University of São Paulo
"Luiz de Queiroz" College of Agriculture

Association of predicted deleterious single nucleotide polymorphisms with carcass traits in meat-type chickens

Priscila Anchieta Trevisoli

Dissertation presented to obtain the degree of Master in Science. Area: Animal Science and Pastures

Piracicaba
2018

Priscila Anchieta Trevisoli
Biotechnologist

Association of predicted deleterious single nucleotide polymorphisms with carcass traits in meat-type chickens

versão revisada de acordo com a resolução CoPGr 6018 de 2011

Advisor:
Prof. Dr. **LUIZ LEHMANN COUTINHO**

Dissertation presented to obtain the degree of Master in Science. Area: Animal Science and Pastures

Piracicaba
2018

Dados Internacionais de Catalogação na Publicação
DIVISÃO DE BIBLIOTECA – DIBD/ESALQ/USP

Trevisoli, Priscila Anchieta

Association of predicted deleterious single nucleotide polymorphisms with carcass traits in meat-type chickens / Priscila Anchieta Trevisoli. - - versão revisada de acordo com a resolução CoPGr 6018 de 2011. - - Piracicaba, 2018.

35 p.

Dissertação (Mestrado) - - USP / Escola Superior de Agricultura "Luiz de Queiroz".

1. Estudo de associação 2. SNPs não sinônimos 3. *Target sequencing* 4. Frangos I. Título

DEDICATION

To the most wonderful parents, Rogério and Francismar, who do not measure efforts on helpping me to persue my dreams during my hole life, and who I will always love.

I dedicate.

ACKNOWLEGDMENTS

First, to Prof. Dr. Luiz Lehmann Coutinho for the opportunity, incentive, trust and guidance.

To Embrapa Suínos e Aves, especially Dr. Mônica Corrêa Ledur for the partnership and substantial collaborations.

To the University of São Paulo (USP), "Luiz de Queiroz" College of Agriculture (ESALQ) and Animal Science Department for the opportunity, support and teaching.

To "Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq)" and São Paulo State Research Foundation (FAPESP) (Process 2016/13589-0) for the scholarships.

To my colleagues from Animal Biotechnology Laboratory and animal science department Aline, Bruna, Clarissa, Fábio, Francisco, Gabriel, Gabriella, Juliana, Mirele, Mayara, Natalia, Otávio, Shilton, Thaís, for all the help, moments of distraction and research. To the lab technicians Jorge, Horacio, Marcela, Nirlei, Pilar and Ricardo for all the patience and help.

To all the RDS girls, that even we meeting in rare occasions, the support, deep conversations and friendship remains the same! To my friends, from UFSCar Ana Carolina, Maria Paula and Marianna, from Goiânia Jéssika e Marcela, from my hometown Fernanda that in some way helped and supported me during these years. To Elis, Mayara and Rayza for all the positivity, happiness and crossfit trainings. And to my dears roommates Auri, Giuliana, Laura and Marina, for the happy and complaining moments.

To my boyfriend Lucas, for immense support, encouragement, comfort, understanding, affection and immeasurable love. I could write for hours… so, Thank you!

To my sisters, Talitha and Marina for the incentive, loughs, talks and affection.

To my parents, for endeless support, kindness, LOVE … and literally everything! Any words or acknowledgments are not enough!

To God for giving me life!

THANK YOU!

Priscila Anchieta Trevisoli

EPIGRAPH

"Science investigates; religion interprets.
Science gives man knowledge, which is power;
religion gives man wisdom, which is control.
Science deals mainly with facts; religion deals mainly with values.
The two are not rivals."

Dr. Martin Luther King Jr.

SUMMARY

RESUMO

Associação de polimorfismos de base única preditos como deletérios com características de carcaça em frangos de corte

O melhoramento genético é o principal responsável pelo aumento da eficiência da produção avícola nas últimas décadas e os programas de melhoramento de aves estão direcionados para um maior rendimento de carne e eficiência alimentar. Dentre as abordagens genômicas, estudos de associação genômica ampla (GWAS) identificaram loci associados com características quantitativas (QTLs) de carcaça em uma população de frangos de corte. Análise de GWAS identifica regiões em desequilíbrio de ligação com possíveis mutações causais e com o objetivo de refinar esses resultados, estudos de associações usando polimorfismos de base única (SNPs) não sinônimos podem ser úteis. O SNP não sinônimo pode ser predito como deletério por meio do *Sorting Intolerant From Tolerant* (SIFT) score quando a alteração de aminoácidos tem o potencial de impactar a função da proteína e consequentemente pode afetar o fenótipo. Portanto, neste estudo, SNPs preditos como deletérios localizados em regiões de QTLs foram identificados e associados com peso e rendimento de coxa, sobrecoxa, gordura abdominal e peito de frangos de corte. Modelo misto foi utilizado, com sexo, incubação e genótipos dos SNPs como efeitos fixos e família como efeito aleatório. De 20 SNPs analisados, seis foram associados significativamente ($p<0,05$) com peso e rendimento de coxa, sobrecoxa e peito, e três deles rs736010549, rs739508259 e rs313532967 estão presentes nos genes *WDR77*, *VWA8* e *BARL*, respectivamente. Estes genes estão relacionados com processos biológicos como via de sinalização de esteroide, receptores de estrogênio e de ácidos biliares. Nossa estratégia permitiu a identificação de potenciais mutações causais associadas com crescimento e desenvolvimento muscular

Palavras-chave: Estudo de associação; SNPs não sinônimos; *Target sequencing*; Frangos;

ABSTRACT

Association of predicted deleterious single nucleotide polymorphisms with carcass traits in meat-type chickens

Breeding has been the mainly responsible for the increase of poultry efficiency in the last decades. The breeding programs are geared towards higher meat yield and feed efficiency. Among the used genomic approaches, genome wide association studies (GWAS) identified quantitative trait loci (QTLs) associated with carcass traits in a meat-type population (TT Reference Population). GWAS analysis identifies variants in linkage disequilibrium with the possible causal mutation and with the aim of refining these results, association study with missense single nucleotide polymorphisms can be useful. A missense SNP can be predicted as deleterious via Sorting Intolerant From Tolerant (SIFT) score when the amino acid change has the potential to impact the protein function and consequently may affects the phenotype. Therefore, in this study, predicted deleterious SNPs within QTLs regions were identified and associated with thigh, drumstick, abdominal fat and breast weight and their yields. Mixed model was used with sex, incubation and SNPs genotypes as fixed effects and family as random effect. From the 20 SNPs analyzed, six were significantly associated ($p < 0.05$) with weight and yield of thigh, breast and drumstick. Three of them s736010549, rs739508259 and rs313532967 are located in the genes *WDR77*, *VWA8* and *BARL*, respectively. These genes are involved in biological process as steroid hormone signaling pathway, estrogen binding, and regulation of cell proliferation. We determined these genes as candidates for muscle growth. Our strategy allowed the identification of potential causal mutations associated with muscle growth and development.

Keywords: Association study; Missense SNPs; Target sequencing; Broilers

1. INTRODUCTION

The constant increase of global population demands a higher production of animal protein in a sustainable way. Since the beginning of the production of broilers in Brazil, the activity has been modernizing in order to improve the avian sector operation, that is, reduce costs and increase productivity (Avila *et al.* 2007). Brazilian poultry farming is a sector of great economic and social importance, and according to '*União Brasileira de Avicultura'* (Ubabef, 2017), the sector contributes to 1.5% of the National Gross Domestic Product and employs more than 3.6 million people.

Since 2004, Brazil is the major chicken meat exporter, and in 2015 the unprocessed chicken meat was the fourth most exported product of Brazil, behind of soybeans, iron ore and crude oil (Avisite, 2016). The Organization for Economic Cooperation and Development (ODCE/FAO, 2016) published a report estimating that until 2024 the increase of world poultry sector will be 24%, which will lead to a global production of 26 million tons more when compared to 2014.

Conventional breeding was the major factor responsible for poultry production efficiency increase over the last decades (Yang *et al.* 1999) and they are focused on performance and carcass traits improvement (Berri *et al.* 2001). Next generation sequencing and the development of high-density single nucleotide polymorphisms (SNP) panels are important genomic tools for the discovery of novels genes and also for alleles identification related to animal production, thus, contributing to a more accurate breeding based on the genotypes (Coutinho *et al.* 2010).

