J. Mar. Biosci. Biotechnol. 2014, p. 102 - 109 Vol. 6, No. 2 [Research Paper]

http://dx.doi.org/10.15433/ksmb.2014.6.2.102 ISSN 2383-5400 (Online)

Molecular adaptation of the CREB-Binding Protein for aquatic living in cetaceans

Jae-Yeon Jeong[1], Ok Sung Chung[2], Young-Joon Ko[1], Kyeong Won Lee[1], Yun Sung Cho[2,3], Jong Bhak[2,3], Hyung-Soon Yim[1]*, Jung-Hyun Lee[1]*

[1]Marine Biotechnology Research Group, Biological Oceanography & Marine Biology Division, Korea Institute of Ocean Science & Technology. Ansan 426-744, Republic of Korea

[2]Personal Genomics Institute, Genome Research Foundation, Suwon 443-270, Republic of Korea

[3]The Genomics Institute, Biomedical Engineering Department, UNIST, Ulsan 689-798, Republic of Korea

(Received 24 December 2014, Revised 7 January 2015, Accepted 12 January 2015)

Abstract Cetaceans (whales, dolphins, and porpoises) are aquatic mammals that experienced drastic changes during the transition from terrestrial to aquatic environment. Morphological changes include streamlined body, alterations in the face, transformation of the forelimbs into flippers, disappearance of the hindlimbs and the acquisition of flukes on the tail. For a prolonged diving, cetaceans acquired hypoxia-resistance by developing various anatomical and physiological changes. However, molecular mechanisms underlying these adaptations are still limited. CREB-binding protein (CREBBP) is a transcriptional co-activator critical for embryonic development, growth control, metabolic homeostasis and responses to hypoxia. Natural selection analysis of five cetacean CREBBPs compared with those from 15 terrestrial relatives revealed strong purifying selection, supporting the importance of its role in mammals. However, prediction for amino acid changes that elicit functional difference of CREBBP identified three cetacean specific changes localized within a region required for interaction with SRCAP and in proximal regions to KIX domain of CREBBP. Mutations in CREBBP or SRCAP are known to cause craniofacial and skeletal defects in human, and KIX domain of CREBBP serves as a docking site for transcription factors including c-Myb, an essential regulator of haematopoiesis. In these respects, our study provides interesting insights into the functional adaptation of cetacean CREBBP for aquatic lifestyle.

Keywords : cetacean, CREB-binding protein, aquatic adaptation, craniofacial, haematopoiesis

Introduction

Cetaceans comprise baleen whales that filter plank-tons from the water using its sieve-like structure baleen and toothed whales (dolphins and porpoises) that hunt large animals with their sharp teeth. Cetaceans are derived from the Indohyus, an extinct semi-aquatic deer-like even-toed ungulates (artiodactyls) 54 million years ago in south Asia [1]. After became fully aquatic, baleen and toothed whales were split just before the Eocene/Oligocene boundary [2]. To adjust in aquatic life, cetaceans underwent major structural changes including streamlined body shape, the relocation of the nostrils toward the top of the cranium, the loss of external ears and the acquisition of hearing through the lower jaw, the alteration of the forelimbs into flippers, the degeneration of the hindlimbs and the development of flukes on the tail [3].

For a prolonged diving, cetaceans have evolved various strategies to cope with limited oxygen supply. They

* Corresponding authors
Phone: +82-31-400-6447/6243 Fax: +82-31-406-2495
E-mail: yimh@kiost.ac / jlee@kiost.ac

exhibit higher capacities for oxygen storage by increasing hemoglobin and myoglobin concentrations, bradycardia and redistribution of blood preferentially to nervous and cardiac tissues with peripheral vasoconstriction, metabolic modifications including greater glycolytic capacity and higher tolerance to lactic acid, and improved neuronal hypoxia tolerance with increased neuroglobin and cytoglobin contents [4]. However, it is not clear how cetaceans deal with hypoxia-related problems such as increased erythropoiesis during dives which causes heart attack and strokes by increasing blood viscosity, increased demand for glucose metabolism, angiogenesis, and oxidative damage after limited oxygen supply and reperfusion.

