

**UNIVERSIDADE FEDERAL DE SÃO CARLOS
BACHARELADO EM BIOTECNOLOGIA**

KARINA SANTOS DE OLIVEIRA

**MECANISMOS DE REGULAÇÃO GÊNICA RELACIONADOS À EFICIÊNCIA
ALIMENTAR EM BOVINOS DA RAÇA NELORE**

SÃO CARLOS - SP

2021

KARINA SANTOS DE OLIVEIRA

**MECANISMOS DE REGULAÇÃO GÊNICA RELACIONADOS À EFICIÊNCIA
ALIMENTAR EM BOVINOS DA RAÇA NELORE**

Trabalho de conclusão de curso apresentado
ao Departamento de Genética e Evolução da
Universidade Federal de São Carlos para
obtenção do título de Bacharel em
Biotecnologia.

Orientador(a): Luciana Correa de Almeida Regitano

Co-Orientador: Marcos Roberto Chiaratti

SÃO CARLOS - SP

2021

RESUMO

O Brasil é um dos principais produtores e exportadores de carne bovina, com rebanhos compostos predominantemente por animais zebuínos da raça Nelore (*Bos indicus*) e seus cruzamentos. Nesse cenário, a genômica pode contribuir na identificação de potenciais biomarcadores de características de interesse agropecuário. Em estudos prévios, *COL1A1*, *CTGF*, *CYP2B6*, *EGR1* e *PRUNE2* foram genes diferencialmente expressos (GDEs) entre grupos de animais contrastantes para o fenótipo consumo alimentar residual em tecido hepático. O objetivo deste trabalho foi selecionar polimorfismos de nucleotídeo único (SNPs) que potencialmente afetam regiões reguladoras nos genes citados e associá-los a diferentes medidas fenotípicas de eficiência alimentar (EA). Para isso, utilizando genótipos imputados de 398 novilhos Nelore, foram selecionados SNPs em uma janela compreendendo desde 2 Kb *upstream* ao sítio de início da transcrição até o final da 3' UTR descrita dos GDEs. Na anotação funcional, realizada com o *software* VEP, foram desconsideradas variantes em regiões intrônicas, sinônimas e as que não são possíveis reguladoras de *splicing*. Para os SNPs selecionados, foi analisado o desequilíbrio de ligação pelos *softwares* PLINK e Haploview, e foi feita a predição *in silico* de sítios de ligação de fatores de transcrição (SLFTs) e de microRNAs (miRNAs) utilizando os *softwares* LASAGNA e RNAhybrid, respectivamente. Um modelo de regressão linear foi aplicado para associar os SNPs às medidas fenotípicas de EA, e os SNPs significativos foram comparados a regiões de *quantitative trait loci* (QTLs) de fenótipos de interesse pecuário provenientes de trabalhos na mesma população experimental e do *Cattle QTL database*. Após anotação funcional, foram selecionados dois SNPs na região *upstream* do gene *CTGF*; três referentes ao *COL1A1*, um em região de *splicing* e dois *upstream* ao gene; 15 para o gene *CYP2B6*, dentre esses, uma variante *missense*, quatro em 5' UTR e três em 3' UTR, *upstream* e *downstream*; cinco para o gene *EGR1*, um desses na região 3' UTR e quatro na região *upstream*; e 33 SNPs referentes ao gene *PRUNE2*, três em região *upstream* e seis em *downstream*, e 24 variantes *missenses*. Dentre esses, os SNPs relativos aos cinco GDEs possuem potencial em influenciar em SLFTs e também foram integrados a regiões de QTLs para diferentes fenótipos bovinos. SNPs dos genes *CYP2B6* e *EGR1* apresentaram potencial em influenciar na ligação de diversos miRNAs bovinos. Por fim, dois SNPs localizados

no gene *CTGF* foram associados com a medida fenotípica eficiência alimentar bruta, um SNP do gene *COL1A1* apresentou associação com as medidas ganho de peso diário e conversão alimentar, e 19 SNPs no *PRUNE2* foram associados com consumo alimentar residual. Os SNPs identificados contribuíram para aumentar o conhecimento sobre a possível influência de regiões regulatórias de genes envolvidos com características fenotípicas de EA.

Palavras-chave: *Bos indicus*. SNP. Expressão gênica. Fator de transcrição. MicroRNA. Fenótipo.

ABSTRACT

Brazil is one of the main beef producers and exporters, with herds predominantly composed of Nellore Zebu animals (*Bos indicus*). In this scenario, genomics can contribute to the identification of potential biomarkers of agricultural interest features. In previous studies, *COL1A1*, *CTGF*, *CYP2B6*, *EGR1* and *PRUNE2* were differentially expressed genes (DEGs) between groups of animals contrasting for the residual feed intake phenotype in liver tissue. The objective of this work was to select single nucleotide polymorphisms (SNPs) that potentially affect regulatory regions in the genes mentioned and associate them with different phenotypic measures of feed efficiency (FE). For this, using imputed genotypes of 398 Nellore steers, SNPs were selected in a window from 2 Kb upstream to the transcription start site to the end of the 3' UTR reduced of the DEGs. In the functional annotation, performed with the VEP software, variants in intronic regions, synonyms and those that are not possible regulatory splicing were not considered. For the selected SNPs, the linkage disequilibrium was analyzed by PLINK and Haploview software, and the *in silico* prediction of transcription factor binding sites (TFBSs) and microRNAs (miRNAs) binding sites was made using the LASAGNA and RNAhybrid software, respectively. A linear regression model was applied to associate SNPs with the phenotypic measures of FE, and significant SNPs were compared to regions of quantitative trait loci (QTLs) of phenotypes of livestock interest from studies in the same experimental population and from the Cattle QTL database. After functional annotation, two SNPs were selected in the upstream region of the *CTGF* gene; three referring to *COL1A1*, one in a splicing region and two in the upstream; 15 for the *CYP2B6* gene, among them, one missense variant, four in 5' UTR and three in 3' UTR, upstream and downstream; five for the *EGR1* gene, one of these in the 3' UTR region and four in the upstream region; and 33 SNPs referring to the *PRUNE2* gene, three in the upstream region and six in the downstream region, and 24 missense variants. Among these, the SNPs for the five DEGs have the potential to influence TFBSs and have also been integrated into QTL regions of bovine phenotypes. SNPs of the *CYP2B6* and *EGR1* genes have the potential to influence the binding of several bovine miRNAs. Finally, two SNPs located in the *CTGF* gene were associated with the phenotypic measure of gross feed efficiency, one SNP in the *COL1A1* gene was associated with the measures of average daily gain and feed conversion, and 19

SNPs in the *PRUNE2* were associated with residual feed intake, respectively. The SNPs identified increase the knowledge about the possible role regulation of genes involved with phenotypic features of FE.

Keywords: *Bos indicus*. SNP. Gene expression. Transcription factor. MicroRNA. Phenotype.

LISTA DE ILUSTRAÇÕES

- Figura 1** - Gráfico dos efeitos das variantes nas regiões dos genes candidatos anotados pelo *software* VEP. **31**
- Figura 2** - Bloco de haplótipo de SNPs obtido pelo Haploview com o desequilíbrio de ligação em função de r^2 . **32**
- Figura 3** - Esquema de um resultado da predição *in silico* de SNPs que possivelmente afetam sítios de ligação de fatores de transcrição. **36**

LISTA DE TABELAS

- Tabela 1.** Modo de Calcular as Medidas Fenotípicas de Eficiência Alimentar. **15**
- Tabela 2.** Exemplos de impactos causados pelos SNPs de acordo com sua localização do genoma. **21**
- Tabela 3.** Número de SNPs selecionados e respectivos efeitos das variantes nas regiões dos genes candidatos anotados pelo *software* VEP. **30**
- Tabela 4.** SNPs selecionados na janela referente a 2 Kb upstream ao início do sítio de transcrição adicionada à sequência consenso dos genes candidatos que potencialmente afetam sítios de ligação de fatores de transcrição de vertebrados expressos em fígado. **33**
- Tabela 5.** Top 5 das diferenças de energias de hibridização resultante da ligação dos microRNAs bovinos aos seus respectivos sítios-alvo influenciadas pela forma alternativa e referência dos SNPs selecionados na região UTR dos genes candidatos. **37**
- Tabela 6.** SNPs dos genes de interesse com associação significativa com as medidas fenotípicas de eficiência alimentar. **41**
- Tabela 7.** SNPs selecionados nas janelas dos genes candidatos integrados a regiões de QTLs do *Cattle QTL Database* e de resultados de GWAS do grupo de pesquisa Biotecnologia Animal. **44**

LISTA DE SIGLAS

bta	<i>Bos Taurus</i>
CA	Conversão Alimentar
Ca	Elemento Químico Cálcio
CAL	Conteúdo de Ácido Lignocérico
CAR	Consumo Alimentar Residual
CMS	Consumo de Matéria Seca
CZM	Conteúdo de Zinco no Músculo
DL	Desequilíbrio de Ligação
EA	Eficiência Alimentar
EAB	Eficiência Alimentar Bruta
FT	Fator de Transcrição
GDE	Gene Diferencialmente Expresso
GI	Gordura Intramuscular
GMD	Ganho Médio Diário
GWAS	Associação de Todo o Genoma (<i>Genome-Wide Association Study</i>)
K	Elemento Químico Potássio
Kb	Kilo (quilo) Pares de Bases = 1.000 pb
MDSCs	Célula-Satélite Derivada de Músculo (<i>Muscle-Derived Satellite Cell</i>)
miRNA	MicroRNA
mRNA	RNA Mensageiro
pb	Par de Bases
PIB	Produto Interno Bruto
PPi / Pi	Pirofosfato Inorgânico / Pirofosfato Ortofosfato
QTL	<i>Loco</i> de Característica Quantitativa (<i>Quantitative Trait Loci</i>)
r / r ²	Coeficiente de Correlação / Coeficiente de Determinação
SLFT	Sítio de Ligação de Fator de Transcrição
SNP	Polimorfismos de Nucleotídeo Único (<i>Single Nucleotide Polymorphism</i>)
TF	Tolerância ao Frio
TSS	Sítio de Início da Transcrição (<i>Transcription Start Site</i>)
UTR	Região não Traduzida (<i>untranslated region</i>)
VEP	<i>Variant Effect Predictor</i>
β	(Letra Grega Beta) Coeficiente de Regressão

SUMÁRIO

1. Introdução	12
2. Revisão de Literatura	13
2.1 A Raça Nelore	13
2.2 Bovinocultura no Brasil	14
2.3 Eficiência Alimentar	15
2.4 Biologia Molecular no Estudo da Eficiência Alimentar de Bovinos	16
2.5 Genes Candidatos	17
2.5.1 Gene Candidato COL1A1	17
2.5.2 Gene Candidato CTGF	18
2.5.3 Gene Candidato CYP2B6	18
2.5.4 Gene Candidato EGR1	19
2.5.5 Gene Candidato PRUNE2	19
2.6 Marcadores Moleculares	20
2.7 Mecanismos de Regulação Gênica	21
2.7.1 Fatores de Transcrição	22
2.7.2 MicroRNAs	23
3. Justificativa	24
4. Hipótese	24
5. Objetivos	24
5.1 Objetivo Geral	24
5.2 Objetivos Específicos	24
6. Metodologia	25
6.1 Metodologia Realizada Anteriormente pelo Grupo de Pesquisa	25
6.1.1 Animais e Coleta das Amostras	25
6.1.2 Extração de DNA e Dados Genotípicos	26
6.1.3 Seleção de Genes Candidatos	27
6.2 Metodologia Utilizada neste Projeto	27
6.2.1 Seleção de SNPs	27
6.2.2 Análise de Desequilíbrio de Ligação	27
6.2.3 Análise in silico de Mecanismos de Regulação	28
6.2.4 Análise de Integração	29
6.2.5 Medidas Fenotípicas e Análise de Associação	29
7. Resultados e Discussão	30
7.1 Seleção de SNPs e Análise de Desequilíbrio de Ligação	30
7.2 Predição in silico de SNPs que Possivelmente Afetam Sítios Alvo de Fatores de Transcrição	32

7.3 Predição in silico de SNPs que Possivelmente Afetam Sítios de Ligação de miRNAs	
37	
7.4 Análise de Associação entre os SNPs Potencialmente Reguladores da expressão gênica e Medidas Fenotípicas de Eficiência Alimentar	41
7.5 Análise de Integração	43
8. Conclusão	46
9. Referências Bibliográficas	47

1. Introdução

O Brasil é o segundo maior produtor e o maior exportador mundial de carne bovina (USDA, 2018), com um rebanho de aproximadamente 215 milhões de cabeças, compostos predominantemente por animais zebuínos da raça Nelore (*Bos indicus*) e seus cruzamentos (BATISTA *et al.*, 2020; ABIEC, 2020; ACNB, 2020). A bovinocultura de corte brasileira conquistou destaque no cenário nacional e internacional (BATISTA *et al.*, 2020), e ao longo de 2020 representou uma participação de 26,6% no PIB do agronegócio brasileiro.

Porém, devido à expansão de áreas de pastagens sobre as regiões de vegetação nativa e à alta emissão de gases de efeito estufa (SEEG, 2019; BATISTA *et al.*, 2020), avanços tecnológicos são necessários para minimizar os impactos ambientais negativos associados à atividade pecuária. A seleção genômica é uma das tecnologias capazes de melhorar o potencial genético dos animais para redução desses impactos. Da mesma forma, estudos relacionando polimorfismos funcionais (SNPs, do inglês *single nucleotide polymorphism*) de genes candidatos à respectiva expressão gênica e a fenótipos de interesse, como a eficiência alimentar (EA) (CHO *et al.*, 2008; CECCHINATO *et al.*, 2014; ABO-ISMAIL *et al.*, 2013; KARISA *et al.*, 2013) podem fornecer informações importantes sobre a modulação dos fenótipos.

A EA é uma das principais características fenotípicas a ser considerada em programas de melhoramento animal, pois pode aumentar o potencial genético dos animais visando um menor consumo de alimento para manter os melhores níveis possíveis de produção. Animais com boas avaliações de EA aumentam a rentabilidade do produtor, gerando economia com os custos da alimentação dos animais, os quais representam um dos maiores custos no sistema de produção (MOORE *et al.*, 2009; LOPES *et al.*, 2011). Além disso, o melhoramento da EA contribui para a redução do impacto ambiental causado pela pecuária de três formas: por meio da resposta correlacionada na diminuição da emissão de metano e de dióxido de carbono (NKRUMAH *et al.*, 2006; VELAZCO *et al.*, 2016), pela redução na produção de esterco (OKINE *et al.*, 2003) e da não necessidade de novas áreas pastoris (BASARAB *et al.*, 2003; MOORE *et al.*, 2009).

Dentre as diferentes medidas de EA, temos a conversão alimentar (CA), calculada pela razão entre o consumo diário de matéria seca (CMS, kg/dia) e o ganho médio diário (GMD, kg/dia), a eficiência alimentar bruta (EAB), e o consumo

alimentar residual (CAR), a medida mais estudada em bovinos de corte, pois apresenta de moderada a alta herdabilidade (KOCH *et al.*, 1963; ARTHUR *et al.*, 2001; CREWS, 2005), mas que é limitada pelo custo elevado de sua mensuração (FERREIRA *et al.*, 2015).

Apesar do grande interesse econômico e ambiental em melhorar a EA, a seleção de biomarcadores para esse fenótipo não é realizada com frequência, pois tal característica é complexa, ou seja, é influenciada por uma combinação de fatores ambientais e mutações em múltiplos genes.

Desse modo, a proposta do presente projeto é detectar SNPs funcionais com potencial de influenciar genes envolvidos com o fenótipo de EA, com o intuito de contribuir para a compreensão de processos biológicos relacionados a esse fenótipo e para a identificação de alvos para a modulação da sua expressão.

2. Revisão de Literatura

2.1 A Raça Nelore

A raça Zebuína Nelore (*Bos indicus*) é descendente do rebanho indiano Ongole, comum na antiga Província de Madras, atual estado de Andhra Pradesh, localizada na costa oriental da Índia. Na Índia, devido a questões religiosas e culturais, os bovinos são utilizados exclusivamente para transporte de carga e produção de leite (ACNB, 2020; SANTOS, 2017).

