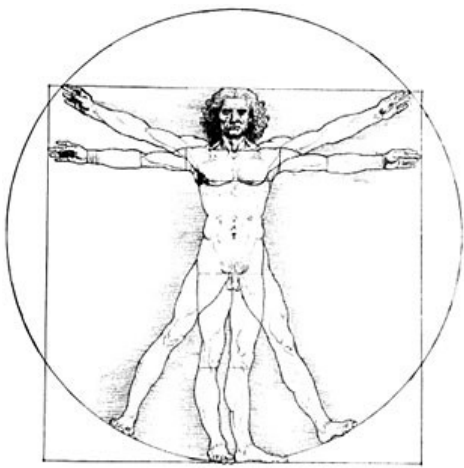


Projetos Genoma



Bianca Zingales

zingales@iq.usp.br

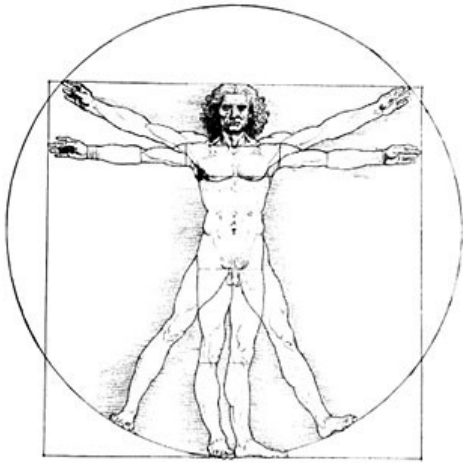


Projeto Genoma Humano - Histórico

1990 - O Projeto Genoma Humano foi lançado pelo **Departamento de Energia dos Estados Unidos** em parceria com os **Institutos Nacionais de Saúde** (National Institutes of Health, NIH) e cerca de 30 laboratórios de todo o mundo.

O Projeto foi concluído em 2003

Custo: US\$ 2,7 bilhões



Projeto Genoma Humano

Objetivos:

- Determinar a sequência de **3 bilhões** de pares de bases do DNA
- Criar **bancos de dados** para armazenar a informação
- Desenvolver ferramentas para a análise dos dados (**Bioinformática**)
- Transferir as tecnologias para o **setor privado**
- Discutir problemas **éticos, legais e sociais** que surgiriam com o projeto.

Estratégia para o Projeto Genoma Humano

Projeto Genoma Humano

Sequenciar o genoma de quem?

Sangue coletado de uma série de doadores anônimos (“pool”)

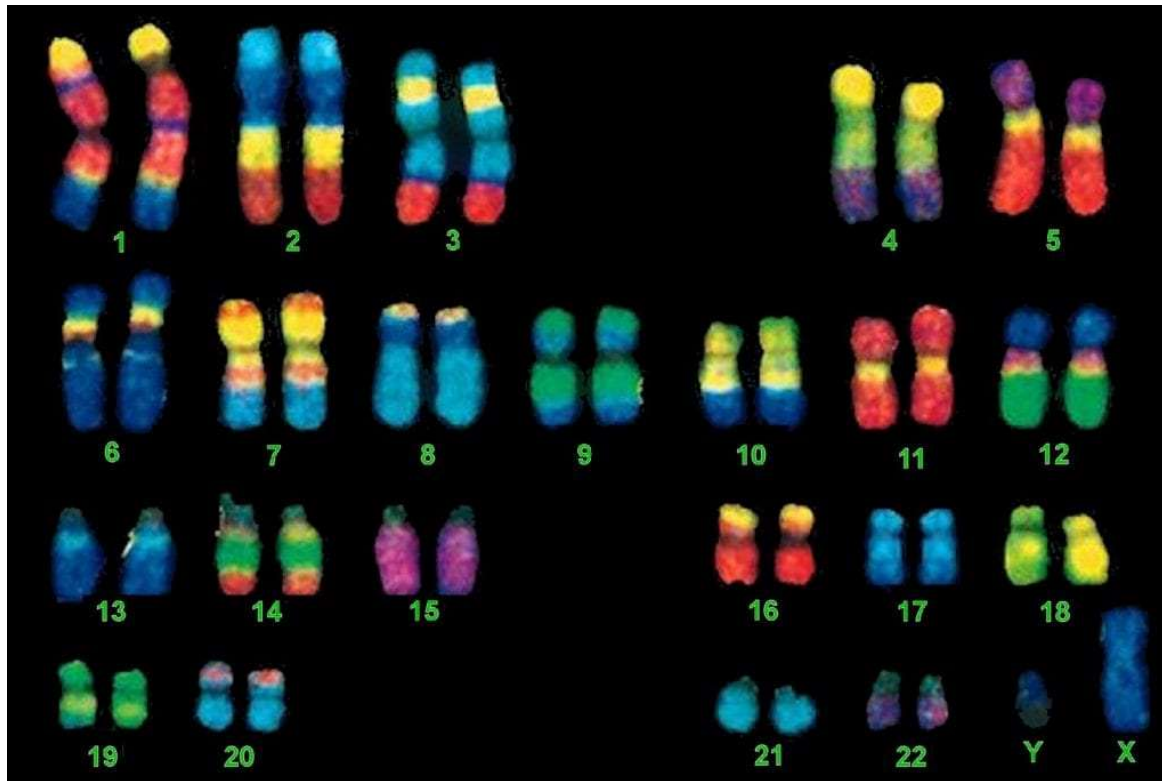


Sequenciamento do “pool” de DNA extraído das células do sangue

Projeto Genoma Humano

<https://www.genome.gov/human-genome-project/Completion-FAQ>

Identificação de **marcadores** conhecidos (genes) de cada cromossomo (23 pares de cromossomos)

