**Runs of Homozygosity in Modern Chicken Revealed by Sequence Data**

Reza Talebi,*1,1 Tomasz Szmatoła,*6,6 Gábor Mészáros,** and Saber Qanbari††,2

*Department of Animal Sciences, Faculty of Agriculture, Bu-Ali Sina University, Hamedan, Iran, 1Department of Systems and Synthetic Biology, Agricultural Biotechnology Research Institute of Iran, Agricultural Research, Education and Extension Organization (AREEO), Karaj, Iran, 6Centre of Experimental and Innovative Medicine, University of Agriculture in Kraków, Al. Mickiewicza 24/28, 30-059 Kraków, Poland, 6Department of Animal Molecular Biology, National Research Institute of Animal Production, Krakowska 1, 32-083 Balice, Poland, **Division of Livestock Sciences (NUW), University of Natural Resources and Life Sciences, Gregor-Mendel Strasse 33, 1180 Vienna, Austria, and ††Leibniz Institute for Farm Animal Biology (FBN), Institute of Genetics and Biometry, 18196 Dummerstorf, Germany

ORCID IDs: 0000-0002-5555-8453 (R.T.); 0000-0003-1588-4198 (T.S.); 0000-0002-5937-0060 (G.M.); 0000-0001-6866-2332 (S.Q.)

**ABSTRACT** Runs of homozygosity (ROH) are chromosomal stretches that in a diploid genome appear in a homozygous state and display identical alleles at multiple contiguous loci. This study aimed to systematically compare the genomic distribution of the ROH islands among five populations of wild vs. commercial chickens of both layer and broiler type. To this end, we analyzed whole genome sequences of 115 birds including white layer (WL, n = 25), brown layer (BL, n = 25), broiler line A (BRA, n = 20), broiler line B (BRB, n = 20) and Red Junglefowl (RJF, n = 25). The ROH segments varied in size markedly among populations, ranging from 0.3 to 21.83 Mb reflecting their past genealogy. White layers contained the largest portion of the genome in homozygous state with an average ROH length of 432.1 Mb (±18.7) per bird, despite carrying it in short segments (0.3-1 Mb). Population-wise inbreeding measures based on Wright’s (Fst) and genomic (FROH) metrics revealed highly inbred genome of layer lines relative to the broilers and Red Junglefowl. We further revealed the ROH islands, among commercial lines overlapped with QTL related to limb development (GREM1, MEOX2), body weight (Meis2a.1, uc_338), eggshell color (GLCCI1, ICA1, UMAD1), antibody response to Newcastle virus (ROBO2), and feather pecking. Comparison of ROH landscape in sequencing resolution demonstrated that a sizable portion of genome of commercial lines segregates in homozygote state, reflecting many generations of assortative mating and intensive selection in their recent history. In contrary, wild birds carry shorter ROH segments, likely suggestive of older evolutionary events.

**KEYWORDS** Inbreeding ROH islands White layer Brown layer Red Junglefowl

Individuals with a recent common ancestor share sizable part of their genomes identical-by-descent (IBD) state. These individuals transmit IBD segments to their progeny, and create runs of homozygosity (ROH) across the genome (McQuillan et al. 2008). ROH were first identified by Broman and Weber in the human genome, whereas Gibson et al. acknowledged their importance for inbreeding calculations (Broman and Weber 1999; Gibson et al. 2006). McQuillan et al. (2008) defined the genomic inbreeding coefficient based on ROH (FROH), which does not depend on allelic frequencies or sampling procedures (Curik et al. 2014; Signer-Hasler et al. 2017). FROH was later shown to be a suitable measure for describing levels of inbreeding in breeds with missing pedigree information (Burren et al. 2016). Therefore, ROH may provide a more accurate measure of inbreeding levels, compared to the pedigree based measurement (Signer-Hasler et al. 2017). While longer haplotypes originate from recent common ancestors, shorter haplotypes inherit from distant
ones (background relatedness), and the sum of all these segments are suggested to be an accurate estimation of the inbreeding level of an individual (Ceballos et al. 2018a,b; Yoshida et al. 2020).

The identification and characterization of ROH islands can further provide insights into population history, structure and demographics over time (Peripolli et al. 2017; Signer-Hasler et al. 2019). ROH islands are widely observed in populations and can be used as a useful tool to identify the phenomenon called “selective sweeps”, genome regions undergone selection pressure (Szmato et al. 2019). ROH islands have been used to localize selection signatures in the genome of farm animals including sheep (Purfield et al. 2017; Mastrangelo et al. 2017; Mastrangelo et al. 2018; Luigi-Sierra et al. 2019; Signer-Hasler et al. 2019), goat (Bertolini et al. 2018; Onzima et al. 2018), horse (Metzger et al. 2015; Grilz-Seger et al. 2018; Ablondi et al. 2019; Grilz-Seger et al. 2019), cattle (Goszczynski et al. 2018; Peripolli et al. 2018; Nandolo et al. 2018; Szmato et al. 2019; Peripolli et al. 2020), pig (Wang et al. 2019; Xu et al. 2019; Xie et al. 2019; Gorssen et al. 2020) and in chicken (Marchesi et al. 2018; Almeida et al. 2019; Strillacci et al. 2018; Zhang et al. 2018; Zhang et al. 2020).

Chicken is a vital livestock for the world food security by producing massive quantities of meat and egg. Chicken also provide an excellent model to investigate the genetics of adaptation, as (1) it provides the landscape of homozygosity in sequence resolution, highlighting several candidate genes co-localized with quantitative trait loci (QTL). However, we were unable to pinpoint the stand-alone ROH islands overlapping among groups of birds as a significant fraction of the genome in modern chicken was carried in homozygous state.

