



UNIONE EUROPEA  
Fondo europeo di sviluppo regionale



REPUBBLICA ITALIANA



REGIONE AUTÒNOMA DE SARDIGNA  
REGIONE AUTONOMA DELLA SARDEGNA



SARDEGNA  
RICERCHE

# Nuove tecnologie diagnostiche non invasive basate sul sequenziamento di nuova generazione

*Malattie Rare – La ricerca per la vita*  
*3 Marzo 2018*

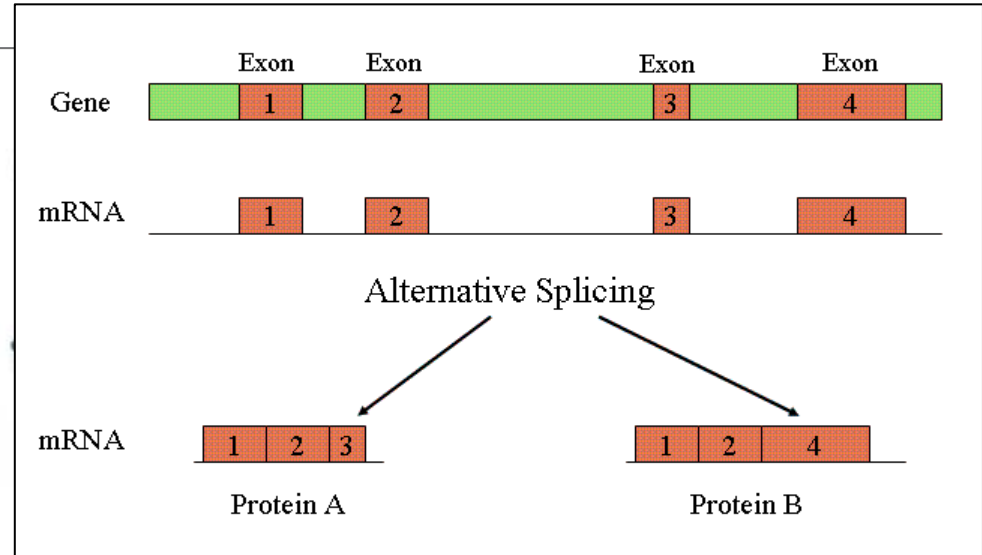
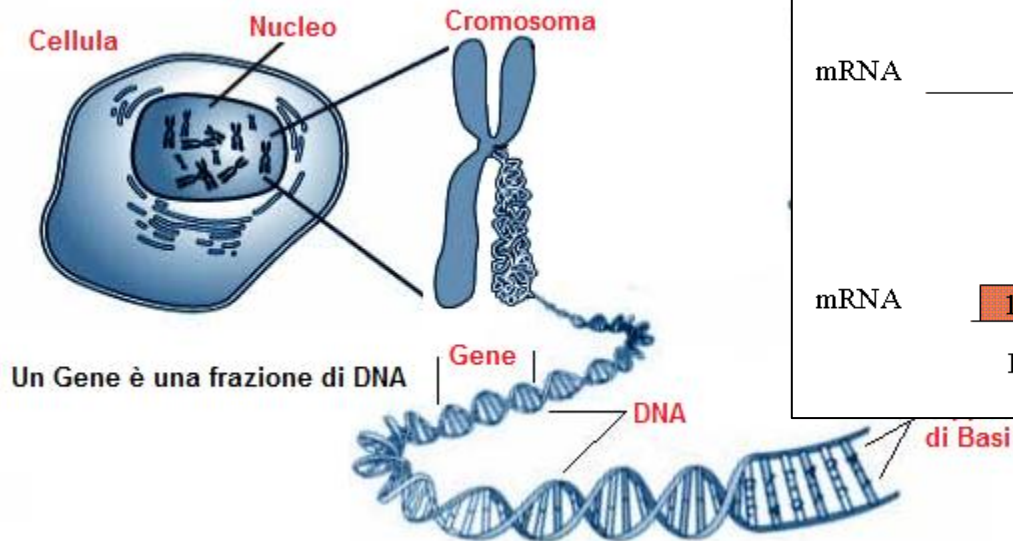
**Paolo Uva**  
CRS4 Biosciences  
NGS Core facility  
Pula CA, Italy



- Sequenziamento Nuova Generazione (NGS)
- Metodi non invasivi:
  - NIPT e recenti progressi per lo studio di malattie monogeniche
  - Altre applicazioni (biopsia liquida)

# Cosa è il DNA

## STRUTTURA DEL DNA

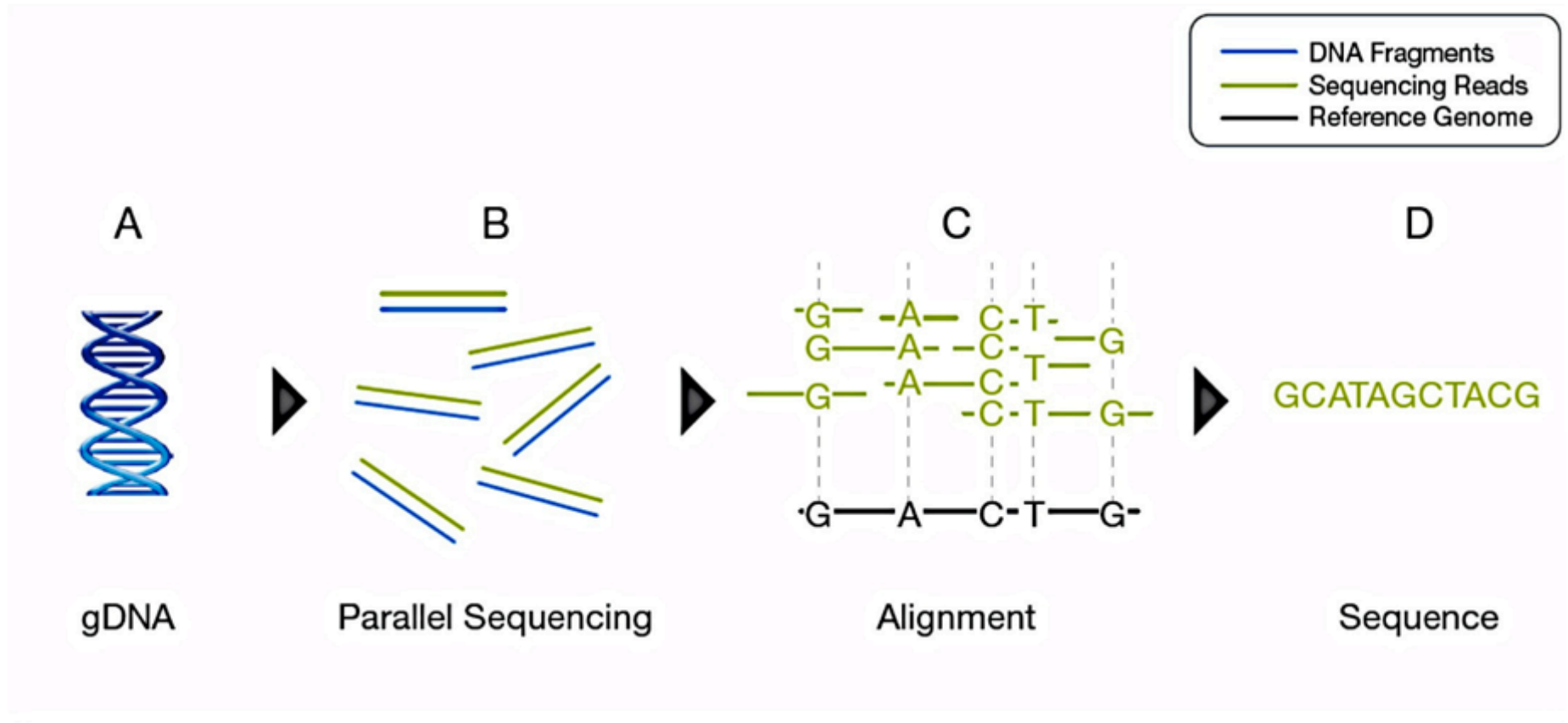


Il DNA umano contiene circa 3 miliardi di paia di basi (A,C,G,T), suddivise in cromosomi.

