Journal of Computational Intelligence in Bioinformatics
ISSN: 0973-385X Volume 2 Number 3 (2009) pp. 177–184

http://www.ripublication.com/jcib.htm

Insulin Like Growth Factor 1 Receptor: Gene Assesment Using *Insilico* Approch

***J. Gaikwad Vishnu, V. N. Ghasghase, #R. S. Kuptekar and S. S. Ghorpade**

Dept. of Biotechnology Engineering, Tatyasaheb Kore Institute of Engineering and Technology, Warananagar, Dist- Kolhapur, MS, INDIA-416113
#Dept. of Biotechnology and Bioinformatics, MGM's College of Computer Science, Near Airport, Nanded
*Author for Correspondence: * vishnu.gaikwad@gmail.com*

Abstract

As the biological data generated abundantly needs to transform these data into mining full information which will be very useful to do further research related to the information. By considering this point of view, we turned on most important global issue of cancer.

In this paper we have present the annotation of IGF1R gene by using Bioinformatics approach. We have generated detailed information about IGF1R gene. The information like total Nucleotide composition, Amino acid composition, no. of coding regions (exons), TATA boxes, Poly A sites, etc. This information is very valuable in further research on cancer like diseases.

Keywords: IGF1R, TATA boxes, ProScan, PolyA site, GENSCAN 1.0.

Introduction

Approximately 2000 kinases are known and more than 90 Protein Tyrosine Kinases (PTKs) have been found in the human genome. They are divided into two classes, receptor and non-receptor PTKs. At present, 58 receptor tyrosine kinases (RTKs) are known, grouped into 20 subfamilies. They play pivotal roles in diverse cellular activities including growth, differentiation, metabolism, adhesion, motility, death (S B Bhise et al., 2004) RTKs are composed of an extracellular domain, which is able to bind a specific ligand, a transmembrane domain, and an intracellular catalytic domain, which is able to bind and phosphorylate selected substrates.

The insulin-like growth factor I (IGF I)' receptor is (Location**:** 15q25-q26) a heterotetrameric transmembrane glycoprotein that binds IGF I with high affinity and IGF I1 with somewhat lower affinity (Rechler, M. M et al., (1985)). In structure and

action, the IGF I receptor is closely related to the insulin receptor, with greater than 50% overall amino acid homology (Rechler, M. M.et al., (1985), Goldfine, I. D. (1987)). Both receptors are composed of two a-subunits and two p-subunits connected by disulfide bonds to form a heterotetrameric complex of generally similar design (Chernausek, S. D et al., (1981)). Both contain tyrosine kinase domains, which are highly homologous (80-95%) in primary amino acid composition. Despite these similarities, the two receptors differ in ligand specificities, tissue distribution, and biological roles.

Insulin/IGF-I action is mediated by the insulin receptor (IR) and the IGF-I receptor (IGF-IR), respectively, which have very similar heterodimeric $\alpha_2ß_2$ structure and belong to the family of receptor tyrosine kinases (Ullrich A et al., 1986). Upon Ligand binding, insulin/IGF-IR undergoes autophosphorylation on tyrosine residues, which activates the intracellular tyrosine kinase of the ß-subunit (Kasuga M et al., 1982).

IGF1 and IGF2 both bind to IGFRI and stimulate multiple signaling cascades, including mitotic, anti-apoptotic, and transformation-associated pathways (Baserga, 2000).

Ligand binding results in receptor PTK autophosphorylation and activation, followed by tyrosine phosphorylation of key substrates including insulin receptor substrate-1 (IRS-1), Shc, and PI3 kinase (PI3K) p85 subunit (reviewed in Leroith *et al*., 1995).

The IGF-1R is implicated in several cancers, (Warshamana-Greene GS et al.,(2005).,Jones HE et al., (2004).) most notably breast cancer. In some instances its anti-apoptotic properties allow cancerous cells to resist the cytotoxic properties of chemotherapeutic drugs or radiotherapy.

So, this gene is important by our point of view. By annotating this gene, we assure in future, this annotated information of the gene will be helpful in designing the drug molecules which act as inhibitors or antagonist. These drug molecules will be better in their specificity.

We have done annotation of this IGFIR gene applying various Bioinformatics tools and publicly available databases.

Materials and Methods

Data Collection

The NCBI houses genome sequencing data in GenBank and an index of biomedical research articles in PubMed Central and PubMed, as well as other information relevant to biotechnology. All these databases are available online through the Entrez search engine. Prosite: Database of protein domains, families and functional sites.

We have collected the nucleotide sequence of IGF1R from GeneBank (NCBI), DDBJ (DNA data bank of Japan) and EMBL (European Molecular Biology Laboratory) are collaborative firms of each other. Fully funded by respective governments.

