Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

# In Silico Analysis of Pathogenic Missense Single Nucleotide Polymorphisms in the Bovine IGF-1 Gene

**Ali Forouharmehr**

forouharmehr.a@lu.ac.ir

**Seyyed Mojtaba Mousavi**
**Narges Nazifı**

**Research Article**

**Keywords:** IGF-1, Single nucleotide polymorphism, Bioinformatics, Pathogenicity, Genetic variation

**Posted Date:** August 6th, 2026

**DOI:** https://doi.org/10.21203/rs.3.rs-10279745/v1

**License:** This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

**Additional Declarations:** No competing interests reported.

# In Silico Analysis of Pathogenic Missense Single Nucleotide Polymorphisms in the Bovine *IGF-1* Gene

Ali Forouharmehr[1]*, Seyyed Mojtaba Mousavi[2], Narges Nazifi[3]

1) (Corresponding author) Department of Animal Science, Faculty of Agriculture, Lorestan University, Khorramabad, Iran. Email address: Forouharmehr.a@lu.ac.ir
2) Department of Animal Science, Faculty of Agriculture, Lorestan University, Khorramabad, Iran. Email address: Mousavi.sym@lu.ac.ir
3) Department of Basic Science, Faculty of Veterinary Medicine, Lorestan University, Khorramabad, Iran. Email address: Nazifi.nrg@lu.ac.ir

## Abstract

Bovine insulin-like growth factor-1is a vital hormone that plays important physiological roles in dairy cows. Therefore, it seems that disruption of this gene may impair dairy production. The current research was performed to investigate the pathogenic effect of missense SNPs in bovine IGF-1 gene. To do so, missense SNPs of bovine IGF-1 gene were collected from database, then deleterious SNPs were identified by PredictSNP tool. Evolutionary conservation of deleterious SNP positions was evaluated by ConSurf server. Also, deleterious SNP effects on stability of bovine IGF-1 protein were investigated by I-Mutant2.0 and MUpro servers, while their effects on protein flexibility were assessed by MEDUSA server. The harmful SNPs were screened by MutPred2 server and the most deleterious SNP was identified based on output of Project HOPE server. Finally, 3D structures of the most deleterious mutant-type and wild-type bovine IGF-1 proteins were modeled and applied for protein- protein docking analysis Our results identified the L105R substitution as the most deleterious of 66 missense SNPs, disrupting the structural and functional integrity of the bovine IGF-1 protein. Furthermore, this substitution impaired the IGF-1–IGF-1 receptor complex by reducing the number of hydrogen bonds. Consequently, we propose this mutation as a potential molecular marker for animal breeding and selection programs.


**Keywords**: IGF-1, Single nucleotide polymorphism, Bioinformatics, Pathogenicity, Genetic variation.

# 1. Introduction

Genetic factors such as single nucleotide polymorphisms (SNP) are categorized as one of the most common variations in genome sequences of mammalians(Venkata Subbiah et al. 2020). Generally, the SNPs can be classified based on their genomic regions: coding SNPs occur in coding regions, while non-coding SNPs are found in non-coding regions. Non-coding SNPs cannot lead to amino acid substitutions, consequently these types of SNPs do not affect the structure and function of a special protein(Kamal et al. 2024). Coding SNPs are divided into two groups: synonymous SNPs which substitute an amino acid with the same amino acid, and non-synonymous SNPs which can substitute an amino acid with different amino acids (missense SNPs) or a stop codon (non-sense SNPs). Therefore, missense SNPs can impress structure and function of a vital protein through amino acid substitution(Kamal et al. 2024). Insulin-like growth factor1(IGF-1) is a physiological growth mediator which has important role in cellular and tissue growth and development(Lupu et al. 2001). Disorders in IGF-1 function have been associated various diseases such as cancers(Oldham and Hafen 2003; Yu and Rohan 2000). This hormone also known as somatotropin C is a protein which is encoded by the *IGF-1* gene, in human, *IGF-1* gene is located on chromosome12, while in cattle this gene is on chromosome 5(Chowdhury 2014). Molecular weight of IGF-1 protein is 7.5 kDa with 170 amino acids in length, this protein is mainly secreted by liver in response to growth hormone stimulation(Bach and Rechler 1992; Chowdhury 2014; Jones and Clemmons 1995). In dairy cattle, IGF-1 hormone plays crucial role in lactation via stimulation of mammary gland development(Zhang et al. 2023). Additionally, this hormone influences quality of milk using regulation of lipid and protein synthesis(Kumar et al. 2023). Besides, IGF-1 is strongly related to muscle development, skeletal growth, and fertility(Soliman et al. 2023). Therefore, it can be concluded that dysfunction of this hormone can lead to economic losses in the dairy cattle industry.

Nowadays, Bioinformatics through providing accurate computational tools and extensive databases is playing pivotal role in interpreting diverse biological data including genomics, transcriptomics and proteomics. Consequently, bioinformatics approaches are being employed to identify association between raw data (e.g., SNPs) and various dysfunctions and diseases.

Hence, in the current study, the most reliable bioinformatics tools were applied to assess missense SNPs in the bovine *IGF- 1* gene that could potentially impair function and structure of IGF-1 protein.

## 2. Material and methods

### 2.1. Data collection

For this study, missense SNPs in bovine *IGF-1* Gene (transcript ID: ENSBTAT00000063328.3) were collected from Ensembl (https://asia.ensembl.org/index.html). Moreover, bovine IGF-1 protein sequence (accession number: G3N086) was retrieved from Uniport (https://www.uniprot.org/).

### 2.2. Prediction of Deleterious Missense SNPs

To classify missense SNPs as deleterious or neutral SNPs, PredictSNP online tool(https://loschmidt.chemi.muni.cz/predictsnp1/) was employed. This predictor can simultaneously apply multi algorithms, including SIFT, MAPP, PhD-SNP, PolyPhen-1, PolyPhen-2 and SNAP.

### 2.3. Prediction of Evolutionary Conservation

In this study, we predicted amino acid conservation in bovine IGF-1 protein by ConSurf server(https://consurf.tau.ac.il/consurf_index.php). This server is able to generate a phylogenic tree from homologous sequences and estimate evolutionary conservation scores of amino acid positions.

