Revista de Biotecnologia & Ciência, v.10, n.1, 2021
ISSN 2238-6629
Universidade Estadual de Goiás

The non-synonymous polymorphism Glu320Pro affects the functional stability and secondary structural conformation of the goat mature GDF9 protein

O polimorfismo não-sinônimo Glu320Pro afeta a estabilidade funcional e a conformação estrutural secundária da proteína GDF9 madura de cabra

Jane Eyre Gabriel[1*]

[1]Colegiado de Medicina, Universidade Federal do Vale do São Francisco, Petrolina, Pernambuco, Brasil.
*Autor correspondente. Jane Eyre Gabriel, Colegiado de Medicina, Universidade Federal do Vale do São Francisco UNIVASF, Petrolina/PE, Brasil - E-mail: janegabriel@ufscar.br

Recebido: 07/07/2020; Aceito: 22/09/2020

RESUMO

Mutações nas sequências de nucleotídeos são frequentemente responsáveis por afetar drasticamente a estabilidade funcional e estrutural de polipeptídeos devido a alterações na suas composições de aminoácidos. O objetivo desse estudo foi avaliar comparativamente os efeitos do polimorfismo Glu320Pro na estabilidade estrutural e funcional, bem como na conformação estrutural secundária da proteína GDF9 madura de cabra por análises comparativas de predição proteica. Uma sequência selvagem de resíduos de aminoácidos da proteína GDF9 de cabra (*Capra hircus*, Q66NC0) pesquisada a partir de um banco de dados e uma sequência mutante carreando a substituição do aminoácido glutamina por prolina na posição 320 (Glu320Pro) foram selecionadas. O polimorfismo Glu320Pro resultou na inferência de danos na estabilidade funcional e estrutural da proteína GDF9 de cabra. Variações na sua estrutura secundária demonstraram a presença de um domínio alfa hélice (H) e um domínio espiral espiralado (C) na posição alvo na proteína GDF9 selvagem e mutante, respectivamente. Tais descobertas predizem que o padrão conformacional diferenciado da proteína GDF9 mutante pode estar exercendo influências sobre sua ação biológica durante a foliculogênese em cabras.

Palavras-chave: Alterações de aminoácidos, Bioinformática, Predição de proteína.

ABSTRACT

Mutations in the nucleotide sequences are often responsible for dramatically affecting the functional and structural stability of polypeptides due to changes in their amino acid compositions. The purpose of this study was to comparatively assess the effects of the polymorphism Glu320Pro on structural and functional stabilities, and secondary structural conformation in the goat mature GDF9 protein by *in silico* comparative protein prediction analyses. A wild sequence of amino acid residues of the goat GDF9 protein (*Capra hircus*, Q66NC0) searched from biological database and a mutant sequence harboring the substitution of the glutamine to proline residues at position 320 (polymorphic Glu320Pro) were selected. The polymorphism Glu320Pro resulted in damages in the functional and structural stabilities of the goat GDF9. Variations in the secondary structure demonstrated the presence of an alpha helix motif (H) and a coiled coil domain (C) in target position of wild and mutant GDF9, respectively. These findings predict that the differentiated conformational pattern of the mutant GDF9 protein may exert influences on its biological action during folliculogenesis in goats.

Keywords: Amino acid changes, Bioinformatics, Protein prediction.

INTRODUCTION

The growth and differentiation factor 9 (GDF9) constitutes a crucial member of the transforming growth factor-beta superfamily highly expressed in oocytes, exerting a pivotal influence on the surrounding somatic

cells, particularly granulosa, cumulus and theca cells during ovarian folliculogenesis in female mammals. The deletion of the gene *gdf9* results in decreased ovary size, halted follicular development at the stage of the primary follicle and the absence of any corpus lutea (DONG et al., 1996). *In vitro* exposure of mammalian ovarian tissue to GDF9 promotes primary follicle progression (NILSSON, 2002) and GDF9 stimulates growth of preantral follicles by preventing granulosa cell apoptosis (ORISAKA et al., 2006). However, GDF9 is expressed in all phases of follicle development with higher levels of expression in small than large antral follicles in goats (SILVA et al., 2005; PRAMOD et al., 2013). In addition to having the highest expression in the ovary, GDF9 gene is also widely expressed in 20 tissues (hypothalamus, pituitary, uterus) in goats, which indicates that it can differentially affect some physiological pathways, metabolism, and phenotypic expression (PAN et al., 2015).