Os primeiros animais da raça Nelore chegaram ao Brasil no ano de 1868 em Salvador-BA. A partir de então, a expansão dos rebanhos da raça Nelore foi ocorrendo pelo Brasil, de modo que os animais importados da Índia entre 1920 e 1963 constituem a base do germoplasma brasileiro da raça Nelore (VOZZI *et al.*, 2006). No Brasil, a raça foi alvo de um intenso melhoramento genético, direcionado quase exclusivamente à produção de carne e teve excelente adaptação ao sistema de produção local, de forma que as estimativas indicam que cerca de 80% do rebanho nacional de gado de corte é Nelore ou anelorado (FREITAS, 2013). A promissora criação de animais Nelore no Brasil se deve à boa adaptabilidade ao clima tropical, alta resistência ao calor, capacidade de aproveitamento de alimentos grosseiros e longevidade reprodutiva (FREITAS, 2013). Além disso, a raça Nelore

apresenta resistência natural a parasitas e à insolação, devido ao tipo de pelagem, à pele escura, fina e resistente (ACNB, 2020).

Sendo assim, o gado Nelore possui papel fundamental para a pecuária nacional, devido ao seu patrimônio genético para a bovinocultura de corte brasileira e pelos índices de desempenho econômicos notáveis (COSTA *et al.*, 2007).

2.2 Bovinocultura no Brasil

Devido a importantes transformações que ocorreram nas últimas décadas no âmbito da bovinocultura de corte brasileira, o país conquistou destaque no cenário nacional e internacional da carne bovina (BATISTA *et al.*, 2020). Atualmente o rebanho brasileiro possui aproximadamente 215 milhões de cabeças, sendo o maior rebanho bovino comercial do mundo, composto predominantemente por animais zebuínos da raça Nelore (*Bos indicus*) (BATISTA *et al.*, 2020; ABIEC, 2020; ACNB, 2020). Nas últimas décadas, houve um aumento expressivo na taxa de produção de carne (IBGE, 2019). De modo que, segundo os cálculos do Cepea (Centro de Estudos Avançados em Economia Aplicada), o produto interno bruto (PIB) do agronegócio brasileiro aumentou consideravelmente ao longo de 2020, alcançando participação recorde de 26,6% no PIB brasileiro, comparado a 20,5% em 2019, ou seja, o país totalizou R\$ 7,45 trilhões em 2020, ao qual R\$ 2 trilhões correspondem à atividade do agronegócio. Assim, o Brasil é conhecido como um dos maiores produtores e exportadores de carne bovina do mundo (USDA, 2019; BATISTA *et al.*, 2020).

Em contrapartida, a bovinocultura de corte brasileira tem muito a melhorar, principalmente em relação à intensa expansão de áreas de pastagens sobre as regiões de vegetação nativa, e nos impactos ambientais negativos associados a esse tipo de atividade, como o aumento da emissão de gases causadores do efeito estufa (SEEG, 2019; BATISTA *et al.*, 2020). Tal cenário resulta em grande preocupação, mas por outra perspectiva, revela o grande potencial que a pecuária bovina de corte brasileira possui para manter sua eficiência produtiva e reduzir os impactos negativos, por meio da adoção de ferramentas genômicas (BATISTA *et al.*, 2020), que podem contribuir na identificação de potenciais biomarcadores para auxiliar os programas de melhoramento genético animal para as características de produção e EA.

2.3 Eficiência Alimentar

A EA em bovinos de corte é uma característica que visa mensurar a produção de carne ou de carcaça em função do consumo alimentar do animal, levando em consideração as exigências de manutenção para preservar os mesmos níveis de produção (MOORE *et al.*, 2009; GOMES, 2015). Tal característica é de grande importância, pois causa o aumento do potencial genético que resulta em animais com um menor consumo alimentar mantendo os mesmos níveis de produção.

Desse modo, o melhoramento da EA propicia o aumento da rentabilidade para o produtor, devido à diminuição dos custos de produção, visto que a alimentação representa um dos maiores custos no sistema de produção (MOORE *et al.*, 2009; LOPES *et al.*, 2011). Além de contribuir com a redução de custos, animais mais eficientes emitem cerca de 28% menos metano na atmosfera comparando-se a animais menos eficientes (NKRUMAH *et al.*, 2006; VELAZCO *et al.*, 2016), produzem menos esterco (OKINE *et al.*, 2003), e diminuem a necessidade da adoção de novas áreas pastoris, o que contribui com a competição por áreas produtivas (BASARAB *et al.*, 2003; MOORE *et al.*, 2009).

O modo de avaliar a EA de bovinos de corte consiste em medir o quanto o animal ingere, e simultaneamente, o quanto ele produz em relação à carne ou carcaça (GOMES, 2015). Diversas medidas são utilizadas para avaliar a EA, como: CA, EAB, e CAR. A Tabela 1 mostra como essas medidas fenotípicas são calculadas.

Tabela 1. Modo de Calcular as Medidas Fenotípicas de Eficiência Alimentar.

Medida Fenotípica	Cálculo ¹
Conversão Alimentar (CA, kg MS / kg ganho)	$CMS (kg/dia) / GMD (kg/dia)$
Eficiência Alimentar Bruta (EAB, kg ganho / kg MS)	$GMD (kg/dia) / CMS (kg/dia)$
Consumo Alimentar Residual (CAR, kg MS / dia)	$CMS observado - CMS estimado$

Fonte: Própria (2021).

¹ KOCH *et al.*, 1963; FERNANDES *et al.*, 2014.

CMS: Consumo de Matéria Seca; GMD: Ganho Médio Diário.

As medidas fenotípicas CA, EAB e CAR são correlacionadas com o fenótipo EA, além disso, as medidas CA e EAB também são correlacionadas com

características de crescimento e padrões de maturidade, como ganho de peso e peso à idade adulta, o que pode comprometer a eficiência dos sistemas a pasto devido ao aumento do tamanho dos animais na fase adulta e das necessidades de manutenção (CAMPOS *et al.*, 2012). Porém a medida CAR é independente, o que possibilita a seleção de animais com menor necessidade de ingestão de energia (CAR negativo) mantendo as exigências produtivas (CAMPOS *et al.*, 2012).

Porém, as diversas medidas de EA são onerosas e tardias, o que dificulta a sua implementação em programas de melhoramento genético animal (MOORE *et al.*, 2009). Contudo, as medidas de EA são características de moderada a alta herdabilidade genômica (OLIVEIRA *et al.*, 2014; ARTHUR *et al.*, 2001; TIZIOTO *et al.*, 2015b), ou seja, são passíveis de serem usadas para seleção (ARTHUR *et al.*, 2001).

2.4 Biologia Molecular no Estudo da Eficiência Alimentar de Bovinos

Existem vários estudos que buscaram fornecer informações relacionadas aos processos biológicos de EA em bovinos. Pesquisas de associação de todo o genoma (GWAS, do inglês *Genome-wide association study*) buscam associações entre SNPs e características fenotípicas complexas e identificam *locos* de características quantitativas (QTL, do inglês *Quantitative trait loci*) (LU *et al.*, 2013). Um estudo de GWAS em várias raças bovinas, com o objetivo de identificar variantes associadas com CAR e características como ganho médio diário e consumo de ração, resultou na associação de um SNP a CAR. Esse SNP, está localizado no cromossomo 8 a montante do gene *GFRA2*, e tem função de influenciar as taxas metabólicas basais (HIGGINS *et al.*, 2018). Ainda, outro estudo mapeou QTLs para as características de CA, CAR e CMS em bovinos de corte, identificando um total de 37 diferentes QTL associadas aos fenótipos, dando destaque a regiões localizadas nos cromossomos BTA 1, 7, 18 e 19, as quais são consistentes com estudos anteriores de mapeamento de QTL para CAR (SHERMAN *et al.*, 2009)

Além disso, pelo sequenciamento de RNA (RNA-seq) todos os genes expressos em um determinado tecido são investigados (WANG *et al.*, 2009), metodologia que torna possível a identificação de genes diferencialmente expressos (GDEs) entre grupos amostrais divergentes para características fenotípicas

complexas. Tizioto *et al.* (2015b), buscaram identificar GDEs em animais geneticamente divergentes para CAR em tecidos hepáticos de bovinos Nelore, tendo sido 112 GDEs anotados identificados entre animais com CAR divergentes, como os genes *COL1A1*, *CTGF*, *CYP2B6*, *PRUNE2* e *EGR1*, GDEs com expressão aumentada no grupo de animais ineficientes. Esses genes estão envolvidos no transporte e ligação de íons, atuam como proteínas de membrana ou transmembrana, e estão relacionados ao transporte e catálise de moléculas e aos mecanismos essenciais de absorção de nutrientes, essenciais para os processos envolvidos com a eficiência alimentar.

2.5 Genes Candidatos

A abordagem de genes candidatos consiste em genes que geralmente possuem significado biológico conhecido, que regulam direta ou indiretamente os processos de um fenótipo específico (GIRI e MOHAPATRA, 2017).

Em estudos prévios, realizados pelo grupo de pesquisa em Biotecnologia Animal da Embrapa Pecuária Sudeste em São Carlos - SP, foram identificados GDEs, como os genes *COL1A1*, *CTGF*, *CYP2B6*, *PRUNE2* e *EGR1*, para grupos de animais contrastantes quanto ao fenótipo de CAR em tecido hepático (TIZIOTO *et al.*, 2015).

2.5.1 Gene Candidato *COL1A1*

O gene Colágeno, Tipo I, Alfa 1, do inglês *collagen, type I, alpha 1* (*COL1A1*), localizado no cromossomo 19 do genoma bovino, atua no movimento celular e na organização da matriz extracelular, por meio da formação do colágeno e de fibrilas de ancoragem (FABREGAT *et al.*, 2017; LI *et al.*, 2016; TIZIOTO *et al.*, 2015b). Em estudos prévios, foi demonstrado que o gene *COL1A1* é regulado positivamente no grupo de animais ineficientes para a medida fenotípica de CAR em bovinos Nelore (TIZIOTO *et al.*, 2015), e também foi identificado como DE em fígado de bovinos Angus selecionados para a medida fenotípica de CAR (CHEN *et al.*, 2011).

2.5.2 Gene Candidato *CTGF*

O gene Fator de Crescimento Do Tecido Conjuntivo, do inglês *connective tissue growth factor (CTGF)*, também conhecido como *CCN2*, pertence a família *CCN* e é caracterizado como uma proteína de matriz extracelular rica em cisteína (BOES *et al.*, 1999). Foi denominado como fator de crescimento de tecido conjuntivo por promover a proliferação de fibroblastos, migração, adesão e formação de matriz extracelular como sugerido pela sua nomenclatura (ABREU *et al.*, 2002; LASKY, 2006). Segundo a literatura, a expressão do gene *CTGF* é importante para o desenvolvimento embrionário no período inicial da vida (JUN e LAU, 2011). Lambi *et al.* (2012), em um estudo com camundongos, demonstrou que os animais com inativação do gene *CTGF* apresentaram diversos dimorfismos esqueléticos e letalidade perinatal. Ainda, doenças como fibrose e crescimento tumoral são vinculados à expressão anormal de *CTGF* em adultos (RAMAZANI *et al.*, 2018). Além disso, *CTGF* atua como indutor de apoptose, está envolvido em processos de cicatrização tecidual, fibrogênese, angiogênese, transdiferenciação celular (LASKY, 2006), e é um gene DE para a via de sinalização de IGF-1 (TIZIOTO *et al.*, 2015b). Tizioto *et al.* (2015) mostrou uma regulação positiva do *CTGF* no grupo de animais ineficientes para a medida fenotípica de CAR em bovinos Nelore.

2.5.3 Gene Candidato *CYP2B6*

O gene Membro 6 da Subfamília B da Família 2 do Citocromo P450, do inglês *cytochrome P450, family 2, subfamily B, polypeptide 6 (CYP2B6)*, está localizado no cromossomo 18 do genoma bovino, é membro da família *CYP* que codifica enzimas com função no metabolismo oxidativo de substâncias como esteróides, ácidos graxos, drogas, síntese de colesterol e moléculas exógenas (NELSON *et al.*, 1996), sendo expresso principalmente no fígado (ZANGER; KLEIN, 2013). Este gene é regulado positivamente no grupo de animais ineficientes para CAR (TIZIOTO *et al.*, 2015b), e atua no metabolismo de degradação da melatonina, que segundo Hatzis *et al.* (2013), atenua o estresse oxidativo de forma a diminuir a peroxidação lipídica em resposta a ativação de células produtoras de colágeno no fígado (LEE *et al.*, 1995). Ainda, a melatonina pode aumentar os níveis de várias enzimas

antioxidantes, e as interações entre ela e as espécies reativas de oxigênio são classificadas como eficazes contra o estresse oxidativo, favorecendo a função da cadeia respiratória mitocondrial (SOLÍS-MUÑOZ *et al.*, 2011; HATZIS *et al.*, 2013).

2.5.4 Gene Candidato *EGR1*

O gene Resposta de Crescimento Precoce 1, do inglês *early growth response 1 (EGR1)*, localizado no cromossomo 7 do genoma bovino, pertence à família EGR. O gene *EGR1* foi encontrado positivamente regulado em grupos de animais ineficientes para a medida fenotípica de CAR (TIZIOTO *et al.*, 2015). Estudos anteriores demonstram que o gene *EGR1* é capaz de regular genes relacionados ao crescimento e diferenciação celular, é ativado em resposta ao estresse oxidativo (PAGEL; DEINDL, 2012; SCHIAFFONATI; TIBERIO, 1997), e interage com GDE associados ao estresse oxidativo (TIZIOTO *et al.*, 2016). Desse modo, alterações na expressão gênica entre animais eficientes e ineficientes podem definir condições de estresse oxidativo, indicando *EGR1* como um potencial regulador acerca dessas alterações (TIZIOTO *et al.*, 2016; TIZIOTO *et al.*, 2015b). Ademais, o gene *EGR1* possivelmente promove a diferenciação de células-satélite derivadas de músculo esquelético bovino (MDSCs), de forma a observar o aumento da expressão de mRNA e proteína de *EGR1* durante a diferenciação de MDSCs, promovendo a diferenciação e inibindo a proliferação, e o aumento da expressão de mRNA e da proteína de *MyoG*, fator de transcrição (FT) específico de tecido muscular envolvido no desenvolvimento e reparação do músculo esquelético, ou seja, *EGR1* pode promover a diferenciação de MDSC por meio da regulação positiva da expressão do gene *MyoG* (ZHANG *et al.*, 2017).

2.5.5 Gene Candidato *PRUNE2*

O gene Homólogo de Prune 2 (*Drosophila*), do inglês *prune homolog 2 (PRUNE2)*, localizado no cromossomo 8 do genoma bovino, atua como uma enzima pirofosfatase, ao qual catalisa a hidrólise do pirofosfato inorgânico (PPi) a ortofosfato (Pi), favorecendo termodinamicamente diversas reações de biossíntese (ZERBINO *et al.*, 2018; KORNBERG, 1962). TIZIOTO *et al.*, 2015 mostrou um aumento na expressão gênica do gene *PRUNE2* no grupo de animais ineficientes para a medida

fenotípica de CAR. Além disso, este gene está envolvido na regulação do processo apoptótico (LI *et al.*, 2011; DU *et al.*, 2017). O processo apoptótico foi relacionado ao fenótipo EA em bovinos *Bos taurus* (KHANSEFID *et al.*, 2017), em tecido hepático, podendo ser encadeado pelo estresse oxidativo (DU *et al.*, 2017), que segundo Tizioto *et al.* (2015b), está aumentado em animais Nelore ineficientes.

2.6 Marcadores Moleculares

Os marcadores moleculares são uma variação específica de DNA entre indivíduos, como inserções, deleções, translocações, duplicações e mutações pontuais, fisicamente ligados a *locos* que determinam características de interesse (SINGH *et al.*, 2014; ALZATE-MARIN *et al.*, 2005), ou a própria variação causal. Esses apresentam como características desejáveis o número ilimitado no genoma, a fácil detecção, herança simples e previsível, que não é afetada pelo ambiente (ALZATE-MARIN *et al.*, 2005). Dentre suas principais aplicações nos programas de melhoramento está a determinação de parentesco para obtenção de informações sobre *pedigree*, e a seleção de animais resistentes a doenças e superiores para características de interesse econômico (SINGH *et al.*, 2014).