Il **sequenziamento** è la lettura delle basi del DNA

# Next-Generation Sequencing (NGS)

## Conceptual overview of Whole-Genome Sequencing (WGS)



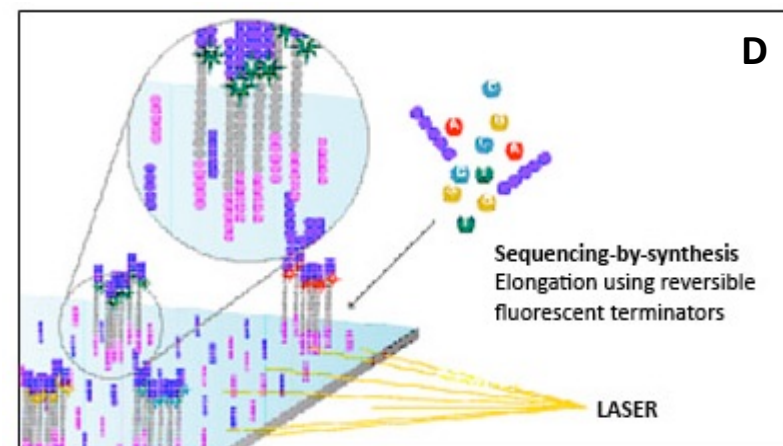
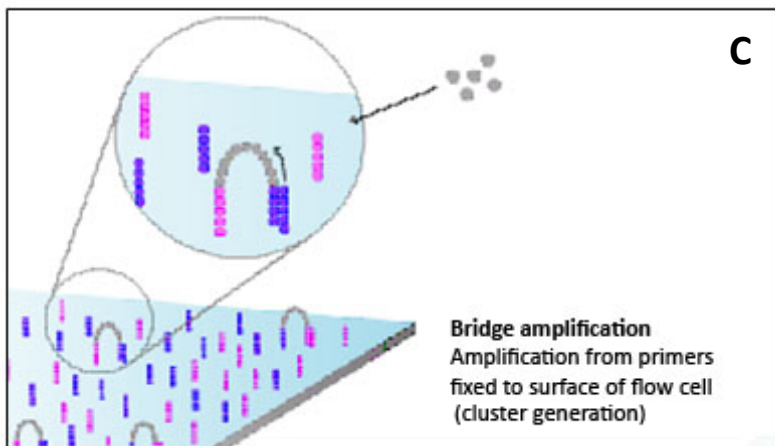
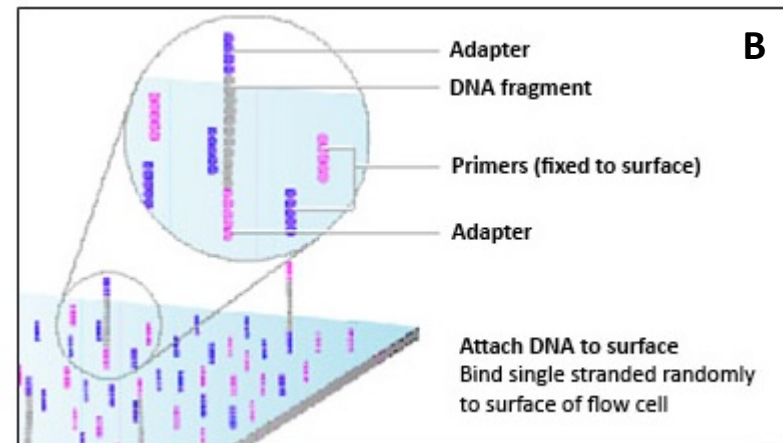
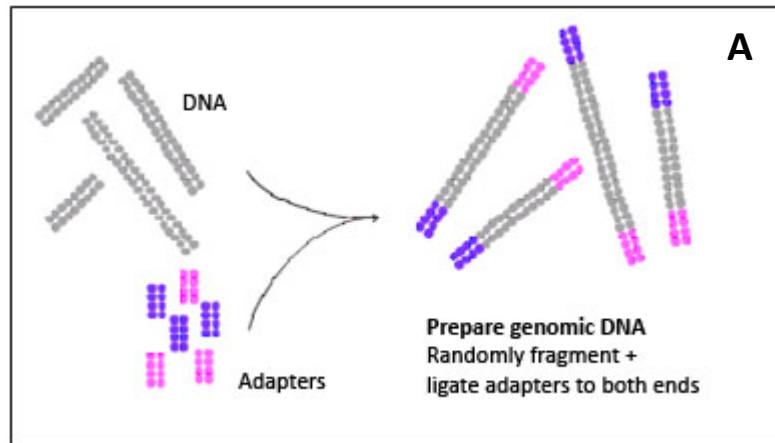
A. Extracted gDNA.

B. gDNA is fragmented into a library of small segments that are each sequenced in parallel.

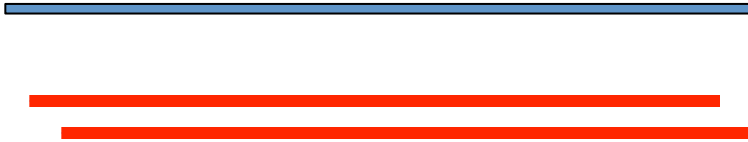
C. Individual sequence reads are reassembled by aligning to a reference genome.

D. The whole-genome sequence is derived from the consensus of aligned reads.

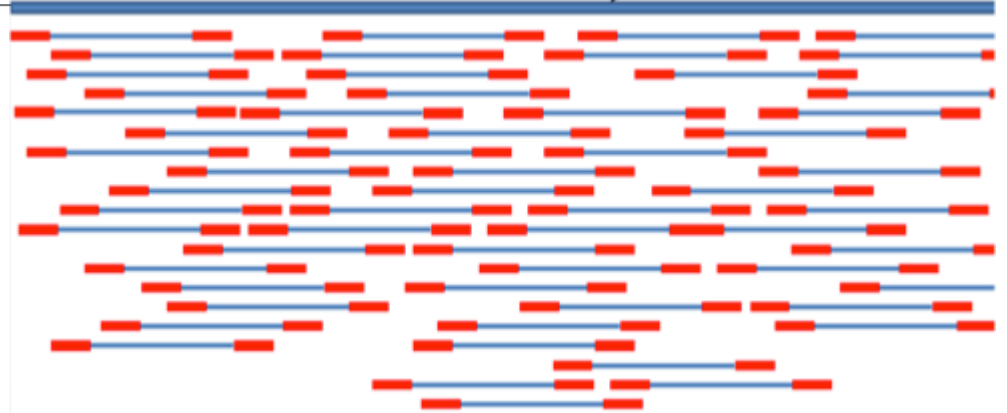
# Next Generation Sequencing



## Sequenziamento classico (Sanger)



## Sequenziamento Nuova Generazione



### Sequenziamento Sanger

- Sequenze lunghe
- Poche sequenze
- **Costo più alto**
- Basso tasso di errore

### Sequenziamento NGS

- Sequenze corte
- Milioni di sequenze
- **Minor costo**
- Ridondanza per compensare errori di lettura

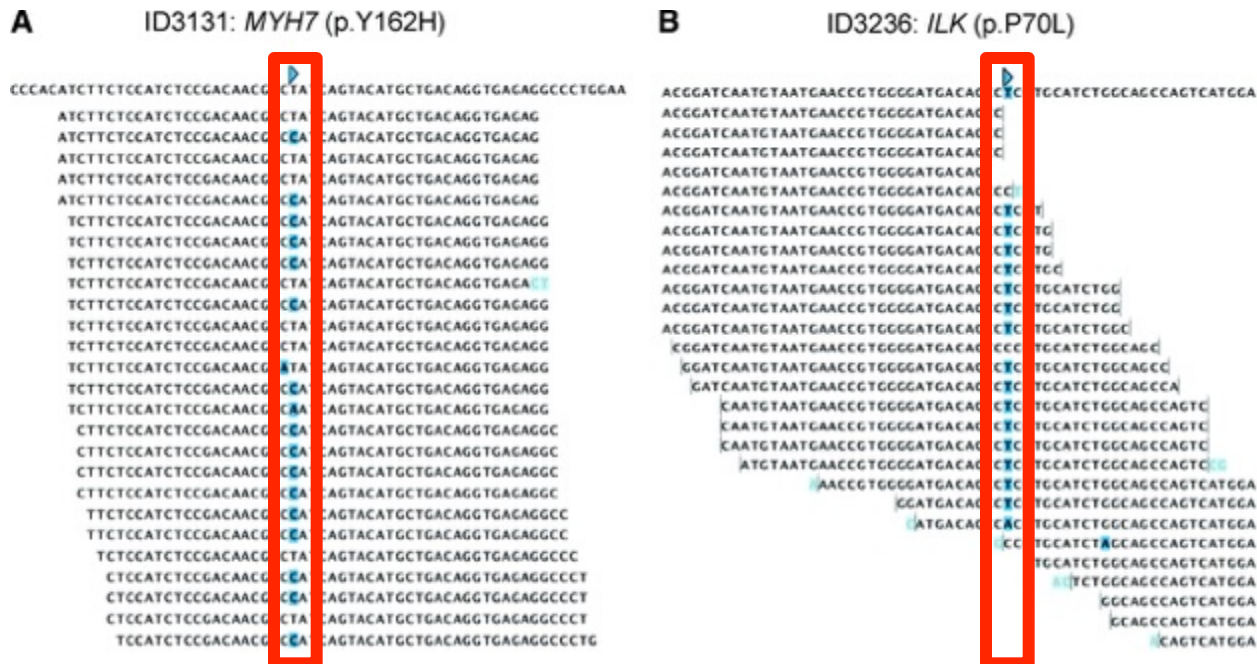




# Dalle sequenze all'identificazione di mutazioni



Analisi



Identificazione delle **varianti**



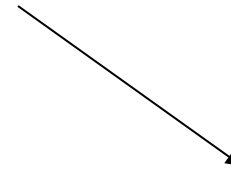
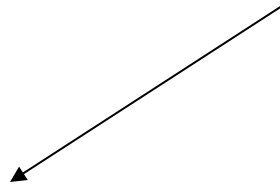


