Assessment of genetic diversity and phylogenetic relationships of Korean native chicken breeds using microsatellite markers

Joo Hee Seo1,2, Jun Heon Lee3, and Hong Sik Kong1,*

Objective: This study was conducted to investigate the basic information on genetic structure and characteristics of Korean Native chickens (NC) and foreign breeds through the analysis of the pure chicken populations and commercial chicken lines of the Hanhyup Company which are popular in the NC market, using the 20 microsatellite markers.

Methods: In this study, the genetic diversity and phylogenetic relationships of 445 NC from five different breeds (NC, Leghorn [LH], Cornish [CS], Rhode Island Red [RIR], and Hanhyup [HH] commercial line) were investigated by performing genotyping using 20 microsatellite markers.

Results: The highest genetic distance was observed between RIR and LH (18.9%), whereas the lowest genetic distance was observed between HH and NC (2.7%). In the principal coordinates analysis (PCoA) illustrated by the first component, LH was clearly separated from the other groups. The correspondence analysis showed close relationship among individuals belonging to the NC, CS, and HH lines. From the STRUCTURE program, the presence of 5 clusters was detected and it was found that the proportion of membership in the different clusters was almost comparable among the breeds with the exception of one breed (HH), although it was highest in LH (0.987) and lowest in CS (0.578). For the cluster 1 it was high in HH (0.582) and in CS (0.368), while for the cluster 4 it was relatively higher in HH (0.392) than other breeds.

Conclusion: Our study showed useful genetic diversity and phylogenetic relationship data that can be utilized for NC breeding and development by the commercial chicken industry to meet consumer demands.

Keywords: Microsatellite Marker; Polymorphism; Genetic Diversity; Phylogenetic Relationships; Korean Native Chicken

INTRODUCTION

About 60 million Korean native chickens (NC) are consumed annually in Korea, which accounts for only 10% of the total chickens consumed [1]. Owing to changes in eating habits, more and more Koreans have started preferring chicken meat, resulting in an increase in consumption [2]. Roasted chicken, Dak galbi (spicy stir-fried chicken), and chicken breast salad utilizing general broiler and imported chicken meat are mainly consumed in Korea. Since NC are used only for seasonal preparations such as Baeksuk (chicken soup with rice), Samgyetang (chicken soup with ginseng), and Braised spicy chicken, the proportion of NC is relatively low. Further, White slender tail roots are disguised as NC, which contributed to its lower consumption percentage [3]. According to the study reported by Lee et al [4] flavor- and taste-contributing amino acids, including aspartic acid, threonine, serine, glycine, alanine, tyrosine, lysine, histidine, and arginine, are significantly higher in the NC breast meat than in other broilers. Native chicken meat is not only light because of its flavory taste and low fat content, but it is also chewy because of solid muscle [5]. Through many research results, the superiority of NC meat has been proved. However, the growth of NC industry is marginal, owing to the consumption of a variety of other chicken breeds in
Korea. The NC breeds such as Hanhyup 3 and Wooirimatdag are commercial chicken breeds developed taking into consideration the merits and economic feasibility of NC, and produced by mating the NC and economically superior and naturalized breeds [1]. This industrialization strategy of using commercial chicken is an effective counterplan for changing consumer demands and market economics. However, for the sustainable growth of NC market, strategies for the improvement should be established based on the precise genetic background for NC breeds. Wooirimatdag was developed from pure NC population at the National Institute of Animal Science (NIAS) in Korea, and various genetic research projects on pure populations have been done and are still underway. Although many researchers have studied the genetic diversity and phylogenetic relationships of NC and foreign breeds using microsatellite (MS) markers, studies on the genetic background or characteristics of commercial chicken lines are still lacking. The MS markers can be suitable for a wide range of applications and have been commonly used as markers for studies of population genetics and paternity since early 1990s. The MS markers are likely to be used continuously in the future since they are comparatively cheap for genotyping and provide more genetic information on population than biallelic markers such as single nucleotide polymorphisms [6]. This experiment was conducted to investigate the basic information on genetic structure and characteristics of NC and foreign breeds through the analysis of the pure chicken populations and commercial chicken lines of the Hanhyup Company which are very popular in the NC market, using the 20 MS markers.

MATERIALS AND METHODS

Sample collection and DNA extraction

Commercial lines of chickens were collected from Hanhyup company, while NC (red, yellow, gray, black, and white) and naturalized breeds such as Rhode Island Red (RIR), Cornish (CS), Leghorn (LH) (which were introduced into Korea from foreign countries a long time ago) were collected from the NIAS in Korea (Table 1).

