

Citation: Sousa MAPd, de Athayde FRF, Maldonado MBC, Lima AOd, Fortes MRS, Lopes FL (2021) Single nucleotide polymorphisms affect miRNA target prediction in bovine. PLoS ONE 16(4): e0249406. https://doi.org/10.1371/journal. pone.0249406

Editor: Junning Yue, University of Tennessee Health Science Center, UNITED STATES

Received: September 14, 2020

Accepted: March 17, 2021

Published: April 21, 2021

Copyright: © 2021 Sousa et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting information files.

Funding: FLL is a PQ-2 Scholar of the National Council for Scientific and Technological Development (CNPq). Studentship to FRFA was financed by the Coordination for the Improvement of Higher Level Education (CAPES). Funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. **RESEARCH ARTICLE**

Single nucleotide polymorphisms affect miRNA target prediction in bovine

Marco Antônio Perpétuo de Sousa^{1©‡}, Flavia Regina Florêncio de Athayde^{1‡}, Mariângela Bueno Cordeiro Maldonado^{1‡}, Andressa Oliveira de Lima^{1©}, Marina Rufino S. Fortes^{2©}, Flavia Lombardi Lopes^{1©}*

1 Department of Production and Animal Health, São Paulo State University (Unesp), School of Veterinary Medicine, Araçatuba, São Paulo, Brazil, 2 School of Chemistry and Molecular Bioscience, The University of Queensland, Brisbane, Queensland, Australia

So These authors contributed equally to this work.

‡ These authors also contributed equally to this work.

* flavia.lopes@unesp.br

Abstract

Single nucleotide polymorphisms (SNPs) can have significant effects on phenotypic characteristics in cattle. MicroRNAs (miRNAs) are small, non-coding RNAs that act as post-transcriptional regulators by binding them to target mRNAs. In the present study, we scanned ~56 million SNPs against 1,064 bovine miRNA sequences and analyzed, in silico, their possible effects on target binding prediction, primary miRNA formation, association with QTL regions and the evolutionary conservation for each SNP locus. Following target prediction, we show that 71.6% of miRNA predicted targets were altered as a consequence of SNPs located within the seed region of the mature miRNAs. Next, we identified variations in the Minimum Free Energy (MFE), which represents the capacity to alter molecule stability and, consequently, miRNA maturation. A total of 48.6% of the sequences analyzed showed values within those previously reported as sufficient to alter miRNA maturation. We have also found 131 SNPs in 46 miRNAs, with altered target prediction, occurring in QTL regions. 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 affect bovine phenotypes and could be of great value for genetic improvement studies, as well as production.

Introduction

Single nucleotide polymorphisms (SNPs), the most common mutation in mammalian genomes, are single-base exchanges that naturally occur in the genome and have long been studied as valuable genetic markers, as they can affect and/or aid the selection of phenotypic traits [1,2]. Current genomic selection of production animals heavily relies on identification and genotyping of molecular markers such as SNPs [3]. Mutation mechanisms result in transitions when there are purine-purine or pyrimidine-pyrimidine exchanges and transversions in the case of purine-pyrimidine or pyrimidine-purine exchanges. SNPs have gained ground over

Competing interests: The authors have declared that no competing interests exist.

microsatellites as useful genetic markers for genomic selection due to the need for high density markers for multifactorial traits [4].

In humans, approximately 70% of the genome is translated into RNA, but 98.5% are noncoding RNAs (ncRNA), i.e. not translated into protein [5]. Among ncRNAs, microRNAs (miRNA) of approximately 22 nucleotides act as post-transcriptional regulators of gene expression, binding primarily to the 3'UTR region of target genes by antisense complementarity in a conformation known as RISC (RNA-induced silence complex). The seed region comprises nucleotides 2–8 of the miRNA and the recognition of the target genes depends mainly on their pairing [6,7]. Following complex formation, there are two ways of translational control: target miRNA cleavage or translational inhibition [8].

In humans, the occurrence of SNPs at target sites of miRNAs on messenger RNAs (mRNAs) has been reported to alter mRNA:miRNA binding, creating or destroying targets, and potentially regulating a wide variety of diseases, including cancer (reviewed by Moszyńska [9]). Notwithstanding, the presence of SNPs can also affect the miRNA sequence itself [10,11]. In bovine, the effects of SNPs in specific miRNA sequences or miRNA target sites were reported as being related to mastitis [12] and fertility [13].

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 these SNPs on target binding, primary miRNA formation, occurrence of miRNAs with altered target prediction in QTL regions, and the evolutionary conservation for each SNP locus.

Materials and methods

Bovine assembly and database

In this study, we used the *Bos taurus* UMD 3.1.1 [14] as the reference genome. SNP positioning of approximately ~56 million SNPs were acquired from the Run5 of the 1000 Bulls Genome Project (1000BGP; <u>http://www.1000bullgenomes.com/</u>) INDELS were not used in this study. Primary and mature miRNA location, positioning, and sequences were obtained from miRbase (www.mirbase.org), release 21 [15–21]. Positions of QTLs were extracted from the Animal QTLdb [22]. Data mining for each mature miRNA that alter target gene prediction was performed in FAANGmine (http://128.206.116.18:8080/faangmine/begin.do) [23].

Mapping of SNPs in miRNA sequences

An in-house Python script (2.7 version) was used to identify all SNPs (Run5 of 1000BGP) that are located within all bovine miRNA sequences available on miRBase (release 21).

