High sequence variation in the exon 10 of *TSHR* gene is associated with flightless-domestic geese

Qiong Peng¹ · Ye Wang¹ · Yaodong Hu¹ · Dan Lan¹ · Daqian He² · Siming Li³ · Zhenping Lin⁴ · Yiping Liu¹

Received: 17 August 2017 / Accepted: 26 July 2018 / Published online: 1 August 2018
© Springer-Verlag GmbH Germany, part of Springer Nature 2018

Abstract

The genetic and molecular mechanisms of the flightless birds without limb modification are rarely reported. To explore the possible reasons for losing flight ability without limb modification, we used the domestic geese as an ideal model to preliminarily study the possible mechanisms for this kind of flightlessness. We compared the sequence variations of the exon 10 of *TSHR* gene between three domesticated geese populations and two wild ancestor populations. The results showed that domestic geese had higher genetic diversity and more complex population structure than their wild ancestors. We did not detect any population expansion in domestic geese population. However, we detected clear relaxed selection signal and positive selection in domesticated geese groups. Furthermore, special phylogenetic relationship of the exon 10 of *TSHR* was observed in domesticated geese groups. Combined with its well-established function on metabolic regulation and photoperiod control, we speculate that relaxed selection of *TSHR* might have effects on flightlessness of domesticated geese.

Keywords Flightlessness · Domestic geese · *TSHR* · Relaxed selection

Qiong Peng and Ye Wang have contributed equally to this work, and shall share as first author.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s13205-018-1371-3>) contains supplementary material, which is available to authorized users.

✉ Yiping Liu
liuyp578@yahoo.com

- ¹ Farm Animal Genetic Resources Exploration and Innovation Key Laboratory of Sichuan Province, Laboratory of Animal Genetics and Breeding, College of Animal Science and Technology, Sichuan Agricultural University, Chengdu Campus, 211 Huiming Road, Wenjiang, Chengdu 611130, Sichuan, People's Republic of China
- ² Institute of Animal Husbandry and Veterinary Sciences, Shanghai Academy of Agricultural Sciences, Shanghai, People's Republic of China
- ³ Institute of Animal Husbandry and Veterinary Sciences, Jiangxi Academy of Agricultural Science, Nanchang, People's Republic of China
- ⁴ Shantou Baisha Research Institute of Original Species of Poultry and Stock, Shantou, Guangdong, People's Republic of China

Introduction

The loss of flight has been considered as a classic case in Darwin's natural selection theory (Darwin and Bynum 2009). So far, 26 families of birds in 17 different orders were found loss of flight with obvious limb modifications (Roff 1994). Though Darwin owned loss of flight to the result of selection in favor of larger bodies and relaxed selection with the absence of predators, the underlying genetic and molecular mechanisms remain unclear. Recently, the Kruglyak's research group applied a predictive and comparative genomics approach to study flightless cormorants (*Phalacrocorax harrisi*), and they found the combined effect of variants in genes regarding to cilia/Ihh signaling which may be responsible for the reduction in the growth of both keel and wings in *P. harrisi* that leading to flightlessness (Burga et al. 2017). Their study provides a new insight into how flightlessness might evolve. However, another important type of flightlessness—loss of flight without limb modification—a new flightlessness model accompanied by human activities has not been widely noticed yet.

Domestic goose is a distinct example for its disappearance of seasonal migration and flight performance, but without obvious limb modification. Domestic goose breeds were deemed to originate from swan goose (*Anser cygnoides*)

and greylag landaise goose (*Anser anser*), as the migrant birds, both of which were famous for excellent ability of long-distance flight (Shi et al. 2006). Recent archaeological discoveries indicated that geese domestication originated around 1700 BC, which means that domestic geese are less than 4000 years old (Ericson and Tyrberg 2004; Mannermaa 2014). This gives us an incredible fact that the time of the domestic geese losing flight ability was much shorter than ratites and penguins which became flightless more than 50 million years, and also shorter than the Galapagos cormorant which became flightless about 2 million years ago (Slack et al. 2006; Kennedy et al. 2009; Mitchell et al. 2014). Interestingly, unlike other flightless birds, domestic geese had no apparent limbs modification which contributed a lot to flightlessness. Obviously, during its domesticated process, the modern goose suffered from strong artificial selection, which imputed have a key role for the flightlessness of domestic geese, for its selection efficiency is more powerful than nature selection. This imply that the flightlessness of domestic geese might be the early status for losing flight of birds—weak flight performance without limb modifications, making domestic geese an ideal model for studying the evolution of loss of flight and artificial selection. Unfortunately, we have limited understanding for the genetic and molecular basis of this process.