### MATERIAL AND METHODS

#### Chicken populations and data preparation

For the purpose of this study we used SNP panel from birds sequenced individually in Qanbari et al. (2019). Briefly, it comprises medium coverage (~10 folds) sequence of the entire genome of 115 chicken samples including white layer (WL, n = 25), brown layer (BL, n = 25), broiler line A (BRA, n = 20), broiler line B (BRB, n = 20) and Red Junglefowl (RFJ, n = 25). For further detail of the SNP calling process we refer to the original study. Markers with minimum 1 minor allele were retained for further analysis. To ensure data quality we excluded the SNPs with genotyping rate < 0.1 and significant deviations from the Hardy–Weinberg equilibrium ($P_{\text{HWE}} < 10^{-6}$).

#### Measuring diversity metrics

The diversity indicators including observed heterozygosity (Ho), and expected heterozygosity (He), were calculated using PLINK v1.9 package (Purcell et al. 2007). Polymorphic marker ratio (PN) that refers to the proportion of polymorphic loci in the target population was also estimated by averaging the proportion of non-missing SNPs per individual for each population.

#### Measuring runs of homozygosity

Analysis of ROH was performed using PLINK v1.9 package (Purcell et al. 2007). The -`roh` module makes ROH calls using a predefined sliding window that scans along an individual’s SNP panel to detect homozygous stretches (Howrigan et al. 2011). The parameters and thresholds to define an ROH were set according to Almeida et al.

![Figure 1 A schematic representation of ROH profile in chicken genome. The profile is given by the total number of homozygous segments and total segment size (Mb). Populations are coded as RJF = red jungle fowl, BL = Brown layer, WL = White layer, BRA = Broiler line A and BRB = Broiler line B.](image-url)
(2019) and Ceballos et al. (2018b). Accordingly, we applied (i) sliding windows of size 50 SNPs across the genome; (ii) the proportion of homozygous overlapping windows set to 0.05; (iii) the minimum number of consecutive SNPs included in a ROH set to 50; (iv) the minimum length of a ROH set to 300 kb; (v) the maximum gap size between consecutive homozygous SNPs set to 1000 kb; (vi) required minimum density to consider a ROH was 1 SNP in 50 Kb; and (vii) a maximum of five SNPs with missing genotypes and up to three heterozygous.

Measuring inbreeding

Two measures of inbreeding coefficient were calculated with PLINK v1.9 (Purcell et al. 2007) for each population.

1. Wright’s inbreeding coefficient (Fis) was calculated by comparing the difference between observed and expected autosomal homozygous genotypes for each sample as follows:

   \[ F_{is} = \frac{\text{Number of observed homozygous loci} - \text{Number of expected homozygous loci}}{\text{Number of nonmissing loci} - \text{Number of expected homozygous loci}} \]

2. Genomic inbreeding coefficients based on ROH (FROH) was estimated for each bird according to McQuillan et al. (2008). Accordingly, FROH was defined as:

   \[ F_{ROH} = \frac{L_{ROH}}{L_{total}} \]

   where \( L_{ROH} \) is the total size of ROH in the genome of each bird.

Table 2 Summary statistics of the runs of homozygosity (ROH) based on length classes

| Class of ROH | Chicken population |
|--------------|--------------------|
|              | Layers             | Broilers          | Wild birds       |
|              | WL (n= 25)         | BL (n= 25)        | BRA (n= 20)      | BRB (n= 20) | RJF (n= 25) |
| 0.3-1Mb      | 9458 (73.9%)       | 4631 (58.8%)      | 2352 (70%)       | 2820 (71.1%) | 1083 (51.6%) |
| 1-2Mb        | 2619 (20.4%)       | 1866 (23.7%)      | 593 (17.6%)      | 718 (18.1%)  | 504 (24.0%)  |
| 2-4Mb        | 685 (5.4%)         | 1001 (12.7%)      | 327 (9.7%)       | 350 (8.8%)   | 330 (15.7%)  |
| 4-8Mb        | 42 (0.3%)          | 335 (4.3%)        | 79 (2.3%)        | 77 (1.9%)    | 153 (7.3%)   |
| 8-10Mb       | 0                 | 23 (0.3%)         | 6 (0.2%)         | 4 (0.1%)     | 17 (0.8%)    |
| 10-16Mb      | 0                 | 15 (0.2%)         | 5 (0.1%)         | 0            | 12 (0.6%)    |
| >16Mb        | 0                 | 1 (0.01%)         | 0               | 0            | 0 (0.05%)    |
| Total N. a   | 12807             | 7872              | 3362            | 3969         | 2100         |
| Mean N. b    | 512.28            | 314.88            | 168.1           | 198.45       | 84           |
| Total L. (Mb) c | 10802           | 10189             | 3526            | 3923         | 3365         |
| Max. (Mb)   d | 7.04             | 21.83             | 13.85           | 9.85         | 17.92        |
| Min. (Mb)   d | 0.30             | 0.30              | 0.30            | 0.30         | 0.30         |

a Total N.: Total number of ROH detected in the population. 

b Mean N.: average number of ROH per individual calculated as the Total N. divided by the number of individuals.

c Total L.: Total length of ROH detected in the population.

d Max., Min.: respectively revealed maximum and minimum length of ROH segments among the total number of ROH.

Distribution of runs of homozygosity

The total number of ROH per chromosome, average length of ROH per chromosome (Mb), and the percentage of chromosomes covered by ROH were estimated in PLINK (–homozyg option). Using an in-house R script the identified ROHs were classified into seven length classes of 0.3–1, 1–2, 2–4, 4–8, 8–10, 10–16 and >16 Mb, respectively.

Detection of autozygosity islands

To investigate genomic regions that were associated with occurrences of ROH within each breed, the fraction of SNPs in ROH was estimated based on the frequency of a SNP in them across individuals. ‘ROH islands’ was identified as a region of adjacent SNPs with an ROH frequency per SNP above the threshold of 1%. These regions were then overlapped with Chicken quantitative trait loci (QTL) database (release 40, 2019, https://www.animalgenome.org/cgi-bin/QTLdb/GG/index).

