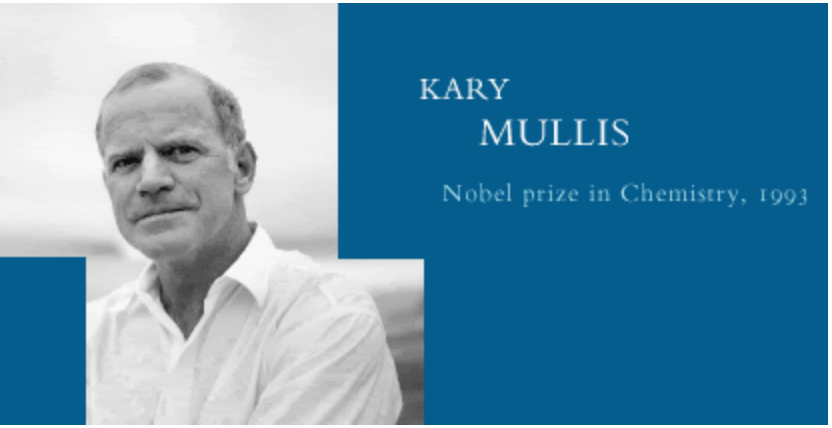
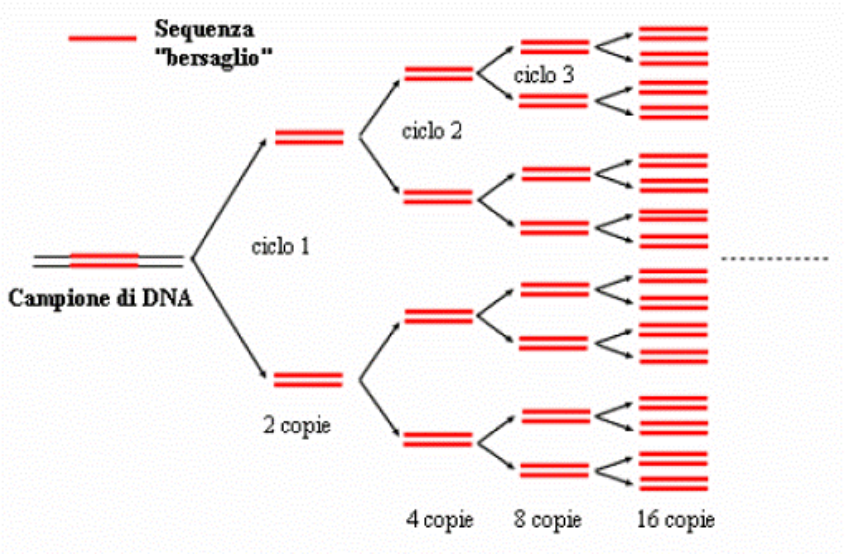
DIAGNOSI GENETICA MOLECOLARE E NUOVE TECNOLOGIE

PCR: REAZIONE A CATENA DELLA POLIMERASI

1985 – Kary Mullis





Applications of PCR

Molecular Identification

- Molecular Archaeology
- Molecular Epidemiology
- Molecular Ecology
- DNA fingerprinting
- Classification of organisms
- Genotyping
- Pre-natal diagnosis
- Mutation screening
- Drug discovery
- Genetic matching
- Detection of pathogens

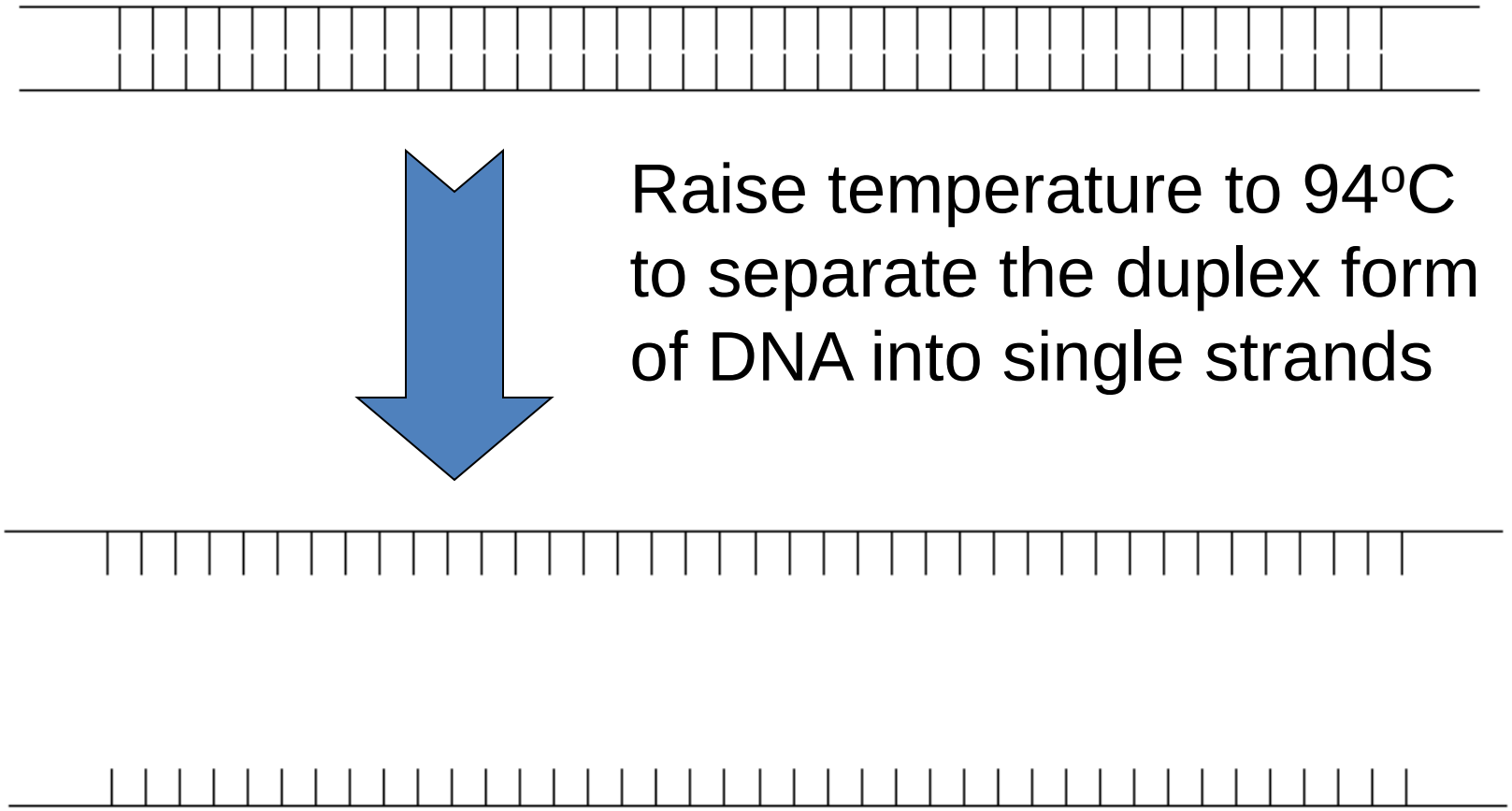
Sequencing

- Bioinformatics
- Genomic cloning
- Human Genome Project

Genetic Engineering

- Site-directed mutagenesis
- Gene expression studies

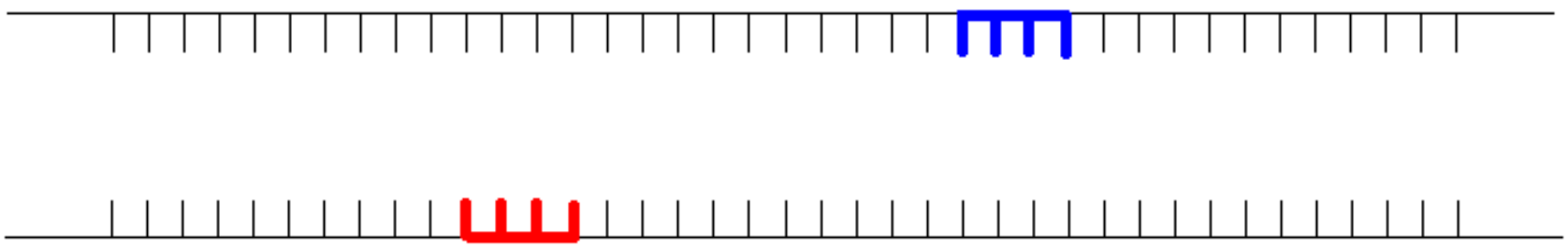
Step 1. Denaturation



Raise temperature to 94^{o}C to separate the duplex form of DNA into single strands

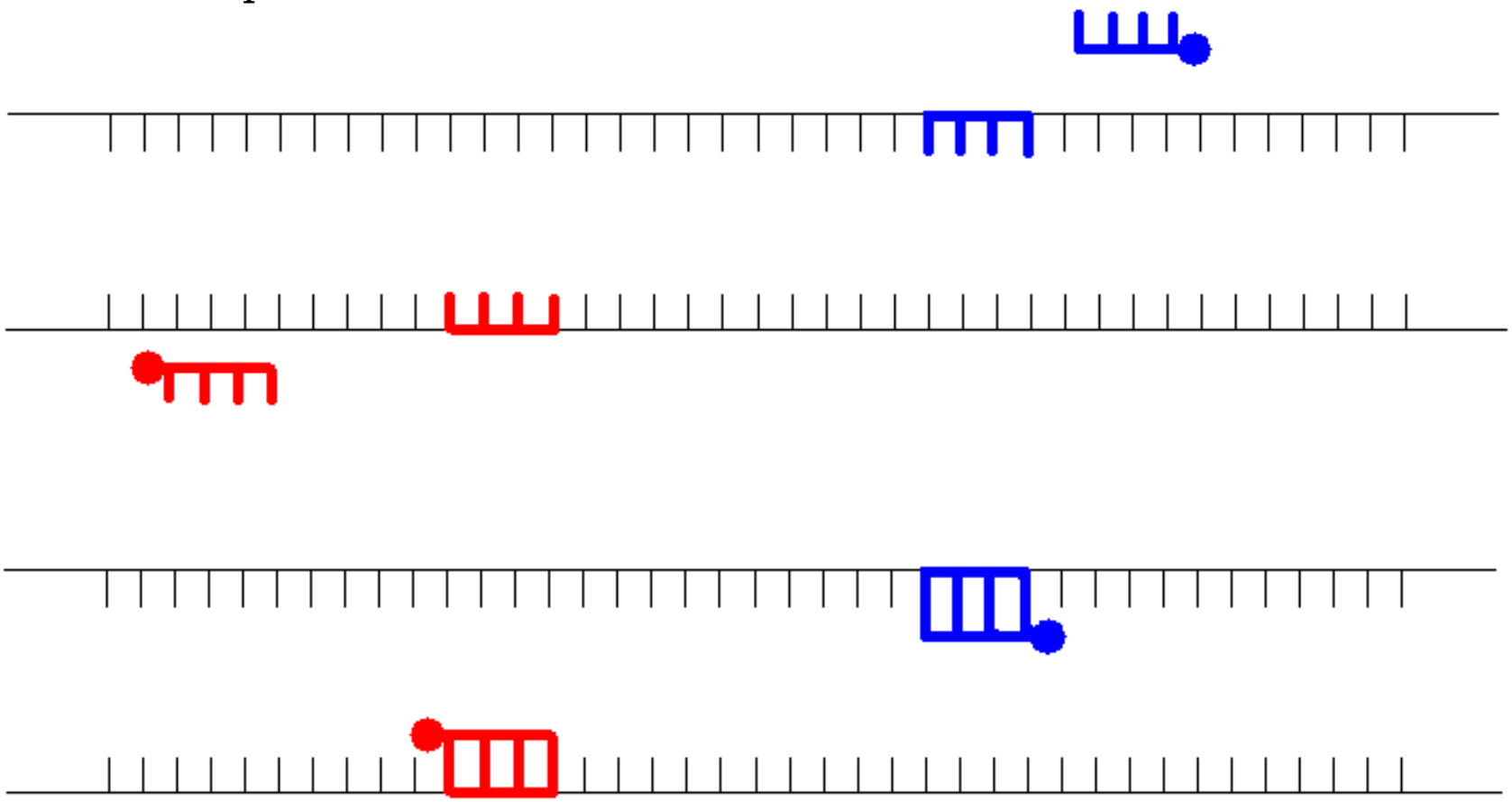
Design primers

- To perform PCR, a 10-20bp sequence on either side of the sequence to be amplified must be known because DNA pol requires a primer to synthesize a new strand of DNA



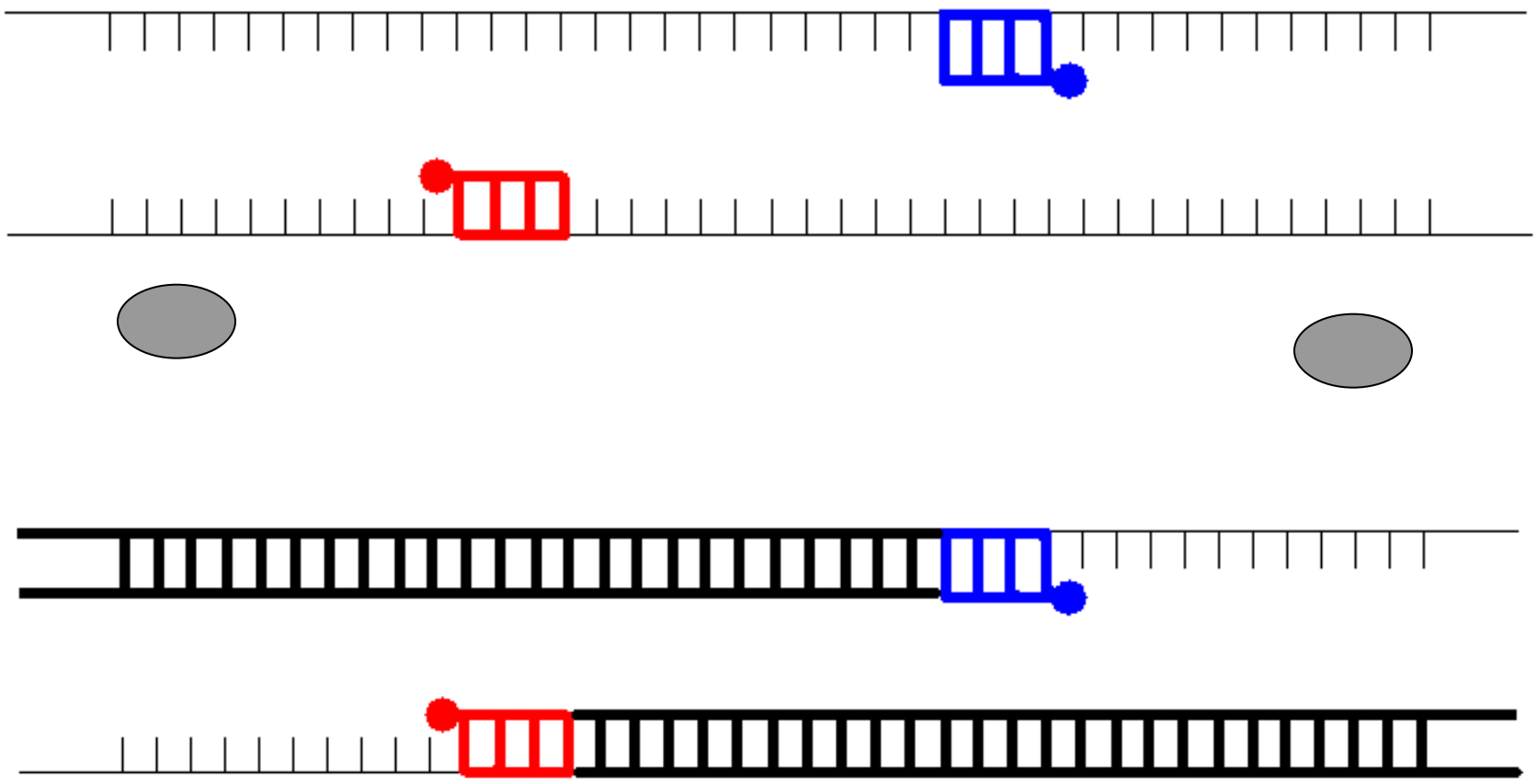
Step 2. Annealing

- Anneal primers at 50-65^{o}C



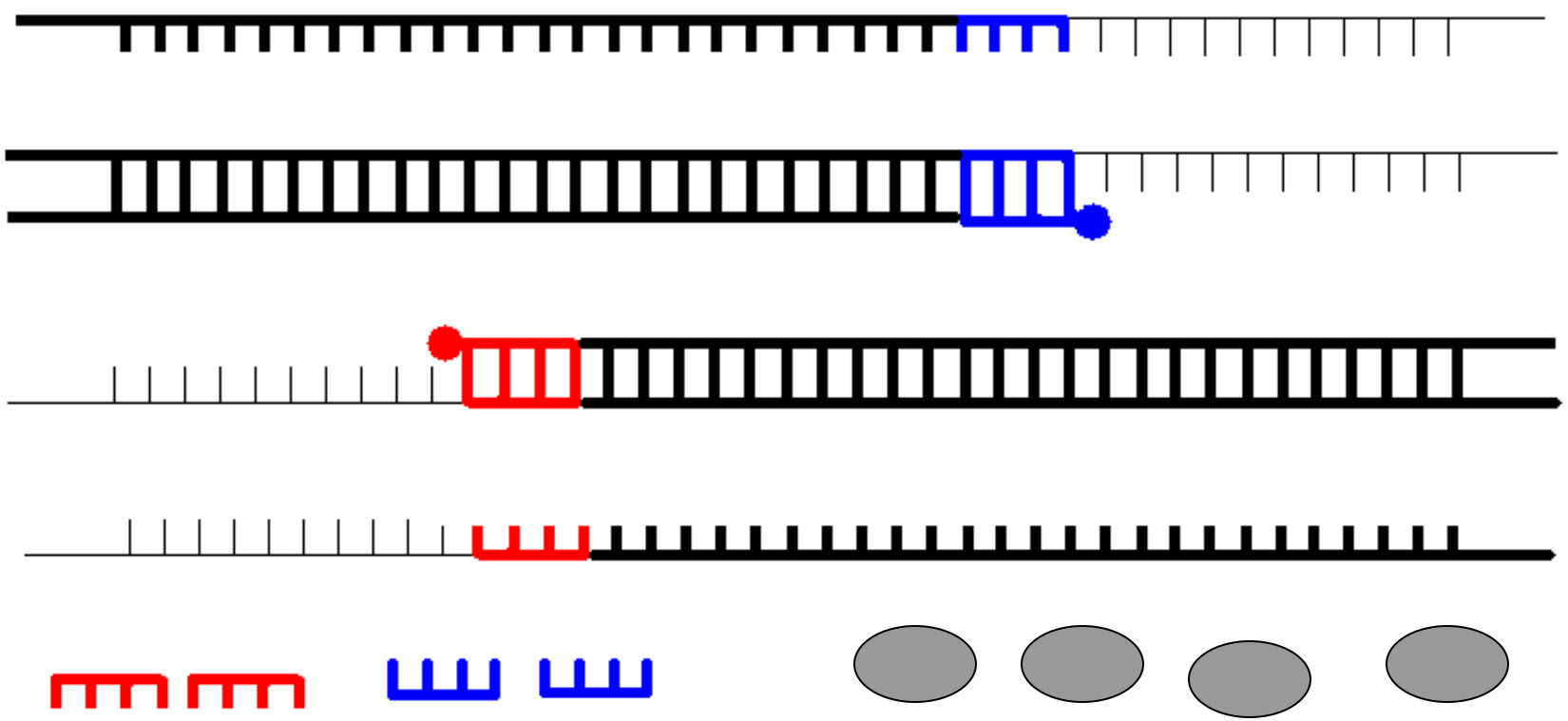
Step 3.Extension

- Extend primers: raise temp to 72°C, allowing Taq pol to attach at each priming site and extend a new DNA strand



Repeat

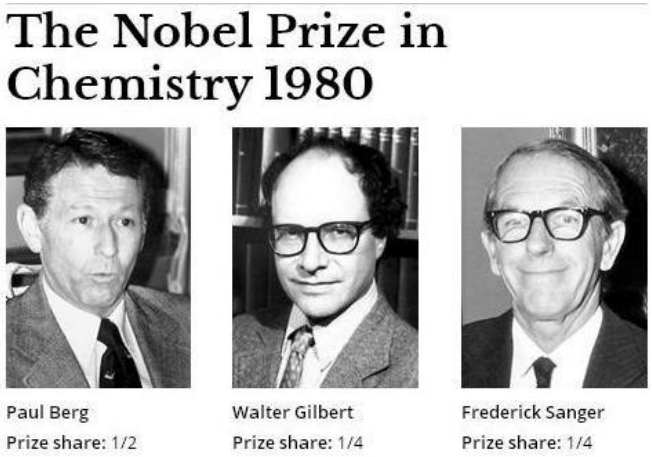
- Repeat the Denature, Anneal, Extension steps at their respective temperatures…

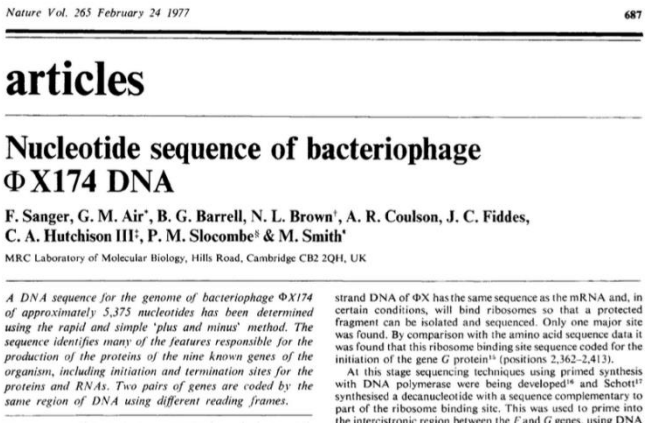


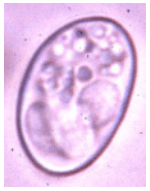
Come si è sviluppata questa nuova tecnologia?

