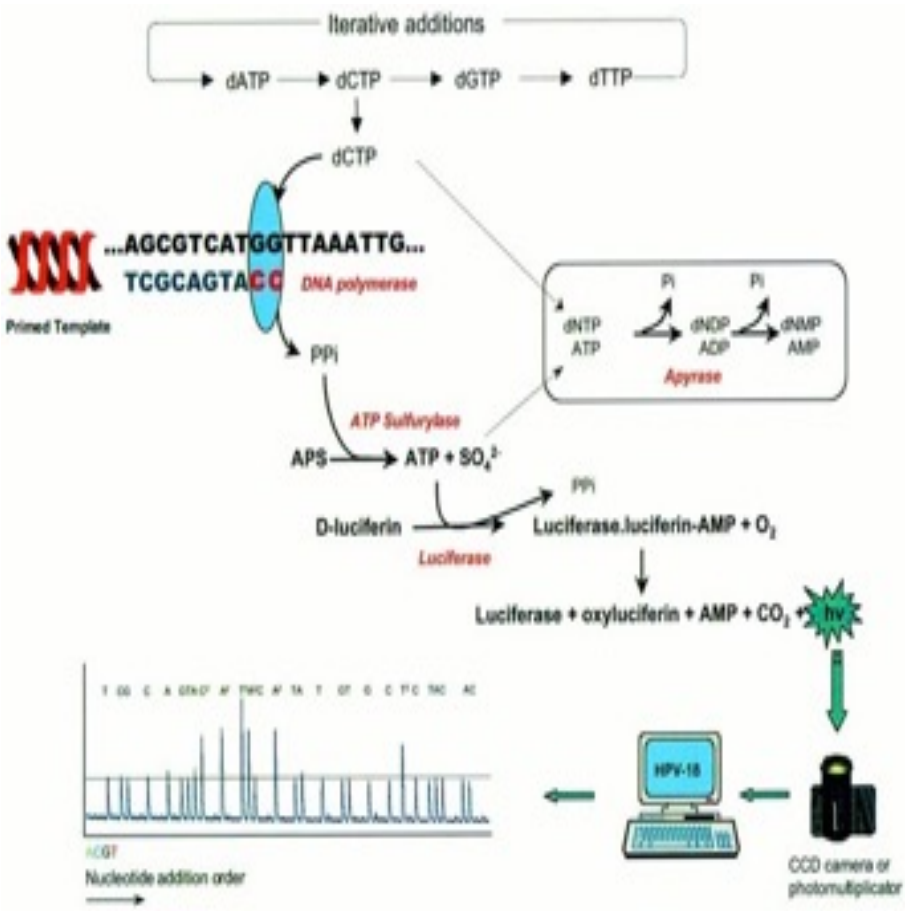




Missense mutation

PYROSEQUENCING

It allows to follow the activity of DNA polymerase during nucleotide incorporation into a DNA strand by analyzing the pyrophosphate released during the process (quantitative method)



- Pyrosequencing
  - 4 nucleotides flow separately
  - If nt incorporation…PPi...light
  - APS + PPi (sulfurylase) → ATP
  - Luciferin + ATP (luciferase) → light + oxyluciferin
  - Amount of light proportional to #nt incorporated
  - Rinse and repeat with next nt

Next Generation Sequencing

Massively parallel or deep sequencing are related terms that describe a DNA sequencing technology which has revolutionised genomic research

Sanger

1 sequence ~ 800 bp (~1 kb)

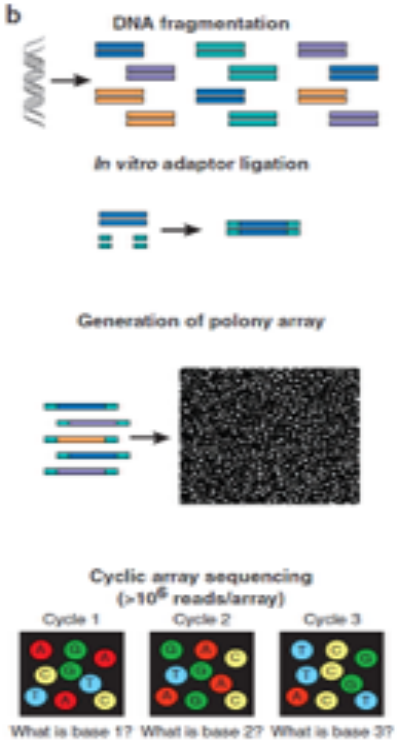
NGS

1 exome ~ 50 Mbp (10 Gb)

1 genome ~ 3 Gbp (0.5 Tb)

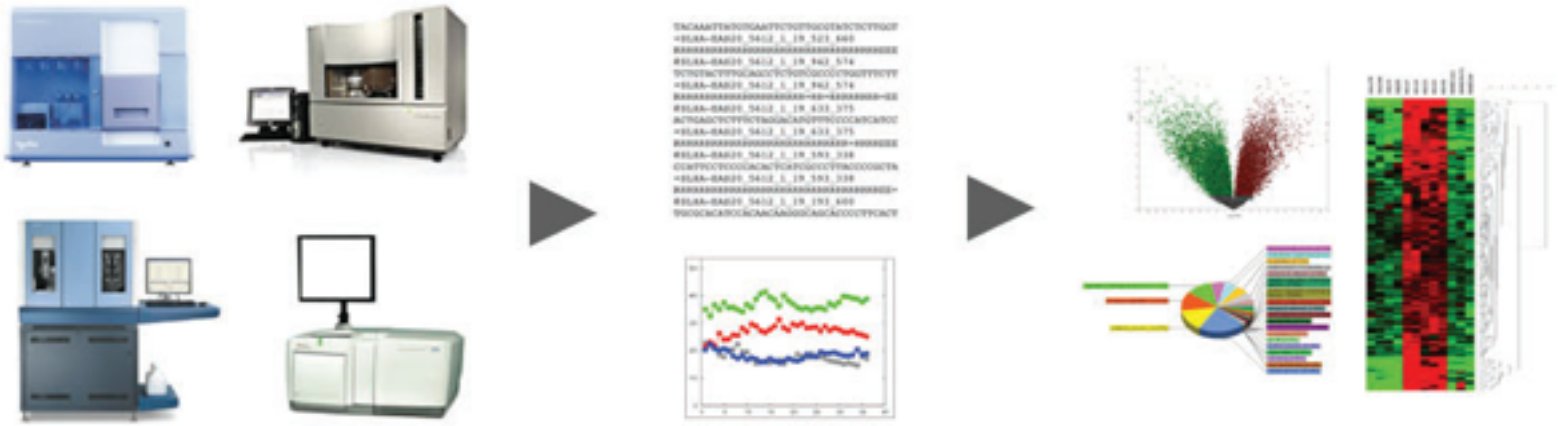
Technology basics

- DNA is fragmented
- Adaptors ligated to fragments
- Several possible protocols yield array of PCR products
  - Emulsion PCR
  - Bridge PCR
- Enzymatic extension with fluorescently tagged nucleotides.
- Cyclic readout by imaging the array.

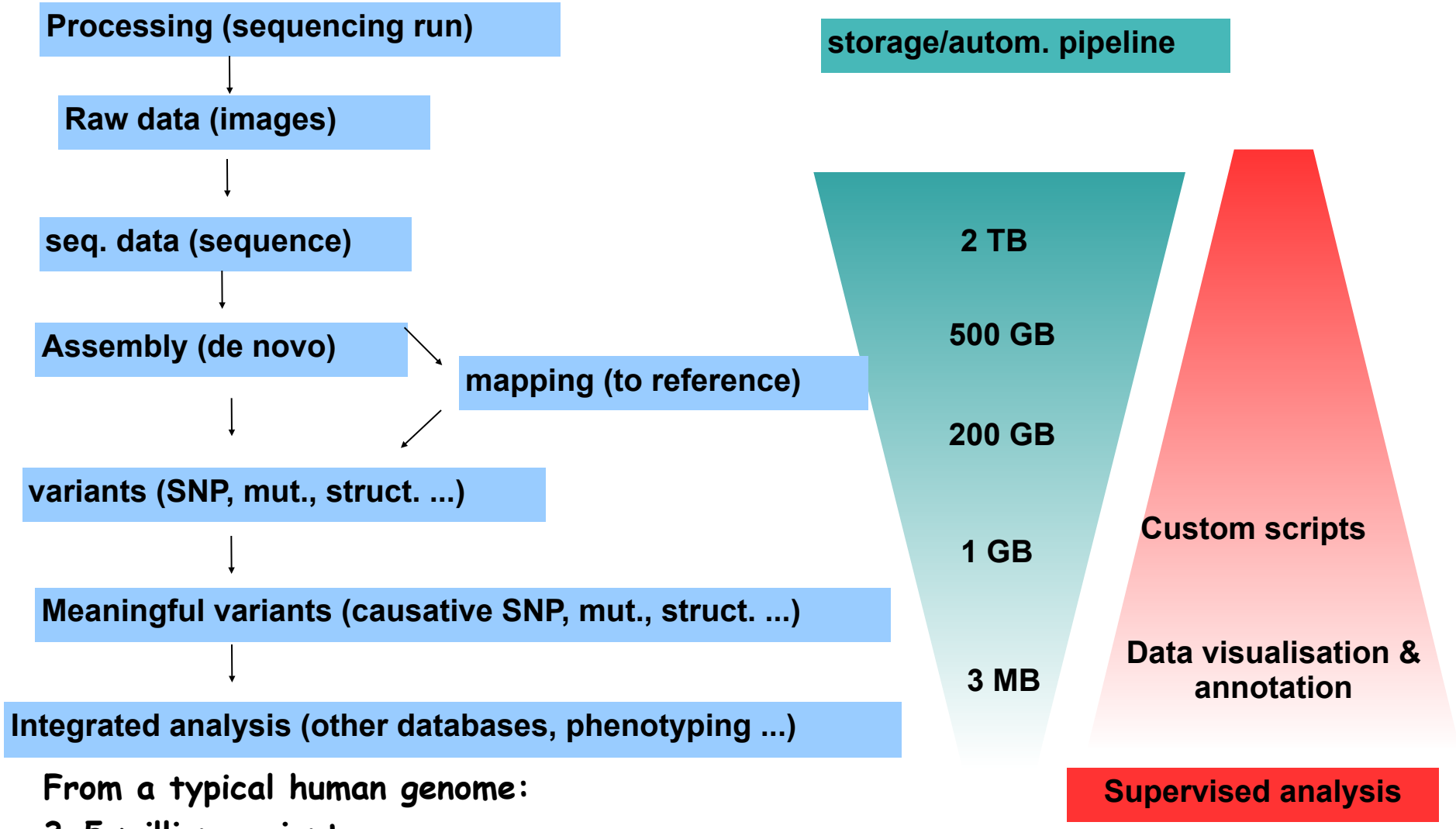


- The raw image data is truly huge: 1-2 TB
- The images are immediately processed into intensity data (spots w/ location and brightness)
- Intensity data is then processed into basecalls

  (A, C, T, or G plus a quality score for each)
- Basecall data is on the order of 5-10 GB/run



- Align sequence reads to reference genome
- Assemble contigs and whole genomes
- CNV/Structural alteration/Mutation/SNP calling/genotyping



From a typical human genome:

3-5 million variants

1000-1300 possibly damaging changes in RefSeq transcripts, including:

- 40-80 stop codons
- splice site changes
- non-synonimous changes

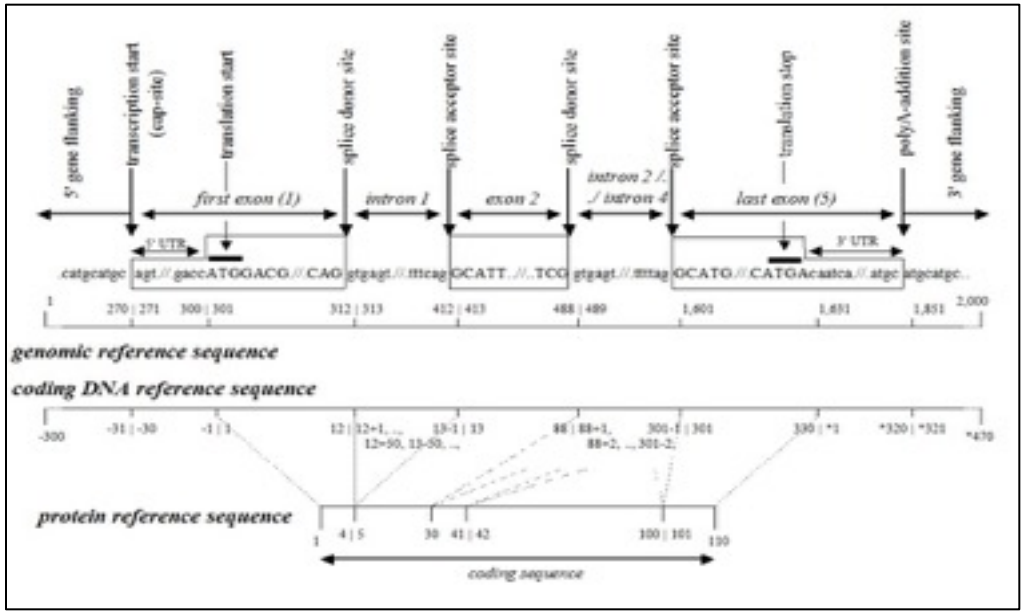
How to decide when a genomic VARIANT is pathogenic

- Functional studies (in vitro)
- Previous evidence: variant found in other patients with the same pathology
- De novo :variant only found in the patient affected by that pathology
- A new change: a variant not identified in 100 healthy controls
- Type of a.a. change: a change modifying the protein structure

Functional studies are the best evidence !!!!

HGVS
HUMAN GENOME VARIATION SOCIETY

http://www.hgvs.org/mutnomen



Genomic DNA

A) **there is no nucleotide 0**; B) **nucleotide 1 is the A of the ATG-translation initiation codon**; C) **the nucleotide 5' of the ATG-translation initiation codon is -1, the previous -2, etc.** D) **the number of the last nucleotide of the preceding exon, a plus sign and the position in the intron, like c.77+1G, c.77+2T**; F) **the number of the first nucleotide of the following exon, a minus sign and the position upstream in the intron, like ..., c.78-2A, c.78-1G**.

Substitutions

A nucleotide substitution is a sequence change where one nucleotide is replaced by one other nucleotide (see Standards - Definition). **Nucleotide substituions are described using a ">".**

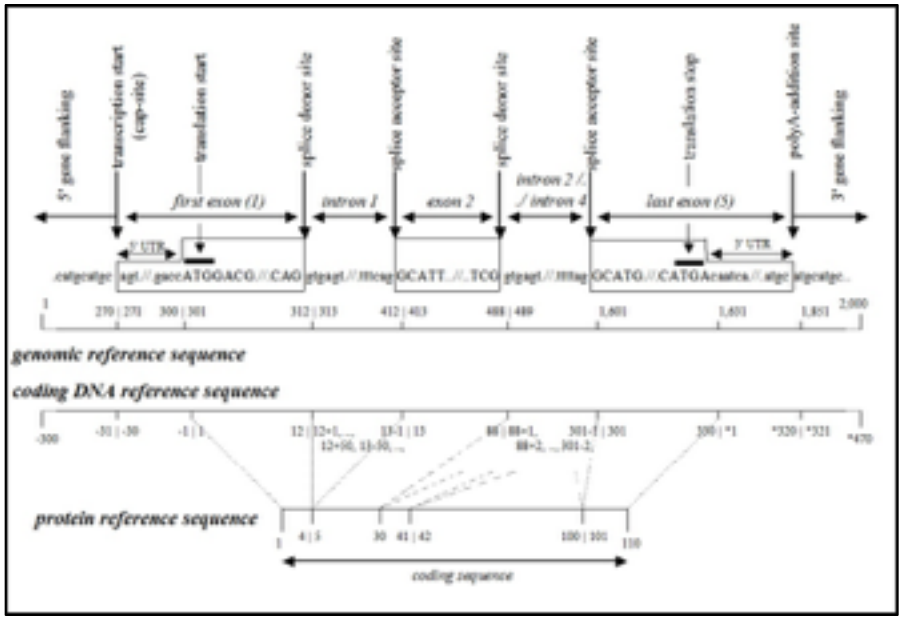
A) **c.76A>C denotes that at nucleotide 76 an A is changed to a C** B) **c.-14G>C denotes a G to C substitution 14 nucleotides 5' of the ATG translation initiation codon** C) **c.88+1G>T denotes the G to T substitution at nucleotide +1 of an intron** D) **c.89-2A>C denotes the A to C substitution at nucleotide -2 of an intron**

Deletions

A nucleotide deletion is a sequence change where one or more nucleotides are removed.
Deletions are described using **"del" after an indication of the first and last nucleotide(s) deleted, separated by a "_" (underscore)**.
A) **c.76_78del (alternatively c.76_78delACT) denotes a ACT deletion from nucleotides 76 to 78**

Duplications

Duplications are designated by **"dup"** after an indication of the first and last nucleotide(s) duplicated.
A) **g.5dupT (or g.5dup, not g.5_6insT) denotes a duplication ("insertion") of the T nucleotide at position 5 in the genomic reference sequence changing ACTCTGTGCC to ACTCTTGTGCC**
B) **c.77_79dup (or c.77_79dupCTG) denotes that the three nucleotides 77 to 79 are duplicated**



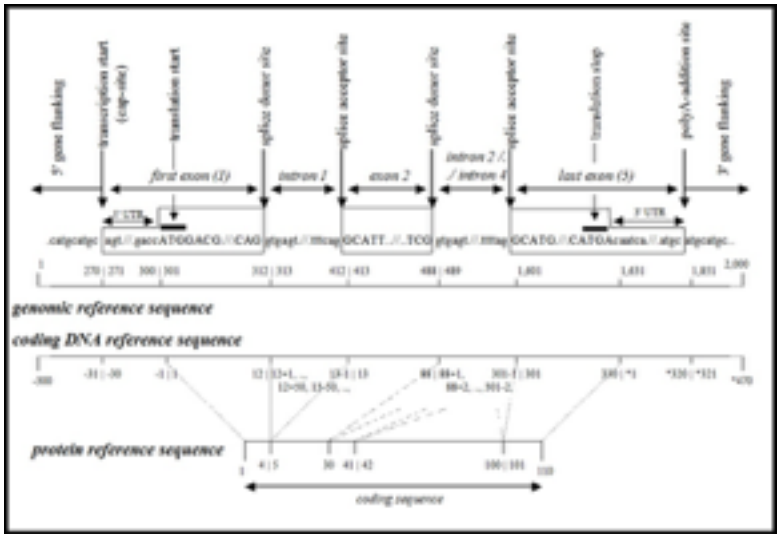
Insertions

Insertions are designated by **"ins"** after an indication of the nucleotides flanking the insertion site, followed by a description of the nucleotides inserted.
For large insertions the number of inserted nucleotides should be mentioned :
A) **c.76_77insT denotes that a T is inserted between nucleotides 76 and 77 of the coding DNA reference sequence**

PROTEINS

Substitutions

Substitutions can be described without using the specific ">"-character which is used on DNA and RNA level (i.e. **p.Trp26Cys instead of p.Trp26>Cys**).



