

# ALGORITMOS

## 2º Capítulo



Mulheres em  
Bioinformática  
& Data Science LA  
*Promovendo a colaboração entre mulheres*

- *Algoritmo de Euclides;*
- *Pesquisa Exaustiva.*

# Introdução



Wikipedia Commons

- Um algoritmo é um **procedimento computacional**;
- Pode ser pensado como um **conjunto de instruções passo a passo** para transformar uma **entrada** numa **saída** de acordo com o enunciado do problema;
- O primeiro algoritmo conhecido é o **algoritmo euclidiano** para calcular o máximo divisor comum, por volta de 400–300 a.C.;
- O estudo moderno de algoritmos remonta ao início da década de 1960, quando a disponibilidade e os **recursos limitados** dos primeiros computadores eram razões convincentes para os usuários se esforçarem para projetar **algoritmos de computador eficientes**.

# Algoritmo de Euclides

$$\text{mdc}(a, b) = \text{mdc}(b, a - b)$$

$$\text{mdc}(630, 135) = 45$$

$$630 - 135 = 495$$

$$495 - 135 = 360$$

$$360 - 135 = 225$$

$$225 - 135 = 90$$

$$135 - 90 = 45$$

$$90 - 45 = 45$$

$$45 - 45 = 0$$



# Técnicas para Projetar Algoritmos de Bioinformática

## 1. Pesquisa Exaustiva:



- É um tipo de **abordagem de Força Bruta**;
- A técnica, do ponto de vista conceitual, representa a abordagem mais simples possível para resolver um problema;
- Técnica que **examina exaustivamente** e considera todas as possíveis soluções candidatas.

# Técnicas para Projetar Algoritmos de Bioinformática

## 1. Pesquisa Exaustiva:

- Você saberia dar um exemplo de algoritmo de força bruta na vida real?
- O algoritmo muda caso haja uma ou mais soluções ótimas para o problema em questão? --> Algoritmos diferentes ou iguais para soluções diferentes?



Mulheres em  
Bioinformática  
& Data Science LA  
*Promovendo a colaboração entre mulheres*

# Técnicas para Projetar Algoritmos de Bioinformática

## 1. Pesquisa Exaustiva:

- Você saberia dar um exemplo de algoritmo de força bruta na vida real?

### Exemplo:

#### Contexto:

1. Tenho um arquivo **multifasta**;
2. Porém, este já havia sido lido por um **script** anterior que, provavelmente, alterou indevidamente alguma sequência no formato .fasta;
3. Agora precisamos encontrar a sequência fora do padrão .fasta desse arquivo.

**Algoritmo:** Procuraria sequência por sequência para verificar se cada uma delas está seguindo o padrão fasta.

# Técnicas para Projetar Algoritmos de Bioinformática

Header —●>VIT\_201s0011g03530.1  
Sequence —●AATTAAGCATAAATACTCACTCTTACCCCCTTATTTTCTTATCTCTCATCACTTTTGGTGCGAAG  
●GACCATGAGAACAAAGCTGCAATGGGTGTAGGGTTCTTCGCAAGGCATGCAGCCAAGACTGCATCA

Header —●>VIT\_201s0011g03540.1  
Sequence —●CAGGTAGCGTGAAGTTAAACCCTAGCGCTTTAGACAAACAGCTGTAGTCACCGCCCACAAACACC  
●AGCCTCTGAGACACCACCTCAAACCTTTCCACTTAAATACACATCCCTCACACCCTTTTCAATTC

Header —●>VIT\_201s0011g03550.1  
Sequence —●CATGCAAAGCTGAACGCGATGCTGTGATTGGTGGTAAGTGGTAGTTGAGTAAATTTGACAGTGAA  
●GCCGAAATGGTAAAAGACTAAGGCTAGAAGTAGAATAACCACTGTTCTTCTCATCACGTGGGCCCCA

