



Universidade de São Paulo
Instituto de Química

Sequenciamento de DNA e PCR

QBQ107

Prof. João Carlos Setubal

PCR

- Polymerase chain reaction
- Reação em cadeia da polimerase
- É um processo **tecnológico**
- Ou seja, inventado por seres humanos para atingir objetivos específicos nossos

Qual polimerase?

- DNA polimerase

Polimerases de DNA: As enzimas que sintetizam DNA

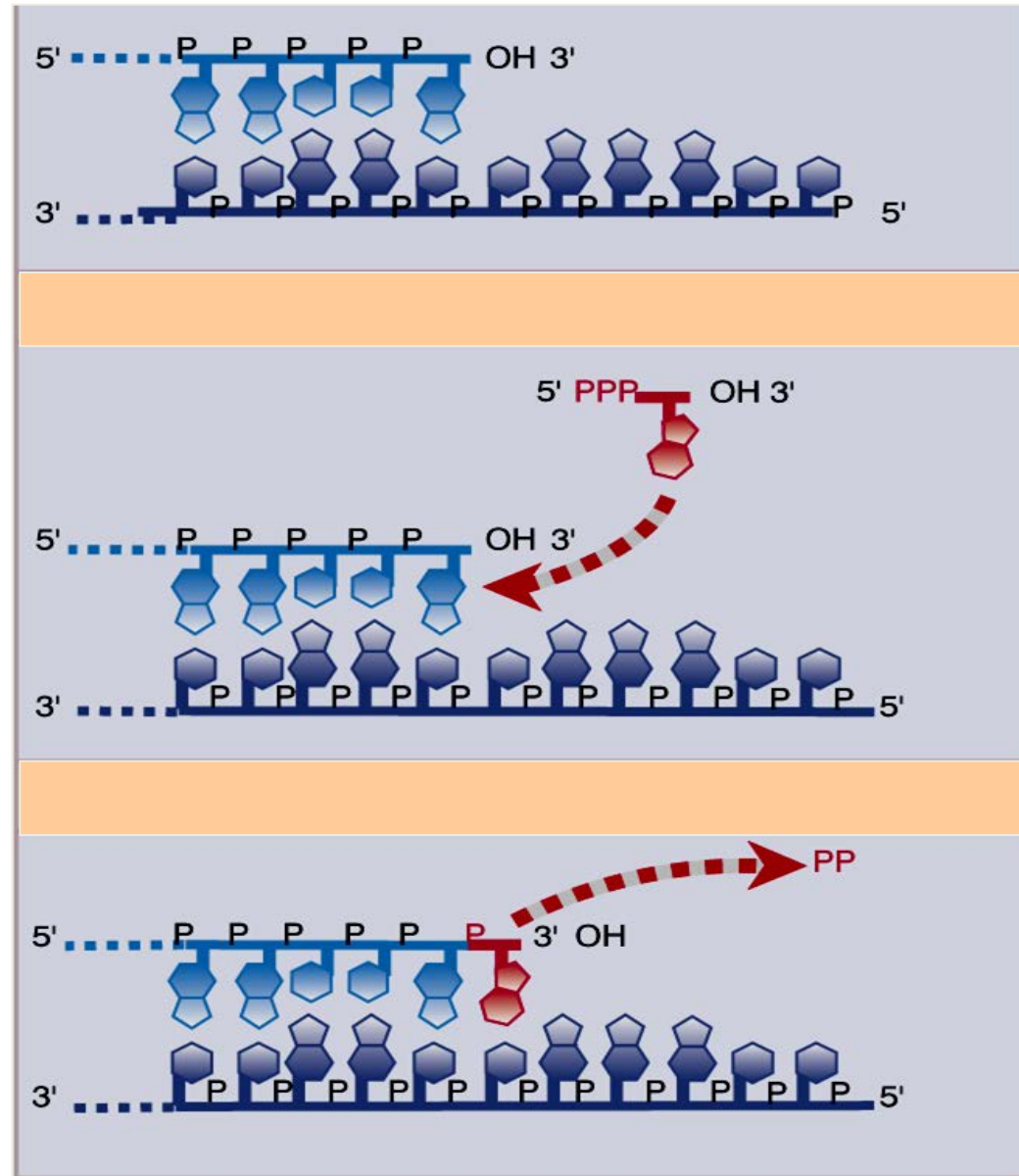
- A síntese de DNA ocorre pela adição de nucleotídeos a extremidade 3'OH da cadeia em crescimento.

- A DNA polimerase requer um *primer* (iniciador) e um molde

- O precursor da síntese é desoxirribonucleosídeo 5' trifosfato

- Sentido da síntese sempre é 5' → 3'

- A replicação é um processo extremamente fiel. As DNA-polimerases tem atividade revisora



O processo de PCR

- Procura **imitar** o processo de replicação de DNA
 - o que explica o uso da DNA polimerase
- Objetivo é aumentar (“**amplificar**”) a quantidade de uma certa sequência de DNA de interesse
 - com mais DNA, é possível fazer mais coisas com esse DNA
- Inventada na década de 80
- Seu inventor ganhou o prêmio Nobel em 1993
 - Kary Mullis (1944 – 2019)

É um processo de crescimento exponencial

- Como na lenda do jogo de xadrez...



- 1 grão na primeira casa
- 2 grãos na segunda casa
- 4 grãos na terceira casa
- ...etc
- Quantos grãos na última casa?

$$2^{n-1}$$

$$2^{63}$$

$$\sim 10^{19}$$

$$1 \text{ trilhão} = 10^{12}$$

10 milhões x 1 trilhão

Ingredientes para PCR

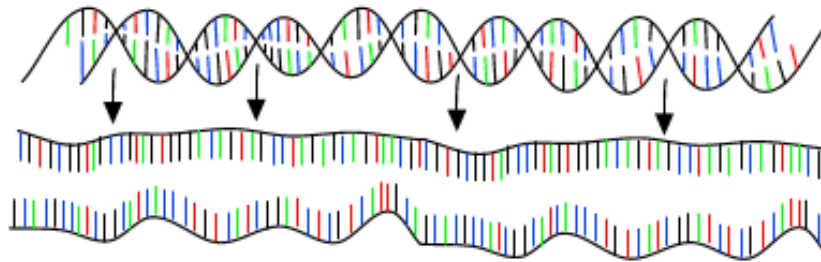
- O DNA **molde** que contém a região do DNA que se deseja amplificar (o alvo)
- Dois **primers** que são complementares às pontas 3' da fita senso e da fita anti-senso do DNA alvo
- Polimerase Taq que funcione a **70 °C**
- *Deoxinucleosídeos trifosfato* (dNTPs) ou seja, nucleotídeos contendo grupos trifosfato
- Tampão (um ambiente químico adequado para ação da polimerase)
- Cations e ions de magnésio e potássio

Ciclos do PCR

- A ideia é fazer o DNA (e os ingredientes) passarem **várias vezes** por **3 passos**
 1. Passo de **desnaturação**, em que DNA **fita dupla** passa a ser de **fita simples**
 2. Passo de **anelamento**, em que os primers se acoplam aos DNAs de fita simples
 3. Passo de **extensão** do DNA, quando a replicação de fato ocorre

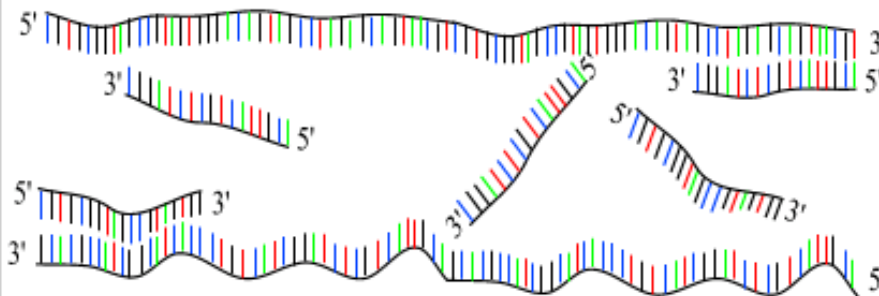
PCR : Polymerase Chain Reaction

30 - 40 cycles of 3 steps :



Step 1 : denaturation

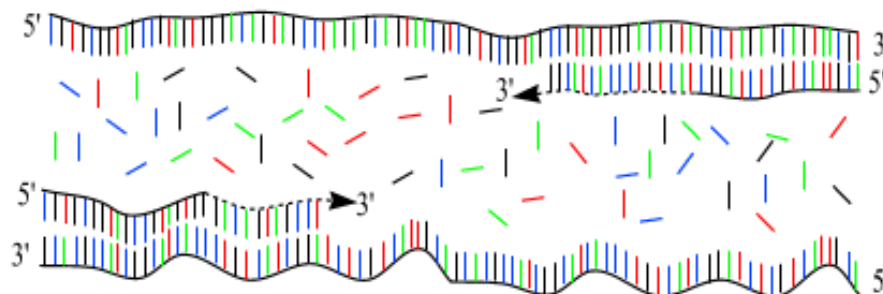
1 minut 94 °C



Step 2 : annealing

45 seconds 54 °C

forward and reverse
primers !!!



Step 3 : extension

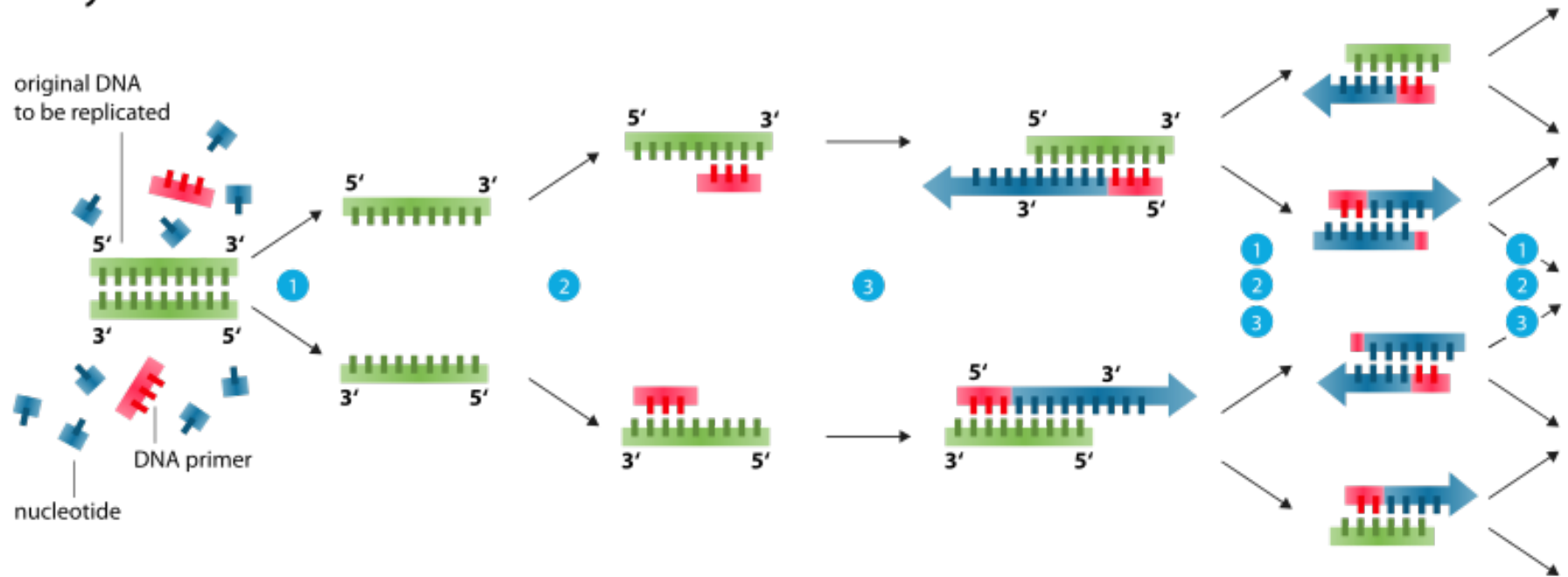
2 minutes 72 °C

only dNTP's

(Andy Vierstraete 1999)

