Elucidating the landscape of genome-wide chromatin interaction sites of the lncRNA TUG1 in bovine cell lines and liver tissue

Rosemarie Weikard, Raghu Bhushan, Doreen Becker, Frieder Hadlich, Carole Charlier, Gabriel Costa Monteiro Moreira, Christa Kuehn

PII: S0888-7543(25)00151-X

DOI: https://doi.org/10.1016/j.ygeno.2025.111135

Reference: YGENO 111135

To appear in: *Genomics*

Received date: 24 May 2025

Revised date: 27 August 2025

Accepted date: 27 September 2025

Please cite this article as: R. Weikard, R. Bhushan, D. Becker, et al., Elucidating the landscape of genome-wide chromatin interaction sites of the lncRNA TUG1 in bovine cell lines and liver tissue, *Genomics* (2024), https://doi.org/10.1016/j.ygeno.2025.111135

Elucidating the landscape of genome-wide chromatin interaction sites of the lncRNA TUG1 in bovine cell lines and liver tissue

Rosemarie Weikard[a], Raghu Bhushan[a], Doreen Becker[a], Frieder Hadlich[a], Carole Charlier[b], Gabriel Costa Monteiro Moreira[b,c], Christa Kuehn[a,d,e,*] christa.kuehn@fli.de

[a]Institute of Genome Biology, Leibniz Institute for Farm Animal Biology (FBN), Dummerstorf, Germany

[b]Unit of Animal Genomics, GIGA-R and Department of Animal Sciences, Faculty of Veterinary Medicine, University of Liège, Liège, Belgium

[c]BREED, Université Paris-Saclay, UVSQ, INRAE, Jouy-en-Josas, France

[d]Agricultural and Environmental Faculty, University of Rostock, Rostock, Germany

[e]Friedrich Loeffler Institute, Greifswald-Riems, Germany

[*]Corresponding author.

Abstract

Long-noncoding RNA (lncRNA) interact with DNA, RNA and proteins to regulate the epigenome and fundamental biological processes. In the bovine genome, the nature and mechanisms of lncRNA interactions with specific chromatin regions are unexplored yet. Here, we aimed to unravel the chromatin interaction sites of the evolutionary conserved lncRNA *TUG1* in the bovine genome using ChIRP-seq (chromatin isolation by RNA precipitation) in two popular bovine cell lines (MDBK, MAC-T) and liver tissue. About half of the genome-wide *TUG1* chromatin occupancy in the genome (3,225, 3,587 and 3977 interaction sites in MDBK, MAC-T and liver, respectively) was associated with protein-coding genes. Observation of numerous concordant *TUG1* chromatin interaction sites between MDBK and MAC-T cells and liver tissue was consistent with the known ubiquitous expression of *TUG1*. Analysis of overlaps between ChIRP-seq peaks and ATAC-seq peaks in MDBK and MAC-T cells pinpointed *TUG1* chromatin interaction sites relevant for modulating chromatin accessibility.

Keywords

lncRNA; TUG1; Chromatin interaction; ChIRP-seq; bovine genome, FAANG, ATAC-seq

1. Introduction

There is clear evidence that the genomes of many organisms are extensively transcribed resulting in numerous long non-coding RNAs (lncRNAs) in addition to protein-coding transcripts. LncRNAs are usually characterized by a size > 200 nucleotides and a lack of an open reading frame, although literature increasingly describes lncRNAs with a coding potential for short peptides and microproteins [1-3]. LncRNAs are a particular class of transcripts, which exert important functions in genome regulation [4-7]. There is increasing evidence that lncRNAs are important modulators involved in various physiological and pathological processes, including gene regulation, cell development, tissue formation, metabolism and diseases [8-13]. LncRNAs are reported to act as molecular elements by interacting with DNA, RNA and proteins to regulate the epigenomic and diverse fundamental biological processes [14]. They are known to display particular individual regulation patterns [15]. In human and mice, lncRNA-chromatin interactions have been shown to regulate crucial cellular mechanisms such as genomic imprinting, dosage compensation, developmental and pathological processes.

However, lncRNA loci are not completely annotated in the bovine genome [16-18], and the nature and their biological functions and mechanisms of interactions with specific chromatin regions are unexplored so far. Whereas there are several studies investigating the regulatory function of lncRNAs in cattle using, e.g. co-expression network analysis following a "guilt-by-association" approach [19-22], there is almost no functional experimental investigation of lncRNA interaction partners. Since lncRNAs show tissue- and age-specific expression patterns and have less cross-species sequence homology compared to protein-coding genes, these lncRNA interaction features require species-specific efforts for their correct annotation and exploration of their respective interaction partners and potential regulatory function.

In this study, we focused on the lncRNA *Taurine Up-Regulated 1 (TUG1)*, which is, in contrast to other lncRNAs, ubiquitously expressed across cell and tissue types and also conserved between a variety of species [23, 24], https://www.ncbi.nlm.nih.gov/gene/55000/ortholog/?scope=9347). *TUG1* was originally identified as a transcript upregulated by taurine, whose function was initially found being associated with retinal development. It is meanwhile understood that *TUG1* plays a much broader functional role in diverse biological and cellular processes. Numerous studies in cancer research have indicated that *TUG1* is significantly associated with tumor development and cell metabolism by regulating cell proliferation, invasion, metastasis, apoptosis, differentiation and drug resistance reviewed by [25]. The authors summarized that *TUG1* appears to mediate these processes primarily by binding with PRC2 to silence downstream target genes and by sponging miRNAs to promote the expression of target genes. In addition, *TUG1* has been shown to be associated with the regulation of metabolic processes [26]. The authors demonstrated that *TUG1* is markedly downregulated in response to high glucose and reported that its overexpression rescued the expression of PPARGC1A, a master transcription regulator of mitochondrial biogenesis. Using ChIRP-seq (Chromatin Isolation by RNA Purification sequencing), the authors identified a *TUG1*-binding element upstream of the *PPARGC1A* gene and showed that its interaction with *TUG1* enhances *PPARGC1A* promoter activity, leading to modulation of mitochondrial bioenergetics in podocytes of diabetic mice. In a follow-up study, the authors provided

clear evidence that the suppression of *TUG1* expression by a high glucose environment is at least partially regulated by binding of ChREBP, the major glucose-responsive transcription factor, and its partner MLX, to the promoter region of *TUG1* [27]. These findings were supported by results [28], which showed that overexpression of *TUG1* significantly promotes fat oxidation, promotes brown remodelling of white adipose tissue and attenuates the development of diabetes by upregulating the SIRT1/AMPK/ACC signalling pathway in a diabetic mouse model. In their detailed analysis of the murine *TUG1* locus, Lewandowski and colleagues [24] demonstrated that the *TUG1* locus plays an essential role in spermatogenesis and male fertility by showing that *TUG1* knockout mice exhibit a sterile phenotype in males. The authors discovered that the mouse *TUG1* locus has three distinct layers of regulatory activities in mice, two non-coding and one coding. They provided evidence for a *cis*-DNA repressor in the *TUG1* genomic region that modulates many neighbouring genes in a broad range of tissues and in an allele-specific manner in the testes. In addition, they discovered a *TUG1* lncRNA that can regulate genes via a *trans*-based function clustered by tissue type, indicating tissue-specific gene dysregulation. Finally, an open reading frame was detected that is evolutionarily conserved including humans and affects mitochondrial membrane potential when overexpressed, suggesting a potential coding function of the *TUG1* locus [24]. Apart from the evolutionarily conserved ORF confirmed in human cells, it is not known whether the various multimodal regulatory properties of the mouse *TUG1* locus are also present in other species and tissues and a corresponding multifaceted function of orthologous lncRNA *TUG1* loci can only be assumed. Reports focusing on the bovine *TUG1* locus and its functional biological network are not known. To elucidate the landscape of this functional network of *TUG1* in the bovine genome, the aim of this study was to reveal the chromatin-interacting DNA partners of the lncRNA *TUG1* in the genome of bovine cells using Chromatin Isolation by RNA Purification sequencing (ChIRP-seq). With this high-throughput method, DNA and proteins bound to a particular target lncRNA (the specific lncRNA interactome) can be detected and mapped genome-wide [29, 30].

2. Materials and methods

2.1. Cell culture

MDBK (Madin-Darby bovine kidney, ATCC- CCL-22) cells were cultured in EMEM (Minimum Eagle's Essential Medium, Merck) containing 10 % FCS (foetal calf serum), 2 % L-glutamine, 1 % NEAA (non-essential amino acids) and 1 % penicillin-streptomycin). MAC-T (bovine mammary alveolar epithelial) cells [31] were cultured in complete DMEM (Dulbecco's Modified Eagle's Media, Merck) supplemented with 10 % FCS, 2 % L-glutamine, 0.1 mM L-methionine, 0.4 mM L-Lysine, 0.01 mM sodium pyruvate, 1 µg/ml prolactin, 1 µg/ml hydrocortisone, 1 µg/ml insulin, and 1 % penicillin-streptomycin. After reaching 80-90 % confluence, the cells were trypsinized. About 10 million cells (1×10^7) were pelleted (1000 rpm, 5 min, 15 °C), washed in PBS and used for one ChIRP experiment each.

2.2. Liver sample collection for Chromatin Isolation by RNA Purification sequencing (ChIRP-seq)

Fresh liver tissue was collected from the EU licensed slaughter house of the Animal facility at the Research Institute for Farm Animal Biology (FBN), Dummerstorf. Samples were collected either during routine culling slaughter processes outside of animal experiments or taken from animals slaughtered within ethically reviewed and approved research projects of the institute according to the Ethics Committee of the Federal State of Mecklenburg-Western Pomerania, Germany (Landesamt für Landwirtschaft, Lebensmittelsicherheit und Fischerei; LALLF).

2.3. Liver sample preparation for ChIRP-seq

Aliquots of 200 mg fresh liver tissue were minced and crushed using a sterile scalpel, mixed thoroughly with 5 ml of ice-cold PBS supplemented with Protease inhibitor cocktail III (PCI III, Merck-Millipore) and homogenized using a pre-cooled Dounce homogenizer (loose pestle, 10 strokes) on ice. The tissue suspension was diluted with 15 ml of ice-cold PBS supplemented with PIC III (PBS-PIC III, Merck-Millipore) and centrifuged for 5 min with 800 x g at 4°C. The pellet was subjected to ChIRP procedure.

2.4. Expression analysis and resequencing

Total RNA was extracted from MDBK cells, MAC-T cells and bovine liver tissue samples using the NucleoSpin RNA II kit (Macherey-Nagel, Düren, Germany). Genomic DNA was carefully eliminated from RNA preparations by repeated on-column RNAse-free DNase I digestion as described in a previous study [32]. The cDNA was synthesized using the Superscript III RNase H-reverse transcriptase (Thermo Fisher) with oligo(dT) and random

hexamer primers according to the manufacturer's instructions. Expression of *TUG1* was analysed by reverse transcription (RT) -PCR using specific primers (Supplementary Table S1) producing a DNA fragment of 171 bp.

The structure of the *TUG1* locus was analysed by PCR on cDNA and genomic DNA from bovine liver tissue. Amplified cDNA fragments were purified using the NucleoSpin Extractkit II (Macherey & Nagel) and verified by sequencing. Resequencing of PCR amplicons was carried out using Sanger Sequencing on an ABI 3500 Genetic Analyzer capillary sequencer (Life Technologies) with primers used for RT-PCR (Supplementary Table S1).