Deletions

Deletions are designated by "del" after a description of the deleted segment, i.e. the first (and last) amino acid(s) deleted.

A) MKLGHQQQCC to MKL___QQCC is described as **p.Gly4_Gln6del (alternatively p.G4_Q6del)**

B) MKLGHQQQCC to MKLGHQQCC is described as **p.Glndel (p.Q8del)**

Duplications

Duplications are designated by "dup" after a description of the duplicated segment, i.e. the first (and last) amino acid(s) duplicated.

A) MKLGHQQQCC to MKLGHQGHQQQCC is described as **p.Gly4_Gln6dup (alternatively p.G4_Q6dup)**

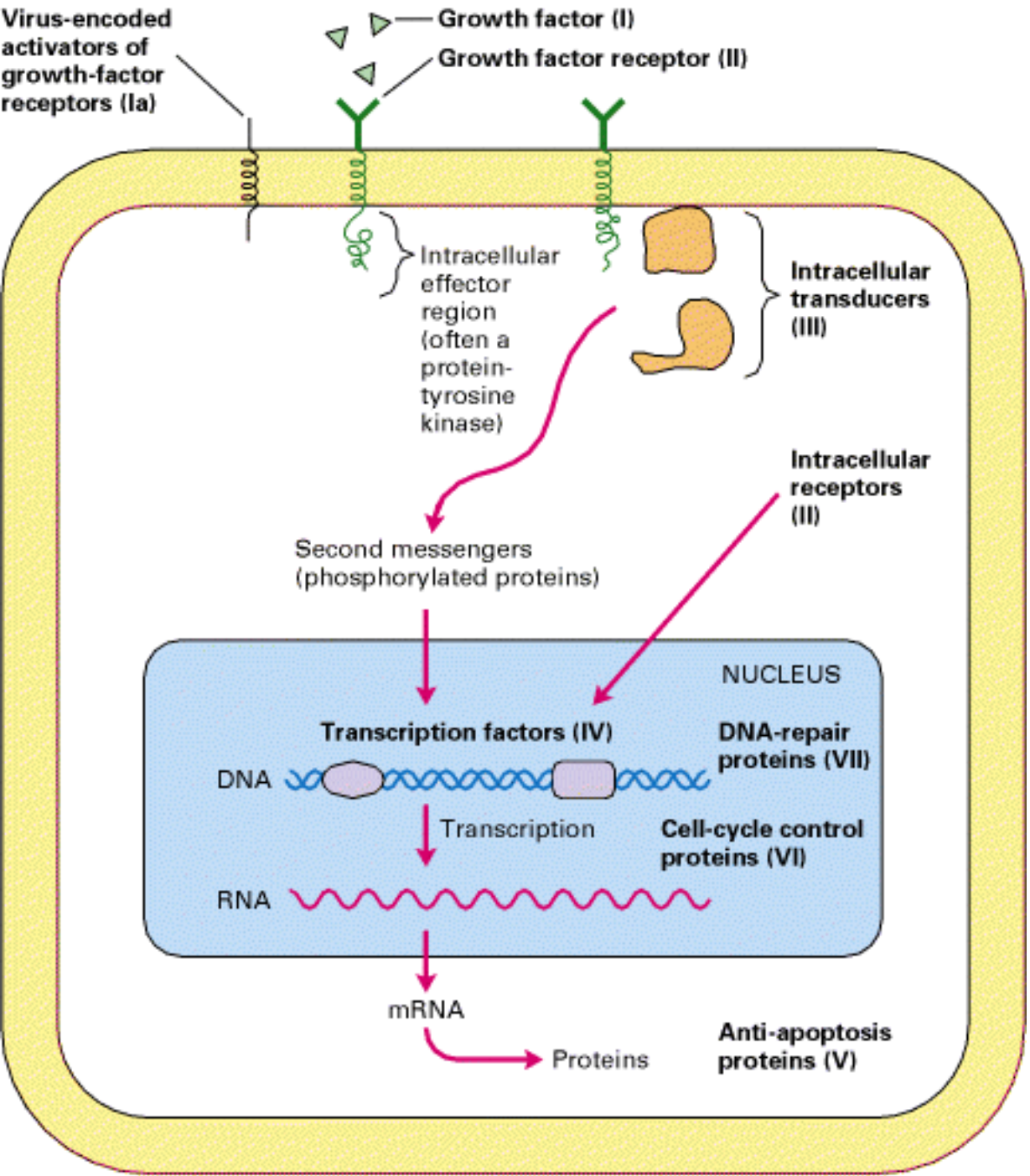
B) MKLGHQQQCC to MKLGHQQQQCC is described as **p.Gln8dup (alternatively p.Q8dup)**

Insertions

Insertions are designated by "ins" after a description of the amino acids flanking the insertion site, followed by a description of the inserted amino acids. When the insertion is large it may be described by its length (e.g. p.Lys2_Leu3ins34).

A) **p.Lys2_Leu3insGlnSer (alternatively p.K2_L3insQS) describes the change from MKLGHQQQCC to MKQSLGHQQQCC**

B) **p.Arg78_Gly79ins23 describes the in-frame insertion of a 23 amino acid sequence**



Oncogenes code for

1. Growth factors
2. Growth factor receptors
3. Transduction proteins
4. Transcription factors
5. Anti-apoptotic proteins
6. Cell-cycle proteins
7. DNA repair proteins

MECHANISMS OF ONCOGENIC ACTIVATION

1. Missense mutations

2. Chromosomal rearrangements- translocations

3. Amplification

4. Insertional mutations ("jumping" genes such as LINE1 or ALU sequence etc)

REPRESENTATIVE ONCOGENES OF HUMAN TUMORS

Oncogene	Type of cancer	Activation mechanism
ABL	Chronic myelogenous leukemia, acute lymphocytic leukemia	Translocation
AKT	Ovarian and pancreatic carcinomas	Amplification
BCL-2	Follicular B-cell lymphoma	Translocation
E2A/pbx1	Acute lymphocytic leukemia	Translocation
HER2	Breast and ovarian carcinomas	Amplification
GIP	Adrenal cortical and ovarian carcinomas	Point mutation
GLI	Glioblastoma	Amplification
GSP	Pituitary and thyroid tumors	Point mutation
HOX-11	Acute T-cell leukemia	Translocation
LYL	Acute T-cell leukemia	Translocation

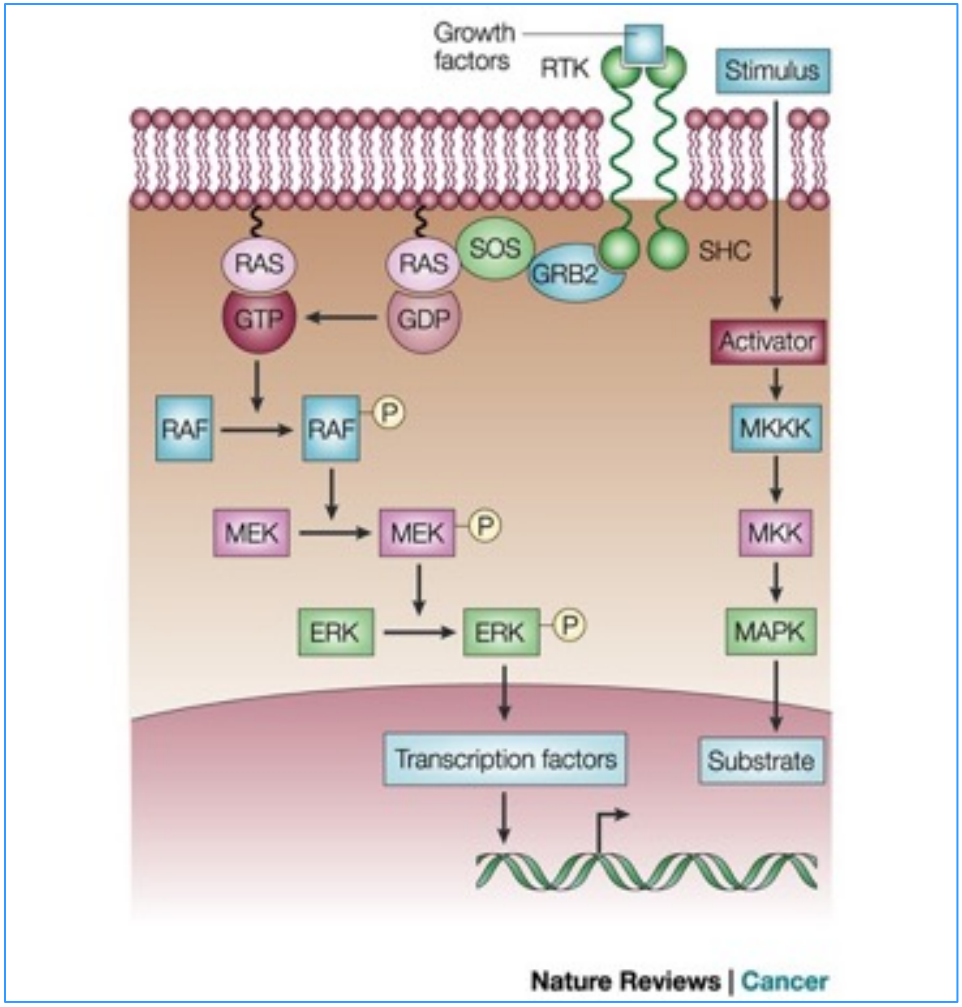
Oncogene	Type of cancer	Activation mechanism
C-MYC	Burkitt's lymphoma	Translocation
	Breast and lung carcinomas	Amplification
L-MYC	Lung carcinoma	Amplification
N-MYC	Neuroblastoma, lung carcinoma	Amplification
PDGFR	Chronic myelomonocytic leukemia	Translocation
PML/RARa	Acute promyelocytic leukemia	Translocation
RAS-H	Thyroid carcinoma	Point mutation
RAS-K	Colon, lung, pancreatic, and thyroid carcinomas	Point mutation
RAS-N	Acute myelogenous and lymphocytic leukemias, thyroid carcinoma	Point mutation
RET	Multiple endocrine neoplasia types 2A and 2B	Point mutation
	Thyroid carcinoma	DNA rearrangement
SMO	Basal cell carcinoma	Point mutation

MISSENSE MUTATIONS

Amino acid substitutions in RAS family proteins

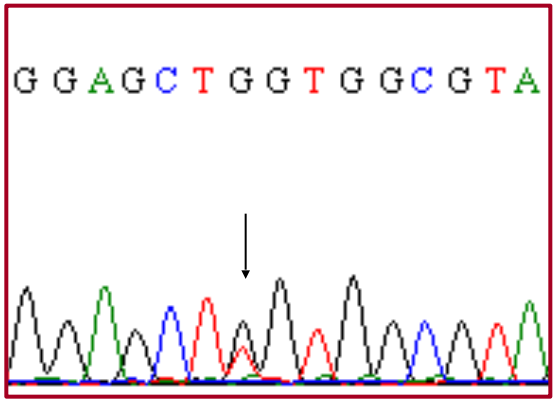
RAS gene	amino acid position 12	59	61	
c-RAS (H, K, N)	Gly	Ala	Gln	normal cells
				cancer cells
H-RAS	Gly	Ala	Leu	lung carcinoma
	Val	Ala	Gln	bladder carcinoma
K-RAS	Cys	Ala	Gln	lung carcinoma
	Val	Ala	Gln	lung carcinoma
	Arg	Ala	Gln	colon carcinoma
N-RAS	Gly	Ala	Lys	neuroblastoma
	Gly	Ala	Arg	lung carcinoma
				Murine sarcoma virus
H-RAS	Arg	Thr	Gln	Harvey strain
K-RAS	Ser	Thr	Gln	Kirsten strain

MAPK pathway



KRAS mutations

Codone 12 G >T



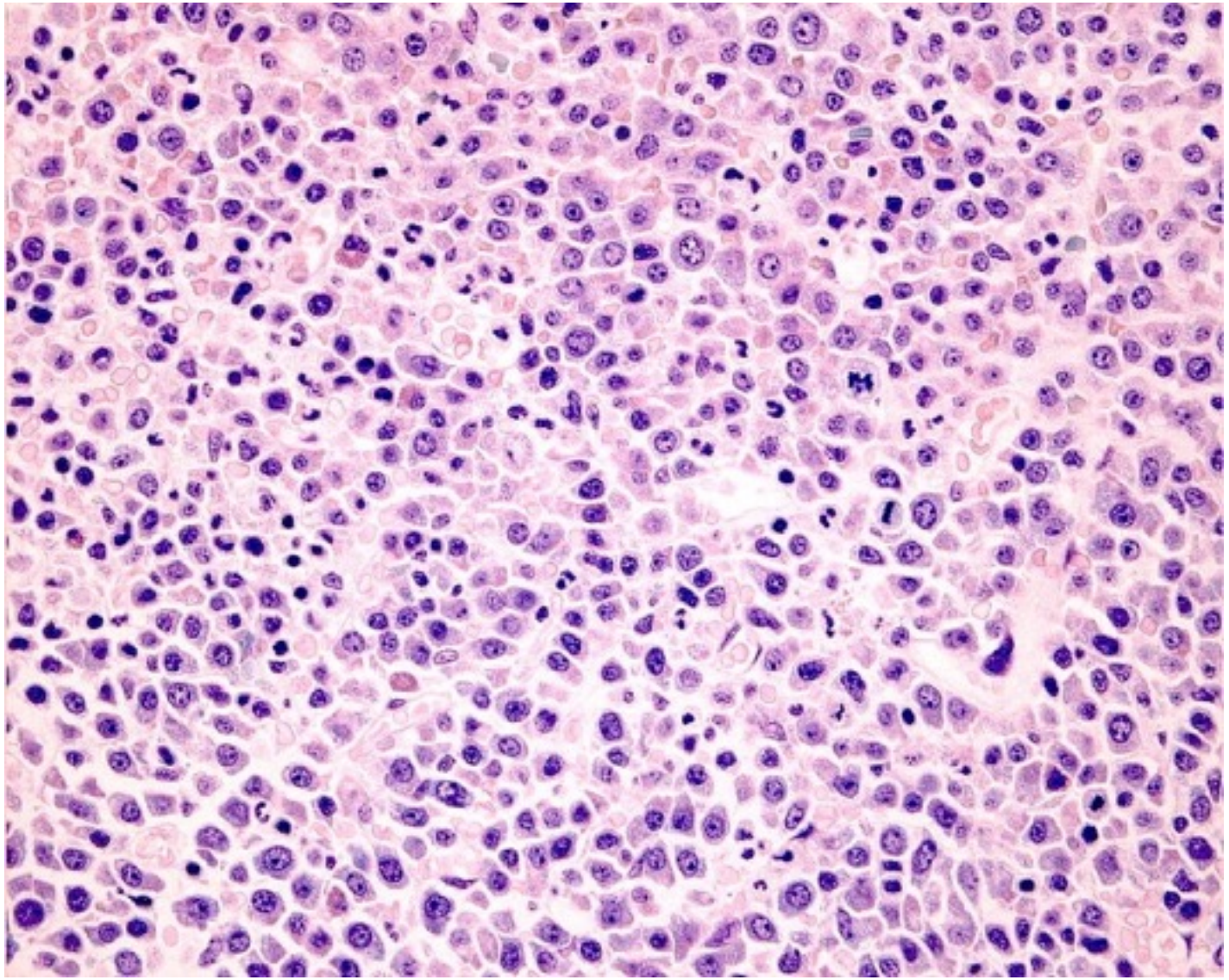
2. CHROMOSOMAL TRANSLOCATIONS

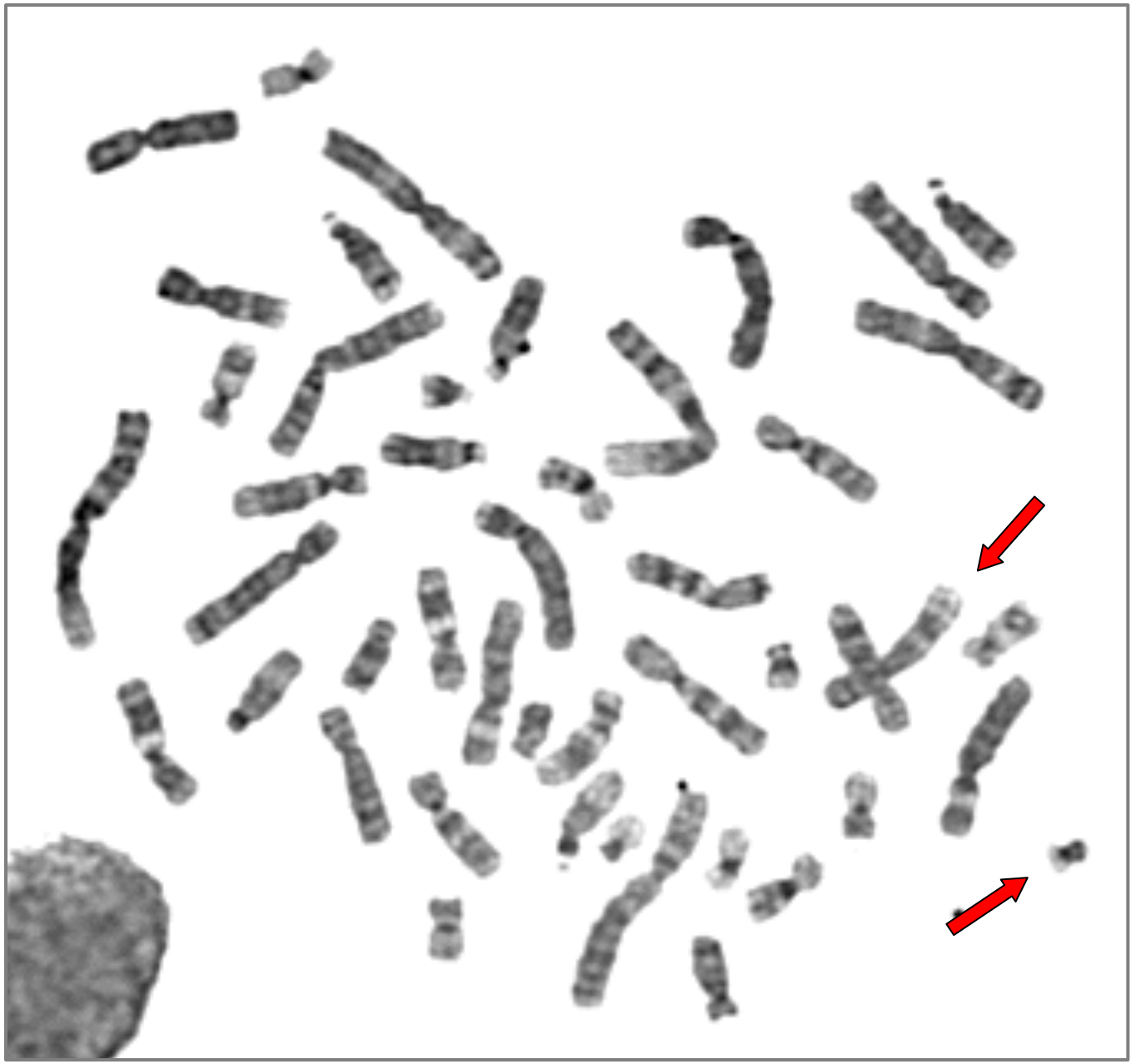
Neoplasm	Translocation	Oncogene
Burkitt lymphoma	t(8;14) 80% of cases t(8;22) 15% of cases t(2;8) 5% of cases	c-MYC[1]
Chronic myelogenous leukemia	t(9;22) 90-95% of cases	BCR-ABL[2]
Acute lymphocytic leukemia	t(9;22) 10-15% of cases	BCR-ABL[2]

