Identification and characterization of ERV transcripts in goat embryos

Ruizhi Deng*, Chengquan Han*, Lu Zhao, Qing Zhang, Beifen Yan, Rui Cheng, Biao Wei, Peng Meng, Tingchao Mao, Yong Zhang and Jun Liu

College of Veterinary Medicine, Northwest A&F University, Key Laboratory of Animal Biotechnology of the Ministry of Agriculture, Yangling, Shaanxi, China

Correspondence should be addressed to Y Zhang or J Liu; Email: zhy1956@263.net or liujun2013@nwsuaf.edu.cn

*(R Deng and C Han contributed equally to this work)

Abstract

Endogenous retroviruses (ERVs), which are abundant in mammalian genomes, can modulate the expression of nearby genes, and their expression is dynamic and stage-specific during early embryonic development in mice and humans. However, the functions and mechanisms of ERV elements in regulating embryonic development remain unclear. Here, we utilized several methods to determine the contribution of ERVs to the makeup and regulation of transcripts during embryonic genome activation (EGA). We constructed an ERV library and embryo RNA-seq library (IVF_2c and IVF_8c) of goat to serve as our research basis. The GO and KEGG analysis of nearby ERV genes revealed that some ERV elements may be associated with embryonic development. RNA-seq results were consistent with the features of EGA. To obtain the transcripts derived from the ERV sequences, we blasted the ERV sequences with embryonic transcripts and identified three lncRNAs and one mRNA that were highly expressed in IVF-8c rather than in IVF-2c (q -value <0.05). Then, we validated the expression patterns of nine ERV-related transcripts during early developmental stages and knocked down three high-expression transcripts in EGA. The knockdown of lncRNA TCONS_00460156 or mRNA HSD17B11 significantly decreased the developmental rate of IVF embryos. Our findings suggested that some transcripts from ERVs are essential for the early embryonic development of goat, and analyzing the ERV expression profile during goat EGA may help elucidate the molecular mechanisms of ERV in regulating embryonic development.

Reproduction (2019) **157** 111–122

Background

Endogenous retroviruses (ERVs), which are also called long terminal repeat (LTR) retrotransposons, comprise approximately 10% of mammalian genomes (Rebollo *et al.* 2012). Although ERVs are considered parasitic elements that are harmful to hosts, ERVs play an important role in the regulation of the gene expression in mammalian genomes (Friedli & Trono 2015, Thompson *et al.* 2016).

To analyze the full transcriptome of repetitive elements in mouse oocytes and at early embryonic stages, researchers used high-throughput sequencing technology and found that the expression of LTR retrotransposons is abundant and stage specific during early development and that specific elements activated at the two-cell stage are essential for embryonic development (Fadloun *et al.* 2013). In the developmental process of human pre-implantation embryos, specific ERV families are transcribed at different developmental stages and regulate the expression levels of numerous functional genes (Goeke *et al.* 2015). Moreover,

numerous ERVs are activated during zygote genome activation in bovine species (Bui *et al.* 2009). Emerging evidence has suggested that ERVs may perform specific and important functions in regulating early embryonic development (Gifford *et al.* 2013).

Endogenous retroviral elements can act as alternative promoters of developmentally expressed genes to form chimeric transcripts, which modulate the expression of host genes during embryonic development (Peaston *et al.* 2004, Batut *et al.* 2013). The expression of murine endogenous retrovirus-like genes, which are essential for the development of mouse embryos, initiates many specific transcripts of two-cell embryos (Kigami *et al.* 2003, Macfarlan *et al.* 2012). Long noncoding RNA (lncRNA) HERVH-derived HPAT5 modulates the acquisition of pluripotency and the formation of inner cell mass in humans (Durruthy-Durruthy *et al.* 2016). ERV-associated lncRNA LincGET is crucial for the cleavage of mouse embryos (Wang *et al.* 2016).

Despite the discovery of numerous transcripts of ERV elements in the early embryos of mammals, the functions and mechanisms of ERV components and their

transcripts in regulating embryonic development have yet to be clarified. In this study, we systematically analyzed and annotated all sequences of ERV elements in goat genomes. We also obtained the ERV-derived transcripts by blasting the early embryonic transcriptome with all ERV sequences in goat genomes. We identified the ERV-derived transcripts that were highly expressed during embryonic genome activation (EGA) and determined several transcripts essential for early embryonic development in goats.

Materials and methods

Ethics statement

This study was performed in strict accordance with the Guidelines for the Care and Use of Animals of Northwest A&F University. All experiments were approved by the Care and Use of Animals Center, Northwest A&F University. All precautions were undertaken to minimize animal suffering.

ERV gathering

Two pipelines, namely, *de novo* prediction and homolog prediction, were established to construct the ERV library, and goat genome data were downloaded from NCBI (NCBI accession number AJPT00000000) as the input data (Fig. 1) (Dong *et al.* 2013). The first pipeline utilized MGEScan-LTR with the default parameters (Rho *et al.* 2007). The second pipeline employed LTRharvest to identify the pairs of putative LTRs that were flanked by target site duplications (TSD)

(Ellinghaus *et al.* 2008). The default settings of the parameters were applied. LTRdigest was used to predict the sequence features of LTR retrotransposon candidates, including primer-binding sites (PBS), protein domains and polypurine tracts (PPT) (Steinbiss *et al.* 2009, Eddy 2011). Both LTRharvest and LTRdigest are based on HMMER (Ellinghaus *et al.* 2008, Steinbiss *et al.* 2009, Eddy 2011). Given that ERV sequences may present as false positives, two key filters were applied to obtain high-confidence full-length ERVs. First, the candidates were removed if the reverse transcriptase (RT) domain was identified to be highly likely non-LTR retrotransposons by searching all repeat libraries in Repbase (Rhead *et al.* 2010). Second, the candidates with less than three standard retroviral protein domains (RH, PR, IN, Gag and RT) were discarded. The outputs of LTRdigest and MGEScan-LTR were inputted to CENSOR for the integrated recognition of candidate ERVs that contained any known repetitive elements (Kohany *et al.* 2006, Rho *et al.* 2007, Ellinghaus *et al.* 2008).