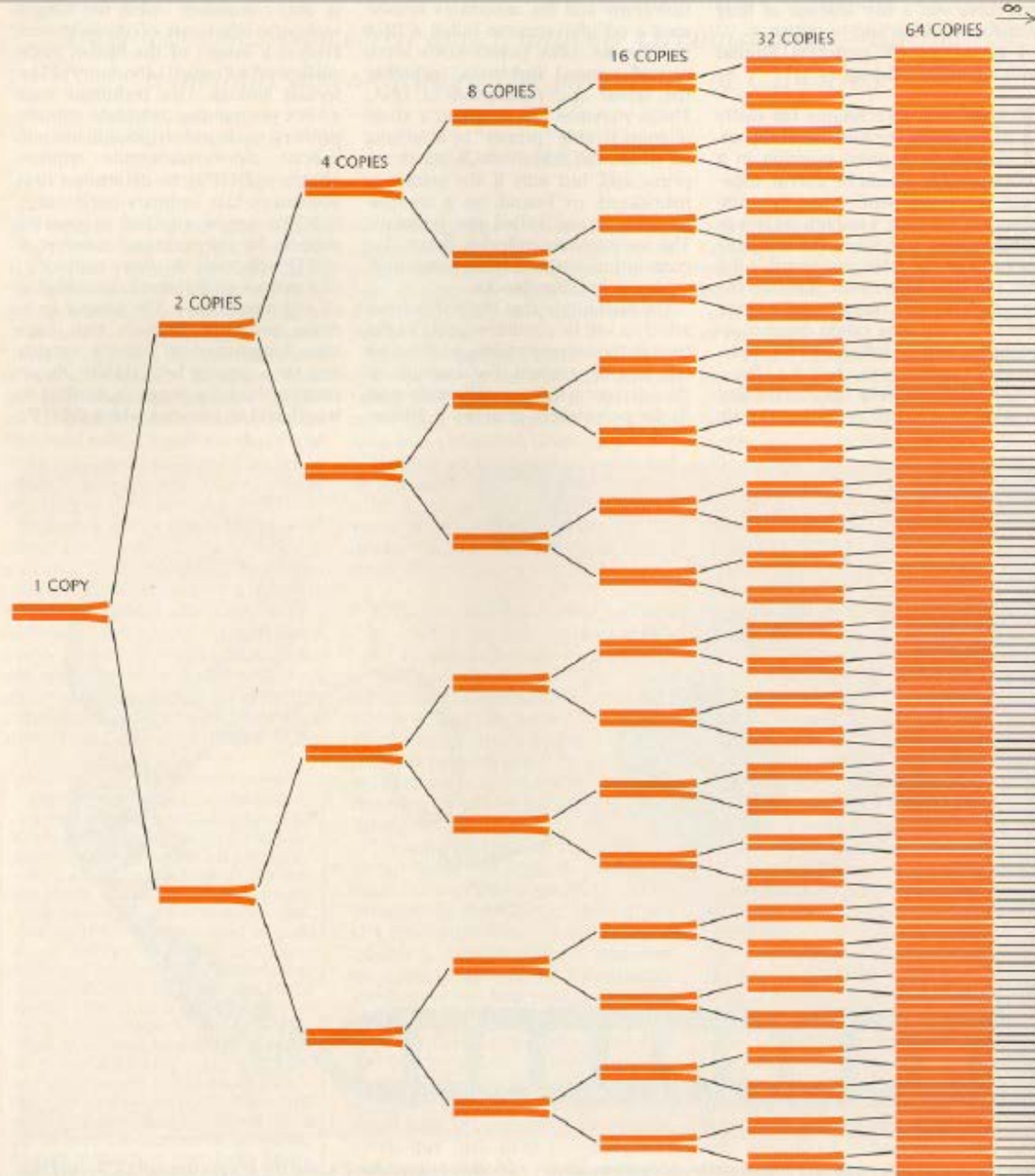
- Matematicamente, se começássemos com **uma única molécula**, teríamos no final
 - Entre 2^{30} e 2^{40} moléculas

Polymerase chain reaction - PCR



- 1 **Denaturation** at 94-96°C
- 2 **Annealing** at ~68°C
- 3 **Elongation** at ca. 72 °C

Fonte: wikipedia



POLYMERASE CHAIN REACTION is a simple technique for copying a piece of DNA in the laboratory with readily available reagents. Because the number of copies increases exponentially, more than 100 billion can be made in only a few hours.

Neste site vc pode ver animações interativas de PCR

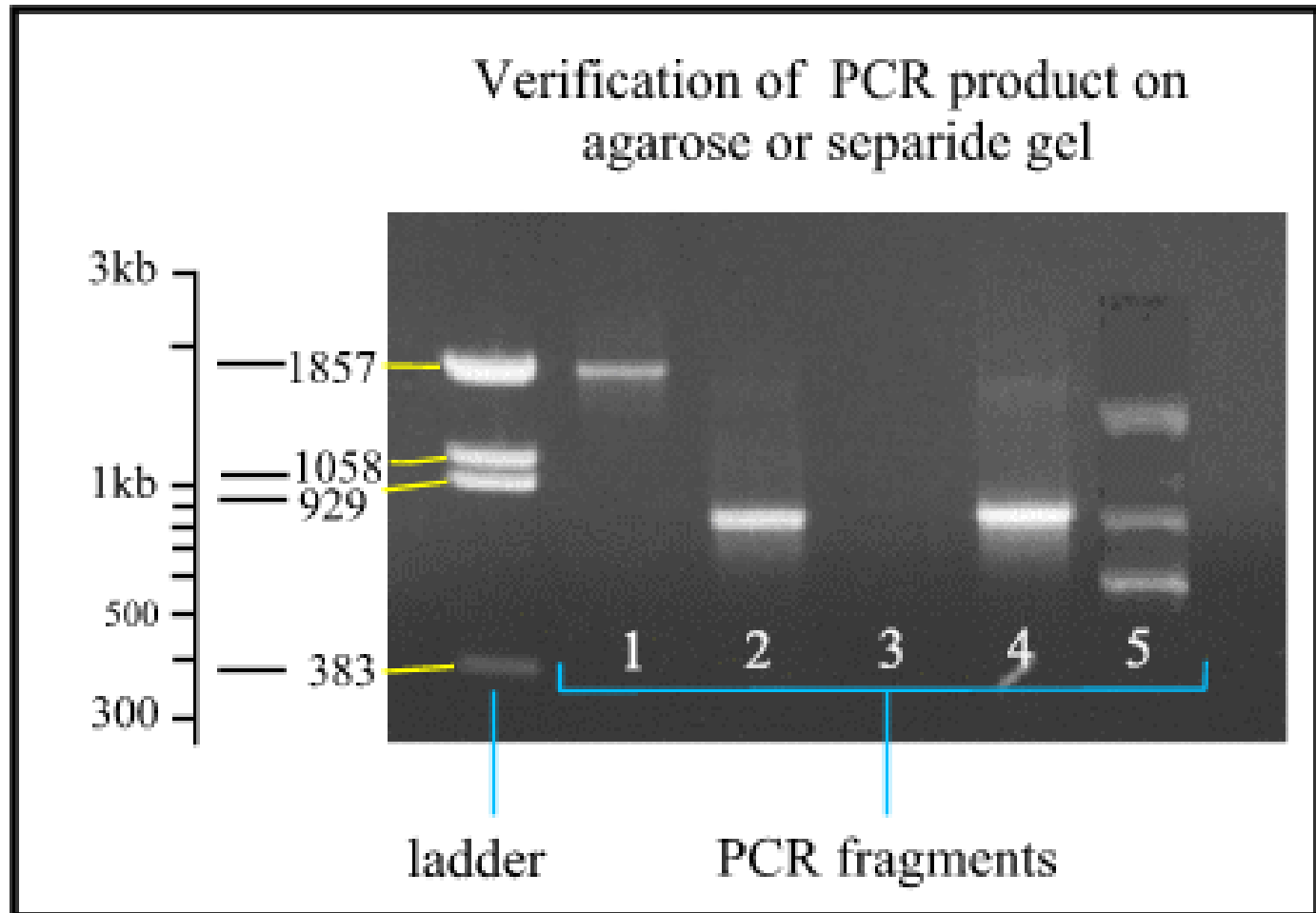
<https://dnalc.cshl.edu/resources/animations/pcr.html>

PCR se faz por máquinas



termocicladores

Nem sempre PCR dá certo



Aplicações

- Aumentar a quantidade de DNA para sequenciamento
- Diagnóstico de doenças genéticas
- Diagnóstico de doenças infecciosas
 - COVID-19
- Identificação de assinaturas genéticas, como em testes de paternidade
- Filogenia de espécies por pequenos trechos de DNA

Sequenciamento de DNA

- PCR aumenta (amplifica) a **quantidade** de um certo fragmento de DNA
- Como **saber a sequência** desse fragmento?

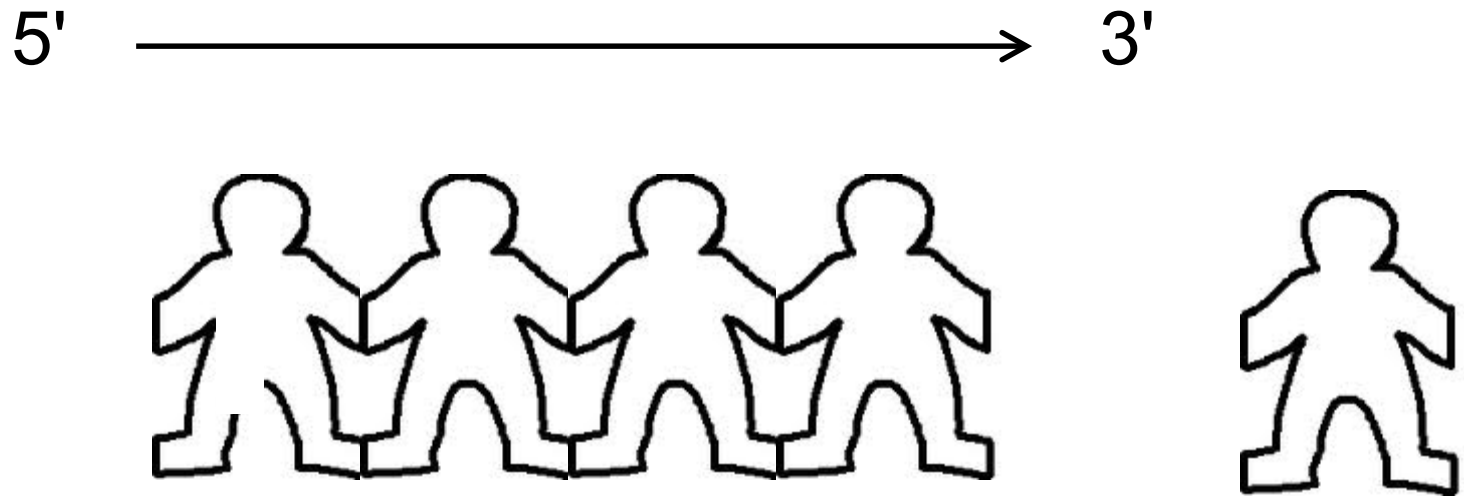
DNA é microscópico

- como saber a composição de DNA?
- não existe microscópio suficientemente poderoso que permita simplesmente “ler” a molécula, percorrendo-a de uma ponta à outra
- métodos **indiretos** são necessários
- todos eles tem que **fragmentar** o DNA para obter resultados (caso já não esteja fragmentado)

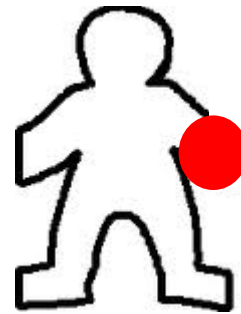
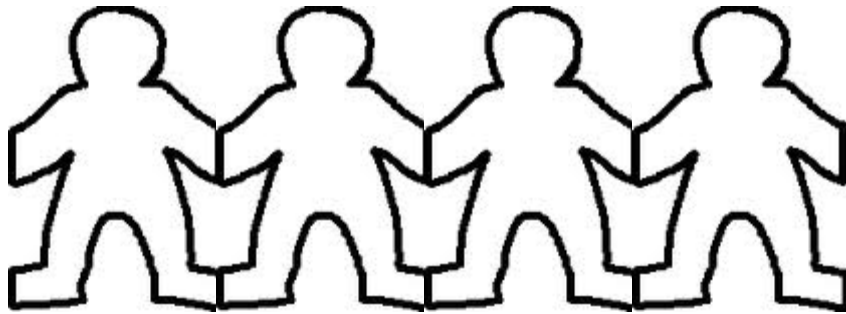
Sequenciamento de DNA

- método **Sanger** (década de 1970)
- está baseado no processo de **replicação**
- Fred Sanger
 - Inglês, 1918 – 2013
 - ganhou 2 vezes o prêmio Nobel de Química
 - uma delas pela tecnologia de sequenciamento

