

Il sequenziamento del DNA: principi e applicazioni presenti e future

Sequenziamento “automatico” del DNA

Maurizio Zini

Biotechnologie
Istituto Zooprofilattico Sperimentale
delle Regioni Lazio e Toscana



UN PO' DI STORIA...

1900

- Riscoperta delle leggi della trasmissione ereditaria dei caratteri, che **Mendel** aveva pubblicato nel 1865 e che avevano come assunto l'esistenza di fattori discreti, *elementen*, i geni, che determinano i caratteri e che si trasmettono da una generazione all'altra.

1902

- **Sutton** ipotizza che i fattori mendeliani, i geni, siano localizzati sui cromosomi.

1910-1911

- **Morgan** inizia a mappare i geni sui cromosomi del moscerino della frutta *D.Melanogaster*

1944

- **Avery MacLeod e McCarty** stabiliscono che il DNA è il materiale ereditario.

1953

- **Watson e Crick** propongono la struttura a doppia elica per il DNA

1956

- Viene stabilito che il patrimonio cromosomico completo dell'uomo è di 46 cromosomi.

1961-1966

- Viene decifrato il codice genetico, ovvero stabilito il rapporto tra le 64 triplette possibili a partire dalle 4 basi nucleotidiche del DNA, e i 20 aminoacidi che formano le proteine.

1967

- Mary Weiss e Green introducono la tecnica dell'ibridazione delle cellule somatiche che rende più agevole la mappatura dei geni umani.

1972

- Berg costruisce la prima molecola DNA ricombinante in vitro utilizzando gli enzimi di restrizione. Vengono realizzati con successo i primi esperimenti di clonazione del DNA.

1973

- Cohen, Chang e Boyer costruiscono il primo batterio ricombinante. Si tiene la prima conferenza sulla mappatura dei geni umani.

1974

- Cohen e Boyer ottengono l'espressione di un gene estraneo trapiantato in un batterio con la tecnica del DNA ricombinante.

1975

- Si tiene la conferenza di Asilomar che propone una moratoria sugli esperimenti con il DNA ricombinante in assenza di linee guida per la sicurezza delle manipolazioni genetiche.

1977

- Per la prima volta un gene umano viene ricombinato e inserito in un batterio per clonare una proteina, la somatostatina **Sanger**, già premio Nobel come inventore del metodo per sequenziare le proteine, sviluppa un metodo nuovo ed efficiente per sequenziare il DNA (contemporaneamente **Gilbert e Maxam** inventano un metodo analogo).

1978

- Boyer costruisce una versione sintetica del gene dell'insulina e lo inserisce in un batterio. La Genentech, fondata nel 1976 dallo stesso Boyer insieme a Swanson, otterrà il permesso di commercializzare l'insulina prodotta per ingegneria genetica nel 1982 e quindi il brevetto

1978-1980

- Vengono sviluppate nuove tecniche basate sull'ibridazione molecolare e l'utilizzazione come marcatori delle variazioni individuali del DNA per mappare fisicamente i geni e le sequenze geniche sui cromosomi. Grazie alle nuove tecniche e alla collaborazione internazionale tra i mappatori, i geni mappati passano da 579 nel 1981 a 1.879 nel 1991: tra questi vengono mappati il gene della corea di Huntington (1983), il gene per la distrofia muscolare (1987) e il gene per la fibrosi cistica (1989).

1980

- Viene concesso il primo brevetto su una forma di vita geneticamente modificata, un microrganismo che si nutre di petrolio. Mullis inventa la tecnica della PCR.

1981

- Viene prodotto il primo animale transgenico.

1983

- Viene inventato il primo sequenziatore automatico (nella lettura dei risultati).

