ANIMAL GENETICS Immunogenetics, Molecular Genetics and Functional Genomics

SHORT COMMUNICATION

doi: 10.1111/age.12670

Genome-wide association study reveals novel genes for the ear size in sheep (*Ovis aries*)

L. Gao*†a, S.-S. Xu‡§a, J.-Q. Yang*†a, M. Shen*† and M.-H. Li‡§

*Institute of Animal Husbandry and Veterinary Medicine, Xinjiang Academy of Agricultural and Reclamation Sciences, Shihezi 832000, China. †State Key Laboratory of Sheep Genetic Improvement and Healthy Breeding, Xinjiang Academy of Agricultural and Reclamation Sciences, Shihezi 832000, China. ‡CAS Key Laboratory of Animal Ecology and Conservation Biology, Institute of Zoology, Chinese Academy of Sciences (CAS), Beijing 100101, China. §College of Life Sciences, University of Chinese Academy of Sciences (UCAS), Beijing 100049, China.

Summary

Variations in ear size can be observed in livestock such as sheep; however, the genetic basis of variable ear size in sheep is still poorly understood. To investigate causative genes associated with ear size in sheep, a genome-wide association study was performed in 115 adult Duolang sheep with different-sized floppy ears using the Ovine Infinium HD BeadChip. We found 38 significant SNPs at the genome-wide or chromosome-wise 5% significance level after Bonferroni correction. The most significant association ($P = 1.61 \times 10^{-6}$) was found at SNP rs402740419, located in the *DCC* gene, which plays a critical role in ear development. Also, we observed two additional significant SNPs, rs407891215 in *PTPRD* and rs407769095 in *SOX5*, both of which are functionally associated with ear developmental processes. Our results are useful for future sheep breeding and provide insights into the genetic basis of ear size development in sheep and other livestock.

Keywords candidate genes, Duolang sheep, ear development, Ovine Infinium HD BeadChip

Sheep (*Ovis aries*) is not only an agriculturally important animal for meat, wool and milk production but also an important large-animal model for human disease, microtia (Jawasreh *et al.* 2016). As part of the auditory system, the external ear plays a vital role in collecting sound as the first step in hearing (Ren *et al.* 2011). Also, ear size is relevant to other physiological functions (e.g., olfactory) and environmental adaptation (Alasti & Van Camp 2009). For example, animals with large ears lose more heat because they have a larger surface area (Blaxter *et al.* 1959). To date, most genetic studies in sheep have targeted economically important traits, such as meat, milk, wool and reproduction, but few investigations have focused on the non-economic characteristics such as ear size. Duolang sheep, a Chinese native breed from the southern region of Xinjiang province, has exceptionally large and floppy ears with a large variation in ear size (Fig. 1a), and provides an excellent model for investigating the genetic basis underlying ear size development in sheep and other mammals.

In this study, we collected data from a total of 115 unrelated adult Duolang sheep (63 ewes and 52 rams) from the Aksu region of Xinjiang province, China. We measured the ear size of each animal by tracing the shape of each ear on plotting paper (Zhang *et al.* 2014). Tissue samples were collected and genomic DNA was extracted using a standard phenol/chloroform method (Köchl *et al.* 2005). The extracted DNA was diluted to 50 ng/μl for genotyping.

Genotyping for all the samples was conducted using the Illumina Ovine Infinium HD SNP BeadChip and yielded a dataset of 606 006 SNPs (genotype and phenotype datasets are available at https://www.animalgenome.org/share/tmp/FJN1519605929.rar.gz). We then implemented quality control on this SNP dataset using PLINK v.1.07 software (Purcell *et al.* 2007). SNPs or individuals meeting any of the following criteria were removed from further analyses: (i) no chromosomal or physical location, (ii) minor allele frequency less than 0.05, (iii) individual call rate less than

Address for correspondence

M.-H. Li, CAS Key Laboratory of Animal Ecology and Conservation Biology, Institute of Zoology, Chinese Academy of Sciences, Beichen West Road No. 1-5, Chaoyang District, Beijing 100101, China.
E-mail: menghua.li@ioz.ac.cn
and
M. Shen, Institute of Animal Husbandry and Veterinary Medicine, Xinjiang Academy of Agricultural and Reclamation Sciences, Shihezi 832000, China.
E-mail: shenmin0993@sina.com

[a]These authors contributed equally to this work.