Um exemplo dessa abordagem é a utilização de marcadores moleculares do tipo SNP, definido como um polimorfismo causado por uma mutação pontual que origina diferentes alelos contendo bases alternativas em uma determinada posição do nucleotídeo dentro de um *loco* (JIN *et al.*, 2016).

O impacto da mutação pontual causada pelos SNPs depende da sua localização no genoma (Tabela 3), assim, os SNPs localizados em regiões codificantes do DNA, podem ser definidos como variantes sinônimas, quando a alteração dos alelos do SNP resulta na codificação do aminoácido esperado (SHASTRY, 2009), ou não sinônimas, quando os SNPs modificam sequências de nucleotídeos, resultando na codificação de um aminoácido diferente do esperado (MCLAREN *et al.*, 2010), como é o caso das variantes *missenses*, que alteram o aminoácido e podem influenciar a estabilidade e função da proteína resultante, e das variantes *nonsenses*, que introduzem códons de parada nas sequências de codificação, causando a terminação prematura da síntese da proteína correspondente (STEFEL *et al.*, 2013; CHU; WEI, 2019).

Ainda, os SNPs podem ser localizados em regiões não codificantes, o que pode afetar a estrutura do mRNA e a expressão gênica (LU *et al.*, 2015), por meio de alterações nos locais de ligação dos FTs (WUNSCH *et al.*, 2005; GUO; JAMISON, 2005), alterações em regiões de *splicing* que afetam o processo de maturação do pré-mRNA (JIAN, BOERWINKLE e LIU, 2014), ou funções importantes na regulação pós-transcricional da expressão gênica, relacionadas a estabilidade, transporte de mRNAs, efetividade da tradução e localização subcelular (MIGNONE *et al.*, 2002; JANSEN, 2001; VELDEN e THOMAS, 1999; BASHIRULLAH, COOPERSTOCK e LIPSHITZ, 2001).

Tabela 2. Exemplos de impactos causados pelos SNPs de acordo com sua localização do genoma.

		Região Codificante			Região Não Codificante
		Sinônima	Nonsense	Missense	
DNA	TTC	TTT	ATC	TCC	- Alteram sítios de ligação de fatores de transcrição e de microRNAs;
mRNA	AAG	AAA	UAG	AGG	
Aminoácido	Lys	Lys	STOP	Arg	- Alteram regiões de <i>Splicing</i> .

Fonte: Própria (2021).

2.7 Mecanismos de Regulação Gênica

Os mecanismos de regulação gênica podem ocorrer em diferentes níveis, desde a transcrição ao processamento da proteína (RAMÍREZ-CLAVIJO; MONTOYA-ORTÍZ, 2013). Contudo, a presença de SNPs nas sequências gênicas pode interferir nestes mecanismos, principalmente sobre os processos que antecedem a transcrição ou em processos pós-transcricionais. Os mecanismos pré-transcricionais baseiam-se na regulação pelos FTs, os quais possuem um papel importante na ativação da maquinaria responsável na região promotora dos genes (MOBLEY, 2019; LUSE, 2013). Enquanto que os mecanismos pós-transcricionais podem ser mediados pelos miRNAs (O'BRIEN *et al.*, 2018).

2.7.1 Fatores de Transcrição

Os FTs são um dos principais elementos reguladores no início da transcrição do gene (SPITZ; FURLONG, 2012). Esses são proteínas que se ligam a sequências reguladoras de DNA, que geralmente estão localizadas na região 5' UTR a montante da sequência dos genes alvo, com a função de modular a taxa de transcrição do gene, podendo aumentá-la ou diminuí-la, afetando a taxa de síntese protéica e, conseqüentemente, levando a alterações na função celular (ADCOCK; CARAMORI, 2009).

O processo de ativação do FT e subsequente ligação a regiões reguladoras, como *enhancers* e promotores, requer o envolvimento de vias de transdução de sinais intracelulares, diferentes sítios de ligação para múltiplos FTs e do posicionamento do nucleossomo (ADCOCK; CARAMORI, 2009; SPITZ; FURLONG, 2012; YAO, 2017).

Ao se ligar na sequência promotora, o FT convoca elementos ativadores de transcrição e proteínas mediadoras com o auxílio da proteína de dobra de DNA, que posiciona os ativadores próximo ao promotor, esses recrutam a RNA polimerase, que se liga aos FTs e ao DNA do promotor do gene, de forma a montar o complexo de pré-iniciação da transcrição (MOBLEY, 2019; LUSE, 2013). Desse modo, é formada a bolha de transcrição, que consiste na dupla hélice de DNA desenrolada próxima ao gene que está sendo transcrito, e o complexo promotor aberto, em que a fita molde de DNA entra em contato com o sítio ativo da RNA polimerase e o produto de RNA complementar à fita molde de DNA é sintetizado (SAINSBURY; BERNECKY; CRAMER, 2015). Se o produto da transcrição codifica uma proteína, a molécula de RNA será traduzido em proteína no processo de tradução.

A presença de SNPs nas regiões *upstream* ao sítio de início da transcrição (TSS) pode influenciar a afinidade de ligação do FT, devido a criação ou modificação dos sítios de ligação do fator de transcrição (SLFTs), e conseqüentemente a expressão gênica. Tal influência tem sido associada a diversos fenótipos de interesse, em doenças humanas como o câncer e diabetes tipo 2, e em bovinos, à EA (NISHIZAKI *et al.*, 2019; FOGARTY *et al.*, 2014; POMERANTZ *et al.*, 2009; STITZEL *et al.*, 2010; ALEXANDRE *et al.*, 2014).

2.7.2 MicroRNAs

Os miRNAs são definidos como subconjuntos de RNAs não codificantes, constituídos por sequências curtas compostas por aproximadamente 22 nucleotídeos de comprimento (O'BRIEN *et al.*, 2018; CAI *et al.*, 2009). De modo geral, a biogênese dos miRNAs resume-se na transcrição de sequências de DNA em miRNAs primários (pri-miRNAs), que são processados em miRNAs precursores (pré-miRNAs) e, por fim, em miRNAs maduros (O'BRIEN *et al.*, 2018). Os miRNAs possuem papel de reguladores pós-transcricionais, por meio da clivagem dos mRNAs alvo ou repressão da tradução dos mesmos (O'BRIEN *et al.*, 2018).

Os primeiros estudos que identificaram a função dos miRNAs foram a partir da descoberta inicial de dois miRNAs, let-7 e lin-4, em um estudo com nematóides (*Caenorhabditis elegans*), em que demonstraram a capacidade de ligação do miRNA a sítios específicos localizados na região 3' UTR do mRNA alvo, agindo como um repressor pós-transcricional (LEE *et al.*, 2002; REINHART *et al.*, 2000). Depois disso, diversos miRNAs foram identificados por análises *in silico*, *in vitro* e *in vivo*, relacionados a tempo de desenvolvimento, diferenciação celular, proliferação, apoptose, tumorigênese e interações hospedeiro-patógeno (TSUCHIYA *et al.*, 2006; SCARIA *et al.*, 2006; LIU *et al.*, 2008).

Relata-se que os miRNAs interagem com as regiões 3' e 5' UTR dos mRNAs alvo, sendo que a especificidade e atividade em alvos de ligação estão relacionadas à homologia das sequências *seed*, que regem a interação entre miRNA e mRNA, para exercer a função de suprimir a expressão (HA; KIM, 2014; BROUGHTON *et al.*, 2016, CAI *et al.*, 2009). No que diz respeito à interação de miRNAs com seus alvos, um único miRNA é capaz de controlar a expressão de mais de um mRNA alvo e cada mRNA pode ser regulado por múltiplos miRNAs (CAI *et al.*, 2009).

Em relação ao impacto provido por SNPs, quer provenientes de mudanças nas sequências codificadoras de miRNA quer em um sítio de ligação de miRNA nas regiões 3' e 5' UTR das sequências de mRNAs, essas variações são capazes de afetar a biogênese e a função do miRNA, fortalecendo ou reduzindo, ou até eliminando, os efeitos de regulação dos mRNAs (CAI *et al.*, 2009).

3. Justificativa

Como a mensuração do fenótipo EA é onerosa e tardia, a identificação de potenciais biomarcadores em regiões reguladoras do genoma bovino que podem influenciar a expressão de genes envolvidos com a EA, como os candidatos *COL1A1*, *CTGF*, *CYP2B6*, *EGR1* e *PRUNE2*, pode auxiliar os programas de melhoramento animal, tanto na priorização de marcadores a serem incluídos nos painéis de SNPs quanto na elaboração de modelos de edição gênica.

4. Hipótese

Variações de sequência de DNA modificam regiões reguladoras e podem afetar medidas fenotípicas de Eficiência Alimentar, conversão alimentar, eficiência alimentar bruta e consumo alimentar residual, em bovinos Nelore.

5. Objetivos

5.1 Objetivo Geral

- Identificar biomarcadores de DNA nos genes *CTGF*, *COL1A1*, *CYP2B6*, *EGR1* e *PRUNE2* que possuam um potencial impacto regulatório sobre sítios de ligação de fatores de transcrição e de miRNAs, associados a medidas fenotípicas de eficiência alimentar em bovinos Nelore.

5.2 Objetivos Específicos

- Selecionar SNPs que possivelmente alteram regiões reguladoras de genes candidatos;
- Analisar *in silico* possíveis mecanismos reguladores que possam ser influenciados por SNPs e afetar a expressão gênica, como SLFTs e miRNAs;
- Associar e integrar dados obtidos na mesma população aos SNPs com o intuito de relacioná-los a fenótipos de interesse agropecuário.

6. Metodologia

6.1 Metodologia Realizada Anteriormente pelo Grupo de Pesquisa

Os itens 6.1.1 a 6.1.3 correspondem à metodologia desenvolvida pelo grupo de pesquisa da Embrapa Pecuária Sudeste, ao qual resultou nos dados para a execução do presente trabalho.

6.1.1 Animais e Coleta das Amostras

Os procedimentos experimentais foram conduzidos de acordo com as Diretrizes do Comitê Institucional de Uso e Cuidado Institucional de Animais da Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA) - CEUA, que aprovou todos os protocolos experimentais (Processo CEUA 01/2013).

Para a seleção dos novilhos que produziram a nossa população por meio de inseminação artificial de matrizes comerciais e puras da raça Nelore, foi utilizado o sêmen de touros que representavam as principais linhagens de criação comercializadas no Brasil. Neste estudo, 398 animais foram utilizados, esses animais cresceram até os 21 meses de idade na fazenda da Embrapa Pecuária Sudeste (São Carlos, São Paulo, Brazil), sendo depois confinados em baias individuais ou coletivas, com período de adaptação de 28 dias seguidos de confinamento por 70 dias (TIZIOTO *et al.*, 2015; NASCIMENTO *et al.*, 2016; LIMA, 2019; Andressa Lima, comunicação pessoal). Os novilhos foram alimentados duas vezes ao dia com 40% de matéria seca e 60% de concentrado, conforme descrito por Oliveira *et al.* (2014).

Todas as medidas fenotípicas e avaliações referentes a características relacionadas à EA coletadas durante o confinamento foram descritas em Oliveira *et al.* (2014) e Nascimento *et al.* (2016). Dentre as medidas fenotípicas, selecionamos ganho médio diário (GMD, kg/d), consumo de matéria seca (CMS, kg/d), taxa de conversão alimentar (CA, CMS/GMD; kg / kg), razão de eficiência alimentar bruta (EAB, GMD/CMS; kg/kg) e consumo alimentar residual (CAR, kg/d).

Os animais foram abatidos em um abatedouro comercial, de acordo com as diretrizes para o Abate Humanitário de Bovinos. No decorrer do abate, as carcaças

foram devidamente identificadas, e amostras de fígado foram coletadas, e congeladas em nitrogênio líquido e mantidas a - 80 °C.

6.1.2 Extração de DNA e Dados Genotípicos

Amostras de sêmen congelado obtidas em centros de inseminação artificial brasileiros foram usadas para extrair DNA por um método padrão de fenol-clorofórmio conforme descrito detalhadamente por Tizioto *et al.* (2012). Os dados da sequência do genoma completo de 26 progenitores da população foram obtidos com o sistema *Illumina HiSeq 2500* (Illumina Inc., San Diego, CA, EUA) com 8-21x de cobertura. Alinhamento, chamada de variantes e controles de qualidade foram realizados de acordo com as recomendações do 1000 *Bulls Genomes Project* (<http://www.1000bullgenomes.com/>) e usados como dados de referência para imputação de SNP.

Enquanto que para a progênie, 5 mL de amostras de sangue foram coletadas e utilizadas para extração de DNA, conforme descrito em Tizioto *et al.* (2012). Todos animais foram genotipados usando o *Illumina BovineHD BeadChip 770k* (Illumina, San Diego, CA, EUA). Os genótipos foram chamados no *software* *Illumina GenomeStudio* e foram descritos em detalhes em estudos anteriores (TIZIOTO *et al.*, 2012; TIZIOTO *et al.*, 2013) e as informações genotípicas foram enriquecidos por meio de imputação conforme descrito em detalhes por Cardoso *et al.* (2021).

A imputação de genótipos consiste no processo de predição de genótipos para *locos* marcadores, como os SNPs, ausentes ou desconhecidos, ou seja, que não foram genotipados em uma população de estudo. Essa é realizada com o intuito de aumentar a quantidade de marcadores que serão utilizados nas análises (MARCHINI; HOWIE, 2010).

Após a imputação, SNPs não informativos, sexuais e com uma frequência de alelo menor inferior a 0,01 foram removidos do conjunto de dados usando o *software* *PLINK* (PURCELL *et al.*, 2007). Após essas etapas de filtragem, um total de 4.813.664 SNPs estava disponível para as análises genéticas (CARDOSO *et al.*, 2021).

6.1.3 Seleção de Genes Candidatos

A seleção dos genes candidatos foi feita a partir da análise dos resultados de expressão gênica diferencial para grupos geneticamente divergentes para a medida fenotípica de CAR, conforme descrito em detalhes por Tizioto *et al.* (2015b). Os GDEs, *COL1A1*, *CTGF*, *CYP2B6*, *EGR1*, e *PRUNE2* foram selecionados por participarem em processos biológicos relacionados à característica EA (TIZIOTO *et al.*, 2015).

6.2 Metodologia Utilizada neste Projeto

6.2.1 Seleção de SNPs

A partir da imputação dos SNPs, descritas por Cardoso *et al.* (2021), para a identificação dos elementos regulatórios nos genes, adotamos uma janela de prospecção correspondente a 2 Kb *upstream* aos TSS de cada um dos genes de interesse, até a região de 3' UTR dos mesmos, usando como base as coordenadas do genoma de *Bos taurus* (ARS-UCD1.2) (ROSEN *et al.*, 2020) disponíveis no banco de dados Ensembl (YATES *et al.*, 2019). A janela de prospecção de SNPs foi escolhida para propiciar a localização de elementos *cis*-reguladores nas regiões promotoras e no interior dos genes de interesse.

O *software* VCFtools (DANECEK *et al.*, 2011) foi utilizado para selecionar os SNPs contidos na janela delimitada para os genes candidatos. Posteriormente, os SNPs selecionados foram anotados utilizando o *software* Variant Effect Predictor (VEP) (MCLAREN *et al.*, 2016), que proporcionou a obtenção de informações sobre a localização e/ou classificação correspondente aos SNPs, e com isso excluir os SNPs localizados em regiões intrônicas que não foram apontados como possíveis reguladores de *splicing* ou foram considerados variantes sinônimas, que não afetam composição de aminoácidos do produto gênico.

6.2.2 Análise de Desequilíbrio de Ligação

Utilizamos o *software* PLINK (PURCELL *et al.*, 2007) para as análises de desequilíbrio de Ligação (DL) entre os SNPs detectados, e o *software* Haploview

(BARRETT *et al.*, 2004) para montagem dos blocos de haplótipos de SNPs. Para as análises de DL, foi calculado o coeficiente de determinação r^2 , sendo considerado DL total quando pares de variantes apresentaram r^2 igual a 1 (ARAÚJO, 2014), e $r^2 \geq 0,80$ como indicativo de DL relevante (KONSTA *et al.*, 2015). Tal análise foi realizada para complementar o entendimento dos resultados (TAKEUCHI *et al.*, 2005). Desse modo, o DL refere-se a uma associação não aleatória de alelos em dois ou mais *loci*, e a importância em ser estimado está no fato de que dois *loci* em DL podem apresentar respostas correlacionadas em sua ação sobre a expressão gênica ou na associação com uma determinada característica fenotípica, de tal forma que a resposta à seleção em um local pode influenciar positiva ou negativamente o outro local (SLATKIN, 2008).

