UNIVERSIDADE ESTADUAL PAULISTA – UNESP
CÂMPUS DE JABOTICABAL

ESTUDO DA FREQUÊNCIA DO CONSUMO DE ÁGUA E DEMAIS ATIVIDADES DE BOVINOS DE CORTE CRIADOS À PASTO

Rafael Nakamura Watanabe
Zootecnista

JABOTICABAL – SP
2022

UNIVERSIDADE ESTADUAL PAULISTA – UNESP

CÂMPUS DE JABOTICABAL

ESTUDO DA FREQUÊNCIA DO CONSUMO DE ÁGUA E DEMAIS ATIVIDADES DE BOVINOS DE CORTE CRIADOS À PASTO

Rafael Nakamura Watanabe

Orientador: Prof. Dr. Danísio Prado Munari

Coorientadora: Prof. Dra. Priscila Arrigucci Bernardes

Tese apresentada à Faculdade de Ciências Agrárias e Veterinárias – Unesp, Câmpus de Jaboticabal, como parte das exigências para obtenção do título de Doutor em Ciência Animal.

JABOTICABAL – SP

2022

W324e

Watanabe, Rafael Nakamura
Estudo da frequência do consumo de água e demais atividades de bovinos de corte criados à pasto / Rafael Nakamura Watanabe. -- Jaboticabal, 2022
60 p. : tabs., fotos

Tese (doutorado) - Universidade Estadual Paulista (Unesp), Faculdade de Ciências Agrárias e Veterinárias, Jaboticabal
Orientador: Danísio Prado Munari
Coorientadora: Priscila Arrigucci Bernardes

1. Expressão gênica. 2. Bovino de corte. 3. Animais comportamento. 4. Zebu. 5. Inteligência artificial. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca da Faculdade de Ciências Agrárias e Veterinárias, Jaboticabal. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

UNIVERSIDADE ESTADUAL PAULISTA

Câmpus de Jaboticabal

CERTIFICADO DE APROVAÇÃO

TÍTULO DA TESE: ESTUDO DA FREQUÊNCIA DO CONSUMO DE ÁGUA E DEMAIS ATIVIDADES DE BOVINOS DE CORTE CRIADOS À PASTO

AUTOR: RAFAEL NAKAMURA WATANABE
ORIENTADOR: DANISIO PRADO MUNARI
COORIENTADORA: PRISCILA ARRIGUCCI BERNARDES

Aprovado como parte das exigências para obtenção do Título de Doutor em GENÉTICA E MELHORAMENTO ANIMAL, pela Comissão Examinadora:

Prof. Dr. DANISIO PRADO MUNARI (Participaçao Virtual)
Departamento de Engenharia e Ciencias Exatas DECEx / FCAV UNESP Jaboticabal

Pós-Doutoranda TAINÃ FIGUEIREDO CARDOSO (Participaçao Virtual)
Embrapa Pecuária Sudeste / São Carlos/SP

Dra. ANA PAULA SBARDELLA (Participaçao Virtual)
Criaris Agropecuaria Inteligente Ltda / Goiânia/GO

Prof. Dr. GUILHERME BATISTA DO NASCIMENTO (Participaçao Virtual)
Centro Universitário de Adamantina - UniFAI / Adamantina/SP

Ph.D. LÚCIO FLÁVIO MACÊDO MOTA (Participaçao Virtual)
Department of Agronomy Food Natural Resources Animals and Environment - DAFNAE / University of Padova - Campus Agripolis

Jaboticabal, 18 de abril de 2022

DADOS CURRICULARES DO AUTOR

RAFAEL NAKAMURA WATANABE - nascido em 24 de setembro de 1990, na cidade de Jaboticabal, São Paulo, filho de Luís Paulo Watanabe e Iraci Maria da Silva Watanabe. Iniciou o curso de Zootecnia em março de 2011 na Universidade Estadual Paulista, Campus de Jaboticabal e obteve o título de bacharel em Zootecnia em março de 2016. Em agosto de 2016, ingressou no curso de mestrado pelo Programa de Pós-graduação em Genética e Melhoramento Animal na Faculdade de Ciências Agrárias e Veterinárias, UNESP, Campus de Jaboticabal sob orientação do Prof. Dr. Danísio Prado Munari e coorientação do Prof. Dr. Guilherme Batista do Nascimento. Foi bolsista CAPES pelo período de agosto de 2016 a julho de 2018. Em agosto de 2018, ingressou no curso de doutorado pelo Programa de Pós-graduação em Genética e Melhoramento Animal na Faculdade de Ciências Agrárias e Veterinárias, UNESP, Campus de Jaboticabal sob orientação do Prof. Dr. Danísio Prado Munari e coorientação da Profa. Dra. Priscila Arrigucci Bernardes. Foi bolsista CAPES de agosto de 2018 a fevereiro de 2022.

Todos os dias, sob todos os pontos de vista, eu vou cada vez melhor

Émile Coué

DEDICO

Ao meu filho Thiago, sem você todo meu esforço seria em vão. Obrigado por ter me ensinado o que é amor.

Aos meus pais Iraci e Luis Paulo, por me proporcionarem a oportunidade de ser quem eu sou, por me ensinarem a ser quem eu sou.

A minha avó Helena, por dedicar-se toda uma vida a ser um exemplo de amor, carinho, esforço e dedicação.

AGRADECIMENTOS

Aos meus pais Luis Paulo e Iraci, por serem a base estrutural da minha vida profissional e pessoal. Por se dedicarem uma vida toda a me proporcionar educação, amor, carinho, afeto e cuidado. Minha mãe, exemplo de amor e cuidado. Meu pai, exemplo de dedicação e família. Ao meu filho Thiago, por ter me dado um significado, uma direção. Único e real amor. A minha avó Helena, por também se dedicar uma vida toda em me oferecer afeto e cuidado, por cuidar não apenas de mim, mas de toda uma família. Símbolo de família e que me ensinou que amor se demonstra de várias maneiras, desde uma comida predileta de um neto, até como se pronunciar “eu te amo”.

Ao meu orientador Prof. Danísio. Um exemplo de profissional e de ser humano. Me ensinou não apenas ser um profissional dedicado, como também um ser humano digno. Uma amostra extremamente acurada de orientador. Seus ensinamentos moldaram quem eu sou hoje e por isso serei eternamente grato. A minha amiga e coorientadora Profa. Priscila. Me ensinou que a dedicação é a chave para um final bem sucedido. Obrigado por me acolher e me ensinar grande parte do que sei hoje. Ao meu amigo e irmão Prof. Guilherme. Peça fundamental para meu crescimento. Obrigado por me ouvir, me acolher e aconselhar nas horas mais escuras. Obrigado por ser uma ótima companhia nos momentos de descontração.

Aos meus amigos que o Departamento de Ciências Exatas me deu. Lucas, Rebeka, Samla, Ana Paula, Letícia, Júlia, Anas Carolinas, Thomaz, Eliéder, Fernando e Leyde. Os dias seriam mais cinzentos sem vocês. Obrigado por todas as conversas na pausa para o café, obrigado por estarem sempre presentes não apenas no café. Vocês são a família que eu pude escolher. Aos meus amigos de infância, Rodrigo, Osmir, Tales, Rafael, Bruno, João, Marcela e William. Vocês, juntamente com meus primos Diego, Klevis, Loiane e Wellington me proporcionaram as melhores lembranças que tenho. A distância não apaga o amor por vocês. Aos meus amigos que convivo sempre, Bigode, Palada, Serginho, Véio, Carol e Poly. Obrigado por proporcionarem dias ímpares que serão lembrados e contados ao meu filho e netos. Obrigado por serem o ponto de equilíbrio nessa vida.

Agradeço a CAPES pelo bolsa concedida (processo número 001) e FAPESP pelo auxílio regular (processo número 2018/20753-7).

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

Sumário

RESUMO - iii
CAPÍTULO 1 – Considerações gerais 1
1 INTRODUÇÃO 1
2 OBJETIVOS 4
2.1 Objetivos gerais 4
2.2 Objetivos específicos 4
3 REVISÃO DE LITERATURA 4
3.1 RNAseq e genes relacionados a características de interesse 4
3.2 Uso de acelerômetros na pecuária 7
3.2.1 Algoritmos de classificação a partir de dados de acelerômetros 9
4. REFERÊNCIAS 11
Capítulo 2 - Gene expression and dispersion profile associated with water consumption frequency and general activity in grazing beef cattle 18
Abstract 18
Implications 19
1. Introduction 19
2. Material and methods 20
2.1 Experimental area 20
2.2 Behavioral observation and selection of extreme animals 20
2.3 RNA-seq 21
2.4 Bioinformatic Analysis 22
3. Results 22
4. Discussion 24
4.1 Water consumption frequency 24
4.2 General activity 28
5. Conclusion 29
References 30
Capítulo 3 - Strategy to Predict High and Low Frequency Behaviors Using Triaxial Accelerometers in Grazing of Beef Cattle 37
Simple Summary: 37
Abstract: 37
1. Introduction 38
2. Materials and Methods 39
2.1. Experimental Area and Animals 39
2.2. Accelerometers and Animal Behavior 40
2.3. Data Processing and Prediction Algorithms 41

2.4. Resampling Methods to Deal with Imbalanced Data 42
3. Results 43
4. Discussion 45
5. Conclusions 48
References 49
Appendices 54

ESTUDO DA FREQUÊNCIA DO CONSUMO DE ÁGUA E DEMAIS ATIVIDADES DE BOVINOS DE CORTE CRIADOS À PASTO

RESUMO - A aplicação da tecnologia na agropecuária auxilia a seleção de animais geneticamente superiores. A utilização de acelerômetros na predição do comportamento animal tem se mostrado como ferramenta com potencial para substituição de mão-de-obra humana, auxiliando na redução de custos para o produtor. Além disso, metodologias de sequência de RNA (RNAseq) para o conhecimento dos genes que são expressos em animais mais adaptados ao meio ambiente são fundamentais para o progresso genético de uma população. Portanto, os objetivos desse trabalho foram: (1) Estudar o diferencial de expressão e dispersão de genes em animais da raça Nelore e F1 (Nelore x Angus) com perfis extremos de frequência de consumo de água individual diário e demais atividades durante o período diurno mantidos em pastagens com diferentes níveis de adubação e com diferentes suplementações; (2) estudar diferentes estratégias na predição do comportamento de animais Nelore criados a pasto, especialmente para atividades de menor frequência, a partir de registros de acelerômetros. Foram observados os comportamentos de pastejo, ruminação, ócio em pé e em decúbito, frequência de consumo de água (WCF), frequência do animal no cocho e andar, de 36 animais Nelore e 33 animais F1, criados à pasto. A WCF foi anotada toda vez que o animal compareceu no bebedouro e as demais atividades a cada 10 minutos. Dez animais extremos (cinco extremos superiores e cinco extremos inferiores, para ambas as composições genéticas) foram selecionados para frequência relativa de WCF (razão entre WCF divido e a soma de todos os comportamentos) e frequência relativa da atividade geral (GA) (razão entre a soma de todos os comportamentos, exceto ócio e soma de todos os comportamentos). Estes animais extremos tiveram seu sangue coletado para subsequente sequenciamento de RNA. Nas análises de predição do comportamento animal, foram considerados nove animais da raça Nelore criados à pasto, acoplados com acelerômetros. Três metodologias de “Machine Learning” (“Random Forest” (RF), “Support Vector Machine” (SVM) e “Naive Bayes Classifier” (NBC)) foram comparadas para predição dos comportamentos de ruminação, pastejo, ócio, WCF, frequência do animal no cocho e andar. A fim de se verificar estratégias para melhor predição dos comportamentos devido ao desbalanceamento dos dados colhidos, foram utilizadas as metodologias de reamostragem “under” e “over-sampling”. As análises funcionais e de enriquecimento mostraram para a análise de diferencial de expressão gênica, um total de 13 e 3 genes distintos relacionados a processos biológicos, componentes celulares e vias de Kegg, para os grupos Nelore para frequência de consumo de água (NWCF) e Nelore para atividade geral (NGA), respectivamente. Para a análise de diferencial de dispersão, três e dois genes distintos foram enriquecidos, relacionados aos componentes celulares para NWCF e NGA, respectivamente. Nos grupos WCF e GA para F1 foram encontrados cinco genes distintos, relacionados a processos biológicos, componentes celulares e vias de Kegg. Os genes encontrados neste estudo foram associados a várias características importantes para a capacidade adaptativa dos animais. Nas predições

do comportamento animal, a acurácia geral foi maior para modelos de RF, sendo a maior para o modelo de RF treinado com dados “over-sampled”. Portanto, em geral, a melhor estratégia para classificar e predizer comportamentos mais frequentes foi com o algoritmo de RF e, quando comportamentos menos frequentes são o principal interesse, a estratégia mais adequada foi utilizar técnica de “over-sampling” para treinamento dos dados.

Palavras-chave: Adaptação, aprendizado de máquina, bovinos de corte, expressão gênica, sensor de movimento.

CAPÍTULO 1 – Considerações gerais

1 INTRODUÇÃO

Os avanços obtidos na produção animal têm se fundamentado diretamente no melhoramento genético animal, biotecnologia da reprodução, nutrição e manejo. Tais tecnologias, quando aplicadas simultaneamente, podem resultar em aumento na produção de carne. No entanto, ainda há muitos desafios que os profissionais dessas áreas precisam enfrentar para suprir a demanda mundial de alimentos, sem causar maiores impactos ao meio ambiente. A utilização de tecnologia aplicada à agropecuária pode ser uma alternativa para superar estes desafios e aumentar a produtividade.

Grande parte da produção de carne no Brasil tem como base a criação extensiva dos animais, em que aproximadamente 162 milhões de hectares são destinados à pastagem (ABIEC, 2020). Assim, na área de nutrição e manejo, muitos estudos têm sido realizados com diferentes sistemas de manejo do solo e dos animais, assim como na avaliação de diferentes suplementos à pasto, com a finalidade de obter animais com elevado desempenho em um sistema economicamente viável e com reduzido impacto ambiental. Isso porque a prática da suplementação contribui para aumentar o ganho de peso por animal por área. Da mesma forma, a adubação de pastagens pode melhorar a qualidade da forragem fornecida, a qual aliada à eficiência de pastejo pode aumentar o desempenho animal e reduzir os impactos ambientais.

O aumento da produtividade com atenção aos danos ambientais também é objetivo de estudo para pesquisadores da área do melhoramento genético animal. Segundo Hayes et al. (2013), três fatores serão responsáveis pela necessidade de mudanças nos objetivos de seleção utilizados recentemente pelos programas de melhoramento animal. O primeiro fator é a necessidade de adicionar novas características no processo de avaliação genética, as quais não faziam parte anteriormente e que poderão ser importantes. O segundo fator é a necessidade de melhor entendimento das características que têm apresentado pequeno ou nenhum ganho genético, como aquelas com elevado custo de mensuração (tais como a conversão alimentar em bovinos). O terceiro fator é a ocorrência cada vez mais eminente do aquecimento global, o qual modificará o ambiente em que a produção

animal é realizada, sendo desejáveis programas de melhoramento genético que possam responder rapidamente a estas mudanças.

Uma alternativa utilizada na adaptação a mudanças ambientais é o cruzamento entre raças ou o uso de raças mais adaptadas ao ambiente. Nos países tropicais, que já sofrem com a temperatura elevada, as principais raças criadas são de origem zebuína, por apresentarem maior adaptação a ambientes mais quentes e a parasitas (EBERHARDT et al., 2009). No Brasil, a produção de sêmen da raça taurina Angus aumentou e, em 2014, esta foi a segunda raça com maior produção de sêmen para bovinos de corte, após a raça Nelore (ASBIA, 2014). Esse fato pode indicar que os produtores estão acasalando reprodutores taurinos com fêmeas zebuínas para melhorar a qualidade da carne e manter a adaptação dos animais (MILLEN e ARRIGONI, 2013). No entanto, as previsões de um ambiente mais quente e de escassez de água deve direcionar os programas de melhoramento genético a selecionarem animais mais adaptados a essas novas condições, sem redução na produtividade. Assim, torna-se necessário obter informações que possam auxiliar na seleção de animais mais resistentes às estas mudanças.

A obtenção de informações que auxiliem a seleção dos animais que sejam mais adaptados às mudanças climáticas é de importância para o futuro da pecuária brasileira. Uma maneira de se obter essas informações é a compreensão dos genes que influenciam a resiliência dos animais à ambientes mais quentes e de baixa disponibilidade de água.

A técnica de sequenciamento de RNA é uma ferramenta capaz de auxiliar na compreensão destes genes. Esta técnica permite a síntese de sequências de RNA com menor custo e em grande quantidade se comparada com o tradicional sequenciamento de Sanger. Dentre as possibilidades que esta técnica permite, destaca-se o perfil de expressão gênica, permitindo mapear e quantificar transcriptomas a fim de se analisar a expressão destes genes em diferentes tecidos (CÁNOVAS et al., 2010).

A compreensão de como complexas redes de expressão gênica se comportam na adaptação às variações do ambiente pode ser verificada por meio do sequenciamento do RNA (MARGUERAT E BÄHLER, 2010), o que, no presente trabalho, é de fundamental importância para o estudo dos genes relacionados a

frequência de consumo de água e atividade em bovinos de corte criados à pasto. O estudo dos genes que se relacionem à essas características, pode trazer conhecimento de como os animais se adaptam a mudanças climáticas de forma eficiente, sem prejudicar a produtividade. O conhecimento destes genes requer medidas fenotípicas viáveis para que estes genes sejam relacionados à estas medidas.

A importância de algumas medidas para indicar a frequência de consumo de água dos animais pode ser observada no estudo de Murphy et al. (2017), os quais objetivaram predizer a demanda de água e sua relação com algumas características de importância econômica em bovinos leiteiros irlandês. Dentre as diversas informações, a frequência de consumo de água pode indicar a eficiência do animal para tolerância térmica, pois ao sofrer estresse calórico, as duas principais respostas dos bovinos são sudorese e aumento da frequência respiratória, em que ambas têm efeito colateral indesejável de esgotar as reservas de água corporal (COLLIER e GEBREMEDHIN, 2015), sendo necessária sua reposição. Além disso, a frequência de consumo de água também poderia influenciar na temperatura ruminal, podendo ser dependente da frequência de ingestão, da temperatura da água e da quantidade de água consumida em cada momento (LEES et al., 2018).

Outro fator que pode auxiliar na adaptação do animal ao ambiente, é a atividade de cada animal ao longo do dia. Segundo Nadeau et al. (2017), comportamento animal pode ser um importante fator na regulação da temperatura animal. No entanto, frequentemente, estes comportamentos podem trazer prejuízos, como a redução no tempo de pastejo, o que afetaria o consumo e consequentemente o desempenho animal. Estudos como estes se tornam mais difíceis em bovinos de corte, pois a obtenção de medidas individuais e diárias de maneira acurada, são onerosas e demanda recursos humanos. Assim, a utilização de novas tecnologias, como acelerômetros, são desejáveis, pois são menos invasivas e podem automatizar a colheita individual diária de dados (Stewart et al., 2017).

2 OBJETIVOS

2.1 Objetivos gerais

Neste trabalho, o objetivo geral foi avaliar os genes diferencialmente expressos e dispersos em bovinos de corte criados à pasto com perfis extremos de atividade e frequência de consumo de água e verificar diferentes estratégias na predição do comportamento destes animais a partir de informações de acelerômetros.

