INNSPUB

International Journal of Biosciences | IJB |
ISSN: 2220-6655 (Print) 2222-5234 (Online)
http://www.innspub.net
Vol. 4, No. 1, p. 399-406, 2014

RESEARCH PAPER OPEN ACCESS

Bioinformatics analysis of *BMPR1B* gene in different species

Ali Reza Eivazi[1*], Mostafa Modarresi[2]

Department of Agriculture, Shahre-Rey Branch, Islamic Azad University, Tehran, I. R. Iran

Key words: Bioinformatics, BMPR1B gene, Genetic diversity, multiple-births and Phylogeny tree.

http://dx.doi.org/10.12692/ijb/4.1.399-406 Article published on January 11, 2014

Abstract

Multiple-births is one of the most animals important trait. Many genetic and environmental factors affect on multiple-births. Bone morphogenetic protein receptor IB (BMPR-IB) is one of the genetic factors affecting this trait in sheep. In this study, 72 different sequences were extracted from NCBI and then some bioinformatics analysis were investigated such as CDS length, stop codon variation, homology search, molecular mass, signal peptide, isoelectric point, tertiary structure, transmembrane, motifs and phylogenetic tree. This gene has been had 4 motifs such as Activin_recp, transmembrane, GS motif and STYKc Protein kinase. Also the results were indicated that BMPR1B in sheep has closely relationship with goat and cow. Multiple alignment showed high conservation of motifs which indicate constancy of this gene during evolution. The results of this study can be used to better recognition of BMPR-1B gene structure in different animals which would be helpful for animal breeding plans.

* **Corresponding Author:** Ali Reza Eivazi ✉ a.eyvazi88@yahoo.com

Introduction

In recent years, improving reproductive traits in sheep has been highly considered by producers. One of these characters is number of offspring per birth. Like ranchers, scientists prefer use of multiple-births animals rather than single birth (Eghbalsaied *et al.*, 2009). Multiple-births in sheep has low heritability coefficient (less than 10%), therefore enough recognition of the genetic and environmental factors influencing this traits very important (Cole and Cupps, 1991).

Today it has been known that three group genes affect on follicles growth and ovulation rate, including ALK9, GDF9 and BMP. All these genes belong to a large family of TGF β that are effective on the expression and secretion of necessary hormones on follicular growth and ovulation rate. Growth factors BMP15 ، BMP6 and GDF9 are produced by oocytes (McNatty *et al.*, 2005), while BMP receptors type IA, IB and type II are located on oocytes and follicular somatic cells (Souza *et al.*, 2002) and also on sheep pituitary cells (Faure *et al.*, 2005).

Previous study on Booroola Merino race and its hybrids indicated that an autosome major gen (FecB or BMPR-1B) has additive effects on ovulation rate and partial dominance on multiple-births. Totally, the sheep which are received one copy of this gene from each parents, can be generated 1.5 oocytes and one lamb more than others in any generation (Davis Gh, 1983). Two decades later, three different groups of researchers discovered that existence of multiple-births in the Booroola Merino race was due to mutation of bone morphogenetic protein receptor gene {serine/threonine receptor kinase (BMPR-1B)} (Souza *et al.*, 2001; Wilson *et al.*, 2001; Chu *et al.*, 2007).

Bone morphogenetic protein family has been identified as transmembrane serine-threonine kinase receptors include two receptors, type I (IA and IB) and type II (Rosenzweig *et al.*, 1995; Naber *et al.*, 2012). BMP receptor type 1B (BMPR1B) is a latent receptor for BMP15, which adjust follicular growth and ovulation rate. This gene is mapped on sheep chromosome 6 and also is called as activin-like kinase 6 (ALK6) (Ho and Bernard, 2009). The one common mutation point in the BMPR-1B is known in Booroola, Garole, Hu, Han, and Javanese sheep (Davis *et al.*, 2002; Davis, 2005; McNatty *et al.*, 2005). An increasing ovulation rate and litter size in ewes who were heterozygous for this mutation and more increase in homozygous ewes were observed.