SEQUENZIAMENTO DEL DNA

1977 – Il DNA è sequenziato per la prima volta da Fred Sanger, Walter Gilbert e Allan Maxam in maniera indipendente. Ottenuta la sequenza completa dell'intero genoma del fago Φ-X174





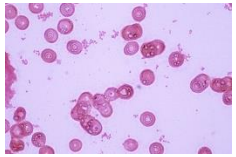


1994: ***H. Influenzae***
1.8 Mbp
(Fleischmann et al.)

2001: ***Homo Sapiens***

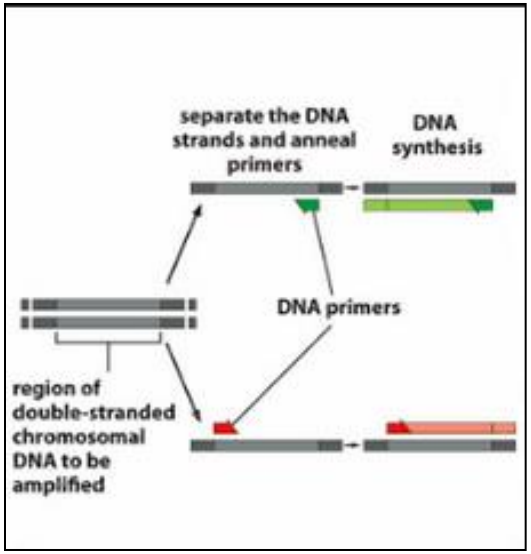
1977

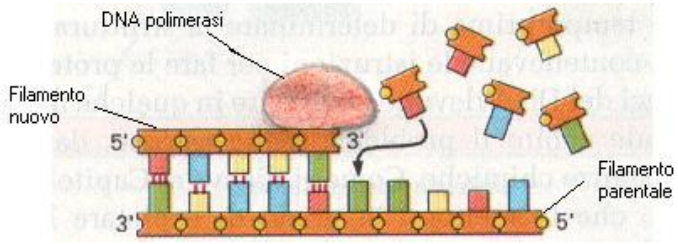
1982: ***lambda virus***
DNA stretches up to 30-40Kbp
(Sanger et al.)

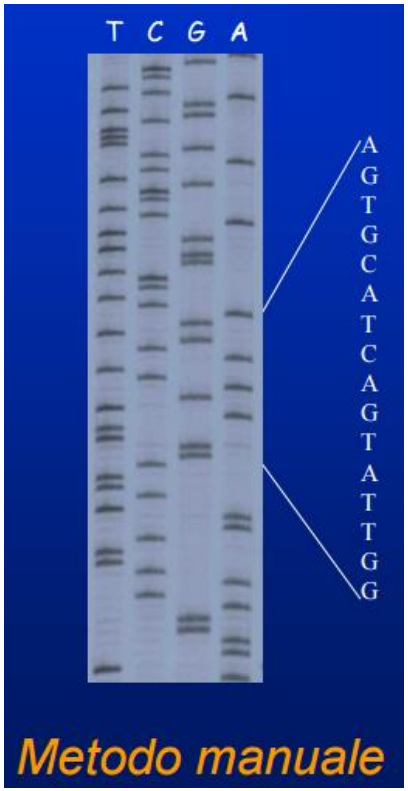


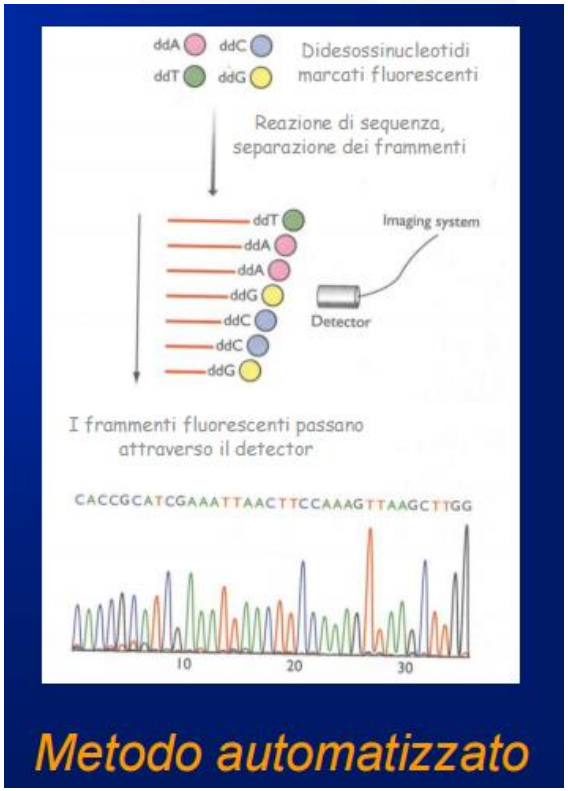
Su cosa si basa il sequenziamento con metodo Sanger?

SEQUENZIAMENTO DEL DNA – Sanger sequencing









Sanger sequencing

- https://www.youtube.com/watch?v=iTBTHmhNNbE

SEQUENZIAMENTO DEL DNA – PROGETTO GENOMA UMANO

1990 - 2003

OBIETTIVI:

-Descrivere completamente il genoma umano mediante il sequenziamento

- Creazione di accurate mappe fisiche dei cromosomi umani

- Sviluppo di tecnologie di supporto

- Creazione di Banche dati per archiviare il dato genetico

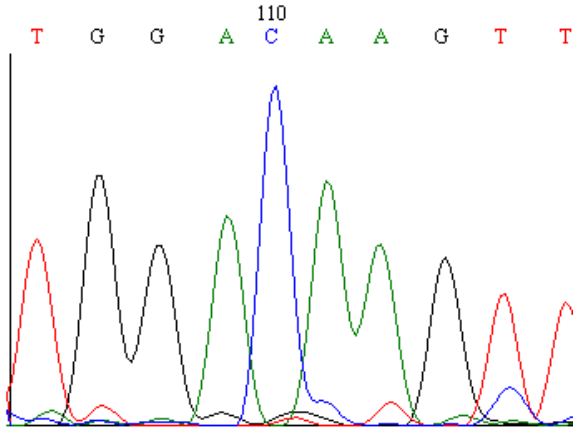
- Identificazione di alterazioni nella sequenza di DNA determinanti lo sviluppo di patologie umane – geni malattia

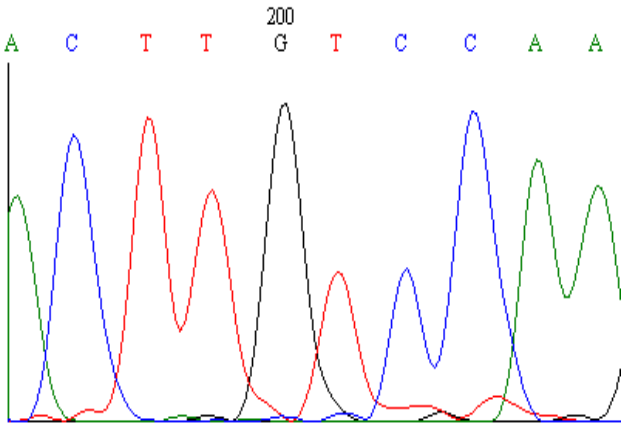
- Comprendere le basi genetiche dell'evoluzione e del funzionamento dell'organismo umano.

MUTAZIONE PUNTIFORME ETEROZIGOTE

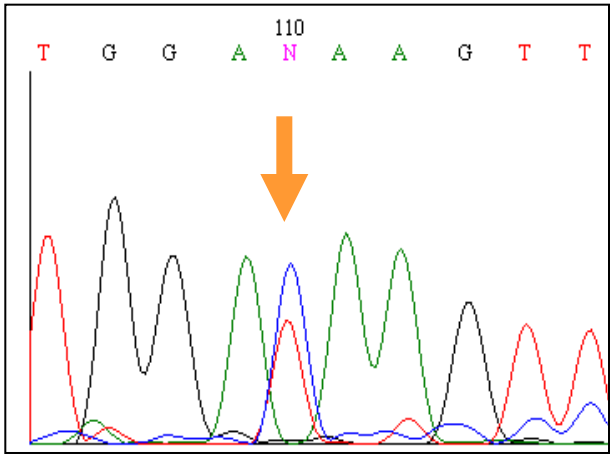
FORWARD **REVERSE**

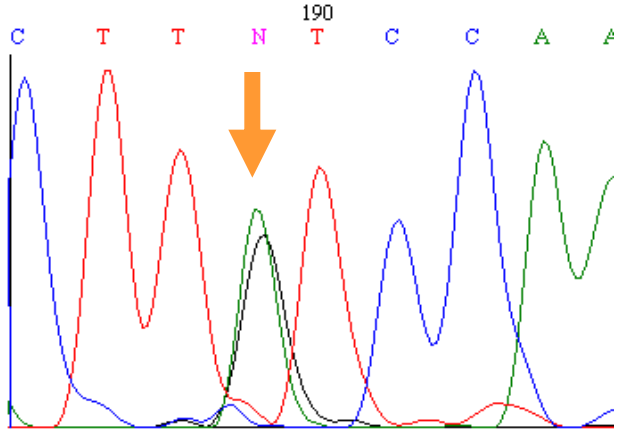
Wild-type





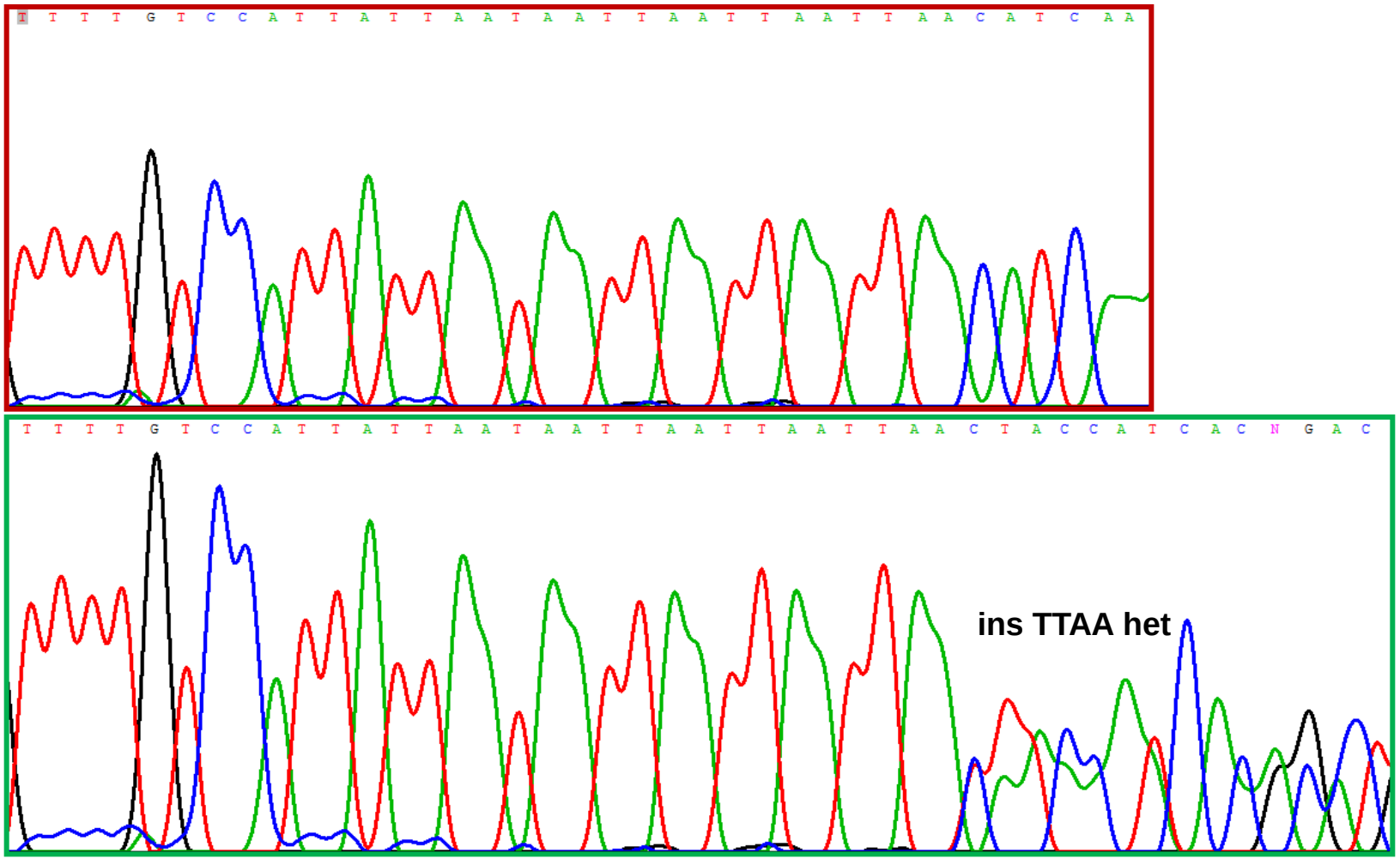
Mutato





SOSTITUZIONE C->T
PRESENZA DI UN DOPPIO PICCO

MUTAZIONE FRAMESHIFT



Diagnosi di paternita

- Metodo

Analisi di VNTR nel nucleo familiare (sequencing)

- Precisione 99,9 %

- **Genetica forense**

D6S434 (202-246)

SEQUENZIAMENTO DEL DNA – PROGETTO GENOMA UMANO

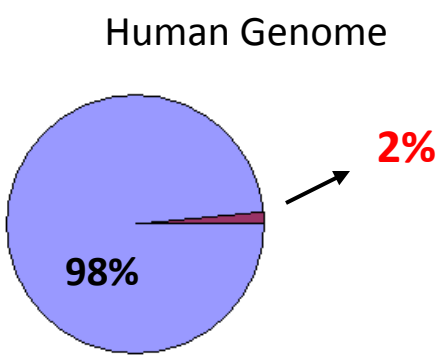
RISULTATI:

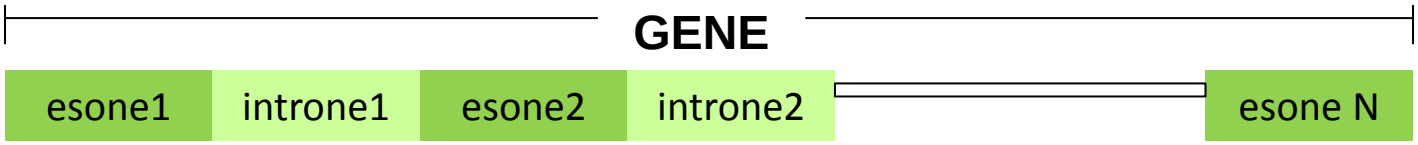
La grandezza totale del genoma umano è di 3.070.000.000 basi di cui 2.843.000.000 sono di eucromatina

Il numero di geni stimato scende da 50.000-100.000 a circa 30.000

Il 60% delle proteine umane presenta similarità di sequenza con proteine presenti in altre specie

Il gene umano medio	
Lunghezza	**100.000 coppie di basi (100 kb)**
Numero di esoni	**7**
Lunghezza dell'esone	**100-200 basi**
Lunghezza dell'introne	**200-2.000 basi**





SEQUENZIAMENTO DEL DNA – PROGETTO GENOMA UMANO

Search NCBI databases Help

Col1A1 Search

Results found in 34 databases for "Col1A1"

Literature

Books	80	books and reports
MeS	4	ontology used for PubMed indexing
NLM Catalog	2	books, journals and more in the NLM Collections
PubMed	1,736	scientific & medical abstracts/citations
PubMed Central	4,720	full-text journal articles

Books Results Page

Health

ClinVar	145	human variations of clinical significance
dbGaP	3	genotype/phenotype interaction studies
GTR	142	genetic testing registry
MedGen	23	medical genetics literature and links
OMIM	89	online mendelian inheritance in man
PubMed Health	3	clinical effectiveness, disease and drug reports

Genomes

Assembly	0	genome assembly information
BioProject	21	biological projects providing data to NCBI
BioSample	11	descriptions of biological source materials
Clone	1,977	genomic and cDNA clones
dbVar	88	genome structural variation studies
Epigenomics	0	epigenomic studies and display tools
Genome	1	genome sequencing projects by organism
GSS	3	genome survey sequences
Nucleotide	614	DNA and RNA sequences
Probe	707	sequence-based probes and primers
SNP	7,558	short genetic variations
SRA	11	high-throughput DNA and RNA sequence read archive
Taxonomy	0	taxonomic classification and nomenclature catalog

Genes

EST	59	expressed sequence tag sequences
Gene	253	collected information about gene loci
GEO DataSets	133	functional genomics studies
GEO Profiles	8,502	gene expression and molecular abundance profiles
HomoloGene	2	homologous gene sets for selected organisms
PopSet	3	sequence sets from phylogenetic and population studies
UniGene	17	clusters of expressed transcripts

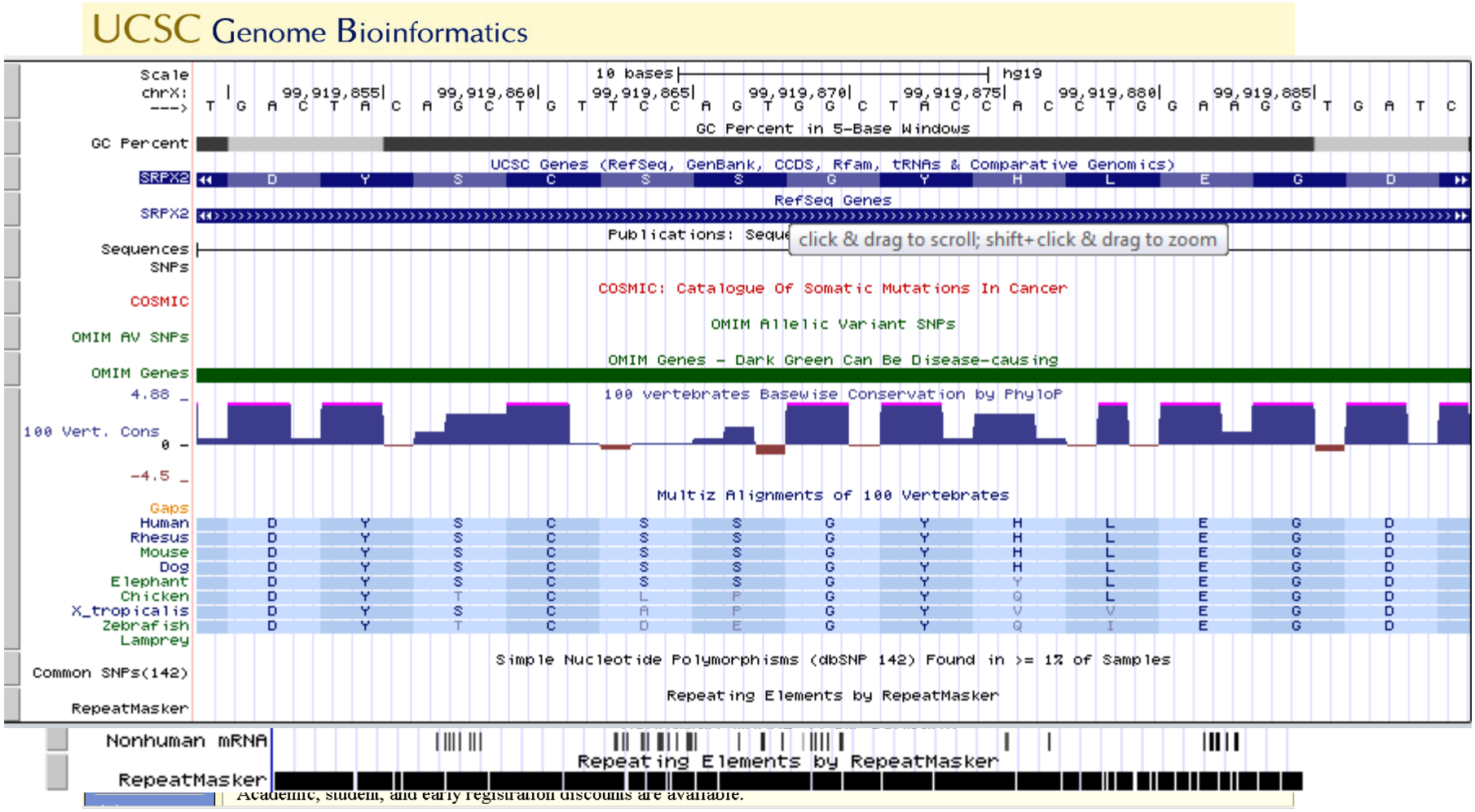
Proteins

Conserved Domains	2	conserved protein domains
Protein	324	protein sequences
Protein Clusters	0	sequence similarity-based protein clusters
Structure	5	experimentally-determined biomolecular structures

