Journal of Applied Pharmaceutical Science Vol. 2 (11), pp. 084-091, November, 2012
Available online at http://www.japsonline.com
DOI: 10.7324/JAPS.2012.21115
ISSN 2231-3354

In silico approach to evaluate evolutionary relatedness of significant Lung cancer proteins

Bhagavathi S, Gulshan Wadhwa, Anil Prakash
Department of Biotechnology & Bio Informatics Centre, Barkatullah University, Bhopal, India - 462026.

ARTICLE INFO

Article history:
Received on: 09/10/2012
Revised on: 18/10/2012
Accepted on: 23/10/2012
Available online: 29/11/2012

Key words:
Lung cancer,
EPAS1,
Thrombomodulin,
Metallothionein and Matrix MetalloProteinase ,
Phylogenetics.

ABSTRACT

Lung cancer is feared to be deadly and is caused due to the involvement of a cascade of proteins significant among which are EPAS1, Thrombomodulin, Metallothionein and Matrix Metallo Proteinase. The evolutionary aspects of these proteins are important to study the extent of conservation and mutation across different generations in diverse species. The rooted and unrooted tree would depict the relatedness among organisms and the ancestor of the organisms involved in the study. The evolutionary histories of groups of species are one of the most widely used tools throughout the life sciences, as well as objects of research with in systematic, evolutionary biology. In every phylogenetic analysis reconstruction produces trees. These trees represent the evolutionary histories of many groups of organisms, bacteria due to horizontal gene transfer and plants due to process of hybridization. Through comparison with several species of healthy organism, one may determine where the defective mutation is located, and then determine how to treat the disease. This strategy may help us to identify the mutations that had occurred in evolutionary conserved residues. We feel this method can be useful for understanding evolutionary rate variation, and for understanding selection variation on different characters. The various tree topologies along with their significant homolog's would enable a thorough understanding of the disease.

INTRODUCTION

Endothelial PAS domain-containing protein 1 is a protein expressed in humans and encoded by the EPAS1 gene (McKnight *et al.,* 1997). This gene encodes the transcription factor HIF2α, which stimulates production of red blood cells and thus increases the concentration of haemoglobin in blood (Beall *et al.,* 2010). The EPAS1 provides novel insight into the regulatory mechanisms of EPAS1 gene expression that contributed to the adaptation of tumour cells under the hypoxic environment. Matrix metalloproteinase-12 (MMP-12) is an enzyme expressed in humans and is encoded by the *MMP12* gene. (Shapiro *et al.,* 1993). The matrix metalloproteinase's (MMPs) consist of 24 known human zinc proteases with essential roles in breaking down components of the extracellular matrix (ECM).

* *Corresponding Author*
Gulshan Wadhwa
Joint Director, Department of Biotechnology, Ministry of Science &Technology, Govt. of India, New Delhi India-110003

MMPs play critical roles in lung organogenesis, but their expression, for the most part, is down regulated after generation of the alveoli. Metallothionein (MT) genes encode low molecules weight cysteine rich protein that binds heavy metals. (Samson and Gedamu, 1998). MT, which is a metal binding protein with high cysteine content and high affinity for heavy metals ions, including Cd, Cu, & Zn has been detected in a number of human tissues and tumours. A high concentration of MT occurs during early development in a variety of tissues, including liver and lung tissues with levels declining to very low concentrations in adult life (Waalkes and Goering, 1990; Hart *et al.,* 1991). MT over expression in lung tumours has not been evaluated adequately and there is only one published study in the literature (Hart and Vacek, 1993). In that study Hart.et.al compared the concentration of Cu, Cd & MT in paired samples of neoplastic and nonplastic lung tissue obtained from 10 patients with primary lung carcinoma, including 2 SCLC samples and 10 non small cell carcinoma samples.

Thrombomodulin which is a receptor for thrombin on the surface of the vascular endothelial cells neutralizes thrombin and the formed thrombin- TM complex activates protein C. TM is not only a thrombin receptor but also an onco developmental antigen, which is found in lung cancers. TM expression in the lung cancer cells appears to vary depending on the cellular conditions. It can be roughly speculated that functionally active TM on lung cancer cells may modulate the biological behaviours of these cells, such as invasiveness and metastatic potential.

The many core phylogenies we present in this article capably provide several orders of magnetic speed up in computing phylogenetic likely hoods. The basic assumption of all phylogenetic analyses is that orthologous genes are being compared. To carry out evolutionary analyses of multigene families, one requires to distinguish orthologs, which have evolved by vertical descent from a common ancestor, from paralogs, which arise by duplication and domain shuffling within a genome (Fitch 1970)**.** Failure to do so can result in functional misclassification and inaccurate molecular evolutionary reconstructions (Doolittle, 1996; Feng, 1997)**.**

The overall similarity (as determined by the BLAST E-value) is often used as a criterion to determine orthology relationships within large data sets such as complete genomes (Rubin *et al.,* 2000; Tatusov *et al.,* 1997; Chervitz et al., 1998; Ruepp *et al.,* 2000) but there is evidence that more rigorous phylogenetic reconstructions are required to confidently determine orthologies (Koski and Golding 2001).