Annotation of genetic variants

SnpEff (v.3.4) (Cingolani et al. 2012) was used to predict the functional effects of the identified mutations based on reference genome

Figure 2 Average percentage of chromosome coverage by runs of homozygosity of minimum length of 0.3 Mb. Populations are coded as WL = White layer, BL = Brown layer, BRA = Broiler line A, BRB = Broiler line B, RJF = Red Junglefowl.
Gallus_gallus-5.0.86. The annotated functional categories included: 5 kb up- and down-stream of a gene, intergenic, missense, synonymous, intronic, 3′ untranslated regions, 5′ untranslated regions, stop gain and stop loss. Variants in the up- and down-stream regions and in the 3′UTR, 5′UTR regions were merged into the single categories. The bedtools’ (v.2.29.2) (Quinlan and Hall 2010) intersect function was used to determine overlap regions between layers (WL and BL), broilers (BRA and BRB), and commercials (WL, BL, BRA and BRB) within ROHislands.

Gene-ontology (GO) enrichment analysis
Gene-ontology (GO) enrichment experiment was conducted using the ClueGO plugin v2.1.7 (http://www.ici.upmc.fr/cluego/) (Bindea et al. 2009). We further used DAVID v6.8 tool (Huang et al. 2009a; Huang et al. 2009b) to focus more on significantly enriched Gene GO terms. Only GO/pathway terms with significant enrichment (corrected P-value, 0.05 of Benjamini-Hochberg) were included in the analysis.

Data availability
Data used in this study has been originally published by Qanbari et al. (2019) and is deposited in the European Nucleotide Archive (ENA) with the accession number: PRJEB30270. Supplemental material available at figshare: https://doi.org/10.25387/g3.13107176.

RESULTS AND DISCUSSION
Genotyping, quality control and genetic diversity
The total genotyping rate ranged from 0.97 to 0.99 and the number of polymorphic sites within breeds ranged from 6′100′640 to 14′734′764 (Table 1). RJF was the most diverse population with over 14 million polymorphic sites retained for the final analyses (Table 1). For each breed, genetic diversity was measured using heterozygosity and Wright's inbreeding coefficient (Fis). RJF breed, revealed the highest polymorphic marker ratio (99.85%) and despite the lowest observed (HO = 0.240 ± 0.032) and expected heterozygosity (HE = 0.249 ± 0.000) we measured a low level of inbreeding (Fis = 0.037) (Table 1). On the contrary, layers represent lowest polymorphic marker ratio and less observed heterozygosity than expected which is attributed to the small number of founders and many generations of mating within closed lines of limited population size, but also partly due to the effect of linked selection (Qanbari et al. 2019). Accordingly, Fis was markedly high in layers in line with a previous report of genetic diversity and inbreeding in commercial white egg layer line (Bortoluzzi et al. 2018). A recent study based on genotypes of a high-density SNP array revealed lower inbreeding (Fis = 0.037 to 0.237) in Chinese indigenous chicken breeds than both Europeans local (Fis = 0.254 to 0.404) and commercial brown egg layers (Fis = 0.144), suggestive of the richer genetic diversity available in indigenous populations (Chen et al. 2019).

Runs of homozygosity
As schematically shown in Figure 1, distinctive clustering patterns reflected the genetic diversity of five chicken populations. The ROH

Table 3 Average genomic inbreeding coefficient (FROH) for different length categories of ROH across five chicken populations

| Length category (Mb) | Layers | Broilers | Wild birds |
|----------------------|--------|----------|------------|
|                      | WL (n= 25) | BL (n= 25) | BRA (n= 20) | BRB (n= 20) | RJF (n= 25) |
| FROH (0.3-1)         | 0.228 | 0.113 | 0.068 | 0.082 | 0.026 |
| FROH (1-2)           | 0.152 | 0.113 | 0.044 | 0.054 | 0.031 |
| FROH (2-4)           | 0.075 | 0.119 | 0.048 | 0.051 | 0.039 |
| FROH (4-8)           | 0.009 | 0.075 | 0.022 | 0.022 | 0.034 |
| FROH (8-10)          | 0.000 | 0.009 | 0.003 | 0.002 | 0.007 |
| FROH (10-16)         | 0.000 | 0.008 | 0.003 | 0.000 | 0.007 |
| FROH (>16)           | 0.000 | 0.001 | 0.000 | 0.000 | 0.001 |
| Total FROH (<0.3)    | 0.46  | 0.44  | 0.19  | 0.21  | 0.14  |
| SD                   | 0.02  | 0.04  | 0.05  | 0.02  | 0.10  |

FROH ≥0.3Mb = overall genomic inbreeding (average of individuals) at ROH threshold of 0.3 Mb, SD = standard deviation of the total FROH calculated based upon individual FROH, which has been shown in File S1.

Figure 3 Inbreeding within chicken populations. The Pearson correlation is presented between two measures of molecular inbreeding metrics. Populations are coded as WL = White layer, BL = Brown layer, BRA = Broiler line A, BRB = Broiler line B, RJF = Red Junglefowl.
Profile (clusters, Figure 1) in wild birds displayed a more variable pattern than domestic birds (Figure 1). Commercial lines exhibited similar average cumulative size as well as average ROH number, indicating extremely homogenous genomes with low genetic diversity (Table 1).

