



**AJUSTE DE ONDAS GENÔMICAS NO SINAL DE INTENSIDADE DE
DADOS DE GENOTIPAGEM PARA A IDENTIFICAÇÃO DE CNVs**

¹FERNANDA C. P. PEREIRA; ²ANDRÉ ROBLES GONÇALVES; ³LUCIANA C. A.
REGITANO; ⁴POLIANA F. GIACHETTO

Nº 13602

RESUMO

A tecnologia de genotipagem baseada no emprego de *chips* de SNPs, tem sido largamente utilizada na detecção de variações no número de cópias (CNVs, do inglês Copy Number Variation) de regiões do genoma, alterações que podem contribuir para a variação fenotípica observada entre indivíduos. Em animais de produção, estudos têm reportado a associação de CNVs com características de importância econômica, tornando a identificação e análise desses marcadores moleculares objeto de estudo na Embrapa. Um fator que interfere na acurácia da identificação das CNVs por meio dessa tecnologia é a ocorrência das chamadas ondas genômicas, definidas como variações na intensidade do sinal de hibridização entre amostras (artefatos), cujos padrões são relacionados ao conteúdo de GC da região. Esse fenômeno pode influenciar a detecção de CNVs pelos algoritmos, resultando em um maior número de falsos positivos. O presente trabalho teve como objetivo avaliar uma metodologia de ajuste da intensidade do sinal para ondas genômicas, na identificação de CNVs em bovinos da raça Canchim, genotipados com um *chip* de SNPs de alta densidade. Foram utilizados 400 dados de genotipagem dos animais, realizadas com o *chip* comercial BovineHD BeadChip (Illumina®). A identificação das CNVs foi feita por meio do *software* PennCNV, que utiliza duas medidas de intensidade de sinal calculadas para cada SNP: LRR e BAF. Para o ajuste da intensidade do sinal de LRR, foi utilizado um modelo de regressão proposto por Diskin et al. (2008), que correlaciona a intensidade do sinal de um dado marcador com o conteúdo de GC dentro de uma janela de 1Mb ao seu redor, e calcula o resíduo para cada marcador como a intensidade ajustada do sinal. Comparando-se os valores de GCWF (índice que mede a magnitude de flutuação da intensidade do sinal ao longo do genoma) das 400 amostras antes e após o ajuste, verificou-se uma variação de -5,197 a 8,694 e -0,006 a 0,008, respectivamente, evidenciando a eficácia do método utilizado na redução dos efeitos das ondas genômicas e no aumento da acurácia na inferência das CNVs.

1. Bolsista CNPq; Graduação em Sistemas de Informação, PUCC, Campinas - SP. fernanda.pereira@colaborador.embrapa.br
2. Colaborador: Graduação em Sistemas de Informação, PUCC, Campinas - SP. andre.goncalves@colaborador.embrapa.br
3. Colaboradora: Pesquisadora, Embrapa Pecuária Sudeste, São Carlos - SP. luciana.regitano@embrapa.br
4. Orientadora: Pesquisadora, Embrapa Informática Agropecuária, Campinas - SP poliana.giachetto@embrapa.br