

Disciplina: Bioinformática
Bio1015



O Sequenciamento genômico

Prof. Macks Wendhell Gonçalves, Msc
mackswendhell@gmail.com

Sequenciamento genômico

- O sequenciamento genômico é uma técnica que permite identificar, na ordem correta, a sequência de nucleotídeos de uma molécula de DNA ou RNA.
- O sequenciamento de DNA, seguido da montagem genômica, permite obter informações sobre:
 - Estrutura e função dos genes
 - Diversidade genética
 - Presença de elementos móveis no genoma
 - Presença de genes adquiridos por transferência lateral
 - Relações evolutivas
 - Construção de mapas metabólicos

Sequenciamento genômico

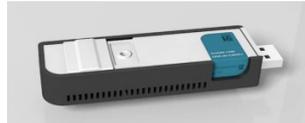
- O que o sequenciamento de DNA genômico não é capaz de responder?
 - Quais genes estão sendo expressos no momento do ensaio
 - Qual o nível da expressão estes genes se encontram
- Para se obter essas informações é necessário o sequenciamento do RNA (cDNA) por meio de técnicas de *RNA-Seq*. Desta forma é possível responder:
 - Se determinado gene é importante a uma situação específica
 - Se o gene sofre algum tipo de regulação da expressão a nível transcricional
 - Este é o caso dos experimentos com ESTs e ORESTES

Métodos de sequenciamento

- **Sequenciamento de 1º Geração**
 - Degradação química – Maxam & Gilbert
 - Interrupção da cadeia (ddNTPs) – Sanger

- **Sequenciamento de 2º Geração**
 - 454 – Roche
 - Illumina – HiSeq, MiSeq

- **Sequenciamento de 3º Geração**
 - Ion Torrent – ABI -Life Technologies
 - Nanopore – GridIon / MiniIon
 - PacBio RS – Pacific Biosciences



Sequenciamento do DNA: 1ª geração



Walter Gilbert

Baseado em hidrólise
química da molécula de
DNA



Frederick Sanger

Sequenciamento
enzimático dideoxi
'término de cadeia'

Sequenciamento do DNA: Maxam-Gilbert

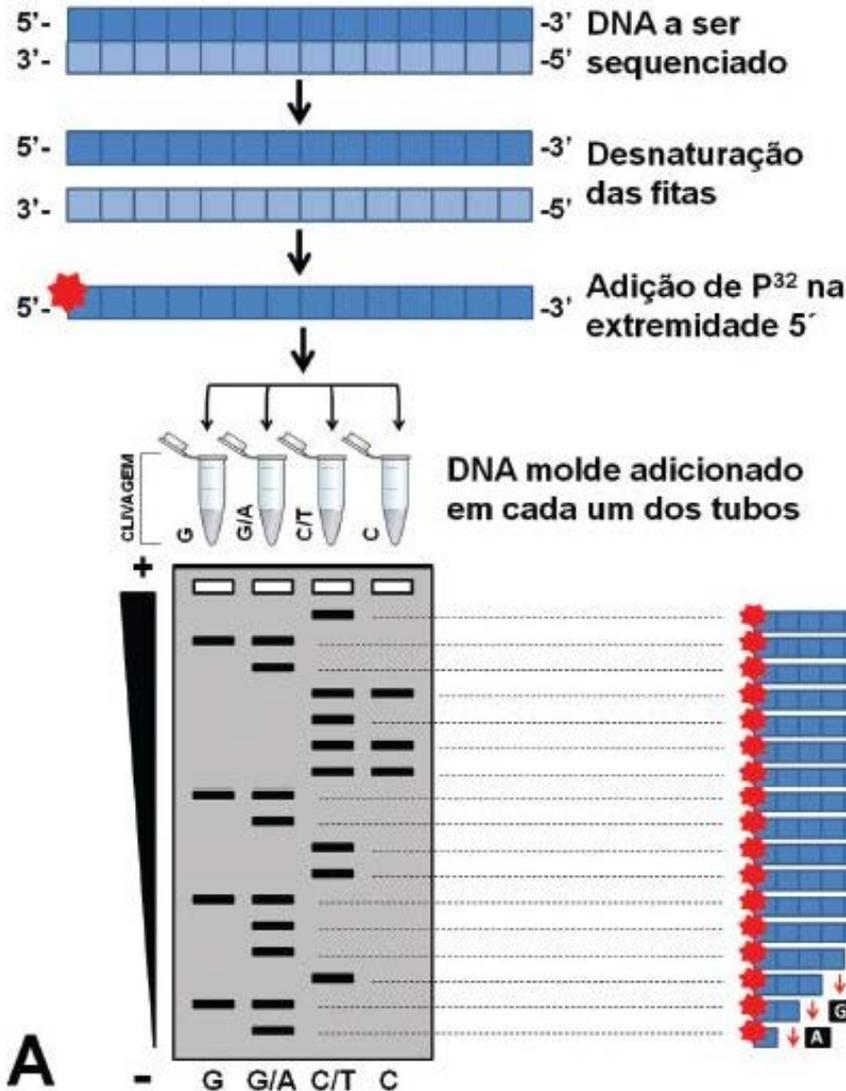
- A técnica desenvolvida por eles utiliza marcação do DNA alvo a ser sequenciado com fósforo radioativo (P^{32}).
- O P^{32} é ligado a um dATP formando um P^{32} -dATP. Essa molécula é incorporada ao DNA molde (extremidades 5' ou 3') pela enzima polinucleotídeo kinase.
- Neste método, o rompimento das pontes de hidrogênio da fita dupla de DNA ocorre pela adição de dimetil sulfato e aquecimento a 90° C.



Walter Gilbert

Descoberta em conjunto com
seu aluno de pós-graduação
Allan M. Maxam

Sequenciamento do DNA: Maxam-Gilbert



A – Após a desnaturação, o DNA molde marcado com P³² é clivado em 4 regiões (antes do “G”, “G/A”, “C/T” e “C”s). Para identificar o nucleotídeo aplica-se o produto dessa reação em 4 canaletas do gel.

B – Representação esquemática evidenciando o último nucleotídeo de cada um dos fragmentos (diferentes tamanhos) do gel de poliacrilamida.

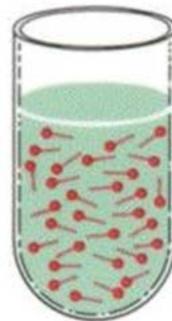
C 5'- [AGTAAGTTAGCCTCAGT] -3'

C – Sequências de nucleotídeos do DNA sequenciado.

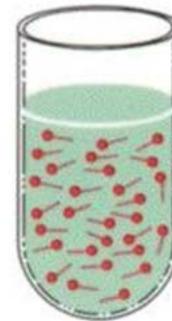
Sequenciamento do DNA: Maxam-Gilbert

Tabela 1. Compostos químicos utilizados na técnica de Maxam-Gilbert e especificidades pelas bases nitrogenadas.