Cell line aliquots of 4 million cells were used to isolate total RNA. After assessing quantity, purity and integrity, total RNA libraries were prepared using the TruSeq® Stranded Total RNA Library Prep Gold kit (Illumina). After library quantification, total RNA libraries were sequenced to obtain 40 million paired-end 150 bp reads. Sequencing was performed using a NovaSeq6000 sequencing system (Illumina).

In addition to the total RNAseq data generated in this study, total RNA-seq data from frozen liver tissue was available from an earlier study (Nolte et al. 2020). In that study, stranded, ribodepleted and indexed libraries had been prepared with the TruSeq Stranded RNA-Ribo-Zero H/M/R Gold Kit (Illumina, San Diego, CA, USA) from total RNA. Paired-end reads had been sequenced (2 × 100 bp) in a multiplexed design on a HiSeq 2500 Sequencing System (Illumina). Applying a guided alignment approach, the reads had been mapped with HISAT2 v.2.1. at the bovine reference genome ARS-UCD1.2.

2.5. Chromatin Isolation by RNA Purification sequencing (ChIRP-seq)

ChIRP-seq of MDBK and MAC-T cells and bovine liver was performed essentially according to the procedure developed by Chu and colleagues [29]. Briefly, the chromatin of the MDBK and MAC-T cells and bovine liver tissue nuclei was extracted, cross-linked and sonicated. Biotinylated anti-sense tiling oligonucleotides (biotinylated) were designed specific to the selected lncRNA *TUG1* and hybridised to the target chromatin regions. The resulting biotin-labelled complexes consisting of lncRNAs, RNA-binding proteins and genomic DNA (RNA interactome) were captured with streptavidin magnetic beads. After separating the interacting DNA from the complexes, the eluted DNA was used for DNA library preparation and subjected to DNA sequencing. Bioinformatic analysis of ChIRP-seq data was performed to identify chromatin interaction sites targeted by *TUG1* lncRNA. Based on the original methological description of the ChIRP procedure [29], a commercial analysis kit for application on cell samples was available from Merck-Millipore (Magna ChIRP RNA Interactome Kit, #17-10494). Some modifications had to be developed for the application to liver tissue samples, which are indicated accordingly in the detailed method description section available as Supplementary Method File 1. For the preparation of ChIRP-seq libraries from the RNA-associated chromatin DNA extracted using the ChIRP procedure, the Qiaseq Ultralow Input Library Kit (Qiagen) was used.

2.6. Design of the ChIRP probes

Before starting with ChIRP-probe (anti-sense tiling oligonucleotides) design, validation of the lncRNA *TUG1* locus model according to the sequence NR_131898 (NCBI annotation) was performed based on the RNA-seq data complemented by PCR amplification and resequencing as described above. For ChIRP probe design, the ChIRP probe designer (https://www.biosearchtech.com/support/tools/design-software/chirp-probe-designer, LGC Biosearch Technologies) was used. The following parameters for probe design were set: length: 20 nt, GC content: 40-50 %, distance between two adjacent probes: about 80 bp. Regions of the targeted *TUG1* lncRNA containing repetitive motifs were identified and masked using the RepeatMasker program, https://www.repeatmasker.org/). These regions and those of previously known or further experimentally verified exon-intron boundaries (splice junctions) of the *TUG1* lncRNA were excluded from ChIRP probe design. Finally, a total of 36 ChIRP probes (Supplementary Table S1) have been designed for *TUG1* and were synthesised with biotin TEG labelling (HPLC-purified, Sigma-Aldrich).

2.7. Bioinformatic analysis of ChIRP-seq

ChIRP-seq data from each ChIRP library were analysed essentially following the pipeline published by Chu et al. in 2011. After quality control of the reads and adapter and quality trimming, sequencing reads from DNA hybridised with ChIRP probe pools as well as the LacZ probe pool and the corresponding input controls were aligned to the bovine genome (ARS-UCD1.2, complemented with the Y chromosome sequence) using BWA-MEM algorithm (BWA program) (https://github.com/lh3/bwa). The existing PCR replicates generated during the preparation of the sequencing libraries due to the low DNA input had to be removed from the sequencing read pools, which was done with the use of Picard tools (Picard Toolkit, 2019, Broad Institute, GitHub Repository, http://broadinstitute.github.io/picard/). Peak calling was performed using MACS2 (FDR < 0.05, applying default parameters, https://pypi.org/project/MACS2/#history). For peak calling, the signal profiles of the respective input samples served as controls. A ChIRP-seq signal was only accepted as a specific *TUG1* interactome signal, if it occurred with both ChIRP probe pools, but not with the input samples. For this purpose, bedtools was used to test for an overlap of the peak coordinates provided by MACS2. These identified signals were subjected to a further filtering for possible overlap with LacZ-ChIRP probe signals, serving as an additional control and to eliminate false-positive hybridisation signals. The remaining ChIRP signals were additionally subjected to comparative visual inspection across all samples and controls using the Interactive Genome Viewer (IGV). Peaks of ChIRP interactome signals were annotated to the regions in the bovine genome (ARS-UCD1.2, NCBI annotation, release 106) using the HOMER analysis software [33] associating them with positions of nearby gene loci](exons, introns, promoters of mRNA encoding proteins, transcription start (-1kb to +100 bp) and termination (-100 bp to 1 kb) sites, short and long noncoding RNAs] and intergenic regions (FDR < 0.05) according to annotations NCBI, release106 and Ensembl, release 108.

2.8. Functional annotation of TUG1 chromatin interaction sites

To understand biological meaning behind *TUG1* chromatin interaction sites (ChIRP signals), their functional annotation was elucidated using the DAVID tool [34], https://davidbioinformatics.nih.gov/home.jsp), version 6.8. Therefore, the lists of annotated genes located in or near these interaction sites in liver tissue, MDBK and MAC-T cells were submitted separately to the DAVID pipeline to predict associated biological functions and assign them to specific gene ontology (GO) terms (molecular functions, biological processes) and biological KEGG pathways (FDR < 0.05).

2.9. Identification of consensus motifs at TUG1 interaction sites

The lists of chromatin interaction sites of *TUG1* for MAC-T and MDBK cells, respectively, were intersected with those of bovine liver samples to identify concordant *TUG1* target genomic regions. Concordance of interaction sites was accepted, if an overlap of at least one base pair between peak boundaries was present. This intersected list was subjected to *de novo* motif search analysis via HOMER v. 4.11 [33]. The analysis applied the findMotifsGenome search with the options parameters -size 200, -len 10 and -S 20.

2.10. ATAC-seq analysis

ATAC-seq (Assay for Transposase-Accessible Chromatin using sequencing) data for MDBK and MAC-T cells were performed in the European Horizon 2020 project BovReg that aimed at the identification of functionally active genomic features relevant to phenotypic diversity and plasticity in cattle (Grant agreement ID: 815668). ATAC-seq libraries were generated using the ATACseq kit (Active Motif) following the protocol provided by the manufacturer, with an input of about 60,000 cells. Briefly, cell nuclei were treated with a hyperactive Tn5 transposase to fragment the DNA and tag it with Illumina-compatible sequencing adapters. After purification, the tagmented DNA was amplified by PCR using Illumina-compatible primers to generate sequencing-ready libraries. Libraries quality assessment was performed using QIAxcel Advanced System (QIAGEN) and quantified by qPCR using the KAPA Library Quantification Kit (Roche). Illumina NovaSeq 6000 instrument was used for sequencing, with a paired-end (PE) protocol (2x 150 bp), targeting a minimum of 50 M fragments/PE reads by sample.

ATAC-seq data analysis was performed using the publicly available BovReg pipeline (https://github.com/BovReg/nf-core-atacseq). Briefly, sequenced libraries (raw fastq files) were initially accessed by the quality of the sequences (Phred score ≥ 28). Alignment of ATAC-seq reads to the bovine genome (ARS-UCD1.2, complemented with the Y chromosome sequence) was performed using BWA-MEM algorithm (BWA program) (https://github.com/lh3/bwa). After alignment, a stringent filtration criteria was applied, removing reads that mapped to mitochondrial DNA, marked as duplicates, not marked as

primary alignments, unmapped, multi-mapped, containing > 4 mismatches, soft-clipped, with insert size ≥ 500 bp (to enrich for nucleosome free and mononucleosome fractions), mapped to different chromosomes, not in forward - reverse orientation and, if only one read of the pair fails the previously listed criteria.

ATAC-seq peak calling was performed by the same pipeline (https://github.com/BovReg/nf-core-atacseq), using MACS2 (https://pypi.org/project/MACS2/#history) narrow peak mode with q-value cutoff for peak detection ≤ 0.05 and hs/2.7e9 as the effective genome size parameter required by MACS2. Peaks were detected using the animals' corresponding whole-genome sequencing as control (purified gDNA control; --control option), as adopted by Yuan and co-workers [35]. The Fraction of Reads in Peaks (FRiP) score was calculated for each cell line and both exhibited FRiP score > 0.3 (MDBK = 0.712765 & MAC-T = 0.531627).

Results and discussion

3.1. Validation of the bovine TUG1 locus model and its expression in MDBK and MAC-T cells and liver tissue

The ChIRP-seq method requires no prior knowledge of the secondary structure or functional domains of the lncRNA of interest. However, the lncRNA sequence should be large enough to design sufficient anti-sense hybridisation probes, and its exon-intron structure as well as possible splice variants of the lncRNA need to be known. Thus, there was an initial requirement to experimentally elucidate and validate the molecular structure of the bovine *TUG1* locus. Starting from hypotheses generated by RNA-seq datasets of the bovine liver transcriptome available from an earlier study (Nolte et al. 2019) and validation by RT-PCR followed by resequencing of amplified fragments, we found that bovine *TUG1* lncRNA has four different transcript variants with a total length from 6,787-7,485 bp and three exons. The transcript variants show little difference from one another, apart from a slight variation of the splice sites between exons leading to variation in exon size (Figure 1). The *TUG1* transcripts show similarity to the *TUG1* lncRNA locus with accession number NR_131898 annotated in the *Bos taurus* genome assembly ARS-UCD 2.0 in the NCBI database, release 106 (BTA17: 69,863,342-69,871,056, 4890 bp) as displayed in Figure 1. However, we found in our RNA-seq datasets that the first and third exons are longer in size compared to those in NR_131898 (Figure 1). The exon-intron structure of the *TUG1* locus and the variation in exon 2 have been validated and confirmed by RT-PCR and resequencing (Figure 1).