A study on chicken found that the thyroid stimulating hormone receptor (*TSHR*) had suffered from striking selective sweep during chicken domestication (Rubin et al. 2010). It is well known that the *TSHR* belongs to the glycoprotein hormone receptor, which has a pivotal role in metabolic regulation and photoperiod control of reproduction in birds and mammals (Hanon et al. 2008; Nakao et al. 2008). In addition, mutations of *TSHR* were reported to have important effects in the regulation of aggression in dog and chicken (Carter et al. 2009; De et al. 2013). Rubin et al. speculated that the selective sweep of *TSHR* might be general in domestic animals, for its close relationship to the absence of the strict regulation of seasonal reproduction (Rubin et al. 2010). As important modern fowl, the geese and chicken had been suffered from similar domestication process during the past thousands of years. Thus, the *TSHR* might also been selected on geese. Combined with its well-established function on metabolic regulation and photoperiod control, we speculated that the possible artificial selection on *TSHR* might have effect on the loss of flight in domestic geese. Therefore, in this study, our goal was to detect whether artificial selection had effects on *TSHR* gene during geese domestication to explore possible molecular reason for flightlessness of domestic geese.

Materials and methods

Sample collection

Blood samples of 73 individuals were collected from three Chinese flightless-domestic geese breeds, including Sichuan white goose (SC), Lion-head goose (S) and Zi goose (Zi), and two wild ancestors including *A. cygnoides* (B) and *A. anser* (H) with good flight performance. Meanwhile, we only took samples that have clear pedigrees to avoid consanguinity. The detailed information of sample collection was listed in Supplementary Table S1. These experiments were approved by the Committee on the Care and Use of Laboratory Animals of the State-Level Experimental Teaching Demonstration Center of Sichuan Agricultural University (Permit No. DKY-S20131105).

DNA amplification and sequencing

Genomic DNA was extracted by standard phenol/chloroform methods. The primer of the exon 10 of *TSHR* (F:5'-ACA TAT CTC AGA CTG AAT TTT ACG C-3'; R:5'-GCT TGG TCT CCT GCT TCC TT-3') used for amplification and sequencing of the exon 10 of *TSHR* were designed according to the genome of Peking duck (GenBank accession no. NW_004676600). Primers were synthesized by Shanghai Sangon Company (Shanghai, China). The estimated length of the target fragment was 2.1 kb. PCR was performed in a mixture volume of 50 μ L containing 100 ng of DNA, 10 mmol/L of Tris-HCl (pH 8.3), 2.5 mmol/L of $MgCl_2$, 50 mmol/L of KCl, 10 mM of each dNTPs, 10 pmol/L of each primer, and 1 unit of Taq polymerase (Takara, Dalian, China). The PCR reaction conditions were as follows: a 7 min preheat at 95 °C, followed by 35 cycles of 40 s at 94 °C, 40 s at 56 °C, and then extension of 72 °C for 5 min. The PCR products were directly sequenced by ABI 3730 sequencer in Shanghai Sangon Company (Shanghai, China).

Data analysis

All target sequences were edited and assembled by BioEdit version 7.1.0 (Hall 1999). Multiple sequence alignment was performed by ClustalX 1.83 (Thompson, Gibson et al. 1997) and variable sites for each gene were analyzed by MEGA7.0 (Kumar, Stecher et al. 2016). The genetic diversity containing number of haplotypes (*H*), haplotype diversity (*Hd*), nucleotide diversity (π), nucleotide substitution rate (θ), the average number of nucleotide differences (*K*), Tajima's D test, and mismatch distribution were calculated using DnaSP5.10 (Librado and Rozas 2009). The historical demographic dynamics for all populations were inferred

from Bayesian skyline analyses implemented in BEAST 1.5.1. In addition, the degree of population differentiation figure was drawn by the STRUCTURE 2.3 (Pritchard, Stephens et al. 2000).

We used Pegasus (Deelman et al. 2005), Plyr (Wickham 2009), and Reshape (McMurdie and Holmes 2013) packages in R program to construct the haplotype network. Bayesian phylogenetic tree was constructed by MrBayes 3.2.6 (Huelsenbeck and Ronquist 2001). The ratio of non-synonymous (d_N) and synonymous (d_S) substitutions and standard errors were performed with a bootstrap of 1000 replicates by MEGA7.0 (Kumar et al. 2016), and using condon-based Z test to do the significance test. The selective pressure detection was computed using the SLAC method in online server (<http://www.datamonkey.org/>) (Pond and Frost 2005).