Chemicals

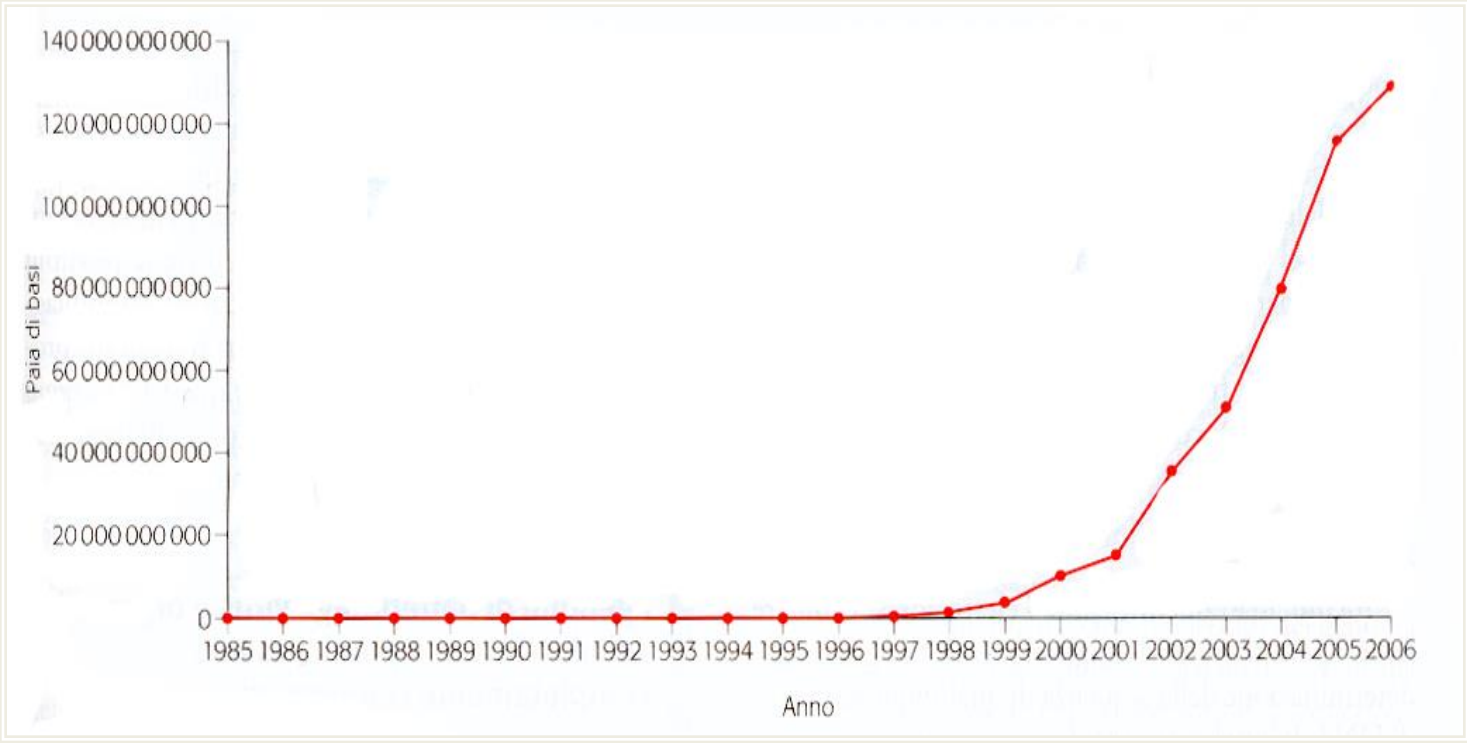
BioSystems	562	molecular pathways with links to genes, proteins and chemicals
PubChem BioAssay	17	bioactivity screening studies
PubChem Compound	0	chemical information with structures, information and links
PubChem Substance	85	deposited substance and chemical information

SEQUENZIAMENTO DEL DNA – PROGETTO GENOMA UMANO



SEQUENZIAMENTO DEL DNA – PROGETTO GENOMA UMANO

Aumento delle sequenze depositate (GenBank, EBi, DBJL) dal 1985 al 2006



POST GENOMICA

Quali le applicazioni del progetto genoma?

APPLICAZIONI PROGETTO GENOMA UMANO

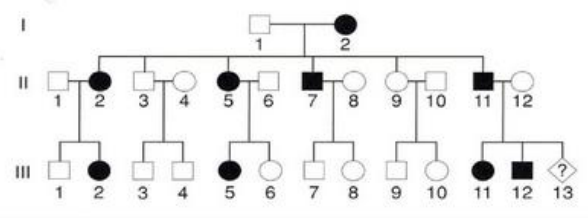
- ✓ Ricerca GENI-MALATTIA in patologie Mendeliane
- ✓ Ricerca di GENI e MUTAZIONI PREDISPONENTI in patologie Multifattoriali e Oncologiche

MALATTIA GENETICA	**MENDELIANA o MONOGENICA** (es. Fibrosi Cistica, Emofilia)	- malattia dovuta alla mutazione di un singolo gene - alto rischio di ricorrenza familiare - ca 7000 note (2000 con difetto biochimico noto)
	MULTIFATTORIALE (es. Diabete, Ipertensione, Problemi cardiovascolari)	- malattia dovuta al coinvolgimento di più geni, molto frequenti - interazione con l'ambiente - basi molecolari e meccanismo patogenetico spesso sconosciuti - predisposizione a sviluppare determinate patologie
	DELLE CELLULE SOMATICHE (es. Vari tipi di tumore)	- a carico generalmente di un unico tipo cellulare - generalmente dovute ad una mutazione ereditata (PREDISPOSIZIONE) + una seconda mutazione nelle cellule somatiche

APPLICAZIONI PROGETTO GENOMA UMANO

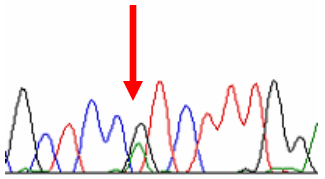
Identificare GENI-MALATTIA

MALATTIA GENETICA



MAPPATURA

IDENTIFICAZIONE DEL GENE
e SCREENING MUTAZIONALE



DIAGNOSI

PREVENZIONE

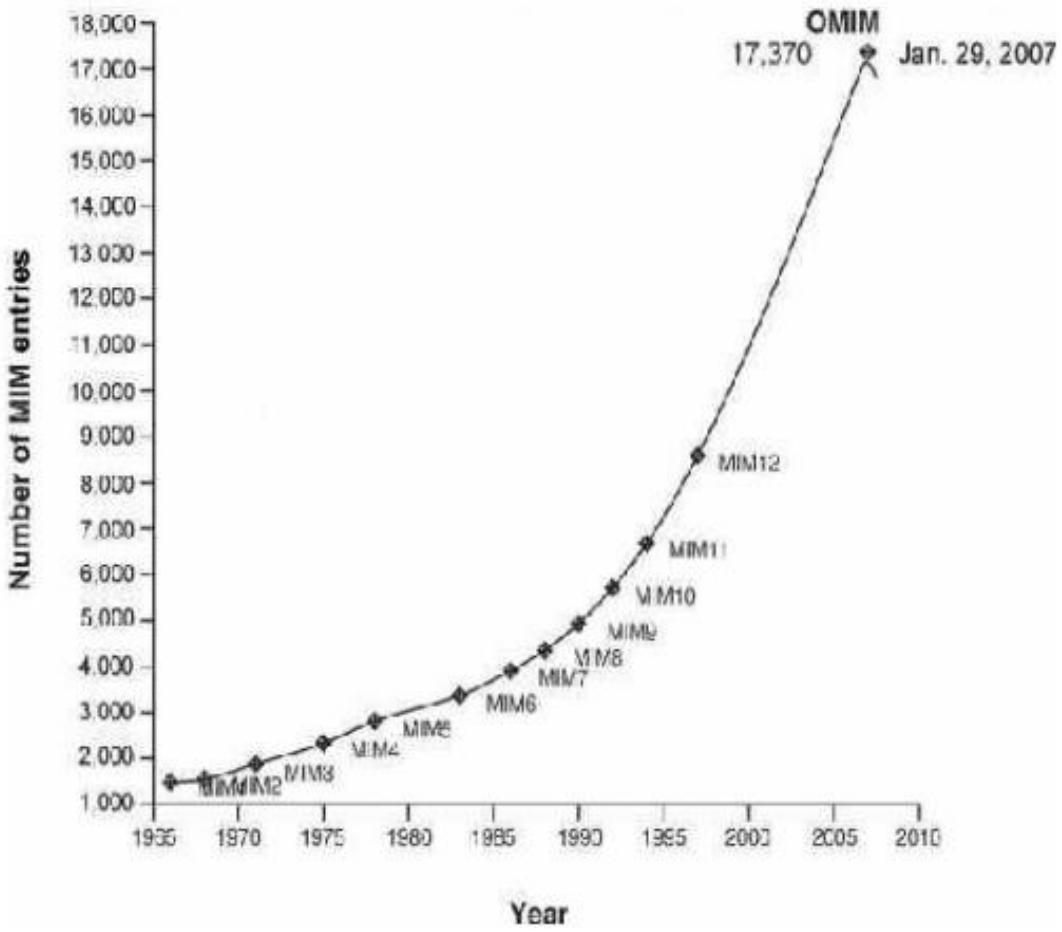
TERAPIA GENICA

DIFETTO BIOLOGICO

TERAPIA

APPLICAZIONI PROGETTO GENOMA UMANO

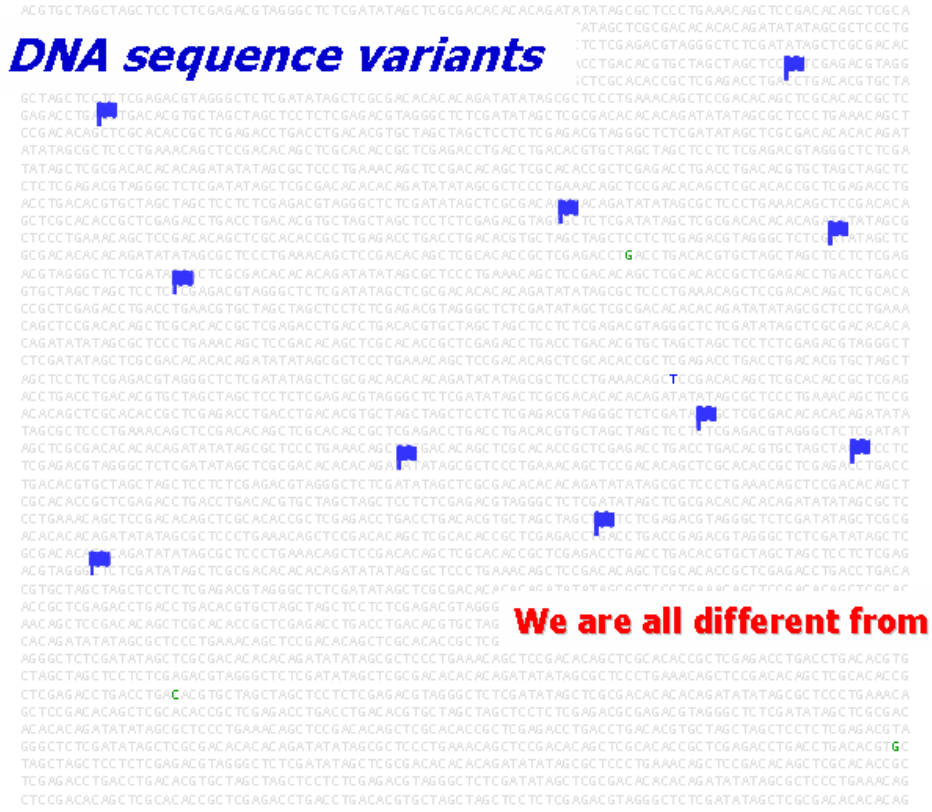
OMIM® **Online Mendelian Inheritance in Man®**
An Online Catalog of Human Genes and Genetic Disorders
Updated 16 October 2015



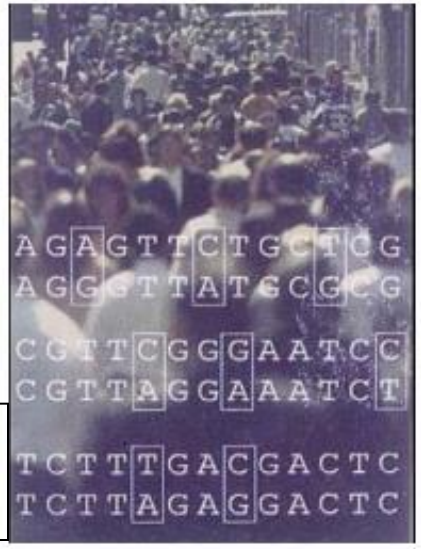
SEQUENZIAMENTO DEL DNA – PROGETTO GENOMA UMANO

Identificare geni e mutazioni che causano <u>PREDISPOSIZIONI</u> a malattie ad alta frequenza nella popolazione (es. diabete, alcune forme tumorali, patologie cardiovascolari)

Due individui qualunque condividono il 99.9% delle sequenze di DNA



Il rimanente 0.1% determina l'unicità dell'individuo



Ogni persona ha un profilo specifico di queste basi che la rende diversa da ogni altra

POST GENOMICA

0.1% delle differenze tra individui

Consistono in variazioni frequenti a livello delle singole basi (conversione di una base in un'altra, delezione, inserzione)

- **SNP** (Single Nucleotide Polymorphisms) -

VARIAZIONI SENZA EFFETTI

Individuo SANO

VARIAZIONI INNOCUE

(es. legate all'aspetto esteriore, alla capacità di arrotolare la lingua, ecc…)

TENDENZA A SVILUPPARE MALATTIE

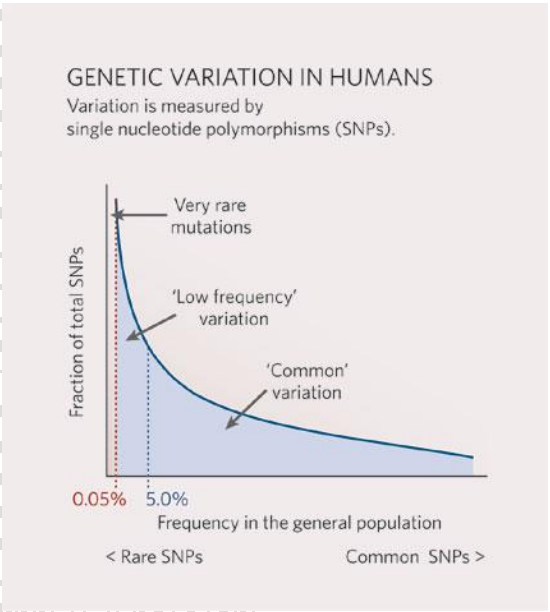
Individuo SANO ma con una proteina con funzionamento alterato

POST GENOMICA

Predisposing

Pathogenic

Neutral



Frequenza SNPs: 1/1000basi

Come vengono utilizzati gli SNPs in ambito medico?

✓ Alcuni polimorfismi genetici relativamente comuni, se associati tra loro e combinati con specifiche componenti ambientali, possono elevare notevolmente il rischio di sviluppare patologie diffuse → GLI SNPs NON SONO CAUSA DI MALATTIE MA POSSONO AIUTARE AD INDIVIDUARE LA PROPENSIONE INDIVIDUALE A CONTRARLE

✓ Alcune patologie hanno un decorso molto diverso in individui portatori di una stessa mutazione causativa; tale fenomeno sembra essere in parte causato dalla presenza di polimorfismi differenti

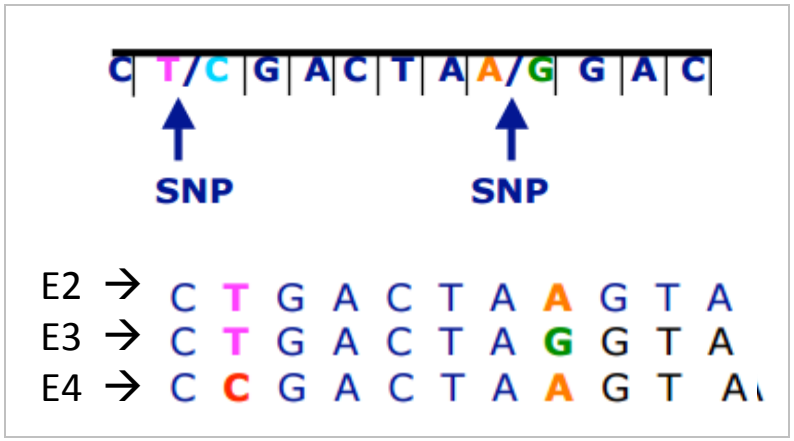
Basandosi sulle informazioni ricavabili dalla costituzione genetica di un individuo è possibile effettuare una **stima del rischio** di sviluppare una determinata patologia durante il corso della vita

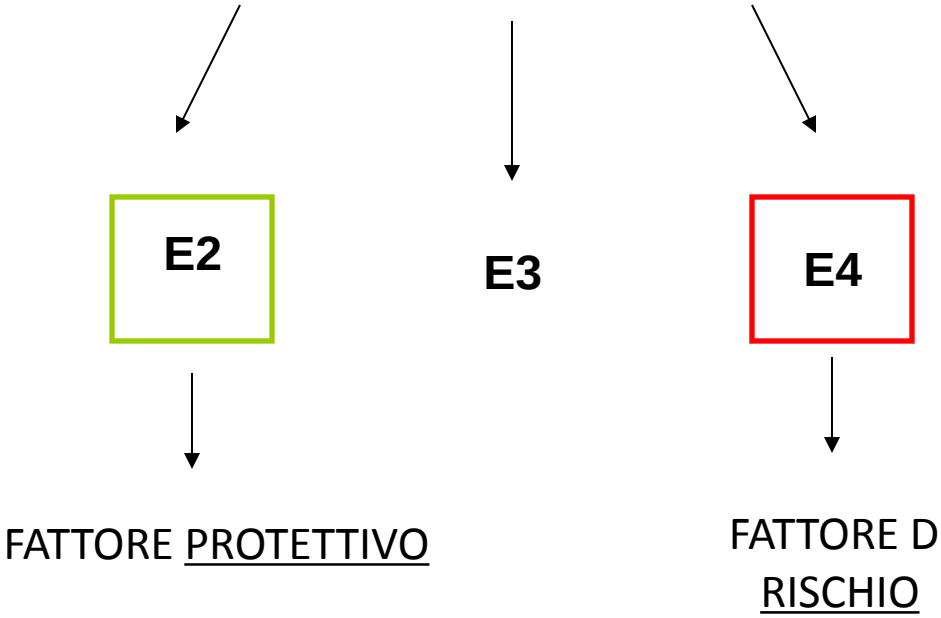
MEDICINA PREDITTIVA- PREVENTIVA

(possibilità di intervenire precocemente, ottimizzazione dei trattamenti, limitazione dei problemi clinici)

STUDI DI ASSOCIAZIONE GENOME-WIDE

ApoE – Apolipoproteina E- → Nella sequenza di questo gene sono presenti 2 SNPs la cui combinazione da origine a 3 diverse possibili sequenze geniche





Alzheimer disease: **APOE***ε4-associated increase in AD risk linked to phospholipid dysregulation.