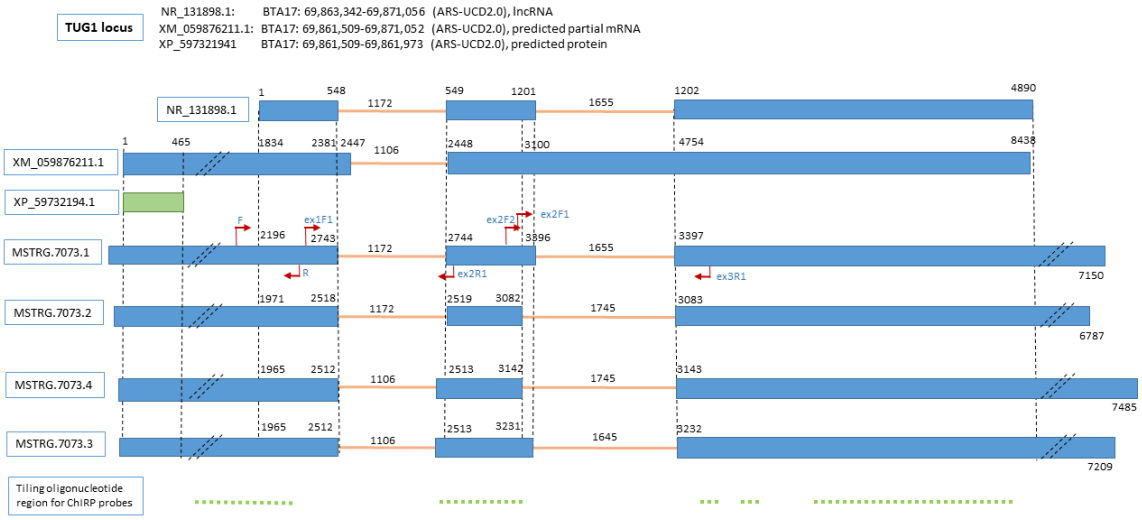

Figure 1. Graphical overview of transcript variants of *TUG1* lncRNA detected in bovine liver tissue

TUG1 transcripts variants MSTRG.7073: 1-4 were found in the liver transcriptome according to bovine liver RNA-seq data from a previous study (Nolte et al. 2019). Blue boxes: exons, orange lines: introns. ChIRP probes were derived from non-repeat-containing regions of the *TUG1* consensus sequence (green dots). Red arrows indicate the location of primers used for PCR and resequencing. Predicted partial mRNA XM_059876211.1 and XP_597321941 annotated as *TUG1* locus model in the *Bos taurus* NCBI annotation, ARS-UCD 2.0, release 106.

As shown in Figure 1, for the *TUG1* locus the latest bovine NCBI genome annotation, ARS-UCD 2.0, release 106, contains an additional model for a predicted partial mRNA with accession number XM_059876211.1, which has a length of 8,438 bp and differs in structure from NR_131898. Retaining intron 2, it encloses only two exons (located on BTA17: 69,861,509-69,871,052). A part of its first exon (465 bp, 69,861,509-69,861,973) is predicted to encode a small protein of 154 amino acids (XP_597321941). This predicted TUG1 protein is conserved between species, which is shown by respective orthologs found in human and mice [24] and is predicted for a variety of other species [23], https://www.ncbi.nlm.nih.gov/gene/55000/ortholog/?scope=9347). There are discrepancies regarding the annotation of the *TUG1* locus in the Ensembl database (ARS-UCD 1.3, release 111). Here, two *TUG1* loci are annotated separately on BTA17: a two-exonic *TUG1*-lncRNA (ENSBTAG00000096942.1, 3,267 bp, 69,865,004-69,871,043) and a *TUG1* mRNA (ENSBTAT00000099970.1), with an exon of 453 bp, encoding a TUG1 protein of 150 amino acids (ENSBTAP00000084776.1, 69,861,521-69,861,973). Both *TUG1* loci are separated by an intergenic genomic region of more than 3 kb according to the Ensembl annotation. However, our experimental data provide clear evidence for a three-exon structure of *TUG1*.

The design of anti-sense tiling oligonucleotides for ChIRP-seq (indicated by the green dots in Figure 1 and described in Supplementary Table S2), relied on non-repeat-containing regions of the *TUG1* sequence, which we had experimentally verified. The potential protein

coding sequence (XP_597321941) is excluded from probe design to make sure that the *TUG1*-DNA interaction signals obtained are attributable to *TUG1* lncRNA.

The expression of *TUG1* in liver and cells of bovine cell lines is shown in Figure 2. Consistent expression of *TUG1* is in line with data from RNA-seq, RT-PCR and respective data from our earlier studies.

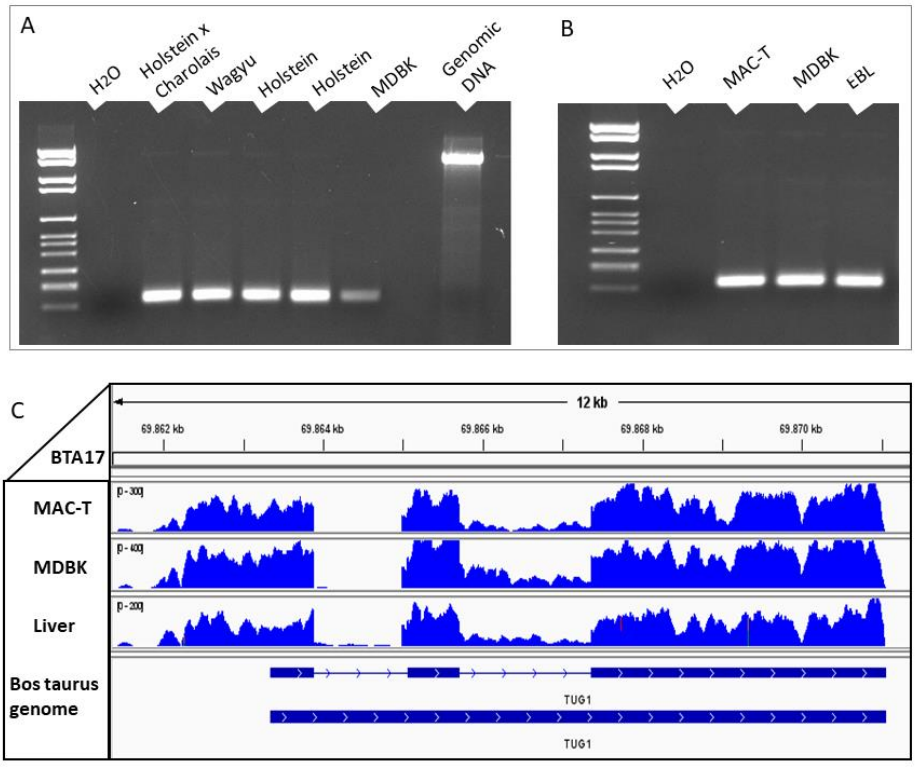

Figure 2. Expression of *TUG1* in bovine cell lines and liver tissue
A: RT-PCR in cDNA from liver tissue samples from different cattle breeds (lanes 2-5) and MDBK (Madin Darby Bovine Kidney, lane 6) cells; B: RT-PCR in cDNA from different bovine cell lines MDBK, MAC-T (bovine mammary alveolar epithelial cells) and EBL (bovine embryonic lung cells, ACC192, as control not used for ChIRP-seq). Primers producing a fragment of 171 bp are given in Supplementary Table S1. C: Total RNA-seq profiles of *TUG1* in MDBK and MAC-T cells and bovine liver.

3.2. Identification of DNA interacting sites of TUG1 lncRNA in MDBK and MAC-T cells and liver tissue using ChIRP-seq

ChIRP-seq was applied to unravel genome-wide chromatin-interacting DNA partners of *TUG1* lncRNA in the bovine genome. Therefore, a total of 36 ChIRP probes (anti-sense tiling oligonucleotides, biotinylated) have been designed for *TUG1* from the non-repeat-containing regions of the *TUG1* consensus sequence (Figure 1) resulting in two

hybridisation pools (even and odd) containing 18 probes each (Supplementary Table S1). In parallel to the hybridisation of the chromatin extracts with these two specific *TUG1* ChIRP probe pools, a separate hybridisation reaction of a sample replicate was carried out along with a separate control with oligonucleotide probes that specifically bind to the LacZ RNA, which is not expressed in mammalian genomes. This approach served as an experimental negative control, and genomic regions indicating specific lncRNA-DNA interaction should not show any signals with the LacZ control. As shown in Supplementary Figure 1, ChIRP probes were able to specifically capture *TUG1* sequence demonstrating that they are suitable for finding *TUG1* chromatin Interaction sites at genome-wide scale. In contrast to the negative control LacZ, strong *TUG1* signals were observed for the even and odd probe pools.

In our final evaluation of the *TUG1* interactome signals in MDBK and MAC-T cells and liver tissue after sequencing of the RNA-associated DNA (ChIRP-seq), a signal was only accepted as a specific ChIRP-seq signal if it occurred with both probe pools and not in the control sample, so that signals due to non-specific probe binding during hybridisation can be excluded. Using this filtering procedure to eliminate non-specific signals, we identified a total of 3,388 ChIRP signals (*TUG1* chromatin interaction sites) in MDBK cells, 3,848 in MAC-T cells and 4,784 in bovine liver tissue in our ChIRP-seq datasets (Supplementary Table S2). These ChIRP signals indicate that *TUG1* occupancy is widely abundant at genomic sites on a genome-wide scale. This supports findings from Chu et al. 2011, who found that lncRNA binding sites are focal, specific and numerous. Most of these *TUG1* chromatin interaction sites could be clearly annotated in relation to mapped gene loci in the bovine genome using the HOMER software: 3,225 in MDBK cells, 3,587 in MAC-T cells and 3,977 in liver (Table 1, Supplementary Table S2). These numerous *TUG1* chromatin interaction sites suggest a multifaceted landscape in terms of its function and impact on multiple cellular processes.

Signal annotation	**MDBK cells**	**MAC-T cells**	**Liver**
Intergenic	1,439	1,493	1,906
Intronic_protein coding	1,493	1,493	1,520
Exonic_protein coding	28	207	178
Promoter-TSS_protein coding	59	126	130
TTS_protein coding	52	77	47
Long noncoding RNA	131	156	153
Short noncoding RNA	9	16	15
Other	14	19	28
Total	3,225	3,587	3,977

Table 1. Number of annotated chromatin interaction sites of TUG1 lncRNA detected genome-wide by ChIRP-seq analysis in bovine cell lines (MDBK and MAC-T) and liver tissue

Peak annotation using HOMER analysis [33] based on the ARS-UCD1.2 assembly (NCBI annotation release 106). TSS: transcription start sites of protein coding genes, TTS: transcription termination sites of protein coding genes. The detailed annotated ChIRP-seq datasets are provided in Supplementary Table S2.

As shown in Figure 3, the *TUG1* chromatin interaction sites are located in different genomic regions and may bind to different transcript sites. However, the distribution pattern appears relatively consistent across the analysed cell types and liver tissue. The majority of ChIRP signals was deciphered in intergenic regions (more than 40 % in all samples analysed, MDBK cells, MAC-T cells and liver) and intronic regions of protein coding genes (38-45 %). It is worth noting that about half of the ChIRP signals obtained in each sample (52 %, 53 % and 47 % in MDBK cells, MAC-T cells and liver, respectively) were annotated within or very close to protein coding gene loci. We found that *TUG1* chromatin interaction sites are predominantly localised in introns, but also in exons, transcription start or termination sites, suggesting a potential regulatory function in gene activity. Considering the low proportion of protein-coding gene sequence in the total bovine genome sequence (about 25 % including exons and introns), this predominant gene-associated distribution pattern of *TUG1* chromatin interaction sites is a high percentage compared to that in the large proportion of intergenic sequences in the genome.

