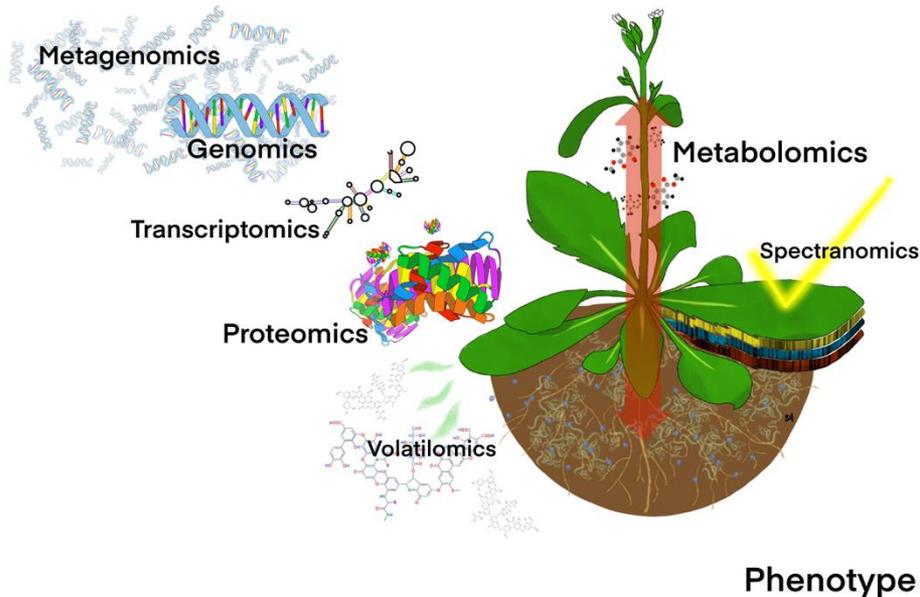


Estudo das Ômicas

Aula 7

LGN0232 – Genética Molecular

Genotype



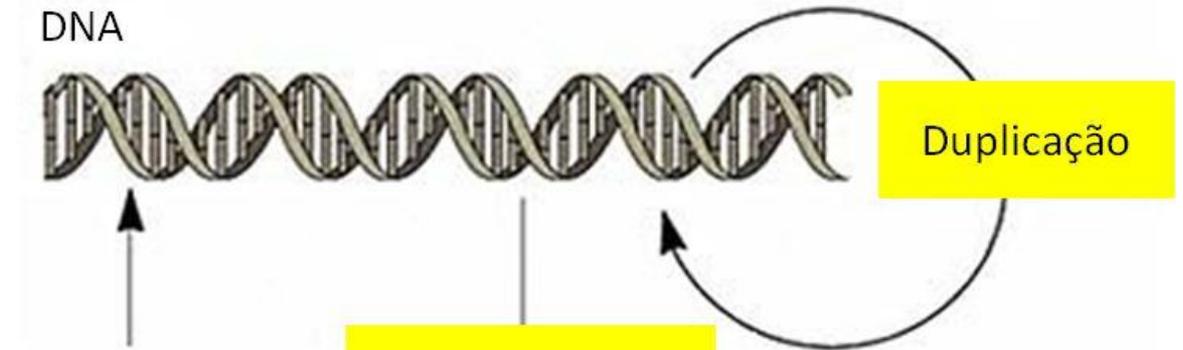
Ilara Budzinski
Departamento de Genética
ilara@usp.br

SUMÁRIO

- Fluxo da informação genética;
- Conceito das ômicas;
- Moléculas avaliadas nas ômicas;
- Aplicação dos conhecimentos adquiridos;
- Estudo de casos.
-

DOGMA DA BIOLOGIA CELULAR

Genoma



Transcriptase Reversa

Transcriptoma



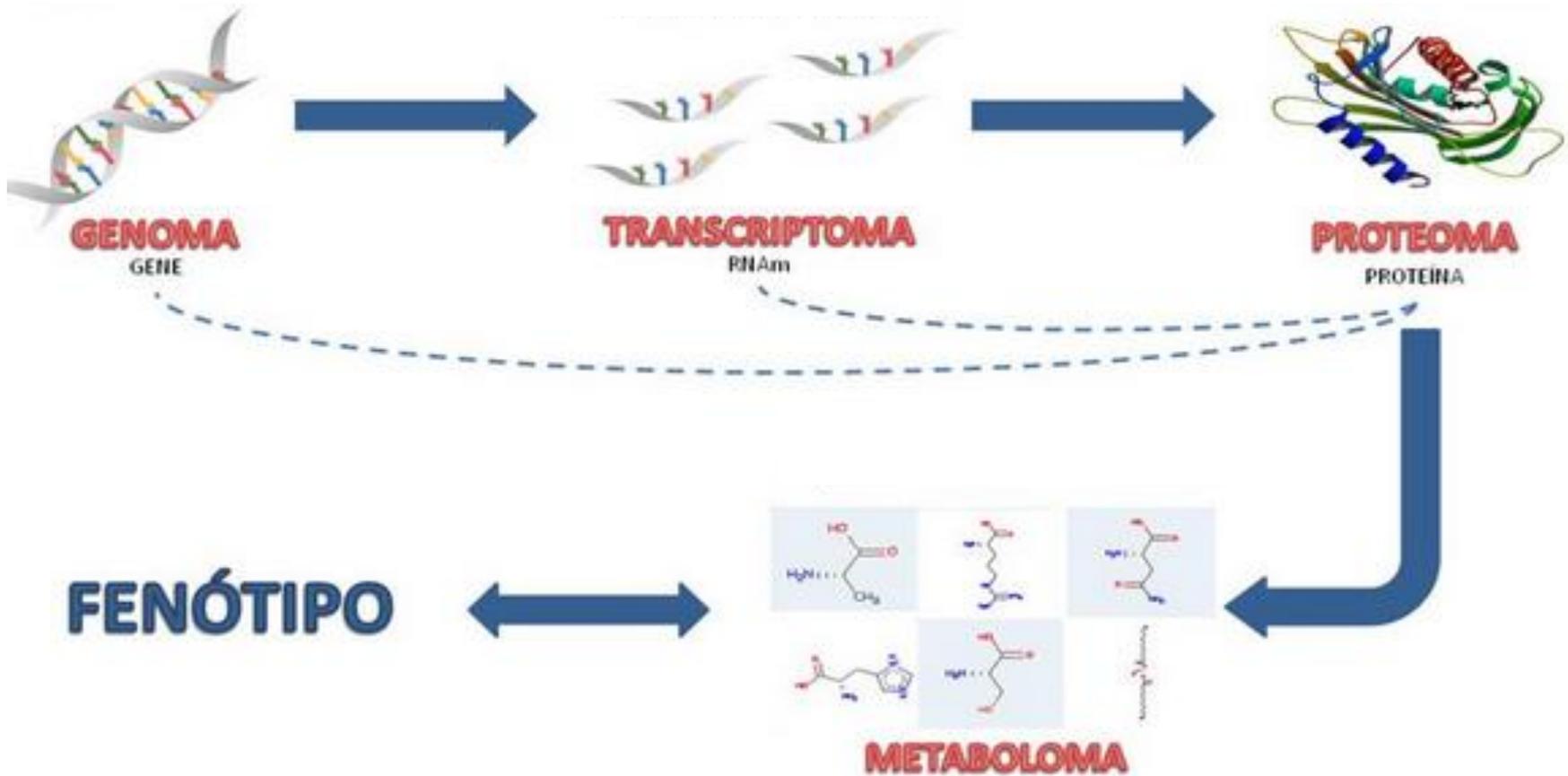
Tradução

Proteoma

Proteína

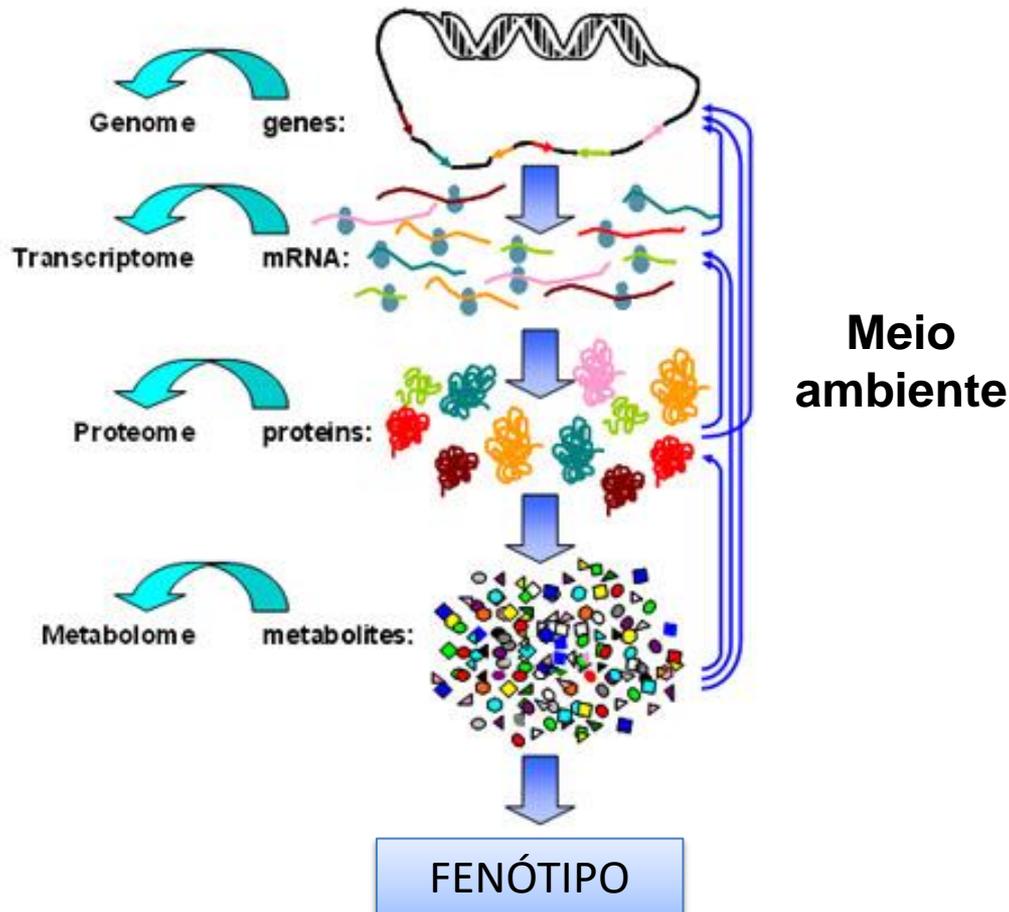


O FENÔMENO DAS ÔMICAS



E tem muito mais: epigenoma, lipidoma, interatoma, etc...

- ✓ Avanços tecnológicos dos últimos anos permitiram o surgimento de uma nova era nas pesquisas: **A Era das Ômicas.**



conjunto de todas as características observáveis
– que são influenciadas tanto por seu genótipo
quanto pelo ambiente

**Em um organismo
somente o genoma
permanece constante,
independente do estágio
de desenvolvimento,
tecido e ou condição
ambiental!**



Diferentes estímulos podem afetar diretamente o transcriptoma, o proteoma e o metaboloma.



DEFININDO ALGUNS CONCEITOS

Genoma: toda a informação hereditária de um organismo que está codificada em seu DNA (ou, em alguns vírus no RNA). Isto inclui tanto os genes como as sequências não-codificadoras.

Transcriptoma: conjunto completo de transcritos (RNAs mensageiros, RNAs ribossômicos, RNAs transportadores e os microRNAs) de um dado organismo, órgão, tecido ou linhagem celular. Portanto, ele é o reflexo direto da expressão dos genes.

Proteoma: conjunto de todas as proteínas em uma célula, organela fluido biológico, tecido ou organismo em um dado momento.

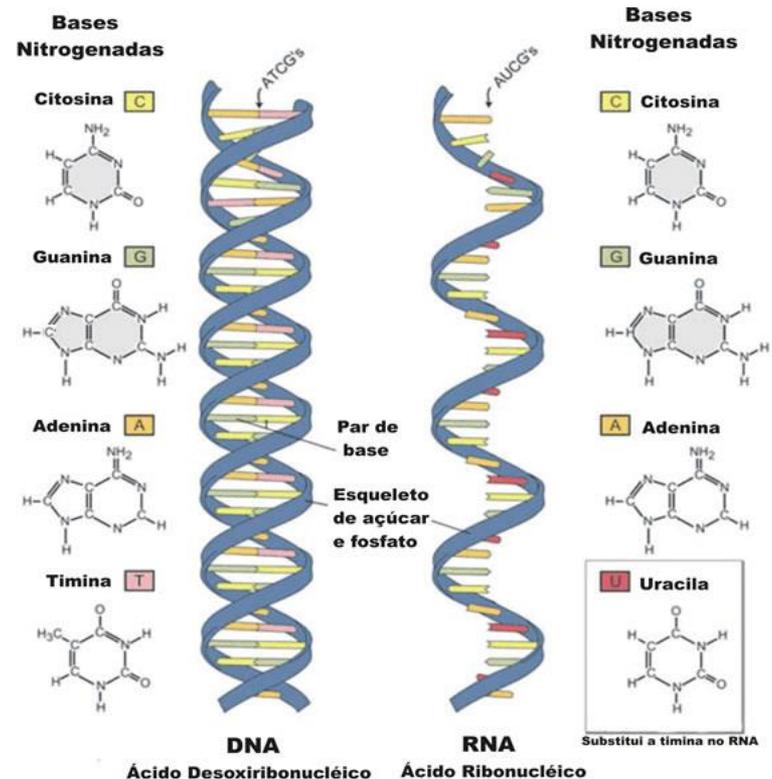
Metaboloma: conjunto de todos os metabólitos em uma célula, fluido biológico, tecido ou organismo.

DNA E RNA:

São ácidos nucleicos, encontrados em todas as células. Estão envolvidos na transmissão de **caracteres hereditários** e na produção de proteínas.

1- O que são?

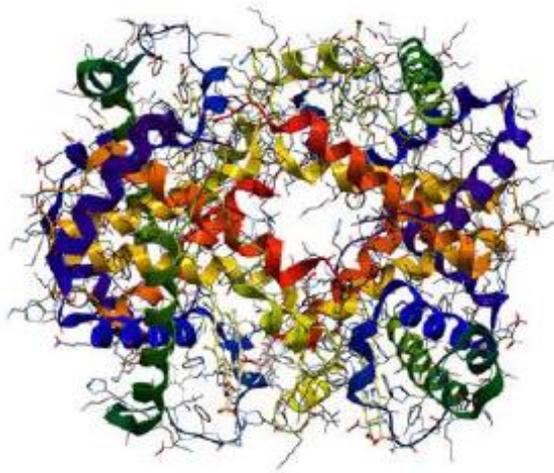
2- Qual a estrutura?



