A Genome-Wide Association Study of Survival to Unexpected Acute Heat Stress in a F2 Chicken Population

H. Asadollahi, R. Vaez Torshizi, A. Ehsani, and A. A. Masoudi

ABSTRACT

Heat stress, or hyperthermia, can have a serious effect on chicken performance in poultry industry in many parts of the world. Both genetics and environment play key roles in the performance of a chicken and, therefore, it is important to consider both factors in addressing heat stress. On a genetic level, genome-wide association studies have become a popular method for studying heat stress in recent years. A population of 202 F2 chickens was reared for 84 days to find genes and genomic regions affecting growth traits and immune system. But, due to unexpected acute increase in temperature at day 83, 182 birds died (case) and 20 birds remained alive (control). At the age of 70 days, blood sample of all birds was collected to extract their DNA, using modified salting out method. All samples were genotyped by a 60 K Single Nucleotide Polymorphism (SNP) chip. Genome-wide association study was carried out by GCTA to identify genes and genomic regions associated with heat stress tolerance. Results indicated a close relationship between 28 SNPs, located on chromosomes 2, 3, 5, 6, 7, 12, 19, 20, and 21 and heat stress tolerance at the level of suggestive significance. Two suggestively significant markers on chromosome 5, namely, GGaluGA273356 and Gga_rs16479429, were located within and 52 Kb downstream of two genes, including MAPKBP1 and SPON1, respectively. Gene ontology analysis indicated that the resistance of chickens to acute increase of temperature might be linked to the function of MAPKBP1 and SPON1 genes and their biological pathways. These results will be useful for understanding the molecular mechanisms of SNPs and candidate genes for heat stress tolerance in chickens and provide a basis for increasing genetic resistance in breeding programs.

Keywords: Gene ontology analysis, GWAS, Hyperthermia, Single nucleotide polymorphism.

INTRODUCTION

In many parts of the world, heat stress, or hyperthermia, can have a serious effect on chicken performance in poultry industry (Lin et al., 2006). Heat stress is a condition in chickens that occurs at high temperatures and will be acute when combined with high relative humidity and low air speed (Ahmad and Sarwar, 2006). Two predisposing factors that affect heat stress include genetics and environment. Literature results indicate that continuous increase in growth rate and feed efficiency in modern broiler breeds coincide with a reduced heat stress tolerance (Lara and Rostango, 2013). In an experiment, muscle damage in fast growing broilers was reported to be associated with an ambient temperature (Zahoor et al., 2016). Additionally, it has been indicated that much of the variation in response to heat stress has apparently a genetic base (Lu et al., 2007). Soleimani et al. (2011) found that commercial broilers were more susceptible to heat stress than their Red Jungle fowl counterparts. Also, Mack et al. (2013) provided evidence that genetic selection is a useful strategy for reducing heat stress response in laying hens.

Evidences demonstrated that genetic progress of heat stress tolerance can be
accelerated through application of genomic information in breeding populations. In recent years, several methods such as genome-wide association studies (Van Goor et al., 2015) and transcriptome comparison (Coble et al., 2014) have been suggested to implement for studying heat stress tolerance. The Genome Wide Association Study (GWAS) has been proposed to be a suitable method for identifying genomic regions affecting the resistance of birds to heat shock (Lamont et al., 2014).

Several Quantitative Trait Loci (QTLs) have been identified to be associated with heat stress tolerance in Holstein cattle (Dikmen et al., 2013) and catfish (Jin et al., 2017). For chickens, few results have been found in literatures and several candidate genes were reported to be associated with heat stress tolerance (Van Goor et al., 2015). Therefore, the objective of the present study was to identify candidate genes related to survival to an unexpected acute increase in temperature, namely, heat stress tolerance, using GWAS in an F2 chicken population.

