ANIMAL GENETICS Immunogenetics, Molecular Genetics and Functional Genomics

doi: 10.1111/age.12252

Detecting signatures of selection in nine distinct lines of broiler chickens

John J. Stainton*, Chris S. Haley*†, Brian Charlesworth‡, Andreas Kranis*§, Kellie Watson§ and Pamela Wiener*

*The Roslin Institute and R(D)SVS, University of Edinburgh, Midlothian EH25 9RG, UK. †MRC Human Genetics Unit MRC IGMM, University of Edinburgh, Western General Hospital, Edinburgh EH4 2XU, UK. ‡Institute of Evolutionary Biology, School of Biological Sciences, University of Edinburgh, Edinburgh EH9 3JT, UK. §Aviagen Ltd, Edinburgh EH28 8SZ, UK.

Summary

Modern commercial chickens have been bred for one of two specific purposes: meat production (broilers) or egg production (layers). This has led to large phenotypic changes, so that the genomic signatures of selection may be detectable using statistical techniques. Genetic differentiation between nine distinct broiler lines was calculated using Weir and Cockerham's pairwise F_{ST} estimator for 11 003 genome-wide markers to identify regions showing evidence of differential selection across lines. Differentiation measures were averaged into overlapping sliding windows for each line, and a permutation approach was used to determine the significance of each window. A total of 51 regions were found to show significant differentiation between the lines. Several lines were consistently found to share significant regions, suggesting that the pattern of line divergence is related to selection for broiler traits. The majority of the 51 regions contain QTL relating to broiler traits, but only five of them were found to be significantly enriched for broiler QTL, including a region on chromosome 27 containing 39 broiler QTL and 114 genes. Additionally, a number of these regions have been identified by other selection mapping studies. This study has identified a large number of potential selection signatures, and further tests with higher-density marker data may narrow these regions down to individual genes.

Keywords genetic differentiation, poultry, selection signatures, selective sweep mapping

Introduction

QTL mapping has previously been utilised to identify genomic regions that influence traits of economic or biological interest in domesticated animals, but this approach requires extensive phenotypic information (Haley 1995). A complementary approach is to use purely genotypic information to find regions of the genome showing signatures of selection, which may also be associated with traits of interest. This approach has been applied to the analysis of pigs (Rubin *et al.* 2012; Ai *et al.* 2013; Wilkinson *et al.* 2013), cattle (Gibbs *et al.* 2009; Stella *et al.*, 2010), chickens (Rubin *et al.* 2010; Elferink *et al.* 2012) and sheep (Kijas *et al.* 2012; Moioli *et al.* 2013; Gutiérrez-Gil *et al.* 2014). Modern chickens (*Gallus gallus domesticus*) were domesticated from red jungle fowl before 6000 B.C. (West & Zhou 1989) and bred primarily for cock fighting and food. Within the last 70 years, chickens used for food were split into separate lines specialised for specific functions: broilers (meat production) or layers (egg production) (Muir *et al.* 2008). Both types of lines have been selected for a wide variety of traits, including welfare and disease resistance traits. However, the main broiler production traits include body weight, growth, feed conversion ratio and fat content, whereas layer selection is focused on reproduction-related traits. This has led to dramatic phenotypic changes, so that the genomic signatures of this selection may be detectable using statistical techniques.

Selection has a number of effects on the genome. Positive selection can cause an advantageous allele to spread throughout the population and become fixed. This causes a reduction of diversity in the population at the selected site. Additionally, genetic variation is reduced at neutral sites linked to those under positive selection by the hitchhiking effect, which results in areas of low variation around the selected gene (Maynard-Smith & Haigh 1974). Another effect can occur between populations. Populations with

Address for correspondence

P. Wiener, The Roslin Institute and R(D)SVS, University of Edinburgh, Midlothian EH25 9RG, UK.
E-mail: pam.wiener@roslin.ed.ac.uk

Accepted for publication 3 October 2014

limited gene flow between them and which experience different selection pressures can undergo genetic differentiation (Wright 1951). Population differentiation measured by differences in allele frequencies between populations can be quantified by the statistic F_{ST} (Wright 1951). When there is no selection, F_{ST} is influenced only by genetic drift, which will affect all loci in a similar way. However, when differentiation has occurred due to positive selection at particular loci across populations, an increase in F_{ST} at markers linked to the targets of selection should be observed (Akey *et al.* 2002).

In earlier applications of F_{ST}-based tests, data simulated under neutral conditions were used to compare with real data (Beaumont & Nichol 1996). However, recent increases in marker density allow another approach. Loci, which are affected by selection, show up as outliers in the tails of the empirical distribution of F_{ST}, which cuts out the need for simulations (Akey *et al.* 2002). This technique has been used to suggest candidate genes associated with skin wrinkling in Shar-Pei dogs (Akey *et al.* 2010) and coat patterns in pigs (Wilkinson *et al.* 2013). In dairy cattle, differentiation tests have been applied to look for genes under selection across breeds including Holstein, Normande and Montbéliarde (Flori *et al.* 2009). Genes showing evidence of differentiation include the *growth hormone* gene (*GH*) and *insulin growth factor* gene (*IGF1*), which are both important in milk production.

In this study, we examined genetic differentiation between nine distinct broiler chicken lines that have been under selection for 40 generations (Andreescu *et al.* 2007), using genome-wide data from ~12 K SNPs to identify genomic regions where selection may have taken place during broiler specialisation. Genetic divergence between broiler lines has previously been identified (Granevitze *et al.* 2009) so that we can expect to detect between-line differentiation. Increased population differentiation in a specific genomic region may indicate a recent selective sweep, as sweeps reduce local variability, which also increases F_{ST} (Charlesworth 1998; Cruickshank & Hahn 2014).

Materials and methods

Animals

The birds used in this study were from nine broiler chicken lines provided by Aviagen, a company primarily involved in broiler breeding. They originated from Plymouth Rock and Cornish lines and have been selected for at least 40 generations for a variety of traits, including growth, feed conversion ratio, reproduction and welfare characteristics. Each of these nine lines has slightly different selection criteria, allowing the production of hybrid broiler lines with different characteristics. Genotypes were available for approximately 200 individuals per line (Table 1).

Table 1 Sample sizes (numbers of individuals) of the broiler lines before and after quality control.

	Original sample size	Sample size after quality control
Line 1	200	65
Line 2	200	70
Line 3	200	70
Line 4	201	72
Line 5	202	67
Line 6	200	68
Line 7	200	71
Line 8	200	71
Line 9	200	62

Data

These lines were genotyped for a total of 12 046 SNPs (hereafter referred to as the 12 K dataset) distributed across the genome. These SNPs are a subset of the 2.8 million SNPs identified in the chicken genome sequencing project (Wong *et al.* 2004) and include the ~6 K markers genotyped by Andreescu *et al.* (2007). The mean distance between markers was 0.13 Mb, with a standard deviation of 0.31 Mb (Powell *et al.* 2012). The markers are well distributed across the genome. However, there were several areas with unusually high or low SNP density; for example, an area on chromosome 3 (110–113 Mb) had a particularly high SNP density, whereas chromosome 16 had very low marker coverage due to the presence of the highly variable major histocompatibility complex. Additionally, several low-density areas were found on chromosomes 1, 2 and 3. A total of 11 988 of the markers had known chromosome locations and were distributed across the 28 autosomes and the Z chromosome (Table S1). The number of SNPs per chromosome ranged from 27 to 2282, mainly due to the large difference in size between the macrochromosomes and microchromosomes in the avian genome (Table S1). All genotyped birds were male. Quality control procedures removed from the dataset SNPs that were missing in >10% of the birds ($n = 283$), and those on the Z chromosome were also removed ($n = 602$), leaving 11 003 SNPs for analysis. Quality control procedures were also used to remove individual animals from the analysis. Closely related individuals, including full sibs, half-sibs and parents, were discarded. The first family member found in the pedigree file was retained for analysis and the rest were removed. This removed a total of 1169 birds across the nine lines. Additionally, 18 individuals with greater than 10% missing SNP calls were removed. This left an average of 68 individuals per line (Table 1).