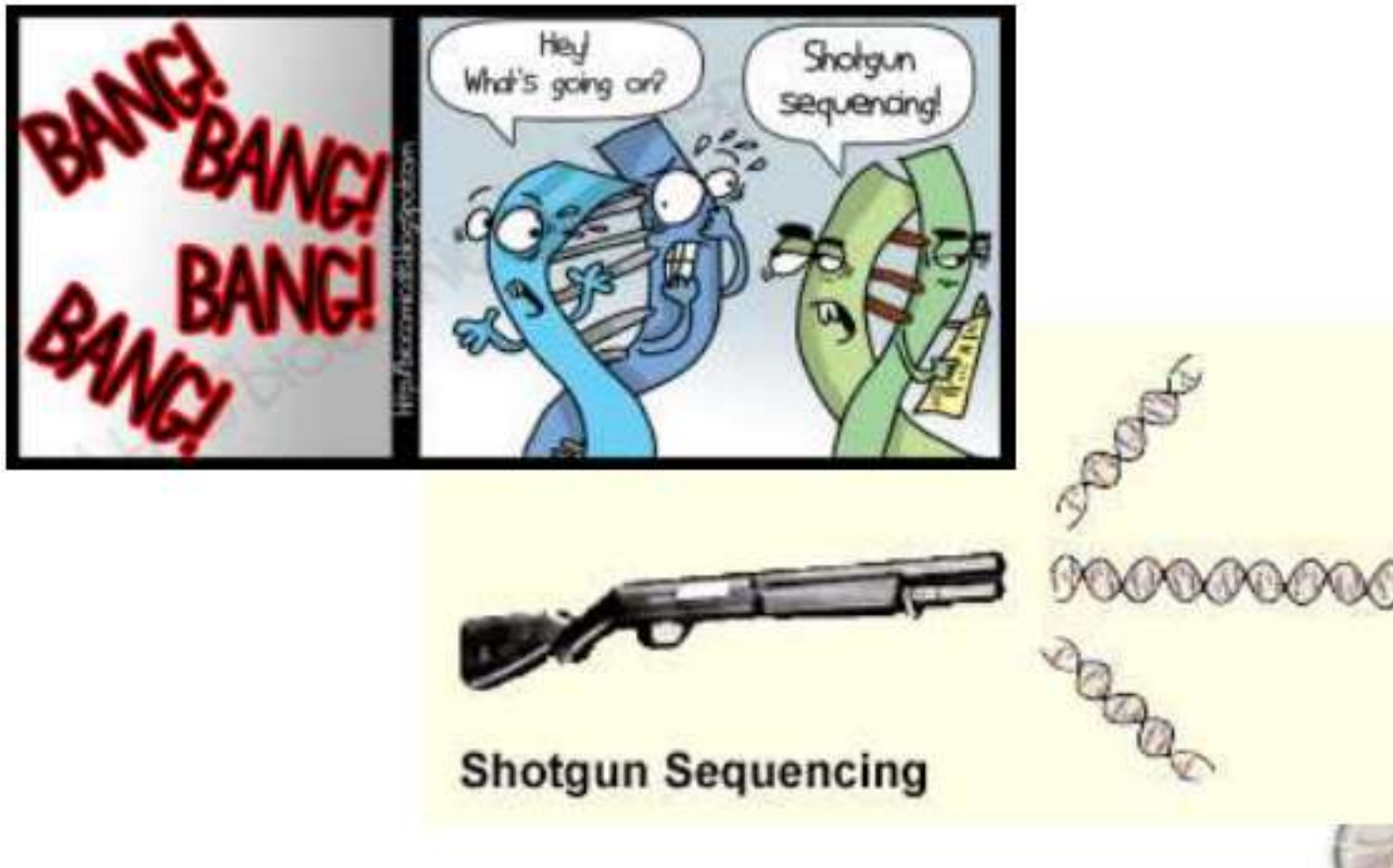


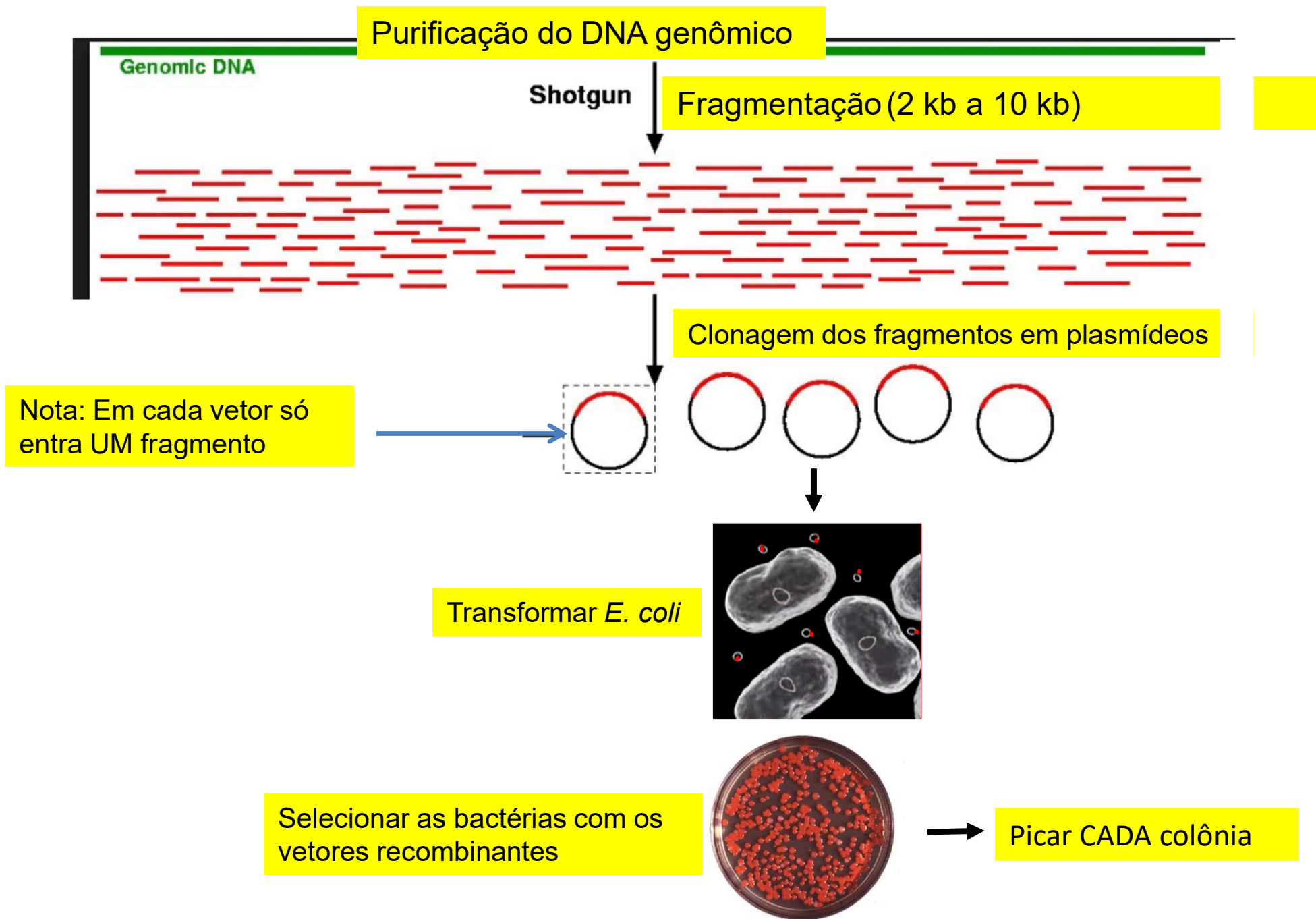
Os marcadores auxiliaram na montagem final do Genoma

Estratégias para Sequenciamento de Genomas

Sequenciamento de Genomas inteiros

Método de “shotgun”

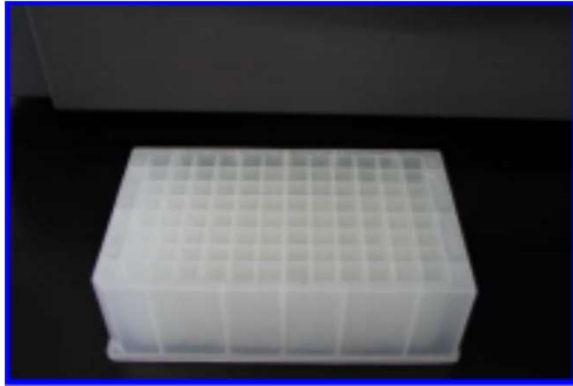




- Fragmentação por ultra-som; nebulizador; etc

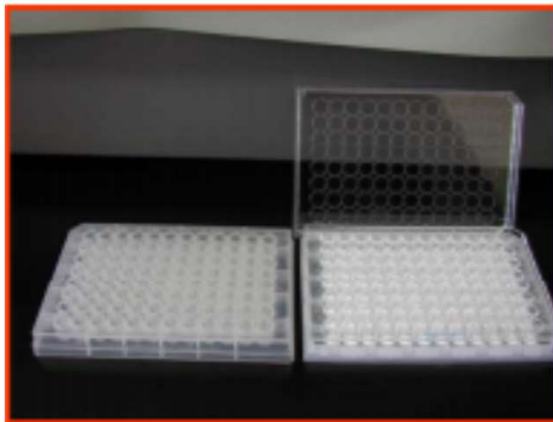
Placas para crescimento das colônias contendo os vetores recombinantes (uso do marcador de seleção!)

Cultivar cada colônia num poço com meio de cultura (+ AMP)



Placas deep well para crescimento dos clones de *E. coli*.

Meio LB Ampicilina. 12 a 16 h. 37°C. 150 rpm.