ERV classification and phylogenetic analysis

MUSCLE (Edgar 2004) was used to build multiple alignments of the RT domain from 645 full-length goat ERVs that identified 26 known viruses of the domain. Poor-quality sequences with $e < 10^{-5}$ and alignment length < 100 bp were manually removed. MEGA6 (Tamura *et al.* 2013) was used to build a neighbor-joining phylogeny that reflected the aforementioned 26 viruses of the RT domain. The parameters were set to 1000 bootstrap replicates, Jones–Taylor–Thorton was implemented as the amino acid substitution model, and the pairwise deletion option was applied. Then, Vmatch was utilized to classify

Figure 1 Overview of ERV-derived transcripts pipeline. Left: identification pipeline for potentially complete ERVs. Using LTRharvest, we identified 64,159 unique LTR candidates within LTR and TSD and annotated the candidates. In addition, we used the other pipeline MEGScan-LTR to identify known complete ERVs. By combining these two pipelines, we obtained 645 potential complete ERVs. All ERVs were classified into 18 families on the basis of the polymorphisms of the RT genes and the similarity of the exogenous retrovirus. Right: lncRNA and mRNA identification pipeline for RNA-Seq. Bottom, ERV-derived transcripts. We used Blastall to blast goat lncRNA and mRNA with potentially complete ERVs.

incomplete RT-domain GERVs on the basis of the similarity of the ERV domain.

RepeatMasker analysis

To generate a systematic annotation of ERVs in the goat genome, the output data of LTRharvest and MGEScan-LTR were compiled to form the goat ERV library (Rho *et al.* 2007, Ellinghaus *et al.* 2008). RepeatMasker (Tarailo-Graovac & Chen 2009) with a library of non-ERV repetitive elements from Repbase (Rhead *et al.* 2010) was used to remove all of the repetitive elements in the ERV library. Then, a custom ERV library that combined our 645 ERV entries and the Repbase ERV library (Rhead *et al.* 2010) was established, and RepeatMasker (Tarailo-Graovac & Chen 2009) was operated on the goat genome with the default parameters.

Gene ontology and pathway analyses

To systematically analyze the function of the goat ERV elements, the goat ERV library contents were uploaded on the Database for Annotation, Visualization and Integrated Discovery (DAVID) for gene ontology (GO) (Huang *et al.* 2009) and pathway (KEGG pathways) analyses (Kanehisa *et al.* 2008). GO terms and pathways with corrected $P < 0.05$ were considered significantly enriched.

RNA-seq and analysis

At present, a strand-specific library could not be generated for the analysis of goat embryo lncRNA because of the lack of available specific rRNA removal kits for goats. In this study, RNA-seq technology was used to analyze early embryonic transcripts, enrich the involved transcripts, and construct a common transcript library. To identify transcripts that were expressed at the early goat embryonic stage, two cDNA libraries were established from early goat embryos at the two developmental stages, namely, two-cell (IVF-2c) and eight-cell embryos (IVF-8c). Each sample contained fifteen embryos and one sample per stage.

The total RNA was isolated from *in vitro* fertilized (IVF) goat embryos at the two-cell and eight-cell stages. We used the Qubit RNA Assay Kit and the Qubit 2.0 Fluorometer (Life Technologies, CA, USA) to detect the RNA concentration. The RNA Nano 6000 Assay Kit of the Bioanalyzer 2100 System (Agilent Technologies) was employed for RNA integrity evaluation, and SMART-Seq v4 Ultra Low RNA Kit for Sequencing (Clontech) was used to obtain the amplified cDNA in accordance with the manufacturer's instructions. The Covaris System was used to shear the cDNA samples prior to library preparation. Then, the sequencing libraries (IVF-2c and IVF-8c) were generated using the NEBNext Ultra DNA Library Prep Kit for Illumina in accordance with the manufacturer's instructions, and index codes were added to attribute sequences to each sample. The index-coded samples were clustered on a cBot Cluster Generation System by using the TruSeq PE Cluster Kit v3-cBot-HS (Illumina) in accordance with the manufacturer's instructions. After cluster generation, the prepared libraries were sequenced on an Illumina HiSeq

2500 platform, and 125 bp paired-end reads were generated. Clean data were obtained by discarding adaptor sequences and low-quality reads. The Phred score (Q30) and GC content of the clean data were calculated. TopHat (version 2.0.12) (Trapnell *et al.* 2009) was used to map the paired-end clean reads to the reference goat genome, and Bowtie (version 2.2.3) was utilized to build the index of the reference genome (Langmead & Salzberg 2012). HTSeq (version 0.6.1) (Anders *et al.* 2015) was used to count the number of reads mapped to each gene. To estimate gene expression levels, the fragments per kilobase of transcript per million fragments mapped (FPKM) of each gene was calculated based on the length of the gene and the number of reads mapped to the gene. Cufflinks (version 2.1.1) (Trapnell *et al.* 2010) was employed to assemble the reads mapped by TopHat2 (version 2.0.12) (Trapnell *et al.* 2009). To obtain a unique transcriptome, Cuffmerge of the Cufflinks (version 2.1.1) package was used to merge the assembled transcript files (GTF format). The assembled transcripts for obtaining the putative lncRNAs were as follows: (1) to remove potentially known transcripts, Cuffcompare was used to compare the merged transcriptome annotated in the ensemble databases; (2) transcripts with < 200 nt were removed; (3) only transcripts with more than two exons were retained and (4) CNCI, CPC and PFAM were used to filter the putative lncRNA transcripts according to the coding potential (Sun *et al.* 2013, Finn *et al.* 2016). After the transcripts were filtered, Cuffdiff was used to calculate the FPKM of the lncRNAs and the coding genes in all libraries (Fig. 1).

Obtained ERV-derived transcripts

We used Blastall (version 2.2.26) (McGinnis & Madden 2004) to blast the early embryonic transcriptome with all ERV sequences in the goat genome and consequently obtained an e -value $< 10^{-5}$ and the ERV-derived transcripts (Fig. 1). To determine any significant difference among the ERV-derived transcripts during EGA, EdgeR (Robinson *et al.* 2010) was used to estimate the significance of the outputs of Blastall (version 2.2.26) (McGinnis & Madden 2004), and threshold was set to q value < 0.05 .

