Genome-Wide Association Scan Suggests Basis for Microtia in Awassi Sheep

K. Jawasreh*, P. J. Boettcher[†,1] A. Stella[‡,§]

*Animal Breeding and Genetics, Department of Animal Production, Faculty of Agriculture, Jordan University of Science and Technology (JUST), Irbid, Jordan

[†]Animal Genetic Resources Branch, Animal Production and Health Division, Agriculture and Consumer Protection Department, Food and Agriculture Organization of the United Nations (FAO), 00153 Rome, Italy

[‡]Parco Tecnologico Padano, 26900 Lodi, Italy

[§]Institute of Agricultural Biology and Biotechnology, National Research Council, Milan, Italy

[1]Corresponding author

Email: paul.boettcher@fao.org

Telephone: +39 06 570 55620

Summary

Microtia, the underdevelopment of the pinna, the structural part of the outer ear, has been observed in many species, including humans, mice, dogs and various livestock. Microtia is relatively prevalent in sheep, observed in multiple breeds including some entire populations, but its genetic basis has not been described. The Awassi sheep, a breed native to Southwest Asia, carries this phenotype and was targeted for molecular characterization via a genome-wide association study. DNA samples were collected from sheep flocks in Jordan, within the native range of the Awassi. Samples from eight affected and twelve normal individuals were genotyped with the Illumina OvineSNP50 ® chip. Haplotype-based analyses failed to identify any major runs of homozygosity associated with the trait. In contrast, a single-locus genome-wide association analysis revealed a solitary statistically significant association ($P = 0.012$, genome wide) with a single-nucleotide polymorphism at base-pair 34,647,499 on OAR23. This marker is adjacent to the gene encoding transcription factor GATA-6, which has been shown to play a critical role in many developmental processes, including chondrogenesis. The lack of extended homozygosity in this region suggests a fairly ancient mutation, and the time of occurrence was estimated to be approximately 860 generations ago. This result suggests that many of the sheep breeds showing this phenotype may share the causative mutation, especially within the sub-group of fat-tailed, wool sheep.

Key words

Introduction

Microtia, the underdevelopment of the pinna, the structural part of the outer ear, has been observed in many species, including humans, mice, dogs and various livestock. Microtia is phenotypically similar to anotia, the condition for which the outer ear is completely absent. Microtia is relatively prevalent in sheep, observed in multiple breeds including some entire populations (e.g. Hernandez, 2013; Mason, 1980; Porter, 2002; Ritzman, 1920). Microtia exists in various forms and is observed both as an apparently isolated trait and as a single characteristic of a more comprehensive syndrome of genetic defects (Harris *et al.*, 1996). In humans, the causative genes for many microtia syndromes have already been identified (Cox *et al.*, 2014), but the molecular genetic basis has not been clarified for the isolated form of the defect, although the Homeobox A2 (*HOXA2*) gene is a strong candidate for some instances of nonsydromic microtia (Brown *et al.* 2013). In sheep, microtia generally seems to be an isolated condition and no studies on the molecular basis of this defect have been reported in the literature.

The Awassi sheep, a breed native to Southwest Asia, is among the breeds in which microtia is observed. The Awassi is a fat-tailed, wool sheep that is primarily used for milk production, particularly the "improved" Awassi strain. The Awassi is the most numerous and most-important sheep breed in Jordan and throughout its native region (Talafha and Ababneh, 2011) and its high milk production has resulted in a much wider international distribution.

As with other breeds, the microtia observed in the Awassi seems to be of the isolated type and affected animals seem to have no other major associated defective trait. Nevertheless, livestock keepers have a preference for the normal, long-eared phenotype. This preference is reportedly

due to a combination of their physical appearance and the fact that earless sheep tend to be more nervous and flighty, perhaps due to decreased auditory acuity, a common occurrence with both isolated and syndromic microtia in humans (Luquetti *et al*., 2012). Microtia in the Awassi appears to be a simply inherited trait. As with other sheep breeds showing this trait, ear size is not a continuously distributed trait, but rather has three distinct phenotypes, normal, earless and small-eared (Figure 1). Researchers have speculated that multiple breeds share the same causative mutation (Lush, 1930). In some breeds, this trait has been shown to follow a classical Mendelian inheritance pattern for a trait affected by a single gene, with normal and earless phenotypes representing homozygous extremes and the short-eared phenotype occurring in heterozygous individuals (Ritzman, 1920; Lush, 1930). This mode of inheritance has not been confirmed in the Awassi.