Malkki H.

Nat Rev Neurol. 2015 Sep 29. doi: 10.1038/nrneurol.2015.180. [Epub ahead of print] No abstract available.

Come vengono utilizzati gli SNPs in ambito medico?

La risposta di singoli individui ai farmaci è spesso molto diversa → ***come si può spiegare?***

Cause Of Death	Number Of Deaths
Heart Disease	652,091
Malignant Neoplasms	559,312
Cerebrovascular Disease	143,579
Lower Respitory Disease	130,933
Accidents	117,809
Adverse Drug Reactions	**106,000**

RISPOSTA DESIDERATA

RISPOSTA DI UN INDIVIDUO AD UN FARMACO

REAZIONE AVVERSA AL FARMACO

(effetti collaterali)

INTERAZIONE TRA FARMACI

VARIABILITA' INDIVIDUALE SU BASE GENETICA

alterazioni di

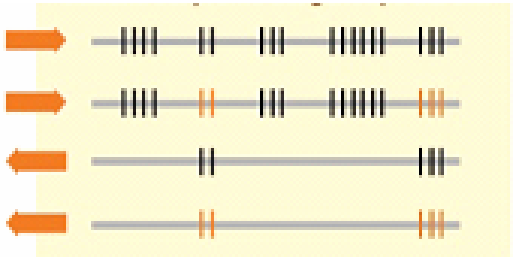
- BERSAGLIO TERAPEUTICO
- METABOLISMO DEL FARMACO
- ASSORBIMENTO, DISTRIBUZIONE, ELIMINAZIONE

POST GENOMICA

SNPs→ Individui diversi possono avere una base diversa in una specifica posizione

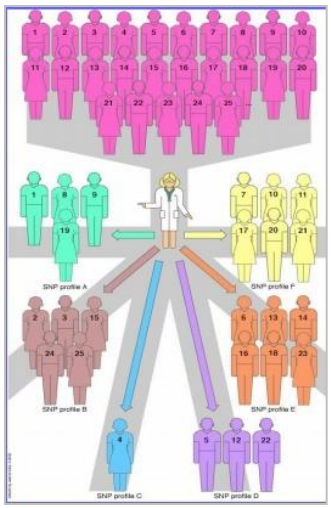
MAPPATURA DEGLI SNPs Posizione degli SNPs nel DNA umano

PROFILO GENOTIPICO SNPs



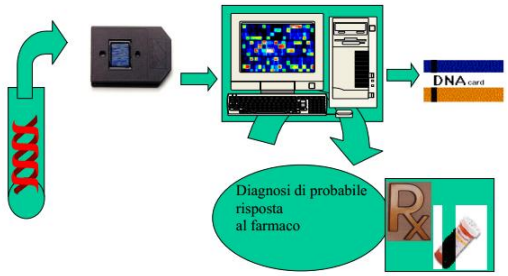
Roses AD Nature 2000, 405:827-865

Raggruppamento degli individui in classi in base al profilo di SNPs



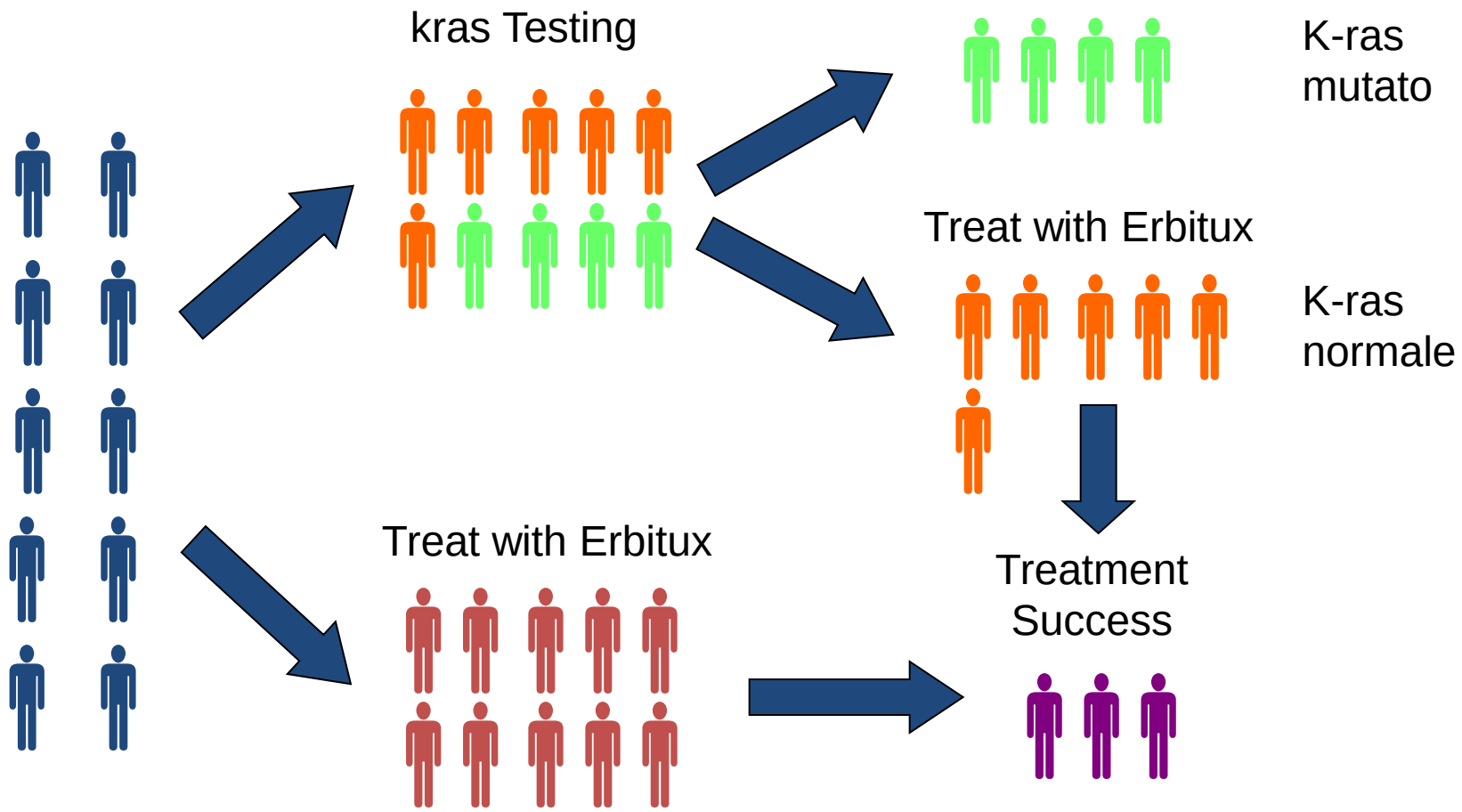
FARMACOGENOMICA

(MEDICINA PERSONALIZZATA)



SNPs microarray → utilizzando un unico chip e un prelievo ematico del paziente si possono analizzare più di 100.000 SNPs in poche ore

Personalized Medicine Reduces Ineffective Treatment in Colon Cancer



Langreth, R. (2008), 'Imclone's Gene Test Battle', Forbes.com, 16May

POST GENOMICA

MEDICINA TRADIZIONALE → MEDICINA GENETICA

DIFFERENTE APPROCCIO DIAGNOSTICO-TERAPEUTICO

MEDICINA TRADIZIONALE

- Epidemiologica
- Diagnosi e terapie uguali per tutti
- **Feno**diagnostica
- **Feno**terapeutica
- Prognostica e Preventiva

MEDICINA GENETICA

- Eziologica
- **Geno**diagnostica
- **Geno**terapeutica
- Predittiva
- Personalizzata

IPOTESI BIOLOGICA

↓

RICERCA DELLE CAUSE

INDAGINE GENETICA

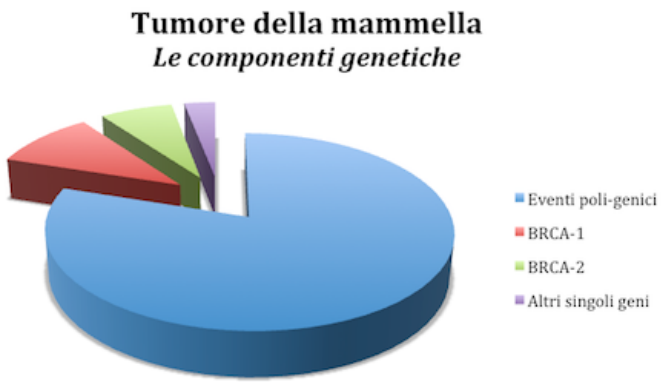
↓

CAUSA BIOLOGICA

TEST GENETICO

Analisi di specifici geni, del loro prodotto o della loro funzione, nonché ogni altro tipo di indagine del DNA o dei cromosomi, finalizzate ad individuare o ad escludere modificazioni del DNA verosimilmente associate a patologie genetiche.

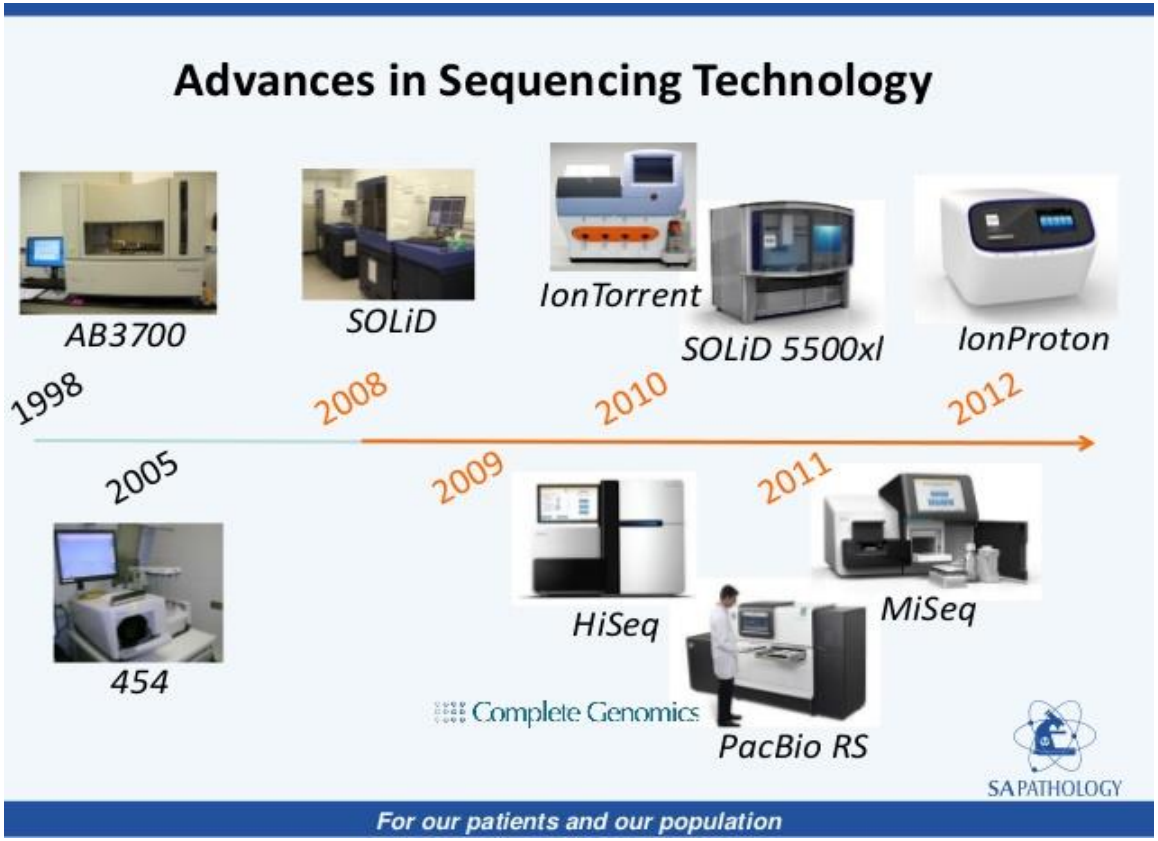
- ✓ test **DIAGNOSTICO** (conferma della clinica)
- ✓ test **PRESINTOMATICO** (per malattie ad esordio tardivo)
- ✓ test **PREDITTIVI** per suscettibilità ad una malattia (es. tumori famigliari del colon o della mammella)
- ✓ test **FARMACOGENETICO** (per la risposta individuale ai farmaci)
- ✓ test **MEDICO-LEGALI** (es. accertamento esclusione paternità)



SEQUENZIAMENTO DEL DNA DI NUOVA GENERAZIONE

NGS – Next Generation Sequencing

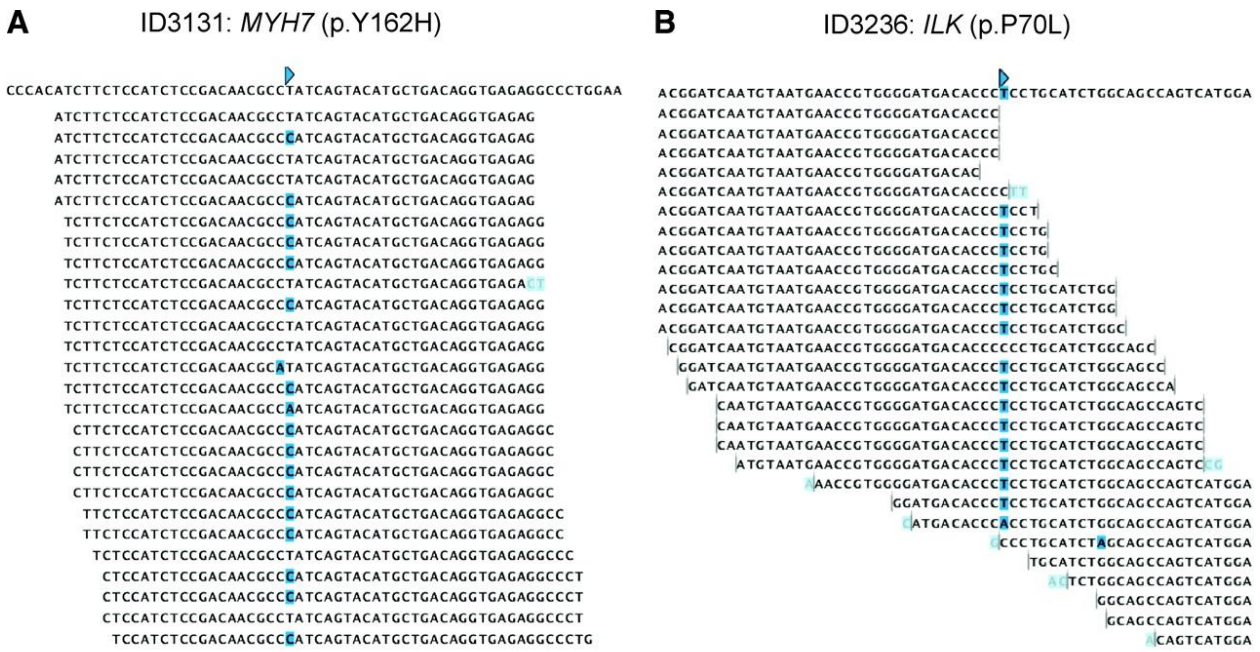
Si basa sul principio del sequenziamento di *cluster* clonali: milioni di analisi di sequenziamento di DNA contemporaneamente sullo stesso campione e su più campioni diversi

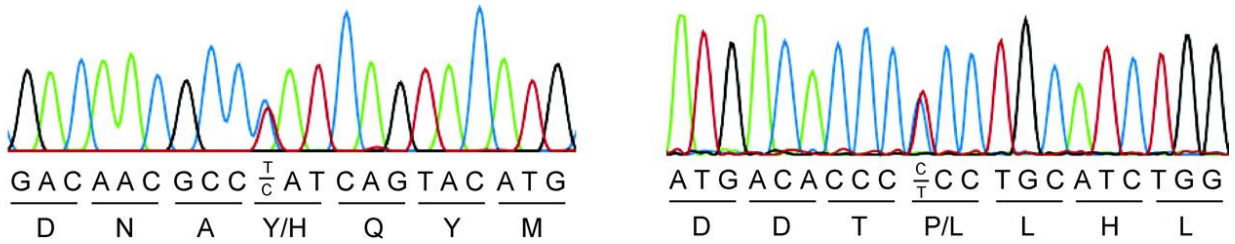


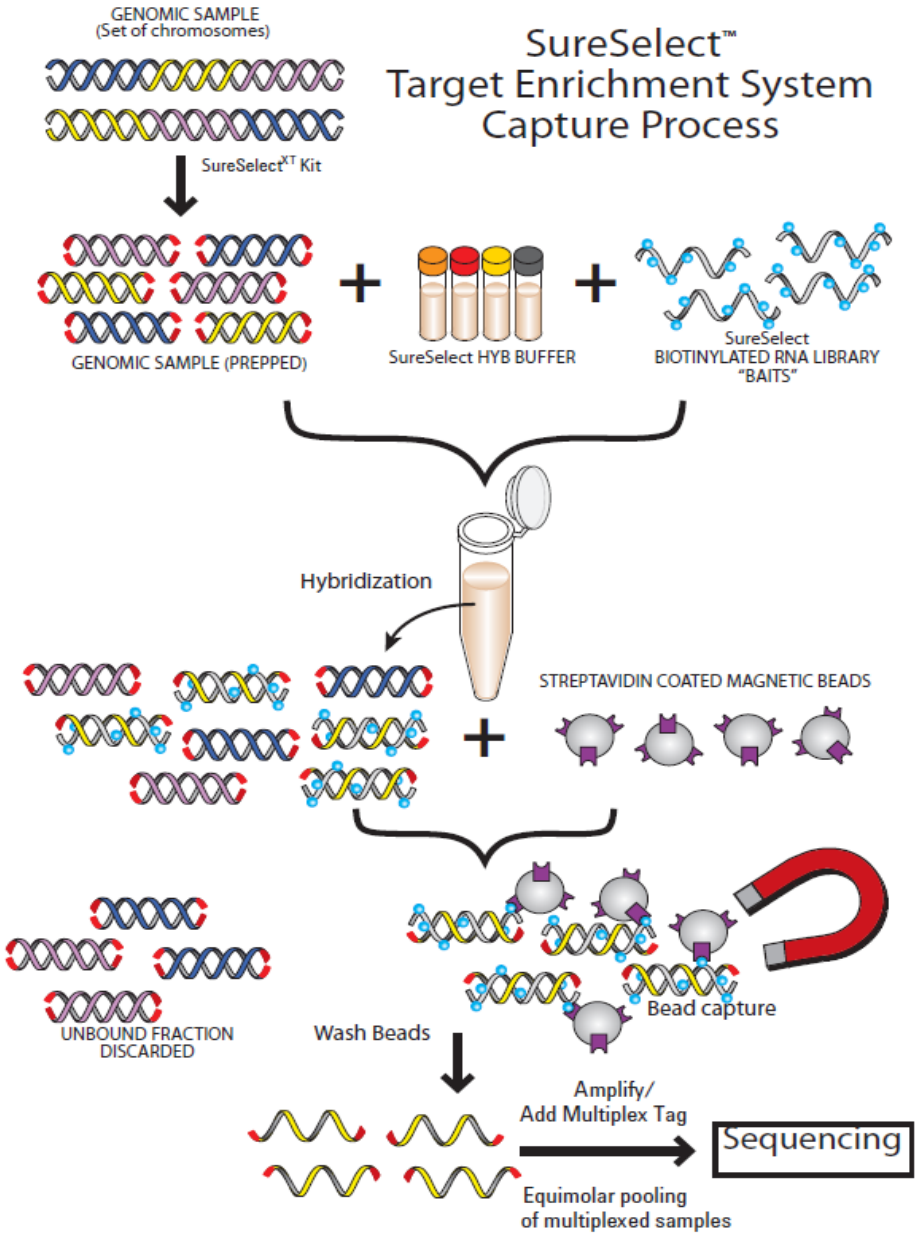
SEQUENZIAMENTO DEL DNA

NGS

Sanger







ARRICCHIMENTO GENICO (ENRICHMENT)

CATTURA GENICA (CAPTURING)

