A Time-dependent Genome-Wide SNP-SNP Interaction Analysis of Chicken Body Weight

Fang-Ge Li
Northeast Agricultural University
Hui Li (lihui@neau.edu.cn)
Northeast Agricultural University https://orcid.org/0000-0001-6282-5559

Research article

Keywords:

Posted Date: September 19th, 2019

DOI: https://doi.org/10.21203/rs.2.10028/v2

Version of Record: A version of this preprint was published on October 23rd, 2019. See the published version at https://doi.org/10.1186/s12864-019-6132-0.

A Time-dependent Genome-Wide SNP-SNP Interaction Analysis of Chickens' Body Weight

Fang-Ge Li[1,2,3], Hui Li[2,3,4,*]

[1]College of Science, Northeast Agricultural University, Harbin, 150030, P. R. China

[2]Key Laboratory of Chicken Genetics and Breeding, Ministry of Agriculture and Rural Affairs, Harbin, 150030, P. R. China;

[3]Key Laboratory of Animal Genetics, Breeding and Reproduction, Education Department of Heilongjiang Province, Harbin, 150030, P. R. China

[4]College of Animal Science and Technology, Northeast Agricultural University, Harbin, 150030, P. R. China

[*]Corresponding author: Hui Li, Email: lihui@neau.edu.cn

Abstract

The important property of model organism's quantitative traits is time-dependent, which have evoked researchers to be absorbed in the relevant studies. However, investigating the genetic interaction network over time is still lacking in the methodological framework. Our study provided a case in the field and provide insights

into the mechanistic basis of epistasis interaction affecting model organismal phenotypes.

The exhaustive genome-wide search for significant SNP-SNP interaction associated with 475 male birds' body weight at multiple time points (birthday body weight, BW0, 7 days body weight, BW1, 21 days body weight, BW3, 35 days body weight, BW5 and 49 days body weight, BW7) were performed. Statistically, 67, 4 and 2 significant SNP pairs associated with BW0, BW1, BW3 were detected respectively, with the significance threshold 8.67×10^{-12} (Bonferroni-adjusted 1%). Meanwhile, no significant SNP pairs associated with BW5 and BW7 were found. The SNP-SNP interaction networks of BW0, BW1, BW3 were built and annotated. On the strength of annotated information and the strict significant threshold, SNP-SNP interaction underpinned the gene-gene interaction, which might occur between chromosomes or in the same chromosome. Comparing and combing the three networks, the results prove that the genetic network of chicken body weight was dynamic and time-dependent.

Author Summary

Focusing on quantitative traits' genetic architecture, epistasis is a biologically plausible feature. Quantitative traits phenotypes variation must result in part from multi-factorial

genetic perturbation of biochemical networks showing highly dynamic, interconnected and non-linear [1,2]. The spotlight we choosing in the study is dynamic. Also, interconnected and non-linear are included. One bird's body weight averagely starts from 44 gram and rises to 2400 gram in her 49 days lifespan in the resource population. We selected five time points (0, 7, 21, 35, 49 days) to carry out the interaction effect test. The test results illustrated significant SNP-SNP interaction effects were easier to be detected in early age than in later days, no significant SNP pairs recurred in different time points. Obviously, the genetic interaction network of chicken body weight be detected in our study is dynamic. For the first three time points, the interaction networks proved SNP-SNP interaction concentrated in some special regions on the chromosomes which would be the results of gene-gene interaction. The characters of significant interaction effect affecting chicken body weight variation were summarized and described, which was a new attempt for quantitative traits to our knowledge.

Introduction

Epistatic interaction (non-linear interactions between segregating loci) is a tough question in contemporary biology, whose role in the genetic architecture of quantitative traits is still obscure and controversial. Researches in Drosophila melanogaster, yeast,

mice, Arabidopsis thaliana and maize show that epistasis is pervasive and is an important factor that determines variation in quantitative phenotypes [1,3]. On the other hand, in the past 15 years, thousands of genome–wide association studies (GWAS) reported lots of single SNP loci showing significant additive effects, however, especially for quantitative traits, the results were challenged for missing heritability and the lack of replication. Identifying epistasis interaction between SNP loci will be a reasonable way to explain a higher proportion of the heritable variance and to resolve the argue.

