

Molecular characterisation, genetic variability and detection of a functional polymorphism influencing the promoter activity of *OXT* gene in goat and sheep

Gianfranco Cosenza¹, Marco Iannaccone^{1*}, Boipuso Alpheus Pico^{1,2}, Daniela Gallo¹, Rosanna Capparelli^{1*} and Alfredo Pauciullo³

¹ Department of Agricultural Sciences, University of Naples ‘Federico II’, Portici (NA), Italy

² Faculty of Agriculture, Science and Technology, North–West University, Private Bag X2046, MMABATHO, 2735, South Africa

³ Department of Agricultural, Forest and Food Science, University of Torino, Grugliasco (TO), Italy

Received 11 October 2016; accepted for publication 1 February 2017; first published online 14 March 2017

The purpose of the study described in this Research Communication was to report the full characterisation of the goat and sheep oxytocin-neurophysin I gene (*OXT*), their promoters and amino acid sequences. Using the genomic DNA as template, we sequenced and compared the whole *OXT* gene (3 exons), plus 958/960 nucleotides at the 5' flanking region and 478/477 nucleotides at the 3' flanking region, in 46 sheep and 24 goats belonging to different breeds/genetic types reared in Italy, Greece and Germany. The comparison of the obtained sequences showed a high degree of genetic variability at these *loci*. In particular, we focused on the SNP g.438T > C as possible example of trans-specific polymorphism. This SNP alters a putative binding site of the transcription factor Oct-1. The set-up of a luciferase assay confirmed that the C variant of this SNP negatively affects the promoter activity of the sheep *OXT* gene. The results of this study suggest that the SNP g.438T > C might be useful to promote association studies with traits/physiological processes controlled by this hormone.

Keywords: *Capra hircus*, *Ovis aries*, *OXT locus*, gene polymorphism, luciferase assay.

Oxytocin is a peptide hormone released into general circulation from the posterior pituitary gland. It is involved in the contraction of smooth muscles, inducing uterine contractions during calving and causing milk ejection (Gimpl & Fahrenholz, 2001). It has been suggested that oxytocin injected before milking may increase milk production, due to an increase in milk flow rate and a decrease in the machine milking time (Lollivier et al. 2002). However, other roles characterise oxytocin: it controls oestrous cycle length, the follicular luteinisation and ovarian steroidogenesis. Oxytocin also acts as a neurotransmitter in the central nervous system and plays a role in various processes like cognition, tolerance, adaptation, and complex sexual and maternal behaviour (Gimpl & Fahrenholz, 2001).

To date, the oxytocin encoding sequences (*OXT*) have been fully reported for many vertebrates, however, incomplete information exists for some ruminants. In particular,

no nucleotide or amino acid information is available in goats.

Despite considerable research efforts and interest among the scientific community on the oxytocin gene, so far, only a few studies on polymorphism detection, gene expression, and association with specific traits have been carried out. For instance, in ruminants, Cosenza et al. (2007) identified polymorphisms in the river buffalo *OXT* gene and Pauciullo et al. (2012a) found a significant association of these SNPs with buffalo milk yield and flow.

In the present study, we report the complete sequences of goat and sheep *OXT* gene, the identification and characterisation of a high degree of genetic variability at these *loci* and we assess their possible effects on the *OXT* promoter activity.

Materials and methods

DNA samples

Using the method of Goossens & Kan (1981) genomic DNA was extracted from individual blood leucocytes of 24 Italian sheep (nine undefined genetic type, 15 Laticauda), 22

*For correspondence; e-mail: m.iannaccone@unina.it; capparel@unina.it

German sheep (10 Merino Landschaf, three Rhönshaf, three Texel, three Ostfriesisches Milchschaft and three Grau Gehörnte Heidschnucke), 14 German goats (four Thuringer Waldziege, four Weiße Deutsche, four Bunte Deutsche Edelziege and two Boer goat cross), two Greek and eight Italian goats of undefined genetic type, randomly chosen.