Statistical analysis: population differentiation

To investigate population differentiation, the pairwise Weir and Cockerham's F_{ST} estimator was calculated for each SNP for each pair of lines (total of 36 pairs) (Weir & Cockerham

1984; Akey *et al.* 2002). To calculate an overall differentiation level for a given line, all pairwise F_{ST} estimates for the line were averaged (eight values per line). The individual pairwise F_{ST} values for each pair of lines were also averaged across all SNPs, yielding a genome-wide average differentiation value for each pair of lines. This allowed us to investigate the overall relatedness between the lines.

Sliding windows

The F_{ST} results were averaged into overlapping sliding windows of SNPs to smooth out stochastic variation in estimates across markers (Weir *et al.* 2005). Two methods were evaluated for the implementation of sliding windows across chromosomes. The first was based on a fixed number of SNPs per window. The second was based on a fixed physical size of each window. The 'fixed SNP number' method used a window with a constant number of SNPs, which moved along each chromosome one SNP at a time. The position of the window was determined by the central SNP. Several window sizes were investigated including 5, 11, 15, 21 and 31 SNPs. A window size of 11 SNPs was chosen, as this balanced the identification of signal and removal of stochastic effects. Larger window sizes diluted the signals of selection, whereas smaller window sizes showed too much variation, making it difficult to distinguish regions of high differentiation from random fluctuations. In this method, the size of the window changed along the chromosome, depending on the density of the SNPs in a genomic region. For the 'fixed size' window method, we used a window size of 840 kbp across each chromosome. The central positions of the windows were spaced 85 kbp apart. This size is approximately the average size and increment used for the fixed SNP method with 11 SNP windows. The distribution of the number of SNPs in a window was investigated. To prevent windows with zero or one SNP from overly influencing results, windows with fewer than two SNPs were discarded. Some windows were identical to the previous window, as no new SNPs had been added or lost from the window. New windows, which were not identical, were referred to as 'unique windows'. This method is slightly different from Randhawa *et al.*'s (2014) sliding window method but generates a similar number of windows, assuming the same window size (results not shown).

The distribution of SNP numbers per window based on the fixed window size method (Fig. 1a) had a smaller tail than did the distribution of window sizes per window using the fixed SNP number method (Fig. 1b), with the majority of windows containing fewer than 30 SNPs and an average of 10.2 SNPs per window. Additionally, some of the window sizes used in the fixed SNP number method were very large (Fig. 1b), whereas fixed size windows with many SNPs were relatively rare (Fig. 1a) and occurred mostly in one genomic region on chromosome 3. We therefore used the fixed window size method for selection analysis. The sliding window average of the pairwise average F_{ST} is referred to as the 'F_{ST}-window' value.

Signatures of selection

Two methods were used to identify regions of selection. The first ('F_{ST}-window distribution method') identified the upper tail of the empirical distribution of F_{ST}-windows, defined as the top 0.5% of F_{ST}-window values for each line. The second method ('permutation threshold method') was used to determine a valid significance threshold for the differentiation results. For this method, we performed circular chromosome-bound permutations (Cabrera *et al.* 2012; Kindt *et al.* 2013), using the eight pairwise F_{ST} comparisons for each SNP/line combination, calculated during the differentiation analysis for each line. This procedure (described below) maintains the order of SNPs within each chromosome, thereby preserving the internal structure of the datasets and conserving the relative distance between SNPs and linkage disequilibrium patterns across the genome. It generates an empirical distribution for the mean F_{ST} in a region for a particular line, under the null hypothesis of no overall mean differentiation, against which our observed values can be compared. After the identifica-

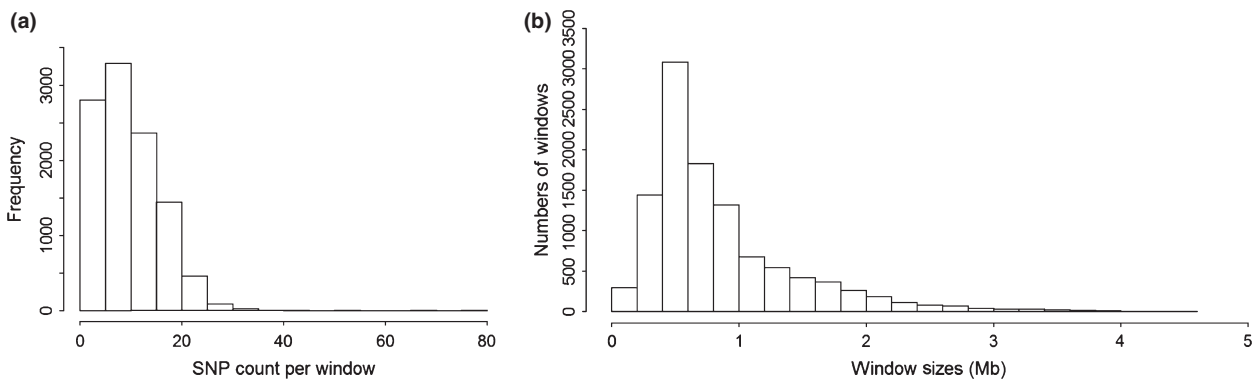

Figure 1 Comparison of fixed window size and fixed SNP number approaches for sliding windows implementation. (a) Distribution of number of SNPs for fixed size windows. (b) Distribution of sizes of windows for fixed SNP number per window.

tion of differentiated regions, we considered two separate sets of results. The first set included regions identified in multiple lines. The second set included regions unique to individual lines. To increase the chances of detecting genuine differentiation due to selection in both sets, only regions spread over more than one unique window were included.

Permutation procedure

Each chromosome was considered to be circular and was permuted separately. The procedure implemented the following algorithm for a given chromosome and a given line: (1) SNPs were numbered based on their position on a given chromosome. (2) For that chromosome, for each of the eight pairwise comparisons associated with that line, a random number between one and the total number of SNPs in the chromosome was generated. The F_{ST} value associated with the first SNP rotated to the position of the SNP with that random number. All other F_{ST} values rotated by the same amount (the position of any SNPs that rotated beyond the final position of the chromosome continued from the beginning of the chromosome). As a different random number was chosen for each of the pairwise comparisons, each of the eight pairwise F_{ST} values for that SNP position was permuted by a different amount. (3) All eight comparisons between lines were averaged into a single overall permuted differentiation value for each SNP for that line. (4) Steps (2)–(3) were repeated for all chromosomes for the current line. (5) Steps (2)–(4) were repeated for each of the other eight lines. (6) Steps (2)–(5) were repeated 100 000 times, giving 100 000 permuted average F_{ST} values for each position/line combination.

To make these permuted results directly comparable to the F_{ST}-window distribution analysis, the average F_{ST} values for each position were again averaged into overlapping sliding windows. These sliding windows were 840 kb long and spaced 85 kb along the genome, as in the F_{ST}-windows analysis. Windows with fewer than two SNPs per window were again discarded. This resulted in a total of 10 498 overlapping windows in both the real data and permutation analysis. These 10 498 overlapping windows are equivalent to 1076 non-overlapping, and hence independent, windows. We thus desired to set our Bonferroni's corrected threshold significance at $P = 0.05/1000$ to take the number of independent windows into account, giving a P-value of approximately 0.00005 for each line. From the permutations, we obtained 100 000 average F_{ST} values for each window. The fifth largest average F_{ST} value for each window was used to set the individual empirical threshold for that window, providing a $5/100\,000 = 0.00005$ threshold for each window.

'Regions' were defined by manually looking for significant windows adjacent to each other. A region could include a small gap of one or two windows to account for windows that were discarded for including fewer than two SNPs. Even with this gap, there is still a large overlap between adjacent windows due to the markers shared between adjacent windows in a sliding window approach. Regions can overlap if there are small (but greater than two windows) gaps. These regions were then compared for overlaps between the lines.