Descriptive summary of ROH number and length classes is presented in Table 2. As expected, White layers carry more ROH with an average number of 512.28 ± 18.76 per bird (ranging from 547 to 471). By contrast, RJFs encompass least ROH average in 84.00 ± 61.02 (ranging from 251 to 6) per bird (see Table 2 and File S1). The average size of ROH tracts observed in Red Junglefowls is comparable to the African indigenous chicken populations from Rwanda and Uganda both genotyped with the chicken Affymetrix 600K Axiom Array (Elbeltagy et al. 2019). Our measure of ROH in broiler populations confirms the dimension reported in previous studies (Marchesi et al. 2018; Almeida et al. 2019). Layers, especially WL carry the lengthy ROH tracts with 432.1 Mb (±18.7 Mb; max. 477.1 Mb; min. 390.8Mb) per bird, whereas the corresponding value was 134.6 Mb (±102.2 Mb; max. 345.4 Mb; min. 3.3 Mb) in wild birds (see File S1). However, the collective length of ROH in WL involved plenty of shorter segments (ROH 1-10Mb) compared to other breeds (Table 2). The largest proportion of the long tracts (ROH >100Mb) was found in BL, reflecting recent inbreeding (Table 2). RJF also revealed several long segments of ROH (Table 2) consistent with the observations in indigenous breeds from Netherlands (Bortoluzzi et al. 2018), Mexico (Strillacci et al. 2018), and China (Zhang et al. 2018). These long ROH tracts likely represent the severe bottlenecks, as traditional populations have experienced a drastic reduction in the effective population size in the recent history and genetic experiments show a slow recovery, despite the potential increase in recent population size (Charlesworth 2009).

The effect of chicken genome heterogeneity in forming ROH segments was also investigated through comparison of ROH on macro- (GGA1 to 5), intermediate- (GGA6 to 10) and microchromosomes (GGA11 to 28) (Figure 2 and Supplementary Table S1). The GGA16 chromosome did not present any ROH segment due to the assembly fallacy. We found the longest ROH segments on macrochromosomes in line with the lower rate of nucleotide diversity and recombination rate (Axelsson et al. 2005; Groenen et al. 2009). Overall, macrochromosomes showed the highest number of common ROH and the longest ROH tracts per individual (Supplementary Table S1). The highest percentage of genome covered by ROH in layers was observed in microchromosomes e.g., in GGA11 (58.21 in WL) and GGA22 (50.12 in BL), (see Figure 2 and Supplementary Table S1), whereas the corresponding chromosomes in broilers and Red Junglefowl were e.g., GGA2 (26.25 in BRB) and GGA5 (18.89 in RJF) (see Figure 2 and Supplementary Table S1).

Genomic inbreeding coefficient estimated From ROH ($F_{ROH}$)

In the absence of pedigree information, numerous studies have documented the usefulness of ROH segments to infer the inbreeding level of an individual (e.g., Purfield et al. 2017; Marchesi et al. 2018; Strillacci et al. 2018; Zhang et al. 2018, among others). We calculated the genomic inbreeding coefficient ($F_{ROH}$) corresponding to the different length classes of ROH segments (Table 3). Consistent with Wright’s inbreeding coefficient, WL was the most inbred population with an $F_{ROH} = 0.46 ± 0.02$ (see Table 3). The genomic inbreeding coefficient ranked in the following order WL > BL > BRB > BRA > RJF for the populations (see Table 3 and File S1), which is consistent with the Wright’s definition of inbreeding as well as the diversity measures in these populations (e.g., Qanbari et al. 2019). Layer lines exhibited markedly greater level of inbreeding in comparison to the broilers and wild birds (Table 3). However, extent of ROH tracts observed in WL and BL is to some extent contradictory. While, a higher proportion of short ROH ($F_{ROH} 0.3-1Mb$) identified in WL

Figure 4 Genome wide distribution of runs of homozygosity (ROH) hotspots. The x-axis represents the SNP genomic coordinate chromosome-wise, and the y-axis shows the proportion of overlapping ROH shared among individuals based upon number in population. Populations are coded as WL = White layer, BL = Brown layer, BRA = Broiler line A, BRB = Broiler line B, RJF = Red Junglefowl.
Table 4 Overlapping region among the ROH islands across commercials (layers, broilers) and Red Junglefowl

| Chicken group | CHR | Start position (bp) | Stop position (bp) | Number of SNPs | Size (bp) | Associated QTL | Associated Gene |
|---------------|-----|---------------------|-------------------|----------------|----------|----------------|----------------|
| Layers        | 2   | 27826712            | 29293102          | 63             | 1466.4   | QTL:24387      | Antibody response to Newcastle virus |
| Broilers      | 1   | 96797489            | 97225788          | 1806           | 428.5    | QTL:24925      | Eggshell color |
|               | 2   | 24570640            | 25113952          | 2198           | 217.3    | GLCCI1, ICA1, UMAD1 |
|               | 2   | 25685016            | 26072852          | 1515           | 387.8    | GLCCI1, ICA1, UMAD1 |
|               | 3   | 133839172           | 134258895         | 994            | 419.7    | QTL:24961      | Eggshell color |
|               | 2   | 143422482           | 143691103         | 1602           | 268.6    | RJFs/Coms, RJFs/BRs, BRs/LRs |
|               | 3   | 84024798            | 84630723          | 1401           | 335.9    | RJFs/Coms, RJFs/BRs, BRs/LRs |
| Red Junglefowl| 1   | 56231495            | 56890000          | 11264          | 658.5    | EGGSHELL strength |
|               | 1   | 57324733            | 59237097          | 29851          | 1912.4   | QTL:57888      | Eggshell strength; Visceral peritoneum pigmentation |
|               | 1   | 59262386            | 60460076          | 24311          | 1197.7   | QTL:57881      | Eggshell strength; Eggshell strength |
|               | 2   | 52312949            | 52315876          | 55             | 2.9      | QTL:57886      | Eggshell strength |
|               | 2   | 52841836            | 52869770          | 24             | 27.9     | QTL:57887      | Eggshell strength |
|               | 2   | 94716992            | 9496946           | 2337           | 252.7    | QTL:57888      | Eggshell strength |
|               | 2   | 95070084            | 9526392           | 676            | 563.3    | QTL:57892      | Eggshell thickness |
|               | 3   | 11705822            | 11714324          | 23             | 8.5      | QTL:57893      | Eggshell thickness |
|               | 3   | 12213570            | 12232316          | 120            | 12.9     | QTL:57894      | Eggshell thickness |
|               | 3   | 86109072            | 86406883          | 3222           | 297.8    | QTL:57895      | Eggshell thickness |
|               | 3   | 89916285            | 90709757          | 7549           | 793.5    | QTL:57896      | Eggshell thickness |
|               | 4   | 18397372            | 18778149          | 923            | 380.8    | QTL:24914      | Average daily gain |
|               | 5   | 41644677            | 42139793          | 3630           | 495.1    | QTL:64552      | Feed intake |
|               | 5   | 42277347            | 42279153          | 7              | 1.8      | QTL:24982      | Age at sexual maturity |
|               | 7   | 274155              | 653362            | 3321           | 379.2    | QTL:24982      | Age at sexual maturity |