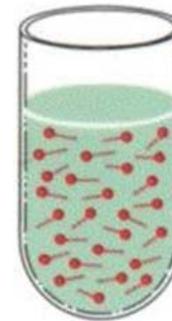
Especificidade de bases	Modificador de bases	Remover de bases	Químico para clivar a fita
G	Dimetil sulfato	Piperidina	Piperidina
A + G	Ácido	Ácido	Piperidina
C + T	Hidrazina	Piperidina	Piperidina
C	Hidrazina + álcali	Piperidina	Piperidina
A > C	Álcali	Piperidina	Piperidina



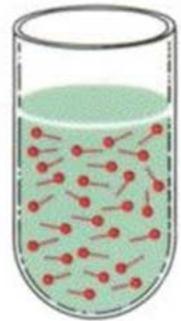
Base G



Base A/G



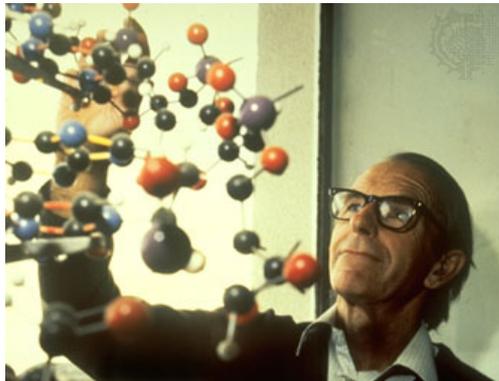
Base T/C



Base C

Sequenciamento DNA: Sanger

- 1975: Propôs melhorias significativas do método de Maxam-Gilbert. Inventou o método dideoxi para o sequenciamento do DNA
- 1977: sequencia, pela primeira vez, o genoma inteiro de um organismo, o bacteriófago phi X174
- 11 genes em 5.386 bases sequenciadas “à mão“
- Ganha o segundo Nobel de química em 1980



1918 - 2013

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 2QH, 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 confirma-

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

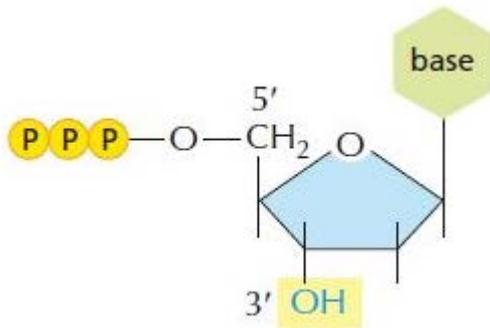
Preparation of the Triphosphate Analogues. The preparation of ddTTP has been described (6, 7), and the material is now commercially available. dda has been prepared by

Sequenciamento DNA: Sanger

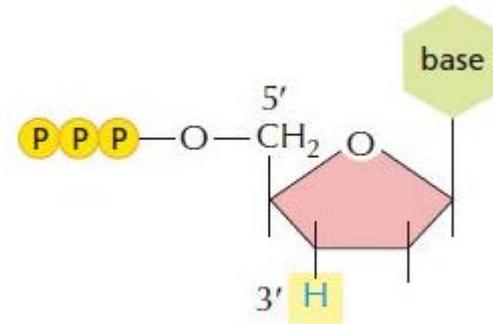
- Assim como a metodologia de Maxam-Gilbert, a técnica desenvolvida por Sanger também utiliza marcação radioativa.
- *A diferença é que a primeira marcava diretamente o DNA molde. Já a de Sanger marca fragmentos de DNA sintetizados a partir de uma fita molde. A síntese desses fragmentos a partir de um DNA molde foi possível graças a PCR.*
- Por fim, é necessário ainda a presença de didesoxinucleotídeos (ddATP, ddCTP, ddGTP e ddTTP), que atuam como terminadores da síntese de DNA.

Sequenciamento DNA: Sanger

- Diferença do desóxinucleotídeo com o didesóxinucleotídeo. As chances de a DNA polimerase reconhecer essas duas moléculas é a mesma.



Desóxinucleotídeo



Didesóxinucleotídeo

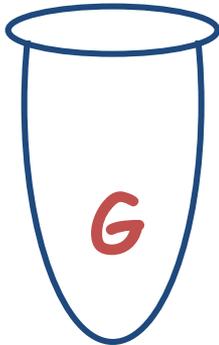
O que faria um nucleotídeo que, ao invés da extremidade 3'OH, tivesse uma extremidade 3'H

???

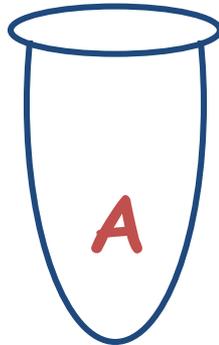
Sequenciamento DNA: Sanger

- DNA molde
- DNA polimerase
- Tampão
- dNTPs

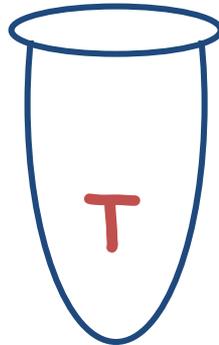
•ddGTPs



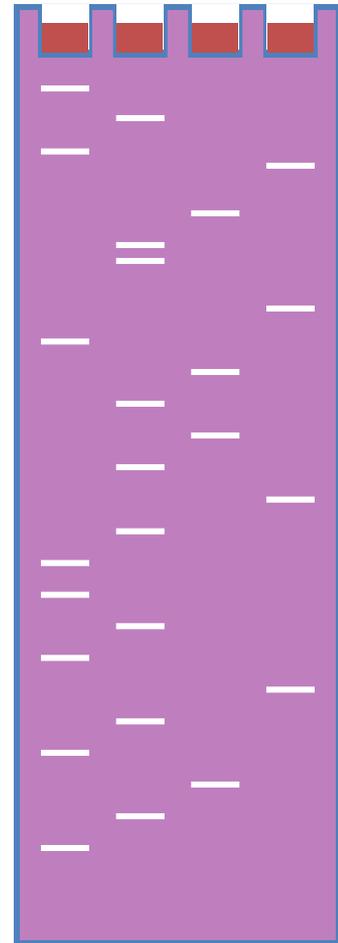
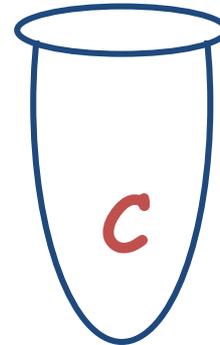
•ddATPs