6.2.3 Análise *in silico* de Mecanismos de Regulação

Após a seleção de SNPs, foi feita a análise *in silico* dos SNPs em potenciais regiões reguladoras, para isso adotamos regiões flangeadoras aos SNPs, contendo 25 pb antes e depois de cada SNP, por meio de um *script* da linguagem de programação Python, empregando como base a sequência referência de *Bos taurus* - ARS-UCD1.2 (YATES *et al.*, 2019; ROSEN *et al.*, 2020).

Utilizamos o *software* LASAGNA (LEE *et al.*, 2014) para prever *in silico* os SNPs selecionados localizados em SLFTs, possibilitando a identificação de alelos, que possivelmente afetam a afinidade de ligação de FT descritos para vertebrados, disponíveis no banco de dados Jaspar (FORNES *et al.*, 2019), empregando os parâmetros padrões do programa. A interferência na afinidade do FT ao seu SLFT pode alterar a atividade do FT sobre a expressão gênica, inibindo-a ou privilegiando-a (NISHIZAKI *et al.*, 2019).

A predição *in silico* dos possíveis sítios de ligação de miRNAs bovinos afetados por SNPs selecionados nas regiões 5' e 3' UTR dos genes candidatos foi realizada pelo *software* RNAhybrid (REHMSMEIER *et al.*, 2004). Para isso, as sequências FASTA de todos os miRNAs bovinos foram recuperadas do banco de dados miRBase (KOZOMARA; BIRGAOANU; GRIFFITHS-JONES, 2018). A ferramenta RNAhybrid é responsável por calcular a energia da hibridização mais favorável entre as sequências do miRNA e a sua respectiva sequência alvo (mRNA). Foram selecionados miRNAs possivelmente influenciados pelos SNPs observando a

variação da energia livre de hibridização na presença da forma referência e alternativa nos respectivos sítios-alvo. Definimos como parâmetro de corte a energia livre de hibridização de -18 kcal/ mol.

6.2.4 Análise de Integração

Foram recuperadas regiões de QTL de fenótipos importantes para a pecuária bovina disponíveis em dois conjuntos de dados diferentes: o banco de dados do genoma bovino *Cattle QTL database* (versão do genoma de referência: UMD_3.1) (<https://www.animalgenome.org/QTLdb/cattle/> baixado em junho de 2019), e também os resultados de GWAS produzidos previamente na mesma população experimental para fenótipos de qualidade de carne (TIZIOTO *et al.*, 2013), conteúdo de minerais (TIZIOTO *et al.*, 2015a), EA (OLIVEIRA *et al.*, 2014), conteúdo e deposição de gordura (CESAR *et al.*, 2014) e também comparamos os SNPs com eQTLs relatados na mesma população no trabalho de CESAR *et al.*, (2018). Todos os QTLs obtidos na versão anterior do genoma de referência bovino foram convertidas para a versão mais atualizada (ARS-UCD1.2) (ROSEN *et al.*, 2020) utilizando a ferramenta Lift Genome Annotations (*UCSC Genome Browser*, disponível em: <https://genome.ucsc.edu/cgi-bin/hgLiftOver>). Assim foi realizada a sobreposição das posições genômicas dos QTLs obtidos desses conjuntos de dados e dos SNPs selecionados neste trabalho a fim de se observar evidências de potencial funcional dessas variantes em relação aos fenótipos de EA.

6.2.5 Medidas Fenotípicas e Análise de Associação

Utilizamos um modelo de regressão linear por meio do *software* RStudio (R Core Team, 2018) para executar as análises de associação, com o objetivo de avaliar a possível influência dos SNPs selecionados sobre as cinco medidas fenotípicas de eficiência alimentar (GPD, CMS, CA, EAB e CAR). Foram incluídos os efeitos fixos de grupo de contemporâneos (contendo a informação de local de confinamento, ano de nascimento, origem e baia) para as medidas fenotípicas GPD, CMS, CA e EA. Além disso, a idade do animal na data do abate foi incluída como covariável para essas medidas escolhidas. A medida fenotípica CAR foi corrigida anteriormente para os efeitos de contemporaneidade necessários (OLIVEIRA *et al.*,

2014). O parâmetro valor de valor $P < 0.05$ foi utilizado para identificar os SNPs com associação significativa dentro de cada característica de interesse.

7. Resultados e Discussão

7.1 Seleção de SNPs e Análise de Desequilíbrio de Ligação

O número total e a anotação funcional de SNPs selecionados na sequência correspondente a janela de 2 Kb *upstream* ao sítio de início da transcrição, adicionada à sequência consenso de cada um dos genes candidatos, estão apresentados na Tabela 3 e na Figura 1.

Tabela 3. Número de SNPs selecionados e respectivos efeitos das variantes nas regiões dos genes candidatos anotados pelo *software* VEP.

	<i>CTGF</i>	<i>COL1A1</i>	<i>CYP2B6</i>	<i>EGR1</i>	<i>PRUNE2</i>
Nº de SNPs submetidos à anotação *	2	14	40	7	283
<i>Upstream</i>	2	2	3	4	3
5' UTR	-	-	4	-	-
<i>Missense</i>	-	-	1	-	24
Região de Splicing	-	1	1	-	-
3' UTR	-	-	3	1	-
<i>Downstream</i>	-	-	3	-	6

Fonte: Própria (2021).

* Total de SNPs selecionados na seleção, anterior à análise de anotação.

Foram encontrados 14 SNPs localizados na região *upstream* nas janelas dos genes *COL1A1*, *CTGF*, *CYP2B6*, *EGR1* e *PRUNE2* (Tabela1). Essas variantes podem influenciar a transcrição dos genes, por meio de alterações nos locais de ligação dos FTs (WUNSCH *et al.*, 2005; GUO; JAMISON, 2005).

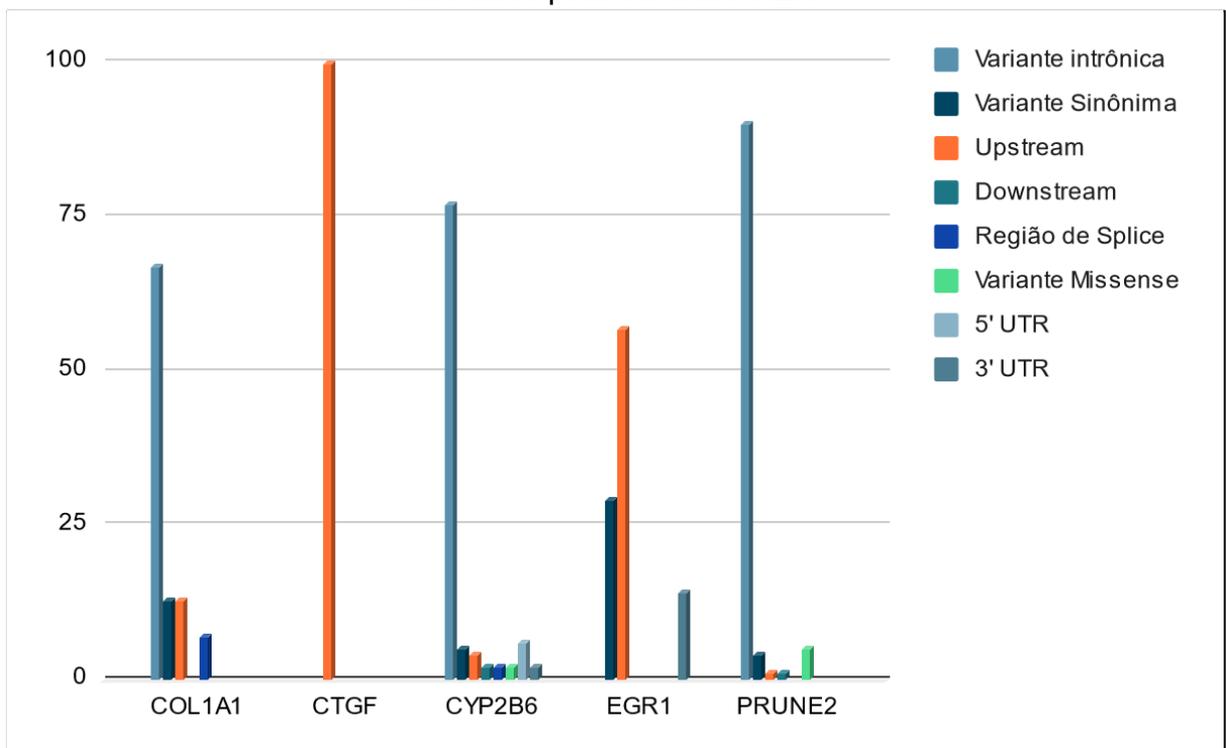
Um total de oito SNPs foram localizados na região UTR dos genes *CYP2B6* e *EGR1* (Tabela 3). Estes SNPs podem desempenhar funções importantes na regulação pós-transcricional da expressão gênica, relacionadas a estabilidade,

transporte de mRNAs, efetividade da tradução e localização subcelular (MIGNONE *et al.*, 2002; JANSEN, 2001; VELDEN e THOMAS, 1999; BASHIRULLAH, COOPERSTOCK; LIPSHITZ, 2001).

Os 25 SNPs correspondentes aos genes *CYP2B6* e *PRUNE2*, anotados como variantes *missenses*, podem modificar a sequência de nucleotídeos, o que resulta na codificação de um aminoácido diferente do esperado (MCLAREN *et al.*, 2010). Tal modificação pode influenciar na estabilidade e função da proteína (STEFL *et al.*, 2013).

Nas janelas de seleção referentes aos genes *COL1A1* e *CYP2B6*, foram selecionados dois SNPs em região de *splicing* que podem influenciar na formação do mRNA por meio do processo de maturação do pré-mRNA (JIAN, BOERWINKLE e LIU, 2014).

Figura 1. Gráfico dos efeitos das variantes nas regiões dos genes candidatos anotados pelo *software* VEP.



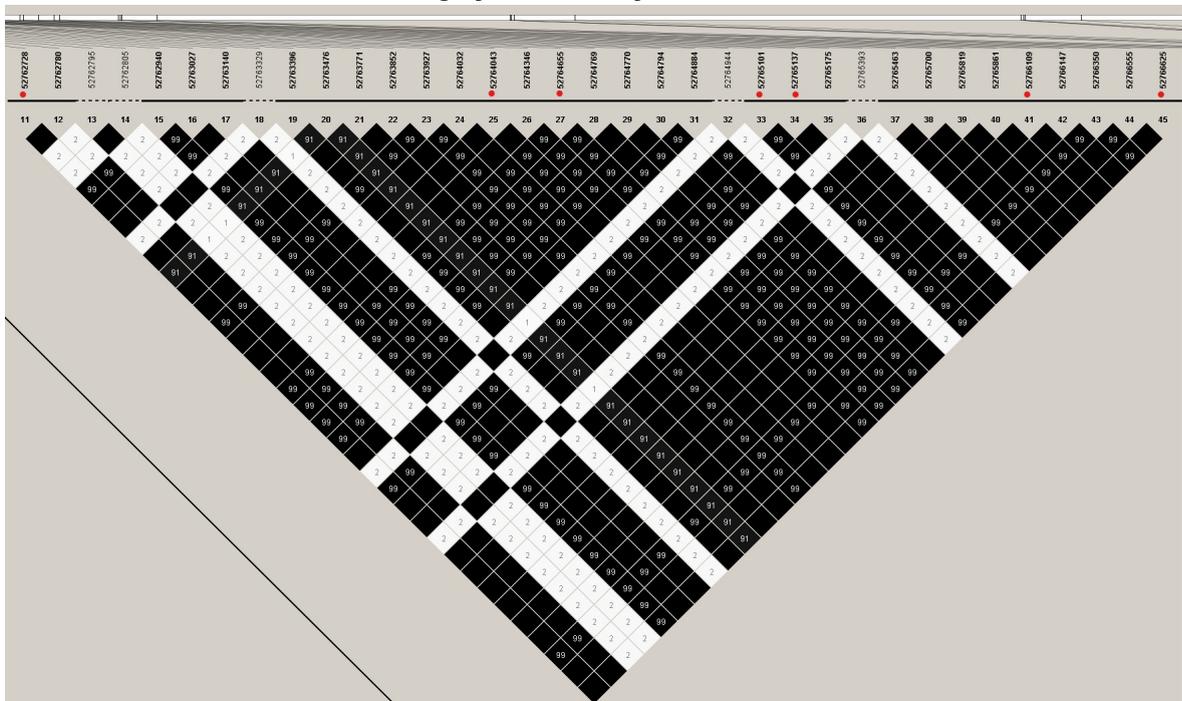
Fonte: Própria (2021).

As análises de DL entre os SNPs selecionados em cada gene foram feitas utilizando o *software* PLINK (PURCELL *et al.*, 2007) e o Haploview (BARRETT *et al.*, 2004) foi utilizado na montagem dos blocos de SNPs. Observamos os valores do coeficiente de determinação r^2 , considerando $r^2 \geq 0,80$ como indicativo de DL

relevante (KONSTA *et al.*, 2015) e $r^2 = 1$ como completo DL (ARAÚJO, 2014). Como resultado, as análises mostraram alto DL dentro de cada gene de interesse.

Tivemos como resultado apenas um bloco de SNPs constituído com o total de SNPs selecionados do gene *PRUNE2*, entre os quais, após a análise de anotação, destacamos sete SNPs anotados como variantes *missenses* do gene *PRUNE2*: g8:52762728G>A, g8:52764043C>T, g8:52764655C>T, g8:52765101A>G, g8:52765137T>A, g8:52766109G>A e g8:52766625T>C (Figura 2). Esses SNPs afetam SLFTs (Tabela 4) e possuem associação significativa com a medida fenotípica CAR (Tabela 6).

Figura 2. Bloco de haplótipo de SNPs obtido pelo Haploview com o desequilíbrio de ligação em função de r^2 .



Fonte: Haploview (BARRETT *et al.*, 2004)

Células em preto representam o completo DL ($r^2=1$), células em preto com escrita em branco representam DL relevante ($r^2 \geq 0,80$). Pontos em vermelho indicam os SNPs candidatos selecionados na região reguladora do gene *PRUNE2*.

7.2 Predição *in silico* de SNPs que Possivelmente Afetam Sítios Alvo de Fatores de Transcrição

A predição *in silico* dos SLFTs possibilitou a identificação de SNPs que potencialmente afetam a específica interação dos FTs aos seus respectivos sítios-alvo. Os sítios de ligação são constituídos por curtas sequências de DNA, que

de modo geral são localizados em regiões *upstream*, promotora, *downstream*, ou codificadora (éxon) em relação ao gene alvo, esses interagem com a maquinaria transcricional mediando o processo de expressão gênica, podendo estimulá-lo ou inibi-lo (STERGACHIS *et al.*, 2013; FERRÃO NETO, 2017; MITSIS *et al.*, 2020).

Na Tabela 4, a coluna "Sequência", mostra como foi interpretada a potencial influência de um SNP sob a afinidade de ligação de um FT. Dependendo da forma referência ou alternativa do alelo, há a ligação ou não de um FT ao correspondente SLFT. Nesse cenário, foram preditos um SNP referente ao gene *CTGF* e um ao gene *COL1A1*, dois SNPs para o gene *CYP2B6*, três SNPs relacionados ao gene *EGR1* e 13 SNPs referentes ao gene *PRUNE2* capazes de afetar SLFTs. Para esses SNPs, a presença do alelo alternativo foi predita a comprometer a afinidade de FTs aos seus sítios-alvos (Tabela 4). Por outro lado, dois SNPs selecionados para o gene *COL1A1*, dois para o gene *EGR1*, quatro para o gene *CYP2B6* e sete SNPs relacionados ao gene *PRUNE2*, potencialmente afetam SLFTs, nos quais na presença do alelo alternativo pode aumentar a afinidade de FTs aos seus sítios-alvo (Tabela 4).

Tabela 4. SNPs selecionados na janela referente a 2 Kb *upstream* ao início do sítio de transcrição adicionada à sequência consenso dos genes candidatos que potencialmente afetam sítios de ligação de fatores de transcrição de vertebrados expressos em fígado.