### 2.4. Assessment of Missense SNPs's Effect on Stability and Flexibility of Bovine IGF-1 protein

In order to evaluate effects of deleterious missense SNPs on bovine IGF-1 protein stability, I-Mutant2.0 severe (https://folding.biofold.org/cgi-bin/i-mutant2.0.cgi) along with MUpro server (https://mupro.proteomics.ics.uci.edu/) were employed. Moreover, flexibility of the bovine IGF-1 protein was assessed using MEDUSA server (https://www.dsimb.inserm.fr/MEDUSA/).

### 2.5. Identification of Harmful Missense SNPs

In the current study, MutPred2 server (http://mutpred.mutdb.org/#qform) was applied to determine association between amino acid substitution and occurrence a disease. In this case, amino acid sequence of bovine IGF-1 protein which contained deleterious mutations was submitted to server with a P-value threshold of 0.05.

### 2.6. Prediction of Missense SNPs effect on Structure of IGF-1 protein

To evaluate effect of missense SNPs on structure of IGF-1 protein, Project HOPE server (https://www3.cmbi.umcn.nl/hope/input/)was applied. This server compares various

terms including; size, hydrophobicity and charge between mutant and wild proteins to identify the most deleterious mutation.

### 2.7. Tertiary Structure Prediction and Protein-protein interaction investigation

To predict tertiary structure of the wild and the most deleterious mutant type of bovine IGF-1 protein, Robetta server (https://robetta.bakerlab.org/) was employed. To perform this modeling amino acid sequences of the proteins were pasted to server and the best model was chosen based on the ranking presented by the server. Additionally, to determine protein-protein interaction among bovine IGF-1 protein and its interactors, STRING server (https://string-db.org/) was applied.

### 2.8. Protein-Protein Docking

In this research, we used ClusPro server (https://cluspro.bu.edu/login.php) to investigate protein-protein interactions between IGF-1 receptor (as the closest interactor), and the most deleterious mutant-type IGF-1. Moreover, to evaluate hydrogen bond formation in docking products, LigPlot+ software was employed.

## 3. Result

### 3.1 Data Collection

As mentioned, in this study, we collected missense SNPs in bovine *IGF-1* Gene (transcript ID: ENSBTAT00000063328.3). For this transcript, various types of SNPs (104 total) have been reported including 66 missense, 30 synonymous, four splice region, one frameshift and 3 stop gained SNPs.

### 3.2. Prediction of Deleterious Missense SNPs

To predict deleterious SNPs among 66 missense SNPs in bovine *IGF-1* gene, we used PredicSNP2 tool. Among these missense SNPs, our analysis identified 18 SNPs were as deleterious and potentially associated with disease (**Table 1**).

### 33. Prediction of Evolutionary Conservation

As mentioned, in this study ConSurf server was used for calculation of evolutionary conservation scores in amino acid positions of bovine IGF-1 protein. Our results confirmed that among 18 amino acid residues of this protein, 14 residues including: C54S, C54R, F73C, I91T, C95W, C95G, C95S, C96Y, F97C, C100Y, C100G, L105R, Y108S and C109G were structural residues, while two residues including G55V and K75Q were classified as functional residues (**Table 2**). As reported in **Table 2**, all residues had a conservation score of ≥7, consequently they were considered as highly conserved residues. As shown **Figure1**, all highly conserved functional residues were exposed, whereas highly conserved structural residues were buried.

### 3.4. Assessment of Missense SNPs's Effect on Stability and Flexibility of Bovine IGF-1 protein

As mentioned, to evaluate effect of missense SNPs on stability of IGF-1 protein, the I-Mutant and MUpro servers were used. According to MUpro results, all 18 SNPs decreased IGF-1 stability (**Table 3**), while the results of I-Mutant revealed that except for C96Y and C100Y the remaining SNPs decreased IGF-1 Stability. Consequently, these two SNPs (C96Y and C100Y) were excluded from subsequent analysis. Besides, our results confirmed that all SNPs located in rigid positions (**Figure 2**).

### 3.5. Identification of Harmful Missense SNPs

In the current study, we used MutPred2 to assess pathogenicity of the remaining missense SNPs. Our results indicated that several missense SNPs disrupt functional features including signal peptide, metal binding site, transmembrane domain and disordered interfaces, while others introduced loop and intrinsic disorder (**Table 4**). Eventually, six missense SNPs including F97C, C95S, C95G, G55V, L105R, C95W and C100G with a score > 0.80, were identified as the most harmful SNPs (**Table 4**). Therefore, except these seven missense SNPs, the other SNPs were excluded from study.

### 3.6. prediction of Missense SNPs effect on Structure of IGF-1 protein

Aa mentioned, in this study we employed Project HOPE server to evaluate structure effects of missense SNPs on bovine IGF-1 protein. Our results revealed that C95S, C95G, L105R and C100G substitutions could decrease hydrophobicity of bovine IGF-1. Moreover, F95C, C100G and C95G substitutions reduced molecular size of the protein, while C95W and L105R substitutions increased molecular size (**Table 5**).

### 3.7. Tertiary Structure Prediction and Protein-protein interaction investigation

To predicted tertiary structure of the wild bovine IGF-1 protein and L105R mutant (as the most deleterious mutation). Tertiary structures of the mentioned protein have been illustrated in **Figure 3**. Moreover, protein-protein interaction results confirmed that 10 proteins including IGF1R, IGFBP3, IGFBP1, IGFBP5, IGFBP2, INSR, IGFBP6, INS, IGF2R and HGF were the closest interactors with bovine IGF-1 protein, respectively **(Figure 4).**

### 3.8. Protein-Protein Docking

Our results revealed that both wild-type IGF1 protein and L105R mutant-type IGF-1 protein could dock to IGF-1 protein with lowest energy of -1089.7 and -1214.5, respectively (**Figure 5**). Moreover, our results confirmed that the wild-type IGF-1 protein bound to IGF-I receptor through 25 hydrogen bonds while L105Rmutant-type IGF-1 bound using 21 hydrogen bonds (**Figure 6**).