QTL and genes within differentiated regions

To investigate QTL previously discovered by other studies, we probed the Animal QTL Database (http://www.animalgenome.org/cgi-bin/QTLdb/index; Hu *et al.* 2013). The database includes a total of 3811 QTL in chickens, representing 297 traits from 191 publications, and the base pair positions reported are on the Galgal4 build of the chicken genome. As our data were recorded on the WASHUC2 build of the chicken genome, we used the UCSC batch coordinate conversion (LIFTOVER) (http://genome.ucsc.edu/cgi-bin/hgLiftOver) with default settings to position the regions on Galgal4. The Chicken QTL database represents the coordinates of QTL as a map property known as 'span'. Span represents a genomic area but does not incorporate statistical confidence. Many QTL had very large spans, and no peak position was available in the database. To narrow down the possible QTL for each region, we assigned a peak position as the midpoint of the QTL span. We searched for the regions identified in this database and included any QTL whose assigned peak position fell within these regions. 'Broiler QTL' were defined as those meeting certain criteria: (1) they were assigned as production QTL in the Chicken QTL database and (2) they were directly related to meat production. For example, growth, feed conversion ratio, body weight and breast muscle weight are all considered to be broiler QTL. A full list of broiler traits is shown in Table S2.

Additionally, the genes in the regions showing high differentiation were investigated with Ensembl/Biomart, using the dataset *Gallus gallus* genes (Galgal4) from the database Ensembl genes 74 (http://www.ensembl.org/biomart). The number of genes in the region and their functions (if available) were recorded.

Significance of QTL results

To test whether certain regions were enriched for broiler QTL, we applied the following approach. For each genomic region that exceeded the permutation-derived significance threshold, a region of that size was randomly assigned 100 000 times to positions on the chicken genome. The number of broiler QTL peak positions present in these simulated regions, based on the Animal QTL Database, was calculated for each iteration. The top 5% of these 100 000 QTL counts were determined for each region size. If the actual number of QTL for a region exceeded its top 5% threshold, it was classed as enriched.

Additionally, to determine whether there were more differentiated regions that contained one or more QTL than would be expected by chance, we calculated the total QTL coverage of the genome to assess the proportion of QTL across the genome that were found within an average-sized differentiation region. To calculate this, we moved along the chicken genome one base pair at a time and recorded whether each position was within 1.599 Mb (the average size of a differentiated region) of a QTL peak position, such that if the differentiation region covered that position, it would also include a QTL position ('overlap position'). The number of base pairs across the genome, which were overlap positions, was calculated and divided by the total number of base pairs in the genome. To calculate the expected number of differentiated regions containing QTL, this proportion was multiplied by the total number of differentiated regions.

Results

Relationships between lines

The genome-wide differentiation (mean F_{ST}) values for several pairs of lines were much lower than for others (Table 2). Lines 1 and 7 have an extremely low differentiation value at 0.015, suggesting they are more closely related than other lines. Lines 2, 4 and 9 appear to be closely related. Lines 3, 6 and 8 also appear to be related, but not as closely as the previous groups. Line 5 does not seem to be closely related to any of the other eight lines. These results are consistent with a tree generated in a previous study with the same lines, based on allele frequencies on chromosomes 1 and 4 (Andreescu *et al.* 2007; see Fig. 2).

Signatures of selection found in multiple lines

Multiple signatures of selection were identified. The F_{ST}-window distribution analysis identified 87 regions as outliers in terms of their pairwise differentiation levels across the nine broiler lines. Of these regions, 14 were shared between at least two lines and spread over more than one unique window on chromosomes 1, 2, 3, 4, 5, 11 and 14. These 14 regions were positioned on the Galgal4 genome build using LIFTOVER (Table S3). The permutation threshold method found a total of 214 regions. Of these regions, 51 were shared between at least two lines and spread over more than one unique window on chromosomes 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 14, 15, 18, 20, 21, 26, 27 and 28 (Table 3). Regions 15 and 16 and regions 24 and 25 were found to overlap, due to the method used for defining regions and the use of overlapping sliding windows. Twelve of the 14 F_{ST}-window regions were also found in the permutation threshold regions. The other two regions were found in fewer lines using the permutation threshold method; because they were identified in only a single line, they are featured in Table 4 (see below) instead of Table 3. A region at (chromosome) 5: 2.52–4.04 Mb was found in lines 4 and 6 in the F_{ST}-window distribution results, but only in line 6 using the permuted thresholds. Another region was found at 11: 1.89–2.65 Mb in lines 4, 7 and 8 by the F_{ST}-window distribution method but only in line 7 in the permutation threshold results. The other 12 regions were present in both sets of results, but some additional lines were gained from the permuted threshold method. For example, a region at 1: 53.51–54.71 Mb was found for lines 4 and 9 in the F_{ST}-window distribution results, but the overlapping Region 2 from the permutation threshold results showed high differentiation in lines 2, 4, 5, 6, 7 and 9.

Thirty-seven of the permutation threshold regions were not identified using the F_{ST}-window distribution method. These include some regions that showed high differentiation across many lines: Region 17 (3: 14.504–16.661 Mb) showed high differentiation for all lines except line 2 and 3 and Region 50 (27: 2.720–4.502 Mb) also showed high differentiation for all lines except 2 and 3. Because most of the F_{ST}-window regions were a subset of the permutation threshold regions, we chose to concentrate on the 51 permutation threshold-based regions.

Several lines were consistently found to have shared significant regions; only nine regions included lines with no

Table 2 The average differentiation values for each pair of lines, with standard errors included in parentheses. Closely related lines, such as lines 1 and 7, have much lower differentiation than do lines with no shared ancestry (Andreescu *et al.* 2007) (Figure 2).

Line	1	2	3	4	5	6	7	8	9
1	X	0.174 (0.0019)	0.178 (0.0021)	0.171 (0.0019)	0.153 (0.0018)	0.160 (0.0019)	0.015 (0.0003)	0.173 (0.0020)	0.174 (0.0019)
2		X	0.162 (0.0019)	0.091 (0.0012)	0.142 (0.0017)	0.158 (0.0018)	0.170 (0.0019)	0.155 (0.0018)	0.093 (0.0012)
3			X	0.153 (0.0017)	0.156 (0.0018)	0.109 (0.0014)	0.173 (0.0020)	0.035 (0.0006)	0.156 (0.0018)
4				X	0.134 (0.0016)	0.148 (0.0017)	0.167 (0.0018)	0.147 (0.0017)	0.056 (0.0008)
5					X	0.134 (0.0016)	0.147 (0.0017)	0.150 (0.0017)	0.137 (0.0016)
6						X	0.155 (0.0018)	0.103 (0.0013)	0.150 (0.0017)
7							X	0.168 (0.0020)	0.171 (0.0019)
8								X	0.146 (0.0017)
9									X

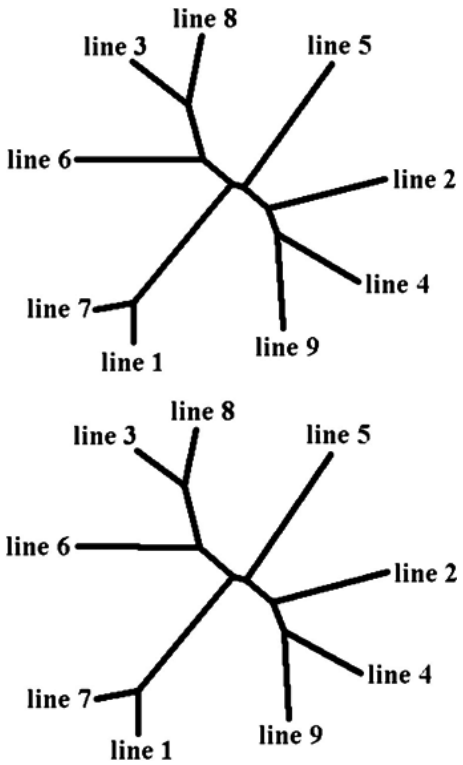

Figure 2 Phylogenetic trees based on marker allele frequencies for chromosome 1 and chromosome 4, obtained using the UPGMA algorithm. Originally produced in Andreescu *et al.* (2007). The Genetics Society of America provided permission to reproduce Figure 2 and retain copyright for this image.

shared branches in the phylogeny shown in Fig. 2. Lines 1 and 7, which are on the same branch, were found to share differentiation signals in 22 regions; nine of which were unique to lines 1 and 7 (the others were shared with additional lines). An additional shared phylogeny branch involves lines 2, 4 and 9. At least two of these lines appear together in 17 regions, five of which were unique to these lines. The last phylogenetic group of lines includes 3, 6 and 8; at least two of these appear together 19 times, and six of these regions are unique to this phylogeny branch. Finally, line 5 appears 10 times, at least once with each of the other lines (most shared with line 7 and least with line 3).