[1] c-MYC is translocated to the IgG locus, which results in its activated expression

[2] BCR-ABL fusion protein is produced, which results in a constitutively active abl kinase

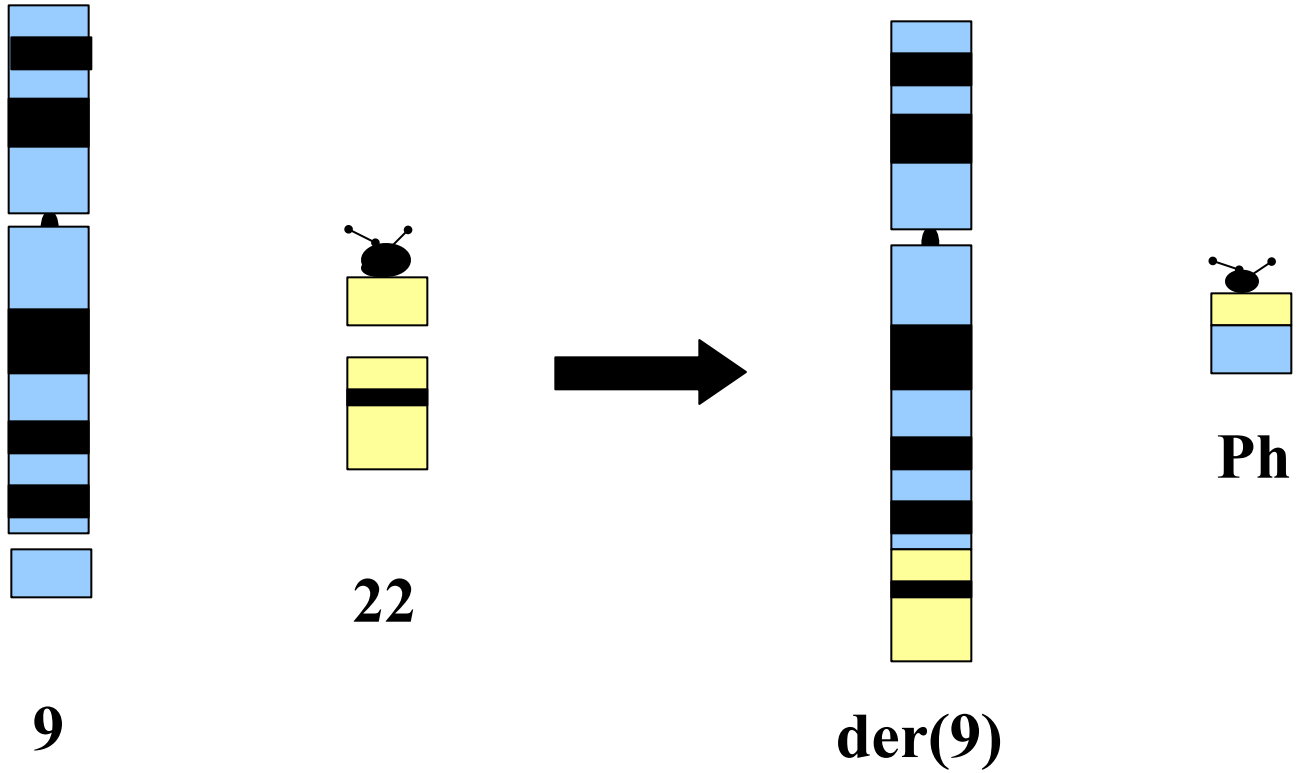
CHRONIC MYELOID LEUKEMIA





t(9;22)(q34;q11)

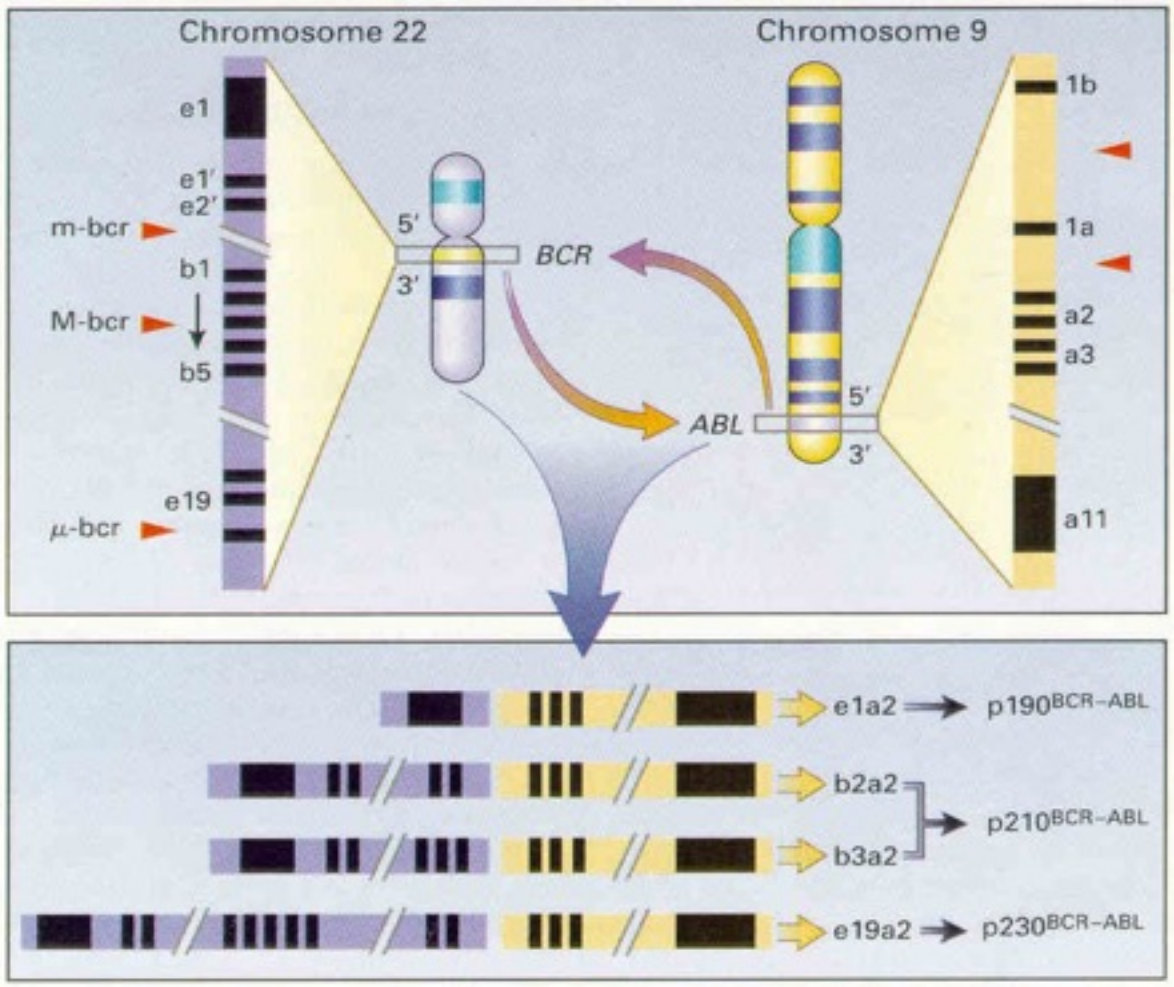
9 22



Chronic Myeloid Leukemia

t(9;22)(q34;q11)

t(9;22)(q34;q11)

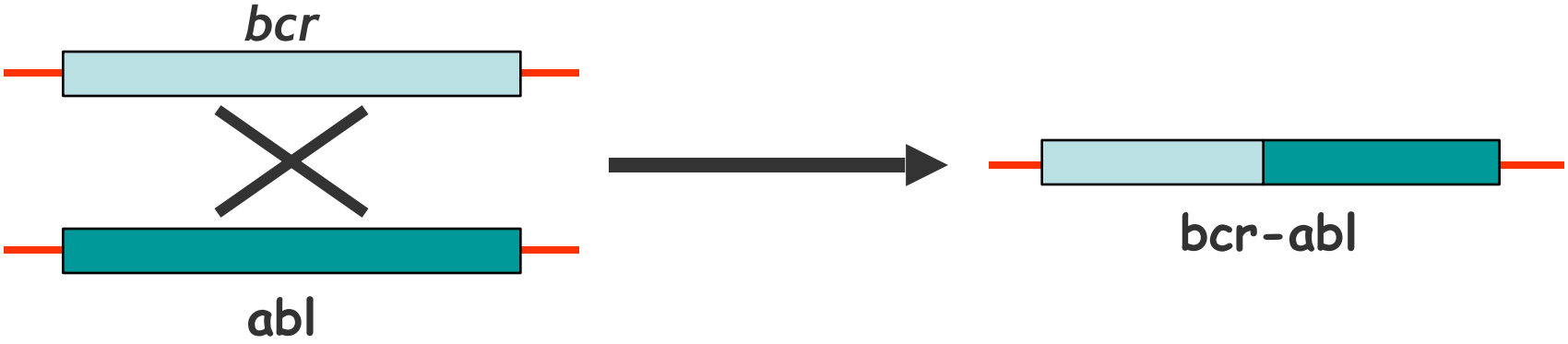


The **ABL1** proto-oncogene encodes a cytoplasmic and nuclear protein tyrosine kinase that has been implicated in processes of cell differentiation, cell division, cell adhesion, and stress response. When it is translocated from chromosome 9 to chromosome 22 it become activated

Chronic myeloid leukemia (CML)

Philadelphia chromosome (small 22)

bcr-abl fusion protein is produced,
which results in a constitutively active abl kinase

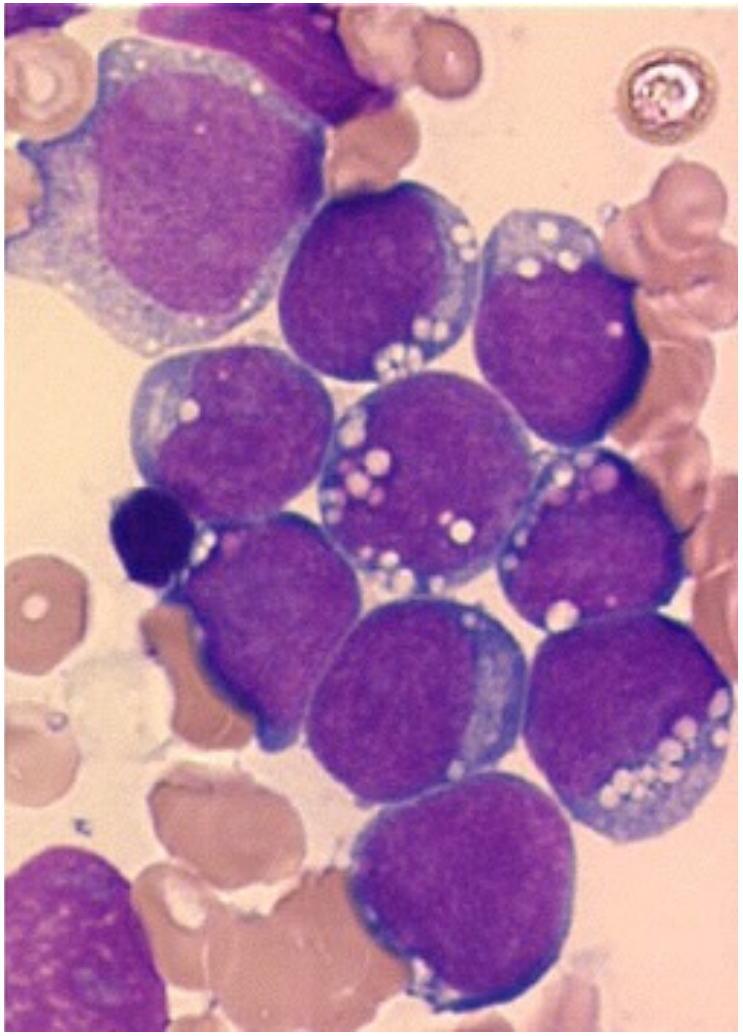


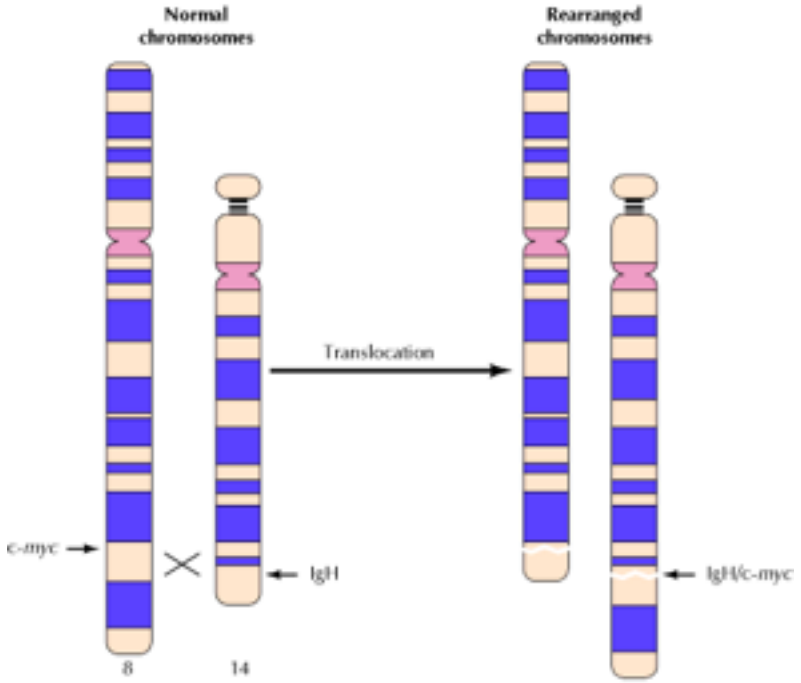
Mechanisms responsible for t(9;22)

- **Random process**
- **Hot spots particularly susceptible to DNA damage (fragile site :** *specific loci that exhibit gaps and breaks on metaphase chromosomes with partial inhibition of DNA synthesis*)
- **Topographic arrangement of chromosomes 9 and 22 in the nucleus**
- **Duplicated genomic sequences close to the BCR and ABL genes**

BURKITT'S LYMPHOMA

- Aggressive lymphoma
- Endemic form in malarian regions (Equatorial Africa). Children 4-7 ya are frequently affected.
- Sporadic form (1-2% of all lymphomas)
- Frequently associated with immunodeficiency (eg. HIV)
- Associated with EBV infection



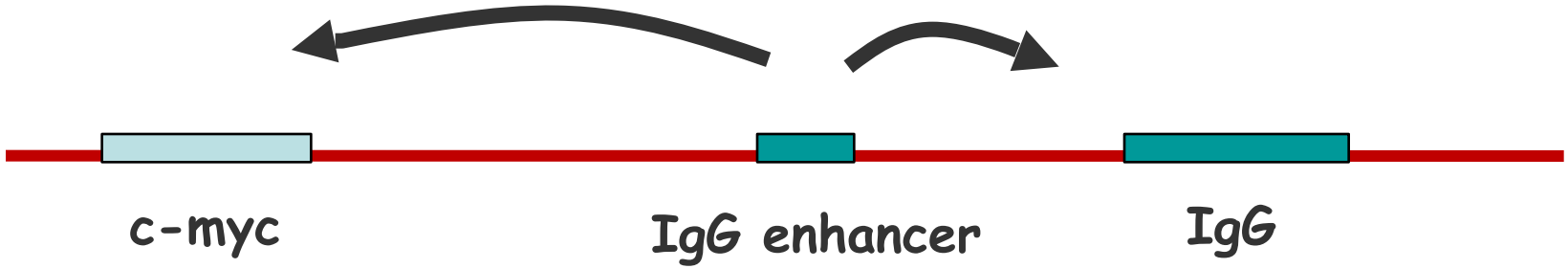


BURKITT LYMPHOMA (B cells)

TRANSLOCATION OF C-MYC

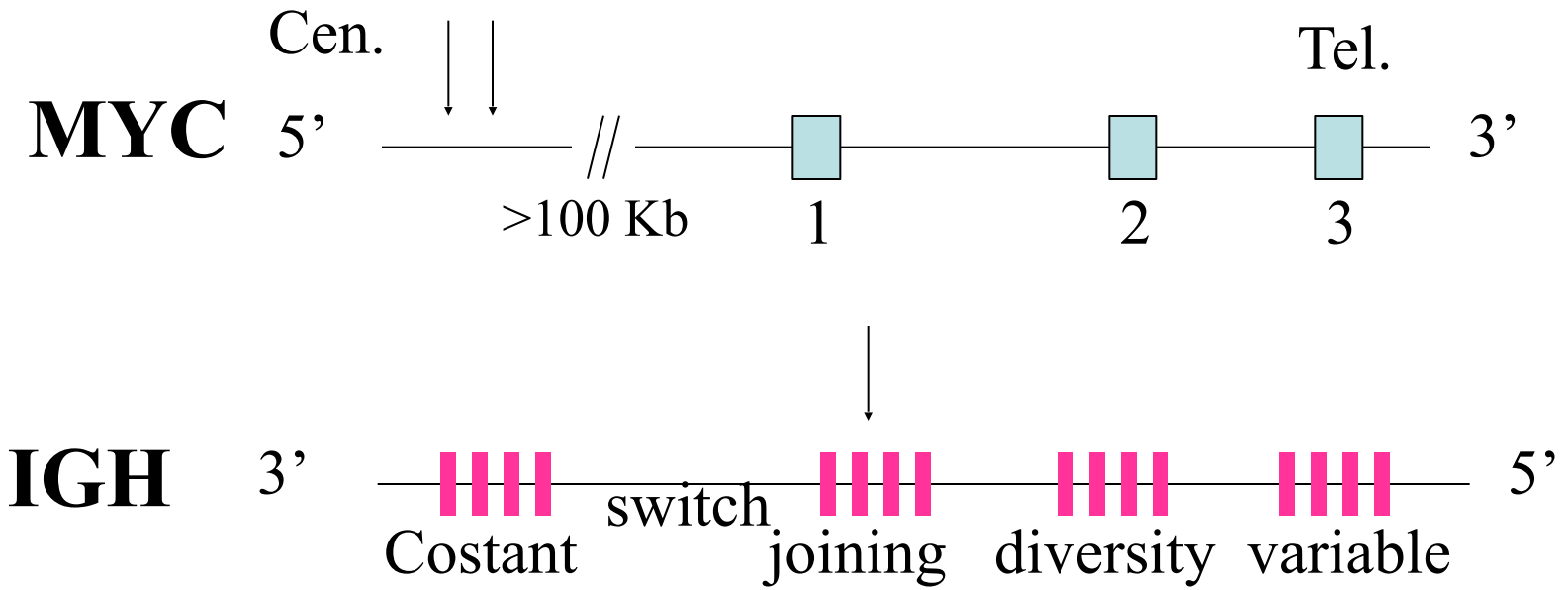
The c-MYC proto-oncogene is translocated from chromosome 8 to the immunoglobulin heavy-chain locus (IgH) on chromosome 14, resulting in abnormal c-myc expression.