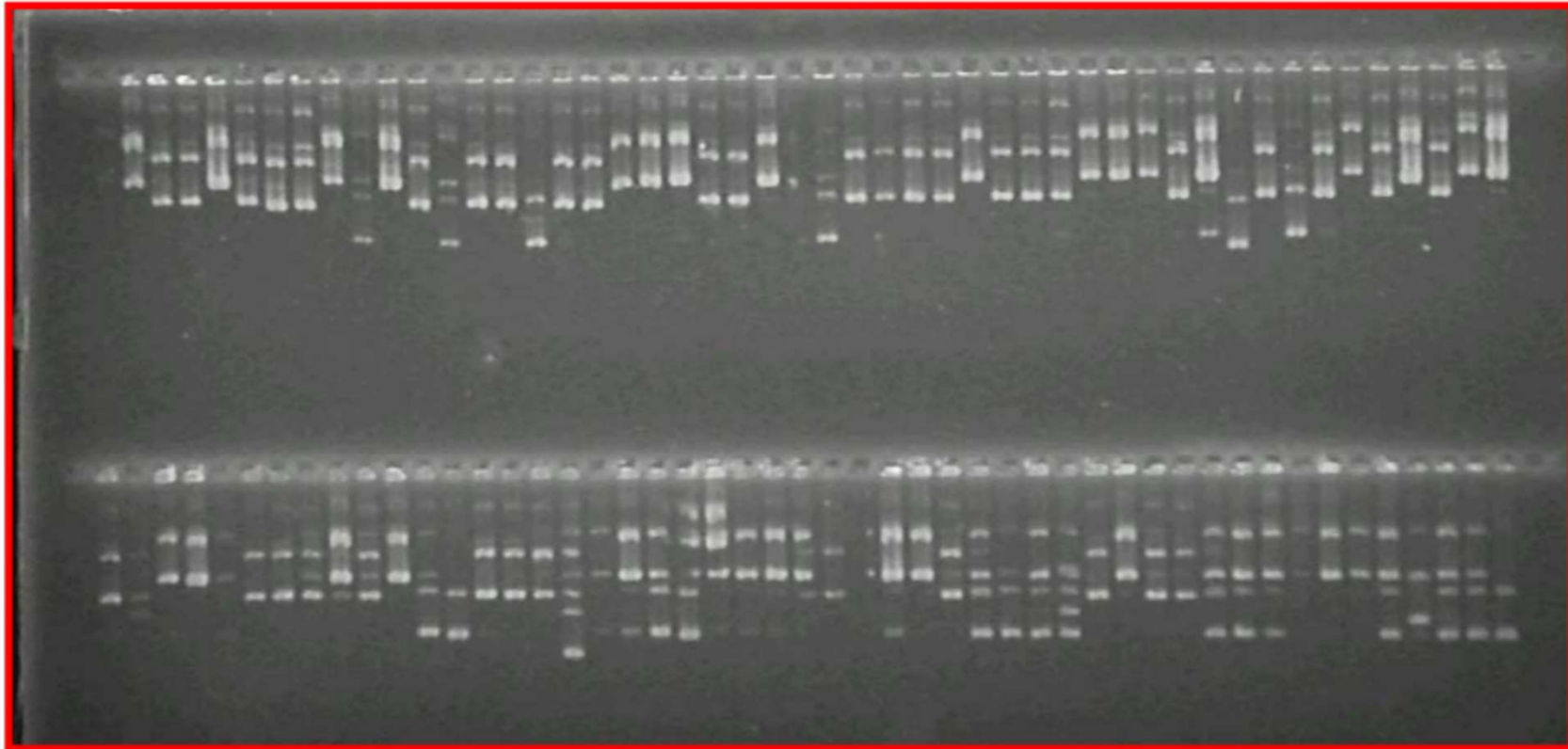
Purificação do DNA de cada vetor

Mini-prep - Lise alcalina

Placa de fundo U
Placa Millipore - filtro para purificar DNA plasmidial na ultima etapa da mini-prep

Eletroforese em gel de agarose

Análise do DNA dos vetores recombinantes



Sequenciamento “shotgun”

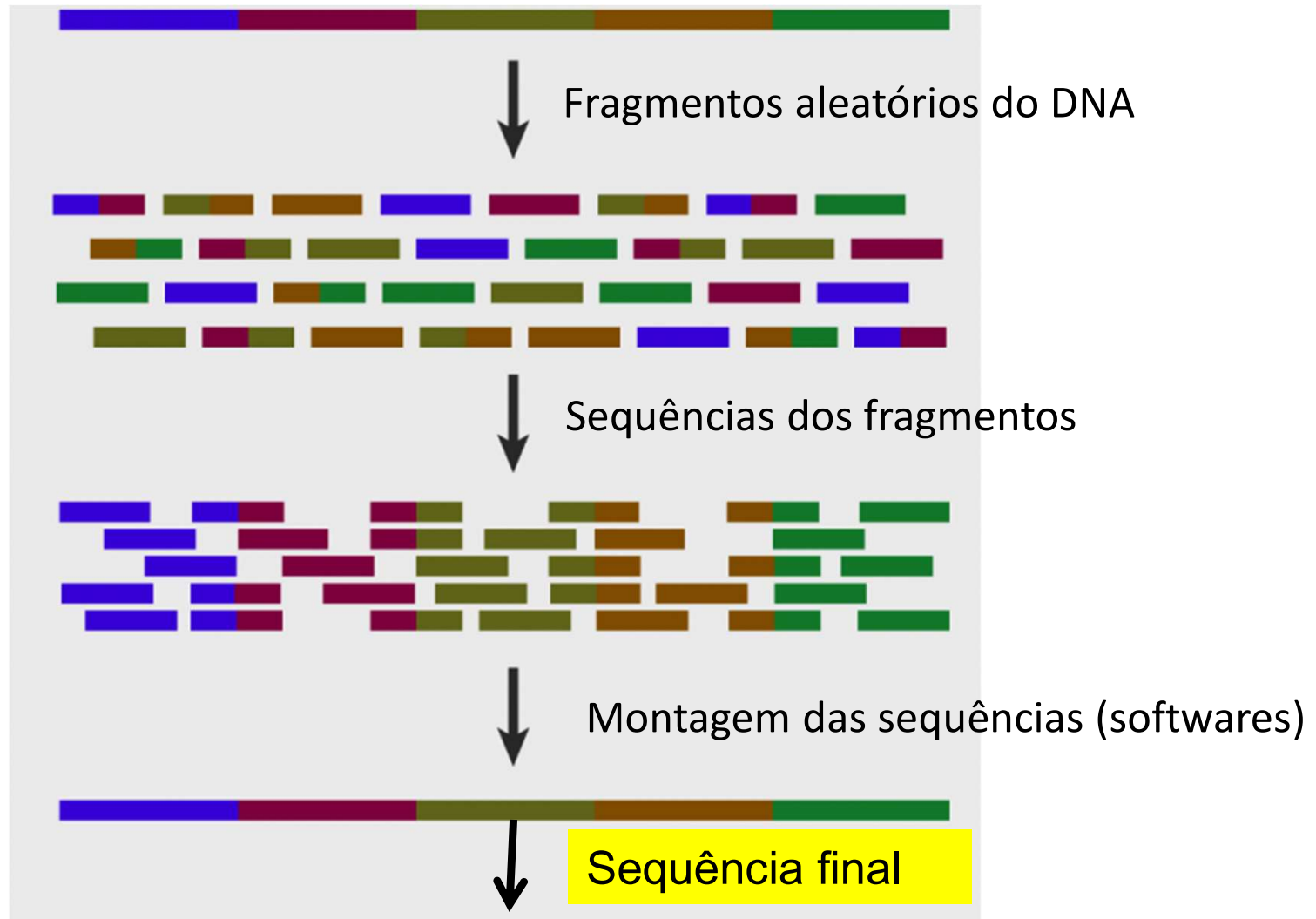
Sequenciar o inserto de cada vetor

Método de Sanger



Montagem das Sequências (Softwares - Ferramentas de Bioinformática)

Exemplo: DNA de um cromossomo



ACGATCGATCGATCGTAATTTATAGCATGCTAGCTACTGACGGGCTTTTACGGCGTTAGATATATATCGATCGATCGATGCTATATAGCGTGACTGATCGTAGCTGT
Genomic Sequence

Depósito da sequência em bancos de dados

Como sequenciar um genoma (pequeno)

Resumo

Etapas Gerais:

- Isolar o DNA do organismo
- Romper o DNA genômico aleatoriamente em muitos fragmentos (ultra-som; nebulizador; etc)
- **Clonar** cada fragmento em um **vetor**
- Determinar a **sequência** de cada fragmento
- Usar um programa computacional para **“juntar”** a sequência dos fragmentos (*Assembly*)
- **Depositar** a sequência em bancos de dados.

