JAST (Journal of Animal Science and Technology) TITLE PAGE
Upload this completed form to website with submission

ARTICLE INFORMATION	**Fill in information in each box below**
Article Type	Research article
Article Title (within 20 words without abbreviations)	Analysis of runs of homozygosity in Yeonsan Ogye chickens using 600K SNP arrays
Running Title (within 10 words)	Analysis of ROH in Yeonsan Ogye chickens using SNP arrays
Author	Jaewon Kim1, Minjun Kim1, Eunjin Cho2, Seung-sook Lee3, Seungchang Kim4, Daehyeok Jin4, Jun Heon Lee1,2
Affiliation	1 Division of Animal and Dairy Science, Chungnam National University, Daejeon 34134, Korea. 2 Department of Bio-AI Convergence, Chungnam National University, Daejeon 34134, Korea. 3 Yeonsan Ogye Foundation, Nonsan 32910, Korea. 4 Animal Genetic Resources Research Center, National Institute of Animal Science, Rural Development Administration, Hamyang 50000, Korea
ORCID (for more information, please visit https://orcid.org)	Jaewon Kim (https://orcid.org/0009-0006-0445-3025) Minjun Kim (https://orcid.org/0000-0002-8173-8431) Eunjin Cho (https://orcid.org/0000-0003-4800-1603) Seung-sook Lee (https:/orcid.org/0000-0002-1898-5448) Seungchang Kim (https://orcid.org/0000-0001-5724-0599) Daehyeok Jin (https://orcid.org/0000-0001-5091-4271) Jun Heon Lee (https://orcid.org/0000-0003-3996-9209)
Competing interests	No potential conflict of interest relevant to this article was reported.
Funding sources State funding sources (grants, funding sources, equipment, and supplies). Include name and number of grant if available.	This research was funded by the project funding number: RS-2021-RD010125(PJ016205) of the Rural Development Administration, South Korea.
Acknowledgements	Not applicable.
Availability of data and material	Upon reasonable request, the datasets of this study can be available from the corresponding author.
Authors' contributions Please specify the authors' role using this form.	Conceptualization: Kim JW, Lee JH. Data curation: Kim MJ, Cho EJ, Lee SS Formal analysis: Kim JW Methodology: Kim JW, Kim MJ, Cho EJ Software: Kim JW Validation: Kim JW, Kim MJ, Cho EJ, Kim SC, Jin DH Investigation: Kim JW, Kim MJ, Cho EJ Writing - original draft: Kim JW Writing - review & editing: Kim JW, Kim MJ, Cho EJ, Lee SS, Kim SC, Jin DH, Lee JH
Ethics approval and consent to participate	This article does not require IRB/IACUC approval because there are no human and animal participants.

CORRESPONDING AUTHOR CONTACT INFORMATION

For the corresponding author (responsible for correspondence, proofreading, and reprints)	**Fill in information in each box below**
First name, middle initial, last name	Jun Heon Lee
Email address – this is where your proofs will be sent	junheon@cnu.ac.kr
Secondary Email address	kimmin6023@gmail.com
Address	213, KTnG, Chungnam National University, 99, Daehak-ro, Yuseong-gu, Daejeon, Republic of Korea
Cell phone number	+82-10-5172-0816
Office phone number	+82-42-821-7031
Fax number	+82-42-825-9754

Abstract

Runs of homozygosity (ROHs) are caused by identical haplotypes inherited from ancestors. ROHs provide useful information regarding the inbreeding rate, demographics, and selection history. The Yeonsan Ogye (YO) breed is an indigenous chicken in Korea that is characterized by a completely black body. In this study, we investigated ROH in the YO genome to determine ROH-based inbreeding coefficients and their correlations with other inbreeding estimators, then analyzed their genetic characteristics. Using 600K single nucleotide polymorphism (SNP) chip information for 189 chickens, we found 20,339 ROHs in the YO population. The average number of ROHs was 107, the total average ROH length was 165 Mb, and the average ROH length was 1.542 Mb. Most ROHs were short (< 8 Mb), suggesting a past population bottleneck. The average inbreeding coefficient (F_{ROH}) calculated based on ROHs was 0.184 and this was correlated with other inbreeding coefficients estimated using allele frequencies. 17 ROH islands were detected and these regions exceeded the threshold of the top 1% of SNPs among SNPs present in ROHs. In the ROH islands, 152 genes were annotated, some of which were genes associated with meat production traits and hyperpigmentation in chickens. A comparison of overlapping regions between ROH islands and quantitative trait loci (QTLs) indicated that most QTLs were related to color traits. These results will help to optimize conservation strategies for the YO breed.

Keywords: local chicken breed, Yeonsan Ogye, runs of homozygosity, genomic inbreeding coefficient, ROH islands, conservation

Introduction

Yeonsan Ogye (YO), a traditional chicken breed in Korea, is designated as Natural Monument No. 265. The YO breed is characterized by distinct physical features, including black feathers, skin, pupils, and bones [1]. For conservation purposes, the YO breed maintained a population of over 1,000 individuals and selected the parent stocks every year without pedigree information. Selection is solely based on external black phenotypic traits, regardless of genetic diversity and inbreeding rates [2]. However, pedigree information is crucial to manage and control inbreeding among individuals, especially for conserved breeds such as the YO [3]. Therefore, a distinct breeding system is needed to prevent potential future inbreeding depression and preserve genetic diversity.

Inbreeding is the mating of related individuals, which increases allele homozygosity in a population and causes inbreeding depression. Inbreeding depression increases the potential for recessive genetic diseases and affects livestock productivity. An inbreeding coefficient, a measure used to estimate inbreeding [4], is defined as the probability that two homozygous alleles in an individual are the same allele derived from a common ancestor (i.e., the probability of identical by descent [IBD]). Traditionally, inbreeding coefficients are calculated using pedigree information [5]. When pedigree information is unavailable, an inbreeding coefficient can be obtained via molecular genetic information using a single nucleotide polymorphism (SNP) array [6]. One of the methods using SNP chip information to measure the level of IBD in livestock is measuring the proportion of runs of homozygosity (ROHs) in the genome [7].

ROH is a contiguous diploid homozygous segment in the genome that is not interrupted by heterozygous alleles [8]. Considering that DNA fragments separate during genomic recombination, the likelihood that long homozygous DNA sequences remain contiguous decreases over generations [9]. Therefore, long ROHs often indicate recent consanguineous mating, whereas short ROHs presumably originated from more distant common ancestors [7]. The inbreeding coefficient based on ROHs (F_{ROH}) can be used to estimate the degree of inbreeding and genetic relatedness among individuals, enabling assessment of the actual level of autozygosity in livestock [10, 11]. Additionally, some ROH characteristics in a population, such as the average ROH length, average number of ROHs, and ROH distribution across the chromosomes, can be used as indicators of various genetic phenomena [11]. ROHs tend to occur in substantial proportions within specific chromosomal regions because of reduced haplotype diversity. These regions, known as ROH islands [7], are used to identify portions of the genome that have undergone selective pressure and are associated with beneficial traits or adaptations [7, 10, 11].

In this study, we used ROH analysis to assess genomic inbreeding and genetic characteristics in the YO population. We confirmed correlations between inbreeding coefficients obtained via ROH analysis and other inbreeding coefficients. Furthermore, we identified features of specific genomic regions with many ROHs in the population.

Materials and Methods

Study Population

This study examined 189 YO chickens. The selected samples were parent stocks used in 2018. Genomic DNA (gDNA) was extracted from chicken blood using PrimePrep™ Genomic DNA Isolation kits (GeNetBio, Daejeon, Korea). The concentration and purity of the isolated gDNA were measured using a NanoDrop2000c spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA). The extracted gDNA was stored at –20°C until use.

SNP genotyping and data filtering

gDNA samples were genotyped by using the 600K Affymetrix Axiom Chicken SNP panel (Affymetrix, Santa Clara, CA, USA), which revealed 546,137 SNPs. For greater accuracy in subsequent analyses, a quality control (QC) procedure was performed using PLINK v.1.9 [12]. Using the "--geno" option, 5,385 SNPs with a calling rate of < 90% were excluded; 5,056 SNPs with Hardy–Weinberg equilibrium test p-values < 10^{-6} were excluded using the "--hwe" option. The "--maf" option, filtering option using minor allele frequency (MAF), was not used for ROH analysis [14]. As a result, 535,696 SNPs were analyzed.

