



Genes hospedeiros e retrocópias no genoma humano exibem perfis transcricionais assimétricos alelo-específicos

Juan Carlo Santos e Silva, Manuella dos Santos Siqueira, Cristina dos Santos Ferreira, Ana Beatriz Garcia, Enrique Medina-Acosta

As retrocópias do genoma são repetições que ocorrem por eventos de duplicação ou retrotransposição. Essas cópias podem mudar o modo como é orquestrada a regulação epigenética e, conseqüentemente, a expressão gênica das regiões onde elas estão inseridas, isto é, os genes hospedeiros. Tais mudanças podem originar eventos peculiares como o *imprinting* genômico, caracterizado pela expressão monoalélica origem parental-dependente. Neste contexto, o objetivo do trabalho foi averiguar os perfis alélicos transcricionais consistentes com *imprinting* genômico em retrocópias, no genoma humano, utilizando abordagens computacionais integrativas. Com esta perspectiva, as análises foram fundamentadas na integração das retrocópias disponíveis na plataforma *UCSC Genome Browser* (versão GRCh38/hg38), com o banco de dados dbSNP que apresenta informações para polimorfismos de nucleotídeo único (SNPs), em conjunto com os estudos de RNA-Seq disponíveis no projeto GTEx (*Genotype Tissue Expression*) para 53 tecidos pós-morte. A fim de evitar ambigüidades da expressão alelo-específica (ASE) dos SNPs, foram excluídas as sequências, centradas no contexto de 29 nucleotídeos que apresentavam mais de uma ocorrência no genoma, e aquelas que ocorrem em regiões de duplicação segmentar. As informações para os SNPs alvo, obtidos após integração, foram classificadas quanto ao status de ASE (monoalélica/bialélica), a partir do processamento em scripts de linguagem R. Desta forma, foi possível analisar 15.491 retrocópias das quais 8.427 possuem 246.269 SNPs informativos. Foram identificadas retrocópias com perfil monoalélico em múltiplos tecidos. Em destaque, foi identificado o retro-*DGKZP1* que possui expressão monoalélica em todos os tecidos informativos, enquanto que o gene no qual ele está inserido, *RP11-65B7.2*, possui expressão bialélica. A investigação da origem parental dos alelos de um SNP (rs3794434) para o retro-*DGKZP1*, em experimentos com dados de RNA-Seq e sua relação com o genótipo de uma família de 17 indivíduos, permite inferir que a expressão monoalélica foi derivada de um processo de editoração de RNA, excluindo o cenário de *imprinting* genômico para esses genes. Análise das 8.427 retrocópias permitirá definir a extensão de editoração e se existem novas retrocópias imprintadas.

Palavras-chave: *imprinting* genômico, expressão alelo-específica, retrocópias.

Instituição de fomento: CNPq, UENF, NUDIM.