



Mais informações:

<https://labbc.es.cena.usp.br/project/sagb/>

Seminário de Afinidades em Genômica e Bioinformática (SAGB)

28 de maio de 2026



Introdução ao seminário

<https://labbc.es.netlify.app/project/sagb/>

Agenda

Periodicidade e local: A cada dois meses approx, na última semana do mês. Um encontro no CENA/USP, o seguinte na ESALQ/USP, e repete.

Metodologia: Em cada encontro duas pessoas liderarão a discussão do artigo selecionado. No final de cada encontro serão definidos os líderes do próximo encontro, tentando que sejam voluntários.

Objetivos

- Criar um espaço de conversa informal e aberto a todos os alunos (graduação e pós-graduação) e pesquisadores do campus LQ, e pessoas externas, em temas relacionados a Genômica e Bioinformática.
- Servir como um instrumento para difundir o interesse e conhecimento de Genômica e Bioinformática no campus LQ
- Fomentar as relações inter-pessoais entres pesquisadores e alunos das duas unidades do campus LQ: CENA e ESALQ

Organização

- Prof. Dr. Claudia Vitorello (ESALQ)
- Prof. Dr. Douglas Silva Domingues (ESALQ)
- Prof. Dr. Diego M. Riaño-Pachón (CENA)



OXFORD

Briefings in Bioinformatics, 2025, 26(6), bbaf576

<https://doi.org/10.1093/bib/bbaf576>

Review

Artificial intelligence in bioinformatics: a survey

Jiyue Jiang^{1,2,‡}, Yunke Li^{1,3,‡}, Shiwei Cao^{1,4,‡}, Yuheng Shan⁵, Yuexing Liu¹, Tianyi Fei^{1,4}, Yule Yu⁶, Yi Feng⁷, Yu Li^{2,8,*}, Yixue Li^{1,7,*},
Jiao Yuan^{1,7,*}

1950's

Artificial intelligence (AI)

Human intelligence exhibited by machines

1980's

Machine learning

AI systems that learn from historical data

2010's

Deep learning

Machine learning models that mimic human brain function

2020's

Generative AI (Gen AI)

Deep learning models (foundation models) that create original content

DATA DELUGE

The billions of terabytes (TB) produced in one year by the SKA telescope (grey) will dwarf today's data sets in genomics and climate science.

Encyclopedia of DNA Elements (ENCODE), 2012

15 TB

US National Climate Assessment (NASA projects), 2013

1,000 TB

Fifth assessment report by the Intergovernmental Panel on Climate Change (IPCC), due 2014

2,500 TB

Square Kilometre Array (SKA), first light due 2020

22,000,000,000 TB per year

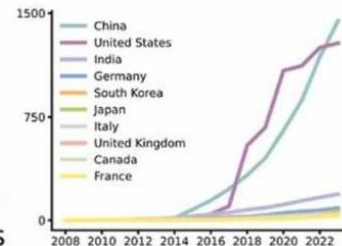
<https://www.nature.com/articles/493473a>

Quantidades massivas de dados sendo geradas em todas as áreas da ciência

The Deluge of Data

- Databases such as the **Sequence Read Archive (SRA)** and **Gene Expression Omnibus (GEO)** play pivotal roles in stress studies across the plant kingdom.

- SRA houses a total of **558,952 transcriptomic** and **1,387,717 genomic** individual accessions for Viridiplantae at the time of writing.



Number of transcriptomic stress-related BioProjects in SRA have shot up in recent years, in China and US particularly

- Metadata includes project's **objectives**, **experimental design** and **data sources**.