**MATERIALS AND METHODS**

**Experimental Design and Management**

All procedures were conducted under the protocols approved by the Laboratory Animal Care Advisory Committee of Tarbiat Modares University. From a reciprocal crosses between fast-growing Arian broiler line (A) and an Iranian Native fowl line (N), F1 chickens were generated by mating of A♂×N♀ and N♂×A♀ birds. The F1 male birds from each reciprocal cross were each mated to four to eight females from non-related families, resulting in a total of 202 F2 chickens of eight half-sib families from three consecutive hatches. Day-old F2 chickens were individually identified with wing tags, and immediately weighed after hatch. The birds were reared on the floor for 7 days under 24 hours light and a brooding temperature of 33°C. This temperature was decreased to 30°C on day 7. On day 8, birds were weighed and moved to individual cages with a temperature of 30°C, which was gradually decreased to reach a final temperature of 22°C, and a 22 hours light and 2 hours dark cycle throughout the experimental period. Chickens did not receive vaccines during rearing period. Feed and water were provided ad libitum. This experiment was conducted to measure chickens body weight and carcass compositions at 84 days of age, but due to unexpected acute increase in the room temperature at the eighty-third night, 182 birds died (case) and 20 birds remained alive (control).

**Genotyping and Quality Control**

At the age of 70 days, DNA from blood samples of all 202 chickens was extracted using the Salting out method and stored at -20°C. These DNA were used to identify Single Nucleotide Polymorphism (SNP) of each bird, using the Illumina chicken 60 K Beadchip containing 54,340 SNP markers provided by Cobb Vantress with cooperation of Arhus University of Denmark. The PLINK (v1.07) software package was used to determine the quality control of the markers (Purcell et al., 2007). After deleting SNPs with a minor allele frequency less than 5%, a Hardy-Weinberg equilibrium test P-value less than 1×10⁻⁶, and a call rate less than 95%, a total of 47,730 SNP markers and 201 individuals remained for final analysis. A multidimensional scaling (MDS) method, which explains observed genetic distance among individuals (Wang et al., 2012), was implemented in PLINK software to evaluate the population structure. In this software, MDS-plot option in conjunction with cluster was used to perform MDS analysis on N×N matrix of genome-wide identity by state pair wise distance. The results of MDS-plot option was used to extract the number of MDS dimensions and it was plotted by R software.
Statistical Analysis

The GWAS analysis for heat stress tolerance was performed by the Mixed Linear Model (MLM), using GCTA software program (Yang et al., 2011) with the 47,730 SNPs passing quality control. The MLM in this software fits the effects of all the SNPs as random (Goddard et al., 2009). The resultant model was as follows:

$$y = Xb + Wq + Sa + e$$  
(1)

Where, $y$ is a vector of discrete binary observations as 0 for the cases (died) and 1 for controls (alive), $b$ is a vector of fixed effects of overall mean, sex (male and female) and hatch of birth (3 hatches), $q$ is a vector of fixed effects of the first dimensions from MDS analysis, $a$ is a vector of random effects of SNP markers with $a \sim N(0, I\sigma_a^2)$, $I$ is identity matrix, and $e$ is a vector of random residual effects with $e \sim N(0, I\sigma_e^2)$. The $X$, $W$, and $S$ are design matrices relating $y$ to $b$, $q$, and $a$, respectively.

For each SNP, an adjusted $P$-value was computed using the False Discovery Rate (FDR) procedure as

$$P_{FDR} = \left(\frac{i}{m}\right) \times Q$$  
(2)

Where, $i$ is the rank of $P$-value, $m$ is the number of total SNPs and $Q$ is the false discovery rate (Benjamini and Hochberg, 1995). Based on this procedure, an adjusted threshold value of $P_{FDR} = 4.99 \times 10^{-5}$ ($\frac{238}{47730} \times 0.01$), with $-\log(P_{FDR}) = 4.30$,  
(3)

and

$$P_{FDR} = 1.83 \times 10^{-3}$$  
(4)

was applied for genome-wide and suggestive levels of significance, respectively, to declare significant association between heat stress tolerance and marker.

Manhattan plots of genome-wide association analyses were produced using R 3.2.2 software with the QQMAN package (Turner, 2014). The positions of significant or suggestive SNPs, were identified by NCBI in chicken reference genome database (Gallus-gallus-4). Names of the genes with or adjacent to the associated SNPs (in 1-Mb segments surrounding each significant SNP) were detected from NCBI (http://www.ncbi.nlm.nih.gov) and Ensembl (ftp://ftpensembl.org/pub/release-73/fasta/gallus_gallus/dna/). The Database for Annotation, Visualization and Integrated Discovery (DAVID: https://david.ncifcrf.gov/) was used to analyze gene ontology (Huang et al., 2009). This database provides a comprehensive set of functional annotation tools for understanding the biological meaning behind large list of genes.

