

Modulo 5

LA FASE ANALITICA

Il laboratorio di diagnostica molecolare in anatomia patologica:
La fase analitica: metodiche

La tecnologia della Polymerase Chain Reaction (PCR)

La tecnologia della **Polymerase Chain Reaction (PCR)** ha rivoluzionato l'attività dei laboratori di ricerca e di diagnostica, trovando applicazioni ed impieghi in svariati campi della medicina e della biologia.

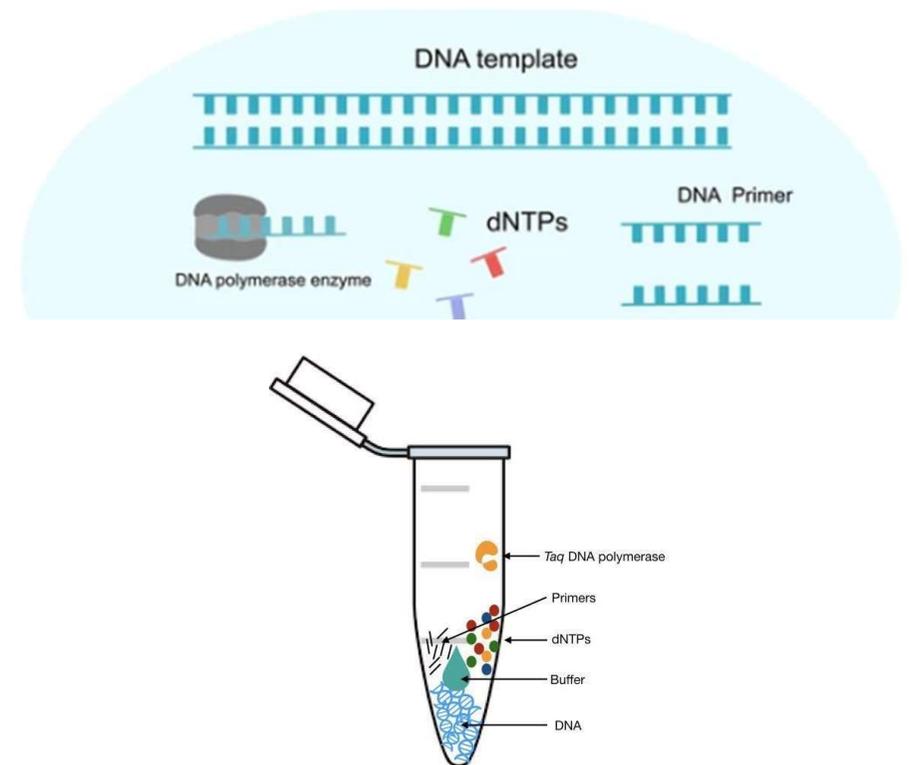
Infatti, viene utilizzata di routine non solo nei laboratori di **genetica, biochimica, microbiologia, oncologia molecolare e virologia**, ma anche in **medicina forense, in botanica** ed in **zoologia**, coprendo tutto il ventaglio delle discipline che hanno in qualche modo a che fare con le scienze della vita.

La PCR è una tecnica che consente di ottenere rapidamente milioni di molecole identiche di DNA a partire da quantità estremamente ridotte dell'acido nucleico → la PCR è una reazione di amplificazione in vitro di uno specifico frammento di DNA per mezzo dell'enzima DNA polimerasi-DNA dipendente.

La tecnologia della Polymerase Chain Reaction (PCR)

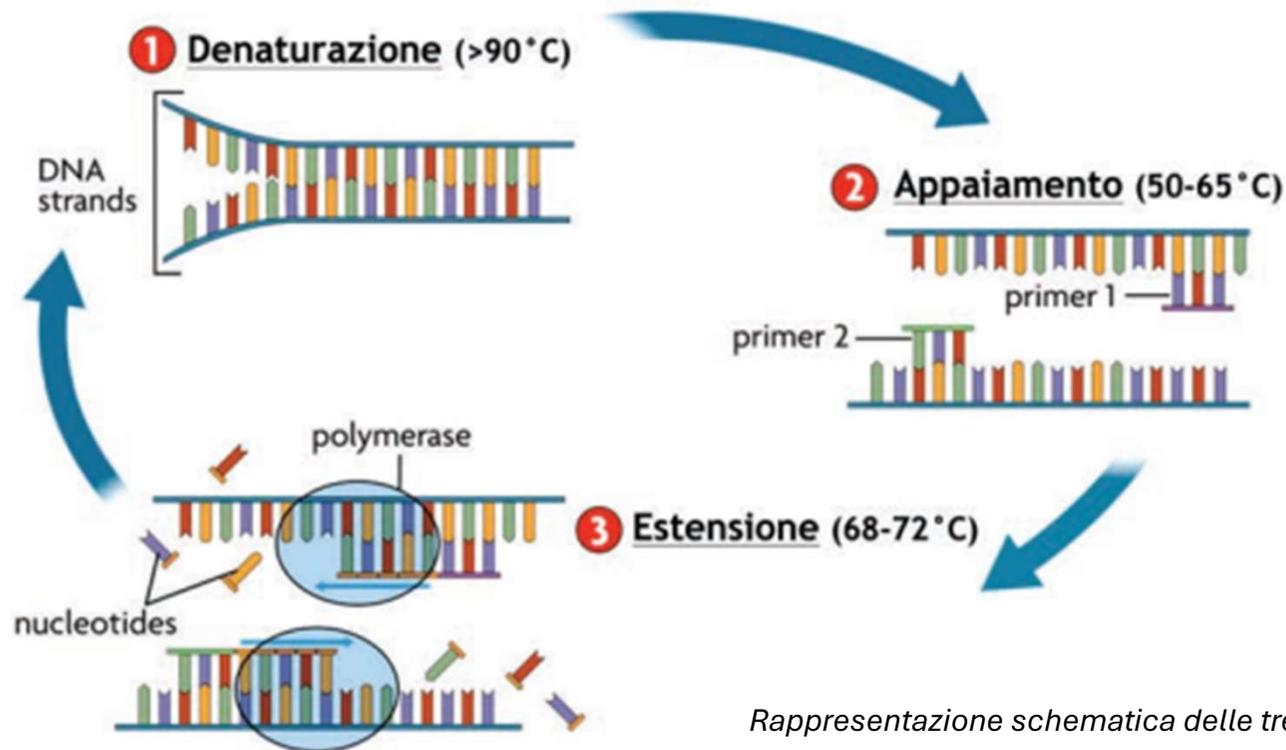
Un requisito indispensabile al realizzarsi della reazione è la **conoscenza delle sequenze alle estremità della regione bersaglio**. Infatti, nella reazione sono coinvolti due oligonucleotidi a singolo filamento (**primer**) complementari uno all'estremità 3' e l'altro all'estremità 5' del segmento di DNA che si vuole amplificare, che costituiscono gli elementi di innesco dell'attività della DNA polimerasi.

Altri elementi coinvolti nella reazione sono i **desossiribonucleotidi (dNTP)** ed il cloruro di magnesio ($MgCl_2$): i primi sono necessari per la sintesi delle nuove eliche ed il secondo rappresenta il cofattore indispensabile alla **DNA polimerasi**.



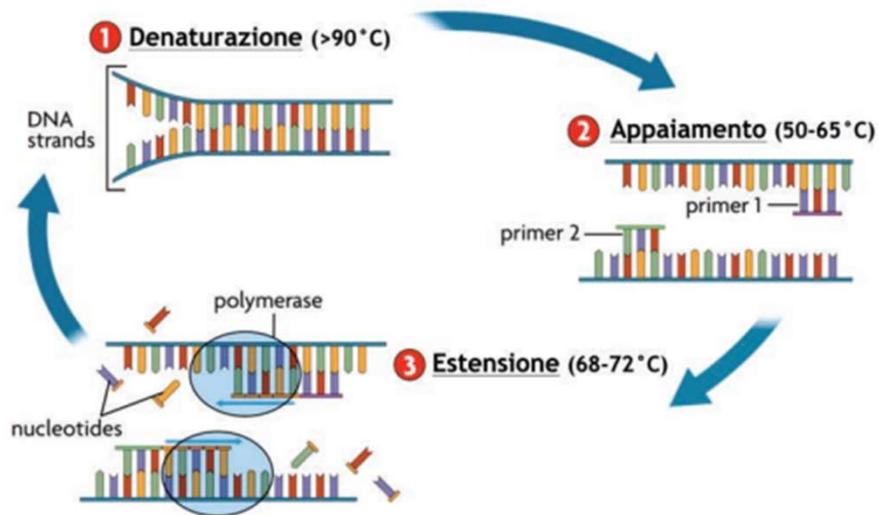
La tecnologia della Polymerase Chain Reaction (PCR)

LA REAZIONE POLIMERASICA A CATENA (PCR) La reazione prevede il succedersi di **cicli di amplificazione** durante i quali si alternano **tre diverse fasi** caratterizzate da specifiche temperature che rendono possibile rispettivamente:



Rappresentazione schematica delle tre fasi del ciclo di PCR

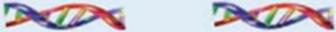
La tecnologia della Polymerase Chain Reaction (PCR)



Rappresentazione schematica delle tre fasi del ciclo di PCR

1. la **denaturazione** (denaturation) della doppia elica del DNA stampo in due singoli filamenti (a temperature superiori a 90°C) per rottura dei ponti idrogeno che tengono appaiate le basi complementari;
2. l'**appaiamento** (annealing) dei primer alle sequenze complementari di DNA a singola elica localizzate alle estremità del frammento bersaglio (ad una temperatura in genere compresa tra i 50 ed i 65°C);
3. l'**estensione** (elongation) dei primer ad opera dell'enzima DNA Polimerasi mediante aggiunta di nucleotidi nella direzione 5'-3' (ad una temperatura compresa tra i 68 e i 72 °C). L'enzima riconosce come innesco il sito in cui è avvenuta l'ibridazione dei primer con i filamenti di DNA complementare e aggiunge i nucleotidi liberi presenti nella soluzione all'estremità 3' dei primer. Si ottiene così la sintesi di una nuova elica complementare al DNA stampo.

La tecnologia della Polymerase Chain Reaction (PCR)

Numero di cicli	Numero di copie del bersaglio (ampliconi)	Bersaglio
1	2	
2	2 ² (4)	
3	2 ³ (8)	
4	2 ⁴ (16)	
5	2 ⁵ (32)
6	2 ⁶ (64)
20	2 ²⁰ (1.048.576)
30	2 ³⁰ (1.073.741.824)

Crescita esponenziale degli ampliconi nel corso dell'amplificazione.

Il numero di nuove molecole di DNA aumenta al succedersi di ogni ciclo.

durante il **primo ciclo** da un singola molecola di DNA si ottengono due molecole, ciascuna costituita da un'elica "originale" che ha fatto da stampo alla sintesi dell'elica "copia";



al **secondo ciclo** ciascuna delle due molecole si denatura, i quattro filamenti di DNA che si ottengono fanno da stampo per l'attività della DNA polimerasi ed alla fine le molecole di DNA diventano quattro. Il processo di amplificazione procede in questo modo, di ciclo in ciclo.



In teoria, quindi, ad ogni ciclo il numero di copie della sequenza bersaglio aumenta in maniera esponenziale; di fatto il numero di copie si duplica ad ogni ciclo fino al raggiungimento di un plateau

La tecnologia della Polymerase Chain Reaction (PCR)

Denaturazione del DNA

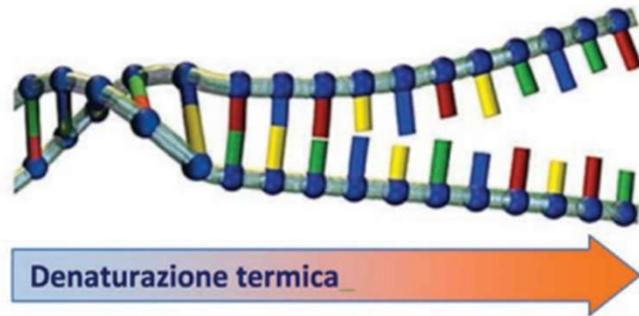
Il primo step nella reazione di PCR consiste nella denaturazione del DNA stampo a doppia elica.

In questa fase occorre ottimizzare due parametri essenziali:

1. la denaturazione ottimale del DNA stampo su cui si effettua la PCR;
2. il mantenimento delle attività della Taq DNA polimerasi.

Normalmente il DNA si trova nella classica conformazione a doppio filamento in cui i due filamenti (strands) del DNA sono tenuti assieme dai **legami a ponte di idrogeno** formati tra le basi azotate complementari (A=T; G=C). Il DNA deve essere portato ad una condizione di singola elica (single-stranded) in modo che successivamente si verifichi l'appaiamento (annealing) alle molecole di primer (anch'esse a singolo filamento)

La tecnologia della Polymerase Chain Reaction (PCR)



Dissociazione dei ponti idrogeno provocata dall'aumento della temperatura

La **rottura dei legami idrogeno** che tengono appaiati i due filamenti può essere ottenuta mediante **un aumento della temperatura**. La separazione provocata dall'instabilità dei legami tra le basi complementari viene indicata con il **termine di fusione**.

In particolare, si definisce “temperatura di fusione” (**melting temperature, T_m**) la temperatura in corrispondenza della quale **il 50% del DNA è svolto in forma di singolo filamento**.

Nel tampone di reazione in cui viene normalmente effettuata la reazione di PCR la temperatura di fusione è solitamente compresa tra 92 e 96°C. L'energia termica necessaria **dipende dalla composizione delle basi** (percentuale delle coppie di basi A=T e G=C), dalla **lunghezza del DNA** e dalla concentrazione salina dei tamponi utilizzati

La tecnologia della Polymerase Chain Reaction (PCR)

La **denaturazione viene favorita** dalla presenza di **concentrazioni saline relativamente alte** (circa 150 mM NaCl). Per ogni aumento della concentrazione dei cationi monovalenti pari ad 1 logaritmo corrisponde un aumento della temperatura di fusione pari a 16,6°C.

La **Taq DNA polimerasi ha solitamente una emivita di 30 min a 95 °C.**

→ Questo fatto limita il **numero di cicli della PCR** ed il **tempo di denaturazione del primo step.**

Infatti, considerando una incubazione di 1 min a 95 °C per ogni ciclo di PCR, il numero di cicli effettuabili non può essere superiore a 30-35. Diminuendo il tempo di denaturazione a 15-30 secondi i cicli di PCR possono solitamente essere aumentati fino a 45. È inoltre possibile ridurre la temperatura di denaturazione dopo i primi 10 cicli di PCR. Ad esempio, per ampliconi di lunghezza inferiore a 3 Kbp si può effettuare la denaturazione a 88 °C (per frammenti di DNA amplificati con meno del 50% di contenuto in G+C)

La tecnologia della Polymerase Chain Reaction (PCR)

Appaiamento dei primer (annealing)

Questo secondo step consiste nella programmazione della temperatura e del tempo di appaiamento dei primer.

La temperatura di annealing (T_a) dei primer dipende dal loro contenuto in G+C, dalla loro lunghezza e quindi dalla temperatura di fusione tra primer e l'elica complementare del DNA stampo.

Considerando primer di lunghezza media di 20 basi, una formula empirica spesso utilizzata per il calcolo della T_m è la seguente: $T_m = [4(G + C) + 2(A + T)]$ °C. Dove G, C, A e T indicano il numero di nucleotidi contenenti le basi azotate guanina, citosina, adenina o timina.

Nel caso che i due primer abbiano T_m diverse generalmente si considera quello con la T_m più bassa. Solitamente si utilizza come temperatura di annealing la formula “ $T_m - 5$ °C”, anche se spesso l'utilizzo diretto della stessa T_m può portare ad avere ottime rese nella reazione di PCR.

La tecnologia della Polymerase Chain Reaction (PCR)

Nel mettere a punto le reazioni di PCR si possono seguire essenzialmente due tipi di criteri riguardo alla T_a :

1. T_a costante durante i cicli → più utilizzata;
2. T_a che diminuisce ciclo dopo ciclo (touch-down)

La strategia di reazione touch-down permette di rendere i primi cicli di PCR estremamente “stringenti”, cioè tali da promuovere l’amplificazione solo di frammenti specifici rendendo instabili eventuali annealing dei primer a sequenze di DNA non perfettamente complementari.



una T_a troppo bassa porta all’annealing dei primer a sequenze non esattamente complementari e quindi all’amplificazione di frammenti non specifici, mentre una T_a troppo alta può ridurre la resa in quanto solo una frazione delle molecole del primer riesce ad innescare la polimerizzazione, a causa dell’elevata instabilità del loro appaiamento con il DNA stampo.

Il tempo di annealing infine non deve essere troppo lungo (in modo da sfavorire appaiamenti a stampi con bassa complementarità); di solito si utilizzano tempi dell’ordine di 30 secondi o meno.

La tecnologia della Polymerase Chain Reaction (PCR)

Estensione dei primer (elongation)

Il terzo step della reazione di PCR è legato alla programmazione della temperatura e del tempo di estensione dei primer. **La temperatura utilizzata è solitamente compresa tra 68 e 72 °C.**

La Taq DNA polimerasi ha un'attività specifica a 37 °C molto simile a quella del frammento di Klenow della DNA polimerasi I di E. coli. Tuttavia l'attività della Taq DNA polimerasi presenta la **massima processività a circa 70°C** e l'estensione dei primer avviene ad una velocità di circa 100 basi/sec. Generalmente 1 minuto è sufficiente per amplificare con una buona resa stampi lunghi circa 1 Kbp. Il tempo di estensione viene quindi calibrato sulla **lunghezza dello stampo da amplificare.**

La tecnologia della Polymerase Chain Reaction (PCR)

Tampone di reazione

Un tampone di reazione (buffer) standard contiene:

- 10-50 mM Tris-HCl pH 8.3,
- fino a 50 mM KCl,
- 1.5 – 4 mM MgCl₂,
- 0.2 – 1 M di ciascun primer,
- 50-200 M di ciascun deossiribonucleotide (dNTP),
- gelatina o BSA (siero albumina bovina) fino a 100 g/ml.

In alcuni casi vengono aggiunti **detergenti** non ionici come Tween-20, Nonidet P-40 o Triton X-100 (0.05-0.10% v/v) che agiscono **stabilizzando la polimerasi, riducono la formazione di strutture secondarie**, ma che possono favorire amplificazioni non specifiche.

Il tampone di reazione viene sempre fornito insieme all'enzima. In alcuni casi il produttore della Taq DNA polimerasi mette in commercio sia tamponi di reazione contenenti MgCl₂ sia privi di esso. Nel caso si debba mettere a punto una PCR è di norma preferibile avere a disposizione il tampone di reazione privo di MgCl₂, ed aggiungere il sale separatamente.

La tecnologia della Polymerase Chain Reaction (PCR)

Fattore	ELEVATA STRINGENZA	BASSA STRINGENZA
Temperatura di annealing		
MgCl ₂		

Effetto del cloruro di magnesio (MgCl₂) e della temperatura di annealing sulla stringenza (specificità) di una reazione di PCR. Le frecce in alto e in basso indicano rispettivamente un aumento o una diminuzione della Temperatura di

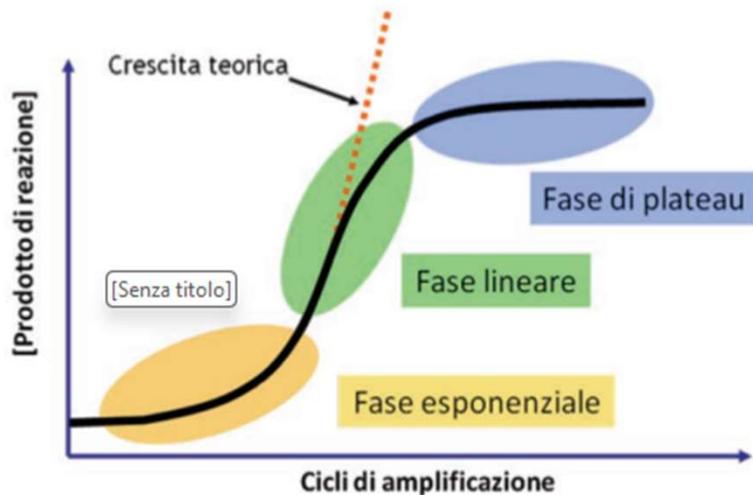
Infatti la concentrazione degli ioni Mg²⁺ ha due ruoli:

1. è un **cofattore** essenziale per la Taq DNA polimerasi per il caricamento dei nucleotidi,
2. la concentrazione di Mg²⁺ **influenza l'appaiamento dei primer** allo stampo. → Maggiore è la concentrazione di MgCl₂, minore è la specificità dell'appaiamento. La variazione delle concentrazioni di MgCl₂ gioca quindi un ruolo simile (ma opposto) a quello della temperatura di annealing e spesso la messa a punto di un'elevata (o bassa) stringenza (specificità) nella PCR è basata tanto sulla Ta quanto sulla concentrazione di MgCl₂.

La tecnologia della Polymerase Chain Reaction (PCR)

Numero di cicli

In generale il numero di cicli è compreso tra 30 e 45. Il numero di cicli di amplificazione necessari ad ottenere una banda visibile su gel di agarosio può dipendere in gran parte dalla concentrazione di DNA iniziale.



“Effetto Plateau” della reazione di PCR causato dalla degradazione dei reagenti e dall’accumulo di prodotti.

Un’amplificazione PCR può essere suddivisa in tre fasi

1. esponenziale: dopo ogni ciclo si verifica un’esatta duplicazione del prodotto accumulato (efficienza di reazione pari al 100%). La reazione segue la formula:

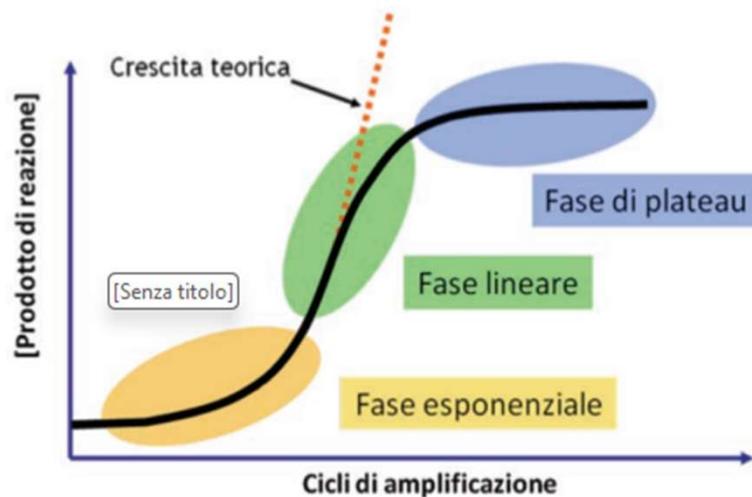
$$a = t \times 2^n$$

a. quantità di amplicone prodotto
t. numero di molecole target iniziali
n. numero di cicli eseguiti

La tecnologia della Polymerase Chain Reaction (PCR)

Numero di cicli

In generale il numero di cicli è compreso tra 30 e 45. Il numero di cicli di amplificazione necessari ad ottenere una banda visibile su gel di agarosio può dipendere in gran parte dalla concentrazione di DNA iniziale.



“Effetto Plateau” della reazione di PCR causato dalla degradazione dei reagenti e dall’accumulo di prodotti.

2. lineare: continua la crescita esponenziale come **fase logaritmica lineare**

Tuttavia, si assiste ad una **graduale riduzione della disponibilità dei reagenti** e dell’efficienza del sistema che porta ad un rallentamento della reazione.

3. plateau: nelle fasi tardive dell’amplificazione il tasso di accumulo di prodotto diminuisce drasticamente a causa di numerosi fattori tra cui la degradazione dei reagenti (dNTP, DNA polimerasi) e l’ inibizione da parte del pirofosfato accumulato (inibizione da prodotto).

La tecnologia della Polymerase Chain Reaction (PCR)

ADDITIVI

Nel caso di sequenze da amplificare **ricche in G+C** può rendersi talvolta necessario aggiungere alla miscela di reazione delle molecole in grado di **favorire la denaturazione della doppia elica del DNA**. Le molecole più frequentemente utilizzate sono:

- dimetilsolfossido (DMSO) (2-10%)
- betaina (0.5-2 M)
- dimetilformamide (DMF)
- urea
- formammide (1-5%)
- glicerolo (in caso di sequenze ricche di strutture secondarie, inoltre protegge l'enzima dalle elevate temperature)
- sieralbumina bovina (BSA) (fino a 0,8 g/ l) in presenza dell'inibitore melanina
- polietilenglicole (PEG)

Il polietilenglicole (PEG) può essere utilizzato per amplificare **DNA in concentrazione molto bassa**, in quanto promuove l'associazione macromolecolare tra DNA e DNA polimerasi attraverso l'esclusione del solvente

La tecnologia della Polymerase Chain Reaction (PCR)

DISEGNO DEI PRIMER

:

Per il disegno dei primer e delle coppie di primer spesso vengono utilizzati dei programmi gratuiti. Alcuni esempi:

- Primer3 URL: <http://biotools.umassmed.edu/bioapps/primer3>; <http://www.cgi>; oppure <http://frodo.wi.mit.edu/>;
- GeneFisher URL: <http://bibiserv.techfak.uni-bielefeld.de/genefisher2>;
- Oligonucleotide calculator URL: <http://www.basic.northwestern.edu/biotools/oligocalc.html>;

La tecnologia della Polymerase Chain Reaction (PCR)

DISEGNO DEI PRIMER

- **Lunghezza dei primer** La lunghezza ottimale di un primer dipende sia dal suo contenuto in A+T, sufficientemente basso da poter avere T_m (e quindi T_a) superiori a 50°C , sia dalla composizione nucleotidica della sua sequenza, in modo che la probabilità di avere siti di annealing diversi da quello voluto sia estremamente bassa (alta specificità).
- **Contenuto in G+C** Il contenuto in G+C determina la T_m , quindi occorre evitare sia valori troppo bassi (bassa T_m) sia valori troppo alti (alta T_m). Solitamente il contenuto in G+C si aggira tra il 40 e il 60%. Si preferisce, in alcuni casi, per aumentare la specificità dell'amplificazione, disegnare primer in cui le ultime basi al terminale 3' abbiano G o C (il terminale 3' è quello che porta all'innescio effettivo della polimerizzazione).
- **Sequenze ripetute** La presenza di regioni a bassa complessità, cioè di sequenze ripetute (sia di singoli nucleotidi sia di regioni di- o tri-nucleotidiche) all'interno di un primer deve essere evitata in quanto può portare ad uno slittamento dell'appaiamento o all'annealing su siti aspecifici (occorre ricordare che i genomi eucarioti sono ricchi di sequenze a bassa complessità) e quindi all'amplificazione di prodotti aspecifici

La tecnologia della Polymerase Chain Reaction (PCR)

DISEGNO DEI PRIMER

Strutture secondarie e formazione di dimeri

La presenza di **strutture secondarie intramolecolari o intermolecolari** può determinare una **diminuzione della resa di amplificazione** o addirittura un'assenza di amplificazione. Le **strutture secondarie** infatti **competono per l'annealing del primer** con lo stampo sulla sequenza di DNA bersaglio, diminuendo drasticamente la concentrazione effettiva di primer disponibile per la reazione di amplificazione. Le strutture secondarie si possono classificare come "hairpin" (forcine), "self dimer" (autodimeri) e "cross dimer" (dimeri crociati):



1. **Hairpin**: le strutture ad hairpin si formano per **l'interazione intramolecolare** tra i nucleotidi del singolo primer. Gli hairpin al terminale 3' sono i meno tollerati in quanto sequestrano direttamente il residuo ossidrilico necessario alla polimerizzazione.
2. **Self dimer**: si formano per **l'interazione intermolecolare** tra due molecole di primer dello stesso tipo nei punti in cui il primer è omologo a se stesso.
3. **Cross dimer**: si formano tra due molecole di primer di tipo diverso nelle regioni di omologia di sequenza

La tecnologia della Polymerase Chain Reaction (PCR)

Controlli Interni Di Reazione

In tutte le tecniche laboratoristiche è necessario allestire un **controllo positivo (C+)** ed un **controllo negativo (C-)**.

