



Molecular characterization, tissue expression and polymorphisms of buffalo *PPARGC1A* gene

Lihua Qiu^{1,*}, Xinyang Fan^{1,*}, Yongyun Zhang^{1,2}, Xiaohong Teng¹, and Yongwang Miao¹

¹Faculty of Animal Science and Technology, Yunnan Agricultural University, Kunming 650201, Yunnan, China

²Teaching Demonstration Center of the Basic Experiments of Agricultural Majors,

Yunnan Agricultural University, Kunming 650201, Yunnan, China

★ These authors contributed equally to this work.

Correspondence: Yongwang Miao (yongwangmiao1@126.com)

Received: 24 March 2020 - Revised: 29 May 2020 - Accepted: 24 June 2020 - Published: 22 July 2020

Abstract. PPARGC1A exerts important functions in activating many nuclear receptors and transcription factors that are related to energy balance. Previous studies have shown that PPARGC1A gene is associated with lactation traits of dairy cattle. However, the functional role of the buffalo PPARGC1A gene is still unknown. In this work, the complete coding sequence (CDS) of buffalo *PPARGC1A* was isolated and characterized for swamp and river buffalo. The CDS length of PPARGC1A for both types of buffalo was the same, which was composed of 2394 nucleotides and encoded a peptide composed of 797 amino acid residues. This protein belonged to a hydrophilic protein and contained one RRM_PPARGC1A domain (AA 674-764) without a signal peptide or a transmembrane domain. The differential expressions of this gene in 10 buffalo tissues in lactation and nonlactation displayed that the PPARGCIA was highly expressed in the muscle, heart, liver, brain and kidney of both non-lactating and lactating periods, but its expression was significantly different in the muscle, heart, liver, small intestine, mammary gland, rumen, spleen and lung between the two periods. Eight single nucleotide polymorphisms (SNPs) were found in buffalo, in which the c.778C >T, c.1257G>A and c.1311G>A were shared by two types of buffalo with similar allele frequencies, while the c.419C>T, c.759A>G, c.920C>A, c.926G>A and c.1509A>T were only observed in river buffalo. The SNP419, SNP920 and SNP926 were non-synonymous, which led to the amino acid changes of p.Ser140Phe, p.Pro307His and p.Arg309Lys. Seven nucleotide differential sites were identified in the PPARGC1A gene between buffalo and other Bovidae species. Phylogenetic analysis indicated that buffaloes were independently clustered into one branch, but they were closely related to the species of the Bos genus. The results indicate that buffalo PPARGC1A is an inducible transcriptional coactivator involved in regulating carbohydrate and fat metabolism. It can exert a functional role in a variety of buffalo tissues and may participate in milk fat synthesis and development in the mammary gland.

1 Introduction

Peroxisome proliferation-activated receptor γ (PPAR γ), as a main modulator of adipocyte differentiation, regulates the expression of genes related to fatty acid and glucose metabolism (Oberkofler et al., 2002). PPAR γ coactivator- 1α (PPARGC1A) interacts with PPAR γ , permitting this protein to interact with various transcription factors, and then plays a role in multiple biological processes (Esterbauer et al., 1999; Knutti et al., 2000). The *PPARGC1A* gene was first identified in the brown fat cDNA library of mice in 1998 (Puigserver et al., 1998). Some researches have shown that PPARGC1A is responsible for activating thermogenesis and oxidative metabolism of brown fat and muscle (Spiegelman et al., 2000). Adenoviral-mediated expression of *PPARGC1A* strongly activates the whole process of key gluconeogenic enzymes in hepatocytes, indicating that PPARGC1A is a pivotal modulator of gluconeogenesis in the liver (Yoon et al., 2001). Cattle *PPARGC1A* gene is located on chromosome 6, containing 13 exons, with a total complete coding sequence

(CDS) length of 2391 bp (Weikard et al., 2005). Some studies have shown that the *PPARGC1A* gene is closely related to milk protein (Pasandideh et al., 2015) and milk fat yield (Weikard et al., 2005; Chen et al., 2017) in cattle and goat. In addition, PPARGC1A is also a key regulator related to cattle intramuscular fat (Ramayo-Caldas et al., 2014).

Domestic buffalo can be divided into swamp buffalo and river buffalo. The former is mainly used for draft, with an annual milk production of 500-600 kg, while the latter is mainly used for milk production, with an annual milk production of about 2000 kg. The content of milk fat and protein in buffalo milk is significantly higher than that in the milk of dairy cattle (D'Ambrosio et al., 2008), which enables good processing characteristics for buffalo milk. Because the PPARGC1A gene has been proven to exert an important function in the expression regulation of genes related to fatty acid and glucose metabolism in some mammals, it can be used as a key candidate gene for lactation in dairy animals. In particular, the single nucleotide polymorphisms (SNPs) in this gene, which have a significant effect on milk production traits, can be used as a marker for assisted selection of buffalo lactation traits. But so far, there are few studies on the buffalo PPARGC1A gene. The purpose of this study is to isolate complete CDS of the buffalo PPARGC1A gene, to describe its molecular characteristics and multi-tissue differential expression in lactating and non-lactating stages, and to detect and characterize the SNPs in the CDS of this gene for two types of buffalo. This work will serve as a molecular basis to bringer further insight into the characteristics, functions and variation of the buffalo PPARGC1A gene.

2 Materials and methods

2.1 Animals and sampling

The heart, liver, spleen, lung, kidney, muscle, mammary tissue samples of adult healthy Binglangjiang buffaloes (river type, n = 3) and Dehong Buffaloes (swamp type, n = 3) were collected for gene isolation. After the buffalo had been slaughtered, each tissue sample was separated immediately, put into a freezing tube and stored in liquid nitrogen.

Eight Binglangjiang buffaloes (about four years old) – four in peak lactation (about 60 d postpartum) and four in dry-off period (about 60 d before parturition) – were selected for the collection of the samples for analysis of tissue differential expression. All buffalo sampled were managed in a similar fashion. After the buffalo were slaughtered, the tissue samples from the heart, liver, spleen, lung, kidney, mammary gland, small intestine, rumen, muscle and brain were immediately culled and stored in a refrigerator at -80 °C until RNA extraction.

Furthermore, the fresh blood samples were collected from 108 Binglangjiang buffaloes and 81 Dehong buffaloes at a local breeding farm for population variation detection. The buffaloes used for sample collection were all adult healthy buffaloes without direct blood relationship.

All procedures for sample collection were performed in accordance with the Guide for Animal Care and Use of Experimental Animals approved by Yunnan Provincial Experimental Animal Management Committee under contract no. 2007-0069.

2.2 RNA extraction and cDNA synthesis

Total RNA was extracted from the buffalo tissues following the manufacturer's instructions for TRIzol reagent (Thermo Fisher Scientific, USA). The RNA was incubated with RNase-free DNase I (TaKaRa, China) to eliminate genomic DNA contamination. RNA quality of different tissues was assessed using agarose gel electrophoresis. Their concentrations were determined with the NANODROP LITE spectrophotometer (Thermo Fisher Scientific). The RNA (3 µg) was synthesized to cDNA through a First Strand cDNA Synthesis Kit (TaKaRa).

2.3 Cloning of the full-length CDS of the PPARGC1A

Base on the mRNA sequence of Bos taurus PPARGC1A (accession no. NM_177945), a pair of primers (Table 1) were designed to clone the whole CDS of the buffalo PPARGC1A gene by Primer Premier 5.0 (Lalitha, 2000). The 25 µL reaction system was as follows: 0.6 µM of each primer, 100 ng of cDNA template (mixed cDNA from each tissue) and 12.5 µL of 2xGoldStar MasterMix (Dye) (CWBIO, Beijing, China). The PCR program initially started with 95 °C denaturation for 10 min, followed by 34 cycles of 95 °C for 30 s, 57.5 °C for 40 s, 72 °C for 90 s, then 72 °C extension for 5 min. The PCR products were detected by agarose gel electrophoresis. After gel extraction, the products were cloned into pMD-18T vector (TaKaRa) and sequenced bidirectionally using the Sanger method using ABI PRISM[®] BigDye[®] Terminator v3.1 Cycle Sequencing Kit (ABI, USA) on an ABI PRISM 3730 DNA sequencer according to the manufacturer's manual.

2.4 Physicochemical characteristics and structure analysis

The obtained sequence of buffalo *PPARGC1A* was checked, proofread and edited by the Lasergene software package (DNAstar Inc., USA). The open reading frame (ORF) was confirmed by Editseq (DNAstar Inc). Then, the homologous search was carried out to identify gene attributes by the BLAST program (https://blast.ncbi.nlm.nih.gov/Blast.cgi, last access: 20 March 2020) in the NCBI database. Physicochemical characteristics, hydropathy, signal peptide and transmembrane region were predicted by the ProtParam tool (https://web.expasy.org/protparam/, last access: 20 March 2020), ProtScale (https://web.expasy.

org/protscale/, last access: 20 March 2020) SignalP-5.0 Server (http://www.cbs.dtu.dk/services/SignalP/, last access: 20 March 2020; Almagro-Armenteros et al., 2019) and TMHMM version 2.0 (http://www.cbs.dtu.dk/services/ TMHMM/, last access: 20 March 2020), respectively. The conserved domain was determined through the Conserved Domain Architecture Retrieval Tool in BLAST. The subcellular localization was analyzed by ProtComp 9.0 (http://linux1.softberry.com/berry.phtml, last access: 16 March 2020). Secondary structures of deduced amino acid (AA) sequences were analyzed by SOPMA (http:// npsa-pbil.ibcp.fr/, last access: 20 March 2020). Biological process and molecular function analysis was further conducted by InterProScan (http://www.ebi.ac.uk/interpro/ search/sequence-search, last access: 21 March 2020).

2.5 RT-qPCR and tissue differential expression

To analyze tissue differential expression, a pair of primers were designed according to the obtained CDS of buffalo PPARGC1A in this work. The relative expression of PPARGC1A in 10 tissues during lactation and non-lactation were assayed by qPCR fluorescent technology using SYBR Premix Ex Taq (Takara) and performed on iQ5 Real Time PCR (Bio-Rad, USA) according to the manufacturers' instructions. The 20 µL reaction system included 2 µL cDNA, 10 µL SYBR Premix Ex Taq, 0.5 µL of 10 µM forward primer, 0.5 µL of 10 µM reverse primer, and 7 µL sterile water. The qPCR amplification was carried out firstly at 95 °C for 30 s, then followed by 40 cycles of 95 °C for 5 s, 60 °C for 20 s and 72 °C for 30 s. The β -actin (ACTB; accession no. NM 001290932) was used as an endogenous reference for normalization of PPARGC1A expression profiles (Table 1). The data of qPCR were analyzed through the $2^{-\Delta\Delta Ct}$ method, where $\Delta Ct = Ct_{PPARGC1A} - Ct_{ACTB}$ and $\Delta \Delta Ct =$ $\Delta Ct - \Delta Ct_{median}$. All the treatments were replicated three times. Significance of PPARGC1A mRNA level in multiple tissues between two periods was determined via Student's t test, which was established at a p < 0.05.

2.6 DNA isolation and polymorphism identification

Genomic DNA from the blood samples was isolated following a previous protocol (Sambrock and Russell, 2001). All primers used to amplify the exons were designed according to the genome sequence of buffalo *PPARGC1A* (accession no. NC_037551; Table 1). The mixture and protocol of PCR reaction were the same as that of clone, only by changing the annealing temperature and extension time. The PCR products were bidirectionally sequenced using the corresponding PCR primers.

