



---

# A ERA DA BIOINFORMÁTICA: SEU POTENCIAL E SUAS IMPLICAÇÕES PARA AS CIÊNCIAS DA SAÚDE

*The effects and applications of bioinformatics on the biomedical area*

Nilberto Dias de Araújo<sup>[a]</sup>, Rodrigo Pessoa de Farias<sup>[a]</sup>, Patrícia Barbosa Pereira<sup>[a]</sup>, Flávia Mota de Figueirêdo<sup>[a]</sup>, Alanna Michely Batista de Moraes<sup>[a]</sup>, Livina Costa Saldanha<sup>[a]</sup>, Jane Eyre Gabriel<sup>[b]</sup>

<sup>[a]</sup> Acadêmico do curso de Bacharelado em Biomedicina ofertado pelas Faculdades Integradas de Patos (FIP), Patos, Paraíba, PB - Brasil.

<sup>[b]</sup> Professora adjunto I do Centro de Ciências e Tecnologia Agroalimentar (CCTA), Universidade Federal de Campina Grande UFCG, Pombal, Paraíba, PB - Brasil, e-mail: eyre.gabriel@gmail.com

---

## Resumo

A bioinformática ou biocomputação associa conhecimentos em distintas áreas do conhecimento a fim de decifrar o código genético contido nas biomoléculas pelo estabelecimento de modelos lógico-matemáticos e estatísticos. Tais estratégias têm assegurado a interpretação e elucidação de eventos biológicos gerados a partir do sequenciamento de genes e proteínas. Bancos de dados de informações biológicas cada vez mais crescentes, o desenvolvimento de novas abordagens para análise e apresentação desses dados e a investigação de novas e complexas perguntas são as forças motrizes da bioinformática. A partir do advento dos projetos genomas, biólogos moleculares passam a empregar ferramentas computacionais capazes de analisar grandes quantidades de dados biológicos, a prever funções dos genes e a demonstrar relações entre genes e proteínas. A internet e sua capacidade de compartilhar dados, além do desenvolvimento de processadores mais rápidos e com maior memória, desempenharam um papel decisivo para a consolidação da bioinformática como potencial campo do conhecimento científico. Entre suas aplicações extensamente descritas na literatura, pode-se destacar a busca pela identificação de proteínas diferencialmente expressas sob condições patológicas como potenciais alvos para a produção de fármacos antitumorais. Assim, a presente revisão de literatura objetiva relatar os eventos que marcaram o surgimento da bioinformática, descrevendo o panorama atual no qual está inserida essa nova área do conhecimento e ressaltando suas potencialidades e implicações às pesquisas na área médica, com ênfase na terapia gênica e na produção de novos fármacos.

**Palavras-chave:** Biologia computacional. Sequenciamento de DNA. Banco de dados biológicos.

## Abstract

*Genomics was developed to simultaneously identify the elements of heredity and to assign biological function to these elements. The completion of the human genome project, and other genome*

*sequencing projects, has spearheaded the emergence of the field of bioinformatics. Despite the inherent complexity of the genome, the invention of just a few molecular tools enabled genomics to flourish into the contemporaneous science. The power of the knowledge emerging from the genome is that the identification of the genetic basis of an inherited disease provides logical strategies to treat those afflicted on an individual basis. Using computer programs to analyze DNA and protein information has become an important area of life science research and development and it is reasonable to assume that in the near future, most common genetic diseases will be identified. Advances in high throughput analytic chemistry and computing technologies make the creation of a vast database of genotype with a genomic database will allow the prescription of drugs to be tailored to an individual's genotype. This same bioinformatic approach, applied to the study of human metabolites, has the potential to identify and validate targets to improve personalized nutritional health. Bioinformatics has long been focusing on the technical challenges and the enormous amount of data from image analysis and comparisons of genes (several hundreds of thousands). Genomic and proteomic technologies will therefore be complementary in their scientific and clinical application. The brief review of literature describes relevant topics concerning the effects and application of the bioinformatics on the biomedical area.*