*Overlapping ROH islands with the putative selective sweeps identified in Qanbari et al. 2019. RJFs/Coms: Red Junglefowls vs. Commercials; RJFs/LRs: Red Junglefowls vs. Layers; RJFs/BRs: Red Junglefowls vs. Broilers; BRs/LRs: Broilers vs. Layers.*
Table 5 Functional enrichment of gene ontology terms among the identified genes within ROH islands

| Breed        | Biological Processes               | GO term                        | Gene                  |
|--------------|-----------------------------------|-------------------------------|-----------------------|
| Commercials  | limb development                  | GO:0060173                    | GREM1, MEOX2          |
|              | negative regulation of apoptotic process | GO:0043066                  | AVEN, ASNS, GREM1     |
|              | cell-cell signaling               | GO:0007267                    | GREM1, TAC1           |
| Layers       | cell differentiation              | GO:0030154                    | AGR2, AGR3, BZW2      |
|              | negative regulation of cell death | GO:0060548                    | AGR2, AGR3            |
| Broilers     | ureteric bud development          | GO:0001657                    | FMN1, GREM1, ROBO2    |
|              | cell migration involved in sprouting angiogenesis | GO:0002042                | GREM1, SPRED1         |
| Red Junglefowl| oogenesis                        | GO:0048477                    | PTN, WASH1, WEE2      |
|              | amino acid transmembrane transport| GO:0003333                    | CGTL, SLC6A12, SLC6A13|
|              | organic acid:sodium symporter activity | GO:0005343                | SLC13A4, SLC6A12, SLC6A13|

(Table 2 and 3), BL carry a higher proportion of long tracts (F_{ROH} > 16Mb), that might be interpreted as an ancient origin of inbreeding in layers than broilers. In fact, given the recent genealogy of close mating, layers are expected to carry long homozygosity tracts, as is reported in the previous studies based on array genotypes (Bortoluzzi et al. 2018). However, long ROHs tend to break into shorter tracts due to the presence of frequent heterozygous gaps in sequencing resolution (e.g., Qanbari 2020; Szmatała et al. 2020), whereas microarrays would likely miss them.

Figure 3 presents a schematic visualization of Pearson correlation between two measures of molecular inbreeding (F_s and F_{ROH}). A moderate to strong correlation was found between F_s and F_{ROH} within all populations, demonstrating that the extent of a genome under ROH can be used fairly accurately to predict the IBD fraction. The correlation between F_s and F_{ROH} was lowest in WL (r = 0.64, P-value = 0.0029) and BL (r = 0.79, P-value < 0.0001), while the highest correlation was observed in RJFs (r = 0.99, P-value < 0.0001). This trend indicates a positive association of the short ROHs with F_s and likely suggestive of ancestral inbreeding in wild birds. Studies have reported high correlation between F_s and F_{ROH} based on short ROH segments in cattle, suggesting ancestral inbreeding back up to 50 generations ago (Ferenčáková et al. 2013; Mastrangelo et al. 2016; Limper 2018). The high correlation observed between two metrics of molecular inbreeding is consistent with the previous reports on chicken (Bortoluzzi et al. 2018; Marchesi et al. 2018; Almeida et al. 2019), cattle (Ferenčáková et al. 2013; Zhang et al. 2015; Mastrangelo et al. 2016; Peripoli et al. 2018) and sheep (Mastrangelo et al. 2017; Purfield et al. 2017; Mastrangelo et al. 2018) as well as human model (Clark et al. 2019). These results confirm the usefulness of the ROH analysis in monitoring differentiation and inbreeding values for further exploitation in chicken breeding programs in the absence of pedigree records.

ROH islands indicative of selection sweeps

ROH islands might be indicative of genomic regions underwent natural and/ or artificial selection. We sought to identify most homozygote variants within ROH islands as candidates of recent adaptation and focused on outlaying SNPs in top 1% of ROHs in each population. Given the variable polymorphism content, homozygosity threshold to call an ROH island was population-specific. For example, 96% of the white layers were homozygote in top 1% of the ROH islands, whereas the corresponding figure in wild birds was only 40%. Accordingly, we set the SNP homozygosity thresholds as > 96%, 96%, 60%, 65%, and 40% in WL, BL, BRA, BRB, and RJF, respectively (Figure 4).

We identified 26 to 58 regions of the genome with a high frequency of ROH occurrence, also known as ROH islands, within individuals of particular breeds. ROH islands were detected more frequent in WL (n = 58) and the less in RJF (n = 26). The ROH islands had the average length in range of 223.5 kb ± 177.2 (WL) to 303.8 kb ± 416.6 (BRA), and average number of consecutive SNPs in range 478.3 ± 742.0 (WL) to 4947.3 ± 7220.3 (RJF). The average length of the ROH islands calculated for all breeds was 405.9 kb ± 103.2, while the average number of consecutive SNPs per region was 2108.6 ± 1576.4. Therefore, ROH islands tended to break into shorter segments (see WL in File S2), due to the presence of short heterozygous gaps within the ROH sequences (Nandolo et al. 2018). A summary statistics is presented in details in File S2 and shown schematically in Figure 4.