Several efforts have focused on identification of single nucleotide polymorphisms SNPs in the gene *gdf9* associating them with reproductive traits in different livestock species, especially goat and sheep. Cambridge and Belclare sheep breeds showed a range of fertility phenotypes due to eight single nucleotide polymorphisms (SNP) across the coding region of the gene *gdf9* (HANRAHAN, 2003). According to SILVA et al. (2011), a new polymorphism in the gene *gdf9* named FecG(E) (Embrapa), which leads to a substitution of a phenylalanine with a cysteine in a conservative position of the mature peptide, is associated with increased ovulation rate and prolificacy in homozygous sheep. Recently, WANG et al. (2019) assumed that non-synonymous Glu320Pro mutation (also known as g.3905A > C) in the gene *gdf9* is the major SNP to affect goat litter size in Shaanbei white cashmere goats, since Glu320Pro was significantly associated with the greatest first-born litter size in goats irrespective of the sample size ($n = 1511$; $P = 0.008$). Thus, this DNA marker emerges as a stronger tool in the marker-assisted selection for breeding in relation to fecundity in ruminant species. Considering that Bioinformatics' tools have provided previous approaches to assess if changes in the nucleotide residues could result in a polymorphism with positive or negative effects on the structure and stability of a determined protein, the curent study aimed to predict the effects of the polymorphism Glu320Pro on the biological function, functional stability and tridimensional structural modelling in the goat mature GDF9 protein by *in silico* comparative protein prediction analyses. Such comparative findings may proportionate new insights into positive structural alterations in the goat mutant GDF9 protein harboring the polymorphism Glu320Pro.

MATERIAL AND METHODS

Search and selection of the amino acid sequence of the goat GDF9 protein from biological database set

In current study, a wild sequence of amino acid residues concerning GDF9 protein of goat *Capra hircus* (Q66NC0) was searched and selected in FASTA format from non-redundant protein sequence database UniProtKB/Swiss-Prot, exhibiting the glutamine (Glu, Q) amino acid at position 320. Moreover, a mutant GDF9 protein sequence harboring a causal polymorphism promoting the presence of the proline (Pro, P) amino acid at same position was also examined under these experimental conditions.

Prediction of the effects of the polymorphism Glu320Pro on stability and biological function of the goat GDF9 protein

Possible impacts of the amino acid substitution Glu320Pro on the biological function of the goat GDF9 protein were estimated using PROVEAN Protein server (Protein Variation Effect Analyzer) (CHOI & CHAN, 2015). Briefly, clustering of BLAST hits is performed by CD-HIT with a parameter of 75% global sequence identity and the top 30 clusters of closely related sequences from the supporting sequence set, which are used to generate the prediction. A delta alignment score is computed for each supporting sequence and the scores are then averaged within and across clusters to generate the final PROVEAN score. If the PROVEAN score is equal to or below a predefined threshold (e.g. -2.5), the protein variant is predicted to have a deleterious effect, and if the PROVEAN score is above the threshold, the variant is predicted to have a neutral effect (CHOI & CHAN, 2015).

Variations in the functional stabilities of the wild and mutant GDF9 protein of goat were predicted using specific empirical rules of the PolyPhen program version2.0 (Polymorphism Phenotyping v2) that validates possible impacts of amino acid substitutions on the structure and function of human proteins using straightforward physical and evolutionary comparative considerations (ADZHUBEI et al., 2010). Scores generated by PolyPhen analyses are classified as "benign" (0.0-0.15), "possibly damaging" (0.15–1.0) or "potentially damaging" (0.85-1.0). Alterations in the structural stability of the proteins of interest were predicted by tool server I-Mutant2.0 based on Gibbs free energy changes (DDG) (CAPRIOTTI et al., 2005). The value of DDG below 0 (<0) indicates decrease in the protein stability, while value higher than 0 (>0) indicates increase in protein stability.

