

Whole Genome Sequencing of Hulunbuir Short-Tailed Sheep for Identifying Candidate Genes Related to the Short-Tail Phenotype

Dafu Zhi,* Lai Da,[†] Moning Liu,* Chen Cheng,* Yukun Zhang,[‡] Xin Wang,* Xiunan Li,* Zhipeng Tian,[‡] Yanyan Yang,[†] Tingyi He,[†] Xin Long,[‡] Wei Wei,* and Guifang Cao*¹

*College of Veterinary Medicine and [‡]College of Life Sciences, Inner Mongolia Agricultural University, Huhhot 010018, People's Republic of China and [†]Inner Mongolia Academy of Agricultural and Animal Husbandry Sciences, Huhhot 010018, People's Republic of China

ABSTRACT The Hulunbuir short-tailed sheep (*Ovis aries*) is a breed native to China, in which the short-tail phenotype is the result of artificial and natural selection favoring a specific set of genetic mutations. Here, we analyzed the genetic differences between short-tail and normal-tail phenotypes at the genomic level. Selection signals were identified in genome-wide sequences. From 16 sheep, we identified 72,101,346 single nucleotide polymorphisms. Selection signals were detected based on the fixation index and heterozygosity. Seven genomic regions under putative selection were identified, and these regions contained nine genes. Among these genes, *T* was the strongest candidate as *T* is related to vertebral development. In *T*, a nonsynonymous mutation at c.G334T resulted in p.G112W substitution. We inferred that the c.G334T mutation in *T* leads to functional changes in Brachyury—encoded by this gene—resulting in the short-tail phenotype. Our findings provide a valuable insight into the development of the short-tail phenotype in sheep and other short-tailed animals.

KEYWORDS

Hulunbuir short-tailed sheep short-tail phenotype whole genome sequencing selective sweep *T*/Brachyury gene Genome Report

Sheep were among the earliest domesticated herbivores. Sheep domestication dates back to the end of the Mesolithic period, ~11,000 yr ago (Chessa *et al.* 2009). Domestication and artificial selection have led to marked changes in sheep behavior, appearance, and other important traits (Megens *et al.* 2008). Most mammal tails are used for balance, communication, and attack. However, in sheep, tails store energy. The fat-tail phenotype is a trait necessary for survival in harsh environments (Poullis 2011). It is exhibited by Hulunbuir short-tailed sheep, which have been bred for the past century by local herders of the Hulunbuir grassland—a world-renowned highland pasture in arid and semi-arid regions of North China characterized by a short frost-free period, long

and cold winter, and dry season constituting a considerable part of the year (Yang *et al.* 2012). The tails of Hulunbuir short-tailed sheep (Figure 1A) are of variable length, and are generally classified into two categories: extremely short, a tail exposing the anus (Figure 1C); and moderately short, a tail covering the anus (Figure 1D).

Several candidate genes—including *HES7* (Bessho *et al.* 2001), *PAX1* (Wilm *et al.* 1998), *T* (Smith 1999), and *WNT3A* (Greco *et al.* 1996)—related to vertebral development in laboratory mice are reported to be related to the short-tail phenotype; however, the determinant genes related to the short-tail phenotype in sheep remain to be identified.

Recently, many genomic regions under putative selection—which may be related to domestication, adaptation, and other important traits—have been reported in chickens, cats, dogs, and pigs (Rubin *et al.* 2010; Axelsson *et al.* 2013; Moon *et al.* 2015; Xiao *et al.* 2016). However, detection of selection signals within a single species or population cannot determine whether the genomic region is under putative selection or related to genetic drift. Genomic regions under putative selection are identified using the fixation index (F_{st}), which is based on significant differences in allele frequencies between two populations. However, F_{st} does not identify the population wherein selection occurs; hence, it cannot be used to determine the direction of selection. Therefore, in the present study, we analyzed heterozygosity (H_p) to identify the specific genomic regions under putative selection. Axelsson *et al.*

Copyright © 2018 Zhi *et al.*

doi: <https://doi.org/10.1534/g3.117.300307>

Manuscript received September 26, 2017; accepted for publication November 27, 2017; published Early Online December 5, 2017.

This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Supplemental material is available online at www.g3journal.org/lookup/suppl/doi:10.1534/g3.117.300307/-/DC1.

¹Corresponding author: College of Veterinary Medicine, Inner Mongolia Agricultural University, No. 306, Zhaowuda Rd., Huhhot 010018, People's Republic of China. E-mail: guifangcao_jmau@126.com

