UNIVERSIDADE ESTADUAL PAULISTA
"JÚLIO DE MESQUITA FILHO"
Campus de Araçatuba

MARCO ANTONIO PERPÉTUO DE SOUSA

Polimorfismos de nucleotídeo único afetam a predição de alvos de microRNAs em bovinos

Araçatuba
2019

MARCO ANTONIO PERPÉTUO DE SOUSA

Polimorfismos de nucleotídeo único afetam a predição de alvos de microRNAs em bovinos

Dissertação apresentada à Faculdade de Medicina Veterinária de Araçatuba da Universidade Estadual Paulista "Júlio de Mesquita Filho" – UNESP, como parte dos requisitos para obtenção do título de Mestre em Ciência Animal.

Orientadora: Flávia Lombardi Lopes

Araçatuba
2019

S725p

Sousa, Marco Antonio Perpétuo de
Polimorfismos de nucleotídeo único afetam a predição de alvos de microRNAs em bovinos / Marco Antonio Perpétuo de Sousa. -- Araçatuba, 2019
47 p. : il., tabs.

Dissertação (mestrado) - Universidade Estadual Paulista (Unesp), Faculdade de Medicina Veterinária, Araçatuba
Orientador: Flávia Lombardi Lopes

1. MicroRNA. 2. Polimorfismos de nucleotídeo único. 3. Bovino. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca da Faculdade de Medicina Veterinária, Araçatuba. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

UNIVERSIDADE ESTADUAL PAULISTA

Câmpus de Araçatuba

CERTIFICADO DE APROVAÇÃO

Título: Polimorfismos de nucleotídeo único afetam a predição de alvos de microRNAs em bovinos

AUTOR: MARCO ANTONIO PERPÉTUO DE SOUSA
ORIENTADORA: FLÁVIA LOMBARDI LOPES

Aprovado como parte das exigências para obtenção do Título de Mestre em CIÊNCIA ANIMAL, área: Medicina Veterinária Preventiva e Produção Animal pela Comissão Examinadora:

Pesquisadora FLÁVIA LOMBARDI LOPES
Departamento de Apoio, Produção e Saúde Animal / Faculdade de Medicina Veterinária - Câmpus de Araçatuba/Unesp

Profa. Dra. MARINA RUFINO SALINAS FORTES
The University of Queensland

Prof. Dr. ADAM TAITI HARTH UTSUNOMIYA
Doutor em Genética pela Faculdade de Ciências Agrárias e Veterinárias - Câmpus de Jaboticabal/Unesp

Araçatuba, 27 de maio de 2019.

UNIVERSIDADE ESTADUAL PAULISTA

Câmpus de Araçatuba

CERTIFICADO DE APROVAÇÃO

Título: Polimorfismos de nucleotídeo único afetam a predição de alvos de microRNAs em bovinos

AUTOR: MARCO ANTONIO PERPÉTUO DE SOUSA
ORIENTADORA: FLÁVIA LOMBARDI LOPES

Aprovado como parte das exigências para obtenção do Título de Mestre em CIÊNCIA ANIMAL, área: Medicina Veterinária Preventiva e Produção Animal pela Comissão Examinadora:

Pesquisadora FLÁVIA LOMBARDI LOPES
Departamento de Apoio, Produção e Saúde Animal / Faculdade de Medicina Veterinária - Câmpus de Araçatuba/Unesp

Profa. Dra. MARINA RUFINO SALINAS FORTES
The University of Queensland

Prof. Dr. ADAM TAITI HARTH UTSUNOMIYA
Doutor em Genética pela Faculdade de Ciências Agrárias e Veterinárias - Câmpus de Jaboticabal/Unesp

Araçatuba, 27 de maio de 2019.

À minha família, com muito amor e carinho, por sua intensa dedicação, e apoio durante a elaboração deste trabalho.

Aos meus filhos Melissa e Nicholas por serem fonte de energia e motivação.

Agradecimentos

À Universidade Estadual Paulista “Júlio de Mesquita Filho”, em especial ao Programa de Pós Graduação em Ciência Animal da Faculdade de Medicina Veterinária – Campus de Araçatuba.

A todos os professores, alunos e colaboradores do Departamento do Apoio, Produção e Saúde Animal (DAPSA).

À Prof. Dra. Flávia Lombardi Lopes pela atenção, dedicação, inspiração e apoio durante o processo de orientação.

Aos meus colegas do Laboratório de Epigenômica pelo apoio, dedicação e companheirismo, que tornaram esta realização possível.

“A ciência nunca resolve um problema sem criar pelo menos outros dez”.

George Bernard Shaw

SOUSA, M. A. P. **Polimorfismos de nucleotídeo único afetam a predição de alvos de microRNAs em bovinos. 2019.** 47 f. Dissertação (Mestrado) – Faculdade de Medicina Veterinária, Universidade Estadual Paulista, Araçatuba, 2019.

RESUMO

O melhoramento genético em bovinos visa a seleção de características para facilitar o manejo, a qualidade da carne, a resistência a doenças e a adaptação ao meio ambiente. Polimorfismos de nucleotídeo único (SNPs) podem gerar grandes efeitos sobre essas características fenotípicas. Os microRNAs são pequenos RNAs não-codificadores que atuam como reguladores da expressão pós-transcricional através de sua ligação a mRNAs alvo. No presente estudo, realizamos o cruzamento de dados entre ~56 milhões de SNPs contra todas as seqüências conhecidas de miRNA bovino e analisamos *in silico*, seus possíveis efeitos. Seguindo a predição dos alvos, mostramos que 82% dos alvos foram alterados como consequência dos SNPs que ocorrem na região de *seed* de miRNAs maduros. Em seguida, identificamos variações na Energia Livre Mínima (MFE) que representam a capacidade de alterar a estabilidade das moléculas e, consequentemente, a maturação dos miRNAs. Também encontramos 129 SNPs em miRNAs, que alteraram sua predição com alvos, ocorrendo em regiões de QTL e, por último, a análise dos escores de conservação evolutiva para cada locus de SNP sugeriu que eles têm uma função biológica conservada através do processo evolutivo. Nossos resultados sugerem que os SNPs em microRNAs têm o potencial de alterar os fenótipos bovinos e são de grande valor para a pesquisa de melhoramento genético, bem como para a produção.

Palavras-chave: microRNA, SNP, bovino.

SOUSA, M. A. P. **Single nucleotide polimorfisms affect miRNA target prediction in bovine. 2019.** 51 f. Dissertação (Mestrado) – Faculdade de Medicina Veterinária, Universidade Estadual Paulista, Araçatuba, 2019.

ABSTRACT

Genetic improvement of cattle is aimed at selection of characteristics to facilitate the handling, quality of the meat, resistance to diseases and adaptation to the environment. Single nucleotide polymorphisms (SNPs) can generate large effects on these phenotypic characteristics. MicroRNAs are small non-coding RNAs that act as regulators of post-transcriptional expression through their binding to target mRNAs. In the present study, we scanned ~56 million SNPs against all known bovine miRNA sequences and analyzed *in silico*, their possible effects. Following target prediction, we show that 82% of targets were altered as a consequence of SNPs that occur in the seed region of mature miRNAs. Next, we identified variations in the Minimum Free Energy (MFE) which represent the capacity to alter molecule stability and, consequently, the maturation of the miRNAs. We have also found 129 SNPs in miRNAs, with altered target prediction, occurring in QTL regions and, lastly, analysis of evolutionary conservation scores for each SNP locus suggested that they have a conserved biological function through the evolutionary process. Our results suggest that SNPs in microRNAs have the potential to alter bovine phenotypes and are of great value for genetic improvement research, as well as production.

Keywords: microRNA, SNP, bovine.

SUMÁRIO

1 INTRODUÇÃO GERAL

1.1 REVISÃO DE LITERATURA

A domesticação de animais e plantas teve início no neolítico e permitiu que as populações humanas, antes nômades, pudessem estabelecer-se formando civilizações, com abundância de alimentos e segurança. Dentre os rebanhos, o bovino destacou-se, pois, sua força de tração e uso em rituais religiosos elevaram sua importância para além do fornecimento de leite, peles e carne. Devido a esta diversidade de vantagens que a criação bovina trouxe ao homem, a preocupação com características que facilitem o manejo, alojamento e pastagem sempre se fez presente, induzindo a seleção artificial destes animais, modificando-os através do tempo (AJMONE-MARSAN et al., 2010)(TANG; HO, 2007).

A compreensão dos processos biológicos em animais e plantas tem se expandido rapidamente através dos estudos relacionados ao sequenciamento genômico. (BENDER, 2004).

Conrad Waddinton usou o termo *epigenética* (epi: acima, sobre) pela primeira vez em meados de 1940 para relacionar as interações entre os genes e o ambiente, indicando as diferentes possibilidades de desenvolvimento de uma célula (NOBLE, 2015), e ao controle da expressão gênica considerando não apenas a sequência primária do DNA, mas sim, a organização da cromatina (BENDER, 2004). Deste modo, a epigenética revolucionou a genética molecular tornando-se imprescindível para o estudo da genômica funcional.

