Proceedings, 10th World Congress of Genetics Applied to Livestock Production

Detecting Signatures of Selection in Lines of Broiler Chickens

J. Stainton*, **C. Haley**†*, **B. Charlesworth**‡, **A. Kranis**§*, **K. Watson**§ **and P. Wiener***.
* The Roslin Institute and R(D)SVS, University of Edinburgh, Midlothian, United Kingdom.
†MRC Human Genetics Unit MRC IGMM, University of Edinburgh Western General Hospital, Edinburgh, United Kingdom ‡Institute of Evolutionary Biology, University of Edinburgh, United Kingdom. §Aviagen Ltd, Edinburgh, United Kingdom

ABSTRACT: Modern commercial chickens have been selected for one of two specific purposes: egg production (layers) and meat production (broilers). Further selection has occurred within these groups and genomic signatures of selection may be detectable by statistical techniques. Regions identified by this approach may include genes associated with production traits.

Weir and Cockerham's pairwise F_{ST} was calculated for genome-wide markers between nine broiler lines and averaged into overlapping sliding windows. The significance of each window was determined by a permutation method. Fifty-one regions were identified as showing significant differentiation, the majority of which contain QTL relating to broiler traits. Five regions were significantly enriched for broiler QTL, including a region on chromosome 27 containing 39 broiler QTL and 114 genes, several of which are functional candidates for association with broiler traits. Further studies with higher density markers may narrow these regions down to individual genes.
Keywords: Broilers; Genetics; Selection

Introduction

QTL mapping can be used to identify traits of interest in domesticated animals but this approach needs extensive phenotypic information (Haley (1995)). A complementary approach uses genotype information to find regions in the genome showing signatures of selection, which may be associated with traits of interest. This approach has been previously applied in the analysis of cattle (Stella et al. (2010)), sheep (Kijas et al. (2012)), pigs (Rubin et al. (2012)) and chickens (Rubin et al. (2010); Elferink et al. (2012)). Domestic chickens (*Gallus gallus domesticus*) are bred for one of two specific purposes: meat (broilers) or egg production (layers). The split into two separate lines occurred approximately seventy years ago (Muir et al. (2008)), and caused dramatic phenotypic changes. Further selection has subsequently occurred within broilers and layers and the genomic signatures of such selection may indicate regions that include genes influencing production traits.

Populations with limited gene flow between them may experience different selection pressures, which can lead to population differentiation. This differentiation is manifested by differences in allele frequencies between populations and can be quantified by the F_{ST} statistic. When selection-related differentiation has occurred, an increase in F_{ST} at markers near the selected gene or genes should be observed. In earlier versions of F_{ST}-based tests, significant results were determined by simulating data under neutral conditions and comparing this to the real data. Recent increases in marker density allow the simulation step to be removed, as loci affected by selection show up as outliers in the tails of the empirical distribution of F_{ST}. This technique has been used to suggest candidate genes associated with coat patterns in pigs (Wilkinson et al. (2013)) and skin wrinkling in Shar-Pei dogs (Akey et al. (2010)). We examined genetic differentiation between nine distinct broiler chicken lines genotyped for ~12k SNP markers to identify genomic regions where selection may have taken place.

Material and Methods

Animals. Nine broiler lines from Aviagen were used in this study. Each of these nine lines has been selected for slightly different criteria, allowing the production of hybrid broiler lines with different characteristics. All birds selected for genotyping were male.

Data. These lines were genotyped for 12,046 SNPs: 11,988 of these had known chromosomal locations and were distributed across the 28 autosomes and the Z chromosome. Quality control removed individuals with greater than 40% missing data. Care was taken to ensure no closely related individuals were included in the analysis. Following quality control, there were approximately 68 individuals analysed per line. SNPs with greater than 10% missing data were also removed, leaving 11,759 SNPs.

Statistical Analysis. Population differentiation was investigated by calculating the pairwise Weir and Cockerham's F_{ST} estimator for each SNP for each pair of lines (Weir et al. (1984); Akey et al. (2002)). An overall differentiation value for each line was calculated by averaging all pairwise values for the line.

