

# Programas para uma biologia integrada

Novos softwares analisam papel de redes de genes no desencadeamento de doenças

Ricardo Aguiar

Com a evolução das técnicas de sequenciamento genômico, uma quantidade cada vez maior de dados é gerada. Desenvolver novas formas de analisar esse grande conjunto de informações é um desafio para a biologia computacional. Essa é a proposta do NetDecoder, software que integra dados e consegue extrair deles informações relevantes. O programa, desenvolvido no ano passado pelo cientista da computação e biólogo molecular brasileiro Edroaldo Lummertz da Rocha quando realizava pós-doutorado na Clínica Mayo, nos Estados Unidos, não analisa apenas o perfil de expressão gênica, ou seja, como cada gene é ativado e pode contribuir individualmente para o desenvolvimento de uma doença. Ele aponta quais redes de interação gênica – conjuntos de genes que se interrelacionam – afetam mais uma determinada condição.

“O diferencial do NetDecoder é permitir a comparação entre essas redes de interação fenótipo-específicas”, diz Rocha, que atualmente faz pós-doutorado na Universidade Harvard. Ao determinar as redes associadas a uma patologia, o programa consequentemente indica quais vias de regulação ou sinalização do organismo estão alteradas em uma população com uma doença genética em relação a um grupo de pessoas saudáveis. “Também é possível comparar diferentes estágios de desenvolvimento de uma

mesma doença e ver se a rede de interação gênica permanece inalterada ao longo do tempo.” O software tem como ponto de partida a identificação de genes expressos (ativados) de forma diferenciada em pessoas com doenças genéticas. Por meio de um algoritmo projetado para integrar informações de banco de dados, ele verifica se o gene ligado à doença tende a interagir com outros genes e a formar uma rede com potencial para afetar alguma via de sinalização.

Para validar a ferramenta computacional, o pesquisador e sua equipe analisaram os transcriptomas – conjunto de RNAs mensageiros que são produzidos em um tecido – de pessoas com câncer de mama, Alzheimer e dislipidemias (distúrbios nos níveis de gordura no sangue). Após identificar as vias alteradas, eles as compararam com a literatura e obtiveram resultados consistentes. Para o câncer de mama, por exemplo, o NetDecoder identificou a via de sinalização Braf, já conhecida por sua relação com a doença. Em pacientes com Alzheimer, encontraram alterações em vias relacionadas ao citoesqueleto celular, comum em doenças neurodegenerativas, e no caso da dislipidemia, alterações em vias metabólicas. Os resultados apareceram em artigo publicado em março na revista científica *Nucleic Acids Research*.

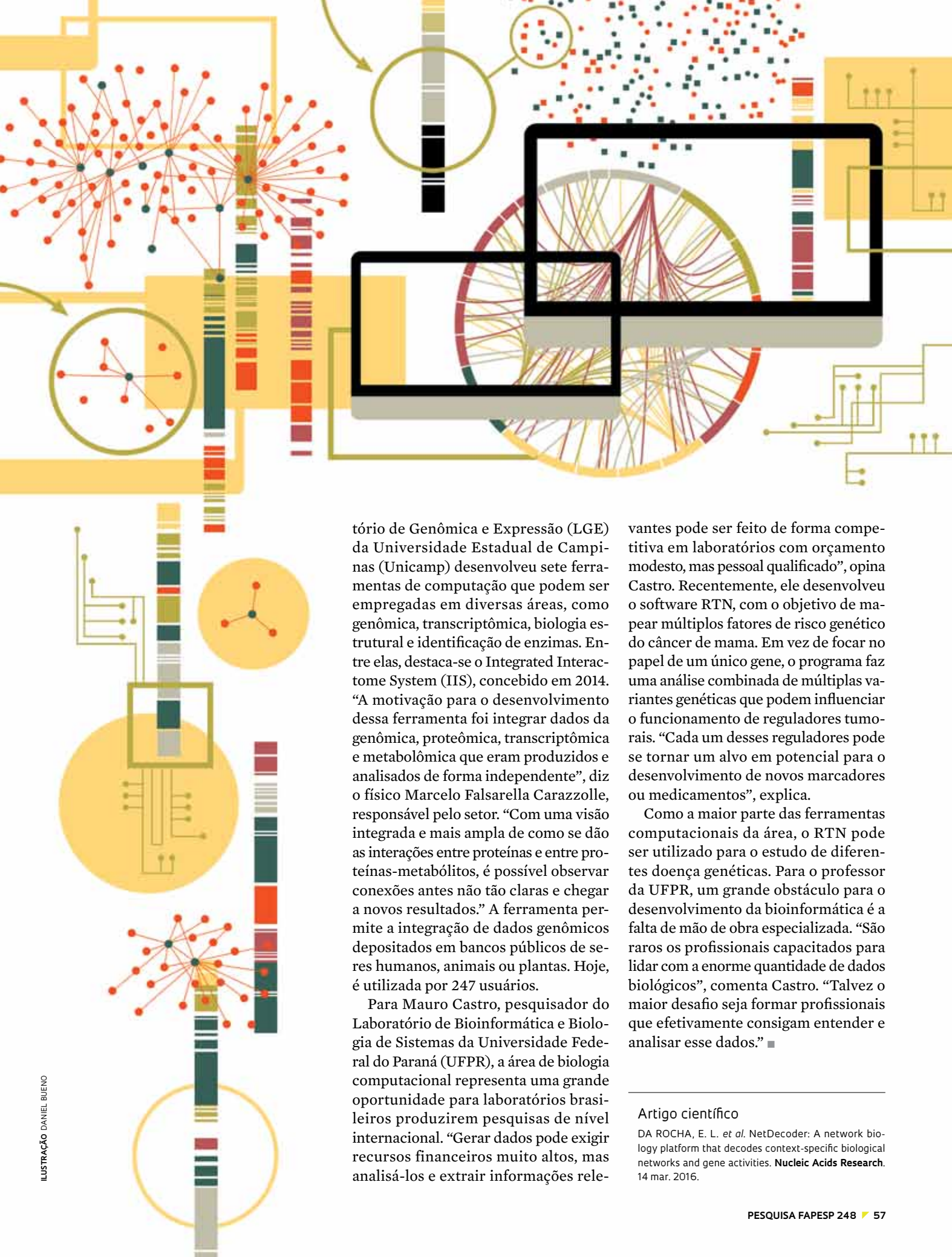
Rocha afirma que a ferramenta computacional pode ser utilizada em estudos