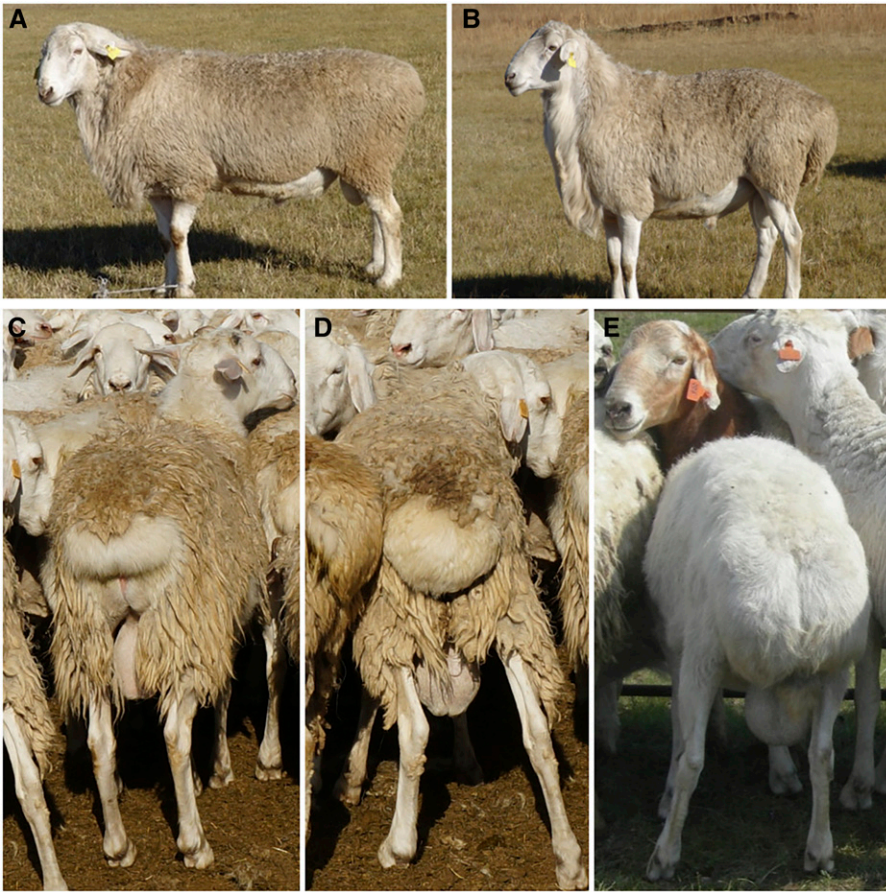


Figure 1 Tail phenotypes of sheep. (A) Hulanbair short-tailed sheep; (B) Barag sheep; (C) extremely short in Hulanbair short-tailed sheep; (D) moderately short in Hulanbair short-tailed sheep; (E) normal tail in Barag sheep.

(2013) combined F_{st} and H_p to locate the genomic regions under putative selection, and the corresponding genes during dog domestication. Here, we applied the methodology of Axelsson *et al.* (2013) to locate the genomic regions related to the short-tail phenotype in Hulanbair short-tailed sheep.

To verify the role of the determinant genes and identify novel genes regulating the short-tail phenotype, we selected sheep with extremely short tails from a population of Hulanbair short-tailed sheep, and performed whole genome sequencing. The short-tail phenotype has previously been reported in Australian Merino sheep (James *et al.* 1991); however, to the best of our knowledge, we are the first to conduct molecular studies of the short-tail phenotype in sheep.

MATERIALS AND METHODS

Ethics statement

All animal care and experiments were conducted in accordance with the *Guide for the Care and Use of Laboratory Animals*, and were approved by the Institutional Animal Care and Use Committee of Inner Mongolia Agricultural University.

Samples

Hulanbair short-tailed sheep and Barag sheep were selected from the Autonomous County of Evenki in the Inner Mongolia Autonomous Region, China. Barag sheep (Figure 1B), which resemble Hulanbair short-tailed sheep but have a normal tail morphology (Figure 1E), served as controls. We randomly selected 100 short-tailed sheep and

100 Barag sheep for tail-length measurement. Caudal vertebrae were collected from 10 short-tailed sheep and 10 Barag sheep. Blood samples were collected from the following sheep—eight unrelated 2-yr-old short-tailed sheep (four males and four females) with extremely short tails, and eight unrelated 2-yr-old Barag sheep (four males and four females) with normal tails were selected for whole genome sequencing; 120 short-tailed sheep and 110 Barag sheep were randomly selected to confirm the mutation site; ~2 ml of blood was collected in EDTA-containing vacutainers and stored in liquid nitrogen (-196°). Genomic DNA was extracted from the whole blood samples using the DNeasy Blood and Tissue Kit (Qiagen, Duesseldorf, Germany).

X-ray analysis and specimen preparation

X-ray images of the caudal vertebrae were captured from the front using the ClearVet DR16 Imaging System. Deposits and muscle were carefully peeled off from the caudal vertebrae. The caudal vertebrae were fixed in ethanol for 5 d, and cleared by immersing in 0.5% NaOH for 2 d. To remove fat, oil, and adipose tissue, the caudal vertebrae were immersed in petrol for 3 d. Data were recorded and saved according to the order of X-ray images.

Sequencing and SNP calling

A sequencing library with an average insert size of 350 bp was constructed for each sample. Sequencing was performed on the Illumina HiSeq 2000 sequencer at Novogene Corporation to generate 150-bp paired-end reads; ~13.1 Gb of high-quality data were generated, achieving an average fivefold genome coverage for each individual. The clean data were aligned to the sheep reference genome