Header —●>VIT\_201s0011g03530.1  
Sequence —●AATTAAGCATAAATACTCACTCTTACCCCCTTATTTTCTTATCTCTCATCACTTTTGGTGCGAAG  
●GACCATGAGAACAAAGCTGCAATGGGTGTAGGGTTCTTCGCAAGGCATGCAGCCAAGACTGCATCA

Header —●>VIT\_201s0011g03540.1  
Sequence —●CAGGTAGCGTGAAGTTAAACCCTAGCGCTTTAGACAAACAGCTGTAGTCACCGCCCACAAACACC  
●AGCCTCTGAGACACCACCTCAAACCTTTCCACTTAAATACACATCCCTCACACCCTTTTCAATTC

Header —●VIT\_201s0011g03550.1  
Sequence —●CATGCAAAGCTGAACGCGATGCTGTGATTGGTGGTAAGTGGTAGTTGAGTAAATTTGACAGTGAA  
●GCCGAAATGGTAAAAGACTAAGGCTAGAAGTAGAATAACCACTGTTCTTCTCATCACGTGGGCCCCA

[https://www.researchgate.net/figure/A-sample-of-the-Multi-FASTA-file\\_fig1\\_309134977](https://www.researchgate.net/figure/A-sample-of-the-Multi-FASTA-file_fig1_309134977)



# Técnicas para Projetar Algoritmos de Bioinformática

## 1. Pesquisa Exaustiva:

- Consiste em tentar uma a uma **todas as soluções possíveis** e retornar uma das soluções candidatas satisfatórias, normalmente a primeira encontrada;
- Uma vez que uma solução é devolvida, os candidatos restantes (se houverem) **são descartados**;
- Se o problema admite exatamente uma solução, descartar os demais candidatos que não podem ser a solução permite **evitar perda de tempo**.



# Técnicas para Projetar Algoritmos de Bioinformática

## 1. Pesquisa Exaustiva:

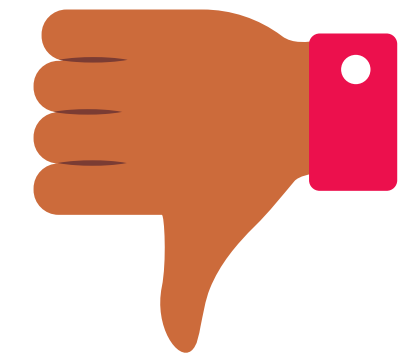
- Quais as DESVANTAGENS da pesquisa exaustiva?
- Quais as VANTAGENS da pesquisa exaustiva?



# Técnicas para Projetar Algoritmos de Bioinformática

## 1. Pesquisa Exaustiva:

- Quais as DESVANTAGENS da pesquisa exaustiva?



1. Enumerar todas as possíveis soluções candidatas pode ser difícil e/ou caro computacionalmente falando;
2. Custo da busca é **proporcional ao número de soluções candidatas** (tornando essa abordagem inadequada para grandes problemas);
3. A **complexidade computacional do pior caso é exponencial** na entrada, tornando essa abordagem inadequada para muitos tipos diferentes de problemas.

# Técnicas para Projetar Algoritmos de Bioinformática

## 1. Pesquisa Exaustiva:

- Quais as VANTAGENS da pesquisa exaustiva?



1. Se o problema tiver uma **única solução**, a **Pesquisa Exaustiva** a achará;
2. Usada para problemas de **pequenos portes**;
3. Pode ser usada como um **teste complementar** útil para verificar se os resultados relatados por algoritmos mais eficientes estão de fato corretos (utilizando pequenos *inputs*, por exemplo).



# Técnicas para Projetar Algoritmos de Bioinformática

## 1. Pesquisa Exaustiva:

- Quais as VANTAGENS da pesquisa exaustiva?



### 1. Usos:

- a. Para **entender o contexto** de um problema e qual seria a melhor solução para este;
- b. Em aplicações críticas onde qualquer erro no algoritmo teria consequências sérias;
- c. Quando estiver usando um computador para provar um teorema matemático.