PROTEÍNAS:

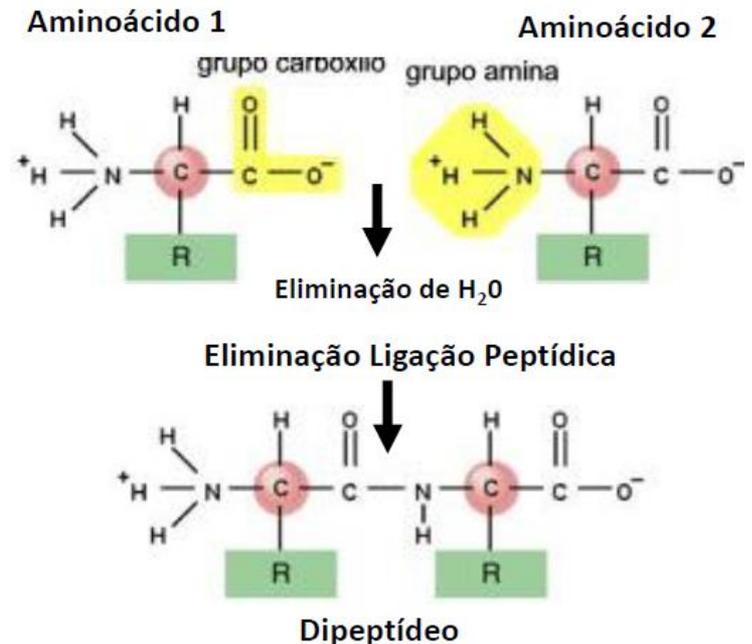
São moléculas de estrutura complexa e massa molecular elevada. São sintetizadas pelos organismos vivos através de ligações peptídicas covalentes entre aminoácidos.

Funções: Enzimas, anticorpos, componentes estruturais.



1- O que são?

2- Qual a estrutura?



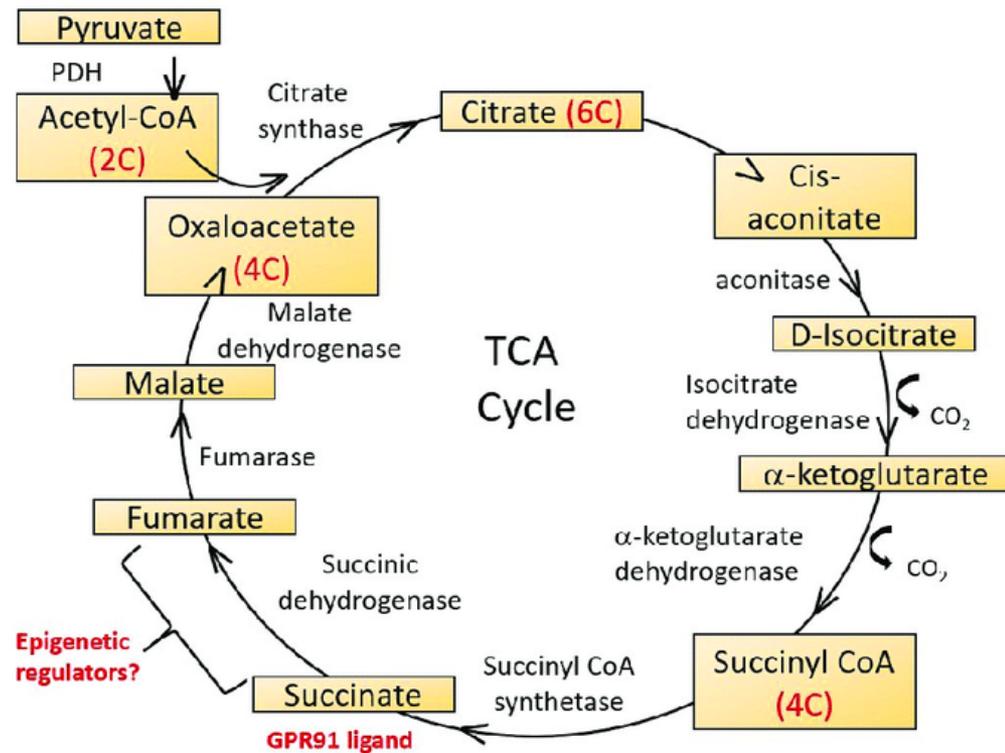
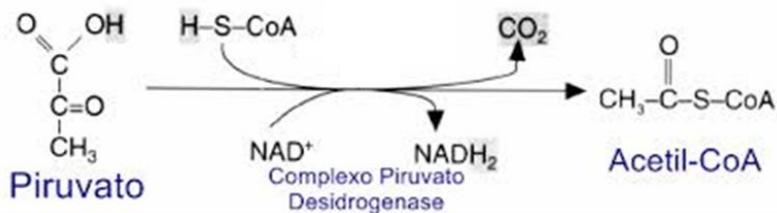
METABÓLITOS:

1- O que são?

2- Qual a estrutura?

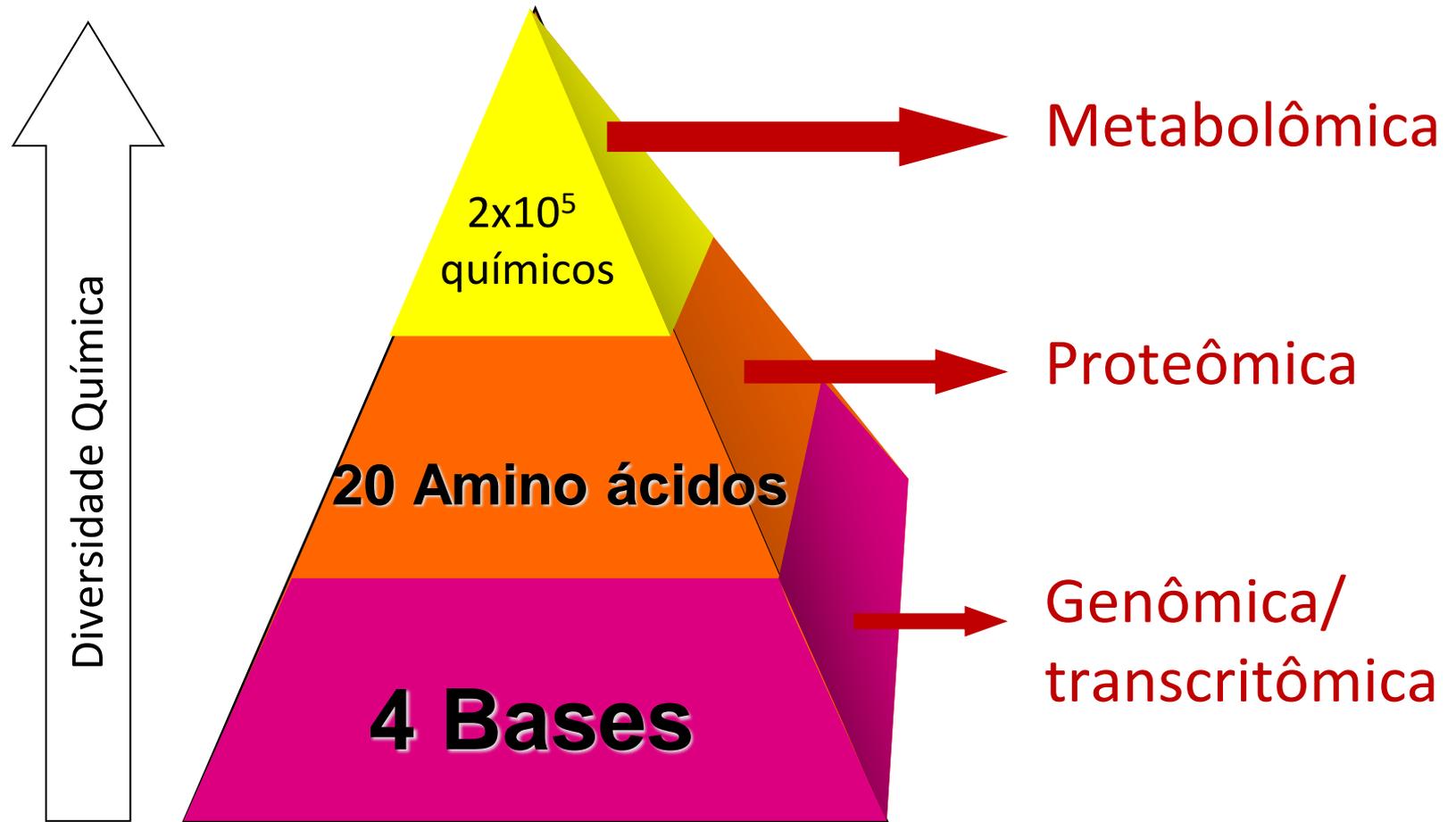
Metabólitos são os intermediários (substratos, cofatores) e produtos do metabolismo!

Carboidratos
Alcoois
Aminoácidos
Ácidos orgânicos
Lipídios



Ciclo do Ácido Cítrico

DIFICULDADE DOS ESTUDOS DAS “ÔMICAS”



A pirâmide da vida

GENÔMICA

- ✓ É o termo atribuído a todo e qualquer estudo do genoma (o material genético dos seres vivos). É constituído pelo DNA.
- ✓ Os estudos de genômica tem por objetivo entender como os genes e a informação genética estão organizados dentro do genoma e como essa organização determina sua função.

➤ PERSPECTIVAS

1. Catálogo da Informação Genômica: tamanho, conteúdo de GC, número de genes, conteúdo de DNA repetitivo e características únicas.

2. Catálogo da Informação Genômica Comparativa: quando uma espécie divergiu?

3. Princípios Biológicos: como os genes estão organizados no genoma? Quais são os mecanismos de evolução do genoma? Qual o papel das alterações epigenéticas?

GENÔMICA

➤ ESTRUTURA GERAL DO GENOMA EUCARIOTO

Genoma nuclear (alto grau empacotamento), mitocondrial (origem endossimbiótica) e cloroplastidial (origem endossimbiótica).

Pode apresentar grande variação no tamanho do genoma (milhares de pb) mas, a variação no número de genes é bem menor.

Presença de:

Ítrons,

Elementos de regulação

Sequências altamente repetitivas (não transcritas)

Sequências moderadamente repetitivas (transcritas)

Ex: transposons e retrotransposons

GENÔMICA

➤ TAMANHO E COMPLEXIDADE DO GENOMA

Table 1: Genome Size and Number of Protein-Coding Genes for a Select Handful of Species

Species and Common Name	Estimated Total Size of Genome (bp)*	Estimated Number of Protein-Encoding Genes*
<i>Saccharomyces cerevisiae</i> (unicellular budding yeast)	12 million	6,000
<i>Trichomonas vaginalis</i>	160 million	60,000
<i>Plasmodium falciparum</i> (unicellular malaria parasite)	23 million	5,000
<i>Caenorhabditis elegans</i> (nematode)	95.5 million	18,000
<i>Drosophila melanogaster</i> (fruit fly)	170 million	14,000
<i>Arabidopsis thaliana</i> (mustard; thale cress)	125 million	25,000
<i>Oryza sativa</i> (rice)	470 million	51,000
<i>Gallus gallus</i> (chicken)	1 billion	20,000-23,000
<i>Canis familiaris</i> (domestic dog)	2.4 billion	19,000
<i>Mus musculus</i> (laboratory mouse)	2.5 billion	30,000
<i>Homo sapiens</i> (human)	2.9 billion	20,000-25,000

GENÔMICA

✓ O QUE GERA VARIABILIDADE GENÉTICA?

Indels (Inserções/Deleções)

Inversões/Translocações

Variação no número de cópias (CNVs)

Regiões repetitivas em tandem (Mini/Microsatélites)

Polimorfismos de nucleotídeo único (SNP)

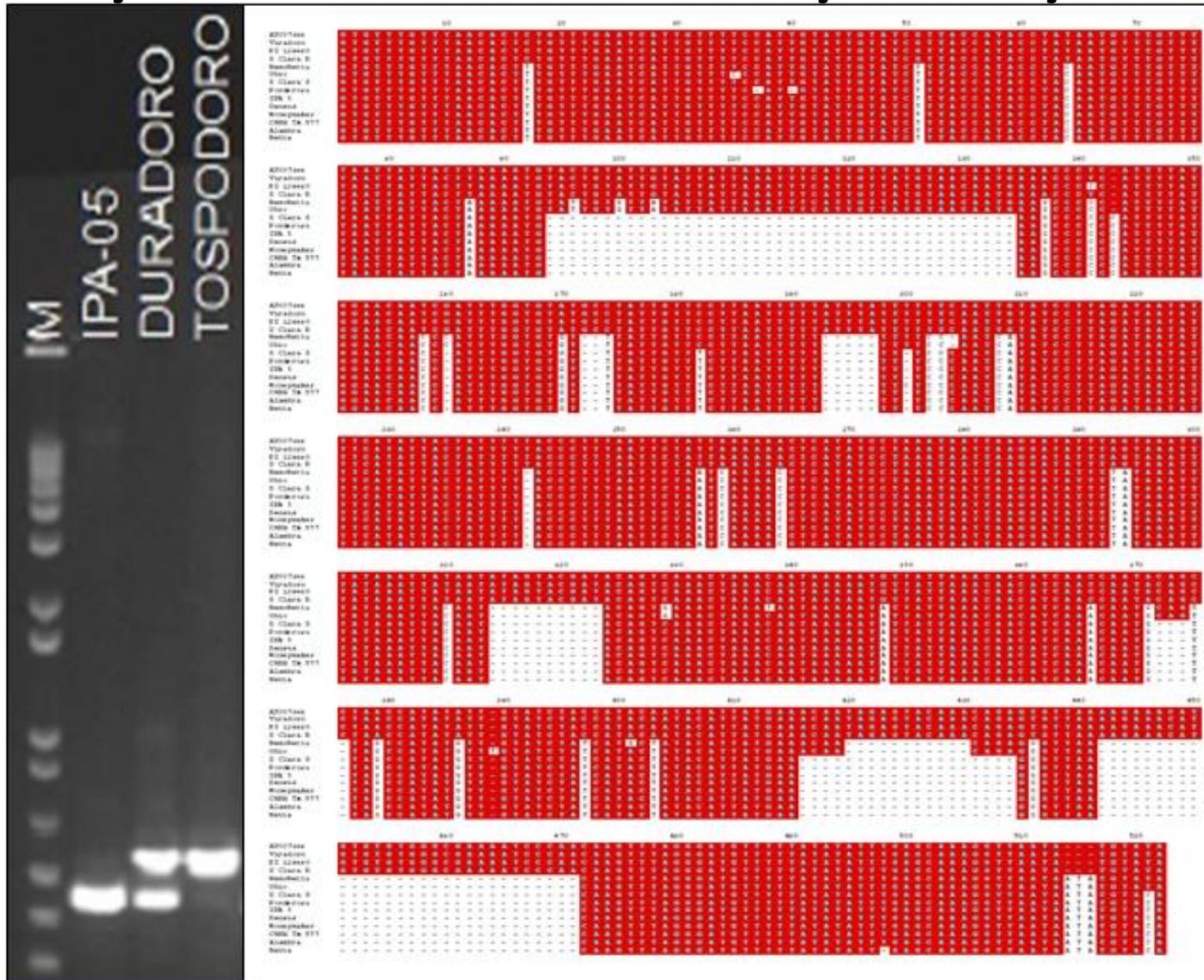
✓ O QUE É UM MARCADOR MOLECULAR?

São características do DNA que diferenciam dois ou mais indivíduos e são herdadas geneticamente (segregação mendeliana).

