

## Next generation sequence variants with potential loss of function in influential sires of the Gyr, Girolando, and Guzerat cattle breeds

Nedenia Bonvino Stafuzza<sup>1</sup>, Adhemar Zerlotini<sup>2</sup>, Michel Eduardo Beleza Yamagishi<sup>2</sup>, Marcos Eli Buzanskas<sup>1,3</sup>, Tatiane Cristina Seleguim Chud<sup>1</sup>, Alexandre Rodrigues Caetano<sup>4</sup>, Danísio Prado Munari<sup>1</sup>, Dorian J. Garrick<sup>5</sup>, Marco Antonio Machado<sup>6</sup>, Marta Fonseca Martins<sup>6</sup>, Maria Raquel Carvalho<sup>7</sup>, John Bruce Cole<sup>8</sup>, Marcos Vinicius Gualberto Barbosa da Silva<sup>6\*</sup>

<sup>1</sup>Departamento de Ciências Exatas, Universidade Estadual Paulista, Faculdade de Ciências Agrárias e Veterinárias, Jaboticabal, São Paulo, Brazil

<sup>2</sup>Embrapa Informática Agropecuária, Campinas, São Paulo, Brazil

<sup>3</sup>Departamento de Zootecnia, Universidade Federal da Paraíba, Areia, Paraíba, Brazil

<sup>4</sup>Embrapa Recursos Genéticos e Biotecnologia, Brasília, Distrito Federal, Brazil

<sup>5</sup>Department of Animal Science, Iowa State University, Ames, IA, United States of America

<sup>6</sup>Embrapa Gado de Leite, Juiz de Fora, Minas Gerais, Brazil

<sup>7</sup>Departamento de Biologia Geral, Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brazil

<sup>8</sup>Animal Genomics and Improvement Laboratory, ARS, USDA, Beltsville, MD, United States of America

\*Corresponding author

M. V. G. B. Silva

Embrapa Gado de Leite – Empresa Brasileira de Pesquisa Agropecuária.

Rua Eugênio do Nascimento 610 – Dom Bosco 36038-330, Juiz de Fora, Minas Gerais, Brazil

Tel.: +55 32 3311 7670

E-mail: [marcos.vb.silva@embrapa.br](mailto:marcos.vb.silva@embrapa.br)

### Abstract

**Background:** Identification and characterization of novel variants in diverse cattle breeds is an important step to understand the molecular mechanisms underlying trait variation. Thus, we detected and annotated into functional categories single nucleotide variants (SNVs) and insertion/deletions (INDELs) in the Guzerat, Gyr, and Girolando cattle breeds.

**Results:** There were 13,253, polymorphic genes in in Guzerat, 13,817 in Gyr, and 12,480 in Girolando, of which 64.84% of genes with SNVs and 31.29% of genes with INDELs were shared among these breeds. The functional enrichment analysis revealed 24, 27, and 28 enriched KEGG pathways (FDR<10%) in Guzerat, Gyr, and Girolando, respectively, of which 14 pathways were common to all breeds. In terms of pathways local to only one breed, three were significantly over-represented only in Guzerat (arachidonic acid metabolism, Fc gamma R-mediated phagocytosis, and aldosterone-regulated sodium reabsorption), six were over-represented only in Gyr (alanine, aspartate and glutamate metabolism, inflammatory mediator regulation of TRP channels, thyroid hormone synthesis, pancreatic secretion, central carbon metabolism in cancer, and choline metabolism in cancer), and eight were over-represented only in Girolando (amino sugar and nucleotide sugar metabolism, inositol phosphate metabolism, vascular smooth muscle contraction, tight junction, regulation of actin cytoskeleton, amoebiasis, small cell lung cancer, and dilated cardiomyopathy). Although the identified polymorphic genes were distributed throughout the genomes, a large number of novel variants were clustered in specific genes. A total of 61 genes were enriched with novel variants common to all breeds, while 349, 404, and 206 genes were enriched with novel variants exclusively in the Guzerat, Gyr, and Girolando breeds, respectively.

**Conclusion:** The genes enriched with novel variants identified in this study are related to many biological processes, providing valuable information about genomic variants that may be responsible for variation in adaptive and economically important traits among these breeds.

**Keywords:** Biological Process, Insertion, Deletion, Novel variants, Pathways, SNVs