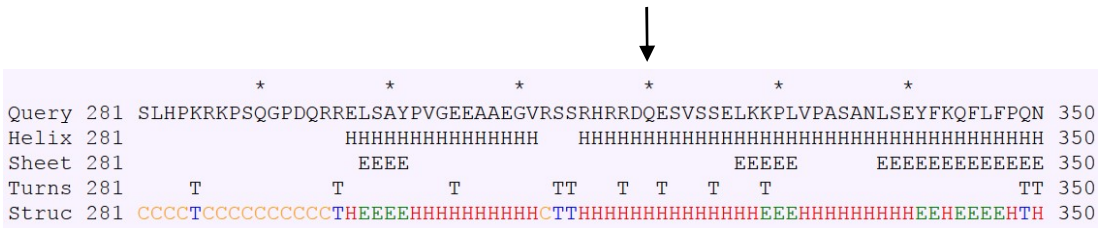
Prediction of the conformational pattern of the secondary structure of the wild and mutant goat GDF9 protein

Conformational pattern of the secondary structure of the wild and mutant goat GDF9 protein were inferred using CFSSP: Chou and Fasman Secondary Structure Prediction server (CHOU & FASMAN, 1974) and Swiss-model (WATERHOUSE et al., 2018). CFSSP server predicts regions of secondary structure from the protein sequence, such as alpha helix, beta sheet, and turns from the amino acid sequence. The output of predicted secondary structure is also displayed in linear sequential graphical view based on the probability of occurrence of alpha helix, beta sheet, and turns. On the other hand, Swiss-model employs advanced remote homology detection methods to build 3D models, predict binding sites of ligands and analyze effects of amino acid variants. In this virtual platform is also available a workbench Swiss-model repository that displays the predicted secondary structure of the proteins modelled and the confidence estimated in this prediction.

RESULTS

The replacement of the amino acid Glu for Pro at position 320 in the goat GDF9 protein provoked a neutral effect on the biological function of this protein based on Provean score value of -1.827. Prediction of the protein stability changes upon single point polymorphism Glu320Pro in goat GDF9 revealed a decrease in the protein stability with free energy change value (DDG) of -0.19. Moreover, the polymorphism Glu320Pro was predicted to be probably or potentially damaging with a score of 0.980 (sensitivity: 0.75; specificity: 0.96) on functional stability of the protein of interest by PolyPhen analyses.

Characterization of the secondary structure of the wild and mutant GDF9 proteins revealed the presence of 296 (65.3%) and 289 (63.8%) alpha helix motifs, respectively. On the other hand, no difference was observed in the extended strand and loop domains between wild and mutant GDF9, showing 265 (58.5%) and 50 (11%) motifs, respectively, for both sequences. Variations in the conformational pattern of the secondary structure were demonstrated in function of the polymorphism Glu320Pro by the presence of an alpha helix motif (H) and a coiled coil domain (C) in wild and mutant GDF9, respectively (Figure 1).



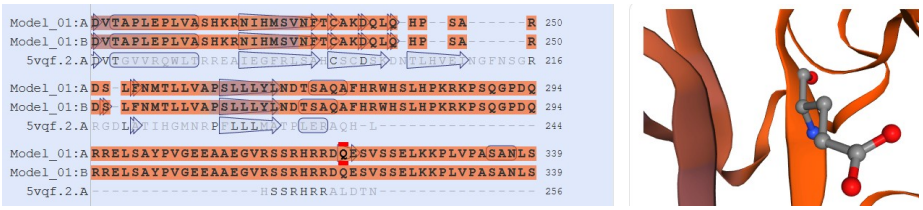
A)