# Explosão Combinatória

- **Números enormes!!!**

Se  $n = 16$  dígitos decimais  $\rightarrow 10^{15}$  instruções  $\rightarrow$  alguns dias

Se  $n = 19$  dígitos decimais  $\rightarrow 10^{18}$  instruções  $\rightarrow \sim 10$  anos

Se  $n = 10$  letras  $\rightarrow 10!$  instruções = 3.628.800 candidatos a considerar  $\rightarrow < 1$  seg

Se  $n = 20$  letras  $\rightarrow 20!$  instruções =  $2.4^{18}$  candidatos a considerar  $\rightarrow 10$ mil anos



# Exemplo das 8 rainhas de xadrez

## 1. Pesquisa Exaustiva:

- Como colocar 8 rainhas de xadrez sem que entre elas sofram nenhuma ameaça de morrer?

Fazer um **loop** sobre cada possível forma de arranjo até encontrar um arranjo que funcione:

64 quadrados (tamaho) e 8 rainhas (dimensão) ->  $64^8 = 281.474.976.710.656$  possibilidades!!!!

# Pergunta

**Como reduzir o tempo e o gasto computacionais?**

# Acelerando buscas por força bruta

## Reduzir o espaço de busca por métodos heurísticos

- Como melhorar?

**Exemplo:** 8 rainhas em 64 quadrados do tabuleiro de xadrez

$64^8 = 281.474.976.656$  possibilidades

- Todas as rainhas são iguais;
- Duas rainhas não podem ser colocadas no mesmo quadrado;

**Então:** 4.429.165.368 possibilidades

- Nenhum arranjo com 2 rainhas na mesma linha ou coluna:

$8! = 40.320$  possibilidades (1/100.000 da estimativa anterior e 1/7.000.000.000 da primeira estimativa) -> < 1/2 seg

De um problema intratável para um problema trivial.

<b>Força Bruta</b>	<b>Pesquisa Exaustiva</b>
Tamanho finito -> Busca por padrões	Soluciona problemas maiores como de permutação e combinatória
Mais consumo de tempo	Menos consumo de tempo
Pesquisa não-uniforme -> Não sabemos a direção da pesquisa	Pesquisa uniforme -> Sabemos a direção da pesquisa
Para buscas por strings 8 rainhas no xadrez	Para ir de um nó para outro Problema do vendedor viajante

# O exemplo do Restriction Mapping Problem (RMP)

- A técnica RMP surgiu antes mesmo do sequenciamento do primeiro genoma.
- **Objetivo:** determinar os pontos no DNA onde **endonucleases de restrição** (enzimas que cortam o DNA) fariam o(s) corte(s) (clivagem).





# O exemplo do Restriction Mapping Problem (RMP)

Letter to the editor

A map of the sites on bacteriophage PM2 DNA for the restriction endonucleases *Hind*III and *Hpa*II ☆

Christine Brack, Helen Eberle, Thomas A. Bickle, Robert Yuan

Show more ▾

+ Add to Mendeley 🔗 Share 📄 Cite

---

[https://doi.org/10.1016/0022-2836\(76\)90016-4](https://doi.org/10.1016/0022-2836(76)90016-4)

[Get rights and content](#)