We found 18 ROH islands uniquely detected in wild birds (see Table 4 and File S2). Among the ROH islands, one and seven overlapping regions were detected respectively in layers and broilers out of which three regions in broilers reported to be in association with traits of economic interest available at the Chicken QTL database (release 40). These regions include genes related to antibody response to Newcastle virus (ROBO2), eggshell color (GLCCI1, ICA1, UMADI), antibody titer to SRBC antigen, body weight at 56 days (Meis2a.1, uc_338) and feather pecking (Table 4). In the same way, the ROH islands detected in RJs overlapped 53 QTL, among them are genes associated with eggshell strength and eggshell thickness (Table 4). The localized panels of ROHs were further compared with the putative selection signatures reported in Qanbari et al. (2019). Our comparison revealed overlap in only two ROH islands located on GGA2 and 5 (see Table 4). The genes located within ROH islands and detected in multiple populations have been presented in File S3.

We further conducted an enrichment test on the gene list located in ROH islands to identify over-represented Gene Ontology (GO) terms (Table 5). In commercial lines, genes located in ROH islands were associated with biological functions related to chicken domestication and evolution such as limb development (GREM1, MEOX2) and negative regulation of apoptotic process (AVEN, ASNS, GREM1) (Table 5). In contrast, the most significant GO term in Red Junglefowl was related to the reproductive traits such as oogenesis (PTN, WASH1, WEE2). These findings show candidate pathways associated with economically important traits and chicken genetic diversity during domestication and recent improvement.

CONCLUSION

This study provided the first systematic comparison of runs homozygosity islands between wild and domestic birds in sequence resolution. We found larger fraction of layers genome segregating in homozygote state, reflecting the recent inbreeding in commercial lines, although some of the long ROH tend to break into smaller tracts. Compared to the homogenous values reported for the commercial lines, wild birds showed important variation in the total length of ROH. Regions with a high frequency of ROH occurrence...
within domestic birds were co-localized with genes implicated in biological functions related to chicken domestication such as a limb development (GREM1, MEOX2), whereas in Red Junglefowl these regions overlapped with genes related to oogenesis. We also found a modest to high correlation between two molecular measurements, substantiating a highly inbred nature of domestic birds relative to their wild ancestors.

ACKNOWLEDGMENTS

RT and SQ conceived and designed the study. RT performed the bioinformatics and analyzed the data. SQ provided the R programming codes. RT and SQ wrote the manuscript. TS, SQ and GM edited the manuscript. All authors approved the manuscript before submission.