The mutation sites were confirmed and outputted using Seqman (DNAstar Inc.) and Mega 6 (Tamura et al., 2013). Allele and genotype frequency and Hardy–Weinberg equilibrium test were carried out using PopGen32 software (Yeh and



Figure 1. PCR results for the buffalo *PPARGC1A* gene. M, λ DNA marker; 1, PCR product for the buffalo *PPARGC1A* gene.

Boyle, 1997). The function influence of non-synonymous substitutions was presumed by the program PANTHER (http: //www.pantherdb.org/, last access: 21 March 2020; Mi et al., 2017). The haplotypes were inferred by PHASE (Stephens et al., 2001). The optimal maximum likelihood model was determined by model selection tests, and then the phylogenetic tree was established based on the Hasegawa-Kishino-Yano model with a bootstrap test of 10 000 replicates.

3 Results

3.1 Cloning and identification of buffalo PPARGC1A

Being consistent with expectations, the PCR products of 2450 bp were obtained (Fig. 1). Sequence prediction showed that the cDNA sequence contained an ORF of 2394 bp. The homology search was performed by the BLAST program in the NCBI database, and the results displayed that the identity between the ORF sequence and the CDS sequences of the PPARGC1A gene in cattle (NM_177945), yak (XM 005897078), bison (XM 010839915), sheep (XM 004009738) and goat (NM 001285631) was 99.12%, 99.21 %, 98.29 %, 99.00 % and 98.83 %, respectively. Sequence analysis displayed that the nucleotide sequence of the PPARGC1A of river buffalo was consistent with that of swamp buffalo. The overall base composition of buffalo PPARGC1A CDS was composed of 29.62 % A, 22.68 % G, 21.85 % T and 25.86 % C. Buffalo PPARGC1A was deduced to encode a protein consisting of 797 amino acids (AAs) (Fig. 2).

	Primers (5' to 3')*	Product length (bp)	Annealing temperature (°C)	Usage
PPARGCIA	F: AACAGCTTGATTGGCGTCAT R: TATTCACCATGCCTCTGTCATCCTT	2450	57.5	CDS isolation
PPARGCIA	F:CCACCGAGAATGAGGCTAGTCCTT R:TTGACAAATGCTCTTCGCTTTATTGCTCCA	223	60	Differential expression
ACTB	F: TGGGCATGGAATCCTG R: GGCGCGATGATCTTGAT	196	60	Differential expression
Exon 1	F: ACAGGTGCCTTCAGTTCA R: CCAAACCCAAGCCCTTCC	254	56.1	SNP detection
Exon 2	F: TTTTCTCCCTGCCTCCTG R: CAAAGCAAGAACCCATTA	300	52.0	SNP detection
Exon 3	F: TACTCATCTCCCAGTGTCA R: AGCCAGAGGCAACTCCAA	300	53.7	SNP detection
Exon 4	F: CTCGCTTTCCCTCCTTCT R: AACCTCCTTGTGACTTCC	204	51.0	SNP detection
Exon 5	F: TTCCCTTTCTTTATGCCT R: CCTCACCACCCTTACCAG	278	53.1	SNP detection
Exon 6-7	F: CTGTTTCCAGTTTCCAAC R: ACACTCATCCATTCAAAA	366	50.2	SNP detection
Exon 8	F: ATCTCAGGGAAGTGAGGAA R: CACCAGGAACATGCTGTTGAG	1158	54.1	SNP detection
Exon 9	F: AGTCATGCTGATAAACTGGGTT R: GGATAAGAGGCACGGAGG	212	55.5	SNP detection
Exon 10	F: AAAATGTAGTCCAAAACC R: TAATCTATGCCCATCACA	309	50.7	SNP detection
Exon 11	F: TGGCATCAGTGTCTTTCC R: ATTCCCATCCTGGTAATC	263	51.2	SNP detection
Exon 12	F: TGCTAATGCTGCCTCACT R: GGTAAAAGGTAGTAATGG	309	48.4	SNP detection
Exon 13	F: AGGGTACATCTGACCTGG R: ATGCCTCTGTCATCCTTAGCC	169	51.1	SNP detection

 Table 1. Primer information on gene isolation, qPCR and polymorphism identification.

* Primer direction (F: forward; R: reverse).

3.2 Characteristics and structures of the PPARGC1A protein

The comparison of physicochemical characteristics of PPARGC1A between buffalo and cattle (accession numbers AC_000163) is shown in Table 2. The physicochemical characteristics of buffalo PPARGC1A were similar to those of cattle. The molecular weight of buffalo PPARGC1A was about 90.49 kDa, and its theoretical isoelectric point was 6.06. This protein belonged to an unstable protein with an instability index (II) of 74.87. Its grand average of hydropathicity was -1.086, illustrating that buffalo PPARGC1A is a hydrophilic protein. Secondary structure analysis showed

that the PPARGC1A consists of 31.37 % α -helix (250 AA), 11.67 % extended strand (93 AA), 4.27 % β turn (34 AA) and 52.70 % random coils (420 AA) (Fig. S1 in the Supplement). Furthermore, the buffalo PPARGC1A contained one conserved RRM_PPARGC1A (AA 674–764) functional domain (Fig. 3) without a signal peptide or a transmembrane region. Cytoplasm/nuclear localization analysis suggested that buffalo PPARGC1A was distributed in nucleus (35.4 %), plasma membrane (21.3 %), Golgi (17.9 %) and mitochondria (10.3 %).