Para Levenson & Sweatt (JONATHAN M. LEVENSON; J.DAVID SWEATT, 2005) a epigenética se trata dos processos que regulam a expressão gênica sem afetar o código genético. Através do controle de sequências gênicas sem alterar a sequência nucleica em si, células geneticamente idênticas são capazes de se distinguir fenotipicamente, dependendo de sua localização e/ou função. As mudanças na expressão gênica podem ser herdadas pelo processo mitótico e ao longo das gerações. A programação epigenética dos gametas, e dos embriões em início de desenvolvimento, é condição vital para o desenvolvimento de um novo organismo (TANG; HO, 2007). Três processos epigenéticos envolvem esta

programação: a ação dos RNA não-codificadores, o remodelamento das histonas e a metilação do DNA. Estes eventos epigenéticos regulam a expressão gênica através do controle da transcrição e/ou tradução. Segundo Lund & Lohuizen (ANDERS H. LUND; MAARTEN VAN LOHUIZEN, 2004), o remodelamento das histonas e o padrão de metilação do DNA modificam a acessibilidade da cromatina para a regulação da transcrição localmente ou globalmente, pelas modificações no DNA e pelas modificações ou rearranjos dos nucleossomos. Já os RNAs não-codificadores podem atuar interferindo na transcrição e também na tradução de genes (TANG; HO, 2007).

1.1.2 MicroRNAs

MicroRNAs (miRNAs) são uma classe de RNA pequenos, não-codificadores, de cadeia simples, que atuam na repressão pós-transcricional, ligando-se na região 3' de genes alvo não traduzidos (ZHAO et al., 2012). Inúmeros processos biológicos são controlados por miRNAs, como o desenvolvimento e diferenciação celular dos organismos, controle do crescimento, apoptose, regulação da resposta imune, entre outros processos fisiológicos, além de processos patológicos como o câncer e doenças degenerativas (BUENO; DE CASTRO; MALUMBRES, 2008; FILIPOWICZ; BHATTACHARYYA; SONENBERG, 2008; KEDDE; AGAMI, 2008; STEFANI; SLACK, 2008; BARTEL, 2009; O'CONNELL; RAO; BALTIMORE, 2012).

Na biogênese dos miRNAs há duas vias: canônica e não canônica. Na via canônica os genes de miRNA localizam-se em regiões intergênicas ou em *clusters.* Na via não canônica os genes de miRNA localizam-se em regiões intrônicas.

Os miRNAs são transcritos pela RNA Polimerase II em um precursor primário chamado pri-miRNA, e na sequência processados, ainda no núcleo. Na via canônica, o pri-miRNA é processado pela enzima RNAse III Drosha e na via não canônica, por um spliceossoma. Em ambos os casos o pri-miRNA forma um *hairpin* de 70 nucleotídeos (SALIMINEJAD et al., 2018). Após transporte deste pré-miRNA para o citoplasma, outra enzima RNAse III, a Dicer, cliva o pré-miRNA em um duplex de aproximadamente 20 nucleotídeos, o tamanho final do miRNA maduro de fita simples (MAUDET et

al., 2014). Este miRNA duplex se associa à proteínas como Argonauta e TRNC6 formando o chamado complex RISC (do inglês “RNA-induced silencing complex”) para exercer a função de repressão de mRNA, que contenham complementariedade de sequência com os miRNAs, impedindo a tradução deste RNAm alvo, ou ainda promovendo a deadenilação e consequente degradação destes alvos (Figura 1) (HUNTZINGER; IZAURRALDE, 2011; EULALIO; SCHULTE; VOGE, 2012).

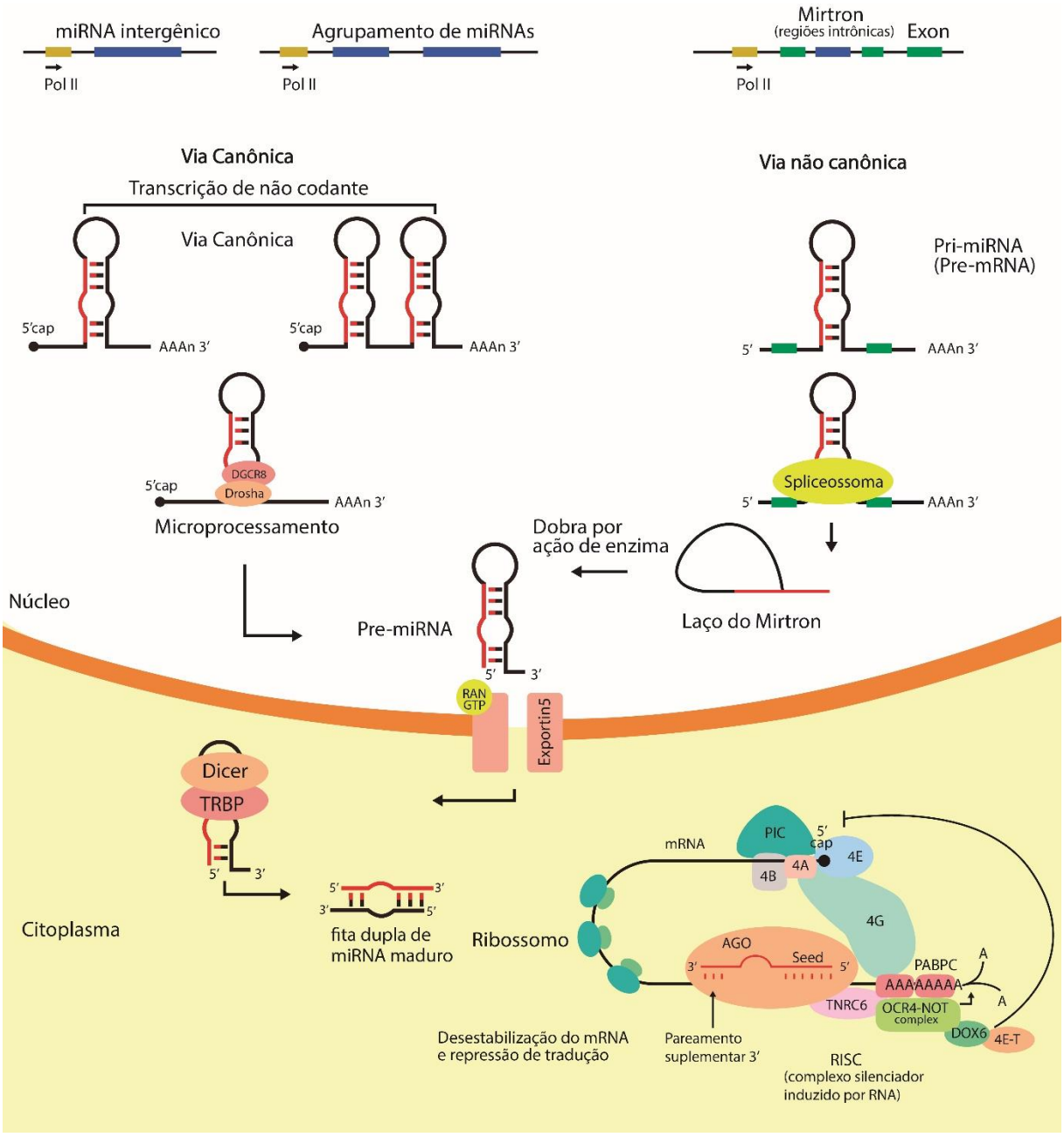


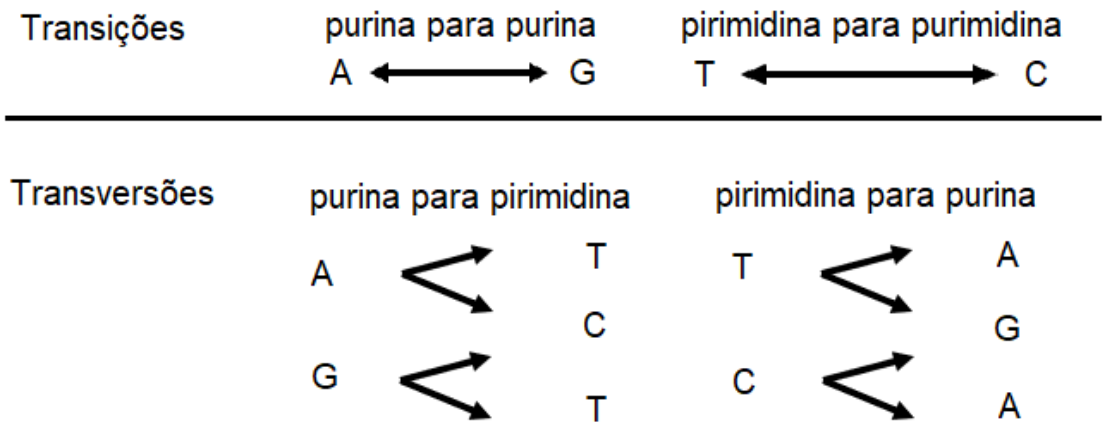
Figura 1 – Biogênese do miRNA e ligação com mRNA