Replicação de DNA

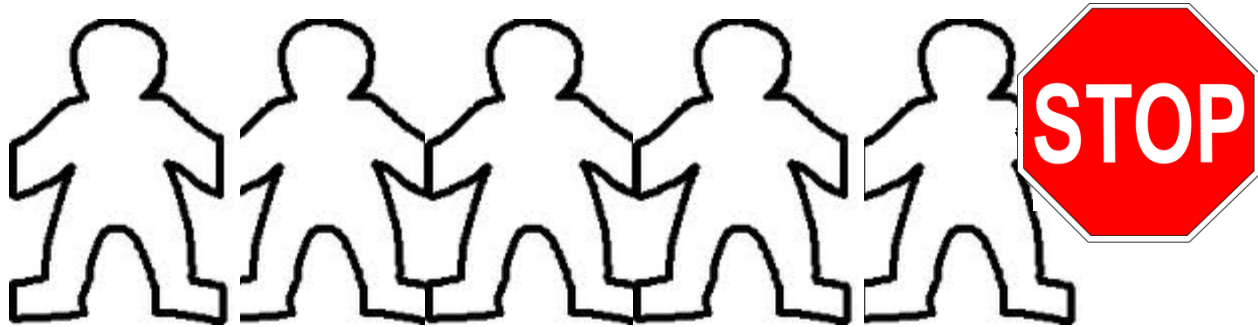


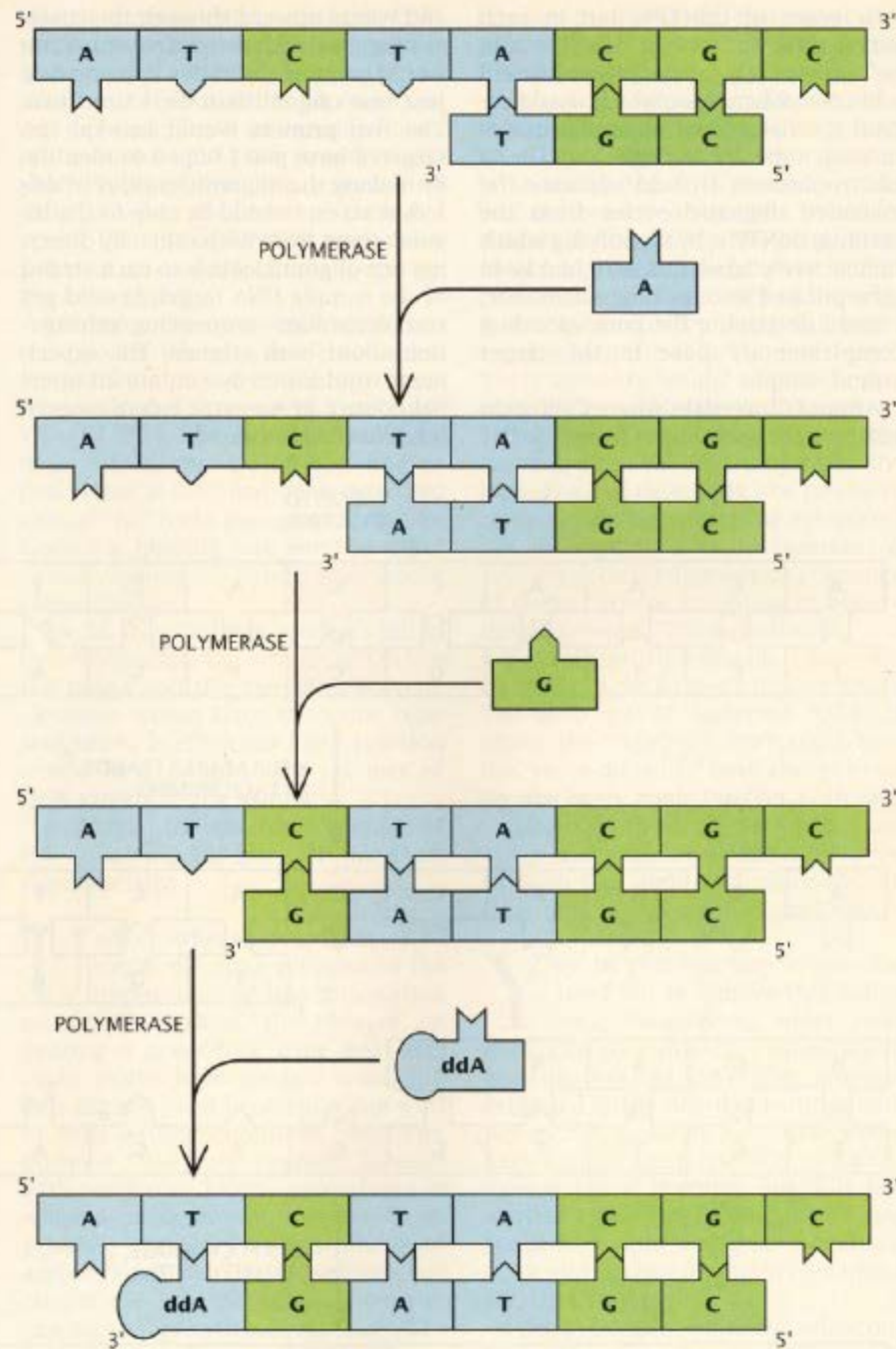
Se um dos nucleotídeos for “defeituoso”...



defeito neste caso significa impossibilidade de prolongar a espinha dorsal
acúcar-fosfato por causa de uma “irregularidade” no fosfato

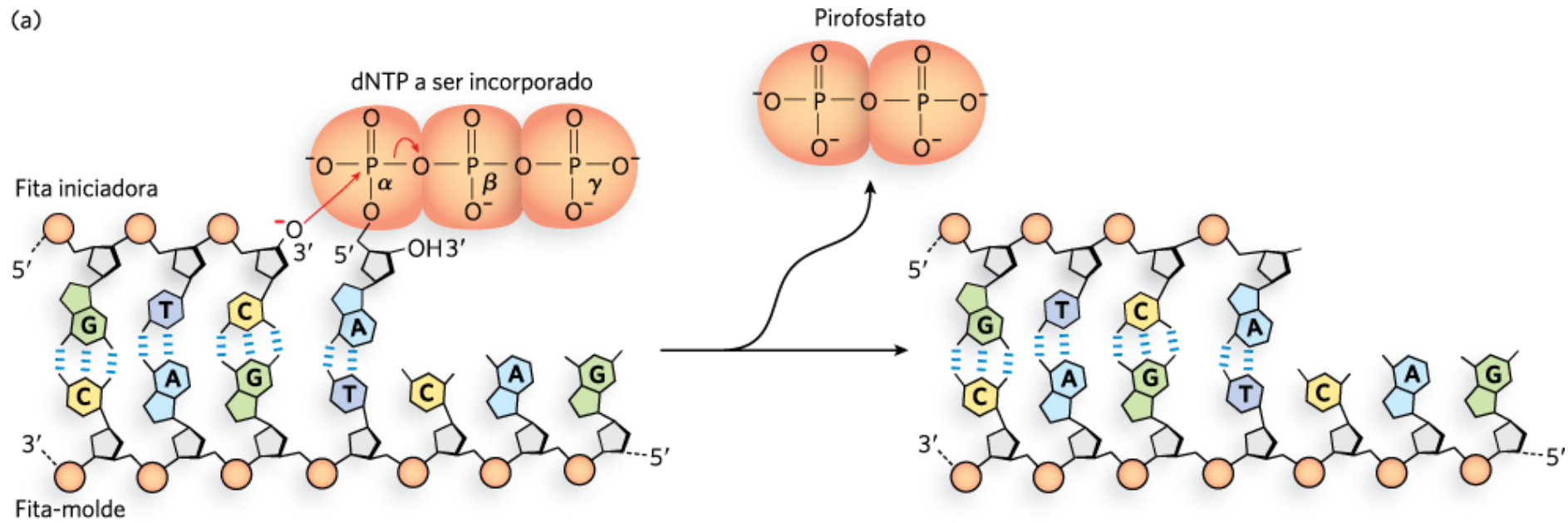
A replicação pára



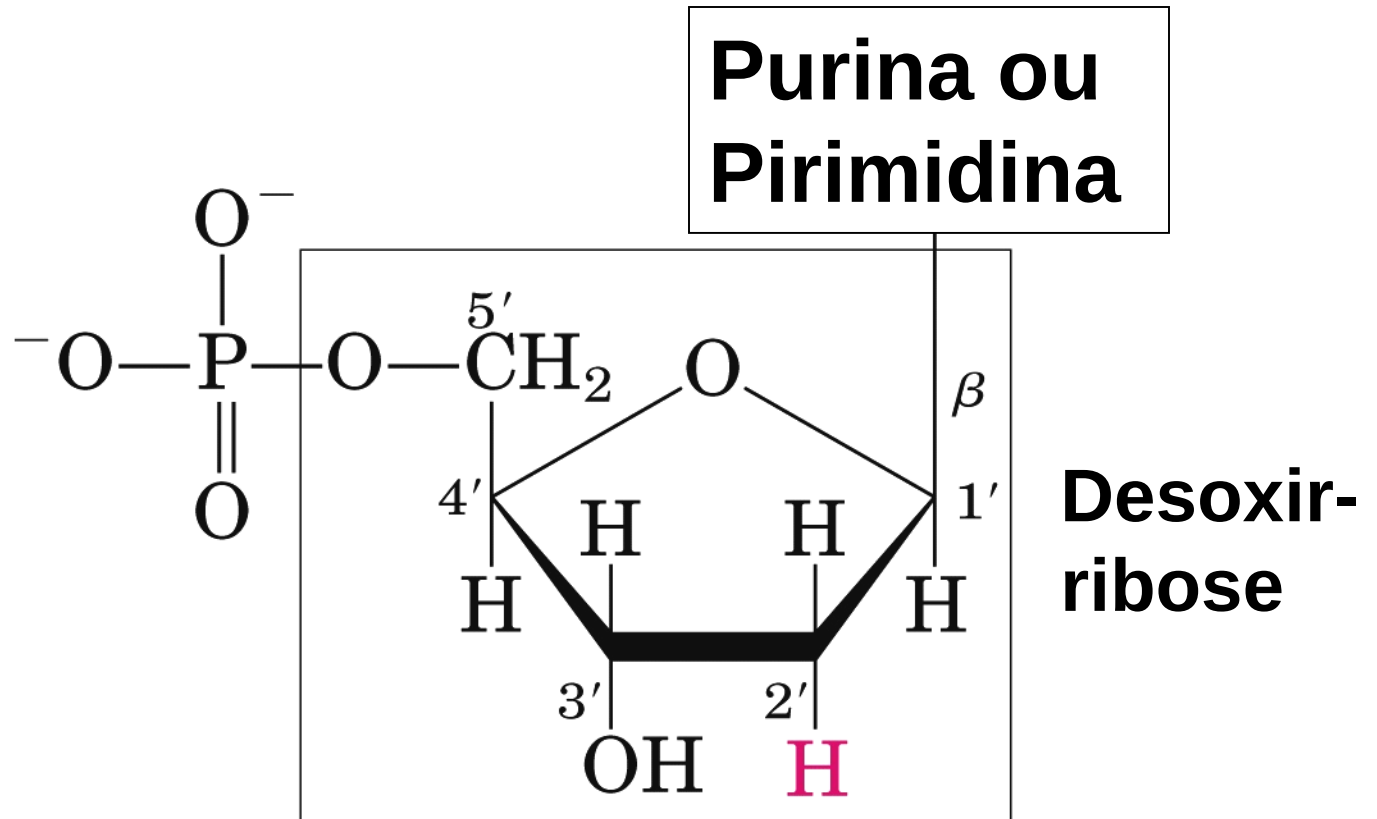


Reação da DNA Polimerase com dNTPs → síntese de DNA

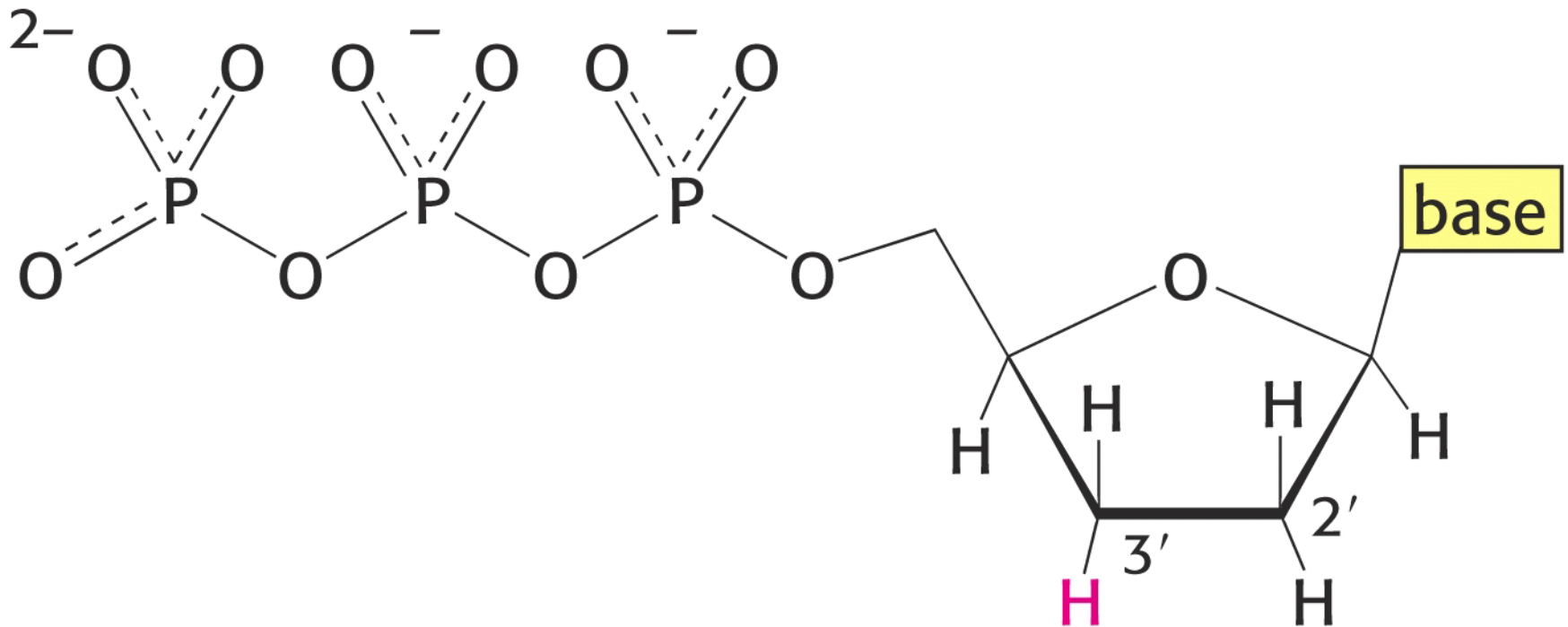
(a)