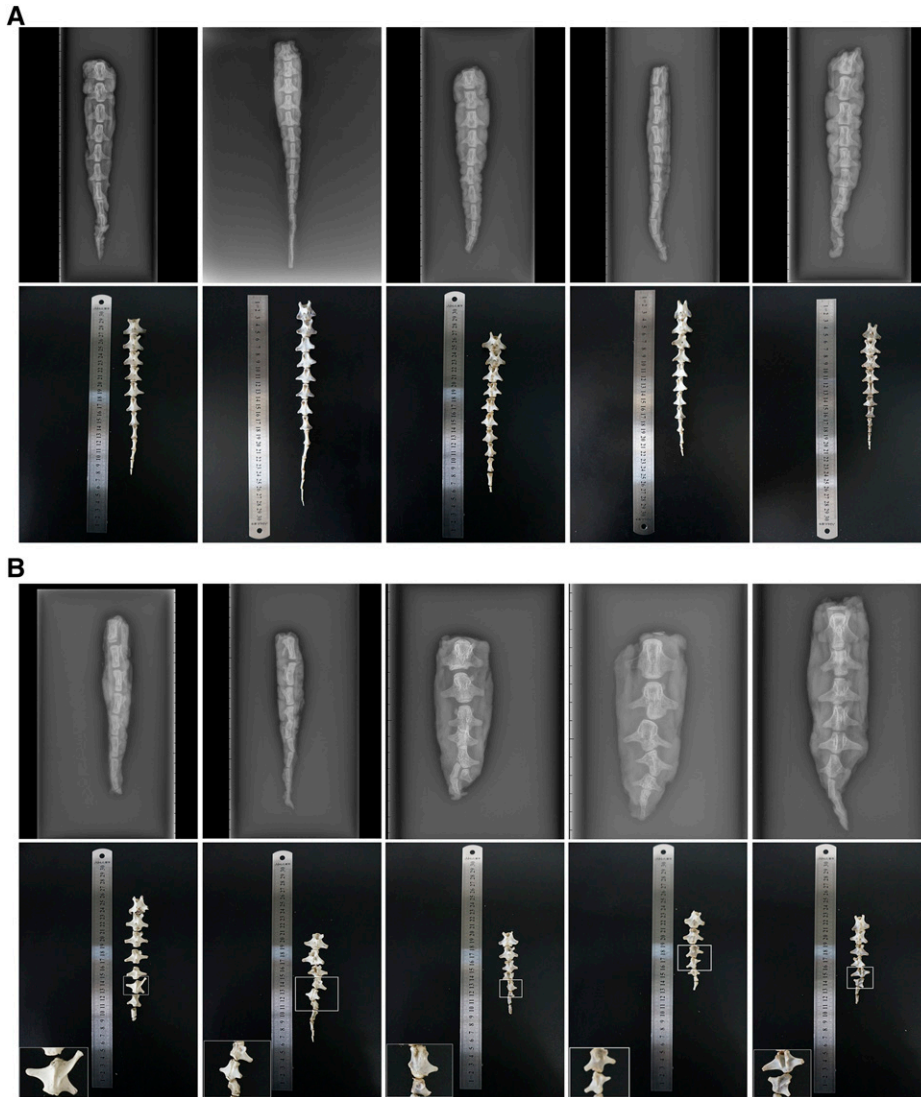


Figure 2 Phenotype of caudal vertebrae in the two populations. (A) Phenotype of the caudal vertebrae in Barag sheep; (B) phenotype of the caudal vertebrae in Hulunbuir short-tailed sheep. Top rows show X-rays and bottom rows show the caudal vertebrae that were placed based on the X-rays in (A and B).

(Oar_v3.1 http://asia.ensembl.org/Ovis_aries/Info/Index) using Burrows-Wheeler Aligner 0.6.1 with parameters set as “mem -t 4 -k 32 -M” (Li and Durbin 2010); duplications were eliminated from the alignments using SAMtools with parameters set as “rmdup.” Single nucleotide polymorphism (SNPs) were detected using SAMtools with parameters set as “mpileup -m 2 -F 0.002 -d 1000” and annotated using ANNOVAR (Li *et al.* 2009; Wang *et al.* 2010). SNPs with low-quality scores ($GQ < 20$) and inter-SNP distance of < 5 bp were filtered.

Selective sweep analysis

A sliding-window approach (100-kb windows sliding in 10-kb steps) was employed for quantifying H_p in the short-tailed sheep and F_{st} between the short-tailed and Barag sheep (Li *et al.* 2014). H_p was calculated using the formula $H_p = 2 \sum nMaj \sum nMin / (\sum nMaj + \sum nMin)^2$, where $\sum nMaj$ and $\sum nMin$ are the sum of nMaj and nMin, respectively. We transformed the H_p values into Z-scores using the formula $ZH_p = (H_p - \mu H_p) / \sigma H_p$, where μH_p is the overall average heterozygosity and σH_p is the SD for all windows within the group. Genetic differentiation between the short-tailed and Barag sheep was measured using

fixation index (F_{st}) with the formula $F_{st} = 1 - p_1 q_1 + p_2 q_2 / 2 p_r q_r$, where p_1, p_2 and q_1, q_2 represent the frequencies of alleles A and a in populations of Hulunbuir short-tailed sheep and Barag sheep, respectively, and p_r and q_r represent the frequencies of alleles A and a, respectively, in the whole population (Clark and Hartl 2007). F_{st} values were Z-transformed using the formula for ZH_p . The H_p and F_{st} data were calculated based on the R language package (<http://www.r-project.org/>). The selected regions were defined as genetic regions in overlapping windows with extremely low ZH_p values ($ZH_p < -5$) and extremely high ZF_{st} values ($ZF_{st} > 4.5$).

T sequencing

The primer pairs for the amplification of *T* exons were designed based on the sheep genome assembly (Oar_v3.1). The primer sequences are provided in Supplemental Material, Table S1. Sanger sequences of PCR duplicates were detected by Invitrogen Corporation, Shanghai, China.