The recent development of dense arrays of single nucleotide polymorphisms (SNP) has facilitated the characterization of simply-inherited traits, including genetic effects (Charlier *et al.*, 2008). The primary objective of this study was to undertake a genome-wide association study to identify genomic region(s) associated with microtia in the Awassi. In addition, the genes identified were evaluated to determine the most plausible site of the causative mutation. Finally, the age of the mutation was estimated according the distribution of alleles at flanking loci, with the aim to provide additional information on the question of a shared genetic basis for microtia in multiple sheep breeds.

Materials and Methods

DNA samples were collected from 20 sheep flocks in northern Jordan, within the native range of the Awassi. Blood was collected from 49 animals, including 19 normal (long-eared), 22 short-eared and eight earless individuals. The blood samples were stored at -20°C until DNA extraction, which was performed by using the E.Z.N.A® Blood DNA kit (OMEGA Bio-tek, Norcross, GA, USA). The quality of the DNA was tested using 1.5% agarose gel electrophoresis and the concentration was determined by spectrophotometry.

Samples from eight earless and 12 normal individuals were genotyped with the Illumina OvineSNP50 ® chip. Edits removed all SNP with call rates <0.90 and minor allele frequency <0.05. In addition, only those SNP with genotypes for all 20 individuals were retained. These edits resulted in a data set of 27,998 SNP for analysis.

The genome-wide association was undertaken based on the concepts of homozygosity mapping. Briefly, the genotypes of earless animals were scanned to identify the length (in SNP) of the longest run of homozygosity in common among all individuals. Then, the "--hap-assoc" command of the PLINK software (Purcell *et al.*, 2007) was used in a case-control design to test the significance of associations of the earless phenotype with the most probable configuration of haplotypes of length L at all locations throughout the genome. This procedure was progressively repeated for all haplotype lengths from L-1 down to 2. In addition, the "--fisher" command of PLINK (Fisher's exact test) was used to test the association of all single SNP with the earless trait. For both the haplotype-based and single SNP analysis, the Bonferroni correction was applied to account for multiple testing.

Upon identification of statistically significant (P < 0.05, genome-wide) associations, the NCBI Map Viewer tool was used to identify the genes in the corresponding genomic region and the most probable candidate gene was determined based on a literature review of gene functions. Finally, the age of the putative mutation (in generations) was estimated by applying a modified version of the approach developed by Genin *et al.* (2004). Briefly, this approach assumes that individuals in a population affected by a given simply-inherited genetic disease or defect have all descended from a common ancestor in which the causative mutation initially occurred. Age of the mutation is estimated based on the lengths of haplotypes surrounding the mutation that are shared by the affected individuals and the rate of recombination between the adjacent loci constituting the haplotypes.

Results and Discussion

A simple scan of genotypes failed to identify any major runs of homozygosity in perfect association with the trait. The length of the longest haplotype shared in a homozygous manner by all earless animals was nine SNP. This haplotype occurred on Chromosome 7 (OAR7), comprising the region from base-pair 61,904,814 (SNP = s51631.1) to 62,972,902 (SNP = OAR7_62972902.1). However, this haplotype was also the most common among normal, long-eared animals, albeit at a smaller frequency (0.42). This difference was not statistically significant (P > 0.05, genome-wide), nor was this the 9-SNP window with the greatest association with the earless phenotype. Table 1 has the results of significance tests for associations between the earless phenotype and haplotypes ranging in length from nine SNP down to two. As indicated in the table, no statistically significant associations were observed. In fact, with the exception of haplotype lengths six and seven, the genomic region with the greatest

association changed with each 1-SNP decrease in haplotype length, providing further evidence for a lack of a true association with any haplotype block.

The situation changed, however, with the single-SNP analysis (Table 1). A statistically significant (genome wide P-value = 0.0124) was observed with an SNP at location 34,647,499 on OAR23 (i.e. OAR23_34647499.1). The alleles at this loci did not segregate in complete linkage with the respective phenotypes, but the major allele among earless animals was much less prevalent among unaffected individuals (p = 0.875 vs. p = 0.083).

Figure 2 is a simple diagram of the region of OAR23 immediately adjacent to OAR23_34647499.1, specifically the region between the two non-significant SNP flanking OAR23_34647499.1. This region includes all or part of three genes, 1) encoding a transfer RNA for arginine, 2) *GATA6*, GATA 6 binding hormone, and 3) *MIB1*, mindbomb E3 ubiquitin protein ligase 1. Among these three genes, *GATA6* is the most likely candidate as the site of the causative mutation, both for its location and its function.