Goat embryo collection and culture

A 40 μ L sperm suspension (2×10^6 spermatozoa/mL) was added to 40–50 cumulus oocyte complexes in a 400 μ L four-well dish containing BO-IVF medium (IVF Bioscience, Falmouth, United Kingdom) and mineral oil. After 18 h of IVF, the cumulus cells and the redundant spermatozoa were removed from the zygotes by using PBS supplemented with 0.1% hyaluronidase. Then, the zygotes were transferred for culture in BO-IVC medium (IVF Bioscience). Embryos were harvested at different times for subsequent experimentation.

Reverse-transcription PCR and quantitative real-time PCR (qPCR)

All primer pairs used for PCR amplification are shown in Table 1. The PrimeScript RT Reagent Kit with gDNA Eraser

Table 1 Primer pairs used for reverse-transcription qPCR.

Gene	Sequence (5'–3')		Size (bp)
	Forward	Reverse	
<i>ACTB</i>	CTGGGACGACATGGAGAAGATC	GCAGGGGTGTTGAAGGTCTC	159
TCONS_00460156	CTTCTGCCAGGGCTCATAA	GCACCAATAAACCTTAGCACAA	87
TCONS_00327228	CTTGCCAGGCTTAATCAGGGG	CACACCAACAAGGCGTTCACAG	107
TCONS_00261366	GGCTCTTCCGCTATTCAACTC	CATACATTCCTGCCAGTCTTCTA	88
TCONS_00584665	CAGAAGCCCAGATTAAGTCCGA	GCCTTACGCTGACTTGCCTCT	84
TCONS_00238769	GGGATGGTCGTCAGCCTCTT	GTCTTACGGGTTTGACATCATCTT	227
TCONS_00644626	ACACCCACAGGTTTAAGATCAAGT	ATAGCCAGGACATGGAAGCAA	318
TCONS_00376395	CACCTGAGACCCGCCATAAT	GTCCCATAGATCCATTGCTACCT	184
XM_013973863.1(<i>CDKN1A</i>)	GAGGCATCATCAGAACTTTGGG	AGGACGGAGGTGGGAATCAA	132
XM_005681879.2(<i>HSD17B11</i>)	GAGGAAACAGCCACCGAATG	GCCTTTGTAGTCCAGAAATGTGC	242

(Takara) was used to convert the total RNA to cDNA. SYBR Premix Ex Taq II and Onestep Real-Time PCR System (Applied Biosystems) were used to conduct reverse transcription qPCR. The reaction volumes were 10 µL of SYBR Premix Ex Taq II, 0.4 µL of Rox, 0.8 µL of 10 µM forward and reverse primers, 2 µL of template cDNA, and 6 µL of dd H₂O for a final volume of 20 µL. The reaction program was set as follows: 95°C for 30 s, followed by 40 two-step cycles of 95°C for 5 s and 60°C for 34 s. 2^{−ΔΔC_t} method was used to calculate the relative gene expression level.

RNA interference

Smart Silencer (RiboBio, Guangzhou, China) comprises three specific small-interfering RNAs (siRNAs) and an antisense oligonucleotide (asON), whereas mix-siRNA (RiboBio, Guangzhou, China) consist only of three specific siRNAs. The siRNA and asON duplex sequences are shown in Table 2. A Narishige micromanipulator on a Nikon inverted microscope was used to microinject 10 pL (20 µM) of mix-siRNA or Smart Silencer into the zygote cytoplasm. The injected embryos were cultured with BO-IVC medium (IVF Bioscience) in an incubator at 38°C under 6% CO₂ atmosphere. Then, the eight-cell embryos were collected for reverse-transcription qPCR to validate the interference efficiency and detect the target gene expression level.

Western blot analysis

The embryos (*n*=100 embryos per pool) were lysed in RIPA buffer, and each sample was separated by 10% SDS-PAGE and electrotransferred to a PVDF membrane. The membranes were blocked with 5% milk for 2 h at room temperature and incubated overnight at 4°C with the following primary

antibodies: anti-CHD1L (1:200 dilution; cat no. ab155669; Abcam), anti-HSD17B11 (1:500 dilution; cat no. ab136109; Abcam) and anti-GAPDH (1:1000 dilution; cat no. ab157156; Abcam). Then, the membranes were washed and incubated with horseradish peroxidase-conjugated anti-Rabbit-IgG secondary antibody at 1:500 dilution (Thermo Scientific) for 1 h at room temperature. Immunoreactive proteins were visualized by autography with the SuperSignal West Pico Chemiluminescent substrate (Thermo Scientific). All experiments were performed in triplicates.

Statistical analysis

All data were analyzed using ANOVA in SPSS 19.0 for Microsoft Windows. The significance level for the statistical analysis was set to *P*<0.05.

Results

Identification of intact ERV sequences

LTRharvest and MGEScan-LTR were used to enrich the ERV sequences in the goat genome (Rho *et al.* 2007, Ellinghaus *et al.* 2008) (Fig. 1). LTRharvest obtained 64,159 ERV candidates, and LTRdigest identified 459 ERV sequences, including PBS, PPT and protein-coding domains. MGEScan-LTR obtained 186 ERVs sequences. A total of 645 intact ERVs, which constituted the goat ERV library, were identified by using LTRharvest and MGEScan-LTR (Supplementary Table 1, see section on supplementary data given at the end of this article). The LTR lengths of the ERVs ranged from 100 bp to 1345 bp, and the total internal lengths ranged from 2580 bp to 20,282 bp (Supplementary Table 2).

Classification and census of goat ERVs

We constructed phylogenetic trees from 436 goat ERVs, including the RT domain, by using MEGA6 (Tamura *et al.* 2013). The goat ERVs were classified into 18 major families, which were labeled as ERV1 to ERV18 (Fig. 2A). We identified 13 ERV families (ERV-1 to -13) with 433 elements as class-I ERVs. Four ERV families (ERV-14 to -17) accounted for 195 elements and

Table 2 SiRNA sequences for RNA interference.

