RESEARCH ARTICLE Indian Journal of Animal Research, Volume 56 Issue 3: 263-269 (March 2022)

The Analysis of Reproduction-related Genes Provides Insights into the Adaptive Evolution of Fecundity Traits in Yangtze Finless Porpoise

Wu Bin[1], Wang Weiping[2], Zhang Huan[3]

10.18805/IJAR.BF-1418

ABSTRACT

Background: Yangtze finless porpoise (YFP, *Neophocaena asiaeorientalis*), is the first class protected animal in China. In order to analyze the adaptive evolution of fecundity traits in YFP, the rapidly evolving gene families of YFP were obtained. At the same time, the major genes controlling ovulation, GDF9, BMP15, FSHβ and FSHR were also analyzed.

Methods: Orthofinder software was employed to search homologous genes based on protein sequence. CAFÉ software was used to obtain the expansion and contraction gene families of YFP. Then, GO terms and pathway enrichment analyses were performed using TBtools software and Swissprot database. PAML package was used to calculate Ka/Ks (*i.e.*, ω). Evolution rate changes in the positive selected genes were examined using the GU99 process in Diverge (v3.0) program.

Result: In YFP, 501 rapid expansion gene families GO enrichment results showed that the reproductive activities related pathways were mainly three significant enrichment process, participation of germline stem cells maintain androgen receptors signaling pathways regulating and male reproductive tract stem cell population to maintain. The most significant GO terms of 220 rapidly contraction gene families associated with reproductive activities mainly consisted of biological processes which were involved in positive regulation of estrogen secretion, mating and estrogen metabolic process. GDF9 and BMP15 genes exhibited purifying selection. However, significant signs of positive selection were detected in FSHR and FSHβ genes, but only FSHβ showed specific changes in the YFP lineage.

Key words: Fecundity traits, *Neophocaena asiaeorientalis*, Ovulation, Positive selection, Rapidly evolving.

INTRODUCTION

Models are often used in the study of population ecology (Mei, 2013; Li, 2017; Wu *et al*. 2018, 2020, 2021), but they often ignore some factors, such as adaptation to reproductive characteristics. Cetaceans (whales, dolphins and porpoises) are a group of mammals adapted to various aquatic habitats, living from ocean to fresh water (Zhou *et al*., 2018). Having evolved from terrestrial ancestors to occupy aquatic niches (Thewissen *et al*., 2007), these mammals, especially their reproductive adaptations fascinate scientists and the public worldwide. For close relatives of the cetacean, fecundity traits are important economic characteristics of cattle and goat which is directly related to the economic benefits of the industry (Lei, 2002). Yangtze finless porpoise (YFP, *Neophocaena asiaeorientalis*) is currently the only cetacean species in Yangtze River (Gao *et al*., 1995). YFP's mating system is polygyny (Hao *et al*., 2006). The mature age of female is 4-6 and that of male is 4.5-7 (Hao *et al*., 2006). The life span of YFP is 20 to 25 years (Zhang *et al*., 1999; Hao *et al*., 2006). Two individuals aged 21 and 23 of finless porpoise (*Nephocaena phocaenoides*) could conceive and breastfeed (Shirakihara *et al*., 1993).

The cetacean oxidative stress-related gene families, such as the peroxiredoxin (PRDX) family, glutathion peroxidase (GPx) and the OGT gene family, had been found to have expanded and it was an adaptive mode for Cetacean tolerance to hypoxia stress. But, neuronal differentiation-related gene families showed contraction, suggesting that they might be associate with smaller brains of porpoises (Sun, 2018). However, there have been no reports on the rapid evolution gene families related to reproductive activity in aquatic mammals such as YFP. And more and more studies have found that many hormones, cytokines and adhesion factors are linked with ovulation through endocrine activities and play a very important role in follicular development and ovulation process, such as: Luteotropic hormone LH), follicle stimulating hormone (FSHβ) (Chu *et al*., 2012; Miyano *et al*., 2014) and follicle stimulating hormone receptors (FSHR) (Lazaros *et al*., 2012), growth and

[1]Jiangsu Key Laboratory for Biodiversity and Biotechnology, College of Life Sciences, Nanjing Normal University, Nanjing 210023, China.
[2]Aquatic Biology Protection and Rescue Center of Jiangxi Province, Nanchang 330000, China.
[3]School of Life Sciences, Nanchang University, Nanchang 330031, China.

