Proceedings, 10th World Congress of Genetics Applied to Livestock Production

Investigation of Candidate Regions Associated With Fat Deposition in Thin and Fat Tail Sheep Breeds

*M. H. Moradi**, A. Nejati-Javaremi†, M. Moradi-Shahrbabak†, K. G. Dodds‡, R. Brauning‡ and J. C. McEwan‡
*Arak University, Arak, Iran, †University of Tehran, Karaj, Iran, ‡AgResearch, Mosgiel, New Zealand

ABSTRACT: The fat tail is a phenotype that divides domesticated sheep into two major groups. The objective of the present study was to refine the map location of candidate regions associated with fat deposition and to determine the nature of the selection occurring in these regions using XP-EHH approach. Zel (thin tail) and Lori-Bakhtiari (fat tail) breed samples were genotyped with a denser set of SNPs in the three candidate regions. Results enabled us to refine the critical intervals from 113kb, 201kb and 2,831kb to 30kb, 38kb and 1,077kb on chromosome 5, 7 and X respectively. Some genes associated with fat metabolisms like *PPP2CA* in Homo sapiens and *EBP* in Bos taurus have previously been reported in these regions on chromosome 5 and X respectively. Core haplotypes in these regions of interest support the hypothesis that the first domesticated sheep were thin tailed and fat tail animals were developed later.
Keywords: Candidate regions; Thin and Fat tail; XP-EHH

Introduction

Fat tailed breeds comprise approximately 25% of the world sheep population and are grazed in a wide range of countries including: northern parts of Africa, the Middle East and Asia(Moradi et al. (2012)).However, to date the loci associated with fat deposition in the tail of these breeds are unknown. Therefore finding the regions associated with fat deposition is one of the most important and challenging areas of research in the countries grazing these breeds.

Genomic scan for finding candidate regions associated with fat deposition in thin and fat tailed breeds have been described previously (Moradi et al. (2012)). In brief, two independent experiments including Iranian and ovine HapMap genotyping data contrasting thin and fat tailed breeds were analyzed and using different statistics especially F_{ST}, three regions on chromosomes 5 (between 47,149-47,263 kb), 7 (between 46,642-46,843 kb) and X (between 58,621-61,452 kb) were confirmed in both data sets.

Unlike F_{ST} that has been used in the previous study, tests based on LD likecross-population EHH(XP-EHH)are multi marker tests (Sabeti et al. (2007)). Densely spaced SNPs give greater power when using statistical tests that rely on linkage disequilibrium (LD), as signals of selection are less likely to be lost. The Illumina Ovine SNP50k BeadChip, while providing uniform genome wide coverage, only has a SNPabout every 56kb: too long to do LD based analysis. Fine mapping, where more SNPs are genotyped in an area of interest, improves the ability to localize causal variants.

Then the aim of this study was to refine the map location of candidate regions associated with fat deposition using XP-EHH approach and to determine the nature of the selection occurring in these regions.

Materials and Methods

Genotyping and quality control. The animal samples and genotyping with the Illumina OvineSNP50k beadchip have been described previously (Moradi et al. (2012)). Once the regions of interest were defined, all known SNPs in each region were examined for suitability for genotyping (http://www.sheephapmap.org/genseq.php) usingSequenom's primer design software (MassARRAY Assay Design 3.1) and probes were designed in multiplex format. The majority of the multiplex assays were located in the first and second plex which contained 71, 62 and 57 SNPs for chromosomes 5, 7 and X respectively. These two SNP assays were used for genotypingby use of the mass-spectrometry based iPlexMassArray platform provided by Sequenom (Sequenom, San Diego, California, USA. http://www.sequenom.com).

After genotyping, all samples with more than 30% missing data and subsequently all loci with more than 15% missing data were excluded. These rejection thresholds were chosen by plotting numbers of animals or loci against percent missing data (Barendse et al. (2009)). For the remaining SNPs those with MAF less than 2% and outlier departure from Hardy-Weinberg equilibrium over all animals of a breed (P-value$<10^{-6}$) were excluded.

Determining of ancestral alleles. Ancestral alleles for the Illumina OvineSNP50 BeadChip were obtained from Dr Clare Gill of Texas A&M University and were available for 30,923 SNPs.To obtain ancestral alleles for the additional Sequenom SNPs described here, 301 base pairs of sequence (1 bp of the alleles plus 150 bases either side of the SNP) were aligned against Bostaurus (cattle), Susscrofa (pig), Equuscaballus (horse), Canisfamiliaris (dog) and Homo sapiens (human) genomes using BLAST (http://blast.ncbi.nlm.nih.gov/Blast). The ancestral allele was taken as the base in the genome sequence at the resulting SNP position.