**Keywords:** *Bioinformatics. DNA sequencing. Gene and protein databases.*

## HISTÓRICO DA BIOINFORMÁTICA

Na era pós-moderna, com o advento e a consolidação da revolução robótica no século XX, foi possível aprimorar a criação de tecnologias computacionais nas mais variadas áreas do conhecimento, o que tem provocado profundas mudanças sociais em âmbito individual e coletivo. Nesse cenário, a bioinformática ou biocomputação surge na década de 1980, a fim de superar as fronteiras das ciências pelo desenvolvimento de novas abordagens capazes de promover a análise e a apresentação de dados biológicos, bem como de garantir a investigação de novos métodos para a resolução de complexas perguntas (1). Inicialmente, a bioinformática foi tida como um paradoxo entre o mundo real e o imaginário, já que inicialmente parecia remoto que biólogos pudessem empregar ferramentas computacionais para desenvolver estratégias a serem empregadas em trabalhos científicos. Contudo, nas últimas décadas a adaptação e os avanços de recursos tecnológicos, especialmente na área de Ciências da Saúde, têm sido uma prioridade para suprir determinadas carências, como promover a criação de programas analíticos capazes de armazenar complexas sequências de genes. Assim, em pouco tempo, a informática tornou-se uma peça essencial aos estudos da área biológica, abrindo novas fronteiras para as análises de genômica e proteômica.

Em linhas gerais, a bioinformática pode ser definida como sendo o estudo da aplicação de

técnicas computacionais e matemáticas para a geração e gerenciamento de bioinformação (1, 2). Buscando tratar dados biológicos brutos, a bioinformática consolida-se como uma nova área do conhecimento, graças à crescente necessidade de desenvolver programas computacionais que permitam reconhecer sequências de genes; prever a configuração tridimensional de proteínas; identificar inibidores de enzimas; organizar e relacionar informação biológica; agrupar proteínas homólogas; estabelecer árvores filogenéticas; analisar experimentos de expressão gênica, entre outros (1, 3, 4).

Seu surgimento teve início a partir dos avanços na área de biologia molecular, especialmente com o advento dos projetos genomas, que geraram volumes intensos de informações obtidas a partir do sequenciamento de fragmentos de ácido desoxirribonucleico (DNA). Diante desse cenário, a bioinformática permite o compartilhamento de dados biológicos com a ajuda da internet pela produção de processadores mais rápidos e eficientes, fatores decisivos para consolidar a bioinformática como importante área do conhecimento científico.

Assim, essa nova disciplina científica com raízes nas ciências da computação, na estatística e na biologia molecular vem sendo desenvolvida especialmente para ordenar os resultados gerados das iniciativas de sequenciamento de genes, que produzem quantidade cada vez maior de dados sobre sequências de DNA e seus produtos proteicos. Desse modo, os biólogos moleculares têm criado reais oportunidades

para utilizar métodos estatísticos capazes de analisar grandes quantidades de dados biológicos, de inferir funções dos genes e de estabelecer relações estruturais entre genes e proteínas.

É indiscutível que a união entre a informatização e a moderna biologia molecular contribuiu contundentemente para o surgimento dessa nova área do conhecimento. Entretanto, o advento da bioinformática teve que esperar muito tempo para realmente se tornar realidade e deixar de ser uma aparente novidade. Apesar da estrutura do DNA ter sido desvendada em 1953 pelos ingleses Watson e Crick (3), foi preciso esperar até meados da década de 1980 para que fosse desenvolvida uma lente de aumento suficientemente boa (uma máquina automatizada) que permitisse a leitura em grandes quantidades do código genético contido nas biomoléculas (5). Do lado da computação foi também preciso um amadurecimento com a revolução da informática pela obtenção de computadores capazes de armazenar cada vez mais informação e de processá-la de modo cada vez mais rápido a custos reduzidos.

## A BIOINFORMÁTICA E SUAS POTENCIALIDADES

Nas últimas décadas, a sociedade contemporânea tem se deparado com os impactos provocados pelos avanços e descobertas oriundos de projetos de pesquisa conduzidos a fim de decifrar a natureza do material genético em uma grande variedade de organismos procarióticos e eucarióticos, objetivando o conhecimento total ou parcial de suas sequências de DNA. Sabe-se que um genoma armazena inúmeros fragmentos de DNA dispostos em posições específicas, denominados genes, cujas informações coordenam e controlam todos os aspectos da síntese de proteínas que, em última instância, atuam como máquinas moleculares moduladoras da homeostase celular (2).

