Application of unsupervised clustering algorithm on a data set of Insulin like Growth Factor Receptor sequences

R Rambabu, Dr. P Srinivasa Rao

Abstract - Clustering can be considered as the most unsupervised learning problem. A component being considered in a clustering algorithm is the distance measure between data points in the given set. In this work, we present a modified single and complete linkage clustering techniques were implemented in python language and applied on a dataset of IGF1R (Insulin like Growth Factor Receptor) protein sequences to generate orthologous clusters of sequences and the phylogenetic trees in each case were presented. From analysis, it was observed that IGF1R_XENLA and PIG sequences formed one group in single linkage whereas IGF1R_XENLA and RAT appeared under one clade in complete linkage algorithm.

Index Terms: Clustering, IGF1R, Insulin, Python, Receptors, Similarity, Sequencing.

———————————— ◆ ————————————

1 INTRODUCTION

A cluster is a collection of objects which have similarities among other objects in the group and are probably dissimilar to the objects belonging to other clusters. Clustering can be considered the most important unsupervised learning problem. The goal of clustering is to determine the intrinsic grouping in a set of unlabeled data [1] [2].Clustering algorithms may be classified as, Exclusive, overlapping, hierarchical and probabilistic clustering. In the first case data are grouped in an exclusive way, so that if a certain datum belongs to a definite cluster then it could not be included in another cluster [3]. The overlapping clustering uses fuzzy sets to cluster data, so that each point may belong to two or more clusters with different degrees of membership. A hierarchical clustering algorithm is based on the union between the two nearest clusters. Finally, the last kind of clustering uses a completely probabilistic approach [4] [5] [6]. An important component of a clustering algorithm is the distance measure between data points.

Human endothelial cells express IR, IGF-1 receptors (IGF-1R), and hybrid receptors (IR/IGF-1R) composed of heterodimers [7] [8]. IGF-1R are more abundant (10-fold higher) than IR [9]. In non endothelial cells, IGF-1R expression modulates insulin signalling by altering the levels of hybrid receptors [10]. It has been reported that type 2 diabetes and obesity are associated with increases in IR/IGF-1R expression in insulin-sensitive tissues [11]. In these pathological states, interventions aimed at down regulating IGF-1R expression may augment endothelial insulin sensitivity. Owing to this importance, in this paper, we report single and average linkage algorithms for a set of IGF1R sequences.

2 MATERIALS AND METHODS

2.1 Sequence Data
The IGF1R protein sequences were extracted from Swiss-Prot database [12] [13]. A data set of 6 IGF1R sequences is downloaded as fasta files. These are further subjected to multiple sequence analysis using clustalw [14] to generate similarity scores among all 6 sequences under study.

2.2 Hierarchical Clustering Algorithms
A similarity matrix is constructed for a given a set of N items:

1. *Each item is added to the cluster there by creating N clusters, each containing just one item. Distances or similarities between the clusters are calculated.*
2. *The most similar clusters are paired and joined as a single cluster.*
3. *Now distances are computed between new cluster group and each of the old ones based on single and average linkage clustering. In single linkage, the distance between one cluster and another cluster is considered to be identical to the shortest distance from any object of one cluster to any object of the other cluster. In complete linkage, the distance between one cluster and another cluster is considered to be equal to the greatest distance from any object of one cluster to any object of the other cluster.*
4. *Repeat steps 2 and 3 until all items are clustered into a single cluster of size N.*

3 RESULTS AND DISCUSSION
Insulin sequences are searched in Uniprot Knowledgebase. The database was accessed from Expert Protein Analysis System (ExPASy) server maintained by Swiss Institute of

Bioinformatics (SIB).It was reported in literature that *Insulin growth factor* and *Insulin* proteins compete to bind with Insulin receptor. Similarly *Insulin growth factor* protein competes to bind with *Insulin Growth Factor 1 Receptor*. An

I increased IGF-1R expression is associated with increased IR/IGF-1R and reduced numbers of IR holoreceptors [15].

