Biologia e Computação: Um Casamento Perfeito

By Editor-in-chief

14 de maio de 2021

Biologia e Computação: Um Casamento Perfeito

Fernanda Stussi D Lage , Fenícia Brito

Revisão: Diego Mariano

BIOINFO – Revista Brasileira de Bioinformática. Edição #01. Julho, 2021.

DOI: [10.51780/978-6-599-275326-02](https://doi.org/10.51780/978-6-599-275326-02)

*“O QUE VOCÊ GANHA QUANDO MISTURA UM
CIENTISTA DA COMPUTAÇÃO COM UM
BIÓLOGO?”*

Ou, para ser mais preciso, o que você ganha quando eles colaboram? Na verdade, você obtém mal-entendidos: dores de cabeça com novas terminologias ou significados diferentes para termos existentes e, às vezes, até uma incapacidade total de entender os termos um do outro [2]. Cada vez mais, essa “mistura” vai além de uma equipe multidisciplinar, podendo ser visto em um único cientista, o Bioinformata.

Os primeiros bioinformatas são anteriores ao aparecimento do termo “Bioinformática” – da mesma forma que, desde 1856, Mendel se utilizava do conhecimento da física e da botânica, duas outras ciências da época, para estudar o que hoje conhecemos como genética. Apesar de não utilizarem o termo “bioinformática” para descreverem seus trabalhos, esses profissionais tiveram uma visão clara de como informática, matemática e a biologia molecular poderiam ser combinadas para responder a perguntas fundamentais nas ciências da vida. Dessa forma, esses cientistas construíram conceitos importantes e fundamentos técnicos para a bioinformática [6].

Ainda na década de 1950, já haviam sido publicados artigos relacionados à Biologia que empregaram computadores eletrônicos digitais em uma simulação de deriva genética [3]. Enquanto isso, o termo 'Bioinformática' em si aparece pela primeira vez apenas na década de 1970.

Como em um bom casamento, os computadores não apenas ampliaram a biologia, eles trouxeram consigo ferramentas e questões completamente novas – conforme observado na Figura 1, como estatísticas, simulação e gerenciamento de dados, que remodelaram completamente a forma como a pesquisa biológica está sendo feita [6].

Figura 1. Bioinformática como uma ciência multidisciplinar.
Adaptado de Institute for Systems Biology [4].

Partindo da ótica da biologia molecular, o motivo pelo qual esse casamento poderia ser considerado perfeito é muito simples: apesar da estrutura do DNA ter sido desvendada em 1953, a informação nela contida não podia ser “lida”. Foi como se tivéssemos descoberto o alfabeto utilizado para escrever “o livro da vida”, mas as “palavras” desse livro estavam com letrinhas tão pequenas, que não éramos capazes de lê-las. Foi preciso esperar até fins da década de 1980 para que aparecesse uma “lente de aumento” suficientemente boa (e automática – na forma de uma máquina) que permitisse a leitura dessas letrinhas em grandes quantidades. Em 1995, uma única máquina dessas já conseguia ler milhares de letrinhas por dia. Do lado da computação, foi também preciso um amadurecimento.

Esse amadurecimento é a tão falada revolução da informática, com computadores sendo capazes de armazenar cada vez mais informação, processá-la de modo cada vez mais rápido, a um custo cada vez menor. É interessante observar que se o sequenciamento automático do DNA tivesse amadurecido mais rapidamente, digamos com 20 anos de antecedência, não haveria computadores com poder suficiente para dar conta dos dados gerados. Na década de 1970, a unidade básica de armazenamento de informação era o kilobyte – 1024 bytes, aproximadamente 1000 letras. Um computador de grande porte daquela época tinha alguns kilobytes de

memória. Com tal memória um computador desses não seria capaz de processar nem sequer o genoma de um vírus de aproximadamente 20 kilobases (ou 20 mil letrinhas), que dirá o genoma humano, com seus 3 bilhões de letrinhas [1].