Carlborg et al.'s study revealed that an apparently major locus for growth in chicken dissected into a genetic network of four interacting loci, which is a proof that epistatic interactions between genes (or QTLs) were important for chicken quantitative traits [4]. Furthermore, our previous studies also provided the evidences that epistasis interaction could be detected and affected the chicken phenotype variation [5,6,7].

In current study, we focused on chicken's body weight whose phenotype data could be analyzed as time series data. Except birthday week, the chicken lifespan was divided into four periods with equal length, 14 days. Five time points' chicken body weight were selected as phenotype value. Significant SNP-SNP interaction associated with

BW0, BW1, BW3, BW5, BW7 were detected with the exhaustive genome-wide test, then the SNP-SNP interaction networks were built and annotated. The results provide further insight into the genetic network controlling body weight in chickens.

Results

Experimental populations and SNP genotyping

A total of 475 male individuals, containing 203 in the lean line and 272 in the fat line, derived from the 11th generation population of Northeast Agricultural University broiler lines divergently selected for abdominal fat content (NEAUHLF) since 1996 were used in the study [6,7].

Genotyping was carried out using chicken 60K SNP chip (57,636 SNPs) manufactured by the Illumina Inc. (San Diego, CA, USA). After quality control, a total of 48,152 SNPs on 28 autosomes, the Z chromosome, linkage groups and 672 SNPs not assigned to any chromosomes in chickens were included in this study (Table 1). Finally, 48,034 SNPs with chromosome position information were filtered for the interaction analysis.

Phenotypic information

The birds were weighed at 0, 7, 21, 35 and 49 days. Phenotypic summary statistics for body weight are listed in Table 2. Body weights are no significant differences between

the lean and fat lines, so we mixed the two lines into one group in the interaction test. The correlation coefficients between the different days' body weights were calculated in the combined population (Table 3). The minimum correlation coefficient, 0.015 near zero, is between BW0 and BW7, that's mean BW0 and BW7 is uncorrelated. The correlation coefficients between BW0 and BW1, BW1 and BW3, BW3 and BW5, BW5 and BW7 are rising and keep at the high level of 0.65~0.69 at the end.

Epistatic analysis of body weight

MatrixEpistasis [8], an ultrafast method, running exhaustive epistasis scan for quantitative traits with covariate adjustment, was applied to the interaction test. Pairwise (two-dimension, SNP-SNP) interaction effect affecting body weight variation were detected. With the significance threshold 8.67×10^{-12} (Bonferroni-adjusted 1%), in all sixty-seven (Table 4), four (Table 5) and two (Table 6) statistically significant SNP pairs associated with BW0, BW1, BW3 were detected, no replicated significant pairs. There are no SNP pairs' p-value is smaller than the threshold in BW5 and BW7.

SNP-SNP interaction network

The plots illustrating the SNP-SNP interaction networks with the significant epistatic effects for chicken body weight were drawn with Cytoscape 3.7.0 software package [9].

The detail analysis of networks was mainly applied to BW0' network, and the networks of BW1 and BW3 was sketched because they are clear and simple.

The SNP-SNP interaction network of BW0 (Figure 1) is constituted of some separated subnets, which containing more than three SNPs are shown in Figure 1. BW0's SNP epistatic interaction network is approximatively 'small world' and scale-free, the major topological features of interaction network in biology. 'Small world' means shorter path and independent subnets, resulting in dense local neighborhoods of interacting genes that interact with each other [1]. The results of gene-gene interaction will be inferred in the next step. The scale-free property of network implies that Gga_rs14184594 is the hub locus with the maximum degree.

Annotation of SNP loci and SNP-SNP interaction network

Annotating genes of the interaction networks, 1Mb length region was designated for each SNP, upstream 0.5Mb and downstream 0.5Mb. The regions will be merged if they intersect. Genes overlapping the regions were retrieved from UCSC (https://genome.ucsc.edu/) (Galgal5). The details were listed in Support Table 1, 2, 3.

Focusing on the network of BW0, the significant interaction SNP-SNP pairs contain 80 single SNPs in which 30 SNPs are located in the Z chromosome. Observing the

annotation information, something interesting emerged out. Many SNPs from the same subnet are neighbors, concentrating in the same region. Therefore, we adjusted the spatial position of SNPs in Figure 1, making SNPs near if they were in the same region. All the subnets include SNPs from the same region, except SubNet_8 and SubNet_9. The phenomenon enhanced the reliability to infer that SNP-SNP interaction would be the results of gene-gene interaction in the correspond region. However, the annotation would generate a gene set in each region, thus the conclusions were the interaction existing between gene sets, that's mean the point to point interaction relationship could not be provided.