Primer design, PCR conditions and sequencing

All primers for amplification and sequencing of sheep and goat *OXT* were designed by DNAsis-Pro software (Hitachi), using the bovine (EMBL acc. no. AB481096) and ovine (EMBL acc. nos. X55131; CP011898) *OXT* gene sequences as templates (online Supplementary Table S1).

PCR reaction mixture and thermal conditions were accomplished according to Cosenza et al. (2007). PCR products were sequenced on both strands at CEINGE – Biotecnologie Avanzate (Naples, Italy).

Bioinformatic and statistical analysis

BLAST analysis (www.ncbi.nlm.nih.gov/BLAST) was used to confirm the sequencing results. Allelic frequencies and Hardy-Weinberg equilibrium (χ^2 test) were calculated. Homology searches, comparison among nucleotide and deduced amino acid sequences and multiple alignments were accomplished using DNAsis-Pro (Hitachi Software Engineering Co., Japan).

Construction of reporter plasmids, transient transfections and luciferase assays

To evaluate the role of the SNP g.438T>C on the sheep *OXT* expression, a 941 bp DNA fragment including the *OXT* promoter region was amplified from homozygous T/T and C/C samples. Furthermore, a nested PCR with two modified internal primers was carried out to provide restriction sites for the *KpnI* (GGTAC!C) and *SmaI* (CCC!GGG) endonucleases. Primers sequences are reported in online Supplementary Table S1.

The nested PCR products were digested with *KpnI* and *SmaI* endonucleases and cloned upstream of a reporter in the pGL3 basic vector (Promega) as previously reported (Cosenza et al. 2016). Later, the new products were sequenced using vector primers RVprimer3 and RVprimer4 (as recommended in the supporting information) to confirm the orientation and integrity of the inserts. HEK293 cells were then transiently transfected with 200 ng of the new plasmids and with 20 ng of pRL-TK vector expressing Renilla luciferase (Promega) using Fugene HD transfection reagent (Promega). A transfection with pGL3 control vector was also performed as control. All transfections were carried out in triplicate and in three independent experiments. After 24 h, the reporter activity of g.438T and g.438C was measured by luminometry on the Victor luminometer (PerkinElmer) using the Dual-Glo luciferase assay system (Promega). In all

transfection, Firefly luciferase values were normalised with the Renilla luciferase expressed by the vector pRL-TK co-transfected as internal control. Promoter activity for each construct was calculated as ratio of normalised reporter activity to pGL3 control vector.

Results and discussion

Structure of goat and sheep oxytocin-neurophysin I gene

We sequenced the genomic DNA of the whole oxytocin-neurophysin I encoding gene (*OXT*) plus 958/960 nucleotides at the 5' flanking region and 478/477 nucleotides at the 3' flanking region respectively in 24 goat (EMBL acc. no. LT592266) and 46 sheep (EMBL acc. no. LT592265) randomly chosen.

The goat *OXT* gene was found to extend over 907 bp, including 510 bp of exonic regions and 397 bp of intronic regions. The gene contains only 3 exons, ranging in size from 153 bp (exon 1) to 202 bp (exon 2), and 2 introns of 307 bp (intron 1) and 90 bp (intron 2). The 5'UT (untranslated) region comprises the first 33 nucleotides of the first exon, while the 3'UT region includes the last 102 nucleotides of the 3rd exon.

The ORF (Open Reading Frame) region codes for the peptide leader and for 106 amino acids of the 'oxytocin-neurophysin I' complex. The first exon encodes a translocator signal (19 amino acid residues encoded by the nucleotides 34–90), the nonapeptide hormone (from nucleotide 91 to 117), the tripeptide processing signal (GKR) (from nucleotide 118 to 126), and the first 9 out of the 94 residues of the neurophysin I; the second exon encodes the central part of neurophysin I and the third exon (155 bp) encodes the COOH-terminal region of neurophysin I.

The nonapeptide is characterised by a disulphide bridge between the Cys residues 1 and 6. This results in a cyclic part made of six-amino acid and a tail of three-residue α -amidated in the COOH-terminal. The isoleucine in position 3 is essential for stimulating *OXT* receptors, whereas the Cys-1 and Tyr-2 in the *OXT* molecule are the principal neurophysin binding residues (Gimpl & Fahrenholz, 2001).