SiRNA ID	Target sequences
TCONS_00460156-SiRNA1	CAAGTTCTCACATCCCTTA
TCONS_00460156-SiRNA2	GGAGCTATACCTTATCTT
TCONS_00460156-SiRNA3	CGCTTAACCTCCCTTTAAA
TCONS_00460156-ASO	AACGTCCACCCTGACTAACC
HSD17B11-SiRNA1	GCAGAAGTTGGAGATGTTA
HSD17B11-SiRNA2	CCTTGGAAGAAGCTGGAAT
HSD17B11-SiRNA3	GCAGACTGCCTACGAATTT

Figure 2 Comprehensive analysis of ERV involvement and chromosome distribution. (A) Phylogenetic tree of goat ERVs constructed by analyzing the polymorphisms of reverse transcriptase. Class III was flanked by classes I and II, and different families with different colors and bootstrap values are shown above the branches with neighbor-joining bootstrap values of >70%. MMTV, mouse mammary tumor virus; GH-G18, Golden Hamster intracisternal A-partide H18; BLV, bovine leukemia virus; HFV, human endogenous retrovirus; ZFERV, zebrafish endogenous retrovirus; FELV, feline leukemia virus; MuLV, murine leukemia virus; MDEV, Mus dunni endogenous virus; GaLV, gibbon ape leukemia virus; BaEV, baboon endogenous virus; KoRV, koala retrovirus; SnRV, snakehead fish retrovirus; SRV-1, Simian type-D retrovirus-1; JSRV, Jaagsiekte sheep retrovirus; HML, human MMTV-like; WDSV, walleye dermal sarcoma virus; HIV-1, human immunodeficiency virus; ALV, avian leukemia virus. (B) Distribution of different classes of ERVs in the chromosome. Each chromosome comprises a similar percentage of different classes except for the Y chromosome, which can only be found in class I. (C) ERV sequence structure feature display. Gag, pol, and pro represent the ERV repeat region.

were designated as class-II ERVs, and one family (ERV-18) represented by 17 elements was classified as class-III ERVs (Table 3). Using RepeatMasker, we annotated the

ERV sequences, tallied the percentage of ERVs located in the goat genome, and determined the statistical data of the ERVs with regard to their family and class

Table 3 Information on ERV elements of goat genomes.

Family	Classification	Neighbor-joining bootstrap	LTR length range (bp)	Copy no.	Internal length range (bp)
ERV1	Class I	0.80	109–604	106	5408–20,282
ERV2	Class I	0.98	101–968	67	5057–20,127
ERV3	Class I	1.00	134–483	5	8768–16,883
ERV4	Class I	0.99	119–635	18	4386–14,079
ERV5	Class I	0.93	100–445	9	5952–13,694
ERV6	Class I	0.86	101–542	27	2580–14,801
ERV7	Class I	0.83	132–495	18	7108–15,825
ERV8	Class I	0.94	103–993	67	3956–16,666
ERV9	Class I	0.94	109–995	19	6723–17,876
ERV10	Class I	0.98	135–661	25	5188–17,138
ERV11	Class I	0.76	106–1046	69	3732–20,183
ERV12	Class I	0.91	125–622	3	9075–13,430
ERV13	Class I	0.77	105–563	60	2877–18,093
ERV14	Class II	1.00	119–299	5	6254–7120
ERV15	Class II	0.84	118–833	56	5262–20,168
ERV16	Class II	0.99	132–1345	47	4681–18,773
ERV17	Class II	0.94	145–482	27	4908–14,917
ERV18	Class III	0.84	161–823	17	3468–19,099

in each chromosome. The percentages of classes I, II, and III varied from 9 to 11%, from 4 to 5% and from 0.5 to 0.7%, respectively, in each chromosome. Only 1.46% of class I was located in the Y chromosome (Fig. 2B). No significant difference was observed in the distribution of families in all chromosomes except for the Y chromosome, which was only enriched in class I. Then, we annotated all the intact ERV sequence with their structural information and found most of the ERV sequences in goat that had distinctive structural features without an envelope structural region (Env-less) (Fig. 2C and Supplementary Table 3).

Enrichment analysis of nearby genes of ERV elements

GO (Huang *et al.* 2009) analysis was conducted to explore the functions of nearby genes. We obtained 48 GO terms that were mainly associated with binding and catalytic activity (Fig. 3A and Supplementary Table 4). Pathway analysis revealed that the nearby genes were enriched in 169 KEGG (Kanehisa *et al.* 2008) pathways and mainly associated with fatty acid metabolism, amino acid metabolism, and peroxidase transport. Among the 169 pathways, only the MAPK, Wnt and cAMP signaling pathways (Fig. 3B and Supplementary Table 4) were related to embryonic development.

Transcriptome analysis of early goat embryos by RNA-seq

The goat IVF-2c and IVF-8c libraries were sequenced using an Illumina Hiseq 2500 platform, and 125 bp of paired-ended reads were consequently generated. A total of 378,170,482 raw reads were obtained. After discarding the adaptor sequences and low-quality reads, we acquired 271,093,190 clean reads. The Q30 of clean reads and GC content ranged from 91.98 to

92.65% and from 43.04 to 45.72% in the two-cell and eight-cell embryos, respectively. We mapped the clean reads to the goat reference genome sequence (CHIR_1.0, NCBI). Approximately 74.89–79.53% of the clean reads in all libraries were mapped to the goat reference genome (Supplementary Table 5). Then, the mapped sequences were assembled using Cufflinks, and 30,148 mRNAs were obtained. We performed strict pipeline filtering to discard non-fitting transcripts on the basis of the lncRNA characteristics. After strict pipeline filtering, we obtained 40,251 transcripts. Then, we utilized CPC, CNCI and PFAM (Sun *et al.* 2013, Finn *et al.* 2016) to predict the protein-coding potentials of these transcripts. Finally, we identified 36,240 lncRNAs. The FPKM scores of the lncRNAs and the mRNAs in the two-cell and eight-cell embryos were calculated by Cuffdiff (2.1.1), and the FPKM distributions of the lncRNAs and the mRNAs were estimated. We found that the lncRNAs showed a wider FPKM distribution and a lower expression level than those of the mRNAs (Trapnell *et al.* 2012) (Fig. 4A and Supplementary Table 6). Basing on the 36,240 lncRNAs and the 30,148 mRNAs, we detected 679 upregulated transcripts and 64 downregulated transcripts between IVF-8c and IVF-2c (Fig. 4B and Supplementary Table 6). The differential expression of the 145 lncRNAs and the 598 mRNAs was measured by systematic cluster analysis (Fig. 4C and Supplementary Table 7). We found several canonical EGA genes, including ZSCAN4 and EIF1A, which showed significantly higher expression levels in IVF-8c than in IVF-2c. These findings validated the robustness of our RNA-seq results.