The output of multiple alignments is given below.

```
sp|O73798|IGF1R_XENLA   YRKVFENFLHNSIFVPRPNRRRRDVLAVGNSTVTSYEKNSTTEDFSNFSDSERDDIEYPF  771
sp|P24062|IGF1R_RAT     YRKVFENFLHNSIFVPRPERRRRDVLQVANTTMSSRSRNTTVADTYNITDPEEFETEYPF  778
sp|Q60751|IGF1R_MOUSE   YRKVFENFLHNSIFVPRPERRRRDVMQVANTTMSSRSRNTTVADTYNITDPEEFETEYPF  778
sp|Q05688|IGF1R_BOVIN   ----------NAIFVPRPERKRREVMQIANTTMSSRSRNTTVLDTYNITDPEELETEYPF  50
sp|P08069|IGF1R_HUMAN   YRKVFENFLHNSIFVPRPERKRRDVMQVANTTMSSRSRNTTAADTYNITDPEELETEYPF  777
sp|Q29000|IGF1R_PIG     -----------------------------------------------------------  0


sp|O73798|IGF1R_XENLA   YETKVDYKWERTVISNLQPFTLYRIDIHSCNHEAEKLGCSASNFVFARTMPAAGADDIPG  831
sp|P24062|IGF1R_RAT     FESRVDNK-ERTVISNLRPFTLYRIDIHSCNHEAEKLGCSASNFVFARTMPAEGADDIPG  837
sp|Q60751|IGF1R_MOUSE   FESRVDNK-ERTVISNLRPFTLYRIDIHSCNHEAEKLGCSASNFVFARTMPAEGADDIPG  837
sp|Q05688|IGF1R_BOVIN   FESRVDNK-ERTVISNLRPFTLYRIDIHSCNHEAEKLGCSASNFVFARTMPAEGADDIPG  109
sp|P08069|IGF1R_HUMAN   FESRVDNK-ERTVISNLRPFTLYRIDIHSCNHEAEKLGCSASNFVFARTMPAEGADDIPG  836
sp|Q29000|IGF1R_PIG     ---------ERTVISNLRPFTLYRIDIHSCNHEAEKLGCSASNFVFARTMPAEGADDIPG  51
                                 *******.*******************************  *******

sp|O73798|IGF1R_XENLA   IVNTKEEDDGVIFLGWPEPLRPNGLILMYEIEYKHQGEVHRECVSRQDYRKNGGIKLVRL  891
sp|P24062|IGF1R_RAT     PVTWEPRPENSIFLKWPEPENPNGLILMYEIKYGSQVEDQRECVSRQEYRKYGGAKLNRL  897
sp|Q60751|IGF1R_MOUSE   PVTWEPRPENSIFLKWPEPENPNGLILMYEIKYGSQVEDQRECVSRQEYRKYGGAKLNRL  897
sp|Q05688|IGF1R_BOVIN   PVTWEPRPENSIFLKWPEPENPNGLILMYEIKYGSQVEDQRECVSRQEYRKYGGAKLNRL  169
sp|P08069|IGF1R_HUMAN   PVTWEPRPENSIFLKWPEPENPNGLILMYEIKYGSQVEDQRECVSRQEYRKYGGAKLNRL  896
sp|Q29000|IGF1R_PIG     PVTWEPRPENSIFLKWPEPENPNGLILMYEIKYGSQVEDQRECVSRQEYRKYGGAKLNRL  111
                        *. :  :  *** ****  *.********:*   * *  :*******.*** ** ** **

sp|O73798|IGF1R_XENLA   PPGNYSAQVQAISLYGNGSWTEMVSFCVKLKPDVRNNILQMVVAIPLALSFLLVGIISIV  951
sp|P24062|IGF1R_RAT     NPGNYTARIQATSLSGNGSWTDPVFFYVPAKTTY-ENFMHLIIALPVAILLIVGGLVIML  956
sp|Q60751|IGF1R_MOUSE   NPGNYTARIQATSLSGNGSWTDPVFFYVPAKTTY-ENFMHLIIALPVAILLIVGGLVIML  956
sp|Q05688|IGF1R_BOVIN   NPGNYTARIQATSLSGNGSWTDPVFFYVQAKTTY-ENFIHLMIALPIAVLLIVGGLVIML  228