1989

- Viene creato il *National Center for Human Genome Research* (NCHGR), guidato da James Watson per mappare e sequenziare tutto il DNA umano entro il 2005

1995

- Viene introdotto nel mercato il primo sequenziatore interamente automatico (anche nel caricamento del campione attraverso capillare)

1977 : Frederick Sanger “inventa” la metodica tutt’ora in uso

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 ϕ X174)

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

Medical Research Council Laboratory of Molecular Biology, Cambridge CB2 3QH, England

Contributed by F. Sanger, October 3, 1977

ABSTRACT A new method for determining nucleotide sequences in DNA is described. It is similar to the “plus and minus” method [Sanger, F. & Coulson, A. R. (1975) *J. Mol. Biol.* 94, 441-448] but makes use of the 2',3'-dideoxy and arabinonucleoside analogues of the normal deoxynucleoside triphosphates, which act as specific chain-terminating inhibitors of DNA polymerase. The technique has been applied to the DNA of bacteriophage ϕ X174 and is more rapid and more accurate than either the plus or the minus method.

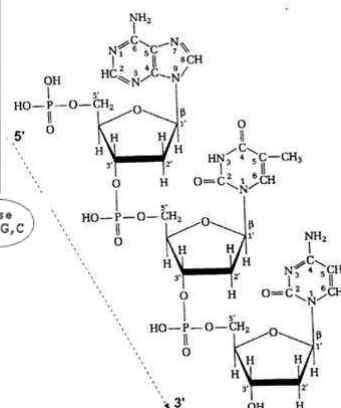
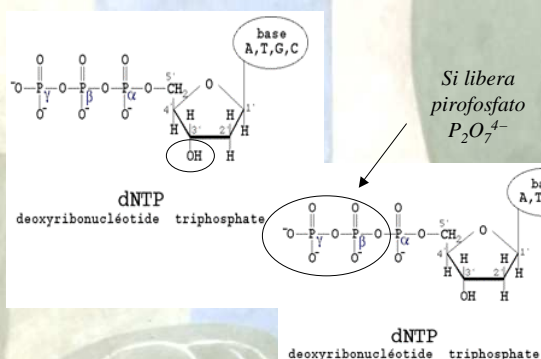
The “plus and minus” method (1) is a relatively rapid and simple technique that has made possible the determination of the sequence of the genome of bacteriophage ϕ X174 (2). It depends on the use of DNA polymerase to transcribe specific regions of the DNA under controlled conditions. Although the method is considerably more rapid and simple than other available techniques, neither the “plus” nor the “minus” method is completely accurate, and in order to establish a sequence both must be used together, and sometimes confirmatory data are necessary. W. M. Barnes (*J. Mol. Biol.*, in press)

a stereoisomer of ribose in which the 3'-hydroxyl group is oriented in *trans* position with respect to the 2'-hydroxyl group. The arabinosyl (ara) nucleotides act as chain terminating inhibitors of *Escherichia coli* DNA polymerase I in a manner comparable to ddT (4), although synthesized chains ending in 3' araC can be further extended by some mammalian DNA polymerases (5). In order to obtain a suitable pattern of bands from which an extensive sequence can be read it is necessary to have a ratio of terminating triphosphate to normal triphosphate such that only partial incorporation of the terminator occurs. For the dideoxy derivatives this ratio is about 100, and for the arabinosyl derivatives about 5000.

METHODS

Preparation of the Triphosphate Analogues. The preparation of ddTTP has been described (6, 7), and the material is now commercially available. ddA has been prepared by McCarty *et al.* (8). We essentially followed their procedure.

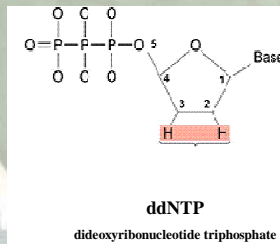
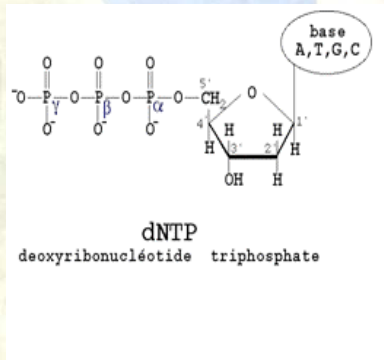
La reazione di polimerizzazione ad opera della DNA polimerasi...