Sliding Windows. To account for stochastic effects across markers, the F_{ST} values were averaged into overlapping sliding windows (Weir et al. (2005)). Two sliding window methods were implemented. The first used a fixed number of SNPs (11) per window. The second used a fixed genomic size (840kbp per window). The central positions of each window were spaced 85kbp apart and windows that contained 2 or fewer SNPs were discarded. Distributions of window sizes in the fixed SNP method and of SNP numbers per window in the fixed size method

were evaluated. Fixed size windows with large numbers of SNPs were relatively rare; therefore the fixed size method was used for analysis, resulting in "F_{ST}-window values."

Signatures of Selection. Signatures of selection were identified using two separate methods. The first method identified the upper tail of the empirical distribution of F_{ST}-window values. This was defined as the top 0.5% of windows of each line ("F_{ST}-distribution method").

A second method was used to determine a valid significance threshold for the F_{ST}-window values, using a circular, chromosome-bound permutation technique (Cabrera et al. (2012); Kindt et al. (2013)) ("permutation threshold method"). For these permuted results to be directly comparable to the actual F_{ST} results, they were averaged into sliding windows using the same criteria as above. For each window in each line, the fifth largest window was used as a significance threshold such that the probability of exceeding this value by chance is approximately 0.00005.

For both methods, regions were defined by manual identification of sequential significant windows. These regions could contain a small gap of two windows. Additional criteria were applied to the results to increase the chance of detecting differentiation specifically due to selection; only regions spread over more than one window and found in more than one line were included.

Composition of regions. Each region was investigated for QTLs and genes using the animal QTL Database (Hu et al. (2013)) and Ensembl Biomart (http://www.ensembl.org/biomart). Regions were also investigated for broiler QTL enrichment by random assignment (100,000 times) to different areas of the genome and recording the number of broiler QTL peak positions found. Any region for which the actual QTL count exceeded the top 5% of random QTL counts was classed as enriched. Finally, results from other selection mapping studies were investigated for overlap with these regions.

Results

A total of 14 regions were identified by the F_{ST}-distribution method while 51 potential selection regions were identified by the permutation threshold method. As the F_{ST}-distribution results were mainly a subset of the permutation threshold results, we focused on the permutation threshold based regions. Several lines were consistently found to have shared selection signatures.

Thirty-six of the 51 regions overlapped QTL in the Animal QTL Database, including QTLs for broiler production traits. Five regions were found to be enriched for broiler QTL (Table 1). Region 50 on chromosome 27 included 39 broiler QTL and 114 genes (Figure 1), and had been identified as a region of high homozygosity in a previous study of two Asian chicken breeds (Fan et al. (2013)). Region 47 overlapped with regions identified by three independent selection mapping studies of chickens (Rubin et al. (2010); Zhang et al. (2012); Fan et al. (2013)).

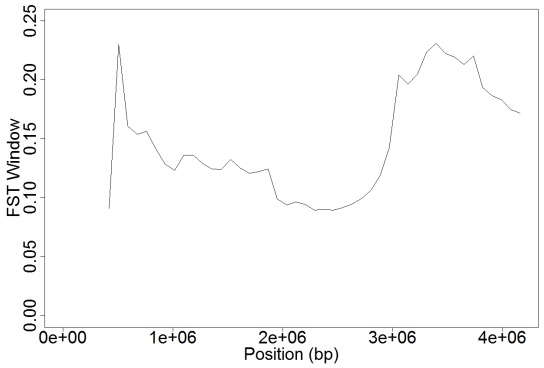

Figure 1. F_{ST}-window values across chromosome 27, which contains region 50

Table 1. Regions enriched for broiler QTL, including the number of QTL present and threshold for each region.