Screening of ERV-derived transcripts

We obtained 165 ERV-derived lncRNAs and 111 ERV-derived mRNAs by blasting the transcript library with the

Figure 3 Enrichment analysis of ERV elements in nearby genes. The input gene number is small, and the significance is determined by $P < 0.05$. (A) GO analysis of the target of ERV upstream–downstream elements within 10kb. (B) KEGG pathway analysis of the target of ERVs upstream–downstream elements within 10kb.

goat ERV library (Supplementary Table 8). We identified four ERV-derived lncRNAs and one ERV-derived mRNA that were differentially expressed between IVF-2c and IVF-8c (q -value < 0.05) (Supplementary Table 9). The expression levels of TCONS_00460156, TCONS_00261366, TCONS_00327228 and cyclin-dependent kinase inhibitor 1A (CDKN1A) in IVF-8c were significantly higher than those in IVF-2c, whereas the expression of TCONS_00565459 in IVF-8c was significantly lower than that in IVF-2c. The expression levels of TCONS_00460156, TCONS_00261366, TCONS_00327228 and CDKN1A were consistent with the features of the EGA process from the two-cell to the eight-cell stage. Therefore, we selected these transcripts as candidate genes to determine whether

their expression profiles were similar with those of EGA genes.

The developmental expression patterns of the differentially expressed transcripts (TCONS_00460156, TCONS_00261366, TCONS_00327228 and CDKN1A) and the highly expressed transcripts in IVF-8c (TCONS_00584665, TCONS_00238769, TCONS_00644626, TCONS_00376395 and 17- β -hydroxysteroid dehydrogenase 11 (HSD17B11)) were examined by reverse-transcription qPCR (Fig. 5), and the results agreed with the RNA-seq results. We found that the expression levels of TCONS_00460156, TCONS_00261366 and HSD17B11 at the eight-cell stage were significantly higher than those at other embryonic stages.

Figure 4 RNA-seq results of IVF-two-cell and IVF-eight-cell embryos. (A) FPKM distributions of lncRNA and mRNA. (B) Volcano plot. Blue dots present upregulated transcripts, red dots denote downregulated transcripts, and green dots indicate no significant difference. (C) Hierarchical clustering analysis of 743 differentially expressed lncRNAs and mRNAs between IVF-2c and IVF-8c. Data are expressed as FPKM. Red: relatively high expression. Blue: relatively low expression.

Specific ERV-derived transcripts are vital for EGA

The characteristics of TCONS_00460156, TCONS_00261366 and HSD17B11 were consistent with those of the EGA gene expression profile, which displayed a higher expression at the eight-cell stage than at other stages during early embryonic development. Therefore, these transcripts were selected as EGA candidates for the subsequent analysis. We conducted RNA interference assay to examine the roles of TCONS_00460156, TCONS_00261366 and HSD17B11 in the early embryonic development of goat. The interference efficiency was approximately 70% when mix-siRNA was injected into the cytoplasm of the zygote, and the protein level of HSD17B11 decreased. Considering that siRNAs and double-strand RNAs failed to effectively knock down the lncRNA, we selected Smart Silencer, which successfully interfered with TCONS_00460156 and TCONS_00261366 (Fig. 6A and D). The downregulation of TCONS_00460156 and HSD17B11 resulted in a developmental arrest at the eight-cell stage of goat embryos, and the blastocyst rate of the downregulated groups was relatively lower than that of the control group. Nonetheless, the

downregulation of TCONS_00261366 did not affect the embryonic development (Fig. 6B and C), whereas the downregulation of TCONS_00460156 decreased the mRNA expression of chromodomain helicase/ATPase DNA binding protein 1-like gene (CHD1L), which is a nearby gene of TCONS_00460156. Moreover, the protein level of CHD1L markedly decreased as depicted by the Western blot images (Fig. 6A and D).

Discussion

We employed ERV-mining pipelines to investigate the genome-wide distribution of intact ERV elements and constructed a transcript library from IVF-2c and IVF-8c embryos by RNA-seq. To determine the role of goat ERV elements during EGA, we obtained ERV-derived transcripts by bidirectionally blasting the two libraries. We identified the transcripts highly expressed during EGA and found that lncRNA TCONS_00460156 and HSD17B11 genes were essential for maintaining a normal embryonic development.

ERVs play an important role in regulating gene expression involved in many physiological

and pathological processes (Rebollo *et al.* 2012, Suntsova *et al.* 2015). A large number of transcripts associated with ERV sequences have been discovered in early embryonic development in mouse and human, but the specific biological function and regulatory mechanism of most transcripts are unknown (Fadloun *et al.* 2013, Gifford *et al.* 2013, Goke *et al.* 2015). At present, information on ERV sequences in the goat genome is unavailable. Therefore, we first enriched all intact ERV sequences in the whole genome of goat and established a comprehensive ERV library to analyze the function of ERVs. We found that class I ERVs were the most diverse and abundant in the goat genome (composed of 13 major families), which is markedly different from other mammalian genomes (Mager & Stoye 2015). Most ERV sequences in goat have distinctive structural features without an envelope structural region (Env-less ERVs) compared with those in other species (Fig. 2C). The absence of the *env* gene blocks the autonomous extracellular replication of retroviruses, thereby causing ERVs to become genomic components of their host's genome. The diversity of ERV sequences in mammalian genomes implies that ERVs may perform distinct functions in different species.

ERV elements containing virus promoter sequences can be used as promoters or enhancers to regulate the expression of nearby genes, and they may transcribe non-encoding RNAs participating in gene expression regulation (Kouamo & Kharche 2015, Suntsova *et al.* 2015). The KEGG and GO enrichment of nearby genes of ERVs indicated that some goat ERV elements were associated with embryonic development (Fig. 3). Then, we focused on the expression profile of ERV during the EGA of goat. Previous studies showed that major EGA

occurs at the eight-cell stage of goat embryos, and the main feature of EGA genes is that their expression levels are significantly higher than those at other stages (Ferrer *et al.* 1995, Pivko *et al.* 1995, Deng *et al.* 2018). Therefore we performed high-throughput sequencing to establish a goat embryonic transcript library (IVF-2c and IVF-8c) and found numerous transcripts activated in IVF_8c embryos, and this observation was consistent with mammalian EGA (Fig. 4B). We identified 145 lncRNAs and 598 mRNAs differentially expressed between IVF-8c and IVF-2c and selected them as candidates for EGA (Fig. 4C). Several genes, such as EIF1A and ZSCAN4, are actively transcribed at the onset of EGA (Ko 2016). Although our RNA-seq data did not have three biological repetitions and sequenced only IVF_2c and IVF_8c, ZSCAN4 and EIF1A were included in the differentially expressed genes, and the mRNA expression profiles of the nine selected genes validated the accuracy of our RNA-seq results (Fig. 5B). In future works, more embryonic stage groups should be prepared to identify more vital genes for EGA.