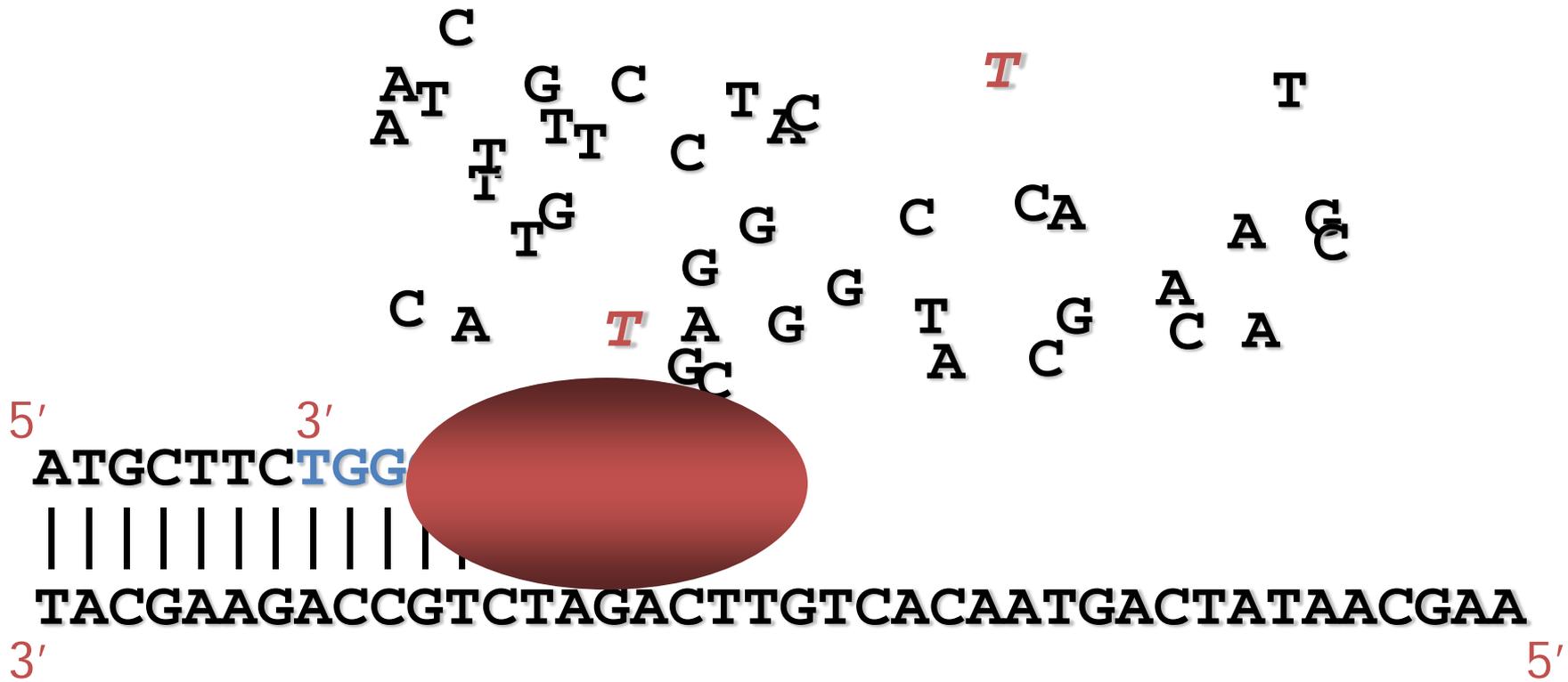
•ddTTPs



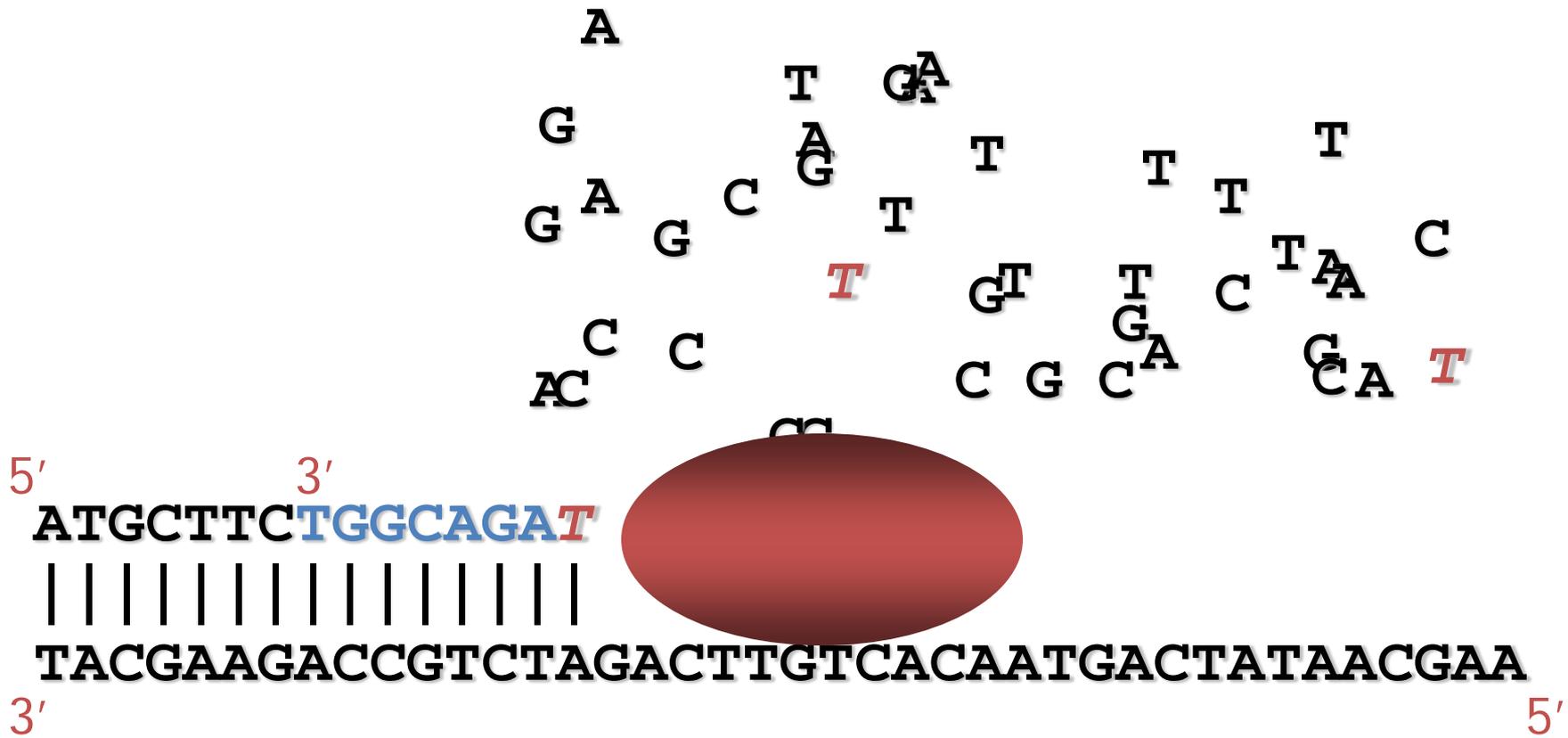
•ddCTPs



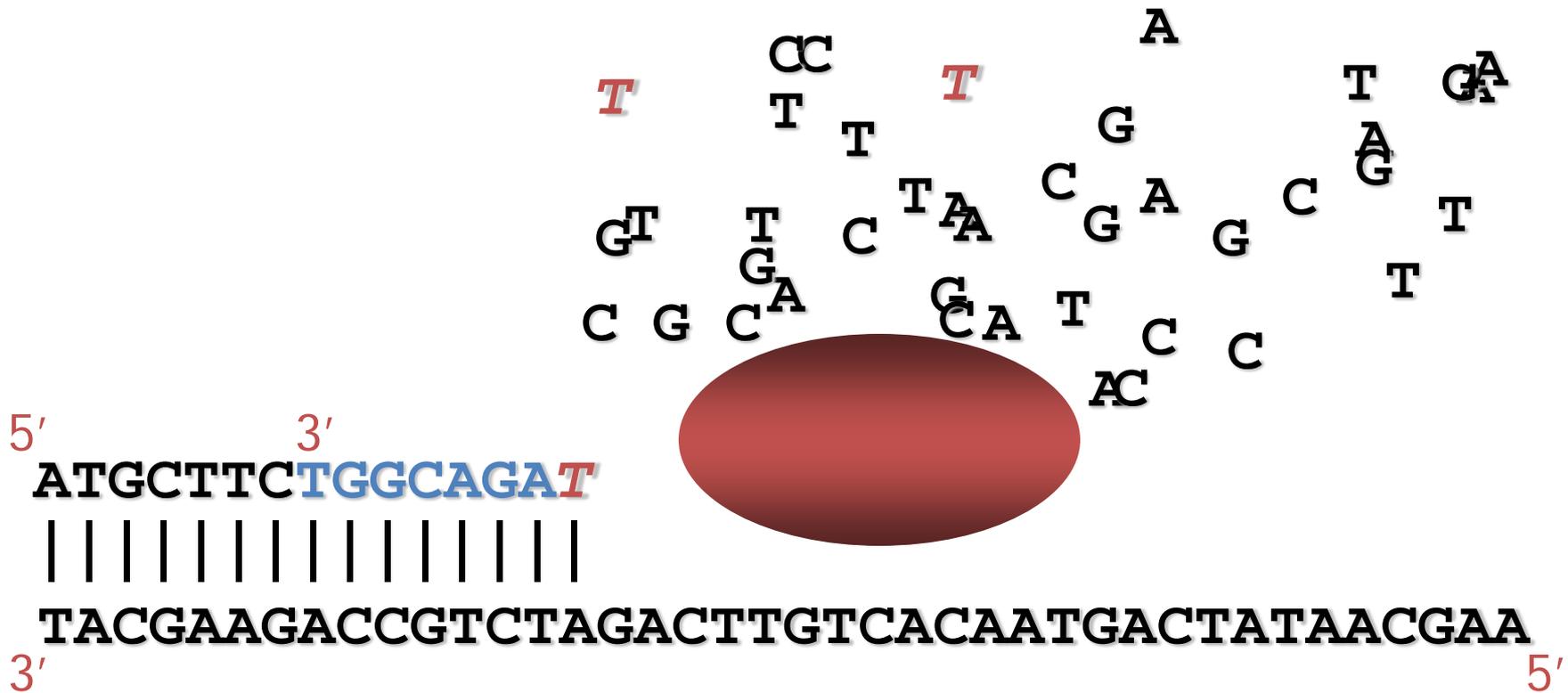
Sequenciamento DNA: Sanger



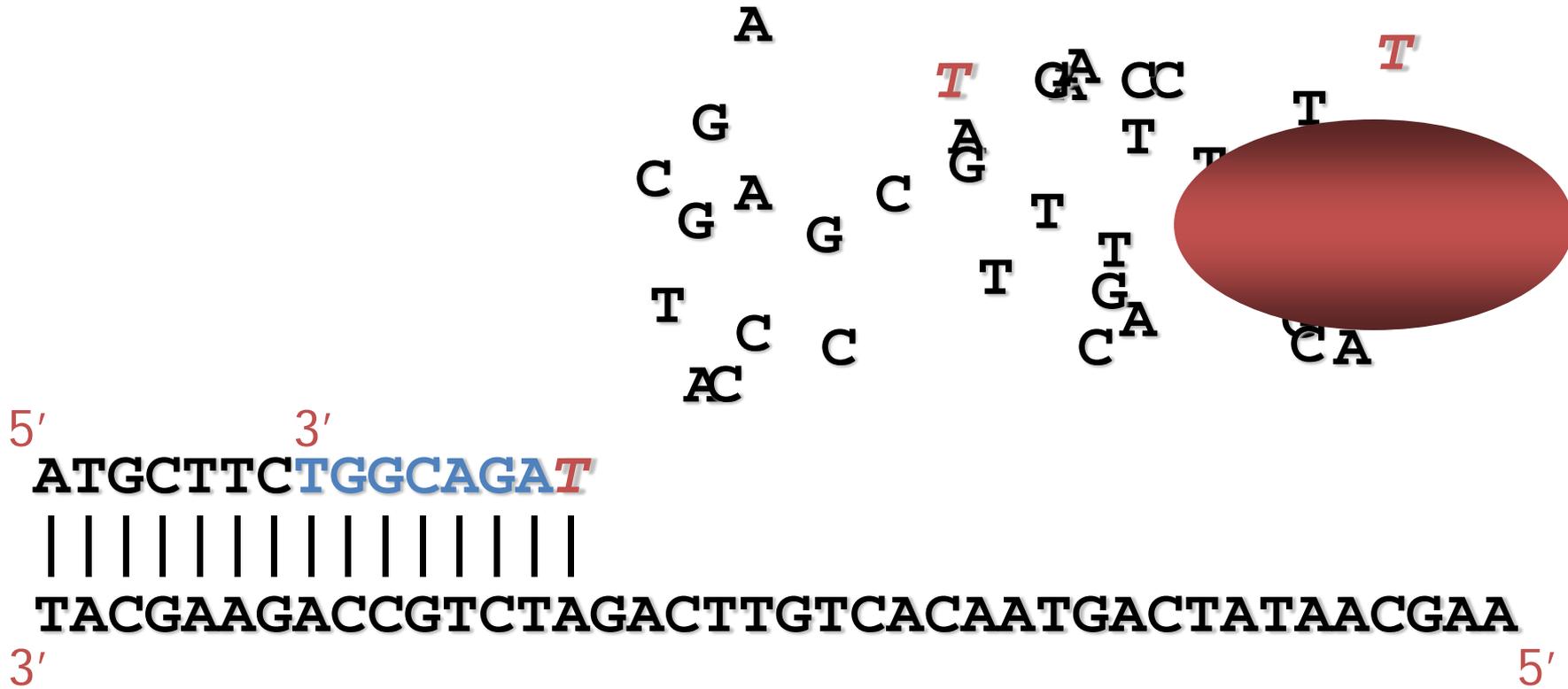
Sequenciamento DNA: Sanger



Sequenciamento DNA: Sanger

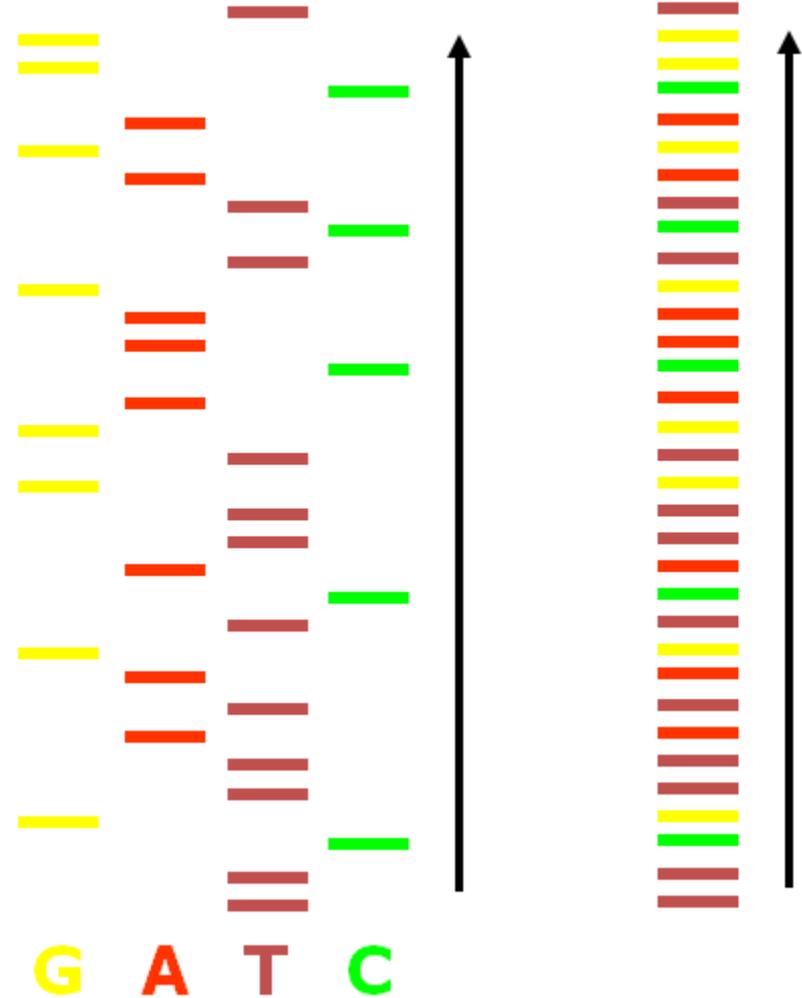


Sequenciamento DNA: Sanger



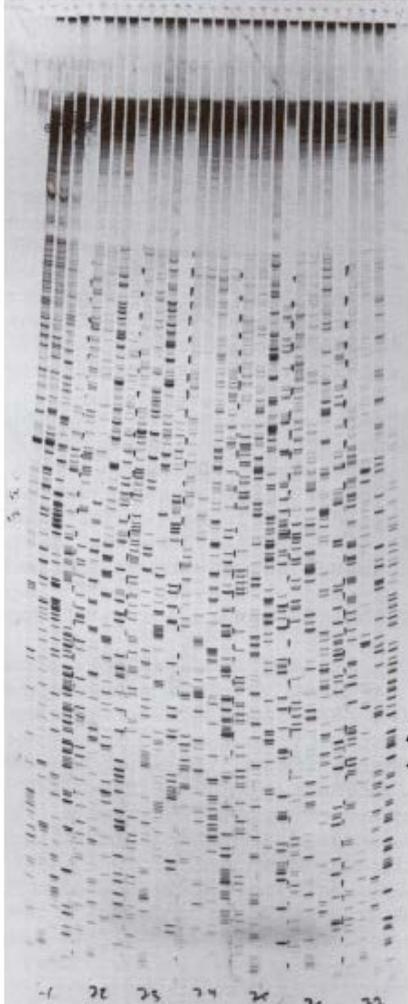
Sequenciamento DNA: Sanger

ATGCTTCT