Corresponding Author: Zhang Huan, School of Life Sciences, Nanchang University, Nanchang 330031, China.
Email: huanzhang@ncu.edu.cn

How to cite this article: Bin, W., Weiping, W. and Huan, Z. (2022). The Analysis of Reproduction-related Genes Provides Insights into the Adaptive Evolution of Fecundity Traits in Yangtze Finless Porpoise. Indian Journal of Animal Research. 56(3): 263-269. DOI: 10.18805/IJAR.BF-1418.

Submitted: 05-08-2021 **Accepted:** 05-11-2021 **Online:** 02-12-2021

differentiation factors-9 (Growth different-I ation factor-9, GDF-9), BMP 15 (Bone morphogenetic protein-15, BMP-15) (Otsuka *et al*., 2014; Migaud *et al*., 2013). Therefore, in order to analysis of the adaptive evolution of fecundity traits in YFP, the present study was performed to rapidly evolving gene families of YFP. At the same time, positive selection test and protein functional differentiation test of GDF9, BMP15, FSHβ and FSHR genes related to the hypothalamo-pituitary-gonadal axis system were carried out to find the specific mutation sites of YFP, thus providing new insights into the adaptive evolution of fecundity traits in order to facilitate current and future breeding protection.

MATERIALS AND METHODS

Data availability

The assembled genome sets are available from NCBI, *Bos taurus* (GCF_002263795.1), *Felis catus* (GCF_000181335.3), *Cavia porcellus* (GCF_000151735.1), *Heterocephalus glaber* (GCF_000247695.1), *Chinchilla lanigera* (GCF_000276665.1), *Castor canadensis* (GCF_001984765.1), *Jaculus jaculus* (GCF_000280705.1), *Nannospalax galili* (GCF_000622305.1), *Camelus bactrianus* (GCF_000767855.1), *Ovis aries* (GCF_002742125.1), *Sus scrofa* (GCF_000003025.6), *Homo sapiens* (GCF_000001405.39), *Mus musculus* (GCF_000001635.27), *Lipotes vexillifer* (GCF_000442215.1), *Phocoena sinus* (GCF_008692025.1), *Neophocaena asiaeorientalis* (GCF_003031525.1).

Orthologous family identification

To define a set of conserved genes for cross-taxa comparison, we employed Orthofinder (v2.3.3) to search homologous genes of 16 species based on protein sequence (Emms *et al*., 2015; 2019). Toolbox for Biologists (TBtools, v1.096) software was used to obtain whole-genome representative CDS sequences from the genome and then representative protein sequences were extracted for subsequent analysis according to the corresponding relationship between the CDS sequence number and the protein sequence number (Chen *et al*., 2020). To identify expanding and contracting gene ortholog groups across the phylogeny. The species tree was obtained from the Orthofinder (v2.3.3) (Emms *et al*., 2017; 2018) and the time of differentiation was obtained from the website Timetree (Kumar *et al.,* 2017). We estimated the gene numbers on internal branches using a random birth and death process model implemented in the software CAFÉ (v3.0), a tool for the statistical analysis of the evolution of the size of gene families (De *et al*., 2006).

GO enrichment and statistical analyses

After the expansion and contraction gene families of the target species were obtained by CAFÉ v3.0 software, the expansion and contraction gene sets were further obtained according to the results of Orthofinderv 2.3.3 software. Then, GO terms and pathway enrichment analyses were performed using TBtools and Swissprot with the default parameters. And the significantly enriched clusters associated with reproductive activities were reported (Benjamini-Hochberg corrected q<.05). Visualization of results were carried out using the R package 'ggplot2' (Hadley *et al.,* 2021).

Positive selection test

PAML package were used to calculate Ka (nonsynonymous substitution rate of nonsynonymous sites), Ks (synonymous substitution rate of synonymous sites) and Ka/Ks (*i.e.*, ω) between the homologous GDF9, BMP15, FSHβ and FSHR genes of 16 species (Yang, 2007). In addition, the ω value between paternal gene and retrogene pair was used to estimate functional constraints with $\omega < 1$ representing the purifying selection, $\omega = 1$ for the neutral selection and $\omega > 1$ for the positive selection (Yang, 2007). Sequence alignment was performed using MEGA X software (Kumar *et al*., 2018).