ROH analysis

ROH analysis of the YO population was conducted using the "--homozyg" option of PLINK v.1.9, which uses a sliding window method that continuously scans an individual's SNP data to identify homozygous regions. The parameter settings have a substantial impact on ROH analysis. However, the default parameter values provided by PLINK may be suboptimal because they depend on factors such as the SNP array density and genomic characteristics of the samples undergoing analysis. As noted by Meyermans et al. [13], this issue can be addressed by choosing a method to determine appropriate parameter values. In this study, we followed the approaches recommended by Meyermans et al. [13] and Gorrsen et al. [14] to determine parameter values used for ROH analysis (Table 1).

The minimum number of SNPs in ROHs (--homozyg-snp) was calculated using formula (1) [10].

$$(1)\ L = \frac{\ln\frac{a}{n_s \cdot n_i}}{\ln(1-\overline{het})},$$

where L is the number of consecutive SNPs constituting an ROH, n_s is the number of SNPs, n_i is the number of individuals, a is the false positive rate of ROHs (set to 0.05), and $\overline{het}$ is the average heterozygosity of all SNPs in the population. The average heterozygosity of the YO was determined using the "--hardy" option of PLINK v.1.9. The size of the sliding window was also set to L [13]. The minimum number of heterozygotes in the sliding window (--homozyg-window-het) and minimum number of missing SNPs in the sliding window (--homozyg-window-missing) were set to 1 and 3, respectively [11].

The maximum gap between SNPs (--homozyg-gap) in an ROH was calculated using the genome coverage proposed by Meyermans et al. [13]. The method for determining gap size based on genome coverage initially involved the artificial generation of an individual in which all SNPs were homozygous. Subsequently, by modifying a specific parameter, the ratio of the total length of detected ROHs in the completely homozygous sample to the length of the autosomal genome was calculated. For an artificially created completely homozygous organism, the entire genome represented a single ROH, and genome coverage referred to the maximum detectable coverage achievable using the same parameter value [13]. To generate completely homozygous individuals, a bim (PLINK extension map) file containing SNP information was used to extract the major alleles. Genome coverage was calculated by varying the maximum gap parameter between SNPs from 1 to 300; other parameters were kept constant with the values shown in Table 1.

The minimum ROH length (--homozyg-kb) was set to 500 kb. This is the minimum length that can be obtained at a density of 600K SNPs without the inclusion of very short ROHs generated by linkage disequilibrium [15]. Other parameters were set to the default values in PLINK. To investigate the genomic characteristics of the YO population using ROHs, the total number and total length of each ROH, as well as the average ROH length, were obtained using the "detectRUNS" package in R [16]. Additionally, the identified ROHs were categorized into five length classes: 0.5–2, 2–4, 4–8, 8–16, and > 16 Mb [15].

Genomic inbreeding value calculation

Four different methods were used to calculate the inbreeding coefficient of the YO population. The first is the inbreeding coefficient based on ROH, calculated using formula (2) [17].

$$(2)\ F_{ROH} = \frac{\sum L_{ROH}}{L_{Auto}},$$

where $\sum L_{ROH}$ is the length of all ROHs in an individual and L_{Auto} is the length of the autosomal genome [17]. **F_{ROH}** was calculated using the ″detectRUNS″ package in R [16].

The second is the inbreeding coefficient based on homozygous SNPs, calculated using formula (3) [18].

$$(3)\ F_{HOM} = \frac{O - E}{L - E},$$

where O is the number of observed homozygous SNPs, E is the number of expected homozygous SNPs, and L is the total number of SNPs in an individual.

The third is the inbreeding coefficient [19] based on diagonal elements of the genomic relationship matrix (GRM), calculated using formula (4) [19].

$$(4)\ F_{GRM} = \frac{1}{m}\sum_{i=1}^{m}\left(\frac{[x_i - E(x_i)]^2}{2p_i(1 - p_i)} - 1\right),$$

where x_i is the number of reference alleles of the i^{th} SNP, m is the total number of SNPs, and p_i is the frequency of the reference allele.

The fourth is the inbreeding coefficient calculated based on the correlation of gametes, calculated using formula (5) [18].

$$(5)\ F_{UNI} = \frac{{x_i}^2 - (1+2p_i)x_i + 2{p_i}^2}{2p_i(1-p_i)},$$

where x_i is the number of reference alleles of the i^{h} SNP, and p_i is the frequency of the reference allele.

F_{HOM}, **F_{GRM}**, and **F_{UNI}** were calculated using the "--ibc" option of GCTA [18]. Pearson correlation coefficients were calculated to determine correlations among the four calculated inbreeding coefficients, using the "ggpbur" package in R.

Detection of ROH islands and gene annotation

The percentage of SNPs located in an ROH region was calculated to identify ROH islands, regions where individuals in a group have common ROHs. The percentages were calculated through the division of the number of corresponding SNPs present in the ROHs of individuals by the total number of individuals. Among the SNPs present in ROHs, the top 1% of SNPs was set as the threshold; a series of adjacent SNPs over the threshold was designated as an ROH island [20]. To search for genes in ROH islands, candidate genes associated with SNPs were annotated using the chicken SNP annotation information (GRCg6a.103) in BioMart [21]. Quantitative trait loci (QTLs) present in ROH islands were identified using the "GALLO" package in R [22]. QTL analysis was performed based on GRCg6a gff (genome annotation file) file information in the QTL database [23].

Results

Optimization for ROH analysis

To determine the optimal value for the maximum gap between SNPs in an ROH, the genome coverage of an artificially created completely homozygous individual was assessed by changing the parameter value from 1 to 300. As a result, the genome coverage showed a notable increase from 4 kb/SNP in 600K SNP data. The maximum gap size of 80 kb/SNP reached 99% coverage; this value was selected as the parameter for the maximum SNP gap in an ROH. After reaching the maximum coverage of 99.34% at a gap size of 220 kb, we observed no further changes in coverage as the gap length increased (Fig. 1).

ROH analysis

ROH analysis identified 20,339 ROHs in the 189 YO population. The average and total lengths of ROHs based on the number of ROHs per individual are shown using scatterplots (Figs. 2A and 2B) and violin plots (Figs. 2C and 2D).

The average ROH length ranged from approximately 2.15 to 0.66 Mb, and the total ROH length varied between about 279 and 16 Mb. The shortest average and total ROH lengths were present in the same individual. The average number of ROHs observed in the YO population was approximately 107, with an average total ROH length of 165 Mb. The average length of an individual ROH was 1.542 Mb (Table 2).

The frequency distribution and average length of ROHs were analyzed for individual chromosomes. Chromosome 1 comprised the largest proportion, approximately 19% of the total length. Generally, the proportion of ROHs decreased with chromosomal length (Fig. 3A). Chromosome 5 had the longest average ROH length at 1.834 Mb, followed by chromosome 2 with 1.833 Mb and chromosome 3 with 1.776 Mb (Fig. 3B).

ROHs were classified into five categories according to length: 0.5‑2, 2‑4, 4‑8, 8‑16, and > 16 Mb. Of the 20,339 ROHs detected in the YO populations, the majority (79%) had a length of ≤ 2 Mb. In contrast, long ROHs with lengths ≥ 8 Mb constituted only 0.01% of the total ROH length. Four ROHs with lengths > 16 Mb were identified; all were on chromosomes 1 and 2. The longest ROH segment (20.85 Mb) was present on chromosome 2 (Fig. 3C). The average ROH lengths in the five categories were 0.98, 2.75, 5.26, 10.08, and 19.19 Mb, respectively (Fig. 3D).

Genomic inbreeding coefficients

In this study, inbreeding coefficients were determined using four equations to confirm the inbreeding of the YO population. F_{ROH} obtained from the ROH information was 0.178. However, all three inbreeding coefficients (F_{HOM}, F_{GRM}, and F_{UNI}) calculated using genomic information had negative average values; F_{GRM} had the lowest value (–0.1344) (Table 3). Examination of correlations among inbreeding coefficients revealed that F_{ROH} had the highest positive correlation (0.47) with F_{HOM}, followed by F_{UNI} (0.2). In contrast, F_{ROH} had a negative correlation (–0.25) with F_{GRM} (Fig. 4).