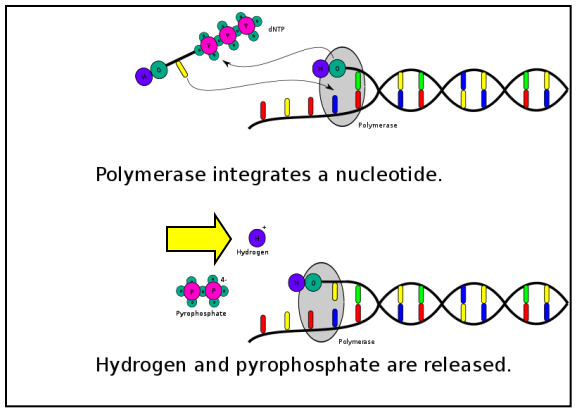
http://www.nbt.nhs.uk/genetics

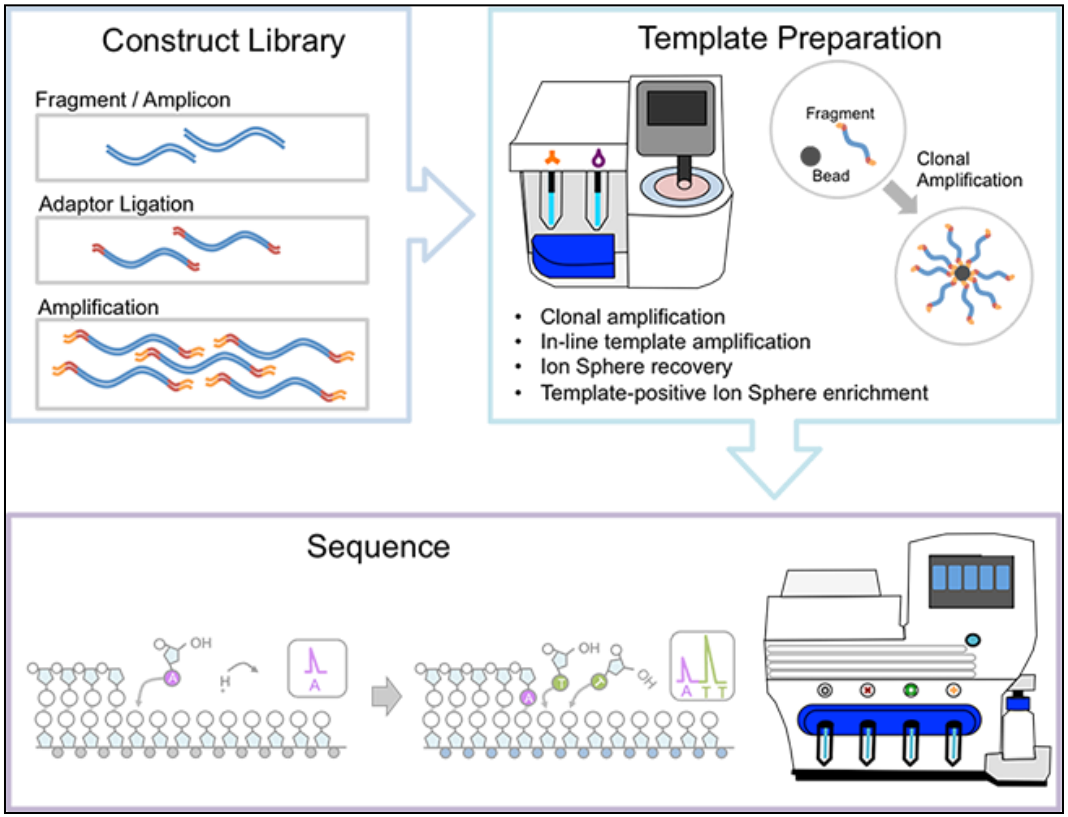
SEQUENZIAMENTO DEL DNA DI NUOVA GENERAZIONE

ION TORRENT

Si basa sul rilevamento di ioni idrogeno liberati in caso di incorporazione di una base

Workflow : Library preparation → Emulsion PCR → Semiconductor Sequencing

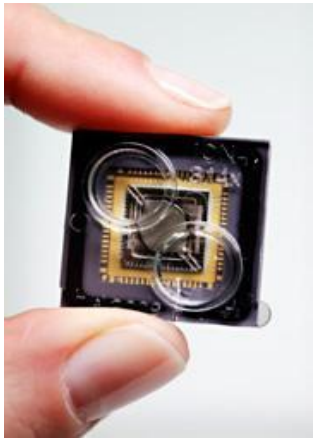


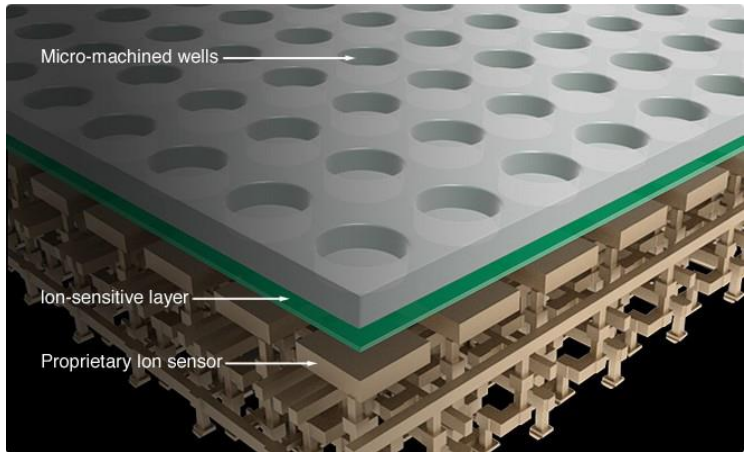


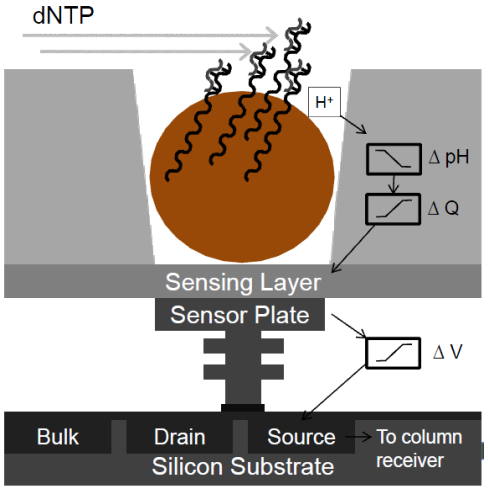
SEQUENZIAMENTO DEL DNA DI NUOVA GENERAZIONE

ION TORRENT

Dove avviene la reazione di sequenziamento?





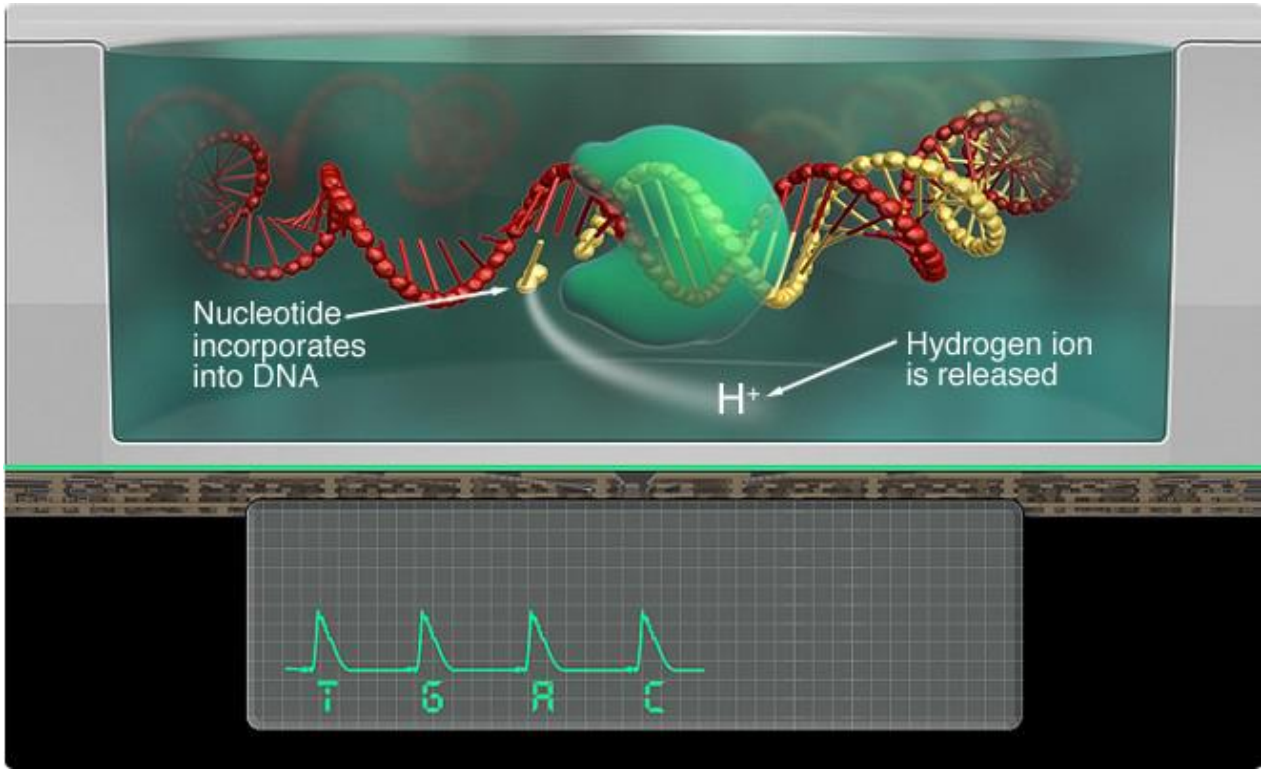


SEQUENZIAMENTO DEL DNA DI NUOVA GENERAZIONE

ION TORRENT

4 nucleotides flow sequentially C G T A

C

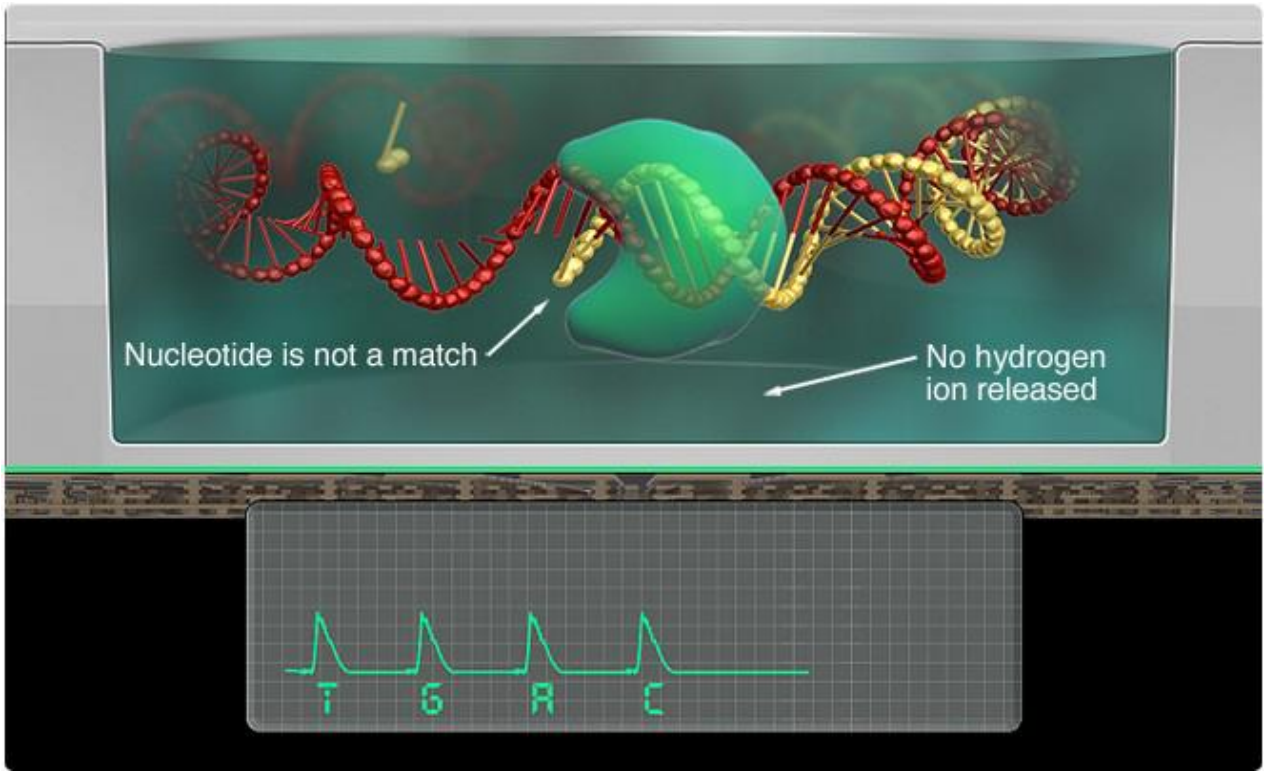


No camera, just a pH sensor

SEQUENZIAMENTO DEL DNA DI NUOVA GENERAZIONE

ION TORRENT

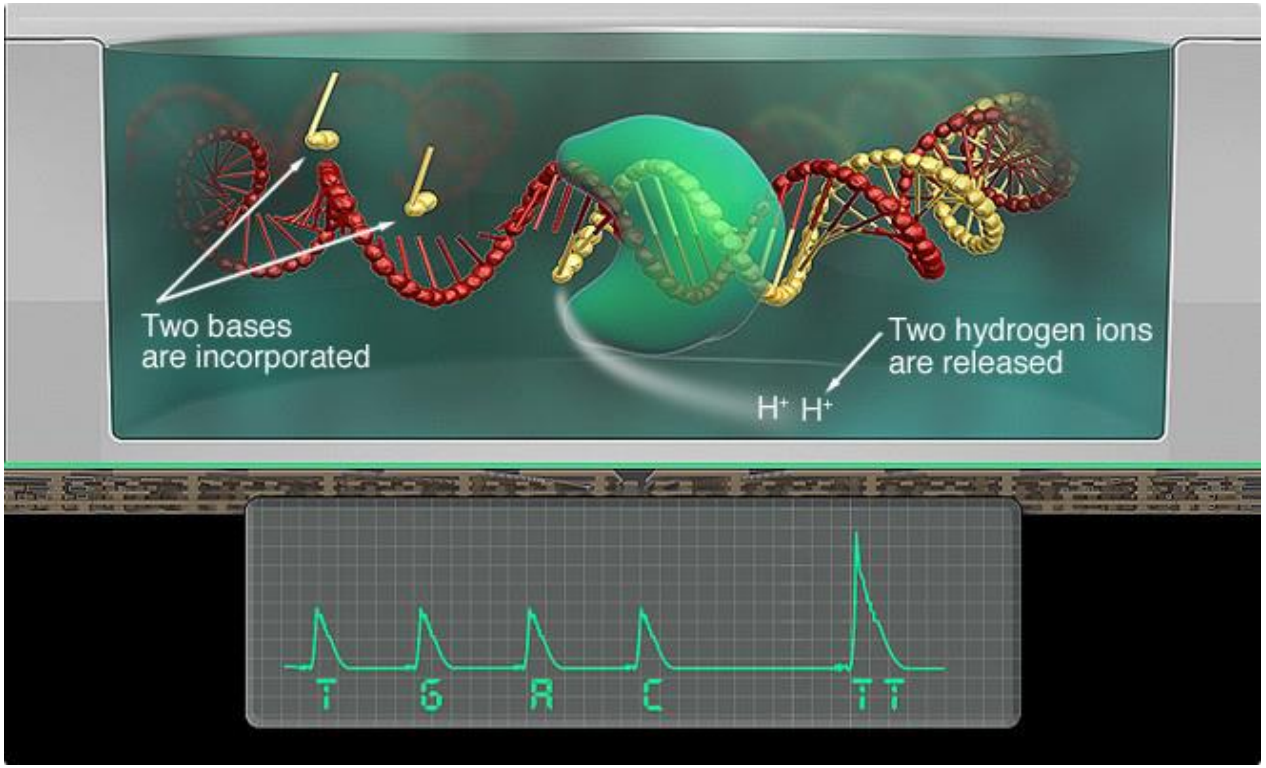
G



SEQUENZIAMENTO DEL DNA DI NUOVA GENERAZIONE

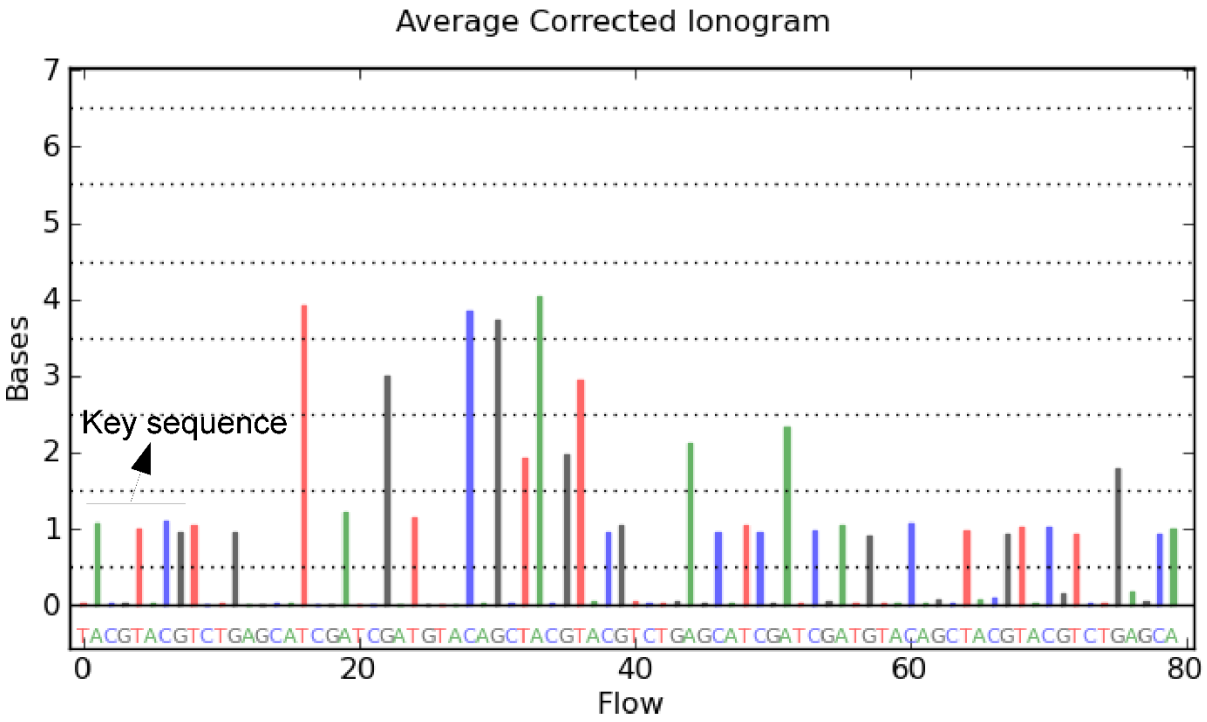
ION TORRENT

T



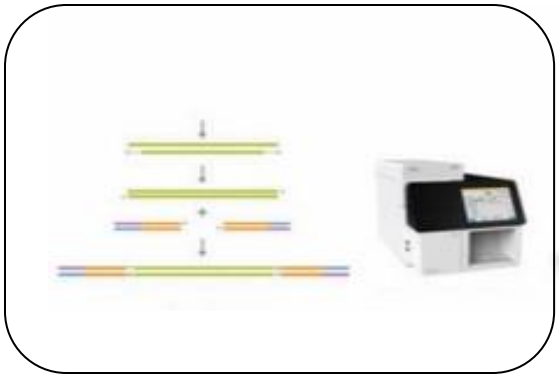
SEQUENZIAMENTO DEL DNA DI NUOVA GENERAZIONE

ION TORRENT



ATCGTGTTTTAGGGTCCCCGGGGTT...