The CREB-binding protein (CREBBP, also known as CBP) is a transcriptional co-activator that interacts with many transcription factors including GATA1, SMADs, NCOAs, IKKs, EP300, SRCAP, and HIF1A [5]. Through its acetyltransferase activity and ability to scaffold many proteins, CREBBP plays essential roles in embryonic development, growth control, immune reaction, and response to hypoxia. Homozygous mutations of this gene causes embryonic lethality with defects in haematopoiesis and blood vessel formation [6] while mutation or loss of one allele is associated with Rubinstein-Taybi syndrome characterized by growth retardation, craniofacial abnormalities, and broad thumbs and toes [7].

To determine whether CREBBP has undergone evolutionary changes associated with adaptation to aquatic lifestyle, we analyzed CREBBP genes from four cetacean species representing three families of the suborder Odontoceti (toothed whales) and one family of the suborder Mysticeti (baleen whales) recently sequenced by our group [8]. In addition, publicly available CREBBP sequences from 15 mammalian genome assemblies were included in our analysis.

Materials and Methods

Nucleotide and protein sequences of the CREBBP gene

In this study, full-length open-reading frame sequences of four cetaceans and fifteen mammals were obtained by BLASTP search from NCBI database with human CREBBP as a query. Species were selected considering phylogenetic position, completeness, and quality of available sequences. Genbank accession numbers for all sequences used are listed in Table 1.

Molecular evolutionary analyses

The phylogenetic tree of the species (Figure 1) used as a guide tree was generated using an iTOL web server [10] and derived from previous study [11]. A PRANK program was applied for multiple sequence alignment for the coding sequences of the 19 mammals [9]. The PRANK was executed with empirical codon model. A CODEML program in PAML 4.5 was used to estimate the rates of synonymous (*dS*) and nonsynonymous substitutions(*dN*) and the *dN*/*dS* ratio (ω) [12]. A *dN/dS* ratio = 1, <1, and >1 indicates neutral evolution, purifying selection, and positive selection at the protein, respectively. The codon frequencies were calculated from the average nucleotide frequencies at the third codon positions (setting CodonFreq = 2 (F3X4)).

One-ratio model (M0 model), which allows only a single *dN*/*dS* ratio for all branches, was used to estimate the general selective pressure among all species. Free-ratio model was used to analyze the *dN*/*dS* ratio along each branch. To further examine potential positive selection, branch-site test was conducted [13]. When examine positive selection of a cetacean branch, remaining lineages were denominated as background branches. It was assumed that the background branches share the same distribution of ω among sites, whereas different values can be applied to the foreground branch. CREBBP genes from four cetacean species were included in the foreground branch. The likelihood of alternative model (fix_omega = 0) and null model (fix_omega = 1) were estimated using CODEML, and then likelihood ratio test (LTR) was conducted. The Bayes Empirical Bayes (BEB) approach was used to calculate the posterior probability of a specific codon site and to identify those most likely to be under positive selection.

Table 1. List of species and protein IDs used in this study

Order	Suborder	Species			Protein ID
		Family	Scientific name	Common name	
Cetacea	Odontoceti	Physeteridae	*Physeter macrocephalus*	Sperm whale	XP_007112741.1
		Delphinidae	*Orcinus orca*	Killer whale	XP_004270285.1
		Lipotidae	*Lipotes vexillifer*	Yangtze river dolphin	XP_007453543.1
	Mysticeti	Balaenopteridae	*Balaenoptera acutorostrata*	Minke whale	XP_007181605.1
Artiodactyla		Bovidae	*Bos taurus*	Cow	NP_001157494.1
		Camelidae	*Vicugna pacos*	Alpaca	XP_006204379.1
Perissodactyla		Equidae	*Equus caballus*	Horse	XP_001499399.3
Carnivora		Canidae	*Canis lupus familiaris*	Dog	XP_003434912.1
		Mustelidae	*Mustela putorius furo*	Ferret	XP_004750593.1
Chiroptera		Vespertilionidae	*Myotis lucifugus*	Microbat	XP_006090495.1
Rodentia		Muridae	*Mus musculus*	Mouse	NP_001020603.1
			Rattus norvegicus	Rat	NP_596872.3
		Bathyergidae	*Heterocephalus glaber*	Naked mole rat	XP_004864823.1
		Caviidae	*Cavia porcellus*	Guinea pig	XP_003477760.1
Primates		Hominidae	*Homo sapiens*	Human	NP_004371.2
			Pan troglodytes	Chimpanzee	XP_523285.2
		Cercopithecidae	*Macaca mulatta*	Macaque	XP_001095225.1
		Hylobatidae	*Nomascus leucogenys*	Gibbon	XP_003269267.1
		Galagidae	*Otolemur garnettii*	Bushbaby	XP_003790880.1