## 4. Discussion

In general, IGF-1 protein is known as a key mediator of growth hormone which plays crucial role in growth, metabolism and fertility of dairy cow. This protein is mainly synthesized by liver and it can bind to receptor tyrosine kinase and activate MAPK and

P13k/Akt pathways, finally this activation leads to differentiation, proliferation and survival of mammary gland cells(Bauman and Vernon 1993; Capuco et al. 2001). Besides, it has been reported, decrease in serum level of IGF-1 hormone during negative balance of energy in early lactation step, can cause infertility (through ovarian dysfunction and decrease of oocyte quality) and increment of mastitis infection(Butler 2003). Therefore, it can be concluded that IGF-1 protein dysfunction can directly impress economics of dairy cow industry, because fertility and lactation are considered as two important factors for success in this industry. The current project was designed to assess the effect of various missense SNP on the structure and consequently the function of bovine IGF-1 protein. This project was started with missense SNP collection from the Ensemble database. Our initial investigation identified 104 various SNPs for bovine *IGF-1* gene (accession number: ENSBTAT00000063328.3). Among these, 66 SNPs were classified as missense SNPs in the database. As mentioned, to identified deleterious SNPs, we employed PredictSNP tool which uses multiple algorithms. To increase selection stringency for deleterious SNPs, a high pathogenicity confidence cutoff ≥80% was considered. As shown in Table1, 18 out of 66 missense SNPs, were classified as deleterious SNPs, therefor our subsequent analysis focused on these 18 SNPs. In the next step, evolutionary conservation of deleterious SNPs in bovine IGF-1 protein were investigated by ConSurf web server. As matter of fact, this server applies multiple sequence alignment and phylogenic tree construction to identify amino acid positions which are crucial to maintain structure and function of a protein(Ashkenazy et al. 2016). It must be mentioned that the ConSurf score≥8 represent positions that can affect structure and function of a protein. As shown in Table 2, all our deleterious SNPs position with a ConSurf score of 9(except for position 37 and 40) were essential for maintaining structure and function of bovine IGF-1 protein. In general, a single amino acid substitution in a protein sequence can decrease its stability(Cheng et al. 2006). Hence investigation of missense SNP effects on stability of a protein is crucial. In this project, we utilized MUpro and I-Mutant tools to assess effect of deleterious SNPs on stability of bovine IGf-1 protein. Our results confirmed that all our deleterious SNPs (except for G55V and C100Y) cloud decrease stability of bovine IGF-1 protein (Table3). Understanding about a protein flexibility is essential to clarify various molecular mechanisms. The MEDUSA server integrates evolutionary data and physicochemical features to predict flexibility of amino acid position(Vander Meersche et al. 2021). As shown in figure 2, all remaining deleterious SNPs were located in rigid regions. To identify pathogenic mutations among our remaining SNPs, MutPred2 server was employed. This server can classify a pathogenic mutation based on its effect on various molecular mechanism through a pathogenicity score ranging from 0-1 (Pejaver et al. 2020). As reported in Table 4, seven mutations including G55V, F97C, C95G, C95S, L105R and C100G with scores >80% could alter molecular mechanisms or gain disorders. In subsequent analysis, we used Project Hope server to investigation of structural and functional effects of the remaining SNPs. As matter of fact. This server merges 3D structure and sequence annotation data to predict effects of a mutation on various features of a protein including charge, size, hydrophobicity and spatial structure in comparison with wild type(Kamal et al. 2024; Venselaar et al. 2010). As mentioned in Table 5, all remaining SNPs were located within a domain which could lead to a dysfunction in IGF-1 protein. Moreover, the server reports revealed that, L105R

substitution was the most deleterious SNPs. This mutant residue not only causes bumps due to increased size but also induces ligands or residues repulsion, owing to changing charge from neutral to positive. It must be mentioned that L105R mutant protein is less hydrophobic than the wild-type IGF-1 protein, consequently its hydrophobic interactions maybe disrupted due to this alteration. Finally, to investigate interaction between L150R mutant-type IGF-1 protein (identified as the most deleterious variant of IGF-1 protein) and IGF-1 receptor, we conducted molecular docking analysis. Comparison between interaction plots of wild-type and L105R mutant-type complexes revealed that, this mutation could reduce the number of hydrogen bonds from 25 to 21. Therefore, it seems that L105R mutation may impair bovine IGF-1/IGF-1 receptor binding.

## Conclusion

In this project, we completely investigated the pathogenicity of missense SNPs in bovine *IGF-1* gene. our results provide strong evidences that L105R mutation leads to dysfunction in bovine IGF-1 protein. Consequently, this mutation is suggested as potential molecular marker for animal breeding programs.

**Data availability**: Not applicable

## Declarations

**Funding declaration:** The authors declare that no funding was received to perform this project.

**Conflict of interest:** The authors declare that there is no conflict of interest.

**Consent for publication:** All authors give consent to publish.

## Reference

Ashkenazy H, Abadi S, Martz E, Chay O, Mayrose I, Pupko T, Ben-Tal N (2016) ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. Nucleic acids research 44, W344–W350

Bach LA, Rechler MM (1992) Insulin-like growth factors and diabetes. Diabetes/metabolism reviews 8, 229–257

Bauman DE, Vernon RG (1993) Effects of exogenous bovine somatotropin on lactation.