L. Qiu et al.: Characterization, expression and polymorphisms of the buffalo PPARGC1A gene

1	ATGG	CGTGC	GACA	TGTG	CAAC	CCAG	GACT	CTG	ГАТG	GAG	ГGAC	CATC	GAG	TGT	GCT	GCT	CTG	GTT	GGT	GAA	GAC	CAG	CCT	CTI	TGC	CCA	GAT	CTT	ССТ	GAA	CTT	GAC	CTT	гсто	AAC	TAGA	CGT	GAA	CGAC	CTTC	GAT	ACA	GAC/	IGCT	ттсто	G 150
1	М	A W	D	M C	N	Q	D	S I	V W	S	D	Ι	Е	С	А	А	L	V	G	Е	D	Q	Р	L	С	Р	D	L	Ρ	Е	L	D	L	S	E	LI) V	Ν	D	L	D	Т	D	S	FL	50
151	GGTG	GACTO	CAAGT	GGTG	CAGI	GAC	CAAT	CAG/	AAAT	AAT/	ATCT	TAAT	CAG	TAC	AAC	AAC	GAG	CCT	TCA	AAC	ATA	TTT	GAG	AAG	GATA	GAT	GAA	GAG	AAT	GAG	GCA	AAC	TTG	CTAC	CGG	TCCI	CAC	AGAG	GAC	4CT(GAC	CAGT	CTC	CTG	TGGAI	006 1
51	G	G L	Κ	W C	S	D	Q	S I	ΞI	Ι	S	Ν	Q	Y	Ν	Ν	Е	Р	S	Ν	Ι	F	Е	Κ	Ι	D	Е	Е	Ν	Е	А	Ν	L	L	A	V I	, T	Е	Т	L	D	S	L	Р	V D	100
301	GAAG	ACGGA	TTGO	ССТС	ATT	GAT	GCAC	TGAG	CAGA	TGG/	AGAT	GTG	ACC	ACC	GAG	AAT	GAG	GCT	AGT	ССТ	TCC	TCC	ATG	CCI	ſĠĂĊ	GGC	ACC	ССТ	CCG	CCT	CAG	GAG	GCA	GAAG	AGC	CGTO	TCT	ACT	TAAC	JAA	CTC	TTA	CTG	GCAC	CAGCO	2 450
101	Е	DG	L	P S	F	D	Α	L 1	ГD	G	D	V	Т	Т	Е	Ν	Е	А	S	Р	S	S	М	Р	D	G	Т	Р	Р	Р	Q	Е	А	Е	Е	P S	L	L	Κ	Κ	L	L	L	A	ΡA	150
451	AACA	CTCAC	CTA/	GTTA	TAAT	GAA'	ГGCA	GTG	GCCT	CAGT	ГАСС	CAG	AAC	CAT	GCA	AAC	CAT.	AAT	CAC	AGG	ATC	CAGA	ACA	AAC	CCCI	GCA	GTT	GTT	AAG	ACC	GAG	AAT	TCA	rgg/	GCA	ATA/	AGC	GAA	GAGC	CAT?	TGT	CAA	CAGO	CAAA	AGCC/	1 600
151	Ν	ΤQ	L	S Y	N	Е	С	s (ΓL	, S	Т	Q	Ν	Н	А	Ν	Н	Ν	Н	R	Ι	R	Т	Ν	Р	А	V	V	Κ	Т	Е	Ν	S	W	S	NF	C A	K	S	Ι	С	Q	Q	Q	ΚP	200
601	CAAA	GACGI	CCGI	GCTC	AGAG	GCTT	CTCA	AGT/	ATCT	GAC	CACA	AAT	GAT	GAC	CCT	ССТ	CAC	ACC	AAA	CCC	ACA	GAG	AAC	CGG	GAAC	AGC	AGC	AGA	GAC	AAA	TGC	ACC	TCC	AAAA	AGA	AGGO	CCA	CAC	ACA	4TCC	CAG	ACA	CAAC	CATC	TACA	A 750
201	Q	R R	Р	C S	Е	L	L	K Y	ΥL	, T	Т	Ν	D	D	Р	Р	Н	Т	Κ	Р	Т	Е	Ν	R	Ν	S	S	R	D	Κ	С	Т	S	Κ	Κ	K A	Н	Т	Q	S	Q	Т	Q	Н	LQ	250
751	GCCA	AACCA	ACAA	CTTT	ATCI	CTT	сстс	TGAG	cccc	AGAC	GTCA	CCA	AAT	GAC	ccc	AAG	GGT	тсс	CCA	TTT	GAG	AAC	CAAG	ACT	TAT	GAA	CGA	ACC	ТТА	AGT	GTG	GAA	CTC	гсто	GAA	CTG	AGG	CCT	AACT	rcc/	ACCC	CACA	ACTO	стс	CTCAT	000
251	А	K P	Т	T L	S	L	Р	L 1	ΓР	E	S	Р	Ν	D	Р	Κ	G	S	Р	F	Е	Ν	Κ	Т	Ι	Е	R	Т	L	S	V	Е	L	S	G	T A	G	L	Т	Р	Р	Т	Т	Р	РH	300
901	AAAG	CCAAC	CAAC	ATAA	CCCI	TTC	AGGG	CTTO	CTCC	AAA	GCTC	GAAG	CCC	TCT	TGC	AAG.	ACT	GTG	GTA	CCT	CCA	CCA	TCC	CAAC	GAAG	GCC	CGG	TAC	AGT	GAG	TCT	TCT	TGT	ACCO	AAG	GCAG	TAA	TTC	CACC	CAAC	GAAC	GGGG	ССТО	GAGC	AGTCI	1050
301	Κ	A N	Q	D N	Р	F	R	A S	S P	K	L	Κ	Р	S	С	Κ	Т	V	V	Р	Р	Р	S	Κ	Κ	Α	R	Y	S	Е	S	S	С	Т	Q	G S	N	S	Т	Κ	Κ	G	Р	Е	Q S	350
1051	GAGT	TGTAC	GCAC	CAGCT	CAGO	CAAG	ACCT	CTGI	ГGCT	CAC	CAGT	GGA	CAC	GAG	GAA	AGG.	AAG	GCC	AAA	CGG	CCC	CAGT	CTC	CGG	GCTO	TTT	GGT	GAC	CAT	GAC	TAT	TGT	CAG	rcg/	TTA	ATTO	CAA	AAT	GGA	4AT/	ICTC	GTT	AGT/	CAT	CACAG	G 1200
351	Е	LΥ	А	Q L	S	Κ	Т	S I	ΓL	. Т	S	G	Η	Е	Е	R	Κ	А	Κ	R	Р	S	L	R	L	F	G	D	Η	D	Y	С	Q	S	I	N S	K	M	Е	Ι	L	V	S	Т	S Q	400
1201	GAGC	TCCAC	GACI	CCAG	ACA	ACTA(GAAA	ATA/	AAGA	TGCC	CCCC	CTCC	TCC	AAC	GGG	TCG	GGG	CAA	ATA	CAC	TCT	TCC	CACA	GAT	ГТСС	GAC	CCG	TGC	TAC	CTG.	AGA	GAG.	ACT	GCAG	AGG	TGAC	CAG	GCA	GGTC	CTC?	1000	CGGC	AGC/	ACCA	GAAA/	1350
401	Е	L H	D	S R	Q	L	Е	Nł	K D	A	Р	S	S	Ν	G	S	G	Q	Ι	Н	S	S	Т	D	S	D	Р	С	Y	L	R	Е	Т	А	Е	V S	R	Q	V	S	Р	G	S	Т	R K	450
1351	CAGC	TCCAA	GACO	CAGGA	AATO	CCGA	GCCG	AGCI	ГGAA	CAAC	GCAC	CTTC	GGT	CAT	CCC	AGT	CAA	GCT	GTT	TTT	GAC	GAC	CAAA	GCA	AGAC	CAAG	ACC	AGT	GAA	CTG	AGG	GAC	AGT	GATI	TCA	GTAA	TGA	ACA	ATTC	CTCC	CAAA	CTA	CCTA	TGT	TTAT/	A 1500
451	Q	LQ	D	Q E	I	R	А	ΕI	L N	K	Η	F	G	Η	Р	S	Q	А	V	F	D	D	Κ	Α	D	Κ	Т	S	Е	L	R	D	S	D	F	S N	ΙE	Q	F	S	Κ	L	Р	M	FΙ	500
1501	AATT	CAGGA	CTAC	CCAT	GGAI	GGC	CTGT	TTG	ATGA	CAG	CGAA	GAT	GAA	AGT	GAT/	AAA	CTG	AAC	TCC	CCT	TGG	GAT	GGC	CACO	GCAG	TCC	TAT	TCA	TTG	TTC	CAT	GTG	TCG	CCTI	CTT	GTTO	TTC	TTT	TAAC	CTC	ICCA	TGT	AGAC	GATT	CTGT	G 1650
501	Ν	S G	L	A M	D	G	L	FΙ	D	S	Е	D	Е	S	D	Κ	L	Ν	S	Р	W	D	G	Т	Q	S	Y	S	L	F	Η	V	S	Р	S	C S	S S	F	Ν	S	Р	С	R	D	S V	550
1651	TCAC	CACCO	CAAAT	CTTT	ATTI	TCT	CAAA	GACO	CCCA	AAGO	GATC	GCGC	TCT	CGT	TCA	AGG	TCC	TTT	TCT	CGA	CAC	CAGG	TCA	TGI	ГТСТ	CGA	TCA	CCA	TAT	TCC.	AGG	TCA.	AGA	ГCA/	GGT	CCCC	CAGG	CAG	TAGA	ATC7	TCT	TCA	AGAT	ICTT	GCTAC	1800
551	S	P P	Κ	S L	F.	S	Q	R I	P Q	R	М	R	S	R	S	R	S	F	S	R	Н	R	S	С	S	R	S	Р	Y	S	R	S	R	S	R	S F	, G	S	R	S	S	S	R	S	С Ү	600
1801	TACT	ATGAC	TCAC	GCCA	CTG	CAGAG	CACC	GCAG	CACA	.CCG/	AAAT	TCG	CCC	CTG	TGC	GCG.	AGGʻ	TCA	CGT	TCA	AGA	TCG	CCC	CAT	rago	CGG	CGG	CCC	AGG	TAT	GAC	AGC	TAC	GAGO	AAT	ACCA	GCA	CGA/	AAGC)CTC	CAAG	AGG	GAAG	GAAT	ACCG	1950
601	Y	ΥE	S	G H	C	R	Н	R 1	ΓН	R	Ν	S	Р	L	С	А	R	S	R	S	R	S	Р	Η	S	R	R	Р	R	Y	D	S	Y	Е	E	Y G) H	E	R	L	Κ	R	Е	E	Y R	650
1951	AGAG	AGTAT	GAGA	AGCG	GGA	TCT	GAAA	GGGG	CCAA	GCAG	GAGC	GAG	AGG	CAG	AGG	CAG	AAG	GCA	ATT	GAA	GAG	CGC	CGT	GTO	GATI	TAT	GTT	GGT	AAA	ATC.	AGA	CCT	GAC	ACAA	CAC	GGAG	CAGA	ACTO	GAGC	<u>}GA(</u>	CCGT	TTT	GAAG	JTTT	TTGG1	2100
651	R	E Y	Е	K R	E	S	Е	R A	A K	Q	R	Е	R	Q	R	Q	Κ	А	Ι	Е	Е	R	R	V	Ι	Y	V	G	Κ	Ι	R	Р	D	Т	Т	R1	È	L	R	D	R	F	Е	V	FG	700
2101	GAAA	TTGAC	GAGT	GCAC	AGTA	AAT	CTGC	GGG/	ATGA	.TGG/	AGAC	CAGC	TAT	GGT	TTC	ATT.	ACC	TAC	CGT	TAT	ACC	CTGT	GAT	GCI	ITTI	GCT	GCT	CTT	GAA	AAT	GGA	TAC.	ACT	ГТGC	GCA	GGTC	GAA	TGA/	AACT	(GAC	CTTC	GAG	CTG7	TACT	TTTGI	2250
701	E	ΙE	Е	C T	V	Ν	L	R I) D	G	D	S	Y	G	F	Ι	Т	Y	R	Y	Т	С	D	A	F	Α	A	L	Е	Ν	G	Y	Т	L	R	R S	N	E	Т	D	F	Е	L	Y	FС	750
2251	GGAC	GCAAC	CAAT	TTTT	CAAC	GTCT/	AACT	ATG	CAGA	.CCT/	AGAT	TCA	AAT	TCA	GAT	GAC	TTT	GAC	CCT	GCT	TCC	CATC	CAAG	AGC	CAAG	TAT	GAC	TCT	CTG	GAT	TTC	GAT.	AGT	ΓTAC	TGA	AAG/	AGC	CCAC	GAGA	AGG	CTTA	CGC.	AGGI	`AA		2394
751	G	R K	Q	F F	K	S	Ν	Y A	A D	L	D	S	Ν	S	D	D	F	D	Р	А	S	Ι	К	S	К	Y	D	S	L	D	F	D	S	L	L	K E	E A	Q	R	S	L	R	R	*		797

Figure 2. Nucleotide sequence of buffalo *PPARGC1A* and its deduced amino acid sequence. The predicted protein sequence is shown immediately above the nucleotide sequence. The RRM_PPARGC1A domain (AA 674–764) is boxed. The stop codon is indicated by an asterisk (*).

Basic physical and chemical properties	Buffalo	Cattle
Formula	C ₃₈₆₅ H ₆₀₇₂ N ₁₁₅₈ O ₁₃₀₂ S ₂₈	C ₃₈₅₇ H ₆₀₅₆ N ₁₁₅₂ O ₁₃₀₂ S ₂₈
Number of amino acids	797	796
Molecular weight	90.49 kDa	90.30 kDa
Isoelectric point (pI)	6.06	5.92
Strongly acidic amino acid (D, E)	122	123
Strongly basic amino acid (K, R)	111	110
Polar amino acid (N, C, Q, S, T, Y)	281	281
Hydrophobic amino acid (A, I, L, F, W, V)	172	172
Instability index (II)	74.87	75.01
Grand average of hydropathicity (GRAVY)	-1.086	-1.082
Aliphatic index	53.14	53.08

3.3 Biological process and molecular function

Predictions showed that buffalo PPARGC1A was involved in the biological process of mitochondrial organization (GO:0007005) and brown fat cell differentiation (GO:0050873). Its molecular functions are mainly nucleic acid binding (GO:0003676), transcription factor binding (GO:0008134), signaling receptor binding (GO:0003712).

3.4 Tissue differential expression of the *PPARGC1A*

The tissue differential expression of the *PPARGC1A* gene was assayed via qPCR in 10 tissues of lactating and non-lactating river buffalo (Fig. 4). The results indicated that the *PPARGC1A* gene was expressed in almost all tissues dur-

ing these two periods, especially in the muscle, heart, liver, brain and kidney. However, the expression of this gene in the brain and kidney did not change significantly between lactation and non-lactation (P > 0.05). The expression of the *PPARGC1A* gene in the heart, liver, spleen and lung during lactation was significantly higher than that during non-lactation (P < 0.05), but its expression in the muscle, mammary gland, small intestine and rumen was on the contrary (Table S1 in the Supplement, Fig. 4).

3.5 Population variation analysis

Five synonymous SNPs were found in the samples of this study, namely c.759A>G, c.778C>T, c.1257G>A, c.1311G>A and c.1509A>T (Table 3). The c.759A>G and



Figure 3. Putative conserved functional domain of buffalo PPARGC1A.



Figure 4. Tissue differential expression of buffalo *PPARGC1A* in 10 tissues during lactation and non-lactation. The values are presented as means \pm SEM; * *P* < 0.05, ** *P* < 0.01, *** *P* < 0.001.

c.1509A>T were found only in river buffalo, and the others were shared by two types of buffalo. The information on each SNP is shown in Table 3. The c.778C and c.1257G were the alleles with high frequency in two types of buffalo. It is worth noting that SNP1311 in swamp buffalo were all heterozygous. The test of Hardy–Weinberg equilibrium showed that SNP778 in both types of buffalo and SNP1311 in swamp buffalo were in dis-equilibrium (P < 0.05).

By analyzing *PPARGC1A* gene sequences of buffalo obtained in this work and the published data of this gene in the NCBI database, we discovered that the number of SNP in buffalo *PPARGC1A* increased to eight. Among them, the c.419C>T, c.920C>A and c.926G>A were from published data in the NCBI database, which were only observed in river buffalo. They are all non-synonymous, resulting in amino acid changes of p.Ser140Phe, p.Pro307His and p.Arg309Lys in the PPARGC1A (Figs. 5 and 6). The prediction showed that the substitutions of p.Ser140Phe and p.Pro307His may affect the function of buffalo PPARGC1A (Table 4).

3.6 Haplotype sequence differences and phylogenetic relationship

Based on the polymorphisms of the *PPARGC1A* gene, a total of 12 haplotypes (B1–B12) were defined in two types of buffalo (Fig. 5). Among them, 10 (B1–B10) (accession numbers MN788077–MN788086) were obtained from the data of this work (Table 5), and the other two were from published data (accession numbers HQ236498 and NW_005785900). Among these haplotypes, B1–B5 were shared by two types of buffalo, and the rest were only found in river buffalo.