<https://pmc.ncbi.nlm.nih.gov/articles/PMC11480249/>

Figure 1 AI in bioinformatics leverages machine learning, reinforcement learning, and deep learning

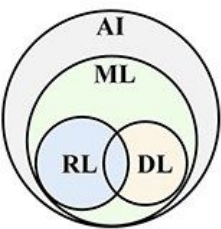
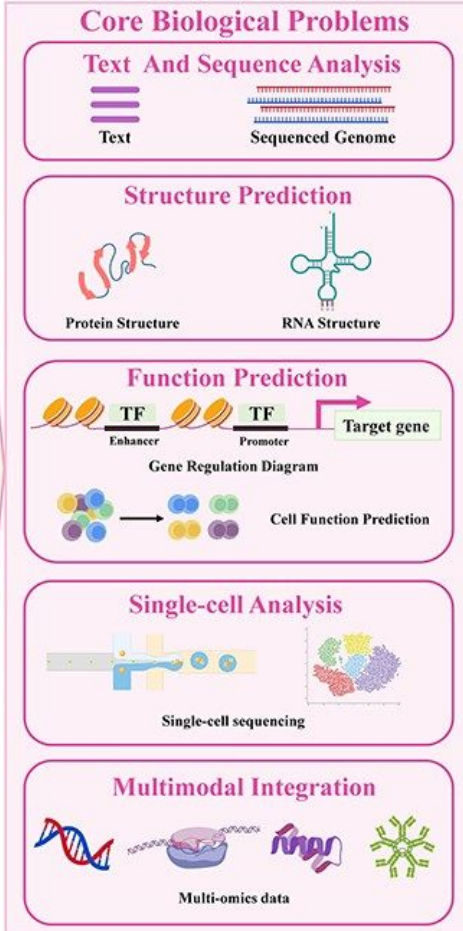