Data availability

The Illumina sequence reads were deposited in the NCBI Sequence Read Archive under the accession number SRP106953. The authors state that

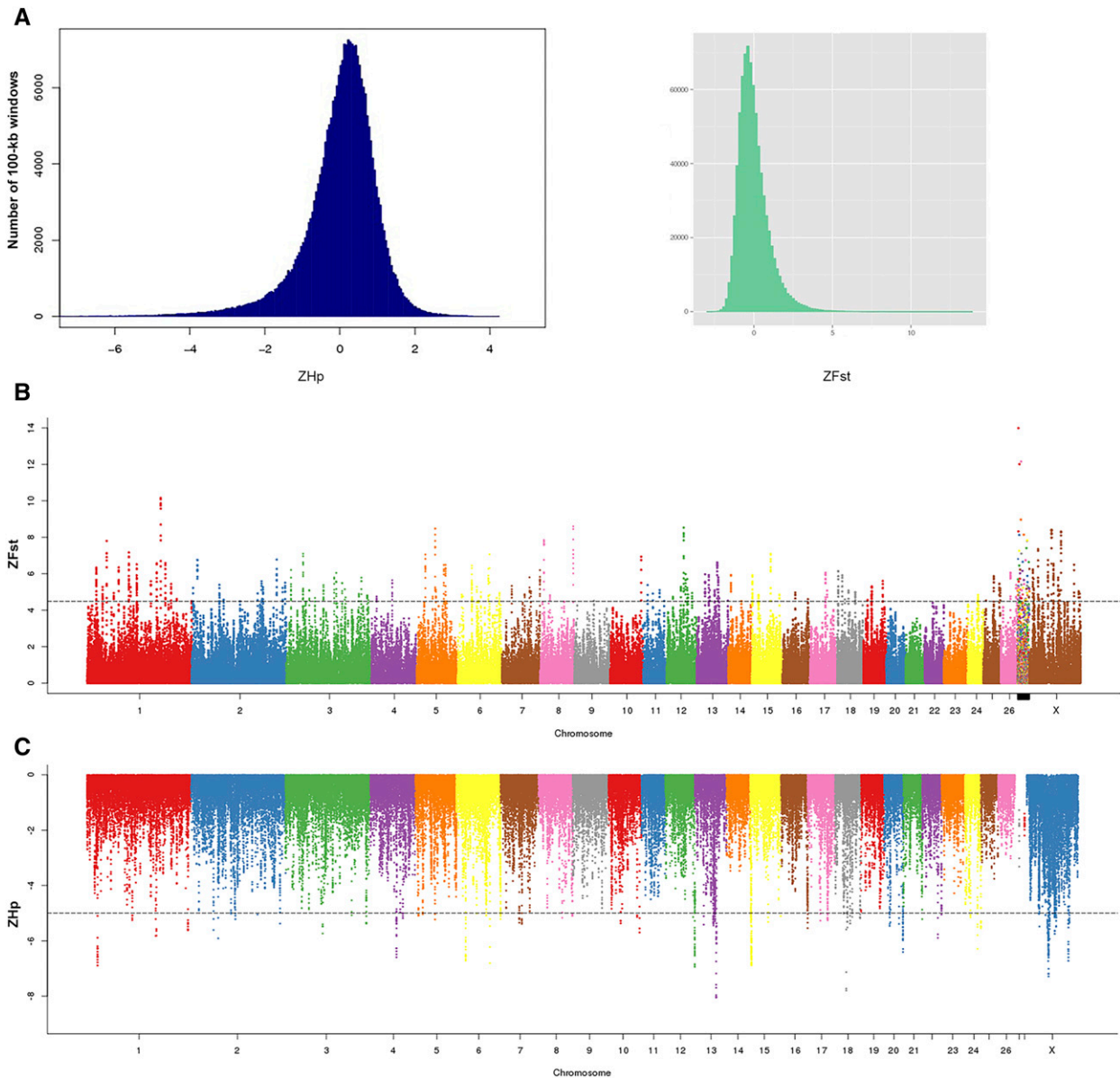


Figure 3 Selective analysis of the sheep genome. (A) Distribution of Z-transformed average heterozygosity in the Hulanbuir short-tailed sheep and Z-transformed average fixation index (ZF_{st}) for autosomal 100-kb windows. (B) Positive end of the (ZF_{st}) distribution plotted along chromosomes (chromosomes are separated by color); horizontal dashed line indicates the cut-off ($ZF_{st} > 4.5$) used for extracting outliers. (C) Negative end of the (ZH_p) distribution plotted along chromosomes (chromosomes are separated by color); horizontal dashed line indicates the cut-off ($ZH_p < -5$) used for extracting outliers.

all data necessary for confirming the conclusions presented in the article are represented fully within the article.

RESULTS

Measurements

The average tail lengths of the Barag and Hulanbuir short-tailed sheep were 26.32 and 16.01 cm, respectively. The average lengths of the caudal vertebrae in the Barag and Hulanbuir short-tailed sheep were 21.12 and 13.86 cm, respectively. Deformed caudal vertebrae were observed in seven Hulanbuir short-tailed sheep (Figure 2).

Genome resequencing

Eight sheep with extremely short tails were selected from the short-tailed sheep, and eight Barag sheep with normal tails were selected. We

performed genome sequencing for the 16 sheep and obtained 232.23 Gb of paired-end DNA reads. Of this, 231.07 Gb (99.50%) consisted of high-quality paired-end reads ($Q20 \geq 94.33\%$ and $Q30 \geq 88.47\%$) that could be aligned to the sheep reference genome (Oar_v3.1) (Table S2). For individual sheep, the alignment rate was 98.13–98.75%. For the reference genome (excluding N's), the average coverage depth was 4.59–5.27 \times . The reads with 1 \times coverage depth (covered by at least one base) accounted for >94.19% (Table S3). The read alignments were normal, and the reads were eligible for variation detection and other analyses.