c-MYC is translocated to the IgG locus, which results in its activated expression



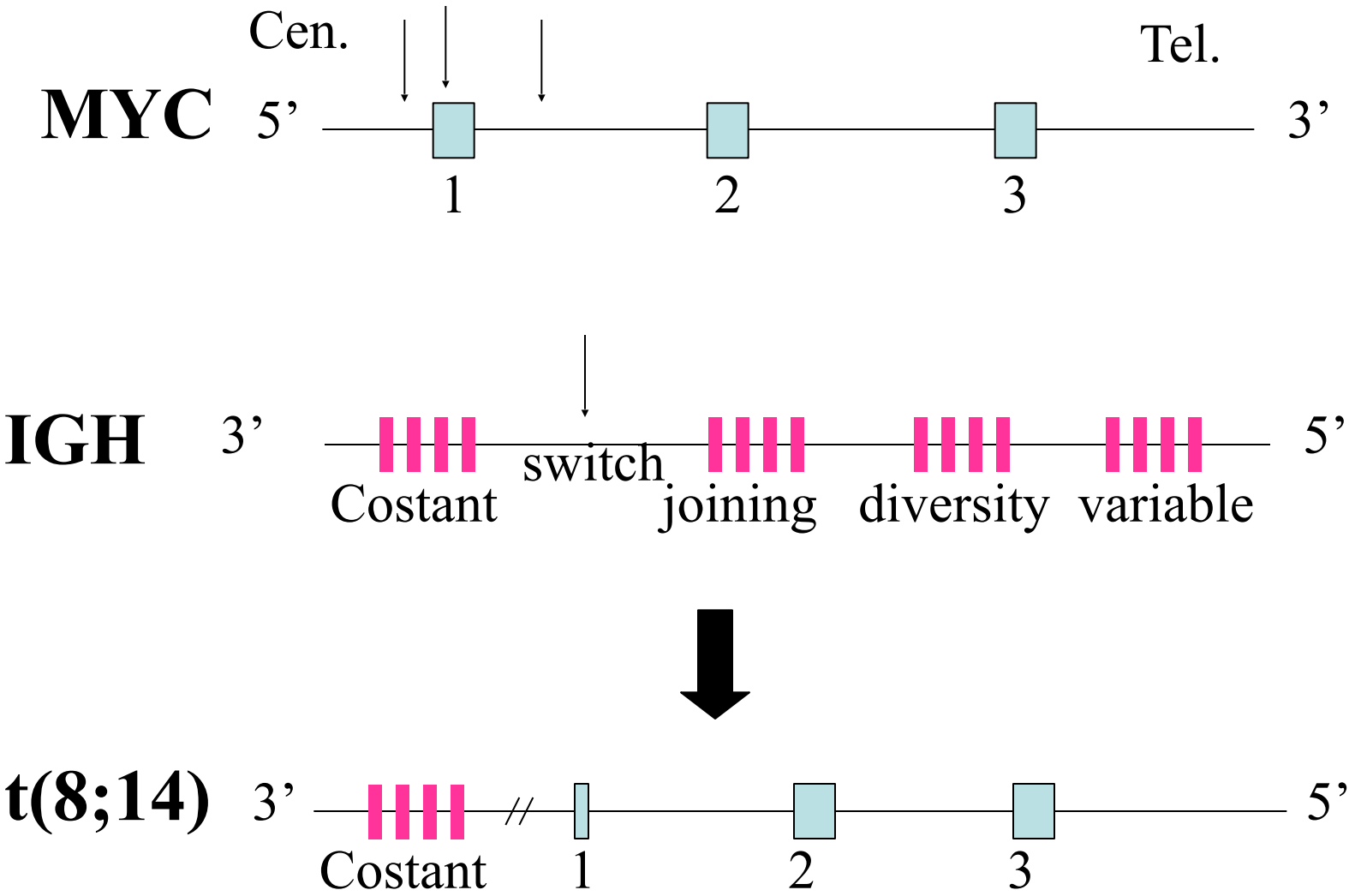
In B lymphocytes c-myc is activated by the IgG enhancer

GM Cooper 2000



Increased expression of c-myc in B-lymphocytes

Proliferation is induced, even in the absence of growth factors.


Cen.
Tel.
MYC
5’
3’
1
2
3
IGH
3’
5’
switch
Costant
joining
diversity
variable
t(8;14)
3’
5’
Costant
1
2
3

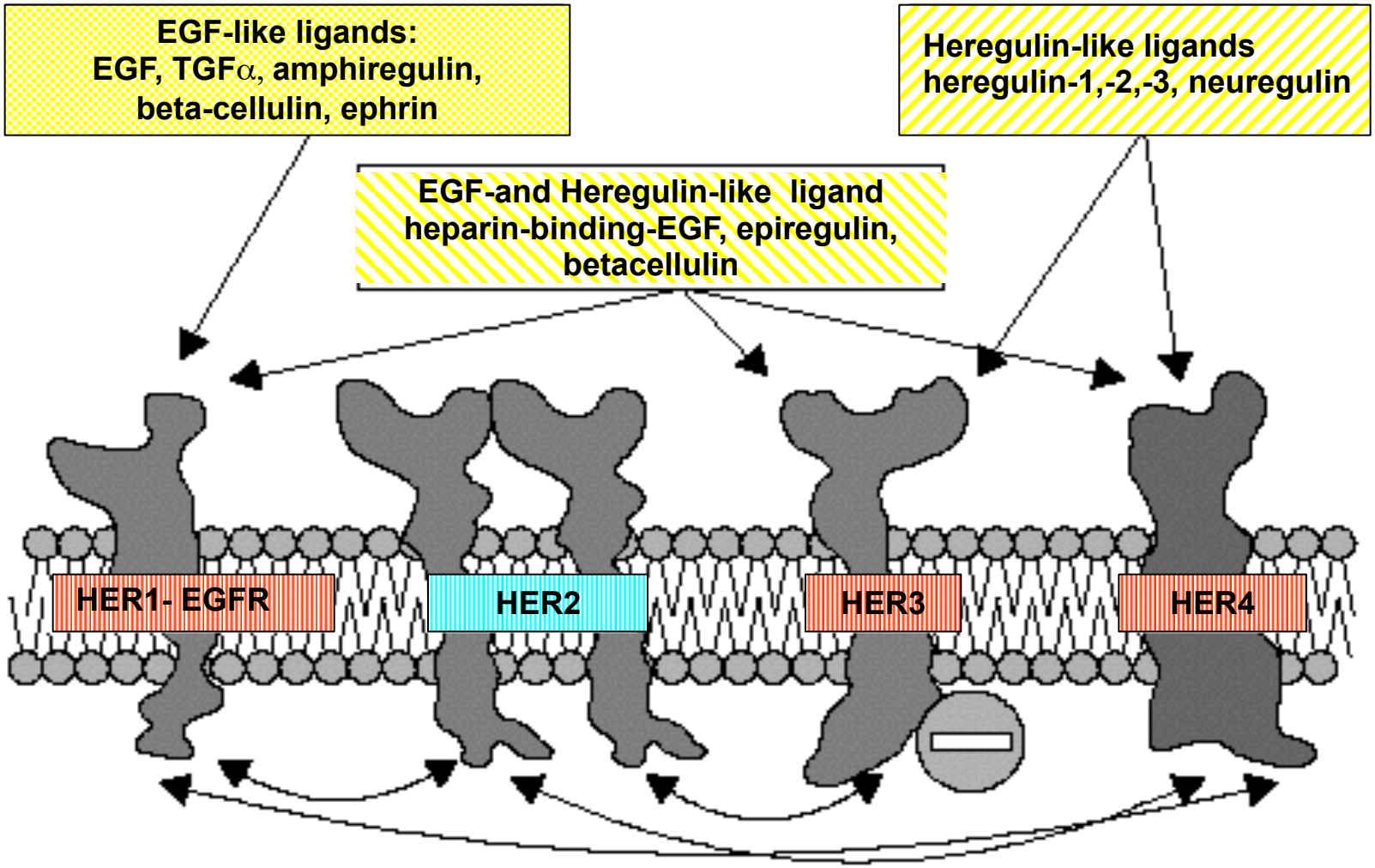
3. GENE AMPLIFICATION

Very frequent event in oncogenes (1000x)

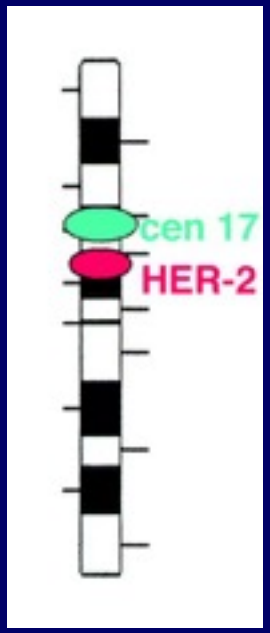
hyperexpression = proliferative selective advantage

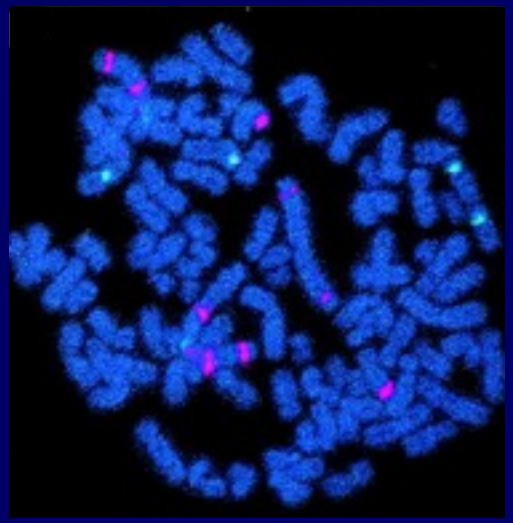
Oncogene	Amplification	Source of tumor
MYC	~20-fold	leukemia /lung carcinoma
MYC	5-1,000-fold	neuroblastoma
MYC	10-20-fold	small-cell lung cancer
ABL	~5-fold	chronic myeloid leukemia
HER2	~30-fold	breast cancer
K-RAS	4-20-fold	colon carcinoma
	30-60-fold	adrenocortical carcinoma

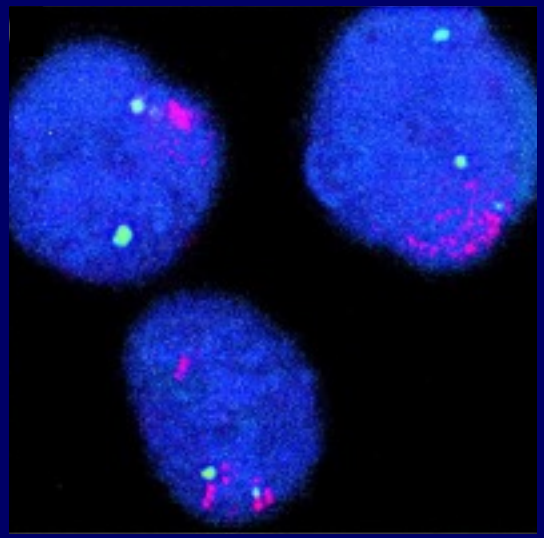
Epidermal Growth Factor Receptor (EGFR)

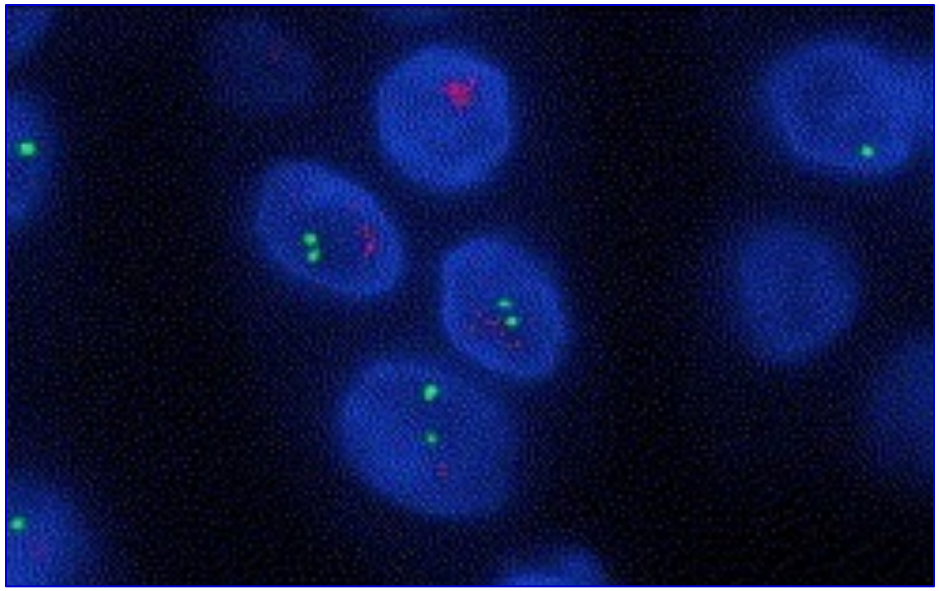


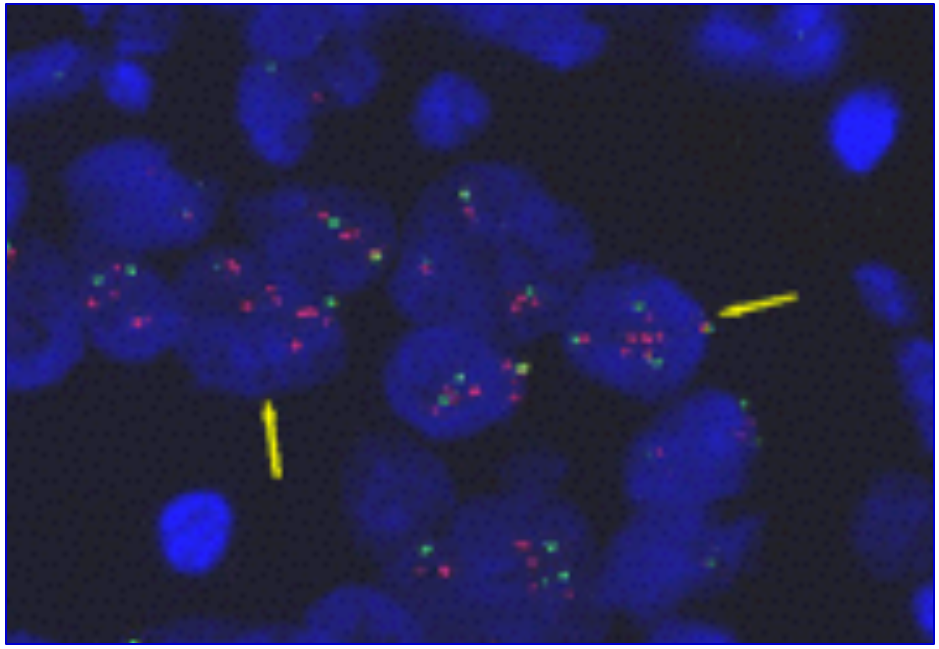
HER-2 gene in Breast Cancer
Dual Color FISH