Detection of ROH islands and functional annotation

An ROH island was designated as an area exceeding 49.20%, which corresponds to the top 1% of SNPs among the 534,705 SNPs in ROHs (Fig. 5). 17 ROH islands were identified on six chromosomes; the shortest ROH island was 233 bp (GGA20) and the longest ROH island was 2,069,982 bp (GGA20) (Table 4). Chromosome 5 had the most ROH islands (eight). In total, 152 genes that have gene symbol were annotated in the ROH islands (Table 4). The ROH island on chromosome 11 had the most annotated genes (39).

QTL regions overlapping the ROH islands were confirmed using the QTL database [26]. QTLs were classified into four categories according to relevant traits: exterior, health, physiology, and production. 30 QTLs were overlapped with ROH islands (Fig. 6A). QTLs were related to exterior traits (68%), production traits (28%), and health traits (3%) (Fig. 6B). The QTL enrichment analysis showed that five QTLs had a false discovery rate (FDR) $\leq$ 0.05; three were exterior-related QTLs and two were production-related QTLs including egg number and age of sexual maturity (Fig. 6C). The QTLs with the highest significant FDR *p*-value were related to skin color and comb color (Table 5).

Discussion

Genetic diversity information of populations is necessary for the development of sustainable conservation strategies in livestock resources. The YO population currently consists of approximately 1,000 individuals. To preserve the YO population, a comprehensive preservation approach includes a dispersed conservation strategy of about 200 subpopulations and keeping the cryopreservation of semen [24]. Previous research conducted that the genetic diversity of the YO population was analyzed using 12 microsatellite markers as part of a continuous effort to minimize inbreeding. This analysis revealed the YO population can be divided into five clusters according to genetic distances among individuals within the population [2]. Although the conservation plan has been continuously improved, the estimation of inbreeding within a population remains challenging because of the absence of pedigree information, which is invaluable for determining inbreeding between individuals. In this study, we estimated the level of inbreeding within the YO population by analyzing ROH.

Optimization for ROH analysis

Various ROH studies have been conducted in livestock. However, the absence of a consensus regarding the definitions of parameters used in ROH analysis is a major challenge [7, 13]. Meyermans et al. [13] observed that appropriate parameter values can vary among species and according to SNP data density. Thus, it is important to establish parameter values that are appropriate for specific SNP chip densities [13]. Here, we adopted the genome coverage method of Meyermans et al. [13] to determine suitable parameter values for the high-density chip used.

The length of the gap between SNPs determines the inclusion of homozygous SNPs within the same ROH segment. As a result, the SNP gap length has a substantial effect on ROH detection. For our analysis, we considered a maximum interval of 80 kb, which resulted in a 99% coverage rate, for the ROH analysis. This interval was smaller than the default PLINK value of 1,000 kb. Moreover, inaccurate SNP density parameters can affect ROH detection in regions with low SNP density [13]. However, the high-density SNP data we used contain an average SNP density of > 2 kb/SNP per chromosome [25]. Consequently, the density does not significantly affect the analysis unless the value decreases to < 2 kb/SNP. Because the default PLINK value of 50 kb/SNP exceeds the average density of high-density SNP arrays, we decided to use this default value.

Additionally, SNP pruning using the MAF during the QC process may affect the ROH results [13]. A SNP with a low MAF could refer to a SNP that nearly fixed within the population, indicating a lack of variation in that specific SNP. Thus, the pruning of rare MAF SNPs affects the detection of continuous homozygous segments. Consequently, the removal of SNPs in the QC process leads to measured ROH lengths that are shorter than their true lengths [13]. Therefore, our analysis used SNP data that did not exclude rare MAF SNPs.

ROH analysis

ROH is a valuable genomic feature that enables the examination of inbreeding and homozygous patterns. We focused on the genome-wide distribution of ROHs and the frequency of ROHs across different length categories. Our analysis revealed ROHs in 27 of the 28 chromosomes, excluding chromosome 16. In the chicken, chromosome 16 is very short (ca. 539 kb) and includes the MHC-B and MHC-Y regions, which contain major histocompatibility complex (MHC) genes with high genetic diversity [25, 26]. The genetic diversity of these regions contributes to the native chicken immune response [26]. Therefore, the absence of ROHs on chromosome 16 was presumably related to its allele diversity and short length (Fig. 3B). The ratio of ROH length to chromosomal length tended to increase with chromosomal length. When ROHs were categorized according to length, most were shorter than 2 Mb; ROHs longer than 8 Mb constituted approximately 1% of all ROHs (Fig. 3C). ROHs exceeding 8 Mb are typically generated by recent inbreeding, whereas ROHs shorter than 8 Mb are derived from common ancestors in more distant generations [9]. These findings indicate that recent inbreeding events were rare in the YO population. Moreover, the presence of numerous shorter ROHs is attributed to the bottleneck effect caused by a past reduction in effective population size [15]. These results are consistent with a previous report concerning a historical bottleneck in the YO population [27].

Genomic inbreeding coefficients

One of our objectives was to compare inbreeding coefficients calculated using different methods. The average inbreeding coefficient based on ROHs was 0.178. Compared with previous research that identified F_{ROH} with 600K SNP chip data in Chinese chicken breeds, the level of F_{ROH} in YO was lower than F_{ROH} of commercial breeds and similar to Chinese indigenous chickens [28]. Therefore, YO has maintained a moderate level of F_{ROH} as indigenous chicken breeds. We assumed that the YO population might have been less affected by inbreeding because the YO population was not subjected to artificial selection based on rigorous criteria conducted on other commercial breeds for production traits. F_{ROH} was calculated as a positive value, while the other inbreeding coefficients calculated by using genomic information had negative values (Table 3). These inbreeding coefficients reflect the expected frequencies of homozygosity based on Hardy–Weinberg equilibrium or the correlation between alleles present in individuals. If the observed level of homozygosity within a population exceeds the expected value, the inbreeding coefficients will have positive values; a negative value indicates observed homozygosity that is lower than expected [6]. Accordingly, inbreeding coefficients based on genomic information suggested that the YO population has levels of homozygosity below predicted values. However, these estimators vary according to allele frequencies present in the population [6, 29]. Inbreeding coefficients based on ROHs, such as $\boldsymbol{F_{ROH}}$, provide more accurate estimates of homozygosity in the genome, compared with inbreeding coefficients based on genomic information. Therefore, $\boldsymbol{F_{ROH}}$ can assess the actual loss of heterozygosity, regardless of allele frequencies, and is more reasonable for small population sizes, such as endangered or conserved populations [4]. Hence, using $\boldsymbol{F_{ROH}}$ could be useful information to detect the inbreeding rate of the YO population, which has no pedigree information, and it would be an effective way to track the variation of the inbreeding rate over generations.

The correlations of $\boldsymbol{F_{ROH}}$ with $\boldsymbol{F_{HOM}}$, and $\boldsymbol{F_{UNI}}$ were all positive, whereas the correlation of with $\boldsymbol{F_{GRM}}$ was negative. This is because $\boldsymbol{F_{ROH}}$ and $\boldsymbol{F_{HOM}}$ give equal weight to all alleles, whereas $\boldsymbol{F_{GRM}}$ gives more weight when rare alleles are homozygous [4]. Therefore, $\boldsymbol{F_{GRM}}$ tends to have a negative correlation with $\boldsymbol{F_{ROH}}$ in a population with more rare minor alleles [29]. In the YO population, the negative correlation between $\boldsymbol{F_{ROH}}$ and $\boldsymbol{F_{GRM}}$ might have been attributed to the lack of MAF pruning during the SNP QC process.

Detection of ROH islands and functional annotation

ROH islands represent genomic regions where genetic diversity associated with selection for specific traits has been reduced, and they indicate genomic regions related to selection [30]. Some ROH islands, such as the island on chromosome 3, contained SNPs that were annotated with long non-coding RNAs and micro RNAs, not coding genes. This finding suggests that these regions might be caused by selection acting on uncharacterized noncoding DNA regions and gene regulatory regions, or fixation occurred because of genetic drift [31].