# 0.1% delle differenze tra due individui



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

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



VARIAZIONI  
SENZA EFFETTI

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



Individuo SANO

TC TCG ATATAGC TC GCG ACAC AC ACAG ATATATAGC GTAGGGCTC TC TATATAGCTC GCG AC ACAC AC AGATATATAGC TC C CTG AAACAGC  
TC CGAC AC AGC TCGC AC ACCGCTC GAG ACC TG ACC TGAC AC G TGC TAGCTAGC TCC TCTC GAG AC G TAGGGC TC TCG ATATAGC TC GCG AC  
AC ACAC AG ATATATAGCGCTCC TGAAAC AGC TCCG AC ACAGCTC GC ACAC CGC TC G AG ACC TGACCTG AC ACG TGC TAGC TAGCTCCTC TC G A  
GAC GTAGGGCTC TC G ATATAGCTC GCG AC ACAC AC AGATATATAGCGC TCC CTG AAACAGCTC CG ACAC AGC TC GCAC AC CGC TCG AGACCTTA  
GCTAGC TCC TC TCG AGACGTAGGGCTC TC G ATATAGCTC GCG AC ACAC AC AGATATTATAGC TCG C GAC AC ACAC AG ATATATAGCG TAGGGC T  
CTC GATATAGC TCG C GAC AC ACAC AG ATATATAGC GCTC CC TGAAAC AGC TCCG AC ACAGCTC GC ACAC CGC TC G AG ACC TGACCTG AC ACG TG  
CTAGCTAGC TCC TC TCG ACG AGACGTAGGGCTC TC G ATATAGCTC GCG AC ACAC AC AGATATATAGCGC TCC CTG AAACAGCTC CG ACAC AGC T  
CGCA ACCGCTC GAG ACC TG ACC TGAC AC G TGC TAGCTAGC TCC TCTC GAG AC G TAGGGC TC TCG ATATAGC TC GCG ACAC AC ACAG ATATATA  
GCGCTCC TGAAAC AGC TCCG AC ACAGCTC GC ACAC CGC TC G AG ACC TTAGCTAGC TCC TCTC GAG AC G TAGGGC TC TCG ATATAGC TC GCG AC  
AC ACAC AG ATATATAGCGCTCC TGAAAC AGC TCCG AC ACAGCTC GC ACAC CGC TC G AG ACC TGACCTG AC ACG TGC TAGC TAGCTCCTC TC G A  
GAC GTTATAGC TCG C GAC AC ACAC AG ATATATAGCG TAGGGC TC TCG ATATAGC TC GCG ACAC AC ACAG ATATATAGCGCTCC TGAAAC AGC T  
CCG AC ACAGCTC GC ACAC CGC TC G AG ACC TGACCTG AC ACG TGC TAGC TAGCTCCTC TC G AG ACG TAGGGC TCTC GATATAGC TCG C GAC A  
CAC AC AG ATATATAGCGCTCC TGAAAC AGC TCCG AC ACAGCTC GC ACAC CGC TC G AG ACC TGACCTG AC ACG TGC TAGC TAGCTCCTC TC G A  
ACG TAGC TCC TC TCG ACG AGACGTAGGGCTC TC G ATATAGCTC GCG AC ACAC AC AGATATATAGCGC TCC CTG AAACAGCTC CG ACAC AGC TC  
CTAGC TCC TCTC GAG AC G TAGGGC TC TCG ATATAGC TC GCG ACAC AC ACAG ATATA TAGCGCTCC TGAAAC AGC TCCG AC ACAGCTC GC ACAC  
CGC TC G AG ACC TGACCTG AC ACG TGC TAGC TAGCTCCTC TC G AG ACG TAG ACG TAGGCTCG C GAC AC ACAC AG ATATATAG  
CGG ACG TAGGGC TC TCG ATATAGC TC GCG ACAC AC ACAG ATATATAGCGCTCC TGAAAC AGC TCCG AC ACAGCTC GC ACAC CGC TC G AG ACC TGACCTG AC A  
CTAGC TCC TCTC GAG AC G TAGGGC TC TCG ATATAGC TC GCG ACAC AC ACAG ATATATAGCGCTCC TGAAAC AGC TCCG AC ACAGCTC GC ACAC  
CGC TC G AG ACC TGACCTG AC ACG TGC TAGC TAGCTCCTC TC G AG ACG TAGGGC TCTC GATATAGC TCG C GAC AC ACAC AG ATATATAGCG TAGG  
GC TCTC GATATAGC TCG C GAC AC ACAC AG ATATATAGCGCTCC TGAAAC AGC TCCG AC ACAGCTC GC ACAC CGC TC G AG ACC TGACCTG AC A  
GTGCTAGC TAGC TC C TC TCG ACG AGACGTAGGGCTC TC G ATATAGCTC GCG AC ACAC AC AGATATATAGCGC TCC CTG AAACAGCTC CG ACAC A  
GC TCGC AC ACCGCTC GAG ACC TG ACC TGAC AC G TGC TAGCTAGC TCC TCTC GAG AC G TAGGGC TC TCG ATATAGC TC GCG ACAC AC ACAG ATAT  
ATAGCGCTCC TGAAAC AGC TCCG AC ACAGCTC GC ACAC CGC TC G AG ACC TTAGCTAGC TCC TCTC GAG AC G TAGGGC TC TCG ATATAGC TC G C  
GAC AC ACAC AG ATATATAGCGCTCC TGAAAC AGC TCCG AC ACAGCTC GC ACAC CGC TC G AG ACC TGACCTG AC ACG TGC TAGC TAGCTCCTC T  
CG AGACGTAGGGCTC TC G ATATAGCTC GCG AC ACAC AC AGATATATAGCGCTAGGGC TCTC GATATAGC TCG C GAC AC ACAC AG ATATATAGCG C  
TCC CTG AAACAGCTC CG ACAC AGC TC GCAC AC CGC TC G AGACCTG AC C TG ACACGTGCTAGC TAGC TCC TC TCG ACG AGACGTAGGGCTC TC G A  
TATAGC TC GCC TCG C GAC AC ACAC AG ATATATAGCGTAGGGC TC TCG ATATAGC TC GCG ACAC AC ACAG ATATATAGCGCTCC TGAAAC AGC T  
CCG AC ACAGCTC GC ACAC CGC TC G AG ACC TGACCTG AC ACG TGC TAGC TAGCTCCTC TC G AG ACG TAGGGC TCTC GATATAGC TCG C GAC A  
CAC AC AGATATATAGCGCTCC TGAAAC AGC TCCG AC ACAGCTC GC ACAC CGC TC G AG ACC TGACCTG AC ACG TGC TAGC TAGCTCCTC TC G A  
CTG AAACAGCTC CG ACAC AGC TC GCAC ACC GCTCG AGACCTG AC C TG ACACGTGCTAGC TAGC TCC TC TCG AGACGTAGGGCTC TC G ATATAGC  
TC GCG ACAC AC ACAG ATATATAGCGCTCC TGAAAC AGC TCCG AC ACAGCTC GC ACAC CGC TC G AG ACC TGACCTG AC ACG TGC TAGCTAGC TC  
CTC TC G AG ACG TAGGGC TCTC GATATAGC TCG C GAC AC ACAC AG ATATATAGCGCTCC TGAAAC AGC TCCG AC ACAGCTC GC ACAC CGC TC G A  
GACCTG AC C TG ACACGTGCTAGC TAGC TCC TC TCG AGACGTAGGGCTC TC G ATATAGCTC GCG AC ACAC AC AGATATATAGCGCTCC CTG AAAC  
AGC TCCG AC ACAGCTC GAG ACC TG ACC TGAC AC G TGC TAGCTAGC TCC TCTC GAG AC G TAGGGC TC TCG ATATAGC TC GCG AC  
AC ACAC AG ATATATAGCGCTCC TGAAAC AGC TCCG AC ACAGCTC GC ACAC CGC TC G AG ACC TGACCTG AC ACG TGC TAGC TAGCTCCTC TC G A  
GACGTAGGGCTC TC G ATATAGCTC GCG AC ACAC AC AGATATATAGCGC TCC CTG AAACAGCTC CG ACAC AGC TC GCAC AC CGC TCG AGACCTG A  
CC TGAC AC G TGC TAGCTAGC TCC TCTC GAG AC G TAGGGC TC TCG ATATAGC TC GCG ACAC AC ACAG ATATATAGCGCTCC TGAAAC AGC TC C

