

# Matemática da vida

MANUEL ALVES FILHO  
manuel@reitoria.unicamp.br

**R**eúna, em um grupo de pesquisa, um engenheiro eletricitista, um engenheiro agrônomo, duas matemáticas e dois biólogos. Depois, deixe que busquem a melhor forma de diálogo e de utilização de seus conhecimentos, naturalmente distintos, em favor de um objetivo comum. O resultado, foco desta reportagem, foi um estudo original cujo saldo acaba de ser publicado pelo periódico científico *Scientific Reports*, pertencente ao grupo que edita a prestigiada revista *Nature*. O artigo sugere que a evolução das sequências biológicas, considerada aleatória, não é tão casual assim. “O que constatamos é que essa aleatoriedade responde a uma estrutura algébrica, que em alguma medida orienta a constituição dessas sequências”, afirma o pesquisador Marcelo Mendes Brandão, do Centro de Biologia Molecular e Engenharia Genética (CBMEG) da Unicamp, um dos autores do trabalho.

Além de Brandão, integram a equipe de pesquisadores o professor Reginaldo Palazzo Júnior, da Faculdade de Engenharia Elétrica e de Computação (FEEC) da Unicamp; e duas de suas ex-orientandas de mestrado e doutorado, Andréa Santos Leite da Rocha e Luzinete Cristina Bonani de Faria. Completam o time dois representantes da Escola Superior de Agronomia Luiz de Queiroz (Esalq-USP): o professor Márcio de Castro Silva Filho e a doutoranda Larissa Spoladore. A linha central da investigação adotada por eles foi o uso de códigos corretores de erros (ECCs, na sigla em inglês), aplicados às sequências de DNA, estrutura que carrega as instruções genéticas dos organismos vivos.

O professor Palazzo Júnior explica que os ECCs são comumente utilizados em transmissão e armazenamento de dados digitais. Assim, estão presentes em equipamentos que compõem o cotidiano das pessoas, como smartphones, TVs e notebooks. A função dos códigos corretores de erros, como o nome indica, é corrigir as falhas que ocorrem durante a transmissão ou armazenamento da informação. A ideia de aplicá-los às sequências biológicas, conforme o docente, veio do fato de haver semelhanças entre os dois sistemas.

Para elucidar melhor esse ponto, Brandão se vale de um exemplo. Imagine um transmissor responsável pelo envio de dados para um satélite, que por sua vez transmite essas informações para um receptor. Enquanto os dados cumprem a sua trajetória, podem ocorrer interferências de diversas ordens, que podem causar a perda de bytes e comprometer a integridade das informações. Nesse caso, os ECCs promovem a correção do erro ao longo do percurso, de modo a garantir que os dados cheguem ao destino de forma fidedigna. “No plano biológico o processo é praticamente o mesmo. O DNA passa a informação para o RNA mensageiro, que por sua vez o transmite para uma proteína. O sistema biológico também possui um mecanismo de correção de erros do DNA para o RNA e deste para a proteína”, detalha o pesquisador.

A despeito das semelhanças entre os sistemas, transpor os códigos de correção de erros do universo da matemática para o da biologia não foi uma tarefa trivial, como observa Luzinete Faria. Inicialmente, segundo ela, as pesquisas tinham o objetivo de obter novos conhecimentos acerca da evolução biológica com vistas ao desenvolvimento, em etapa posterior, de plantas transgênicas. O *insight* para a realização do trabalho aconteceu durante uma palestra do professor Márcio de Castro, no 56º Congresso Brasileiro de Genética, em Guarujá, no ano de 2010.

Ao apresentar uma sequência gerada pelo código corretor de erro, havia a introdução de um códon de terminação da tradução (TGA) na sequência de uma proteína ao invés do códon TGG [responsável pelo aminoácido triptofano]. Este “erro” gerado pelo código provocaria o término da síntese de uma proteína, inviabilizando a sua função. Ao término da apresentação,

Estudo sugere que a evolução das sequências biológicas não é tão aleatória quanto se acredita e que responde a uma estrutura algébrica



Fotos: Antonio Scarpinetti

Marcelo Brandão, pesquisador do CBMEG: “O que constatamos é que essa aleatoriedade responde a uma estrutura algébrica, que em alguma medida orienta a constituição dessas sequências biológicas”



Da esq. para a dir., Reginaldo Palazzo Júnior, Luzinete Cristina Bonani de Faria e Andréa Santos Leite da Rocha, integrantes do grupo de pesquisa: resultados têm gerado uma saudável polêmica no meio científico

o professor Everaldo Barros, da Universidade Católica de Brasília, fez uma observação que foi fundamental para a realização do trabalho. “Esta alteração gerada pelo código não poderia representar o aminoácido triptofano em organismos ancestrais ou que poderiam representar uma exceção à universalidade do código genético?”, questionou.