Butler WR (2003) Energy balance relationships with follicular development, ovulation and fertility in postpartum dairy cows. Livestock production science 83, 211–218

Capuco A, Wood D, Baldwin R, Mcleod K, Paape M (2001) Mammary cell number, proliferation, and apoptosis during a bovine lactation: relation to milk production and effect of bST. Journal of dairy science 84, 2177–2187

Cheng J, Randall A, Baldi P (2006) Prediction of protein stability changes for single-site mutations using support vector machines. Proteins: Structure, Function, and Bioinformatics 62, 1125–1132

Chowdhury S (2014) Novel genes expressed in the pancreatic islets and regulated by IGF-I: McGill University (Canada)

Jones JI, Clemmons DR (1995) Insulin-like growth factors and their binding proteins: biological actions. Endocrine reviews 16, 3–34

Kamal MM, Mia MS, Faruque MO, Rabby MG, Islam MN, Talukder MEK, Wani TA, Rahman MA, Hasan MM (2024) In silico functional, structural and pathogenicity analysis of missense single nucleotide polymorphisms in human MCM6 gene. Scientific Reports 14, 11607

Kumar S, Dahiya S, Magotra A, Ratwan P, Bangar Y (2023) Influence of single nucleotide polymorphism in the IGF-1 gene on performance and conformation traits in Munjal sheep. Zygote 31, 70–77

Lupu F, Terwilliger JD, Lee K, Segre GV, Efstratiadis A (2001) Roles of growth hormone and insulin-like growth factor 1 in mouse postnatal growth. Developmental biology 229, 141–162

Oldham S, Hafen E (2003) Insulin/IGF and target of rapamycin signaling: a TOR de force in growth control. Trends in cell biology 13, 79–85

Pejaver V, Urresti J, Lugo-Martinez J, Pagel KA, Lin GN, Nam H-J, Mort M, Cooper DN, Sebat J, Iakoucheva LM (2020) Inferring the molecular and phenotypic impact of amino acid variants with MutPred2. Nature communications 11, 5918

Soliman A, Ahmad S, Hamed N, Alyafei F, Alaaraj N, Soliman N (2023) Maternal, placental, and fetal Insulin-Like Growth Factor-I (IGF-I) and IGF Binding proteins (IGFBPs) in Diabetic pregnancies: Effects on fetal growth and birth size. World Journal of Advanced Research and Reviews 17, 287–295

Vander Meersche Y, Cretin G, de Brevern AG, Gelly J-C, Galochkina T (2021) MEDUSA: prediction of protein flexibility from sequence. Journal of molecular biology 433, 166882

Venkata Subbiah H, Ramesh Babu P, Subbiah U (2020) In silico analysis of non-synonymous single nucleotide polymorphisms of human DEFB1 gene. Egyptian Journal of Medical Human Genetics 21, 66

Venselaar H, Te Beek TA, Kuipers RK, Hekkelman ML, Vriend G (2010) Protein structure analysis of mutations causing inheritable diseases. An e-Science approach with life scientist friendly interfaces. BMC bioinformatics 11, 548

Yu H, Rohan T (2000) Role of the insulin-like growth factor family in cancer development and progression. Journal of the National Cancer Institute 92, 1472–1489

Zhang W, Lin S, Jiao Z, An L, Xie T, Wu J, Zhang L (2023) The Mouse CircGHR Regulates Proliferation, Differentiation and Apoptosis of Hepatocytes and Myoblasts. Genes 2023, 14, 1207.

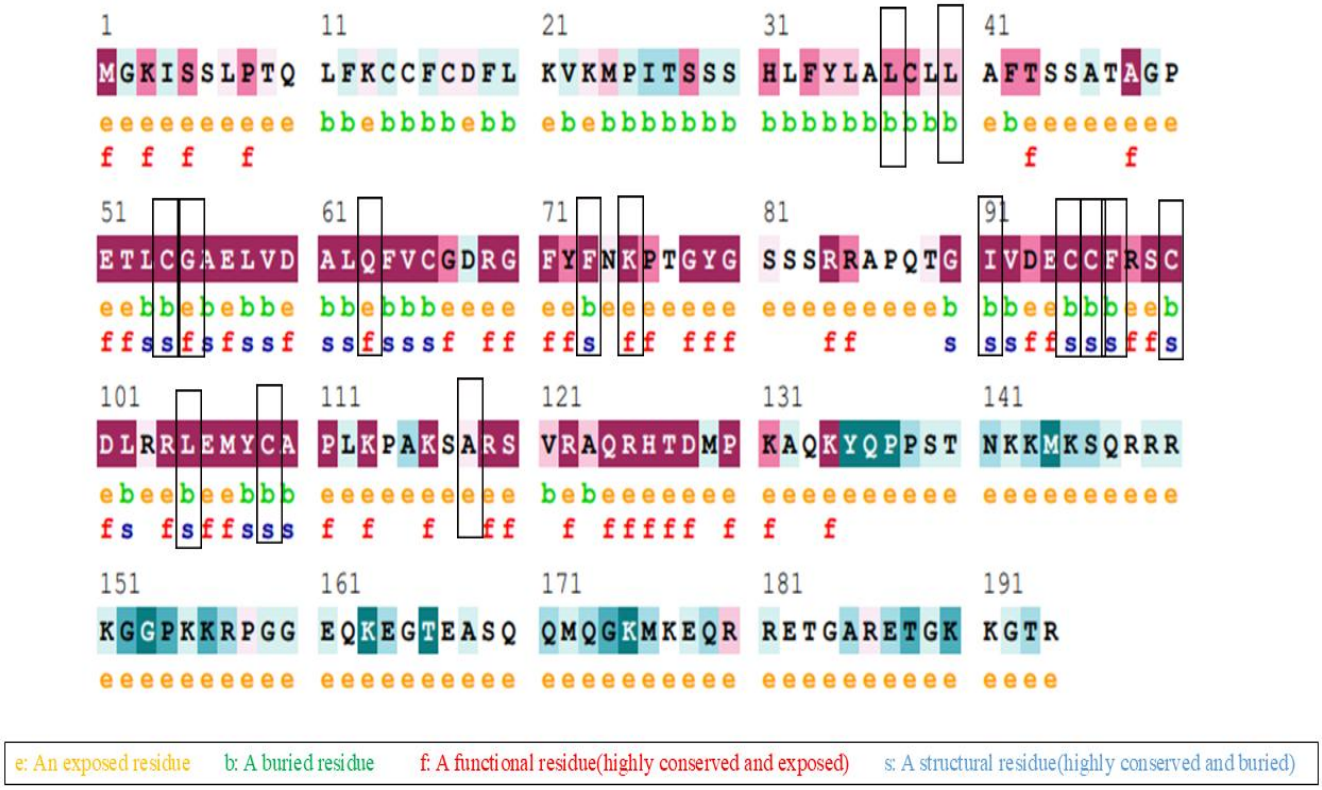



**Figure1:** The conservation analysis of IGF-1 protein residues, the deleterious positions have been highlighted by rectangles.