Reconstruction of haplotypes, Core SNP alleles and haplotype frequencies in candidate regions.A pair of haplotypes for each chromosome and also haplotypes for candidate regions of interest were reconstructed using fastPHASE version 1.2.3 and PHASE 2.1 respectively (Stephens et al. (2001)). Then the haplotypes for each region were fed into SWEEP v.1.1 (Sabeti et al. (2002)) to detect core regions based on the EHH statistic and to obtain haplotype frequencies in these regions.

Calculation of XP-EHH.A pair of haplotypes that reconstructed for each animal in the sample was used to calculate XP-EHH as in Sabeti et al. (2007). Scripts for computing this test were obtained from the Pritchard lab website (http://hgdp.uchicago.edu/Software/). For calculation of XP-EHH tests in this research, the physical position of each SNP was obtained from OAR true chromosomes (ver.1.0, as at 5/2008) from CSIRO.

Study of identified genes in candidate regions. The regions of interest in O. aries were compared to the corresponding areas in H. sapiens and B. Taurus, as their genome is better annotated, by BLAT search using the UCSC Genome Browser on Human Feb. 2009 (GRCh37/hg19) and Cow Oct. 2007 (Baylor 4.0/bosTau4) assembly. In order to discover this,Sheep Genome Browser OAR v2.0(http://www.livestockgenomics.csiro.au/cgi-bin/gbrowse/oarv2.0/) was applied. The Biological Process (BP), Molecular Functions (MF) and KEGG Pathway of the entire identified genes in candidate regions were determined using DAVID (http://david.abcc.ncifcrf.gov/).

Results and Discussion

Data mining. After data mining and considering SNPs with ancestral information for XP-EHH calculation, a total of 36, 29 and 38SNPs were applied for final analysis in our regions of interest on chromosome 5, 7 and X respectively. The average SNP intervals were relatively consistent and the overall average distance between adjacent SNPs was about 24kb, 17kb and 80kb for candidate regions on chromosome 5, 7 and X respectively, compared to about 60kb and 115kb for autosomes and chromosome X on the Illumina Ovine SNP50k BeadChip respectively.

Cross-population EHH (XP-EHH). To identify selective sweeps in which the selected allele has approached or achieved fixation in a subpopulation, but remains polymorphic in the population as a whole, the standardized cross population extent of haplotype homozygosity scores (XP-EHH) were plotted against genomic locations (Figure 1). The results revealed clear peaks in all of our regions of interest, suggesting that selected alleles approached fixation or have risen to near fixation in favour of fat tail breed on chromosome 5 and X while on chromosome 7 the frequency of alleles have risen close to fixation in thin tail breed. These results are based on linkage disequilibrium around given SNP and are in agreement with the results of F_{ST} and median homozygosity plots (Moradi et al. (2012).

Sabeti et al. (2007) considered a region as a candidate for selection in the human HapMap Phase 2 dataset when the 2 population XP-EHH was above 4.34. This score is in the 99.9 percentile for thin versus fat tail breeds. As shown in figure 1, we found evidence of selection with |XP-EHH| value > 4.34 on chromosomes 5 (47,141,229-47,171,110 bp), 7 (46,604,500-46,642,359 bp) and X (59,187,456-60,264,325 bp).

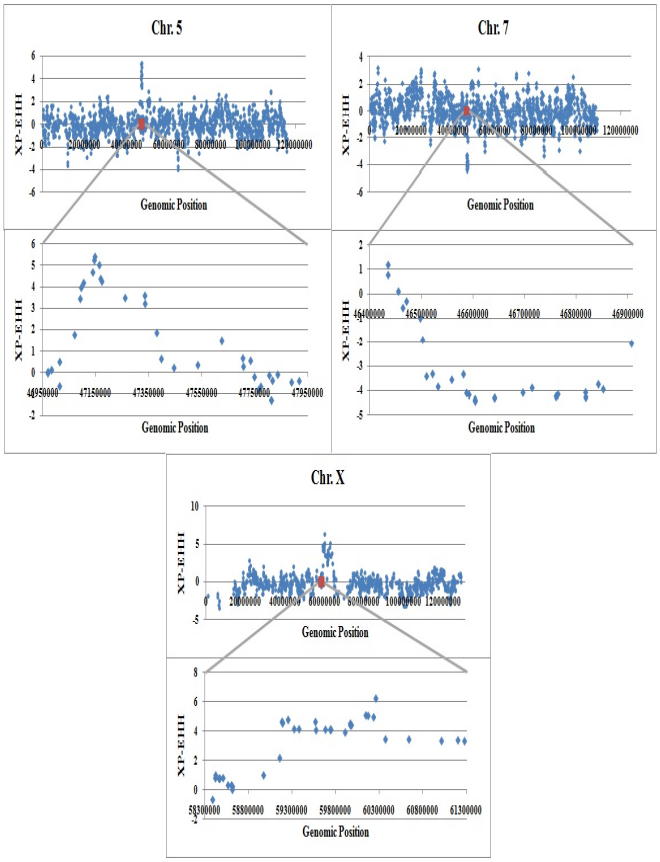

Figure 1- Plot of XP-EHH relation to genomic position (bp) for thin and fat tailed breeds on whole chromosome (upper) and candidate region of interest (lower) in different chromosomes; High positive values suggest selection in fat tailed population and negative values selection in thin tailed.