2.2 Objetivos específicos

(1) Estudar o diferencial de expressão e dispersão de genes em animais da raça Nelore e F1 (Nelore x Angus) com perfis extremos de frequência de consumo de água individual diário e atividades durante o período diurno mantidos em pastagens com diferentes níveis de adubação e com diferentes suplementações;

(2) estudar diferentes estratégias na predição do comportamento de animais Nelore criados a pasto, especialmente para atividades de menor frequência, a partir de registros de acelerômetros.

3 REVISÃO DE LITERATURA

3.1 RNAseq e genes relacionados a características de interesse

De acordo com a Organização Meteorológica Mundial (World Meteorological Organization - WMO, 2014), haverá aumento na temperatura média do Brasil de 4º C até o final do século XXI. Esse incremento na temperatura pode afetar diretamente a produção pecuária brasileira, afetando não só o animal, como também as pastagens (Nääs et al., 2010). Segundo a ABIEC (2020), apenas 14% do total de abates no Brasil são resultado de confinamento, portanto grande parte do gado de corte brasileiro depende de pastagens. Sendo assim, a utilização de animais que sejam mais adaptados ao clima tropical brasileiro é fundamental para o progresso produtivo de carne de bovinos no país.

A água é o recurso natural mais importante para produção e sobrevivência, desempenhando papel importante em processos fisiológicos e é responsável por manter a temperatura corporal dos animais (Lardner et al., 2005). Além da água, outro importante fator a ser considerado são os recursos naturais que são utilizados na produção de bovinos de corte. A atividade geral de um animal pode ser um fenótipo a ser considerado, pois um animal mais ativo, também precisa de mais energia para se

manter (Campbell et al., 2008), resultando em maior consumo destes recursos. Portanto, o conhecimento dos genes que influenciem animais que necessitam de maior quantidade de água e que são mais ativos, pode auxiliar a seleção de animais mais sustentáveis e produtivos.

Os avanços nas tecnologias da genética molecular, como técnicas de sequenciamento de alto rendimento, promovem o benefício de se conhecer e estudar genes e vias metabólicas associadas com características de interesse econômico e doenças que podem afetar a produtividade animal (Muers, 2013). Estudos com estas técnicas, relacionando a expressão de genes e consumo alimentar residual (Tizioto et al., 2015), qualidade da carne (Fonseca et al., 2020, Fonseca et al., 2017), eficiência alimentar (Oliveira et al., 2014), estresse térmico (Pires et al., 2019), características reprodutivas (Nascimento et al, 2018) são reportados na literatura, evidenciando a importância do estudo da expressão dos genes para o progresso produtivo da pecuária mundial.

Em análises de diferencial de expressão de genes, a tecnologia de sequenciamento de RNA, ou RNAseq, tem grande destaque. Algumas vantagens dessa técnica são a possibilidade do sequenciamento do RNA sem o conhecimento prévio do mesmo, alto nível de reprodutibilidade de dados, reduzindo a necessidade de grande quantidade de réplicas do experimento e a possibilidade de identificar e quantificar a expressão de isoformas e transcritos desconhecidos (Costa-Silva et al., 2017). Porém o enfoque principal na tecnologia de RNAseq é a análise de diferencial de expressão de genes, que permite o conhecimento de variações fenotípicas entre duas ou mais condições específicas. Nesta análise, os genes expressos nas condições determinadas são quantificados e são identificados aqueles que tem diferença estatística significativa em condições de ambientes distintos, possibilitando maior conhecimento dos mecanismos genéticos dos indivíduos que promovem diferenças fenotípicas observadas (McDermaid et al., 2019).

A divisão de grupos em condições divergentes é relevante para se conhecer genes que influenciem características relacionadas à adaptação dos animais. A eficiência em se perder calor corporal para o ambiente é uma das características observadas em animais adaptados ao estresse térmico. Em fêmeas da raça Nelore, foi observado que as mais eficientes apresentaram maior taxa de sudorese quando

comparado com as menos eficientes em perder calor corporal, enquanto as vacas menos eficientes apresentaram maior temperatura retal e maior amplitude na temperatura vaginal. Outro fator importante para a eficiência da perda de calor corporal é a capacidade da manutenção celular das HSPs (Heat Shock Protein), mesmo em altas temperaturas (Hooper et al., 2019).

Os genes da família HSP são genes expressos em animais que possuem a habilidade de se adaptar de maneira mais eficaz a mudanças de temperatura ambiental. Esses genes podem ser considerados na seleção de animais, quando há temperaturas baixas ou elevadas, para melhor adaptação e produtividade (Kumar et al., 2015). Em animais Nelore e Caracu, foi observado o diferencial de expressão de genes da família HSP, em diferentes condições ambientais de temperatura e umidade, mostrando que esses fatores ambientes têm influência na expressão destes genes (Pires et al., 2019). Neste mesmo estudo, os resultados de parâmetros fisiológicos de temperatura retal e da pele, além do diferencial de expressão dos genes da família HSP, evidenciaram que animais da raça Nelore são mais adaptados ao estresse térmico se comparados com animais da raça Caracu.

Manifestações comportamentais foram observadas no estresse térmico causado por altas temperaturas ambientais, como redução do consumo alimentar, uma vez que este é responsável por grande parte da produção de energia em forma de calor. Um decréscimo de 10 a 35% do consumo alimentar é observado quando os animais estão em ambientes com temperaturas acima de 35º C (Conrad, 1985). O consumo de água é influenciado pela fração de matéria seca ingerida pelos animais e de temperatura ambiental. Esse consumo é maior em animais taurinos se comparado com animais zebuínos (consomem apenas 60% do que é consumido por taurinos). Os animais zebuínos mantem seu consumo em temperaturas ambientais mais elevadas (Colditz and Kellaway ,1972).

A termorregulação é um mecanismo de adaptação às mudanças de temperatura do ambiente, responsável pelo balanço entre a produção e perda de calor, responsável por manter constante a temperatura corporal do animal (Bernabucci et al., 2010). Os mecanismos que controlam a adaptação do animal à temperatura do ambiente diferem entre as raças de bovino de corte. Devido ao maior período de adaptação em climas tropicais, animais zebuínos (*Bos taurus indicus*) são mais

adaptados a mudanças de temperatura e conseguem regular de maneira mais eficaz sua temperatura corporal em resposta ao estresse térmico que os animais taurinos (*Bos taurus taurus*) (Hansen, 2004; Pereira et al., 2014).

Os animais zebuínos mais adaptados a ambientes com temperaturas elevadas, conseguem reduzir a produção de calor interna e aumentar a capacidade de perda de calor para o ambiente. As baixas taxas metabólicas, devido a redução do crescimento e produção de leite, são fatores importantes para a termotolerância se comparados com raças européias (Hansen, 2004). Em estudo com as raças Gir, Girolando, Nelore, Sindi e Indubrasil, observou-se menor adaptabilidade às condições climáticas na raça Gir e melhor índices fisiológicos nas raças Girolando e Sindi, quando expostas à estresse térmico (Cardoso et al., 2015).

3.2 Uso de acelerômetros na pecuária

O comportamento dos animais, se monitorados individualmente e de forma contínua, pode servir como indicativo de bem-estar e saúde dos mesmos (Barwick et al.; 2018). Porém, em grande parte dos trabalhos, a atividade animal é realizada com a observação ou monitoramento por vídeo, os quais dificultam a obtenção dos dados, devido a demanda de recursos humanos e dificuldade de obtenção de imagens de animais a pasto. Portanto, a utilização de sensores que mensurem automaticamente a atividade dos animais possui o potencial de auxiliar na obtenção desta informação em sistemas extensivos, sendo assim de importância para o país, uma vez que grande parte dos animais são mantidos à pasto. Trabalhos que utilizaram acelerômetros a fim de se avaliar a frequência de consumo de água (WILLIAMS et al., 2017; ROLAND et al., 2018) e classificação de comportamento em bovinos de corte, bovinos de leite, ovelhas e suínos (WATANABE et al., 2008; WOLFGER et al., 2015; RAYAS-AMOR et al., 2017) já foram reportados na literatura. Os autores ressaltaram a necessidade de maiores estudos e melhor adaptação de algoritmos para predição mais adequada de comportamento e frequência de consumo de água dos animais.

Dados provenientes de acelerômetros podem oferecer um sistema viável para monitoramento do comportamento animal para posterior avaliação de seu bem-estar. Os dados provenientes dos acelerômetros, usados em modelos de aprendizado de máquina, podem predizer de forma acurada o comportamento do animal e, com isso,

a variação destes comportamentos (Robert et al., 2009). O uso de acelerômetros triaxiais para predição do comportamento animal gera dados de aceleração da movimentação em três eixos: eixo x – movimentos horizontais; eixo y – movimentos longitudinais e eixo z – movimentos verticais. Estes dados precisam de pré-processamento para serem usados como variáveis preditoras nos algoritmos de predição de comportamento (Brown et al., 2013). Na fase de pré-processamento, os dados são editados a fim de se remover frequências de sinal que podem atrapalhar a predição do comportamento animal. Estes então são divididos em grupos de treinamento e validação do modelo de predição. Nos dados de treino, as variáveis respostas de comportamento são consideradas no modelo a fim de se auxiliar o algoritmo a predizer o comportamento e, nos dados de validação, o algoritmo busca predizer o comportamento animal, sem o auxílio das variáveis resposta (Riaboff et al., 2019).

Os comportamentos de ruminação e alimentação também são indicadores de bem-estar animal e estes podem apontar risco de doenças e medidas de produtividade (Diosdado et al., 2015). Estudos com acelerômetros em vacas leiteiras sugerem resultados promissores na detecção destes comportamentos relacionados à alimentação, porém com maior necessidade de ajustes de sensibilidade e futuros estudos para melhor detecção (Zambelis, Wolfe e Vasseur, 2019).

Em vacas leiteiras, o uso de acelerômetros para detecção do estro tem sido amplamente utilizado. A detecção do estro é reportada na literatura de diversas maneiras. Acelerômetros acoplados na perna traseira das vacas registraram as mudanças no número de vezes que os animais estavam em decúbito ventral, em comparação a atividade geral dentro do período pré-parto. Essa mudança no comportamento pode ser indicativa de estro em vacas no período pré-parto (Jensen, 2012). Schirmann et al. (2013) utilizando acelerômetros acoplados em colares nos pescoços de vacas próximas ao parto identificou mudanças no comportamento de ruminação e alimentação com reduções no consumo alimentar e o tempo gasto na ruminação horas antes do momento do parto.

A laminite é uma reposta à dor, que pode ser causada por vários tipos de patologias (Alsaaod et al., 2019). Em vacas leiteiras, a detecção de laminite pode ser feita com a classificação da locomoção das vacas, com o uso de escores (Van Nuffel

et al., 2015) ou apenas pela observação do movimento dos animais no momento da ordenha. Estes métodos de detecção de laminites não são eficazes, devido à falta de regularidade e até mesmo a ausência de qualquer tipo de método para detecção (Dolecheck e Bewley, 2018). O uso de acelerômetros auxilia na detecção de laminite por meio do monitoramento das atividades dos animais. Vacas que apresentam laminite, tendem a ficar mais tempo em decúbito ventral em casos mais graves, auxiliando a detecção desta doença (Thorup et al., 2015).

O comportamento de sucção no momento da amamentação em bezerros de corte tem papel importante no rápido desenvolvimento de animais saudáveis (Kour et al., 2018) e tem influência direta no ganho de peso médio diário, pois o leite recebido de sua progenitora é uma das únicas fontes de nutrientes (Valros et al., 2002). Acelerômetros são oriundos de tecnologia de baixo custo e eficiente para identificação deste comportamento em bezerros de corte criados à pasto, além do tempo e duração da amamentação de cada bezerro. Com o uso desta tecnologia, os produtores podem planejar de maneira eficaz o manejo destes animais, para otimizar o ganho de peso dos bezerros e determinar o período ideal de desmame dos bezerros, levando a fêmea a retornar mais rápido para produção (Kour et al., 2018). Sendo assim, a utilização de acelerômetros na pecuária tem se mostrado como ferramenta importante na detecção de doenças, fertilidade e bem-estar dos animais, auxiliando a produtividade animal.

3.2.1 Algoritmos de classificação a partir de dados de acelerômetros

Os métodos de aprendizado de máquina têm sido utilizados juntamente com os dados de acelerômetros para predição do comportamento animal. Os algoritmos de aprendizado de máquina podem criar um modelo de predição a partir de informações prévias, sem a necessidade de conhecimento prévio das relações entre as variáveis preditoras e alvos (Kotsiantis et al., 2007). Esses métodos também possibilitam a utilização de variáveis que não apresentam distribuição normal (Gianola et al., 2011). Dentre os algoritmos reportados na literatura, destaca-se o “Random Forest” (RF) na predição e classificação do comportamento animal (Kleanthous et al., 2018). Este algoritmo consiste na combinação de várias árvores de decisão independentes para uma predição mais acurada e estável. A combinação de diferentes árvores de decisão para obtenção de um único resultado consiste na criação de vários modelos diferentes

a partir do algoritmo implementado. Cada árvore de decisão recebe uma condição e, caso essa condição seja atendida, o fluxo é continuado até a próxima árvore, até a finalização da análise (Breinman, 2001). Se comparado com métodos de RF e árvores de decisão, o algoritmo de RF mostra-se superior em acurácia na classificação e predição do comportamento de touros criados à pasto (Abell et al., 2017).

Outro método utilizado é o “Naive Bayes Classifier” (NBC) é um modelo probabilístico simples baseado na aplicação do teorema de Bayes para classificação de variáveis preditas a partir de probabilidades *a priori*. Este método parte do princípio da independência entre as variáveis preditoras e contribui de forma independente na probabilidade de predição das variáveis respostas (Jensen, 1996). Apesar de menor robustez, este algoritmo apresenta desempenho competitivo com algoritmos mais complexos (Domingos e Pazzani, 1997). Em comparação com o algoritmo de RF, NBC mostrou-se similar na análise de sobrevivência de vacas Holstein para a segunda lactação (Van der Heide, et al., 2019). Porém, em comparação com outro algoritmo amplamente utilizado em classificação de comportamento animal, o “Support Vector Machine” (SVM), o NBC mostrou-se inferior em acurácia na predição do comportamento de decúbito ventral, ficar em pé e comer em estudo com vacas leiteiras da raça Holstein (Benaissa et al., 2017). Zehner et al. (2019) ao estudar predição de parição e baseado no comportamento digestivo em vacas leiteiras, observaram que o algoritmo NBC não se mostrou adequado para aplicação na prática, porém as variáveis preditoras de comportamento de ruminação tiveram valores preditos consideráveis.

Dificuldades de classificação de variáveis recentes têm sido abordados com o método de SVM (Vapnik, 1995). Este método estatístico de aprendizagem pode ser aplicado em variáveis classificatórias e regressão. O algoritmo de SVM transforma as variáveis preditoras em um espaço de recursos de alta dimensão. Após isso, um hiperplano é construído entre duas classes de pontos de dados, para que a margem que os separa, seja maximizada. A melhor capacidade de generalização do classificador é devido ao princípio de minimização de risco estrutural presente no SVM, quando são utilizados dados desconhecidos. Este algoritmo tem sido um dos métodos de aprendizado de máquina mais utilizado pois, tem bom desempenho em aplicações reais de reconhecimento de padrões, classificação de imagens e

categorização de texto (Burges, 1998; Vert, 2007). Atividades como andar, dormir, deitar, ficar em pé, ruminar e se alimentar utilizando o SVM para predição destas atividades, com dados de acelerômetros em bovinos são reportados na literatura, apontando a necessidade de futuros estudos (Martiskainen et al., 2009; Hokkanen et al., 2011).

Em algoritmos de classificação de variáveis, uma dificuldade que pode levar ao viés na habilidade de predição do modelo, é o desbalanceamento de dados, em que as classes majoritárias podem ser favorecidas, se comparadas com as classes com menor quantidade de observações (Chen et. al., 2004). Devido ao curto período de tempo e intermitência que bovinos bebem água (Delagarde e Lamberton, 2015; Dutta et al., 2015), este comportamento tende a ser menos frequente que outros comportamentos observados, levando a necessidade de balanceamento de dados para predições mais acuradas em algoritmos de classificação (Roland et al., 2018). Alguns métodos de reamostragem de dados são utilizados para lidar com dados desbalanceados. Estes métodos são utilizados para reequilibrar a quantidade de observações em cada classe, a fim de se minimizar o efeito da distribuição distorcida de classes no processo de aprendizagem do algoritmo (López et al., 2013). Alguns métodos de reamostragem são mais comumente usados como o “over-sampling”, em que é criado criada maior quantidade de amostras para classes de minoria de dados (Chawla et al., 2002), e o método de “under-sampling” em que é descartado amostras das classes majoritárias, afim de se equilibrar as observações das classes (Tahir et al., 2009).

4 REFERÊNCIAS

ABELL, K. M.; THEURER, M. E.; LARSON, R. L.; WHITE, B. J.; HARDIN, D. K.; RANDLE, R. F. Predicting bull behavior events in a multiple-sire pasture with video analysis, accelerometers, and classification algorithms. **Computers and Electronics in Agriculture**, v. 136, p. 221-227, 2017.

ALSAAOD, M.; FADUL, M.; STEINER, A. Automatic lameness detection in cattle. **The veterinary journal**, v. 246, p. 35-44, 2019.

ASSOCIAÇÃO BRASILEIRA DE INDÚSTRIAS EXPORTADORAS DE CARNE (ABIEC), 2020. Beef Report. Disponível em: <http://abiec.com.br/publicacoes/beef-report-2020/>. Acesso em: 10 de julho de 2021.

ASSOCIAÇÃO BRASILEIRA DE INSEMINAÇÃO ARTIFICIAL – ASBIA. Index ASBIA 2014. Disponível em <http://www.asbia.com.br/index> Acesso 02 jun. 2019.

BARWICK, J.; LAMB, D. W.; DOBOS, R.; WELCH, M.; TROTTER, M. Categorising sheep activity using a tri-axial accelerometer. **Computers and Electronics in Agriculture**, v. 145, p. 289-297, 2018.

BENAISSA, S., TUYTTENS, F., PLETS, D., DE PESSEMIER, T., TROGH, J., TANGHE, E., ... & SONCK, B. Behaviours recognition using neck-mounted accelerometers in dairy barns. **In 8th European Conference on Precision Livestock Farming** (EC-PLF 2017) (pp. 69-76), 2017.

BREIMAN, L. Random forests. **Machine learning**, v. 45(1), p. 5-32, 2001.

BROWN, D. D.; KAYS, R.; WIKELSKI, M.; WILSON, R.; KLIMLEY, A. P. Observing the unwatchable through acceleration logging of animal behavior. **Animal Biotelemetry**, v. 1(1), p. 1-16, 2013.

BURGES, C. J. A tutorial on support vector machines for pattern recognition. **Data mining and knowledge discovery**, v. 2(2), p. 121-167, 1998.

CAMPBELL, N. A.; REECE, J. B.; URRY, L. A.; CAIN, M. L.; WASSERMAN, S. A.; MINORSKY, P. V.; JACKSON, R. B. "Energy Requirements Are Related to Animal Size, Activity, and Environment," **in Campbell Biology**, 8th ed. (San Francisco: Benjamin Cummings, 2008), 870 P., 2008.

CÁNOVAS, A.; RINCON, G.; ISLAS-TREJO, A.; WICKRAMASINGHE, S.; MEDRANO, J. F. SNP discovery in the bovine milk transcriptome using RNA-Seq technology. **Mammalian genome**, v. 21(11), p. 592-598, 2010.