MATERIALS & METHODS

DNA sequence data and Homology search

Protein sequences of Metallothionein, Thrombomodulin, EPAS1, Matrix Metallo Proteinase belonging to different genera were retrieved from NCBI (available at www.ncbi.nlm.nih.gov) and saved in fasta file format for further analysis. It was submitted to NCBI-BLAST server to retrieve closest related homologs.

Multiple sequence alignment:

The homologs with significant score were processed by converting the GenPept format to the FASTA format and compiled together in a single text file. The ClustalW (Thompson *et al.,* 1997) algorithm was used to perform an initial multiple sequence alignment. This alignment was again manually edited and realigned using ClustalW with default parameters for Gap Opening, Gap Extension Penalty and DNA weight matrix to obtain optimal global sequence alignment. Phylogenetic trees were then built using this multiple sequence alignment file.

Phylogenetic tree using parsimony

Phylogenetic tree using parsimony was calculated using PHYLIP's dnapars/protpars algorithm. The tool used for finding the tree is a standalone tool known as Sea View 4.0 (Galtier *et al.,* 1996)**.**

Proml

This module of PHYLIP makes use of the Maximum likelihood method for the construction of phylogenetic tree. This module was chosen over the Maximum Parsimony and Distance methods as it is more efficient and reliable, despite taking greater time to process the data. The method uses probability calculations to generate a tree that best accounts for variation in a set of sequences.

Consense

Consense module of PHYLIP reads a file of computer-readable trees and prints out a consensus tree. The tree printed out has at each fork a number indicating how many times the group which consists of the species to the right of (descended from) the fork occurred. Tree View was used to view the final outtree obtained from Consense.

RESULT AND DISCUSSION

Homology search of epas1 using blast

The sequence of EPAS1 was subjected to similarity search using BLAST tool in order to retrieve closely related species which is tabulated in Table 1.

PHYLOGENETIC TREE CONSTRUCTION OF EPAS1

The closest related homologs based on high percentage of similarity with respect to EPAS1 domain protein gene associated with lung cancer were subjected to phylogenetic analysis using PHYLIP Both unrooted and rooted tress were constructed and shown in both Fig.1 and Fig.2. The result produced by the phylip software (Fig.1. and 2.) reveals different distinct groups based on relative branch lengths. The comparison of nucleotide sequences strictly separated the rodents (mouse) from the rabbit which might be due to the deletion mutations at genetic level **(**Pagani *et al.,* 2005). But that didn't necessarily mean that the nucleotide function is different in this three species. So the observed divergence between rodents and non-rodents probably reflected the differences in pattern of the expression between the two groups. The determination of the rodents branching point enables us to root the tree with primate group which include *Homosapiens* (Vuillaumier et al., 1995).

If we see through the un-rooted tree as shown in the figure 1 we can conclude that *Bos taurus* and Sheep are overlapped with each other with respect to rooted dendrogram. Therefore the parsimony tree obtained from extensive search of PHYLIP with *Xenopus tropicalis* was deeply separated from the rabbits and rodents. The topology was the same as the Neighbour joining tree obtained with or without Distance correction. Therefore Long Branch attraction of *Xenopus* & rodents were rather unlikely (Felsenstein, 1973). Next we assume to determine *Bos taurus* and sheep branch with each other which enables us to root the tree with primate group. This assumption was validated with bootstrap parsimony analysis.

In the same way Dog (*Canis familiaris*) is being rooted from the bear and another group namely frog and toad were also rooted from the rodents. Therefore, it can be concluded that primates, rodents were more closely related to mammals which includes toads, frogs and bear. This was due to mainly the divergence of sequences in mammalian species. Thus it can be postulated that remaining homologies were under a strong pressure and therefore had critical role in severity of disease development.

Table .1 : Homologs of EPAS1 using BLAST tool.

Index	Organism	Scientific Name	Reference no	Generic ID
1	Human	*Homo sapiens*	262527236	NM_001430.4
2	Bear (Giant Panda)	*Ailuropoda melanoleuca*	301753235	XM_ 002912437.1
3	Dog	*Canis familiaris*	73970125	XM_531807.2
4	Rodents Mouse	*Mus musculus*	226061947	NM_010431.2
5	Chicken	*Gallus gallus*	46048878	NM_204807.1
6	Breed white leg horn	*Gallus gallus*	5051909	AF_129813.1
7	Frog	*Xenopus Tropicalis*	49250452	BC_ 074648.1
8	Toad	*Xenopus Silurana*	54020802	NM_001005647.1
9	Cow	*Bos taurus EPAS1*	31340626	NM_ 74725.2
10	Sheep	*Ovis aries metallothionein*	164600782	EU 340261.1
11	Pig	*Sus scrofa EPAS1*	148228021	NM_001097420.1
12	Rabbit	*Oryctolagus cuniculus*	291386846	NM_002709731.1

Table. 2 : Homologs of MMP12 using BLAST tool.