Prediction of functional effects of cetacean-specific amino acid changes

Sequences for all data sets were aligned using CLUSTAL omega to show cetacean-specific amino acid changes [14]. To predict functional outcome of these alterations on the bases of sequence and structure, PolyPhen-2 (Polymorphism Phenotyping v2) analysis was employed [15]. As the program was initially devised for detecting mutations in human genes, the functional effect of each variation is classified as “probably damaging”, “possibly damaging”, or “benign”.

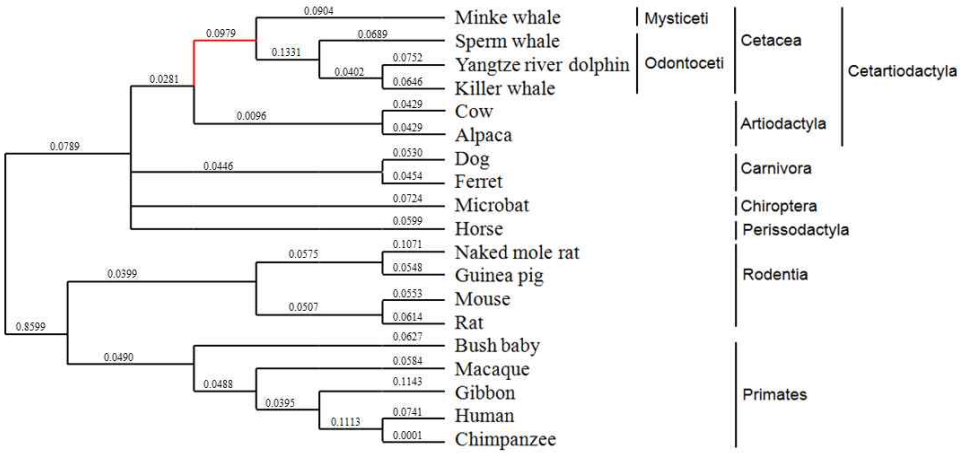

Fig. 1. The ω values of CREBBP genes in distinct evolutionary lineages of cetaceans and other mammals using a phylogenetic tree derived from Liang et al. [11]. The cetacean branch (red colored) was used as foreground branch shown in Table 2. The ω values of individual branches shown are based on the free-ratio model.

