

Study of in Silico Phylogenetic Analysis of Vertebrate Protein Neuroglobin

Gedam K.R¹, Patel P.R², Khamankar D.B³

¹Research scholar, Department of zoology, Lokmanya Tilak Mahavidyalaya wani, Yawatmal, Maharashtra, India

^{2,3}Department of zoology, Lokmanya Tilak Mahavidyalaya wani, Yawatmal Maharashtra, India

Abstract

Neuroglobin protein is found in invertebrate and vertebrate bodies, which provide oxygen to the central nervous system of organisms. Ngb also performs a role in the enhancement of program cell death (Apoptosis). If Ngb does not work properly it affects the function of nerves in the body, and it may cause Hypoxia, cerebral stroke, and in severe conditions, nerve cells die. We observed Ngb having some moderate variation in several amino acids, but without affecting protein function. we also concluded that phylogenetic tree analysis that proteins show some similarity in the different species of vertebrates.

Keyword: Neuroglobin (Ngb), protein, vertebrate

Introduction

Neuroglobin is a member of the globin family, it performs functions like other globin proteins. Ngb discovered approximately 10 years ago in vertebrates' nervous system. It is present in different locations of bodies like nerve tissue, the brain, and the retina. Protein is made up of 151 amino acids. the molecular mass of a protein is 17 kDa. Ngb found 20% to 25% similarity with myoglobin and hemoglobin. It shows affinity with oxygen. it prevents hypoxic conditions, and brain damage, interacts with G protein, it also performs a role in the signal transduction pathway. Ngb consists of a structure with 4 introns and 5 exons but never the less shows the classical 3 on 3 globin fold.

Phylogeny of Neuroglobin and frequency of amino acid in *Homo sapience*:

Modern evolutionary theory says, all organisms are different but they evolved from the same ancestor. This relationship is called *phylogeny*, it is represented by a *phylogenetic tree*. In Ngb the most prominent protein is Leucine, and after that Alanine, glutamic acid, arginine & valine. lastly, tyrosine, aspartic acid, histidine, asparagine, tryptophan, phenylalanine, cysteine, and methionine, are found in protein in very low frequency.

Analysis of the phylogenetic tree of Neuroglobin

The phylogenetic evolutionary tree was constructed by the **UPGMA method** (Unweighted Pair Group Method with Arithmetic Mean). The overall branch length of the optimal tree is = 1.97807612 is obtained. The distance of the evolutionary phylogenetic tree is constructed by using the Poisson correction method, in the units of the number of aa substitutions per site. In this analysis, 37 aa are in sequence. So total of

137 positions in the final dataset. Evolutionary analyses were conducted in **MEGA7**. Phylogenetic tree is key of evolutionary studies and It is important to understand species relationship.

Distance Based Method: The most popular method is UPGMA , neighbour joining,(NJ) , and that optimize the additivity of a distance tree (FM and ME)

UPGMA Method:

This method follows a clustering procedure :

1. Suppose all species are in initial stages are in a respective member of a same group.
2. in this method we recalculate the distance of the joint pair by taking the average, and joint two very closest clusters of species.
3. This procedure perform in repetitive manner until all species are connect in a same or single cluster.
4. Strictly speaking , this algorithm is phenetic , which does not aim to reflect evolutionary descent . it assigns randomized molecular clock ,and equal weight on the distance.

WPGMA :

it is a quite similar to the UPGMA method but it assign different value of weight on the distance. As compared to WPGMA method UPGMA is very simple , fast and extensively used in literature. However it behaves poorly at most cases where the above presumptions are not met.

Neighbor joining method :

This method gives fast and accurate results as compare to the UPGMA method . but in this method we makes only one tree not more. Remaining trees avoid it. which might be as good as NJ trees , if not significantly better .

