



# V WORKSHOP DE PÓS-GRADUAÇÃO EM ZOOTECNIA E EM CIÊNCIA ANIMAL DO ESTADO DE MATO GROSSO DO SUL

03 A 04 de Outubro de 2018

Local: FAMEZ/UFMS

## ANÁLISES PRELIMINARES DE CONTROLE DE QUALIDADE DE DADOS GENÔMICOS EM UM BANCO DE DADOS DE BÚFALAS LEITEIRAS

Karina Rosa da SILVEIRA\*<sup>1</sup>, Jéssica Cristina Gonçalves dos SANTOS<sup>1</sup>, Leonardo de Oliveira SENO<sup>1</sup>, Humberto TONHATI<sup>2</sup>, Rusbel Raul Aspilcueta BORQUIS<sup>1</sup>

\*Autor correspondente: karinasilveira50@gmail.com

<sup>1</sup>Universidade Federal da Grande Dourados – Faculdade de Ciências Agrárias

<sup>2</sup>Universidade Estadual Paulista - Faculdade de Ciências Agrárias e Veterinárias

A técnica de genotipagem em diferentes escalas, acompanhado do uso de marcadores SNPs (*Single Nucleotide Polymorphism*) está sujeita a inconsistências e a avaliação da qualidade das amostras e marcadores garante maior confiabilidade das análises estatísticas, além de diminuir o número de informações irrelevantes otimizando a análise computacional dos dados, sem afetar a precisão dos resultados. Assim, o objetivo deste trabalho é analisar diferentes níveis dos critérios de controle de qualidade de marcadores SNPs em um banco de dados de búfalas com aptidão leiteira. Dados genômicos, pertencentes a três fazendas, foram cedidos pela Faculdade de Ciências Agrárias e Veterinárias da Universidade Estadual Paulista (UNESP), campus Jaboticabal. Os animais foram genotipados com o chip 90K Axion® Buffalo Genotyping (116169 marcadores), tendo como referência o genoma bovino, visto que, ainda não há chip específico para a espécie *Bubalus bubalis*. Para adequar os dados fornecidos ao formato solicitado pelo programa usado para consistência dos marcadores e exclusão dos cromossomos sexuais foi utilizado o *software* Cygwin. O controle de qualidade (CQ) do banco de dados contendo 322 animais (amostras) foi feito por meio do *software* PLINK 1.07, no qual os dados foram filtrados com base na frequência do alelo menor (*minor allele frequency* – MAF), porcentagem de SNPs válidos na população (*call rate*) e equilíbrio de Hardy–Weinberg (EHW). Foram adotados dois níveis de CQ sendo 1) 0,05 MAF; 0,95 *Call rate* (menos rigoroso) e 2) 0,02 MAF; 0,98 *Call rate* (mais rigoroso) e para ambos EHW de  $10^{-6}$ . Concluída a análise, obteve-se para o nível 1 o total de 58666 marcadores (autossômicos) viáveis e, para o nível 2, 47191 marcadores. A partir destes resultados, será verificada a influência do controle de qualidade dos dados genômicos sobre o estudo da estrutura da população em questão e suas implicações no melhoramento de búfalas leiteiras.

**Palavras-chave:** SNP, MAF, Equilíbrio de Hardy-Weinberg, *call rate*, genômica

**Agradecimentos:** UFGD, CAPES