

Perfil transcricional de RNAs não codificantes em amostras do tecido subcutâneo dorsal adiposo de duas raças de suínos

Transcriptional profile of the non-coding RNA in the dorsal adipose subcutaneous tissue samples of two swine breeds

RESUMO

A crescente demanda por alimentos reflete na necessidade de abordagens para aumento da produtividade e qualidade de alimentos, onde a genética possui papel protagonista. Pesquisadores no mundo todo têm se dedicado em estudos de RNAs reguladores para o entendimento de sua contribuição em diferentes organismos. Este trabalho objetivou analisar a expressão diferencial de transcritos (mRNA, ncRNA e lncRNA) entre duas raças de suínos, através de métodos computacionais de análise de transcriptoma. Para isso, foram coletadas as informações e mapeadas a partir de um genoma de referência. Posteriormente, foram identificados os transcritos presentes e analisado a expressão diferencial dos mesmos, selecionando os 10 RNAs regulados positivamente e 10 regulados negativamente. Com os resultados, foi possível detectar RNAs codificantes e não codificantes com um nível de expressão diferencial semelhante e possivelmente intrínsecos à raça de suínos. Assim, essas análises poderão trazer base para futuras pesquisas para abranger a co-expressão desses RNAs e confirmar se estão na mesma via de regulação.

PALAVRAS-CHAVE: Genômica funcional. Bioinformática. RNAs regulatórios.

ABSTRACT

The growing demand for food reflects the need for an approach to increase productivity and quality of food, where genetics has a leading role. Researchers around the world have dedicated themselves to studies of regulatory RNAs to understand their contribution in different organisms. This work aimed to analyze a differential expression of transcripts (mRNA, ncRNA and lncRNA) between two pig breeds, using computational methods of transcriptome analysis. For this, the information was collected and mapped from a reference genome. The transcripts present and analyze their differential expression, determining the 10 positively regulated and 10 negatively regulated RNAs. With the results, it was possible to detect coding and non-coding RNAs with a similar level of differential expression and possibly intrinsic to the swine breed, these analyzes provide a basis for future research to cover the co-expression of these RNAs and confirm whether they are on the same path of regulation.

KEYWORDS: Functional genomic. Bioinformatics. Regulatory RNA.

Larissa Fernanda Quintino

larif.q@live.com

Universidade Tecnológica Federal do Paraná, Dois Vizinhos, Paraná, Brasil

Nédia de Castilhos Ghisi

nediaighisi@utfpr.edu.br

Universidade Tecnológica Federal do Paraná, Dois Vizinhos, Paraná, Brasil

Henrique Moura Dias

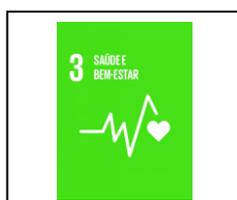
diash@usp.br

Universidade de São Paulo, São Paulo, São Paulo, Brasil

Recebido: 04 set. 2020.

Aprovado: 01 out. 2020.

Direito autoral: Este trabalho está licenciado sob os termos da Licença Creative Commons-Atribuição 4.0 Internacional.



INTRODUÇÃO

De acordo com a ONU, a população mundial irá aumentar significativamente em 2050, ultrapassando os 9 bilhões de habitantes no mundo (ONU, 2015). Com o aumento da demanda por alimentos (FAO, 2017), ferramentas de genética serão um dos principais pilares para a melhoria da cadeia produtiva da suinocultura (KHATIB, 2014).

Nos últimos anos, estudos mostraram a importância de RNAs não codificantes (ncRNAs) em mecanismos de regulação de células animais (DO et al., 2017). Essas moléculas podem influenciar em diversas características de importância produtiva como qualidade do leite, susceptibilidade à mastite (TONG et al., 2017), deposição de gordura e crescimento muscular (WANG et al., 2015).

Os ncRNAs são caracterizados como um conjunto de transcritos que não detêm capacidade de codificar para proteínas (RAO, 2017; BORGA; MEERAN; FASSAN, 2019). Diferenças como tamanho e função são algumas das características de distinção entre essas moléculas, que podem compor sequências maiores ou menores de 200 nucleotídeos. Dada a diversidade de ncRNAs, o local e a dinâmica celular de sua atuação também podem variar (BORGA; MEERAN; FASSAN, 2019). Os ncRNAs se dividem em dois grandes grupos, estruturais (RNA ribossômico – rRNA e RNA de Transferência - tRNA) e reguladores (pequenos RNAs e RNAs longos) (MENCK; VAN SLUYS, 2017).