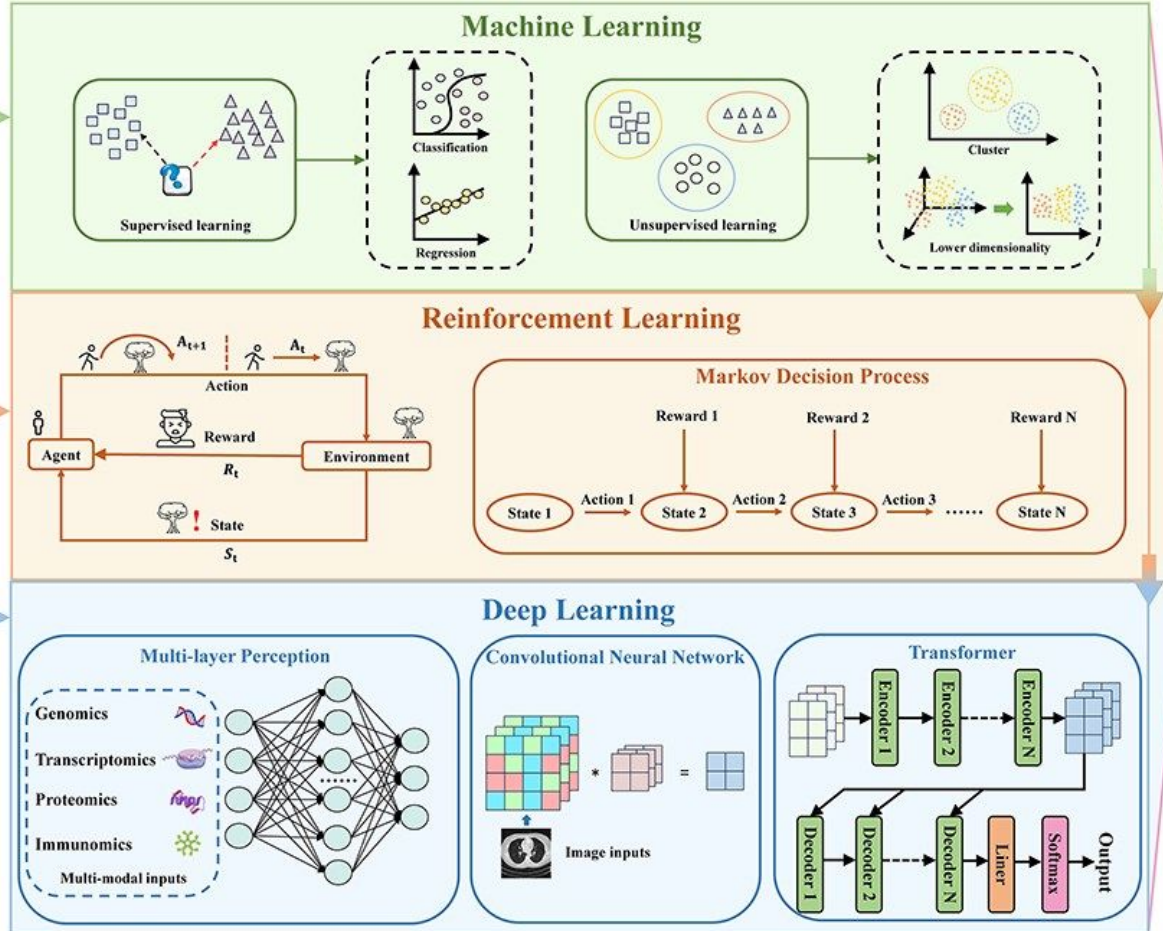
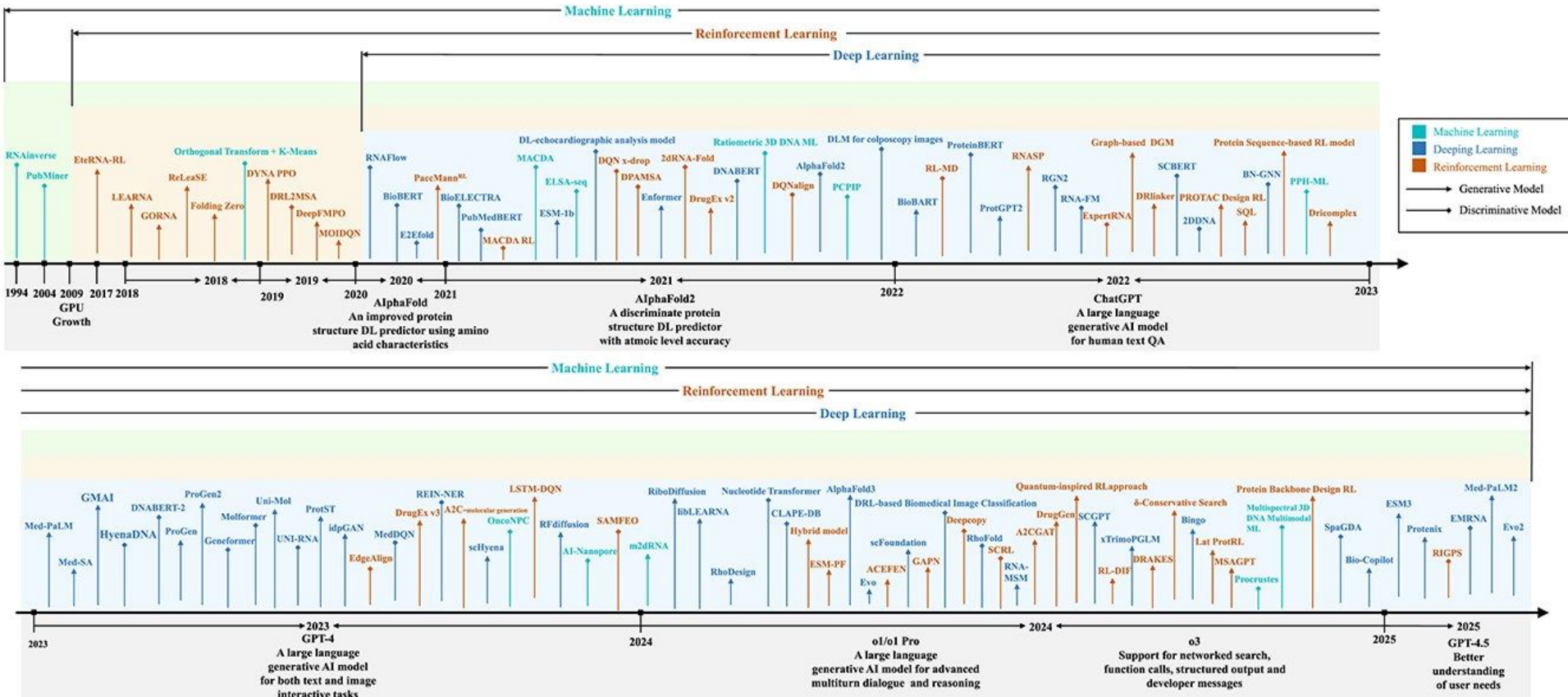
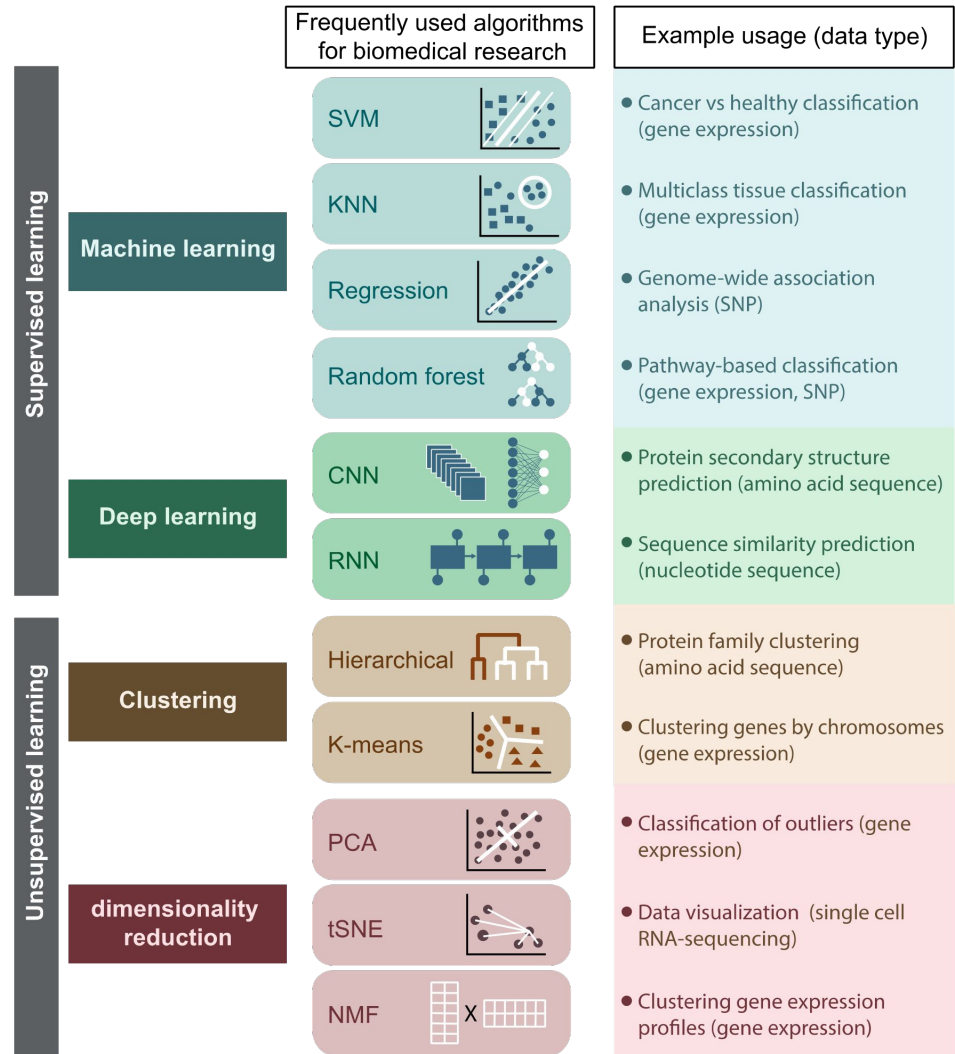


