Phylogenetic Analysis of Neutral Ceramidase

Nirmal Rajah. R
Independent Scholar
Dubai, United Arab Emirates
Po Box. 22252

Rufus Auxillia
Research scholar
Department of Botany and Microbiology
Lady Doak College, Madurai

ABSTRACT

Neutral ceramidase (NCDase) plays an important role in cell regulation, and has been observed in most of the organism across different groups. In this study the molecular evolution of NCDase in different organisms was studied. The phylogenetic tree relationships among the organisms were explored by comparing the sequences of their NCDase. Mulitle sequence alignment of NCDase protein sequence of different source organisms revealed sequence level homology and highly conserved regions level. A conserved hexapeptide sequence GDVSPN within a large conserved domain was observed frequently in the sequences studies, further study of the sequences indicated the presence of Ser/Thr rich mucin like domains in the sequences of vertebrate NCDase enzymes, but was totally absent in invertebrates and bacteria, confirming the specific domain was acquired separately in mammals during the course of evolution. Comparisons of similarity values and inspection of phylogenetic tree was used to better trace the evolutionary history of Neutral ceramidase and resolve relationships between the organisms.

Keywords
Molecular evolution, Protein Evolution, Cladogram, Phylogenetics

1. INTRODUCTION
Sphingolipids are multifunctional biomodulator in cells. Ceramide is a backbone lipid of sphingomyelin and glycosphingolipids which mediates cell differentiation and apoptosis [15]. Neutral ceramidase, a key enzyme of sphingolipid metabolism [10] that hydrolyzes ceramide to sphingosine. Ceramidases are classified into three categories based on their primary structure and PH Optima. Alikali, acid & neutral. Neutral ceramidase has been classified based on the pH (6.5-8.5) of the reaction it catalyzes [15]. NCDase is present within in the mitochondria and plasma membrane "raft "microdomains. [4]. The NCDase is a glycoprotein with N glycans) [16]

Mucins are characterized by a Tandem Repeat (TR) domain, composed of TR s rich in serine and threonine residues that act as glycosylation sites [6]. Ncdase containing High mannose type N- glycans transported to golgi apparatus from ER region to form a complex with Mucins, where the mucins are highly O glycosylated. It helps the Ncdase to retain its integral membrane structure in the raft region [7].and also mucin expression provides a protective barrier against pathogenic bacterial colonization [6]

In this communication an attempt to resolve the evolutionary relationship between the NCDases in different organisms which was made by comparing the sequences among different classes and presence of highly conserved domain also a CDase active site with possible role in reverse CDase activity and Mucin Box a specific Serine – Threoinine rich domain in the sequences *Insilco.*

2. MATERIALS AND METHODS
The Neutral ceramidase (NCDase) sequences of different source organisms available in NCBI (http://www.ncbi.nlm.nih.gov) were downloaded after a BLAST search against a query Neutral Ceramidase sequence of Mus musculus. A total of 30 sequences from three different classes (Mammalia, Reptilia, Aves, Arthropoda, and Pisces) were taken for analysis. Guidance server (http://guidance.tau.ac.il/) [8] using guidance algorithm and MUSCLE Algorithm for Multiple Sequence Alignment

2.1. MEGA -Molecular Evolutionary Genetics Analysis
MEGA V 6.0 [14] was used for multiple sequence alignment. The cladograms were constructed using MEGA 6.0 by Neighbor Joining (NJ) and UPGMA (Unweighted Pair Group Method with Arithmetic Mean) method. Conserved domain search was done using MEGA 6.0

2.2. NETOGLYC 4.0
NetOGlyc 4.0 server [12] was used to identify the mucin type GalNAc O-glycosylation sites in the study sequences. [8]

(http://www.cbs.dtu.dk/services/NetOGlyc/)