Index	Organism	Scientific Name	Reference no	Generic ID
1	Human	*Homo sapiens*	261878521	NM_0024264
2	Rodents-Mouse	*Mus musculus*	115392137	NM_0086053
3	Cow	*Bos taurus*	331284140	NM_0012066401
4	Rat	*Rattus norvegicus*	291327536	NM_05369632
5	Rabbit	*Oryctolagus cuniculus*	130487231	NM_001082771
6	Pig	*Sus scrofa*	153791307	NM_ 0010999381
7	Gorilla	*Troglodytes*	94961199	DQ4822301
8	Cat	*Mustela putorius*	355702773	JP0134461
9	Dog	*Canis lupus familiaris*	345799780	XM_8495012
10	Chimpanzee	*Pantroglodytes*	332837594	XM_5087243
11	Orangutan	*Pongo abelii*	94961200	DQ4822311

Table. 3 : Homologs of Metallothionein using BLAST tool.

Index	Organism	Scientific Name	Reference no	Generic ID
1	Mouse	*Mus musculus*	74149618	AK_165111
2	Rhesus Macaque	*Macaca mulata*	307938348	NM_ 0011958301
3	Human	*Homosapiens*	71274112	NM _0059462
4	Chimpanzee	*Pantroglodytes*	114664341	NW_001226160
5	Sumatram Orugunatum	*Pongo abelii*	241026459	NM_0029588781
6	Rat	*Rattus norvegicus*	8394161	NM_0173211
7	Cow	*Bos taurus*	164519003	NM_0011130041
8	Pig	*Sus scrofa*	164544	MZ_95151
9	Sheep	*Ovis aries*	1342	X07951
10	Bony fishes	*Pelteobagrus fulvidracus*	159137606	EU_1246611
11	Salmoniformes	*Salmo salar*	213513466	NM_0011413381
12	Cypriniformes	*Danio rerio*	238814362	NM_0011614701
13	Chicken	*Gallus gallus*	147901435	NM_0010975381
14	Birds	*Coturnix*	57903623	AY8664091
15	Lizards	*Zootoca vivipara*	78057034	AM0873991
16	Snakes	*Elaphaquatuor*	78057022	AM 0873931
17	Frog	*Xenopus levis*	602417	XL 014649

Table. 4 : Homologs of Thrombomodulin using BLAST tool.

Index	Organism	Scientific Name	Reference no	Generic ID
1	Human	*Homo sapiens*	194097423	NM_0026192
2	Mouse	*Mus musculus*	26100940	AK_0828851
3	Rat	*Rattus norvegicus*	2511547	AF_0227421
4	Cow	*Bos taurus*	163762	M146571
5	Pig	*Sus scrofa*	157399570	EF6926421
6	Dog	*Canis lupus familiaris*	55741804	NM_0010069531
7	Rabbit	*Oryctolagus cuniculus*	126723033	NM_0010821441
8	Rhesus monkey	*Macaca mulatta*	302565499	NM_0011944851
9	Frog	*Xenopus laevis*	148223392	NM_0010951151
10	Chicken (birds)	*Gallus gallus*	313661359	NM_0011994511

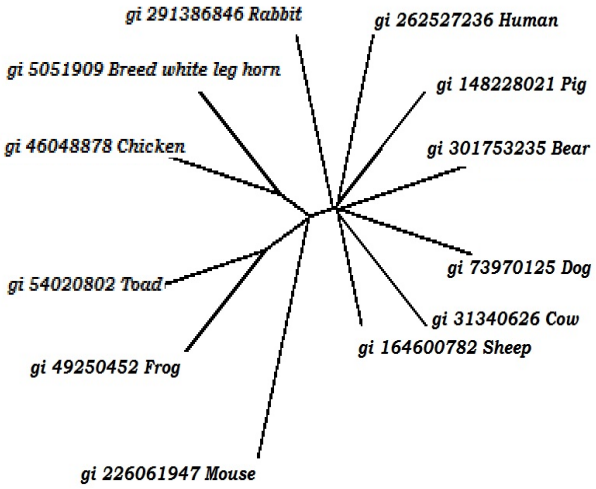

Fig. 1: Unrooted Tree Of Epas1.