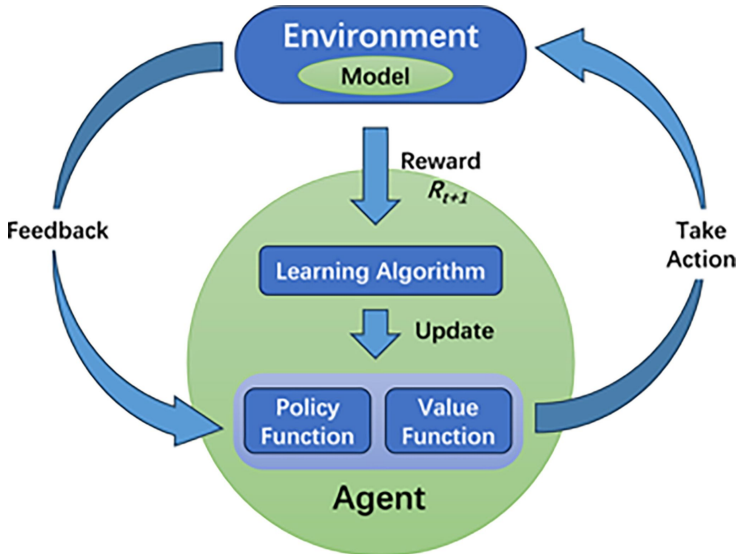
Figure 2 The evolution of AI in bioinformatics from 2004 to 2025