Results

Sequence variation and molecular diversity

The integrated exon 10 of *TSHR* gene with 1467 bp length was surveyed in this study. Among the five populations,

polymorphic sites of B, H, and SC were relatively lower than that of S and Zi populations (Table S2). Furthermore, the polymorphic sites of SC were slightly lower than that of B population. A large number of mutations were found in Zi population, while no mutation was found in H population. In the 67 domestic geese samples, a total of 146 polymorphic sites were identified. However, only four of them were found in the six wild ancestors. Compared to the domestic geese, the wild ancestors had much fewer polymorphisms. Comprehensively, the domestic geese had higher genetic diversity than the wild ancestors. An insertion Gly-405 and a deletion Asn-406 of exon 10 of *TSHR* gene were simultaneously found in H, SC, and Zi populations, and an insertion Phe-388 was exclusively found in S and Zi groups (Fig. S1).

Expansion of the domestic geese population

The mismatch distribution analysis provides a mean of detecting the genetic signal of a historical demographic expansion in a species or population. The results of population (S, SC, Zi, and the combined domestic geese groups) size change were demonstrated in Fig. 1. The pairwise mismatch distribution curve of the combined domestic geese

Fig. 1 Analysis of the pairwise mismatch distribution of lion-head goose (a), Sichuan white goose (b), Zi goose (c), and the combined domestic geese groups (d). Unimodal distribution represents population expansion; bimodal or multimodal distributions represent

no population expansion. X-axis, number of pairwise differences; Y-axis, frequency of mismatches; Exp, expected (vertical bars); Obs, observed (solid line)

Fig. 2 Bayesian skyline plots for lion-head goose (a), Sichuan white goose (b), Zi goose (c), and the combined domestic geese groups (d). The Y-axis represents the effective population size and the X-axis

represents generation time (million years ago). The mean estimate is enclosed within the 95% highest posterior densities (HPD). The black line indicates the median; the shadow part shows the 95% HPD

Fig. 3 Bayesian clustering plot of five geese populations. The best clustering solution: $K=3$. Individual assignment probabilities of five geese populations to three theoretical genetic ancestry groups using STRUCTURE software. Populations were separated by black thin vertical line

groups and Zi population displayed a bimodal and multi-modal distribution, respectively, which was not consistent with the population expansion model. The S showed that apparent unimodal distribution indicates that the S populations were expanded in the latest historical period.

Bayesian skyline plots (BSP) revealed a relatively explicit demographic history for the same derived four populations.

Fig. 4 Network profile of 28 haplotypes of the exon 10 of *TSHR* gene in five geese groups. Each Rome number corresponds to the corresponding haplotype individuals. Circled areas are proportional to haplotype frequencies. Each breed of domestic geese is marked by a specific color

The combined domestic geese groups showed slight trend of increasing population size after remaining a stable period of time, SC group was nearly at a stable population size. Zi group presented a trend from rise to smooth change after a long period of stability. It was noteworthy that S group underwent a distinct population expansion (Fig. 2).

Population and haplotype structure

Bayesian clustering implemented in STRUCTURE revealed the degree of population differentiation. In our study, the

optimal clustering solution was obtained at $K=3$ (Fig. 3). The B, H, and SC populations almost shared the same cluster, which showed high membership proportions to the cluster, and could be clearly distinguished. While the S and Zi populations comprised by three clusters showed mixed distribution.

A total of 28 haplotypes were detected in exon 10 of *TSHR* gene. The distribution of haplotypes was similar with bayesian clustering results (Fig. 4). Though the star-like topologies of haplotypes, the B, H, and SC populations were relatively concentrated in haplotype I, III, and XVIII,

Fig. 5 Phylogenetic tree of 73 goose and other aves is built by Bayesian method based on the exon 10 of *TSHR* gene. The *Eptesicus fuscus* is regarded as the outgroup. These samples in this study were marked with a background color

while the S and Zi populations showed a complex distribution which was consistent with their structure plot.

