Acta Scientific Biotechnology

Volume 2 Issue 3 June 2021

Research Article

Genome Wide Comparative Analysis of Igf1 Gene Using Bioinformatic Tools

Abdul Manan*

Department of Bioinformatics and Biotechnology, Government College University Faisalabad, Faisalabad, Pakistan

***Corresponding Author:** Abdul Manan, Department of Bioinformatics and Biotechnology, Government College University Faisalabad, Faisalabad, Pakistan.

Received: April 16, 2021

Published: May 29, 2021

Abstract

Insulin like growth factor IGF1 is also known as somatostadin C. It is involved in somatic growth of cell like muscle, kidney or cartilage cells. IGF1 is primarily expressed in the liver. This gene has significance growth of individuals. IGF family is involved in cellular signaling and it is also responsible for cell division. Mutation in igf1 gene can cause Laron dwarfism and acromegaly. This study is performed to understand genomic variations and for sequence analysis. Freely available online tools such as pfam, gene structure display server, MEME motif, Serial Cloner and Mega 7.0 were used to perform sequence analysis. The results from this analysis showed that IGF1 is a polymorphic gene and amino acid variation can be seen among igf1 gene of different species. This study could be further use for studying the function of igf1 gene.

Keywords: Igf1; Insulin Like Growth Factor; Gene; Database

Introduction

IGF1 insulin like growth factor 1 is expressed in many cells postnatal and fetal tissues. In mammals, somatic growth is caused through signalling of insulin like growth factor 1 [1]. The synergistic interaction between insulin like growth factor 1 and growth hormone cause somatic growth in postnatals. The result of over expression of IGF1 is malignant formation and under expression of IGF1 also cause transformation of malignant cells. IGF1 gene is also known as somatostadin C. IGF1 plays an important role in growth of organism and maintenance of biological systems. Cell differentiation, proliferation and migration responses are stimulated by IGF1. This gene has links in pathophysiology of many cancers. During development of muscles, IGF1 also induces many mytogenic responses. In most of the cases, insulin like growth factor 1 act as differentiation factor and mitogen in the cells. Many tissues produces IGF1. IGF1 expression have paracrine and endocrine functions modes [2].

Alternative splicing of IGF1 gene produce different mRNA transcripts and IGF1 gene is precursor of many proteins. These proteins have different extension peptides and different structures. However, mature peptide of these precursor proteins is same as that of IGF1 due to common parts from same gene. Insulin like growth factor 1 bind to many receptors for their activation. IGF1 gene is located on chromosome number 12 of human and it is spread over the region of 84690 bp DNA. In primates and mammals, the regions of IGF1 gene are highly conserved. Due to the presence of 7 exons in IGF1 gene heterogenous mRNA transcripts are arisen due to many transcription sites. Due to polyadenylation signals and alternative splicing, heterogenous mRNAs are produced. Then all these heterogenous mRNAs encodes for multiple peptides precursors which go through many post translational modifications. There are two classes of mRNA transcripts which are transcribed from IGF1 gene [3]. These are classified as class 1 and class 2. In class 1 initiation sites are present on promotor 1 and in class 2, promotor two is

used as leader exon. Different mRNA transcripts are produced by alternative splicing of exon 5. Due to different combination of terminal exons and promotors, different transcripts of IGF1 gene are produced. It has also been demonstrated that paracrine IGF1 synthesis is associated with promotor 1 and it has also an affect on the interaction of IGF binding proteins. There is huge number of transcripts in the tissue which are induced by promotor 1 while promotor 2 transcripts are only expressed in kidney and liver [4].

To date, mechanism, and physiology of IGF1 transcripts is unclear and function of spliced transcripts is also unknown. However, these transcripts are responsible for many stimuli and responses across the cell. It depends upon the condition that which transcript will be produced during transcription. These conditions may include prostate cancer, endometriosis, or muscle damage due to exercise. The precursor which are transcribed by IGF1 gene are IGF1-IEa, IGF1-IEb and IGF1-IEc. Single peptide insulin like growth factor 1 prohormone is produced by alternation splicing of 1, 2, 3 exons at 5' end. At exon 1, there are 4 sites that can start transcription and it can produce class 1 mRNA which contains almost 3 signal peptides. Due to alternative splicing, regulation of insulin like growth factor 1 gene occurs at many levels. Bioavailability and bioactivity of IGF1 is influenced by alternative splicing of its exons [5].

In 1986, IGF1 cDNA sequence of human was cloned. The results of sequencing revealed that number of nucleotides in IGF1 cDNA are around 4989 and it encodes of 1367 amino acids precursors. An α chain is produced by cleavage of signal peptide residue. 408 residues of tyrosine kinase in cytoplasmic domain and a transmembrane sequence is produced. An alternative form of human insulin like growth factor one is identified which has deletion of CAG base pair. This deletion also substituted the Arg amino acid to Thr and Gly. In monomers of subfamily of isulin receptors, 11 different regions are present which are responsible for extracellular protein expression. The domains of IGF1 gene are classified as L1, L2, Cis-rich, fibronectin type3 and tyrosine kinase domains [6].

