

ARTICLE

Genome-wide association study reveals that the *IBSP* locus affects ear size in cattleJiafei Shen^{1,2,6}, Xiaoting Xia^{1,6}, Luyang Sun¹, Xiaohui Ma¹, Bizhi Huang³, Quratulain Hanif⁴, Ningbo Chen¹, Kaixing Qu⁵, Jicai Zhang³, Hong Chen¹, Yu Jiang¹ and Chuzhao Lei¹

© The Author(s), under exclusive licence to The Genetics Society 2023

Ear size is a classical model for hot climate adaptation following the evolution, but the genetic basis of the traits associated with ear size remains to be elucidated. Here, we performed a genome-wide association study on 158 cattle to explain the genetic mechanism of ear size. One region on BTA6 between 36.79 and 38.80 Mb included 50 suggestive SNPs and 4 significant SNPs that were significantly associated with ear size. The most significant locus ($P = 1.30 \times 10^{-8}$) was a missense mutation (T250I) on the seventh exon of integrin-binding sialoprotein (*IBSP*), which had an allele substitution effect of 23.46 cm² for ear size. Furthermore, this mutation will cause changes in the three-dimensional structure of the protein. To further identify genes underlying this typical feature, we performed a genome scan among nine cattle breeds with different ear sizes by using SweeD. Results suggested that *IBSP* was under positive selection among four breeds with relatively large ear sizes. The expression levels of *IBSP* in ear tissues of large- and small-ear cattle were significantly different. A haplotype diversity survey of this missense mutation in worldwide cattle breeds strongly implied that the origin of this missense mutation event was *Bos taurus*. These findings have important theoretical importance for the exploration of major genes associated with ear size and provide important molecular markers for the identification of cattle germplasm resources.

Heredity (2023) 130:394–401; <https://doi.org/10.1038/s41437-023-00614-9>

INTRODUCTION

The ear is primarily responsible for converting sound waves into nerve impulses that are translated in the brain into auditory perception in animals (Van 2003). The central nervous system in the brain then translates these impulses into a spoken language, birds chirping, or the cow's moo. The ear serves other vital functions, like heat radiation (Law et al. 1979), sense of balance (Ekdale 2016; Horwitz et al. 2011; Torres and Giráldez 1998), and external signals in the mood besides auditory perception (Reefmann et al. 2009). The ear consists of three sections, i.e., external, middle, and inner ears (Webster 1966). The external ear includes an external auditory canal (Alvord and Farmer 1997), an eardrum, and a visible pinna that reside outside of the head and can collect sound and direct it down the ear canal. Over time, mammals adapt to their respective environmental conditions and evolve with an extraordinary variety of ear shapes and sizes. For example, animals adapt to high-temperature and hot environments and evolve with large ears to combat heat waves, such as African elephants. African elephants adjust their body temperature through their large ears, providing shade to reduce the loss of water in the body (Phillips and Heath 1992). Furthermore, a general observation includes large ear sizes in indicine cattle, such as Brahman and Burmese cattle, which

harbor in harsh hot environments. By contrast, most animals inhabiting temperate or cold regions, such as the Arctic fox, tend to have small ears. The small size of the ear can help mammals to reduce heat dissipation and maintain their body temperature. Likewise, cattle adapting to cold environments, such as Yanbian, Yakut, and Tibetan cattle, share small ears.

Size and orthotropism are important conformational characteristics of ears with huge diversity in different species. With the development of the whole-genome resequencing technology, the genome-wide association study (GWAS) has been applied to reveal SNPs associated with complex ear traits in livestock. Previous studies showed that significant quantitative trait loci (QTLs) for pig ear size are located on *Sus scrofa* chromosomes (SSCs) 5 and 7 (Ma et al. 2009; Wei et al. 2007). A missense mutation (*G32E*) in the *PPARD* gene on SSC7 is considered to be responsible for the differences in porcine ear size (Ren et al. 2011; Wilkinson et al. 2013). The QTL fine mapping on SSC5 revealed that *LEMD3* is one of the most important candidates for porcine ear size (Zhang et al. 2014). Moreover, GWAS revealed that *DCC*, *SOX5*, and *PTPRD* are potential candidate genes for ear size in sheep (Gao et al. 2018). Furthermore, genome fragments containing *WIF1* and *HMG2* appear to control ear size in pigs and dogs (Vaysse et al. 2011).

¹Key Laboratory of Animal Genetics, Breeding and Reproduction of Shaanxi Province, College of Animal Science and Technology, Northwest A&F University, Yangling, Shaanxi 712100, China. ²International Institutes of Medicine, The Fourth Affiliated Hospital, Zhejiang University School of Medicine, N1 Shangcheng Road, Yiwu, Zhejiang 322000, China. ³Yunnan Academy of Grassland and Animal Science, Kunming, Yunnan 650212, China. ⁴Computational Biology Laboratory, Agricultural Biotechnology Division, National Institute for Biotechnology and Genetic Engineering, Faisalabad 577, Pakistan. ⁵Academy of Science and Technology, Chuxiong Normal University, Chuxiong, Yunnan 675000, China. ⁶These authors contributed equally: Jiafei Shen, Xiaoting Xia. Associate editor Christine Baes ✉email: yu.jiang@nwafu.edu.cn; leichuzhao1118@nwafu.edu.cn

Received: 17 February 2022 Revised: 16 March 2023 Accepted: 17 March 2023

Published online: 4 April 2023

However, no current report exists on the ear size of cattle, but the available resources of cattle are rich worldwide.

Cattle are not only important for their meat and milk production but also represent an important large-animal model for the study of some important economic traits (Herath et al. 2006). Ear sizes among cattle breeds are diverse. *Bos indicus* breeds, such as Brahman cattle, have large and lop ears, whereas *Bos taurus* breeds have small and erect ears. Overall, ear size is an important distinguishing characteristic among subspecies. China has a wide geographical latitude span, with taurine and indicine cattle breeds distributed from north to south, and is a natural cattle breed resource bank for studying ear size. Yunling cattle is a typical hybrid cow with variable ear size, and therefore, it is an ideal model for studying the genetic basis of ear size traits. In the current study, GWAS is performed to screen potential candidate genes, which are further systematically analyzed to identify the associated loci and the candidate mutation in a particular gene or region. Our findings provide new insights into the genetic basis of the ear size of cattle and provide important molecular markers for the identification of cattle germplasm resources.

MATERIALS AND METHODS

Ethics statement

In accordance with the recommendation of the Regulations for the Administration of Affairs Concerning Experimental Animals of China, the Institutional Animal Care and Use Committee of Northwest A&F University approved all animal experiments.