Senza effetti

Patogenetica

Innocua

# Anomalia genetica



Genoma



Capitolo:  
Cromosoma



Paragrafo:  
Microdelezioni  
Microduplicazioni



Lettera:  
Mutazioni  
puntiformi

Cariotipo



Microarray



Sequenziamento



# CRS4 – NGS Core Facility

## Pula – Polaris

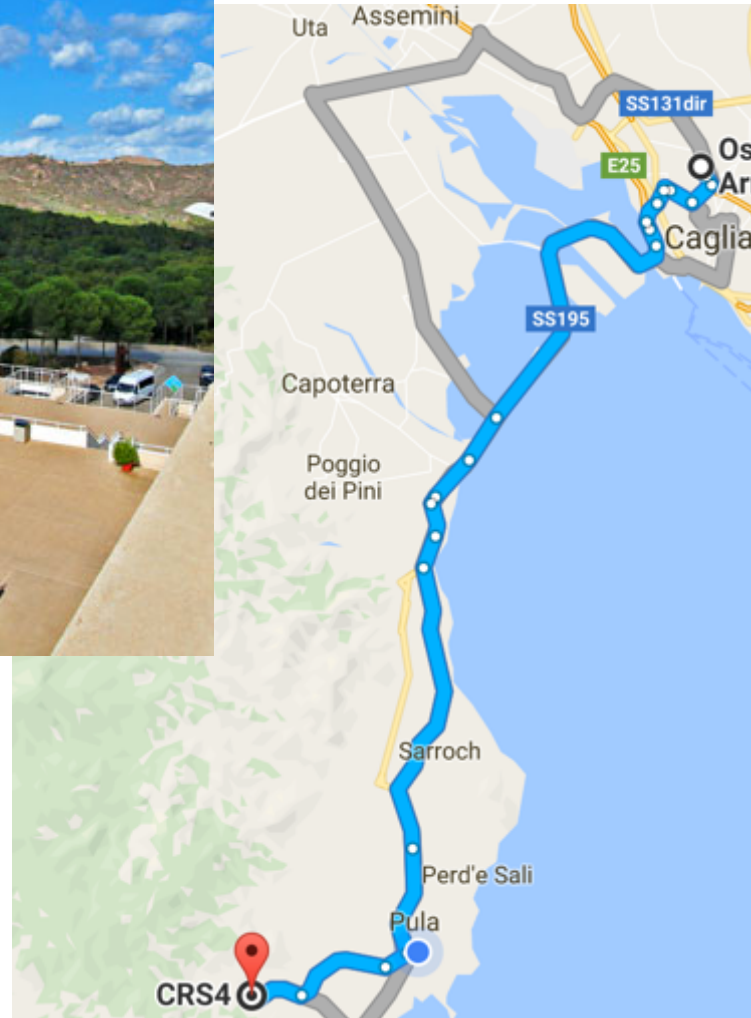
- High-throughput NGS facility
  - La più **grande** piattaforma di sequenziamento NGS in Italia
  - X000 genomi, x000 RNA-Seq, x00 esomi
- Centro di calcolo
  - Tra i top 5 in Italia, **3200+ processori**, **Pb spazio disco**
- Piattaforma di analisi
  - Metodi per la gestione e l'analisi dei dati di sequenziamento



# Dove siamo



**CRS4 NGS Core Facility**  
Parco Scientifico e Tecnologico Polaris  
09010 – Pula, Cagliari



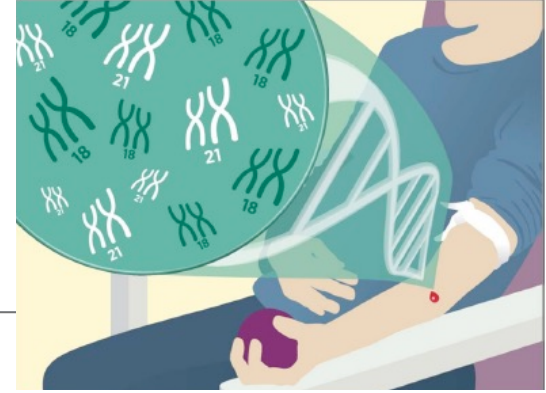
# Test pre-natale non invasivi per malattie genetiche

---

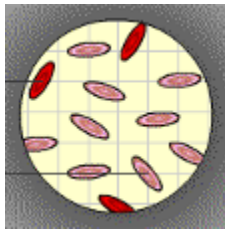
	<b>Falsi positivi</b>
Età materna	5-25%
Screening biochimico	5%
Traslucenza nucale	5%
<i>Test integrato</i>	2-5%
<b>DNA fetale circolante</b>	<1%



# Dov'è presente il DNA fetale?

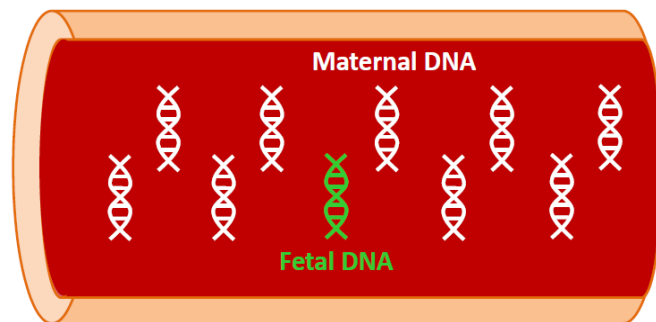


- Le cellule di origine fetale sono presenti nel sangue materno



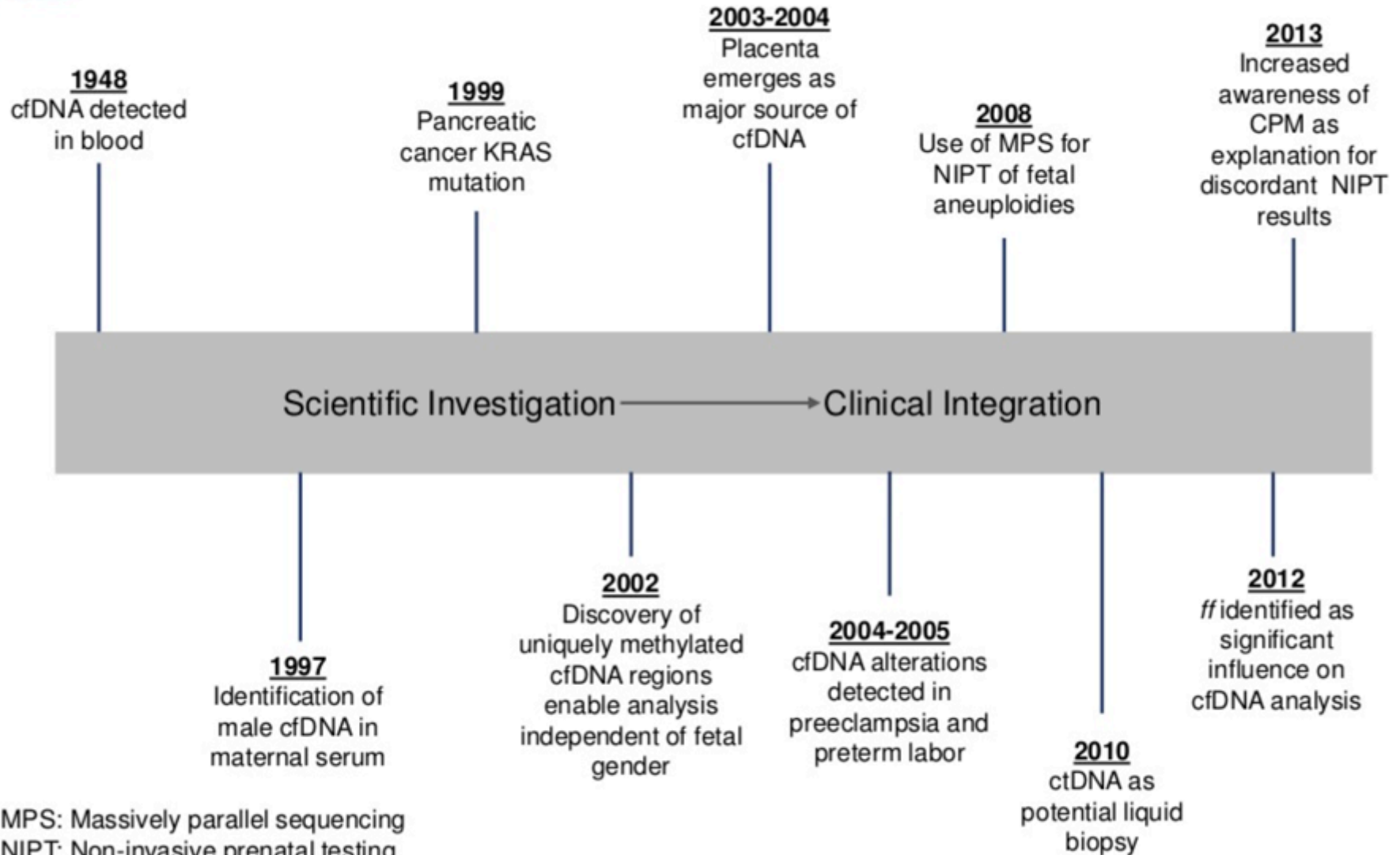
- Separazione delle cellule materne e fetali
- Poche cellule fetali disponibili: necessario arricchimento

- **DNA fetale cell-free (cffDNA)** nel sangue materno



- Il cffDNA è presente in quantità maggiore rispetto al DNA nelle cellule fetali circolanti (20-100x)
- Rappresenta ~10% del DNA circolante totale
- Rimosso dal sangue materno in < 24h