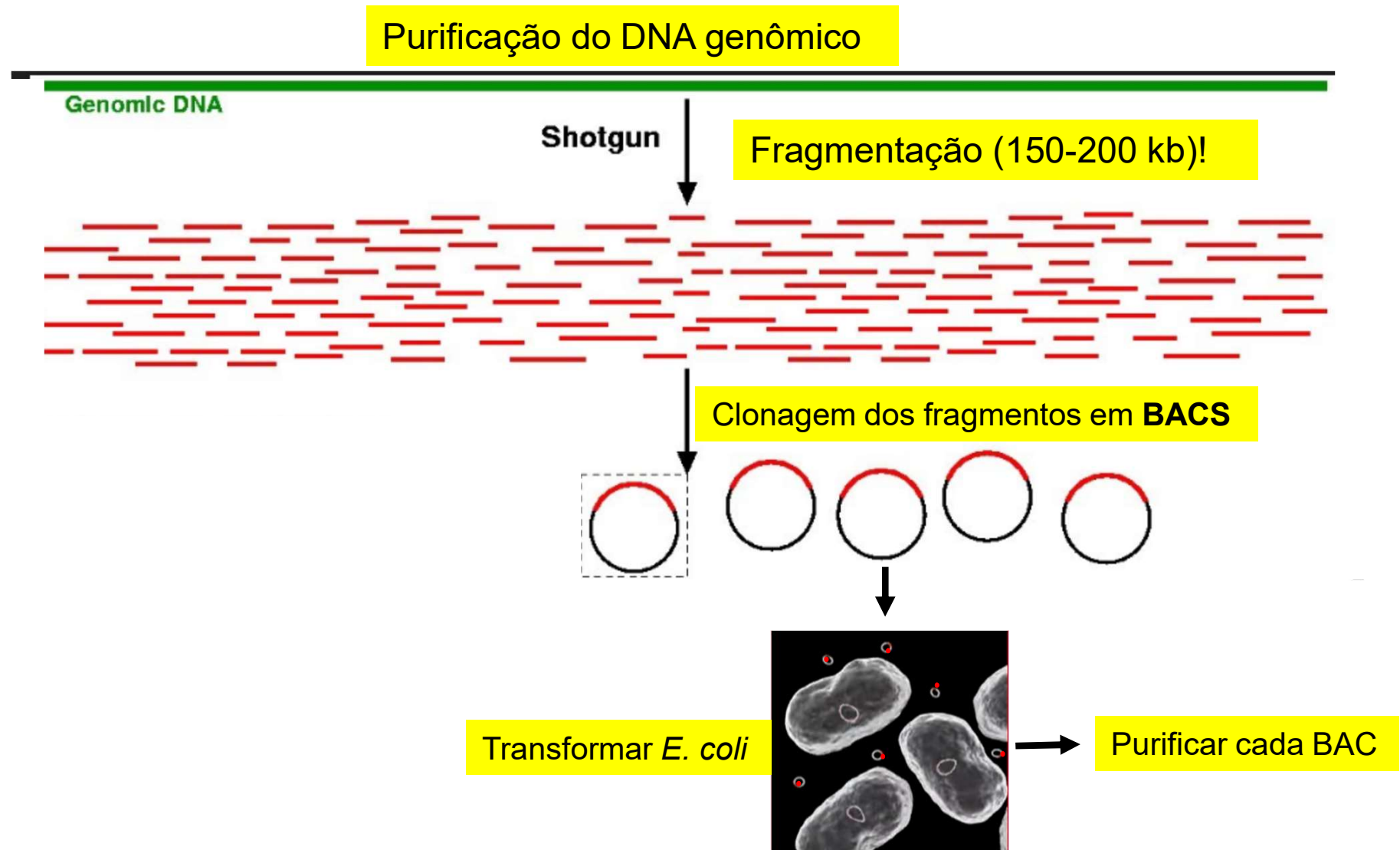
Para genomas grandes

Sequenciamento “shotgun” hierárquico

Hierarquia (definição): organização baseada na ordem de prioridade entre os elementos de um conjunto

Genoma Humano

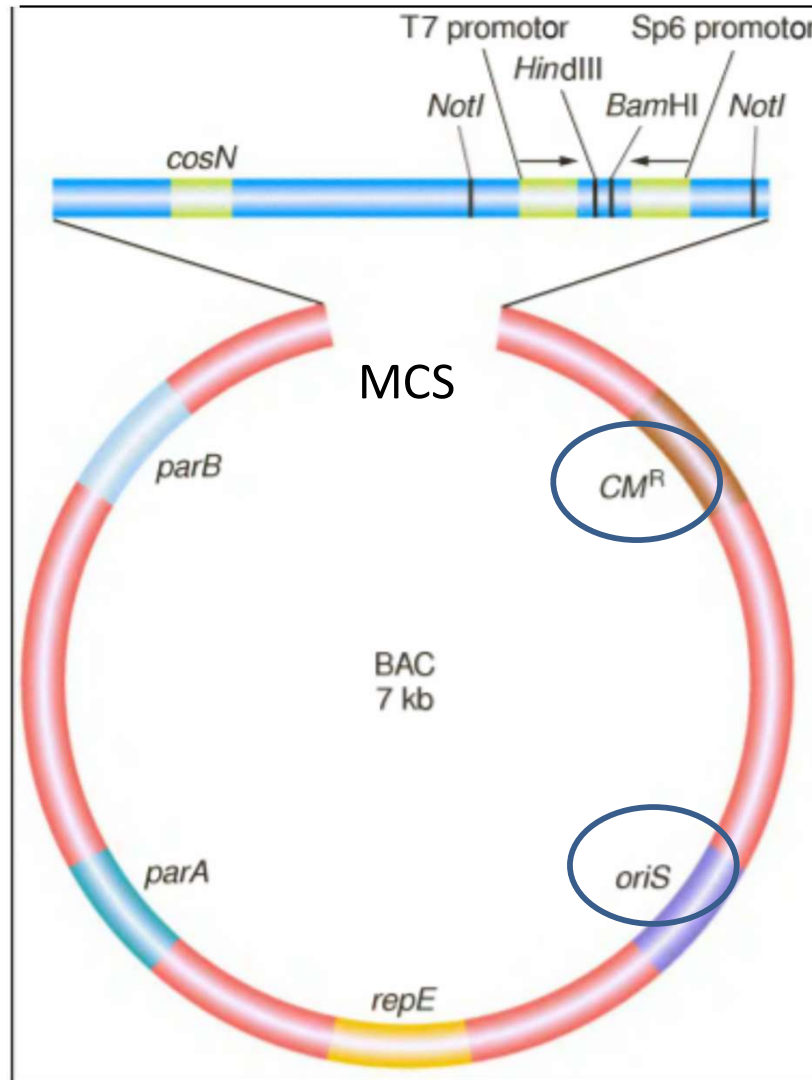
Sequenciamento “shotgun” hierárquico



Vetores BACs – Bacterial Artificial Chromosomes – permitem clonar insertos de 100.000 a 300.000 pb. Replicam-se em *E. coli*.

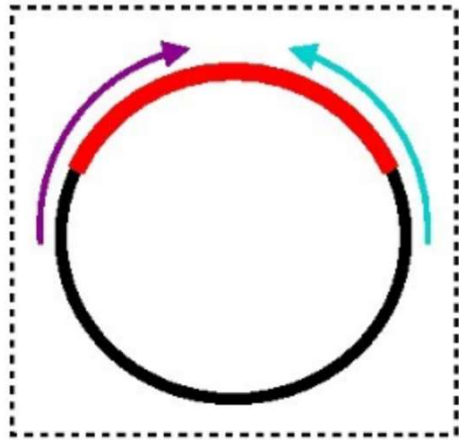
BAC

Inserto de 100 a 300 Kb

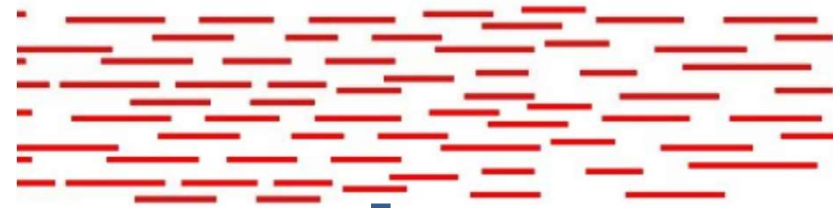


Sequenciamento de **cada extremidade de cada BAC**

DNA de BAC purificado



Fragmentação do inserto de **cada BAC**



Clonagem dos fragmentos de **CADA BAC** em plasmídeos



Sequenciar o inserto de **cada** plasmídeo



Montagem da sequência de **cada BAC**

As sequências dos BACs são reunidas in sílico com ferramentas de Bioinformática.