3. PHYLOGENETIC ANALYSIS
Neutral ceramidase enzyme was identified in genomes of different taxonomic orders; recent studies [5] indicate that the genetic information was conserved from bacteria to mammals. With an aim to gain insight on the evolution and point of origin of different domains in Neutral ceramidase in higher animals, the NCDase sequence information was analyzed and a phylogenetic tree of the organisms belonging to different classes was constructed. A total of 30 homologous protein sequences of Neutral Ceramidase enzyme from different source organisms along with were considered for *in silico* analysis. Only the sequences with at least 65% sequence similarity were chosen and aligned using MUSCLE [1] at Guidance server with 100 bootstrap repeats, the sequences were strongly aligned with high confidence scores. Historical insertions were deleted and the sequences were used to construct a phylogenetic tree by the (distance matrix) - Neighbor Joining (NJ) method [9], as it applies the Minimum Evolution principle (ME) at each step in the clustering process considering the non constant rate of the molecular clock. The evolutionary history was deduced with the NJ method with the bootstrap consensus tree inferred from 1000 replicates. Another phylogenetic tree constructed using (Un weighted Pair Group Method with Arithmetic Mean) UPGMA method which joins tree branches based on the criterion of greatest similarity among pairs and averages of joined pairs tree, and constructs the tree based on sequence dissimilarity and divergence based on the molecular clock [9] (**Figure 2**)

3. RESULTS

In this communication an attempt to trace the evolutionary history of NCDase has been made. Multiple sequence alignment of the 30 sequences showed significant homology between the sequences of Neutral Ceramidase from different species. Phylogenetic analysis was carried out on the amino acid sequences of NCDase enzyme proteins. The phylogenetic tree was calculated based on Neighbor Joining method (**Figure.1**) Analysis of the same sequences using UPGMA method predicteded a similar topology (**Figure.2**). The resulting phylogenetic tree is very well supported for most branches except for the divergence of Tarsius from primates along with Ovis. But given the fact that at molecular level the rate of evolution varies and the NCDase in Tarsius might have evolved distinctively due to unique selection pressures that resulted in accelerated evolution of NCDase in Tarsius.

Otherwise the relationships between other taxa for each NCDase orthologs are resolved in a manner that it agrees with previous phylogenic analyses, indicating a reliable evolutionary reconstruction [5]

Initial multiple sequence alignment and search for conserved sequences in the sequences revealed the presence of a highly conserved hexapeptide sequence **GDVSPN** in all the sequences analyzed. Previous studies indicate that it act as an active site with an additional role in the activity of neutral ceramidase and also in reverse CDase activity [3] which explains the conservation of the active site sequence and also playing a critical role in the NCDase activity, hence confirming the signature of natural selection in the evolution of Neutral ceramidase.

Domain analysis of the sequences studied revealed the presence of ser thr rich mucin like domains [17] exclusively in the vertebrate NCDase sequences and not in invertebrate and bacterial enzyme proteins, supporting the place of mammals in a separate branch from other classes in the constructed tree, The presence of serine threonine rich domain (Mucin box) that helps the localization of the enzyme to the surface of the plasma membrane as a type II integral membrane. [16]. we analyzed all the 30 sequences for Ser/Thr-rich regions as well as regions predicted to be highly *O*-glycosylated by NetOGlyc (http://www.cbs.dtu.dk). Only the glycosylated sites predicted with a confidence score of above **0.5** was taken in for consideration. [12] The sites are observed predominantly in the vertebral NCDase sequences (**Table.1**) and are more or less absent in bacterial, invertebrate and in the embryophyte sequences, this indicates that the specific domain might have been acquired during the later stages of vertebrate evolution.

4. DISCUSSION

In this communication we have tried to show the evolutionary history of neutral ceramidase in various different species by constructing an phylogenetic tree based on its amino acid sequences, NCDase sequences from 30 different species were aligned using MUSCLE algorithm, which revealed high sequence similarity between the sequences suggesting that the amino acid information of the protein is being preserved across species. The evolution of NCDase was deduced by constructing a phylogenetic tree using distance matrix methods [9] [11] measuring the genetic distances between the sequences, the resulted tree can be used to illustrate both the evolutionary relationship of the organisms as well as the protein itself. The tree constructed using both the algorithms resulted in approximately the same tree of life in agreement with the generally agreed evolutionary sequences interpreted by other comparative sequence data sets. However the

placement of Tarsius along with Ovis branch using NJ method if true can be explained through a differential rate of molecular evolution.