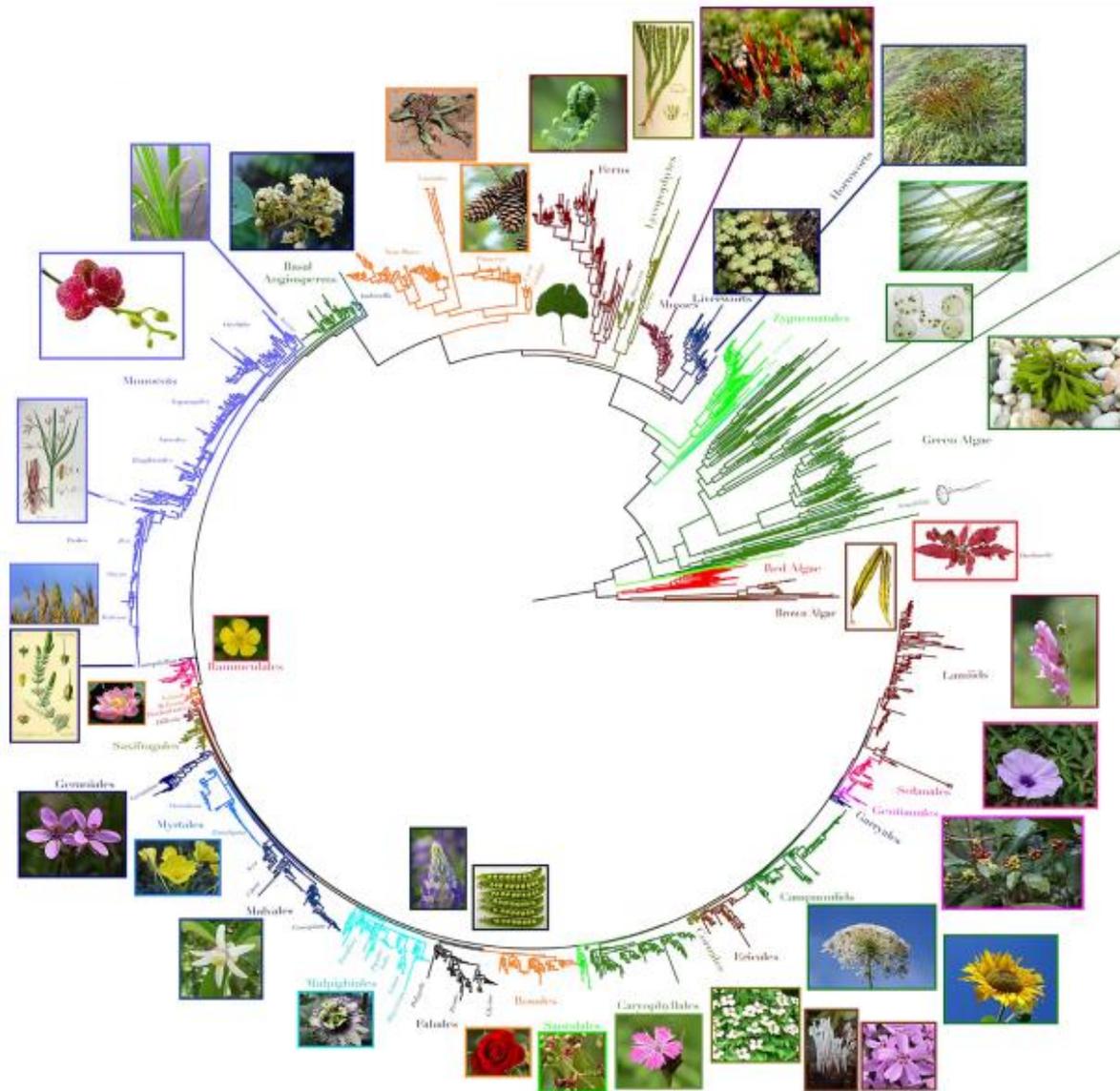


GENÔMICA - EXEMPLO

Marcador - seleção de tomatesiros resistentes à doença vira-cabeça.



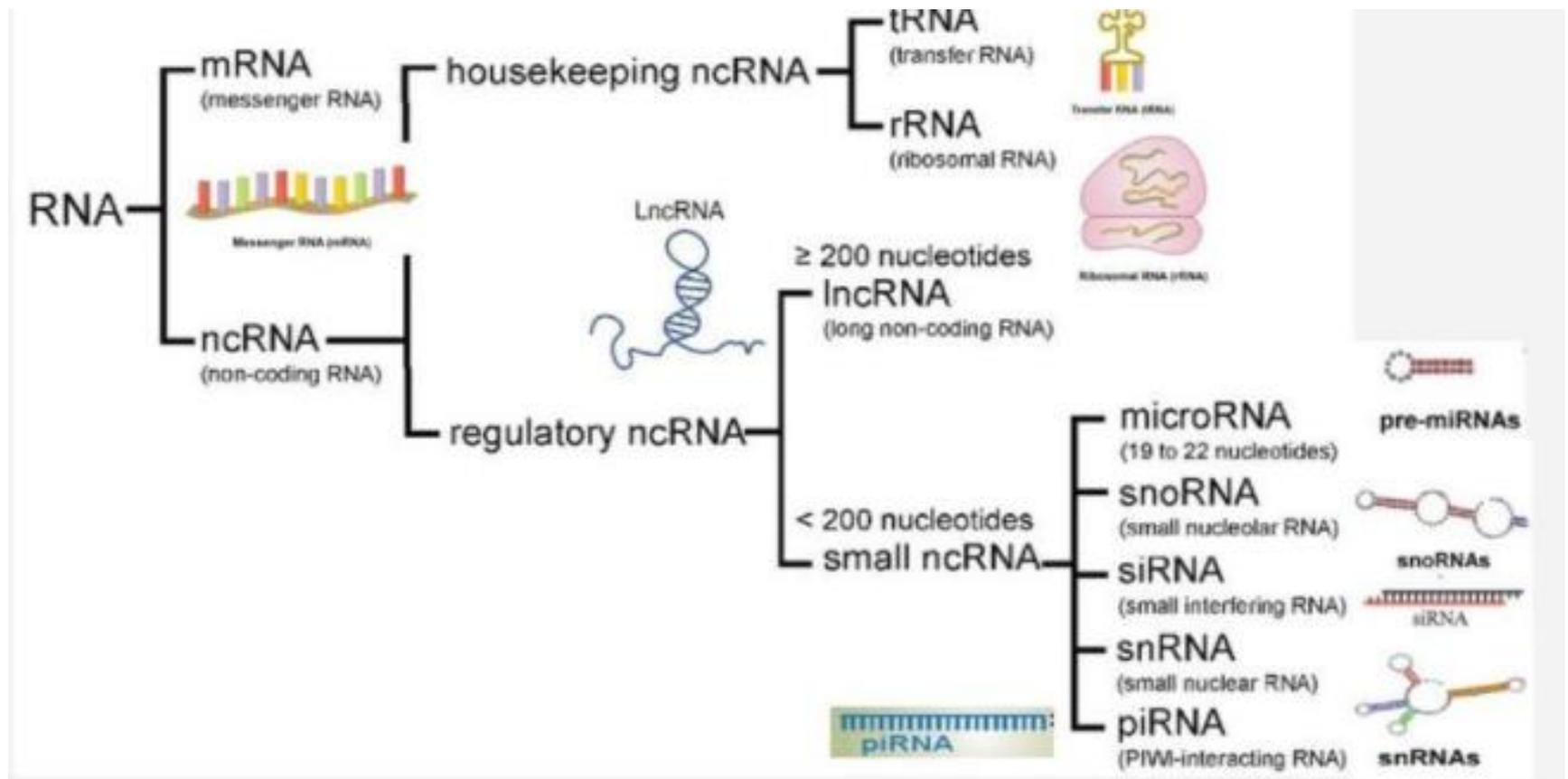
GENÔMICA



Plastid phylogeny for 1,879 species of green plants, based on 80 protein-coding genes (DOI: 10.1002/ppp3.10159)

TRANSCRIPTÔMICA

- É o estudo do transcriptoma, ou seja, do conjunto completo de todas as moléculas de RNA expressas em uma célula, tecido ou organismo, em um determinado momento (condição fisiológica ou desenvolvimento).



TRANSCRIPTÔMICA

➤ PORQUÊ ESTUDAR O TRANSCRIPTOMA?

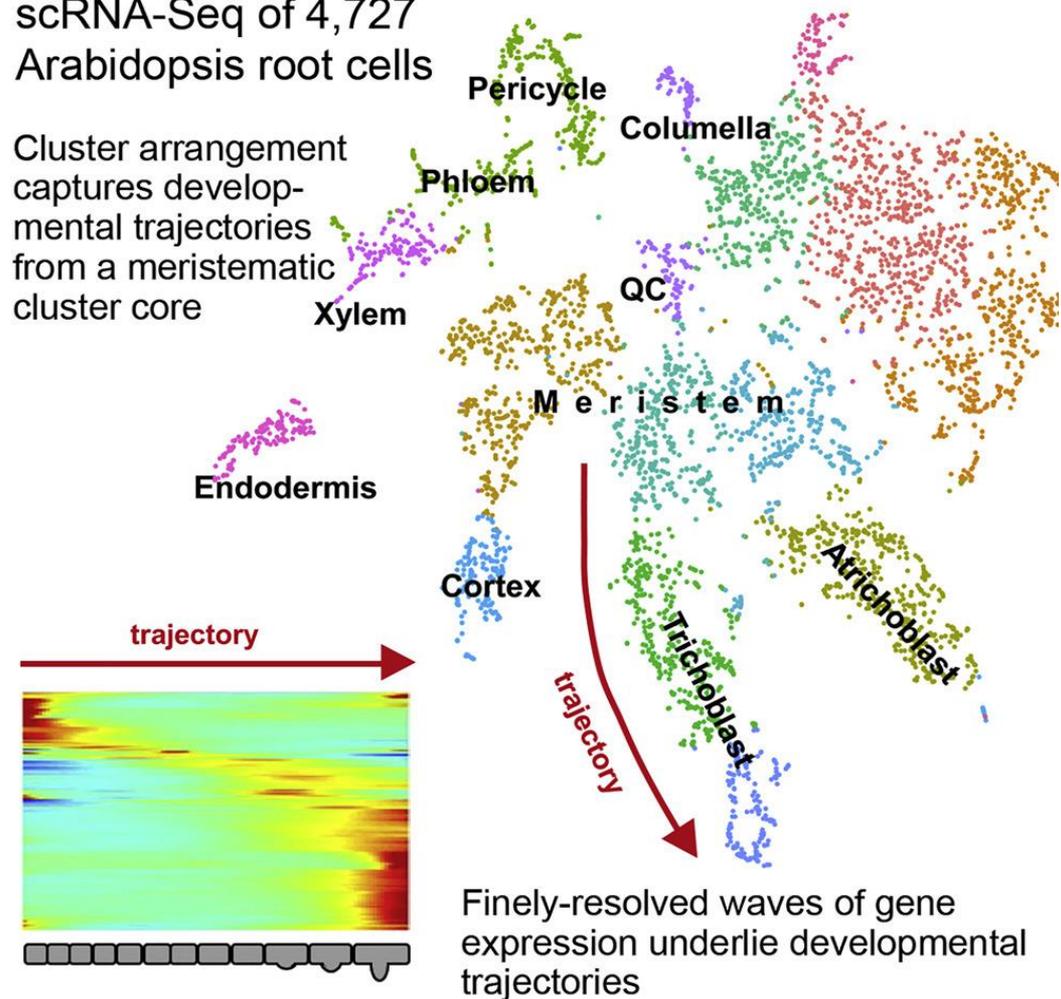
- **Catalogar todas as espécies de RNA (mRNAs, RNAs não-codificantes e pequenos RNAs);**
- **Determinar a estrutura da transcrição de genes, em termos dos seus locais de início, 5' e 3', padrões de splicing e de outras modificações pós-transcricionais;**
- **Quantificar alteração dos níveis de expressão sob diferentes condições;**
- **Identificar o modelo de expressão espaço-tempo dos transcritos e como eles variam entre os diferentes tecidos e sob diferentes condições;**

TRANSCRIPTÔMICA

Spatio-temporal Developmental Trajectories in the Arabidopsis Root Revealed Using High-Throughput Single-Cell RNA Sequencing.

scRNA-Seq of 4,727 Arabidopsis root cells

Cluster arrangement captures developmental trajectories from a meristematic cluster core



TRANSCRIPTÔMICA

Hanfu wild (HC)



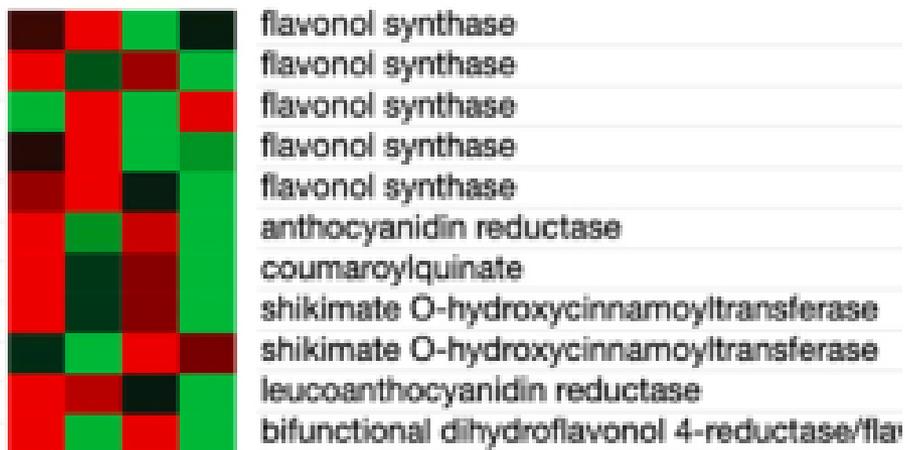
Hanfu mutant (HM)