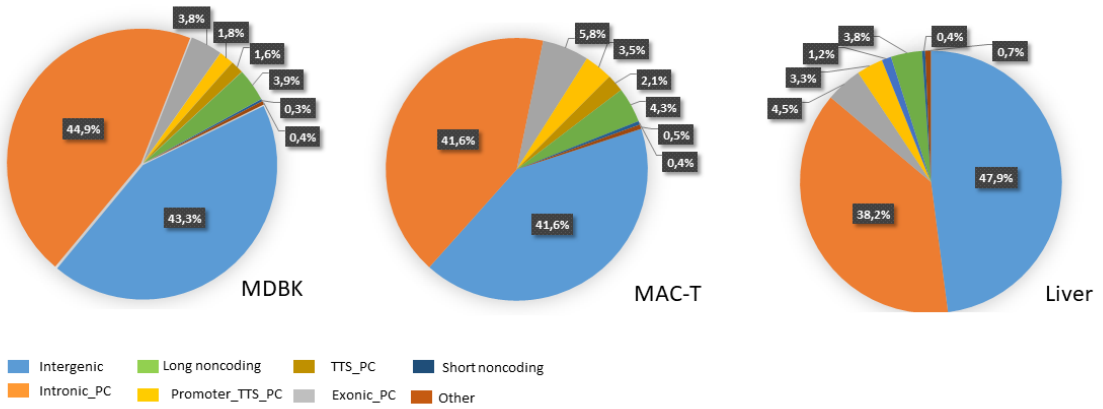

Figure 3. Classification and distribution of annotated *TUG1* chromatin interaction sites (%) in the bovine genome

Annotation using HOMER based analysis on the *Bos taurus* ARS-UCD1.2 assembly (NCBI annotation, release 106). TSS: transcription start sites of protein coding genes, TTS: transcription termination sites of protein coding genes, long noncoding: long noncoding RNA, short noncoding: short noncoding RNA. The detailed annotated ChIRP-seq datasets are provided in Supplementary Table S2.

The analysis of the chromosomal distribution of *TUG1* chromatin interaction sites clearly shows that they are distributed across all chromosomes of the bovine genome (Figure 4). It is striking that chromosomes 18 and 19 have the highest number of chromatin interaction sites in MDBK and MAC-T cells as well as in liver tissue. It is known that these chromosomes are very rich in protein coding genes compared to other bovine chromosomes [36]. There seems to be a correlation of *TUG1* chromatin interaction sites with gene density. Obviously, there is a strong overrepresentation of *TUG1* chromatin interaction sites in or near genomic regions containing protein-coding genes, as already has been

concluded from the classification analysis of chromatin interaction sites at a genome-wide scale (Figure 3).

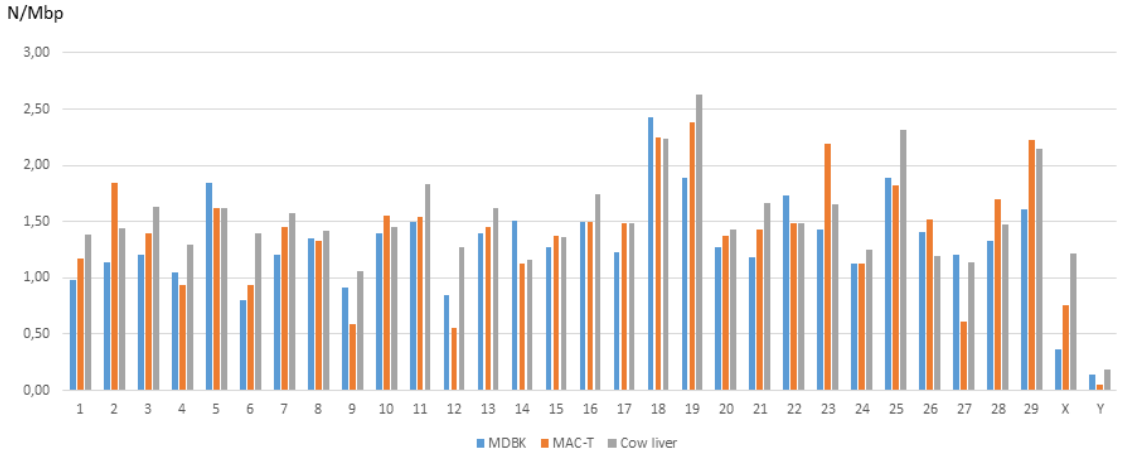

Figure 4. Chromosomal distribution of *TUG1* chromatin interaction sites detected in MDBK and MAC-T cells and bovine liver tissue

N/Mbp: Normalization of the number of signals was performed according to chromosomal length.

3.3. ChIRP-seq analysis in MDBK and MACT cell and liver tissue samples

Inspection of ChIRP-seq datasets from MDBK and MAC-T cells and liver tissue revealed numerous concordant matches of *TUG1* chromatin interaction sites in the three sample types analysed, which is in agreement with the known ubiquitous expression of *TUG1* in many cell and tissue types and its conservation across species ([23, 24], https://www.ncbi.nlm.nih.gov/gene/55000/ortholog/?scope=9347).

Pairwise comparison of the genome-wide annotated *TUG1* chromatin interaction sites in the three ChIRP datasets revealed numerous common signals. While the intersection between the ChIRP signals in MDBK cells and those in liver tissue revealed 407 concordant sites, the intersection between signals in MAC-T cells and liver tissue showed only 157 matches (Supplementary Table S3). The overlap between the signals of both cell types resulted in 266 overlapping sites (Supplementary Table S3). Finally, the cross-sample intersection of annotated chromatin interaction sites across all three ChIRP-seq datasets resulted in a list of 97 sites present in all three datasets (Supplementary Table S4). As expected, most of these positionally matched genomic sites are located in or near protein-coding loci. These concordant *TUG1* chromatin interaction sites at DNA level in different bovine cell types suggest conserved regulatory processes mediated by *TUG1* lncRNA across different bovine cell types/tissue and indicate a functional molecular relationship. Genomic regions and subsequent putative regulatory processes modulated by *TUG1*–chromatin interaction are broadly diversified. A few examples are selected from this list of conserved *TUG1* chromatin interaction sites shown in Figures 5 – 8 and discussed further.

Inspection of *TUG1* ChIRP signals with respect to exonic regions of protein coding genes highlighted a concordant chromatin interaction site in a genome region overlapping the last exon of *PIGT* (Figure 5) and the transcription termination site of *SPINT3*. The protein PIGT is as a component of the glycosylphosphatidylinositol transamidase complex essential for transfer of glycosylphosphatidylinositol to proteins [37], whereas SPINT3 is a serine peptidase inhibitor [38]. Another concordant *TUG1* chromatin interaction site (Figure 6) was observed in the first exon of *FNIP2*, encoding a protein that binds to the tumor suppressor folliculin and to AMPK-activated protein kinase and may play a role in cellular metabolism and nutrient sensing [39]. A further conserved exonic *TUG1* chromatin interaction site was observed in *DNAJC13* (Figure 5) that encodes a protein with co-chaperone function for heat-shock proteins and controls trafficking of multiple GPCRs through the endosomal/lysosomal pathway [40].

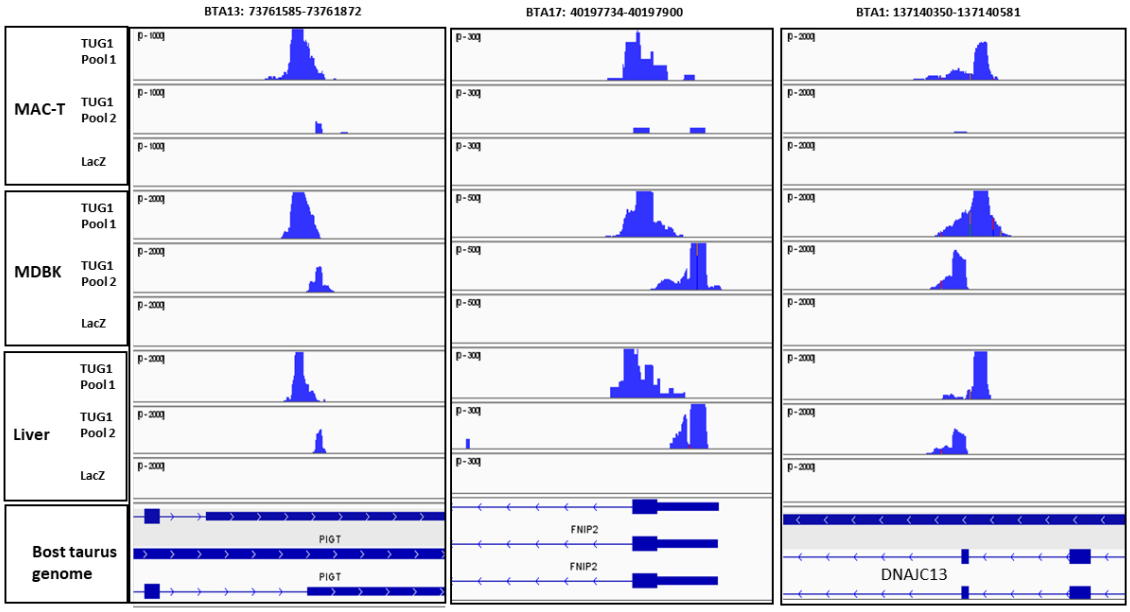

Figure 5. Concordant exonic chromatin interaction sites of *TUG1* lncRNA detected genome-wide by ChIRP-seq analysis across bovine cell lines MDBK and MAC-T and liver tissue

Bos taurus genome: *Bos taurus* NCBI genome annotation, release 106; TUG1 Pool 1 and Pool 2: *TUG1* ChIRP probe pools 1 and 2; LacZ: LacZ probe Pool (control).

Examples with prominent concordant *TUG1* chromatin interaction signals at intronic genomic sites (Figure 6) are *ZNF341* encoding a protein that acts as a transcriptional activator for cytokine-mediated *STAT3* expression involved in regulation of immune homeostasis [41], and *CNTN4*, whose protein product is involved in neuronal network formation and plasticity [42], as well as *ABHD12B* encoding an endocarabinoid lipase that

participates in lipid metabolism by catalyzing 2-arachidonoyl glycerol [43].

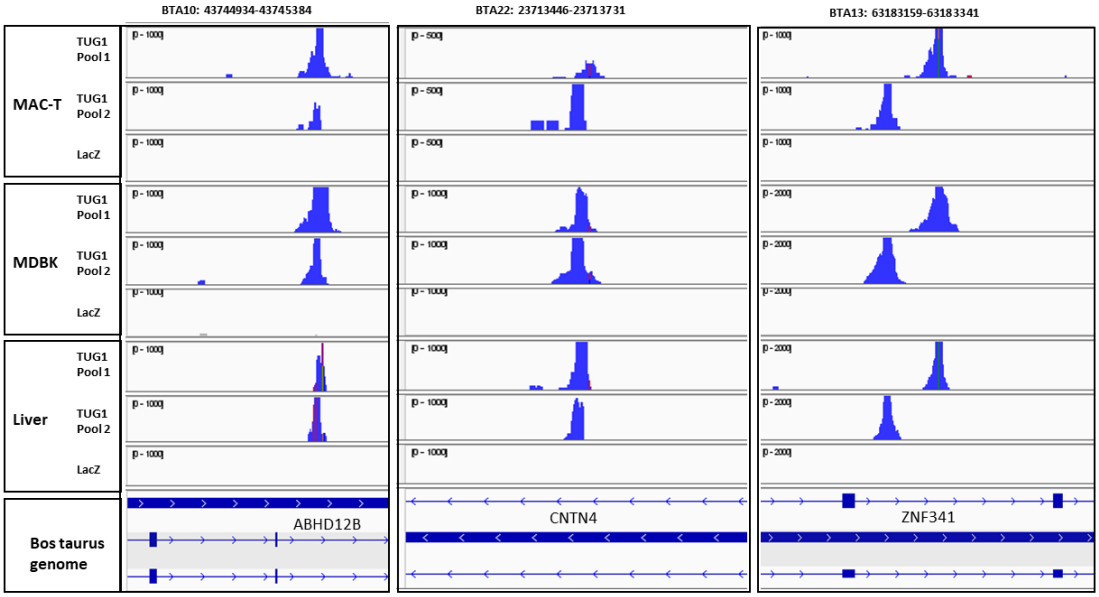

Figure 6. Concordant intronic chromatin interaction sites of *TUG1* lncRNA detected genome-wide by ChIRP-seq analysis across bovine cell lines MDBK and MAC-T and liver tissue

Bos taurus genome: *Bos taurus* NCBI genome annotation, release 106; TUG1 Pool 1 and Pool 2: *TUG1* ChIRP probe pools 1 and 2; LacZ: LacZ probe Pool (control).

