Pakistan J. Zool., vol. 46(4), pp. 1139-1145, 2014.

Molecular Cloning and Evolutionary Analysis of *FUCA1* Gene in Bats

Haikui Chen[1§], Tianyun Zhang[2§] Shouhe Lin[3] and Xiaohong Cao[1]*

[1] *Department of Life Science, Beifang University of Nationalities, Yinchuan 750021, China*
[2] *The First Hospital, The Chinese People's Liberation Army (PLA), Lanzhou 730050, China*
[3] *Department of math and information science, Beifang University of Nationalities, Yinchuan 750021, China*
[§]**These two authors contributed equally to this work**

Abstract.- Bats are mammals of the order Chiroptera, which represents about 20% of all mammalian species worldwide. Bats have evolved various reproductive strategies, such as delayed sperm maturation, sperm storage, delayed fertilization, delayed implantation and delayed development. The strategies of delayed sperm maturation allow female bats, especially hibernating species, to reproduce successfully. Alpha-L-Fucosidase 1 gene (*FUCA1*) has been proved to be involved in the processes of sperm maturation and sperm-egg interactions in mammals. It is still not clear how the *FUCA1* gene evolved in bats. Therefore, it is important to determine the evolutionary pattern of *FUCA1* gene and positively selected sites which affect the reproductive activity of bats. In this work, we cloned the coding regions of *FUCA1* from 7 species of bats and performed an evolutionary study of this gene. Although selection pressure varied among different lineages, *FUCA1* gene underwent positive selection in the ancestral species leading to Old World fruit bats. Our results demonstrated that *FUCA1* is basically conserved in hibernating bats but has evolved within the ancestral species of Old World fruit bats to adapt their survival environments.

Keywords: *FUCA1*, evolutionary analysis, bat, molecular cloning, glycosidase, glycoside hydrolase, fucosidase.

INTRODUCTION

Bats are one of the most extraordinary mammals orders because of their adaptations to flight (Yin *et al.,* 2011), and mammals of the order Chiroptera, which represent about 20% of the mammals and geographically distributed widespread on earth (Bininda-Emonds and Simmons, 2002). Different species of bat are widely distributed in various climate and landform conditions all over the world. During the evolutionary process, bats have evolved many strategies for their survive in complex geographical environments, adverse climates. Many bat species have evolved some special reproductive strategies, such as delayed fertilization, delayed implantation, sperm storage and delayed development. The reproductive strategies allow many bat species, especially hibernating bats, reproduce during the period of foodless winter.

Alpha-L-fucosidases is a kind of lysosomal enzyme that can catalyze the removal of fucose from glycolipids, glycoproteins and oligosaccharides. It has been proved that alpha-L-fucosidases also play an important role in transglycosylation process in mammals (Abascal and Skalaban, 1998). Until now, there are two distinct glycoside hydrolase (GH) families of alpha-L-fucosidases that have been identified: alpha-L-fucosidases that catalyze the hydrolysis using a retaining mechanism belonging to GH family 29 (GH29) (Cantarel *et al.,* 2009), and the others which could catalyze the hydrolysis belong to the novel GH family 95. Deficiency of mammalian alpha-L-fucosidases has been proved to cause the fucosidosis, which results in the lethal accumulation of glycoproteins and fucosylated glycosphingolipids in the nervous system (Delves, 1998). Until now, the *FUCA1* gene is the only member of alpha-L-fucosidases that has been widely studied. In humans, the *FUCA1* gene has received wide attention in the development of storage disease. The expression of *FUCA1* increased dramatically during the development of hepatocellular carcinoma, thus it could be used as a biological marker of for diagnose of hepatocellular carcinoma. The *FUCA1* is present in the plasma membrane of mammalian sperm. It has been documented that FUCA1 could interact with glycoconjugates of the surface of the eggshell in bulls, and its biological functions are conserved across mammalian species (Srivastava *et al.,* 1986). Up to date, we have realized that FUCA1 plays an

* Corresponding author: address:cxh8892@sohu.com.
0030-9923/2014/0004-1139 $ 8.00/0

important role in gametes recognition in several mammalian species, such as humans and rats (Cordero and Merino, 2001). Although several FUCA1 genes have been identified in mammals, the evolution of FUCA1 genes remain uncharacterized in bats.