RESULTS

The population structure of $F_2$ chickens was identified through MDS analysis. The first and second components of MDS were used to create two-dimensional plot for individuals. The result is shown in Figure 1. Six different subpopulation groups with different dispersions were identified, indicating the chickens within each half-sib family were clustered together. These results suggest that MDS analysis largely support the consensus clustering outcome in this study.

The Manhattan plot of the $P$ values calculated from the genome-wide association study for heat stress tolerance for all SNPs is presented in Figure 2. No significant association was found between heat stress tolerance and SNP markers at the genome-wide significance level ($P > 4.99 \times 10^{-5}$). A close relationship was found between 28 SNPs, located on chromosomes 2, 3, 5, 6, 7, 12, 19, 20, and 21 and heat stress tolerance at the level of suggestive significance ($P < 1.83 \times 10^{-3}$).
Figure 1. Population structure identified by multidimensional scaling analysis. The scale on the X-axis is the first MDS compound, and the scale on the Y axis is the second MDS compound. Each color identifies a half-sib family.

Figure 2. Manhattan chart heat resistance Thermometer for chicken chromosomes. The X axis, the marker position on the chromosome, and the Y axis represent the negative logarithm of the $P$-value 10 base.

The 10 more significant SNPs with $P$ values $\leq 9.5 \times 10^{-4}$, the name of chromosome and the gene surrounding each SNP, the position of each marker on the chromosome, and the closest gene to the desired marker are presented in Table 1.

Among the identified markers, the SNP position of Gga_rs16039304 was not detected on chromosome 2. On these chromosomes, four SNPs (GGaluGA155695, GGaluGA155135, GGaluGA155136 and GGaluGA149469) were located within a 22.77 Mb region. These SNPs were clustered in a region between 60.98 and 83.75 Mb and located 39 Kb downstream, 4 Kb upstream of $FHOD3$, $IKZF1$ genes, and within $IKZF1$ and $LOC107051657$ genes, respectively. The GGaluGA238855 SNP located at 106.87 Mb on chromosome 3 was 8 Kb upstream of
Table 1. Location of and gene information for 10 SNPs association with heat stress tolerance.

<table>
<thead>
<tr>
<th>SNP</th>
<th>Chr.</th>
<th>Position (bp)</th>
<th>P-value</th>
<th>Gene</th>
<th>Distance (Kb)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gga_rs16039304</td>
<td>2</td>
<td>-</td>
<td>0.000304981</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>GGaluGA155695</td>
<td>2</td>
<td>83749423</td>
<td>0.000332619</td>
<td>FHOD3</td>
<td>39 kb</td>
</tr>
<tr>
<td>GGaluGA155135</td>
<td>2</td>
<td>80910368</td>
<td>0.000360726</td>
<td>IKZF1</td>
<td>4 kb</td>
</tr>
<tr>
<td>GGaluGA155136</td>
<td>2</td>
<td>80918742</td>
<td>0.000634819</td>
<td>IKZF1</td>
<td>within</td>
</tr>
<tr>
<td>GGaluGA149469</td>
<td>2</td>
<td>60975377</td>
<td>0.000945898</td>
<td>LOC107051657</td>
<td>within</td>
</tr>
<tr>
<td>GGaluGA238855</td>
<td>3</td>
<td>106874244</td>
<td>0.000849221</td>
<td>MSRA</td>
<td>8 kb</td>
</tr>
<tr>
<td>GGaluGA273356</td>
<td>5</td>
<td>6474615</td>
<td>0.000593381</td>
<td>SPON1</td>
<td>52 kb</td>
</tr>
<tr>
<td>Gga_rs16479429</td>
<td>5</td>
<td>25109295</td>
<td>0.000878839</td>
<td>MAPKB1</td>
<td>within</td>
</tr>
<tr>
<td>Gga_rs14623035</td>
<td>7</td>
<td>28801482</td>
<td>0.000615910</td>
<td>EN1</td>
<td>127 kb</td>
</tr>
<tr>
<td>Gga_rs14119331</td>
<td>19</td>
<td>3792047</td>
<td>0.000472838</td>
<td>MYL10</td>
<td>35 kb</td>
</tr>
</tbody>
</table>

Chr. is chromosome. U and D: Represent Upstream and Downstream of gene, respectively.

