

QBQ 1354 – 2023

# Sequenciamento de DNA

## Sequenciamento de DNA

determinar a ordem das bases nitrogenadas adenina (A), guanina (G), citosina (C) e timina (T) da molécula de DNA

Obter a sequência completa de fragmentos de DNA:

- produtos de PCR
- clones de DNA em plasmídeos
- genomas completos

# Métodos de sequenciamento de DNA

- **Método de Sanger:** terminação controlada da síntese de DNA com didesoxirribonucleotídeos
- **Metodologias de sequenciamento de alto desempenho (*Next Generation Sequencing*, NGS)**



## The Nobel Prize in Chemistry 1980

Paul Berg, Walter Gilbert, Frederick Sanger

The Nobel Prize in Chemistry 1980

Nobel Prize Award Ceremony

Paul Berg

Walter Gilbert

Frederick Sanger



Paul Berg



Walter Gilbert



Frederick Sanger

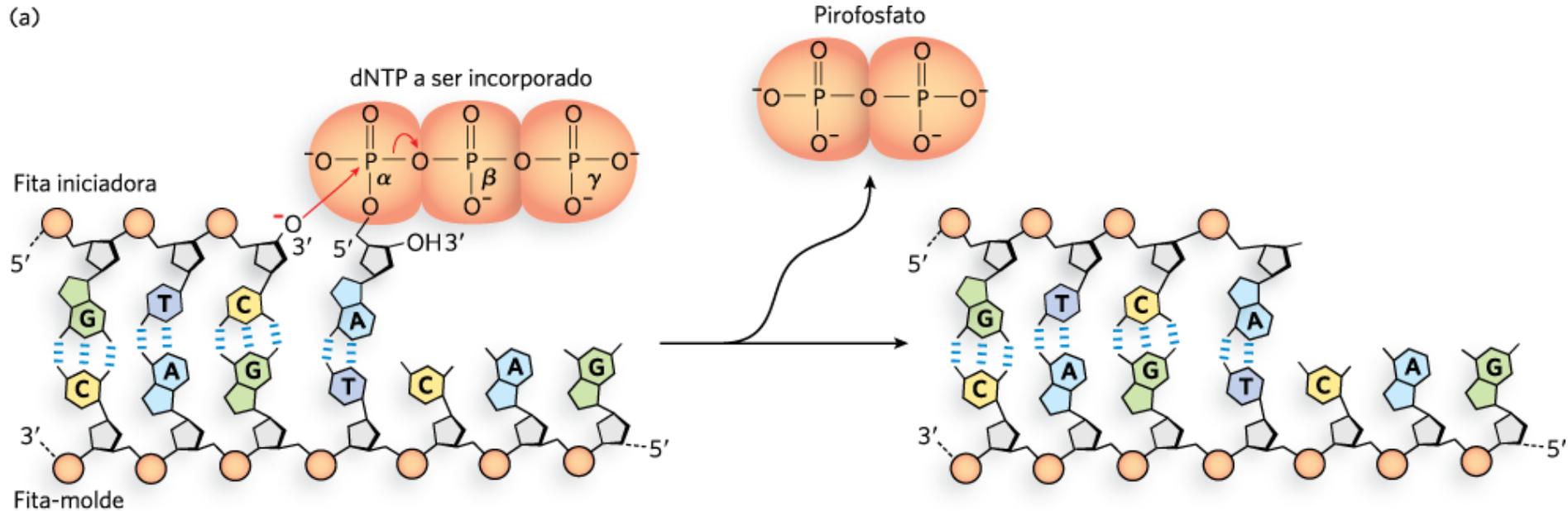
The Nobel Prize in Chemistry 1980 was divided, one half awarded to Paul Berg "for his fundamental studies of the biochemistry of nucleic acids, with particular regard to recombinant-DNA", the other half jointly to Walter Gilbert and Frederick Sanger "for their contributions concerning the determination of base sequences in nucleic acids".

# **Método de Sanger:** terminação controlada da síntese de DNA com didesoxirribonucleotídeos

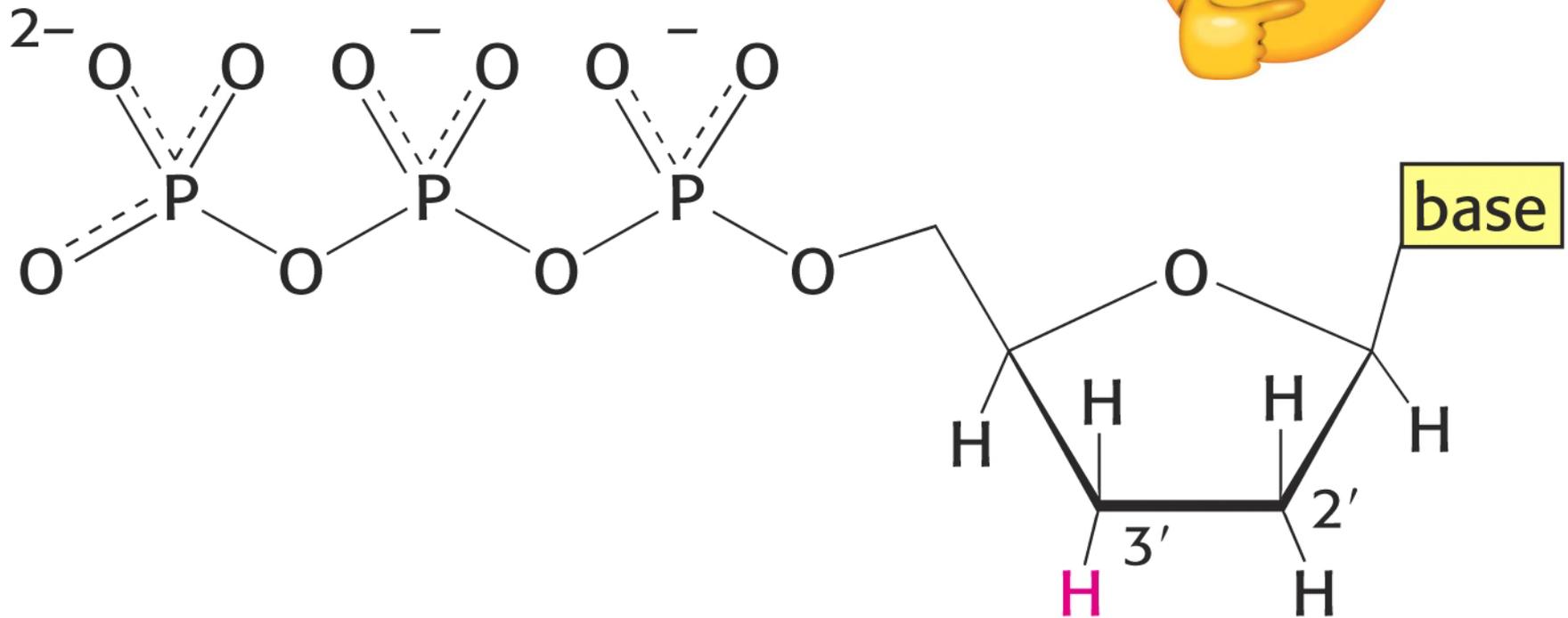
- Manual
- Automatizado

# Relembrando.....

Reação da **DNA Polimerase** com dNTPs →  
síntese de DNA



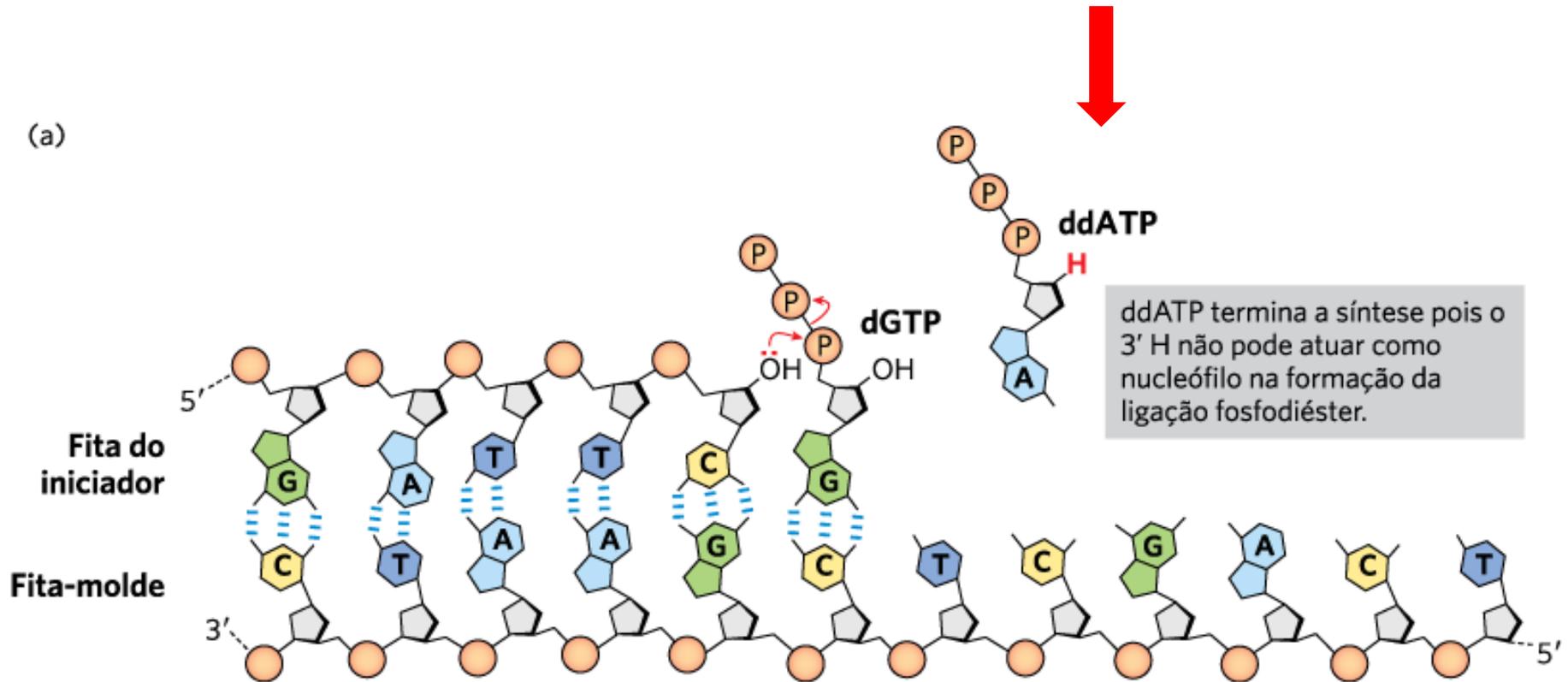
O que ocorre se for utilizado **dd**NTP na síntese de DNA?



2', 3' **didesoxi**ribonucleotídeo trifosfato (ddNTP)

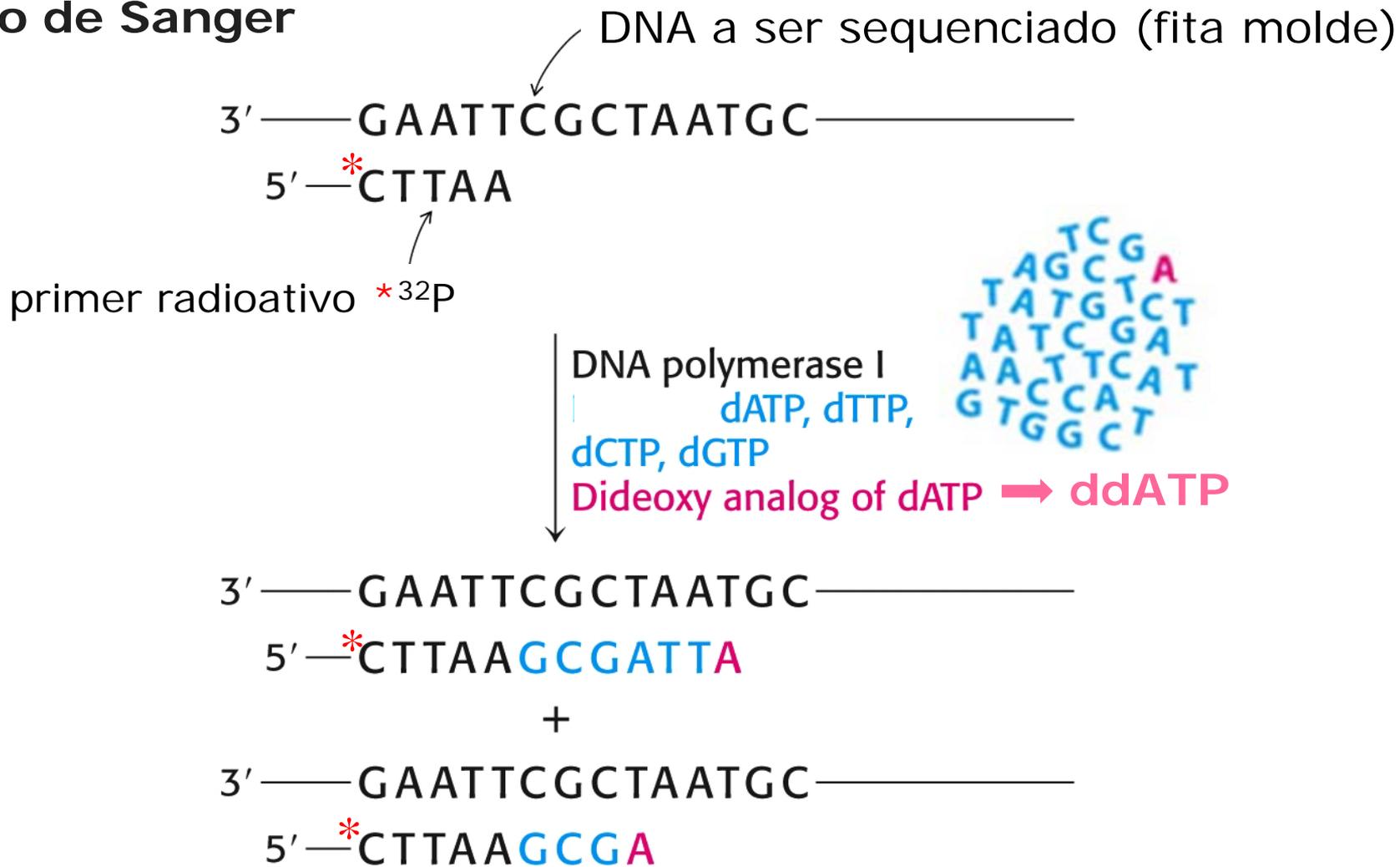
Reação da DNA Polimerase com dNTPs + ddNTPs → interrupção da síntese de DNA após incorporação do ddNTP