LITERATURE CITED

Ablondi, M., Å. Viklund, G. Lindgren, S. Eriksson, and S. Mikko, 2019 Signatures of selection in the genome of Swedish warmblood horses selected for sport performance. BMC Genomics 20: 717.
Almeida, O., G. Moreira, F. M. Rezende, C. Boschiero, J. de Oliveira Peixoto et al., 2019 Identification of signature selections involved in performance traits in a paternal broiler line. BMC Genomics 20: 449. https://doi.org/10.1186/s12864-019-5811-1
Axelsson, E., M. T. Webster, N. G. Smith, D. W. Burt, and H. Ellegren, 2014 Identification of genes linked to traits in the chicken. Nat. Genet. 46: 1217–1223. https://doi.org/10.1038/ng.2928
Bertolini, F., T. F. Cardoso, G. Marras, E. L. Nicolazzi, M. F. Rothschild et al., 2018 Genome-wide patterns of homozygosity reveal clues about the population history and adaptation of goats. J. Anim. Sci. 96: 3026–3037. https://doi.org/10.1093/jas/sky016
Brome, K. W., and J. L. Weber, 1999 Long homozygous chromosomal segments in reference families from the centre d’Etude du polymorphisme humain. Am. J. Hum. Genet. 65: 1493–1500. https://doi.org/10.1086/302661
Burren, A., M. Neuditschko, H. Hafner, M. Frischknecht, I. Reber et al., 2016 Genetic diversity analyses reveal first insights into breed-specific selection signatures within Swiss goat breeds. Anim. Genet. 47: 727–739. https://doi.org/10.1111/age.12476
Ceballos, F., C. P. K. Joshi, D. W. Clark, M. Ramsay, and J. F. Wilson, 2018a Runs of homozygosity: windows into population history and trait architecture. Nat. Rev. Genet. 19: 220–234. https://doi.org/10.1038/nrg.2017.109
Ceballos, F., C. S. Hazelhurst, and M. Ramsay, 2018b Assessing runs of homozygosity: a comparison of SNP Array and whole genome sequence low coverage data. BMC Genomics 19: 106. https://doi.org/10.1186/s12864-018-4489-0
Charlesworth, B., 2009 Fundamental concepts in genetics: effective population size and patterns of molecular evolution and variation. Nat. Rev. Genet. 10: 195–205. https://doi.org/10.1038/nrg2526
Chen, L., X. Wang, D. Cheng, K. Chen, Y. Fan et al., 2019 Population genetic analyses of seven Chinese indigenous chicken breeds in a context of global breeds. Anim. Genet. 50: 82–86. https://doi.org/10.1111/age.12732
Cingolani, P., A. Platts, L. L. Wang, M. Coon, T. Nguyen et al., 2012 A program for annotating and predicting the effects of single nucleotide polymorphisms, SnPEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2, iso-3. Fly (Austin) 6: 80–92. https://doi.org/10.4161/fly.19695
Clark, D. W., Y. Okada, K. Moore, D. Mason, N. Pirastu et al., 2019 Associations of autozygosity with a broad range of human phenotypes. Nat. Commun. 10: 4957. https://doi.org/10.1038/s41467-019-12283-6
Curik, I., M. Ferencakovic, and J. Soelkner, 2014 Inbreeding and runs of homozygosity: a possible solution to an old problem. Livest. Sci. 166: 26–34. https://doi.org/10.1016/j.livsci.2014.05.034
Elbelhtagi, A. R., F. Bertolini, D. S. Fleming, A. Van Goor, C. M. Ashwell et al., 2019 Natural selection footprints among African chicken breeds and village ecotypes. Front. Genet. 10: 376. https://doi.org/10.3389/fgene.2019.00376
Elferink, M. G., H. J. Megens, A. Vereijken, X. Hu, R. P. Crooijmans et al., 2012 Signatures of selection in the genomes of commercial and non-commercial chicken breeds. PLoS One 7: e32720. https://doi.org/10.1371/journal.pone.0032720
Ferencaković, M., E. Hamzić, B. Gredler, T. R. Solberg, G. Klemetsdal et al., 2013 Estimates of autozygosity derived from runs of homozygosity: empirical evidence from selected cattle populations. J. Anim. Breed. Genet. 130: 286–293. https://doi.org/10.1111/jbg.12012
Gibson, J., N. E. Morton, and A. Collins, 2006 Extended tracts of homozygosity in outbred human populations. Hum. Mol. Genet. 15: 789–795. https://doi.org/10.1093/hmg/ddi493
Gorsens, W., R. Meyermans, N. Buys, and S. Janssens, 2020 SNP genotypes reveal breed substructure, selection signatures and highly inbred regions in Piétrain pigs. Anim. Genet. 51: 32–42. https://doi.org/10.1111/age.12888
Goszczynski, D., A. Molina, E. Terán, H. Morales-Durand, P. Ross et al., 2018 Runs of homozygosity in a selected cattle population with extremely inbred bulls: Descriptive and functional analyses revealed highly variable patterns. PLoS One 13: e0200069. https://doi.org/10.1371/journal.pone.0200069
Grilz-Seger, G., M. Mesarić, M. Cotman, M. Neuditschko, T. Druml et al., 2018 Runs of homozygosity and population history of three horse breeds with small population size. J. Equine Vet. Sci. 71: 27–34. https://doi.org/10.1016/j.jevs.2018.09.004
Grilz-Seger, G., T. Druml, M. Neuditschko, M. Mesarić, M. Cotman et al., 2019 Analysis of ROH patterns in the Noriker horse breed reveals signatures of selection for coat color and body size. Anim. Genet. 50: 334–346. https://doi.org/10.1111/age.12797
Groenen, M. A., P. Wallberg, M. Foglio, H. H. Cheng, H. J. Megens et al., 2009 A high-density SNP-based linkage map of the chicken genome reveals sequence features correlated with recombination rate. Genome Res. 19: 510–519. https://doi.org/10.1101/gr.086538.108
Howrigan, D. P., M. A. Simonson, and M. C. Keller, 2011 Detecting autozygosity through runs of homozygosity: A comparison of three autozygosity detection algorithms. BMC Genomics 12: 460. https://doi.org/10.1186/1471-2164-12-460
Huang, D. W., B. T. Sherman, and R. A. Lempicki, 2009a Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res. 37: 1–13. https://doi.org/10.1093/nar/gkn923
Huang, D. W., B. T. Sherman, and R. A. Lempicki, 2009b Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat. Proteo. 4: 44–57. https://doi.org/10.1038/nprot.2008.211
Limper, J., 2018 Genomic inbreeding estimation and effective population size of four South African dairy breeds (Doctoral dissertation, University of Pretoria).
Luigi-Sierra, M. G., T. F. Cardoso, A. Martinez, A. Pons, L. A. Bermejo et al., 2019 Low genome-wide homozygosity in 11 Spanish ovine breeds. Anim. Genet. 50: 501–511. https://doi.org/10.1111/age.12832
Marchesi, J., M. E. Buzanskas, M. E. Cantão, A. Ibelli, J. O. Peixoto et al., 2018 Relationship of runs of homozygosity with adaptive and production traits in a paternal broiler line. Animal 12: 1126–1134. https://doi.org/10.1017/S1751731117002671
Mastrangelo, S., E. Ciani, M. T. Sardina, G. Sottile, F. Pilla et al., 2018 Runs of homozygosity reveal genome-wide autozygosity in Italian sheep breeds. Anim. Genet. 49: 71–81. https://doi.org/10.1111/age.12634
Mastrangelo, S., M. Tolone, M. T. Sardina, G. Sottile, A. M. Sutera et al., 2017 Genome-wide scan for runs of homozygosity identifies potential
candidate genes associated with local adaptation in Valle del Belice sheep. Genet. Sel. Evol. 49: 84. https://doi.org/10.1186/s12711-017-0360-z

Mastrangelo, S., M. Tolone, R. Di Gerlando, L. Fontanesi, M. T. Sardina et al., 2016 Genomic inbreeding estimation in small populations: evaluation of runs of homozygosity in three local dairy cattle breeds. Animal 10: 746–754. https://doi.org/10.1017/S1751731115002943

McQuillan, R., A. L. Leutenegger, R. Abdel-Rahman, C. S. Franklin, M. Pericic et al., 2008 Runs of homozygosity in European populations. Am. J. Hum. Genet. 83: 359–372. https://doi.org/10.1016/j.ajhg.2008.08.007

Metzger, J., M. Karwath, R. Tonda, S. Beltran, L. Águeda et al., 2015 Runs of homozygosity reveal signatures of positive selection for reproduction traits in breed and non-breed horses. BMC Genomics 16: 764. https://doi.org/10.1186/s12864-015-1977-3