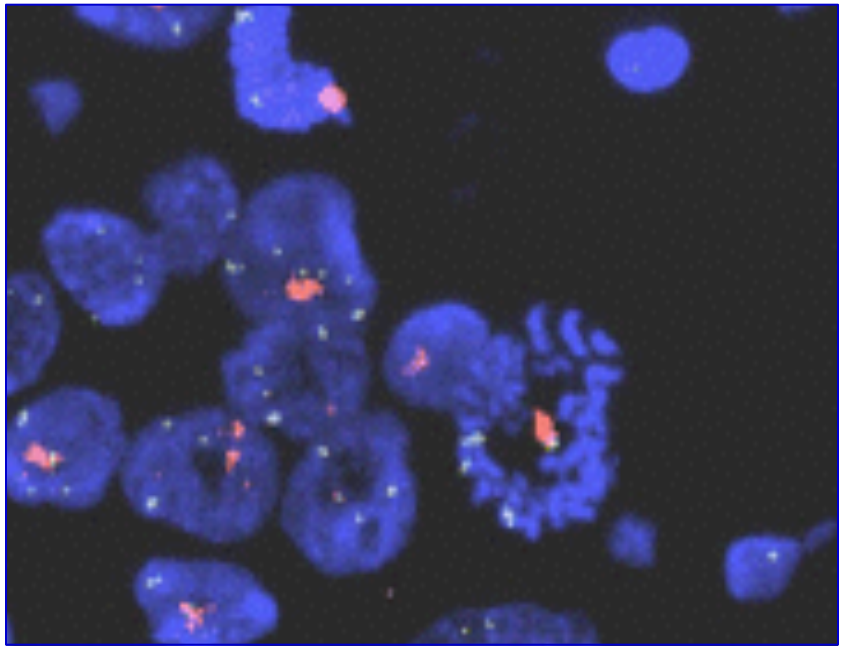






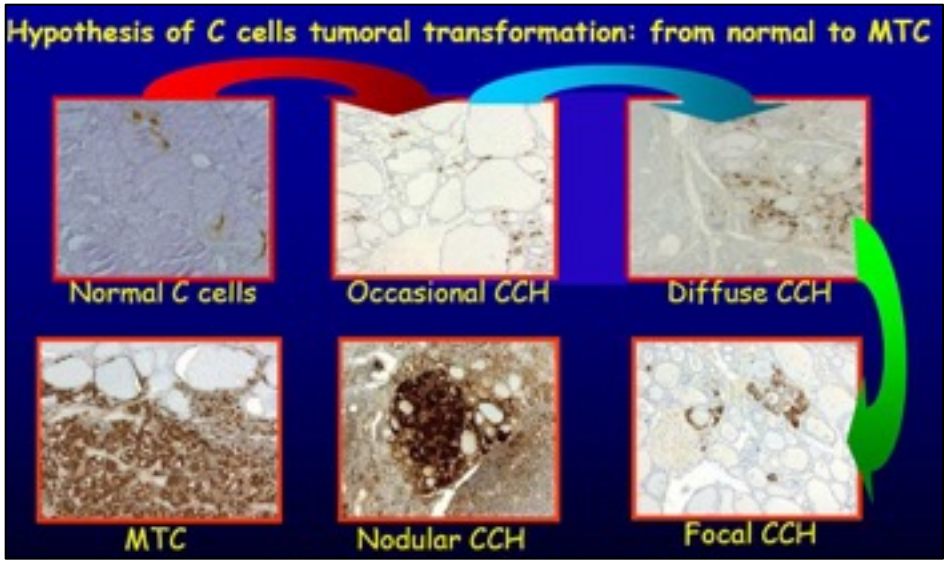


HER-2 amplification



Oncogenes and hereditary tumors

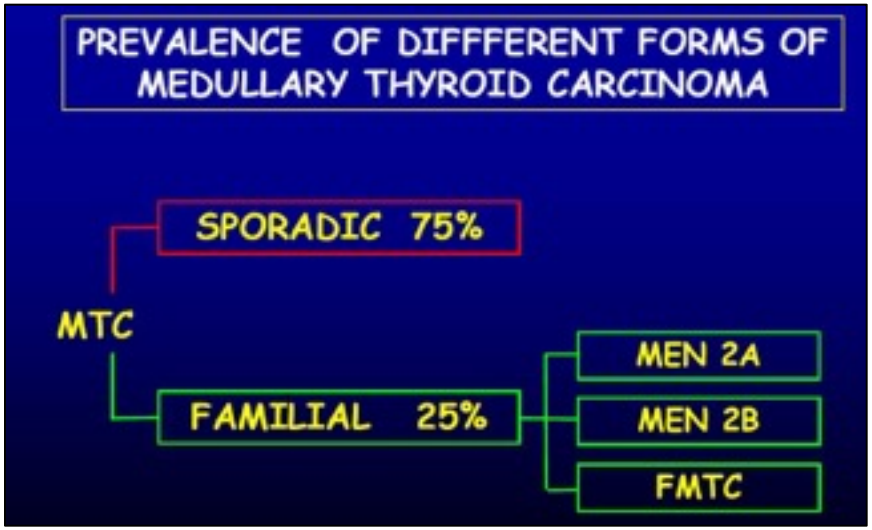
Medullary thyroid carcinoma can be **sporadic** or **familial**



Familial cases are found in

Multiple Endocrine Neoplasia II (MEN II)

Autosomic dominant pattern



Oncogenes and hereditary tumors

Multiple Endocrine Neoplasia: MEN II

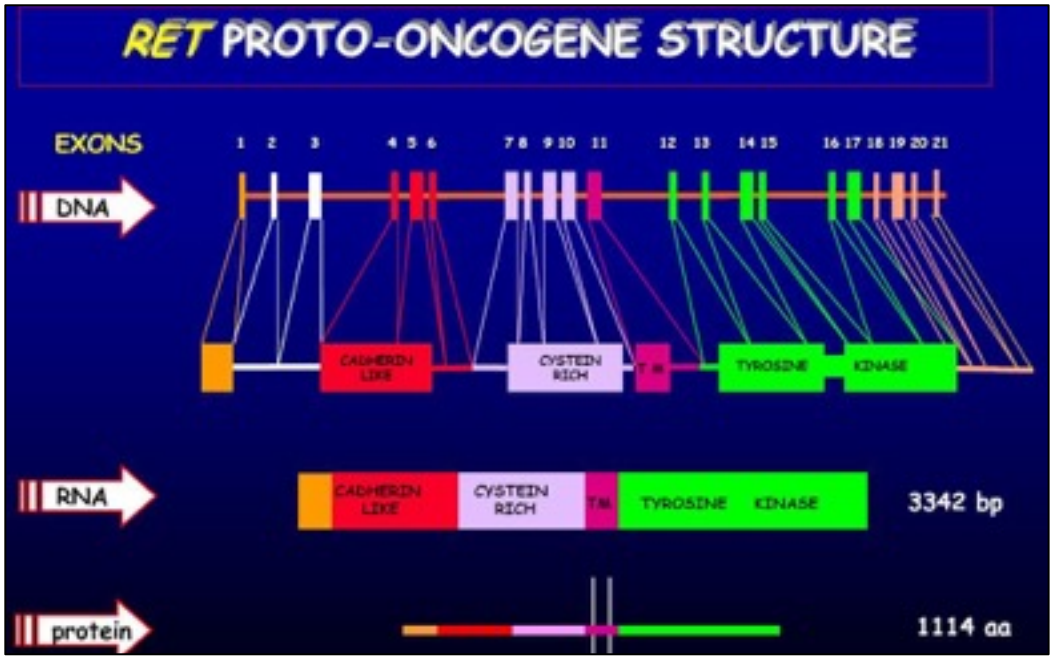
	PATHOLOGY	PREVALENCE
II A	Medullary thyroid carcinoma	100%
	Phaeochromocytoma	50%
	Parathyroid adenoma	30%
	Coutaneous lichen amyloidosis	<10%
II B	Medullary thyroid carcinoma	100%
	Mucosal neurinomas	100%
	Marfanoid habitus	65%
	Phaeochromocytoma	45%
	Corneal nerve hypertrophy	50%
FMTC	Medullary thyroid carcinoma	100%

MEN II *A* and B

FMTC

Autosomic dominant

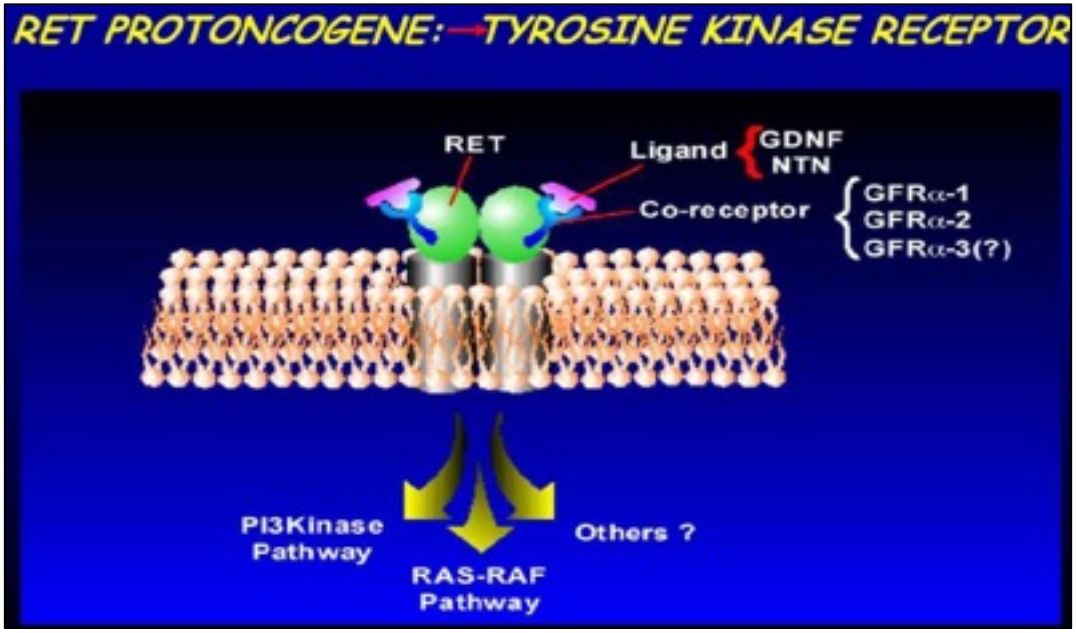
RET oncogene



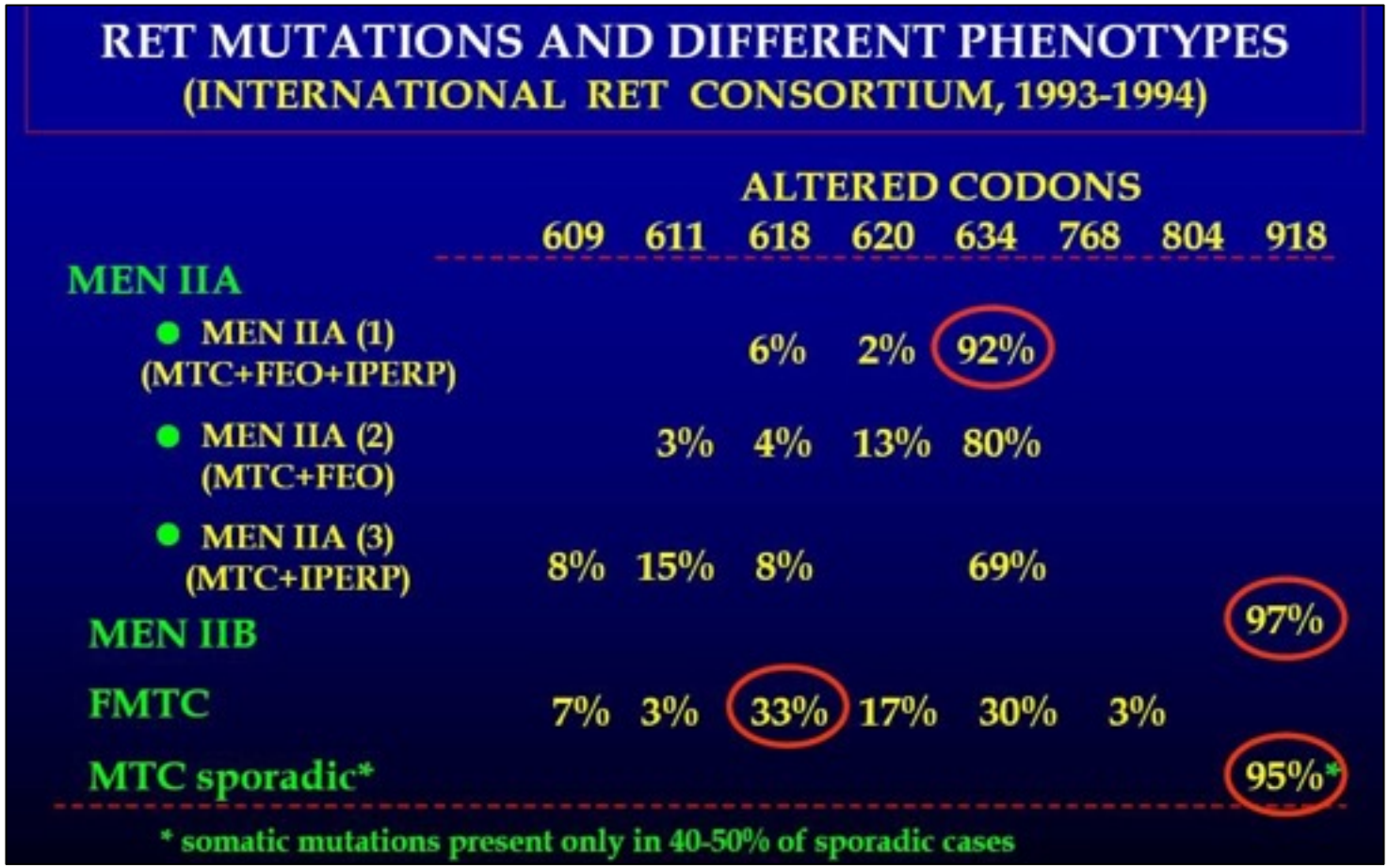
The RET oncogene encodes a **receptor tyrosine kinase** for members of the glial cell line-derived neurotrophic factor family

a) loss of function mutations are associated with the development of Hirschsprung's disease (congenital megacolon : occurs when part or all of the large intestine or parts of the gastrointestinal tract have no ganglion cells and cannot function)

b) gain of function mutations are associated with the development of various types of human cancer, including medullary thyroid carcinoma, multiple endocrine neoplasias type 2A and 2B, pheochromocytoma and parathyroid hyperplasia



RET germline gain-of-function mutations and phenotypes



Noonan Syndrome

It is a genetic multisystem disorder characterised by
1) distinctive facial features, 2) developmental delay, 3) learning difficulties, 4) short stature, 5) congenital heart disease, 6) renal anomalies, 7) lymphatic malformations, and 8) bleeding difficulties

Autosomal dominant

Prevalence of 1 in 1000-2500

Mutations that cause Noonan syndrome alter genes encoding proteins with roles in the **RAS–MAPK pathway**, leading to pathway dysregulation

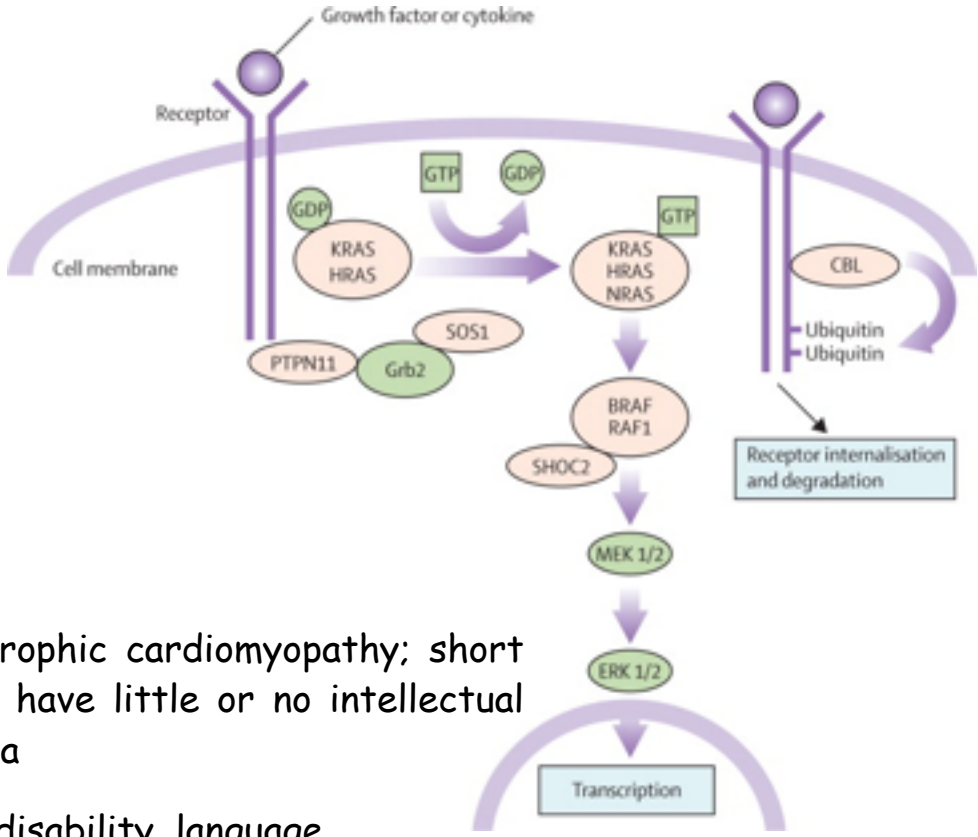
Mutations in Noonan Syndrome

KRAS/NRAS 2% of cases-mild gain of function

PTPN11 50% of cases- gain of function

SOS-1 10% of cases

RAF1-BRAF 5-10% of cases



Genotype-phenotype correlations

PTPN11 pulmonary stenosis; less hypertrophic cardiomyopathy; short stature; patients with N308D and N308S have little or no intellectual disability; juvenile myelomonocytic leukaemia

SOS1 less short stature, less intellectual disability, language delays similar to cardiofaciocutaneous syndrome

RAF1 more hypertrophic cardiomyopathy, more naevi, lentigines, café au lait spots

KRAS more severe cognitive delay ; similar to cardiofaciocutaneous syndrome

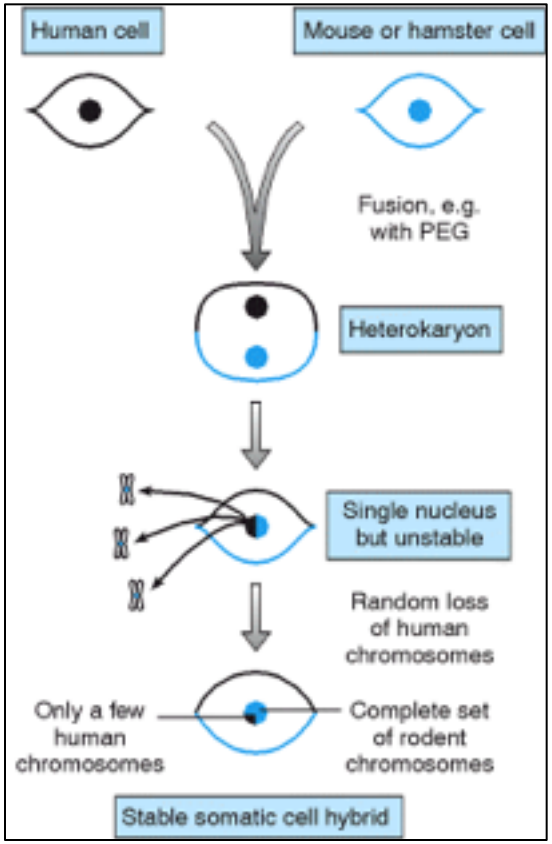
TUMOR SUPPRESSOR GENES

Cell fusion experiments showed that the transformed phenotype could be corrected in vitro by fusion of the transformed cell with normal cell (Stanbridge, 1982)

This provided evidences that tumorigenesis involves not only activated oncogenes but also loss of functions mutations in other genes

Tumor suppressor genes (TS) can have a variety of functions.
Their products mainly inhibit cell proliferation

Both alleles of a TS gene must be inactivated to change the behaviour of the cells



Knudson and the Retinoblastoma model

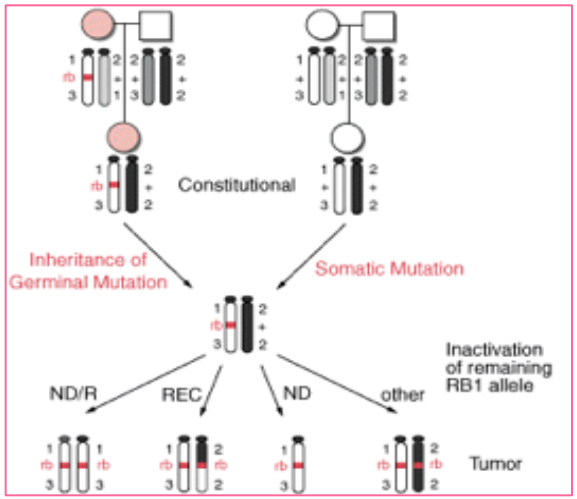
Retinoblastoma is a rare , aggressive childhood tumour of the retinal cells