(a)



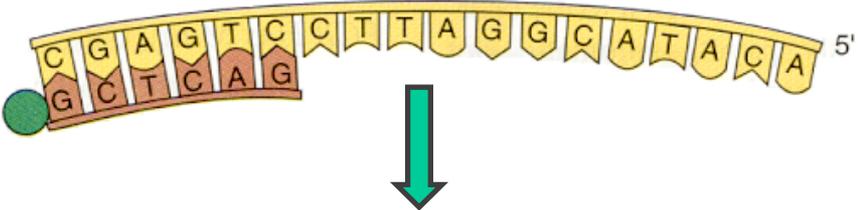
**Este é o princípio do Método de Sanger:**  
terminação controlada da síntese de DNA com ddNTP

# Método de Sanger



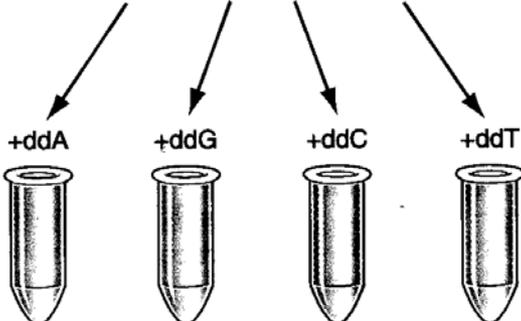
- Novas cadeias de DNA sintetizadas serão separadas por eletroforese em gel de poliacrilamida com **resolução para separar fragmentos de DNA com 1 nucleotídeo de diferença**
- Cadeias sintetizadas marcadas com <sup>32</sup>P (permite a detecção após separação por eletroforese)

# Método de Sanger

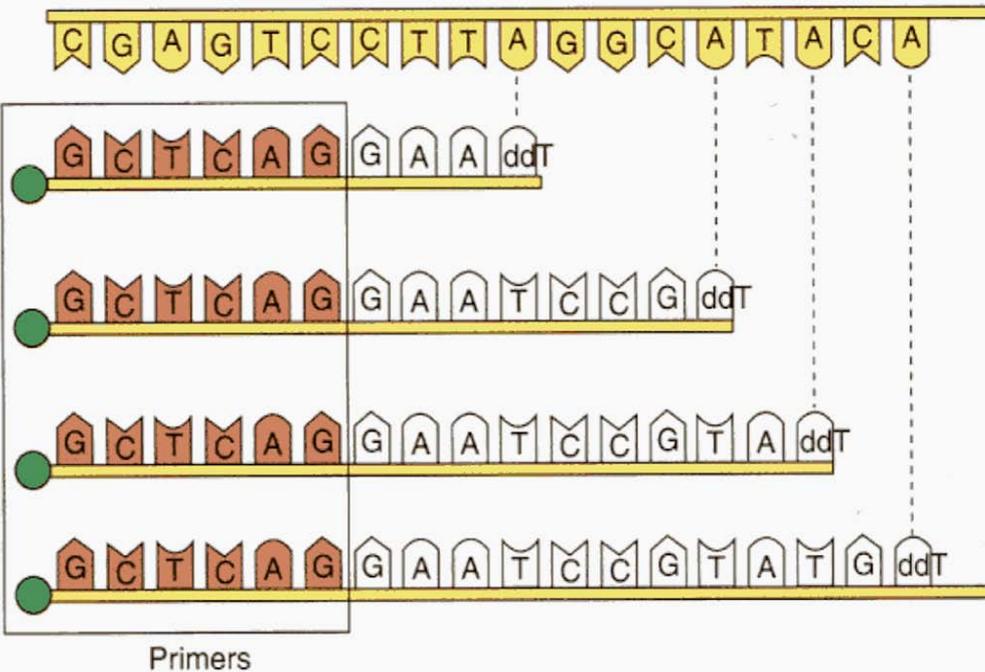


Desnaturação da dupla fita  
Anelamento do "primer radioativo"

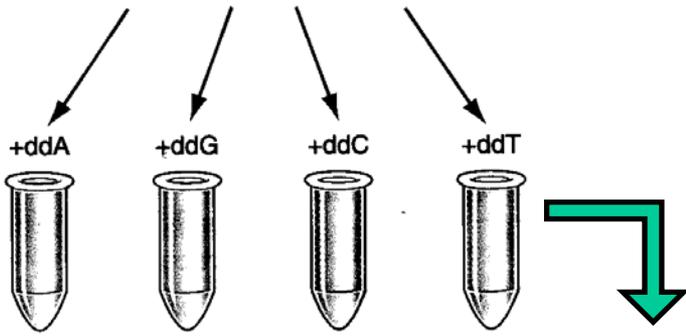
dNTP's and DNA polymerase Adição da enzima e dNTPs



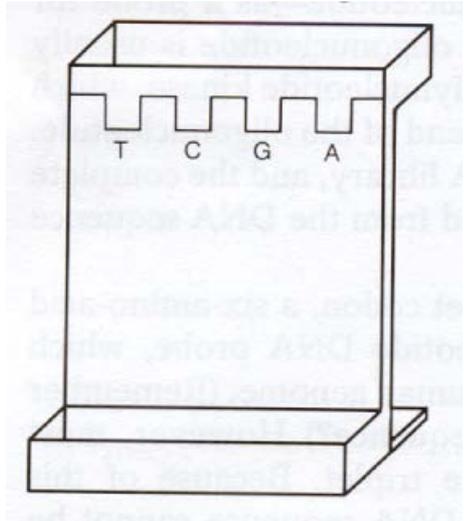
Terminação da síntese: adição dos ddNTPs



dNTP's and DNA polymerase

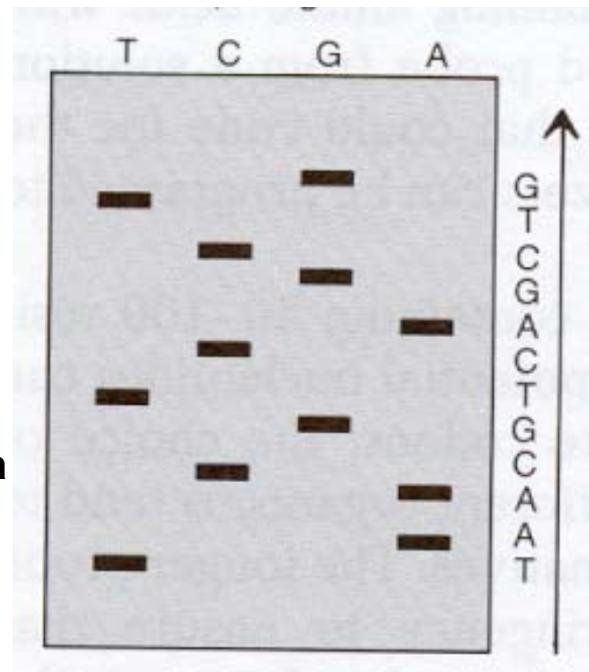


Eletroforese em gel de poliacrilamida desnaturante



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

**Autorradiografia**



Sequência complementar ao DNA molde

**Método de Sanger:**  
 terminação controlada da  
 síntese de DNA com  
 dideoxirribonucleotídeos

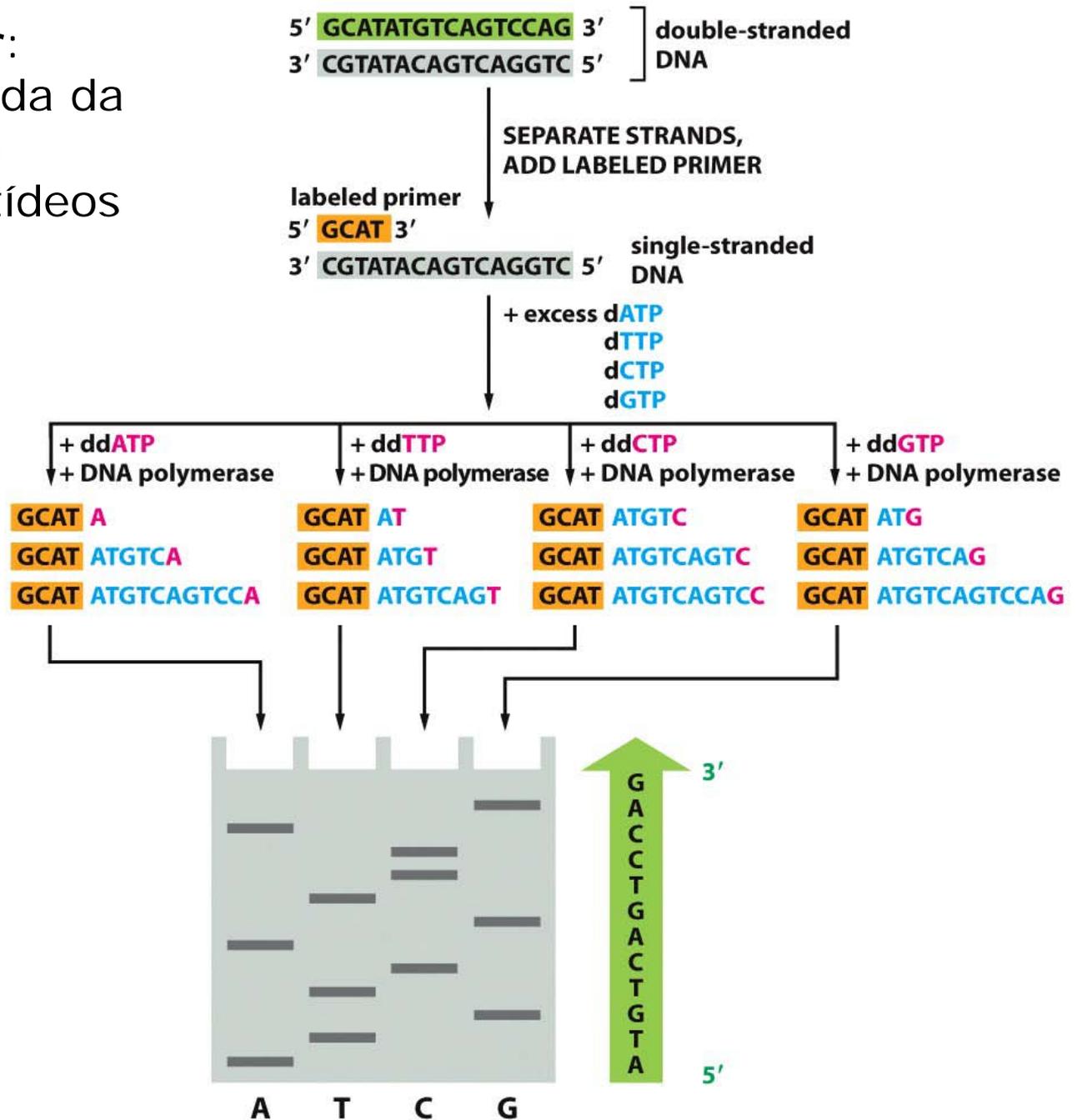
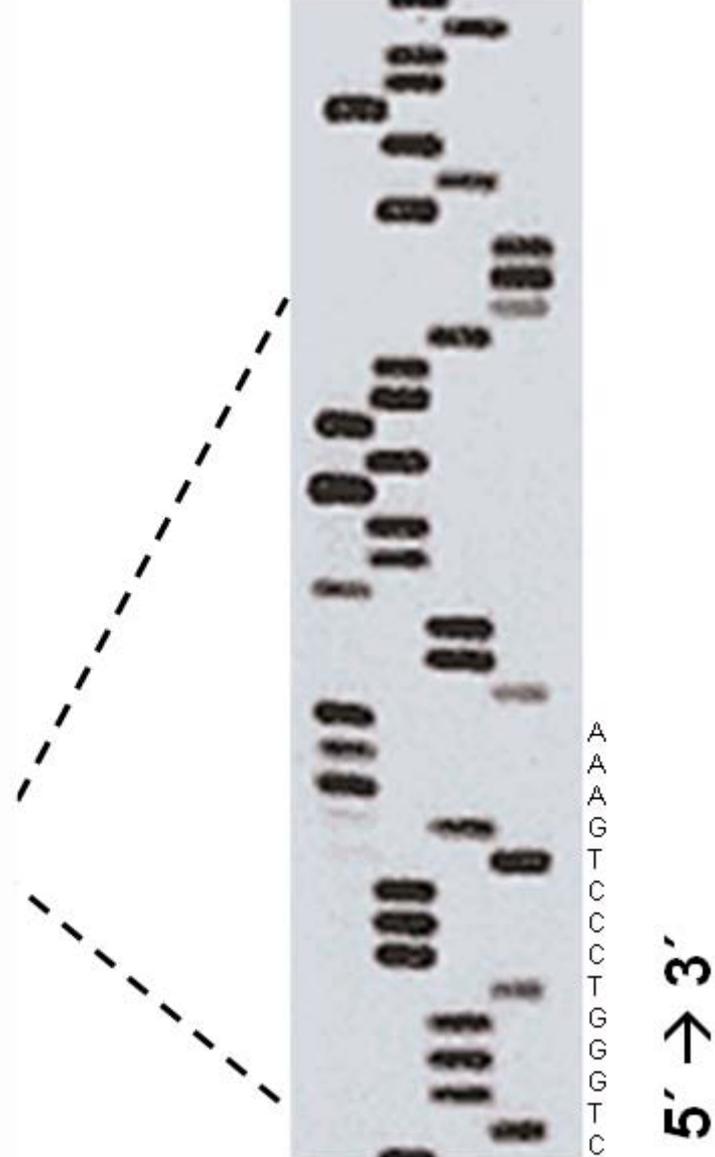