SNP and selective sweep

In the 16 sheep, we identified 72,101,346 SNPs, of which 492,307 SNPs were localized to coding regions, and responsible for 205,408 nonsynonymous nucleotide substitutions (202,338 missense, 2905 stop-gains, and

■ Table 1 Overlapping genes identified using ZH_p and ZF_{st}

Chr	Interval Location	ZF _{st}	ZH _p	Candidate Gene	Annotation	Gene Location
1	69800001–69950000	4.9462	-5.4862	<i>FNBP1L</i>	Formin-binding protein 1-like	69805549–69915406
1	69800001–69950000	4.7863	-5.9525	<i>BCAR3</i>	Breast cancer anti-estrogen resistance 3	69930328–70015474
6	37440001–37580000	4.9422	-5.0057	<i>LCORL</i>	Ligand-dependent nuclear receptor corepressor-like	37365236–37452332
8	87770001–87890000	8.4551	-5.1642	<i>T</i>	T, Brachyury homolog	87796143–87805552
15	16940001–17060000	4.7804	-6.5384	<i>SLC35F2</i>	Solute carrier family 35, member F2	16854039–16994650
15	16940001–17060000	4.6566	-5.9519	<i>RAB39A</i>	RAB39A, member RAS oncogene family	166995339–17027728
15	17090001–17100000	4.7602	-6.0915	<i>CUL5</i>	Cullin 5	17088942–17210813
X	77440001–77470000	4.8223	-5.1215	<i>IRAK1</i>	Interleukin-1 receptor-associated kinase 1	77447119–77450460
X	77440001–77470000	4.8801	-5.0186	<i>MECP2</i>	Methyl-CpG-binding protein 2	77461651–77467940

165 stop-losses) (Table S4). According to the SNP densities in the genome and reference genome assembly, the number of SNPs in the sliding window was <20 beginning from 100 kb; this trend persisted. Additionally, the regions containing <20 SNPs were stable. Therefore, the width of the sliding window was selected as 100 kb. Alignments were performed between the two populations, using a 50%-overlapping interval as the step length. For each window, F_{st} was calculated and Z-transformed into a ZF_{st} value that obeyed standard normal distribution (Figure 3A). The larger the ZF_{st} value, the greater the genetic differentiation between the populations. H_p was evaluated for the short-tailed sheep population with the same sliding-window parameters, and it was Z-transformed to standard normal distribution for the population (Figure 3A). Based on the Z value, we observed considerable genetic differentiation in many genomic regions between the populations, including those unrelated to tail morphology. To determine

whether certain genetically different genomic regions were related to tail morphology, we overlapped the small interval of ZH_p and large interval of ZF_{st}. We set ZF_{st} = 4.5 as the threshold for identifying the genomic regions under putative selection and selected windows with ZF_{st} values greater than this threshold (Figure 3B). Moreover, we set ZH_p = -5 as the threshold and selected windows with ZH_p values less than this threshold (Figure 3C). The overlapping regions were identified and combined into seven candidate genomic regions (20–160 kb). These candidate regions covered a genomic length of 790 kb and contained nine genes (Table 1).

Of these nine annotated genes, *T*, which is a gene related to development of the spine in mice (Beddington *et al.* 1992), harbored 10 mutations, including eight synonymous and two nonsynonymous mutations. The two nonsynonymous substitutions in *T* were considered putative mutations, including G-to-T transversion in exon two (c.G334T) and