Considering the fact that BMPR-1B gene firstly recognized in Booroola Merino and importance of multiple births in sheep, gene diversity and polymorphism of this gene is mainly centralized in sheep (Polley *et al.*, 2010; Luong *et al.*, 2011; Ganai *et al.*, 2012).

In the present study, the bioinformatics analysis was used to study BMPR-1B gene diversity, structure, polymorphism, function and phylogenetic relationships of BMPR-1B in several animals using large amounts of stored biological data in bioinformatical databases.

Material and methods

Data collection and information resources*:*

In this study, the 72 different entries were used with their protein and DNA sequences which selected from National Center for Biotechnology Information (NCBI) from URL: (http://www.ncbi.nih.gov). All studied CDs sequences were complete.

Identification of motifs and homology determination

MEME software (http://meme.nbcr.net) was used to identify protein sequence motifs. Parameters of this analysis consist of: number of repetitions (any), Maximum number of motifs to find (4) and optimum width of motif was ≥15. Also TMHMM inter membrane motifs were identified by using of TMHMM Server (http://www.cbs.dtu.dk/services/TMHMM) with default settings. Then, the secondary structure and homology of BMPR-1B gene were examined by using

the Jalview software. Finally, protein 3D structure was calculated using SWISS-MODEL software and was observed using Pymol software.

Phylogenetic analysis

Genetic evolution is a scientific process that gives us possibility of determining the evolutionary history of species group or sequences. Evolutionary relationships among a set of sequences can be studied using phylogenetic methods, so the conformity of all protein sequences of BMPR1B domain were done. MAFFT software was used to obtain L-INS-I, that is one of the most accurate methods of MSA (Multiple protein Sequence Alignment). Phylogenetic tree was drawn by using MEGA5 software with 100 replications of Bootstrap analysis.

Result and discussion

Variation between stop codons

As shown in Table 1, the terminal codons of 31 sequences from five different species were analyzed. The results were shown that the TGA was terminal codon for all sequences. It was compatible with the results of Sun *et al.* (2005), who has suggested the TAG terminal codon found in primitive eukaryotes, TAA in simpler eukaryotes such as yeast and invertebrates and finally the TGA exists in complex eukaryotes, especially in the vertebrates. These different stop codon have different efficiencies, TAA and TGA have highest and lowest performance respectively (Tate *et al.*, 1999). Variation in terminal codon will cause phenotypic variation, for example, TGA causes variation through effect on protein expression and activation (Zheng *et al.*, 2010).

Table 1. Polymorphic information of BMPR1Bgene for each species.

species	N	h	*H*d	*K*	π	π_s	S	SP	PIP	NS2	NP2	Stop codon	CDS length
Homo sapiens	11	2	0.327	375.709	0.24898	0.24898	1148	0	1148	0	1148	TGA	1509(9) 1599(2)
Ovis aries	4	4	1.000	2.500	0.00166	0.00166	5	5	0	5	0	TGA	1509
Capra hircus	9	6	0.889	4.167	0.00276	0.00276	18	17	1	17	1	TGA	1509
Sus scrofa	4	3	0.833	2.667	0.00177	0.00177	5	4	1	4	1	TGA	1509
Equus caballus	3	1	0.000	0.000	0.0000000000	0.00000	0	0	0	0	0	TGA	1509
all	31	16	0.903	213.475	0.14147	0.14147	1215	5	1210	5	112	-	-

Note: N (No. of sequence), h (No. of haplotypes), Hd (haplotype diversity) and K (average number of nucleotide differences), π (nucleotide diversity), π_s(synonymous nucleotide diversity), S(No. of polymorphic sites), SP(singleton variable sites), PIP(parsimony informative sites), NS2 (No. of singleton variable sites with 2 variants) and NP2 (No. of parsimony-informative sites with 2 variants).