The penetration and invasion of ERVs in vertebrate genomes are ubiquitous, providing a source of genetic variation that strongly influences the biological characteristics and evolution of host species (Jern & Coffin 2008). The ERV expression reflects the successive reprogramming of the embryonic genome during the transition from an oocyte to an embryo and in preimplantation embryos (Peaston *et al.* 2004, Batut *et al.* 2013). To investigate the ERV transcripts during EGA, we utilized the embryonic transcript library and the ERV library to perform a bidirectional blast. We obtained 165 lncRNAs and 111 mRNAs derived from ERVs in goat IVF-2c and IVF-8c embryos. This result

Figure 5 Expression profile of transcripts. β -Actin functioned as the control. The error bars represent S.E.M. Thirty embryos from each stage were used, and three experimental replications were performed. GFF, goat fetal fibroblast; MII, metaphase II; 2C, two-cell; 8C, eight-cell; MO, morula; BL, blastocyst.

Figure 6 Function of ERV-derived transcripts in the goat embryonic development. (A) SiRNA and Smart Silencer efficiently mediated the mRNA and lncRNA knockdown, respectively, and decreased the mRNA expression of CHD1L, which is the target gene of TCONS_00460156. Embryos were collected at the eight-cell stage for reverse-transcription qPCR analysis. Error bars represent s.e.m. Thirty embryos from each stage were used in each group. (B) Embryos injected with the control SiRNA were successfully developed to hatched blastocyst stage. SiRNA-depleted embryos arrested at the eight-cell stage. Scale bar, 100 μ m. A minimum of three experimental replications were performed. (C) Effects of TCONS_00460156 and HSD17B11 knockdown on the development of embryos to the blastocyst stage. (D) Knockdown of HSD17B11 and TCONS_00460156 considerably decreased the protein levels of HSD17B11 and CHD1L, respectively. Embryos were collected at the eight-cell stage for Western blot analysis. Approximately 100 eight-cell embryos were used in each group for GAPDH.

indicated that ERV elements were involved in forming chimeric mRNAs with a subset of host genes and transcribed to lncRNAs. Approximately two-thirds of the lncRNAs identified in mammals are related to ERVs and other transposable elements, and the LTR of ERVs generally acts as a promoter of lncRNAs (Kapusta *et al.* 2013). Moreover, the expression levels of lncRNAs derived from transposable elements in early embryos and embryonic stem cells were significantly higher than those of other sequences in adult cells and tissues, suggesting the important role of lncRNAs derived from transposable elements in regulating pluripotency (Kelley & Rinn 2012, Fadloun *et al.* 2013).

EGA is an important step for early embryos to establish the totipotent state of each blastomere, and the delay of EGA initiation leads to embryo blocking (Kigami *et al.* 2003). In our study, the downregulation of TCONS_00460156 decreased the expression level of CHD1L, which is a nearby gene of TCONS_00460156, ultimately blocking the IVF-8c embryo and reducing the blastocyst formation rate (Fig. 6C and D). CHD1L is a chromatin enzyme that is critical for early development, and chromatin modifiers are vital in preimplantation embryos (Snider *et al.* 2013). The regulatory mechanisms of ERV, lncRNA, and targeted genes should be further investigated. In

addition, the downregulation of HSD17B11 caused a phenomenon similar to that of TCONS_00460156 in early goat embryos. HSD17B11, which is a specific protein enriched in the lipid droplet-enriched fraction, is distributed to lipid droplets (Fujimoto *et al.* 2004, Yokoi *et al.* 2007). Deficient lipid droplets cannot store extranuclear free histones, which are harmful to embryos during early development (Li *et al.* 2012). We found that the LTR of ERVs was directly related to the 3'UTR of HSD17B11. We assumed that the LTR of goat ERVs might control the expression of HSD17B11, which mainly functions by affecting the formation of lipid droplets in the embryos. Considerable effort should be devoted to investigating the role of HSD17B11 in embryonic development.

Our study systematically analyzed the distribution and structural characteristics of goat ERVs and provided their expression profiles during goat EGA. In addition, we identified TCONS_00460156 and HSD17B11 to be essential for goat embryonic development. As the first functional ERV-associated mRNA and lncRNA revealed in goat embryos, TCONS_00460156 and HSD17B11 provided insights into ERV functions during EGA. Our data might facilitate the understanding of the physiological roles of ERV-associated RNAs for embryonic development.

Supplementary data

This is linked to the online version of the paper at <https://doi.org/10.1530/REP-18-0336>.

Declaration of interest

The authors declare that there is no conflict of interest that could be perceived as prejudicing the impartiality of the research reported.

Funding

This work was supported by the National Natural Science Foundation of China (Grant No. 31772689) and the Innovation Project of Science and Technology in Shaanxi Province (Grant No. 2015KTCQ02-18).