Accepted for publication 14 March 2018

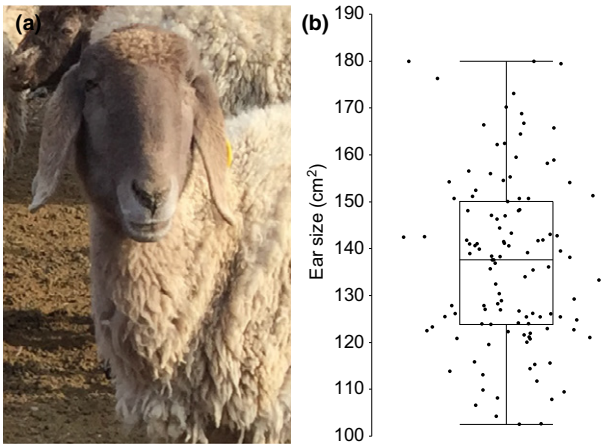

Figure 1 (a) Duolang sheep. (b) A box-dotplot of ear size of each individual in the Duolang sheep population studied. The box indicates the 25th, 50th, and 75th percentiles, whereas the bottom and top lines indicate minimum and maximum values respectively.

0.95, (iv) missing genotype frequency for SNPs more than 0.05 and (v) *P*-value for Hardy-Weinberg equilibrium less than 0.001 (Peng *et al.* 2017). Also, we implemented sex screening using the 'check-sex' option in PLINK v1.07. After filtering, a total of 498 909 SNPs and 115 individuals were retained in the working dataset for the genome-wide association study (GWAS).

We performed the GWAS using a two-step approach via the general linear model and genome-wide efficient mixed-model analysis (GEMMA) as detailed by Xu *et al.* (2017). The first seven principal components from a multi-dimensional scaling analysis were used as covariates to account for the biases caused by the population stratification in the GWAS (Xu *et al.* 2017). The models are detailed below.

(1) Residual calculation:

$$Y = \mu + \boldsymbol{X}s + \boldsymbol{T}c + e, \qquad \text{model (1)}$$

where y is the area of ear trait, μ is overall population mean, s is sex effect (male and female are 1 and 2 respectively), c is the first seven principal components effect, e is the residual and $\boldsymbol{X}$ and $\boldsymbol{T}$ are the corresponding matrix vectors.

(2) Genome-wide efficient mixed-model analysis (GEMMA):

$$e = \mu^* + \boldsymbol{G}k + e^*, \qquad \text{model (2)}$$

where e from the residual in model (1) is taken as the phenotype, μ^* is mean value of the population, k is the SNP effect, $\boldsymbol{G}$ is the genotype matrix vector and e^* is the residual. Further, we performed pairwise tests of linkage disequilibrium between the target SNPs and their flanking SNPs within approximately 1 Mb upstream and downstream using PLINK v1.07. Regional association plots were generated using R v3.2.2 (http://www.r-project.org).

Ear size varied from 102.55 to 179.95 cm^2 with a mean value equal to 137.15 cm^2 and a coefficient of variation equal to 132.75% (Fig. 1b). We observed a genomic control factor lambda close to 1.00, suggesting well-controlled population stratification, as indicated in the quantile–quantile (Q–Q) plot (Fig. S1). In the GWAS, we identified