Conversion of coordinates to Galgal4 assembly

Of the 51 significantly differentiated regions, 33 were converted directly from the WASHUC2 (Galgal3) chicken genome assembly to the newer chicken genome assembly, Galgal4, using LIFTOVER. Sixteen of the 18 remaining regions were converted by submitting their start and end points separately. For the other two regions (22 and 40), either the start or end points could not be successfully converted by LIFTOVER. To estimate the position of these missing coordinates, genes nearby were identified in the WASHUC2 genome build and positioned in the Galgal4 build. The closest gene to the end coordinate of region 22 is *SLITRK2.* The end coordinate of this gene in Galgal4 (19 134 524 bp) was used as the missing coordinate for this region, which is likely to be a conservative estimate (i.e. a smaller region), as the 3' end of the gene is found ~0.39 Mb upstream of the end of region 22 according to WASHUC2. For region 40, the closest gene in the WASHUC2 build is the *coatomer protein complex, subunit gamma* gene (*COPG*). As the start coordinate of this gene is only 912 bp away from the missing coordinate, this was used as an estimate of the start position for the region (9 090 877 bp according to Galgal4).

Details of all identified regions positioned on the Galgal4 genome build can be found in Tables S4 and S5, including all QTL and genes found within each region. A few selected examples of these regions are described below (with Galgal4 coordinates). These include the most highly differentiated region, the regions shared by the most lines and the largest region of physical size.

Region 22 (4: 17.84–19.134 Mb) – the region with the highest differentiation

This 1.30-Mb region showed significant differentiation in four lines: 1, 3, 6 and 7. Line 1 contains the window with the highest differentiation across the genome (Fig. 3a). There were no QTL found with a peak position in this region. However, several QTL had spans which overlapped with this region, including body weight, shank growth and visceral fat weight. Shank traits were found in multiple studies.

There were three regions shared by seven broiler lines; these are described below.

Region 15 (2: 139.95–142.128 Mb) – region shared by the majority of lines (1)

This 2.2-Mb region showed significant differentiation in lines 1, 2, 4, 5, 7, 8 and 9 (Fig. 3b). One QTL for body weight at 200 days had its peak position in this region (Table S4). The spans for several other QTL reported in two studies overlapped the region; these include drumstick and thigh muscle weight, fat distribution, blood cell count, breast muscle weight and body weight.

Region 17 (3: 14.50–16.66 Mb) – region shared by the majority of lines (2)

The second region present in seven lines is 2.17 Mb in size, and appears in lines 1, 4, 5, 6, 7, 8 and 9 (Fig. 3c). No QTL were found with a peak position in this region, although a number of QTL had spans that overlapped with this region, including body weight, abdominal fat weight, breast muscle percentage and egg weight. Body weight and abdominal fat QTL were found in multiple studies.

Table 3 Regions displaying putative signatures of selection that were found in multiple lines. The criteria for inclusion in this table are as follows: windows in the region exceeded their permuted significance threshold, the region was found in more than one line and the region was spread over more than one unique window. The 95% significance threshold is included for the five regions for which the number of QTL exceeded that expected by chance. Regions also found by the F_{ST}-window distribution method are highlighted in bold.

Region no.	Chromosome	Start (Mb)	End (Mb)	Size (Mb)	Lines	QTL	Broiler QTL	Genes	95% threshold for broiler QTL number
1	1	2.74	4.52	1.77	3, 5, 8	0	0	24	
2	**1**	**52.13**	**54.71**	**2.59**	**2, 4, 5, 6, 7, 9**	**17**	**5**	**34**	
3	1	102.21	103.65	1.43	1, 4, 7	1	1	13	
4	1	130.48	132.07	1.59	1, 7	2	1	27	
5	1	133.97	136.13	2.16	2, 3, 4, 8, 9	1	1	21	
6	**1**	**140.67**	**142.08**	**1.41**	**4, 6, 8, 9**	**8**	**3**	**6**	
7	1	145.22	146.81	1.59	2, 8	3	1	9	
8	**1**	**177.11**	**178.95**	**1.85**	**2, 3, 8**	**1**	**1**	**27**	
9	1	187.21	188.28	1.08	2, 9	5	2	12	
10	**2**	**16.71**	**19.49**	**2.78**	**2, 3, 4, 5, 7, 9**	**2**	**1**	**34**	
11	2	24.14	25.98	1.85	1, 4, 5, 7, 9	0	0	19	
12	2	102.42	103.60	1.18	3, 6, 7	6	5	19	
13	2	111.56	113.96	2.41	1, 6	22	4	18	
14	2	122.56	123.89	1.33	5, 8	2	1	10	
15	2	139.95	142.13	2.18	1, 2, 4, 5, 7, 8, 9	1	1	20	
16	**2**	**141.88**	**143.47**	**1.59**	**1, 2, 4, 7, 9**	**0**	**0**	**4**	
17	3	14.50	16.66	2.16	1, 4, 5, 6, 7, 8, 9	0	0	72	
18	**3**	**34.21**	**35.66**	**1.46**	**3, 8**	**0**	**0**	**17**	
19	3	38.14	39.36	1.22	1, 6	3	0	19	
20	3	50.58	51.84	1.26	1, 5	5	0	23	
21	3	58.32	59.90	1.58	1, 9	8	1	17	
22	**4**	**17.84**	**19.14**	**1.30**	**1, 3, 6, 7**	**0**	**0**	**7**	
23	4	19.56	20.73	1.17	1, 7	0	0	18	
24	**4**	**57.62**	**59.21**	**1.59**	**2, 3, 8, 9**	**5**	**1**	**13**	
25	4	58.64	59.69	1.06	3, 8	2	1	12	
26	4	63.62	64.79	1.17	2, 8	6	4	17	
27	**4**	**84.77**	**87.19**	**2.42**	**1, 7**	**8**	**2**	**47**	
28	**5**	**27.42**	**29.09**	**1.67**	**1, 7**	**3**	**2**	**37**	
29	5	55.37	56.90	1.53	1, 2, 3, 4, 6	21	6	29	
30	6	10.48	13.11	2.64	1, 3, 4, 7, 8, 9	0	0	52	
31	7	12.03	13.78	1.75	1, 7	1	1	29	
32	7	17.52	18.93	1.41	1, 7	19	18	31	7
33	7	23.20	24.80	1.60	4, 9	5	4	10	
34	7	31.04	32.21	1.18	4, 9	0	0	4	
35	8	26.34	28.01	1.67	1, 3, 6, 7	18	9	50	9
36	9	8.27	9.78	1.51	3, 8	5	4	22	
37	10	4.66	5.91	1.25	2, 3	0	0	22	
38	11	0.00	1.37	1.37	2, 4, 9	0	0	75	
39	**11**	**13.04**	**14.79**	**1.75**	**1, 3, 7, 8**	**0**	**0**	**11**	
40	12	9.09	10.31	1.22	6, 8	1	1	21	
41	13	12.42	13.65	1.23	3, 8	13	7	41	7
42	13	15.61	17.22	1.61	2, 6	3	3	48	
43	14	0.26	1.77	1.50	2, 4	0	0	35	
44	**14**	**9.80**	**10.85**	**1.05**	**1, 7, 9**	**1**	**1**	**10**	
45	15	1.34	2.77	1.42	1, 5, 7	1	0	16	
46	18	7.41	9.00	1.59	1, 4, 7	14	9	24	9
47	20	6.17	7.34	1.17	3, 8	1	0	25	
48	21	1.30	2.47	1.17	1, 7	0	0	48	
49	26	3.43	4.61	1.18	1, 7	3	2	59	
50	27	2.72	4.50	1.78	1, 4, 5, 6, 7, 8, 9	46	39	114	9
51	28	0.00	1.49	1.49	1, 2, 7	0	0	81	