Protein functional differentiation test

Evolution rate changes in the positive selected genes were examined using the GU99 process in Diverge V. 3.0 program to predict amino acid sites associated with functional differentiation (Gu *et al*., 2013). θ indicates the likelihood of functional differentiation between two sets of genes (Thompson *et al*., 2018). In fact, GDF9 and BMP15 genes exhibited purifying selection. For FSHR gene, no functional differentiation was found between the clusters of 16 species. Only FSHβ gene needs to be further expanded for protein functional differentiation test. For FSHβ gene, the above gene family analysis showed that YFP and *Phocoena sinus* had two homologous genes. By comparing these four sequences, it was found that they could be divided into two groups, which constituted one of the YFP and one of *Phocoena sinus* (XP_024620387.1 and XP_032497072.1) and the other two constituted another group (XP_024620395.1 and XP_032496507.1). Therefore, protein sequence sets A and B were respectively constructed by the two distinct sequences groups described above and another 32 common sequences downloaded from the FSHβ gene NCBI Orthologs, as shown in Table 1.

RESULTS AND DISCUSSION

Rapidly evolving gene families

In 16 species, orthogroups number of genes was 324580 and percentage of genes in orthogroups was 97.8%; number of orthogroups was 19498; number of species-specific orthogroups was 752; number of genes in species-specific orthogroups was 4002; mean orthogroup size was 16.3; median orthogroup size was 16; number of single-copy orthogroups was 8809. In YFP, 501 gene families showed rapidly expansion and 220 gene families showed rapidly contraction, which was consistent with the number of expanding and contracting gene families as shown in Fig 1. For YFP, 860 genes were obtained; 270 genes were lost; 18777 genes were no change; in general, YFP showed genes expansion.

Table 1: Protein sequence sets of FSHβ genes from 34 species.

Cetacea			Primates		
Neophocaena asiaeorientalis; *Phocoena sinus*	A set XP_024620387.1 XP_032497072.1	*Homo sapiens*	NP_001369218.1	*Macaca nemestrina*	XP_011722257.1
	B set XP_024620395.1 XP_032496507.1	*Pan paniscus*	XP_003830474.1	*Cercocebus atys*	XP_011917459.1
Orcinus orca	XP_004263999.1	*Papio anubis*	XP_009184656.1	*Rhinopithecus bieti*	XP_017708175.1
Delphinapterus leucas	XP_022439434.1	*Gorilla gorilla*	XP_004050915.3	*Piliocolobus tephrosceles*	XP_023086149.1
Lagenorhynchus obliquidens	XP_026970928.1	*Chlorocebus sabaeus*	XP_008001624.1	*Theropithecus gelada*	XP_025214147.1
Globicephala melas	XP_030720683.1	*Rhinopithecus roxellana*	XP_010363997.1		
Lipotes vexillifer	XP_007460994.1	*Macaca mulatta*	XP_001087914.1		
Cetacean distant species		**Cetacean relatives**			
Cavia porcellus	NP_001166419.1	*Trichechus manatus latirostris*	XP_004369799.1	*Sus scrofa*	NP_999040.1
Heterocephalus glaber	XP_004851851.1	*Zalophus californianus*	XP_027437293.1	*Felis catus*	XP_003993171.1
Jaculus jaculus	XP_004660900.1	*Panthera tigris altaica*	XP_007082072.1	*Bubalus bubalis*	NP_001277871.1
Nannospalax galili	XP_008849438.1	*Oryctolagus cuniculus*	NP_001075640.1	*Camelus bactrianus*	XP_010963037.1
Mus musculus	NP_032071.1	*Canis lupus familiaris*	XP_542546.3	*Ovis aries*	NP_001009798.1

Functional enrichment analysis of 1158 genes (genes of rapidly expansion 501 gene families) was executed. Only clusters associated with reproductive activities enrichment with $p<0.05$ were listed (Fig 2). The most significant GO terms mainly consisted of biological processes which were involved in germ-line stem cell population maintenance, regulation of androgen receptor signaling pathway and male germ-line stem cell population maintenance. This could be a genetic adaptation for YFP to remain fertile when it was old. According to the conversion of body length and age, the females of YFP in Poyang Lake may still be able to reproduce at the age of 18-19. YFP, meanwhile, had 220 gene families showed rapidly contraction.

Functional enrichment analysis of 726 genes (220 gene families of YFP rapidly contraction and the gene set was the corresponding *Mus musculus* genes) was executed. Only clusters associated with reproductive activities enrichment with $p<0.05$ were listed (Fig 3). The most significant GO terms mainly consisted of biological processes which were involved in positive regulation of estrogen secretion, mating and estrogen metabolic process. Ireland *et al*. believed that estrogen was mainly secreted by dominant follicles on the ovary (Ireland *et al*., 1984) and the conception rate of female animals might be affected by the concentration of estrogen secreted by dominant follicles (Kiewisz *et al*., 2011). This might be related to YFP's polygyny mating system (Hao *et al*., 2006), the fact that no YFP gave birth to more than 1 baby per birth (Li, 2017) and the pregnancy rate of YFP in Poyang Lake was as high as 70% (Mei, 2013).

