



PROVA DIDÁTICA – PR-001

ABORDAGENS DIDÁTICAS NO ENSINO DE BIOINFORMÁTICA

CANDIDATO: DR. GABRIEL RODRIGUES COUTINHO PEREIRA

INSCRIÇÃO: 1701857

SUMÁRIO

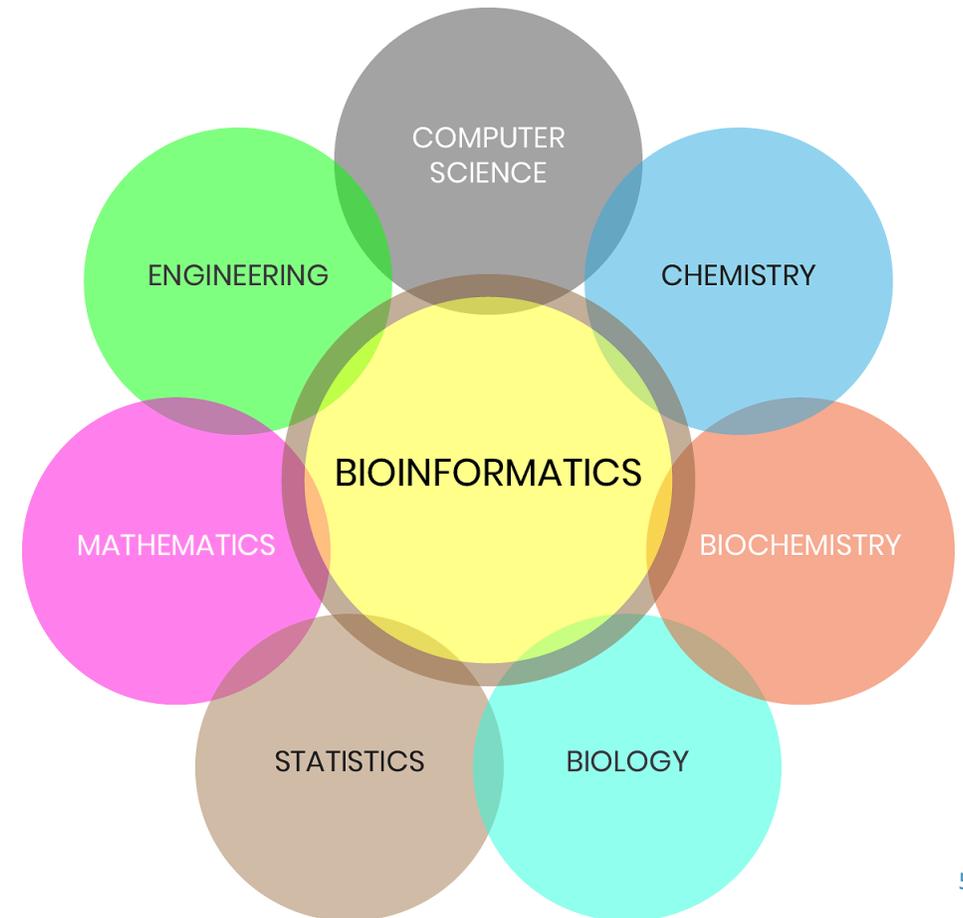
1. Introdução (definição, origem e aplicações);
2. Desafios no ensino de bioinformática;
3. Práticas comuns de ensino em bioinformática;
4. Metodologias ativas;
5. Aprendizagem assistida por tecnologia.



INTRODUÇÃO

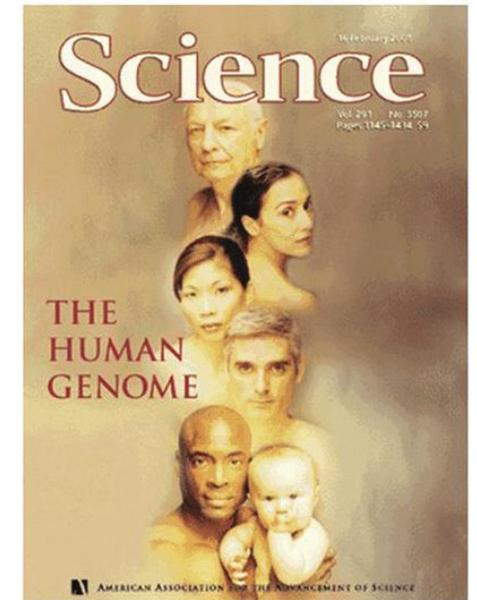
DEFINIÇÃO

- Campo interdisciplinar.
- Desenvolvimento e aplicação de ferramentas para gerenciamento, análise e interpretação de dados biológicos.
- Forma de experimentação: *in silico*.