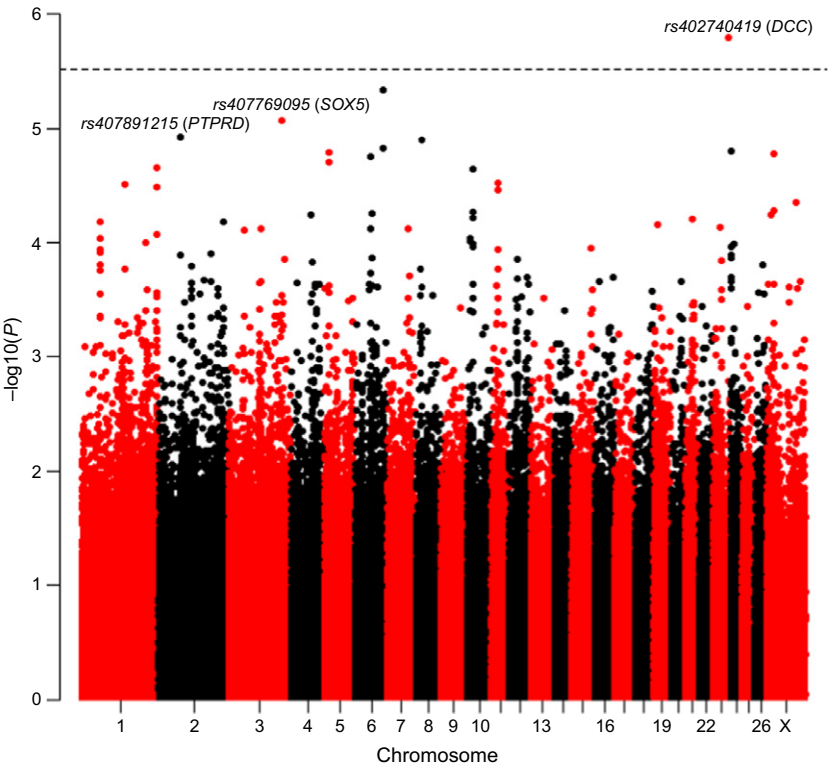

Figure 2 Manhattan plot of genome-wide association study. The 5% genome-wide significant threshold value after Bonferroni correction ($P < 3.0 \times 10^{-6}$) is indicated by the dashed line.

Table 1 Genome-wide and chromosome-wise significant SNPs associated with the ear size in sheep.

Chr.	SNP	Position	*P*-value	Relevant genes
1	rs412709734	268826726	2.19E-05	/
2	rs407891215	77506632	1.19E-05	*PTPRD*
3	rs407769095	190172938	8.53E-06	*SOX5*
5	rs426157058	14056565	1.61E-05	/
5	rs409318795	14058707	1.97E-05	/
6	rs419602052	98893802	4.65E-06	*CDS1*
6	rs407927499	98809086	1.49E-05	*CDS1*, *NKX6-1*
6	rs425276090	54563840	1.79E-05	/
6	rs426546475	54567720	1.79E-05	/
6	rs403626070	54621581	1.79E-05	/
6	rs420696987	54631007	1.79E-05	/
6	rs420816139	58462609	5.59E-05	*KLB*
6	rs420267941	58469389	5.59E-05	*KLB*
6	rs419111284	54730804	7.49E-05	/
8	rs404265961	24231329	1.25E-05	/
10	rs426442861	19289402	2.25E-05	*CAB39L*
10	rs400985598	19295808	5.42E-05	*CAB39L*
10	rs428617163	19268683	6.13E-05	*CAB39L*
10	rs400764128	8454210	9.27E-05	/
10	rs411021804	8452774	9.68E-05	/
11	rs403548001	23400675	2.99E-05	*RAP1GAP2*
11	rs400982572	23397576	3.49E-05	*RAP1GAP2*
11	rs402636338	23409928	1.14E-04	*RAP1GAP2*
19	rs427733907	12189222	7.01E-05	*SCN11A*
21	rs425831823	23301139	6.25E-05	*NELL1*
23	rs402740419[1]	52464180	1.61E-06	*DCC*
23	rs401042464	25841996	7.33E-05	*DSG2*
23	rs161364326	25842069	7.33E-05	*DSG2*
24	rs426458593	2885150	1.56E-05	*NAA60*
24	rs410447958	11949057	1.04E-04	*SHISA9*
24	rs416611112	2620760	1.10E-04	/
24	rs400373831	2892191	1.29E-04	*C24H16orf90*
24	rs424908971	2606388	1.37E-04	/
26	rs417053761	29881524	1.56E-04	*UNC5D*
X	rs425764457	23447275	1.65E-05	/
X	rs408624046	104283220	4.43E-05	*DCAF12L2*
X	rs410535498	23446965	5.33E-05	/
X	rs430752624	11310724	5.65E-05	*GPM6B*

/, no gene identified or annotated for the SNP.
[1]Genome-wide significant SNP after Bonferroni correction.

one SNP (rs402740419) at the genome-wide level ($P < 3.0 \times 10^{-6}$) and 38 SNPs at the chromosome-wise level (Table S1) 5% significance level after Bonferroni correction (Fig. 2, Table 1). The most significant SNP, rs402740419, is located within the *DCC* gene (Fig. 2), which is associated with ear development (Matilainen *et al.* 2007). Also, two additional SNPs, rs407891215 and rs407769095, were significant at the chromosome level and are located within genes *PTPRD* and *SOX5* (Fig. 2) respectively, both of which are functionally associated with ear development (Girotto *et al.* 2011; Edea *et al.* 2017). We did not find strong linkage disequilibrium between the significant SNPs and their flanking SNPs (Fig. 3), which indicates that the target SNPs could be the causative SNPs for ear size or due to very old selection events. Of the genome-wide and chromosome-wise significant SNPs (Fig. 2, Table 1), we did not observe significant association with genes *TRNAR-UCU*, *GATA6* and *MIB1*, which have been suggested to be associated with microtia in Awassi sheep (Jawasreh *et al.* 2016).