Study of core haplotypes and their ancestral status in candidate regions. To evaluate the haplotype frequencies and the status of selected alleles (derived or ancestral), the core SNP alleles and their haplotype frequencies were investigated in the selected regions.

For chromosome 5 (Table 1), we defined a core region of 26k where XP-EHH values were at their highest. There are 5 genotyped SNPs in this region. The SNPs defined 14 core haplotypes (denoted haplotype 1 to 14) in thin tailed but only 6 core haplotypes in fat tailed sheep. As

shown in table 1 there is a common haplotype (haplotype 1) with frequency of 90% in fat tailed sheep whereas its frequency for thin tailed is 15%. In contrast, the common haplotype in thin tail breed is haplotype 8 with a frequency of 31% while it is almost absent (2%) in the fat tail breed. The interesting result in this region is that all the SNPs in the common haplotype for the fat tail breed are derived SNPs whereas all SNPs in the common haplotype for the thin tail breed are ancestral.

Table 1: Core SNP alleles and haplotype frequencies in the candidate region on chromosome 5 for fat tail (Lori) and thin tail (Zel) breeds

	Core SNP alleles					Core haplotype frequencies	
SNPs	C/T	A/G	T/C	A/G	A/C	Fat tail breed	Thin tail breed
Ancestral allele	T	G	T	G	C		
Haplotype 1	**C**	**A**	**C**	**A**	**A**	0.90	0.15
Haplotype 2	-	G	-	-	-	0.01	0.06
Haplotype 3	-	G	T	-	-	0.01	-
Haplotype 4	-	G	-	G	C	-	0.12
Haplotype 5	T	G	-	-	-	-	0.13
Haplotype 6	T	-	T	-	-	0.01	-
Haplotype 7	T	G	T	-	-	0.04	0.11
Haplotype 8	**T**	**G**	**T**	**G**	**C**	0.02	0.31
Other Haplotypes						-	0.14

The results of chromosome 7 and X were in consistent with chromosome 5 (data not shown) and revealed that in almost all cases, derived alleles have been under selection pressure in the fat tail breed.This is consistent with selection for a new mutation in these breeds. Given that fat tailed breeds are now prevalent in the Fertile Crescent, where sheep were originally domesticated, while thin tailed sheep breeds are predominant in peripheral areas and also considering this fact that the wild ancestor of sheep is thin tail, it has been assumed that the first domesticated sheep were thin tailed and fat tail was developed later (Davidson (2006)). Our results are in agreement with this hypothesis.

Study of the identified genes associated with fat metabolisms in candidate regions. The genes obtained by orthology with H. Sapiens and B. Taurus, and their function studied in this research. The results showed that some genes associated with fat metabolisms like *PPP2CA* in H. sapiens and *EBP* in B. taurus have been reported previously in these rejoins on chromosome 5 and X respectively. *PPP2CA* have a variety of roles in different biological process such as cellular lipid metabolic, membrane lipid metabolic and sphingolipid metabolic process while EBPget participate inlipid metabolic process. It seems as the annotation of the ovine genome becomes more complete, all genes located in the candidate regions will be identified. Promising targets can then be verified by further experimentation.

Conclusion

Our results provide an assessment of how the selection has affected the patterns of variation in candidate regions associated with fat deposition in thin and fat tail sheep breeds. These results do offer hope that the causal mutations and the mode of inheritance of this trait will soon be discovered by further experimentation.

Literature Cited

Barendse, W., Harrison, B.E., Bunch, R.J.et al. (2009). BMC Genomics. 10: 178

Davidson.A. (2006). 2nd edition. Edited by Tom Jaine. Oxford University Press, UK

Moradi, M. H., Nejati-Javaremi, A., Moradi-Shahrbabak, M.et al. (2012). BMC Genetics. 13:10

Sabeti, P. C., Varilly, P., Fry, B.et al. (2007). Nature. 449(7164):913-U912

Sabeti, P. C., Reich, D. E., Higgins, J. M. et al. (2002). Nature. 419: 832-837

Stephens, M., Smith, N.J., and Donnelly, P. (2001). Am. J. Hum.Genet. 68(4): 978-89