- **60% of the cases are unilateral and onset by 30 months age**
- **40% of the cases are bilateral and onset by 14 months and these children can develop other tumors during their lives (generally osteosarcoma)**

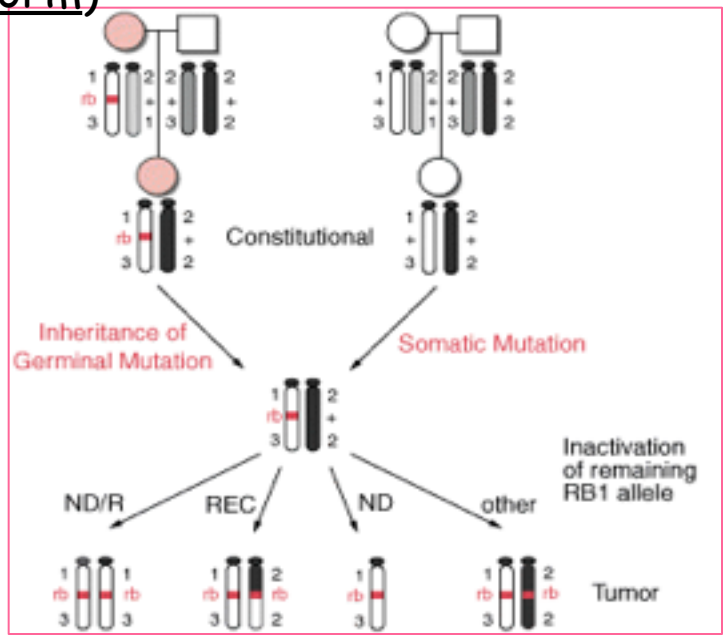
Knudson suggested that two successive and independent alterations were necessary to turn a normal cell into a tumour cell

Knudson's two «hits» hypothesis

In unilateral form the alterations are both acquired by the cells after birth (sporadic form)

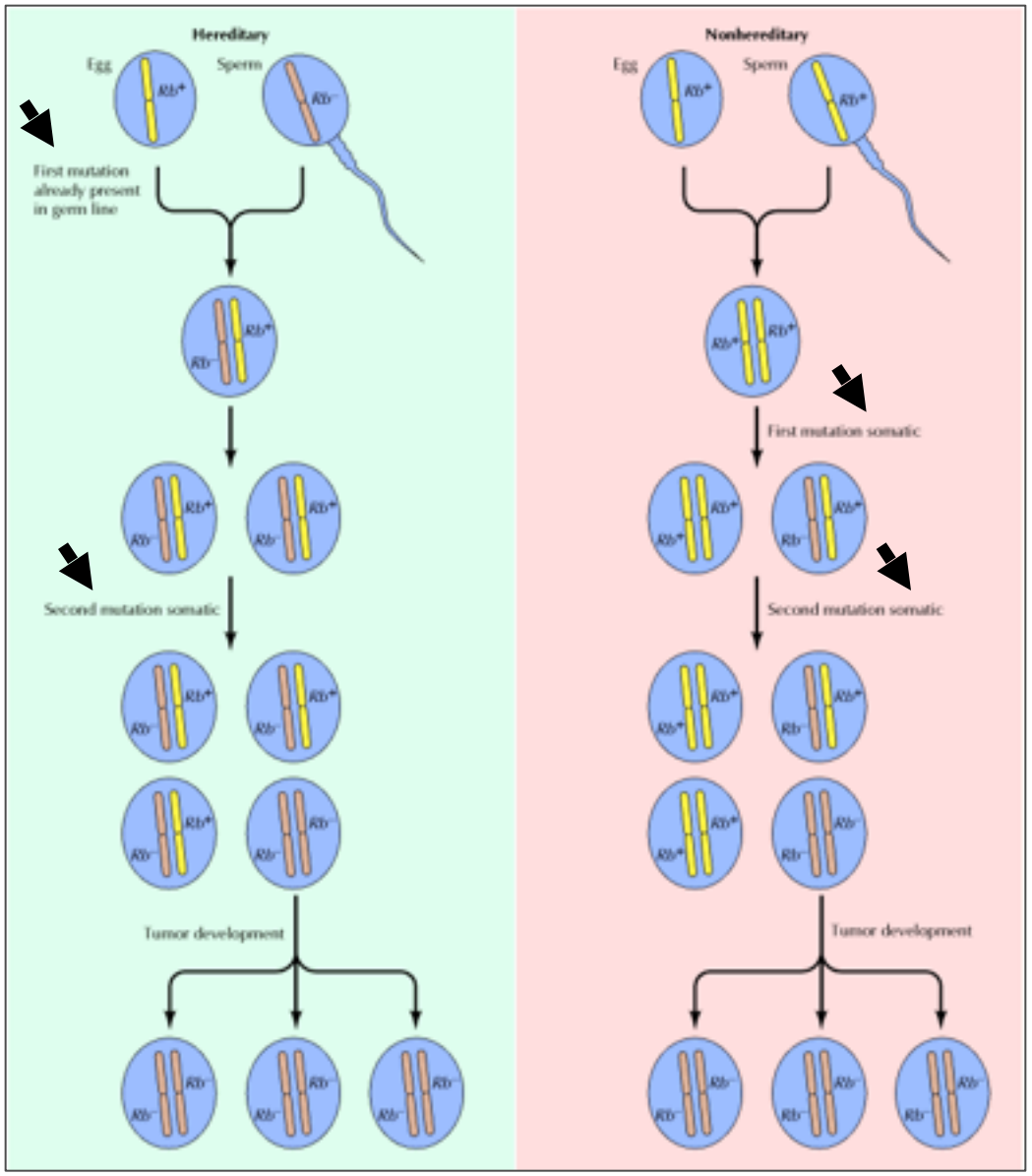


In bilateral form the first alteration is inherited whereas the second is acquired after birth (inherited form)



Susceptibility to Retinoblastoma

Autosomic dominant



Hereditary retinoblastoma:
- a defective copy of the Rb gene (Rb^-) is inherited from the affected parent
- a second somatic mutation, which inactivates the single normal Rb+ copy in a retinal cell, then leads to the development of retinoblastoma.

Non hereditary retinoblastoma: -
- two normal Rb+ genes are inherited, and retinoblastoma develops only if two somatic mutations inactivate both copies of Rb in the same cell

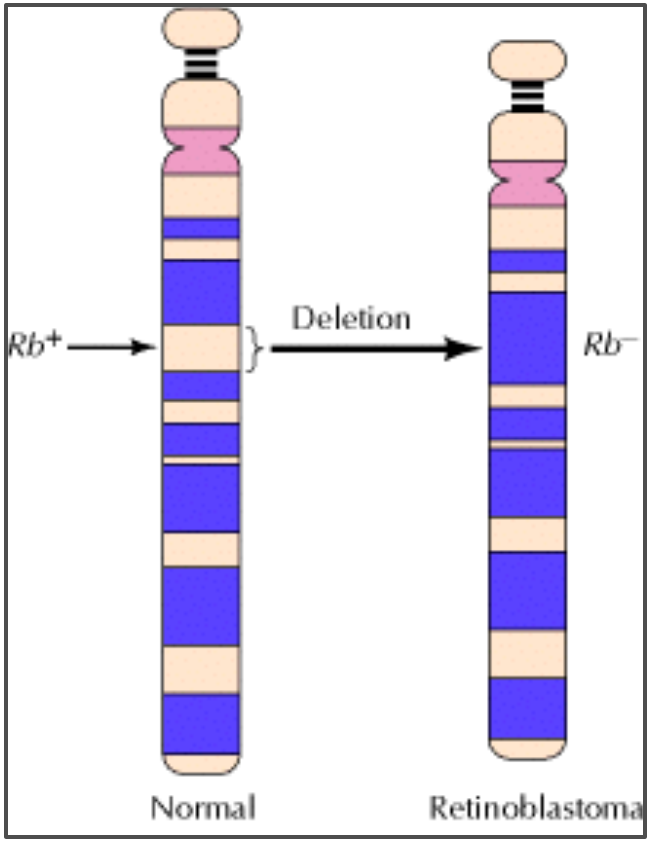
Investigations in Retinoblastoma families allowed to localize the gene on chromosome 13q14

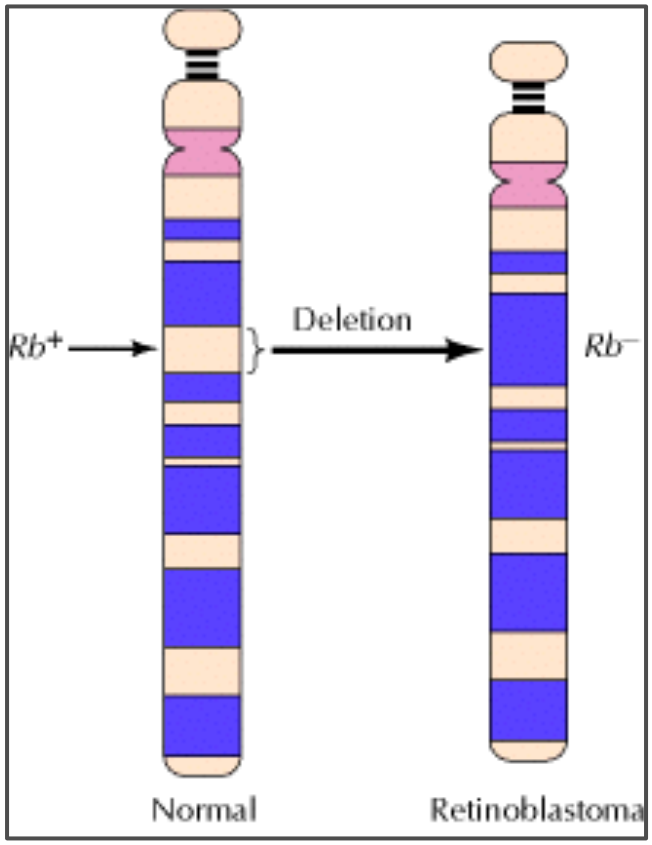
- Linkage analysis by using polymorphic markers mapping on chromosome 13q14

- 5-10% showed **constitutional deletions** involving part or the entire 13q14 band including RB1 locus

Cavenee and colleagues compared blood and tumour samples from the same patients: tumours cells were apparently homozygous for some markers because they have acquired a somatic deletion (the second hit)

Cavenee et al., 1983 proved Knudson's hypothesis and established a paradigm for all the investigations on TSs





Retinoblastoma (OMIM 180200) is a model of human hereditary tumours

Susceptibility is in linkage with mutant allele (missense mutations, deletions etc) at high penetrance (80-90%)

So far identified 930 type of germline alterations

Inactivation of a tumor-suppressor genes

- Both the alleles loss their function
- Mechanisms: **mutations, deletions, methylation**

Loss of heterozygosity (LOH)

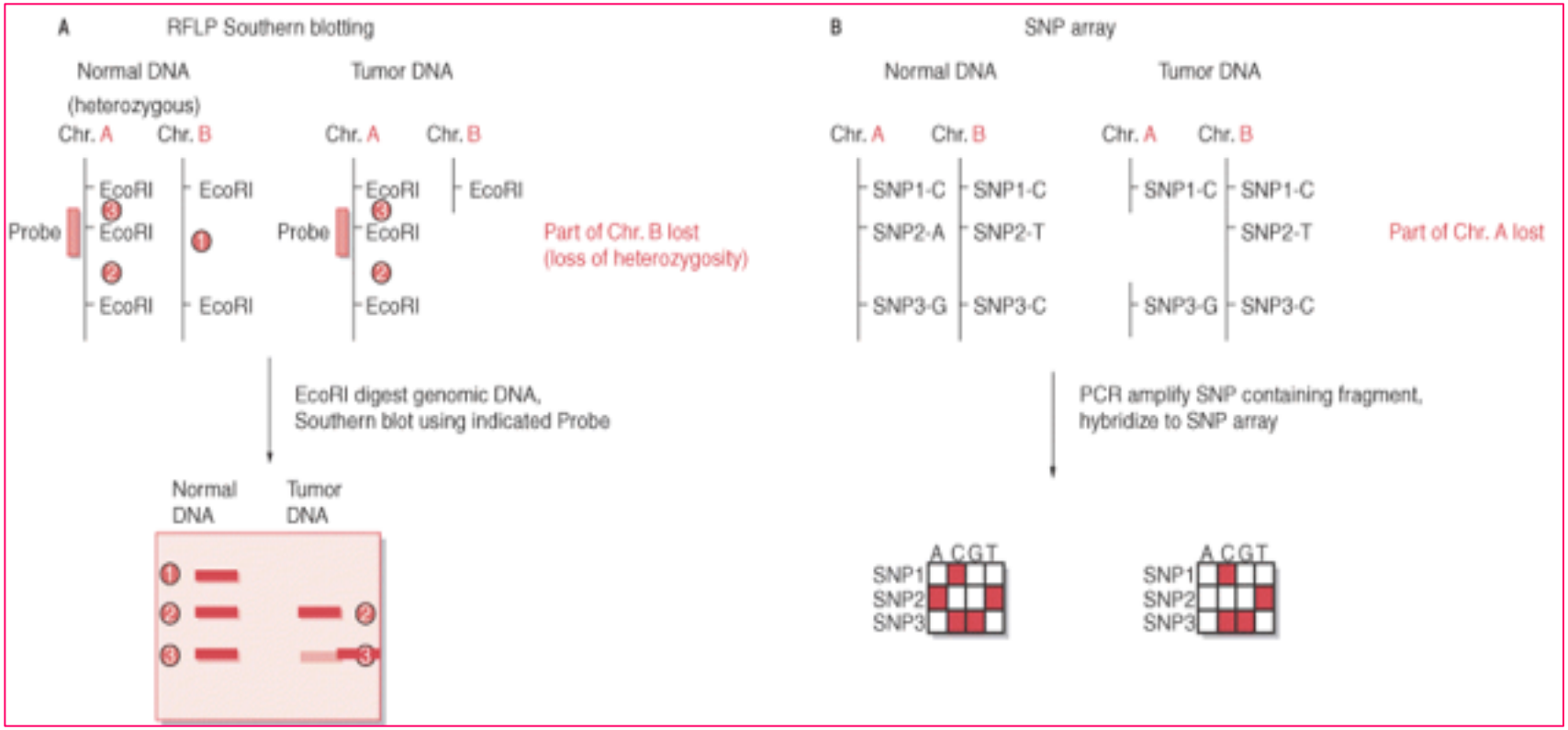
Loss of function of one allele of the gene when the other **has been already inactivated** by mutation/deletion/methylation

Heterozygosity is lost for the onset of a second mutation,deletion or methylation

Molecular analysis:

- cytogenetics
- DNA polymorphisms

Loss of heterozygosity (LOH)



Heterozygote : two different alleles for a locus

LOH : comparison of tumor DNA and germline DNA with the loss of one allele

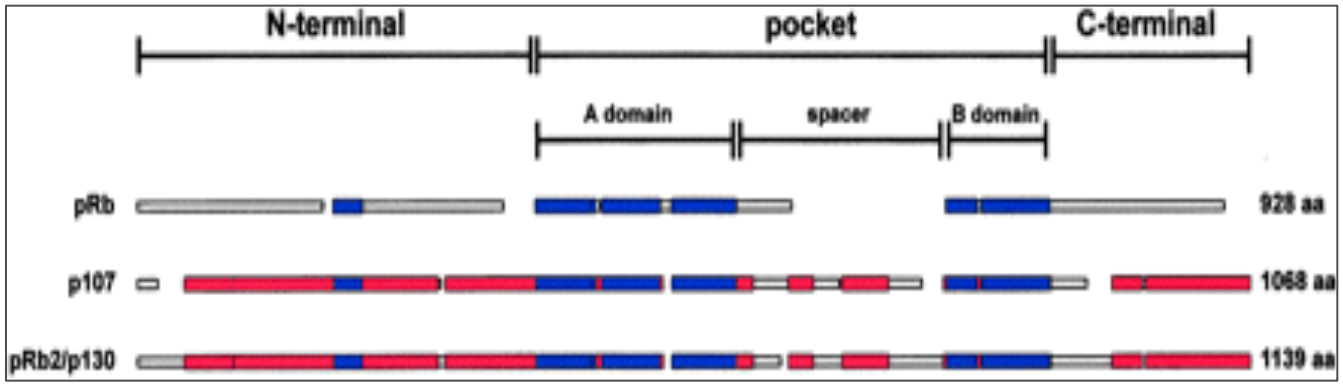
Tumor Suppressor Genes and Associated Protein Function

Gene	Associated inherited cancer syndrome	Cancers with somatic mutations	Presumed function of protein
RB1	Familial retinoblastoma	Retinoblastoma, osteosarcoma, SCLC, breast, prostate, bladder, pancreas, esophageal, others	Transcriptional regulator; E2F binding
TP53	Li-Fraumeni syndrome	Approximately50% of all cancers (rare in some types, such as prostate carcinoma and neuroblastoma)	Transcription factor; regulates cell cycle and apoptosis
p16	Familial melanoma, Familial pancreatic carcinoma	Approximately 25-30% of many different cancer types (eg, breast, lung, pancreatic, bladder)	Cyclin-dependent kinase inhibitor (ie, Cdk4 and Cdk6)
	Familial melanoma	Approximately 15% of many different cancer types	Regulates Mdm-2 protein stability and hence p53 stability; alternative reading frame of *p16/INK4a gene*
APC	Familial adenomatous polyposis coli (FAP), Gardner syndrome,	Colorectal, desmoid tumors, thyroid cancers, stomach cancers	Regulates levels of β-catenin protein in the cytosol; binding to microtubules
BRCA1	Inherited breast and ovarian cancer	Ovarian (~10%), rare in breast cancer	DNA repair; complexes with Rad 51 and BRCA2; transcriptional regulation
BRCA2	Inherited breast (both female andmale), pancreatic cancer, ?others?	Rare mutations in pancreatic, ?others/	DNA repair; complexes with Rad 51 and BRCA1
WT-1	WAGR, Denys-Drash Syndrome	Wilms' tumor	Transcription factor

