

# Proprietà di Biopolimeri

a.a. 2023-2024  
Prof. R. Urbani

**Determinazione della struttura  
primaria di Acidi Nucleici**

# Storia della sequenza del DNA

Avery propone il DNA  
come materiale genetico

1944

Holley sequenza il tRNA  
di lievito

1965

**METODO SANGER**  
(terminazione di catena)  
**METODO MAXAM-GILBERT**  
(degradazione chimica)

1977

Parziali modificazioni chimiche del DNA e successiva lisi  
della catena nel sito adiacente al nucleotide modificato

**KARY MULLIS**  
Introduce la PCR

1983

1869

Miescher osserva  
per la prima volta il DNA

1953

Watson e Crick determinano la  
struttura della doppia elica

1970

Vengono sviluppate le tecniche  
per la sintesi degli oligonucleotidi e per  
la degradazione chimica del DNA. Viene  
introdotta l'elettroforesi su gel per la  
Separazione di frammenti di DNA

1980

Il DNA genomico viene clonato nel fago M13  
o in vettori plasmidici, nascono i primi  
Programmi informatici di analisi dei dati,  
vengono sviluppate  
nuove tecnologie per il sequenziamento

1989

**BANCHE DATI!!!!!!!**

**AUTOMAZIONE PARZIALE**  
Vengono sviluppate le prime  
Apparecchiature per il sequenziamento  
che impiegano sistemi  
per la rilevazione della fluorescenza.

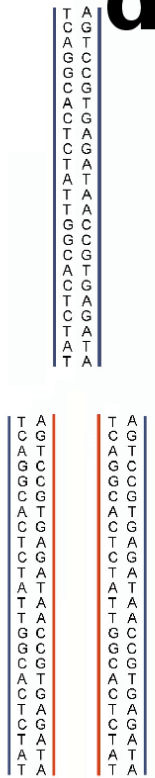
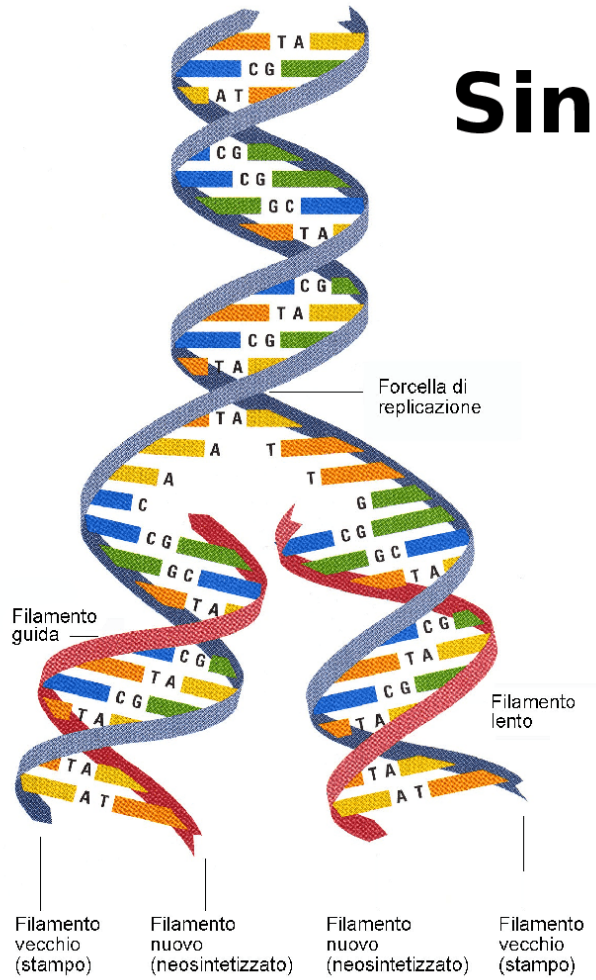
Automazione completa  
Sequenziamento completo  
del genoma umano

2003

**Metodo di SANGER per sequenziare  
oligonucleotidi  
(Dideoxynucleotide chain termination)**

*Metodo di terminazione di catena  
(o dei dideossinucleotidi)*

# Sintesi (replicazione) del DNA

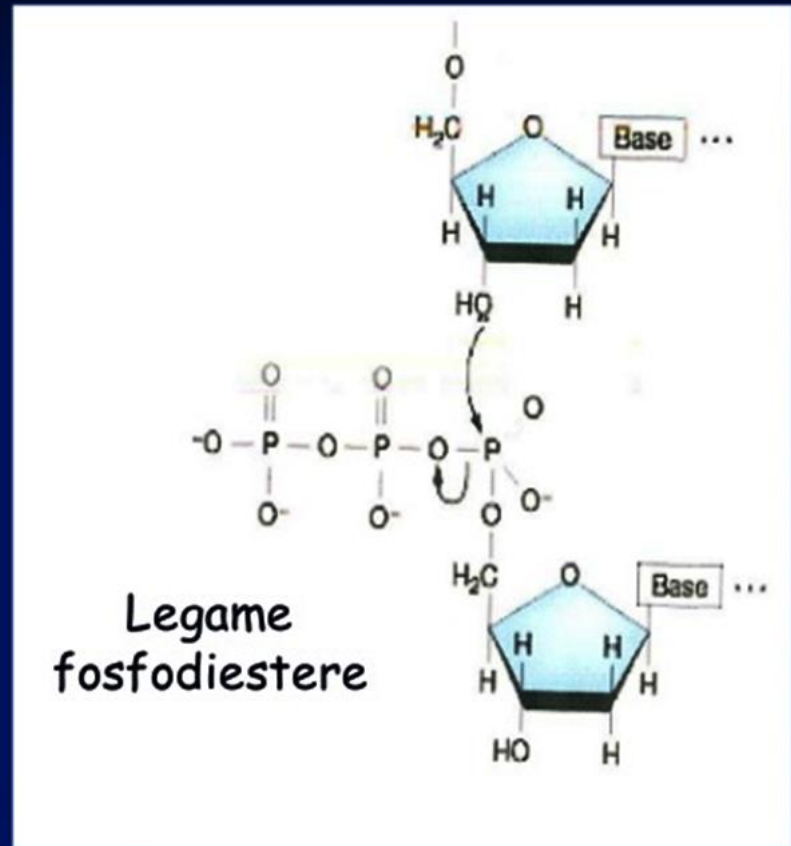


Meccanismo di replicazione del DNA

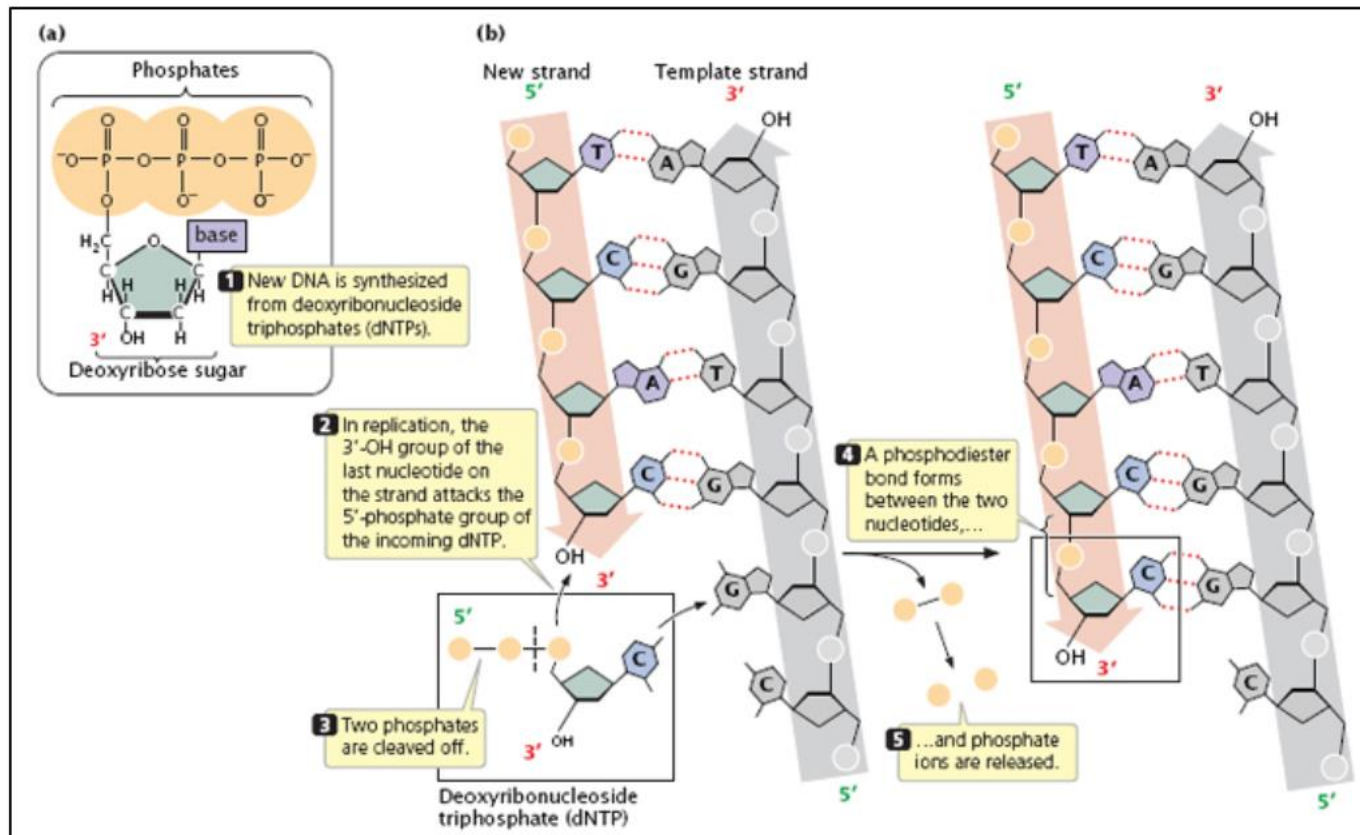
Prodotti della replicazione del DNA

# Formazione del legame fosfodiester del DNA

La reazione avviene per un attacco nucleofilo da parte del gruppo ossidrilico 3' del nucleotide all'atomo di fosforo  $\alpha$  di un nucleoside 5'-trifosfato, causando il rilascio del pirofosfato e la formazione del legame. Tale reazione in vivo è catalizzata dalla DNA polimerasi



l'enzima DNA polimerasi, *in vivo*, è responsabile della replicazione del DNA che avviene durante la mitosi.

