

RESEARCH ARTICLE

Open Access



Detection of genomic signatures of recent selection in commercial broiler chickens

Weixuan Fu¹, William R Lee² and Behnam Abasht^{1*}

Abstract

Background: Identification of the genomic signatures of recent selection may help uncover causal polymorphisms controlling traits relevant to recent decades of selective breeding in livestock. In this study, we aimed at detecting signatures of recent selection in commercial broiler chickens using genotype information from single nucleotide polymorphisms (SNPs). A total of 565 chickens from five commercial purebred lines, including three broiler sire (male) lines and two broiler dam (female) lines, were genotyped using the 60K SNP Illumina iSelect chicken array. To detect genomic signatures of recent selection, we applied two methods based on population comparison, cross-population extended haplotype homozygosity (XP-EHH) and cross-population composite likelihood ratio (XP-CLR), and further analyzed the results to find genomic regions under recent selection in multiple purebred lines.

Results: A total of 321 candidate selection regions spanning approximately 1.45 % of the chicken genome in each line were detected by consensus of results of both XP-EHH and XP-CLR methods. To minimize false discovery due to genetic drift, only 42 of the candidate selection regions that were shared by 2 or more purebred lines were considered as high-confidence selection regions in the study. Of these 42 regions, 20 were 50 kb or less while 4 regions were larger than 0.5 Mb. In total, 91 genes could be found in the 42 regions, among which 19 regions contained only 1 or 2 genes, and 9 regions were located at gene deserts.

Conclusions: Our results provide a genome-wide scan of recent selection signatures in five purebred lines of commercial broiler chickens. We found several candidate genes for recent selection in multiple lines, such as *SOX6* (*Sex Determining Region Y-Box 6*) and *cTR* (*Thyroid hormone receptor beta*). These genes may have been under recent selection due to their essential roles in growth, development and reproduction in chickens. Furthermore, our results suggest that in some candidate regions, the same or opposite alleles have been under recent selection in multiple lines. Most of the candidate genes in the selection regions are novel, and as such they should be of great interest for future research into the genetic architecture of traits relevant to modern broiler breeding.

Keyword: Chickens, Selection signatures, Breeding, Commercial broilers

Abbreviations: CLR, Composite likelihood ratio; EHH, Extended haplotype homozygosity; iHH, Integrated extended haplotype homozygosity; iHS, Integrated haplotype score; LD, Linkage disequilibrium; MAF, Minor allele frequency; QTL, Quantitative trait loci; XP-CLR, Cross-population composite likelihood ratio; XP-EHH, Cross-population extended haplotype homozygosity

* Correspondence: abasht@udel.edu

¹Department of Animal and Food Sciences, University of Delaware, Newark, DE 19716, USA

Full list of author information is available at the end of the article



Background

Artificial selection is the primary factor in the domestication and breeding history of livestock species. Modern broiler (meat-type) chickens have been under strong artificial selection, mostly for traits of economic importance for farmers, such as growth rate, feed efficiency and body composition [1]. Comparing a modern broiler chicken cross, Ross 308, with a broiler population that had not been subjected to artificial selection since 1957 [i.e., Athens-Canadian Random-bred Control (ACRBC) strain], Havenstein *et al.* (2003) found that the average body weight at 42 days of age increased from 539 g in 1957 as represented by the ACRBC strain to 2,672 g in 2001 as represented by the Ross 308 strain and that the feed conversion ratio decreased from 2.34 to 1.43 over the same time period [2]. The authors indicated that genetic selection contributed 85–90 % of the improvement in growth rate over the past 45 years. These dramatic phenotypic changes imply that the frequencies of the underlying causal polymorphisms themselves have been altered by selection for performance in these traits during the intervening time period. Thus, detecting the genomic footprints of artificial selection should help researchers to identify the causal polymorphisms underlying phenotypic changes and to better understand the biological and genetic mechanism controlling these traits.

In broiler chicken genetic stocks, the traits of most relevance for recent decades of breeding, such as feed efficiency, growth rate and meat yield, are complex traits that are controlled by many genes. Consequently, it is highly likely that selection for these traits has worked simultaneously on multiple causal genes across the genome. Therefore, high throughput methods are required to screen the whole genome for signatures of recent selection. With the availability of high throughput genotyping tools, such as high-density SNP arrays and next-generation sequencing, it has become possible to conduct genome-wide studies for detection of genomic footprints of artificial selection.

Using whole-genome re-sequencing and high-density SNP chips, respectively, Rubin *et al.* (2010) and Elferink *et al.* (2012) have investigated selection signatures in large numbers of chicken breeds using Z-transformed pooled heterozygosity (ZHp) scores. This statistic estimates local heterozygosity depression in chromosomal regions [3, 4] and has been appropriately applied for detecting alleles that have swept to fixation or near-fixation by long-term directional selection or during domestication [5]. However, modern broiler chicken breeding practices that have a more recent selection history, and have been employed to select for a suite of more sophisticated traits, such as feed efficiency and meat yield with different selection priorities in different specialized component lines, would not be expected to leave such common and distinct changes.

Therefore, most signatures of more recent selection are likely yet to be uncovered in the genome of modern broiler chickens.

In contrast with the ZHp method, it has been suggested that methods based on extended haplotype homozygosity (EHH) [6] or change in allele frequency spectrum can be more useful for detecting signatures of recent selection in animal breeds [5, 7]. These population genetics methods are developed to find frequent alleles and long-range haplotypes with high frequency, which are indicative of chromosomal regions under recent selection [6]. In dairy cattle, Qanbari *et al.* (2010) adopted the relative EHH method to detect signatures of positive selection in Holstein–Friesian cattle using a 50K SNP array [7]. Zhang *et al.* (2012) applied the same method to detect selection signatures in two broiler chicken lines divergently selected for abdominal fat content and reported the *PC1/PCSK1* region as the most likely candidate region to have a causal effect on abdominal fat weight [8, 9]. To detect selection signatures in Fleckvieh cattle, Qanbari *et al.* (2014) applied two statistical methods, the integrated haplotype score (iHS) [10] and the composite likelihood ratio (CLR) [11, 12], and found that many candidate regions were relevant to coat coloring pattern, neurobehavioral functioning and sensory perception [13].

Concerned that the EHH and iHS methods may have insufficient power to identify the alleles that have been more recently strongly selected and swept to near-fixation or fixation [14], Sabeti *et al.* (2007) developed the cross-population extended haplotype homozygosity (XP-EHH) test to detect signatures of recent selection by comparing EHH between two different populations regardless of whether or not the favored allele had reached fixation. Also, the single-population CLR method does not take advantage of larger differences in allele frequencies between two breeds and is very sensitive to SNP ascertainment bias. To overcome these limitations, Chen *et al.* (2010) developed the cross-population composite likelihood ratio (XP-CLR) test [15].