Supplementary material

Proposed names	Gene Locus	Protein accession #	RNA accession#	Exons	Chr #	ORF length	Amino acid length	Start of Genomic Location	Conserved domains in protein sequence
IGF1	LOC100418882	NP_000609.1	NM_000618.5	7	12	462	153	102217316	L1, L2, Cis-rich, fibronectin type3 and tyrosine kinase

Materials and Methods

IGF1 gene identification

Protein and DNA sequences of igf1 gene were obtained from NCBI (www.ncbi.nlm.nih.gov). Protein blast was performed using protein sequence of igf1 gene of human. Alignment was performed with other species. Variants which had longest open reading frames were selected, and their E value was carefully checked. To eliminate the sequences that do not contain conserved motifs, the sequence were processed in online tools such as (http://pfam.janelia.org/) pfam database and (http://smart.embl-heidelberg.de/) SMART database. The processing of sequences through these databases gave those sequences that had conserved domains. Information such as accession numbers, chromosomal location, gene locus, number of amino acids and length of open reading frames were obtained from NCBI website.

Figure 1: Significant matches of motifs of igf1 gene from different species using pfam.

Identification of structure of gene

Structure of gene was obtained from gene structure display server which is an online freely available tool. It was used to determine the exon and introns in igf1 gene.

Figure 2: Schematic diagram of IGF1 gene.

IGF1 gene have 7 exons and it is located on chromosome number 12. It is located on q arm of chromosome. Molecular weight of igf1 gene is 7690 D.

Figure 3: Structure of IGF1 protein obtained from PDB.

Figure 4: 3D structure of insulin like growth factor obtained from PDB.

Phylogenetic analysis

Mega 7.0 software is used for phylogenetic analysis. Neighbour-joining and maximum likelihood methods were performed for analysis of phylogeny of igf1 gene. Basic local alignment was performed by taking the protein sequence of human igf1 gene. Then sequences were selected by considering E-value of sequences. From the Genebank databases sequence of closest family members of igf1 gene were obtained. These members were igf2 and igf3. Same screening mechanism was performed to screen out sequences having conserved motifs using pfam database. The phylogenetic tree was created using Mega 7.0 software. Percentage of bootstrap value is denoted at the node of each branch. 5 percent divergence was observed in each amino acids of each sequence and this value is denoted in the form of bar at the bottom of the tree.

14 species were screened out for analysis of igf1 gene and 10 species were selected for igf2 gene. For neighbour-joining, 15 species were selected for igf3 gene. The protein sequence of all these species were obtained form NCBI database. These protein sequences were process in Mega 7.0 software to create evolutionary tree.

Figure 5: Phlogenetic analysis of igf1 gene with its other family members.

Identification of conserved motif

For identification of conserved motifs in igf1 gene, MEME program is used. MEME is an online freely available tool which is used to identify the conserve motive in the sequence of gene. Number of motifs was selected as 10 and motif distribution was selected as

independent [7], which means it could be any repetition of motif. For the display of exons and introns in igf1 gene, gene structure display server was used.

Figure 6: The height of amino acid names represents the conservation of the residues.

Figure 7: Conserved motif location of igf1 gene is displayed. it is drawn the basis of p-value.

Gene duplications and chromosomal location

IGF1 gene is located on chromosome number 12 in human genome. It is present on the q arm. The band of igf1 gene is denoted as 12q23.2. On chromosome number 12, the starting location of IGF1 gene is from 102,395,874 bp to 102,481,744 bp.

Restriction sites in igf1 gene were identified using serial cloner tool. The results are shown in figure below.

Figure 8

Ontology analysis of IGF1 gene

Ontology analysis of igf1 gene was performed. In humans, IGF1 is mainly produced in the liver. This gene effects the growth hormone production. Igf1 gene is involved in growth of body and it is also responsible for growth of every cells mainly kidney, liver, cartilage, and muscle cells [8]. Igf1 secretion also stimulates DNA synthesis. As name indicates, insulin like growth factor has effects like insulin. Igf1 binds to two receptors in the cell. One is insulin receptor IR and the other is insulin like growth factor receptors. Intracellular signalling is produced when IGF1 bind to IGF1-R [9]. Through the expression of tyrosine kinase, Igf1 stimulate and activate its binding to insulin like growth factor receptor. Transductions cascades are used for further intracellular signalling. In the survival of the cell, IGF1 play an important role [10,11]. It also mediate the activity of paracrine gland. IGF1 shows metabolic effect by signalling the cell about sufficiency of nutrients to start the cell division. Igf1 play an important role in inhibition of apoptosis and stimulate the production of many other proteins in the cell. The closest member of igf1 is igf2 which also binds to insulin like growth factor 2 receptors [12]. Igf2 receptor is also known as mannose-6-phosphate receptor. It doesn't have transduction capacity, but it plays an important role by binding to igf1 receptors [13-17].