Resumo do Sequenciamento "shotgun" hierárquico

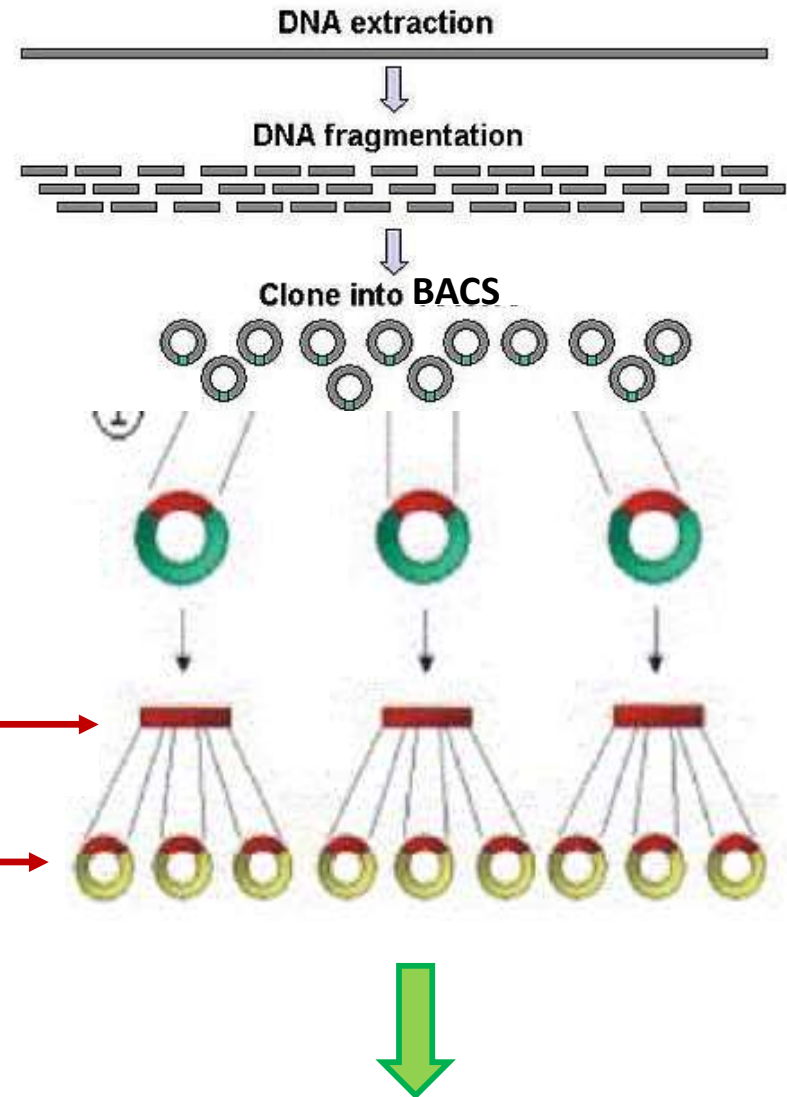
Fragmentos de DNA de 100.000 a 300.000 pb

Clonagem em BACs

Sequenciamento parcial dos BACs

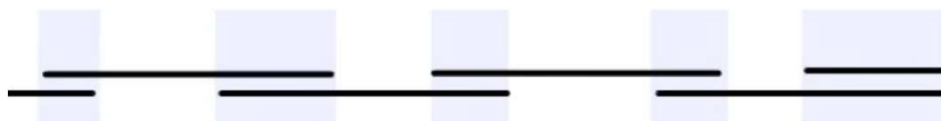
Fragmentar inserto de cada BAC

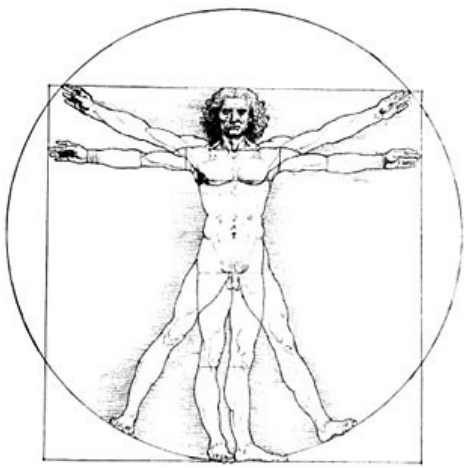
Clonar em plasmídeos



Sequenciar inserto de cada plasmídeo

Montagem das seqüências





Para o Genoma Humano

Os marcadores conhecidos (genes) de cada cromossomo auxiliaram a montagem final das seqüências de cada cromossomo



O que fazer com a sequência nucleotídica de 3 bilhões de bases?

```
. atgtcgtgct gccgagcggg ggtgaatgag ccagtaactc ccagctccgc ttcattcattg
. gagtctgacg accaaatcgc gtcgaagcag aagggtaacg aaccctcat cgaagactat
. agcatcatcg cccccggatt ctccatcaag ggctccgctg cgagttcga cgctgtggaa
. caaaataaga gctccgtgag tggacgcttt tctattcccg tctcatggca taatttggca
. tattcggcaa acggcacgaa aattctttgc ggccctcacag gaacagcgtt accgtcacga
. tgccttgctg tgatgggatc cagcgggtgc ggcaagacga cttttctcaa tgctatctct
. gaccgactta aaacctcgcg tacccttaag ctgacagggg aacgccagct gggggacttg
. gagtacaagc gtcattaccg caggatggtt ggttttgtgg cgcaagacga cattctctca
. ccacgggcaa caccgaaga ttcccttcgc ttttcgctgc gcgtgaggcg tggcacaagc
. ataagtgaag cgaataaatt tgttgaggaa actttggagg aattacgcct tgtccactgc
. cgtgagacca ttgttggcat ccctggcctt gtctctggtc tttcaggtgg tgaacgcaa
. cgcacaagta ttggagtgga gctcatttgc gatcctaaaa ttctcttgcg ggatgaacc
. acctctggtc tggactccgt gacatctgtg aagattgtgc atcttctgaa taacattgcc
. cgaacaggcc gcacgggtgat ttacaccatt caccagccca ctgctgagac attgacgtac
. tttgatgatc tcatgcttct cactgggggt cgatgtgctt accatggcac gatggcaaaa
. tctgtggaat actttgagtc catcggattc ccctgtcctg aacgatatac gccaaagcga
. ttctttatga agttgctcca agatccagaa attccaagg tactggttaa aaaatggaag
. agctatctaa aacacgggtg gagaacccca catacaaccg cggttgagct aaatcccaat
. ccctccgagt ctcccaccgc gaagaatatt gaaagctacc ttagtaggtt tgggagcacc
. tcgggtatcc aattccagga gctttttcgt cgtttttcca tagatctcag tcgcaatcat
. gtatacattt tttcacattt tatacaggct gccttctttg cagtaattgt gggctctcata
. tttctgaatg ttaaagatga tttagctggt atgcaagatc gcgagggagt ttttttatg
. gtaacgatga atcgggctat ggggcagact tttatcatgg tcaactcctt tatgcaagat
. aagcctttgt acgtgcggga gcaaatggtt ggctcactt ccccttttat tttcttttta
. tcaaaaacc tgggtggagt tccaatgcgc gtgttttttg cttttcttga gtgctgtatt
. ttatactgga tgggtgggtt ttaccgccag gcaggagctt tttttacta ctttgcggtc
. atcgcgctgc ttactgaagt ggccctcgggt cttgggtttg ccattggtgc cacgtttaa
. agtttggtcg ttgcttccgg taccgcgccc gtgattttgc tgccgcttgc catggtcggg
. ggtcttttgg cgaacacaga tcgacttcat ccgtattggt actggttggg gaaaccatcc
. tttattcgtc aggcctatat tcttcttgcg cgcaatgaat ttaagcatat cgaccacatt
. cgggtgtgat gtagaggcaa accaccgggc tactgtaaag ataagccca aaacggcgag
. gatattctgc gccaaacttg gtttcagcag aagcaatatg aaagctggat tttgtggcta
. actcttgcgc ttttatatat tgctttccgc ggttgggccc ttatttccct gtactctgcc
. gcgcgtacaa agtttttag
```

ANOTAR

Anotação

A sequência completa de um genoma sozinha tem pouco valor se não for submetida a análises adicionais.

