*Animal* (2018), **12:6**, pp 1126–1134 © The Animal Consortium 2017
doi:10.1017/S1751731117002671

# Relationship of runs of homozygosity with adaptive and production traits in a paternal broiler line

J. A. P. Marchesi[1a], M. E. Buzanskas[2], M. E. Cantão[3], A. M. G. Ibelli[3], J. O. Peixoto[3], L. B. Joaquim[1], G. C. M. Moreira[4], T. F. Godoy[4], A. P. Sbardella[1], E. A. P. Figueiredo[3], L. L Coutinho[4], D. P. Munari[1] and M. C. Ledur[3†]

*[1]Departamento de Ciências Exatas, Faculdade de Ciências Agrárias e Veterinárias, Universidade Estadual Paulista (Unesp), Via de Acesso Prof. Paulo Donato Castelane, S/N – Vila Industrial, Jaboticabal, São Paulo 14884-900, Brazil; [2]Departamento de Zootecnia, Centro de Ciências Agrárias – Campus II, Universidade Federal da Paraíba (UFPB), Rodovia BR 079 – Km 12, Areia, Paraíba 58397-000, Brazil; [3]Embrapa Suínos e Aves, Rodovia BR 153, Km 110, Distrito de Tamanduá, Caixa Postal 321, Concórdia, Santa Catarina 89715-899, Brazil; [4]Departamento de Zootecnia, Escola Superior de Agricultura 'Luiz de Queiroz', Universidade de São Paulo, Av. Pádua Dias 11, Piracicaba, São Paulo 13419-900, Brazil*

(Received 29 March 2017; Accepted 15 September 2017; First published online 25 October 2017)

*Genomic regions under high selective pressure present specific runs of homozygosity (ROH), which provide valuable information on the genetic mechanisms underlying the adaptation to environment imposed challenges. In broiler chickens, the adaptation to conventional production systems in tropical environments lead the animals with favorable genotypes to be naturally selected, increasing the frequency of these alleles in the next generations. In this study, ~1400 chickens from a paternal broiler line were genotyped with the 600 K Affymetrix® Axiom® high-density (HD) genotyping array for estimation of linkage disequilibrium (LD), effective population size ($N_e$), inbreeding and ROH. The average LD between adjacent single nucleotide polymorphisms (SNPs) in all autosomes was 0.37, and the LD decay was higher in microchromosomes followed by intermediate and macrochromosomes. The $N_e$ of the ancestral population was high and declined over time maintaining a sufficient number of animals to keep the inbreeding coefficient of this population at low levels. The ROH analysis revealed genomic regions that harbor genes associated with homeostasis maintenance and immune system mechanisms, which may have been selected in response to heat stress. Our results give a comprehensive insight into the relationship between shared ROH regions and putative regions related to survival and production traits in a paternal broiler line selected for over 20 years. These findings contribute to the understanding of the effects of environmental and artificial selection in shaping the distribution of functional variants in the chicken genome.*

**Keywords:** animal breeding, *Gallus gallus*, inbreeding, oxidative stress, selection

## Implications

The use of genomic data to evaluate selection signals in livestock may provide new and precise insights about the effects of natural or artificial selection for adaptive traits. In our study, chickens that acquired favorable fitness traits in tropical environment had a reduction in the variability of genomic regions associated with survival. Through the study of runs of homozygosity, that are regions of the genome where the copies inherited from parents are identical, it was possible to identify candidate genes that might be helping the adaptation and resistance to heat stress in tropical environments, and those related to production traits under selection in this paternal broiler line.

## Introduction

The poultry industry is one of the main chains responsible for animal protein production (meat and eggs) for human consumption (OECD-FAO, 2015). The increased consumer demand has forced the poultry industry to develop methods to increase their production output, offering chicken meat at competitive prices in the worldwide market. With the intensification of selection to obtain more productive genotypes for economic traits, the modern broiler chickens have experienced notable phenotypic and genetic changes (Rubin *et al.*, 2010). In addition, the challenging environments of the production system have also contributed for shaping the genome of these chickens by having them to adapt to the newly imposed conditions (Badyaev, 2005).

[a] Present address: Departamento de Genética, Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Bairro Vila Monte Alegre, 14049-900, Ribeirão Preto, São Paulo, Brazil.
[†] E-mail: monica.ledur@embrapa.br

Chicken raised in tropical regions suffer from heat stress effects because the temperatures are high in the whole year. Besides that, these effects are intensified due to the high density of broilers used in current production systems (Türkyilmaz, 2008). Biochemical and physiological events associated with hyperthermia can potentially promote reactive oxygen species (ROS) formation, disruption of internal energy balance, lipid peroxidation and, consequently, immunosuppression (Pamok *et al*., 2009). It has been observed that elevated environmental temperatures affect the development of a specific immune response in chickens. Quinteiro-Filho *et al*. (2010) reported that heat stress activates the chicken hypothalamic–pituitary–adrenal axis, increasing corticosterone serum levels and, consequently, influencing the quality of the intestinal-immune barrier, thereby allowing pathogenic bacteria to migrate through the intestinal mucosa generating inflammatory infiltrate. Therefore, the conditions imposed by the environmental changes that chickens have undergone in their evolutionary course will define their response to adaptation (Badyaev, 2005).

The constant selective pressure for one specific trait can lead to the reduction in the variability around the genomic regions associated with this trait. Therefore, the identification of those regions may allow a better understanding about their biological importance. In broilers, the presence of genomic signatures related to selective pressure has been extensively studied, and several methods have been used to identify these segments under selection, such as Z-transformed heterozygosity and cross-population extended haplotype homozygosity, which measures the reduction in diversity of a genomic region (Fu *et al*., 2016).

A potential alternative to detect signatures of selection for complex traits is the identification of long stretches of consecutive homozygous genotypes in the genome, also known as runs of homozygosity (ROH) (Fleming *et al*., 2016). The presence of ROH can provide useful information about the population selection history, as well as enable a better understanding of genotype–phenotype relationships through the discovery of genes or genomic regions that are or have been under selective pressure (Zhang *et al*., 2015). Recently, ROH extracted from single nucleotide polymorphism (SNP) arrays have been used to study the population history of chickens. Fleming *et al*. (2016), working with Ugandan and Rwandan chicken ecotypes, showed that these populations have alleles under selective pressure from their environment, which may aid in adaptation to harsh environments.

Increased homozygous levels are associated with a reduction in effective population size ($N_e$), which is a measure of diversity within a population that allows the estimation of the rate of inbreeding per generation and the loss of genetic variation (Zhu *et al*., 2013). It is already well established that high-density SNP panels aid estimates of population parameters, such as $N_e$, intergenerational inbreeding, and linkage disequilibrium (LD) between loci within a population. Molecular approaches based on ROH and SNPs may help to avoid several disadvantages of using pedigrees to analyze inbreeding. Hence, the present study was undertaken to identify conserved regions based on ROH using genomic data to investigate their relationship with functional variants in a paternal broiler line.