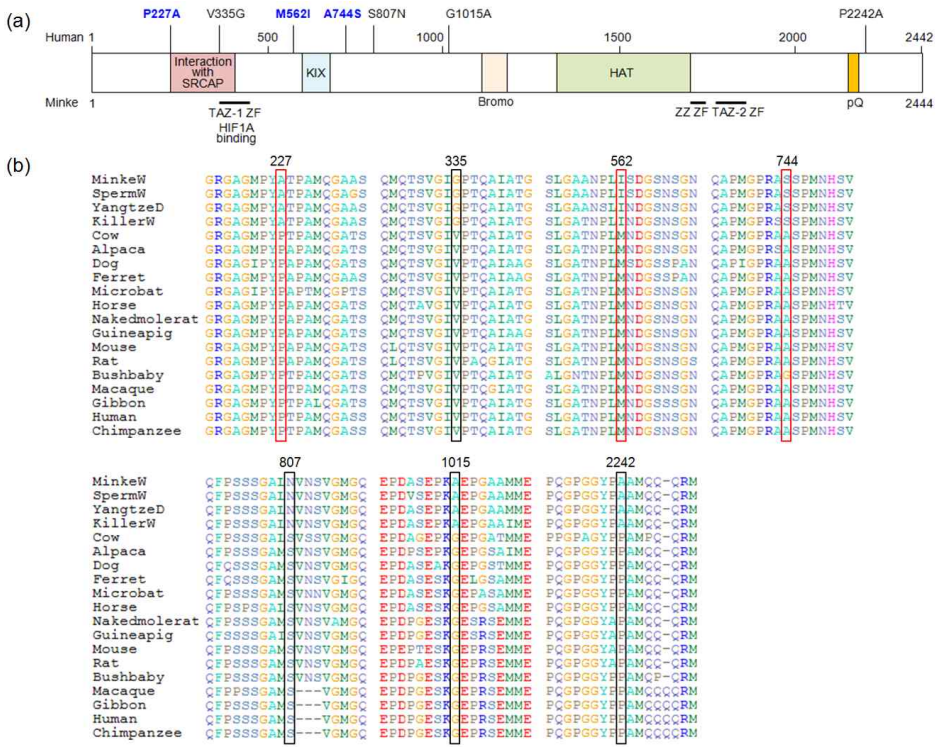

Fig. 2. Structure and amino acid changes of CREBBP. (a) Schematic representation of domains and amino acid changes in CREBBP. Amino acid changes predicted to have functional effects on CREBBP were marked in blue. (B) Protein sequence alignment of four cetacean and 15 mammalian CREBBPs. Cetacean-specific amino acid changes were marked in boxes.

Results and Discussion

Molecular evolution of CREBBP

We obtained the complete coding sequences of CREBBP from four cetaceans representing three families of the suborder Odontoceti (Sperm whale, Killer whale, Yangtze river dolphin) and one family of the suborder Mysticeti (minke whale). They all

show typical features of CREBBP, including the domain that interact with SRCAP (Snf2-related CREBBP activator protein), the CREB-binding domain KIX (kinase-inducible domain interacting), a Bromodomain that recognize monoacetylated lysine, Histone acetyltransferase (HAT) domain, C-terminal poly-glutamine track, and two TAZ-type and one ZZ-type zinc fingers (Fig. 2A).

To access the evolutionary pressure on the CREBBP gene, branch and branch-site models were performed with the likelihood framework, using the species tree shown in Fig. 1 as the working topology. One-ratio model analyses of all mammals (19 sequences) showed that all of them shared the same estimated ω of 0.05917 (Table 2), which indicating the existence of strong functional constraints on mammalian CREBBP. The free ratio model analysis also showed similar ω value (0.0979), which is no evidence of positive selection. Analysis using branch-site model also showed that CREBBP is highly conserved in mammals with no positively selected site detected in the cetacean lineage. However, statistical methods based on a comparison of the *dN/dS* value may not suitable for analyzing conservative proteins since they do not represent the possibility that adaptation may result from very few amino acid changes [16].

Table 2. Estimate of the selective pressure on the CREBBP gene of the order cetacean

Models		Parameter estimate	lnL[a]	2ΔL[b] (*P*-value)	Positively selected sites
Branch model	One-ratio model (M0)	ω=0.05917	-30067.4439		
	Free-ratio model (M1)	ω=0.0979	-30042.4992	49.89 (P^c < 0.0001)	
Branch site models	Null	P_0[d]=0.95583, P_1[e]=0.04380, P_{2a}[f]=0.00035, P_{2b}[g]=0.00002, ω_0=0.03299, ω_1=1.0	-29269.5		
	Alternative	P_0[d]=0.95583, P_1[e]=0.04380, P_{2a}[f]=0.00035, P_{2b}[g]=0.00002, ω_0=0.03299, ω_1=1.0, ω_2=1.0	-29269.5	0 (P^c = 1)	None

[a]lnL is the log-likelihood score; [b]Likelihood ratio test (LRT) to detect positive selection;