Annotation

As the biological data generated abundantly needs to transform these data into mining full information which will be very useful to do further research relate to the information. By considering this point of view, we turned on most important global issue of cancer. As we were students of bioinformatics, we inspired to annotate IGFIR gene. We also took into consideration our PC configuration; we had Pentium4, 512 RAM. For this we decided to use popular online tools and databases for gene annotation e.g. ORF Finder: graphical analysis tool which finds all open reading frames of a selectable minimum size in a user's sequence or in a sequence already in the database, available at http://www.ncbi.nlm.nih.gov/projects/gorf/.

This tool identifies all open reading frames using the standard or alternative genetic codes. The deduced amino acid sequence can be saved in various formats and searched against the sequence database using the WWW BLAST server available at http://blast.ncbi.nlm.nih.gov/Blast.cgi.

ProScan- find promoter sequences from given DNA sequence available at http://www- bimas.cit.nih.gov/molbio/proscan/, GENSCAN 1.0 available at http://bioinformatics.ubc.ca/resources/tools/genscan this server provides access to the program Genscan for predicting the locations and exon-intron structures of genes in genomic sequences from a variety of organisms. This server can accept sequences up to 1 million base pairs (1 Mbp) in length. GeneAtlas available at

http://128.252.166.31/gene/main.jsp;jsessionid=2739071214885982959:

By using above tools and databases we proceed for annotation of IGF1R gene.

Results and discussion

(1) We first calculated nucleotide contents of IGF1R gene nucleotide sequence. Total length of the gene is **11242 base pairs**. Molecular Weight is **3417222.00 Daltons**, of single stranded and Molecular Weight of double stranded is **6833894.00 Daltons**, the **G+C** content of is **49.92%** and **A+T** content is **50.08%**

Table:

Nucleotide	**Number**	**Mol%**
A	2738	24.36
C	2861	25.45
G	2751	24.47
T	2892	25.72

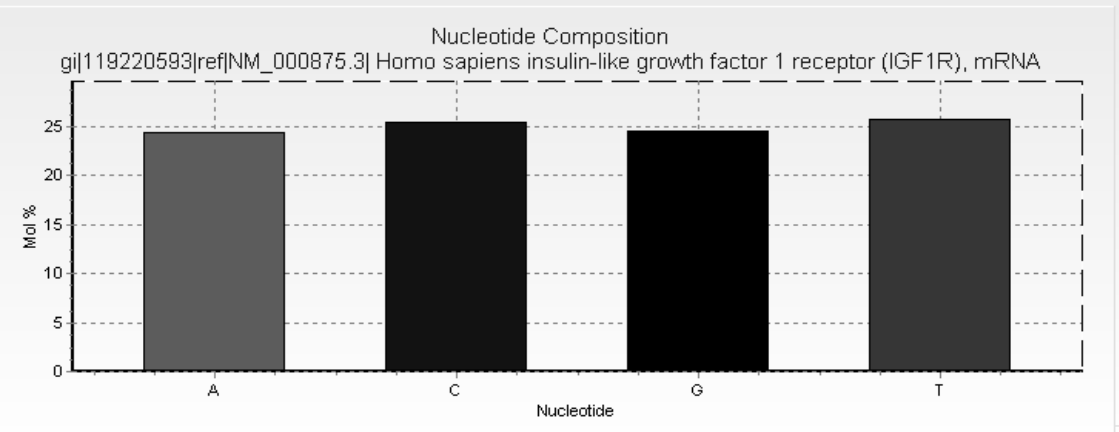

Figure 1: Graph of A, T, G, C composition of gene is generated by BioEdit.

We also calculated the amino acid composition of translated the nucleotide sequence of the IGF1R gene by using BioEdit. We found that the **Leucine L** is having high number i.e. **116 residues**. If we think about the molecular percentage of Leucine in the protein sequence is **8.49%.** Another amino acid is **Glutamic acid (Glu E)** with number **113 residues** and **8.27%** in total. The detail amino acid composition is given in the table below.