Prediction of miRNA targets

The miRmap [24] package was adapted to input both miRNA reference and SNP altered sequences for target prediction through Python scripts (3.6.4 version), and using 13,345 3'UTR bovine target sequences available within the same package. Target prediction was performed for all miRNAs containing SNPs. To validate microRNA targets predicted by the miRmap software, we also used the algorithm employed by the TargetScan software [25] to perform target prediction. We used the 3' UTR database available on the miRmap package against our miRNA reference sequences and sequences altered by SNP. Common targets, predicted by softwares, are described in Fig 2 and in supplementary S1 File. 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 [26] was utilized. This tool can calculate changes in MFE, which can affect the formation of mature miR-NAs, 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. The threshold 2.0 applied was previously reported by Gong et al. as sufficient to alter the MFE [27].

QTLs

An in-house Python script (version 2.7) was used to identify which SNPs, with altered target prediction, were located in QTL mapped regions available in Animal QTLdb [22].

PhyloP scores

PhyloP scores are used to measure non-neutral nucleotide substitution rates, indicating conservation and acceleration independently at each nucleotide [28]. To calculate PhyloP scores [29], we performed a BLAST [30] search (e-value $\leq 10^{-3}$) [31] of all primary miRNA sequences against the human genome, then the identified human sequences were located on the UCSC Genome Browser [32,33]. Once sequences were located, an alignment was performed using MUSCLE [34], 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.

Results

SNPs in miRNA primary sequences

We performed a scan in 1,064 miRNA primary sequences and found 452 miRNA sequences with SNPs. Out of 452 miRNAs with SNPs, 193 of those had SNP(s) located within miRNA's mature sequences, with 91 of them within seed regions. Some miRNAs harbored more than one SNP per sequence. Nucleotide changes were comprised of 47% transitions and 53% transversions (Fig 1).

Changes in target prediction

Out of 193 miRNA mature sequences with SNPs, 91 had SNPs occurring within their seed region, and 70 of these 91 (77%) had changes in target prediction as a consequence of the nucleotide present. In regards to target prediction alteration observed for the 70 miRNAs, 48% of all targets were newly created, and 35.1% were lost (Fig 2a and 2b), as a consequence of the presence of a SNP. Less than one third 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 (1,101 vs 1,238 average target alteration, respectively). The list of target genes for reference and altered miRNAs is in <u>S1 File</u>.

(a)





https://doi.org/10.1371/journal.pone.0249406.g001

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 ($\Delta\Delta G$) between reference and altered sequences of 917 miRNA sequences (Fig 3A) using the RNAfold Program [26]. 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 and 3C) in 434 primary miRNAs. In the present study, we found 58% transitions and 42% transversions, and statistical analysis indicated that there was a significant difference regarding MFE variation between these two types (-0.83 vs -1.61, transitions and transversions, respectively; Welch Two Sample t-test p-value<2.348e-05).



Fig 2. Changes in target prediction. Prediction of gained, lost and unaltered targets of mature miRNAs containing SNP (a). Venn diagram depicting all created, lost and maintained targets (b).

https://doi.org/10.1371/journal.pone.0249406.g002

Presence of miRNAs with SNPs in QTL regions

Focusing on the 70 mature miRNAs with altered target prediction, we crossed their positions with that of known QTL regions available on Animal QTLdb (Table 1). We found 46 traits that had one or more SNPs in 47 different miRNA sequences within their locations.

PhyloP conservation scores for SNPs within primary sequences

To investigate the conservation level of each altered base occurring in miRNA primary sequences, we obtained the PhyloP score based on the alignment of 100 vertebrates [29] relative to human genome. Positive, neutral and negative values are employed to indicate slow,





https://doi.org/10.1371/journal.pone.0249406.g003

neutral and fast-evolving variations, respectively. Of the 917 primary miRNAs sequences containing SNPs, we were able to confidently align 214 bovine sequences to homologous sequences in humans, in order to obtain the phyloP score of each SNP position (Fig 4). All miRNAs containing a SNP in their seed region had a positive phyloP score, regardless of target alteration.

Trait categories	miRNA
Reproduction	
Calving ease	bta-miR-181c
	bta-miR-23a
	bta-miR-2881
	bta-let-7e
	bta-miR-574
	bta-miR-2900
	bta-miR-199c
	bta-miR-2339
	bta-miR-584
	bta-miR-2886
Calving ease (maternal)	bta-miR-2284i
	bta-miR-22850
	bta-miR-551a
	bta-miR-2325c
Calf size	bta-miR-2414
	bta miR-22840
	bta-miR-2387
Birth index	bta miP 2328 5p
bitti index	bta-miR-2901
	bta-miR-2899
Stillbirth	hts_miR_199a_5n
	bta-miR-371
Scrotal circumference	bta miP 2418
scrotarence	bta-miR-2418
Udder swelling score	hta_miR_2293
ouder swelling score	bta-miR-2295
Calving to conception interval	bta-miR-2419-5p
Age at puberty	bta-miR-6121-3p
Calving index	bta-miR-2480
Length of productive life	bta miR 2400
Nilk and duction	Uta-IIIIK-2292
Milk production Milk caproic acid content	
	bta-miR-2284r
	bta miP 1
	bta-miR-2307
	bta-miR-544b
	bta-miR-7858
	bta-miR-449b
	bta-miR-877
	bta-miR-6531
	bta-miR-2284c
	bta-miR-219
Milk palmitoleic acid content	bta-miR-2286
	bta-miR-2284y
	bta-miR-2354
	bta-miR-488
	bta-miR-669
	bta-miR-6532
Milk butyric acid content	bta-miR-2284u
	bta-miR-181b
	bta-miR-2344
	Dta-miR-2345