Gene evolution can be explained by the signatures of purifying selection, positive selection and neutral evolution. Purifying selection may be of great interests because it may help to detect regions of functional constraints; positive selection in the functional importance regions may contribute to important functional changes. Although several *FUCA1* genes have been identified in mammals, the evolution of *FUCA1* genes remain uncharacterized in bats. In this work, we cloned the coding region of the *FUCA1* gene from 7 bat species of bat and compared them with other orthologs. Molecular phylogenetic trees of *FUCA1* in bats were constructed by different methods. In this study, we sought to analyze the variation in selective pressures across sites and lineages in *FUCA1* gene.

MATERIALS AND METHODS

Sample collections

Seven bat species (*Rousettus leschenaulti*, *Cynopterus sphinx*, *Hipposideros armiger*, *Rhinolophus ferrumequinum*, *Rhinolophus pusillus*, *Tadarida teniotis*, *Murina leucogaster*) were selected in this study, which cover the main categories of order Chiroptera. *FUCA1* cDNA sequences were downloaded from NCBI Genbank and Ensembl databases, including the human (*Homo sapiens,* accession number: NM_000147), little brown bat (*Myotis lucifugus,* accession number: ENSMLUG00000005281) and large flying fox (*Pteropus vampyrus,* accession number: ENSPVAG00000013615), respectively.

Gene cloning and sequencing

Total RNAs were exacted from liver tissues of seven bat species using the RNAiso kit (Takara, Japan), and then were reverse transcribed to cDNA with SuperScript III (Invitrogen, USA) according to the manufacturer's instructions. One pair of primers (Table I, F1/R1) were designed on the basis of conserved regions of other mammal *FUCA1* genes to obtain about 1000 bp coding sequences, and another pair of the primers (Table I, F2/R2) were designed to get the other 600bp coding sequences from *FUCA1* genes. The PCR protocol of F1/R1 involved denaturation at 95°C for 5 min, then 31 cycles of 95°C for 30 s, 51°C for 30 s, and 72°C for 1 min 20s. The another amplification conditions were: 95°C for 5 min, then 31 cycles of 95°C for 30 s, 53°C for 30 s, and 72°C for 1 min. Finally, PCR products were all extended at 72°C for 10 mins.

Table I.- Primers sequences used in this study

Primer	Sequence
F1	5'-ATG(T/C)GGT(T/C)(T/C)CC(G/C)GAAAT-3'
R1	5'-CT(T/G)ACTGTCTG (A/G)ACCAGTTCCGA(A/C) (A/C)TGAT -3'
F2	5'-AACTG(T/C)TC(T/C)TG(T/C)C(A/G)CCATGG-3'
R2	5'-CT(T/C)CACGCCCATCATCT(T/G)TGTA -3'

The 900 bp and 600 bp products from PCR of was electrophoresed in the 0.8% agarose gel and stained with ethidium bromide. The PCR products were ligated into pGEM-T-Easy vectors (Promega, WI) and cloned. More than five positive clones of each species were chosen for sequencing, by Majorbio (Shanghai, China). All sequences have been submitted to NCBI Genbank with the accession NO. (KF793809, KF793810, KF793811, KF793812, KF793813, KF793814, KF793815, KF793816, KF793817).

Sequence alignment and homology modeling

Multiple alignment for the *FUCA1* genes was performed using ClustalX 1.83 software (Larkin *et al.,* 2007). The Kimura 2-parameter model and the poisson correction model were used to calculate the pairwise genetic distances for nucleotide and amino acid sequences of bat *FUCA1* genes by MEGA 5.02, respectively (Kumar and Tamura, 2004). The DNA sequences were deduce to amino acid sequences of bat *FUCA1* genes. Protein molecular weight, topology and pI were acquired from the ProSA-web and NQ-Flipper page V2 respectively (Wiederstein and Sippl, 2007). Trans-membrane segments and protein orientation were predicted by using the TMpred and SOSUI software (Notredame *et al.,* 2000; Weichenberger *et al.,* 2008). Next, phylogenetic analyses were performed by using

MEGA 5.02 with Maximum Likelihood (Juckes-Cantor model) and Minimum Evolution methods (Juckes-Cantor model) by using *FUCA1* gene sequences of 10 mammal species.