The methionine initiation codon is indicated at nucleotides +34/+36 and the translation stop codon TGA is located at nt 54–56 of the exon 3. All splice junctions follow the 5' GT/3' AG splice rule. Moreover, the polyadenylation signal (AAATAAA) and polyadenylation site were identified in the 3'UT region (3rd exon), located between the nucleotides +76/+78 and +86/92 from the terminator codon, respectively. This structure shows a homology of 98% with the sheep *OXT* sequence (EMBL LT592265 present work) (online Supplementary Fig. S1).

Analysis of the intronic, exonic and 3' flanking regions

The comparison of the intronic sequences for the 46 sheep under investigation showed only two and three perfect repeats of the monomer 5'-TCCC-3' downstream the

nucleotide 253 of the intron 1. Conversely, the goat *OXT* showed higher genetic variation being characterised by 6 polymorphic sites located at the intron 1: one transition (g.1151T>C), 4 transversions (g.1194T>A, g.1221C>A, g.1333C>A and g.1381C>G) and one deletion (g.1289_1300delC) (online Supplementary Fig. S1). None of the detected mutations affects canonic splicing sites, therefore no influence on the *OXT* gene expression is expected for these polymorphisms.

The comparison of the exonic sequences of the investigated sheep showed 3 synonymous substitutions. The first was located at the position 186 of the exon 2 (g.1616C>G) and the other two at the position 62 and 147 of the exon 3, respectively (g.1784G>A and g.1879T>A). The last two mutations are located in the 3' UTR (online Supplementary Fig. S1). Since the presence of the cytosine in position 186 of the exon 2 characterises also other published *OXT* sequences of sheep (EMBL acc. nos. X55131.1, NM_001009801.2, X16052.1), as well as the sequences of other species, its presence might be indicative of an ancestral condition. For this reason, the sheep *OXT* allele containing the cytosine was named *OXT A*, whereas the changed form was named *OXT B*.

The mutation g.1879T>A is very interesting because it took place inside the polyadenylation site (TTTTTT). Therefore, it is reasonable to hypothesise that the g.1879T>A transversion might, directly or indirectly, influence the stability of the mRNA and, consequently, the amount of oxytocin produced.

Regarding our investigation in goats, only one polymorphic site was detected at exon level. However, it is a transition g.1814C>T located at the 3' UTR. The BLAST alignment with the other sequences available in EMBL showed that the cytosine might be the ancestral condition. Therefore, the goat *OXT* allele containing the cytosine was named *OXT A*, whereas the changed form was named *OXT B*.

The analysis of the 3' flanking region in both the species showed one element 207 bp long (from 2087 to 2293 and from 2086 to 2291 in sheep and goat, respectively, online Supplementary Fig. S1) with a similarity of about 80% with the annotated Bov-tA3 sequences (EMBL no. AC150888). The 3' flanking region also showed genetic variability, in terms of duplication and SNPs (g.1930_1934dupGGCTC and g.2054C>T in goats; g.2052G>C and g.2123G>T in sheep). It is known that the 3' end is an important region for the control of the gene expression. Therefore, any polymorphism falling in this region may putatively affect binding sites, for example seed sequences for microRNAs (miRNA), which are known to play a fundamental role in the regulation of gene expression in many physiological pathways including the lactation.

OXT promoter analysis

Compared to the buffalo (EMBL acc. nos. AM234538; AM234539) and similarly to bovine, the goat and sheep

OXT gene promoter sequences (958 and 960 nucleotides, respectively) are characterised by a deletion of 26 bp which occurs between the nucleotides –892/893 (numbering is relative to the first nucleotide of the sheep first exon, +1). Most probably, this event was originated by a 16-nt insertion (TTTCTTTTTTTTTTTA) with a 10-nt duplication of an adjacent sequence (AATTATTTTC). Moreover, both the goat and sheep sequences are characterised, unlike that of buffalo and bovine, by the insertion and deletion of two decamers, between nt –506/–495 and –305/–304, respectively (online Supplementary Fig. S2).