References

- Anders S, Pyl PT & Huber W 2015 HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31** 166–169. (<https://doi.org/10.1093/bioinformatics/btu638>)
- Batut P, Dobin A, Plessy C, Carninci P & Gingeras TR 2013 High-fidelity promoter profiling reveals widespread alternative promoter usage and transposon-driven developmental gene expression. *Genome Research* **23** 169–180. (<https://doi.org/10.1101/gr.139618.112>)
- Bui LC, Esvikov AV, Khan DR, Archilla C, Peynot N, Henaut A, Le Bourhis D, Vignon X, Renard JP & Duranthon V 2009 Retrotransposon expression as a defining event of genome reprogramming in fertilized and cloned bovine embryos. *Reproduction* **138** 289–299. (<https://doi.org/10.1530/REP-09-0042>)
- Deng M, Wan Y, Liu Z, Ren C, Zhang G, Pang J, Zhang Y & Wang F 2018 Long noncoding RNAs exchange during zygotic genome activation in goat. *Biology of Reproduction* **99** 707–717. (<https://doi.org/10.1093/biolre/roy118>)
- Dong Y, Xie M, Jiang Y, Xiao N, Du X, Zhang W, Tosser-Klopp G, Wang J, Yang S, Liang J *et al.* 2013 Sequencing and automated whole-genome optical mapping of the genome of a domestic goat (*Capra hircus*). *Nature Biotechnology* **31** 135–141. (<https://doi.org/10.1038/nbt.2478>)
- Durruthy-Durruthy J, Sebastiano V, Wossidlo M, Cepeda D, Cui J, Grow EJ, Davila J, Mall M, Wong WH, Wysocka J *et al.* 2016 The primate-specific noncoding RNA HPAT5 regulates pluripotency during human preimplantation development and nuclear reprogramming. *Nature Genetics* **48** 44–52. (<https://doi.org/10.1038/ng.3449>)
- Eddy SR 2011 Accelerated profile HMM searches. *PLOS Computational Biology* **7** e1002195. (<https://doi.org/10.1371/journal.pcbi.1002195>)
- Edgar RC 2004 MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **5** 1–19. (<https://doi.org/10.1186/1471-2105-5-1>)
- Ellinghaus D, Kurtz S & Willhoeft U 2008 LTRharvest, an efficient and flexible software for de novo detection of LTR retrotransposons. *BMC Bioinformatics* **9** 18. (<https://doi.org/10.1186/1471-2105-9-18>)
- Fadloun A, Le Gras S, Jost B, Ziegler-Birling C, Takahashi H, Gorab E, Carninci P & Torres-Padilla M-E 2013 Chromatin signatures and retrotransposon profiling in mouse embryos reveal regulation of LINE-1 by RNA. *Nature Structural and Molecular Biology* **20** 332–338. (<https://doi.org/10.1038/nsmb.2495>)
- Ferrer F, Garcia C, Villar J & Arias M 1995 Ultrastructural study of the early development of the sheep embryo. *Anatomia Histologia Embryologia: Journal of Veterinary Medicine Series C* **24** 191–196. (<https://doi.org/10.1111/j.1439-0264.1995.tb00034.x>)
- Finn RD, Coghill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, Potter SC, Punta M, Qureshi M, Sangrador-Vegas A *et al.* 2016 The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Research* **44** D279–D285. (<https://doi.org/10.1093/nar/gkv1344>)
- Friedli M & Trono D 2015 The developmental control of transposable elements and the evolution of higher species. *Annual Review of Cell and Developmental Biology* **31** 429–451. (<https://doi.org/10.1146/annurev-cellbio-100814-125514>)
- Fujimoto Y, Itabe H, Sakai J, Makita M, Noda J, Mori M, Higashi Y, Kojima S & Takano T 2004 Identification of major proteins in the lipid droplet-enriched fraction isolated from the human hepatocyte cell line HuH7. *Biochimica et Biophysica Acta (BBA): Molecular Cell Research* **1644** 47–59. (<https://doi.org/10.1016/j.bbamcr.2003.10.018>)
- Gifford WD, Pfaff SL & Macfarlan TS 2013 Transposable elements as genetic regulatory substrates in early development. *Trends in Cell Biology* **23** 218–226. (<https://doi.org/10.1016/j.tcb.2013.01.001>)
- Goeke J, Lu X, Chan Y-S, Ng H-H, Ly L-H, Sachs F & Szczerbinka I 2015 Dynamic transcription of distinct classes of endogenous retroviral elements marks specific populations of early human embryonic cells. *Cell Stem Cell* **16** 135–141. (<https://doi.org/10.1016/j.stem.2015.01.005>)
- Goke J, Lu X, Chan YS, Ng HH, Ly LH, Sachs F and Szczerbinka I 2015 Dynamic transcription of distinct classes of endogenous retroviral elements marks specific populations of early human embryonic cells. *Cell Stem Cell* **16** 135–141. (<https://doi.org/10.1016/j.stem.2015.01.005>)
- Huang DW, Sherman BT & Lempicki RA 2009 Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols* **4** 44–57. (<https://doi.org/10.1038/nprot.2008.211>)
- Jern P & Coffin JM 2008 Effects of retroviruses on host genome function. *Annual Review of Genetics* **42** 709–732. (<https://doi.org/10.1146/annurev.genet.42.110807.091501>)
- Kanehisa M, Araki M, Goto S, Hattori M, Hirakawa M, Itoh M, Katayama T, Kawashima S, Okuda S, Tokimatsu T *et al.* 2008 KEGG for linking genomes to life and the environment. *Nucleic Acids Research* **36** D480–D484. (<https://doi.org/10.1093/nar/gkm882>)
- Kapusta A, Kronenberg Z, Lynch VJ, Zhuo X, Ramsay L, Bourque G, Yandell M & Feschotte C 2013 Transposable elements are major contributors to the origin, diversification, and regulation of vertebrate long noncoding RNAs. *PLoS Genetics* **9** e1003470. (<https://doi.org/10.1371/journal.pgen.1003470>)
- Kelley D & Rinn J 2012 Transposable elements reveal a stem cell-specific class of long noncoding RNAs. *Genome Biology* **13** R107. (<https://doi.org/10.1186/gb-2012-13-11-r107>)
- Kigami D, Minami N, Takayama H & Imai H 2003 MuERV-L is one of the earliest transcribed genes in mouse one-cell embryos. *Biology of Reproduction* **68** 651–654. (<https://doi.org/10.1095/biolreprod.102.007906>)
- Ko MS 2016 Zygotic genome activation revisited: looking through the expression and function of Zscan4. *Current Topics in Developmental Biology* **120** 103–124. (<https://doi.org/10.1016/bs.ctdb.2016.04.004>)
- Kohany O, Gentles AJ, Hankus L & Jurka J 2006 Annotation, submission and screening of repetitive elements in Repbase: RepbaseSubmitter and Censor. *BMC Bioinformatics* **7** 474. (<https://doi.org/10.1186/1471-2105-7-474>)
- Kouamo J & Kharche SD 2015 A comparative study of parthenogenetic activation and in vitro fertilization of in vitro matured caprine oocytes. *Iranian Journal of Veterinary Research* **16** 20–24.
- Langmead B & Salzberg SL 2012 Fast gapped-read alignment with Bowtie 2. *Nature Methods* **9** 357–U354. (<https://doi.org/10.1038/nmeth.1923>)
- Li Z, Thiel K, Thul PJ, Beller M, Kuhnlein RP & Welte MA 2012 Lipid droplets control the maternal histone supply of Drosophila embryos. *Current Biology* **22** 2104–2113. (<https://doi.org/10.1016/j.cub.2012.09.018>)
- Macfarlan TS, Gifford WD, Driscoll S, Lettieri K, Rowe HM, Bonanomi D, Firth A, Singer O, Trono D & Pfaff SL 2012 Embryonic stem cell potency fluctuates with endogenous retrovirus activity. *Nature* **487** 57–63. (<https://doi.org/10.1038/nature11244>)
- Mager DL & Stoye JP 2015 Mammalian endogenous retroviruses. *Microbiology Spectrum* **3** MDNA3-0009-2014. (<https://doi.org/10.1128/microbiolspec.MDNA3-0009-2014>)
- McGinnis S & Madden TL 2004 BLAST: at the core of a powerful and diverse set of sequence analysis tools. *Nucleic Acids Research* **32** W20–W25. (<https://doi.org/10.1093/nar/gkh435>)
- Peaston AE, Esvikov AV, Graber JH, de Vries WN, Holbrook AE, Solter D & Knowles BB 2004 Retrotransposons regulate host genes in mouse oocytes and preimplantation embryos. *Developmental Cell* **7** 597–606. (<https://doi.org/10.1016/j.devcel.2004.09.004>)