Anotação do genoma é o processo de associar às sequências alguma informação biológica.

Etapas da anotação:

- Identificar elementos no genoma com potencial codificador de proteínas e RNAs
- Identificar porções do genoma que não codificam proteínas
- Associar a esses elementos informação biológica

Uso de ferramentas de Bioinformática

Características do Genoma Humano

- ❑ Tamanho **~3 bilhões** de pb
- ❑ Cerca de **30.000 genes codificadores de proteínas**
- ❑ Perfazem de **1,5 a 2% do genoma total**

- ❑ **98% do genoma constituído por DNA não codificante (ncDNA)**
- ❑ **99% do genoma não apresenta diferenças entre indivíduos**

DNA não codificante de proteínas do Genoma Humano (98% do genoma)

Constituído por:

- Espécies de rRNA e tRNA
- Pseudogenes – cópias inativas de genes codificadores de proteínas
- Introns e regiões não traduzidas do mRNA (26%)
- Sequências de DNA regulatórias da expressão gênica
- Elementos móveis (transposons)
- **Sequências de DNA repetitivo (50%) (Ex: microssatélites)**

Complexidade humana

A complexidade humana não deriva do **número de genes**

Deriva de:

- combinações de partes dos genes para obtenção de **diferentes produtos (splicing alternativo do mRNA)**;
- **modificações químicas** dos seus componentes;
- **processos de regulação** da expressão gênica e de modificações pós-traducionais.

Mapa genético do macaco bonobo é 98,7% igual ao humano



Será possível?

Evolução do custo do sequenciamento de um Genoma Humano

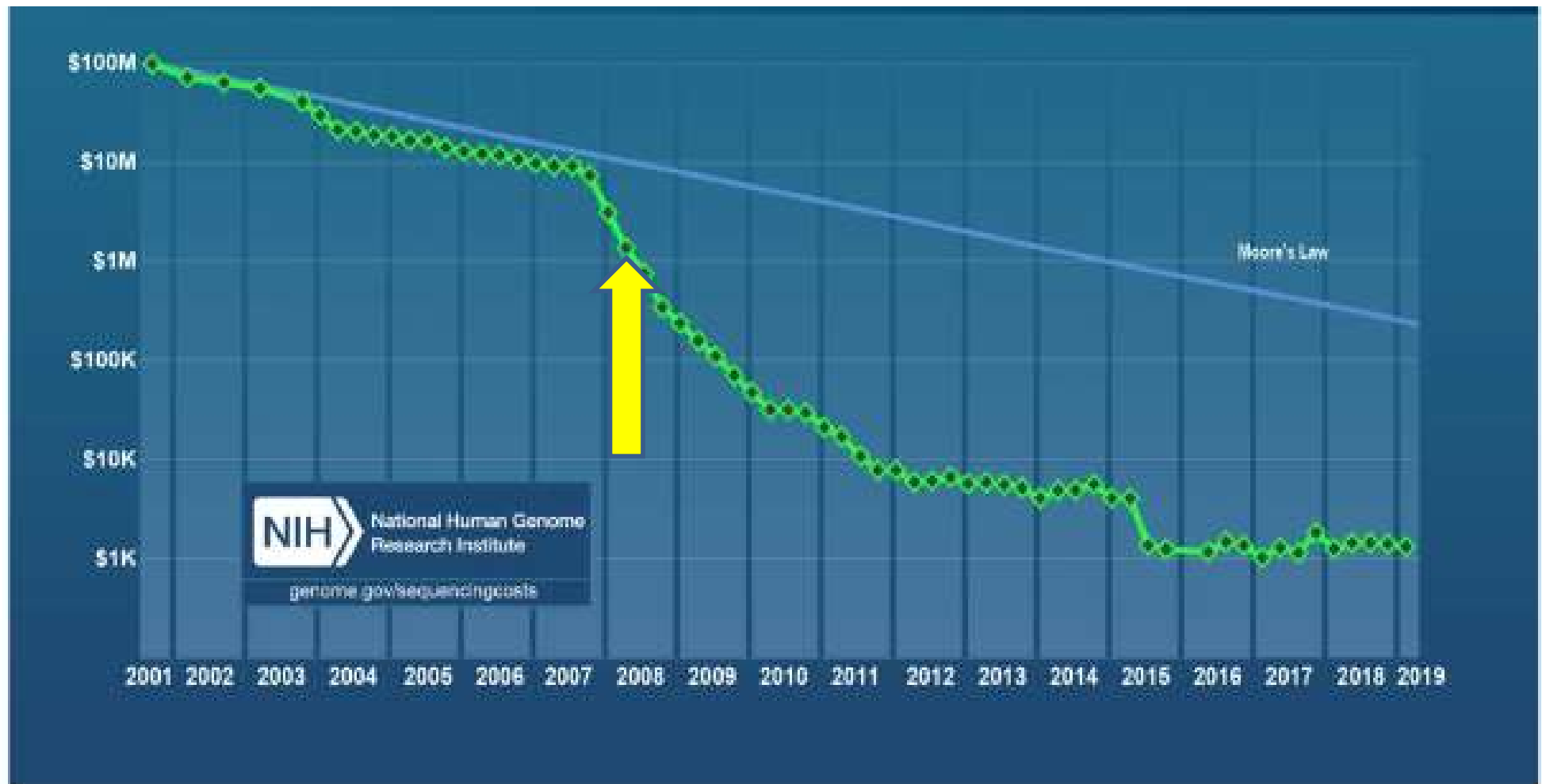
1990 - 2003 O custo total de produzir o genoma humano por um consórcio internacional foi de **US\$3 bilhões** ; **levou 13 anos**

Em 2006 o rascunho do genoma do chimpanzé custou **US\$22 milhões**

Em 2009 o sequenciamento do genoma humano custaria US\$1 a 10 milhões

2020 Atualmente as novas tecnologias permitem sequenciar o genoma humano por até **US\$1.200**, em **uma semana**

Evolução do custo do sequenciamento de um genoma humano



Mudança dos métodos de sequenciamento

<http://www.genome.gov/sequencingcosts/>

Dados de 2001 a 2007 sequenciamento pelo método de Sanger (*first generation sequencing platforms*).

A partir de 2008, uso de diferentes plataformas de sequenciamento



National Center for
Biotechnology Information

Os dados dos genomas de todos os organismos estão depositados em Bancos de Dados **Públicos**

The National Center for Biotechnology Information (NCBI)

garante o acesso a toda a informação biomédica e genômica

<https://www.ncbi.nlm.nih.gov/>

Aplicações de Dados do Projeto Genoma Humano

- Doenças: Entendimento, diagnóstico e prognóstico
- Desenho de fármacos para um grupo de pacientes (farmacogenômica)
- Terapia gênica
- Identificação de indivíduos
- Pesquisa básica (expressão gênica, mutações, RNAs não codificadores, etc.)

2023 - Pangenoma

Sequências genéticas quase completas de 47 homens e mulheres: afro-americanos, caribenhos, sul-americanos, leste da Ásia e oeste da África.