In order to explore the sequence differences of the *PPARGC1A* gene between buffalo and other animals in

Bovidae, all the haplotype sequences of buffalo in this work were compared with the published homologous sequences of other species of Bovidae. The accession numbers in the NCBI database of the representative haplotypes for each species are XM 010806009, AC 000163, XM 005897078, NW_011494708, NW_014639015, XM_012131592, NC_019463, NW_011942373, HM600810, NC_030813, EU304457. XM_005980241, NM_001286596 and NW_005815902. The sequence differences of nucleotide and its corresponding amino acid among all the species are shown in Figs. 5 and 6, respectively. There were seven nucleotide differences located at c.1041, c.1161, c.1175, c.1255, c.1308, c.1600 and c.1728 of this gene between buffalo and other species of Bovidae, including three nucleotide differences (c.1175, c.1255 and c.1600) which led to the amino acid differences of the PPARGC1A (the corresponding amino acids in buffalo PPARGC1A were p.392Met, p.419Ser and p.534His, respectively) (Fig. 6). It is noteworthy that the PPARGC1A gene of buffalo and the Ovis genus was three nucleotides longer (from c.1858 to c.1860) than that of Bos (with an Arg insertion at p.620 in buffalo and Ovis).

Based on the haplotype sequences of buffalo, cattle, yak, bison, sheep, goat and chiru, a phylogenetic tree was established with a homologous sequence of mouse as an outgroup (Fig. 7). The phylogenetic analysis showed that buffalo, *Bos* and *Ovis* gathered on their own independent clades with high supports. The genetic relationship between buffalo and the species of *Bos* is closer than that of the species in *Ovis*.

4 Discussion

In this work, the whole CDS of the *PPARGC1A* was cloned from two types of buffalo. The PPARGC1A for both river and swamp buffalo was all composed of 797 amino acid residues, and its basic physicochemical properties were similar to those of cattle. The prediction of subcellular localization showed that the buffalo PPARGC1A was distributed not only in the nucleus and plasma membrane, but also in the mitochondria and Golgi, indicating that buffalo PPARGC1A may exert a biological function in the nucleus, plasma membrane, mitochondria and Golgi. Previous studies have shown that PPARGC1A contains a RRM_PPARGC1A domain and is an inducible transcriptional coactivator that can interact with a variety of transcription factors, which are related to various biological processes, including adaptive thermogenesis, glucose/fatty acid metabolism, skeletal muscle fiber

Population	SNP	Genotype	Frequency	Allele	Frequency	P value*
River buffalo	c.759A>G	AA	0.500	А	0.7083	0.9012
		AG	0.417	G	0.2917	
		GG	0.083			
	c.778C>T	CC	0.750	С	0.7500	0.0002
		CT	0.000	А	0.2500	
		TT	0.250			
	c.1257G>A	GG	0.667	С	0.8333	0.5556
		GA	0.333	Т	0.1667	
		AA	0.000			
	c.1311G>A	GG	0.083	G	0.4167	0.2542
		GA	0.667	С	0.5833	
		AA	0.250			
	c.1509A>T	AA	0.917	С	0.9583	1.0000
		AT	0.083	G	0.0417	
		TT	0.000			
Swamp buffalo	c.759A>G	AA	1.000	А	1.0000	_
		AG	0.000	G	0.0000	
		GG	0.000			
	c.778C>T	CC	0.778	С	0.7778	0.0007
		CT	0.000	А	0.2222	
		TT	0.222			
	c.1257G>A	GG	0.667	С	0.8333	0.6326
		GA	0.333	Т	0.1667	
		AA	0.000			
	c.1311G>A	GG	0.000	G	0.5000	0.0047
		GA	1.000	С	0.5000	
		AA	0.000			
	c.1509A>T	AA	1.000	С	1.0000	-
		AT	0.000	G	0.0000	
		TT	0.000			

 Table 3. Genetic information on the SNPs found in two types of buffalo.

* P value of Hardy–Weinberg equilibrium test.

type switching and cardiac development (Spiegelman et al., 2000; Yoon et al., 2001; Mortensen et al., 2006). It has been confirmed that PPARGC1A can interact with steroid receptors and activate them (Knutti et al., 2000). It can also coordinate the expression of genes participated in fatty acid metabolism (Dominy et al., 2010) and is closely related to milk fat synthesis (Weikard et al., 2005; Chen et al., 2017). In this study, it is predicted that buffalo PPARGC1A also contains a RRM PPARGC1A domain, and its molecular functions are mainly involved in nucleic acid binding, signal receptor binding, transcription factor binding and transcriptional coregulator activity. The sequence consistency, physicochemical properties, and structure of buffalo PPARGC1A were similar to those of cattle. Based on the above results, it is speculated that buffalo PPARGC1A is also associated with the regulation of carbohydrate and lipid metabolism, and it exerts a function in adipose tissue, skeletal muscle, heart, liver and mammary gland. These results also indicate that buffalo PPARGC1A may have similar functions to other mammals, especially the Bovidae species.

Previous studies have shown that the mouse *PPARGC1A* gene is highly expressed in brown fat, heart, kidney and brain, but it is lower in the liver and the lowest in white adipose tissue (Puigserver et al., 1998). And the human *PPARGC1A* is highly expressed in the heart, skeletal muscle and kidney, but its expression in the liver, brain, pancreas and perirenal adipose tissue is low (Esterbauer et al., 1999). In this study, the *PPARGC1A* gene was highly expressed in the muscle, heart, liver, brain and kidney of both lactating and non-lactating buffaloes, indicating that this gene exerts a key role in these buffalo tissues under various physiological conditions. The expression levels of buffalo *PPARGC1A* in the