CARDOSO, C. C.; PERIPOLLI, V.; AMADOR, S. A.; BRANDÃO, E. G.; ESTEVES, G. I. F.; SOUSA, C. M. Z.; ... & MCMANUS, C. Physiological and thermographic response to heat stress in zebu cattle. **Livestock Science**, v. 182, p. 83-92, 2015.

CHAWLA, N. V.; BOWYER, K. W.; HALL, L. O.; KEGELMEYER, W. P. SMOTE: synthetic minority over-sampling technique. **Journal of artificial intelligence research**, v. 16, p. 321-357, 2002.

CHEN, C.; LIAW, A.; BREIMAN, L Using random forest to learn imbalanced data. **University of California, Berkeley**, v. 110(1-12), p. 24, 2004.

COLDITZ, P. J.; KELLAWAY, R. C. The effect of diet and heat stress on feed intake, growth, and nitrogen metabolism in Friesian, F1 Brahman× Friesian, and Brahman heifers. **Australian Journal of Agricultural Research**, v. 23(4), p.717-725, 1972.

COLLIER, R. J.; GEBREMEDHIN, K. G. Thermal biology of domestic animals. **Annual Review of Animal Biosciences**, v. 3, n. 1, p. 513-532, 2015.

CONRAD, J. H. (1985). Feeding of farm animals in hot and cold environments. In: M.K. Yousef (Editor), **Stress Physiology in Livestock**, CRC Press, Boca Raton, FL, v.2, p205-226, 1985.

COSTA-SILVA, J.; DOMINGUES, D.; LOPES, F. M. RNA-Seq differential expression analysis: An extended review and a software tool. **PloS one**, v. 12(12), p. e0190152, 2017.

DE OLIVEIRA, P. S.; CESAR, A. S.; DO NASCIMENTO, M. L.; CHAVES, A. S.; TIZIOTO, P. C.; TULLIO, R. R.; ... & REGITANO, L. C. Identification of genomic regions associated with feed efficiency in Nelore cattle. **BMC genetics**, v. 15(1), p. 1-10, 2014.

DELAGARDE, R.; LAMBERTON, P. Daily grazing time of dairy cows is recorded accurately using the Lifecorder Plus device. **Applied Animal Behaviour Science**, v. 165, p. 25-32, 2015.

DIOSDADO, J. A. V.; BARKER, Z. E.; HODGES, H. R.; AMORY, J. R.; CROFT, D. P.; BELL, N. J.; CODLING, E. A. Classification of behaviour in housed dairy cows using an accelerometer-based activity monitoring system. **Animal Biotelemetry**, v. 3(1), p. 1-14, 2015

DOLECHECK, K., BEWLEY, J. Animal board invited review: Dairy cow lameness expenditures, losses and total cost. **Animal**, v. 12(7), p. 1462-1474, 2018.

DOMINGOS, P.; PAZZANI, M. On the optimality of the simple Bayesian classifier under zero-one loss. **Machine Learning**, v. 29, p. 103–130, 1997.

DUTTA, R.; SMITH, D.; RAWNSLEY, R.; BISHOP-HURLEY, G.; HILLS, J.; TIMMS, G.; HENRY, D. Dynamic cattle behavioural classification using supervised ensemble classifiers. **Computers and electronics in agriculture**, v. 111, p. 18-28, 2105.

EBERHARDT, B. G.; SATRAPA, R. A.; CAPINZAIKI, C. R.; TRINCA, L. A.; BARROS, C. M. Influence of the breed of bull (Bos taurus indicus vs. Bos taurus taurus) and the breed of cow (Bos taurus indicus, Bos taurus taurus and crossbred) on the resistance of bovine embryos to heat. **Animal reproduction science**, v. 114(1-3), p. 54-61, 2009.

FONSECA, L. F. S.; DOS SANTOS SILVA, D. B.; GIMENEZ, D. F. J.; BALDI, F.; FERRO, J. A.; CHARDULO, L. A. L.; DE ALBUQUERQUE, L. G. Gene expression profiling and identification of hub genes in Nellore cattle with different marbling score levels. **Genomics**, v. 112(1), p. 873-879, 2020.

FONSECA, L. F. S.; GIMENEZ, D. F. J.; DOS SANTOS SILVA, D. B.; BARTHELSON, R.; BALDI, F.; FERRO, J. A.; & ALBUQUERQUE, L. G. Differences in global gene expression in muscle tissue of Nellore cattle with divergent meat tenderness. **BMC genomics**, v. 18(1), p. 1-12, 2017.

GIANOLA, D.; OKUT, H.; WEIGEL, K. A.; ROSA, G. J. Predicting complex quantitative traits with Bayesian neural networks: a case study with Jersey cows and wheat. **BMC genetics**, v. 12(1), p. 1-14, 2011.

HANSEN, P. J. Physiological and cellular adaptations of zebu cattle to thermal stress. **Animal reproduction science**, v. 82, p. 349-360, 2004.

HAYES, B. J.; LEWIN, H. A.; GODDARD, M. E. The future of livestock breeding: genomic selection for efficiency, reduced emissions intensity, and adaptation. **Trends in Genetics**, v. 29, p. 206-214, 2013.

HOKKANEN, A. H.; HÄNNINEN, L.; TIUSANEN, J.; PASTELL, M. Predicting sleep and lying time of calves with a support vector machine classifier using accelerometer data. **Applied Animal Behaviour Science**, v. 134(1-2), p. 10-15, 2011.

HOOPER, H. B.; TITTO, C. G.; GONELLA-DIAZA, A. M.; HENRIQUE, F. L.; PULIDO-RODRÍGUEZ, L. F.; LONGO, A. L. S.; ... & TITTO, E. A. L. Heat loss efficiency and HSPs gene expression of Nellore cows in tropical climate conditions. **International journal of biometeorology**, v. 63(11), p. 1475-1486, 2019.

JENSEN, F. An Introduction to Bayesian Networks. **Springer**, New York, 1996.

JENSEN, M. B. Behaviour around the time of calving in dairy cows. **Applied animal behaviour science**, v. 139(3-4), p. 195-202, 2012.

KLEANTHOUS, N.; HUSSAIN, A.; MASON, A.; SNEDDON, J.; SHAW, A.; FERGUS, P.; ... & AL-JUMEILY, D. Machine learning techniques for classification of livestock behavior. In **International Conference on Neural Information Processing** (p. 304-315). Springer, Cham, 2018.

KOTSIANTIS, S. B.; ZAHARAKIS, I.; PINTELAS, P. Supervised machine learning: A review of classification techniques. **Emerging artificial intelligence applications in computer engineering**, v. 160(1), p. 3-24, 2007.

KOUR, H., PATISON, K. P., CORBET, N. J., & SWAIN, D. L. (2018). Validation of accelerometer use to measure suckling behaviour in Northern Australian beef calves. Applied Animal Behaviour Science, v. 202, p. 1-6, 2018.

KUMAR, R. R.; GOSWAMI, S.; SHARMA, S. K.; KALA, Y. K.; RAI, G. K.; MISHRA, D. C.; ... & RAI, R. D. Harnessing next generation sequencing in climate change: RNA-Seq analysis of heat stress-responsive genes in wheat (Triticum aestivum L.). **Omics: a journal of integrative biology**, v. 19(10), p. 632-647, 2015.

LARDNER, H. A.; KIRYCHUK, B. D.; BRAUL, L.; WILLMS, W. D.; YAROTSKI, J. The effect of water quality on cattle performance on pasture. **Australian Journal of Agricultural Research**, v. 56(1), p. 97-104, 2005.

LEES, A. M.; LEES, J. C.; LISLE, A. T.; SULLIVAN, M. L.; GAUGHAN, J. B. Effect of heat stress on rumen temperature of three breeds of cattle. **International Journal of Biometeorology**, v. 62(2), p. 207-215, 2018.

LÓPEZ, V.; FERNÁNDEZ, A.; GARCÍA, S.; PALADE, V.; HERRERA, F. An insight into classification with imbalanced data: Empirical results and current trends on using data intrinsic characteristics. **Information sciences**, v. 250, p. 113-141, 2013.

MARGUERAT, S.; BÄHLER, J. RNA-seq: from technology to biology. **Cellular and molecular life sciences**, v. 67, n. 4, p. 569-579, 2010.

MARTISKAINEN, P.; JÄRVINEN, M.; SKÖN, J. P.; TIIRIKAINEN, J.; KOLEHMAINEN, M.; MONONEN, J. Cow behaviour pattern recognition using a three-dimensional accelerometer and support vector machines. **Applied animal behaviour science**, v. 119(1-2), p. 32-38, 2009.

MCDERMAID, A.; MONIER, B.; ZHAO, J.; LIU, B.; MA, Q. Interpretation of differential gene expression results of RNA-seq data: review and integration. **Briefings in bioinformatics**, v. 20(6), p. 2044-2054, 2019.

MILLEN, D. D.; ARRIGONI, M. De B.; Drivers of change in animal protein production systems: Changes from 'traditional'to 'modern'beef cattle production systems in Brazil. **Animal Frontiers**, v. 3, n. 3, p. 56-60, 2013.

MUERS, M. Sequencing for disease architecture. **Nature Reviews Genetics**, v. 14(8), p. 518-518, 2013.

MURPHY, E.; DE BOER, I. J. M.; VAN MIDDELAAR, C. E.; HOLDEN, N. M.; SHALLOO, L.; CURRAN, T. P.; UPTON, J. Water footprinting of dairy farming in Ireland. **Journal of cleaner production**, v. 140, p. 547-555, 2017.

NÄÄS, I. D. A.; ROMANINI, C. E. B.; SALGADO, D. D. A.; LIMA, K. A. O.; VALE, M. M. D.; LABIGALINI, M. R.; ... & MOURA, D. J. D. Impact of global warming on beef cattle production cost in Brazil. **Scientia Agricola**, v. 67, p. 01-08, 2010.

NADEAU, C. P.; URBAN, M. C.; BRIDLE, J. R. Climates past, present, and yet-to-come shape climate change vulnerabilities. **Trends in Ecology & Evolution**, v. 32, n. 10, p. 786-800, 2017.

NASCIMENTO, A. V. D.; ROMERO, Â. R. D. S.; UTSUNOMIYA, Y. T.; UTSUNOMIYA, A. T. H.; CARDOSO, D. F.; NEVES, H. H. R.; ... & GRISOLIA, A. B. Genome-wide association study using haplotype alleles for the evaluation of reproductive traits in Nelore cattle. **PloS one**, v. 13(8), p. e0201876, 2018.

PEREIRA, A. M. F.; TITTO, E. L.; INFANTE, P.; TITTO, C. G.; GERALDO, A. M.; ALVES, A.; ... & ALMEIDA, J. A. Evaporative heat loss in Bos taurus: Do different cattle breeds cope with heat stress in the same way? **Journal of Thermal Biology**, v. 45, p. 87-95, 2014.

PIRES, B. V.; STAFUZZA, N. B.; LIMA, S. B. G. P. N. P.; NEGRÃO, J. A.; PAZ, C. C. P. Differential expression of heat shock protein genes associated with heat stress in Nelore and Caracu beef cattle. **Livestock Science**, v. 230, p. 103839, 2019.

RAYAS-AMOR, A. A.; MORALES-ALMARÁZ, E.; LICONA-VELÁZQUEZ, G.; VIEYRA-ALBERTO, R.; GARCÍA-MARTÍNEZ, A.; MARTÍNEZ-GARCÍA, C. G.; ... & MIRANDA-DE LA LAMA, G. C. Triaxial accelerometers for recording grazing and ruminating time in dairy cows: An alternative to visual observations. **Journal of Veterinary Behavior**, v. 20, p. 102-108, 2017.

RIABOFF, L.; AUBIN, S.; BEDERE, N.; COUVREUR, S.; MADOUASSE, A.; GOUMAND, E.; ... & PLANTIER, G. Evaluation of pre-processing methods for the prediction of cattle behaviour from accelerometer data. **Computers and Electronics in Agriculture**, v. 165, p. 104961, 2019.

ROBERT, B.; WHITE, B. J.; RENTER, D. G.; LARSON, R. L. Evaluation of three-dimensional accelerometers to monitor and classify behavior patterns in cattle. **Computers and Electronics in Agriculture**, v. 67(1-2), p. 80-84, 2009

ROLAND, L.; SCHWEINZER, V.; KANZ, P.; SATTLECKER, G.; KICKINGER, F.; LIDAUER, L.; ... & IWERSEN, M. Evaluation of a triaxial accelerometer for monitoring selected behaviors in dairy calves. **Journal of dairy science**, v. 101(11), p. 10421-10427, 2018.

SCHIRMANN, K.; CHAPINAL, N.; WEARY, D. M.; VICKERS, L.; VON KEYSERLINGK, M. A. G. Rumination and feeding behavior before and after calving in dairy cows. **Journal of dairy science**, v. 96(11), p. 7088-7092, 2013.

STEWART, M.; WILSON, M. T.; SCHAEFER, A. L.; HUDDART, F.; SUTHERLAND, M. A. The use of infrared thermography and accelerometers for remote monitoring of dairy cow health and welfare. **Journal of dairy science**, v. 100(5), p. 3893-3901, 2017.

TAHIR, M. A.; KITTLER, J.; MIKOLAJCZYK, K.; YAN, F. A multiple expert approach to the class imbalance problem using inverse random under sampling. **In International workshop on multiple classifier systems** (pp. 82-91). Springer, Berlin, Heidelberg, 2009.

THORUP, V. M.; MUNKSGAARD, L.; ROBERT, P. E.; ERHARD, H. W.; THOMSEN, P. T.; FRIGGENS, N. C. Lameness detection via leg-mounted accelerometers on dairy cows on four commercial farms. **Animal**, v. 9(10), p. 1704-1712, 2015.

TIZIOTO, P. C.; COUTINHO, L. L.; DECKER, J. E.; SCHNABEL, R. D.; ROSA, K. O.; OLIVEIRA, P. S.; ... & REGITANO, L. C. Global liver gene expression differences in Nelore steers with divergent residual feed intake phenotypes. **BMC genomics**, v. 16(1), p. 1-14, 2015.

VALROS, A. E.; RUNDGREN, M.; ŠPINKA, M.; SALONIEMI, H.; RYDHMER, L.; ALGERS, B. Nursing behaviour of sows during 5 weeks lactation and effects on piglet growth. **Applied Animal Behaviour Science**, v. 76(2), p. 93-104, 2002.

VAN DER HEIDE, E. M. M.; VEERKAMP, R. F.; VAN PELT, M. L.; KAMPHUIS, C.; ATHANASIADIS, I.; DUCRO, B. J. Comparing regression, naive Bayes, and random forest methods in the prediction of individual survival to second lactation in Holstein cattle. **Journal of dairy science**, v. 102(10), p. 9409-9421, 2019.

VAN NUFFEL, A.; ZWERTVAEGHER, I.; PLUYM, L.; VAN WEYENBERG, S.; THORUP, V. M.; PASTELL, M.; ... & SAEYS, W. Lameness detection in dairy cows: Part 1. How to distinguish between non-lame and lame cows based on differences in locomotion or behavior. **Animals**, v. 5(3), p. 838-860, 2015.

VAPNIK, V.; GUYON, I.; HASTIE, T. Support vector machines. **Machine Learning**, v. 20(3), p. 273-297, 1995.

VERT, J. P. Kernel methods in genomics and computational biology. **In Kernel methods in bioengineering, signal and image processing** (pp. 42-63). IGI Global, 2007.

WATANABE, N.; SAKANOUE, S.; KAWAMURA, K.; KOZAKAI, T. Development of an automatic classification system for eating, ruminating and resting behavior of cattle using an accelerometer. **Grassland science**, v. 54(4), p. 231-237, 2008

WILLIAMS, L. R.; BISHOP-HURLEY, G. J.; ANDERSON, A. E.; SWAIN, D. L. Application of accelerometers to record drinking behaviour of beef cattle. **Animal Production Science**, v. 59(1), p. 122-132, 2017.

WOLFGER, B.; TIMSIT, E.; PAJOR, E. A.; COOK, N.; BARKEMA, H. W.; ORSEL, K. Accuracy of an ear tag-attached accelerometer to monitor rumination and feeding behavior in feedlot cattle. **Journal of animal science**, v. 93(6), p. 3164-3168, 2015.

ZAMBELIS, A., WOLFE, T., VASSEUR, E. Validation of an ear-tag accelerometer to identify feeding and activity behaviors of tiestall-housed dairy cattle. **Journal of dairy science**, v. 102(5), p. 4536-4540, 2019.

ZEHNER, N.; NIEDERHAUSER, J. J.; SCHICK, M.; UMSTATTER, C. Development and validation of a predictive model for calving time based on sensor measurements of ingestive behavior in dairy cows. **Computers and Electronics in Agriculture**, v. 161, p. 62-71, 2019.

Capítulo 2 - Gene expression and dispersion profile associated with water consumption frequency and general activity in grazing beef cattle

Abstract

Selecting animals able for increasing the sustainability of the production system and less degrading the environment is an essential part of the implementation of programs to intensify sustainable livestock. Knowing genes that influence beef cattle water consumption frequency and general activity in grazing can help the animal selection and genetic constitution knowledge about these traits. Therefore, the aim of this work was to verify the differential gene expression (DE) and dispersion (DD) profile associated with water consumption frequency (WCF) and general activity (GA) in grazing beef cattle. The animals' behavior was monitored and recorded for 12 hours daily for four days. For RNA sequence, 10 extreme animals (five animals with greater frequency and five with less frequency) were selected for WCF and GA for Nelore and F1 (Nelore x Angus) animals. The functional and enrichment analysis showed, for DE analysis, a total of ten and three distinctive genes related to biological process, cellular components and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways for Nelore water consumption frequency (NWCF) and Nelore general activity (NGA), respectively. No enriched genes were found for F1 groups. For the DD analysis, three and two distinctive genes were enriched, related to cellular components for NWCF and NGA, respectively. In the F1 animals' water consumption frequency (F1WCF) and F1 animals' general activity (F1GA) groups, four and five distinctive genes related to biological process, cellular components, and KEGG pathways, respectively, were found. Genes found in this study related to water consumption frequency and general activity were associated with several important traits for the adaptive capacity of animals. These genes can elucidate and help to select animals that are more adapted to challenging environments and degrade less the environment.

Keywords

Adaptation, Angus, functional annotation, enrichment analysis, Nelore

Implications

Forecasts of water and resource scarcity worldwide must be taken into consideration regarding beef cattle production. Knowledge of genes that influence the water consumption and energy resources by farm animals can help to select animals that are more adapted to future sustainability needs, promoting environments that are less degraded by production. Genes found in this study were associated with several important traits for the adaptive capacity of animals. These genes can elucidate and help to select animals that are more adapted to challenging environments and which degrade less the environment.

1 Introduction

Water is the most important factor for the survival and production of livestock animals. It plays a fundamental role in all physiological processes, directly or indirectly, in body temperature maintenance, and it is part of 50 to 80% of the animal's liveweight (Lardner et al., 2005). Livestock must reconcile its global food production without damaging the environment (Godfray et al., 2010), since global water scarcity is a threat to human society development (Mekonnen and Hoekstra, 2016). Another important point to be taken into consideration regarding the preservation of the environment is the use of environmental resources in the beef cattle production. For this, the measurement of the general activity of an animal should be considered, since a more active animal needs more energy to maintain itself (Reece et al., 2009), resulting in greater consumption of natural resources.

Knowing the genes that influence animals that need more water and are more active can help to understand the physiology of these animals, as well as help in the selection of more sustainable and productive animals. RNA sequencing (RNA-Seq) is a method for analyzing gene expression profiling through the full sequencing of a whole transcriptome, resulting in more useful information. When the expression patterns of two physiological conditions or stages are compared, differences in the magnitudes of average gene expression (differential gene expression, DE) and gene expression variance (differential gene dispersion, DD) can be observed.