Fonte: adaptado de SALIMINEJAD et. al., 2018

1.1.3 Polimorfismos de nucleotídeo único

Os Polimorfismos de nucleotídeo único (SNPs) representam a forma mais abundante de variação genética em ambos os genomas de plantas e animais (Jiang, 2008). Essas variações têm como base as alterações mais elementares da molécula de DNA, ou seja, mutações em bases únicas da cadeia de bases nitrogenadas (adenina (A) ou timina (T) ou citosina (C) ou guanina (G)) (CAETANO, 2009). Quimicamente, os nucleotídeos podem ser agrupados em purinas (A - G) e pirimidinas (C - T). SNPs dentro dos grupos são chamados transições e aqueles entre os grupos são chamados transversões. Portanto, há duas transições possíveis (C ↔ T e A ↔ G) e quatro transversões possíveis (A ↔ C, G ↔ T, A ↔ T, C ↔ G) (Figura 2), sem considerar as direções das mutações. As transições ocorrem mais frequentemente nos genomas de plantas e animais (JIANG et al., 2008).

Figura 2 – Diferenças entre transições e transversões



Fonte: Elaborado pelo autor

Meuwissen et al. em 2001 propôs o estudo denominado “Genome Wide Selection” (GWS) (MEUWISSEN, T. H. E.; HAYES, B. J.; GODDARD, 2001). Este modelo de estudo usa um grupo de genótipos definidos por um conjunto de SNPs para selecionar fenótipos de interesse através consideráveis cálculos matemáticos (SEIDEL, JR., 2010). Desta forma, estes polimorfismos são um recurso útil na criação de mapeamentos genômicos assim como estudos populacionais. Entre as vantagens do uso de SNPs

como marcadores genéticos em comparação com os microssatélites, pode-se destacar o fato de que ocorrem em grandes quantidades, são herdados de forma estável, e são facilmente testados com o uso da maioria das tecnologias disponíveis (BEUZEN; STEAR; CHANG, 2000).

Portanto, SNPs são importantes marcadores que vinculam os genes às mudanças fisiológicas normais, doenças, resposta a agentes patogênicos, respostas a produtos químicos, medicamentos, vacinas e outros agentes (RILEY et al., 2005; KIM; MISRA, 2007). O estudo dos SNPs também é importante em programas de melhoramento genético de plantas e bovinos, pois a informação obtida pode ser usada para localizar loci que afetam caracteres quantitativos, identificar regiões cromossômicas sob seleção, história da população estudada e caracterizar/gerenciar recursos genéticos e diversidade (RAFALSKI, 2000; DU; CLUTTER; LOHUIS, 2007).

Diferenças de nucleotídeos em regiões codantes podem alterar a composição final de proteínas, ou ainda SNPs nas regiões promotoras de genes podem causar ganhos/perdas de elementos de resposta e resultar em uma regulação diferencial da transcrição (JIANG et al., 2007). No entanto, o possível efeito de SNPs em controladores pós-transcricionais como miRNAs ainda necessita ser elucidado.

1.2 OBJETIVO

O presente trabalho tem como objetivo mapear os SNPs que ocorrem em sequências de microRNAs de bovinos e analisar seus possíveis efeitos na ligação com seus alvos e na estrutura primária do miRNA, visando a criação de um banco de dados de regiões de interesse para futura análise da expressão de alvos entre raças e/ou indivíduos polimórficos e fenotipicamente distintos.

2 - CAPÍTULO 1 - SINGLE NUCLEOTIDE POLYMORPHISMS AFFECT MIRNA TARGET PREDICTION IN BOVINE

Marco Antonio Perpétuo de Sousa[1], Flavia Regina Florencio de Athayde[1], Mariângela Bueno Cordeiro Maldonado[1], Natan Amorim Souza Gomes de Moraes[1], Gilberto Chiantelli Ferreira[1], Jeremy F. Taylor[2] Flavia Lombardi Lopes[1].

[1]Departament of support, production and animal health of Sao Paulo University State, Araçatuba, Brazil;

[2]Division of Animal Sciences, University of Missouri, Columbia, MO, USA

* Corresponding author: Flavia Lombardi Lopes

Tel: + 55 18 36360032

Address: School of Veterinary Medicine, Rua Clóvis Pestana, 793 - Jd. Dona Amélia,

Universidade Estadual Paulista (UNESP), Araçatuba, SP, Brazil, 16050-680

E-mail: flavia.lopes@unesp.br

2.1 Abstract

Background: Genetic improvement of cattle is aimed at selection of traits to facilitate the handling, quality of the meat, resistance to diseases and adaptation to the environment. Single nucleotide polymorphisms (SNPs) can generate large effects on these phenotypic characteristics. MicroRNAs are small non-coding RNAs that act as regulators of post-transcriptional expression through their binding to target mRNAs. In the present study, we scanned ~56 million SNPs against all known bovine miRNA sequences and analyzed *in silico*, their possible effects.

Results: Following target prediction, we show that 82% of targets were altered as a consequence of SNPs that occur in the seed region of mature miRNAs. Next, we identified variations in the Minimum Free Energy (MFE) which represent the capacity to alter molecule stability and, consequently, the maturation of the miRNAs 48.63% of the sequences analyzed showed values within those reported as sufficient to alter maturation. We have also found 129 SNPs in miRNAs, with altered target prediction, occurring in QTL regions and, lastly, analysis of evolutionary conservation scores for each SNP locus suggested that they have a conserved biological function through the evolutionary process.

Conclusions: Our results suggest that SNPs in microRNAs have the potential to alter bovine phenotypes and could be of great value for genetic improvement research, as well as production.

Keywords: microRNA, SNP, bovine

2.2 Background

Concerns about the genetic improvement of cattle dates back to the time of domestication of these animals, including characteristics that facilitate management, meat production, milk and adaptation to the environment[1]. Single nucleotide polymorphisms (SNPs) are single-base exchanges that occur naturally in the genome and have long been studied as valuable genetic markers, as they can affect phenotypic traits [2]. SNPs can occur in both genic and intergenic regions [3], with the former resulting in alterations that can be more easily related to consequent phenotype effects, i.e. changes in protein or alteration of cis-acting elements.

MicroRNAs (miRNA) are small, non-coding RNAs of approximately 22 nucleotides that act as post-transcriptional regulators of gene expression, binding to 3' UTR target genes by antisense complementarity in a conformation known as RISC (RNA-induced silence complex) [4]. Following complex formation, there are two ways of translational control: target miRNA cleavage or translational inhibition [5].

In humans, the occurrence of SNPs at target sites of miRNAs on messenger RNAs (mRNAs) may alter the mRNA:miRNA binding, creating or destroying targets, and potentially regulating a wide variety of diseases, including cancer [6]. Notwithstanding, the presence of SNPs can also affect the miRNA sequence itself [7][8]. In bovine, the effects of SNPs in specific miRNA sequences or miRNA target sites were reported as being related to mastitis [9] and fertility [10].

Through the use of bioinformatic tools and the largest bovine SNP database currently available (1000 Bulls Genome Project), we sought to

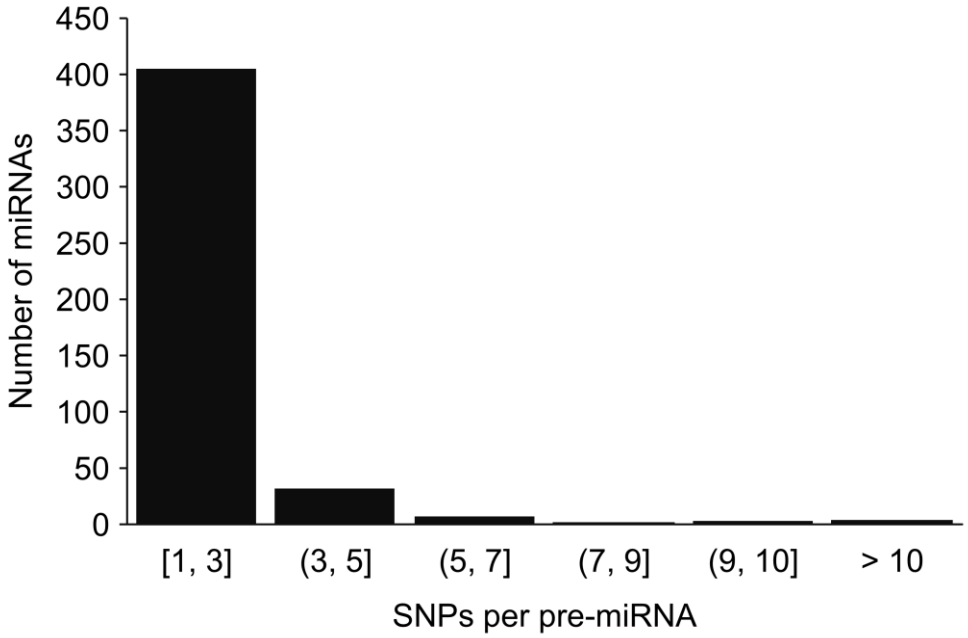
identify SNPs that occur in miRNA sequences and to perform *in silico* predictions on the effects of the presence of these SNPs on target binding and primary miRNA formation.