## Material and methods

*Chicken population*

The chickens used in this study were a sample composed by five hatches from a paternal broiler line called TT, made in 2007, to represent the whole TT line. This line was developed by the Embrapa Swine and Poultry National Research Center, in Concordia, Santa Catarina State – South of Brazil. The TT pure line has been maintained under multi-trait selection, in open sided poultry houses, since 1992. In 1991, the TT founder population size was 89 sires pen-mated to 10 dams each. From 1992 to 1996, 50 sires were pen-mated to 10 dams each every year, to produce the next generations. In 1997, the selected sires were 31, and each mated to five dams. In the year 2000, there were 25 sires mating five dams each, being the population structure used to date.

The main traits under selection in TT pure line are body weight (BW) feed conversion, breast size (mass selection), carcass and cuts yield, abdominal fat (indirect selection based on non-selected sibs), fertility, hatchability of fertile eggs and chick viability (independent culling levels). The TT performance level in 1992, as hatched average weight at 35 days of age was 1435 g, and individual feed conversion rate (FCR), from 36 to 42 days of age, of the 300 mass selected males was 2.397. In 2007, as hatched average weight at 42 days of age was 2457 g, and individual FCR from 43 to 49 days of age was 2.798. Note that the age of selection for BW has changed from 35 to 42 days of age in the described period, as well as the FCR evaluation period, which has changed from 36 to 42 days since 1992 to 43 to 49 days of age in 2007.

The sample used in this study was called 'TT Reference Population', because it was developed for genomic studies and had about 92 traits measured and evaluated in ~1400 individuals, which were all slaughtered in 2008. Details on this population can be found in Ledur *et al*. (2012) and Fornari *et al*. (2014). In brief, one sire from each sire family (25) and one dam from each dam family (125) were chosen to generate the TT Reference Population. Matings between close relatives were avoided, and conducted in a proportion of one male to five females to produce about 1600-day-old chicks of both sexes in five hatches. The TT Reference Population founders that had their offspring genotyped represented 20 from the original 25 sires, and 93 from the original 125 dams. The complete and accurate genealogy is available since the foundation of the TT line, totaling 2007 pedigree records distributed in 17 generations. Although the average hottest temperature was 35.0°C in December and the coolest 0.2°C in July from 1992 to 2008, the temperatures are highly variable in this region, and the extremes of the temperature in this period ranged from −3.6°C to 37.5°C (Agrometeorological Station of the Embrapa Swine and Poultry National Research Center).

*Genotyping data*

Blood samples from 1430 chickens (652 males and 778 females) were collected, of which 113 (20 sires and 93 dams) are the progenitors (born in 2007) and the remaining are full- and half-siblings (born in 2008). The extraction of the genomic DNA was performed using PureLink® Genomic DNA (Invitrogen, Carlsbad, CA, USA) kit. The extracted DNA was quantified using a Qubit® 2.0 Fluorometer (Thermo Fisher Scientific, Waltham, MA, USA) and diluted to a concentration of ~10 ng/ul. Samples were genotyped using the 600 K Affymetrix® Axiom® high-density (HD) genotyping array at the 'Centro de Genomica Funcional, ESALQ', University of São Paulo, Brazil.

The genotypes were initially read and edited with the Axiom™ Analysis Suite (Affymetrix®) software using a DishQC removal value of 0.82. The SNPs quality control was performed in PLINK v.1.9 (http://pngu.mgh.harvard.edu/purcell/plink/) excluding SNPs with call rate lower than 98%, significant deviations ($P < 10^{-6}$) from the Hardy–Weinberg equilibrium (HWE), and minor allele frequency (MAF) lower than 2%. Samples presenting call rate lower than 90% were also excluded. For the analysis, only the SNPs with positions attributed to autosomal chromosomes (based on the *Gallus gallus* v. 4.0 genome reference) were considered.

*Linkage disequilibrium and effective population size*

The LD estimates were performed with PLINK v.1.9, using the $r^2$ calculated between marker pairs separated by <5 Mb on each chromosome, as described by Hill and Robertson (1968), with the equation $r^2 = \frac{D^2}{f(A)f(a)f(B)f(b)}$, in which $f(A)$, $f(a)$, $f(B)$ and $f(b)$ are the respective frequencies of alleles $A$, $a$, $B$ and $b$, and $D = f(AB) - f(A)f(B)$.

The LD was estimated separately for each category of the avian chromosomes: macrochromosomes (GGA1 to 5), intermediate chromosomes (GGA6 to 10) and microchromosomes (GGA11 to 38), according to the chromosome size classification of the International Chicken Genome Sequencing Consortium (2004). To visualize the pattern for each category of LD, the $r^2$ values between all pairs of SNPs within each chromosome group were arranged in ascending order, grouped according to their pairwise physical distance in classes of 50 Kb, started from 0 to 5000 Kb, and then the mean was calculated for all of these classes and plotted using the R software (https://www.r-project.org/).

The effective population size ($N_e$) for different periods of the population history were estimated with the SNP1101 package (Sargolzaei, 2014), based on the expectation of $r^2$ in different distances, with the equation described by Sved (1971): $N_e = \left[\left(\frac{1}{E(r^2)}\right) - 1\right] \times \frac{1}{4c}$, where $c$ is the distance between two SNPs, in Morgans, estimated for each chromosome in the LD analysis. The $E(r^2)$ is the expected $r^2$ at distance $c$, calculated as $N\left(r^2\right) = \frac{1}{1+4N_e c}$. Each genetic distance ($c$) corresponds to a value of $t$ generations in the past calculated as $t = \frac{1}{2c}$ (Hayes *et al.*, 2003). The $N_e$ was investigated in 6 points: 5, 10, 20, 50, 100 and 200 generations ago.

*Runs of homozygosity analysis*

The ROH were identified using PLINK v.1.9, and calculated for each individual. For the ROH analysis, segments with at least 100 SNPs with a minimum length of 1000 Kb, density of 1 SNP/50 Kb, accepting the presence of one heterozygous SNP, the absence of one SNP in a window of 100 SNPs, and allowed one gap within the 1000 Kb were considered. The population ROH frequency was calculated using all identified ROH. The search for shared ROH, which are consensus regions among individuals, was made by looking for how many individuals had a ROH with at least the same start and end positions. Those overlapped homozygous regions were grouped and called shared ROH, which were identified using a Perl homemade script.

*Functional analyses of runs of homozygosity*

Genes located within shared ROH regions present in more than 1% of the population were searched using the UCSC genome browser (Galgal 4) and used in the enrichment analysis. This threshold was chosen based on the assumption that ROH occurring in more than 1% of the population has less probability of being randomly assigned. Gene ontology (GO) enrichment analysis was performed using David 6.8 database (https://david.ncifcrf.gov/home.jsp) and the results were considered statistically significant if the false discovery rate (FDR) was lower than 0.20. The REVIGO software (http://revigo.irb.hr/) was used for visualization and removal of redundant terms of the GO enrichment results. A gene network of predicted functional proteins was constructed using the STRING 10.0 software (http://string-db.org/).