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

(Continued)

Trait categories	miRNA
Milk decenoic acid content	bta-miR-2318
	bta-miR-584
	bta-miR-2369
Milk myristic acid content	bta-miR-2285y
	bta-miR-22850
	bta-miR-2400
Milk capric acid content	bta-miR-2895
	bta-miR-1287
	bta-miR-2285K
305-day milk yield	bta-miR-1839
	bta-miR-2309
Milk oleic acid content	
	bta-miR-22851
Mills manietalais a sid contant	hts miD 1256
Milk myristoleic acid content	bta miR 1721
Mille sheles he at the second second second	http://D.155
Milk alpha-lactalbumin percentage	bta-miR-155
Mille launa lais a sid comtant	hts miD 2422h
	bta-miR-1434-5p
Milk vield	bta_miR_22850
wink yield	bta-miR-584
Milk beta-casein percentage	bta-miR-126-5p
Milk conjugated linoleic acid content	hta-miR-2465
Conformational	
Pody weight (weaping)	hta miD 2285h
Body weight (wearing)	bta-miR-22850
	bta-miR-1584-3p
	bta-miR-8550
	bta-miR-2285u
	bta-miR-6519
	bta-miR-2324
	bta-miR-6530
	bta-miR-128
Body weight (yearling)	bta-miR-302b
Residual feed intake	bta-miR-22841
Somatic call score	bta miR 30d
	http://D.2466.2m
Average daily gain	bta-miR-2466-3p
Body weight (birth)	bta-m1R-2362
Conformation score	bta-miR-2402
Non-return rate (EBV)	bta-miR-2421
Meat characteristics	
Shear force	bta-miR-2284ab
	bta-miR-2446
	bta-miR-455-5p
	Dta-miR-6536
	bta-miR-001
	bta-miR-2398

Table 1. (Continued)

(Continued)

Trait categories	miRNA
Intramuscular fat	bta-miR-2425-3p bta-miR-1291 bta-miR-2429 bta-miR-2430 bta-miR-2433 bta-miR-2434
Tridecylic acid content	bta-miR-2313-3p bta-miR-6518 bta-miR-2284d bta-miR-2316 bta-miR-144
Margaric acid content	bta-miR-2885 bta-miR-2406 bta-miR-2408
Subcutaneous fat	bta-miR-2489
Muscle iron content	bta-miR-6533
Trans-15-C18:1 fatty acid content	bta-miR-2305
Monounsaturated fatty acid content	bta-miR-653
Cis-10 Heptadecenoic acid content	bta-miR-2381
Environmental adaptation	
Cold tolerance	bta-miR-6120-3p bta-miR-2284y bta-miR-7691 bta-miR-2325a
Heat tolerance	bta-miR-2397-5p
Disease tolerance	
Tick resistance	bta-miR-2284z bta-miR-199a-5p
Bovine tuberculosis susceptibility	bta-miR-2883

Table 1. (Continued)

https://doi.org/10.1371/journal.pone.0249406.t001

Discussion

We were able to contrast the largest SNP database available in the bovine species to date, with all available bovine miRNA sequences. The publicly available SNP databases, Bovine Genome Database [35] and SNPchiMp v.3 [36,37], that gathers data from SNPs used on several chips produced by Illumina and Affymetrix, have a combined total of 2 million SNP annotations, whereas the dataset used in this study, Run 5 of the 1000BGP, has more than 56 million SNP annotations.

In our study, SNPs within seed sequences greatly affected target prediction, with less than 20% of similar targets between miRNAs containing the reference or altered alleles [38]. Our prediction demonstrated a large number of targets for each miRNA, as observed in humans, and given that 90% of human genes are estimated to be regulated by miRNAs [39], we suggest that SNPs within miRNAs may have a significant impact on cattle phenotypes. Target gain and loss due to the presence of SNPs in microRNA sequences or miRNA binding sites have been previously reported for 41 human miRNAs, affecting approximately 50% of their predicted target gain in 1 of the evaluated miRNAs [40]. In another study, 30 million SNPs were mapped to 1048 human pre-miRNAs, and target prediction detected a wide variation between reference and altered miRNAs, with an average gain of 1,790 targets among the three miRNAs with the highest gains, and of 1,576 targets lost among the three with the highest losses [27]. This



Fig 4. PhyloP conservation scores. Frequency distribuition of PhyloP scores for each SNP position in miRNA sequences. Positive, neutral and negative values are employed to indicate slow, neutral and fast-evolving variations, respectively.

https://doi.org/10.1371/journal.pone.0249406.g004

previously study also validated 4 miRNAs that lost targets and 1 miRNA that gained targets, by luciferase assays, confirming the prediction. Our results in cattle suggest that the same biological mechanisms are applicable to both species as we demonstrate a major impact on target prediction due to the occurrence of SNPs in the miRNAs seed region. Changes in prediction of miRNA targets as a consequence of SNPs are not limited to the animal kingdom. In plants, a 26% target gain was observed for 13 miRNAs with SNPs in their seed region [41].