In Brazil, the Chicken breeding program of EMBRAPA Swine and Poultry developed a meat-type population (TT Reference Population) for genomic studies. The population was originated from a pure parental meat-type line (TT) that has been under selection since 1992 (Fornari *et al.* 2014). The selection aimed to improve body weight, feed conversion, carcass part yield, viability, carcass yield, fertility, hatchability and also reduce abdominal fat (Venturini *et al.* 2014). The TT Reference Population was developed with the mating of 20 males with 92 females (1:5) in 5 hatches, resulting in approximately 1,500 offsprings (Grupioni *et al.* 2017).

In previous studies of our group, several genomic association studies were conducted. Quantitative trait loci (QTL) were mapped for tibia length, width and weight (Ragognetti *et al.* 2015). Single nucleotide polymorphism (SNP) in the calpain 1 (*CAPN1*) gene was associated with body weight at 42 days, eviscerated weight and leg weight, and, other SNP in the calpain 3 (*CAPN3*) gene was associated with shear force, lightness content, and meat

water loss by drip and thawing (Felício *et al.* 2013). A SNP in the osteoprotegerin (*OPG*) gene was associated with body weight at 21 days, leg muscle yield and tibia weight (Fornari *et al.* 2014). Grupioni and collaborators (2017) associated one SNP located in the runt related transcription factor 2 (*RUNX2*) gene with several carcass traits (e.g. chilled femur weight, body weight at 21 days, wing weight, and thigh wing weight), and one SNP located in the TNF superfamily member 11 (*TNFSF11*) gene was associated with heart weight, body weight, weight gain from 35 to 41 days and feed conversion from 35 to 41 days.

Our research group in partnership with Embrapa Swine and Poultry National Research Center have been conducting several studies on poultry genomics, as quantitative trait loci (QTL) mapping, identification of candidate genes and association of single nucleotide polymorphism (SNP) for important economic traits as mentioned above, as well, genome wide association analysis (GWAS). Regarding GWAS, our recent study using 600k SNP chip, identified genomic windows associated with weight and yields of breast, thigh, drumstick and abdominal fat traits (Moreira, et al., unpublished data).

Genome wide association studies (GWAS) implies that the studied genetics variations (SNPs) are in linkage disequilibrium with the causative mutation, and consequently does not identify the causal mutation (Spain and Barrett 2015). Therefore, single association studies are helpful for the causal mutation discovery. SNP is the most common and simple genetic variation in genome (Ahmadian *et al.* 2000; Vignal *et al.* 2002). SNP can occur in intergenic, up and downstream of the genes, intron, 5'UTR and 3'UTR regions (non-coding) and exons (coding). When located in exon it can be classified as synonymous (amino acid is not changed) and missense (amino acid is changed), and missense SNPs can be predicted as tolerated and deleterious by SIFT tool (Ng and Henikoff 2003). The prediction is based on the degree of the proteins evolutionary conservation (Ng and Henikoff 2003; Kumar *et al.* 2009) and means that the amino acid change probably affects the protein structure and consequently, protein function.

The most recognized example of a missense SNP affecting animal production is the $Arg^{615} \rightarrow Cys^{615}$ mutation in *CRC* (calcium release channel) gene in pigs. This nucleotide change affects the skeletal muscle calcium regulation causing halothane sensitivity which has a high impact on meat quality (Harbitz *et al.* 2009). In chickens, some studies reported missense SNPs associations with body weight, body size and abdominal fat weight traits (Han *et al.* 2012), semi eviscerated carcass weight, eviscerated carcass weight and leg muscle weight (Wang *et al.* 2015), and body weight on different days of age, breast muscle color and fat traits (Nie *et al.* 2010).

However, there are no studies in literature that performed this association analysis with predicted deleterious and carcass traits in chickens. And, as described above predicted deleterious SNP association studies may be an important step for causal mutation discovery. Therefore, our studied focused in predicted deleterious SNPs association analysis with weight and yield of thigh, abdominal fat, breast and drumstick traits, in a meat-type population TT. We identified six deleterious SNPs associated with the studied traits and we suggested some candidate genes for carcass traits and muscle growth, but we are aware that further functional studies are necessary to confirm our findings.

References

Ahmadian, A., B. Gharizadeh, A. C. Gustafsson, F. Sterky, P. Nyrén *et al.*, 2000 Single-Nucleotide Polymorphism Analysis by Pyrosequencing. Anal. Biochem. 280: 103–110.

Avisite. A Revista do AviSite Dez/2015. a.VIII: 99:26-27.

Avila, V. S., A. Kunz, D. P. Paiva, F. R. . Jaenisch, I. M. Trevisoli *et al.*, 2007 Boas práticas de produção de frangos de corte. - Portal Embrapa.:

Berri, C., N. Wacrenier, N. Millet, and E. Le Bihan-Duval, 2001 Effect of Selection for Improved Body Composition on Muscle and Meat Characteristics of Broilers from Experimental and Commercial Lines. Poult. Sci. 80: 833–838.

Botstein, D., and N. Risch, 2003 Discovering genotypes underlying human phenotypes: past successes for mendelian disease, future approaches for complex disease. Nat. Genet. 33: 228–237.

Coutinho, L. L., M. F. do Rosário, and E. C. Jorge, 2010 *Biotecnologia animal*. Instituto de Estudos Avançados da Universidade de São Paulo.

Felício, A. M., C. Boschiero, J. C. C. Balieiro, M. C. Ledur, J. B. S. Ferraz *et al.*, 2013 Identification and association of polymorphisms in CAPN1 and CAPN3 candidate genes related to performance and meat quality traits in chickens. Genet. Mol. Res. 12: 472–482.

Fornari, M. B., R. Zanella, A. M. Ibelli, L. T. Fernandes, M. E. Cantão *et al.*, 2014 Unraveling the associations of osteoprotegerin gene with production traits in a paternal broiler line. Springerplus 3: 682.

Grupioni, N. V., N. B. Stafuzza, A. B. Carvajal, A. M. G. Ibelli, J. O. Peixoto *et al.*, 2017 Association of RUNX2 and TNFSF11 genes with production traits in a paternal broiler line. Genet. Mol. Res. 16:.

Han, R., Y. Wei, X. Kang, H. Chen, G. Sun *et al.*, 2012 Novel SNPs in the PRDM16 gene and their associations with performance traits in chickens. Mol. Biol. Rep. 39: 3153–3160.

Harbitz, I., T. Kristensen, M. Bosnes, S. Kran, and W. Davies, 2009 DNA sequence of the skeletal muscle calcium release channel cDNA and verification of the Arg615→ Cys615 mutation, associated with porcine malignant hyperthermia, in Norwegian Landrace pigs. Anim. Genet. 23: 395–402.

Kumar, P., S. Henikoff, and P. C. Ng, 2009 Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat. Protoc. 4: 1073–1081.

Ng, P. C., and S. Henikoff, 2003 SIFT: Predicting amino acid changes that affect protein function. Nucleic Acids Res. 31: 3812–3814.

Nie, Q. H., M. X. Fang, L. Xie, X. Shen, J. Liu *et al.*, 2010 Associations of ATGL gene polymorphisms with chicken growth and fat traits. J. Appl. Genet. 51: 185–191.

OECD/FAO. Organização para o Desenvolvimento e Cooperação Econômica/ Organização das Nações Unidas para Alimentação e Agricultura. Agricultural Outlook", OECD Agriculture Statistics 1 2 (database), Disponível em: <http://dx.doi.org/10.1787/agr-outl-data-en> Acesso em 28 Janeiro de 2018.

Ragognetti, B. N. N., N. B. Stafuzza, T. B. R. da Silva, T. C. S. Chud, N. V. Grupioni *et al.*, 2015 Genetic parameters and mapping quantitative trait loci associated with tibia traits in broilers. Genet. Mol. Res. 14: 17544–17554.

Spain, S. L., and J. C. Barrett, 2015 Strategies for fine-mapping complex traits. Hum. Mol. Genet. 24Spain, S: R111–R119.

Ubabef, 2016. Produção sustentável garante ao Brasil liderança nas exportações. Revista Avicultura Brasil (União Brasileira de Avicultura). 1:4-7. Disponível em: < http://abpa-br.com.br/files/publicacoes/938d713b69d9f25901b1d810f038272b.pdf> Acesso em 14 de Março de 2018.