 ATGCTTCTG
 ATGCTTCTGG
 ATGCTTCTGGC
 ATGCTTCTGGCA
 ATGCTTCTGGCAG
 ATGCTTCTGGCAGA
 ATGCTTCTGGCAGAT
 ATGCTTCTGGCAGATC
 ATGCTTCTGGCAGATCT
 ATGCTTCTGGCAGATCTG
 ATGCTTCTGGCAGATCTGA
 ATGCTTCTGGCAGATCTGAA
 ATGCTTCTGGCAGATCTGAAC
 ATGCTTCTGGCAGATCTGAACA
 ATGCTTCTGGCAGATCTGAACAG
 ATGCTTCTGGCAGATCTGAACAGT
 ATGCTTCTGGCAGATCTGAACAGTG
 ATGCTTCTGGCAGATCTGAACAGTGT
 ATGCTTCTGGCAGATCTGAACAGTGTT
 ATGCTTCTGGCAGATCTGAACAGTGTTA
 ATGCTTCTGGCAGATCTGAACAGTGTTAC
 ATGCTTCTGGCAGATCTGAACAGTGTTACT
 ATGCTTCTGGCAGATCTGAACAGTGTTACTG
 ATGCTTCTGGCAGATCTGAACAGTGTTACTGA
 ATGCTTCTGGCAGATCTGAACAGTGTTACTGAT
 ATGCTTCTGGCAGATCTGAACAGTGTTACTGATA
 ATGCTTCTGGCAGATCTGAACAGTGTTACTGATAT
 ATGCTTCTGGCAGATCTGAACAGTGTTACTGATATT
 ATGCTTCTGGCAGATCTGAACAGTGTTACTGATATTG
 ATGCTTCTGGCAGATCTGAACAGTGTTACTGATATTGC
 ATGCTTCTGGCAGATCTGAACAGTGTTACTGATATTGCT
 ATGCTTCTGGCAGATCTGAACAGTGTTACTGATATTGCTT