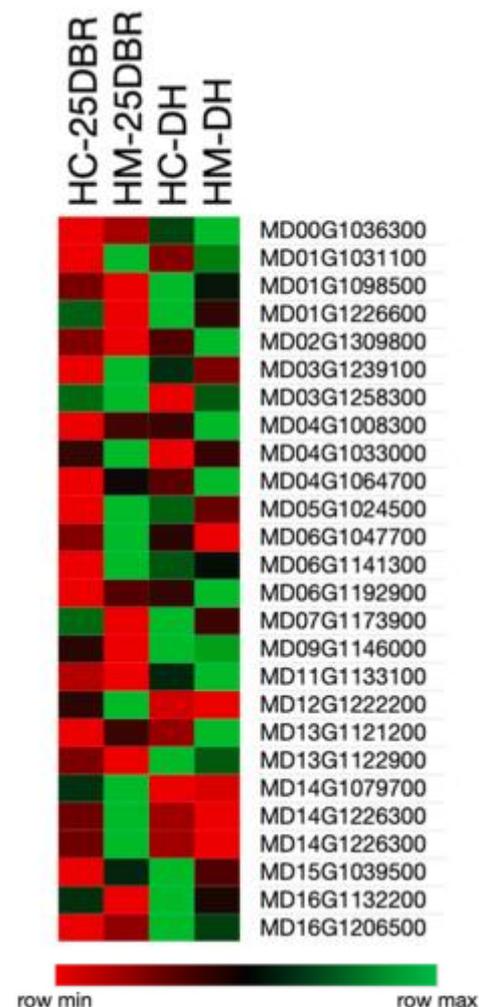
Amadurece 20 dias antes do HC

HC-25DBR
HM-25DBR
HC-DH
HM-DH

Biossíntese de flavonóides



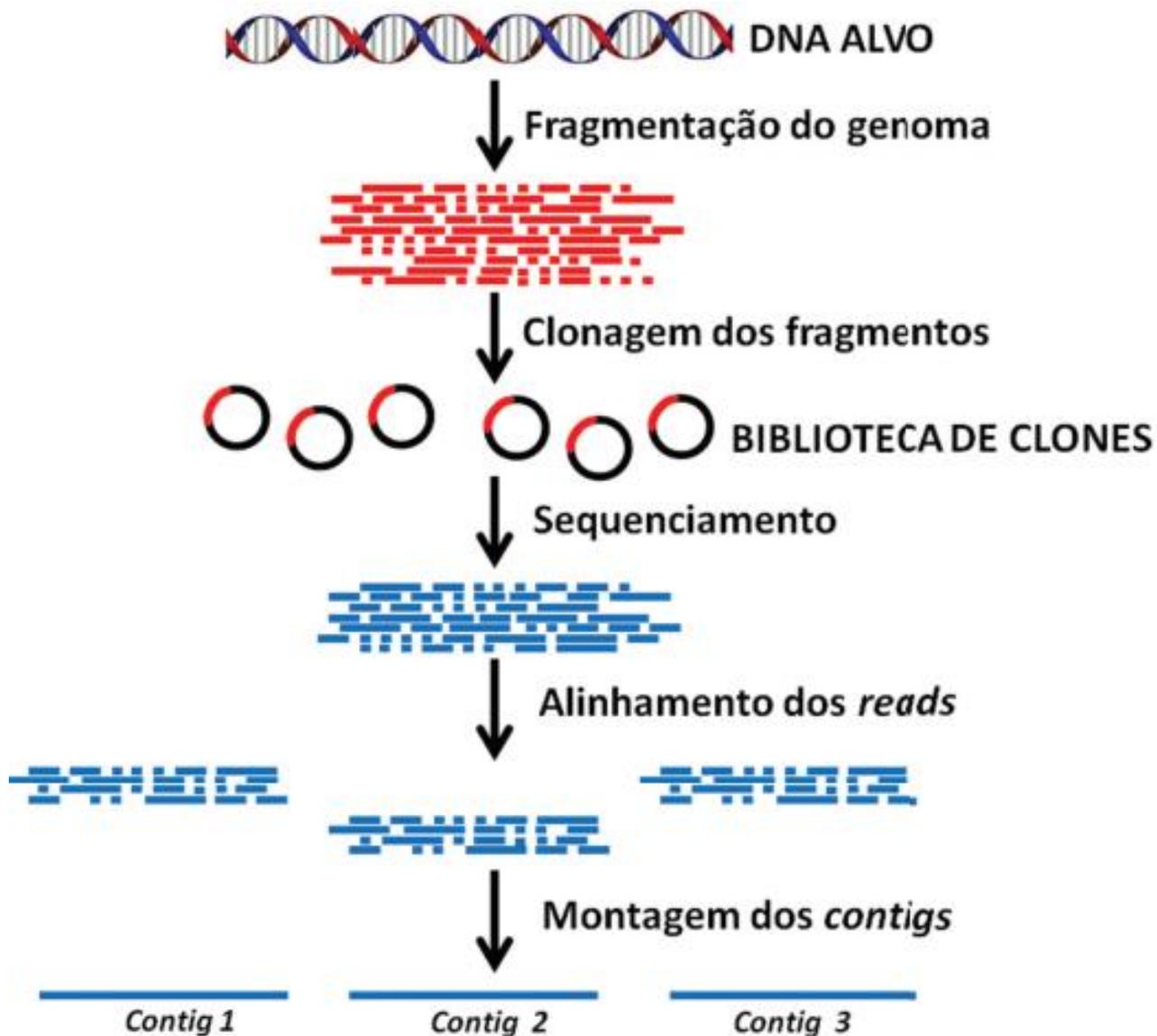
Fatores de transcrição



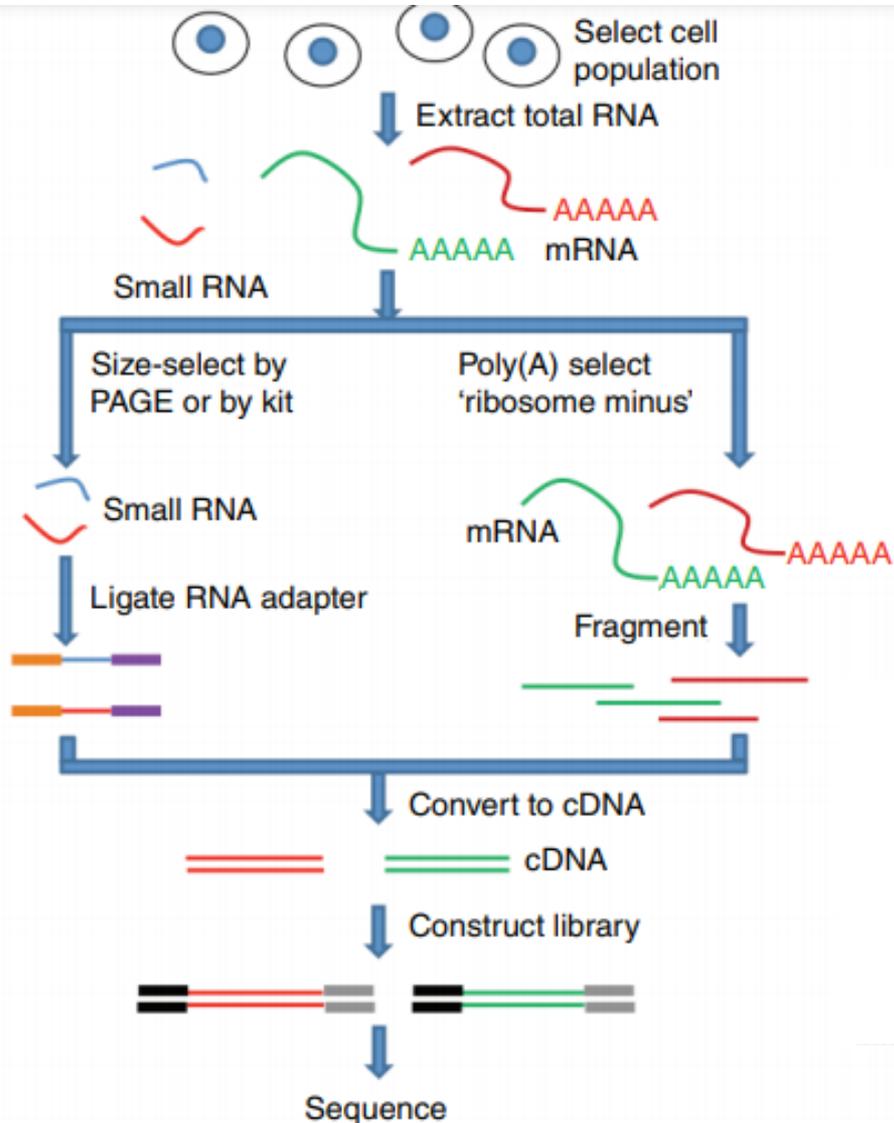
DE ONDE VEM O DNA/RNA A SER SEQUENCIADO ?



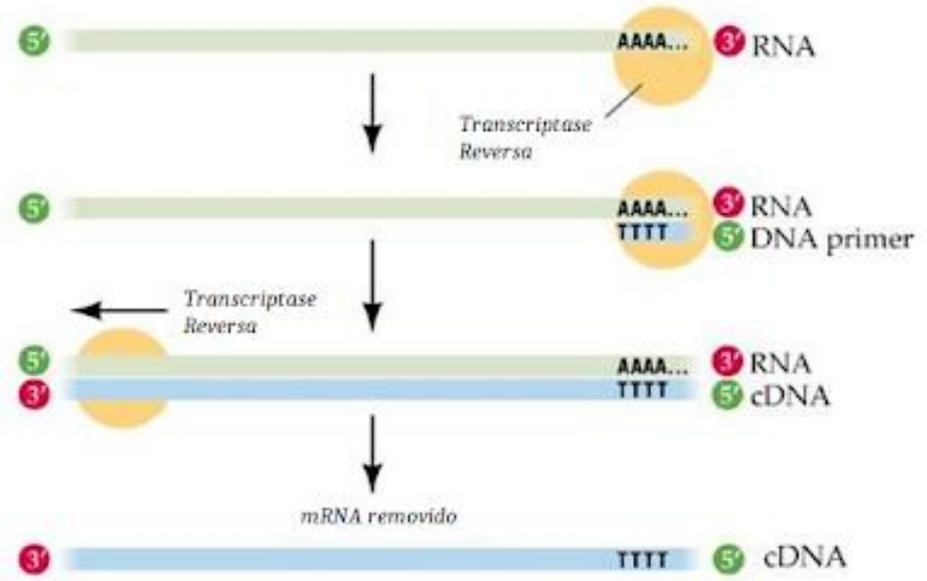
SEQUÊNCIAMENTO DNA



SEQÜENCIAMENTO RNA



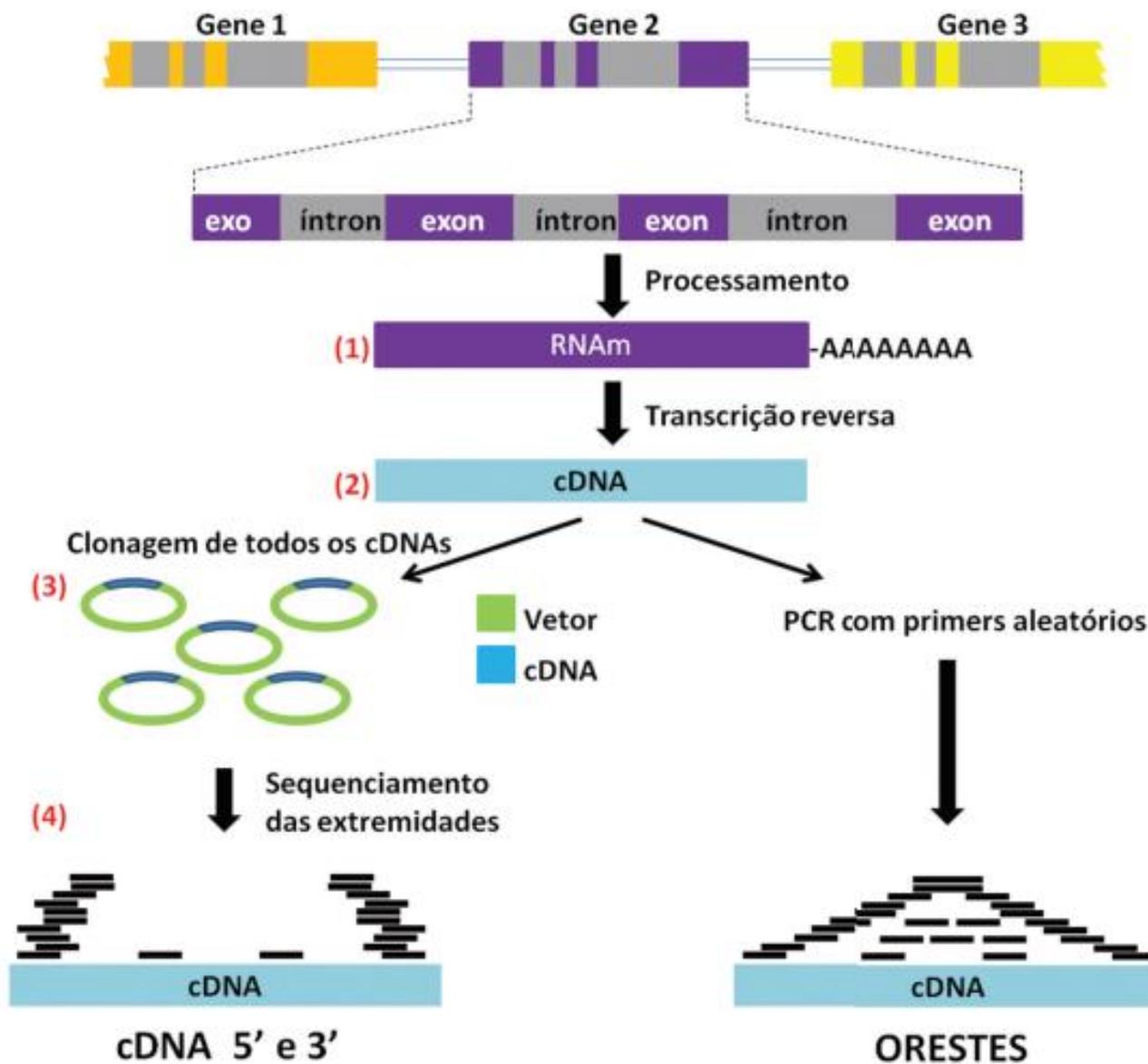
cDNA (DNA complementar)



*Transcriptase reversa

É uma DNA polimerase RNA dependente, ou seja, sintetiza uma cadeia de DNA a partir de um molde de RNA.

SEQUÊNCIAMENTO RNA



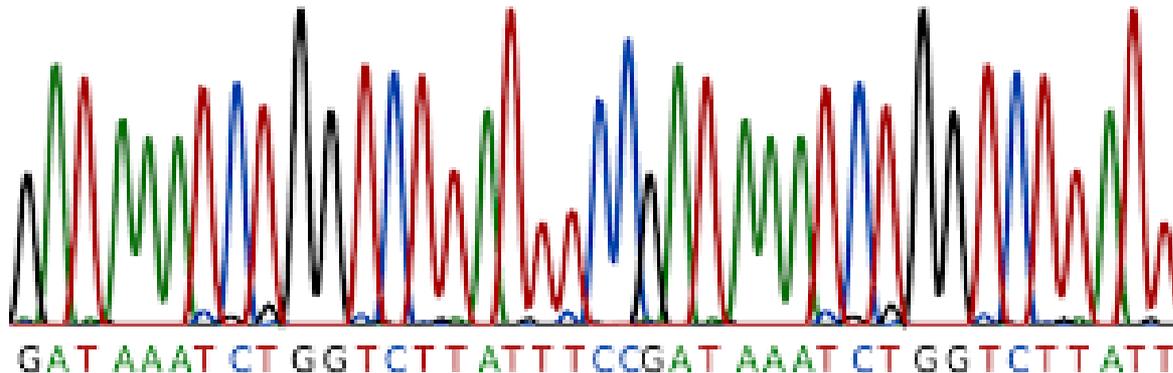
**MAS COMO ESTUDAR O GENOMA E O
TRANSCRIPTOMA????**



SEQUENCIAMENTO

- O sequenciamento é uma técnica que permite identificar, na ordem correta, a sequência de nucleotídeos de uma molécula de DNA ou RNA, visando conhecer a informação genética contida nesta estrutura.
- As metodologias responsáveis por tal façanha fornecem, para cada uma das bases determinadas, uma informação referente a sua qualidade (confiabilidade).

SEQUENCIAMENTO DE DNA



SEQUÊNCIAMENTO

➤ COMO SEQUENCIAR?