La DNA polimerasi catalizza la reazione che porta al legame fosfo-diester tra gruppo **gruppo OH (al 3')** del nucleotide incorporato e gruppo fosfato (al 5') del nucleotide entrante.

Chimica del metodo Sanger

I dideossiribonucleotidi



La mancanza del gruppo OH al 3' non consente l'estensione del filamento di DNA...

Se proviamo ad eseguire una “specie di PCR” ma :

- 1) con un singolo primer (quindi un solo filamento sarà interessato da questa reazione...)
- 2) Con dATP dCTP dTTP (deossiribonucleotidi)
- 3) Con ddGTP (di-deossiribonucleotide)

STOP!

5' **AATGCCCTGAT** a a c a c t **g**3' filamento nascente

3' **TTACGGGACTATTGTGTGACTGCTA**..... filamento stampo

In tutti i filamenti nascenti la polimerizzazione si bloccherà al nucleotide 20...

Otterremo solo filamenti da 20 nucleotidi che termineranno tutti con una G.

Ma...

..se usiamo una miscela di nucleotidi così composta :

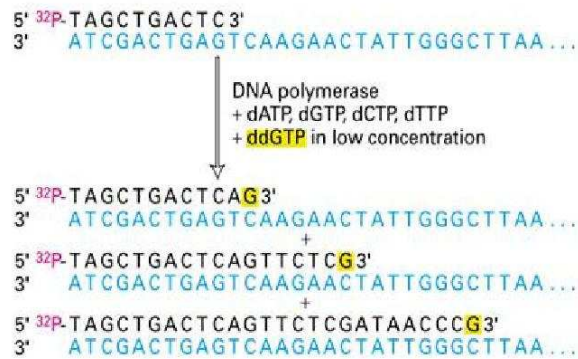
dATP dCTP dTTP (deossiribonucleotidi al 100%)

dGTP al 99%

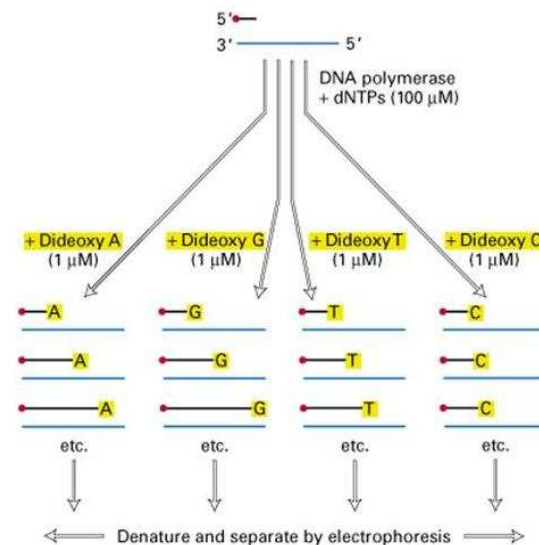
ddGTP all'1%

Ad ogni
incorporazione di una
G statisticamente
diamo la possibilità ad
un buon numero di
filamenti di proseguire
nella estensione...

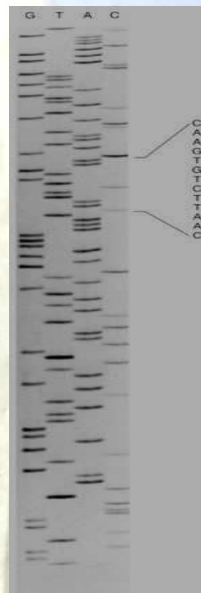
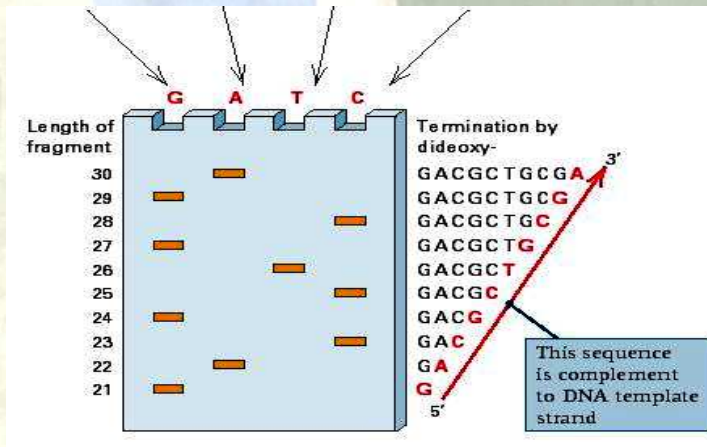
(b)