Sample collection and genome sequencing

Individuals used for ear size trait comprised 158 adult female cattle aged around 2 years old (39 Brahman and 119 Yunling cattle). The Yunling cattle, a typical hybrid cattle breed, was bred by the Yunnan Academy of Grassland and Animal Science. Ear tissue samples were collected from adult females. The standard phenol–chloroform protocol was used to extract the genomic DNA from the ear tissue (Green et al. 2012). A total of 158 paired-end libraries with an insert size of 350 bp were constructed and sequenced using the Illumina NovaSeq. The length of the reads was 150 bp. The sequence data used in this paper were obtained from published papers where detailed information about sampling and sequencing was available (Chen et al. 2020).

Measurement and calculation of ear size

In this study, we used the “living pixel method” to collect the ear shape photo with a standard scale of 475 cattle (39 Brahman, 119 Yunling, 65 Simmental, 52 Burmese, 100 Wenshan, and 100 Dabieshan cattle). All individuals were female, aged 2 years, and had similar body sizes in one breed. The head of the cattle was restrained, and the scale-labeled left ear was flattened for the picture. Digital images were taken using a high-pixel camera in a horizontal position. The ear size of each animal was measured using Photoshop CS6 (Adobe, USA). First, the graphs of the ear range and 1 cm^2 in the scale were obtained through the “magnetic lasso” tool, and pixel numbers corresponding to the above parameters were obtained through the “histogram” tool. The ear size could be obtained through the “measuring record” tool and the pixel ratio column.

Read mapping and SNP calling

Default parameters were used to map clean reads to the cattle reference assembly ARS-UCD1.2 (Rosen et al. 2020) by using BWA-MEM (Li and Durbin 2009). Duplicate reads were filtered using the “REMOVE_DUPLICATES = true” option of Picard tools. The “Haplotype-Caller”, “GenotypeGVCFs”, and “SelectVariants” arguments of the Genome analysis toolkit 3.8 (Nekrutenko and Taylor 2012) were used for calling raw SNPs. The average alignment rate and coverage were 99.54% and 5.61×, respectively. The argument “VariantFiltration” of the same software was applied to all raw SNPs with the following options: DP < 303 (1/3-fold total sequence depth for all individuals), DP > 2727 (threefold of total sequence depth for all individuals), QD < 2, FS > 60, MQ < 40.0, MQRankSum < -12.5, ReadPosRankSum < -8.0, and 1/3× < mean sequence depth (for all individuals) < 3×. In addition, the

haplotype-phase inference and missing allele imputation were produced using Beagle v.4.1 (Browning and Browning 2007) to carry out the GWAS further. Based on about 41M autosomal SNPs, we estimated the eigenvectors by using the smartPCA of the EIGENSOFT v5.0 package (Patterson et al. 2006) to adjust the population structure in GWAS. The principal component 1 based on the genotype matrix separated the Brahman cattle from Yunling cattle, and this finding was in accordance with those in a previous study (Chen et al. 2020).

GWAS analysis

Based on 158 sequenced genomes, 13,057,965 SNPs (MAF > 0.10, missing rate > 0.1) were used in GWAS for the ear size trait. The primary association analysis was carried out using the genome-wide efficient mixed-model association (GEMMA) software package (Zhou and Stephens 2012). The mixed linear model assumed the following model:

$$y = X\alpha + S\beta + K\mu + \varepsilon$$

where y is a vector of phenotypes, α is a vector of fixed effects representing marker effects, β is a vector of fixed effects representing nonmarker effects, and μ is a vector of unknown random effect. X , S , and K represent the incidence matrices relating α , β , and μ , respectively, and ε represents a vector of random residual effects. The top three PCs and feeding regimes were defined as the S matrix. The kinship matrix calculated from nucleotide polymorphism was defined as the K matrix.

The secondary association analysis used a multiple linear regression model using PLINK v. 1.90b6.21 (Purcell et al. 2007), combining body height, cross high, head length, head width, and top three PCs as covariates to perform GWAS on the ear size phenotype. The number of linkage disequilibrium (LD)-pruned SNPs (750,367) was defined as the effective number of independent SNPs and calculated using PLINK (-indep-pairwise 50 5 0.2) to estimate the correction required for multiple testing. Therefore, the significance and suggestive threshold were defined as approximately 5×10^{-8} (0.05/750,367) and 1×10^{-6} (1/750,367), respectively. Thresholds were widely used by numerous studies.

Identification of candidate genes in the GWAS-associated loci

We used the following strategy to narrow down our findings. First, PLINK (Purcell et al. 2007) was used to estimate candidate regions by using the pairwise LD correlation ($r^2 > 0.6$) between SNPs related to ear size characteristics, and the count of suggestive SNPs < 3 was used as false positives and removed. Second, SNPs (leading SNPs) with P_{wald} values < 1×10^{-6} were characterized as candidate SNPs, and the functional annotation of suggestive associated SNPs was carried out in accordance with the *Bos taurus* reference genome in the package ANNOVAR version (Wang et al. 2010).

Selective sweep analysis

The genomes of groups with big and small ears were compared on the basis of the phenotype relationship of ear size to identify the signatures of positive selection. First, the genome-wide distribution of F_{ST} values was estimated using the VCFtools (Danecek et al. 2011) with 100 kb window size and 50 kb increment to investigate pairwise genetic differentiation among groups with big and small ears. To further identify the selection signals within each cattle breed, we used SweeD, an open-source tool for the rapid detection of selective sweeps in whole genomes (Pavlidis et al. 2013). We also calculated the SweeD for 14 cattle breeds (i.e., Angus, Brahman, Burmese, Hanwoo, Hereford, Holstein, Kazakh, Ji'an, Longlin, Mishima, Mongolian, Muturu, Simmental, and Yanbian cattle) by using the SweepFinder2 (DeGiorgio et al. 2016) with a sliding window approach (100 kb windows with 50 kb increments). Only the strong regions overlapping both methods (SweeD: top 1%, F_{ST} : top 1%) were defined as candidate regions under positive selection. A custom perl script was used to annotate the regions under selection on the basis of the *Bos taurus* reference genome (ARS-UCD1.2).