sp|P08069|IGF1R_HUMAN   NPGNYTARIQATSLSGNGSWTDPVFFYVQAKTGY-ENFIHLIIALPVAVLLIVGGLVIML  955
sp|Q29000|IGF1R_PIG     NPGNYTARIQATSLSGNGSWTEPVFFYVQAKTTY-ENFIHLIIALPVAVLLIVGGLVIML  170
                         ****:*.:.** ** ******.  * *  *      :*::::::*:*:*: ::: *:: ::

sp|O73798|IGF1R_XENLA   CFVFKKRNSNRLGNGVLYASVNPEYFSAAEMYVPDKWEVPREKITMNRELGQGSFGMVYE  1011
sp|P24062|IGF1R_RAT     YVFHRKRNNSRLGNGVLYASVNPEYFSAADVYVPDEWEVAREKITMNRELGQGSFGMVYE  1016
sp|Q60751|IGF1R_MOUSE   YVFHRKRNNSRLGNGVLYASVNPEYFSAADVYVPDEWEVAREKITMNRELGQGSFGMVYE  1016
sp|Q05688|IGF1R_BOVIN   YVFHRKRNSSRLGNGVLYASVNPEYFSAADVYVPDEWEVAREKITMSRELGQGSFGMVYE  288
sp|P08069|IGF1R_HUMAN   YVFHRKRNNSRLGNGVLYASVNPEYFSAADVYVPDEWEVAREKITMSRELGQGSFGMVYE  1015
sp|Q29000|IGF1R_PIG     YVFHRKRNNSRLGNGV--------------------------------------------  186
                         ...:.*** .. ******

sp|O73798|IGF1R_XENLA   GIAKGVVKDEAETKVAIKTVNEAASMRERIEFLNEASVMKEFNCHHVVRLLGVVSQGQPT  1071
sp|P24062|IGF1R_RAT     GVAKGVVKDEPETRVAIKTVNEAASMRERIEFLNEASVMKEFNCHHVVRLLGVVSQGQPT  1076
sp|Q60751|IGF1R_MOUSE   GVAKGVVKDEPETRVAIKTVNEAASMRERIEFLNEASVMKEFNCHHVVRLLGVVSQGQPT  1076
sp|Q05688|IGF1R_BOVIN   GVAKGVVKDEPETRVAIKTVNEAASMRERIEFLNEASVMKEFNCHHVVRLLGVVSQGQPT  348
sp|P08069|IGF1R_HUMAN   GVAKGVVKDEPETRVAIKTVNEAASMRERIEFLNEASVMKEFNCHHVVRLLGVVSQGQPT  1075
sp|Q29000|IGF1R_PIG     -----------------------------------------------------------  186
```

```
sp|O73798|IGF1R_XENLA  LVIMELMTRGDLKSYLRSLRPDTESNSG-QPTPSLKKMIQMAGEIADGMSYLNANKFVHR 1130
sp|P24062|IGF1R_RAT    LVIMELMTRGDLKSYLRSLRPEVE-NNLVLIPPSLSKMIQMAGEIADGMAYLNANKFVHR 1135
sp|Q60751|IGF1R_MOUSE  LVIMELMTRGDLKSYLRSLRPEVEQNNLVLIPPSLSKMIQMAGEIADGMAYLNANKFVHR 1136
sp|Q05688|IGF1R_BOVIN  LVIMELMTRGDLKSYLRSLRPEMEN-NPVLAPPSLSKMIQMAGEIADGMAYLNANKFVHR 407
sp|P08069|IGF1R_HUMAN  LVIMELMTRGDLKSYLRSLRPEMEN-NPVLAPPSLSKMIQMAGEIADGMAYLNANKFVHR 1134
sp|Q29000|IGF1R_PIG    ----------------------------------------------------------- 186


sp|O73798|IGF1R_XENLA  DLAARNCMVTEDFTVKIGDFGMTRDIYETDYYRKGGKGLLPVRWMSPESLKDGVFTTNSD 1190
sp|P24062|IGF1R_RAT    DLAARNCMVAEDFTVKIGDFGMTRDIYETDYYRKGGKGLLPVRWMSPESLKDGVFTTHSD 1195