Apart from the above differences, the organisation of the goat and sheep *OXT* promoter region is very similar to the homologous gene regions in the other two species, including a modified RNA polymerase binding TATA box (TTAAA) located between the nucleotides –24/–29.

The pentamers TGACC and TCACC, known to constitute one half of the oestrogen- and the thyroid hormone-responsive element (Beato, 1989), are present in the proximal flanking region of sheep and goat *OXT* gene (indicated by bold letters in online Supplementary Fig. S2).

A highly conserved DNA element exists at about 160 nucleotides upstream of the transcriptional initiation site, a 'composite hormone response element' (monomeric orphan receptor binding site, CATAACCTTGACC), conserved across species (Gimpl & Fahrenholz, 2001), including the binding sequence sites of the oestrogen response elements (ERE), the COUP-TF sites and the steroidogenic factor-1 (SF-1) (Burbach et al. 2001).

A putative activator protein-1 (AP-1) site (GTGACGAA), with a single difference from the one indicated by Ivell et al. (1990) (GTGACTAA) and from the consensus sequences reported by Spandidos et al. (1988), is indicated by a box at nucleotides –618/–627. Other potential binding sites for AP-1, a CCAAT/enhancer binding protein-alpha (C/EBP alpha) and a transcription factor Oct-1 in the *OXT* promoter sequence (identified by TRANSFAC® 7.0 database) are also boxed in online Supplementary Fig. S2.

The comparative analysis of the goat promoter sequences showed 8 polymorphic sites, 7 transitions (g.154G>A, g.156T>C, g.438T>C, g.524A>G, g.579G>A, g.705C>T, g.827G>A) and 1 transversions (g.937G>T) which falls only 2 nucleotides far (–22) from the TATA box.

Similarly, 6 SNPs were identified in the promoter region of sheep *OXT* gene, 3 transversions (g.171T>A, g.698C>A, g.959G>T) and 3 transitions (g.43A>G, g.438T>C, g.919G>A) in addition to a variability of a mono- (g.97A [6]) and di-nucleotide repeats (g.653AG[2]). Two of these mutations could be responsible for alteration of gene expression. In fact, the first (g.438T>C) would alter the putative binding site of the transcription factor Oct-1, whereas the second mutation (g.959G>T) would involve the CpG-potential first splice donor site.

It is very interesting to notice that the comparison of sequences for the two examined species showed that both are characterised by the same polymorphism, the transition g.438T>C. This mutation could have been realised

following a convergent evolution (allelic lineages evolve similar features independently in separate lineages), or a genetic introgression (allelic lineages are horizontally transferred either from recipient species to donor species or in both directions). However, another possibility is that this mutation represents a Trans-Specific Polymorphism (TSP), which refers to the occurrence of identical or similar alleles in related species.

TSP is a crucial evolutionary mechanism responsible for sharing adaptive genetic variation across taxa. Although most studies dealing with TSP have concentrated only on the MHC *loci*, a few examples described in other immune genes suggest that TSP is a common and general evolutionary phenomenon. In fact, it has been documented in ABO blood system, retroviral restriction factor (TRIM5 encoding gene), zinc-finger antiviral protein (ZC3HAV1 gene) and Ladinin-1 encoding gene (Teixeira et al. 2015) in primates, or in complementary sex determiner gene in Hymenoptera (Lechner et al. 2014).

Therefore, the identification of this polymorphism could represent the first report of a DNA-level trans-specific polymorphism in goat and sheep and also the first in a hormone coding gene.

Genotype distribution of goat and sheep for the SNP g.438T > C is reported in Table 1.