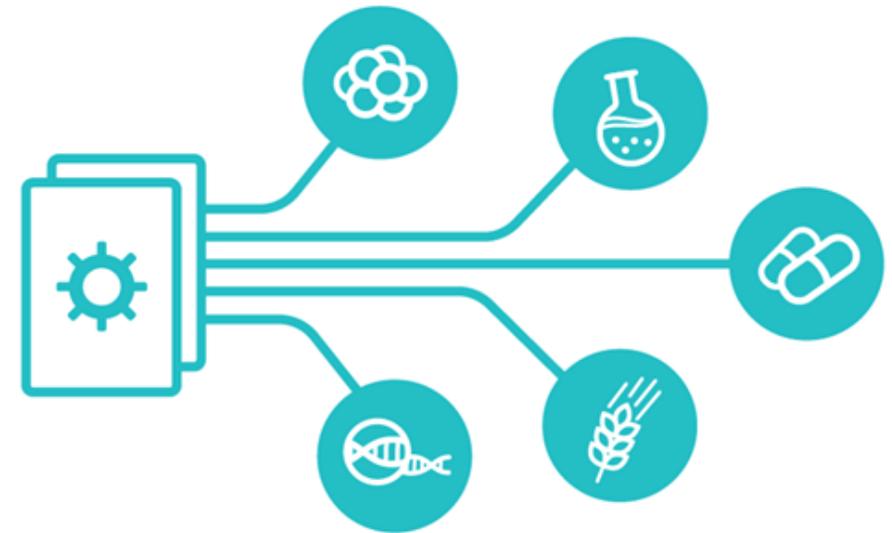
ORIGEM

- 1960': Atlas of Protein Sequence and Structure (Margaret Dayhoff);
- 1970': Emerge da necessidade de armazenar e interpretar dados de sequenciamento de DNA na década (1ª gen.);
- 1990': Surgimento do termo bioinformática;
- 2000's: Consolidação:
 - Projeto Genoma Humano (1990-2003);
 - Sequenciamento de nova geração (NGS);
 - Avanços no processamento computacional.



PRINCIPAIS APLICAÇÕES

- Construção de banco de dados biológicos;
- Montagem de genoma;
- Análise de sequências;
- Filogenética;
- Bioinformática estrutural;
- Análise da expressão gênica;
- Análise genética e populacional;
- Biologia de sistemas;
- Bioinformática de imagens;
- Planejamento de fármacos.



<http://www.support.illumina.com/content/illumina-marketing/en/informatics/specialized-bioinformatics-applications.html>

PRINCIPAIS DESAFIOS NO PROCESSO DE APRENDIZAGEM

- Campo interdisciplinar;
- Público-alvo com formação heterogênea;
- Deficiência na formação previa;
- Equilibrar o conteúdo para diferentes perfis;



Biólogo

x



Cientista da
computação

OUTROS DESAFIOS NO PROCESSO DE APRENDIZAGEM

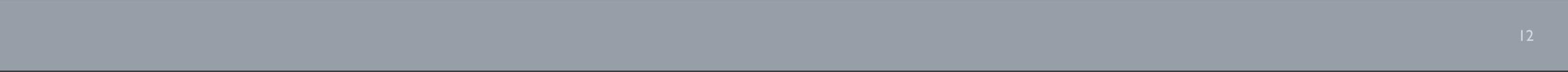
- Rápida evolução da área;
- Desatualização/descontinuação de ferramentas e protocolos;
- Reduzida disponibilidade de materiais didáticos prontos (e.g. livros textos).



COMO ENSINAR BIOINFORMÁTICA DE FORMA EFICIENTE?

INTENSO DEBATE NA LITERATURA...

- “Um currículo para bioinformática” – Altman (1998);
- “Educação em bioinformática: perspectivas e desafios” – Ranganathan (2005);
- “Uma perspectiva global sobre a evolução das necessidades de treinamento em bioinformática e ciência de dados.” – Attwood *et al.* (2019);
- “Conceitos Interdisciplinares e Transferíveis na Educação em Bioinformática” – Johnston, Slater e Cazier (2022);
- “Grandes Desafios na Educação e Formação em Bioinformática” – Işık *et al* (2023).