Venturini, G. C., V. A. R. Cruz, J. O. Rosa, F. Baldi, L. El Faro *et al.*, 2014 Genetic and phenotypic parameters of carcass and organ traits of broiler chickens. funpecrp.com.br Genet. Mol. Res. Genet. Mol. Res 13: 10294–10300.

Vignal, A., D. Milan, M. SanCristobal, and A. Eggen, 2002 A review on SNP and other types of molecular markers and their use in animal genetics. Genet. Sel. Evol. 34: 275–305.

Wang, Y. C., R. R. Jiang, X. T. Kang, Z. J. Li, R. L. Han *et al.*, 2015 Identification of single nucleotide polymorphisms in the ASB15 gene and their associations with chicken growth and carcass traits. funpecrp.com.br Genet. Mol. Res. Genet. Mol. Res 14: 11377–11388.

Yang, A., D. A. Emmerson, E. A. Dunnington, and P. B. Siegel, 1999 Heterosis and Developmental Stability of Body and Organ Weights at Hatch for Parental Line Broiler Breeders and Specific Crosses Among Them. 78: 942–948.

2 ASSOCIATION OF PREDICTED DELETERIOUS SINGLE NUCLEOTIDE POLYMORPHISMS WITH CARCASS TRAITS IN MEAT-TYPE CHICKENS

ABSTRACT

Breeding programs have been focused mainly in performance traits improvement and in the last years, genomic approaches have been used for the understanding of genomic regions associated with these traits. Genome wide association (GWAS) study was performed, and quantitative trait loci were identified (QTL) for abdominal fat, drumstick, thigh and breast traits in chickens, however, this methodology assumes that the studied variants are in linkage disequilibrium with the causal mutation and consequently not identifying it. In attempt to identify causal mutations in candidate genes for carcass traits in broilers, we selected 20 predicted deleterious SNPs (SIFT score) within the mentioned QTLs to associate with weights of thigh, abdominal fat, breast weight and drumstick traits and their yields. Associated SNPs were tested for additive, dominance and allele substitution effects. From the 20 SNPs analyzed, we identified six SNPs with significant association ($p < 0.05$) with carcass traits, and three are highlighted here. The SNP rs736010549 was associated with drumstick weight and yield with significant additive and dominance effects. The SNP rs739508259 was associated with thigh weight and yield, and with significant additive and allele substitution effects. The SNP rs313532967 was associated with breast weight and yield. The three SNPs that were associated with carcass traits (rs736010549, rs739508259 and rs313532967) are located in the coding regions of the *WDR77*, *VWA8* and *BARL* genes, respectively. These genes are involved in biological process such as steroid hormone signaling pathway, estrogen binding, and regulation of cell proliferation. Our strategy allowed the identification of putative casual mutations associated with muscle growth.

Keywords: Association study; Missense SNPs; Target sequencing; Broilers

2.1 BACKGROUND

Chicken is an important source of protein for human nutrition and a model system in growth and developmental biology (Ellegren 2005). The complete genome sequence of a Red Jungle Fowl female (*Gallus gallus gallus*), that is considered the ancestor of domestic chicken (*G. g. domesticus*) (Abplanalp 1992; Cassoli 2007; Dodgson *et al.* 2011), was completed in 2004 (Hillier *et al.* 2004) and opened the opportunity to explore the molecular control of complex phenotypes such as growth and muscle deposition among other.

High throughput sequencing of several chicken lines allowed the identification of millions of single nucleotide polymorphisms (SNPs) in the chicken genome (Rubin *et al.* 2010; Boschiero *et al.* 2018) and the development of high density SNP panels (Kranis *et al.* 2013). SNPs are the most common and frequent DNA variant, with approximately 5 SNPs per kilobase (kb) in chicken (Rubin *et al.* 2010). When located in coding and regulatory regions

of genes, they may affect traits of economic interest in animal models and livestock species (Roux *et al.* 2014).

High-density SNP panels were used in genome wide association studies (GWAS) to identify genomic regions associated with quantitative traits such as body weight (Gu *et al.* 2011a), fatness traits (Sun *et al.* 2013), breast and leg muscle weight, wing weight (Xie *et al.* 2012), carcass and eviscerated weight (Liu *et al.* 2013).

GWAS relies on the linkage disequilibrium of the genetic variant present in the SNP panel and the casual mutation, so further studies are necessary to identify the mutation responsible for the phenotype of interest. In attempt to solve this issue, statistical evidences as association studies combining with functional annotations of genes and genetic variants are important to determine the causal mutation (Spain and Barrett 2015). A SNP that occurs in coding regions can be classified as missense when the coded amino acid is changed, or synonymous, when the coded amino acid remains the same. Thereby, missense SNPs can be predicted as deleterious or tolerated by SIFT tool [Sorting Intolerant From Tolerant, (Ng and Henikoff 2003)]. Changes at well-conserved positions tend to be deleterious due to the assumption that important amino acids will be conserved in the protein family. (Ng and Henikoff 2002, 2003). Thus, based on sequence homology SIFT predict a score that less than 0.05 is classified deleterious. When a SNP is predicted as a deleterious mutation, it means that the change of amino acids probably affects the protein structure and function, and consequently, can potentially alter the phenotype.

Using missense SNPs, some previous studies identified associations with body weigh at hatch, semi-eviscerated carcass weight, eviscerated carcass weight, leg muscle weigh and carcass weight (Wang *et al.* 2015), abdominal fat weight, body weight at different ages and body size traits (Han *et al.* 2012). However, there are no studies in the literature using predicted deleterious SNPs for association studies to identify casual mutations. Therefore, in this study we used previously developed whole genome sequence and GWAS information to identified predicted deleterious SNPs in QTL regions. Furthermore, we tested the association of these SNP with carcass traits in order to identify potential causal mutations in broilers.

2.2 METHODS AND MATERIALS

2.2.1 Ethics statement

In this study, all experimental protocols that used animals were performed in agreement with the resolution number 010/2012 approved by the Embrapa Swine and Poultry

Ethics Committee on Animal Utilization to ensure compliance with international guidelines for animal welfare.

2.2.2 Experimental Population

The TT reference population used for this study was generated from an Embrapa broiler line called TT. The TT line has been under selection since 1992, for many generations and several traits, with the goals to increase body weight and carcass yield, improve viability, fertility, hatchability, feed conversion, and reduce abdominal fat (Rosário et al. 2009). The TT Reference Population is an expansion of the TT line and was developed from crossing 20 males and 92 females (1:5) in five hatches, yielding approximately 1,500 chickens (Cruz et al. 2015; Marchesi et al. 2017). From this population we selected 237 offspring and 45 parental chickens (12 males and 33 females) for target sequencing. The offspring were selected based on the following criteria: (1) descendent of one of the 14 parental males that we have the whole genome sequencing data; (2) from families that have between 5 to 7 animals; (3) and from three incubations.

2.2.3 Phenotype measurement

Body weight at 42 days of age (BW42) was measured six hours after fasting. Blood samples were collected for DNA extraction during the bleeding. After bleeding, feathers were removed mechanically after a hot water bath (60ºC for 45 s). The carcass cuts as breast weight (BTW), thigh weight (THW), drumstick weight (DRW) and abdominal fat weight (ABFW) were individually measured in grams. Drumstick yield (DR%), abdominal fat yield (ABF%), thigh yield (TH%) and breast yield (BT%) were calculated as a percentage of live body weight at 42 days of age (BW42). More details about the slaughter and phenotypes measurements are available at (Venturini et al. 2014; Cruz et al. 2015).

2.2.4 SNPs selection and custom amplicon design

Predicted deleterious SNPs were selected from whole genome re-sequence data previously generated from 14 of the parental animals of the population used in this study (TT Reference population). Sequences were generated with a Illumina HiSeq and SNP identified using SAMtools v.1.2 software (Li *et al.* 2009). Further details about library preparation, sequencing and filtering are available in Moreira *et al.* (2015) and Boschiero *et al.* (2018). SNP functional annotation was performed using VEP (Variant Effect Predictor, McLaren *et*

al. 2016) and deleterious prediction was based on SIFT score prediction (Sorting Intolerant From Tolerant, Ng and Henikoff 2003). All the SNPs identified are available at EVA-EMBL database (https://www.ebi.ac.uk/ena/data/view/PRJEB25004).

In addition, a GWAS was performed by Moreira *et al.* (unpublished data) in the same meat-type chicken population (TT), and identified some QTLs associated with BTW, BT%, THW, TH%, DRW, DR%, ABFW and ABF% traits using high-density SNP chip (600K) data.