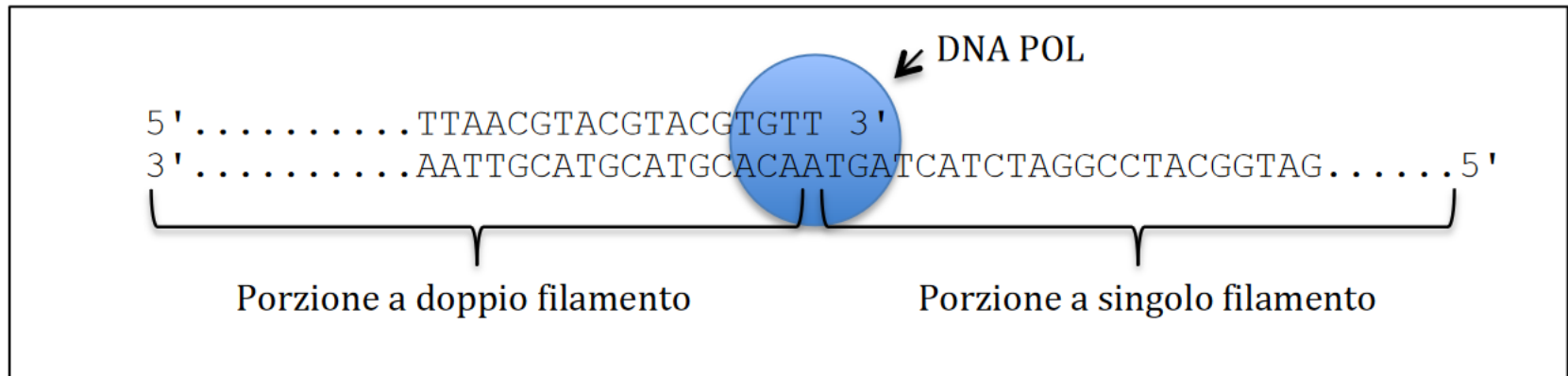


Schema che raffigura l'incorporazione di un nuovo nucleotide da parte della DNA Polimerasi.

La PCR (*Polymerase Chain Reaction*, Mullis, 1983) ricostruisce *in vitro* uno specifico passaggio della duplicazione cellulare: la ricostituzione (sintesi) di un segmento di DNA "completo" (a doppia elica) a partire da un filamento a singola elica.

Per iniziare la reazione occorre che il DNA *stampo* venga denaturato, cioè le due singole eliche che costituiscono il DNA devono essere completamente separate; la DNA polimerasi per poter funzionare ha bisogno di un innesco (*primer*), ovvero di una regione a doppio filamento seguita da una regione a singolo filamento

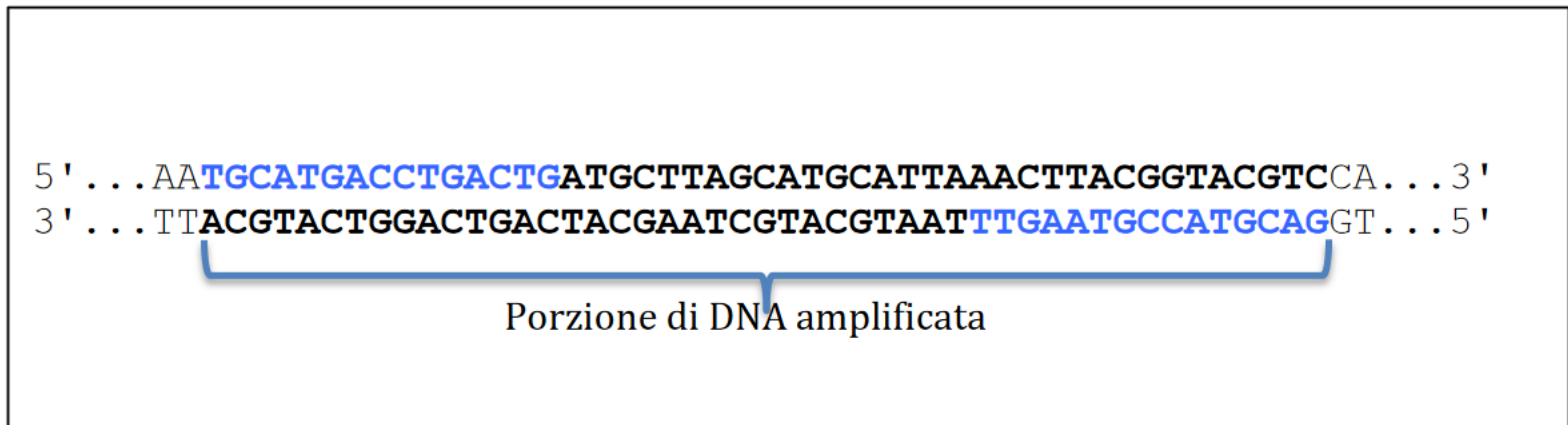
L'enzima, a partire dall'innesco, incorpora un nucleotide dopo l'altro, secondo la complementarità delle basi dell'elica a singolo filamento (che fa da stampo), muovendosi in direzione 5' > 3'



*Esempio di innesco per la DNA Polimerasi*

Nella PCR per formare questo innesco a forma di "scalino" si utilizza un breve frammento di DNA a singola elica (oligonucleotide o **primer**) lungo 15-30 bp, in grado di appaiarsi in una specifica regione di una delle due eliche del DNA, secondo la complementarietà delle basi.

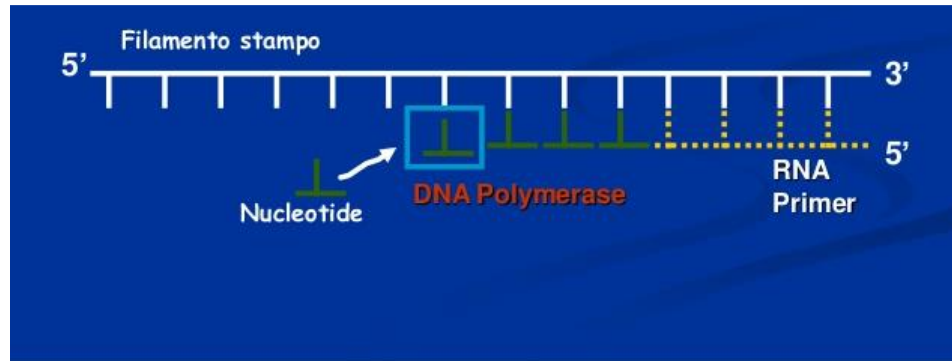
Le regione di DNA che viene amplificata è quella compresa tra i due oligonucleotidi.



La PCR è una reazione ciclica costituita da 25--35 cicli ciascuno costituito da 3 passaggi successivi:

- denaturazione del DNA
- appaiamento degli oligonucleotidi (primer)
- polimerizzazione (montaggio dei nucleotidi da parte della DNA polimerasi).





Nessuna DNA-polimerasi è in grado di iniziare la sintesi di un filamento "da zero". Per questa ragione le DNA-polimerasi necessitano di un oligonucleotide (*primer*) complementare al filamento template.

Il *primer* funge da innesco della reazione di amplificazione perché fornisce un gruppo **idrossilico (-OH) in posizione 3'** a cui la DNA polimerasi legherà il primo nucleotide del filamento neosintetizzato.