The *DCC* gene (*deleted in colorectal cancer*) encodes a netrin receptor, which is involved in mediating the transition from proliferation to terminal differentiation in various tissues (Keino-Masu *et al.* 1996). This gene is expressed in different parts of the otic epithelium and plays a critical role in ear development by co-functioning with the *netrin 1* gene, which has been shown to be involved in inner ear semicircular duct formation (Matilainen *et al.* 2007). The *PTPRD* (*protein tyrosine phosphatase, receptor type D*) gene belongs to a member of the protein tyrosine phosphatase (PTP) family, which regulates a variety of cellular processes including cell growth, differentiation, mitotic cycle and oncogenic transformation (Girotto *et al.* 2011). This gene has been identified as playing an important role in inner ear development and hearing function (Girotto *et al.* 2011, 2014). *SOX5* (*SRY-box 5*) encodes a protein that belongs to the SOX (SRY-related HMG-box) transcription factor family, which contributes to the activation of the chondrocyte

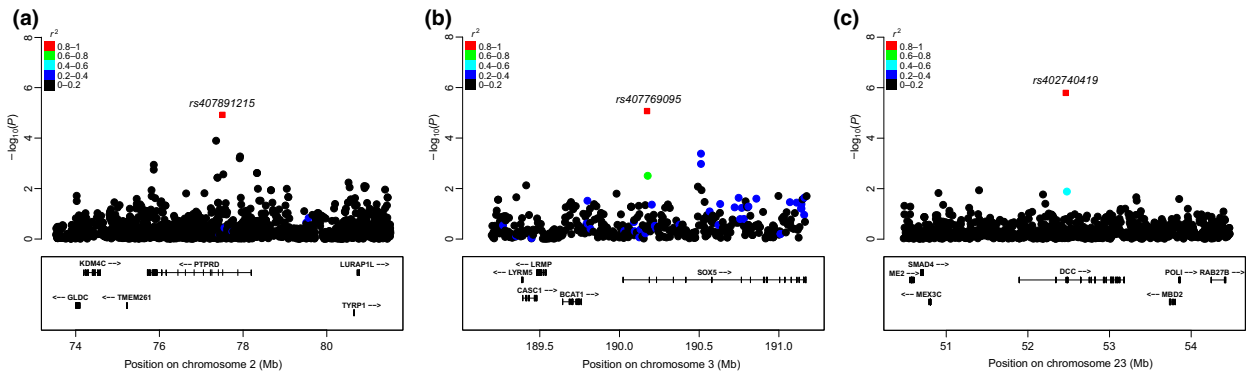

Figure 3 Plots of regional association results for (a) rs407891215, (b) rs407769095 and (c) rs402740419 (red squares). Different colours represent the r^2 values of pair-wise linkage disequilibrium estimates, and functional genes in this region are plotted in the box.

program (Lefebvre *et al.* 1998; Smits *et al.* 2001). Thus, *SOX5* plays an important physiological role in chondrogenesis and cartilage formation and could be a strong candidate gene for the size and morphology of ears in sheep (Li *et al.* 2012; Edea *et al.* 2017).

In conclusion, we identified for the first time three novel candidate genes (*DCC*, *PTPRD* and *SOX5*) for the variation in ear size of sheep. Our results will help molecular-based breeding programs in sheep and also provide insights into the genetic mechanisms of microtia in human. Future investigations based on comparisons between breeds displaying different ear size would be particularly relevant.

Acknowledgements

This work was supported by the Special Project for Innovation of Genetic Resources & Exploration and Utilization of Functional Genes of the Xinjiang Production and Construction Corps (2012BB044) and the Academician Foundation Program of the Xinjiang Production and Construction Corps (2015AG019).