For SNPs selection, we overlapped all predicted deleterious SNPs identified in parental animals with the genomic windows identified in GWAS analysis that explained a high percentage of the additive genetic variance associated with the studied traits (Additional file 1). After, the overlapped SNPs were analyzed in Haploview software (Barrett *et al.* 2005) using Tagger pairwise method and a tagging threshold of pairwise $r^2 < 0.3$. This analysis was conducted to avoid the selection of markers accounting for the same effect due to the linkage disequilibrium (LD).

Due to the sequencing read size, we defined 150 bp around each predicted deleterious SNP as target region, with the variant located in the middle of the region. These regions were selected for target sequencing, and the amplicons designed were performed using DesignStudio online platform (Illumina Technology).

2.2.5 Target sequencing

Genomic DNA was extracted using PureLink® Genomic DNA kit (Invitrogen, Carlsbad, CA, USA) and quantified using Qubit® 2.0 Fluorometer (Thermo Fisher Scientific, Waltham, MA, USA). DNA integrity was evaluated in 1% agarose gel. Library preparation was performed according to Truseq® Custom Amplicon Low Input Kit Reference Guide (Illumina Technology). Libraries were quantified with quantitative real time PCR, using KAPA® Library Quantification kit (KAPA Biosystem) and fragments size evaluated using either Bioanalyzer® (Agilent Technologies) or Fragment Analyzer (Advanced Analytical Technologies). Paired-end sequencing with a read length of 150 bp was performed on a MiniSeq™ (Illumina Technology).

2.2.6 Sequencing data analyses, variant calling and functional annotation

Sequencing data analysis were conducted for the 282 chickens (offspring and parental generations). Raw sequencing data were aligned against the chicken reference genome

*Gallus_gallus*5.0 (NCBI) with BWA v.0.7.15 program, using BWA-MEM algorithm. For the SNP calling, we used SAMtools v.1.3.1 program (Li *et al.* 2009), with *mpileup* option (Li 2011), and mapping and base qualities Phred ≥ 20. The variant calling was performed with all 282 animals together. After the initial variant identification, the following filtering options were applied: INDEL removal, minor allele frequency (MAF) ≥ 0.05, SNP call rate ≥ 0.7, biallelic locus, sequencing depth ≥ 15 and Phred score quality ≥ 40.

After variant calling and filtration, the remained SNPs were annotated using VEP tool version 91 (McLaren *et al.* 2016) available on Ensemble v. 91 website (Zerbino *et al.* 2018), and the SIFT score was predicted.

2.2.7 Linkage disequilibrium analysis

This analysis was conducted to avoid predicted deleterious SNPs in strong LD for association analyses. Therefore, predicted deleterious SNPs were visualized in Haploview program (Barrett *et al.* 2005) and adjacent SNPs with $r^2 > 0.3$ were considered having a strong LD, and consequently one of them was randomly excluded for association analyses. After, for each SNP the Hardy-Weinberg equilibrium (HWE) were statically analyzed in Haploview.

2.2.8 Association analysis

For association analysis, Proc Mixed Procedure was used on SAS 9.4 Studio online platform (Statistical Analysis System Institute Inc., Cary, NC). Association analysis was performed with all 20 predicted deleterious SNPs together for each carcass trait, and because of that correction for multiple tests was not necessary. The model used was:

$$y = X\beta + Wa + Zu + e$$

Where y is the vector of observations for the measured phenotype; X is the incidence matrix relating the fixed effects to β; β is the vector of fixed effects, which included sex and incubation; W is the genotype matrix (coded as 0, 1 and 2; 0 and 2 for homozygous and 1 for heterozygous) for all 20 deleterious SNPs and a is the vector of SNPs fixed effects. Z is the incidence matrix relating u to y; u is the vector of the family random effect; and e is the vector of residual effects. For the weight traits, BW42 was used as a covariate. Association was considered significant at $p < 0.05$ for the F test.

Contrasts were used to compare the mean performance of one homozygote against the other and to estimate additive and allelic substitution effects. Similarly, dominance effect was estimated through the contrast of the mean performance of the heterozygote against the mean

performance of both homozygotes. Contrasts for additive and dominance effect were orthogonal between them. These analyses were performed under the same linear model detailed above with Proc Mixed Procedure on SAS 9.4 Studio online platform, considering the SNPs that presented the three genotypes (0, 1 and 2). Estimates and contrasts were set based on the methodology defined by Falconer and Mackay (1996). Effects were considered significant for $p < 0.05$ in the F test.

2.3 RESULTS

2.3.1 Phenotype measures

The summary statistics for BW42, THW, TH%, ABFW, ABF%, DRW, DR%, BTW and BT% are given in Table 1.

Table 1. Number of animals (N), mean, standard deviation (SD), minimum and maximum values for body weight at 42 days of age (BW42), thigh weight (THW), thigh yield (TW%), abdominal fat weight (ABFW), abdominal fat yield (ABF%), drumstick weight (DRW), drumstick yield (DR%), breast weight (BTW) and breast yield (BT%).

Phenotype	N	Mean	SD	Minimum	Maximum
BW42 (g)	237	2219.75	254.87	1310.00	2816.00
THW (g)	237	203.75	30.27	110.80	277.80
TH%	237	9.16	0.63	7.26	11.64
ABFW (g)	237	50.93	14.63	19.00	91.00
ABF%	237	2.29	0.61	1.01	4.25
DRW (g)	237	300.10	44.63	161.60	419.20
DR%	237	13.80	0.89	11.81	16.35
BTW (g)	237	495.39	61.93	260.00	660.00
BT%	237	22.32	1.31	18.28	26.51

2.3.2 SNP selection and amplicon design

The whole genome re-sequencing from 14 parental chickens of the studied meat-type population identified approximately 11 million SNPs across the genome, and after functional annotation 4,708 of them were predicted as deleterious SNPs. As the result of the overlap between these 4,708 SNPs and the six selected genomic windows identified in GWAS, 89 predicted deleterious SNPs were kept for the further step. After the selection of SNPs with $r^2>0.3$, 19 uncorrelated SNPs remained for the amplicon design using DesignStudio online platform. The final amplicon panel had 99% of coverage.

2.3.3 Sequencing and variant calling

Libraries sequencings from MiniSeq produced an average number of raw reads of 298,791.91 per sample. The average of overall mapping rate of the raw reads against to *Gallus_gallus*5.0 (NCBI) genome assembly was 99.74%. After variant calling, 1,957 variants (including SNPs and INDELs) were initially detected, and 195 SNPs remained after filtration. The average depth of the remained SNPs was 7,961.21 reads.

The functional annotation was performed for the 195 SNPs. As shown in Table 2, 29 SNPs were annotated as novel variants. From the 195 SNPs, 26% were in intronic regions, 20% were classified as missense and 14% were synonymous variants.

As already mentioned, missense SNPs can be predicted as deleterious or tolerated based on SIFT score (Ng and Henikoff 2003). In our study, from the 56 missense SNPs identified, 26 were classified as tolerated, 21 as deleterious, and 9 had no prediction (Table 2).

Table 2. Number of novel and existing variants, and classification of functional annotation of SNPs performed with Variant Effect Predictor (VEP) online platform.

Variant	**Number of SNPs**
Novel	29
Existing	166

Variant	**SIFT Prediction**	**Number of SNPs**
Missense	Deleterious	21
	Tolerated	26
	No prediction	9
Intron		70
Intergenic		25
Synonymous		45
5' URT		1
Downstream gene		32
Upstream gene		35
Splice		6

2.3.4 Linkage disequilibrium verification

As presented in Figure 1, only one region of adjacent SNPs had $r^2 > 0.3$ ($r^2 = 1.0$, black square – chr26:3747344 and chr26: 3747346). Thus, chr26:3747344 was excluded, remaining 20 predicted deleterious SNPs for further analysis.