# 14 sono le DNA polimerasi eucariotiche !!!

**Tabella 6.1 Le DNA polimerasi eucariotiche\***

Nome	Funzione
<b>Replicasi ad alta fedeltà</b>	
Pol $\alpha$ (alfa)	Innesco della sintesi del DNA durante la replicazione e la riparazione
Pol $\delta$ (delta)	Replicazione del DNA del filamento leader (e ritardato?) durante la replicazione e la riparazione (BER, DSBR, MMR, NER)
Pol $\epsilon$ (epsilon)	Replicazione del DNA del filamento ritardato durante la replicazione e la riparazione (BER, DSBR, NER)
Pol $\gamma$ (gamma)	Replicazione e riparazione del DNA mitocondriale
<b>Riparazione ad alta fedeltà</b>	
Pol $\beta$ (beta)	BER, DSBR
Pol $\eta$ (eta)	Sintesi translesione del DNA (replicazione relativamente accurata di fronte ai dimeri timina-timina)
<b>Riparazione incline all'errore</b>	
Pol $\zeta$ (zeta)	Sintesi translesione del DNA (aggiramento dei dimeri di timina)
Pol $\theta$ (teta)	Riparazione dei legami crociati interfilamento
Pol $\iota$ (iota)	Sintesi translesione del DNA (richiesta durante la meiosi)
Pol $\kappa$ (kappa)	Sintesi translesione del DNA (delezioni e sostituzioni di basi), DSBR (unione non omologa delle estremità)
Pol $\lambda$ (lambda)	Sintesi translesione del DNA
Pol $\mu$ (mu)	DSBR (unione non omologa delle estremità)
Pol $\nu$ (nu)	Riparazione di legami crociati nel DNA?
Rev1	Sintesi di siti abasici (l'attività deossicitidil trasferasica inserisce C in un nucleotide privo di una base)

# 5 sono le DNA polimerasi batteriche !!!

## DNA FOCUS 6.1

### Le DNA polimerasi batteriche

Nel batterio *E. coli* ci sono cinque DNA polimerasi principali: le DNA polimerasi I, II, III, IV e V. Insieme, queste cinque polimerasi svolgono la stessa serie di funzioni svolta dalle 14 DNA polimerasi eucariotiche.

#### DNA polimerasi I

La DNA polimerasi I è stata la prima polimerasi ad essere purificata e caratterizzata ed è usata estesamente in biologia molecolare per la sua facile reperibilità e per le sue proprietà uniche. È la polimerasi più abbondante in *E. coli*, ma non è l'enzima responsabile della maggior parte della replicazione mentre ha un ruolo nella rimozione del primer, nel riempimento delle interruzioni fra i frammenti di Okazaki e nella via di riparazione per escissione dei nucleotidi (vedi Sezione 7.6). La DNA polimerasi I ha due subunità. Una subunità, chiamata frammento di Klenow, ha attività polimerasica 5'→3', mentre l'altra ha attività esonucleasica sia 3'→5' che 5'→3'. L'insieme delle due subunità è chiamato oloenzima. L'oloenzima ha la capacità esclusiva di iniziare la replicazione a livello di un nick (un legame fosfodiesterico spezzato) nell'ossatura di zucchero-fosfato del DNA. Questa proprietà è sfruttata in laboratorio con una tecnica chiamata "nick translation". Anche il frammento di Klenow è largamente usato in biologia molecolare,

per esempio, per marcare il DNA in una tecnica chiamata "random priming" (vedi Strumenti 8.5).

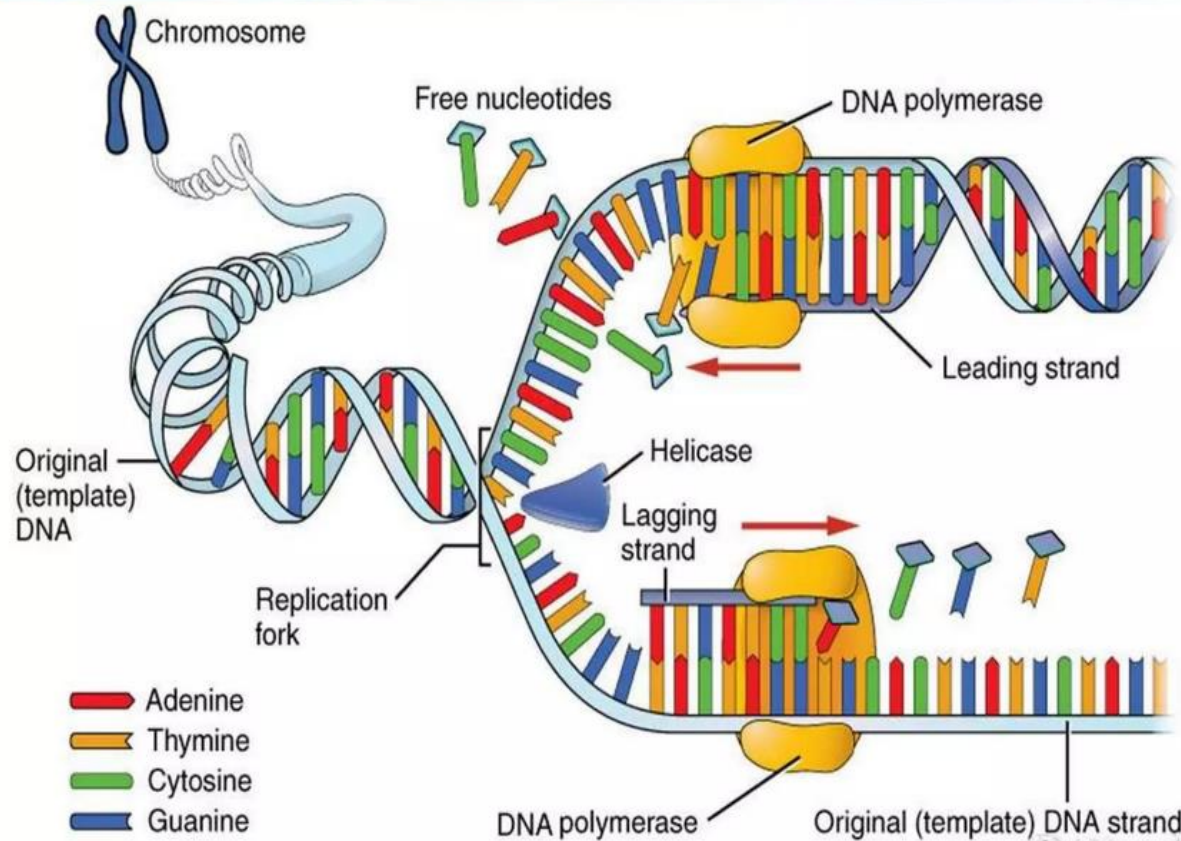
#### DNA polimerasi III

Fu una sorpresa per i ricercatori quando si accorsero che la DNA polimerasi I, la polimerasi più abbondante, non era la polimerasi replicativa principale in *E. coli*. In effetti, un enzima molto meno abbondante, la DNA polimerasi III, catalizza la replicazione del genoma. L'oloenzima contiene 10 subunità polipeptidiche diverse. La subunità  $\alpha$  ha attività di replicasi e la subunità  $\epsilon$  ha attività di correzione delle bozze (esonucleasi 3'→5'). Anche la DNA polimerasi III ha un ruolo nelle vie di riparazione per escissione dei nucleotidi (vedi Sezione 7.6).

#### DNA polimerasi II, IV e V

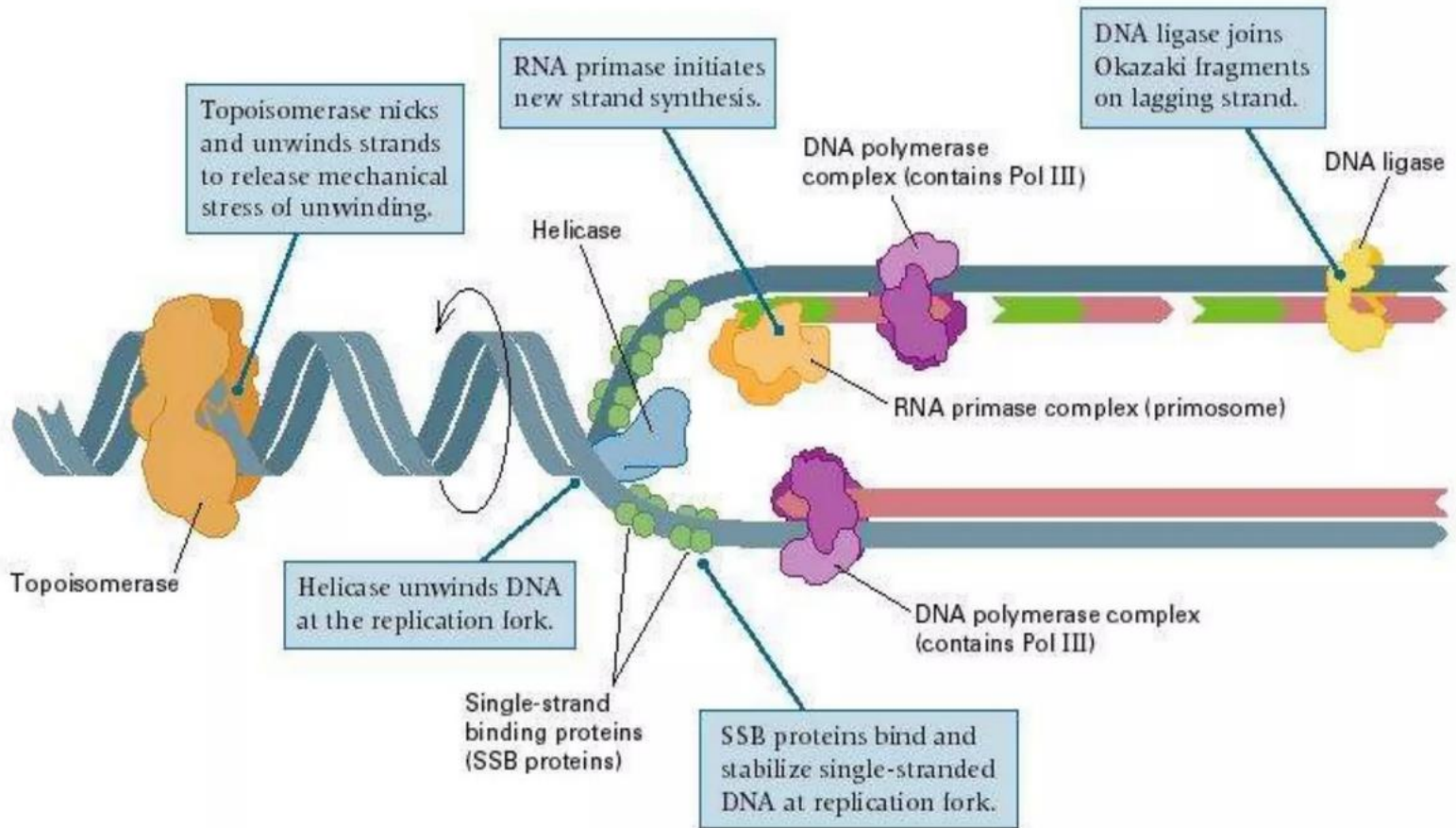
La DNA polimerasi II è coinvolta nei meccanismi di riparazione del DNA (vedi Sezione 7.4), mentre le DNA polimerasi IV e V mediano la sintesi translesione del DNA. La DNA polimerasi IV, chiamata anche DinB, è codificata dal gene *dinB*. La polimerasi V, nota anche come complesso UmuD'<sub>2</sub>C, è codificata dall'operone *umuDC*. Entrambe le polimerasi possono aggirare il danno al DNA che ha bloccato la replicazione da parte della DNA polimerasi III. Queste polimerasi possono avere un ruolo nella mutagenesi adattiva, in quanto tendono a compiere errori.