Evolutionary analysis

The strength of natural selection acting on a protein-coding gene can be measured by the ratio of nonsynonymous substitutions rate (*dN*) to the synonymous substitutions per site (*dS*). If *dN/dS*=1, it represents that the gene is under neutral evolution; if *dN/dS>1*, it is under positive selection; if *dN/dS*<1, it is under purifying selection. To investigate the variable selective pressures in the *FUCA1* gene of bats, maximum likelihood analysis of the sequence evolution was performed with CODEML program implemented in the PAML 4.0 package (Yang, 2007). Three models were used in the branch analysis: 1) Two ratio model, which assumes the different ω ratios for the foreground branches and background branches, respectively; 2) Free ratio model, which assumes independent ω ratios for all branches. To test if the ratio was deviated from the expected one, we performed the likelihood ratio tests (LRT) by comparing with 3) one ratio model, which assigns an equal ω ratio for all branches in the phylogeny; The LRT statistic was determined by chi-squared test by P-value < 0.05. A false discovery rate of 10% was used to adjust the significance.

Since most amino acids are under functional constraints and adaptive evolution only affects few sites at a few time points, models of variable selective pressures among amino acid sites were used to detect for the presence and identification of sites. We herein used the branch-site model (Yang, 2005) nested in CODEML program to detect positively selected sites, which allows ω to vary among sites along particular lineages, with the null hypothesis that all sites have similar ω value**.**

RESULTS

Sequence analysis and phylogenetic tree construction

In this work, the coding regions of *FUCA1* gene were obtained from 9 bat species. We generated alignments of *FUCA1* coding sequences by the Clustal X (Larkin *et al.,* 2007) comprising 10 taxa, spanning 1,407 nucleotides, and equating to 469 amino acids, at positions 1–469 as referenced to the mature human *FUCA1* amino acid sequence. No stop codons that may have been caused by insertion/deletion mutations were detected in the partial coding sequences of 7 bat species we sequenced. As the result, the coding sequences were highly conserved across the 10 species we examined. Next, we compared the pairwise sequence similarities of *FUCA1* amino acid sequences. The average sequence similarity between *Pteropus vampyrus* and other Old World fruit bats was 95.7 %, whereas the similarity was only 87.5 % between the *Pteropus vampyrus* and hibernating bats (Table II). Moreover, compared with humans, more amino acid substitutions were found in non-hibernating bats than hibernating bats. Next, phylogenetic trees were constructed by using maximum likelihood and minimum evolution methods (Figs. 2, 3), and they show identical topologies. As the phylogenetic trees shown, Pterropodidea, Rhinolophidae and Vespertilionidea lineages are present in the right positions according to the advanced molecular phylogenetic reconstruction of the lineages of Chiroptera, and the boot strap probabilities are relatively high.

Evolutionary analysis

The free ratio test results showed that the free ratio model, which allows the dN/dS ratio to vary among branches, fits significantly better than the one-ratio model, in which dN/dS was fixed across all branches (log likelihood test statistic ($2\Delta\ell$) = 47.6, *d.f.* = 16, *P value* < 0.001). The ω values for all branches in the phylogenetic tree ranged from 0 to 0.90, and the ω values for the ancestral branches of Old World fruit bats were 0.87. These results indicated that *FUCA1* gene might undergo accelerated evolution in the ancestral species of Old World fruit bats. We then used the two-ratio branch model (in which *dN/dS* ratios vary between pterropodidae clade branches and other branchesto test for the evidence of selection changes on the ancestral species of Old World bats. The two-ratio models (with the Old World fruit bat ancestral branch fixed as foreground) was a significantly better fit than the one-ratio model (ω value of pterropodidae clade branches is 0.89 and ω value of

Table II.- Similarity of *FUCA1* sequences between bats and humans.