2.3 Results

SNPs in miRNA primary sequences

We performed a scan in 1064 miRNA primary sequences and found 452 miRNA sequences with SNPs, 263 produced by canonical pathway and the remaining 189 by non-canonical pathway (Fig 1a). Some miRNAs harbored more than one SNP per sequence. Nucleotide changes were comprised of 42% transversions and 58% transitions (Fig 1b).

(a)



(b)

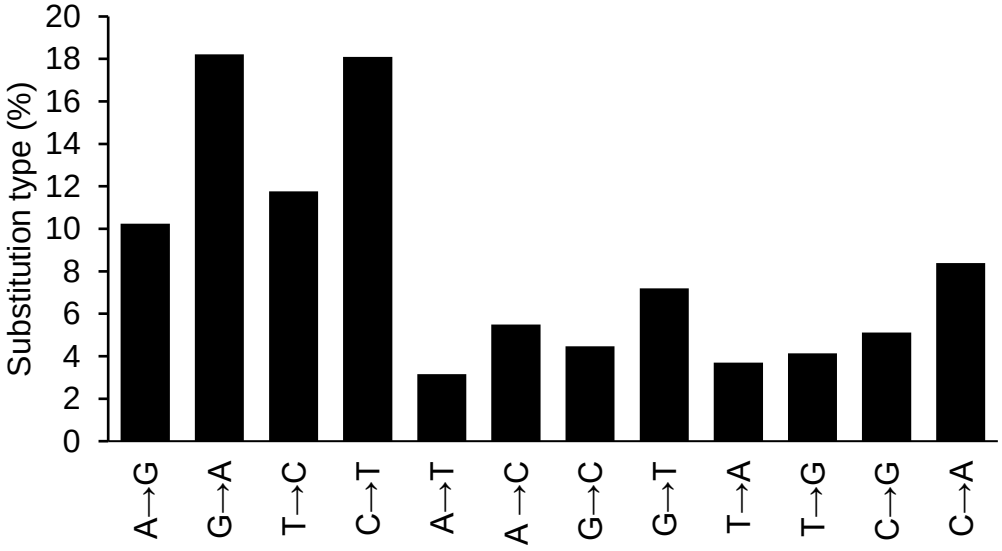
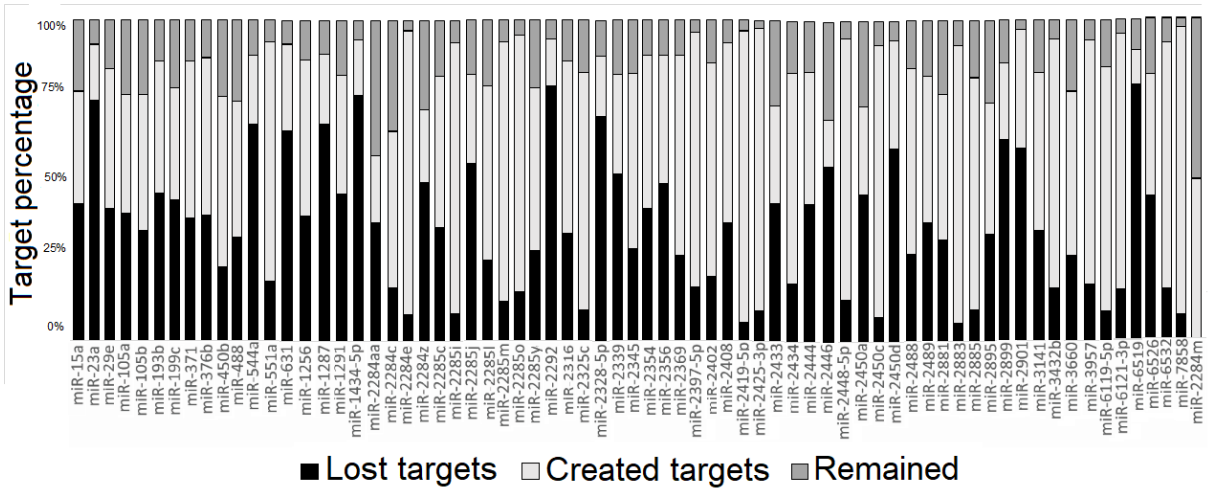


Fig 1 – Frequency distributions of SNPs per miRNA primary sequences (a) and nucleotide substitution type (b) in primary sequences. Arrows indicate the direction of change, based on the reference genome.

Changes in target prediction

Out of 193 miRNA mature sequences with SNPs, 70 had alterations in target prediction. Another 21 mature miRNA had SNPs in their seed region, however, no changes were observed in target prediction. Out of all predicted targets for these 70 miRNAs, 48.2% were created and 33.8% were lost (Fig 2a and b) as a consequence of the presence of a SNP. Only 18% of the targets were unaltered by the presence of SNP within the seed region of the miRNAs (Fig 2b). No significant difference in target alteration was observed between transitions and transversions (1171 vs 1347 average target alteration, respectively).

(a)



(b)

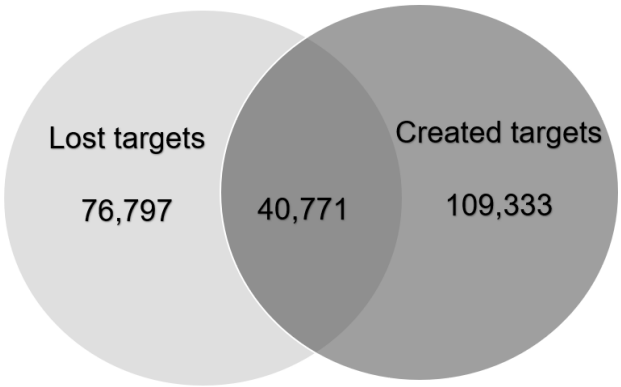
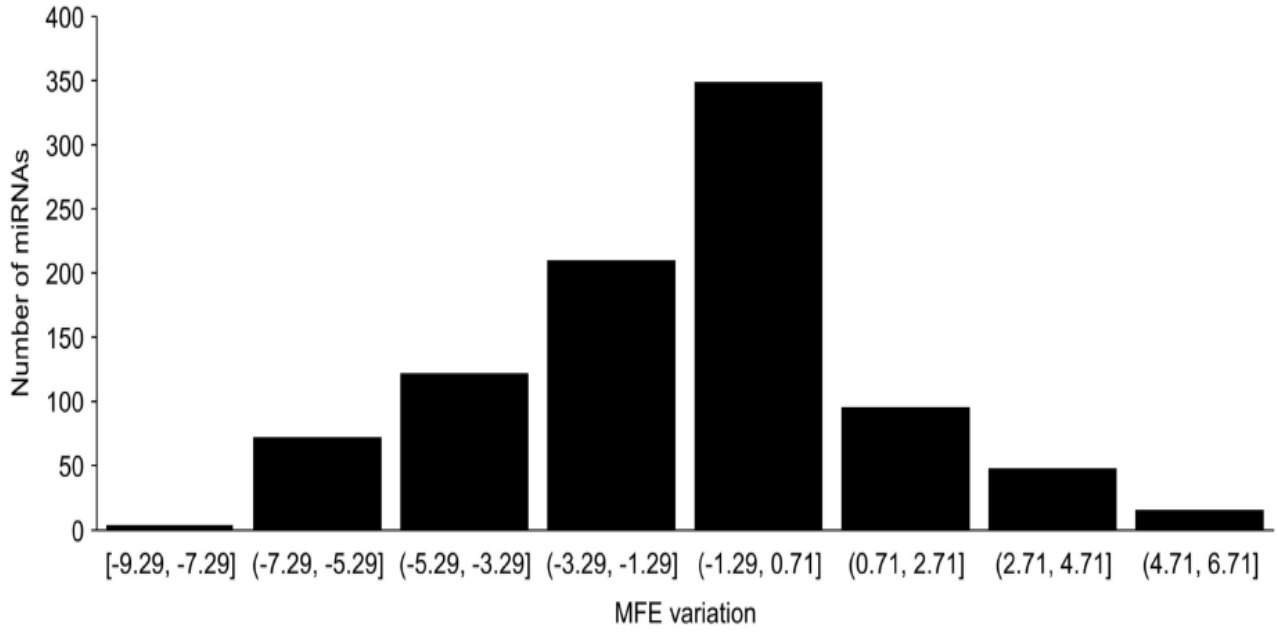


Fig 2 - Predicted target created and lost targets (a) of each miRNA mature sequence. Venn diagram containing all lost and created targets (b).