In terms of promoter regions, exemplary concordant *TUG1* chromatin interaction sites were found near to the transcription start sites of *SNX24* and *DEGS1* (Figure 7). The protein SNX24 is predicted to enable phosphatidylinositol phosphate binding activity and to be involved in protein transport at several stages of intracellular trafficking [44]. *DEGS1* encodes a membrane fatty acid desaturase, which catalyzes the last step in the ceramide synthesis pathway and is essential for mitochondria-associated membrane integrity [45].

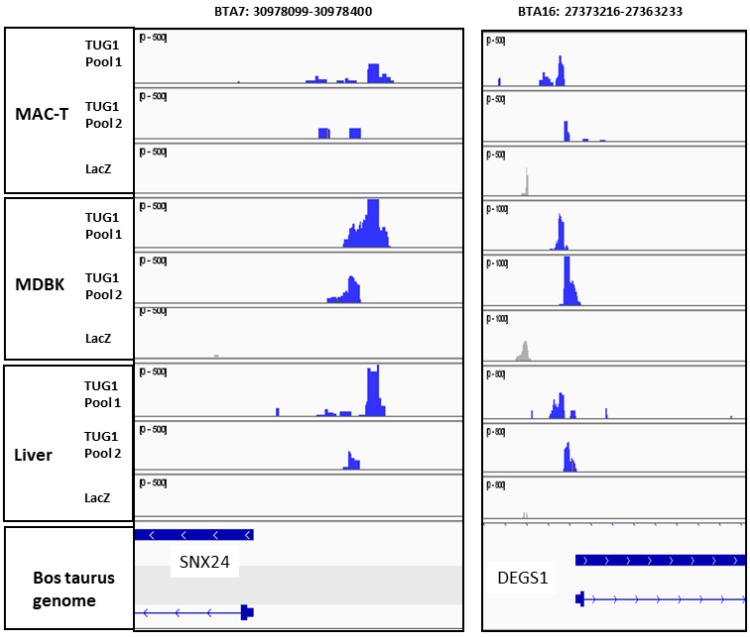

Figure 7. Concordant chromatin interaction sites of *TUG1* lncRNA detected genome-wide in promoter regions by ChIRP-seq analysis across bovine cell lines MDBK and MAC-T and liver tissue

Bos taurus genome: Bos taurus NCBI genome annotation, release 106; TUG1 Pool 1 and Pool 2: *TUG1* ChIRP probe pools 1 and 2; LacZ: LacZ probe Pool (control).

When looking for reports on the interaction partners of *TUG1* at transcript, protein or genomic level, we found that a *TUG1*-binding element was identified proximal to the *PPARGC1A* gene, encoding the master transcription regulator of mitochondrial biogenesis, via ChIRP-seq in mouse podocyte chromatin [26]. In addition, previous studies have found a functional *TUG1-PPARGC1A* association at the transcriptional level in mouse fat and human muscle cells [28, 46]. In the liver dataset of our study, a promoter-associated *TUG1* chromatin interaction site in the liver was found near *PPARGC1A* in a similar genomic location as has been reported in the mouse podocyte study (Figure 8). Although this ChIRP signal was not one of the most highly enriched in the liver, the *PPARGC1A* promoter region attracts particular interest as a potential *TUG1* interacting partner.

Another study reported binding between *TUG1* and *CDC27* that was validated by RNA immunoprecipitation [47]. The authors showed that *TUG1* knockdown resulted in down-regulation of the CDC27 locus at mRNA and protein level in rat alveolar epithelial cells. Furthermore, down-regulation of lncRNA *TUG1* mitigated pulmonary fibrosis by inhibiting CDC27 expression and subsequent inactivation of the PI3K/Akt/mTOR pathway with impact on inflammatory response and autophagy. In our study we detected clear signals in *CDC27* in the liver dataset and not in the two bovine cell types (Figure 8). The encoded CDC27 protein is a component of the anaphase promoting complex/cyclosome that controls the cell cycle.

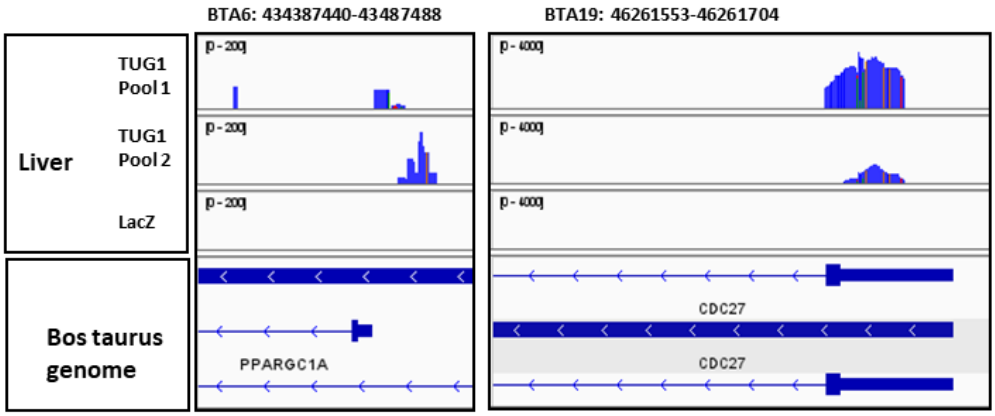

Figure 8. Specific chromatin interaction sites of *TUG1* lncRNA detected by ChIRP-seq analysis in liver tissue

Bos taurus genome: *Bos taurus* NCBI genome annotation, release 106; TUG1 Pool 1 and Pool 2: *TUG1* ChIRP probe pools 1 and 2; LacZ: LacZ probe Pool (control).

3.4. Functional annotation of TUG1 chromatin interaction sites in the liver and MAC-T and MDBK cells

Biological interpretation of *TUG1* chromatin interaction sites in the liver via their functional annotation analysis using DAVID 6.8 (https://david.ncifcrf.gov/summary.jsp) was performed with a list containing 2,271 gene loci based on detailed annotation information retrieved from the genome-wide HOMER annotation analysis of ChIRP signals (Supplementary Table S2). The results revealed a total of 58 significantly enriched (FDR < 0.05) gene ontology (GO) terms and biological KEGG pathways (Supplementary Table S5). The functional annotation chart contained 21 different KEGG pathways, comprising a broad range of cellular and metabolic processes. The most significantly enriched pathways are displayed in Figure 9. Axon guidance, Parathyroid hormone synthesis, secretion and action as well as Calcium signalling are at the top. Ranking the enrichment score of pathways associated with *TUG1*-chromatin interaction sites in the GO term categories showed that of the 13 Molecular function terms, ATP binding, Metal ion binding and Calcium ion binding are the most enriched. The *TUG1* chromatin interaction sites showed to be associated with the GO term RNA polymerase II core promoter proximal region sequence-specific DNA binding. This indicates interaction of *TUG1* with a transcription factor recognition sequence in cis location to a transcription start site of a gene that is transcribed by RNA polymerase II. It supports one of the molecular principles described for lncRNA-guided regulatory mechanisms of transcription [7].

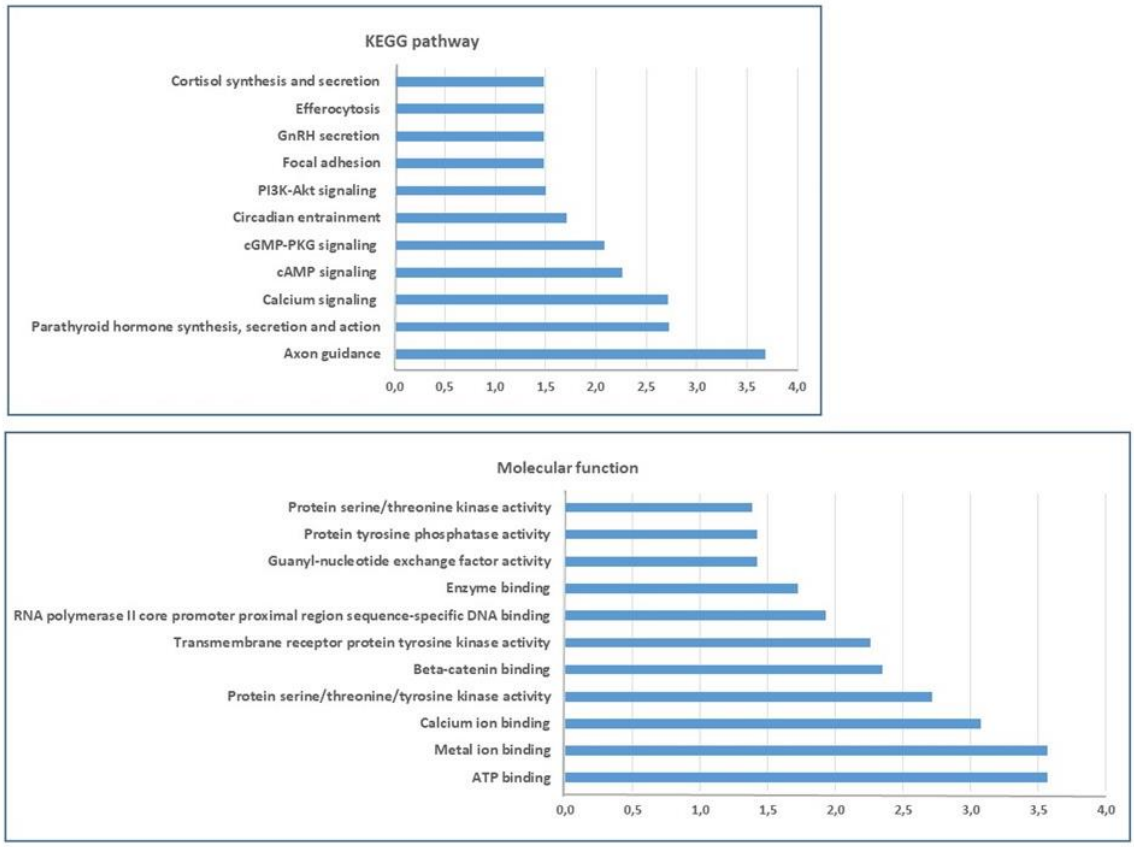

Figure 9. Functional annotation of *TUG1* chromatin interaction sites in the liver
The functional annotation analysis was performed using the DAVID tool, version 6.8, which is based on the comprehensive DAVID knowledge database (FDR < 0.05). X-axis: -log10FDR.