SEQUENZIAMENTO DEL DNA DI NUOVA GENERAZIONE

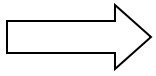


Technology	Sanger Sequencing	Next Generation Sequencing			
Manufacturer	Applied Biosystems	Roche 454	Illumina	Life Technologies	
Model	ABI 3730XL	GS FLX Titanium XL+	HiSeq 2000 dual flow cell	SOLiD 4 System	Ion PGM
Bases per RUN	~ 96 Kb	700 Mb	600 Gb	100 Gb	1 Gb
Time per RUN	2 h	~1 day	~11 days	~14 days	4.5 h
Reads per RUN	96	1 million	6 billions (paired-end)	1.4 billions	5 millions
Reads length	up to 1000 bp	up to 1000 bp (mode 700 bp)	2*100 bp	2*50 bp	up to 400 bp

Produzione di una mole enorme di dati:

-Da 1 Gb a diversi Tb di dati grezzi prodotti per corsa

- il processamento delle sequenze tramite *pipeline* informatiche richiede molta capacità di calcolo (CPU)



INFRASTRUTTURE INFORMATICHE ADEGUATE

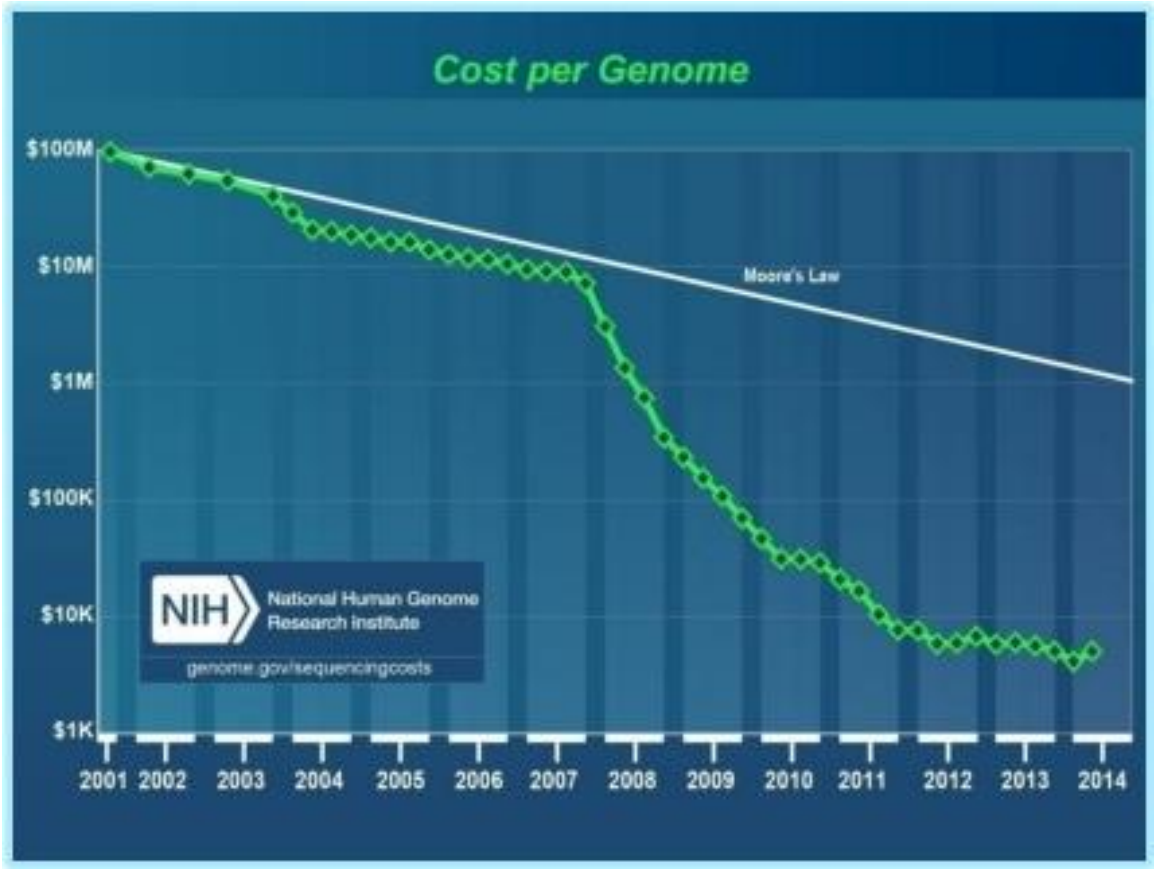
VIDEO

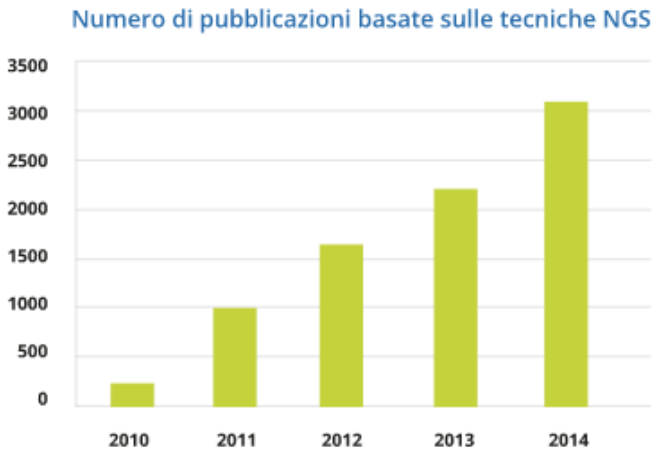
- https://www.youtube.com/watch?annotation_id=annotation_228575861&feature=iv&src_vid=womKfikWlxM&v=fCd6B5HRaZ8

POST GENOMICA

2001: *Human Genome Project*

2.7G$, 11 years





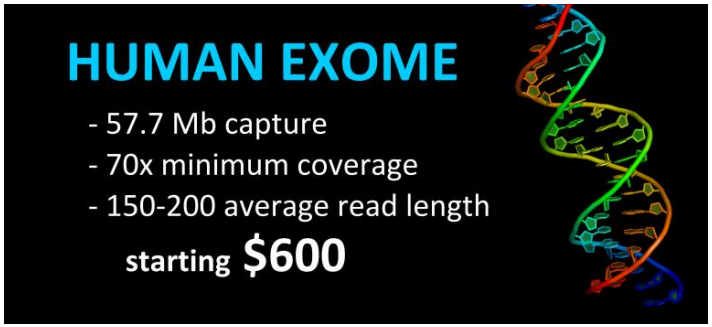
Stiamo andando verso la capacità di sequenziare il genoma di un individuo in 48h con una spesa di circa 1000 euro

NGS technology approaches

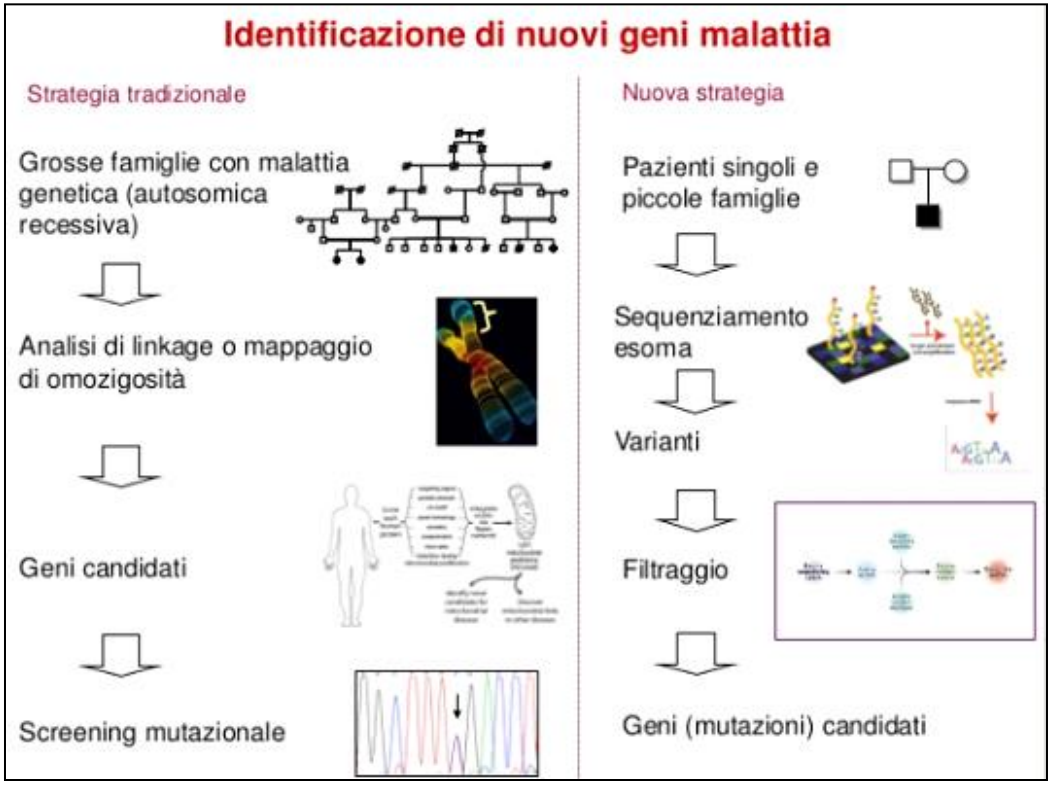
Target	Bases in the target	Median coverage	Bases to be sequenced	Expected variants
GS	3.100.000.000	30x	>120Gb	3.000.000
ES	50.000.000	100x	10Gb	30.000
Large panel	1.500.000	200x	1Gb	1.000
Small panel	50.000	300x	0.05Gb	30

POST GENOMICA

WHOLE EXOME SEQUENCING (**WES**) → sequenziamento di tutti i tratti del genoma che codificano proteine (< 2% del genoma di un individuo); si stima che l'esoma debba contenere circa l'**85%** delle mutazioni associate a patologia.



NUOVO APPROCCIO PER L'INDIVIDUAZIONE DI NUOVI GENI-MALATTIA



✓ I primi successi di questa tecnica risalgono al 2009 quando sono stati diagnosticati 3 pazienti affetti da malattie rare

✓ Identificazione del gene-malattia nel 25% dei casi analizzati

Implementazione TEST GENETICO (in termini di sensibilità, riduzione dei costi e dei tempi di analisi) per **TEST GENETICI MALATTIA-SPECIFICI**

Applicazioni:

- ✓ Diagnosi molecolare in malattie associate a mutazioni in **geni di grandi dimensioni** non approcciabili mediante sequenziamento Sanger

- ✓ Diagnosi molecolare per malattie a **ereditarietà multigenica**: il target resequencing è destinato a diventare una pratica standard dato l'alto beneficio in termini di costo-detection rate della tecnologia NGS (es. cardiomipatia, ritardo mentale X-linked, retinite pigmentosa):

  - Pannelli che includono numerosi geni, selezionati in base alla correlazione con la patologia in esame. Approccio indicato per lo studio di malattie ad elevata eterogeneità genetica

Osteogenesis Imperfecta

Coverage ~97%

- COL1A1
- COL1A2
- IFITM5
- SERPINF1
- SERPINH1
- BMP1
- CRTAP
- FKBP10
- PLOD2
- PPIB
- LEPRE1
- SP7
- ATP6V0A2
- CREB3L1
- PLS3
- LRP5
- LRP6
- SMPD3
- SERPINB5
- BMP1
- TMEM38B
- WNT1

OSTEOGENESI IMPERFETTA: malattia delle ossa fragili; è una patologia rara genetica caratterizzata da una forte eterogeneità genetica (causata cioè da 16 potenziali geni).

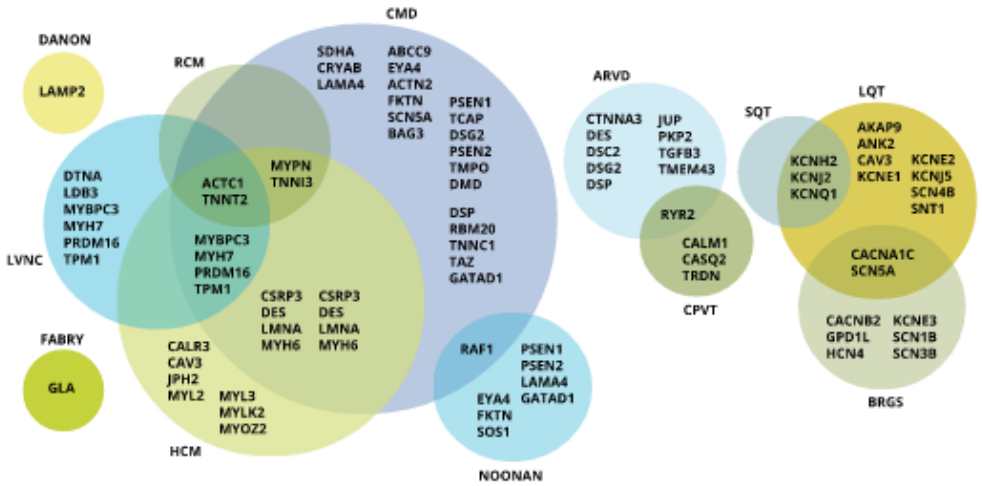
Costi per analisi NGS

OSTEOGENESIS IMPERFECTA (PANEL 2)	• PLOD2 • IFITM5 • CRTAP • LEPRE1 • PPIB • SERPINH1 • FKBP10 • SP7 • SERPINF1 • BMP1 • WNT1 • CREB3L1 • TMEM38B • PLS3	14 genes (Sequencing)	1900

Costi per Sequenziamento Sanger

CRTAP (CARTILAGE-ASSOCIATED PROTEIN)	605497	**OSTEOGENESIS IMPERFECTA, TYPE 7, OI7**	610682		550

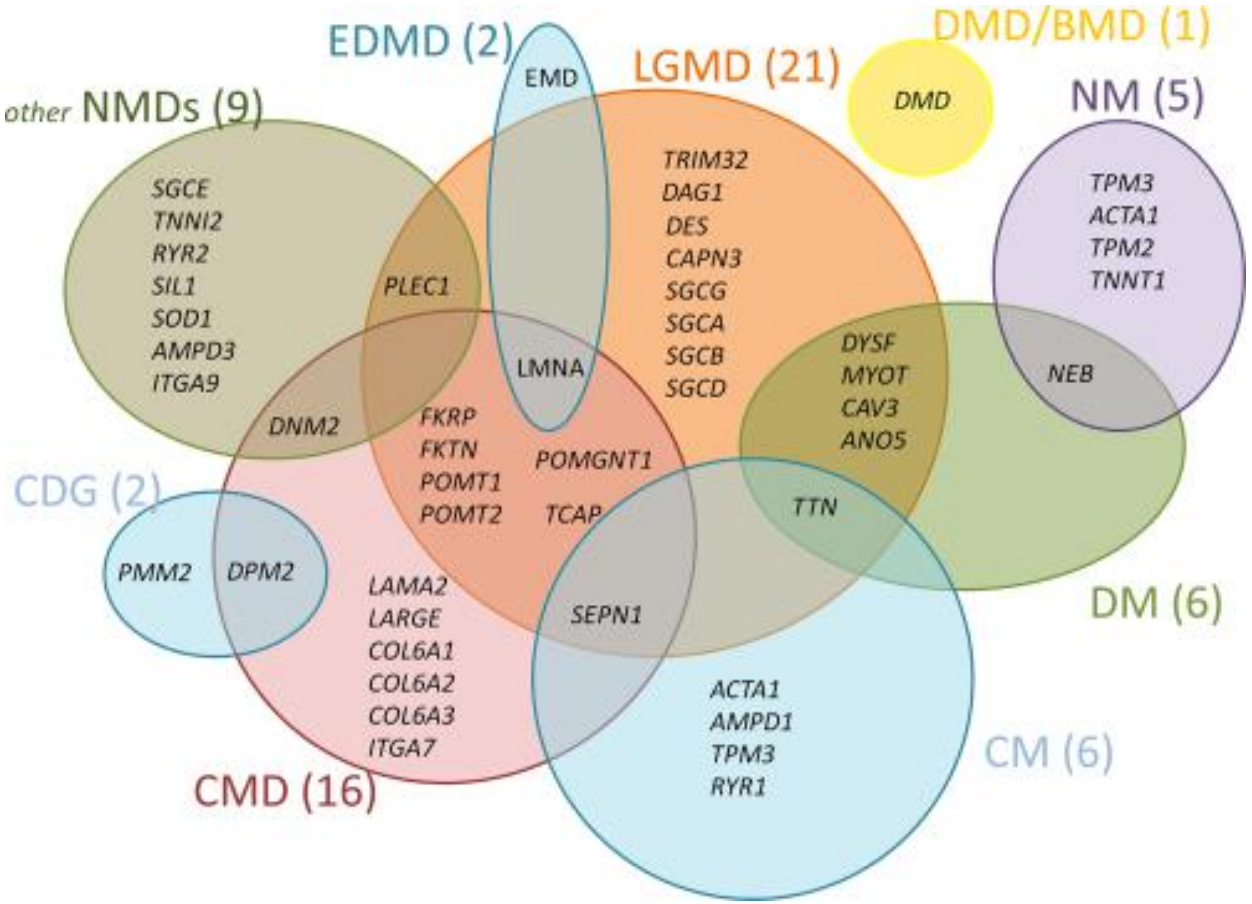
ETEROGENEITA' GENETICA DELLE CARDIOMIOPATIE



TEST	GENI	NOTE
CARDIOMIOPATIE STRUTTURALI (62 geni)	ABCC9 ACTC1 ACTN2 ANKRD1 BAG3 BRAF CALR3 CAV3 CRYAB CSRP3 DES DMD DSC2 DSG2 DSP DTNA EYA4 FKTN GATAD1 GLA JPH2 JUP KRAS LAMA4 LAMP2 LDB3 LMNA MIB1 MYBPC3 MYH6 MYH7 MYL2 MYL3 MYOZ2 MYPN NEXN NF1 NNT NRAS PKP2 PLN PRDM16 PRKAG2 PTPN11 RAF1 RBM20 RIT1 RYR2 SCN5A SGCD SOS1 TAZ TCAP TGFB3 TMEM43 TMPO TNNC1 TNNI3 TNNT2 TPM1 TTN VCL	Il pannello comprende geni implicati nel/nelle: cardiomiopatie (CM) dilatative, CM ipertrofiche, CM da non compattazione del ventricolo destro, CM restrittiva ed altre condizioni, tra cui la sindrome di Noonan, la malattia di Fabry e la malattia di Danon
CARDIOMIOPATIA DILATATIVA (35 geni)	ABCC9 ACTC1 ACTN2 ANKRD1 BAG3 CRYAB CSRP3 DES DMD DSG2 DSP EYA4 FKTN GATAD1 LAMA4 LDB3 LMNA MYBPC3 MYH6 MYH7 MYPN NEXN PLN RAF1 RBM20 SCN5A SGCD TCAP TMPO TNNC1 TNNI3 TNNT2 TPM1 TTN VCL	Il pannello rileva circa il 50% dei casi ereditari di cardiomiopatia dilatativa
CARDIOMIOPATIE DA NON-COMPATTAZIONE DEL VENTRICOLO SINISTRO (11 geni)	ACTC1 DTNA LDB3 MIB1 MYBPC3 MYH7 NNT PRDM16 TAZ TNNT2 TPM1	Il pannello rileva circa il 44% dei casi ereditari di cardiopatie da non-compattazione del ventricolo sinistro

NGS technology and NMDs

Ankala et al., A comprehensive genomic approach for neuromuscular diseases gives a high diagnostic yield
(Ann Neurol 2015; 77: 206-214)



Comprehensive NMD Targeted Gene panel

TABLE 1. Genes Included in the Different NGS Panels Offered

DMD	LGMD	CMD	NMD[a]
DMD	*CAPN3*	*COL6A1*	*ACTA1*
	CAV3	*COL6A2*	*AMPD1*
	DYSF	*COL6A3*	*AMPD3*
	LMNA	*FKRP*	*ANO5*
	MYOT	*FKTN*	*DES*
	SGCA	*ITGA7*	*EMD*
	SGCB	*LAMA2*	*NEB*
	SGCD	*LARGE*	*PLEC*
	TCAP	*POMGNT1*	*PMM2*
	SGCG	*POMT1*	*RYR1*
	TRIM32	*POMT2*	*RYR2*
		SEPN1	*SIL1*
			TTN
			TNNI2
			TNNT1
			TPM2
			TPM3

[a]NMD panel includes the genes listed under it, as well as all the other genes in the DMD, LGMD, and CMD lists. CMD = congenital muscular dystrophy; DMD = Duchenne muscular dystrophy; LGMD = limb-girdle muscular dystrophy; NMD = neuromuscular disease.