## WHAT IS DNA POLYMERASE?



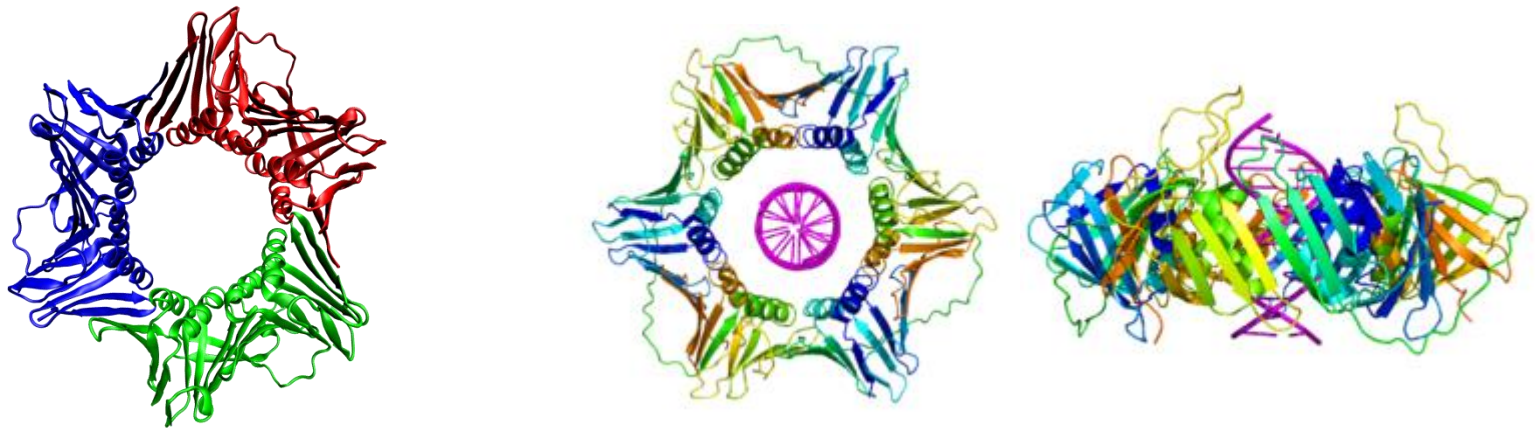
# ENZYMES INVOLVED IN DNA REPLICATION

- This is the list of Enzymes involved in DNA Replication.
  - DNA Helicase
  - DNA Polymerase
  - DNA clamp
  - Single-Strand Binding (SSB) Proteins
  - Topoisomerase / DNA Gyrase
  - DNA Ligase
  - Primase



Struttura proteica della **pinza scorrevole** di PCNA umana  
(DNA clamp o sliding clamp)

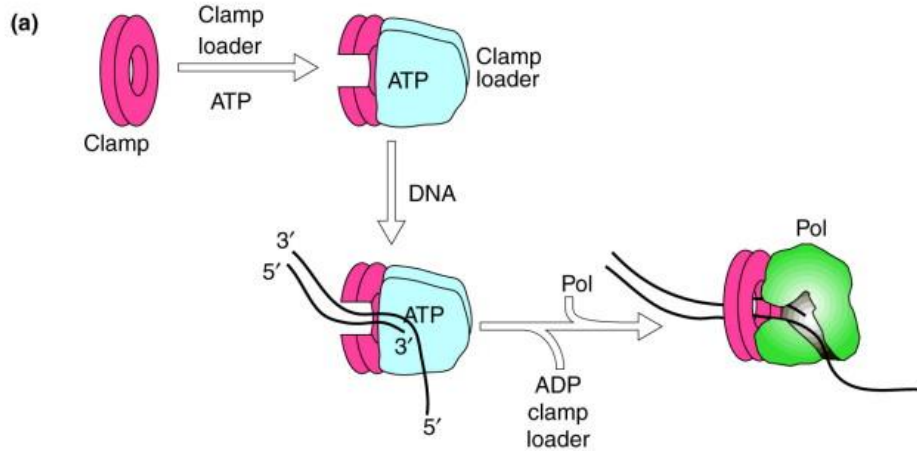
fattore di *processamento* per la DNA-polimerasi, promuove e controlla l'azione della polimerasi durante la replicazione, impedendo alla polimerasi di staccarsi dal filamento di DNA (interazioni specifiche forti)



la presenza della **pinza scorrevole** aumenta notevolmente il numero di nucleotidi che la polimerasi può aggiungere e quindi la velocità di replicazione del filamento di DNA (fino a 1.000 volte rispetto ad una polimerasi priva del clamp)

PCNA = Proliferating Cell Nuclear Antigen (antigene di proliferazione nucleare)

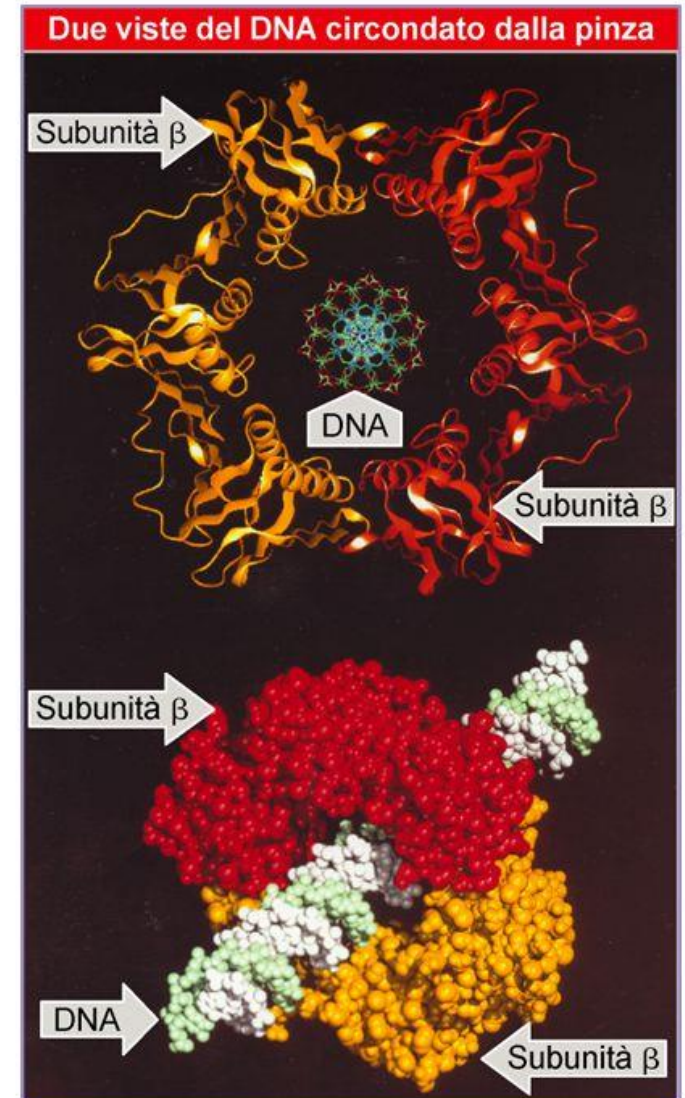
La pinza scorrevole è una proteina  $\alpha+\beta$  (due subunità nei batteri, tre negli eucarioti) che vengono assemblate in un multimero ad anello (toroidale) il quale circonda completamente la doppia elica del DNA



batteri

virus

eucarioti





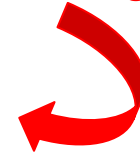
# articles

## Nucleotide sequence of bacteriophage $\Phi$ X174 DNA

F. Sanger, G. M. Air\*, B. G. Barrell, N. L. Brown†, A. R. Coulson, J. C. Fiddes, C. A. Hutchison III‡, P. M. Slocombe§ & M. Smith\*

MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK

L'articolo descrive il primo sequenziamento di un genoma



## Metodo enzimatico di Sanger

*Proc. Natl. Acad. Sci. USA*  
Vol. 74, No. 12, pp. 5463-5467, December 1977  
Biochemistry

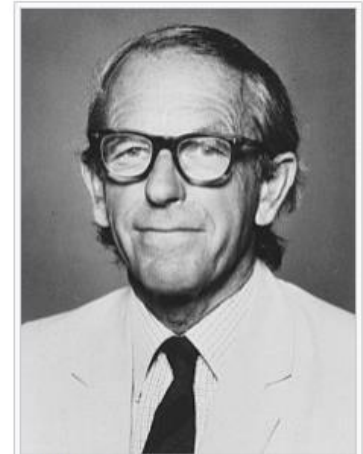
### DNA sequencing with chain-terminating inhibitors

(DNA polymerase/nucleotide sequences/bacteriophage  $\phi$ X174)

F. SANGER, S. NICKLEN, AND A. R. COULSON

Medical Research Council Laboratory of Molecular Biology, Cambridge CB2 2QH, England

*Contributed by F. Sanger, October 3, 1977*



Frederick Sanger



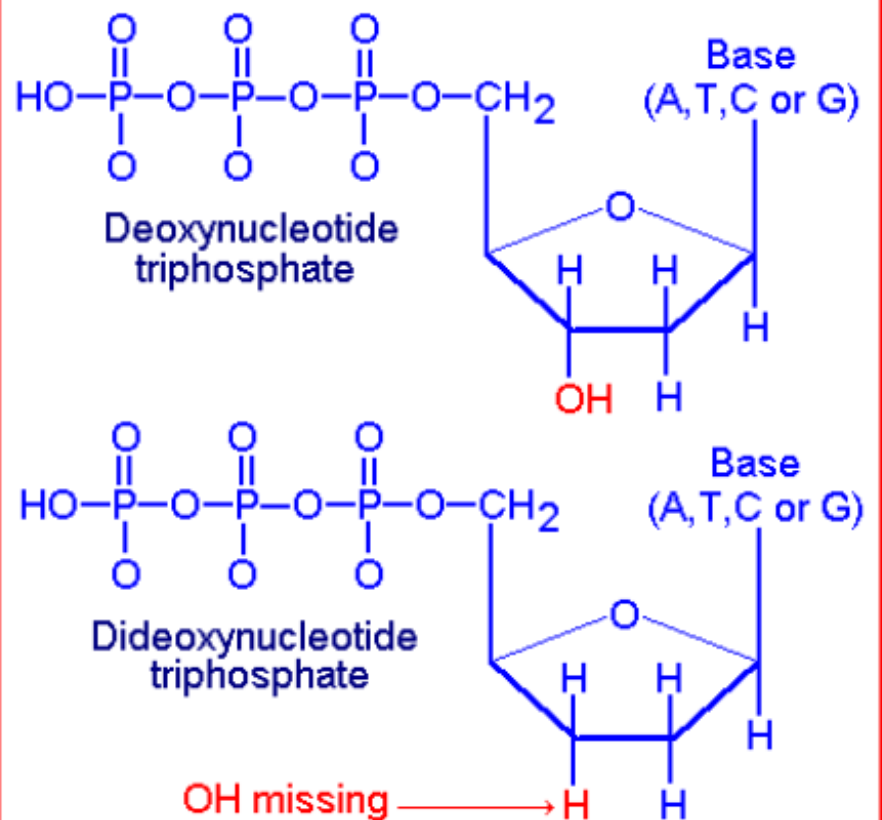
Premio Nobel per la chimica 1958

Premio Nobel per la chimica 1980

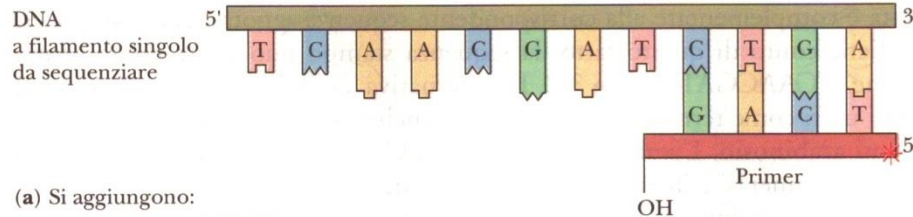
1918 - 2013

## Nucleotidi terminatori (ddNTP)

- Nucleotidi che bloccano l'allungamento del filamento di acido nucleico: la mancanza del gruppo ossidrilico in 3' impedisce l'attacco al gruppo fosforico del nucleotide successivo
- Nucleotidi marcati con  $^{32}\text{P}$  (in passato) o molecole fluorescenti (recentemente)



# Metodo di SANGER di terminazione di catena (o dei dideozinucleotidi)

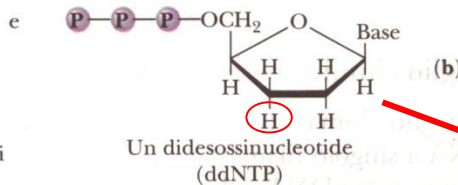


(a) Si aggiungono:

Un dNTP è marcato radioattivamente

- dATP
- dTTP
- dCTP
- dGTP

+ DNA polimerasi



(c)

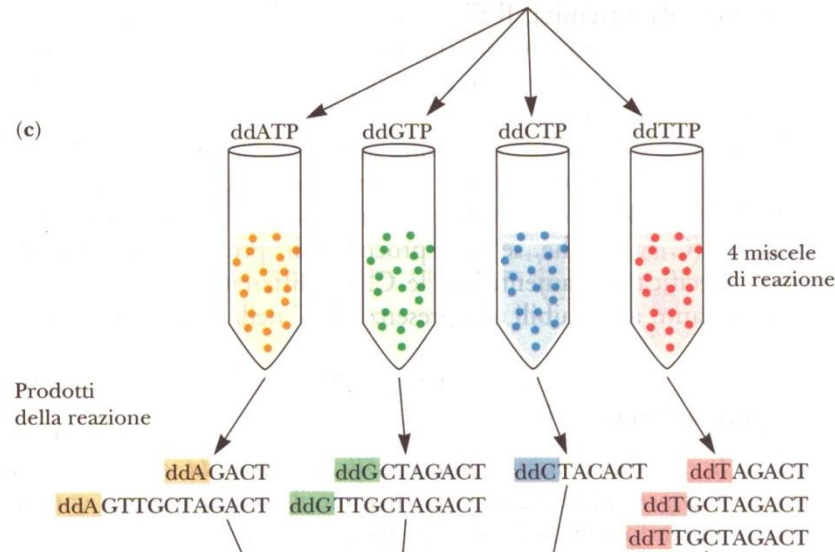
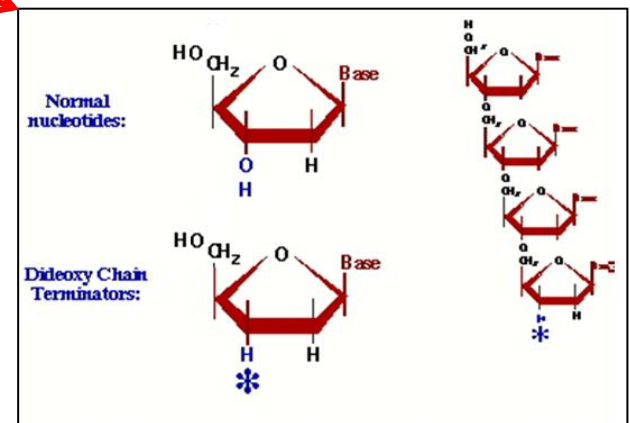
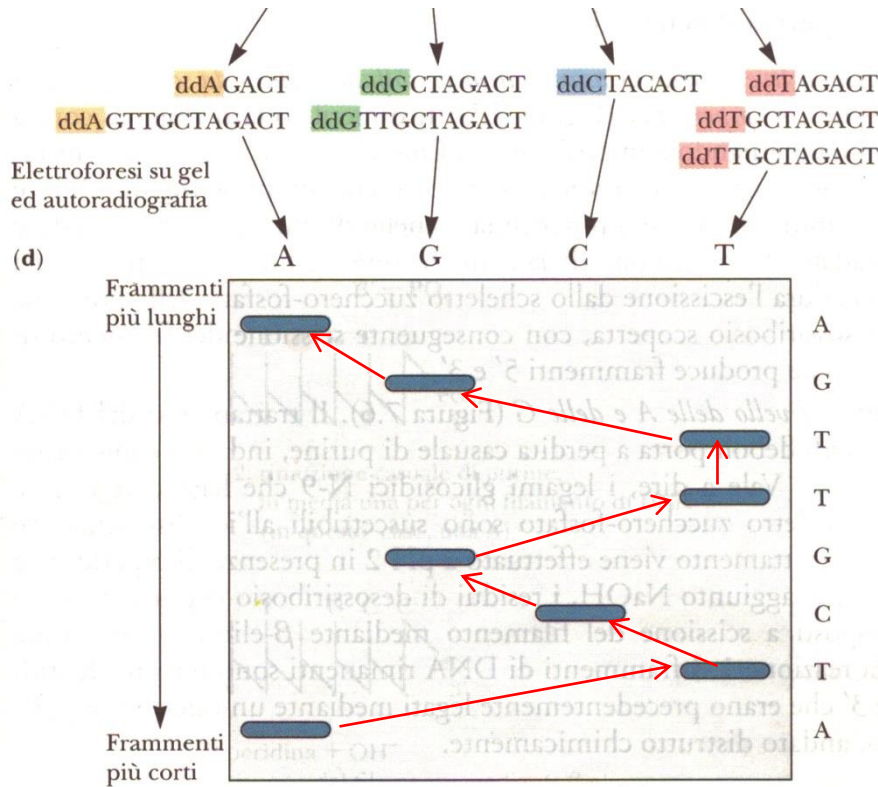


Figura 7.3 Il metodo di sequenziamento del DNA di terminazione della catena, o metodo dei dideozinucleotidi. (a) Reazione della DNA polimerasi. (b) Struttura di un dideozinucleotide. (c) Quattro miscele di reazione, contenenti i nucleosidi trifosfati, con l'aggiunta, in ciascuna reazione, di un solo dideozinucleoside trifosfato.