Aprendizado de máquina em Bioinformática

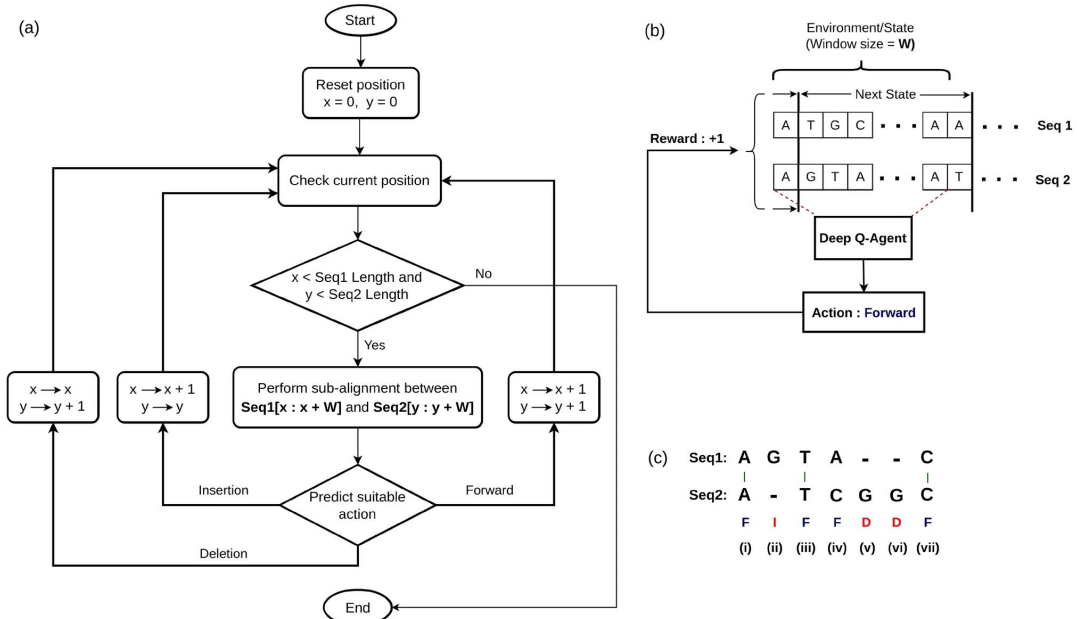


Aprendizado de reforço em Bioinformática



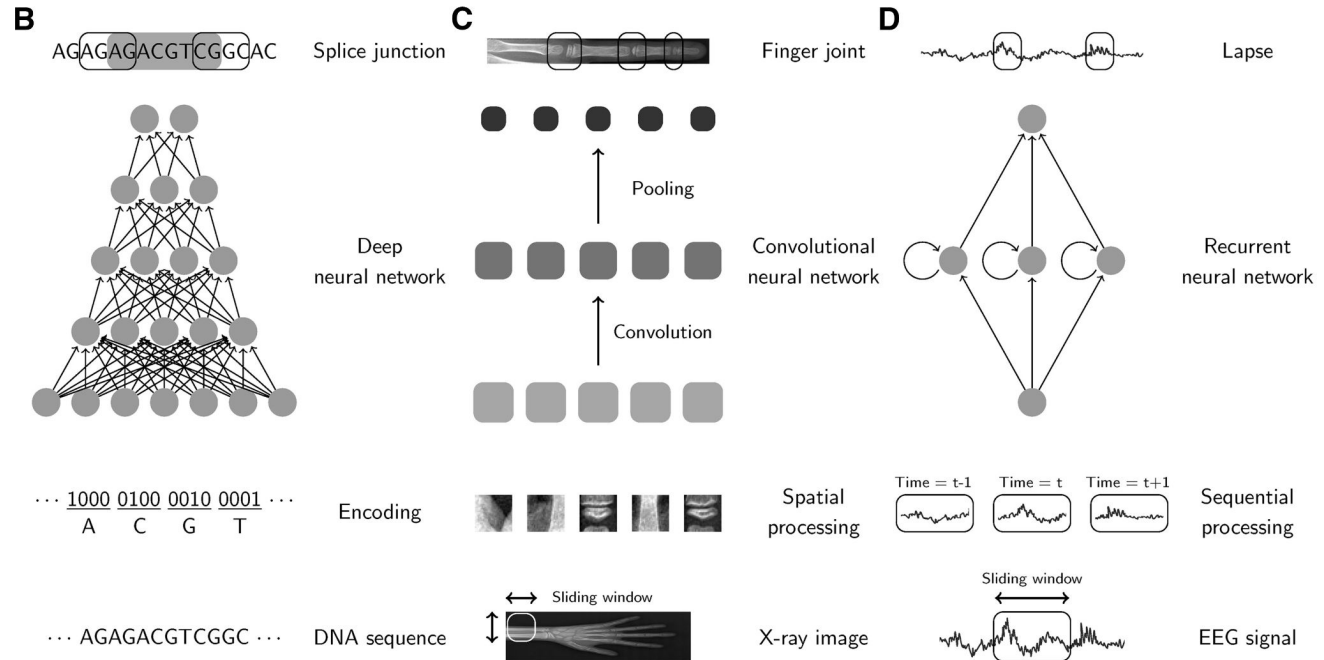
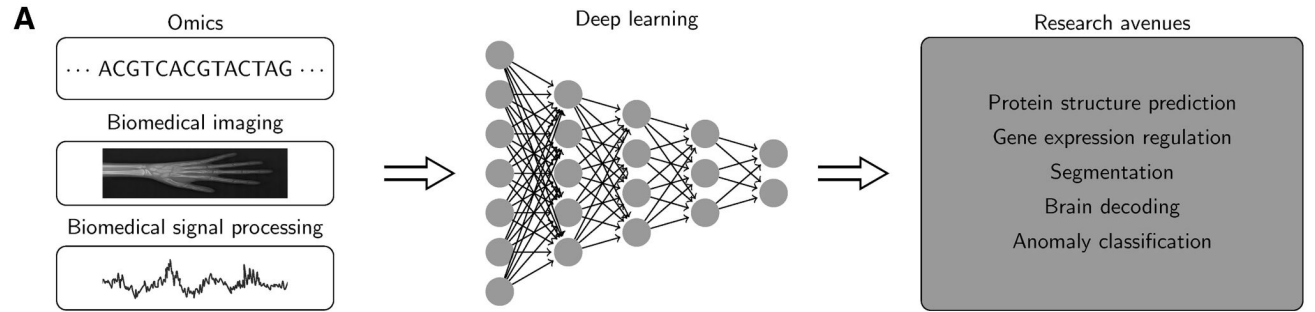
<https://wires.onlinelibrary.wiley.com/doi/full/10.1002/wcms.1723>