Sheep *OXT* gene expression analysis

To test in vitro the effects of the SNP g.438T > C on the sheep *OXT* promoter activity, a 941 bp region from TT and CC homozygous subjects was amplified and cloned upstream a reporter gene, generating two different constructs. These constructs were used to transiently transfect HEK293 cells and the luciferase activity was measured. As shown in Fig. 1, the luciferase activity was lower when the cells were transfected with CC, indicating that the nucleotide C in position 438 negatively affects the promoter activity of the sheep *OXT* gene.

This polymorphic site falls into a potential binding site for the transcription factor Oct-1. In particular, the presence of cytosine removes the consensus site for this transcription factor (*in silico* analysis performed by Transfact 7.0 database). The protein Oct-1 is a multifunctional molecule and it belongs to transcription factors of the POU family. Presumably, it participates in the hormonal regulation of various genes, mediating the effect of hormones on gene activation or suppression (Sytina & Pankratova, 2003). In addition, mutations in the consensus sequences of the ubiquitous Oct-1 transcription factor are reported to reduce hormonal induction in different gene promoters, such as in the mouse β -casein encoding gene (*CSN2*) promoter (Dong & Zhao, 2007). Therefore, based on these considerations and keeping in mind the result of the in vitro reporter assay, it is possible to suppose an effective role of this promoter-binding site in the expression of the oxytocin gene.

The knowledge of the transcriptional and translational effects of *OXT* gene polymorphisms will provide new

Table 1. Genotyping data and allele frequencies of the SNP g.438T > C at the promoter region of the *OXT* gene in sheep and goat

Position	Sheep genotype distribution				Allele frequency		Goat genotype distribution				Allele frequency	
	CC	CT	TT	TOT	C	T	CC	CT	TT	TOT	C	T
g.438T > C	Obs.	1	7	38	0.0978	0.9022	–	3	21	24	0.0625	0.9375
	Exp.	0.4402	8.1196	37.4402			0.0938	2.8125	21.0938			
	$\chi^2 = 0.8746$						$\chi^2 = 0.1067$					

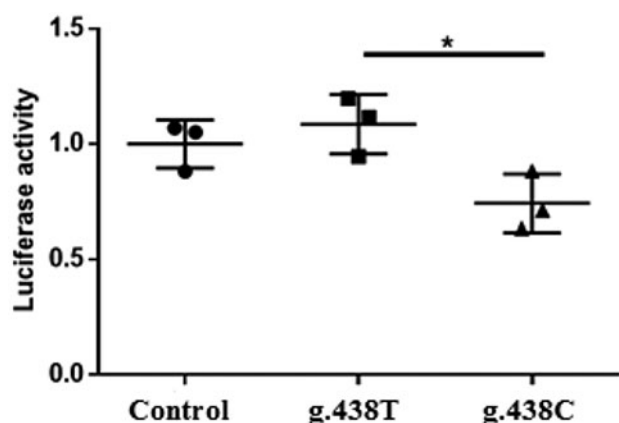


Fig. 1. The g.438T > C SNP influences the activity of the promoter OXT in sheep. Luciferase activity measured in HEK293 cells transfected with pGL control, g.438T and g.438C. pRL-TK was co-transfected and used for normalising luciferase activity. Data are shown as mean \pm SEM (the mean is presented as the longest horizontal line, whereas the two smaller lines indicate the limit of SEM) and are representative of three independent experiments (circle, square and triangle represent the three replicates of the different conditions). Data elaboration was performed using GRAPHPAD PRISM6 software. * $P < 0.05$, Student's *t*-test.

opportunities to select the best dairy goats and sheep for all the physiological processes controlled by this hormone, such as, for example, uterine smooth muscle contraction during labour, oestrous cycle length, follicle luteinisation and ovarian steroidogenesis, cognition, tolerance, adaptation and complex sexual and maternal behaviour. The detected polymorphisms represent a good opportunity to carry out studies focused on the identification of significant association with milk yield and flow in ovine and caprine species, as already performed in other species for OXT (Pauciullo et al. 2012a), or other genes candidate for quali-quantitative variations of these traits (Pauciullo et al. 2012b).

Supplementary material

The supplementary material for this article can be found at <https://doi.org/10.1017/S0022029917000097>.