The amino acid contents of the protein of our gene is as fallows

IGF1R_HUMAN Insulin-like growth factor 1 receptor precursor (Insulin-like growth factor I receptor) (IGF-I receptor) (CD221 antigen) [Contains: Insulin-like growth factor 1 receptor alpha chain; Insulin-like growth factor 1 receptor beta chain]

Length = 1367 amino acids

Molecular Weight = 154785.13 Daltons

Table:

Amino Acid	**Number**	**Mol%**
Ala **A**	72	5.27
Cys **C**	44	3.22
Asp **D**	62	4.54
Glu E	**113**	**8.27**
Phe **F**	47	3.44
Gly **G**	91	6.66
His **H**	23	1.68
Ile **I**	73	5.34
Lys **K**	69	5.05
Leu L	**116**	**8.49**
Met **M**	39	2.85
Asn **N**	85	6.22

Pro **P**	81	5.93
Gln **Q**	35	2.56
Arg **R**	80	5.85
Ser **S**	97	7.10
Thr **T**	69	5.05
Val **V**	87	6.36
Trp **W**	21	1.54
Tyr **Y**	63	4.61

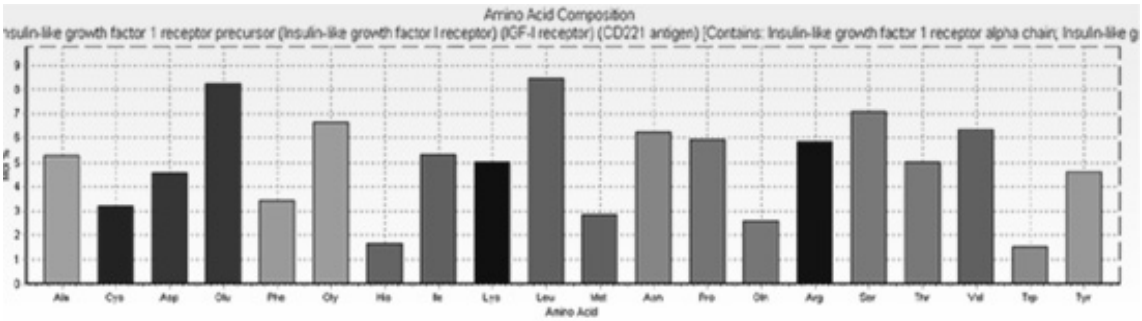

Figure 2: Graph showing Amino Acid composition in IGF1R (BioEdit).

(2) We also found six possible reading frames of IGF1R gene using ORF finder. This will give the six possible frames (three from each end). This six reading frames state that gene has expressed by one of the above six reading frames.

(3) The promoter sequence of the IGF1R gene is predicted using Proscan tool.

Promoter predictions for 1 eukaryotic sequence with score cutoff 0.80 (transcription start shown in larger, bold and underlined font):

Promoter predictions for IGF1R Gene sequence:

Start End Score Promoter Sequence

1245 1295 0.97

TTGTCCTTCCTAAAAAACCTTCGCCTCATCCTAGGAGAGG**A**GCAGCTAGA

1442 1492 0.91

AGTGACGGGGACTAAAGGGCGCCAAAGCAAAGGGGACATA**A**ACACCAGGA

1468 1518 0.82

GCAAAGGGGACATAAACACCAGGAACAACGGGGAGAGAGC**C**TCCTGTGAA

2587 2637 1.00

ACTCCATCTTTTTAAAGTGGCCGGAACCTGAGAATCCCAA**T**GGATTGATT

5602 5652 1.00

GATGCCTTTTTATAAATACATCCCCCATCCCTGCTCCCAC**C**TGCCCCTTT

6634 6684 0.94

TCAATCACTGTAGAAAAGCCCCATTATGAATTTAAATTTC**A**AGGAAAGGG

8925 8975 0.99

CACTCCCTTCTAAAACACAGGCGCCCTCCTGGTGACAGTG**A**CCCGCCGTG

11010 11060 0.96

ATCCTGTTTATATAAAAAATCAGTAGATGAAAAAAATTTC**A**AAATGTTTT

(4) Another common aspect in the gene prediction is to prediction of coding regions (exons) in gene sequence. We have done it using NCBI's GENEATLAS. Below is the result of GENEATLAS, there are total 21 coding regions (exons) present in our gene.

TYPE	functioning gene
STRUCTURE	308,79 kb 21 Exon(s)

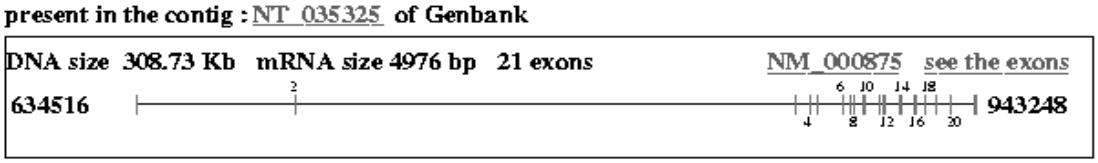

Figure 3: Graphical representation of the gene.

