

### **Análise de regiões ricas em heterozigose em diferentes populações de bovinos**

Rita Carolina Gaia<sup>1</sup>, Henrique Alberto Mulim<sup>2</sup>, Luiz F. Brito<sup>3</sup>, Pamela Carla Machado<sup>1</sup>, Emanueli de Fátima Pereira da Silva<sup>1</sup>, Gabriele Ratke Morgan<sup>1</sup>, Leticia Sikorski Caldeira<sup>1</sup>, José Bento Sterman Ferraz<sup>4</sup>, Lais Grigoletto<sup>4</sup>, Marcio Ribeiro Silva<sup>5</sup>, Luís Fernando Batista Pinto<sup>2</sup>, Victor Breno Pedrosa<sup>1</sup>

<sup>1</sup>Departamento de Zootecnia, Universidade Estadual de Ponta Grossa, Ponta Grossa, Paraná, Brasil.

<sup>2</sup>Departamento de Zootecnia, Escola de Medicina Veterinária e Zootecnia da UFBA, Salvador, Bahia, Brasil

<sup>3</sup>Departamento de Zootecnia, Purdue University, West Lafayette, Indiana, Estados Unidos

<sup>4</sup>Departamento de Zootecnia, Faculdade de Zootecnia e Engenharia de Alimentos da USP, Pirassununga, São Paulo, Brasil

<sup>5</sup>Melhor Animal e Katayama Agropecuária Ltda., Guararapes, São Paulo, Brasil

**Resumo:** O objetivo desse estudo foi identificar regiões ricas em heterozigose, existentes no genoma de diferentes populações de bovinos, verificar seu comprimento e analisar os genes associados a estas regiões. Foram avaliados 16 rebanhos distintos, de 15 raças diferentes. Para a observação das regiões ricas em heterozigose foi utilizado o pacote detecRUNS, que detectou 3.576 regiões em heterozigose dentre as populações avaliadas. A população da raça Montana se destacou por apresentar o maior número de regiões ricas em heterozigose, com 1.702 corridas, além de apresentar o maior comprimento de heterozigose, com 4.25 MB, responsável pela expressão de 114 genes. A população da raça Nelore, genotipada com o chip de 35k, apresentou 662 regiões ricas em heterozigose e nos bovinos da mesma raça com chip de 50k foram encontradas 675 regiões. A menor região com concentração de heterozigose foi encontrada em bovinos da raça Nelore, com o comprimento de cerca de 0,7 MB, responsável pela expressão de somente um gene, o *KHDRBS2*, habitualmente relacionado a fertilidade. Grande parte dos genes encontrados nas regiões ricas em heterozigose são relacionados a imunidade, o que pode contribuir com a adaptação dos animais a agentes estressores.

**Palavras-chave:** corridas de heterozigose, diversidade genômica, melhoramento genético

### **Analysis of heterozygous-enriched regions in different populations of cattle**

**Abstract:** The aim of this study was to identify regions rich in heterozygosity, existing in the genome of different populations of cattle, verify their length and analyze the genes associated with these regions. Sixteen different herds of 15 different breeds were evaluated. For the observation of regions rich in heterozygosity, the detecRUNS package was used, which detected 3,576 regions in heterozygosity among the populations evaluated. The Montana breed population stood out for having the largest number of regions rich in heterozygosity, with 1,702 runs, in addition to having the longest heterozygosity length, with 4.25 MB, responsible for the expression of 114 genes. The population of the Nelore breed, genotyped with the 35k chip, showed 662 regions rich in heterozygosity and in the same breed cattle with the 50k chip, 675 regions were found. The smallest region with concentration of heterozygosity was found in Nelore cattle, with a length of about 0.7 MB, responsible for the expression of only one gene, *KHDRBS2*, usually related to fertility. A large part of the genes found in regions rich in heterozygosity are related to immunity, which may contribute to the animals' adaptation to stressors.

**Keywords:** animal breeding, genomic diversity, runs of heterozygosity

### **Introdução**

Os estudos de regiões ricas em heterozigose (RRH) se caracterizam por identificar as regiões do genoma com alta variabilidade e podem prover valiosas informações sobre a diversidade das populações e histórico evolutivo (Samuels et al., 2016). Este tipo de abordagem ainda foi pouco explorado, principalmente em rebanhos bovinos, o que faz com que os estudos de RRH sejam inovadores, especialmente com um grande número de populações. A identificação de regiões em heterozigose no genoma de animais cruzados e puros, são de extrema importância pois estas regiões afetam características relacionadas a produção animal. Por essa razão, o objetivo desse estudo foi identificar em 16 rebanhos

bovinos, regiões do genoma com alta variabilidade, através das análises de regiões ricas em heterozigose, além de detectar os genes associados a estas regiões.