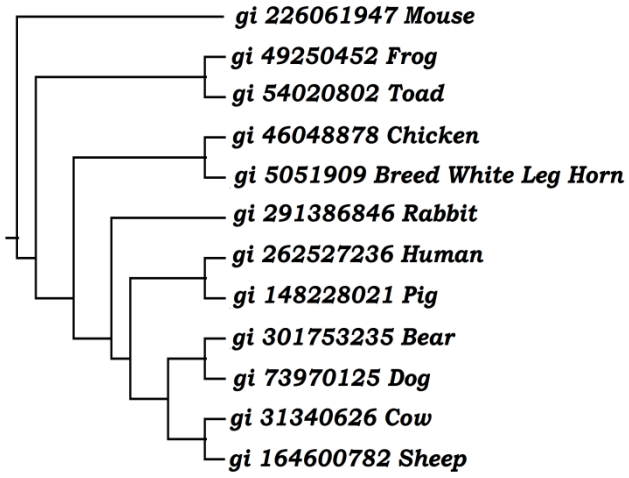

Fig. 2: Rooted Tree of Epas1.

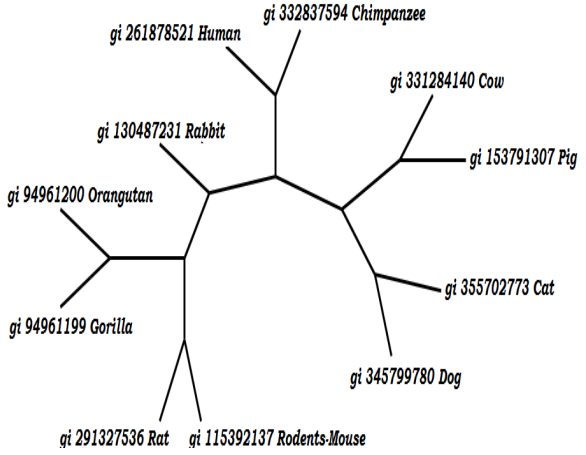

Fig. 3: Unrooted Tree of Mmp12.

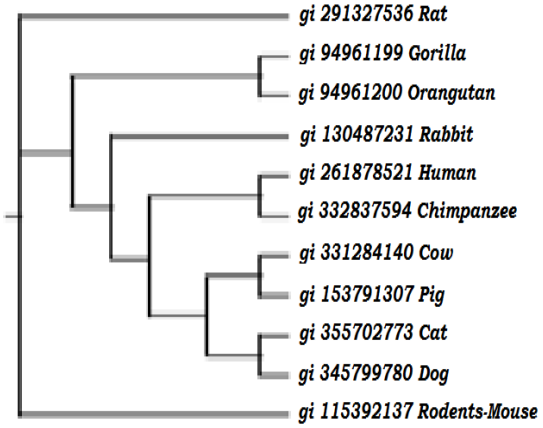

Fig. 4: Rooted Tree of Mmp12.

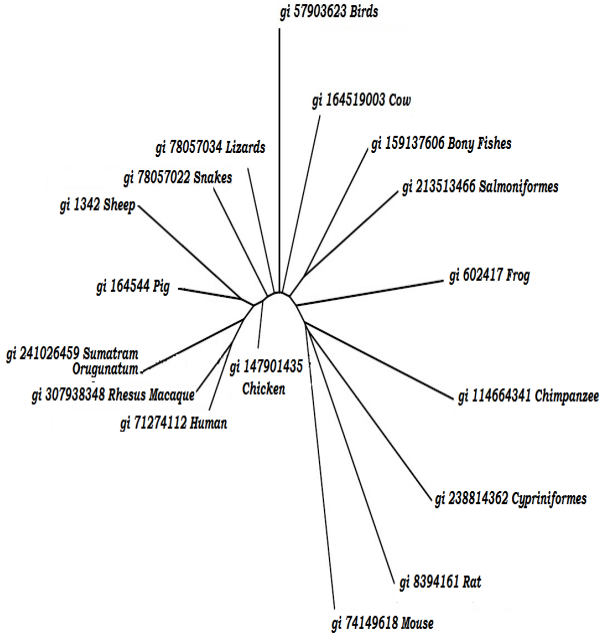

Fig. 5: Unrooted Tree of Metallothionein.

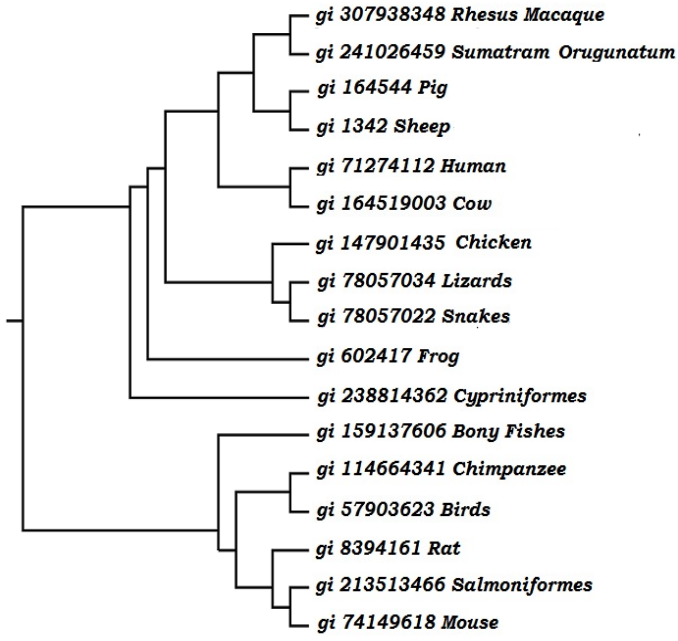

Fig.6: Rooted Tree of Metallothionein.