RNA extraction and quantitative PCR

Ear tissue samples from ten Simmental cattle with small ear sizes and ten Brahman cattle with large ear sizes were used for qRT-PCR analysis. Total RNA was extracted using the PrimeScript™ RT reagent Kit with gDNA Eraser (Takara, Beijing, China). The first-strand cDNA was synthesized from 1 µg of total RNA using the Omniscript reverse transcriptase kit (Qiagen, Germany). qPCR was performed in a 20 µL reaction volume containing 2 µL of 3-fold diluted cDNA, 10 µL of 2× Power SYBR Green Master mix (Applied

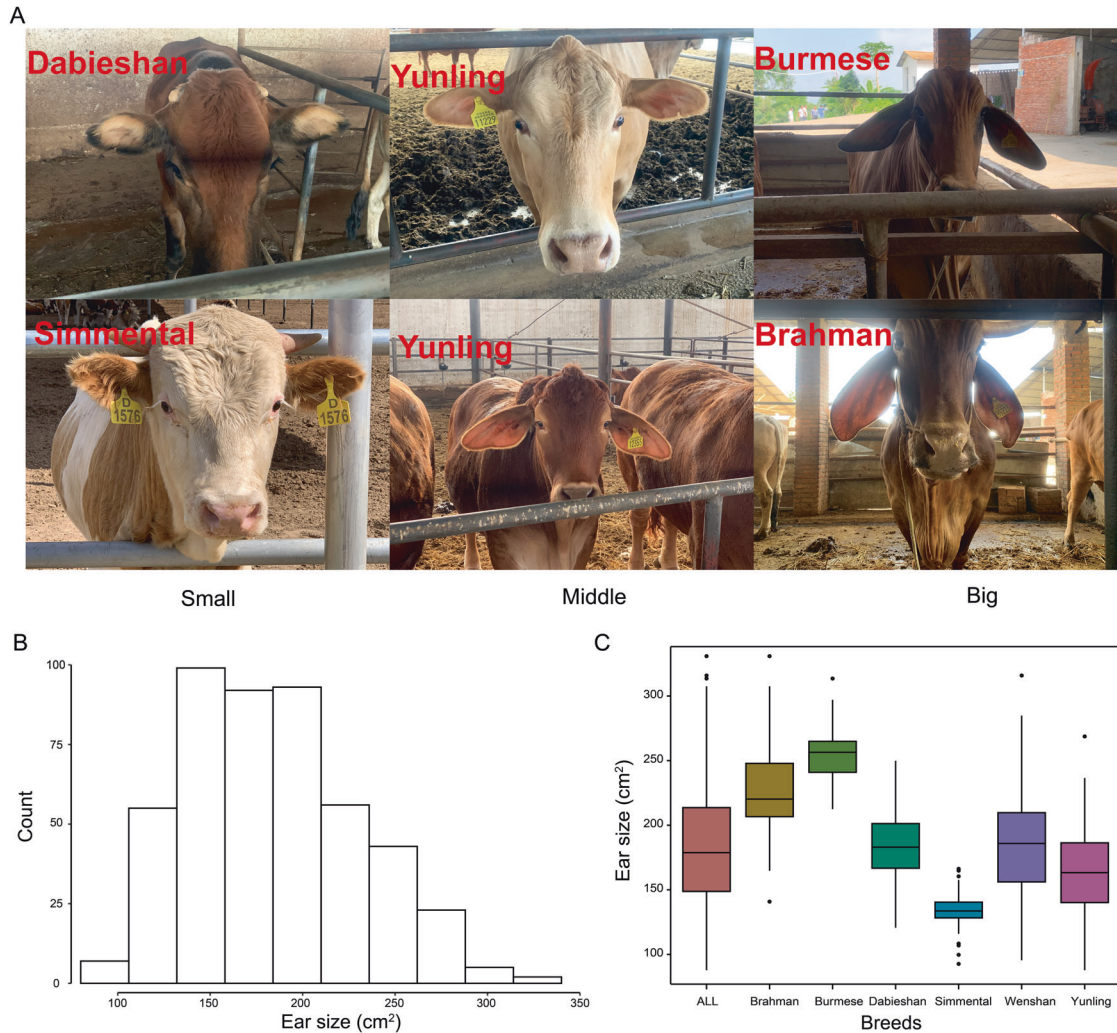


Fig. 1 Statistical analysis of ear size. **A** Ear diversity of different cattle breeds, the first column was smaller ear size cattle (Simmental and Dabieshan cattle), and the second column represents hybrid cattle, while the third column shows the large ear size *Bos indicus* cattle (Brahman and Burmese cattle). **B** Histogram plot for 475 ear size data, the vertical axis on the left represents number of individuals in different sections. **C** Boxplot analysis for ear size among all individuals and six cattle breeds.

Biosystems, USA), and 150 nM of each primer, with the following thermocycling conditions: 10 min at 95 °C followed by 40 cycles at 95 °C for 15 s, 60 °C for 1 min and 72 °C for 30 s. Then, a dissociation curve analysis was performed. β -actin was used as an internal control (Livak and Schmittgen 2002). Gene-specific primers and probes were designed using the primer premier 5.0 software (forward primer CAGCAGCAACAGCACA-GAGG and reverse primer CATTGGTGCCTGTTGTTTCAT). qPCR was performed using the 7500 Fast real-time PCR System (Applied Biosystems, USA). The experiment was repeated three times for each group.

RESULTS

Ear size variation trait analyses

Ear sizes among different cattle individuals and breeds are evidently different. As a general observation, the ear size of *Bos indicus* in the Indo-Pak region was larger than that of *Bos taurus* in the world, whereas hybrid cattle breeds depicted diversified ear size (Fig. 1A). According to this observational phenomenon, ear size data were collected and calculated from 475 adult female cattle (Supplementary Table S1), including six representative breeds (i.e., Burmese, Brahman, Simmental, Yunling, Wenshan, and Dabieshan cattle) by using the pixel method (Yang et al. 2013). The overall ear size data ranged from 87.81 to 330.85 cm². *Bos indicus* breeds (i.e., Brahman and Burmese cattle) had the

largest mean values, which were 226.75 and 254.06 cm², respectively. By contrast, *Bos taurus* (i.e., Simmental cattle) had the smallest mean value, which was 134.62 cm². The sizes of hybrid breed (i.e., Yunling cattle) and Chinese indicine (i.e., Wenshan and Dabieshan cattle) cattle were relatively scattered. The Yunling, Wenshan, and Dabieshan cattle had average values of 164.21, 186.31, and 184.37 cm², respectively (Table 1). The ear size distribution of all individuals showed a unimodal distribution according to the histogram and density plot (Fig. 1B) and could be well distinguished in three distinct intervals, i.e., big (>230 cm²), middle (<230 and >148 cm²), and small (<148 cm²) ears with 25 and 75% quantiles of all individuals' boxplot analysis. The boxplot and *F* statistic were used to analyze the variance in the ear size among six cattle breeds. Results showed that the ear size of breeds followed the order: *Bos indicus* > hybrid (Yunling) and Chinese indicine (Wenshan and Dabieshan cattle) > *Bos taurus* (Fig. 1C). Moreover, the *F* statistic results suggested that ear sizes among cattle breeds were significantly different, supporting the above statement.