Gene	Posição	Alelos	FT	Sequência
Informações ^a			LASAGNA ^b	
<i>CTGF</i>	9:69891302	G/T	<i>Gfi1</i>	AGAA t CCCTG
			<i>Mafb</i>	Gc T GATTC
<i>COL1A1</i>	19:36456642	C/G	<i>Foxa2</i>	CAACG c TGATTC
			<i>REL</i>	CG g TGATTCC
<i>CYP2B6</i>	18:50290773	C/T	<i>REST</i>	TTGAGCA c AAGGGA
			<i>RELA</i>	GTGAATCT Cc
	18:50291018	C/G	<i>NFKB1</i>	GTGAATCT Cc C
			<i>ESR1</i>	CCAGGTCACCTCAGACT A
18:50291051	C/T	<i>ESR2</i>	CCAGGTCACCTCAGACT A	

Fonte: Própria (2021).^a Informações básicas sobre os SNPs selecionados: gene, posição no genoma e os alelos em sua forma referência / alternativa (Alelos); ^b Resultados LASAGNA: Possíveis FTs e as referentes sequências dos sítios-alvo criados pelos SNPs identificados em letra minúscula em negrito.

Continuação da Tabela 4.:

Tabela 4. SNPs selecionados na janela referente a 2 Kb *upstream* ao início do sítio de transcrição adicionada à sequência consenso dos genes candidatos que potencialmente afetam sítios de ligação de fatores de transcrição de vertebrados expressos em fígado.

Gene	Posição	Alelos	FT	Sequência
Informações ^a			LASAGNA ^b	
<i>CYP2B6</i>	18:50301918	C/T	<i>HNF4A</i>	TGGCCGAAGT c G
<i>EGR1</i>	7:49824536	T/C	<i>CTCF</i>	t CCCCACTAGGT
			<i>REST</i>	TTGAGCA c AAGGGA
	7:49825415	A/G	<i>HNF4A</i>	ATCCCAAGG g CCA
			<i>PPARG</i>	AAG Gg CCACTAGCTCCTAGT
	7:49829451	A/G	<i>FOXC1</i>	GCCCTGT a
	8:52762269	A/G	<i>REST</i>	GAGGTGTGTAAGGT g CCTAA
	8:52762340	C/G	<i>Prrx2</i>	c ATTA
	8:52762708	C/G	<i>Myb</i>	CA c AGTTG
<i>PRUNE2</i>	8:52762728	G/A	<i>FOXA1</i>	TT a CACAA
			<i>HLF</i>	AGTT a CACAAAT
	8:52762795	T/C	<i>FOXL1</i>	GTG c CATA
			<i>SRF</i>	TG c CATATCTGG
	8:52763329	C/T	<i>Pax2</i>	AGT c ACTC
	8:52763476	C/T	<i>HLF</i>	TG t TACACCAGA
	8:52764032	C/T	<i>FOXL1</i>	t TGCCATA
	8:52764043	C/T	<i>CREB1</i>	TG a cATCT
8:52764655	C/T	<i>NFE2L1:: MafG</i>	c ATGAC	
8:52764769	A/C	<i>Mafb</i>	G c TGAAGA	
8:52765101	A/G	<i>RXRA:: VDR</i>	ATTTCA a TAGTTTCA	
8:52765137	T/A	<i>REST</i>	G a GCTCTTCACAGGTCTCT	

Fonte: Própria (2021).^a Informações básicas sobre os SNPs selecionados: gene, posição no genoma e os alelos em sua forma referência / alternativa (Alelos); ^b Resultados LASAGNA: Possíveis FTs e as referentes sequências dos sítios-alvo criados pelos SNPs identificados em letra minúscula em negrito.

Continuação da Tabela 4.:

Tabela 4. SNPs selecionados na janela referente a 2 Kb *upstream* ao início do sítio de transcrição adicionada à sequência consenso dos genes candidatos que potencialmente afetam sítios de ligação de fatores de transcrição de vertebrados expressos em fígado.

Gene	Posição	Alelos	FT	Sequência
	Informações^a		LASAGNA^b	
	8:52765819	T/C	YY1	GCCAtG
	8:52766109	G/A	BRCA1	TCAaCAC
PRUNE2			SPI1	CTGAAc
	8:52766625	T/C	NR2F1	GGACATCTGAAcCT
			PPARG	CTGGGACATCTGAAcCTCCT

Fonte: Própria (2021).^a Informações básicas sobre os SNPs selecionados: gene, posição no genoma e os alelos em sua forma referência / alternativa (Alelos); ^b Resultados LASAGNA: Possíveis FTs e as referentes sequências dos sítios-alvo criados pelos SNPs identificados em letra minúscula em negrito.

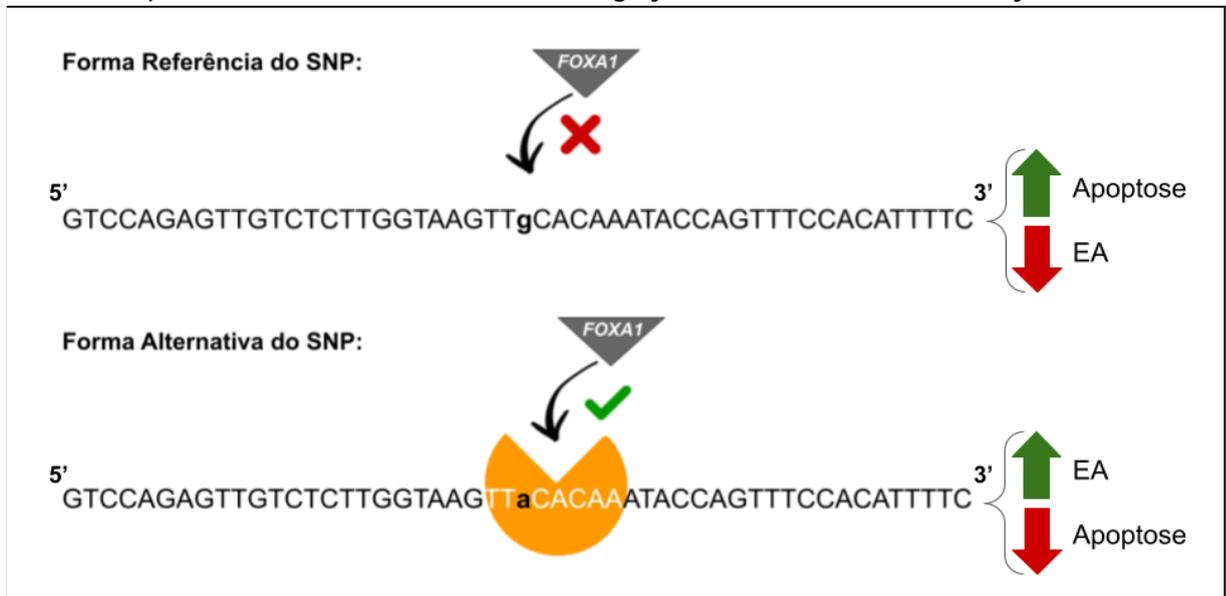
A presença do alelo T, do SNP g7:49824536T>C, localizado na região *upstream* do gene *EGR1* possivelmente proporciona um aumento de afinidade do FT *CTCF* ao sítio-alvo; enquanto a presença do alelo C possivelmente proporciona um aumento de afinidade do FT *REST* ao sítio-alvo. Segundo a literatura, *EGR1* e *CTCF* estão envolvidos na regulação da expressão do gene *MyoG* durante a diferenciação das células satélites do músculo esquelético bovino (ZHANG *et al.*, 2017). Além disso, SEKIYA *et al.* (2019) demonstrou que o fator de transcrição *CTCF* é necessário para transcrever o gene *EGR1* no início da fase G1 do ciclo celular por meio da organização da estrutura da cromatina. *EGR1* e *REST* podem regular a expressão do mRNA e a atividade do promotor Ca(v)3.2 da subunidade do canal de Ca²⁺, e com isso, a corrente de Ca²⁺, de modo a implicar na regulação da homeostase cardíaca e neural Ca²⁺ (VAN LOO *et al.*, 2012). Ademais, outros estudos sugerem uma possível regulação do TF *REST* sobre a expressão de *EGR1* (WELLS *et al.*, 2011; ZHANG *et al.*, 2016).

O alelo alternativo do SNP g18:50291051C>T, localizado na região 3' UTR do gene *CYP2B6*, possui potencial em criar um SLFT para o FT *ESR1*. LO *et al.*, (2010), demonstrou que em linhagens de células de câncer de mama humano, o *knockdown* do gene *ESR1* é associado a uma diminuição significativa da expressão

gênica de *CYP2B6*, o qual é altamente expresso em tumores de mama positivos para *ESR1*.

Já o alelo A do SNP g8:52762728G>A, localizado no gene candidato *PRUNE2*, possivelmente cria um SLFT para o FT *FOXA1* (Figura 3). Em um estudo em humanos, *FOXA1* foi constatado como um gene supressor do processo de apoptose em células hepáticas tumorais (YUAN *et al.*, 2020). E segundo a literatura, o gene *PRUNE2* está envolvido na regulação do processo apoptótico, o qual consiste na destruição celular por meio da ativação de genes que exercem essa função (LI *et al.*, 2011; DU *et al.*, 2017). Tal processo, relacionado à EA em bovinos *Bos taurus* (KHANSEFID *et al.*, 2017), em tecido hepático, pode ser desencadeado por eventos relacionados ao estresse oxidativo (DU *et al.*, 2017), que segundo Tizioto *et al.* (2015b), está aumentado em animais Nelore com maiores valores de CAR (ineficientes). Assim, propomos que a expressão do gene *PRUNE2* em tecido hepático pode indicar um processo apoptótico provocado por estresse oxidativo (Andressa Lima, comunicação pessoal), o qual pode ser suprimido por *FOXA1*, e assim favorecer a EA.

Figura 3. Esquema de um resultado da predição *in silico* de SNPs que possivelmente afetam sítios de ligação de fatores de transcrição.



Fonte: Própria (2021).

Região flanqueadora do SNP g8:52762728G>A para os alelos na forma referência e alternativa relativo ao gene *PRUNE2* do SNP g8:52762728G>A, na qual o alelo alternativo cria um sítio de ligação para o fator de transcrição *FOXA1*, o qual é candidato a suprimir o processo apoptótico provocado pela expressão do gene *PRUNE2*.

7.3 Predição *in silico* de SNPs que Possivelmente Afetam Sítios de Ligação de miRNAs

A predição *in silico* verificou a possibilidade de os SNPs selecionados nas regiões 5' e 3' UTR dos genes candidatos afetarem sítios de ligação para miRNAs de bovinos (Tabela 5). A influência dos alelos referência e alternativo dos SNPs, sobre os sítios de ligação de miRNAs, foi analisada por meio da variação da energia livre de hibridização predita para a ligação do miRNA ao seu respectivo sítio-alvo. Como já mencionamos, definimos como parâmetro limite a energia livre de hibridização predita de -18 kcal/ mol. Assim, consideramos que há ligação do miRNA ao seu respectivo sítio se a energia livre for menor que -18 kcal/mol, e então, esse possui um maior potencial de atuar na repressão da expressão gênica.

Tabela 5. Top 5 das diferenças de energias de hibridização preditas para a ligação dos microRNAs bovinos aos seus respectivos sítios-alvo, comparando-se a forma alternativa e referência dos SNPs selecionados na região UTR dos genes candidatos.

SNP ^a	miRNA ^b	Energia ALT ^c	Energia REF ^d	Diferença de Energia ^e
CYP2B6				
g18:50290773C>T	bta-miR-2285s	-17,8	-20,2	-2,40
	bta-miR-1298	-19,1	-16,5	2,60
	bta-miR-12016	-20,8	-16,9	3,90
	bta-miR-6120-3p	-19,2	-16,8	2,40
	bta-miR-190b	-18	-15,3	2,70
g18:50291018C>G	bta-miR-2885	-19,8	-17	2,80
	bta-miR-10163-3p	-20,4	-17,8	2,60
	bta-miR-12038	-18,6	-15,8	2,80
	bta-miR-2285bi	-18,1	-15,3	2,80
	bta-miR-551a	-19,9	-16,3	3,60

Fonte: Própria (2021).

^a Posição do SNP no genoma; ^b microRNAs bovinos; ^c Energia de hibridização predita para a forma alternativa do SNP; ^d Energia de hibridização predita para a forma referência do SNP; Diferença entre a energia de hibridização predita para a forma referência do SNP e a forma alternativa. Destacados em **negrito** encontram-se os alelos preditos para ligação mais estáveis entre miRNA-mRNA.

Continuação da Tabela 5.:

Tabela 5. Top 5 das diferenças de energias de hibridização previstas para a ligação dos microRNAs bovinos aos seus respectivos sítios-alvo, comparando-se a forma alternativa e referência dos SNPs selecionados na região UTR dos genes candidatos.

SNP ^a	miRNA ^b	Energia ALT ^c	Energia REF ^d	Diferença de Energia ^e
CYP2B6				
g18:50291051C>T	bta-miR-345-5p	-18,9	-15,6	3,30
	bta-miR-144	-20	-16,2	3,80
	bta-miR-2370-5p	-17,3	-19,6	-2,30
	bta-miR-2358	-20,8	-16,3	4,50
	bta-miR-151-5p	-20,6	-17,1	3,50
g18:50307954A>T	bta-miR-2396	-15,9	-19,5	-3,60
	bta-miR-26a	-18,3	-14,6	3,70
	bta-miR-2284e	-19,3	-15,8	3,50
	bta-miR-11988	-19,9	-16,7	3,20
	bta-miR-2370-5p	-16,9	-20,5	-3,60
g18:50307956C>T	bta-miR-11999	-18,6	-14,9	3,70
	bta-miR-574	-20,8	-17	3,80
	bta-miR-2285db	-18,1	-13,6	4,50
	bta-miR-424-3p	-19,4	-15,9	3,50
	bta-miR-6520	-19,5	-15	4,50
g18:50307959T>C	bta-miR-23a	-16,3	-18,7	-2,40
	bta-miR-654	-16	-20,3	-4,30
	bta-miR-25	-17,1	-20,7	-3,60
	bta-miR-411b	-15,9	-19,8	-3,90
	bta-miR-2380	-18,8	-16,3	2,50

Fonte: Própria (2021).

^a Posição do SNP no genoma; ^b microRNAs bovinos; ^c Energia de hibridização prevista para a forma alternativa do SNP; ^d Energia de hibridização prevista para a forma referência do SNP; Diferença entre a energia de hibridização prevista para a forma referência do SNP e a forma alternativa. Destacados em **negrito** encontram-se os alelos previstos para ligação mais estáveis entre miRNA-mRNA.

Continuação da Tabela 5.:

Tabela 5. Top 5 das diferenças de energias de hibridização previstas para a ligação dos microRNAs bovinos aos seus respectivos sítios-alvo, comparando-se a forma alternativa e referência dos SNPs selecionados na região UTR dos genes candidatos.

SNP ^a	miRNA ^b	Energia ALT ^c	Energia REF ^d	Diferença de Energia ^e
<i>CYP2B6</i>				
g18:50308024G>A	bta-miR-6120-3p	-16,1	-20,9	-4,80
	bta-miR-541	-17,5	-21,4	-3,90
	bta-miR-767	-16,5	-20,1	-3,60
	bta-miR-2344	-16	-21,3	-5,30
	bta-miR-2457	-16,4	-21,3	-4,90
<i>EGR1</i>				
g7:49829451A>G	bta-miR-2408	-21,5	-15,4	6,10
	bta-miR-2319a	-21	-15,7	5,30
	bta-miR-1264	-19,7	-14,8	4,90
	bta-miR-339a	-20	-15,8	4,20
	bta-miR-1307	-17,9	-19,2	1,30

Fonte: Própria (2021).

^a Posição do SNP no genoma; ^b microRNAs bovinos; ^c Energia de hibridização prevista para a forma alternativa do SNP; ^d Energia de hibridização prevista para a forma referência do SNP; Diferença entre a energia de hibridização prevista para a forma referência do SNP e a forma alternativa. Destacados em **negrito** encontram-se os alelos preditos para ligação mais estáveis entre miRNA-mRNA.

Todos os SNPs localizados nas regiões UTR nos genes candidatos *CYP2B6* e *EGR1*, em sua forma referência e alternativa, possivelmente afetam sítios de miRNAs de bovinos.