Nandolo, W., Y. T. Utusunomiya, G. Mézéras, M. Wurzinger, N. Khayadzadeh et al., 2018 Misidentification of runs of homozygosity islands in cattle caused by interference with copy number variation or large intermarker distances. Genet. Sel. Evol. 50: 43. https://doi.org/10.1186/s12711-018-0414-x

Onzima, R. B., M. R. Upadhyay, L. F. Brito, M. Bosse, M. S. Lund, and G. Sahana, 2015 Estimates of autozygosity through runs of homozygosity in Gyr (Bos indicus) dairy cattle. BMC Genomics 19: 34. https://doi.org/10.1186/s12864-017-4365-3

Peripolli, E., N. B. Stafuzza, D. P. Munari, M. Silva, A. Lima, R. Irgang et al., 2018 Assessment of runs of homozygosity islands and estimates of genomic inbreeding in Gyr (Bos indicus) dairy cattle. BMC Genomics 19: 34. https://doi.org/10.1186/s12864-017-4365-3

Peripolli, E., N. B. Stafuzza, S. T. Amorim, M. de Lemos, L. Grigoletto et al., 2020 Genome-wide scan for runs of homozygosity in the composite Montana Tropical beef cattle. J. Anim. Breed. Genet. 137: 155–165. https://doi.org/10.3390/ani9121024

Purcell, S., B. Neale, K. Todd-Brown, L. Thomas, M. A. Ferreira et al., 2007 PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. 81: 559–575. https://doi.org/10.1086/519795

Purcell, D. C., S. McParland, E. Wall, and D. P. Berry, 2017 The distribution of runs of homozygosity and selection signatures in six commercial meat sheep breeds. PLoS One 12: e0176780. https://doi.org/10.1371/journal.pone.0176780

Quanbari, S., C. J. Rubin, K. Maqbool, S. Weigend, A. Weigend et al., 2019 Genetics of adaptation in modern chicken. PLoS Genet. 15: e1007989. https://doi.org/10.1371/journal.pgen.1007989

Quanbari, S., 2020 On the extent of linkage disequilibrium in the genome of farm animals. Front. Genet. 10: 1304. https://doi.org/10.3389/fgene.2019.01304

Quinlan, A. R., and I. M. Hall, 2010 BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26: 841–842. https://doi.org/10.1093/bioinformatics/btp033

Signer-Hasler, H., A. Burren, M. Neuditschko, M. Frischknecht, D. Garrick et al., 2017 Population structure and genomic inbreeding in nine Swiss dairy cattle populations. Genet. Sel. Evol. 49: 83. https://doi.org/10.1186/s12711-017-0358-6

Signer-Hasler, H., A. Burren, P. Ammann, C. Drögemüller, and C. Flury, 2019 Runs of homozygosity and signatures of selection: a comparison among eight local Swiss sheep breeds. Anim. Genet. 50: 512–525. https://doi.org/10.1111/age.12828

Strillacci, M. G., V. E. Vega-Murillo, S. I. Román-Ponce, F. López, M. C. Cozzi et al., 2018 Looking at genetic structure and selection signatures of the Mexican chicken population using single nucleotide polymorphism markers. Poult. Sci. 97: 791–802. https://doi.org/10.3382/ps/pex374

Szmatoła, T., A. Gurgul, I. Jasiełczuk, T. Żabek, K. Ropka-Molik et al., 2019 A comprehensive analysis of runs of homozygosity of eleven cattle breeds representing different production types. Animals (Basel) 9: 1024. https://doi.org/10.3390/ani9121024

Szmatoła, T., A. Gurgul, I. Jasiełczuk, W. Fu, and K. Ropka-Molik, 2020 A detailed characteristics of bias associated with long runs of homozygosity identification based on medium density SNP microarrays. J Genomics 8: 43–48. https://doi.org/10.7150/jgen.39147

Wang, L., Y. Mu, L. Xu, K. Li, J. Han et al., 2019 Genomic analysis reveals specific patterns of homozygosity and heterozygosity in inbred pigs. Animals (Basel) 9: 314. https://doi.org/10.3390/ani9060314

Xie, R., L. Shi, J. Liu, T. Deng, L. Wang et al., 2019 Genome-wide scan for runs of homozygosity identifies candidate genes in three pig breeds. Animals (Basel) 9: 518. https://doi.org/10.3390/ani9080518

Xu, Z., H. Sun, Z. Zhang, Q. Zhao, B. S. Olasege et al., 2019 Assessment of autozygosity derived from runs of homozygosity in Jinhua pigs disclosed by sequencing data. Front. Genet. 10: 274. https://doi.org/10.3389/fgene.2019.00274

Yoshida, G. M., P. Cáceres, R. Marín-Nahuelpi, B. F. Koop, and J. M. Yáñez, 2020 Estimates of autozygosity through runs of homozygosity in farmed Coho salmon. Genes (Basel) 11: 490. https://doi.org/10.3390/genes11050490

Zhang, J., C. Nie, X. Li, Z. Ning, Y. Chen et al., 2020 Genome-wide population genetic analysis of commercial, indigenous, game, and wild chickens using 600K SNP microarray data. Front. Genet. 11: 543294. https://doi.org/10.3389/fgene.2020.543294

Zhang, M., W. Han, H. Tang, G. Li, M. Zhang et al., 2018 Genomic diversity dynamics in conserved chicken populations are revealed by genome-wide SNPs. BMC Genomics 19: 598. https://doi.org/10.1186/s12864-018-4973-6

Zhang, Q., B. Guldbrandtsen, M. Bosse, M. S. Lund, and G. Sahana, 2015 Runs of homozygosity and distribution of functional variants in the cattle genome. BMC Genomics 16: 542. https://doi.org/10.1186/s12864-015-1715-x

Communicating editor: D.-J. de Koning