Five breeds of NC (red, yellow, gray, black, and white) reconstructed by NC conservation project (1994) were used. However, the genetic origins of commercial lines were not identified because it is confidential information of the company.

Blood samples from wing veins were collected from 445 individuals in tubes containing ethylenediaminetetraacetic acid for DNA extraction. Samples were stored at –20°C, and genomic DNA was extracted using the manual extraction method.

Microsatellite markers and polymerase chain reaction amplification

Initially, 28 MS markers were genotyped, and all MS markers were labeled with a capillary-based dye—FAM (blue), VIC (green), or NED (yellow)—to enable genotyping. From these results, 20 highly polymorphic MS markers were selected based on the number of alleles, expected heterozygosity ($H_{exp}$) and polymorphic information content (PIC) values (Table 2). All 445 DNA samples were amplified using a GeneAmp PCR system 9700 (Applied Biosystems, Foster City, CA, USA). The amplifications were carried out using 15 μL reaction mixtures containing genomic DNA (5 to 20 ng/μL), each primer (10 pM), 2.5 mM dNTP, 10X Maxima Hot Start Taq buffer, 25 mM MgCl₂, and 1.5 U Thermo Scientific Maxima Hot Start Taq DNA polymerase (Thermo Scientific, Waltham, MA, USA). The initial denaturation was performed at 95°C for 10 min, followed by 30 cycles of 30 s at 95°C, 30 s at 55°C, and 30 s at 72°C. The final extension step was at 72°C for 15 min and then cooled to 8°C.

Genotyping

The amplified DNA was genotyped using an automated Genetic Analyzer 3130xl (Applied Biosystems, USA). The genotyping reaction contained 1 μL of polymerase chain reaction (PCR) products, 8.9 μL of Hi-Di formamide, and 0.1 μL of GeneScan-500LIZ size standard in 10 μL total volume. The genotyping results were obtained using GeneMapper V 4.1 (Applied Biosystems, USA).

Statistical analysis

The genotyped data were analyzed using Cervus V 3.0 program [7] and Excel MS toolkit version 3.1 [8] to calculate allele frequencies at each locus for each population, $H_{exp}$, observed heterozygosity ($H_{obs}$), and PIC values. The amount of inbreeding-like effects within subpopulations (genetic distance $[F_{is}]$), among subpopulations (within inbreeding $[F_{in}]$), and within the entire population (total inbreeding $[F_{it}]$) were analyzed by $F$-statistics [9]. The STRUCTURE program [10] was used to analyze the Proportion of membership of the studied breeds. The neighbor joining method [11] was used to construct a phylogenetic tree. The principal coordinates analysis (PCoA) was carried out for 5 breeds using GenAlEx 6.4 program. The correspondence analysis, which is a weighted principal component analysis, was performed using the allele frequency data for the individuals of all five breeds and the 20 MS markers using GENETIX software [12].

RESULTS AND DISCUSSION

The $H_{exp}$, $H_{obs}$, and PIC values for the five chicken breeds are
summarized in Table 3. Among the 20 MS markers selected, the number of alleles was confirmed in the range of 5 (ADL0268) to 20 (MCW0127). Total number of alleles with 20 markers was 250, and the mean number of alleles (MNA) was 12.5. The $H_{obs}$ ranged from 0.349 (GCT0016) to 0.683 (MCW0145). MCW0104 showed the highest $H_{exp}$ and PIC values of 0.878 and 0.866 respectively, while ROS0083 showed the lowest $H_{exp}$ and PIC values of 0.742 and 0.708 respectively. Estimation of genotypic diversity in heterozygosity and PIC value informativeness of MS markers were previously used for determining animal breed selection [13]. For the animal traceability, PIC>0.5 and $H_{exp}$>0.6 are the most reasonable informative locus for application in genetics [14]. In this study, 15 MS markers selected were highly informative for the five chicken breeds and are appropriate for discrimination as well.