sp|Q60751|IGF1R_MOUSE  DLAARNCMVAEDFTVKIGDFGMTRDIYETDYYRKGGKGLLPVRWMSPESLKDGVFTTHSD 1196
sp|Q05688|IGF1R_BOVIN  DLAARNCMVAEDFTVKIGDFGMTRDIYETDYYRKGGKGLLPVRWMSPESLKDGVFTTHSD 467
sp|P08069|IGF1R_HUMAN  DLAARNCMVAEDFTVKIGDFGMTRDIYETDYYRKGGKGLLPVRWMSPESLKDGVFTTYSD 1194
sp|Q29000|IGF1R_PIG    ----------------------------------------------------------- 186


sp|O73798|IGF1R_XENLA  VWSFGVVLWEIATLAEQPYQGMSNEQVLRFVMEGGLLEKPDNCPDMLFELMRMCWQFNPK 1250
sp|P24062|IGF1R_RAT    VWSFGVVLWEIATLAEQPYQGLSNEQVLRFVMEGGLLDKPDNCPDMLFELMRMCWQYNPK 1255
sp|Q60751|IGF1R_MOUSE  VWSFGVVLWEIATLAEQPYQGLSNEQVLRFVMEGGLLDKPDNCPDMLFELMRMCWQYNPK 1256
sp|Q05688|IGF1R_BOVIN  VWSFGVVLWEIATLAEQPYQGLSNEQVLRFVMEGGLLDKPDNCPDMLFELMRMCWQYNPK 527
sp|P08069|IGF1R_HUMAN  VWSFGVVLWEIATLAEQPYQGLSNEQVLRFVMEGGLLDKPDNCPDMLFELMRMCWQYNPK 1254
sp|Q29000|IGF1R_PIG    ----------------------------------------MLFELMRMCWQYNPK 201
                                                               ************.***


sp|O73798|IGF1R_XENLA  MRPSFLEIISSIKDELDPGFKEVSFFYSEENKPPDTEELD----LEAENMESIPLDPSCA 1306
sp|P24062|IGF1R_RAT    MRPSFLEIIGSIKDEMEPSFQEVSFYYSEENKPPEPEELEMELELEPENMESVPLDPSAS 1315
sp|Q60751|IGF1R_MOUSE  MRPSFLEIIGSIKDEMEPSFQEVSFYYSEENKPPEPEELEMELEMEPENMESVPLDPSAS 1316
sp|Q05688|IGF1R_BOVIN  MRPSFLEIISSVKDEMEAGFREVSFYYSEENKPPEPEEL----DLEPENMESVPLDPSAS 583
sp|P08069|IGF1R_HUMAN  MRPSFLEIISSIKEEMEPGFREVSFYYSEENKLPEPEEL----DLEPENMESVPLDPSAS 1310
sp|Q29000|IGF1R_PIG    MRPSFLEIISSIKDEMEPGFREVSFYYSEENKPPEPEEL----DLEPENMESVPLDPSAS 257
                       ********.*.*.*:*.*.*.*:****.*:*****.*:.***    .:* *****:***** .:


sp|O73798|IGF1R_XENLA  ---LQNSEHHAGHKSENG--PGVVVLRASFDERQPYAHMNGGRKNERALPLPQSSAC    1358
sp|P24062|IGF1R_RAT    SASLPLPERHSGHKAENG--PGVLVLRASFDERQPYAHMNGGRANERALPLPQSSTC    1370
sp|Q60751|IGF1R_MOUSE  SASLPLPERHSGHKAENGPGPGVLVLRASFDERQPYAHMNGGRANERALPLPQSSTC    1373
sp|Q05688|IGF1R_BOVIN  SASLPLPDRHSGHKAENGPGPGVLVLRASFDERQPYAHMNGGRKNERALPLPQSSTC    640
sp|P08069|IGF1R_HUMAN  SSSLPLPDRHSGHKAENGPGPGVLVLRASFDERQPYAHMNGGRKNERALPLPQSSTC    1367
sp|Q29000|IGF1R_PIG    SSSLPLPDRHSGHKAENGPGPGVLVLRASFDERQPYAHMNGGRKNER---------    304
                          *   :::*:***:***  ***:******************** ***
```

