

SILVA; Bruna Pereira Martins da ¹, FANALLI; Simara Larissa ², FREITAS; Felipe André Oliveira ³, AFONSO; Juliana ⁴, COUTINHO; Luiz Lehmann ⁵

RESUMO

Polimorfismos de Nucleotídeo Único ou SNPs (*Single Nucleotide Polymorphisms*) são o tipo mais simples e abundante de variação genética encontrada no genoma. São originados de mutações pontuais que ocorrem no DNA e podem estar presentes em regiões codificantes (exons) e não codificantes (ítrons e regiões intergênicas). As tecnologias de Sequenciamento de Nova Geração (NGS) como a aplicada à sequências de RNA (RNA-Seq), quando associadas a ferramentas computacionais, são um excelente recurso para busca e identificação de marcadores moleculares do tipo SNP em regiões codificantes. Esses estudos de polimorfismos dentro de regiões codificantes são fundamentais, pois abordam as mutações que ocorrem em regiões transcritas dos genes, aumentando a chance de encontrar variantes funcionais. Esses marcadores moleculares do tipo SNP permitem buscar variações nas sequências gênicas que possam estar associadas a um fenótipo de interesse. O objetivo deste estudo foi identificar polimorfismos do tipo SNP a partir do sequenciamento do transcriptoma (conjunto dos transcritos, RNA mensageiro) do músculo de suínos. Foram utilizados 72 machos da raça *Large White*, imunocastrados e halotano homozigoto-negativo (NN) que foram abatidos com 133.9 ± 9.4 kg de peso vivo e 169 dias de idade. Foi realizada a extração de RNA total do tecido muscular *Longissimus dorsi* seguida pelo sequenciamento do mRNA por meio de análise de RNA-Seq. A partir dos dados de RNA-Seq, foi realizada a identificação dos SNPs em regiões codificantes do DNA por meio de ferramentas da bioinformática. Como primeiro passo foi realizada uma filtragem para remover a lista de *contigs* dos arquivos com extensão .bam, resultantes do sequenciamento, e em seguida foi realizada a chamada de SNPs pelo programa GATK (*Genome Analysis Toolkit*) versão 4.1.9.0. Como passo final, foram realizadas filtragens dos dados segundo os parâmetros: (1) somente SNP; (2) profundidade (DP ≥ 10); (3) qualidade da chamada de SNPs (QUAL ≥ 30); (4) remoção de monomórficos; e (5) remoção de SNPs com baixa frequência alélica (*Minor Allele Frequency*, MAF $\leq 0,05$). Por fim, a anotação dos SNPs que passaram pelos filtros de qualidade foi feita com o uso da ferramenta VEP (*Variant Effect Predictor*, Ensembl). Como resultados foram encontrados um total de 77.043 polimorfismos do tipo SNP, distribuídos entre os cromossomos autossômicos e DNA mitocondrial. Destes, 3.035 marcadores foram classificados como deletérios, marcadores decorrentes de mutações não sinônimas, que resultam na alteração do códon e podem gerar modificações na composição e/ou ausência de aminoácidos, e consequentemente, modificações nas proteínas. Por outro lado, 16.207 variantes foram classificadas como tolerantes, marcadores decorrentes de mutações que alteram o códon, mas não alteram o aminoácido que será traduzido, não alterando a proteína. Além desses, 12.155 foram considerados novos SNPs, ou seja, não foram identificados até o momento por outros autores. Os resultados deste estudo exploratório trazem uma contribuição relevante para estudos futuros relacionados a identificação e associação de variações (já conhecidas e novas) no padrão de expressão de genes que possam estar relacionados a características de interesse, contribuindo com o melhoramento genético dessa espécie.

PALAVRAS-CHAVE: Melhoramento genético e reprodução animal, alteração de nucleotídeo,

¹ Pós-graduanda - FZEA/USP, brunamartins@usp.br

² Pós-graduanda - FZEA/USP, simarafanalli@usp.br

³ Pós-graduando - ESALQ/USP, felipeoliveirafreitas@usp.br

⁴ Doutora - EMBRAPA, juafonsobio@gmail.com

⁵ Professor Doutor - ESALQ/USP, llcoutinho@usp.br

¹ Pós-graduanda - FZEA/USP, brunamartins@usp.br

² Pós-graduanda - FZEA/USP, simarafanalli@usp.br

³ Pós-graduando - ESALQ/USP, felipeoliveirafreitas@usp.br

⁴ Doutora - EMBRAPA, juafonsobio@gmail.com

⁵ Professor Doutor - ESALQ/USP, llcoutinho@usp.br