PRÁTICAS COMUNS DE ENSINO EM BIOINFORMÁTICA

AVALIAÇÃO DIAGNÓSTICA INICIAL – IDENTIFICAÇÃO DE LACUNAS

- Questões básicas de conhecimento prévio dos alunos em biologia, programação e estatística;
 - Biologia molecular (e.g., estrutura do DNA, transcrição e tradução, sequenciamento);
 - Programação (e.g., estrutura de dados, lógica de programação, sintaxe básica);
 - Estatística (e.g., conceitos de média, desvio padrão, distribuição normal).
- Identificar a formação acadêmica prévia dos estudantes;
- Mapear o quadro de disciplinas cursadas;



Google Forms



Microsoft Forms

NIVELAMENTO INICIAL



NIVELAMENTO INICIAL

Fundamentos em Biologia Molecular:

- Teórica: Aulas, vídeos explicativos, estudos dirigidos, livro texto e leitura complementar;
- Conteúdo: Estrutura e função de DNA, RNA e proteínas; replicação, transcrição e tradução; conceitos de genética e genômica, métodos de sequenciamento.

Introdução à Programação:

- Teórico-prático: Aulas teóricas atreladas à prática com exercícios de programação;
- Conteúdo: Lógica de programação, sintaxe básica, estruturas de dados (listas, dicionários, arrays), e noções de manipulação de arquivos de texto, ambiente Linux.

Estatística Básica:

- Teórico-prático: Aulas teóricas atreladas a práticas com uso de ferramentas em Python;
- Conteúdo: Conceitos de probabilidade, distribuições estatísticas, testes de hipóteses, e noções de análise exploratória de dados (**revisão**).

CURVA DE APRENDIZADO

- Início dificultado;
- Ambiente de aprendizado eficiente;
- Ferramentas *low-code* e *no-code*;
- Interface gráfica;
- Foco na aplicação prática;
- Motivação e engajamento.



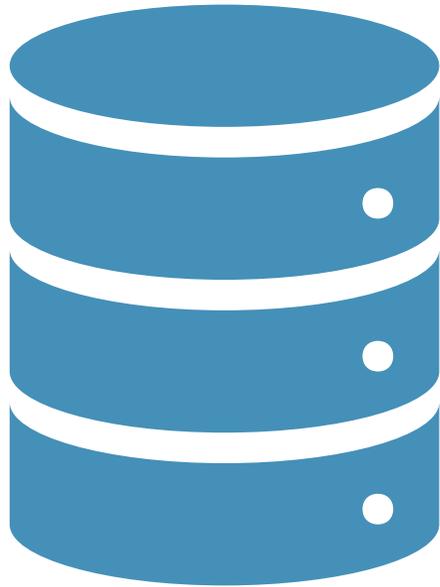
■ Atividades práticas:

- *learning by doing*;

■ Laboratórios multiusuários:

- e.g., LIGs (UFRJ).





INTRODUÇÃO AOS PRINCIPAIS BANCOS DE DADOS

LINGUAGENS DE PROGRAMAÇÃO E SCRIPT PRINCIPAIS

Python

Versátil, fácil aprendizado e amplo suporte

Biopython

Análise de dados de sequenciamento de DNA, RNA e proteínas.

R

Rico ecossistema para análise de dados genômicos

Bioconductor

Bash

Essencial para automação de tarefas e manipulação de *strings* em Linux:

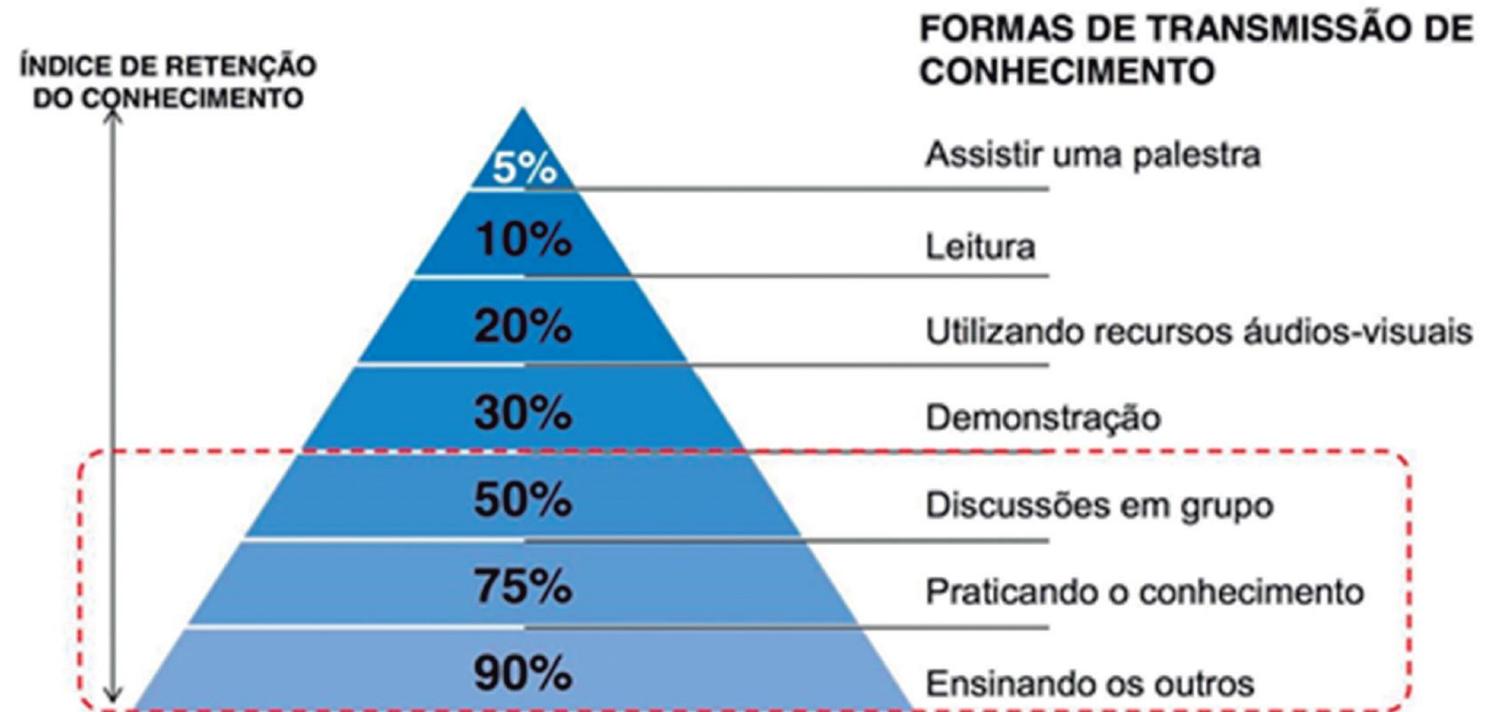
- Arquivos FASTA ou FASTQ.



ABORDAGENS DIDÁTICAS EM BIOINFORMÁTICA

■ Metodologias ativas:

- Discutir, praticar e ensinar;
- Aprendizado autônomo e participativo;
- Estimulam o pensamento crítico e a iniciativa;
- Alunos no centro do processo de aprendizagem;
- Construir o conhecimento.



https://www.researchgate.net/publication/326642446_A_Cultura_Maker_em_Prol_da_Inovacao_nos_Sistemas_Educacionais/figures?lo=1