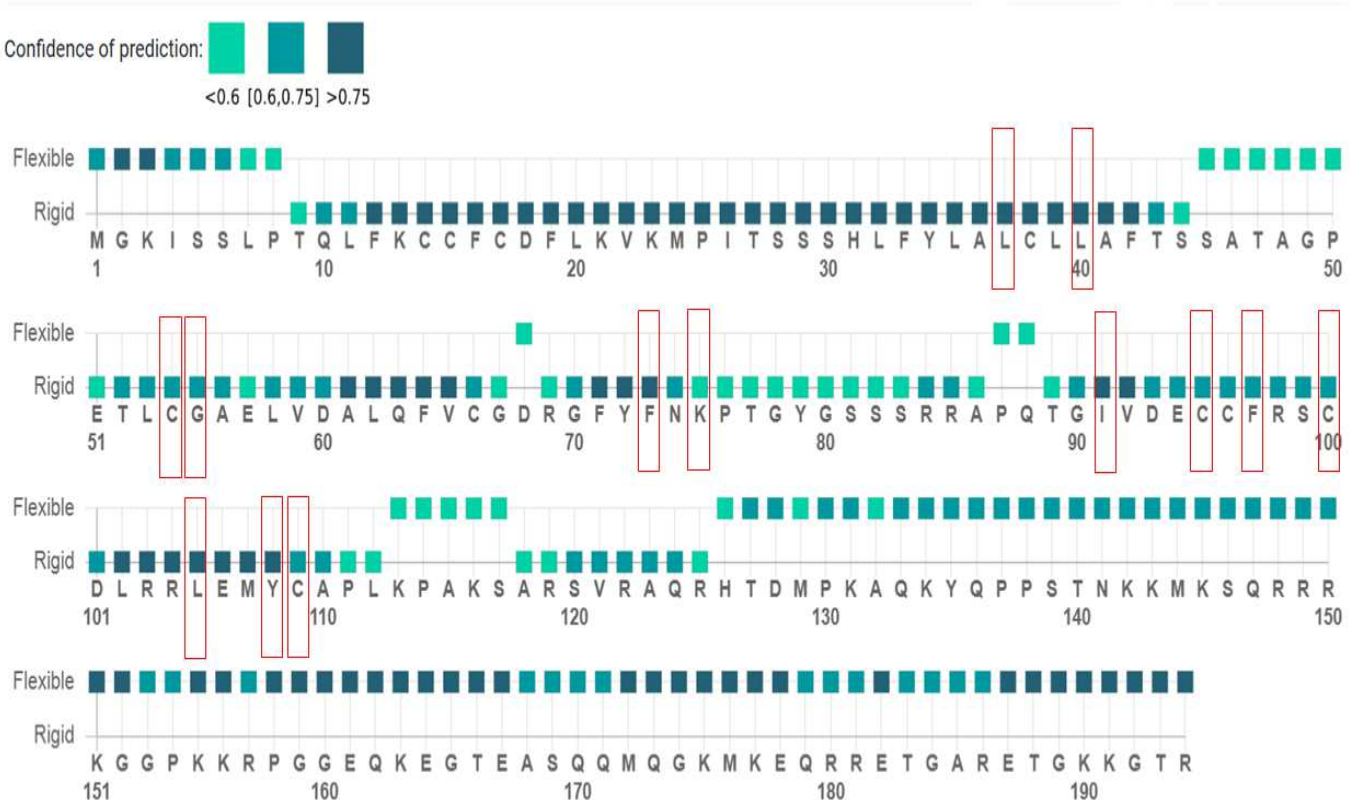



**Figure 2**: Prediction of flexibility in IGF-1 protein residues position. As highlighted all SNP positions were rigid.

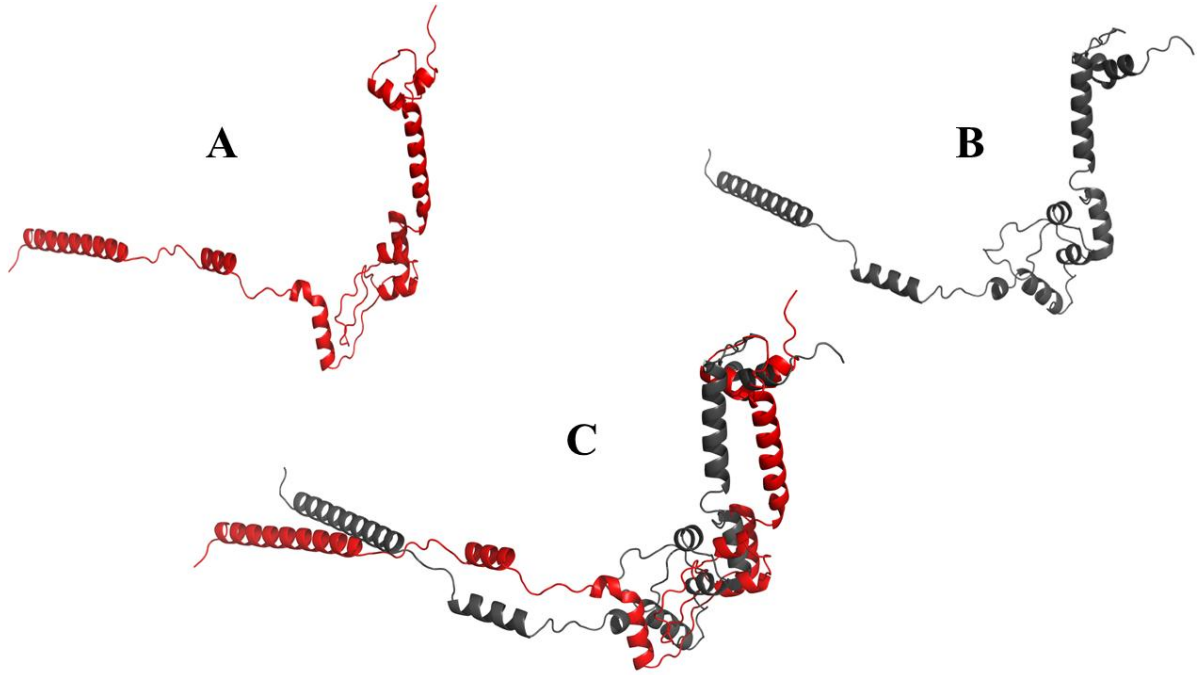



**Figure3**: **Tertiary structure models of bovine IGF-1 protein**. (A) Wild-type structure. (B) L105R mutant structure. (C) Structural alignment of wild-type (red) and L105R mutant (black) structures. Structures were visualized using PyMol software.