Genome-wide association studies for ear size trait

Based on 13,057,965 autosomal SNPs derived from 158 published resequencing cattle data (Supplementary Table S2) (Chen et al. 2020),

Table 1. Descriptive statistics of ear size in cattle breeds.

Breeds	Origin	Number	Maximum	Minimum	Mean	CV (%)
Simmental	<i>Bos taurus</i>	65	166.26	92.67	134.62	10.53
Brahman	<i>Bos indicus</i>	39	330.85	140.86	226.75	15.68
Burmese	<i>Bos indicus</i>	52	313.53	212.46	254.06	8.50
Yunling	<i>Bos indicus</i> × <i>Bos taurus</i>	119	268.71	87.81	164.21	18.88
Wenshan	Chinese indicine	100	315.87	95.47	186.31	22.28
Dabie mountain	Chinese indicine	100	249.86	120.63	184.37	13.75

CV coefficients of variation.

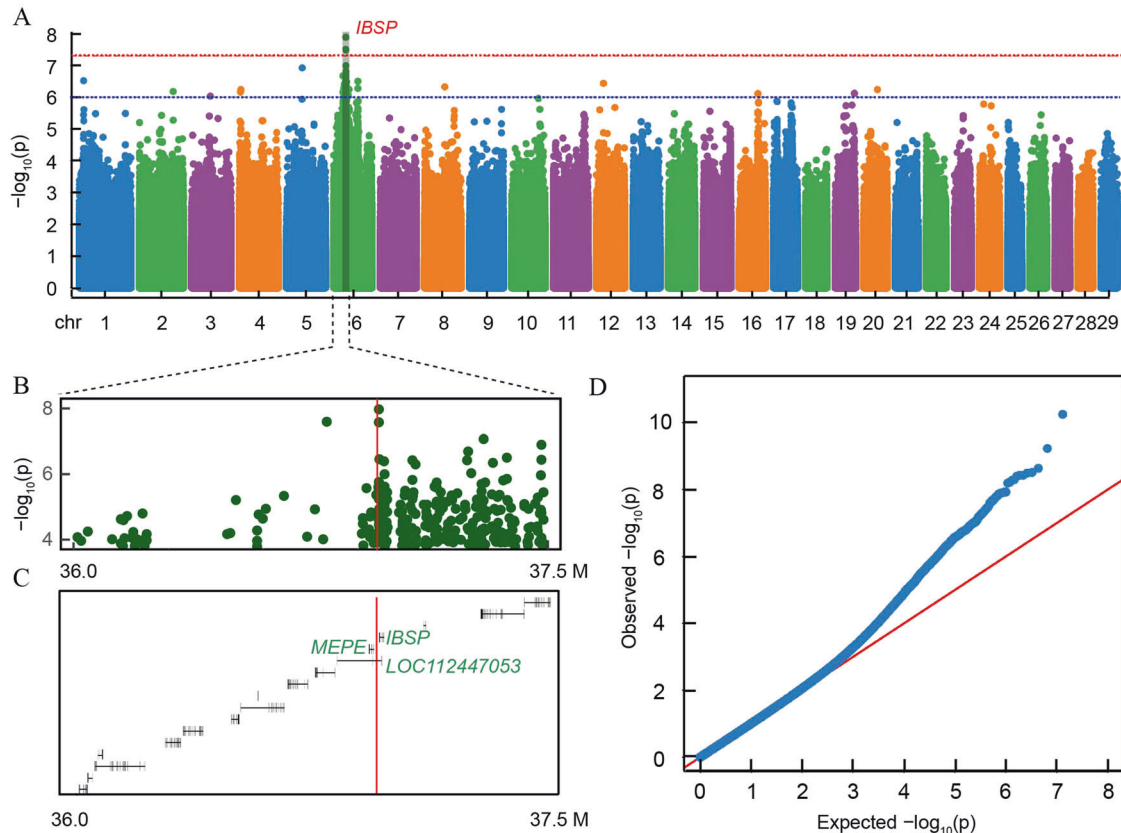


Fig. 2 Genome-wide association study of ear size. **A** Manhattan plot for ear size GWAS, **B, C** plot for regional association result for *IBSP*, the functional genes in this region were plotted in the box, and **D** Quantile–Quantile plot observed and expected p value (expressed as $-\log_{10}(p)$) of the GWAS for ear size.

a GEMMA (Zhou and Stephens 2012) was used in the primary GWAS to identify the significant loci. Figure 2A shows the Manhattan plot for the GWAS. Red and blue horizontal lines represent the Bonferroni-adjusted genome-wide significant and suggestive threshold. A total of 4 significant SNPs and 50 suggestive SNPs showed genome-wide associations with ear size (Supplementary Table S3). However, most of the significant potential SNPs were located on BTA6 (36.79–38.80 Mb) (Fig. 2B). The significant locus ($P = 1.30 \times 10^{-8}$) was located in the integrin-binding sialoprotein (*IBSP*) (Fig. 2C). The Quantile–Quantile plot (QQ-plot) in Fig. 2D showed the observed and expected P values of GWAS for ear size. The red line represents the distribution of SNPs under the null hypothesis of no association of SNPs with the trait of interest. The strong deviation of the observed from the expected P values for QQ-plots indicated more SNPs significantly associated with all ear size traits than would be expected by chance.

Considering that body size traits and ear size data could share an underlying genetic basis, the correlation was measured and analyzed for each trait in Supplementary Fig. S1. The Pearson

correlation coefficient was used to assess the linear relationships between ear size and body size traits. Results showed that the body height, cross high, head length, and head width had weakly-to-moderately positive correlations with the ear size trait. The multiple linear regression model was analyzed using these related traits as covariates for the secondary GWAS by using PLINK (Purcell et al. 2007). The autosomal SNP scan for ear size revealed associated markers (Supplementary Fig. S2). A total of 293 SNPs were observed on the potential region, most of which were found on BTA6. The most significant SNP ($P = 5.74 \times 10^{-11}$) was located on *IBSP*. Two GWAS results strongly suggested that *IBSP* had a strong correlation with ear size and might be a key candidate gene influencing cattle ear size.