Weighted Neighbor – joining method (weighbor):-

In this method two term is observed 1) additive term (of external branches) and a Positive term (of internal branches) that qualifies the implication of joining the pair . this method gives minimum weight to the longer distance in the distance matrix and the resulting trees are less sensitive to specific biases than NJ and relatively immune to the “long branches attraction /distraction” drawbacks observed with other method .

Fitch -Margoliash (FM) and Minium Evolution (ME) Method :

This method is proposed by two scientist Fitch and Margoliash in the year 1967. This method is used for fitting trees to distance matrices. This method is used to find out least squared fit of all observed pair- wise distance to the expected distance of a tree.

The ME method is also used to find out the trees with the minimum sum of branch length. But instead of using all the pair wise distance as FM.

Character – Based Method :

Distance based methods are more rapid and less computationally intensive than character based method .but the actual characters are discarded ones the distance matrix is derived.on the other hand , character based method makes used of all known evolutionary information.

Maximum Parsimony (MP):

The MP method gives simple explanation of data is preferred because it requires the few evidences .the MP trees is the one with few evolutionary changes for all sequence to derived from a common ancestor . MP works faster than ML(maximum Likellywood) .

Maximum Likellywood (ML):

ML method also used each position in an alignment and evaluates all possible trees . this method calculated each trees . likelihood is determined by evaluating the probability that a certain evolutionary model has generated the observed data .

Rooting trees :

Root is the common ancestor of the species under study . most phylogenetic method do not locate the root of a tree and the unrooted trees only reflect the relationship among species but not the evolutionary path . sometime a more commonly used method is to evaluated the rooting by an out groups.

A dendrogram of Ngb constructed by using **MEGA7** reveals the formation of an outgroup comprising *Equus caballus* (horse), *myotis lucifugus* (little brown bat), *Heterocephalus globe* (sand puppy), and *Xenopus tropicalis* (tropical clawed frog).

Observation and result :

The sequence of *Homo sapiens*, *Pan troglodytes* (common chimpanzee), *Macaca mulatta* (rhesus macaque), *Papio Anubis* (Anubis baboon), *Microcebus murinus* (grey mouse lemur), *Odocoileus virginianus* (whitetail), *Oryctolagus cuniculus* (European rabbit), *Sus scrofa* (wild boar) and *Otolemur garnettii* (Garnett's greater galago), show moderate similarity with each other and a cluster.

The sequences of *Bos grunniens* (domestic yak), *Bos taurus* (cow), and *Pantholops hodgsonii* (Tibetan antelope), show considerable similarity with each other and form clusters.