Fosfato



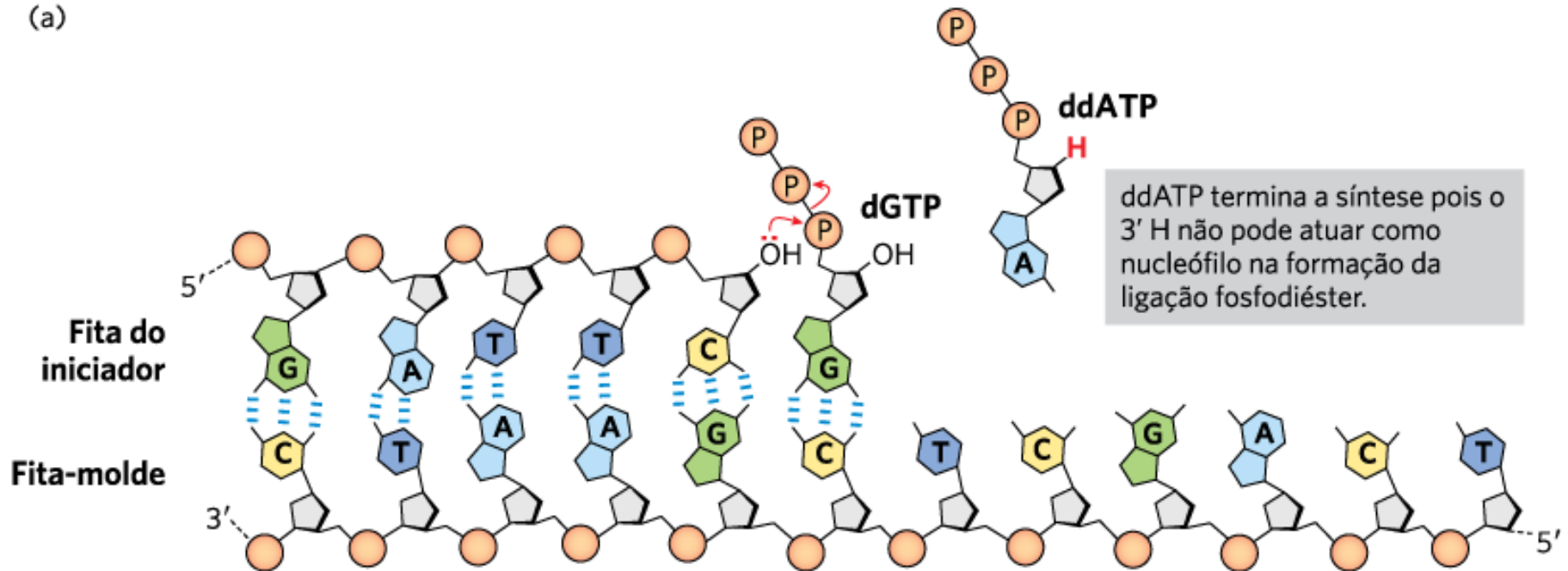
Desoxirribonucleotídeo



**2', 3' dideoxirribonucleotídeo
trifosfato (ddNTP)**

Reação da DNA Polimerase com dNTPs + ddNTPs → interrupção da síntese de DNA

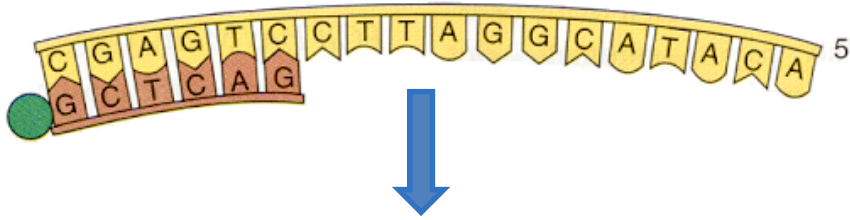
(a)



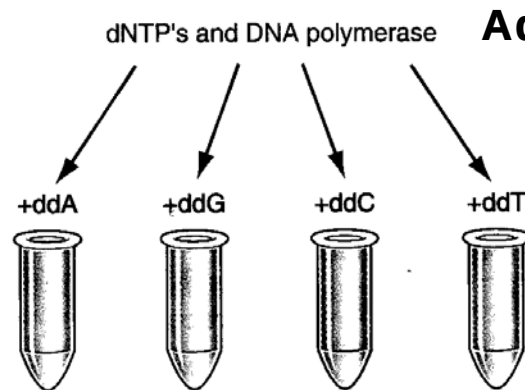
Truques

- Existem ddNTPs específicos para cada base
 - ddA
 - ddC
 - ddG
 - ddT
- É possível fazer com que esses ddNTPs sejam **fluorescentes** (para que possam ser detectados)

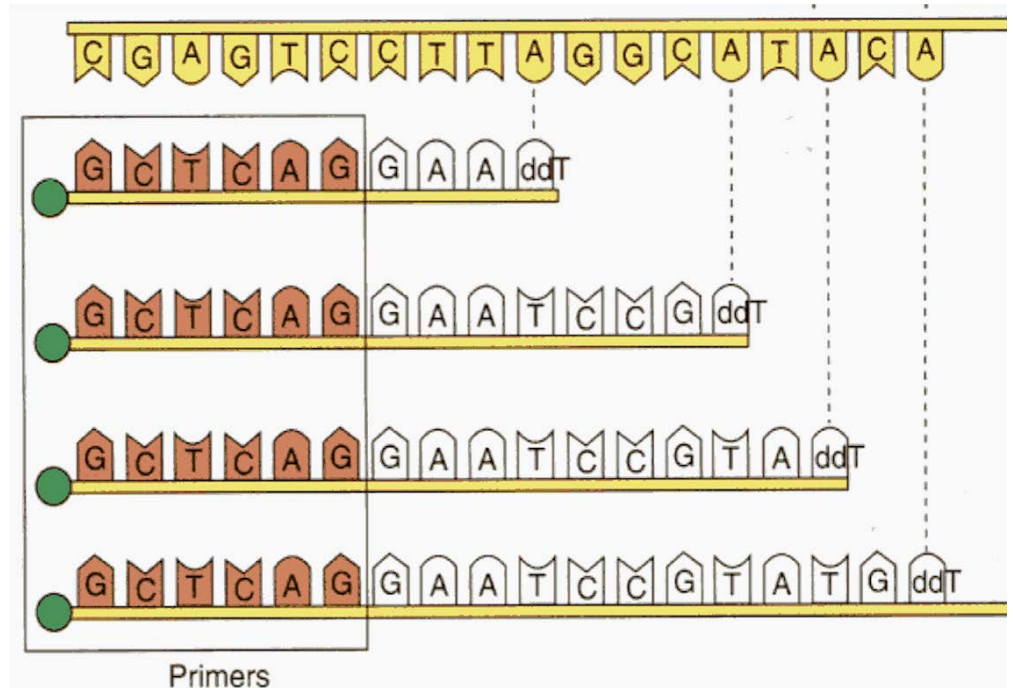
Método de Sanger



Desnaturação da dupla fita
Anelamento do “primer radioativo”



Terminação da síntese: adição dos ddNTPs



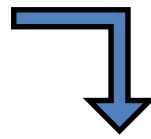
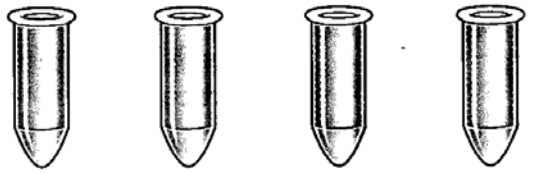
dNTP's and DNA polymerase

+ddA

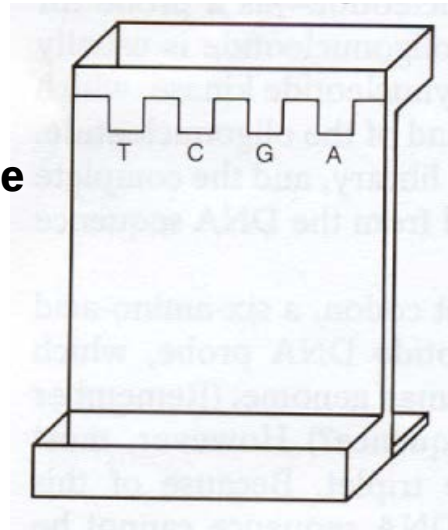
+ddG

+ddC

+ddT

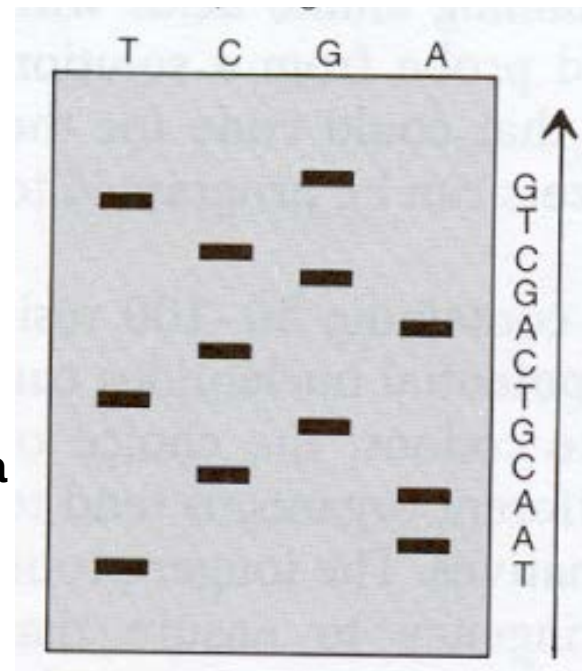
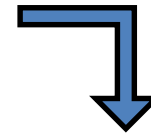


**Eletroforese em gel de
poliacrilamida
desnaturante**



Autorradiografia

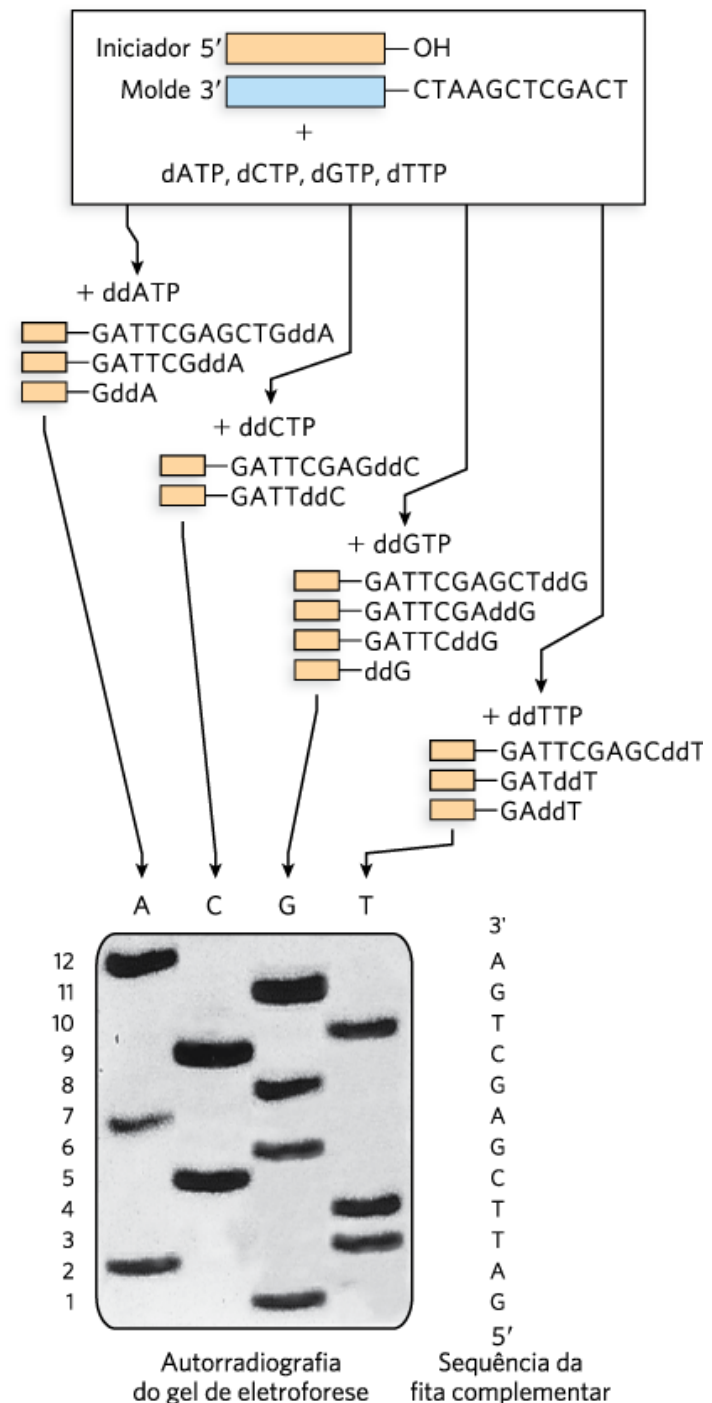
Eletroforese em gel de
poliacrilamida: separação de
fragmentos de DNA diferindo
por 1 nucleotídeo no tamanho