Validation on known variants: analytical sensitivity and specificity 100%

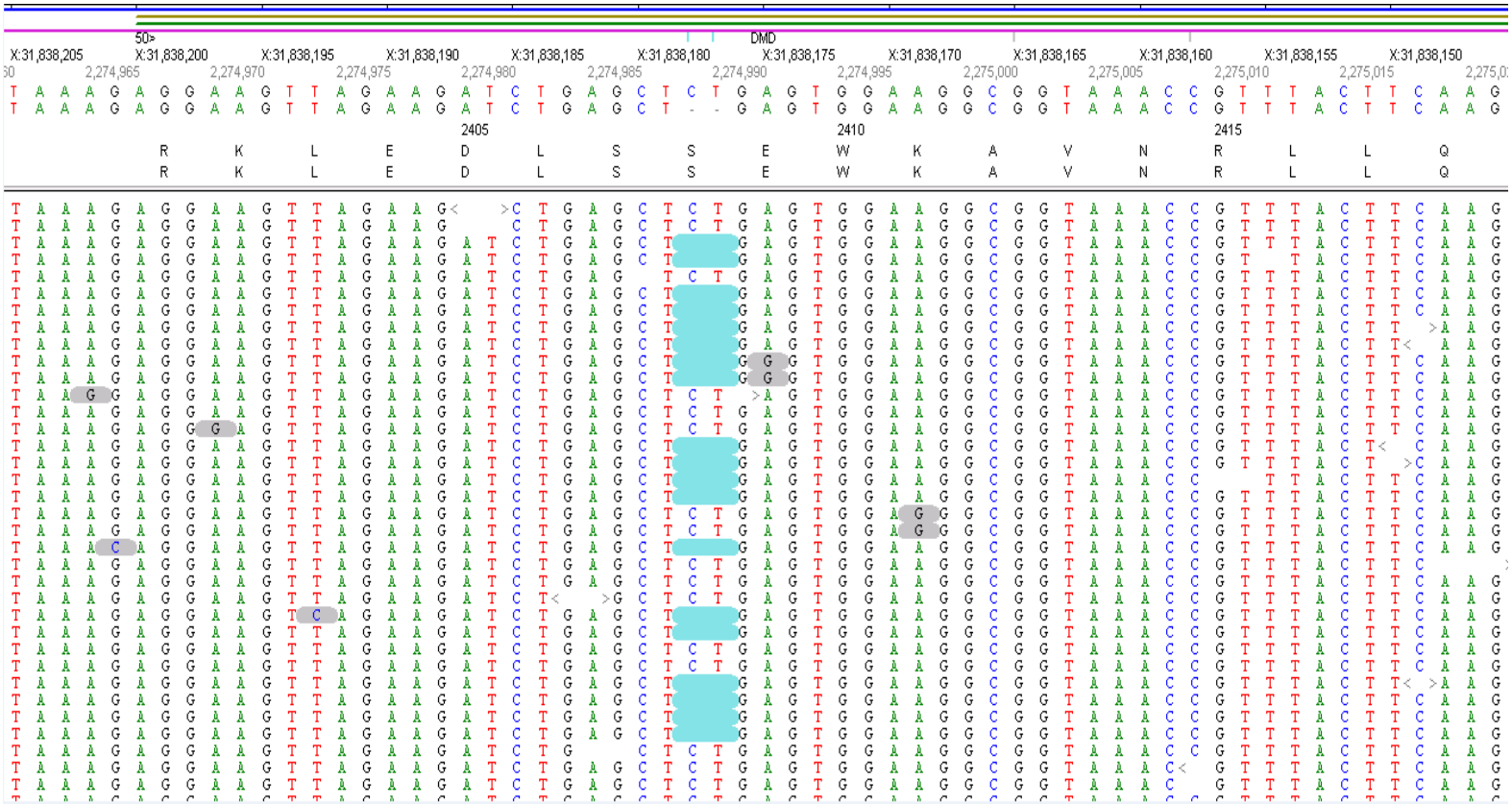
The average coverage for the all genes in the comprehensive NMD targeted gene panel assay was > 200x.

The low-coverage exons totaled 68, which involve 22 genes

NGS + aCGH + sanger fill-in of low coverage regions

Duchenne muscular dystrophy

FEMALE CARRIER 2619/11 DMD c.7223_7224delCT



Exome sequencing

has the advantage of being unbiased regarding what set of genes is analyzed, enabling parallel interrogation of most of the genes in the human genome.

Differently from gene panels, exploring selected regions, diagnostic exome can discover novel disease genes

Exome sequencing

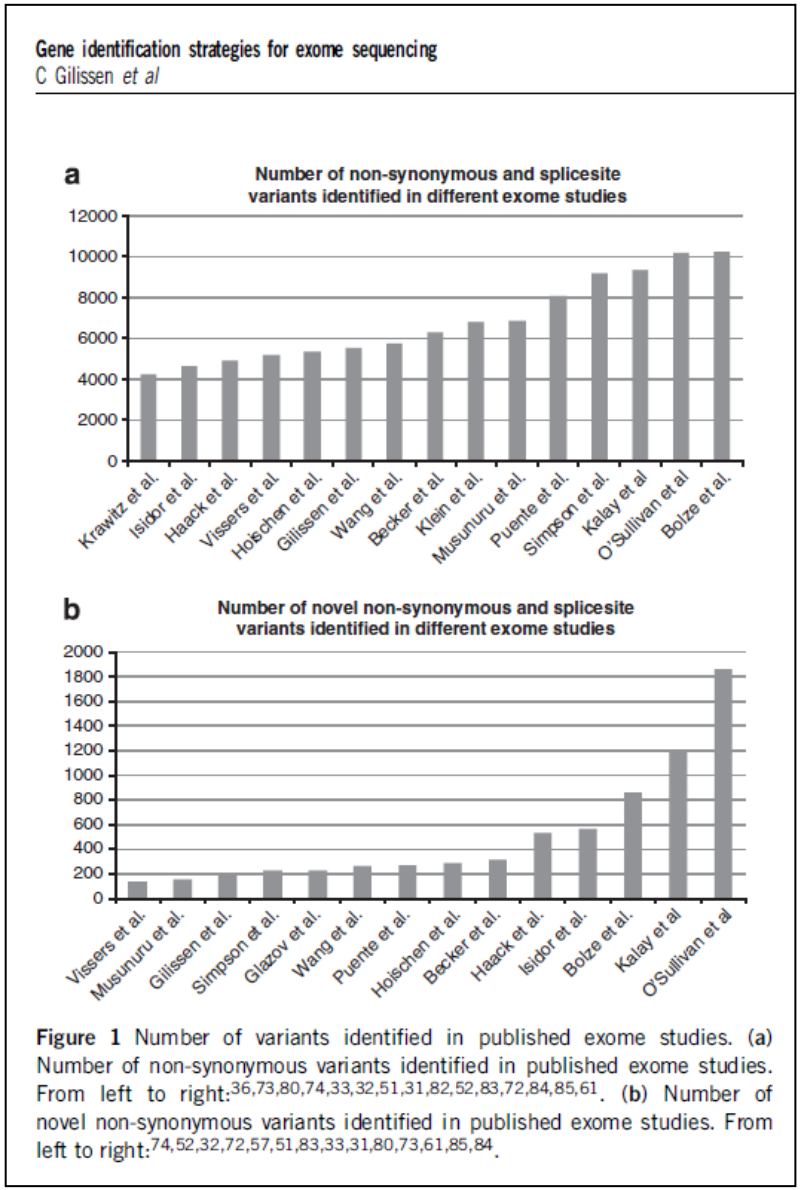


Figure 1 Number of variants identified in published exome studies. **(a)** Number of non-synonymous variants identified in published exome studies. From left to right:[36,73,80,74,33,32,51,31,82,52,83,72,84,85,61]. **(b)** Number of novel non-synonymous variants identified in published exome studies. From left to right:[74,52,32,72,57,51,83,33,31,80,73,61,85,84].

Number of variants per sequenced exome:

20.000 - 50.000

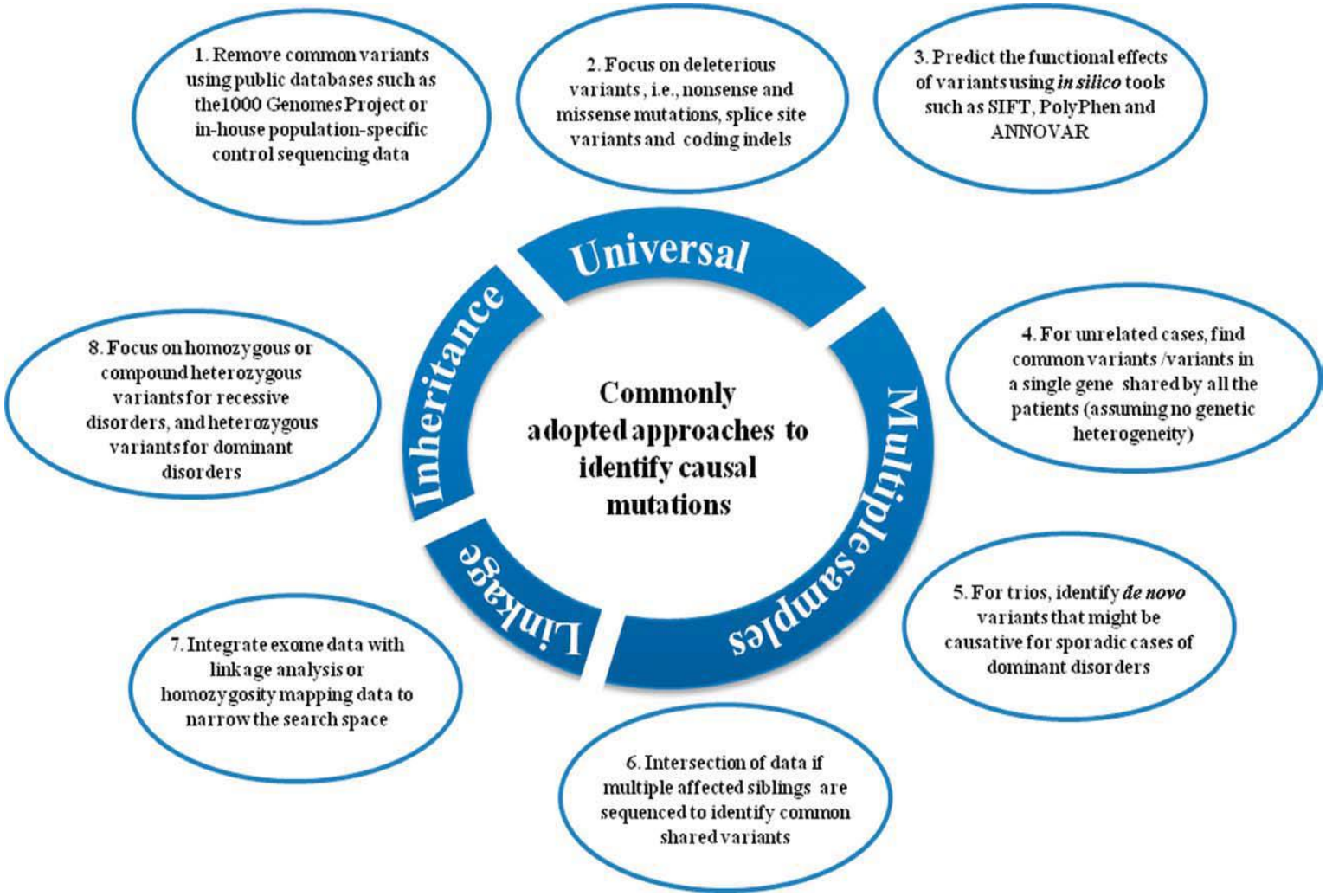
Filtering for:
quality criteria
non coding variants
synonymous variant

Potential disease causing variants: 5000

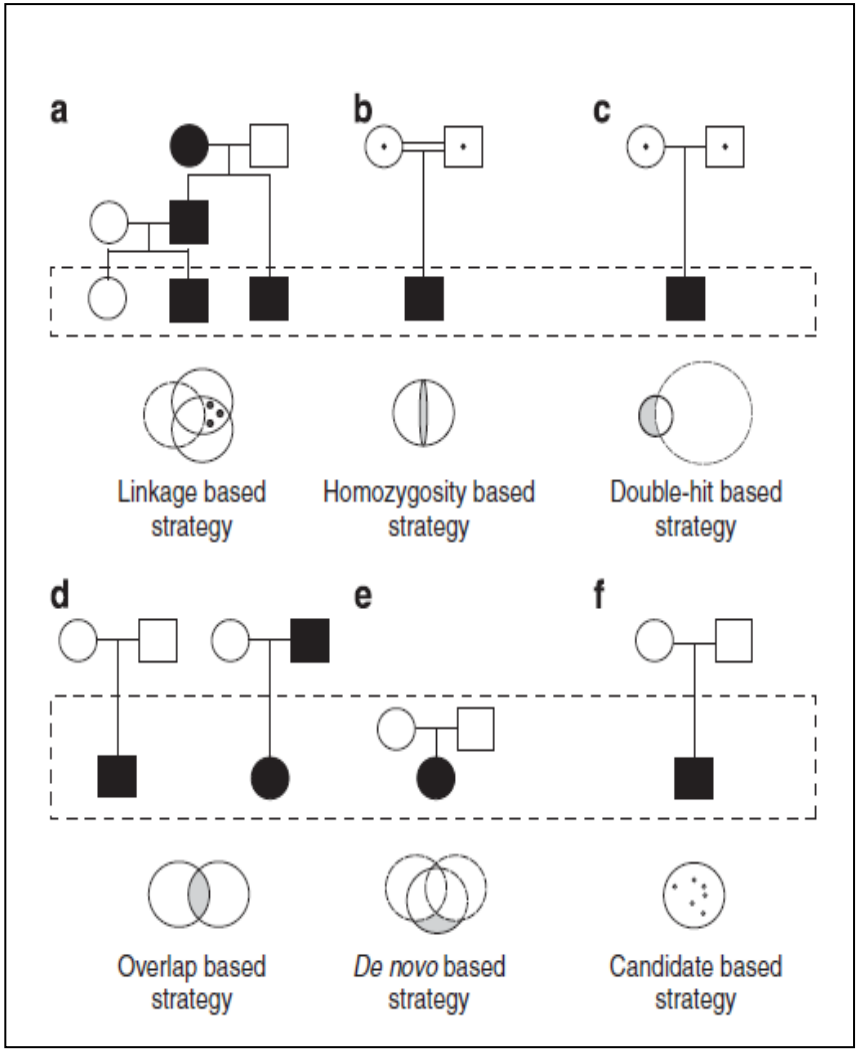
Filtering for:
Known polymorphic variants

Private, non-synonymous, splice-site variants
150-500

BIOLOGICAL INTERPRETATION OF DATA AND IDENTIFICATION OF CAUSAL MUTATION



DISEASE GENE IDENTIFICATION STRATEGIES FOR EXOME SEQUENCING:



a:LINKAGE BASED
shared variations among affected family members

b: HOMOZYGOUS BASED
variants prioritized by their presence in large homozygous regions

c: DOUBLE-HIT
select genes carrying homozygous as well as compound heterozygous variants

d: OVERLAP BASED
Mutations in the same genes in multiple unrelated individuals with a similar phenotype

d: *DE-NOVO* BASED
Family based exome sequencing searching for de novo variations

e: CANDIDATE BASED
Prioritation based on gene function, most damaging change

NGS strategies comparison

Advantages and disadvantages of different MPS strategies

Strategies	Targeted sequencing	Exome sequencing	Genome sequencing
Coverage depth	++	+	+
Range of mutation : Large rearrangement, ...	++	+	+++
Interpretation	+++	+++	+++(exon)/+(intron)
Ethical issue	+	+++ or + (targeted analysis)	+++ or + (targeted analysis)
Data re-analysis for novel disease genes	NO	YES	YES
Cost (2015)	+	++	+++

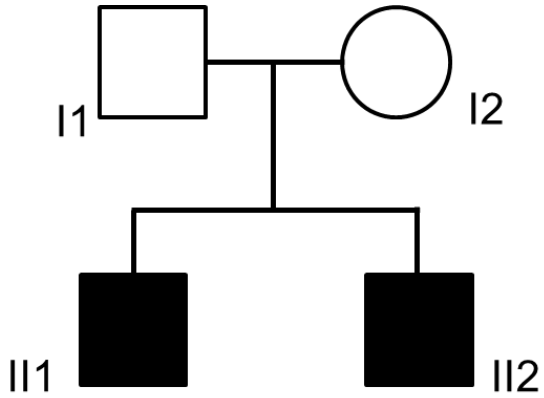
Costs: the cost for targeted sequencing of about 150 genes is usually equivalent to the cost for analysing 3-4 genes of medium size by sanger sequencing.

CLINICAL CASE

Two brothers (40 and 44 years) presenting with a distal axonal motor neuropathy and a focal autonomic dysfunction, which main manifestation was retrograde ejaculation.

One of the two sibs also suffered from enhanced gastro-intestinal motility and bradycardia. First symptoms arose around 27 years of age and presentation was with cramps, progressive difficulties in feet dorsiflexion, lower limbs distal muscle weakness, distal muscle atrophy and skin laxity.

EMG and muscular biopsy showed neurogenic patterns. EMG: axonal motor neuropathy. CPK values: 415-478 U/L (nv 0-190).



From 2003 to 2013:

Sanger sequencing excluded CMT1A-1B-2A, CMTX, LMNA, Kennedy disease, HSPB1 e HSPB8 related neuropathies

LMND GENE PANEL

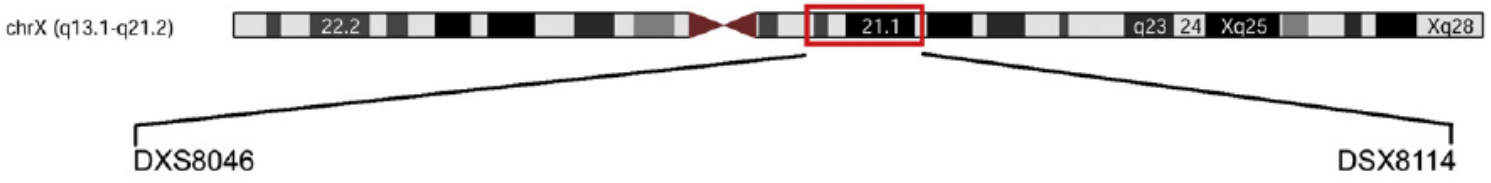
65 genes:

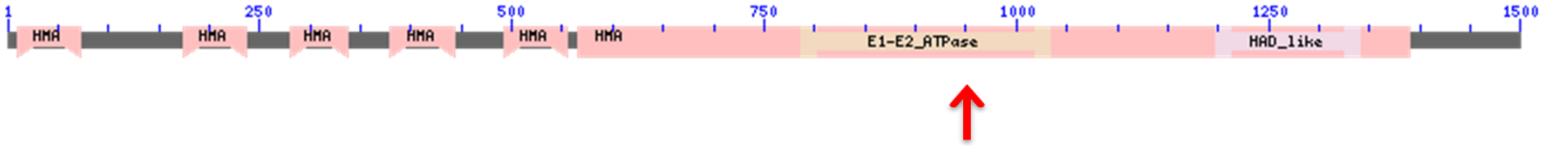
21 ALS-ASSOCIATED GENES

13 SMA-ASSOCIATED GENES

16 HMN-ASSOCIATED GENES

14 CMT-ASSOCIATED GENES





HEMIZYGOSITY FOR THE MISSENSE VARIATION c.2972C>A (p.A991D) IN EXON 15 OF THE ATP7A IN BOTH AFFECTED BROTHERS; HETEROZYGOSITY of their unaffected mother for the same variant.

X-linked recessive inheritance

ATP7A gene (ATPase, Cu++transporting, alpha polypeptide), localized on Xq21.1, encodes a trasmembrane copper transporter P-type ATPase and controls cellular copper homeostasis.

The copper is a cofactor of several enzymes, useful energetic cellular metabolism, iron metabolism and collagen maturation.

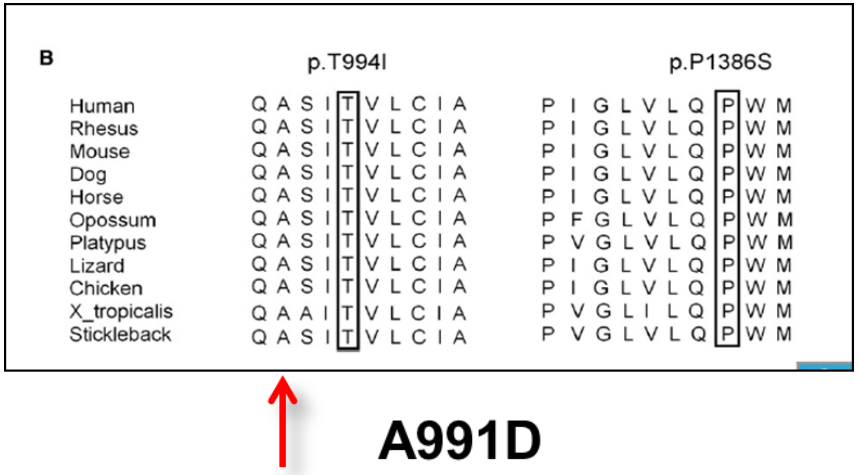
ATP7A -RELATED PHENOTYPES

MENKES DISEASE
Severe infantile neurodegenerative disease

«OCCIPITAL HORN SYNDROME» OR «X-LINKED CUTIS LAXA»
Milder variant with skeletal and skin abnormalities and autonomic dysfunction

ISOLATED DISTAL MOTOR NEUROPATHY
In 2010, ATP7A missense mutations were identified in two families with isolated X-linked dHMN (SMAX3) (Kennerson ML et al.) No additional clinical signs were present.