Genetic differentiation of the mutation in *IBSP* among groups with different ear sizes

In order to further screen candidate genes for ear size differentiation, genetic differentiation between Yunling cattle

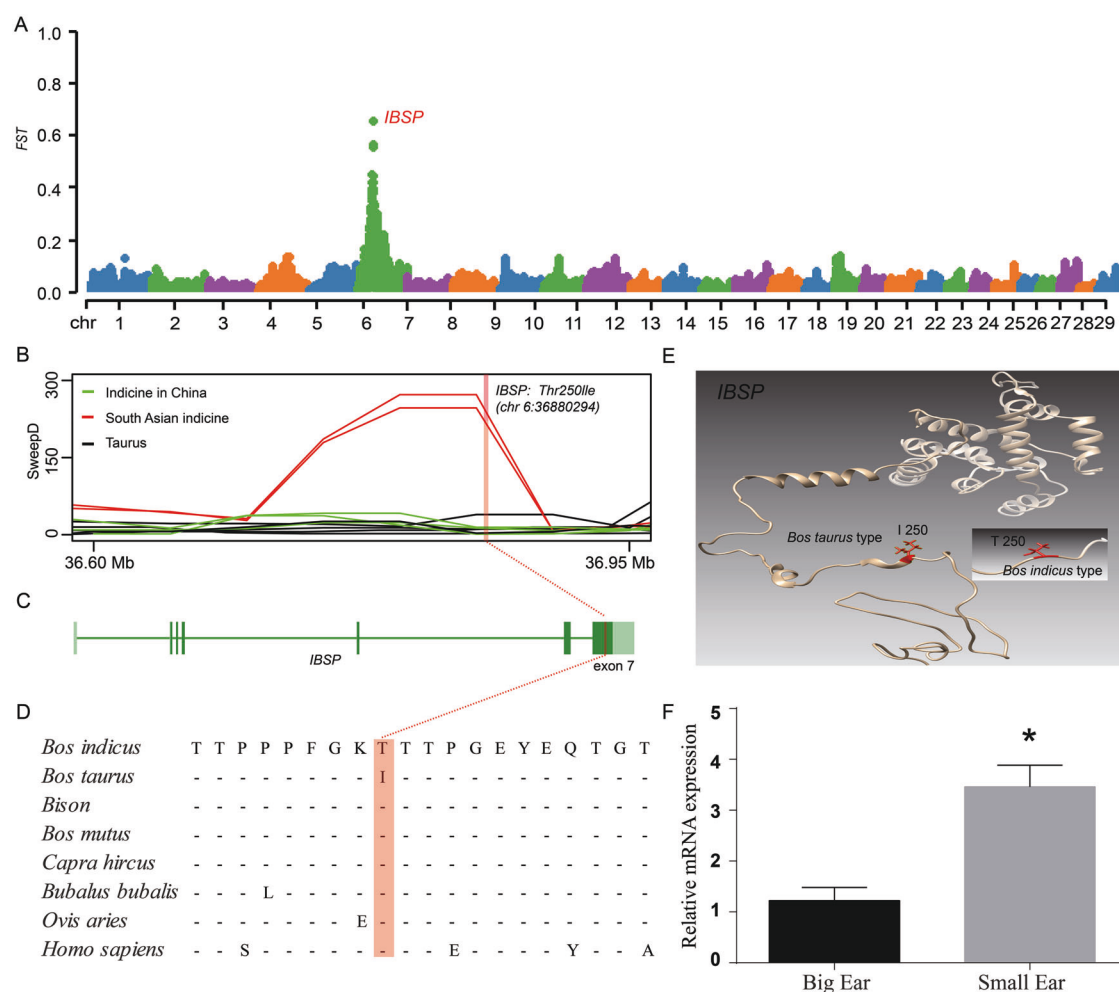


Fig. 3 Selection and conservative analysis for IBSP. A Manhattan plot for F_{st} analysis of big and small ear among Yunling cattle group. The red box indicated the location of *IBSP*. **B** Line charts of SweepD in 9 cattle breeds on the gene regions of *IBSP*. **C** Gene structure of *IBSP*. **D** Conservative prediction of the mutant *IBSP* protein with its ortholog proteins in Bovidae. **E** 3D protein structure prediction of *IBSP*. **F** The mRNA expression of *IBSP* between big- and small-ear groups.

populations with ear sizes larger than 230 cm² and smaller than 148 cm² was measured using a pairwise fixation index (F_{st}) (Weir and Cockerham 1984). The two populations of Yunling cattle had the same genetic background and were similar, but the difference in ear size was significant. By annotating significant regions, the most significant loci ($F_{st} = 0.68$) (Fig. 3A) were located on the seventh exon of *IBSP* (Fig. 3C). At the same time, we also conducted similar genetic differentiation index analysis for different cattle breeds with significant differences in ear size, namely, two Zebu populations living in India and southern China, both belonging to Zebu with similar genetic background and significant differences in ear size. By annotating significant regions ($F_{st} > 0.29$, empirical $P < 0.005$), *IBSP* was also found in the most significant region on BTA6 (position: 36,840,001–36,890,000, $F_{st} = 0.43$) (Supplementary Fig. S3). Nine cattle breeds were used for Sweep Detector (SweepD) (Pavlidis et al. 2013) analyses to further investigate the differentiation of these mutation loci across diverse cattle breeds (Supplementary Fig. S4). Results showed that the SweepD value on these loci was observed to be near zero in *Bos taurus* (i.e., Angus, Kazakh, Hanwoo, Hereford, Holstein, Mishima, Simmental, and Yanbian cattle) and over 200 in South Asian indicine cattle (i.e., Brahman and Burmese cattle), indicating that *IBSP* might be selected in indicine breeds (Fig. 3B). In accordance with the functional annotation, a missense

mutation was found on the seventh exon of *IBSP* (Threonine-250→Isoleucine, T250I) (Fig. 3C). To further evaluate the functional impact of the variants, we aligned the mutant *IBSP* protein with its ortholog proteins in Bovidae (Fig. 3D) and other diverse vertebrates (Supplementary Fig. S5). The comparison revealed that T250 was a conserved amino acid mutation and varied among all the other animals we examined except *Bos taurus* (T250I), lesser mouse-deer (T250S), and white-lipped deer (T250A). We also have counted PhastCons and PhyloP of this mutation from UCSC. The score of 100 vertebrates conservation by PhastCons was 1, and the score of 100 vertebrates Basewise Conservation by PhyloP was 5.44393. These three mutant species had small ears, and this finding was consistent with our present hypothesis. We calculated the LD values of the SNPs, which were shared in the region of *IBSP*. We observed a strong linkage in this region (Supplementary Fig. S5). Furthermore, the prediction-based 3D structure of the protein predicted by the website <https://seq2fun.dcmdb.med.umich.edu/I-TASSER/> showed that the mutation site of *IBSP* could change its protein structure (Fig. 3E). Next, we compared the total mRNA levels of *IBSP* between small and big ear groups by using quantitative real-time PCR. Results showed that the mRNA expression of *IBSP* was significantly different (Fig. 3F). Taken together, our results suggest that T250I was a plausible candidate mutation for the *IBSP* sweep in cattle ear size.