NF-1	Neurofibromatosis type 1	Melanoma, neuroblastoma	p21ras-GTPase
NF-2	Neurofibromatosis type 2	Schwannoma, meningioma, ependymoma	Juxtamembrane link to cytoskeleton
VHL	von-Hippel Lindau syndrome	Renal (clear cell type), hemangioblastoma	Regulator of protein stability
MEN-1	Multiple endocrine neoplasia type 1 Endocrine tumors of the pancreas	Parathyroid adenoma, pituitary adenoma, enocrine tumors of the pancreas	Not known
PTCH	Gorlin syndrome, hereditary basal cell carcinoma syndrome	Basal cell skin carcinoma, medulloblastoma	Transmembrane receptor for sonic hedgehog factor; negative regulator of smoothened protein
PTEN/MMAC1	Cowden syndrome; sporadic cases of juvenile polyposis syndrome	Glioma, breast, prostate, follicular thyroid carcinoma, head and neck squamous carcinoma	Phosphoinositide 3-phosphatase; protein tyrosine phosphatase
DPC4	Familial juvenile polyposis syndrome	Pancreatic(~50%), approximately 10–15% of colorectal cancers, rare in others	Transcriptional factor in TGF-β signaling pathway
E-CAD	Familial diffuse-type gastric cancer; lobular breast cancer	Gastric (diffuse type), lobular breast carcinoma, rare in other types (eg, ovarian)	Cell-cell adhesion molecule
LKB1/STK1	Peutz-Jeghers syndrome	Rare in colorectal, not known in others	Serine/threonine protein kinase
SNF5/INI1	Rhabdoid predisposition syndrome renal or extra-renal malignant rhabdoid tumors), choroid plexus carcinoma medulloblastoma; central primitive neuroectodermal tumors)	Rare in rhabdoid tumors, choroid plexus carcinoma, medulloblastoma	Member of the SWI/SNF chromatin ATP-dependent remodeling complex
EXT1 / EXT2	Hereditary multiple exostoses	Not known	Glycosyltransferase; heparan sulfate chain elongation
TSC1 / TSC2	Tuberous sclerosis	Not known	Not known; cytoplasmic vesicle localization
MSH2, MLH1 PMS1, PMS2, MSH6	Hereditary non-polyposis colorectal cancer	Colorectal, gastric, endometrial	DNA mismatch repair

Rb

Nuclear phosphoprotein 105 kDa = p105Rb / pRb
The RB family: p105, p107, p130

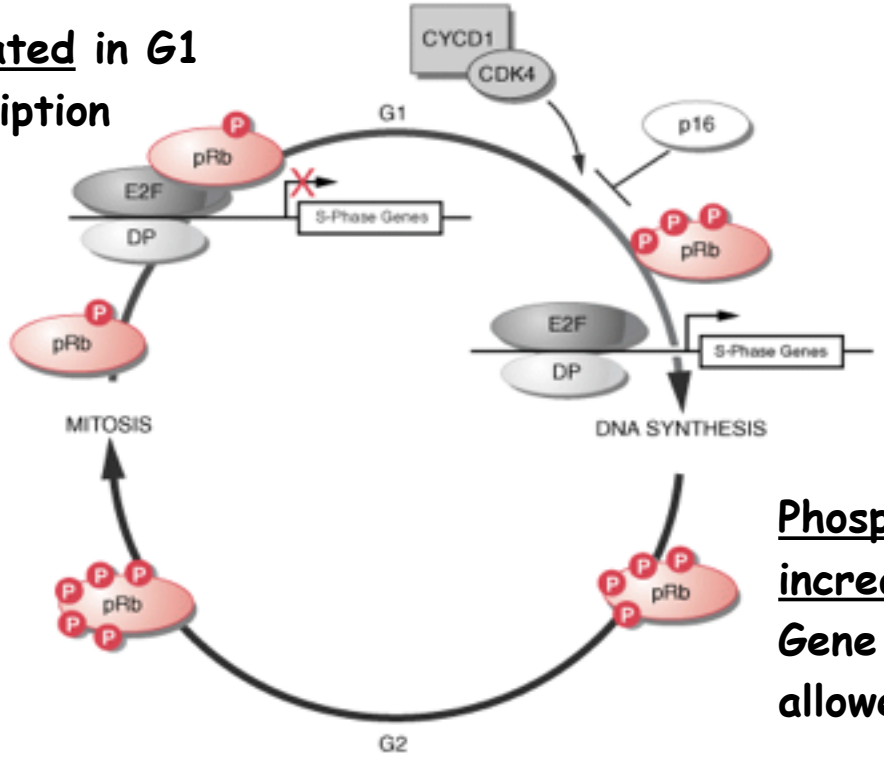


Homologies among the RB family proteins.

Blue boxes, homologous regions among the three proteins

pRb is <u>hypophosphorylated</u> in G1

Suppression of transcription



<u>Phosphorylation decreases</u> in mitosis

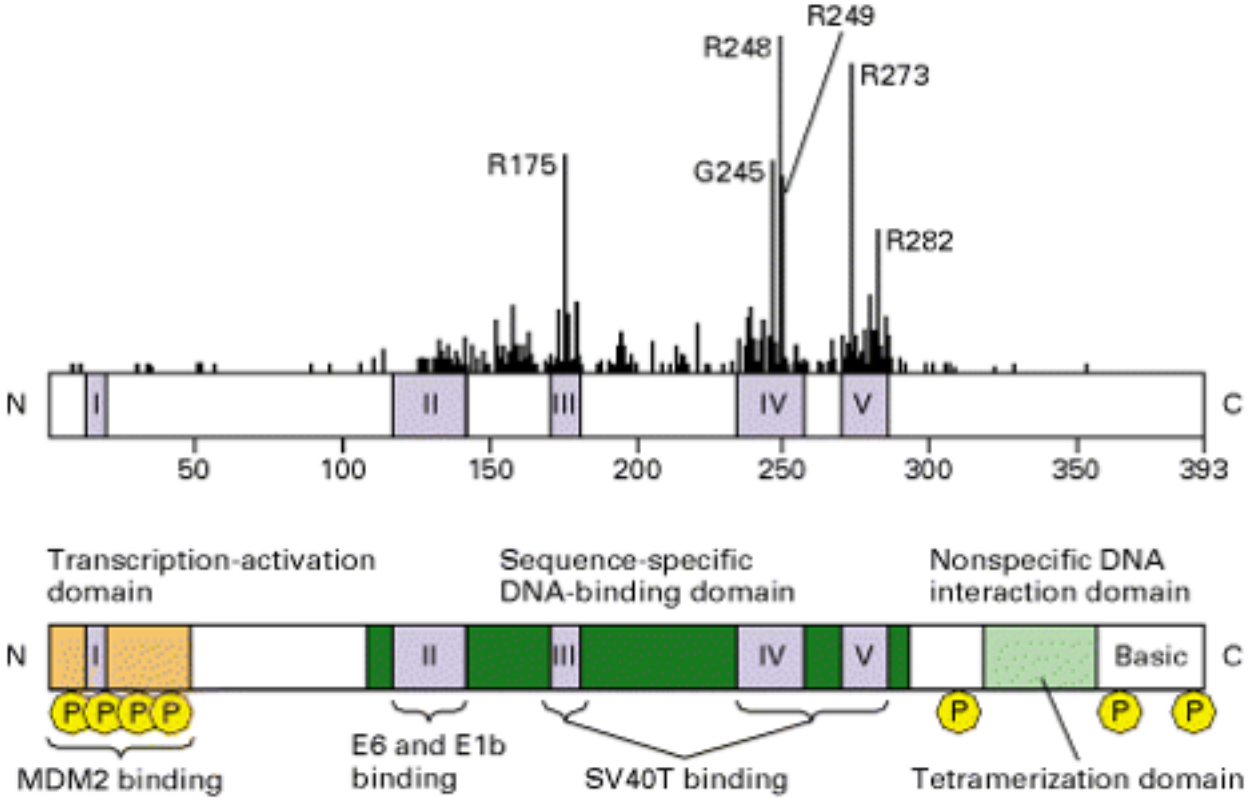
<u>Phosphorylation increases</u> in S phase

Gene transcription is allowed

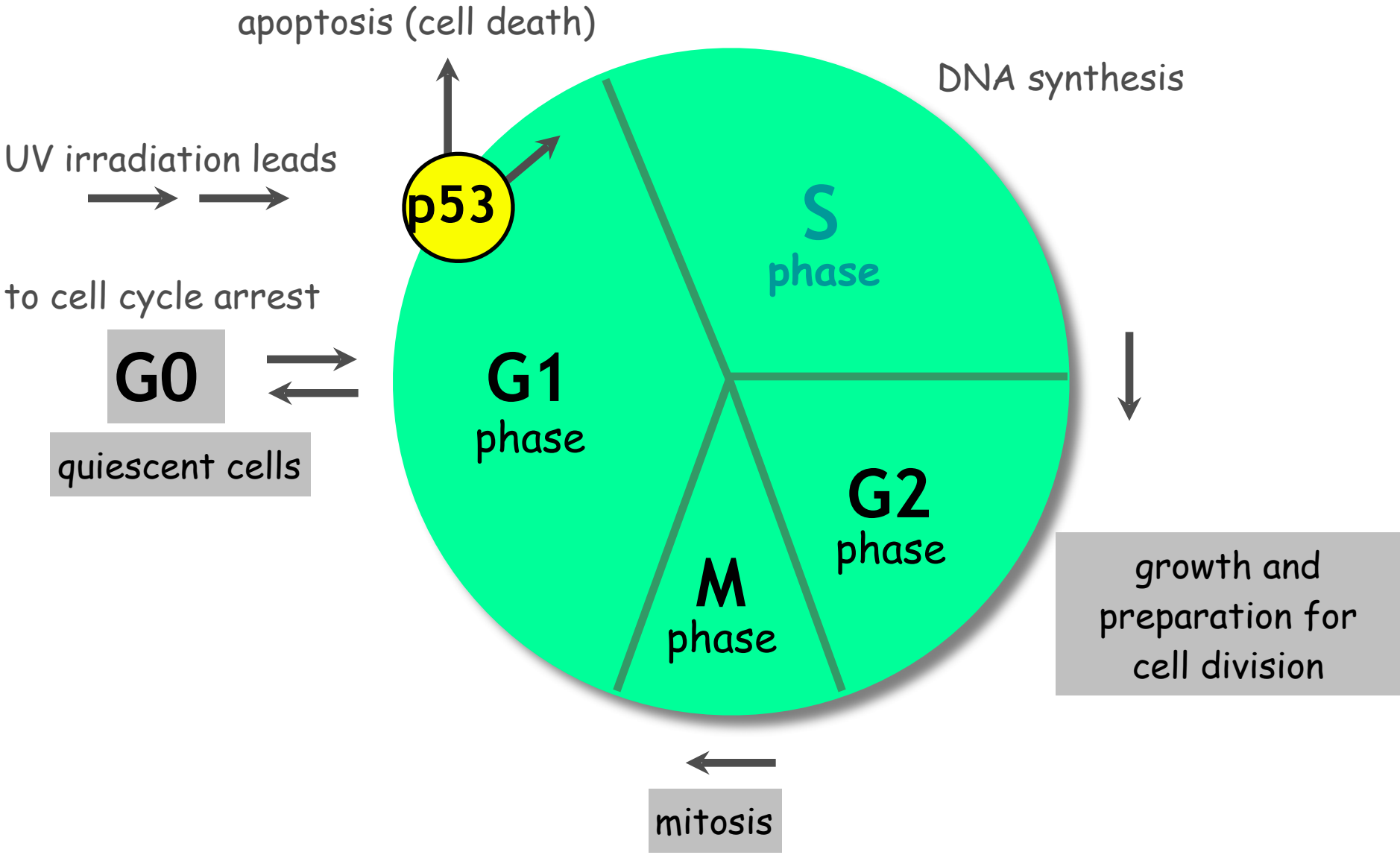
p53

"Guardian of the genome"

- somatic mutations very common in different types of human tumours
- germline mutations associated with Li-Fraumeni syndrome
- p53 is a transcription factor involved in DNA repair
- it stops cell-cycle to repair DNA; when it is mutated cells lose this function (loss of function mutations)

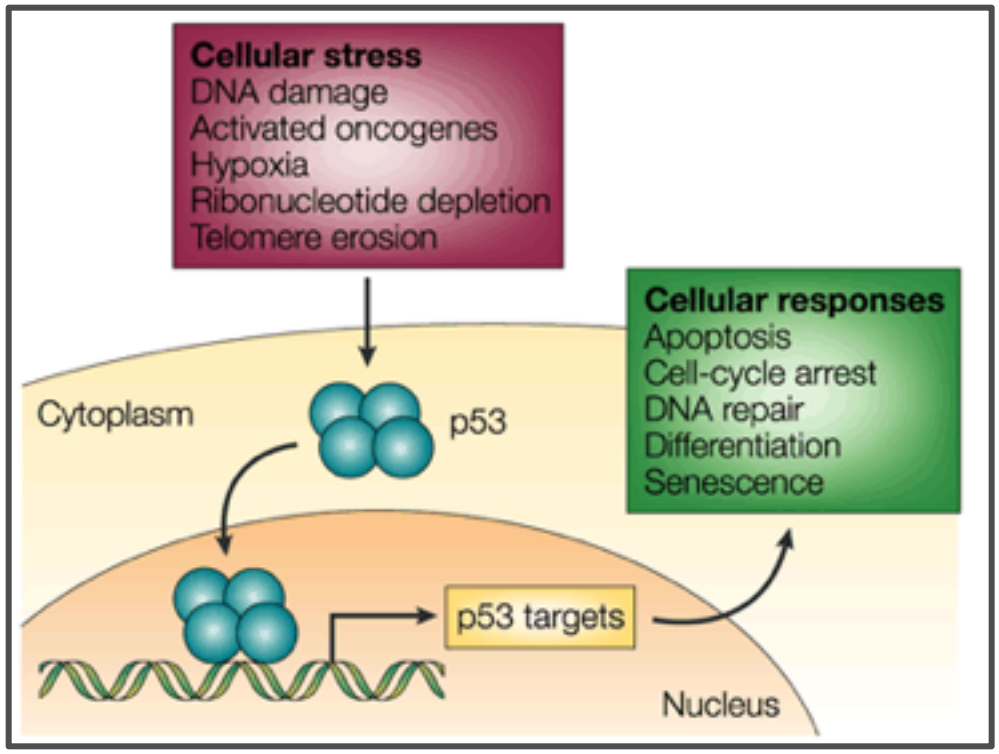


The role of p53 in the cell cycle



THE p53-MEDIATED RESPONSE

- In non-stressed cells p53 exists at a very low concentration



- Under stress conditions, the p53 protein accumulates in the cell
- Binds in its tetrameric form to p53-response elements
- Induces the transcription of genes involved in cell-cycle control, apoptosis, DNA repair, differentiation and senescence

- The loss of p53 tumour-suppressor activity by mutation/deletion or inhibition allows the proliferation of the damaged cells under the stress conditions.

- This uncontrolled proliferation can lead to tumour development.

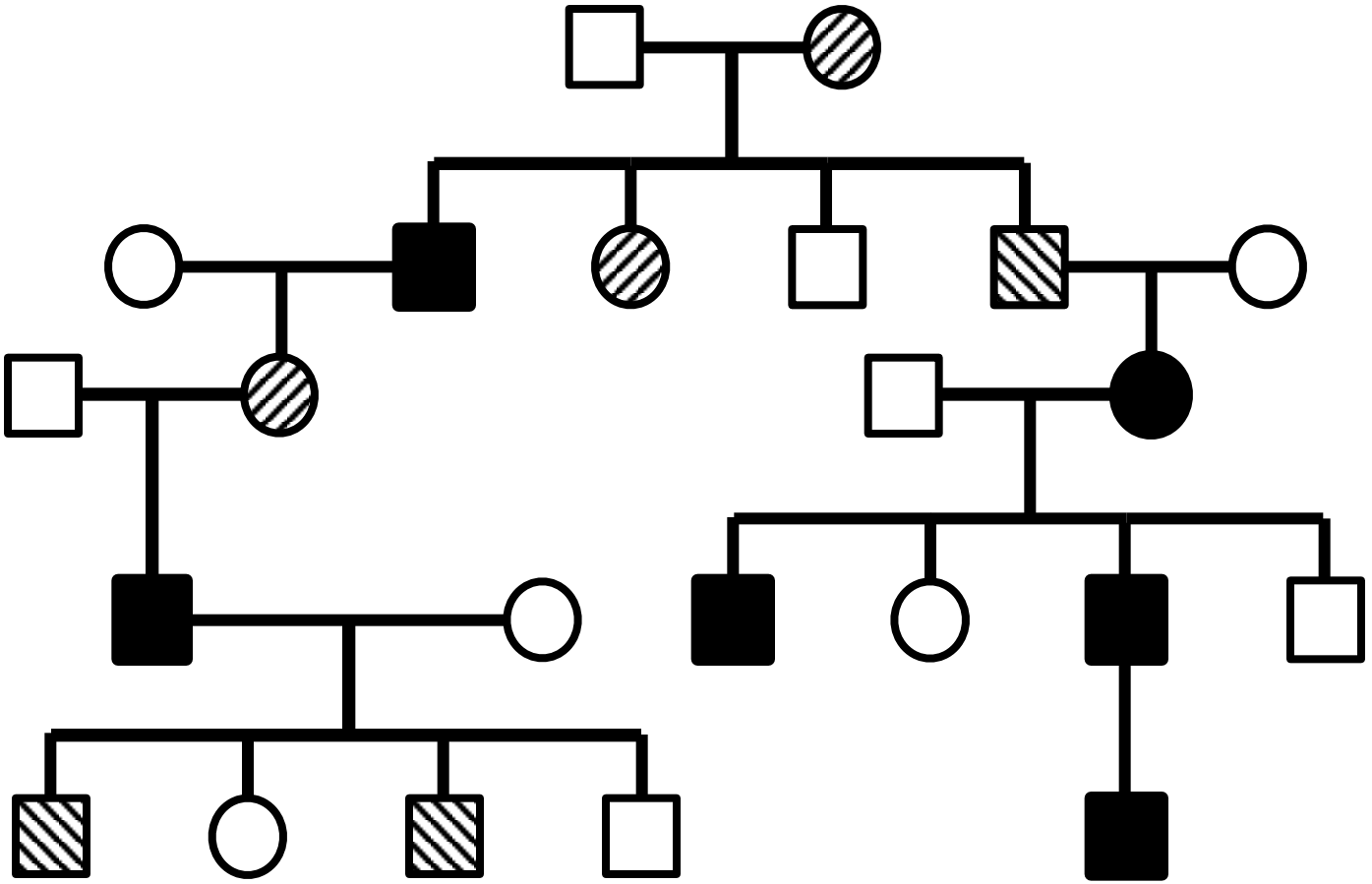
Chène, Nat Rev Cancer 3: 102 (2003)

In human sporadic tumors the most frequent p53 mutations are missense mutations in DNA binding domain that can confer a gain of function (dominant negative)

Li-Fraumeni syndrome

- It is a very rare syndrome described by Li and Fraumeni (1969)
- Families with frequent early onset sarcomas and osteosarcomas, breast cancers, brain tumours, and leukemias (SBLA)
- In 1990 p53 germline mutations were associated with this syndrome
- Transmission is autosomic dominant
- KO Mice for p53 show Li-Fraumeni tumours

Li-Fraumeni syndrome - caused by mutations in the p53 gene



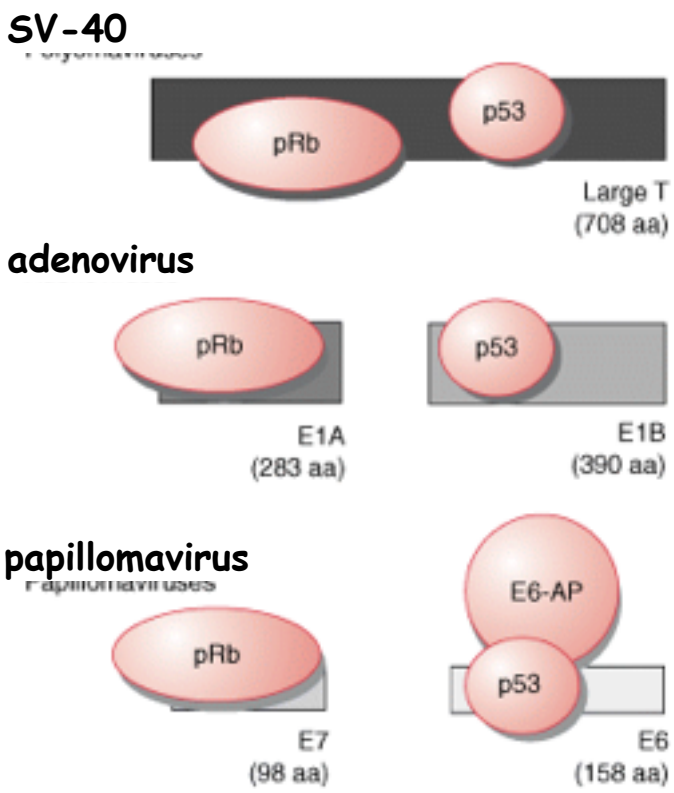
Breast cancer

Sarcoma

Other malignant neoplasms

There are multiple neoplasms (SBLA) in Li-Fraumeni families that are inherited in an autosomal dominant fashion

Interactions between tumor-suppressor gene products and proteins encoded by DNA tumor viruses



- Large T antigen from polyomaviruses (such as simian virus 40 binds both the retinoblastoma (pRb) and p53 proteins.
- For the **adenoviruses** and the high-risk human **papillomaviruses** (HPV types 16 and 18), different viral protein products complex with pRb and p53.
- A cellular protein known as E6-associated protein (E6-AP) cooperates with the HPV E6 protein to complex and degrade p53.