*Inbreeding estimates based on genomic and pedigree data*

Different analyses were performed to estimate the inbreeding coefficient using genomic and pedigree data. The genealogical inbreeding coefficient ($F_{PED}$) was estimated using the ENDOG v4.8 software (http://pendientedemigracion.ucm.es/info/ prodanim/html/JP_Web.htm) considering all pedigree information that had been recorded since the foundation of the population. To estimate individual genomic inbreeding coefficients using the ROH data ($F_{ROH}$), the length of the genome covered by ROH was divided by the total chicken autosomal genome length covered by the SNP array (931 280 Kb, Gallus gallus v. 4.0 NCBI, June 2016), as described by Ferencakovic *et al.* (2012). The genomic SNP-by-SNP inbreeding coefficient ($F_{SNP}$) for each individual i ($F_{SNP}(i)$) was obtained based on the excess of SNP homozygosity, and calculated as $F_{SNP}\ (i) = (OH_i - EH)^2/(EH)$, where OH is the observed number of homozygotes and EH is the expected number of homozygotes. The inbreeding coefficients obtained by the three methods were compared using Pearson's correlation.

*Declaration of ethics in the use of animals*

This study was conducted with the approval of the Ethics Committee for Animal Use (CEUA) from the Embrapa Swine and Poultry National Research Center under the protocol no. 011/2010.

## Results

In the quality control, 70 552 SNPs failed in the call rate, 121 199 failed the MAF criteria, 13 434 were excluded based on HWE test and 20 256 were present on chromosomes not included in the analysis: 15 305 in the sex chromosomes, 160 in the linkage groups and 4791 in undefined positions, remaining 355 520 SNPs for the analyses (Supplementary Table S1).

*Linkage disequilibrium and effective population size*

The LD declined as the distance between markers increased in all chromosome groups as shown in Figure 1. The mean $r^2$ converged to 0.30 approximately at 500 to 550 Kb in macro, 350 to 400 Kb in intermediate and 150 to 200 Kb in microchromosomes. The average LD between adjacent SNPs in all autosomes was 0.37 and micro, intermediate and macrochromosomes presented the lowest, second lowest and highest mean $r^2$, respectively.

The $N_e$ obtained in our study was estimated from five to 200 generations ago (Figure 2). The decay observed in

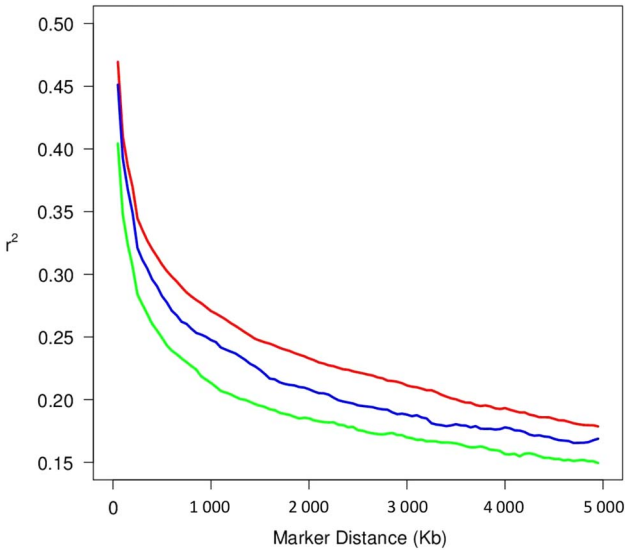


**Figure 1** Decay of average pairwise linkage disequilibrium ($r^2$) over the distance between single nucleotide polymorphisms (SNPs) in the different chromosome size classes (macrochromosomes in red, intermediated chromosomes in blue and microchromosomes in green) in a paternal broiler line.

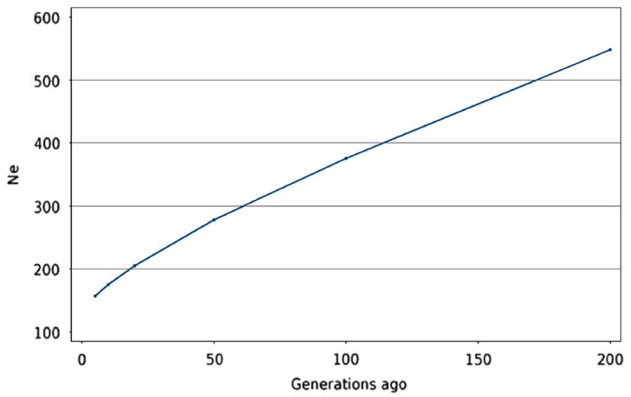


**Figure 2** Estimated effective population size ($N_e$) over time in a paternal broiler line.

Figure 2 indicates that the $N_e$ of the ancestral population based on 200 past generations was much larger, with 548 individuals, compared with the most current generations. In the last five generations, $N_e$ ranged from 113 (current population) to 157 chickens.

*Inbreeding coefficient estimates*

The average, minimum and maximum inbreeding coefficients based on pedigree records ($F_{PED}$) were equal to 0.0423, 0.00 and 0.0981, respectively. Based on ROH observations ($F_{ROH}$), the average, minimum and maximum inbreeding coefficients were 0.0739, 0.00 and 0.2136, respectively. The estimates of inbreeding coefficient obtained using the SNP-by-SNP ($F_{SNP}$) method ranged from −0.1669 to 0.1394 with an average of −0.0185. The $F_{SNP}$ sometimes can be negative, when EH is higher than $OH_i$.

The correlation between $F_{ROH}$ and $F_{SNP}$ was moderately high (0.63), however, the correlation between $F_{PED}$ with $F_{ROH}$ and $F_{SNP}$ were low (0.06 and 0.05, respectively).

*Runs of homozygosity*

A total of 1279 chickens presented ROH segments with an average of 12.87 ROH per animal and a maximum number of 30 ROH. The average length and the longest ROH were 5.34 and 59.24 Mb, respectively. The macrochromosomes had the highest number of shared ROH among chickens and the largest mean ROH length. Among them, the chromosome GGA1 presented the highest number of shared ROH (1895) in the largest number of animals in the population (1172), whereas GGA2 presented the largest mean ROH length (7.52 Mb). The chromosome GGA16 did not present any ROH segment according to the parameters used (Table 1). The average amount of the genome covered by ROH in the studied population was 7.22% per individual, ranging from 1% to 20.87%.

The total number of homozygous segments was 9414, of which 6843 were observed in more than 1% of the population. Most of the identified ROH presented a small size (1 to 5 Mb), with the great majority of the segments (~64%) with frequency between 1% and 5% in the population. Segments with size between 5 and 10 Mb were the second most abundant and presented the highest number of ROH in more than 1% of the population (Figure 3). A ROH of 11.72 Mb located on GGA2 (77 084 081 to 88 808 596 bp) was the largest segment, with a frequency of ~5% in the population. The most frequent shared segment (~28%) identified with 5.21 Mb of length was also located in GGA2 (69 807 237 to 75 014 591 bp). This ROH contains 12 genes of which seven are novel genes, one is a microRNA (*gga-mir-6545*) and the others are from the cadherin superfamily: cadherin 9 (*CDH9*), cadherin 10 (*CDH10*), cadherin 12 (*CDH12*) and cadherin 18 (*CDH18*).