Se usiamo un primer marcato con radioattivo (³⁵S o ³²P)



Si ottengono dei frammenti di DNA di lunghezza diversa, che vengono poi separati elettroforeticamente su gel di poliacrilamide.



Fine anni 70

Utilizzo di un primer marcato con radioattivo, difficoltà nella corretta esecuzione dell'elettroforesi su gel e sviluppo di autoradiografie.

Quattro reazioni per ogni filamento...

Lettura massima di sequenze di 150 nucleotidi...

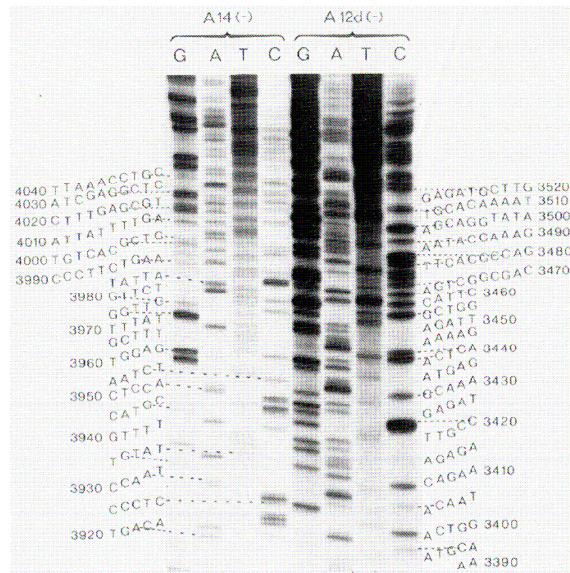
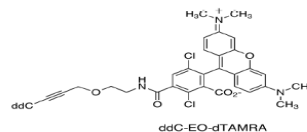
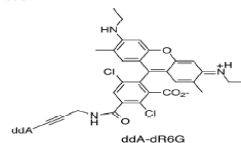
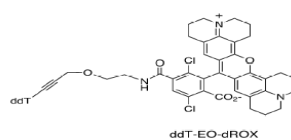
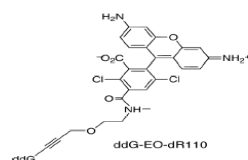


FIG. 1. Autoradiograph of the acrylamide gel from the sequence determination using restriction fragments A12d and A14 as primers on the complementary strand of ϕ X174 DNA. The inhibitors used were (left to right) ddGTP, ddATP, ddTTP, and araCTP. Electrophoresis was on a 12% acrylamide gel at 40 mA for 14 hr. The top 10 cm of the gel is not shown. The DNA sequence is written from left to right and upwards beside the corresponding bands on the radioautograph. The numbering is as given in ref. 2.

Anni 1986 : introduzione dei marcatori fluorescenti Primi sequenziatori automatici in commercio

I fluorocromi sono molecole grandi e ricche di doppi legami, che se colpiti da luce sono in grado di assorbire questa energia ed emettere a loro volta una luce di una con una lunghezza d'onda caratteristica per ciascuno di loro.

Terminator	Dye Label
A	dichloro[R6G]
C	dichloro[TAMRA]
G	dichloro[R110]
T	dichloro[ROX]



I fluorocromi : un altro passo avanti

Unica reazione in una sola provetta, ed unica corsa;
Ad ogni nucleotide è associata una diversa
"colorazione" (cioè una emissione di luce di lunghezza
d'onda diversa):

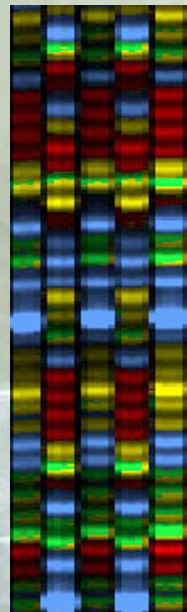
ddTTP
ddATP
ddGTP
ddCTP

Questo sistema aumenta la produttività ed elimina
la variabilità di corsa su gel di poliacrilammide.