Many coding genes were present in the ROH islands. The ROH island region on chromosome 5 (2.085–3.522 Mb) was reported in indigenous chicken breeds from various countries [30, 32, 33]. This region contained 11 genes, including the *ANO5* and *NELL1* genes associated with production-related traits. *ANO5* is involved in muscle tissue development and estrogen production, whereas *NELL1* is associated with skeletal tissue formation. These genes have important roles in the high body weight gain of broiler chickens [34, 35]. ROH islands containing genes associated with the unique phenotype, a completely black color in YO, were also discovered. The *MC1R* gene in the ROH island on chromosome 11 is involved in melanin synthesis and influences feather coloration in chickens [36]. Furthermore, the ROH island region on chromosome 20 overlapped a region that is likely involved in skin pigmentation in the Korean native chicken [37]. Genes within this ROH island, such as *GNAS* and *RBM38*, might be associated with visceral peritoneum hyperpigmentation in chickens [38, 39]. *GNAS* gene encodes G protein α-subunit protein ($G_s\alpha$) that interact with various G protein-coupled receptor (GPCR) and stimulate the upregulation of cAMP by adenyl cyclase [40]. *MC1R* is the gene that encodes the GPCR melanocortin-1 receptor, which is coupled with $G_s\alpha$ [41]. Therefore, overactivity of the cAMP pathway caused by the *GNAS* mutation could give rise to hyperpigmentation [42]. Mutations in the promoter region of the *GNAS* gene contribute to skin pigmentation in chickens [39]. Based on these findings, the ROH islands on chromosomes 11 and 20 are presumed to have arisen through selection based on phenotypic traits in the YO population.

Comparative analysis of ROH islands and QTLs from the QTL database revealed that a strong association between ROH islands and exterior trait QTLs. Especially among the six exterior trait QTLs that overlapped ROH islands, five QTLs were related to color. Furthermore, the QTL enrichment analysis showed that QTLs associated with skin and comb color traits were significantly enriched in the ROH islands in YO. This evidence could suggest that ROH islands are highly connected with the color of YO, and the long-term selection for a black exterior appearance in the YO breed has critically affected the formation of ROH islands.

Conclusion

This study used ROH analysis of genomic information from the YO population to assess the inbreeding coefficient and identify traces of selection. ROH-based inbreeding coefficients are suitable for the conservation of populations without pedigree information, such as the YO population. We found significant correlations between the ROH-based inbreeding coefficient and other inbreeding coefficients, thus validating the use of ROH analysis as a reliable measure of inbreeding. Importantly, the ROH-based inbreeding coefficient was not affected by allele frequencies, conferring an advantage over other methods, which may be influenced by genetic variation within a population. Furthermore, this study identified the characteristics of ROH islands that were associated with exterior appearance and production traits in YO chickens. These results provide useful information for the establishment of effective conservation strategies of the YO population.

Acknowledgments

Not applicable.

Funding source

This research was funded by the project funding number RS-2021-RD010125(PJ016205) of the Rural Development Administration, South Korea.

References (Vancouver or NLM style)

1. Cho SH, Lee S-S, Seo D, Manjula P, Lee SH, Lee JH. Gender test of unhatched egg in Yeonsan Ogye using PCR technique. Journal of Animal Breeding and Genomics 2019;3. https://doi.org/10.12972/jabng.20190008.

2. Cho SH, Lee S-S, Manjula P, Kim M, Lee SH, Lee JH, et al. Population structure analysis of Yeonsan Ogye using microsatellite markers. Journal of Animal Science and Technology 2020;62:790–800. https://doi.org/10.5187/jast.2020.62.6.790.

3. Jones OR, Wang J. Molecular marker-based pedigrees for animal conservation biologists. Animal Conservation 2010;13:26–34. https://doi.org/10.1111/j.1469-1795.2009.00324.x.

4. Alemu SW, Kadri NK, Harland C, Faux P, Charlier C, Caballero A, et al. An evaluation of inbreeding measures using a whole-genome sequenced cattle pedigree. Heredity 2020;126:410–23. https://doi.org/10.1038/s41437-020-00383-9.

5. Wright S. Coefficients of Inbreeding and Relationship. The American Naturalist 1922;56:330–8. https://doi.org/10.1086/279872.

6. Caballero A, Villanueva B, Druet T. On the estimation of inbreeding depression using different measures of inbreeding from molecular markers. Evolutionary Applications 2020;14:416–28. https://doi.org/10.1111/eva.13126.

7. Peripolli E, Munari DP, Silva MVGB, Lima ALF, Irgang R, Baldi F. Runs of homozygosity: current knowledge and applications in livestock. Animal Genetics 2016;48:255–71. https://doi.org/10.1111/age.12526.

8. Gibson J, Morton NE, Collins A. Extended tracts of homozygosity in outbred human populations. Human Molecular Genetics 2006;15:789–95. https://doi.org/10.1093/hmg/ddi493.

9. Howrigan DP, Simonson MA, Keller MC. Detecting autozygosity through runs of homozygosity: A comparison of three autozygosity detection algorithms. BMC Genomics 2011;12. https://doi.org/10.1186/1471-2164-12-460.

10. Purfield DC, Berry DP, McParland S, Bradley DG. Runs of homozygosity and population history in cattle. BMC Genetics 2012;13. https://doi.org/10.1186/1471-2156-13-70.

11. Ceballos FC, Joshi PK, Clark DW, Ramsay M, Wilson JF. Runs of homozygosity: windows into population history and trait architecture. Nature Reviews Genetics 2018;19:220–34. https://doi.org/10.1038/nrg.2017.109.

12. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. OUP Academic 2015. https://doi.org/10.1186/s13742-015-0047-8.

13. Meyermans R, Gorssen W, Buys N, Janssens S. How to study runs of homozygosity using PLINK? A guide for analyzing medium density SNP data in livestock and pet species. BMC Genomics 2020;21. https://doi.org/10.1186/s12864-020-6463-x.

14. Gorssen W, Meyermans R, Janssens S, Buys N. A publicly available repository of ROH islands reveals signatures of selection in different livestock and pet species. Genetics Selection Evolution 2021;53. https://doi.org/10.1186/s12711-020-00599-7.

15. Kirin M, McQuillan R, Franklin CS, Campbell H, McKeigue PM, Wilson JF. Genomic Runs of Homozygosity Record Population History and Consanguinity. PLOS ONE 2010;5:e13996. https://doi.org/10.1371/journal.pone.0013996.

16. Biscarini F, Cozzi P, Gaspa G, Marras G. detectRUNS: an R Package to Detect Runs of Homozygosity Heterozygosity in Diploid Genomes. (2019). Available online at: https://cran.r-project.org/web/packages/detectRUNS/vignettes/detectRUNS.vignette.html#references (accessed January 1, 2021).

17. McQuillan R, Leutenegger A-L, Abdel-Rahman R, Franklin CS, Pericic M, Barac-Lauc L, et al. Runs of Homozygosity in European Populations. The American Journal of Human Genetics 2008;83:359–72. https://doi.org/10.1016/j.ajhg.2008.08.007.

18. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: A Tool for Genome-wide Complex Trait Analysis. The American Journal of Human Genetics 2011;88:76–82. https://doi.org/10.1016/j.ajhg.2010.11.011.

19. VanRaden PM. Efficient Methods to Compute Genomic Predictions. Journal of Dairy Science 2008;91:4414–23. https://doi.org/10.3168/jds.2007-0980.

20. Purfield DC, McParland S, Wall E, Berry DP. The distribution of runs of homozygosity and selection signatures in six commercial meat sheep breeds. PLOS ONE 2017;12:e0176780. https://doi.org/10.1371/journal.pone.0176780.

21. Kinsella R, Kähäri A, Haider S, Zamora J, Proctor G, Spudich G, et al. Ensembl BioMarts: a hub for data retrieval across taxonomic space. Database 2011; bar030. https://doi.org/10.1093/database/bar030.

22. Fonseca PAS, Suárez-Vega A, Marras G, Cánovas Á. GALLO: An R package for genomic annotation and integration of multiple data sources in livestock for positional candidate loci. GigaScience 2020;9. https://doi.org/10.1093/gigascience/giaa149.

