

Filogenética Molecular Aplicada

Extensão PUC-PR

Curitiba Nov/2023

O material completo do curso, incluindo os arquivos usados nas aulas práticas, está disponível no repositório:

<https://github.com/matheusbianconi/filogeneticamolecular>

Prática 04

Árvores de famílias gênicas e relações de homologia

(1) Identificação de genes ortólogos e parálogos

1. IDENTIFICAÇÃO DE GENES ORTÓLOGOS E PARÁLOGOS

Ao compararmos sequências de genes específicos entre diferentes espécies, nota-se que não necessariamente existe uma relação um-para-um para todos os genes, ou seja, que cada espécie tem apenas uma cópia de cada gene herdado do ancestral comum mais recente entre elas. Eventos de duplicação gênica (e de genomas inteiros) são frequentes durante a evolução, o que leva à existência de mais de uma cópia de um determinado gene dentro do mesmo genoma. Esses eventos de duplicação podem ser recentes (por exemplo, restritos à espécie em questão ou ao gênero), de modo que só aquele grupo terá múltiplas cópias daquele gene, mas também podem ser mais antigos, e anteriores à origem de uma determinada linhagem. Neste caso, todas as linhagens descendentes carregarão múltiplas cópias daquele gene (note, porém, que eventos de perda gênica também são frequentes, logo, não necessariamente as múltiplas cópias serão observadas em todas as linhagens descendentes). Tais genes que são relacionados evolutivamente devido a um evento de duplicação gênica são chamados genes parálogos. Genes em diferentes espécies que descendem de um mesmo gene no ancestral comum mais recente daquelas espécies (i.e. são relacionados evolutivamente devido a um evento de especiação) são chamados genes ortólogos.

Um aspecto importante na comparação de genes ortólogos entre diferentes espécies é que a filogenia desses genes não necessariamente seguirá a filogenia da espécie. Embora é esperado que a maioria dos genes produzirá uma filogenia que recapitula a filogenia da espécie, sabe-se que haverá casos em que a filogenia do gene diferirá da filogenia da espécie. Isso ocorre devido a

diversos fenômenos biológicos, tais como a convergência evolutiva e a transferência horizontal de genes, e também a eventos que ocorrem em nível populacional durante o processo de formação de novas espécies, tais como o sorteamento incompleto de polimorfismos ancestrais e a introgressão (hibridação introgressiva). Nestes casos, estudos genéticos mais detalhados com amostras de múltiplas populações das espécies são necessários para determinar a causa biológica das discrepâncias filogenéticas observadas.

Exercício 1

Você recebeu as sequências de quatro genes do genoma nuclear de nove espécies de plantas da família das gramíneas e de uma espécie que servirá como grupo externo ('metadados.tsv'). Você deverá (1) determinar quantos eventos de duplicação gênica ocorreram durante a história evolutiva de cada um dos genes, e (2) identificar as espécies que são possivelmente poliplóides. Para identificar os eventos de duplicação, é sugerido que antes você faça uma árvore filogenética com um marcador do cloroplasto ('01_sequencias_matk_cloroplasto.fasta'), para determinar as relações filogenéticas esperadas entre as espécie.