

# Identificação de CNVs em bovinos Canchim, a partir de dados de gentipagem de SNPs com chips de alta densidade

André Robles Gonçalves<sup>1</sup>  
Poliana Fernanda Giachetto<sup>2</sup>

Variações no número de cópias do *Copy Number Variation* (CNVs) são regiões do genoma caracterizadas pelo ganho ou perda de sequências entre 2 ou mais indivíduos de uma população. Em humanos, têm sido observado que as CNVs podem estar envolvidas na variabilidade fenotípica e mesmo na susceptibilidade de indivíduos a doenças (ROVELET-LECRUX et al., 2006). Várias publicações têm relatado o efeito das CNVs sobre a expressão gênica e na associação com síndromes e doenças complexas, como o autismo, esquizofrenia e doença de Parkinson (HENRICHSEN et al., 2009). Em animais de produção, tem sido observado que as CNVs podem afetar características fenotípicas de importância econômica, como a produção de leite, qualidade de carne e resistência dos animais a parasitas (SEROUSSI et al., 2010). Dada a sua importância, várias técnicas têm sido desenvolvidas para a detecção de CNVs, incluindo o resequenciamento de genomas e os *chips* de genotipagem de SNPs. Dados de genotipagem de bovinos utilizando *chips* de DNA de alta densidade são largamente gerados na Empresa Brasileira de Pesquisa Agropecuária (Embrapa), para estudos de associação genótipo – fenótipo. O objetivo desse estudo foi utilizar uma ferramenta *open source*, o CNstream (ALONSO et al., 2010), para a identificação de CNVs a partir de dados de genotipagem de bovinos por meio de chips de SNPs da plataforma Illumina. Foram utilizados dados de 400 animais (bovinos Canchim),

---

<sup>1</sup> Sistemas de Informação/PUC-Campinas, andreroblesg@hotmail.com

<sup>2</sup> Embrapa Informática Agropecuária, poliana.giachetto@embrapa.br

participantes de um programa de melhoramento da Embrapa Pecuária Sudeste, genotipados com o BovineHD BeadChip (Illumina). Um total de 5.789 CNVs foram detectadas, sendo caracterizadas 4.536 deleções e 1.253 duplicações de regiões do genoma, cujo tamanho variou de 2.251 a 99.494bp. Vários autores têm observado uma grande variabilidade entre os dados gerados pelos algoritmos de identificação de CNVs, assim como uma elevada taxa de falsos positivos e falsos negativos associada aos métodos (TSUANG et al., 2010). O mesmo conjunto de amostras foi analisado também com a ferramenta PennCNV (WANG et al., 2007), uma das mais utilizadas para a análise de CNVs a partir de dados provenientes da plataforma Illumina de genotipagem e alguns resultados divergentes foram encontrados, os quais podem ser atribuídos às limitações metodológicas de cada ferramenta. Nós sugerimos a inclusão de mais de uma ferramenta nos pipelines de descoberta e análise de CNVs, para uma predição mais destas por experimentos de bancada.

## Agradecimentos

Embrapa Informática Agropecuária, Embrapa Pecuária Sudeste

## Referências

- ALONSO, A.; JULIÀ, A.; TORTOSA, R.; CANALETA, C.; CAÑETE, J. D.; BALINNA, J.; BALSÀ, A.; TORNERO, J.; MARSAL, S. CNstream: a method for the identification and genotyping of copy number polymorphisms using Illumina microarrays. **BMC bioinformatics**, v. 11, p. 264, 2010. Doi:10.1186/1471-2105-11-264
- HENRICHSEN, C. N.; CHAIGNAT, E.; REYMOND, A. Copy number variants, diseases and gene expression. **Human Molecular Genetics**, v. 18, p. R1-8, 2009.
- ROVELET-LECRUX, A.; HANNEQUIN, D.; RAUX, G.; LE, M. N.; LAQUERRIÈRE, A.; VITAL, A.; DUMANCHIN, C.; FEUILLETTE, S.; BRICE, A.; VERCELLETTO, M.; DUBAS, F.; FREBOURG, T.; CAMPION, D. APP locus duplication causes autosomal dominant early-onset alzheimer disease with cerebral amyloid angiopathy. **Nature Genetics**, v. 38, n. 1, p. 24–26, Jan. 2006.

SEROUSSI, E.; GLICK, G.; SHIRAK, A.; YAKOBSON, E.; WELLER, J. I.; EZRA, E.; ZERON, Y. Analysis of copy loss and gain variations in Holstein cattle autosomes using BeadChip SNPs. **BMC Genomics**, v. 11, p. 673, Nov. 2010. Doi:10.1186/1471-2164-11-673

TSUANG, D. W.; MILLARD, S. P.; ELY, B., CHI, P.; WANG, K.; WENDY, H.; KIM, S.; BRKANAC, Z.; YU, C.-E. The effect of algorithms on copy number variant detection. **PLoS ONE**, v. 5, n. 12, p. e14456, 2010. Doi:10.1371/journal.pone.0014456

WANG, K.; LI, M.; HADLEY, D.; LIU, R.; GLESSNER, J.; GRANT, S.; HAKONARSON, H.; BUCAN, M. PennCNV: an integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. **Genome Research**, v. 17, n. 11, p. 1665-1674, Nov. 2007.