In this study, we applied both XP-EHH and XP-CLR methods to five commercial broiler purebred populations, including three male lines and two female lines, to detect the signatures of recent selection in these commercial broiler stocks. The findings here help to improve our understanding of the biological mechanisms controlling economically important traits in modern commercial broiler chickens.

Methods

Animals and data preparation

A total of 565 chickens from five commercial purebred lines were genotyped using the 60K SNP Illumina iSelect chicken array [16]. Blood samples were collected from a wing vein and the samples were shipped on dry ice to

DNA LandMarks (Saint-Jean-sur-Richelieu, Quebec, Canada) for DNA extraction and genotyping with the 60K SNP Illumina iSelect chicken array. All genotyped birds were males and were sampled from 3 male (broiler sire) lines, ML1, ML2 and ML3, and 2 female (broiler dam) lines, FL1 and FL2. In total, 318 birds were sampled from male lines: 24 ML1, 256 ML2 and 38 ML3 chickens; and 247 birds were sampled from female lines: 126 FL1 and 121 FL2 chickens. The FL1, FL2 and ML1 chickens as well as a portion of ML2 (ML2_0; $n = 96$) chickens were elite sires randomly sampled from three overlapping generations. Another portion of ML2 genotyped chickens (ML2_1; $n = 160$) was a random sample of the progeny of the ML2_0 elite sires. The ML3 genotyped chickens were random samples of male chickens from this purebred population. Male and female lines originated from different breeds, i.e. the male line from Cornish, a meat type breed, and the female lines from White Rock, a dual-purpose breed. Each of these five lines came from a different source to Heritage Breeders, and all lines, except FL1, have been reproductively isolated for more than 40 generations. A one-time crossbreeding with ML2 and then backcrossing with FL1 happened early in the history of FL1, and the resulting new FL1 population has been reproductively isolated for more than 25 generations. In each generation within each purebred line, approximately 50 to 80 male and 500 to 800 female birds have been selected for reproducing the next generation. More details about genetic diversity in ML2, FL1 and FL2 can be found in a previous study, where the three lines were labeled as B, C and D, respectively [17]. Since allele frequency of SNPs and linkage disequilibrium (LD) among SNPs were highly consistent between the two sampled generations of ML2, ML2_0 and ML2_1 (unpublished data in our laboratory) and the methods we used for detecting recent selection signatures relied on allele frequency and LD information, we combined data from ML2_0 and ML2_1 in the current study.

Although male lines (ML1, ML2 and ML3) have shared similar selection objectives, which have been primarily focused on increasing growth rate, feed efficiency and breast muscle yield, the relative magnitude of selection pressure on these major traits varied among the three lines: ML1 has been more heavily selected for rapid growth, ML2 for high breast meat yield and ML3 for improved feed efficiency. Therefore, we expect artificial selection has unequally increased the frequency of alleles controlling these traits among the male lines, and some alleles may have been selected in opposite direction between male and female lines. The expected differences in the allele frequency among these purebred lines make it possible to apply the two population comparison based methods, XP-EHH and XP-CLR, in our study.

The 60K SNP Illumina iSelect chicken array contains a total of 57,636 SNPs [16]. For the purpose of this study, we used only SNPs with assigned positions on the current chicken genome based on the latest reference genome (*Gallus gallus* 4.0 UCSC, May 2012). We excluded SNPs with a call rate $< 90\%$ or Mendelian inconsistency > 0.001 and SNPs that were monomorphic among all the purebred lines. We also excluded SNPs on chromosomes 16 and W and two linkage groups, as there were too few SNPs in the 60K SNP Illumina iSelect chicken array for these chromosomes. After quality control, 48,950 SNPs were used in subsequent analyses of the five populations (Table 1).

Since a linkage map was required for the XP-CLR method, we calculated the genetic positions of all the markers in the 60K SNP Illumina iSelect chicken array using a subset of markers with known genetic positions based on the male linkage map previously provided by Groenen *et al.* [18]), and assuming that the recombination rates between two markers were uniformly distributed. We used BEAGLE (Version 3.3.2) [19] to impute missing genotypes, phase the chromosomes and identify haplotype structure at the candidate selection regions in each purebred line.

The XP-EHH test

The XP-EHH test uses the integrated EHH (iHH) of a core SNP in two populations, A and B, rather than two alleles in a single population. The unstandardized XP-EHH statistic can be calculated as [14]:

$$\text{unstandardizedXP-EHH} = \ln \left(\frac{iHH_A}{iHH_B} \right) \quad (1)$$

where iHH_A and iHH_B are the integrated EHH of a given core SNP in population A and B, respectively. A large positive value of XP-EHH suggests either selection in population A or a negative value in population B.

We used the software developed by Pickrell *et al.* [20] to estimate unstandardized XP-EHH statistics for all SNPs (after quality control) in all five purebred lines with cross-population comparison of each purebred line with the four remaining lines: for example, ML1 vs. ML2, ML3, FL1 or FL2 (four cross-population tests for each line). The unstandardized XP-EHH statistics were

Table 1 Quality control of genotype data

Total SNPs	Quality control					SNPs used
	NI ¹	MI ²	UG ³	MO ⁴	LCR ⁵	
57,636	1,507	1,478	873	4,292	536	48,950

Note: ¹SNPs on GGA16, W and two linkage groups (LGE22C19W28_E50C23 and LEG64) or SNPs with unknown positions on Galgal4; ²SNPs with Mendelian inconsistency; ³Ungenotyped SNPs; ⁴SNPs monomorphic among all the 5 purebred lines; ⁵SNPs with low call rate

standardized using their means and variances in each purebred comparison. Because previous studies found that the standardized XP-EHH statistics follows the standard normal distribution [14, 21, 22], *P*-values of SNPs were estimated using the standard normal distribution. For each purebred comparison, we determined the candidate regions under positive selection by clustering the significant core SNPs (*P*-value < 0.05) with a distance of less than 200 kb.

The XP-CLR test

To confirm selection signatures detected by the XP-EHH analysis, we applied the XP-CLR test based on the change in the allele frequency spectrum, since it has the advantage of enlarging signals to allow the resolution of more precise regions [15]. The XP-CLR test [15] was also adopted for the five purebred lines by cross-population comparison of each line with the four remaining lines as reference populations using the XP-CLR 1.0 software available at http://genetics.med.harvard.edu/reich/Reich_Lab/Software.html (last accessed Jan. 3, 2016). The grid points at the putative selected allele positions were set along each chicken chromosome with a spacing of 2 kb, and sliding window size was set as 0.5 cM around the grid points. To reduce the contribution of SNPs in high LD to the likelihood function, the cut-off level of absolute pairwise correlation coefficient of two SNPs was set to 0.9 for estimation of the weight factor (*w* [15]).