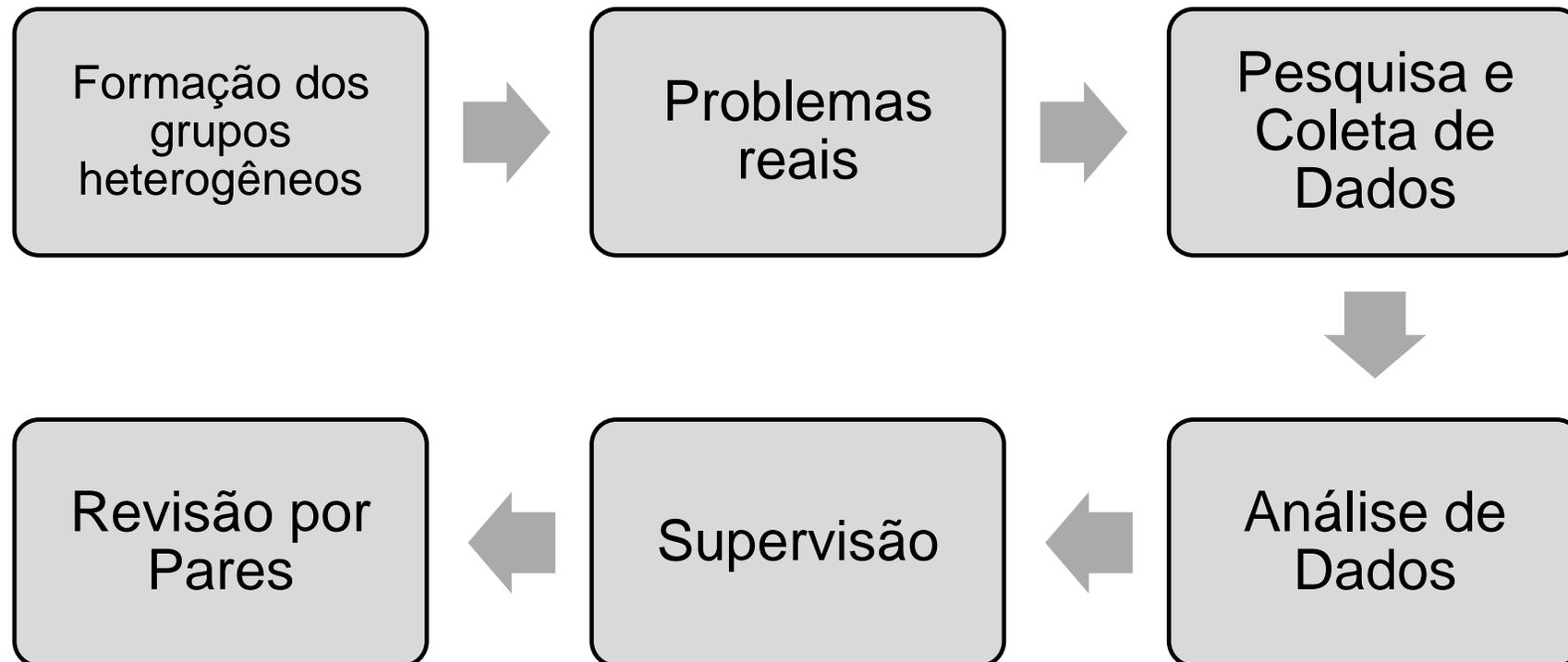
METODOLOGIAS ATIVAS COM APLICAÇÃO EM BIOINFORMÁTICA

1. Aprendizagem baseada em projetos;
2. Sala de aula invertida;
3. “*Gamificação*”;
4. Estudo de caso;
5. Seminários;
6. Discussões dirigidas.

Complementares



1. APRENDIZAGEM BASEADA EM PROJETOS

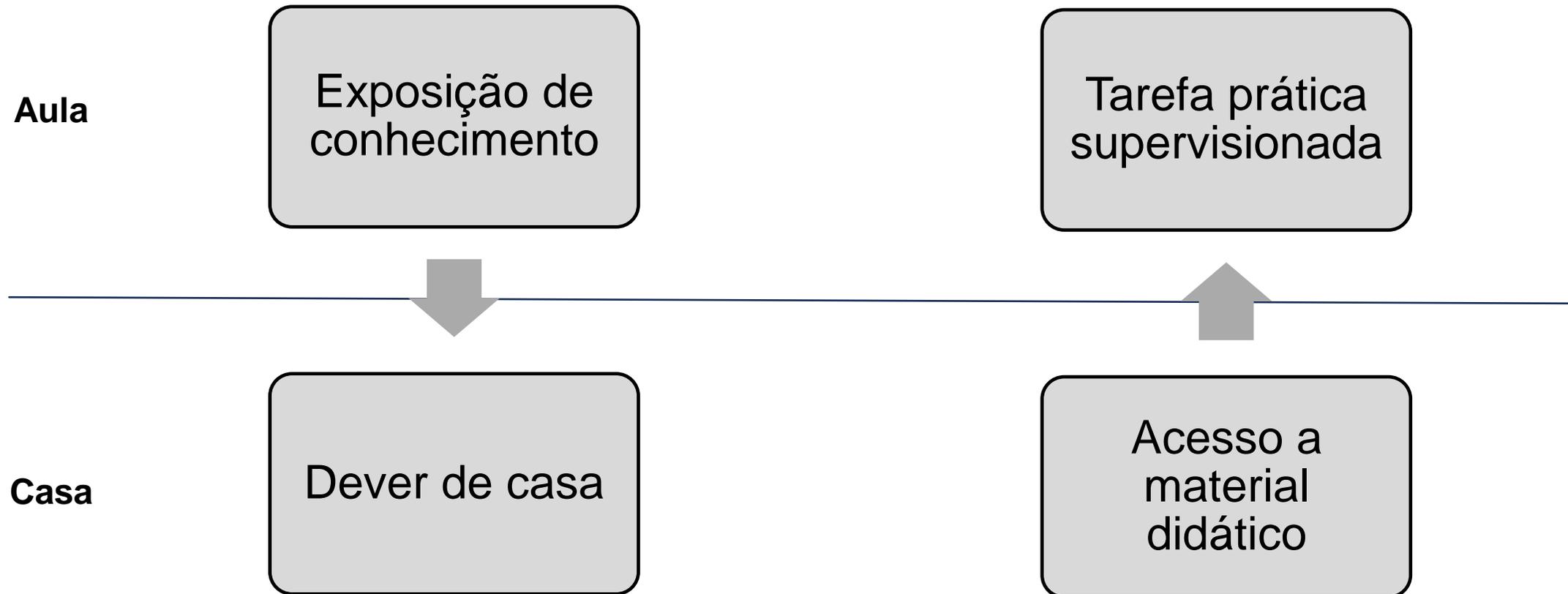


2. SALA DE AULA INVERTIDA

- Conteúdo teórico;
- Tutoriais;
- Prática simplificada.