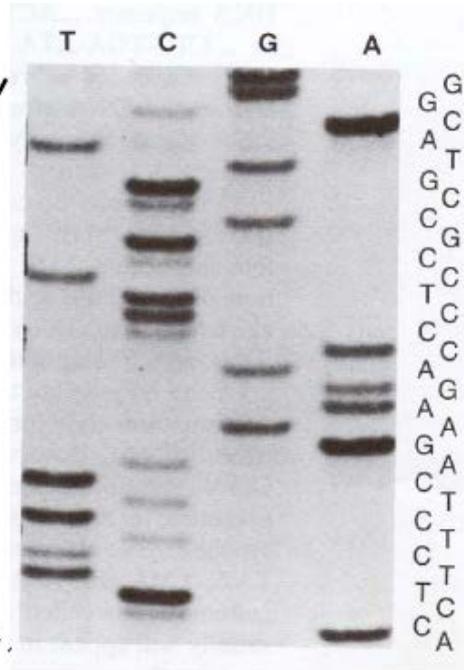


Sequenciamento DNA: Sanger

- **Sanger e o fago Phi 174**



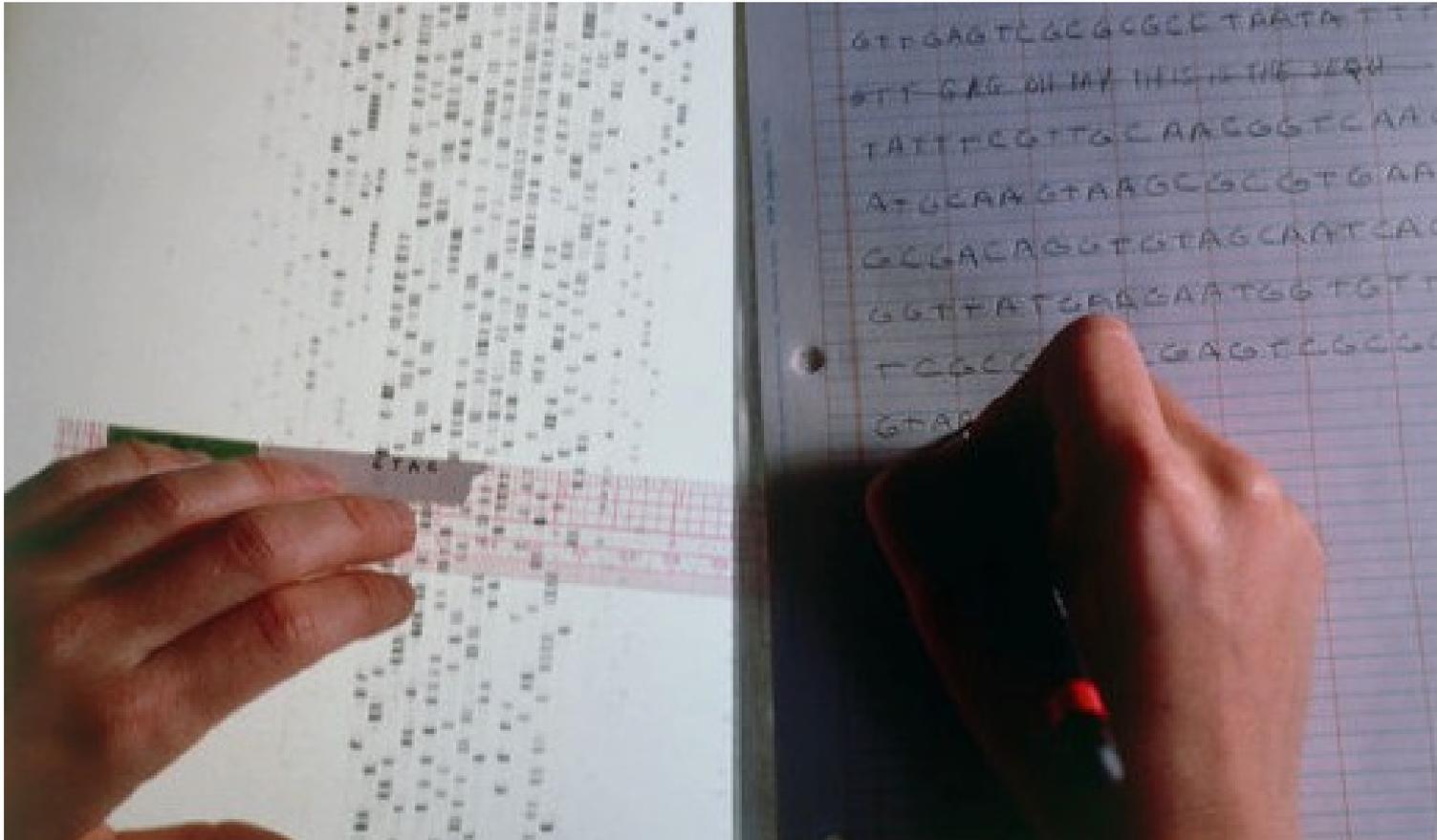
Autorradiograma de um gel de sequenciamento de DNA



De fato, esse tipo de sequenciamento manual continuou acontecendo por décadas...

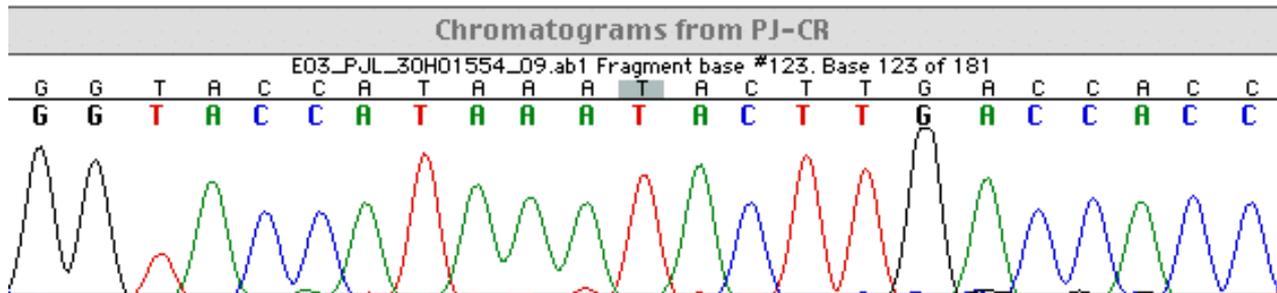
Sequenciamento DNA: Sanger

- **Desvantagem**



Sequenciamento DNA: Método Semiautomático

- A partir do aprimoramento do sequenciamento de Sanger surgiu o método de sequenciamento semiautomático.
- A principal modificação foi, no lugar de compostos radioativos, adicionar aos dideoxinucleotídeos corantes capazes de emitir fluorescência quando excitados em comprimento de onda específico.
- Como os dideoxi passaram a ser marcados com fluoróforos, agora as 4 reações passaram a poder ocorrer no mesmo tubo, aumentando 4x a velocidade do ensaio.



ABI 3730xl



Sequenciamento DNA: Automatizado

- Nos métodos automatizados os géis de difícil manuseio foram substituídos por finíssimos capilares preenchidos com gel onde os fragmentos de DNA são separados em altíssima velocidade.
- Após a eletroinjeção os fragmentos começam a migrar até passarem por um feixe de laser que excita os fluoróforos em cada dNTP fazendo com que esses emitam uma fluorescência característica de cada nucleotídeo.
- Um detector registra esta fluorescência e a transmite para um computador que possui um software capaz de converter fluorescência em picos coloridos, reconhecendo os 4 nucleotídeos.