Os mecanismos envolvidos em características intrínsecas à carne de suínos, como a gordura intramuscular, incluem a associação de RNAs longos não codificantes (lncRNA). Apesar de serem pouco descritos na literatura, os lncRNAs, podem regular a via de sinalização de adipocitoquinas, deposição e degradação de gordura (ZOU et al., 2018). A mesma relação ocorre com pequenos RNAs, como o caso de micro RNAs (miRNA) no músculo de suínos que atuam na dinâmica da adipogênese (SUN et al., 2019), metabolismo de lipídios e crescimento muscular (WANG et al., 2015).

Portanto, o objetivo do trabalho foi analisar a expressão diferencial de transcritos, com enfoque em mRNA, ncRNA, e lncRNA de acordo com a raça de suínos (*Sus scrofa*), através de métodos computacionais de análise de transcriptoma.

METODOLOGIA

Os dados selecionados para as análises derivam de experimentos realizados por Albuquerque et al. (2020) depositados no banco de dados público *National Center for Biotechnology Information - NCBI* (<https://www.ncbi.nlm.nih.gov/>) sob o código de identificação GSE145956. A biblioteca de RNAseq compõem o sequenciamento de RNA por metodologia Illumina Hiseq 2000 *paired* de duas raças de suínos (Alentajano e Bisaro) com oito réplicas biológicas. Todas as corridas foram submetidas a análise de qualidade, com remoção de sequências com baixa qualidade e adaptadores, com auxílio das ferramentas FastQC e Trim Galore na plataforma Galaxy Europe (<https://usegalaxy.eu/>). O mapeamento foi feito com a ferramenta Hisat2 (KIM; LANGMEAD; SALZBERG, 2015) com parâmetros de análise padrão utilizando o genoma de referência *Sus scrofa* 11.1 (GCA_000003025.6) depositado no banco de dados Ensembl (WARR et al., 2020). Os arquivos de

mapeamento (extensão BAM) gerados pelo Hisat2 foram utilizados para quantificação de cada transcrito, através da ferramenta Htseq-count (ANDERS; PYL; HUBER, 2014) com sua respectiva indexação pelo arquivo de anotação Sus_scrofa.Scrofa11.1.100 (WARR et al., 2020). As análises de expressão diferencial foram realizadas com o pacote de análises edgeR (ROBINSON; MCCARTHY; SMYTH, 2009).

RESULTADOS E DISCUSSÃO

Os dados de contagem de *reads* de mRNA, ncRNA e lncRNA, foram submetidas a análise de expressão diferencial, a fim de verificar o perfil transcricional sob a influência das diferenças entre as raças na deposição de gordura subcutânea dorsal. Em análise dos dados de expressão diferencial, foi possível identificar maior número de transcritos diferencialmente expressos para mRNAs codificantes, em comparação a transcritos de ncRNAs e lncRNA.

Ao todo foram encontrados 21.303 transcritos de mRNA codificante, 8.979 transcritos de ncRNAs e 9.374 transcritos de lncRNA. Um resumo detalhado da análise de transcritos para cada classe são apresentados no Quadro 1. Nota-se que os lncRNAs (classe específica dentro de ncRNA) apresenta maior número de transcritos que ncRNA (classe maior e mais genérica), isso se deve ao fato de que existem lncRNAs que são derivados de porções gênicas codificantes, sendo assim processados pelo maquinário de regulação pós-transcricional. A abundância e expressão diferencial de ncRNAs e lncRNAs, indica uma importante participação dessas classes de RNAs em processos regulatórios que podem ser intrínsecos a nível intraespecífico.

Quadro 1 – Total de transcritos diferencialmente expressos por classe, em análise comparativa entre as raças Alentajano e Bisaro.

	Positivamente	Normal	Negativamente	N° Transcritos
Codificante	2760	15942	2601	21303
ncRNA	211	8572	196	8979
lncRNA	185	9086	103	9374
Total	3156	33600	2900	39656

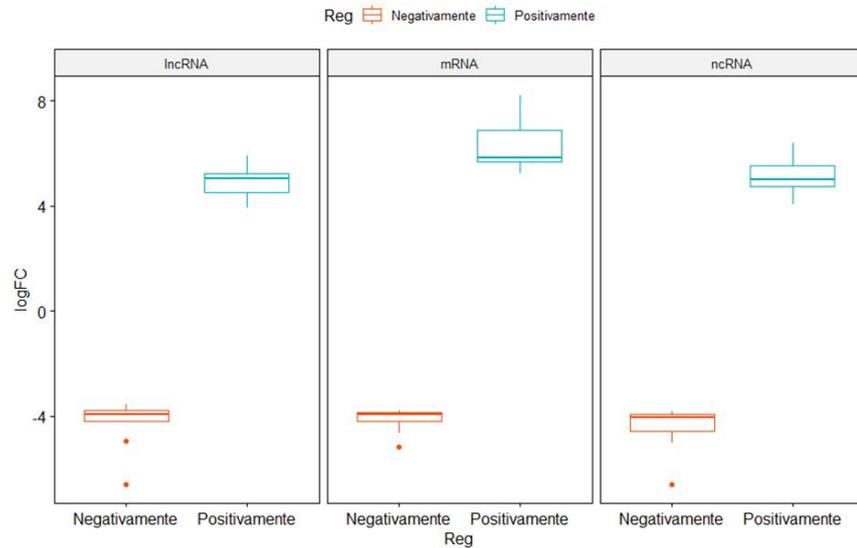
Fonte: Autoria própria (2020).