23. Hu Z-L, Park CA, Wu X-L, Reecy JM. Animal QTLdb: an improved database tool for livestock animal QTL/association data dissemination in the post-genome era. Nucleic Acids Research 2012;41:D871–9. https://doi.org/10.1093/nar/gks1150.

24. Kim M, Cho E, Cho S, Choo H, Jin D, Lee JH. A study on the conservation status of Korean native chicken populations. Journal of Animal Breeding and Genomics 2022;6. https://doi.org/10.12972/jabng.20220015.

25. Kranis A, Gheyas AA, Boschiero C, Turner F, Yu L, Smith S, et al. Development of a high density 600K SNP genotyping array for chicken. BMC Genomics 2013;14:59. https://doi.org/10.1186/1471-2164-14-59.

26. Manjula P, Kim M, Cho S, Seo D, Lee JH. High Levels of Genetic Variation in MHC-Linked Microsatellite

Markers from Native Chicken Breeds. Genes 2021;12:240. https://doi.org/10.3390/genes12020240.

27. Cho Y, Kim J-Y, Kim N. Comparative genomics and selection analysis of Yeonsan Ogye black chicken with whole-genome sequencing. Genomics 2022;114:110298. https://doi.org/10.1016/j.ygeno.2022.110298.

28. Zhang J, Nie C, Li X, Ning Z, Chen Y, Jia Y, et al. Genome-Wide Population Genetic Analysis of Commercial, Indigenous, Game, and Wild Chickens Using 600K SNP Microarray Data. Frontiers in Genetics 2020;11. https://doi.org/10.3389/fgene.2020.543294.

29. Zhang Q, Calus MP, Guldbrandtsen B, Lund MS, Sahana G. Estimation of inbreeding using pedigree, 50k SNP chip genotypes and full sequence data in three cattle breeds. BMC Genetics 2015;16. https://doi.org/10.1186/s12863-015-0227-7.

30. Gao C, Du W, Tian K, Wang K, Wang C, Sun G, et al. Analysis of Conservation Priorities and Runs of Homozygosity Patterns for Chinese Indigenous Chicken Breeds. Animals 2023;13:599. https://doi.org/10.3390/ani13040599.

31. Qanbari S, Gianola D, Hayes B, Schenkel F, Miller S, Moore S, et al. Application of site and haplotype-frequency based approaches for detecting selection signatures in cattle. BMC Genomics 2011;12. https://doi.org/10.1186/1471-2164-12-318.

32. Cendron F, Perini F, Mastrangelo S, Tolone M, Criscione A, Bordonaro S, et al. Genome-Wide SNP Analysis Reveals the Population Structure and the Conservation Status of 23 Italian Chicken Breeds. Animals 2020;10:1441. https://doi.org/10.3390/ani10081441.

33. Strillacci MG, Vega-Murillo VE, Román-Ponce SI, López FJR, Cozzi MC, Gorla E, et al. Looking at genetic structure and selection signatures of the Mexican chicken population using single nucleotide polymorphism markers. Poultry Science 2018;97:791–802. https://doi.org/10.3382/ps/pex374.

34. Elferink MG, Megens H-J, Vereijken A, Hu X, Crooijmans RPMA, Groenen MAM. Signatures of Selection in the Genomes of Commercial and Non-Commercial Chicken Breeds. PLOS ONE 2012;7:e32720. https://doi.org/10.1371/journal.pone.0032720.

35. Sun M, Sui Y, Li L, Su W, Hao F, Zhu Q, et al. Anoctamin 1 Calcium-Activated Chloride Channel Downregulates Estrogen Production in Mouse Ovarian Granulosa Cells. Endocrinology 2014;155:2787–96. https://doi.org/10.1210/en.2013-2155.

36. Schwochow D, Bornelöv S, Jiang T, Li J, Gourichon D, Bed'Hom B, et al. The feather pattern autosomal barring in chicken is strongly associated with segregation at the MC1R locus. Pigment Cell & Melanoma Research 2021;34:1015–28. https://doi.org/10.1111/pcmr.12975.

37. Cha J, Jin D, Kim J-H, Kim S-C, Lim JA, Chai H-H, et al. Genome-wide association study revealed the genomic regions associated with skin pigmentation in an Ogye x White Leghorn F2 chicken population. Poultry Science 2023;102:102720. https://doi.org/10.1016/j.psj.2023.102720.

38. Luo C, Qu H, Wang J, Wang Y, Ma J, Li C, et al. Genetic parameters and genome-wide association study of hyperpigmentation of the visceral peritoneum in chickens. BMC Genomics 2013;14. https://doi.org/10.1186/1471-2164-14-334.

39. Wang H, Chen M, ou L Zhang L, Zhang C, Chen X, Zhang X. GNAS gene promoter mutation in chicken and the correlation with skin color traits. Acta Veterinaria et Zootechnica Sinica 2016;47:2354-2361

40. Weinstein LS, Liu J, Sakamoto A, Xie T, Chen M. Minireview: GNAS: Normal and Abnormal Functions. Endocrinology 2004;5459–5464. https://doi.org/10.1210/en.2004-0865.

41. Cui Y, Miao Y, Cao L, Guo L, Cui Y, Yan C, et al. Activation of melanocortin-1 receptor signaling in melanoma cells impairs T cell infiltration to dampen antitumor immunity. Nature Communications 2023;14. https://doi.org/10.1038/s41467-023-41101-3.

42. Speeckaert R, Van Gele M, Speeckaert MM, Lambert J, van Geel N. The biology of hyperpigmentation syndromes. Pigment Cell & Melanoma Research 2014;27:512–24. https://doi.org/10.1111/pcmr.12235.

Tables and Figures

Table 1. Parameters for ROH detection using PLINK v1.9.

Parameter	**PLINK1.9 command**	**Value**
Size of sliding window (Number of SNPs)	--homozyg-window-snp	39
Minimum number of heterozygotes within the sliding window	--homozyg-window-het	1
Minimum number of missing SNPs within the sliding window	--homozyg-window-missing	3
Minimum number of SNPs within ROH	--homozyg-snp	65
Maximum gap (kb) between SNPs within ROH	--homozyg-gap	80
Minimum length (kb) of ROH	--homozyg-kb	500

ROH, runs of homozygosity; SNP, single nucleotide polymorphism.

Table 2. Identified average total ROH length (S_{ROH}), average ROH length (L_{ROH}), and average number of ROHs in the YO population.

S_{ROH} mean ± SD(Mb)	L_{ROH} mean ± SD(Mb)	Number of ROHs mean ± SD
165.649 ± 32.364	1.542 ± 0.274	107.6 ± 15.828

ROH, runs of homozygosity; YO, Yeonsan Ogye; SD, standard deviation.

Table 3. Calculated inbreeding coefficients for the YO population.

Population	F_{HOM} mean ± SD	F_{GRM} mean ± SD	F_{UNI} mean ± SD	F_{ROH} mean ± SD
YO	-0.0089 ± 0.0854	-0.1229 ± 0.0915	-0.0087 ± 0.0448	0.1785 ± 0.0349

ROH, runs of homozygosity; YO, Yeonsan Ogye; F_{HOM}, inbreeding coefficient based on homozygous SNPs; F_{GRM}, inbreeding coefficient based on the diagonal elements of the genomic relationship matrix; F_{UNI}, inbreeding coefficient based on uniting gametes; F_{ROH}, inbreeding coefficient based on ROH; SD, standard deviation.

Table 4. Information of ROH island regions and annotated genes in YO population.