(5) Online gene annotation tool LIBRA, we here predicted primers within our sequence

Primer3 result

PRIMER PICKING RESULTS FOR gi|119220593|ref|NM_000875.3| Homo sapiens insulin-like growth factor 1 receptor (IGF1R), mRNA

No mispriming library specified
Using 0-based sequence positions

OLIGO	start	len	tm	gc%	any	3'	seq
LEFT PRIMER	2661	20	59.98	50.00	5.00	2.00	CACAAGTTGAGGATCAGCGA
RIGHT PRIMER	3448	20	60.00	50.00	3.00	1.00	GTGGACGAACTTATTGGCGT

SEQUENCE SIZE: **11242**
INCLUDED REGION SIZE: **11242**

(6) PolyA site and TATA box have been predicted by using GeneBuilder tool of WebGene

Table:

Poly A site prediction		**TATA box prediction**	
Position	Pattern	Position	Pattern
27	CAAATAAATA	606	TGAATAAAGA
128	GCAATTAATT	636	TCTATAAGTG
254	CGCTTAAAAA	636	TCTATAAGTG
545	GCAATTTAAG	636	TCTATAAGTG
606	TGAATAAAGA	1402	TGCATTTAAG
1403	GCATTTAAGC	1621	GTTATCAGGC
1474	TAAATACAAG	2105	TCTATAATAG
2108	ATAATAGAGT	2105	TCTATAATAG
2206	ATAATAAACG	2115	AGTATATAGG
		2184	TTTATAAAAG
		2184	TTTATAAAAG
		2184	TTTATAAAAG
		2229	TGTGTATATG

Conclusion

In this paper we have present the annotation of IGF1R gene by using Bioinformatics approach. We have generated detailed information about IGF1R gene. The information like total Nucleotide composition, Amino acid composition, no. of coding regions (exons), TATA boxes, Poly A sites, etc. This information is very valuable in further research on cancer like diseases.

References

[1] Baserga R. (2000). *Oncogene*, 19: 5574-5581. MEDLINE

[2] Chernausek, S. D., Jacobs, S., and Van Wyk, J. J. (1981) Biochemistry 20,

[3] Goldfine, I. D. (1987) Endocr. Reu. 8,235-255 7345-7350

[4] Jennifer M. Gross and Douglas Yee, Departments of Medicine and Pharmacology, University of Minnesota Cancer Center, 420 Delaware Street SE, cMinneapolis, MN, USA

[5] Jones HE, Goddard L, Gee JM, Hiscox S, Rubini M, Barrow D, Knowlden JM, Williams S, Wakeling AE, Nicholson RI (2004). "Insulin-like growth factor-I receptor signalling and acquired resistance to gefitinib (ZD1839; Iressa) in human breast and prostate cancer cells". *Endocr. Relat. Cancer* 11 (4): 793–814. doi: 10.1677/erc.1.00799. PMID 15613453

[6] Kasuga M, Karlsson FA, Kahn CR: Insulin stimulates the phosphorylation of the 95,000-dalton subunit of its own receptor. *Science* 215:185–187,

1982[Abstract/Free Full Text] 7. Leroith D, Werner H, Beitner-Johnson D, Roberts JrCT. (1995). *Endocr. Rev.,* 16: 143-163. MEDLINE

[7] RH Gunby et al. *Oncogenic Fusion Tyrosine Kinases As Molecular Targets for Anti-Cancer Therapy*. Anti-Cancer Agents in Medicinal Chemistry, 2007; 7:594-611.

[8] Rechler, M. M., and Nissley, S. P. (1985) Annu. Reu. Physiol. 47, 425-442

[9] S B Bhise, Abhijit D. Nalawade and Hitesh Wadhawa, *Role of protein tyrosine kinase inhibitors in cancer therapeutics*. Indian Journal of Biochemistry & Biophysics, 2004 Dec; 41: 273-280. ISSN 0301-1208.

[10] Ullrich A, Gray A, Tam AW, Yang-Feng T, Tsubokawa M, Collins C, Henzel W, Bon TL, Kathuria S, Chen E, Jacobs S, Francke U, Ramachandran J, Fujita-Yamaguchi Y: Insulin-like growth factor receptor primary structure: comparison with insulin receptor suggests structural determinants that define functional specificity. *EMBO J* 5:2503–2512, 1986[Medline]

[11] Warshamana-Greene GS, Litz J, Buchdunger E, García-Echeverría C, Hofmann F, Krystal GW (2005). "The insulin-like growth factor-I receptor kinase inhibitor, NVP-ADW742, sensitizes small cell lung cancer cell lines to the effects of chemotherapy". Clin. Cancer Res. 11 (4): 1563–71. doi:10.1158/1078-0432.CCR-04-1544. PMID 15746061.