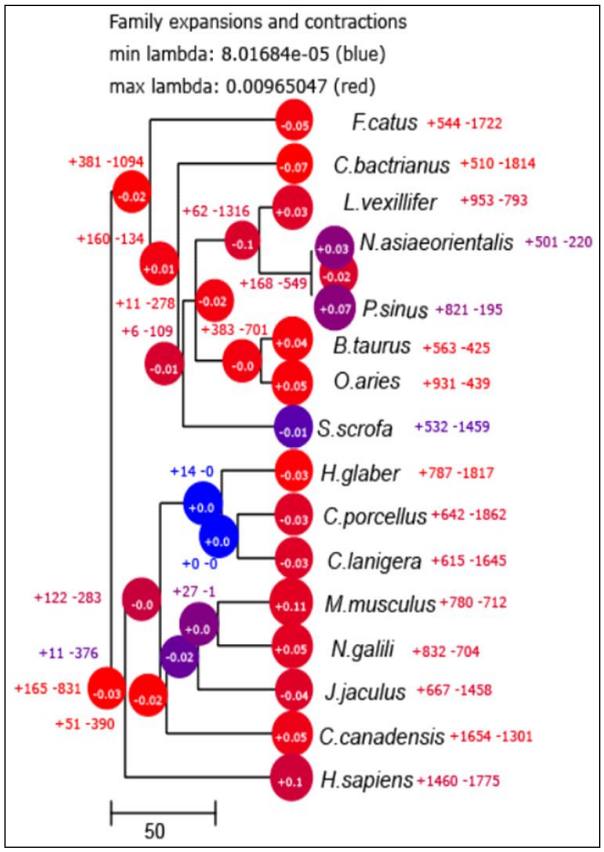

Fig 1: Orthologous family analysis.

Positive selection genes and protein functional differentiation

According to the branch site model detection, GDF9 and BMP15 genes exhibited purifying selection. For FSHR gene, 11 positive selection sites with a posterior probability greater than 0.5 with experience were found. For FSHβ gene, YFP, *Phocoena sinus* and *Bos taurus* all had two homologous genes. Therefore, FSHβ genes could form 8 sequence combinations. No positive selection sites were found in one combination and positive selection sites were found in the other seven combinations. Only FSHβ gene needs to be further expanded for protein functional differentiation test, seeing Materials and Methods for detailed description. For B FSHβ gene set, no functional differentiation was found between the clusters. For A FSHβ gene set, there were 1-4 functional differentiation sites in cetaceans, as shown in Table 2. Further analysis revealed that the L at position 40 is a specific site for the YFP, Phocoena sinus and baiji, as shown in Fig 4-5. The FSHβ protein sequence of YFP and *Phocoena sinus* had fragment insertions, as shown in Fig 6. YFP, *Phocoena sinus* and *Bos taurus* all had two homologous genes for FSHβ, but only FSHβ protein sequence of YFP and *Phocoena sinus* had fragment insertions. Further analysis revealed that the L at position 40 was a specific site for the YFP, *Phocoena sinus* and baiji.