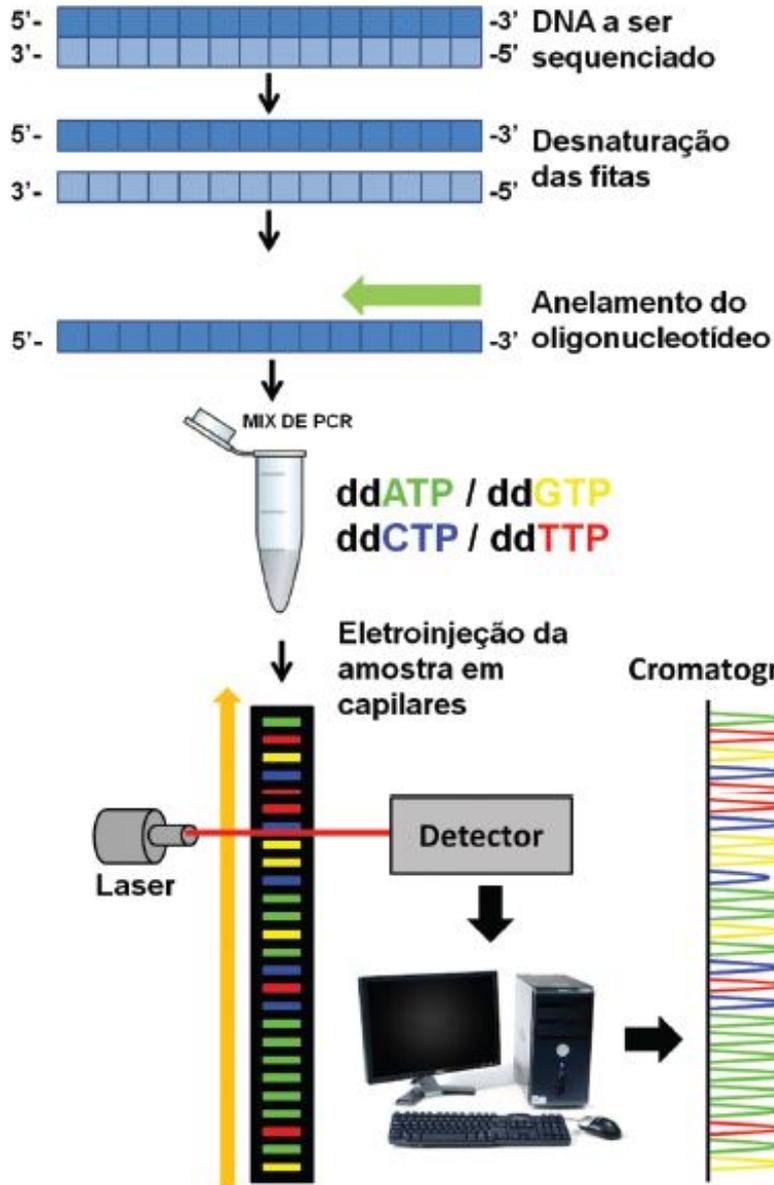
 A

 G

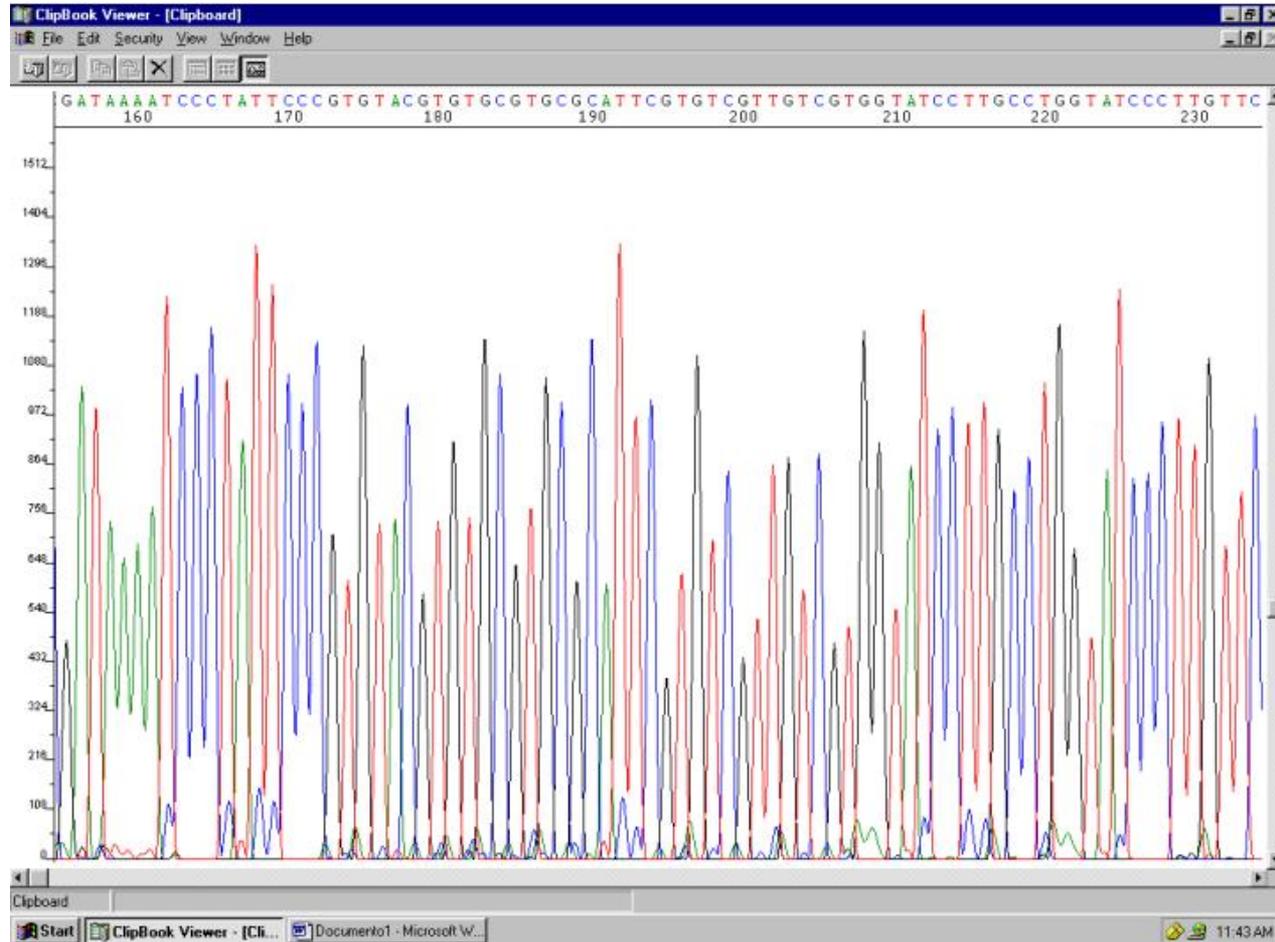
 C

 T

Sequenciamento DNA: Automatizado



Sequenciamento DNA: Automatizado



Sequenciamento DNA: Automatizado

- Modelos de sequenciadores automáticos que fazem uso da técnica de Sanger.