One of the reported variant (p.T994I) is within the same a-helix-TH domain of the novel mutation identified in our family, only 3 residues apart.



T994I e P1386S

dHMN

Distal motor neuropathy

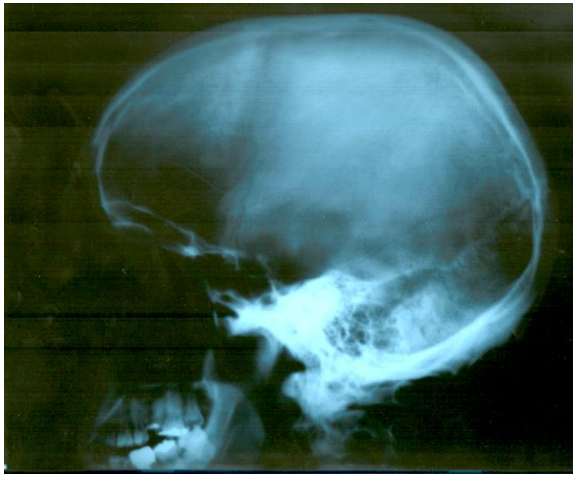
Other ATP7A variants

OHS

Autonomic dysfunction
NO distal neuropathy

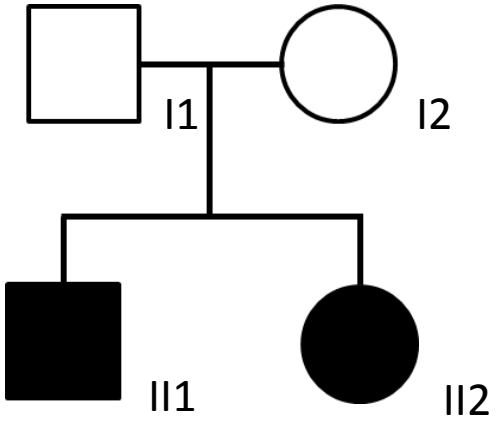
Skin laxity
Inguinal ernia
Occipital exostosis

ATP7A A991D



Occipital Exostosis

"Reverse phenotyping":



Family with congenital myopathy

I1 – I2:
non consanguineous healthy parents

II1:

- Normal ar birth; independent walk at 17 mo; **frequent falls and mild muscular weakness**; tiptoe walking (specific surgery at age 13 yrs).
- Height and weight below 50p
- **myalgia**
- Muscle biopsy (5 yrs): «compatible with congenital myopathy with fiber I predominance and hypotrophy of some of them; Negative IIC for dystrophin, alfa-sarcoglycan, caveolin 3. Immunoblot for calpain 3, dysferlin.
- **CPK: 294 UI/L**
- 15 yrs: the boy is able to walk with widen base, climb stairs with unilateral hand rail support, stand up from the floor with Gower's manouvre, jump.
- **MRC scale 4 in all muscles except for 4- - in limb girdle muscles**. No muscle hypotrophy
- Cardiac and respiratory functions are normal

II2:

- At birth weak suction with gastroesophageal reflux, and frequent deglutition problems
- independent walk at 18 mo; **frequent falls and mild muscular weakness**; no tiptoe walking
- delayed speech development
- Height and weight below 50p
- **Myalgia**
- **CPK: 344 UI/L**
- 7 yrs: the girl is able to walk with widened base, climb stairs with bilateral hand rail support, jump and stand up from the floor with incomplete Gower's manouvre
- **MRC scale 4 in all muscles except for 4- - in limb girdle muscles.** Mild muscle hypotrophy
- Mild lumbar rigid spine; finger hyperlaxity
- Cardiac and respiratory function are normal

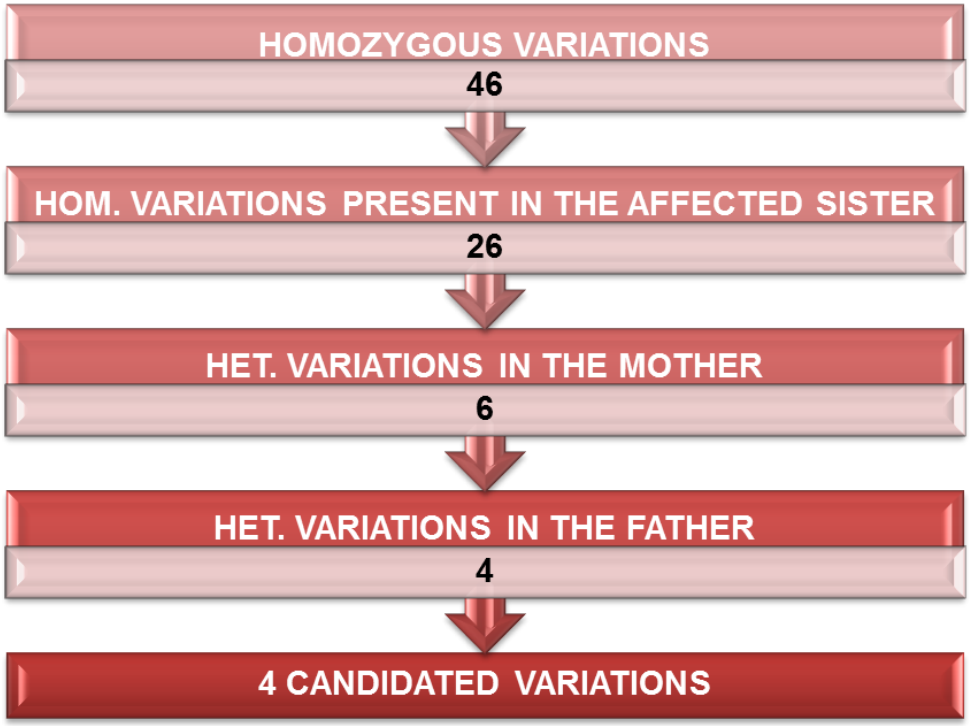
Whole Exome sequencing performed by

deCODE genetics

All variations
6913
(5440het -1473hom)

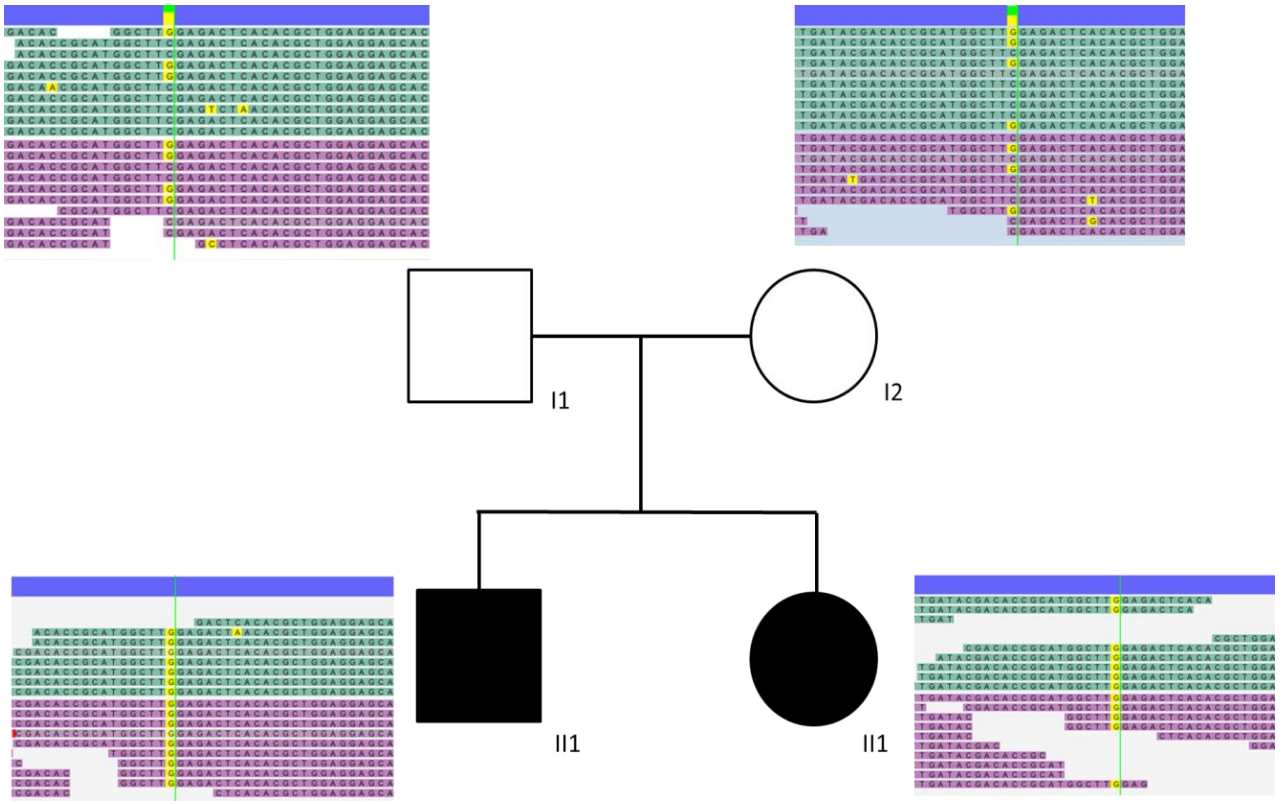
vep consequence filter 571
(525het -46hom)

variants prioritization with SEQUENCE MINER



Validated by Sanger sequencing

CHR	GENE	POSITION	REF	CALL	INDEX	CASE	FATHER	MOTHER	FUNCTION	AA	PREDICTION
19	WTIP	34984389	G	C	hom	hom	het	het	splice region	-	polymorphism
19	ATP4A	36051324	G	A	hom	hom	het	het	missense	H244Y	DAMAGING
19	**RYR1**	**39076790**	**C**	**G**	**hom**	**hom**	**het**	**het**	**missense**	**F4976L**	**DAMAGING**
19	PSG6	43414725	TACTC	T	hom	hom	het	het	splice region	-	polymorphism



RYR1 – huge gene composed of 106 exons
(heavy traditional Sanger Sequencing)

Muscular biopsy: no mentioned the presence of clear «cores»

The RYR1 gene was previously screened in the index case only for hot spot regions.
This RYR1 pathogenic variant correlates with the clinical phenotype

Revaluation of patients in which only hot spot regions have been screened due to gene complexity

DIAGNOSTIC EXOME: CHALLENGES

Technical limitations :

- Not all the exons in the genome are targeted. Many genomic regions cannot be captured and sequences, including many exons. >97% of exons are targeted
- About 10% of exons that are targeted may not be well covered
- Certain mutation types are not detectable, such as large rearrangements, CNVs, mitochondrial genome mutations, trinucleotide repeat expansios, epigenetic effects.

When novel variants/novel disease gene are identified, a complex post-test workflow is needed to prove disease causing significance.

Test Genetici, BENEFICI e RISCHI

- ✓ formulazione o conferma della diagnosi clinica
- ✓ disponibilità di informazioni prognostiche
- ✓ prevenzione di possibili complicanze (MEDICINA PREDITTIVA)
- ✓ potenziale beneficio terapeutico, TRATTAMENTO PERSONALIZZATO
- ✓ riduzione di indagini inutili

- ✓ conoscenza di condizioni prive di alcun approccio terapeutico
- ✓ difficoltà interpretative del dato, confusione tra rischio e certezza, tra mutazione e malattia
- ✓ quali risultati comunicare? A chi?
- ✓ nuove problematiche etiche (i risultati ottenuti rappresentano una condizione permanente dell'individuo e possono influenzarne le scelte riproduttive e avere importanti ricadute sul soggetto in esame e su altri membri della famiglia)

Per una corretta gestione del test e del dato genetico:

✓ PRESENZA DI **NORMATIVE** NAZIONALI E INTERNAZIONALI A TUTELA DEL PAZIENTE E DEL DATO GENETICO

COUNCIL OF EUROPE

Convention on Human Rights and Biomedicine (1997)

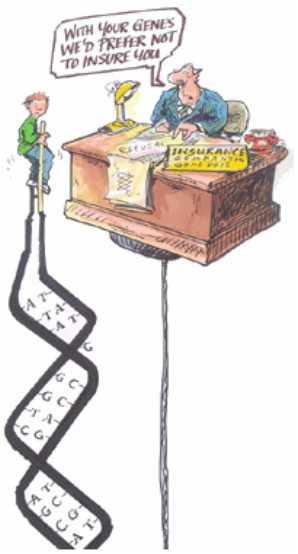
Articolo 11 - proibisce qualsiasi forma di discriminazione nei confronti di chiunque sulla base del proprio patrimonio genetico.

Articolo 12 – i test genetici possono essere impiegati solo per scopi medici o per ricerca.

AUTORIZZAZIONE 11 dicembre 2014.

Autorizzazione generale al trattamento dei dati genetici.
(Autorizzazione n. 8/2014).

IL GARANTE PER LA PROTEZIONE
DEI DATI PERSONALI



✓ PRESENZA DI LINEE GUIDA CHE STABILISCONO A CHI APPLICARE UN TEST GENETICO E PER QUALE FINE

Per una corretta gestione del test e del dato genetico:

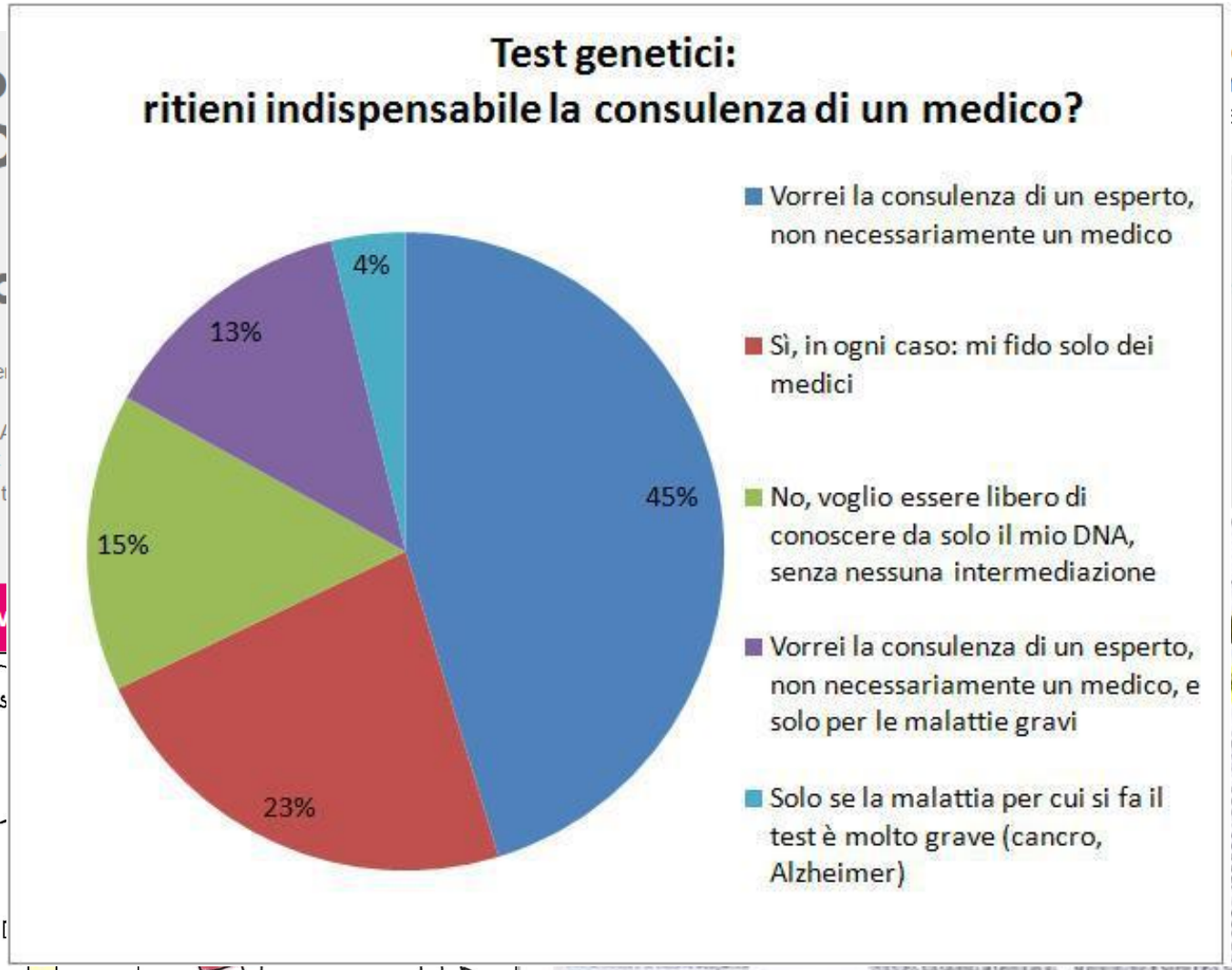
✓ **CONSENSO INFORMATO** ALL'ESECUZIONE DEI TEST GENETICI

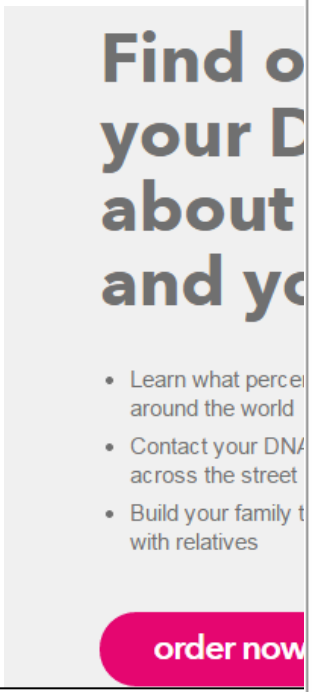
Secondo la legge italiana, richiedono esplicitamente il consenso informato i trattamenti terapeutici, chirurgici e diagnostici, compresi i test genetici, e le sperimentazioni cliniche (D.L. 24 giugno 2003, n. 211) e il paziente deve inoltre acconsentire al trattamento dei propri dati personali (D.L. 30 giugno 2003, n. 196). Come indicazione generale, un consenso scritto deve essere redatto in maniera chiara e comprensibile, deve contenere la descrizione dell'intervento medico, la spiegazione dei rischi prevedibili derivanti dall'esecuzione della prestazione prevista, illustrazioni delle tecniche e degli eventuali materiali da utilizzare, la descrizione dei benefici e dei rischi che possono derivare da eventuali e le potenziali complicazioni.

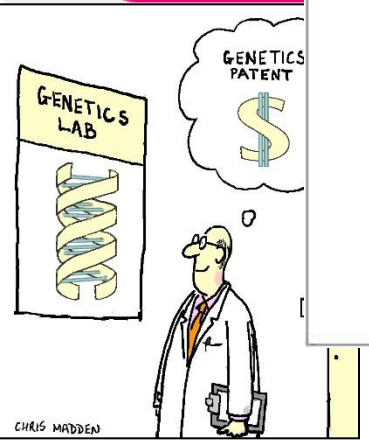
✓ **CONSULENZA GENETICA (PRE-TEST e POST-TEST):**

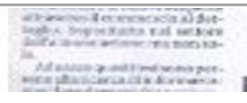
- comprendere le informazioni mediche, inclusa la diagnosi, la prognosi e le terapie disponibili
- rendersi conto del contributo ereditario alla malattia e del rischio di ricorrenza
- prendere le decisioni che sembrano appropriate in rapporto al rischio di ricorrenza, ai progetti familiari, agli standard etici e religiosi e ad agire in accordo con queste decisioni
- ottenere il miglior possibile adattamento alla malattia (in un soggetto affetto) o al rischio di ricorrenza

COMMERCIALIZZAZIONE DEL TEST GENETICO









Niente nella vita va temuto, deve essere solamente compreso. Ora è tempo di comprendere di più, così possiamo temere di meno.
(Marie Curie)

> *La medicina ha fatto così tanti progressi che ormai più nessuno è sano.*