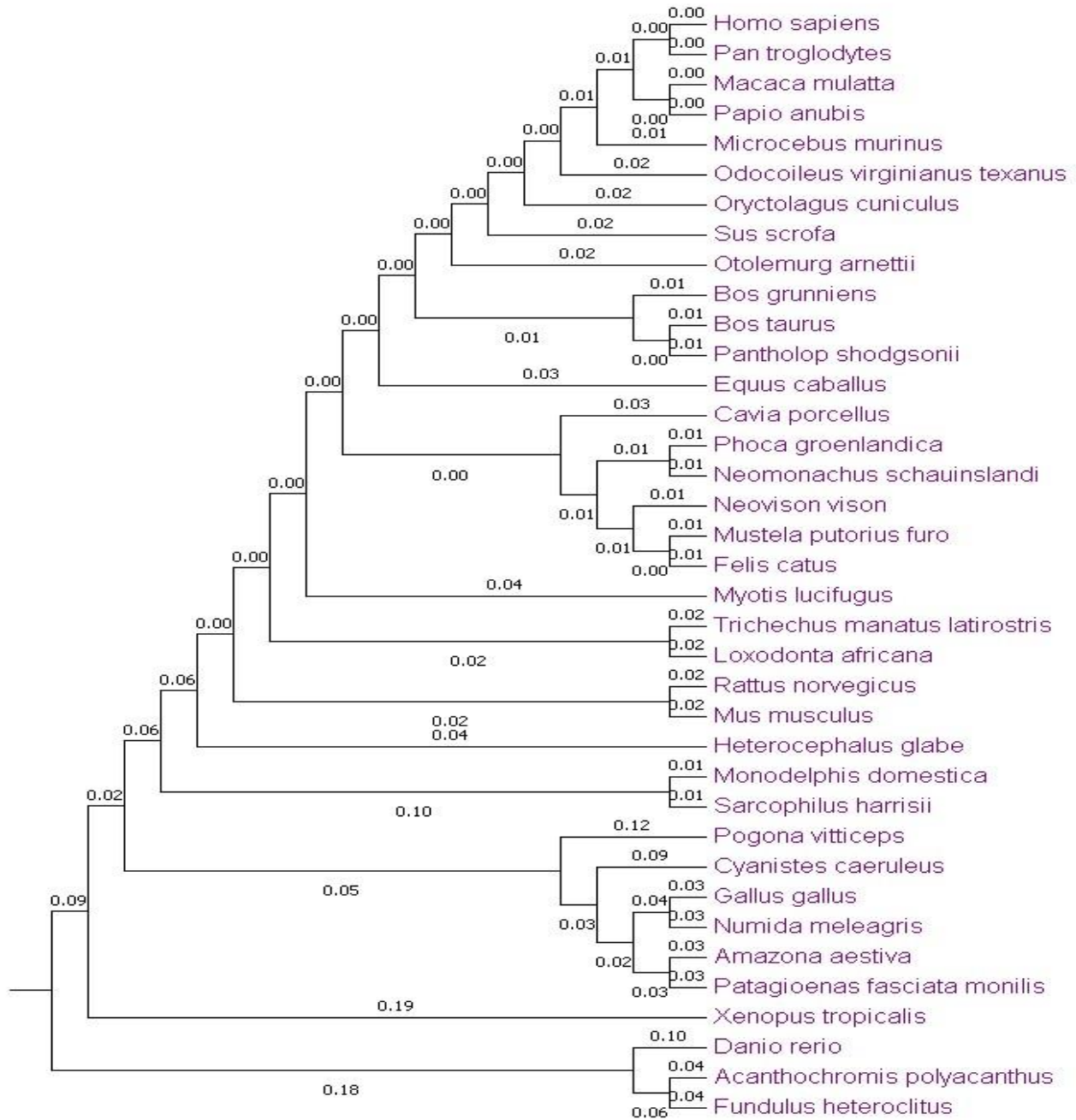
The sequences of *Cavia porcellus* (guinea pig), *Phoca groenlandica* (harp seal), *Neomonachus schauinslandi* (Hawaiian monk seal), *Neovison vison* (American mink), *Mustela putorius furo* (European polecat), and *Felis catus* (Domestic cat) show considerable similarity with each other and form a cluster. The sequence of *Rattus norvegicus* (brown rat), and *Mus musculus* (house mouse) show absolute homology and form a cluster.

The sequences of *Monodelphis domestica* (gray short-tailed opossum) and *Sarcophilus harrissii* (Tasmanian devil) show absolute homology with each other and form a cluster.

The sequences of *Pogona vitticeps* (bearded dragon), *Cyanistes caeruleus* (Eurasian blue tit), *Gallus gallus* (red jungle fowl), *Numida meleagris* (guinea fowl bird), *Amazona aestival* (turquoise-fronted parrot), and *Patagioenas fasciata* (band tail pigeon), show moderate similarity with each other and form a cluster.

The sequence of *Danio rerio* (zebrafish), *Acanthochromis polycanthus* (spiny chromis damselfish), and *Fundulus heteroclitus* (Atlantic killifish) show considerable similarities with each other and form a cluster.