O mapa do Genoma Humano **renovado** representa uma ferramenta valiosa para identificar **variações genéticas** associadas a doenças e assim promover esquemas de tratamento.

Exemplo

Prognóstico de Doenças

BRCA2 (breast cancer type 2 susceptibility protein)





BRCA2 (breast cancer type 2 susceptibility protein)

BRCA2 - Proteína envolvida no reparo de quebras de DNA dupla fita.

Gene *BRCA2* - pertence à família de supressores de tumor.
Contém 27 exons

cDNA (obtido a partir do mRNA maduro) = 10.254 pb
Proteína = 3.418 amino ácidos

BRCA2 “mutada” não repara o DNA e aumenta o risco de câncer de mama, ovário etc.

O sequenciamento do gene *BRCA2* em diferentes indivíduos mostrou mais de 800 mutações (inserção ou deleção de nucleotídeos)

<https://www.cancer.gov/about-cancer/causes-prevention/genetics/brca-fact-sheet>



Mutações em BRCA2

Tipo de mutação - varia com o grupo das populações

cDNA = 10.254 pb

Population or subgroup

Mutation(s) Position and Type

→ Posição

Ashkenazi Jewish

6174delT

Dutch

5579insA

Finns

8555T>G, 999del5, IVS23-2A>G

French Canadians

8765delAG

Germans

C61G

Hungarians

9326insA

Icelandics

999del5

Italians

8765delAG

Northern Irish

6503delTT

Pakistanis

3337C>T

Scottish

6503delTT

Slovenians

IVS16-2A>G

Spanish

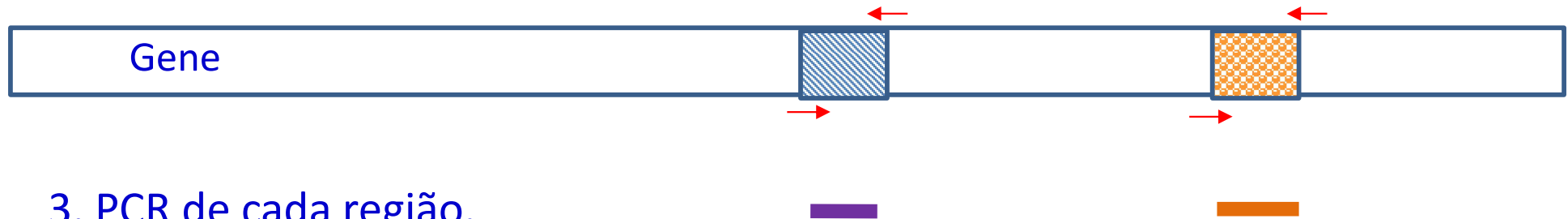
3034delAAAC(codon936), 9254del5

Swedish

4486delG

TESTE para Mutações em BRCA2

1. DNA extraído do sangue ou saliva
2. Desenho de Primers flanqueando **exons** que apresentam mutações



3. PCR de cada região.

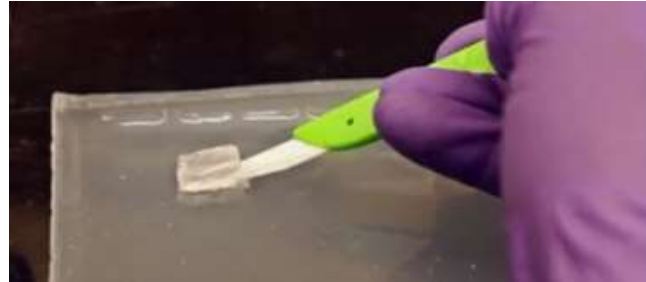
4. Eletroforese em gel de agarose

<http://www.cancer.gov/about-cancer/causes-prevention/genetics/brca-fact-sheet>

<https://www.einstein.br/especialidades/oncologia/exames-tratamentos/testes-brca1-brca2>

TESTE para Mutações em BRCA2

5. Remoção de cada produto da PCR do gel de agarose



6. Purificação do DNA de cada produto da PCR

7. **Sequenciamento de cada produto da PCR**

8. **Comparar** a sequência obtida com a sequência da mesma região do **gene normal. (A sequência do gene normal está no banco de dados).**

9. Laudo: Há ou não há mutações.

10. Se houver mutações há **Probabilidade/Propensão** para câncer de mama e outros tipos de câncer.

11. A/O paciente decide o que fazer....

TESTE para Mutações em BRCA2 (cont.)

Custo nos USA: 1 sítio ~US\$400 ; vários sítios ~ \$3,000

No Brasil: varia de R\$ 3 mil a R\$ 5 mil.

<https://www.uai.com.br/app/noticia/saude/2015/09/25/noticias-saude,187014/cancer-teste-genetico-pode-salvar-vidas.shtml>

Problemas éticos, legais e sociais

Ética

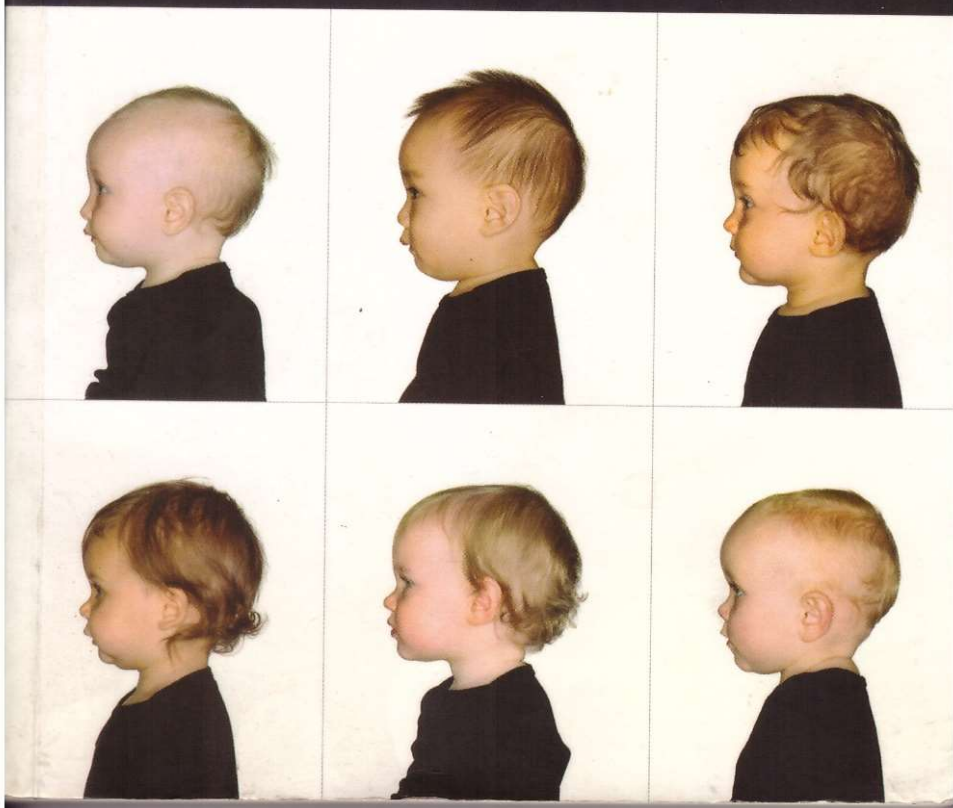
<https://www.scielo.br/pdf/spp/v14n3/9771.pdf>

Como usar os dados originados do Projeto Genoma Humano?