SALA DE AULA TRADICIONAL vs SALA DE AULA INVERTIDA



3. “GAMIFICAÇÃO”

- Aplicação de elementos de jogos;
- Desafios de bioinformática (e.g. alinhamento de sequência);
- Recompensa atribuída;
- Estimula a competição/colaboração;
- **Fornecer *feedback* imediato;**
- ❖ *Quizzes* interativos: Kahoot!



<https://kahoot.it/>



KAHOOT!

What percentage of teens said that it's important for them to follow the news and current events?

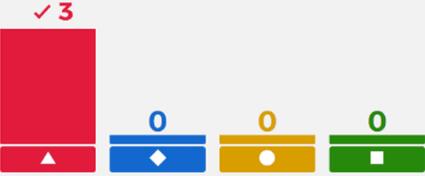


15

0 Answers

▲ 14%	◆ 33%
● 52%	■ 78%

What is the name of this platform?



3

0

0

0

Show media

End game

▲ Kahoot! ✓	◆ Quizzoot! ✗
● Toohak! ✗	■ Kangaroo! ✗

1/1

kahoot.it Game PIN: 7292927

Fixação de
conteúdo:

ao final da
aula;

Atrair
atenção (“quebra”):

intercalado
com conteúdo
teórico denso.

SUGESTÕES DE QUANDO APLICAR “GAMIFICAÇÃO”

ESTUDO DE CASO + SEMINÁRIO + DISCUSSÃO

1

- Artigo científico em bioinformática;

2

- Proposição de uma tarefa específica:
 - apontar falhas no artigo;
 - aplicação de metodologias ensinadas.

3

- Apresentação de seminário;

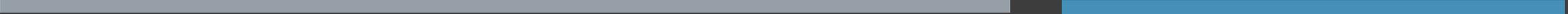
4

- Justificativa e argumentação;

5

- Supervisão e discussão.

Modelagem molecular (FF-UFRJ)

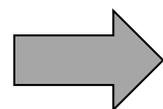


APRENDIZAGEM ASSISTIDA POR TECNOLOGIA

AMBIENTE VIRTUAL DE APRENDIZAGEM (AVA)



Ambiente tradicional



AVA

Fóruns de discussão;
Quizzes (fixação de conteúdo);
Slides da aula;
Tutoriais;
Arquivos compartilhados (e.g. inputs);
Calendários.



Centralização



Information

Flag question

Multiple Choice

One of the options within Moodle is the 'Multiple Choice' question. This can be used in a variety of different ways as shown below.

Question 1

Not complete

Marked out of 1.00

Flag question

Simple textual multiple choice

One of the easiest question types is a multiple choice question, which has 1 correct answer and many wrong 'distractor' answers.

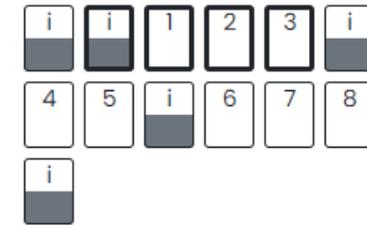
Moodle is an example of a VLE. What does the acronym VLE stand for?

Select one:

- a. Very Long Experience
- b. Virtual Learning Experience
- c. Virtual Learning Environment
- d. Virtual Lesson Engine