Phylogenetic tree of vertebrate protein Neuroglobin



Discussion and Conclusion:

Predicted Structure Content:

1. Alpha helix : 61.6%
2. Beta Sheet : 3.3%
3. Coil: 35.1%

Folding class coefficient:

1. Alpha : 0.912797
2. Beta : 0.85355
3. Mixed : 0.90427

Post Translational Modification: We observed some post translational modification on protein neuroglobin like Acetylation , C-Mannosylation , Carbamoylation, Deamination , Dimethylation, farnesylation, Geranyl – Geranylation , Glucosylation , Hydroxylation , Methylation , Myristoylation , Palmitoylation, phosphorylation, and Trimethylation.

Protein Analysis:

Neuroglobin amino acid in human – 151
Molecular weight (Daltons) 16933.457
Mean amino acid weight (Daltons) 112.142
Average hydrophobicity -0.0655629
Ratio of hydrophilicity to hydrophobicity 1.0501
Percentage of Hydrophilic amino acid 49.0066
Percentage of hydrophobic amino acid 50.9934
Ratio of % hydrophilic to % hydrophobic is 0.961039
Mean beta hydrophobic moment is 0.199122
Mean helix hydrophobic moment is 0.194925
Basic amino acid : 15
Acidic Amino acid : 19
Estimated pI for protein is 6.2
Total linear charge density is 0.238411
Protein solubility : 1.48033
No.of buried Amino acid in neuroglobin 36

Summary:

Neuroglobin show high affinity with oxygen. number research performed on this protein to protect CNS. and also protective role for some cases of apoptotic process (Program cell death). Due to lack of concentration of Ngb in body it reduce the viability of invertebrate neurons under hypoxia. The final conclusion of this study is the sequence of Amino acid of neuroglobin show slight to moderate differences without affecting the function . it also concluded from the phylogenetic tree analysis that the neuroglobin in different species also show strong to moderate sequence similarities among studied vertebrates.

REFERENCES

1. Bae, Y. A., Cai, G. B., Kim, S. H., Zo, Y. G., & Kong, Y. (2009). Modular evolution of glutathione peroxidase genes in association with different biochemical properties of their encoded proteins in invertebrate animals. *BMC Evolutionary Biology*, 9, 1-13.
2. Boron, I., Russo, R., Boechi, L., Cheng, C. H. C., Di Prisco, G., Estrin, D. A., ... & Nadra, A. D. (2011). Structure and dynamics of Antarctic fish neuroglobin assessed by computer simulations. *IUBMB life*, 63(3), 206-213.
3. Brittain, T. (2012). The anti-apoptotic role of neuroglobin. *Cells*, 1(4), 1133-1155.
4. Burmester, T., & Hankeln, T. (2004). Neuroglobin: a respiratory protein of the nervous system. *Physiology*, 19(3), 110-113.
5. Dröge, J., & Makołowski, W. (2011). Phylogenetic analysis reveals wide distribution of globin X. *Biology Direct*, 6, 1-5.

6. Ganfornina, M. D., Gutie´ rrez, G., Bastiani M, and Sanchez D. (2018) A phylogenetic Analysis of the Lipocalin Protein Family *Department of Biology*, 114-115.
7. Ganfornina, M. D., Guti´ errez, G., & Bastiani, M. (2000). A phylogenetic analysis of the lipocalin protein family. *Molecular biology and evolution*, 17(1), 114-126.
8. <http://www.NCBI.com>
9. <http://www.ncbi.nlm.nih.gov>
10. Hundahl, C. A., Luuk, H., Ilmjärv, S., Falktoft, B., Raida, Z., Vikesaa, J., ... & Hay-Schmidt, A. (2011). Neuroglobin deficiency exacerbates Hif1A and c-FOS response, but does not affect neuronal survival during severe hypoxia in vivo. *PloS one*, 6(12), e28160.
11. Storz, J. F., Opazo, J. C., & Hoffmann, F. G. (2011). Phylogenetic diversification of the globin gene superfamily in chordates. *IUBMB life*, 63(5), 313-322.