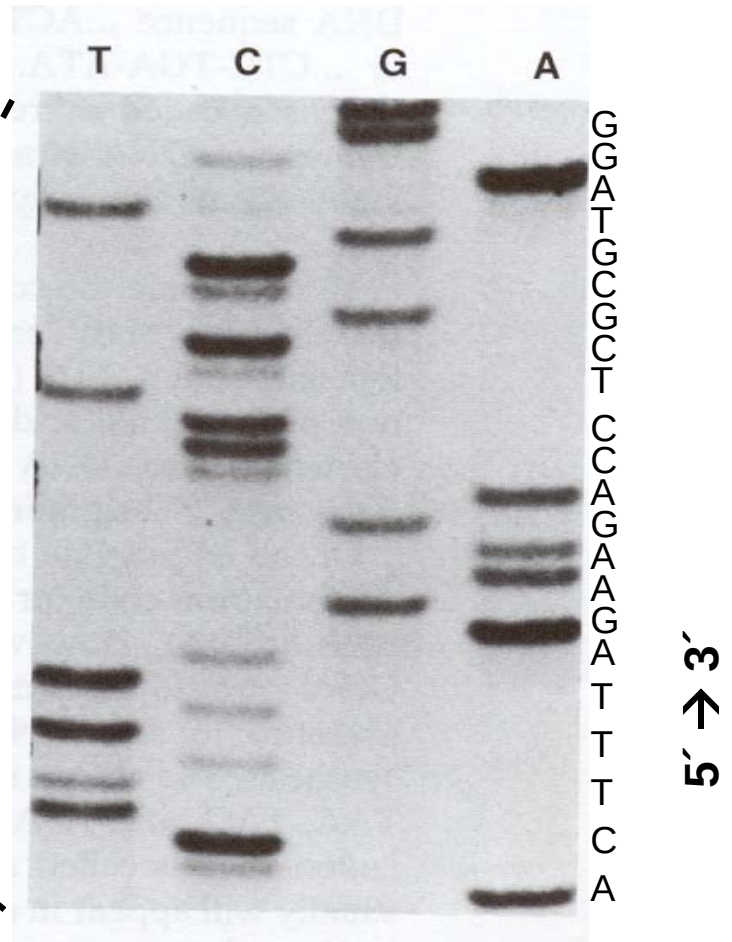
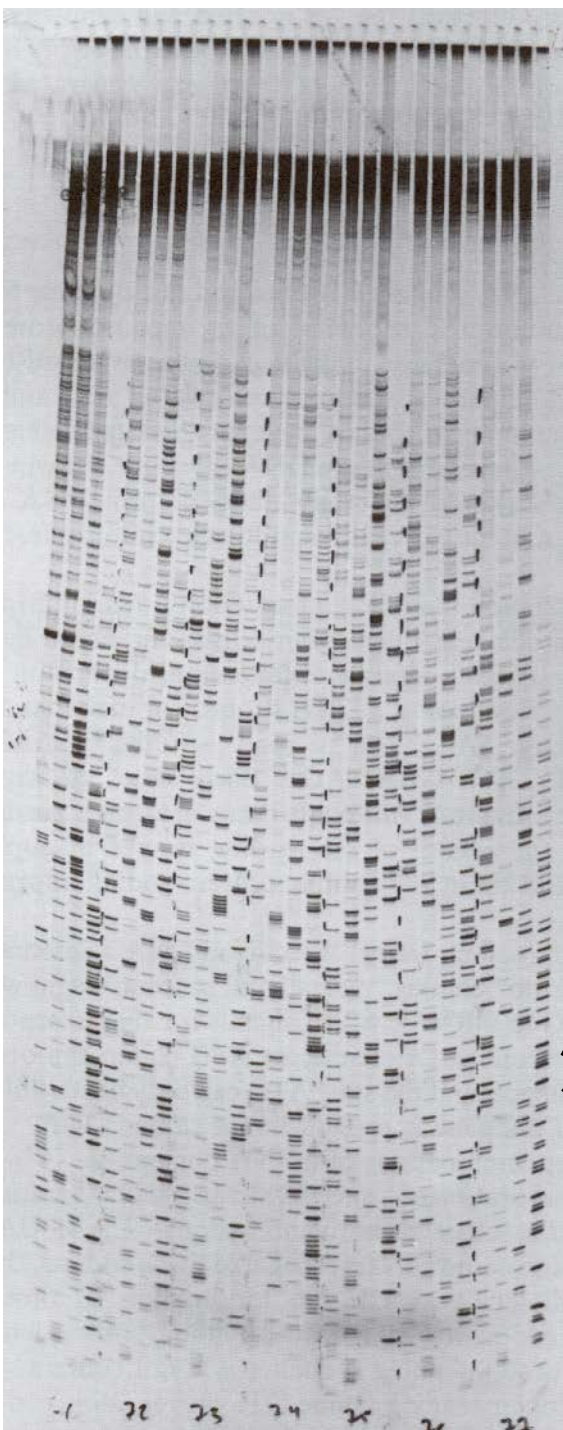
**Sequência complementar
ao DNA molde**

Método de Sanger:

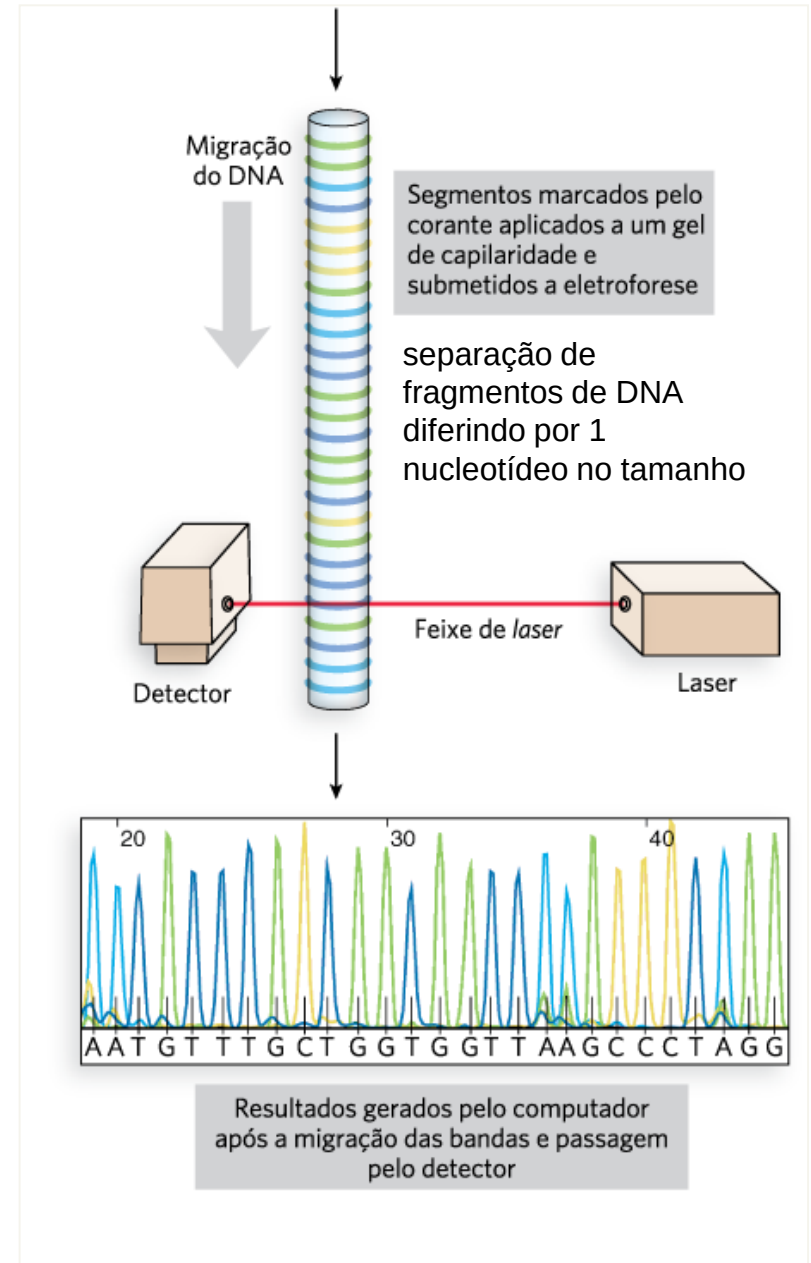
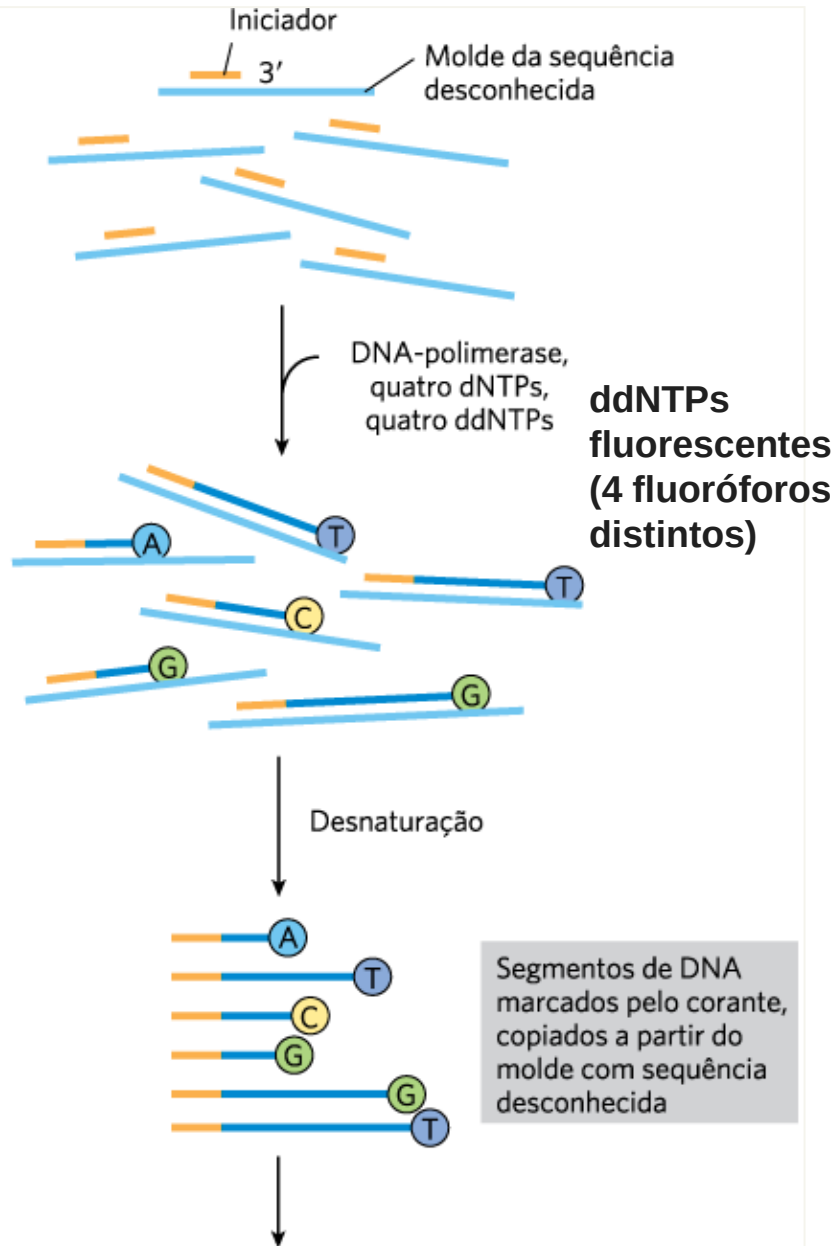
terminação controlada da síntese de DNA com dideoxirribonucleotídeos



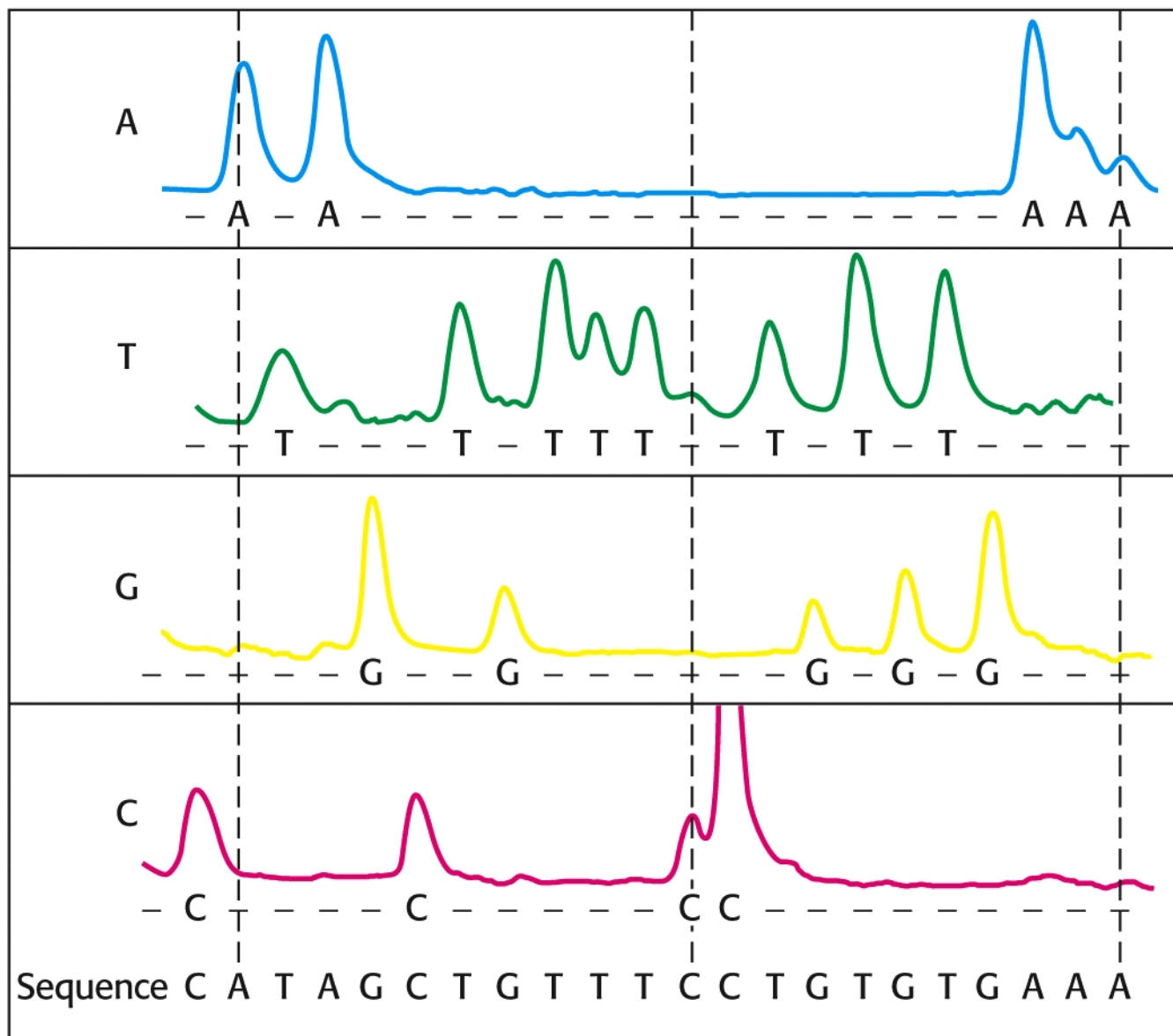
Autorradiograma de um gel de sequenciamento de DNA