Por exemplo, a energia livre de hibridização prevista da ligação do miRNA bta-miR-106b ao seu respectivo sítio-alvo, para a forma alternativa do SNP g18:50308024G>A, localizado na região 5' UTR do gene *CYP2B6*, é de -16,2 kcal/mol, enquanto que com a sequência da forma referência é de -19,2 kcal/mol, ou seja, por influência da forma referência do SNP a ligação é prevista como mais estável do miRNA bta-miR-106b ao seu respectivo sítio-alvo. O miRNA bta-miR-106b foi regulado positivamente em estudo de bovinos com alto valor de

CAR, além disso sua expressão em fígado de novilhos mostrou-se altamente influenciada pela dieta (AL-HUSSEINI *et al.*, 2015; ROMAO *et al.*, 2012).

O miRNA bta-mir-23a teve sua energia livre de hibridização predita alterada pelo alelo C do SNP g18:50307959T>C, com energia livre de hibridização de -16,3 kcal/mol, tornando-se instável a predição de ligação do miRNA ao seu sítio de ligação, enquanto que na presença do alelo T desse SNP temos uma ligação predita como mais estável entre o bta-mir-23a e o gene *CYP2B6*, com uma energia livre de hibridização predita de -18,7 kcal/mol, o que acarretaria em uma diminuição da expressão do gene candidato e poderia auxiliar em um aumento de EA. Em um estudo sobre adipogênese bovina no músculo esquelético, a expressão de bta-mir-23a foi regulada negativamente durante a diferenciação celular do músculo esquelético bovino fetal, o que indica que o miRNA atua como um anti-adipogênico no processo de adipogênese intramuscular precoce, de forma a interferir no marmoreio da carne bovina e conseqüentemente na qualidade da mesma (GUAN *et al.*, 2017). Segundo o estudo de Nascimento *et al.* (2016), animais eficientes apresentaram menores níveis de gordura intramuscular em comparação com os animais ineficientes, e constatou-se que os animais eficientes possuíam carne dura, o que influencia na palatabilidade e no valor de mercado.

Do mesmo modo, a energia livre de hibridização predita para a ligação do miRNA bta-miR-339a ao seu respectivo sítio-alvo, influenciada pela forma referência do SNP g7:49829451A>G, localizado na região 3' UTR do gene *EGR1*, é de -15,8 kcal/mol, enquanto que com a sequência da forma alternativa do SNP é de -20,0 kcal/mol, desse modo, com a presença do alelo alternativo G temos uma ligação predita mais estável. Segundo Oliveira *et al.* (2018), o miRNA bta-miR-339a/b foi DE em bovinos com valores extremos para CAR, sendo regulado negativamente no fígado de animais mais eficientes. Ainda, em outros estudos, bta-miR-339a/b foi apontado como mais expresso em tecido adiposo bovino, e *EGR1* indicado por potencialmente regular o metabolismo e deposição de gordura em bovinos de corte, afetando as funções dos adipócitos, por meio da interação com genes envolvidos na via de sinalização (GU *et al.*, 2007; HUANG *et al.*, 2017).

Adicionalmente, o miRNA bta-miR-1307, com a mediação do SNP em sua forma alternativa a energia livre de hibridização predita é -17,9 kcal/mol, diminuindo a ligação do miRNA ao seu sítio, frente ao alelo referência, com a energia livre de hibridização predita é de -19,2 kcal/mol, aumentando a estabilidade da ligação do

miRNA ao gene *EGR1*. Em um estudo sobre miRNAs de bovinos divergentes para CAR, o miRNA bta-mir-1307 foi descrito como um isomiRs expresso nas sequências de miRNA-seq em fígado de animais ineficientes para CAR (AL-HUSSEINI *et al.*, 2015).

7.4 Análise de Associação entre os SNPs Potencialmente Reguladores da expressão gênica e Medidas Fenotípicas de Eficiência Alimentar

Nas análises de associação entre os SNPs e as medidas fenotípicas de EA, foram consideradas significativas as associações com valor $P < 0.05$ (Tabela 6). A magnitude da influência predita do SNP sobre a medida fenotípica com a qual ocorreu uma associação significativa é dada pelo coeficiente de regressão β (SHARMA *et al.*, 2018).

Tabela 6. SNPs dos genes de interesse com associação significativa com as medidas fenotípicas de eficiência alimentar.

Gene	SNP	Variante	Feno.	Valor P	β
CTGF	g9:69891302G>T	Upstream	EAB	0.028	-0.018
	g9:69891686C>T	Upstream	EAB	0.028	-0.018
COL1A1	g19:36456412A>G	Upstream	GPD	0.033	-0.169
			CA	0.0004	1.628
PRUNE2	g8:52762269A>G	Missense	CAR	0.039	0.252
	g8:52762340C>G	Missense	CAR	0.011	-0.511
	g8:52762632C>T	Missense	CAR	0.039	0.252
	g8:52762675G>A	Missense	CAR	0.039	-0.252
	g8:52762728G>A	Missense	CAR	0.039	-0.252
	g8:52762780T>G	Missense	CAR	0.039	-0.252
	g8:52763140G>T	Missense	CAR	0.039	0.252
	g8:52763476C>T	Missense	CAR	0.043	0.248
g8:52763771T>A	Missense	CAR	0.039	-0.252	

Fonte: Própria (2021).

Gene: Nomenclatura do gene de interesse; SNP: Posição no genoma de referência e os alelos Referência > Alternativo dos SNPs; Variante: Anotação dos SNPs; Alelos: Forma referência/ alternativa; Freq.: Frequência dos alelos; Feno.: Medidas fenotípicas de eficiência alimentar (GPD, kg/dia; CMS, kg/dia; EAB, kg ganho/kg MS; CA, kg MS/kg ganho; CAR, kg MS/dia); Valor P: Valor P significativos; β : valores do coeficiente de regressão referente ao alelo alternativo dos SNPs.

Continuação da Tabela 6.:

Tabela 6. SNPs dos genes de interesse com associação significativa com as medidas fenotípicas de eficiência alimentar.

Gene	SNP	Variante	Feno.	Valor P	β
PRUNE2	g8:52764043C>T	Missense	CAR	0.039	0.252
	g8:52764346A>G	Missense	CAR	0.039	0.252
	g8:52764655C>T	Missense	CAR	0.039	0.252
	g8:52764769A>C	Missense	CAR	0.039	0.252
	g8:52765101A>G	Missense	CAR	0.039	0.252
	g8:52765137T>A	Missense	CAR	0.039	-0.252
	g8:52765819T>C	Missense	CAR	0.039	-0.252
	g8:52765861C>A	Missense	CAR	0.039	-0.252
	g8:52766109G>A	Missense	CAR	0.039	-0.252
	g8:52766625T>C	Missense	CAR	0.039	-0.252

Fonte: Própria (2021).

Gene: Nomenclatura do gene de interesse; SNP: Posição no genoma de referência e os alelos Referência > Alternativo dos SNPs; Variante: Anotação dos SNPs; Alelos: Forma referência/ alternativa; Freq.: Frequência dos alelos; Feno.: Medidas fenotípicas de eficiência alimentar (GPD, kg/dia; CMS, kg/dia; EAB, kg ganho/kg MS; CA, kg MS/kg ganho; CAR, kg MS/dia); Valor P: Valor P significativos; β : valores do coeficiente de regressão referente ao alelo alternativo dos SNPs.

Nesse contexto, os SNPs g9:69891302G>T e g9:69891686C>T, localizados na região *upstream* do gene *CTGF*, em completo DL ($r^2=1$), que possuem associação significativa com a medida fenotípica EAB, cujos alelos alternativos influenciam negativamente a característica (β negativo). O alelo T do SNP g9:69891302G>T possivelmente cria um sítio de ligação para o FT *Gfi1* (Tabela 4), o qual é descrito na literatura como um inibidor da transcrição do gene *CTGF* (XIAN *et al.*, 2017). O gene *CTGF* apresentou uma diminuição de expressão no grupo de animais eficientes para CAR (TIZIOTO *et al.*, 2015b), contudo esse resultado contradiz nossos resultados de associação, nos quais o alelo T foi associado negativamente a EAB, o que pode ser decorrente da baixa correlação genética entre CAR e EAB (FERNANDES *et al.*, 2014; FREITAS NETO, 2013). Contudo, devido ao limitado número de animais com dados de expressão em tecido hepático utilizados no estudo de Tizioto *et al.*, (2015b) que apresentam o alelo T, infelizmente, não foi possível confirmar o impacto do SNP sobre a expressão do *CTGF*.

A CA é um índice de EA capaz de identificar animais eficientes no rebanho, porém por ser uma variável correlacionada com ganho de peso, a seleção por ela pode resultar no aumento do tamanho adulto dos animais e das exigências nutricionais, o que implica na maior necessidade de alimentos (BISHOP *et al.*, 1991; ARTHUR *et al.*, 2004). Contudo, nossos resultados mostraram que o alelo G do SNP g19:36456412A>G, localizado na região *upstream* do gene *COL1A1*, foi associado negativamente com GPD ($\beta = -0.169$, valor $P = 0.033$), e com um aumento de CA ($\beta = 1.628$, valor $P = 0.0004$). O gene *COL1A1* possui expressão aumentada em grupos de animais ineficientes (CAR positivo) (TIZIOTO *et al.*, 2015). Apesar da correlação entre CAR e CA ser moderada, como descrito por (ARTHUR *et al.*, 2001; LANCASTER *et al.*, 2009; KELLY *et al.* 2010) é possível que o alelo alternativo do SNP g19:36456412A>G que beneficia CA ($\beta = 1.628$), possa influenciar na expressão de *COL1A1* resultando em um baixo valor de CAR e consequentemente favorecendo a EA como um todo.

Por fim, as sete variantes *missenses* do bloco de SNPs do gene *PRUNE2* (Figura 2), possuem associação significativa com a medida fenotípica CAR (valor $P = 0.039$) (Tabela 6). Como já mencionamos, o gene *PRUNE2* está envolvido na regulação do processo apoptótico (LI *et al.*, 2011), que é relacionado à EA em bovinos *Bos taurus* (KHANSEFID *et al.*, 2017). Em tecido hepático, o processo apoptótico pode ser encadeado por estresse oxidativo (DU *et al.*, 2017), sendo mostrado que a expressão de *PRUNE2* está aumentada em fígado de animais Nelore com maiores valores de CAR (TIZIOTO *et al.*, 2015b), o que corresponde aos animais ineficientes. Desse modo, a expressão de *PRUNE2* em tecido hepático pode indicar um processo apoptótico provocado por estresse oxidativo (Andressa Lima, comunicação pessoal), aumentando os valores do fenótipo CAR, diminuindo a EA. Novos estudos devem ser realizados para investigar o impacto dessas variantes sobre a expressão do gene *PRUNE2*.

7.5 Análise de Integração

A análise de integração de dados entre os SNPs selecionados e regiões de QTLs de fenótipos relacionados a pecuária bovina, possibilitou um enriquecimento para fenótipos de qualidade de carne, fisiologia bovina, características do leite bovino, entre outros.

Como já mencionado, para isso, os SNPs foram sobrepostos com regiões de QTL do genoma bovino provenientes do arquivo *Cattle QTL database* (ROSEN *et al.*, 2020), e com os resultados do nosso grupo de pesquisa de QTLs identificados por GWAS (TIZIOTO *et al.*, 2013; TIZIOTO *et al.*, 2015a). Os resultados são apresentados na Tabela 7.

Tabela 7. SNPs selecionados nas janelas dos genes candidatos integrados a regiões de QTLs do *Cattle QTL Database* e de resultados de GWAS do grupo de pesquisa Biotecnologia Animal.

Genes Candidatos	SNPs integrados	QTLs
<i>COL1A1</i>	14 *	- Tamanho da panturrilha; - Facilidade de parto; - Gordura intramuscular; - Conteúdo de ácido lignocérico; - Teor de ácido butírico no leite; - Teor de ácido cáprico no leite; - Teor de ácido caprílico no leite; - Teor de ácido linoleico conjugado no leite; - Teor de ácido mirístico no leite; - Teor de ácido oleico no leite; - Teor de ácido palmítico no leite; - Teor de ácido palmitoleico no leite; - Circunferência escrotal; - Contagem de inchaço do Úbere;
<i>CTGF</i>	2 **	- Coloração de gordura - espectro vermelho (AFAT);
<i>CYP2B6</i>	40 *	- Índice de nascimento; - Suscetibilidade a <i>Mycobacterium paratuberculosis</i> ;
<i>EGR1</i>	7 *	- Tolerância ao frio;
<i>PRUNE2</i>	283 *	- Teor de ácido laurico no leite; - Teor de ácido butírico no leite; - Conteúdo de zinco no músculo;
	1 ***	- Conteúdo mineral K;
	1 ***	- Conteúdo mineral Ca;

Fonte: Própria (2021).

* *Cattle QTL database*; ** TIZIOTO *et al.*, 2013; *** TIZIOTO *et al.*, 2015a;

Foram identificados 346 SNPs selecionados nas janelas dos genes candidatos *COL1A1*, *CYP2B6*, *EGR1* e *PRUNE2*, localizados em regiões de QTLs

do banco de dados *Cattle QTL* (ROSEN *et al.*, 2020) para 21 características; e quatro SNPs nos genes candidatos *CTGF* e *PRUNE2*, integrados em regiões de QTL da mesma população deste estudo, para coloração de gordura (TIZIOTO *et al.*, 2013) e conteúdo mineral (TIZIOTO *et al.*, 2015a).

Sete SNPs relativos ao gene *EGR1* foram integrados em uma região de QTL para a característica tolerância ao frio (TF), 14 SNPs do gene *COL1A1* foram integrados em QTLs de gordura intramuscular (GI) e conteúdo de ácido lignocérico (CAL), e 283 SNPs selecionados na região do *PRUNE2* em QTL de conteúdo de zinco no músculo (CZM).

Referente ao QTL de TF, YOUNG *et al.*, (1981), relatou que o estresse ocasionado pelo frio pode resultar em alterações hormonais e adaptativas, como o aumento da taxa metabólica de repouso, resultando no direcionamento da energia oriunda da alimentação para suprir as necessidade energéticas para a manutenção da temperatura corporal e não para o crescimento do animal. Desse modo, a TF pode beneficiar a EA, de forma a selecionar animais que são menos afetados pelas mudanças oriundas do estresse ao frio.

Já com relação às características GI e CAL, o conteúdo de ácidos graxos desempenha papel significativo em processos biológicos como adipogênese, de forma a auxiliar no entendimento da regulação do tecido adiposo em bovinos de corte, que é um fator determinante na qualidade da carne e no seu valor nutricional (ROMAO *et al.*, 2014; OLIVEIRA *et al.*, 2019). O ácido lignocérico, com cadeia carbonada com mais de 20 carbonos, pode ser encontrado em gordura animal e vegetal (LEOPOLDINO JÚNIOR, 2011). E, com relação ao QTL de CZM, o mineral zinco é caracterizado como um inibidor das proteínas calpastatina e calpaína, essas possuem um papel importante no processo de maturação da carne bovina (KOOHMARAIE, 1990). Assim, níveis de concentração aumentados de zinco no músculo bovino podem colaborar com a diminuição da função da calpastatina e da calpaína e, conseqüentemente, ocasionar a diminuição da maciez da carne (KOOHMARAIE, 1990).

Os SNPs g19:36456642C>G e g19:36462243C>T, selecionados na janela do gene *COL1A1*, foram integrados ao QTL bovino de GI. Em estudos com suínos, o gene *COL1A1* foi identificado como um gene DE relacionado à matriz extracelular em grupos de animais extremos para teor de GI, e também foi um dos genes com maior expressão no tecido adiposo, podendo desempenhar um papel importante

neste tecido (LIM *et al.*, 2016; XING *et al.*, 2019). Ainda, em estudo com frangos, o gene *COL1A1* foi identificado como um biomarcador para células de gordura intramuscular (LI *et al.*, 2020). Além disso, o SNP g19:36456642C>G, em sua forma referência, possui potencial em criar o SLFT para o FT *Foxa2* (Tabela 4), que segundo a literatura, em modelos de camundongos, foi identificado como regulador na diferenciação e metabolismo de adipócitos e da glicose (WOLFRUM *et al.*, 2003). O tecido adiposo modula vários processos relacionados a EA, como o metabolismo lipídico e a homeostase da glicose (MCKENNA *et al.*, 2018).