Chr	No. SNPs	Physical position (bp)	Length (bp)	Gene Symbol	No. Genes
3	265	32,927,377 - 33,454,824	527,447	-	-
4	240	14,268,782 - 14,785,480	516,698	*GAB3*, *SMARCA1*	2
	443	64,170,233 - 65,011,778	841,545	*SGCZ*, *DLC1*, *TRMT9B*, *LONRF1*, *PAICS*, *PPAT*, *AASDH*, *CRACD*, *CEP135*, *EXOC1*, *EXOC1L*	11
5	291	2,135,723 - 3,492,810	1,357,087	*NAV2*, *PRMT3*, *LEUTX*, *SLC6A5*, *NELL1*, *ANO5*, *SLC17A6*, *FANCF*, *GAS2*, *SVIP*, *ANO3*	11
	289	28,913,241 - 29,409,863	496,622	*RAD51B*, *TMEM229B*, *PLEKHH1*, *PIGH*, *ARG2*, *VTI1B*, *ZFYVE26*, *PLEK2*, *EIF2S1*, *ATP6V1D*, *MPP5*, *GPHN*	12
	198	29,489,091 - 29,849,491	360,400	*GPHN*, *BMF*, *SRP14*, *EIF2AK4*, *GPR176*, *FSIP1*, *THBS1*	7
	231	30,415,825 - 30,786,010	370,185	*RYR3*, *FMN1*, *GREM1*	3
	57	30,886,135 - 31,013,813	127,678	*RASGRP1*, *FAM98B*, *SPRED1*	3
	36	31,045,059 - 31,144,065	99,006	-	-
	145	48,848,363 - 49,162,160	313,797	*DLK1*, *BEGAIN*, *WDR25*	3
	2	49,320,624 - 49,326,092	5,468	-	-
11	816	18,792,907 - 20,207,704	1,414,797	*FANCA*, *SPIRE2*, *TCF25*, *MC1R*, *TUBB3*, *DEF8*, *DBNDD1*, *GAS8*, *URAH*, *CDH1*, *TANGO6*, *HAS3*, *CHTF8*, *UTP4*, *SNTB2*, *PDF*, *NIP7*, *TMED6*, *TERF2*, *CYB5B*, *NFAT5*, *NOB1*, *WWP2*, *PSMD7*, *ZFHX3*, *DHX38*, *DHODH*, *IST1*, *ZNF821*, *ATXN1L*, *AP1G1*, *PHLPP2*, *TAT*, *TERF2IP*, *KARS*, *ADAT1*, *GABARAPL2*, *CHST6*, *TMEM231*	39
14	449	76,321 - 681,755	605,434	*PDXDC1*, *NTAN1*, *RRN3*, *RSL1D1*, *GSPT2*, *SNX29*, *CPPED1*	7
	191	6,285,034 - 6,526,791	241,757	*GNG13*, *CHTF18*, *RPUSD1*, *MSLN*, *NARFL*, *HAGHL*, *HAGH*, *FAHD1*, *MEIOB*, *HS3ST6*, *MSRB1*	11
20	672	10,862,696 - 11,770,350	907,654	*PRELID3B*, *TUBB1*, *CTSZ*, *NELFCD*, *GNAS*, *NPEPL1*, *STX16*, *APCDD1L*, *VAPB*, *RAB22A*, *C20orf85*, *PMEPA1*, *PCK1*, *RBM38*, *RAE1*, *SPO11*, *BMP7*	17
	2	12,015,574 - 12,015,807	233	*CSTF1*	1
	1,372	12,028,923 - 14,098,905	2,069,982	*FAM210B*, *MC3R*, *CBLN3*, *DOK5*, *PFDN4*, *BCAS1*, *TSHZ2*, *ZFP64*, *SALL4*, *ATP9A*, *NFATC2*, *KCNG1*, *MOCS3*, *DPM1*, *ADNP*, *PARD6B*, *RIPOR3*, *PTPN1*, *CEBPB*, *TMEM189*, *UBE2V1*, *SNAI1*, *RNF114*, *SPATA2*, *SLC9A8*, *B4GALT5*	26

ROH, runs of homozygosity; YO, Yeonsan Ogye; Chr, chromosome; SNP, single nucleotide polymorphism.

Table 5. Significant QTLs in ROH islands in YO population.

QTL	**No. QTLs**	**No. QTLs in database**[1]	***p*-value**	**FDR adjusted *p*-value**	**Trait**[2]
Skin color	90	223	3.88E-101	1.16E-99	Exterior
Comb color	76	138	3.72E-98	5.58E-97	Exterior
Egg number	48	476	2.01E-22	2.01E-21	Production
Visceral peritoneum pigmentation	4	17	2.22E-04	1.67E-03	Exterior
Age at sexual maturity	2	5	3.26E-03	1.96E-02	Production

1) Chicken QTL data in Animal QTLdb

2) Chicken QTLdb trait class

QTL, quantitative trait loci; ROH, runs of homozygosity; YO, Yeonsan Ogye; FDR, false discovery rate.

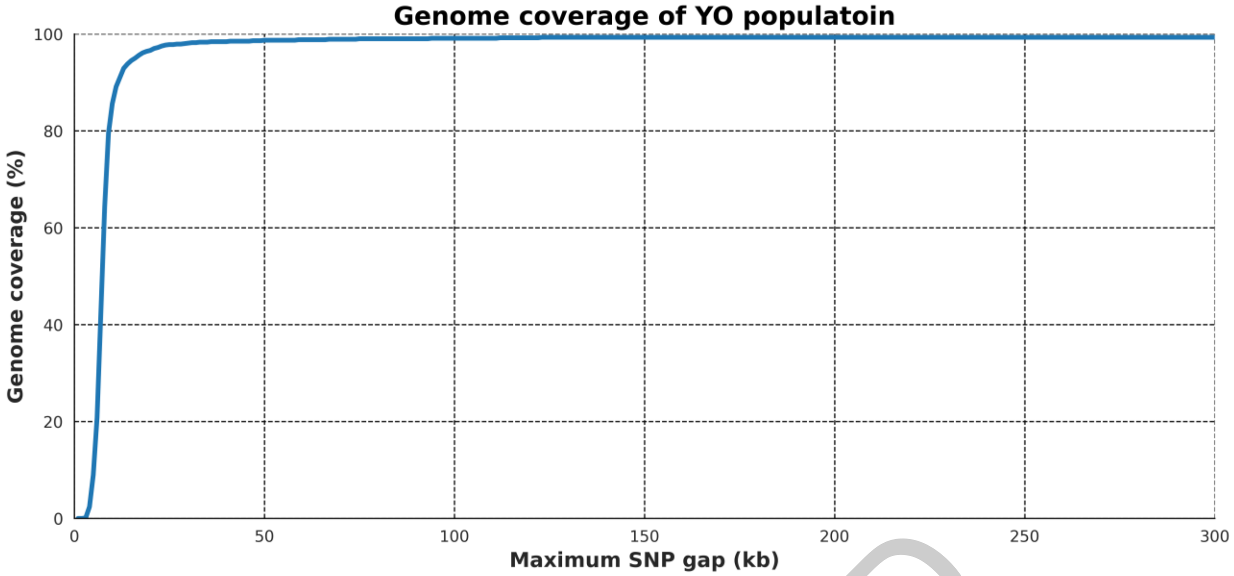

Fig. 1. Changes in genome coverage (%) according to the maximum SNP gap (kb) used in ROH analysis with the 600K SNP chip. The genome coverage is notably increased from 4 kb/SNP and reached 99% coverage at 80 kb/SNP. SNP, single nucleotide polymorphism; YO, Yeonsan Ogye.