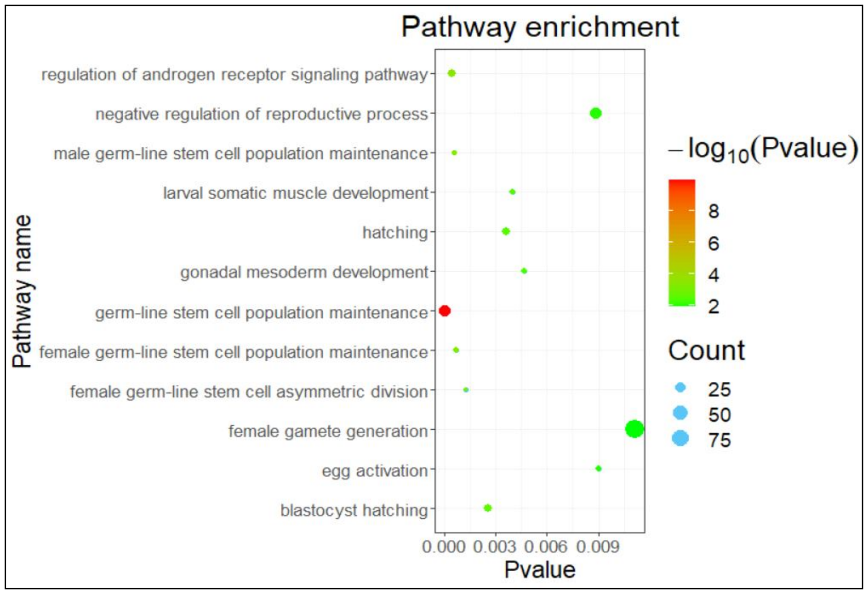

Fig 2: GO and pathway enrichment analysis of rapidly expansion genes.
(Only clusters associated with reproductive activities enrichment with p<.05 were listed).

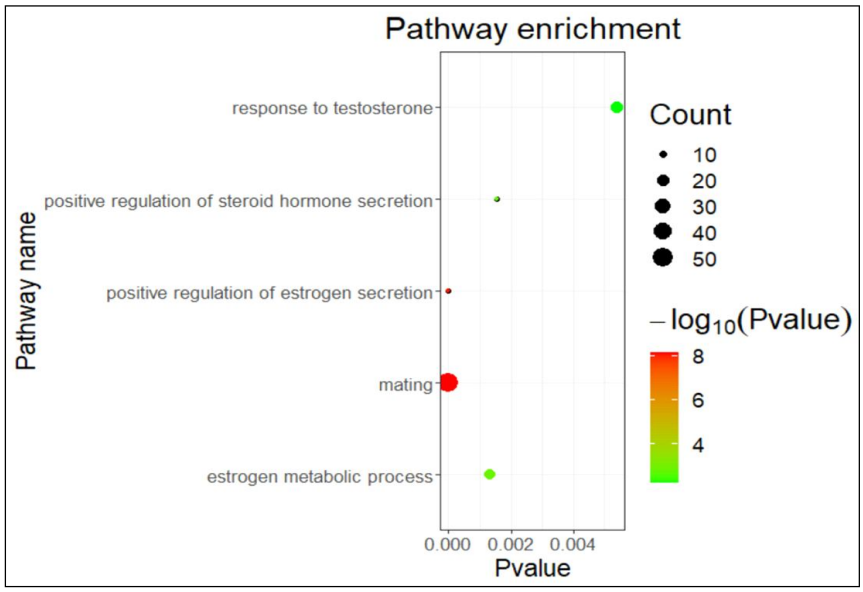

Fig 3: GO and pathway enrichment analysis of rapidly contraction genes
(Only clusters associated with reproductive activities enrichment with p<.05 were listed).

The relationship between genes of GDF9, BMP15, FSHβ, FSHR and ovulation number is the main object of research on multiple fertility traits in cattle and sheep. Similar to previous studies examining positive selection, the majority of genes investigated in this study were under purifying selection, which is not surprising given the evolutionary constraint on protein coding genes. However, signs of significant positive selection were detected in 2 genes (FSHR and FSHβ). This suggests that FSHR and FSHβ evolved fast along all YFP lineages examined and their positive selection might be associated with fecundity traits. The mutation rate of FSHR gene exon 10 was significantly different between twin and single bovine cattle (Lei *et al*., 2004). It was inferred that FSHR gene exon 10 was related with the major gene that controls the high prolificacy of goat and the individuals with FSHR gene mutation showed high lambing number, indicating that the high expression of FSHβ in the ovary promoted the high ovulation number, which was the result of the mutual regulation between FSHR and FSHβ gene (Ji, 2007).

Table 2: Functional divergence estimated in FSHβ gene in different clusters of mammals.

Clusters	θa	SEb	LRTc	N (0.9) d
Cetacea/primates	-0.025245	0.023681	23.862782	4
Cetacea/cetacean distant species	-0.074536	0.033527	nan	1
Cetacea/cetacean relatives	-0.098359	0.031419	5.265424	0