### Material e Métodos

Foram utilizadas informações genóticas de 16 rebanhos bovinos, em 15 raças distintas. Os dados foram oriundos de quatro bases de dados distintas: Purdue University – com dados dos animais cruzados Angus x Simental (ANGSIM - 487), Pecuária Katayama – com dados dos animais Nelore (NEL50 – 192, NEL35 – 209), Universidade de São Paulo – com dados dos animais da raça composta Montana (MON – 271), e da plataforma WIDDE (<http://widde.toulouse.inra.fr/widde/>) fornecendo os dados das raças Angus (ANG - 99), Borgou (BOR – 158), Brahman (BRM – 70), Criolo de Guadalupe (CGU – 140), Charolês (CHL – 62), Gir (GIR – 50), Hereford (HFD – 61), Holandês (HOL – 137), Jersey (JER – 84), Limousin (LMS – 87), Senepol (SEN – 153) e Santa Gertrudis (SGT – 55). Para o controle de qualidade foi considerada a remoção de genótipos com call rate ( $>0.9$ ), SNPs duplicados, não-autossomais ou sem posição definidas. Tal limpeza possui importância para assegurar que nenhum outro fator além dos parâmetros escolhidos interferisse na detecção das corridas. Para a observação das regiões ricas em heterozigose, foi utilizado o pacote detecRUNS (Biscarini et al., 2019) e a metodologia baseada em SNPs consecutivos. Os parâmetros admitidos foram:

- O número mínimo de SNPs consecutivos que constitui uma RRH de 20 SNPs
- O tamanho mínimo de um RRH de 500 KB
- Permitindo o número mínimo de 2 genótipos homozigóticos e 1 perdido.
- O gap entre SNPs consecutivos não pôde ser maior que 1000 kb;

Além disso, as regiões genômicas que apresentaram incidência de RRH em pelo menos 10% da população foram separadas para posterior análises funcionais através do pacote GALLO (Fonseca et al., 2020).

### Resultados e Discussão

Ao todo foram encontradas 3.576 RRH's nos 16 rebanhos avaliados, em que a distribuição de RRH por população pode ser observada na Figura 1. As populações que apresentaram os maiores números de RRH foram as populações MON com 1.702 corridas, NEL35 com 662 e a população NEL50 com 675. As menores quantidade de corridas foram encontradas nas populações CGU com 2 RRH e BRM com 4 RRH encontradas.

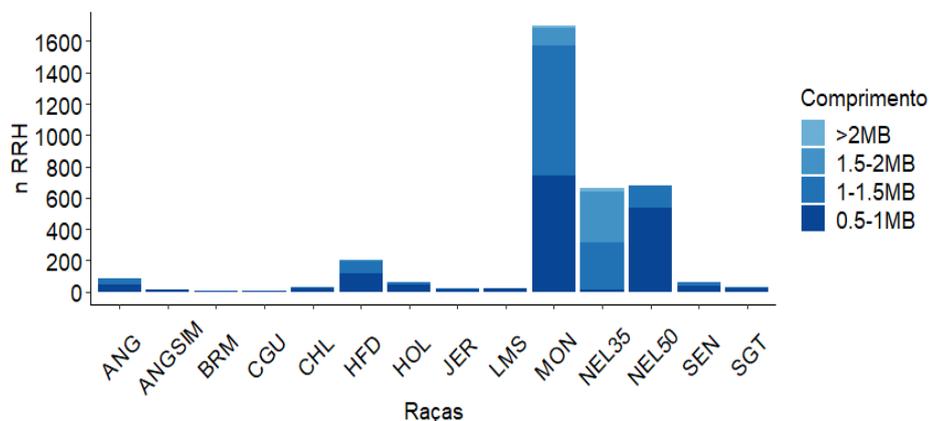
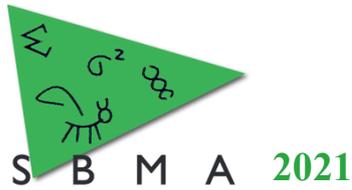