	111 345	222 000 567	222 222 345	222 666 234	333 222 567	444 112 890	444 222 456	$666 \\ 111 \\ 012$	777 112 890	777 333 678	777 444 567	777 555 789	777 778 890	888 333 567	888 666 567	888 888 678	999 122 901	999 222 567	999 333 123	999 333 456	999 777 678	999 788 901	999 888 567	111 000 000 678	111 000 344 901	111 111 222 678	111 111 566 901	111 111 666 234	111 111 777 456	111 111 999 234	111 222 000 789	111 222 555 567	111 222 666 123	111 222 889 890
B1	ATG	AAC	TTT	GCG	GCA	TCT	CTT	CCG	GCC	ACA	CTA	CCA	CTG	GAA	ACT	ACA	CCT	AGG	TCT	CCA	TCC	AAG	GCC	TGT	CCT	CTC	TCG	ATT	ATG	ACA	CAC	TCG	CAA	CCC
B2																																		
B3																																A		
B4													Τ																					
B5													Τ																					
B6																				• • •														
B7					• • •	• • •					• • •	• • •			• • •	• • •	• • •	• • •	• • •	• • •	• • •	• • •	• • •	• • •	• • •			• • •			• • •	A	• • •	• • •
B8	• • •	• • •	• • •	• • •	• • •	• • •	• • •	• • •	• • •	• • •	• • •	•••	Τ	• • •	• • •	• • •	• • •	•••	• • •	• • •	• • •	• • •	• • •	• • •	• • •	• • •	• • •	• • •	• • •	• • •	• • •	A	• • •	• • •
B9	• • •	• • •	• • •	• • •	• • •	• • •	• • •	• • •	• • •	• • •	• • •	G	• • •	• • •	• • •	• • •	• • •	• • •	• • •	• • •	• • •	• • •	• • •	• • •	• • •	• • •	• • •	• • •	• • •	• • •	• • •	• • •	• • •	• • •
BIU D11	• • •	• • •	• • •	• • •	• • •	• • •	• • •	• • •	• • •	• • •	• • •	6	• • •	• • •	• • •	• • •		• • •	• • •	• • •	• • •	• • •	• • •	• • •	• • •	• • •	• • •	• • •	• • •	• • •	• • •	• • •	• • •	• • •
DII DII	• • •	• • •	• • •	• • •	• • •		• • •	• • •	• • •	• • •	• • •	• • •	• • •	• • •	• • •	• • •	. A.	. A.	• • •	• • •	• • •	• • •	• • •	• • •	• • •	• • •	• • •	• • •	• • •	• • •	• • •	• • •	• • •	• • •
oattle hanl	• • •	• • •	• • •		• • •	• 1•	• • •	• • •	• • •	• • •	• • •	• • •		• • •		• • •	• • •	• • •	• • •	• • •		• • •	• • •	• • •	· · · ·	• • •		• • •	· · ·	• • •	· · · · т	 C	• • •	• • •
cattle_hap1	•••	• • •	•••		•••	• • •	с	• • •	• • •	•••	•••	• • •	• • •	•••	•••	•••	• • •	•••	•••	• • •	•••	• • •	• • •	• • •	0	• • •		• • •	. C.	•••	 Т	с	•••	• • •
vak han1	•••		•••	A	•••		0	• • •	• • •	•••	•••			• • •	• • •				• • •	• • •			• • •		c	• • •	A		. с.	•••	•••	с	•••	
hison han1				A																							A		. C.			C		
sheep hap1					G					G				G							G			G	C	T	A		. C.	. T.		C	G	. T.
sheep hap2					G					G				G						G	G			G	C	T	A		. C.	. T.		С	G	. T.
sheep hap3					G		G			G				G							G			G	C	T	A		. C.	. T.		С	G	. T.
sheep hap4					G	. T.				G				G						G	G			G	C	T	A		.с.	. T.		С	G	.Т.
goats_hap1	???		C		G									G	A				C		G			G	C	T	A		.С.	.Т.		С	G	
goats_hap2			C		G			A						G					C		G			G	C	T	A		.С.	.T.		С	G	
goats_hap3	Τ		C		G									G					C		G	G		G	C	T	A		.С.	.Т.		С	G	
chiru_hap1										G				G							G		Α	G	C	T	A	G	.с.	.T.	T	С	G	
chiru_hap2									. A.	G	Τ			G		C					G		Α	G	C	T	A	G	.С.	.T.	T	С	G	
chiru_hap3						.T.				G				G							G		Α	G	C	T	A	G	.С.	.Т.	T	С	G	
																											000	000	000	000	000	000	000	
	111	111	111	111	111	111	111	111	111	111	111	111	111	111	111	111	111	111	111	111	111	111	111	111	111	111	222	222	222	222	222	222	222	
	111 333	111 333 011	111 333	111 333	111 333 777	111 333	111 333	111 555	111 555	111 555	111 666	111 666	111 666	111 666	111 777	111 777 222	111 888 444	111 888 555	111 888 555	111 888 556	111 999	111 999	$ \begin{array}{c} 111 \\ 999 \\ 222 \end{array} $	111 999	111 999 777	111 999	222 000	222 000	222 111	222 222	222 222 445	222 333	222 333	
	111 333 000 678	111 333 011 901	111 333 333 345	111 333 555 456	111 333 777 567	111 333 888 456	111 333 999 678	111 555 000 789	111 555 555 567	111 555 999 789	111 666 000	111 666 000 345	111 666 333 345	111 666 666 678	111 777 000 234	111 777 222 678	111 888 444 012	111 888 555 234	111 888 555 567	111 888 556 890	111 999 111 234	111 999 222 123	111 999 222 789	111 999 333 345	111 999 777 234	111 999 999 678	222 000 444 123	222 000 555 012	222 111 111 345	222 222 000 678	222 222 445 890	222 333 000 234	222 333 222 012	
B1	111 333 000 678 ACT	111 333 011 901 GCG	111 333 333 345 CCC	111 333 555 456 CTC	111 333 777 567 GCC	111 333 888 456 AAC	111 333 999 678 GGT	111 555 000 789 GGA	111 555 555 567 444	111 555 999 789 TTC	111 666 000 012 CAT	111 666 000 345 GTG	111 666 333 345 CCA	111 666 666 678 TTA	111 777 000 234 TCA	111 777 222 678 TCA	111 888 444 012 AAT	111 888 555 234 TGC	111 888 555 567 GCG	111 888 556 890 AGG	111 999 111 234 GAA	111 999 222 123 CAC	111 999 222 789 AGG	111 999 333 345 AAG	111 999 777 234 TCT	111 999 999 678 AGG	222 000 444 123 GGT	222 000 555 012 AGA	222 111 111 345 TGC	222 222 000 678 ACT	222 222 445 890 TGT	222 333 000 234 GAT	222 333 222 012 TCC	
B1 B2	111 333 000 678 ACT	111 333 011 901 GCG	111 333 333 345 CCC	111 333 555 456 CTC	111 333 777 567 GCC	111 333 888 456 AAC	111 333 999 678 GGT	111 555 000 789 GGA	111 555 555 567 AAA	111 555 999 789 TTC	111 666 000 012 CAT	111 666 000 345 GTG	111 666 333 345 CCA	111 666 666 678 TTA	111 777 000 234 TCA	111 777 222 678 TCA	111 888 444 012 AAT	111 888 555 234 TGC	111 888 555 567 GCG	111 888 556 890 AGG	111 999 111 234 GAA	111 999 222 123 CAC	111 999 222 789 AGG	111 999 333 345 AAG	111 999 777 234 TCT	111 999 999 678 AGG	222 000 444 123 GGT	222 000 555 012 AGA	222 111 111 345 TGC	222 222 000 678 ACT	222 222 445 890 TGT	222 333 000 234 GAT	222 333 222 012 TCC	
B1 B2 B3	111 333 000 678 ACT	1111 333 011 901 GCG A	111 333 333 345 CCC 	111 333 555 456 CTC	111 333 777 567 GCC	111 333 888 456 AAC	111 333 999 678 GGT 	111 555 000 789 GGA	111 555 555 567 AAA 	111 555 999 789 TTC	111 666 000 012 CAT	111 666 000 345 GTG 	111 666 333 345 CCA	111 666 666 678 TTA	111 777 000 234 TCA	111 777 222 678 TCA	111 888 444 012 AAT 	111 888 555 234 TGC 	111 888 555 567 GCG 	111 888 556 890 AGG	111 999 111 234 GAA 	111 999 222 123 CAC	111 999 222 789 AGG	111 999 333 345 AAG	111 999 777 234 TCT 	111 999 999 678 AGG	222 000 444 123 GGT	222 000 555 012 AGA 	222 111 111 345 TGC	222 222 000 678 ACT	222 222 445 890 TGT	222 333 000 234 GAT 	222 333 222 012 TCC 	
B1 B2 B3 B4	1111 333 000 678 ACT 	111 333 011 901 GCG A	111 333 333 345 CCC 	111 333 555 456 CTC 	1111 333 777 567 GCC 	1111 333 888 456 AAC 	111 333 999 678 GGT 	111 555 000 789 GGA 	111 555 555 567 AAA 	111 555 999 789 TTC 	1111 6666 000 012 CAT 	111 666 000 345 GTG 	111 666 333 345 CCA 	1111 666 666 678 TTA 	1111 777 000 234 TCA 	1111 777 222 678 TCA 	1111 888 444 012 AAT 	1111 888 555 234 TGC 	111 888 555 567 GCG 	111 888 556 890 AGG 	111 999 111 234 GAA 	111 999 222 123 CAC 	111 999 222 789 AGG 	1111 999 333 345 AAG 	1111 999 777 234 TCT 	111 999 999 678 AGG 	222 000 444 123 GGT 	222 000 555 012 AGA 	222 111 111 345 TGC 	222 222 000 678 ACT 	222 222 445 890 TGT 	222 333 000 234 GAT 	222 333 222 012 TCC 	
B1 B2 B3 B4 B5	1111 333 000 678 ACT 	1111 333 011 901 GCG A	1111 333 333 345 CCC 	111 333 555 456 CTC 	1111 333 777 567 GCC 	1111 333 888 456 AAC 	111 333 999 678 GGT 	111 555 000 789 GGA 	111 555 555 567 AAA 	111 555 999 789 TTC 	1111 666 000 012 CAT 	111 666 000 345 GTG 	111 666 333 345 CCA 	1111 666 678 TTA 	111 777 000 234 TCA 	1111 777 222 678 TCA 	1111 888 444 012 AAT 	111 888 555 234 TGC 	111 888 555 567 GCG 	111 888 556 890 AGG 	111 999 111 234 GAA 	111 999 222 123 CAC 	111 999 222 789 AGG 	111 999 333 345 AAG 	1111 999 777 234 TCT 	111 999 999 678 AGG 	222 000 444 123 GGT 	222 000 555 012 AGA 	222 111 111 345 TGC 	222 222 000 678 ACT 	222 222 445 890 TGT 	222 333 000 234 GAT 	222 333 222 012 TCC 	
B1 B2 B3 B4 B5 B6	1111 333 000 678 ACT 	1111 333 011 901 GCG A 	1111 333 345 CCC 	111 333 555 456 CTC 	1111 333 777 567 GCC 	1111 333 888 456 AAC 	111 333 999 678 GGT 	111 555 000 789 GGA 	111 555 555 567 AAA 	1111 555 999 789 TTC 	1111 666 000 012 CAT 	111 666 000 345 GTG 	111 666 333 345 CCA 	111 666 666 678 TTA 	111 777 000 234 TCA 	1111 777 222 678 TCA 	111 888 444 012 AAT 	111 888 555 234 TGC 	111 888 555 567 GCG 	111 888 556 890 AGG 	1111 999 1111 234 GAA 	111 999 222 123 CAC 	111 999 222 789 AGG 	111 999 333 345 AAG 	1111 999 777 234 TCT 	111 999 999 678 AGG 	222 000 444 123 GGT 	222 000 555 012 AGA 	222 111 111 345 TGC 	222 222 000 678 ACT 	222 222 445 890 TGT 	222 333 000 234 GAT 	222 333 222 012 TCC 	
B1 B2 B3 B4 B5 B6 B7	1111 333 000 678 ACT 	1111 333 011 901 GCG A A	1111 333 345 CCC 	111 333 555 456 CTC 	1111 333 777 567 GCC 	1111 333 888 456 AAC 	1111 333 999 678 GGT 	111 555 000 789 GGA T	111 555 555 567 AAA 	1111 555 999 789 TTC 	1111 666 000 012 CAT 	111 666 000 345 GTG 	111 666 333 345 CCA 	111 666 666 678 TTA 	111 777 000 234 TCA 	111 777 222 678 TCA 	1111 888 444 012 AAT 	1111 888 555 234 TGC 	111 888 555 567 GCG 	111 888 556 890 AGG 	111 999 111 234 GAA 	111 999 222 123 CAC 	111 999 222 789 AGG 	111 999 333 345 AAG 	1111 999 777 234 TCT 	111 999 999 678 AGG 	222 000 444 123 GGT 	222 000 555 012 AGA 	222 111 345 TGC 	222 222 000 678 ACT 	222 222 445 890 TGT 	222 333 000 234 GAT 	222 333 222 012 TCC 	
B1 B2 B3 B4 B5 B6 B7 B8	1111 333 000 678 ACT 	1111 333 011 901 GCG A A	1111 333 345 CCC 	1111 333 555 456 CTC 	1111 333 777 567 GCC 	1111 333 888 456 AAC 	1111 333 999 678 GGT 	111 555 000 789 GGA T	111 555 567 AAA 	1111 555 999 789 TTC 	111 666 000 012 CAT 	1111 6666 000 345 GTG 	1111 666 333 345 CCA 	111 666 666 678 TTA 	111 777 000 234 TCA 	111 777 222 678 TCA 	1111 888 444 012 AAT 	1111 888 555 234 TGC 	111 888 555 567 GCG 	111 888 556 890 AGG 	111 999 111 234 GAA 	111 999 222 123 CAC 	111 999 222 789 AGG 	111 999 333 345 AAG 	111 999 777 234 TCT 	111 999 999 678 AGG 	222 000 444 123 GGT 	222 000 555 012 AGA 	222 111 345 TGC 	222 222 000 678 ACT 	222 222 445 890 TGT 	222 333 000 234 GAT 	222 333 222 012 TCC 	
B1 B2 B3 B4 B5 B6 B7 B8 B9	1111 333 000 678 ACT 	1111 333 011 901 GCG A A A	1111 333 345 CCC 	1111 333 555 456 CTC 	1111 333 777 567 GCC 	1111 333 888 456 AAC 	1111 333 999 678 GGT 	111 555 000 789 GGA 	111 555 555 567 AAA 	1111 555 999 789 TTC 	1111 666 000 012 CAT 	1111 666 000 345 GTG 	1111 666 333 345 CCA 	1111 666 666 678 TTA 	111 777 000 234 TCA 	111 777 222 678 TCA 	1111 888 444 012 AAT 	1111 888 555 234 TGC 	111 888 555 567 GCG 	111 888 556 890 AGG 	111 999 111 234 GAA 	111 999 222 123 CAC 	111 999 222 789 AGG 	111 999 333 345 AAG 	111 999 777 234 TCT 	111 999 999 678 AGG 	222 000 444 123 GGT 	222 000 555 012 AGA 	222 111 111 345 TGC 	222 222 000 678 ACT 	222 222 445 890 TGT 	222 333 000 234 GAT 	222 333 222 012 TCC 	
B1 B2 B3 B4 B5 B6 B7 B8 B9 B10	1111 333 000 678 ACT 	1111 333 011 901 GCG A A	1111 333 345 CCC 	1111 333 555 456 CTC 	1111 3333 777 567 GCC 	1111 333 888 456 AAC 	1111 333 999 678 GGT 	111 555 000 789 GGA T 	111 555 555 567 AAA 	111 555 999 789 TTC 	1111 666 000 012 CAT 	1111 666 000 345 GTG 	1111 666 333 345 CCA 	1111 666 666 678 TTA 	1111 777 0000 234 TCA 	1111 777 222 678 TCA 	111 888 444 012 AAT 	1111 888 5555 234 TGC 	111 888 555 567 GCG 	111 888 556 890 AGG 	111 999 111 234 GAA 	111 999 222 123 CAC 	111 999 222 789 AGG 	111 999 333 345 AAG 	111 999 777 234 TCT 	111 999 999 678 AGG 	222 000 444 123 GGT 	222 000 555 012 AGA 	222 111 345 TGC 	222 222 000 678 ACT 	222 222 445 890 TGT 	222 333 000 234 GAT 	222 333 222 012 TCC 	
B1 B2 B3 B4 B5 B6 B7 B8 B9 B10 B11	1111 333 000 678 ACT 	1111 333 011 901 GCG A A A	1111 333 345 CCC 	111 333 555 456 CTC 	1111 3333 777 567 GCC 	1111 333 888 456 AAC 	1111 3333 999 678 GGT 	111 555 000 789 GGA T 	111 555 555 567 AAA 	111 555 999 789 TTC 	1111 666 000 012 CAT 	111 666 000 345 GTG 	111 666 333 345 CCA 	111 666 678 TTA 	111 777 000 234 TCA 	1111 777 222 678 TCA 	111 888 444 012 AAT 	111 888 555 234 TGC 	1111 8888 5555 567 GCG 	111 888 556 890 AGG 	111 999 111 234 GAA 	111 999 222 123 CAC 	111 999 222 789 AGG 	111 999 333 345 AAG 	1111 999 777 234 TCT 	111 999 999 678 AGG 	222 000 444 123 GGT 	222 000 555 012 AGA 	222 111 345 TGC 	222 222 000 678 ACT 	222 222 445 890 TGT 	222 333 000 234 GAT ????	222 333 222 TCC ???	
