

INTRODUÇÃO À EVOLUÇÃO MOLECULAR: O MODELO DE JUKES-CANTOR

Aluno: Adriana Cruz Martins
Orientador: Sérgio Bernardo Volchan

Introdução

O objetivo principal do estudo da denominada *filogenia molecular* é a reconstrução de árvores filogenéticas que descrevem as relações evolucionárias entre diferentes organismos. Um dos maiores desafios deste estudo consiste na determinação da *distância genética* (ou evolucionária) entre os organismos estudados, ou seja, uma medida de seu “parentesco” evolucionário. Diante desta dificuldade, foram propostos e desenvolvidos inúmeros modelos probabilísticos que avaliam o número de mutações acumuladas entre seqüências homólogas do DNA dos organismos desde sua separação de um ancestral comum, o que permite inferir a distância entre eles com base nas diferenças observadas atualmente.

Neste projeto, analisamos o mais simples destes modelos: o *modelo de Jukes-Cantor* no qual se assume, entre outras hipóteses, que cada substituição de nucleotídeos do DNA possui a mesma probabilidade de ocorrência e que os sítios do DNA são independentes. Para esta análise, ao longo do projeto, procuramos nos familiarizar com os principais conceitos biológicos relacionados à filogenia assim como estudar a aplicação de alguns conceitos matemáticos, particularmente relacionados à probabilidade, nas teorias de evolução molecular.

A teoria neutra da Evolução Molecular

Mutações neutras são aquelas que não são necessariamente responsáveis pelo aparecimento de características adaptativas (um exemplo são as alterações “silenciosas” que modificam os aminoácidos, mas não afetam a produção e função da proteína correspondente). Em 1968, o geneticista japonês Motoo Kimura propôs que, a nível molecular, tais mutações são mais freqüentes que os demais tipos de mutações (mutações vantajosas ou prejudiciais) e que são fixadas por variações genéticas aleatórias (“*deriva gênica*”). Esta nova idéia evolucionária sugeria, por exemplo, que as chamadas “mutações vantajosas” possuiriam pouca contribuição para a variabilidade genética das populações por serem extremamente raras e se fixarem muito rapidamente (pela seleção natural).

Segundo Kimura, grande parte das diferenças observadas entre seqüências de nucleotídeos de espécies diferentes seria resultado da fixação de mutações neutras por deriva gênica. Esta chamada *teoria neutra da evolução*, com o auxílio de modelos matemáticos apropriados e da noção de “relógio molecular” (ver abaixo), é uma das principais ferramentas para a compreensão moderna da evolução molecular.

O relógio molecular

A substituição dos nucleotídeos do DNA ao longo do tempo pode ser vista como um processo estocástico. Uma versão muito utilizada no estudo da evolução molecular é o *modelo de Wright-Fisher*, que prediz que a probabilidade de fixação de uma nova mutação em uma população de N indivíduos diplóides é dada por $1/2N$. No contexto deste modelo, seja k a taxa de fixação de mutações, definida por:

$$k = (\text{n}^\circ \text{ médio de mutações por geração}) \cdot (\text{fração de mutações que se fixam}).$$

Por outro lado, o número médio de mutações por geração é determinado pelo produto entre número de gametas produzido por geração $2N$ e a probabilidade de ocorrência de uma mutação por geração, u . Considerando que a fração de mutações que se fixam é $1/2N$, obtemos:

$$k = (2 \cdot N \cdot u) \cdot \left(\frac{1}{2 \cdot N} \right), \text{ isto é, } k = u.$$

Este é um dos resultados mais importantes da teoria neutra da evolução. De acordo com este modelo a taxa na qual as mutações neutras ocorrem é igual à taxa de substituições de nucleotídeos do DNA. O termo “relógio molecular” foi introduzido em 1965, por Emile Zuckerkandl e Linus C. Pauling, para ilustrar esta acumulação de substituições de nucleotídeos, supostamente a uma taxa constante, durante a evolução das moléculas e auxiliar na construção de árvores filogenéticas, por exemplo. Neste sentido, pode-se observar que as moléculas são capazes de determinar seu “tempo evolucionário” através do acúmulo de substituições de nucleotídeos, funcionando, portanto, como um verdadeiro “documento histórico” da evolução. É importante observarmos, entretanto, que a hipótese da ocorrência de substituições a uma taxa constante é uma simplificação, uma vez que não leva em consideração uma série de outros fatores, como por exemplo, os tipos de organismos analisados, o que pode complicar substancialmente a análise.

O modelo de Jukes-Cantor

O primeiro e mais simples modelo desenvolvido com o objetivo de determinar a distância evolucionária entre seqüências de DNA foi proposto em 1969 por Thomas H. Jukes e Charles R. Cantor. Neste modelo, assume-se que as transições (trocas entre bases do tipo purinas: adenina e guanina) ocorrem com a mesma probabilidade que as demais substituições de nucleotídeos (transversões). Através deste modelo, obtém-se uma fórmula que determina o número d estimado de substituições de nucleotídeos por sítio ou distância evolucionária:

$$d = -\frac{3}{4} \ln \left(1 - \frac{4}{3} p \right).$$

Aqui, p é a proporção de diferenças de nucleotídeos observada entre as duas seqüências (número de substituições observadas dividido pelo número total de sítios de nucleotídeos comparados). Esta fórmula leva em consideração as substituições que não são evidentes (não observadas, e.g., substituições múltiplas), convertendo a proporção p em uma distância evolucionária d .

A área de estudo da evolução molecular é bastante ativa e objeto de interesse de diversas pesquisas científicas. Portanto, já foram desenvolvidos modelos mais complexos e sofisticados para realizar esta mesma correção de uma forma mais precisa.

Referências bibliográficas

- 1 – BROWN, T. A. **Genomes**. 2.ed. Oxford: Wiley-Liss, 2002. 572p.
- 2 – GRIMMETT, G. R. & STIRZAKER, D. R. **Probability and Random Processes**. 2.ed. Oxford: Oxford University Press, 1992. 600p.
- 3 – KLEIN, J. & TAKAHATA, N. **Where do we come from? The molecular evidence for Human Descent**. 1.ed. Berlin: Springer, 2001. 462p.
- 4 – TAVARÉ, S. & ZEITOUNI, O. **Lectures on Probability Theory and Statistics**. New York: Springer Verlag, 2004. 314p.