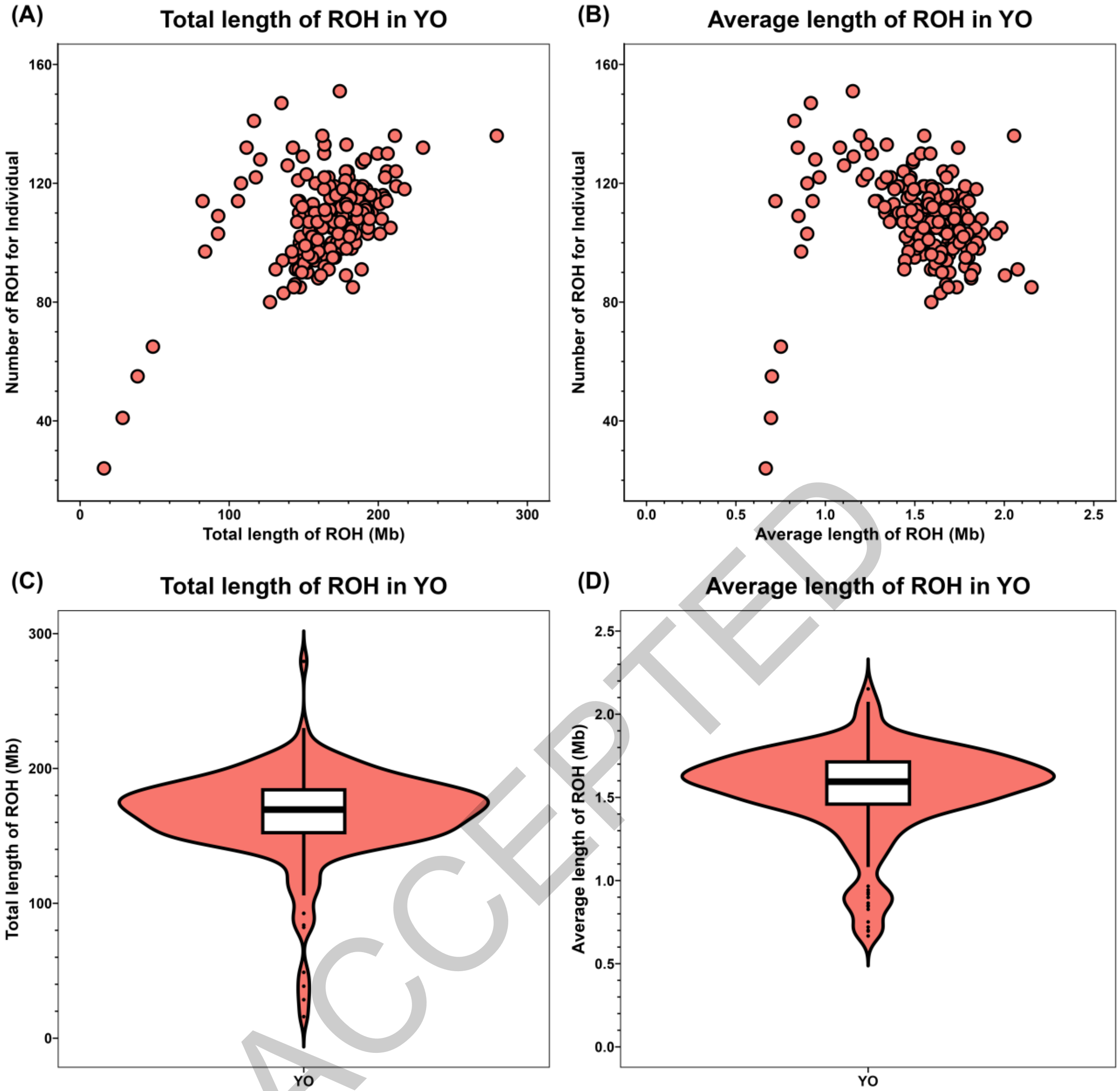

Fig. 2. Individual and population ROH values for the YO population. (A) The number of ROHs and total length of ROHs (Mb) in YO individuals, (B) the number of ROHs and average length of ROHs (Mb) in YO individuals, (C) total length of the average ROH in the YO population, and (D) average ROH length in the YO population. ROH, runs of homozygosity; YO, Yeonsan Ogye.

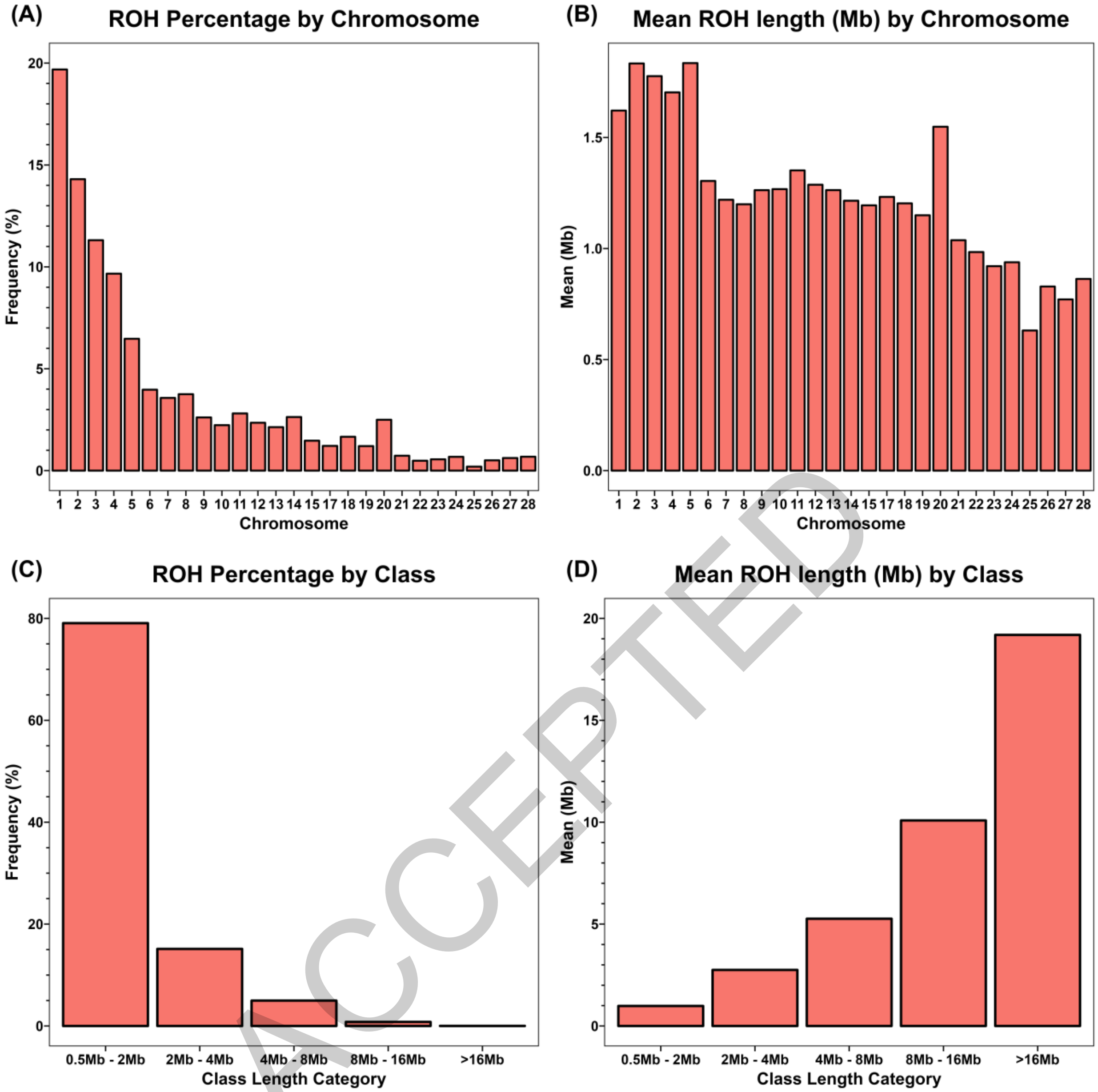

Fig. 3. Distribution of ROHs according to chromosome and ROH length category in the YO population. (A) ROH percentages in individual chromosomes, (B) average ROH length according to chromosome, (C) ROH percentages in different length categories, and (D) average ROH length (Mb) according to length categories. ROH, runs of homozygosity; YO, Yeonsan Ogye.