The cross lines in Subnet_1 accounted the interactions between chr19: (3823581, 5935922) and chrz: (65912281, 67063604), chr19: (1728331, 3504813) and chrz: (65912281, 67063604), which would be the signal of gene set (INIP, GNG10, SMC2, PTGR1, TXN, MUSK, LPAR1) on the Z chromosome interacting with gene sets on the chromosome 19. SubNet_2 claimed the gene set (GLDC, TYRP1, MPDZ, NFIB, ZDHHC21, CER1, PSIP1) on the Z chromosome interacting with the gene set (IGFBP1, IGFBP3, TNS3, SLC12A7) near the hub SNP on the chromosome 2. SubNet_3 and SubNet_4 proofed the interaction effect could happen in the same chromosome.

Furthermore, SNPs in SubNet_4 are all in the same region, neighborhoods genes interacting with each other. The inferred gene sets interactions were shown in Table 7. Functional annotation of genes performed by DAVID bioinformatics resources 6.8 (https://david.ncifcrf.gov/home.jsp), eight terms, including Chromosome, Nucleus, Phosphoprotein, Acetylation, DNA-binding, nucleosome, Nucleosome core, Histone H5, were significantly ($P < 0.05$) enriched.

In brief, our study stated the genetic interaction network of chicken body weight is time-dependent and the epistatic interaction effect is dynamic. The most active time of interaction effect is birthday, then the effects declining, and at the time of 5 and 7 weeks the effects are difficulty to be detected. The results might imply that the interaction effect among genes would be active in the early days.

Discussion

To our knowledge, the study was conducted from a new perspective, detecting the interaction effects affecting quantitative traits phenotype variation at multiple time points with SNP data. Basing on the recognition of phenotype data continuous changing, the key feature of quantitative traits, we tried to discuss the quantitative traits' genetic network would be similar or dissimilar at different periods, which was the orientation

of the study. Many studies have certified that interaction effects would be tested, however, the study of whether the genetic network is time-dependent is lacking. Our research provided an example and manifested the genetic network is time-dependent, which could make up the lacking at some extent.

Chicken (*Gallus gallus*), a vertebrate, is a model organism and agricultural specie, whose body weight is a typical quantitative trait and easy to be measured. Broiler body weight's heritability, in males, estimates ranged from 0.29 to 0.37 [9, 10], a medium level. As a consequence, broiler body weight is a suitable quantitative trait for detecting interaction effect and figuring out the feature of genetic network.

Male individuals' body weight data used in the study were derived from NEAUHLF, a broiler line. Body weights did not show significant differences between the lean and fat lines in the resource population, thus the phenotype is not be divided by lines. Besides, larger sample number will improve the interaction test power. However, stratification in genomic and population might occur, because of the long-time artificial selection pressure. In order to handle these problems, MatrixEpistasi with covariate adjustment was selected as the statistical method which can remove confounding bias [8]. The other advantage of MatrixEpistasi is ultra-computational speed, the critical factor for SNP-

SNP interaction mapping at the genome-wide level.

Testing multiple hypotheses cause the significance threshold p-value (8.67×10^{-12}, Bonferroni-adjusted 1%) is far smaller than 1%. The significant test results heavily depended on the significance threshold that would be arbitrary. Although some effects would be ignored, the strict threshold enhanced the confidence of results. With the strict threshold, the detection results showed the interaction effect were totally different at different time points. It's a hint that time point is an important factor in the quantitative trait's interaction test. For the individuals, it's easy to find out interaction effect in birthday, while it's difficulty for 35 days and 49 days, which shows that the cooperation between genes would be closer in the early days than later days. From the perspective of data driving, the correlations between BW0 and other days body weight were relatively small, which was one the reasons of the different results. More important and interesting, the results would be strong evidences that the genetic regulation network would be different at different time points. Carlborg et al [11] have found the similar conclusion in the chicken.