A gene list with 756 genes identified in the ROH regions was analyzed using the GO enrichment analysis. A total of 72 GO-terms was enriched with biological processes (FDR lower than 0.20; Supplementary Material S1), and filtered for redundancy of terms, producing a list of significant GO-terms related

**Table 1** *Summary of runs of homozygosity (ROH) by chromosome in a broiler population*

| Chromosomes | Length (Mb)[1] | Number of ROH | Number of animals | Mean ROH length (Mb) | Mean SNP[2] |
|---|---|---|---|---|---|
| GGA1 | 196.20 | 1895 | 1172 | 6.13 | 1890 |
| GGA2 | 149.56 | 1328 | 1129 | 7.52 | 1784 |
| GGA3 | 111.30 | 947 | 951 | 5.82 | 1863 |
| GGA4 | 91.28 | 758 | 966 | 6.36 | 1813 |
| GGA5 | 59.83 | 530 | 724 | 6.13 | 1836 |
| GGA6 | 35.47 | 377 | 577 | 4.85 | 1981 |
| GGA7 | 36.95 | 477 | 669 | 4.78 | 1840 |
| GGA8 | 29.96 | 206 | 402 | 4.44 | 1594 |
| GGA9 | 24.09 | 355 | 537 | 3.81 | 1802 |
| GGA10 | 20.44 | 375 | 642 | 3.44 | 1906 |
| GGA11 | 20.22 | 219 | 327 | 4.28 | 2047 |
| GGA12 | 19.95 | 252 | 375 | 3.99 | 1997 |
| GGA13 | 18.41 | 194 | 327 | 4.26 | 1764 |
| GGA14 | 15.60 | 198 | 334 | 3.41 | 1859 |
| GGA15 | 12.76 | 190 | 390 | 3.21 | 1642 |
| GGA16 | 0.65 | – | – | – | – |
| GGA17 | 10.96 | 169 | 269 | 2.78 | 1640 |
| GGA18 | 11.05 | 163 | 226 | 3.51 | 1973 |
| GGA19 | 9.98 | 145 | 244 | 2.90 | 1588 |
| GGA20 | 14.11 | 164 | 262 | 3.67 | 1707 |
| GGA21 | 6.86 | 89 | 136 | 1.72 | 1543 |
| GGA22 | 4.73 | 24 | 42 | 1.29 | 1475 |
| GGA23 | 5.79 | 79 | 137 | 2.19 | 1613 |
| GGA24 | 6.28 | 127 | 289 | 1.99 | 1534 |
| GGA25 | 2.91 | 7 | 12 | 1.78 | 1162 |
| GGA26 | 5.31 | 71 | 123 | 1.79 | 1361 |
| GGA27 | 5.66 | 41 | 63 | 1.87 | 1412 |
| GGA28 | 4.97 | 34 | 60 | 2.15 | 1411 |

SNP = single nucleotide polymorphism.
[1]Chromosome length.
[2]Mean SNP per ROH.

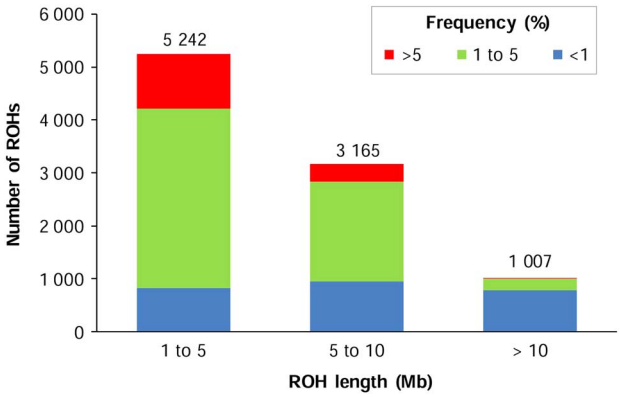


**Figure 3** Distribution of runs of homozygosity (ROH) identified by length and frequency class in a paternal broiler line.

to selective pressure for environmental stressors, morphological pattern and immunity (Figure 4). The most relevant GO-terms related to environmental stressors were: cell surface receptor signaling pathway (GO:0007166), positive regulation of Mitogen-Activated Protein Kinases (MAPK) cascade (GO:0043410), cellular response to hypoxia (GO:0071456), canonical Wnt signaling pathway (GO:0060070) and cellular response to toxic substance (GO:0097237), all grouped into the most significant supercluster cellular response to hypoxia (Figure 4).

The gene network showed some main interaction nodes grouped by important genes, such as the cell division cycle 42 (*CDC42*), RAP1A, member of RAS oncogene family (*RAP1A*), ras homolog family member C (*RHOC*) and elongation factor Tu GTP binding domain containing 2 (*EFTUD2*), interacting with several other genes identified in the ROH (Figure 5). Since many interactions were found with these genes, it could be an indication that their variants are possibly involved in important biological processes related to survival traits in this selected population.

## Discussion

The effects of LD are important to understand the variations in the genome, particularly when selection is being carried out (Bosse *et al.*, 2012). In our study, the extension of the LD related to the physical distance between markers was smaller in microchromosomes than in macrochromosomes and in intermediate chromosomes (Figure 1), which is in accordance with the results obtained by other authors (Qanbari *et al.*, 2010; Fu *et al.*, 2015). According to the International Chicken Genome Sequencing Consortium (2004) the difference among the extension of the LD in these chromosome classes occurs because the microchromosomes present higher recombination rates (6.4 cM/Mb), followed by the intermediate chromosomes (3.9 cM/Mb), and the macrochromosomes (2.8 cM/Mb). The average LD found in the current study between adjacent markers in all autosomes (0.37) was comparable with that found by Fu *et al.* (2015) in pure lines (from 0.34 to 0.40). When purebred and crossbred commercial broiler lines were compared, the extent and consistency of LD for short distances between markers (0 to 10 Kb) were remarkably high, but the crossbred lines had a considerably lower level of LD than the purebred lines (Fu *et al.*, 2015).