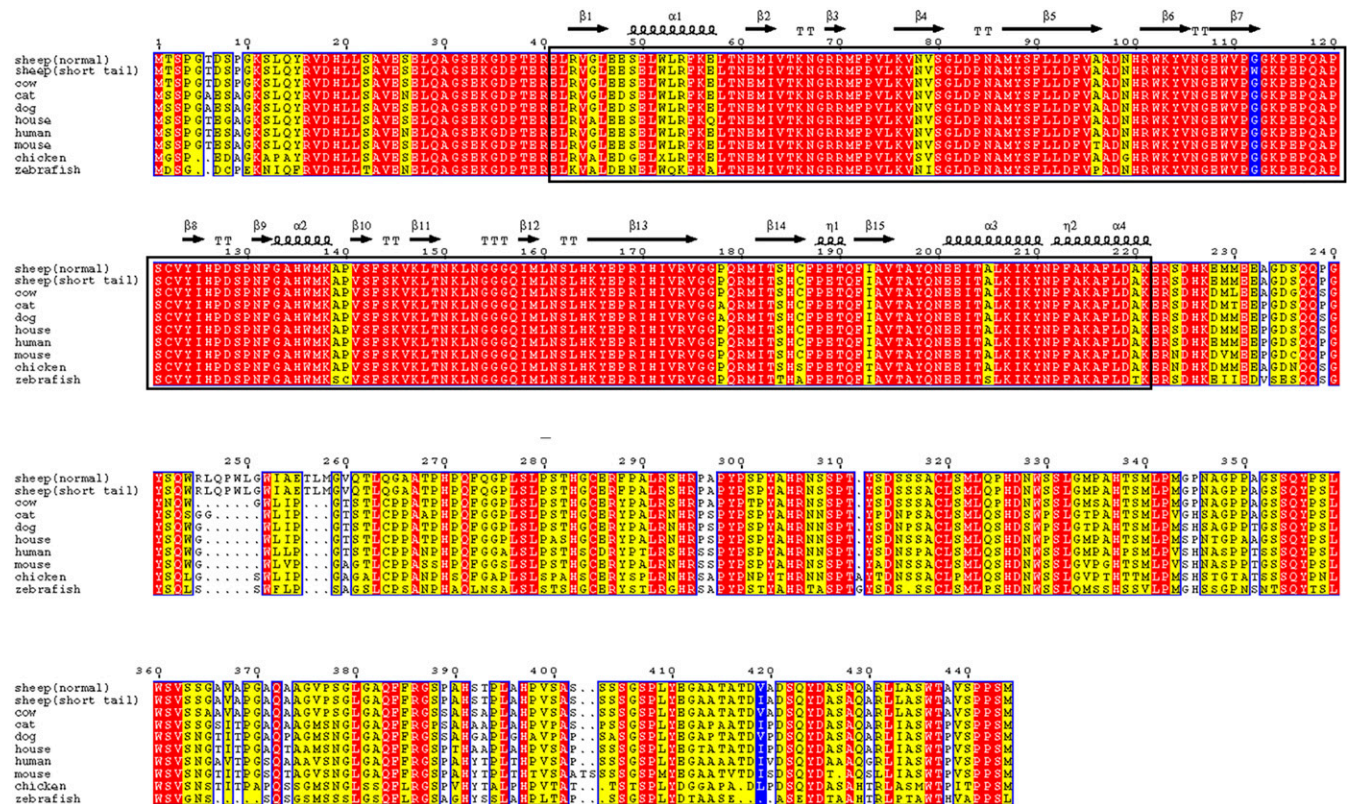


Figure 4 Alignment of amino acid sequences of Brachyury among animals. Brachyury is a protein encoded by *T*. Amino acid residues where p.G112W and p.V419I were located are indicated in blue. Identical amino acids, conserved substitutions, and semiconserved substitutions are indicated in red, yellow, and white respectively. Dots represent gaps in the alignment.

■ **Table 2 Summary of genotypic data in the two populations**

Breed	N	c.G334T			c.G1255A		
		(G/G)	(G/T)	(T/T)	(G/G)	(G/A)	(A/A)
Hulunbair short-tailed sheep	120	17	103	0	81	0	39
Barag sheep	110	110	0	0	97	0	13

G-to-A transition in exon nine (c.G1255A), which corresponded to glycine-to-tryptophan substitution at amino acid residue 112 (p.G112W) and valine-to-isoleucine substitution at amino acid residue 419 (p.V419I), respectively (Figure 4).

Validation using Sanger sequencing

The identified nonsynonymous nucleotide substitutions in *T* were verified using Sanger sequencing with a large sample size in 120 short-tailed and 110 Barag sheep. Variation data were deposited in GenBank under the accession number MF996360. We detected the heterozygous form of c.G334T in the short-tailed sheep, but not in the Barag sheep. However, we detected c.G1255A in both populations (Table 2). We inferred that the c.G334T mutation in *T* is the primary cause of the short-tail phenotype. However, this does not exclude the possibility that other genes control the short-tail phenotype.

DISCUSSION

Hulunbair short-tailed sheep are fat-tailed sheep, with large quantities of adipose deposited in their tail regions. Most previous research studies on sheep tails have focused on adipose deposition (Moradi *et al.* 2012; Wang *et al.* 2014; Yuan *et al.* 2017). Liu *et al.* (2015) suggested that tail length in short-tailed sheep is related to adipose deposition. However, we believe that tail length is related to the length of the caudal vertebrae. This prediction was supported by our observation of deformed vertebrae in the tails of seven Hulunbair short-tailed sheep. The specific molecular mechanism functioning in the short-tail phenotype remains to be elucidated; however, we believe that the mutated *T* gene plays an important role in regulating this phenotype. Of the nine investigated genes, *BCAR3* (Agthoven *et al.* 1998), *FNBPI1* (Huett *et al.* 2009), *IRAK1* (Li *et al.* 2015), *CUL5* (Byrd *et al.* 1997), and *RAB39A* (Seto *et al.* 2013) are related to the immune system; *MECP2* (Tao *et al.* 2009) and *SLC35F2* (Sonuga-Barke 2013) are related to neurodevelopment; and *LCORL* (Signer-Hasler *et al.* 2012) is related to skeletal frame size. Only the *T* gene is related to development of the spine, and this gene seems to be associated with a short-tail phenotype in mice (Beddington *et al.* 1992).

We identified seven candidate genomic regions that included *T*. This gene was further verified using Sanger sequencing. *T* encodes a transcription factor named Brachyury, which is the key regulator of mesoderm formation during early development. The Brachyury protein is an important functional transcription factor, in which ~180 amino acid residues located near the N-terminus display DNA-binding activity; this is the T domain. This region can specifically bind to a 5-bp functional domain in DNA (TCACA) (Kispert and Herrmann 1993; Palena *et al.* 2007; Fernando *et al.* 2010). Brachyury is specifically expressed in the notochord of the mesoderm during gastrulation, and it regulates growth and development of the embryonic notochord. However, Brachyury is not expressed during mid-to-late pregnancy (Sangoi *et al.* 2011).

As early as 1927, Dobrovolskaia-Zavadskaia described the phenotype of a Brachyury mutation in mice (Kavka and Green 1997). Mice heterozygous for the Brachyury mutation had short and slightly curved tails. Homozygous or compound heterozygous mice died *in utero* after ~10 d of the embryonic period because of failure to form the dorsal

cord and allantois (Showell *et al.* 2004). The short-tail phenotype in mice was first discovered in 1927, but it was not until 1990 that the *T* gene was cloned (Herrmann *et al.* 1990). Pennimpede *et al.* (2012) characterized an inducible miRNA-based on an *in vivo* knockdown mouse model of *T*, which exhibited skeletal defects.