# Discovery of cfDNA: historical developments



MPS: Massively parallel sequencing  
 NIPT: Non-invasive prenatal testing  
 CPM: Confined placental mosaicism  
 ff: fetal fraction

Modified from: Taglauer, E.S., Wilkins-Haug, L. and Bianchi, L.W. 2014. Review: cell-free fetal DNA in the maternal circulation as an indication of placental health and disease. *Placenta* **35**, S64-68

# Discovery of cfDNA: historical developments

**1948**  
cfDNA detected  
in blood

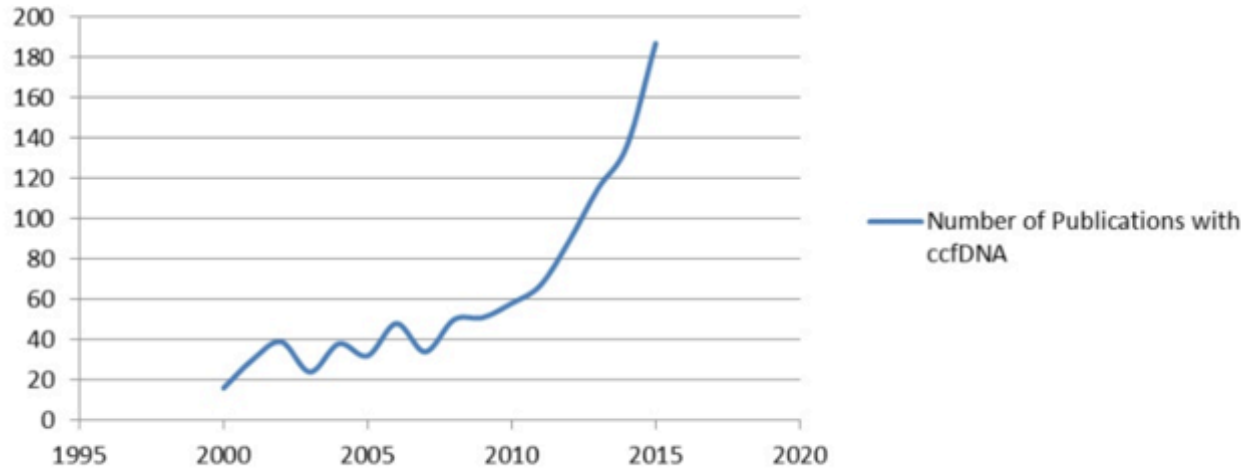
**1999**  
Pancreatic  
cancer KRAS  
mutation

**2003-2004**  
Placenta  
emerges as  
major source of  
cfDNA

**2008**  
Use of MPS for  
NIPT of fetal  
aneuploidies

**2013**  
Increased  
awareness of  
CPM as  
explanation for  
discordant NIPT  
results

**Number of PubMed Publications**



ation

**2012**  
*ff* identified as  
significant  
influence on  
cfDNA analysis

**2010**  
ctDNA as  
potential liquid  
biopsy

MPS: Massively parallel sequencing  
NIPT: Non-invasive prenatal testing  
CPM: Confined placental mosaicism  
*ff*: fetal fraction

Modified from: Taglauer, E.S., Wilkins-Haug, L. and Bianchi, L.W. 2014. Review: cell-free fetal DNA in the maternal circulation as an indication of placental health and disease. *Placenta* **35**, S64-68

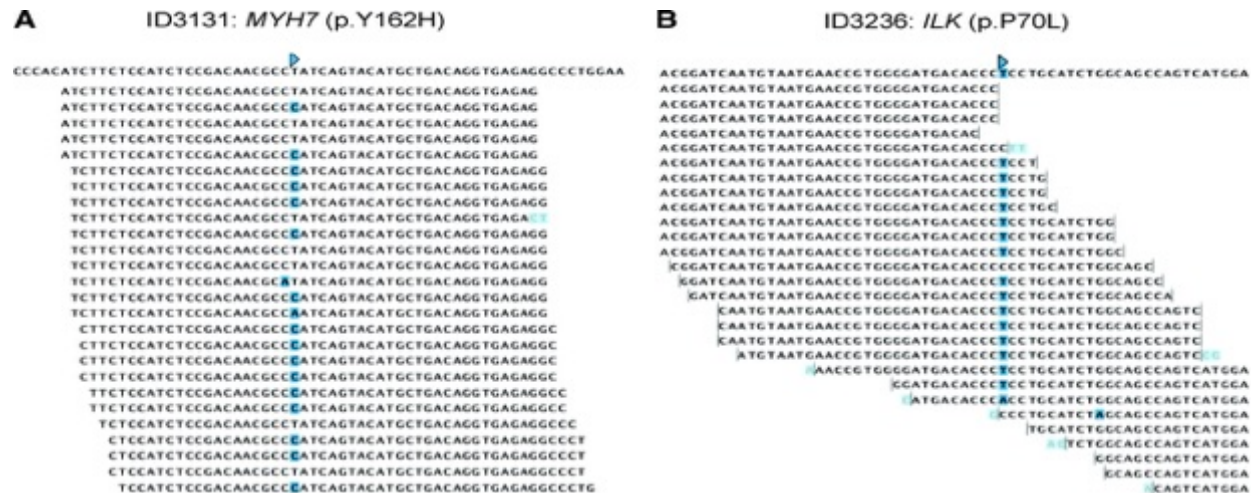
# Fattori chiave per NIPT

## Not Invasive Prenatal Testing

- Frazione fetale: la percentuale di DNA fetale presente all'interno del plasma materno è un fattore determinante per l'accuratezza del NIPT. Percentuali inferiori al 5-10% rendono il test poco affidabile

➔ determinazione del DNA fetale

- Coverage/profondità di sequenziamento: numero medio di letture per base. Maggiore il coverage, più accurato (e più costoso) il test



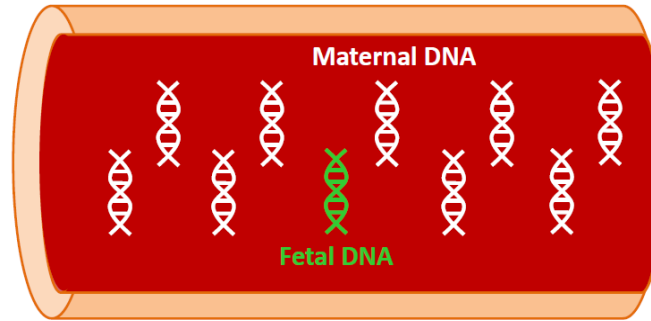
# Fattori chiave per NIPT

## Not Invasive Prenatal Testing

---

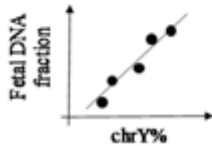
- Frazione fetale: la percentuale di DNA fetale presente all'interno del plasma materno è un fattore determinante per l'accuratezza del NIPT. Percentuali inferiori al 5-10% rendono il test poco affidabile
  - ➔ determinazione del DNA fetale
- Coverage/profondità di sequenziamento: numero medio di letture per base. Maggiore il coverage, più accurato (e più costoso) il test
- Analisi bioinformatica: lo sviluppo di nuove soluzioni per l'analisi dei dati consente di aumentare l'accuratezza del test
  - ➔ NIPT di seconda generazione

# Determinazione della frazione fetale



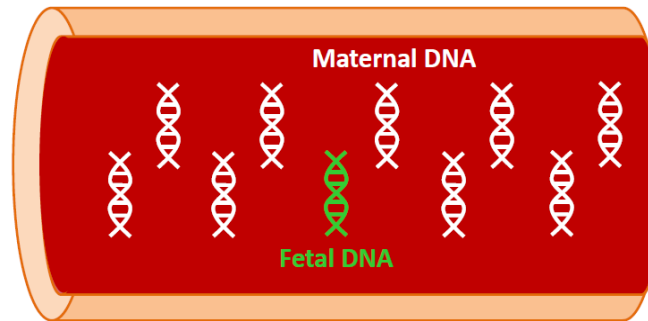
Approaches for fetal DNA  
fraction determination

## a. Y chromosome-based



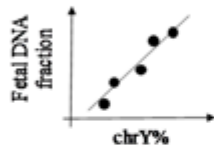
Proporzione di sequenze che allineano sul  
chr Y (solo per feti di sesso maschile)

# Determinazione della frazione fetale

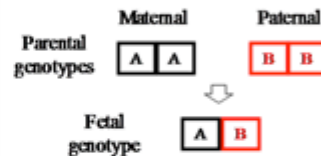


Approaches for fetal DNA fraction determination

## a. Y chromosome-based



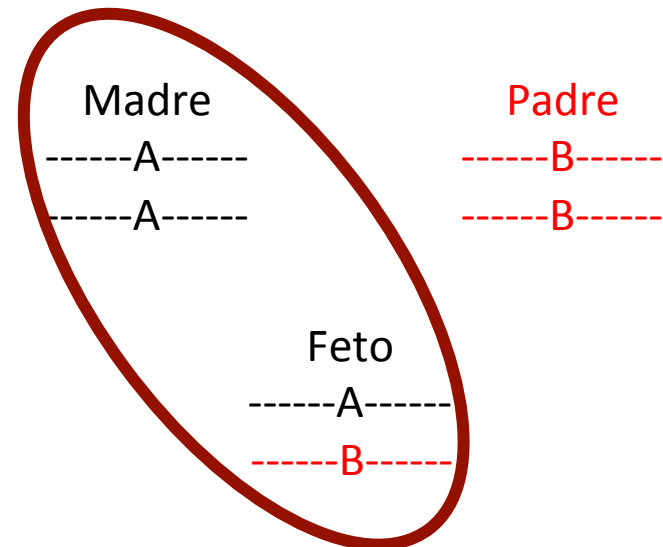
## b. SNP-based



$$\text{Fetal DNA fraction} = \frac{2\sum B}{\sum A + \sum B}$$

Extended versions:

