

NEW SINGLE NUCLEOTIDE POLYMORPHISMS IN GUZERÁ BREED REVEALED BY WHOLE-GENOME RE-SEQUENCING

¹Izinara Rosse Cruz, ²Juliana Assis Geraldo, ²Francislon Silva de Oliveira, ²Laura Rabelo Leite, ³Flávio Araújo, ⁴Adhemar Zerlotini, ⁵Beatriz C. Lopes, ⁶Wagner A. Arbex, ⁶Marco Antônio Machado, ⁶Maria Gabriela Campolina Diniz Peixoto, ⁶Rui da Silva Verneque, ⁶Marta Fonseca Martins, ⁶Marcos Vinícius G.B. Silva, ⁷Roney Santos Coimbra, ¹Maria Raquel Santos Carvalho, ⁸Guilherme Oliveira

¹Universidade Federal de Minas Gerais-UFMG, Brazil, ²Universidade Federal de Minas Gerais-UFMG/Genomics and Computational Biology Group - FIOCRUZ-MG/Center for Excellence in Bioinformatics – FIOCRUZ-MG, Brazil, ³Genomics and Computational Biology Group - FIOCRUZ-MG, Brazil, ⁴EMBRAPA Informática Agropecuária - SP, Brazil, ⁵EPAMIG – MG, Brazil, ⁶EMBRAPA Dairy Cattle – MG, Brazil, ⁷Biosystems Informatics, Research Center René Rachou - FIOCRUZ, Brazil, ⁸Genomics and Computational Biology Group - FIOCRUZ-MG/Center for Excellence in Bioinformatics – FIOCRUZ-MG, Brazil

Background: Currently, livestock accounts for 9% of Brazil's GDP and the traditional breeding has provided a gain in production of 1% per year. With the proposal to increase the efficiency of breeding programs, genotyping chips containing thousands of SNPs as putative markers for traits of interest were developed. These chips contain variations observed in various commercial breeds, but do not include the Guzerá that contributes significantly to milk production and is well adapted to local climate conditions. In this context, the objective of our work was to sequence and assemble the Guzerá genome in order to identify race-specific variations that might be used in breeding programs. **Results:** Our group performed the genome assembly of Guzerá breed that was performed using mate-paired libraries, with 1-2 and 3-4 kb inserts, by SOLiD v3 e v4 second-generation sequencing. The sequences were mapped to the publicly available reference genome of *Bos taurus* (UMD 3.1) using the LifeScope software. The average depth of coverage achieved from mapping was 26X (ranging from 14.5 to 52.3) for each chromosome and 87% of the reference genome was covered. A list of putative SNPs was generated from the mapped reads, using the diBayes SNP Detection module included with LifeScope. These SNPs were filtered according to the following criteria: (1) SNPs with an overall quality less than 20 were removed; (2) minimum rate of 20% for the new allele; (3) variants with too low or too high read depths were removed: the minimum of 4 reads of depth and the maximum as the mean read depth + 3 times the standard deviation; (4) variants within 5 bp of each other were removed. 4.040.476 SNPs (54%) remained after the filter and they were compared with dbSNP138. It was observed that 51.5% of them were new. **Conclusions:** A significant number of distinct SNPs were observed between the Guzerá and Hereford breeds. These differences may be associated with the production traits and adaptability that characterize the zebu breeds and may be useful in breeding programs. **Supported by:** CAPES, CNPq (307975/2010-0, 309312/2012-4), FAPEMIG (CBB-1181/0, TCT 12.093/10, REDE-56/11), NIH-USA (TW007012), CAPES/CDTS-FIOCRUZ. **Acknowledgments:** PDTIS-FIOCRUZ - Platforms RPT04B (Bioinformatics) and RPT01F (NGS).

Keywords: SNPs; animal breeding; next-generation sequencing