In terms of location, *GATA6* is closer to OAR23_34647499.1 than the transfer RNA gene, whereas *MIB1*, although closer to OAR23_34647499.1 than *GATA6*, includes SNP OAR23_34788536.1, which has similar allelic distributions in affected and normal animals. In terms of function, *GATA6* is a highly plausible candidate as the location of the causative mutation. GATA6 is a transcription factor that has been shown to play a critical role in many developmental processes, including chondrogenesis (Alexandrovich *et al*. 2008). Chondrogenesis is the process by which cartilage is formed and cartilage is the tissue that

provides shape to the outer ear. Moreover, GATA-6 has been demonstrated (in mice) to be highly expressed within the pre-cartilaginous primordia in the first branchial arch of the embryo, which gives rise to the outer ear, among other facial structures (Brewer *et al.* 2002). Therefore, in addition to its close proximity to the SNP with a significant association with the earless phenotype, both the function of GATA-6 and its temporal and spatial expression are consistent with a role in outer-ear development.

The lack of extended homozygosity in this region suggests a fairly ancient mutation. As indicated in the Materials and Methods section, a modified version of the approach of Genin *et al.* (2004) was used to estimate mutation age. A modification was necessary, because the approach of Genin *et al.* (2004) assumes the presence of a run of homozygosity of at least one SNP on each side of the putative causative mutation. To make the necessary calculations, we assumed that the putative mutation occurred at the midpoint of *GATA6* and that recombination between this site and the two flanking assayed SNP (i.e. OAR23_34225387.1 and OAR23_34647499.1) had historically occurred at random. We then determined the locations on each side of the putative mutation between which the probability of no historical recombination was 0.50 and based the calculations on the assumed presence of complete homozygosity between these two sites. We further assumed a universal recombination rate within sheep of 1% per Mb.

According to these assumptions, the time of occurrence of the mutation was estimated to be approximately 860 generations ago. Considering a generation interval of 3.5 years (Willis 1998), this places the occurrence of the mutation to about 3000 years in the past. This result suggests that many of the sheep breeds showing this phenotype may share the causative mutation, but

perhaps not all earless breeds. Sharing of the mutation would be more probable within the sub-group of fat-tailed, wool sheep, as the Awassi is of this type and the divergences of fat- vs. thin-tailed (Moradi *et al.* 2012) and wool vs. hair sheep (Chessa *et al.* 2009) are believed to have occurred previous to this time.

Although *GATA6* is a strong candidate for the site of the causative mutation for the earless defect in Awassi sheep, additional analysis based on sequencing would be valuable. Sequencing would help confirm the results observed here, identify the precise point of the mutation and allow the proposal of a precise physiological and biochemical basis for the defect.

The present results provide a basis for determining if this putative causative mutation is shared by other sheep breeds exhibiting the earless condition and offer *GATA6* as a candidate gene to be studied in other species in which microtia is observed, including humans.

Acknowledgments

References

1. Alexandrovich A., Qureishi A., Coudert A.E., Zhang L. Grigoriadis A.E., Shah A.M., Brewer A.C. & Pizzey J.A. (2008) A Role for GATA-6 in vertebrate chondrogenesis. Developmental Biology 314, 457–470.

2. Brewer A., Nemer G., Gove C., Rawlins F., Nemer M., Patient R. & Pizzey J. (2002) Widespread expression of an extended peptide sequence of GATA-6 during murine embryogenesis and non-equivalence of RNA and protein expression domains. Mechanisms of Development 119 (Suppl 1), S121-S129.

3. Brown K.K., Viana L.M., Helwig C.C., Artunduaga M.A., Quintanilla-Dieck L., Jarrin P., Osorno G., McDonough B., DePalma S., Eavey R., Seidman J.G. & Seidman C. E. (2013). HOXA2 Haploinsufficiency in dominant bilateral microtia and hearing loss. Human Mutation 34, 10.

4. Charlier C., Coppieters, W., Rollin F., Desmecht D., Agerholm J. S., Cambisano N., Carta E., Dardano S., Dive M., Fasquelle C., Frennet J.-C., Hanset R., Hubin X., Jorgensen C., Karim L., Kent M., Harvey K., Pearce B.R., Simon P., Tama N., Nie H., Vandeputte S., Lien S., Longeri M., Fredholm M., Harvey R.J. & Georges, M. (2008) Highly effective SNP-based association mapping and management of recessive defects in livestock. Nature Genetics 40, 449–54.