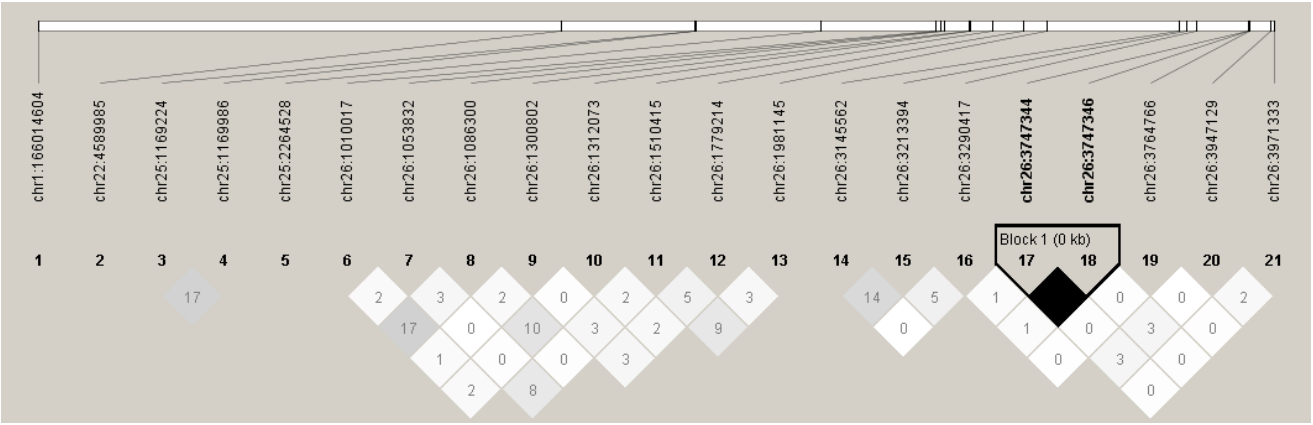

Figure 1. Linkage disequilibrium plot of the filter remained deleterious SNPs. The number inside of each square represents r^2 value (x 100). The gradient color also represents r^2 values, white is 0 and black is 1.

The detailed information of the selected 20 predicted deleterious SNPs about genome position, SNP ID, located gene, alleles and genotypes frequencies, HWE test and SIFT score are presented in Tables 3. Two SNPs do not have *rs* ID, and four genes were considered as novel genes, therefore, the ensemble gene ID was also presented. Seven SNPs did not have any animal genotyped with the alternative homozygous and, five SNPs were significant for HWE test.

Table 3. Deleterious SNPs selected for the association analyses with carcass traits.

GGA[1]	Position	SNP ID	Gene Symbol	Ensembl Gene ID	R/A[2]	Allele Frequency		Genotype Frequency						HW p-value[6]	SIFT[7]
								HREF[3]		HT[4]		HALT[5]			
						R	A	Freq	N	Freq	N	Freq	N		
1	166,014,604	rs739508259	*VWA8*	ENSGALG00000016955	G/C	0.588	0.412	0.364	86	0.449	106	0.186	44	0.6262	0.01
22	4,589,985	rs314536739	*ANXA4*	ENSGALG00000038783	C/T	0.738	0.262	0.531	126	0.413	98	0.054	13	0.4782	0.02
25	1,169,224	rs312547749	Novel Gene	ENSGALG00000027316	C/T	0.435	0.565	0.181	43	0.506	120	0.312	74	0.2798	0.00
25	1,169,986	rs737797683	*CRNN*	ENSGALG00000027316	G/C	0.808	0.192	0.666	158	0.282	67	0.050	12	0.2771	0.00
25	2,264,528	rs739048621	Novel Gene	ENSGALG00000014643	G/A	0.947	0.052	0.894	212	0.105	25	0.000	0	0.7875	0.00
26	1,010,017	c.482C>T	*MYBPH*	ENSGALG00000000164	C/T	0.764	0.236	0.616	146	0.295	70	0.088	21	0.0071*	0.04
26	1,053,832	c.383C>T	*CEPT1*	ENSGALG00000000142	C/T	0.941	0.059	0.881	209	0.118	28	0.000	0	0.7402	0.02
26	1,086,300	rs312325687	Novel Gene	ENSGALG00000000104	A/G	0.639	0.361	0.443	105	0.392	93	0.164	39	0.0124	0.01
26	1,300,802	rs314560661	*AHCYL1*	ENSGALG00000000329	T/C	0.941	0.059	0.881	209	0.118	28	0.000	0	0.7875	0.01
26	1,312,073	rs14297872	*STRIP1*	ENSGALG00000037995	C/T	0.932	0.068	0.865	205	0.135	32	0.000	0	0.6067	0.01
26	1,510,415	rs741234441	Novel Gene	ENSGALG00000000477	C/T	0.605	0.395	0.320	76	0.569	135	0.109	26	2.101E-6*	0.01
26	1,779,214	rs733369312	*PPP1R15B*	ENSGALG00000000611	C/A	0.624	0.376	0.299	71	0.649	154	0.050	12	5.457E-13*	0.00
26	1,981,145	rs731705610	*CNTN2*	ENSGALG00000000653	C/T	0.947	0.053	0.894	212	0.105	25	0.000	0	0.9367	0.02
26	3,145,562	rs738655377	Novel Gene	ENSGALG00000028858	A/G	0.863	0.137	0.725	172	0.274	65	0.000	0	1.0E-4*	0.00
26	3,213,394	rs736010549	*WDR77*	ENSGALG00000040864	A/T	0.810	0.190	0.628	149	0.362	86	0.008	2	0.1768	0.01
26	3,290,417	rs737237434	*DDX20*	ENSGALG00000001504	A/G	0.780	0.220	0.605	143	0.347	82	0.046	11	0.8671	0.01
26	3,747,346	rs14300225	*PTPN22*	ENSGALG00000021656	C/T	0.084	0.916	0.004	1	0.160	38	0.835	198	0.7153	0.00
26	3,764,766	rs739340698	*AP4B1*	ENSGALG00000035295	C/T	0.935	0.065	0.873	207	0.122	29	0.004	1	1.0	0.02
26	3,947,129	rs313532967	*BARL*	ENSGALG00000002170	A/G	0.736	0.264	0.493	117	0.485	115	0.021	5	5.410E-5*	0.00
26	3,971,333	rs741234600	*SYCP1*	ENSGALG00000002511	A/C	0.950	0.050	0.881	209	0.118	28	0.000	0	0.5256	0.03

[1]GGA: *Gallus gallus* chromosome; [2] R: reference allele and A: alternative allele; [3] HREF: homozygous of reference allele. [4]HT: heterozygous. [5] HALT: homozygous of alternative allele. [6]HW p-value: Hardy-Weinberg. [7]SIFT score predicted in functional annotation. *Significant P-value <0.05.

2.3.5 Association analysis, additive and dominance effects

From the 20 predicted deleterious SNPs studied, six were significantly associated (p-value<0.05) with at least one carcass trait. Three SNPs (rs737797683, rs313532967 and rs741234600) were associated with breast traits; other three (rs739508259, rs312325687 and rs741234600) were associated with thigh traits; and one SNP (rs736010549) was associated with drumstick weight. No SNP was associated with abdominal fat traits. Detailed results for all association tests (p-values) are presented in Table 4.

Table 4. P-values of association analyses results between carcass traits and SNPs.

SNP ID	N[1]	BW42	BTW	BT%	THW	TH%	DRW	DR%	ABFW	ABF%
rs739508259	234	0.2221	0.8332	0.8003	0.0098*	0.0146*	0.7289	0.6781	0.2133	0.1428
rs314536739	234	0.2297	0.2395	0.1313	0.1342	0.1393	0.4548	0.3859	0.1078	0.5586
rs312547749	234	0.3428	0.1884	0.2211	0.8707	0.7415	0.3945	0.3945	0.5279	0.5510
rs737797683	234	0.1798	0.0532	0.0346*	0.0573	0.0487*	0.1075	0.0959	0.7872	0.7077
rs739048621	234	0.6947	0.5004	0.5365	0.5458	0.5398	0.5261	0.5375	0.3414	0.4976
c.482C>T	234	0.6547	0.2735	0.3831	0.7965	0.6477	0.7273	0.8108	0.9809	0.8872
c.383C>T	234	0.0042*	0.8124	0.4050	0.3775	0.1619	0.3114	0.2233	0.9014	0.2758
rs312325687	234	0.1343	0.1050	0.1199	0.0354*	0.0220*	0.9174	0.9382	0.9506	0.9583
rs314560661	234	0.8758	0.2043	0.3390	0.3678	0.4045	0.4196	0.3606	0.5637	0.9232
rs14297872	234	0.0304*	0.5964	0.8217	0.2631	0.1209	0.4663	0.4815	0.8762	0.5523
rs741234441	234	0.2136	0.4656	0.4490	0.2668	0.2490	0.6970	0.7647	0.6032	0.9682
rs733369312	234	0.6968	0.9697	0.9677	0.1303	0.2975	0.1121	0.0505	0.6662	0.5475
rs731705610	234	0.1467	0.2037	0.1149	0.2577	0.1953	0.1164	0.1230	0.0669	0.1613
rs738655377	234	0.8992	0.1334	0.1881	0.1141	0.1789	0.8429	0.9745	0.4098	0.5450
rs736010549	234	0.9026	0.0677	0.1090	0.5474	0.6187	0.0038*	0.0052*	0.1523	0.3421
rs737237434	234	0.6837	0.1330	0.1863	0.0611	0.0885	0.5636	0.5137	0.8907	0.8882
rs14300224	234	0.8352	0.5980	0.4824	0.8976	0.9416	0.7702	0.6967	0.0934	0.1217
rs739340698	234	0.1450	0.2839	0.2143	0.4592	0.4109	0.4663	0.5506	0.7801	0.8390
rs313532967	234	0.6840	0.0144*	0.0234*	0.9630	0.9307	0.6828	0.5735	0.1644	0.1666
rs741234600	234	0.5306	0.0229*	0.0208*	0.0086*	0.0093*	0.2248	0.3559	0.3680	0.371