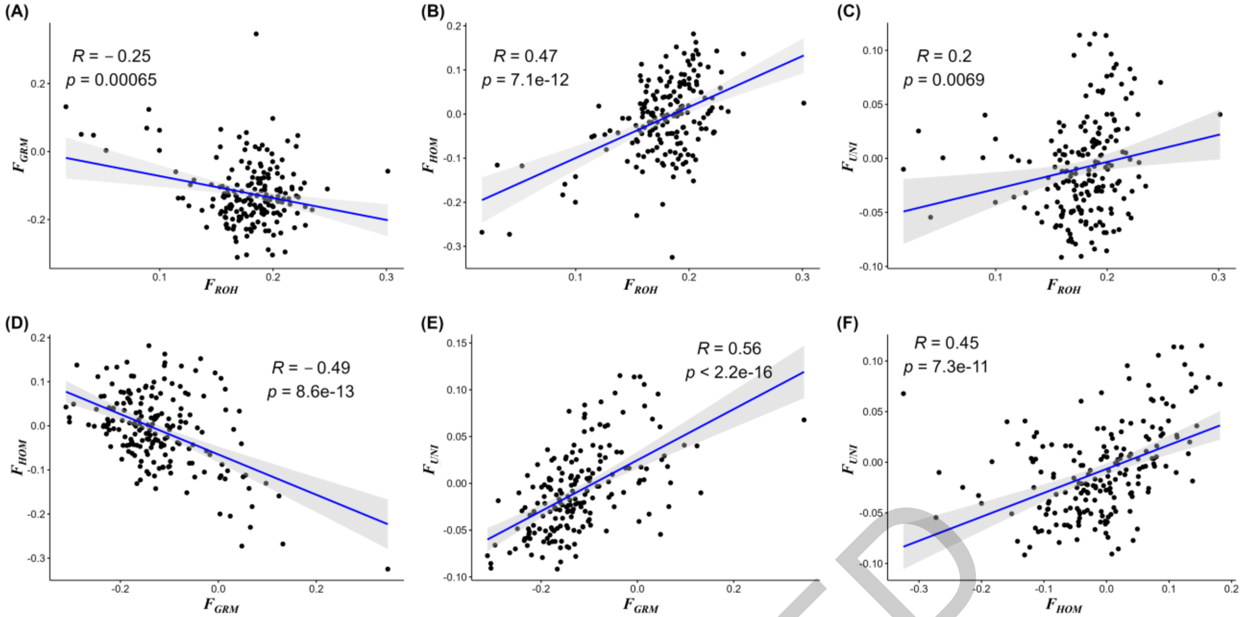

Fig. 4. Comparison of correlations among inbreeding coefficients for the YO population. **F_{HOM}** is the inbreeding coefficient based on homozygous SNPs. **F_{GRM}** is the inbreeding coefficient based on diagonal elements of the genomic relationship matrix. **F_{UNI}** is the inbreeding coefficient based on gamete union. **F_{ROH}** is the inbreeding coefficient based on runs of homozygosity. A and D show negative correlations between **F_{GRM}** with **F_{ROH}** and **F_{HOM}**, respectively. YO, Yeonsan Ogye; SNP, single nucleotide polymorphism.

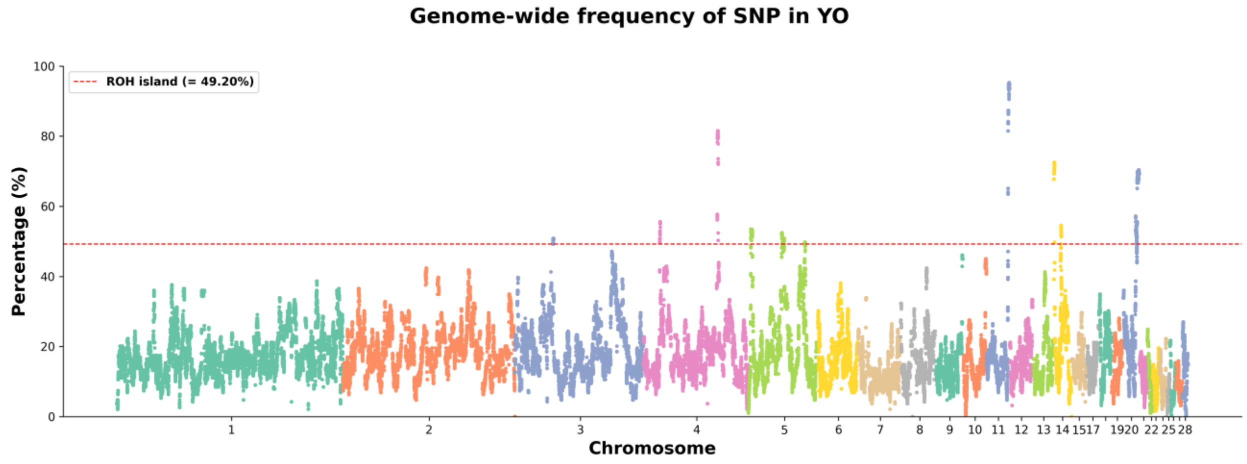

Fig. 5. Genome-wide frequency of SNPs in ROHs. The red dashed line is the threshold of an ROH island (49.20%), the frequency of SNPs in ROHs in the top 1%. SNP, single nucleotide polymorphism; ROH, runs of homozygosity.

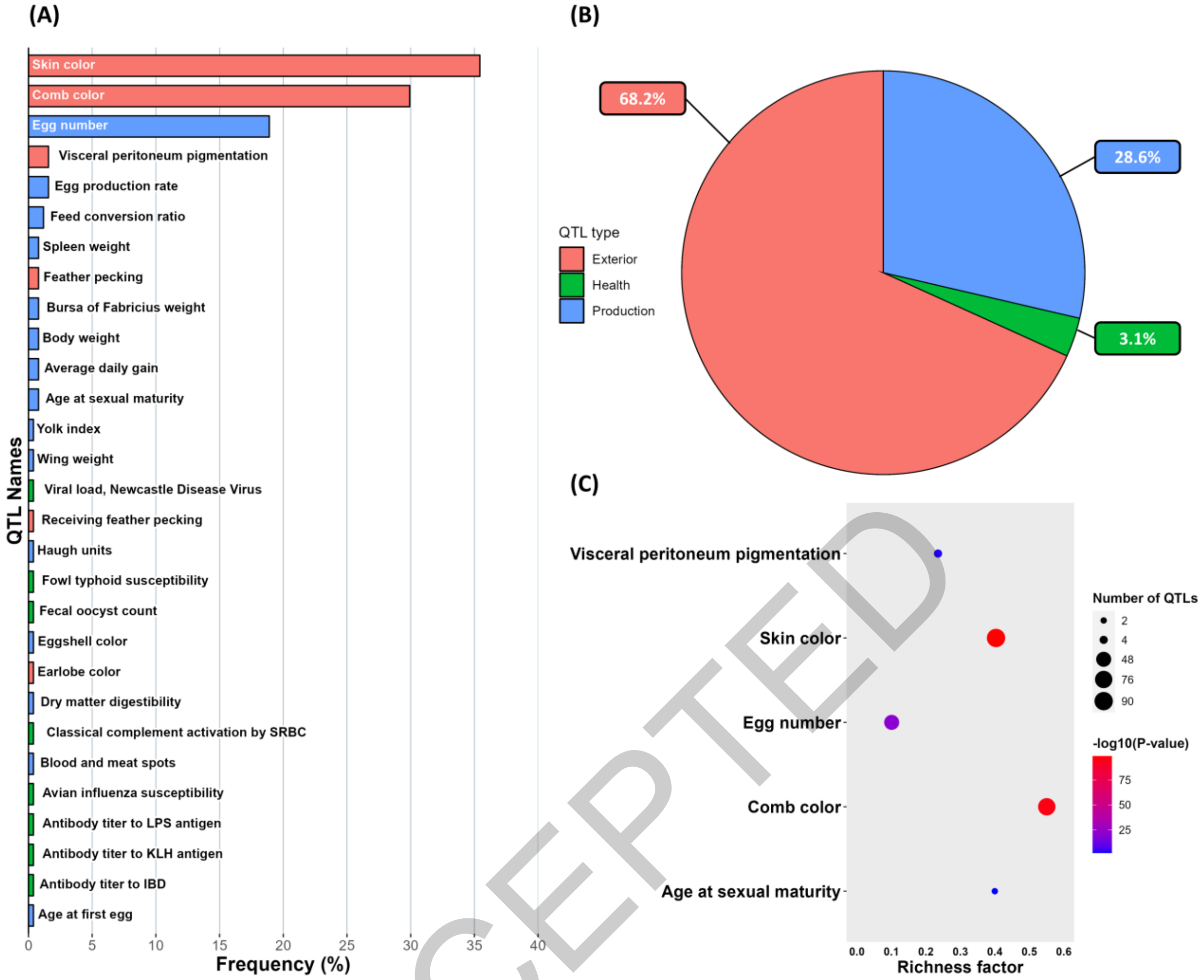

Fig. 6. QTLs identified in ROH islands and QTL enrichment analysis results. (A) Ratios of related QTL names in ROH islands, (B) ratios of related QTL types in ROH islands, and (C) enrichment analysis results for QTLs with a false discovery rate < 0.05 in ROH islands. QTL, quantitative trait loci; ROH, runs of homozygosity.