Check

Quiz navigation

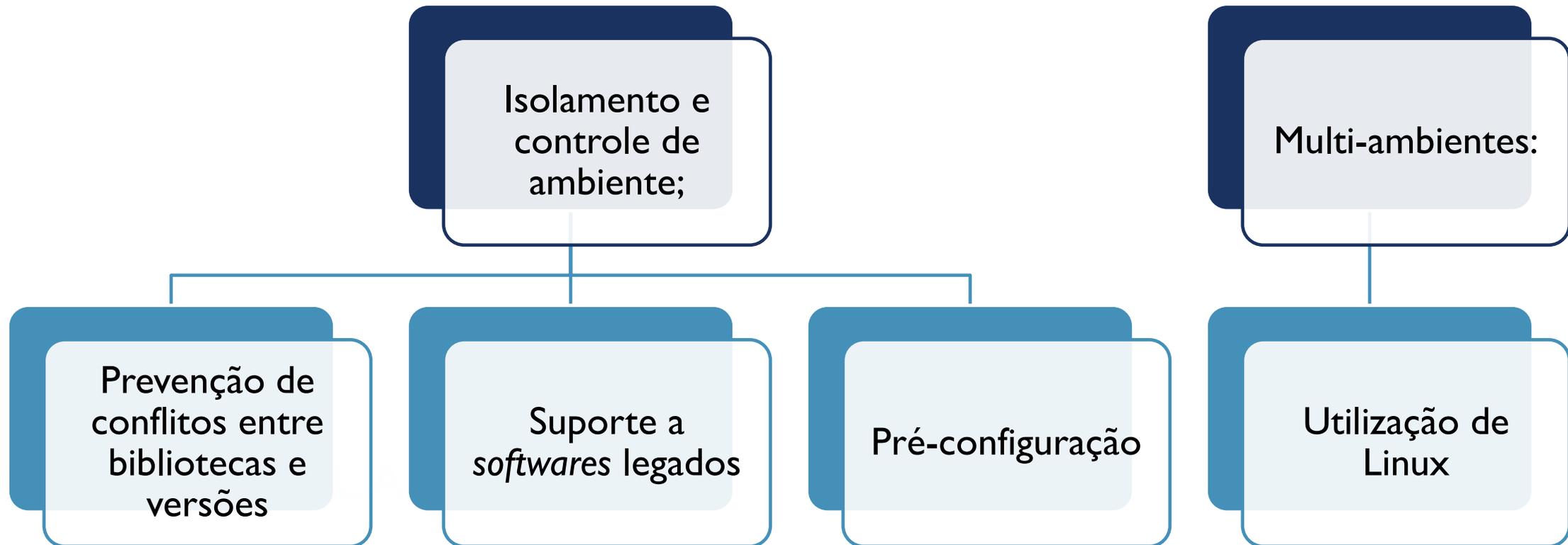


Finish attempt ...

VIRTUALIZAÇÃO



<https://www.oracle.com/br/virtualization/virtualbox/>



APRENDIZADO BASEADO EM NOTEBOOKS

- Execução modular do código (bloco-a-bloco);
- Integração entre o código, explicação e visualização;
- Registra a lógica do desenvolvimento;
- Compatível com Python, R e Julia.

HSPFragment {#sec:searchio-hspfragment}

`HSPFragment` represents a single, contiguous match between the query and hit sequences. You could consider it the core of the object model and search result, since it is the presence of these fragments that determine whether your search have results or not.

In most cases, you don't have to deal with `HSPFragment` objects directly since not that many sequence search tools fragment their HSPs. When you do have to deal with them, what you should remember is that `HSPFragment` objects were written with to be as compact as possible. In most cases, they only contain attributes directly related to sequences: strands, reading frames, alphabets, coordinates, the sequences themselves, and their IDs and descriptions.

These attributes are readily shown when you invoke `print` on an `HSPFragment`. Here's an example, taken from our BLAST search:

```
In [67]: from Bio import SearchIO
blast_qresult = SearchIO.read('data/my_blast.xml', 'blast-xml')
blast_frag = blast_qresult[0][0][0] # first hit, first hsp, first fragment
print(blast_frag)

Query: gi|8332116|gb|BE037100.1|BE037100 MP14H09 MP Mesembryanthemum cr...
Hit: gi|731339628|ref|XM_010682658.1| PREDICTED: Beta vulgaris subsp...
Query range: [86:679] (1)
Hit range: [80:677] (1)
Fragments: 1 (597 columns)
Query - TTGGCCATGAAAACGTGATCAATTGGCCGTGGCTAATATGATCGATTCCGATATCAATGA~~~TGTAG
      |
Hit - TTGGCCATGAAAACGTGAGCAAATGGCGTTGGCTAATTTGATAGATTATGATATGAATGA~~~TGTAG
```



GOOGLE COLAB



The screenshot displays the Google Colab interface. On the left, a code cell contains Python code for checking available RAM. The code is as follows:

```
[ ] 1 from psutil import virtual_memory
2 ram_gb = virtual_memory().total / 1e9
3 print('Your runtime has {:.1f} gigabytes of available RAM\n'.format(ram_gb))
4
5 if ram_gb < 20:
6     print('Not using a high-RAM runtime')
7 else:
8     print('You are using a high-RAM runtime!')
```

Below the code cell, there is a section titled "Ambientes de execução com maior duração" (Execution environments with longer duration). The text explains that all Colab execution environments are redefined after a period, which is shorter if the environment is not running code. Colab Pro and Pro+ users have access to longer execution environments compared to the standard Colab.