Region 50 (27: 2.72–4.506 Mb) – region shared by the majority of lines (3)

The final region found in seven lines is 1.78 Mb in size and appears in lines 1, 4, 5, 6, 7, 8 and 9 (Fig. 3d). This is a QTL- and gene-rich area, overlapping 59 QTL spans and including 107 genes. Of the 59 QTL spans, 46 had a peak position that fell within this region (Table S4). The QTL include traits such as body weight, carcass protein content,

Table 4 Line-specific regions displaying putative signatures of selection. The criteria for inclusion in this table are as follows: windows in the region exceeded their permuted significance thresholds, the region was spread over more than one unique window and the region was unique to a single line. The 95% significance threshold is included for the three regions, which exceeded the number of QTL expected by chance. Regions also found by the F_{ST}-window distribution method are highlighted in bold.

Region no	Chromosome	Start (Mb)	End (Mb)	Size (Mb)	Line	QTL	Broiler QTL	Genes	95% threshold for broiler QTL number
1	1	5.36	6.37	1.01	5	0	0	14	
2	1	10.20	11.68	1.48	1	2	2	7	
3	1	12.70	13.70	1.00	6	1	0	11	
4	1	15.06	16.00	0.94	2	0	0	15	
5	1	27.70	28.93	1.22	5	5	3	8	
6	1	31.13	32.48	1.35	7	0	0	8	
7	1	34.56	36.15	1.59	8	10	10	27	
8	1	55.56	56.81	1.25	6	4	4	22	
9	1	61.33	62.68	1.35	6	3	2	26	
10	1	64.88	66.06	1.18	5	8	4	18	
11	1	73.03	74.35	1.32	2	3	3	16	
12	1	81.19	82.60	1.41	9	2	0	15	
13	1	84.71	85.88	1.17	4	4	2	11	
14	1	86.06	87.97	1.91	5	7	4	19	
15	1	107.52	109.11	1.59	6	1	1	25	
16	1	109.35	110.16	0.80	3	0	0	15	
17	1	117.45	118.37	0.92	3	1	1	13	
18	1	119.14	120.39	1.25	2	0	0	16	
19	1	122.09	123.24	1.15	8	3	1	15	
20	1	150.42	151.99	1.57	5	11	8	2	
21	1	190.49	191.81	1.32	9	0	0	2	
22	2	0.60	1.87	1.27	2	0	0	21	
23	2	3.91	5.33	1.42	6	16	9	46	
24	2	20.51	21.43	0.92	5	0	0	14	
25	2	31.04	32.24	1.19	7	2	0	23	
26	2	55.06	55.99	0.93	3	0	0	15	
27	2	65.20	66.36	1.17	2	5	0	10	
28	2	69.63	70.67	1.04	8	0	0	1	
29	2	70.90	72.63	1.73	8	1	0	4	
30	2	99.48	100.42	0.94	4	4	4	8	
31	2	106.17	107.26	1.09	1	1	1	11	
32	2	120.45	122.04	1.59	8	3	0	19	
33	2	146.68	147.93	1.25	9	0	0	7	
34	3	0.02	1.36	1.34	1	0	0	15	
35	3	0.86	1.95	1.09	2	0	0	6	
36	3	2.13	3.40	1.27	4	6	3	37	
37	3	2.98	3.91	0.93	8	11	7	27	
38	3	10.86	12.38	1.52	5	0	0	24	
39	3	12.73	14.15	1.43	5	0	0	16	
40	3	16.25	17.67	1.42	8	3	0	40	
41	3	19.96	20.88	0.92	9	8	3	7	
42	3	45.94	47.28	1.34	2	7	3	7	
43	3	48.81	49.73	0.92	5	1	1	10	
44	3	70.34	71.45	1.10	5	7	4	11	
45	3	78.55	79.54	0.99	4	0	0	15	
46	4	6.15	7.07	0.92	5	0	0	1	
47	4	22.25	23.25	1.00	3	0	0	8	
48	4	53.21	54.21	1.00	9	8	3	17	
49	4	55.93	57.11	1.18	1	0	0	15	
50	4	68.00	68.91	0.91	2	5	4	17	
51	4	80.58	81.88	1.31	5	6	3	19	
52	4	83.08	84.26	1.18	1	7	4	10	
53	5	1.34	2.42	1.08	6	0	0	34	
54	**5**	**2.52**	**4.04**	**1.53**	**6**	**0**	**0**	**21**	
55	5	29.61	31.10	1.50	6	13	3	16	
56	6	7.02	8.44	1.42	2	2	2	18	
57	6	31.13	32.13	1.00	5	13	9	22	9

Table 4 (Continued)

Region no	Chromosome	Start (Mb)	End (Mb)	Size (Mb)	Line	QTL	Broiler QTL	Genes	95% threshold for broiler QTL number
58	7	22.11	23.19	1.08	4	0	0	13	
59	8	5.71	7.05	1.34	4	2	2	40	
60	8	10.68	11.76	1.08	2	8	3	26	
61	9	15.20	16.96	1.76	1	0	0	58	
62	9	16.67	17.43	0.75	1	8	8	15	7
63	9	19.93	21.02	1.09	6	0	0	15	
64	9	22.29	23.29	1.00	9	7	1	44	
65	10	1.75	2.76	1.01	2	0	0	57	
66	10	2.20	3.22	1.02	1	0	0	23	
67	10	7.68	8.63	0.94	2	0	0	15	
68	10	15.57	16.83	1.26	6	2	1	17	
69	10	17.82	19.15	1.33	5	18	11	40	11
70	**11**	**1.72**	**2.65**	**0.93**	**7**	**2**	**0**	**46**	
71	12	5.31	6.56	1.26	5	7	6	13	
72	12	8.09	9.00	0.91	4	1	1	21	
73	14	10.13	11.48	1.35	6	4	4	3	
74	15	7.17	8.07	0.90	6	6	5	36	
75	17	3.99	4.97	0.98	6	0	0	27	
76	17	9.48	10.40	0.92	7	0	0	25	
77	18	6.00	7.74	1.74	6	2	2	54	
78	19	0.00	1.23	1.23	5	5	0	51	
79	19	2.75	4.10	1.36	8	5	4	25	
80	20	0.60	2.17	1.57	2	5	5	45	
81	20	9.62	10.92	1.30	6	4	4	101	
82	20	13.22	14.14	0.91	5	0	0	21	
83	22	2.59	3.94	1.35	4	10	4	39	
84	23	3.48	5.12	1.64	9	9	5	73	
85	24	0.00	0.90	0.90	4	1	0	36	
86	24	0.66	1.66	1.00	5	0	0	22	
87	26	0.00	1.50	1.50	6	6	3	83	

growth rate, shank weight percentage and abdominal fat weight (Table S4). QTL for body weight and abdominal fat weight were found by multiple studies.

Region 10 (2: 16.711–19.493 Mb) – largest region in terms of physical size

A 2.78-Mb region on chromosome 2 was the largest region found in the permutation threshold analysis (Fig. 3b). Significant differentiation in this region was found for lines 2, 3, 4, 5, 7 and 9. One broiler QTL for feed intake had a peak position within this region. A total of 34 genes were found in this region (Table S5).

Significance of QTL results

Approximately 82% of the genome was located within 1.599 Mb of any QTL. Thus, with a total of 51 differentiated regions, we expect 42 regions on average to contain QTL by chance. As only 36 of the differentiated regions contained one or more QTL, we did not observe more regions containing QTL than would be expected by chance. However, five regions (32, 35, 41, 46 and 50) were found to be enriched for the numbers of QTL (the number of broiler QTL exceeded or equalled the top 5% threshold) (Table 3).