B)

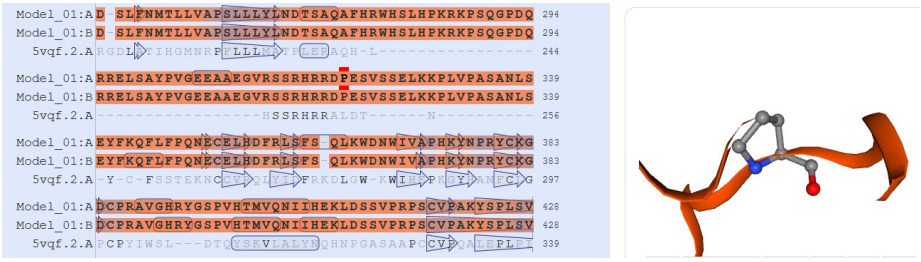
Figure 1. Partial alignments concerning secondary structural pattern of the goat GDF9 protein by CFSSP server analyses. (A) Wild GDF9 harboring the amino acid glutamine (Glu, Q) in position 320. (B) Mutant GDF9 harboring the amino acid proline (Pro, P) in same position. Arrows indicate the target position 320 marked by substituting amino acid residues.

Tridimensional structural models of the wild and mutant GDF9 protein were randomly built using as template a PDB molecule of crystal structure of pro-transforming growth factor beta-1 (ID:5vqf.2.A) deposited in Swiss-model repository (Figure 2). This template was randomly chosen during modelling analyses and have a 19.44% identity to wild and mutant GDF9 sequences. Notably, the replacement of a single amino acid residue (Glu320Pro) provoked modelling with maxima confidence in 320 (71%) and 329 residues (73%) in the wild and mutant GDF9 sequences, respectively, generated from single highest scoring template. Meanwhile, 54% of both wild and mutant GDF9 protein sequences were meaningfully predicted in disordered regions. Specially, the

substitution Glu320Pro in the protein GDF9 was detected in admittedly disordered regions during structural modelling, without the presence of complementary of none amino acid residue to template sequence at the target position. Moreover, QMEAN values of -7.13 and -6.84 were calculated for wild and mutant GDF9, respectively, indicating low modelling quality for both predictive structural models. Particularly, the presence of amino acid Glu at position 320 in the wild GDF9 protein had a QMEAN value of 0.51. Surprisingly, the replacement of a single amino acid Pro at same position in the mutant protein demonstrated a change in the QMEAN value of 0.15.



A)



B)

Figure 2. Amino acid sequence alignments and tridimensional structural models predicted for wild (A) and mutant (B) GDF9 protein by Swiss-model analyses.

DISCUSSION

The replacement of a single amino acid residue Glu320Pro provoked decreases and probably damages on functional stability of goat GDF9 protein, even biological functional was remained unaltered under these experimental conditions. Potentially distinct biochemical proprieties of the glutamine (Glu) and proline (Pro) residues may explain such findings. While glutamine is a nonessential amino acid in mammals, characterized as a charge-neutral, polar amino acid containing a negatively charged R group with a side chain similar to that of glutamic acid, except the carboxylic acid group is replaced by an amide (NELSON & COX, 2005). Particularly, proline imino acid shares many properties with the aliphatic group and its nitrogen atom is covalently bound within the molecule five-membered ring, a feature that markedly restricts the phi angular range in peptide bond formation at this locus in a peptide or protein (MORGAN & RUBENSTEIN, 2013). Furthermore, variations in the functional and biological stability of mutant goat GDF9 protein may be associated by translating rare codon sequences. Glutamine amino acid is specified by GAA and GAG codon sequences, whereas CCU, CCC, CCA, and CCG codon sequences codify proline imino acid (NELSON & COX, 2005). Notably, theoretical studies have hypothesized that the presence of a rare codon in a silent mutation marked by a synonymous polymorphism in a complex mammalian membrane transport protein alters the substrate specificity; affecting the timing of cotranslational folding and resulting in altered function in mature proteins (KIMCHI-SARFATY et al., 2007). Within this context, PONNALA (2010) have also hypothesized that clusters of rare codons have negatively affected the efficiency and accuracy of bacterial protein production. Thus, polymorphisms may be altering the frequencies with which codons and their specific tRNAs are originally present and required within intracellular environment, promoting inefficient gene expression, premature termination of the synthesized protein or misincorporation of amino acids.