Sequenciamento de Primeira Geração (CLONAGEM)

- Degradação química – Maxam & Gilbert
- Interrupção da cadeia (ddNTPs) – Sanger

Sequenciamento de Segunda Geração (AMPLIFICAÇÃO CLONAL)

- **HiSeq, MiSeq , HiScan SQ - Illumina**
- 454 –Roche
- Solid – Applied Biosystem
- **Ion Torrent – ABI – Life Technologies**

Sequenciamento de Terceira Geração (SINGLE MOLECULE)

- **Nanopore – Gridlon/Minilon**
- Heliscope - Helicos Biosciences
- **PacBio RS – Pacific Biosciences**

SEQUÊNCIAMENTO DE DNA

➤ TECNOLOGIAS

↳ Dependente da evolução das técnicas

• 1ª GERAÇÃO:

1977 - Método de Maxam & Gilbert (hidrólise química)

Método de Sanger - Método de terminação di-deoxi

* Usado no sequenciamento do genoma humano. Principal método de sequenciamento por 30 anos

SEQUÊNCIAMENTO DE DNA

➤ MÉTODO SANGER

Reação de Sequenciamento

Amostra (cDNA; DNA)

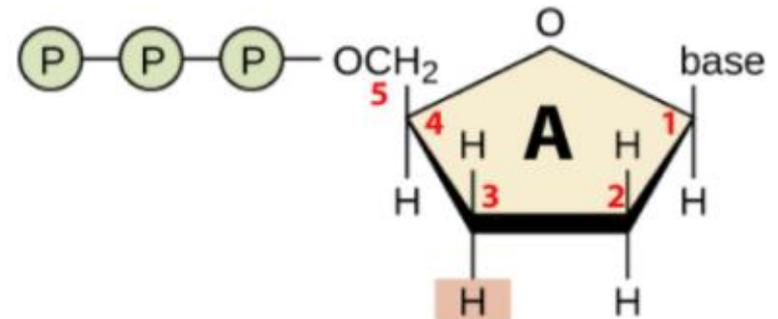
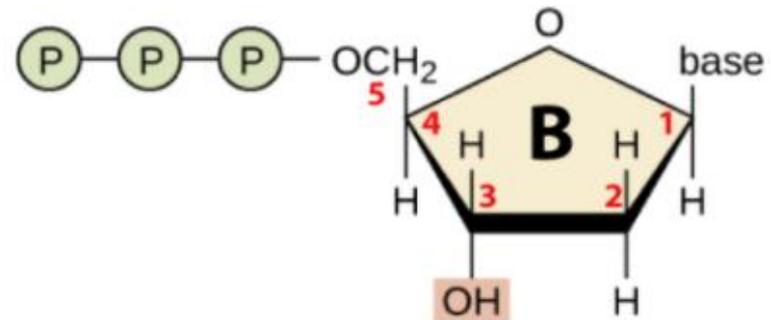
dNTPs (dATP, dTTP, dCTP, dGTP)

Tampão da enzima

Enzima DNA Polimerase

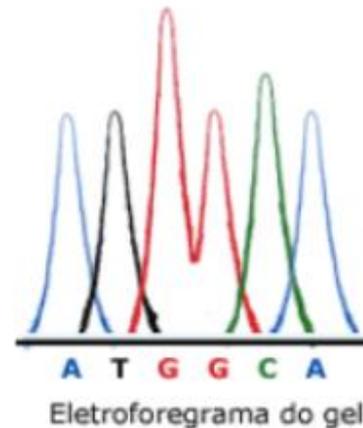
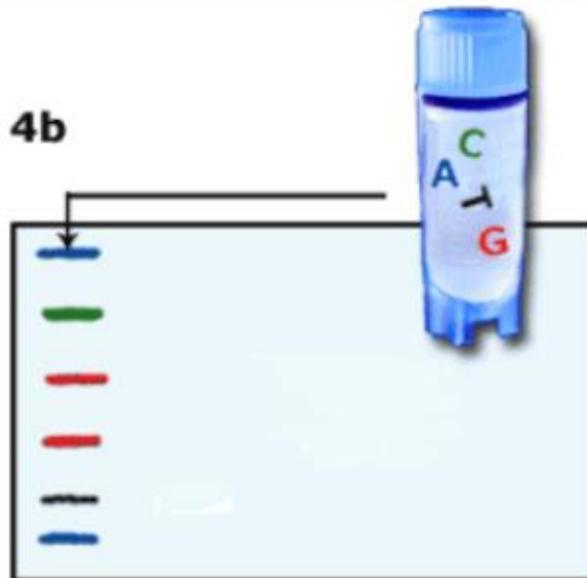
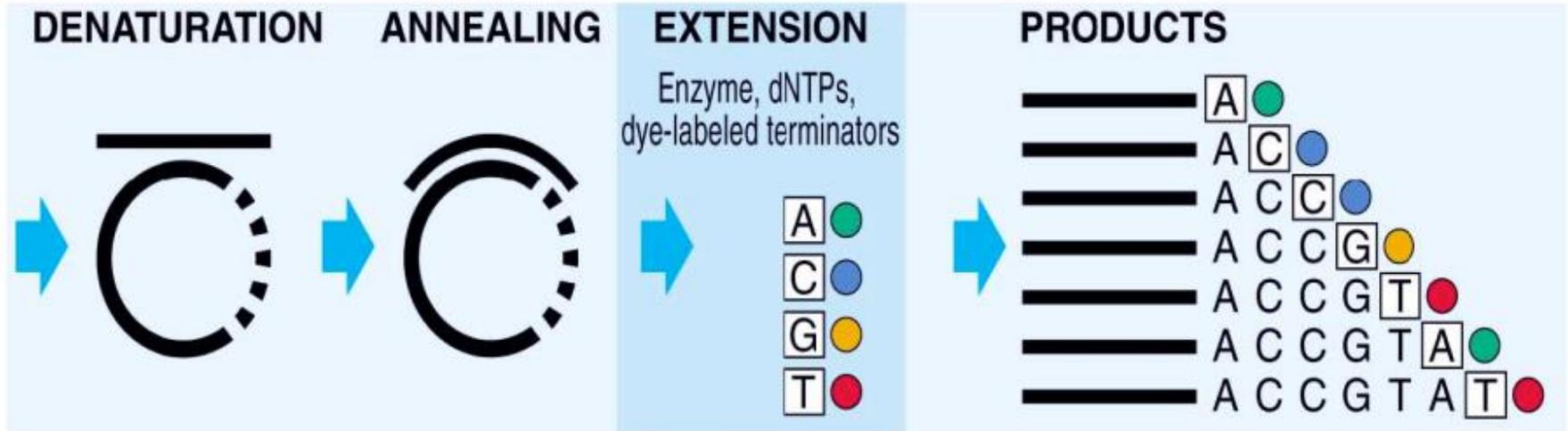
Iniciador (S ou AS)

ddNTPs (ddATP, ddTTP, ddCTP, ddGTP)



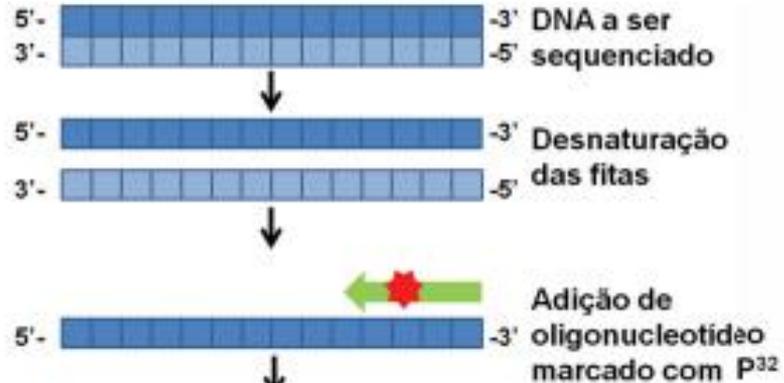
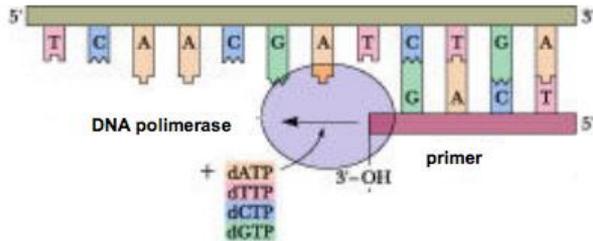
SEQUÊNCIAMENTO DE DNA

➤ MÉTODO SANGER



SEQUÊNCIAMENTO DE DNA

➤ MÉTODO SANGER



Frederick Sanger

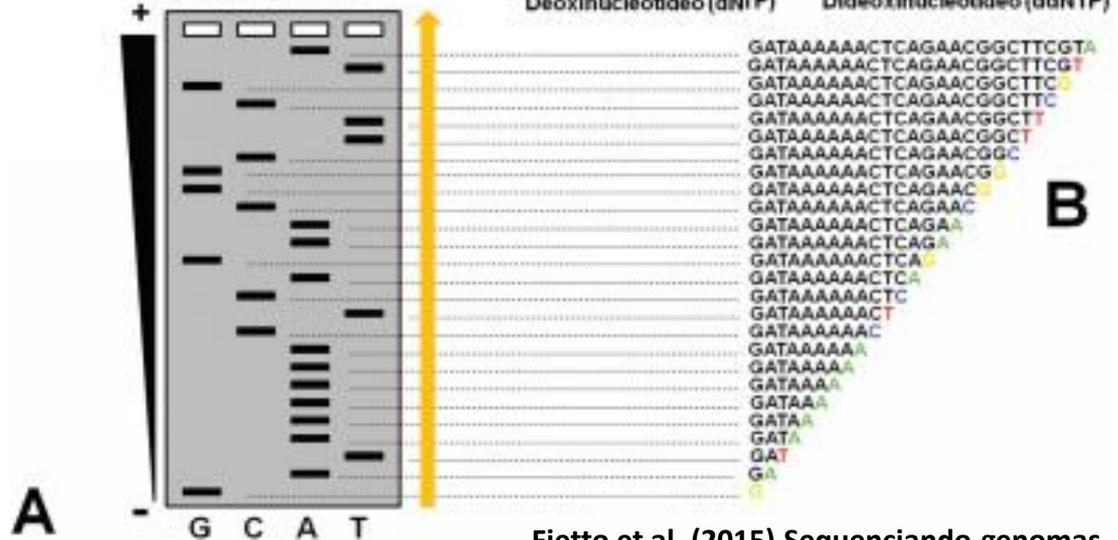
MIX DE PCR



Deoxinucleotídeo (dNTP)



Dideoxynucleotídeo (ddNTP)



SEQUÊNCIAMENTO DE DNA

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.

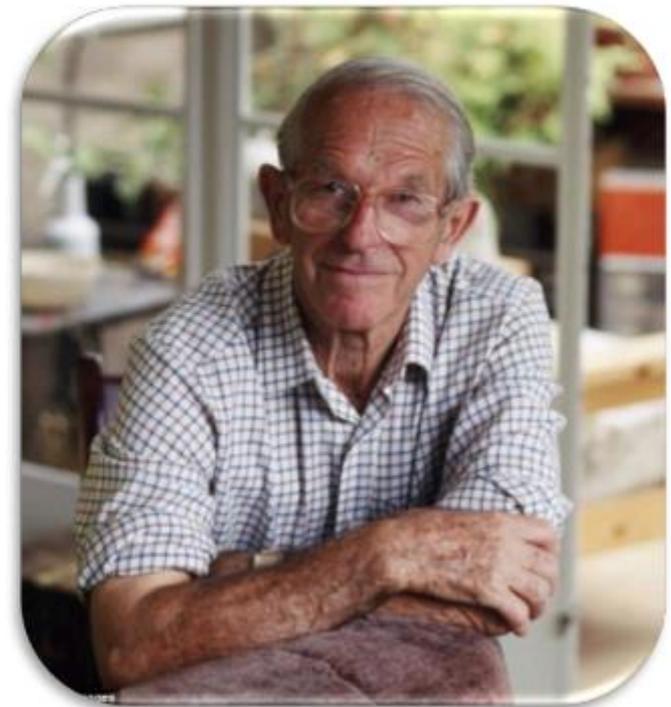
The "plus and minus" method (1) is a relatively rapid and simple technique that has made possible the determination of the sequence of the genome of bacteriophage ϕ X174 (2). It depends on the use of DNA polymerase to transcribe specific regions of the DNA under controlled conditions. Although the method 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

Frederick Sanger
13 Aug 1918 – 19 Nov 2013



- Prêmio Nobel Prize de Química em **1958** e **1980**
- Publicou o método de sequenciamento em **1977**

SEQUÊNCIAMENTO DE DNA

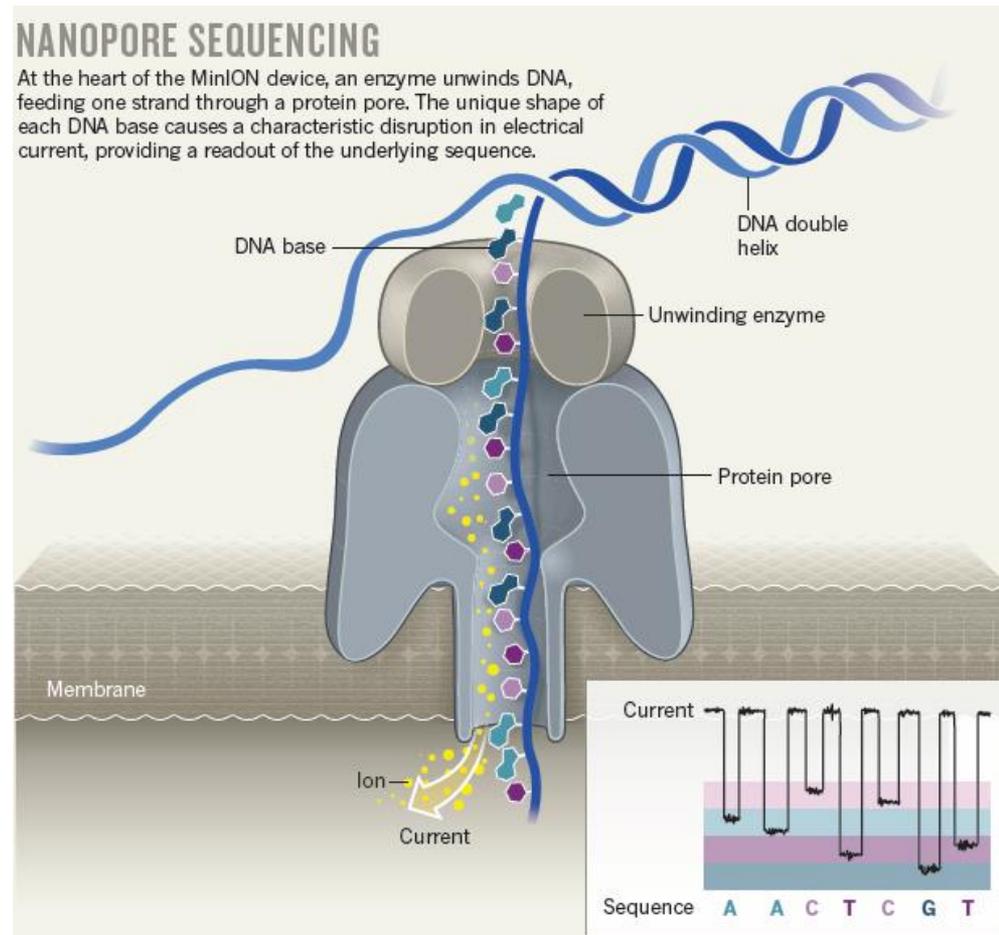
Empresa	Sequenciador	Total de bases por corrida (Gb)	Tamanho das sequências (pb)	Tempo de corrida (h)	Preço do equipamento (US\$)
Illumina	iSeq 100	1,2	2 x 150	9,5 – 19	190.900
	MiniSeq	7,5	2 x 150	4 – 24	50.000
	MiSeq	15	2 x 300	4 – 55	99.000
	NextSeq 550	120	2 x 150	12 – 30	285.000
	NextSeq 1000	120	2 x 150	29	215.000
	NextSeq 2000	300	2 x 150	24 – 48	335.000
	NovaSeq	6000	2 x 150	13 – 45	985.000
Thermo Fisher	SOLiD 5500xl	95	2 x 60	144	595.000
	SOLiD 5500xl Wildfire	240	2 x 50	240	70.000
	SOLiD 5500	48	2 x 60	144	349.000
	SOLiD 5500 Wildfire	120	2 x 50	240	70.000
	Ion Gene Studio S5 System	15	600	19	120.000
	Ion Gene Studio S5 Plus System	30	600	10 – 11,5	150.000
	Ion Gene Studio S5 Prime System	50	600	6,5	180.000
Pacific Biosystems	Sequel	20	1.000 – 20.000	20	350.000
	Sequel II	160	15.000 – 20.000	30	495.000