Species	1	2	3	4	5	6	7	8	9	10
1. *Homo sapiens*										
2. *Pteropus vampyrus*	0.84									
3. *Cynopterus sphinx*	0.83	0.95								
4. *Rousettus leschenaulti*	0.83	0.96	0.95							
5. *Hipposideros armiger*	0.85	0.87	0.86	0.86						
6. *Rhinolophus ferrumequinum*	0.85	0.87	0.86	0.86	0.93					
7. *Rhinolophus pusillus*	0.85	0.86	0.85	0.86	0.92	0.97				
8. *Myotis lucifugus*	0.85	0.85	0.85	0.85	0.87	0.88	0.87			
9. *Tadarida teniotis*	0.85	0.86	0.86	0.85	0.87	0.88	0.87	0.91		
10. *Murina leucogaster*	0.86	0.86	0.85	0.85	0.87	0.88	0.87	0.96	0.91	

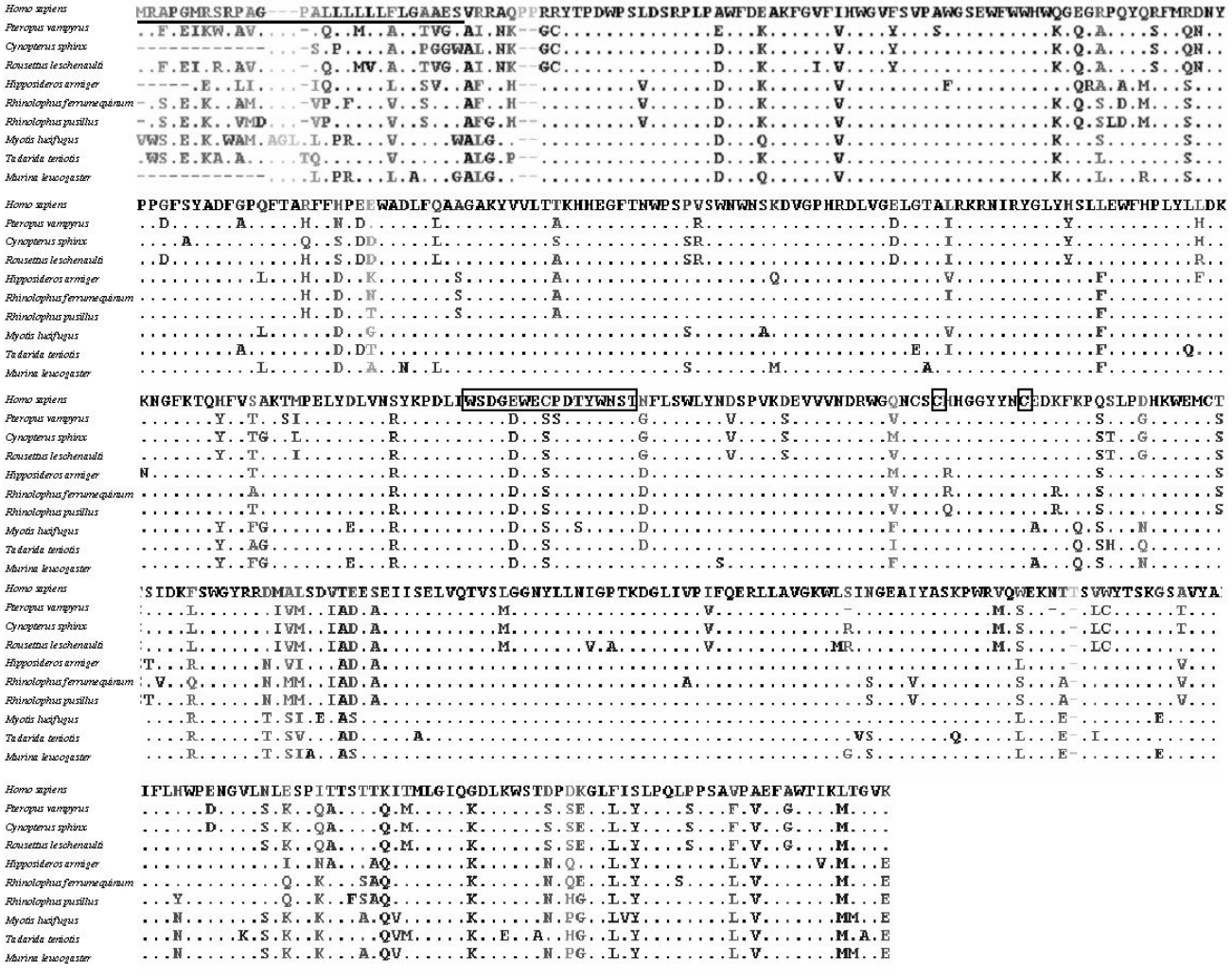

Fig. 1. Alignment of FUCA1 amino acid sequences in bats. The predicted trans-membrane region was underlined. The predicted catalytic nucleophile and catalytic acid are boxed. The dissimilar amino acid residues are in black and the similar amino acid residues are in gray.

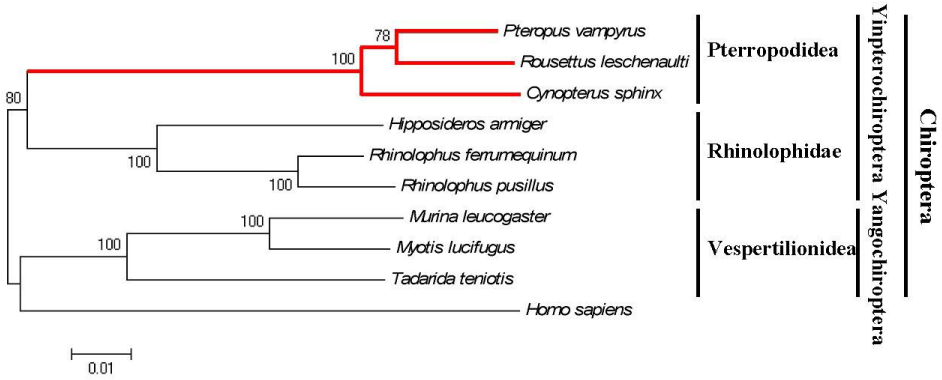

Fig. 2. The maximum likelihood tree based on complete sequences of FUCA1 genes from Chiroptera lineages. The values on the nodes indicate the numbers of bootstrap support. The lineages with positive selection were shown in red.

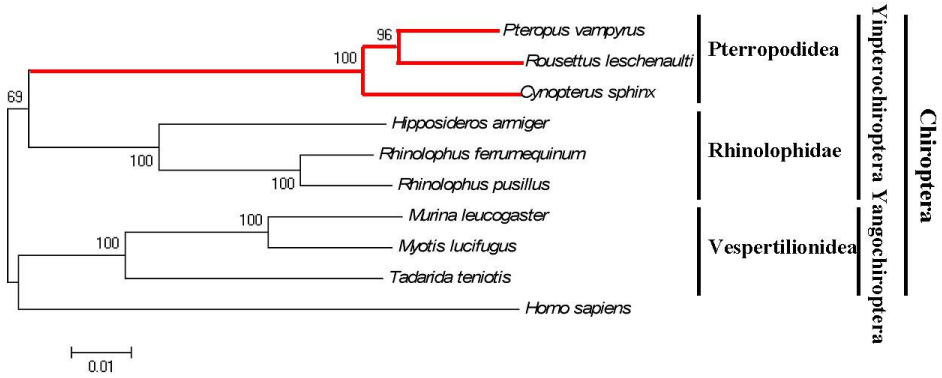

Fig. 3. The minimum-evolution tree based on complete sequences of FUCA1 genes from Chiroptera lineages. The values on the nodes indicate the numbers of bootstrap support. The lineages with positive selection were shown in red.

other branches is 0.25, respectively). Also, the one-ratio model could be rejected by the two-ratio model that fixed the Old World ancestral branch as foreground ($2\Delta\ell = 23$, *d.f.* = 1, *P value* < 0.01).

Next, the branch-site model was applied to detect evidence of positive selection on the Old World fruit bats ancestral branch. The Old World fruit bats ancestral branch was considered as the foreground branch for the small dataset and all other branches in the phylogenetic tree were assumed as the background branches. The result showed that the ω values of the positively selected sites in the foreground was 3.39 for the Old World fruit bats ancestral branch, and the branch-site model A was significantly better than the null model ($2\Delta\ell = 5.2$, *d.f.* = 1, *P value* = 0.023). The branch-site model A revealed that 21 sites suffered positive selection along the ancestral species of the Old World fruit bats (Table III).