Adicionalmente, o SNP g8:52765819T>C, selecionado na janela do gene *PRUNE2*, integrado ao QTL bovino que afeta o CMZ, em sua forma referência possui potencial em criar o SLFT para o FT *YY1* (Tabela 4). O FT *YY1* é classificado como uma proteína dedo de zinco a qual é estabilizada estruturalmente pelo íon zinco (ISALAN, 2013), essa possui papel importante no desenvolvimento e regeneração do músculo esquelético (CHEN *et al.*, 2019). Em estudos com camundongos, a deleção condicional de *YY1* resultou no não desenvolvimento do músculo, e a deleção induzida foi prejudicial para o processo de regeneração muscular (CHEN *et al.*, 2019). A implicação do gene *PRUNE2* no desenvolvimento muscular também foi relatada.

8. Conclusão

A partir desse trabalho foram identificadas variantes, por meio de uma janela de prospecção de SNPs, com potencial regulatório sobre a expressão dos genes candidatos previamente relacionados a medidas fenotípicas de eficiência alimentar em animais Nelore (*CTGF*, *COL1A1*, *CYP2B6*, *EGR1* e *PRUNE2*). Polimorfismos dos genes *CTGF*, *COL1A1* e *PRUNE2* foram associados a características fenotípicas de importância produtiva.

Os SNPs identificados, priorizados pela anotação, contribuíram com o conhecimento de novas regiões reguladoras envolvidas com características de eficiência alimentar. Esses são candidatos para validação *in vitro*, e dessa forma garantir sua futura aplicabilidade em programas de melhoramento animal, como a inclusão em chips de genotipagem, de forma a beneficiar o produtor.

9. Referências Bibliográficas

ABIEC. **A Associação Brasileira das Indústrias Exportadoras de Carnes**. 2020. Disponível em: <http://abiec.com.br/BrazilianBeef.aspx>. Acesso em: 10 dez. 2020.

ABO-ISMAIL, M. K. et al. Identification of single nucleotide polymorphisms in genes involved in digestive and metabolic processes associated with feed efficiency and performance traits in beef cattle^{1,2}. **Journal Of Animal Science**, v. 91, n. 6, p. 2512-2529, 2013.

ABREU, J. G. et al. (2002). Connective-tissue growth factor (CTGF) modulates cell signalling by BMP and TGF- β . **Nat Cell Biol** 4, 599–604, 2002.

ADCOCK, I. M.; CARAMORI, Gaetano. Transcription Factors. **Asthma And Copd**, p. 373-380, 2009.

AL-HUSSEINI, W. et al. Characterization and Profiling of Liver microRNAs by RNA-sequencing in Cattle Divergently Selected for Residual Feed Intake. **Asian-Australasian Journal Of Animal Sciences**, v. 29, n. 10, p. 1371-1382, 2015.

ALEXANDRE, P. A. et al. Bovine NR113 gene polymorphisms and its association with feed efficiency traits in Nelore cattle. **Meta Gene**, v. 2, p. 206-217, 2014.

ALZATE-MARIN, A.L. et al. Seleção assistida por marcadores moleculares visando ao desenvolvimento de plantas resistentes a doenças, com ênfase em feijoeiro e soja. **Fitopatologia Brasileira**, v.30, p.333-342, 2005.

ARAÚJO, R. O. **Densidade e Análise de Desequilíbrio de Ligação de Polimorfismos em Genes Candidatos em Bovinos da Raça Girolando**. 2014. 171 f. Tese (Doutorado) - Curso de Ciências Animais, Universidade de Brasília, Brasília, 2014.

ARTHUR, P. F. et al. Genetic and phenotypic variance and covariance components for feed intake, feed efficiency, and other postweaning traits in Angus cattle. **Journal Of Animal Science**, v. 79, n. 11, p. 2805-2812, 2001.

ARTHUR, P.F.; ARCHER, J.A.; HERD, R.M. Feed intake and efficiency in beef cattle: Overview of recent Australian research and challenges for the future. **Aust. J. Exp. Agric.**, v. 44, p. 361-369, 2004.

ACNB. **ASSOCIAÇÃO DOS CRIADORES DE NELORE DO BRASIL** . 2020. Disponível em: <<http://www.nelore.org.br>>.

BARRETT, J. C. et al. Haploview: analysis and visualization of LD and haplotype maps. **Bioinformatics**, v. 21, n. 2, p.263-265, 2004.

BASARAB, J. A. et al. Residual feed intake and body composition in young growing cattle. **Can. J. Anim. Sci.** 83, 189–204, 2003.

BASHIRULLAH, A.; COOPERSTOCK, R. L.; LIPSHITZ, H. D. Spatial and temporal control of RNA stability. **Proceedings Of The National Academy Of Sciences**, v. 98, n.13, p. 7025-7028, 2001.

BATISTA, E. L. da S. et al. Cenários para Intensificação da Bovinocultura de Corte Brasileira. Ed.Igc/UFMG, Belo Horizonte, p.1-65, 2020. Disponível em: <http://www.csr.ufmg.br/brasilpec>. Acesso em: 05 set. 2020.

BISHOP, M. D. et al. Divergent selection for postweaning feed conversion in Angus beef cattle: II. Genetic and phenotypic correlations and realized heritability estimate. **J. Anim. Sci.**, v. 69, p. 4360-4367, 1991.

BOES, M. et al. Connective Tissue Growth Factor (IGFBP-rP2) Expression and Regulation in Cultured Bovine Endothelial Cells. **Endocrinology**, v.140, n.4, p.1575–1580, 1999.

BROUGHTON J. P. et al. Pairing beyond the Seed Supports MicroRNA Targeting Specificity. **Mol Cell**. v. 64, p.320–33, 2016.

CAI, Y. et al. A Brief Review on the Mechanisms of miRNA Regulation. **Genomics, Proteomics & Bioinformatics**, v. 7, n. 4, p. 147-154, 2009.

CAMPOS, M. M. et al. **Eficiência alimentar: ferramenta para aumento de bioeficiência em gado de leite**. In: FERNANDE, Elizabeth Nogueira et al. Desafios para a sustentabilidade ambiental, social e econômica da cadeia produtiva do leite. Juiz de Fora: Embrapa Gado de Leite, 2012. Cap. 8. p. 1-189.

CARDOSO, T. F. et al. Multi-Omics Approach Reveals miR-SNPs Affecting Muscle Fatty Acids Profile in Nelore Cattle. **Genes**, v. 12, n. 1, p. 67, 2021.

CECCHINATO, A. et al. Candidate gene association analysis for milk yield, composition, urea nitrogen and somatic cell scores in Brown Swiss cows. **Animal**, v. 8, n. 7, p. 1062-1070, 2014.

CEPEA - **CENTRO DE ESTUDOS AVANÇADOS EM ECONOMIA APLICADA** (Piracicaba). PIB do Agronegócio Brasileiro. Disponível em: <https://www.cepea.esalq.usp.br/br/pib-do-agronegocio-brasileiro.aspx>. Acesso em: 18 nov. 2020.

CESAR, A.S.M. et al. Genome-wide association study for intramuscular fat deposition and composition in Nelore cattle. **Bmc Genetics**, v. 15, n. 1, p. 39, 2014.

CESAR, A.S.M. et al. Identification of putative regulatory regions and transcription factors associated with intramuscular fat content traits. **BMC Genomics** 19, 499, 2018.

CHEN, F. et al. YY 1 regulates skeletal muscle regeneration through controlling metabolic reprogramming of satellite cells. **The Embo Journal**, v. 38, n. 10, p. 111-111, 2019.

CHEN, Y. et al. Global gene expression profiling reveals genes expressed differentially in cattle with high and low residual feed intake. **Animal Genetics**, v. 42, n. 5, p. 475-490, 2011.

CHO, S. et al. Identification of genetic polymorphisms in FABP3 and FABP4 and putative association with back fat thickness in Korean native cattle. **Bmb Reports**, v. 41, n. 1, p. 29-34, 2008.

CHU, D.; WEI, L. Nonsynonymous, synonymous and nonsense mutations in human cancer-related genes undergo stronger purifying selections than expectation. **Bmc Cancer**, v. 19, n. 1, p. 111-111, 2019.

COSTA, D. P. B. da. et al. (2007). **CARACTERÍSTICAS DE CARÇAÇA DE NOVILHOS INTEIROS NELORE E F1 NELORE X HOLANDÊS**. *Ciência Animal Brasileira*, 8(4), 685-694. Recuperado de <https://www.revistas.ufg.br/vet/article/view/2689>

CREWS, D. H. Jr. Genetics of feed utilization and national cattle evaluation: A review. **Genet. Mol. Res.** v.4, p.152-165, 2005.

DANECEK, P. et al. The variant call format and VCFtools. **Bioinformatics**, v. 27, n. 15, p.2156-2158, 2011.

DU, X. et al. Elevated apoptosis in the Liver of Dairy Cows with Ketosis. **Cellular Physiology and Biochemistry**, v.43, n.2, p.568-578, 2017.

FABREGAT, A. et al. The Reactome Pathway Knowledgebase. **Nucleic Acids Research**, v.46, p.649-655, 2017.

FERNANDES, S.R. et al. Características de desempenho e eficiência alimentar de touros Purunã em crescimento de três classes de consumo alimentar residual. **Arq. Bras. Med. Vet. Zootec.**, Curitiba, v. 66, n. 1, p. 268-276, 2014.

FERRÃO, N. A. **Predição Computacional de Sítios de Ligação de Fatores de Transcrição Baseada em Gramáticas Regulares Estocásticas**. 2017. 161 f. Dissertação (Mestrado) - Curso de Bioinformática, Universidade de São Paulo, São Paulo, 2017.

FERREIRA, F. A. et al. **Consumo alimentar residual em bovinos de corte**. *Nutritime Revista Eletrônica*, Viçosa, v.12, n.6, p.4368-9006, 2015. Disponível em: www.nutritime.com.br. Acesso em: 17 abr. 2021.

FOGARTY, M. P. et al. Identification of a Regulatory Variant That Binds FOXA1 and FOXA2 at the CDC123/CAMK1D Type 2 Diabetes GWAS Locus. **Plos Genetics**, v. 10, n. 9, p. 1004633-1004642, 2014.

FORNES, O. et al. JASPAR 2020: update of the open-access database of transcription factor binding profiles. **Nucleic Acids Research**, v. 48, n. 1, p.87-92, 2019.

FREITAS NETO, M. D. de. **AVALIAÇÃO DO DESEMPENHO E DA EFICIÊNCIA ALIMENTAR DE BOVINOS NELORE**. 2013. 57 f. Tese (Doutorado) - Curso de Ciência Animal, Escola de Veterinária e Zootecnia, Universidade Federal de Goiás, Goiânia, 2012.

FREITAS, G. **Nelore: Conheça mais sobre a raça que representa 80% do gado de corte brasileiro.** 2013. Elaborado por equipe BeefPoint. Disponível em: <https://www.beefpoint.com.br/>. Acesso em: 10 fev. 2021.

GIRI P., MOHAPATRA B. (2017) Candidate Gene. Em: Vonk J., Shackelford T. (eds) Encyclopedia of Animal Cognition and Behavior. **Springer, Cham.**

GOMES, R. C.. **Entendendo a Eficiência Alimentar.** Publicações Embrapa Gado de Leite, Campo Grande, p. 22-24, fev. 2015.

GU, Z. et al. Identification and characterization of microRNAs from the bovine adipose tissue and mammary gland. **Febs Letters**, V. 581, n. 5, p. 981-988, 2007.

GUAN, L. et al. Bta-miR-23a involves in adipogenesis of progenitor cells derived from fetal bovine skeletal muscle. **Scientific Reports**, v. 7, n. 1, p. 43716-43722, 2017.

GUO, Y.; JAMISON, D. C. The distribution of SNPs in human gene regulatory regions. **Bmc Genomics**, v. 6, n. 1, 2005.

HA, M.; KIM, V. N. Regulation of microRNA biogenesis. **Nature Reviews Molecular Cell Biology**, v. 15, n. 8, p. 509-524, 2014.

HATZIS G., et al. Melatonin attenuates high fat diet-induced fatty liver disease in rats. **World J Hepatol.** v.5, n.4, p. 160, 2013.

HIGGINS, M. G. et al. GWAS and eQTL analysis identifies a SNP associated with both residual feed intake and GFRA2 expression in beef cattle. **Scientific Reports**, v. 8, n. 1, p. 1–12, 2018.

HUANG, W. et al. Global transcriptome analysis identifies differentially expressed genes related to lipid metabolism in Wagyu and Holstein cattle. **Scientific Reports**, v. 7, n. 1, p. 5278-5285, 2017.

IBGE. **Instituto Brasileiro de Geografia e Estatística.** 2019. Disponível em: <https://www.ibge.gov.br/>. Acesso em: 10 dez. 2020.

ISALAN, M. Zinc Fingers. Encyclopedia Of Biological Chemistry, p. 575-579, 2013. **Elsevier.**

JANSEN, R. mRNA localization: message on the move. **Nature Reviews Molecular Cell Biology** , v. 2, n. 4, p. 247-256, 2001.

JIAN, X.; BOERWINKLE, E.; LIU, X. In silico prediction of splice-altering single nucleotide variants in the human genome. **Nucleic Acids Research** , v. 42, n. 22, p. 13534-13544, 2014.

JIN, Y. et al. Catfish genomic studies: progress and perspectives. **Genomics In Aquaculture**, p. 73-104, 2016.

JUN, JI., LAU, L. Taking aim at the extracellular matrix: CCN proteins as emerging therapeutic targets. **Nat Rev Drug Discov** **10**, p.945–963, 2011.

KARISA, B. K. et al. Candidate genes and single nucleotide polymorphisms associated with variation in residual feed intake in beef cattle. **Journal Of Animal Science**, v. 91, n. 8, p. 3502-3513, 2013.

KELLY, A.K. et al. Effect of divergence in residual feed intake on feeding behavior, blood metabolic variables, and body composition traits in growing beef heifers. **J. Anim. Sci.**, v.88, p.109-123, 2010.

KHANSEFID, M. et al. (2017). Gene expression analysis of blood, liver, and muscle in cattle divergently selected for high and low residual feed intake. **Journal of Animal Science**, v.95, n.11, p.4764–4775, 2017.

KOCH, R. M.; SWIGER, L. A.; CHAMBERS, D. Efficiency of feed use in beef cattle. **Journal of Animal Science**, Savoy, v.22, p.486-494, 1963.

KONSTA, O. D. et al. An in silico Approach Reveals Associations between Genetic and Epigenetic Factors within Regulatory Elements in B Cells from Primary Sjögren's Syndrome Patients. **Frontiers In Immunology** , v. 6, 2015.

KOOHMARAIE, M. Inhibition of postmortem tenderization in ovine carcasses through infusion of zinc. **Journal Of Animal Science**, v. 68, n. 5, p. 1476-1483, 1990.

KORNBERG, A. (1962) On the metabolic significance of phosphorolytic and pyrophosphorolytic reactions. In: Horizons in Biochemistry (Kasha, M. ePullman, B; Eds.) p.251-264, **Academic Press**, New York.

KOZOMARA, A.; BIRGAOANU, M.; GRIFFITHS-JONES, S.. MiRBase: from microRNA sequences to function. **Nucleic Acids Research** , v. 47, n. 1, p.155-162, 2018.

LAMBI, A.G., et al. The Skeletal site-specific role of connective tissue growth factor in prenatal osteogenesis. **Dev. Dyn.** 241, 1944–1959, 2012.

LANCASTER, P.A. et al. Phenotypic and genetic relationships of residual feed intake with performance and ultrasound carcass traits in Brangus heifers. **J. Anim. Sci.**, v.87, p.3887-3896, 2009.

LASKY, J. A.. CONNECTIVE TISSUE GROWTH FACTOR. **Encyclopedia Of Respiratory Medicine**, p. 553-557, 2006.

LEE K.S. et al. Activation of hepatic stellate cells by TGF alpha and collagen type I is mediated by oxidative stress through c-myb expression. **J Clin Invest.** v.96, n. 5, p.2461, 1995.

LEE, C.; HUANG, C. H. LASAGNA-Search 2.0: Integrated transcription factor binding site search and visualization in a browser. **Bioinformatics**, v. 30, n. 13, p. 1923–1925, 2014.

LEE, Y. et al. MicroRNA maturation: stepwise processing and subcellular localization. **The Embo Journal**, v. 21, n. 17, p. 4663-4670, 2002.