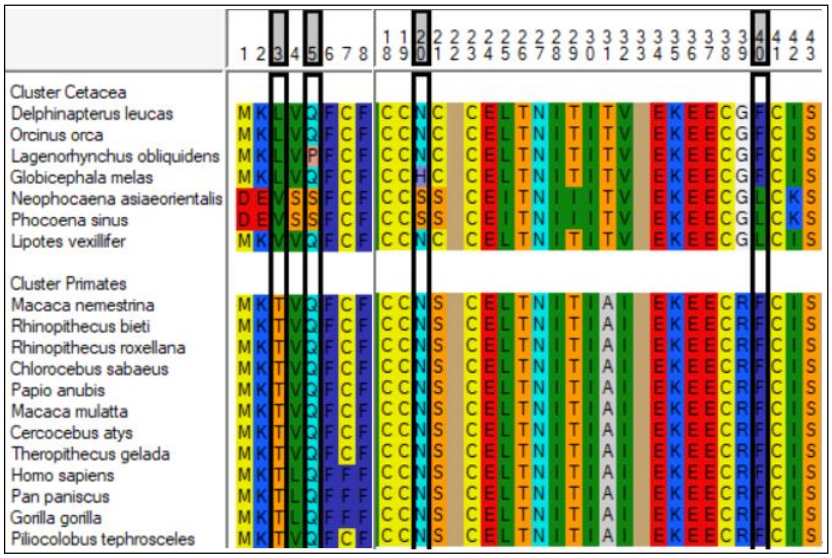

Fig 4: Functional divergence estimated in FSHβ gene in Cetacea/Primates.

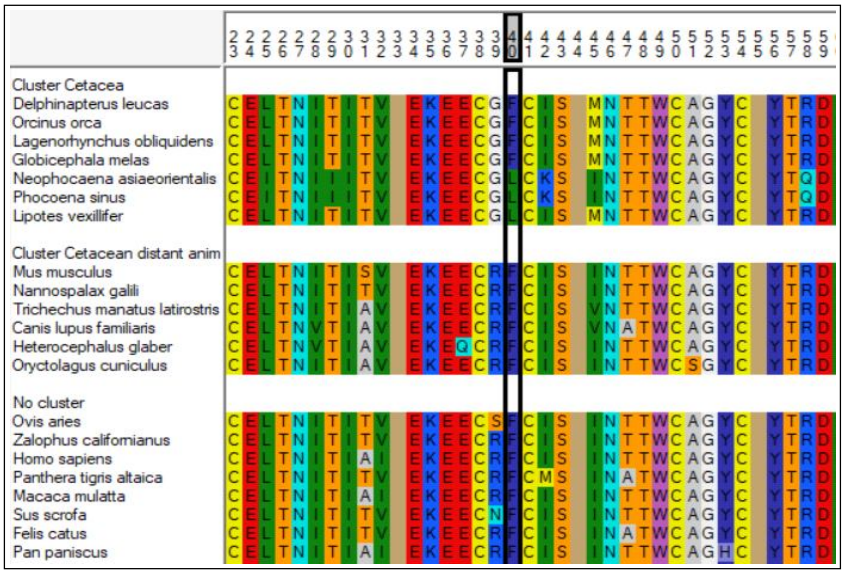

Fig 5: Functional divergence estimated in FSHβ gene in Cetacea/ Cetacean distant species.

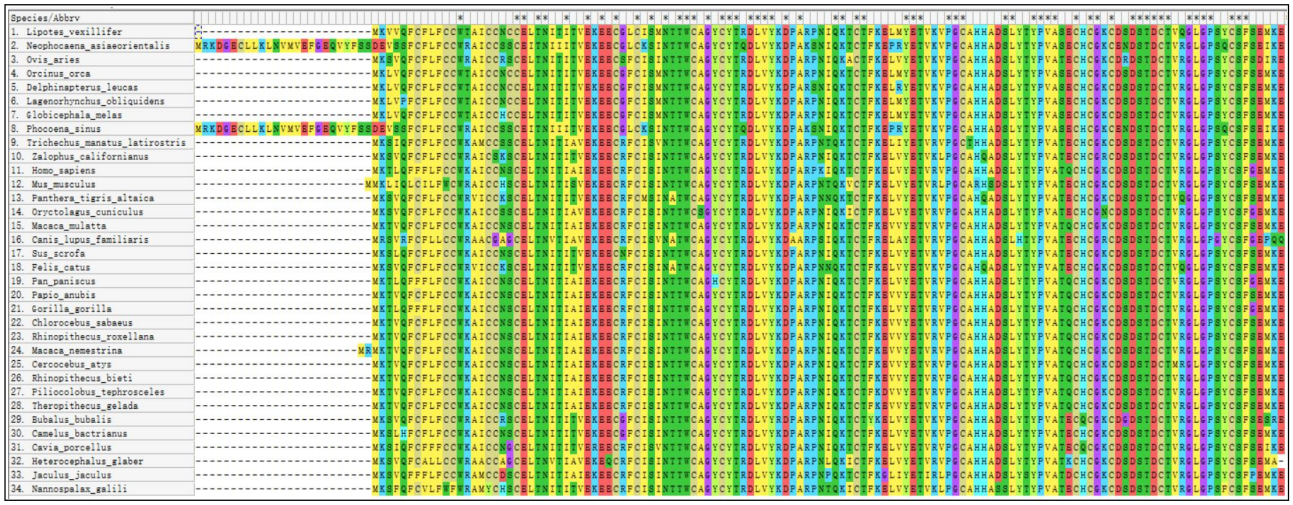

Fig 6: Protein alignment of FSHβ in 34 mammalian species.