B1 B2 B3 B4 B5 B6 B7 B8 B9 B10 B11 B12	1111 333 000 678 ACT 	1111 333 011 901 GCG A A A A	1111 3333 3333 345 CCCC 	1111 3333 5555 456 CTC 	1111 3333 777 567 GCC 	1111 3333 8888 456 AAC 	1111 333 999 678 GGT 	1111 5555 0000 789 GGA T 	1111 5555 5555 567 AAA 	1111 5555 9999 789 TTC 	1111 6666 0000 012 CAT 	1111 6666 0000 345 GTG 	1111 6666 3333 345 CCA 	111 666 678 TTA 	1111 777 0000 234 TCA 	1111 7777 2222 678 TCA 	1111 8888 444 012 AAT 	1111 8888 5555 234 TGC 	1111 8888 5555 567 GCG 	1111 8888 5556 890 AGG 	1111 9999 1111 234 GAA 	111 999 222 123 CAC 	111 999 222 789 AGG 	1111 9999 3333 345 AAG 	1111 9999 7777 234 TCT 	1111 9999 9999 678 AGG 	222 000 444 123 GGT 	2222 0000 5555 012 AGA 	2222 1111 1111 345 TGC 	222 222 000 678 ACT 	222 222 445 890 TGT 	2222 333 0000 234 GAT ????	222 333 222 012 TCC ????	
B1 B2 B3 B4 B5 B6 B7 B8 B9 B10 B11 B12 cattle_hap1	1111 3333 000 678 ACT 	1111 3333 011 901 GCG A A A A A	1111 3333 3333 345 CCC 	1111 3333 5555 456 CTC 	1111 3333 7777 5667 GCC 	1111 3333 888 456 AAC 	1111 333 999 678 GGT 	1111 5555 0000 789 GGA T	1111 5555 5555 567 AAA 	1111 5555 9999 789 TTC 	1111 6666 0000 012 CAT G	1111 6666 0000 345 GTG 	1111 6666 3333 345 CCA 	1111 6666 678 TTA 	1111 777 0000 2334 TCA 	1111 7777 2222 678 TCA 	1111 8888 444 012 AAT 	1111 8888 5555 234 TGC 	1111 8888 5555 567 GCG 	1111 888 5556 890 AGG 	1111 9999 1111 234 GAA 	1111 9999 2222 123 CAC 	1111 9999 2222 789 AGG 	1111 9999 3333 345 AAG 	1111 9999 7777 2344 TCT 	1111 9999 9999 678 AGG 	2222 0000 444 123 GGT 	2222 0000 5555 012 AGA 	2222 1111 1111 345 TGC 	2222 2222 0000 678 ACT 	2222 2222 445 890 TGT 	2222 333 0000 234 GAT ????	2222 3333 2222 012 TCC 	
B1 B2 B3 B4 B5 B6 B7 B8 B9 B10 B11 B12 cattle_hap1 cattle_hap2	1111 3333 0000 678 ACT 	1111 3333 011 901 GCG A A A A A	1111 3333 345 CCCC 	1111 3333 5555 456 CTC 	1111 3333 7777 5667 GCC 	1111 3333 888 456 AAC 	1111 3333 9999 678 GGT 	1111 5555 0000 789 GGA T 	1111 5555 5555 567 AAA 	1111 5555 9999 TTC 	1111 6666 0000 012 CAT G G	1111 6666 0000 3455 GTG 	1111 6666 3333 345 CCA 	1111 6666 6666 678 TTA 	1111 777 0000 2334 TCA 	1111 7777 2222 678 TCA 	1111 8888 444 012 AAT 	1111 8888 5555 234 TGC 	1111 8888 5555 567 GCG 	1111 8888 5556 8900 AGG 	1111 9999 1111 234 GAA GAA 	1111 9999 2222 123 CAC 	1111 9999 2222 7899 AGG 	1111 9999 3333 345 AAG 	1111 9999 7777 234 TCT 	1111 9999 678 AGG 	2222 0000 4444 1233 GGT 	2222 0000 5555 0122 AGA G G	2222 1111 345 TGC 	2222 2222 0000 678 ACT 	2222 2222 445 890 TGT 	2222 333 0000 234 GAT ????	2222 3333 2222 012 TCC ???? 	
B1 B2 B3 B4 B5 B6 B7 B8 B9 B10 B11 B12 cattle_hap1 cattle_hap2 yak_hap1	1111 3333 0000 678 ACT 	1111 3333 0111 901 GCG A A A A A A	1111 3333 345 CCCC 	1111 3333 5555 4566 CTC 	1111 3333 7777 5677 GCC 	1111 3333 8888 456 AAC 	1111 3333 9999 678 GGT 	1111 5555 0000 789 GGA 	1111 5555 5555 567 AAA 	1111 5555 9999 TTC 	1111 6666 000 012 CAT G.T. G G C	1111 6666 0000 345 GTG 	1111 6666 3333 345 CCA G G C.G	1111 6666 678 TTA 	1111 777 000 234 TCA 	1111 7777 2222 6788 TCA 	1111 8888 444 012 AAT 	1111 8888 5555 2344 TGC 	1111 888 5555 567 GCG 	1111 888 556 890 AGG 	1111 9999 1111 234 GAA GG	1111 9999 2222 123 CAC 	1111 9999 2222 789 AGG 	1111 9999 3333 345 AAG 	1111 9999 7777 2344 TCT 	1111 9999 678 AGG 	2222 0000 444 1233 GGT 	2222 0000 5555 012 AGA G G 	2222 1111 345 TGC 	2222 2000 6788 ACT 	2222 2222 445 890 TGT 	2222 333 0000 234 GAT ????	2222 3333 2222 012 TCC ???? 	
B1 B2 B3 B4 B5 B6 B7 B8 B9 B10 B11 B12 cattle_hap1 cattle_hap1 bison_hap1 bison_hap1	1111 3333 000 678 ACT 	1111 3333 0111 901 GCG A A A A A A A	1111 3333 345 CCCC 	1111 3333 5555 4566 CTC 	1111 3333 7777 5677 GCC 	1111 3333 8888 456 AAC 	1111 3333 9999 678 GGT 	1111 5555 0000 789 GGA 	1111 5555 5555 567 AAA 	1111 5555 9999 TTC 	1111 6666 000 012 CAT G.T. G G G.	1111 6666 0000 3455 GTG 	1111 6666 3333 345 CCA 	1111 6666 678 TTA 	1111 777 000 234 TCA 	1111 7777 2222 6788 TCA 	1111 8888 4444 012 AAT 	1111 888 5555 234 TGC 	1111 8888 5555 567 GCG 	1111 8888 5566 8900 AGG 	1111 9999 1111 234 GAA ????	1111 9999 2222 123 CAC ????	1111 9999 2222 789 AGG ????	1111 9999 3333 345 AAG 	1111 9999 7777 2344 TCT 	1111 9999 6788 AGG 	2222 0000 444 1233 GGT 	2222 0000 5555 0122 AGA G G G	2222 1111 345 TGC 	2222 2000 6788 ACT 	2222 2222 445 890 TGT 	2222 333 0000 234 GAT ???? ????	222 333 222 012 TCC ???? 	
B1 B2 B3 B4 B5 B6 B7 B8 B9 B10 B11 B12 cattle_hap1 cattle_hap1 vak_hap1 bison_hap1 sheep_hap2	1111 3333 000 678 ACT 	1111 333 011 901 GCG A A A A A	1111 3333 345 CCCC 	1111 3333 5555 4566 CTCC 	1111 333 777 567 GCC T T.	1111 3333 8888 456 AAC T T.	1111 3333 9999 678 GGT 	1111 555 000 789 GGA 	1111 5555 5555 5677 AAA 	1111 5555 9999 TTC 	1111 6666 0000 012 CAT G.T G.T G.T G.T G.T G.	1111 6666 0000 3455 GTG 	1111 6666 3333 345 CCA 	1111 666 678 TTA 	1111 777 000 234 TCA 	1111 7777 2222 678 TCA 	1111 888 444 012 AAT 	1111 888 5555 234 TGC 	1111 888 5555 567 GCG 	1111 888 556 890 AGG 	1111 9999 1111 234 GAA ????	1111 9999 2222 1233 CAC ????	1111 9999 2222 789 AGG ????	1111 9999 3333 345 AAG 	1111 9999 7777 234 TCT 	1111 9999 9099 6788 AGG 	2222 000 444 123 GGT 	2222 0000 5555 012 AGA G G G	2222 1111 345 TGC 	2222 2000 678 ACT 	2222 2222 445 890 TGT 	2222 333 000 234 GAT ???? ????	222 333 222 012 TCC ???? 	
B1 B2 B3 B4 B5 B6 B7 B8 B9 B10 B11 B12 cattle_hap1 cattle_hap1 cattle_hap1 sheep_hap1 sheep_hap3	1111 3333 0000 678 ACT 	1111 333 011 901 GCG A A A A A	1111 3333 345 CCCC 	1111 333 555 456 CTC 	1111 333 777 567 GCC T T.	1111 3333 8888 456 AAC T T.	1111 3333 9999 678 GGT 	1111 555 000 789 GGA 	1111 5555 5555 567 AAA 	1111 5555 9999 TTC 	1111 6666 0000 012 CAT G G G G G	1111 6666 0000 345 GTG 	1111 6666 3333 345 CCA G G G G G G	1111 666 678 TTA 	1111 777 000 234 TCA 	1111 7777 2222 678 TCA 	1111 888 444 012 AAT 	1111 888 5555 234 TGC 	1111 888 555 567 GCG 	1111 888 556 890 AGG 	1111 9999 1111 234 GAA ???? 	1111 9999 2222 123 CAC ????	1111 9999 2222 789 AGG ????	1111 9999 333 345 AAG 	1111 9999 7777 234 TCT 	1111 9999 9799 678 AGG 	2222 000 444 123 GGT 	2222 0000 5555 012 AGA G G G G G G	222 111 111 345 TGC 	2222 2000 678 ACT 	2222 2222 445 890 TGT 	2222 333 000 234 GAT ???? 	222 333 222 012 TCC ???? ????	
B1 B2 B3 B4 B5 B6 B7 B8 B9 B10 B11 B12 cattle_hap1 cattle_hap1 cattle_hap1 bison_hap1 sheep_hap3 sheep_hap3 sheap_hap4	1111 3333 0000 678 ACT 	1111 333 011 901 GCG A A A A A A	1111 333 333 345 CCC 	1111 333 555 456 CTC 	1111 333 777 567 GCC T T	1111 333 888 456 AAC T T T	1111 3333 9999 678 GGT 	1111 5555 0000 789 GGA 	1111 5555 567 AAAA 	1111 5555 9999 789 TTC 	1111 6666 0000 012 CAT G G G G G	1111 6666 0000 345 GTG 	1111 6666 3333 345 CCA G G G G G G G G	1111 666 678 TTA 	1111 777 000 234 TCA 	1111 7777 2222 678 TCA 	1111 888 444 012 AAT 	1111 888 555 234 TGC 	1111 888 555 567 GCG 	1111 8888 5566 8900 AGG 	1111 9999 1111 234 GAA ??? 	1111 9999 2222 123 CAC ????	1111 9999 2222 789 AGG ????	1111 9999 3333 345 AAG 	1111 9999 7777 234 TCT 	1111 9999 9799 678 AGG 	2222 0000 444 123 GGT 	2222 0000 5555 012 AGA G G G G G G G G	2222 1111 1111 345 TGC 	2222 2222 0000 678 ACT 	2222 2222 445 890 TGT 	2222 333 000 234 GAT ???? 	222 333 222 012 TCC ???? .G. .G. .G. 	
B1 B2 B3 B4 B5 B6 B7 B8 B9 B10 B11 B12 cattle_hap2 yak_hap1 bison_hap1 sheep_hap3 sheep_hap3 sheep_hap4 shap1	1111 333 000 678 ACT 	1111 333 011 901 GCG A A A A	1111 333 333 345 CCC 	1111 333 555 456 CTC 	1111 333 777 567 GCC T T	1111 333 888 456 AAC T T T	1111 3333 9999 678 GGT 	1111 5555 0000 789 GGA 	1111 5555 5555 567 AAA 	1111 5555 9999 TTC 	1111 6666 0000 012 CAT G G G G G	1111 6666 0000 345 GTG 	1111 6666 3333 345 CCA 	1111 666 678 TTA 	1111 777 0000 234 TCA 	1111 7777 2222 678 TCA 	1111 888 444 012 AAT 	1111 888 5555 234 TGC 	1111 8888 5555 5677 GCG 	1111 8888 5566 8900 AGG 	1111 9999 1111 234 GAA ???? 	1111 9999 2222 123 CAC ????	1111 9999 222 789 AGG ????	1111 9999 3333 345 AAG 	1111 9999 7777 234 TCT 	1111 9999 678 AGG 	2222 0000 444 123 GGT 	2222 0000 5555 012 AGA G G G G G G G 	2222 1111 1111 345 TGC 	2222 2222 0000 678 ACT 	2222 2222 445 890 TGT 	2222 333 0000 234 GAT ???? ????	222 333 222 012 TCC ???? 	
B1 B2 B3 B4 B5 B6 B7 B8 B9 B10 B11 B12 cattle_hap1 cattle_hap2 yak_hap1 bison_hap1 sheep_hap3 sheep_hap3 sheep_hap4 goats_hap2	1111 333 000 678 ACT 	1111 333 011 901 GCG A 	1111 3333 3333 345 CCC 	1111 333 555 456 CTC 	1111 3333 7777 567 GCC T T 	1111 333 888 456 AAC T T T	1111 333 999 678 GGT 	1111 5555 0000 789 GGA T T	1111 5555 5555 567 AAA 	1111 5555 9999 789 TTC 	1111 6666 0000 0122 CAT G G G G	1111 6666 0000 345 GTG 	1111 6666 3333 345 CCA 	1111 666 666 678 TTA 	1111 777 0000 234 TCA 	1111 7777 2222 6788 TCA 	1111 8888 444 012 AAT 	1111 8888 5555 2344 TGC 	1111 8888 5555 5677 GCG 	1111 8888 5566 8900 AGG 	1111 9999 1111 234 GAA ????	1111 9999 2222 1233 CAC ???? 	1111 9999 222 7899 AGG ????	1111 9999 3333 345 AAG 	1111 9999 7777 234 TCT 	1111 9999 678 AGG 	2222 0000 444 123 GGT GG G G G G G 	2222 0000 5555 012 AGA G G G G G G G 	2222 1111 345 TGC 	2222 2222 0000 6788 ACT 	2222 2222 445 890 TGT 	2222 333 0000 234 GAT ???? ???? 	2222 3333 2222 012 TCC ???? 	
B1 B2 B3 B4 B5 B6 B7 B8 B9 B10 B11 B12 cattle_hap1 cattle_hap1 cattle_hap2 yak_hap1 bison_hap1 sheep_hap3 sheep_hap3 sheep_hap4 goats_hap2 goats_hap3	1111 333 000 678 ACT 	1111 333 011 901 GCG A A 	1111 333 333 345 CCC 	1111 333 555 456 CTC 	1111 3333 7777 567 GCC T 	1111 333 888 456 AAC T T	1111 333 999 678 GGT 	1111 5555 0000 789 GGA T 	1111 5555 5555 567 AAA 	1111 5555 9999 7899 TTC 	1111 6666 0000 012 CAT G G G G G	1111 6666 0000 345 GTG 	1111 6666 3333 345 CCA 	1111 6666 6666 678 TTA 	1111 777 000 234 TCA 	1111 7777 2222 6788 TCA 	1111 8888 444 012 AAT 	1111 8888 5555 2344 TGC 	1111 8888 5555 5667 GCG 	1111 8888 5566 8900 AGG 	1111 9999 1111 234 GAA ???? 	1111 9999 2222 1233 CAC ???? 	1111 9999 2222 789 AGG ???? 	1111 9999 3333 345 AAG 	1111 9999 7777 234 TCT 	1111 9999 678 AGG 	2222 0000 444 123 GGT 	2222 0000 5555 012 AGA G G G G G G G G G 	2222 1111 1111 345 TGC 	2222 2222 0000 6788 ACT 	222 222 445 890 TGT 	2222 333 0000 234 GAT ???? ???? 	2222 3333 2222 012 TCC ???? 	
B1 B2 B3 B4 B5 B6 B7 B8 B9 B10 B11 B12 cattle_hap1 cattle_hap2 yak_hap1 bison_hap1 sheep_hap1 sheep_hap3 sheep_hap4 goats_hap1 goats_hap3 chiru hap1	1111 333 000 678 ACT 	1111 333 011 901 GCG A A A A A A A	1111 333 333 345 CCC 	1111 333 555 456 CTC 	1111 333 777 567 GCC 	1111 333 888 456 AAC T T T	1111 333 999 678 GGT 	1111 5555 0000 789 GGA 	1111 5555 5555 567 AAA 	1111 5555 9999 7899 TTC 	1111 6666 0000 012 CAT G G G G G	1111 6666 0000 3455 GTG 	1111 6666 3333 345 CCA 	1111 6666 678 TTA 	1111 7777 0000 234 TCA 	1111 7777 2222 6788 TCA 	1111 8888 4444 012 AAT 	1111 8888 5555 234 TGC 	1111 8888 5555 5667 GCG 	1111 8888 5566 8900 AGG 	1111 9999 1111 234 GAA ???? 	1111 9999 2222 1233 CAC ???? 	1111 9999 2222 7899 AGG ???? 	1111 9999 3333 345 AAG 	1111 9999 7777 234 TCT 	1111 9999 6788 AGG 	2222 0000 444 123 GGT 	2222 0000 5555 012 AGA G G G G G 	2222 1111 1111 345 TGC 	2222 2222 0000 6788 ACT 	222 222 445 890 TGT 	2222 333 0000 234 GAT ???? ???? 	2222 3333 2222 012 TCC ???? 	
B1 B2 B3 B4 B5 B6 B7 B8 B9 B10 B11 B12 cattle_hap2 yak_hap1 bison_hap1 sheep_hap3 sheep_hap3 sheep_hap3 sheep_hap3 goats_hap1 goats_hap1 goats_hap1 chiru_hap2	1111 333 000 678 ACT 	1111 333 011 901 GCG A A 	1111 333 333 345 CCC 	1111 333 555 456 CTC 	1111 3333 7777 5667 GCC T. T. T. T. T.	1111 333 888 456 AAC T T T	1111 333 999 678 GGT 	1111 5555 0000 789 GGA 	1111 5555 5555 567 AAA 	1111 5555 9999 7899 TTC 	1111 6666 0000 012 CAT G G G G G	1111 6666 0000 3455 GTG 	1111 6666 3333 345 CCA 	1111 6666 678 TTA 	1111 7777 0000 234 TCA 	1111 7777 2222 6788 TCA 	1111 8888 4444 012 AAT 	1111 8888 5555 234 TGC 	1111 8888 5555 5677 GCG 	1111 8888 5566 8900 AGG 	1111 9999 1111 234 GAA ???? 	1111 9999 2222 1233 CAC ???? 	1111 9999 2222 7899 AGG ???? 	1111 9999 3333 345 AAG 	1111 9999 7777 234 TCT 	1111 9999 6788 AGG 	2222 0000 444 123 GGT G G G G G G 	2222 0000 5555 012 AGA G G G G G G G G G G 	2222 1111 1111 345 TGC 	2222 2222 0000 678 ACT 	2222 2222 445 890 TGT 	2222 333 0000 234 GAT ???? ????	2222 3333 2222 012 TCC ???? 	