MSRA gene. Two SNPS, namely, GGaluGA273356 and Gga_rs16479429, located within a 18.63 Mb region between 6.47 and 25.11 Mb on chromosome 5 were 52 Kb downstream and within SPON1 and MAPKB1 genes, respectively. Two SNPs, one on chromosome 7 (Gga_rs14623035) and the other on chromosome 19 (Gga_rs14119331), were located at 28.80 and 3.79 Mb and were found to be 127 and 35 Kb downstream of EN1 and MYL10 genes, respectively.

**DISCUSSION**

In the case of heat stress, the physiological responses of different organisms to the increase in ambient temperature is heat gain of body and increment in the blood flow of the skin (Sandercock *et al.*, 2001). These can be regarded as indicators of the reaction of body to regulate its production through elimination of excess body heat (Hocking *et al.*, 1994). Genes regulation related to activation of Heat Shock Factors (HSFs) and their transcription activators have been identified as cellular responses to different types of stresses (Sorger, 1991). Heat shock proteins have been suggested to be the key cellular defense mechanisms in chickens during exposure to hot environments (Cedraz *et al.*, 2017). Also, the c-Jun amino-terminal Kinases (JNK) and P38 Mitogen-Activated Protein Kinase (MAPK), which are primarily activated in heat shocks in cells, regulate cellular activities such as gene expression, mitosis, differentiation, and cell survival (Tibbles and Woodgett, 1999; Bogoyevitch *et al.*, 2010). These two kinase proteins, which are of the subgroups of the MAPK family (Koul *et al.*, 2013), are involved in various complex biological processes and strongly respond to stress signals, including osmotic stress (Vassalli *et al.*, 2012). Hao *et al.* (2018) suggested that the MAPK signaling was involved in the protective role of HSP70 against heat stress-induced injury in the small intestine of rats.

In the present study, several SNPs were identified to be suggestively associated with heat stress tolerance using information from a population of F2 chickens. Gene ontology indicated that these SNPs were within or nearby FHOD3, IKZF1, LOC107051657, MSRA, SPON1, MAPKB1, EN1 and MYL10 genes. Two SNPs, Gga_rs16479426 and GGaluGA273356, located on chromosome 5, were found to be within and close to the MAPKB1 and SPON1 genes, respectively. The MAPKB1 gene plays an important role in initiating and activating the JNK enzymes and, therefore, protein kinases of MAPK in heat stress environment (Fan *et al.*, 2014). The prevention of oxidative stress by the c-Jun N-terminal Kinase due to heat shocks has been reported by Courtial *et al.* (2017) in corals and human cells. Also, a key transcription factor in the Nuclear Factor kappa B (NF-kB) signaling pathway, which regulates cell cycle survival, apoptosis, and cell transformation, has been
suggested for MAPKBP1 gene in literature (Fu et al., 2015). Several studies indicated that NF-κB signaling is essential for resistance to heat stress-induced early stage apoptosis in human (Janus et al., 2011; Liu et al., 2015), chickens (Akbarian et al., 2016), pigs (Ganesan et al., 2017) and cattle (Kumar and Singh, 2019).

Furthermore, several studies have reported more than one member of the MAPK family in phosphorylation and HSF-1 activity (Chu et al., 1996), and in some studies, it has been confirmed that part of the HSF-1 activity in the body is regulated by MAPK (He et al., 1998). HSF-1 is a monomer inactive in a complex with heat shock proteins, including HSP40, HSP70, and HSP90. With heat increment, the HSF-1 is released from the protective complex and, after being transferred to the nucleus, it is phosphorylated and attached to the DNA containing the heat shock elements. The most important HSF-1 target genes are also HSP genes (Shamovsky and Nudler, 2008). HSF-1 proteins increase the level of HSP-encoded proteins (Kim et al., 2005), and are also a molecule that is used to reduce or denature defective proteins, as well as protect the cell against heat stress (Zhang et al., 2002). It should be noted that among the heat shock proteins, the HSP70 protein plays a central and key role in cellular repair and protection against heat or other stresses (Lindquist, 1986). It seems that the MAPKBP1 gene plays an important role in increasing the heat resistance in F2 populations through regulating the activity of JNK protein kinases, consequently activating HSF-1 and increasing the level of heat shock proteins.