CONCLUSION

Results from rapidly evolving gene families of YFP showed that there might be trade-offs between longer reproductive life, higher pregnancy rates and “only one baby born per birth”. Even though most of the genes were under purifying selection, FSHβ gene in YFP lineages exhibited specific changes, which likely contributed to their fecundity traits. In future, longitudinal studies of cetaceans are needed to explore the breeding-associated genes and pathways and then *in vitro* functional assays are required to confirm their roles in YFP fecundity traits analyses. Construction of FSHβ and other genes mutants in model animals can be to test the relationship between genes and “only one baby born per birth” of YFP.

REFERENCES

Chen, C.J., Chen, H. and Zhang, Y., *et al.* (2020). TBtools: An integrative toolkit developed for interactive analyses of big biological data. Molecular Plant. 13(8): 1194-1202.

Chen, M.M., Zheng, Y. and Hao, Y.J., *et al.* (2016). Parentage-based group composition and dispersal pattern studies of the Yangtze Finless Porpoise Population in Poyang Lake. International Journal of Molecular Sciences. 17: 1268. doi: 10.3390/ijms17081268.

Chu, C., Zhou, J.A., Zhao, Y., *et al.* (2012). Expression of FSH and its co-localization with FSH receptor and GnRH receptor in rat cerebellar cortex. Journal of Molecular Histology. 44(1): 19-26.

De, B.T., Cristianini, N. and Demuth, J.P. *et al.* (2006). CAFE: A computational tool for the study of gene family evolution. Bioinformatics. 22(10): 1269-1271.

Emms, D.M. and Kelly, S. (2019). OrthoFinder: Phylogenetic orthology inference for comparative genomics, Genome Biology. 20: 238. https://github.com/davidemms/OrthoFinder.

Emms, D.M. and Kelly, S. (2018). STAG: Species Tree Inference from All Genes. bioRxiv https://doi.org/10.1101/267914.

Emms, D.M. and Kelly, S. (2017). STRIDE: Species tree root inference from gene duplication events Mol. Biol. Evol. 34(12): 3267-3278.

Emms, D.M. and Kelly, S. (2015). OrthoFinder: Solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. Genome Biology. 16: 157.

Gu, X., Zou, Y. and Su, Z., *et al.* (2013). An update of DIVERGE software for functional divergence analysis of protein family. Mol. Biol. Evol. 30(7): 1713-1719.

Gao, A. and Zhou, K. (1995). Geographical variation of external measurements and three subspecies of *Neophocaena phocaenoides* in Chinese waters. Acta Theriol. Sin. 15: 81-92.

Hadley, W., Winston, C. and Lionel, H. *et al.* (2021). Package ‘ggplot2’. Create Elegant Data Visualisations Using the Grammar of Graphics. Version 3.3.4.

Hao, Y.J., Wang, D. and Zhang, X.F. (2006). Review on breeding biology of Yangtze finless porpoise (*Neophocaena phocaenoides asiaeorientalis*). Acta Theriologica Sinica of China. 26(2): 191-200.

Ireland, J.J., Fogwell, R.L., Oxender, W.D., Ames, K., Cowley, J.L. (1984). Production of estradiol by each ovary during the estrous cycle of cows. J. Anim. Sci. 59: 764-771.

Ji, Y.T. (2007). Study on Candidate Genes of Fecundity and Gene Expression in Goat. Yangling: Northwest A and F University, China.

Kumar, S., Stecher, G. and Li, M., *et al.* (2018). MEGA X: Molecular evolutionary genetics analysis across computing platforms. Molecular Biology and Evolution. 35: 1547-1549.

Kumar, S., Stecher G. and Suleski, M., *et al.* (2017). TimeTree: A resource for timelines, timetrees and divergence times. Molecular Biology and Evolution. 34: 1812-1819. DOI: 10.1093/molbev/msx116.