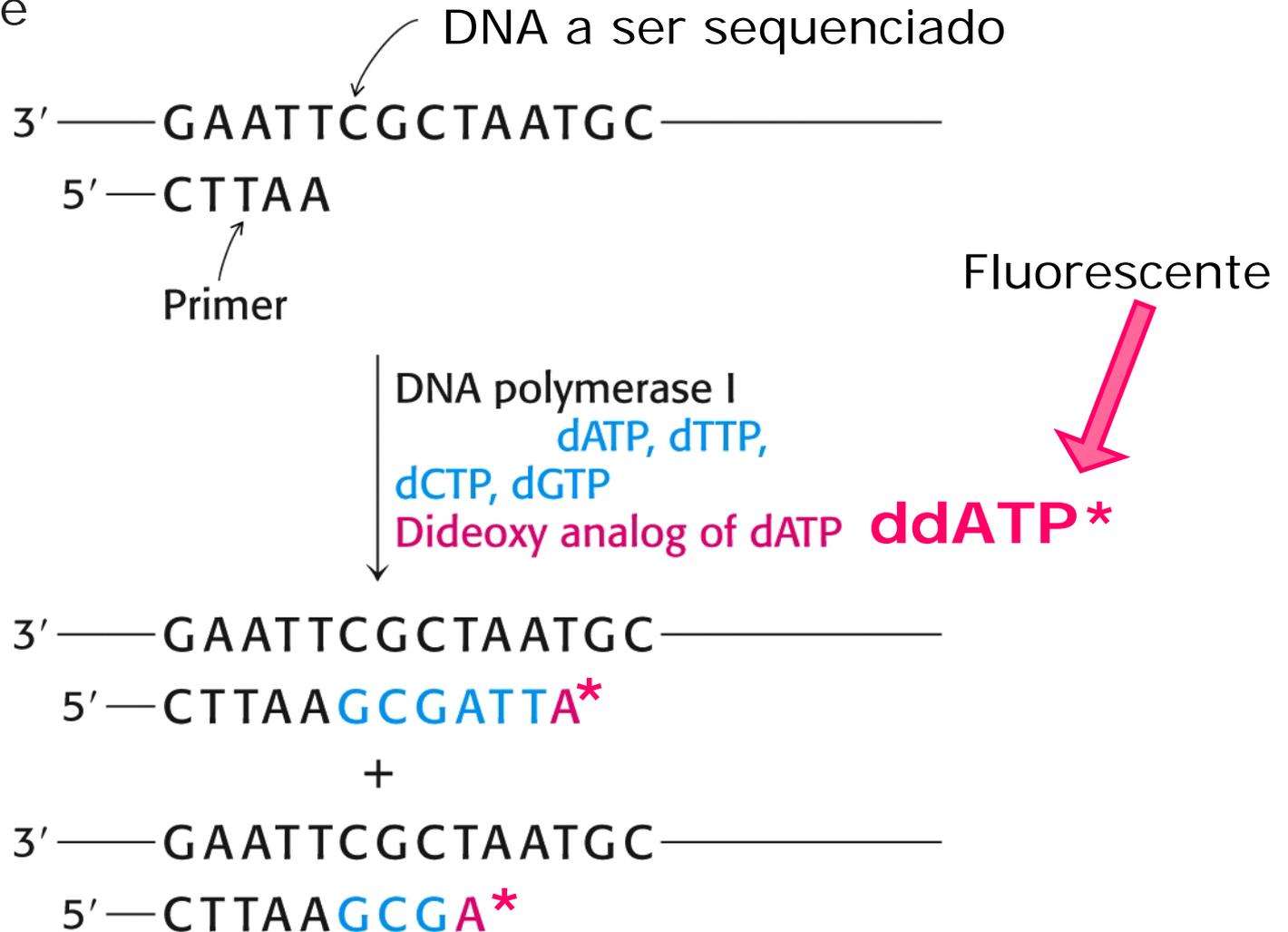
Figure 10-21 Essential Cell Biology 3/e (© Garland Science 2010)

# Autorradiograma de um gel de sequenciamento de DNA



**A automatização** do sequenciamento pelo Método de Sanger foi possível a partir da utilização de ddNTPs acoplados a grupos fluorescentes

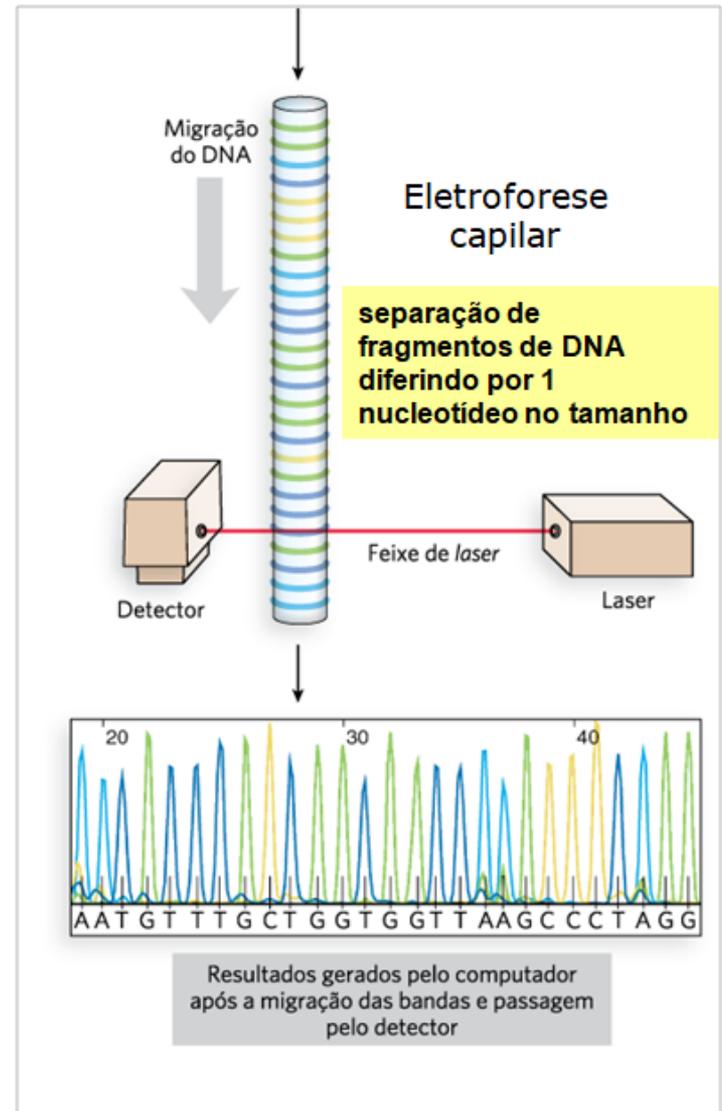
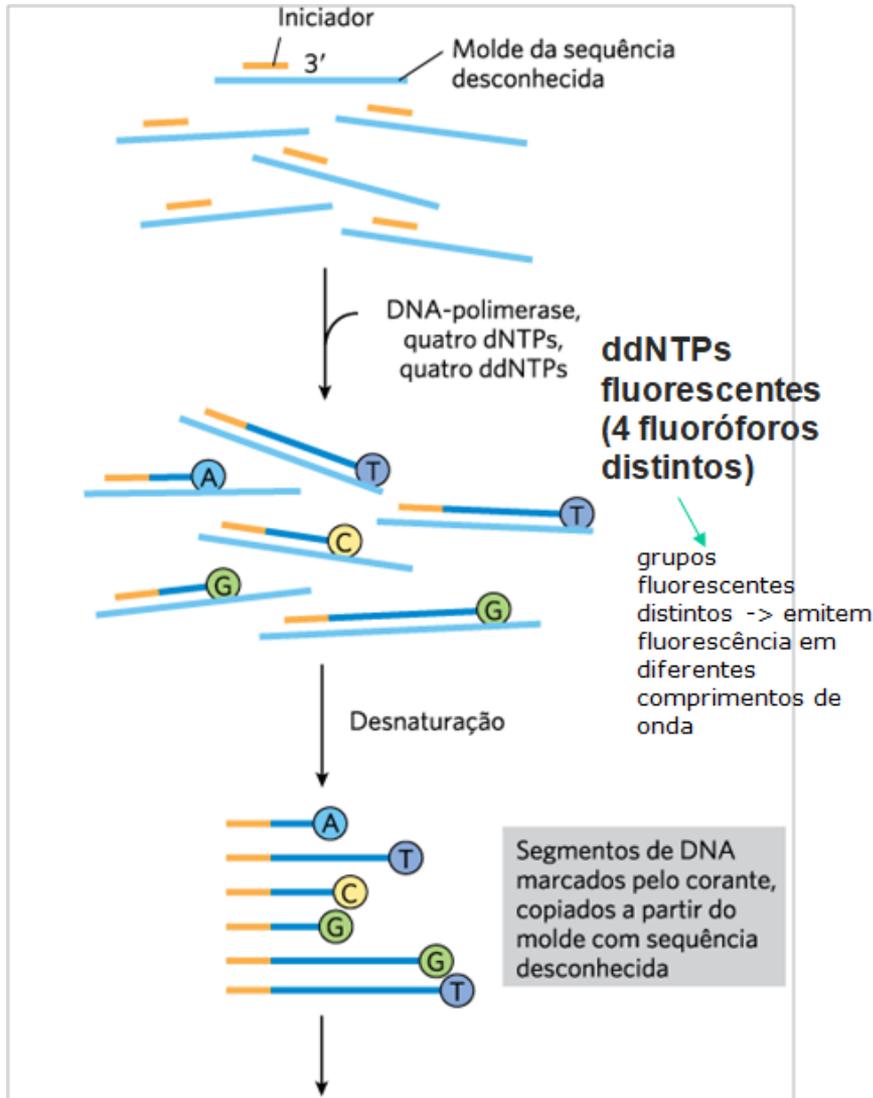
# Sequenciamento pelo Método de Sanger



Novas cadeias de DNA serão separadas por eletroforese capilar com **resolução para separar fragmentos de DNA com 1 nucleotídeo de diferença –detecção por fluorescência**