The typical analysis of RNA-Seq data focuses on the finding of genes that show DE between groups (Zarek et al., 2017; Pires et al., 2019). However, gene expression is naturally a stochastic process, and in some cases its variation could be significantly

influenced by the animals' physiological state (Jong et al., 2019) and may represent a new source of information about gene regulation as important as changes in mean levels of gene expression observed in differential expression analysis (Wang and Zhang, 2011).

Studies in cattle regarding water consumption and its importance (Lardner et al., 2013; Ahlberg et al., 2018; Wagner and Engle, 2021) and genes related to heat tolerance and water consumption (Beatty et al., 2006), are reported in the literature. However, there is a need to study gene expression profiles associated with water consumption and with the general activity of animals and genes associated with this behavior. Therefore, the aim of this work was to verify the differential gene expression (DE) and dispersion (DD) profile associated with water consumption frequency (WCF) and general activity (GA) in grazing beef cattle.

2 Material and methods

2.1 Experimental area

The experiment was carried out in the Forage Crops and Grasslands section of Universidade Estadual Paulista (Unesp), Jaboticabal, São Paulo, Brazil. The total area is divided into 24 paddocks, seeded with Urochloa brizantha (Hochs tex A. Rich) Stapf cv. Marandu (Marandu grass) in 2001. The grazing mean height was 25 cm, using the continuous grazing method with a variable stocking rate. The region's climate is humid subtropical, with dry winters and rainy summers. The average annual temperature is 22.3°C, with a maximum average of 29.1°C and a minimum average of 16.9°C. The annual average relative humidity is 71.2% and the wettest quarter of the year is the first one (January, February, and March), with an average rainfall of 628.8 mm, equivalent to 44.2% of the rainfall annual total.

2.2 Behavioral observation and selection of extreme animals

In this experiment, 36 Nelore and 33 F1 (Nelore x Angus) animals were used, divided into the 24 paddocks of the experimental area used in this study. The animals' behaviors were monitored for four days daily (March, 4-7th, 2020, maximum average temperature of 31.5ºC and minimum average of 17.25ºC), 12 hours per day (6 AM to 6 PM). The animals were identified with ribbons of different colors and separated by paddocks. The behavior observation (i.e., grazing, rumination, standing idle, lying idle, frequency of the animal in the trough and walking) was evaluated by direct visual

observation (Hughes and Reid, 1951) undertaken every 10 min (Martín and Bateson, 1986). The water consumption frequency (WCF), as indicative of water consumption (Perissinotto et al., 2005; Wagner and Engle, 2021), was performed by recording all the moments when the identified animal went to the drinking fountain, thus recording the hour and minute at that moment. For subsequent analyses, the relative frequency of WCF was used, calculated for each animal, as the ratio between the total number of times the animal went to the drinking fountain by the sum of the other behaviors (noted every ten minutes plus every time the animal went to the drinking fountain) observed. The general activity (GA) of each animal was analyzed by its relative frequency, which is the ratio between the sum of all behaviors excluding idle behaviors and the sum of all behaviors.

Extremes for WCF and GA were selected (N = 10 for each phenotype and for each genetic composition) and the statistical difference between the extreme groups (all the analyses considered the low frequency groups versus high frequency groups) for each genetic composition (Nelore for WCF (NWCF), Nelore for GA (NGA), F1 for WCF (F1WCF) and F1 for GA (F1GA)) was verified by analysis of variance using the least squares method. The assumptions of residuals' normality and homogeneity of variances were verified. It was verified whether there was a statistical difference (p-value < 0.05) in the average daily weight gain in the extreme groups by the Student's t-test.

2.3 RNA-seq

Blood samples from the jugular vein puncture were collected from the 69 fasting animals observed, after the behavioral observation period (8 and 9 March 2020). The blood samples were stored in RNA storage tubes (PAXgene®) under refrigeration at -80°C. The total RNA of each of the samples was extracted at the Laboratory of EMBRAPA Pecuária Sudeste with the "Blood RNA Kit" (PAXgene®), following the manufacturer's recommendations. The isolated total RNA was quantified by spectrophotometry (NanoDrop®, Wilmington, DE, USA) and its quality evaluated by Bioanalyzer 2100 (Agilent, Palo Alto, CA, USA). All samples had RNA Integrity Number (RIN) values greater than 8.0. Libraries were prepared using the TruSeq SBS Kit v3-HS (200 cycles) sequenced on HiScanSQ (Illumina, San Diego, CA, USA). These

procedures were performed at the Functional Genomics Center (USP, Campus “Luiz de Queiroz”, SP, Brazil).

2.4 Bioinformatic Analysis

The quality of the sequences was verified with the FastQC software (Andrews, 2010). The removal of low Phred quality score and adapters of the reads were made using the Trimmomatic software (Bolger et al., 2014). Reads with nucleotide quality scores less than three and read lengths less than 30 bp were discarded. The sequences were aligned with the bovine reference genome version ARS-UCD 1.2 with the HiSat2 algorithm (Kim et al., 2015) by using default parameters. The counting of readings was performed by the FeatureCounts package (Liao et al., 2014) of the R software. Normalization of readings and analysis of differential expression (DE) of genes was performed by the DESeq2 package (Love et al., 2014) of the R software. The differential dispersion analysis (DD) was performed by the MDseq R software package (Ran and Daye, 2017), genes with an adjusted p-value < 0.1 were considered as differentially expressed and dispersed genes. Functional and enriched analyzes were performed by the DAVID Bioinformatic Resources 6.8 platform (Huang et al., 2008), and pathways with p-value <0.05 were considered enriched.

3 Results

The analysis of variance showed that there was a statistically significant difference (p-value < 0.05) for phenotype calculated frequencies between the selected extreme groups. Student's t statistical test showed that there was not statistically significant (p-value > 0.05) difference between the average daily weight gain between the extreme studied groups in each breed.

A total of 27 and 15 DE (Appendix 1), and 25 and 23 DD (Appendix 2) genes were found significantly (p-value <0.1) differentially expressed between extremes for NWCF and NGA, respectively. In the F1 animals, a total of five DE genes were observed for both F1WCF and F1GA groups. In the DD analysis, 21 and 41 genes were found significantly (p-value <0.1) differentially expressed for F1WCF andF1GA genes, respectively.

The Functional and enrichment analysis identified for DE analysis (Table 1), a total of ten and three distinctive genes related to biological processes, cellular components and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways, for

NWCF and NGA, respectively. No enriched genes were found for F1 groups. For the DD analysis (Table 2), three and two distinctive genes were enriched, related to cellular components for NWCF and NGA, respectively. In the F1WCF and F1GA groups, four and five distinctive genes were found, related to biological process, cellular components and KEGG pathways, respectively.

Table 1. Functional annotation genes description, Gene Ontology (GO) term and main pathway enriched Kegg for the differential expression analysis in Nelore and F1 groups

Access Number	Term	N	P-value	Bonferroni	Benjamini	FDR	Genes
NWCF - Up-regulation							
Biological Process							
GO:0010608	Posttranscriptional regulation of gene expression	2	3.2E-3	7.9E-2	8.3E-2	8.3E-2	LOC52033; GCNT2
GO:0010718	Positive regulation of epithelial to mesenchymal transition	2	8.1E-3	1.9E-1	1.1E-1	1.1E-1	LOC52033; GCNT2
GO:0007179	Transforming growth Factor beta receptor signaling pathway	2	1.5E-2	3.2E-1	1.1E-1	1.1E-1	LOC52033; GCNT2
GO:0051897	Positive regulation of protein kinase B signaling	2	1.7E-2	3.6E-1	1.1E-1	1.1E-1	LOC52033; GCNT2
GO:0006486	Protein glycosylation	2	2.2E-2	4.4E-1	1.1E-1	1.1E-1	LOC52033; GCNT2
GO:0070374	Positive regulation of ERK1 and ERK2 cascade	2	3.7E-2	6.3E-1	1.6E-1	1.6E-1	LOC52033; GCNT2
Cellular Component							
GO:0016021	Integral component of membrane	7	6.5E-3	8.7E-2	9.1E-2	9.1E-2	ATP6V0A4; CD96; LOC520336; GCNT2; KLRD1; NCR1; TMEM242
Molecular Function							
GO:0008109	N-acetyllactosaminide beta-1,6-N-acetylglucosaminyltransferase activity	2	1.1E-3	9.1E-3	9.2E-3	9.2E-3	LOC52033; GCNT2
Kegg							
	Natural killer cell mediated cytotoxicity	2	4.6E-2	4.0E-1	5.0E-1	5.0E-1	KLRD1; NCR1
NWCF - Down-regulation							
Biological Process							
GO:0006955	Immune response	2	4.6E-2	9.0E-1	1.0E0	1.0E0	VPREB1; THBS1
NGA - Up-regulation							
Cellular Component							
GO:0005887	Integral component of plasma membrane	3	6.4E-2	7.3E-1	1.0E0	1.0E0	LOC50686; SLC16A6; STOM

Table 2. Functional annotation genes description. Gene Ontology (GO) term and main pathway enriched Kegg for the differential dispersion analysis in Nelore and F1 groups

Access Number	Term	N	P-value	Bonferroni	Benjamini	FDR	Genes
NWCF - Up-regulation							
Cellular Component							
GO:0005654	Nucleoplasm	3	6.0E-2	6.1E-1	9.1E-1	9.1E-1	XRCC6; CDCA7; ZC3H14
NGA - Down-regulation							
Cellular Component							
GO:0071439	Clathrin complex	2	3.1E-3	1.3E-1	1.4E-1	1.4E-1	VPS41; CLTA
F1WCF - Up-regulation							
Biological Process							
GO:0061028	Establishment of endothelial barrier	2	6.1E-3	3.3E-1	3.9E-1	3.9E-1	RAPGEF2; PPP1R16B
Cellular Component							
GO:0005886	Plasma membrane	4	6.9E-2	9.1E-1	1.0E0	1.0E0	ATP6V0A2; RAPGEF2; GOT2; PPP1R16B
F1GA - Up-regulation							
Biological Process							
GO:0006935	Chemotaxis	2	8.8E-2	1.0E0	1.0E0	1.0E0	CCR4; CCR5
Molecular Function							
GO:0016493	C-C chemokine receptor activity	2	2.2E-2	6.3E-1	9.7E-1	9.7E-1	CCR4; CCR5
Kegg							
	Cytokine-cytokine receptor interaction	3	6.0E-2	8.9E-1	1.0E0	1.0E0	CCR4; CCR5; IFNAR2
	Amino sugar and nucleotide sugar metabolism	2	8.5E-2	9.6E-1	1.0E0	1.0E0	CMAS; LOC537017

4 Discussion

4.1 Water consumption frequency

4.1.1 Glucose metabolism

The LOC520336 (N-acetyllactosaminide beta-1,6-N-acetylglucosaminyl-transferase, isoform C) and GCNT2 (glucosaminyl (N-acetyl) transferase 2, I-branching enzyme (I blood group)) genes were founded as upregulated in NWCF group in the DE analysis. The LOC520336 gene, associated with DNA regions containing Single Nucleotide Polymorphism (SNP) assigned to this gene, related with Danish Holstein milk oligosaccharides synthesis (Poulsen et al., 2019). The GCNT2 gene were reported in the literature in genome-wide association studies associated with feet and leg conformation traits in Chinese Holstein cattle (Abdalla et al., 2021). The GCNT2 gene

was reported as a highly expressed gene on a slow-growing Xinghua chickens (Claire D'Andre et al., 2013) and on abomasal mucus biosynthesis, in nematode-infected Holstein calves (Rinaldi et al., 2011). Animals under heat stress may present higher water consumption (Blackshaw and Blackshaw, 1994) and greater glucose demand (Hamzaoui et al., 2020). This gene is involved with glycol-metabolism and in our study, GCNT2 were upregulated in the NWCF group, which could suggest a possibility of Nelore animals with lower water consumption frequency show differences of glucose metabolism compared to Nelore animals with higher water consumption frequency.
According to Honorat et al. (2017) the GOT2 gene (glutamic-oxaloacetic transaminase 2) encodes a mitochondrial enzyme, glutamate oxaloacetate transaminase 2, that performs an essential role in the intracellular NAD(H) redox balance during glycolysis (van Karnebeek et al., 2019). In a study with cold tolerance in Martes zibellina species, was pointed that this gene suffered a relaxation of selective pressure due to its evolutionary rate being three times greater than other carnivore relatives. The authors suggested that this species can obtain energy in a shorter time by the metabolism of lipids and amino acids while increasing their adaptation to cold weather (Xia et al., 2021). This gene was founded to be upregulated in F1WCF group for DD analysis, suggesting that probably the animals that had a lower frequency of water consumption needed an improved glucose metabolism to compensate for the lower water consumption, considering that water plays an important role in glycolysis (Mego, 1986).

4.1.2 Kidney phisiology

The ATP6V0A4 (ATPase H+ transporting V0 subunit a4), an up-regulated gene on NWCF in the DE analysis encodes a component of the vacuolar ATPase (V-ATPase), that is a multisubunit enzyme that mediates acidification of intracellular compartments of eukaryotic cells, and which is involved in normal transport of vector acid into urine by the kidneys (Stover et al., 2002). The mutations associated with this gene are related to preserved hearing and renal tubular acidosis (Nagara et al., 2014). This gene was observed to be differentially expressed in Large White pigs, when considering the backfat thickness phenotype (Wang et al., 2019). The ATP6V0A4 gene was up-regulated in animals that had a lower water consumption frequency in both

breeds considering DE and DD, probably due to tubular renal acidosis being controlled by the acid-base balance that water provides in the bloodstream (Karet, 2011).

4.1.3 Immune system

The CD96 (Cluster of Differentiation 96) gene was identified as statistically differentially expressed up-regulated in the low frequency NWCF in the DE analysis. This gene was reported in the literature as a high expressed gene in susceptible goats (Bhuiyan et al., 2017) and cattle (Araujo et al., 2009) to nematodes infection. It has also been reported in genomic analysis studies and selection signals in Iraqi (Kasarda et al., 2021) and Slovak (Alshawi et al., 2019) cattle as a gene associated with the immune system (Kasarda et al., 2021), as well as the KLRD1 (killer cell lectin-like receptor subfamily D, member 1) gene, also related to the immune response, expressed in natural killer cells (Bongen et al., 2018).

4.1.4 Adaptation to the environment

A study with camels characterizing the genetic content, organization and variability of genetic regions associated with cytotoxicity receptor genes highlighted the NCR1 (natural cytotoxicity triggering receptor 1) gene (Futas et al., 2019). These animals are known to survive in challenging environments of high temperatures and lack of water and food (Jirimutu et al., 2012). As phylogenetically most closely related species (Futas et al., 2019), the upregulation of this gene found in the NWCF group, in DE analysis, can also be explained because those animals seem more adapted to water scarcity.

The XRCC6 (X-ray repair cross complementing 6) founded to be up-regulated for NWCF group in DD analysis, is a gene associated with DNA repair (Reinardy et al., 2013). In a study regarding predictor genes for feed efficiency in Nelore bulls, the XRCC6 gene was found to be a co-expressed gene on high feed efficiency animals (Chen et al., 2021). In our study, animals that had a lower WCF can be considered efficient, as with less water consumption they still proved to be productively efficient since there was no difference in the average daily weight gain between the studied groups.

The PPP1R16B (protein phosphatase 1 regulatory subunit 16B) is a gene related to innate viral immunity (Rahman et al., 2017). This gene was verified in a study with Welsh Sheep breed, as a strong selected gene in its genome region, associated with adaptative ability of these animals in challenging environments (Sweet-Jones et al., 2021). This gene is also founded in a study with methane production levels in dairy cows. The authors reported that the biological process involving this gene could affect the digestive process, controlling the water transport within the intestine (Pszczola et al., 2018). This gene was founded up-regulated in F1WCF group, in the DD analysis, suggesting that these animals had adaptative ability in an adverse environment and could use water more effectively, in more challenging environments such as the one considered in this experiment.

4.1.5 Fat deposition

In cattle, THBS1 (thrombospondin 1) is reported in the literature as a gene induced during natural luteolysis or induced by prostaglandin F2 alpha (PGF2a); natural anti-angiogenic compound, pro-apoptotic and regulated in luteal endothelial cells (Farberov and Meiden, 2016; Basavaraja et al., 2017; Farberov and Meiden, 2018). This gene was differentially expressed and down-regulated in cattle infected with tuberculosis (Klepp et al., 2019) and was highly expressed on Korean castrated steers, been involved in the deposition of intramuscular fat after castration (Ju Park et al., 2018). Nutrition studies indicated that water consumption is inversely associated with body fat deposition in humans (Daniels and Popkin, 2010; Milla-Tobarra et al., 2016). In the present study, the THBS1 gene is down-regulated in animals with a lower frequency of water consumption in DE analysis, probably because when less water is consumed, there is less need for the expression of this gene related to fat deposition.

4.1.6 Blood pressure maintenance

A gene related to blood pressure in a study performed with humans (He et al., 2013), CDCA7 (cell division cycle associated 7), is found to be up-regulated on NWCF group in DD analysis. Study associating water intake and blood pressure in mice showed an increase in water consumption as the blood pressure increased to a high salt diet (Chen et al., 2006). Dehydrated rats presented a rise in plasma vasopressin

hormone, responsible for maintenance of normal systolic blood pressure (Woods and Johnston, 1983). As a group that showed a lower water consumption frequency, dispersion analysis of CDCA7 gene may be due to the maintenance of blood pressure while the animals consume less water. The RNA samples were extracted from the blood of the animals studied, therefore they represent the RNAs present in the blood.

4.2 General activity

4.2.1 Adaptation to the environment

For the NGA group, the enrichment analysis showed the up-regulated SLC16A6 (solute carrier family 16 member six) gene, in the DE analysis. This gene is associated with normal hearing function (Girotto et al., 2014) and height in humans and length in Zebrafish species by modulating monocarboxylic acid transportation in several tissues (Karanth and Schlegel, 2019). A study with whole blood transcriptome analysis in cattle adaptation to sub-arctic conditions, revealed the SLC16A6 gene to be up-regulated in Holstein cattle. The authors suggested that, as a solute carrier protein gene related, this can be associated with maintaining internal homeostasis in response to low temperatures, by increasing the ion uptake rate (Pokharel et al., 2019). Animals of this study had showed an upregulation of this gene, probably because with less general activity, and therefore, less energy requirement, managed to maintain themselves and produce the same as the animals with higher general activity frequency, suggesting that they are more adapted to the environment.

A study with the indigenous African cattle highlighted the adaptive capacity of those animals in the face of environmental pressures and human selection. Studying the genomic diversity, the authors found the STOM (stomatin) gene putatively positively selected in the studied animals (Kim et al., 2017). This gene was observed up-regulated on NGA group in the DE analysis, possibly also because they have greater adaptive capacity.

The genes VPS41 (VPS41 subunit Of HOPS complex) and CLTA (clathrin light chain A) were founded to be down-regulated in NGA group, as the DD analysis showed. A study with homozygosity segments in West African Djallonké Sheep, showed the VPS41 as a gene related to adaptation in the challenging environment of Sub-Saharan West Africa. This gene was part of a functional cluster of a genomic region that shows that, as this breed was not subject to selection programs for

production traits, this gene was favored by natural selection process for adaptation (Álvarez et al., 2020). The CLTA gene is a protein associated with intracellular transfer of membrane by coated vesicles (Pearse, 1976). This gene was reported on the literature associated with prion protein deposition in sheep (Filali et al., 2014) and it was found in a study with runs of homozygosity associated with local adaptation in Valle del Belice sheep (Mastrangelo et al., 2017).