A evolução conjunta dessas duas ciências também pode ser observada quanto à quantidade e a integração de dados que podem ser estudados. Enquanto o sequenciamento de uma única proteína ou gene poderia ter sido objeto de uma tese de doutorado até o início dos anos 90, um estudante de doutorado agora pode analisar o genoma coletivo de muitas comunidades microbianas durante seus estudos de pós-graduação. Hoje em dia, podemos modelar computacionalmente organismos inteiros – com todas as suas relações metabólicas – e sua interação com outros seres vivos e o ambiente – com todas as categorias moleculares consideradas simultaneamente.

Inclusive, se Carl von Linné (também conhecido como Carl Linnaeus), o botânico sueco e pai da taxonomia, vivesse hoje, ele seria um bionformata. Como botânico, ele poderia criar uma base de dados para organizar o que sabemos sobre os genótipos e fenótipos de culturas e plantas modelo. Ou ele poderia trabalhar com o *Gene Ontology Consortium* (<http://www.geneontology.org/>) para criar vocabulários compartilhados que unificariam o conhecimento biológico entre os organismos. Assim como o *Systema Naturae de Linné*, esses bancos de dados são contribuições intelectuais importantes para a nossa compreensão da vida.

Ao misturar a Biologia com a Ciência da Computação, se combinam grandes coleções de dados com bancos de dados e estatísticas, um mapa de referência para a biologia. Esse mapa, por um lado, se parece com o *Google Street View*, permitindo observar em *zoom* cada subárea da Biologia – como, por exemplo, a bioquímica estudando as reações químicas que acontecem no interior de uma célula – e de longe permitindo observar as interações dessas várias subáreas – buscando entender a quantidade de dados e sequências estruturais que foram gerados em vários níveis de sistemas biológicos.

Por outro lado, esse mapa, não está no nível de resolução fornecido pelo *Google Street View*. Ao contrário, é um mapa como as cartas náuticas utilizadas por Colombo, fornecendo um esboço geral, mas muitas áreas estão incompletas e algumas partes importantes podem estar ausentes e aguardando para serem descobertas. Mas mesmo com todas essas deficiências, o mapa ainda é um guia indispensável. O atlas da vida, fornecido pela biologia computacional, forma o pano de fundo para o planejamento, execução e interpretação de todos os experimentos de pequena escala que investigam áreas ainda não mapeadas visando expandir os limites de conhecimento biológico [6].

Referências

[1] ARAÚJO, Nilberto Dias de; FARIAS, Rodrigo Pessoa de; PEREIRA, Patrícia Barbosa; FIGUEIRÊDO, Flávia Mota de; MORAIS, Alanna Michely Batista de; SALDANHA, Livina Costa; GABRIEL, Jane Eyre. A era da bioinformática: seu potencial e suas implicações para as ciências da saúde. **Estudos de Biologia**, [s.l.], v. 30, n. 70/72, p. 143-148, 27 nov. 2008. Pontifícia Universidade Católica do Paraná – PUCPR. <http://dx.doi.org/10.7213/rev.v30i70/72.22819>.

[2] BENTLEY, Peter J.. Why Biologists and Computer Scientists Should Work Together. **Lecture Notes In Computer Science**, [s.l.], p. 3-15, 2002. Springer Berlin Heidelberg. http://dx.doi.org/10.1007/3-540-46033-0_1.

[3] DINIZ, W. J. S.; CANDURI, F. REVIEW-ARTICLE Bioinformatics: an overview and its applications. **Genetics and Molecular Research**, v. 16, n. 1, 2017.

[4] GAUTHIER, J. et al. A brief history of bioinformatics. **Briefings in Bioinformatics**, v. 20, n. 6, p. 1981-1996, 27 nov. 2019.

[5] HAGEN, J. B. The origins of bioinformatics. **Nature Reviews Genetics**, v. 1, n. 3, p. 231-236, dez. 2000.

[6] MARKOWETZ, Florian. All biology is computational biology. **Plos Biology**, [s.l.], v. 15, n. 3, p. 1-4, 9 mar. 2017. Public Library of Science (PLOS). <http://dx.doi.org/10.1371/journal.pbio.2002050>.

[7] What is Systems Biology? Institute for Systems Biology. Disponível em: <<https://isbscience.org/about/what-is-systems-biology/>>. Acesso em 01 de maio de 2020.