Results and Discussion

In many researches, insulin like growth factor 1 gene is describes as candidate for growth. In this study structure analysis of gene was performed which showed that it has 7 exons and this gene is located on chromosome number 7. SMART and Pfam analysis

indicated that igf1 gene has 11 conserved regions. It has multiple domains. These domains are L domains, tyrosine kinase domains and cis-rich type 3 domains. It has two closely related members which are igf2 and igf3. The structure of insulin like growth factor is similar to insulin and it plays an important role in the growth of children. Mutation in the igf1 gene cause growth related disorders. Mutation in igf1 gene cause Laron dwarfism and acromegaly in humans.

Conclusion

IGF1, also known as somatostadin C is important gene for growth. The conclusion of this study is that insulin like growth factor gene is a polymorphic gene and by aligning with sequence of different species, it is seen that amino acid variation is present between different gene. Igf1 has structural similarity with insulin. The defect in igf1 could cause Laron dwarfism and acromegaly. This study could bring insight to existing knowledge of Igf1 gene.

Bibliography

1. Chew SL., *et al*. "An alternatively spliced human insulin-like growth factor-I transcript with hepatic tissue expression that diverts away from the mitogenic IBE1 peptide". *Endocrinology* 136.5 (1995): 1939-1944.

2. A D., *et al*. "Analysis of insulin growth factor (IGF-1) gene in some selected monogastrics: An Insilico approach". *IJMBOA* 3.2 (2018).

3. Barton ER., *et al*. "The insulin-like growth factor (IGF)-I E-peptides are required for isoform-specific gene expression and muscle hypertrophy after local IGF-I production". *Journal of Applied Physiology* 108.5 (2010): 1069-1076.

4. Selvakumar P and Pandey A. "Comparative studies on inulinase synthesis by Staphylococcus sp. and Kluyveromyces marxianus in submerged culture". *Bioresource Technology* 69.2 (1999): 123-127.

5. Yang T-L., *et al*. "Comprehensive association analyses of IGF1, ESR2, and CYP17 genes with adult height in Caucasians". *European Journal of Human Genetics* 16.11 (2008): 1380-1387.

6. Yang SY and Goldspink G. "Different roles of the IGF-I Ec peptide (MGF) and mature IGF-I in myoblast proliferation and differentiation". *FEBS Letter* 522 (2002): 156-160.

7. Ge Y., *et al*. "Identification of differentially expressed genes in cervical cancer by bioinformatics analysis". *Oncology Letter* (2018).

8. López-Bermejo A., *et al*. "Insulin resistance is associated with increased serum concentration of IGF-binding protein-related protein 1 (IGFBP-rP1/MAC25)". *Diabetes* 55.8 (2006): 2333-2339.

9. Baxter RC. "Insulin-like growth factor (IGF)-binding proteins: interactions with IGFs and intrinsic bioactivities". *American Journal of Physiology - Endocrinology and Metabolism* 278.6 (2000): E967-976.

10. Duguay SJ., *et al*. "Mutational analysis of the insulin-like growth factor I prohormone processing site". *Journal of Biological Chemistry* 270.29 (1995): 17566-17574.

11. Cohen P. "Problems with Reclassification of Insulin-Like Growth Factor I Production and Action Disorders". *The Journal of Clinical Endocrinology and Metabolism* 91.11 (2006): 4235-4236.

12. Rotwein P., *et al*. "Organization and sequence of the human insulin-like growth factor I gene. Alternative RNA processing produces two insulin-like growth factor I precursor peptides". *Journal of Biological Chemistry* 261.11 (1986): 4828-4832.

13. Brisson BK and Barton ER. "New Modulators for IGF-I Activity within IGF-I Processing Products". *Frontiers in Endocrinology* (2013): 4.

14. Philippou A., *et al*. "The Complexity of the IGF1 Gene Splicing, Posttranslational Modification and Bioactivity". *Molecular Medicine* 20.1 (2014): 202-214.

15. Lund PK. "Insulin-like growth factor I: molecular biology and relevance to tissue-specific expression and action". *Recent Progress in Hormone Research* 49 (1994): 125-148.

16. Metzger F., *et al*. "Separation of fast from slow anabolism by site-specific PEGylation of insulin-like growth factor I (IGF-I)". *Journal of Biological Chemistry* 286.22 (2011): 19501-19510.

17. Yuan Y., *et al*. "Gene expression profiles and bioinformatics analysis of insulin-like growth factor-1 promotion of osteogenic differentiation". *Molecular Genetics and Genomic Medicine* 7.10 (2019): e00921.

Volume 2 Issue 3 June 2021