Effects of SNPs on miRNA hairpin structure

In order to investigate the influence of a SNP in the hairpin structure of the miRNAs, we calculated the energy change (ΔΔG) between reference and altered sequences (Fig 3A) using the RNAfold Program [11]. We observed that the presence of SNPs can alter hairpin conformation due to loss or acquisition of base pairing, consequently altering free energy (Fig 3B, 3C). In our study we found 58% transitions and 42% transversions, and statistical analysis indicated that there was a significant difference regarding the MFE variation between the two types (-0.83 vs -1.61, transitions and transversions, respectively; p-value<2.348e-05).

(a)



(b)

(c)

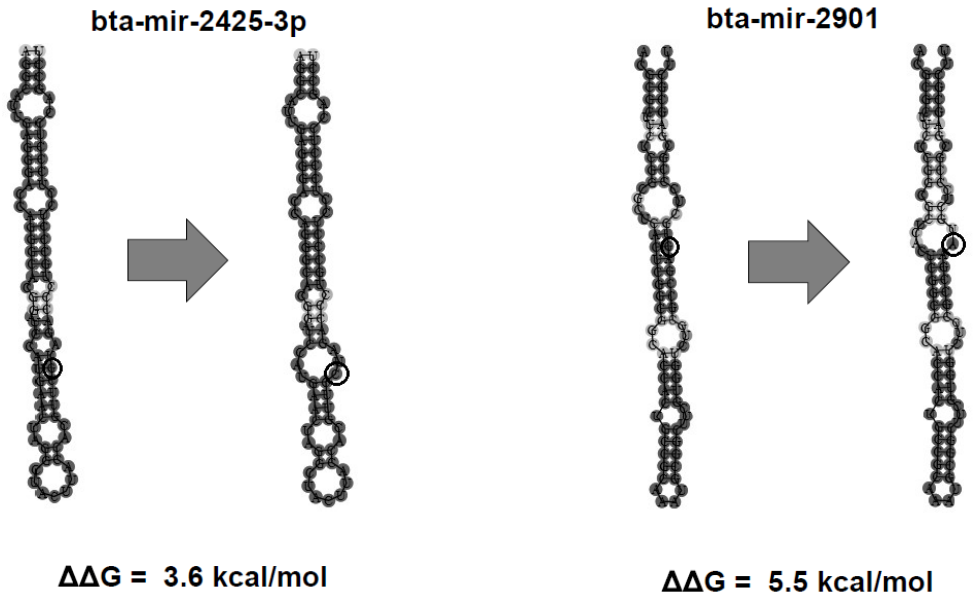


Fig 3 – Frequency distribution of energy variation between reference and altered miRNA sequences (a). Illustration of bta-mir-2425-3p with a G to C transversion (b) and bta-mir-2901 with a G to A transition (c) and the alterations in conformation.

Presence of miRNAs with SNPs in QTL regions

Focusing on mature miRNAs with altered target prediction, we crossed their positions with that of known QTL regions (Table 1). We found that 45 traits had one or more SNP in miRNA sequence within their locations.

Table 1 - Mature miRNA with SNPs affecting target prediction in QTL regions

Trait categories	Number of miRNAs
Reproduction traits	
Calving ease	11
Calving ease (maternal)	5
Calf size	4
Birth index	3
Stillbirth	2
Scrotal circumference	2
Udder swelling score	2
Calving to conception interval	1
Age at puberty	1
Calving index	1
Length of productive life	1
Milk production traits	
Milk caproic acid content	11
Milk palmitoleic acid content	6
Tridecylic acid content	6
Milk butyric acid content	4
Margaric acid content	3
Milk decenoic acid content	3
Milk myristoleic acid content	3
Milk capric acid content	3
305-day milk yield	3
Milk oleic acid content	2
Milk alpha-lactalbumin percentage	2
Milk lauroleic acid content	2
Milk yield	2
Milk beta-casein percentage	1
Milk conjugated linoleic acid content	1
Monounsaturated fatty acid content	1
Trans-15-C18:1 fatty acid content	1
Cis-10 Heptadecenoic acid content	1
Conformational traits	
Body weight (weaning)	9
Body weight (yearling)	2
Residual feed intake	2
Somatic cell score	1
Average daily gain	1
Body weight (birth)	1
Conformation score	1
Non-return rate (EBV)	1
Meat production traits	
Shear force	7
Intramuscular fat	6

Subcutaneous fat	1
Muscle iron content	1
Environmental adaptation traits	
Cold tolerance	4
Heat tolerance	1
Disease tolerance traits	
Tick resistance	2
Bovine tuberculosis susceptibility	1

PhyloP Conservation scores for SNPs within primary sequences

To investigate the conservation level of each altered base occurring in miRNA primary sequences, we obtain the PhyloP score based on the alignment of 100 vertebrates [12]. Positive, neutral and negative values are employed to indicate slow, neutral and fast-evolving variations, respectively. We aligned and obtained the phyloP score of 214 sequences of primary miRNAs (Fig 4). All miRNAs containing a SNP in their seed region, had a positive phyloP score, regardless of target alteration.

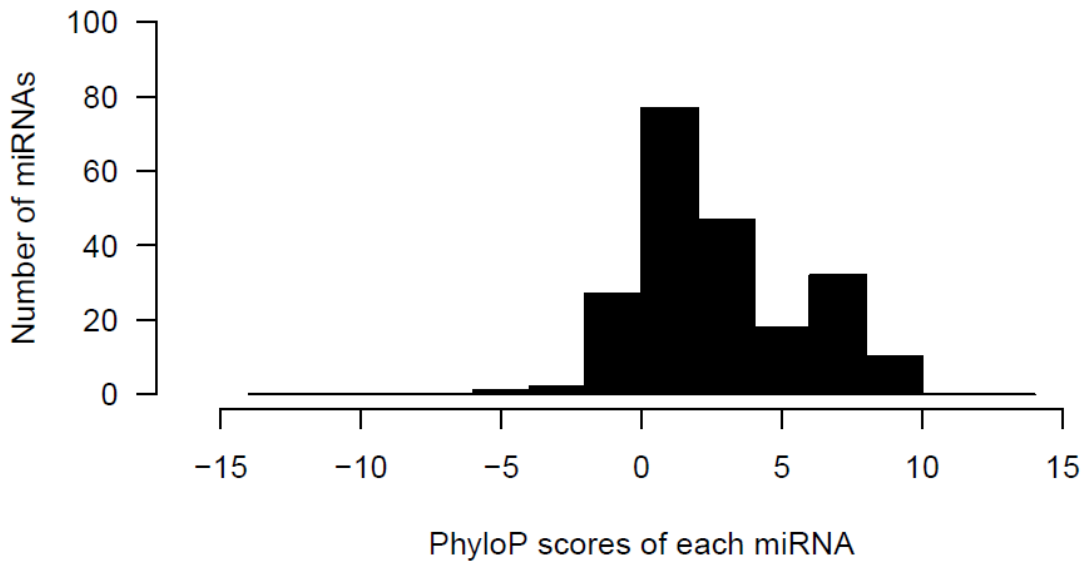


Fig 4 – Frequency distribuition of PhyloP scores for each SNP position in miRNA sequences.

2.4 Discussion

We were able to contrast the largest SNP database available in the bovine species to date, with all available miRNA sequences. The publicly available SNP databases, Bovine Genome Database [13] and SNPchiMp v.3 [14][15], that gathers data from SNPs used on several chips produced by Illumina and Affymetrix, have a combined total of 2 million SNP annotations. The dataset used in this study, Run 5 of the 1000BGP, has more than 56 million SNP annotations.

Guo *et al*. have recently demonstrated that transversions have more pronounced regulatory effects than transitions [16], though their effects on non-coding sequences is yet to be elucidated. Likewise, we found a significant increase in MFE variation in transversions when compared to transitions, suggesting different effects of transitions and transversions on non-coding RNAs.

In humans, a single miRNA can regulate hundreds of thousands of genes, and it is estimated that they regulate 90% of all genes in the genome [17]. The impact of SNPs in miRNA sequences or in a miRNA target sites is related to loss or gain of target genes in humans [18] and plants [19]. In our study, SNPs within seed sequences greatly affected target prediction, with less than 20% of targets remaining unaltered between reference and altered alleles. [20]. Interestingly, 21 miRNA sequences with SNPs in their seed region that did not alter their binding to targets, and in all these sequences the SNP was located at the last nucleotide of the seed, suggesting that SNPs in this position do seem to affect target prediction. Further, the influence of SNPs on the divergence of traits between cattle breeds, or within breed, are

well reported in several studies, as revised by Seidel [21] . In fact, a GWAS analysis detected the presence of SNPs in a QTL region linked to calving ease, without the presence of protein-coding genes at this site, but rather of a miRNA [22]. In our study, we have identified 140 miRNAs with SNPs in previously identified QTL regions. Although these loci are large, spanning several Kbs within chromosomes, 129 of these miRNAs associated with QTLs presented altered target prediction as a consequence of a SNP within their seed region, which could translate directly in significant protein differences in individuals harboring different bases.