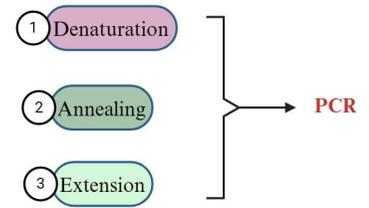
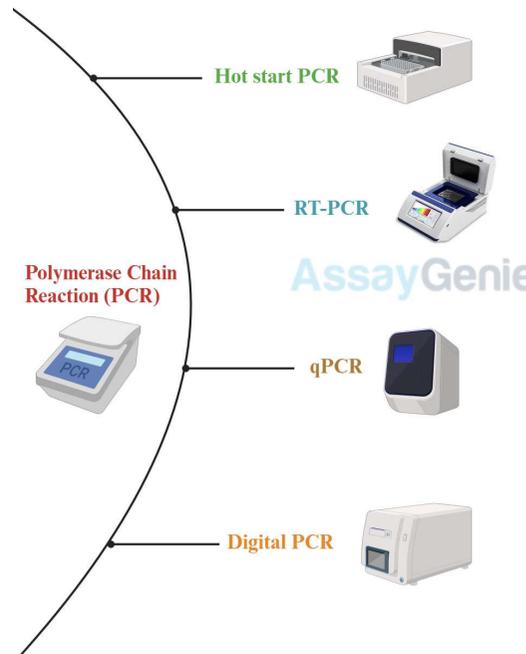


controllo positivo → La positività di questo campione attesta la corretta esecuzione della reazione di amplificazione. Il controllo positivo è costituito da un campione di DNA/RNA contenente la sequenza bersaglio. Sono sufficienti poche molecole di target per ottenere un segnale positivo dopo l'amplificazione. E' preferibile che il C+ non abbia troppe copie bersaglio, al fine di evitare che costituisca una pericolosa fonte di contaminazione (generando quantità enormi e non necessarie di sequenze di DNA)

controllo negativo → La negatività di questo campione attesta l'assenza di contaminazione della miscela di amplificazione. Il controllo negativo può essere costituito da un campione sicuramente negativo o in alternativa da acqua distillata. Per controllare la specificità, può essere utile includere anche controlli contenenti sequenze di DNA che non corrispondono alla sequenza da amplificare

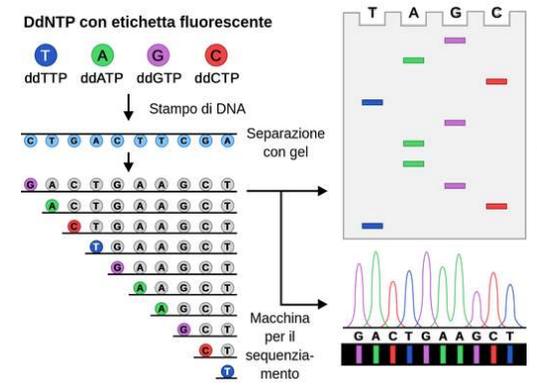
Dalla PCR a

PCR: Types and Applications

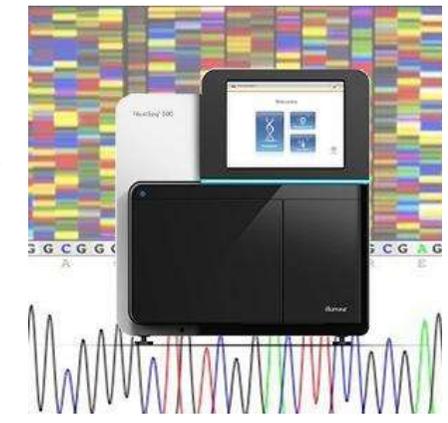


- Applications**
- **Diagnosis:** Detecting infectious diseases
 - **Research:** Amplifying DNA for cloning
 - **Forensics:** Analyzing DNA in criminal investigations and paternity testing.
 - **Drug Development:** Screening for drug targets, drug efficacy, and pharmacogenomics.
 - **Environmental Monitoring:** Identifying microorganisms in soil, water, and air samples.

Sequenziamento Sanger



Sequenziamento NGS



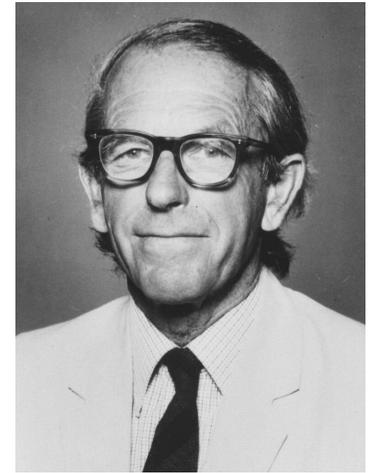
Il Sequenziamento Sanger

Sequenziare significa determinare l'ordine esatto dei monomeri che formano una biomolecola. Nel caso del DNA, il sequenziamento permette di ricostruire l'ordine dei nucleotidi (A, C, G e T) che formano una molecola di acido desossiribonucleico

CHE COS'È IL SEQUENZIAMENTO SECONDO SANGER?

Il sequenziamento secondo Sanger, detto anche “metodo di terminazione della catena”, è una tecnica che permette di determinare la sequenza nucleotidica di un DNA. Il metodo prende il nome dal suo ideatore, il due volte premio Nobel Frederick Sanger, che lo sviluppò insieme ai suoi colleghi nel 1977.

1° prova di sequenziamento: (1965) per sequenziamento di 76 nucleotidi
→ tempo impiegato circa 3 anni

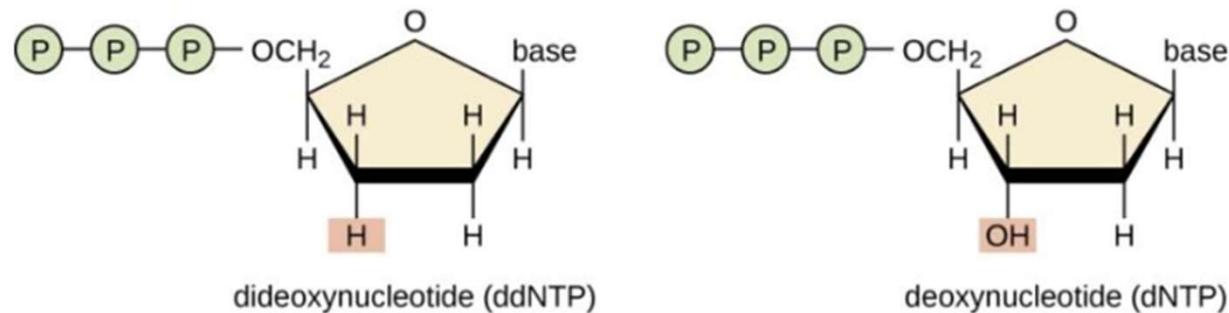


Frederick Sanger

Il Sequenziamento Sanger

Funzionamento del Sequenziamento Sanger

Il metodo di Sanger o **metodo della terminazione della catena** consiste nell'utilizzo di nucleotidi (dNTP) modificati chiamati **dideozinucleotidi (ddNTPs)**. I ddNTPs si differenziano dai dNTP per l'assenza del gruppo idrossilico sul carbonio 3' della molecola.



A destra: deossinucleotide trifosfato. A sinistra: dideozinucleotide trifosfato

Quando un dideoziribonucleotide viene aggiunto alla catena da una **DNA polimerasi** ad alta processività questo comporta la cessazione dell'allungamento della catena perché non può avvenire la formazione del legame fosfodiesterico.

Il Sequenziamento Sanger

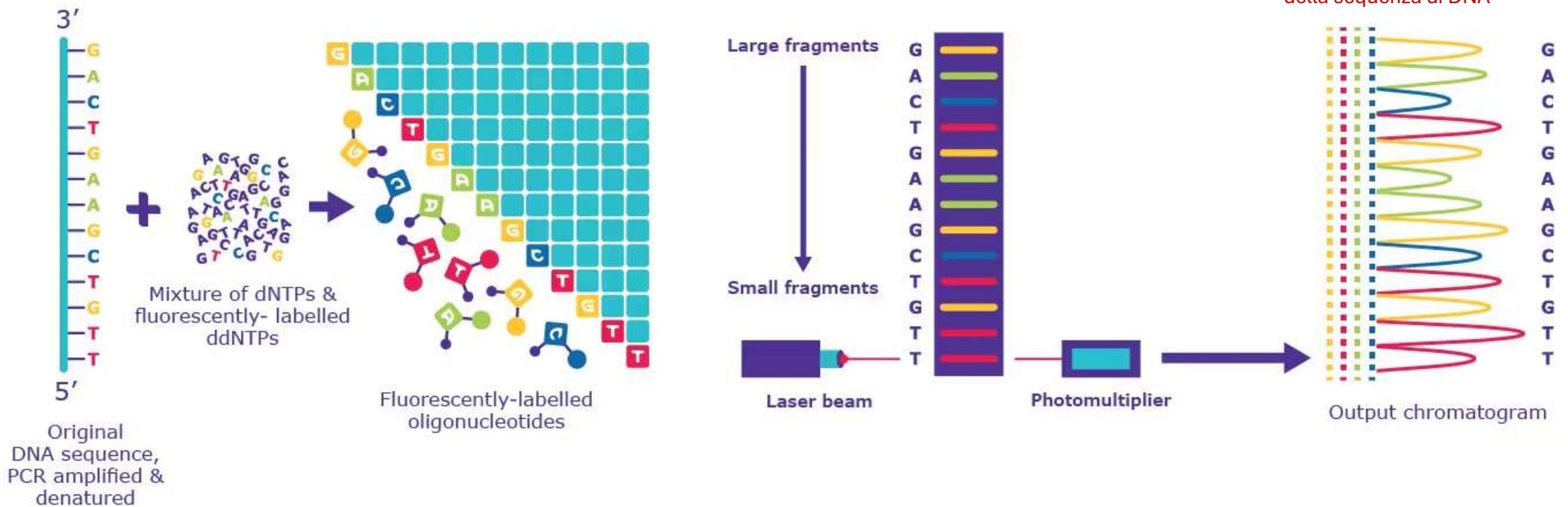
PASSAGGI DEL SEQUenziAMENTO SECONDO SANGER

Per ottenere il sequenziamento del DNA con il metodo di Sanger i passaggi principali sono tre.

1. PCR Con Terminazione Della Catena

2. Elettroforesi su gel per la separazione su base dimensionale

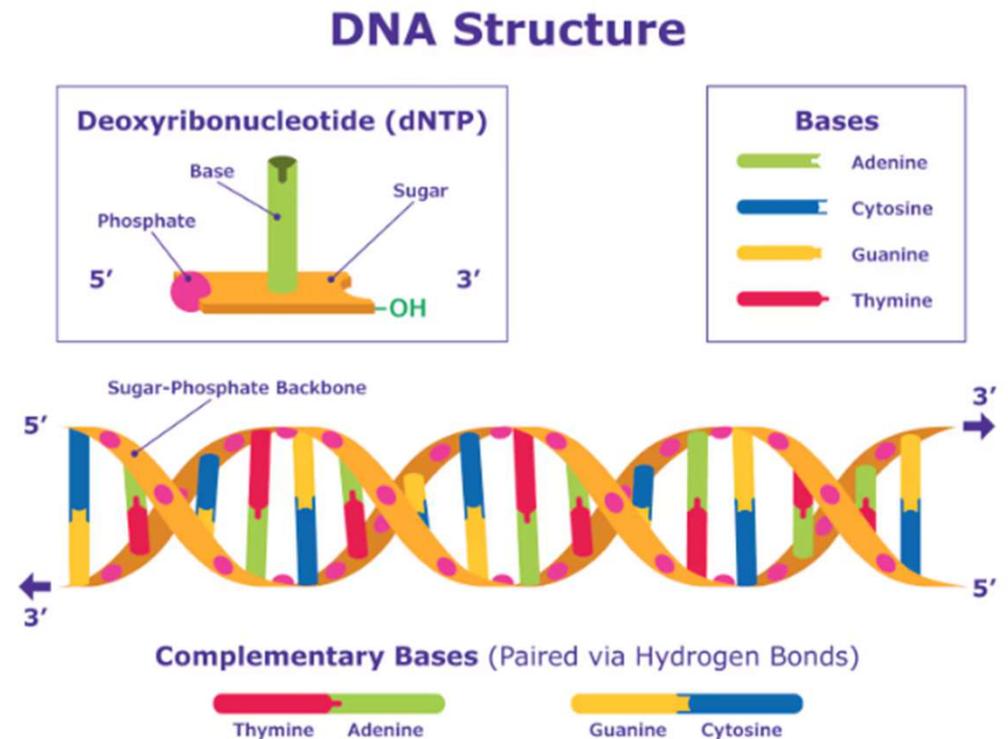
3. Analisi del gel e determinazione della sequenza di DNA



Il Sequenziamento Sanger

1. PCR CON TERMINAZIONE DELLA CATENA

La sequenza di DNA di interesse viene usata come stampo per la **PCR con terminazione della catena**, che funziona come una PCR standard tranne che per una differenza fondamentale: **l'aggiunta di nucleotidi (dNTP) modificati chiamati didesossiribonucleotidi (ddNTP)**. Durante la fase di estensione della PCR standard, la DNA polimerasi aggiunge i dNTP al filamento di DNA crescente catalizzando la formazione di un legame fosfodiesterico tra il gruppo 3'-OH libero dell'ultimo nucleotide e il gruppo 5'-fosfato del successivo.



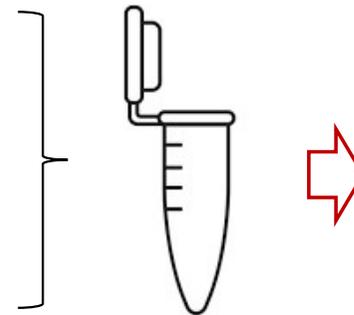
Schematizzazione della struttura del DNA. Il DNA è una molecola composta da due filamenti che si avvolgono l'uno sull'altro formando una doppia elica. Ogni filamento è composto da una stringa di molecole chiamate desossiribonucleotidi (dNTP).

Il Sequenziamento Sanger

1. PCR CON TERMINAZIONE DELLA CATENA

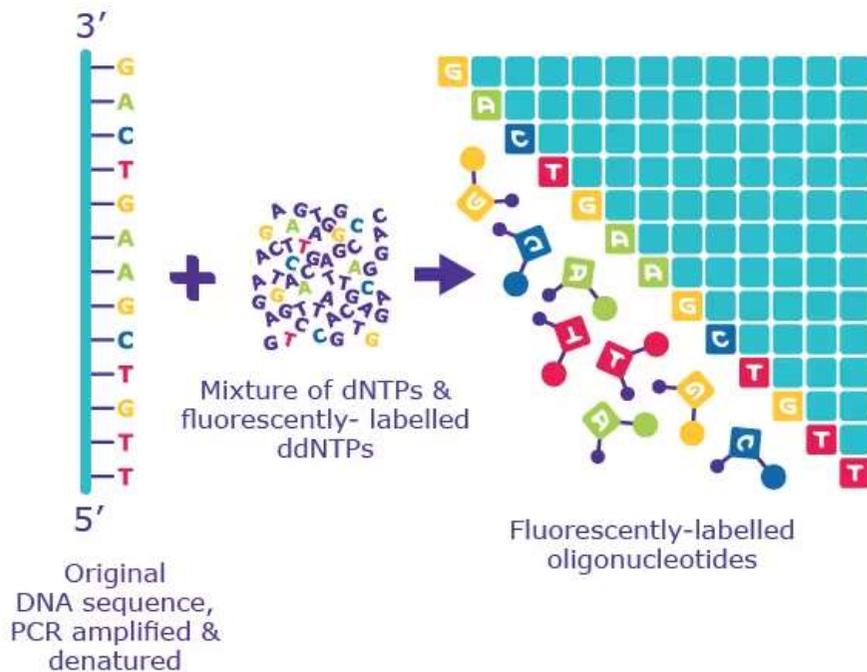
Nella PCR con terminazione della catena, si utilizza una miscela costituita dai **normali dNTP** insieme a una **bassa concentrazione di ddNTP**. Poiché i ddNTP sono privi del gruppo 3'-OH necessario per la formazione de legame fosfodiesterico, quando la DNA polimerasi incorpora casualmente un ddNTP, l'estensione cessa.

- DNA stampo
- un primer per iniziare la reazione di polimerizzazione
- enzima Taq-DNA polimerasi
- Deossinucleotidi (dATP, dCTP, dGTP, dTTP)
- Dideossinucleotidi (ddNTPs)



Il Sequenziamento Sanger

1. PCR CON TERMINAZIONE DELLA CATENA



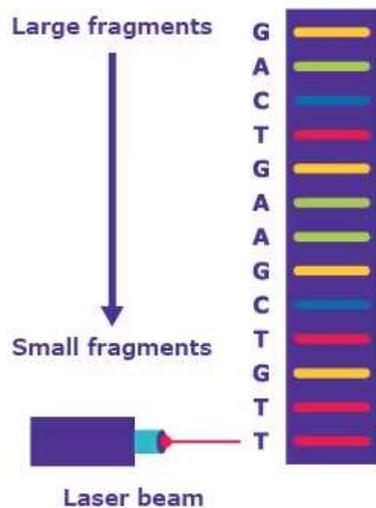
Il risultato di questo tipo di PCR è rappresentato da una miscela di oligonucleotidi che sono copie della sequenza di DNA di interesse, ma che **terminano a una lunghezza casuale** (n) con un 5'-ddNTP.

Nel sequenziamento con metodo Sanger di tipo **manuale**, si allestiscono 4 reazioni di PCR, ciascuna con l'aggiunta di un solo tipo di ddNTP (ddATP, ddTTP, ddGTP e ddCTP).

Invece, nel sequenziamento **automatico** si aggiungono tutti i ddNTP nella stessa miscela di reazione, ognuno **legato a un marcatore fluorescente diverso**.

Il Sequenziamento Sanger

2. ELETTROFORESI SU GEL PER LA SEPARAZIONE SU BASE DIMENSIONALE



Nel secondo passaggio, **gli oligonucleotidi a catena tronca vengono separati in base alle dimensioni grazie all'elettroforesi su gel**. I campioni vengono perciò caricati a un'estremità della matrice di gel, a cui viene poi applicata una corrente elettrica. Il **DNA è carico negativamente**, quindi gli oligonucleotidi saranno attratti verso l'elettrodo positivo che si trova all'estremità opposta del gel.

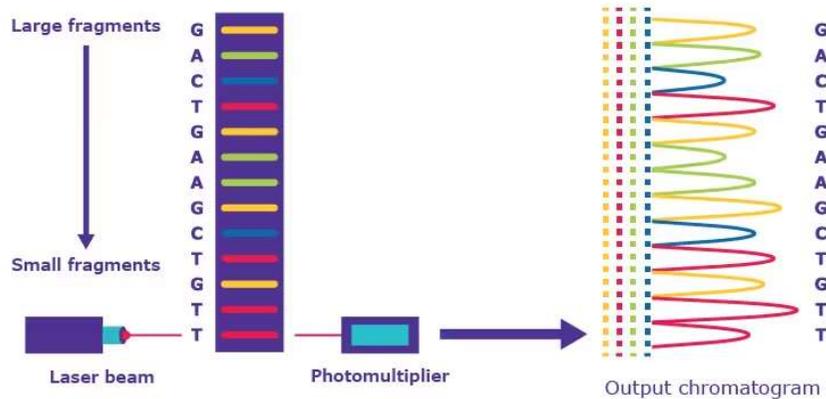
Dato che tutti i frammenti di DNA presentano la stessa carica per unità di massa, la **velocità** alla quale si muovono gli oligonucleotidi **dipende solo dalle loro dimensioni**. Più il frammento è piccolo, meno resistenza incontrerà spostandosi nel gel e quindi più velocemente avanzerà, con il risultato che i frammenti si ordineranno dai più piccoli ai più grandi leggendo il gel dal basso verso l'alto.

Nel sequenziamento **manuale**, gli oligonucleotidi ottenuti da ciascuna delle quattro reazioni di PCR vengono aggiunti a quattro diverse corsie di un gel, così è noto quali oligonucleotidi corrispondono a ciascun ddNTP.

Nel sequenziamento **automatico**, tutti gli oligonucleotidi vengono sottoposti a corsa elettroforetica su un unico gel capillare all'interno dell'apparecchiatura di sequenziamento.

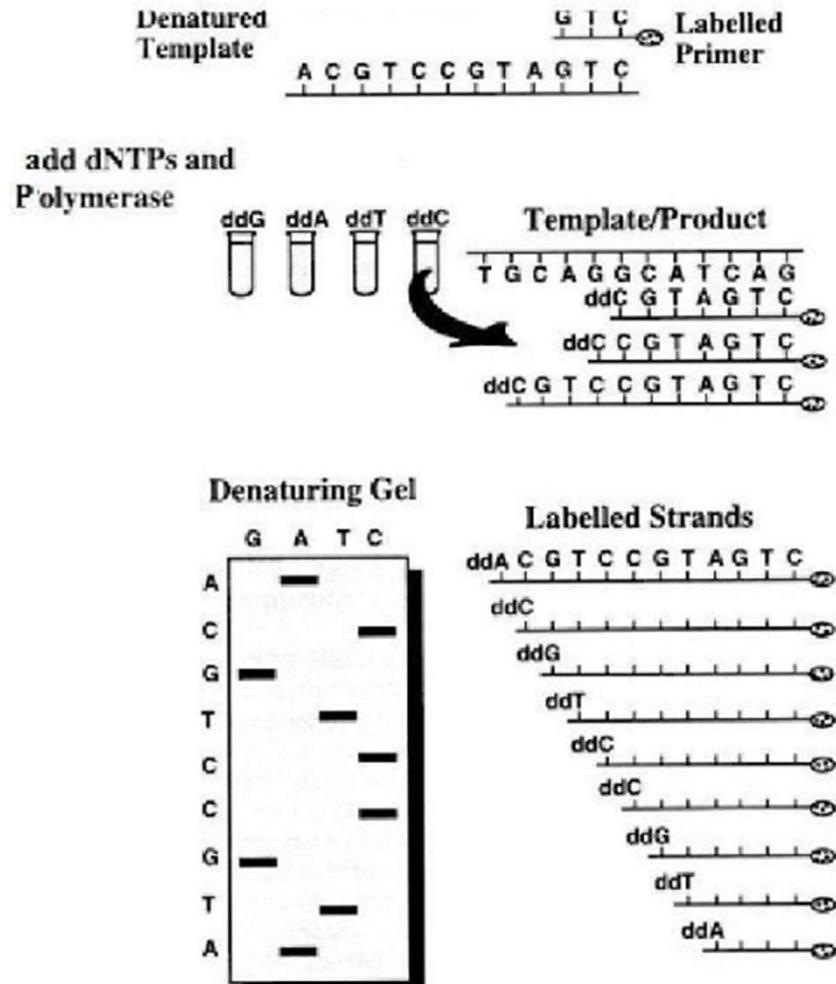
Il Sequenziamento Sanger

3. ANALISI DEL GEL E DETERMINAZIONE DELLA SEQUENZA DI DNA



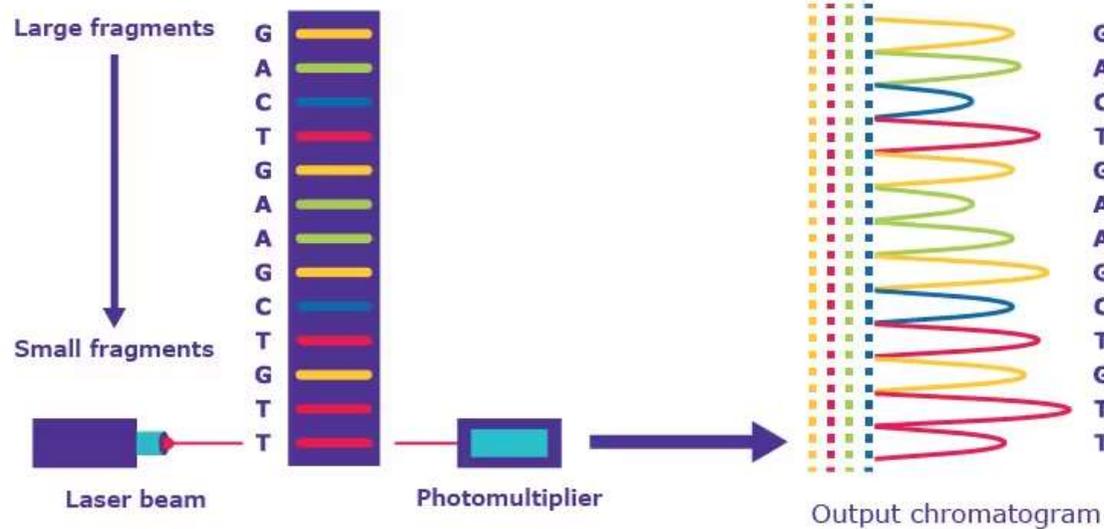
L'ultimo passaggio prevede la **lettura del gel per determinare la sequenza del DNA in esame**. Dato che la **DNA polimerasi** sintetizza il DNA solo nella **direzione da 5' a 3'** a partire da un primer fornito, ogni ddNTP terminale corrisponderà a uno specifico nucleotide nella sequenza originale (per es., il frammento più corto deve terminare a livello del primo nucleotide a partire dall'estremità 5', il secondo frammento più corto deve terminare a livello del secondo nucleotide a partire dall'estremità 5' e così via). Pertanto, leggendo le bande del gel dalla più piccola alla più grande è possibile determinare la sequenza da 5' a 3' del filamento di DNA di partenza.

Il Sequenziamento Sanger



Nel sequenziamento **manuale** si leggono tutte e quattro le corsie del gel contemporaneamente, spostandosi dal basso verso l'alto, e utilizzando la corsia per determinare l'identità del ddNTP terminale di ogni banda. Per esempio, se la banda più in basso è presente nella colonna che corrisponde al ddGTP, allora il frammento più piccolo della PCR termina con ddGTP e il primo nucleotide a partire dall'estremità 5' della sequenza di partenza corrisponde a una guanina (G).

Il Sequenziamento Sanger



Nel sequenziamento **automatico**, un computer legge tutte le bande del gel capillare, in ordine, usando la fluorescenza per assegnare l'identità di ogni ddNTP terminale (le cosiddette "call").

In breve, i frammenti sono caricati in un unico pozzetto e separati per elettroforesi; durante l'elettroforesi un laser eccita i "tag" (marcatori) fluorescenti di ciascuna banda e un computer rileva la risultante luce emessa. Dato che ognuno dei quattro ddNTP è dotato di un "tag" fluorescente diverso, la luce emessa è indicativa dell'identità del ddNTP terminale. Si ottiene così un **chromatogramma**, dove è mostrato il picco di fluorescenza di ogni nucleotide lungo la catena del DNA stampo, con aree proporzionali all'intensità di emissione.

Il Sequenziamento Sanger

SEQUENZIAMENTO CON METODO SANGER *VERSUS* PCR

Il sequenziamento con metodo Sanger e la PCR si avvalgono degli stessi materiali di partenza e possono essere usati congiuntamente.

La **PCR** si usa per amplificare il **DNA nella sua interezza**. Sebbene si possano generare accidentalmente anche frammenti di svariate lunghezze (per es., la DNA polimerasi potrebbe staccarsi), l'obiettivo è duplicare l'intera sequenza di DNA. A tal fine, gli "ingredienti" sono il DNA target, i nucleotidi, il primer a DNA e la DNA polimerasi.

Nel **sequenziamento con metodo Sanger** lo scopo è generare oligonucleotidi di **tutte le lunghezze possibili**, dalla più corta fino a quella più lunga e completa del DNA target. Ecco perché, oltre ai materiali di partenza per la PCR, sono necessari anche i didesossinucleotidi.

Il Sequenziamento Sanger

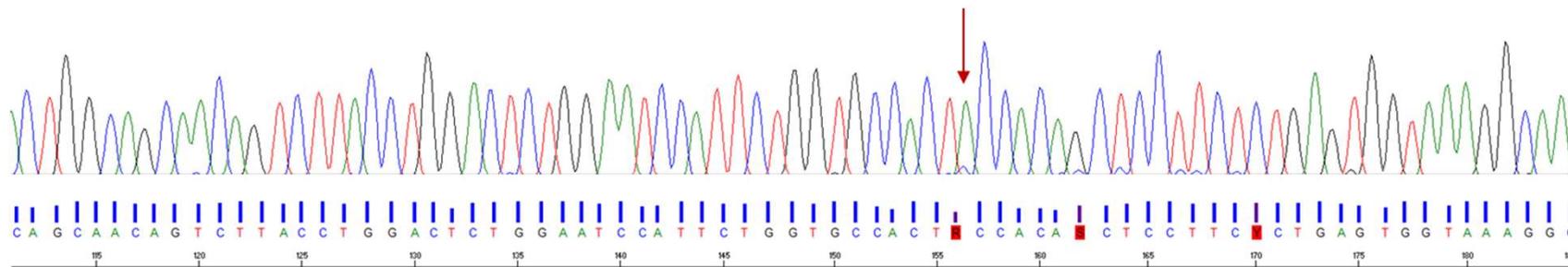
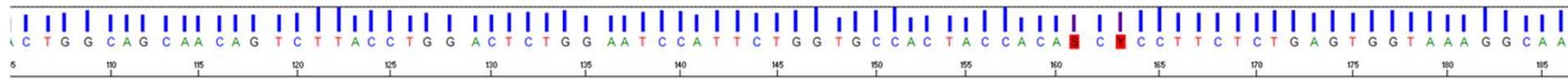
SEQUENZIAMENTO CON METODO SANGER *VERSUS* PCR

Il **sequenziamento con metodo Sanger** e la **PCR** possono **essere abbinati** quando si genera il materiale di partenza per un protocollo di sequenziamento secondo Sanger. La PCR può servire a **creare molte copie del DNA da sequenziare**.

Disporre di più di uno stampo da cui partire può rendere più efficiente il protocollo di Sanger. Se la sequenza target è lunga 1.000 nucleotidi ed è disponibile una sola copia dello stampo, sarà necessario molto tempo per generare i 1.000 frammenti dotati di tag. Invece se sono disponibili molte copie dello stampo, in teoria servirà meno tempo per generarli.