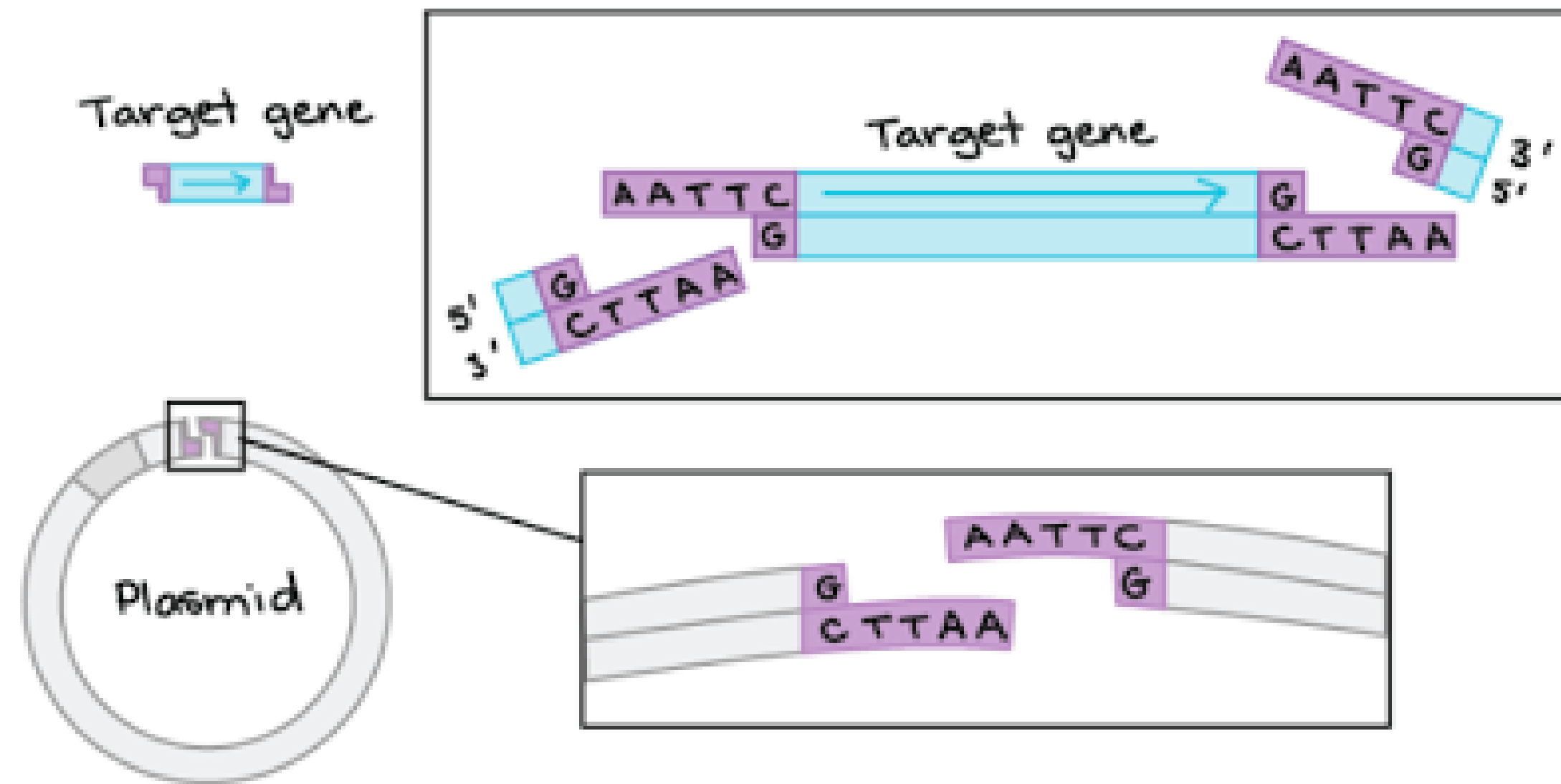
---

## Abstract

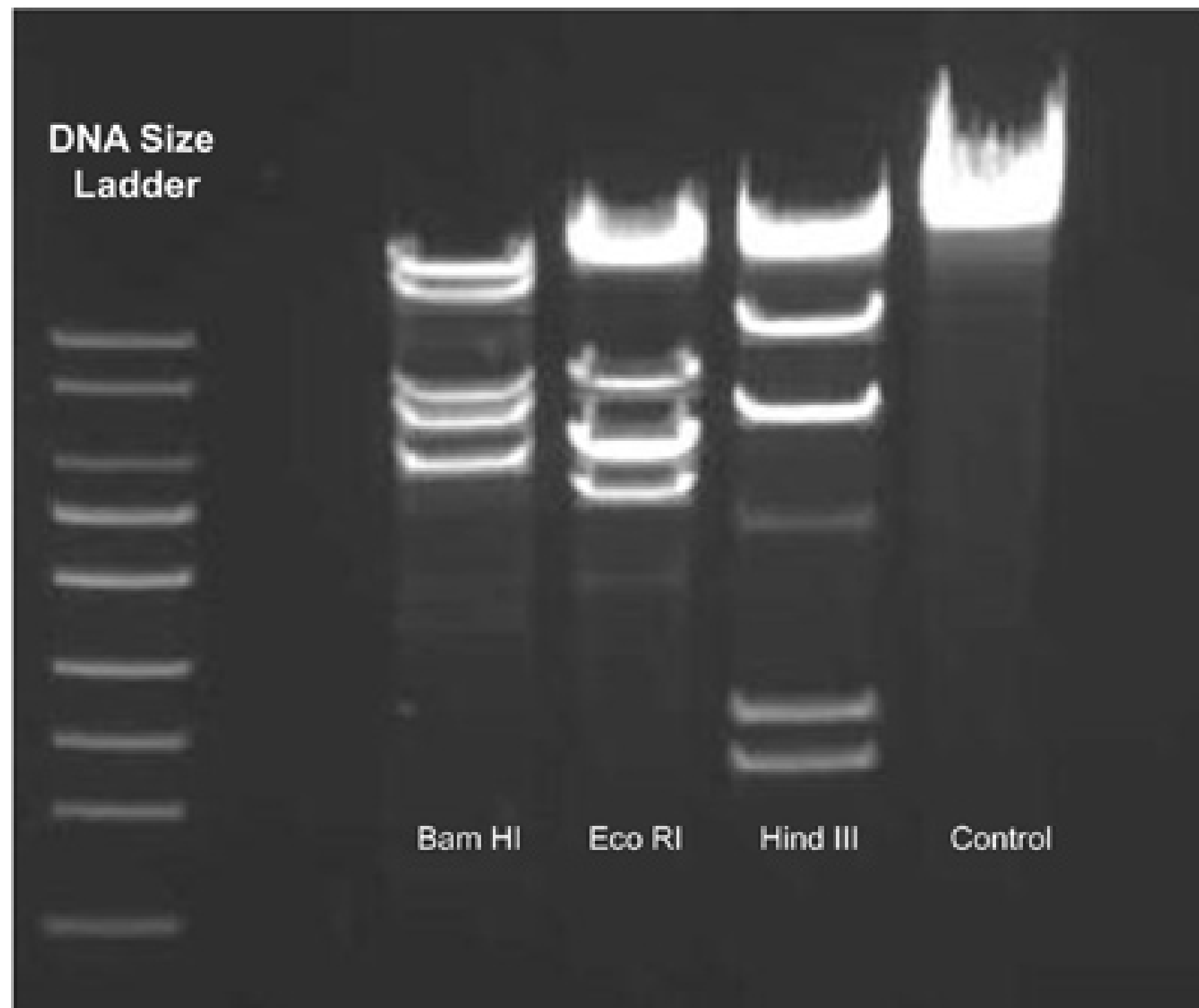
The map of the seven sites for the restriction endonuclease *Hind*III<sup>†</sup> and the single site for endo R.*Hpa*II on PM2 DNA was determined. This map was oriented with respect to the denaturation map of this DNA (Brack *et al.*, 1975) by partial denaturation mapping of the fragments. A new method for localizing restriction fragments by DNA-DNA hybridization and electron microscopy is described.