Presence of SNPs in primary miRNAs sequences may also alter the production of mature miRNAs. This is due to the alterations in recognition by their cleavage proteins, which identify the conformation of the secondary structure of the molecule and not simply the nucleotide sequence [23][24]. Association between the presence of SNPs in the pre-miRNA flanking region and cancer has been reported in humans [25]. Alteration in MFE in miRNA primary sequences, caused by SNPs, are reported and validated in carp [26]. Values ranging from 2.1 to 7.1 were reported as sufficient to alter the production of mature miRNAs [27]. In our study 48.63% of miRNAs primary sequences with SNPs showed ΔΔG greater than these values. MFE alteration between wild-type and SNP-type pre-miRNA-2467 has already been reported in a study with five bovine breeds (Angus, Jersey, Holstein-Friesian, Limousin, Hereford), where a loss of stability resulted from the formation of a new loop in the hairpin structure following the loss of base pairing [28]

Evolutionary conservation of miRNAs, as well the conservation of their target genes, have already been evidenced in previous studies [29][30]. In this study we decided not to limit ourselves to determining the conservation of miRNA sequences based on homologous occurrence in other vertebrates, but rather to infer conservation relevance by retrieving phyloP scores. PhyloP scores with positive values represent conservation, while scores with negative values represent rapid evolution. Our analysis showed that ~86% of the loci of 214 aligned miRNA sequences where SNPs occurred had scores greater than 0, suggesting a conserved biological function through the evolutionary process, where natural selection favored individuals who maintained this conserved position, increasing their frequency in populations. Therefore, a change in these conserved positions could indicate unfavorable phenotypes, influenced by altered maturation of primary miRNAs or misregulated post-transcriptional repression promoted by mature miRNAs.

2.5 Conclusions

The present study is the first to perform a global analysis of SNPs that occur in miRNAs in cattle. Through the use of computational tools, we have created a database that can aid in the investigation of epigenetic regulation of phenotypes in cattle. We demonstrated that SNPs can alter target prediction for miRNAs, as well as pre-miRNA processing, increasing or decreasing mature miRNA production, which in turn could affect production traits in cattle.

2.6 Methods

Bovine Assembly and Database

In this study we used the *Bos taurus* UMD 3.1.1 [31] as the reference genome. SNP positioning of approximately ~56 million SNPs were acquired from the Run5 of the 1000 Bulls Genome Project (BGP; http://www.1000bullgenomes.com/). Primary and mature miRNA location, positioning, and sequences were obtained from miRbase (www.mirbase.org), release 21 [32][33][34][35][36][37][38]. Positions of QTLs were extracted from the Animal QTLdb [39].

Mapping of SNPs in miRNA sequences

An in-house Python (2.7) script was used to identify all SNPs (Run5 of BGP) that are located within all bovine miRNA sequences available on miRBase (release 21).

Prediction of miRNA targets

The miRmap [40] package was adapted for target prediction through Python (3.6.4) scripts, using 13,345 3`UTR bovine sequences available within the same package. Target prediction was performed for all miRNAs containing SNPs. Reference sequences (similar to UMD 3.1.1) and alternate sequences (alternate allele from Run5) were then classified in regards to target gene prediction as: 1) targets were created as a result of SNP; 2) targets were lost; and 3) targets were unaltered. Welch Two Sample t-test

was employed to compare transitions and transversions in regards to target alteration.

Prediction of energy change in primary miRNA sequences

To calculate variations in Minimum Free Energy (MFE) between reference and altered sequences of primary miRNAs, the program RNAFold - ViennaRNA Package 2.0 [11] was utilized. This tool can calculate changes in MFE, which can affect the formation of mature miRNAs, based on alterations to the hairpin structure of primary miRNAs caused by unpairing of bases within the double strand section of the hairpin. Welch Two Sample t-test was used to compare transitions and transversions in regards to the MFE change induced.

PhyloP Scores

PhyloP scores are used to measure nonneutral nucleotide substitution rates, indicating conservation and acceleration independently at each nucleotide [41]. To calculate PhyloP scores [12], we performed a BLAST [42] search (e-value $\leq 10^{-3}$) [43] of all primary miRNA sequences against the human genome, then the identified human sequences were located on the UCSC Genome Browser [44][45]. Once sequences were located, an alignment was performed using MUSCLE [46], followed by manual search of the corresponding human base pair to the original bovine SNP position. Lastly, the conservation track Cons 100 Verts (phyloP100way), available on the UCSC genome browser, was used to obtain the PhyloP value to each position of interest.

2.7 Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

The datasets generated and/or analysed during the current study are not publicly available due were provide by the internacional private consortium 1k Bulls Genomes Project but are available from the corresponding author upon reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

Financial support: Coordination for the Improvement of Higher Level Education (CAPES); studentship #2015/20557-5 and #2016/07584-6, São Paulo Research Foundation (FAPESP). National Research Initiative grants number 2010-65205-20414 and 2017-67015-26760 from the USDA National Institute of Food and Agriculture.

Authors´ contributions

All authors read and approved the final manuscript.

Acknowledgements

Not applicable.

2.8 References

1. Ajmone-Marsan P, Lenstra JA, Fernando Garcia J, The Globaldiv Consortium. On the origin of cattle: how aurochs became domestic and colonized the world Attenuation of the inflammatory phenomena in the transition period of dairy cows View project Climate Genomics for Farm Animal Adaptation View project. Evol Anthropol. 2010;19:148–57.

2. Orita M, Iwahana H, Kanazawat H, Hayashi K, Sekiya T. Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphisms (mobility shift of separated strands/point mutation/riction franment length polymorphism). Genetics. 1989;86(April):2766–70.

3. Caetano AR. Marcadores SNP: Conceitos básicos, aplicações no manejo e no melhoramento animal e perspectivas para o futuro. Rev Bras Zootec. 2009;38 (SUPPL.) 1:64–71.

4. Bartel DP, Lee R, Feinbaum R. MicroRNAs: Genomics, Biogenesis, Mechanism, and Function. Cell. 2004;116:281–97.

5. Wienholds E, Plasterk RHA. MicroRNA function in animal development. FEBS Lett. 2005;579(26):5911–22.

6. Moszyńska A, Gebert M, Collawn JF, Bartoszewski R. SNPs in microRNA target sites and their potential role in human disease. Open Biol. 2017;7(170019)1-13.

7. Bradshaw G, Haupt LM, Aquino EM, Lea RA, Sutherland HG, Griffiths LR. Single Nucleotide Polymorphisms in MIR143 Contribute to Protection Against Non-Hodgkin Lymphoma (NHL) in Caucasian Populations. Genes (Basel). 2019;10:(185)1-20.

8. Wang S, Zhu H, Ding B, Feng X, Zhao W, Cui M, et al. Genetic variants in microRNAs are associated with cervical cancer risk. Mutagenesis. 2019;XX:1–7.

9. Ju Z, Wang C, Wang X, Yang C, Zhang Y, Sun Y, et al. The effect of the SNP g.18475 A>G in the 3′UTR of NCF4 on mastitis susceptibility in dairy cattle. Cell Stress Chaperones. 2018;23:385–391.

10. Gao Q, Ju Z, Zhang Y, Huang J, Zhang X, Qi C, et al. Association of TNP2 gene polymorphisms of the bta-miR-154 target site with the semen

quality traits of Chinese Holstein bulls. PLoS One. 2014;9(1):1–9.

11. Lorenz R, Bernhart SH, Höner zu Siederdissen C, Tafer H, Flamm C, Stadler PF, et al. ViennaRNA Package 2.0. Algorithms Mol Biol. 2011;6(26):1-14.

12. Hubisz MJ, Pollard KS, Siepel A. Phastand Rphast: Phylogenetic analysis with space/time models. Brief Bioinform. 2011;12(1):41–51.

13. Elsik CG, Unni DR, Diesh CM, Tayal A, Emery ML, Nguyen HN, et al. Bovine genome database: New tools for gleaning function from the Bos taurusgenome. Nucleic Acids Res. 2016;44:(D1)834–839.

14. Nicolazzi EL, Picciolini M, Strozzi F, Schnabel RD, Lawley C, Pirani A, et al. SNPchiMp: A database to disentangle the SNPchip jungle in bovine livestock. BMC Genomics. 2014;15(123):1–6.

15. Nicolazzi EL, Caprera A, Nazzicari N, Cozzi P, Strozzi F, Lawley C, et al. SNPchiMp v.3: Integrating and standardizing single nucleotide polymorphism data for livestock species. BMC Genomics. 2015;16(283):1–6.