Table 1. Predicted Glycosylated Sites

Organism	Sites	Organism	Sites
Drosophila melanogaster	558, 568	Nasonia vitripennis	299, 303
Triticum aestivum	299,330, 380	Dictyostelium discoideum	434
Danio rerio	46,51,53,55,46	Alligator mississippiensis	48,49,50,51,52,53,59,60,61
Macaca fascicularis	50,51,2,58,61,62,64,66,67,68,70,73,74,76,78,79,80,82,84,85,88,89,90	Chlorocebus sabaeus	44,50,51,52,58,61,62,64,66,67,68,70,73,74,76,78,80,82,84,85,88,89,90
Pan paniscus	50,51,52,54,58,61,62,64,67,68,70,73,74,76,78,79,80,82,84,85,88,90	Tarsius syrichta -	50,51,52,54,58,61,62,64,70,72,73,76,78,79
Homo Sapiens	50,51,52,54,58,61,62,67,68,70,73,74,76,78,79,80,82	Mus musculus	50, 51, 52, 55, 56, 57, 58, 64, 66
Myotis davidii	64,65,67,71,73,75,80,81,86,87,89,93,94,95,101,102,103	Tupaia chinensis	44, 46, 51, 52, 54, 55, 61, 62, 64, 66, 67, 70, 72
Balaenoptera acutorostrata scammoni	42,46,47,53,57,62,64,66,67,42,46,47,53,57,62,64,66,67	Callithrix jacchus	48,49,50,52,53,56,59,60,62,64,65,66,68,70,71,74,76,77,78,80,82
Equus caballus	44, 47 48, 49, 51, 55, 57, 63, 67,70, 71, 76, 77	Felis cattus	44,49,51,54,55,58,59,61,64,65,67,70,71,73,79,80,81
Microtus ochrogaster	50,51,52,58,60,61,63,64,67,50,51,52,58,60	Condylura cristata	44,46,47,52,53,,55,56,57,62,63,67,68,69
Oryctolagus cuniculus	50,51,52,54,58,64,66,70,50,51,52,54	Camelus ferus	44,48,49,51,52,55,58,59,64,65,66,72,74,77,79,81,82

Vicugna pacos	44,48,49,51,5 2,55,58,59,64, 65,66,72,74,7 7,79,81,82	Otolemur garnettii	44,46,51,52, 56,58,62,64, 68,70,76,80, 82,85,87,89, 90
Jaculus jaculus	50,51,52,55,5 8,60,61,64	Pantholops hodgsonii	42,49,50,52, 56,59,65,68, 72,73,74,77, 79,81,82
Ictidomys tridecemline atus	44,51,52,54,5 5,58,61,64,67, 70	Ovis aries	51,52,54,58, 61,67,68,70, 74,75,76,79, 81,83,84
Canis lupus familiaris	44,48,49,51,5 2,53,54,55,58, 59,61,64,65,6 7,70,71,73,75, 80,81		

5. CONCLUSION

In this communication we have tried to construct a phylogentic tree of life using the differences in the a amino acid sequences of NCDase, a protein that plays a critical role in cell regulation and found widely in most of the species, to trace the historical changes throughout the course of evolution. An algorithm that employs minimum evolution principle to achieve the genetic distances observed now. The presence of a highly conserved hexapeptie active site sequence (**Figure.3**), relatively unaltered in all the species indicate that the protein has been subjected to severe selection pressure that the active site remains more or less unaltered over the long course of evolution. The presence of ser thr rich mucin like domains **[17]** exclusively in vertebrate NCDase sequences and its total absence in invertebrate sequences can be used to track the changes of the protein itself over the course of time. .Comparing more sequences from other organisms and using algorithms that applies other principles would help to construct a highly defined tree of life.

6. ACKNOWLEDGMENTS

7. REFERENCES

[1] Edgar, Robert C. "MUSCLE: multiple sequence alignment with high accuracy and high throughput." Nucleic acids research 32.5 (2004): 1792-1797.

[2] Felsenstein J. Confidence limits on phylogenies: An approach using the bootstrap. Evolution (1985) 39:783-791.