SEQUÊNCIAMENTO DE DNA

➤ 3ª OU 4ª GERAÇÃO

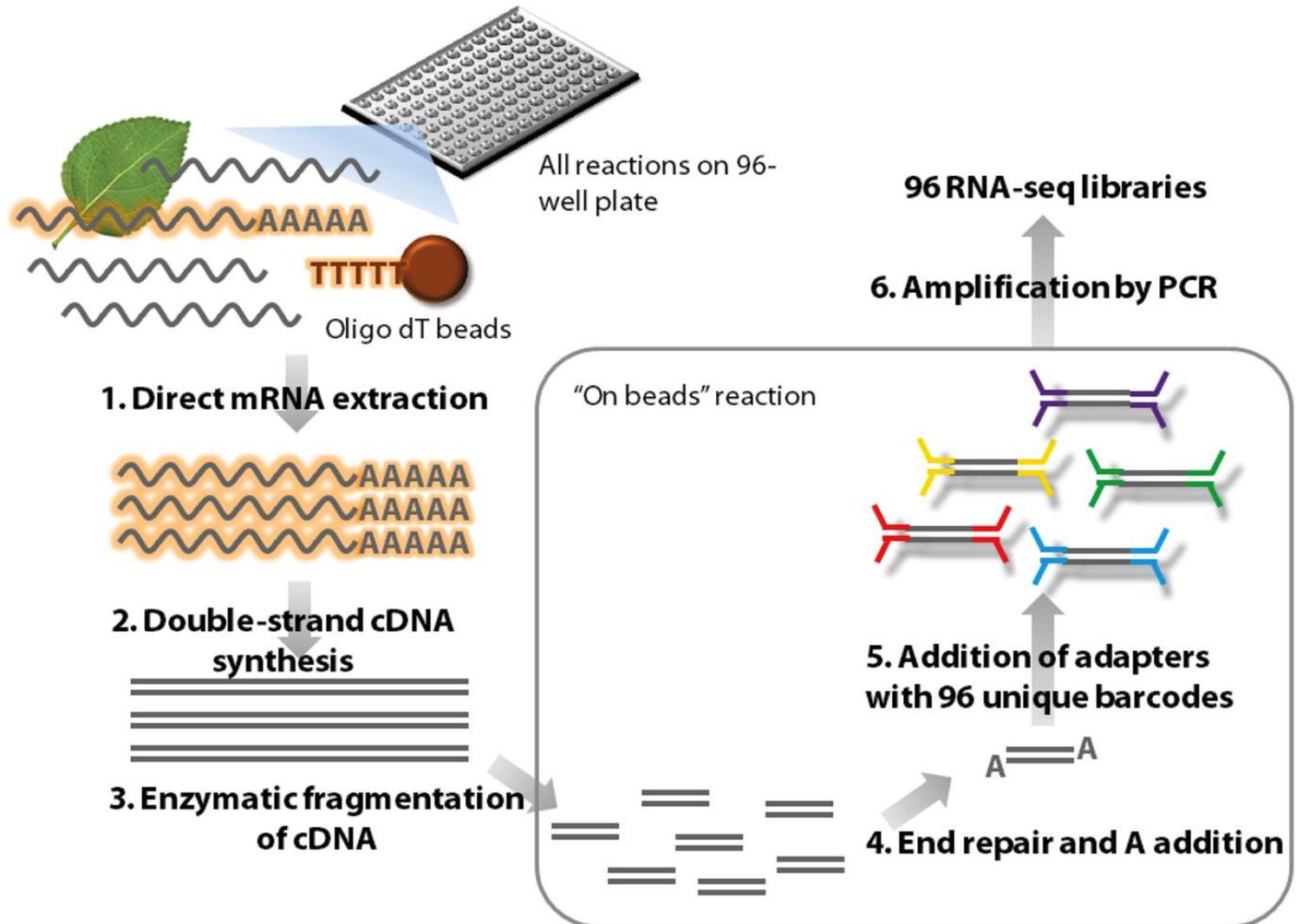
Pesquisadores brasileiros sequenciam genoma do coronavírus por meio da tecnologia Nanopore

Pesquisadores do Brasil e da Universidade de Oxford sequenciaram o genoma do vírus 2019 n-CoV do [primeiro caso confirmado da doença em São Paulo](#). O trabalho, geralmente feito em 15 dias pelos cientistas, foi realizado em apenas dois dias, de acordo com a Fundação de Amparo à Pesquisa do Estado de S. Paulo (Fapesp).



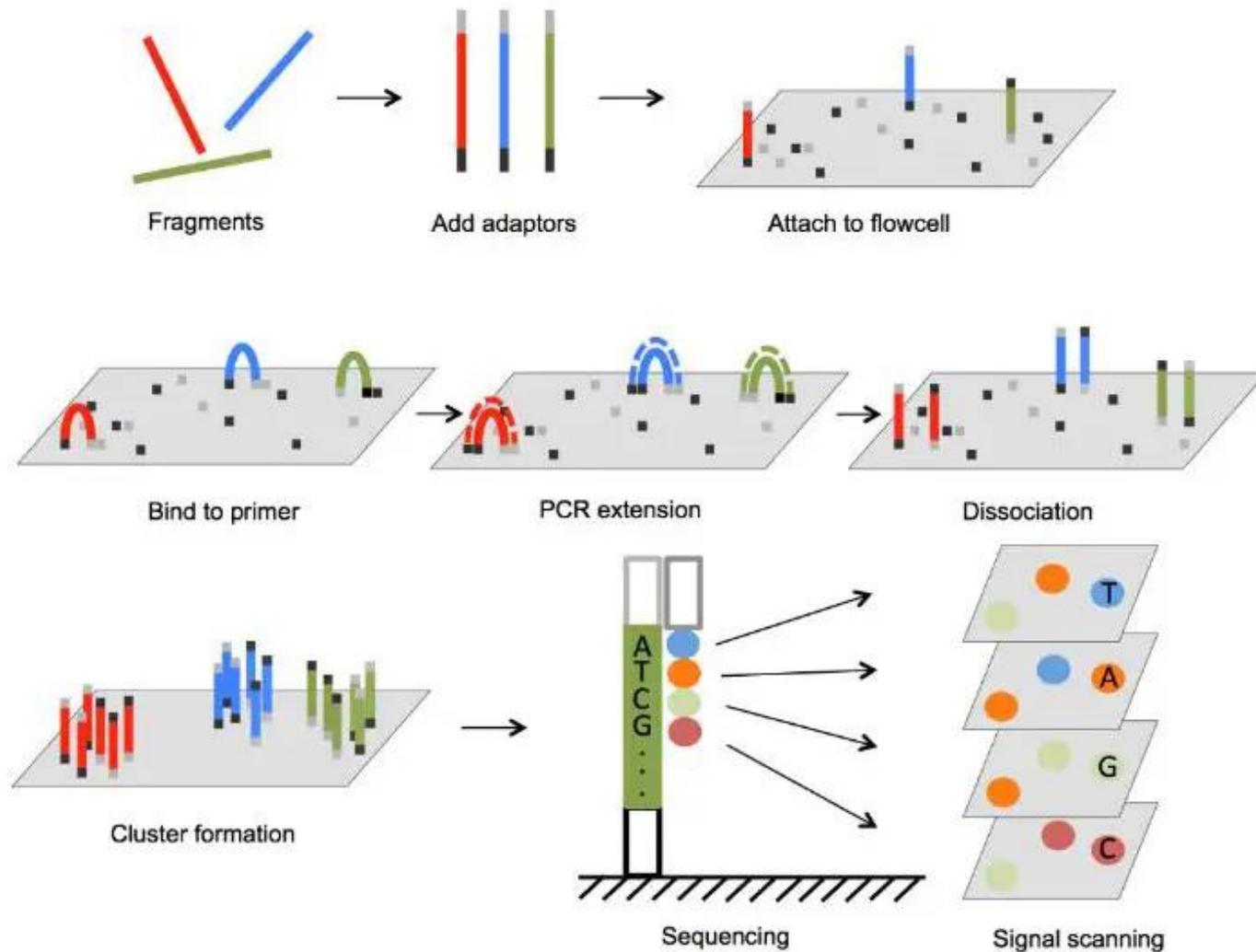
SEQUÊNCIAMENTO DE RNA

➤ PLATAFORMA ILLUMINA

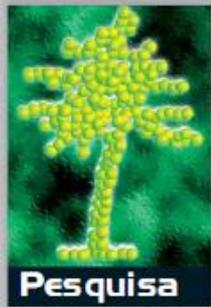


SEQUÊNCIAMENTO DE RNA

➤ PLATAFORMA ILLUMINA



VALE A PENA LER!!!



As ômicas: Integrando a bioinformação

O papel da bioinformática em expansão

Dr. Eliseu Binneck

Consultor/Pesquisador na área de Bioinformática

Embrapa Soja, Londrina - PR.

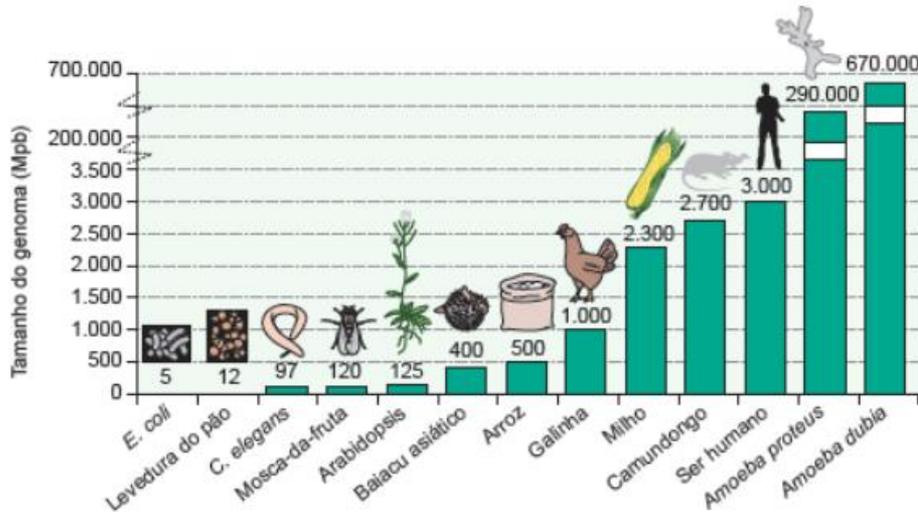
binneck@cnpso.embrapa.br

Imagens cedidas pelo autor

PROTEÔMICA

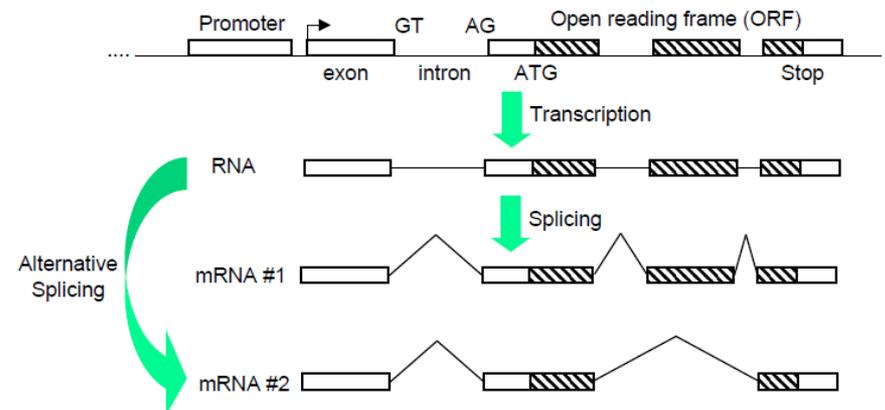
➤ POR QUE ESTUDAR AS PROTEÍNAS?

Desempenham funções essenciais (replicação, transcrição, síntese de carboidratos e lipídeos) e são versáteis (moléculas estruturais, sinalizadoras, enzimas).



Como esses genes são “usados” ?

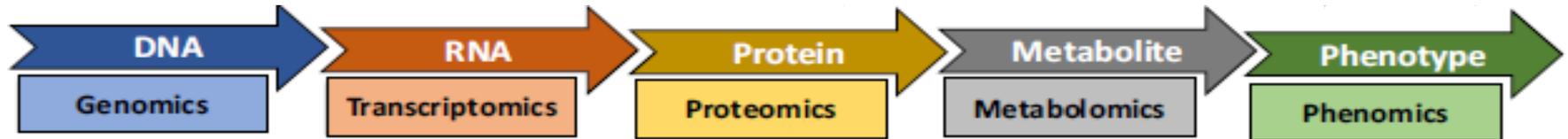
- *Splicing* alternativo
- Modific. pós transcricionais
- Modific. pós- traducionais



PROTEÔMICA

➤ POR QUE ESTUDAR AS PROTEÍNAS?

As proteínas fornecem informações, em nível molecular, sobre a variabilidade genética que é efetivamente expressa pelo genoma !!!