We identified two loci of nonsynonymous mutations in *T* using genome sequencing in short-tailed sheep, and we localized the c.G334T mutation in the T domain of the Brachyury gene. Specific regions in the T domain are highly sensitive to structural changes caused by mutations, and these mutations may result in conformational changes in the protein. By combining X-ray diffraction and analysis of the DNA-binding domain in *Xenopus*, Muller observed high similarity between amino acids in this region and those in contact with DNA (Müller and Herrmann 1997). In the present study, we inferred that the p.G112W mutation affects the binding ability of the T domain to DNA.

We did not detect the c.T334T mutation in the short-tailed sheep population, consistent with the theory that homozygotes or compound heterozygotes result in embryonic lethality (Buckingham *et al.* 2013). The c.G334G mutation was detected in individuals with normal tails. Wu *et al.* (2010) observed that, when mice with moderately long tails were mated with short-tailed mice, the offspring showed short-tail and normal-tail phenotypes; however, the genotypes of the offspring were not analyzed. This may explain the phenomenon of the c.G334G mutation in individuals with normal tails in the short-tailed sheep population. Herdsmen continue to breed this group and eliminate individuals with nonideal tail types and other individuals exhibiting poor phenotypes; we propose that this explains the nonconformity of genotypic frequencies with genetic theory.

In conclusion, the results of our present study suggest that the c.G334T mutation in *T* directly results in the short-tail phenotype in sheep. The candidate genes identified in our study provide the basis for understanding the molecular mechanism of the short-tail phenotype in sheep and other short-tailed animals.

ACKNOWLEDGMENTS

We thank Wang Yu of the Bureau of Animal Husbandry of Evenki for sample collection. This work was supported by the Technology Innovation Project of the Inner Mongolia Agricultural University (NDPYTD 2010-6) and the Key Program of the Department of Science and Technology, Inner Mongolia Autonomous Region (201602057).

LITERATURE CITED

- Agthoven, T. V., T. L. A. V. Agthoven, A. Dekker, P. J. V. D. Spek, L. Vreede *et al.*, 1998 Identification of BCAR 3 by a random search for genes involved in antiestrogen resistance of human breast cancer cells. *EMBO J.* 17: 2799–2808.
- Axelsson, E., A. Ratnakumar, M. L. Arendt, K. Maqbool, M. T. Webster *et al.*, 2013 The genomic signature of dog domestication reveals adaptation to a starch-rich diet. *Nature* 495: 360.
- Beddington, R. S., P. Rashbass, and V. Wilson, 1992 Brachyury—a gene affecting mouse gastrulation and early organogenesis. *Dev. Suppl.* 114: 157.
- Bessho, Y., R. Sakata, S. Komatsu, K. Shiota, S. Yamada *et al.*, 2001 Dynamic expression and essential functions of Hes7 in somite segmentation. *Genes Dev.* 15: 2642.
- Buckingham, K. J., M. J. McMillin, M. M. Brassil, K. M. Shively, K. M. Magnaye *et al.*, 2013 Multiple mutant T alleles cause haploinsufficiency of Brachyury and short tails in Manx cats. *Mamm. Genome* 24: 400–408.
- Byrd, P. T., C. M. McConville, A. D. Smith, P. R. Cooper, and A. M. Taylor, 1997 Identification and analysis of expression of human VACM-1, a cullin gene family member located on chromosome 11q22–23. *Genome Res.* 7: 71–75.
- Chessa, B., F. Pereira, F. Arnaud, A. Amorim, F. Goyache *et al.*, 2009 Revealing the history of sheep domestication using retrovirus integrations. *Science* 324: 532–536.