[3] Galadari, S., Wu, B. X., Mao, C., Roddy, P., El Bawab, S., & Hannun, Y. X. (2006). Identification of a novel amidase motif in neutral ceramidase. Biochem. J, 393, 687-695..

[4] Houben, E., Holleran, W. M., Yaginuma, T., Mao, C., Obeid, L. M., Rogiers, V., & Uchida, Y. (2006). Differentiation-associated expression of ceramidase isoforms in cultured keratinocytes and epidermis. Journal of lipid research, 47(5), 1063-1070. (2006): 1063-1070.

[5] Ito, Makoto, (2002) Molecular evolution of Neutral Ceramidase : From Bacteria to Mammals In Ceramide Signaling (41-45) Landes Bioscience, Texas, U.S.A

[6] Maher, D. M., Gupta, B. K., Nagata, S., Jaggi, M., & Chauhan, S. C. (2011). Mucin 13: structure, function, and potential roles in cancer pathogenesis.Molecular Cancer Research, 9(5), 531-537.

[7] Nakagawa, T., Tani, M., Sueyoshi, N., & Ito, M. (2011). The Mucin Box and Signal/Anchor Sequence of Rat Neutral Ceramidase Recruit Bacterial Sphingomyelinase to the Plasma Membrane. Bioscience, biotechnology, and biochemistry, 75(5), 987-990.

[8] Penn, O., Privman, E., Ashkenazy, H., Landan, G., Graur, D., & Pupko, T. (2010). GUIDANCE: a web server for assessing alignment confidence scores. Nucleic acids research, 38(suppl 2), W23-W28.

[9] Saitou N. and Nei M.The neighbor-joining method: A new method for reconstructing phylogenetic trees. Molecular Biology and Evolution (1987) 4:406-425.

[10] Shayman, James A. "Sphingolipids." Kidney international 58.1 (2000): 11-26.

[11] Sneath P.H.A. and Sokal R.R. (**1973**). Numerical Taxonomy. Freeman, San Francisco.

[12] Steentoft, C., Vakhrushev, S. Y., Joshi, H. J., Kong, Y., Vester-Christensen, M. B., Schjoldager, K. T. B., ... & Clausen, H. (2013). Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. The EMBO journal, 32(10), 1478-1488..

[13] Swofford, D. L., Olsen, G. J., Waddell, P. J., & Hillis, D. M. (1996). {Phylogenetic inference}.

[14] Tamura K., Stecher G., Peterson D., Filipski A., and Kumar S.MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Molecular Biology and Evolution (2013)30: 2725-2729.

[15] Tani, M., Okino, N., Mitsutake, S., Tanigawa, T., Izu, H., & Ito, M. (2000). Purification and Characterization of a Neutral Ceramidase from Mouse Liver a single protein catalyzes the reversible reaction in which ceramide is both hydrolyzed and synthesized. journal of biological chemistry, 275(5), 3462-3468

[16] Tani, Motohiro, Hiroshi Iida, and Makoto Ito. "O-glycosylation of mucin-like domain retains the neutral ceramidase on the plasma membranes as a type II integral membrane protein." Journal of Biological Chemistry 278.12 (2003): 10523-10530.