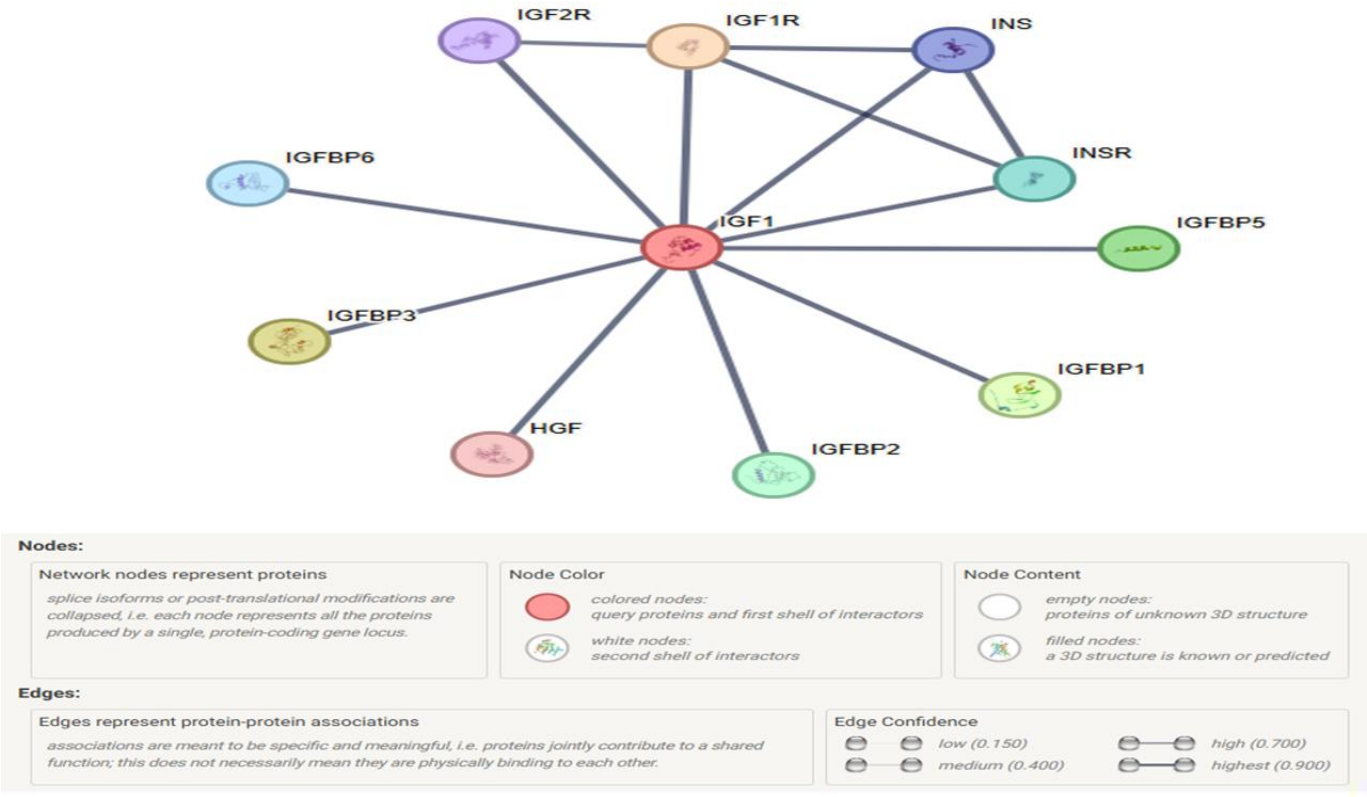


**Figure 4**: Protein-protein interaction network of bovine IGF-1 protein.

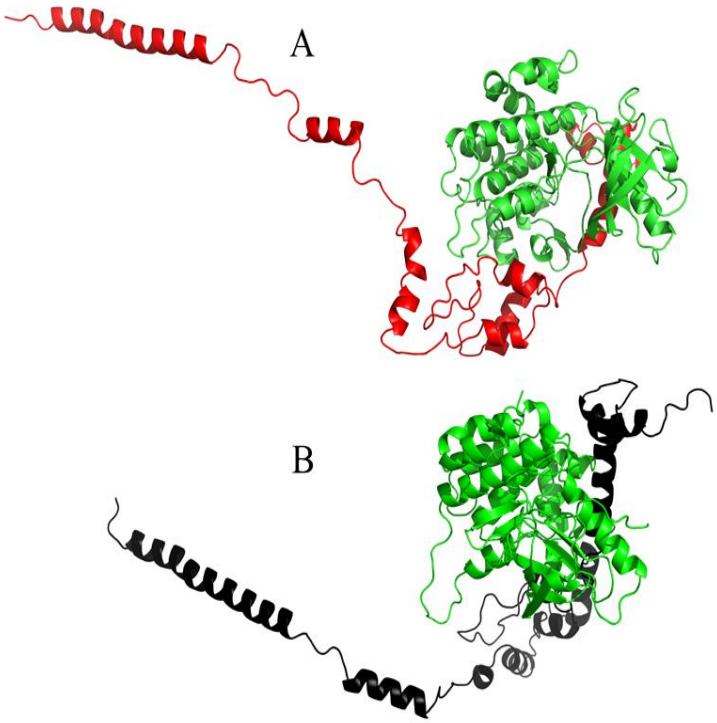



**Figure 5**: **Molecular docking of IGF-1/IGF-1 receptor complexes**. (**A**) Wild-type IGF-1 (red) bound to IGF-1 receptor (green). (**B**) L105R mutant IGF-1 (black) bound to IGF-1 receptor (green).