Variation in CDS *length within and among animal species*

The profile of CDS region for BMPR-1B gene within and between species has been shown in Table 1. CDS length for all species except *Homo sapiens* were equal 1509 bp. In *Homo sapiens* nine sequences had 1509 bp length and two sequences had 1599 bp length. In these two cases, using the database from http://www.signalpeptide.de, it was determined that MGWLEELNWQLHIFLLLILLSMHTRANFLDN amino acid sequences were added to the protein N-terminal. Here, 25 elementary amino acids (1-25) play the role as signal peptide (Fig.2). Typically the signal peptides are sequences of peptide with 5-30 bp lengths of amino acids that are important in secretion pathways to organelles such as the endoplasmic reticulum, Golgi apparatus or out of cells. In this study, the signal peptides have been added to the target protein

for secretion into the extracellular environment. Variation in CDS usually occurs due to deletion, addition and proteins different splicing and then polymorphism subsequently is created. Species that have greater diversity in their CDS have higher phenotypic variations. And this feature can be used to study the relationship between genetic variations and existing phenotype.

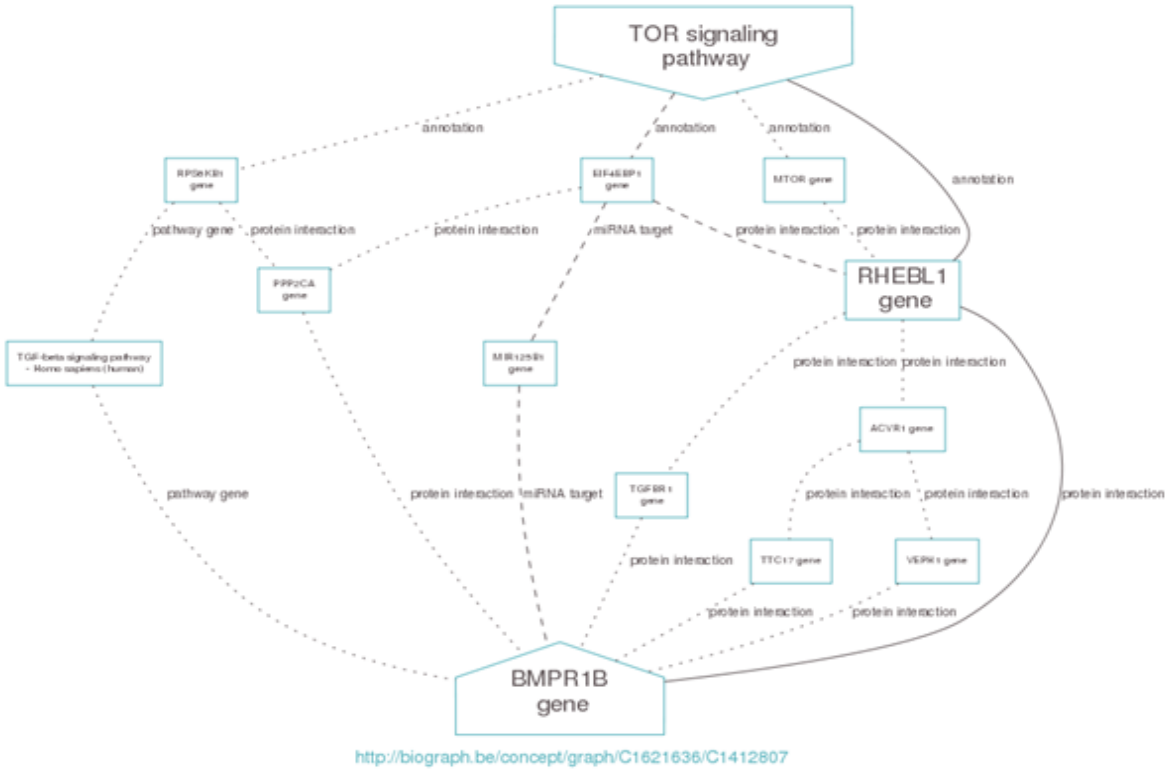

Fig. 1. BMPR1Bbiosynthesis pathway in animals.

