Integration of GWAS, CNV and selection signature reveals candidate genes for abdominal fat regulation in chickens

G.C.M. Moreira[1], T.F. Godoy[1], C. Boschiero[1], A.S.M. Cesar[1], J.M. Reecy[2], M.C. Ledur[3], D. Garrick[4], V.H. da Silva[5,6,7], R.P.M.A. Crooijmans[5], M.A.M. Groenen[5], L.L. Coutinho[1]

[1] University of São Paulo (USP) / Luiz de Queiroz College of Agriculture (ESALQ), Piracicaba, São Paulo, Brazil
llcoutinho@usp.br (Corresponding Author)
[2] Iowa State University (ISU), Department of Animal Science, Ames, Iowa, United States.
[3] Embrapa Suínos e Aves, Concórdia, Santa Catarina, Brazil
[4] School of Agriculture, Massey University, Ruakura, Hamilton, New Zealand
[5]Animal Breeding and Genomics Centre, Wageningen University & Research, 6709 PG Wageningen, The Netherlands
[6]Department of Animal Ecology (AnE), Netherlands Institute of Ecology (NIOO-KNAW), 6708 PB Wageningen, The Netherlands
[7]Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, Ulls väg 26, Ultuna, Uppsala, Sweden

Summary

Carcass fat content is an economically important trait in commercial chickens. The use of genome-wide high density SNPs may improve the power and resolution to identify QTLs, putative candidate genes and copy number variations (CNVs), for selection programs. The main goal of this study was to identify genomic windows and putative candidate genes for carcass fat content. We checked the overlap of QTL with regions demonstrating signatures of selection and inherited CNVs identified in the same population. A total of 497 42 day-old chickens from the EMBRAPA F_2 Chicken Resource Population developed for QTL studies were genotyped with the 600K SNP genotyping array (Affymetrix®), and phenotyped for carcass fat content weight (CFCW) and carcass fat content on a dry matter basis (CFCDM). After quality control, a total of 480 samples and 371,557 SNPs annotated in autosomal chromosomes (GGA1-28) based on Gallus_gallus-5.0 (NCBI) were kept for further analysis. GWAS analyses were performed with GenSel software using BayesB method (π=0.9988) to identify genomic windows associated with CFCW or CFC%. We identified 15 genomic windows associated with CFC% on GGA1, 7, 15, 20 and 28, and from those, we identified two adjacent windows on GGA7 considered as the same QTL explaining 1.31 and 2.18% of the genetic variance for CFCW and CFC%, respectively. This QTL overlapped with one regions previsiouly know to regulate abdominal fat in chickens and the QTL region encompassed two putative candidate genes overlapping with signatures of selection and inherited CNVs. Our findings are helpful to better understand the genetic regulation of fatness in chickens.

Keywords: cnv, gwas, selection signatures, 600k snp genotyping array

Introduction

Excess fat deposition in chickens, especially in the abdominal region, the skin and the carcass, is a negative factor for production because it reduces feed efficiency and also the nutritional value of chicken parts and, consequently, their commercial value (Jennen et al. 2004, Zhou et al. 2007). Higher fat deposition has been observed in chickens selected for rapid growth and muscle deposition.

The discovery of genomic regions associated with CFC% in chickens is a first step in the identification of putative candidate genes (PCGs) responsible for this trait. Nones et al. (2012) mapped 11 QTL for carcass fat content (six for CFCW and five for CFC%) using the EMBRAPA F_2 Chicken Resource Population and microsatellite markers; the same population used in this study. However, the QTL regions spanned tens of centimorgans (cM).

The use of a high density SNP panel provides higher resolution for QTL mapping and identification of copy number variations (CNVs). Thus, the main goal of this study was to identify genomic windows and putative candidate genes for carcass fat content traits. Additionally, we also used sequencing data from the parental animals for identification of selection signatures and integrated all the data to identify putative candidate genes.

Material and methods

Samples, genotyping and phenotype measure

A total of 497 42 days-old chickens from the EMBRAPA F_2 Chicken Resource Population developed for QTL studies were genotyped with the 600K SNP genotyping array (Affymetrix®). After genotyping, the samples were filtered for DishQC≥0.82 and call rate≥90% and the SNPs were limited to those with autosomal locations on chromosomes (GGA1-28) based on Gallus_gallus-5.0 (NCBI), then filtered for call rate≥98%, MAF≥2% and HWE p-value (P<0,000001). A total of 480 samples and 371,557 SNPs were kept for further analysis.

In this population, fat (ether extract) was measured by near-infrared reflectance spectroscopy (NIRS), estimated as percentage of the sample's weight (250 g of ground and homogenized carcass). Carcass fat content weight (CFCW) was estimated multiplying the percentage in the sample by body weight at 42 days-of-age. Carcass fat content, dry matter basis (CFC%) was estimated dividing sample fat by carcass dry matter content and after, multiplying by 100. Find more details are in Nones et al. (2012).