Il Sequenziamento Sanger

Sequenza NON mutata

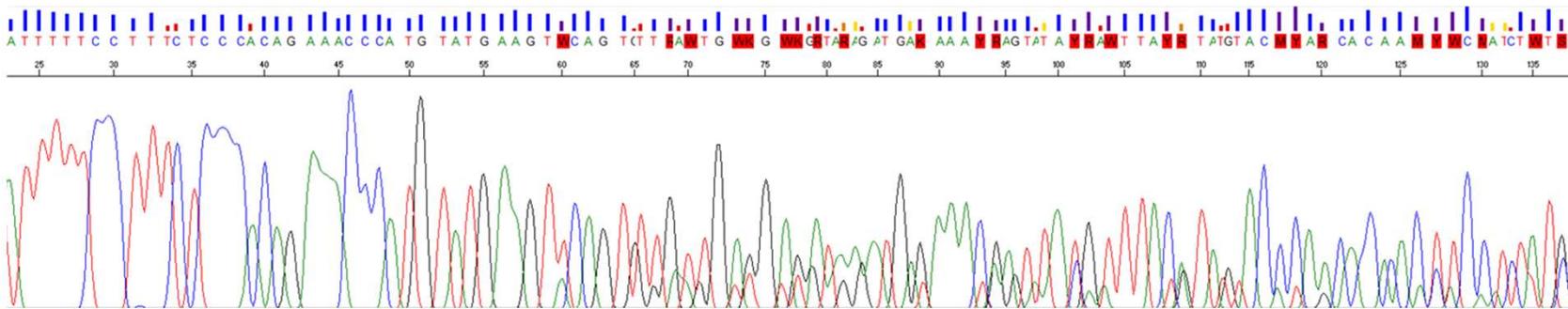
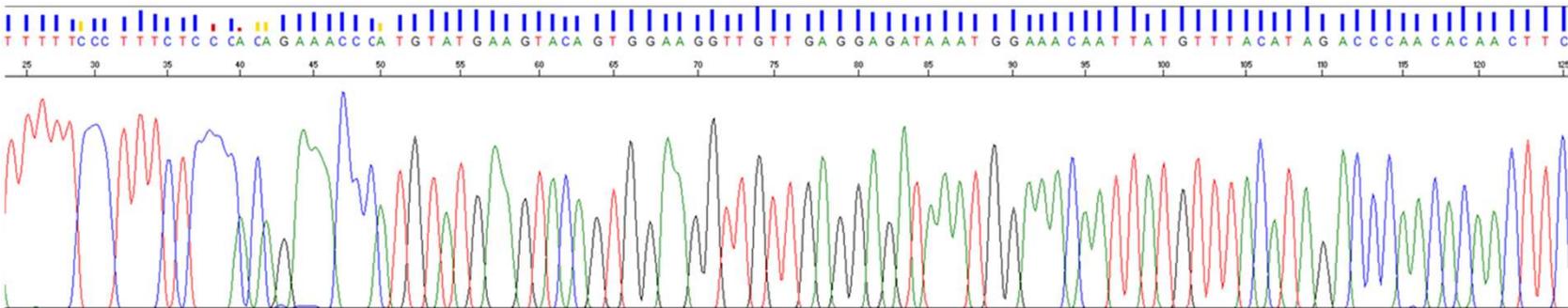


Sequenza con presenza di mutazione puntiforme



Il Sequenziamento Sanger

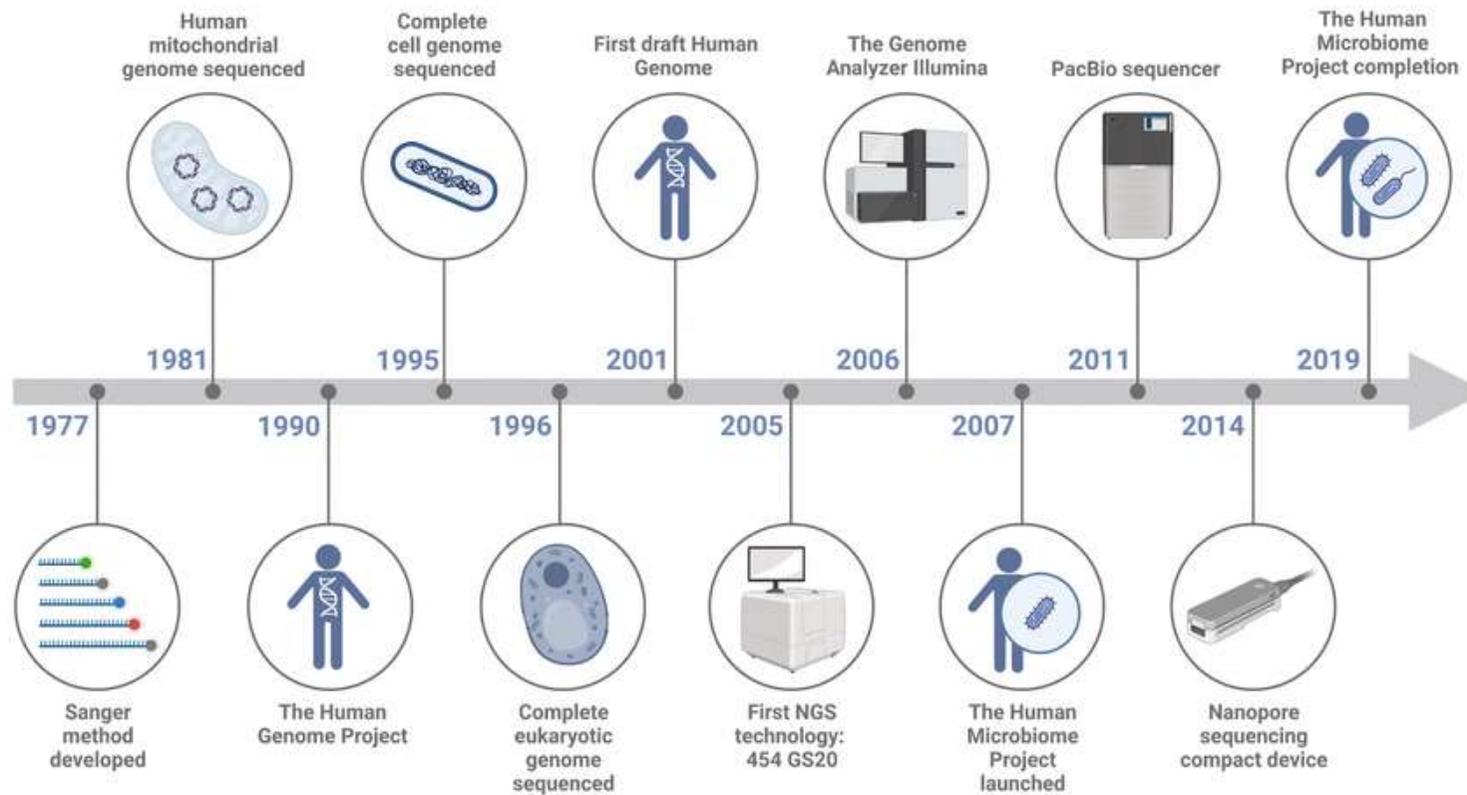
Sequenza NON mutata



Sequenza con presenza di delezione



Dal Sequenziamento Sanger al Sequenziamento di seconda generazione



Il Sequenziamento di seconda generazione o Next Generation Sequencing (NGS)

La definizione di **sequenziamento di seconda generazione o Next Generation Sequencing (NGS)** comprende diverse tecnologie correlate che consentono la copertura, tramite **sequenziamento massivo parallelo o sequenziamento profondo**, di una specifica regione del genoma o dell'intero genoma di un dato organismo.

Il sequenziamento NGS utilizza un procedimento molto simile a quello del Sequenziamento Sanger; tuttavia la principale differenza risiede nella **quantità** di campione che è possibile sequenziare.

Mentre il metodo Sanger permette di sequenziare un frammento di DNA di modeste dimensioni (700 – 1000 paia di basi), attraverso l'utilizzo della tecnologia **NGS** si è in grado di sequenziare più frammenti di DNA **simultaneamente, in una sola corsa, in un tempo breve**, con un basso prezzo per base sequenziata. I sistemi di sequenziamento adoperati nella tecnologia di seconda generazione sono molteplici, ognuno di questi è sviluppato da un'azienda differente

Il Sequenziamento di seconda generazione o Next Generation Sequencing (NGS)

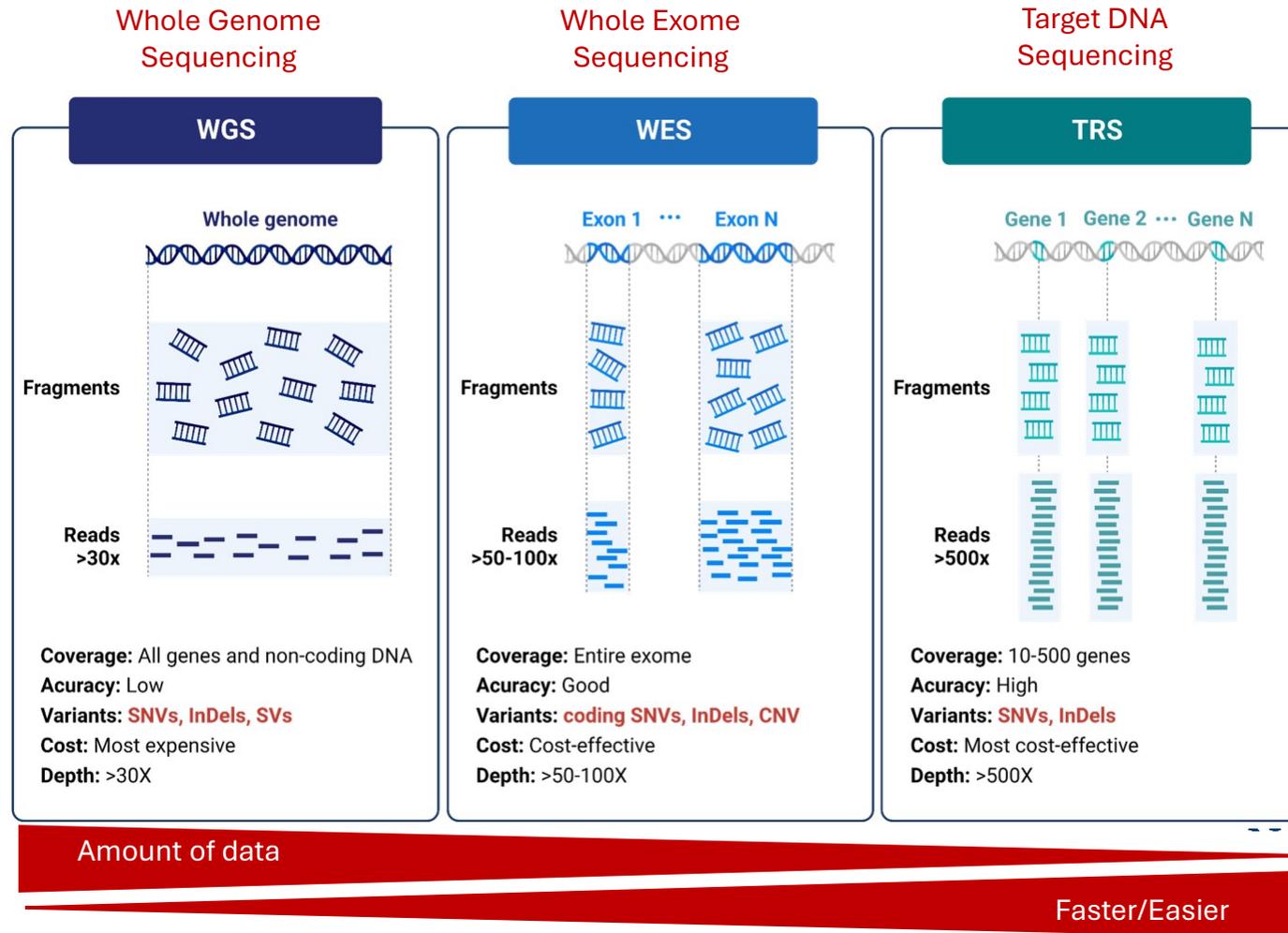
	Sanger Sequencing	Targeted NGS
Benefits	<ul style="list-style-type: none">• Fast, cost-effective sequencing for low numbers of targets (1–20 targets)• Familiar workflow	<ul style="list-style-type: none">• Higher sequencing depth enables higher sensitivity (down to 1%)• Higher discovery power*• Higher mutation resolution†• More data produced with the same amount of input DNA‡• Higher sample throughput
Challenges	<ul style="list-style-type: none">• Low sensitivity (limit of detection ~15–20%)• Low discovery power• Not as cost-effective for high numbers of targets (> 20 targets)• Low scalability due to increasing sample input requirements	<ul style="list-style-type: none">• Less cost-effective for sequencing low numbers of targets (1–20 targets)• Time-consuming for sequencing low numbers of targets (1–20 targets)

<https://youtu.be/fCd6B5HRaZ8>

Il Sequenziamento di seconda generazione o Next Generation Sequencing (NGS)

TIPOLOGIE DI TEST NGS. I test genetici basati sull'NGS possono essere suddivisi in base all'ampiezza della porzione di genoma analizzata.

DNA SEQUENCING



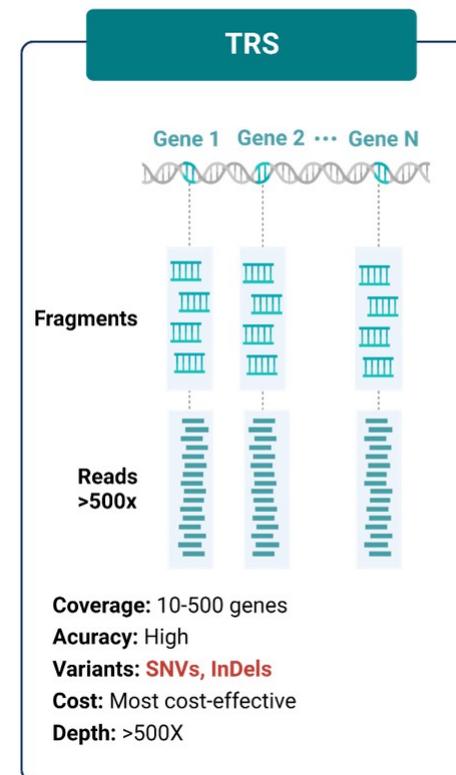
Il Sequenziamento di seconda generazione o Next Generation Sequencing (NGS)

TIPOLOGIE DI TEST NGS. I test genetici basati sull'NGS possono essere suddivisi in base all'ampiezza della porzione di genoma analizzata.

DNA SEQUENCING

Target DNA Sequencing

- ✓ Dataset ridotti, analisi computazionale più agile
- ✓ Scalabile (si possono lavorare più campioni per corsa)
- ✓ Tempi più brevi compatibili con TAT
- ✓ Minori costi



Amount of data

Faster/Easier

Il Sequenziamento di seconda generazione o Next Generation Sequencing (NGS)

TIPOLOGIE DI TEST NGS. I test genetici basati sull'NGS possono essere suddivisi in base all'ampiezza della porzione di genoma analizzata.

The number of genes covered by commercial panels is highly variable

up to >500 genes in panels that also permit to calculate the Tumor Mutational Burden (TMB)

approximately 16 genes (the genes considered by current Livelli Essenziali di Assistenza -LEAs- in the frame of the SSN)

Table 2: DNA content included in TruSight Oncology 500 and TruSight Oncology 500 High-Throughput

ABL1	BCR	CHEK1	EPHA7	FGF23	GSK3B	IDH2	MAP3K1	NF2	PIK3CA	RAD51D	SMAD4	TGFB2
ABL2	BIRC3	CHEK2	EPHA1	FGF3	H3F3A	IFNGR1	MAP3K13	NFE2L2	PIK3CB	RAD52	SMARCA4	TMEM127
ACVR1	BLM	CIC	ERBB2	FGF4	H3F3B	INHBA	MAP3K14	NFKB1A	PIK3CD	RAD54L	SMARCB1	TM6RSS2
ACVR1B	BMPRIA	CREBBP	ERBB3	FGF5	H3F3C	JNPP4A	MAP3K4	NKX2-1	PIK3CG	RAF1	SMARCD1	TNFAIP3
AKT1	BRAF	CRKL	ERBB4	FGF6	HGF	JNPP4B	MAPK1	NKX3-1	PIK3R1	RANBP2	SMC1A	TNFRSF14
AKT2	BRCA1F	CRLF2	ERCC1	FGF7	HIST1H1C	INSR	MAPK3	NOTCH1	PIK3R2	RARA	SMC3	TOP1
AKT3	BRCA2F	CSF1R	ERCC2	FGFR1	HIST1H2BD	IRF2	MAX	NOTCH2	PIK3R3	RASA1	SMO	TOP2A
ALK	BRD4	CSF3R	ERCC3	FGFR2	HIST1H3A	IRF4	MCL1	NOTCH3	PIM1	RB1	SNCAIP	TP53
ALOX12B	BRIP1	CSNK1A1	ERCC4	FGFR3	HIST1H3B	IRS1	MDC1	NOTCH4	PLCG2	RBM10	SOC1	TP63
ANKRD11	BTG1	CTCF	ERCC5	FGFR4	HIST1H3C	IRS2	MDM2	NPM1	PLK2	RECOL4	SOX10	TRAF2
ANKRD26	BTK	CTLA4	ERG	FH	HIST1H3D	JAK1	MDM4	NRAS	PMAIP1	REL	SOX17	TRAF7
APC	CTNFB	CTNNA1	ERRF1	FLCN	HIST1H3E	JAK2	MED12	NRG1	PMS1	RET	SOX2	TSC1
AR	CALR	CTNNA1	ESR1	FLII	HIST1H3F	JAK3	MEF2B	NSD1	PMS2	RFWD2	SOX9	TSC2
ARAF	CARD11	CUL3	ETS1	FLT1	HIST1H3G	JUN	MEN1	NTRK1	PNKRC1	RHEB	SPEN	TSHR
ARFIP1	CASP8	CUX1	ETV1	FLT3	HIST1H3H	KAT5A	MET	NTRK2	POLD1	RHOA	SPOP	UZAF1
ARID1A	CBFB	CXCR4	ETV4	FLT4	HIST1H3I	KDM5A	MGA	NTRK3	POLE	RICTOR	SPTA1	VEGFA
ARID1B	CB1	CYLD	ETV5	FOXK1	HIST1H3J	KDM5C	MITF	NUP93	PPARG	RIT1	SRC	VHL
ARID2	CCND1	DAXX	ETV6	FOXL2	HIST2H3A	KDM6A	MLH1	NUTM1	PPMD	RNF43	SRSF2	VTCN1
ARID5B	CCND2	DCUN1D1	EWSR1	FOXO1	HIST2H3C	KDR	MLL	PAK1	PPP2R1A	RGS1	STAG1	WSP3
ASXL1	CCND3	DDR2	EZH2	FOXO1	HIST2H3D	KEAP1	MLL2	PAK3	PPP2R2A	RPS6KA4	STAG2	WT1
ASXL2	CCNE1	DDX41	FAM123B	FRS2	HIST3H3	KEL	MPL	PAK7	PPP3C	RPS6KB1	STAT3	XIAP
ATM	CD274	DHX15	FAM175A	FUBP1	HLA-A	KIF5B	MRE11A	PALB2	PRDM1	RPS6KB2	STAT4	XPO1
ATR	CD276	DICER1	FAM46C	FYN	HLA-B	KIT	MSH2	PARK2	PREX2	RPTOR	STAT5A	XRC2
ATRX	CD74	DIS3	FANCA	GABRA6	HLA-C	KLIF4	MSH3	PARP1	PRKARIA	RUNX1	STAT5B	YAP1
AURKA	CD78A	DNAJB1	FANCC	GATA1	JNFB1A	KLHL6	MSH6	PAX3	PRKCI	RUNX1T1	STK11	YES1
AURKB	CD78B	DNM11	FANCD2	GATA2	HRNR9K	KMT2B	MST1	PAX5	PRKDC	HYBP	STK40	ZBTB2
AXIN1	CD273	DNMT3A	FANCE	GATA3	H0XB13	KMT2C	MST1R	PAX7	PRSS8	SDHA	SUFU	ZBTB7A
AXIN2	CDH1	DNMT3B	FANCF	GATA4	IGF1	KMT2D	MTOR	PAX8	PTCHI	SDHAF2	SUZ12	ZFXK3
AXL	CDK12	DOT1L	FANCG	GATA6	IGFIR	KRAS	MUTYH	PBRM1	PTEN	SDHB	SYK	ZNF217
B2M	CDK4	E2F3	FANCI	GEN1	IGF2	LAMP1	MYB	PDCD1	PTN1	SDHC	TAF1	ZNF703
BAP1	CDK6	EED	FANCL	GID4	IKBKE	LATS1	MYC	PDCD1L2	PTPRD	SDHD	TBX3	ZRSR2
BARD1	CDK8	EGFL7	FAS	GLI1	IKZF1	LATS2	MYCL1	PDGFRA	PTPRS	SETBP1	TCEB1	
BBC3	CDKN1A	EGFR	FAT1	GNAT1	IL10	LMG1	MYCN	PDGFBR	PTPRT	SETD2	TCF3	
BCL10	CDKN1B	EIF3AK	FBXW7	GNAS	IL11	LMNB1	MYCN	PDGFRA	PTPRT	SETD2	TCF3	
BCL2	CDKN2A	EIF4A2	FGF1	GNAS	IL11	LMNB1	MYCN	PDGFRA	PTPRT	SETD2	TCF3	
BCL2L1	CDKN2B	EIF4E	FGF8	GNAS	IL11	LMNB1	MYCN	PDGFRA	PTPRT	SETD2	TCF3	
BCL2L11	CDKN2C	EML4	FGF9	GNAS	IL11	LMNB1	MYCN	PDGFRA	PTPRT	SETD2	TCF3	
BCL2L2	CEBPA	EP300	FGF10	GNAS	IL11	LMNB1	MYCN	PDGFRA	PTPRT	SETD2	TCF3	
BCL6	CENPA	EPCAM	FGF14	GNAS	IL11	LMNB1	MYCN	PDGFRA	PTPRT	SETD2	TCF3	
BCOR	CHD2	EPHA3	FGFR19	GRIN2A	ID3	MAP2K2	NEGR1	PIK3C2G	RAD51B	SMAD2	TFRC	
BCORL1	CHD4	EPHA5	FGF2	GRM3	IDH1	MAP2K4	NF1	PIK3C3	RAD51C	SMAD3	TOFBI	

a. Large rearrangements (non-level CNVs) detected for BRCA1 and BRCA2.
Content shaded in gray is analyzed for CNV detection.

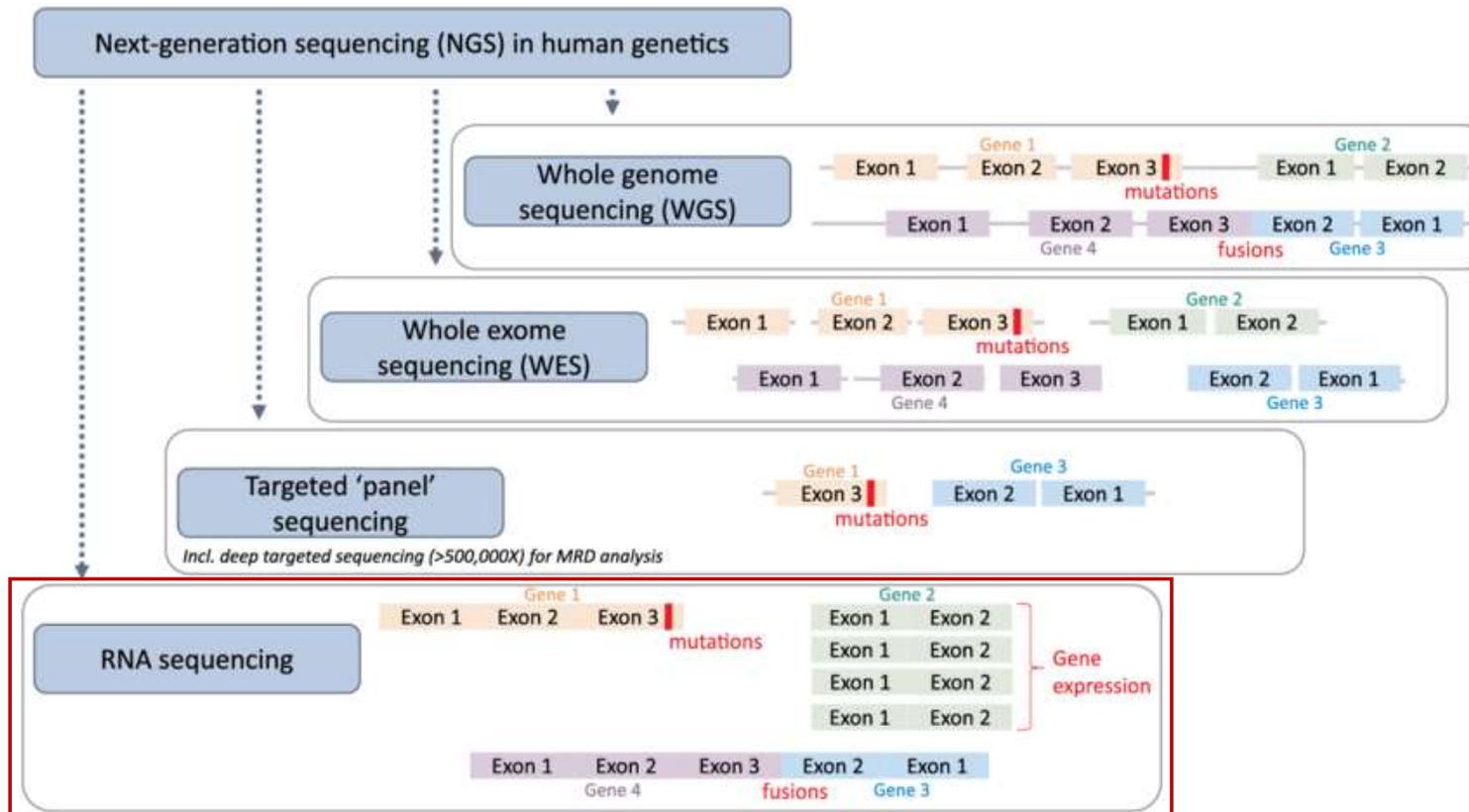
ALK	BRAF	EGFR	ERBB2
FGFR3	HRAS	IDH1-2	NRAS
KIT	KRAS	MET	ROS1
PDGFRA	PIK3CA	RET	POLE

Amount of data

Faster/Easier

Il Sequenziamento di seconda generazione o Next Generation Sequencing (NGS)

TIPOLOGIE DI TEST NGS. I test genetici basati sull'NGS possono essere suddivisi in base all'ampiezza della porzione di genoma analizzata.



Il Sequenziamento di seconda generazione o Next Generation Sequencing (NGS)

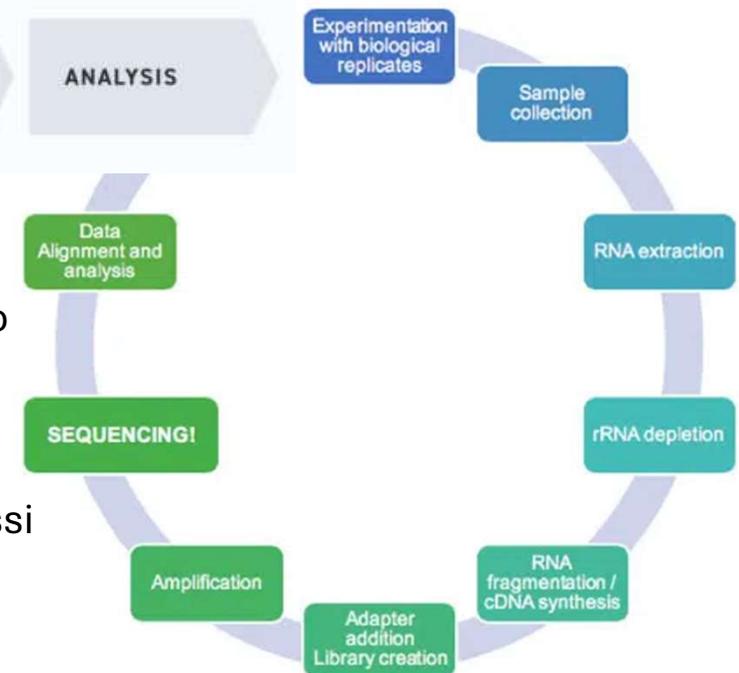
TIPOLOGIE DI TEST NGS. I test genetici basati sull'NGS possono essere suddivisi in base all'ampiezza della porzione di genoma analizzata.