This work was supported by the RARECA- PSR 214 project of the Campania Region (Italy). We thank Prof. Dr Georg Erhardt (Institute for Animal Breeding and Genetics, Justus Liebig University, Gießen, Germany) for providing German sheep and goat samples.

Conflict of interest

None of the authors has any conflict of interest to declare.

References

- Beato M 1989 Gene regulation by steroid hormones. *Cell* **56** 335–344
- Burbach JPH, Luckman SM, Murphy D & Gainer H 2001 Gene regulation in the magnocellular hypothalamo-neurohypophysial system. *Physiological Reviews* **81** 1197–1267
- Cosenza G, Pauciullo A, Mancusi A, Nicodemo D, Di Palo R, Zicarelli L, Di Berardino D & Ramunno L 2007 Mediterranean river buffalo oxytocin-neurophysin I (OXT) gene: structure, promoter analysis and allele detection. *Italian Journal of Animal Science* **6**(Suppl. 2) 303–306
- Cosenza G, Iannaccone M, Pico BA, Ramunno L & Capparelli R 2016 The SNP g.1311T>C associated with the absence of β -casein in goat milk influences CSN2 promoter activity. *Animal Genetics* **47** 615–617
- Dong B & Zhao FQ 2007 Involvement of the ubiquitous Oct-1 transcription factor in hormonal induction of beta-casein gene expression. *Biochemical Journal* **401** 57–64
- Gimpl G & Fahrenholz F 2001 The oxytocin receptor system: structure, function, and regulation. *Physiological Reviews* **81** 630–683
- Goossens M & Kan YW 1981 DNA analysis in the diagnosis of hemoglobin disorders. *Methods in Enzymology* **76** 805–817
- Ivell R, Hunt N, Abend N, Brackman B, Nollmeyer D, Lamsa JC & McCracken JA 1990 Structure and ovarian expression of the oxytocin gene in sheep. *Reproduction, Fertility and Development* **2** 703–711
- Lechner S, Ferretti L, Schöning C, Kinuthia W, Willemssen D & Hasselmann M 2014 Nucleotide variability at its limit? Insights into the number and evolutionary dynamics of the sex-determining specificities of the honey bee *Apis mellifera*. *Molecular Biology and Evolution* **31** 272–287
- Lollivier V, Guinard-Flament J, Ollivier-Bousquet M & Marnet PG 2002 Oxytocin and milk removal: two important sources of variation in milk production and milk quality during and between milkings. *Reproduction Nutrition Development* **42** 173–186
- Pauciullo A, Cosenza G, Steri R, Coletta A, Jemma L, Feligini M, Di Berardino D, Macciotta NP & Ramunno L 2012a An association between Single Nucleotide Polymorphisms at the Oxytocin locus and milk yield and flow in Italian Mediterranean river buffalo. *Journal of Dairy Research* **79** 150–156
- Pauciullo A, Cosenza G, Steri R, Coletta A, La Battaglia A, Di Berardino D, Macciotta NNP & Ramunno L 2012b A single nucleotide polymorphism (SNP) in the promoter region of River buffalo Stearoyl CoA Desaturase gene (SCD) is associated with milk yield. *Journal of Dairy Research* **79** 429–435
- Spandidos DA, Nichols RAB, Wilkie NM & Pintzas A 1988 Phorbol ester responsible H-ras1 gene promoter contains multiple TPA-inducible/AP-1-binding consensus sequence elements. *FEBS Letters* **240** 191–195
- Sytina EV & Pankratova EV 2003 Transcription Factor Oct-1: plasticity and multiplicity of functions. *Molecular Biology* **37** 637–648
- Teixeira JC, de Filippo C, Weihmann A, Meneu JR, Racimo F, Dannemann M, Nickel B, Fischer A, Halbwax M, Andre C, Atencia R, Meyer M, Parra G, Pääbo S & Andrés AM 2015 Long-term balancing selection in LAD1 maintains a missense trans-species polymorphism in humans, chimpanzees, and bonobos. *Molecular Biology and Evolution* **32** 1186–1196