MEGA BACE

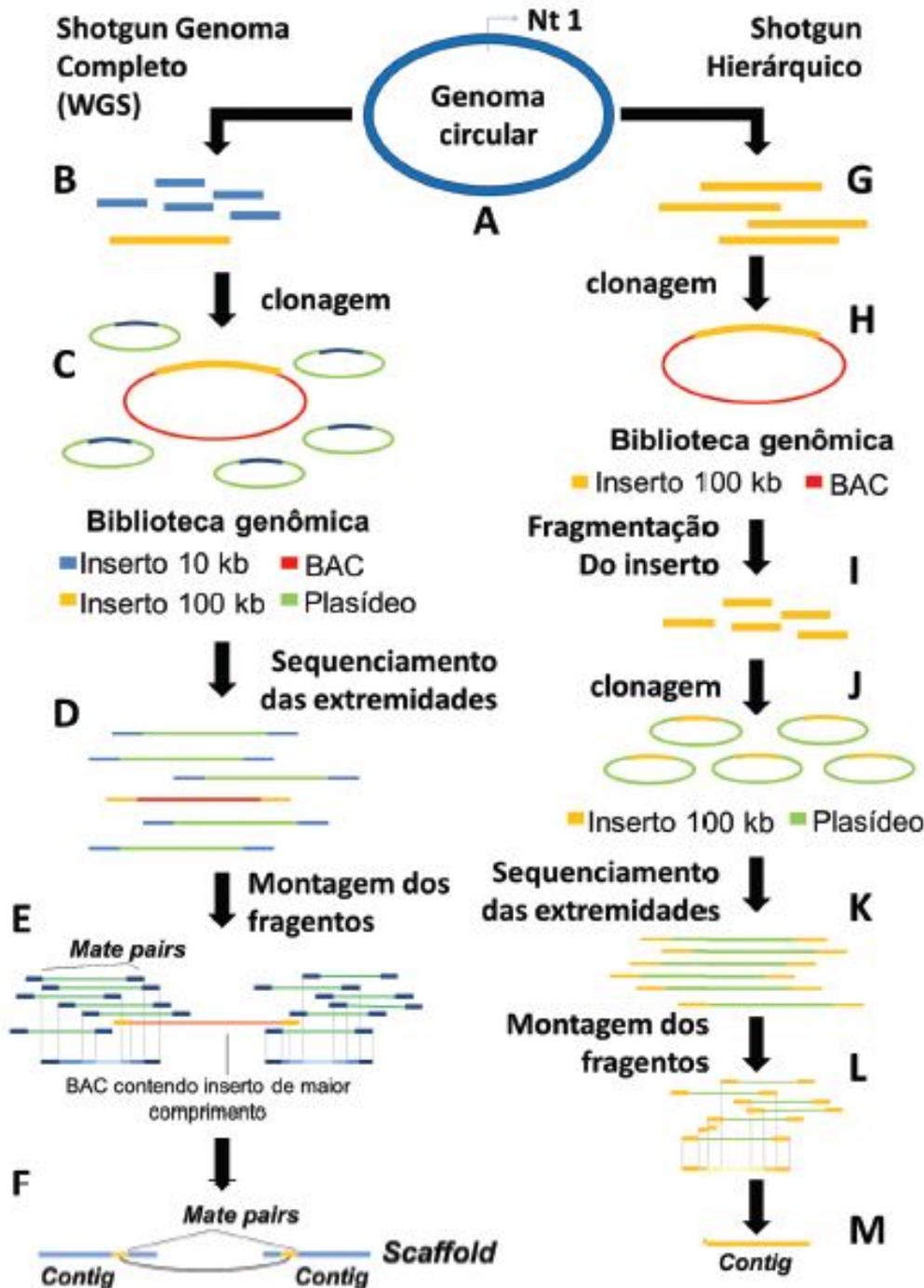


ABI PRISM 3700



ABI PRISM 3100

Sequenciamento DNA: Shotgun

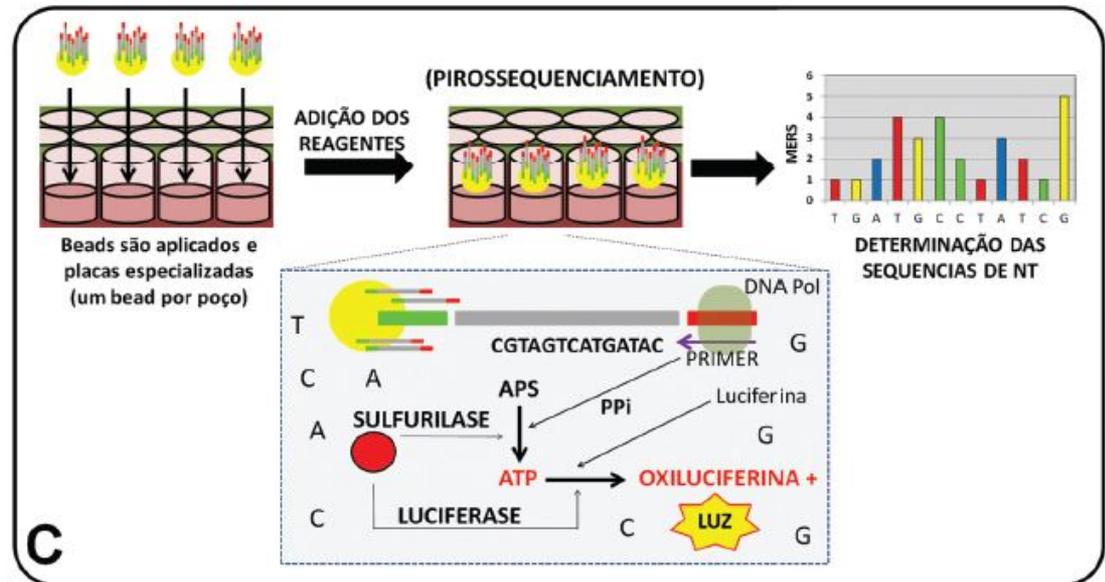
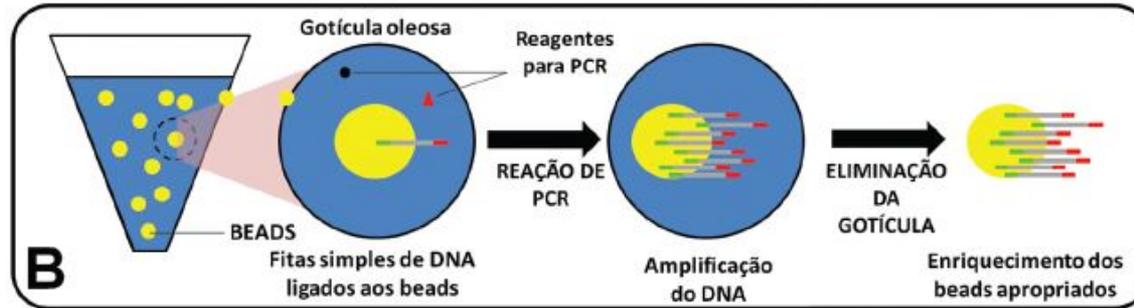
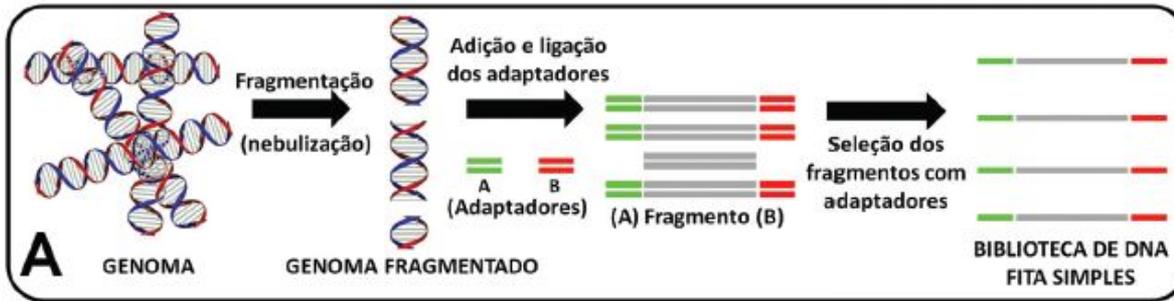


O genoma humano foi sequenciado utilizando essa metodologia de sequenciamento.

Sequenciamento de Nova Geração: 2º geração

- Após a publicação do rascunho do genoma humano, houve um avanço nas tecnologias de sequenciamento culminando no surgimento dos “sequenciadores de segunda geração”.
- **Principais plataformas:**
 - Plataforma 454 (Roche)
 - Plataforma Illumina
 - Plataforma SOLID

Sequenciamento de 2º geração: Plataforma 454



Sequenciamento de 2º geração:

- Após a divulgação do sequenciamento do genoma humano surge os sequenciadores de 2º geração reduzindo significativamente os custos.