Alinhamento de pares de sequências



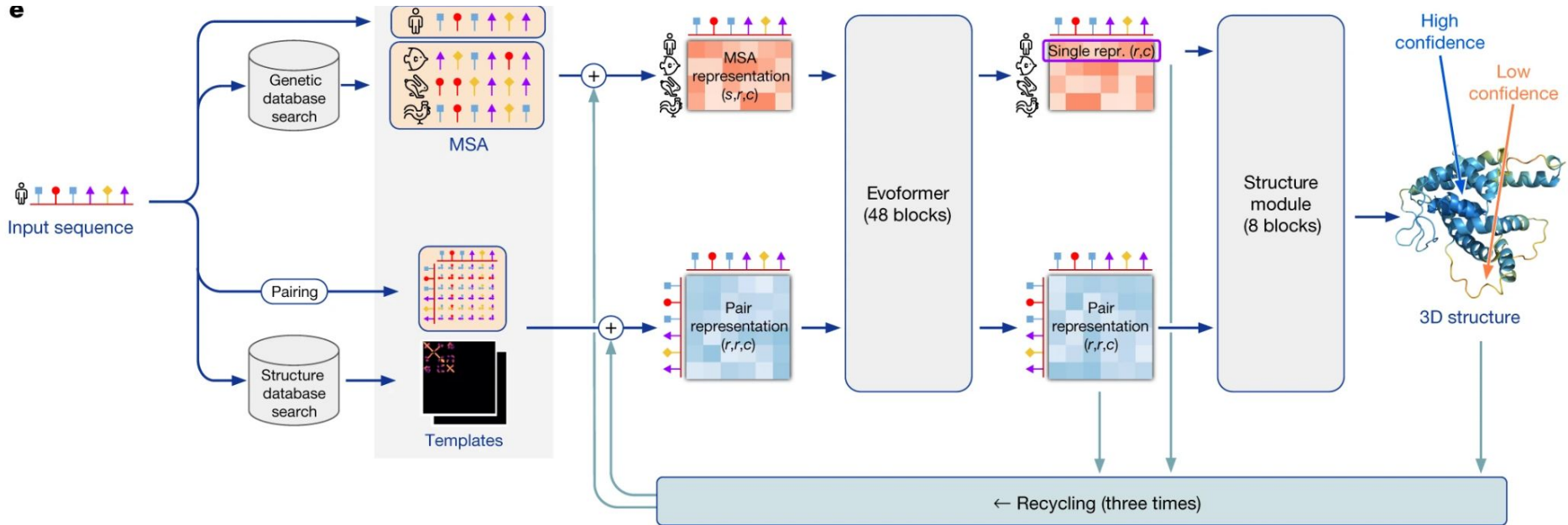
<https://www.nature.com/articles/s41598-023-29277-6>