Figure 5. Nucleotide differences of the haplotype sequences among some species of Bovidae. Number represents the position of coding region. Dots (.) denote identity with the B1. Nucleotide substitutions are denoted by different letters. Missing information is denoted by a question mark (?). Horizontal line (–) represents the deletion in the sequences. These definitions apply to the following figure as well.

Table 4. Functional effect of non-synonymous substitution on buffalo PPARGC1A.

SNP	Substitution	Preservation time	Message
c.419C>T	p. Ser140Phe	456	Probably damaging
c.920C>A	p. Pro307His	456	Probably damaging
c.926G>A	p. Arg309Lys	361	Possibly damaging

muscle, heart, liver, small intestine, mammary gland, rumen, spleen and lung between the two periods were significantly different, indicating that the expression of this gene in these tissues was regulated by physiological state. In the present study, the expression of the *PPARGC1A* gene was detected in the mammary gland of lactating buffalo, and it is speculated that this gene also plays a role in milk fat synthesis of buffalo. It is noteworthy that buffalo *PPARGC1A* was expressed in the mammary gland with a higher expression level during the dry-off period than that during peak lactation. This may be due to the fact that the *PPARGC1A* gene also participates

L. Qiu et al.: Characterization, expression and polymorphisms of the buffalo PPARGC1A gene

	1123333333334455666677
	4440022238991313124607
	50207967968289094405654
B1	MSLAPRSKACIMTSPKHNRKRCS
B2	
B3	
B4	
B5	
B6	
B7	
B8	
B9	
B10	
B11	HK?
B12	. F
Cattle_hap1	C
Cattle_hap2	VT.PDC
Yak_hap1	T. P D. – C
Bison_hap1	
Sheep_hap1	A G. TIPL. D S
Sheep_hap2	AG. TIPL. DS
Sheep_hap3	VAG. TIPL. DS
Sheep_hap4	. F A G. TIPL. D S
Goats_hap1	?AG.TIP.RDSS
Goats_hap2	AG. TIPDSS
Goats_hap3	L AE. G. TIP DS SY.
Chiru_hap1	A. TGVTIPD
Chiru_hap2	DA. TGVTIPD
Chiru_hap3	. F A. TGVTIP D

Figure 6. Differences of amino acid sequences corresponding to the haplotypes of *PPARGC1A* in the species of Bovidae.

Table 5. Haplotype information on the buffalo PPARGC1A gene.

	Haplotype	Actual frequency	Expected frequency
B1	ACGGA	0.143	0.196
B2	ACGAA	0.310	0.267
B3	ACAGA	0.095	0.054
B4	ATGGA	0.071	0.095
B5	ATGAA	0.119	0.095
B6	ACGGT	0.024	0.011
B7	ACAAA	0.024	0.055
B8	ATAGA	0.048	0.024
B9	GCGGA	0.071	0.062
B10	GCGAA	0.095	0.094

in the degradation and remodeling of the mammary gland tissue in the non-lactating buffaloes. It has been reported that there are quantitative trait loci influencing milk traits in dairy cows on BTA 6, and the *PPARGC1A* gene is located in the middle of BTA 6 (Khatib et al., 2007). Therefore, this gene is also qualified as a key functional candidate gene influencing milk production traits (Weikard et al., 2005). In addition, by constructing the network of genes involved in milk fat synthesis, it is revealed that bovine PPARGC1A plays a pivotal regulatory function in the network of milk fat synthesis in the mammary gland (Bionaz and Loor, 2008).