Functional annotation of *TUG1* chromatin interaction sites in MAC-T cells from the list of 2,191 gene loci annotated adjacent to ChIRP signals using HOMERr (Supplementary Table S2) detected a total of 48 significantly enriched (FDR < 0.05) GO terms and biological KEGG pathways (Supplementary Table 5). The functional annotation chart table included 11 KEGG biological pathways, some of which were also highlighted in the functional annotation of the liver tissue, such as Axon guidance, Calcium signalling, PI3K-Akt signalling and Focal adhesion as being highest significantly enriched pathways (Figure 10). However, distinct pathways enriched in MAC-T cells were the Rap1 signalling and the MAPK pathways besides cancer-associated pathways. The connection of *TUG1* with development and progression of various types of cancer has been widely reported. With regard to the 37 GO term categories found, enrichment of *TUG1*-chromatin interaction sites displayed association with the Molecular function terms ATP binding, Metal ion binding and Calcium ion binding that were also observed in the liver. In addition, Actin binding, Cadherin binding and Chromatin DNA binding were enriched (Figure 10). The latter GO term points to a regulatory function of *TUG1* lncRNA in gene expression via chromatin DNA interaction and binding of transcription factors.

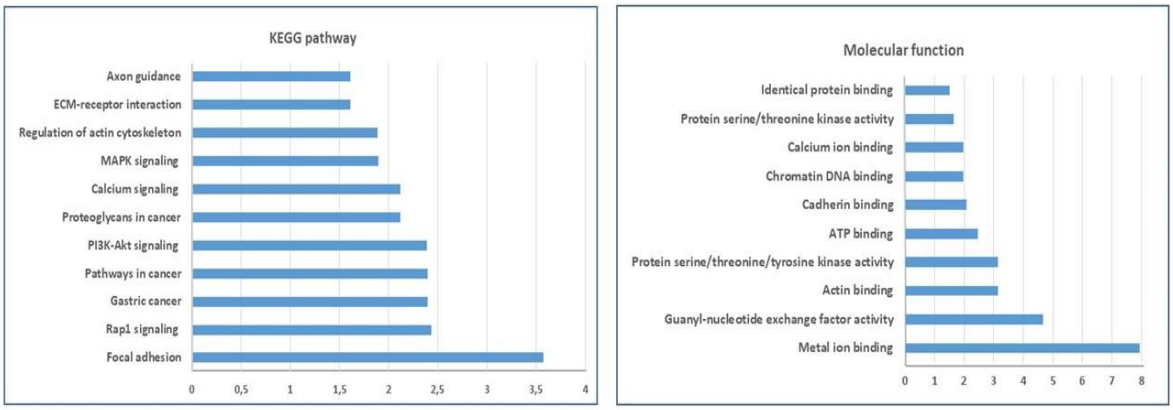

Figure 10. Functional annotation of *TUG1* chromatin interaction sites in MAC-T cells
The functional annotation analysis was performed using the DAVID tool, version 6.8, which is based on the comprehensive DAVID knowledge database (FDR < 0.05). X-axis: -log10FDR.

Functional annotation of the *TUG1* interactome sites in MDBK cells based on 2,093 gene loci annotated near ChIRP signals (Supplementary Table S2) revealed a total of 71 significantly enriched (FDR < 0.05) biological categories, comprising 49 GO terms and 22 biological KEGG pathways (Supplementary Table 5). As can be seen from the Figure 11, some of the biological categories, which were observed in the liver and MAC-T cells, are also evident for MDBK, such as KEGG pathways Calcium-, Rap1-, MAPK-, and cAMP signalling, Focal adhesion and cortisol synthesis as well as the GO terms Metal ion, Cadherin, ATP and Calcium ion binding and Protein serine/threonine/tyrosine kinase activity. Unique for MDBK cells are molecular functions, such as Wnt-protein binding and GTPase activator activity (Figure 11).

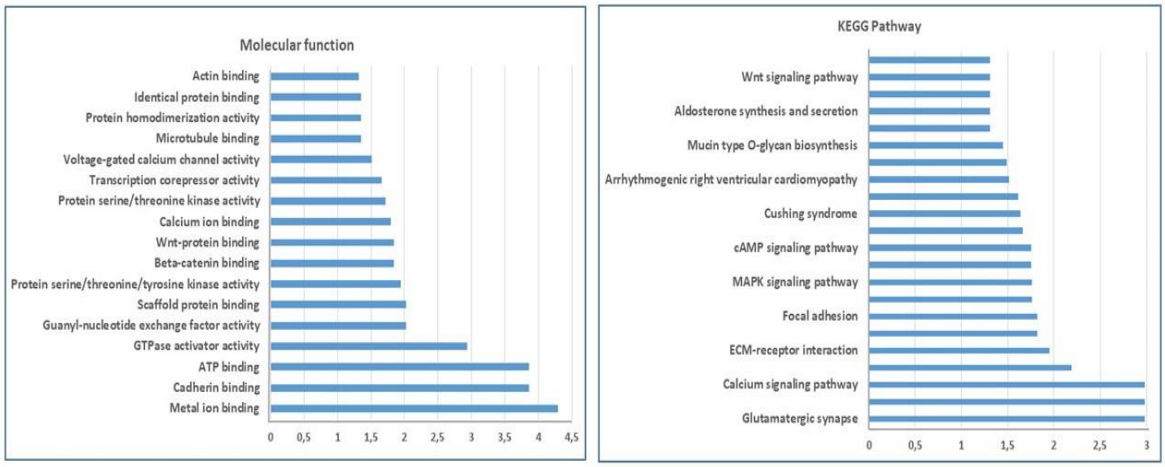

Figure 11. Functional annotation of *TUG1* chromatin interaction sites in MDBK cells
The functional annotation analysis was performed using the DAVID tool, version 6.8, which is based on the comprehensive DAVID knowledge database (FDR < 0.05). X-axis: -log10FDR.

3.5. Identification of consensus interaction motifs

Chu et al. 2011 showed that genome-scale collections of lncRNA chromatin interaction sites can be used to discover the underlying DNA sequence binding motifs. To search for potential DNA binding motifs of *TUG1* in our ChIRP-seq datasets, only peaks detected with ChIRP pool 1 and pool 2 were included in the sample-specific peak lists for each of the three samples (liver, MDBK and MAC-T cells). Then, we established a final dataset with sequences from peaks that were shared between liver and at least one of the two cell lines. This provided one final conservative consensus sequence list as input for the screening for enrichment of DNA binding motifs. Merging of ChIRP signals that were shared either in MDBK and liver or by MAC-T cells and liver revealed a list containing a total of 596 overlapping *TUG1* chromatin interaction sites (Supplementary Table S6). Motif analysis of sequences at these interaction sites using the HOMER software did not identify any sequence motifs for known binding or transcription factors. However, using the option of novel interaction motif search implemented in the HOMER software revealed eight potential motifs with a significant enrichment against the entire bovine genome sequence as background. Of them, as the best guess, the novel motif CATGCCTCAG exhibited the strongest enrichment signal ($p < 1E-29$) for *TUG1* interaction. This sequence motif only weakly resembles other known transcription factors, e.g. TFAP2A (Figure 12). Other potential DNA interaction motifs are YTGCCTGCCT ($p < 1E-18$) and VATAGGGWCC ($p < 1E-17$), which all have to be confirmed in further studies.

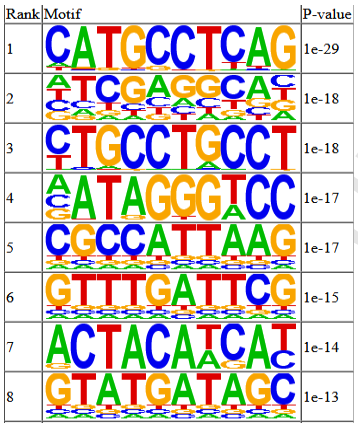

Figure 12. Graphical depiction of predicted novel *TUG1* interaction motifs for potential binding sites of *TUG1* associated regulatory factors in the genome of MDBK and MAC-T cells

The motif analysis was performed using the HOMER software (version v4.11)

3.6. Overlap of TUG1 chromatin interaction sites identified by ChIRP-seq with chromatin accessibility sites analysed with ATAC-seq

LncRNAs are known to exhibit regulatory functions in histone modification and thus, are involved in chromatin conformation itself via histone-modifying complexes [48].

Furthermore, lncRNAs interact in accessible chromatin conformation as prerequisite for the regulation of target genes through the binding of transcription factors. In our study, we wanted to know to what extent the *TUG1* chromatin interaction sites identified by ChIRP-seq correspond to open chromatin positions determined by ATAC-seq. The ATAC-seq technique assesses the overall accessibility of chromatin throughout the genome and identifies potential binding sites for regulatory DNA-associated proteins, eg. transcription factors or directly lncRNA-DNA interaction sites in open chromatin regions. In consequence, we hypothesized that *TUG1* binding to genomic sites identified by ChIRP-seq might simultaneously highlight coincident genomic regions marked by open chromatin as identified by ATAC-seq, should highlight potential regions with direct regulatory function in gene expression. Intersection of our ChIRP-seq datasets obtained for MAC-T and MDBK cells and ATAC-seq datasets available from the European BovReg project for the same cell lines revealed a total of 455 overlapping regions for MAC-T cells and 335 for MDBK cells (Supplementary Table S7) and thus mapped active *TUG1* chromatin interaction sites in regions of accessible chromatin at bovine genome scale. Integration of ChIRP-seq, ATAC-seq datasets is shown for examples of overlapping signal profiles in Figure 13.

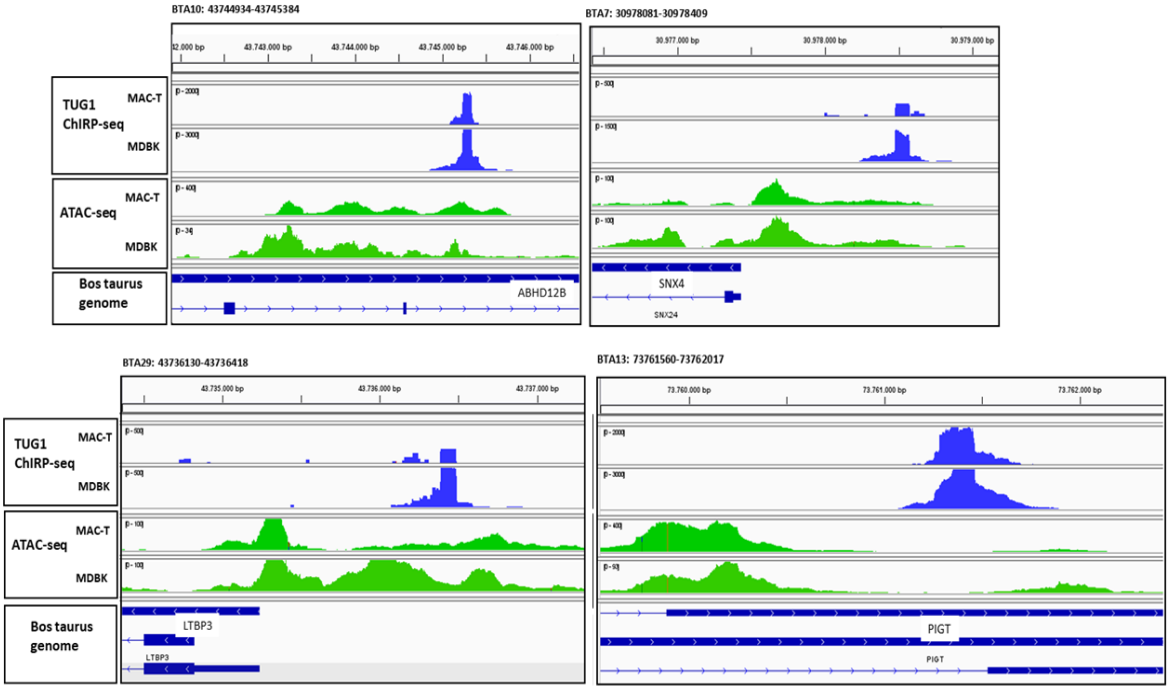

Figure 13. Co-localisation of *TUG1* chromatin interaction and chromatin accessibility sites in MDBK and MAC-T cells

Blue: *TUG1* ChIRP-seq signals in MAC-T and MDBK cells (Probe pool 1). Green: ATAC-seq signals in MAC-T and MDBK cells. Bos taurus genome: *Bos taurus* NCBI genome annotation, release 106.