# Sequenciamento pelo Método de Sanger

Automatização através do uso de ddNTPs marcados com fluorocromos distintos



# ddNTPs fluorescent

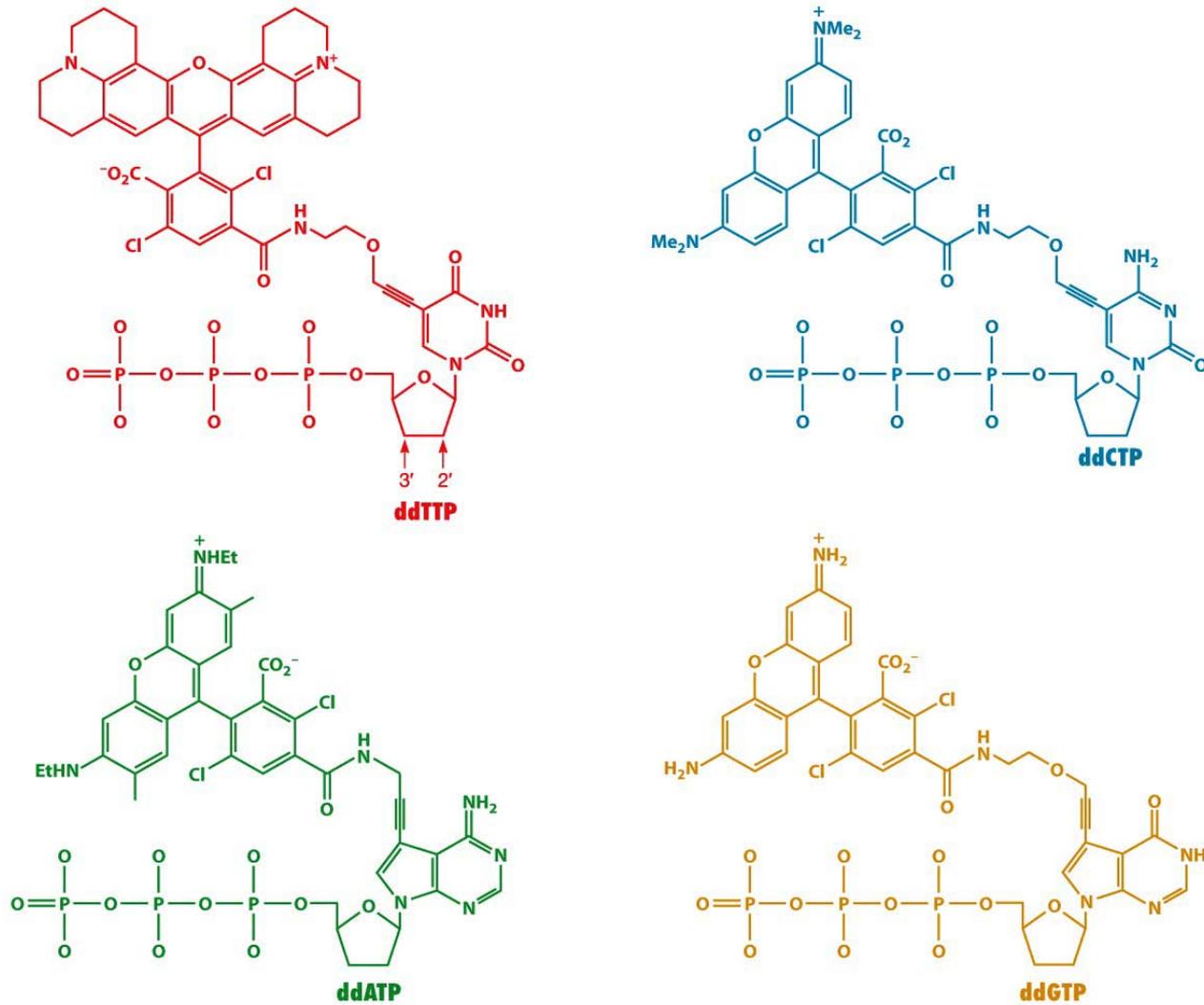
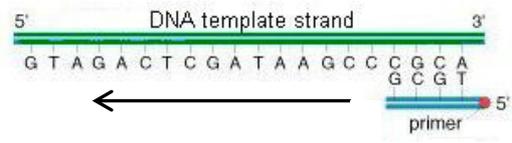
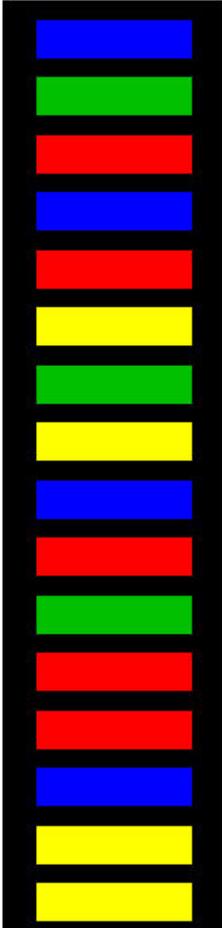
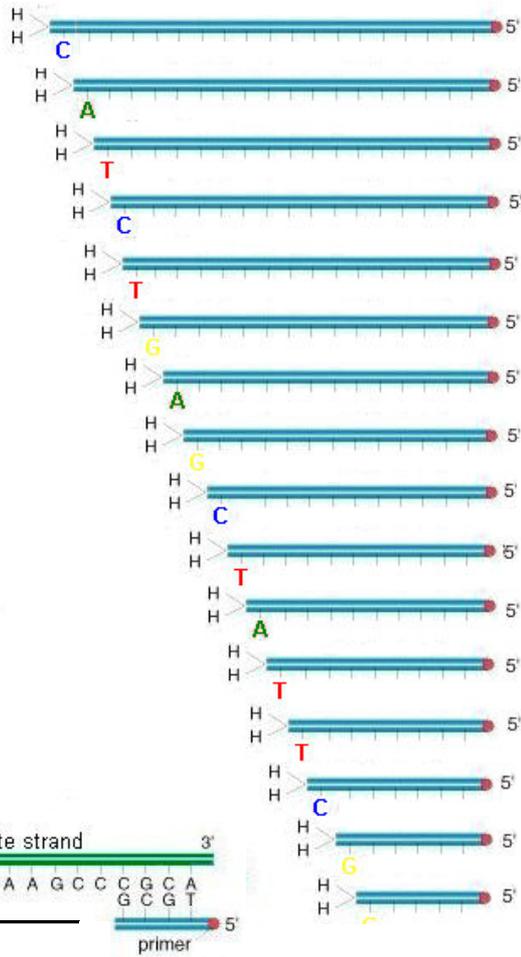


Figure 10-10  
*Genes and Genomics: A Short Course (3e)*  
© 2007 W. H. Freeman and Company

5' GTAGACTCGATAAGCCCGCA 3'



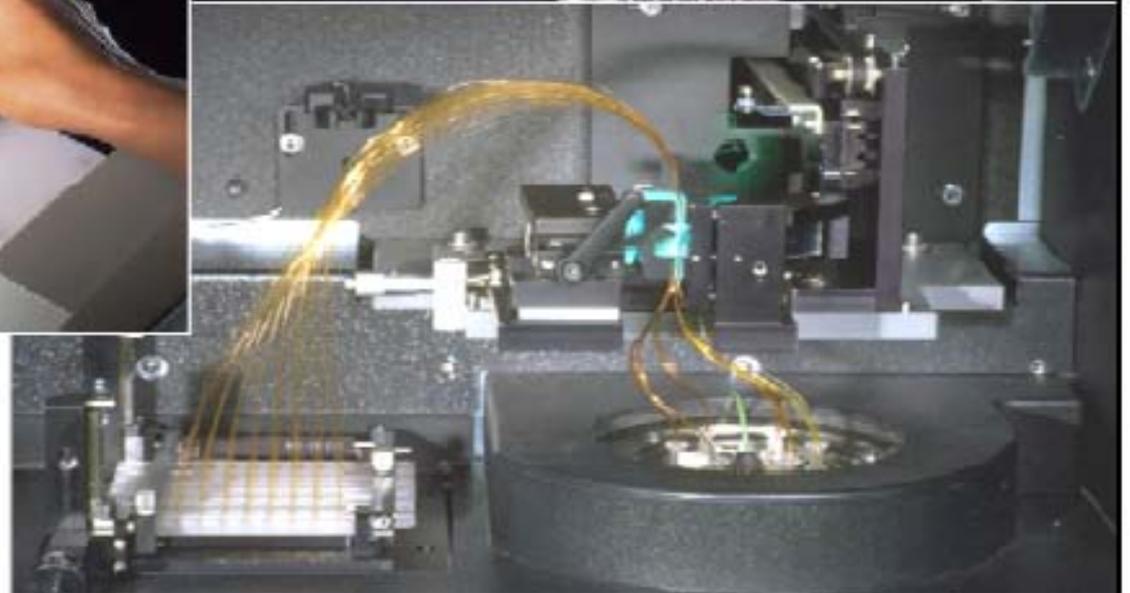
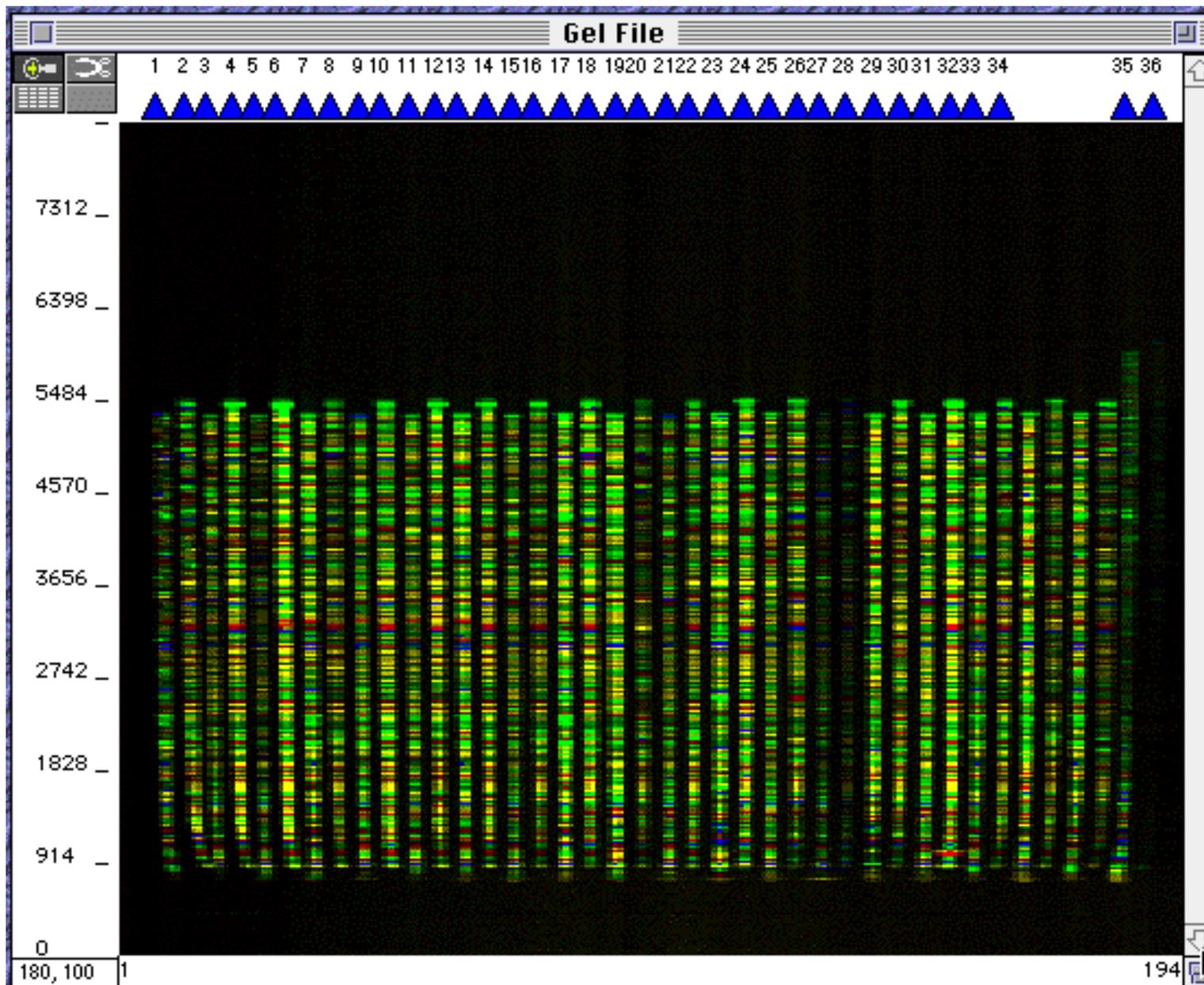
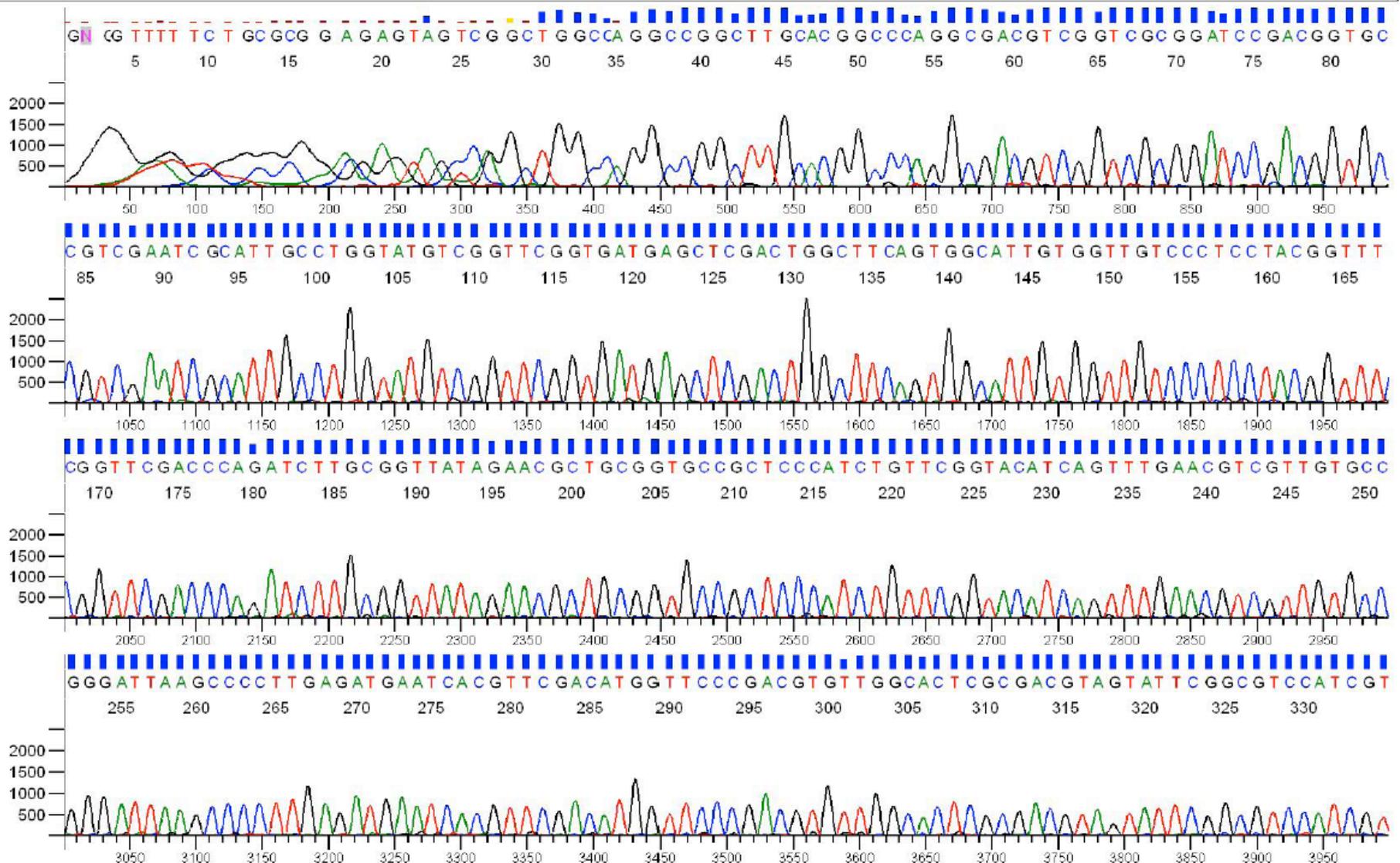