In this study, the LD was used to estimate $N_e$ instead of the homozygosity of individual markers, because it has the advantage of using the recombination rate, which is more controllable than the mutation rate (Hayes *et al.*, 2003). As a result, the latest $N_e$ can be estimated because the recombination rates are much higher than mutation rates (Hayes *et al.*, 2003). The $N_e$ estimates were substantially higher than those reported by Andreescu *et al.* (2007) in broiler (from 50 to 200) and by Qanbari *et al.* (2010) in egg-laying chicken lines, which confirms the larger $N_e$ estimates for broiler than for layer lines observed by Muir *et al.* (2008). The $N_e$ estimated in our study for the last five generations was 157 animals (Figure 2). Qanbari *et al.* (2010) found $<70$ and 50 individuals for brown and white layers, respectively, in five generations ago and pointed out the existence of low inbreeding coefficients in the population. According to these authors, both lines studied, whether commercial or experimental, showed clear evidence of $N_e$ decay over the generations. In our study, the latest generation (parental generation) presented $N_e$ of 113 animals, which is higher than the

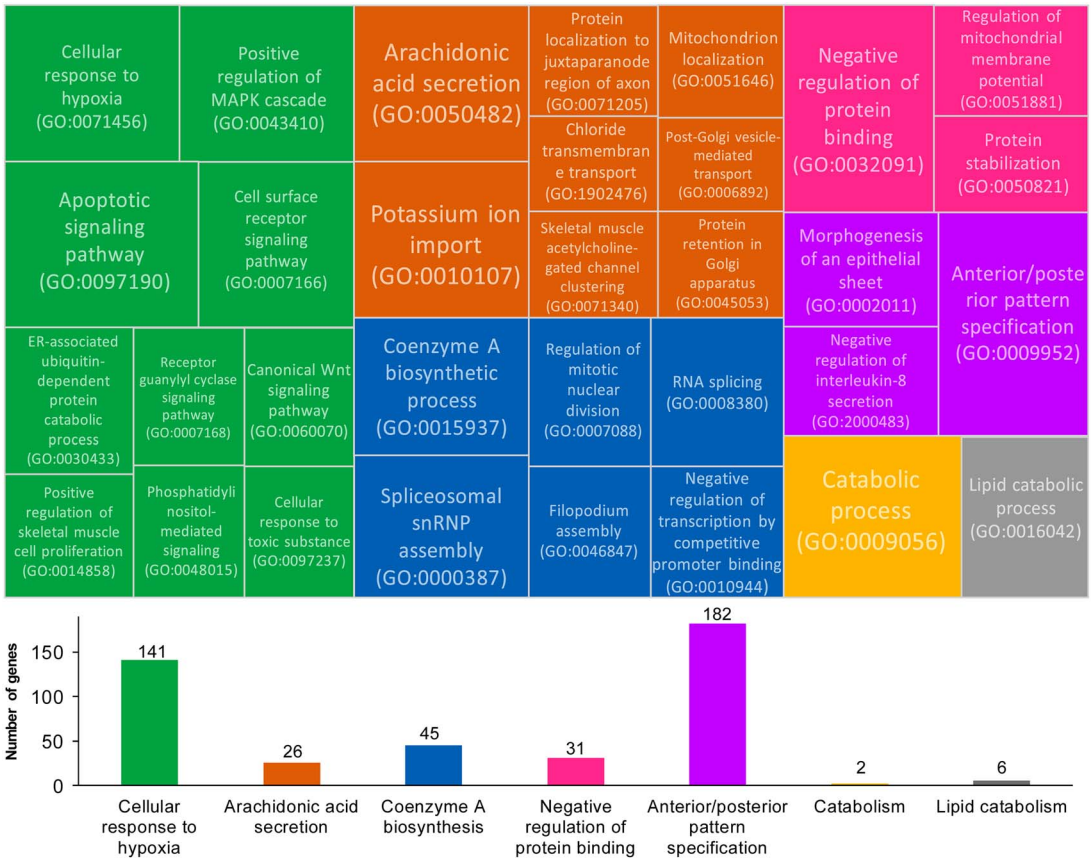


**Figure 4** 'Treemap' of biological processes based on Gene Ontology terms statistically significant (false discovery rate < 0.20) for genes identified in shared runs of homozygosity (ROH) in the broiler population. Each rectangle is a single cluster representative. The representatives are joined into 'superclusters' of related terms, visualized with different colors.

recommended by the Food and Agriculture Organization of the United Nations (2004) for conservation purpose, of a minimum of 50 individuals for any livestock species.

In the studied population, the inbreeding coefficient was low in the three methodologies used, evidencing that the matings were well controlled to avoid inbreeding. The low correlation between inbreeding estimates based on pedigree and genomic data is expected for several reasons. According to Ferencakovic *et al.* (2012), $F_{PED}$ methodology does not consider the stochastic nature of recombination and fails to capture the influence of relatedness among founders from the base population. Furthermore, $F_{PED}$ describes IBD status with respect to a rather poorly defined founder generation considered to be unrelated. For example, the estimated human inbreeding from ROH data ($F_{ROH}$) resulting from the mating of first cousins would be 0.0625, with a SD of 0.0243, whereas the average $F_{PED}$ from the same parents is always the same (0.0625) (Carothers *et al.*, 2006). This variance increases with each meiosis, and it is even possible for offspring of third cousins to be more autozygous than offspring of second cousins (McQuillan *et al.*, 2008). Also, $F_{PED}$ assumes that the entire genome is selection-neutral and does not account for potential bias resulting from selection (Curik *et al.*, 2002). The TT line has been under multi-trait selection since 1992, explaining part of the low correlation found between $F_{PED}$ and $F_{ROH}$ in this study (0.06). Although matings among close relatives are avoided, some degree of inbreeding is generated, since TT is a closed population. Low correlation between $F_{PED}$ and $F_{ROH}$ were also observed in several species, as discussed by Zanella *et al.* (2016), who also found this correlation to be low, of 0.015, for Large White pigs. Therefore, the use of $F_{ROH}$ becomes more efficient in the detection of the real percentage of the genome that is IBD than traditional methodologies.

Regarding ROH, a significant correlation between the sharing of ROH and regions putatively under selection suggests that some shared ROH are the result of a combination of inbreeding and selection (Zhang *et al.*, 2015). In addition, it is possible that ROH carry genes or variants under positive selection (Bosse *et al.*, 2012). In this study, the most representative supercluster identified by the GO-terms enrichment was related to response to hypoxia (Figure 4), which is intimately related to oxidative stress. This indicates that the individuals in the analyzed population have possibly experienced selective pressure on their genomes to events that are necessary to deal with oxidative stress, such as selection for rapid growth rate and adaptation. In the last decades, the intense selection for heavier and faster growing broilers, and the intensification of the production system have increased the metabolism of fast-growing chickens; as consequence, metabolic disorders became a worldwide concern. These metabolic problems affect primarily the cardiovascular and skeletal muscle systems, causing more economic losses than infectious agents (Julian, 2005). The increase in