Conformational characterization of the secondary structure of the wild and mutant goat GDF9 proteins revealed decreases in amount of alpha helix motifs in the mutant Glu320Pro-harboring sequence (Figure 1). The linkage of proline to other amino acids through the amino group contributes to various bends and kinks in the shape

of the protein, without which the protein could affect function properly. When proline is in a peptide bond, it does not have a hydrogen on the α amino group, so it cannot donate a hydrogen bond to stabilize an alpha helix or a beta sheet, and in practice the presence of a proline inhibit really the formation of alpha helix domains (BALDWIN, 2008). Thus, proline is often found at the end of α helix or in turns or loops and tend to be excluded from alpha helices and beta sheets. In contrast to other amino acids that exist almost exclusively in the *trans*- form in polypeptides, proline can readily adopt a cis configuration as well as a trans-configuration in response to subtle influences, presumably differences in local charge distribution (BALDWIN, 2008). In one simplified view, proline disrupts protein secondary structure by inhibiting the backbone to conform to an alpha helix or beta-sheet domains (MORGAN & RUBENSTEIN, 2013). Thus, the arguments above-mentioned may explain a decrease in the amount of alpha helix motifs showed in the conformational pattern of the mutant GDF9 protein.

Interestingly, differentiated conformational pattern in the secondary structure of mutant GDF9 protein was revealed with the presence of a coiled coil in the target position in contrast to alpha helix domain typical in the wild sequence (Figure 1). Such findings reinforce evidences that glutamate and proline residues have especially high and poor helix-forming propensities, respectively (PACE & SCHOLTZ, 1998). A coiled coil is a structural motif in proteins in which 2–7 alpha-helices are coiled together like the strands of a rope (dimers and trimers are the most common types) (LIU et al., 2006). Theoretical evidences suggest that the random coiled coil domain characterize as a class of conformations that indicate an absence of regular secondary structure (NELSON & COX, 2005). Tridimensional structural models predicted a difference in the QMEAN values for wild and mutant GDF9, indicating low quality for both predictive structural models (Figure 2). Moreover, in particular amino acid proline imposes its own kind of secondary structure with a confined phi angle that overrides other forms of the secondary structure. Within this context, the presence of coiled coil domain in mutant GDF9 could affect the spatial modifications in its unique geometric shape and exert a positive influence on its biological action during folliculogenesis in goats.

Failures in the functional protein dynamics generally provoke inferences and conformational changes about protein folding and thus, the chemistry of amino acid side chains is critical to protein structure, since these side chains can bond with one another to hold a length of protein in a certain shape or conformation. Over the past decade, *in silico* analyses have allowed to assess the impacts of single nucleotide polymorphisms on the structural conformation pattern and functional stability in different mature proteins (COSTA et al., 2013; COSTA et al., 2018; GUZZI et al., 2017; GUZZI et al., 2019). Within this context, the findings described herein proportionate findings of the potential impacts of the polymorphism Glu320Pro on biological functionality and stability and secondary structural pattern of the mature GDF9 protein in goat by *in silico* predictions.

CONCLUSION

Based on the essentially predictive character of the *in silico* comparative analyses, the findings described herein suggest that the replacement of the amino acid Glu for Pro at position 320 (polymorphism Glu320Pro) in goat GDF9 protein provoked decreases in the protein stability and potentially damages on its functional stability as well as a differentiated conformational pattern in the tridimensional structural modelling of the mutant GDF9 protein in contrast to the wild sequence.

ACKNOWLEDGMENTS

This study was supported by the Bioinformatics and Computational Biology Group “BIO in BYTES”.