Automação do Sequenciamento pelo Método de Sanger

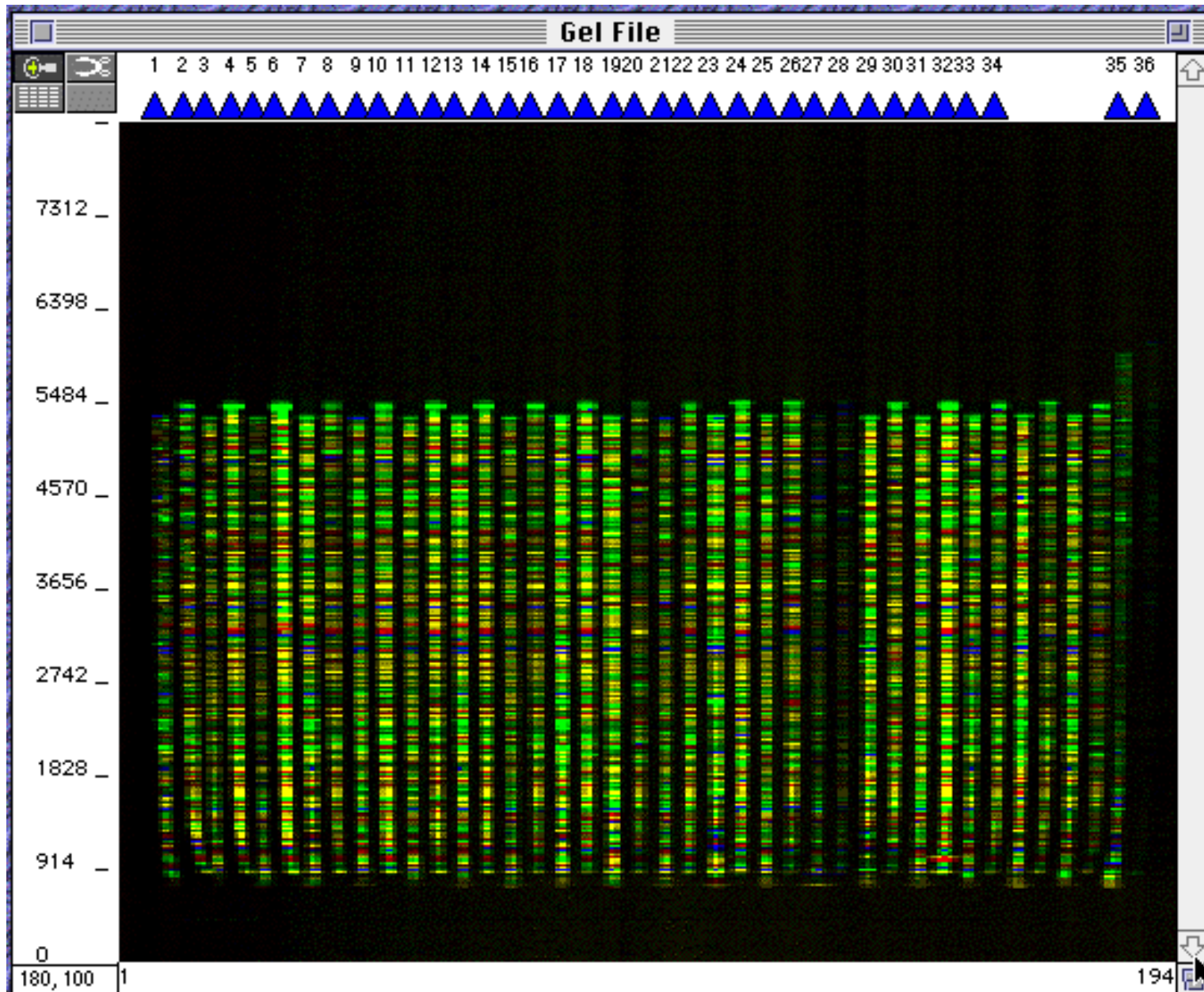


Intensidade de fluorescência

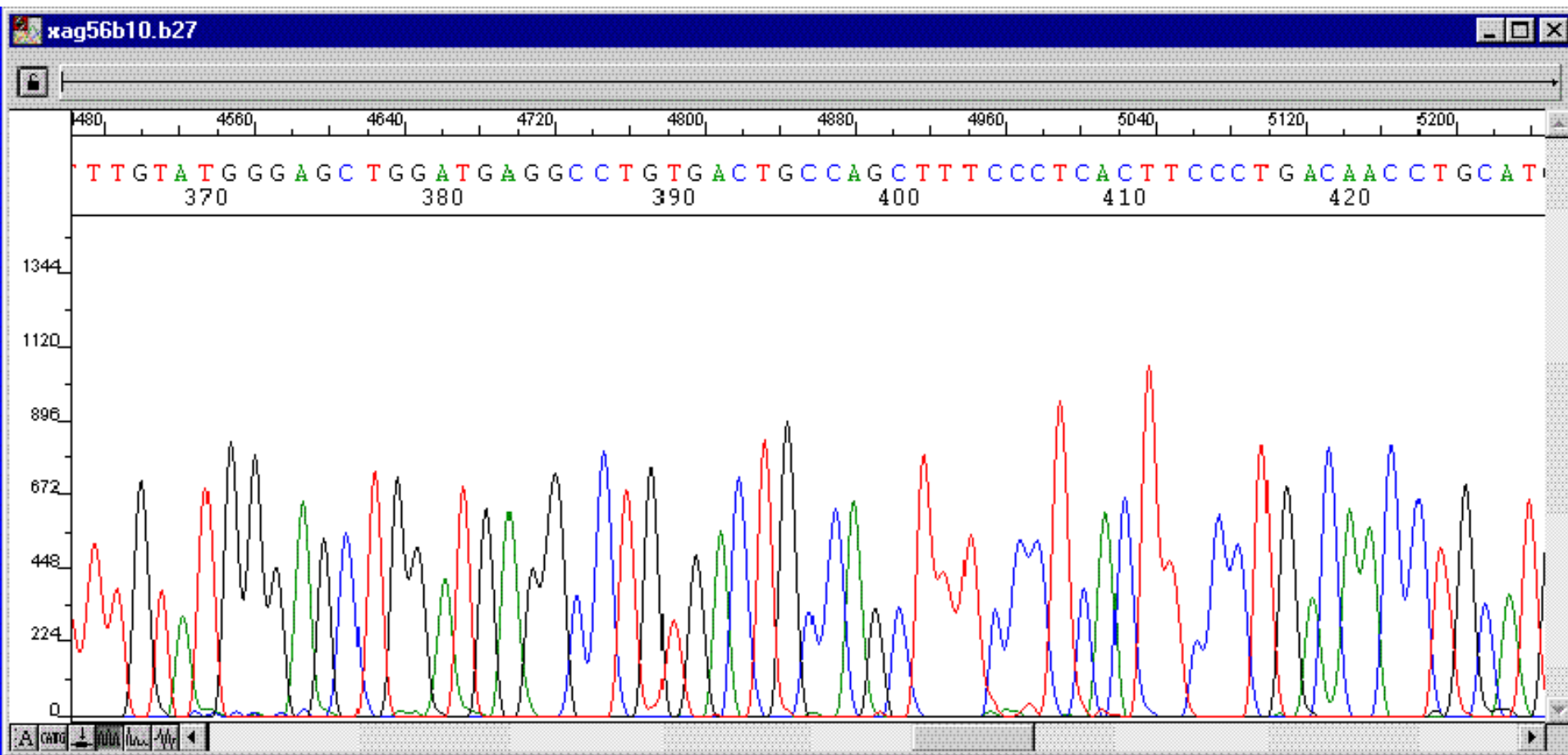


Tamanho em nucleotídeos (bases)

Imagem da detecção por fluorescência no sequenciamento automatizado: cada amostra em um capilar



Cromatograma do sequenciamento automatizado pelo método de Sanger



Tamanho médio das sequências geradas → 700 – 1000 pb

Tecnologias atuais de sequenciamento

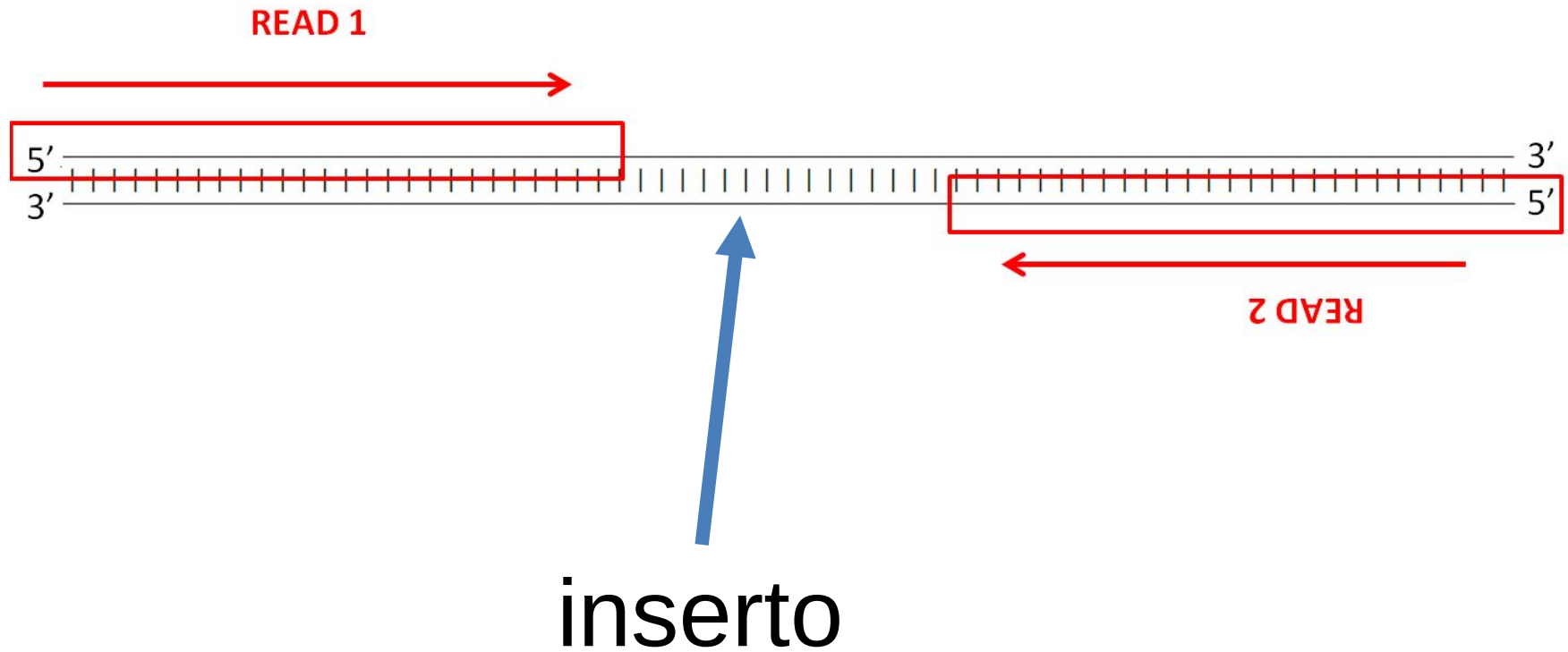
- Illumina
 - sequenciamento por síntese
 - líder de mercado
- Pacific Biosciences (pacbio)
- Oxford Nanopore

Illumina

máquina	tamanho de read (bp)	número de reads por corrida (milhões)
iSeq	2 x 150	4
miniSeq	2 x 150	25
miSeq	2 x 300	25
nextSeq	2 x 150	400

acurácia: 99,9%

Paired-end reads



PacBio

- leituras podem chegar a 100 kb
- acurácia: 87%

Oxford Nanopore

- tamanho do read depende de escolhas do usuário
- de alguns kb até 500 kb
- acurácia: 92 a 97%