[1] Sample sizes. * Significant at $p < 0.05$. BW42: body weight at 42 days; BTW: breast weight; BT%: breast yield; THW: thigh weight; TW% thigh yield; DRW: drumstick weight; DR%: drumstick yield; ABFW: abdominal fat weight; ABF%: abdominal fat yield.

Additive and dominance effects were estimated only for associated SNPs that presented the three genotypes, consequently rs741234600 was not considered on these analyses. We deemed significant effects with P-value <0.05. Allele substitution effect test was performed only for significant additive effects (Table 5). Additive and allele substitution effects for rs739508259 and rs312325687 associated with THW and TH% were significant. Additive and dominance effects for rs736010549 associated with DRW and DR% were significant, but not for allele substitution test.

Table 5. P-values, estimates and standard error for additive, dominance and allele substitution effects for SNPs with respective associated traits in broilers.

SNP ID	Association		Additive Effect		Dominance Effect		Allele Substitution Effect	
	Trait	P-value	P	E (se)	P	E (se)	P	E (se)
rs739508259	THW	0.0098**	0.0029*	-5.76 (1.906)	0.7616	-0.79 (2.615)	0.0062*	5.12 (1.850)
	TH%	0.0146*	0.0042*	-0.25 (0.085)	0.6723	-0.04 (0.117)	0.0064*	0.23 (0.082)
rs737797683	BT%	0.0346*	0.0104*	-0.49 (0.188)	0.2308	-0.25 (0.212)	0.0506	0.30 (0.155)
	TH%	0.0487*	0.1966	-0.19 (0.146)	0.0144*	-0.41 (0.167)	-	-
rs312325687	THW	0.0354*	0.0124*	6.67 (2.630)	0.9484	-0.19 (2.950)	0.0126*	-6.56 (2.5853)
	TH%	0.0220*	0.0069*	0.32 (0.116)	0.9387	0.01 (0.132)	0.0063*	-0.32 (0.113)
rs736010549	DRW	0.0038*	0.0009*	-13.84 (4.132)	0.0016*	-13.80 (4.324)	0.2394	2.32 (1.964)
	DR%	0.0052*	0.0016*	-0.59 (0.183)	0.0017*	-0.61 (0.192)	0.3520	0.08 (0.088)
rs313532967	BTW	0.0144*	0.6572	-2.43 (5.475)	0.2326	6.92 (5.783)	-	-
	BT%	0.0234*	0.6708	-0.11 (0.255)	0.2634	0.30 (0.269)	-	-
rs741234600	BTW	-	-	-	-	-	0.0602	9.44 (4.999)
	BT%	-	-	-	-	-	0.0491*	0.46 (0.233)
	THW	-	-	-	-	-	0.0136*	10.09 (4.058)
	TH%	-	-	-	-	-	0.0121*	0.46 (0.182)

THW: thigh weight; TW% thigh yield; DRW: drumstick weight; DR%: drumstick yield; BTW: breast weight; BT%: breast yield. P, E and Se are p-values, estimates and standard errors for the respective analysis.

2.4 DISCUSSION

The identification of genetic markers associated with carcass weight and yield traits has been the focus of several studies due to the economic importance of these traits in broiler production. With the main goal of finding potentials causal mutations for carcass traits, this study selected predicted deleterious SNPs in chicken QTLs to be evaluated as potential causal mutations in our TT Reference Population. Here we discuss the impact of the predicted deleterious SNP and its association with carcass traits and prospect harboring genes as possible candidates for carcass traits in broilers.

The SNP rs739508259 is located in the von willebrand factor A domain containing 8 (*VWA8*) gene and within the GGA-1 at 166 Mb genomic window identified in the GWAS analysis. This region was associated with DRW and DR%, explaining 3.20 and 2.79 of the additive genetic variance respectively (Additional File 1). This SNP is a G>C nucleotide change with minor allele (C) frequency of 0.42. The nucleotide change causes the amino acids substitution of glutamine to histidine. The rs739508259 was associated with THW and TH% and had significant additive and allele substitution effects for both traits. On average, for each C allele in the animal's genotype, an increase of 5.12g was observed for THW and 0.24% for TH%, compared to the GG genotype. The window on GGA-1 at 166 Mb also explained 0.20% and 0.14% of the additive genetic variance for THW and TH%, respectively. However, these proportions were not enough to be considered significant (Additional File 1).

In mice *VWA8* gene is highly expressed in skeletal muscle, has ATPase domains, mitochondrial targeting sequences and is a mitochondrial protein (Luo *et al.* 2017). More studies are necessary in attempt to relate this gene with muscle growth in chickens.

The predicted deleterious SNP rs736010549 is located in the WD repeat domain 77 (*WDR77*) gene. Furthermore, is located within the GGA-26 at 3 Mb genomic window identified in GWAS analysis, and this region was associated with breast weight (BRW) and breast yield (BR%), representing 0.53 and 0.86 of the additive genetic variance respectively (Additional File 1). This polymorphism results in an amino acid change from serine to cysteine (A/T allele substitution), with the minor allele (T) frequency of 0.19. This SNP was associated with drumstick weight (DRW) and drumstick yield (DR%) traits, and also significant for additive and dominance effects tests for both traits. The window on GGA-26 at 3 Mb explained 0.22% and 0.25% of the additive genetic variance for DRW and DR%,

respectively. However, these proportions were not enough to be considered significant (Additional File 1).

The WD repeat domain 77 (*WDR77*) gene belongs to the WD repeat proteins that is characterized by multiple protein interactions capacity (Friesen *et al.* 2001). The protein p44 (also named as methylosome protein 50, MEP50) is coded by *WDR77*, is an androgen receptor (AR) coactivator by multiprotein complex formation (Hosohata *et al.* 2003). In humans, p44 was associated with inhibition of prostate cancer cell growth as coactivator of AR (Zhou *et al.* 2006; Gu *et al.* 2011b) and with breast cancer growth mediated trough estrogen and its receptor (Peng *et al.* 2010).

Several studies showed the inhibitory action of androgenic steroids in chickens growth (Fennell *et al.* 1990; Fennell and Scanes 1992; Esquivel-Hernandez *et al.* 2016), which is a possible consequence of the androgen receptor or estrogen receptor aromatization (Fennel *et al.* 1996; Callewaert *et al.* 2010). Kong *et al.* (2017) studied different expressed (DE) genes in a selected and unselected broiler breeds, and among their results, they suggested that inhibited AR was predicted to be an effective regulatory factor for DE genes in selected breed, corroborating with previous cited studies. That way, we suggest that this SNP (rs736010549) associated with DRW and DR%, may alters the conformation of p44, decreasing the AR activation and so contributing to growth in chickens.

The predicted deleterious rs313532967 is located in the bile acid receptor-like (*BARL*) gene and within the GGA-26 at 3 Mb genomic window identified in GWAS analysis. This region was associated with BTW and BT%, representing 0.53 and 0.86 of the additive genetic variance respectively (Additional File 1). This SNP is an A>G change, resulting in the amino acid change of asparagine to serine, and the minor allele (G) frequency is 0.264. The HWE test was significative, and this may be due to our finite population, or indicating that this locus may be under selection or inbreeding.