A study with tropical adaptation from Creole cattle showed positive selection for IFNAR2 (interferon alpha and beta receptor subunit 2). The authors suggested that this gene plays an important role on key immune functions and highlighted that the candidate regions related to this gene could be important for responses against tropical diseases and parasite infections (Pitt et al, 2017). This gene was founded to be up-regulated in F1GA group for DD analysis due to the fact that these animals was raised in the same tropical conditions as the cited article. The same could be found in a study with genomic and selection signatures related to adaptation in Eritrean indigenous cattle (Goitom et al., 2021).

The genes found in this study related to the frequency of water consumption in animals are associated with glucose metabolism, kidney physiology, immune system, adaptive capacity, fat deposition, blood pressure maintenance and efficiency in taking advantage of the consumed water. The genes found for the general activity of animals are related to adaptation to environmental temperature and challenging environments. The adaptive mechanisms evidenced by the two genetic groups can, in part, be explained by the genes related to the studied traits.

5 Conclusion

The genes found in this study related to the frequency of water consumption in animals are associated with glucose metabolism, kidney physiology, immune system, adaptive capacity, fat deposition, blood pressure maintenance and efficiency in taking advantage of the consumed water. The genes found for the general activity of animals are related to adaptation to environmental temperature and challenging environments. The adaptive mechanisms evidenced by the two genetic groups can, in part, be explained by the genes related to the studied traits.

Ethics approval

The study was conducted according to the guidelines of the National Council for Animal Experiment Control, and approved by the Ethics Committee for Use of Animals (CEUA) of Universidade Estadual Paulista (Unesp) (protocol code #001081/2019, approved on 14 th February, 2019).

Declaration of interest

The authors declare no conflict of interest.

Financial support statement

This work was founded by Coordination for the Improvement of Higher Education Personnel (CAPES—grant number 001) and The Sao Paulo Research Foundation (FAPESP—grant numbers: and 2015/16631-5, 2018/20753-7 and 2018/11953-2).

References

Abdalla, I.M., Lu, X., Nazar, M., Arbab, A.A.I., Xu, T., Yousif, M.H., Mao, Y., Yang, Z., 2021. Genome-Wide Association Study Identifies Candidate Genes Associated with Feet and Leg Conformation Traits in Chinese Holstein Cattle. Animals 11, 2259. doi.org/10.3390/ani11082259.

Ahlberg, C.M., Allwardt, K., Broocks, A., Bruno, K., McPhillips, L., Taylor, A., Clint, R.K., Calvo-Lorenzo, M.S., Richards, C.J., Place, S.E., DeSilva, U., VanOverbekE, D.L., Mateescu, R.G., Kuehn, L.A., Weaber, R.L., Bormann, J.M., Rolf, M.M., 2018. Environmental effects on water intake and water intake prediction in growing beef cattle. Journal of Animal Science 96, 4368-4384.

Alshawi, A., Essa, A., Al-Bayatti, S., Hanotte, O., 2019. Genome analysis reveals genetic admixture and signature of selection for productivity and environmental traits in Iraqi cattle. Frontiers in Genetics 10, 609. doi.org/10.3389/fgene.2019.00609.

Álvarez, I., Fernández, I., Traoré, A., Pérez-Pardal, L., Menéndez-Arias, N.A., Goyache, F., 2020. Genomic scan of selective sweeps in Djallonké (West African Dwarf) sheep shed light on adaptation to harsh environments. Scientific Reports 10, 2824. doi.org/10.1038/s41598-020-59839-x.

Andrews, S., 2010. Babraham bioinformatics-FastQC a quality control tool for high throughput sequence data. Retrieved on 5 September 2020 from https://www.bioinformatics.babraham.ac.uk/projects/fastqc.

Araujo, R.N., Padilha, T., Zarlenga, D., Sonstegard, T., Connor, E.E., Van Tassel, C., Lima, W.S., Nascimento, E., Gasbarre, L.C., 2009. Use of a candidate gene array to delineate gene expression patterns in cattle selected for resistance or susceptibility to intestinal nematodes. Veterinary Parasitology 162, 106-115.

Basavaraja, R., Przygrodzka, E., Pawlinski, B., Gajewski, Z., Kaczmarek, M.M., Meidan, R., 2017. Interferon-tau promotes luteal endothelial cell survival and inhibits specific luteolytic genes in bovine corpus luteum. Reproduction 154, 559-568.

Beatty, D.T., Barnes, A., Taylor, E., Pethick, D., McCarthy, M., Maloney, S.K., 2006. Physiological responses of Bos taurus and Bos indicus cattle to prolonged, continuous heat and humidity. Journal of Animal Science 84, 972-985.

Beltman, M.E., Forde, N., Lonergan, P., Crowe, M.A., 2013. Altered endometrial immune gene expression in beef heifers with retarded embryos. Reproduction, Fertility and Development 25, 966-970.

Bhuiyan, A.A., Li, J., Wu, Z., Ni, P., Adetula, A.A., Wang, H., Zhang, C., Tang, X., Bhuyan, A.A., Zhao, S., Du, X., 2017. Exploring the genetic resistance to gastrointestinal nematodes infection in goat using RNA-sequencing. International Journal of Molecular Sciences 18, 751. doi.org/10.3390/ijms18040751.

Blackshaw, J.K., Blackshaw, A.W., 1994. Heat stress in cattle and the effect of shade on production and behaviour: a review. Australian Journal of Experimental Agriculture 34, 285-295.

Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114-2120.

Bongen, E., Vallania, F., Utz, P.J., Khatri, P., 2018. KLRD1-expressing natural killer cells predict influenza susceptibility. Genome medicine 10, 45. doi.org/10.1186/s13073-018-0554-1.

Bricarello, P.A., Zaros, L.G., Coutinho, L.L., Rocha, R.A., Kooyman, F.N.J., De Vries, E., Gonçalves, J.R.S., Lima, L.G., Pires, A.V., Amarante, A.F.T., 2007. Field study on nematode resistance in Nelore-breed cattle. Veterinary parasitology 148, 272-278.

Chen, W., Alexandre, P., Ribeiro, G., Fukumasu, H., Sun, W., Reverter, A., Li, Y., 2021. Identification of Predictor Genes for Feed Efficiency in Beef Cattle by Applying Machine Learning Methods to Multi-Tissue Transcriptome Data. Frontiers in Genetics 12, 103. doi.org/10.3389/fgene.2021.619857.

Chen, Y., Oroszi, T.L., Morris, M., 2006. Salt consumption increases blood pressure and abolishes the light/dark rhythm in angiotensin AT1a receptor deficient mice. Physiology & Behavior 88, 95-100.

D'Andre, H.C., Paul, W., Shen, X., Jia, X., Zhang, R., Sun, L., Zhang, X., 2013. Identification and characterization of genes that control fat deposition in chickens. Journal of Animal Science and Biotechnology 4, 43. doi.org/10.1186/2049-1891-4-43.

Daniels, M.C., Popkin, B.M., 2010. Impact of water intake on energy intake and weight status: a systematic review. Nutrition Reviews 68, 505-521.

de Jong, T.V., Moshkin, Y.M., Guryev, V., 2019. Gene expression variability: the other dimension in transcriptome analysis. Physiological genomics 51, 5, 145-158.

Ekman, A., Ilves, M., Iivanainen, A., 2012. B lymphopoiesis is characterized by pre-B cell marker gene expression in fetal cattle and declines in adults. Developmental & Comparative Immunology 37, 39-49.

Ekman, A., Niku, M., Liljavirta, J., & Iivanainen, A., 2009. Bos taurus genome sequence reveals the assortment of immunoglobulin and surrogate light chain genes in domestic cattle. BMC immunology 10, 22. doi.org/10.1186/1471-2172-10-22.

Farberov, S., Meidan, R., 2016. Thrombospondin-1 affects bovine luteal function via transforming growth factor-beta1-dependent and independent actions. Biology of Reproduction 94, 25. doi.org/10.1095/biolreprod.115.135822.

Farberov, S., Meidan, R., 2018. Fibroblast growth factor-2 and transforming growth factor-beta1 oppositely regulate miR-221 that targets thrombospondin-1 in bovine luteal endothelial cells. Biology of Reproduction 98, 366-375.

Futas, J., Oppelt, J., Jelinek, A., Elbers, J.P., Wijacki, J., Knoll, A., Burger, P.A., Horin, P., 2019. Natural killer cell receptor genes in camels: Another mammalian model. Frontiers in Genetics, 10, 620. doi.org/10.3389/fgene.2019.00620.

Girotto, G., Vuckovic, D., Buniello, A., Lorente-Canovas, B., Lewis, M., Gasparini, P., Steel, K.P., 2014. Expression and replication studies to identify new candidate genes involved in normal hearing function. PLoS One 9, e85352.

Godfray, H.C.J., Crute, I.R., Haddad, L., Lawrence, D., Muir, J.F., Nisbett, N., Pretty, J., Robinson, S., Toulmin, C., Whiteley, R., 2010. The future of the global food system. Philosophical Transactions of the Royal Society 365, 2769-2777.

Goitom, S., Gicheha, M.G., Njonge, F.K., Kiplangat, N., 2021. Landscape genomics and selection signatures of local adaptation of Eritrean indigenous cattle along environmental gradients. Tropical Animal Health and Production 53, 423. doi.org/10.1007/s11250-021-02837-6.

Hamzaoui, S., Caja, G., Such, X., Albanell, E., Salama, A.A.K., 2020. Milk production and energetic metabolism of heat-stressed dairy goats supplemented with propylene glycol. Animals 10, 2449. doi.org/10.3390/ani10122449.

He, J., Kelly, T.N., Zhao, Q., Li, H., Huang, J., Wang, L., Jaquish, C.E., Sung, Y.J., Shimmin, L.C., Lu, F., Mu, J., Hu, D., Ji, X., Shen, C., Guo, D., Ma, J., Wang, R., Shen, J., Li, S., Chen, J., Mei, H., Chen, C., Chen, S., Chen, J., Li, J., Cao, J., Lu, X., Wu, X., Rice, T.K., Gu, C.C., Schwander, K., Hamm, L.L., Liu, D., Rao, D.C., Hixson, J.E., Gu D., 2013. Genome-wide association study identifies 8 novel loci associated with blood pressure responses to interventions in Han Chinese. Circulation: Cardiovascular Genetics 6, 598-607.

Honorat, J.A., Nakatsuji, Y., Shimizu, M., Kinoshita, M., Sumi-Akamaru, H., Sasaki, T., Takata, K., Koda T., Namba A., Yamashita K., Sanda E., Sakaguchi M., Kumanogoh A., Shirakura T., Tamura M., Sakoda S., Mochizuki, H., Okuno T., 2017. Febuxostat ameliorates secondary progressive experimental autoimmune encephalomyelitis by restoring mitochondrial energy production in a GOT2-dependent manner. PloS One 12, e0187215.

Huang D.W., Sherman B.T., Lempicki R.A., 2008. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Research 37, 1-13.

Jirimutu, Wang, Z., Ding, G., Chen, G., Sun, Y., Sun, Z., Zhang, H., Wang, L., Hasi, S., Zhang, Y., Li, J., Shi, Y., Xu, Z., He, C., Yu, S., Li, S., Zhang, W., Batmunkh, M., Ts, B., Narenbatu, Unierhu, Bat-Ireedui, S., Gao, H., Baysgalan, B., Li, Q., Jia, Z., Turigenbayila, Subudenggerile,

Narenmanduhu, Wang Z., Wang, J., Pan, L., Chen, Y., Ganerdene, Y., Dabxilt, Erdemt, Altansha, Altansukh, Liu, T., Cao, M., Aruuntsever, Bayart, Hosblig, He, F., Zha-ti, A., Zheng, G., Qiu F., Sun, Z., Zhao, L., Zhao, W., Liu, B., Li, C., Chen, Y., Tang, X., Guo, C., Liu, W., Ming, L., Temuulen, Cui, A., Li, Y., Gao, J., Li, J., Wurentaodi, Niu, S., Sun, T., Zhai, Z., Zhang, M., Chen, C., Baldan, T., Bayaer, T., Li, Y., Meng, H., 2012. Genome sequences of wild and domestic Bactrian camels. Nature Communications 3, 1202. doi: 10.1038/ncomms2192.

Jong, T.V., Moshkin, M., Gurvey, V., 2019. Gene expression variability: the other dimension in transcriptome analysis. Physiological Genomics, 51, 145-158. doi:10.1152/physiolgenomics.00128.2018

Karanth, S., Schlegel, A., 2019. The monocarboxylate transporter SLC16A6 regulates adult length in zebrafish and is associated with height in humans. Frontiers in Physiology 9, 1936.

Karet, F.E., 2011. Disorders of water and acid-base homeostasis. Nephron Physiology 118, 28-34.

Kasarda, R., Moravčíková, N., Olšanská, B., Mészáros, G., Vostrý, L., Vostrá-Vydrová, H., Lehocká, K., Prišťák, J., Candrák, J., 2021. The evaluation of genomic diversity and selection signals in the autochthonous Slovak Spotted cattle. Czech Journal of Animal Science 66, 251-261.

Kim, D., Langmead, B., Salzberg, S.L., 2015. HISAT: a fast spliced aligner with low memory requirements. Nature methods 12, 357-360.

Kim, J., Hanotte, O., Mwai, O.A., Dessie, T., Bashir, S., Diallo, B., Agaba, M., Kim, K., Kwak, W., Sung, S., Seo, M., Jeong, H., Kwon, T., Taye, M., Song, K., Lim, D., Cho, S., Lee, H., Yoon, D., Oh, S.J., Kemp, S., Lee, H., Kim, H., 2017. The genome landscape of indigenous African cattle. Genome biology 18, 34. doi.org/10.1186/s13059-017-1153-y.

Klepp, L.I., Eirin, M.E., Garbaccio, S., Soria, M., Bigi, F., Blanco, F.C., 2019. Identification of bovine tuberculosis biomarkers to detect tuberculin skin test and IFNγ release assay false negative cattle. Research in Veterinary Science 122, 7-14.

Lardner, H.A., Braul, L., Schwartzkopf-Genswein, K., Schwean-Lardner, K., Damiran, D., Darambazar, E., 2013. Consumption and drinking behavior of beef cattle offered a choice of several water types. Livestock science 157, 577-585.

Lardner, H.A., Kirychuk, B.D., Braul, L., Willms, W.D., Yarotski, J., 2005. The effect of water quality on cattle performance on pasture. Australian Journal of Agricultural Research 56, 97-104.

Liao, Y., Smyth, G.K., Shi, W., 2014. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 30, 923-930.

Love, M.I., Huber, W., Anders, S., 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology 15, 550. doi.org/10.1186/s13059-014-0550-8.

Mastrangelo, S., Tolone, M., Sardina, M.T., Sottile, G., Sutera, A.M., Di Gerlando, R., Portolano, B., 2017. Genome-wide scan for runs of homozygosity identifies potential candidate genes associated with local adaptation in Valle del Belice sheep. Genetics Selection Evolution 49, 84. doi.org/10.1186/s12711-017-0360-z.

Mego, J. L. (1986). The role of water in glycolysis. Biochemical Education, 14(3), 130-131.

Mekonnen, M.M., & Hoekstra, A.Y., 2016. Four billion people facing severe water scarcity. Science Advances 2, e1500323.

Milla-Tobarra, M., García-Hermoso, A., Lahoz-García, N., Notario-Pacheco, B., Lucas-de la Cruz, L., Pozuelo-Carrascosa, D.P., García-Meseguer, M.J., Martínez-Vizcaíno, V., 2016. The association between water intake, body composition and cardiometabolic factors among children: the Cuenca study. Nutricion Hospitalaria 33, 19-26.

Nagara, M., Voskarides, K., Elouej, S., Zaravinos, A., Riahi, Z., Papagregoriou, G., Kefi, R., Boussetta, K., Deltas, C., Abdelhak, S., Tinsa, F., 2014. A novel splice-site mutation in ATP6V0A4 gene in two brothers with distal renal tubular acidosis from a consanguineous Tunisian family. Journal of Genetics 93, 859-863.

Neil A. Campbell, Jane B. Reece, Lisa A. Urry, Michael L. Cain, Steven A. Wasserman, Peter V. Minorsky, and Robert B. Jackson, "Energy Requirements Are Related to Animal Size, Activity, and Environment," in Campbell Biology, 8th ed. (San Francisco: Benjamin Cummings, 2008), 870.

O'brien, A.M.P., Höller, D., Boison, S.A., Milanesi, M., Bomba, L., Utsunomiya, Y.T., Carvalheiro, R., Neves, H.H.R., Silva, M.V.B. da, VanTassell, C.P., Sonstegard, T.S., Mészáros, G., Ajmone-Marsan, P., Garcia, F., Sölkner, J., 2015. Low levels of taurine introgression in the current Brazilian Nelore and Gir indicine cattle populations. Genetics Selection Evolution 47, 31. doi.org/10.1186/s12711-015-0109-5.

Park, S.J., Beak, S., Jung, D.J.S., Kim, S.Y., Jeong I.H., Piao, M.Y., Kang, H.J., Fassah, D.M., Na, S.W., Yoo, S.P., Baik, M., 2018. Genetic, management, and nutritional factors affecting intramuscular fat deposition in beef cattle—a review. Asian-Australasian Journal of Animal Sciences 31, 1043-1061.

Pearse, B.M., 1976. Clathrin: a unique protein associated with intracellular transfer of membrane by coated vesicles. Proceedings of the National Academy of Sciences 73, 1255-1259.

Pires, B.V., Stafuzza, N.B., Lima, S.B.G.P.N.P., Negrão, J.A., Paz, C.C.P., 2019. Differential expression of heat shock protein genes associated with heat stress in Nelore and Caracu beef cattle. Livestock Science, 230, 103839. doi.org/10.1016/j.livsci.2019.103839.

Pitt, D., Bruford, M.W., Barbato, M., Orozco-terWengel, P., Martínez, R., Sevane, N., 2019. Demography and rapid local adaptation shape Creole cattle genome diversity in the tropics. Evolutionary Applications 12, 105-122.

Pokharel, K., Weldenegodguad, M., Popov, R., Honkatukia, M., Huuki, H., Lindeberg, H., Peippo, J., Reilas, T., Zarovnyaev, S., Kantanen, J., 2019. Whole blood transcriptome analysis reveals footprints of cattle adaptation to sub-arctic conditions. Animal Genetics 50, 217-227.

Poulsen, N.A., Robinson, R.C., Barile, D., Larsen, L.B., Buitenhuis, B., 2019. A genome-wide association study reveals specific transferases as candidate loci for bovine milk oligosaccharides synthesis. BMC Genomics 20, 404. doi.org/10.1186/s12864-019-5786-y.

Pszczola, M., Strabel, T., Mucha, S., Sell-Kubiak, E., 2018. Genome-wide association identifies methane production level relation to genetic control of digestive tract development in dairy cows. Scientific Reports 8, 15164. doi.org/10.1038/s41598-018-33327-9.

Rahman M.M., Bagdassarian E., Ali M.A.M., McFadden G., 2017. Identification of host DEAD-box RNA helicases that regulate cellular tropism of oncolytic Myxoma virus in human cancer cells. Scientific Reports 7, 15710. doi.org/10.1038/s41598-017-15941-1.