Aprendizado Profundo em Bioinformática



Aprendizado Profundo em Bioinformática

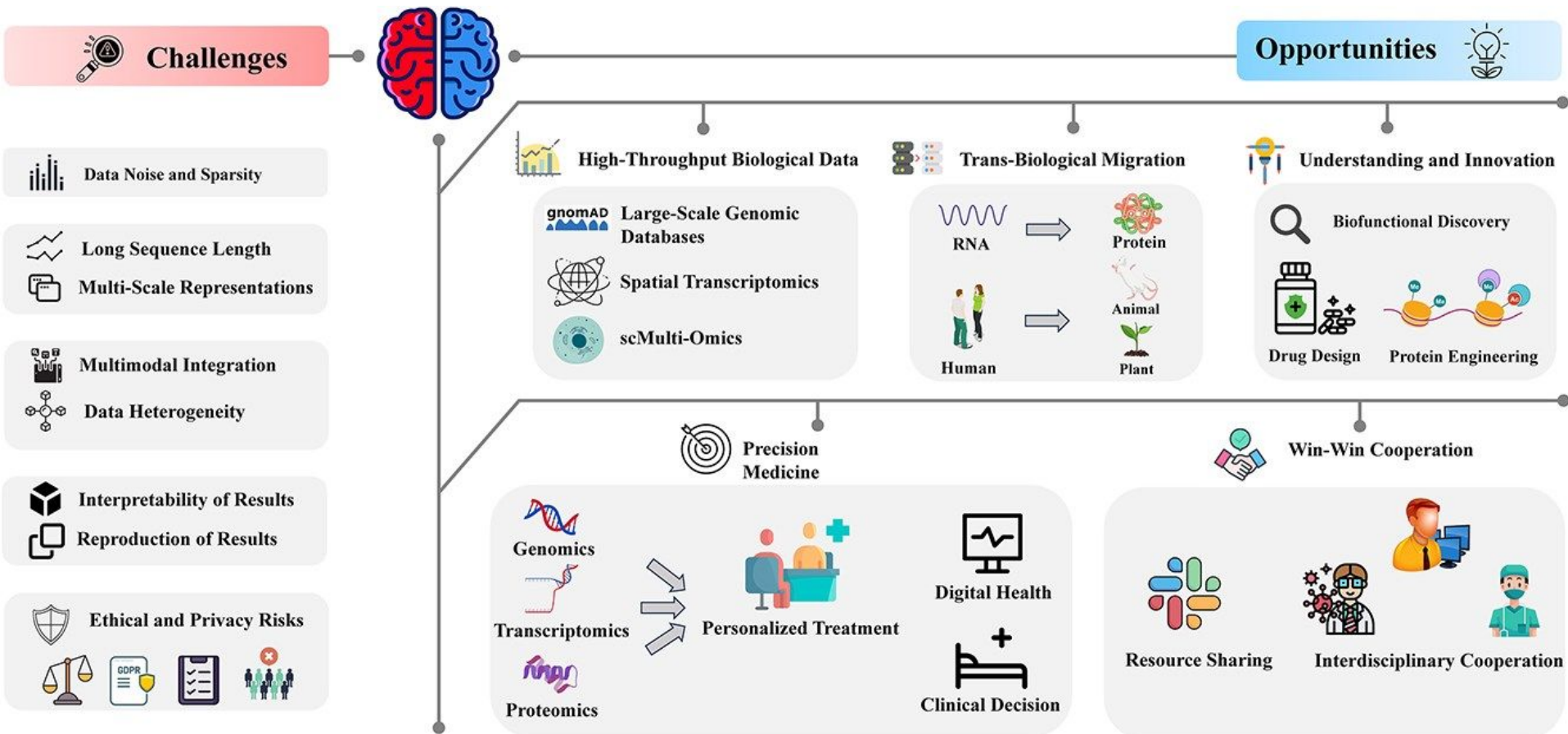
AlphaFold2



Limitaciones técnicas

- Infraestrutura (CPU vs GPU)
- Recursos computacionais: Velocidade e relação com tamanho das "perguntas"
- Tempo para treinamento
- Governança de dados e validação

Desafios



Seminário de Afinidades em Genômica e Bioinformática (SAGB)

Obrigado!