DISCUSSION

In the present study, we cloned the coding region of *FUCA1* from 7 bat species, and attempted to characterize the evolutionary pattern of this gene in bats. The ω value provides a sensitive measure of selective pressure on the protein. Most amino acids are under functional constraints and adaptive evolution only affects few sites at a few time points. Although reproduce-related genes are one of the important pools for adaptive evolution, evolutionary analyses have revealed that *FUCA1* gene has

Table III.- Branch-site model A test for positively selected sites in the ancestral branches of the Old World fruit bats.

Branch-site model	np	LRT	$2\Delta\ell$ (*p value*)	Positively selected sites
Model A (alternative hypothesis) for the ancestral species of Old World fruit bats	21	-5123.010	5.2 (0.023)	6M(0.69) 29E(0.76) 34A(0.89) 38R(0.72) 39R(0.69) 70F(0.75) 94R(0.72) 97S(0.50) 98D(0.71) 152V(0.94) 170E(0.70) 186H(0.72) 255D(0.75) 260D(0.92) 313M(0.72) 334L(0.69) 353I(0.72) 380V(0.73) 390W(0.67) 451P(0.64) 460A(0.69)
Model A (null hypothesis) for the ancestral species of Old World fruit bats	22	-5120.445		Not allowed

Np, number of parameters; LRT, likelihood ratio test.

suffered strong purifying selection in mammals. The process of gametes recognition is very important for the mammalian fertilization. Bats colonize and occupy vastly diverse habitats, with different reproductive strategies, such as delayed sperm maturation, sperm storage, delayed fertilization, and so on (Geiser and Ruf, 1995; Dausmann *et al.,* 2004). The knowledge about *FUCA1* gene in bat have not been recognized or annotated until now. It might be a potential target for the study of reproductive strategies and hibernating strategies of mammals. Here, we showed an positive selection of *FUCA1* gene in the ancestral species of Old World fruit bats.

It has been documented that *FUCA1* gene plays an important role in gametes recognition in several mammalian species, such as human, rat, amphibians (Brewis and Moore, 1997; Seppala and Koistinen, 2007). Since the *FUCA1* gene is important for reproductive strategies of bats, a strong functional constraint of *FUCA1* gene was shown in mammals, which was caused by strong purifying selection. However, we also noticed that the ω values at some sites of the *FUCA1* sequences in the Old World fruit bats were relatively higher than other bats, and many positive selected sites along the protein sequences have been identified in this work. These amino acid residues are responsible for the substrate size selection at the substrate entry channel in *FUCA1*. The evolutionary analysis implied that positive selection had stressed on the Pteropodidate lineages, suggesting that *FUCA1* had accelerated changes in the lineage of Old World fruit bats after the split of Old World fruit bats and other bats. Phylogenetic analysis can help us to address whether *FUCA1* has evolved within species to adapt to the new survival environments. According to previous study about ancestors of bats, heterothermy (hibernation and daily torpor) is the ancestral states of bats (Jones *et al.,* 2002; Yuan *et al.,* 2011). Compared with the Old World fruit bats (non-hibernating bats), hibernating bats living in limited food, long-term cooling areas, usually feed on insects and require high energy to survive. So, the reproductive strategies, such as delayed sperm maturation and delayed fertilization, are very important for these hibernating bats (Geiser and Ruf, 1995; Dausmann *et al.,* 2004). Whereas the Old World fruit bats who live in mild-climate environment with abundant food have their own reproductive strategies, which makes them to control the efficiency of reproduction. Therefore, it may be postulated that gene has evolved within species for the adaptation of reproductive strategies.

In summary, we cloned and sequenced the *FUCA1* gene from 7 bats and examined the evolution pattern of *FUCA1* in bats. The results showed an strong selection constraint leading to mammal *FUCA1* gene while positive selection has acted on the sequences of ancestral species of the Old World fruit bats. Twenty-one positively selected sites have been identified and all have been fixed. *FUCA1* gene might have evolved within the Old World fruit bat species to adapt newly emerged challenging environments.