de qualquer tipo de doença humana. Há outros softwares que também analisam redes de genes, mas, segundo o pesquisador, uma particularidade do NetDecoder é permitir a comparação entre redes de interação que atuam em pacientes com dois fenótipos distintos, com manifestações clínicas diferentes de uma mesma doença. Essa particularidade torna possível descobrir novas relações entre genes, vias de sinalização (sistema de comunicação que coordena as ações básicas das células) e as patologias estudadas. “Assim podemos gerar novos e melhores alvos para tratamentos”, diz Rocha. “É mais fácil produzir medicamentos que atuem em uma via do que em um gene específico.”

## CRESCIMENTO EXPONENCIAL

A biologia computacional, embora relativamente recente, apresentou nos últimos anos um crescimento exponencial. A vertiginosa expansão da área pode ser vista pelo projeto de colaboração internacional Bioconductor, coordenado por pesquisadores do Fred Hutchinson Cancer Research Center, dos Estados Unidos. Iniciado em 2001, a iniciativa reúne e dá acesso a diversos softwares de código aberto para análise de dados genômicos. Em 2002, contava com apenas 20 ferramentas registradas. Neste ano, contabiliza mais de 1.200.

Desde o ano 2000, quando foi criado, o setor de bioinformática do Labora-



tório de Genômica e Expressão (LGE) da Universidade Estadual de Campinas (Unicamp) desenvolveu sete ferramentas de computação que podem ser empregadas em diversas áreas, como genômica, transcriptômica, biologia estrutural e identificação de enzimas. Entre elas, destaca-se o Integrated Interactome System (IIS), concebido em 2014. “A motivação para o desenvolvimento dessa ferramenta foi integrar dados da genômica, proteômica, transcriptômica e metabolômica que eram produzidos e analisados de forma independente”, diz o físico Marcelo Falsarella Carazzolle, responsável pelo setor. “Com uma visão integrada e mais ampla de como se dão as interações entre proteínas e entre proteínas-metabólitos, é possível observar conexões antes não tão claras e chegar a novos resultados.” A ferramenta permite a integração de dados genômicos depositados em bancos públicos de seres humanos, animais ou plantas. Hoje, é utilizada por 247 usuários.

Para Mauro Castro, pesquisador do Laboratório de Bioinformática e Biologia de Sistemas da Universidade Federal do Paraná (UFPR), a área de biologia computacional representa uma grande oportunidade para laboratórios brasileiros produzirem pesquisas de nível internacional. “Gerar dados pode exigir recursos financeiros muito altos, mas analisá-los e extrair informações rele-

vantes pode ser feito de forma competitiva em laboratórios com orçamento modesto, mas pessoal qualificado”, opina Castro. Recentemente, ele desenvolveu o software RTN, com o objetivo de mapear múltiplos fatores de risco genético do câncer de mama. Em vez de focar no papel de um único gene, o programa faz uma análise combinada de múltiplas variantes genéticas que podem influenciar o funcionamento de reguladores tumorais. “Cada um desses reguladores pode se tornar um alvo em potencial para o desenvolvimento de novos marcadores ou medicamentos”, explica.

Como a maior parte das ferramentas computacionais da área, o RTN pode ser utilizado para o estudo de diferentes doença genéticas. Para o professor da UFPR, um grande obstáculo para o desenvolvimento da bioinformática é a falta de mão de obra especializada. “São raros os profissionais capacitados para lidar com a enorme quantidade de dados biológicos”, comenta Castro. “Talvez o maior desafio seja formar profissionais que efetivamente consigam entender e analisar esse dados.” ■

#### Artigo científico

DA ROCHA, E. L. *et al.* NetDecoder: A network biology platform that decodes context-specific biological networks and gene activities. *Nucleic Acids Research*. 14 mar. 2016.