For each cross-population comparison, the cutoff threshold of 0.5 % XP-CLR scores was applied to determine windows with strong signals across the whole genome. We then determined the candidate selection regions by clustering these windows, such that windows with genetic distances less than 1 cM constitute a candidate selection region. The selection regions detected by both statistical methods for each purebred line were determined as candidate regions under positive selection. A Karyogram layout of candidate selection regions detected by both tests was created using the ggbio R package [23].

In comparison with human populations, modern livestock breeds generally have much smaller effective population sizes due to animal breeding programs [24–26]. To minimize false discovery due to genetic drift resulted from the small effective population size, the candidate selection regions shared by multiple purebred lines were considered as high-confidence selection regions, and these regions were chosen to identify candidate genes using the genomic database search engine BioMart (<http://www.biomart.org/>).

Results

In total, 1,079 putative selection regions were detected with *P*-values < 0.05 using XP-EHH test (Additional file 1: Table S1), and 1,018 putative selection regions were detected using the criterion of a 0.5 % cutoff of XP-CLR

scores (Additional file 2: Table S2). Regions detected using XP-EHH overlapped 31.53 % of the regions that were identified using XP-CLR (Additional file 3: Table S3). Even though 328 overlapped regions (i.e., detected by both methods) were presented on Additional file 3: Table S3, some regions detected by either the XP-EHH or XP-CLR tests were wide enough to overlap with more than one region detected by the other test. Therefore, in total, 224 and 321 unique regions were detected using XP-EHH and XP-CLR tests, respectively. In each line, approximately 11.09 % of the chicken genome was covered by regions detected by XP-EHH methods, while approximately 2.58 % of the chicken genome was covered by regions detected by XP-CLR methods. The overlapped regions (shared by both methods) only represented approximately 1.45 % of the chicken genome in each line. Selection regions detected by XP-EHH were much wider, mainly because the EHH test is an LD-based method, and LD is expected to extend over longer distances in regions under recent selection [17]. For example, Fig. 2a and 2b represent the results of XP-EHH and XP-CLR tests on GGA5, which show candidate regions detected by XP-CLR tests were overall narrower and perhaps more accurate than those detected by the XP-EHH tests. Thus, to narrow down regions that overlapped between the two methods, we considered the 321 regions based on the XP-CLR test as the candidate selection regions. Their ranges are presented in Additional file 3: Table S3 and visualized in Fig. 1.

By further examining these 321 regions, we identified 42 regions that were shared by two or more purebred lines (Additional file 4: Table S4) and considered them as high-confidence selection regions. Figure 2c and 2d represent the results of XP-EHH and XP-CLR tests in one of the 42 regions (GGA5: 31.06–31.82 Mb) shared by ML1 and ML3. To further narrow down these high-confidence selection regions, only common regions shared by two or more purebred lines were counted in the overlapped regions. Of these 42 common regions, 20 were 50 kb or less while 4 regions were larger than 0.5 Mb. Using BioMart, 91 genes could be found in the 42 regions (Additional file 4: Table S4) among which 9 regions were located at gene deserts and 19 regions only harbored 1 or 2 genes. For the 9 regions located at gene deserts, the genes closest to them (± 100 kb) are listed on Additional file 4: Table S4.

To gain insight into population differences in the overlapped candidate regions, we constructed haplotypes and estimated haplotype frequencies in these regions in each population (Additional file 5: Table S5). This analysis was performed only for the high-confidence selection regions that contained at least 5 informative SNPs in our genotype data (14 out of 42 regions). Figure 3 represents the results of haplotype analysis in four selection regions containing 10 to 20 SNPs in our genotype data. As

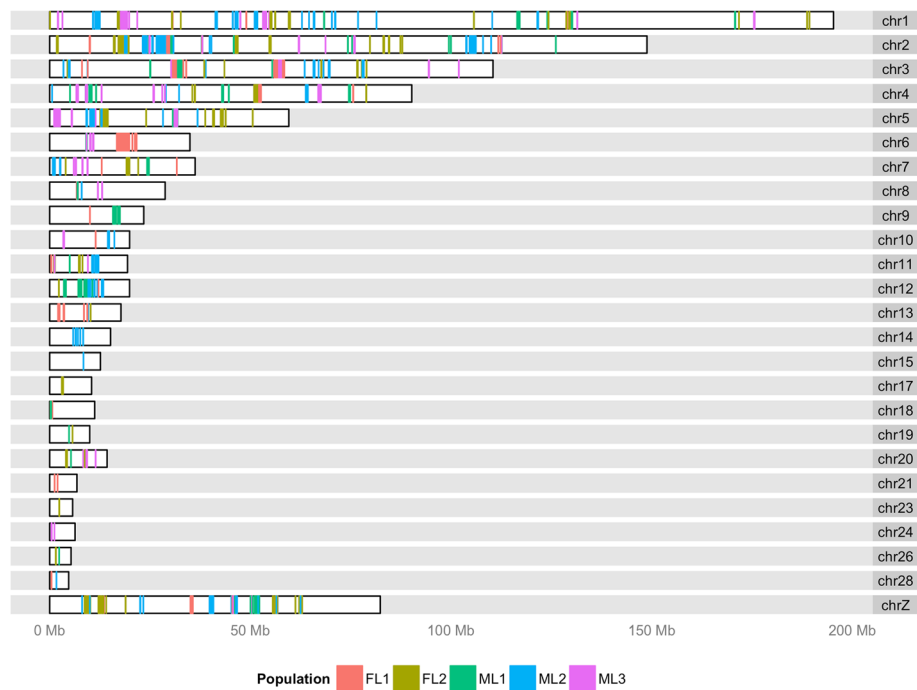


Fig. 1 Candidate selection regions detected by XP-EHH and XP-CLR tests. For each purebred line, the overlapped regions detected by the XP-EHH and XP-CLR tests were presented based on the ranges from XP-CLR test. Each population is denoted by a different color