[17] Zuckerkandl E. and Pauling L. Evolutionary divergence and convergence in proteins. Edited in Evolving Genes and Proteins by V. Bryson and H.J. Vogel, (1965):97-166. Academic Press, New York.
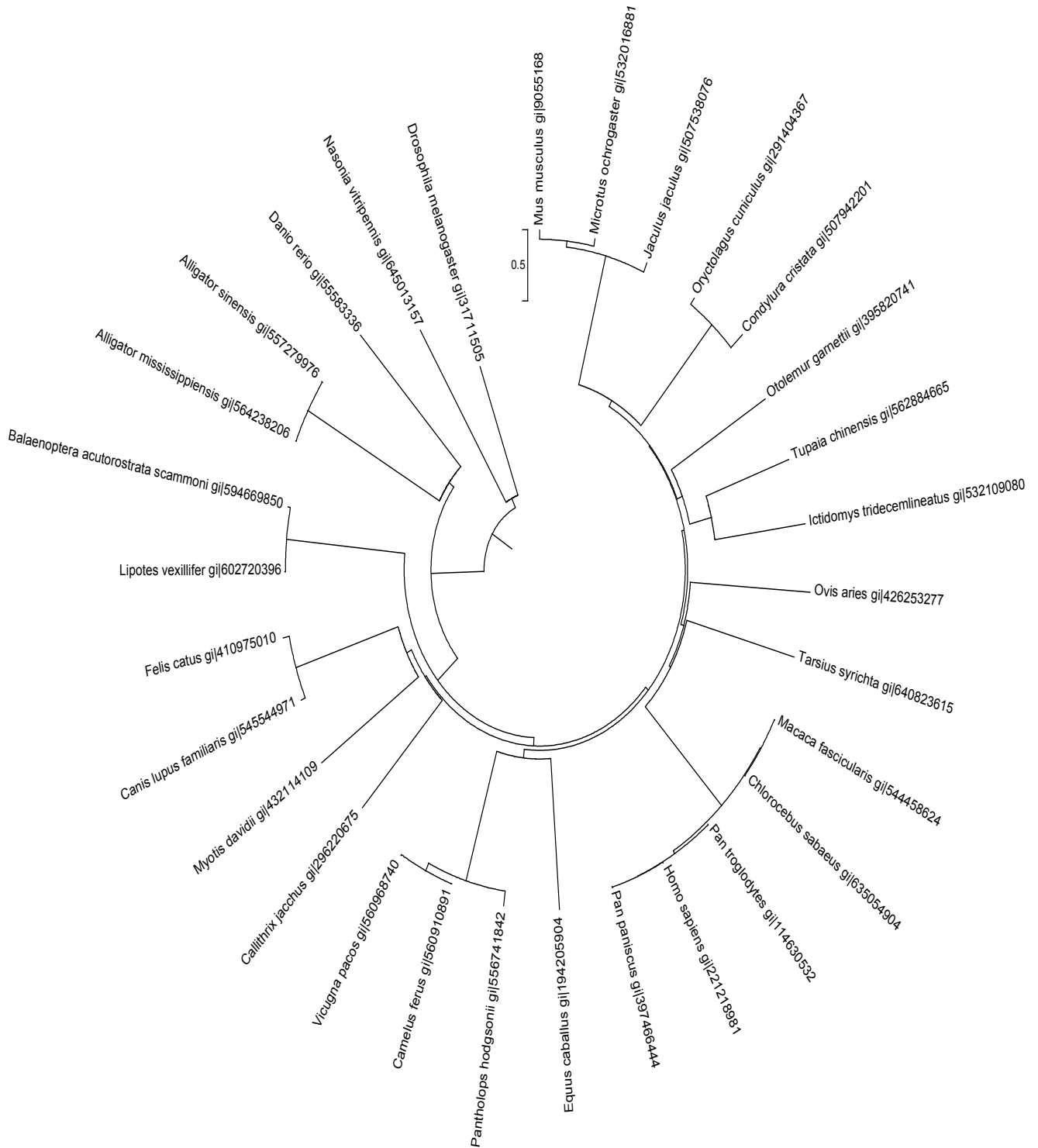
Fig 1: Figure. 1. Evolutionary relationships of taxa (Neighbor-Joining method)