The *BARL* gene have a DNA-binding domain of Farnesoid X receptor (FXR) family. This domain in humans was intensively studied and when it is activated by bile acids it can regulate bile acids synthesis, conjugation and transport, consequently impacting in lipid and glucose metabolism (Claudel *et al.* 2005; Preidis *et al.* 2017). When bile acids are released in the ileum, induces the synthesis of fibroblast growth factor (FGF-19) which stimulates hepatic protein and glycogen synthesis (Kir *et al.* 2011). In an interesting work in broilers, Lai *et al.* (2018) demonstrated that dietary supplementation of swine bile acids for broiler chickens influences their growth performance and carcass characteristics as reduction of abdominal fat,

increase of carcass weight, eviscerated weight and leg weight. Therefore, our study indicates that BARL gene can be involved in growth and carcass development in chicken.

The SNP rs741234600 is located in synaptonemal complex protein 1 (*SYCP1*) gene and within the GGA-26 at 3 Mb genomic window identified in GWAS analysis. This region was associated with BTW and BT%, representing 0.53 and 0.86 of the additive genetic variance respectively (Additional File 1). In our study, this SNP was associated with BTW, BT%, THW and TH%, and is an A>C nucleotide change, resulting in the amino acid substitution of lysine to threonine, and the minor allele (C) frequency is 0.050. This polymorphism was significant for allele substitution effect for BT%, THW, TH%. Hence, each C allele on average contribute to more 0,46%, 10,09g and 0,45% for BT%, THW and TH% respectively. In mammals, the protein produced by *SYCP1* gene is the major component of the transverse filaments of synaptonemal complexes (SC) which links lateral and central elements during meiosis (Costa *et al.* 2005). Consequently, mutations of SC components can lead to meiotic arrest and aneuploidy (Page and Hawley 2004).

The SNP rs737797683 is located in the cornulin (*CRNN*) gene, within the GGA-25 at 1 Mb genomic window identified in GWAS analysis. This region was associated with BT%, BR% and ABF% representing 0.24, 0.24 and 0.23 of the additive genetic variance respectively (Additional File 1). This SNP was associated with BT% and TH%. This SNP in a G>C change, resulting in the amino acid change of aspartic acid to histidine, and the minor allele (C) frequency is 0.192. This polymorphism was significant for additive effect for BT% trait, significant for dominance effect for TH%. *CRNN* codes a Ca2+ binding protein present in upper layer of squamous epithelia, contributing to epidermal differentiation (Contzler *et al.* 2005). This gene was associated with atopic dermatitis (Trzeciak *et al.* 2017) and can play an important role in the esophageal cancer development (Xu *et al.* 2000).

The SNP rs312325687 is located in the cryptochrome 4 (*CRY4*) gene and within the GGA-26 at 1 Mb genomic window identified in GWAS analysis, previously associated with ABFW and ABF%, representing 1.06 and 0.54 of the additive genetic variance respectively (Additional File 1). This SNP was associated with THW and TH% and is an A>G change, resulting in the amino acid change of aspartic acid to glycine, being the minor allele (G) frequency equal to 0.361. This polymorphism had significant additive and allele substitution effects for both traits. The *CRY4* gene have been studied recently in chickens and were found expression in visual pigment cells, retinal ganglion cells (Watari et al. 2012) and pineal gland (Kubo et al. 2006) corroborating on its action as circadian photoceptor in photosensitive tissues (Ozturk et al. 2009).

As presented above the genes *BARL*, *WDR77* and *VWA8* were selected as candidate genes for muscle growth and development in chickens. *SYCP1*, *CRNN*, *CRY4* were not selected as candidate genes for muscle growth or carcass development in this study, because there is no information available in the literature to support this. More studies with these genes are necessary to understand their relationship with carcass and muscle growth.

Is pertinent to note that although rs738655377 was not significantly associated with any of the phenotypes tested, none of the animal were homozygous for the alternative allele and the HWE test was significant. This observation provides evidence for a lethal polymorphism when in homozygosity. This variant is within a novel gene (ENSGALG00000028858) that has gene ontology terms related to oxidoreductase activity.

In conclusions, our study identified 20 predicted deleterious SNPs in different QTLs associated with carcass traits and succeeded in associating six of them with phenotypes related to muscle growth. Three predicted deleterious SNPs associated were located in genes that we consider candidate genes for carcass and muscle weight, and development. The main limitation of our study is that it is difficult to determine if the identified mutation is the causative mutation or are in linkage disequilibrium with the real causal mutation. Under these circumstances, *in silico* tests and functional studies are necessary to contribute in causal mutation discovery.

REFERENCES

Abplanalp, H., 1992 Inbred lines as genetic resources of chickens. Poult. Sci. Rev. 4: 29–39.

Barrett, J. C., B. Fry, J. Maller, and M. J. Daly, 2005 Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 21: 263–265.

Boschiero, C., G. C. M. Moreira, A. A. Gheyas, T. F. Godoy, G. Gasparin *et al.*, 2018 Genome-wide characterization of genetic variants and putative regions under selection in meat and egg-type chicken lines. BMC Genomics 19: 83.

Callewaert, F., M. Sinnesael, E. Gielen, S. Boonen, and D. Vanderschueren, 2010 Skeletal sexual dimorphism: relative contribution of sex steroids, GH-IGF1, and mechanical loading. J. Endocrinol. 207: 127–134.

Cassoli, C. S. da S., 2007 Identificação e análise de etiquetas de seqüências expressas (ESTs) na hipófise e hipotálamo de Gallus gallus. Tese - Esc. Super. Agric. Luiz Queiroz.

Claudel, T., B. Staels, and F. Kuipers, 2005 The Farnesoid X receptor: a molecular link between bile acid and lipid and glucose metabolism. Arterioscler. Thromb. Vasc. Biol. 25: 2020–30.

Contzler, R., B. Favre, M. Huber, and D. Hohl, 2005 Cornulin, a New Member of the Fused Gene'' Family, Is Expressed During Epidermal Differentiation. J. Invest. Dermatol. 124: 990–997.

Costa, Y., R. Speed, R. Öllinger, M. Alsheimer, C. A. Semple *et al.*, 2005 Two novel proteins recruited by synaptonemal complex protein 1 (SYCP1) are at the centre of meiosis. J. Cell Sci. 118: 2755 LP-2762.

Dodgson, J. B., M. E. Delany, and H. H. Cheng, 2011 Poultry genome sequences: progress and outstanding challenges. Cytogenet. Genome Res. 134: 19–26.

Ellegren, H., 2005 The avian genome uncovered. Trends Ecol. Evol. 20: 180–186.

Esquivel-Hernandez, Y., R. E. Ahumada-Cota, M. Attene-Ramos, C. Z. Alvarado, P. Castañeda-Serrano *et al.*, 2016 Making things clear: Science-based reasons that chickens are not fed growth hormones. Trends Food Sci. Technol. 51: 106–110.

Falconer, D. S., and T. F. C. Mackay, 1996 *Introduction to Quantitative Genetics*. Pearson Education Limited.

Fennel, M. J., S. V. Radecki, J. A. Proudman, and C. G. Scanes, 1996 The Suppressive Effects of Testosterone on Growth in Young Chickens Appears to be Mediated via a Peripheral Androgen Receptor; Studies of the Anti-Androgen ICI 176,334. Poult. Sci. 75: 763–766.

Fennell, M. J., A. L. Johnson, and C. G. Scanes, 1990 Influence of androgens on plasma concentrations of growth hormone in growing castrated and intact chickens. Gen. Comp. Endocrinol. 77: 466–475.

Fennell, M. J., and C. G. Scanes, 1992 Inhibition of Growth in Chickens by Testosterone, 5 - Dihydrotestosterone, and 19-Nortestosterone. Poult. Sci. 71: 357–366.

Friesen, W. J., A. Wyce, S. Paushkin, L. Abel, J. Rappsilber *et al.*, 2001 A Novel WD Repeat Protein Component of the Methylosome Binds Sm Proteins*.

Gu, X., C. Feng, L. Ma, C. Song, Y. Wang *et al.*, 2011a Genome-wide association study of body weight in chicken F2 resource population. PLoS One 6: e21872.

Gu, Z., L. Zhou, S. Gao, and Z. Wang, 2011b Nuclear transport signals control cellular localization and function of androgen receptor cofactor p44/WDR77. PLoS One 6: e22395.

Han, R., Y. Wei, X. Kang, H. Chen, G. Sun *et al.*, 2012 Novel SNPs in the PRDM16 gene and their associations with performance traits in chickens. Mol. Biol. Rep. 39: 3153–3160.