Sob esse aspecto, os dados gerados pela bioinformática podem ser considerados como a forma mais intramolecular de colher informações sobre os seres vivos pela análise detalhada de seu código genético. É fascinante pensar que uma célula de qualquer organismo vivo possa ser codificada, decodificada e recodificada em caracteres, e que tais informações possam ser interpretadas e alteradas pelos cientistas. Efetivamente, a transformação da natureza humana passa por implicações de

alcance incalculável graças à aplicação direta dessas ferramentas, considerando que esses procedimentos possibilitam mudanças na linha evolutiva das espécies. Atualmente, a contribuição da informática ao desenvolvimento da era genômica pode ser avaliada pela rapidez e otimização na decodificação de genes e de suas regiões regulatórias – por exemplo, uma nova sequência de DNA com cerca de 12 mil bases pode ser decifrada em um minuto, enquanto há alguns anos a mesma tarefa levaria mais ou menos um ano (6). De uma maneira geral, uma aplicação direta da bioinformática nos estudos de genômica refere-se à identificação de possíveis diferenças nas sequências gênicas que possam favorecer o desenvolvimento de ferramentas para melhor diagnóstico de doenças e anomalias. As diferenças localizadas nas sequências nucleotídicas são submetidas a ferramentas estatísticas, que poderão definir as possibilidades de adoecimento ou não dos indivíduos de uma população, em que cada ser vivo se transforma em informações, dotados de instruções sobre seu próprio processo de funcionamento (3, 7).

Um organismo vivo contém em torno de 50 mil diferentes tipos de proteínas, que devem assumir sua conformação específica para ser capaz de exercer corretamente sua função biológica. Assim, os estudos de proteômica envolvem a caracterização de conjuntos de proteínas que intervêm nos processos biológicos de uma espécie, sendo que nesse estudo o principal objetivo da bioinformática é determinar a composição, estrutura e funções das proteínas, e, assim, conhecer a função dos genes anteriormente descritos pelos Projetos Genomas (7). O desafio é determinar a função e o papel biológico de cada uma das sequências de proteínas que foram identificadas. Se os genes – Genoma – são os portadores das instruções que permitem a “construção” de um determinado organismo, as proteínas – Proteoma – são as responsáveis pela estrutura e pelo funcionamento dos eventos envolvidos na homeostase celular.

Não seria exagero afirmar que a bioinformática é uma das peças-chave para o sucesso da análise de dados dos projetos anteriormente mencionados. O fortalecimento de pesquisas acerca desse tema é estratégico e necessário para impulsionar significativamente relevantes descobertas científicas nessa área do conhecimento. Em função deste quadro, a análise computacional de dados existentes torna-se um exercício de suma importância para viabilizar a extração de informações pertinentes e concretizar as potencialidades de se possuir o genoma completo

ou parcial de um organismo. Há desse modo, uma relação de motivação científica e tecnológica muito forte entre a descoberta do DNA, estudos de transgenia, materialidade do conhecimento propiciada pela biologia molecular e pela genômica, e o desenvolvimento da bioinformática.

Vista com certa indulgência, inicialmente essa nova área do conhecimento encontrou poucos adeptos para sua aplicação entre os cientistas. Entretanto, a busca por específicos sítios de restrição de fragmentos de DNA, a análise da composição de sequências de genes contra um banco de dados, mesmo que ainda reduzido, logo foram transformadas em ferramentas indispensáveis para os avanços na área de biologia computacional. É provável que somente após a dramática multiplicação, em tamanho e em número, dos bancos de dados de sequências resultantes dos projetos genomas, a bioinformática tenha sido amplamente reconhecida como uma disciplina fundamental para a nova era da genética, da biotecnologia e da biologia de sistemas (Figura 1). Modelagem molecular, bioestatística e outros campos passaram por um grande desenvolvimento, ao mesmo tempo em que havia um crescente

interesse dos cientistas das áreas de computação pela aplicação de metodologias, algoritmos, estruturação e mineração de bancos de dados em um campo ainda inexplorado da biologia, induzindo, assim, a tão necessária multidisciplinaridade.

