GENOME-WIDE ASSOCIATION STUDY FOR SOME MORPHOLOGICAL TRAITS IN IRANIAN BUFFALO

Mahdi Mokhber* *
a: Department of animal science, Faculty of Agriculture, Urmia University, Urmia, Iran
*Corresponding author: m.mokhber@urmia.ac.ir (Mahdi Mokhber)

ABSTRACT
A genome-wide association study (GWAS) was carried out on Iranian water buffaloes with the objective of detecting genomic regions associated with some morphological traits. For this proposal, a set of 376 water buffalo with phenotypic data from Azeri (N=246), Khuzestani (N=109) and Mazandarani (N=20) breeds were genotyped with Axiom® Buffalo Genotyping Array which contains 89,988 single nucleotide polymorphism (SNP) markers. Detecting genomic regions that associated with type traits was done using GenABEL package in R. In total, 8, 9, 9, 5, 7, 5, 12, 8 and 7 genomic regions were detected that maybe associated with body length (BL), heart girth (HG - represented the circumference of the chest just behind the forelimbs), chest depth (CD), height-at-withers (HAW), and hip width (HIW), pin width (PW), hip to pin length (HP) and head size (HS) and horn length (HL) morphological traits, respectively. Annotation of these regions using the UMD3.1 Bostaurus Genome Assembly was performed to find putative candidate genes within the selected regions. Among the candidate genes identified in these regions GRB10, EGR1, ACAT2, SLC25A25, NRP1, FOXC2 and ABR genes are related to biological cell growth pathways and BMP2K, EGR1, CCIN, FRMPD1, ARPC5L, DNM1, FOXC2, ABR and MPP1 are related skeletal bone development biological pathways such as bone mineralization. Further investigation of these regions with additional data is necessary to confirm the associations and fine-map the QTL regions.

Keywords: Genome wide association study, single nucleotide polymorphisms, Azeri, Khuzestani and Mazandarani buffalo breeds.

INTRODUCTION
The water buffalo (Bubablusbubalis) reared in many regions of the world, in tropical and subtropical countries and have very important in the rural economies in these areas in Asia (Misra and Tyagi, 2010). Water buffalo provide about 97 million tons (12.9%) and 3.6 million tons (4.5%) of the world’s total milk and red meat production (FAO, 2014). Therefore, utilize new technologies and genomics area research is require to obtaining genetic progress in production traits. In recent years, research in genomics area has been increased with development of high-density SNP arrays in many livestock species (Cole et al., 2009). Recently, because of availability of a commercial buffalo SNP chip array for water buffalo, research in genomic such as GWAS and genomic selection, is possible for water buffalo (Iamartino et al., 2013; Peters et al., 2012). To date, several studies were used GWAS for some economically important traits in domestic animals including cattle (Hayes et al., 2009; Pryce et al., 2010; Sahana et al., 2010; Lu et al., 2013), poultry (Yuan et al., 2015; Gu et al., 2011), pig (Bergfelder-Drüing et al., 2015; Le et al., 2017) and water buffalo by Bovine 50K and 777K SNP Beadchips (Wu et al., 2013; Borquis et al., 2014) and specific buffalo 90K SNP array(Iamartino et al., 2013; Camargo et al., 2015; El-Halawany et al., 2017).
The objective of the current study was to perform a GWAS for some morphological traits in Iranian buffalo using with Axiom® Buffalo Genotyping 90K Array, to detect genomic regions associated with these traits.

MATERIALS & METHODS

A set of 376 water buffalo with phenotypic morphological data from Azeri (N=246), Khuzestani (N=109) and Mazandarani (N=20) breeds were genotyped with Axiom® Buffalo Genotyping Array which contains 89,988 single nucleotide polymorphism (SNP) markers. Phenotypic morphological data including body length (BL), heart girth (HG - represented the circumference of the chest just behind the forelimbs), chest depth (CD), height-at-withers (HAW), and hip width (HIW), pin width (PW), hip to pin length (HP) and head size (HS) and horn length (HL) that measured by unique valuator. Genotype quality control was accomplished using Plink software (Purcell et al., 2007). Genotypes were checked for Individual call rate, SNP call rate and minor allele frequency (MAF). Then, Individual and SNPs with more than 5% missing data and MAF less than 10% were excluded. After quality control (QC) all Individual with a total of 65,185SNPs remained for upcoming analysis. Former to carrying out the GWAS, Multi Dimension Scaling (MDS) was implemented to survey genetic homogeneity and population sub-structure of the dataset using GenABEL package in R (The R project website, http://www.r-project.org/). Also, the effective factors on studied morphological traits to put into main model were surveyed in SAS software, version 9.4 (SAS 2014, SAS Institute Inc., Cary, NC, USA). Detecting genomic regions that associated with morphological traits was done by GenABEL package in R (The R project website, http://www.r-project.org/). Based on SAS and MDS analysis results, final model determined by driving significant fixed effect on each studied traits and applying adjustment for the stratification by use of the principal components of genetic variation. The final model was:

$$MorphRec_{ijkl} = \mu + Farm_i + Age_j + sex_l + PC1 + PC2 + PC3 + Polygenic_k + e_{ijkl}$$

where $MorphRec_{ijkl}$ is phenotypic value of each type traits; $\mu$ is the overall population mean for every trait; $Farm_i$ is the fixed effect of the $i$th farm or herd; $Age_j$ is the fixed effect of the $j$th Age in recording; $sex_l$ is the fixed effect of the $l$th sex; $PC1$ to $3$ are principal components; $Polygenic_k$ is the effect of the $k$th genotype; and $e_{ijkl}$ is the residual random effect.

Associated genomic regions were used to extract genes from the corresponding areas in UMD3.1 Bos Taurus Genome Assembly using Biomart (www.ensembl.org/biomart/martview). Finally, DAVID (Huang et al., 2009) Enrichment Map Cytoscape plug-in (Merico et al., 2010) were used to perform a gene ontology analysis and construct networks, respectively.

RESULTS AND DISCUSSION

The MDS analysis results was revealed that three Azeri, Khuzestani and Mazandarani buffalo breeds are genetically distinct and moreover some stratification exist within each breed, then stratification have been arose in studied population. Therefore, structured association
analysis was preferred for association model with identified fixed effects that have significant effect on studied traits. In total, 8 (on chromosomes BTA2, 6, 7, 8, 11, 15, 20 and 23), 9 (on chromosomes BTA4[2 region], 6, 12, 16, 17, 18 and 20[2 region]), 9 (on chromosomes BTA1[3 region], 4[2 region], 13, 27[2 region] and 28), 5 (on chromosomes BTA3, 8, 9 13 and 27), 7 (on chromosomes BTA3[2 region], 4, 11[2 region], 17 and 28), 5 (on chromosomes BTA3, 6, 8[2 region] and 19), 12 (on chromosomes BTA1, 8[3 region], 12, 22, 28[3 region], X[2 region]), 8 (on chromosomes BTA3, 4.8 and 9[4 region]) and 7 (on chromosomes BTA3[3 region], 4, 19, 20,21) genomic regions were detected with suggestive signals of association with body length (BL), heart girth (HG - represented the circumference of the chest just behind the forelimbs), chest depth (CD), height-at-withers (HAW), and hip width (HIW), pin width (PW), hip to pin length (HP) and head size (HS) and horn length (HL) morphological traits, respectively. The 500 Kbp genomic regions flanking the significant SNPs regions were investigated for genes that may be related to studied traits. Annotation of these regions using the UMD3.1 Bostaurus Genome Assembly was performed to find putative candidate genes within the selected regions. A total of 101 genes were perceived within these regions. Among the candidate genes identified in these regions GRB10, EGR1, ACAT2, SLC25A25, NRP1, FOXC2 and ABR genes are related to biological cell growth pathways and BMP2K, EGR1, CCIN, FRMPD1, ARPC5L, DNMI, FOXC2, ABR and MPP1 are related skeletal bone development biological pathways such as bone mineralization, that maybe related to studied traits. Finally, gene ontology analysis was performed using DAVID to detecting biological network among provided genes and the Enrichment Map Cytoscape plug-in to construct networks. However, there no significant biological pathway was found in gene ontology analysis.

CONCLUSIONS (FACULTATIVE)

Present study is the opening GWAS in Iranian water buffalo. Despite of small sample size, some of detected SNPs were associated with biological cell growth and skeletal bone development pathways. However, further investigation on these regions with additional data is necessary to perform more association and functional studies to confirm the associations and fine-mapping regions and reveal the implication of these genes in marker-assisted selection (MAS).

REFERENCES CITED

EnsemblBioMart: Ensembl online genome data base BioMart Tool.
The R project website, http://www.r-project.org/