The evolutionary history was inferred using the Neighbor-Joining method **[9]**. The bootstrap consensus tree inferred from 1000 replicates is taken to represent the evolutionary history of the taxa analyzed **[2]**. Branches corresponding to partitions reproduced in less than 50% bootstrap replicates are collapsed. The evolutionary distances were computed using the Poisson correction method **[17]** and are in the units of the number of amino acid substitutions per site. The analysis involved 41 amino acid sequences. All positions containing gaps and missing data were eliminated. There were a total of 441 positions in the final dataset. Evolutionary analyses were conducted in MEGA6 **[14]**
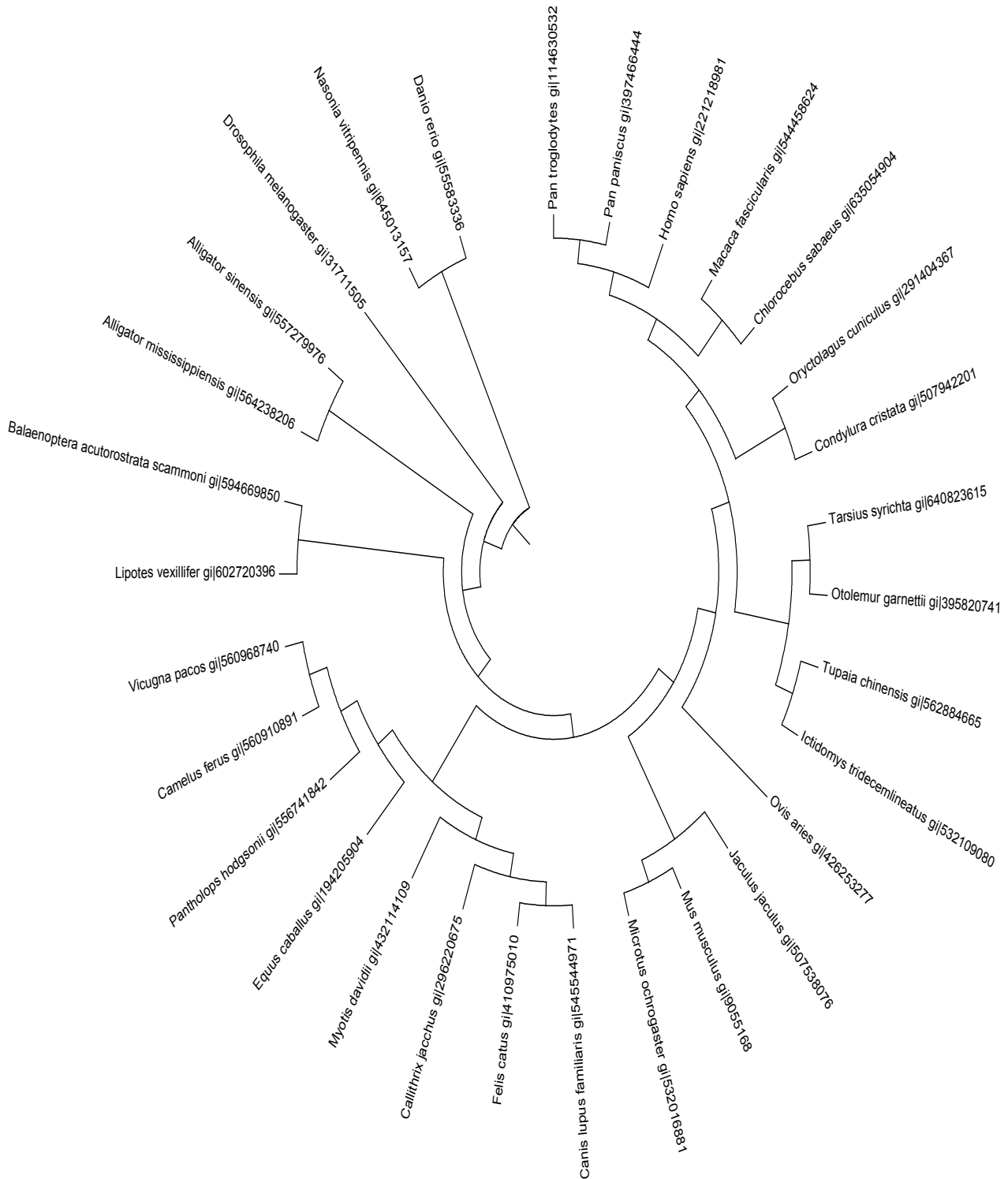
Figure.2 Evolutionary relationships of taxa (UPGMA)

The evolutionary history was inferred using the UPGMA method [11]. The optimal tree with the sum of branch length = 20.65974347 is shown. The evolutionary distances were computed using the Poisson correction method [17] and are in the units of the number of amino acid substitutions per site.