Figura 1. Distribuição das regiões em heterozigose nas diferentes populações de bovinos

De modo geral, as maiores porcentagens das RRH foram classificadas como sendo de comprimento entre 0.5-1MB (45,13%) e de 1-1.5MB (41,11%). Sendo a porcentagem de RRH maiores que 2MB somente encontrada em 1% das RRH observadas. A população a apresentar a maior quantidade de RRH foi a população MON, apresentando mais de 47% das RRH encontradas em todo estudo. Resultado esse



esperado, uma vez que a população MON é uma população composta, que propõe na sua formação combinar grupos definidos por características biológicas, fisiológicas, de crescimento e reprodução, cruzando animais *Bos taurus indicus* e *Bos taurus taurus* (Peripolli et al., 2020).

A maior região de RRH foi encontrada na população MON com 4.25 MB de comprimento e abrange 20,30% da população. Tal região está associada a 114 genes, dos quais 71 são proteínas codificantes. A menor região com concentração de RRH foi encontrada nos animais NEL50 no cromossomo 23 com o comprimento de cerca de 0,7 MB. Tal região é responsável pela expressão de somente um gene, o *KHDRBS2*. Alguns estudos já demonstraram que este gene está relacionado com a fertilidade e performance reprodutiva de bovinos, como apresentado por Reverter et al. (2016) e De León et al. (2019). Apesar disso, a maioria dos genes encontrados nas regiões ricas em heterozigose apresentaram-se com alguma relação a imunidade dos animais sejam em vias ou processos biológicos em respostas a alguns desafios.

#### Conclusão

A quantidade e o padrão de aparecimento das regiões ricas em heterozigose foram específicas de cada população. Foram encontradas 3.576 regiões ricas em heterozigose, em que a população a apresentar a maior quantidade de regiões ricas em heterozigose foi a de bovinos da raça Montana, com total de 1.702. Concentrações de regiões ricas em heterozigose, com maior comprimento, foram observadas na população Montana com 4,35 MB e a menor na população Nelore (0,7 MB), associada ao gene *KHDRBS2*, habitualmente relacionado a fertilidade. A maioria dos genes encontrados nas regiões ricas em heterozigose foram relacionados a imunidade animal, fato que pode contribuir com a adaptação a agentes estressores.

#### Agradecimentos

A Katayama Pecuária Ltda., Universidade de São Paulo e Purdue University pelo fornecimento dos bancos de dados. Ao grupo de estudo LeMA – Laboratório de Estudos em Melhoramento Animal, por todo auxílio na pesquisa.

#### Literatura citada

- Biscarini F, Cozzi P, Gaspa G, Marras G (2019) detectRUNS: an R package to detect runs of homozygosity and heterozygosity in diploid genomes. *CRAN.R*. doi:1.
- Fonseca PAS, Suárez-Vega A, Marras G, Cánovas Á (2020) GALLO: An R package for genomic annotation and integration of multiple data sources in livestock for positional candidate loci. *GigaScience* **9**, 1–9. doi:10.1093/gigascience/giaa149.
- De León C, Manrique C, Martínez R, Rocha JF (2019) Research Article Genomic association study for adaptability traits in four Colombian cattle breeds. *Genetics and Molecular Research* **18**,. doi:10.4238/gmr18373.
- Peripolli E, Stafuzza NB, Amorim ST, Lemos MVA, Grigoletto L, Kluska S, Ferraz JBS, Eler JP, Mattos EC, Baldi F (2020) Genome-wide scan for runs of homozygosity in the composite Montana Tropical ® beef cattle. *Journal of Animal Breeding and Genetics* **137**, 155–165. doi:10.1111/jbg.12428.
- Reverter A, Porto-Neto LR, Fortes MRS, McCulloch R, Lyons RE, Moore S, Nicol D, Henshall J, Lehnert SA (2016) Genomic analyses of tropical beef cattle fertility based on genotyping pools of Brahman cows with unknown pedigree1. *Journal of Animal Science* **94**, 4096–4108. doi:10.2527/jas.2016-0675.
- Samuels DC, Wang J, Ye F, He J, Levinson RT, Sheng Q, Zhao S, Capra JA, Shyr Y, Zheng W, Guo Y (2016) Heterozygosity Ratio, a Robust Global Genomic Measure of Autozygosity and Its Association with Height and Disease Risk. *Genetics* **204**, 893–904. doi:10.1534/genetics.116.189936.