The summary file output of clustalw contains all the necessary that is required to perform clustering. The pair wise scores generated by the program are given as input to the python based single linkage and average linkage clustering algorithm.

Percent Identity Matrix - created by Clustal2.1

1:	sp|O73798|IGF1R_XENLA	100.00	75.02	74.82	77.88	75.67	71.48
2:	sp|P24062|IGF1R_RAT	75.02	100.00	99.71	96.08	96.04	96.36
3:	sp|Q60751|IGF1R_MOUSE	74.82	99.71	100.00	95.94	95.98	96.05
4:	sp|Q05688|IGF1R_BOVIN	77.88	96.08	95.94	100.00	97.66	97.70
5:	sp|P08069|IGF1R_HUMAN	75.67	96.04	95.98	97.66	100.00	98.68
6:	sp|Q29000|IGF1R_PIG	71.48	96.36	96.05	97.70	98.68	100.00

3.1 Single linkage clustering

The scores output from clustalw analysis were given as input to python program to cluster sequences based on single linkage algorithm. The output is given below.

```
Distance Calculation Method: ********** Single
Linkage Clustering ***********
Objects:
IGF1R_XENLA
IGF1R_RAT
IGF1R_MOUSE
IGF1R_BOVIN
IGF1R_HUMAN
IGF1R_PIG
0  75.02  74.82  77.88  75.67  71.48
0  99.71  96.08  96.04  96.36
0  95.94  95.98  96.05
0  97.66  97.70
0  98.68
0


Objects:
('IGF1R_XENLA', 'IGF1R_PIG', 71.48)
IGF1R_RAT
IGF1R_MOUSE
IGF1R_BOVIN
IGF1R_HUMAN
0  75.02  74.82  77.88  75.67
0  99.71  96.08  96.04
0  95.94  95.98
0  97.66
0


Objects:
(('IGF1R_XENLA',     'IGF1R_PIG',       71.48),
'IGF1R_MOUSE', 74.82)
IGF1R_RAT
IGF1R_BOVIN
IGF1R_HUMAN
0  75.02  77.88  75.67
0  96.08  96.04
0  97.66
0


Objects:
((('IGF1R_XENLA',    'IGF1R_PIG',      71.48),
'IGF1R_MOUSE', 74.82), 'IGF1R_RAT', 75.02)
IGF1R_BOVIN
IGF1R_HUMAN
0  77.88  75.67
0  97.66
0


Objects:
(((('IGF1R_XENLA',    'IGF1R_PIG',     71.48),
'IGF1R_MOUSE',  74.82),  'IGF1R_RAT',  75.02),
'IGF1R_HUMAN', 75.67)
IGF1R_BOVIN
0  77.88
0


Objects:
```

```
(((((('IGF1R_XENLA',    'IGF1R_PIG',     71.48),
'IGF1R_MOUSE',  74.82),  'IGF1R_RAT',  75.02),
'IGF1R_HUMAN', 75.67), 'IGF1R_BOVIN', 77.88)
```

3.2 Complete linkage clustering

The outcome of clusters generated by the average linkage clustering method resulted in the following.

```
Distance Calculation Method: ******** Complete
Linkage Clustering ********
MULTIPLE  SEQUENCES:  Insulin  like  Growth
Factor-1 Receptor
IGF1R_XENLA
IGF1R_RAT
IGF1R_MOUSE
IGF1R_BOVIN
IGF1R_HUMAN
IGF1R_PIG
0  75.02  74.82  77.88  75.67  71.48
0  99.71  96.08  96.04  96.36
0  95.94  95.98  96.05
0  97.66  97.70
0  98.68
0


MULTIPLE  SEQUENCES:  Insulin  like  Growth
Factor-1 Receptor

('IGF1R_XENLA', 'IGF1R_RAT', 75.02)
IGF1R_MOUSE
IGF1R_BOVIN
IGF1R_HUMAN
IGF1R_PIG
0  87.26  86.98  85.86  83.92
0  95.94  95.98  96.05
0  97.66  97.70
0  98.68
0


MULTIPLE  SEQUENCES:  Insulin  like  Growth
Factor-1 Receptor
(('IGF1R_XENLA',      'IGF1R_RAT',      75.02),
'IGF1R_MOUSE', 87.26499999999999)
IGF1R_BOVIN
IGF1R_HUMAN
IGF1R_PIG
0  91.46  90.92  89.98
0  97.66  97.70
0  98.68
0


MULTIPLE  SEQUENCES:   Insulin   like   Growth
Factor-1 Receptor
((('IGF1R_XENLA',      'IGF1R_RAT',      75.02),
'IGF1R_MOUSE',          87.26499999999999),
'IGF1R_BOVIN', 91.46)
IGF1R_HUMAN
IGF1R_PIG
0  94.29  93.84
0  98.68
0
MULTIPLE   SEQUENCES:   Insulin   like   Growth
Factor-1 Receptor
```

```
(((('IGF1R_XENLA',     'IGF1R_RAT',      75.02),
'IGF1R_MOUSE',                87.26499999999999),
'IGF1R_BOVIN',      91.46),      'IGF1R_HUMAN',
94.28875)
IGF1R_PIG
0  96.26
0
```

```
MULTIPLE   SEQUENCES:   Insulin   like   Growth
Factor-1 Receptor
((((('IGF1R_XENLA',     'IGF1R_RAT',      75.02),
'IGF1R_MOUSE',                87.26499999999999),
'IGF1R_BOVIN',      91.46),      'IGF1R_HUMAN',
94.28875), 'IGF1R_PIG', 96.26125)
```