Many SNPs' positions on the chromosome are neighboring in the SNP-SNP interaction network. We suggested it should be the signal of gene-gene interaction in the

correspond region. LD (Linkage Disequilibrium) or QTL (Quantitative Trait Locus) information were not introduced in the study, because the way to test LD is not correlated with epistatic interaction test and gene-gene results would more accurate than QTL-QTL interaction usually. Actually, gene-gene interaction results were based on the gene sets, and the really gene-gene interaction would be verified by biological experiments.

In the future, more research will be needed to study to get more recognition of the quantitative traits' genetic network.

Methods

The method is implemented in R and available at https://github.com/fanglab/MatrixEpistasis.The statistical model

$$p = \alpha + \beta_1 G_{.s} + \beta_2 G_{.t} + \beta_3 G_{.s} G_{.t} + \sum_v \gamma_v C_v + \varepsilon$$

Where α is the overall mean of the quantitative phenotype, $\beta_1, \beta_2, \beta_3$ and γ_v are, respectively, the regression coefficients for the main genetic additive effect, interaction effect and covariates, and ε is a normal variable with zero mean and ξ^2variance [8]. In this model, the phenotype with both main genetic additive effects and covariates adjusted, the size of the effect that the interaction term is the regression coefficient β_3,

205 therefore the hypotheses H_0: $\beta_3 = 0$ and H_1: $\beta_3 \neq 0$, that's mean tests of interaction

206 correspond to testing whether the regression coefficient β_3 equals zero or not.

207

208

209 **Acknowledgements**

210 The authors acknowledge the members of the Poultry Breeding Group of the College

211 of Animal Science and Technology at the Northeast Agricultural University for

212 managing the birds and collecting the data.

213

214 **Funding**

215 This research was supported by the National Natural Science Foundation (No.

216 31601935) and China Agriculture Research System (No. CARS-41).

217

218 **Authors' contributions**

219 FGL analyzed and interpreted the data, drafted, and wrote the manuscript. HL led the

220 conception and design of the study and helped write the manuscript. All authors

221 submitted comments on drafts, and read and approved the final manuscript.

Availability of data and materials

The chicken 60 k SNP data used in the the article, including the additional files, have been deposited into Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo/) with the identifier GSE58551.

Ethics approval and consent to participate

All the animal work, the care and use of experimental animals, followed the guidelines established by the Ministry of Science and Technology of the People's Republic of China (Approval number: 2006–398) and were approved by the Laboratory Animal Management Committee of Northeast Agricultural University.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Mackay T F C. Epistasis and quantitative traits: using model organisms to study gene–gene interactions[J]. Nature Reviews Genetics, 2014, 15(1): 22.

2. Kauffman, SA. The Origins of Order. Oxford University Press; New York: 1993.

3. Huang W, Richards S, Carbone MA, Zhu D, Anholt RR, et al. (2012) Epistasis dominates the genetic architecture of Drosophila quantitative traits. Proc Natl Acad Sci U S A 109: 15553–15559.

4. Carlborg O, Jacobsson L, Ahgren P, Siegel P, Andersson L (2006) Epistasis and the release of genetic variation during long-term selection. Nat Genet 38: 418– 420.

5. Hu G, Wang SZ, Wang ZP, Li YM, Li H (2010) Genetic epistasis analysis of 10 peroxisome proliferator-activated receptor gamma-correlated genes in broiler lines divergently selected for abdominal fat content. Poult Sci 89: 2341–2350.

6. Li F, Hu G, Zhang H, Wang S, Wang Z, Li H. Epistatic effects on abdominal fat content in chickens: results from a genome-wide SNP-SNP interaction analysis. PLoS One. 2013;8(12):e81520. Published 2013 Dec 5. doi:10.1371/journal.pone.0081520

7. Zhang H, Yu JQ, Yang LL, et al. Identification of genome-wide SNP-SNP interactions associated with important traits in chicken. BMC Genomics.

2017;18(1):892. Published 2017 Nov 21. doi:10.1186/s12864-017-4252-y

8. Zhu S, Fang G, Berger B. Matrix Epistasis: ultrafast, exhaustive epistasis scan for quantitative traits with covariate adjustment[J]. Bioinformatics, 2018, 1: 8.

9. Maniatis G, Demiris N, Kranis A, et al. Comparison of inference methods of genetic parameters with an application to body weight in broilers[J]. Archives Animal Breeding, 2015.

10. Mebratie W, Madsen P, Hawken R, et al. Multi-trait estimation of genetic parameters for body weight in a commercial broiler chicken population[J]. Livestock Science, 2018, 217: 15-18.