PROTEÔMICA:

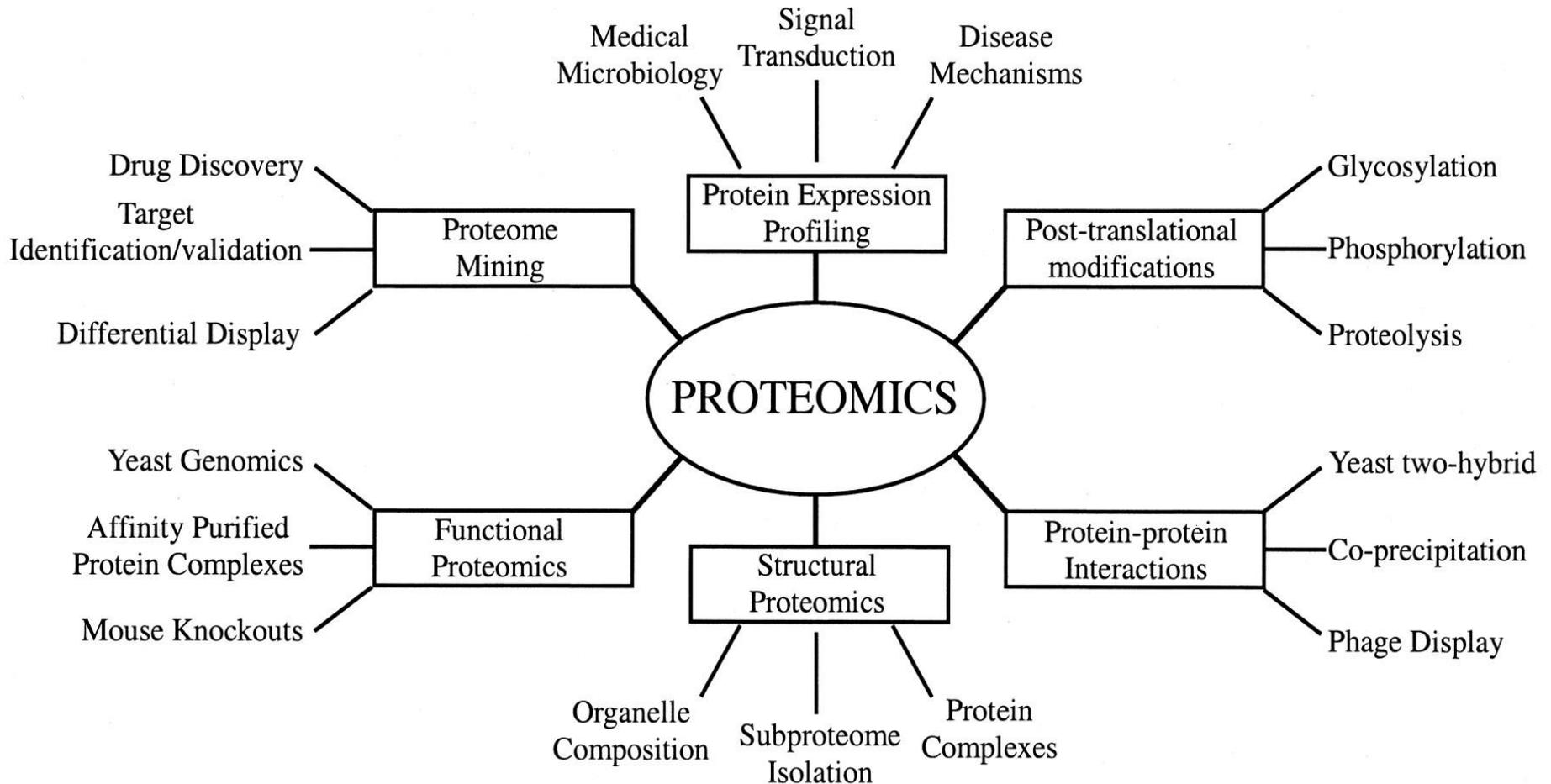
Metodologia, conjunto de técnicas utilizadas para se estudar o Proteoma.

PROTEOMA:

Conjunto de todas as proteínas de uma célula, organela, tecido, ou organismo em um dado momento/ condição.

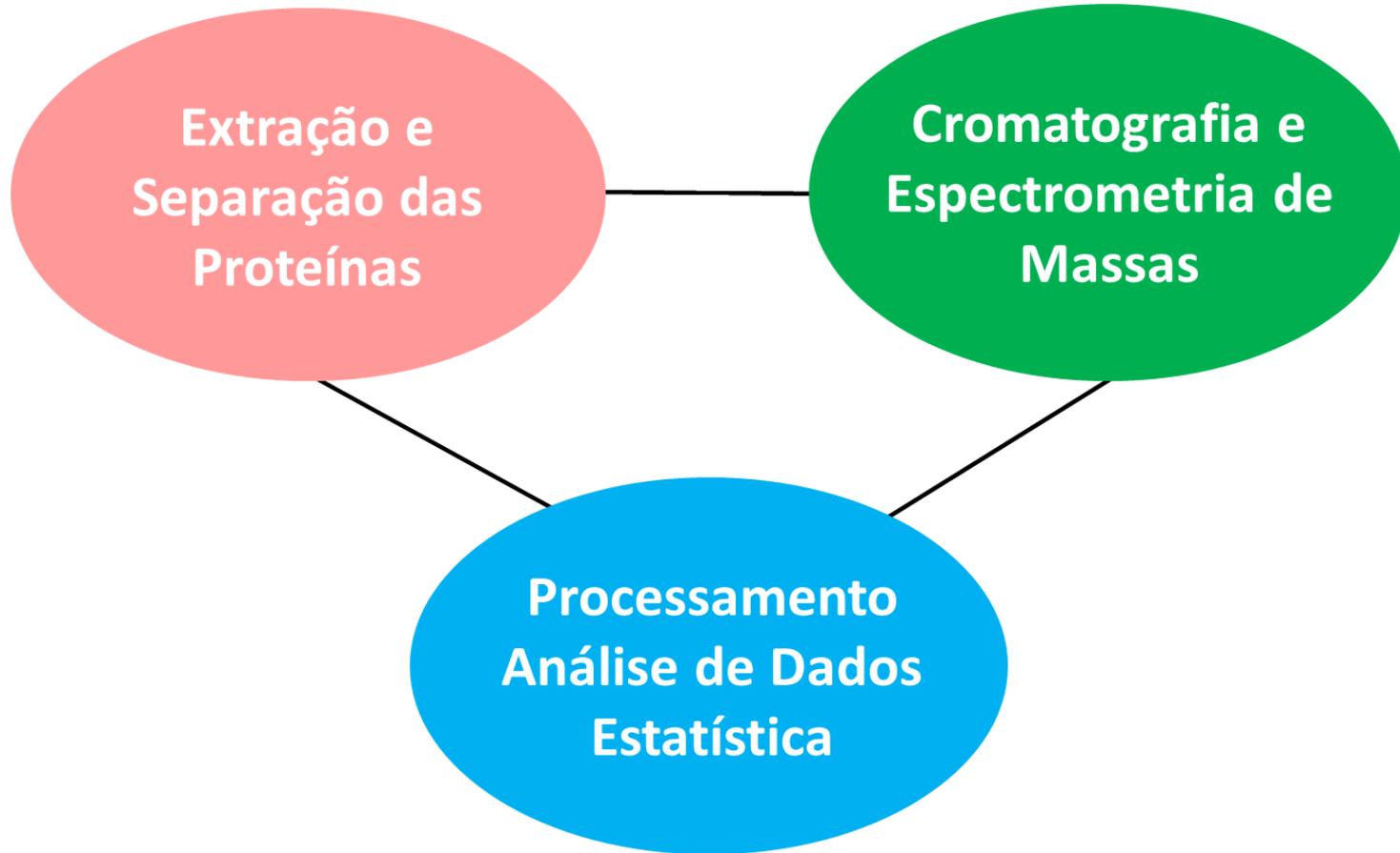
PROTEÔMICA

➤ POSSIBILIDADES:



PROTEÔMICA

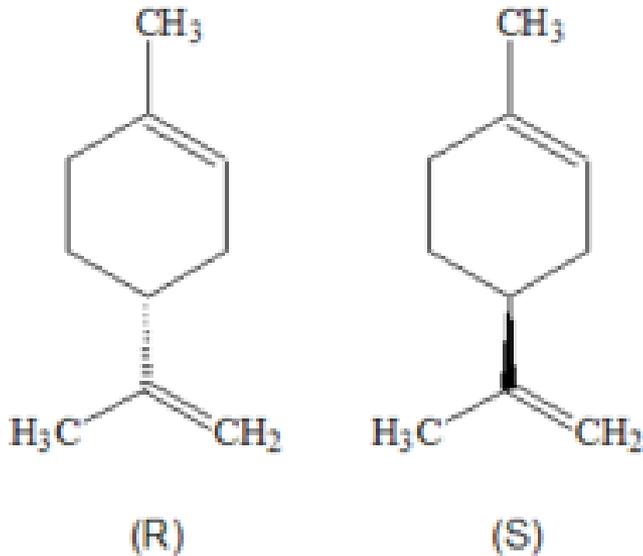
➤ ETAPAS:



METABOLÔMICA

➤ METABÓLITOS TEM VÁRIAS FUNÇÕES:

Estrutural, sinalização, estimulantes, inibidores enzimáticos, cofatores de enzimas, defesa e interação com outros organismos (feromônios, pigmentos e o aromas).



(*R*)-limoneno e (*S*)-limoneno:
aroma da laranja e do limão

METABOLÔMICA

METABOLÔMICA

Metodologia, conjunto de técnicas utilizadas para se estudar o Metaboloma.

METABOLÔMA:

Conjunto de todos os metabólitos de uma célula, organela, tecido, ou organismo em um dado momento/ condição.

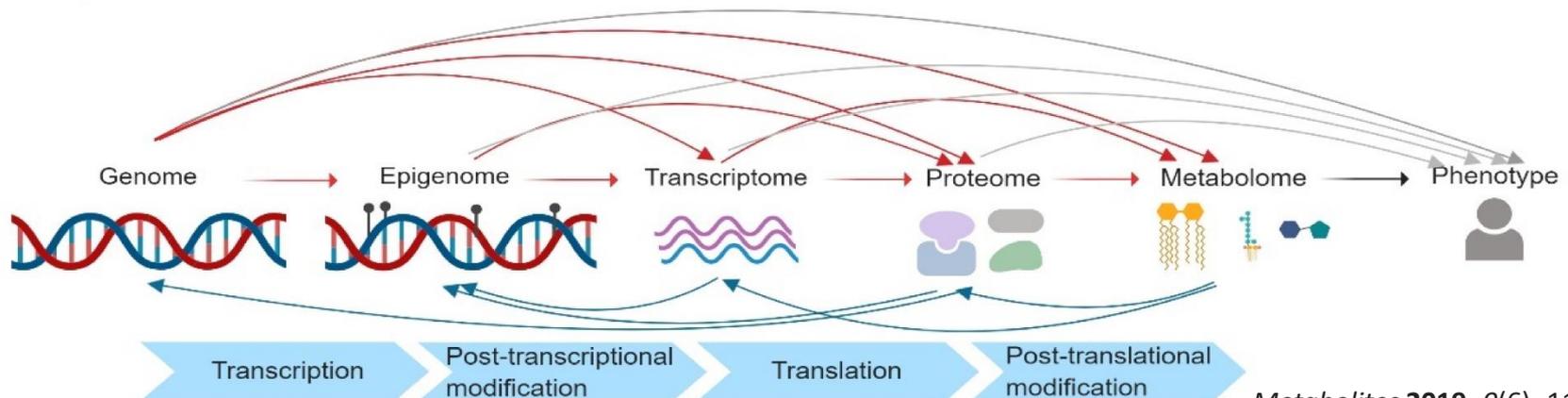
METABOLÔMICA

➤ PORQUE ESTUDAR OS METABÓLITOS:

Seus níveis representam a resposta final do sistema biológico às mudanças internas e externas.

Metabolomics – the link between genotypes and phenotypes

Oliver Fiehn 2002, *Plant Mol. Biol.*



METABOLÔMICA

SCIENCE ADVANCES | RESEARCH ARTICLE

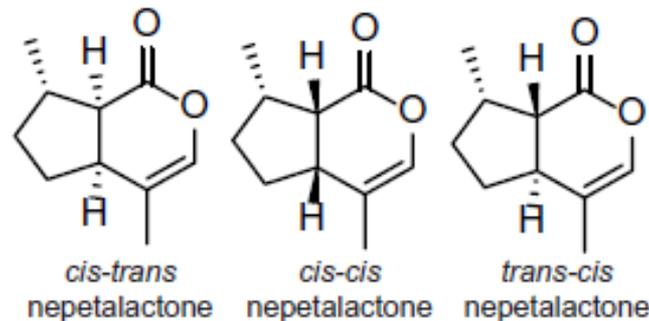
BIOSYNTHESIS

Lichman *et al.*, *Sci. Adv.* 2020; 6 : eaba0721 13 May 2020

The evolutionary origins of the cat attractant nepetalactone in catnip



Gênero *Nepeta* L.



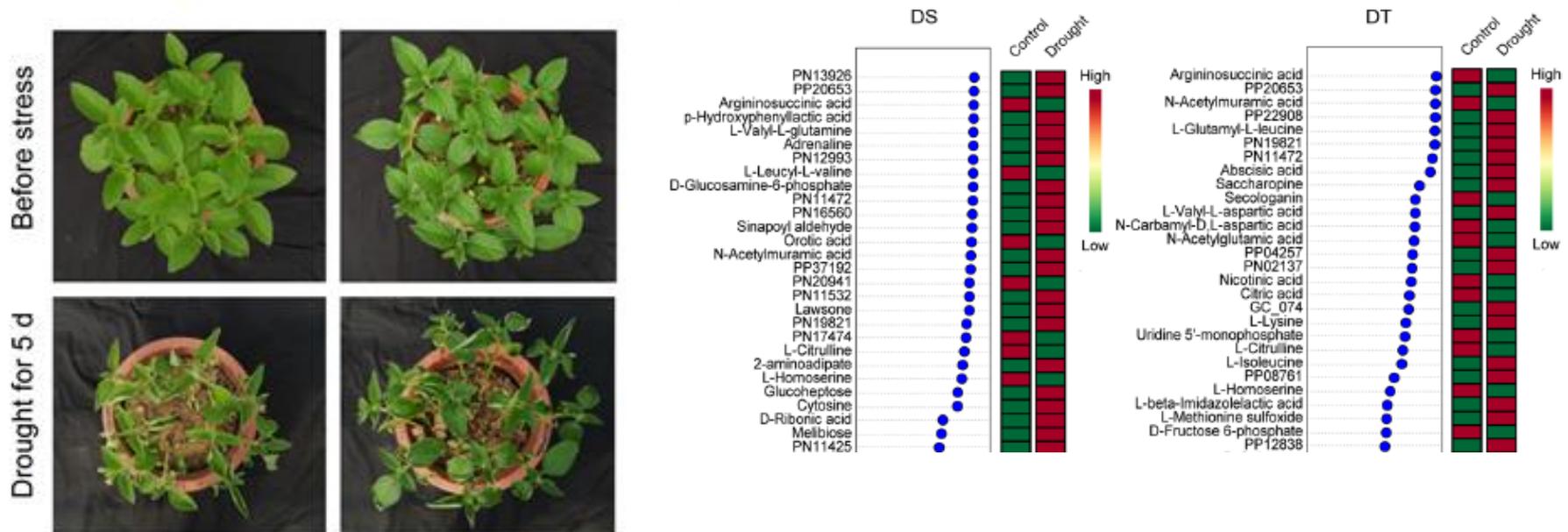
Nepetalactona (terpenoides)

METABOLÔMICA

Transcriptomic and metabolomic profiling of drought-tolerant and susceptible sesame genotypes in response to drought stress

Jun You¹, Yujuan Zhang^{1,2}, Aili Liu¹, Donghua Li¹, Xiao Wang³, Komivi Dossa^{1,4}, Rong Zhou¹, Jingyin Yu¹, Yanxin Zhang¹, Linhai Wang¹ and Xiurong Zhang^{1*} 

Background: Sesame is an important **oil crop** due to its high oil, antioxidant, and protein content. **Drought stress** is a major abiotic stress that affects sesame production as well as the quality of sesame seed. **To reveal the adaptive mechanism of sesame** in response to **water deficient** conditions, **transcriptomic and metabolomics** were applied in drought-tolerant (DT) and drought-susceptible (DS) sesame genotypes.