RNA SEQUENCING



Circa l'80% del RNA cellulare è RNA ribosomiale
→ la maggior parte dei protocolli NGS per RNA prevedono

- Ribodeplezione (eliminazione rRNA)
 - Selezione polyA
- per favorire il rilevamento anche dei trascritti poco espressi



Il Sequenziamento di seconda generazione o Next Generation Sequencing (NGS)

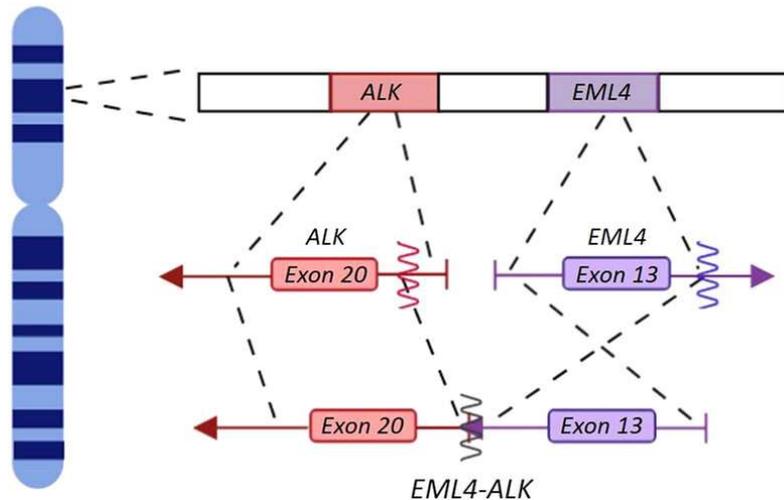
Rilevazione di

- ✓ Traslocazioni
- ✓ Alterazioni di splicing
- ✓ **Fusioni**

Geni maggiormente coinvolti sono: *ALK*, *ROS1*, *RET* e *NTRK* (fusioni); *MET* (exon 14 skipping)

RNA SEQUENCING

Chromosome 2

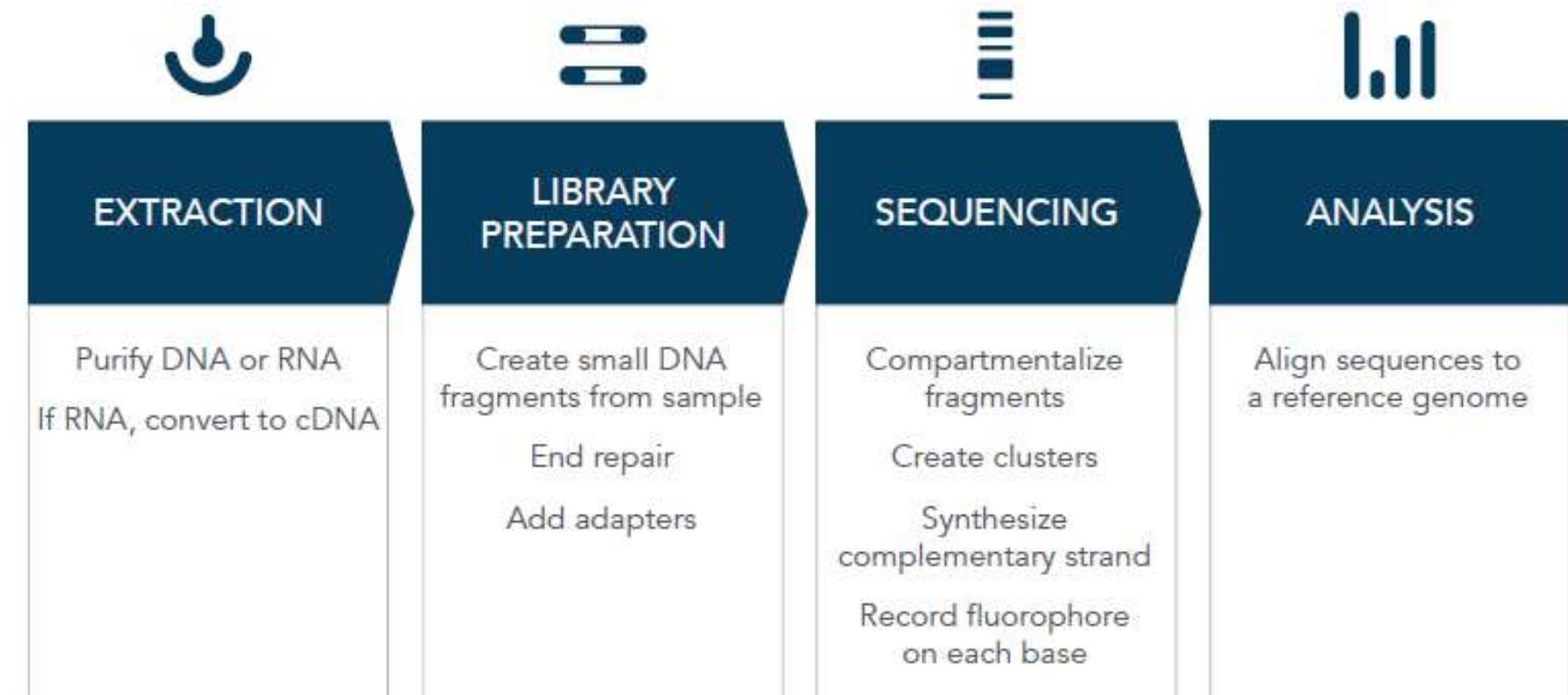


DNA-based NGS: la rilevazione di fusioni può essere complessa per la presenza di lunghe sequenze introniche.

RNA-based NGS: consente di rilevare solo quei prodotti di fusione che hanno una maggiore probabilità di generare una proteina 'fusa'.

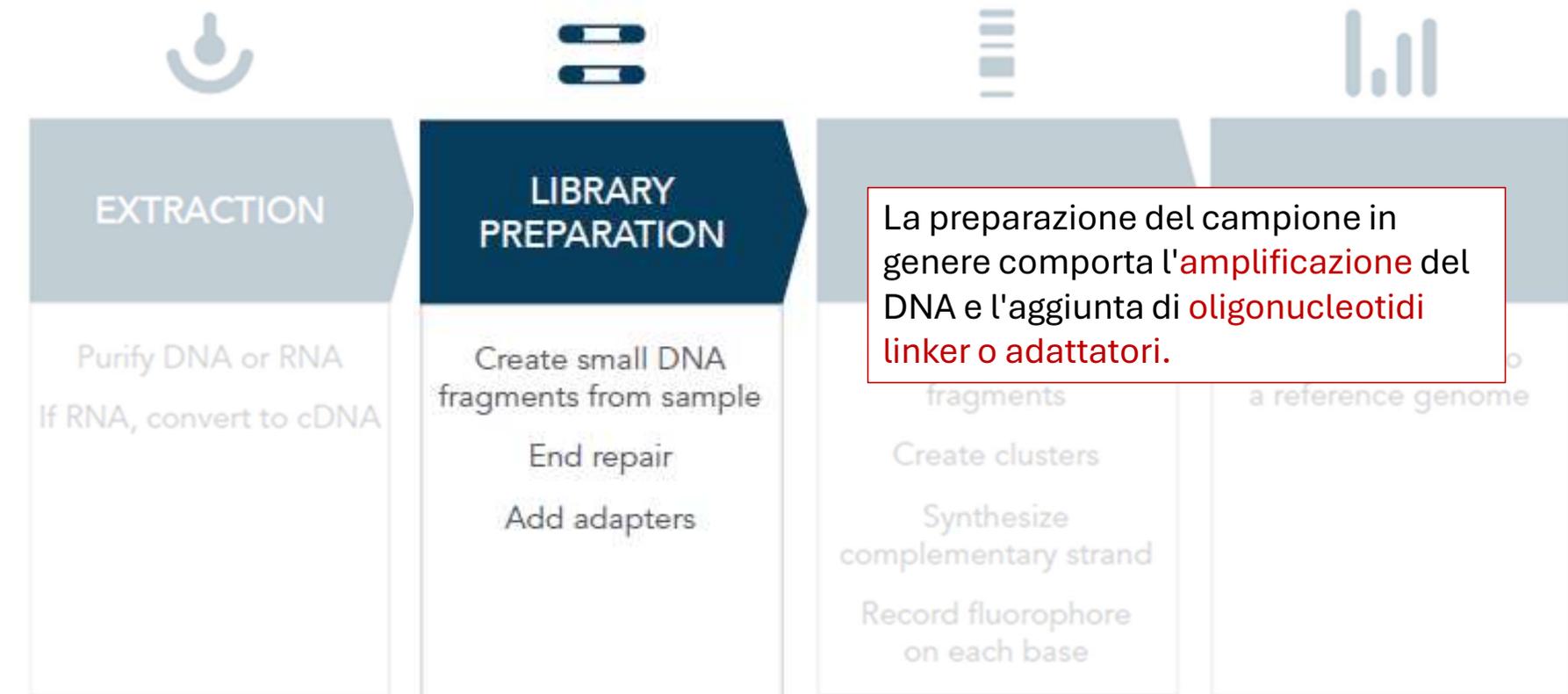
Il Sequenziamento di seconda generazione o Next Generation Sequencing (NGS)

Le piattaforme più diffuse integrano nel flusso di lavoro dell'NGS molti dei passaggi critici, tra cui la (1) **preparazione del campione o della libreria**, (2) **la generazione dei cluster**, e il **sequenziamento** e (4) **l'analisi dei dati**.



Il Sequenziamento di seconda generazione o Next Generation Sequencing (NGS)

Le piattaforme più diffuse integrano nel flusso di lavoro dell'NGS molti dei passaggi critici, tra cui la (1) **preparazione del campione o della libreria**, (2) **la generazione dei cluster**, e il sequenziamento e (3) **l'analisi dei dati**.



Il Sequenziamento di seconda generazione o Next Generation Sequencing (NGS)

Le piattaforme più diffuse integrano nel flusso di lavoro dell'NGS molti dei passaggi critici, tra cui la (1) **preparazione del campione o della libreria**, (2) **la generazione dei cluster**, e il sequenziamento e (3) **l'analisi dei dati**.



- ✓ La **generazione di un cluster** a partire da ogni sequenza di DNA avviene quando il DNA legato covalentemente al linker si ibrida a una **superficie solida** per l'amplificazione a ponte mediante PCR o con metodi alternativi come la PCR in emulsione.
- ✓ Sono disponibili molti metodi per **sequenziare il DNA**, tra cui sequenziamento mediante ligazione, sequenziamento per sintesi, pirosequenziamento e sequenziamento basato su semiconduttore ionico.
- ✓ Ciascun metodo comporta fasi di reazione e processi chimici differenti dai quali dipenderanno la lunghezza di ciascuna sequenza (lunghezza delle reads), il tasso di errore e il costo dei reagenti.

SEQUENCING

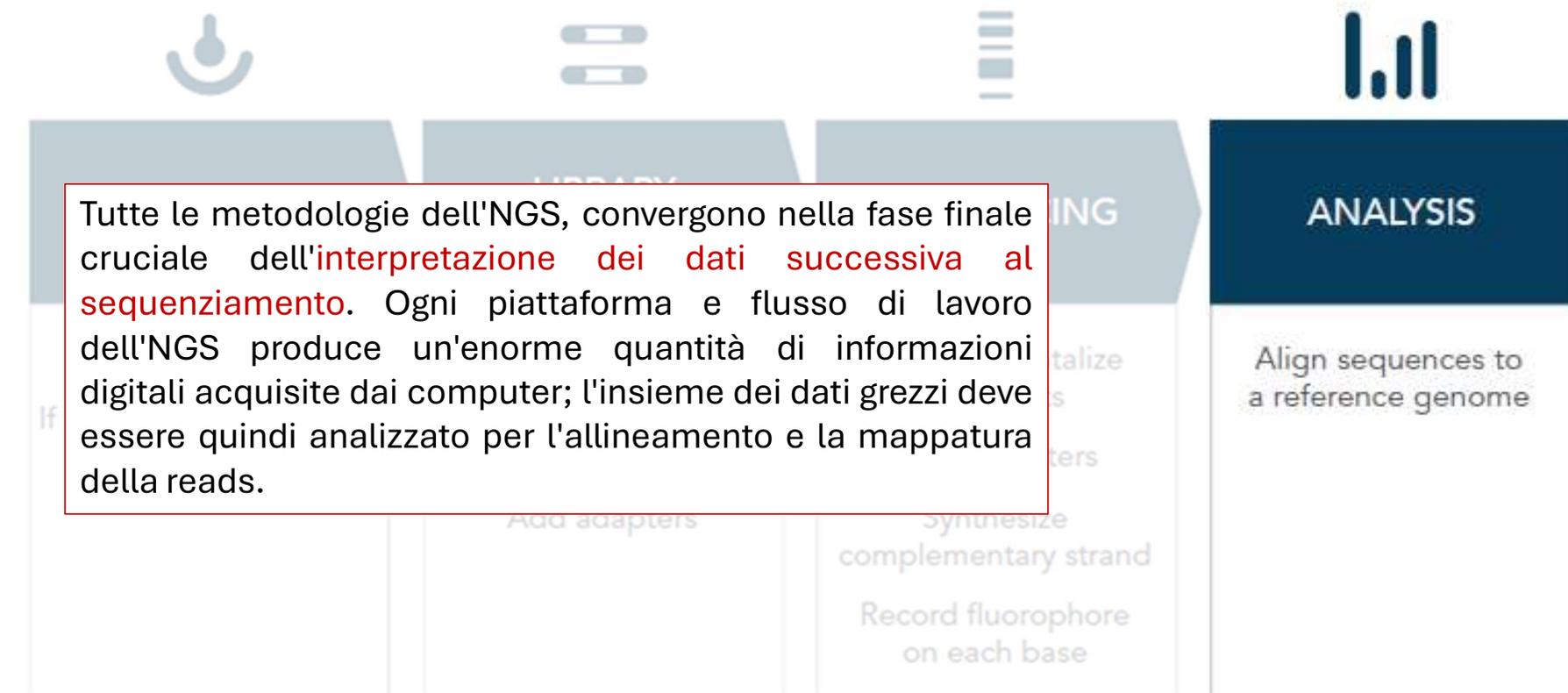
Compartmentalize fragments
Create clusters
Synthesize complementary strand
Record fluorophore on each base

ANALYSIS

Align sequences to a reference genome

Il Sequenziamento di seconda generazione o Next Generation Sequencing (NGS)

Le piattaforme più diffuse integrano nel flusso di lavoro dell'NGS molti dei passaggi critici, tra cui la (1) **preparazione del campione o della libreria**, (2) **la generazione dei cluster**, e il sequenziamento e (3) **l'analisi dei dati**.



Il Sequenziamento di seconda generazione o Next Generation Sequencing (NGS)



1. **frammentazione** e/o riduzione delle sequenze bersaglio ad una lunghezza desiderata
2. **legame di oligonucleotidi adattatori** alle estremità dei frammenti di DNA bersaglio
3. **arricchimento mediante PCR** delle sequenze bersaglio e generazione delle library

Il Sequenziamento di seconda generazione o Next Generation Sequencing (NGS)

LIBRARY PREPARATION

- Create small DNA fragments from sample
- End repair
- Add adapters

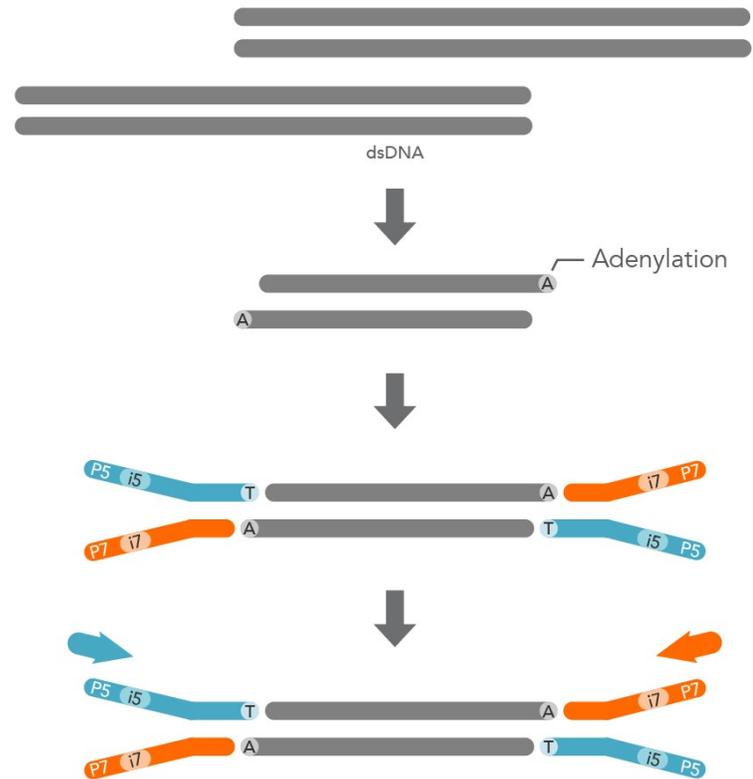
LIBRARY PREPARATION

Fragmentation

End repair and A-tailing

Ligation

PCR amplification

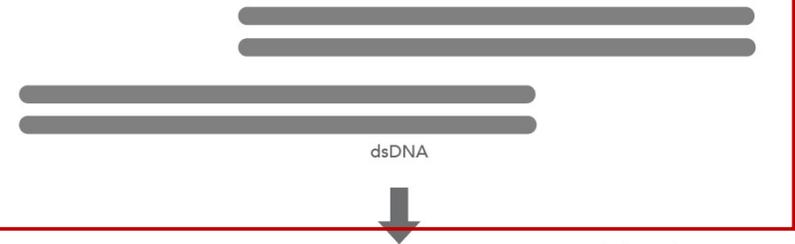


Il Sequenziamento di seconda generazione o Next Generation Sequencing (NGS)



LIBRARY PREPARATION

Fragmentation



1. La **frammentazione** dei campioni di DNA è **casuale**.

- ✓ Frammentazione fisica del DNA: rottura casuale acustica usando ultrasuoni, es. Covaris (100-500bp)
- ✓ Frammentazione chimica del DNA: con calore e cationi di metalli bivalenti (magnesio o zinco)
- ✓ Frammentazione del DNA con metodi enzimatici con DNasi I o altre endonucleasi di restrizione

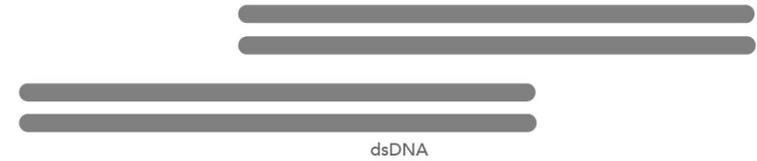
Il Sequenziamento di seconda generazione o Next Generation Sequencing (NGS)

LIBRARY PREPARATION

LIBRARY PREPARATION

- Create small DNA fragments from sample
- End repair
- Add adapters

Fragmentation



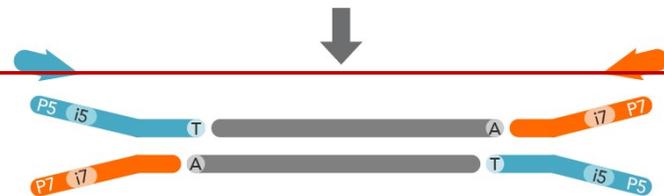
End repair and A-tailing



Ligation



PCR amplification



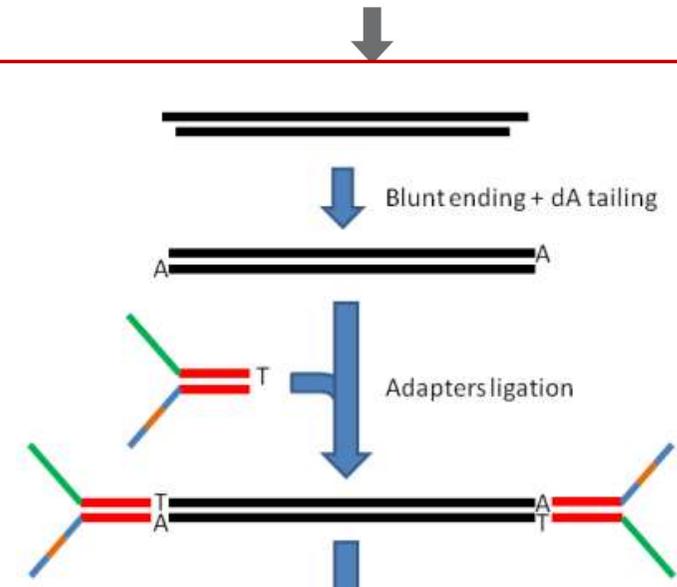
Il Sequenziamento di seconda generazione o Next Generation Sequencing (NGS)



LIBRARY PREPARATION

LIBRARY PREPARATION

- **End repair:** I frammenti ottenuti sono modificati per ottenere estremità blunt
- **A-tailing:** aggiunta della coda **PoliA** → all'estremità blunt si aggiunge una coda A, che serve poi a legare l'adattatore, che contiene una sequenza base di T complementare al coda poliA.



Il Sequenziamento di seconda generazione o Next Generation Sequencing (NGS)



LIBRARY PREPARATION

Create small DNA fragments from sample

End repair

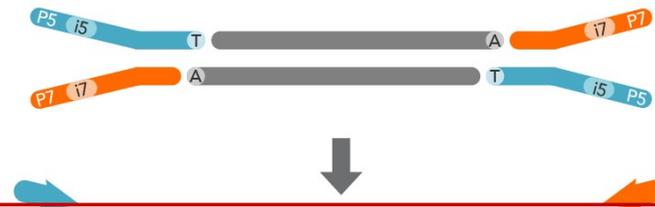
Add adapters

LIBRARY PREPARATION

Gli adattatori presentano **tre** regioni.

1. la sequenza che consente ai frammenti della libreria di legare gli **oligonucleotidi** presenti sul supporto di sequenziamento
2. le **sequenze indice**, una **regione di 6 – 8 paia** di basi, che permette di distinguere i vari adattatori, permettendo di distinguere i **diversi campioni** e quindi di sequenziarne più di un campione.
3. i siti di legame dei **primer di sequenziamento**.

Ligation



Il Sequenziamento di seconda generazione o Next Generation Sequencing (NGS)



LIBRARY PREPARATION

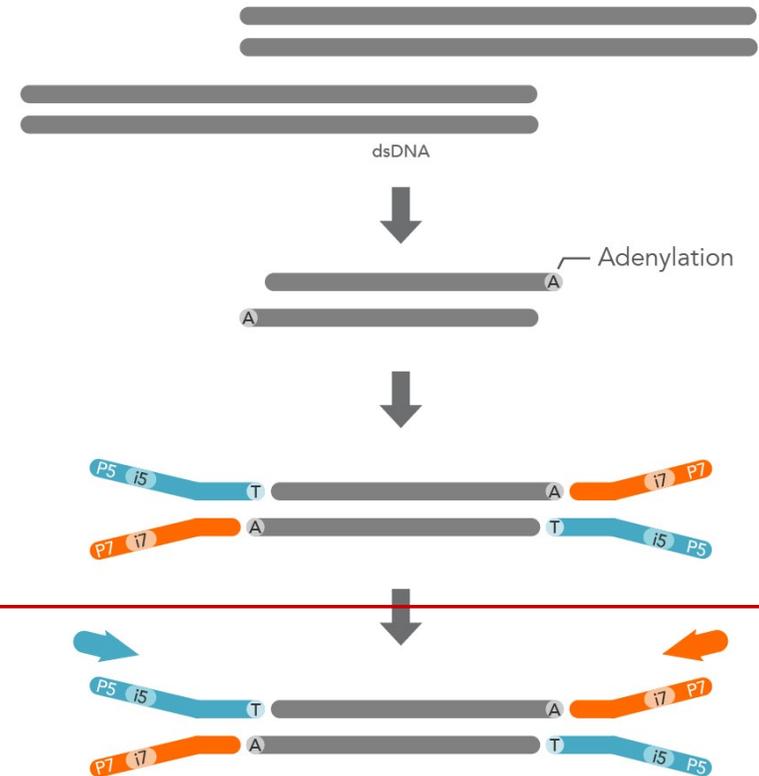
LIBRARY PREPARATION

Le **dimensioni dei frammenti di DNA bersaglio nella libreria finale** è un parametro fondamentale per la costruzione delle librerie.

Peculiarità del NGS è la lettura di molte sequenze, ma di **ridotta lunghezza**.

La lunghezza delle sequenze lette dipende dalla piattaforma.

Es. il sequenziamento Illumina prevede una lunghezza ottimale delle library di 300-500 bp, complessiva di adattatori.



Il Sequenziamento di seconda generazione o Next Generation Sequencing (NGS)



Libreria basata su ampliconi

Arricchimento delle regioni di interesse **usando una reazione di PCR**: coppie di primers specifici per le regioni di interesse e DNA Pol.

Libreria basata su ibridazione e cattura

Il workflow prevede una fase aggiuntiva, l'**arricchimento** delle sequenze di interesse.

Le regioni desiderate vengono «catturate» tramite ibridazione con oligonucleotidi esca e successivamente selezionate tramite purificazione

Il Sequenziamento di seconda generazione o Next Generation Sequencing (NGS)



LIBRARY
PREPARATION

HYBRIDIZATION CAPTURE

Create sma
fragments fro
End ref
Add adap

L'**arricchimento** delle sequenze di interesse è eseguito dopo aver creato le librerie: le regioni d'interesse della libreria sono 'catturate' usando delle sonde (sequenze oligonucleotidiche) coniugate alla **biotina** (all'estremità 5'). Dopo la cattura le sequenze legate alla sonda con la biotina (sequenze di interesse) sono separate da quelle non legate (sequenze di NON interesse) utilizzando delle biglie di streptavidina, che legheranno tutto quello che è biotinilato.

Overlap tiling (typical for array-based probes)



End-to-end tiling (IDT)



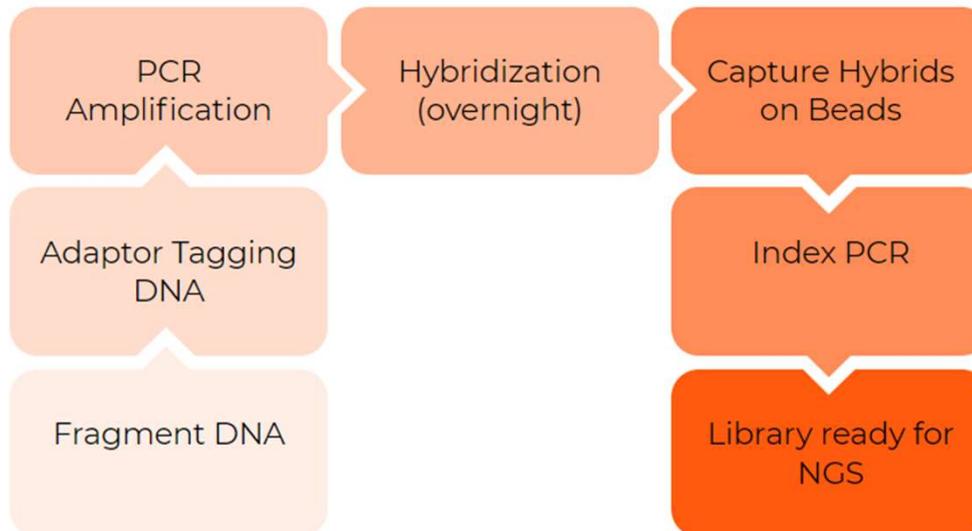
- ✓ **Sonde end to end**
- ✓ **Sonde overlap** aggiunta di una extra sequenza oltre alla regione target di interesse per garantire una copertura sempre totale, permettono di evitare anche lunghe sequenze ripetute.

Il Sequenziamento di seconda generazione o Next Generation Sequencing (NGS)

Amplicon Sequencing Library Preparation Workflow



A Typical Hybrid Capture Target Enrichment workflow



- ✓ Amplicon sequencing library preparation workflow è **rapido**, facile, più economico, **adatto per casi in cui c'è necessità di risposta rapida (diagnostica)**.
- ✓ Lavora bene anche con basse quantità di DNA.
- ✓ Utilizzato in pannelli con pochi geni.
- ✓ Usato per rilevare mutazioni somatiche anche rare, variazioni del copy number, SNPs, indels.

- ✓ Hybrid capture può valutare un'ampia gamma di geni, adatto per **pannelli estesi**, **richiede più tempo**,
- ✓ valuta anche alterazioni strutturali non note (es nuove fusion)
- ✓ adatto per ricerca.