Ao se debruçar sobre a pergunta, o professor Márcio de Castro verificou que sim, em organismos primitivos o aminoácido triptofano pode ser codificado pelo códon TGA. Estas observações foram identificadas em outras sequências descritas no artigo, que foram cruciais para o estudo sobre a origem e evolução do código genético. Em discussão com o grupo da Unicamp, esta proposição converteu-se naquilo que popularmente é classificado de “pulo do gato”.

Os pesquisadores passaram, então, a fazer uma espécie de regressão com base nas informações geradas pelos ECCs. Antes, porém, os cientistas tiveram obviamente que construir esses códigos, de maneira a identificar as sequências biológicas. “O que nós percebemos foi que, para cada sequência biológica, sempre aparecia uma diferen-

ça de posição de um nucleotídeo [composto que auxilia nos processos metabólicos das células]. Em outras palavras, identificamos uma estrutura algébrica ligada a essas sequências, com definição matemática muito bem fundamentada”, diz Luzinete Faria.

A pergunta seguinte a ser respondida, acrescenta Andrea Rocha, era o que significavam essas diferenças de posição. A conclusão foi que as sequências biológicas sempre apresentavam um erro, que na linguagem da biologia equivale a uma mutação. Ao lançarem um olhar em retrospectiva, por meio de um estudo filogenético, método que traça a história evolutiva de uma espécie, os pesquisadores constataram que o código revelava uma sequência ancestral. Dito de modo simplificado, o código indicava para uma mutação que ocorreria há milhões de anos e que, por questões evolutivas desconhecidas, foi extinta ao longo do tempo.

A constatação foi posteriormente referendada por um teste feito com uma sequência de câncer de mama, como relata Brandão. Ao aplicarem os códigos corretores de erro a ela, os pesquisadores obtiveram, como resposta biológica, o ponto em que

a transcrição iria parar. “Quando voltamos para a sequência do código, agora no âmbito da matemática, encontramos uma região de triptofano [aminoácido que compõe as proteínas dos seres vivos]. O interessante é que quando vamos aos organismos primordiais, identificamos uma sequência semelhante. Portanto, não enxergamos somente a parte matemática indicando que o código era antigo. Também constatamos isso fisicamente”, reforça.

Para compreender melhor o que significa essa descoberta, Brandão cita o comentário feito pelo editor da *Scientific Reports*, segundo o qual o estudo contribuirá para mudar o pensamento sobre a filogenia molecular. Obviamente, por se tratar de uma proposição nova, que aponta para a possibilidade de a aleatoriedade da evolução biológica não ser tão aleatória assim, a pesquisa desenvolvida pelos cientistas da Unicamp e Esalq vinha causando polêmica no meio científico, antes mesmo da publicação do artigo. “Tudo que é novo encontra alguma resistência. Isso é natural. Agora, vamos aguardar as contribuições da comunidade científica ao nosso estudo”, pondera o professor Palazzo Júnior. “Isso sem contar que o trabalho é 100% brasileiro, o que sempre provoca alguma desconfiança por parte da comunidade científica dos países centrais”, completa Andrea Rocha.

Mas, afinal, qual a importância dos resultados alcançados pelo grupo de pesquisadores da conexão Campinas-Piracicaba? Quem responde é Luzinete Faria. De acordo com ela, a identificação das mutações ocorridas nas sequências biológicas através do tempo abre caminho para a manipulação genética. “O que está no horizonte é promovermos, por exemplo, intervenções para evitar ou corrigir mutações que possam levar a diferentes doenças”, antevê, com a concordância de Brandão.

Os pesquisadores fazem questão de assinalar, entretanto, que a ciência ainda está longe de atingir esse patamar. “Nós identificamos somente um código, mas podem e devem existir muitos outros. Nós conseguimos responder a algumas perguntas, mas estas suscitaram muitas outras, que ainda aguardam por solução. O que nós afirmamos internamente ao grupo é que demos o primeiro passo para o cumprimento de uma longa jornada”, resume Andrea Rocha.