We found a significant increase in the variation of MFE in the transversions when compared to transitions. This variation was also reported by Guo et al. [42], however, their study considered only SNPs in coding regions, so the present study suggests that the type of substitution also interferes with the regulatory effects of non-coding RNAs. Presence of SNPs in primary miRNAs sequences, such as observed in 434 miRNAs in our study, may also alter the production of mature miRNAs. This is due to alterations in recognition by their cleavage proteins, which identify the conformation of the secondary structure of the molecule and not simply the nucleotide sequence [43,44]. In fact, association between the presence of SNPs in the pre-miRNA flanking region and cancer has been reported in humans [45]. Alteration in MFE in miRNA primary sequences, caused by SNPs, are also reported and validated in carp (Cyprinus carpio) [46]. Positive values greater than 2.0 account for reducing mature miRNA production and negative values for increasing mature miRNA production in humans and viruses (reviewed by Gong et al. 2012) [27]. In our study 48.6% of miRNAs primary sequences with SNPs showed $\Delta\Delta G$ greater than this value. MFE alteration between reference and altered premiRNA-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 a loss of base pairing [47].

In our study, we have identified 125 miRNAs with SNPs in previously identified QTL regions. These loci are large, spanning several Kbs within chromosomes, therefore caution

should be taken when discussing point variations in such large regions. Notwithstanding, 47 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 in non-coding regions of the genome. 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 [48], however non-protein-coding sequences, outside of promoter regions, have been largely overlooked as possible functional variations in cattle. One of the few studies associating such regions to phenotypes is a published GWAS analysis that detected the presence of two SNPs within the same LD block (separated by 80 kb) that, together, were related to calving ease. This region is devoid of protein coding genes, but contains a miRNA, bta-mir-1256 [49]. Interestingly, our prediction of target gene binding indicated that a SNP present in bta-miR-1256 causes this miRNA to target LCORL for downregulation. The LCORL gene encodes the PALI1 and PALI2 proteins, essential for the activation of the PRC2.1 and PRC2.2 complexes and regulation of the suppressive histone mark H3K27me3, essential for the normal development, as demonstrated in mice [50]. The LCORL gene is highly conserved among mammals and is related to skeletal height and size in humans [51-53], body size in horses [54-57], dogs [58] and pigs [59], body weight in chickens [60,61] and sheep [62]. Previous studies in cattle have linked the LCORL gene with body conformation indexes in Nordic red cattle [63], carcass weight [64] and organ weight [65] in Simmental cattle, growth in Quinchuan cattle [66] and stature in Limousin, Hereford and Angus [67], thus, suggesting that the acquisition of regulation of LCORL, caused by a SNP in btamiR-1256, may affect overall size and therefore calving ease.

In addition, Purfield et al. [31] detected a SNP close to the *IGFBP-2* gene in a QTL related to calving difficulty. Human growth hormone (GH) regulates the expression of IGF (insulin-like growth factor) and the deficiency of GH / IGF1 results in short stature in children. IGFBP (Insulin-like growth factor-binding protein) belong to a group of 6 proteins that bind to IGF and have as main function to regulate the access of IGF1 and 2 to the IGF1 receptor [68,69].

Evolutionary conservation of miRNAs, as well the conservation of their target genes, have already been evidenced in previous studies [70,71]. 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 the 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 suggest unfavorable phenotypes, influenced by altered maturation of primary miRNAs or misregulated post-transcriptional repression promoted by mature miRNAs.

Among the miRNAs with altered target-binding prediction, we highlight miR-15a and 23a, which have their binding to target genes previously validated in cattle. MiR-15a targets the *CD163* gene and was recently identified by Ju *et al.*, where increased expression levels of these miRNAs in the udder of Holsteins were associated with mastitis [72]. Interestingly, our study revealed that a SNP present in the seed region in miR-15a is predicted to alter its binding with *CD163*, which could influence immune response to infection. In 2017, Guan et al. identified the *ZNF423* gene as a target for miR-23a, affecting cellular differentiation in bovine fetal muscle. Authors reported that the level of expression of miR-23a is low during cell differentiation, while the forced overexpression of a miR-mimic 23a leads to a reduction in the accumulation of lipids [73]. Gupta el at. (2010) showed that *ZNF423* acts as a positive regulator of PPARy

expression, promoting the differentiation of undifferentiated cells into adipocytes. They have also demonstrated that the reduction of *ZNF423* expression, through small RNAs, negatively regulates the rate of adipocyte differentiation [74]. In the present study, a SNP in the miR-23a resulted in the loss of *ZNF423* as a target, and in the gain of *ZNF521* which, interestingly, was previously reported as a negative regulator of *ZNF423* in mice [75]. Thus, the loss of *ZNF423* as a target for miR-23a, in addition to downregulation of *ZNF521*, caused by the presence of a SNP in the miR-23a seed region, can result in increased adipogenesis and could impact meat marbling.