Imagem da detecção por fluorescência no sequenciamento automatizado: cada amostra em um capilar

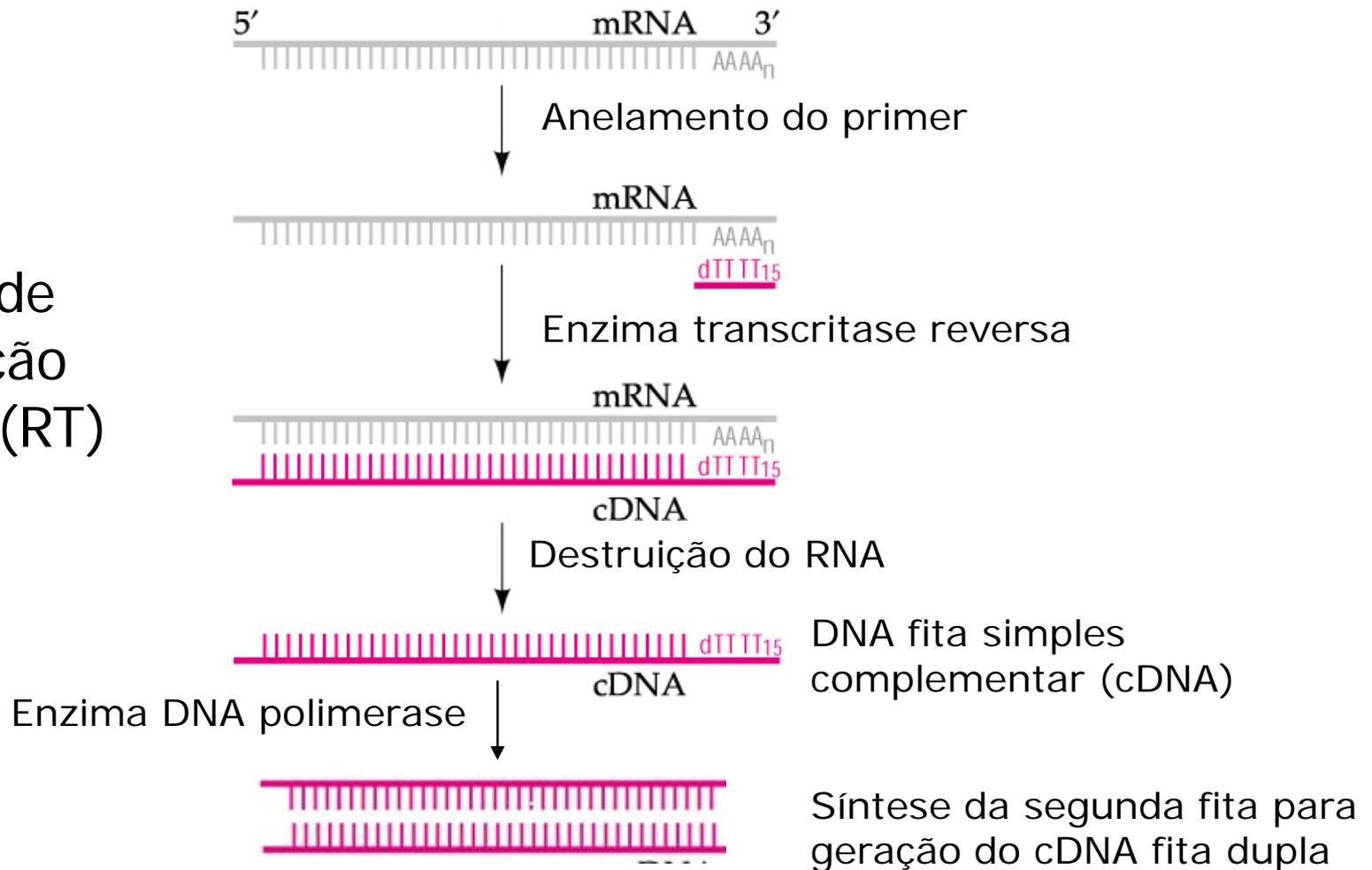


# Eletoferograma do sequenciamento automatizado pelo método de Sanger



Também é possível determinar a ordem das bases nitrogenadas adenina (A), guanina (G), citosina (C) e uracila (U) da molécula de **RNA**. Porém é necessário transcrever a molécula de RNA em DNA antes do sequenciamento propriamente dito.

## Reação de transcrição reversa (RT)



Atividade online

**Sequenciamento de Sanger**

<http://www.bch.cuhk.edu.hk/vlab2/animation/dna/index.html>

# Limitações no Sequenciamento Sanger

Baixa capacidade: sequencia um único fragmento de DNA por vez.

Processo mais lento, normalmente levando várias horas para gerar dados de sequência.

Custo elevado por base sequenciada

Adequado para sequenciamento de genomas menores ou sequenciamento direcionado de genes específicos

# **Metodologias de sequenciamento de alto desempenho**

***(Next Generation Sequencing, NGS)***

# Características do NGS

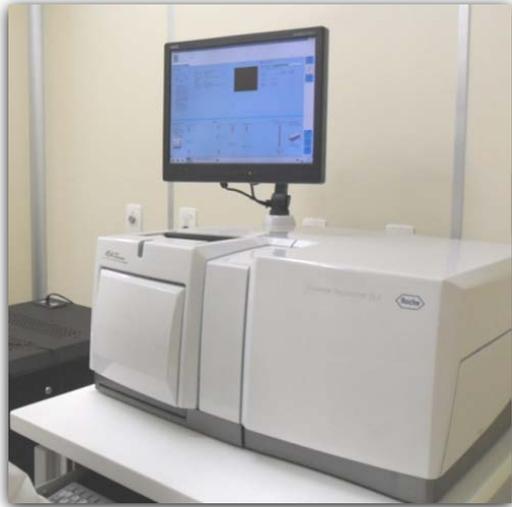
- **Alto rendimento.** Pode sequenciar simultaneamente milhões a bilhões de fragmentos de DNA de forma paralelizada.
- **Altamente passível de multiplexação,** permitindo sequenciamento eficiente de muitas amostras em uma única execução.
- **Custo por base reduzido:** Torna-se rentável especialmente para projetos de sequenciamento em grande escala.
- Mais adequado para sequenciamento de novo de genomas grandes e complexos.

## Desvantagens em relação ao sequenciamento de Sanger

Leituras mais curtas, normalmente variando de 50 a 600 pares de bases

Embora altamente preciso, o NGS pode estar sujeito a certos tipos de erros, principalmente em regiões homopoliméricas e no final das leituras.

# Equipamentos/Tecnologias de NGS



454 - Roche



Illumina HiSeq



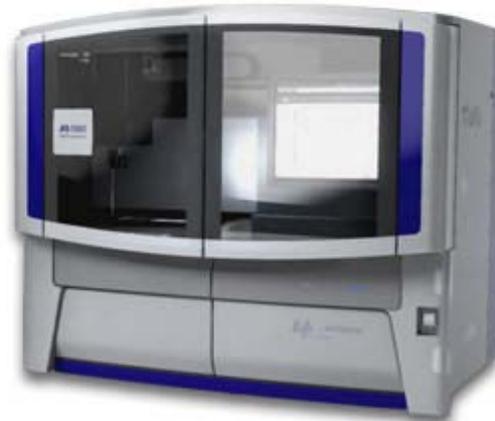
Illumina MiSeq



Illumina  
NextSeq500



PacBio



SoLiD

Ion Proton

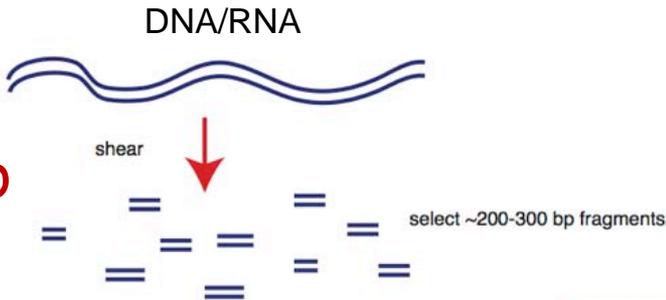


# Tecnologias de sequenciamento de DNA de alto desempenho

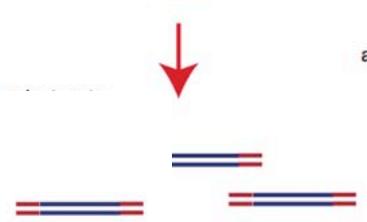
<b>Tecnologia</b>	<b>Química do Sequenciamento</b>	<b>Tipo de preparação do molde</b>	<b>Vantagens/Desvantagens</b>
Illumina	Síntese com DNA polimerase detecção de fluorescência	Amplificação por ponte	<ul style="list-style-type: none"><li>•Alta capacidade</li><li>•Tamanho de sequencia limitado (&lt; 1000 pb)</li></ul>
Ion Torrent	Síntese com DNA polimerase detecção de protons liberados na síntese	PCR em emulsão	<ul style="list-style-type: none"><li>• Alta capacidade</li><li>• Custo reduzido</li><li>• Menor acurácia</li><li>•Tamanho de sequencia limitado (&lt; 1000 pb)</li></ul>
Pacific Biosciences	Síntese com DNA polimerase detecção da fluorescência	Não utiliza PCR. Sequenciamento direto da molécula	<ul style="list-style-type: none"><li>• sequencias longas (&gt; 10.000 pb)</li></ul>
Oxford Nanopore	Condutância da bases a medida que o atravessa um nanoporo proteico	Não utiliza PCR. Sequenciamento direto da molécula	<ul style="list-style-type: none"><li>• sequencias longas (&gt; 10.000 pb)</li><li>• detecta modificação químicas nas bases</li><li>• baixo custo</li><li>• capacidade limitada</li></ul>

# Illumina - sequenciamento por polinização

Fragmentação



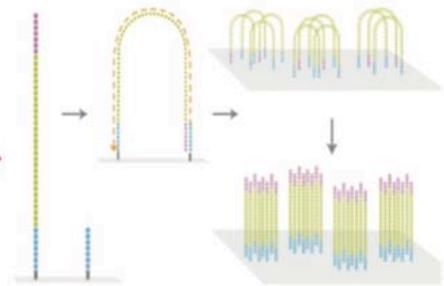
Ligação de adaptadores



apply to flowcell



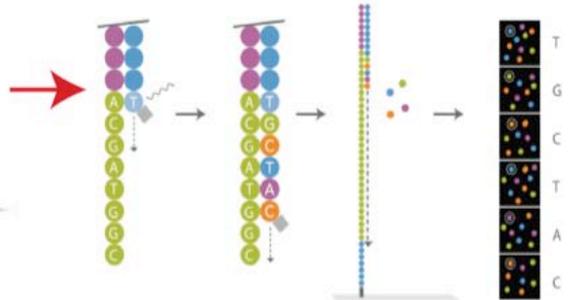
cluster generation by solid phase PCR (bridge amplification)