5. Cox T.C., Camci E.D., Vora S., Luquetti D.V. & Turner E.E. (2014) The genetics of auricular development and malformation: new findings in model systems driving future directions for microtia research. European Journal of Medical Genetics 57, 394-401.

6. Genin E., Tullio-Pelet A., Begeot F., Lyonnet S. & Abel L. (2004) Estimating the age of rare disease mutations: the example of Triple-A syndrome. Journal of Medical Genetics 41, 445-449.

7. Harris J., Kallen B. & Robert E. (1996) The epidemiology of anotia and microtia. Journal of Medical Genetics. 33, 809–813.

8. Hernández I., Rodríguez J.V., Romero O., Hernández J.S., Macías A., López H. & Herrera J.G. Guadalupe. (2013) Morphometric characterization of Creole sheep without ear of the Sierra North State of Puebla-Mexico. International Research Journal of Biological Sciences 2, 1-8.

9. Luquetti D.V., Heike C.L., Hing A.V., Cunningham M.L. & Cox T.C. (2012) Microtia: epidemiology and genetics. American Journal of Medical Genetics 158A(1),124-139.

10. Lush J.L. (1930) Earlessness in Karakul sheep. The Journal of Heredity 21, 107-112.

11. Mason I.L. (1980) Prolific Tropical Sheep. FAO Animal Production and Health Paper 17. Food and Agriculture Organization of the United Nations. Rome.

12. Moradi M. H., Nejati-Javaremi A., Moradi-Shahrbabak M., Dodds K. G. & McEwan, J. C. (2012). Genomic scan of selective sweeps in thin and fat tail sheep breeds for identifying of candidate regions associated with fat deposition. BMC Genetics 13, 10.

13. Porter V. (2002) Mason's World Dictionary of Livestock Breeds, Types and Varieties. Fifth Edition. C.A.B International, Wallingford, United Kingdom.

14. Purcell S., Neale B., Todd-Brown K., Thomas L., Ferreira M.A.R., Bender D., Maller J., Sklar P., de Bakker P.I.W., Daly M.J. & Sham P.C. (2007) PLINK: a toolset for whole-genome association and population-based linkage analysis. American Journal of Human Genetics, 81, 559-575.

15. Ritzman E.G. (1920) Breeding earless sheep. Journal of Heredity 11, 238-240.

16. Talafha A.Q. & Ababneh M.M. (2011) Awassi sheep reproduction and milk production: review. Tropical Animal Health and Production 43, 1319-1326.

17. Willis M.B. 1998. Dalton's Introduction to Practical Animal Breeding, 4th edition. Blackwell Sciences Ltd. Oxford, U.K.

Table 1 Locations of the haplotypes of length 1 to 9 with the most significantly significant associations with the earless phenotype in the Awassi sheep.

Haplotype length (SNP)	Chromosome (OAR)	Base pair range	P-value
9	20	11,456,378 to 12,226,961	NS
8	3	191,077,623 to 191,697,552	NS
7	16	28,733,052 to 29,187,392	NS
6	16	28,733,052 to 29,012,867	NS
5	2	66,335,370 to 66,675,089	NS
4	6	50,284,784 to 50,419,833	NS
3	4	55,768,294 to 55,853,752	NS
2	1	93,124,341 to 93,219,257	NS
1	23	34,647,499	0.0124

Figure captions

Figure 1 Examples of normal (A), short-eared (B) and earless (C) Awassi sheep.

Figure 2 Map of single nucleotide polymorphisms (♦) genes (■) in the region of OAR23 where

a statistically significant association with the earless phenotype was observed.

Figure 1

A

B

C

Figure 2

SNP:
Location: 34,225,387
MAF affected: 0.44
MAF normal: 0.46
Genome wide *P*-value: 1.00

transfer RNA arginine (anticodon UCU)

GATA6 - GATA binding protein 6

SNP: OAR23_34647499.1
Location: 34,647,499
MAF affected: 0.88
MAF normal: 0.08
Genome wide *P*-value: 0.124

SNP: OAR23_34788536.1
Location: 34,788,536
MAF affected: 0.44
MAF normal: 0.54
Genome wide *P*-value: 1.00

MIB1 - mindbomb E3 ubiquitin protein ligase 1