Il Sequenziamento di seconda generazione o Next Generation Sequencing (NGS)

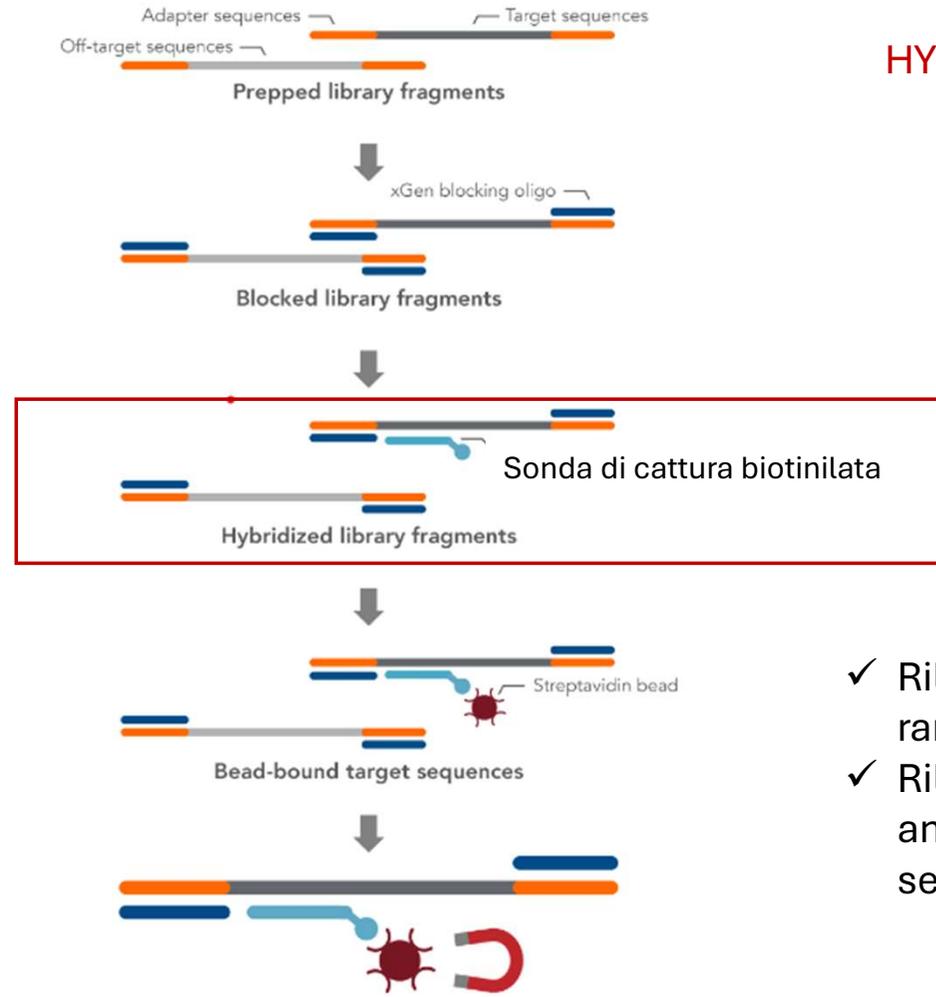
Target capture workflow

Add blocking oligos

Hybridize targets to capture probes

Incubate with magnetic streptavidin beads

Isolate targets with magnet

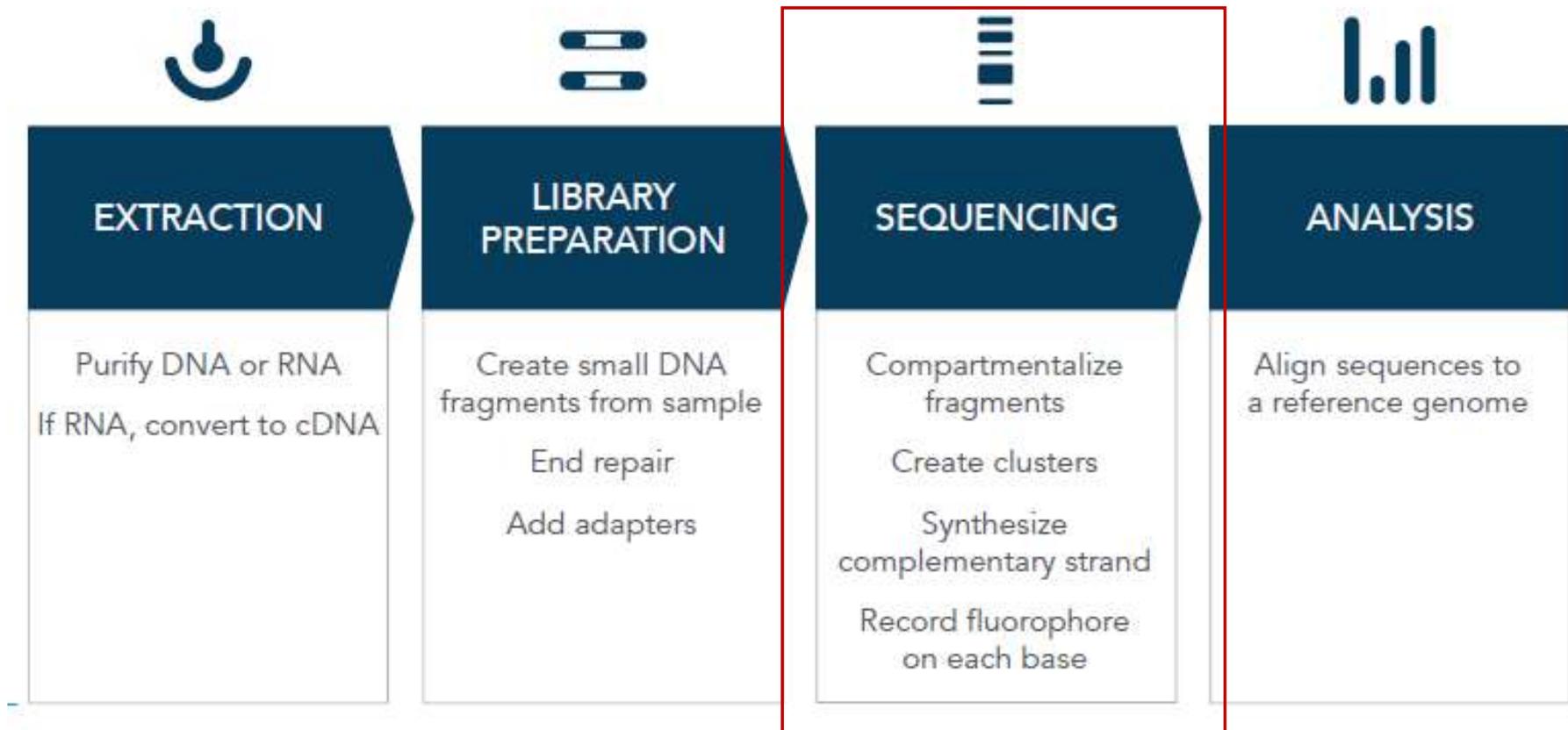


HYBRIDIZATION CAPTURE

- ✓ Rilevazione di varianti rare
- ✓ Rilevazione di varianti anche nel caso di sequenze complesse

Il Sequenziamento di seconda generazione o Next Generation Sequencing (NGS)

Le piattaforme più diffuse integrano nel flusso di lavoro dell'NGS molti dei passaggi critici, tra cui la (1) **preparazione del campione o della libreria**, (2) **la generazione dei cluster**, e il sequenziamento e (3) **l'analisi dei dati**.



Il Sequenziamento di seconda generazione o Next Generation Sequencing (NGS)

Le piattaforme più diffuse integrano nel flusso di lavoro dell'NGS molti dei passaggi critici, tra cui la (1) **preparazione del campione o della libreria**, (2) **la generazione dei cluster**, e il sequenziamento e (3) **l'analisi dei dati**.



- ✓ La **generazione di un cluster** a partire da ogni sequenza di DNA avviene quando il DNA legato covalentemente al linker si ibrida a una superficie solida per l'amplificazione a ponte mediante PCR o con metodi alternativi come la PCR in emulsione.
- ✓ Sono disponibili molti metodi per sequenziare il DNA, tra cui sequenziamento mediante ligazione, sequenziamento per sintesi, pirosequenziamento e sequenziamento basato su semiconduttore ionico.
- ✓ Ciascun metodo comporta fasi di reazione e processi chimici differenti dai quali dipenderanno la lunghezza di ciascuna sequenza (lunghezza delle reads), il tasso di errore e il costo dei reagenti.

SEQUENCING

Compartmentalize fragments
Create clusters
Synthesize complementary strand
Record fluorophore on each base

ANALYSIS

Align sequences to a reference genome

Il Sequenziamento di seconda generazione o Next Generation Sequencing (NGS)

LA GENERAZIONE DEI CLUSTER E IL SEQUENZIAMENTO

illumina®



	MiSeq	NextSeq	HiSeq 2500	HiSeq X Ten
Output	15 Gb	120 GB	1000 GB	1800 GB
Number of Reads	25 Million	400 Million	4 Billion	6 Billion
Read Length	2x300 bp	2x150 bp	2x125 bp (2x250 update mid-2014)	2x150 bp
Cost	\$99K	\$250K	\$740K	\$10M

5/29/2014 IIT Indore Source: Illumina 15

ion torrent
by *life technologies™*



Ion PGM

- 3 types of chips
- 200 or 400 bp reads
- Up to 5.5 million reads / Ion 318 chip
- 4 – 7 h run time

Ion S5 System			Ion S5 XL System		
					
Simple workflow for panels, microbes, exomes, and transcriptomes			Simple, rapid workflow for panels, microbes, exomes, and transcriptomes		
Ion 520 Chip	Ion 530 Chip	Ion 540 Chip	Ion 520 Chip	Ion 530 Chip	Ion 540 Chip
Final Reads 3–5 million	Final Reads 15–20 million	Final Reads 60–80 million	Final Reads 3–5 million	Final Reads 15–20 million	Final Reads 60–80 million

Le tecnologie di sequenziamento a disposizione sono molteplici; le più diffuse sono il sequenziamento con **terminatori reversibili (Illumina)** e il sequenziamento **‘ion semiconductor’ (IonTorrent, life Technologies)**.

Entrambe le metodiche presentano un buon livello di flessibilità riguardo al numero di campioni analizzabili in un'unica corsa, con tempistiche abbastanza variabili relativamente l'analisi dei risultati.

Il Sequenziamento di seconda generazione o Next Generation Sequencing (NGS)

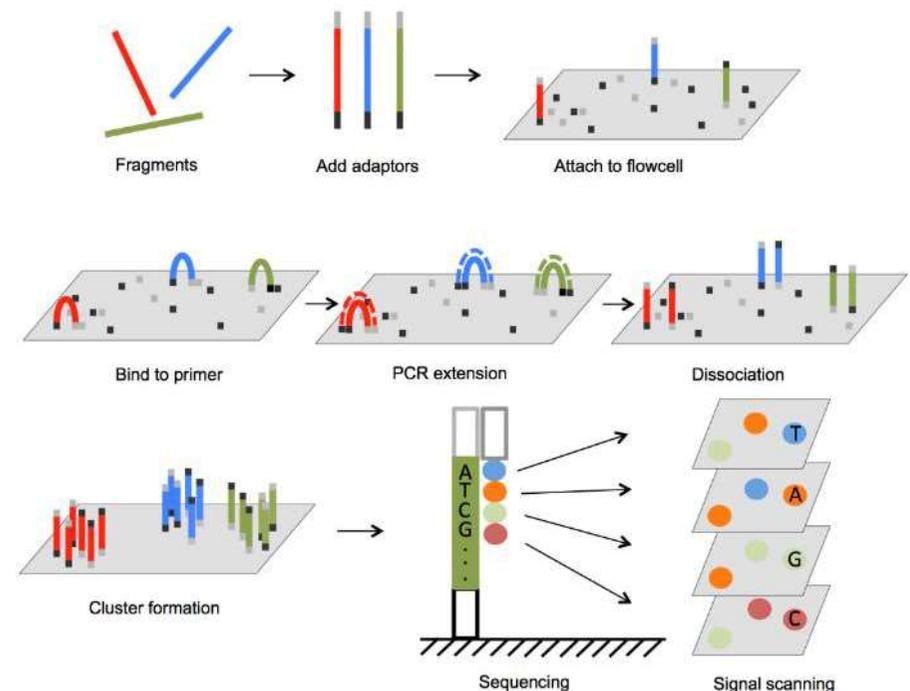
LA GENERAZIONE DEI CLUSTER E IL SEQUENZIAMENTO

Le librerie vengono usate per generare un DNA a singolo filamento (ssDNA) che servirà come stampo per il successivo sequenziamento.

Le due tecnologie più diffuse (Illumina e IonTorrent) prevedono due differenti modalità di generazione del ssDNA:

Illumina: la libreria è dispensata su un supporto solido (Flow Cell) sul quale sono stati sintetizzati e ibridati primers complementari agli adattatori.

Tramite PCR vengono creati gruppi (**clusters**) di ssDNA monoclonali per il successivo sequenziamento.

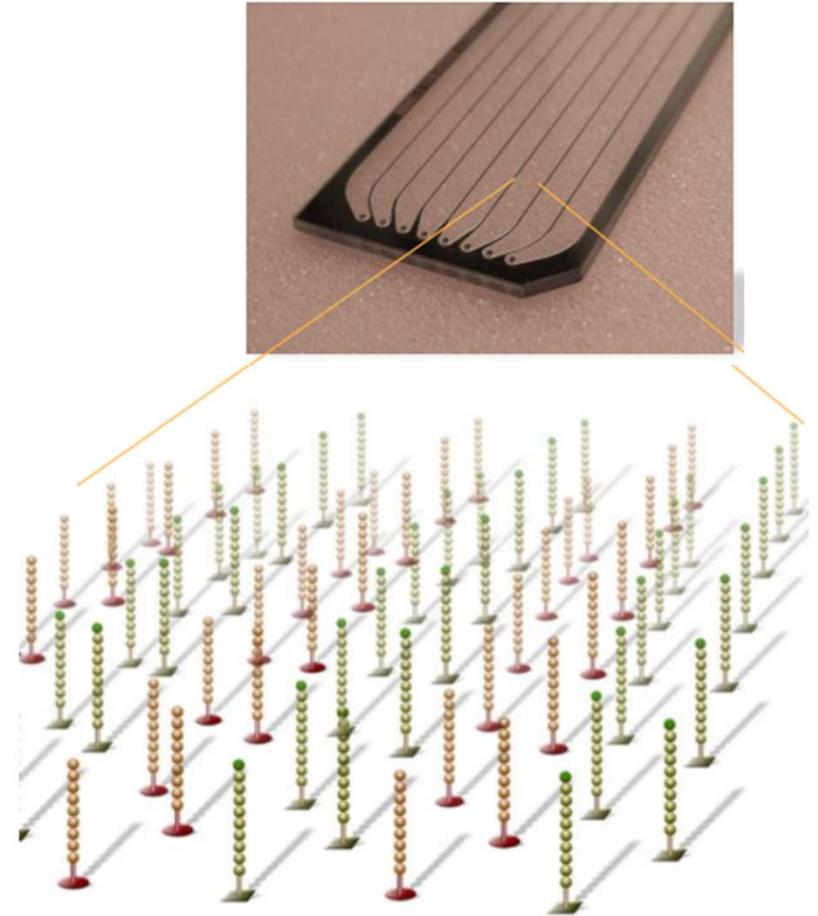


Il Sequenziamento di seconda generazione o Next Generation Sequencing (NGS)

LA GENERAZIONE DEI CLUSTER E IL SEQUenziAMENTO

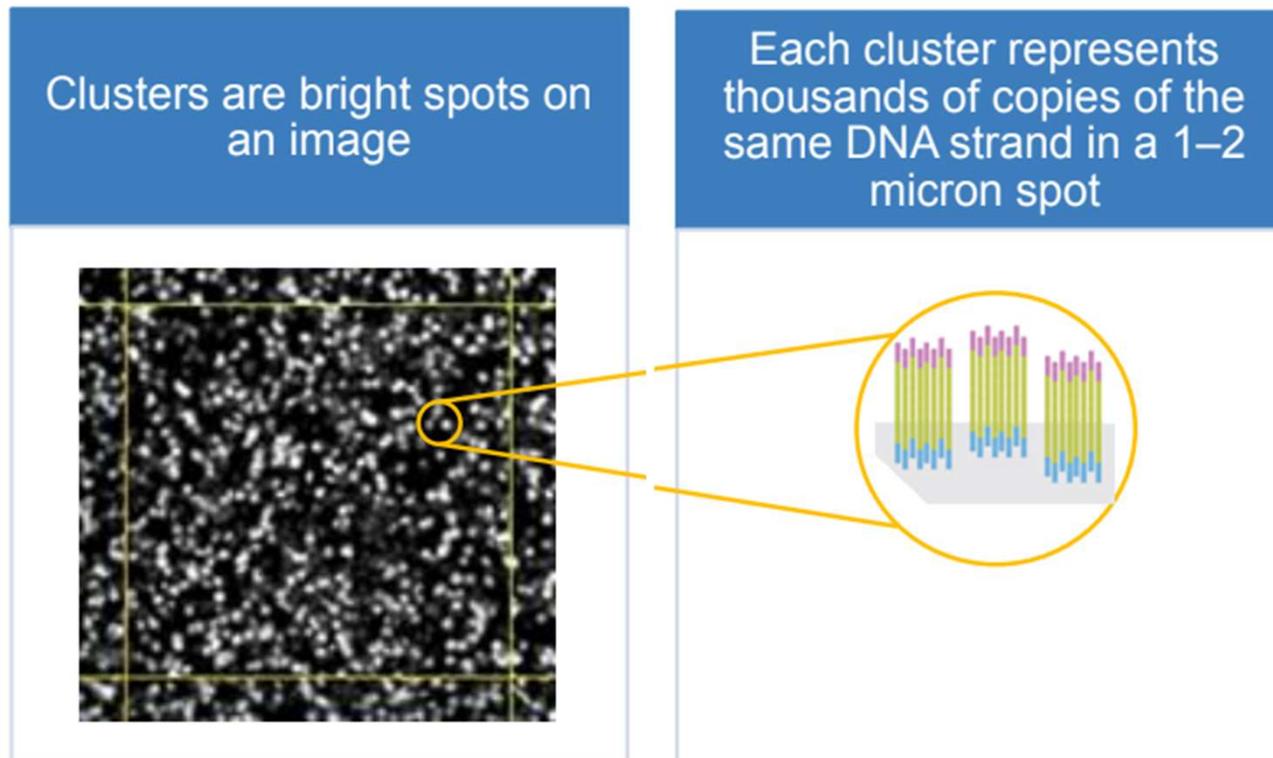
La chimica di sequenziamento si basa un processo di sequenziamento tramite sintesi -**Sequencing by Synthesis (SBS)** :

- primers complementari agli adattatori vengono coniugati ai ssDNA generati precedentemente;
- tramite polimerasi, vengono aggiunti **nucleotidi trifosfato terminatori coniugati con fluorofori**
- viene letta la fluorescenza emessa;
- i **nucleotidi trifosfato terminatori** hanno la peculiarità di essere **'reversibili'**: dopo una prima fase nella quale estensione viene inibita, è possibile proseguire con la sintesi del filamento e quindi con il sequenziamento tramite la modificazione del 3'OH della catena di DNA crescente.



Il Sequenziamento di seconda generazione o Next Generation Sequencing (NGS)

LA GENERAZIONE DEI CLUSTER E IL SEQUENZIAMENTO

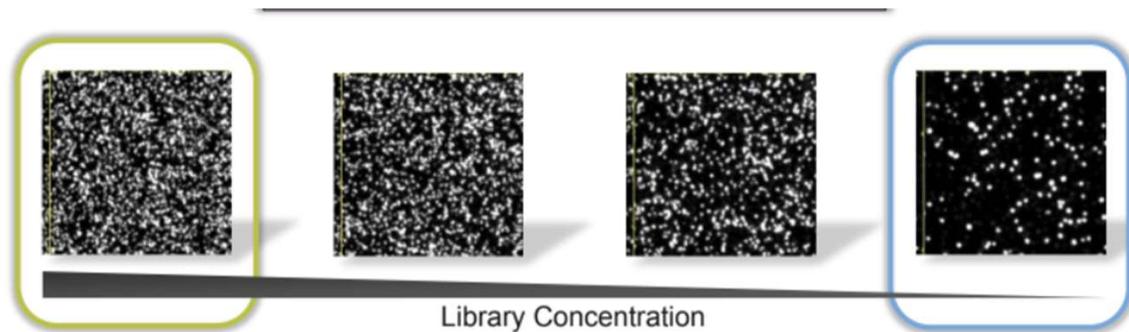
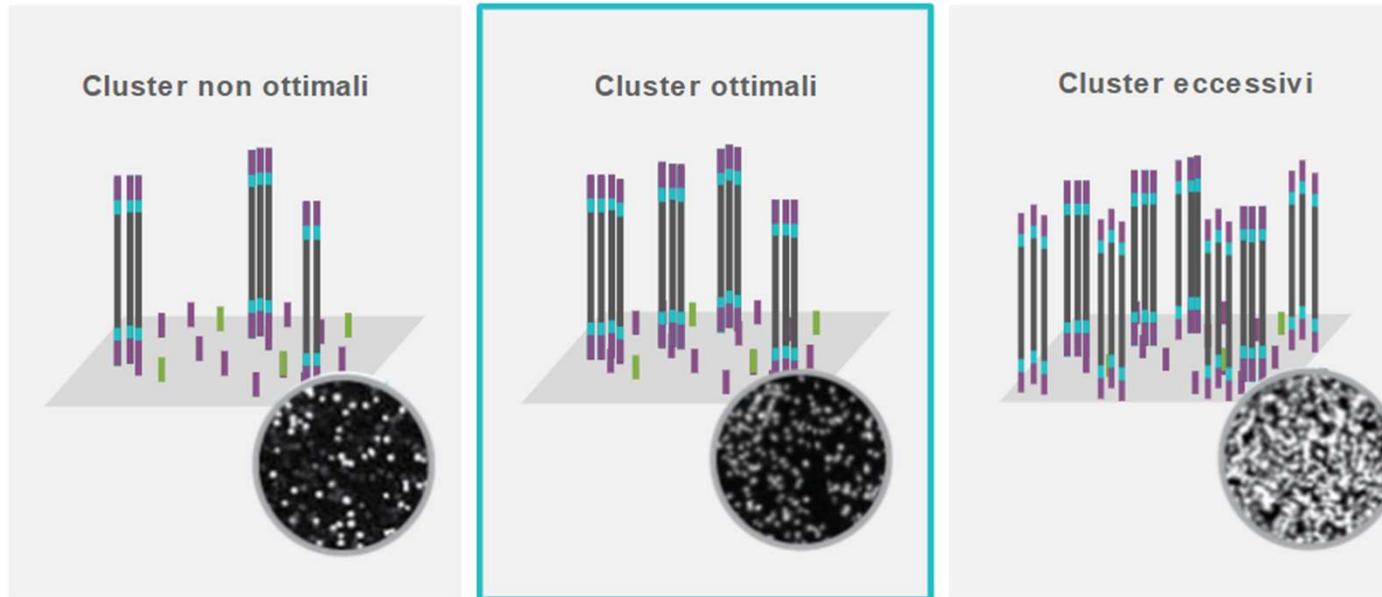


illumina®

Il Sequenziamento di seconda generazione o Next Generation Sequencing (NGS)

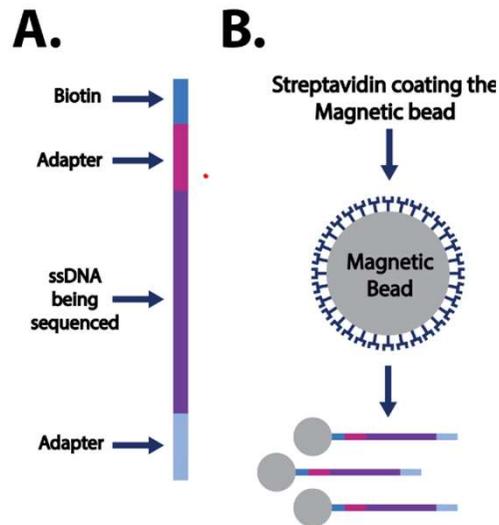
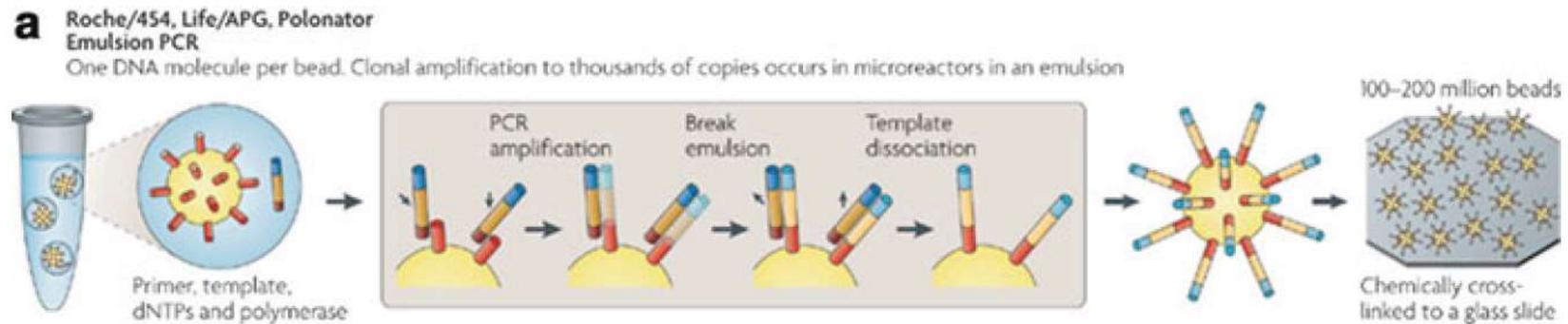
LA GENERAZIONE DEI CLUSTER E IL SEQUENZIAMENTO

La generazione dei cluster ottimizzata sulla cella a flusso determina la qualità dei dati e la resa complessiva dei dati. Per ottenere la densità dei cluster ottimale eseguire la quantificazione delle librerie prima del sequenziamento



Il Sequenziamento di seconda generazione o Next Generation Sequencing (NGS)

LA GENERAZIONE DEI CLUSTER E IL SEQUENZIAMENTO



IonTorrent: la libreria è sottoposta ad un processo di **PCR ad emulsione**, nella quale piccole gocce di soluzione oleosa creano un insieme di micro-reattori di PCR: ogni singola molecola di libreria entra in contatto con una biglia magnetica ricoperta di primers complementari agli adattatori e, al termine della PCR, le biglie saranno ricoperte di ssDNA monoclonali pronti per il sequenziamento

Il Sequenziamento di seconda generazione o Next Generation Sequencing (NGS)

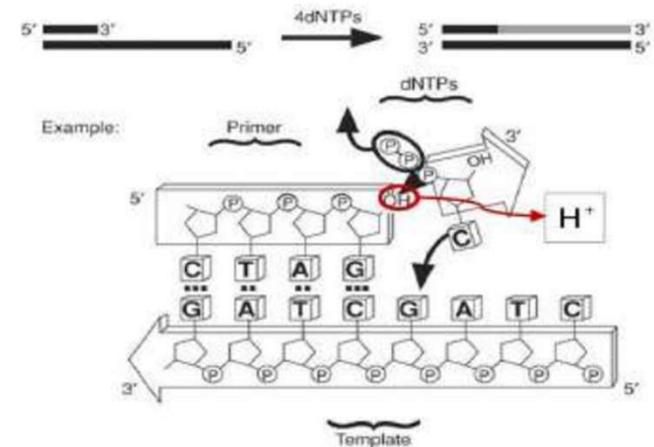
LA GENERAZIONE DEI CLUSTER E IL SEQUENZIAMENTO

IonTorrent: durante il processo di aggiunta di un nucleotide, viene liberato un **protone**, il quale genera **una variazione di pH** che viene rilevata dallo strumento come una variazione di potenziale.

Le basi azotate vengono dispensate una per volta, e la quantità di protoni rilasciati è proporzionale al numero di basi presenti.