Genetic diversity between and among species

Multiple alignments for 72 sequences of the nucleotide and protein from different species were performed using BioEdit and MEGA5.0 software (Fig. 3). The results showed that studied sequences had overall high similarity to each others. Sheep sequence (NM_001009431.1) had highest similarity with corresponding sequences in cow (NP_001098798.1) and European rabbit (NP_001098798.1) with 99% identity. Also this sequence had least similarity (70%) with the corresponding sequences in the toad (NP_001081209.1) and zebra fish (NP_001004585.1). Then the CDS sequences (Table 1) were used for DnaSP software analyses. Among the 31 examined sequences, different species had 1509 sites (excluding sites with gaps / missing data) which 1215 sites of them were polymorphic (segregating) sites including 5 singleton variable sites and 1210 parsimony informative sites. All singleton variables had two variants while 1210 Parsimony informative sites including 1093 with two variants, 112 with three variants and 5 with four variants. Also for all sequences observed that nucleotide diversity (π) equal 0.14147 and average number of nucleotide differences (k) equal 213.475. In addition, the nucleotide diversity and polymorphic species are listed separately at Table 1, too. Different species have different variety of genes for BMPR1B. Usually higher genetic variation is more useful for natural selection in nature. In studied samples, average number of nucleotide differences and nucleotide diversity were highest in *Homo sapiens* and lowest in *Equus caballus*. Secondary structure of sheep protein sequence (NM_001009431.1) was calculated using Jalview and SOPMA software and determined that 34.66% of the BMPR1B protein sequence was as

Alpha helix, 13.75% as extended strand and 50.00% as Random coil. Thus, random coil is dominance of structural elements in this gene.

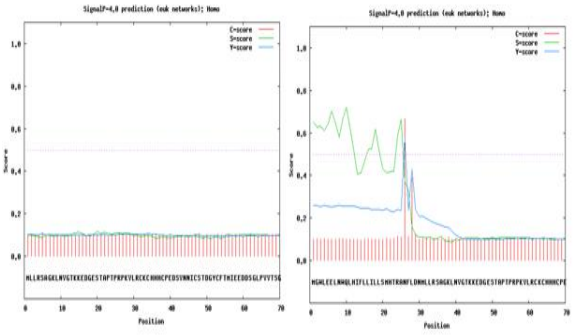

Fig. 2. Signal peptide analysis for BMPR1B proteins. Left, CDS within 1599 bp from human BMPR1B with accession number NM_001256793and right with 1509 bp, accession number NM_001256792.

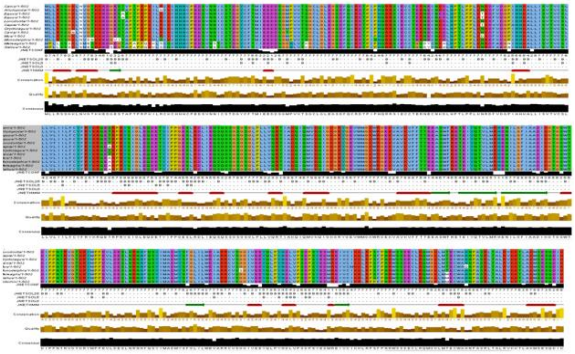

Fig. 3. Multiple alignment and secondary structure of BMPR1B gene in different species.

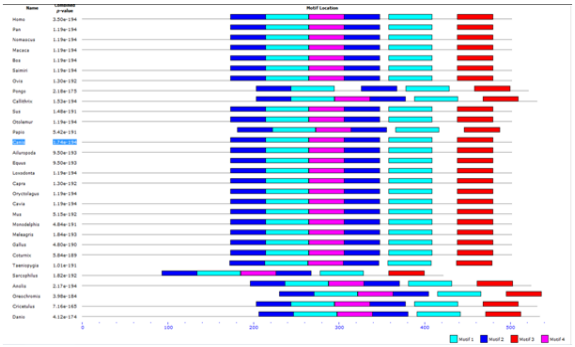

Fig. 4. Motifs for BMPR1B proteins. The MEME motifs are shown as different-colored boxes.