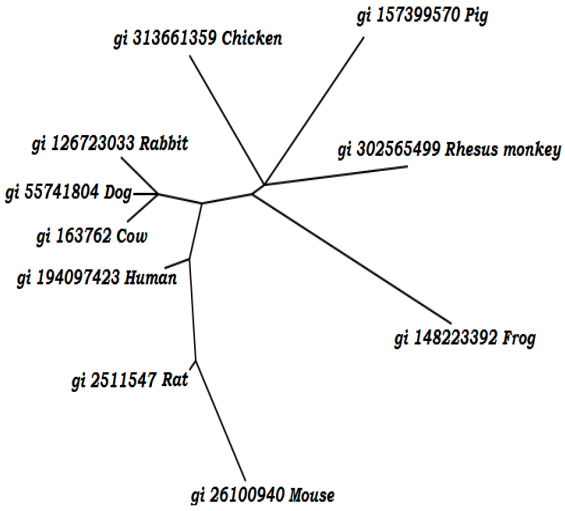

Fig. 7 : Unrooted Tree of Thrombomodulin**.**

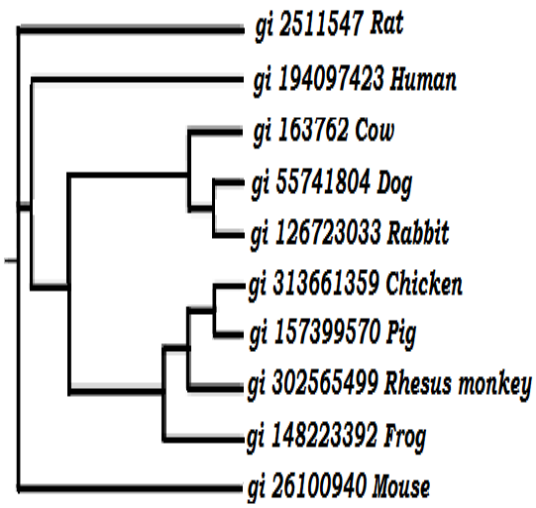

Fig.8. Rooted Tree of Thrombomodulin.

Phylogenetic Analyses Result of Matrix Metalloproteinase 12

The top 11 homologs of MMP12 with high sequence similarity were chosen for further steps.

Phylogenetic Tree Construction

With the PHYLIP package, we had analyzed 11 sequences of highest similarity and inferred their phylogenetic relationship with respect to matrix metalloproteinase gene associated with lung cancer. With the protpars program unrooted tree was obtained. Fig.3 and rooted tree dendrogram tree was also obtained as shown in Fig.4.

The result produced by the phylip software reveals different distinct groups based on relative branch lengths. The comparison of nucleotides sequences strictly separated the rat, mouse which might be due to the deletion mutation at genetic level. The top organisms in the gene data base reveals five distinct groups based on their relative branch lengths. One group included the Mouse (*Mus musculus*), Second group include the Human (*Homo sapiens*), Third group includes the Rat (*Rattus norvegicus*), the fourth and fifth contained the even toed ungulates i.e., Cow (*Bos taurus*) and Pig (*Sus scrofa*) and other taxa which include primates. The comparison of amino acids in MMP12 sequences strictly separated the Rat from the mouse but that doesn't necessarily mean that the protein function is different from these two species **it** might be due to the deletion mutations at the genetic level.

Next we assume to determine the unexpected Rabbit – Rodent divergence which is very similar to Human-Chimpanzee having a typical pseudo stratified surface epithelium (Plopper *et al.,* 1983) in which ciliated epithelial cells were the abundant cell types (Jeffrey, 1983) From the phylogenetic point of view we can conclude that Rabbit appears to be closer to primates than rodents(Graur et al., 1996)**.**The comparison of Human and chimpanzee in the unrooted tree fig. 3 could be based upon the previous studies of Human chimpanzee transcription comparison which showed 39% of silent sites in Protein coding regions were under purifying selection (Hellmann et al., 2003) and that of humans the average substitution rates of silent sites was 30% lower in functional genes (Bustamante et al., 2002**)** than in pseudo genes, Pseudogenes are dysfunctional relatives of genes that have lost their protein-coding ability or are otherwise no longer expressed in the cell (Vanin, 1985). Also, the experimental evidence showed that the percentage of synonymous substitutions involved was comparable to that suggested to be under purifying selection in comparative inter and intra species studies (Burrati et al., 2001). Thus it can be postulated that the remaining homologies were under a strong pressure and therefore had a critical role in the severity of disease development.