Foram selecionados dez transcritos com maior fold change (positivo e negativo) para cada classe (mRNA, ncRNA e lncRNA), com base no FDR - valor p ajustado (> 0,05). É possível assumir que o perfil dos dez transcritos com maior logFC (*fold change*) positivamente e negativamente entre mRNA, ncRNA e lncRNA não se diferem e apresentam menor variação tratando dos transcritos negativamente (Figura 1). Uma análise mais refinada, onde sugere-se inferência da co-expressão de transcritos, pode auxiliar na identificação de padrões na rede de interação de expressão entre transcritos codificantes e não codificantes.

Nossas análises sugerem que a maioria dos transcritos selecionados como diferencialmente mais ou menos expressos (negativamente e positivamente), apresentam expressão característica da raça, ou seja pode ter uma contribuição significativa da particularidade do genoma de cada raça, com exceção do ncRNA

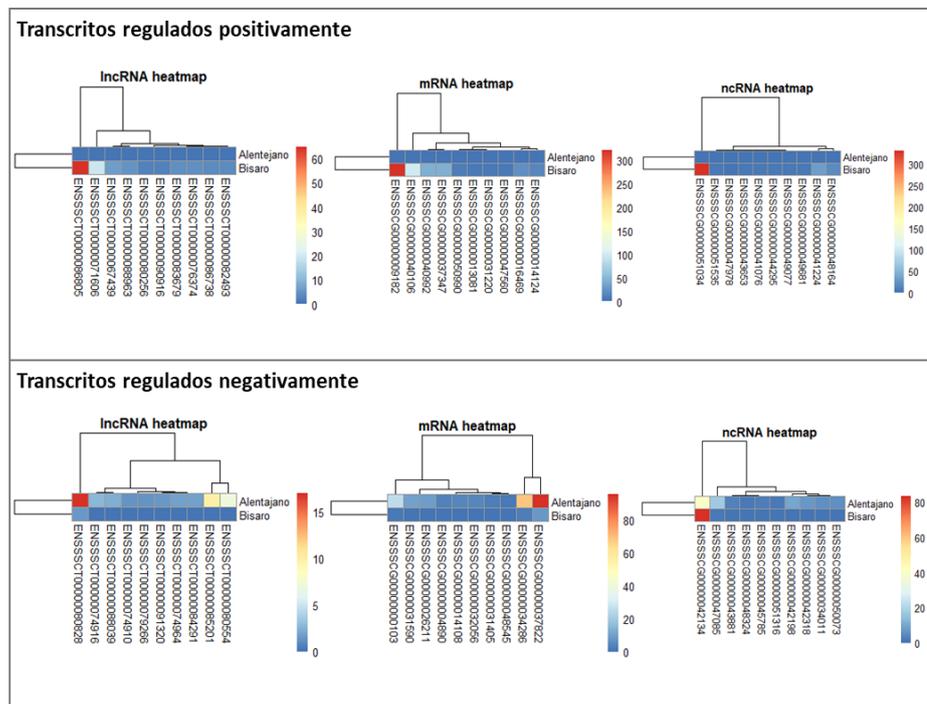
“ENSSSCG00000042134” que possui expressão significativa para ambas raças de suínos analisadas (Figura 2).

Figura 1 – Boxplot do logFC de transcritos regulados negativamente e positivamente na análise de comparação entre raças de suínos Alentajano x Bisaro. Transcritos filtrados com FDR = 0,05.



Fonte: Autoria própria (2020).

Figura 2 – Heatmap com o perfil da expressão de transcritos de RNA para duas raças de suínos. Transcritos filtrados com FDR = 0,05. Valores de expressão determinados em CPM (counts per million).



Fonte: Autoria própria (2020).

CONCLUSÃO

Muitos transcritos de lncRNA ainda permanecem sem identificação em bancos de dados, tal como *Ensembl*. Isto ressalta a importância de estudar essas moléculas e sua funcionalidade em organismos animais. A identificação dessas moléculas, bem como outros ncRNAs, poderão atingir uma visão abrangente dos mecanismos de regulação transcricional em animais. Análises de bioinformática podem contribuir significativamente na prospecção e identificação de alvos associados às vias de regulação específicas, como visto em tecidos adiposos de suínos em análise comparativa entre raças de uma mesma espécie. Sugere-se que a partir destes dados pilotos, sejam realizadas análises de redes de co-expressão com intuito de identificar conjuntos regulatórios destes transcritos não codificantes associados à genes, para possível classificação funcional *in silico*.