Phylogenetic relationships of geese

We investigated the phylogenetic relationship of 73 goose samples based on Bayesian inference (BI) method using *Eptesicus fuscus* as the outgroup (Fig. 5). To get a comprehensive understanding of the phylogenetic relationship of geese in birds, other aves were employed to build the phylogenetic tree, and the reference sequences were shown in Supplementary Table S3. Consequently, four clades were obviously observed in our phylogenetic tree. It was interesting that the domestic geese were all in Clade I and clustered closely with Peking duck, *Coturnix japonica*, and chicken. In addition, birds that belong to *Passeriformes* were clustered eminently into Clade II. However, Clade III was a complex cluster, containing the birds of *trogoniformes*, *sphenisciformes*, *falconiformes*, *gruiformes*, *charadriiformes*, *apodiformes*, *caprimulgiformes*, *coliiiformes*, *pelecaniformes*, *procellariiformes*, *gaviiformes*, *strigiformes*, *pterocliiformes*, and *piciformes*. In Clade III, there was no apparent clustering for each order. However, orders of wader and scansores had a fuzzy trend for clustering. For instance, orders of *gruiformes*, *gaviiformes*, *sphenisciformes*, and *charadriiformes* were clustered closely in Clade III, and orders of *psittaciformes*, *coraciiformes*, and *piciformes* were clustered closely. Interestingly, all birds in Clade IV were flightless, and the relationship of Clade IV was relative distant than the rest clades.

Test for selection

We first performed neutral test for each individual population and two combined groups, wild ancestor (B and H) and domestic geese (S, SC, and Zi) (Table S4). In our results, neutral test for all three domestic geese populations was negative, but not statistically significant. While, when combined S, SC, and Zi together as the domestic geese group, the neutral test was significantly negative. Those results reminded us that population expansion or

directional selection might be happened recently. We can exclude population expansion based on our previous analysis. Thus, we next carried out a synonymous (d_S) and non-synonymous (d_N) substitution analysis to detect the possible selection pressure (Table S5). Among the three domestic geese populations, $d_N/d_S > 1$ were both found in S and Zi population without statistically significant. To further study the selection pressure, an SLAC method was also employed to detect selection pressure for each codon (Table 1). Finally, we detected a significant positive selection on codon 227 and 305, and a significant negative selection on codon 405.

Discussion

Flightless birds had a general feature of changing size and proportion of limbs, which played an outstanding role in favor of Darwin's theory of natural selection. Heterochrony, the relative change in the rate or timing of developmental events among species, has been thought as the major reason of limb modifications that leading the flightlessness of birds (Roff 1994). A recent study on Galapagos cormorant proposed a genetic and molecular model to explain the underlying mechanisms of flightlessness (Burga et al. 2017). This study found that the perturbations of cilia/Ihh signaling might be responsible for the reduction in growth of both keel and wings that leading to the flightlessness of Galapagos cormorant. However, this explanation was not suitable for flightless birds without apparent limbs change. Virtually, nothing is known about the underlying genetic and molecular mechanisms of this kind of flightlessness. Thus, we chose the typical representative animal, domestic geese, as our research target to preliminarily study the underlying mechanisms. In our result, higher level of genetic diversity was found in exon 10 of *TSHR* gene in domestic geese populations. Besides, compared to wild ancestors, domestic geese had more haplotypes and more complex population structure, indicating a higher genetic diversity. The neutral test of domestic geese group was all significantly negative, which reminded us population expansion or directional selection of exon 10 of *TSHR* which might happen recently. However, the mismatch distribution and BSP results found no apparent population expansion in the domestic geese populations, which showing a clear evidence of directional selection. The further analysis of d_N/d_S revealed a result of non-significant positive selection, but two significant positive selection sites were still detected. Consequently, our results revealed a clear evidence of relaxed selection of exon 10 of *TSHR* gene, which indicating that relaxation of selective constraints played a key role in the evolution of flightlessness of domestic geese. Interestingly, our results

Table 1 Selection pressure detection of each codon of the exon 10 of *TSHR* gene

Codon	$d_N-d_S^a$	p value
227	0.926401	0.027873
305	1.01092	0.016911
405	-2.23567	0.000111

^a $d_N-d_S > 0$ positively selected sites; $d_N-d_S < 0$ negatively selected sites (0.1 significance level)

were similar with Shen's research, who studied on rapidly flying birds, weakly flying, and flightless birds, and highlighted that birds with degraded flight ability would like to accumulate more non-synonymous nucleotide substitutions relative to synonymous substitutions on mtDNA (Shen et al. 2009). The authors also raised that functional constraints on locomotion played an important role in keeping flight. In our study, the relaxed selection signal in the exon 10 of *TSHR* was detected in flightless-domestic geese. As we know, *TSHR* has an important impact in metabolic and photoperiod control of seasonal reproduction in domestic chicken. Thus, we speculate that relaxed selection of *TSHR* gene might have effects on the flightlessness of domestic geese. In addition, the *TSHR* occurred glycine-to-arginine substitution during the evolution of domestic chicken, which may be related to the absence of the strict regulation of seasonal reproduction that is the classical feature of domestic chicken (Rubin et al. 2010). Similarly, in our study, insertion and deletion of amino acid in exon 10 of *TSHR* were found in different geese populations. In addition, these frameshift variants are not reported in other species. We speculated that these frameshift variants in exon 10 of *TSHR* may affect the function of *TSHR* and be relative to the change of behaviour.