Cerca de 1.000 projetos já foram finalizados no mundo ou estão em fase de conclusão, sendo responsáveis pela disponibilização de uma vasta quantidade de dados biológicos, tais como: identificação, sequenciamento e descrição da estrutura tridimensional de proteínas e das similaridades dos genes dentro de um genoma e entre diferentes espécies (8, 4). No Brasil, esse fenômeno pode ser acompanhado de perto pela efetiva execução e andamento de diversos projetos do Programa Genoma, acabando por consolidar uma nova concepção do desenho institucional da pesquisa no País. Além disso, tais iniciativas têm contribuído para o estabelecimento de uma cultura de interação e de integração entre as universidades e as empresas, essencial quando se quer enfrentar o desafio necessário e impositivo da transformação do conhecimento no desenvolvimento de patentes e em riqueza.

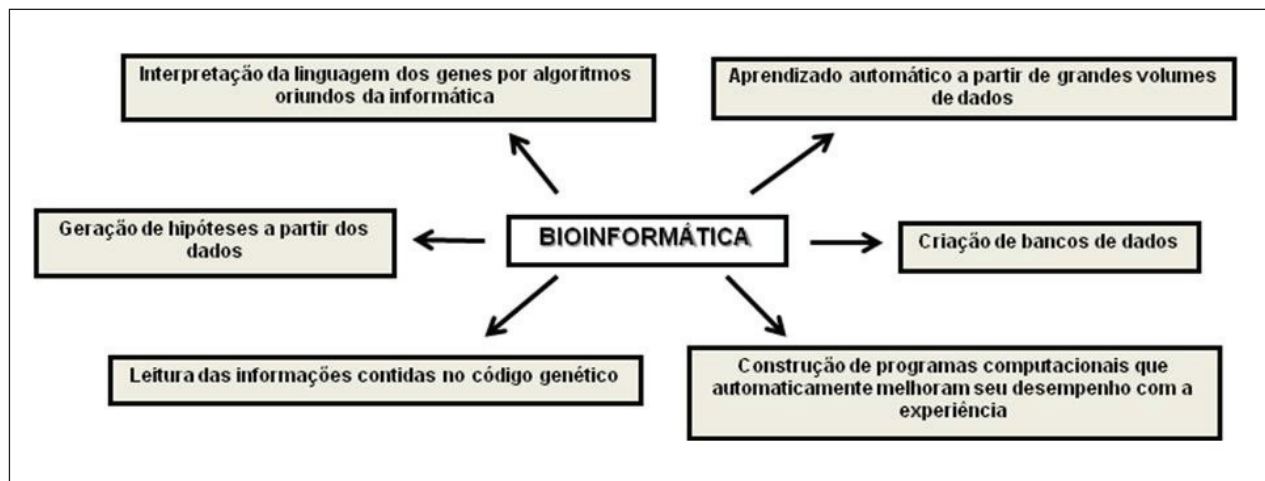


FIGURA 1 - Associação de conhecimentos entre biologia molecular e informática, que culminou com o advento da bioinformática

Atualmente, ênfase especial está se deslocando progressivamente do acúmulo de dados para a sua interpretação, visto que estudos de caracterização genômica têm gerado um grande volume de dados que precisam ser caracterizados por análises experimentais dispendiosas e pelo emprego de ferramentas computacionais sofisticadas.

## A BIOINFORMÁTICA E SUAS APLICAÇÕES MÉDICAS

Desde a descoberta da estrutura do DNA e do desenvolvimento da biologia molecular, o homem vem buscando novas alternativas que o auxiliem na busca por maior fluidez nos trabalhos

de pesquisa, acelerando a geração de processos e produtos. A bioinformática tem contribuído para avanços na pesquisa básica, mas também na ciência aplicada, promovendo a elucidação das bases moleculares envolvidas em importantes eventos celulares e garantindo sua aplicação direta nos mais variados setores da sociedade contemporânea. A identificação de alvos proteicos com potencial para serem modificados diretamente pela interação com fármacos poderá minimizar os sintomas ou as causas de uma doença (9-12). A identificação de moléculas que poderão atuar como princípio ativo na elaboração de produtos terapêuticos, inibindo ou acelerando certas reações bioquímicas, poderá apresentar efeitos sobre a cura ou a atenuação da doença. Na década de 1990, as grandes multinacionais do setor farmacêutico têm focalizado seus esforços na seleção de potenciais moléculas, visando à produção de novos medicamentos na área médica ou mesmo na agropecuária (10).