Selective proliferation

↓

Tumour growth

↓

GATEKEEPER

a gene responsible of the homeostatsis of a specific cell type

Gatekeepers are frequently tumor-suppressor genes

- They have been mainly identified studying hereditary tumours
- Each cell types have specific gatekeeper
- They are frequently mutated in the hereditary and sporadic form of the same tumour type

- **APC for colonocytes**
- **NF2 for Schwann cells**
- **VHL for renal cells**

GENETIC MARKERS

A genetic marker is a known DNA sequence containing a variant (heterozygotes)

- Genetic markers are helpful to identify an association between a hereditary disease and the genetic cause
- Characteristics: traceable, associated with a specific locus, highly polymorphic (no information in homozygous)

POLYMORHISMS

- 1975- Restriction Fragment Length Polymorphisms (RFLP)
- 1995- Single Nucleotide

REPEATED SEQUENCES

- 1985- Variable Number of Tandem Repeats (VNTR)- MINISATELLITE
- 1990- Short Tandem Repeats- MICROSATELLITE

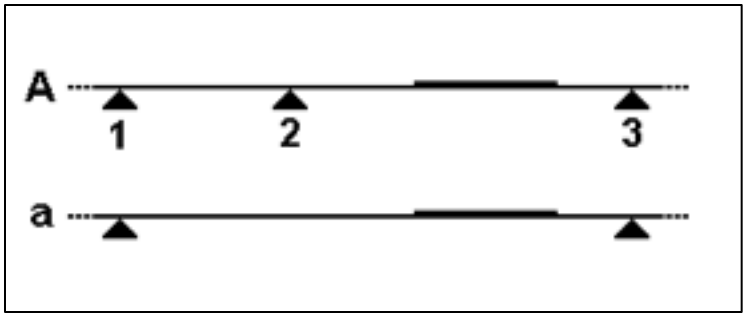
FIRST GENERATION GENETC MARKERS:

RFLP AND VNTR (MINISATELLITES)

DNA variations that can be identified with:

Restriction enzymes and egel-electrophoresis

RFLP (Restriction Fragment Length Polymorphism)

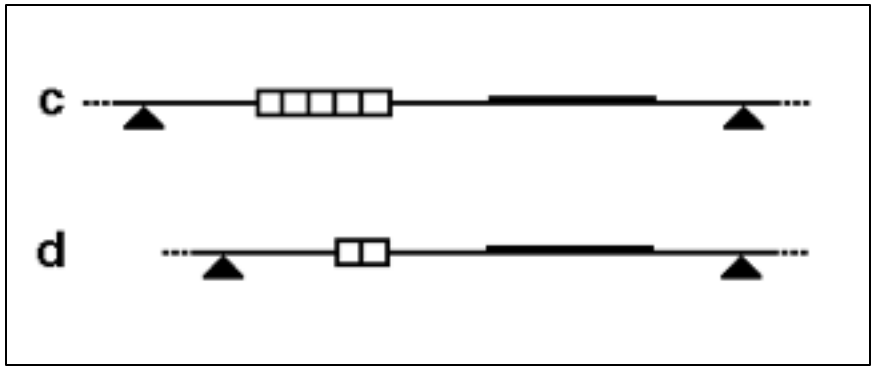


Single nucleotide variation in the two alleles

VNTR (Variable Number of Tandem Repeats)

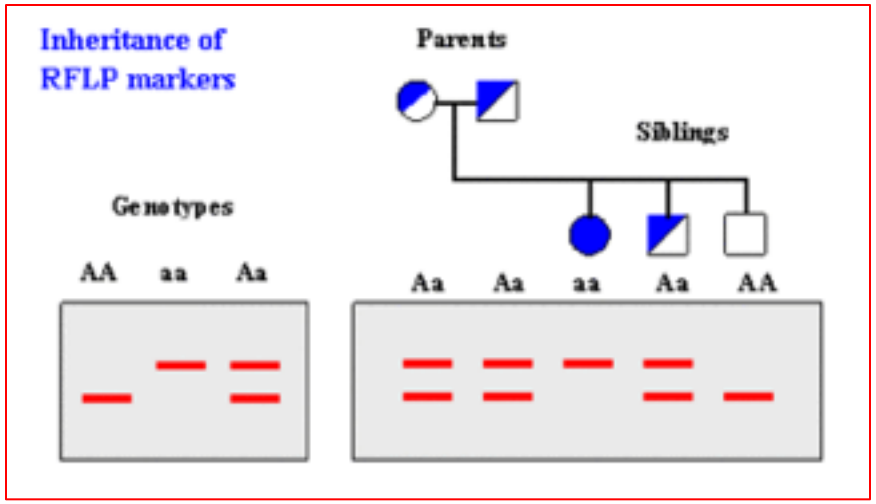
Variation in the number of repeats in the two alleles

50-70 bp repeated sequences, very frequent in sub-telomeric regions

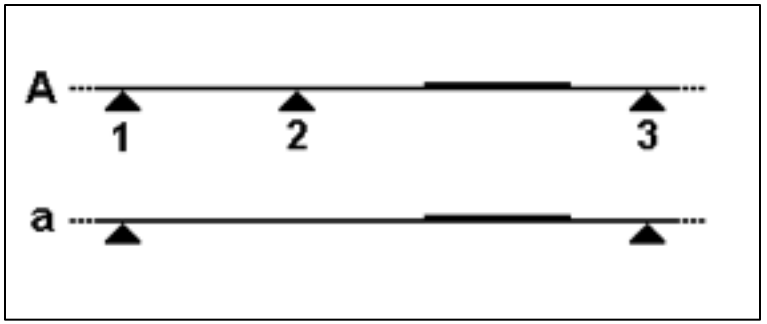


RFLP Analysis

Gel separation



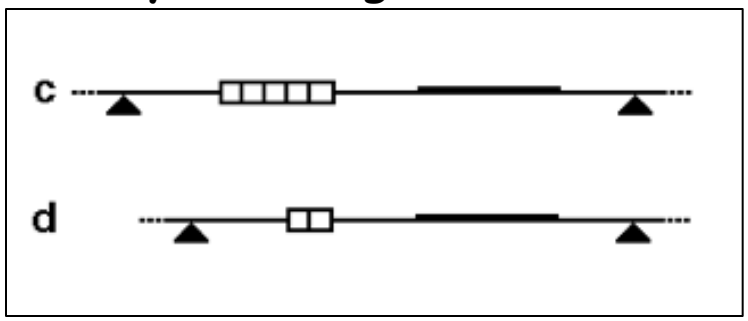
Enzymatic digestion



Informativity max 50%

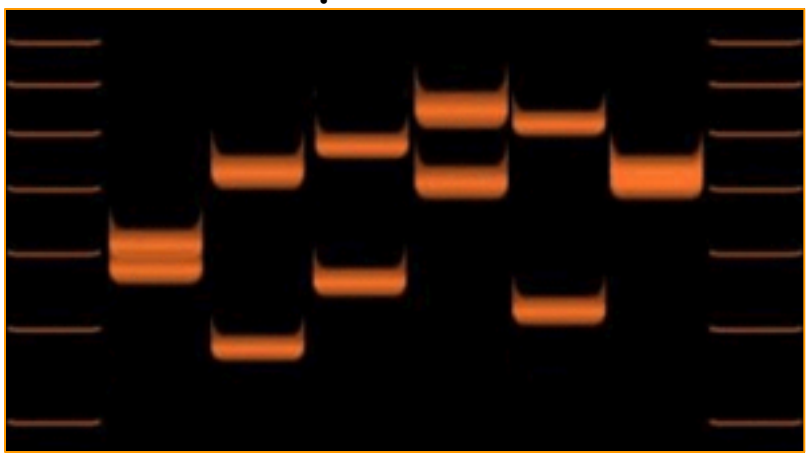
VNTR analysis

Enzymatic Digestion



Informativity max 80%

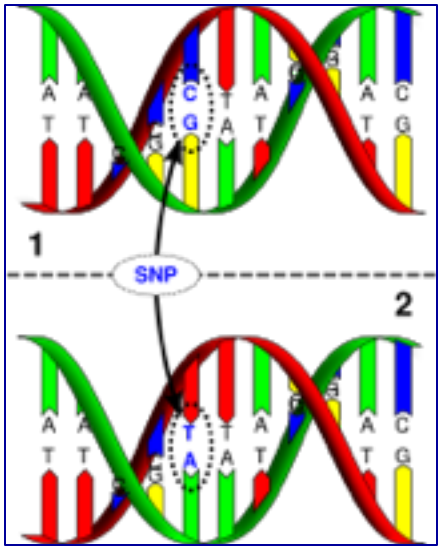
Gel separation



Every change in the DNA sequence (single base mutations, deletions, translocations and inversions) can cause a change in the RFLP

This can be used to follow the disease segregation in patients with the same RFLP pattern

Human Genome Project results led to use Single Nucleotide Polymorphism (SNPs) instead of RFLP



SNP = is a single base DNA variant (A,T,C o G) among individuals of the same species

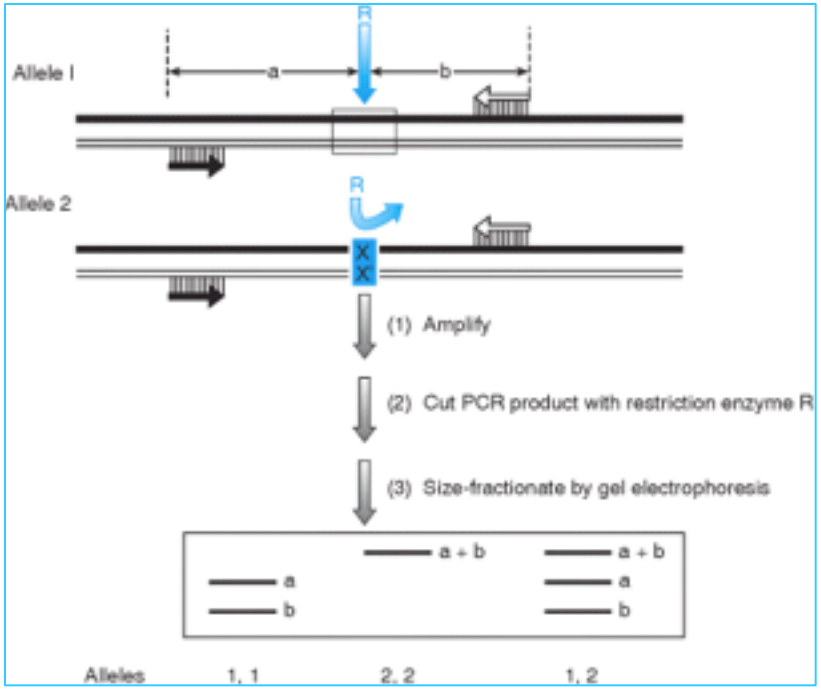
A DNA variant can be considered a SNP when can be found in at least 1% of the individuals

GENETIC MARKERS OF SECOND GENERATION:

SNPs AND MICROSATELLITEs

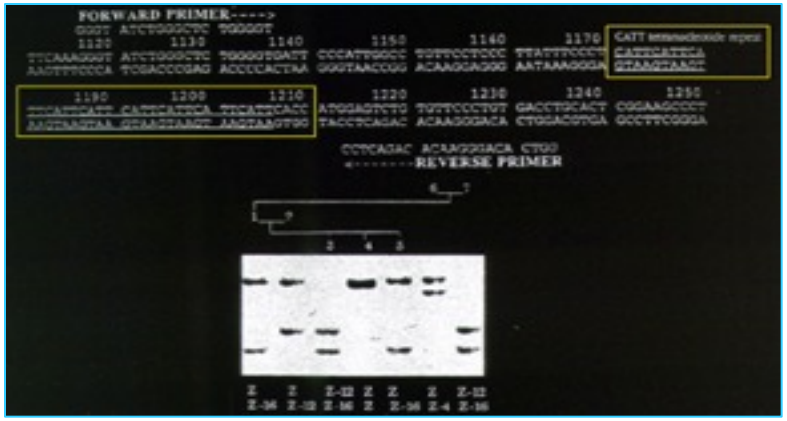
Can be used with PCR

SNPs



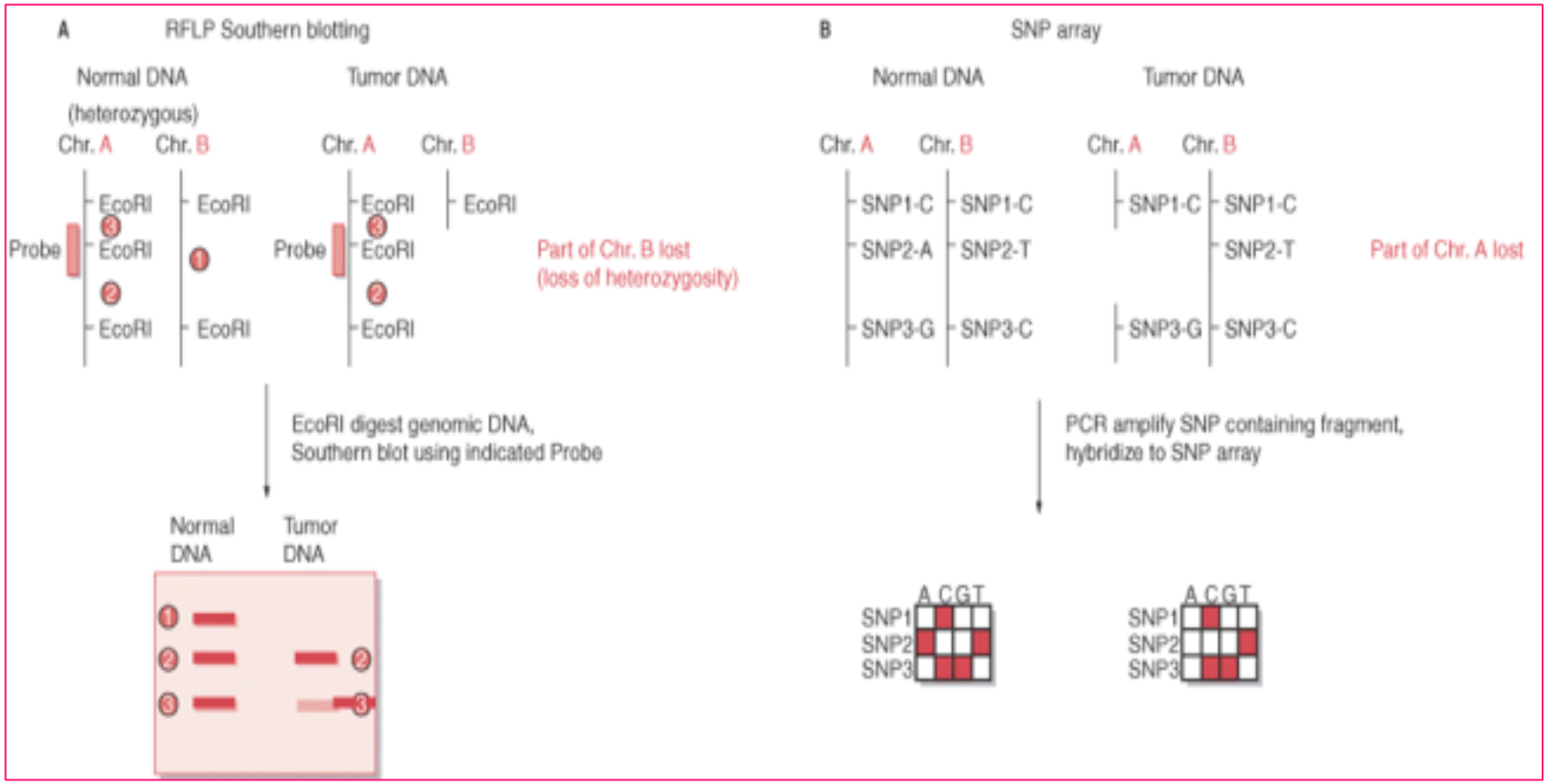
Microsatellites

Simple Short Repeats (SSR)



1-6 bp short repeats
frequently (CA)n

Loss of heterozygosity (LOH)



Loss of heterozygosity (LOH)

Germline DNA is compared with tumour DNA in heterozygote individuals:

LOH is found when in tumour DNA one band disappears

Deletions detected with microsatellites