Region Number	Chr	Start, Mb	End, Mb	Broiler QTL	0.95 Threshold
32	7	17.522	18.927	18	7
35	8	26.344	28.014	9	9
41	13	12.418	13.648	7	7
46	18	7.410	9.002	9	9
50	27	2.720	4.502	39	9

Discussion

The overlapping sliding window method implemented here uses a fixed size of the genome rather than a fixed number of SNPs. However, there were variable SNP densities in different areas of the genome, resulting in considerable variation of the physical size of SNP-based windows. The fixed window size method accounts for variable SNP density across the genome.

Previous studies have used the empirical tails of the F_{ST} distribution to identify regions where selection may have occurred. As there is a large variation in the number of SNPs per window, the empirical tails are more likely to include windows with fewer SNPs, as these regions will show greater variance and more extreme differentiation values. The permutation threshold method was utilised to calculate a significance threshold for every window, which reduced this bias towards low SNP density windows.

Region 50 on chromosome 27 is particularly interesting, as it contains peak positions for 39 broiler QTL, reported in 12 independent studies of broiler traits. Two insulin-like growth factor binding genes are found in this region and several growth-regulation-related genes (*GH1*, *GHRHR* and *CRHR1*) are located within 1Mb of the region, which may account for broiler QTL (Nones et al. (2012)). Region 50 was present in seven out of nine lines, suggesting selection across lines for traits related to broiler production.

Conclusions

A large number of possible selection signatures were identified by this study of broiler chicken lines. Many of these regions contain QTL associated with broiler chicken

traits in previous studies. These regions may be narrowed down to individual genes using higher density marker data.

Akey, J. M., Ruhe, A. L., Akey, D. T., et al. (2010) *Proc. Natl. Acad. Sci. U.S.A.,* 107(3), pp. 1160-1165.

Akey, J. M., Zhang, G., Zhang, K., et al. (2002) *Genome Res.,* 12(12), pp. 1805-1814.

Cabrera, C. P., Navarro, P., Huffman, J. E., et al. (2012) *G3 (Bethesda),* 2(9), pp. 1067-1075.

Elferink, M. G., Megens, H.-J., Vereijken, A., et al. (2012) *PLoS One,* 7(2), pp. e32720.

Fan, W.-L., Ng, C. S., Chen, C.-F., et al. (2013) *Genome Biol. Evol.,* 5(7), pp. 1376-1392.

Haley, C. S. (1995) *Trends Genet.,* 11(12), pp. 488-492.

Hu, Z.-L., Park, C. A., Wu, X.-L., et al. (2013) *Nucleic Acids Res.,* 41(D1), pp. D871-D879.

Kijas, J. W., Lenstra, J. A., Hayes, B., et al. (2012) *PLoS Biol.,* 10(2), pp. e1001258.

Kindt, A. S. D., Navarro, P., Semple, C. A. M., et al. (2013) *BMC Genomics,* 14(1), pp. 108.

Muir, W. M., Wong, G. K.-S., Zhang, Y., et al. (2008) *Proc. Natl. Acad. Sci. U.S.A.,* 105(45), pp. 17312-17317.

Nones, K., Ledur, M. C., Zanella, E. L., et al. (2012) *Anim. Genet.,* 43(5), pp. 570-576.

Rubin, C.-J., Megens, H.-J., Barrio, A. M., et al. (2012) *Proc. Natl. Acad. Sci. U.S.A.,* 109(48), pp. 19529-19536.

Rubin, C.-J., Zody, M. C., Eriksson, J., et al. (2010) *Nature,* 464, pp. 587-591.

Stella, A., Ajmone-Marsan, P., Lazzari, B., et al. (2010) *Genetics,* 185(4), pp. 1451-1461.

Weir, B. S., Cardon, L. R., Anderson, A. D., et al. (2005) *Genome Res.,* 15(11), pp. 1468-1476.

Weir, B. S. and Cockerham, C. C. (1984) *Evolution,* 38(6), pp. 1358-1370.

Wilkinson, S., Lu, Z. H., Megens, H.-J., et al. (2013) *PLoS Genet.,* 9(4), pp. e1003453.

Zhang, H., Hu, X., Wang, Z., et al. (2012) *PLoS One,* 7(7), pp. e40736.