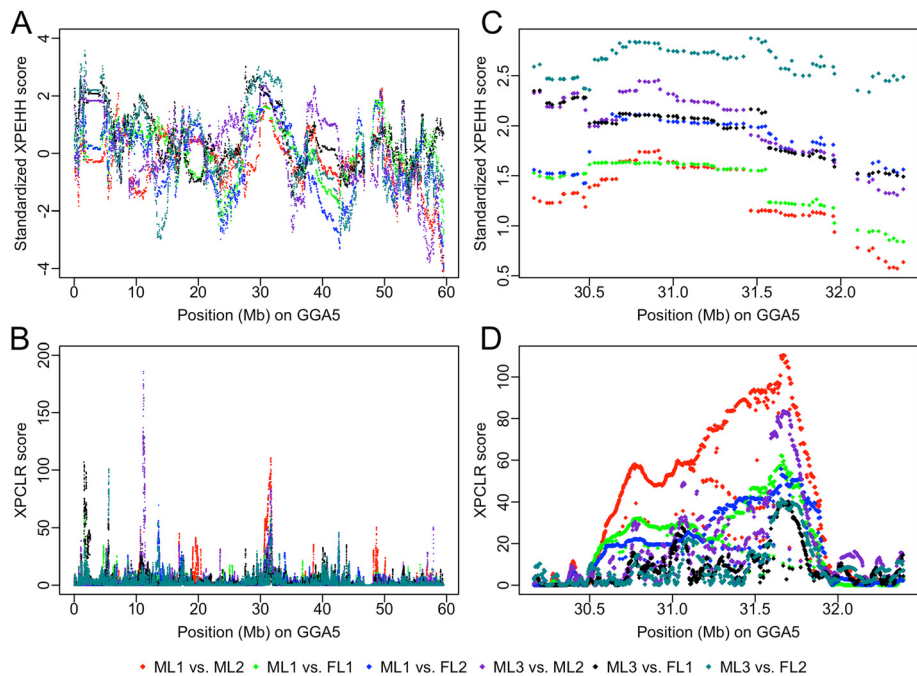


Fig. 2 XP-EHH and XP-CLR scores on GGA5. **A** and **B**: The results of the XP-EHH (**a**) and XP-CLR (**b**) statistics on the whole GGA5 using multiple population comparisons. **c** and **d**: The results of the XP-EHH (**a**) and XP-CLR (**b**) statistics in a candidate selection region (GGA5: 31.06-31.82 Mb) shared by ML1 and ML3. The dots in the Fig. 2b and 2d represent the XP-CLR scores of sliding windows, and the dots in the Fig. 2a and 2c represent the standardized XP-EHH scores of SNPs. Each comparison of ML1 or ML3 against the other 3 lines is denoted by a different color

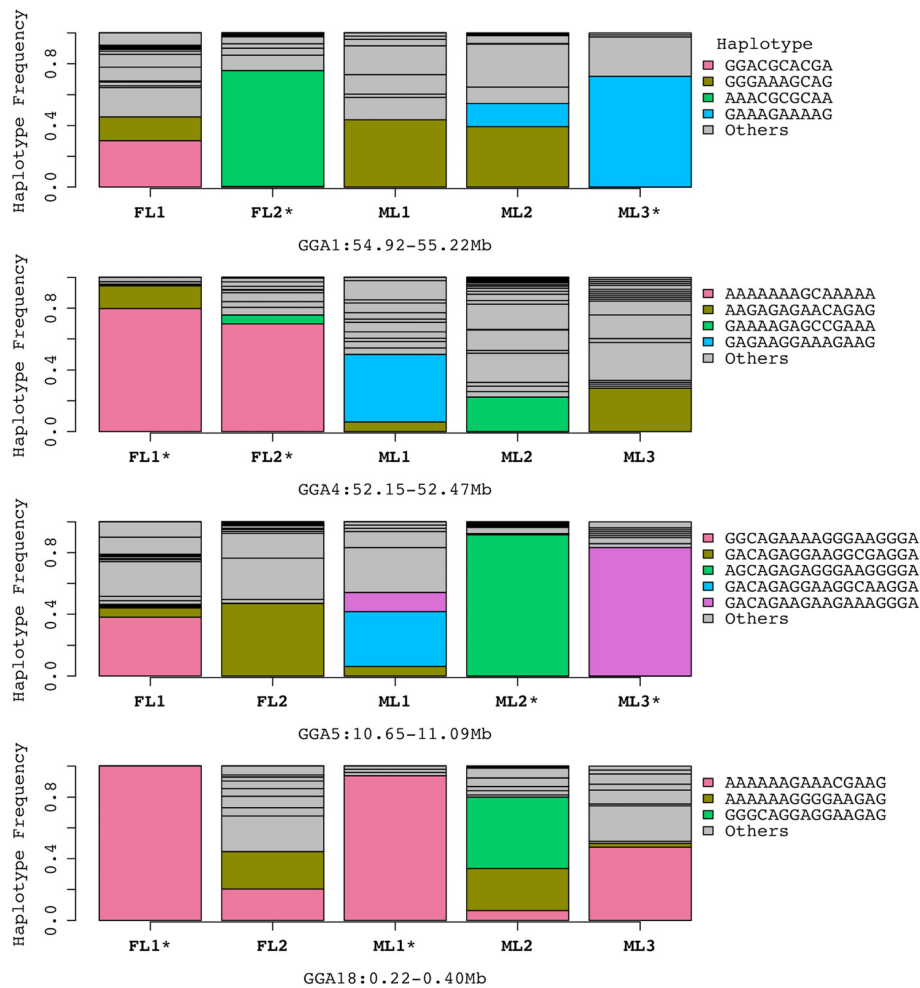


Fig. 3 Haplotype frequencies of SNPs in four selection regions detected in multiple purebred lines of broiler chickens. The various colors excluding grey refer to the major haplotypes identified in the 5 purebred lines. At each denoted region, a selection signature was detected in the purebred lines marked with “**”

demonstrated in Fig. 3, haplotypes with high frequencies were detected in the denoted genomic regions. For example, in a selection region on GGA4 (52.15-52.47 Mb), the same haplotype showed high frequency in FL1 and FL2, although the range of this region was more than 300 kb. Another interesting example is a ~240 kb region on GGAZ (45.49-45.73 Mb). In this region, all three male lines had the same major haplotype, but each female line had a different major haplotype (Additional file 5: Table S5).

Discussion

In modern broiler breeding, the practice of selective mating is utilized to influence the expression of economically important traits in subsequent generations. Through such selection, the “beneficial” alleles tend to become more frequent in populations over time. In our study, we applied XP-EHH and XP-CLR tests to detect the genomic regions under recent selection by measuring the characteristics of

extended haplotype homozygosity and changes in the allele frequency spectrum. By cross-population comparisons of five commercial broiler purebred lines, we identified the genomic regions that are most likely to harbor genes related to traits of economic importance in broiler chickens.

It should be mentioned that both bottleneck events and genetic drift have potential to influence the results of selection signature studies such as this one. Based on records from Heritage Breeders, bottleneck events did not occur in the lines used in the present study for more than 40 generations. To minimize false discovery due to genetic drift, only 42 of the candidate selection regions that were shared by 2 or more purebred lines were considered as high-confidence selection regions in our study. Another potential limitation is inherent to cross-population methods, which may fail to detect a selection signature where the desirable allele has been under similar level of positive or negative selection pressure in all

these purebred lines that were studied. However, this limitation should not be a major concern because the relative magnitude of selection pressure on major traits, growth rate, feed efficiency and breast muscle yield, varied among the 5 purebred lines. Also, unlike in male lines, selection for reproduction traits has been emphasized in female lines.