Ran, D., Daye, Z.J., 2017. Gene expression variability and the analysis of large-scale RNA-seq studies with the MDSeq. Nucleic Acids Research 45, e127. doi.org/10.1093/nar/gkx456.

Rao, M.S., Van Vleet, T.R., Ciurlionis, R., Buck, W.R., Mittelstadt, S.W., Blomme, E.A.G., Liguori, M.J., 2019. Comparison of RNA-Seq and microarray gene expression platforms for the toxicogenomic evaluation of liver from short-term rat toxicity studies. Frontiers in Genetics 9, 636. doi.org/10.3389/fgene.2018.00636.

Reinardy, H.C., Syrett, J.R., Jeffree, R.A., Henry, T.B., Jha, A.N., 2013. Cobalt-induced genotoxicity in male zebrafish (Danio rerio), with implications for reproduction and expression of DNA repair genes. Aquatic Toxicology 126, 224-230.

Rinaldi, M., Dreesen, L., Hoorens, P.R., Li, R.W., Claerebout, E., Goddeeris, B., Vercruysse, J., Broek, W.V.D., Geldhof, P., 2011. Infection with the gastrointestinal nematode Ostertagia ostertagi in cattle affects mucus biosynthesis in the abomasum. Veterinary Research 42, 61. doi.org/10.1186/1297-9716-42-61.

Sanders, J.O., 1980. History and development of Zebu cattle in the United States. Journal of Animal Science 50, 1188-1200.

Stover, E.H., Borthwick, K.J., Bavalia, C., Eady, N., Fritz, D.M., Rungroj, N., Giersch, A.B.S., Morton, C.C., Axon, P.R., Akil, I., Al-Sabban, E.A., Baguley, D.M., Bianca, S., Bakkaloglu, A., Bircan, Z., Chauveau, D., Clermont, M-J., Guala, A., Hulton, S.A., Kroes, H., Li Volti, G., Mir, S., Mocan, H., Nayir, A., Ozen, S., Rodriguez Soriano, J., Sanjad, S.A., Tasic, V., Taylor, C.M., Topaloglu, R., Smith, A.N., Karet, F.E., 2002. Novel ATP6V1B1 and ATP6V0A4 mutations in autosomal recessive distal renal tubular acidosis with new evidence for hearing loss. Journal of Medical Genetics 39, 796-803.

Stromberg, B.E., Averbeck, G.A., 1999. The role of parasite epidemiology in the management of grazing cattle. International Journal for Parasitology 29, 33-39.

Sweet-Jones, J., Lenis, V.P., Yurchenko, A.A., Yudin, N.S., Swain, M., Larkin, D.M., 2021. Genotyping and Whole-Genome Resequencing of Welsh Sheep Breeds Reveal Candidate Genes and Variants for Adaptation to Local Environment and Socioeconomic Traits. Frontiers in Genetics, 12, 612492. doi.org/10.3389/fgene.2021.612492.

van Karnebeek, C.D., Ramos, R.J., Wen, X.Y., Tarailo-Graovac, M., Gleeson, J.G., Skrypnyk, C., Brand-Arzamendi, K., Karbassi, F., Issa, M.Y., van der Lee, R., Drögemöller, B.I., Koster, J., Rousseau J., Campeau, P.M., Wang, Y., Cao, F., Li, M., Ruiter J., Ciapaite, J., Kluijtmans, L.A.J., Willemsen, M.A.A.P., Jans, J.J., Ross C.J., Wintjes, L.T., Rodenburg, R.J., Huigen, M.C.D.G., Jia, Z., Waterham, H.R., Wasserman, W.W., Wanders, R.J.A., Verhoeven-Duif, N.M., Zaki, M.S., Wevers, R.A., 2019. Bi-allelic GOT2 mutations cause a treatable malate-aspartate shuttle-related encephalopathy. The American Journal of Human Genetics 105, 534-548.

Wagner, J.J., Engle, T.E., 2021. Invited Review: Water consumption, and drinking behavior of beef cattle, and effects of water quality. Applied Animal Science 37, 418-435.

Wang, B., Li, P., Zhou, W., Gao, C., Liu, H., Li, H., Niu, P., Zhang, Z., Li, Q., Zhou, J., Huang, R., 2019. Association of twelve candidate gene polymorphisms with the intramuscular fat content and average backfat thickness of Chinese Suhuai pigs. Animals 9, 858. doi.org/10.3390/ani9110858.

Wang, Z., Gerstein, M., Snyder, M., 2009. RNA-Seq: a revolutionary tool for transcriptomics. Nature Reviews Genetics 10, 57-63.

Wang, Z., Zhang, J., 2011. Impact of gene expression noise on organismal fitness and the efficacy of natural selection. Proceedings of the National Academy of Sciences, 108, 16, E67-E76.

Woods, R.L., Johnston, C.I., 1983. Contribution of vasopressin to the maintenance of blood pressure during dehydration. American Journal of Physiology-Renal Physiology 245, F615-F621.

Xia, T., Zhang, L., Sun, G., Yang, X., Zhao, C., Zhang, H., 2021. Insights into cold tolerance in sable (Martes zibellina) from the adaptive evolution of lipid metabolism. Mammalian Biology 101, 861-870.

Yim, S.H., Chung, Y.J., Jin, E.H., Shim, S.C., Kim, J.Y., Kim, Y.S., Hu, H.J., Shin, S.H., Pae, H.O., Zouali, M., Chung, H.T., 2011. The potential role of VPREB1 gene copy number variation in susceptibility to rheumatoid arthritis. Molecular immunology 48, 1338-1343.

Zarek, C.M., Lindholm-Perry, A.K., Kuehn, L.A. et al. Differential expression of genes related to gain and intake in the liver of beef cattle. BMC Res Notes 10, 1 (2017). https://doi.org/10.1186/s13104-016-2345-3

Capítulo 3[1] - Strategy to Predict High and Low Frequency Behaviors Using Triaxial Accelerometers in Grazing of Beef Cattle

Simple Summary: Monitoring animal activity in production systems is an important tool for obtaining information on health, production, and reproduction. In this study, we evaluated the use of accelerometers with different strategies to predict the grazing behavior of Nelore cattle. This research was conducted in an environment both more challenging and representative of the practices adopted in livestock production systems in Brazil. The results of this study showed that the use of the Random Forest algorithm, together with techniques for resampling the training data of the models, classified the studied behaviors with high accuracy, especially for important, and less frequent activities such as water consumption frequency.

Abstract: Knowledge of animal behavior can be indicative of the well-being, health, productivity, and reproduction of animals. The use of accelerometers to classify and predict animal behavior can be a tool for continuous animal monitoring. Therefore, the aim of this study was to provide strategies for predicting more and less frequent beef cattle grazing behaviors. The behavior activities observed were grazing, ruminating, idle, water consumption frequency (WCF), feeding (supplementation) and walking. Three Machine Learning algorithms: Random Forest (RF), Sup-port Vector Machine (SVM) and Naïve Bayes Classifier (NBC) and two resample methods: under and over-sampling, were tested. Overall accuracy was higher for RF models trained with the over-sampled dataset. The greatest sensitivity (0.808) for the less frequent behavior (WCF) was observed in the RF algorithm trained with the under-sampled data. The SVM models only per-formed efficiently when classifying the most frequent behavior (idle). The greatest predictor in the NBC algorithm was for ruminating behavior, with the over-sampled training dataset. The results showed that the behaviors of the studied animals were classified with high accuracy and specificity when the RF algorithm trained with the resampling methods was used. Resampling training datasets is a strategy to be considered, especially when less frequent behaviors are of interest.

[1] Artigo aceito para publicação no periódico “Animals” em 23 de novembro de 2021.

Keywords: machine learning; Naïve Bayes Classifier; Nelore; Random Forest; Support Vector Machine

1 Introduction

Monitoring and accessing animal behavior are important tasks in ensuring the success of an animal production system. The animal's behavior monitored individually and continuously can serve as an indication of its welfare and health [1]. Rumination and feeding behaviors of dairy cattle can indicate productivity measures [2]. How much time animals spend lying down can help estrus detection in cows [3]. By observing how animals walk or how much time they spend lying down can also help to detect and prevent lameness [4]. However, monitoring animal behavior is often carried out by human observation or video monitoring, which makes it difficult to obtain data, due to the demand for human resources [5], as well as the fact that sometimes access to the animals is not easy [6]. Therefore, using accelerometers that automatically measure the animal's activity has the potential to obtain this information, especially in extensive systems, where access to the animals is more difficult.

There are few studies on systems that help to classify less frequent behaviors such as drinking [7], although these behaviors are indicative of animal health and growth [8]. Drinking behavior has been reported in studies with a controlled environment [9] and in general the authors suggested further studies and adaptations for better drinking behavior predictions [10]. Additionally, the majority of results reported for animal behavior prediction using accelerometer information is with a taurine genetic composition [11–14], which can be different from a zebu genetic composition, due to the differences in temperament [15].

Less frequent behavior classes can lead to a bottleneck in the classification algorithms' performance [16]. These less frequent types of behavior increase detection difficulty due to the infrequency and casualness, which results in misclassification of these classes [17]. They can lead to imbalanced data, which refers to a dataset with one or some of the classes having a greater number of observations than the others. The most prevalent class is called the majority class, while the class with less frequent observation is called the minority class [18]. To deal with imbalanced datasets, some resampling techniques are used to rebalance the number of observations in order to facilitate the effect of skewed class distribution in the learning process of a prediction

algorithm. These resampling methods are more versatile as they are independent of the classification algorithm [19].

Therefore, the aim of this study was to provide strategies for predicting more and less frequent bovine behaviors, using over and under-sampling training data and comparing three distinct Machine Learning classification algorithms, using information from triaxial accelerometers on pasture raised animals.

2 Materials and Methods

All the procedures used followed the Ethical Principles for Animal Experimentation stated by the National Council for Animal Experiment Control and were approved by the Ethics Committee for Use of Animals (CEUA) of Universidade Estadual Paulista (Unesp), under protocol #001081/2019.

2.1. Experimental Area and Animals

The experiment was carried out in the Forage Crops and Grasslands section of Uni-versidade Estadual Paulista (Unesp), Jaboticabal, São Paulo, Brazil. The total area was di-vided into 24 paddocks, seeded with *Trochlea brizantha* (Hochs tex A. Rich) Stapf cv. Ma-randu (Marandu grass) in 2001. The grazing mean height was 25 cm, using the continuous grazing method with a variable stocking rate. The region's climate was humid sub-tropical, with dry winters and rainy summers. The average annual temperature was 22.3 °C, with a maximum average of 29.1 °C and a minimum average of 16.9 °C. The annual average relative humidity was 71.2% and the wettest quarter of the year was the first one (January, February, and March), with an average rainfall of 628.8 mm, equivalent to 44.2% of the total annual rainfall. In this experiment, eight Nelore (Bos indicus) animals (343 ± 27 kg), were finished in pastures and were provided a high level of supplementation. The quantity of dietary supplementation provided daily, to reach the animals' requirements corresponded to 2% of the animals' body weight, During the dry season. ambient average temperature, maximum average temperature, and minimum average temperature were 26.1 °C, 34.0 °C and 99 18.5 °C, respectively, and the rainfall was 156.8 mm, distributed over 12 days.

The animals were previously adapted to using tags coupled to custom halters. The tags attached to the halters were kept on the animals (appendix 3) for 28 days,

which corresponded to the periods of adaptation (25 days) and behavioral observation of the animals. The accelerometers were adjusted on the halter so that it was possible to obtain the information accurately but without causing physical harm to the animals. The animals were observed daily during supplementation and, after the end of the experiment, were evaluated in the corral.

2.2. Accelerometers and Animal Behavior

The tags used in this experiment were provided by the Ovi-Bovi® company (Minsk, Belarus) and consist of triaxial accelerometers using a microelectromechanical system (MEMS) (model LIS2DE12; ST Microelectronics® (Plan-les-Ouates, Switzerland)), weighing 80 g, with dimensions of 105 mm × 60 mm × 22 mm, and attached to a custom halter and placed on the underjaw of young bulls to detect their movements (appendix 3). The accelerometers provide movement information along three axes: (X [horizontal movements—side to side], Y [longitudinal movements—front to back] and Z [vertical movements—up and down]). The information from the accelerometers was transmitted in 6 s window size (approximately 0.167 Hz) time intervals and collected through a wireless system (band of 433 MHz) and later stored in the cloud (Ovi-Bovi® company (Minsk, Belarus) server). The window size was determined following the manufacturer's recommendations and based on previous results, taking into account the biology of the behaviors studied, battery life and data loss due to their collision at the time of transmission.

The behavioral observations of the animals were carried out during a 12 h per day period (6am to 6pm) for two consecutive days (24 September 2019 and 1 October 2019). The animal's behavior was noted whenever the animal changed its behavior, registering the time when it occurred. The behavior activities observed were grazing, ruminating (noted whenever rumination was observed, whether standing or lying down), idle (lying or standing), water consumption frequency (WCF), feeding (supplementation) and walking. A description of each behavior is provided in appendix 4. When the animal changed behavior, the time was noted according to each activity by the animal.

2.3. Data Processing and Prediction Algorithms

All accelerometer data was processed using the R base package (version 4.0.0, RStudio, Boston, MA, USA) [20]. The raw data from the accelerometers was accessed by the Ovi-Bovi® (Minsk, Belarus) tag provider server. This data consisted of tag identification, information of time and date and the variables of the movement axes, x, y and z (transformed into gravity unit g = 9.18 m s−2), which totalized 101,144 records for each variable. In addition to the variables provided by the accelerometers, the predictor variables of signal magnitude area (SMA), signal vector magnitude (SVM), movement variation, energy, entropy, pitch, roll, and inclination were calculated based on information from the three movement axes, according to Alvarenga et al. [21]. The equations for the calculations of these variables are presented in supplementary appendix 5. Additionally, the meteorological variables provided by the agroclimatological station of Universidade Estadual Paulista (Unesp), Jaboticabal, São Paulo, Brazil were considered as predictors in the models. The weather station was located 800 m away from the experimental site. The variables were air temperature, relative humidity, wind speed, wind direction, solar radiation, and maximum wind gust, provided in 10 min daily time intervals.

The prediction of animal behaviors through accelerometer information was evaluated using Random Forest (RF) [22], Support Vector Machine (SVM) [23] and Naïve Bayes Classifier (NBC) [24] algorithms. The first algorithm, RF, was performed by the random-Forest R package [25] considering 500 trees (ntree), five variables randomly sampled as candidates at each split (mtry) with the predictors' importance being taken into account (importance). The second and third algorithms were performed by the e1071 R package [26]. The SVM models were performed using a classification method type, a radial kernel type with a 0.1 g value and a cost of constraints violation of 10. The NBC models were built with default function arguments. All the algorithms considered the raw accelerometer data, transformed into gravity units, calculated variables as mentioned and meteorological variables as predictors and the animals' behaviors, as a response variable. Window size considered for predictions was 6 s with no overlapping window stride.

The dataset was divided into training (70% of the original dataset), wherein the predictor variables and all the observed animal behavior were considered, and test

(30% of the original dataset) datasets, wherein only the predictor variables were included. The accelerometer data were compared on a 6 s basis with the observation data. Each algorithm was trained to classify the six behaviors considered. As the observations data were carried out by noting the time that the animal changed its behavior and the window size of each event recorded by the accelerometer was 6 s, the observation data was replicated until the behavior change to compare with the accelerometer data. To compare the prediction ability of each model, the sensitivity (1), specificity (2), accuracy (3), and Kappa coefficient [27], which compares the observed accuracy with the expected accuracy (random chance), were calculated for the test dataset using the confusionMatrix function of caret R package [28].

$$sensitivity = \frac{true\ positive}{(true\ positive + false\ negative)} \quad (1)$$

$$specificity = \frac{true\ negative}{(true\ negative + false\ positive)} \quad (2)$$

$$accuracy = \frac{(true\ positive + true\ negative)}{(true\ positive + true\ negative + false\ positive + false\ negative)} \quad (3)$$

where true positive was the number of instances in which the animal behavior of interest was correctly classified after testing; false negative was the number of instances in which the animal behavior of interest was observed visually but was classified incorrectly as some other animal behavior; false positive was the number of instances in which the animal behavior of interest was incorrectly classified but not observed; and true negative was the number of instances in which the animal behavior of interest was correctly classified as not being observed.

2.4. Resampling Methods to Deal with Imbalanced Data

To deal with an imbalanced dataset (Figure 1) that can impair the predictive ability of the studied methods, two resampling methods were used in the training dataset. The over-sampling method [29], which eliminated the damage caused by skewed distribution by creating new minority class samples, and the under-sampling method [30], which also eliminated the damage caused by skewed distribution, but by removing the intrinsic samples in the majority class. The functions upSample and downSample from the caret R package [28] were used to add additional samples to

the minority classes with replacements to make the class distributions equal and to discard samples randomly so that all classes had the same frequency as the minority class, respectively.

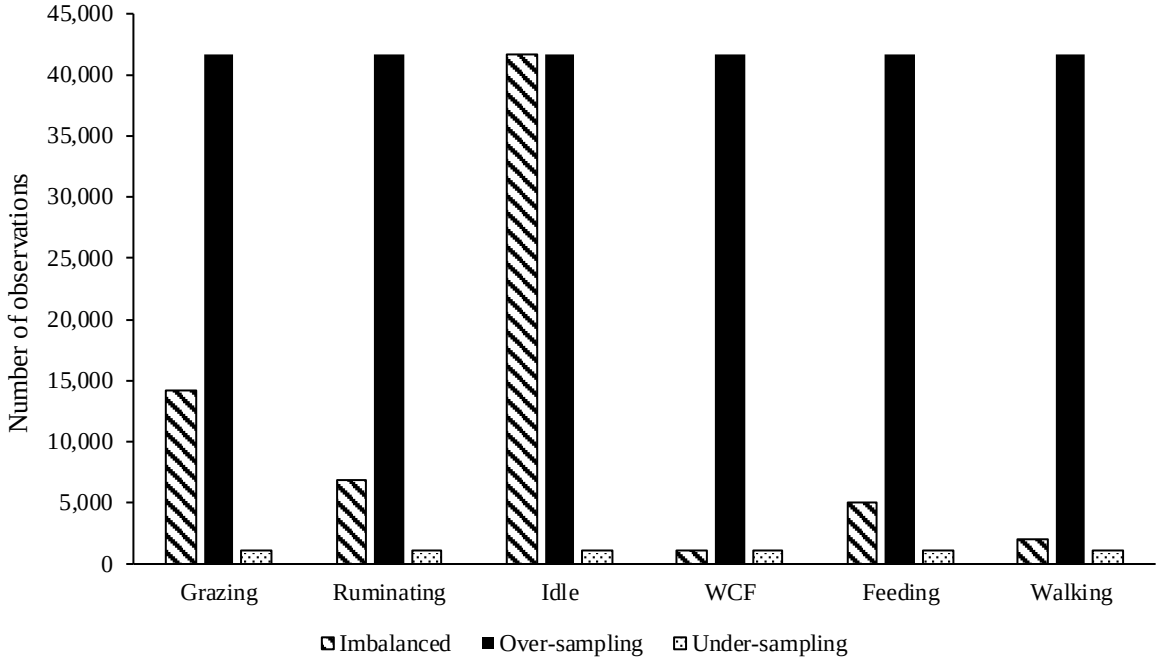

Figure 1. Number of observations for the imbalanced, over and under-sampled training datasets for grazing, ruminating, idle, WCF, feeding and walking behaviors in grazing beef cattle.

