

Short title: genetic variability at goat and sheep *OXT* gene

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

Gianfranco Cosenza¹, Marco Ianaccone^{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

Table S1. Oligonucleotide primers sequence and positions

Amplified regions	Nucleotide position *	EMBL acc. nos	Primers sequence (5'-3') **	TM	PCR product size (kb)
Partial promoter - partial exon 1	2134-2151	AB481096	Forward ATCTGCTGACCACTCTCC	62.6	1.170
	Complementary to: 811- 828	X55131	Reverse GTGCTGGACCTCGACGTG	65.6	
Partial exon 1 - partial exon 3	748 - 763	X55131	Forward GGCCTCCTGGCGTTGA	64.8	0.80
	Complementary to: 1523 - 1542	X55131	Reverse CTCCCTCTGTAATCATCCCC	65.5	
Exon 3 - partial 3' flanking region	1411 - 1427	X55131	Forward GGCGTCTCTCTGTGCAG	62.7	0.65
	Complementary to: 51496027 - 51496011	CP011898	Reverse CAAGGGTCTGACAGGGG	63.5	
Sheep <i>OXT</i> promoter	26-45	LT592265	Forward TCATGGCAACACTGACCAAC	63.6	0.94
	Complementary to: 949-966	LT592265	Reverse CTCGGGTCTGGCCTTTTA	62.2	
Sheep <i>OXT</i> promoter - nested fragment***	52-72	LT592265	Forward CTGAGGT ACC CTCTCTTTCTT	64.8	0.90
	Complementary to: 939-955	LT592265	Reverse CCTTTTAAG CCCCGGGGG	65.5	

* Numbering agrees with the nucleotide sequence of the corresponding EMBL acc. no. ** Primers were designed by means of DNAsis-Pro (Hitachi Software Engineering Co., Japan); ***The nested primers were modified by changing a GT to CC at the position 9-10 for the forward and the C to a G at position 15 for the reverse (modified nucleotides are shown in bold and restriction sites are underlined).

		<u>Peptide leader</u>	
	+1	M A G S S L	
-2	-----kc	AGTCCCCGGCCCGGAGCCCAGCGCTCTGCACCATGGCAGGTTCCAGCCT	50
-2	-----g	-----	50
		<u>Oxytocin</u>	
51	A C C L L G L L A L T S A	C Y I Q N C P } TGCTACATTGAGAACTGCCC	110
51	CGCCTGCTGCCTGCTCGGCCTCCTGGCGTTGACCTCCGCT		110
		<u>Neurophysin 1</u>	
111	L G } G K R { A V L D L D V R T	gtgagcgccccgcctc	160
111	CCTGGGCGGCAAGCGTCCGGTCTGGACCTCGACGTGCGCAGG		160
161	gtcccggtggctctcggggctggcagggccgctgccacagggcgcccccgccgccccctt		220
161	-----g--g-----y-----c-----		220
221	tcccgccctgaccgcggtaccggccccacctagcctgggaatcgagggagcggaggagctt		280
221	-----yg-----m-----		280
281	ttgactgcctccttcgaccaattctgagcccaaagagagcggggaagaccgcatctc		340
281	-----g-----c--g-----		340
341	ccgcgctcctcaggccgcccccgcccggctcagccccccaccccacaggggtctccc		400
341	-----m-----		400
401	tccccggccggtccccctcccgccccggctcatccccgcctcccgccag TGTCTCCCCT		460
401	-----s-----		460
461	C G P G G K G R C F G P S I C C G D E L		
461	GCGCCCCGGGGCAAAGGCCGCTGCTTCGGGCCAGCATCTGCTGCGGGGACGAGCTGG		520
461	-----		520
521	G C F V G T A E A L R C R E E N Y L P S		
521	GCTGCTTCGTGGGCACGGCCGAGGCGCTGCGCTGCCGAGAGGAGAACTACCTGCCGTGCG		580
521	-----A-----		580
		Q	
581	P C Q S G Q K P C G S G G R C A A A G I		
581	CCTGCCAGTCCGGCCAGAAGCCCTGCGGGAGCGGGGCGCTGCGCCGCCGCGGSATCT		640
581	-----C-----		640
641	C C S P		
641	GCTGCAGCCCGG gtgagtcggcaggggcccagagcggggccggggctccggaaccagggcg		700
641	-----g-----		700
701	ggccgggcccaggggtggccctgactcggcgctctctctgtgcag ACGGCTGCCACGCGGATC		760
701	-----		760
761	P A C D P E A A F S Q H } *		
761	CCGCTGCGACCCCGAGGCCGCTTCTCCAGCACTGAGACCGRCCGGCCCCTGACACCA		820
761	-----G-----		820
		<u>Polyadenylation signal</u>	
821	TCGGAGCACAGCCCTCACTCCCTCTGTAATCATCCCCAGGAATTATGACAATGAAATAAA		880
821	-----G--A---Y-----		880
		<u>Polyadenylation site</u>	
881	GCCTTTTTTWTCCCCTCC aacaagcctcgtgtctgagtgctcagaacggggagggagggct		940
881	-----T-----		940
941	ttgggggaattcaagatcctcggctctcggcgcttc aacggaagtagagccgcggtcca		999
941	-----c-----a-----tc-----		1000
1000	ggcgccatcctataggggtggatggaagaaggggtggaggtaccacacggcgggctccc		1059
1001	-----g-----g-----		1060

```

1060  aggcatatggaaatagccccastttccaactccaaacaggctggacataggaggaaCTT  1119
1061  -----a-----c-----g--y-----  1120

1120  GGTGGCTCAGACGGTAAAGAGTTTGCCTGCAGTKCGAGAGACCCGAATTCGATCCCTGGG  1179
1121  -----G-----G-----  1180

1180  TTGGGAAGACACCCTAGAGAAGCGGCAACCCACTCCAGTATTCTGGCTTGGAAAATCCCA  1239
1181  -----T-----T-----  1240

1240  TGGACGGAGAGCCTGGCTGGCTACCGTCCATGGGGGTGCGCAAAGCGTCGGACACGACTC  1299
1241  -----  1299

1300  AGCGACTTCACTTTCTTTCACTTTctgaggtctaaagggagagttgggcagacaatgaca  1359
1300  -----  1359

1360  agggctctgacagggga  1375 Sheep (EMBL acc. no. LT592265)
1360  -----  1375 Goat (EMBL acc. no. LT592266)