LEOPOLDINO JÚNIOR, I. **Perfil de ácidos graxos na carne de cordeiros Santa Inês alimentados com gordura protegida e vitamina E**. 2011. 153 f. Dissertação (Mestrado) - Curso de Zootecnia, Produção e Nutrição de Ruminantes, Universidade Federal de Lavras, Lavras, 2011.

LI, J. et al. Identification of diverse cell populations in skeletal muscles and biomarkers for intramuscular fat of chicken by single-cell RNA sequencing. **Bmc Genomics**, v. 21, n. 1, p. 1234-1238, 2020.

LI, J.; DING, Y.; LI, A.. Identification of COL1A1 and COL1A2 as candidate prognostic factors in gastric cancer. **World Journal of Surgical Oncology**, v. 14, n. 297, p. 1- 5, 201, 2016.

LI, S. et al. The expression and localization of Prune2 mRNA in the central nervous system. **Neuroscience Letters**, v. 503, n. 3, p. 208-214, 2011.

LIM, K. S. et al. Identification of differentially expressed genes in longissimus muscle of pigs with high and low intramuscular fat content using RNA sequencing. **Animal Genetics**, v. 48, n. 2, p. 166-174, 2016.

LIMA, A. O. (2019). **Genes e Variantes Genéticas na Regulação da Eficiência Alimentar de Gado Nelore**. Ph.D. thesis, Federal University of São Carlos, São Carlos.

Andressa Lima. Prune Homolog 2 with BCH Domain (PRUNE2) gene expression affects feed efficiency-related traits in Nelore steers. Comunicação Pessoal.

LIU, J. Control of protein synthesis and mRNA degradation by microRNAs. **Curr. Opin. Cell Biol.**, v. 20, p.214-221, 2008.

LO, R. et al. Estrogen receptor-dependent regulation of CYP2B6 in human breast cancer cells. *Biochimica Et Biophysica Acta (Bba) - Gene Regulatory Mechanisms*, v. 1799, n. 5-6, p. 469-479, 2010.

LOPES, L. S. et al. **VIABILIDADE ECONÔMICA DA TERMINAÇÃO DE NOVILHOS NELORE E RED NORTE EM CONFINAMENTO NA REGIÃO DE LAVRAS-MG**. *Ciência e Agrotecnologia*, Lavras, v. 35, n. 4, p. 774-780, ago. 2011.

LU, D. et al. Genome-wide association analyses for growth and feed efficiency traits in beef cattle1. **Journal Of Animal Science**, v. 91, n. 8, p. 3612-3633, 2013.

LU, Y. et al. IFNL3 mRNA structure is remodeled by a functional non-coding polymorphism associated with hepatitis C virus clearance. **Scientific Reports**, 2015.

LUSE, D. S. The RNA polymerase II preinitiation complex: Through what pathway is the complex assembled? **Transcription**, v.5, n. NOV, 2013.

MARCHINI, J.; HOWIE, B. Genotype imputation for genome-wide association studies. **Nature Reviews Genetics**, London, v. 11, n. 7, p. 499-511, 2010.

MCKENNA, C. et al. Residual feed intake phenotype and gender affect the expression of key genes of the lipogenesis pathway in subcutaneous adipose tissue of beef cattle. **Journal Of Animal Science And Biotechnology**, 2018.

MCLAREN, W. et al. Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor. **Bioinformatics**, v. 26, p.2069-2070, 2010.

MIGNONE, F. et al. Untranslated regions of mRNAs. **Genome Biology**, v. 3, n. 3, 2002.

MITIS, T. et al. Transcription factors and evolution: an integral part of gene expression (review). **World Academy Of Sciences Journal**, v. 2, n. 1, p. 3-8, 2020.

MOBLEY, A. S.. Induced Pluripotent Stem Cells. **Neural Stem Cells And Adult Neurogenesis**, p. 67-94, 2019.

MOORE, S. S.; MUJIBI, F. D.; SHERMAN, E. L.. Molecular basis for residual feed intake in beef cattle. **Journal Of Animal Science** , v. 87, n. 14, p. 41-47, 2009.

NASCIMENTO, M. L. et al. Feed efficiency indexes and their relationships with carcass, non-carcass and meat quality traits in Nelore steers. **Meat Science**, v. 116, p. 78-85, jun. 2016.

Nelson, D. R. et al. P450 superfamily: update on new sequences, gene mapping, accession numbers and nomenclature. **Pharmacogenetics**, v.6, n.1, p.1-42, 1996.

NISHIZAKI, S. S. et al. Predicting the effects of SNPs on transcription factor binding affinity. **Bioinformatics**, [S.L.], v. 321, n. 123, p. 123-321, 2019.

NKRUMAH, J. D. et al. Relationships of feedlot feed efficiency, performance, and feeding behavior with metabolic rate, methane production, and energy partitioning in beef cattle. **J. Anim. Sci.** v.84, p.145-153, 2006.

O'BRIEN, Jacob et al. Overview of MicroRNA Biogenesis, Mechanisms of Actions, and Circulation. **Frontiers In Endocrinology**, 2018.

OKINE, E. K. et al. Residual feed intake – What is it and how does it differ from traditional concepts of feed utilization. **Canadian Society of Animal Science Conference**, p.10-13, 2003.

OLIVEIRA, P. S. N. de et al. An integrative transcriptome analysis indicates regulatory mRNA-miRNA networks for residual feed intake in Nelore cattle. **Scientific Reports**, v. 8, n. 1, p. 17072-17079, 2018.

OLIVEIRA, P. S. N. de et al. Co-Expression Networks Reveal Potential Regulatory Roles of miRNAs in Fatty Acid Composition of Nelore Cattle. **Frontiers In Genetics**, v. 10, n. 651, p. 1-14, 2019.

OLIVEIRA, P. S. N. de et al. Identification of genomic regions associated with feed efficiency in Nelore cattle. **Bmc Genetics**, 2014.

PAGEL, JI.; DEINDL, E. Disease progression mediated by egr-1 associated signaling in response to oxidative stress. **Int J Mol Sci**. v.13, n.10, p.13104, 2012.

POMERANTZ, M. M. et al. The 8q24 cancer risk variant rs6983267 shows long-range interaction with MYC in colorectal cancer. **Nat Genet**. v.41, n.8, p.882-4, 2009.

PURCELL, S. et al. PLINK: a toolset for whole-genome association and population-based linkage analysis. **American Journal of Human Genetics**, 81, 2007.

R Core Team (2018). **R: A language and environment for statistical computing**. R Foundation for Statistical Computing, Vienna, Austria.

RAMAZANI, Y. et al. Connective tissue growth factor (CTGF) from basics to clinics. **Matrix Biol**. 2018.

RAMÍREZ-CLAVIJO, S.; MONTOYA-ORTÍZ, G.. **Gene expression and regulation**. In: JM, Anaya. Autoimmunity: from bench to bedside. Bogota: El Rosario University Press, 2013. p. 1-855. Disponível em: <https://www.ncbi.nlm.nih.gov/books/NBK459447/>. Acesso em: 22 abr. 2021.

REHMSMEIER, M. et al. Fast and effective prediction of microRNA/target duplexes. **Rna** , v. 10, n. 10, p.1507-1517, 2004.

REINHART, B. J. et al. The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans*. **Nature**, v.403, p.901-906, 2000.

ROMAO, J. et al. MicroRNAs in bovine adipogenesis: genomic context, expression and function. **Bmc Genomics**, v. 15, n. 1, p. 137, 2014.

ROMAO, J. M. et al. Altered MicroRNA Expression in Bovine Subcutaneous and Visceral Adipose Tissues from Cattle under Different Diet. **Plos One**, v. 7, n. 7, p. 1-10, 2012.

ROSEN, B. et al. De novo assembly of the cattle reference genome with single-molecule sequencing. **Gigascience**, v. 9, n. 3, p. 1-9, 2020.

SAINSBURY, S.; BERNECKY, C.; CRAMER, P. Structural basis of transcription initiation by RNA polymerase II. **Nature Reviews Molecular Cell Biology**, v. 16, n. 3, p. 129–143, 2015.

SANTOS, P. M.. **EVOLUÇÃO DA RAÇA NELORE NO BRASIL**. 2017. 29 f. TCC (Graduação) - Curso de Bacharelado em Zootecnia, Universidade Federal de Goiás, Jataí, 2017.

SCARIA, V. et al. Host-virus interaction: a new role for micromnas. **Retrovirology**, v. 3, n. 1, p. 68, 2006.

SCHIAFFONATI, Luisa; TIBERIO, Laura. Gene expression in liver after toxic injury: analysis of heat shock response and oxidative stress-inducible genes. **Liver**, v. 17, n. 4, p. 183-191, 2008.

SEEG. **Sistema de Estimativas de Emissões e Remoções de Gases de Efeito Estufa**. 2019. Disponível em: <http://seeg.eco.br/>. Acesso em: 15 nov. 2020.

SEKIYA, T. et al. Involvement of CTCF in transcription regulation of EGR1 at early G1 phase as an architecture factor. **Scientific Reports**, [S.L.], v. 9, n. 1, p. 329, 23 jan. 2019.

SHARMA, A. et al. Jasmonic Acid Seed Treatment Stimulates Insecticide Detoxification in Brassica juncea L. **Frontiers In Plant Science**, v. 9, n. 444, p. 401-417, 2018.

SHASTRY, B. S.. SNPs: impact on gene function and phenotype. **Methods In Molecular Biology**, v.578, p.3-22, 2009.

SHERMAN, E. L. et al. Fine mapping quantitative trait loci for feed intake and feed efficiency in beef cattle1. **Journal Of Animal Science**, v. 87, n. 1, p. 37-45, 2009.

SINGH, Umesh et al. Molecular markers and their applications in cattle genetic research: a review. **Biomarkers And Genomic Medicine**, v. 6, n. 2, p. 49-58, 2014.

SLATKIN, M. Linkage disequilibrium - understanding the evolutionary past and mapping the medical future. **Nature Reviews Genetics**, v. 9, n. 6, p. 477-485, 2008.

SOLÍS-MUÑOZ, Pablo et al. Melatonin improves mitochondrial respiratory chain activity and liver morphology in ob/ob mice. **Journal Of Pineal Research**, v. 51, n. 1, p. 113-123, 2011.

SPITZ, F.; FURLONG, E. E. M. Transcription factors: from enhancer binding to developmental control. **Nature Reviews Genetics**, v. 13, n. 9, p. 613–626, 2012.

STEFL, S. et al. Molecular Mechanisms of Disease-Causing Missense Mutations. **Journal Of Molecular Biology**, v. 425, n. 21, p. 3919-3936, 2013.

STERGACHIS, A. B. et al. Exonic Transcription Factor Binding Directs Codon Choice and Affects Protein Evolution. **Science**, v. 342, n. 6164, p. 1367-1372, 2013.

STITZEL, M. L. et al. Global epigenomic analysis of primary human pancreatic islets provides insights into type 2 diabetes susceptibility loci. **Cell Metab**. 2010.

TAKEUCHI, F. et al. Linkage Disequilibrium Grouping of Single Nucleotide Polymorphisms (SNPs) Reflecting Haplotype Phylogeny for Efficient Selection of Tag SNPs. **Genetics**, v. 170, n. 1, p. 291-304, 2005.

TIZIOTO, P. C et al. Gene expression differences in Longissimus muscle of Nelore steers genetically divergent for residual feed intake. **Scientific Reports**, v. 6, n. 1, p.1-12, 2016.

TIZIOTO, P. C et al. Global liver gene expression differences in Nelore steers with divergent residual feed intake phenotypes. **Bmc Genomics**, v. 16, n. 1, 2015b.

TIZIOTO, P. C. et al. Detection of quantitative trait loci for mineral content of Nelore longissimus dorsi muscle. **Genetics Selection Evolution**, v. 47, n. 1, p. 15, 2015a.

TIZIOTO, P. C. et al. Genome scan for meat quality traits in Nelore beef cattle. **Physiological Genomics**, v. 45, n. 21, p. 1012-1020, 2013.

TIZIOTO, P.C. et al. Candidate genes for production traits in Nelore beef cattle. **Genetics And Molecular Research**, v. 11, n. 4, p. 4138-4144, 2012.

TSUCHIYA, S. et al. MicroRNA: biogenetic and functional mechanisms and involvements in cell differentiation and cancer J. **Pharmacol. Sci.**, v.101, p. 267-270, 2006.

USDA. **United States Department of Agriculture**. 2018. Disponível em: <<https://www.nass.usda.gov/>>. Acesso em: jun. 2019.

USDA. **United States Department of Agriculture**. 2019. Disponível em: <<https://www.nass.usda.gov/>>. Acesso em: jun. 2019.

VAN LOO, K. M. J. et al. Transcriptional Regulation of T-type Calcium Channel CaV3.2. **Journal Of Biological Chemistry**, v. 287, n. 19, p. 15489-15501, 2012.

VELAZCO, J. I. et al. Daily methane emissions and emission intensity of grazing beef cattle genetically divergent for residual feed intake. **Animal Production Science**, v. 57, n. 4, p. 627-635, 2016.

VELDEN, A.; THOMAS, A. The role of the 5' untranslated region of an mRNA in translation regulation during development. **The International Journal Of Biochemistry & Cell Biology**, v. 31, n. 1, p. 87-106, 1999.

VOZZI, P. A. et al. Structure and genetic variability in Nelore (*Bos indicus*) cattle by pedigree analysis. **Genetics And Molecular Biology**, v. 29, n. 3, p. 482-485, 2006.

WANG, Z. et al. RNA-Seq: a revolutionary tool for transcriptomics. **Nature Reviews Genetics**, v. 10, n. 1, p. 57-63, 2009.

WELLS, Timothy et al. Transcription Mapping of Embryonic Rat Brain Reveals EGR-1 Induction in SOX2+ Neural Progenitor Cells. **Frontiers In Molecular Neuroscience**, v. 4, p. 6, 2011.

WOLFRUM, C. et al. Role of Foxa-2 in adipocyte metabolism and differentiation. **Journal Of Clinical Investigation**, v. 112, n. 3, p. 345-356, 2003.

WUNSCH, A. et al. Single-nucleotide polymorphisms in the promoter region influence the expression of the human follicle-stimulating hormone receptor. **Fertility And Sterility**, v. 84, n. 2, p. 446-453, 2005.

XIAN, G. et al. Simvastatin attenuates macrophage-mediated gemcitabine resistance of pancreatic ductal adenocarcinoma by regulating the TGF- β 1/Gfi-1 axis. **Cancer Letters**, v. 385, p. 65-74, 2017.

XING, K. et al. Comparative adipose transcriptome analysis digs out genes related to fat deposition in two pig breeds. **Scientific Reports**, v. 9, n. 1, p. 12.925-12.930, 2019.

YAO, J. Imaging Transcriptional Regulation of Eukaryotic mRNA Genes: Advances and Outlook. **Journal of Molecular Biology**, v. 429, n. 1, p. 14–31, 2017.

YATES, Andrew D et al. Ensembl 2020. **Nucleic Acids Research** , p.682-688, 2019.

YOUNG, B. A. et al. Cold Stress as it Affects Animal Production. **Journal Of Animal Science**, v. 52, n. 1, p. 154-163, 1981.

YUAN, Z. et al. FOXA1 Promotes Cell Proliferation and Suppresses Apoptosis in HCC by Directly Regulating miR-212-3p/FOXA1/AGR2 Signaling Pathway. **Oncotargets And Therapy**, v.13, p.5231-5240, 2020.

ZANGER, U. M.; KLEIN, K.. Pharmacogenetics of cytochrome P450 2B6 (CYP2B6): advances on polymorphisms, mechanisms, and clinical relevance. **Frontiers In Genetics**, v. 4, n. 24, 2013.

ZERBINO, D. R. et al. Ensembl 2018. **Nucleic Acids Research** , v. 46, p. 754-761, 2018.

ZHANG, D. *et al.* Inhibition of REST Suppresses Proliferation and Migration in Glioblastoma Cells. **International Journal Of Molecular Sciences**, v. 17, n. 5, p. 664, 2016.

ZHANG, W. et al. Transcription factor EGR1 promotes differentiation of bovine skeletal muscle satellite cells by regulating MyoG gene expression. **Journal Of Cellular Physiology**, v. 233, n. 1, p. 350-362, 2017.