Figure 7. Phylogenetic tree constructed by using maximum likelihood method (Hasegawa-Kishino-Yano model). Bootstrap confidences which are adjacent to nodes are based on 10 000 replicates.

So far, multiple polymorphic sites have been found in the cattle PPARGC1A gene. In dairy cows, there were significant association between c.1790+514G>A, c.1892+19T>C and c.1892+19G>A of PPAGC1A gene and milk fat yield (Weikard et al., 2005; Schennink et al., 2009). However, there are few reports about the SNPs of the PPARGC1A gene in buffalo. In this work, a total of eight SNPs were found in buffaloes, of which three were non-synonymous. And it was predicted that two non-synonymous substitutions (c.419C>T and c.920C>A) led to changes in amino acids of p.Ser140Phe and p.Pro307His, which seriously affected the function of buffalo PPARGC1A. The Ser is a polar AA, while the Phe is a hydrophobic AA. The Pro is a non-polar AA, while the His is a basic AA. These substitutions belong to amino acid substitutions with different physicochemical properties. In addition, the 140Ser is a potential O-glycosylation site and this substitution may lead to the loss of an O-glycosylation site of buffalo PPARGC1A. These may cause changes in the structure or function of buffalo PPARGC1A, while SNPs, which cause synonymous changes, are thought to act sometimes by altering translation efficiency and thus can influence traits (Zhou et al., 2018). Whether the SNPs identified in this study, especially the nonsynonymous SNPs, have any effect on the function of buffalo PPARGC1A and the lactation traits of buffalo needs to be verified by further expanding the sample size.

The alignment indicated that the CDS length of buffalo *PPARGC1A* was 3 bp longer than that of *Bos* but the same as that of *Ovis*. In addition, there were three amino acid differential sites in the PPARGC1A between buffalo and other species of Bovidae. It is speculated that these differences in PPARGC1A may lead to functional differences between buffalo and other species of Bovidae. The phylogenetic tree showed that buffalo had a closer genetic relationship with the species of *Bos*. This indicates that there may be little difference in the function of PPARGC1A between buffalo and the species of *Bos*. In addition, there were 13 and 9 differential nucleotides in the *PPARGC1A* gene between buffalo and *Bos* and between buffalo and *Ovis*, respectively, which can be used as molecular markers to distinguish buffalo from other species of Bovidae.

5 Conclusions

The length of PPARGC1A CDS for both types of buffalo was the same, which encoded a peptide composed of 797 amino acid residues with the same physicochemical properties and molecular functions. Buffalo PPARGC1A contains one RRM_PPARGC1A domain without a signal peptide or a transmembrane domain and is an inducible transcriptional coactivator related to the regulation of carbohydrate and lipid metabolism. It can function in a variety of tissues and performs a critical function in the milk fat synthesis of the mammary gland. Eight SNPs were found in two types of buffalo. However, only the p.Ser140Phe and p.Pro307His caused by c.419C>T and c.920C>A may affect the function of buffalo PPARGC1A. The mechanism of buffalo PPARGC1A on milk traits is still unknown and needs to be further analyzed. This work will lay a preliminary foundation for further understanding the structure and function of the buffalo PPARGC1A.

Data availability. The original data used in this study are available from the corresponding author upon request.

Supplement. The supplement related to this article is available online at: https://doi.org/10.5194/aab-63-249-2020-supplement.

Author contributions. YM conceived and designed the research. LQ and XF performed the material preparation and experiments. LQ, YZ, and XT performed the data collection and analysis. LQ, XF and YM drafted the manuscript. All authors read and approved the final manuscript. **Competing interests.** The authors declare that they have no conflict of interest.

Financial support. This research has been supported by the National Natural Science Foundation of China (grant nos. 31760659 and 31460582) and the Natural Science Foundation Key Project of Yunnan Province, China (grant nos. 2014FA032 and 2007C0003Z).

Review statement. This paper was edited by Steffen Maak and reviewed by two anonymous referees.

References

- Almagro-Armenteros, J. J., Tsirigos, K. D., Sønderby, C. K., Petersen, T. N., Winther, O., Brunak, S., von Heijne, G., and Nielsen, H.: SignalP 5.0 improves signal peptide predictions using deep neural networks, Nat. Biotechnol., 37, 420–423, https://doi.org/10.1038/s41587-019-0036-z, 2019.
- Bionaz, M. and Loor, J. J.: Gene networks driving bovine milk fat synthesis during the lactation cycle, BMC Genomics, 9, 366, https://doi.org/10.1186/1471-2164-9-366, 2008.
- Chen, Z., Luo, J., Sun, S., Cao, D., Shi, H., and Loor, J. J.: miR-148a and miR-17-5p synergistically regulate milk TAG synthesis via PPARGC1A and PPARA in goat mammary epithelial cells, RNA Biol., 14, 326–338, https://doi.org/10.1080/15476286.2016.1276149, 2017.
- D'Ambrosio, C., Arena, S., Salzano, A. M., Renzone, G., Ledda, L., and Scaloni, A.: A proteomic characterization of water buffalo milk fractions describing PTM of major species and the identification of minor components involved in nutrient delivery and defense against pathogens, Proteomics, 8, 3657–3666, https://doi.org/10.1002/pmic.200701148, 2008.
- Dominy, J. E., Lee, Y., Gerhart-Hines, Z., and Puigserver, P.: Nutrient-dependent regulation of PGC-1 α 's acetylation state and metabolic function through the enzymatic activities of Sirt1/GCN5, Biochim. Biophys. Acta, 1804, 1676–1683, https://doi.org/10.1016/j.bbapap.2009.11.023, 2010.
- Esterbauer, H., Oberkofler, H., Krempler, F., and Patsch, W.: Human peroxisome proliferator activated receptor gamma coactivator 1 (PPARGC1) gene: cDNA sequence, genomic organization, chromosomal localization, and tissue expression, Genomics, 62, 98–102, https://doi.org/10.1006/geno.1999.5977, 1999.
- Khatib, H., Zaitoun, I., Wiebelhaus-Finger, J., Chang, Y. M., and Rosa, G. J. M.: The association of bovine PPARGC1A and OPN genes with milk composition in two independent Holstein cattle populations, J. Dairy Sci., 90, 2966–2970, https://doi.org/10.3168/jds.2006-812, 2007.
- Knutti, D., Kaul, A., and Kralli, A.: A tissue-specific coactivator of steroid receptors, identified in a functional genetic screen, Mol. Cell Biol., 20, 2411–2422, https://doi.org/10.1128/MCB.20.7.2411-2422.2000, 2000.
- Lalitha, S.: Primer Premier 5, Biotech Software & Internet Report, 1, 270–272, https://doi.org/10.1089/152791600459894, 2000.
- Mi, H., Huang, X., Muruganujan, A., Tang, H., Mills, C., Kang, D., and Thomas, P. D.: PANTHER version 11: expanded annotation data from Gene Ontology and Reactome pathways, and

data analysis tool enhancements, Nucleic Acids Res., 45, D183–D189, https://doi.org/10.1093/nar/gkw1138, 2017.

- Mortensen, O. H., Frandsen, L., Schjerling, P., Nishimura, E., and Grunnet, N.: PGC-1 α and PGC-1 β have both similar and distinct effects on myofiber switching toward an oxidative phenotype, Am. J. Physiol.-Endoc. M., 291, E807–E816, https://doi.org/10.1152/ajpendo.00591.2005, 2006.
- Oberkofler, H., Esterbauer, H., Linnemayr, V., Strosberg, A. D., Krempler, F., and Patsch, W.: Peroxisome proliferator-activated receptor (PPAR) gamma coactivator-1 recruitment regulates PPAR subtype specificity, J. Biol. Chem., 277, 16750–16757, https://doi.org/10.1074/jbc.m200475200, 2002.
- Pasandideh, M., Mohammadabadi, M. R., Esmailizadeh, A. K., and Tarang, A.: Association of bovine *PPARGC1A* and *OPN* genes with milk production and composition in Holstein cattle, Czech J. Anim. Sci., 60, 97–104, https://doi.org/10.17221/8074-CJAS, 2015.
- Puigserver, P., Wu, Z., Park, C. W., Graves, R., Wright, M., and Spiegelman, B. M.: A cold-inducible coactivator of nuclear receptors linked to adaptive thermogenesis, Cell, 92, 829–839, https://doi.org/10.1016/S0092-8674(00)81410-5, 1998.
- Ramayo-Caldas, Y., Fortes, M. R. S., Hudson, N. J., Porto-Neto, L. R., Bolormaa, S., Barendse, W., Kelly, M., Moore, S. S., Goddard, M. E., Lehnert, S. A., and Reverter, A.: A marker-derived gene network reveals the regulatory role of PPARGC1A, HNF4G, and FOXP3 in intramuscular fat deposition of beef cattle, J. Anim. Sci., 92, 2832–2845, https://doi.org/10.2527/jas.2013-7484, 2014.
- Sambrock, J. and Russell, D.: Molecular cloning: a laboratory manual, 3rd Edn., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 6.4–6.12, 2001.
- Schennink, A., Bovenhuis, H., Léon-Kloosterziel, K. M., Van-Arendonk, J. A. M., and Visker, M. H. P. W.: Effect of polymorphisms in the *FASN*, *OLR1*, *PPARGC1A*, *PRL* and *STAT5A* genes on bovine milk-fat composition, Anim. Genet., 40, 909– 916, https://doi.org/10.1111/j.1365-2052.2009.01940.x, 2009.

- Spiegelman, B. M., Puigserver, P., and Wu, Z.: Regulation of adipogenesis and energy balance by PPARg and PGC-1, Int. J. Obesity, 4, S8–S10, https://doi.org/10.1038/sj.ijo.0801492, 2000.
- Stephens, M., Smith, N., and Donnelly, P.: A new statistical method for haplotype reconstruction from population data, Am. J. Hum. Genet., 68, 978–989, https://doi.org/10.1086/319501, 2001.
- Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S.: MEGA6: molecular evolutionary genetics analysis version 6.0, Mol. Biol. Evol., 30, 2725–2729, https://doi.org/10.1093/molbev/mst197, 2013.
- Weikard, R., Kühn, C., Goldammer, T., Freyer, G., and Schwerin, M.: The bovine *PPARGC1A* gene: molecular characterization and association of an SNP with variation of milk fat synthesis, Physiol. Genomics, 21, 1–13, https://doi.org/10.1152/physiolgenomics.00103.2004, 2005.
- Yeh, F. C. and Boyle, T. B. J.: Population genetic analysis of codominant and dominant marker and quantitative traits, Belg. J. Bot., 129, 157–163, 1997.
- Yoon, J. C., Puigserver, P., Chen, G., Donovan, J., Wu, Z., Rhee, J., Adelmant, G., Stafford, J., Kahn, C. R., Granner, D. K., Newgard, C. B., and Spiegelman, B. M.: Control of hepatic gluconeogenesis through the transcriptional coactivator PGC-1, Nature, 413, 131–138, https://doi.org/10.1038/35093050, 2001.
- Zhou, Z. Y., Hu, Y., Li, A., Li, Y. J., Zhao, H., Wang, S. Q., Otecko, N. O., Zhang, D., Wang, J. H., Liu, Y., Irwin, D. M., Qin, Y., and Zhang, Y. P.: Genome wide analyses uncover allele-specific RNA editing in human and mouse, Nucleic Acids Res., 46, 8888–8897, https://doi.org/10.1093/nar/gky613, 2018.