Phylogenetic Analysis Result of Metallothionein

The top 17 sequences with high sequence similarity chosen for further steps are represented in Table 3.

Phylogenetic Tree Construction

With the PHYLIP package, we had analyzed 17 sequences of highest similarity and inferred their phylogenetic relationship with respect to metallothionein gene associated with lung cancer. With the protpars program unrooted tree was obtained. Fig.5 and rooted tree dendrogram tree was also obtained as shown in Fig.6

The results produced by the phylip software reveal different distinct groups based on relative branch lengths. The comparison of nucleotides sequences strictly separated the rat, mouse and *Salmoniformes* which might be due to the deletion mutation at genetic level .But that didn't necessarily mean that the nucleotide function is different in this three species. A functional region might likely have a lower evolutionary constraint as compared to other region which could be argued on the basis, that most of the mutations are neutral. On the other hand two less divergent taxonomic group might have a fewer substitutions but

these substitutions might occur in functionally important region. So the observed divergence between rodents and non-rodents probably reflected the differences in pattern of the expression between the two groups.

The determination of the rodents branching point enables us to root the tree with non eutherian group which includes rhesus (*Macaca mulata*) **.**If we see through the un-rooted dendrogram we can conclude that *Xenopus levis* and *Elapha quatuor* are overlapped with each other with respect to rooted dendrogram. Similarly *Pongo abelii* and Rhesus macaque are also overlapped with each other due to substitution saturation.

So far from the phylogenetic view we can conclude that non primates appear to be closer to rodents than primates. Next we assume to determine if the two outgroup *Xenopus levis* and *salmosalar* metallothionein as in figure.6 were substitution saturated with respect to mammals **(** Novacek, 1992).

This assumption was validated by the result of saturation tests where pair wise comparisons of substitutions as inferred in Maximum parsimony tree were plotted against pair wise raw differences (Felsenstein, 1998). Therefore the parsimony tree obtained from extensive search of PHYLIP *with Xenopus levis* outgroup was deeply separated from all other members.

The topology was the same as the Neighbour joining tree obtained with or without Distance correction. Therefore Long Branch attraction of *Xenopus* and other members were rather unlikely. The comparison of human and chimpanzee nucleotide based upon the previous studies of human chimpanzee transcript comparison which showed 50% of silent sites in nucleotide coding regions which were under purifying selection and that in humans, the average substitution rates of silent sites were 50% lower in functional genes than in differentially expressed genes. On the other hand we can compare the lizards and snakes which forms monophyletic group.

This assumption was validated with bootstrap parsimony analysis. On the Other hand we can conclude that pig metallothionein seems closer to sheep metallothinien than rodents. So far from the Phylogenetic point of view we can conclude that rodents appeared to be closer to primates. Next we assume to determine *Bos taurus* and *Homo sapiens* branch with each other which enable us to root the tree with primate group. In the same way *Gallus gallus* metallothinein is being un rooted from lizards and snakes.

Another group namely *Cypriniformes* and *Xenopus levis* were substitution saturated with respect to primates. So, therefore over all it can be concluded that primates, rodents were more closely related to mammals which includes lizards, frogs and snakes. This was due to mainly the divergence of sequences in mammalian species. Thus it can be postulated that remaining homologies were under a strong pressure and therefore had critical role in severity of disease development.

Phylogenetic Results of Thrombomodulin

The top 10 sequences with high sequence similarity were chosen for further steps.

Phylogenetic Tree Construction

With the PHYLIP package, we had analyzed 10 sequences of highest similarity and inferred their phylogenetic relationship with respect to thrombomodulin gene associated with lung cancer. With the protpars program unrooted tree was obtained. Fig.7 and rooted tree dendrogram tree was also obtained as shown in Fig.8

The results produced by the phylip software reveals different distinct groups based on relative branch lengths as in figure 8. The comparison of nucleotides sequences strictly separated the rat and mouse which might be due to the deletion mutation at genetic level **.**But that didn't necessarily mean that the nucleotide function is different in this three species. A functional region might likely have a lower evolutionary constraint as compared to other region which could be argued on the basis, that most of the mutations are neutral. On the other hand two less divergent taxonomic group might have a fewer substitutions but these substitutions might occur in functionally important region. So the observed divergence between rodents and non-rodents probably reflected the differences in pattern of the expression between the two groups. The determination of the rodents branching point enables us to root the tree with primate group which include *Homosapiens***.**

The top organisms in the gene data base reveals five distinct groups based on their relative branch lengths. One group included the Mouse (*Mus musculus*), Second group include the Human (*Homo sapiens*), Third group includes the Rat (*Rattus norvegicus*), the fourth and fifth contained the even toed ungulates i.e., Cow (*Bos taurus*) and Pig (*Sus scrofa*) and other taxa which include primates and carnivores. If we see through the un-rooted tree without branch length as shown in the figure 7., we can conclude that Rat and mouse are overlapped with each other when compared to rooted dendrogram and rat and mouse are showing a slight distance which is due to deletion mutation at genetic level.