11.Carlborg O, Kerje S, Schütz K, Jacobsson L, Jensen P, et al. (2003) A global search reveals epistatic interaction between QTL for early growth in the chicken. Genome research 13: 413–21.

Figures

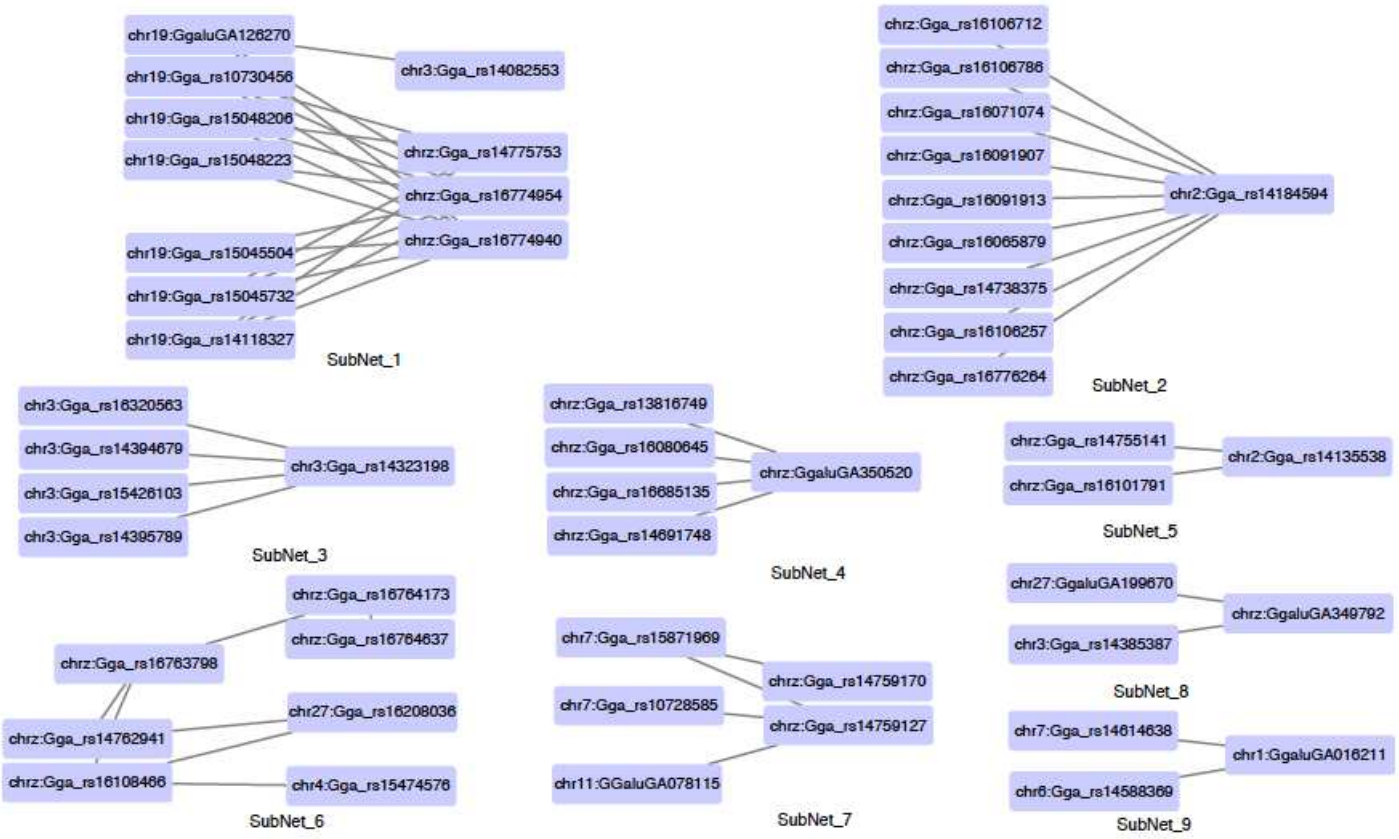

Figure 1

Epistatic SNP-SNP interaction network of birthday body weight (BW0) in NEAUHLF. One node represents one SNP whose name and chromosome number are shown in the rectangle. Significant SNP-SNP interactions were connected by the edge.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- tables.pdf