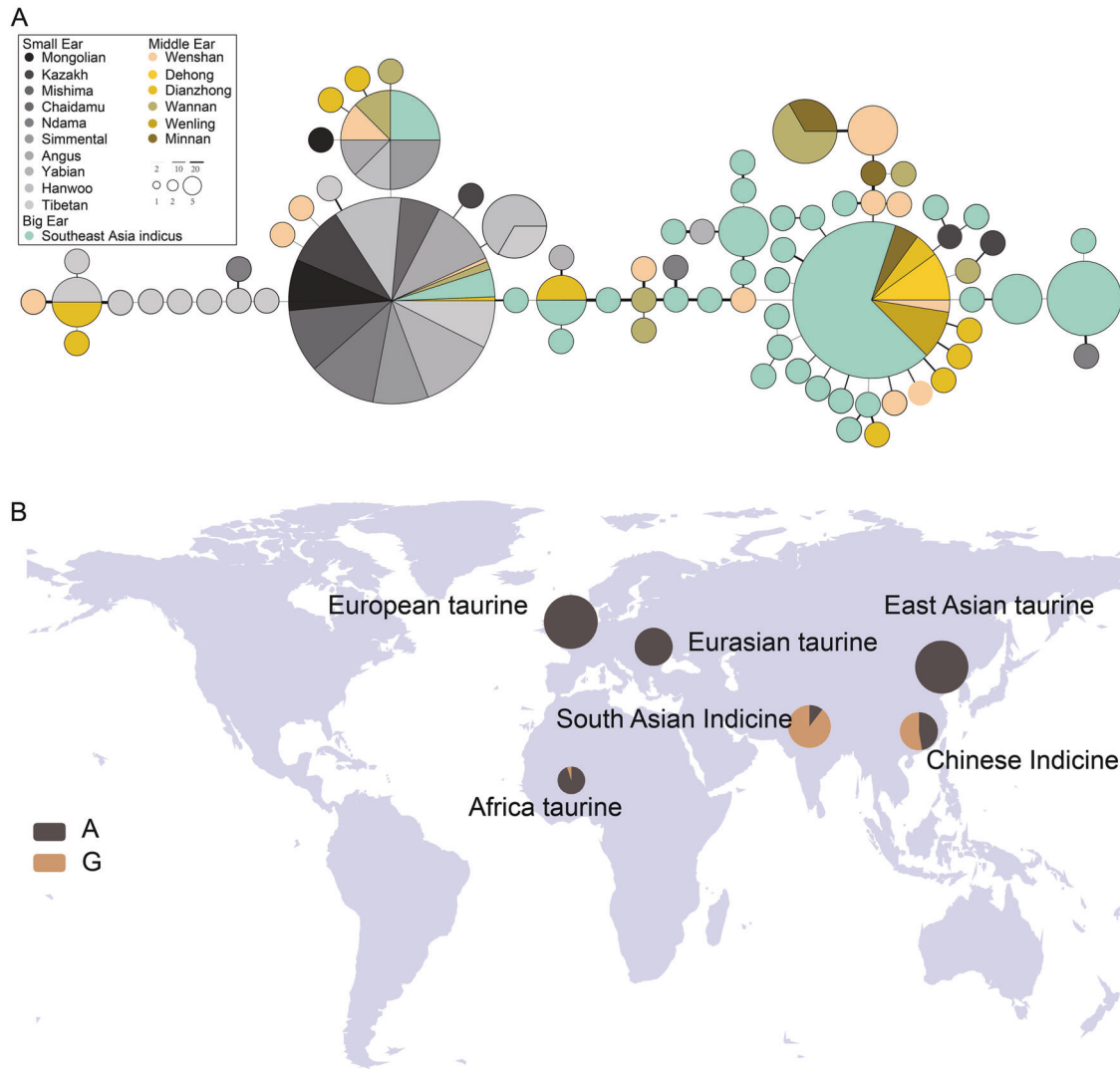


Fig. 4 Haplotype network analysis and geographical distribution of the mutation among different cattle breeds. A Haplotype network analysis for the surrounding region of *IBSP* mutation site among different cattle breeds. **B** The mutation allele frequency distribution of six ancestral cattle groups. The size of the circle represents the count of the samples. The orange and brown colors represent G and A alleles, respectively.

Allele frequency of *IBSP* mutation among different cattle breeds indicating the origin

In this study, a region with strong linkage (BTA6: 36,851,626–36,929,220) was selected from the surrounding region of *IBSP* mutation site for haplotype network analysis, and the results showed that all samples formed two obvious branches. Cattle breeds with small ear sizes (Yanbian, Hanwoo, Mishima, Tibetan cattle, etc.) and cattle breeds (Brahman, Burma, Sahiwal, Nelore, Tharparkar, and Gir cattle) with larger ear sizes were clearly divided into one branch, respectively, while the breeds (Wannan, Wenshan, Wenling, Dehong, Dianzhong, and Minnan cattle) with middle ear size were mixed in these two branches (Fig. 4A). At the same time, we genotyped the DNA sequences of *IBSP* with 394 cattle representing *Bos indicus*, *Bos taurus*, and hybrid cattle to investigate the genotype frequency of the mutation loci across diverse cattle breeds. Two genotypes (A and G) of the *IBSP* mutation loci were found (Supplementary Table S4). The A allele of *IBSP* mutation occurred at low frequency in Burmese (0.06), Brahman (0.30), and Lingnan (0.40) cattle. In Wenshan, Shigatse, Weining, Hainan, Dianzhong, Jinjiang, Ji'an, Wannan, and Luxi cattle, the A allele of the two mutations showed a moderate frequency close to 0.5. By contrast, the A allele was almost equal to 1 in Yanbian, Kazakh, and Simmental cattle and other *Bos taurus*. At the same time,

we searched the mutation frequency of *IBSP* among different cattle breeds in the world from the Bovine Genome Variation Database and Selective Signatures (<http://animal.nwsuaf.edu.cn/code/index.php/BosVar>) (Chen et al. 2019). The database contained 24 South Asian indicine, 19 Chinese indicine, 37 East Asian taurine, 38 European taurine, 19 Eurasian taurine, and 10 Africa taurine cattle. Results showed that the A allele frequency of *IBSP* was lowest in South Asian indicine cattle (0.10), and the frequency was 0.47 in Chinese indicine cattle. However, the A frequency of *IBSP* was near or equal to 1 in East Asian, European, and Eurasian taurine cattle (Fig. 4B). The allele distribution indicated that the A allele of *IBSP* might have originated in *Bos taurus*. Given the infiltration of South Asian indicine cattle, the Chinese indicine cattle caused a reduction in the A allele frequency. We also found the G allele in African taurine cattle at low frequency, which might be due to the large introduction of South Asian indicine cattle after the African rinderpest in 1890 (Phoofolo 1993).