Kiewisz, J., Kaczmarek, M.M. and Morawska, E. *et al.* (2011). Estrus synchronization affects WNT signaling in the porcine reproductive tract and embryos. Theriogenology. 76: 1684-1694.

Li, Y.T. (2017). Study on the habitat selection, environmental capacity and population viability of the yangtze finless porpoises in Tian-e-zhou semi-natural *ex situ* reserve. Wuhan: Institute of Hydrobiology, The Chinese Academy of Sciences.

Lazaros, L.A., Hatzi, E.G. and Pamporaki, C.E. *et al*. (2012). The ovarian response to standard gonadotrophin stimulation depends on FSHR, SHBG and CYP19 gene synergism. Journal of Assisted Reproduction and Genetics. 29(11): 1185-1191.

Lei, X.Q., Chen, H., Yuan, Z.F., *et al*. (2004). Single nucleotide polymorphism in exon 10 of bovine FSHR gene and its relationship with twin traits [J]. Chinese Journal of Biochemistry and Molecular Biology. 20: 34-37.

Lei, X.Q. (2002). DNA Molecular Markers on Bovine and Ovine Fecundity Traits. Yangling: Northwest A and F University, China.

Miyano, Y., Tahara, S. and Sakata, I., *et al*. (2014). Regulation of LH/FSH expression by secretoglobin 3A2 in the mouse pituitary gland. Cell and Tissue Research. 356(1): 253-260.

Migaud, M., Mullen, M.P. and Hanrahan, J.P. *et al*. (2013). Investigation of Prolific Sheep from UK and Ireland for Evidence on Origin of the Mutations in BMP15 (FecXG, FecXB) and GDF9 (FecGH) in Belclare and Cambridge Sheep. PloS One. 8(1): e53172.

Mei, Z.G. (2013). Study on population dynamics and Endangered Mechanism of a Yangtze finless porpoise [Dissertation]. Wuhan: Institute of Hydrobiology, Chinese Academy of Sciences.

Otsuka, F., McTavish, K.J., Shimasaki, S. (2011). Integral role of GDF-9 and BMP-15 in ovarian function. Molecular Reproduction and Development. 78(1): 9-21.

Sun, D. (2018). Population genomics of Finless Porpoise and Adaptive Mechanism of Bone Microstructure of Cetaceans. Nanjing: Nanjing Normal University, China.

Shirakihara, M., Takemura, A. and Shirakhara, K. (1993). Age, growth and reproduction of the finless porpoise. *Nephocaena phocaenoides*, in the coastal waters of western Kyuhu, Japan. Marine Mammal Science. 9(4): 392-406.

Thompson, C.E., Freitas, L.B. and Salzano, F.M. (2018). Molecular evolution and functional divergence of alcohol dehydrogenases in animals, fungi and plants. Genetics and Molecular Biology. 41(1): 341-354.

Thewissen, J.G.M., Cooper, L.N. and Clementz, M.T. *et al*. (2007). Whales originated from aquatic artiodactyls in the Eocene epoch of India. Nature. 450: 1190-1194.

Wu, B., Wang, H.H. and Fu, H.Y. *et al*. (2018). Estimating some population parameters and stock assessment of Dark Sleeper *Odontobutis potamophila* in the Gaosha River, Wuyuan County, Jiangxi Province. Indian Journal of Animal Research. 52(5): 664-668.

Wu, B., Wang, W.P. and Wang, H.H. *et al*. (2020). A retrospective Analysis on the Population Viability of the Yangtze River Dolphin or Baiji (*Lipotes vexillifer*). Indian Journal of Animal Research. DOI: 10.18805/ijar.B-1238.

Wu, B., Xu, J.G. Wang, J.M. *et al*. (2021). Population viability analysis of Yangtze Finless Porpoise in the Yangtze main steam suggesting that a total ban on productive fishing could be decisive. Indian Journal of Animal Research. DOI: 10.18805/IJAR.B-1353.

Yang, Z. (2007). PAML 4.0: Phylogenetic analysis by maximum likelihood. Molecular Biology and Evolution. 24 (8): 1586-1591.

Zhou, X.M, Guang, X.M, Sun, D. *et al*. (2018). Population genomics of finless porpoises reveal an incipient cetacean species adapted to freshwater. Nature Communications. 9: 1276.

Zhang, X.F, Wang, K.X. (1999). Viability analysis of finless porpoise populations in the Yangtze River. Chinese Acta Ecologica Sinica. 19(4): 529-533.