So far from the phylogenetic point of view we can conclude that rodents are closer to primates. Next we assume to determine the two outgroups carnivores and *Bovidae* as in figure.8, were substitution saturated with respect to primates. This assumption was validated by the result of saturation tests where pair wise comparisons of substitutions as inferred in Maximum parsimony tree were plotted against pair wise raw differences. Therefore the parsimony tree obtained from extensive search of PHYLIP as in fig.7 with *Bos taurus* was deeply separated from the rabbits and rodents. The topology was the same as the Neighbour joining tree obtained with or without Distance correction (Gulshan Wadhwa *et al.,* 2008) Therefore Long Branch attraction of *Xenopus* & rodents were rather unlikely. On the other hand we can compare also the Chicken with Pig which are slightly closer to each other. In the same way we can conclude that Pig (*Sus scrofa*) seems closer to monkey (primates) than rodents. Next we assume to determine *Canis lupus* and *Lagomorph* branch with each other which enable us to root the tree with primate group. This assumption was validated with bootstrap parsimony analysis. So, therefore over all it can be concluded that carnivores, lagomorphs

were more closely related to primates which include humans, monkey. This was due to mainly the divergence of sequences rodent species. Thus it can be postulated that remaining homologies were under a strong pressure and therefore had critical role in severity of disease development.

Moreover, as thrombomodulin is associated with tumorigenesis and belongs to a family of disease-relevant protein that can be targeted by different drugs (Wadhwa *et al.,* 2009) it represents a promising approach for the development of novel anticancer therapies. Thrombomodulin is not only a thrombin receptor but also an onco developmental antigen, which is found in lung cancers. Recent large-scale studies of individuals within a population have demonstrated that there is widespread variation in copy number in many gene families. In addition, there is increasing evidence that the variation in gene copy number can give rise to substantial phenotypic effects. In some cases, these variations have been shown to be adaptive. These observations show that a full understanding of the evolution of biological function requires an understanding of gene gain and gene loss. Accurate, robust evolutionary models of gain and loss events are, therefore, required **(**Ryan *et al.,* 2012).

CONCLUSION

The above work is an *in-silico* work; which will provide an insight into the evolutionary relatedness among the key proteins involved in Lung cancer. We strongly believe that this work can serve as an predicted model and can be useful to understand their evolutionary aspects and their function in other organisms in Lung Cancer. The *in-silico* approach helps researchers by giving them an bird's view of the task so that they can advance towards wet lab procedures that will enable effective treatment of the disease.

ACKNOWLEDGEMENT

One of the Author (S.Bhagavathi) is thankful to the Corresponding Author Dr. Gulshan Wadhwa for guiding and correcting various documents with attention & care and also thanks to Dr. Anil Prakash, Barkatullah University, Bhopal, for extending his support.

REFERENCES

Beall C.M, Cavalleri G.L, Deng L , Elston R.C, Gao Y, Knight J, Li C, Li J.C. Natural selection on EPAS1 (HIF2alpha) associated with low haemoglobin concentration in Tibetan highlanders." Proc Natl Acad Sci U S A 2010; 107(25):11459-64.

Bustamante C.D, Neilson R, Harlt D.L. A maximum likelihood method for analyzing pseudogene evolution: implications for silent site evolution in humans and rodents.Mol.Biol.Evol.2002; 19,110-117.

Buratti E, Dork T, Zuccato E, Pagani F., Romano M & Baranello FE. Nuclear factor TDP-43 and SR proteins promote in vitro and in vivo CFTR exon 9 skipping. EMBO J.2001; 20(7): 1774-1784.

Chervitz S.A, Aravind L, Sherlock G, Ball C.A, Koonin E.V, Dwight S.S, Harris M.A, Dolinski K, Mohr S, Smith T. Comparison of the complete protein sets of worm and yeast: orthology and divergence. Science 1998; 282:2022-2028.

Doolittle R.F, Feng D.F, Tsang S, Cho G, Little. Determining divergence times of the major kingdoms of living organisms with a protein clock. Science 1996; 271:470-477.

Fitch WM. Distinguishing homologous from analogous proteins. Syst Zool 1970; 19:99-113.

Feng D.F, Cho G, Doolittle R.F. Determining divergence times with a protein clock: update and re evaluation. Proc Natl Acad Sci USA 1997; 94:13028-13033.

Franco Pagani, Lewandowska M.A, Stuani C, Parvizpur A, Baralle F.E. Genome.Res.2005; 13,831-837.