Line-specific signatures of selection

As well as investigating regions which appeared in a minimum of two lines, regions identified for only a single line were also examined using the permutation threshold method. A total of 87 regions were found that were significantly differentiated in only one line, had a greater F_{ST} value than the permuted threshold and were found in two or more unique windows (Table 4). These regions were placed on the Galgal4 genome build, as described above. Line 5 had the largest number of line-specific regions (18 regions), followed by line 6 (16 regions). Lines 3 and 7 contained the fewest number of line-specific regions, with only four each. Three regions exceeded or equalled their permuted threshold for broiler QTL. These included region 57, a 1-Mb region on chromosome 6 identified in line 5; region 62, a 0.75-Mb region on chromosome 9 identified in line 1 and region 69, a 1.33-Mb region on chromosome 10 identified in line 5.

Details of all identified regions positioned on the Galgal4 genome build can be found in Tables S6 and S7, including all QTL and genes found within each region. Some selected examples of these regions are described below. These include the largest region of physical size, the region containing the most QTL peak positions and the most gene-rich region.

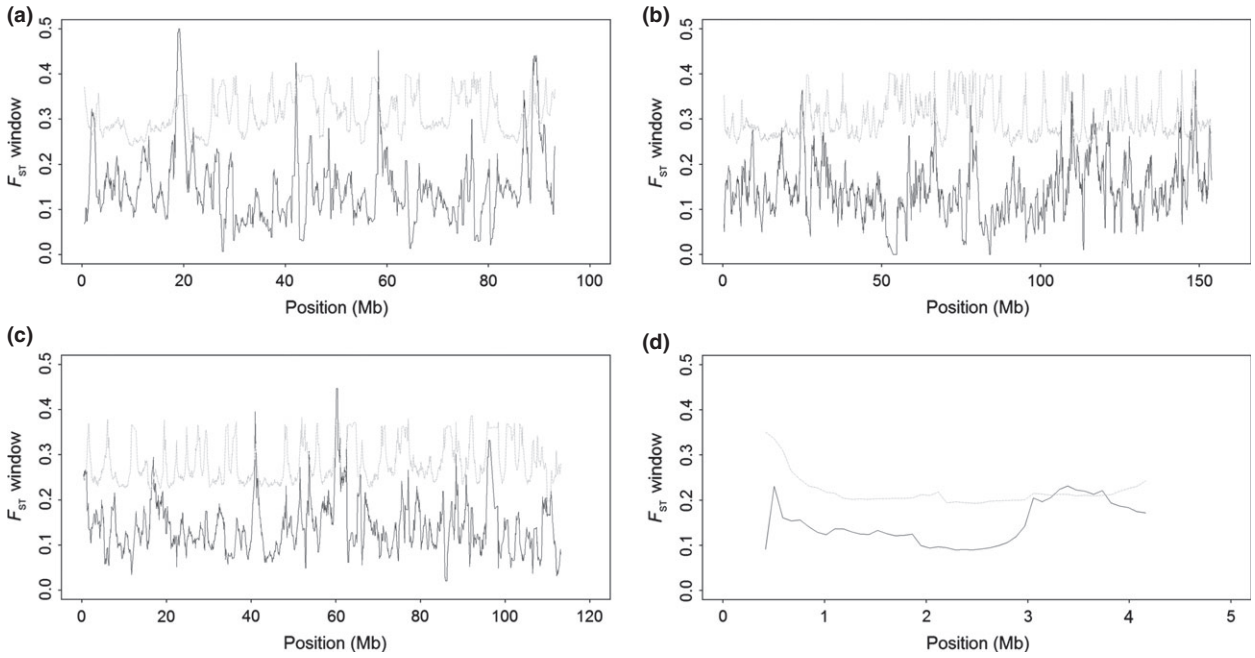

Figure 3 Graphs of position against F_{ST}-window in line 1 for chromosomes (a) 4, (b) 2, (c) 3 and (d) 27. The dashed line represents the permuted significance threshold for each window.

Region 14 (1: 86.06–87.97 Mb) – largest region in terms of physical size

The largest region of physical size is a 1.91-Mb region found on chromosome 1 in line 5. Seven QTL, each from an independent study, were found in this region, four of which were broiler QTL. These include body weight, visceral fat weight, pectoralis minor percent and shank weight percentage. A total of 19 genes were found in this region.

Region 69 (10: 17.82–19.15 Mb) – region with greatest number of QTL peak positions

The region with the greatest number of QTL peak positions, and also statistically enriched for QTL, is a 1.33-Mb region on chromosome 10 found in line 5. A total of 18 QTL were found in this region, 11 of which were broiler QTL. These include seven body weight QTL, found in five independent studies, as well as QTL for abdominal fat percentage, carcass weight and growth. Additionally, 40 genes were found in this region.

Region 81 (20: 9.62–10.92 Mb) – region containing the greatest number of genes

The most gene-enriched region is a 1.30-Mb region on chromosome 20, which was found in line 6. There were a total of 101 genes overlapping with this region. Four broiler QTL, all identified in the same study, were found with peak positions in this region. These include QTL for drumstick and thigh weight, shank length and wing weight.

Discussion

Artificial selection has produced the distinct lines of domesticated chickens we see today and has drastically increased the efficiency of food production. We have identified possible signatures of selection between nine broiler lines by analysing genome-wide SNP data. A total of 51 regions were discovered that showed evidence of differential selection across more than one line, whereas 87 regions showed evidence of differential selection in a single line.

Methodological issues

The overlapping sliding window method used here is slightly different from the majority of sliding window methods. Usually, a fixed number of SNPs is used as the size of a window, and the window moves along the genome one SNP at a time (Stella *et al.*, 2010; Wilkinson *et al.* 2013). However, in this dataset, there were extremely variable SNP densities in different areas of the genome, which would have resulted in substantial variation in the physical size of SNP-based windows and thus invalidated the main approach of comparing average F_{ST} values across windows. The fixed window size method applied here allows the SNP density to vary across the genome. To prevent the results from being affected by windows with very low SNP density, any windows with fewer than two SNPs were discarded.

Previous studies have used the empirical tails of the F_{ST} distribution to identify regions of interest (Akey *et al.* 2002).

However, due to large variation in the number of SNPs per window described above, the empirical tails of the F_{ST} distribution are more likely to include windows with few SNPs, as these will show greater variance and hence more extreme F_{ST} values. This could lead to a bias in the regions selected as significant. A circular permutation method (as suggested by Cabrera *et al.* 2012 and Kindt *et al.* 2013) was therefore utilised to calculate a significance threshold for every window for all nine lines, reducing this bias towards low SNP windows and also preserving the structure of the individual chromosomes, the relative distance between SNPs and the patterns of linkage disequilibrium.

Fifty-one regions that were differentiated in at least two lines were identified by the permutation threshold method, whereas 14 regions were identified by the F_{ST}-window distribution method. The mean SNP count per window for the 51 permutation threshold regions found in multiple lines was 12.37 (SE 0.78), significantly greater than the 5.65 (SE 1.22) SNPs/window for the 14 regions identified by the F_{ST}-window distribution ($P < 0.001$), consistent with the above prediction. The average SNP count across all windows in the genome (with at least two SNPs) fell between the average values for the regions identified by the two methods (10.16, SE 0.062). All of the F_{ST}-window distribution regions overlapped with regions identified using the permutation threshold prior to the imposition of additional criteria, although there were some differences with respect to the lines that were represented in a given region.

Composition of differentiated regions

Thirty-six of the 51 differentiated regions found in multiple lines based on the permutation threshold results overlapped with QTL from the Animal QTL Database, including QTL for broiler production traits. Region 50 on chromosome 27 is particularly interesting in that it contains peak positions for 39 broiler QTL, well in excess of the number expected by chance. This region had already been highlighted because it was highly differentiated in seven of the nine lines, suggesting selection across broiler lines for production-related characteristics. In this region, lines 1, 4, 7 and 9 showed the largest allele frequency differences from other lines and levels of differentiation substantially greater than their genome-wide averages. The related pairs of lines (1 and 7, 4 and 9) show similar allele frequencies (Table 2). This suggests that the main selection in this region has been on this group of lines.