AGRADECIMENTOS

Agradeço a Universidade Tecnológica Federal do Paraná – UTFPR pela oportunidade.

REFERÊNCIA BIBLIOGRÁFICA

ALBUQUERQUE, André et al. Comparative Transcriptomic Analysis of Subcutaneous Adipose Tissue from Local Pig Breeds. **Genes**, [S.l.], v. 11, n. 4, 15 abr. 2020. MDPI AG. <http://dx.doi.org/10.3390/genes11040422>.

ANDERS, S.; PYL, P. T.; HUBER, W. HTSeq--a Python framework to work with high-throughput sequencing data. **Bioinformatics**, [S.l.], v. 31, n. 2, p. 166-169, 25 set. 2014. Oxford University Press (OUP). <http://dx.doi.org/10.1093/bioinformatics/btu638>.

BORGA, Chiara; MEERAN, Syed M.; FASSAN, Matteo. Non-coding RNAs, a real Next-Gen Class of Biomarkers. **Non-coding Rna Research**, [S.l.], v. 4, n. 3, p. 80-81, set. 2019. Elsevier BV. <http://dx.doi.org/10.1016/j.ncrna.2019.10.001>.

DO, Duy N. et al. Transcriptome Analysis of Non-Coding RNAs in Livestock Species: elucidating the ambiguity. **Applications Of Rna-seq And Omics Strategies - From Microorganisms To Human Health**, [S.l.], 13 set. 2017. InTech. <http://dx.doi.org/10.5772/intechopen.69872>.

FAO. **The future of food and agriculture – Trends and challenges**. Roma, 2017.

KHATIB, H. **Molecular and quantitative animal genetics**. Hoboken: Wiley-blackwell, 2014. 336 p.

KIM, Daehwan; LANGMEAD, Ben; SALZBERG, Steven L. HISAT: a fast spliced aligner with low memory requirements. **Nature Methods**, [S.l.], v. 12, n. 4, p.

357-360, 9 mar. 2015. Springer Science and Business Media LLC.
<http://dx.doi.org/10.1038/nmeth.3317>.

MENCK, Carlos F. M.; VAN SLUYS, Marie-anne. **Genética Molecular Básica: dos genes aos genomas**. [S. I.]: Guanabara Koogan, 2017. 528 p.

ONU. **World Population Prospects: key findings and advance tables**. New York: United Nations, 2015. 59 p. Disponível em:
https://population.un.org/wpp/Publications/Files/Key_Findings_WPP_2015.pdf.
Acesso em: 07 jul. 2020.

RAO, M.R.S. **Long Non Coding RNA Biology**. [S. I.]: Springer, 2017. 323 p.

ROBINSON, M. D.; MCCARTHY, D. J.; SMYTH, G. K.. EdgeR: a bioconductor package for differential expression analysis of digital gene expression data. **Bioinformatics**, [S.I.], v. 26, n. 1, p. 139-140, 11 nov. 2009. Oxford University Press (OUP). <http://dx.doi.org/10.1093/bioinformatics/btp616>.

SUN, Yan et al. Profiling and characterization of miRNAs associated with intramuscular fat content in Yorkshire pigs. **Animal Biotechnology**, [S.I.], v. 31, n. 3, p. 256-263, 25 abr. 2019. Informa UK Limited.
<http://dx.doi.org/10.1080/10495398.2019.1573191>.

TONG, Chao et al. Identification and characterization of long intergenic noncoding RNAs in bovine mammary glands. **Bmc Genomics**, [S.I.], v. 18, n. 1, 19 jun. 2017. Springer Science and Business Media LLC.
<http://dx.doi.org/10.1186/s12864-017-3858-4>.

WANG, Zhixiu et al. Identification of Genes Related to Growth and Lipid Deposition from Transcriptome Profiles of Pig Muscle Tissue. **Plos One**, [S.I.], v. 10, n. 10, 27 out. 2015. Public Library of Science (PLOS).
<http://dx.doi.org/10.1371/journal.pone.0141138>.

WARR, Amanda et al. An improved pig reference genome sequence to enable pig genetics and genomics research. **Gigascience**, [S.I.], v. 9, n. 6, 1 jun. 2020. Oxford University Press (OUP). <http://dx.doi.org/10.1093/gigascience/giaa051>.

ZOU, Cheng et al. Identification and Functional Analysis of Long Intergenic Non-coding RNAs Underlying Intramuscular Fat Content in Pigs. **Frontiers In Genetics**, [S.I.], v. 9, 27 mar. 2018. Frontiers Media SA.
<http://dx.doi.org/10.3389/fgene.2018.00102>.