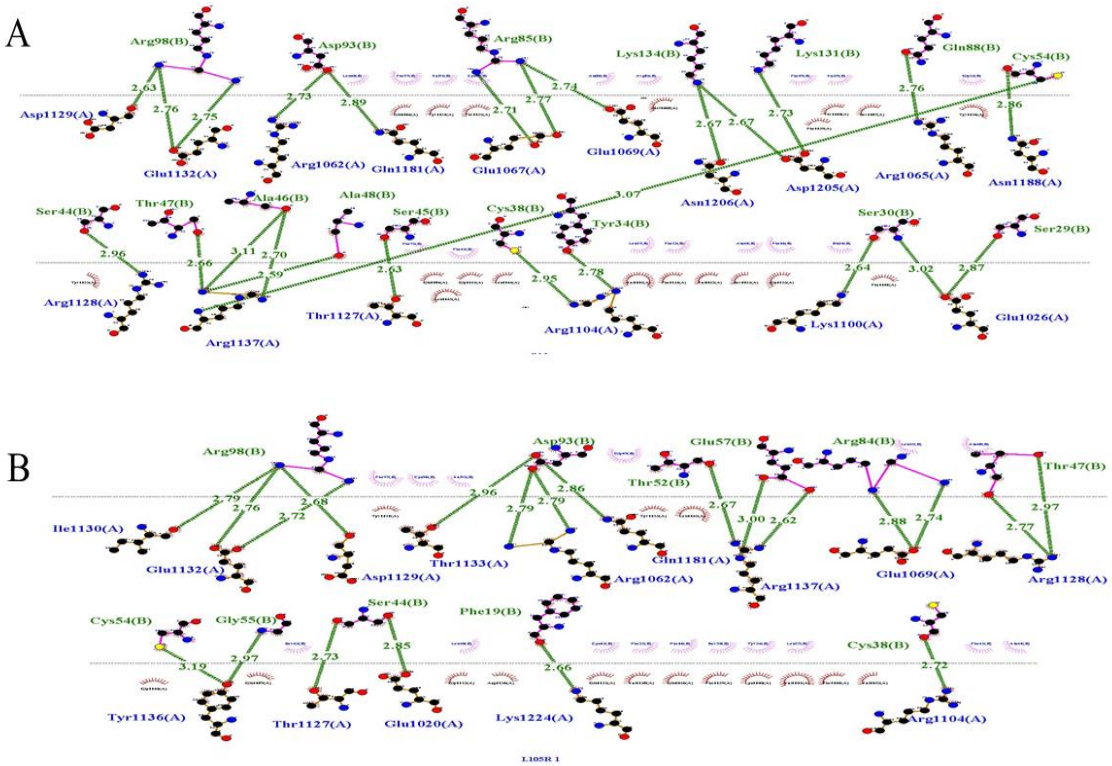



**Figure 6**: **Protein-protein interaction plots of IGF-1/IGF-1 receptor complexes**. (**A**) Wild-type IGF-1–IGF-1 receptor complex. (**B**) L105R mutant IGF-1–IGF-1 receptor complex. Hydrogen bonds are highlighted in green.

Table 1: List of deleterious SNPs in *IGF1*-Gene

| NO | Variant ID | Alleles | Amino acid substitution | Predict SNP | SIFT | MAPP | PhD-SNP | PolyPhen-1 | PolyPhen-2 | SNAP | Average of pathogenicity confidence |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | rs460963763 | A/C | L37R | D* | D | D | D | D | D | D | 82% |
| 2 | rs478704767 | A/G | L40P | D | D | D | D | D | D | D | 81% |
| 3 | rs472547819 | C/G | C54S | D | D | D | D | D | D | D | 83% |
| 4 | rs434743998 | A/G | C54R | D | D | D | D | D | D | D | 82% |
| 5 | rs452475196 | C/A | G55V | D | D | D | D | D | D | D | 84% |
| 6 | rs482754265 | A/C/G | F73C | D | D | D | D | D | D | D | 82% |
| 7 | rs450717758 | T/G | K75Q | D | D | D | D | D | D | D | 80% |
| 8 | rs455759048 | A/G | I91T | D | D | D | D | D | D | D | 80% |
| 9 | rs456928487 | G/C | C95W | D | D | D | D | D | D | D | 82% |
| 10 | rs464204228 | A/C/T | C95G | D | D | D | D | D | D | D | 81% |
| 11 | rs464204228 | A/C/T | C95S | D | D | D | D | D | D | D | 84% |
| 12 | rs437108565 | C/T | C96Y | D | D | D | D | D | D | D | 82% |
| 13 | rs468436354 | A/C | F97C | D | D | D | D | D | D | D | 80% |
| 14 | rs458376971 | C/T | C100Y | D | D | D | D | D | D | D | 82% |
| 15 | rs477050967 | A/C | C100G | D | D | D | D | D | D | D | 82% |
| 16 | rs462399135 | A/C | L105R | D | D | D | D | D | D | D | 84% |
| 17 | rs439669768 | T/G | Y108S | D | D | D | D | D | D | D | 82% |
| 18 | rs457016084 | A/C | C109G | D | D | D | D | D | D | D | 81% |

*Deleterious SNP

| Table 2: The results of functional and structural analysis of the deleterious SNPs | | | | |
|---|---|---|---|---|
| NO | Variant ID | Amino acid substitution | ConSurf results | |
| | | | Functional/Structural | Score |
| 1 | rs460963763 | L37R | - | 8 |
| 2 | rs478704767 | L40P | - | 7 |
| 3 | rs472547819 | C54S | Structural | 9 |
| 4 | rs434743998 | C54R | Structural | 9 |
| 5 | rs452475196 | G55V | Functional | 9 |
| 6 | rs482754265 | F73C | Structural | 9 |
| 7 | rs450717758 | K75Q | Functional | 9 |
| 8 | rs455759048 | I91T | Structural | 9 |
| 9 | rs456928487 | C95W | Structural | 9 |
| 10 | rs464204228 | C95G | Structural | 9 |
| 11 | rs464204228 | C95S | Structural | 9 |
| 12 | rs437108565 | C96Y | Structural | 9 |
| 13 | rs468436354 | F97C | Structural | 9 |
| 14 | rs458376971 | C100Y | Structural | 9 |
| 15 | rs477050967 | C100G | Structural | 9 |
| 16 | rs462399135 | L105R | Structural | 9 |
| 17 | rs439669768 | Y108S | Structural | 9 |
| 18 | rs457016084 | C109G | Structural | 9 |