Candidate selection regions

We compared the genes in the candidate selection regions in our study with those from two previous studies on detecting selection sweeps in chickens. Among 91 genes in the 42 regions in our study, only two genes (*SOX6* and *GJD2*) are in genomic regions detected by Rubin *et al.* (2010) in commercial broilers. Also, only two genes in our list (*GAS7* and *STXBP6*; Additional file 4: Table S4) are among 366 genes (based on Ensembl gene ID) detected by Elferink *et al.* (2012). This low extent of overlap with previous studies is likely related to the different methods that we used for detecting selection signatures in the present study. As mentioned before, we aimed at detecting signatures of recent selection using the cross-population methods, whereas the ZHp method used in two previous studies is primarily focused on detecting older selection signatures such as those accumulated during domestication. For better comparison, we estimated ZH scores over sliding 5-marker windows on autosomes using data from our study (Additional file 6: Supplemental file). This analysis led to the detection of 41 selection regions containing 81 genes, including 31 genes detected in broilers from two previous studies (Additional file 7: Table S7 and Additional file 8: Table S8) although our resource populations were much different from those in the two previous studies. The most possible reason of high overlap using ZH scores is that some older selection signatures during chicken domestication were shared with commercial broilers used in our study as well as two previous studies.

In our results from two cross-population methods (XP-EHH and XP-CLR), 42 regions were detected by both methods in multiple populations, which might indicate that gene(s) in these regions have been independently selected in multiple populations, i.e., parallel selection. Of these 42 regions, 13 regions were shared among male lines, 5 regions were shared among female lines and 24 regions were shared among both male and female lines. In addition to the significant overlap between the suites of selected traits among all lines selected for broiler performance (especially for growth rate, meat yield, and feed efficiency), the shared selection regions among male and female lines suggests that alleles with pleiotropic effects have been under recent selection in these regions, i.e., alleles that are positively correlated with growth related traits and negatively with reproduction traits, or vice versa. Alternatively, the regions shared by male and female lines

may contain closely linked genes impacting both growth and reproduction traits. To further analyze the high-confidence selection regions, we examined population differences in haplotype structure and frequencies and identified the major haplotype (the haplotype with the highest frequency) within each population at each denoted region. The results showed that, in some candidate selection regions, the same major haplotype is shared by multiple lines (Fig. 3 and Additional file 5: Table S5). However, in 9 out of 14 candidate selection regions presented in Additional file 5: Table S5, such as GGA1: 54.92-55.22 Mb and GGA13: 9.38-9.44 Mb, the major haplotype varied greatly among the five purebred lines. The difference in major haplotype may represent high diversity of genetic background among these purebred lines [17]. Alternatively, it is possible that selection has acted on different alleles of a gene in these purebred lines. For example, previous studies found that fertility was reduced in chickens under strong selection for body weight due to the negative genetic correlation between reproduction and growth traits [27–29]. Overall, selection for reproduction traits has been more emphasized in female lines, whereas selection for high feed-efficiency and increased skeletal muscle growth has been the major focus in male lines. Thus, the frequency of alleles benefiting reproduction traits but adversely affecting growth traits are expected to be relatively higher in female lines as compared with the male lines. Some candidate genes with potential pleiotropic effects, such as *STXBP6* and *cTR*, in the high-confidence selection regions are discussed below.

Candidate genes in regions detected in multiple populations

In the 42 high-confidence candidate selection regions detected by both methods (XP-EHH and XP-CLR) in two or more purebred lines, we identified several genes related to growth, development, feed efficiency and reproduction in chickens (Table 2). Only a few of them are mentioned below to discuss their potential involvements in controlling these traits.

Myosin heavy chain 13 (MYH13)

MYH13 is located in a candidate selection region on GGA18 (0.22-0.40 Mb), and other four genes of the *myosin heavy chain* (MyHC) family (*MYH1A*, *MYH1B*, *MYH1C*, *MYH1E*) are located very close to this region. Previous studies found that MyHC genes play important roles in skeletal muscle development [30–32], and the polymorphisms in *MYH3* were significantly associated with growth and body composition traits in Qinchuan cattle [33, 34].

Table 2 A partial list of candidate genes in or near the 42 high-confidence selection regions detected in multiple purebred lines of broiler chickens

Gene name	Gene symbol	Function or association
<i>Thyroid hormone receptor beta</i>	<i>cTR</i>	Growth, development and homeostasis
<i>Sex Determining Region Y-Box 6</i>	<i>SOX6</i>	Development of chondrocytes and skeletal muscle
<i>Actin, alpha, cardiac muscle 1</i>	<i>ACTC1</i>	Muscle development
<i>Syntaxin binding protein 6</i>	<i>STXBP6</i>	Bone allocation and fecundity traits
<i>Myosin heavy chain 13</i>	<i>MYH13</i>	Skeletal muscle development
<i>Calpastatin</i>	<i>CAST*</i>	Growth and meat quality

Note: *this gene is located close to a candidate selection region detected in a gene desert area on GGAZ. Information about chromosomal locations of the 42 candidate selection regions and the full list of candidate genes in or near these regions can be found in Additional file 4: Table S4

Sex determining region Y-Box 6 (*SOX6*)

SOX6 is located in a candidate selection region (GGA5: 10.65–11.09 Mb) detected in two male lines, ML2 and ML3. This gene encodes a Sry-related transcription factor that promotes early chondroblast differentiation and plays a critical role in differentiation and proliferation of chondrocytes as well as normal fiber type differentiation of fetal skeletal muscle in mice [35–37].

Proprotein convertase subtilisin/kexin type 1 (*PCSK1*) and calpastatin (*CAST*)

Although no gene was found inside a candidate selection region around 56.76 Mb on GGAZ due to its small size (8 kb), this region is consistent with findings from two previous studies [8, 9] in which a selection signature was detected using two chicken lines divergently selected for abdominal fat content for 11 generations. Of note, a previously known candidate gene for fatness in chickens, *PCSK1* [8], is located close to this region. Another gene close to this region is *CAST*, which encodes calpastatin, a specific inhibitor of an endogenous calpain. The calpain family plays an important role in embryonic development and muscle growth [38–40]. Many studies have found that polymorphisms in *CAST* are significantly associated with growth traits and meat quality traits in livestock animals [41–46].