ACKNOWLEDGEMENTS

This work was supported by the National

Technology support project (No.2012BAK17B07), and the National Natural Science Foundation of China (No.31060217)

Competing interests

The authors have declared that no competing interests exist.

REFERENCES

ABASCAL, I. AND SKALABAN, S. R., 1998. Alteration of the isoform composition of plasma-membrane-associated rat sperm alpha-L-fucosidase during late epididymal maturation: comparative characterization of the acidic and neutral isoforms. *Biochem. J.,* **333**: 201.

BININDA-EMONDS, O. R. AND SIMMONS, N. B., 2002. A phylogenetic supertree of the bats (Mammalia: Chiroptera). *Biol. Rev.,* **77**: 223-259.

BREWIS, I.A. AND MOORE, H.D., 1997. Molecular mechanisms of gamete recognition and fusion at fertilization. *Hum. Reprod.,* **12**(11 Suppl): 156-165.

CANTAREL, B.L., COUTINHO, P.M., RANCURE, C., 2009. The Carbohydrate-Active EnZymes database (CAZy): an expert resource for glycogenomics. *Nucl. Acids Res.,* **37**(suppl 1): D233-D238.

CORDERO, O.J. AND MERINO, A., 2001. Cell surface human alpha-L-fucosidase. *Europ. J. Biochem,* **268**: 3321-3331.

DAUSMANN, K.H., GLOS, J. AND GANZHORN, J.U., 2004. Physiology: hibernation in a tropical primate. *Nature,* **429**: 825-826.

DELVES, P. J., 1998. The role of glycosylation in autoimmune disease. *Autoimmunity*, **27**: 239-253.

GEISER, F. AND RUF, T., 1995. Hibernation versus daily torpor in mammals and birds: physiological variables and classification of torpor patterns. *Physiol. Zool.,* **48**: 935-966.

JONES, K.E., PURVIS, A. AND MACLARNON, N., 2002. A phylogenetic supertree of the bats (Mammalia: Chiroptera). *Biol. Rev.,* **77**: 223-259.

KUMAR, S. AND TAMURA, K., 2004. MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. *Brief. Bioinform.,* **5**: 150-163.

LARKIN, M.A., BLACKSHIELDS, G. AND BROWN, N.P., 2007. Clustal W and Clustal X version 2.0. *Bioinformatics.,* **23**: 2947-2948.

NOTREDAME, C., HIGGINS, D.G. AND HERINGA, T., 2000. T-Coffee: A novel method for fast and accurate multiple sequence alignment. *J. mol. Biol.,* **302**: 205-217.

SEPPALA, M. AND KOISTINEN, H., 2007. Glycosylation related actions of glycodelin: gamete, cumulus cell, immune cell and clinical associations. *Hum. Reprod. Update,* **13**: 275-287.

SRIVASTAVA, P. N., ARBTAN, K. AND TAKEI, G. H., 1986. α-L-fucosidase from bull seminal plasma: its purification and acrosome reaction promoting property. *Biochem. biophys. Res. Commun.,* **137**: 1061-1068.

WEICHENBERGER, C.X., BYZIA, P. AND SIPPL, M.J., 2008. Visualization of unfavorable interactions in protein folds. *Bioinformatics,* **24**: 1206-1207.

WIEDERSTEIN, M. AND SIPPL, M. J., 2007. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucl. Acids Res.,* **35**(suppl 2): W407-W410.

YANG, Z., 2005. The power of phylogenetic comparison in revealing protein function. *Proc. natl. Acad. Sci. USA,* **102**: 3179-3180.

YANG, Z., 2007. PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evolut.,* **24**: 1586-1591.

YIN, J., WANG, H., RACEY, P. AND ZHANG, S., 2011. Microanatomy of the fishing bat skin. *Pakistan J. Zool.,* **43**: 387-392.

YUAN, L., ZHAO, X., LIN, B., ROSSITER, S., HE, L., ZUO, X. AND HE, G., 2011. Adaptive evolution of leptin in heterothermic bats. *PLoS one,* **6**: e27189.

(Received 18 November 2013, revised 19 February 2014)