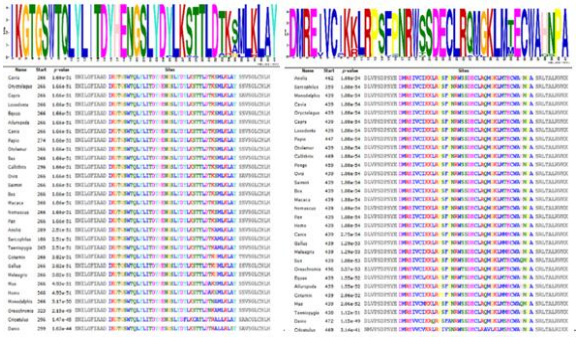

Fig. 5. MEME motifs for BMPR1B proteins.

Hydrophobicity and Hydrophilicity profile, Phylogenetic tree and SMART analysis

Hydrophobicity analysis of the NM_001256793 sequence by Prot Scale software showed that large parts of the protein are hydrophilic, since this protein is secretary and it's reasonable. In the other hand, signal peptide sequence in the N-terminal region caused a hydrophobilic peak. Analysis of this protein by SMART software leads to detect 4 active areas of amino acid sequences (30–110: Activin_recp, 126–148: transmembrane, 174–204: GS motif and 204–493: STYKc Protein kinase). Activin receptor is receptor-type kinases of Ser/Thr type family. These receptors have a transmembrane domain and a specific hydrophilic Cys-rich ligand-binding domain. Terminal portion of the extra cellular domain is protected. These receptors have 9 amino acids cysteine box with a consensus sequences CCX {4-5} CN (Attisano and Wrana, 1996). In this research, consensus sequence was as CCTERNECN. GS motif is engrossed an area of over 30 amino acids. It had a helix-loop-helix domain structure and phosphorylation activities. STYKc Protein kinase which made up the bulk of the protein sequence had phosphotransferase activity. Also MEME software analysis used to finding conserved motifs (Figs. 4 and 5). ProtParam software data analysis indicated that the sequence with Accession number NM_001009431.1 had 56.9304 KDa molecular weight and isoelectric point 7.78. In proteins with signal peptides, 3.3669 KDa added to molecular weight. Tertiary structure of the protein was calculated with the SWISS-MODEL software (Fig. 6) and high similarities were observed among the tertiary structures in human, monkeys, goats, cattle and sheep. The analysis by TMHMM software indicated that this gene in sheep had one transmembrane domain among amino acid 127 to 149 (Data not showed).

Phylogenetic tree of amino acid sequences was designed to determine the evolutionary relationships among different species (Figure7). Phylogenetic tree was divided into two different branches, one branch is

the fish *Danio rerio* with accession number NM_001145996.1 and other branch involve the another sequences. This subcategory also divided into two branches, birds and mammals. It was observed that in the sheep, goats and cows had most closely relationship. Also, human, monkey and chimpanzee were placed in a same directory.

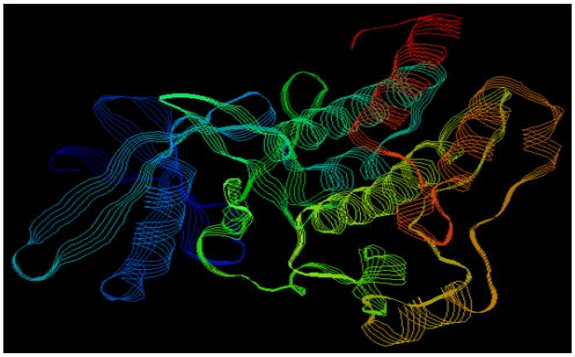

Fig. 6. The 3D structure of Sheep BMPR1B protein (NM_001009431.1) predicted by SWISS-MODEL software.

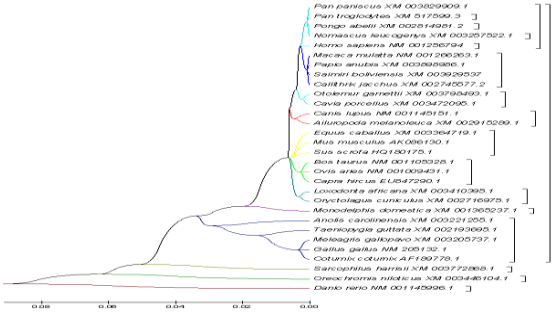

Fig. 7. The phylogenetic tree of BMPR1B genes from various species using the MEGA 5.0 software.