```

Figure S1. Homology between the complete nucleotide (nt) sequences of ovine and caprine Oxytocin-Neurophysin I encoding (*OXT*) gene. Numbering is relative to the first nucleotide of the first exon (+1), dashes represent nt identical to those in upper lines and the stop codon is symbolized by *. Artiodactyla retroposon sequence is in uppercase and italic. Peptide leader, Polyadenylation signal and Polyadenylation site are underlined.

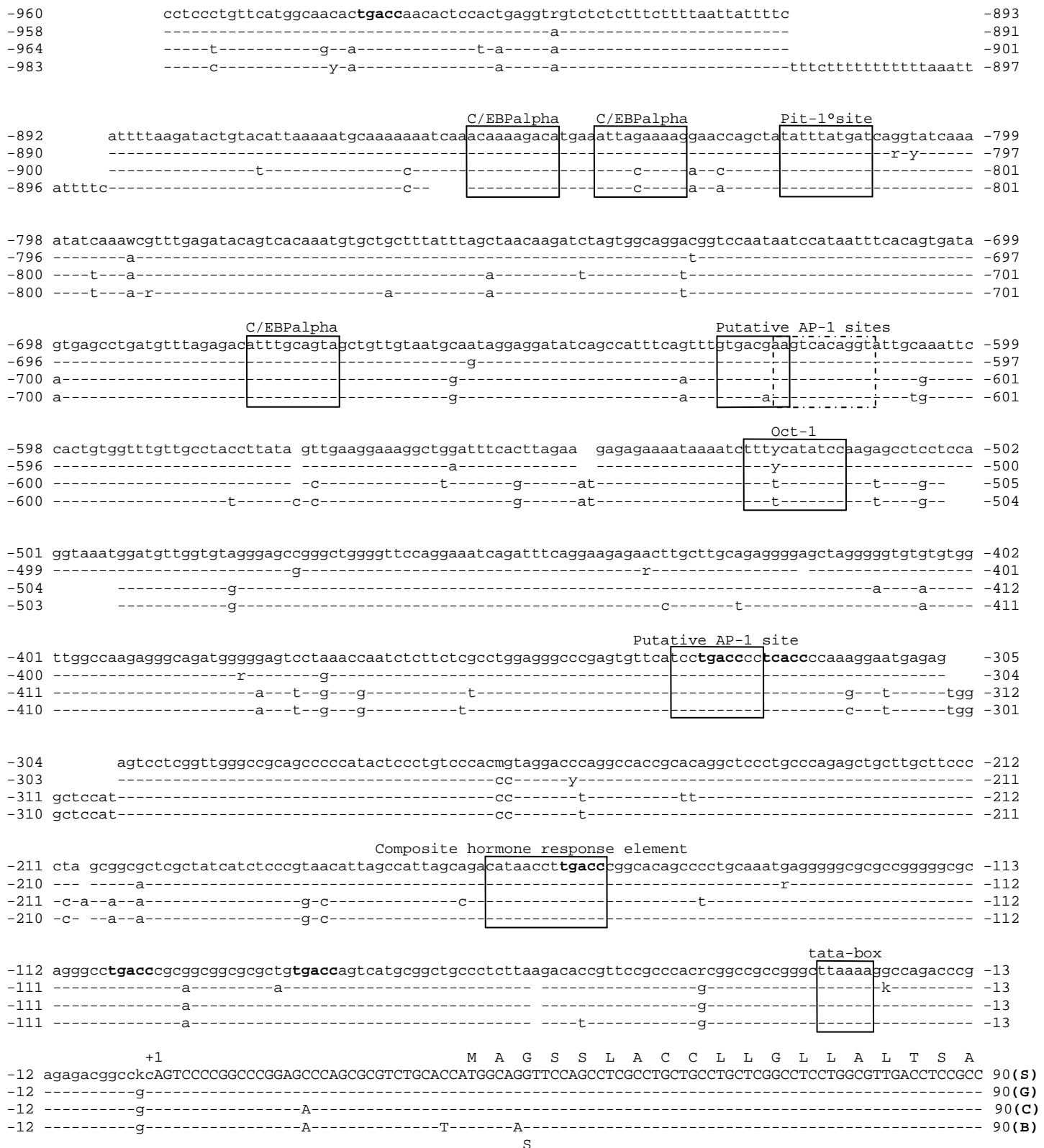


Figure S2. Homology between the nucleotide (nt) sequences of the 5' flanking region and partial exon 1 (3'UTR and peptide leader encoding region) of sheep (S) (EMBL acc. no. LT592265), goat (G) (EMBL acc. no. LT592266), cattle (C) (EMBL acc. no. AB481096) and buffalo (B) (A and B alleles, EMBL acc. nos. AM234538; AM234539) *OXT* gene. Numbering is relative to the first nucleotide of the first exon (+1) and dashes represent nt identical to those in upper lines. Congruent

and putative factors are in shaded bold letters and boxed. 3'UTR and peptide leader encoding region are in uppercase.