Actin, alpha, cardiac muscle 1 (*ACTC1*) and Syntaxin binding protein 6 (*STXBP6*)

Another muscle-related gene, *ACTC1*, was found in one of the selection regions on GGA5 (31.06–31.82 Mb, Fig. 3). This gene encodes cardiac muscle alpha actin in chickens and plays an important role in fetal development as well as cell survival, differentiation and development of muscle [47–50]. *STXBP6* is another gene in this candidate selection region on GGA5. A previous study has indicated *STXBP6* had potential pleiotropic effect on bone tissue and fecundity traits in chickens [51]. Interestingly, this selection sweep, which was detected in two male lines (ML1 and ML3), was also found in a previous study in

table egg layer breeds of chickens [52]. One possible reason why this selection sweep is shared by broiler (meat-type) and layer (egg-type) chickens is that both genes, *ACTC1* and *STXBP6*, may influence body weight in broilers and in layers. It should be mentioned that breeders improved meat production in broilers by selection on high body weight at an early age (<8 weeks of age) while they improved feed efficiency and egg production in layers by selection on low body weight at a late age (>24 weeks of age) [53–57]. Alternatively, this shared selection sweep may be explained by the pleiotropic effect of *STXBP6* on both bone tissue and fecundity traits.

Thyroid hormone receptor beta (*cTR*)

cTR was found in a selection region on GGA2 (37.90–38.07 Mb). Thyroid hormone can regulate animal growth, development and homeostasis [58], and its receptor mediates thyroid hormone actions [59]. Mice with homozygous mutant *cTR* gene manifest low weight gain and decreased bone development compared to normal mice [60]. In a ~40 kb-length candidate selection region (GGA2: 38.03–38.07 Mb) which was detected in 4 purebred lines (FL1, FL2, ML2 and ML3), there are 3 SNPs in our dataset, which construct the same major haplotype (AAA) in two female lines and ML2, but the major haplotype (GGG) in ML3 is completely different. It should be mentioned that among the five purebred lines of chickens used in our study, ML3 is the most feed-efficient line. It is possible that in this candidate region, the same allele of *cTR* has been selected in two female lines and ML2 while an alternative allele has been selected in ML3. This assumption may be further supported considering diverse functions of thyroid hormone: it has been reported that thyroid hormone also plays a critical role in fertility, but excessive amounts of this hormone in hyperthyroidism has a negative effect on reproduction in humans [61, 62]. Therefore, the pleiotropic effects of thyroid hormone on reproduction and growth traits may explain why the receptor gene, *cTR*, may have been under selection among both female and male lines.

Conclusions

In this study, we identified novel candidate regions for recent selection in broiler chickens. Based on the biological function of genes in the candidate regions, several genes, such as *SOX6* and *cTR*, have possibly made large contributions to economically important traits in chickens. Our findings suggest that recent selection in broiler breeding has had large impact on frequency of genes controlling economically important traits, such as weight gain, muscle mass, feed efficiency and reproduction. Finally, since most of the candidate genes identified in the present study are novel and have probably been under recent selection, they should be of great interest for future research into the genetic architecture of traits relevant to modern broiler breeding.

Additional files

Additional file 1: Table S1. Putative selection regions detected by XP-EHH test (XLSX 99 kb)

Additional file 2: Table S2. Putative selection regions detected by XP-CLR test (XLSX 92 kb)

Additional file 3: Table S3. Candidate selection regions detected by both XP-EHH and XP-CLR tests (XLSX 73 kb)

Additional file 4: Table S4. High-confidence selection regions detected by both methods (XP-EHH and XP-CLR) in 2 or more populations (XLSX 41 kb)

Additional file 5: Table S5. Major haplotypes and their frequencies detected in the purebred lines at the high-confidence selection regions that span at least 5 SNPs in our genotype data (XLSX 12 kb)

Additional file 6: Methods and results of detecting selection signatures using ZH scores. (DOCX 32 kb)

Additional file 7: Table S7. Candidate selection regions on autosomes detected using ZH scores (XLSX 33 kb)

Additional file 8: Table S8. Genes in regions detected using ZH scores that overlap with previous studies (XLSX 34 kb)

Acknowledgments

We would like to thank Heritage Breeders for providing SNP genotype data; and USDA Chicken GWMAS Consortium, Cobb Vantress, and Hendrix Genetics for access to the developed 60K SNP Illumina iSelect chicken array.

Funding

The study was financially supported by the University of Delaware.

Availability of data and materials

The datasets supporting the conclusions of this article are included within the article and its additional files.

Authors' contributions

WF and BA contributed to the preparation of the manuscript and to the scientific discussions. WF contributed to the genetic analyses of the study, BA conceived of the study, and WL and BA participated in its management, overall design and coordination. WL participated in revising the manuscript. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Consent for publication

Not applicable.

Ethics approval and consent to participate

The University of Delaware Agricultural Animal Care and Use Committee approved the animal protocol used for this scientific study.

Author details

¹Department of Animal and Food Sciences, University of Delaware, Newark, DE 19716, USA. ²Maple Leaf Farms, Inc, Leesburg, IN 46538, USA.