Ainda, as contribuições da bioinformática na área médica são ilimitadas graças aos avanços promovidos nas áreas de manipulação e terapia gênica com a identificação detalhada de novos genes e de seus mecanismos regulatórios. Por outro lado, ferramentas computacionais poderão permitir a detecção molecular de parasitas, acarretando no emprego direto no diagnóstico e na epidemiologia de inúmeras patologias (13). Entre as aplicações das ferramentas de biologia molecular e bioinformática merece ser destacado o sucesso no desenvolvimento de medidas terapêuticas ou preventivas contra tumores e agentes infecciosos (1, 8). Dessa forma, novos procedimentos terapêuticos também poderão assegurar que genes ausentes ou defeituosos sejam restituídos, ou, ao menos, suprir o seu produto. Em contrapartida, se um gene está sendo superativado, sua expressão poderia ser desativada. Todas as essas implicações aqui elencadas demonstram o grande potencial das ferramentas de biologia computacional no diagnóstico de doenças e na elaboração de novas alternativas terapêuticas em áreas afins.

Alternativamente, tais estratégias poderão proporcionar melhor entendimento dos processos evolutivos e das relações filogenéticas. Variações nas sequências de DNA conferem marcas individuais, tais como impressões digitais, úteis na identificação e no estabelecimento de relações de parentesco e ancestralidade entre diferentes espécies. As informações obtidas na análise comparativa detalhada de indivíduos pertencentes a uma mesma espécie são

de grande valia na antropologia, fornecendo dados acerca das variações históricas quanto ao tamanho e padrões de migração (1, 14).

O alcance e os limites de novos aspectos assumidos pela bioinformática, especialmente ao associar técnicas e instrumentos da informática e da biologia molecular, acarretam em impactantes efeitos sobre o campo das ciências biológicas e suas práticas de pesquisa. Espera-se que nas próximas décadas as aplicações dos conhecimentos gerados pela bioinformática possam se tornar mais factíveis, desenvolvendo-se mais rapidamente da pesquisa básica para a prática industrial e clínica padrão.

## CONSIDERAÇÕES FINAIS

A bioinformática integra essencialmente o desenvolvimento de programas computacionais para tratar dados biológicos preexistentes e identificar sequências de genes, prever a configuração tridimensional das proteínas, identificar inibidores enzimáticos, promover o agrupamento de proteínas, estabelecer árvores filogenéticas e analisar experimentos da expressão gênica. A bioinformática caracteriza-se como ferramenta indispensável na ordenação e agrupamento dos resultados gerados pelas análises de sequenciamento de genes, que produzem uma quantidade cada vez maior de dados sobre a composição de DNA, RNA e proteínas. Desse modo, os biólogos moleculares passaram a utilizar métodos estatísticos capazes de analisar grandes quantidades de dados biológicos, a inferir funções de genes e a demonstrar relações entre genes e proteínas.

Sob esse aspecto, a bioinformática tem sido objeto de intenso interesse nos tempos atuais, caracterizando-se como uma das mais importantes ferramentas para os profissionais da área da Saúde graças às grandes descobertas e aos avanços gerados especialmente em âmbito biomédico. Seu valor é indispensável no que tange à manutenção das ciências que caminham concomitantes à tecnologia, principalmente à pesquisa biológica. Para compreender a grandeza dessa nova área do conhecimento, é preciso correlacioná-la com algumas habilidades necessárias e indispensáveis ao seu uso. O aprofundado conhecimento de informática, o domínio da língua inglesa, o êxito das pesquisas básicas na área de biologia molecular, a crescente otimização de técnicas moleculares mais sensíveis que surgem a cada dia, dentre outros, exigem que