DISCUSSION

The external ear of cattle is composed of cartilage and thin layers of skin. The function of cartilage in the external ear is a stent, which determines the size and shape of the ear (Bos et al. 2018).

This is different from the candidate genes that have been reported to affect ear size in pigs, dogs, sheep, and other species, such as *PPARD*, *LEMD3*, *DCC*, *SOX5*, *PTPRD*, *WIF1*, *HMG2A*, etc. In this study, we identified *IBSP* as the main candidate gene affecting bovine ear size, and it is acceptable that the same trait is regulated by different candidate genes in different species. There are great genetic differences among different species, and ear size, as a quantitative trait, is regulated by multiple genes. In this study, two GWAS methods were used to identify the gene, and the hypothesis that the gene underlies the ear phenotype was supported by structural analysis. The *IBSP* has been mentioned many times in numerous studies of cartilage and chondrocytes (Fukui et al. 2008; Komori 2010, 2017, 2018; Lui et al. 2019). The main *IBSP*-knockout mice show a short body, reduced chondrocyte proliferation, and impaired cartilage absorption. At the same time, the lack of *IBSP* can alter bone growth and the formation and mineralization of primary bone (Bouletour et al. 2014). The protein encoded by *IBSP* is the main structural protein in the bone matrix, and nearly 12% non-collagenous proteins are from *IBSP* in human osteoblasts, hypertrophic chondrocytes, and osteoclasts (Duncan et al. 2011; Kim et al. 1994; Wuttke et al. 2001). *IBSP* is related to tissue development and cell growth, which may relate to its mechanism in promoting or inhibiting the growth and proliferation of bovine ear chondrocytes.

We have discovered and validated one missense mutation at *IBSP* in *Bos taurus*, of which T250I occurs in a well-defined domain and quite conserved site (Supplementary Fig. S5). In nearly 70 vertebrate species except *Bos taurus* (T250I), lesser mouse-deer (T250S), and white-lipped deer (T250A), this mutation is conserved. We have found this missense mutation in *IBSP* that alters the 3D structure of the expressed protein, which may change the function of the protein. The analysis of the mRNA expression suggests that *IBSP* is highly expressed in small-ear cattle breeds and minimally expressed in big-ear cattle breeds, which may also be caused by this mutation. At the same time, the statistical analysis of the *IBSP* missense genotype among different cattle breeds shows that the GG genotype is presented in *Bos indicus* with a large ear size and that the AA genotype is presented in *Bos taurus*, which has a small ear size. Three genotypes are presented in hybrid cattle, which lives in the south of China and has complex ear sizes. Through the mutation frequency of the *IBSP* missense variant among different cattle breeds, we have found that the lowest frequency of the A allele is in South Asian indicine cattle and that the frequency is near or equal to 1 in East Asian, European, and Eurasian taurine cattle. Moreover, the mutations of the A allele are found in cattle with small ear sizes living in a cold environment, whereas the relatively large ear size of cattle living in tropical or subtropical regions has almost no mutation frequency. We speculate that the missense mutation of *IBSP* may be the main factor that affects the cattle ear size and that the mutations of the A allele in *IBSP* may originate from *Bos taurus*. The difference in ear size may be due to the variable environmental conditions in which they are located. The large ear area can help *Bos indicus* to dissipate heat temperature, and a small ear size may be helpful for *Bos taurus* to keep their temperature in a cold environment. The temperature in southern China is not high or low. Thus, hybrid cattle have diverse ear-size phenotypes. Meanwhile, the mutation site had a higher genetic differentiation index in the populations with significant ear size, which means that ear size may also be influenced by species differentiation. In conclusion, the *IBSP* loci related to cattle ear size is identified by GWAS, and the missense mutation in *IBSP* (T250I) is speculated to be the candidate mutation that affects ear size among cattle breeds. These findings have important theoretical significance for the exploration of major genes of ear size traits and provide important molecular markers for the identification of cattle germplasm resources.

DATA AVAILABILITY

The raw whole-genome sequencing data were reported in our previous study (Chen et al. 2020) and are available at the NCBI Short Read Archive under the BioProject accession number PRJNA555741.