With regard to the total number of intersecting sites of *TUG1* chromatin interaction and chromatin accessibility, it is again observed that these overlapping sites are predominantly localised in regions in or near genes, 66 % in MAC-T cells and 60 % in MDBK cells. In contrast, 25 % (MAC-T cells) or 33 % (MDBK cells) of overlapping TUG1 chromatin interaction and open chromatin sites were found in intergenic regions. In relation to the total number of *TUG1* chromatin interaction sites (3,587 for MAC-T and 3,225 for MDBK cells), the overlapping sites to chromatin accessibility correspond to 12.7 % for MAC-T cells and 10.4% for MDBK cells, respectively. The concordant *TUG1* chromatin interaction sites and chromatin accessibility in MDBK and MAC-T cells, respectively, pinpoint, which of the identified genome-wide *TUG1* chromatin interaction sites are potentially functionally significant. They point towards sites of regulation of active gene expression and that *TUG1* may be involved in the corresponding molecular and biological processes associated with the respective gene loci in MAC-T and MDBK cells. The limited number of overlaps between ATAC-seq and ChIRP-seq peaks highlights *TUG1* interaction sites outside of open chromatin, however, it is in line with the assumption by Zhou et al. [25] that *TUG1* seems to modulate many cellular processes via silencing target genes.

3. Conclusions

Given the potential biological significance of lncRNAs as functional elements of the genome, it can be expected that they are involved at crucial steps in the gene regulation of cellular and cell-type specific processes and the expression/appearance of certain phenotypic traits and complex characteristics. A comprehensive study of genome-wide lncRNA-chromatin interactions can therefore not only help to elucidate their functional role, but also provide important clues as to how lncRNAs contribute to the expression of specific phenotypes or the pathogenesis of complex diseases through epigenetic modifications [49]. Our ChIRP-seq study focused on the *TUG1* lncRNA, one of the most conserved lncRNAs between species and ubiquitously expressed in cell and tissue types, and has experimentally uncovered *TUG1* lncRNA chromatin interaction sites in two widely used bovine cell lines and bovine liver tissue for the first time and thus, contributes to the functional annotation of non-coding elements in the bovine genome. These results will form the basis for gaining deeper insights into the *cis/trans* modes of *TUG1* regulatory networks in forthcoming studies. The data and results obtained in our study will help to discover and analyze the lncRNA-phenome relationships and open novel avenues in future investigations. Furthermore, the established experimental pipeline can be applied to identify the interactomes of further lncRNAs in other cell lines and tissues of interest.

CRediT authorship contribution statement

Rosemarie Weikard: Writing – review & editing, Methodology, Investigation, Formal analysis, Conceptualization, Funding acquisition, Project administration. **Christa Kuehn:** Writing – review & editing, Formal analysis, Methodology, Validation, Conceptualization, Funding acquisition, Resources, Project administration. **Doreen Becker:** Investigation,

Review & editing. **Raghu Bhushan:** Investigation, Review & editing. **Frieder Hadlich:** Software, Formal analysis, Review & editing. **Carole Charlier:** Methodology, Investigation, Review & editing. **Gabriel Costa Monteiro Moreira**: Methodology, Investigation, Review & editing.

Declaration of competing interest

The authors declare that there are no conflicts of interests.

Funding
This study was funded by the German Research Foundation (DFG grant numbers: KU 771/8-1 and WE 1786/5-1) and received funding from the European Union's Horizon 2020 research and innovation program for the BovReg project (grant number 815668).

Data availability
The data generated and analyzed in this study are included as figures and tables or included in supplementary tables as indicated appendix A. The liver RNA-seq data used in this study was already used in a previous study (Nolte et al. 2019) stored in the Functional Annotation of Animal Genomes (FAANG) database (https://data.faang.org/dataset) under project number PRJEB34570. The ATAC-seq datasets for MDBK and MAC-T cells are available from the FAANG data portal (https://data.faang.org/home) under the European Horizon 2020 project BovReg (Grant agreement ID 815668) and deposited under project number PRJEB51163. The ChIRP-seq datasets for MDBK, MAC-T and liver (BioSamples SAMEA119609839, SAMEA119609840, SAMEA119609841) are publicly available at ENA under project number PRJEB 96303.

Ethics Statement
For collecting liver samples used to generate RNA-seq data, the project (Nolte et al. 2019) was reviewed and approved by Animal care and experimental procedures followed the guidelines of the German Law of Animal Protection. The protocols were approved by the Animal Protection Board of the Institute for Farm Animal Biology (FBN) as well as by the Animal Care Committee of the State Mecklenburg-Western Pomerania, Germany (State Office for Agriculture, Food Safety and Fishery; LALLF M-V/ Rostock, Germany, TSD/7221.3-2.1-010/03).

Appendix A: Supplementary data
The supplementary data comprises Supplementary Method File 1 (docx), Supplementary Figure 1 (png) and Supplementary Tables 1-7 (xlsx). **Method File 1**: ChIRP-seq method file. **Supplementary Figure 1.** Proof for signal enrichment using *TUG1* specific anti-sense tiling oligonucleotides in a ChIRP experiment with MDBK and MAC-T cells and liver tissue. **Table S1:** Anti-sense tiling oligonucleotides designed for ChIRP-seq and primer sequences used for PCR amplification and resequencing. **Table S2:** ChIRP-seq data sets for bovine liver, MDBK and MAC-T cells. **Table S3:** Intersection of *TUG1* chromatin interaction sites in bovine cells and liver. **Table S4:** Concordant *TUG1* chromatin interaction sites across bovine cells and liver. **Table S5:** Functional annotation of *TUG1* chromatin interaction sites. **Table S6:** Concordant dataset of liver and cells for motif analysis. **Table S7:** Overlap of ChIRP-seq and ATAC-seq in bovine cells.

Acknowledgements
The authors thank Simone Wöhl, Manuela Lötzke, Sebastian Gaedecke (FBN Dummerstorf) for their excellent technical work in the lab and Sébastien Dupont as well as the GIGA Genomics facility at ULiège.

Declaration of generative AI and AI assisted technologies

The authors declare that they did not use generative AI and AI-assisted technologies in the writing process.

References

[1] G. Housman, I. Ulitsky, Methods for distinguishing between protein-coding and long noncoding RNAs and the elusive biological purpose of translation of long noncoding RNAs, Biochim. Biophys. Acta - Gene Regulatory Mechanisms, 1859 (2016) 31-40.

[2] Y. Xiao, Y. Ren, W. Hu, A.R. Paliouras, W. Zhang, L. Zhong, K. Yang, L. Su, P. Wang, L. Y., M. Ma, L. Shi, Long non-coding RNA-encoded micropeptides: functions, mechanisms and implications, Cell Death Discovery, 10 (2024) 450.

[3] A.L. Tornesello, A. Cerasuolo, N. Starita, S. Amiranda, T.P. Cimmino, P. Bonelli, F.M. Tuccillo, F.M. Buonaguro, L. Buonaguro, M.L. Tornesello, Emerging role of endogenous peptides encoded by non-coding RNAs in cancer biology, Noncoding RNA Research, 10 (2024) 231-241.

[4] L. Statello, C.J. Guo, L.L. Chen, M. Huarte, Gene regulation by long non-coding RNAs and its biological functions, Nature Reviews Molecular Cell Biology 22 (2021) 96-118.

[5] J.L. Rinn, H.Y. Chang, Genome regulation by long noncoding RNAs, Annual Review of Biochemistry, 81 (2012) 145-166.

[6] A.E. Kornienko, P.M. Guenzl, D.P. Barlow , F.M. Pauler, Gene regulation by the act of long non-coding RNA transcription, BMC Biology 11 (2013) 59.

[7] P.P. Mattick, P. Carninci, S. Carpenter, H.Y. Chang, L.L. Chen, R. Chen, C. Dean, M.E. Dinger, K.A. Fitzgerald, T.R. Gingeras, M. Guttman, T. Hirose, M. Huarte, R. Johnson, C. Kanduri, P. Kapranov, J.B. Lawrence, L. J.T., J.T. Mendell, T. Mercer, K. Moore, S. Nakagawa, J.L. Rinn, D. Spector, I. Ulitsky, Y. Wan, J.E. Wilusz, M. Wu, Long non-coding RNAs: definitions, functions, challenges and recommendations, Nature Reviews Molecular Cell Biology, 24 (2023) 430-447.

[8] K.C. Wang, H.Y. Chang, Molecular mechanisms of long noncoding RNAs. , Molecular Cell 43 (2011) 904–914

[9] L. Ulitsky, D.P. Bartel, LincRNAs: genomics, evolution, and mechanisms, Cell, 154 (2013) 26-46.

[10] J.W. Kornfeld, J.C. Bruening, Regulation of metabolism by long, non-coding RNAs, Frontiers in Genetics 5(2014) 5.

[11] X.Y. Zhao, J.D. Lin, Long noncoding RNAs: a new regulatory code in metabolic control, Trends Biochemical Sciences 40 (2015) 585-596.

[12] E.K. Robinson, S. Covarrubias, S. Carpenter, The how and why of lncRNA function: An innate immune perspective, Biochimica Biophysica Acta Gene Regulatory Mechanisms 1863 (2020) 194419.

[13] K. Muret, C. Désert, L. Lagoutte, M. Boutin, F. Gondret, T. Zerjal, S. Lagarrigue, Long noncoding RNAs in lipid metabolism: literature review and conservation analysis across species, BMC Genomics, 20 (2019) 882.

[14] A.M. Khalil, M. Guttman, M. Huarte, M. Garber, A. Raj, D. Rivea Morales, K. Thomas, A. Presser, B.E. Bernstein, A. van Oudenaarden, A. Regev, E.S. Lander, J.L. Rinn, Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression, Proceedings of the National Academy of Sciences of the United States of America, 106 (2009) 11667–11672

[15] T. Nojima, N.J. Proudfoot, Mechanisms of lncRNA biogenesis as revealed by nascent transcriptomics, Nature Reviews Molecular Cell Biology, 23 (2022) 389–406.

[16] R. Weikard, W. Demasius, C. Kuehn, Mining long noncoding RNA in livestock, Animal Genetics, 48 (2017) 3-18.

[17] B. Kosinska-Selbi, M. Mielczarek, J. Szyda, Review: Long non-coding RNA in livestock, Animal, 14 (2020) 1-11.

[18] S. Lagarrigue, M. Lorthiois, F. Degalez, D. Gilot, T. Derrien, LncRNAs in domesticated animals: from dog to livestock species, Mammalian Genome, 33 (2022) 248-270.