3 Results

In general, overall accuracy was higher for RF models, being the greatest for the RF model trained with over-sampled data. The lowest overall accuracy was observed in the NBC model trained with over-sampled records, which was the only method in which the over-sampling showed negative effects on behavior classification, since for the RF and NBC algorithms the training with over-sampled data promoted the highest results (Table 1). The same patterns can be observed for the Kappa coefficient, where the highest values were for the RF algorithm and, with the exception of the NBC algorithm, the highest values were observed when the training was performed with over-sampled data.

Table 1. Sensitivity, specificity, overall accuracy and Kappa coefficient for the Machine Learning algorithms and resampled training datasets for the studied behaviors.

Algorithm and Resample Training Datasets	**Behaviors**						**Overall Accuracy**	**Kappa Coefficient**
	Grazing	**Ruminating**	**Idle**	**WCF**	**Feeding**	**Walking**		
Random Forest								
Imbalanced							0.880	0.789
Sensitivity	0.816	0.876	0.957	0.278	0.688	0.501		
Specificity	0.960	0.996	0.823	0.999	0.991	0.998		

Over-sampling							0.920	0.865
Sensitivity	0.894	0.938	0.952	0.590	0.860	0.700		
Specificity	0.966	0.995	0.917	0.999	0.990	0.997		
Under-sampling							0.647	0.505
Sensitivity	0.644	0.901	0.580	0.808	0.768	0.797		
Specificity	0.892	0.921	0.942	0.948	0.925	0.949		
Support Vector Machine								
I*mbalanced*							0.611	0.078
Sensitivity	0.039	0.131	0.995	0.021	0.100	0.027		
Specificity	0.998	0.999	0.059	0.999	0.999	0.999		
Over-sampling							0.611	0.078
Sensitivity	0.039	0.131	0.995	0.021	0.100	0.027		
Specificity	0.998	0.999	0.059	0.999	0.999	0.999		
Under-sampling							0.267	0.075
Sensitivity	0.970	0.201	0.066	0.222	0.096	0.204		
Specificity	0.120	0.994	0.992	0.994	0.994	0.992		
Naïve Bayes Classifier								
Imbalanced							0.367	0.100
Sensitivity	0.284	0.700	0.392	0.000	0.105	0.083		
Specificity	0.922	0.605	0.608	0.999	0.971	0.971		
Over-sampling							0.179	0.072
Sensitivity	0.122	0.865	0.065	0.138	0.360	0.157		
Specificity	0.962	0.422	0.963	0.933	0.850	0.950		
Under-sampling							0.362	0.124
Sensitivity	0.121	0.580	0.423	0.015	0.402	0.126		
Specificity	0.958	0.702	0.727	0.993	0.815	0.954		

The proportions for each behavior of the total observations were equal to 20%, 10%, 59%, 1%, 7% and 3% for grazing, ruminating, idle, WCF, feeding and walking, respectively. The greatest sensitivity (0.808) for the less frequent behavior (WCF) was observed in the RF algorithm trained with under-sample data (0.808), followed by the same algorithm trained with over-sample and imbalanced data (0.590). Similar results were noted for the second less frequent behavior (walking). Feeding behavior presented a greater proportion of true positive observations when the RF with over-sampled training was used (0.860), followed by under-sampled (0.768) and imbalanced (0.688) training in the same method.

Considering the NBC algorithm to classify feeding behavior, training with re-sampled records showed better results than training with imbalanced data, with a difference ranging from 0.255 to 0.297 for sensitivity (Table 1). Classifying ruminating behavior by the RF algorithm was better for re-sampled training, however, the training with imbalanced data also presented a high sensitivity. The NBC algorithm showed higher sensitivity prediction for ruminating behavior when training datasets were with over-sampled (0.865) and imbalanced (0.700) records,

followed by the under-sample records (0.580). Grazing and Idle behaviors resulted in a greater proportion of true positive observations in the trained over-sampled and imbalanced RF algorithm, followed by the RF algorithm trained with under-sample data (Table 1).

The SVM models only performed efficiently when classifying the two more frequent behaviors (imbalanced and over-sampled training datasets for idle and under-sampled training dataset for grazing). The lowest proportion of true negative observations was found in SVM models when classifying the most frequent behavior (idle) with imbalanced and over-sampled training datasets (0.059 for both models), followed by the classification for grazing behavior using the under-sampled training dataset (0.120).

4 Discussion

The highest overall accuracy values observed for the RF algorithm (Table 1) corroborate the accuracy results, comparing RF with other Machine Learning algorithms, found in a study that classified behaviors in wild animals [31] and human behavior [32] using accelerometer information. The RF classification algorithm is highly capable at selecting and classifying predictor variables and at discriminating between predicted variables. This RF feature becomes important to evaluate information derived from accelerometers because these generate large amounts of data, which consume more time to select relevant variables [33] and lead to error prone and subjective tasks [34]. Therefore, due to the great-er stability compared to SVM [35], the overall accuracy results of the current study were higher when RF was used.

When compared to SVM, the NBC algorithm showed the lowest overall accuracies. This lower result was also found in a study classifying cows' behavior, comparing classification algorithms, using accelerometer information [36]. The overall accuracy result for the SVM algorithm of the present study (ranging from 0.267 to 0.611) was lower than that observed by the aforementioned study. According to Douglas et al. [37], SVM algorithms can be more suitable for complex classification tasks, especially in the training algorithm process. When trying to classify sow-activity using accelerometer data, Escalante et al. [38] found the

lowest performance using the NBC algorithm, compared to SVM and RF. In our study, the SVM algorithm did not perform well, in general, when classifying the behaviors studied, as observed by the aforementioned authors.

Due to the short period of time and intermittence that cattle drink water [39,40], this behavior tends to be less frequent than other observed behaviors and often this action is not considered in the classification analyses, even if this behavior was observed and noted [12,41]. In the current study, the less frequent behavior of WCF presented better results than those found in the literature [10], when the RF algorithm was used, especially when trained with resampled datasets. Similarly, when classifying WCF using resampled training and the RF algorithm, the sensitivity and specificity were greater than those found by Williams et al. [9], assessing the classification of drinking water behavior in cattle in periods of time less than or equal to 10 s. Although these authors reported higher true positive rate results when time periods longer than 10 s were observed, the experiment was conducted in a more controlled environment and used, in addition to accelerometers, a water flow meter. The results found in the present study were obtained in an extensive rearing system environment, a widely used practice, therefore representative of Brazilian regions and of the greater difficulty in handling the animals. The better results for predicting lower frequency behaviors, found in the current study, may help in future studies to monitor animals' health and welfare and also in genetic breeding programs.

Performance in predicting true positives for feeding and walking behaviors using NBC and SVM algorithms with the three training datasets considered was higher than the percentage of correct classifications in a study classifying sow-activity using accelerometer data [38]. When the RF algorithm was considered, walking activity had better prediction results when resampling methods were used. Feeding behavior had a higher correct classification when all three RF datasets were used. Even in a more controlled environment, with more data collected and the animals' behavior being monitored by video cam-eras, the results of the present study were, in general, greater than those found in the aforementioned study. The resampling strategy used in this paper can lead to better behavior classifications.

Resampling training with the RF algorithm showed higher sensitivity results than those observed with unbalanced data in predicting rumination, WCF, feeding and walking activities. The same was observed when the NBC algorithm was considered, as well as for prediction of rumination, WCF and walking with the under-sampled dataset in the SVM algorithm. The slight decrease or no sensitivity gain observed for grazing and idle behaviors when the resample methods were used, especially for RF and NBC, may have occurred due to the random sampling of these majority behavior classes, leading to a de-crease in their true positive rates. Balancing the database to equalize the number of observations should only be performed when the class of interest is the minority one [42]. Idle behavior had higher sensitivity results when considering the imbalanced and un-der-sampled datasets for the RF algorithm. According to Escalante et al. [38], passive activities are more difficult to classify due to noise generated when eventual animal movements happen. Thus, the RF algorithm can deal better with noisy measurements when more observations for idle behavior are considered.

Considering the SVM algorithm with the imbalanced training dataset, grazing behavior had low sensitivity and high specificity, while for idle behavior the sensitivity was high and the specificity low. This pattern was repeated within each behavior for training with the under and over-sampled datasets. A possible explanation is due to the fact that the SVM algorithm confused the classification of grazing behavior with idle behavior. Martiskainen et al. [43] observed that the SVM algorithm confused some of the behaviors studied, especially in similar activities, according to the position of the accelerometer attached to the cows. The differences observed in each movement pattern of a given behavior can interfere in the movement's classification [44].

According to Zughrat et al. [45], the under-sampling technique can drastically reduce the number of support vectors in an SVM algorithm, leading to less computational de-mand, resulting in a performance gain if compared with the over-sampling technique. In the present study, the under-sampled dataset used with SVM, showed higher sensitivity results for the majority of classified behaviors, however for idle behavior this was not observed, probably also due to the confusion in the classification of behaviors, where the algorithm may have classified idle

behavior as grazing when the under-sampled training dataset was used. Using this resampling technique, data from the majority class was re-moved, which may also have influenced this result. The opposite may have occurred for the oversampled and imbalanced training datasets. Using resampling methods in training datasets can promote little or no gain in the predictive performance of this algorithm [46].

When the under-sampling technique was considered, due to the fact that the observations of the majority behavioral classes were reduced, the sensitivity of these classes was impaired. In general, the RF algorithm was the one that best managed to classify the studied behaviors, together with the over-sampling training dataset, as it increases the number of observations for the minority behavior classes without impairing the classification of the majority classes. However, it should be taken into account that when increasing the amount of information there is greater computational cost for the analyses and a greater amount of time is needed to accomplish them.

5 Conclusions

The results showed that the behaviors of the studied animals were classified with great accuracy and specificity when the RF algorithm trained with the resampling methods was used. Therefore, in general, the best strategy to classify and predict more frequent behaviors was using the RF algorithm, and when less frequent behaviors are the main interest, the most appropriate strategy would be using the over-sampling technique for training the data.

Author Contributions: Conceptualization, R.N.W., P.A.B., E.P.R., R.A.R. and D.P.M.; methodology, R.N.W., P.A.B. and E.P.R.; experiment conduction, R.N.W., L.G.B., T.R.B. and R.W.T.; software, R.N.W., P.A.B. and E.P.R.; formal analysis, R.N.W.; investigation, R.N.W. and P.A.B.; data curation, R.N.W.; writing—original draft preparation, R.N.W.; writing—review and editing, R.N.W. and L.G.B.; visualization, R.N.W., P.A.B.; supervision, D.P.M., R.A.R. and P.A.B.; project administration, D.P.M.; funding acquisition, D.P.M. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by Coordination for the Improvement of Higher Education Personnel (CAPES—grant number 001) and The Sao Paulo

Research Foundation (FAPESP—grant numbers: 2018/20753-7 and 2015/16631-5).

Institutional Review Board Statement: The study was conducted according to the guidelines of the National Council for Animal Experiment Control, and approved by the Ethics Committee for Use of Animals (CEUA) of Universidade Estadual Paulista (Unesp) (protocol code #001081/2019, ap-proved on 14 February 2019).

Data Availability Statement: The data that support the findings of this study is available upon reasonable request contacting the corresponding author.

Acknowledgments: The authors would like to acknowledge everyone who helped in collecting data and conducting the experiment.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Barwick, J.; Lamb, D.W.; Dobos, R.; Welch, M.; Trotter, M. Categorising sheep activity using a tri-axial accelerometer. Comput. Electron. Agric. 2018, 145, 289–297, doi: 10.1016/J.COMPAG.2018.01.007.

2. Diosdado, J.A.V.; Barker, Z.E.; Hodges, H.R.; Amory, J.R.; Croft, D.P.; Bell, N.J.; Codling, E.A. Classification of behaviour in housed dairy cows using an accelerometer-based activity monitoring system. Anim. Biotelemetry 2015, 3, 1–14, doi:10.1186/S40317-015-0045-8.

3. Jensen, M.B. Behaviour around the time of calving in dairy cows. Appl. Anim. Behav. Sci. 2012, 139, 195–202, doi: 10.1016/J.APPLANIM.2012.04.002.

4. Thorup, V.M.; Munksgaard, L.; Robert, P.E.; Erhard, H.W.; Thomsen, P.T.; Friggens, N.C. Lameness detection via leg-mounted accelerometers on dairy cows on four commercial farms. Animal 2015, 9, 1704–1712, doi:10.1017/S1751731115000890.

5. Penning, P.D. A technique to record automatically some aspects of grazing and ruminating behaviour in sheep. Grass Forage Sci. 1983, 38, 89–96, doi:10.1111/j.1365-2494. 1983.tb01626. x.

6. Buerkert, A.; Schlecht, E. Performance of three GPS collars to monitor goats' grazing itineraries on mountain pastures. Comput. Electron. Agric. 2009, 65, 85–92, doi: 10.1016/J.COMPAG.2008.07.010.

7. Hill, T.M.; Suarez-Mena, F.X.; Hu, W.; Dennis, T.S.; Schlotterbeck, R.L.; Timms, L.L.; Hulbert, L.E. Technical Note: Evaluation of an ear-attached

movement sensor to record rumination, eating, and activity behaviors in 1-month-old calves. Prof. Anim. Sci. 2017, 33, 743–747, doi:10.15232/pas.2017-01623.

8. Miller-Cushon, E.K.; Devries, T.J. Invited review: Development and expression of dairy calf feeding behaviour. Can. J. Anim. Sci. 2015, 95, 341–350, doi:10.4141/CJAS-2014-163.

9. Williams, L.R.; Moore, S.T.; Bishop-Hurley, G.J.; Swain, D.L. A sensor-based solution to monitor grazing cattle drinking behaviour and water intake. Comput. Electron. Agric. 2020, 168, 105141, doi: 10.1016/J.COMPAG.2019.105141.

10. Roland, L.; Lidauer, L.; Sattlecker, G.; Kickinger, F.; Auer, W.; Sturm, V.; Efrosinin, D.; Drillich, M.; Iwersen, M. Monitoring drinking behavior in bucket-fed dairy calves using an ear-attached tri-axial accelerometer: A pilot study. Comput. Electron. Agric. 2018, 145, 298–301, doi: 10.1016/J.COMPAG.2018.01.008.

11. Rayas-Amor, A.A.; Morales-Almaráz, E.; Licona-Velázquez, G.; Vieyra-Alberto, R.; García-Martínez, A.; Martínez-García, C.G.; Cruz-Monterrosa, R.G.; Miranda-de la Lama, G.C. Triaxial accelerometers for recording grazing and ruminating time in dairy cows: An alternative to visual observations. J. Vet. Behav. 2017, 20, 102–108, doi: 10.1016/J.JVEB.2017.04.003.

12. Tamura, T.; Okubo, Y.; Deguchi, Y.; Koshikawa, S.; Takahashi, M.; Chida, Y.; Okada, K. Dairy cattle behavior classifications based on decision tree learning using 3-axis neck-mounted accelerometers. Anim. Sci. J. 2019, 90, 589–596, doi:10.1111/ASJ.13184.

13. Riaboff, L.; Couvreur, S.; Madouasse, A.; Roig-Pons, M.; Aubin, S.; Massabie, P.; Chauvin, A.; Bédère, N.; Plantier, G. Use of predicted behavior from accelerometer data combined with GPS data to explore the relationship between dairy cow behavior and pasture characteristics. Sensors 2020, 20, 4741, doi:10.3390/s20174741.

14. Sprinkle, J.E.; Sagers, J.K.; Hall, J.B.; Ellison, M.J.; Yelich, J.V.; Brennan, J.R.; Taylor, J.B.; Lamb, J.B. Predicting cattle grazing behavior on rangeland using accelerometers. Rangel. Ecol. Manag. 2021, 76, 157–170, doi: 10.1016/J.RAMA.2020.10.001.

15. Hearnshaw, H.; Morris, C. Genetic and environmental effects on a temperament score in beef cattle. Aust. J. Agric. Res. 1984, 35, 723–733, doi:10.1071/AR9840723.

16. Zheng, Z.; Cai, Y.; Li, Y. Oversampling method for imbalanced classification. Comput. Inform. 2015, 34, 1017–1037.

17. Haixiang, G.; Yijing, L.; Shang, J.; Mingyun, G.; Yuanyue, H.; Bing, G. Learning from class-imbalanced data: Review of methods and applications. Expert Syst. Appl. 2017, 73, 220–239, doi: 10.1016/J.ESWA.2016.12.035.

18. Yijing, L.; Haixiang, G.; Xiao, L.; Yanan, L.; Jinling, L. Adapted ensemble classification algorithm based on multiple classifier system and feature selection for classifying multi-class imbalanced data. Knowl. -Based Syst. 2016, 94, 88–104, doi: 10.1016/J.KNOSYS.2015.11.013.

19. López, V.; Fernández, A.; García, S.; Palade, V.; Herrera, F. An insight into classification with imbalanced data: Empirical results and current trends on using data intrinsic characteristics. Inf. Sci. 2013, 250, 113–141, doi: 10.1016/j.ins.2013.07.007.

20. RStudio. Team RStudio: Integrated Development Environment for R. 2020. Available online: https://www.rstudio.com/categories/integrated-development-environment/ (accessed on 02 January 2020).

21. Alvarenga, F.A.P.; Borges, I.; Palkovič, L.; Rodina, J.; Oddy, V.H.; Dobos, R.C. Using a three-axis accelerometer to identify and classify sheep behaviour at pasture. Appl. Anim. Behav. Sci. 2016, 181, 91–99, doi: 10.1016/J.APPLANIM.2016.05.026.

22. Breiman, L. Random forests. Mach. Learn. 2001, 45, 5–32, doi:10.1023/A:1010933404324.

23. Vapnik, V.; Guyon, I.; Hastie, T. Support vector machines. Mach. Learn. 1995, 20, 273–297.

24. Hand, D.J.; Yu, K. Idiot's Bayes—Not so stupid after all? Int. Stat. Rev. 2001, 69, 385–398, doi:10.1111/J.1751-5823. 2001.TB00465X.

25. Liaw, A.; Wiener, M. Classification and regression by randomForest. R News 2002, 2, 18–22.

26. Meyer, D.; Dimitriadou, E.; Hornik, K.; Weingessel, A.; Leisch, F.; Chang, C.C.; Lin, C.C. e1071: Misc Functions of the De-partment of Statistics, Probability Theory Group (Formerly: E1071), TU Wien [R package version 1.7-9]. Comprehensive R Archive Network (CRAN). 2021. Available online: https://CRAN.R-project.org/package=e1071 (accessed on 5 August 2021). Meyer, D.; Dimitriadou, E.; Hornik, K.; Weingessel, A.; Leisch, F.; Chang, C.C.; Lin, C.C. Package 'e1071'. R Journal 2019, 1–66.

27. Kraemer, H.C. Extension of the kappa coefficient. Biometrics 1980, 36, 207–216.

28. Kuhn, M. Building predictive models in R using the caret package. J. Stat. Softw. 2008, 28, 1–26, doi:10.18637/JSS.V028.I05.

29. Chawla, N.V.; Bowyer, K.W.; Hall, L.O.; Kegelmeyer, W.P. SMOTE: Synthetic minority over-sampling technique. J. Artif. Intell. Res. 2002, 16, 321–357, doi:10.1613/JAIR.953.

30. Tahir, M.A.; Kittler, J.; Mikolajczyk, K.; Yan, F. A multiple expert approach to the class imbalance problem using inverse random under sampling. In Lecture Notes in Computer Science, Proceedings of the 8th International Workshop on Multiple Classifier Systems (MCS), Reykjavik, Iceland, 10–12 June 2009; Benediktsson, J.A., Kittler, J., Roli, F., Eds.; Springer: Berlin/Heidelberg, Germany, 2009; Volume 5519, pp. 82–91.