16. Guo C, McDowell IC, Nodzenski M, Scholtens DM, Allen AS, Lowe WL, et al. Transversions have larger regulatory effects than transitions. BMC Genomics. 2017;18(394):1-9.

17. Latini A, Ciccacci C, Novelli G, Borgiani P. Polymorphisms in miRNA genes and their involvement in autoimmune diseases susceptibility. Immunol Res. 2017;65(4):811–827.

18. Gong J, Tong Y, Zhang H-M, Guo A-Y. miRNASNP: a database of miRNA related SNPs and their effects on miRNA function. BMC Bioinformatics. 2012;13 (Suppl 18):A2.

19. Ling J, Luo Z, Liu F, Mao Z, Yang Y, Xie B. Genome-wide analysis of microRNA targeting impacted by SNPs in cucumber genome. BMC Genomics. 2017;18(1):1–12.

20. Duan R, Pak CH, Jin P. Single nucleotide polymorphism associated with mature miR-125a alters the processing of pri-miRNA. Hum Mol Genet. 2007;16(9):1124–1131.

21. Seidel, Jr. GE. Brief introduction to whole-genome selection in cattle using single nucleotide polymorphisms. Reprod Fertil Dev. 2010;22:138–44.

22. Purfield DC, Bradley DG, Kearney JF, Berry DP. Genome-wide association study for calving traits in Holstein-Friesian dairy cattle. Animal.

2014;8(2):224–35.
23. Schanen BC, Li X. Transcriptional regulation of mammalian miRNA genes. Genomics. 2011;97:1–6.
24. Sommer SS, Li H, Sarkis DA, Feng J, Noltner K, Rossi JJ, et al. SNPs in human miRNA genes affect biogenesis and function. Rna. 2009;15:1640–1651.
25. Bensen JT, Tse CK, Nyante SJ, Barnholtz-Sloan JS, Cole SR, Millikan RC. Association of germline microRNA SNPs in pre-miRNA flanking region and breast cancer risk and survival: The Carolina Breast Cancer Study. Cancer Causes Control. 2013;24(6):1099–1109.
26. Zhu YP, Xue W, Wang JT, Wan YM, Wang SL, Xu P, et al. Identification of common carp (Cyprinus carpio) microRNAs and microRNA-related SNPs. BMC Genomics. 2012;13(413)1-12.
27. Gong J, Tong Y, Zhang HM, Wang K, Hu T, Shan G, et al. Genome-wide identification of SNPs in MicroRNA genes and the SNP effects on MicroRNA target binding and biogenesis. Hum Mutat. 2011;33(1):254–263.
28. Łukaszewicz A, Basiak S, Proskura WS, Dybus A. Nucleotide Substitution in 3’ Arm of Bovine MIR-2467 in Five Cattle Breeds. Anim Biotechnol. 2015;26(4):276–278.
29. Fehlmann T, Laufer T, Backes C, Kahramann M, Alles J, Fischer U, et al. Large-scale validation of miRNAs by disease association, evolutionary conservation and pathway activity. RNA Biol. 2019;16:93–103.
30. Friedman RC, Farh KKH, Burge CB, Bartel DP. Most mammalian mRNAs are conserved targets of microRNAs. Genome Res. 2009;19:92–105.
31. Zimin A V., Delcher AL, Florea L, Kelley DR, Schatz MC, Puiu D, et al. A whole-genome assembly of the domestic cow, Bos taurus. Genome Biol. 2009;10(4)1-10.
32. Griffiths-Jones S. The microRNA Registry. Nucleic Acids Res. 2003;32:110-111.
33. Ambros V, Bartel B, Bartel DP, Burge CB, Carrington JC, Chen X, et al. A uniform system for microRNA annotation. RNA. 2003;9(3)277-279.
34. Kozomara A, Birgaoanu M, Griffiths-Jones S. miRBase: from microRNA sequences to function. Nucleic Acids Res. 2019;47(D1):155-162.
35. Kozomara A, Griffiths-Jones S. MiRBase: Integrating microRNA

annotation and deep-sequencing data. Nucleic Acids Res. 2011;39:152-157.
36. Kozomara A, Griffiths-Jones S. MiRBase: Annotating high confidence microRNAs using deep sequencing data. Nucleic Acids Res. 2014;42:68-73.
37. Griffiths-Jones S, Saini HK, Van Dongen S, Enright AJ. miRBase: Tools for microRNA genomics. Nucleic Acids Res. 2008;36:154-158.
38. Griffiths-Jones S. miRBase: microRNA sequences, targets and gene nomenclature. Nucleic Acids Res. 2005;1(34):140-144.
39. Hu ZL, Park CA, Reecy JM. Building a livestock genetic and genomic information knowledgebase through integrative developments of Animal QTLdb and CorrDB. Nucleic Acids Res. 2019;47(D1):701-710.
40. Vejnar CE, Zdobnov EM. MiRmap: Comprehensive prediction of microRNA target repression strength. Nucleic Acids Res. 2012;40(22):11673-11683.
41. Pollard KS, Hubisz MJ, Rosenbloom KR, Siepel A. Detection of nonneutral substitution rates on mammalian phylogenies. Genome Res. 2010;20:110–21.
42. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990;215:403–10.
43. Pinhal D, Bovolenta LA, Moxon S, Oliveira AC, Nachtigall PG, Acencio ML, et al. Genome-wide microRNA screening in Nile tilapia reveals pervasive isomiRs' transcription, sex-biased arm switching and increasing complexity of expression throughout development. Sci Rep. 2018;8(1):1-18.
44. James Kent W, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, et al. The human genome browser at UCSC. Genome Res. 2002;12(6):996-1006.
45. Karolchik D, Hinrichs AS, Kent WJ. The UCSC genome browser. Curr Protoc Hum Genet. 2011;1.4:1.4.1-1.4.33.
46. Edgar RC. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 2004;32:1792–1797.

APÊNDICES

REFERÊNCIAS DA INTRODUÇÃO GERAL

AJMONE-MARSAN, P. et al. **On the origin of cattle: how aurochs became domestic and colonized the world Attenuation of the inflammatory phenomena in the transition period of dairy cows View project Climate Genomics for Farm Animal Adaptation View project.** Evolutionary Anthropology, v. 19, p. 148–157, 2010.

ANDERS H. LUND; MAARTEN VAN LOHUIZEN. **Epigenetics and cancer.** GENES & DEVELOPMENT, v. 18, n. 3, p. 2315–2335, 2004.

BARTEL, D. P. **MicroRNAs: Target Recognition and Regulatory Functions.** Cell, v. 136, n. 2, p. 215–233, 2009.

BENDER, J. **Dna Methylation and Epigenetics.** Annual Review of Plant Biology, v. 55, n. 1, p. 41–68, 2004.

BEUZEN, N. D.; STEAR, M. J.; CHANG, K. C. **Molecular markers and their use in animal breeding.** Veterinary Journal, v. 160, n. 1, p. 42–52, 2000.

BRADSHAW, G. et al. **Single Nucleotide Polymorphisms in MIR143 Contribute to Protection Against Non-Hodgkin Lymphoma (NHL) in Caucasian Populations.** Genes, v. 10, n. 3, p. 185, 2019.

BUENO, M. J.; DE CASTRO, I. P.; MALUMBRES, M. **Control of cell proliferation pathways by microRNAs.** Cell Cycle, v. 7, n. 20, p. 3143–3148, 2008.

CAETANO, A. R. **Marcadores SNP: Conceitos básicos, aplicações no manejo e no melhoramento animal e perspectivas para o futuro.** Revista Brasileira de Zootecnia, v. 38, n. SUPPL. 1, p. 64–71, 2009.

DU, F. X.; CLUTTER, A. C.; LOHUIS, M. M. **Characterizing linkage disequilibrium in pig populations.** International Journal of Biological Sciences, v. 3, n. 3, p. 166–178, 2007.

EULALIO, A.; SCHULTE, L. N.; VOGE, J. **The mammalian microRNA response to bacterial infections.** RNA Biology, v. 9, n. 6, p. 742–750, jun.2012.

FILIPOWICZ, W.; BHATTACHARYYA, S. N.; SONENBERG, N. **Mechanisms of post-transcriptional regulation by microRNAs: Are the**

answers in sight? Nature Reviews Genetics, v. 9, n. 2, p. 102–114, fev.2008.

HUNTZINGER, E.; IZAURRALDE, E. **Gene silencing by microRNAs: Contributions of translational repression and mRNA decay.** Nature Reviews Genetics, v. 12, n. 2, p. 99–110, fev.2011.

JIANG, Z. et al. **A novel type of sequence variation: Multiple-nucleotide length polymorphisms discovered in the bovine genome.** Genetics, v. 176, n. 1, p. 403–407, fev.2007.