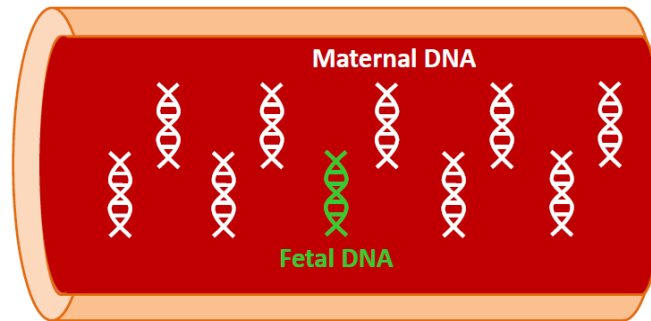
1. High-depth targeted sequencing (*FetalQuant*);
2. Shallow-depth sequencing, coupled with maternal genotypes (*FetalQuant<sup>SD</sup>*).



La frazione di alleli paterni fornisce una stima della frazione fetale



# Determinazione della frazione fetale



Approaches for fetal DNA fraction determination

**Metilazione del DNA:** modifica delle Citosine con impatto sulla trascrizione DNA → RNA.

La metilazione di una regione genomica varia da tessuto a tessuto, e può essere usata per la stima della frazione fetale.

## d. Methylation-based

### 1. Differentially methylated markers

Maternal cells

Four white circles representing unmethylated markers on a horizontal line.

Placental cells

Four black circles representing methylated markers on a horizontal line.

Maternal plasma

A circle containing a mixture of white and black circles representing the markers in maternal plasma.

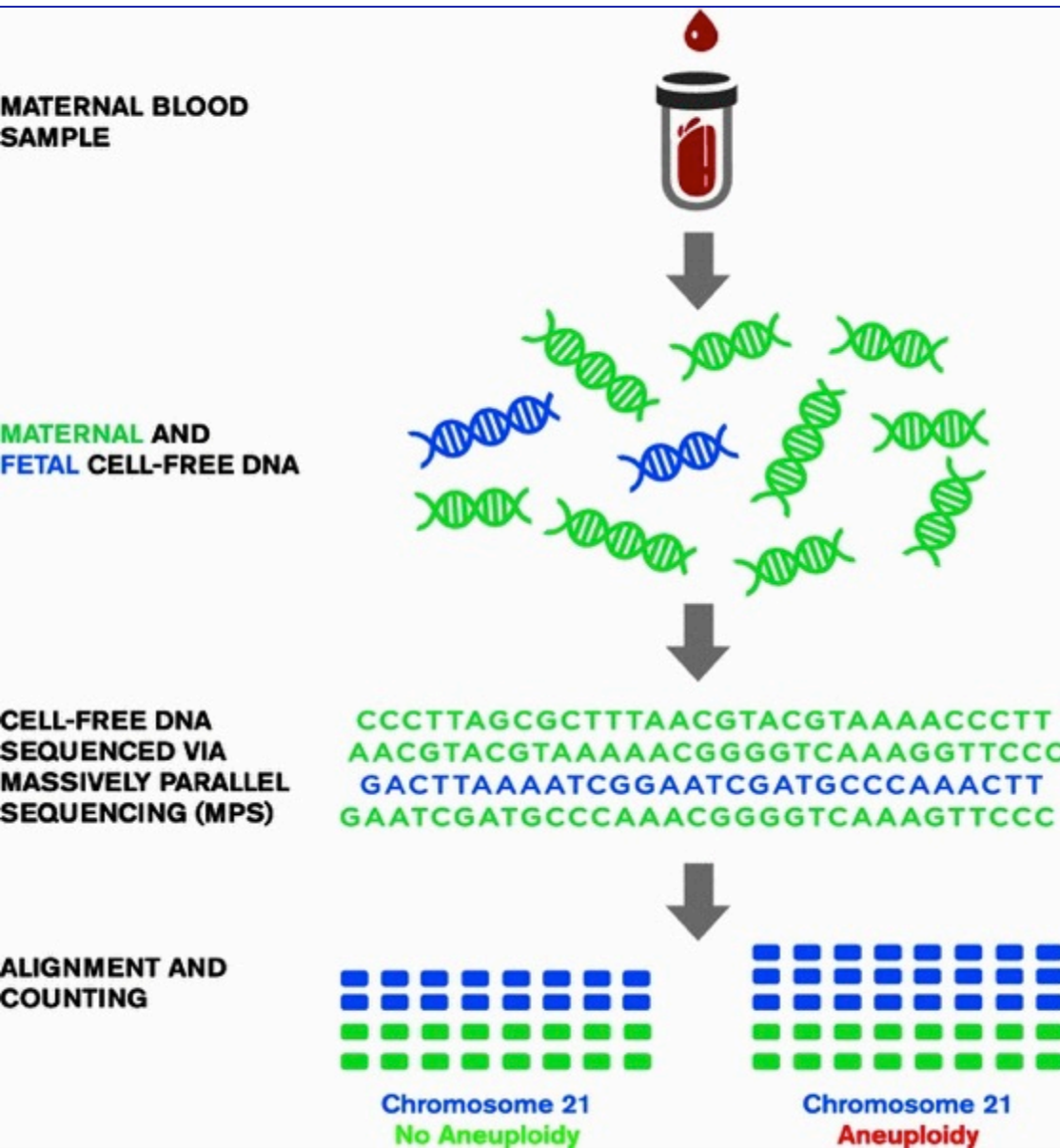
$$\text{Fetal DNA fraction} = \frac{\sum M}{\sum M + \sum U}$$