Conclusion

In this study according to results of the 31 different studied CDS sequences, it is determined that TGA was terminal codon for all sequences .CDS length for all species were equal 1509 bp as shown in Table 1 except *Homo sapiens*. In *Homo sapiens* nine sequences had 1509 bp length and two sequences had 1599 bp length. Multiple alignments for 72 sequences of the nucleotide and protein from different species showed that studied sequences had overall high similarity to each others. Sheep sequence had highest similarity with corresponding sequences in cow and European rabbit with 99% identity while it had least similarity (70%) with the corresponding sequences in the toad and zebra fish. Also, every species have variation in BMPR1B gene sequence and average number of nucleotide differences and nucleotide diversity were highest value in *Homo sapiens* and lowest value in *Equus caballus*. Other results showed that secondary structure of sheep BMPR1B protein included 34.66% of Alpha helix, 13.75% of extended strand and 50.00% of Random coil. Thus, random coil is dominance of structural elements in this gene. Results showed that large parts of the protein are hydrophilic and it led to detect of 4 active areas of amino acid sequences. Also high tertiary structure similarities were observed among the tertiary structures in human, monkeys, goats, cattle and sheep. Phylogenetic tree was divided into two different branches, which one branch was the fish *Danio rerio* and other branch involved the remained sequences and this shown similarity among amino acid sequences of this gene in many species.

Reference

Attisano L, Wrana JL. 1996. Signal transduction by members of the transforming growth factor-beta superfamily. Cytokine Growth Factor Review **7,** 327-339.
http://dx.doi.org/10.1016/S1359-6101(96)00042-1

Chu MX, Liu ZH, Jiao CL, He YQ, Fang L, Ye SC, Chen GH, Wang JY. 2007. Mutations in BMPR-IB and BMP-15 genes are associated with litter size in Small Tailed Han sheep (Ovis aries). Journal of animal science **85,** 598-603.
http://dx.doi.org/10.2527/jas.2006-324

Cole HH, Cupps PT, 1991. Reproduction in Domestic Animals. Academic Press Inc.

Davis GH. 2005. Major genes affecting ovulation rate in sheep. Genetics Selection Evolution 37 Suppl **1,** S11-23.
http://dx.doi.org/ 10.1051/gse:2004026

Davis GH, Galloway SM, Ross IK, Gregan SM, Ward J, Nimbkar BV, Ghalsasi PM, Nimbkar C, Gray GD, Subandriyo, Inounu I, Tiesnamurti B, Martyniuk E, Eythorsdottir E, Mulsant P, Lecerf F, Hanrahan JP, Bradford

GE, Wilson T. 2002. DNA Tests in Prolific Sheep from Eight Countries Provide New Evidence on Origin of the Booroola (FecB) Mutation. Biology of reproduction **66,** 1869-1874.
http://dx.doi.org/ 10.1095/biolreprod66.6.1869

Davis Gh KRW. 1983. Segregation of a major gene influencing ovulation rate in progeny of Booroola sheep in commercial and research flocks. New Zealand Society of Animal Production.

Eghbalsaied S, Ghaedi K, Forouzanfar M, Hajian M, Hosseini SM, Nasr-e-Esfahan M. 2009. Science and technology of farm animal Transgenesis. Yakhteh Medical Journal **11,** 78-87.

Faure M-O, Nicol L, Fabre S, Fontaine J, Mohoric N, McNeilly A, Taragnat C. 2005. BMP-4 inhibits follicle-stimulating hormone secretion in ewe pituitary. Journal of Endocrinology **186,** 109-121.
http://dx.doi.org/10.1677/joe.1.05988

Ganai TAS, Misra SS, Shabir M. 2012. Polymorphism analysis of BMPR1B gene by forced RFLP and PCR-SSCP techniques and expression of the mutation in introgressed sheep. Tropical Animal Health and Production **44,** 277-283.
http://dx.doi.org/10.1007/s11250-011-0015-y

Ho CC, Bernard DJ. 2009. Bone Morphogenetic Protein 2 Signals via BMPR1A to Regulate Murine Follicle-Stimulating Hormone Beta Subunit Transcription. Biology of reproduction **81,** 133-141.
http://dx.doi.org/10.1095/biolreprod.108.074211