References

Alasti F. & Van Camp G. (2009) Genetics of microtia and associated syndromes. *Journal of Medical Genetics* **46**, 361–9.

Blaxter K., Graham N. & Wainman F. (1959) Environmental temperature, energy metabolism and heat regulation in sheep. III. The metabolism and thermal exchanges of sheep with fleeces. *Journal of Agricultural Science* **52**, 41–9.

Edea Z., Hong J.K., Jung J.H., Kim D.W., Kim Y.M., Kim E.S., Shin S.S., Jung Y.C. & Kim K.S. (2017) Detecting selection signatures between Duroc and Duroc synthetic pig populations using high-density SNP chip. *Animal Genetics* **48**, 473–7.

Girotto G., Pirastu N., Sorice R. *et al.* (2011) Hearing function and thresholds: a genome-wide association study in European isolated populations identifies new loci and pathways. *Journal of Medical Genetics* **48**, 369–74.

Girotto G., Vuckovic D., Buniello A., Lorente-Canovas B., Lewis M., Gasparini P. & Steel K.P. (2014) Expression and replication studies to identify new candidate genes involved in normal hearing function. *PLoS ONE* **9**, e85352.

Jawasreh K., Boettcher P.J. & Stella A. (2016) Genome-wide association scan suggests basis for microtia in Awassi sheep. *Animal Genetics* **47**, 504–6.

Keino-Masu K., Masu M., Hinck L., Leonardo E.D., Chan S.S., Culotti J.G. & Tessier-Lavigne M. (1996) *Deleted in colorectal cancer* (*DCC*) encodes a netrin receptor. *Cell* **87**, 175–85.

Köchl S., Niederstätter H. & Parson W. (2005) DNA extraction and quantitation of forensic samples using the phenol–chloroform method and real-time PCR. *Methods in Molecular Biology* **297**, 13–30.

Lefebvre V., Li P. & de Crombrugghe B. (1998) A new long form of *Sox5* (L-Sox5), *Sox6* and *Sox9* are coexpressed in chondrogenesis and cooperatively activate the type II collagen gene. *EMBO Journal* **17**, 5718–33.

Li P., Xiao S., Wei N., Zhang Z., Huang R., Gu Y., Guo Y., Ren J., Huang L. & Chen C. (2012) Fine mapping of a QTL for ear size on porcine chromosome 5 and identification of *high mobility group AT-hook 2* (*HMGA2*) as a positional candidate gene. *Genetics Selection Evolution* **44**, 6.

Matilainen T., Haugas M., Kreidberg J.A. & Salminen M. (2007) Analysis of netrin 1 receptors during inner ear development. *International Journal of Developmental Biology* **51**, 409–13.

Peng W.F., Xu S.S., Ren X. *et al.* (2017) A genome-wide association study reveals candidate genes for the supernumerary nipple phenotype in sheep (*Ovis aries*). *Animal Genetics* **48**, 570–9.

Purcell S., Neale B., Todd-Brown K., Thomas L., Ferreira M.A., Bender D. & Sham P.C. (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics* **81**, 559–75.

Ren J., Duan Y., Qiao R. *et al.* (2011) A missense mutation in *PPARD* causes a major QTL effect on ear size in pigs. *PLoS Genetics* **7**, e1002043.

Smits P., Li P., Mandel J., Zhang Z., Deng J.M., Behringer R.R., de Crombrugghe B. & Lefebvre V. (2001) The transcription factors L-Sox5 and Sox6 are essential for cartilage formation. *Developmental Cell* **1**, 277–90.

Xu S.S., Ren X., Yang G.L. *et al.* (2017) Genome-wide association analysis identifies the genetic basis of fat deposition in the tails of sheep (*Ovis aries*). *Animal Genetics* **48**, 560–9.

Zhang L.C., Liang J., Luo W.Z., Liu X., Yan H., Zhao K.B., Shi H.B., Zhang Y.B., Wang L.G. & Wang L.X. (2014) Genome-wide scan reveals *LEMD3* and I on SSC5 as the candidates for porcine ear size. *PLoS ONE* **9**, e102085.

Supporting information

Additional supporting information may be found online in the supporting information tab for this article:

Figure S1 Q–Q (quantile–quantile) plot. Grey and black rings represent association statistics before and after correction for population stratification respectively.

Table S1 Chromosome-wise significance threshold at the 5% level after Bonferroni correction.