31. Tatler, J.; Cassey, P.; Prowse, T.A.A. High accuracy at low frequency: Detailed behavioural classification from accelerometer data. J. Exp. Biol. 2018, 221, jeb184085, doi:10.1242/JEB.184085.

32. Gjoreski, H.; Bizjak, J.; Gjoreski, M.; Gams, M. Comparing deep and classical machine learning methods for human activity recognition using wrist accelerometer. In Proceedings of the IJCAI-16 Workshop on Deep Learning for Artificial Intelligence (DLAI), New York, NY, USA, 10 July 2016.

33. Körting, T.S.; Garcia Fonseca, L.M.; Câmara, G. GeoDMA—Geographic data mining analyst. Comput. Geosci. 2013, 57, 133–145, doi: 10.1016/J.CAGEO.2013.02.007.

34. Belgiu, M.; Drăgu, L. Random Forest in remote sensing: A review of applications and future directions. ISPRS J. Photogramm. Remote Sens. 2016, 114, 24–31, doi: 10.1016/J.ISPRSJPRS.2016.01.011.

35. Sheykhmousa, M.; Mahdianpari, M.; Ghanbari, H.; Mohammadimanesh, F.; Ghamisi, P.; Homayouni, S. Support vector machine versus random forest for remote sensing image classification: A meta-analysis and systematic review. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 2020, 13, 6308–6325, doi:10.1109/JSTARS.2020.3026724.

36. Benaissa, S.; Tuyttens, F.A.M.; Plets, D.; Cattrysse, H.; Martens, L.; Vandaele, L.; Joseph, W.; Sonck, B. Classification of inges-tive-related cow behaviours using RumiWatch halter and neck-mounted accelerometers. Appl. Anim. Behav. Sci. 2019, 211, 9–16, doi: 10.1016/J.APPLANIM.2018.12.003.

37. Douglas, P.K.; Harris, S.; Yuille, A.; Cohen, M.S. Performance comparison of machine learning algorithms and number of independent components used in fMRI decoding of belief vs. disbelief. Neuroimage 2011, 56, 544–553, doi: 10.1016/J.NEUROIMAGE.2010.11.002.

38. Escalante, H.J.; Rodriguez, S.V.; Cordero, J.; Kristensen, A.R.; Cornou, C. Sow-activity classification from acceleration patterns: A machine learning approach. Comput. Electron. Agric. 2013, 93, 17–26, doi: 10.1016/J.COMPAG.2013.01.003.

39. Delagarde, R.; Lamberton, P. Daily grazing time of dairy cows is recorded accurately using the Lifecorder Plus device. Appl. Anim. Behav. Sci. 2015, 165, 25–32, doi: 10.1016/J.APPLANIM.2015.01.014.

40. Dutta, R.; Smith, D.; Rawnsley, R.; Bishop-Hurley, G.; Hills, J.; Timms, G.; Henry, D. Dynamic cattle behavioural classification using supervised ensemble classifiers. Comput. Electron. Agric. 2015, 111, 18–28, doi: 10.1016/J.COMPAG.2014.12.002.

41. Watanabe, N.; Sakanoue, S.; Kawamura, K.; Kozakai, T. Development of an automatic classification system for eating, ru-minating and resting behavior of cattle using an accelerometer. Grassl. Sci. 2008, 54, 231–237, doi:10.1111/J.1744-697X.2008.00126X.

42. Kamminga, J.W.; Meratnia, N.; Havinga, P.J.M. Dataset: Horse movement data and analysis of its potential for activity recognition. In Proceedings of the 2nd Workshop on Data Acquisition to Analysis (DATA 2019), Prague, Czech Republic, 26–28 July 2019; pp. 22–25.

43. Martiskainen, P.; Järvinen, M.; Skön, J.P.; Tiirikainen, J.; Kolehmainen, M.; Mononen, J. Cow behaviour pattern recognition using a three-dimensional accelerometer and support vector machines. Appl. Anim. Behav. Sci. 2009, 119, 32–38, doi:10.1016/J.APPLANIM.2009.03.005.

44. Ravi, N.; Dandekar, N.; Mysore, P.; Littman, M.L. Activity recognition from accelerometer data. In Proceedings of the 17th Conference on Innovative Applications of Artificial Intelligence (IAAI-17), Pittsburgh, PA, USA, 9–13 July 2005; pp. 1541–1546.

45. Zughrat, A.; Mahfouf, M.; Yang, Y.Y.; Thornton, S. Support vector machines for class imbalance rail data classification with bootstrapping-based over-sampling and under-sampling. In Proceedings of the 19th World Congress of the International Federation of Automatic Control (IFAC), Cape Town, South Africa, 24–29 August 2014; Volume 47, pp. 8756–8761.

46. Bermejo, P.; Gámez, J.A.; Puerta, J.M. Improving the performance of Naive Bayes multinomial in e-mail foldering by in-troducing distribution-based balance of datasets. Expert Syst. Appl. 2011, 38, 2072–2080, doi: 10.1016/J.ESWA.2010.07.146

Appendices

Appendix 1. Differentially expressed genes

Differential Expression Analysis			
NWCF			
Up-regulated genes			
Ensembl gene id	**Gene Symbol**	**log2FoldChange**	**Adjusted Pvalue**
ENSBTAG00000014441	CD96	0.728610479	0.00084
ENSBTAG00000044419	-	3.61887337	0.06344
ENSBTAG00000051030	CYP2J2	5.159699616	0.000186
ENSBTAG00000052514	LOC100336869	2.162794413	7.70E-10
ENSBTAG00000052658	NKG2A	1.521539519	0.000352
ENSBTAG00000046389	KLRD1	1.368787237	7.97E-06
ENSBTAG00000004263	ATP6V0A4	0.952547715	0.017114
ENSBTAG00000006497	-	1.727581868	0.06344
ENSBTAG00000007201	TMEM242	0.648459821	0.001592
ENSBTAG00000004641	PLEKHA6	1.477179667	0.001553
ENSBTAG00000000297	ZNF330	0.594963804	0.004862
ENSBTAG00000013055	LOC786126	0.717344068	0.06344
ENSBTAG00000021438	ZNF821	0.767318435	0.06344
ENSBTAG00000045529	NCR1	1.143185456	0.00084
ENSBTAG00000053383	-	1.441907999	0.022235
ENSBTAG00000039326	LOC520336	0.893555714	0.06344
ENSBTAG00000039513	GCNT2	1.266752909	0.005486
ENSBTAG00000054008	LOC521568	3.285776327	0.06344
Down-regulated genes			
ENSBTAG00000049044	-	-3.168149872	0.068373
ENSBTAG00000014750	EPB41L4A	-2.134740718	0.095572
ENSBTAG00000002006	THBS1	-1.459268623	0.000653
ENSBTAG00000037832	ASPHD2	-0.799635556	0.00877
ENSBTAG00000017305	-	-0.814591071	0.0046
ENSBTAG00000055240	-	-0.957017775	0.000557
ENSBTAG00000050871	LOC100139881	-2.173800184	0.030507
ENSBTAG00000052414	-	-2.341079713	6.07E-16
ENSBTAG00000054086	-	-0.849490309	0.095572
NGA			
Up-regulated genes			
ENSBTAG00000004587	DUSP6	0.588085353	0.040476
ENSBTAG00000038375	STOM	0.701346954	0.033554
ENSBTAG00000048122	CFD	0.592988003	0.039602
ENSBTAG00000039340	SCN4B	1.784315536	0.098326
ENSBTAG00000002411	CTNND1	0.889660349	0.082081
ENSBTAG00000038650	MS4A7	0.741645785	0.07309
ENSBTAG00000000598	CST3	0.611216938	0.048506
ENSBTAG00000054085	SELENOP	0.654369902	0.043207
ENSBTAG00000000336	LOC112442378	0.928456268	0.098326

ENSBTAG00000050944	LOC100848752	0.937352233	0.000593
ENSBTAG00000039413	-	0.998537744	0.002453
ENSBTAG00000003074	SLC16A6	0.711354471	0.036168
ENSBTAG00000015829	ENPP5	0.800612317	0.033554
Down-regulated genes			
ENSBTAG00000001257	AGTPBP1	-1.02866366	0.012043
ENSBTAG00000054702	-	-1.044900475	0.048506
F1WCF			
Up-regulated genes			
ENSBTAG00000051633	-	1.601546755	0.000907
Down-regulated genes			
ENSBTAG00000055211	-	-2.637713957	0.033304
ENSBTAG00000021160	AP1M2	-2.1069205	0.098319
ENSBTAG00000006566	FFAR2	-1.871935337	0.057678
ENSBTAG00000053421	-	-1.601824277	0.007536
F1GA			
Up-regulated genes			
ENSBTAG00000017500	KCNK12	3.178081285	0.087072
ENSBTAG00000053653	-	3.579374663	0.074815
Down-regulated genes			
ENSBTAG00000010155	LOC504773	-1.720935297	0.008245
ENSBTAG00000053829	-	-0.810111731	6.35E-17
ENSBTAG00000024901	LOC509881	-1.790103584	0.049863

Appendix 2. Differentially dispersed genes

Differential Dispersion Analysis			
NWCF			
Up-regulated genes			
Ensembl gene id	**Gene Symbol**	**Log2FoldChange**	**Adjusted Pvalue**
ENSBTAG00000054647	RPE	2.823837387	0.016256147
ENSBTAG00000015776	MFSD14A	5.019901968	0.009704501
ENSBTAG00000033735	WASHC4	5.662035904	4.27E-05
ENSBTAG00000000260	ZNRF2	4.361857459	0.096280671
ENSBTAG00000004329	RELL1	4.571787431	0.096280671
ENSBTAG00000014522	SMIM14	4.598266016	0.048585315
ENSBTAG00000046503	NCOA7	4.741277661	0.023056337
ENSBTAG00000044420	-	2.35430165	0.058056248
ENSBTAG00000002271	CDADC1	4.640616442	0.041209631
ENSBTAG00000015334	ZHX1	4.94420076	0.013676184
ENSBTAG00000048843	UBE2W	5.436318277	0.00152099
ENSBTAG00000021240	DCSTAMP	4.546263452	0.064745861
ENSBTAG00000019581	PLEKHF2	5.894724797	0.000130008
ENSBTAG00000008438	ARFIP1	4.810547053	0.024329349

ENSBTAG00000018200	IL15	4.706725758	0.034648569
Down-regulated			
ENSBTAG00000049044	-	-5.916083776	0.004045711
ENSBTAG00000003718	HACL1	-5.734477896	0.000611895
ENSBTAG00000008552	-	-5.391626492	0.001721351
ENSBTAG00000003458	CDCA7	-6.22956489	3.58E-05
ENSBTAG00000006103	XRCC6	-5.273969396	0.076797317
ENSBTAG00000003531	ABCA13	-6.968531233	0.036229607
ENSBTAG00000051942	-	-4.888025117	0.016256147
ENSBTAG00000030453	ZC3H14	-5.397291049	0.00152099
ENSBTAG00000016568	LUC7L	-6.203865815	3.58E-05
ENSBTAG00000002374	ARMC5	-4.149826269	0.066983552
NGA			
Up-regulated genes			
ENSBTAG00000026758	LOC788425	4.963047193	0.016270074
ENSBTAG00000019033	CD84	5.445572389	0.001495545
ENSBTAG00000020294	PTPN6	6.277528661	0.048771622
ENSBTAG00000007305	VPS41	4.734108786	0.02926821
ENSBTAG00000000296	IMPDH1	4.460931086	0.078837737
ENSBTAG00000021130	MFSD10	6.666587565	0.016270074
ENSBTAG00000001137	CLTA	4.566433947	0.053643797
ENSBTAG00000018436	UBE2J1	4.570036395	0.059719544
ENSBTAG00000046819	-	3.095602422	5.10E-161
ENSBTAG00000007079	LCP1	6.476134859	0.028405009
ENSBTAG00000004863	RIC3	2.856645134	4.65E-11
ENSBTAG00000014612	DOCK2	5.225520839	9.75E-08
ENSBTAG00000038627	-	1.409690209	0.001415087
ENSBTAG00000010668	GEMIN7	4.600951855	0.050448835
ENSBTAG00000008518	MED25	4.878487537	0.01712087
ENSBTAG00000009354	EVI2A	5.873821748	0.029029981
ENSBTAG00000012024	SLC29A2	4.56523316	0.059719544
Down-regulated genes			
ENSBTAG00000051663	-	-2.444458453	0.078837737
ENSBTAG00000003901	DNMT3B	-5.615002787	0.000765174
ENSBTAG00000020166	ZNFX1	-6.862162465	0.01126434
ENSBTAG00000052247	TPPP	-6.572696596	9.56E-06
ENSBTAG00000015019	STN1	-5.648132528	0.027270518
ENSBTAG00000048950	ULBP11; ULBP13; ULBP15	-5.121393642	0.010908714
F1WCF			
Up-regulated genes			
ENSBTAG00000004557	DHRS9	4.890134792	0.025389782
ENSBTAG00000048514	-	5.782740431	0.001727784
ENSBTAG00000019478	NAAA	5.252105447	0.00305743
ENSBTAG00000053318	CLEC17A	5.476629636	0.002660172
ENSBTAG00000002148	RAB3D	6.302594873	1.74E-05

ENSBTAG00000004920	LOC783202	5.100910742	0.027605586
ENSBTAG00000010057	LOC508646	6.919468913	0.025859763
ENSBTAG00000020936	NOD2	4.501264304	0.081138475
ENSBTAG00000026375	RMI2	3.011995359	0.000183703
Down-regulated genes			
ENSBTAG00000018986	ACSL4	-4.452531371	0.091059697
ENSBTAG00000048244	LOC784052	-7.552192574	0.001306664
ENSBTAG00000005861	CCNT1	-4.801526611	0.027184543
ENSBTAG00000054128	-	-7.925732956	0.014713704
ENSBTAG00000031395	C6H4orf48	-2.136652279	9.07E-06
ENSBTAG00000044021	TOGARAM2	-4.994509429	0.021676158
ENSBTAG00000000653	PPP1R16B	-6.343022447	1.37E-05
ENSBTAG00000012450	RAPGEF2	-4.42115879	0.027605586
ENSBTAG00000007272	ATP6V0A2	-5.037157251	0.007223495
ENSBTAG00000007172	GOT2	-5.401564623	0.001674875
ENSBTAG00000053747	-	-5.156504251	9.56E-05
ENSBTAG00000018937	HMOX2	-5.129961776	0.005663025
F1GA			
Up-regulated genes			
ENSBTAG00000044021	TOGARAM2	5.185050658	0.015129994
ENSBTAG00000052568	-	6.307993895	0.000498714
ENSBTAG00000000728	-	2.94399357	0.001664157
Down-regulated genes			
ENSBTAG00000015212	IFNAR2	-6.55954641	1.93E-06
ENSBTAG00000013937	SPICE1	-4.447656731	0.070574226
ENSBTAG00000014337	EIF2S3	-5.873114375	0.085352468
ENSBTAG00000009949	DARS1	-4.811296202	0.017601201
ENSBTAG00000005891	FCRL1	-6.476618661	2.46E-06
ENSBTAG00000004987	NUP210L	-4.671713311	0.033197033
ENSBTAG00000023551	MOB3C	-4.881408388	0.014432308
ENSBTAG00000010136	CMAS	-4.961206264	0.011655864
ENSBTAG00000051809	STMP1	-4.670338441	0.031992409
ENSBTAG00000016070	LINGO2	-2.35008332	0.044833862
ENSBTAG00000025859	ZCCHC7	-4.357967769	0.084302238
ENSBTAG00000007659	CORO2A	-6.52987414	2.09E-06
ENSBTAG00000011824	OGN	-4.428098255	0.091078286
ENSBTAG00000020498	ACP1	-4.359983047	0.084302238
ENSBTAG00000040334	MEX3D	-4.351348886	0.085352468
ENSBTAG00000054976	MCTP1	-2.763188329	2.65E-16
ENSBTAG00000011176	TPRKB	-5.017059787	0.011655864
ENSBTAG00000034662	-	-4.557971893	0.084302238
ENSBTAG00000009065	MAP3K5	-4.607679806	0.037609857
ENSBTAG00000005851	VTA1	-4.289347106	0.09799541
ENSBTAG00000011608	KIAA1191	-3.280085853	6.03E-07
ENSBTAG00000000371	DIS3L	-6.29116263	0.037609857
ENSBTAG00000039462	PCLAF	-6.094412718	0.083117618

ENSBTAG00000000988	BRCA2	-4.592289229	0.051874593
ENSBTAG00000008310	TMEM9B	-5.535395868	0.000666038
ENSBTAG00000012417	ARL14EP	-4.7147965	0.03082586
ENSBTAG00000049036	-	-5.323310968	0.015129994
ENSBTAG00000004920	LOC783202	-5.392180702	0.009177829
ENSBTAG00000016513	ZNF331	-4.456604378	0.07624006
ENSBTAG00000005008	WSB1	-5.962957753	0.084302238
ENSBTAG00000007409	-	-4.338207794	0.085352468
ENSBTAG00000034883	CCR4	-4.376020497	0.084302238
ENSBTAG00000008099	CCR5	-5.037311402	0.014432308
ENSBTAG00000015571	GCLC	-4.699883447	0.031346682
ENSBTAG00000020595	KCTD20	-5.890945839	0.084302238
ENSBTAG00000003892	LOC537017	-5.795881037	0.000214583
ENSBTAG00000007847	EED	-6.025608232	0.011655864
ENSBTAG00000012654	CHCHD1	-4.356609334	0.084302238

Appendix 3. Male Nelore with an accelerometer attached to the custom halter in the mandibular region.

Appendix 4. Description for each animal behavior studied.

Animal behavior	Description
Grazing	Animals searching for food while walking short distances with their head down, without picking food up with their mouth; standing still with their head down while apprehending food with their mouth; and chewing either with their head down or their head up, while stationary.
Ruminating	Animals chewing and swallowing a ruminal bolus.
Idle	Animals being idle. Lying down in any resting position or standing up on all four legs, without locomotion.
WCF	Animals attending the water fountain.
Feeding	Animals located in the feeding supplement zone, ingesting dietary supplement.
Walking	Animals walking from one position to another, without searching for food or ruminating.

WCF: Water Consumption Frequency.

Appendix 5. Predictor variables created from x, y, and z axis acceleration variables

Predictor variables	**Equation**
SMA	$\lvert X_i\rvert + \lvert Y_i\rvert + \lvert Z_i\rvert$
SVM	$\sqrt{X_i^2 + Y_i^2 + Z_i^2}$
Movement variation	$\lvert X_{i+1} - X_i\rvert + \lvert Y_{i+1} - Y_i\rvert + \lvert Z_{i+1} - Z_i\rvert$
Energy	$(X_i^2 + Y_i^2 + Z_i^2)^2$
Entropy	$\big(1 + (X_i + Y_i + Z_i)\big)^2 x \ln(1 + (X_i + Y_i + Z_i)^2)$
Pitch (degrees)	$\tanh^{-1}\left(-X_i / \left(\sqrt{Y_i^2 + Z_i^2}\right)\right) x\ 180/\pi$
Roll (degrees)	$atan2(Y_i Z_i) x\ 180/\pi$
Inclination (degrees)	$\tan^{-1}\left((X_i^2 + Y_i^2)/Z_i\right) x\ 180/\pi$

SMA – signal magnitude area; SVM – signal vector magnitude [22]