Received: 10 March 2016 Accepted: 22 August 2016

Published online: 26 August 2016

References

- Crawford RD. Poultry Breeding and Genetics. New York: Elsevier Science Publishing Company Inc.; 1990
- Havenstein GB, Ferret PR, Qureshi MA. Growth, livability, and feed conversion of 1957 versus 2001 broilers when fed representative 1957 and 2001 broiler diets. *Poult Sci*. 2003;82:1500–8.
- Rubin C-J, Zody MC, Eriksson J, Meadows JRS, Sherwood E, Webster MT, Jiang L, Ingman M, Sharpe T, Ka S, Hallböök F, Besnier F, Carlborg O, Bed'hom B, Tixier-Boichard M, Jensen P, Siegel P, Lindblad-Toh K, Andersson L. Whole-genome resequencing reveals loci under selection during chicken domestication. *Nature*. 2010;464:587–91.
- Elferink MG, Megens H-J, Vereijken A, Hu X, Crooijmans RPMA, Groenen M a M. Signatures of selection in the genomes of commercial and non-commercial chicken breeds. *PLoS One*. 2012;7, e32720.
- Utsunomiya YT, Pérez O'Brien AM, Sonstegard TS, Van Tassell CP, do Carmo AS, Mészáros G, Sölkner J, Garcia JF. Detecting loci under recent positive selection in dairy and beef cattle by combining different genome-wide scan methods. *PLoS One*. 2013;8, e64280.
- Sabeti PC, Reich DE, Higgins JM, Levine HZP, Richter DJ, Schaffner SF, Gabriel SB, Platko JV, Patterson NJ, McDonald GJ, Ackerman HC, Campbell SJ, Altshuler D, Cooper R, Kwiatkowski D, Ward R, Lander ES. Detecting recent positive selection in the human genome from haplotype structure. *Nature*. 2002;419:832–7.
- Qanbari S, Pimentel ECG, Tetens J, Thaller G, Lichtner P, Sharifi A R, Simianer H. A genome-wide scan for signatures of recent selection in Holstein cattle. *Anim Genet*. 2010;41:377–89.
- Zhang H, Hu X, Wang Z, Zhang Y, Wang S, Wang N, Ma L, Leng L, Wang S, Wang Q, Wang Y, Tang Z, Li N, Da Y, Li H. Selection signature analysis implicates the PC1/PCSK1 region for chicken abdominal fat content. *PLoS One*. 2012;7, e40736.
- Zhang H, Wang S-Z, Wang Z-P, Da Y, Wang N, Hu X-X, Zhang Y-D, Wang Y-X, Leng L, Tang Z-Q, Li H. A genome-wide scan of selective sweeps in two broiler chicken lines divergently selected for abdominal fat content. *BMC Genomics*. 2012;13:704.
- Voight BF, Kudaravalli S, Wen X, Pritchard JK. A map of recent positive selection in the human genome. *PLoS Biol*. 2006;4, e72.
- Nielsen R, Williamson S, Kim Y, Hubisz MJ, Clark AG, Bustamante C. Genomic scans for selective sweeps using SNP data. *Genome Res*. 2005;15:1566–75.
- Williamson SH, Hubisz MJ, Clark AG, Payseur BA, Bustamante CD, Nielsen R. Localizing recent adaptive evolution in the human genome. *PLoS Genet*. 2007;3, e90.
- Qanbari S, Pausch H, Jansen S, Somel M, Strom TM, Fries R, Nielsen R, Simianer H. Classic selective sweeps revealed by massive sequencing in cattle. *PLoS Genet*. 2014;10, e1004148.
- Sabeti PC, Varilly P, Fry B, Lohmueller J, Hostetter E, Cotsapas C, Xie X, Byrne EH, McCarroll SA, Gaudet R, Schaffner SF, Lander ES, Frazer KA, Ballinger DG, Cox DR, Hinds DA, Stuve LL, Gibbs RA, Belmont JW, Boudreau A, Hardenbol P, Leal SM, Pasternak S, Wheeler DA, Willis TD, Yu F, Yang H, Zeng C, Gao Y, Hu H, et al. Genome-wide detection and characterization of positive selection in human populations. *Nature*. 2007;449:913–8.
- Chen H, Patterson N, Reich D. Population differentiation as a test for selective sweeps. *Genome Res*. 2010;20:393–402.
- Groenen MAM, Megens H-J, Zare Y, Warren WC, Hillier LW, Crooijmans RPMA, Vereijken A, Okimoto R, Muir WM, Cheng HH. The development and characterization of a 60K SNP chip for chicken. *BMC Genomics*. 2011;12:274.
- Fu W, Dekkers JCM, Lee WR, Abasht B. Linkage disequilibrium in crossbred and pure line chickens. *Genet Sel Evol*. 2015;47(1):11.
- Groenen MAM, Wahlberg P, Foglio M, Cheng HH, Megens H, Crooijmans RPMA, Besnier F, Lathrop M, Muir WM, Wong GK-S, Gut I, Andersson L. A high-density SNP-based linkage map of the chicken genome reveals