---

# O exemplo do Restriction Mapping Problem (RMP)



# O exemplo do Restriction Mapping Problem (RMP)



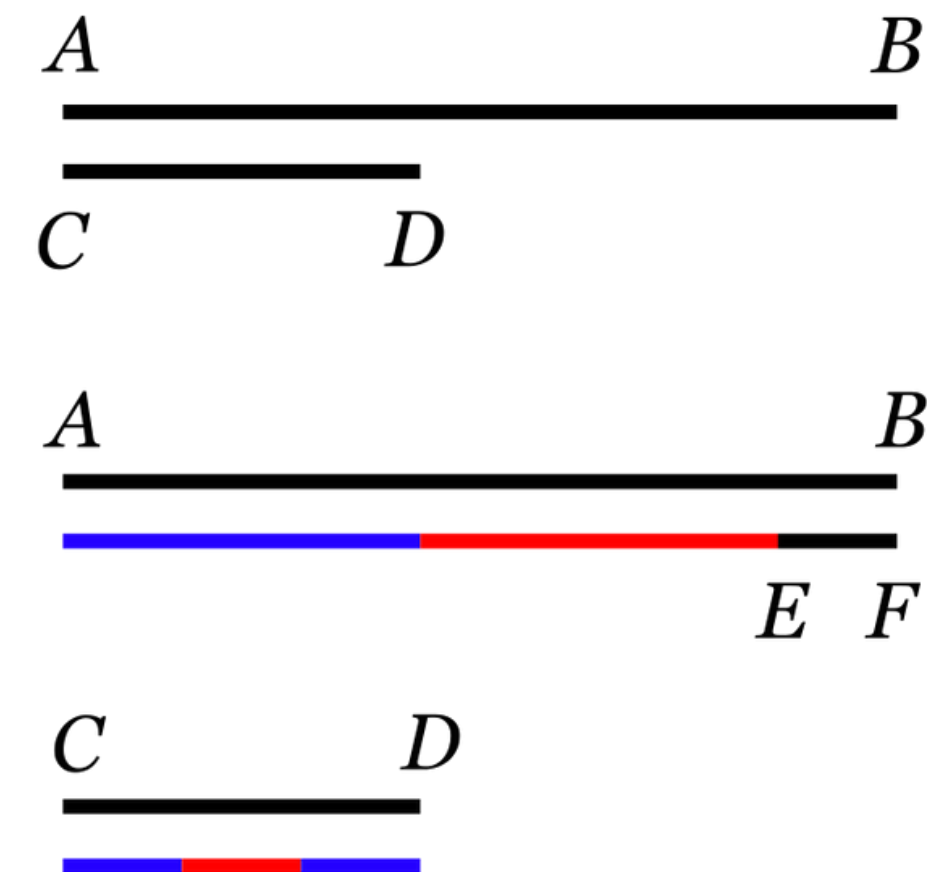
A **velocidade** com que uma molécula de DNA se move através de um gel de agarose durante o processo de eletroforese é **inversamente proporcional** ao seu **tamanho**.

Então, esta informação pode ser explorada para determinar as posições dos sítios de clivagem em uma molécula de DNA.

# O exemplo Restriction Mapping Problem (RMP)

- **Problema:** Dadas apenas distâncias pareadas entre um conjunto de pontos, o problema de mapeamento de restrição requer a recuperação das posições dos pontos.

## Algoritmo de Euclides



$EF$  es la mayor medida común

[Wikipedia Commons](#)

## Referências:

Encyclopedia of Bioinformatics and Computational Biology - ABC of Bioinformatics (2018).

<https://docs.optaplanner.org/6.2.0.CR2/optaplanner-docs/html/exhaustiveSearch.html>

<https://askanydifference.com/difference-between-brute-force-and-exhaustive-search/>

[https://www.wikiwand.com/pt/Algoritmo\\_de\\_Euclides](https://www.wikiwand.com/pt/Algoritmo_de_Euclides)

[https://www.wikiwand.com/pt/Busca\\_por\\_for%C3%A7a\\_bruta](https://www.wikiwand.com/pt/Busca_por_for%C3%A7a_bruta)

[https://www.youtube.com/watch?v=rTZsui\\_ftAQ](https://www.youtube.com/watch?v=rTZsui_ftAQ)

<https://www.youtube.com/watch?v=zv5Zm8xxMWA>

<https://www.youtube.com/watch?v=hrMgfwi75zY>