REFERENCES

- Alvord LS, Farmer BL (1997) Anatomy and orientation of the human external ear. *J Am Acad Audiol* 8:383–390
- Bos EJ, Pluemeekers M, Helder M, Kuzmin N, van der Laan K, Groot M-L et al. (2018) Structural and mechanical comparison of human ear, alar, and septal cartilage. *Plast Reconstr Surg Glob Open* 6:e1610
- Bouletour W, Boudiffa M, Wade-Gueye NM, Bouët G, Cardelli M, Laroche N et al. (2014) Skeletal development of mice lacking bone sialoprotein (BSP)-Impairment of long bone growth and progressive establishment of high trabecular bone mass. *PLoS One* 9(5):e95144
- Browning SR, Browning BL (2007) Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *Am J Hum Genet* 81(5):1084–1097
- Chen N, Fu W, Zhao J, Shen J, Chen Q, Zheng Z et al. (2020) BGVD: Integrated Web-database for Bovine Sequencing Variations and Selective Signatures. *Genom. Proteom. Bioinform* 18(2):186–193
- Chen Q, Zhan J, Shen J, Qu K, Hanif Q, Liu J et al. (2020) Whole-genome resequencing reveals diversity, global and local ancestry proportions in Yunling cattle. *J Anim Breed Genet* 137(6):641–650
- Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA et al. (2011) The variant call format and VCFtools. *Bioinformatics* 27(15):2156–2158
- DeGiorgio M, Huber CD, Hubisz MJ, Hellmann I, Nielsen R (2016) SweepFinder2: increased sensitivity, robustness and flexibility. *Bioinformatics* 32(12):1895–1897
- Duncan EL, Danoy P, Kemp JP, Leo PJ, McCloskey E, Nicholson GC et al. (2011) Genome-wide association study using extreme truncate selection identifies novel genes affecting bone mineral density and fracture risk. *PLoS Genet* 7(4):e1001372
- Ekdale EG (2016) Form and function of the mammalian inner ear. *J Anat* 228(2):324–337
- Fukui N, Miyamoto Y, Nakajima M, Ikeda Y, Hikita A, Furukawa H et al. (2008) Zonal gene expression of chondrocytes in osteoarthritic cartilage. *Arthritis Rheum* 58(12):3843–3853
- Gao L, Xu SS, Yang JQ, Shen M, Li MH (2018) Genome-wide association study reveals novel genes for the ear size in sheep (*Ovis aries*). *Anim Genet* 49(4):345–348
- Green MR, Sambrook J (2012) Molecular cloning: a laboratory manual 4th ed. Cold Spring Harbor Laboratory Press: New York, NY, USA
- Herath S, Dobson H, Bryant C, Sheldon I (2006) Use of the cow as a large animal model of uterine infection and immunity. *J Reprod Immunol* 69(1):13–22
- Horwitz GC, Risner-Janiczek JR, Jones SM, Holt JR (2011) HCN channels expressed in the inner ear are necessary for normal balance function. *J Neurosci* 31(46):16814–16825
- Kim RH, Shapiro HS, Li JJ, Wrana JL, Sodek J (1994) Characterization of the human bone sialoprotein (BSP) gene and its promoter sequence. *Matrix Biol* 14(1):31–40
- Komori T (2010) Regulation of bone development and extracellular matrix protein genes by RUNX2. *Cell Tissue Res* 339(1):189–195
- Komori T (2017) Roles of Runx2 in Skeletal Development. *Adv Exp Med Biol* 962:83–93
- Komori T (2018) Runx2, an inducer of osteoblast and chondrocyte differentiation. *Histochem Cell Biol* 149(4):313–323
- Law MP, Ahier R, Field S (1979) The effect of prior heat treatment on the thermal enhancement of radiation damage in the mouse ear. *Br J Radiol* 52(616):315–321
- Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 25(14):1754–1760
- Livak KJ, Schmittgen TD (2002) Analysis of relative gene expression data using real-time quantitative PCR. *Methods* 25(4):402–408
- Lui JC, Yue S, Lee A, Kikani B, Temycky A, Barnes KM et al. (2019) Persistent Sox9 expression in hypertrophic chondrocytes suppresses transdifferentiation into osteoblasts. *Bone* 125:169–177
- Ma J, Qi W, Ren D, Duan Y, Qiao R, Guo Y et al. (2009) A genome scan for quantitative trait loci affecting three ear traits in a White Duroc Chinese Erhualian resource population. *Anim Genet* 40(4):463–467
- Nekrutenko A, Taylor J (2012) Next-generation sequencing data interpretation: enhancing reproducibility and accessibility. *Nat Rev Genet* 13(9):667–672
- Patterson N, Price AL, Reich D (2006) Population structure and eigenanalysis. *PLoS Genet* 2(12):e190
- Pavlidis P, Živković D, Stamatakis A, Alachiotis N (2013) SweeD: likelihood-based detection of selective sweeps in thousands of genomes. *Mol Biol Evol* 30(9):2224–2234

- Phillips PK, Heath JE (1992) Heat exchange by the pinna of the African elephant (*Loxodonta africana*). *Comp Biochem Physiol Comp Physiol* 101(4):693–699
- Phoofolo P (1993) Epidemics and revolutions: the rinderpest epidemic in late nineteenth-century Southern Africa. *Past Present* 138:112–143
- Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D et al. (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 81(3):559–575
- Reefmann N, Kaszás FB, Wechsler B, Gygas L (2009) Ear and tail postures as indicators of emotional valence in sheep. *Appl Anim Behav Sci* 118(3–4):199–207
- Ren J, Duan Y, Qiao R, Yao F, Zhang Z, Yang B et al. (2011) A missense mutation in PPAR α causes a major QTL effect on ear size in pigs. *PLoS Genet* 7(5):e1002043
- Rosen BD, Bickhart DM, Schnabel RD, Koren S, Elsik CG, Tseng E et al. (2020) De novo assembly of the cattle reference genome with single-molecule sequencing. *Gigascience* 9(3):giaa021
- Van Hemmen JL (2003) Auditory system. *Biol. Cybern* 89:317
- Torres M, Giráldez F (1998) The development of the vertebrate inner ear. *Mech Dev* 71(1–2):5–21
- Vaysse A, Ratnakumar A, Derrien T, Axelsson E, Pielberg GR, Sigurdsson S et al. (2011) Identification of genomic regions associated with phenotypic variation between dog breeds using selection mapping. *PLoS Genet* 7(10):e1002316
- Wang K, Li M, Hakonarson H (2010) ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res* 38(16):e164
- Webster DB (1966) Ear structure and function in modern mammals. *Am Zool* 6(3):451–466
- Wei W, De Koning D, Penman J, Finlayson H, Archibald A, Haley C (2007) QTL modulating ear size and erectness in pigs. *Anim Genet* 38(3):222–226
- Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. *Evolution* 38(6):1358–1370
- Wilkinson S, Lu ZH, Megens H-J, Archibald AL, Haley C, Jackson IJ et al. (2013) Signatures of diversifying selection in European pig breeds. *PLoS Genet* 9(4):e1003453
- Wuttke M, Müller S, Nitsche DP, Paulsson M, Hanisch F-G, Maurer P (2001) Structural characterization of human recombinant and bone-derived bone sialoprotein functional implications for cell attachment and hydroxyapatite binding. *J Biol Chem* 276(39):36839–36848
- Yang H, Zhang X, Hou L, Hong Y, Ren J, Guo Y (2013) Measurement of the ear area of pig in vivo with the pixel method. *Acta Agric Univ Jiangxi* 35(5):1024–1029
- Zhang L, Liang J, Luo W, Liu X, Yan H, Zhao K et al. (2014) Genome-wide scan reveals LEMD3 and WIF1 on SSC5 as the candidates for porcine ear size. *PLoS One* 9(7):e102085
- Zhou X, Stephens M (2012) Genome-wide efficient mixed-model analysis for association studies. *Nat Genet* 44(7):821–824

ACKNOWLEDGEMENTS

We would like to thank High-Performance Computing (HPC) of Northwest A&F University (NWAUFU) for providing computing resources. The study was supported by the National Natural Science Foundation of China (31872317) and China Agriculture Research System of MOF and MARA (Grant No. CARS-37).

AUTHOR CONTRIBUTIONS

YJ and CZL designed the study and supported the funding. JFS and XTX curated and analyzed the data. JFS wrote the original manuscript. QH reviewed and edited the manuscript. LYS, XHM, NBC, and HC reviewed the manuscript. BZH, KXQ, and JCZ organized sampling and conducted fieldwork. All authors commented on the manuscript and gave final approval for publication.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41437-023-00614-9>.

Correspondence and requests for materials should be addressed to Yu Jiang or Chuzhao Lei.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.