In addition to the MAPKBP1 gene, SPON1 gene has been also introduced as an important regulator of response to heat stress during heat increment, and reduces the heat shock stress. In current study, the GGaluGA273356 SNP, located 52 Kb downstream of the SPOD1 gene, was found to be suggestively associated with heat stress tolerance. SPOD1 gene encodes Spondin 1 (F-spondin), which is an extracellular matrix protein. It was shown that over-expression of Spondin 1 protein suppresses endogenous levels of amyloid β and reduces Aβ plaque deposition in mice (Hafez et al., 2012). The cytotoxic formation of amyloid β fibrils by heat shock proteins on lipid membranes in response to heat stress has been reported by Sakono et al. (2013). Coble et al. (2014) reported a down-regulation for SPOD1 in response to the heat stress treatment.

For other genes, no functions related to heat stress tolerance have been reported in chickens. Hamzić et al. (2015) reported that the FHOD gene was significantly associated with percentage of β2-globulin in blood plasma. The IKZF1 gene was suggested to be a central regulator of hematopoiesis that contributes to nearly every level of B cell differentiation and function (Sellars et al., 2011). For LOC107051657 gene, no function has been reported in the literatures. Xue et al. (2017) have reported MYL10 gene to be associated with the growth and development of muscle in chickens. For EN1 and MSRA genes, the development of cell organ limbs transcription regulation and the binding of specific DNA sequences (Simon et al., 2004) and protein regeneration and response to oxidative stress (Stadtman et al., 2002) have been reported.

In general, the results of the present study suggest that the resistance of some of our F2 chickens population to acute increase of temperature may be linked to the function of MAPKBP1 and SPON1 genes and their biological pathways. However, this conclusion is based on suggestive association of the SNPs and heat stress tolerance, and needs more realistic models to identify potential genes in LD with significantly associated SNPs. If the findings are confirmed, the results will be useful for understanding the molecular roles and mechanisms of SNPs and candidate genes to heat stress tolerance in chickens and provide a basis for increasing genetic resistance in breeding programs.

ACKNOWLEDGEMENTS
This work was funded by Tarbiat Modares University, Tehran, Iran. Genotyping of the birds was supported by Aarhus University, Denmark. The authors would like to thank Dr. Just Jensen for financial support of the bird's genotyping. Access to the developed 60 K SNP Illumina chicken array was kindly provided by the USDA Chicken GWMAS Consortium, Cobb Vantress, and Hendrix Genetics.

REFERENCES


17. He, B., Meng, Y. H. and Mivechi, N. F. 1998. Glycogen Synthase Kinase 3β and


آوري شده بود، آنها به روش بهينه شده نمکي استخراج شد. هم نمونه های استخراج شده توسط تراشه تجاري 60 هزار راپيي چندشکلي تک نوکلتويدي تعدين زنوبيت شيدند. پيونش زنومي اين دادهها برای GCTA شناسايي زنها و مناطق زنومي مرتبط با مقاومت به تنش گرمابي با استفاده از نرم افزار انجام شد. نتایج نشان داد كه ارتباط معناداري بین 28 چندشکلي تک نوکلتويدي با صفت مقاومت به تنش گرمابي در سطح بيشهادي وجود دارد. دو شناسگر معناداري روي كروموزوم 5، بعنی 56 و GGaluGA273356، Gga_rs16479429 با فاصله 25 كيلو باز پايين دست ZNQ در مهارت 1, MAPKBP1 از 22 كيلو باز پايين دست ZNQ درن ZNQQ MAPKBP1 سالما. بررسی هستشي شناسي ZNQ نشان داد كه مقاومت به تنش گرمابي در جوجهها با ZNQ و ميرهاي بيوپزديکي آنها در ارتباط است. اين نتایج برای درك مکانيسم مولکولي چندشکلي هاي تک نوکلتويدي و زنهاي نامزد برای مقاومت به تنش گرمابي مفيد بود و مي- توانند در برنامه هاي اصلاحي برای بهبود مقاومت زنهاي استفاده شود.