Entrambe le metodiche produrranno un insieme di sequenze di DNA che prendono il nome di READS.

ion torrent
by life technologies™



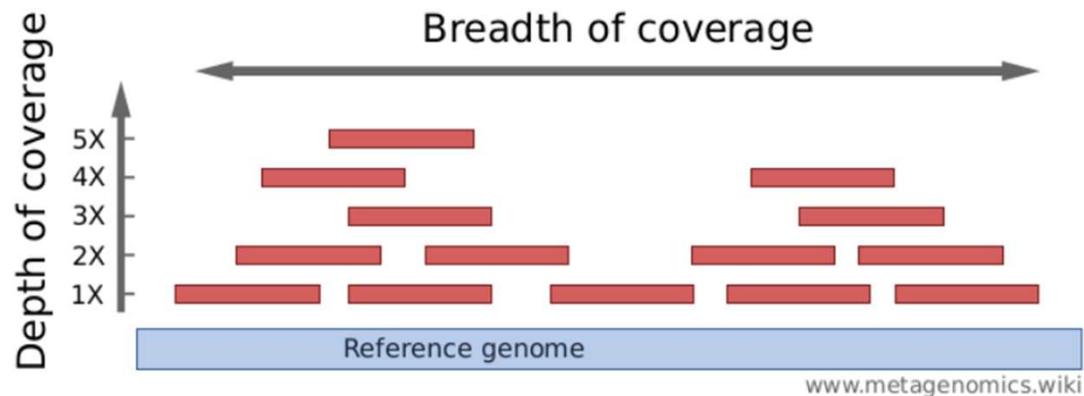
Il Sequenziamento di seconda generazione o Next Generation Sequencing (NGS)

LA GENERAZIONE DEI CLUSTER E IL SEQUENZIAMENTO

Reads: sequenza ottenuta da un singolo cluster originato a partire da un frammento di DNA o cDNA

Coverage: % di copertura del target di interesse

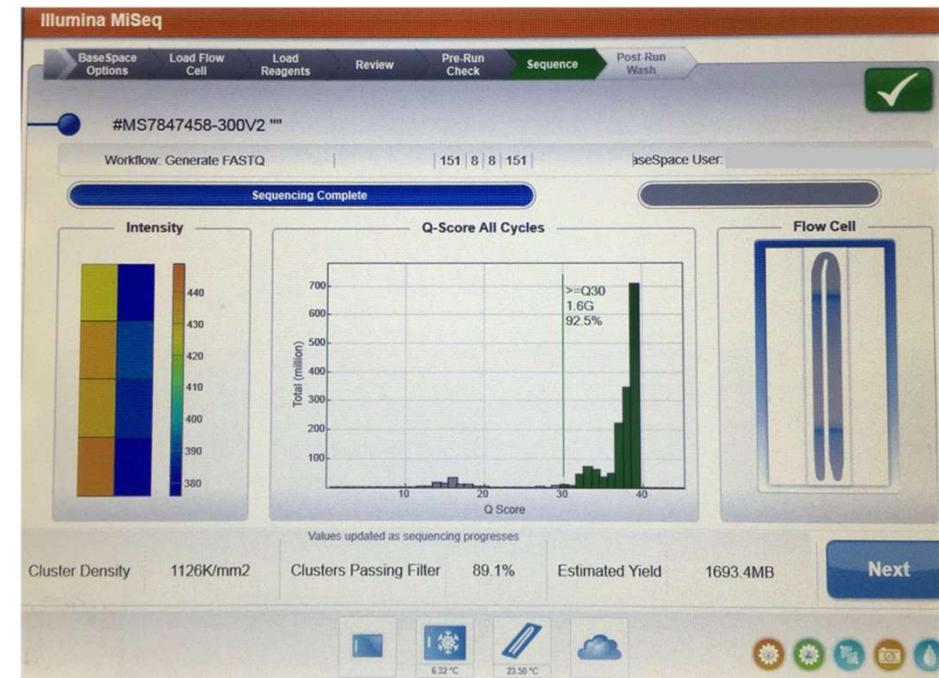
Depth Coverage: numero di volte in cui una base del target di interesse viene letta in maniera indipendente. Indicata con una X: ex. 500X coverage depth



Il Sequenziamento di seconda generazione o Next Generation Sequencing (NGS)

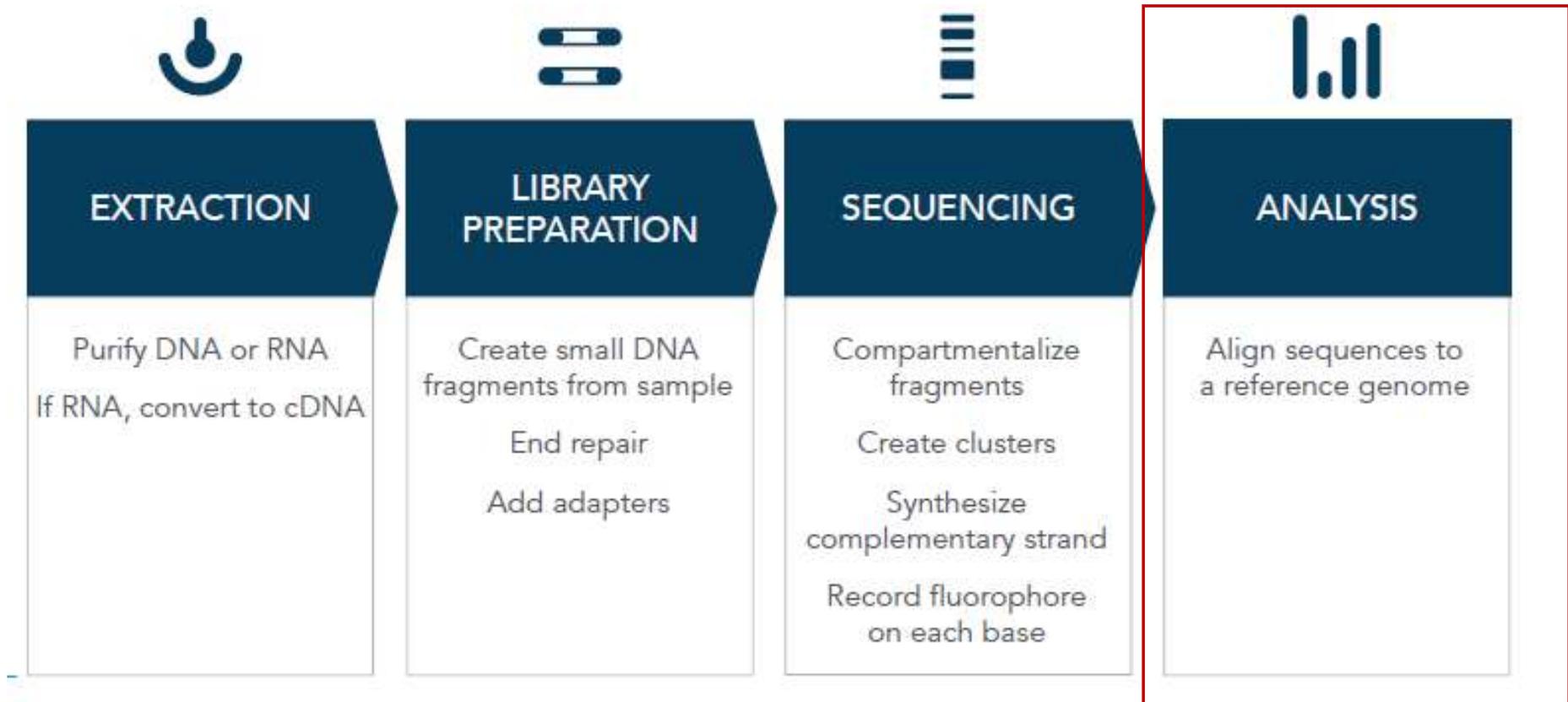
LA GENERAZIONE DEI CLUSTER E IL SEQUENZIAMENTO

- **Metriche** di qualità a livello di **run**:
 - cluster density
 - cluster passing filter
 - numero di reads
 - qualità media delle basi (Q score)
 - Throughput
- **Controllo positivo**: per valutare la qualità della corsa: libreria di controllo di PhiX. Da includere in ogni run → consente di valutare il tasso di errore
- **Controllo negativo**: → consente per valutare la cross-contaminazione tra una run e l'altra (possibile, ma contenuta)
- **Metriche** di qualità per **campione**:
 - dimensione media delle reads
 - numero di reads
 - coverage stimato :oltre il 20X
 - qualità media delle basi (Q score). oltre il 30



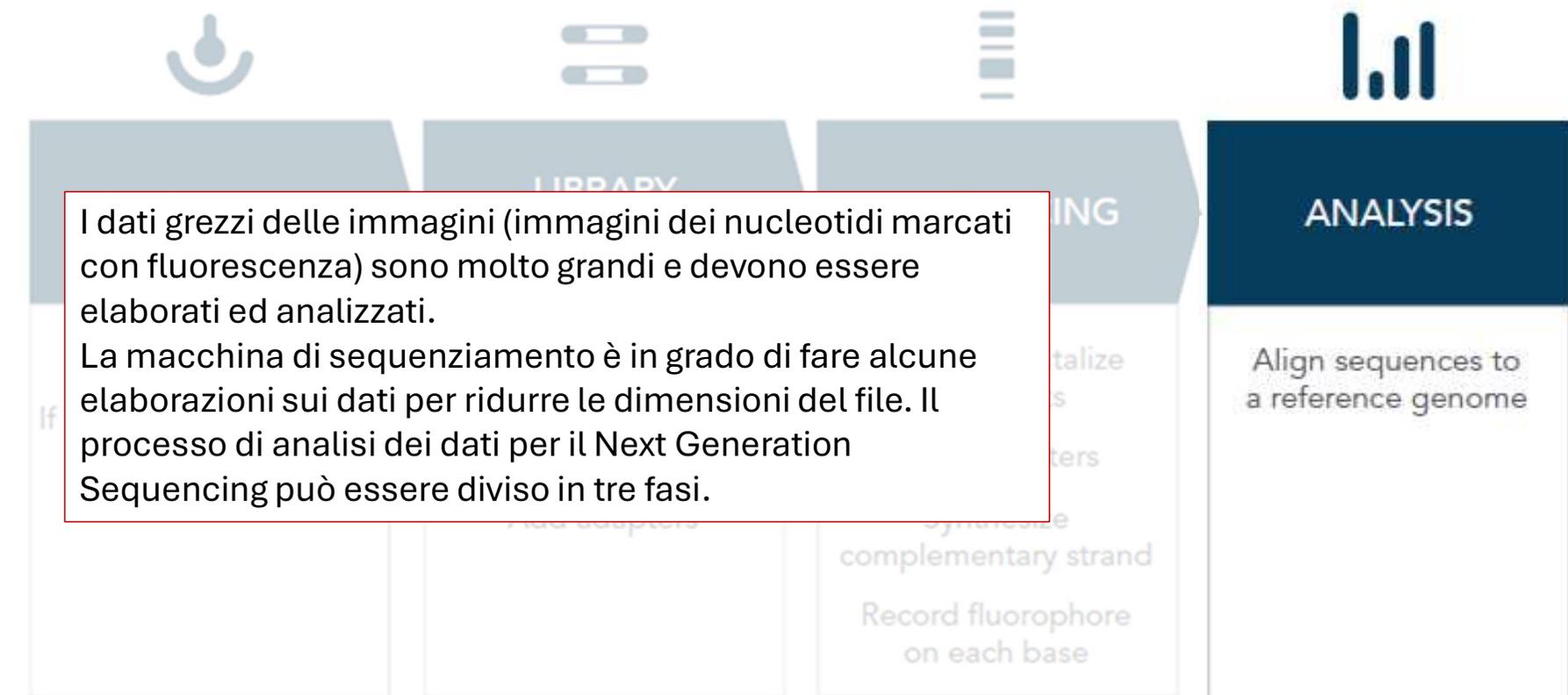
Il Sequenziamento di seconda generazione o Next Generation Sequencing (NGS)

Le piattaforme più diffuse integrano nel flusso di lavoro dell'NGS molti dei passaggi critici, tra cui la (1) **preparazione del campione o della libreria**, (2) **la generazione dei cluster**, e il sequenziamento e (3) **l'analisi dei dati**.



Il Sequenziamento di seconda generazione o Next Generation Sequencing (NGS)

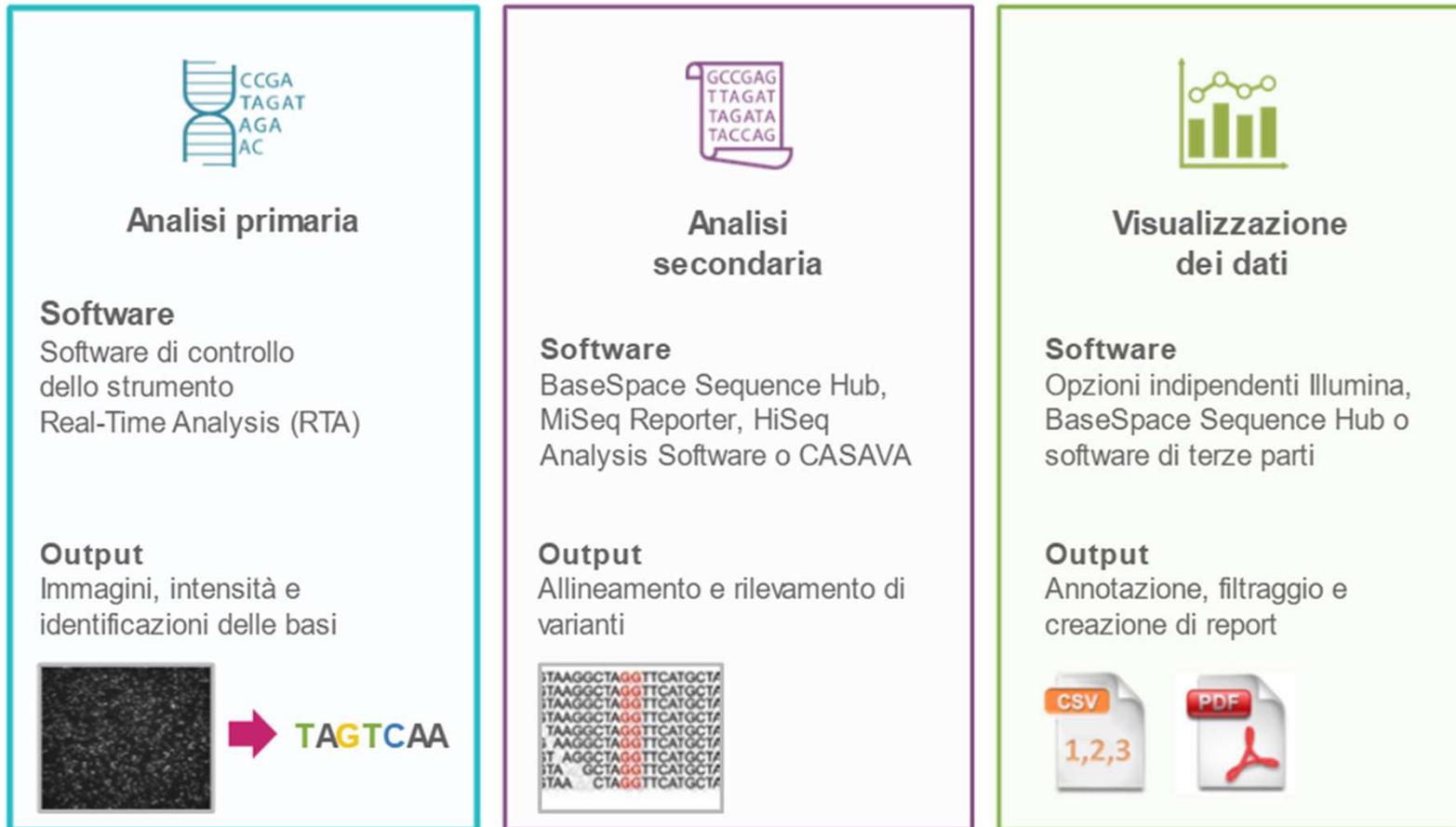
Le piattaforme più diffuse integrano nel flusso di lavoro dell'NGS molti dei passaggi critici, tra cui la (1) **preparazione del campione o della libreria**, (2) **la generazione dei cluster**, e il sequenziamento e (3) **l'analisi dei dati**.



Il Sequenziamento di seconda generazione o Next Generation Sequencing (NGS)

ANALISI DEI DATI

Questo processo richiede l'intervento di un bioinformatico esperto o di un software che lo possa sostituire



Il Sequenziamento di seconda generazione o Next Generation Sequencing (NGS)

ANALISI DEI DATI

	Analisi Bioinformatica di I livello	Analisi Bioinformatica di II livello	Analisi Bioinformatica di III livello
Demultiplexing			
Produzione dei file FastQ			
Allineamento			
Chiamata delle varianti			
Annotazione			
Priorizzazione delle varianti			

- **Analisi primaria:** → generazione delle sequenze (reads) e la valutazione della loro qualità.

Include tutti i passi necessari per **nominare o identificare ogni base**. Oltre a identificare le basi, la macchina di sequenziamento assegna anche un **punteggio di qualità per ciascuna delle basi**. I risultati sono comunemente salvati come un file FASTQ, contenente:

1. gli identificatori,
2. i nucleotidi assegnati (A, G, T o C) chiamati anche «letture»
3. il punteggio di qualità Phred associato.

Il Sequenziamento di seconda generazione o Next Generation Sequencing (NGS)

ANALISI DEI DATI

Se a un nucleotide è assegnata la categoria N, questo significa che la macchina non può determinare il nucleotide esatto.

Il punteggio di **qualità Phred** si riferisce **alla probabilità di un'errata assegnazione della base**. L'analisi primaria è tipicamente eseguita nella macchina di sequenziamento automaticamente dopo ogni ciclo.

Quality score (Phred score): è una misura della qualità con cui viene chiamato un nucleotide in una data posizione (si traduce nella percentuale di errore).

Table 1: Quality Scores and Base Calling Accuracy

Phred Quality Score	Probability of Incorrect Base Call	Base Call Accuracy
10	1 in 10	90%
20	1 in 100	99%
30	1 in 1,000	99.9%
40	1 in 10,000	99.99%
50	1 in 100,000	99.999%

Il Sequenziamento di seconda generazione o Next Generation Sequencing (NGS)

ANALISI DEI DATI

	Analisi Bioinformatica di I livello	Analisi Bioinformatica di II livello	Analisi Bioinformatica di III livello
Demultiplexing			
Produzione dei file FastQ			
Allineamento			
Chiamata delle varianti			
Annotazione			
Prioritizzazione delle varianti			

- **Analisi secondaria:** viene eseguita dopo l'analisi primaria. Quando si vogliono sequenziare diversi campioni insieme nello stesso ciclo (per esempio appartenenti a pazienti o esperimenti diversi) è possibile assegnare **una specifica etichetta a ciascuno di essi**. L'**etichetta**, nota anche come codice a barre, è una breve sequenza di DNA che viene aggiunta all'**adattatore** per differenziare le letture di ogni campione. Anche l'etichetta verrà sequenziata, e sarà grazie all'identificazione della sequenza specifica dell'adattatore di ogni campione che è possibile separare gli uni dagli altri. Questa tecnica è anche **chiamata multiplexing e ha il grande vantaggio di abbassare il costo di sequenziamento e ottenere un campione più grande**. Il primo passo da effettuare prima di eseguire l'analisi secondaria è quello di tagliare l'etichetta e gli adattatori, in quanto queste sequenze non hanno un significato biologico.

Il Sequenziamento di seconda generazione o Next Generation Sequencing (NGS)

ANALISI DEI DATI

	Analisi Bioinformatica di I livello	Analisi Bioinformatica di II livello	Analisi Bioinformatica di III livello
Demultiplexing			
Produzione dei file FastQ			
Allineamento			
Chiamata delle varianti			
Annotazione			
Prioritizzazione delle varianti			

Lo scopo principale dell'analisi secondaria è quello di assemblare tutte le brevi sequenze di DNA (chiamate anche letture) per poter **interpretare i dati di sequenziamento**, attraverso **l'allineamento delle reads** prodotte **ad un genoma di riferimento** (il genoma attualmente utilizzato è il GCRh37 o hg19).

Prima di questo riassettaggio, le letture "grezze" dalla macchina sono spesso valutate e filtrate per qualità per produrre i migliori risultati, rimuovendo le letture che hanno un basso punteggio di qualità Phred.

Quando il riassettaggio viene eseguito da zero senza alcun genoma di riferimento, si parla di assemblaggio de novo. Tuttavia, quando c'è un genoma di riferimento disponibile, il processo è molto più semplice perché possiamo semplicemente allineare tutte le letture al genoma di riferimento.

Il Sequenziamento di seconda generazione o Next Generation Sequencing (NGS)

ANALISI DEI DATI

	Analisi Bioinformatica di I livello	Analisi Bioinformatica di II livello	Analisi Bioinformatica di III livello
Demultiplexing			
Produzione dei file FastQ			
Allineamento			
Chiamata delle varianti			
Annotazione			
Prioritizzazione delle varianti			

Normalmente avremmo diverse letture che mappano la stessa area del genoma e sono spesso indicate come "**profondità di lettura**". La profondità di lettura misura quante volte quell'area è coperta da letture diverse; per esempio, una profondità di lettura di 10 implica che ci sono 10 letture che si sovrappongono nella stessa area genomica.

Infine è eseguita l'**identificazione delle varianti**, che consiste in un processo di determinazione accurata delle **variazioni (o differenze) tra un campione e il genoma di riferimento**. Queste possono essere sotto forma di varianti di un singolo nucleotide (SNV), piccoli inserimenti o rimozioni (chiamati indel), o varianti strutturali più grandi di categorizzazioni come le trasversioni, le traslocazioni e le varianti del numero di copie.

Il Sequenziamento di seconda generazione o Next Generation Sequencing (NGS)

ANALISI DEI DATI

	Analisi Bioinformatica di I livello	Analisi Bioinformatica di II livello	Analisi Bioinformatica di III livello
Demultiplexing			
Produzione dei file FastQ			
Allineamento			
Chiamata delle varianti			
Annotazione			
Prioritizzazione delle varianti			

- **Analisi terziaria:** consiste nell'interpretazione del significato delle varianti

Il Sequenziamento di seconda generazione o Next Generation Sequencing (NGS)

ANALISI DEI DATI

Le varianti identificate vengono **annotate** con l'effetto predetto sulla proteina e con i dati presenti nei database pubblici.

Annotazione con database pubblici

- Clinici (OMIM, ClinVar, MedGen, CIViC, ICGC Simple Somatic Mutations);
- Frequenza nella popolazione (1000 Genomes project, NHLBI ESP6500, ExAC, gnomAD, PGVD);
- Annotazione funzionale (RefSeq, GWAS catalog, GTEx, ACMG IF, ExAC LoF, dbNFSP, ENCODE)

Le varianti somatiche annotate dovrebbe essere prioritizzate, ovvero filtrate secondo specifici parametri e riportate considerando:

- la patogenicità : varianti somatiche patogenetiche o potenzialmente tali nei database clinici oncologici;
- la frequenza della variante somatica nella popolazione;
- Varianti non riportate nei database clinici ma con conseguenze funzionali predette clinicamente significative.

Il Sequenziamento di seconda generazione o Next Generation Sequencing (NGS)

ANALISI DEI DATI

Guideline > [Genet Med. 2015 May;17\(5\):405-24. doi: 10.1038/gim.2015.30. Epub 2015 Mar 5.](#)

Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology

[Sue Richards](#)¹, [Nazneen Aziz](#)², [Sherri Bale](#)³, [David Bick](#)⁴, [Soma Das](#)⁵, [Julie Gastier-Foster](#)⁶, [Wayne W Grody](#)⁷, [Madhuri Hegde](#)⁸, [Elaine Lyon](#)⁹, [Elaine Spector](#)¹⁰, [Karl Voelkerding](#)⁹, [Heidi L Rehm](#)¹¹, [ACMG Laboratory Quality Assurance Committee](#)

Affiliations + expand

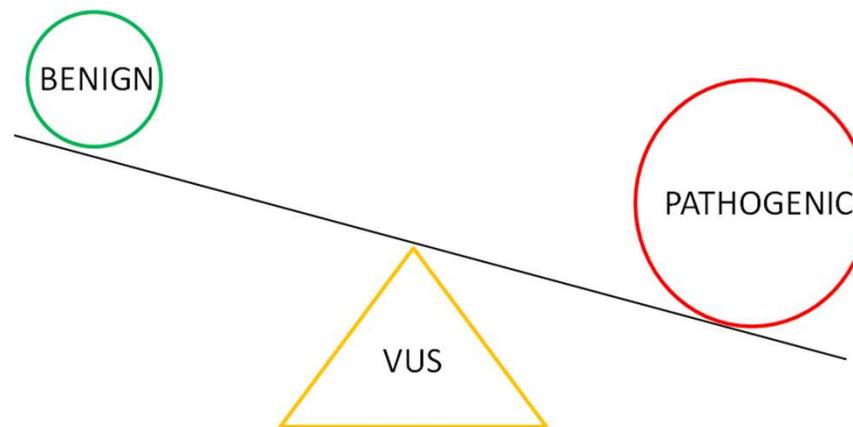
PMID: 25741868 PMCID: PMC4544753 DOI: 10.1038/gim.2015.30

Il Sequenziamento di seconda generazione o Next Generation Sequencing (NGS)

ANALISI DEI DATI



- varianti benigne
- varianti verosimilmente benigne
- varianti di incerto significato (VUS)
- varianti verosimilmente patogenetiche
- varianti patogenetiche



Guideline | Genet Med. 2015 May;17(5):405-24. doi: 10.1038/gim.2015.30. Epub 2015 Mar 5.

Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology

Sue Richards¹, Nazreen Aziz², Sherril Bale³, David Bick⁴, Soma Das⁵, Julie Gasten-Foster⁶, Wayne W Grody⁷, Madhuri Hegde⁸, Elaine Lyon⁹, Elaine Spector¹⁰, Karl Voelkerding⁹, Heidi L Walim¹¹, ACMG Laboratory Quality Assurance Committee

Affiliations: + expand
PMID: 25741868 | PMCID: PMC4344753 | DOI: 10.1038/gim.2015.30

Il Sequenziamento di seconda generazione o Next Generation Sequencing (NGS)

ANALISI DEI DATI

	Benign		Pathogenic			
	Strong	Supporting	Supporting	Moderate	Strong	Very Strong
Population Data	MAF is too high for disorder <i>BA1/BS1</i> OR observation in controls inconsistent with disease penetrance <i>BS2</i>			Absent in population databases <i>PM2</i>	Prevalence in affecteds statistically increased over controls <i>PS4</i>	
Computational And Predictive Data		Multiple lines of computational evidence suggest no impact on gene /gene product <i>BP4</i> Missense in gene where only truncating cause disease <i>BP1</i> Silent variant with non predicted splice impact <i>BP7</i>	Multiple lines of computational evidence support a deleterious effect on the gene /gene product <i>PP3</i>	Novel missense change at an amino acid residue where a different pathogenic missense change has been seen before <i>PM5</i> Protein length changing variant <i>PM4</i>	Same amino acid change as an established pathogenic variant <i>PS1</i>	Predicted null variant in a gene where LOF is a known mechanism of disease <i>PVS1</i>
Functional Data	Well-established functional studies show no deleterious effect <i>BS3</i>		Missense in gene with low rate of benign missense variants and path. missenses common <i>PP2</i>	Mutational hot spot or well-studied functional domain without benign variation <i>PM1</i>	Well-established functional studies show a deleterious effect <i>PS3</i>	
Segregation Data	Non-segregation with disease <i>BS4</i>		Co-segregation with disease in multiple affected family members <i>PP1</i>	Increased segregation data →		
De novo Data				<i>De novo</i> (without paternity & maternity confirmed) <i>PM6</i>	<i>De novo</i> (paternity & maternity confirmed) <i>PS2</i>	
Allelic Data		Observed in <i>trans</i> with a dominant variant <i>BP2</i> Observed in <i>cis</i> with a pathogenic variant <i>BP2</i>		For recessive disorders, detected in <i>trans</i> with a pathogenic variant <i>PM3</i>		
Other Database		Reputable source w/out shared data = benign <i>BP6</i>	Reputable source = pathogenic <i>PP5</i>			
Other Data		Found in case with an alternate cause <i>BP5</i>	Patient's phenotype or FH highly specific for gene <i>PP4</i>			

Guideline | Genet Med. 2015 May;17(5):405-24. doi: 10.1038/gim.2015.30. Epub 2015 Mar 5.

Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology

Sue Richards¹, Nazreen Aziz², Sherril Bale³, David Bick⁴, Soma Das⁵, Julie Gasten-Foster⁶, Wayne W Grody⁷, Madhuri Hegde⁸, Elaine Lyon⁹, Elaine Spector¹⁰, Karl Voelkerding⁹, Heidi L Walim¹¹, ACMG Laboratory Quality Assurance Committee

Affiliations: + expand
PMID: 25741668 | PMCID: PMC4344753 | DOI: 10.1038/gim.2015.30

Pros and Cons of NGS *versus* Individual Molecular Tests

Pros and Cons of Next Generation Sequencing (NGS)

Pros

- Multiple genes can be tested simultaneously for mutations or disease markers.
- Previously unknown disease markers may be discovered.
- Additional information may help with the development of new and more effective treatments.

Cons

- NGS takes longer (2-3 weeks) and is more expensive (thousands of dollars) than testing for a single marker.
- NGS generates a large amount of complex data of which the results may be difficult to interpret.

Ask your doctor if NGS testing will benefit your care.

Pros and Cons of NGS vs Individual Molecular Tests

Pros of NGS:

- ✓ Possibilità di analizzare più geni contemporaneamente: risparmio di tempo, soldi, materiale.
- ✓ Può dare informazioni anche su altri geni potenzialmente targettabili.
- ✓ Maggiori informazioni che possono aumentare la conoscenza su determinate tipologie di tumore.

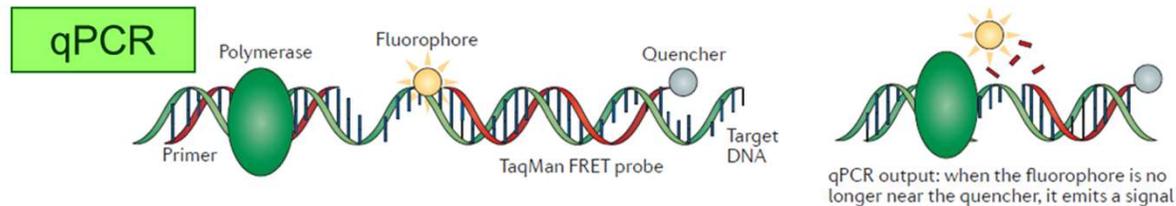
Cons of NGS:

- ✓ Tempi più lunghi almeno una settimana.
- ✓ Costi dell'analisi
- ✓ Maggiore difficoltà interpretazione dati

Alternatives to NGS analysis



Competitive alternatives to NGS in the clinic?



Incorporation of a double-stranded DNA-specific dye or by the release of a TaqMan FRET (fluorescence resonance energy transfer) probe. Point-of-care applications: genotyping, gene expression analysis, CNV assays and pathogen detection.