Amplificação Clonal ("bridge amplification")

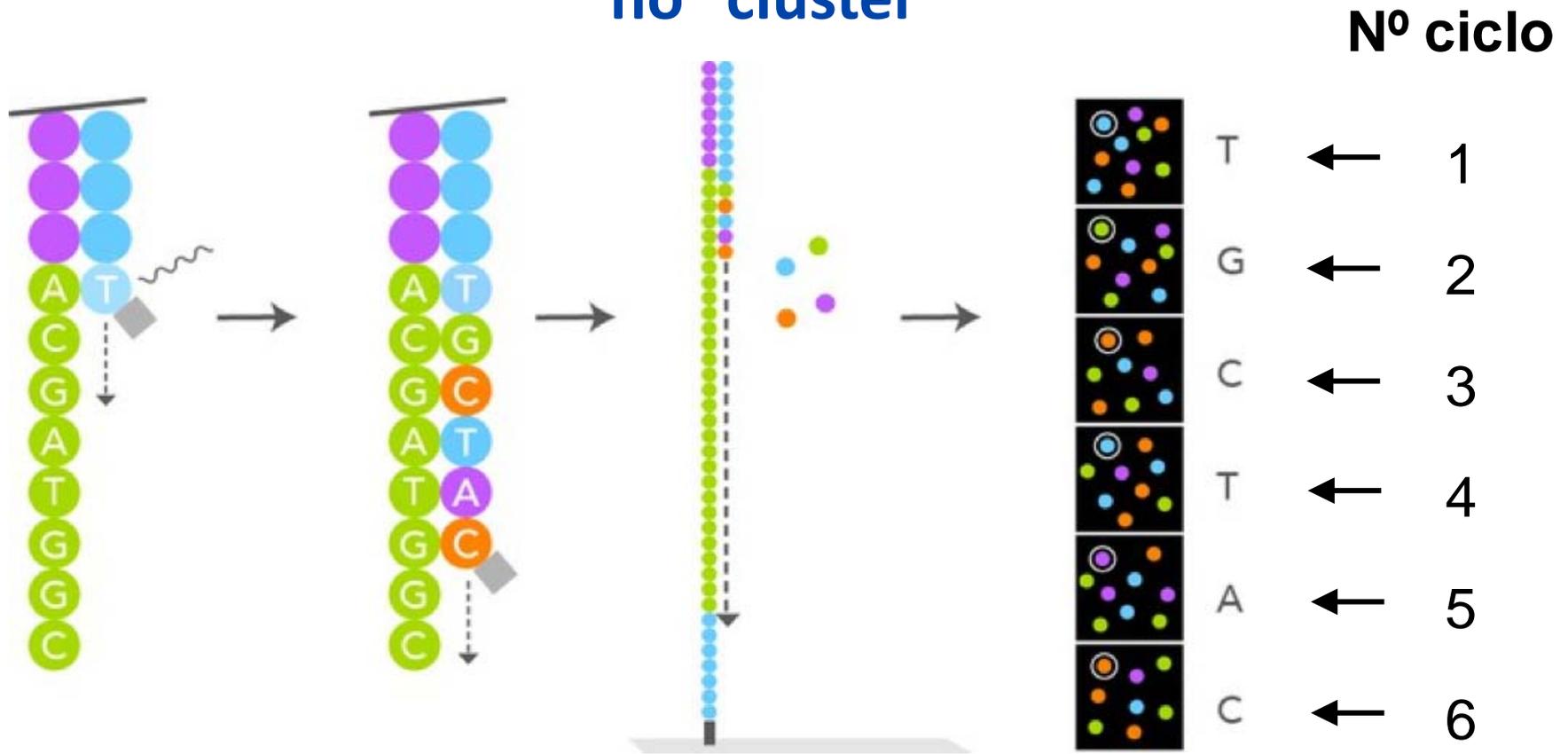


sequencing by synthesis with reversible terminators



Sequenciamento por síntese

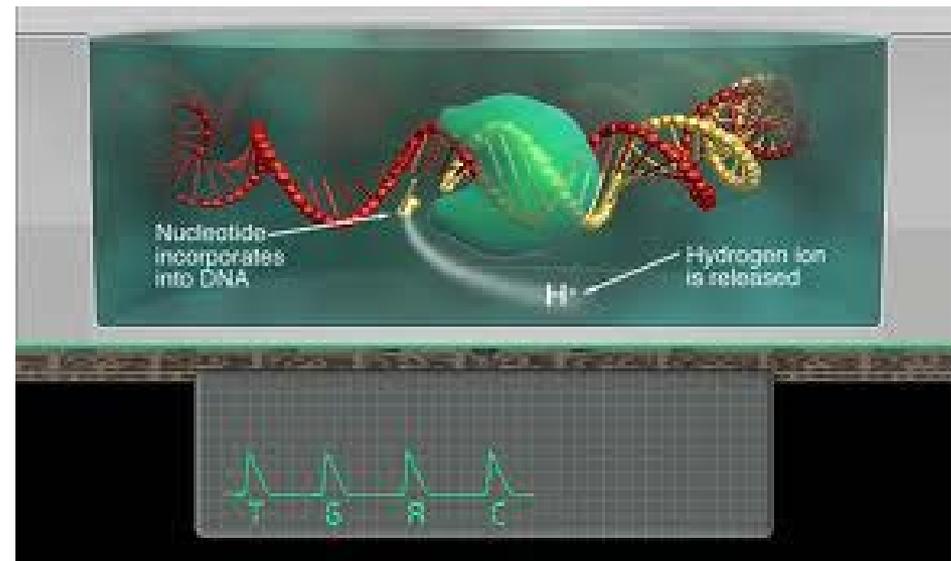
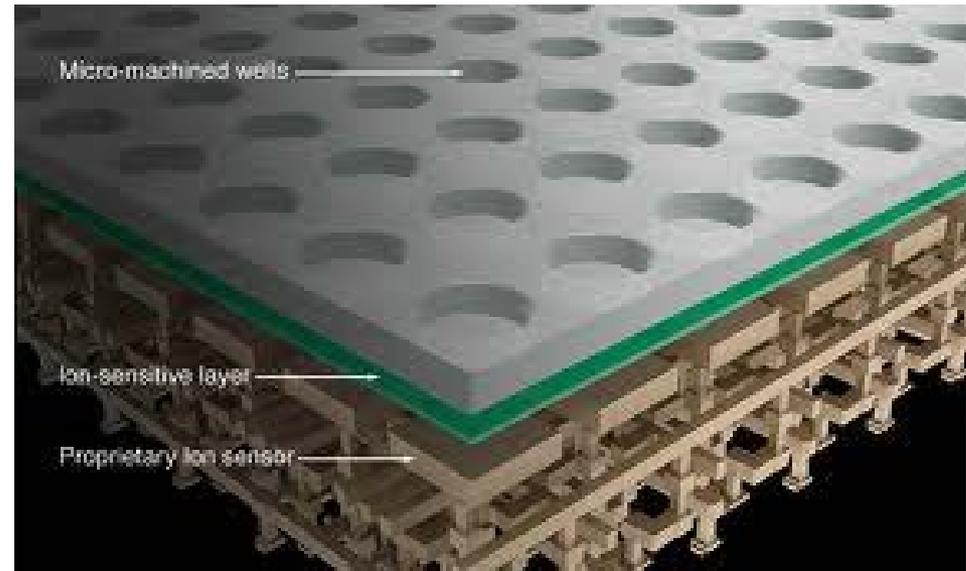
# Análise estatística das fluorescências emitidas em cada ciclo permite determinar a sequência do DNA presente no "cluster"



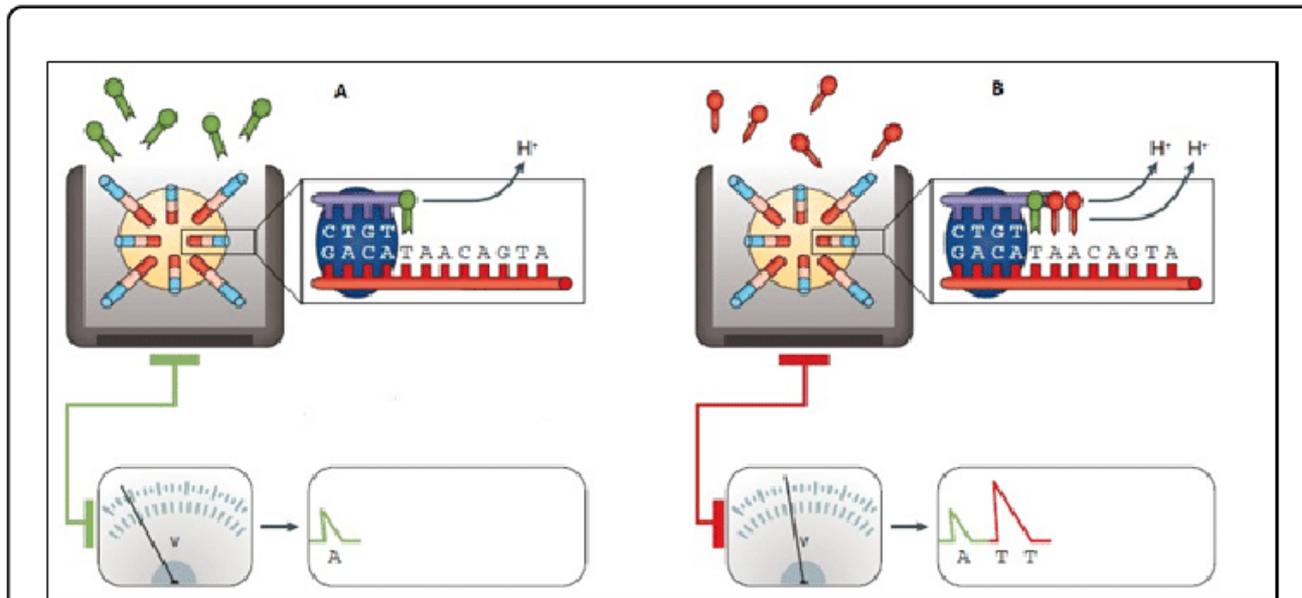
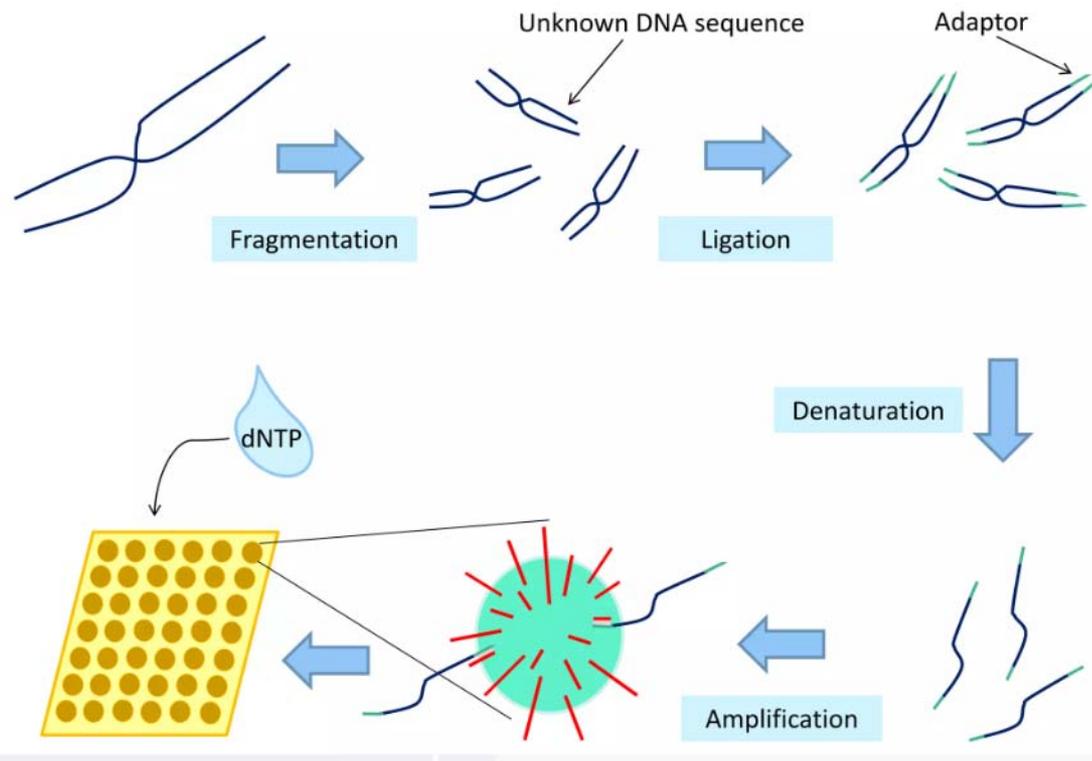
Bilhões de "clusters" em cada corrida