The QTL in region 50 were reported in 12 independent studies of broiler production traits, which included body weight at various days (Ambo *et al.* 2009; Podisi *et al.* 2011), abdominal fat weight percentage (Campos *et al.* 2009; Ankra-Badu *et al.* 2010), shank weight percentage (Baron *et al.* 2011), drumstick and thigh weight percentage (Ankra-Badu *et al.* 2010) and carcass fat and protein content (Nones *et al.* 2012). Two insulin-like growth factor (IGF) binding protein genes are found in this region: *insulin-like growth factor 2 mRNA binding protein 1* (*IGF2BP1*), which binds to the mRNA of *insulin-like growth factor 2* (*IGF2*) and other genes and regulates their translation, and *insulin-like growth factor binding protein 4* (*IGFBP4*), which binds IGFs and alters their interaction with cell surface receptors. Furthermore, there are several genes located within 1 Mb of Region 50 that have been suggested to account for broiler QTL, including *growth hormone 1* (*GH1*), *growth hormone-releasing hormone receptor* (*GHRHR*) and *corticotropin releasing hormone receptor 1* (*CRHR1*) (Nones *et al.* 2012). Our findings suggest that one or more of these growth-related genes may have been differentially selected in the nine broiler lines.

Two other genes of note were found in the differentiated regions detected in multiple lines. *LEPR (leptin receptor* gene, Region 35) encodes the receptor for the leptin hormone, which is secreted by adipose tissue and plays a key role in appetite and energy expenditure in mammals. SNPS in *LEPR* were previously found to differ in allele frequencies between broilers and layers (Twito *et al.* 2011). *MXD4 (MAX dimerisation protein 4,* Region 27) is a tumour suppressor gene and is involved in cell growth. It was recently identified in a GWAS study as a positional candidate gene associated with abdominal fat weight in chickens (Sun *et al.* 2013).

The majority of the 51 differentiated regions found in multiple lines also overlapped with regions showing evidence of selection in one or more previous studies of chickens (Rubin *et al.* 2010; Zhang *et al.* 2012a,b; Fan *et al.* 2013). These include Regions 15 and 16, the first of which was mentioned above because it was shared by seven of the nine broiler lines. These regions overlapped regions of high homozygosity in a set of domestic lines and a subset of commercial broilers (Rubin *et al.* 2010) and in two Asian breeds (Fan *et al.* 2013). Region 47 was also notable in that it overlapped with regions identified in three independent studies: regions of high homozygosity in commercial broilers (Rubin *et al.* 2010) and two Asian breeds (Fan *et al.* 2013) and a region showing allele frequency differences between lines under divergent selection for abdominal fat (Zhang *et al.* 2012a). As observed for region 50, the primary allele frequency differences in most of the other QTL-enriched regions were observed between lines 1 and 7 and lines 4 and 9.

Fifty-six of the 87 regions that were differentiated in individual lines overlapped with QTL from the Animal QTL Database. Of these 56 regions, 47 contain QTL associated with broiler production traits. Several of these regions also overlap with regions showing evidence of selection in previous chicken studies. Region 87 (line 6) on chromosome 26 overlaps with a region displaying high linkage disequilibrium values in two lines under divergent selection for abdominal fat (Zhang *et al.* 2012a) and a region of high homozygosity in commercial broilers (Rubin *et al.* 2010). Regions 30 (line 4), 37 (line 8), 52 (line 1), 61 (line 1) and 81 (line 6) overlap with regions displaying significant

relative extended haplotype homozygosity in the lines divergently selected for abdominal fat (Zhang *et al.* 2012b). Line 5 contains the largest number of line-specific differentiated regions. This is consistent with the overall line relationships (Table 2, Fig. 2), which show that line 5 is not closely related to any of the other lines.

A number of candidate genes known to be related to growth were found in the line-specific differentiated regions. The *growth hormone-releasing hormone receptor* gene (*GHRHR*; region 22, line 2) encodes the receptor that binds growth hormone-releasing hormone, which causes the production and release of growth hormone. Mutations in this gene have been known to cause isolated growth hormone deficiency, which is associated with small size in mice (Godfrey *et al.* 1993). *Insulin-like growth factor 1 receptor* (*IGF1R*; region 68, line 6) is activated by insulin-like growth factor 1, which is important in growth. Mice lacking the *IGF-1* receptor die during development and have massively reduced body mass compared to normal mice (Liu *et al.* 1993). Finally, *insulin-like growth factor 2 mRNA binding protein 3* (*IGF2BP3*; region 25, line 7) has been known to repress translation of insulin-like growth factor II, a growth-promoting hormone, during late development (Nielsen *et al.* 1999).

The majority of the above-mentioned QTL studies investigated broiler-layer crosses. Therefore, these QTL represent differentiation between broiler and layer lines but not necessarily between different broiler lines. The absence of broiler QTL for some of the regions identified in these studies may reflect this bias towards regions differentiated between broilers and layers.

Future studies

Due to the high density of QTL in the chicken genome, we did not find more regions containing at least one QTL than would be expected by chance. This may be partially due to the relatively large average size of differentiated regions (1.599 Mb), which can be reduced in future studies by use of a denser SNP chip (i.e. the recently developed 600 K SNP chip; Kranis *et al.* 2013). A denser SNP chip may also allow the detection of additional regions that were not found in this analysis.

Conclusions

In summary, a large number of possible selection signatures were identified by this study of broiler chicken lines. QTL associated with broiler chicken traits from previous studies have been found in some of these regions. In particular, a region on chromosome 27 was highly differentiated between most of the broiler lines and included a large number of QTL for broiler traits as well as genes associated with growth regulation. Further tests with higher-density marker data may allow these regions to be narrowed down to individual genes.

Acknowledgements

This work was supported by funding from the BBSRC (core funding to the Roslin Institute and a CASE studentship), Biosciences KTN and Aviagen Ltd.

References

Ai H., Huang L. & Ren J. (2013) Genetic diversity, linkage disequilibrium and selection signatures in Chinese and Western pigs revealed by genome-wide SNP markers. *PLoS ONE* **8**, e56001.

Akey J.M., Zhang G., Zhang K, Jin L. & Shriver M.D. (2002) Interrogating a high-density SNP map for signatures of natural selection. *Genome Research* **12**, 1805–14.

Akey J.M., Ruhe A.L., Akey D.T., Wong A.K., Connelly C.F., Madeoy J., Nicholas T.J. & Neff M.W. (2010) Tracking footprints of artificial selection in the dog genome. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 1160–5.

Ambo M., Moura A.S.A M.T., Ledur M.C. *et al.* (2009) Quantitative trait loci for performance traits in a broiler × layer cross. *Animal Genetics* **40**, 200–8.

Andreescu C., Avendano S., Brown S.R., Hassen A., Lamont S.J. & Dekkers J.C.M. (2007) Linkage disequilibrium in related breeding lines of chickens. *Genetics* **177**, 2161–9.

Ankra-Badu G., Shriner D., Le Bihan-Duval E. *et al.* (2010) Mapping main, epistatic and sex-specific QTL for body composition in a chicken population divergently selected for low or high growth rate. *BMC Genomics* **11**, 107.

Baron E.E., Moura A.S.A.M.T., Ledur M.C. *et al.* (2011) QTL for percentage of carcass and carcass parts in a broiler x layer cross. *Animal Genetics* **42**, 117–24.

Beaumont M.A. & Nichol R.A. (1996) Evaluating loci for use in the genetic analysis of population structure. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **263**, 1619–26.

Cabrera C.P., Navarro P., Huffman J.E. *et al.* (2012) Uncovering networks from genome-wide association studies via circular genomic permutation. *G3: Genes|Genomes|Genetics* **2**, 1067–75.