1997

The ABI Prism 3700 /3730

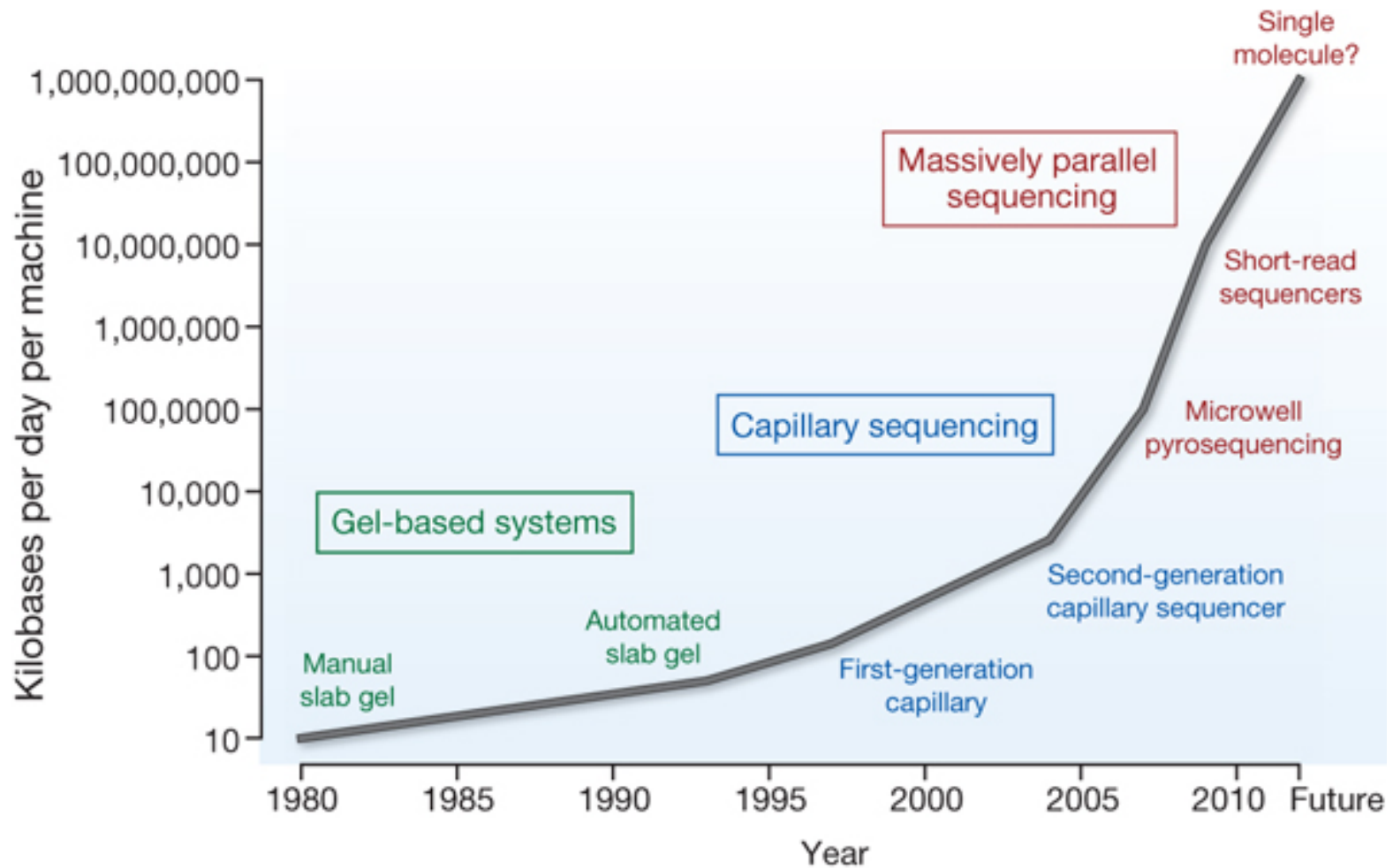


2017

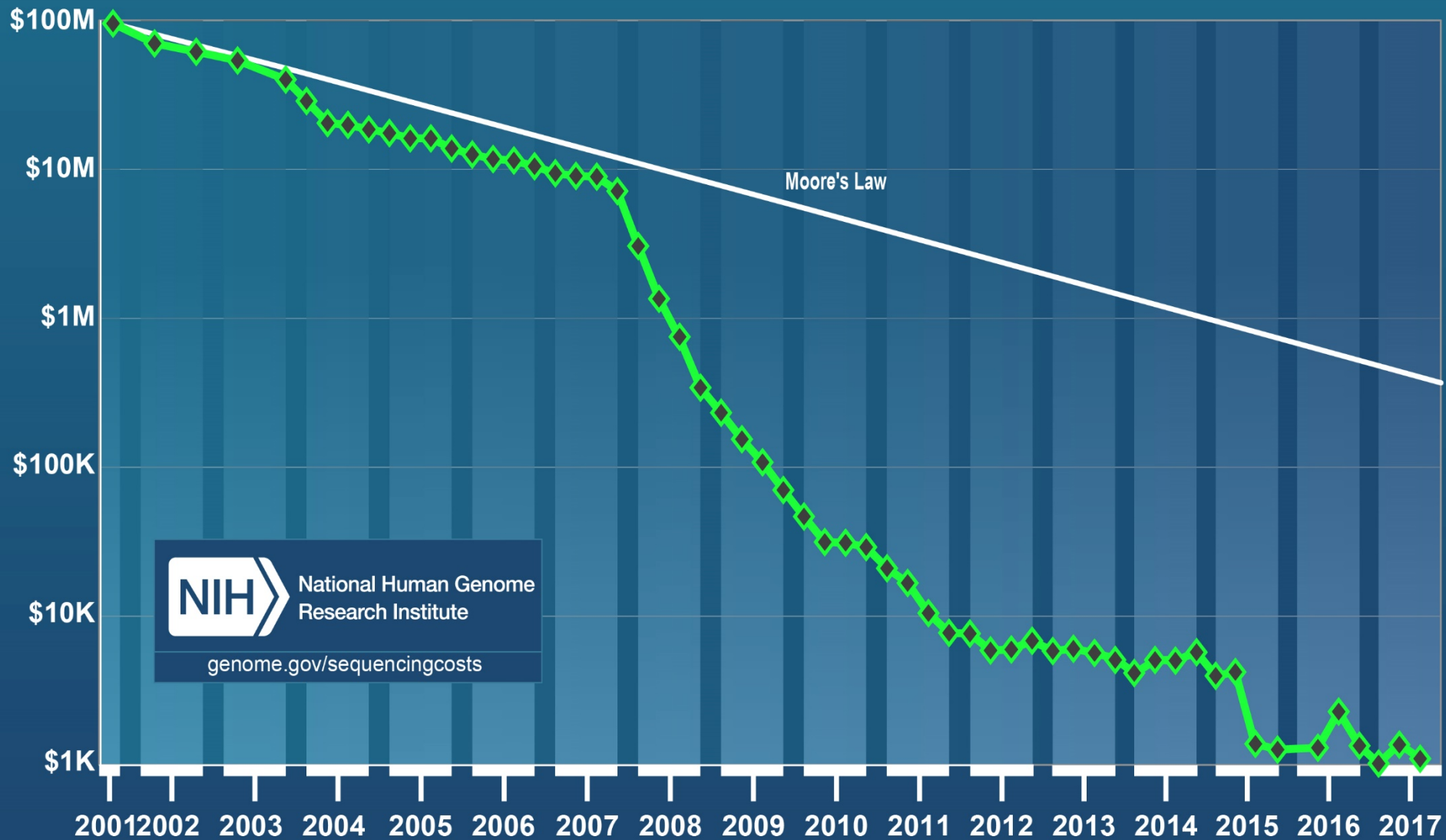
oxford nanopore minION



Evolução da tecnologia de sequenciamento de DNA



Cost per Genome



Como obter genomas inteiros?

- Por montagem

Como os genomas são sequenciados

DNA genômico



Fragmentação



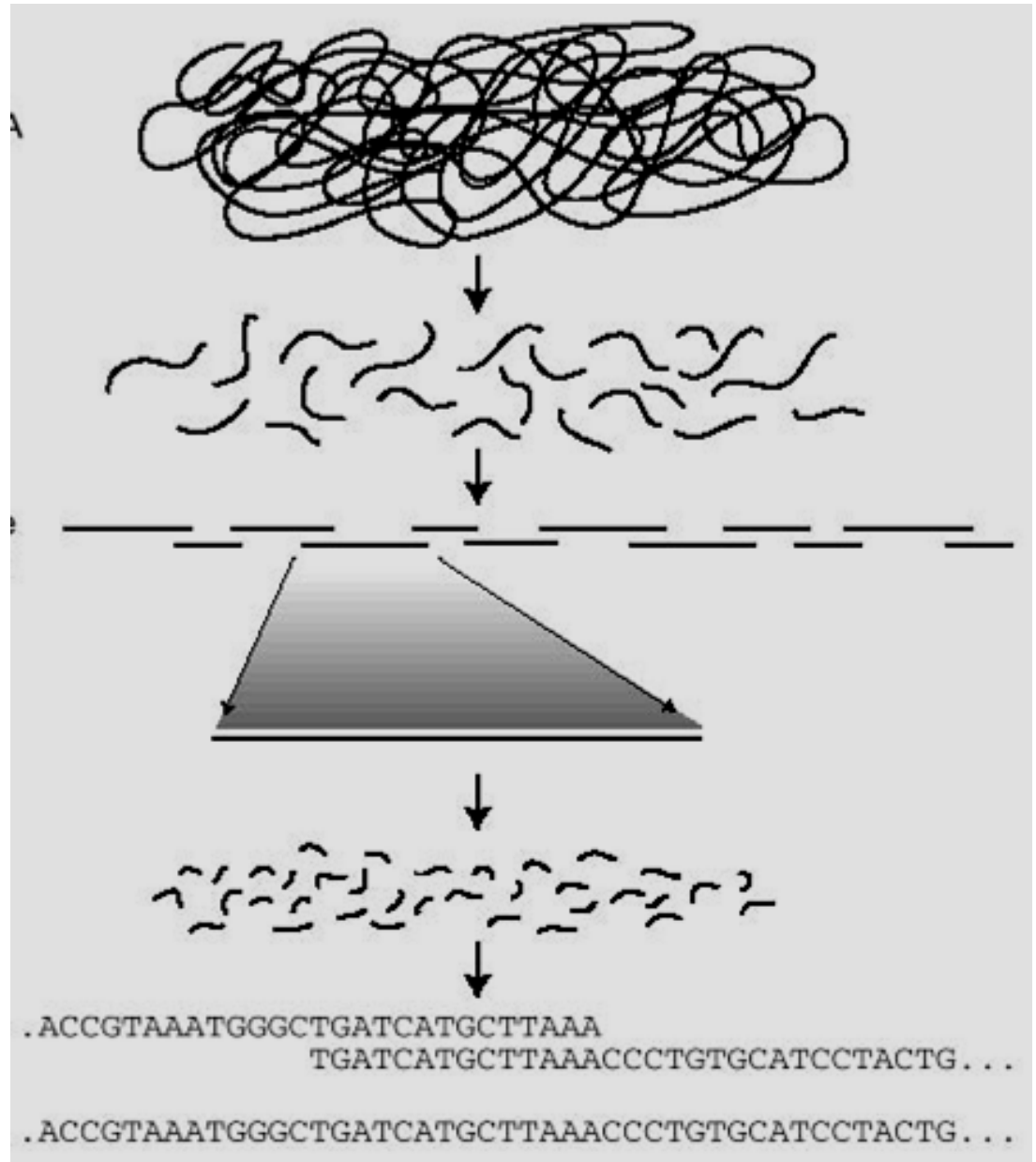
**Amplificação
dos fragmentos**



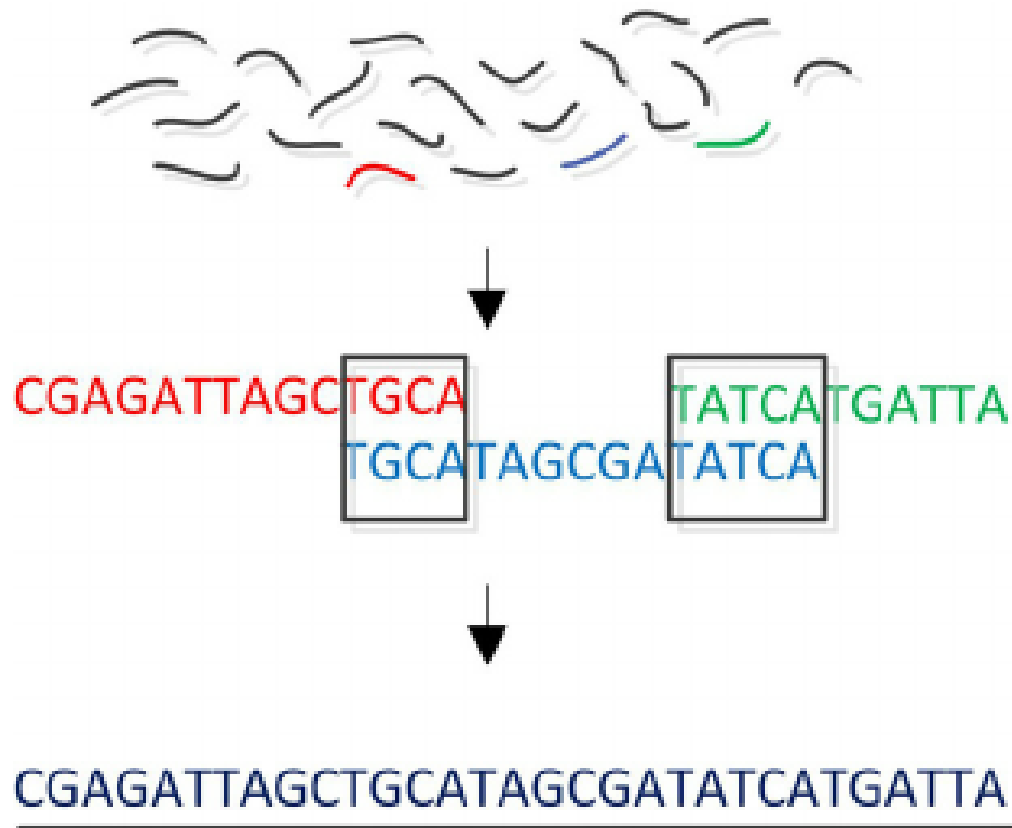
Sequenciamento



Montagem (*in silico*)



Montagem



Pergunta

- se sequenciamento pode ter erros, como atingir a sequência correta?
 - dica: a montagem ajuda

Aplicações de sequenciamento

- **Medicina**

- Genoma humano

- Primeiro sequenciamento

- ano 2000, a um custo de 3 bilhões de dólares

- Hoje

- Centenas de milhares de genomas foram sequenciados

- Custo abaixo de mil dólares

- Transcritoma humano

- Diagnóstico de doenças infecciosas

- Medicina personalizada

- Agricultura, pecuária, veterinária, ecologia, criminalística, antropologia, arqueologia, etc