Luong HTT, Chaplin J, McRae AF, Medland SE, Willemsen G, Nyholt DR, Henders AK, Hoekstra C, Duffy DL, Martin NG. 2011. Variation in BMPR1B, TGFRB1 and BMPR2 and control of dizygotic twinning. Twin Research and Human Genetics **14,** 408.
http://dx.doi.org/

McNatty KP, Galloway SM, Wilson T, Smith P, Hudson NL, O'Connell A, Bibby AH, Heath DA, Davis GH, Hanrahan JP, Juengel JL. 2005. Physiological effects of major genes affecting ovulation rate in sheep. Genetics, selection, evolution GSE 37 Suppl **1,** S25-38.
http://dx.doi.org/10.1051/gse:2004029

Naber HH, Wiercinska E, Pardali E, Laar T, Nirmala E, Sundqvist A, Dam H, Horst G, Pluijm G, Heckmann B, Danen EJ, Dijke P. 2012. BMP-7 inhibits TGF-β-induced invasion of breast cancer cells through inhibition of integrin β3 expression. Cell Oncol **35,** 19-28.
http://dx.doi.org/10.1007/s13402-011-0058-0

Polley S, De S, Brahma B, Mukherjee A, Vinesh P, Batabyal S, Arora JS, Pan S, Samanta AK, Datta TK. 2010. Polymorphism of BMPR1B, BMP15 and GDF9 fecundity genes in prolific Garole sheep. Tropical Animal Health and Production **42,** 985-993.
http://dx.doi.org/10.1007/s11250-009-9518-1

Rosenzweig BL, Imamura T, Okadome T, Cox GN, Yamashita H, ten Dijke P, Heldin CH, Miyazono K. 1995. Cloning and characterization of a human type II receptor for bone morphogenetic proteins. Proceedings of the National Academy of Sciences **92,** 7632-7636.

Souza C, Campbell B, McNeilly A, Baird D. 2002. Effect of bone morphogenetic protein 2 (BMP2) on oestradiol and inhibin A production by sheep granulosa cells, and localization of BMP receptors in the ovary by immunohistochemistry. Reproduction **123,** 363-369.
http://dx.doi.org/10.1530/rep.0.1230363

Souza C, MacDougall C, Campbell B, McNeilly A, Baird D. 2001. The Booroola (FecB) phenotype is associated with a mutation in the bone morphogenetic receptor type 1 B (BMPR1B) gene. Journal of Endocrinology **169,** R1-R6.
http://dx.doi.org/10.1677/joe.0.169R001

Sun J, Chen M, Xu J, Luo J. 2005. Relationships Among Stop Codon Usage Bias, Its Context, Isochores, and Gene Expression Level in Various Eukaryotes. Journal of molecular evolution **61,** 437-444.
http://dx.doi.org/10.1007/s00239-004-0277-3

Tate WP, Mansell JB, Mannering SA, Irvine JH, Major LL, Wilson DN. 1999. UGA: a dual signal for 'stop' and for recoding in protein synthesis. Biochemistry. Biokhimiia **64,** 1342-1353. h.

Wilson T, Wu X-Y, Juengel JL, Ross IK, Lumsden JM, Lord EA, Dodds KG, Walling GA, McEwan JC, O'Connell AR, McNatty KP, Montgomery GW. 2001. Highly Prolific Booroola Sheep Have a Mutation in the Intracellular Kinase Domain of Bone Morphogenetic Protein IB Receptor (ALK-6) That Is Expressed in Both Oocytes and Granulosa Cells. Biology of reproduction **64,** 1225-1235.
http://dx.doi.org/10.1095/biolreprod64.4.1225

Zheng H, Li X, Zhou R, Li L, Guo X, Kang J, Li D. 2010. Bioinformatics analysis of tyrosinase-related protein 1 gene (TYRP1) from different species. Frontiers of Agriculture in China **4,** 109-115.
http://dx.doi.org/10.1007/s11703-009-0081-3