- Fernando, R. I., M. Litzinger, P. Trono, D. H. Hamilton, J. Schlom *et al.*, 2010 The T-box transcription factor Brachyury promotes epithelial-mesenchymal transition in human tumor cells. *J. Clin. Invest.* 120: 533–544.
- Greco, T. L., S. Takada, M. M. Newhouse, J. A. McMahon, A. P. McMahon *et al.*, 1996 Analysis of the vestigial tail mutation demonstrates that Wnt-3a gene dosage regulates mouse axial development. *Genes Dev.* 10: 313.
- Hartl, D. L., and A. G. Clark, 2007 *Principles of Population Genetics*, Ed. 4. Sinauer Associates, Sunderland, MA.
- Herrmann, B. G., S. Labeit, A. Poustka, T. R. King, and H. Lehrach, 1990 Cloning of the T gene required in mesoderm formation in the mouse. *Nature* 343: 617.
- Huett, A., A. Ng, Z. Cao, P. Kuballa, M. Komatsu *et al.*, 2009 A Novel hybrid yeast-human network analysis reveals an essential role for FNBPI1 in antibacterial autophagy. *J. Immunol.* 182: 4917.
- James, P. J., R. W. Ponzoni, D. R. Gare, and K. S. Cockrum, 1991 Inheritance of short tailedness in South Australian Merinos. *Proceedings of the Australian Association of Animal Breeding and Genetics*, Vol. 9, pp. 404–407.
- Kavka, A. I., and J. B. A. Green, 1997 Tales of tails: Brachyury and the T-box genes. *Biochim. Biophys. Acta* 1333: 73–84.
- Kispert, A., and B. G. Herrmann, 1993 The Brachyury gene encodes a novel DNA binding protein. *EMBO J.* 12: 4898–4899.
- Li, C., S. Huang, S. Mo, N. Zhang, L. Zhou *et al.*, 2015 Susceptibility of autoimmune diseases in three polymorphisms of infection-associated gene IRAK1. *J. Infect. Dev. Ctries.* 9: 614–623.
- Li, H., and R. Durbin, 2010 Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* 26: 589–595.
- Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan *et al.*, 2009 The sequence alignment/map format and SAMtools. *Bioinformatics* 25: 2078–2079.
- Li, M., S. Tian, C. K. L. Yeung, X. Meng, Q. Tang *et al.*, 2014 Whole-genome sequencing of Berkshire (European native pig) provides insights into its origin and domestication. *Sci. Rep.* 4: 4678.
- Liu, Z., H. Wang, R. Liu, W. U. Mingming, S. Zhang *et al.*, 2015 Genome-wide detection of selection signatures of distinct tail types in sheep populations. *Acta Vet. Zootech. Sin.* 46: 1721–1732.
- Megens, H. J., R. P. Crooijmans, C. M. San, X. Hui, N. Li *et al.*, 2008 Biodiversity of pig breeds from China and Europe estimated from pooled DNA samples: differences in microsatellite variation between two areas of domestication. *Genet. Sel. Evol.* 40: 103–128.
- Moon, S., T. H. Kim, K. T. Lee, W. Kwak, T. Lee *et al.*, 2015 A genome-wide scan for signatures of directional selection in domesticated pigs. *BMC Genomics* 16: 1–12.
- Moradi, M. H., A. Nejati-Javaremi, M. Moradi-Shahrbabak, K. G. Dodds, and J. C. McEwan, 2012 Genomic scan of selective sweeps in thin and fat tail sheep breeds for identifying of candidate regions associated with fat deposition. *BMC Genet.* 13: 1–15.
- Müller, C. W., and B. G. Herrmann, 1997 Crystallographic structure of T domain-DNA complex of the Brachyury transcription factor. *Nature* 389: 884–888.
- Palena, C., D. E. Poley, K. Y. Tsang, R. I. Fernando, M. Litzinger *et al.*, 2007 The human T-box mesodermal transcription factor Brachyury is a candidate target for T-cell-mediated cancer immunotherapy. *Clin. Cancer Res.* 13: 2471–2478.
- Pennimpede, T., J. Proske, A. König, J. A. Vidigal, M. Morkel *et al.*, 2012 *In vivo* knockdown of Brachyury results in skeletal defects and urorectal malformations resembling caudal regression syndrome. *Dev. Biol.* 372: 55–67.
- Pourlis, A. F., 2011 A review of morphological characteristics relating to the production and reproduction of fat-tailed sheep breeds. *Trop. Anim. Health Prod.* 43: 1267.
- Rubin, C. J., M. C. Zody, J. Eriksson, J. R. Meadows, E. Sherwood *et al.*, 2010 Whole-genome resequencing reveals loci under selection during chicken domestication. *Nature* 464: 587.
- Sangoi, A. R., J. Karamchandani, B. Lane, J. P. Higgins, R. V. Rouse *et al.*, 2011 Specificity of brachyury in the distinction of chordoma from clear cell renal cell carcinoma and germ cell tumors: a study of 305 cases. *Mod. Pathol.* 24: 425–429.
- Seto, S., K. Sugaya, K. Tsujimura, T. Nagata, T. Horii *et al.*, 2013 Rab39a interacts with phosphatidylinositol 3-kinase and negatively regulates autophagy induced by lipopolysaccharide stimulation in macrophages. *PLoS One* 8: e83324.
- Showell, C., O. Binder, and F. L. Conlon, 2004 T-box genes in early embryogenesis. *Dev. Dyn.* 229: 201.
- Signer-Hasler, H., C. Flury, B. Haase, D. Burger, H. Simianer *et al.*, 2012 A genome-wide association study reveals loci influencing height and other conformation traits in horses. *PLoS One* 7: e37282.
- Smith, J., 1999 T-box genes: what they do and how they do it. *Trends Genet.* 15: 154–158.
- Sonuga-Barke, E. J. S., 2013 Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. *Lancet* 381: 1371–1379.
- Tao, J., K. Hu, Q. Chang, H. Wu, N. E. Sherman *et al.*, 2009 Phosphorylation of MeCP2 at Serine 80 regulates its chromatin association and neurological function. *Proc. Natl. Acad. Sci. USA* 106: 4882–4887.
- Wang, K., M. Li, and H. Hakonarson, 2010 ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* 38: e164.
- Wang, X., G. Zhou, X. Xu, R. Geng, J. Zhou *et al.*, 2014 Transcriptome profile analysis of adipose tissues from fat and short-tailed sheep. *Gene* 549: 252–257.
- Wilm, B., E. Dahl, H. Peters, R. Balling, and K. Imai, 1998 Targeted disruption of Pax1 defines its null phenotype and proves haploinsufficiency. *Proc. Natl. Acad. Sci. USA* 95: 8692–8697.
- Wu, B., Y. Shao, B. Chen, C. Liu, Z. Xue *et al.*, 2010 Identification of a novel mouse brachyury (T) allele causing a short tail mutation in mice. *Cell Biochem. Biophys.* 58: 129–135.
- Xiao, X., S. Xin, X. S. Hu, Z. Yan, Y. C. Liu *et al.*, 2016 Whole genome sequencing identifies a missense mutation in HES7 associated with short tails in Asian domestic cats. *Sci. Rep.* 6: 31583.
- Yang, Q., Z. Qin, W. Li, and B. Xu, 2012 Temporal and spatial variations of vegetation cover in Hulun Buir grassland of inner Mongolia, China. *Arid Land Res. Manage.* 26: 328–343.
- Yuan, Z., E. Liu, Z. Liu, J. W. Kijas, C. Zhu *et al.*, 2017 Selection signature analysis reveals genes associated with tail type in Chinese indigenous sheep. *Anim. Genet.* 48: 55–66.

Communicating editor: D. J. de Koning