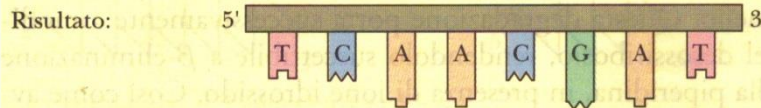
Poiché il DNA sintetizzato contiene un gruppo marcato (radio) può essere distinto dal template



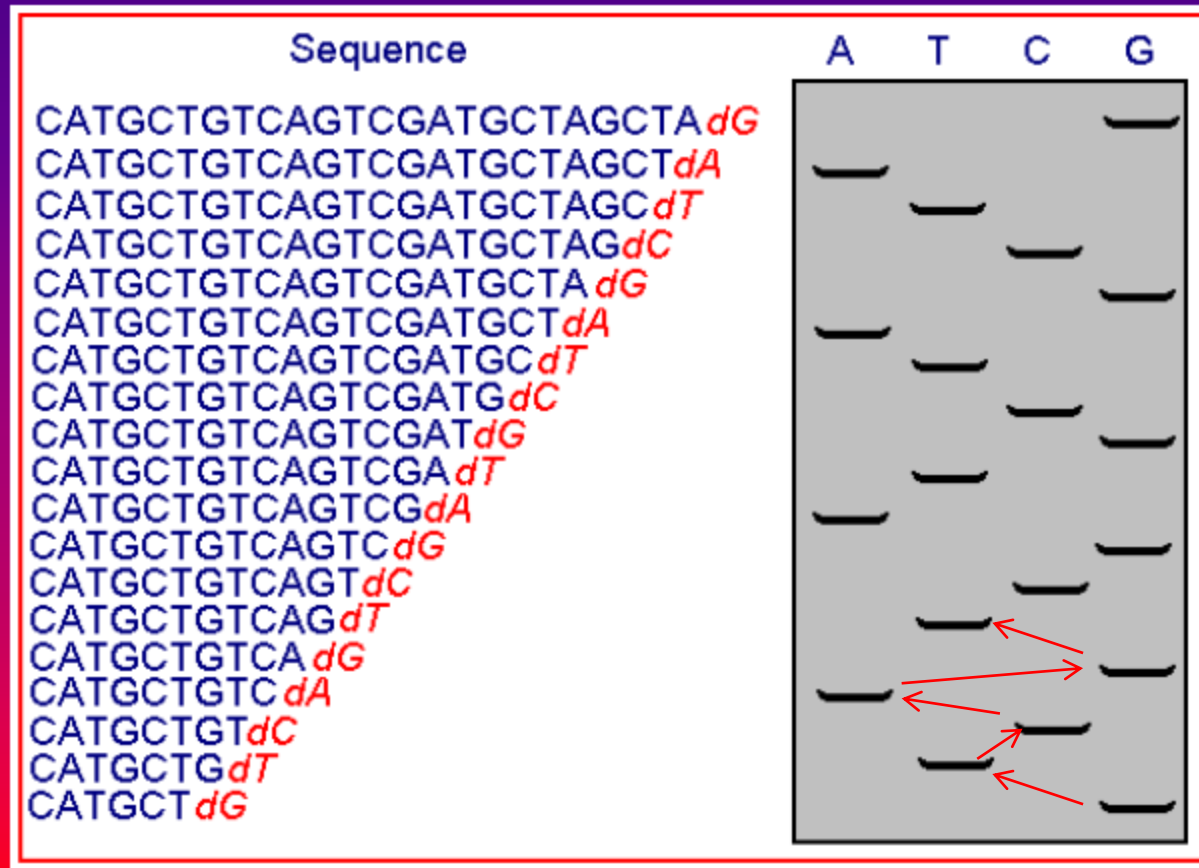
Elettroforetogramma. Si noti che la sequenza nucleotidica che si legge nel gel dal basso verso l'alto riflette l'ordine con cui i nucleotidi sono stati aggiunti dalla DNA polimerasi.

Letture della sequenza dal basso verso l'alto: -A-T-C-G-T-T-G-A-

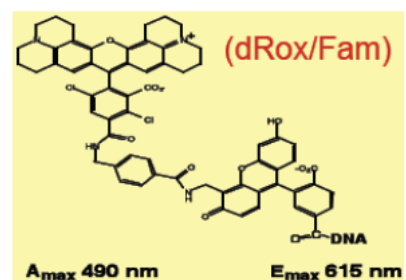
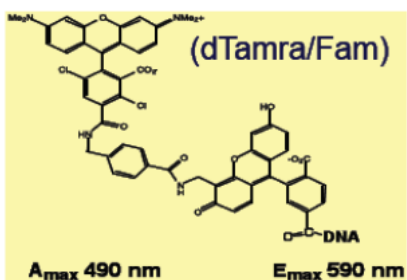
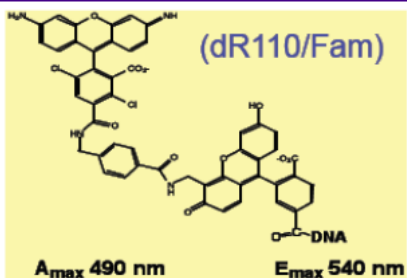
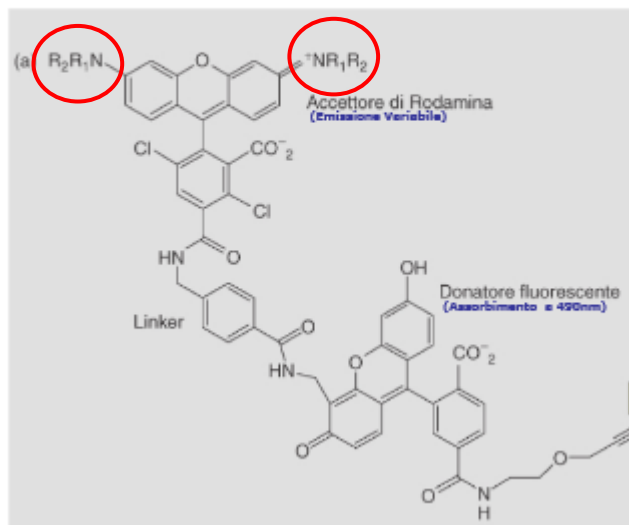
La sequenza complementare è il filamento stampo originale (3'→5'): -T-A-G-C-A-A-C-T-



# Produzione dei frammenti (metodo di Sanger)



## Fluorocromi (BigDye™)



I quattro fluorocromi hanno differenti proprietà di emissione  
In funzione della natura dei gruppi R.

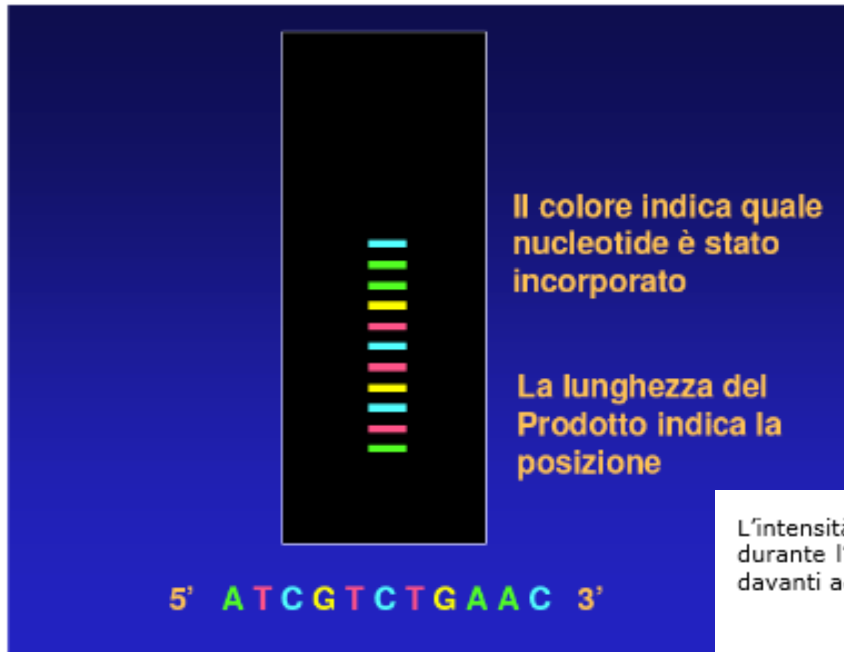
Fluorocromo	Max eccitazione	Max emissione
<b>A</b> FITC	495	519
Alexa 488	499	520
Oregon Green	513	533
PI	538	603
TRITC	552	578
Alexa 568	577	603
Texas Red	595	613
Cy5	648	665
Alexa 660	663	691
<b>B</b> Clorofilla a	430	670
Clorofilla b	460	650
<b>C</b> CFP	430	474
GFP	494	510
YFP	520	535
DsRed	553	585