MAYANA ZATZ

GENÉTICA

escolhas que nossos avós não faziam



SUMÁRIO

Prefácios

- Um futuro que já é presente, por Jorge Forbes 11
Direito em um mundo em transformação,
por Adriana Diaféria 15

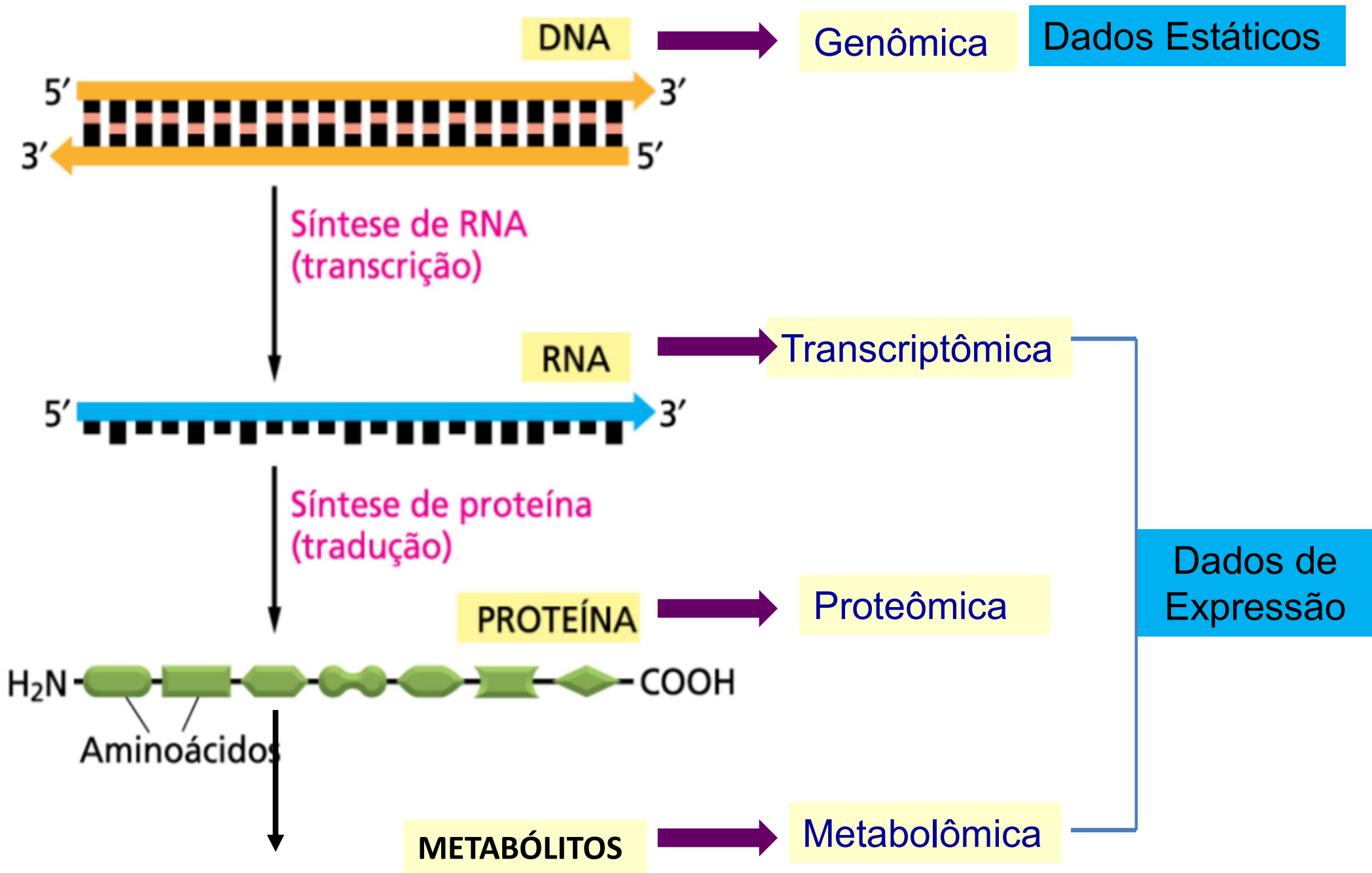
Introdução 21

1. Paternidade ou o direito de “não saber” 37
2. Confidencialidade: esses genes são meus 49
3. Diagnóstico pré-natal ou o que você vai ser quando crescer 61
4. Gêmeos, trigêmeos e até mais 75
5. Menino ou menina: o que você faria se pudesse escolher o sexo do bebê? 83
6. Embriões salvadores ou o motivo pelo qual eu nasci 93
7. Projeto Genoma ou o que esperar de nossos genes 103
8. Loira ou morena, alta ou baixa, atleta ou cientista: os genes fúteis 119

O mundo das “ÔMICAS”

https://edisciplinas.usp.br/pluginfile.php/4449003/mod_resource/content/1/Aula%207%20-%20ESTUDOS%20DAS%20%27MICAS%20GEN%27MICA%20VS%20TRANSCRIPT%27MAS%20E%20METAGEN%27MICA.pdf

O mundo das “ÔMICAS”



Transcriptoma

Objetivos

- identificar **todos** os RNAs expressos (transcritos) em um organismo ou tecido **numa determinada condição**.

Abordagem: Sequenciamento

Transcriptoma permite avaliar QUAIS são os genes diferencialmente expressos e as diferenças QUANTITATIVAS de expressão

Aplicação: Identificação de marcadores de tumor e câncer.
Permitem o diagnóstico e prognóstico

Proteômica

Análise global das **proteínas** expressas em uma amostra biológica (organismo, tecido ou célula) em determinada condição

- Análise de cada proteína por Espectrometria de Massas (MALDI-ToF)
- Dedução da sequência de aminoácidos de cada mancha
- **Comparação com dados do Genoma**

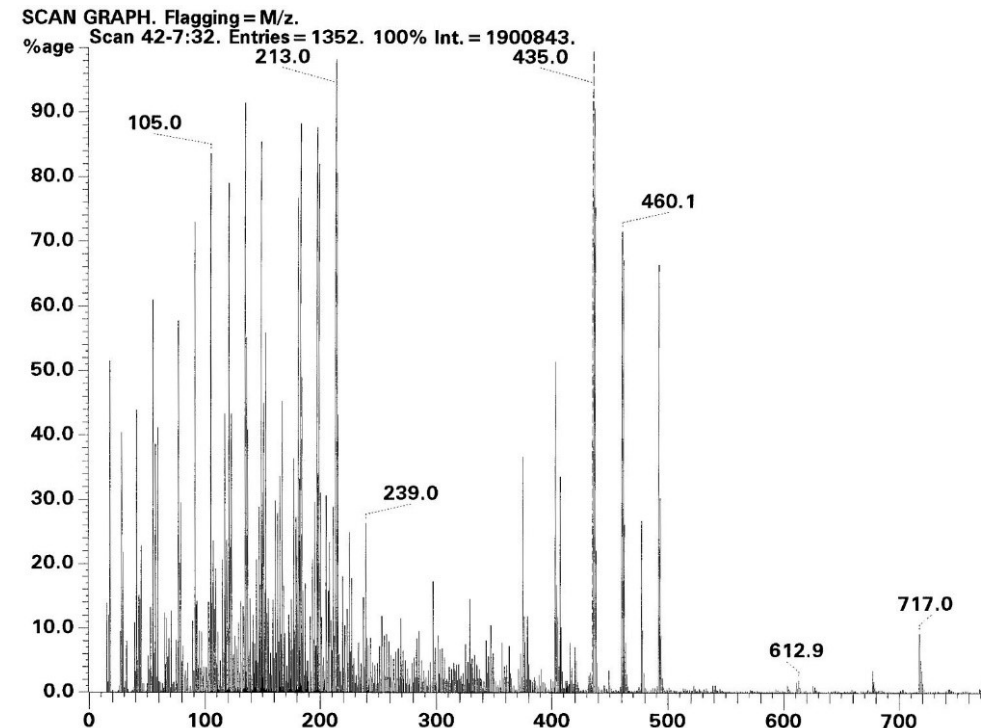
Exemplo: ANÁLISE PROTEÔMICA DE VARIEDADES DE FEIJÃO. Qual é mais nutritiva?

Metabolômica

Análise global dos **metabólitos** presentes em uma amostra biológica (organismo, tecido ou célula).

Técnicas

- Espectrometria de massa (GC-MS)
- Espectroscopia de ressonância magnética nuclear (RMN)



Exemplo: metabólitos presentes na célula APÓS a administração de um determinado fármaco

FIM DA DISCIPLINA

OBRIGADA PELA ATENÇÃO

BOA PROVA!!!!!!