Followed by the clusters generated by the single linkage and complete linkage analysis methods, phylogenetic trees were constructed using python code.
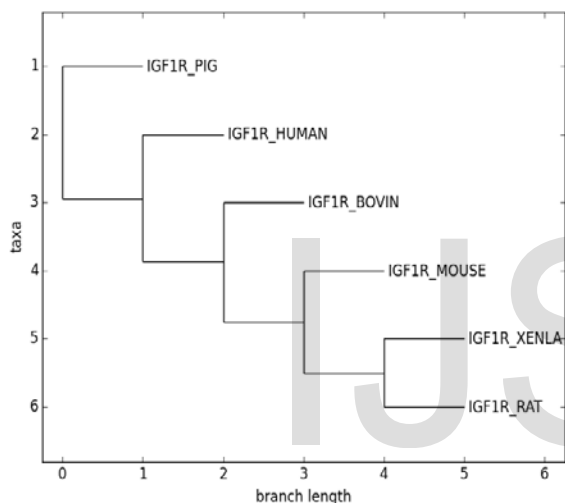
Figure 1: Phylogenetic tree obtained by single linkage method.

From Figure 1, it is evidenced that the single linkage clustering method clusters groups of similar sequences one after the other and the resultant phylogenetic tree can be viewed to understand the concept.
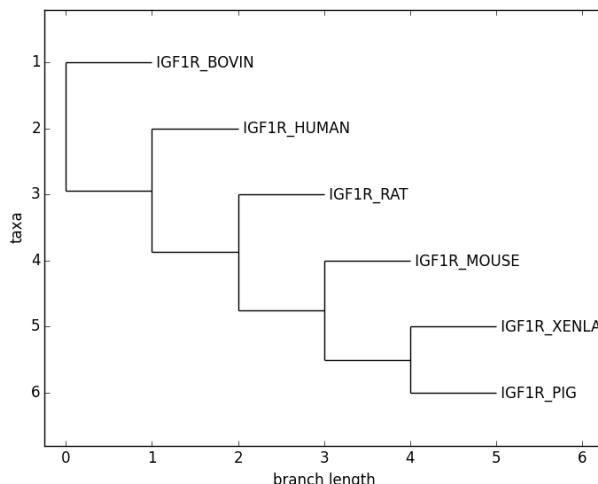
Figure 2: Phylogenetic tree obtained by complete linkage method.

From figures 1 and 2, it was observed that IGF1R_XENLA and PIG sequences formed one group in single linkage whereas IGF1R_XENLA and RAT appeared under one clade in complete linkage algorithm. Remaining sequences appeared as separate clades. Hence, in order to effectively cluster a set of IGF1R sequences, clustering by average linking technique was employed.

In our study, we employed Euclidean distance as distance measure. In this method, we considered the distance between one cluster and another cluster to be equal to the shortest distance from any member of one cluster to any member of the other cluster.

4 CONCLUSION

Clustering techniques were implemented to study the relationship among IGF1R sequences based on the fact that analysis of gene data represents the important step towards identification of groups of genes that manifest similar expression pattern. Single linkage and average linkage clustering algorithms written in python language are used to generate clusters of orthologous insulin sequences, which resulted in a dendrogram.The pair wise scores generated from clustalw are given as input to the python based single linkage and average linkage clustering algorithm and the phylogenetic trees are reported.