**Figura 13.4**

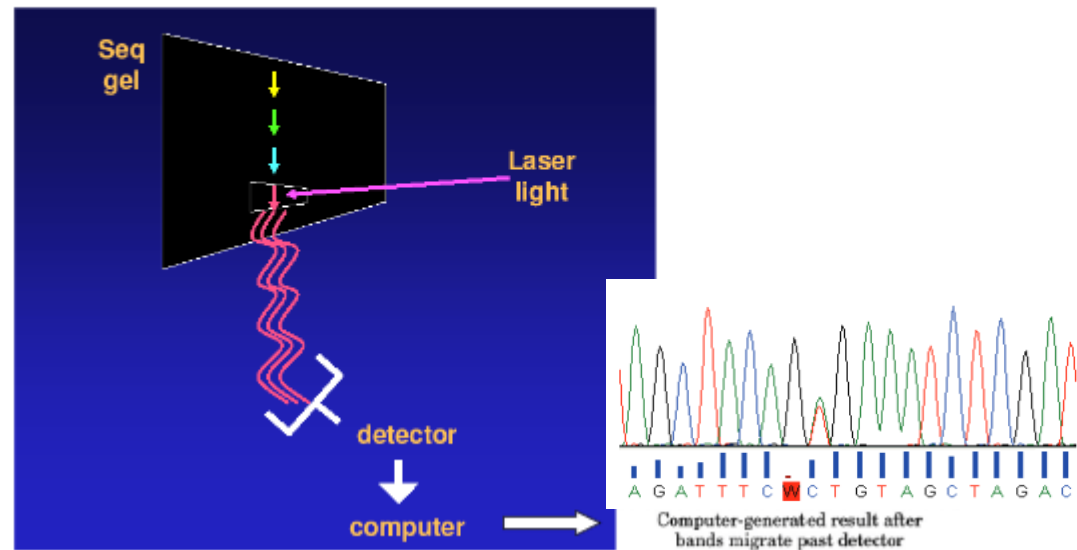
Massimi di eccitazione e di emissione per alcuni fluorocromi di sintesi (A), pigmenti (B) e proteine fluorescenti (C).

Nel caso dei **BigDye Terminators** sono i ddNTPs ad essere legati al fluorocromo

# Dye Terminator Labelling

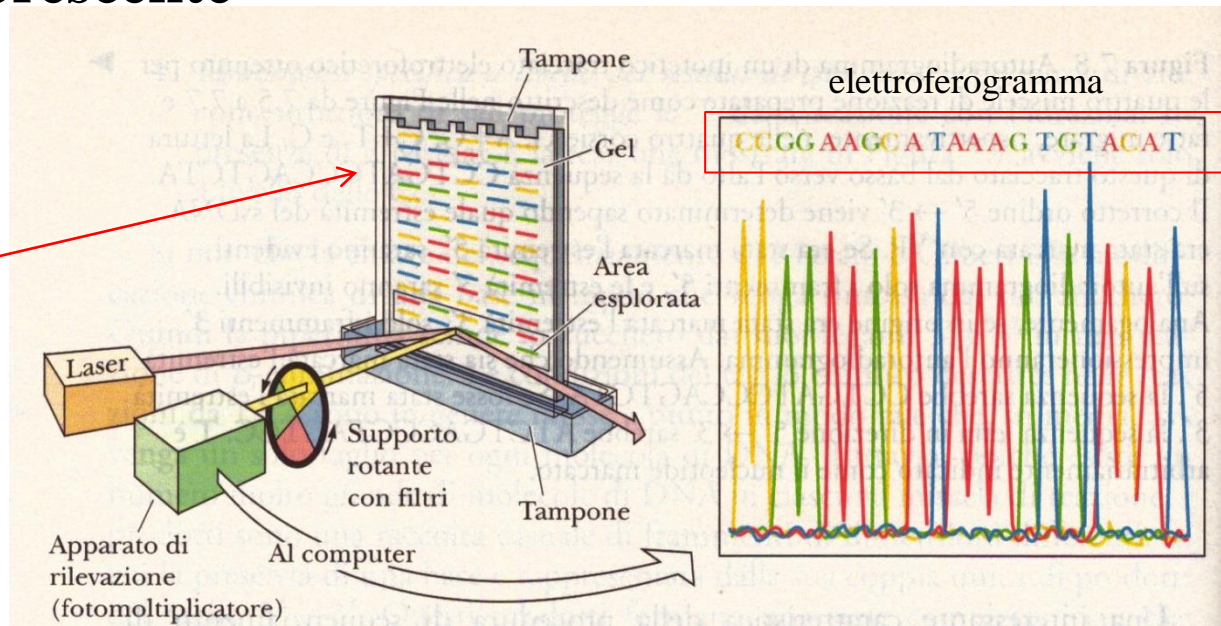


L'intensità e la lunghezza d'onda delle emissioni fluorescenti vengono captate durante l'elettroforesi quando i frammenti di DNA, eccitati dalla luce laser, passano davanti ad un rivelatore.



# Marcatura fluorescente

È possibile marcare il *primer* con un composto fluorescente diversamente colorato per ciascuna delle quattro miscele di reazione di terminazione