● M: methylated

○ U: unmethylated

### 2. Plasma DNA tissue mapping

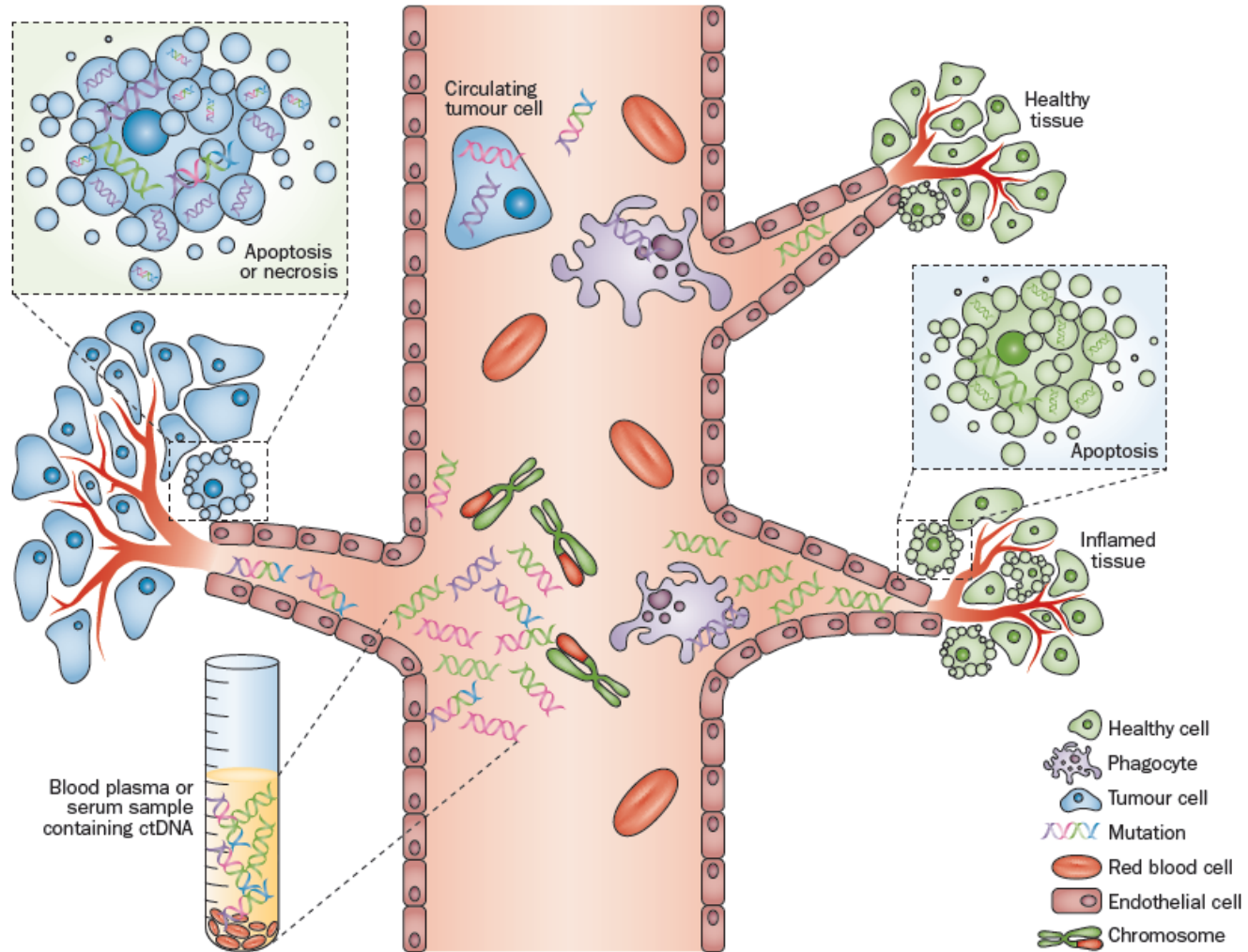
# NIPT per le aneuploidie



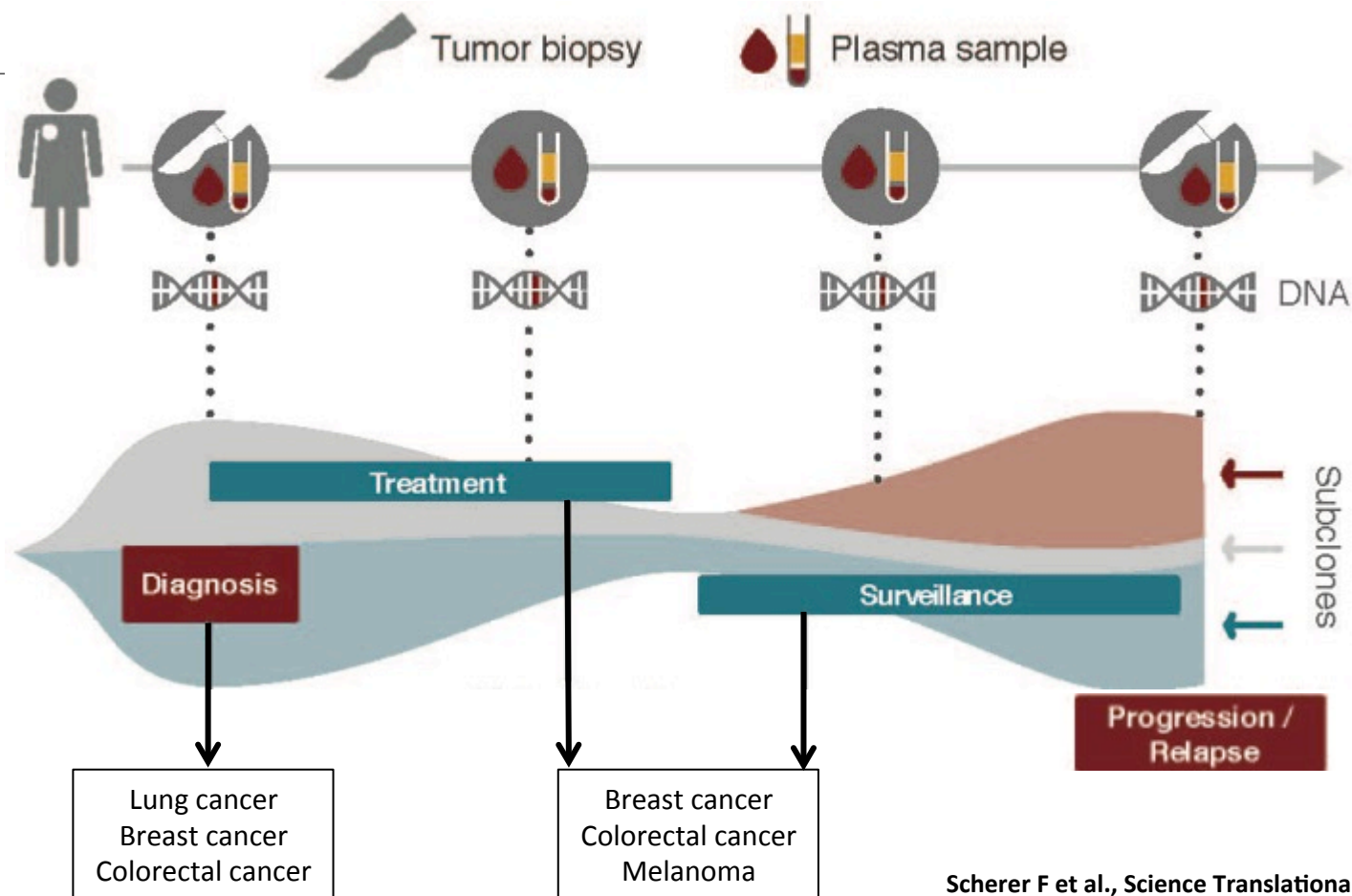
Conta delle sequenze

- Feto con trisomia contribuisce con un 50% in più di cfDNA dal cromosoma trisomico
- Milioni di reads necessarie per identificare lievi variazioni di cfDNA
- Non distingue il DNA materno dal fetale

# DNA tumorale circolante (ctDNA)



# Applicazioni cliniche della biopsia liquida

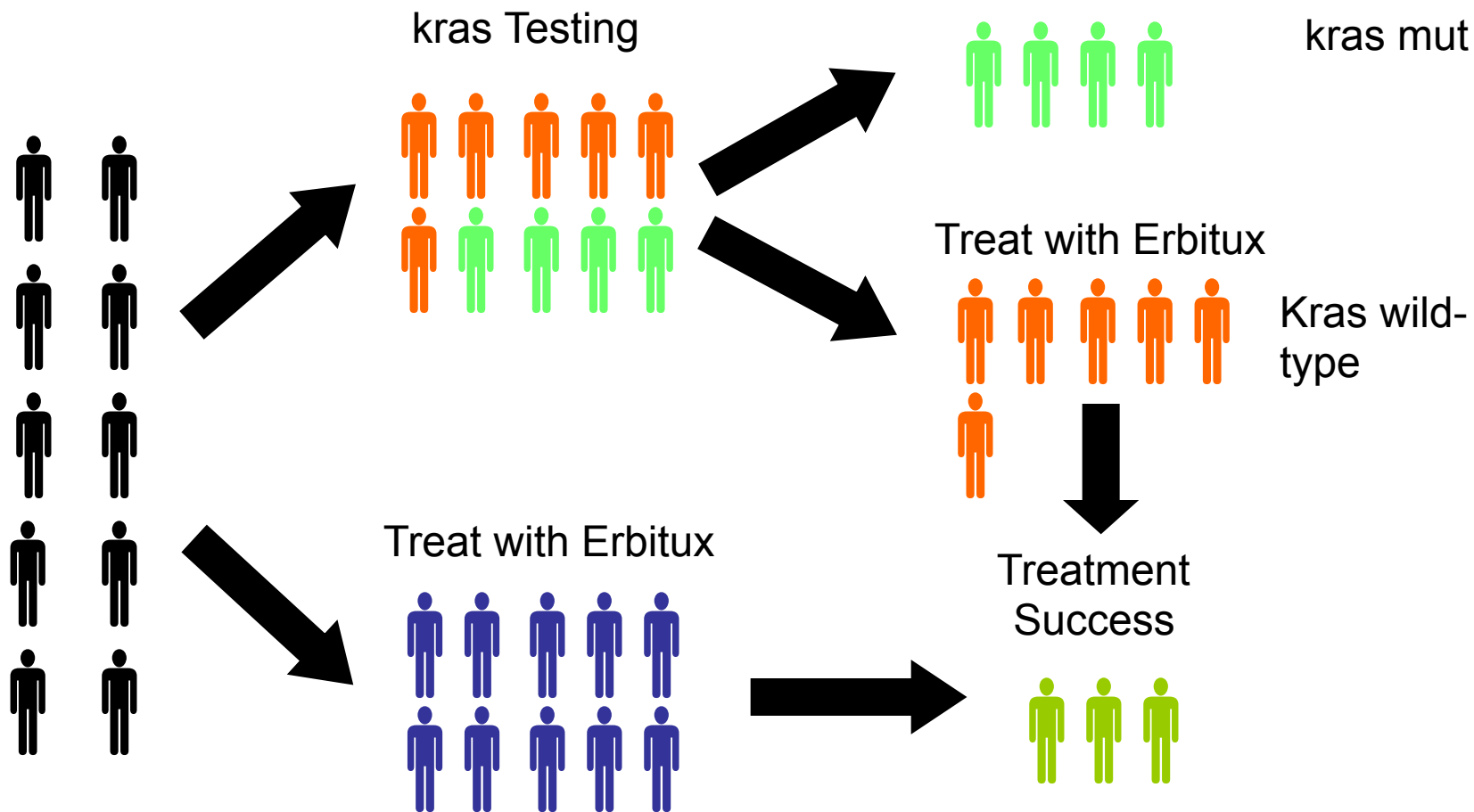


Circulating tumor DNA as a liquid biopsy for cancer, Heitzer et al., *Clin Chem*, 2015