Hillier, L. W., W. Miller, E. Birney, W. Warren, R. C. Hardison *et al.*, 2004 Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. Nature 432: 695–716.

Hosohata, K., P. Li, Y. Hosohata, J. Qin, R. G. Roeder *et al.*, 2003 Purification and Identification of a Novel Complex Which Is Involved in Androgen Receptor-Dependent Transcription. Mol. Cell. Biol. 23: 7019–7029.

Kir, S., S. A. Beddow, V. T. Samuel, P. Miller, S. F. Previs *et al.*, 2011 FGF19 as a postprandial, insulin-independent activator of hepatic protein and glycogen synthesis. Science 331: 1621–4.

Kong, B.-W., N. Hudson, D. Seo, S. Lee, B. Khatri *et al.*, 2017 RNA sequencing for global gene expression associated with muscle growth in a single male modern broiler line compared to a foundational Barred Plymouth Rock chicken line. BMC Genomics 18:.

Kranis, A., A. A. Gheyas, C. Boschiero, F. Turner, L. Yu *et al.*, 2013 Development of a high density 600K SNP genotyping array for chicken. BMC Genomics 14: 59.

Lai, W., W. Huang, B. Dong, A. Cao, W. Zhang *et al.*, 2018 Effects of dietary supplemental bile acids on performance, carcass characteristics, serum lipid metabolites and intestinal enzyme activities of broiler chickens. Poult. Sci. 97: 196–202.

Li, H., 2011 A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. Bioinformatics 27: 2987–2993.

Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan *et al.*, 2009 The Sequence Alignment/Map format and SAMtools. Bioinformatics 25: 2078–2079.

Liu, R., Y. Sun, G. Zhao, F. Wang, D. Wu *et al.*, 2013 Genome-Wide Association Study Identifies Loci and Candidate Genes for Body Composition and Meat Quality Traits in Beijing-You Chickens (B. A. White, Ed.). PLoS One 8: e61172.

Luo, M., A. E. Mengos, W. Ma, J. Finlayson, R. Z. Bustos *et al.*, 2017 Characterization of the novel protein KIAA0564 (Von Willebrand Domain-containing Protein 8). Biochem. Biophys. Res. Commun. 487: 545–551.

McLaren, W., L. Gil, S. E. Hunt, H. S. Riat, G. R. S. Ritchie *et al.*, 2016 The Ensembl Variant Effect Predictor. Genome Biol. 17: 122.

Moreira, G. C. M., T. F. Godoy, C. Boschiero, A. Gheyas, G. Gasparin *et al.*, 2015 Variant discovery in a QTL region on chromosome 3 associated with fatness in chickens. Anim. Genet. 46: 141–147.

Ng, P. C., and S. Henikoff, 2002 Accounting for human polymorphisms predicted to affect protein function. Genome Res. 12: 436–46.

Ng, P. C., and S. Henikoff, 2003 SIFT: Predicting amino acid changes that affect protein function. Nucleic Acids Res. 31: 3812–3814.

Page, S. L., and R. S. Hawley, 2004 THE GENETICS AND MOLECULAR BIOLOGY OF THE SYNAPTONEMAL COMPLEX. Annu. Rev. Cell Dev. Biol. 20: 525–558.

Peng, Y., Y. Li, L. L. Gellert, X. Zou, J. Wang *et al.*, 2010 Androgen receptor coactivator p44/Mep50 in breast cancer growth and invasion. J. Cell. Mol. Med. 14: 2780–9.

Preidis, G. A., K. H. Kim, and D. D. Moore, 2017 Nutrient-sensing nuclear receptors PPARα and FXR control liver energy balance. J. Clin. Invest. 127: 1193–1201.

Roux, P.-F., M. Boutin, C. Désert, A. Djari, D. Esquerré *et al.*, 2014 Re-sequencing data for refining candidate genes and polymorphisms in QTL regions affecting adiposity in chicken. PLoS One 9: e111299.

Rubin, C.-J., M. C. Zody, J. Eriksson, J. R. S. Meadows, E. Sherwood *et al.*, 2010 Whole-genome resequencing reveals loci under selection during chicken domestication. Nature 464: 587–591.

Spain, S. L., and J. C. Barrett, 2015 Strategies for fine-mapping complex traits. Hum. Mol. Genet. 24Spain, S: R111–R119.

Sun, Y., G. Zhao, R. Liu, M. Zheng, Y. Hu *et al.*, 2013 The identification of 14 new genes for meat quality traits in chicken using a genome-wide association study. BMC Genomics 14: 458.

Trzeciak, M., M. Sakowicz-Burkiewicz, M. Wesserling, J. Gleń, D. Dobaczewska *et al.*, 2017 Altered Expression of Genes Encoding Cornulin and Repetin in Atopic Dermatitis. Int. Arch. Allergy Immunol. 172: 11–19.

Wang, Y. C., R. R. Jiang, X. T. Kang, Z. J. Li, R. L. Han *et al.*, 2015 Identification of single nucleotide polymorphisms in the ASB15 gene and their associations with chicken growth and carcass traits. funpecrp.com.br Genet. Mol. Res. Genet. Mol. Res 14: 11377–11388.

Xie, L., C. Luo, C. Zhang, R. Zhang, J. Tang *et al.*, 2012 Genome-Wide Association Study Identified a Narrow Chromosome 1 Region Associated with Chicken Growth Traits (Z. Liu, Ed.). PLoS One 7: e30910.

Xu, Z., M.-R. Wang, X. Xu, Y. Cai, Y.-L. Han *et al.*, 2000 Novel Human Esophagus-Specific Gene C1orf10: cDNA Cloning, Gene Structure, and Frequent Loss of Expression in Esophageal Cancer. Genomics 69: 322–330.

Zerbino, D. R., P. Achuthan, W. Akanni, M. R. Amode, D. Barrell *et al.*, 2018 Ensembl 2018. Nucleic Acids Res. 46: D754–D761.

Zhou, L., H. Wu, P. Lee, and Z. Wang, 2006 Roles of the androgen receptor cofactor p44 in the growth of prostate epithelial cells. J. Mol. Endocrinol. 37: 283–300.

ADDITIONAL FILE

Additional File 1. Characterization of selected genomic windows identified in genome wide association analysis for the studied traits.

Trait	GGA_Mb[1]	Start – end positions (Chr_SNP)	SNPs[2]	%GV[3]	PPA (p>0)[4]
DRW	1_166	1_166000511 - 1_166999195	390	3.20	0.92
	26_3	26_3000141 - 26_3998650	998	0.22	0.94
DR%	1_166	1_166000511 - 1_166999195	390	2.79	0.93
	26_3	26_3000141 - 26_3998650	998	0.25	0.95
THW	1_166	1_166000511 - 1_166999195	390	0.20	0.71
	22_4	22_4000760 - 22_4676714	1035	0.54	0.95
	26_1	26_1002598 - 26_1999851	662	0.11	0.84
TH%	1_166	1_166000511 - 1_166999195	390	0.14	0.71
	22_4	22_4000760 - 22_4676714	1035	0.57	0.97
	25_1	25_1000996 - 25_1982441	691	0.11	0.85
	26_1	26_1002598 - 26_1999851	662	0.14	0.85
BRW	25_1	25_1000996 - 25_1982441	691	0.24	0.88
	25_2	25_2001192 - 25_2887176	512	0.81	0.88
	26_3	26_3000141 - 26_3998650	998	0.53	0.96
BR%	25_1	25_1000996 - 25_1982441	691	0.24	0.90
	25_2	25_2001192 - 25_2887176	512	0.6	0.84
	26_3	26_3000141 - 26_3998650	998	0.86	0.98
ABFW	26_1	26_1002598 - 26_1999851	662	1.06	0.95
ABF%	25_1	25_1000996 - 25_1982441	691	0.23	0.87
	26_1	26_1002598 - 26_1999851	662	0.54	0.92

[1] Map position based on Gallus_gallus-5.0 NCBI assembly; [2] Number of SNPs per region; [3] % of genetic variance explained by the window; [4] Posterior probability of association (PPA) as reported by Onteru et al. (2013). THW: thigh weight; TW% thigh yield; ABFW: abdominal fat weight; ABF%: abdominal fat yield; DRW: drumstick weight; DR%: drumstick yield; BTW: breast weight; BT%: breast yield.