A phenome search carried out through the FAANG (Functional Annotation of Animal Genomes) Consortium [23], using our 70 miRNAs with altered target prediction, pointed to a role for miR-29e on immune response. Lawless et al. [76] predicted 2,491 target genes for differentially expressed miRNAs identified after infection by *Streptococcus uberis* in bovine mammary epithelial cells. A pathway analysis highlighted MAPK signalling, Cytokine-Cytokine Receptor Signalling and the JAK-STAT Signalling Pathway as important pathways for mastitis response in these cells. Of the 2,491 target genes predicted by the authors, 93 genes were predicted as targets for mir-29e, which was upregulated at 4 hours of infection and downregulated at 6 hours. When we compare these targets with the prediction performed in our study, the presence of a SNP in the miR-29e seed region results in the loss of 52 of these 93 target genes and a gain of 4 target genes, suggesting that a SNP in the seed region of this miR-29e miRNA can alter more than 50% of the target genes for this miRNA, as predicted by Lawless et al., potentially affecting important pathways related to innate immunity in response to gram-positive bacterial infection.

Finally, the present study points to an approach focused on differences in genomic sequences that are not present in coding genes, but can indirectly generate differences in the phenotype of individuals, through epigenetically regulated processes. In fact, our group has previously identified SNPs that, through the creation or destruction of CpG sites, may mechanistically induce a quantitative difference in methylation, which in turn modulates gene expression and consequently phenotypes [77]. Out of 1,064 bovine miRNAs presently described in miRbase, only 265 have target genes that have been experimentally validated [78] and, of these, only 3.7% have targets within their own chromosome. This further highlights the importance of considering point changes away from protein-coding regions, as those can affect genes located in distant loci and indirectly explain phenotypic variation between animals through epigenetic regulation.

Conclusions

In the present study we demonstrated that SNPs in miRNA sequences can alter their target prediction, as well as pre-miRNA processing, possibly increasing or decreasing mature miRNA production, which in turn could affect production traits in cattle.

Supporting information

S1 File. List of target genes predicted for reference and altered miRNAs. (XLSX)

Acknowledgments

Authors would like to thank Dr. Jeremy F. Taylor for providing access to the SNP annotation from the international private consortium 1k Bulls Genomes Project.

Author Contributions

- **Conceptualization:** Marco Antônio Perpétuo de Sousa, Mariângela Bueno Cordeiro Maldonado, Marina Rufino S. Fortes, Flavia Lombardi Lopes.
- **Data curation:** Marco Antônio Perpétuo de Sousa, Flavia Regina Florêncio de Athayde, Mariângela Bueno Cordeiro Maldonado, Andressa Oliveira de Lima, Marina Rufino S. Fortes, Flavia Lombardi Lopes.
- Formal analysis: Marco Antônio Perpétuo de Sousa, Marina Rufino S. Fortes, Flavia Lombardi Lopes.
- Funding acquisition: Flavia Lombardi Lopes.
- Investigation: Marco Antônio Perpétuo de Sousa, Andressa Oliveira de Lima.
- Methodology: Marco Antônio Perpétuo de Sousa, Flavia Regina Florêncio de Athayde, Mariângela Bueno Cordeiro Maldonado, Marina Rufino S. Fortes, Flavia Lombardi Lopes.
- Project administration: Marco Antônio Perpétuo de Sousa.
- Resources: Marina Rufino S. Fortes, Flavia Lombardi Lopes.
- **Software:** Marco Antônio Perpétuo de Sousa, Flavia Regina Florêncio de Athayde, Mariângela Bueno Cordeiro Maldonado.
- **Supervision:** Flavia Regina Florêncio de Athayde, Andressa Oliveira de Lima, Marina Rufino S. Fortes, Flavia Lombardi Lopes.
- Visualization: Marco Antônio Perpétuo de Sousa, Marina Rufino S. Fortes, Flavia Lombardi Lopes.
- Writing original draft: Marco Antônio Perpétuo de Sousa.
- Writing review & editing: Flavia Regina Florêncio de Athayde, Mariângela Bueno Cordeiro Maldonado, Andressa Oliveira de Lima, Marina Rufino S. Fortes, Flavia Lombardi Lopes.

References

- 1. 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 framment length polymorphism). Genetics. 1989; 86: 2766–2770.
- Bourneuf E, Otz P, Pausch H, Jagannathan V, Michot P, Grohs C, et al. Rapid Discovery of De Novo Deleterious Mutations in Cattle Enhances the Value of Livestock as Model Species. Sci Rep. 2017; 7:11466: 1–19. https://doi.org/10.1038/s41598-017-11523-3 PMID: 28904385
- 3. Wiggans GR, Cole JB, Hubbard SM, Sonstegard TS. Genomic Selection in Dairy Cattle: The USDA Experience *. Annu Rev Anim Biosci. 2017; 5: 1–19.
- 4. Vignal A. A review on SNP and other types of molecular markers and their use in animal genetics. Genet Sel Evol. 2002; 34: 275–305. https://doi.org/10.1186/1297-9686-34-3-275 PMID: 12081799
- Gayon J. From Mendel to epigenetics: History of genetics. Comptes rendus—Biol. 2016; 339: 225–230. https://doi.org/10.1016/j.crvi.2016.05.009 PMID: 27263362
- Bartel DP, Lee R, Feinbaum R. MicroRNAs: Genomics, Biogenesis, Mechanism, and Function. Cell. 2004; 116: 281–297. https://doi.org/10.1016/s0092-8674(04)00045-5 PMID: 14744438
- Saliminejad K, Khorram Khorshid HR, Soleymani Fard S, Ghaffari SH. An overview of microRNAs: Biology, functions, therapeutics, and analysis methods. J Cell Physiol. 2018; 234: 5451–5465. <u>https://doi. org/10.1002/jcp.27486</u> PMID: 30471116
- Wienholds E, Plasterk RHA. MicroRNA function in animal development. FEBS Lett. 2005; 579: 5911– 5922. https://doi.org/10.1016/j.febslet.2005.07.070 PMID: 16111679
- Moszyńska A, Gebert M, Collawn JF, Bartoszewski R. SNPs in microRNA target sites and their potential role in human disease. Open Biol. 2017; 7: 1–13. https://doi.org/10.1098/rsob.170019 PMID: 28381629