To get a comprehensive understanding of phylogenetic relationship of *TSHR* gene, we constructed a Bayesian inference tree containing most orders of birds. In our results, all domestic geese were clustered into Clade I. Interestingly, this clade was totally composed by birds that had suffered from artificial selection. Theoretically, domestic geese should be clustered into Clade III, and clustered together with other waterfowl, or clustered together with in Clade IV with flightless birds. The unexpected cluster pattern suggested us that flightlessness resulted by domestication had stronger effect than that of nature selection. Due to the limitation of sample size and studied gene, we cannot get a more precise and comprehensive understanding about the mechanisms of flightlessness of domestic geese. In addition, it is hard for us to clarify whether those frameshift mutations affected the function of *TSHR*; this is another limitation in this study. However, our studies remind us that studying the flightlessness of domesticated birds without limb modification might be helpful to reveal the evolutionary mechanisms of flightlessness of birds.

Acknowledgements This work was funded by the National Natural Science Foundation of China (Nos. 31172181 and 31660663) and the National Waterfowl Industrial Technology System (Nos. CARS-43-4 and 2016YFD0500510).

Compliance with ethical standards

Conflict of interest The authors declare that they have no competing interests.

References

- Burga A, Wang W et al (2017) A genetic signature of the evolution of loss of flight in the Galapagos cormorant. *Science* 356(6341):eaal3345
- Carter GR, Scott-Moncrieff JC et al (2009) Serum total thyroxine and thyroid stimulating hormone concentrations in dogs with behavior problems. *J Vet Behav Clin Appl Res* 4(6):230–236
- Darwin C, Bynum WF (2009) The origin of species by means of natural selection: or, the preservation of favored races in the struggle for life. AL Burt
- De GB, Grommen SV et al (2013) Hatching the cleidoic egg: the role of thyroid hormones. *Front Endocrinol* 4:63
- Deelman E, Singh G et al (2005) Pegasus: a framework for mapping complex scientific workflows onto distributed systems. *Sci Program* 13(3):219–237
- Ericson PGP, Tyrberg T (2004) The early history of the Swedish avifauna. A review of the subfossil record and early written sources
- Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl Acids Symp Ser*
- Hanon EA, Lincoln GA et al (2008) Ancestral TSH mechanism signals summer in a photoperiodic mammal. *Curr Biol* 18(15):1147–1152
- Huelsenbeck JP, Ronquist F (2001) MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 17(8):754–755
- Kennedy M, Valle CA et al (2009) The phylogenetic position of the Galápagos Cormorant. *Mol Phylogenet Evol* 53(1):94–98
- Kumar S, Stecher G et al (2016) “MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol* 33(7):1870
- Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25(11):1451–1452
- Mannermaa K (2014) *Goose: domestication*. Springer, New York
- McMurdie PJ, Holmes S (2013) phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* 8(4):e61217
- Mitchell KJ, Llamas B et al (2014) Ancient DNA reveals elephant birds and kiwi are sister taxa and clarifies ratite bird evolution. *Science* 344(6186):898–900
- Nakao N, Ono H et al (2008) Thyrotrophin in the pars tuberalis triggers photoperiodic response. *Nature* 452(7185):317
- Pond SLK, Frost SD (2005) Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. *Bioinformatics* 21(10):2531–2533
- Pritchard JK, Stephens M et al (2000) Inference of population structure using multilocus genotype data. *Genetics* 155(2):945–959
- Roff DA (1994) The evolution of flightlessness: is history important?. *Evol Ecol* 8(6):639–657
- Rubin CJ, Zody MC et al (2010) Whole-genome resequencing reveals loci under selection during chicken domestication. *Nature* 464(7288):587
- Shen Y-Y, Shi P et al (2009) Relaxation of selective constraints on avian mitochondrial DNA following the degeneration of flight ability. *Genome Res* 19(10):1760–1765
- Shi XW, Wang JW et al (2006) Mitochondrial DNA cleavage patterns distinguish independent origin of Chinese domestic geese and western domestic geese. *Biochem Genet* 44(5–6):237–245
- Slack KE, Jones CM et al (2006) Early penguin fossils, plus mitochondrial genomes, calibrate avian evolution. *Mol Biol Evol* 23(6):1144–1155
- Thompson JD, Gibson TJ et al (1997) The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* 25(25):4876–4882
- Wickham H (2009) plyr: Tools for splitting, applying and combining data. R package version 0(1 9):651