Dalla PCR a

PCR: Types and Applications

Hot start PCR 

RT-PCR 

qPCR 

Digital PCR 

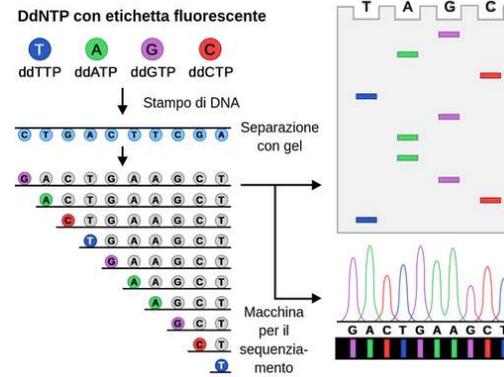
Polymerase Chain Reaction (PCR)

- 1 Denaturation
 - 2 Annealing
 - 3 Extension
- PCR

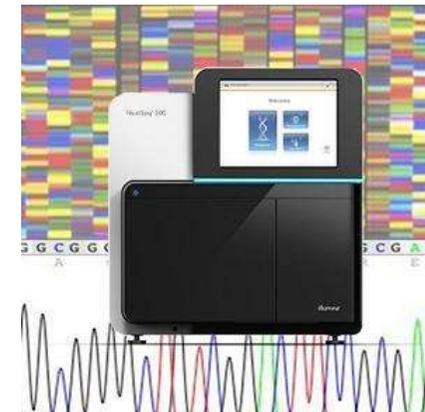
Applications

- **Diagnosis:** Detecting infectious diseases
- **Research:** Amplifying DNA for cloning
- **Forensics:** Analyzing DNA in criminal investigations and paternity testing.
- **Drug Development:** Screening for drug targets, drug efficacy, and pharmacogenomics.
- **Environmental Monitoring:** Identifying microorganisms in soil, water, and air samples.

Sequenziamento Sanger



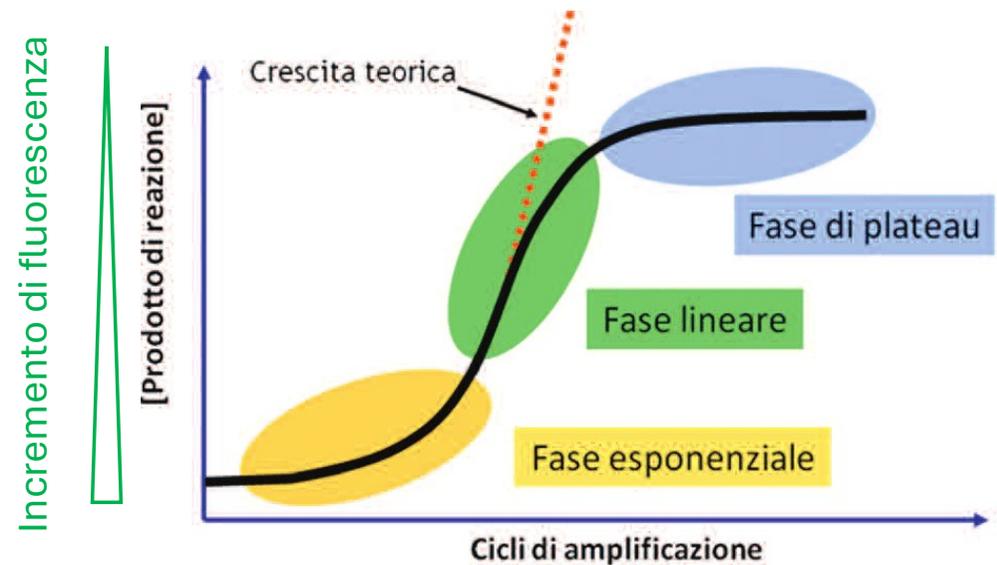
Sequenziamento NGS



PCR Real Time

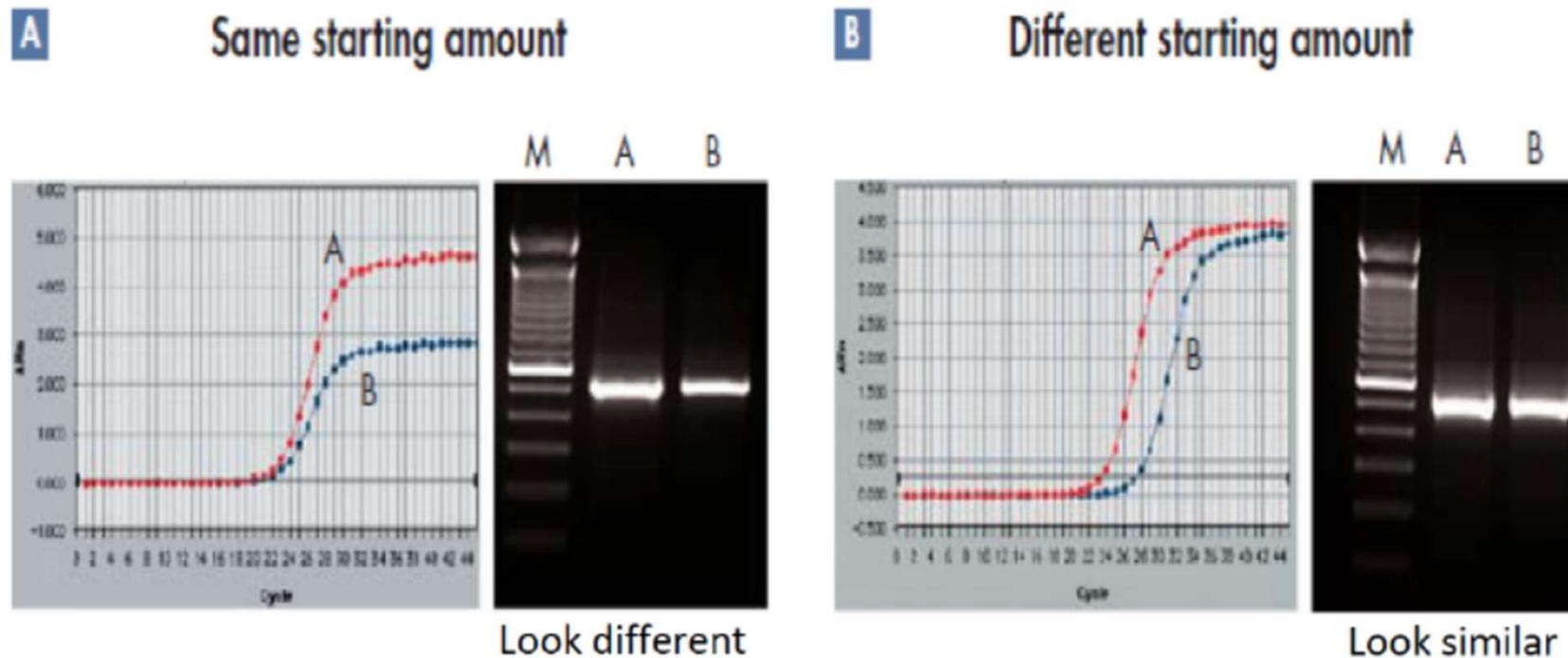
La PCR real-time (conosciuta anche come qPCR, ovvero Quantitative PCR, da non confondere con RT-PCR) è un metodo che simultaneamente **amplifica** e **quantifica il DNA**.

- Rilevamento della fluorescenza associata all'amplificazione durante ogni ciclo di PCR
- Il prodotto di PCR non viene analizzato su gel di agarosio
- Analisi del prodotto di fluorescenza tramite computer



PCR Real Time con sonda Taqman

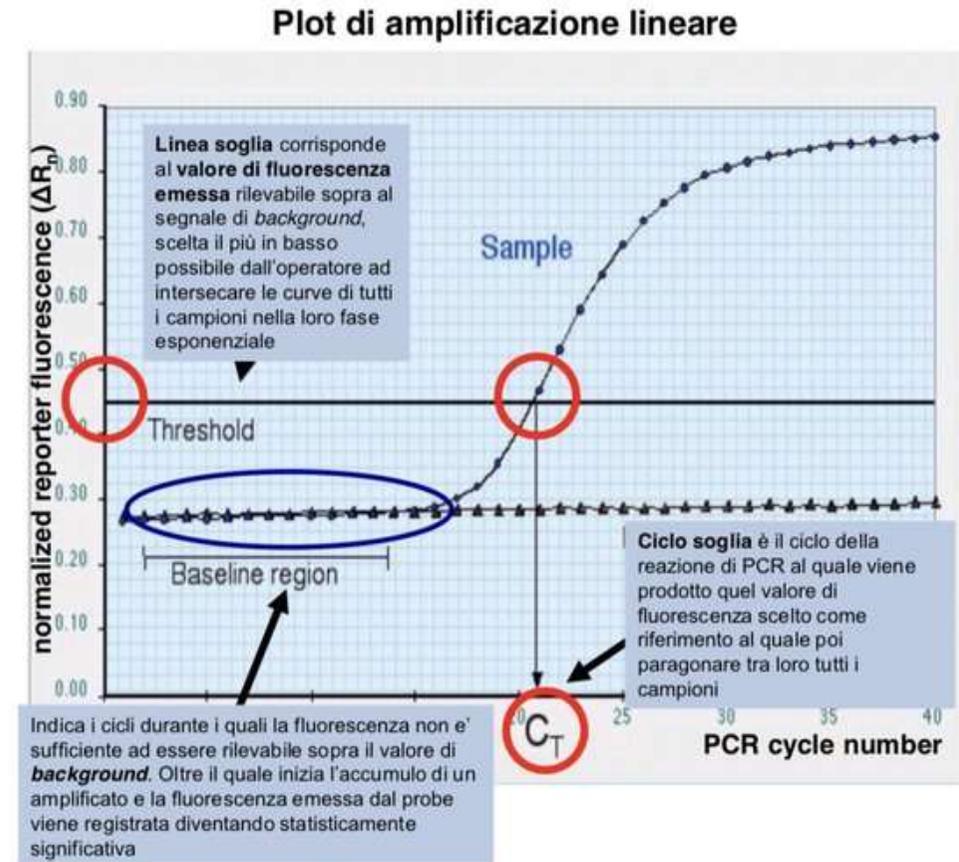
Analisi quantitativa mediante PCR end-point



Nella PCR convenzionale, il prodotto del DNA amplificato viene rilevato in un'analisi del punto finale. Nella PCR in tempo reale, l'accumulo del prodotto di amplificazione viene misurato man mano che la reazione progredisce, in tempo reale, con la quantificazione del prodotto dopo ogni ciclo.

PCR Real Time con sonde Taqman

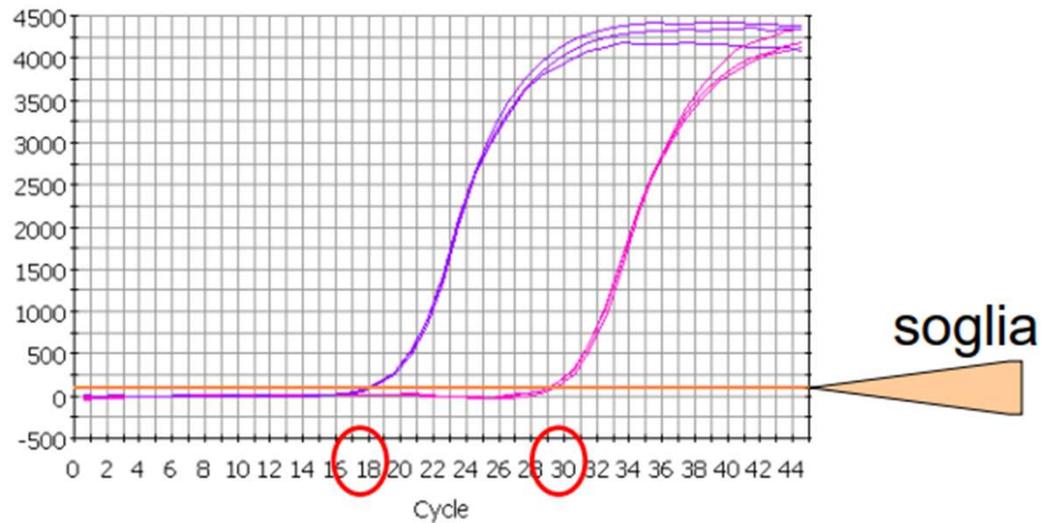
- La Real-time qPCR è la più sensibile ed affidabile tecnica per la quantificazione di acidi nucleici (qPCR, RT-qPCR)
- Si basa sull'**identificazione** e sulla **quantificazione** in tempo reale della **fluorescenza emessa** da una molecola reporter
- Questo avviene durante la fase di accumulo (fase esponenziale) del prodotto di PCR e ad ogni ciclo di amplificazione, consentendo così di monitorare l'amplificazione durante la fase iniziale e soprattutto durante la **fase esponenziale** della reazione di PCR



PCR Real Time

PCR Quantitativa - Real-Time PCR → principio

Maggiore è il numero delle “molecole stampo” presenti all’inizio della reazione e minore sarà il numero di cicli necessari per raggiungere un determinato valore minimo di ammontare di prodotto (Cycle threshold (Ct) - ciclo soglia).



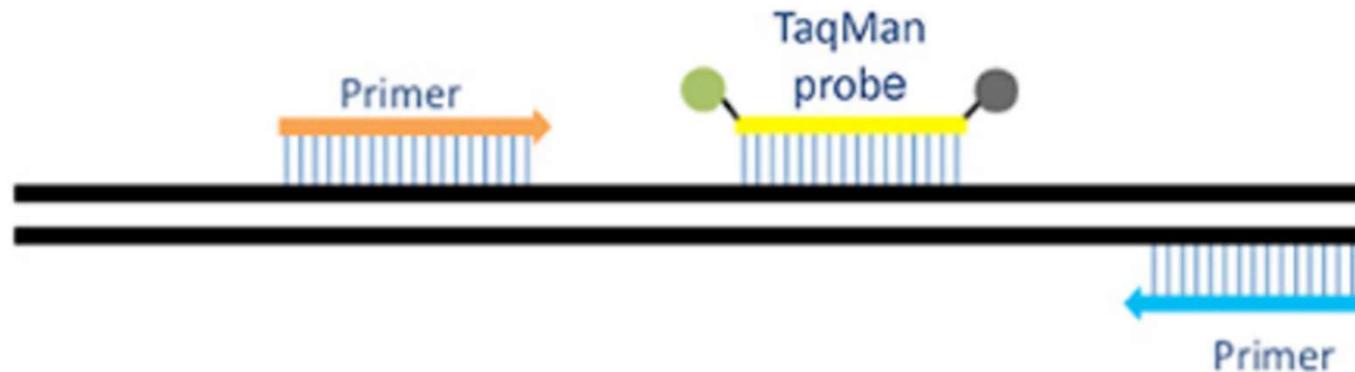
PCR Real Time

La chimica della Real Time PCR

- DNA binding dyes: **SYBR green**  sequenza indipendente
- Fluorescent PCR Primer- and Probe-Based Chemistries:
 - Hydrolysis (**TaqMan**) probes o 5' nuclease assay
 - Molecular beacons
 - Dual hybridization probes
 - Eclipse probes
 - Amplifluor assays
 - Scorpions PCR primers
 - LUX PCR primers
 - Qzyme PCR primers sequenza specifiche

PCR Real Time con sonde Taqman

La **sonda di tipo TaqMan** è un oligonucleotide che, come i primers della PCR, viene disegnato per essere **complementare alla sequenza bersaglio da amplificare**.



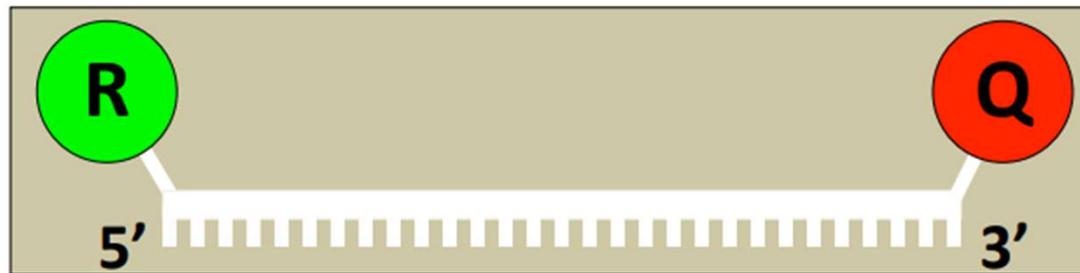
La sonda è disegnata in modo da ibridarsi all'interno del frammento amplificato nella reazione di PCR.

PCR Real Time con sonde Taqman

Rilevazione del prodotto di PCR mediante saggio Taqman

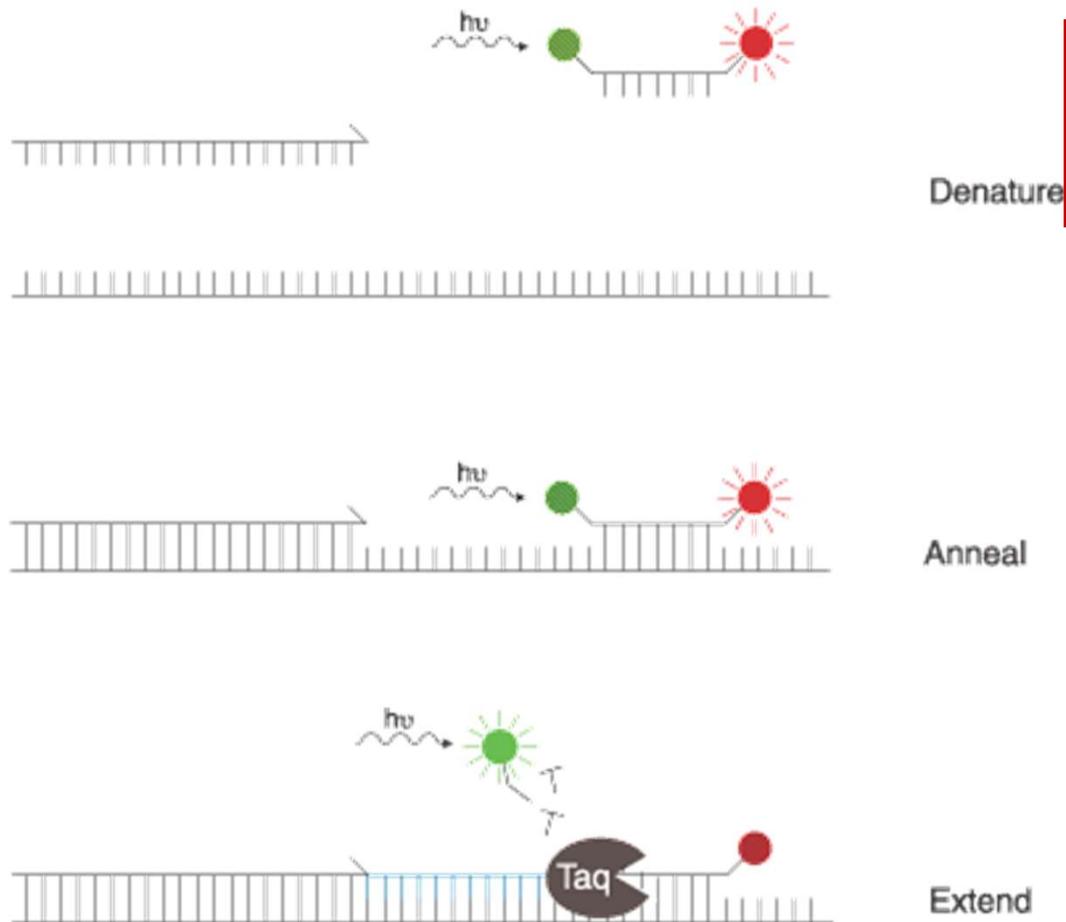
5' REPORTER (R): fluorocromo ad alta energia che emette fluorescenza

3' QUENCHER (Q): fluorocromo a bassa energia che maschera la fluorescenza del reporter



Se **R** e **Q** si trovano vicini, **Q** spegne l'effetto di **R** perché i fotoni di **R** vengono assorbiti da **Q**

PCR Real Time con sonde Taqman



La RT-PCR è attualmente la **metodica single gene gold standard** per l'analisi di mutazioni puntiformi e/o di piccole inserzioni/delezioni su DNA e ctDNA.

L'analisi permette, mediante amplificazione con sonde **sequenza-specifiche marcate con FAM** (target) ed **HEX** (controllo endogeno wild-type), di rilevare basse percentuali di allele mutato, anche in presenza di elevate quantità di DNA genomico wild-type, arrivando ad un limite di rilevazione (limit of detection, LOD) anche inferiore allo 0,5%.

L'amplificazione del **gene di controllo interno** permette di verificare la corretta esecuzione della procedura di amplificazione e l'eventuale presenza di inibitori, che possono causare dei **falsi negativi**.

PCR Real Time con sonda Taqman

Colorectal Cancer	Non-Small Cell Lung Cancer	Thyroid Cancer	Cutaneous Melanoma	Hepatocellular Carcinoma	Cholangiocarcinoma	Breast Cancer	Cervical Cancer	Glioma	Head and Neck Cancer
									
Clinical biomarkers included in EasyPGX® assay portfolio									
KRAS, BRAF, NRAS, DPYD, UGT1A1, MSI, NTRK, PIK3CA	KRAS, BRAF, ALK, ROS1, RET, MET, NTRK, EGFR, MSI	HRAS, KRAS, BRAF, NRAS, RET, NTRK, MSI, PPARG, ALK	BRAF, NTRK, MSI	MSI, NTRK	IDH1-2, MSI, NTRK, DPYD	DPYD, UGT1A1-2, MSI, NTRK, PIK3CA	HPV, MSI, NTRK	IDH1-2, NTRK, MSI, MGMT	MSI, DPYD, HPV, NTRK, HRAS

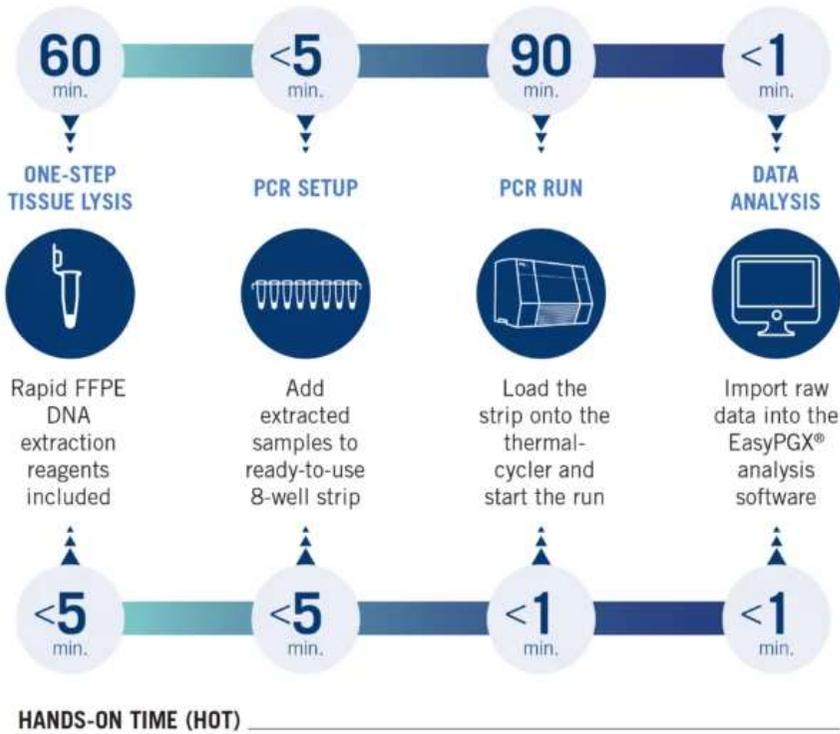
PCR Real Time con sonda Taqman

Tempo di esecuzione 5 min

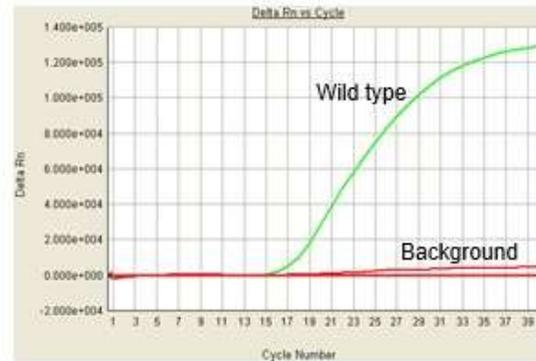
Amplificazione 90 min

Lettura 5 min

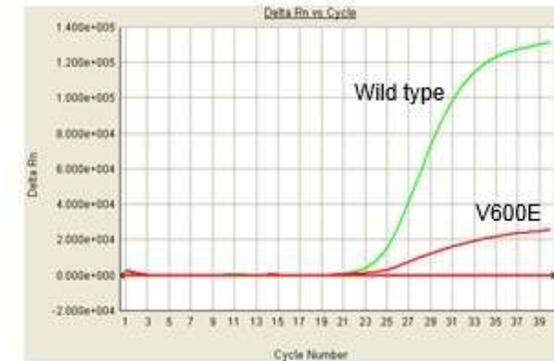
TURNAROUND TIME (TAT)



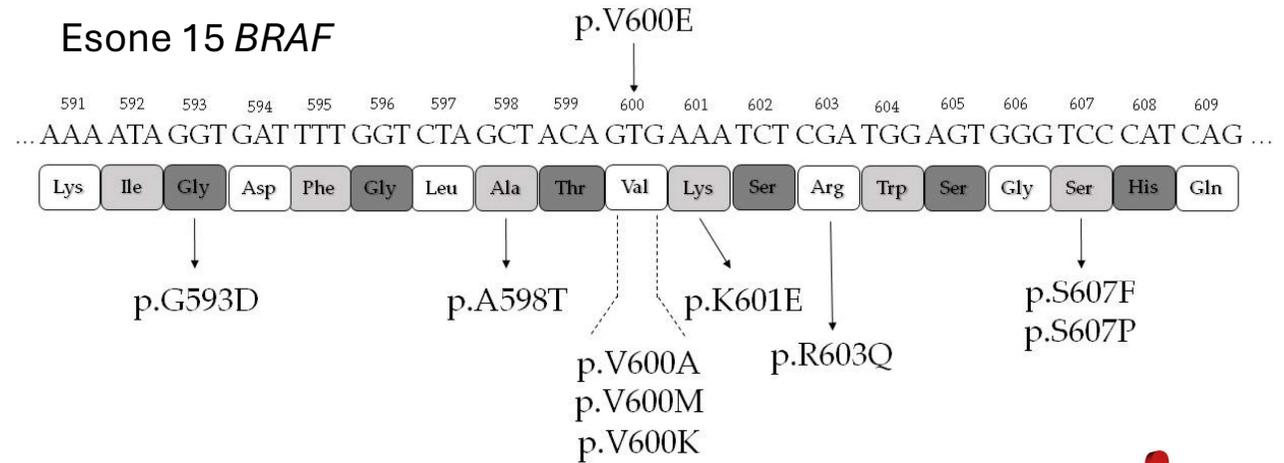
No Mutation Detected



V600E Mutation Detected



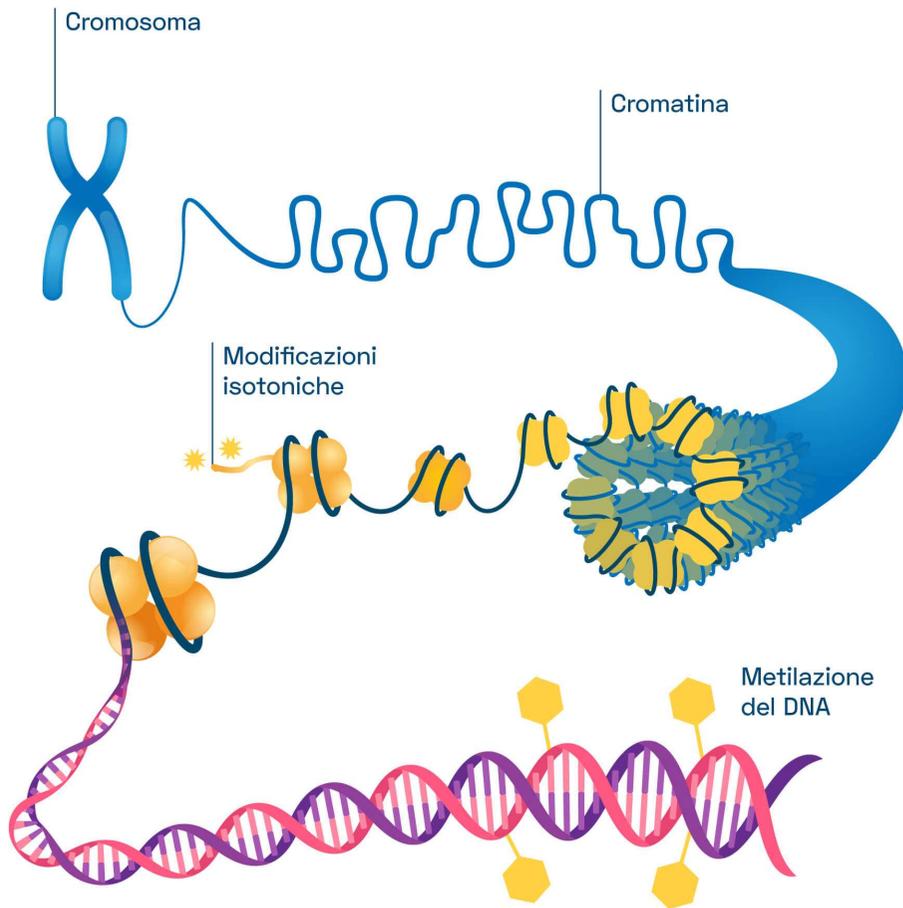
Esone 15 *BRAF*



NGS/Sanger unico campione vedo tutte le mutazioni
qPCR ho bisogno di più saggi uno per ogni mutazione!