- 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. https://doi.org/10.3390/genes10030185 PMID: 30818878
- 11. 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; 1–7.
- 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. https://doi.org/10.1007/s12192-017-0848-z PMID: 29476341
- 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: 1–9. https://doi.org/10.1371/journal.pone.0084355 PMID: 24416221
- Zimin AV., 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. <u>https://doi.org/10.1186/gb-2009-10-4-r42</u> PMID: 19393038
- Ambros V, Bartel B, Bartel DP, Burge CB, Carrington JC, Chen X, et al. A uniform system for microRNA annotation. RNA. 2003. https://doi.org/10.1261/rna.2183803 PMID: 12592000
- Griffiths-Jones S. The microRNA Registry. Nucleic Acids Res. 2003. <u>https://doi.org/10.1093/nar/gkh023 PMID: 14681370</u>
- 17. Griffiths-Jones S. miRBase: microRNA sequences, targets and gene nomenclature. Nucleic Acids Res. 2005. https://doi.org/10.1093/nar/gkj112 PMID: 16381832
- Griffiths-Jones S, Saini HK, Van Dongen S, Enright AJ. miRBase: Tools for microRNA genomics. Nucleic Acids Res. 2008. https://doi.org/10.1093/nar/gkm952 PMID: 17991681
- Kozomara A, Griffiths-Jones S. MiRBase: Integrating microRNA annotation and deep-sequencing data. Nucleic Acids Res. 2011. https://doi.org/10.1093/nar/gkq1027 PMID: 21037258
- Kozomara A, Griffiths-Jones S. MiRBase: Annotating high confidence microRNAs using deep sequencing data. Nucleic Acids Res. 2014. https://doi.org/10.1093/nar/gkt1181 PMID: 24275495
- Kozomara A, Birgaoanu M, Griffiths-Jones S. miRBase: from microRNA sequences to function. Nucleic Acids Res. 2019. https://doi.org/10.1093/nar/gky1141 PMID: 30423142
- 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. <u>https://doi.org/10.1093/nar/gky1084</u> PMID: 30407520
- 23. Giuffra E, Tuggle CK. Functional Annotation of Animal Genomes (FAANG): Current Achievements and Roadmap. Annu Rev Anim Biosci. 2019; 7: 65–88. <u>https://doi.org/10.1146/annurev-animal-020518-114913 PMID: 30427726</u>
- 24. Vejnar CE, Zdobnov EM. MiRmap: Comprehensive prediction of microRNA target repression strength. Nucleic Acids Res. 2012. https://doi.org/10.1093/nar/gks901 PMID: 23034802
- Agarwal V, Bell GW, Nam JW, Bartel DP. Predicting effective microRNA target sites in mammalian mRNAs. Elife. 2015; 4: 1–38. https://doi.org/10.7554/eLife.05005 PMID: 26267216
- Lorenz R, Bernhart SH, Höner zu Siederdissen C, Tafer H, Flamm C, Stadler PF, et al. ViennaRNA Package 2.0. Algorithms Mol Biol. 2011. https://doi.org/10.1186/1748-7188-6-26 PMID: 22115189
- Gong J, Tong Y, Zhang HM, Wang K, Hu T, Shan G, et al. Genome-wide identification of SNPs in Micro-RNA genes and the SNP effects on MicroRNA target binding and biogenesis. Hum Mutat. 2012; 33: 254–263. https://doi.org/10.1002/humu.21641 PMID: 22045659
- Pollard KS, Hubisz MJ, Rosenbloom KR, Siepel A. Detection of nonneutral substitution rates on mammalian phylogenies. Genome Res. 2010; 20: 110–121. <u>https://doi.org/10.1101/gr.097857.109</u> PMID: 19858363
- Hubisz MJ, Pollard KS, Siepel A. Phastand Rphast: Phylogenetic analysis with space/time models. Brief Bioinform. 2011; 12: 41–51. https://doi.org/10.1093/bib/bbg072 PMID: 21278375
- **30.** Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990; 215: 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2 PMID: 2231712
- Pinhal D, Bovolenta LA, Moxon S, Oliveira AC, Nachtigall PG, Acencio ML, et al. Genome-wide micro-RNA screening in Nile tilapia reveals pervasive isomiRs' transcription, sex-biased arm switching and increasing complexity of expression throughout development. Sci Rep. 2018. https://doi.org/10.1038/ s41598-018-26607-x PMID: 29844338
- James Kent W, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, et al. The human genome browser at UCSC. Genome Res. 2002. https://doi.org/10.1101/gr.229102 PMID: 12045153 Article published online before print in May 2002.

