

A Ciência e os caminhos do desenvolvimento

## Identificação de lncRNAs com perfil de expressão monoalélica associados às regiões controladoras de *imprinting* genômico em humanos

Amanda Pereira Vasconcelos, Cristina dos Santos Ferreira, Enrique Medina-Acosta

O *imprinting* genômico é um tipo de herança epigenética caracterizado pelo silenciamento alelo-específico, resultando em um perfil monoalélico de expressão transcricional de origem parental-dependente. As marcas epigenéticas que regulam a expressão gênica imprintada incluem a metilação do DNA e o enriquecimento de modificações químicas em histonas nas regiões diferencialmente metiladas (DMRs). O nosso grupo de pesquisa identificou recentemente 125 gDMR candidatas. Considerando que o mecanismo de regulação do *imprinting* genômico envolve também a expressão monoalélica de RNAs longos não codificantes (lncRNAs), esse trabalho objetiva identificar lncRNAs com perfil de expressão monoalélica numa extensão de até 4,6Mb ao redor de cada DMR candidata a fim de oferecer mais informações sobre essas regiões supostamente controladoras de *imprinting*. Para averiguação dos perfis alélicos dos lncRNA estamos implementando um pipeline computacional de análise utilizando estudos de transcriptoma (RNA-Seq). Para estimar a expressão alelo-específica (ASE) quantificaremos a fração alélica pelo número de *reads* através de SNPs nos lncRNAs utilizando experimentos de RNA-Seq do banco de dados do projeto GTEx. Os valores de ASE obtidos serão analisados por algoritmos de linguagem R de programação e integrados pela plataforma UCSC Genome Browser versão GRCh37/hg38. Para análise das modificações nas histonas serão explorados os dados disponíveis no Roadmap Epigenomics Browser do NIH via scripts em R. Dentre as análises realizadas foram anotadas as posições físicas (4,6Mb ao redor) das 125 iDMRs obtendo genes, pseudogenes e lncRNA desta região que filtrados pelo seu biotipo, tal como classificados no Ensembl Genome Browser. No total, anotamos 7030 lncRNAs, compreendendo 61.202 ASE sites (SNPs, MAF  $\geq 0.01$ ) no projeto GTEx. Os ASE sites foram validados como únicos no genoma (ausência de parálogos) utilizando o programa ShortMatch em R. O seguinte passo consiste na avaliação dos perfis de expressão alélica em múltiplos tecidos e a averiguação do enriquecimento de modificações de histonas nos lncRNA com expressão monoalélica.

Palavras-chave: epigenômica, *imprinting* genômico, lncRNA

Instituição de fomento: CNPq