- Pivko J, Grafenau P & Kopečný V 1995 Nuclear-fine-structure and transcription in early goat embryos. *Theriogenology* **44** 661–671. ([https://doi.org/10.1016/0093-691X\(95\)00246-5](https://doi.org/10.1016/0093-691X(95)00246-5))
- Rebollo R, Farivar S & Mager DL 2012 C-GATE – catalogue of genes affected by transposable elements. *Mobile DNA* **3** 9. (<https://doi.org/10.1186/1759-8753-3-9>)
- Rhead B, Karolchik D, Kuhn RM, Hinrichs AS, Zweig AS, Fujita PA, Diekhans M, Smith KE, Rosenbloom KR, Raney BJ *et al.* 2010 The UCSC Genome Browser database: update 2010. *Nucleic Acids Research* **38** D613–D619. (<https://doi.org/10.1093/nar/gkp939>)
- Rho M, Choi J-H, Kim S, Lynch M & Tang H 2007 De novo identification of LTR retrotransposons in eukaryotic genomes. *BMC Genomics* **8** 90. (<https://doi.org/10.1186/1471-2164-8-90>)
- Robinson MD, McCarthy DJ & Smyth GK 2010 edgeR: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26** 139–140. (<https://doi.org/10.1093/bioinformatics/btp616>)
- Snider AC, Leong D, Wang QT, Wysocka J, Yao MW & Scott MP 2013 The chromatin remodeling factor Chd1l is required in the preimplantation embryo. *Biology Open* **2** 121–131. (<https://doi.org/10.1242/bio.20122949>)
- Steinbiss S, Willhoeft U, Gremme G & Kurtz S 2009 Fine-grained annotation and classification of de novo predicted LTR retrotransposons. *Nucleic Acids Research* **37** 7002–7013. (<https://doi.org/10.1093/nar/gkp759>)
- Sun L, Luo H, Bu D, Zhao G, Yu K, Zhang C, Liu Y, Chen R & Zhao Y 2013 Utilizing sequence intrinsic composition to classify protein-coding and long noncoding transcripts. *Nucleic Acids Research* **41** e166. (<https://doi.org/10.1093/nar/gkt646>)
- Suntsova M, Garazha A, Ivanova A, Kaminsky D, Zhavoronkov A & Buzdin A 2015 Molecular functions of human endogenous retroviruses in health and disease. *Cellular and Molecular Life Sciences* **72** 3653–3675. (<https://doi.org/10.1007/s00018-015-1947-6>)
- Tamura K, Stecher G, Peterson D, Filipski A & Kumar S 2013 MEGA6: molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution* **30** 2725–2729. (<https://doi.org/10.1093/molbev/mst197>)
- Tarailo-Graovac M & Chen N 2009 Using RepeatMasker to identify repetitive elements in genomic sequences. *Current Protocols in Bioinformatics* Chapter 4 Unit 4.10. (<https://doi.org/10.1002/0471250953.bi0410s25>)
- Thompson PJ, Macfarlan TS & Lorincz MC 2016 Long terminal repeats: from parasitic elements to building blocks of the transcriptional regulatory repertoire. *Molecular Cell* **62** 766–776. (<https://doi.org/10.1016/j.molcel.2016.03.029>)
- Trapnell C, Pachter L & Salzberg SL 2009 TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25** 1105–1111. (<https://doi.org/10.1093/bioinformatics/btp120>)
- Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, Salzberg SL, Wold BJ & Pachter L 2010 Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nature Biotechnology* **28** 511–517. (<https://doi.org/10.1038/nbt.1621>)
- Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL & Pachter L 2012 Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature Protocols* **7** 562–578. (<https://doi.org/10.1038/nprot.2012.016>)
- Wang J, Li X, Wang L, Li J, Zhao Y, Bou G, Li Y, Jiao G, Shen X, Wei R *et al.* 2016 A novel long intergenic noncoding RNA indispensable for the cleavage of mouse two-cell embryos. *EMBO Reports* **17** 1452–1470. (<https://doi.org/10.15252/embr.201642051>)
- Yokoi Y, Horiguchi Y, Araki M & Motojima K 2007 Regulated expression by PPARalpha and unique localization of 17beta-hydroxysteroid dehydrogenase type 11 protein in mouse intestine and liver. *FEBS Journal* **274** 4837–4847. (<https://doi.org/10.1111/j.1742-4658.2007.06005.x>)

Received 29 June 2018

First decision 16 August 2018

Revised manuscript received 29 October 2018

Accepted 5 November 2018