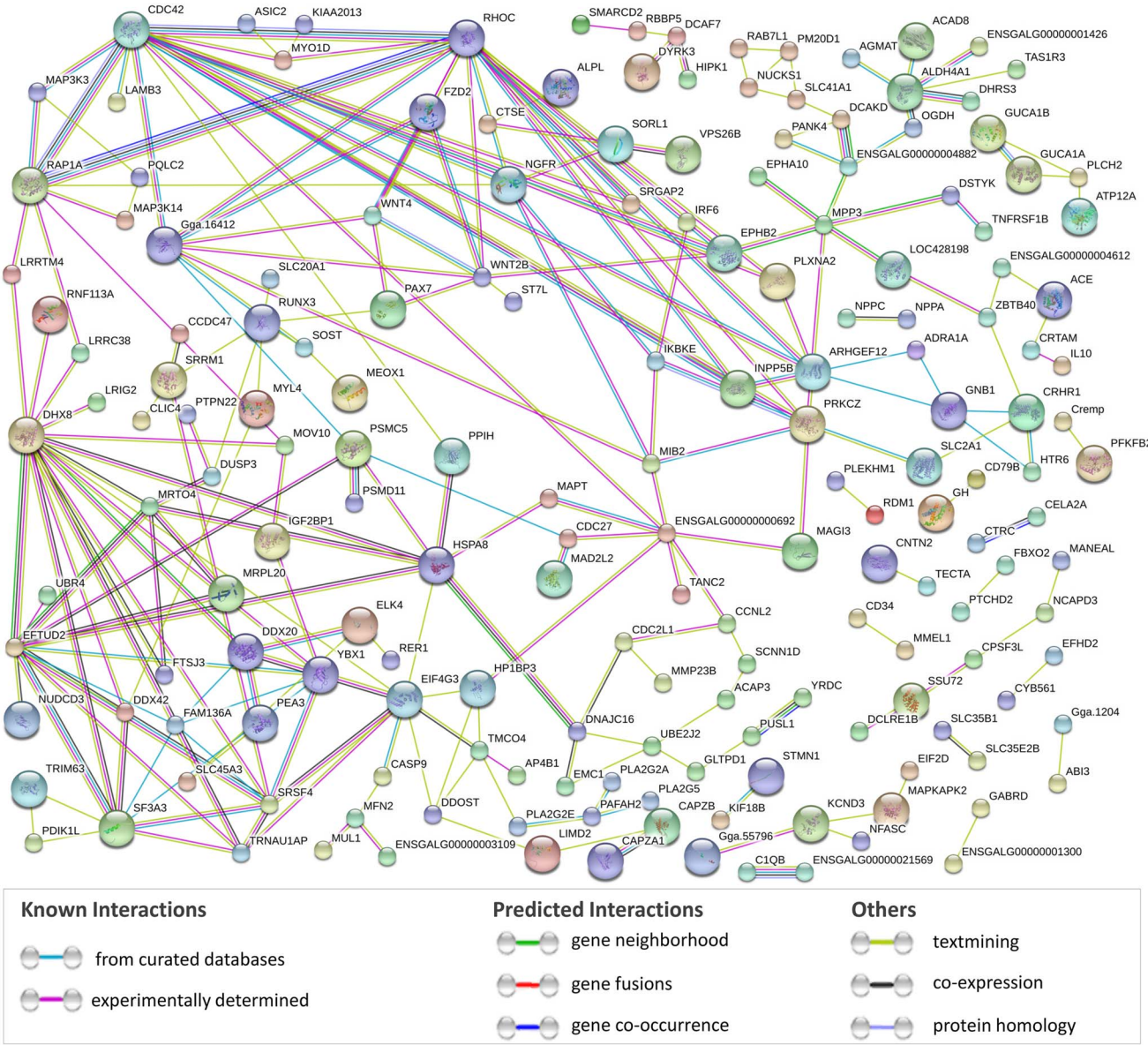


**Figure 5** Network of predicted functional proteins for the genes located within runs of homozygosity identified in the broiler population created using STRING. The lines represent the predicted interactions between genes.

metabolic disorders, such as heart failure syndrome, ascites and lungs and heart edemas can be directly related to an insufficient oxygen supply (Scheele, 1997).

In response to heat stress, several physiological and biochemical events have the potential to promote ROS in the cell, especially in the mitochondria (Mujahid *et al.*, 2006). High levels of these molecules result in the disturbance of balance between oxidation and antioxidant defense systems, causing lipid peroxidation and oxidative damages to proteins and DNA (Lin *et al.*, 2006). In the enrichment analysis, several GO-terms related to CoA biosynthesis were identified (Figure 4). This enzyme helps to transfer fatty acids from the cytoplasm to mitochondria. The arachidonic acid secretion supercluster was also well represented in the enrichment analysis (Figure 4). This molecule is abundant in the brain, muscles and liver, and previous studies have demonstrated that the mobilization of arachidonic acid occurs in response to oxidative stress and is profoundly different from one cell to another (Balboa and Balsinde, 2006). The GO-term Positive regulation of MAPK cascade (GO:0043410) was also enriched in our study. The activation of different members of the MAPK is involved in directing cellular responses to a diverse array of stimuli, as the response to heat stress (Gorostizaga *et al.*, 2005). The identification of several genes related to these processes may indicate that associated genetic variants have undergone a positive selection for greater efficiency in maintaining cellular homeostasis. Fleming *et al.* (2016), studying Ugandan and Rwandan chicken ecotypes, identified genes in ROH and selection signature regions related to kinases, calcium activity and oxidative stress responses, and related these to selective pressure on genes involved in mechanisms used by chickens to tolerate the effects of a

challenging environment. Some of the genes identified by Fleming *et al.* (2016) were also present in the ROH regions found in our population, such as the dispatched RND transporter family member 3 (*DISP3*) and syntaxin 12 (*STX12*) genes.

Mujahid *et al.* (2005) have suggested that oxidative stress resulting from exposure to heat can lead to cytotoxicity and the generated ROS have been demonstrated to be mediators of cellular injury. Anwar *et al.* (2008) reported that broilers under heat stress had the relative weight of their major immunological organs, such as spleen, thymus and bursa of Fabricius, decreased. Aengwanich (2008) also found that the Bursa of Fabricius in heat-stressed broilers was atrophied when examined under histopathological observation. The Bursa of Fabricius is an important oval-like shaped gland located at the proctodaeal region of the cloaca in chickens, and is known to be a primary lymphoid organ where immunologically competent cells are produced, whereas it is also a secondary or peripheral lymphoid organ which produces antibodies (Tsuji and Miypshi, 2001).

In this study, GO-terms as negative regulation of interleukin-8 secretion (GO: 2000483), associated with the immune system, were identified (Figure 4). Several genes involved in immune system processes have also been identified in the network analysis (Figure 5). One of the genes with the highest number of predicted interactions is the *CDC42,* reported as critical for interstitial dendritic cell migration during immune responses (Harada *et al.*, 2012), and overexpressed in chicken macrophages infected with *Salmonella enterica* (Zhang *et al.*, 2008). Other genes related to immune response are the *RAP1A,* which positively regulates T cells via integrin activation (Sebzda *et al.*, 2002), and the gene *EFTUD2,* that controls the alternate splicing of the myeloid differentiation primary response 88 (*MyD88*) innate immunity signaling adaptor to modulate the extent of the innate immune response (Arras *et al.*, 2014).