JIANG, Z. et al. **The complementary neighborhood patterns and methylation-to-mutation likelihood structures of 15,110 single-nucleotide polymorphisms in the bovine genome.** Genetics, v. 180, n. 1, p. 639–647, jun.2008.

JONATHAN M. LEVENSON; J.DAVID SWEATT. **Epigenetic mechanisms in memory formation.** NATURE REVIEWS | NEUROSCIENCE, v. 6, February, p. 108–118, fev.2005.

KEDDE, M.; AGAMI, R. **Interplay between microRNAs and RNA-binding proteins determines developmental processes.** Cell Cycle, v. 7, n. 7, p. 899–903, apr.2008.

KIM, S.; MISRA, A. **SNP genotyping: technologies and biomedical applications.** Annual review of biomedical engineering, v. 9, p. 289–320, mar.2007.

MAUDET, C. et al. **Functional high-throughput screening identifies the miR-15 microRNA family as cellular restriction factors for Salmonella infection.** Nature Communications, v. 5, n. 5718, p. 1-13, aug.2014.

MEUWISSEN, T. H. E.; HAYES, B. J.; GODDARD, M. E. **Prediction of Total Genetic Value Using Genome-Wide Dense Marker Maps.** Genetics, v. 157, p. 1819–1829, jan.2001.

O'CONNELL, R. M.; RAO, D. S.; BALTIMORE, D. **microRNA Regulation of Inflammatory Responses.** Annual Review of Immunology, v. 30, n. 1, p. 295–312, jan.2012.

RAFALSKI, A. **Applications of single nucleotide polymorphisms in crop genetics**. Current Opinion in Plant Biology, v. 5, n. 2, p. 94–100, apr.2002.

RILEY, J. H. et al. **The use of single nucleotide polymorphisms in the isolation of common disease genes.** Pharmacogenomics, v. 1, n. 1, p. 39–

47, 2000.

SALIMINEJAD, K. et al. **An overview of microRNAs: Biology, functions, therapeutics, and analysis methods.** Journal of Cellular Physiology, v. 234, n. 5, p. 5451–5465, jun.2018.

SEIDEL, JR., G. E. B**rief introduction to whole-genome selection in cattle using single nucleotide polymorphisms.** Reproduction Fertility and Development, v. 22, n. 1, p. 138–144, 2010.

STEFANI, G.; SLACK, F. J. **Small non-coding RNAs in animal development.** Nature Reviews Molecular Cell Biology, v. 9, n. 3, p. 219–230, mar.2008.

TANG, W. Y.; HO, S. M. **Epigenetic reprogramming and imprinting in origins of disease.** Reviews in Endocrine and Metabolic Disorders, v. 8, n. 2, p. 173–182, jul.2007.

ZHAO, C. et al. **MiRNA-dysregulation associated with tenderness variation induced by acute stress in Angus cattle.** Journal of Animal Science and Biotechnology, v. 3, n. 1, p. 1–8, 2012.

ANEXO

NORMAS DA REVISTA

BMC Genetics Submission Guidelines

Preparing your manuscript

The information below details the section headings that you should include in your manuscript and what information should be within each section.

Please note that your manuscript must include a 'Declarations' section including all of the subheadings (please see below for more information).

Title page

The title page should:

- present a title that includes, if appropriate, the study design
- list the full names and institutional addresses for all authors
    - if a collaboration group should be listed as an author, please list the Group name as an author. If you would like the names of the individual members of the Group to be searchable through their individual PubMed records, please include this information in the “Acknowledgements” section in accordance with the instructions below
- indicate the corresponding author

Abstract

The Abstract should not exceed 350 words. Please minimize the use of abbreviations and do not cite references in the abstract. The abstract must include the following separate sections:

- **Background:** the context and purpose of the study
- **Results:** the main findings
- **Conclusions:** a brief summary and potential implications

Keywords

Three to ten keywords representing the main content of the article.

Background

The Background section should explain the background to the study, its aims, a summary of the existing literature and why this study was necessary.

Results

This should include the findings of the study including, if appropriate, results of statistical analysis which must be included either in the text or as tables and figures.

Discussion

For research articles this section should discuss the implications of the findings in context of existing research and highlight limitations of the study. For study protocols and methodology manuscripts this section should include a discussion of any practical or operational issues involved in performing the study and any issues not covered in other sections.

Conclusions

This should state clearly the main conclusions and provide an explanation of the importance and relevance of the study to the field.

Methods

The methods section should include:

- the aim, design and setting of the study
- the characteristics of participants or description of materials
- a clear description of all processes, interventions and comparisons. Generic names should generally be used. When proprietary brands are used in research, include the brand names in parentheses
- the type of statistical analysis used, including a power calculation if appropriate

List of abbreviations

If abbreviations are used in the text they should be defined in the text at first use, and a list of abbreviations can be provided.

Declarations

All manuscripts must contain the following sections under the heading 'Declarations':

- Ethics approval and consent to participate
- Consent for publication
- Availability of data and material
- Competing interests
- Funding
- Authors' contributions
- Acknowledgements

- Authors' information (optional)

Please see below for details on the information to be included in these sections.

If any of the sections are not relevant to your manuscript, please include the heading and write 'Not applicable' for that section.

Ethics approval and consent to participate

Manuscripts reporting studies involving human participants, human data or human tissue must:

- include a statement on ethics approval and consent (even where the need for approval was waived)
- include the name of the ethics committee that approved the study and the committee's reference number if appropriate

Studies involving animals must include a statement on ethics approval.

See our editorial policies for more information.

If your manuscript does not report on or involve the use of any animal or human data or tissue, please state "Not applicable" in this section.

Consent for publication

If your manuscript contains any individual person's data in any form (including any individual details, images or videos), consent for publication must be obtained from that person, or in the case of children, their parent or legal guardian. All presentations of case reports must have consent for publication.

You can use your institutional consent form or our consent form if you prefer. You should not send the form to us on submission, but we may request to see a copy at any stage (including after publication).

See our editorial policies for more information on consent for publication.

If your manuscript does not contain data from any individual person, please state "Not applicable" in this section.

Availability of data and materials

All manuscripts must include an 'Availability of data and materials' statement. Data availability statements should include information on where data supporting the results reported in the article can be found including, where applicable, hyperlinks to publicly archived datasets analysed or generated during the study. By data we mean the minimal dataset that would be

necessary to interpret, replicate and build upon the findings reported in the article. We recognise it is not always possible to share research data publicly, for instance when individual privacy could be compromised, and in such instances data availability should still be stated in the manuscript along with any conditions for access.

Data availability statements can take one of the following forms (or a combination of more than one if required for multiple datasets):

- The datasets generated and/or analysed during the current study are available in the [NAME] repository, [PERSISTENT WEB LINK TO DATASETS]
- The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.
- All data generated or analysed during this study are included in this published article [and its supplementary information files].
- The datasets generated and/or analysed during the current study are not publicly available due [REASON WHY DATA ARE NOT PUBLIC] but are available from the corresponding author on reasonable request.
- Data sharing is not applicable to this article as no datasets were generated or analysed during the current study.
- The data that support the findings of this study are available from [third party name] but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of [third party name].
- Not applicable. If your manuscript does not contain any data, please state 'Not applicable' in this section.

Competing interests

All financial and non-financial competing interests must be declared in this section.

See our editorial policies for a full explanation of competing interests. If you are unsure whether you or any of your co-authors have a competing interest please contact the editorial office.

Please use the authors initials to refer to each authors' competing interests in this section.

If you do not have any competing interests, please state "The authors declare that they have no competing interests" in this section.

Funding

All sources of funding for the research reported should be declared. The role of the funding body in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript should be declared.

Authors' contributions

The individual contributions of authors to the manuscript should be specified in this section. Guidance and criteria for authorship can be found in our editorial policies.

Please use initials to refer to each author's contribution in this section, for example: "FC analyzed and interpreted the patient data regarding the hematological disease and the transplant. RH performed the histological examination of the kidney, and was a major contributor in writing the manuscript. All authors read and approved the final manuscript."

Acknowledgements

Please acknowledge anyone who contributed towards the article who does not meet the criteria for authorship including anyone who provided professional writing services or materials.

Authors should obtain permission to acknowledge from all those mentioned in the Acknowledgements section.

See our editorial policies for a full explanation of acknowledgements and authorship criteria.

If you do not have anyone to acknowledge, please write "Not applicable" in this section.

Authors' information

This section is optional.

You may choose to use this section to include any relevant information about the author(s) that may aid the reader's interpretation of the article, and understand the standpoint of the author(s). This may include details about the authors' qualifications, current positions they hold at institutions or societies, or any other relevant background information. Please refer to authors using their initials. Note this section should not be used to describe any competing interests.

Endnotes

Endnotes should be designated within the text using a superscript lowercase letter and all notes (along with their corresponding letter) should be included

in the Endnotes section. Please format this section in a paragraph rather than a list.

References

Examples of the Vancouver reference style are shown below.

See our editorial policies for author guidance on good citation practice.