REFERENCES

ADZHUBEI, I.A.; SCHMIDT, S.; PESHKIN, L.; RAMENSKY, V.E.; GERASIMOVA, A.; BORK, P.; KONDRASHOV, A.S.; SUNYAEV, S.R. A method and server for predicting damaging missense mutations. **Nature Methods**, Oxford, v.*7,* n.4, p.248–249, 2010.

BALDWIN, R.L. The search for folding intermediates and the mechanism of protein folding. **Annual Review of Biophysics**, Palo Alto, v.37, p.1-21, 2008.

CAPRIOTTI, E.; FARISELLI, P.; CASADIO, R. I-Mutant2.0: predicting stability changes upon mutation from the protein sequence or structure. **Nucleic Acids Research**, Oxford, v.33, p.W306-W310, 2005.

CHOI, Y.; CHAN, A.P. PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels. **Bioinformatics**. Oxford, v.31, n.16, p.2745-2747, 2015.

CHOU, P.Y.; FASMAN, G.D. Conformational parameters for amino acids in helical, β-sheet, and random coil regions calculated from proteins. **Biochemistry**, Berkeley, v.13, n.2, p.211–222, 1974.

COSTA, F.F.P.; LIDANI, K.C.F.; NUNES, S.L.; GOUVEIA, J.J.S. ; GABRIEL, J.E. Phylogenetical clustering among domestic ruminants from diacylglycerol-acyltransferase-1 lipogenic enzyme and biochemical characterization of its single nucleotide polymorphism K232A in bovine *Bos taurus*. **Animal Biology Journal**, New York, v.4, p.1-10, 2013.

COSTA, F.F.P.; NOGUEIRA, J.F.; GOUVEIA, J.J.S.; GABRIEL, J.E. Characterizing the polymorphism K232A of the diacylglycerol-acyltransferase-1 lipogenic enzyme of bovine *Bos taurus* using in silico comparative protein prediction analyses. **Brazilian Journal of Biology**, São Carlos, v.78, n.2, p.389-390, 2018.

DONG, J.; ALBERTINI, D.; NISHIMORI, K.; KUMAR, T.; LU, N.; MATZUK, M. Growth differentiation factor-9 is required during early ovarian folliculogenesis. **Nature**, London, v.383, n.6600, p.531-535, 1996.

GUZZI, A.F.; OLIVEIRA, F.S.L.; AMARO, M.M.S.; TAVARES-FILHO, P.F.; GABRIEL, J.E. Predicting the effects of the single nucleotide polymorphism A122V on the CXC chemokine receptor type 1 of bovine *Bos taurus* (Artiodactyla: Bovidae) by *in silico* analyses. **Biotemas**, Florianópolis, v.30, p.1-6, 2017.

GUZZI, A.F.; OLIVEIRA, F.S.L.; AMARO, M.M.S.; TAVARES-FILHO, P.F.; Gabriel, J.E. In silico prediction of the functional and structural consequences of the non-synonymous single nucleotide polymorphism A122V in bovine CXC chemokine receptor type 1. **Brazilian Journal of Biology**, São Carlos, v.1, p.1-8, 2019.

HANRAHAN, J. Mutations in the genes for oocyte-derived growth factors GDF9 and BMP15 are associated with both increased ovulation rate and sterility in Cambridge and Belclare sheep (*Ovis aries*). **Biology of Reproduction**, Oxford, v.70, n.4, p.900-909, 2003.

KIMCHI-SARFATY, C.; OH, J.M.; KIM, I.W.; SAUNA, Z.E.; CALCAGNO, A.M.; AMBUDKAR, S.V.; GOTTESMAN, M.M. A "silent" polymorphism in the MDR1 gene changes substrate specificity. **Science**, New York, v.315, n.5811, p.525–528, 2007.

LIU, J.; ZHENG, Q.; DENG, Y.; CHENG, C.S.; KALLENBACH, N.R.; LU, M. A seven-helix coiled coil. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v.103, n.42, p.15457–15462, 2006.