Felsenstein J. Maximum likelihood estimation of evolutionary trees from continuous characters American journal of Human Genetics 1973; 25:471-492.

Felsenstein J. Cases in which parsimony or compatibility methods will be positively misleading. Syst. Zool. 1978; 27,401-410,(1978).

Galtier N, Gouy M, Gautier C. SEAVIEW and PHYLO_WIN: two graphic tools for sequence alignment and molecular phylogeny. Comput Appl Biosci 1996; 12: 543-548.

Graur D.L, Duret , Gouy M. Phylogenetic position of the order Lagomorpha (rabbits. Hares and allies). Nature 1996; 739:333-335.Gulshan Wadhwa and Debgiti Ghosh, Cross Species Comparison of Cystic Fibrosis Transmembrane Conductance Regulator Protein (CFTR), Anvesa 2008; 3(2), 1-13.

Hart B.A, Voss G.W, Garvey J.S. Metallothionein in rat lung during post natal development. Biol Neonate 1991; 59:236-43.

Hart B.A, Voss G.W ,Vacek P.M. Metallothionein in human lung carcinoma. Cancer Lett 1993; 75:121-8.

Hellmann I, Zollner S , Enard W, Ebersberger I, Nickel. B, Paabo S. Selection on human genes as revealed by comparisons to chimpanzee cDNA. Genome . Res 2003; 13, 831-837.

Jeffrey P.K Morphological features of airway surfaces epithelial cells and glands. Am. Rev. Respir. Dis. 1983 ; 128, S14-S20.

Jayakanthan M, Gulshan Wadhwa , Madhan Mohan T, Arul L, Balasubramanian P and Sundar D, Computer-Aided Drug Design for Cancer-Causing H-Ras p21 Mutant Protein ,Letters in Drug Design and Discovery; 2009, 6, 14-20.

Koski L.B, Golding G.B. The closest BLAST hit is often not the nearest neighbor. J Mol Evol. 2001, 52:540-542.

Novacek M.J. Mammalian phylogeny: shaking the tree. Nature. 1992; 356, 121-125.

Plopper C.G, Mariassy A.T, Wilson D.N, Alley J.L, Nishio S.J. and Nerresheim P. Comparison of non-ciliated tracheal epithelial cells in six mammalian species: Ultra structure and population densities. Exp.lung. Res. 1983; 5, 281-294.

Rubin G.M, Yandell M.D, Wortman J.R, Gabor Miklos G.L, Nelson C.R, Hariharan I.K, Fortini M.E, Li P.W, Apweiler R, Fleischmann W. Comparative genomics of the eukaryotes. Science. 2000; 287:2204-2215.

Ruepp A, Graml W, Santos-Martinez M.L, Koretke K.K, Volker C, Mewes H.W, Frishman D, Stocker S, Lupas A.N, Baumeister W: The genome sequence of the thermo acidophilic scavenger Thermoplasma acidophilum. Nature. 2000; 407:508-513.

Ryan M. Ames, Daniel Money, Vikramsinh P. Ghatge , Simon Whelan and Simon C. Lovell. Determining the evolutionary history of gene families Bioinformatics 2012; 28(1): 48-55

Shapiro S.D, Kobayashi D.K, Ley T.J. Cloning and characterization of a unique elastolytic metalloproteinase produced by human alveolar macrophages. J. Biol. Chem. 1993; 268 (32): 23824–9.

Samson S.L and Gedamu L. Molecular analysis of metallothionein gene regulation. Progr. Nucleic Acid Res. Mol.Biol. 1998; 59:257-288.

Tian H , McKnight SL, Russell DW. "Endothelial PAS domain protein 1 (EPAS1), a transcription factor selectively expressed in endothelial cells". Genes Dev 11, 1997 (1): 72–82.

Tatusov R.L, Koonin E.V, Lipman D.J. A genomic perspective on protein families. Science. 1997, 278:631-637.

Thompson J.D, Gibson, T.J., Plewniak F., Jeanmougin F., Higgins, D.G.: The CLUSTALX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Research. 1997, 24:4877-4882.

Vuillaumier S, Messai H, Pinta V, Kaltenboeck B, Perret C, Denamur E. Phylogenetic footprinting and DNase I hypersensitivity site analysis within the CFTR promoter region reveal a Camp response element. Pediatr. Pulmonol. 1995; 20, 426.

Vanin E.F .Processed pseudogenes: characteristics and evolution. Annu. Rev. Genet. 1985;19: 253–72.

Waalkes M.P, Goering P.L. Metallothionein and other Cd binding proteins: recent developments. Chem Res Toxicol. 1990; 3:281-8.

How to cite this article:

Bhagavathi S, Gulshan Wadhwa, Anil Prakash, *In silico* approach to evaluate evolutionary relatedness of significant Lung cancer proteins. J App Pharm Sci. 2012; 2 (11): 084-091.