Another gene with a great number of predicted interactions and that interacts with the *CDC42* and *RAP1A* genes (Figure 5) is the *RHOC,* which is important in protrusion formation in migrating cells, and was also identified as acting in the response to heat stress in a chicken hepatocellular carcinoma cell line (Sun *et al.*, 2015). Other 31 genes located in the ROH regions in our population (Supplementary Table S2) were also identified by Sun *et al.* (2015) as differentially expressed in their transcriptome analysis in response to heat stress in chicken hepatocellular carcinoma cell line. Out of those, the solute carrier family 39 member 6, BCL2 associated athanogene 1, NOP2/Sun RNA methyltransferase family member 2, molybdenum cofactor sulfurase and elongator acetyltransferase complex subunit 2 are located within the longest ROH identified in our population (GGA2: 77 084 081 to 88 808 596 bp). These results reinforce the hypothesis that the ancestors of the studied population had to adapt to the high environmental temperatures.

In the most frequent ROH of the studied population (GGA2: 69 807 237 to 75 014 591 bp), four coding genes of the Cadherin protein family, the *CDH9*, *CDH10*, *CDH12* and *CDH18*, involved in the skeletal muscle differentiation were identified (Donalies *et al.*, 1991). In special, the *CDH12* gene was associated with heat stress in chicken cell lines (Sun *et al.*, 2015). Several quantitative trait loci (QTL) related to production traits curated in the Chicken QTL database have already been reported in this ROH segment. The most important traits were breast muscle weight (QTL: 1904), BW at 42 days of age (QTL: 6899) and troponin T concentration (QTL: 1906). Probably, the candidate genes observed in this frequent ROH region are involved with breast and BW, which are important selection criteria applied in our population. Figueiredo *et al.* (2003), evaluating the genetic gain of the TT population during 15 generations of selection, found an improvement in BW of 758 g and in average breast area of 31.95 $cm^2$. In summary, the results showed that the genome of this broiler line had undergone a strong homogenizing effect at loci related to important production traits under selection, and possibly improving their adaptability to survive in new breeding environments.

## Conclusion

The use of ROH to evaluate selective pressure in chicken populations can provide new insights into the population history and the discovery of genomic regions and candidate genes involved in the expression of adaptive and production traits. It can be highlighted that this chicken population may have had to adapt to survive in environments with high temperatures, usually found in the tropics, and intensified by the high density of broilers in breeding facilities. This adaptation led to the homozygosity of genomic regions that harbor genes involved in the maintenance of homeostasis and the immune system, favoring the tolerance in this type of environment. Also, the conserved regions identified based on ROH are possibly related to the production traits under selection in this paternal broiler line.

## Acknowledgments

The authors are grateful to Alexandre L. Tessmann for technical assistance. This study was supported by the Brazilian Agricultural Research Corporation – Embrapa (project no. 01.11.07.002.04.02), by the thematic project (2014/08704-0) from São Paulo Research Foundation (FAPESP) and by the 'Conselho Nacional de Desenvolvimento Científico e Tecnológico' (CNPq, project no. 449564/2014-2). JAPM, APS and LBJ received a scholarship from the 'Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES' and GCMM and TFG from FAPESP. DPM and LLC are CNPq fellows.

## Supplementary material

To view supplementary material for this article, please visit https://doi.org/10.1017/S1751731117002671

## References

Aengwanich W 2008. Pathological changes and the effects of ascorbic acid on lesion scores of bursa of fabricius in broilers under chronic heat stress. Research Journal of Veterinary Sciences 1, 62–66.

Andreescu C, Avendano S, Brown SR, Hassen A, Lamont SJ and Dekkers JCM 2007. Linkage disequilibrium in related breeding lines of chickens. Genetics 177, 2161–2169.

Anwar B, Khan SA, Aslam A, Maqbool A and Khan KA 2008. Effects of ascorbic acid and acetylsalicylic acid supplementation on the performance of broiler chicks exposed to heat stress. Pakistan Veterinary Journal 24, 109–112.

Arras LD, Laws R, Leach SM, Pontis K, Freedman JH, Schwartz DA and Alper S 2014. Comparative genomics RNAi screen identifies Eftud2 as a novel regulator of innate immunity. Genetics of Immunity 197, 485–496.

Badyaev AV 2005. Stress-induced variation in evolution : from behavioural plasticity to genetic assimilation. Proceedings of the Royal Society B 272, 877–886.

Balboa MA and Balsinde J 2006. Oxidative stress and arachidonic acid mobilization. Biochimica et Biophysica Acta 1761, 385–391.

Bosse M, Megens H-J, Madsen O, Paudel Y, Frantz LAF, Schook LB, Crooijmans RPMA and Groenen MAM 2012. Regions of homozygosity in the porcine genome : consequence of demography and the recombination landscape. PLoS Genetics 8, e1003100.

Carothers AD, Rudan I, Kolcic I, Polasek O, Hayward C, Wright AF, Campbell H, Teague P, Hastie ND and Weber JL 2006. Estimating human inbreeding coefficients : comparison of genealogical and marker heterozygosity approaches. Annals of Human Genetics 70, 666–676.

Curik I, Sölkner J and Stipic N 2002. Effects of models with finite loci, selection, dominance, epistasis and linkage on inbreeding coefficients based on pedigree and genotypic information. Journal of Animal Breeding and Genetics 119, 101–115.

Donalies M, Cramer M, Ringwald M and Starzinski-Powitz A 1991. Expression of M-cadherin, a member of the cadherin multigene family, correlates with differentiation of skeletal muscle cells. Proceedings of the National Academy of Sciences 88, 8024–8028.

Ferencakovic M, Hamzic E, Gredler B, Solberg TR, Klemetsdal G, Curik I and Sölkner J 2012. Estimates of autozygosity derived from runs of homozygosity : empirical evidence from selected cattle populations. Journal of Animal Breeding and Genetics 130, 286–293.

Figueiredo EAP, Rosa PS, Scheuermann GN, Jaenisch FRF, Schmidt GS, Ledur MC, Brentano L and Costa CAF 2003. Genetic gain in body weight, feed conversion, and carcass traits in White Plymouth Rock broiler strain Embrapa 021. Paper presented at the IX World Conference on Animal Production, 26–31 October, Porto Alegre, Brazil.

Fleming DS, Koltes JE, Markey AD, Schmidt CJ, Ashwell CM, Rothschild MF, Persia ME, Reecy JM and Lamont SL 2016. Genomic analysis of Ugandan and Rwandan chicken ecotypes using a 600 K genotyping array. BMC Genomics 17, 407.

Food and Agriculture Organization of the United Nations 2004. Secondary guidelines for development of national farm animal genetic resources management plans. FAO, Rome, Italy.

Fornari MB, Zanella R, Ibelli AMG, Fernandes LT, Cantão ME, Thomaz-Soccol V, Ledur MC and Peixoto JO 2014. Unraveling the associations of osteoprotegerin gene with production traits in a paternal broiler line. SpringerPlus 3, 682.

Fu W, Dekkers JCM, Lee WR and Abasht B 2015. Linkage disequilibrium in crossbred and pure line chickens. Genetics Selection Evolution 47, 1–12.