Genome-wide association analysis (GWAS) and overlap with known QTLs

The GWAS analyses were performed using GenSel software (Fernando & Garrick, 2013) to identify 1-Mb associated windows. We considered sex and hatch as fixed effects and BW42 as a covariate for CFCW in the linear model. We adopted $\pi = 0.9988$ in the BayesB model with 41,000 Markov Chain Monte Carlo (MCMC) interactions with the first 1,000 interactions discarded. The markers were allocated to 947 1-Mb non-overlapping windows. We defined as significant the windows that explained more than 0.5% of the genetic variance.

We checked the overlap of our significant genomic windows with known QTLs in chickens, using the information available at the Chicken QTL database (release 33, 2017). We used the BED file from the Chicken QTL database.

Overlapping with selection signature regions and inherited CNVs

We compared the significant windows with selection signature regions identified in a previous study that was performed with 28 parental chickens (14 TT and 14 CC) from the population analyzed in our study (Boschiero et al., 2017, unpublished results). That analysis was performed using SNPs identified from sequencing data using Fst method (Weir & Cockerham, 1984) to estimate the divergence between populations (TT vs CC). Those analyses were performed based on Gallus_gallus-4.0 assembly. We used CrossMap tool (http://crossmap.sourceforge.net/) to convert the selection signature regions coordinates (genomic positions) to the new chicken genome assembly (Gallus_gallus-5.0, NCBI).

We compared putative candidate genes with inherited CNVs identified in the same population (EMBRAPA F_2 Chicken Resource Population) using genotypic information from 600K SNP genotyping array (Godoy et al., 2017, unpublished results). CNVs were called based on Gallus_gallus-5.0 assembly and using PennCNV (Wang et. al., 2007). That analysis integrated the Log R Ratio (LRR) and B Allele Frequency (BAF) values with family information (sire, dam and offspring). The LRR values were corrected for genomic wave bias of SNP-arrays (Diskin et. al., 2008), considering a genomic window of 1 kb.

In the filter step, two types of CNVs were removed from the result: (i) CNVs smaller than 1 kb and (ii) duplicated CNVs (same event in the same parental). The CNV detected in the offspring that was identified in at least one of the parents was defined as inherited.

Results and discussion

Genome-wide association analysis (GWAS) and overlap with known QTLs

We identified 15 genomic windows associated with CFCW or CFC% in chickens on GGA1, 7, 15, 20 and 28, with the proportion of genetic variance explained ranging from 0.54% to 1.71% and the posterior probability of association (PPA) ranging from 0.67 to 0.85. From the 15 windows, four were associated with both traits, seven were associated only with CFCW and 4 were associated only with CFC%. Among those windows associated with both traits, two adjacent windows on GGA7 were also considered as representing the same QTL and their respective proportion of genetic variance explained were summed, as in Table 1.

Table 1 - Characterization of GGA7 at 35 - 37 Mb window.

GGA (Mb)[1]	SNP window (first – last position)	#SNP/ window	Proportion of genetic variance explained by the SNP window (trait)
7 (35 – 36)	35,001,761 – 36,898,384	643	1.31 (CFCW)
			2.18 (CFC%)

[1] Map position based on Gallus_gallus-5.0, NCBI assembly.

The QTL located on GGA7 at 35 - 36 Mb harbored 36 annotated genes and overlapped with 11 different published QTL for production traits, including one related to fatness (abdominal fat weight, QTL #2167; Park et al., 2006). This overlap corroborates our finding that this QTL is associated with fat deposition in chickens.

Overlap with selection signature regions, inherited CNVs and putative candidate genes (PCGs)

The QTL located on GGA7 (Table 1) overlapped a selection signature region that harbors two genes: *NR4A2* and *GPD2*. The *NR4A2* gene belongs to a receptor superfamily (NR4A) and according to Han et al. (2012), these receptors may regulate hepatic glucose and, consequently, lipid metabolism. The *GPD2* gene is a mitochondrial dehydrogenase that affects gluconeogenesis and glucose homeostasis (Madiraju et al.; 2014). Brown et al. (2002) reported a reduction of 40% in the weight of white adipose tissue in *GPD2* knock-out mice. The overlap of these genes with this selection signature region might indicate that selection have affected glucose homeostasis, lipid metabolism and consequently lipid storage.

Additionally, this QTL overlapped with 11 inherited CNVs and two of those overlapped with *NR4A2* and *GPD2* genes. Both CNVs are partial duplications that were inherited from dam to offspring, but were not present in the sire. The CNV that overlapped with *NR4A2* was detected in two different families, while the CNV that overlapped with *GPD2* was detected in only one family. The CNV sizes were 14,460 bp (36,214,618 - 36,229,078 bp) and 10,733 bp (36,273,167 – 36,283,900 bp), respectively. Figure 1 shows the overlap between a selection signature region, two CNVs and the two candidate genes within the QTL mapped on GGA7.