Amplification of the MET receptor drives resistance to anti-EGFR therapies in colorectal cancer, Bardelli et al., *Cancer Discov.*, 2013

Detection of Circulating Tumor DNA in Early- and Late-Stage Human Malignancies, Bettgowda et al., *Sci Transl Med.*, 2014

# La conoscenza del profilo genomico consente di ottimizzare le strategie terapeutiche: Medicina di Precisione

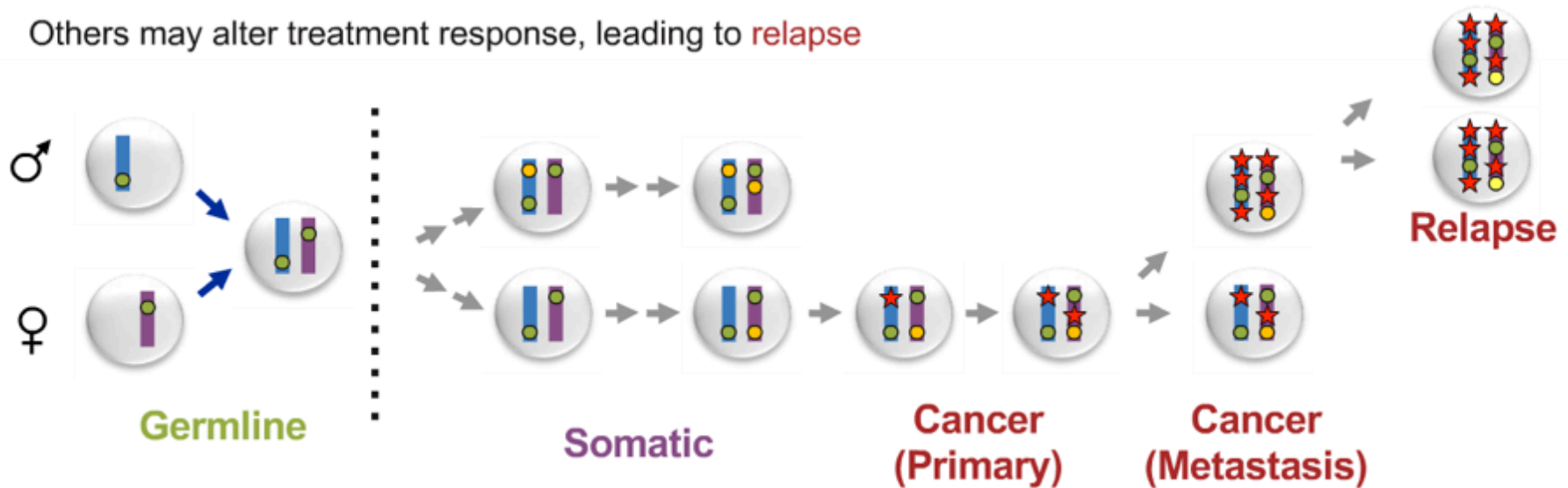


# NGS per lo studio dell'evoluzione *temporale* del tumore

- ▶ WGS is a **snapshot**
- ▶ Certain mutations reflect paternal and/or maternal **germline** variation
- ▶ Additional **somatic** mutations accumulate through life
- ▶ “**Driver**” mutations cause **cancer**, “*passenger*” mutations are carried along
- ▶ **Additional drivers evolve** and diversify the cancer
- ▶ Some alter aggressiveness...
- ▶ ...which may be **treatable**
- ▶ Others may alter treatment response, leading to **relapse**

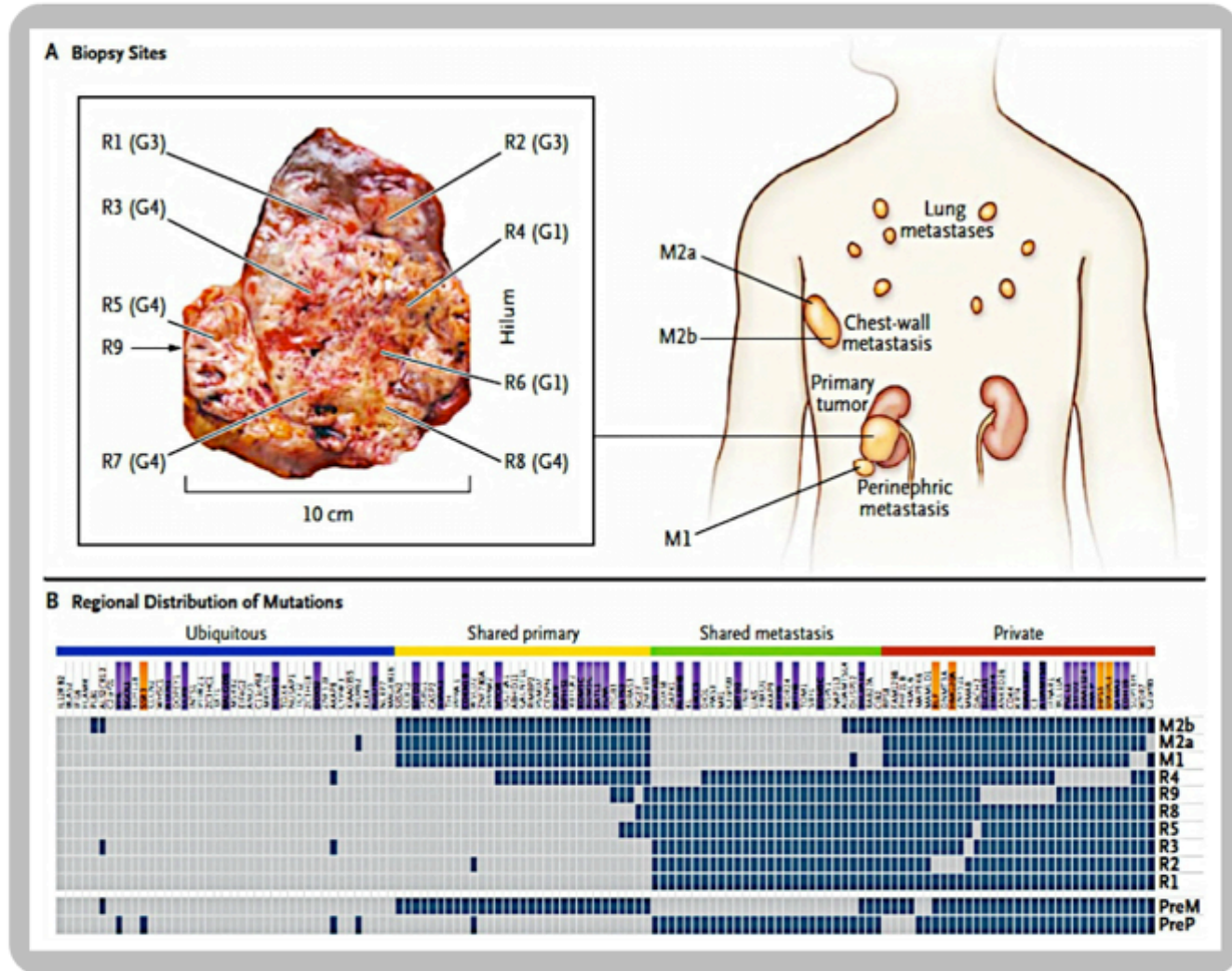
Cancer genomes are not static.

In cancer, one snapshot is not enough.





# NGS per lo studio dell'evoluzione *spaziale* del tumore: eterogeneità tumorale e metastatica





# Progetto NIASMIC

POR FESR 2014/2020 – “RICERCA SCIENTIFICA, SVILUPPO TECNOLOGICO E INNOVAZIONE” Azioni Cluster Top-Down

---

Sviluppo di un protocollo non invasivo di analisi del profilo genetico del DNA tumorale circolante (ctDNA), estratto da biopsia liquida, per un ampio spettro di tumori.

*Durata: 2018-2020*

## PARTNERS

- BGT-Inversol
  - Kitos
  - Nurex
  - Yenetics
  - Ardea
  - Microtec
- 
- Osp. Oncologico, Cagliari
  - Policlinico Gemelli, Roma



UNIONE EUROPEA  
Fondo europeo di sviluppo regionale



REPUBBLICA ITALIANA



REGIONE AUTÒNOMA DE SARDIGNA  
REGIONE AUTONOMA DELLA SARDEGNA



SARDEGNA  
RICERCHE

# Futuro del sequenziamento NGS

- **Sviluppo tecnologico:** rapida evoluzione della tecnologia; investimenti per tecnologie più efficienti: riduzione dei tempi e dei costi, aumento qualità dei risultati

## The GRAIL approach



GRAIL is poised to detect cancer early in asymptomatic individuals by combining high-intensity sequencing of unprecedented breadth and depth with the techniques of modern data science. Through one of the largest clinical trial programs ever pursued in genomic medicine, GRAIL will create vast datasets to develop our products and demonstrate their clinical utility.

*Costo per il sequenziamento di 1 genoma*

2000  
\$ 100 Milioni



2018  
\$ 1000



2030  
????



Grazie per l'attenzione