434

435

| Table 3: The results of deleterious SNPs effect on IGF-1 protein stability | | | | | | |
|---|---|---|---|---|---|---|
| NO | Variant ID | Amino acid substitution | I-Mutant | | MUpro | |
| | | | Stability | DDG (kcal/mol) | Stability | DDG (kcal/mol) |
| 1 | rs460963763 | L37R | decrease | -0.93 | decrease | -1.6686468 |
| 2 | rs478704767 | L40P | decrease | -0.95 | decrease | -2.0099046 |
| 3 | rs472547819 | C54S | decrease | -1.19 | decrease | -1.2715878 |
| 4 | rs434743998 | C54R | decrease | -1.15 | decrease | -0.93101523 |
| 5 | rs452475196 | G55V | decrease | -0.84 | increase | 0.22557721 |
| 6 | rs482754265 | F73C | decrease | -2.90 | decrease | -1.2257966 |
| 7 | rs450717758 | K75Q | decrease | -0.59 | decrease | -0.9074037 |
| 8 | rs455759048 | I91T | decrease | -3.24 | decrease | -1.7065422 |
| 9 | rs456928487 | C95W | decrease | -0.50 | decrease | -0.97228795 |
| 10 | rs464204228 | C95G | decrease | -1.56 | decrease | -1.8768055 |
| 11 | rs464204228 | C95S | decrease | -0.77 | decrease | -1.310688 |
| 12 | rs437108565 | C96Y | increase | 0.12 | decrease | -1.0418833 |
| 13 | rs468436354 | F97C | decrease | -2.02 | decrease | -0.89818623 |
| 14 | rs458376971 | C100Y | increase | 0.40 | decrease | -0.89385889 |
| 15 | rs477050967 | C100G | decrease | -1.37 | decrease | -1.7712997 |
| 16 | rs462399135 | L105R | decrease | -0.59 | decrease | -1.8234098 |
| 17 | rs439669768 | Y108S | decrease | -0.82 | decrease | -0.99588314 |
| 18 | rs457016084 | C109G | decrease | -1.26 | decrease | -1.4211852 |

436

437

438

439

440

441

Table 4: The results of Mutpred2 analysis of the deleterious SNPs

| NO | Variant ID | Amino acid substitution | Mutpred2 score | Molecular mechanisms | Probability |
|---|---|---|---|---|---|
| 1 | rs450717758 | K75Q | - | - | - |
| 2 | rs472547819 | C54S | 0.606 | Gain of Intrinsic disorder | 0.47 |
| 3 | rs482754265 | F73C | 0.690 | Altered Transmembrane protein | 0.32 |
| 4 | rs455759048 | I91T | 0.728 | Gain of Intrinsic disorder | 0.33 |
| 5 | rs478704767 | L40P | 0.750 | Altered Signal peptide | 0.10 |
| 6 | rs457016084 | C109G | 0.774 | Altered Disordered interface | 0.33 |
| 7 | rs434743998 | C54R | 0.779 | Gain of Intrinsic disorder | 0.36 |
| 8 | rs460963763 | L37R | 0.781 | Altered Signal peptide | 0.03 |
| 9 | rs439669768 | Y108S | 0.796 | Gain of Intrinsic disorder | 0.55 |
| 10 | rs468436354 | F97C | 0.833 | Altered Metal binding | 0.47 |
| 11 | rs464204228 | C95S | 0.835 | Altered Metal binding | 0.25 |
| 12 | rs452475196 | G55V | 0.842 | Altered Transmembrane protein | 0.37 |
| 13 | rs464204228 | C95G | 0.846 | Altered Metal binding | 0.25 |
| 14 | rs462399135 | L105R | 0.884 | Gain of Intrinsic disorder | 0.44 |
| 15 | rs456928487 | C95W | 0.891 | Gain of Loop | 0.27 |
| 16 | rs477050967 | C100G | 0.904 | Altered Metal binding | 0.28 |

464

| Table 5: | | | |
|---|---|---|---|
| Variant | Wild and Mutant substitution | Amino acid properties | Location |
| F97C |  | The mutant residue is smaller; this might lead to loss of interactions | The mutation is located within a domain; it might disturb protein function. |
| C95S |  | The wild-type residue is more hydrophobic than the mutant residue, hydrophobic interactions, either in the core of the protein or on the surface, will be lost. | The mutation is located within a domain; it might disturb protein function. |
| G55V |  | The mutant residue is bigger; this might lead to bumps.<br>The torsion angles for this residue are unusual. only glycine is flexible enough to make these torsion angles, mutation into another residue will force the local backbone into an incorrect conformation and will disturb the local structure. | The mutation is located within a domain; it might disturb protein function. |
| C95G |  | The mutant residue is smaller than the wild-type residue, this might lead to loss of interactions. The wild-type residue is more hydrophobic than the mutant residue. It can disrupt hydrophobic interactions, either in the core or on the surface of protein. | The mutation is located within a domain; it might disturb protein function. |
| L105R |  | The mutant residue is bigger than the wild-type residue, it might lead to bumps. The wild-type residue charge was neutral, while the mutant residue charge is positive. this can cause repulsion of ligands or other residues with the same charge. The wild-type residue is more hydrophobic than the mutant residue, it can disrupt hydrophobic interactions, either in the core or on the surface of protein | The mutation is located within a domain; it might disturb protein function. |
| C95W |  | The mutant residue is bigger than the wild-type residue, it might lead to bumps. | The mutation is located within a domain; it might disturb protein function. |

| | | | |
|---|---|---|---|
| C100G |  | The mutant residue is smaller than the wild-type residue, this might lead to loss of interactions. The wild-type residue is more hydrophobic than the mutant residue, it can disrupt hydrophobic interactions, either in the core or on the surface of protein | The mutation is located within a domain; it might disturb protein function. |
| W: Wild M: Mutant | | | |

494

495

496

497

498