os profissionais que irão lidar com essa tecnologia sejam altamente multidisciplinares (4). Embora muitos desafios estejam sendo superados no que tange a maximizar a acessibilidade da população brasileira para essa nova área do conhecimento, o fator mais agravante nesse panorama refere-se à escassez de mão de obra qualificada, bem como a centralização de instituições com condições adequadas para o desenvolvimento de pesquisas nessa área de bioinformática, que ainda estão restritas a determinadas regiões do território nacional. No Brasil, assim como acontece em outras áreas de pesquisa, os investimentos em bioinformática estão concentrados na região Sudeste pela existência anterior de mão de obra qualificada e de investimentos constantes por parte do Estado de São Paulo nessa área do conhecimento. No entanto, iniciativas dos governos federal e estaduais têm empenhado esforços para combater essa concentração, tornando possível o desenvolvimento de projetos em biotecnologia em outros Estados brasileiros, o que deve fazer com que a demanda por conhecimento em bioinformática se distribua de forma igualitária em território nacional. Nesse sentido, a integração entre os Estados deverá promover a melhoria no desenvolvimento de tecnologias, aumentando a potencialidade desta poderosa ferramenta, a bioinformática, na evolução da ciência em favor da humanidade, e provocando impactos e avanços nos mais variados setores da sociedade.

## REFERÊNCIAS

1. Lesk AM Introdução à bioinformática. Porto Alegre: Artmed; 2008.
2. Alberts B, Bray D, Jonhson A, Lewis J, Raff M, Roberts K et al. Fundamentos da biologia celular: uma introdução à biologia molecular da célula. Porto Alegre: Artmed; 2005.
3. Griffith A, Muller JF, Suzuki JH, Lewontin RC, Gelbart WM, Wessler SR. Introdução à genética. Rio de Janeiro: Guanabara Koogan; 2006.
4. Cattley S, Arthur JW. BioManager: the use of a bioinformatics web application as a teaching tool in undergraduate bioinformatics training. *Brief Bioinform.* 2007;8(6):457-65.
5. Setúbal JC. 2003. A origem e o sentido da bioinformática. [acesso em mar. 2008]. Disponível em: <http://www.comciencia.br/reportagens/bioinformatica/bio01.shtml>.
6. Vogt C. 2003. Bionformática, genes e inovação. Disponível em: <http://www.comciencia.br/reportagens/bioinformatica/bio01.shtml>. [acesso em: 21 mar. 2008].
7. Krenn V, Petersen I, Morawietz L, Häupl T, Köpenick A, Dietel M, et al. Tissue arrays in proteomics. *Path Res Pract.* 2004;200:95-103.
8. Voit EO, Radivoyevitch T. Biochemical systems analysis of genome-wide expression data. *Bioinformatics.* 2000;16(11):1023-37.
9. Somogyi R, Greller LD. The dynamics of molecular networks: applications to therapeutic discovery. *Drug Discov Today.* 2001;6(24):1267-77.
10. Watkins SM, Hammock BD, Newman JW, German JB. Individual metabolism should guide agriculture toward foods for improved health and nutrition. *Am J Clin Nutr.* 2001;74(3):283-6.
11. Grant EP, Pickard MD, Briskin MJ, Gutierrez-Ramos JC. Gene expression profiles: creating new perspectives in arthritis research. *Arthritis Rheum.* 2002;46(4):874-84.
12. Häupl T, Krenn V, Stuhlmüller B, Radbruch A, Burmester G. Perspectives and limitations of gene expression profiling in rheumatology: new molecular strategies. *Arthritis Res Therapy.* 2004;6(4):140-6.
13. Herzel H, Beule D, Kielbasa S, Korbel J, Sers C, Malik A, et al. Extracting information from cDNA arrays. *Chaos.* 2001;11:98-107.
14. Heller RA, Schena M, Chai A, Shalon D, Bedilion T, Gilmore J, et al. Discovery and analysis of inflammatory disease-related genes using cDNA microarrays. *Proc Natl Acad Sci U S A.* 1997;94(6):2150-5.

Recebido: 06/04/2007

Received: 04/06/2007

Aprovado: 22/12/2007

Approved: 12/22/2007