Fu W, Lee WR and Abasht B 2016. Detection of genomic signatures of recent selection in commercial broiler chickens. BMC Genetics 17, 122.

Gorostizaga A, Brion L, Maloberti P, Paz C, Maciel FC, Podestá EJ and Paz C 2005. Heat shock triggers MAPK activation and MKP-1 induction in Leydig testicular cells. Biochemical and Biophysical Research Communications 327, 23–28.

Harada Y, Tanaka Y, Terasawa M, Pieczyk M, Habiro K, Katakai T, Hanawa-Suetsugu K, Kukimoto-Niino M, Nishizaki T, Shirouzu M, Duan X, Uruno T, Nishikimi A, Sanematsu F, Yokoyama S, Stein JV and Kinashi T 2012. DOCK8 is a Cdc42 activator critical for interstitial dendritic cell migration during immune responses. Blood 119, 4451–4461.

Hayes BJ, Visscher PM, Mcpartlan HC and Goddard ME 2003. Novel multilocus measure of linkage disequilibrium to estimate past effective population size. Genome Research 13, 635–643.

Hill WG and Robertson A 1968. Linkage disequilibrium in finite populations. Theoretical and Applied Genetics 38, 226–231.

International Chicken Genome Sequencing Consortium 2004. Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. Nature 432, 695–716.

Julian RJ 2005. Production and growth related disorders and other metabolic diseases of poultry – a review. The Veterinary Journal 169, 350–369.

Ledur MC, Peixoto JO, Nones K and Coutinho LL 2012. Applied genomics: the Brazilian experience. In World's Poultry Congress, 24., 2012, Salvador Abstract. Salvador: WSPA, 2012. 1 CD-ROM. Poultry Science Journal, 68, supl. 1, 2012.

Lin H, Decuypere E and Buyse J 2006. Acute heat stress induces oxidative stress in broiler chickens. Comparative Biochemistry and Physiology 144, 11–17.

McQuillan R, Leutenegger AL, Abdel-Rahman R, Franklin CS, Pericic M, Barac-Lauc L, Smolej-Narancic N, Janicijevic B, Polasek O, Tenesa A, MacLeod AK, Farrington SM, Rudan P, Hayward C, Vitart V, Rudan I, Wild SH, Dunlop MG, Wright AF, Campbell H and Wilson JF 2008. Runs of homozygosity in European populations. The American Journal of Human Genetics 83, 359–372.

Muir WM, Wong GK-S, Zhang Y, Wang J, Groenen MAM, Crooijmans RPMA, Megens H-J, Zhang H, Okimoto R, Vereijken A, Jungerius A, Albers GAA, Lawley CT, Delany ME, MacEachern S and Cheng HH 2008. Genome-wide assessment of worldwide chicken SNP genetic diversity indicates significant absence of rare alleles in commercial breeds. Proceedings of the National Academy of Sciences 105, 17312–17317.

Mujahid A, Sato K, Akiba Y and Toyomizu M 2006. Acute heat stress stimulates mitochondrial superoxide production in broiler skeletal muscle, possibly via downregulation of uncoupling protein content. Poultry Science 85, 1259–1265.

Mujahid A, Yoshiki Y, Akiba Y, Toyomizu M 2005. Superoxide radical production in chicken skeletal muscle induced by acute heat stress. Poultry Science 84, 307–314.

Organisation for Economic Co-operation and Development (OECD)-Food and Agriculture Organization (FAO) 2015. OECD-FAO agricultural outlook 2015-2024. OECD Publishing, Paris, France.

Pamok S, Aengwanich W and Komutrin T 2009. Adaptation to oxidative stress and impact of chronic oxidative stress on immunity in heat-stressed broilers. Journal of Thermal Biology 34, 353–357.

Qanbari S, Hansen M, Weigend S, Preisinger R and Simianer H 2010. Linkage disequilibrium reveals different demographic history in egg laying chickens. BMC genetics 11, 103.

Quinteiro-Filho WM, Ribeiro A, Ferraz-de-Paula V, Pinheiro ML, Sakai M, Sá LRM, Ferreira AJP and Palermo-Neto J 2010. Heat stress impairs performance parameters, induces intestinal injury, and decreases macrophage activity in broiler chickens. Poultry Science 89, 1905–1914.

Rubin C, Zody MC, Eriksson J, Meadows JRS, Sherwood E, Webster MT, Jiang L, Ingman M, Sharpe T, Ka S, Hallbo F, Besnier F, Carlborg O, Bed'hom B, Tixier-Boichard M, Jensen P, Siegel P, Lindblad-Toh K, Andersson L 2010. Whole-genome resequencing reveals loci under selection during chicken domestication. Nature 464, 587–591.

Sargolzaei M 2014. SNP1101 User's Guide. Version 1.0 Canada: University of Guelph.

Scheele CW 1997. Pathological changes in metabolism of poultry related to increasing production levels. Veterinary Quarterly 19, 127–130.

Sebzda E, Bracke M, Tugal T, Hogg N and Cantrell DA 2002. Rap1A positively regulates T cells via integrin activation rather than inhibiting lymphocyte signaling. Nature Immunology 3, 252–258.

Sun L, Lamont SJ, Cooksey AM, Mccarthy F, Tudor CO, Derita RM, Rothschild M, Ashwell C, Persia ME and Schmidt CJ 2015. Transcriptome response to heat stress in a chicken hepatocellular carcinoma cell line. Cell Stress and Chaperones 20, 939–950.

Sved JA 1971. Linkage disequilibrium of chromosome segments. Theoretical Population Biology 2, 125–141.

Tsuji T and Miypshi M 2001. A scanning and transmission electron microscopic study of the lymphoreticular framework in the chicken fabricius' bursa. Medical Bulletin of Fukuoka University 28, 63–75.

Türkyilmaz MK 2008. The effect of stocking density on stress reaction in broiler chickens during summer. Turkish Journal of Veterinary & Animal Sciences 32, 31–36.

Zanella R, Peixoto JO, Cardoso FF, Cardoso LL, Biegelmeyer P, Cantao ME, Otaviano A, Freitas MS, Caetano AR and Ledur MC 2016. Genetic diversity analysis of two commercial breeds of pigs using genomic and pedigree data. Genetics Selection Evolution 48, 24.

Zhang Q, Guldbrandtsen B, Bosse M, Lund MS and Sahana G 2015. Runs of homozygosity and distribution of functional variants in the cattle genome. BMC Genomics 16, 542.

Zhang S, Lillehoj HS, Kim C-H, Keeler CL, Babu JU and Zhang MZ 2008. Transcriptional response of chicken macrophages to *Salmonella enterica* serovar enteritidis infection. Animal Genomics for Animal Health 132, 141–151.

Zhu M, Zhu B, Wang YH, Wu Y, Xu L, Guo LP, Yuan ZR, Zhang LP, Gao X, Gao HJ, Xu SZ and Li JY 2013. Linkage disequilibrium estimation of Chinese beef simmental cattle using high-density SNP panels. Asian-Australasian Journal of Animal Sciences 26, 772–779.