5 REFRENNCES

[1] Eisen, M. B., Spellman, P. T., Brown, P. O., & Botstein, D. (1998). Cluster analysis and display of genome-wide expression patterns. *Proceedings of the National Academy of Sciences*, *95*(25), 14863-14868.

[2] Rao, S. G., & Govardhan, A. (2015). Investigation of Validity Metrics for Modified K-Means Clustering Algorithm. *i-Manager's Journal on Computer Science*, *3*(2), 33.

[3] Ben-Dor, A., Shamir, R., & Yakhini, Z. (1999). Clustering gene expression patterns. *Journal of computational biology*, *6*(3-4), 281-297.

[4] Rambabu, R., & Rao, P. S. (2016). A Group Average Cluster Analysis of Few IGF1R Sequences using Modified Group Average Link Clustering Algorithm. *International Journal of Computer Applications*, *150*(11).

[5] Shamir, R., & Sharan, R. (2002). Algorithmic approaches to clustering gene expression data. In *In*.

[6] Jiang, T., Xu, Y., & Zhang, M. Q. (2002). *Current topics in computational molecular biology*. MIT Press.

[7] Rambabu, R., & Rao, P. S. (2016). Linear Regression Analysis and Validation Studies of Insulin-Like Growth Factor (IGF-1) Receptor Inhibitors. *i-Manager's Journal on Information Technology*, *5*(3), 31.

[8] Belfiore, A., Frasca, F., Pandini, G., Sciacca, L., & Vigneri, R. (2009). Insulin receptor isoforms and insulin receptor/insulin-like growth factor receptor hybrids in physiology and disease. *Endocrine reviews*, *30*(6), 586-623.

[9] Chisalita, S. I., Nitert, M. D., & Arnqvist, H. J. (2006). Characterisation of receptors for IGF-I and insulin; evidence for hybrid insulin/IGF-I receptor in human coronary artery endothelial cells. *Growth hormone & IGF research*, *16*(4), 258-266.

[10] Ulanet, D. B., Ludwig, D. L., Kahn, C. R., & Hanahan, D. (2010). Insulin receptor functionally enhances multistage tumor progression and conveys intrinsic resistance to IGF-1R targeted therapy. *Proceedings of the National Academy of Sciences*, *107*(24), 10791-10798.

[11] Federici, M., Porzio, O., Lauro, D., Borboni, P., Giovannone, B., Zucaro, L., ... & Sesti, G. (1998). Increased Abundance of Insulin/Insulin-Like Growth Factor-I Hybrid Receptors in Skeletal Muscle of Obese Subjects Is Correlated with In Vivo Insulin Sensitivity 1. *The Journal of Clinical Endocrinology & Metabolism*, *83*(8), 2911-2915.

[12] Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R. D., & Bairoch, A. (2003). ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic acids research*, *31*(13), 3784-3788.

[13] www.expasy.org

[14] Lipman, D. J., & Pearson, W. R. (1985). Rapid and sensitive protein similarity searches. *Science*, *227*(4693), 1435-1441.

[15] Muniyappa, R., & Sowers, J. R. (2012). Endothelial insulin and IGF-1 receptors: when yes means NO. *Diabetes*, *61*(9), 2225-2227.

6 AUTHOR'S PROFILE

Mr. R Rambabu is working as Associate Professor and Head of the Department, Dept. of CSE, Rajamahendri Institute of Engineering & Technology, Rajahmundry. He is pursuing Ph.D. (CSE) from JNTUK. He received his M.Tech (IT) from Andhra University. His areas of interest are Bioinformatics, Data Mining, Computer Networks, Software Engineering, and Mobile Computing technologies.

Dr. Peri Sriniasa Rao is a Professor at the Department of Computer Science & Systems Engineering, Andhra University Visakhapatnam. His research interest is in the areas of Image Processing, Queuing Applications, Bioinformatics and Computer Networks. At AU additionally, he is also a Chairman Board of Studies and Expert Member in CSE for AICTE and NBA. He guided 15 students for Ph.D. Degree. He is a Life Time member of CSI.