$F$-statistics [9] were estimated in a fixation index as genetic differentiation ($F_{st}$), the global heterozygote deficit among five chicken breeds ($F_{it}$), and the heterozygote deficit within the breed/line ($F_{is}$) among the 20 MS markers (Table 3). Among these markers, estimation of fixation index has been discovered for $F_{st}$, $F_{it}$, and $F_{is}$ values ranging from 0.048 to 0.182, 0.200 to 0.551, and 0.111 to 0.442, respectively. The estimated mean value of the total inbreeding ($F_{it}$), within line inbreeding ($F_{is}$) and genetic distance were 0.308, 0.229, and 0.102 respectively. The $F$-statistics values for $F_{st}$, $F_{it}$, and $F_{is}$ were high (0.182, 0.551, and 0.442, respectively) for the GCT0016 marker. While the lowest values for genetic

| No | Markers | Chr | Primer | Size | Dye |
|----|---------|-----|--------|------|-----|
| 1  | ADL0259 | 9   | F CTCATTGACAGGGTGAATTCGTCATTCTCTTCAGGTTGCT | 107-147 | VIC |
| 2  | ADL0268 | 1   | F CTCACCCCTCTTCAGAAGAATGT | 105-117 | PET |
| 3  | ADL0292 | 5   | F CGAATGCCGACGAAATTT | 110-1140 | FAM |
| 4  | ADL0293 | 17  | F GATACACGTCTTGTTTGCATTCT | 107-277 | PET |
| 5  | ADL0304 | 18  | F GGGGAACACTGTTGAAATGGT | 138-162 | PET |
| 6  | ADL0317 | 4   | F AGTGTGTTCAAGCGCCTCAT | 178-204 | FAM |
| 7  | GCT0016 | 9   | F TCCAGGTTCTTCAGGTTG | 108-168 | VIC |
| 8  | LEI0074 | 26  | F GACCTGGATCGATGATGTTG | 132-244 | VIC |
| 9  | LEI0094 | 4   | F GATCCACAGTAGTAGCTGC | 232-282 | FAM |
| 10 | LEI0141 | 2   | F GCAATGGATGATGATGTTG | 218-244 | PET |
| 11 | MCW0029 | 5   | F GTGGACCACTTGGTTGCT | 131-187 | VIC |
| 12 | MCW0063 | 2   | F GGTCCCAAAAGCTTCTAACGT | 128-150 | FAM |
| 13 | MCW0087 | 2   | F ATTCCTGCCGCACTGTTG | 265-289 | FAM |
| 14 | MCW0104 | 13  | F GTTCAGCAATCGCTAGTG | 192-232 | FAM |
| 15 | MCW0127 | 3   | F GATTTTCAGGAGTGGAGTAGT | 226-274 | FAM |
| 16 | MCW0145 | 1   | F ACCTTTCTGCTCAGGTTG | 178-214 | FAM |
| 17 | MCW0228 | 10  | F TATTTTCACACTGCTGATCAG | 220-248 | PET |
| 18 | MCW0264 | 2   | F AGTCTCCGACGCCTGCTGAACT | 225-243 | FAM |
| 19 | ROS0013 | 5   | F TCTTCTTCAAGCAGAAGACG | 220-262 | VIC |
| 20 | ROS0083 | 13  | F CATTACAGAGGATGGGTTACC | 108-130 | VIC |
The statistical analysis of heterozygosity ($H_{ob}$ and $H_{exp}$), polymorphism information content (PIC), and F-statistics value using selected 20 microsatellite markers among 5 chicken breeds.

| Locus       | No. of alleles | $H_{ob}$ | $H_{exp}$ | PIC  | $F_{st}$ (t) | $F_{is}$ (F) | $F_{st}$ (t) |
|------------|---------------|---------|---------|------|-------------|-------------|-------------|
| ADL0259    | 14            | 0.675   | 0.874   | 0.862| 0.081       | 0.223       | 0.156       |
| ADL0268    | 5             | 0.516   | 0.762   | 0.725| 0.117       | 0.310       | 0.215       |
| ADL0292    | 11            | 0.562   | 0.751   | 0.714| 0.152       | 0.267       | 0.137       |
| ADL0293    | 10            | 0.497   | 0.778   | 0.745| 0.139       | 0.360       | 0.256       |
| ADL0304    | 9             | 0.488   | 0.784   | 0.753| 0.092       | 0.375       | 0.308       |
| ADL0317    | 11            | 0.611   | 0.834   | 0.814| 0.143       | 0.272       | 0.152       |
| GC10016    | 12            | 0.349   | 0.806   | 0.78 | 0.182       | 0.551       | 0.442       |
| LEI0074    | 13            | 0.525   | 0.831   | 0.82 | 0.053       | 0.362       | 0.323       |
| LEI0094    | 18            | 0.597   | 0.831   | 0.81 | 0.06        | 0.283       | 0.236       |
| LEI1014    | 14            | 0.466   | 0.851   | 0.832| 0.087       | 0.440       | 0.384       |
| MCW0029    | 15            | 0.673   | 0.857   | 0.843| 0.111       | 0.210       | 0.111       |
| MCW0063    | 11            | 0.577   | 0.763   | 0.731| 0.128       | 0.248       | 0.136       |
| MCW0087    | 14            | 0.555   | 0.878   | 0.865| 0.055       | 0.365       | 0.328       |
| MCW0104    | 17            | 0.631   | 0.878   | 0.866| 0.125       | 0.275       | 0.171       |
| MCW0127    | 20            | 0.595   | 0.863   | 0.848| 0.093       | 0.327       | 0.259       |
| MCW0145    | 11            | 0.683   | 0.841   | 0.821| 0.048       | 0.200       | 0.160       |
| MCW0228    | 12            | 0.585   | 0.796   | 0.769| 0.108       | 0.273       | 0.191       |
| MCW0264    | 9             | 0.596   | 0.824   | 0.803| 0.097       | 0.284       | 0.207       |
| ROS0013    | 16            | 0.62    | 0.837   | 0.817| 0.075       | 0.259       | 0.197       |
| ROS0083    | 8             | 0.551   | 0.742   | 0.708| 0.094       | 0.276       | 0.202       |
| Total/mean | 250/12.5      | 0.568   | 0.820   | 0.7963| 0.102       | 0.308       | 0.229       |