- sequence features correlated with recombination rate. *Genome Res.* 2009;19:510–9.
19. Browning BL, Browning SR. A unified approach to genotype imputation and haplotype-phase inference for large data sets of trios and unrelated individuals. *Am J Hum Genet.* 2009;84:210–23.
 20. Pickrell JJK, Coop G, Novembre J, Kudaravalli S, Li JZ, Absher D, Srinivasan BS, Barsh GS, Myers RM, Feldman MW. Signals of recent positive selection in a worldwide sample of human populations. *Genome Res.* 2009;19:826–37.
 21. Ma Y, Zhang H, Zhang Q, Ding X. Identification of selection footprints on the X Chromosome in pig. *PLoS One.* 2014;9:e94911.
 22. Zhao FP, Wei CH, Zhang L, Liu JS, Wang GK, Zeng T, Du LX. A genome scan of recent positive selection signatures in three sheep populations. *J Integr Agric.* 2016;15:162–74.
 23. Yin T, Cook D, Lawrence M. ggbio: an R package for extending the grammar of graphics for genomic data. *Genome Biol.* 2012;13:R77.
 24. Reich DE, Cargill M, Bolk S, Ireland J, Sabeti PC, Richter DJ, Lavery T, Kouyoumjian R, Farhadian SF, Ward R, Lander ES. Linkage disequilibrium in the human genome. *Nature.* 2001;411:199–204.
 25. Hayes B, Visscher P. Novel multilocus measure of linkage disequilibrium to estimate past effective population size. *Genome Res.* 2003;13:635–43.
 26. Tenesa A, Navarro P, Hayes BJ, Duffy DL, Clarke GM, Goddard ME, Visscher PM. Recent human effective population size estimated from linkage disequilibrium. *Genome Res.* 2007;17:520–26.
 27. Liu G, Dunnington EA, Siegel PB. Correlated responses to long-term divergent selection for eight-week body weight in chickens: growth, sexual maturity, and egg production. *Poult Sci.* 1995;74:1259–68.
 28. Anthony NB, Dunnington E a, Siegel PB. Egg Production and Egg Composition of Parental Lines and F1 and F2 Crosses of White Rock Chickens Selected for 56-Day Body Weight. *Poult Sci.* 1989;68:27–36.
 29. Jaap RG, Muir FV. Erratic Oviposition and egg defects in broiler-type Pullets. *Poult Sci.* 1968;47:417–23.
 30. Lagrutta AA, McCarthy JG, Scherzinger CA, Heywood SM. Identification and developmental expression of a novel embryonic myosin heavy-chain gene in chicken. *DNA.* 1989;8:39–50.
 31. Lyons G, Ontell M, Cox R. The expression of myosin genes in developing skeletal muscle in the mouse embryo. *J Cell Biol.* 1990;111:1465–76.
 32. Tajsharghi H, Kimber E, Kroksmark AK, Jerre R, Tulinius M, Oldfors A. Embryonic myosin heavy-chain mutations cause distal arthrogryposis and developmental myosin myopathy that persists postnatally. *Arch Neurol.* 2008;65:1083–90.
 33. Wang L, Liu X, Niu F, Wang H, He H, Gu Y. Single nucleotide polymorphisms, haplotypes and combined genotypes in MYH3 gene and their associations with growth and carcass traits in Qinchuan cattle. *Mol Biol Reports.* 2013;40:417–26.
 34. Niu F, Wang L, Liu X, Wang H, Yang J. Genetic diversity of MYH3 gene associated with growth and carcass traits in Chinese Qinchuan cattle. *Mol Biol Rep.* 2013;40:5635–43.
 35. Smits P, Li P, Mandel J, Zhang Z, Deng JM, Behringer RR, de Crombrughe B, Lefebvre V. The transcription factors L-Sox5 and Sox6 are essential for cartilage formation. *Dev Cell.* 2001;1:277–90.
 36. Smits P, Dy P, Mitra S, Lefebvre V. Sox5 and Sox6 are needed to develop and maintain source, columnar, and hypertrophic chondrocytes in the cartilage growth plate. *J Cell Biol.* 2004;164:747–58.
 37. Hagiwara N, Yeh M, Liu A. Sox6 is required for normal fiber type differentiation of fetal skeletal muscle in mice. *Dev Dyn.* 2007;236:2062–76.
 38. Goll DE, Thompson VF, Taylor RG, Christiansen JA. Role of the calpain system in muscle growth. *Biochimie.* 1992;74:225–37.
 39. Huang J, Forsberg NE. Role of calpain in skeletal-muscle protein degradation. *Proc Natl Acad Sci.* 1998;95:12100–5.
 40. Arthur JS, Elce JS, Hegadorn C, Williams K, Greer PA. Disruption of the murine calpain small subunit gene, *Capn4*: calpain is essential for embryonic development but not for cell growth and division. *Mol Cell Biol.* 2000;20:4474–81.
 41. Lindholm-Perry AK, Rohrer GA, Holl JW, Shackelford SD, Wheeler TL, Koohmaraie M, Nonneman D. Relationships among calpastatin single nucleotide polymorphisms, calpastatin expression and tenderness in pork longissimus1. *Anim Genet.* 2009;40:713–21.
 42. Hu Y-D, Zhang Z-R, Zhu Q. Identification and Association of the Single Nucleotide Polymorphisms in Calpastatin (CAST) Gene with Carcass Traits in Chicken. *J Anim Vet Adv.* 2011;10:2968–74.
 43. Allais S, Journaux L, Levézuel H, Payet-Duprat N, Raynaud P, Hocquette JF, Lepetit J, Rousset S, Denoyelle C, Bernard-Capel C, Renand G. Effects of polymorphisms in the calpastatin and μ -calpain genes on meat tenderness in 3 French beef breeds. *J Anim Sci.* 2011;89:1–1.
 44. Byun SO, Zhou H, Forrest RHJ, Frampton CM, Hickford JGH. Association of the ovine calpastatin gene with birth weight and growth rate to weaning. *Anim Genet.* 2008;39:572–3.
 45. Greguła-Kania M. Effect of calpastatin gene polymorphism on lamb growth and muscling. *Ann Anim Sci.* 2012;12:63–72.
 46. Tait RG, Shackelford SD, Wheeler TL, King DA, Casas E, Thallman RM, Smith TP, Bennett GL. μ -Calpain, calpastatin, and growth hormone receptor genetic effects on preweaning performance, carcass quality traits, and residual variance of tenderness in Angus cattle selected to increase minor haplotype and allele frequencies. *J Anim Sci.* 2014;92:456–66.
 47. Wu T, Zhang Z, Yuan Z, Lo LJ, Chen J, Wang Y, Peng J. Distinctive Genes Determine Different Intramuscular Fat and Muscle Fiber Ratios of the longissimus dorsi Muscles in Jinhua and Landrace Pigs. *PLoS One.* 2013;8:e53181.
 48. Gunning P, O'Neill G, Hardeman E. Tropomyosin-based regulation of the actin cytoskeleton in time and space. *Physiol Rev.* 2008;88:1–35.
 49. Kuninger D, Kuzmickas R, Peng B, Pintar JE, Rotwein P. Gene discovery by microarray: identification of novel genes induced during growth factor-mediated muscle cell survival and differentiation. *Genomics.* 2004;84:876–89.
 50. McNally E, Dellefave L. Sarcomere mutations in cardiogenesis and ventricular noncompaction. *Trends Cardiovasc Med.* 2009;19:17–21.
 51. Johnsson M, Rubin C, Höglund A, Sahlqvist A, Jonsson KB, Kerje S, Ekwall O, Kämpe O, Andersson L, Jensen P. The role of pleiotropy and linkage in genes affecting a sexual ornament and bone allocation in the chicken. *Mol Ecol.* 2014;23:2275–86.
 52. Qanbari S, Strom TM, Haberer G, Weigend S, Gheyas AA, Turner F, Burt DW, Preisinger R, Gianola D, Simianer H. A High Resolution Genome-Wide Scan for Significant Selective Sweeps: An Application to Pooled Sequence Data in Laying Chickens. *PLoS One.* 2012;7:e49525.
 53. Muir WM, Aggrey SE. *Poultry Genetics, Breeding, and Biotechnology.* Cambridge: CABI Publishing; 2003.
 54. Thiruvankadan AK, Panneerselvam S, Prabakaran R. Layer breeding strategies: an overview. *Worlds Poult Sci J.* 2010;66:477–502.
 55. Thiruvankadan A, Prabakaran R, Panneerselvam S. Broiler breeding strategies over the decades: an overview. *Worlds Poult Sci J.* 2011;67:309–36.
 56. Wolc A, Stricker C, Arango J, Settar P, Fulton JE, O'Sullivan NP, Preisinger R, Habier D, Fernando R, Garrick DJ, Lamont SJ, Dekkers JCM. Breeding value prediction for production traits in layer chickens using pedigree or genomic relationships in a reduced animal model. *Genet Sel Evol.* 2011;43:5.
 57. Lacin E, Yildiz A, Esenbuga N, Macit M. Effects of differences in the initial body weight of groups on laying performance and egg quality parameters of Lohmann laying hens. *Czech J Anim Sci.* 2008;53:466–71.
 58. Zhou H, Mitchell AD, McMurtry JP, Ashwell CM, Lamont SJ. Insulin-like growth factor-I gene polymorphism associations with growth, body composition, skeleton integrity, and metabolic traits in chickens. *Poult Sci.* 2005;84:212–9.
 59. Lazar MA. Thyroid hormone receptors: multiple forms, multiple possibilities. *Endocr Rev.* 1993;14:184–93.
 60. Kaneshige M, Kaneshige K, Zhu XG, Dace A, Garrett L, Carter TA, Kazlauskaite R, Pankratz DG, Wynshaw-Boris A, Refetoff S, Weintraub B. Mice with a targeted mutation in the thyroid hormone β receptor gene exhibit impaired growth and resistance to thyroid hormone. *Proc Natl Acad Sci.* 2000;97:13209–14.
 61. Dittrich R, Beckmann MW, Oppelt PG, Hoffmann I, Lotz L, Kuwert T, Mueller A. Thyroid hormone receptors and reproduction. *J Reprod Immunol.* 2011;90:58–66.
 62. Krassas GE, Poppe K, Glinoe D. Thyroid function and human reproductive health. *Endocr Rev.* 2010;31:702–55.