[c]P-values are calculated for lnL values of two models (One-ratio and Free-ratio models, Null and Alternative models) using the likelihood ratio test.; [d]Proportion of sites that are under purifying selection (ω_0 < 1) on both background and foreground branches; [e]Proportion of sites that are under neutral evolution (ω_1 = 1) on both background and foreground branches; [f]Proportion of sites that are under purifying selection (ω_0 < 1) on background branches and under positive selection ($\omega_2 \geq$ 1) on the foreground branch; [g]Proportion of sites that are under neutral evolution (ω_1 = 1) on background branches and under positive selection ($\omega_2 \geq$ 1) on the foreground branch.

Prediction of the functional effects of cetacean-specific amino acid changes in CREBBP

To provide insight into the unique cetacean phenotypes, we investigated cetacean-specific amino acid changes in CREBBP using omega (Fig. 2B and data not shown). A total of 7 amino acid changes unique to cetacean CREBBPs were identified, and 6 of them were located in the N-terminal half of the protein (Fig. 2A). The effects of these changes on the function of the protein were predicted by computational analysis (PolyPhen2) [15], resulting in two sites as "possibly damaging" and one site as "probably damaging" (Table 3). Interestingly, Pro227Ala, the "possibly damaging" site, was localized in the beginning of the interaction domain with SRCAP, and the other two "damaging sites", Met562Ile and Ala744Ser, were located in adjacent to KIX domain (Fig. 1A). Considering that mutations in SRCAP cause Floating-Harbor syndrome characterized by short stature and facial abnormalities [17] similar to the defects observed with CREBBP mutations, cetacean-specific amino acid change within this domain might have functional effects on the craniofacial adaptations in cetaceans. KIX domain of CREBBP was initially identified as the specific and minimal region that was sufficient to bind and interact with phosphorylated CREB and then activate transcription [18]. KIX has been also shown to interact with many other transcription factors such as c-Myb, mixed lineage leukemia protein (MLL), c-Jun, p53, breast cancer 1 (BRCA1), signal transducers and activators of transcription (STAT1), and sterol responsive element-binding protein (SREBP) [18]. C-Myb has essential role in the regulation of haematopoiesis [19] and mutations in MLL result in facial defects and short statue [20]. Since the KIX domain is a highly conserved, independently folding region, any change within this domain would have been detrimental to the function of CREBBP. Instead, the cetacean-specific amino acid changes predicted as damaging in proximity to the KIX domain might have contributed the alteration of the binding specificity or affinity of cetacean CREBBP to its targets without affecting its intrinsic activity.

It is worth mentioning that a cetacean-specific change, Val335Gly, is in close proximity to the TAZ-1 zinc finger domain (amino acid 347-433) and Gly1015Ala is next to the conserved lysine residue that is modified by acetylation. TAZ-1 domain is known to be responsible for hypoxia-induced HIF1A binding, the master regulator of hypoxia response in all metazoan species [21]. Although not predicted as damaging, it is intriguing to see if this change has any functional significance in the alteration of hypoxia-mediated responses in cetaceans.

Table 3. Unique amino acid changes in Cetacean CREBBP analyzed by PolyPhen 2.

Position (human)	Human	Cetacea	PolyPhen2	
			Prediction	Score
227	P	A	Possibly damaging	0.524
335	V	G	Benign	0.110
562	M	I	Possibly damaging	0.756
744	A	S	Probably damaging	0.990
807	S	N	Benign	0.226
1015	G	A	Benign	0.043
2242	P	A	Benign	0.018

Conclusion

The aquatic adaptation of the cetaceans from their terrestrial relatives is one of the most remarkable events throughout mammalian evolution. Analysis of the selective pressure on CREBBP gene using the non-synonymous to synonymous rate ratio ω demonstrated that CREBBP has been under purifying selection. Protein sequence alignment showed seven cetacean-specific amino acid changes in highly conserved sites, and three of them were predicted to have functional effects on cetacean CREBBP. Considering the localization of these changes within or in the vicinity of important domains, there are possibilities that these molecular adaptation of CREBBP might have influenced the morphological and physiological changes of cetacean such as craniofacial alteration and haematopoiesis during their transition from land to ocean.