$H_{ob}$, observed heterozygosity; $H_{exp}$, expected heterozygosity; PIC, polymorphism information content; $F_{st}$, genetic distance; $F_{is}$, total inbreeding; $F_{st}$, within inbreeding.

The genetic divergences among the populations based on allele frequencies were calculated according to $D_{A}$ genetic distance [11]. The genetic distances of 5 chicken breeds were in the range of 0.131 (HH and NC) to 0.545 (RIR and LH) (Table 5) and neighbour-joining tree of genetic relationship among 5 chicken breeds using $D_{A}$ genetic distance is shown in Figure 1. HH and NC were grouped into the same branch.

Figure 2 illustrates the population relationships based on the PCoA using individual multi-locus genotypes of 20 MS markers. Contribution to the variance of the principal components was over 90% including the third ingredient. The first two components contributed 47.85% and 33.56% respectively, and the third component contributed 12.42%. Notably, by the first component, LH was confirmed to be clearly separated from the other groups. CS was confirmed near the NC and HH by the second component. On the other hand, it was separated from HH and NC under the third component.

The correspondence analysis was performed on the five populations using allele frequencies of the 20 MS markers to summarize the breed relationships (Figure 3). The first four factors contributed to a total of 99.99% with values of 39.68%, 25.03%, 19.53%, and 15.75% respectively. Figure 3 shows close relationship among individuals which belong to the NC (pink plots), CS

Table 4. Mean number of alleles (MNA), heterozygosity ($H_{ob}$ and $H_{exp}$), and polymorphism information content (PIC) observed across 20 microsatellite loci for each population

| Population | MNA    | $H_{ob}$ | $H_{exp}$ | PIC  |
|------------|--------|---------|---------|------|
| HH         | 9.90   | 0.612   | 0.785   | 0.755|
| NC         | 10.05  | 0.630   | 0.796   | 0.767|
| LH         | 5.35   | 0.420   | 0.586   | 0.537|
| CS         | 6.50   | 0.494   | 0.721   | 0.674|
| RIR        | 4.40   | 0.464   | 0.580   | 0.520|
| Mean       | 7.240  | 0.524   | 0.694   | 0.650|

$H_{ob}$, observed heterozygosity; $H_{exp}$, expected heterozygosity; HH, Hanhyup (Commercial line); RIR, Rhode Island Red; LH, Leghorn; CS, Cornish; NC, Korean Native Chicken.

Table 5. $D_{A}$ genetic distance matrix estimated from the frequencies of 20 microsatellite loci among the five chicken breeds

| Population | HH    | RIR   | LH    | CS    | NC    |
|------------|-------|-------|-------|-------|-------|
| HH         | -     | 0.301 | -     | -     | -     |
| RIR        | 0.301 | -     | 0.545 | -     | -     |
| LH         | 0.434 | 0.545 | -     | -     | -     |
| CS         | 0.190 | 0.396 | 0.421 | -     | -     |
| NC         | 0.131 | 0.337 | 0.403 | 0.203 | -     |

HH, Hanhyup (Commercial line); RIR, Rhode Island Red; LH, Leghorn; CS, Cornish; NC, Korean native chicken.
Table 6 illustrates the results of the analysis performed using STRUCTURE program. The presence of 5 clusters was detected and the proportion of membership in the different clusters was almost comparable among the breeds with the exception of one breed (HH), although it was highest in LH (0.987), and lowest in CS (0.578). For the cluster 1 it was high in HH (0.582) and in CS (0.368), while for the cluster 4 it was relatively higher in HH (0.392) than other breeds. The reason for this seems to be due to the variety of commercial pure lines by the improvement goals (for egg or meat production, etc). However, further research is needed to determine why HH line did not form a single cluster with high proportion of membership.