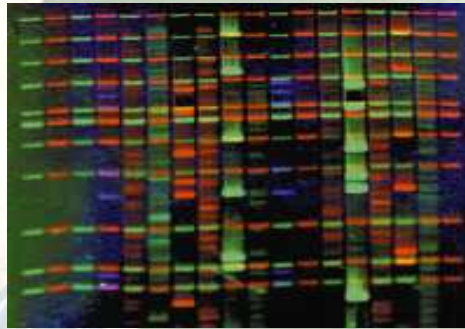
Il sequenziamento automatizzato con marcatori fluorescenti

Coniugando a ciascun ddNTP
un diverso marcatore fluorescente, è
possibile effettuare le quattro reazioni di
sequenziamento in un unico tubo da saggio
e caricare il tutto in un solo pozzetto di gel

ddA 
ddT 
ddC 
ddG 

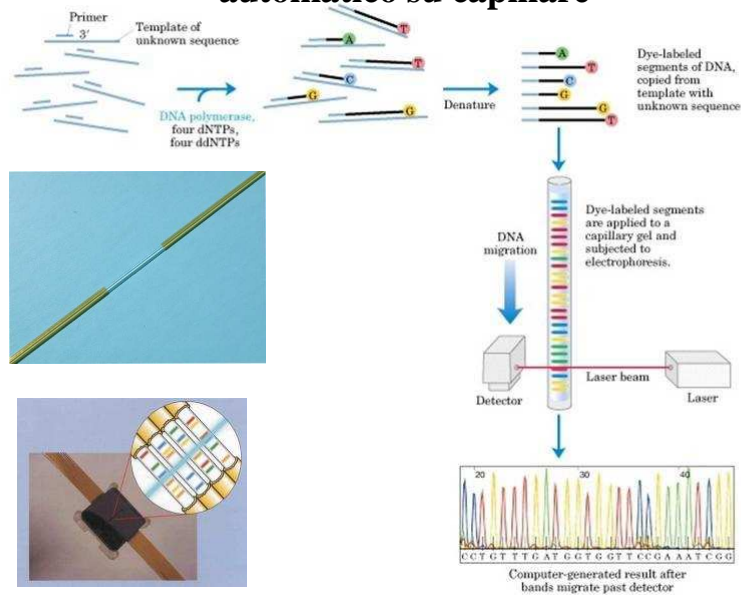


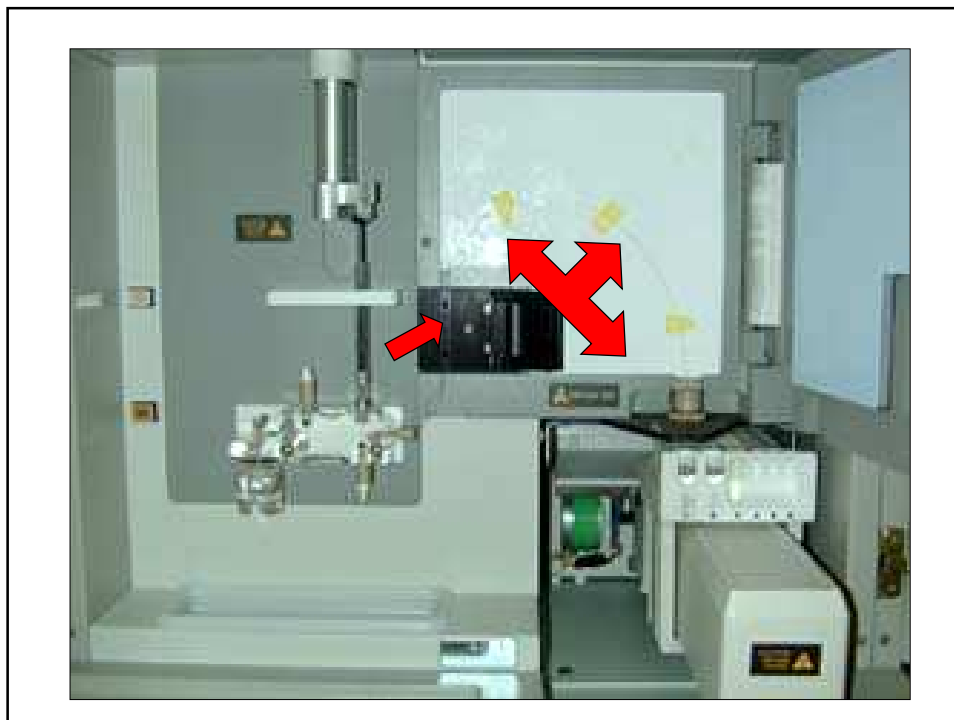
Seconda metà anni 80



Su un gel su supporto a lastra il caricamento avviene manualmente. La migrazione dei vari frammenti è seguita rilevando le emissioni in fluorescenza a diverse lunghezze d'onda dei diversi fluorocromi dopo l'eccitazione provocata dal laser.

1995 Una ulteriore evoluzione: il sequenziamento automatico su capillare



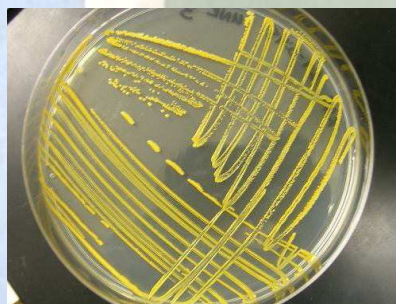


Sequenziamento automatico con metodo Sanger

Le fasi

- ★ Estrazione di DNA (dipende dalla matrice)
- ★ PCR
- ★ Purificazione del campione di DNA (per eliminare Sali, nucleotidi liberi, ma soprattutto residui dei primers usati per l'amplificazione), a volte è necessario purificare la banda di DNA dal gel di agarosio.
- ★ Quantificazione (la quantità di DNA da usare per il Cycle seq. è un fattore critico)