The analysis involved 30 amino acid sequences. All positions containing gaps and missing data were eliminated. There were a total of 668 positions in the final dataset. Evolutionary analyses were conducted in MEGA6 [14].

```
Drosophila      262.............KFVGAFCSSNLGDVSPNIMGPKCSISGNECDLLTSRCPTGEGD.CFA
Nasonia         269.............DFVAAFASTNLGDVSPNTRGPKCEFSGQHC.TVQYTCP.GKKEMCFA
Dictyostelium   279.............PFIAAFGQSNEGDVSPNTRGPTCRD.GKPCDYKTSTCN.GKVEECWA
Triticum        332.............KFVSAFCQSNCGDVSPNVLGTFCIDTHLPCDFNHSTCN.GKNELCYG
Danio           298.............PFVAGFSSSNLGDSSPNIRGPVCVNTGLKCDYINSSCPVGGKKACIA
Alligator       310.............SFVAAFASSNLGDVSPNTKGPHCINTGESCDNPQSYCPVGGAKMCMA
Alligator       310.............SFVAAFASSNLGDVSPNTKGPHCINTGETCDNPQSYCPVGGAKMCMA
Mus             316.............PFVAGFASSNLGDVSPNILGPHCVNTGESCDNDKSTCPNGGPSMCMA
Microtus        316.............PFVAGFASSNLGDVSPNILGPHCTNTGESCDNEKSTCPVGGPSMCMA
Equus           343.............SYVAAFASSNLGDVSPNTLGPHCINTGESCDNPNSSCPIGGPNMCIA
Jaculus         316.............PFVAGFASSNLGDVSPNILGPHCVNTGESCDNPNSTCPIGGSSMCIA
Myotis          353.............PFVAAFASSNLGDVSPNILGPRCANTGESCDNANSTCPIGGATMCMA
Condylura       319.............PYVAAFASSNLGDVSPNILGPHCVDTGESCDNANSSCPLGGPNMCMA
Canis           331.............PYVAAFASSNLGDVSPNILGPHCVNTGESCDNANSSCPIGGSSMCMA
Felis           331.............PYVAAFASSNLGDVSPNILGPRCVNTGDSCDNANSSCPIGGSSMCIA
Camelus         332.............PYVAAFASSNLGDVSPNILGPHCINTGESCDNANSSCPVGGPSMCIA
Vicugna         332.............PYVAAFASSNLGDVSPNILGPHCINTGESCDNANSSCPVGGPSMCIA
Pantholops      332.............PYVAAFASSNLGDVSPNVLGPHCTNTGESCENANSSCPIGGASMCVA
Ovis            334.............PYVAAFASSNLGDVSPNVLGPHCTNTGESCENANSSCPIGGASMCVA
Lipotes         317.............PYVAAFASSNLGDVSPNILGPHCINTGESCDNANSTCPIGGPSMCIA
Balaenoptera    317.............PYVAAFASSNLGDVSPNILGPHCINTGESCDNANSTCPIGGPSMCIA
Tupaia          322.............PYVAAFASSNLGDVSPNILGPHCTNTGESCDNANSTCPIGGPGMCVA
Oryctolagus     319.............PFVAAFASSNLGDVSSNIVGPHCINTGESCDNANSSCPIGGPNMCIA
Otolemur        339.............PYVAAFASSNLGDVSPNILGPHCINTGESCDNANSTCPIGGPGMCMA
Tarsius         328.............PFVAAFASSNLGDVSPNVLGPHCINTGESCDNANSTCPIGGPSMCIA
Callithrix      338.............PFVAAFASSNLGDVSPNILGPHCINTGESCDNPNSTCPIGGPSMCIA
Homo            340.............PFVAAFASSNLGDVSPNILGPRCINTGESCDNANSTCPIGGPSMCIA
Pan             340.............PFVAAFASSNLGDVSPNILGPRCINTGESCDNANSTCPIGGPSMCIA
Pan             340.............PFVAAFASSNLGDVSPNILGPRCINTGESCDNANSTCPIGGPSMCIA
Macaca          340.............PFVAAFASSNLGDVSPNILGPRCINTGESCDNANSTCPIGGSNMCIA
Chlorocebus     340.............PFVAAFASSNLGDVSPNILGPRCINTGESCDNANSTCPIGGSNMCIA
```

Figure 3. Alignment of the neutral ceramidases from various organisms showing the highly conserved sequence, GDVSPN within the larger conserved amidase domain, NXGDVSPNXXG