LH is known as the typical egg breed, while RIR is known as the dual purpose breed and CS as meat type breed [15]. According to this study, the CS breed was genetically close to NC and HH commercial lines. From these results CS, NC, and HH commercial lines seem to be genetically close because they have been improved for the same purpose (meat type breed).

The expected probabilities of identity values were calculated for random individuals (PI), random half-sib (PI_{half-sibs}), and random sibs (PI_{sibs}) using 10, 15, and 20 MS markers and their estimates were $3.88 \times 10^{-60}$, $4.86 \times 10^{-39}$, and $1.38 \times 10^{-12}$, respectively.
when calculated using 20 MS markers (Table 7). Seo et al [16] reported that random individuals (PI), random half-sib (PI half-sibs), and random sibs (PI sibs) with 15 markers were estimated as $7.98 \times 10^{-29}$, $2.88 \times 10^{-20}$, and $1.25 \times 10^{-8}$ respectively. On comparing the two results, it was confirmed that the expected probabilities of identity values from this study were higher than those of Seo’s Report [16]. When the domestic chicken market scale is considered, it seems to be sufficient if 15 markers are used. These results in this study indicated that these MS markers will be used to aid the conservation, traceability, and future improvement of chicken lines in Korea.

**CONFLICT OF INTEREST**

We certify that there is no conflict of interest with any financial organization regarding the material discussed in the manuscript.

**ACKNOWLEDGMENTS**

This work was supported by Golden Seed Project. (2013010-05-1-WT251).

**REFERENCES**

1. Lee KW, Oh JD, Lee JA, et al. Estimation of genetic characteristic and cumulative power of discrimination using the microsatellite markers in Korean native chicken. Korean J Poult Sci 2010;37:81-7.
2. Heo KN. The importance of the origin of genetic resources and the preservation of our Korean native chickens. Korean Poult J 2013;12:130-3.
3. Oh JD, Lee KW, Seo OS, et al. Estimation of genetic characteristics and cumulative power of discrimination in Korean native chicken and Korean native commercial chicken. J Life Sci 2010;20:1086-92.
4. Lee KH, Kim HJ, Lee HJ, Kang MG, Jo CU. A study on components related to flavor and taste in commercial broiler and Korean native chicken meat. Korean J Food Preserv 2012;19:385-92.
5. Lee MJ, Heo KN, Choi HC, Hong EC, Kim CD. The performance test in crossbreds of Korean native chickens for the establishment of new lines. Korean J Poult Sci 2014;41:39-44.
6. Dawson DA, Ball AD, Spurgin LG, et al. High-utility conserved avian
microsatellite markers enable parentage and population studies across a wide range of species. BMC Genomics 2013;14:1-22.
7. Marshall TC, Slate J, Kruuk LE, Pemberton JM. Statistical confidence for likelihood-based paternity inference in natural populations. Mol Ecol 1998;7:639-55.
8. Park SDE. Trypanotolerance in West African cattle and the population genetic effects of selection [Ph.D. Thesis]. Dublin, Ireland: University of Dublin; 2001.
9. Wright S. The interpretation of population structure by F-statistics with special regard to systems of mating. Evolution 1965;19:395-420.
10. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. Genetics 2000;155:945-59.
11. Saitou N, Nei M. The neighbor-joining method: A new method for reconstructing phylogenetic tree. Mol Biol Evol 1987;4:406-25.
12. Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F. GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II, Montpellier (France) [cited 2015 Feb 7]. Available from http://www.univ-montp2.fr/
13. Berthouly C, BedHorn B, Tixier-Boichard M, et al. Using molecular markers and multivariate methods to study the genetic diversity of local European and Asian chicken breeds. Anim Genet 2008;39:121-9.
14. Botstein D, White RL, Skolnick M, Davis RW. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J Hum Genet 1980;32:314-31.
15. Korea Food Science Association. Food science technology dictionary. Gwangil Press; 2008.
16. Seo DW, Hoque MR, Choi NR, et al. Discrimination of Korean native chicken lines using fifteen selected microsatellite markers. Asian-Australas J Anim Sci 2013;26:316-22.