Acknowledgement

This study was supported by an in-house program (PE99212) from KIOST (Korea Institute of Ocean Science & Technology).

References

1. Thewissen, J. G., Cooper, L. N., Clementz, M. T., Bajpai, S., and Tiwari, B. N. 2007. Whales originated from aquatic artiodactyls in the Eocene epoch of India. *Nature.* **450**, 1190–1194.
2. Uhen, M. D. 2007. Evolution of marine mammals: back to the sea after 300 million years. *Anat. Rec.* (Hoboken). **290**, 514-522.
3. Reidenberg, J. S. 2007. Anatomical adaptations of aquatic mammals. *Anat. Rec.* (Hoboken). **290**, 507-513.
4. Kooyman, G. L. 2009. Diving physiology. *Encyclopedia of marine mammals. 2nd Ed.* 327-332.
5. Blobel, G. A. 2002. CBP and p300: versatile coregulators with important roles in hematopoietic gene expression. *J. Leukoc. Biol.* **71**, 545-556.
6. Oike Y., Takakura, N., Hata, A. et al., 1999. Mice homozygous for a truncated form of CREB-binding protein exhibit defects in hematopoiesis and vasculo-angiogenesis. *Blood.* **93**, 2772-2779.
7. Petrij, F., Giles, R. H., Dauwerse, H. G. et al. 1995. Rubinstein-Taybi syndrome caused by mutations in the transcriptional co-activator CBP. *Nature.* **376**, 348-351.
8. Yim, H. S., Cho, Y. S., Guang, X. et al., 2014. Minke whale genome and aquatic adaptation in cetaceans. *Nat. Genet.* **46**, 88-92.
9. Loytynoja, A. and Goldman, N. 2005. An algorithm for progressive multiple alignment of sequences with insertions. *Proc. Natl. Acad. Sci.* **30**, 10557-10562.
10. Letunic, I. and Bork, P. 2006. Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics* **23**, 127-128.
11. Liang, L., Shen, Y. Y., Pan, X. W. et al. 2013. Adaptive Evolution of the Hox Gene Family for Development in Bats and Dolphins. *PLoS ONE* **8(6)**, e65944.
12. Yang, Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24**, 1586–1591.
13. Zhang, J., Nielsen. R. and Yang, Z. 2005. Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. *Mol. Biol. Evol.* **22**, 2472–2479.
14. Sievers, F. and Higgins, D. G. 2014. Clustal Omega, accurate alignment of very large numbers of sequences. *Methods Mol. Biol.* **1079**, 105-116.
15. Adzhubei, I. A., Schmidt, S., Peshkin, L. et al. 2010. A method and server for predicting damaging missense mutations. *Nat. Methods* **7**, 248-249.
16. Yang, Z. and Bielawski, J. P. 2000. Statistical methods for detecting molecular adaptation. *Trends Ecol. Evol.* **15**, 496-503.
17. Hood, R. L., Lines, M. A., Nikkel, S. M. et al. 2012. Mutations in SRCAP, encoding SNF2-related

CREBBP activator protein, cause Floating-Harbor syndrome. *Am. J. Hum. Genet.* **90**, 308-313.

18. Thakur, J.K., Yadav, A. and Yadav, G. 2014. Molecular recognition by the KIX domain and its role in gene regulation. *Nucleic Acids Res.* **42**, 2112-2125.
19. Mucenski, M. L., McLain, K., Kier, A.B. et al. 1991. A functional c-myb gene is required for normal murine fetal hepatic hematopoiesis. *Cell.* **65**, 677-689.
20. Yu, B. D., Hess, J. L., Horning, S. E. et al. 1995. Altered Hox expression and segmental identity in Mll-mutant mice. *Nature.* **378**, 505-508.
21. Semenza, G. L. 2012. Hypoxia-inducible factors in physiology and medicine. *Cell.* **148**, 399-408.