[19] W. Nolte, R. Weikard, R.M. Brunner, E. Albrecht, H.M. Hammon, A. Reverter, C. Kühn, Identification and annotation of potential function of regulatory antisense long non-coding RNAs related to feed efficiency in bos taurus bulls, International Journal of Molecular Sciences, 21 (2020) 3292.

[20] R. Weikard, F. Hadlich, H.M. Hammon, D. Frieten, C. Gerbert, C. Koch, G. Dusel, C. Kuehn, Long noncoding RNAs are associated with metabolic and cellular processes in the jejunum mucosa of pre-weaning calves in response to different diets, Oncotarget, 9 (2018) 21052-21069.

[21] P.A. Alexandre, A. Reverter, R.B. Berezin, L.R. Porto-Neto, G. Ribeiro, M.H.A. Santana, J.B.S. Ferraz, H. Fukumasu, Exploring the Regulatory Potential of Long Non-Coding RNA in Feed Efficiency of Indicine Cattle, Genes, 11 (2020) 22.

[22] X. Yang, ., X. Ma, C. Mei, L. Zan, A genome-wide landscape of mRNAs, lncRNAs, circRNAs and miRNAs during intramuscular adipogenesis in cattle, BMC Genomics, 23 (2022) 691.

[23] T.L. Young, T. Matsuda, C.L. Cepko, The noncoding RNA taurine upregulated gene 1 is required for differentiation of the murine retina, Current Biology, 15 (2005) 501–512.

[24] J.P. Lewandowski, G. Dumbović, A.R. Watson, T. Hwang, E. Jacobs-Palmer, N. Chang, h.C. Muc, K.M. Turner, C. Kirby, N.D. Rubinstein, A.F. Groff, s.S.C. Liapi, C. Gerhardinger, r.A. Beste, P.P. Pandolfi, J.G. Clohessy, H.E. Hoekstra, M. Sauvageau, J.L. Rinn, The Tug1 lncRNA locus is essential for male fertility, Genome Biology, 21 (2020) 237.

[25] H. Zhou, L. Sun, F. Wan, Molecular mechanisms of TUG1 in the proliferation, apoptosis, migration and invasion of cancer cells, Oncology Letters, 18 (2019) 4393-4402.

[26] J. Long, S.S. Badal, Z. Ye, Y. Wang, B.A. Ayanga, D.L. Galvan, N.H. Green, B.H. Chang, P.A. Overbeek, F.R. Danesh, Long noncoding RNA Tug1 regulates mitochondrial bioenergetics in diabetic nephropathy, Journal of Clinical Investigation, 126 (2016) 4205-4218.

[27] J. Long, D.L. Galvan, K. Mise, Y.S. Kanwar, L. Li, N. Poungavrin, P.A. Overbeek, B.H. Chang, F.R. Danesh, Role for carbohydrate response element-binding protein

(ChREBP) in high glucose-mediated repression of long noncoding RNA Tug1, Journal Biological Chemistry, 295 (2020) 15840-15852.

[28] Y. Zhang, Y. Ma, M. Gu, Y. Peng, lncRNA TUG1 promotes the brown remodeling of white adipose tissue by regulating miR 204 targeted SIRT1 in diabetic mice, International Journal Molecular Medicine, 46 (2020) 2225-2234,.

[29] C. Chu, K. Qu, F.L. Zhong, S.E. Artandi, H.Y. Chang, Genomic Maps of Long Noncoding RNA Occupancy Reveal Principles of RNA-Chromatin Interactions, Molecular Cell, 44 (2011) 667-678.

[30] C. Chu, Q.C. Zhang, S.T. da Rocha, R.A. Flynn, M. Bharadwaj, J.M. Calabrese, T. Magnuson, E. Heard, H.Y. Chang, Systematic discovery of Xist RNA binding proteins, Cell, 161 (2015) 404-416.

[31] H.T. Huynh, G. Robitaille, J.D. Turner, Establishment of bovine mammary epithelial cells (MAC-T): an in vitro model for bovine lactation, Experimental Cell Research, 197 (1991) 191-199.

[32] W. Nolte, R. Weikard, R.M. Brunner, E. Albrecht, H.M. Hammon, A. Reverter, C. Kuehn, Biological Network Approach for the Identification of Regulatory Long Non-Coding RNAs Associated With Metabolic Efficiency in Cattle, Frontiers in Genetics, 10 (2019) 19.

[33] S. Heinz, C. Benner, N. Spann, E. Bertolino, Y.C. Lin, P. Laslo, J.X. Cheng, C. Murre, H. Singh, C.K. Glass, Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities, Molecular Cell 38 (2010) 576-589.

[34] D.W. Huang, B.T. Sherman, R.A. Lempicki, Systematic and integrative analysis of large gene lists using DAVID Bioinformatics Resources, Nature Protocols 4(2009) 44-57.

[35] C. Yuan, L. Tang, T. Lopdell, V.A. Petrov, C. Oget-Ebrad, G.C.M. Moreira, J.L. Gualdrón Duarte, A. Sartelet, Z. Cheng, M. Salavati, D.C. Wathes, M.A. Crowe, GplusE Consortium, W. Coppieters, M. Littlejohn, C. Charlier, T. Druet, M. Georges, H. Takeda, An organism-wide ATAC-seq peak catalog for the bovine and its use to identify regulatory variants, Genome Research, 33 (2023) 1848.

[36] A. Kommadath, H. Nie, M.A. Groenen, M.F. te Pas, R.F. Veerkamp, M.A. Smits, Regional regulation of transcription in the bovine genome. , PLoS One, 6 (2011) e20413.

[37] D. Li, Structure and Function of the Glycosylphosphatidylinositol Transamidase, a Transmembrane Complex Catalyzing GPI Anchoring of Proteins, Subcellular Biochemistry, 104 (2024) 425-458.

[38] A. Clauss, M. Persson, H. Lilja, Å. Lundwall, Three genes expressing Kunitz domains in the epididymis are related to genes of WFDC-type protease inhibitors and semen coagulum proteins in spite of lacking similarity between their protein products, BMC Biochemistry 11 (2011) 1255.

[39] M. Baba, S.B. Hong, N. Sharma, M.B. Warren, N. M.L, A. Iwamatsu, D. Esposito, W.K. Gillette, R.F.r. Hopkins, J.L. Hartley, M. Furihata, S. Oishi, W. Zhen, T.R.J. Burke, W.M. Linehan, L.S. Schmidt, B. Zbar, Folliculin encoded by the BHD gene interacts with a binding protein, FNIP1, and AMPK, and is involved in AMPK and mTOR signaling, Proceedings of the National Academy of Sciences of the U S A., 103 (2006) 15552-15557.

[40] B. Novy, A. Dagunts, T. Weishaar, E.E. Holland, H. Adoff, E. Hutchinson, M. De Maria, M. Kampmann, N.G. Tsvetanova, B.T. Lobingier, An engineered trafficking biosensor reveals a role for DNAJC13 in DOR downregulation, Nature Chemical Biology, 21 (2025) 360-370.

[41] S. Frey-Jakobs, J.M. Hartberger, M. Fliegauf, C. Bossen, M.L. Wehmeyer, J.C. Neubauer, A. Bulashevska, M. Proiett, P. Fröbel, C. Nöltner, L. Yang, J. Rojas-Restrepo, N. Langer, S. Winzer, K.R. Engelhardt, C. Glocker, D. Pfeifer, A. Klein, A.A. Schäffer, I. Lagovsky, I. Lachover-Roth, V. Béziat, A. Puel, J.L. Casanova, B. Fleckenstein, S. Weidinger, S.S. Kilic, B.Z. Garty, A. Etzioni, B. Grimbacher, ZNF341 controls STAT3 expression and thereby immunocompetence, Science Immunology 3(2018).

[42] A. Oguro-Ando, R.A. Bamford, W. Sital, J.J. Sprengers, A. Zuko, J.M. Matser, H. Oppelaar, A. Sarabdjitsingh, M. Joëls, J.P.H. Burbach, M.J. Kas, Cntn4, a risk gene for neuropsychiatric disorders, modulates hippocampal synaptic plasticity and behavior, Translational Psychiatry, 11 (2021) 106.

[43] A. Joshi, M. Shaikh, S. Singh, A. Rajendran, A. Mhetre, S.S. Kamat, Biochemical characterization of the PHARC-associated serine hydrolase ABHD12 reveals its preference for very-long-chain lipids, Journal Biological Chemistry, 293 (2018) 16953-16963.

[44] A. Ravussin, A. Brech, S.A. Tooze, H. Stenmark, The phosphatidylinositol 3-phosphate-binding protein SNX4 controls ATG9A recycling and autophagy, Journal Cell Science 134 (2021) jcs250670.

[45] L. Planas-Serra, N. Launay, L. Goicoechea, B. Heron, C. Jou, N. Juliá-Palacios, M. Ruiz, S. Fourcade, C. Casasnovas, C. De La Torre, A. Gelot, M. Marsal, P. Loza-Alvarez, À. García-Cazorla, A. Fatemi, I. Ferrer, M. Portero-Otin, E. Area-Gómez, A. Pujol, Sphingolipid desaturase DEGS1 is essential for mitochondria-associated membrane integrity, Journal of Clinical Investigation, 133 (2023) e162957.

[46] A.J. Trewin, J. Silver, H.T. Dillon, P.A. Della Gatta, L. Parker, D.S. Hiam, Y.P. Lee, M. Richardson, G.D. Wadley, S. Lamon, Long non-coding RNA Tug1 modulates mitochondrial and myogenic responses to exercise in skeletal muscle, BMC Biology, 20 (2022) 164.

[47] F. Qi, Z.D. Lv, W.D. Huang, S.C. Wei, X.M. Liu, W.D. Song, LncRNA TUG1 promotes pulmonary fibrosis progression via up-regulating CDC27 and activating PI3K/Akt/mTOR pathway, Epigenetics, 18 (2023) 2195305.

[48] Y. Long, X. Wang, D.T. Youmans, T.R. Cech, How do lncRNAs regulate transcription?, Science Advances, 3 (2017) eaao2110.

[49] G. Zhang, Y. Lan, A. Xie, J. Shi, H. Zhao, I. Xu, S. Zhu, T. Luo, T. Zhao, Y. Xiao, X. Li, Comprehensive analysis of long noncoding RNA (lncRNA)-chromatin interactions reveals lncRNA functions dependent on binding diverse regulatory elements, Journal Biological Chemistry 294 (2019) 15613-15622.

CRediT authorship contribution statement

Rosemarie Weikard: Writing – review & editing, Methodology, Investigation, Formal analysis, Conceptualization, Funding acquisition, Project administration. **Christa Kuehn:** Writing – review & editing, Formal analysis, Methodology, Validation, Conceptualization, Funding acquisition, Resources, Project administration. **Doreen Becker:** Investigation, **R**eview & editing. **Raghu Bhushan:** Investigation, Review & editing. **Frieder Hadlich:** Software, Formal analysis, Review & editing. **Carole Charlier:** Methodology, Investigation, **R**eview & editing. **Gabriel Costa Monteiro Moreira**: Methodology, Investigation, **R**eview & editing.

Declaration of competing interest

The authors declare that there are no conflicts of interests.

Highlights

- Chromatin isolation by RNA precipitation (ChIRP-seq) in two cell lines and liver
- First description of *TUG1* lncRNA chromatin interaction sites in bovine
- *TUG1* chromatin interaction sites predominantly in or near protein-coding genes
- Numerous *TUG1* chromatin interaction sites concordant in cells and liver tissue
- *TUG1* seems to modulate many cellular processes via silencing target genes