Epigenomica – studio dello stato di metilazione del DNA



Metilazione del DNA

Una delle modificazioni epigenetiche più comuni è la **metilazione** del DNA.

La metilazione del DNA consiste nel trasferimento di un **gruppo metilico** dalla **S-adenosilmetionina** ai **residui di citosina** posti in **posizione 5'** nei **dinucleotidi CpG**, mediante una reazione catalizzata da una **DNA metiltransferasi**. Generalmente solo la metilazione dei dinucleotidi CpG raggruppati in cluster di 0,5-3 kb, denominati **isole CpG**, ha una funzione biologica. Le isole CpG sono presenti per lo più nelle regioni promotrici e/o nei primi esoni del 50-60% di tutti i geni umani. A seconda di quanto sono metilati i siti CpG presenti nei promotori, i geni possono essere trascritti o meno.

Epigenomica – studio dello stato di metilazione del DNA

La **metilazione del DNA** a livello degli elementi regolatori dei geni, come i **promotori e gli enhancer**, generalmente **sopprime la loro funzione**, attraverso l'inibizione diretta o indiretta del legame dei fattori di trascrizione. La metilazione, infatti, mantenendo la cromatina in uno stato compatto, stabilizza i cromosomi e reprime l'espressione genica.

L'ipermetilazione, localizzata a livello di specifici promotori, determina il silenziamento dei geni oncosoppressori, degli inibitori degli oncogeni e dei geni coinvolti nella regolazione del ciclo cellulare o nella riparazione del DNA.

Epigenomica – studio dello stato di metilazione del DNA

PROMOTORE DEL GENE *MGMT*

La **farmacoepigenetica** studia la diversa risposta ai farmaci sulla base delle modificazioni epigenetiche, come la metilazione del DNA, le modificazioni istoniche e gli RNA non codificanti.

Esempio: **risposta alla terapia a base di alchilanti in glioblastoma**

La **temozolomide** è approvata per i pazienti affetti da **glioblastoma** (GBM) e la sua efficacia in questa malattia è correlata ad un **biomarcatore predittivo**: la **metilazione del promotore di O6-metilguanina-DNA metiltransferasi (*MGMT*)**.

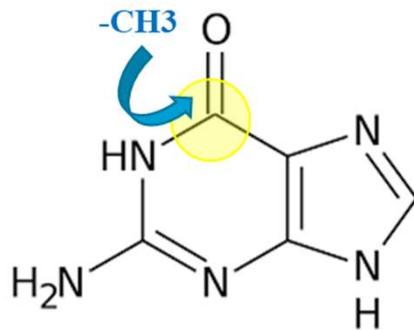
- ✓ La **bassa espressione di *MGMT*** correla con una **umentata sopravvivenza libera** da malattia in pazienti con glioblastoma (GBM) trattati con temozolomide (TMZ)
- ✓ La metilazione del promotore di *MGMT* ha dimostrato di essere un fattore prognostico favorevole nei pazienti con GBM trattati con TMZ: la **sopravvivenza** risulta **umentata** in presenza di **p*MGMT* ipermetilato**.

Lo stato di metilazione del promotore di *MGMT* è il fattore prognostico più forte per l'outcome nei pazienti con glioblastoma alla diagnosi ed è un potente predittore di risposta alla chemioterapia alchilante.

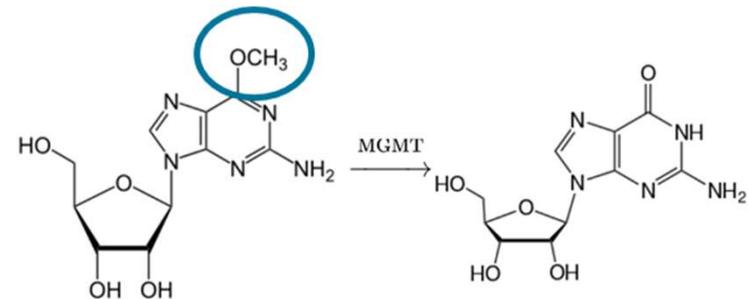
Epigenomica – studio dello stato di metilazione del DNA

PROMOTORE DEL GENE *MGMT*

Temozolomide è un agente alchilante: aggiunge un gruppo metilico, alla guanina



MGMT è un enzima coinvolto nella riparazione dei danni a carico del DNA, è in grado di rimuovere anche i gruppi metilici aggiunti da TMZ

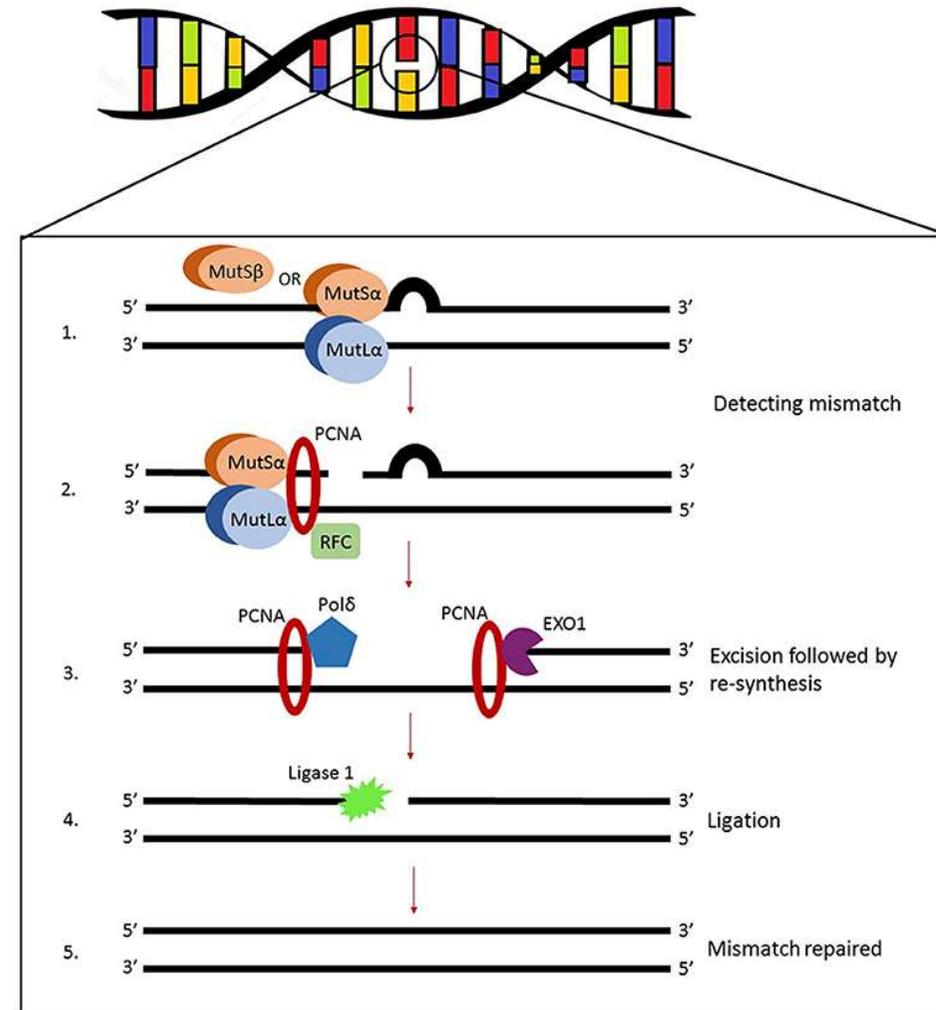


Il grado di metilazione, ottenibile con **pirosequenziamento**, è stato proposto come fattore prognostico indipendente per stratificare i pazienti con GBM trattati con TMZ. Considerato il **9% di metilazione** come cut-off tra tumori metilati e non, i pazienti con una percentuale di metilazione media superiore al 29%, calcolata su 12 siti CpG di *MGMT*, mostravano un esito clinico migliore, con valori di progression-free survival e overall survival significativamente più alti di quelli dei pazienti con metilazione di *MGMT* compresa tra il 9% e il 29%.

Epigenomica – studio dello stato di metilazione del DNA

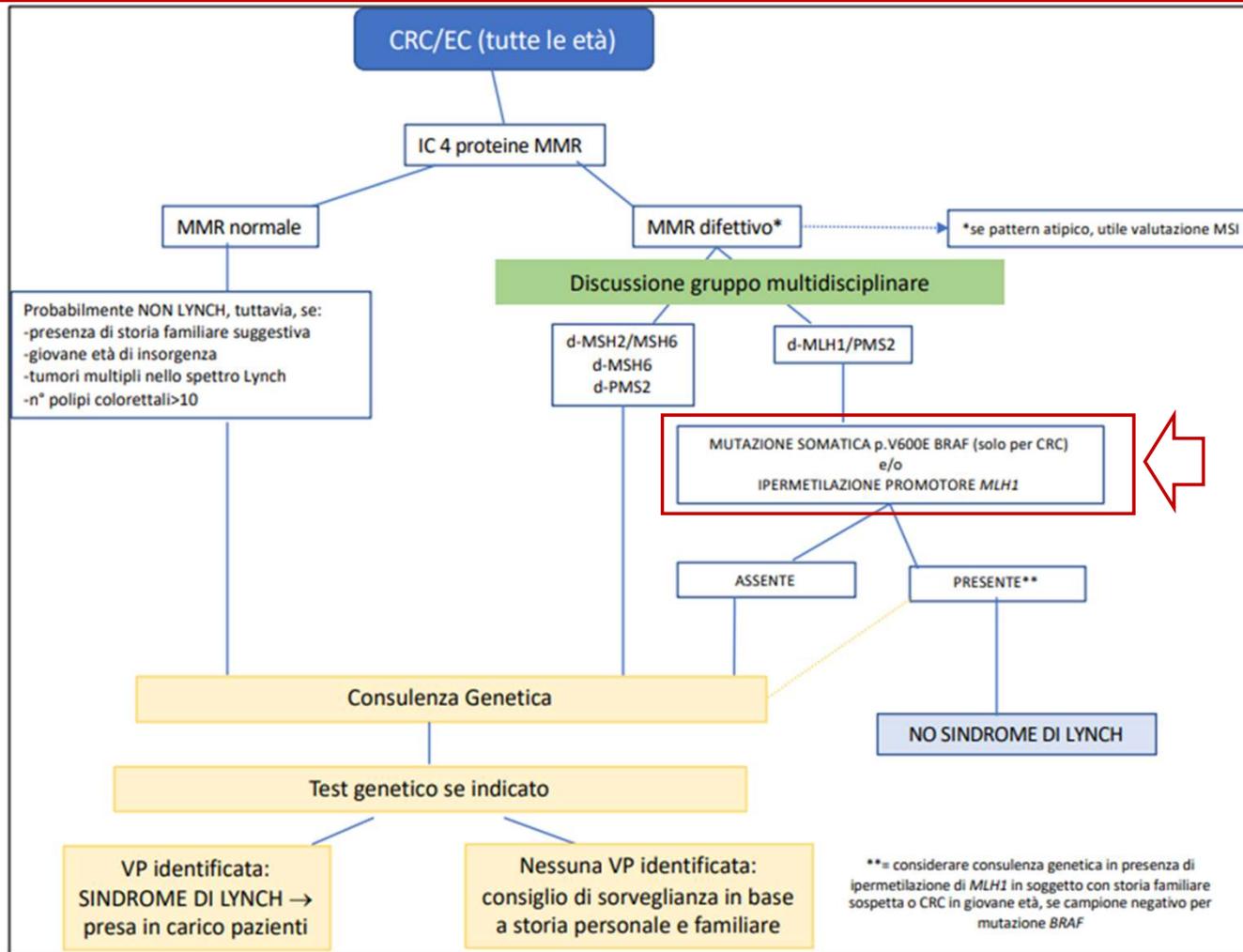
PROMOTORE DEL GENE *MLH1*

- ✓ *MLH1* è un gene che codifica per l'omonima proteina, parte del complesso mismatch repair (MMR)
- ✓ La metilazione del gene porta alla mancata produzione della proteina MLH1 e di conseguenza a un complesso MMR non attivo (*instabilità dei microsatelliti*)
- ✓ Studiare la metilazione del gene *MLH1* è fondamentale per capire l'origine del fenotipo MMRd (*sporadica vs ereditaria/sindrome di Lynch*): MMRd di tipo sporadico sono in genere associate a ipermetilazione del promotore *MLH1*

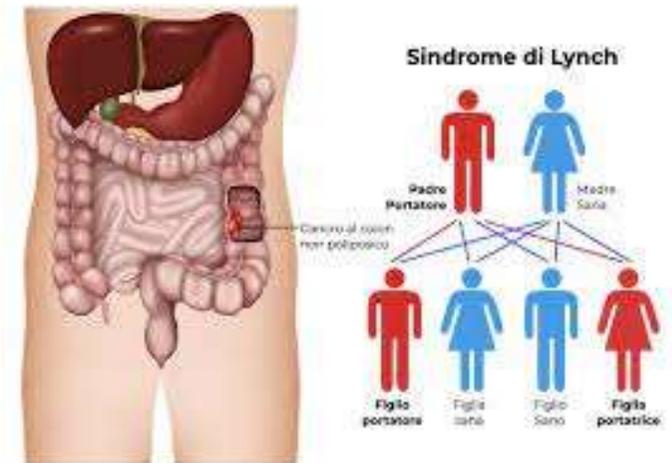


Epigenomica – studio dello stato di metilazione del DNA

PROMOTORE DEL GENE *MLH1*



Studio del profile di metilazione del promotore del gene *MLH1* è fondamentale per il raggiungimento di diagnosi di sindrome di Lynch



Epigenomica – studio dello stato di metilazione del DNA

CONVERSIONE CON SODIO BISOLFITO

GGTCAGTGAC/mCG

↓ Bisulfite conversion
U

Unmethylated Cytosine C is transformed to Uracil U by the bisulfite treatment, whereas methylated Cytosine mC remains unchanged.

GGTUAGTGAU/mCG

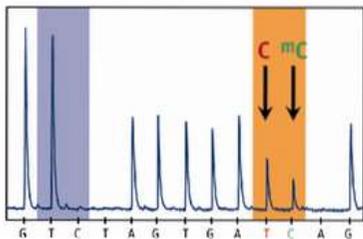
↓ PCR amplification
T

The products of the PCR arise from DNA that originally contained either an unmethylated Cytosine at a given position (the proportion of which is represented by T in the final product) or a methylated Cytosine (the proportion represented by C).

GGTTAGTGAT/C G

↓ Pyrosequencing analysis

Any Cytosine not followed by a Guanine is unmethylated; therefore it should show full conversion in all copies to Thymine. This is a useful quality control to confirm that the transformation process went to completion (blue column). Reference peaks (all peaks outside the coloured columns) confirm that the assay was performed at the correct positions. Methylation at each CpG site is individually quantified (orange column).



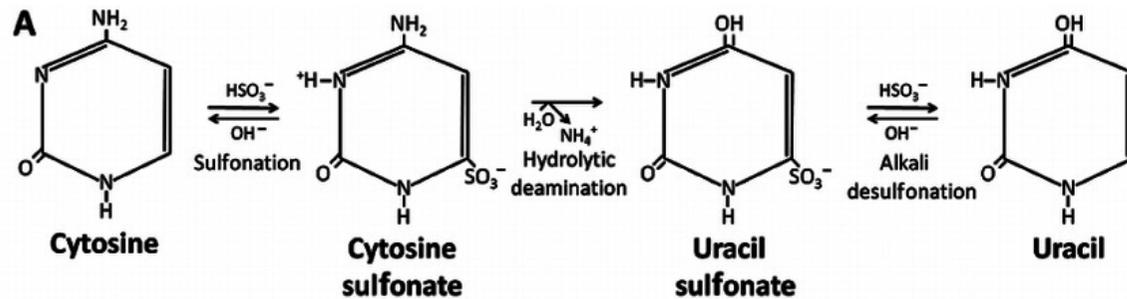
$$\% \text{ methylation} = \frac{\text{C peak height} \times 100}{\text{C peak height} + \text{T peak height}}$$

Conversione con sodio bisolfito

- Fondamentale sia per analisi di regioni (o pannelli di geni) che di singoli geni
- Gli attuali sistemi di sequenziamento non sono in grado di discriminare tra una **C** e una **mC**
- Il trattamento del DNA con **bisolfito** converte le citosine in uracili (**C** → **U**) attraverso una reazione di **deamminazione**. **Tutti gli uracili in fase di sequenziamento saranno letti come T**
- **mC è resistente alla conversione a U quindi rimarrà C**
- Fase cruciale è la **completa conversione** di tutte le C → conversioni parziali possono influire sul risultato finale del profilo di metilazione
- Il DNA convertito diventa template per diverse tipologie di metodiche

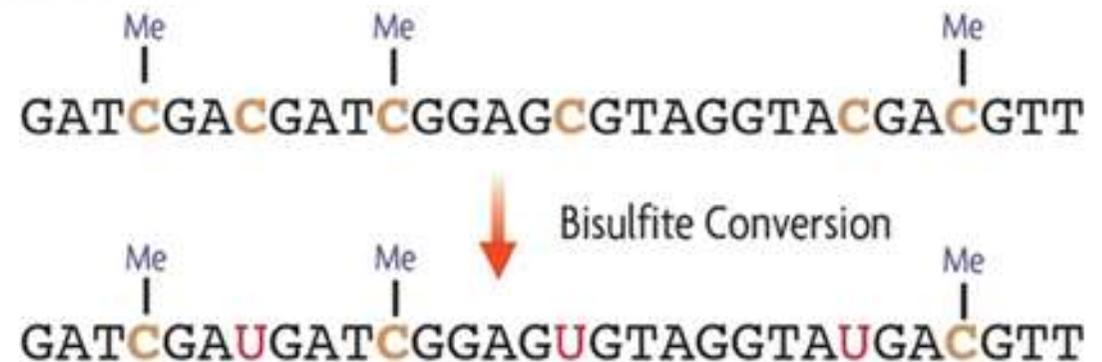
Epigenomica – studio dello stato di metilazione del DNA

CONVERSIONE CON SODIO BISOLFITO



B

	Original sequence	Sequence after bisulfite treatment
Unmethylated DNA	A-T-C-G-G-T-C-A-T-C-G-C-A-T	A-T-U-G-G-T-U-A-T-U-G-U-A-T
Methylated DNA	A-T-C-G-G-T-C-A-T-C-G-C-A-T	A-T-C-G-G-T-U-A-T-C-G-U-A-T



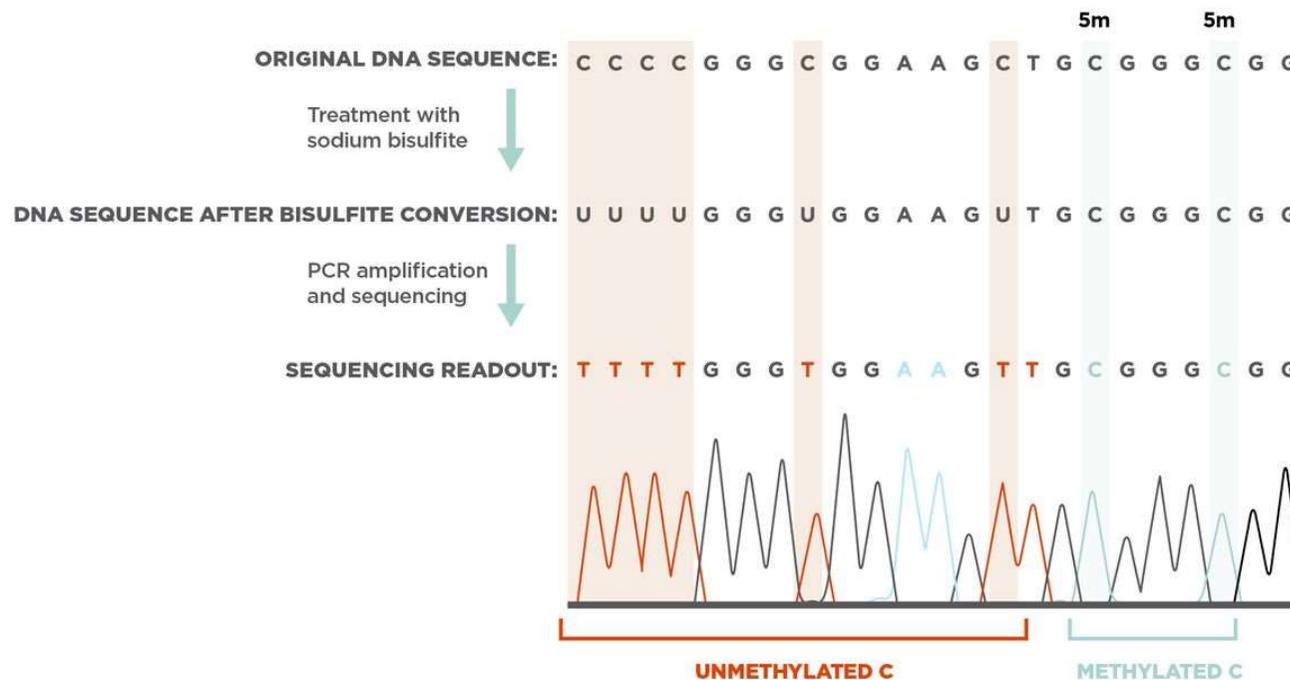
1. La **conversione** si esegue su DNA denaturato (ssDNA), ad alte temperature e basso pH
2. Prima di procedere con le applicazioni successive il DNA convertito deve essere **purificato** per rimuovere i Sali di bisolfito e reagenti chimici

Epigenomica – studio dello stato di metilazione del DNA

CONVERSIONE CON SODIO BISOLFITO

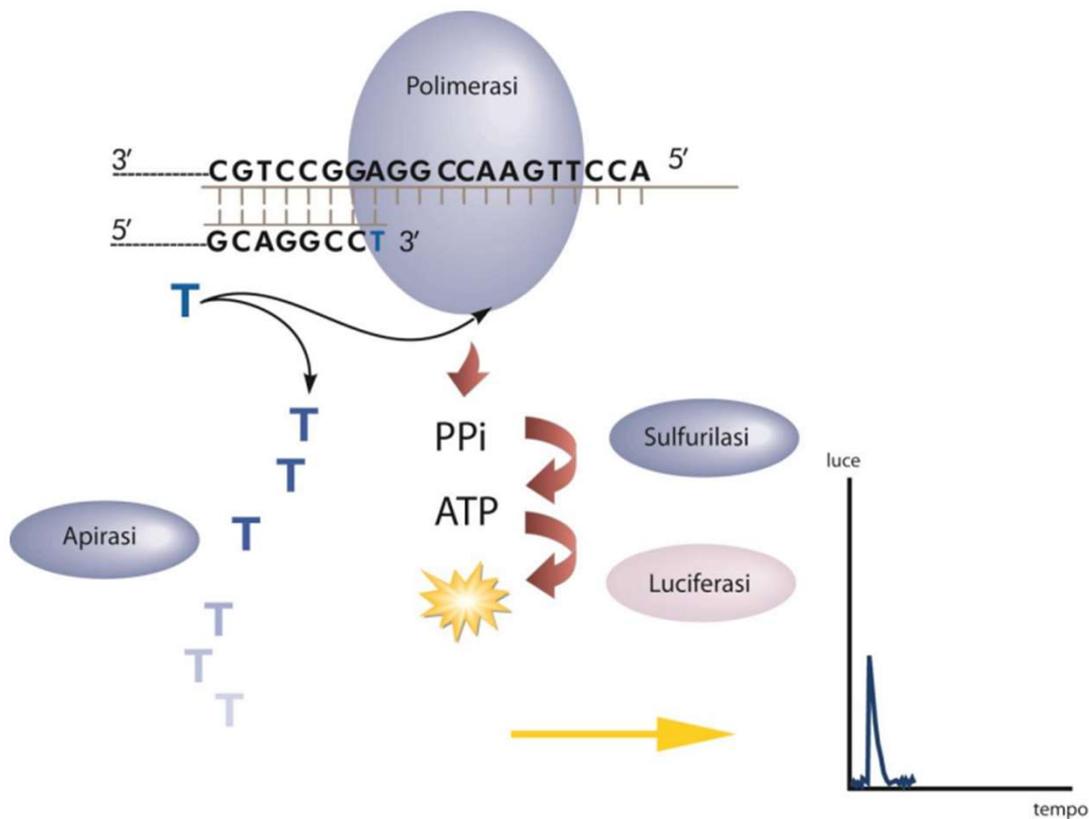
3. **PCR + sequenziamento:** la PCR amplifica le regioni di interesse, che diventano poi il template per un sequenziamento sanger

- I primers sono disegnati vicino alle isole CpG, per cui a volte il disegno può essere complesso
- l'amplificazione di lunghe sequenze convertire può essere complicato (max frammenti 100-300bp)



Epigenomica – studio dello stato di metilazione del DNA

CONVERSIONE CON SODIO BISOLFITO



Utilizza una **DNA Pol** e una miscela di **dNTPs** somministrati uno alla volta

Quando il dNTP è incorporato nel filamento nascente, viene rilasciato **un gruppo pirofosfato (PPi)**

Il pirofosfato è usato da un **enzima sulfurilasi** per produrre **ATP** (1 PP= 1 ATP)

L'ATP è usato dall'**enzima luciferasi** per ossidare la luciferina ed emettere **luce**

Ad ogni nucleotide incorporato si ha l'emissione di un **segnale luminoso**

I nucleotidi non incorporati e l'ATP prodotto è rimosso **dall'enzima apirasi**, per procedere all'iniezione della seconda tipologia di dNTPs

Epigenomica – studio dello stato di metilazione del DNA

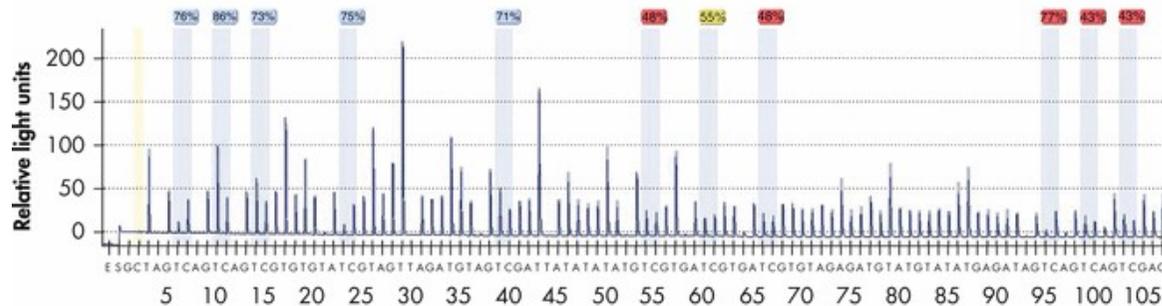
CONVERSIONE CON SODIO BISOLFITO

Pirosequenziamento

- ✓ Usato per analisi di un **numero limitato di sequenze**
- ✓ Dopo la conversione e l'amplificazione attraverso PCR della regione di interesse si procede con il pirosequenziamento di una piccola regione (< **100 bp**)
- ✓ Il grado di metilazione per ogni isola CpG è stimata facendo **il rapporto tra GTP (C) e ATP (U) incorporati**: ottengo un dato quantitativo
- ✓ E' in grado di identificare anche **piccole differenze di metilazione (<5%)**
- ✓ Adatta anche per **campioni eterogenei** (es neoplasia) dove solo una frazione di cellule ha il gene di interesse metilato in modo anomalo

Sequence to analyze:

TTGYGTTYGYGTTTGTGAYGTTTAGGTTTTAGATTTGGTGGTYGATTTTTATTATATTAGGYGTTTAYGTAYGTGTAGAAGATG
TTTATGTATAATTTGAGATGYGYGGYGGAGGAAAATTGTTTGGTTAGGTAT



Epigenomica – studio dello stato di metilazione del DNA

CONVERSIONE CON SODIO BISOLFITO

PCR con High Resolution Melting

Il DNA convertito e successivamente amplificato, è usato come template

Si esegue una PCR usando un **intercalante fluorescente** (SYBR o Eva Green), che quando legato al dsDNA, ha elevata fluorescenza.

Quando il prodotto di PCR è sottoposto **all'incremento di temperatura**, e la struttura ds si dissocia l'intercalante è rilasciato in soluzione e la **fluorescenza diminuisce**.

Il **DNA metilato** (mantiene le C) ha una **temperatura di melting maggiore** rispetto a quello non metilato (U)

Il grado di metilazione è correlato al profilo di melting del campione
Possibile vedere **differenze piccole (5-10%)**, anche in campioni eterogenei