- Karolchik D, Hinrichs AS, Kent WJ. The UCSC genome browser. Curr Protoc Hum Genet. 2011. <u>https://doi.org/10.1002/0471142905.hg1806s71 PMID: 21975940</u>
- Edgar RC. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 2004; 32: 1792–1797. https://doi.org/10.1093/nar/gkh340 PMID: 15034147
- 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: D834–D839. https://doi.org/10.1093/nar/gkv1077 PMID: 26481361
- 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: 1–6.
- 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: 1–6.
- 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: 1124–1131. <u>https://doi.org/10.1093/hmg/ddm062</u> PMID: 17400653
- Latini A, Ciccacci C, Novelli G, Borgiani P. Polymorphisms in miRNA genes and their involvement in autoimmune diseases susceptibility. Immunol Res. 2017; 65: 811–827. <u>https://doi.org/10.1007/s12026-017-8937-8 PMID: 28741258</u>
- 40. 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: A2. https://doi.org/10.1186/1471-2105-13-s18-a2
- 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: 1–12. https://doi.org/10.1186/s12864-016-3406-7 PMID: 28049423
- 42. 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: 1.
- 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. <u>https://doi.org/10.1261/rna.1560209</u> PMID: 19617315
- 44. Schanen BC, Li X. Transcriptional regulation of mammalian miRNA genes. Genomics. 2011; 97: 1–6. https://doi.org/10.1016/j.ygeno.2010.10.005 PMID: 20977933
- 45. 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: 1099–1109. <u>https://doi.org/10.1007/s10552-013-0187-z PMID: 23526039</u>
- 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: 1.
- 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: 276–278. https://doi.org/10.1080/10495398. 2015.1014491 PMID: 26158459
- Seidel GE Jr.. Brief introduction to whole-genome selection in cattle using single nucleotide polymorphisms. Reprod Fertil Dev. 2010; 22: 138–144. https://doi.org/10.1071/RD09220 PMID: 20003856
- 49. Purfield DC, Bradley DG, Kearney JF, Berry DP. Genome-wide association study for calving traits in Holstein-Friesian dairy cattle. Animal. 2014; 8: 224–235. https://doi.org/10.1017/S175173111300195X PMID: 24256561
- 50. Conway E, Jerman E, Healy E, Koseki H, Bracken AP, Conway E, et al. A Family of Vertebrate-Specific Polycombs Encoded by the LCOR / LCORL Genes Balance PRC2 Subtype Article A Family of Vertebrate-Specific Polycombs Encoded by the LCOR / LCORL Genes Balance PRC2 Subtype Activities. Mol Cell. 2018; 70: 408–421. https://doi.org/10.1016/j.molcel.2018.03.005 PMID: 29628311
- Sovio U, Bennett AJ, Millwood IY, Molitor J, Reilly PFO, Nicholas J, et al. Genetic Determinants of Height Growth Assessed Longitudinally from Infancy to Adulthood in the Northern Finland Birth Cohort 1966. PLoS Genet. 2009; 5: 1–8. https://doi.org/10.1371/journal.pgen.1000409 PMID: 19266077
- Pryce JE, Hayes BJ, Bolormaa S, Goddard ME. Polymorphic Regions Affecting Human Height Also Control Stature in Cattle. Genetics. 2011; 187: 981–984. https://doi.org/10.1534/genetics.110.123943 PMID: 21212230
- Carty CL, Johnson NA, Hutter CM, Reiner AP, Peters U, Tang H, et al. Genome-wide association study of body height in African Americans: the Women 's Health Initiative SNP Health Association Resource (SHARe). Hum Mol Genet. 2012; 21: 711–720. https://doi.org/10.1093/hmg/ddr489 PMID: 22021425