- ★ Cycle Sequencing (almeno 2 sequenziamenti per campione per ottenere una sequenza "consensus")
- ★ Purificazione dei prodotti di cycle sequencing (per eliminare i sali ma soprattutto i fluorocromi "liberi")
- ★ Sequenziamento automatico
- ★ Analisi dei dati (software e banche dati)

Gli stafilococchi meticillino-resistenti (MRSA)



Gram positivi, potenzialmente patogeni, possono anche essere produttori di enterotossine.

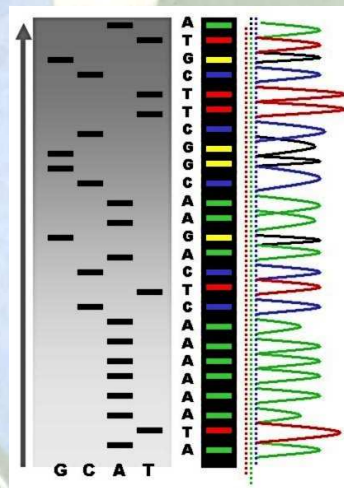
Sanità animale e alimentare, una applicazione: lo spa TYPING

Consiste nella tipizzazione di un singolo locus, la regione polimorfica X del gene per la proteina A di *S.aureus*. Questa regione presenta un polimorfismo che consiste nella ripetizione in numero variabile di una unità di 24pb.

Sequenziare questa regione dopo averla amplificata per PCR, consente di assegnare un codice univoco per identificare (tipizzare) il ceppo batterico in esame.

Un ulteriore vantaggio offerto riguarda la possibilità di collezionare dati continuamente con lo scopo di tenere sotto controllo il diffondersi di un'infezione, e la possibilità di sviluppare un algoritmo elettronico che individui automaticamente il pericolo dell'insorgenza di epidemie.

1977



...e poi?

Grazie per l'attenzione!