METABOLÔMICA

You et al. *BMC Plant Biology* (2019) 19:267

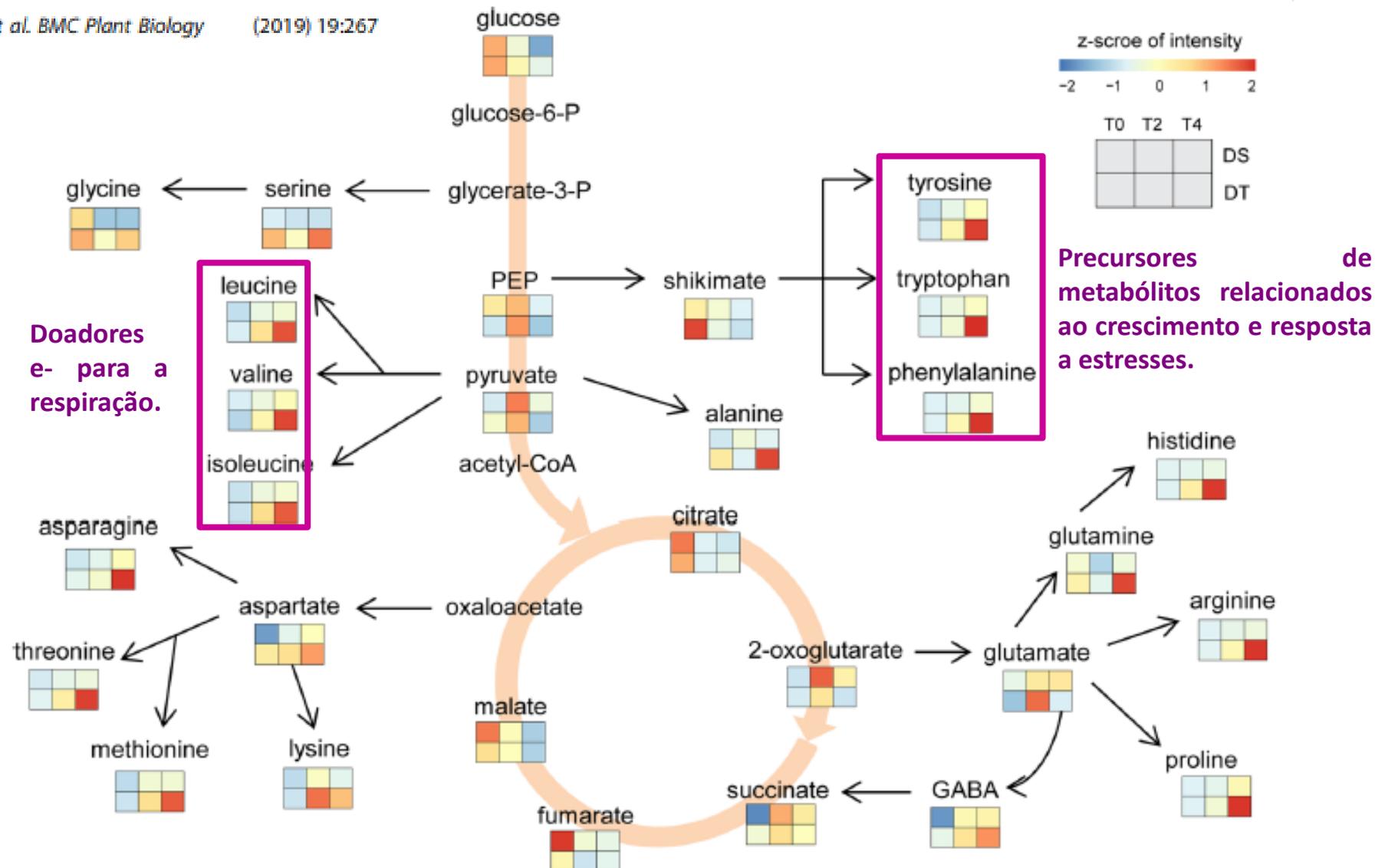
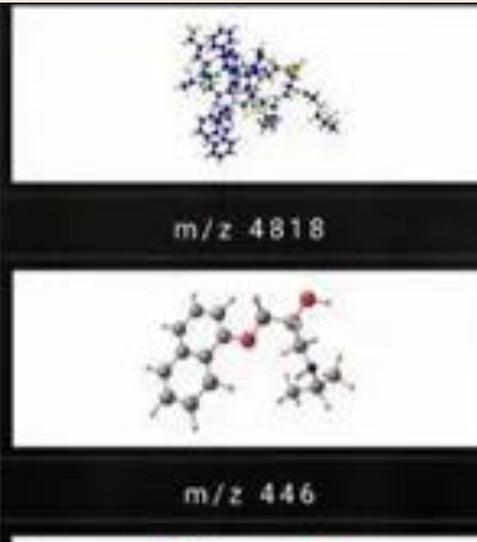


Fig. 6 Metabolic changes involved in amino acid pathways and tricarboxylic acid cycle under drought stress in DT and DS. Each graph represents the normalized intensity of corresponding metabolite in three sampling points from two genotypes

IDENTIFICANDO PROTEÍNAS E METABÓLITOS



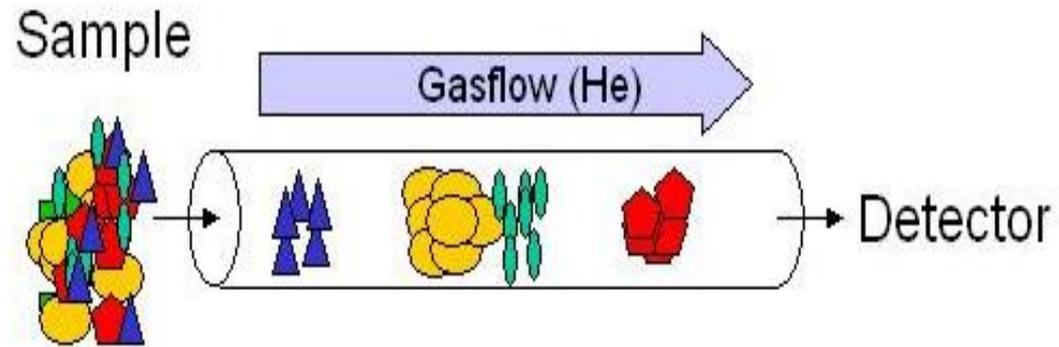
WHO WE ARE.
TECHNOLOGY TO



IDENTIFICANDO PROTEÍNAS E METABÓLITOS

➤ CROMATOGRAFIA ASSOCIADA A ESPECTROMETRIA DE MASSAS

CROMATOGRAFIA

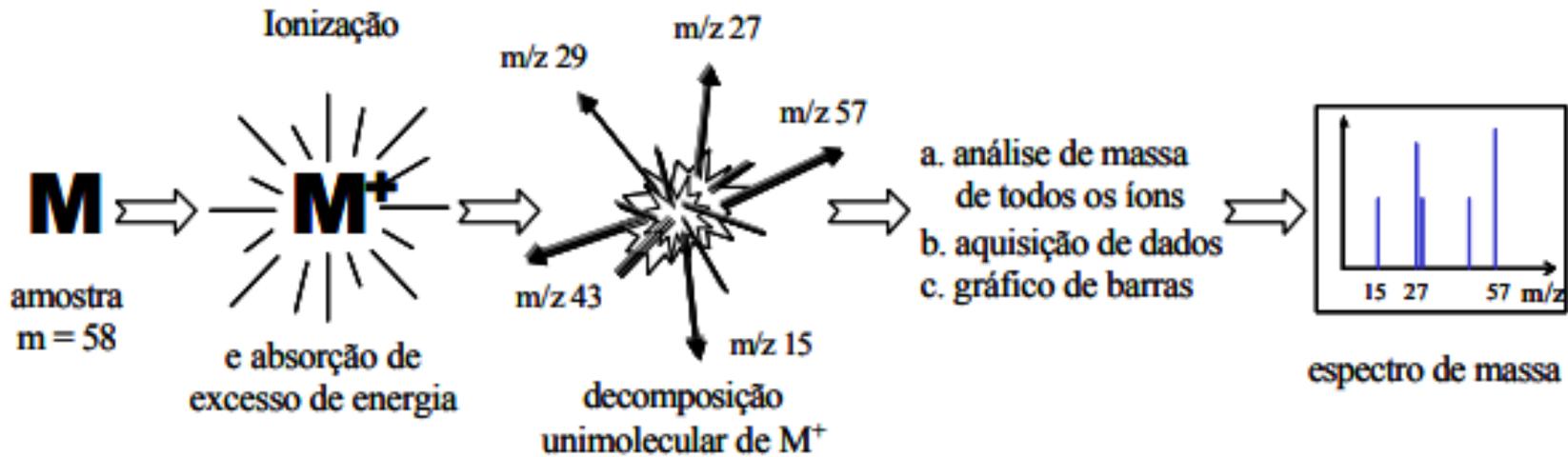


ESPECTROMETRIA DE MASSAS

Ferramenta analítica para determinar a composição molecular de uma amostra, através da relação **massa/carga (m/z)** da molécula ionizada.

IDENTIFICANDO PROTEÍNAS E METABÓLITOS

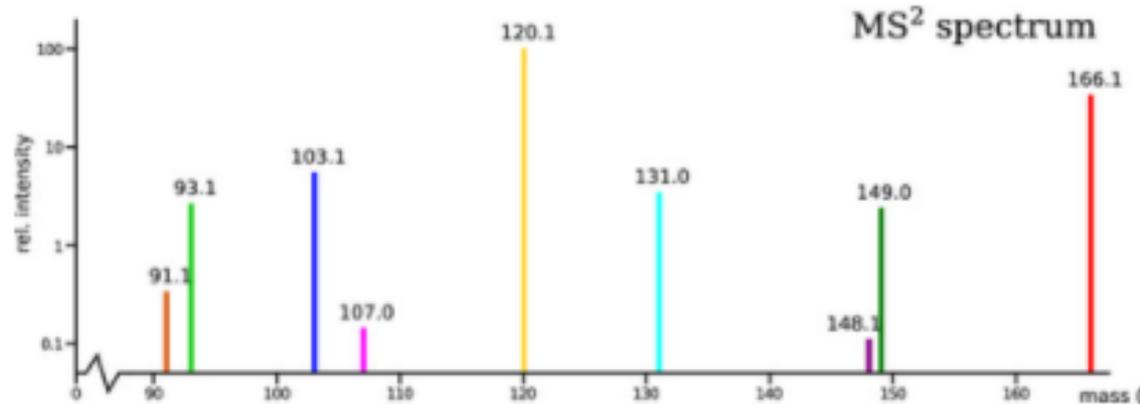
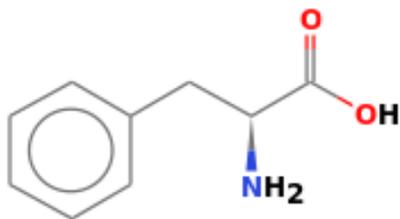
ESPECTROMETRIA DE MASSAS



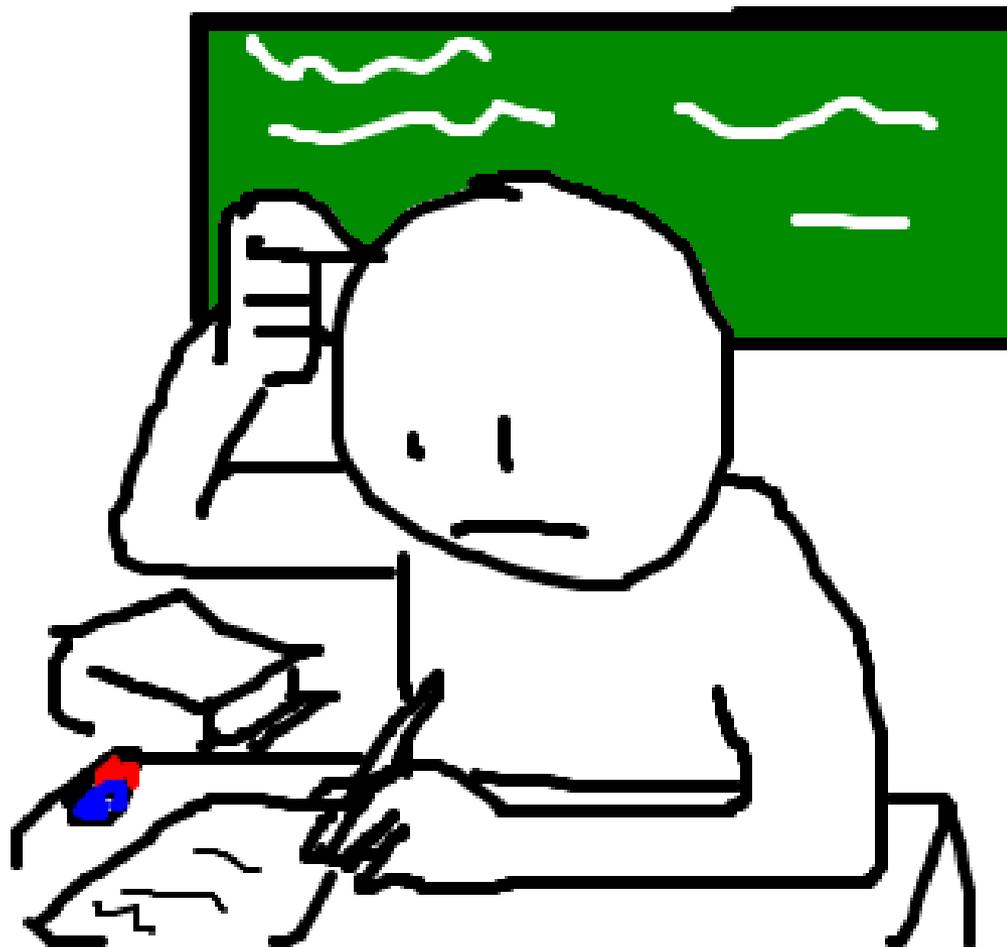
FENILALANINA

Fórmula: $C_9H_{11}NO_2$

Peso molecular: 165.1891



ATIVIDADE



EM GRUPO (mínimo 2 - máximo 5)

1. ESCOLHA UM DOS TÓPICOS ABAIXO

**ESTRESSE POR
FRIO**

DESENVOLVIMENTO

DÉFICIT HÍDRICO

**RESISTÊNCIA À
PATÓGENO**

**DEFICIÊNCIA
NUTRICIONAL**

ESTRESSE SALINO

2. ESCOLHA UMA CULTURA

QUESTÕES

2. **Discuta a relevância do tópico escolhido pelo grupo.**
3. **Elabore um ensaio para discutir os mecanismos genéticos envolvidos no estudo de caso selecionado. Leve em consideração:**
 - a) **Quais serão os tratamentos;**
 - b) **Qual será a ômica utilizada. Discorra sobre o porquê da escolha, destacando as vantagens e desvantagem da escolha;**
 - c) **Quais resultados você espera observar;**
 - d) **Cite baseado na literatura pelo menos duas moléculas (genes transcritos, proteínas, metabólitos) já descritos relacionados ao seu estudo de caso.**

ENTREGA – 16/10 (e-mail)

ESTUDO DIRIGIDO

1. Definição de ômicas;
2. Diferenças entre genômica, transcriptômica, proteômica e metabolômica;
3. Moléculas envolvidas em cada ômica;
4. Exemplos de aplicação dos estudos.

LEITURA RECOMENDADA

Era da ômicas

Genética Molecular Básica: Dos Genes aos Genomas

Carlos F.M. Menck & Marie-Anne Van Sluys
Capítulo 13