Testes genéticos baseados em PCR e/ou sequenciamento são cada vez mais comuns

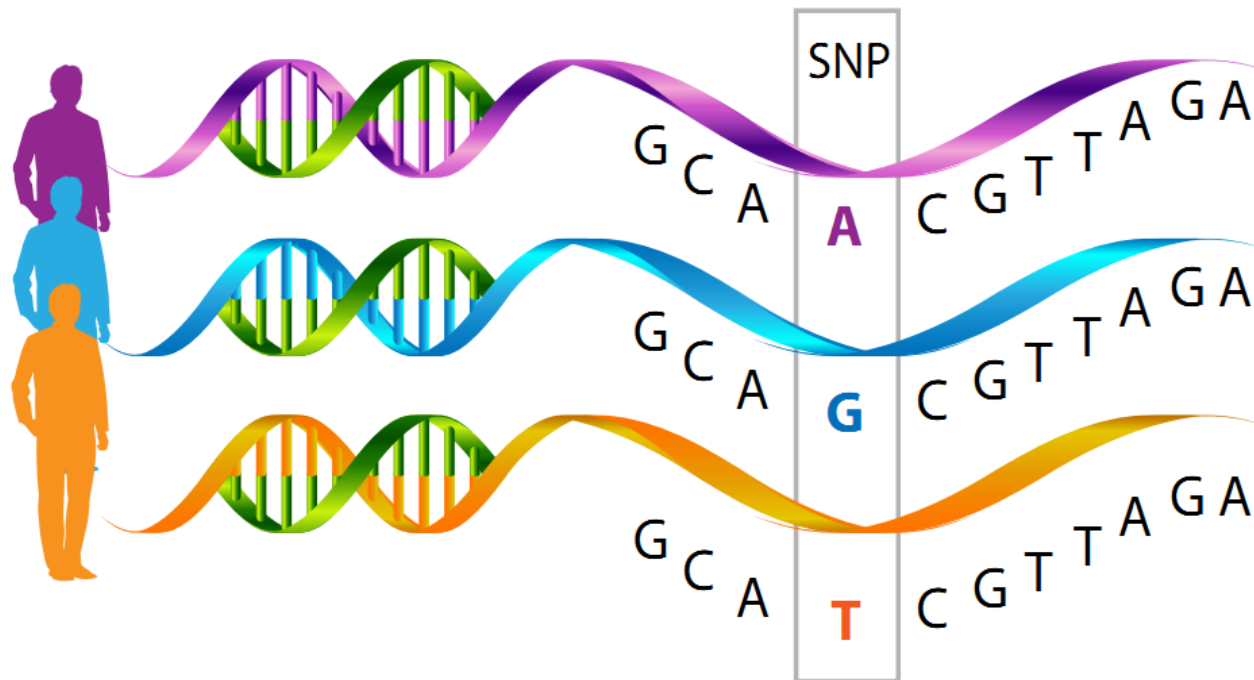
- genes associados a doenças
- intolerância à lactose
- exoma
- etc

Ministério da saúde do Reino Unido
planeja sequenciar 5 milhões de
genomas humanos nos próximos 5
anos

3 de outubro de 2018

SNP

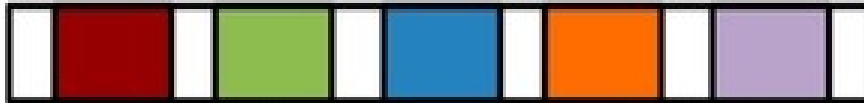
- Single Nucleotide Polymorphism
- polimorfismo de um único nucleotídeo



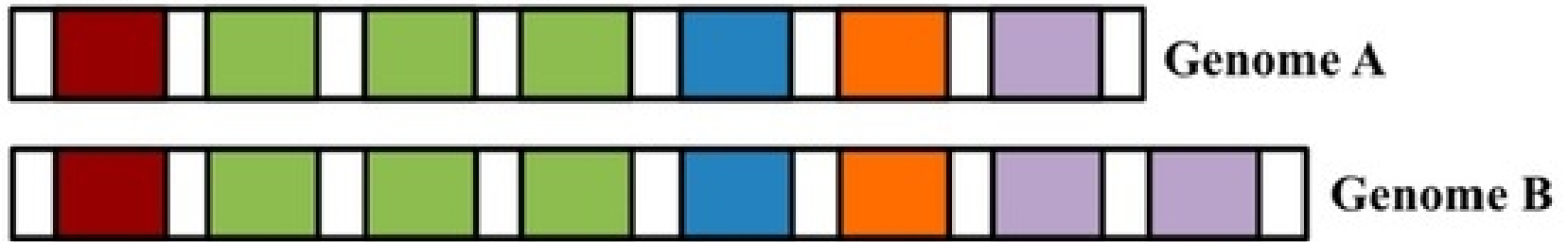
Outras variações

- **indel** = INsertion / DELetion
 - trechos pequenos de DNA inseridos ou removidos num genoma quando comparado com outro
- **número variável de repetições**
 - CNV = copy number variation

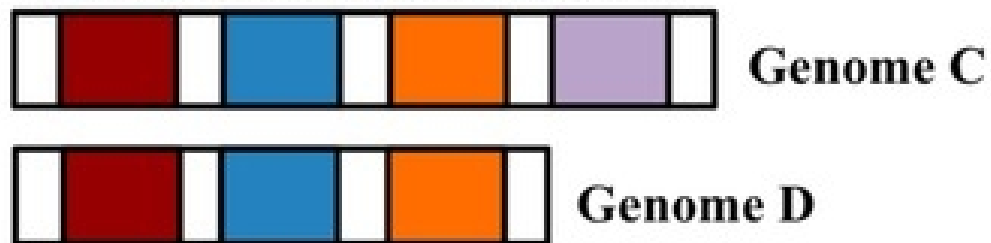
a. Normal reference



b. Copy number variation



c. Presence absence variation



Haplótipo

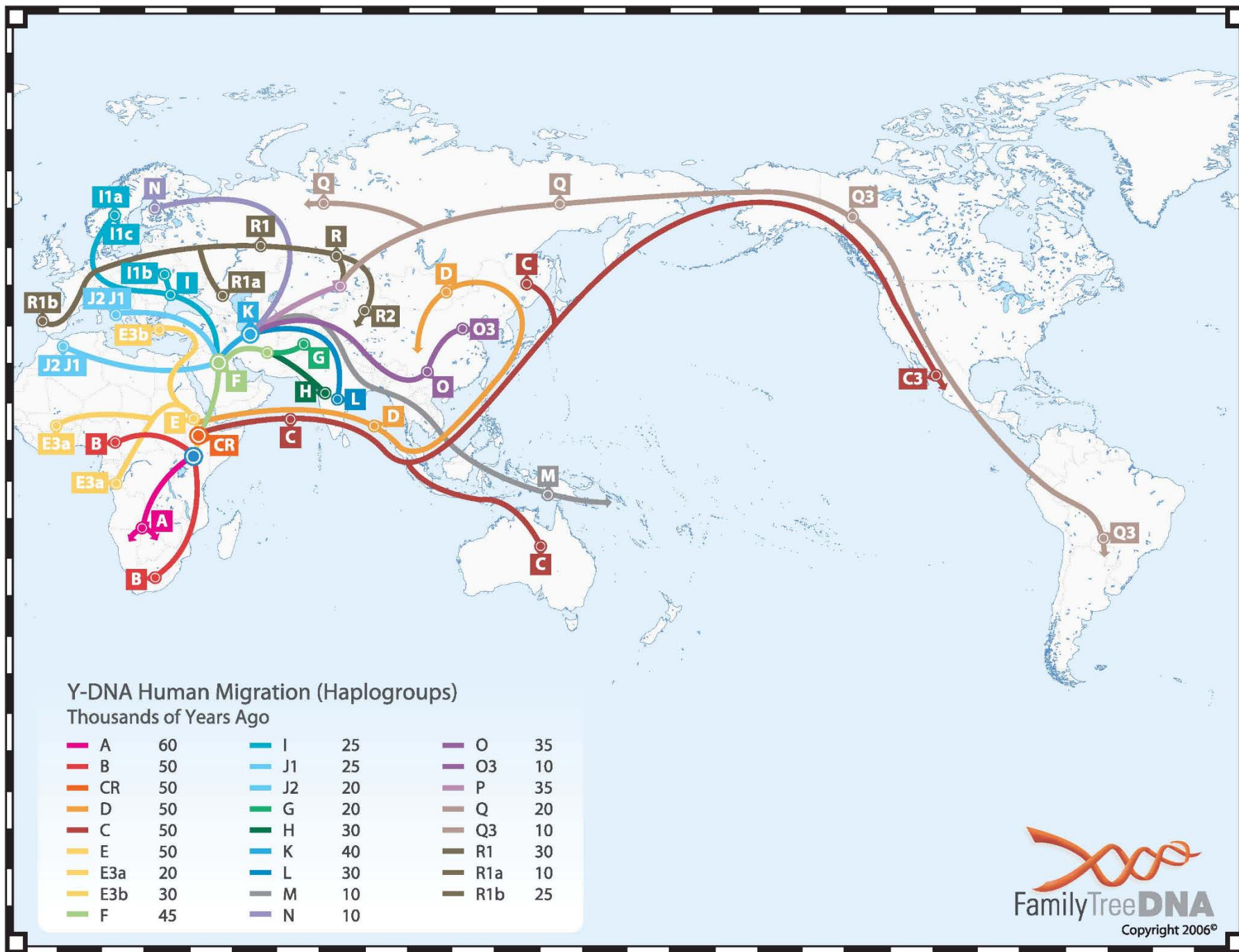
- São blocos de variações que caracterizam populações
 - as variações são **herdadas** “em bloco”

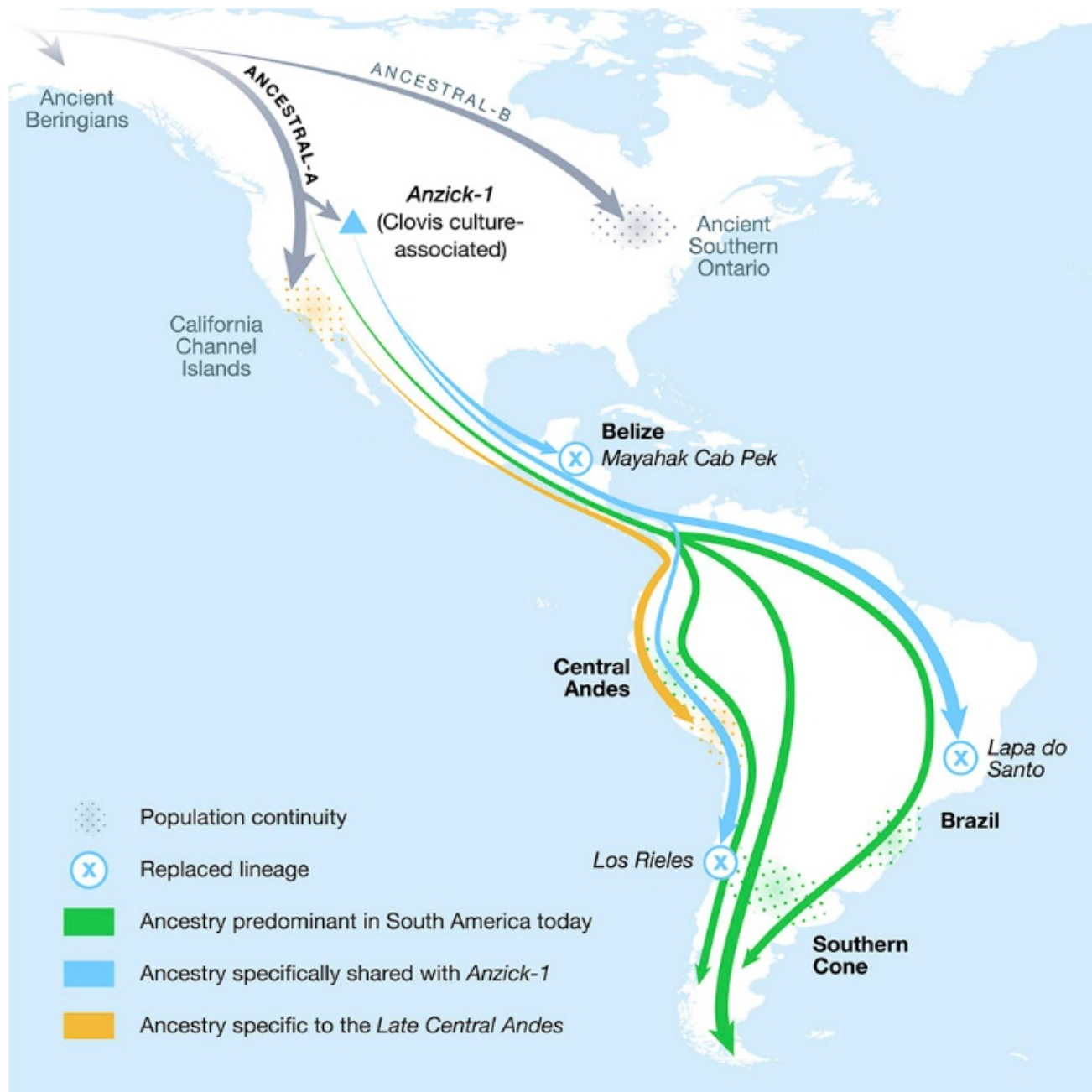
Sequenciamento e comparação de sequências genômicas de indivíduos



Identificação de **SNP** (polimorfismo de único nucleotídeo) em genomas. Grupos de SNPs marcadores são compilados em um haplótipo e podem ser utilizados para identificação de indivíduos pelo sequenciamento de regiões definidas de seus genomas.

O sequenciamento de genomas ou trechos de genomas de diferentes indivíduos de diferentes grupos étnicos em diferentes regiões do mundo tem permitido a reconstrução das grandes migrações do ser humano pelo planeta, remontando há milhares de anos atrás. Nos próximos 2 slides apresento duas dessas reconstruções





Medicina personalizada (ou medicina de precisão)

- **tamoxifen** é uma droga que costumava ser prescrita para mulheres com câncer de mama
- observação empírica: **não funcionava** para **65% das mulheres**
- motivo: mulheres com uma **mutação no gene CYP2D6** não conseguem produzir **a proteína que processa a droga**
- hoje em dia: verificar por PCR + sequenciamento se a paciente tem essa mutação (antes de receitar a droga)

PCR ou sequenciamento?

- Embora essa pergunta seja comum, ela está **mal formulada**
- PCR é amplificação
- para sabermos o que contém a molécula amplificada, temos que **sequenciá-la**
- Então na verdade a pergunta deveria ser:
 - PCR de um pedaço seguido de sequenciamento desse pedaço
 - ou
 - sequenciamento do genoma inteiro?

Resposta

- PCR depende de molécula alvo, previamente conhecida (por exemplo, um gene de interesse)
- PCR+sequenciamento da molécula alvo é muito **mais barato** do que sequenciamento do genoma inteiro
- Mas sequenciamento fornece **muito mais informação**