# Ion Torrent – Life technologies

- Não utiliza fluoróforos.
- Baixo custo
- Próton liberado após incorporação de nucleotídeo causa mudança de pH que é detectada pelo equipamento



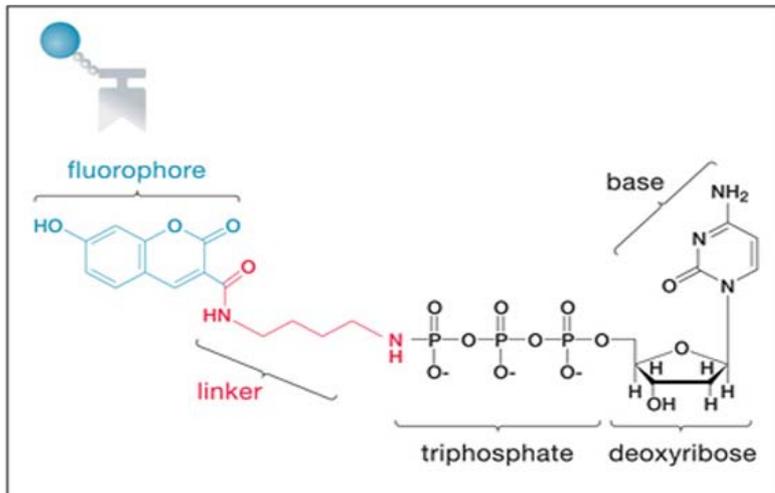
# Ion Torrent





# Single Molecule Real Time (SMRT™) DNA sequencing technology – Pacific Bioscience

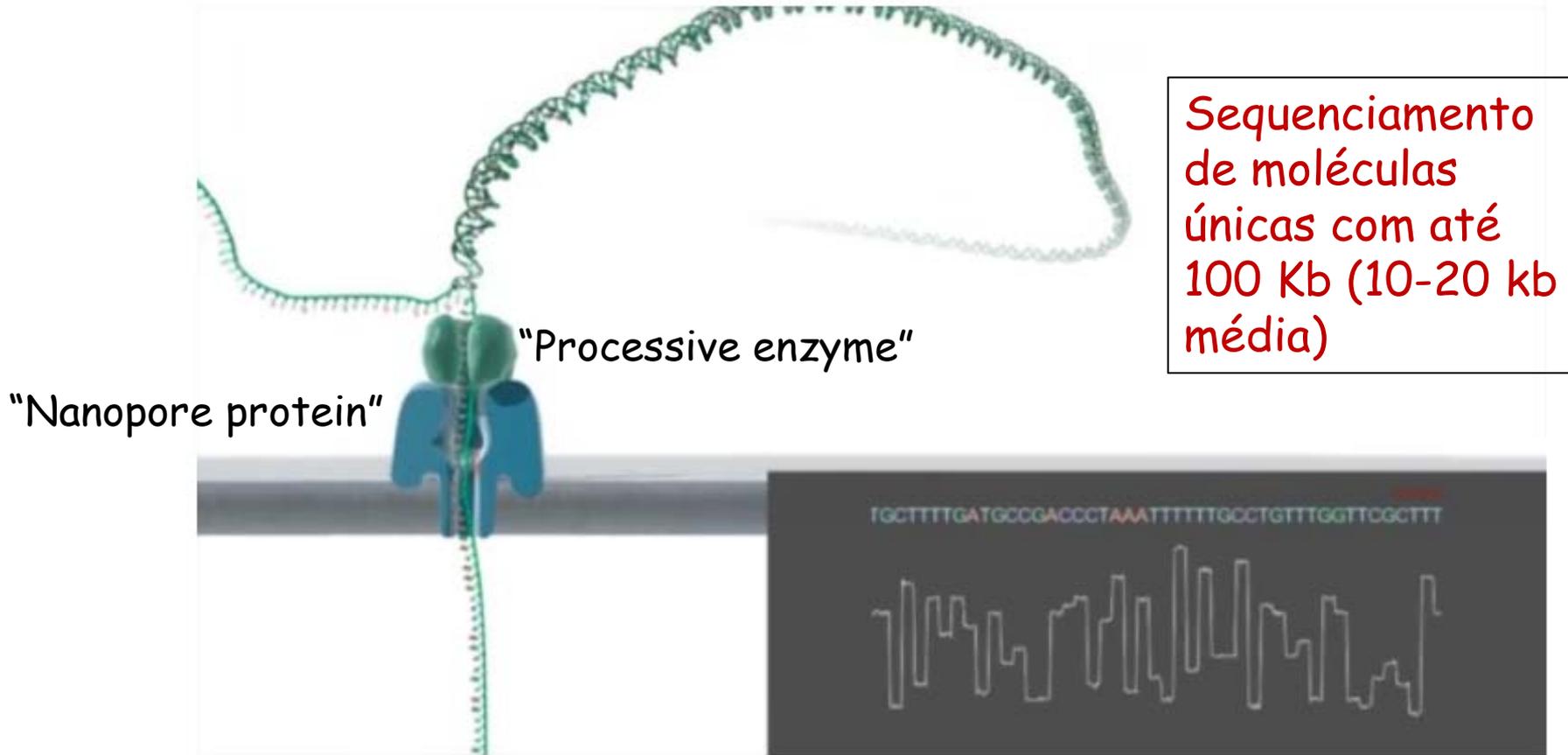
Síntese contínua, com detecção do nucleotídeo incorporado em tempo real



- Nucleotídeos marcados com fluoróforos no fosfato gama.
- fluoróforo liberado após a formação da ligação fosfodiester
- **Sequencias longas (> 20 kb)**



# Nanopore sequencing™ – Oxford Nanopore



Nucleotídeos da cadeia da DNA são detectados a medida que a fita simples passa pelo nanoporo devido a mudança na condutância dentro do canal

Permite detectar modificações químicas nas bases do DNA/RNA:

- metilação de DNA
- edição de RNAs

<https://www.youtube.com/watch?v=RcP85JHLmnl>

# Algumas aplicações do sequenciamento de DNA de nova geração

- Obter a sequência completa de genomas
- Identificar mutações no genoma de tumores
- Obter a sequência de genes expressos (transcritos/RNAs) em uma célula/tecido (transcritoma)

# Como os genomas são sequenciados atualmente

DNA genômico



Fragmentação



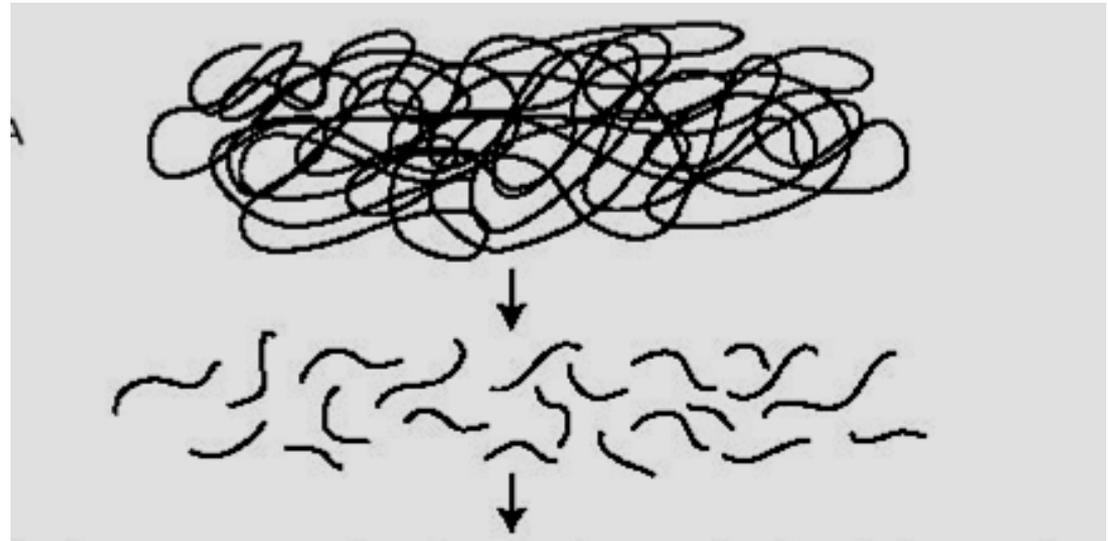
Amplificação dos fragmentos



Sequenciamento NGS



Montagem (in silico)



# Regiões repetitivas no DNA dificultam a montagem

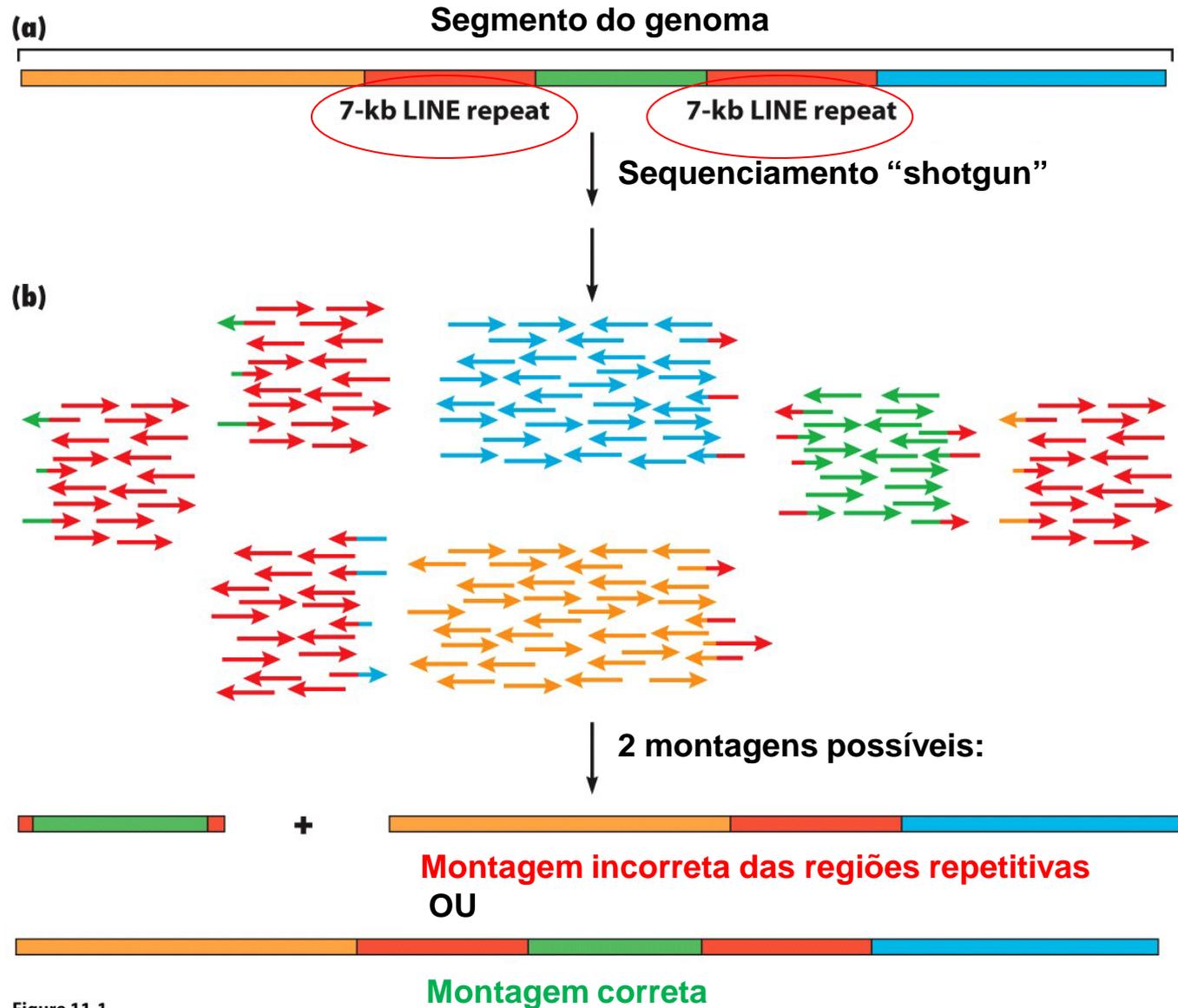
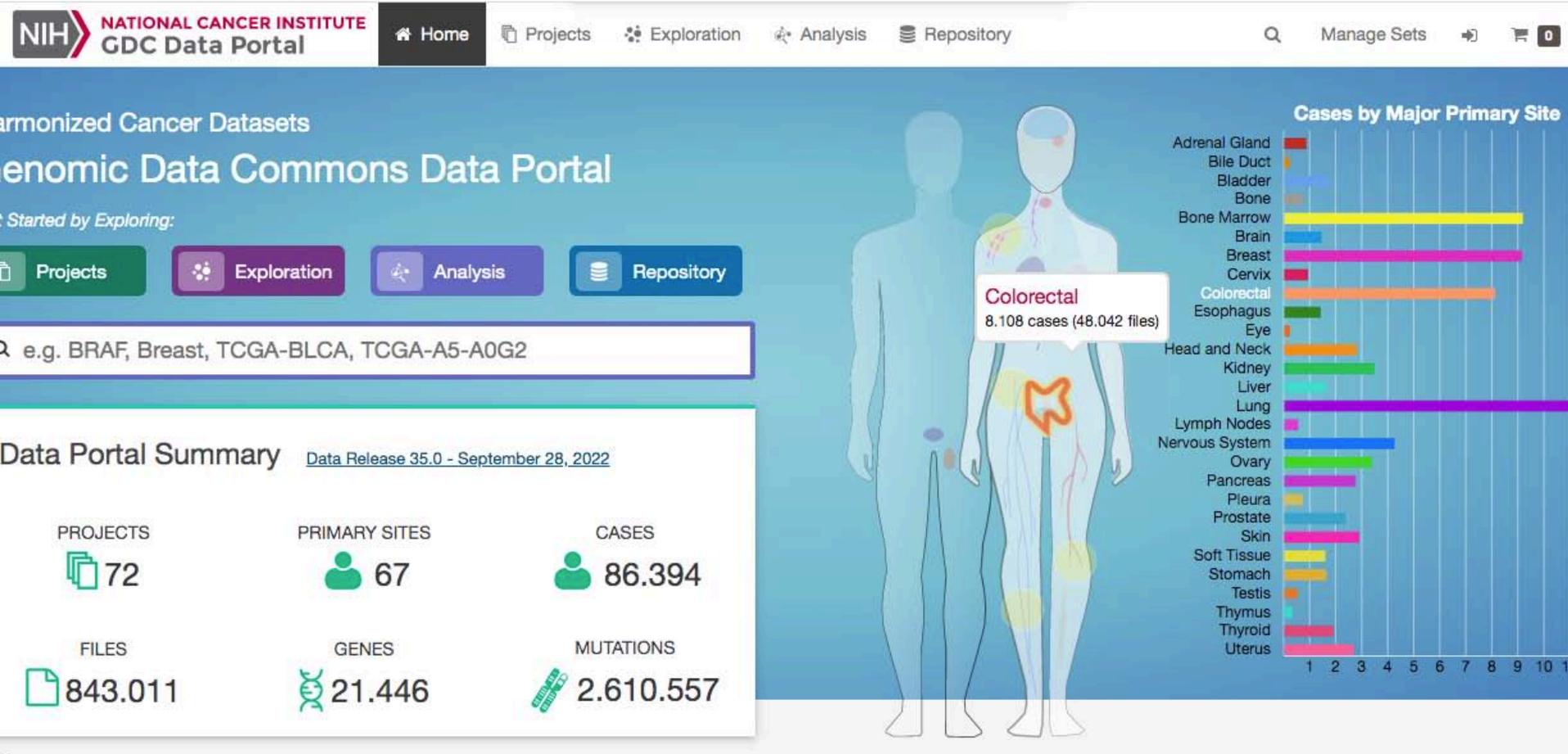