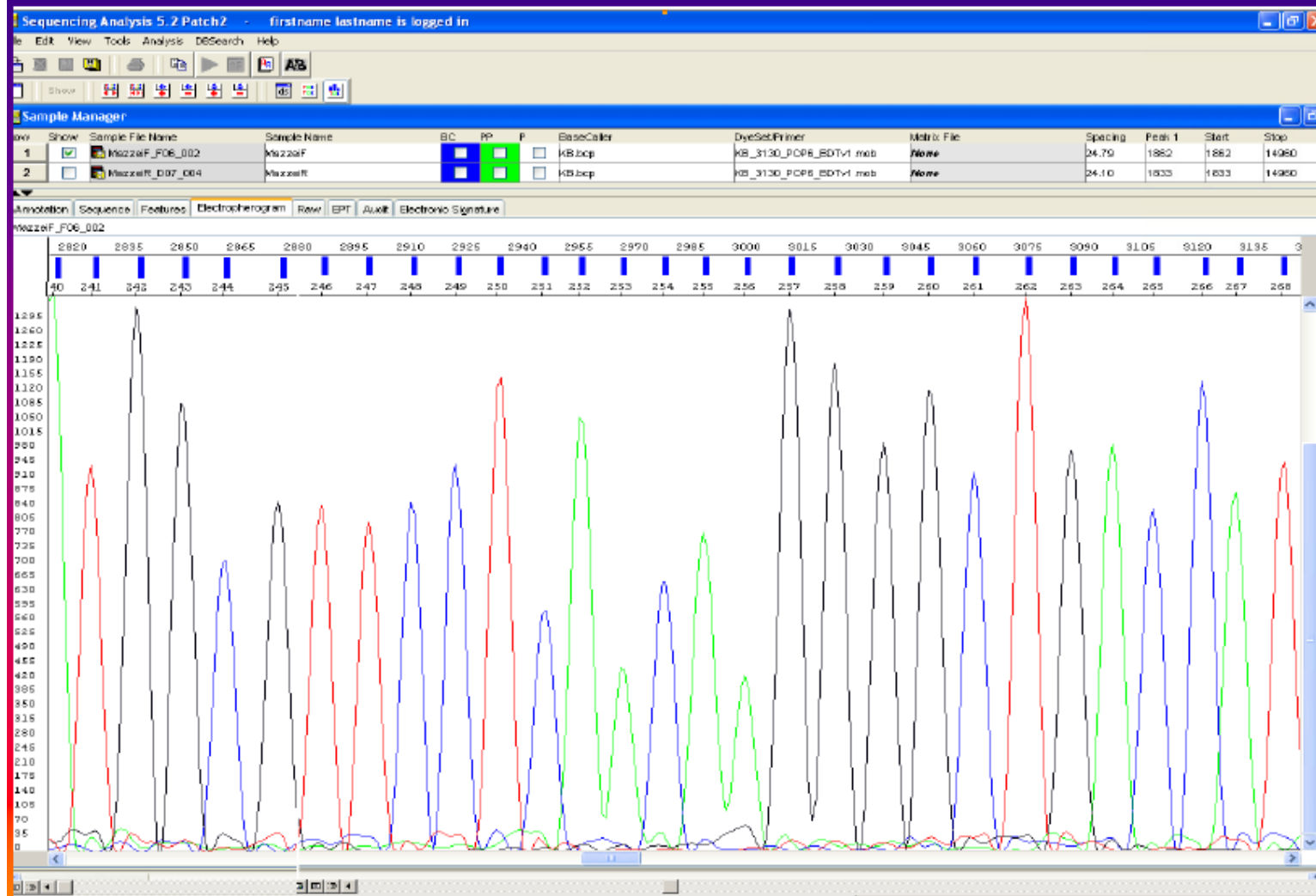


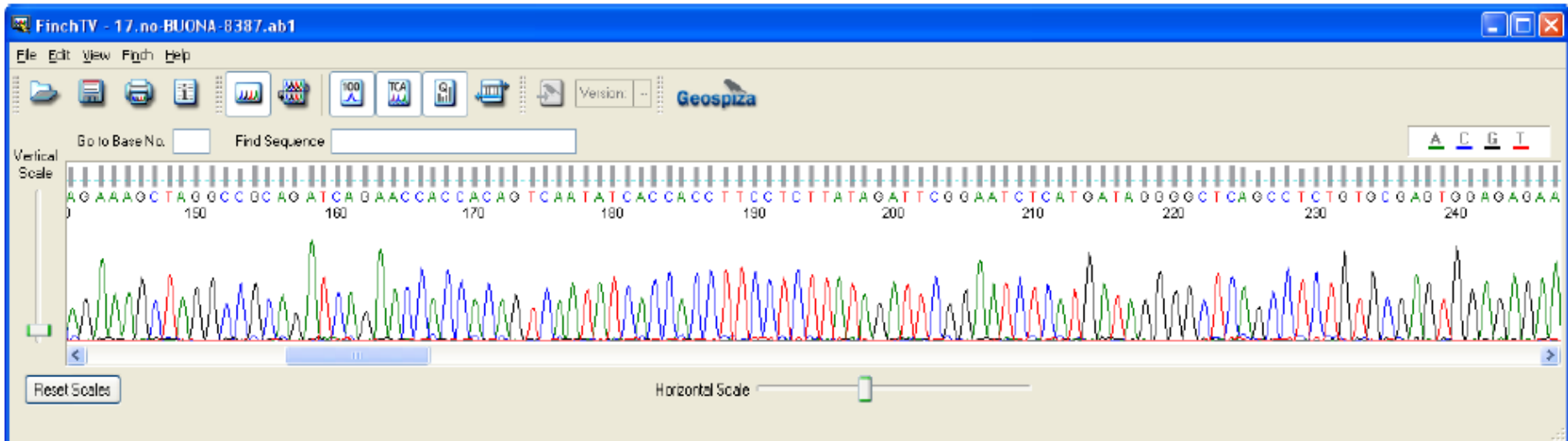
**Figura 7.10** Illustrazione schematica della metodologia impiegata nel sequenziamento automatico con marcatura fluorescente del DNA. Si allestiscono quattro reazioni, una per ciascuna base, ed in ciascuna di esse il primer viene marcato ad una estremità con uno dei quattro diversi coloranti fluorescenti; i coloranti funzionano in questo protocollo di sequenziamento come un codice di colori, specifici per le basi (viene usato un solo colorante in ciascuna reazione contenente il relativo didesossinucleotide). Le quattro miscele di reazione vengono poi combinate, e fatte migrare in una sola corsia del gel. Pertanto, ciascuna corsia rappresenta un differente esperimento di sequenziamento. Quando i frammenti di dimensioni differenti procedono verso il basso nel gel, un raggio laser eccita il colorante nell'area di rilevazione. L'energia emessa passa attraverso un filtro rotante colorato e viene rilevata da un fluorimetro. Il colore della luce emessa identifica la base finale del frammento.

<10,000 basi



# ELETTROFEROGRAMMA





# GenBank

- Banca dati ad accesso libero (internet)
- Circa **240 milioni** di sequenze geniche (animali e vegetali)
- Permette il confronto della propria sequenza in tempo reale: motore di ricerca (programma blast) individua le sequenze con il più elevato grado di omologia
- Chiunque può depositarvi le proprie sequenze

## BLAST Assembled Genomes

Choose a species genome to search, or [list all genomic BLAST databases](#).

- [Human](#)
- [Mouse](#)
- [Rat](#)
- [Arabidopsis thaliana](#)
- [Oryza sativa](#)
- [Bos taurus](#)
- [Danio rerio](#)
- [Drosophila melanogaster](#)
- [Gallus gallus](#)
- [Pan troglodytes](#)
- [Microbes](#)
- [Apis mellifera](#)

## Basic BLAST

Choose a BLAST program to run.

<a href="#">nucleotide blast</a>	Search a <b>nucleotide</b> database using a <b>nucleotide</b> query <i>Algorithms:</i> blastn, megablast, discontinuous megablast
<a href="#">protein blast</a>	Search <b>protein</b> database using a <b>protein</b> query <i>Algorithms:</i> blastp, psi-blast, phi-blast
<a href="#">blastx</a>	Search <b>protein</b> database using a <b>translated nucleotide</b> query
<a href="#">tblastn</a>	Search <b>translated nucleotide</b> database using a <b>protein</b> query
<a href="#">tblastx</a>	Search <b>translated nucleotide</b> database using a <b>translated nucleotide</b> query



National Institutes of Health  
Turning Discovery Into Health

 An official website of the United States government [Here's how you know](#) ▾

 National Library of Medicine  
National Center for Biotechnology Information



GenBank

Nucleotide ▾

Search

GenBank ▾

Submit ▾

Genomes ▾

WGS ▾

Metagenomes ▾

TPA ▾

TSA ▾

INSDC ▾

Documentation ▾

Other ▾

## GenBank Overview

### What is GenBank?

GenBank<sup>®</sup> is the NIH genetic sequence database, an annotated collection of all publicly available DNA sequences ([Nucleic Acids Research, 2013 Jan;41\(D1\):D36-42](#)). GenBank is part of the [International Nucleotide Sequence Database Collaboration](#), which comprises the DNA DataBank of Japan (DDBJ), the European Nucleotide Archive (ENA), and GenBank at NCBI. These three organizations exchange data on a daily basis.

A GenBank release occurs every two months and is available from the [ftp site](#). The [release notes](#) for the current version of GenBank provide detailed information about the release and notifications of upcoming changes to GenBank. Release notes for [previous GenBank releases](#) are also available. GenBank growth [statistics](#) for both the traditional GenBank divisions and the WGS division are available from each release.

An [annotated sample GenBank record](#) for a *Saccharomyces cerevisiae* gene demonstrates many of the features of the GenBank flat file format.

### GenBank Resources

[GenBank Home](#)

[Submission Types](#)

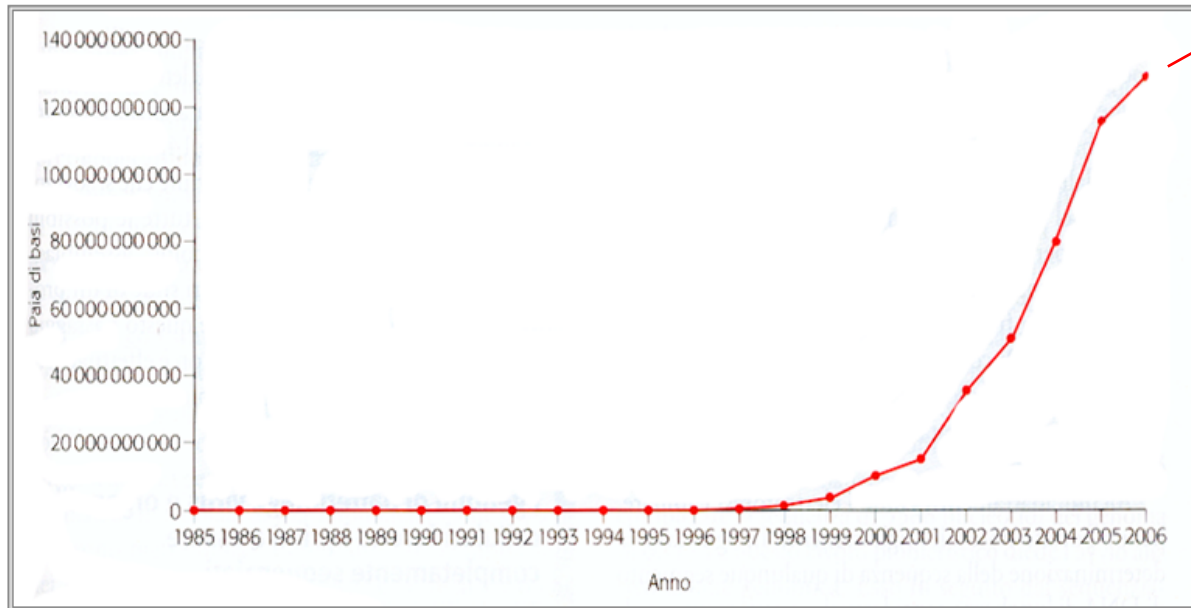
[Submission Tools](#)

[Search GenBank](#)

[Update GenBank Records](#)

# SEQUENZIAMENTO DEL DNA – PROGETTO GENOMA UMANO

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



Dal 1982 a oggi, il numero di basi in GenBank è **raddoppiato** ogni 18 mesi circa.

Nell' **ottobre 2022** si sono raggiunte le **201 562 963 366 851** basi relative a **240 539 282** sequenze

## Next Generation Sequencing (NGS)

- L'insieme delle tecnologie di sequenziamento degli acidi nucleici che hanno in comune la capacità di sequenziare, in parallelo, milioni di frammenti di DNA
- Potenzialità di produrre, in un'unica seduta di analisi, una quantità di informazioni genetiche milioni di volte più grande

**3 miliardi di dollari e 13 anni di lavoro (1990-2003) di 20 diverse università e centri di ricerca** si sono resi necessari per completare nel 2003 la caratterizzazione di un genoma umano per lo "*Human genome project*". La prima bozza del genoma si è avuta nel 2000 e il completamento della sequenza nel 2003

Oggi, a pochi anni di distanza, le tecnologie NGS più avanzate consentono il sequenziamento di un genoma umano **in pochi giorni al costo di circa 1.000 dollari !**