- Signer-hasler H, Flury C, Haase B, Burger D, Simianer H, Leeb T, et al. A Genome-Wide Association Study Reveals Loci Influencing Height and Other Conformation Traits in Horses. PLoS One. 2012; 7: 3– 8. https://doi.org/10.1371/journal.pone.0037282 PMID: 22615965
- 55. Tetens J, Widmann P., Khun C., Taller G. A genome-wide association study indicates LCORL / NCAPG as a candidate locus for withers height in German Warmblood horses. Anim Genet. 2013; 44: 467–471. https://doi.org/10.1111/age.12031 PMID: 23418885
- **56.** Boyko AR, Brooks SA, Behan-braman A, Castelhano M, Corey E, Oliveira KC, et al. Genomic analysis establishes correlation between growth and laryngeal neuropathy in Thoroughbreds. BMC Genomics. 2014; 15: 1–9.
- 57. Taylor P, He S, Zhang L, Li W, He S, Zhang L, et al. BIEC2-808543 SNP in the LCORL Gene is Associated with Body Conformation in the Yili Horse BIEC2-808543 SNP in the LCORL Gene is Associated with Body Conformation in the Yili Horse. Anim Biotechnol. 2015; 26: 289–291. <u>https://doi.org/10.1080/10495398.2014.995303 PMID: 26158461</u>
- Vaysse A, Ratnakumar A, Derrien T, Axelsson E, Pielberg GR, Hansen MST, et al. Identification of Genomic Regions Associated with Phenotypic Variation between Dog Breeds using Selection Mapping. PLoS Genet. 2011; 7: 1–21. https://doi.org/10.1371/journal.pgen.1002316 PMID: 22022279
- Rubin C, Megens H, Martinez A, Maqbool K, Sayyab S. Strong signatures of selection in the domestic pig genome. PNAS. 2012; 109: 19529–19536. <u>https://doi.org/10.1073/pnas.1217149109</u> PMID: 23151514
- Liu Q, Wang H, Zhu L, Hu H, Sun Y. Genome-wide identification and analysis of miRNA-related single nucleotide polymorphisms (SNPs) in rice. Rice. 2013; 6: 1–10.
- Liu R, Sun Y, Zhao G, Wang H. Identification of loci and genes for growth related traits from a genomewide association study in a slow- 3 fast-growing broiler chicken cross. Genes Genomics. 2015. <u>https:// doi.org/10.1007/s13258-015-0314-1</u>
- Kijas JW. Haplotype-based analysis of selective sweeps in sheep. NRC Res Press. 2014; 437: 433– 437.
- Sahana G, Höglund JK, Guldbrandtsen B, Lund MS. Loci associated with adult stature also affect calf birth survival in cattle. BMC Genet. 2015; 16:47: 1–12. https://doi.org/10.1186/s12863-015-0202-3 PMID: 25935543
- 64. Chang T, Xia J, Xu L, Wang X, Zhu B, Zhang L, et al. A genome-wide association study suggests several novel candidate genes for carcass traits in Chinese Simmental beef cattle. Physiol Genomics. 2018; 50: 523–531. https://doi.org/10.1152/physiolgenomics.00022.2018 PMID: 29676954
- 65. An B, Xia J, Chang T, Wang X, Miao J, Xu L, et al. Genome-wide association study identifies loci and candidate genes for internal organ weights in Simmental beef cattle. Physiol Genomics. 2018; 50: 523– 531. https://doi.org/10.1152/physiolgenomics.00022.2018 PMID: 29676954
- Han YJ, Chen Y, Liu Y, Liu XL. Sequence variants of the LCORL gene and its association with growth and carcass traits in Qinchuan cattle in China. J Genet. 2017; 96: 9–17. <u>https://doi.org/10.1007/s12041-016-0732-0 PMID: 28360384</u>
- 67. Doyle JL, Berry DP, Veerkamp RF, Carthy TR, Walsh SW, Evans RD, et al. Genomic Regions Associated With Skeletal Type Traits in Beef and Dairy Cattle Are Common to Regions Associated With Carcass Traits, Feed Intake and Calving Difficulty. Front Genet. 2020; 11: 1–16.
- Clemmons DR. Role of IGF-binding proteins in regulating IGF responses to changes in metabolism. J Mol Endocrinol. 2018; 61: T139–T169. https://doi.org/10.1530/JME-18-0016 PMID: 29563157
- Bach LA. IGF-binding proteins. J Mol Endocrinol. 2018; 61: T11–T28. https://doi.org/10.1530/JME-17-0254 PMID: 29255001
- Friedman RC, Farh KKH, Burge CB, Bartel DP. Most mammalian mRNAs are conserved targets of microRNAs. Genome Res. 2009; 19: 92–105. https://doi.org/10.1101/gr.082701.108 PMID: 18955434
- Fehlmann T, Laufer T, Backes C, Kahramann M, Alles J, Fischer U, et al. Large-scale validation of miR-NAs by disease association, evolutionary conservation and pathway activity. RNA Biol. 2019; 16: 93– 103. https://doi.org/10.1080/15476286.2018.1559689 PMID: 30567465
- 72. Ju Z, Huang J, Jiang Q, Wang C, Wang X, Zhao S. Identification of bta-miR-15a ~ 16a cluster expression, localization and regulated target in Holsteins. Mol Cell Probes. 2018; 40: 8–12. https://doi.org/10.1016/j.mcp.2018.05.003 PMID: 29859243
- Guan L, Hu X, Liu L, Xing Y, Zhou Z, Liang X, et al. bta-miR-23a involves in adipogenesis of progenitor cells derived from fetal bovine skeletal muscle. Sci Rep. 2017; 1–12.
- 74. Gupta RK, Arany Z, Seale P, Mepani RJ, Ye L, Conroe HM, et al. Transcriptional control of preadipocyte determination by Zfp423. Nature. 2010; 464: 619–623. https://doi.org/10.1038/nature08816 PMID: 20200519

- 75. William N. A, Martin M. F, Helen X. Y, Zhao L, Kenichi N, Francesca G, et al. Direct Transcriptional Repression of Zfp423 by Zfp521 Mediates a Bone Morphogenic Protein-Dependent Osteoblast versus Adipocyte Lineage Commitment Switch. Mol Cell Biol. 2014; 34: 3076–3085. https://doi.org/10.1128/ MCB.00185-14 PMID: 24891617
- 76. Lawless N, Foroushani ABK, Mccabe MS, Farrelly CO, Lynn DJ. Next Generation Sequencing Reveals the Expression of a Unique miRNA Profile in Response to a Gram-Positive Bacterial Infection. PLoS One. 2013; 8: 1–13. https://doi.org/10.1371/journal.pone.0057543 PMID: 23472090
- 77. Maldonado BC, DR Neto NB, Nagamatsu ST, Id B SK, Id M SD, Id T JF, et al. Identification of bovine CpG SNPs as potential targets for epigenetic regulation via DNA methylation. PLoS One. 2019;September,: 1–19. https://doi.org/10.1371/journal.pone.0222329 PMID: 31513639
- 78. Chou C, Shrestha S, Yang C, Chang N, Lin Y, Liao K, et al. miRTarBase update 2018: a resource for experimentally validated microRNA-target interactions. Nucleic Acids Res. 2018; 46: 296–302. https:// doi.org/10.1093/nar/gkx1067 PMID: 29126174