Figure 11-1  
*Genes and Genomics: A Short Course (3e)*  
© 2007 W.H. Freeman and Company

# The Cancer Genome Atlas (TCGA)



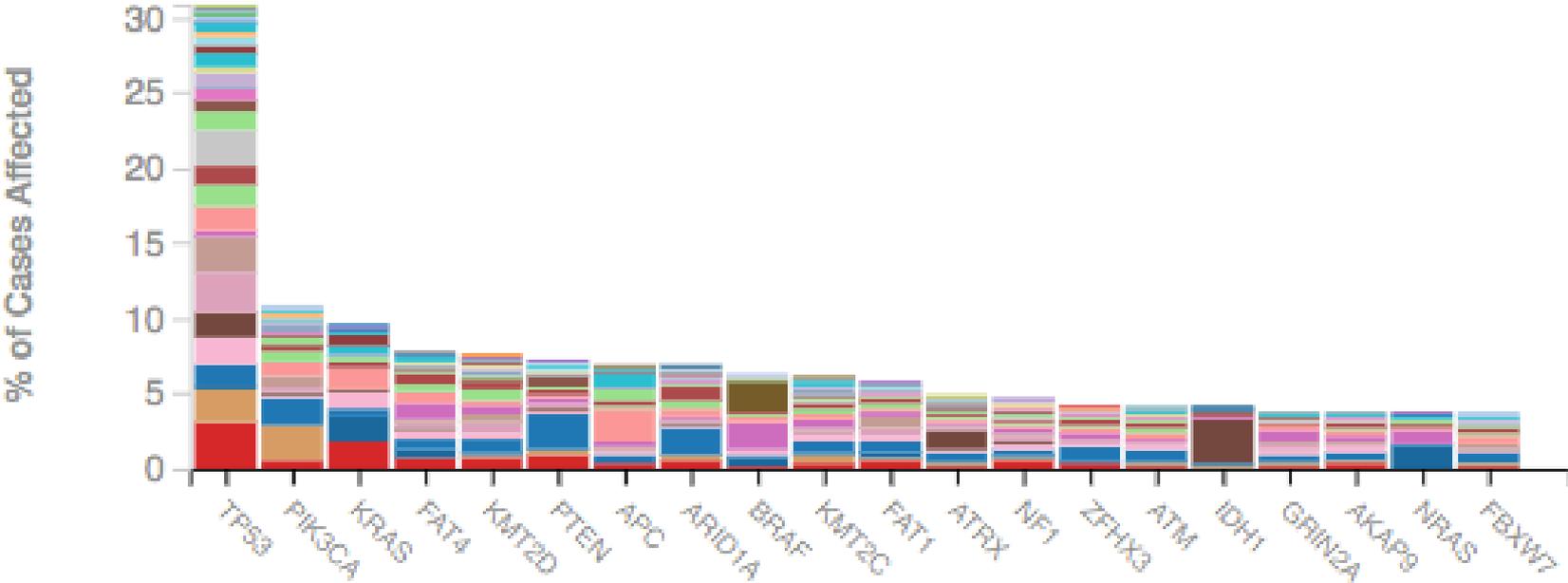
<https://portal.gdc.cancer.gov/>

# Mutações mais frequentes em tumores humanos

Frequencia de mutações mais frequentes em 86.394 casos analisados no TCGA

### Top Mutated Cancer Genes in Selected Projects ?

% of Cases Affected     # of Cases Affected



Sequenciamento de transcritomas - RNAseq

# Transcritoma:

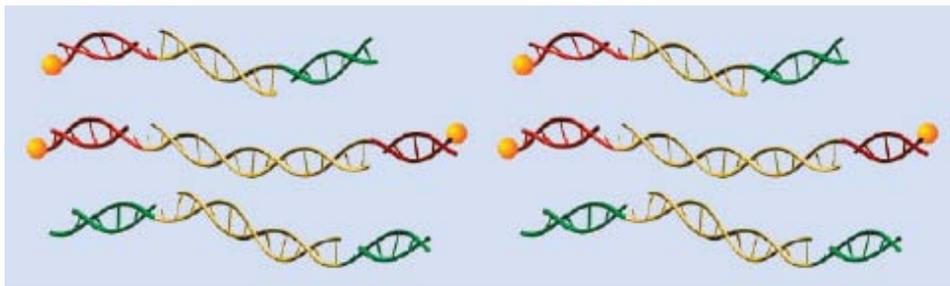
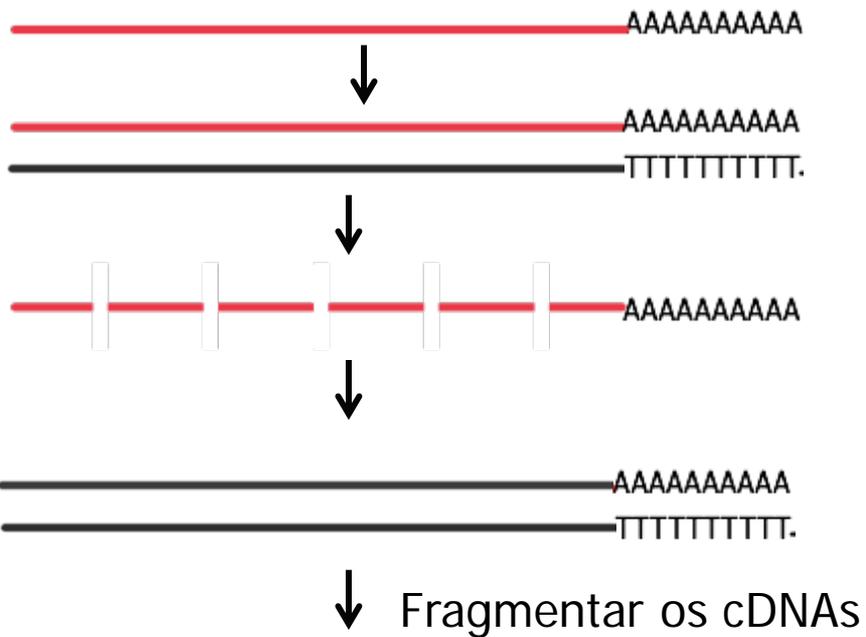
Conjunto de todas moléculas de RNA (transcritos) existentes nas células em um dado momento/condição

- Dinâmico
- Variável

Tão importante quanto o conteúdo gênico é **como, onde e quando** cada um dos genes é expresso (transcrito)!

# RNAseq

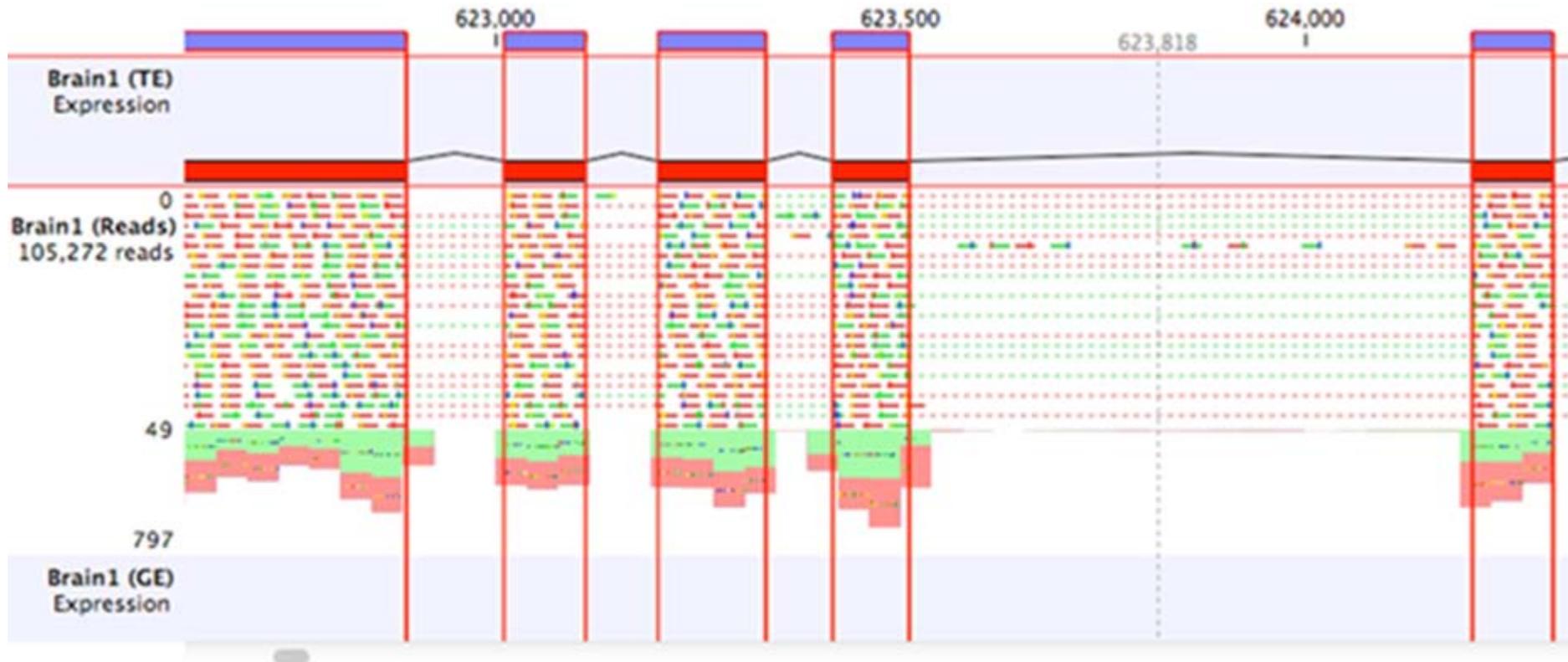
## Preparação da biblioteca de cDNA para sequenciamento NGS



1. Purificação do RNA
2. Síntese do cDNA com transcriptase reversa
3. Degradação do mRNA
4. Síntese da segunda fita do cDNA fita dupla
5. Ligação de adaptadores
6. Sequenciamento NGS: bilhões de sequências (*reads*) são geradas

# RNAseq

Mapeamento *in silico* das sequências no respectivo genoma



Estimativa da expressão pela abundância de sequências obtidas no RNAseq

# Abundância de transcritos avaliadas em diferentes amostras por RNAseq