MORGAN, A.A.; RUBENSTEIN, E. Proline: the distribution, frequency, positioning, and common functional roles of proline and polyproline sequences in the human proteome. **PLoS One**, San Francisco, v.8, n.1, p.e53785, 2013.

NELSON, D.; COX, M. **Lehninger Principles of Biochemistry**. (4th ed.); W.H. Freeman and Company, New York, 2005, 1216p.

NILSSON, E. Growth and Differentiation Factor-9 stimulates progression of early primary but not primordial rat ovarian follicle development. **Biology of Reproduction**, Oxford, v.67, n.3, p.1018-1024, 2002.

ORISAKA, M.; ORISAKA, S.; JIANG, J.; CRAIG, J.; WANG, Y.; KOTSUJI, F.; TSANG, B. Growth Differentiation Factor 9 is antiapoptotic during follicular development from preantral to early antral stage. **Molecular Endocrinology**, Oxford, v.20, n.10, p.2456-2468, 2006.

PACE, C.N.; SCHOLTZ, J.M. A helix propensity scale based on experimental studies of peptides and proteins. **Biophysical Journal**, Cambridge, v.75, p.422-427, 1998.

PAN Z.Y.; DI R.; TANG Q.Q.; JIN H.H.; CHU M.X.; HUANG D.W.; HE J.N.; LIU Q.Y.; HU W.P.; WANG X.Y.; YAO Y.X.; LIU L.; SONG C.L. Tissue-specific mRNA expression profiles of GDF9, BMP15, and BMPR1B genes in prolific and nonprolific goat breeds. **Czech Journal of Animal Science**, Praga, v.60, p.452e8, 2015.

PONNALA, L. Detecting slow-translating regions in *E.coli*. **International Journal of Bioinformatics Research and Applications**, Atlanta, v.6, p.522–530, 2010.

PRAMOD, R.K.; SHARMA, S.K.; SINGHI, A.; PAN, S.; MITRA, A. Differential ovarian morphometry and follicular expression of BMP15, GDF9 and BMPR1B influence the prolificacy in goat. **Reproduction in Domestic Animal**, Malden, v.48, p.803e9, 2013.

SILVA, B.D.; CASTRO, E.A.; SOUZA, C.J.; PAIVA, S.R.; SARTORI, R.; FRANCO, M.M.; AZEVEDO, H.C.; SILVA, T.A.; VIEIRA, A.M.; NEVES, J.P.; MELO, E.O. A new polymorphism in the Growth and Differentiation Factor 9 (GDF9) gene is associated with increased ovulation rate and prolificacy in homozygous sheep. **Animal Genetics**, Oxford, v.42, n.1, p.89-92, 2011.

SILVA, J.R.; VAN DEN HURK, R.; VAN TOL, H.T.; ROELEN, B.A.; FIGUEIREDO, J.R. Expression of growth differentiation factor 9 (GDF9), bone morphogenetic protein 15 (BMP15), and BMP receptors in the ovaries of goats. **Molecular Reproduction and Development**, Oxford, v.70, p.11e9, 2005.

WANG, X.; YANG, Q.; WANG, K.; YAN, H.; PAN, C.; CHEN, H.; LIU, J.; ZHU, H.; QU, L.; LAN, X. Two strongly linked single nucleotide polymorphisms (Q320P and V397I) in GDF9 gene are associated with litter size in cashmere goats. **Theriogenology**, New York, v.125, p.115-121, 2019.

WATERHOUSE, A.; BERTONI, M.; BIENERT, S.; STUDER, G.; TAURIELLO, G.; GUMIENNY, R.; HEER, F.T.; DE BEER, T.A.P.; REMPFER, C.; BORDOLI, L.; LEPORE, R.; SCHWEDE, T. SWISS-MODEL:homology modelling of protein structures and complexes. **Nucleic Acids Research**, Oxford, v.46, p.W296-W303, 2018.