Campos R.L.R., Nones K., Ledur M.C. *et al.* (2009) Quantitative trait loci associated with fatness in a broiler–layer cross. *Animal Genetics* **40**, 729–36.

Charlesworth B. (1998) Measures of divergence between populations and the effect of forces that reduce variability. *Molecular Biology and Evolution* **15**, 538–43.

Cruickshank T.E. & Hahn M.W. (2014) Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. *Molecular Ecology* **23**, 3133–57.

Elferink M.G., Megens H.-J., Vereijken A., Hu X., Crooijmans R.P.M.A. & Groenen M.A.M. (2012) Signatures of selection in the genomes of commercial and non-commercial chicken breeds. *PLoS ONE* **7**, e32720.

Fan W.-L., Ng C.S., Chen C.-F. *et al.* (2013) Genome-wide patterns of genetic variation in two domestic chickens. *Genome Biology and Evolution* **5**, 1376–92.

Flori L., Fritz S., Jaffrézic F., Boussaha M., Gut I., Heath S., Foulley J.-L. & Gautier M. (2009) The genome response to artificial selection: a case study in dairy cattle. *PLoS ONE* **4**, e6595.

Gibbs R.A., Taylor J.F., Van Tassell C.P. *et al.* (2009) Genome-wide survey of SNP variation uncovers the genetic structure of cattle breeds. *Science* **324**, 528–32.

Godfrey P., Rahal J.O., Beamer W.G., Copeland N.G., Jenkins N.A. & Mayo K.E. (1993) GHRH receptor of little mice contains a missense mutation in the extracellular domain that disrupts receptor function. *Nature Genetics* **4**, 227–32.

Granevitze Z., Hillel J., Feldman M., Six A., Eding H. & Weigend S. (2009) Genetic structure of a wide-spectrum chicken gene pool. *Animal Genetics* **40**, 686–93.

Gutiérrez-Gil B., Arranz J.J., Pong-Wong R., García-Gámez E., Kijas J. & Wiener P. (2014) Application of selection mapping to identify genomic regions associated with dairy production in sheep. *PLoS ONE* **9**, e94623.

Haley C.S. (1995) Livestock QTLs – bringing home the bacon? *Trends in Genetics* **11**, 488–92.

Hu Z.-L., Park C.A., Wu X.-L. & Reecy J.L. (2013) Animal QTLdb: an improved database tool for livestock animal QTL/association data dissemination in the post-genome era. *Nucleic Acids Research* **41**, D871–9.

Kijas J.W., Lenstra J.A., Hayes B. *et al.* (2012) Genome-wide analysis of the world's sheep breeds reveals high levels of historic mixture and strong recent selection. *PLoS Biology* **10**, e1001258.

Kindt A.S.D., Navarro P., Semple C.A.M. & Haley C.S. (2013) The genomic signature of trait-associated variants. *BMC Genomics* **14**, 108.

Kranis A., Gheyas A., Boschiero C. *et al.* (2013) Development of a high density 600K SNP genotyping array for chicken. *BMC Genomics* **14**, 59.

Liu J.-P., Baker J., Perkins A.S., Robertson E.J. & Efstratiadis A. (1993) Mice carrying null mutations of the genes encoding insulin-like growth factor I (Igf-1) and type 1 IGF receptor (Igf1r). *Cell* **75**, 59–72.

Maynard-Smith J. & Haigh J. (1974) The hitch-hiking effect of a favourable gene. *Genetical Research* **23**, 23–35.

Moioli B., Scatà M.C., Steri R., Napolitano F. & Catillo G. (2013) Signatures of selection identify loci associated with milk yield in sheep. *BMC Genetics* **14**, 76.

Muir W.M., Wong G.K.-S., Zhang Y. *et al.* (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 of the United States of America* **105**, 17312–7.

Nielsen J., Christiansen J., Lykke-Andersen J., Johnsen A.H., Wewer U.M. & Nielsen F.C. (1999) A family of insulin-like growth factor ii mRNA-binding proteins represses translation in late development. *Molecular and Cellular Biology* **19**, 1262–70.

Nones K., Ledur M.C., Zanella E.L. *et al.* (2012) Quantitative trait loci associated with chemical composition of the chicken carcass. *Animal Genetics* **43**, 570–6.

Podisi B.K., Knott S.A., Dunn I.C., Law A.S., Burt D.W. & Hocking PM. (2011) Overlap of quantitative trait loci for early growth rate, and for body weight and age at onset of sexual maturity in chickens. *Reproduction* **141**, 381–9.

Powell J.E., Kranis A., Floyd J., Dekkers J.C., Knott S. & Haley C.S. (2012) Optimal use of regression models in genome-wide association studies. *Animal Genetics* **43**, 133–43.

Randhawa I.A., Khatkar M., Thomson P. & Raadsma H.W. (2014) Composite selection signals can localize the trait specific genomic regions in multi-breed populations of cattle and sheep. *BMC Genetics* **15**, 34.

Rubin C.-J., Zody M.C., Eriksson J. *et al.* (2010) Whole-genome resequencing reveals loci under selection during chicken domestication. *Nature* **464**, 587–91.

Rubin C.-J., Megens H.-J., Barrio A.M. *et al.* (2012) Strong signatures of selection in the domestic pig genome. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 19529–36.

Stella, A., Ajmone-Marsan, P., Lazzari, B. & Boettcher, P. (2010) Identification of Selection Signatures in Cattle Breeds Selected for Dairy Production. *Genetics* **185**, 1451–61.

Sun Y., Zhao G., Li R., Zheng M., Hu Y., Wu D., Zhang L., Li P. & Wen J. (2013) The identification of 14 new genes for meat quality traits in chicken using a genome-wide association study. *BMC Genomics* **14**, 458.

Twito T., Madeleine D., Perl-Treves R., Hillel J. & Lavi U. (2011) Comparative genome analysis with the human genome reveals chicken genes associated with fatness and body weight. *Animal Genetics* **42**, 642–9.

Weir B.S. & Cockerham C.C. (1984) Estimating *F*-statistics for the analysis of population structure. *Evolution* **38**, 1358–70.

Weir B.S., Cardon L.R., Anderson A.D., Nielsen D.M. & Hill W.G. (2005) Measures of human population structure show heterogeneity among genomic regions. *Genome Research* **15**, 1468–76.

West B. & Zhou B.-X. (1989) Did chickens go north? New evidence for domestication. *World's Poultry Science Journal* **45**, 205–18.

Wilkinson S., Lu Z.H., Megens H.-J., Archibald A.L., Haley C., Jackson I.J., Groenen M.A., Crooijmans R.P., Ogden R. & Wiener P. (2013) Signatures of diversifying selection in European pig breeds. *PLoS Genetics* **9**, e1003453.

Wong G., Liu B., Wang J. *et al.* (2004) A genetic variation map for chicken with 2.8 million single-nucleotide polymorphisms. *Nature* **432**, 717–22.

Wright S. (1951) The genetical structure of populations. *Annals of Eugenics* **15**, 323–54.

Zhang H., Hu X., Wang Z. *et al.* (2012a) Selection signature analysis implicates the *PC1/PCSK1* region for chicken abdominal fat content. *PLoS ONE* **7**, e40736.

Zhang H., Wang S.-Z., Wang Z.-P. *et al.* (2012b) A genome-wide scan of selective sweeps in two broiler chicken lines divergently selected for abdominal fat content. *BMC Genomics* **13**, 704.

Supporting information

Additional supporting information may be found in the online version of this article.

Table S1 Statistics of the 12 K SNP dataset.

Table S2 List of QTL defined as broiler traits found in the chicken QTL database.

Table S3 14 regions identified by the F_{ST}-window distribution analysis, shared between more than one line and spread over at least two unique windows.

Table S4 QTL peak positions found for multiple lines in Regions 1–51.

Table S5 Genes overlapping with Regions 1–51 for multiple lines.

Table S6 QTL peak positions found for single lines in Regions 1–87.

Table S7 Genes overlapping with Regions 1 – 87 for single lines.