On the right side of the interface, a "Recursos" (Resources) panel is visible, showing the following information:

- RAM do sistema: 1.2 / 12.7 GB
- Disco: 28.8 / 225.8 GB

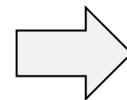
There are also two small line graphs below the resource usage, one for RAM and one for the disk, showing usage over time.

- Computação em nuvem;
- Recursos sob demanda;
- Utilização de GPUs.

FERRAMENTAS DE VISUALIZAÇÃO

		Amino Acid		Chain name		-----Coordinates-----		
		Element		Sequence Number		X	Y	Z
ATOM	1	N	ASP L	1		4.060	7.307	5.186
ATOM	2	CA	ASP L	1		4.042	7.776	6.553
ATOM	3	C	ASP L	1		2.668	8.426	6.644
ATOM	4	O	ASP L	1		1.987	8.438	5.606
ATOM	5	CB	ASP L	1		5.090	8.827	6.797
ATOM	6	CG	ASP L	1		6.338	8.761	5.929
ATOM	7	OD1	ASP L	1		6.576	9.758	5.241
ATOM	8	OD2	ASP L	1		7.065	7.759	5.948

//
Element position within amino acid



PyMOL Viewer

1E/101 106 111 116 /F/101 all A S H L C

DC DC DC DG DT DG DC DA DC DT DC DA DA DT DG DC DA DA DT DG 1B01

Actions:

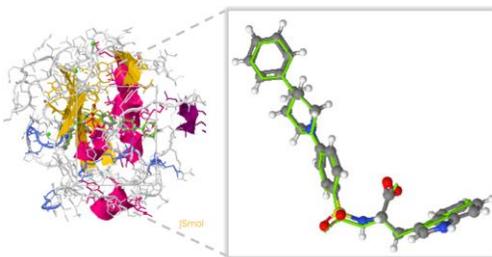
- zoom
- orient
- center
- origin
- drag
- preset
- find
- align
- generate
- assign sec. struc.
- rename object
- duplicate object
- delete object
- hydrogens
- remove waters
- state
- masking
- sequence
- movement
- compute
- MovH FKHC FKTB MVSZ
- Ctrl MvAZ Orig Clip MovZ
- CtSh PkAt Cent Menu
- SnglClk MovA DrgM PKTB
- Db1Clk

SERVIDORES ONLINE

<https://dockthor.Incc.br/v2/>



Rank	Compound	Score	T. Energy	RMSD
1	ligand_d227b0f4db_1	-10.526	-17.080	0.498
11	run28_model1	-10.526	-17.080	0.498
12	run13_model2	-9.916	-15.655	3.088
13	run8_model8	-9.910	-14.596	2.310
14	run14_model12	-9.523	-13.062	2.242
15	run22_model16	-9.823	-11.899	4.023
16	run24_model24	-9.557	-9.017	3.038
17	run1_model26	-9.626	-8.830	3.513
18	run27_model23	-9.425	-6.470	2.317
19	run15_model22	-9.395	-5.727	4.663



A screenshot of the DockThor web interface. The top navigation bar includes "Home", "Docking", "References", "About", and "Support". The main content area is titled "Upload your protein file" and features a "New COVID-19 resources" section with dropdown menus for "Target", "Variant", and "Structure". There are also buttons for "Add file" and "Select Test File".

- Interface gráfica e intuitiva;
- Não necessitam de instalação;
- Integra visualização e execução.

SERVIDORES ONLINE

<https://dockthor.Incc.br/v2/>



- Independe da infraestrutura local

RECAPITULANDO

1. Introdução (definição, origem e aplicações);
2. Desafios no ensino de bioinformática;
3. Práticas comuns de ensino em bioinformática;
4. Metodologias ativas;
5. Aprendizagem assistida por tecnologia.

LEITURA COMPLEMENTAR

- Conceitos Interdisciplinares e Transferíveis na Educação em Bioinformática” – Johnston, Slater e Cazier (2022).
<https://www.frontiersin.org/journals/education/articles/10.3389/feduc.2022.826951/full#B5>
- Ensinando bioinformática por meio da análise do SARS-CoV-2 – Poličar *et. al* (2024).
https://academic.oup.com/bioinformatics/article/40/Supplement_1/i20/7700875

“Feliz é aquele que transfere o que sabe e aprende o que ensina.”

Cora Coralina



DÚVIDAS?