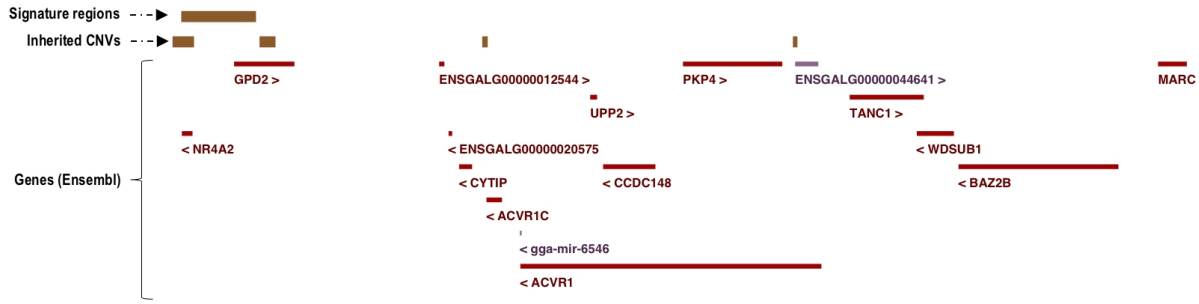

Figure 1 – Plot of one small interval within the QTL mapped on GGA7 between 35 and 36 Mb showing the overlapping of the signature of selection region, inherited CNVs and putative candidate genes.

Conclusions

We identified a total of 15 QTLs that were associated with carcass fat content traits in chickens. One of these QTL, covering 2 Mb on GGA7, was associated with CFCW and CFC%. For this QTL, we identified two putative candidate genes for fat deposition, which overlapped with a selection signature region and two inherited CNVs. These results obtained using an F_2 Chicken Resource Population developed for QTL mapping are important to achieve a better understand of the genetic regulation of fatness in chickens. The use of our findings in poultry breeding programs may improve selection accuracy consequently leading to the production of chickens with lower fat deposition in the carcass.

List of References

Brown, L. J., R. A. Koza, C. Everett,, M. L. Reitman, L. Marshall, L. A. Fahien, L. P. Kozak, M. J. MacDonald, 2002. Normal thyroid thermogenesis but reduced viability and adiposity in mice lacking the mitochondrial glycerol phosphate dehydrogenase. J. Biol. Chem. 277: 32892-32898.

Chicken QTL database (release 33). Available at: <http://www.animalgenome.org/cgi-bin/QTLdb/GG/index>. Assessed in September 18, 2017.

Diskin, S. J., M. Li, C. Hou, S. Yang, J. Glessner, H. Hakonarson, M. Bucan, J.M. Maris, K. Wang, 2008. Adjustment of genomic waves in signal intensities from whole-genome SNP genotyping platforms. Nucleic Acids Res., 36: e126.

Fernando, R. L., D. J. Garrick, 2013. Implementing a QTL detection study (GWAS) using genomic prediction methodology. Methods Molecular Biology, 1019: 275–298.

Han, Y.-F., & G.-W. Cao, 2012. Role of nuclear receptor NR4A2 in gastrointestinal inflammation and cancers. World Journal of Gastroenterology : WJG, 18(47): 6865–6873.

Jennen, D. G. J., L. J. Vereijken, H. Bovenhuis, R. P. M. Crooijmans, A. Veenendaal, J. J. Van der Poel, M. a. M. Groenen, 2004. Detection and localization of quantitative trait loci affecting fatness in broilers. Poultry Science, 3(83): 295–301.

Madiraju, A. K., D. M. Erion, Y. Rahimi, X.-M. Zhang, D. Braddock, R. A. Albright, et al., 2014. Metformin suppresses gluconeogenesis by inhibiting mitochondrial glycerophosphate dehydrogenase. Nature, 510(7506): 542–546.

Nones, K., M.C. Ledur, E. L. Zanella, C. Klein, L.F.B. Pinto, A.S.A.M.T. Moura, L.L. Coutinho, 2012. Quantitative trait loci associated with chemical composition of the chicken carcass, Animal Genetics, 43:570–576.

Park, H.-B., L. Jacobsson, P. Wahlberg, P.B. Siegel, L. Andersson, 2006. QTL analysis of body composition and metabolic traits in an intercross between chicken lines divergently selected for growth. Physiological Genomics, 2(26): 216–223.

Zhou, H., C.M. Evock-Clover, J.P. McMurtry, C.M. Ashwell, S.J. Lamont, 2007. Genome-wide linkage analysis to identify chromosomal regions affecting phenotypic traits in the chicken. Poultry Science, 2(86): 267-276.

Wang, K., M. Li, D. Hadley, R. Liu, J. Glessner, S.F. Grant, H. Hakonarson, M. Bucan, 2007. PennCNV: an integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. Genome Res. 17:1665–1674.

Weir, B. S., C. C. Cockerham, 1984. Estimating F-statistics for the analysis of population structure. Evolution, 38:1358–1370.